A Set Theoretic Approach to Lifting Procedures for $0, 1$ Integer Programming

Mark Zuckerberg

Submitted in partial fulfilment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2004

ABSTRACT

A Set Theoretic Approach to Lifting Procedures for 0, 1 Integer Programming

Mark Zuckerberg

A new lifting procedure for 0, 1 integer programming problems is introduced in which variables are appended to correspond to each logical statement that can be made about the vectors in the feasible region. It is shown that this lifting generalizes the liftings of Sherali and Adams [SA90] and Lovász and Schrijver [LS91], and that its features generalize the features of those liftings. The new larger lifting provides a broader conceptual framework for 0, 1 integer programming that is only incompletely exploited by the techniques based on the older liftings. We suggest several polynomial time algorithms in particular that take advantage of the larger lifting by tailoring their choice of new variables to the structure of the feasible set itself. One notable feature of these algorithms is that for large classes of problems, including the "Set Covering Problem", they produce in polynomial time a linear system of polynomial size each of whose solutions is guaranteed to satisfy all linear constraints on the feasible set whose coefficients are in $\{0, 1, \ldots, k\}$.

Contents

Acknowledgements

I would like to express my sincere gratitude for all of the support, encouragement and suggestions that were so graciously offered by family, friends, colleagues and professors, and that are ultimately responsible for this work. I would like to thank the Department of Industrial Engineering and Operations Research at Columbia University's School of Engineering and Applied Science for providing a stimulating yet extremely friendly and supportive environment for learning. I am grateful to have shared the company and friendship of all of the students and staff.

In particular I want to express my appreciation to Professor Donald Goldfarb, who first recruited me to Columbia, and to my thesis advisor, Professor Daniel Bienstock. Professors Goldfarb and Bienstock have been my teachers, patrons and friends throughout my years at Columbia, and I am deeply grateful for their friendship and for their help. I would like to thank Professor Bienstock in particular for first directing me to this line of research, for his collaboration, and for his unhalting supply of suggestions, help and encouragement. I learned a great deal, and I enjoyed it too. I would also like to thank Professors Egon Balas of Carnegie Mellon University, Cliff Stein of Columbia University and Dr. Sanjeeb Dash Of I.B.M. T.J. Watson Research Center, as well as Professors Goldfarb and Bienstock for serving on my thesis committee and for their helpful comments and suggestions.

My parents and my wife's parents have been responsible for shouldering most of the financial burden of this undertaking, but their support goes far beyond the financial. To them we quite literally owe everything.

I suppose that it is standard fare to thank one's wife for her patience, support and encouragement, and I certainly owe a debt of gratitude to my wife Rivka for that. But this only begins to tell the story. Rivka is indeed my better half and whatever is mine is hers. Our children, ages 2, 4 and 6, also deserve my thanks for ceaseless hours of fun – and work.

I thank the Almighty for providing me with all of the above, and for altogether providing me with boundless opportunities. May He grant me the wisdom to use them.

Preface

Introduction

The general integer programming problem is to find the minimum of a function over the set of integer vectors that satisfy a given collection of constraints. In particular, a linear integer programming problem is a problem of the form

$$\text{minimize } \left\{ c^T x \text{ subject to } Ax \geq b,\ x \in Z^n \right\} \tag{1}$$

where A is an $m \times n$ matrix of real numbers, $b \in R^m$, $c \in R^n$ and Z^n is the set of integer points in R^n. The special case where x is restricted to belong to $\{0, 1\}^n$ is known as $0, 1$ linear integer programming. Optimization problems concerning yes/no decisions can often be modeled as $0, 1$ linear integer programming problems, and in particular, many combinatorial and graph theoretical optimization problems can be modeled in this manner.

These problems are NP-hard, and have long been recognized as extremely difficult, though various approaches exist for approximately (and, on occasion, exactly) solving them. A classical enumerative approach is "branch and bound", in which the feasible region is broken up into progressively smaller pieces and one uses approximations of the optimal value of the function taken over these smaller regions in order to provide increasingly better bounds on the optimal value of the function taken over the entire region. Another standard tool is polyhedral optimization, in which the integrality constraints are dropped, turning the problem into a linear program (for which efficient algorithms are known). If dropping the integrality constraints yields the "convex hull" of the original feasible set (i.e. the smallest convex set that includes the original feasible set), then the optimal function value taken over the relaxation is the same as the optimal function value taken over the original feasible set itself. In general, the feasible region of the relaxation produced by dropping the integrality constraints is considerably larger than the convex hull. However, there exist a number of "cutting plane algorithms" that cut down the relaxation (by appending new valid linear constraints on the original feasible region) so as to better approximate the convex hull. See

[S86] and [W98].

Another approach, conceptually related to the cutting plane approach, that has attracted interest recently is that of the "lifting algorithms", in which one appends new variables (with certain associated constraints attached) to the problem in a systematic way and then seeks to solve the expanded "lifted" problem (see [SA90]). Stated loosely, the larger "lifted" formulations tend to describe the feasible region more comprehensively, and thus lifting the problem in this way can often make the problem easier to solve. Lovász and Schrijver [LS91] and more recently Lasserre [Las01] have shown that the lifting procedures can be used to impose certain semidefinite constraints. The lifting procedures are also related to "disjunctive programming" (see [BCC93]) in which the feasible region is seen as a union of sets (we will see more on this later).

In this work we describe a new kind of lifting in which variables are appended for each logical statement that may be made about the vectors in the feasible space (we will quantify this idea further in the following section). We show that the liftings that have been described in the earlier literature are all subsumed by this larger lifting. Further, we show that this larger lifting puts all of the specific properties of the older liftings (including their associated semidefinite constraints) in a broader and more natural context, and that much of the potential implicit in the larger lifting goes untapped by the older liftings. In particular, we introduce several algorithms that systematically incorporate variables of this new stripe in ways that reflect the specific structure of the feasible region. There are significant gains to be realized in doing this. For example, for large classes of problems one can produce with these algorithms a linear system in polynomial time and of polynomial size whose solutions are guaranteed to satisfy all valid linear constraints on the feasible region whose coefficients are in $\{0, 1, \ldots, k\}$, where k is fixed. (We show in the following section that there is actually a considerably stronger characterization of constraints guaranteed to be satisfied.) This larger lifting will also further clarify the connection between lifting theory and disjunctive programming.

Overview of the Thesis

One of the classical approaches for approximately solving linear integer programs in general, and for solving them exactly in certain special cases, is polyhedral optimization. (In what follows we will limit our discussion to bounded polyhedral sets, as these are the only sets of interest in $0, 1$ programming.) The first fundamental result underlying this approach is the fact that a linear function attains its minimum over a bounded polyhedral set at one of its vertices. (A polyhedral set is the subset of R^n that satisfies some finite system of linear constraints.) Observe now that for any finite set of points $Q \subset R^n$, the convex hull of Q is a bounded polyhedral set whose vertices all belong to Q. Thus, in particular, where the set of integer points that satisfy $Ax \geq b$ is denoted P, then the function $c^T x$ is minimized over $Conv(P)$ at one of those integer points. This integer point therefore minimizes $c^T x$ over the subset P of $Conv(P)$ as well.

The second fundamental result concerns "polynomial time separation oracles". A polynomial time separation oracle (on R^n) for a subset Q of R^n is a function that takes a point $x \in R^n$ as input, and outputs, in time polynomial in the size of some given representation of Q, a "yes" if $x \in Q$, or otherwise a vector $u \in R^n$ such that $u^T y \geq \beta$, $\forall y \in Q$ but for which $u^T x < \beta$. The second fundamental result states if a polynomial time separation oracle exists for a set $Q \subseteq R^n$, then linear functions can be minimized over Q in polynomial time. (By "polynomial time", as above, we mean polynomial in the size of the given representation of Q.) We can therefore conclude that, where the set of integer points that satisfy $Ax \geq b$ is denoted P, and where by "polynomial" we mean polynomial in the size of an encoding of the matrix $A|b$, if a polynomial time separation oracle exists for $Conv(P)$, then we can solve the linear integer programming problem in polynomial time.

Given a polyhedral set $\{x \in R^n : Hx \geq h\}$, it is clearly possible to separate over this set in polynomial time. Thus general linear programming problems can be solved in polynomial time. It has also long been known that any bounded polyhedral set in R^n is the convex hull of a finite set of points, and vice-versa. Thus where P is, as above, the set of integer points that satisfy $Ax \geq b$, there must exist some representation of $Conv(P)$ as $\{x \in R^n : Hx \geq h\}$, over which we can separate in polynomial time (in the size of $H|h$). The two sticking points, however, are that we do not know this representation in advance, and that the size of the matrix $H|h$ may be exponentially larger than that of $A|b$. Observe however, that $\{x \in R^n : Ax \geq b\} \supseteq Conv(P)$ and, as a linear system, we can optimize over this set in polynomial time, and thereby obtain a lower bound on the minimum over $Conv(P)$. Moreover, any constraint $d^T x \geq \alpha$ that is valid for $Conv(P)$, and is not implied by the constraints $Ax \geq b$ can be appended to the system, tightening the formulation, and improving the approximation. This gives rise to various "cutting plane algorithms" that, through a number of techniques and heuristics, seek to derive such valid constraints, as well as a body of theoretical work characterizing some situations in which a formulation can be known to be convex hull defining for its set of integer points. For details see [S86] and [W98].

A different method for dealing with integer programs is that of "lifting" the underlying set P to a higher dimension. The lifting methods append additional variables to the original formulation, and then place new constraints on the "lifted" vector. The basic idea is that a lifting of the set $Conv(P)$ to a higher dimension may yield a set that has fewer facets and is easier to characterize than the original representation. We will see examples of sets with an exponential number of facet defining inequalities all of which are satisfied by a lifting with a polynomial number of constraints, and similar examples have been known for some time (see references cited in the introduction to [LS91]).

Given a $0, 1$ linear integer programming problem with feasible region P, Sherali and Adams proposed a lifting technique [SA90] which by its n'th "level" (the procedure is exponential in the value of its level), produces a system of linear constraints in the "lifted vector" with a solution set whose projection on the original variables will be exactly the convex hull of P. One of the noteworthy features of their technique is that it can be applied to polynomial $0, 1$ programs as well. Their procedure can be thought of as a strengthening of the "convexification" procedure described by Balas ([B74], [B79], see [BCC93]), which is also guaranteed to obtain the convex hull by its n'th level. "Convexification" was originally conceived as an application of "disjunctive programming", i.e. problems in which the feasible set is cast as a union of polyhedra, and the general notion of disjunctive programming, in one form or another, reverberates through much of the theory of liftings (as we will see). Lovász and Schrijver [LS91] introduced a semidefinite constraint that can be applied to these liftings, and they also generalized the theory that underlies them. More recently, Lasserre [Las01] introduced an algorithm for general polynomial programming whose application to $0, 1$ integer programming strengthens the semidefinite constraint of [LS91], and replaces the linear constraints of the earlier procedures with semidefinite constraints of the same flavor as that of [LS91] (see [Lau01]).

In Chapter 1 we will outline the basic ideas that underlie all of these algorithms, and we will review in detail the broader theory developed in [LS91], as it serves in many ways as the motivation for what follows. No new results are presented, but the presentation and, in most cases, the proofs, have been altered. In Chapter 2 we will present a new interpretation and derivation of the results of Chapter 1, by which the liftings described in the first chapter take on a much more natural meaning. At the end of the chapter we will show how this new interpretation suggests a much broader lifting, to $O(2^{2^n})$ dimensions.

This larger lifting, which is described in Chapter 3, is based on the notion that each coordinate of the $0, 1$ vectors that belong to a set can be thought of as saying something about the point of which it is a coordinate. For example $y_i = 1$ says that the i'th coordinate of y is one, or equivalently, $y \in \{y \in \{0, 1\}^n : y_i = 1\}$. But there are many other "things" that may be said about a point as well, and for each such "thing" we can append a coordinate that identifies whether or not the statement holds for that point. The logical and set theoretic structure of P can thus be captured in the behavior of its lifted vectors.

Suppose, for example, that for every point $y \in P \subseteq \{0,1\}^n$ for which either $y_1 = 0$ and $y_2 = 1$, or $y_1 = 1$ and $y_2 = 0$ we must have $y_3 = 1$. This is a logical constraint of the form

$$y_1 \; XOR \; y_2 \Rightarrow y_3 \tag{2}$$

(where the expression "XOR" means "exclusive or"). (For more on logical programming, see for example [H00], [BH01].) Define the set

$$q = \{y \in P : \text{exactly one of the two coordinates } y_1 \text{ and } y_2 \text{ has value } 1\}. \tag{3}$$

Given a vector $y = (y_1, \ldots, y_n) \in P$ and a set $r \subseteq P$, define now the $0,1$ valued function $y[r]$ which will take the value 1 if and only if $y \in r$. We can then think of $y[q]$ as the coordinate of a lifting of the vector y that "says" whether or not y is indeed such that exactly one of its two coordinates y_1 and y_2 has value 1. (Technically, $y[q]$ is the boolean function $y_1 \; XOR \; y_2$.) Note also that by this definition, each variable y_i, $i \in \{1, \ldots, n\}$, can be thought of as $y[\{y \in P : y_i = 1\}]$.) Thus since P has been assumed to be such that wherever $y \in q$ then $y_3 = 1$, it follows that for each $y \in P$ we have $y[q] \leq y_3$. Thus $y[q] \leq y_3$ is a linear condition on the lifted vector that encodes the logical condition (2). We could also note that where we define

$$u = \{y \in P : y_1 = y_2 = 1\} \tag{4}$$

and

$$v = \{y \in P : y_1 = 1 \text{ or } y_2 = 1\} \tag{5}$$

then it is easy to see that

$$y_1 + y_2 - y[u] = y[v], \quad \text{and} \tag{6}$$

$$y[v] - y[u] = y[q]. \tag{7}$$

It is evident that by way of such constraints, an array of linear relationships can be constructed connecting the new variables with each other and with the original variables. We will see an example of this point shortly.

Define now

$$Y_i^P = \{y \in P : y_i = 1\} \tag{8}$$

so that

$$u = Y_1^P \cap Y_2^P, \quad v = Y_1^P \cup Y_2^P, \quad q = (Y_1^P \cup Y_2^P) - (Y_1^P \cap Y_2^P), \text{ and} \tag{9}$$

$$y_i = y\left[\{y \in P : y_i = 1\}\right] = y[Y_i^P]. \tag{10}$$

Thus the relationships (6) and (7) say that for each $y \in P$,

$$y[Y_1^P] + y[Y_2^P] - y[Y_1^P \cap Y_2^P] = y[Y_1^P \cup Y_2^P] \tag{11}$$

and

$$y\left[(Y_1^P \cup Y_2^P) - (Y_1^P \cap Y_2^P)\right] = y[Y_1^P \cup Y_2^P] - y[Y_1^P \cap Y_2^P]. \tag{12}$$

The quantities $y[\cdot]$ thus seem to behave as measures, and we will develop this connection fully in Chapters 2 and 3. Note moreover that it follows from (2) that $q \subseteq Y_3^P$, so the relationship $y[q] \leq y_3$ also reflects the measure theoretic property that for all sets A, B with $A \subseteq B$, the measure of A is less than or equal to the measure of B. It is easy to see that in general, for any sets $r, s \subseteq P,\ r \subseteq s$, we will have $y[r] \leq y[s]$, and for any two disjoint sets $t, w \subseteq P$ we will have $y[t] + y[w] = y[t \cup w]$. Note also that considering expressions (11) and (12), the constraint $y[q] \leq y_3$ is very similar to the valid polynomial constraint

$$y_1 + y_2 - 2y_1y_2 \leq y_3. \tag{13}$$

Thus the constraints that may be imposed on the lifted vectors of P can be thought of as measure theoretic constraints, and they are closely connected with the logical constraints and the polynomial constraints that may be imposed on the vectors of P. An important difference however between the measure theoretic and the logical/polynomial constraints, is that unlike the latter two, the measure theoretic constraints on the lifted vectors are linear, and they carry over to convex combinations of the (liftings of the) points of P as well. Say for example that P satisfies (2) as above, and define

$$\hat{P} = \left\{(y_1, \ldots, y_n, y[q]) \in \{0,1\}^{n+1} : (y_1, \ldots, y_n) \in P\right\}. \tag{14}$$

Any point $\hat{x} = (x_1, \ldots, x_n, x[q]) \in [0,1]^{n+1}$ such that $\hat{x} \in Conv(\hat{P})$ will also satisfy $x[q] \leq x_3$. More generally, given $y \in P$, let $\bar{y}$ be the lifting of the vector y obtained by appending a coordinate of value $y[u]$ for each set $u \in \mathcal{Q}$, where $\mathcal{Q}$ is some collection of subsets of P; let $\hat{P} = \{\bar{y} : y \in P\}$, and let $\alpha^T \bar{y} \geq \beta$ be any valid constraint on $\hat{P}$. Then any vector $\bar{x}$ of the same dimension as $\bar{y}$, such that $\bar{x} \in Conv(\hat{P})$, will satisfy $\alpha^T \bar{x} \geq \beta$ as well. (It is important to note that though in our example we denoted the appended coordinate of $\bar{x}$ as $x[q]$, the value $x[q]$ is not a function of $(x_1, \ldots, x_n)$, (in contradistinction to $y[q]$, which is a function of $(y_1, \ldots, y_n)$). The only constraint that we have placed on the numbers $x_1, \ldots, x_n, x[q]$ is that the vector $(x_1, \ldots, x_n, x[q])$ must belong to the convex hull of $\hat{P}$.)

We have thus seen that the lifting approach provides a means by which we may constrain candidate vectors for membership in $Conv(P)$ with *linear* constraints that reflect abstract logical characteristics of the set P. But it is obvious that the effectiveness of the additional variables and constraints will depend upon the quality of the network of relationships that connect the new variables to the old. For example, where P satisfies the logical constraint (2) as above, and say that P is given by

$$P = \{y \in \{0,1\}^n : Ay \geq b\} \tag{15}$$

for some matrix A and vector b, then if we merely append the single variable $x[q]$ and the single constraint $x[q] \leq x_3$ to the linear relaxation

$$\bar{P} = \{x \in [0,1]^n : Ax \geq b\} \tag{16}$$

it is clear that for any $x \in \bar{P}$ we could always choose $x[q] = x_3$, and so we will not eliminate any points from $\bar{P} - P$. We will see however that a careful choice of new variables and constraints, guided by the structure of P, can be used to replace an exponentially large number of facet defining constraints on the original system with a polynomially large number of constraints on the lifted system.

This brings us to one of the key features by which our work differs from its predecessors. The Sherali-Adams, Lovász-Schrijver and Lasserre algorithms can all also be understood within the framework that we have described, though this is not the way that they were originally conceived. They can all be interpretted as processes that methodically append variables corresponding to logical properties of vectors in $\{0,1\}^n$ (or equivalently, to sets in $\{0,1\}^n$), and which then impose linear or semidefinite constraints on the new variables. But viewed from this perspective, we will see that all of them limit themselves to appending variables solely for a particular class of subsets of $\{0,1\}^n$. Specifically, we will see that the variables that are appended in these algorithms all correspond (explicitly or implicitly) to sets of the form

$$\bigcap_{j \in J} Y_j \cap \bigcap_{j \in \bar{J}} Y_j^c \tag{17}$$

where J and $\bar{J}$ are disjoint subsets of $\{1, \ldots, n\}$, and $Y_j = \{y \in \{0,1\}^n : y_j = 1\}$. Note moreover that there are still exponentially many such sets, and so a polynomial time algorithm could at most select a sample of sets of this form. But in this selection process itself, these algorithms also all follow the same procedure, and the procedure that they follow is completely independent of the structure of the given feasible region $P \subseteq \{0,1\}^n$. Regardless of P, they first consider all intersections of ≤ 2 sets of the form (17), then all intersections of ≤ 3 sets of the form (17), and so on. Thus while these algorithms may be understood as applications of the framework that we have outlined, they are quite limited applications.

In Chapter 3 we will develop the more general mathematics that characterizes the larger liftings that we have described. We will show how this lifting establishes a natural connection between the algebra of subsets of the feasible region $P \subseteq \{0,1\}^n$, the measures on that algebra, and the convex hull of P, and we will see how the mathematical properties that characterized the procedures described in the first two chapters are special applications. We will also indicate a way in which these results can be generalized to arbitrary countable sets $P \subseteq R^n$.

In Chapter 4 we will focus specifically on the semidefinite constraint that was introduced in [LS91], and which finds broader application in [Las01]. We will see how this constraint also can be put into the larger context introduced in Chapter 3. This larger context will shed a good deal of light on where, why and how positive semidefiniteness can (or cannot) be used to advantage, and it will considerably broaden the possibilities for its application.

In Chapters 5 and 6 we will turn our attention to using the new machinery to develop new algorithms. These algorithms, which are also guaranteed to eventually obtain the convex hull of the feasible region $P \subseteq \{0,1\}^n$, will not select their new variables in any rigid manner, rather they will use the specific structure of P as their guide in selecting new variables. Consider, for example, $P \subseteq \{0,1\}^n$ defined as the set of points $y \in \{0,1\}^n$ satisfying the system of linear constraints given by the full circulant matrix:

$$y_1 + y_2 + y_3 \geq 1 \tag{18}$$

$$y_1 + y_2 + y_4 \geq 1 \tag{19}$$

$$y_1 + y_3 + y_4 \geq 1 \tag{20}$$

$$y_2 + y_3 + y_4 \geq 1. \tag{21}$$

Recall the definition $Y_j = \{y \in \{0,1\}^n : y_j = 1\}$, and define

$$R_1 = (Y_1 \cup Y_2 \cup Y_3) \tag{22}$$

$$R_2 = (Y_1 \cup Y_2 \cup Y_4) \tag{23}$$

$$R_3 = (Y_1 \cup Y_3 \cup Y_4) \tag{24}$$

$$R_4 = (Y_2 \cup Y_3 \cup Y_4) \tag{25}$$

$$Q_{1,1} = Y_1,\ Q_{1,2} = Y_2,\ Q_{1,3} = Y_3 \tag{26}$$

$$Q_{2,1} = Y_1,\ Q_{2,2} = Y_2,\ Q_{2,3} = Y_4 \tag{27}$$

$$Q_{3,1} = Y_1,\ Q_{3,2} = Y_3,\ Q_{3,3} = Y_4 \tag{28}$$

$$Q_{4,1} = Y_2,\ Q_{4,2} = Y_3,\ Q_{4,3} = Y_4 \tag{29}$$

and note that P can also be described by

$$P = R_1 \cap R_2 \cap R_3 \cap R_4. \tag{30}$$

For each $i = 1, \ldots, 4$, define the sets

$$T(\{i,1\}) := \bigcap_{\substack{\bar{i}=1 \\ \bar{i}\neq i}}^{4} R_{\bar{i}} \cap Q_{i,1} \tag{31}$$

$$T(\{i,2\}) = \bigcap_{\substack{\bar{i}=1 \\ \bar{i}\neq i}}^{4} R_{\bar{i}} \cap Q_{i,1}^c \cap Q_{i,2} \tag{32}$$

$$T(\{i,3\}) = \bigcap_{\substack{\bar{i}=1 \\ \bar{i}\neq i}}^{4} R_{\bar{i}} \cap Q_{i,1}^c \cap Q_{i,2}^c \cap Q_{i,3}. \tag{33}$$

Note now that for each $i = 1, \ldots, 4$,

$$R_i = Q_{i,1}\ \cup\ (Q_{i,1}^c \cap Q_{i,2})\ \cup\ (Q_{i,1}^c \cap Q_{i,2}^c \cap Q_{i,3}) \tag{34}$$

and the sets in the union are pairwise disjoint. Thus for each $i = 1, \ldots, 4$, the sets P can be partitioned as

$$P = T(\{i,1\}) \cup T(\{i,2\}) \cup T(\{i,3\}) \tag{35}$$

and each set $Y_l^P = Y_l \cap P$ can also be partitioned as

$$Y_l^P = \left(T(\{i,1\}) \cap Y_l^P\right)\ \cup\ \left(T(\{i,2\}) \cap Y_l^P\right)\ \cup\ \left(T(\{i,3\}) \cap Y_l^P\right). \tag{36}$$

Note also that for each i, j, $T(\{i,j\}) \subseteq Q_{i,j}$, so that $T(\{i,j\}) \cap Q_{i,j} = T(\{i,j\})$. For each $Q_{i,j}$, define now

$$Q_{i,j}^P = Q_{i,j} \cap P. \tag{37}$$

Thus where for any $y = (y_1, \ldots, y_n) \in P$, and any set $t \subseteq P$ we define, as above, $y[t] = 1$ if $y \in t$ and zero otherwise, then (noting that $y[P] \equiv 1$,) for any $y \in P$, it must be that for each $i = 1, \ldots, 4$,

$$y[T(\{i,1\})] + y[T(\{i,2\})] + y[T(\{i,3\})] = y[P] = 1 \tag{38}$$

$$y[T(\{i,1\}) \cap Y_l^P] + y[T(\{i,2\}) \cap Y_l^P] + y[T(\{i,3\}) \cap Y_l^P] = y[Y_l^P] = y_l \tag{39}$$

$$y[T(\{i,j\}) \cap Q_{i,j}^P] = y[T(\{i,j\})]. \tag{40}$$

(With regard to the use of the term $Q^P_{i,j}$ rather than $Q_{i,j}$ in (40), it may be noticed that $T(\{i,j\})$ is in fact a subset of $Q^P_{i,j}$ as well. However, even were $T(\{i,j\})$ to be only a subset of $Q_{i,j}$ and not of P, we could still write (40) since y was assumed to belong to P.) One last point to notice is that for any valid constraint $\sum_{j=1}^n \alpha_j y_j \geq \beta$ on P, and any set $t \subseteq P$, the constraint

$$\sum_{j=1}^{n} \alpha_j y[t \cap Y_j^P] \geq \beta y[t] \tag{41}$$

is valid for every $y \in P$. To see this note that if $y \notin t$ then every term in expression (41) has value 0, and if $y \in t$, then $y[t] = 1$ and each $y[t \cap Y_j^P] = y[Y_j^P] = y_j$, so that (41) reduces to $\sum_{j=1}^n \alpha_j y_j \geq \beta$, which we have assumed to be valid.

Observe that the constraint

$$y_1 + y_2 + y_3 + y_4 \geq 2 \tag{42}$$

is valid for P. We will now show how one of our algorithms will guarantee that (42) will in fact be satisfied. What follows is not an exact description of the algorithm (it actually differs in some of the details), but an illustration of some of its main ideas.

Beyond the initial variables $x_1, \ldots, x_n$, the algorithm will introduce new variables $x[P] = 1$, and $x[T(\{i,j\})]$ for each $i = 1, \ldots, 4$ and each $j = 1, 2, 3$, as well as variables $x[T(\{i,j\}) \cap Y_l^P]$ for each i, j and each $l = 1, \ldots, 4$. Considering that for any $x \in Conv(P)$, if we denote the lifted version of a vector by the "bar" sign, the lifted vector $\bar{x}$ is in the convex hull of the lifted vectors $\bar{y}$, we can therefore impose constraints on $\bar{x}$ that parallel relationships (38), (39), (40) and (41). These are namely,

$$x[T(\{i,1\})] + x[T(\{i,2\})] + x[T(\{i,3\})] = x[P] = 1 \tag{43}$$

$$x[T(\{i,1\}) \cap Y_l^P] + x[T(\{i,2\}) \cap Y_l^P] + x[T(\{i,3\}) \cap Y_l^P] = x[Y_l^P] = x_l \tag{44}$$

$$x[T(\{i,j\}) \cap Q^P_{i,j}] = x[T(\{i,j\})] \tag{45}$$

$$x[T(\{i,j\}) \cap Y_1^P] + x[T(\{i,j\}) \cap Y_2^P] + x[T(\{i,j\}) \cap Y_3^P] \geq x[T(\{i,j\})] \tag{46}$$

$$x[T(\{i,j\}) \cap Y_1^P] + x[T(\{i,j\}) \cap Y_2^P] + x[T(\{i,j\}) \cap Y_4^P] \geq x[T(\{i,j\})] \tag{47}$$

$$x[T(\{i,j\}) \cap Y_1^P] + x[T(\{i,j\}) \cap Y_3^P] + x[T(\{i,j\}) \cap Y_4^P] \geq x[T(\{i,j\})] \tag{48}$$

$$x[T(\{i,j\}) \cap Y_2^P] + x[T(\{i,j\}) \cap Y_3^P] + x[T(\{i,j\}) \cap Y_4^P] \geq x[T(\{i,j\})]. \tag{49}$$

Consider now $i = j = 1$. By (45) and the definition of the sets $Q^P_{i,j}$, we have

$$x[T(\{1,1\} \cap Y_1^P] = x[T(\{1,1\})] \tag{50}$$

and by (49) we have

$$x[T(\{1,1\}) \cap Y_2^P] + x[T(\{1,1\}) \cap Y_3^P] + x[T(\{1,1\}) \cap Y_4^P] \geq x[T(\{1,1\})]. \qquad (51)$$

Putting these two together we get

$$x[T(\{1,1\}) \cap Y_1^P] + x[T(\{1,1\}) \cap Y_2^P] + x[T(\{1,1\}) \cap Y_3^P] + x[T(\{1,1\}) \cap Y_4^P] \geq 2x[T(\{1,1\})]. \qquad (52)$$

It is not hard to check that a similar argument will show that for all $j = 1, \ldots, 3$ we will also have

$$x[T(\{1,j\}) \cap Y_1^P] + x[T(\{1,j\}) \cap Y_2^P] + x[T(\{1,j\}) \cap Y_3^P] + x[T(\{1,j\}) \cap Y_4^P] \geq 2x[T(\{1,j\})]. \qquad (53)$$

Thus by (43), (44) and (53),

$$x_1 + x_2 + x_3 + x_4 = \sum_{l=1}^{4} \sum_{j=1}^{3} x[T(\{1,j\}) \cap Y_l^P] = \qquad (54)$$

$$\sum_{j=1}^{3} \sum_{l=1}^{4} x[T(\{1,j\}) \cap Y_l^P] \geq 2 \sum_{j=1}^{3} x[T(\{1,j\})] = 2 \qquad (55)$$

which establishes (42).

It is evident that the choice of sets $T(\{i,j\})$ and $T(\{i,j\}) \cap Y_l^P$, corresponding to which we assigned the new variables, depended on the specific structure of P, and these sets bear little similarity to the sets of the form (17) that are considered by the algorithms of the first two chapters. It should be noted that one of the other algorithms that we will suggest does in fact introduce variables that correspond mostly (but not entirely) to sets of the form (17), but that algorithm will not limit itself to first consider the intersections of pairs, and then triples, and so on. When applied to the feasible region given by the general full circulant matrix,

$$P = \{y \in \{0,1\}^n : \sum_{\substack{\bar{i}=1 \\ \bar{i} \neq i}}^{n} y_{\bar{i}} \geq 1,\ i = 1, \ldots, n\}, \qquad (56)$$

that algorithm will choose at its initial level a selection of sets of the form (17) (among some others) that are obtained by intersections of $n-2$ sets of the form Y_l^P and $(Y_l^P)^c$. The algorithms of the first two chapters will also eventually consider these sets, but not until their $n-2$ level, which requires exponential time. We will see in Chapter 5 that the Sherali-Adams and the semidefinitely constrained Lovász-Schrijver operators (and even another theoretical operator that is vastly more powerful than either of them), when applied to the feasible region P given by (56), will not in fact guarantee that the valid constraint

$$\sum_{i=1}^{n} x_i \geq 2 \tag{57}$$

will be satisfied until their $n-2$ level (so they will not satisfy (57) in polynomial time) despite the presence of the semidefinite constraint. Most of our algorithms, however, will guarantee that (57) is satisfied at their initial (nontrivial) level. This is particularly noteworthy considering the fact that our algorithms will obtain this result without the imposition of any semidefinite constraints.

The satisfaction of constraint (57) in the full circulant case (56) is an example of a more general feature that is shared by most of the algorithms that we will suggest. But before we describe the property explicitly, we need some background.

As these algorithms are guided by the set theoretic structure of the feasible region P, they will focus primarily on characterizations of the feasible region in terms of unions, intersections and complementations of sets Y_i (as in (30)), rather than as the set of integer solutions to systems of linear constraints. Where $f(\cdot)$ is an indexing function, let the sets $M_{f(\cdot)}$ be all either of the form Y_l or Y_l^c for some $l \in \{1, \ldots, n\}$. The main algorithm presented in Chapter 5 will address sets P of the form

$$P = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(\cdot)} \bigcap_{i_2=1}^{m_2(\cdot)} \bigcup_{j_2=1}^{t_2(\cdot)} \cdots \bigcap_{i_h=1}^{m_h(\cdot)} \bigcup_{j_h=1}^{t_h(\cdot)} M_{f(i_1,j_1,\cdots,i_h,j_h)} \tag{58}$$

where each t_l is a function of $i_r,\ r \leq l$ and $j_r,\ r < l$, and each m_l is a function of i_r and $j_r,\ r < l$. This definition of P corresponds essentially to an arbitrary logical program. The special case

$$P = \bigcap_{i=1}^{m} \bigcup_{j=1}^{t(i)} Y_{f(i,j)} \tag{59}$$

where $f(i,j)$ maps into $\{1, \ldots, n\}$, characterizes feasible regions for "set covering problems" (which will be defined soon), and the more general case

$$P = \bigcap_{i=1}^{m} \bigcup_{j=1}^{t(i)} M_{f(i,j)} \tag{60}$$

can also be easily characterized as the solution to a system of m linear constraints. The case (58), however, cannot in general be characterized as the solution to a polynomial size (in the length of the set theoretic description) system of linear constraints. The algorithms of Chapter 6 will only address problems of the forms (59) and (60), but they will address them in a different and arguably more interesting way.

To make matters definite, we will be referring in the forthcoming discussion to the main algorithm of Chapter 5, but similar statements will hold for the algorithms of Chapter 6 as well. Given a set P described as per (58), for each fixed k, the main algorithm of Chapter 5 will produce, at level k, a system of linear constraints $D\bar{x} \geq a$ in the lifted vector $\bar{x}$ whose size is polynomial in the length of the set theoretic description (58) of P. The projection of the solution set to this system on the original variables, i.e. the set of subvectors $\{x \in R^n : D\bar{x} \geq a\}$, will be denoted as P^{A_k}, and will constitute a valid relaxation of $Conv(P)$ (i.e. $P^{A_k} \supseteq Conv(P)$). Clearly we can optimize over P^{A_k} in polynomial time (in the length of the set theoretic description of P), since for each $c \in R^n$,

$$\min\left\{c^T x : x \in P^{A_k}\right\} \;=\; \min\left\{\bar{c}^T \bar{x} : D\bar{x} \geq a\right\}, \tag{61}$$

where $\bar{c}$ is the lifting of c to the same dimension as $\bar{x}$ that is obtained by assigning the value 0 to all of the new coordinates ("padding c with zeroes"), and $\min\{c^T x : x \in P^{A_k}\}$ is a lower bound for the value of the integer program $\min\{c^T x : x \in P\}$.

Let us pose now the following definition.

Definition 0.1 *Given an inequality*

$$\alpha^T x \geq \beta,\ \alpha \geq 0,\ 0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{|support(\alpha)|} \tag{62}$$

we will say that the pitch *of the inequality, to be denoted $\pi(\alpha, \beta)$ is*

$$\pi(\alpha, \beta) = \min\left\{k : \sum_{j=1}^{k} \alpha_j \geq \beta\right\}. \tag{63}$$

Stated loosely, the pitch of an inequality may be thought of as a measure of how positive a $0,1$ vector needs to be in order for the inequality to be satisfied. Note that every valid inequality $\alpha^T x \geq \beta,\ (\alpha, \beta) \geq 0$, on a set $P \subseteq \{0,1\}^n$ has pitch $\leq n$, and that where P can be written as (58) with each $M_{f(\cdot)}$ of the form Y_l, then it can be shown that every valid constraint on P is dominated by constraints of the form $\alpha^T x \geq \beta,\ (\alpha, \beta) \geq 0$. In the general case also, we can always introduce n extra variables $x'_i = 1 - x_i,\ i = 1, \ldots, n$ and then write any constraint $\alpha^T x \geq \beta$ as $\tilde{\alpha}^T x + (\alpha')^T x' \geq \beta',\ (\tilde{\alpha}, \alpha', \beta') \geq 0$. Constraint (57) is an example of a valid pitch 2 inequality.

Now we are ready to describe the general property of the algorithms to which we referred above. We will show that where P is of the form (58) with each $M_{f(\cdot)}$ of the form Y_l, then all constraints of pitch $\leq k$ that are valid for P (and for the general P of the form (58), all constraints of pitch $\leq k$ that are valid for a particular relaxation of P) – and there may be exponentially many – are valid for the set P^{A_k}. This characterization shows that our algorithms produce relaxations of $Conv(P)$ that telescope to $Conv(P)$ in a concrete manner, with measurable and concrete improvements at each stage. Again this result holds even without the imposition of semidefinite constraints.

Define now

$$P^k = \{x \in [0,1]^n : \alpha^T x \geq \beta,\ \forall \text{ valid } \alpha^T x \geq \beta \text{ on } P \text{ with } (\alpha, \beta) \geq 0,\ \pi(\alpha, \beta) \leq k\}. \quad (64)$$

Thus if P is indeed of the form (58) with each $M_{f(\cdot)}$ of the form Y_l, then for each k,

$$Conv(P) \subseteq P^{A_k} \subseteq P^k \quad (65)$$

which implies that for each $c \in R^n$,

$$\min\left\{c^T x : x \in P\right\} \geq \min\left\{c^T x : x \in P^{A_k}\right\} \geq \min\left\{c^T x : x \in P^k\right\}. \quad (66)$$

One measure of the significance of this result is as follows. Note first that a "set covering problem" is a problem of the form

$$\min\left\{c^T x : Ax \geq e,\ x \in \{0,1\}^n\right\} \quad (67)$$

where A is a $0,1$ matrix and e is the vector of all 1's. Note that for any set covering problem, its feasible region is of the form $\bigcap_i \bigcup_j Y_{f(i,j)}$, and therefore belongs to the class of sets P for which $P^{A_k} \subseteq P^k$.

Note also that for any set $\bar{P}$ given by a system of linear constraints $\bar{P} = \{x \in R^n_+ : Ax \geq b\}$, where A is an $m \times n$ matrix, the *rank* 1 *Chvátal-Gomory inequalities* of the system are the inequalities

$$\lceil \lambda^T Ax \rceil \geq \lceil \lambda^T b \rceil, \quad \lambda \in R^m_+ \quad (68)$$

where for any vector $x = (x_1, \ldots, x_n)$, the expression $\lceil x \rceil = (\lceil x_1 \rceil, \ldots, \lceil x_n \rceil)$, and each $\lceil x_i \rceil$ is the smallest integer $\geq x_i$. It is easy to see that all of the rank 1 Chvátal-Gomory inequalities are valid for the set $P = \{x \in Z^n_+ : Ax \geq b\}$. We will refer to the original inequalities $Ax \geq b$ as the rank 0 Chvátal-Gomory inequalities. The set of points $\bar{P}^{C-G}$ that satisfies *every* Chvátal-Gomory inequality of rank ≤ 1 is referred to as the Chvátal-Gomory closure of $\bar{P}$, and obviously,

$$Conv(P) \subseteq \bar{P}^{C-G} \subseteq \bar{P}. \quad (69)$$

The rank 2 Chvátal-Gomory inequalities are the inequalities

$$\lceil \lambda^T A^1 x \rceil \geq \lceil \lambda^T b^1 \rceil, \quad \lambda \geq 0 \quad (70)$$

where the matrix A^1 is such that for each i'th row A_i^1 of the matrix, the inequality $(A_i^1)^T x \geq b_i^1$ is a Chvátal-Gomory inequality of rank ≤ 1. In general, the rank r Chvátal-Gomory inequalities are the inequalities of the form

$$\lceil \lambda^T A^{r-1} x \rceil \geq \lceil \lambda^T b^{r-1} \rceil, \quad \lambda \geq 0 \tag{71}$$

where the matrix A^{r-1} is such that for each i'th row A_i^{r-1} of the matrix, the inequality $(A_i^{r-1})^T x \geq b_i^{r-1}$ is a Chvátal-Gomory inequality of rank $\leq r-1$. The rank r Chvátal-Gomory closure $\bar{P}^{C-G(r)}$ of $\bar{P}$ is the set of points that satisfies every Chvátal-Gomory inequality of rank $\leq r$, and obviously

$$Conv(P) \subseteq \bar{P}^{C-G(r)} \subseteq \bar{P}^{C-G(r-1)} \subseteq \cdots \subseteq \bar{P}^{C-G(0)} = \bar{P}. \tag{72}$$

Given a set covering problem

$$(SC) \;:\; \min \left\{ c^T x : Ax \geq e, \; x \in \{0,1\}^n \right\}, \tag{73}$$

denote the feasible region as P, and denote the continuous relaxation $\{x \in [0,1]^n : Ax \geq e\}$ as $\bar{P}$. Define now

$$\bar{c}^r(SC) := \min \left\{ c^T x : x \in \bar{P}^{C-G(r)} \right\} \tag{74}$$

$$\hat{c}^k(SC) := \min \left\{ c^T x : x \in P^k \right\} \tag{75}$$

$$\tilde{c}^k(SC) := \min \left\{ c^T x : x \in P^{A_k} \right\} \tag{76}$$

$$c^*(SC) := \min \left\{ c^T x : x \in P \right\}. \tag{77}$$

It can be shown ([BZ03]) that for each fixed $\epsilon > 0$, and each fixed integer r, there exists a fixed integer k such that for every set covering problem (SC),

$$\hat{c}^k(SC) \geq (1-\epsilon)\bar{c}^r(SC), \tag{78}$$

so that

$$c^*(SC) \geq \tilde{c}^k(SC) \geq \hat{c}^k(SC) \geq (1-\epsilon)\bar{c}^r(SC). \tag{79}$$

Thus for each fixed r and ϵ, we can use the algorithm to find in polynomial time a lower bound on $c^*(SC)$ that can be no worse than a factor of $1-\epsilon$ times the lower bound $\bar{c}^r(SC)$ provided by the rank r Chvátal-Gomory closure.

One final point that should be noted is that while we have referred to the algorithms of Chapters 5 and 6 as producing systems of linear constraints, semidefinitely constrained versions of the algorithsm can also be defined. These algorithms are in fact intended to maximize the effect of the semidefinite constraints, as we will see in Chapter 4 that the effectiveness of positive semidefiniteness depends on how well the appended variables and their associated sets characterize the structure of the feasible region. A formal study of how to take full advantage of positive semidefiniteness in the context of the framework developed in Chapters 3 and 4 remains an object for future research, but we will begin to show in Chapter 6 how the specific structure of the algorithms of that chapter contributes to the effectiveness of the semidefinite constraints.

Most importantly, however, these algorithms and the theory that underlies them point the way to a different paradigm for addressing integer and logical programs.

Road Map

Though each chapter builds on the material of the previous chapters, the chapters are also to a large extent self-contained, and one can in principle read much of the text out of order. The first chapter of the thesis is an overview of earlier material and though our presentation of that material differs in a several respects from their original presentations, the reader familiar with that material can skip that chapter.

The second chapter reinterprets the older material and motivates the new lifting, though once the new lifting is introduced in Chapter 3 there is very little further reference to Chapter 2.

Sections 3.1.1, 3.2, 3.4 and 3.6.2 of Chapter 3 introduce and describe the basic properties of the larger lifting that are used most extensively in the following chapters.

Chapter 4 discusses positive semidefiniteness in detail in the context of the larger lifting, though the only extensive use of positive semidefiniteness in the subsequent chapters is in Section 6.6, and there also the only required prior reading is Section 4.1.1.

Chapters 5 and 6 present two different classes of algorithms and they are largely independent of one another, though the aforementioned sections (at least) of Chapters 3 and 4 should be read first. Within Chapter 5 itself, the main algorithm is described in the beginning of Section 5.4 and the pitch k result is proven in Section 5.5, though to properly understand the algorithm one ought to read Sections 5.1 and 5.2.1 – 5.2.4 at least. The main algorithm of Chapter 6 along with the associated pitch k result is located in Section 6.4.1, but Sections 6.1 – 6.3 are recommended prior reading. Section 6.6 contains a positive semidefiniteness result for the algorithms of Chapter 6, and as noted above, one ought to read Section 4.1.1 first.

Chapter 1

A Survey of Lift and Project Operators

1.1 Convexification

1.1.1 Basic Concepts

Given a set $P \subseteq \{0,1\}^n$, the vector $x \in R^n$ belongs to its convex hull, $Conv(P)$, if and only if it can be written as a convex combination of points in P. Thus

$$x \in Conv(P) \text{ iff there exist numbers } \lambda_p \geq 0 \text{ for each } p \in P \text{ such that}$$

$$x = \sum_{p \in P} \lambda_p p, \text{ where } \sum_{p \in P} \lambda_p = 1. \tag{1.1}$$

This expression is a linear system, with variables $\{\lambda_p : p \in P\}$ and x, that describes $Conv(P)$ explicitly, though the description relies on having a list of the points of P in advance.

Assume now that the set P is the set of integer points belonging to some polytope $\bar{P} \subseteq [0,1]^n$ (i.e. $P = \bar{P} \cap \{0,1\}^n$). Then there exists an $m \times n$ matrix A and a vector b such that for each $y \in \{0,1\}^n$, $y \in P$ iff $Ay \geq b$. It follows then that for any number $\lambda_y > 0$,

$$y \in P \text{ iff } A(\lambda_y y) \geq \lambda_y b. \tag{1.2}$$

This fact will allow us to remove the dependency on P from the summation above, as the only points $y \in \{0,1\}^n$ for which positive numbers λ_y can exist that satisfy $A(\lambda_y y) \geq \lambda_y b$ are those that belong to P. (For those $y \notin P$, demanding that $A(\lambda_y y) \geq \lambda_y b$ ensures that $\lambda_y = 0$.) We can therefore rewrite the above statement as follows

Lemma 1.1 *Where $P \subseteq \{0,1\}^n$ and $x \in R^n$, $x \in Conv(P)$ iff there exist numbers $\lambda_y \geq 0$ for each $y \in \{0,1\}^n$ such that*

$$x = \sum_{y\in\{0,1\}^n} \lambda_y y, \text{ where } A(\lambda_y y) \geq \lambda_y b, \text{ and } \sum_{y\in\{0,1\}^n} \lambda_y = 1. \ \square \tag{1.3}$$

This is an explicit linear system in the $2^n + n$ variables $x_1, \ldots, x_n, \{\lambda_y : y \in \{0,1\}^n\}$ with $m(2^n)+2$ constraints, and the projection of the solution set on the variables $x_1, \ldots, x_n$ gives the convex hull of P exactly.

Obviously this system is too large for us to want to deal with, but the method used to generate it can be relaxed, and this is the idea behind "convexification".

Convexification seeks to ensure that x can be written as a convex combination of points (each of which satisfy $Ay \geq b$) that have values of zero or one in some fixed sized subset of their coordinates. It is easy to see that if the size of this subset is a fixed constant then the resulting linear system will be of a size polynomial in the size of the linear system that defined $\bar{P}$. Typically the fixed subset of coordinates will be some single coordinate i, and the procedure will thus attempt to decompose any x for which $0 < x_i < 1$ into a convex combination of the two vectors v and w, each of which satisfy $Ay \geq b$, and such that $v_i = 1$ and $w_i = 0$. More precisely, it will look to decompose any x $(0 \leq x_i \leq 1)$ into the sum of two vectors $\hat{v}$ and $\hat{w}$ such that

$$A\hat{v} \geq \lambda b \text{ and } A\hat{w} \geq (1-\lambda)b \tag{1.4}$$

and such that

$$\hat{v}_i = \lambda \quad \text{and} \quad \hat{w}_i = 0. \tag{1.5}$$

(Note that if we are to have $x = \lambda v + (1-\lambda)w$ we must have $\lambda = x_i$. Note also that where $x_i = 1$ then $\lambda = 1$, and the decomposition is $\hat{v} = x, \ \hat{w} = 0$, and similarly where $x_i = 0$ then $\lambda = 0$, and the decomposition is $\hat{v} = 0, \ \hat{w} = x$.) Equivalently, it attempts to decompose the $n+1$ dimensional vector $(1, x)$ into the sum of two vectors $(\hat{v}_0, \hat{v})$ $(\hat{w}_0, \hat{w})$ with $\hat{v}_0 = \hat{v}_i = x_i$ (and therefore $\hat{w}_i = 0$ and $\hat{w}_0 = 1 - \hat{v}_0$) such that each of which satisfy the homogenized system

$$(-b \mid A) \binom{y_0}{y} \geq 0. \tag{1.6}$$

To this end append $n+1$ new variables forming the new vector $(\hat{v}_0, \hat{v})$, with

$$\hat{v}_0 = \hat{v}_i = x_i \text{ and } (-b \mid A) \binom{\hat{v}_0}{\hat{v}} \geq 0 \tag{1.7}$$

and demand that

$$(-b \mid A) \left(\binom{1}{x} - \binom{\hat{v}_0}{\hat{v}} \right) \geq 0 \tag{1.8}$$

(i.e. $x - \hat{v}$ qualifies as the vector $\hat{w}$).

This can be made more general, but first let us suggest the following definition.

Definition 1.2 *Given $P \subseteq \{0,1\}^n$, and any convex set $Q \subseteq [0,1]^n$ such that $Q \cap \{0,1\}^n = P$, define $\bar{K}(Q) \subset R^{n+1}$ by*

$$\bar{K}(Q) = Cone\Big(\{x \in [0,1]^{n+1} : x_0 = 1, (x_1, \ldots, x_n) \in Q\}\Big) \tag{1.9}$$

and define $K(Q) := K(P) \subset \{0,1\}^{n+1}$ by

$$K(P) = \{y \in \{0,1\}^{n+1} : y_0 = 1, (y_1, \ldots, y_n) \in P\} \tag{1.10}$$

Note that

$$\bar{K}(Conv(P)) = Cone(K(P)), \tag{1.11}$$

that $\bar{K}(Q) \cap \{0,1\}^{n+1} = K(P) \cup \{0\}$, and that a polynomial time separation oracle exists for $\bar{K}(Q)$ if and only if one exists for Q. Note also that there is a one to one correspondence between the convex sets $Q \subseteq [0,1]^n$ with $Q \cap \{0,1\}^n = P$ and the cones $\bar{K} \subseteq Cone(y \in \{0,1\}^{n+1} : y_0 = 1)$ with $\bar{K} \cap \{0,1\}^{n+1} = K(P) \cup \{0\}$, via the functions $\bar{K}(Q)$ (of Definition 1.2) and $Q(\bar{K}) := \{x \in [0,1]^n : (1,x) \in \bar{K}\}$. Before we continue let us also point out that Lemma 1.1 can be recast somewhat more cleanly in a conic framework as follows.

Lemma 1.3 *Where $P \subseteq \{0,1\}^n$, $Q \subseteq [0,1]^n$ and $Q \cap \{0,1\}^n = P$, then $(x_0, x) \in R^{n+1}$ belongs to $Cone(K(P))$ iff there exist numbers $\lambda_y \geq 0$ for each $y \in \{0,1\}^n$ such that*

$$(x_0, x) = \sum_{y \in \{0,1\}^n} \lambda_y(1, y), \text{ where } \lambda_y(1, y) \in \bar{K}(Q) \text{ for all } y. \ \Box \tag{1.12}$$

We are now in a position to give a more general definition for convexification. Given a point $x \in Q$, convexification seeks to decompose $(1, x)$ into two vectors $(\hat{v}_0, \hat{v})$ and $(\hat{w}_0, \hat{w})$ such that $\hat{v}_0 = \hat{v}_i = x_i$ and such that each of these two vectors belong to $\bar{K}(Q)$. Formally,

Definition 1.4 *For any convex set $Q \subseteq [0,1]^n$, define the "convexification operators with respect to coordinate i" as follows:*

$$C_i(Q) = \{x \in Q : \text{ either } x_i \in \{0,1\}, \text{ or }$$

$$\exists v, w \in Q \text{ s.t. } x_i v + (1 - x_i) w = x, \ v_i = 1, w_i = 0\} \tag{1.13}$$

and where x and $\hat{v}$ are each construed as $n+1$ dimensional vectors,

$$M_i(\bar{K}(Q)) = \{(x, \hat{v}) \in R^{2n+2} : \hat{v}_0 = \hat{v}_i = x_i, \ \hat{v} \in \bar{K}(Q), \ x - \hat{v} \in \bar{K}(Q)\} \tag{1.14}$$

and

$$N_i(\bar{K}(Q)) = \{x \in R^{n+1} : \exists (x, \hat{v}) \in M_i(\bar{K}(Q))\}. \tag{1.15}$$

Note that the convexity of Q implies that for any $(x_0, \ldots, x_n) \in N_i(\bar{K}(Q))$ for which $x_0 = 1$ we must have $(x_1, \ldots, x_n) \in Q$, and therefore, since $N_i(\bar{K}(Q))$ is a cone, we conclude that $N_i(\bar{K}(Q)) \subseteq \bar{K}(Q)$. The subset of Q that satisfies the convexification requirement on coordinate i is $C_i(Q)$, and this is the projection of $M_i(\bar{K}(Q))$ on its $1, \ldots, n$ coordinates, intersected with the hyperplane $x_0 = 1$. The set $N_i(\bar{K}(Q))$ is the projection of $M_i(\bar{K}(Q))$ on its first $n+1$ coordinates, and therefore $C_i(Q)$ is just the projection of $N_i(\bar{K}(Q)) \cap \{x \in R^{n+1} : x_0 = 1\}$ on its $1, \ldots, n$ coordinates. Since $N_i(\bar{K}(Q))$ is a cone and a subset of $Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1\})$ (so that a point $x \in N_i(\bar{K}(Q))$ can be such that $x_0 = 0$ only if $x = 0$) it is easy to conclude that

$$N_i(\bar{K}(Q)) = \bar{K}(C_i(Q)) \tag{1.16}$$

(as per Definition 1.2). Both $M_i(\bar{K}(Q))$ and $N_i(\bar{K}(Q))$ are cones, and they are easier sets to work with than is $C_i(Q)$. In later sections we will therefore be focusing primarily on them. In general we will also be writing M_i and N_i as functions of sets of the form $\bar{K} \subseteq Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1\})$, and we will suppress the dependence on Q.

Another equivalent way to view the convexification operator, and this is the context in which it was originally developed by Balas ([B74]), is in the setting of "disjunctive programming". The feasible set P is the disjoint union

$$P = (P \cap \{y \in \{0,1\}^n : y_i = 1\}) \ \cup \ (P \cap \{y \in \{0,1\}^n : y_i = 0\}) \tag{1.17}$$

and therefore every point $x \in Conv(P)$ can be written as a convex combination of points

$$v \in Conv(P \cap \{y \in \{0,1\}^n : y_i = 1\}) \text{ and } w \in Conv(P \cap \{y \in \{0,1\}^n : y_i = 0\}). \tag{1.18}$$

(More generally, where $P = \bigcup_i P_i$ then any $x \in Conv(P)$ can be written as a convex combination of points x^i, each of which belongs to $Conv(P_i)$, and conversely.) Obviously v and w must both satisfy every necessity condition of P, and we must have $v_i = 1$ and $w_i = 0$. (This interpretation will play a significant role in the algorithms to be introduced later.)

We have seen that where Q is a polyhedral set, then so is $M_i(\bar{K}(Q))$. This is the case that interests us most, but it should be noted that a parallel result holds for any convex set $Q : Q \cap \{0,1\}^n = P$ for which a polynomial time separation oracle exists.

The following is an adaptation of a lemma of Lovász and Schrijver.

Lemma 1.5 *Let $Q \subseteq [0,1]^n$ be a convex set for which a polynomial time separation oracle exists, then polynomial time separation oracles exist for $M_i(\bar{K}(Q))$, $N_i(\bar{K}(Q))$ and for $C_i(Q)$.*

Proof: We have already noted that Q has a polynomial time separation oracle if and only if $\bar{K}(Q)$ has one also. Consider a candidate point $(x', \hat{v}') \in R^{2n+2}$ for membership in $M_i(\bar{K}(Q))$. Check first if $\hat{v}'_i = \hat{v}'_0 = x'_i$. If this fails then we trivially obtain a separating hyperplane. Otherwise, run $\bar{K}(Q)$'s separation oracle on x', on $\hat{v}'$, and on $x' - \hat{v}'$. They all pass iff $(x', v') \in M_i(\bar{K}(Q))$. If $(x', \hat{v}') \notin M_i(\bar{K}(Q))$, then at least one of these must fail and return a hyperplane separating the point that failed from $\bar{K}(Q)$, i.e. it must return some inequality $a^T y \geq \beta$ satisfied by every point in $\bar{K}(Q)$ but violated by that point. Clearly any inequality that x' or $\hat{v}'$ must satisfy for membership in $\bar{K}(Q)$ must also be satisfied by $(x, \hat{v}')$ for membership in $M_i(\bar{K}(Q))$. It is also clear that for any point $(x, \hat{v})$ to belong to $M_i(\bar{K}(Q))$ it must also satisfy $a^T(x - \hat{v}) \geq \beta$, and this expression can be trivially recast as a valid inequality for $M_i(\bar{K}(Q))$. Thus the failure of $x' - \hat{v}'$ to belong to $\bar{K}(Q)$ also yields a violated valid inequality for $M_i(\bar{K}(Q))$. We conclude that a polynomial time separation oracle exists for $M_i(\bar{K}(Q))$, and by general results (see [GLS81]) a polynomial time separation oracle must also exist for $N_i(\bar{K}(Q))$ as it is a projection of $M_i(\bar{K}(Q))$. Finally, if a polynomial time separation oracle exists for $M_i(\bar{K}(Q))$ then it is easy to see that one exists for $M_i(\bar{K}(Q)) \cap \{x \in R^{2n+2} : x_0 = 1\}$, and $C_i(Q)$ is just a projection of this set. □

One more trivial detail that ought to be pointed out explicitly is that convexification does not cut off any points from $Conv(P)$.

Lemma 1.6 *Let Q be a convex set in $[0,1]^n$ such that $Q \cap \{0,1\}^n = P$, and let $C_i(Q)$ be defined as above, then*

$$Conv(P) \subseteq C_i(Q) \subseteq Q \tag{1.19}$$

Proof: It is clear that $C_i(Q)$ preserves convexity as a convex combination of points in $C_i(Q)$ must belong to Q by the convexity of Q, and must also satisfy that it is a convex combination of points in Q with a zero or one in coordinate i, as it is defined to be a convex combination of points that are themselves convex combinations of points with a zero or one in the i coordinate. It is also clear that every point in P must belong to $C_i(Q)$ since $P \subseteq Q$ and any $p \in P$ trivially satisfies the convexification requirement. The lemma follows. □

It should also be noted that though we have stated these results with respect to the operator C_i and sets of the form $Q \subseteq [0,1]^n$, it is easy to recast them in terms of N_i and sets of the form $\bar{K} \subseteq Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1\})$ as well (recall Definitions 1.2 and 1.4). The following two lemmas are analogs of Lemmas 1.5 and 1.6.

Lemma 1.7 *Let $K \subseteq \{y \in \{0,1\}^{n+1} : y_0 = 1\}$, and let $\bar{K} \subseteq Cone(\{y \in \{0,1\}^{n+1} : y_0 =$*

1}) be a cone satisfying $\bar{K} \cap \{0,1\}^{n+1} = K \cup \{0\}$, then for any $i = 1, \ldots, n$,

$$Cone(K) \subseteq N_i(\bar{K}) \subseteq \bar{K}. \; \square \tag{1.20}$$

Lemma 1.8 *Let $\bar{K} \subseteq Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1\})$ be a cone for which a polynomial time separation algorithm exists, then for any $i = 1, \ldots, n$, a polynomial time separation algorithm exists for $N_i(\bar{K})$ as well.* $\square$

If the convexification procedure is performed simultaneously for all n coordinates, i.e. we demand that for each fractional coordinate j there are points v^j with $v^j_j = 1$ and w^j with $w^j_j = 0$ belonging to Q such that

$$x = x_j v^j + (1 - x_j) w^j \tag{1.21}$$

then we obtain an operator that we will refer to as C^0. (This is essentially the same as the N^0 operator of Lovász and Schrijver [LS91], which will also be defined here). Formally,

Definition 1.9 *Given a convex set $Q \subseteq [0,1]^n$, the set $C^0(Q)$ is the set of $x \in Q$ such that for each $i = 1, \ldots, n$, either $x_i \in \{0,1\}$ or there exists a vector $v^i \in R^n$ satisfying*

$$v^i_i = 1 \tag{1.22}$$

$$v^i \in Q \tag{1.23}$$

$$\frac{x - x_i v^i}{1 - x_i} \in Q. \tag{1.24}$$

As above, where $\bar{K}$ is understood to mean $\bar{K}(Q)$, we will define $N^0(\bar{K})$ to be the set of vectors $x \in R^{n+1}$ for which there exist n vectors $\hat{v}^i \in R^{n+1}, \; i = 1, \ldots, n$ such that

$$\hat{v}^i_0 = \hat{v}^i_i = x_i, \; i = 1, \ldots, n \tag{1.25}$$

$$\hat{v}^i \in \bar{K}, \; i = 1, \ldots, n \tag{1.26}$$

$$x - \hat{v}^i \in \bar{K}, \; i = 1, \ldots, n. \tag{1.27}$$

As above, $C^0(Q)$ is the projection of $N^0(\bar{K}) \cap \{x \in R^{n+1} : x_0 = 1\}$ on the $1, \ldots, n$ coordinates.

Stated in another way,

Lemma 1.10

$$N^0(\bar{K}) = \{x \in R^{n+1} : x = Ye_0\} \tag{1.28}$$

where e_0 *is the unit vector for the zero coordinate and* Y *is an* $n+1 \times n+1$ *matrix satisfying*

$$Y_{j,0} = Y_{j,j} = Y_{0,j} \tag{1.29}$$

$$Ye_j \in \bar{K} \ , \ \forall j = 1, \ldots, n \tag{1.30}$$

$$Y(e_0 - e_j) \in \bar{K} \ , \ \forall j = 1, \ldots, n. \tag{1.31}$$

Proof: The zero column of Y is the vector x, and the j'th column of Y is the vector $\hat{v}^j$. □

In parallel to the definition of M_i given above, we give the following definition (again following Lovász and Schrijver [LS91]).

Definition 1.11 *Define the set* $M^0(\bar{K})$ *as the set of* $n+1 \times n+1$ *matrices* Y *satisfying constraints 1.29, 1.30, and 1.31. Note that the projection of* $M^0(\bar{K})$ *on its zero'th column's coordinates is* $N^0(\bar{K})$*.*

1.1.2 Repeated Convexification

To give some geometric insight into the meaning of convexification, consider a diagram of a possible choice of $\{v^i\}$ and $\{w^i\}$ for a given x in two dimensions.

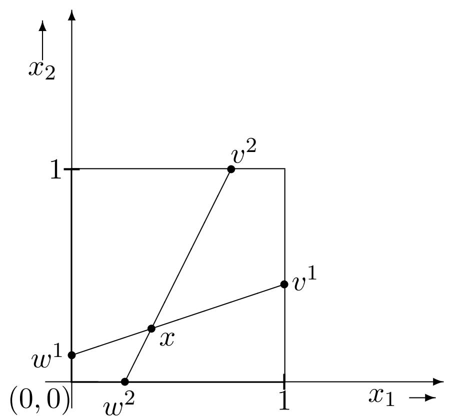

Figure 1

The simultaneous convexification operator $C^0(Q)$ will only allow choices of v^i and w^i that themselves belong to the set Q. If it is impossible to draw such lines with endpoints in Q then the point x will be eliminated.

Naturally, the procedure can be repeated for the new vectors v^j and w^j. For each coordinate i, we can demand that there exist vectors $v^i(v^j)$ and $w^i(v^j)$ in Q such that $[v^i(v^j)]_i = 1$ and $[w^i(v^j)]_i = 0$ and such that v^j can be written as a convex combination of the two. We can do something similar for w^j as well. It is important to notice that

$$[v^i(v^j)]_j = [w^i(v^j)]_j = 1 \tag{1.32}$$

since otherwise the convex combination of the two could not have a j coordinate equal to 1. Thus these new vectors will have a zero or one in at least two of their coordinates. Doing this repeatedly we would thus expect to eventually obtain x as a convex combination of points that are $0, 1$ in all of their coordinates. This is in fact the case and we will have more to say about this later. For the meantime, however, we shall use this idea to prove a stronger result that does not depend on simultaneous convexification. The following theorem is due to [BCC93], but with a different proof.

Theorem 1.12 *Let $P \subseteq \{0,1\}^n$, let $Q \subseteq [0,1]^n$ be a convex set for which $Q \cap \{0,1\}^n = P$, and let $C_i(Q)$ be as in Definition 1.4, then*

$$C_n(C_{n-1}(\cdots C_1(Q))\cdots) = Conv(P). \tag{1.33}$$

Proof: Define

$$C^j(Q) = C_j(C_{j-1}(\cdots C_1(Q))\cdots). \tag{1.34}$$

We know that for all $x \in C^1(Q)$, x can be written as a convex combination of points in Q for which their first coordinate is 0 or 1. Assume now that all points in $C^j(Q)$ can be written as convex combinations of points in Q having 0 or 1 in their first j coordinates, and consider

$$C^{j+1}(Q) = C_{j+1}(C^j(Q)). \tag{1.35}$$

By definition, all points in this set can be written as convex combinations of points in $C^j(Q)$ with zeroes or ones in their $j+1$ coordinates. But any such point v with, say $v_{j+1} = 1$, can itself be written as a convex combination of points from Q with zeroes or ones in their first j coordinates. Moreover, all of those points must have a 1 in their $j+1$ coordinate, or else v, which is a convex combination of those points could not have a 1 in its $j+1$'st coordinate. A parallel statement holds if $v_{j+1} = 0$. We thus conclude that all points in $C^{j+1}(Q)$ can be written as convex combinations of points from Q with zeroes or ones in their first $j+1$ coordinates. The theorem follows by induction. □

The conic analog of Theorem 1.12 is as follows.

Theorem 1.13 *Let $P \subseteq \{0,1\}^n$, let $Q \subseteq [0,1]^n$ be a convex set for which $Q \cap \{0,1\}^n = P$, let $C_i(Q)$ be as in Definition 1.4, and let $\bar{K}(Q)$ and $K(P)$ be as in Definition 1.2, then*

$$N_n(N_{n-1}(\cdots N_1(\bar{K}(Q)))\cdots) = Cone(K(P)). \tag{1.36}$$

Proof: This follows from Theorem 1.12 and from expressions (1.11) and (1.16). □

Consider, for example the polytope Q defined as follows:

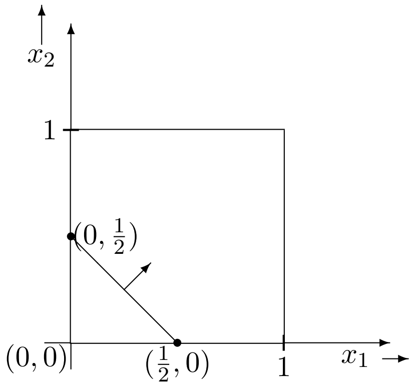

Figure 2a

The set $C_1(Q)$ will be

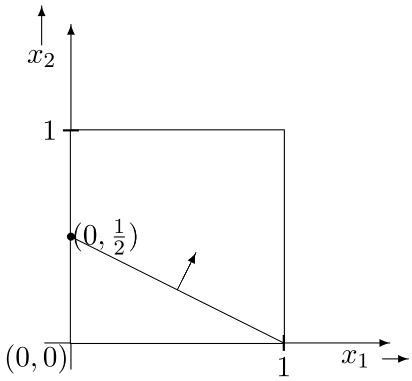

Figure 2b

The set $C^2(Q) = C_2(C_1(Q))$ will be

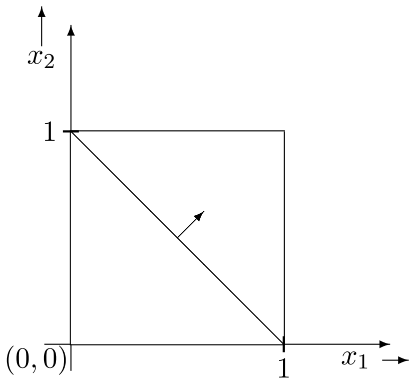

Figure 2c

and this is the convex hull of $Q \cap \{0,1\}^n$.

1.1.3 Beyond Convexification

As was noted earlier, the conic interpretation of convexification is the easier one to work with, and from here on our focus will shift to the conic framework. The fundamental insight that allows one to go beyond mere convexification is that where $K \subseteq \{y \in \{0,1\}^{n+1} : y_0 = 1\}$, and $\bar{K} \subseteq Cone(y \in \{0,1\}^{n+1} : y_0 = 1\}$ is such that $\bar{K} \cap \{0,1\}^{n+1} = K \cup \{0\}$, and where the matrix Y is as defined in Lemma 1.10, there are simple but nonobvious ways to further constrain the matrix Y without cutting off any points from $Cone(K)$. The proof of the following theorem is trivial, but it has a deeper meaning that we will discuss in the next section and chapter.

Theorem 1.14 *Let $K \subseteq \{y \in \{0,1\}^{n+1} : y_0 = 1\}$, and let $\bar{K} \subseteq Cone(y \in \{0,1\}^{n+1} : y_0 = 1\}$ be such that $\bar{K} \cap \{0,1\}^{n+1} = K \cup \{0\}$, and let the matrix Y be as in Lemma 1.10. The following two additional constraints on the matrix Y are consistent with all points in $Cone(K)$,*

1. *Y is symmetric.*
2. *Y is positive semidefinite.*

Proof: For any point $x \in K$, let

$$Y = xx^T. \tag{1.37}$$

Under this choice of Y, the vector x satisfies all the requirements for membership in $N^0(\bar{K})$ as well as the two additional constraints of the theorem. These two constraints (as well as the $N^0(\bar{K})$ constraints) are thus consistent with every point of K, and therefore with every point of $Cone(K)$ as well (since both are preserved under nonnegative combinations, as are the $N^0(\bar{K})$ constraints). □

Definition 1.15 *Define the set $M(\bar{K})$ to be the set of symmetric $n+1 \times n+1$ matrices Y with $Ye_o = Diag(Y)$ satisfying*

$$Ye_i \in \bar{K},\ Y(e_0 - e_i) \in \bar{K},\ \forall i = 1, \ldots, n. \tag{1.38}$$

Define $N(\bar{K})$ to be the set

$$\{x \in R^{n+1} : x = Ye_0,\ Y \in M(\bar{K})\}. \tag{1.39}$$

Define the set $M^+(\bar{K})$ to be the set of matrices in $M(\bar{K})$ that satisfy the additional requirement,

$$Y \succeq 0. \tag{1.40}$$

Define the set $N^+(\bar{K})$ to be the set

$$\{x \in R^{n+1} : x = Ye_0,\ Y \in M^+(\bar{K})\}. \tag{1.41}$$

Geometrically, these extra requirements on the matrix Y say that the lines that we drew in Figure 1 cannot be so freely pivoted. Before we demonstrate this explicitly, we will review first the geometric meaning of the matrices Y. Observe from Definition 1.9 and Lemma 1.10 that the simultaneous convexification requirements for a point $x \in R^n$ with respect to a convex set $Q \subseteq [0,1]^n$ are equivalent to the requirement that there exist an $n+1 \times n+1$ matrix Y with $Ye_0 = (1, x)$, satisfying

$$Y_{j,0} = Y_{j,j} = Y_{0,j} \tag{1.42}$$

$$Ye_j \in \bar{K}(Q)\ ,\ \forall j = 1, \ldots, n \tag{1.43}$$

$$Y(e_0 - e_j) \in \bar{K}(Q)\ ,\ \forall j = 1, \ldots, n. \tag{1.44}$$

Specifically, given a matrix Y with $Ye_0 = (1, x)$, satisfying constraints (1.42) - (1.44), for each $j \in \{1, \ldots, n\}$ such that $0 < x_j < 1$, define the vectors v^j and w^j in R^n by

$$Ye_j = x_j(1, v^j),\ \ Y(e_0 - e_j) = (1 - x_j)(1, w^j). \tag{1.45}$$

Then

$$x = x_j v^j + (1 - x_j)w^j,\ \forall j \in \{1, \ldots, n\} \text{ s.t. } 0 < x_j < 1 \tag{1.46}$$

and constraints (1.42) - (1.44) say that $x \in Q$, and that for each $j = 1, \ldots, n$ for which $0 < x_j < 1$,

$$v^j_j = 1,\ w^j_j = 0,\ v^j, w^j \in Q. \tag{1.47}$$

Conversely, for any $x \in Q$ that can be written as the convex combination

$$x = x_j v^j + (1 - x_j) w^j \tag{1.48}$$

of vectors $v^j, w^j \in Q$ with $v^j_j = 1,\ w^j_j = 0$, for each $j \in \{1, \ldots, n\}$ such that $0 < x_j < 1$, then the column vectors defined by $Ye_0 = (1, x),\ Ye_j = x_j(1, v^j)$ for each j such that $0 < x_j < 1$, by $Ye_j = 0$ where $x_j = 0$, and by $Ye_j = (1, x)$ where $x_j = 1$, satisfy constraints (1.42) - (1.44).

The matrices Y with $Y_{0,0} = 1$ that satisfy constraints (1.42) - (1.44), where we construe the column Ye_0 of the matrix as the vector $(1, x)$, are thus the matrices whose columns are the vectors $x_j(1, v^j)$ for some choice of vectors v^j defining a valid simultaneous convexification of x (and whose j'th column is $(1, x)$ where $x_j = 1$, and 0 where $x_j = 0$). Turning our attention now to Figure 1, we indicated there that given a convex set $Q \subseteq [0, 1]^2$, the simultaneous convexification requirements for the point x drawn in that diagram were that there be a way to select points v^1 with $v^1_1 = 1$, w^1 with $w^1_1 = 0$, v^2 with $v^2_2 = 1$ and w^2 with $w^2_2 = 0$, all of which belong to Q, and such that x lies on the line between v^1 and w^1 as well as on the line between v^2 and w^2. It is clear from the diagram that a choice of v^1 fixes the choice of w^1, as does the choice of v^2 fix the choice of w^2, but notice that the choices of v^1 and v^2 are independent. Notice now that, as above, in terms of the matrix Y, the column $Ye_1 = x_1(1, v^1)$, and the column $Ye_2 = x_2(1, v^2)$. Thus if we were to also impose the symmetry requirement on the matrix Y, we will therefore have $Y_{1,2} = Y_{2,1}$, or in terms of the vectors v^i,

$$v^2_1 = \left(\frac{x_1}{x_2}\right) v^1_2. \tag{1.49}$$

Thus (since we are already given $v^1_1 = 1 = v^2_2$,) the choice of v^1_2 will fix all four points v^1, v^2, w^1, w^2. So for example, where $x = (3/8, 1/4)$ and we choose (arbitrarily) $v^1_2 = 1/2$, then we must have

$$v^2_1 = \left(\frac{3}{8}/\frac{1}{4}\right) \times \frac{1}{2} = \frac{3}{4}. \tag{1.50}$$

The following figure depicts the consequent points v^1, v^2, w^1, w^2. Obviously to require $Y \succeq 0$ will restrict the choices still further.

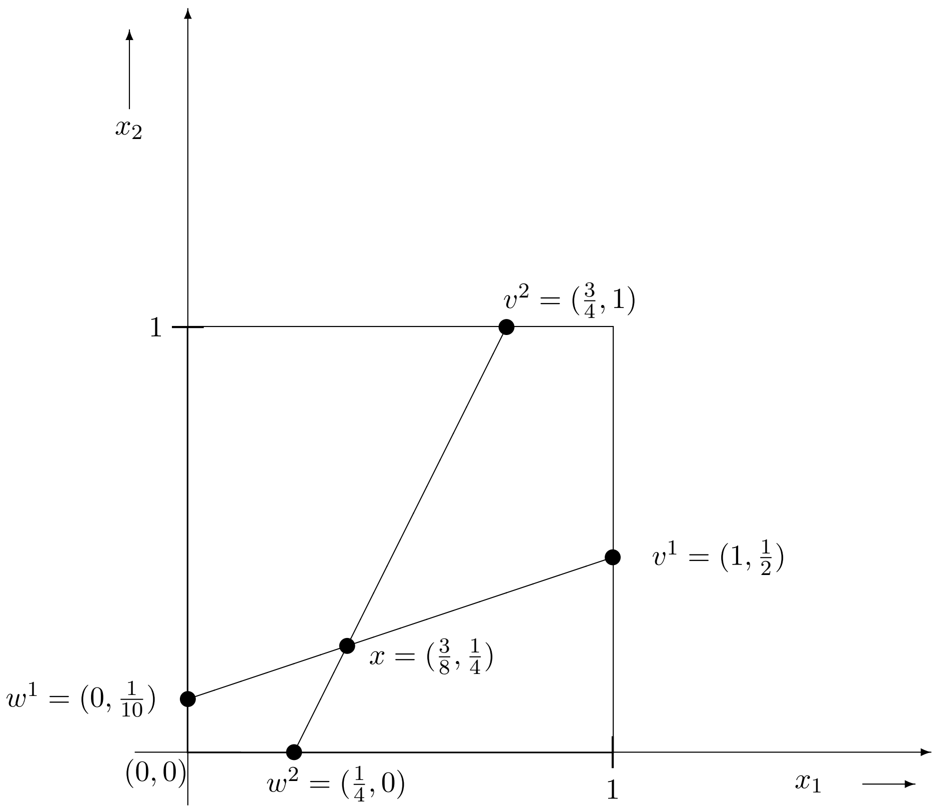

Figure 3

1.2 The N, N^+ and $\bar{N}$ Operators

The N and N^+ operators were introduced by Lovász and Schrijver. The $\bar{N}$ operator, as it will be described in this section, is the Sherali Adams operator applied to sets $P \subseteq \{0,1\}^n$ that are defined by a system of linear constraints. In this section we will describe the Lovász and Schrijver interpretation of these three operators. Lasserre's algorithm applies to general polynomial programs, of which $0, 1$ integer programs are a special case (i.e. $x_i^2 = x_i,\ i = 1, \ldots, n$). We will also begin to show here that Lasserre's algorithm as applied to $0, 1$ integer programs can be seen as a generalization of the N^+ and $\bar{N}$ operators. In the next chapter we will reanalyze these operators from a different perspective. The original interpretation of the Sherali Adams operator will be indicated there, and the Lasserre algorithm will also be examined more formally. The treatment we will be giving here of these operators parallels that given by Lovász and Schrijver in Section 3 of their paper, but with changes in presentation and proofs.

1.2.1 The Lattice L

Definition 1.16 *A partially ordered set T such that for any two elements $p, q \in T$ there exists a unique least upper bound, and a unique greatest lower bound (both in T) is called a lattice. The least upper bound of p and q is denoted $p \vee q$ and is referred to as the "join" of p and q, and the greatest lower bound is denoted $p \wedge q$, and is referred to as the "meet" of p and q (see [Ro64]).*

Lemma 1.17 *Let S be a set containing n elements $\{s_1, \ldots, s_n\}$, then the powerset $\mathcal{P}(S)$ partially ordered by inclusion is a lattice. We will refer to this lattice as L.*

Proof: Given $A \subseteq S$ and $B \subseteq S$, then $A \cup B \subseteq S$ and is an upper bound on both A and B (partial ordering by inclusion) as it includes both of them. Let $C \subseteq T$ be such that $C \geq A,\ C \geq B$, then $A \subseteq C,\ B \subseteq C \Rightarrow A \cup B \subseteq C \Rightarrow A \cup B \leq C$. The proof for meets is similar. □

The 2^n elements in $\{0,1\}^n$ can be thought of as incidence vectors for sets in $\mathcal{P}(S)$, where a set

$$B = \{s_i : i \in \beta \subseteq \{1, \ldots, n\}\} \tag{1.51}$$

is represented by the vector $v \in \{0,1\}^n$ with a 1 in coordinates $i \in \beta$ and 0 elsewhere. Similarly the set $\{y \in \{0,1\}^{n+1} : y_0 = 1\}$ can also be thought of as the set of incidence vectors for the sets of $\mathcal{P}(S)$ with the zero coordinate corresponding to the empty set (so the incidence vector for every set will have a 1 in its zero'th coordinate).

Let $\bar{\bar{z}}^B$ be the vector in $\{y \in \{0,1\}^{n+1} : y_0 = 1\}$ that is the incidence vector for the set $B \subseteq S$. So B has a coordinate for the empty subset and each singleton subset of S that indicate whether or not B contains those singleton or empty sets as a subset. Consider now expanding $\bar{\bar{z}}^B$ as follows.

Definition 1.18 *Number the elements of $\mathcal{P}(S)$, i.e. establish a one-to-one correspondence between the subsets of S and the numbers $0, 1, \ldots, 2^n - 1$. For each $B \subseteq S$ define $z^B \in \{0,1\}^{2^n}$ to be the vector with a 1 in coordinate k iff the set $A_k \subseteq S$ corresponding to the number k is a subset of B*

Thus the vector that we called $\bar{\bar{z}}^B$ is the projection of z^B on its coordinates that correspond to the empty and singleton subsets.

Definition 1.19 *Given a countable partially ordered set T, and some one-to-one correspondence l between the elements of the set and the numbers $0, 1, \ldots, |T| - 1$ (if T is finite, otherwise the correspondence will be with the nonnegative integers), so that each set element is identified uniquely as l_i for some $i \in \{0, 1, \ldots, |T| - 1\}$, the zeta matrix Z of T is the $|T| \times |T|$ matrix defined by*

$$Z_{i,j} = \begin{cases} 1 : l_i \leq l_j \\ 0 : l_i \not\leq l_j \end{cases} \tag{1.52}$$

where the inequality refers to the partial ordering T.

The vectors z^B are thus just the columns of the zeta matrix of the lattice L. We will refer to these vectors as "zeta vectors".

We will now show that the zeta matrix Z of L is nonsingular. But for completeness, as this is true in greater generality for zeta matrices, we will prove a stronger lemma.

Lemma 1.20 *The zeta matrix of any countable locally finite partially ordered set T that contains a zero element is nonsingular.*[1]

Proof: "Locally finite" means that for any two elements $p, q \in T$, the interval $[p, q] = \{t \in T : p \leq t \leq q\}$ is a finite set. A zero element is an element that is less or equal to every element in T. Consider the following numbering procedure for T. Let l_0 be the zero element (Notice that every finite lattice has a zero, namely $\bigwedge_{t \in T} t$), and let $\{l_0, l_1, l_2 \ldots\}$ be a complete listing of T. Let $t_0 = l_0$. Reset $T := T - \{t_0\}$. Begin the following procedure with $i = \bar{i} = 1$.

1. Consider the intersection of the open interval (t_0, l_i) with T. By hypothesis this set is finite. If it is empty then select $t_{\bar{i}} = l_i$ and go to step (3). Otherwise, as a nonempty finite partially ordered set, $(t_0, l_i) \cap T$ must contain a minimal element r. Set $t_{\bar{i}} = r$.

2. Reset $T := T - \{r\}$, $\bar{i} := \bar{i} + 1$, and go to step (1).

3. Reset $T := T - \{l_i\}$, $\bar{i} := \bar{i} + 1$, and if there is a $k > i$ such that $l_k \in T$ then let $i := \min \{k > i : l_k \in T\}$ and go to step (1); otherwise stop.

[1] By nonsingular we mean that where we denote the zeta matrix as Z, there exists a unique matrix M of real numbers such that for all numberings of the rows and columns as $0, 1, 2, \ldots$, and for each pair of nonnegative integers i, j,

$$\sum_{k=0}^{\infty} M_{i,k} Z_{k,j} = \delta_{i,j} = \sum_{k=0}^{\infty} Z_{i,k} M_{k,j}. \tag{1.53}$$

It is easy to see that the sequence generated, $\{t_0, t_1, \ldots\}$, is a complete listing of T, and that by construction $i \leq j \Rightarrow t_i \not> t_j$. Thus if the rows and columns of the matrix are numbered according to $t_0, t_1, \ldots$, i.e. if we refer to Z_{t_i,t_j} as $Z_{i,j}$, then for $i \neq j$

$$Z_{i,j} = 1 \Rightarrow t_i < t_j \Rightarrow i < j \tag{1.54}$$

so the only nonzero entries in Z aside from those on the diagonal are those in the upper triangular part (by this numbering). Moreover $Z_{i,i} = 1, \forall i$, so Z is an upper triangular matrix with ones along the diagonal. It is now easy to construct the unique matrix M that satisfies

$$\sum_{k=0}^{\infty} M_{i,k} Z_{k,j} = \delta_{i,j}, \ \forall i, j. \tag{1.55}$$

This matrix M moreover is also upper triangular with ones along the diagonal, and thus by the same reasoning it is easy to construct the unique matrix Y that satisfies

$$\sum_{k=0}^{\infty} Y_{i,k} M_{k,j} = \delta_{i,j}, \ \forall i, j. \tag{1.56}$$

But Y is also upper triangular with ones along the diagonal, and therefore for each pair of nonnegative integers i, j with $j < i$, $Y_{i,j} = 0 = Z_{i,j}$, and if $j \geq i$ then

$$Y_{i,j} = \sum_{k=0}^{j} Y_{i,k}\delta_{k,j} = \sum_{k=0}^{j} Y_{i,k}\left(\sum_{l=0}^{\infty} M_{k,l} Z_{l,j}\right) = \sum_{k=0}^{j} Y_{i,k}\left(\sum_{l=0}^{j} M_{k,l} Z_{l,j}\right) = \tag{1.57}$$

$$\sum_{l=0}^{j}\left(\sum_{k=0}^{j} Y_{i,k} M_{k,l}\right) Z_{l,j} = \sum_{l=0}^{j}\left(\sum_{k=0}^{\infty} Y_{i,k} M_{k,l}\right) Z_{l,j} = \sum_{l=0}^{j} Z_{l,j}\delta_{i,l} = Z_{i,j} \tag{1.58}$$

so $Y = Z$ and we can indeed conclude that

$$\sum_{k=0}^{\infty} M_{i,k} Z_{k,j} = \delta_{i,j} = \sum_{k=0}^{\infty} Z_{i,k} M_{k,j}, \ \forall i, j \tag{1.59}$$

where the rows and columns of the matrices are numbered according to $t_0, t_1, \ldots$. Consider finally an arbitrary numbering $n_0, n_1, n_2 \ldots$ of the elements of T. Let f be a one to one mapping from the nonnegative integers onto the nonnegative integers that satisfies $t_{f(l)} = n_l$ for all integers $l \geq 0$; let i and j be nonnegative integers, and let $i' = f(i)$ and let $j' = f(j)$. Then

$$\sum_{k=0}^{\infty} M_{n_i,n_k} Z_{n_k,n_j} \tag{1.60}$$

is just a rearrangement of the sum

$$\sum_{k=0}^{\infty} M_{t_{i'},t_k} Z_{t_k,t_{j'}} = \sum_{k=0}^{j'} M_{t_{i'},t_k} Z_{t_k,t_{j'}} + \sum_{k=j'+1}^{\infty} 0 \tag{1.61}$$

which is obviously absolutely convergent. Thus all rearrangements are equal (see Chapter 3 of [Ru64]) and therefore for each pair of nonnegative integers i, j,

$$\sum_{k=0}^{\infty} M_{n_i,n_k} Z_{n_k,n_j} = \sum_{k=0}^{\infty} M_{t_{i'},t_k} Z_{t_k,t_{j'}} = \delta_{i',j'} = \delta_{i,j}. \tag{1.62}$$

The same reasoning shows that

$$\sum_{k=0}^{\infty} Z_{n_i,n_k} M_{n_k,n_j} = \delta_{i,j} \tag{1.63}$$

as well. □

Corollary 1.21 *The zeta matrix of the lattice L is invertible.* □

Definition 1.22 *The inverse matrix M of a zeta matrix Z of a partially ordered set T is known as the Möbius matrix of T.*

For more on lattices, see, for example, [Ro64].

Notation: From here on, we will begin to ignore the specific numbering of the lattice elements. Thus given a zeta vector or matrix we will refer to its coordinates by the names of their corresponding lattice elements rather than by their numbered positions. So if say, p and q are lattice elements, we will refer to "the p coordinate" of the vector, or "the q coordinate".

Observe that the matrix $z^r(z^r)^T$ satisfies

$$(z^r(z^r)^T)_{p,q} = \begin{cases} 1 : p \le r \text{ and } q \le r \\ 0 : \text{ otherwise} \end{cases} \tag{1.64}$$

but this means that

$$(z^r(z^r)^T)_{p,q} = \begin{cases} 1 : p \vee q \le r \\ 0 : \text{ otherwise} \end{cases} \tag{1.65}$$

We conclude that $(z^r(z^r)^T)_{p,q} = (z^r)_{p \vee q}$. Moreover, as this relationship is linear,

$$(\sum_{r \in L} \alpha_r z^r (z^r)^T)_{p,q} = (\sum_{r \in L} \alpha_r z^r)_{p \vee q}. \tag{1.66}$$

Observe also that since Z is nonsingular, every $x \in R^L$ (i.e. with a coordinate for each element of L) can be written as $x = \sum_{r \in L} \alpha_r z^r$.

Definition 1.23 *For every $x \in R^L$ define the $|L| \times |L|$ matrix W^x by*

$$W^x_{p,q} = x_{p \vee q}. \tag{1.67}$$

The following lemma is now clear.

Lemma 1.24 *For any $x = \sum_{r \in L} \alpha_r z^r$,*

$$W^x = \sum_{r \in L} \alpha_r z^r (z^r)^T. \ \Box \tag{1.68}$$

Where $p \in L$, denote by m^p the p'th row of the Möbius matrix, i.e. the row of M for which $(m^p)^T z^p = 1$. The matrices W^{z^r} have the following inverse-type relationship with the rows of the Möbius matrix.

Lemma 1.25 *Let a and b be vectors in R^L, and let p and r belong to L. Then*

$$a^T W^{z^r} b = \left((z^r)^T a\right)\left((z^r)^T b\right) \tag{1.69}$$

In particular,

$$(m^p)^T W^{z^r} m^p = \delta_{p,r}. \ \Box \tag{1.70}$$

In general,

Lemma 1.26 *Where $x = \sum_{r \in L} \alpha_r z^r$ and $p \in L$, we have*

$$(m^p)^T W^x m^p = \alpha_p. \ \Box \tag{1.71}$$

The previous two lemmas imply the following lemma.

Lemma 1.27 *The vector $x \in R^L$ belongs to the cone generated by the zeta vectors (i.e. the columns of Z) iff $W^x \succeq 0$*

Proof: Write $x = \sum_{r \in L} \alpha_r z^r$ (this expression exists and is unique by the nonsingularity of Z). If $W^x \succeq 0$ then by Lemma 1.26, $\alpha_r \geq 0, \ \forall r$ so x is in the cone of the z^r vectors. Conversely, if $\alpha \geq 0$ then for any $v \in R^L$,

$$v^T W^x v = \sum_{r \in L} \alpha_r \left(v^T z^r (z^r)^T v\right) = \sum_{r \in L} \alpha_r (v^T z^r)^2 \geq 0. \ \Box \tag{1.72}$$

Observe that by Definition 1.23, every entry of the matrix W^x is one of the coordinates of the vector x. This gives the following lemma.

Lemma 1.28 *For every pair of vectors a, b from R^L, there exists a unique vector, to be denoted $a \bigvee b$, such that*

$$a^T W^x b = \left(a \bigvee b\right)^T x \tag{1.73}$$

for all $x \in R^L$.

Proof: Each $(W^x)_{p,q} = x_{p \vee q}$ entry is multiplied in the expression $a^T W^x b$ by $a_p b_q$. Thus for any given r, x_r will be multiplied by

$$\sum_{p,q \in L : p \vee q = r} a_p b_q \tag{1.74}$$

Denote the vector with this expression as its r entry as $a \bigvee b$. It is clear that $a^T W^x b = (a \bigvee b)^T x$ for all $x \in R^L$. Uniqueness is also clear as $u^T z^r = (a \bigvee b)^T z^r$, $\forall r \in L \Rightarrow u = a \bigvee b$, as the vectors z^r constitute a basis for R^L. □

Let us now state this as a formal definition.

Definition 1.29 *For every pair of vectors a, b from R^L, define the vector $a \bigvee b \in R^L$ by*

$$\left(a \bigvee b\right)_r = \sum_{p,q \in L : p \vee q = r} a_p b_q. \tag{1.75}$$

Lemma 1.30 *The binary operator $\bigvee$ on $R^L \times R^L$ is commutative, associative, and distributive. Furthermore,*

$$e_p \bigvee e_q = e_{p \vee q} \tag{1.76}$$

where e_p is the unit vector corresponding to the lattice element p.

Proof: For each zeta vector z^r,

$$(e_p)^T W^{z^r} e_q = (e_p)^T z^r (z^r)^T e_q = z^r_p z^r_q = z^r_{p \vee q} \Rightarrow \tag{1.77}$$

$$\left(e_p \bigvee e_q\right)^T z^r = (e_{p \vee q})^T z^r \tag{1.78}$$

Since this is true for every z^r and the $\{z^r\}$ constitute a basis we conclude that $e_p \bigvee e_q = e_{p \vee q}$. The remainder of the lemma is clear by construction. □

At this point let us summarize what we know about the cone of zeta vectors.

Definition 1.31 *Define*

$$H = \{x \in R^L : x = Z\alpha,\ \alpha \in R^L,\ \alpha \geq 0\}. \tag{1.79}$$

Lemma 1.32 *The following are equivalent:*

1. $x \in H$
2. $Mx \geq 0$
3. $W^x \succeq 0$
4. $(a \bigvee a)^T x \geq 0, \ \forall a \in R^L$

Proof: The only part of the statement that has not yet been proven explicitly is the equivalence of $Mx \geq 0$ with the rest, so that the polar H^* is generated by the rows of M. But this follows trivially from the fact that Z and M are inverse to one another.

$$x = Z\alpha, \ \alpha \geq 0 \Rightarrow Mx = MZ\alpha = \alpha \geq 0 \tag{1.80}$$

and conversely,

$$Mx = \alpha \geq 0 \Rightarrow x = ZMx = Z(Mx) = Z\alpha, \ \alpha \geq 0. \ \Box \tag{1.81}$$

Thus the polar cone H^* of H can be generated either from the rows m^p of M, or from the vectors of the form $a \bigvee a$. It therefore follows that the rows m^p of M are generated by vectors of the form $a \bigvee a$ and conversely. In fact, $m^p = m^p \bigvee m^p$, and more generally,

Lemma 1.33

$$m^p \bigvee m^q = \delta_{p,q} m^p \tag{1.82}$$

Proof: For every z^r,

$$(m^p)^T W^{z^r} m^q = (m^p)^T z^r (z^r)^T m^q = \tag{1.83}$$

$$\delta_{p,r}\delta_{q,r} = \delta_{p,q}\delta_{p,r} = \delta_{p,q}(m^p)^T z^r \ \Box \tag{1.84}$$

Corollary 1.34 *The set of idempotents of the operator* $\bigvee$ *is exactly the set*

$$\{x \in R^L : x = \sum_{t \in T \subseteq L} m^t\} \tag{1.85}$$

for some subset T *of* L.

Proof: Since M is nonsingular, every $x \in R^L$ can be written as $x = \sum_{r \in L} \beta_r m^r$ so by Lemmas 1.30 and 1.33,

$$x \bigvee x = \sum_{r \in L} \beta_r^2 m^r = \sum_{r \in L} \beta_r m^r = x \text{ iff } \beta_r \in \{0,1\}, \ \forall r \in L. \ \Box \tag{1.86}$$

Lovász and Schrijver give one further characterization of the cone H in the form of a "remark". We will see a simple proof of their statement later on (Corollaries 2.24 and 3.20, and Lemma 3.29).

Theorem 1.35 *The vector $x \in R^L$ with $x_\emptyset = 1$ belongs to the cone H iff there exists a probability measure $\mathcal{X}$ on some measure space $(\Omega, \mathcal{W})$, and sets $A_1, \ldots, A_n$ in $\mathcal{W}$, such that for every $r \in L$,*

$$\mathcal{X}(\bigcap_{i: s_i \in r} A_i) = x_r. \ \Box \tag{1.87}$$

1.2.2 Exponential Lifts and the N Operator

We now return to the set $P \subseteq \{0,1\}^n$. Recall that any vector z^r is an expansion of a vector in $\{0,1\}^n$ to 2^n dimensions, and that each vector z^r is the expansion of exactly one such point in $\{0,1\}^n$. Thus the points of P are each projections of exactly one vector z^r to their singleton set coordinates (and the points of $K(P) \subseteq \{0,1\}^{n+1}$, as it was defined in Definition 1.2, are the projections of those same vectors to their empty set and singleton set coordinates).

Definition 1.36 *Let $P \subseteq \{0,1\}^n$ and let $K(P)$ be as in Definition 1.2. Define $K^e(P) \subseteq \{0,1\}^{|L|}$ to be the set of zeta vectors z^r whose projection to singleton sets belongs to P (equivalently, whose projection to the empty set and singleton sets belongs to $K(P)$). For the forthcoming discussion assume also that, where H is as in Definition 1.31, $\bar{K}^e(P) \subseteq H$ is any cone satisfying*

$$\bar{K}^e(P) \cap \{0,1\}^{|L|} = K^e(P) \cup \{0\}. \tag{1.88}$$

Where the dependence is clear or irrelevant we will drop the "P" from the notation and write simply K, K^e and $\bar{K}^e$.

The cone of K^e is thus trivial to characterize. It is the set of nonnegative combinations of those z^r vectors whose projection is in P. Formally,

Lemma 1.37 *Writing $x = \sum_{r \in L} \alpha_r z^r$, x belongs to $Cone(K^e)$ iff $\alpha \geq 0$, and $\alpha_r = 0$ wherever $z^r \notin K^e$. Equivalently,*

$$Cone(K^e) = \{x \in R^L : x \in H, \ (m^r)^T x = 0 \ \forall z^r \notin K^e\} \tag{1.89}$$

and the cone of $K \subseteq \{0,1\}^{n+1}$ is the projection of this cone to the empty set and singleton set coordinates. $\square$

So for $x = \sum_{r \in L} \alpha_r z^r$ we need to have $\alpha \geq 0$ and $\alpha_r = 0\ \forall z^r \notin K^e$, but this is equivalent to saying that for all $r \in L$, $\alpha_r z^r$ must be a nonnegative multiple of an element of K^e. Or equivalently,

Lemma 1.38

$$Cone(K^e) = \{x = \sum_{r \in L} \alpha_r z^r \in R^L : \alpha_r z^r \in Cone(K^e)\ \forall r \in L\} \tag{1.90}$$

Proof: Given $x = \sum_{r \in L} \alpha_r z^r$ it is trivial that if each $\alpha_r z^r \in Cone(K^e)$ then $x \in Cone(K^e)$. Conversely if some $\alpha_r z^r \notin Cone(K^e)$ then either $\alpha_r < 0$ or if $\alpha_r > 0$ then it must be that $z^r \notin Cone(K^e)$. In either case, by Lemma 1.37, $x \notin Cone(K^e)$ $\square$.

Claim 1.39 *Let $\bar{K}^e$ be as in Definition 1.36. For any $r \in L$, $\alpha_r \in R$,*

$$\alpha_r z^r \in Cone(K^e) \quad \textit{iff} \quad \alpha_r z^r \in \bar{K}^e. \tag{1.91}$$

Proof: The zeta vectors z^r are all nonzero, so by hypothesis $z^r \in K^e$ iff $z^r \in \bar{K}^e$. If $\alpha_r = 0$ then the claim is trivial. If $\alpha_r < 0$ then by hypothesis and Lemma 1.37, $\alpha_r z^r$ belongs to neither $Cone(K^e)$ nor $\bar{K}^e$, and if $\alpha_r > 0$ then

$$\alpha_r z^r \in Cone(K^e) \Leftrightarrow z^r \in Cone(K^e) \Leftrightarrow z^r \in \bar{K}^e \Leftrightarrow \alpha_r z^r \in \bar{K}^e. \tag{1.92}$$

Lemma 1.40

$$Cone(K^e) = \{x \in R^L : W^x m^p \in \bar{K}^e,\ \forall p \in L\}. \tag{1.93}$$

Stated another way, since the m^p generate the cone H^*,

$$Cone(K^e) = \{x \in R^L : W^x h^* \in \bar{K}^e,\ \forall h^* \in H^*\}. \tag{1.94}$$

Proof: Where $x = \sum_{r \in L} \alpha_r z^r$,

$$W^x m_p = \sum_{r \in L} \alpha_r z^r (z^r)^T m^p = \alpha_p z^p. \tag{1.95}$$

The lemma follows directly now from the claim and Lemma 1.37. $\square$

Definition 1.41 *Given $x \in R^L$ denote the subvector made up of only those coordinates corresponding to the empty set, the singleton subsets, and the pairs subsets as $\bar{x}$, and denote* *the subvector made up of only the empty set and singleton subset coordinates as $\bar{\bar{x}}$. Denote the square submatrix of W^x whose rows and columns correspond to the empty and singleton subsets as $W^{\bar{\bar{x}}}$. (Note that any entry in that submatrix is of the form $x_{p\vee q}$ where p and q are both either singletons or empty. Thus $p \vee q$ is an empty, singleton, or pairs subset, and is therefore a coordinate of $\bar{x}$.)*

Lemma 1.42 *Let Y be a $n+1 \times n+1$ symmetric matrix with $Diag(Y) = Ye_0$. Then*

$$Y = W^{\bar{\bar{x}}} \tag{1.96}$$

for some $x \in R^L$.

Proof: Write $x_\emptyset = Y_{0,0}$. For each singleton subset $\{s_i\}$, write $x_{\{s_i\}} = Y_{i,i}$ and for each pair $\{s_i, s_j\}$, write $x_{\{s_i,s_j\}} = Y_{i,j}$ (this is well-defined since $Y_{i,j} = Y_{j,i}$). We claim that we now have $Y = W^{\bar{\bar{x}}}$ where the zero'th row and column correspond to the empty set, and the i'th row and column correspond to the singleton set $\{s_i\}$. For this to hold we must have $Y_{i,j} = x_{\{s_i,s_j\}}$ where i and j are both not zero, and we must have

$$Y_{i,0} = x_{\{s_i\}\vee\emptyset} = x_{\{s_i\}} \tag{1.97}$$

where $i \neq 0$, and

$$Y_{0,j} = x_{\{s_j\}\vee\emptyset} = x_{\{s_j\}} \tag{1.98}$$

where $j \neq 0$, and $Y_{0,0} = x_{\emptyset\vee\emptyset} = x_\emptyset$. Our construction guarantees that all of these criteria are met. □

We are now interested in relaxing the definition of $Cone(K^e)$ in Lemma 1.40 to obtain an approximation of its $n+1$ dimensional projection, $Cone(K)$. Certainly we can replace the cone $\bar{K}^e$ with the cone, to be denoted $\bar{K}$, that is its projection on its empty set and singleton set coordinates, as this cone has the same relationship with $Cone(K)$ as $\bar{K}^e$ has with $Cone(K^e)$.

Definition 1.43 *Denote the collection of empty set and singleton set coordinates by I.*

Observe that the projection of H on I is $Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1\})$.

Here, and for the remainder of this chapter, let $\bar{K} \subseteq Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1\})$ be a cone that satisfies

$$\bar{K} \cap \{0,1\}^{n+1} = K \cup \{0\}. \tag{1.99}$$

Lemma 1.44 *For any projected zeta vector $\bar{\bar{z}}^r \in \{0,1\}^{n+1}$ and any number α_r,*

$$\alpha_r \bar{\bar{z}}^r \in Cone(K) \;\; \textit{iff} \;\; \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{1.100}$$

Proof: The points $\bar{\bar{z}}^r$ are all nonzero, so by hypothesis $\bar{\bar{z}}^r \in K$ iff $\bar{\bar{z}}^r \in \bar{K}$. Clearly if $\alpha = 0$ then $\alpha_r \bar{\bar{z}}^r$ is in both $Cone(K)$ and $\bar{K}$, and if $\alpha_r < 0$ then it is in neither (as both cones are contained in R_+^{n+1}), and if $\alpha_r > 0$ then

$$\alpha_r \bar{\bar{z}}^r \in Cone(K) \Leftrightarrow \bar{\bar{z}}^r \in Cone(K) \Leftrightarrow \bar{\bar{z}}^r \in \bar{K} \Leftrightarrow \alpha_r \bar{\bar{z}}^r \in \bar{K}. \;\Box \tag{1.101}$$

We will also relax W^x to $W^{\bar{x}}$, but it is less obvious what should be the relaxation of the term H^* in Lemma 1.40. A simple suggestion would be to try the polar of the projection of H on the empty set and singleton set coordinates.

Definition 1.45 *Given a set $T \subseteq R^L$, let its projection on its I coordinates be denoted $T|_I$. Consider now the intersection of T with the subspace Sp^I generated by the vectors of R^L that have zeroes in all but their I coordinates. The projection of this intersection to the I coordinates will be denoted T_I (so these are the projections to I coordinates of the points of T that are zero outside of their I coordinates).*

Lemma 1.46 *The projection $H|_I$ is just the set*

$$Cone(\{\bar{\bar{z}}^r : r \in L\}) = \tag{1.102}$$

$$Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1) \tag{1.103}$$

with polyhedral representation

$$\{x \in \;\; R^{n+1} : x \geq 0, x_i \leq x_0, \;\; \forall i = 1, \ldots, n\}. \tag{1.104}$$

The polar cone is therefore generated by the vectors

$$\{e_i, e_0 - e_i : i = 1, \ldots, n\}. \tag{1.105}$$

Proof: The first statement is obvious from the discussion at the beginning of Section 1.2.1. As for the second statement, it is clear that any vector $x \in H|_I$ satisfies $x \geq 0, x_i \leq x_0$ as this is true for every $y \in \{0,1\}^{n+1} : y_0 = 1$. Conversely, any nonnegative x satisfying $x_i \leq x_0, \;\; i = 1, \ldots, n$, can be decomposed into a nonnegative combination of vectors from

$\{y \in \{0,1\}^{n+1} : y_0 = 1\}$ as follows. Arrange the $1, \ldots, n$ coordinates of x such that $x_1 \geq x_2 \geq \cdots$, and say that the last nonzero coordinate of x is the k'th. Then

$$x = \sum_{i=0}^{k} \lambda_i y^i \tag{1.106}$$

where $y^i \in \{0,1\}^{n+1}$ has a 1 in coordinates $0, \cdots, i$ and zeroes elsewhere, and $\lambda_i = x_i - x_{i+1}$ (if $i = n$ then define $x_{n+1} = 0$). $\square$

The relaxation obtained where we replace $\bar{K}^e$ by $\bar{K}$, W^x by $Y = W^{\bar{x}}$, and H^* by $(H|_I)^*$ is thus

$$\{\bar{x} : W^{\bar{x}} h^* \in \bar{K},\ \forall h^* \in (H|_I)^*\} = \tag{1.107}$$

$$\{\bar{x} : W^{\bar{x}} e_i \in \bar{K},\ W^{\bar{x}}(e_0 - e_i) \in \bar{K},\ i = 1, \ldots, n\}. \tag{1.108}$$

Lemma 1.47 *Where $M(\bar{K})$ is as in Definition 1.15, then*

$$M(\bar{K}) = \{W^{\bar{x}} : W^{\bar{x}} e_i \in \bar{K},\ W^{\bar{x}}(e_0 - e_i) \in \bar{K},\ i = 1, \ldots, n\}. \tag{1.109}$$

Proof: By Lemma 1.42, this is just

$$\{Y \in R^{(n+1)\times(n+1)} :$$

$$Y = Y^T,\ Diag(Y) = Ye_0,\ Ye_i \in \bar{K},\ Y(e_0 - e_i) \in \bar{K}, i = 1, \ldots, n\}. \tag{1.110}$$

But this is exactly $M(\bar{K})$. Note that the projection to I coordinates, which is $N(\bar{K})$ (and is obtained by taking the $\emptyset$ column of Y) can also be written

$$\{\bar{\bar{x}} : W^{\bar{x}} e_i \in \bar{K},\ W^{\bar{x}}(e_0 - e_i) \in \bar{K}\}. \square \tag{1.111}$$

1.2.3 The $N(\bar{K}, \bar{K'})$ and Lasserre Operators

Lovász and Schrijver also describe the following variation of the N procedure. Consider first the following variation of Lemma 1.40.

Lemma 1.48 *If $\bar{K}^e \subseteq H$ is any cone satisfying*

$$\bar{K}^e \cap \{0,1\}^{|L|} = K^e \cup \{0\} \tag{1.112}$$

and $\bar{K'}^e \subseteq H$ is any cone satisfying

$$\bar{K'}^e \cap \{0,1\}^{|L|} \supseteq K^e \cup \{0\} \tag{1.113}$$

Then

$$Cone(K^e) = \{x \in R^L : W^x(k')^* \in \bar{K}^e,\ \forall (k')^* \in (\bar{K'}^e)^*\}. \tag{1.114}$$

Proof:

$$\bar{K'}^e \subseteq H \Rightarrow H^* \subseteq (\bar{K'}^e)^* \Rightarrow \tag{1.115}$$

$$\{x \in R^L : W^x (k')^* \in \bar{K}^e,\ \forall (k')^* \in (\bar{K'}^e)^*\} \subseteq \tag{1.116}$$

$$\{x \in R^L : W^x h^* \in \bar{K}^e,\ \forall h^* \in H^*\} = Cone(K^e) \tag{1.117}$$

Conversely, for any $z^r \in K^e$, we have $z^r \in \bar{K'}^e$ so that for any $(k')^* \in (\bar{K'}^e)^*$,

$$W^{z^r}(k')^* = z^r (z^r)^T (k')^* \tag{1.118}$$

is a nonegative multiple of z^r, and therefore must belong to $\bar{K}^e$. So

$$Cone(K^e) \subseteq \{x \in R^L : W^x (k')^* \in \bar{K}^e,\ \forall (k')^* \in (\bar{K'}^e)^*\} \tag{1.119}$$

and the lemma follows. □

If we relax $(\bar{K'}^e)^*$ to the polar of its projection $\bar{K'}$ on the I coordinates in the same manner as we did for H^* then we obtain a stronger operator than N, defined by

$$\{\bar{x} : W^{\bar{x}}(k')^* \in \bar{K},\ \forall (k')^* \in \bar{K'}^*\}. \tag{1.120}$$

The projection of this set on the I coordinates is

$$\{x \in R^{n+1} : \exists Y \in R^{(n+1)\times(n+1)} \text{ such that } Y(k')^* \in \bar{K},\ \forall (k')^* \in \bar{K'}^*\} \tag{1.121}$$

where Y is symmetric with $Diag(Y) = Ye_0$. This is the set that Lovász and Schrijver refer to as $N(\bar{K}, \bar{K'})$.

Lemma 1.49

$$Cone(K) \subseteq N(\bar{K}, \bar{K'}) \subseteq N(\bar{K}) \tag{1.122}$$

Proof: All references to matrices Y should be assumed to refer to $n+1 \times n+1$ symmetric matrices with $Diag(Y) = Ye_0$.

$$N(\bar{K}, \bar{K'}) = \tag{1.123}$$

$$\{x \in R^{n+1} : \exists Y \in R^{(n+1)\times(n+1)} \text{ such that } Y(k')^* \in \bar{K},\ \forall (k')^* \in \bar{K'}^*\} \subseteq \tag{1.124}$$

$$\{x \in R^{n+1} : \exists Y \in R^{(n+1)\times(n+1)} \text{ such that } Yh^* \in \bar{K},\ \forall h^* \in (H|_I)^*\} = \tag{1.125}$$

$$N(\bar{K}) \tag{1.126}$$

because $\bar{K}'^* \supseteq (H|_I)^*$. And since, by construction, every $\bar{\bar{z}}^r \in K$ also belongs to $\bar{K}'$ (as well as $\bar{K}$),

$$W^{\bar{z}^r}(k')^* = \bar{\bar{z}}^r(\bar{\bar{z}}^r)^T(k')^* \quad (1.127)$$

(where $\bar{z}^r$ is the projection of z^r to empty set, singleton and pairs coordinates) is a non-negative multiple of $\bar{\bar{z}}^r$ (for every $(k')^* \in (\bar{K}')^*$), and therefore belongs to $\bar{K}$. So $N(\bar{K}, \bar{K}')$ does not cut off any points from $Cone(K)$. □

Notice that $W^{\bar{x}}$ is a principal minor of W^x and therefore a necessary condition for $\bar{x}$ to be a projection of an x that is even in H (let alone in K) is for $W^{\bar{x}}$ to be positive semidefinite. But after relaxing the procedure we have no guarantee of this. So the procedure will be further strengthened by insisting on positive semidefiniteness, and this gives us the $N^+(\bar{K})$ and $N^+(\bar{K}, \bar{K}')$ procedures.

Letting $\bar{K} = Cone(K)$, we can see that the main idea behind the $N(\bar{K}, \bar{K}')$ procedure is that in addition to the valid conditions $W^{\bar{x}}e_i \in Cone(K)$ and $W^{\bar{x}}(e_0 - e_i) \in Cone(K)$, for all $(k')^* \in \bar{K}'^*$ we must also have $W^{\bar{x}}(k')^* \in Cone(K)$. Thus any necessary condition for membership in $Cone(K)$ can be placed on $W^{\bar{x}}(k')^*$. In particular there must be a way to append coordinates (corresponding to doubles) to the vector $W^{\bar{x}}(k')^*$ such that the $n+1 \times n+1$ matrix that *it* defines will be positive semidefinite. This is the essential idea behind the Lasserre operator (as applied to $0,1$ integer programming). There are a number of additional details, however, to which we will return later.

1.2.4 The $\bar{N}$ Operator

To understand this operator we need to know what the Möbius matrix M looks like. The p'th row of M, where $p \in L$, has zeroes in all entries r such that $|r| < |p|$. For $|r| = |p|$ the only nonzero entry is at $r = p$, and that entry is 1. For $|r| = |p| + 1$, the nonzeroes are only at those r for which $r \supset p$, and they are all -1. For $|r| = |p| + 2$, again the nonzeroes are only at those r for which $r \supset p$, but here they are all 1. This pattern then continues, alternating between 1's and -1's (i.e. $(-1)^{|r|-|p|}$), with $r \not\supseteq p$ always yielding value 0. For example, where $|S| = 3$ and the set $\{s_i, s_j\}$ is represented as i, j, the matrix M is as follows.

	$\emptyset$	1	2	3	1,2	1,3	2,3	1,2,3
$\emptyset$	1	−1	−1	−1	1	1	1	−1
1	0	1	0	0	−1	−1	0	1
2	0	0	1	0	−1	0	−1	1
3	0	0	0	1	0	−1	−1	1
1,2	0	0	0	0	1	0	0	−1
1,3	0	0	0	0	0	1	0	−1
2,3	0	0	0	0	0	0	1	−1
1,2,3	0	0	0	0	0	0	0	1

Definition 1.50 *Define the vector* $m^{[p,q]} \in R^L$ *by*

$$m^{[p,q]}_r = \begin{cases} m^p_r : r \subseteq q \\ 0 : \ \textit{otherwise} \end{cases} \tag{1.128}$$

So $m^{[p,q]}$ is the same as m^p, but with all entries $r : r \not\subseteq q$ zeroed out. For example,

$$m^{[\emptyset,\{s_1,s_2\}]} = (1,-1,-1,0,1,0,0,0) \tag{1.129}$$

in the above case.

Observe that for any $r \in L$,

$$(m^{[p,q]})^T z^r = \sum_{t \in L : t \subseteq r} m^{[p,q]}_t = \sum_{t \in L : t \subseteq r,\ t \subseteq q} m^p_t = (m^p)^T z^{r \wedge q} = \delta_{p, r\wedge q} \geq 0 \tag{1.130}$$

so $m^{[p,q]} \in H^*$. Moreover, they are idempotent with respect to the operator $\bigvee$ since

$$(m^{[p,q]})^T W^{z^r} m^{[p,q]} = ((m^{[p,q]})^T z^r)^2 = (m^{[p,q]})^T z^r \Rightarrow \tag{1.131}$$

$$m^{[p,q]} \bigvee m^{[p,q]} = m^{[p,q]}. \tag{1.132}$$

Recall that the vectors $\{e_i : i = 1, \ldots, n\}$ and $\{e_0 - e_i : i = 1 \ldots, n\}$ in R^{n+1} generated the cone $(H|_I)^*$ (where the zero coordinate corresponds to the emptyset and the i'th coordinate corresponds to the i'th singleton set, $i = 1, \ldots, n$). Expressed in terms of the lattice L, these are the vectors $\{e_{\{s_i\}}\}$ and $\{e_\emptyset - e_{\{s_i\}}\}$. Notice that these are exactly the (nonzero) vectors $m^{[p,q]}$ where $|q| \leq 1$. In particular,

$$e_{\{s_i\}} = m^{[\{s_i\},\{s_i\}]} \tag{1.133}$$

and

$$\{e_\emptyset - e_{\{s_i\}}\} = m^{[\emptyset,\{s_i\}]}. \tag{1.134}$$

Notice now that by Lemma 1.30, given some $m^{[p,q]}$, $|q| = k$,

$$m^{[p,q]} \bigvee e_{\{s_i\}} \tag{1.135}$$

shifts every nonzero entry in $m^{[p,q]}$ from position r to position $r \vee \{s_i\}$, and

$$m^{[p,q]} \bigvee (e_\emptyset - e_{\{s_i\}}) = m^{[p,q]} - m^{[p,q]} \bigvee e_{\{s_i\}}. \tag{1.136}$$

Looking at the matrix above, it is not difficult to see that these expressions are themselves of the form $m^{[p',q']}$ for some q' with $|q'| \leq k+1$, but here is a formal proof.

Lemma 1.51 *Let $p, q \in L$. Consider the following four cases:*

Case 1: *If $p \subseteq q$, and $s_i \notin q$ then*

$$m^{[p,q]} \bigvee e_{\{s_i\}} = m^{[p\vee\{s_i\},q\vee\{s_i\}]} \tag{1.137}$$

$$m^{[p,q]} \bigvee (e_\emptyset - e_{\{s_i\}}) = m^{[p,q]} - m^{[p,q]} \bigvee e_{\{s_i\}} = m^{[p,q\vee\{s_i\}]}. \tag{1.138}$$

Case 2: *If $p \subseteq q$, and $s_i \in p$ then*

$$m^{[p,q]} \bigvee e_{\{s_i\}} = m^{[p,q]} \tag{1.139}$$

$$m^{[p,q]} \bigvee (e_\emptyset - e_{\{s_i\}}) = 0. \tag{1.140}$$

Case 3: *If $p \subseteq q$, and $s_i \in q - p$ then*

$$m^{[p,q]} \bigvee e_{\{s_i\}} = 0 \tag{1.141}$$

$$m^{[p,q]} \bigvee (e_\emptyset - e_{\{s_i\}}) = m^{[p,q]}. \tag{1.142}$$

Case 4: *If $p \not\subseteq q$, then*

$$m^{[p,q]} \bigvee e_{\{s_i\}} = m^{[p,q]} \bigvee (e_\emptyset - e_{\{s_i\}}) = 0. \tag{1.143}$$

Proof:

Case 1: As we noted, $m^{[p,q]} \bigvee e_{\{s_i\}}$ shifts each r'th entry to the $r \vee \{s_i\}$ position. The nonzero entries can therefore be only in positions r such that $p \vee \{s_i\} \subseteq r \subseteq q \vee \{s_i\}$, and since $s_i \notin q$, the value mapped into any such r'th position is exactly that which was in the $r - \{s_i\}$'th position of $m^{[p,q]}$. Thus the first (according to the lattice partial order) nonzero will be of value 1 and in position $p \vee \{s_i\}$. The next nonzeroes will all be of value -1 and in positions $r : p \vee \{s_i\} \subseteq r \subseteq q \vee \{s_i\}$ with $|r| = |p| + 2$. Subsequent nonzeroes will be in $r : p \vee \{s_i\} \subseteq r \subseteq q \vee \{s_i\}$ with signs alternating for each unit increase in the cardinality of r. This vector is exactly $m^{[p \vee \{s_i\}, q \vee \{s_i\}]}$.

Consider now

$$m^{[p,q]} \bigvee (e_\emptyset - e_{\{s_i\}}) = m^{[p,q]} - m^{[p,q]} \bigvee e_{\{s_i\}} = m^{[p,q]} - m^{[p \vee \{s_i\}, q \vee \{s_i\}]} \tag{1.144}$$

and observe that since $s_i \notin q$, the nonzero entries of $m^{[p \vee \{s_i\}, q \vee \{s_i\}]}$ are all of value zero in $m^{[p,q]}$, and conversely the nonzero entries of $m^{[p,q]}$ are all of value zero in $m^{[p \vee \{s_i\}, q \vee \{s_i\}]}$. Thus the nonzero entries of (1.144) are of exactly one of the following two nonoverlapping types:

1. A 1 in position p, with subsequent nonzeroes in positions $r : p \subseteq r \subseteq q$ with signs alternating for every unit increase in $|r|$.
2. A -1 in position $p \vee \{s_i\}$, with subsequent nonzeroes in positions $r : p \vee \{s_i\} \subseteq r \subseteq q \vee \{s_i\}$ with signs alternating for every unit increase in $|r|$.

This defines the vector $m^{[p, q \vee \{s_i\}]}$.

Case 2: Equation 1.139 follows from the fact that no nonzeroes are shifted in this case, and (1.140) follows directly from (1.139).

Case 3: In this case the only change effected by $m^{[p,q]} \vee e_{\{s_i\}}$ is to shift the value of each r'th position for which $p \subseteq r \subseteq q,\ s_i \notin r$ to the $r \vee \{s_i\}$'th position. But for any such r, $r \vee \{s_i\} \subseteq q$, and *its* value in $m^{[p,q]}$ is also nonzero. In particular its value is of opposite sign to the value in the r'th position, and thus shifting will result in a value of zero. Conversely, any r'th position for which $p \vee \{s_i\} \subseteq r \subseteq q$, will have its value cancelled by the shifting of the value of the $r - \{s_i\}$'th position into the r'th position. This establishes (1.141), and equation (1.142) follows directly from (1.141).

Case 4: Trivial. □

The following lemma is a consequence of Lemma 1.51.

Lemma 1.52 *Every vector $m^{[p,q]}$, $|q| \leq k \geq 2$ satisfies*

$$m^{[p,q]} = m^{[\hat{p},\hat{q}]} \bigvee m^{[s,t]} \tag{1.145}$$

for some $\hat{p}, \hat{q}, s, t$ *satisfying* $|\hat{q}| \leq k-1$, *and* $|t| \leq 1$. *Conversely, for every* $\hat{p}, \hat{q}, s, t \in L$ *with* $|\hat{q}| \leq k$ *and* $|t| \leq 1$,

$$m^{[\hat{p},\hat{q}]} \bigvee m^{[s,t]} = m^{[p,q]} \tag{1.146}$$

for some $p, q \in L$ *with* $|q| \leq k+1$. □

Before proceeding to a corollary, we need a definition.

Definition 1.53 *Where* J *and* J' *are subsets of* L, x *is a* $|J|$ *dimensional vector in* $R^J := R^L|_J$, *and* y *is a* $|J'|$ *dimensional vector in* $R^{J'}$*,the vector*

$$x \bigvee y \tag{1.147}$$

is defined to be the vector $(x,0) \bigvee (y,0) \in R^L$ *where* $(x,0)$ *and* $(y,0)$ *are the* $|L|$ *dimensional vectors obtained by appending coordinates – all of value zero – to* x *for each* $r \in L$ *such that* $r \notin J$, *and to* y *for each* $r \in L$ *such that* $r \notin J'$. *Therefore for any* $u \in R^L$,

$$\left(x \bigvee y\right)^T u = (x,0)^T W^u (y,0) \tag{1.148}$$

Observe that $x \bigvee y$ *can only have a nonzero coordinate* r *where* $r = j \vee j'$ *for some* $j \in J$ *and* $j' \in J'$.

Corollary 1.54 *The set* H^l, $(l \leq n)$, *defined to be the* l*-fold product*

$$(H|_I)^* \bigvee (H|_I)^* \bigvee \cdots \bigvee (H|_I)^* = \tag{1.149}$$

$$\left\{v \in R^L : v = y^1 \bigvee y^2 \bigvee \cdots \bigvee y^l\right\} \tag{1.150}$$

where each $y^i \in (H|_I)^*$, *is the cone generated by the vectors*

$$\{m^{[p,q]} : |q| \leq l\} \tag{1.151}$$

Proof: The vectors $v \in H^l$ are l-fold $\bigvee$-products of nonnegative linear combinations of the vectors $m^{[s,t]}, |t| = 1$ (Lemma 1.46), and therefore by Lemma 1.52, they all belong to $Cone(\{m^{[p,q]} : |q| \leq l\})$. Conversely, also by Lemma 1.52, any vector $m^{[p,q]}$, $|q| \leq l$, can be decomposed as an l-fold $\bigvee$-product of vectors $m^{[s,t]}, |t| \leq 1$, all of which belong to $(H|_I)^*$ (note that $m^{[\emptyset,\emptyset]} = e_\emptyset \in (H|_I)^*$, and recall that $e_\emptyset$ is the identity for $\bigvee$). □

We can now rewrite Lemma 1.47 in terms of $\bigvee$ notation as follows.

Lemma 1.55

$$N(\bar{K}) = \left(\bar{K}^* \bigvee (H|_I)^*\right)^* |_I \tag{1.152}$$

Proof:

$$N(\bar{K}) = \{\bar{\bar{x}} : W^{\bar{x}}h^* \in \bar{K},\ \forall h^* \in (H|_I)^*\} = \tag{1.153}$$

$$\{\bar{\bar{x}} : k^* W^{\bar{x}} h^* \geq 0,\ \forall k^* \in \bar{K}^*,\ \forall h^* \in (H|_I)^*\} \tag{1.154}$$

Now observe that for all $k^* \in \bar{K}^*$ and $h^* \in (H|_I)^*$,

$$k^* W^{\bar{x}} h^* = \left(k^* \bigvee h^*\right)^T x \tag{1.155}$$

for every expansion x of $\bar{x}$ to R^L since even when we expand k^* and h^* to R^L they remain zero at all but their I coordinates, so the remaining entries of W^x and the remaining entries of x are irrelevant. Therefore

$$N(\bar{K}) = \left\{\bar{\bar{x}} : \left(k^* \bigvee h^*\right)^T x \geq 0,\ \forall k^* \in \bar{K}^*,\ h^* \in (H|_I)^*\right\} = \tag{1.156}$$

$$\left\{\bar{\bar{x}} : x \in \left(\bar{K}^* \bigvee (H|_I)^*\right)^*\right\} = \tag{1.157}$$

$$\left(\bar{K}^* \bigvee (H|_I)^*\right)^* |_I \ \square \tag{1.158}$$

Before we can describe nested N operators we will need one more lemma. The lemma states that if we take the projection of a cone $V \subseteq R^L$ on its I coordinates, then the polar of that projection is the intersection of the polar of the original set V with the subspace Sp^I (defined in Definition 1.45), projected on its I coordinates (which are the only coordinates of that intersection that can be nonzero).

Lemma 1.56 *For any cone $V \subseteq R^L$, $(V|_I)^* = (V^*)_I$.*

Proof: It suffices to show that a vector $\bar{y}$ is in the polar of the projection of V on its I coordinates iff the extension y of $\bar{y}$ obtained by appending all of the non-I coordinates to $\bar{y}$ all with value zero (so that $y \in Sp^I$) is in the polar of V. If $\bar{y}$ is in the polar of the projection, then for any $x \in V$, where $\bar{x}$ is its projection to I,

$$\bar{y}^T \bar{x} \geq 0 \Rightarrow y^T x = (\bar{y}, 0)^T x = \bar{y}^T \bar{x} \geq 0. \tag{1.159}$$

Conversely, if there is a vector $x \in V$ whose projection $\bar{x}$ is such that $\bar{y}^T \bar{x} < 0$ then

$$y^T x = (\bar{y}, 0)^T x = \bar{y}^T \bar{x} < 0. \ \square \tag{1.160}$$

To reduce clutter, let us use the notation

$$H^1 = (H|_I)^* \tag{1.161}$$

as per the definition above in Corollary 1.54. The repeated N operator thus satisfies

$$N^2(\bar{K}) = N\left(N(\bar{K})\right) = \left((N(\bar{K}))^* \bigvee H^1\right)^* |_I. \tag{1.162}$$

By the lemma,

$$(N(\bar{K}))^* = \left(\left(\bar{K}^* \bigvee H^1\right)^* |_I\right)^* = \tag{1.163}$$

$$\left(\left(\bar{K}^* \bigvee H^1\right)^{**}\right)_I = \left(\bar{K}^* \bigvee H^1\right)_I \tag{1.164}$$

and so

$$N\left(N(\bar{K})\right) = \left(\left(\bar{K}^* \bigvee H^1\right)_I \bigvee H^1\right)^* |_I \supseteq \tag{1.165}$$

$$\left(\bar{K}^* \bigvee H^1 \bigvee H^1\right)^* |_I = \tag{1.166}$$

$$\left(\bar{K}^* \bigvee H^2\right)^* |_I \tag{1.167}$$

where the containment follows from the fact that a polar of a smaller set is a larger set.

Lemma 1.57 *For each $l \geq 1$,*

$$N^l(\bar{K}) \supseteq \left(\bar{K}^* \bigvee H^l\right)^* |_I \tag{1.168}$$

Proof: It is clear from the definition of N that for any two cones K_1 and K_2 such that $K_1 \subseteq K_2$ we have $N(K_1) \subseteq N(K_2)$, and therefore by induction, for any j, $N^j(K_1) \subseteq N^j(K_2)$. We have shown that the lemma holds for $l \leq 2$. Assume now that it holds for some $l \geq 2$. Notice that the reasoning employed in deriving the result for $l = 2$ did not depend on the superscript 1 of H. Therefore by induction and by the same reasoning as above,

$$N^{l+1}(\bar{K}) = N\left(N^l(\bar{K})\right) \supseteq N\left(\left(\bar{K}^* \bigvee H^l\right)^* |_I\right) \supseteq \left(\bar{K}^* \bigvee H^l \bigvee H^1\right)^* |_I = \tag{1.169}$$

$$\left(\bar{K}^* \bigvee H^{l+1}\right)^* |_I \ \square \tag{1.170}$$

Definition 1.58 *Where l is a positive integer $\leq n$, $\bar{N}^l(\bar{K})$ is defined by*

$$\bar{N}^l(\bar{K}) = \left(\bar{K}^* \bigvee H^l\right)^* |_I. \tag{1.171}$$

Lemma 1.59

$$Cone(K) \subseteq \bar{N}^l(\bar{K}) \subseteq N^l(\bar{K}) \tag{1.172}$$

Proof: The second inclusion has already been shown. As for the first, for any $\bar{\bar{z}}^r \in K$, consider the lifting $z^r \in R^L$, then for any $k \in \bar{K}^*$ and $h \in H^l$,

$$\left(k \bigvee h\right)^T z^r = (k,0)^T W^{z^r} h = (k,0)^T z^r (z^r)^T h = \lambda (k,0)^T z^r, \ \lambda = (z^r)^T h \geq 0 \tag{1.173}$$

since H^l is generated by vectors in H, and

$$\lambda(k,0)^T z^r = \lambda k^T \bar{\bar{z}}^r \geq 0 \tag{1.174}$$

since $\bar{\bar{z}}^r \in K \subseteq \bar{K}$. □

We will now characterize explicitly the vectors belonging to $\bar{N}^l(\bar{K})$ in similar terms to those used in the previous sections.

The cone H^l is generated by the vectors $m^{[p,q]}, \ |q| \leq l$, all of which are zero in all coordinates $r : |r| > l$, so any vector in this cone is also zero in all of those coordinates. Note further that by the definition of $\bigvee$ on vectors in R^J (Definition 1.53), for any $k \in \bar{K}^*, \ h \in H^l$,

$$\left(k \bigvee h\right)^T x = (k,0)^T W^x h \tag{1.175}$$

where $(k,0)$ is zero in all non-I coordinates. So the only relevant part of W^x is the rows corresponding to I, and the columns corresponding to $r : |r| \leq l$. The relevant coordinates of x are therefore those of the form $r \vee t$ where $|r| \leq l$ and $|t| \leq 1$, i.e. the relevant coordinates are those r with $|r| \leq l+1$. Let us denote by $\hat{x}$ the projection of x on these coordinates, by $W^{\hat{x}}$ the relevant rectangular submatrix of W^x, and by $\tilde{h}$ the projection of h on the coordinates $r : |r| \leq l$. Therefore, for any $x \in R^L$,

$$\left(k \bigvee h\right)^T x = k^T W^{\hat{x}} \tilde{h} \tag{1.176}$$

So the points x in the polar of $\bar{K}^* \bigvee H^l$ are those for which $W^{\hat{x}}\tilde{h} \in \bar{K}$ for each $h \in H^l$. Since H^l is generated by the vectors $m^{[p,q]}, \ p \subseteq q, \ |q| \leq l$ we have the following characterization.

Lemma 1.60 *Let $l \geq 1$ be a fixed integer, and let the vectors $\tilde{m}^{[p,q]}$ be the projections of the vectors $m^{[p,q]}$ on the coordinates $r : |r| \leq l$. Then*

$$\bar{N}^l(\bar{K}) = \{\bar{\bar{x}} : W^{\hat{x}} \tilde{m}^{[p,q]} \in \bar{K}, \ \forall p,q \in L \text{ such that } p \subseteq q, \ |q| \leq l\} \tag{1.177}$$

and if a polynomial time separation oracle exists for $\bar{K}$ then one exists for $\bar{N}^l(\bar{K})$ as well.

Proof: For any fixed l, $W^{\hat{x}}$ is of polynomial size, and there are only polynomially many pairs $p,q \in L$ with $p \subseteq q, \ |q| \leq l$. □

Example: Where $n = 4$ and $l = 2$ and where we represent the variables $x_{\{s_1,s_2,s_3\}}$ as $x_{1,2,3}$, and $x_\emptyset$ as x_0, then the matrix $W^{\hat{x}}$ is as follows.

	$\emptyset$	1	2	3	4	1,2	1,3	1,4	2,3	2,4	3,4
$\emptyset$	x_0	x_1	x_2	x_3	x_4	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{2,3}$	$x_{2,4}$	$x_{3,4}$
1	x_1	x_1	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,2,3}$	$x_{1,2,4}$	$x_{1,3,4}$
2	x_2	$x_{1,2}$	x_2	$x_{2,3}$	$x_{2,4}$	$x_{1,2}$	$x_{1,2,3}$	$x_{1,2,4}$	$x_{2,3}$	$x_{2,4}$	$x_{2,3,4}$
3	x_3	$x_{1,3}$	$x_{2,3}$	x_3	$x_{3,4}$	$x_{1,2,3}$	$x_{1,3}$	$x_{1,3,4}$	$x_{2,3}$	$x_{2,3,4}$	$x_{3,4}$
4	x_4	$x_{1,4}$	$x_{2,4}$	$x_{3,4}$	x_4	$x_{1,2,4}$	$x_{1,3,4}$	$x_{1,4}$	$x_{2,3,4}$	$x_{2,4}$	$x_{3,4}$

and the vectors $\tilde{m}^{[p,q]},\ p \subseteq q,\ |q| \leq 2$ are those of the form

$$e_0,\ e_i,\ e_{i,j},\ e_0 - e_i,\ e_i - e_{i,j},\ e_0 - e_i - e_j + e_{i,j}. \ \Box \tag{1.178}$$

We have now completed the survey of the convexification and Lovász Schrijver methodologies. At this point we will attempt to understand why they work, so that we can see if they can be further generalized.

Chapter 2

Analysis of the Operators

In this chapter we will redevelop the operators of the previous chapter from a different perspective, and we will see that the new approach suggests a much more general framework. Note first that given a sum $\sum_{\lambda\in\Lambda} a_\lambda$, a *partial sum* is defined to be a sum of the form

$$\sum_{\lambda\in\Gamma\subseteq\Lambda} a_\lambda. \tag{2.1}$$

Note now that for any vector $\bar{\bar{x}} \in [0,1]^n$ there exist scalars λ_y for each $y \in \{0,1\}^n$, such that

$$\bar{\bar{x}} = \sum_{y\in\{0,1\}^n} \lambda^y y, \quad \sum_{y\in\{0,1\}^n} \lambda^y = 1. \tag{2.2}$$

We will see that for any $\bar{\bar{x}} \in [0,1]^n$, and any lifting of $\bar{\bar{x}}$ to $x \in R^L$, where L is the lattice defined in Lemma 1.17, there is a representation (2.2) of $\bar{\bar{x}}$ for which each coordinate $x_r,\ r \in L$ has a value which is a partial sum of $\sum_{y\in\{0,1\}^n} \lambda^y$, i.e.

$$x_r = \sum_{y\in Q(r)\subseteq\{0,1\}^n} \lambda^y \tag{2.3}$$

for some subset $Q(r)$ of $\{0,1\}^n$. But we will also see that there are lifted coordinates $x_r,\ r \in L$ only for a small portion of the partial sums of $\sum_{y\in\{0,1\}^n} \lambda^y$. This provides the first indication that a more comprehensive lifting should append a coordinate for each partial sum.

We will see through the course of the chapter that the notion of partial summation is central to understanding the operators of Chapter 1, and that partial summation provides a perspective from which all of the results of Chapter 1 arise naturally. But as we indicated, and as we will see in greater detail, the possibilities for using partial summation, and its underlying idea of decomposition, are not exhausted by the procedures of Chapter 1, and these ideas themselves will suggest a broader lifting. In particular we will see near the end of the chapter that decomposition and partial summation have a natural measure theoretic interpretation, as do the vectors $m^{[p,q]}$ of the first chapter (Definition 1.50), and we will see that the measure theoretic connection will also point the way to a broader lifting, as well as to a generalization (which we will begin to explore in this chapter) of the vectors $m^{[p,q]}$.

2.1 The Partial Sum Interpretation

2.1.1 Introduction

The fundamental idea underlying both convexification and the Lovász Schrijver operators is as follows. Considering that, by definition, any $\bar{\bar{x}} \in H|_I = Cone(\{y \in \{0,1\}^{n+1} : y_0 = 1\})$ can be written (not uniquely) as

$$\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r, \ \alpha \geq 0 \tag{2.4}$$

(recall from Section 1.2.1 that the projections $\bar{\bar{z}}^r$ of the zeta vectors on the emptyset and singleton set coordinates are exactly the vectors $\{y \in \{0,1\}^{n+1} : y_0 = 1\}$) the question of whether or not $\bar{\bar{x}} \in Cone(K)$, where $K \subseteq \{y \in \{0,1\}^{n+1} : y_0 = 1\}$, is the question of whether or not there exists such a representation (2.4) of $\bar{\bar{x}}$ for which $\alpha_r = 0$ wherever $\bar{\bar{z}}^r \notin K$. Note that where the cone $\bar{K} \subseteq H|_I$ is such that $\bar{K} \cap \{0,1\}^{n+1} = K \cup \{0\}$, then for each $k \in \bar{K}^*$, we can multiply $\bar{\bar{x}}$ by k, and enforce that

$$0 \leq k^T \bar{\bar{x}} = \sum_{r \in L} \alpha_r k^T \bar{\bar{z}}^r. \tag{2.5}$$

Clearly this constraint must be satisfied if there is indeed a representation (2.4) for which $\alpha_r = 0$ wherever $\bar{\bar{z}}^r \notin K$, since $k^T \bar{\bar{z}}^r \geq 0$ for all $\bar{\bar{z}}^r \in K$. But the converse does not hold, since though $k^T \bar{\bar{z}}^r \geq 0, \ \forall k \in \bar{K}^*$, iff $\bar{\bar{z}}^r \in K$, we may still have some of the $\alpha_r > 0$ even where $k^T \bar{\bar{z}}^r < 0$, and the sum can still be nonnegative, so long as the negative contributions of $\alpha_r k^T \bar{\bar{z}}^r, \ \bar{\bar{z}}^r \notin K$ are offset by positive contributions from terms $\alpha_r k^T \bar{\bar{z}}^r, \ \bar{\bar{z}}^r \in K$.

We will see that these operators can all be understood as attempts to consider *partial sums* of the sum (2.4) defining the vector $\bar{\bar{x}}$, and to enforce membership in $\bar{K}$ for those partial sums. That is, given the vector $\bar{\bar{x}}$, they seek to find vectors $\bar{\bar{x}}'$ such that there exists a representation

$$\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r, \ \alpha \geq 0 \tag{2.6}$$

for which

$$\bar{\bar{x}}' = \sum_{r \in T \subseteq L} \alpha_r \bar{\bar{z}}^r \tag{2.7}$$

for some $T \subseteq L$. (We will see, however, that none of these operators is actually guaranteed to find such vectors $\bar{\bar{x}}'$ so long as we insist on the $\alpha \geq 0$ requirement. But if this requirement is eliminated then such vectors can be found.) The operators then multiply these vectors by $k \in \bar{K}^*$ and enforce that

$$0 \leq k^T \bar{\bar{x}}' = \sum_{r \in T \subseteq L} \alpha_r k^T \bar{\bar{z}}^r. \tag{2.8}$$

Naturally where this is repeated over $T_1, \ldots T_j \subseteq L$ satisfying $T_1 \cup \cdots \cup T_j = L$, then (even where we cannot ensure that the representation is via an α that is ≥ 0) we obtain considerably stronger conditions than the original

$$0 \leq k^T \bar{\bar{x}} = \sum_{r \in L} \alpha_r k^T \bar{\bar{z}}^r. \tag{2.9}$$

(One way to think of partial summation is as an extended version of abstract disjunctive programming, in that the partial sums are meant to belong to the cone of particular subsets of the integer points. We will see more on this later.)

In the extreme case where the sets T_i each contain exactly one element $r \in L$, by the definition of $\bar{K}$ we obtain

$$0 \leq k^T \bar{\bar{x}}' = \alpha_r k^T \bar{\bar{z}}^r, \ \forall k \in \bar{K}^* \Rightarrow \alpha_r > 0 \text{ only if } \bar{\bar{z}}^r \in K \tag{2.10}$$

and even without a general apriori assumption that $\alpha \geq 0$, we still have $e_0 \in \bar{K}^*$ and $e_0^T \bar{\bar{z}}^r = 1 \ \forall r$, which implies that in this case $\alpha_r \geq 0 \ \forall r$. So if this were to be repeated for all elements $r \in L$ then this would indeed guarantee that $\bar{\bar{x}} \in Cone(K)$.

2.1.2 N and N^+

To obtain the partial sums we first construct the expression

$$\sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T m, \ \alpha \geq 0 \tag{2.11}$$

where m is a vector that satisfies

$$(\bar{\bar{z}}^r)^T m \in \{0, 1\}, \ \forall r \in L \tag{2.12}$$

so that

$$\sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T m = \sum_{r : (\bar{\bar{z}}^r)^T m = 1} \alpha_r z^r, \ \alpha \geq 0. \tag{2.13}$$

Conceptually there are two steps in the procedure. The first is to figure out what the matrix

$$\sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T, \ \alpha \geq 0 \tag{2.14}$$

looks like, and the second is to find vectors m such that

$$(\bar{\bar{z}}^r)^T m \in \{0,1\},\ \forall r \in L. \tag{2.15}$$

The convexification operator (in conic form) N^0, and the Lovász Schrijver operators N and N^+ both use

$$\{e_1, \ldots, e_n, e_0 - e_1, \ldots, e_0 - e_n\} \tag{2.16}$$

as the vectors m. It is easy to see that these all satisfy $(\bar{\bar{z}}^r)^T m \in \{0,1\},\ \forall r \in L$.

Lemma 2.1

$$\sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T e_i = \sum_{r \in L: s_i \in r} \alpha_r \bar{\bar{z}}^r \tag{2.17}$$

$$\sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T (e_0 - e_i) = \sum_{r \in L: s_i \notin r} \alpha_r \bar{\bar{z}}^r \ \square \tag{2.18}$$

All three of these operators check for each $i = 1, \ldots, n$, that the part of x made up of linear combinations of vectors $\bar{\bar{z}}^r$ where each r contains s_i (so that $\bar{\bar{z}}^r_i = 1$) belongs to $\bar{K}$, and that the part of x made up linear combinations of vectors $\bar{\bar{z}}^r$ where each r does not contain s_i (so that $\bar{\bar{z}}^r_i = 0$) belongs to $\bar{K}$.

The difference between these operators lies in the other conceptual part of the procedure, namely how to characterize the matrices

$$Y = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T,\ \alpha \geq 0. \tag{2.19}$$

The convexification operator N^0 notes the following.

1. $Ye_0 = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r = x$
2. $Y_{i,i} = Y_{i,0} = Y_{0,i}$ (since for all r, $\bar{\bar{z}}^r_0 = 1$ and $(\bar{\bar{z}}^r_i)^2 = \bar{\bar{z}}^r_i$)

The N operator notes additionally

3. $Y = Y^T$

(It follows from Lemma 1.42 that with this addition the matrices of the form

$$\sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T \tag{2.20}$$

are completely characterized.) The N^+ operator adds the further requirement

4. $Y \succeq 0$

It should be noted that this is still not enough to characterize the matrices

$$Y = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T, \ \alpha \geq 0 \tag{2.21}$$

completely (positive semidefiniteness is a necessary but not sufficient condition for submatrices of W^x (Definition 1.23) to correspond to projections of vectors in $x \in H$ (Definition 1.31).

The conclusion is that these three operators are all of the same type. All of them seek to break up x into the same "pieces" (partial sums), and then check those pieces for membership in $\bar{K}$. The difference is that N and N^+ do this more rigorously. They observe that not just any vectors can serve as these pieces of x; those pieces are interrelated, and they place constraints accordingly.

The $N(\bar{K}, \bar{K}')$ operator is a strengthening of the other type. Instead of multiplying the matrices Y by $\{e_i\}$ and $\{e_0 - e_i\}$, and checking for membership of the product in $\bar{K}$, it multiplies by vectors generating a cone ($\bar{K'}^*$) that includes the vectors $\{e_i\}$ and $\{e_0 - e_i\}$ (subject to the condition that $Cone(K) \subseteq \bar{K}'$). The strongest such choice is generally $N(\bar{K}, \bar{K})$ as $\bar{K}$ is typically the smallest cone we know of (with a polynomial time separation oracle) containing $Cone(K)$, and so it will yield the largest $\bar{K'}^*$. The problem, as was pointed out by Lovász and Schrijver, is that even if polynomial time separation oracles exist for $\bar{K}$ and $\bar{K}'$, there is no guarantee that one exists for $N(\bar{K}, \bar{K}')$.

2.1.3 Reinterpreting $\bar{N}$

The $\bar{N}$ operator is a bit more interesting.

Lemma 2.2 *Let $\mathcal{M}^1$ denote the set of vectors m that satisfy*

$$(\bar{\bar{z}}^r)^T m \in \{0, 1\}, \ \forall r \in L \tag{2.22}$$

The cone generated by the elements of $\mathcal{M}^1$ is the cone generated by

$$\{e_i, e_0 - e_i : i = 1, \ldots, n\} \tag{2.23}$$

i.e. the vectors of $\mathcal{M}^1$ can all be generated from the e_i and $e_0 - e_i$ vectors.

(Actually, there are no vectors in $\mathcal{M}^1$ other than e_i and $e_0 - e_i$, but we have no need to prove that.)

Proof: Any vector that has a $0, 1$ dot product with every $\bar{\bar{z}}^r$ must belong to the polar of the cone generated by the vectors $\bar{\bar{z}}^r$, and we have seen already (Lemma 1.46) that this cone is generated by $\{e_i, e_0 - e_i : i = 1, \ldots, n\}$. □

So there are no other (relevant) vectors in $\mathcal{M}^1$ besides e_i and $e_0 - e_i$. But observe that for a matrix of the form

$$Y = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T, \ \alpha \geq 0 \tag{2.24}$$

the j'th column is

$$Ye_j = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r (\bar{\bar{z}}^r)^T e_j = \sum_{r \in L : s_j \in r} \alpha_r \bar{\bar{z}}^r \tag{2.25}$$

which is the partial sum of x over the lattice elements containing s_j (i.e. over the points in $\{y : y \in \{0,1\}^{n+1}, y_0 = 1\}$ for which $y_j = 1$). Moreover, the i, j entry of that matrix is

$$e_i Y e_j = \sum_{r \in L : s_j \in r} \alpha_r \bar{\bar{z}}^r_i = \tag{2.26}$$

$$\sum_{r \in L : s_j, s_i \in r} \alpha_r \tag{2.27}$$

by the definition of the vectors $\bar{\bar{z}}^r$. Considering that each

$$x_i = \sum_{r \in L : s_i \in r} \alpha_r \tag{2.28}$$

and

$$x_0 = \sum_{r \in L} \alpha_r \tag{2.29}$$

it therefore makes sense to think of each x_i as a partial sum of x_0, and then

$$Y_{i,j} = \sum_{r \in L : s_j, s_i \in r} \alpha_r \tag{2.30}$$

is also a partial sum, which we will denote $x_{\{s_i\} \cup \{s_j\}}$, or more briefly as $x_{i,j}$, as it similarly represents the contribution to x_0 of those α_r where r contains both s_i and s_j (i.e. of those points with a 1 in positions i and j).

But then there is no reason to settle for defining only variables $x_{i,j}$. We can define variables for other partial sums as well.

Definition 2.3 *Given* $\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r$, *for each* $q \in L$ *define*

$$x_q = \sum_{r \in L : q \subseteq r} \alpha_r. \tag{2.31}$$

Where we are given $\bar{\bar{x}}$ *but not* α, *there can be many possible choices for* x_q. *Note that technically* x_q *is a function of* α, *but we will usually suppress the dependence notation, and write merely* x_q. *Denote the vector in* R^L *with each* q*'th coordinate equal to* x_q *as* x.

Lemma 2.4

$$x_{\{s_i\}} = \bar{\bar{x}}_i \tag{2.32}$$

$$x_\emptyset = \bar{\bar{x}}_0 \tag{2.33}$$

and therefore x is a lifting of $\bar{\bar{x}}$. □

Clearly, as representations $\bar{\bar{x}} = \sum_{r\in L} \alpha_r \bar{\bar{z}}^r$ are not unique, neither are the possible choices for x_q. Ideally we would like to restrict ourselves to choices for x_q that arise from representations in which $\alpha \geq 0$, but as we noted, we do not have any guaranteed way of doing this. For the case of $\bar{\bar{z}}^r$, however, we know exactly what the new variables will look like when α is constrained to be ≥ 0.

Lemma 2.5 *For all $p \in L$, there is a unique representation*

$$\bar{\bar{z}}^p = \sum_{r\in L} \alpha_r \bar{\bar{z}}^r, \alpha \geq 0 \tag{2.34}$$

namely, $\alpha_p = 1,\ \alpha_r = 0\ \forall r \neq p$. Thus for $q \in L$, the unique choice for z_q^p (w.r.t. representations $\alpha \geq 0$) is

$$z_q^p = \begin{cases} 1 : q \subseteq p \\ 0 : \ \textit{otherwise} \end{cases} \tag{2.35}$$

Proof: This says that the vectors $\bar{\bar{z}}^r$ are the extreme rays of the cone they generate. To see this, consider

$$\bar{\bar{z}}^t = \sum_{r\in L} \alpha_r \bar{\bar{z}}^r,\ \alpha \geq 0 \tag{2.36}$$

and assume that at least two coefficients $\alpha_p, \alpha_q > 0$ (if only one coefficient is positive then the result is trivial). We must have

$$\sum_{r\in L} \alpha_r = \bar{\bar{z}}_0^t = 1. \tag{2.37}$$

Since $\bar{\bar{z}}^p \neq \bar{\bar{z}}^q$, there must be some $u \in L$ such that, say, $\bar{\bar{z}}_u^p = 1$ while $\bar{\bar{z}}_u^q = 0$, so by construction we would have

$$0 < \alpha_p \leq \sum_{r\in L} \alpha_r \bar{\bar{z}}_u^r = \bar{\bar{z}}_u^t \leq \sum_{r\in L, r\neq q} \alpha_r = 1 - \alpha_q < 1 \tag{2.38}$$

contradicting the fact that $\bar{\bar{z}}^t$ is a $0,1$ vector. □

(Note that a similar statement to that of the lemma would hold for any set of $0,1$ vectors all of which had a 1 in some given location.)

Now that we have generalized the lifted variables of the N operator, we will construct a related generalization of the matrices Y. The original matrices Y had a j'th column for each partial sum of x taken over the lattice elements containing s_j (i.e. over the points in $\{y : y \in \{0,1\}^{n+1}, y_0 = 1\}$ for which $y_j = 1$). For a given $l > 1$, for each $q \in L,\ |q| \leq l$, we will now append a column to Y representing the partial sum of x taken over the lattice elements containing q (or, in other words, corresponding to the points in $\{y : y \in \{0,1\}^{n+1}, y_0 = 1\}$ for which $y_j = 1$ for each $j \in \{1, \ldots, n\} : s_j \in q$).

Definition 2.6 *Given* $\bar{\bar{x}} = \sum_{r\in L} \alpha_r \bar{\bar{z}}^r$*, define the matrix* $\bar{X}^l(\bar{\bar{x}})$ *to be the matrix with rows corresponding to the empty set and to each singleton* $s_1, \ldots, s_n$*, and a column for each* $q \in L$ *for which* $|q| \leq l$ *where the* q *column of* $\bar{X}^l(\bar{\bar{x}})$ *is defined to be*

$$\sum_{r\in L:q\subseteq r} \alpha_r \bar{\bar{z}}^r. \tag{2.39}$$

Again, where we are given $\bar{\bar{x}}$ but not α there will be many possible matrices $\bar{X}^l(\bar{\bar{x}})$. As above, technically $\bar{X}^l$ is a function of α. If the dependence is clear, though, we will suppress all dependence notation and write simply $\bar{X}^l$.

Lemma 2.7 *Where* $p, q \in L,\ |p| \leq 1,\ |q| \leq l$ *then given a representation* $\bar{\bar{x}} = \sum_{r\in L} \alpha_r \bar{\bar{z}}^r$,

$$\bar{X}^l_{p,q} = x_{p\cup q}. \tag{2.40}$$

Proof:

$$X^l_{p,q} = \sum_{r\in L:q\subseteq r} \alpha_r \bar{\bar{z}}^r_p = \sum_{r\in L:q\subseteq r, p\subseteq r} \alpha_r = x_{p\cup q}\ \square \tag{2.41}$$

So the expanded matrices generalize the rule

$$Y_{i,j} = x_{\{s_i\}\cup\{s_j\}}. \tag{2.42}$$

Lemma 2.8 *Given* $\bar{\bar{x}} = \sum_{r\in L} \alpha_r \bar{\bar{z}}^r$*, then where* $x \in R^L$ *is an in Definition 2.3 and* z^r *is as in Definition 2.5,*

$$x = \sum_{r\in L} \alpha_r z^r. \tag{2.43}$$

Proof: This is a direct consequence of the way that we constructed the expansions of the vectors. For any $q \in L$,

$$\sum_{r\in L} \alpha_r z^r_q = \sum_{r\in L:q\subseteq r} \alpha_r = x_q.\ \square \tag{2.44}$$

Definition 2.9 *Given $\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r$, define the matrix $X(\bar{\bar{x}}) \in R^{L \times L}$ to be the matrix with a row and a column for each $q \in L$ where the q'th column of X is defined to be*

$$\sum_{r \in L : q \subseteq r} \alpha_r z^r. \tag{2.45}$$

Let $X^l(\bar{\bar{x}})$ be the square submatrix of $X(\bar{\bar{x}})$ with rows and columns corresponding to $q \in L : |q| \leq l$.

Again, where we are not given α, this matrix is not defined uniquely by $\bar{\bar{x}}$, and again we will suppress all dependence notation where it is not needed for clarity, and write simply X and X^l.

Consider now that for any $q \in L$,

$$\{r \in L : q \subseteq r\} = \{r \in L : 1 = z^r_q = (z^r)^T e_q\} \tag{2.46}$$

where e_q is the q'th unit vector in R^L. Thus

$$Xe_q = \sum_{r \in L : q \subseteq r} \alpha_r z^r = \sum_{r \in L : (z^r)^T e_q = 1} \alpha_r z^r = \tag{2.47}$$

$$\sum_{r \in L} \alpha_r z^r (z^r)^T e_q. \tag{2.48}$$

Since this is true for all q we conclude

Lemma 2.10

$$X = \sum_{r \in L} \alpha_r z^r (z^r)^T \ \square \tag{2.49}$$

This generalizes the $n + 1 \times n + 1$ matrices Y from above. It is also clear that

$$X_{p,q} = x_{p \cup q} \tag{2.50}$$

so that this is the matrix that Lovász and Schrijver denoted W^x, and that $\bar{X}^l$ is made up of the first $1 + n$ rows of X^l.

Note also that given any choice of $x \in R^L$, since we have seen already that the vectors z^r form a basis of R^L, there is an α such that $x = \sum_{r \in L} \alpha_r z^r$ (though we have no guarantee that $\alpha \geq 0$). So we can always add new variables corresponding, say, to all $q \in L : |q| \leq l+1$, and then we can use those values to fill in the entries of $\bar{X}^l$, and be guaranteed that the resulting matrix is made up of the first $1 + n$ rows of X^l for some representation $\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r$, i.e. we can be sure that the resulting matrix is in fact of the form $\bar{X}^l$. Formally,

Lemma 2.11 *Given $G \subseteq L$, and $J = \{p \cup q : p, q \in G\}$, and a vector $\hat{x}$ with coordinates corresponding to J, say that the square matrix A with rows and columns corresponding to G satisfies*

$$A_{p,q} = \hat{x}_{p \cup q}, \quad \forall p, q \in G. \tag{2.51}$$

Then, where hat indicates projection on the J coordinates, and tilde indicates projection on the G coordinates, there exists an $\alpha \in R^L$ (not necessarily unique) such that

$$\hat{x} = \sum_{r \in L} \alpha_r \hat{z}^r \tag{2.52}$$

and

$$A = \sum_{r \in L} \alpha_r \tilde{z}^r (\tilde{z}^r)^T. \tag{2.53}$$

In particular, for any selection of numbers x_q, $|q| \leq l+1$, the matrix with rows corresponding to the empty set and the singleton sets, and columns corresponding to each $r : |r| \leq l$, whose u, v entry is $x_{p \cup q}$ is a matrix of the form $\bar{X}^l$. Note also that where $\hat{x}$ is a vector with coordinates $q : |q| \leq l+1$, the matrix $\bar{X}^l$ is a unique function of $\hat{x}$ (regardless of α). Similarly, where $\hat{x}$ is a vector with coordinates $q : |q| \leq 2l$, the matrix X^l is a unique function of $\hat{x}$. □

(Note, however, that despite the formal functional dependence, we will usually suppress the dependence notation.)

Thus where we have added sufficiently many coordinates to $\bar{\bar{x}}$ to write the matrix $\bar{X}^l$ we are assured that this matrix is the first $1 + n$ rows of

$$\sum_{r \in L} \alpha_r \tilde{z}^r (\tilde{z}^r)^T \tag{2.54}$$

where $\tilde{z}^r$ is the projection of z^r to the coordinates $q \in L : |q| \leq l$, for some α.

At this point we are no longer restricted to the vectors e_i and $e_0 = e_i$ to obtain partial summations. In fact, for the full matrix $X = \sum_{r \in L} \alpha_r z^r (z^r)^T$, the rows m_p of the Möbius matrix of the lattice L satisfy

$$(z^r)^T m_p = \delta_{r,p} \tag{2.55}$$

and therefore $X m_p$ is the partial sum made up of a single contribution, namely $\alpha_p z^p$.

We have seen in the previous chapter (expression 1.130) that the vectors $m^{[u,v]}$ all satisfy

$$(m^{[u,v]})^T z^r \in \{0, 1\}, \ \forall r \in L \tag{2.56}$$

so those among these vectors whose nonzeroes are all in coordinates $q : |q| \leq l$ will satisfy

$$(\tilde{m}^{[u,v]})^T \tilde{z}^r \in \{0, 1\}, \ \forall r \in L \tag{2.57}$$

where the tilde indicates projection to the coordinates $q \in L : |q| \leq l$. Specifically, all $\tilde{m}^{[u,v]} : |v| \leq l$ qualify.

Lemma 2.12 *Let* $\bar{\bar{x}} = \sum_{r \in R} \alpha_r \bar{\bar{z}}^r$ *where the double bar, as usual, indicates that these vectors have coordinates corresponding only to the empty set and singletons. Then where* $u, v \in L, \ |v| \leq l$*,*

$$\bar{X}^l \tilde{m}^{[u,v]} = \sum_{r \in L : u = r \cap v} \alpha_r \bar{\bar{z}}^r \tag{2.58}$$

where the tilde symbol indicates that the vector has coordinates corresponding to $q \in L : |q| \leq l$*.*

Proof: The matrix $\bar{X}^l$ is the first $1 + n$ rows of X^l, which satisfies

$$X^l = \sum_{r \in L} \alpha_r \tilde{z} (\tilde{z}^r)^T \Rightarrow \tag{2.59}$$

$$X^l \tilde{m}^{[u,v]} = \sum_{r \in L} \alpha_r \tilde{z} (\tilde{z}^r)^T \tilde{m}^{[u,v]}. \tag{2.60}$$

But by construction

$$(\tilde{z}^r)^T \tilde{m}^{[u,v]} = (z^r)^T m^{[u,v]} = \tag{2.61}$$

$$\sum_{t \in L : t \subseteq r} m_t^{[u,v]} = \sum_{t \in L : t \subseteq r, t \subseteq v} m_t^u = \sum_{t \in L : t \subseteq r \cap v} m_t^u = \tag{2.62}$$

$$(z^{r \cap v})^T m^u = \delta_{u, r \cap v} \Rightarrow \tag{2.63}$$

$$X^l \tilde{m}^{[u,v]} = \sum_{r \in L : u = r \cap v} \alpha_r \tilde{z}^r. \tag{2.64}$$

Taking projections on the first $1 + n$ rows gives the lemma. □

Corollary 2.13

$$\{\bar{\bar{x}} \in R^{1+n} : \exists \bar{X}^l \ s.t. \ \bar{X}^l \tilde{m}^{[u,v]} \in \bar{K}\} =$$

$$\{\bar{\bar{x}} \in R^{1+n} : \exists \alpha \in R^L \ s.t. \ \bar{\bar{x}} = \sum_{r \in R} \alpha_r \bar{\bar{z}}^r \ and \sum_{r \in L : u = r \cap v} \alpha_r \bar{\bar{z}}^r \in \bar{K}\} \ \Box \tag{2.65}$$

The intersection of the sets (2.65) over all $m^{[u,v]}$, $|v| \leq l$ is $\bar{N}^l(\bar{K})$. Note that where $u \subseteq v$, the set $\{r \in L : u = r \cap v\}$ are those r that from among the s_j in v, contain exactly those s_j that are in u. For example, if

$$u = \{s_1, s_2\}, \quad v = \{s_1, s_2, s_3, s_4\} \tag{2.66}$$

then $\{r \in L : u = r \cap v\}$ is the set of lattice elements that contain s_1 and s_2 but not s_3 or s_4. So the set

$$\{r \in L : \{s_1, s_2\} = r \cap \{s_1, s_2, s_3, s_4\}\} \tag{2.67}$$

is the set of lattice elements whose corresponding points in $\{0,1\}^n$ have the configuration $(1,1,0,0)$ in their first four coordinates. In words,

Corollary 2.14 *The set $\bar{N}^l(\bar{K})$ is made up of those points in $\bar{\bar{x}} \in R^{1+n}$ for which a representation exists, $\bar{\bar{x}} = \sum_{r \in R} \alpha_r \bar{\bar{z}}^r$, such that for every subset of size l or smaller of the coordinates $1, \ldots, n$, and every configuration of $0, 1$ values on each such subset, the part of x (via that representation) made up of those $\bar{\bar{z}}^r$ that possess that configuration is a vector belonging to $\bar{K}$.* □

Note that it is not actually necessary to consider every $\leq l$ sized subset of the coordinates. It is easy to see that it suffices to consider merely the size l subsets.

For any given $u, v \in L$, $u \subseteq v$, consider the $0,1$ points $\bar{\bar{z}}^r$ that have 1's in their u coordinates and 0's in their $v - u$ coordinates. These are those $\bar{\bar{z}}^r$ for which

$$\prod_{s_i \in u} \bar{\bar{z}}_i^r \prod_{s_i \in v-u} (1 - \bar{\bar{z}}_i^r) = 1 \tag{2.68}$$

while for all other $\bar{\bar{z}}^r$ this product is zero. So therefore,

$$\sum_{r \in L : u = r \cap v} \alpha_r \bar{\bar{z}}^r = \sum_{r \in L} \alpha_r \left(\prod_{s_i \in u} \bar{\bar{z}}_i^r \prod_{s_i \in v-u} (1 - \bar{\bar{z}}_i^r) \right) \bar{\bar{z}}^r. \tag{2.69}$$

So by Lemma 2.12, demanding that

$$\bar{X}^l \tilde{m}^{[u,v]} \in \bar{K} \tag{2.70}$$

is the same as demanding that

$$\sum_{r \in L} \alpha_r \left(\prod_{s_i \in u} \bar{\bar{z}}_i^r \prod_{s_i \in v-u} (1 - \bar{\bar{z}}_i^r) \right) \bar{\bar{z}}^r \in \bar{K}. \tag{2.71}$$

Notice also that for any $k \in \bar{K}^*$,

$$\{r : \left(\prod_{s_i \in u} \bar{\bar{z}}_i^r \prod_{s_i \in v-u} (1 - \bar{\bar{z}}_i^r) \right) k^T \bar{\bar{z}}^r \geq 0\} = \tag{2.72}$$

$$\{r : u = r \cap v, k^T \bar{\bar{z}}^r \geq 0\} \cup \{r : u \neq r \cap v\}. \tag{2.73}$$

The inequality in the first expression is a valid polynomial inequality for all points $\bar{\bar{z}}^r \in K$. So attempting to establish that those $\bar{\bar{z}}^r$ for which $r : u = r \cap v$ that contribute to x satisfy the linear inequality $k^T \bar{\bar{z}}^r \geq 0$ is the same as attempting to establish that those $\bar{\bar{z}}^r$ that contribute to x satisfy the polynomial inequality

$$\left(\prod_{s_i \in u} \bar{\bar{z}}_i^r \prod_{s_i \in v-u} (1 - \bar{\bar{z}}_i^r) \right) k^T \bar{\bar{z}}^r \geq 0. \tag{2.74}$$

This is reminiscent of the original definition given by Sherali and Adams for their procedure, in which they linearize inequalities of this form. We will see more on this connection soon, but we will not go through the motions of proving formal equivalence.

With this qualitative characterization of $\bar{N}$ in mind, let us compare $\bar{N}^l(\bar{K})$ to the repeated N operator, $N^l(\bar{K})$. Let us first consider the case $l = 2$. Given $\bar{\bar{x}} \in R^{n+1}$, then $\bar{\bar{x}} \in N(N(\bar{K}))$ iff there exists a representation

$$\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r \tag{2.75}$$

such that for each $i = 1, \ldots, n$,

$$\sum_{r \in L : s_i \in r} \alpha_r \bar{\bar{z}}^r \in N(\bar{K}) \text{ and} \tag{2.76}$$

$$\sum_{r \in L : s_i \notin r} \alpha_r \bar{\bar{z}}^r \in N(\bar{K}). \tag{2.77}$$

But $\sum_{r \in L : s_i \in r} \alpha_r \bar{\bar{z}}^r \in N(\bar{K})$ itself means that there exists a representation

$$\sum_{r \in L : s_i \in r} \alpha_r \bar{\bar{z}}^r = \sum_{r \in L} \beta_r \bar{\bar{z}}^r \tag{2.78}$$

such that for each $j = 1, \ldots, n$,

$$\sum_{r \in L : s_j \in r} \beta_r \bar{\bar{z}}^r \in \bar{K} \text{ and} \tag{2.79}$$

$$\sum_{r \in L : s_j \notin r} \beta_r \bar{\bar{z}}^r \in \bar{K}. \tag{2.80}$$

The N^2 procedure does not require the representation β to be the same as the representation α, but if it did, then this would mean that for each i and $j = 1, \ldots, n$,

$$\sum_{r \in L : s_i \in r, s_j \in r} \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{2.81}$$

$$\sum_{r \in L: s_i \in r, s_j \notin r} \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{2.82}$$

$$\sum_{r \in L: s_i \notin r, s_j \in r} \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{2.83}$$

$$\sum_{r \in L: s_i \notin r, s_j \notin r} \alpha_r \bar{\bar{z}}^r \in \bar{K}. \tag{2.84}$$

But this is exactly $\bar{N}^2(K)$. The difference between the two is thus that $\bar{N}^2$ insists that the representations α and β must be the same, while N^2 does not. It is thus possible that a vector $\bar{\bar{x}}$ for which no representation $\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r$ exists that satisfies the four constraints above may nevertheless belong to $N^2(\bar{K})$ so long as appropriate β representations exist. The situation is similar for higher l as well. Thus $\bar{N}^l$ and N^l both look for the same partial sums, but N^l is far less consistent in the way that it constructs these partial sums.

2.1.4 Polynomial Constraints

Until this point we have been defining $\bar{N}$ with respect to integer sets that are construed as the $0, 1$ solutions for systems of linear constraints. The following theorem shows that polynomial constraints can be used just as well.

Theorem 2.15 *Let K and K^e be as in Definition 1.36. The polynomial inequality*

$$\sum_{V \subseteq \{1,\ldots,n\}} \beta_V \prod_{i: i \in V} \bar{\bar{z}}_i \geq 0 \tag{2.85}$$

is valid for every $\bar{\bar{z}}^r \in K$ iff the linear inequality

$$\sum_{v \in L} \beta_v x_v \geq 0 \tag{2.86}$$

(where x_v is as in Definition 2.3) under the one to one correspondence

$$V \subseteq \{1, \ldots, n\} \leftrightarrow v \in L : v = \bigcup_{i: i \in V} \{s_i\} \tag{2.87}$$

is valid for every $x \in K^e$.

Proof:

$$\sum_{V \subseteq \{1,\ldots,n\}} \beta_V \prod_{i: i \in V} \bar{\bar{z}}_i = \sum_{v \in L} \beta_v z_v^r \tag{2.88}$$

for every $r \in L$, and (2.86) is valid for every x in the cone of $z^r \in K^e$ iff it is valid for every $z^r \in K^e$ (since the z^r are extreme rays of the cone they generate). □

Thus in lifting $\bar{\bar{x}}$ to R^L we obtain the opportunity to enforce a linear inequality for each valid polynomial inequality (that does not involve powers other than 1) on K. Naturally, so long as there are a sufficient number of rows defined in the (restrictions of the) matrices X^l, we will be able to enforce these linear inequalities on each column of the matrix as well. (One can check that the linear constraints $k^T\bar{\bar{x}}' \geq 0,\ k \in \bar{K}^*$ applied to the partial sum column vectors $\bar{\bar{x}}' = \bar{X}^l\tilde{m}^{[u,v]}$ correspond by this reasoning to the polynomial constraints $(\prod_{s_i \in u} \bar{\bar{z}}_i^r \prod_{s_i \in v-u}(1 - \bar{\bar{z}}_i^r))k^T\bar{\bar{z}}^r \geq 0$.) Thus for example if K is the set of solutions from among $\{y \in \{0,1\}^{n+1} : y_0 = 1\}$ to a system of quadratic constraints of the form

$$\sum_{i,j=0}^{n} \beta_{i,j} x_i x_j \geq 0 \tag{2.89}$$

and $\bar{K}$ is the relaxation of K defined as the set of points in R^{n+1} that satisfies the constraints (2.89) along with the constraints $0 \leq x_h \leq x_0,\ h = 1, \ldots, n$, then the appropriate adaptation of $\bar{N}^l(\bar{K})$ is to form the submatrix $\hat{X}^l$ with rows for the empty set, the singletons and the doubles, and columns for each $q \in L : |q| \leq l$, and to enforce the linear inequalities

$$\sum_{i,j} \beta_{i,j} x_{i,j} \geq 0 \tag{2.90}$$

on the vectors $\hat{X}^l\tilde{m}^{[u,v]}$. (The subscript indices "i" and "j" each refer to a lattice element, "i" to s_i if $i \geq 1$ and to $\emptyset$ if $i = 0$, and similarly for "j", and the index i, j refers to the lattice element that is the union of the elements corresponding to i and j, so that $2, 3$ would refer to the lattice element $\{s_2, s_3\}$, and $0, 2$ would refer to the lattice element $\{s_2\}$, and $0, 0$ would refer to the lattice element $\emptyset$.) So where $l = n$ and where $x \in R^L$ is the (unique) vector corresponding to $\hat{X}^n$, and x is represented (uniquely) as $x = \sum_{r \in L} \alpha_r z^r$, then each vector in $y \in \{0,1\}^n$, corresponding to a lattice element r, is naturally a configuration of 0's and 1's in its n coordinates, and there is therefore some $m^{[u,v]}$ (as per Corollary 2.14) such that $\hat{X}^n m^{[u,v]} = \alpha_r \hat{z}^r$, i.e. the partial sum of $\hat{x}$ that is contributed by the single point $y \in \{0,1\}^n$ (where the hat indicates projection on the empty set, singletons and doubles coordinates). So applying the constraints

$$0 \leq (\hat{X}^n m^{[u,v]})_0 = \alpha_r \hat{z}_\emptyset^r = \alpha_r \tag{2.91}$$

implies that $\alpha \geq 0$, and then applying

$$0 \leq \sum_{i,j} \beta_{i,j} (\hat{X}^n m^{[u,v]})_{i,j} = \sum_{i,j} \beta_{i,j} (\alpha_r \hat{z}^r)_{i,j} \tag{2.92}$$

for each of the constraints defining $\bar{K}$ ensures that either $\alpha = 0$ or that $\bar{\bar{z}}^r \in K$ (where the double bar indicates projection on the empty set and singletons coordinates), so that in either case

$$\alpha_r \bar{\bar{z}}^r \in Cone(K) \tag{2.93}$$

which implies that $\bar{\bar{x}} \in Cone(K)$ as well. So this adaptation is also guaranteed to satisfy $\bar{N}^n(\bar{K}) = Cone(K)$.

2.1.5 Two Stepping Stones to the Lasserre Operator

Observe that $\bar{N}^l$ makes no specific attempt at ensuring that the α representations satisfy $\alpha \geq 0$. We noted that we do not have as yet any tools that will guarantee this (for $l < n$), but positive semidefiniteness can be used at least as a necessity condition. Based on the above characterization of the difference between $\bar{N}^l$ and N^l we can construct a new operator, to be denoted $\bar{N}^+$ such that $(\bar{N}^+)^l$ is stronger than $(N^+)^l$, and such that $(\bar{N}^+)^l$ has the same relationship to $(N^+)^l$ as does $\bar{N}^l$ to N^l.

In addition to constructing the matrix $\bar{X}^l$, this operator will also construct the matrix $\hat{X}^{l-1}$ with columns corresponding to $q \in L : |q| \leq l-1$, but with rows corresponding to the pairs as well (so both of these matrices are determined by the same coordinates $x_q,\ q : |q| \leq l+1$). Notice that each column of the matrix, $\hat{X}^{l-1}$, and more generally each vector $y = \hat{X}^{l-1}\tilde{m}^{[p,q]},\ |q| \leq l-1$, is a vector with a coordinate for the empty set, each singleton and each pair. Each such vector therefore uniquely determines a matrix W^y with $W^y_{u,v} = y_{u \cup v}$ where $|u|,\ |v| \leq 1$. This operator will, in addition to requiring that $\bar{X}^l \tilde{m}^{[p,q]} \in \bar{K}$ for all $q : |q| \leq l$, also require that all of the vectors $y = \hat{X}^{l-1}\tilde{m}^{[p,q]},\ |q| \leq l-1$ satisfy $W^y \succeq 0$. Formally,

Definition 2.16 *Let $\hat{x}$ be an expansion of the vector $\bar{\bar{x}}$ with coordinates corresponding to all $q \in L : |q| \leq l+1$. Let $\bar{X}^l$ and X^l be as in Definitions 2.6 and 2.9 respectively, and let $\hat{X}^{l-1}$ be the submatrix of X^{l-1} with rows corresponding to the empty set, singletons, and doubles, so that both $\hat{X}^{l-1}$ and $\bar{X}^l$ are unique functions of $\hat{x}$. Recall also that where $\bar{x}$ is a vector with coordinates corresponding to the empty set, the singletons and the pairs, the matrix $W^{\bar{x}}$ is the $(1+n) \times (1+n)$ matrix with entries $W^{\bar{x}}_{p,q} = \bar{x}_{p \cup q}$, where $|p|, |q| \leq 1$. Define the set $(\bar{N}^+)^l(\bar{K})$ to be the set of points $\bar{\bar{x}} \in R^{n+1}$ that satisfy*

1. *$\exists \hat{x}$ s.t. $\bar{X}^l(\hat{x})\tilde{m}^{[p,q]} \in \bar{K},\ \forall q$ with $|q| \leq l$ and*

2. *$W^{\hat{X}^{l-1}(\hat{x})\tilde{m}^{[p,q]}} \succeq 0,\ \forall q$ with $|q| \leq l-1$*

Observe that $(\bar{N}^+)^1(\bar{K}) = N^+(\bar{K})$.

(Naturally we could tailor the procedure to handle polynomial constraints by considering matrices with more rows. The same will hold for the next procedure to be introduced. Note also that we could have referred to the matrix $W^{\bar{x}}$ as $X^2(\bar{x})$.)

Let us now analyze $(N^+)^l$ in the same manner as we analyzed N^l above. Again let us first consider the case $l = 2$. A point $\bar{\bar{x}} \in N^+(N^+(\bar{K}))$ iff there exists a representation

$$\bar{\bar{x}} = \sum_{r \in L} \alpha_r \bar{\bar{z}}^r \tag{2.94}$$

such that for each $i = 1, \ldots, n$,

$$\sum_{r \in L: s_i \in r} \alpha_r \bar{\bar{z}}^r \in N^+(\bar{K}) \text{ and} \tag{2.95}$$

$$\sum_{r \in L: s_i \notin r} \alpha_r \bar{\bar{z}}^r \in N^+(\bar{K}) \tag{2.96}$$

and such that the corresponding expansion $\bar{x} = \sum_{r \in L} \alpha_r \bar{z}^r$ of $\bar{\bar{x}}$ to include pairs coordinates (where $\bar{z}^r$ is the expansion of $\bar{\bar{z}}^r$ to include pairs coordinates) satisfies

$$W^{\bar{x}} \succeq 0. \tag{2.97}$$

But $\sum_{r \in L: s_i \in r} \alpha_r \bar{\bar{z}}^r \in N^+(\bar{K})$ itself means that there exists a representation

$$\sum_{r \in L: s_i \in r} \alpha_r \bar{\bar{z}}^r = \sum_{r \in L} \beta_r \bar{\bar{z}}^r \tag{2.98}$$

such that for each $j = 1, \ldots, n$,

$$\sum_{r \in L: s_j \in r} \beta_r \bar{\bar{z}}^r \in \bar{K} \text{ and} \tag{2.99}$$

$$\sum_{r \in L: s_j \notin r} \beta_r \bar{z}^r \in \bar{K} \tag{2.100}$$

and such that the expansion

$$y = \sum_{r \in L} \beta_r \bar{z}^r \tag{2.101}$$

satisfies

$$W^y \succeq 0. \tag{2.102}$$

As above the $(N^+)^2$ procedure does not require the representation β to be the same as the representation α, but if it did, then this would mean that for each i and $j = 1, \ldots, n$,

$$\sum_{r \in L: s_i \in r, s_j \in r} \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{2.103}$$

$$\sum_{r \in L: s_i \in r, s_j \notin r} \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{2.104}$$

$$\sum_{r \in L: s_i \notin r, s_j \in r} \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{2.105}$$

$$\sum_{r \in L: s_i \notin r, s_j \notin r} \alpha_r \bar{\bar{z}}^r \in \bar{K} \tag{2.106}$$

(so this says exactly that there must be a matrix $\bar{X}^2$ satisfying

$$\bar{X}^2 \tilde{m}^{[p,q]} \in \bar{K}, \ \forall q : |q| \leq 2 \tag{2.107}$$

as above,) and moreover the matrices determined by the vectors

$$\sum_{r \in L: s_j \in r} \alpha_r \bar{z}^r \quad \text{and} \tag{2.108}$$

$$\sum_{r \in L: s_j \notin r} \alpha_r \bar{z}^r \tag{2.109}$$

must be positive semidefinite. But (2.108) and (2.109) are just the column vectors $\hat{X}^1 e_j$ and $\hat{X}^1(e_0 - e_j)$ whose entries are determined uniquely by $\bar{X}^2$. Thus this says that the matrices determined by the vectors $\hat{X}^1 m^{[p,q]}$, $|q| \leq 1$ must be positive semidefinite. But this is now exactly $(\bar{N}^+)^2$ as we defined it. It is easy to see that the situation is the same (by induction) for any $l > 2$ as well. So, as above, the difference between $(N^+)^l$ and $(\bar{N}^+)^l$ lies in the fact that $(\bar{N}^+)^l$ is more consistent in the way it treats partial sums.

A still stronger operator, in some ways more in the spirit of N^+, would be obtained if we constructed the full matrix X^l and demanded $X^l \succeq 0$.

Definition 2.17 *Let $\hat{x}$ be the expansion of $\bar{\bar{x}}$ obtained by appending coordinates for all $q \in L : |q| \leq 2l$. Thus X^l is a unique function of $\hat{x}$. Define*

$$(\bar{N}^*)^l(\bar{K}) = \{\bar{\bar{x}} \in R^{n+1} : \exists \hat{x} \ s.t. \ \bar{X}^l(\hat{x}) \tilde{m}^{[p,q]} \in \bar{K}, \ \forall q : |q| \leq l, \ X^l \succeq 0\} \tag{2.110}$$

Observe that $(\bar{N}^*)^1(\bar{K}) = N^+(\bar{K})$.

The following theorem, which states that this new operator is stronger than $(\bar{N}^+)^l$, will be proven later (Lemma 4.27).

Theorem 2.18

$$Cone(K) \subseteq (\bar{N}^*)^l(\bar{K}) \subseteq (\bar{N}^+)^l(\bar{K}) \ \square \tag{2.111}$$

2.1.6 The Lasserre Operator

Lasserre's operator (as applied to 0, 1 integer programs, see [Lau01]) is a strengthening of $\bar{N}^*$ obtained by replacing the linear constraints of $\bar{N}^*$ with the semidefinite constraints suggested in the previous chapter's discussion of $N(\bar{K}, \bar{K}')$. It can also be thought of as generalizing the spirit of $\bar{N}^+$. Specifically, where, as above, we have expanded $\bar{\bar{x}}$ to the vector $\hat{x}$ with coordinates for all $q \in L : |q| \leq 2l$, $l \geq 2$, then the matrix X^l is uniquely defined. Moreover for any valid constraint (on K^e), $k^T x \geq 0$, (recall from Theorem 2.15 that these correspond to the valid polynomial constraints on K,) where k has nonzero coordinates corresponding only to $q \in L$ such that $|q| \leq l$, and where $\tilde{k}$ is the restriction of k to those coordinates,

$$X^l(\hat{x})\tilde{k} = \sum_{r \in L} \alpha_r \tilde{z}^r (\tilde{z}^r)^T \tilde{k} = \sum_{r \in L} \left(\alpha_r (\tilde{z}^r)^T \tilde{k} \right) \tilde{z}^r \tag{2.112}$$

so for any $\bar{\bar{x}}$ that belongs to the cone of K, and can therefore be represented as

$$\sum_{r \in L : \bar{\bar{z}}^r \in K} \alpha_r \bar{\bar{z}}^r, \alpha \geq 0 \tag{2.113}$$

we must have, for $\hat{x}$ corresponding to that representation,

$$X^l(\hat{x})\tilde{k} = \sum_{r \in L : \bar{\bar{z}}^r \in K} \left(\alpha_r (\tilde{z}^r)^T \tilde{k} \right) \tilde{z}^r, \ \alpha_r (\tilde{z}^r)^T \tilde{k} \geq 0, \ \forall r \tag{2.114}$$

so that its projection belongs to $Cone(K)$, and we can therefore enforce the necessary linear(ized) constraints on the vector $X^l(\hat{x})\tilde{k}$. This is the polynomial inequality version of the $N(\bar{K}, \bar{K}')$ operator applied to the matrix X^l. But as we observed in the previous chapter (Section 1.2.3), we can also conclude that the *matrix* implied by the vector $X^l(\hat{x})\tilde{k}$ must also be positive semidefinite. If $l \geq 2$ then the matrix whose rows and columns are indexed by $q \in L$ with $|q| \leq \lfloor \frac{l}{2} \rfloor$ is uniquely determined by this vector. Moreover for any $\tilde{m}^{[u,v]}$, $|v| \leq \lfloor \frac{l}{2} \rfloor$, where we represent $\bar{\bar{x}}$ as $\sum_{r \in L} \alpha_r \bar{\bar{z}}^r$, and where double tilde represents projection on the coordinates $q \in L : |q| \leq \lfloor \frac{l}{2} \rfloor$, we have (using the notation $X^t(x)$ to mean the matrix with rows and columns corresponding to $q \in L : |q| \leq t$ determined by the vector x)

$$X^{\lfloor \frac{l}{2} \rfloor}(X^l(\hat{x})k) \succeq 0 \Rightarrow \tag{2.115}$$

$$0 \leq \left(\tilde{\tilde{m}}^{[u,v]} \right)^T \left(X^{\lfloor \frac{l}{2} \rfloor}(X^l(\hat{x})\tilde{k}) \right) \tilde{\tilde{m}}^{[u,v]} = \tag{2.116}$$

$$\left(\tilde{\tilde{m}}^{[u,v]} \right)^T \left(\sum_{r \in L} \left(\alpha_r (\tilde{z}^r)^T \tilde{k} \right) \tilde{\tilde{z}}^r (\tilde{\tilde{z}}^r)^T \right) \tilde{\tilde{m}}^{[u,v]} = \tag{2.117}$$

$$\sum_{r \in L : u = r \cap v} \alpha_r (\tilde{z}^r)^T \tilde{k} = \tilde{k}^T X^l(\hat{x}) \tilde{m}^{[u,v]} \tag{2.118}$$

Thus the semidefinite constraints, $X^{\lfloor \frac{l}{2} \rfloor}(X^l(\hat{x})k) \succeq 0$, dominate the linear constraints $\tilde{k}^T X^l(\hat{x})\tilde{m}^{[u,v]} \geq 0$, where $|v| \leq \lfloor \frac{l}{2} \rfloor$.

This gives the basic idea of Lasserre's algorithm, but it is possible to be somewhat more efficient in the number of variables we need to append in order to define semidefinite constraints that can replace the linear constraints. The details are as follows. If $\tilde{z}^r$ is the projection of z^r on the coordinates $q \in L : |q| \leq g$, for some $g \geq 0$, and $k^T \tilde{z}^r \geq 0$ is valid for the lifted K, then so long as we lift $\bar{\bar{x}}$ to have coordinates for each $q \in L : |q| \leq t$ for some $t \geq g + 2$, then we will be able to define the rectangular matrix $\hat{X}(\hat{x})$ with columns for each $q \in L : |q| \leq g$, and rows for each $q \in L : |q| \leq t - g$. The vector $\hat{X}(\hat{x})k$, with a coordinate for each $q \in L : |q| \leq t - g$, is therefore defined, and since $t - g \geq 2$, it will imply a square matrix with rows and columns for each $q \in L : |q| \leq \lfloor \frac{t-g}{2} \rfloor$, and we can demand that this matrix be positive semidefinite (in addition to demanding that the square matrix implied by $\hat{x}$ is positive semidefinite). Formally we have the following definition.

Definition 2.19 *Let K^e be the lifting of the set $P \subseteq \{0,1\}^n$ to $\{0,1\}^{|L|}$ defined in Definition 1.36, and assume $k_1, \ldots, k_m \in R^L$ are such that $K^e \cup \{0\}$ is the set of integer solutions for the system of constraints $k_i^T x \geq 0$, where each k_i is nonzero only in coordinates $q \in L : |q| \leq g_i$. Assume $2l \geq \max_{i=1}^m g_i$. Let the lifting $\hat{x}$ of $\bar{\bar{x}}$ be formed by appending coordinates for each $q \in l : |q| \leq 2l + 2$. For each $i \in \{1, \ldots, m\}$, define $\hat{X}^i(\hat{x})$ to be the matrix with entries fixed by $\hat{x}$, with columns for each $q \in L : |q| \leq g_i$, and rows for each $q \in L : |q| \leq 2l + 2 - g_i$, and define*

$$X^{l+1-\lceil \frac{g_i}{2} \rceil}(\hat{X}^i(\hat{x})k_i) \tag{2.119}$$

to be the (largest) square matrix we can generate from the vector $\hat{X}^i(\hat{x})k_i$. Then where $2l \geq \max_{i=1}^m g_i$, the Lasserre operator (as per [Lau01]) at level l is defined by

$$La^l(\{k_1, \ldots, k_m\}) = \{\bar{\bar{x}} \in R^{n+1} : \exists \hat{x} \text{ such that } X^{l+1}(\hat{x}) \succeq 0,$$

$$X^{l+1-\lceil \frac{g_i}{2} \rceil}(\hat{X}^i(\hat{x})k_i) \succeq 0, \; i = 1, \ldots, m\} \tag{2.120}$$

In particular, where $g_i = 1$ then the square matrix implied by $\hat{X}^i(\hat{x})k_i$ is $X^l(\hat{X}^i(\hat{x})k_i)$ which has a row and a column for each $q \in L$, $|q| \leq l$. In the same manner as we saw above, we will now show that constraining this matrix to be positive semidefinite will imply that for every $v \in L : |v| \leq l$, we will have $k_i^T \bar{X}^l \tilde{m}^{[u,v]} \geq 0$. Note first that had we lifted $\hat{x}$ by further appending coordinates for each $q \in L : |q| \leq 4l + 2$ to obtain the vector $\dot{x}$, then the matrix $\hat{X}^i(\hat{x})$ will be the submatrix of the matrix $X^{2l+1}(\dot{x})$ defined by its columns $q : |q| \leq 1$. Thus where $\dot{x}$ is represented as $\sum_{r \in L} \alpha_r \dot{z}^r$, and where we denote the projection of each vector z^r on the coordinates $q : |q| \leq 2l+1$ as $\ddot{z}^r$, and we let $\ddot{k}_i$ be the lifting of k_i obtained by appending coordinates for each $q : 2 \leq |q| \leq 2l+1$ all of value zero, then

$$\hat{X}^i(\hat{x})k_i = X^{2l+1}(\dot{x})\ddot{k}_i = \sum_{r\in L} \alpha_r \ddot{z}^r (\ddot{z}^r)^T \ddot{k}_i = \sum_{r\in L} \left(\alpha_r (\ddot{z}^r)^T \ddot{k}_i\right) \ddot{z}^r = \sum_{r\in L} \left(\alpha_r (\bar{\bar{z}}^r)^T k_i\right) \ddot{z}^r \tag{2.121}$$

(where double bar indicates projection on empty set and singleton coordinates), and therefore where tilde denotes projection on coordinates $q : |q| \leq l$,

$$X^l(\hat{X}^i(\hat{x})k_i) = \sum_{r\in L} \left(\alpha_r (\bar{\bar{z}}^r)^T k_i\right) \tilde{z}^r (\tilde{z}^r)^T \tag{2.122}$$

so that positive semidefiniteness implies that for each $v : |v| \leq l$,

$$0 \leq \left(\tilde{m}^{[u,v]}\right)^T \left(X^l(\hat{X}^i(\hat{x})k_i)\right) \tilde{m}^{[u,v]} = \left(\tilde{m}^{[u,v]}\right)^T \left(\sum_{r\in L} \left(\alpha_r (\bar{\bar{z}}^r)^T k_i\right) \tilde{z}^r (\tilde{z}^r)^T\right) \tilde{m}^{[u,v]} = \tag{2.123}$$

$$\sum_{r\in L : u = r\cap v} \alpha_r (\bar{\bar{z}}^r)^T k_i = k_i^T \left(\bar{X}^l \tilde{m}^{[u,v]}\right). \tag{2.124}$$

(The proof for the general case $g_i \geq 1$ is similar.) This now proves the following theorem.

Theorem 2.20 *As usual, let $K \subseteq \{y \in \{0,1\}^{n+1} : y_0 = 1\}$ and let $\bar{K}$ be a cone contained in Cone($\{y \in \{0,1\}^{n+1} : y_0 = 1\}$) such that $\bar{K} \cap \{0,1\}^{n+1} = K \cup \{0\}$, and let $\bar{K}^*$ be the polar cone of $\bar{K}$. Then the Lasserre operator*

$$La^l(\bar{K}) = \{\bar{\bar{x}} \in R^{n+1} : \exists \hat{x} \text{ s.t. } X^{l+1}(\hat{x}) \succeq 0,\ X^l(\hat{X}(\hat{x})k) \succeq 0,\ \forall k \in \bar{K}^*\} \tag{2.125}$$

refines $\bar{N}^*(\bar{K})$. □

Note that where $k_1, \ldots, k_m$ generate $\bar{K}^*$, then, by (2.122), the latter condition is equivalent to the condition $X^l(\hat{X}(\hat{x})k_i) \succeq 0,\ i = 1, \ldots, m$.

This completes the survey and reinterpretation of the existing operators.

2.2 The Idempotents of $\vee$

Observe that the vectors $\tilde{m}^{[p,q]}$ are not the only ones in general that satisfy

$$m^T \tilde{z}^r \in \{0,1\},\ \forall r \in L. \tag{2.126}$$

Say for example that $n = 3$ and that we have appended coordinates corresponding to each pair, and consider the vector m

	$\emptyset$	1	2	3	1,2	1,3	2,3
m	0	0	0	1	1	−1	−1

This vector is not of the form $\tilde{m}^{[p,q]}$, $|q| \leq 2$, nor does it belong to their cone, but nevertheless one can easily see by inspection that it satisfies

$$m^T \tilde{z}^r \in \{0,1\}, \ \forall r \in L \tag{2.127}$$

(where the tilde indicates projection on empty set, singletons and pairs coordinates).

The following lemma gives a characterization of the vectors with this property.

Lemma 2.21 *Let the operator $\bigvee$ be as defined in Definition 1.29. A vector $m \in R^L$ satisfies*

$$m^T z^r \in \{0,1\}, \ \forall r \in L \text{ iff } m \bigvee m = m. \tag{2.128}$$

Let $G \subseteq L$. Those vectors m satisfying (2.128) that have zeroes in all but their G coordinates constitute the set of vectors that satisfy

$$\hat{m}^T \hat{z}^r \in \{0,1\}, \ \forall r \in L \tag{2.129}$$

where the hat indicates projection on the G coordinates.

Proof:

$$m \bigvee m = m \text{ iff } m^T W^{z^r} m = m^T z^r, \ \forall r \in L \text{ iff} \tag{2.130}$$

$$m^T z^r (z^r)^T m = m^T z^r, \ \forall r \in L \text{ iff} \tag{2.131}$$

$$m^T z^r \in \{0,1\}, \ \forall r \in L. \tag{2.132}$$

Furthermore, a vector $\hat{m}$ satisfies (2.129) iff the vector $m \in R^L$ obtained by padding $\hat{m}$ with zeroes in the $L - G$ locations satisfies $m^T z^r \in \{0,1\}, \ \forall r \in L$. □

Notice that given an expanded vector $\bar{x}$ with coordinates corresponding to the empty set, the singletons and the pairs, the positive semidefiniteness condition $W^{\bar{x}} \succeq 0$ is equivalent to the infinite set of constraints

$$\left(\bar{\bar{a}} \bigvee \bar{\bar{a}}\right)^T \bar{x} \geq 0 \tag{2.133}$$

for every $\bar{\bar{a}} \in R^{1+n}$ (see Lemma 1.28).

Lemma 2.22 *Let $J \subseteq L$ be the collection of lattice elements $q : |q| \leq 2$ (i.e. the empty set, the singletons and the pairs) The expanded vector $\bar{x}$ with coordinates corresponding to J can be represented as*

$$\bar{x} = \sum_{r \in L} \alpha_r \bar{z}^r, \ \alpha \geq 0 \tag{2.134}$$

iff for all vectors $a \in R^L$ such that $a \bigvee a$ is zero in every non-J coordinate ($a \bigvee a \in Sp^J$ in the terminology of Definition 1.45), we have

$$\left(\left(a \bigvee a\right)|_J\right)^T \bar{x} \geq 0. \tag{2.135}$$

Proof: The set of $\bar{x}$ that can be represented as $\bar{x} = \sum_{r \in L} \alpha_r \bar{z}^r, \ \alpha \geq 0$ is the set $H|_J$ (where H is as in Definition 1.31). By Lemma 1.56, the polar cone

$$(H|_J)^* = (H^*)_J. \tag{2.136}$$

Since H^* is generated by the vectors $a \bigvee a, \ a \in R^L$ (Lemma 1.31), this gives the lemma. □

Thus positive semidefiniteness of $W^{\bar{x}}$ is the relaxation of the condition of the lemma to only consider vectors $a \in Sp^I$, where $I \subseteq L$ is the collection of lattice elements made up of the empty set and the singletons alone (which guarantees $a \bigvee a \in Sp^J$ by Lemma 1.30). So positive semidefiniteness could be strengthened if we were to test this condition on more vectors $a \bigvee a, \ a \notin Sp^I$. In particular the idempotents

$$a \in Sp^J : a \bigvee a = a \tag{2.137}$$

satisfy $a \bigvee a \in Sp^J$ and therefore qualify. Thus positive semidefiniteness would be strengthened by insisting that for all such idempotents

$$\bar{a}^T \bar{x} \geq 0 \tag{2.138}$$

where the bar indicates projection on the J coordinates.

Example: Consider the idempotent m mentioned above at the beginning of the section, and the expanded vector $\bar{x}$

	$\emptyset$	1	2	3	1, 2	1, 3	2, 3
$\bar{x}$	1	5/6	1/3	3/4	1/6	3/4	1/4
m	0	0	0	1	1	−1	−1

The matrix $W^{\bar{x}} =$

	$\emptyset$	1	2	3
$\emptyset$	1	5/6	1/3	3/4
1	5/6	5/6	1/6	3/4
2	1/3	1/6	1/3	1/4
3	3/4	3/4	1/4	3/4

is positive semidefinite, but $m^T x < 0$. $\square$

Thus this test strengthens the positive semidefiniteness condition and is not specific to any particular K. Notably this test can also be performed without appending any new coordinates to $\bar{x}$. Obviously these vectors could also be used to multiply $\bar{X}^l$ in the same manner as $m^{[p,q]}$, and we could then check the product for membership in $\bar{K}$ as we did with $m^{[p,q]}$.

We will show now that there is a much more fundamental way to characterize these vectors.

Recall that by Theorem 1.35, if $x \in R^L$ belongs to the cone H then there exists a measure $\mathcal{X}$ on a measure space $(\Omega, \mathcal{W})$, and sets $A_i \in \mathcal{W},\ i = 1, \ldots, n$ such that for every $r \in L$,

$$\mathcal{X}\left(\bigcap_{i:s_i \in r} A_i\right) = x_r. \tag{2.139}$$

Lemma 2.23 *Let $x \in R^L$, and suppose that there exists a measure $\mathcal{X}$ on a measure space $(\Omega, \mathcal{W})$, and sets $A_i \in \mathcal{W},\ i = 1, \ldots, n$ such that for every $r \in L$,*

$$\mathcal{X}\left(\bigcap_{i:s_i \in r} A_i\right) = x_r. \tag{2.140}$$

Then for any vector of the form $m^{[p,q]}$ we have

$$x^T m^{[p,q]} = \mathcal{X}\left(\bigcap_{i:s_i \in p} A_i \cap \bigcap_{j:s_j \in q-p} A_j^c\right). \tag{2.141}$$

Proof: By elementary measure theory,

$$\mathcal{X}\left(\bigcap_{i:s_i \in p} A_i \cap \bigcap_{j:s_j \in q-p} A_j^c\right) = \tag{2.142}$$

$$\mathcal{X}\left(\bigcap_{i:s_i\in p} A_i \cap \left(\bigcup_{j:s_j\in q-p} A_j\right)^c\right) = \tag{2.143}$$

$$\mathcal{X}\left(\bigcap_{i:s_i\in p} A_i\right) - \mathcal{X}\left(\bigcap_{i:s_i\in p} A_i \cap \left(\bigcup_{j:s_j\in q-p} A_j\right)\right) = \tag{2.144}$$

$$\mathcal{X}\left(\bigcap_{i:s_i\in p} A_i\right) - \mathcal{X}\left(\bigcup_{j:s_j\in q-p}\left(\bigcap_{i:s_i\in p} A_i \cap A_j\right)\right) = \tag{2.145}$$

$$\mathcal{X}\left(\bigcap_{i:s_i\in p} A_i\right) - \sum_{j:s_j\in q-p}\mathcal{X}\left(\left(\bigcap_{i:s_i\in p} A_i\right)\cap A_j\right) + \tag{2.146}$$

$$\sum_{j_1,j_2:s_{j_1},s_{j_2}\in q-p}\mathcal{X}\left(\left(\bigcap_{i:s_i\in p} A_i\right)\cap A_{j_1}\cap A_{j_2}\right) - \cdots + \cdots \tag{2.147}$$

$$+(-1)^k \sum_{j_1,\ldots,j_k:s_{j_m}\in q-p,\ m=1,\ldots,k}\mathcal{X}\left(\bigcap_{i:s_i\in p} A_i \cap \bigcap_{m=1,\ldots,k} A_{j_m}\right) - \cdots + \cdots \tag{2.148}$$

$$\mathcal{X}\left(\bigcap_{i:s_i\in p} A_i \cap \bigcap_{j:s_j\in q-p} A_j\right) = \tag{2.149}$$

$$x_p - \sum_{j:s_j\in q-p} x_{p\cup\{s_j\}} + \sum_{j_1,j_2:s_{j_1},s_{j_2}\in q-p} x_{p\cup\{s_{j_1}\}\cup\{s_{j_2}\}} - \cdots + \cdots \tag{2.150}$$

$$+(-1)^k \sum_{j_1,\ldots,j_k:s_{j_1},\ldots,s_{j_k}\in q-p} x_{p\cup\{s_{j_1}\}\cup\cdots\cup\{s_{j_k}\}} - \cdots + \cdots x_q = \tag{2.151}$$

$$x^T m^{[p,q]} \ \square \tag{2.152}$$

Corollary 2.24 *Let $x \in R^L$, and suppose that there exists a measure $\mathcal{X}$ on a measure space $(\Omega, \mathcal{W})$, and sets $A_i \in \mathcal{W}$, $i = 1, \ldots, n$ such that for every $r \in L$,*

$$\mathcal{X}\left(\bigcap_{i:s_i\in r} A_i\right) = x_r \tag{2.153}$$

Then $x \in H$.

Proof: Each v'th row, m^v, of the Möbius matrix is itself of the form $m^{[v,s]}$ (where $s = \bigcup_{i=1}^n \{s_i\}$), and thus the lemma implies that $x^T m^v \geq 0$ as any measure is nonnegative by definition. Thus $Mx \geq 0$ which implies that $x \in H$. This proves half of Theorem 1.35. □

Definition 2.25 *A collection of sets that is closed under finite unions, intersections and complements is said to be an algebra. If it is closed under all countable unions as well then it is called a σ-algebra.*

Definition 2.26 *Let Ω be a set. Given any collection of subsets $\{A_1, \ldots, A_n\}$, where $n \in Z_+ \cup \{\infty\}$, the intersection of all (σ-)algebras containing the sets $\{A_1, \ldots, A_n\}$ is said to be the (σ-)algebra $\mathcal{A}$ generated by those sets.*

For more details see Chapter 1 of [F99].

Lemma 2.27 *Assume Ω is a countable set. The σ-algebra $\mathcal{A}$ generated by the subsets $\{A_1, A_2, \ldots\}$ of Ω is the collection of all sets that can be written as countable unions of sets of the form*

$$A^V = \bigcap_{i \in V} A_i \cap \bigcap_{j \in Z - V} A_j^c \tag{2.154}$$

for $V \subseteq Z$ (including the empty set), where Z is the set of positive integers. The sets of the form (2.154) are referred to as the atoms of the algebra $\mathcal{A}$. The σ-algebra generated by the finite collection, $\{A_1, \ldots, A_n\}$, $n < \infty$, is the algebra generated by that collection (and in this case it makes no difference if Ω is countable).

Proof: Sets of the form A^V can be thought of as collections of points that satisfy a particular assignment of membership in the sets $\{A_i\}$, i.e. the points of Ω that belong to a particular A^V are those that belong to exactly those A_i such that $i \in V$ and to no other A_i. Obviously every point in Ω has exactly one such assignment, so each point belongs to exactly one such set, and the sets are disjoint. Thus

$$\Omega = \bigcup_{V : A^V \neq \emptyset} A^V \tag{2.155}$$

and the union is disjoint and countable (since Ω is countable by assumption). Obviously the collection of countable unions of sets A^V is closed under countable unions, and by (2.155) it is also closed under complementations, and therefore under intersections as well. Moreover for each A_i, every $\omega \in A_i$ belongs to some $A^V \subseteq A_i$, so A_i belongs to the collection as well. We conclude that the collection is indeed a σ-algebra containing $\{A_i\}$ and it is clear that it must be a subcollection of every σ-algebra containing $\{A_i\}$. Similar reasoning shows that (regardless of the countability of Ω) the algebra generated by $\{A_1, \ldots, A_n\}$, $n < \infty$ is the collection of all finite unions of sets

$$A^V = \bigcap_{i \in V \subseteq \{1, \ldots, n\}} A_i \cap \bigcap_{j \in \{1, \ldots, n\} - V} A_j^c \tag{2.156}$$

which is the σ-algebra described above, as the collection of sets A^V is finite in this case. $\square$

Theorem 2.28 *For any $x \in H$, and corresponding*

$$(\Omega, \mathcal{W}, \mathcal{X}), \text{ and } \{A_1, \ldots, A_n\} \subseteq \mathcal{W} \tag{2.157}$$

every set in the algebra $\mathcal{A}$ generated by $\{A_1, \ldots, A_n\}$ has $\mathcal{X}$ measure equal to $m^T x$ for some vector $m \in R^L$ satisfying

$$m^T z^r \in \{0,1\}, \ \forall r \in L. \tag{2.158}$$

Conversely, for every vector m that satisfies $m^T z^r \in \{0,1\}, \ \forall r \in L$, we have that $m^T x$ is the $\mathcal{X}$ measure of some set in the algebra $\mathcal{A}$.

Proof: Let $s = \{s_1, s_2, \ldots, s_n\}$. Then by Lemma 2.23 and Theorem 1.35, each atom A^V has $\mathcal{X}$ measure equal to

$$x^T m^{[v,s]} = x^T m^v \tag{2.159}$$

where $v = \bigcup_{i \in V}\{s_i\}$, and m^v is the v'th row of the Möbius matrix. By disjointness of the atoms, and the additivity property of measures and Lemma 2.27, the measure of any set A in the algebra $\mathcal{A}$ generated by $A_1, \ldots, A_n$ is therefore

$$\sum_{v:A^V \subseteq A} x^T m^v = x^T \left(\sum_{v:A^V \subseteq A} m^v \right) \tag{2.160}$$

and conversely any sum of the form (2.160) is the measure of the set that is the union of the atoms that the sum is taken over. Note that there are 2^n possible sets $V \subseteq \{1, \ldots, n\}$ corresponding to the 2^n elements in L, and 2^{2^n} possible sets $A \in \mathcal{A}$, one for each possible collection of distinct subsets $V \subseteq \{1, \ldots, n\}$ (or equivalently, one for each distinct collection of lattice elements). So there is a one to one correspondence between the sets $A \in \mathcal{A}$ and the subsets of L, so that

$$\{ \sum_{v:A^V \subseteq A} m^v : A \in \mathcal{A}\} = \{ \sum_{v \in T \subseteq L} m^v : T \subseteq L\}. \tag{2.161}$$

But by Corollary 1.34, the set of vectors that can be written as

$$\sum_{v \in T \subseteq L} m^v \tag{2.162}$$

is exactly the set of idempotents of $\bigvee$. □

Thus the vectors m that satisfy $m^T z^r \in \{0,1\}, \ \forall r \in L$ are the vectors that describe the measure of sets in the algebra $\mathcal{A}$ in terms of x ($x \in H$). For example, given $x \in H$ and corresponding $(\Omega, \mathcal{W}, \mathcal{X})$, and $\{A_1, \ldots, A_n\} \subseteq \mathcal{W}$, the measure of the set

$$(A_1^c \cap A_2) \cup A_3 \tag{2.163}$$

is, in terms of x,

$$x_{\{s_2\}} - x_{\{s_1,s_2\}} + x_{\{s_3\}} - x_{\{s_2,s_3\}} + x_{\{s_1,s_2,s_3\}} \tag{2.164}$$

so the corresponding m vector (where $\{s_i\}$ location is denoted i) is

	$\emptyset$	1	2	3	1, 2	1, 3	2, 3	1, 2, 3
m	0	0	1	1	-1	0	-1	1

Observe now that though we defined x_q for all $q \in L$ as

$$x_q = \sum_{r \in L : q \subseteq r} \alpha_r \tag{2.165}$$

which is a partial sum of

$$x_\emptyset = \sum_{r \in L} \alpha_r \tag{2.166}$$

such identifications define only a small subset of the collection of possible partial sums of (2.166).

Lemma 2.29 *Given*

$$x \in R^L : x = \sum_{r \in L} \alpha_r z^r \tag{2.167}$$

the collection of partial sums of $x_\emptyset = \sum_{r \in L} \alpha_r$ *is*

$$\{x^T m : m \text{ s.t. } m^T z^r \in \{0,1\},\ \forall r \in L\} \tag{2.168}$$

Proof: For any $q \in L$,

$$x^T m^q = \alpha_q \tag{2.169}$$

so that the collection of all partial sums of $x_\emptyset$ is the collection of all numbers

$$x^T \left(\sum_{q \in T \subseteq L} m^q \right). \ \square \tag{2.170}$$

So the partial sums of $x_\emptyset$ are just the measures of the sets of $\mathcal{A}$ (where $x \in H$), and the vectors m that satisfy $m^T z^r \in \{0,1\},\ \forall r \in L$ are the vectors that describe these measures and partial sums in terms of $x \in R^L$.

Thus the central object of our concern, which is the partial sums, is in one to one correspondence with the algebra $\mathcal{A}$. The vectors m that satisfy $m^T z^r \in \{0,1\},\ \forall r \in L$, which are also in one to one correspondence with $\mathcal{A}$, are what allow us to describe the 2^{2^n} partial sums in terms of vectors in R^L. A more natural approach would thus be to shift our focus from the lattice L to the algebra $\mathcal{A}$. We will do this by adopting a more comprehensive expansion of the vector $\bar{\bar{x}}$, raising its dimension to $O(2^{2^n})$ by introducing variables for every partial sum of x and not only for those corresponding to the lattice elements of L. We will also see that this more general framework can provide a natural way to describe and analyze subsets of $\{0,1\}^n$.

Chapter 3

Algebraic Representation

In this chapter we will broaden the framework developed in the previous chapter by lifting to $O(2^{2^n})$ dimensions. As we indicated in the Preface, the general idea that will govern this lifting will be to append variables to encode every possible "description" of a vector $y \in P \subseteq \{0,1\}^n$. Logical properties of the set P could then find expression as linear relationships between the new variables. Recall our example from the Preface in which P had the logical property that for each $y \in P$, wherever exactly one of the first two coordinates of y has the value 1 then the third coordinate of y also has value 1. This property, stated logically as y_1 XOR $y_2 \Rightarrow y_3$, could then be encoded as $y[y_1 \text{ XOR } y_2] \leq y_3$, where $y[y_1 \text{ XOR } y_2]$ is a new $0,1$ variable encoding the "state" that y is such that exactly one of its first two coordinates has value 1.

In the first section, drawing on the equivalence between logical expressions and set theoretic expressions, we will implement this general idea in the form of a lifting that, given $P \subseteq \{0,1\}^n$, assigns a variable to each subset of P. Each "description" of a vector $y \in P$ will be thought of as the set of points of P for which that description holds, and will be assigned a variable. In particular, each of the original variables y_i, $i = 1, \ldots, n$, which "describes" whether or not the i'th coordinate of y has value 1, will be thought of as the variable $y[\{y \in P : y_i = 1\}]$. The lifted vectors can thus be thought of set functions on the algebra of subsets of P, with the original vector $(y_1, \ldots, y_n)$ as the n function values $y[\{y \in P : y_i = 1\}]$, $i = 1, \ldots, n$.

In the second section we will establish the connection with measure theory. In particular we will show that, given $P \subseteq \{0,1\}^n$, a vector $x \in R^n$ belongs to $Conv(P)$ if and only if that vector can be lifted to a set function that is a probability measure on the algebra of subsets of P, i.e. if and only if there exists a probability measure χ on the algebra of subsets of P, such that for each $i = 1, \ldots, n$, $\chi(\{y \in P : y_i = 1\}) = x_i$. We will also indicate a way to generalize this result to the case where P is a countably infinite set.

Sections 3.3, 3.4 and 3.5 describe the basic mathematics that govern this lifting and show how to use the lifting to characterize the convex hull of subsets of $\{0,1\}^n$ in a variety of ways. The tools outlined in these sections will be used repeatedly in the later work. The concept of "signed measure consistency", in particular, which refers to the situation where a lifted vector $\bar{x}$ is "consistent with" (i.e. it can be lifted to) an additive (though not necessarily nonnegative) set function (a "signed measure") on the algebra of subsets of P, will be discussed in Section 3.3, and will prove crucial in Chapter 4.

In Section 3.3 we will also describe, based on the measure theoretical characterization of convex hulls of subsets of $\{0,1\}^n$, a "proof by picture" method (essentially implicit in the work of [LS91]), for establishing that a system of inequalities is convex hull defining for its integer hull. We will also see in that section our first example of how an intelligent lifting can be used to replace an exponentially large system of constraints with a polynomially large system of constraints.

The "delta vectors" of Section 3.4 reflect the ways in which the measures of various sets in the algebra can be used to identify the measures of other sets in the algebra. For example, for any measure χ, the measure $\chi(A \cup B) = \chi(A) + \chi(B) - \chi(A \cap B)$, so the measures of the sets A, B and $A \cap B$ determine the measure of the set $A \cup B$, and the relationship between them is captured by the vector $(1, 1, -1)$ where the first coordinate corresponds to the set A, the second to the set B and the third to the set $A \cap B$. It is important to note that this relationship is independent of the specific choice of measure; it reflects a relationship between the sets themselves. Given a collection of sets $\mathcal{Q}$ from the algebra of subsets of P, the delta vectors $\mu^{\mathcal{Q}}(q)$, where q is a set in the algebra, are the vectors that (in the same manner as the vector $(1, 1, -1)$ of our example) describe how to obtain the measure of the set q in terms of the measures of the sets of $\mathcal{Q}$. These vectors can be thought of as describing the measure theoretical relationship between q and the sets of $\mathcal{Q}$. We will see that the delta vectors represent a considerable generalization of the $m^{[p,q]}$ vectors (Definition 1.50) of the previous two chapters. The $m^{[p,q]}$ vectors, which are essential to all of the algorithms of the first two chapters, are in fact the delta vectors for a particular collection of sets $\mathcal{Q}$ from the algebra of subsets of $\{0,1\}^n$, and a particular choice of sets q in that algebra.

Section 3.5 presents a generalization of the delta vectors which will be useful in the generalization of the Lasserre algorithm that will be described in Chapter 4.

In Section 3.6 we will discuss "measure preserving operators". Let $P \subseteq \{0,1\}^n$, and let T be a function that maps measures on the algebra of subsets of P (or vectors that are consistent with measures, i.e. vectors that can be lifted to measures) into measures on the algebra of subsets of P (or vectors that are consistent with measures). Then considering the equivalence between membership of a vector x in the convex hull of P, and consistency with a measure on the algebra of subsets of P, we can validly constrain liftings $\bar{x}$ of x by demanding that $T(\bar{x})$ also be consistent with a measure on the algebra of subsets of P. Thus all constraints that can be applied to $\bar{x}$ may also be applied to $T(\bar{x})$. We will see that this idea can be seen to underlie both the concept of partial summation developed in Chapter 2 as well as the methodology of the Lasserre algorithm. We will indicate in Section 3.6 and in Section 4.2 of the next chapter generalizations of these two specifically, as well as other directions that measure preserving operators may take. A full study of the effectiveness of measure preserving operators, however, remains an object for future research.

The focus of the chapter, however, is not on the generalizations of the algorithms of the previous chapters per se. It is on the development of a broader framework that can be applied in the form of completely different algorithms.

3.1 Fundamentals

3.1.1 The Algebra $\mathcal{P}$

Consider $(y_1, \ldots, y_n) \in P \subseteq \{0,1\}^n$. We would like to find a way to encode everything that can be said about the variables $y_1, \ldots, y_n$, in the form of new variables. But first we have to quantify the notion of a "statement that can be made about the vector y", or equivalently, a "state" that the vector y can have.

Note that a variable y_i can be thought of as a boolean function representing the "state" $y_i = 1$. If y_i is indeed 1 then the variable has value *true*, represented as 1, and if it is 0 then it has value *false*, represented as 0. Thus in a similar manner we might introduce a $0,1$ variable $y_{i,j}$ representing the "state" "y_i and y_j are both 1". Thus $y_{i,j}$ would be a boolean function on y_i and y_j having the value

$$y_i \text{ AND } y_j \tag{3.1}$$

where again *true* is represented by 1 and *false* is represented by 0. Similarly we might introduce $y_{i \text{ or } j}$ representing the "state" "y_i or y_j are 1". Thus $y_{i \text{ or } j}$ would be a boolean function on y_i and y_j having the value

$$y_i \text{ OR } y_j. \tag{3.2}$$

Thus where the vectors y that we are considering belong to some $P \subseteq \{0,1\}^n$, the broadest definition of a "state" or of a "statement that can be made about y" is as some condition that holds on some subset of the possible vectors $y \in P \subseteq \{0,1\}^n$. In principle there are therefore "states" and "statements" corresponding to every subset of P, or equivalently, to every boolean function on P (as for any subset of P we can define a boolean function that is *true* exactly on that subset). The paradigm that we will be using is therefore to introduce a new variable corresponding to every boolean function on y, or equivalently, to every subset of P. Given $y \in P \subseteq \{0,1\}^n$, for each subset of P there will be a variable corresponding to the boolean function that holds *true* on exactly that subset, with that variable having a value of 1 iff y belongs to that subset.

Observe that the subset of P on which the boolean function $y_i = 1$ holds true is (naturally enough) the set

$$\{y \in P \subseteq \{0,1\}^n : y_i = 1\}. \tag{3.3}$$

Definition 3.1 *Denote the sets*

$$\{y \in P \subseteq \{0,1\}^n : y_i = 1\} \tag{3.4}$$

by the name Y_i^P. *Denote* $Y_i^{\{0,1\}^n}$ *as just* Y_i.

Definition 3.2 *Denote the subset algebra of* $P \subseteq \{0,1\}^n$ *as* $\mathcal{P}$. *Denote the subset algebra of* $\{0,1\}^n$ *as* $\mathcal{A}$.

Definition 3.3 *Two sets* U *and* V *contained in* $\{0,1\}^n$ *will be said to be* P*-equivalent if*

$$U \cap P = V \cap P. \tag{3.5}$$

The most convenient way to interpret the new variables that we are appending is as corresponding to the subsets of P, but they can also be understood as corresponding to the logical statements that can be made about vectors in P, or as corresponding to the set theoretic expressions that involve $Y_1^P, \ldots, Y_n^P$, as we will show now.

3.1.2 Logical Representation

Definition 3.4 *A logical expression*

$$\theta(y_1, \ldots, y_n) \tag{3.6}$$

is an expression entailing the variables $y_1, \ldots, y_n$, *and the operators AND, OR and NOT, such that the expression defines a boolean function on* $\{0,1\}^n$. *For example,*

$$\theta(y_1, y_2, y_3) = NOT(y_1\ OR\ NOT(y_2))\ AND\ y_3 \tag{3.7}$$

is a logical expression.

A restricted logical expression

$$\theta^P(y_1, \ldots, y_n), \tag{3.8}$$

defined to be the logical expression $\theta(y_1, \ldots, y_n)$ *with the values* $\{y_1, \ldots, y_n\}$ *restricted to belong to* $P \subseteq \{0,1\}^n$*, will be referred to as a "P-logical expression". Similarly, a set theoretic expression*

$$\Theta(Y_1, \ldots, Y_n) \tag{3.9}$$

is defined to be an expression entailing the sets $Y_1, \ldots, Y_n$*, unions, intersections and complements with respect to* $\{0,1\}^n$ *that defines a set in* $\{0,1\}^n$*. The "P-set theoretic expression"*

$$\Theta^P(Y_1^P, \ldots, Y_n^P) \tag{3.10}$$

is defined to be the expression Θ *applied to the sets* $\{Y_i^P\}$*, but with complementation taken with respect to* P*. Note that* $\Theta^P(Y_1^P, \ldots, Y_n^P)$ *is a set contained in* P*.*

Remark 3.5 *Exchanging* y_i *with* Y_i^P*,* $i = 1, \ldots, n$*, AND with* $\cap$*, OR with* $\cup$ *and NOT with complementation with respect to* P *yields a one to one correspondence between* P*-logical expressions and* P*-set theoretic expressions. Moreover, the subset of* P *for which a given* P*-logical expression holds true is exactly the set defined by the corresponding* P*-set theoretic expression, and conversely.*

Proof: The first part of the statement is clear. As for the second part,

$$\{y \in P : y_i = 1\} = Y_i^P,\ \{y \in P : (y_i \text{ AND } y_j) = 1\} = Y_i^P \cap Y_j^P, \tag{3.11}$$

$$\{y \in P : (y_i \text{ OR } y_j) = 1\} = Y_i^P \cup Y_j^P,\ \{y \in P : \text{NOT}(y_i) = 1\} = P - Y_i^P \tag{3.12}$$

and the statement follows from induction. □

Corollary 3.6 *For all set theoretic expressions,* $\Theta(Y)$*,*

$$\Theta^P(Y_1^P, \ldots, Y_n^P) = \Theta(Y_1, \ldots, Y_n) \cap P \tag{3.13}$$

Proof: Where we denote the set theoretic expression that corresponds to $\theta(y)$ by $\Theta(Y)$,

$$\{y \in P : \theta(y) = 1\} = \{y \in \{0,1\}^n : \theta(y) = 1\} \cap P \tag{3.14}$$

by definition, and by the remark,

$$\{y \in P : \theta(y) = 1\} = \Theta^P(Y_1^P, \ldots, Y_n^P). \ \square \tag{3.15}$$

Remark 3.7 *The algebra generated by $Y_1^P, \ldots, Y_n^P$ (where P is treated as the universal set) is $\mathcal{P}$.*

Proof: The atoms of the algebra are the sets of the form

$$\bigcap_{i \in V \subseteq \{1,\ldots,n\}} Y_i^P \cap \bigcap_{j \in \{1,\ldots,n\}-V} (Y_j^P)^c. \tag{3.16}$$

But each such set is exactly the intersection of P with the single point with 1's in its V coordinates and 0's in its other coordinates. So each such set is either empty, or is composed of the single point with 1's in exactly its V coordinates. Since there is such a set for every V, it follows that the atoms are exactly the points of P (and the empty set if $P \neq \{0,1\}^n$), so the collection of all unions of atoms is the collection of all subsets of P. □

These two remarks imply the following remark.

Remark 3.8 *Every boolean function on P can be represented by a P-logical expression (not uniquely).*

Proof: We need to show that for every subset A of P there exists a P-logical expression that holds true exactly on A. As in the proof of Remark 3.7, each subset $A \subseteq P$ can be written as the union of the atoms corresponding to the points of A. Thus A can be defined by a P-set theoretic expression entailing $Y_1^P, \ldots, Y_n^P$, and by Remark 3.5, the subset of P on which the corresponding logical expression holds true is exactly A. □

Thus for every subset $A \subseteq P$ there exists a P-logical expression $\theta_A^P(y_1, \ldots, y_n)$ that (where y is restricted to belong to P,) holds true exactly for the points $y \in A$. Equivalently, there exists a P-set theoretic expression

$$\Theta_A^P(Y_1^P, \ldots, Y_n^P) = A. \tag{3.17}$$

In general, for every subset $Q \subseteq \{0,1\}^n$, letting $P = \{0,1\}^n$ we see that there exists a set-theoretic expression,

$$\Theta_Q(Y_1, \ldots, Y_n) = Q \tag{3.18}$$

and by Remark 3.5

$$\Theta_Q^P(Y_1^P, \ldots, Y_n^P) = Q \cap P. \tag{3.19}$$

Definition 3.9 *Two logical expressions $\theta(y)$ and $\phi(y)$ are said to be equivalent if*

$$\theta(y) = \phi(y) \ \forall y \in \{0,1\}^n \tag{3.20}$$

i.e. if the subset of $\{0,1\}^n$ *on which one holds true and the subset of* $\{0,1\}^n$ *on which the other holds true are the same. The logical expressions* θ *and* ϕ *will be said to be* P*-equivalent if*

$$\theta^P(y) = \phi^P(y),\ \forall y \in P \subseteq \{0,1\}^n \tag{3.21}$$

i.e. if

$$\{y \in \{0,1\}^n : \theta(y) = 1\} \cap P = \{y \in \{0,1\}^n : \phi(y) = 1\} \cap P \tag{3.22}$$

Similarly, two set theoretic expressions $\Theta(Y)$ *and* $\Phi(Y)$ *(where* Y *represents an* n*-tuple of sets) will be said to be equivalent if*

$$\Theta(Y_1, \ldots, Y_n) = \Phi(Y_1, \ldots, Y_n) \tag{3.23}$$

and they will be said to be P*-equivalent if*

$$\Theta^P(Y_1^P, \ldots, Y_n^P) = \Phi^P(Y_1^P, \ldots, Y_n^P). \tag{3.24}$$

Remark 3.10 *Two logical expressions* $\theta(y_1, \ldots, y_n)$ *and* $\phi(y_1, \ldots, y_n)$ *are equivalent if and only if the corresponding set theoretic descriptions* $\Theta(Y_1, \ldots, Y_n)$ *and* $\Phi(Y_1, \ldots, Y_n)$ *define the same set. Two logical expressions* $\theta(y_1, \ldots, y_n)$ *and* $\phi(y_1, \ldots, y_n)$ *are* P*-equivalent if and only if the corresponding set theoretic descriptions* $\Theta^P(Y_1^P, \ldots, Y_n^P)$ *and* $\Phi^P(Y_1^P, \ldots, Y_n^P)$ *define the same set, and this happens if and only if*

$$\Theta(Y_1, \ldots, Y_n) \cap P = \Phi(Y_1, \ldots, Y_n) \cap P. \ \square \tag{3.25}$$

Thus two nonequivalent logical expressions may define the same boolean function (and the same subset of P) so long as they are P-equivalent. P-equivalence of two nonequivalent logical expressions means that if we restrict our attention to the vectors $y \in P$ then these two expressions describe the same set. Equivalently, the corresponding set theoretic descriptions in terms of $\{Y_i^P\}$ (with P as the universal set) define the same set. For example if all points in P that have a 1 in their y_1 coordinate also have a 1 in either their y_2 or y_3 coordinate, then

$$Y_1^P = Y_1^P \cap (Y_2^P \cup Y_3^P). \tag{3.26}$$

Thus all of the logical structure of P, i.e. the logical equivalences that are specific to the vectors of P, is captured by the nonequivalent but P-equivalent expressions.

Given $P \subseteq \{0,1\}^n$, we will thus append variables to the vectors $y \in P$ corresponding to every P-equivalence class of logical expressions (or alternatively, of set theoretic expressions). These variables will be assigned a value of 1 where the logical expression holds for that point y (where y belongs to the set defined by the corresponding set theoretic expression), and 0 otherwise. In principle we could append variables for every $\{0,1\}^n$-equivalence class (i.e. every class of equivalent expressions), but for any point in P, P-equivalent expressions are all true or false together (they define the same set), so the variables assigned to P-equivalent expressions would be assigned all the same value anyway.

Another important point to observe is that while it is convenient and intuitive to think of the algebra $\mathcal{P}$ as the subset algebra of P, it can be viewed in a broader way as well. The entire structure of $\mathcal{P}$ is determined by its atoms, and the only distinguishing characteristic of the atoms of $\mathcal{P}$ in particular is the fact that a particular subset of them is empty. But there is no significance to the fact that the nonempty atoms are comprised of exactly one point, and that that point belongs to $\{0,1\}^n$.

Lemma 3.11 *Let Π be the algebra generated by sets $W_1, \ldots, W_n$ contained in some Ω, and let the atoms*

$$\bigcap_{i \in V \subseteq \{1,\ldots,n\}} W_i \cap \bigcap_{j \in \{1,\ldots,n\}-V} W_j^c \tag{3.27}$$

be empty iff the point $y \in \{0,1\}^n$ with ones in exactly its V coordinates does not belong to P. Then $\mathcal{P}$ is isomorphic to Π.

Proof: Every set in Π can be represented as a union of nonempty atoms of Π, and as the atoms are all disjoint, this representation is unique. Similarly every set in $\mathcal{P}$ can be represented uniquely as a union of nonempty atoms of $\mathcal{P}$. Let T be the collection of subsets of $\{1,\ldots,n\}$ that satisfy that the point with 1's in exactly those coordinates belongs to P. Then any $A \in \mathcal{P}$ can be written uniquely as

$$A = \bigcup_{V \subseteq \{1,\ldots,n\}: V \in \tau} (Y^P)^V \tag{3.28}$$

for some $\tau \subseteq T$, where $(Y^P)^V$ are atoms of $\mathcal{P}$. Similarly any $W \in \Pi$ can be written uniquely as

$$W = \bigcup_{V \subseteq \{1,\ldots,n\}: V \in \tau'} W^V \tag{3.29}$$

where W^V are atoms of $\mathcal{P}$, for some $\tau' \subseteq T$. Consider the function

$$f : \mathcal{P} \to \Pi \tag{3.30}$$

defined by

$$f(A) = f\left(\bigcup_{V \subseteq \{1,\ldots,n\}: V \in \tau} (Y^P)^V \right) = \bigcup_{V \subseteq \{1,\ldots,n\}: V \in \tau} W^V. \tag{3.31}$$

It is clear that this function is a one to one correspondence and that $f(A \cup B) = f(A) \cup f(B)$. Moreover,

$$f(A^c) = f\left(\bigcup_{V \subseteq \{1,\ldots,n\}: V \in T - \tau} (Y^P)^V \right) = \bigcup_{V \subseteq \{1,\ldots,n\}: V \in T - \tau} W^V = (f(A))^c. \tag{3.32}$$

Thus $f(A \cap B) = f(A) \cap f(B)$ as well, and f is an isomorphism. □

3.2 Zeta Vectors for $\mathcal{P}$

The expanded vectors $\zeta(z)$ of the points $z \in P$, where the coordinates of $\zeta(z)$ are indexed by the sets $p \in \mathcal{P}$, therefore satisfy

$$\zeta(z)_p = \begin{cases} 1 : z \in p \\ 0 : \text{ otherwise} \end{cases} \tag{3.33}$$

Our primary interest will be in vectors of this form, but these vectors can be put in a wider context by noting that they are dual zeta vectors of the algebra $\mathcal{P}$ (or Π) partially ordered by inclusion.

Remark 3.12 *The algebra $\mathcal{P}$ partially ordered by inclusion is a lattice. The dual zeta vectors of this lattice are the vectors $\zeta^q(P)$, where $q \in \mathcal{P}$, that satisfy*

$$(\zeta^q(P))_p = \begin{cases} 1 : q \subseteq p \\ 0 : \textit{ otherwise} \end{cases} \tag{3.34}$$

(These are the zeta vectors of the algebra ordered by reverse inclusion, see [Ro64].) Thus the expanded vectors $\zeta(z)$ are just the dual zeta vectors $\zeta^q(P)$ where q is the set made up of the single point z. □

Stated in terms of the more general framework, the vectors $\zeta(z)$ are the dual zeta vectors $\zeta^q(P)$ where q is the atom that corresponds to the point z, i.e. it is the atom that belongs to each W_i iff $z_i = 1$.

Notation: The dual zeta vectors $\zeta^r(P)$ are functions of P. Nevertheless, to avoid clutter we will denote them simply as ζ^r (and the expression $\zeta^r(\cdot)$ will generally be given a different meaning) so long as it is clear which P they depend on. Also, although technically these vectors are dual zeta vectors for the algebra ordered by inclusion, they are zeta vectors of the algebra ordered by reverse inclusion, and we will refer to them throughout as just the zeta vectors.

The zeta vectors encode all of the inclusion relationships in the algebra. Observe that each inclusion relationship can also be thought of as a logical implication valid for points in P. For example, an inclusion

$$Y_1^P \cap Y_2^P \subseteq (Y_3^P \cup Y_4^P) \tag{3.35}$$

means that for all points $z \in P$,

$$z_1 = 1 \text{ and } z_2 = 1 \Rightarrow z_3 = 1 \text{ or } z_4 = 1. \tag{3.36}$$

The following lemma shows how some set theoretic relationships manifest themselves as numerical relationships between zeta vectors.

Definition 3.13 *Define $\mathcal{S}^P \in \mathcal{P}$ be the collection of sets in $\mathcal{P}$ that contain a single point (in the wider framework, $\mathcal{S}^P$ is the collection of nonempty atoms). For the case $P = \{0,1\}^n$, denote $\mathcal{S}^P$ as $\mathcal{S}$.*

Lemma 3.14 *Let $r \subseteq P$ be any nonempty set in $\mathcal{P}$, then where u and v are sets in $\mathcal{P}$, and recalling that complementation is with respect to the universal set P,*

1. $u \subseteq v \Rightarrow \zeta_u^r \leq \zeta_v^r$ *i.e.* $\zeta_u^r = 1 \Rightarrow \zeta_v^r = 1$
2. $\zeta_u^r = 1 \Rightarrow \zeta_{u^c}^r = 0$
3. $\zeta_{u\cap v}^r = 1$ *iff* $\zeta_u^r = 1$ *and* $\zeta_v^r = 1$

putting these three together yields

4. *if* $u \subseteq v$ *then* $\zeta_u^r = 1 \Rightarrow \zeta_v^r = 1 \Rightarrow \zeta_{v^c}^r = 0$ *and therefore*
5. *if* $u \subseteq v$ *then* $\zeta_u^r = \zeta_{u\cap v}^r + \zeta_{u\cap v^c}^r$

and if $r \in \mathcal{S}^P$ then

6. $\zeta_u^r = 1$ *iff* $\zeta_{u^c}^r = 0$ *and therefore*
7. $\zeta_v^r = \zeta_{u\cap v}^r + \zeta_{u^c\cap v}^r$ *for all u and v in $\mathcal{P}$.*

Note that (2) holds not only for u and u^c but for any any u and w such that $u \cap w = \emptyset$ (so long as $r \neq \emptyset$), but (6) holds in general only for u and u^c The following generalization of (7), however, holds for any mutually exclusive pair u and w (for $r \in \mathcal{S}^P$).

8. $\zeta^r_{u\cup w} = \zeta^r_u + \zeta^r_w$, *and more generally*

9. $\zeta^r_{u\cup v} = \zeta^r_u + \zeta^r_v - \zeta^r_{u\cap v}$.

Proof: The first four statements are trivial from the definitions.

(5): If $\zeta^r_u = 0$ then clearly the right side of the equation is also 0. If $\zeta^r_u = 1$ then by (4), $\zeta^r_{v^c} = 0 \Rightarrow \zeta^r_{u\cap v^c} = 0$ (by (3)), and $\zeta^r_{u\cap v} = \zeta^r_u$ since $u \subseteq v$.

(6), (7), (8): We have already seen that in the case of $r \in \mathcal{S}^P$, ζ^r_u has the value 1 iff the single point that constitutes r belongs to the set u. (In general, where $r \notin \mathcal{S}^P$, if the set r overlaps u but is not contained in u then neither u nor its complement will contain r, but if r contains only one point then this cannot happen.) (6) and (8) are therefore a consequence of the fact that a point belongs to a disjoint union iff it belongs to exactly one of the elements of the union. (7) is a consequence of (3) and (6).
(9): Observe that $u \cup v = u \cup (u^c \cap v)$ and that u and $u^c \cap v$ are mutually exclusive. So by (8),

$$\zeta^r_{u\cup v} = \zeta^r_u + \zeta^r_{u^c\cap v} = \text{ (by (7))} \tag{3.37}$$

$$\zeta^r_u + \zeta^r_v - \zeta^r_{u\cap v}. \ \Box \tag{3.38}$$

The following theorem shows that for $\mathcal{S}^P$, property (8) can be turned around to provide a sufficient condition for a 0, 1 vector in $R^{\mathcal{P}}$ to be a zeta vector for a set $r \in \mathcal{S}^P$. More generally we will show that (8), coupled with a nonnegativity constraint, defines the cone of the zeta vectors of $\mathcal{S}^P$. Before we come to the theorem, however, we will need a claim, but first we need a definition.

Definition 3.15 *Given a vector $\chi \in R^{\mathcal{P}}$, where $r \in \mathcal{P}$, define the vector $\chi^{\mathcal{P}\cap r} \in R^{\mathcal{P}}$ by*

$$\chi^{\mathcal{P}\cap r}_u = \chi_{u\cap r} \tag{3.39}$$

for all $u \in \mathcal{P}$. Vectors $\chi^{\mathcal{P}\cap r}$ will also be referred to simply as χ^r.

Claim 3.16 *If $\chi \in R^{\mathcal{P}}$ satisfies that for all $u, v \in \mathcal{P}$ we have*

$$u \cap v = \emptyset \Rightarrow \chi_{u\cup v} = \chi_u + \chi_v \tag{3.40}$$

then for any $r \in \mathcal{S}^P$ and any $u \in \mathcal{P}$ we have

$$\chi_{r\cap u} = \begin{cases} \chi_r : r \subseteq u \\ 0 : \ \textit{otherwise} \end{cases} \tag{3.41}$$

In other words,

$$\chi^{\mathcal{P}\cap r} = \chi_r \zeta^r. \tag{3.42}$$

Proof: If $r \subseteq u$ then it is clear that $\chi_{r\cap u} = \chi_r$, as $r \cap u = r$. Note now that if a vector χ satisfies (3.40), then letting $u \in \mathcal{P}$, since we always have $u \cap \emptyset = \emptyset$,

$$\chi_u = \chi_{u\cup\emptyset} = \chi_u + \chi_\emptyset \Rightarrow \tag{3.43}$$

$$\chi_\emptyset = 0. \tag{3.44}$$

Thus if $r \not\subseteq u$, then since r is composed of a single point,

$$r \cap u = \emptyset \Rightarrow \chi_{r\cap u} = \chi_\emptyset = 0. \ \square \tag{3.45}$$

Observe also that for arbitrary $u, v \in \mathcal{P}$, (3.40) implies (as in the proof of property (9) above) that

$$\chi_{u\cup v} = \chi_u + \chi_{u^c\cap v} = \chi_u + \chi_v - \chi_{u\cap v}. \tag{3.46}$$

Notation: The set of all linear combinations of a collection $\{v_i\}$ of vectors will be denoted $Span(\{v_i\})$.

Theorem 3.17 *A nonzero vector $\gamma \in \{0,1\}^{\mathcal{P}}$ satisfies $\gamma = \zeta^r$ for some $r \in \mathcal{S}^P$ iff for all disjoint $u, v \in \mathcal{P}$,*

$$\gamma_{u\cup v} = \gamma_u + \gamma_v. \tag{3.47}$$

Moreover, a vector $\chi \in R^{\mathcal{P}}$ satisfies that $\chi \in Span(\{\zeta^r : r \in \mathcal{S}^P\})$ iff for all disjoint $u, v \in \mathcal{P}$,

1. $\chi_{u\cup v} = \chi_u + \chi_v$.

The vector χ belongs to the cone, $Cone(\{\zeta^r : r \in \mathcal{S}^P\})$ iff it satisfies the additional condition,

2. $\chi \geq 0$.

The vector χ belongs to the convex hull, $Conv(\{\zeta^r : r \in \mathcal{S}^P\})$ iff the additional condition

3. $\chi_P = 1$.

holds as well. The vector χ *belongs to the affine hull* $Af(\{\zeta^r : r \in \mathcal{S}^P\})$ *iff Conditions (1) and (3) hold.*

Proof: We will first prove the statement about the linear span and the cone, and then the statements about the convex hull and the affine hull, and then the statement about the $0,1$ vectors.

It follows from Lemma 3.14 that any vector in the span satisfies condition (1), and any vector in the cone clearly satisfies condition (2). As for the converse, note first that for any $u \in \mathcal{P}$,

$$u = \bigcup_{r \in \mathcal{S}^P} (u \cap r) \tag{3.48}$$

where the union is disjoint, since $\mathcal{S}^P$ is the collection of all the single point sets and any set is the disjoint union of the points that it contains. Thus by repeated application of (1),

$$\chi_u = \sum_{r \in \mathcal{S}^P} \chi_{u \cap r} = \sum_{r \in \mathcal{S}^P} \chi_u^{\mathcal{P} \cap r} = \sum_{r \in \mathcal{S}^P} \chi_r \zeta_u^r \tag{3.49}$$

by Claim 3.16, and thus we conclude that

$$\chi = \sum_{r \in \mathcal{S}^P} \chi_r \zeta^r \tag{3.50}$$

and thus χ belongs to the linear span of the zeta vectors of $\mathcal{S}^P$, and if nonnegativity is assumed, then it belongs to their cone as well. As for affineness, every affine combination χ of $\zeta^r, r \in \mathcal{S}^P$ must satisfy $\chi_P = 1$, as $\zeta_P^r = 1, \ \forall r \in \mathcal{S}^P$. Conversely, if χ satisfies condition (1) as well as $\chi_P = 1$, then by the reasoning above,

$$1 = \chi_P = \sum_{r \in \mathcal{S}^P} \chi_{P \cap r} = \sum_{r \in \mathcal{S}^P} \chi_r \tag{3.51}$$

and thus the combination that yields χ is indeed affine, and if χ is also nonnegative then the combination is convex. Moving now to the case of the $0,1$ vectors, notice that if $\chi \geq 0$, then for any $u \in \mathcal{P}$,

$$\chi_P = \chi_{u \cup u^c} = \chi_u + \chi_{u^c} \geq \chi_u \tag{3.52}$$

so that

$$\chi \neq 0 \Rightarrow \chi_P > 0. \tag{3.53}$$

Thus if γ is nonzero and is a $0,1$ vector that satisfies (1), then it must satisfy (2) and (3) as well and therefore must belong to the convex hull

$$Conv\left(\{\zeta^r : r \in \mathcal{S}^P\}\right) \subseteq [0,1]^{\mathcal{P}}. \tag{3.54}$$

But since γ is $0,1$, and no $0,1$ point can be a convex combination of other points in the hypercube, we conclude that γ must itself belong to the set $\{\zeta^r : r \in \mathcal{S}^P\}$. $\Box$

Definition 3.18 *A set function $f : \mathcal{W} \to R^1 \cup \{\infty, -\infty\}$, where $\mathcal{W}$ is an algebra of sets belonging to some universal set Ω, is said to be additive if for all disjoint sets $u, v \in \mathcal{W}$, $f(u \cup v) = f(u) + f(v)$. The function f is said to be σ-additive if $\mathcal{W}$ is a σ-algebra, and for any pairwise disjoint countable union of sets $\{u_i, i \geq 1\} \subseteq \mathcal{W}$ we have $f(\bigcup_{i=1}^{\infty} u_i) = \sum_{i=1}^{\infty} f(u_i)$. Obviously if $\mathcal{W}$ is finite then any additive set function is also σ-additive. A σ-additive set function on $\mathcal{W}$ is also known as a* signed measure *on $\mathcal{W}$. If f is also nonnegative then f is said to be a* measure, *and if, in addition, $f(\Omega) = 1$ then f is said to be a* probability measure.

For formal details, see, for example, Chapter 1 of [F99].

Corollary 3.19 *The vector χ belongs to the span of the zeta vectors of $\mathcal{S}^P$ iff χ, when viewed as a set function on $\mathcal{P}$ (defined by $\chi(u) = \chi_u$, $\forall u \in \mathcal{P}$) is a finite signed measure on $\mathcal{P}$. The vector χ belongs to the cone of the zeta vectors of $\mathcal{S}^P$ iff χ defines a finite measure on $\mathcal{P}$. The vector χ belongs to the convex hull of the zeta vectors of $\mathcal{S}^P$ iff χ defines a probability measure on $\mathcal{P}$.* $\Box$

Observe that the zeta vectors of the lattice L introduced by Lovász and Schrijver (Definition 1.17) are projections of the zeta vectors $\{\zeta^r(\{0,1\}^n) : r \in \mathcal{S}\}$ (recall that $\mathcal{S}^P = \mathcal{S}$ where $P = \{0,1\}^n$) on the sets of the form

$$\bigcap_{i \in V \subseteq \{1,\ldots,n\}} Y_i \tag{3.55}$$

(recall that the sets Y_i are defined by $Y_i = \{y \in \{0,1\}^n : y_i = 1\}$), and that the cone H (Definition 1.31) is the cone of the projections of these zeta vectors. Recall that the terms of the lattice L correspond to the subsets $V \subseteq \{1,\ldots,n\}$, and that the zeta vector z^V has a 1 in exactly those coordinates that correspond to sets $W \subseteq V$. If we rename the coordinates according to the mapping

$$W \to \bigcap_{i \in W \subseteq \{1,\ldots,n\}} Y_i \tag{3.56}$$

then it is evident that the values assigned by z^V to the original coordinates are the same as those assigned by $\zeta^{r(V)}$ to the new coordinates (where $r(V)$ is the atom corresponding to the point with 1's in exactly its V coordinates). For example, if $n = 3$ and $V = \{1,2\}$, then the zeta vector $(z^V)^T$ is

	$\emptyset$	1	2	3	1,2	1,3	2,3	1,2,3
$(z^{\{1,2\}})^T$	1	1	1	0	1	0	0	0

Renaming the coordinates according to the mapping (3.56), we obtain

$\{0,1\}^n$	Y_1	Y_2	Y_3	$Y_1 \cap Y_2$	$Y_1 \cap Y_3$	$Y_2 \cap Y_3$	$Y_1 \cap Y_2 \cap Y_3$
1	1	1	0	1	0	0	0

and these values do indeed correspond to the values of $\zeta^{(1,1,0)}$ (the zeta vector of the atom corresponding to the point $(1,1,0)$, we suppressed the dependence on $\{0,1\}^n$ in the notation) in the indicated coordinates (as they identify the sets to which $(1,1,0)$ belongs).

Thus this proves the following corollary, which is half of Theorem 1.35. The other half of that theorem has already been proven in the previous chapter in a different context (Corollary 2.24). It is also easy to derive the other half within the context developed here, and it will be a consequence of Lemma 3.29.

Recall that $\mathcal{A}$ is the algebra generated by the sets Y_i, $i = 1, \ldots, n$, i.e. it is $\mathcal{P}$ where $P = \{0,1\}^n$.

Corollary 3.20 *Let L be as in Lemma 1.17 and let L be indexed by the subsets $V \subseteq \{1, \ldots, n\}$. Let H be as in Definition 1.31, and assume that the vector $x \in H$ satisfies $x_\emptyset = 1$. Then there exists a probability measure χ on the algebra $\mathcal{A}$ such that for all $V \subseteq \{1, \ldots, n\}$,*

$$\chi\left(\bigcap_{i \in V} Y_i\right) = x_V. \tag{3.57}$$

Proof: If $x \in H$, then its lifting $\chi \in R^{\mathcal{P}}$ is in the cone of $\{\zeta^r : r \in \mathcal{S}\}$, and $x_\emptyset = 1$, after renaming coordinates, means that $\chi_{\{0,1\}^n} = 1$. Theorem 3.17 now implies that χ is a probability measure on $\mathcal{A}$. □

Observe also that the zeta vectors of the sets in $\mathcal{S}^P$ are indicator measures; ζ^r has value 1 for each q that contains r and 0 for each q that does not.

Until now we have been dealing exclusively with finite algebras, as P is a finite set, and this will continue to be the case throughout the coming chapters. For the purposes of future work, however, we remark that these results can be generalized to sets P of countably infinite size. In this generalization the notions of lifting and projecting are also replaced by more general mappings to and from a different space.

Theorem 3.21 *Let $P \subseteq R^n_+$ be countably large, and let $\mathcal{P}$ be the powerset of P (obviously $\mathcal{P}$ is a σ-algebra). For each pair of sets $u, v \in \mathcal{P}$ define*

$$\zeta^u(v) = \begin{cases} 1 : u \subseteq v \\ 0 : \ otherwise \end{cases} \tag{3.58}$$

(so that ζ^u can be thought of as a $0,1$ valued function on $\mathcal{P}$). Let g be a function that maps nonnegative real valued functions on $\mathcal{P}$ into the n dimensional extended reals, satisfying $g(\zeta^{\{x\}}) = x$ for every point $x \in P$, and

$$g\left(\sum_{i=1}^{\infty} \alpha_i h_i\right) = \sum_{i=1}^{\infty} \alpha_i g(h_i) \tag{3.59}$$

for all series $\sum_{i=1}^{\infty} \alpha_i h_i$, for which $\alpha \geq 0$, each h_i is of the form $\zeta^{\{x\}}, x \in P$, and the series is pointwise convergent to finite numbers. Then χ is a measure on $\mathcal{P}$ satisfying $\chi(\{y\}) < \infty$, $\forall y \in P$ iff there exist nonnegative scalars λ_y for each $y \in P$ such that

$$\chi = \sum_{y \in P} \lambda_y \zeta^{\{y\}} \tag{3.60}$$

(where χ is the pointwise limit). Note that the expression is well-defined since for each $u \in \mathcal{P}$, every $\lambda_y \zeta^{\{y\}}(u) \geq 0$. The set function χ is a probability measure on $\mathcal{P}$ iff those scalars can also be chosen such that $\sum_{y \in P} \lambda_y = 1$. Moreover $x \in R^n$ can be written as a countable convex combination of points in P iff there exists a probability measure χ on $\mathcal{P}$ that satisfies $g(\chi) = x$.

Proof: The assumptions of the theorem allow us to essentially reuse the demonstration from the $P \subseteq \{0,1\}^n$ case. The key issue in what follows is the fact that an infinite series of nonnegative terms is invariant under reordering (see, for example, Chapter 3 of [Ru64]). If we have $\chi = \sum_{y \in P} \lambda_y \zeta^{\{y\}}$, then for any pairwise disjoint sequence $\{u_j : j \geq 1\} \subseteq \mathcal{P}$,

$$\chi\left(\bigcup_{j=1}^{\infty} u_j\right) = \sum_{y \in P} \lambda_y \zeta^{\{y\}}\left(\bigcup_{j=1}^{\infty} u_j\right) = \sum_{y \in \bigcup_{j=1}^{\infty} u_j} \lambda_y = \text{ (by disjointness of the } u_j) \tag{3.61}$$

$$\sum_{j=1}^{\infty} \sum_{y \in u_j} \lambda_y = \sum_{j=1}^{\infty} \sum_{y \in P} \lambda_y \zeta^{\{y\}}(u_j) = \sum_{j=1}^{\infty} \chi(u_j) \tag{3.62}$$

and χ is clearly nonnegative, so χ is a measure, and $\chi(\{y\}) = \lambda_y < \infty$, $\forall y \in P$. If, additionally, $\sum_{y \in P} \lambda_y = 1$ then

$$\chi(P) = \sum_{y \in P} \lambda_y \zeta^{\{y\}}(P) = \sum_{y \in P} \lambda_y = 1. \tag{3.63}$$

Conversely, if χ is a measure on $\mathcal{P}$ with $\chi(\{y\} < \infty,\ \forall y \in P$, then for every $u \in \mathcal{P}$,

$$\chi(u) = \chi\left(\bigcup_{y \in u}\{y\}\right) = \sum_{y \in u}\chi(\{y\}) = \sum_{y \in P}\zeta^{\{y\}}(u)\chi(\{y\}) \Rightarrow \tag{3.64}$$

$$\chi = \sum_{y \in P}\chi(\{y\})\zeta^{\{y\}} \tag{3.65}$$

and each $\chi(\{y\})$ is finite by assumption and nonnegative since χ is a measure. If, additionally, $\chi(P) = 1$ then the expression above also implies that $\sum_{y \in P}\chi(\{y\}) = 1$. Finally, if $x \in R^n$ can be written

$$x = \sum_{y \in P}\lambda_y y,\ \lambda_y \geq 0\ \forall y,\ \sum_{y \in P}\lambda_y = 1 \tag{3.66}$$

then consider (what we now know is) the probability measure $\chi = \sum_{y \in P}\lambda_y\zeta^{\{y\}}$, and note that χ is a nonnegative finite real-valued function on $\mathcal{P}$. We therefore have

$$g(\chi) = g\left(\sum_{y \in P}\lambda_y\zeta^{\{y\}}\right) = \sum_{y \in P}\lambda_y g(\zeta^{\{y\}}) = \sum_{y \in P}\lambda_y y = x. \tag{3.67}$$

Conversely if there exists a probability measure χ on $\mathcal{P}$ for which $g(\chi) = x$, then there exist nonnegative scalars $\{\lambda_y :\ y \in P\}$ with $\sum_{y \in P}\lambda_y = 1$ such that

$$\chi = \sum_{y \in P}\lambda_y\zeta^{\{y\}} \tag{3.68}$$

and, as above, χ is a nonnegative finite real-valued function on $\mathcal{P}$, so

$$x = g(\chi) = g\left(\sum_{y \in P}\lambda_y\zeta^{\{y\}}\right) = \sum_{y \in P}\lambda_y g(\zeta^{\{y\}}) = \sum_{y \in P}\lambda_y y.\ \Box \tag{3.69}$$

One straightforward example is where $P \subseteq Z^n_+$ (so P is countable) with g defined as follows: Let

$$Y_i^j = \{y \in P : y_i \geq j\},\ i = 1, \ldots, n,\ j = 1, 2, \ldots \tag{3.70}$$

then for any $h : \mathcal{P} \to R^n_+$, let $g(h)$ be the point in R^n (extended) with

$$[g(h)]_i = \sum_{j=1}^{\infty} h(Y_i^j),\ i = 1, \ldots, n. \tag{3.71}$$

Remark 3.22 *Another generalization of $\mathcal{P}$, more in line with the prior development, would be as follows. Let $\{F_1, F_2 \ldots\}$ be a collection of subsets of a countable set in R^n. Let $\mathcal{P}$ be the σ-algebra generated by $\{F_i\}$. The role occupied by the points of P in the first statement* *of the theorem would now be played by the atoms of the algebra $\mathcal{P}$. With regard to g and the last statement of the theorem, assume $P \subseteq R_+^n$ is in one-to-one correspondence with the nonempty atoms of $\mathcal{P}$, and that g maps the zeta function of any any atom to its corresponding point in P.* $\square$

3.3 Measure and Signed Measure Consistency

Definition 3.23 *Let $\mathcal{Q} \subseteq \mathcal{P}$ be some subset of the algebra $\mathcal{P}$, and let $\tilde{\chi} \in R^{\mathcal{Q}}$ have coordinates corresponding only to those sets that belong to $\mathcal{Q}$. Then $\tilde{\chi}$ is said to be $\mathcal{P}$-signed-measure consistent if there exists a signed measure χ on $\mathcal{P}$ such that for all $q \in \mathcal{Q}$, the signed measure $\chi(q) = \tilde{\chi}_q$ (or, where χ is written as a vector in $R^{\mathcal{P}}$, the projection of χ on its $\mathcal{Q}$ coordinates is $\tilde{\chi}$). If such an χ can be chosen where χ is a measure on $\mathcal{P}$, then $\tilde{\chi}$ will be said to be $\mathcal{P}$-measure consistent. Where $P = \{0,1\}^n$, $\mathcal{P}$-measure and $\mathcal{P}$-signed-measure consistency will be referred to simply as measure and signed measure consistency.*

Speaking loosely, the following lemma states that a vector $\tilde{\chi}$ defined on some collection $\mathcal{Q}$ of sets in $\mathcal{P}$ is (signed) measure consistent iff the (signed) measure that it assigns to every set is the sum of the (signed) measures that it assigns to each of the nonempty atoms that constitute that set.

Lemma 3.24 *A vector $\tilde{\chi} \in R^{\mathcal{Q}}$ (i.e. the coordinates of $\tilde{\chi}$ correspond to the elements of $\mathcal{Q}$) is $\mathcal{P}$-signed-measure consistent iff there exists a number $\tilde{\chi}_r$ for every $r \in \mathcal{S}^P$ such that for each $q \in \mathcal{Q}$,*

$$\tilde{\chi}_q = \sum_{r \in \mathcal{S}^P : r \subseteq q} \tilde{\chi}_r. \tag{3.72}$$

(If $\emptyset \in \mathcal{Q}$ then the empty sum that would therefore equal $\tilde{\chi}_\emptyset$ should be understood to mean that $\tilde{\chi}_\emptyset = 0$.) The vector $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff these numbers can be chosen to be nonnegative.

Proof: Clearly additivity requires that the (signed) measure of any set is the sum of the nonempty atoms that comprise it. Conversely, suppose that there exist numbers $\tilde{\chi}_r, r \in \mathcal{S}^P$ satisfying the condition, then define the set function χ as follows. For each set $u \in \mathcal{P}$ assign

$$\chi(u) = \sum_{r \in \mathcal{S}^P : r \subseteq u} \tilde{\chi}_r. \tag{3.73}$$

Now for any two disjoint sets $u, v \in \mathcal{P}$,

$$u \cup v = \bigcup_{r \in \mathcal{S}^P : r \subseteq u} r \cup \bigcup_{r' \in \mathcal{S}^P : r \subseteq v} r' \tag{3.74}$$

and the union is disjoint, so it is clear that χ is additive and is therefore a signed measure. If, in addition, those numbers $\tilde{\chi}_r, r \in \mathcal{S}^P$ are all nonnegative, then by this definition the set function χ is nonnegative as well, and is therefore a measure. $\square$

Observe that $\mathcal{P} \subseteq \mathcal{A}$ (recall that $\mathcal{A}$ is the subset algebra of $\{0,1\}^n$) and thus a set function f on $\mathcal{P}$ can be thought of as a set function on $\mathcal{A}$ (i.e. a vector indexed by $\mathcal{A}$) for which we have identified only the function values of the sets that are also in $\mathcal{P}$. But recall that $\mathcal{A}$ is in one to one correspondence with the nonequivalent set theoretic expressions, $\Theta(Y_1, \dots, Y_n)$, and that for every $q \in \mathcal{A}$ there exists an expression

$$\Theta_q(Y_1, \dots, Y_n) = q. \tag{3.75}$$

Set theoretic expressions can be framed in terms of Y_i^P as well, and recall that that same expression framed in terms of Y_i^P (with P as the universal set) satisfies

$$\Theta_q^P(Y_1^P, \dots, Y_n^P) = q \cap P. \tag{3.76}$$

Thus a set function χ' on $\mathcal{P}$ has a natural representation as a vector χ indexed by $\mathcal{A}$, with coordinates corresponding to every nonequivalent expression $\Theta_q(Y)$, with value

$$\chi_q = \chi'(\Theta_q^P(Y_1^P, \dots, Y_n^P)) = \chi'(q \cap P). \tag{3.77}$$

Formally,

Definition 3.25 *As a notational convenience and to create a more unified framework, where $\mathcal{Q} \subseteq \mathcal{A}$ is the collection of sets*

$$q = \Theta_q(Y_1, \dots, Y_n), \ \forall q \in \mathcal{Q} \tag{3.78}$$

and $\mathcal{Q}' \subseteq \mathcal{P}$ is the collection of sets

$$q' = \Theta_q^P(Y_1^P, \dots, Y_n^P), \ \forall q \in \mathcal{Q} \tag{3.79}$$

we will allow ourselves to refer to set functions χ' on $\mathcal{P}$ defined on the collection of sets $\mathcal{Q}' \subseteq \mathcal{P}$ by a vector in $R^{\mathcal{Q}}$ with coordinates corresponding to $\mathcal{Q}$. This vector will be referred to as "representation of χ' w.r.t. $\mathcal{Q}$". The value of the q coordinate will be the function value on the set $\Theta_q^P(Y^P) = q \cap P$. Obviously such a vector can only describe a set function on $\mathcal{P}$ if it assigns the same value to all of the P-equivalent sets in $\mathcal{Q}$. Such vectors will be said to be "$\mathcal{P}$-set-function consistent". Thus a vector $\tilde{\chi} \in R^{\mathcal{Q}}$ will be said to be $\mathcal{P}$-(signed-)measure *consistent (i.e. it describes a signed measure on $\mathcal{P}$) if it is $\mathcal{P}$-set-function consistent and if the vector $\tilde{\chi}'$ defined on the collection $\mathcal{Q}' \in \mathcal{P}$ with*

$$\tilde{\chi}'_{q\cap P} = \tilde{\chi}_q, \ \forall q \in \mathcal{Q} \tag{3.80}$$

is $\mathcal{P}$-(signed-)measure consistent.

The following lemma adapts Lemma 3.24 for the new definition of (signed) measure consistency.

Lemma 3.26 *A vector $\tilde{\chi} \in R^{\mathcal{Q}}$ (where $\mathcal{Q} \subseteq \mathcal{A}$) is $\mathcal{P}$-signed-measure consistent iff there exist numbers χ_r for each $r \in \mathcal{S}^P$ such that for each $q \in \mathcal{Q}$,*

$$\tilde{\chi}_q = \sum_{r\in\mathcal{S}^P : r\subseteq q} \chi_r. \tag{3.81}$$

If these numbers are all nonnegative then $\tilde{\chi}$ is $\mathcal{P}$-measure consistent.

Proof: Two sets are P-equivalent iff they overlap on all of the atoms subset to P (i.e. on $\mathcal{S}^P$). So if such numbers exist then the $\tilde{\chi}$ values for P-equivalent sets are defined by the same sum, and therefore must be equal. Thus the vector $\tilde{\chi}'$ defined on the sets $\mathcal{Q}' \in \mathcal{P}$ where $\mathcal{Q}'$ is defined as the collection of sets

$$\{q \cap P : q \in \mathcal{Q}\} \tag{3.82}$$

for which

$$\tilde{\chi}'_{q'} = \tilde{\chi}_q, \ q' = q \cap P \tag{3.83}$$

is well defined and satisfies (where $q' = q \cap P$)

$$\tilde{\chi}'_{q'} = \tilde{\chi}_q = \sum_{r\in\mathcal{S}^P : r\subseteq q} \chi_r = \sum_{r\in\mathcal{S}^P : r\subseteq q'} \chi_r \tag{3.84}$$

since any $r \in \mathcal{S}^P : r \subseteq q$ also satisfies that $r \subseteq q \cap P = q'$ since every $r \in \mathcal{S}^P$ is contained in P. As for the converse, if $\tilde{\chi} \in R^{\mathcal{Q}}$ assigns the same values to all P-equivalent sets in $\mathcal{Q}$, and if the vector $\tilde{\chi}'$ defined on the sets $\mathcal{Q}' \in \mathcal{P}$ as above is $\mathcal{P}$-signed-measure consistent, then there exist numbers χ_r for each $r \in \mathcal{S}^P$ such that for all $q' \in \mathcal{Q}'$,

$$\tilde{\chi}'_{q'} = \sum_{r\in\mathcal{S}^P : r\subseteq q'} \chi_r. \tag{3.85}$$

Thus for each $q \in \mathcal{Q}$, where we denote $q' = q \cap P$,

$$\tilde{\chi}_q = \tilde{\chi}_{q'} = \sum_{r\in\mathcal{S}^P : r\subseteq q'} \chi_r = \sum_{r\in\mathcal{S}^P : r\subseteq q} \chi_r \tag{3.86}$$

by the same reasoning as above. □

Lemma 3.27 *Let $\mathcal{Q} \subseteq \mathcal{A}$, and let P^c refer to the set $\{0,1\}^n - P$. A vector $\tilde{\chi} \in R^{\mathcal{Q}}$ is $\mathcal{P}$-measure consistent iff the vector*

$$(\tilde{\chi}, \chi_{P^c}) \in R^{|\mathcal{Q}|+1} \tag{3.87}$$

obtained by appending the single coordinate χ_{P^c} with value

$$\chi_{P^c} = 0 \tag{3.88}$$

is measure consistent.

Note that we assumed in the statement of the theorem that $P^c \notin \mathcal{Q}$. If $P^c \in \mathcal{Q}$ then $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff $\tilde{\chi}$ is measure consistent and is such that $\tilde{\chi}_{P^c} = 0$.

Proof: By Lemma 3.24, to show that $(\tilde{\chi}, \chi_{P^c})$ is measure consistent we need to show that there exist nonnegative numbers χ_r for each $r \in \mathcal{S}$ such that for each $q \in \mathcal{Q}$, and for the additional set $P^c = \{0,1\}^n - P$,

$$\tilde{\chi}_q = \sum_{r \in \mathcal{S}: r \subseteq q} \chi_r \tag{3.89}$$

$$0 = \chi_{P^c} = \sum_{r \in \mathcal{S}: r \subseteq P^c} \chi_r. \tag{3.90}$$

But if $\tilde{\chi}$ is $\mathcal{P}$-measure consistent, then we already know that there exist nonnegative numbers χ_r for each $r \in \mathcal{S}^P$ such that for each $q \in \mathcal{Q}$,

$$\tilde{\chi}_q = \sum_{r \in \mathcal{S}^P: r \subseteq q} \chi_r. \tag{3.91}$$

Recall that $\mathcal{S}^P$ is the collection of all of the sets that contain a single point of P, and that $\mathcal{S}$ is the collection of all of the sets that contain a single point from $\{0,1\}^n$. Thus $\mathcal{S}^P \subseteq \mathcal{S}$, and if we assign these values to the χ_r, $r \in \mathcal{S}^P$, and we assign a value of zero to each χ_r, $r \in \mathcal{S} - \mathcal{S}^P$ (i.e. to each $r : r \subseteq P^c$) then both conditions will be satisfied. Conversely if $(\tilde{\chi}, \chi_{P^c})$ with $\chi_{P^c} = 0$ is measure consistent, then there exist nonnegative numbers χ_r for each $r \in \mathcal{S}$ such that for each $q \in \mathcal{Q}$, and for the additional set $P^c = \{0,1\}^n - P$,

$$\tilde{\chi}_q = \sum_{r \in \mathcal{S}: r \subseteq q} \chi_r \tag{3.92}$$

$$0 = \chi_{P^c} = \sum_{r \in \mathcal{S}: r \subseteq P^c} \chi_r. \tag{3.93}$$

By nonnegativity, the latter condition implies that for all $r \in \mathcal{S} : r \subseteq P^c$, i.e. for all $r \in \mathcal{S} - \mathcal{S}^P$, we must have $\chi_r = 0$. Thus the first condition becomes

$$\tilde{\chi}_q = \sum_{r \in \mathcal{S}^P : r \subseteq q} \chi_r \tag{3.94}$$

which implies that $\tilde{\chi}$ is $\mathcal{P}$-measure consistent. □

Remark 3.28 *By Definition 3.25, any set function defined on the sets $\Theta_q^P(Y^P)$ can be represented by the vector $\tilde{\chi}$ indexed by the expressions $\Theta_q(Y)$, and by Lemma 3.27, $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff it is consistent with a measure χ on $\mathcal{A}$ for which $\chi(P^c) = 0$. Thus another way to think about $\mathcal{P}$-measure consistency is as follows. Given a vector $\tilde{\chi}$ indexed by set theoretic expressions entailing sets Y_i^P, then $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff after renaming the indices according to those same set theoretic expressions, but this time of sets Y_i, the vector $\tilde{\chi}$ is $\mathcal{A}$-measure consistent with a measure χ for which $\chi(P^c) = 0$. (In particular, a vector $(x_1, \ldots, x_n) \in R^n$ belongs to $Conv(P)$ iff there is a probability measure χ on $\mathcal{A}$ for which $\chi(Y_i) = x_i,\ i = 1, \ldots, n$, and for which $\chi(P^c) = 0$.) Thus for any $P \subseteq \{0,1\}^n$, given any vector $\tilde{\chi}$ indexed by set theoretic expressions entailing sets Y_i^P, a necessary condition for the $\mathcal{P}$-measure consistency of $\tilde{\chi}$ is that the vector $\tilde{\chi}$, when its coordinates are construed as being indexed by those same expressions but of sets Y_i, be $\mathcal{A}$-measure consistent. Thus $\mathcal{A}$-measure consistency can be thought of as a relaxation of $\mathcal{P}$-measure consistency. We will see shortly that given $\mathcal{A}$-measure consistency, it is, on occasion, a simple matter to guarantee $\mathcal{P}$-measure consistency as well. We will see in Chapter 4, however, that there are circumstances in which $\mathcal{A}$-measure consistency does little to contribute toward guaranteeing $\mathcal{P}$-measure consistency.* □

Lemma 3.29 *A vector $\tilde{\chi} \in R^{\mathcal{Q}}$, where $\mathcal{Q} \subseteq \mathcal{A}$ is the collection of sets*

$$q = \Theta_q(Y_1, \ldots, Y_n),\ \forall q \in \mathcal{Q} \tag{3.95}$$

is measure consistent iff there exist sets $W_1, \ldots, W_n$ belonging to some universal set Ω such that for some measure X on the algebra generated by $W_1, \ldots, W_n$,

$$X(\Theta_q^\Omega(W_1, \ldots, W_n)) = \tilde{\chi}_q,\ \forall q \in \mathcal{Q}. \tag{3.96}$$

(The superscript Ω indicates that Ω is treated as the universal set with respect to the set theoretic expression.)

Proof: If $\tilde{\chi}$ is measure consistent then the sets $\{W_i\}$ and measure X clearly exist (just let $W_i = Y_i$ and $\Omega = \{0,1\}^n$). Conversely, suppose that such sets and measure exist. By Lemma 3.11 the algebra generated by $\{W_i\}$ is isomorphic to $\mathcal{P}$ for some $P \subseteq \{0,1\}^n$ (with $\{W_i\}$ corresponding to $\{Y_i^P\}$, and P corresponding to Ω). So all we need to show is that if there exists a $\mathcal{P}$ for which the vector $\tilde{\chi}'$ defined on the collection $\mathcal{Q}'$ of sets

$$q' = \Theta_q^P(Y_1^P, \ldots, Y_n^P),\ \forall q \in \mathcal{Q} \tag{3.97}$$

such that

$$\tilde{\chi}'_{q'} = \tilde{\chi}_q \tag{3.98}$$

is $\mathcal{P}$-measure consistent, then $\tilde{\chi}$ is measure consistent. But notice that $\tilde{\chi}$ is just the representation of $\tilde{\chi}'$ w.r.t. $\mathcal{Q}$, so if $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent then so is $\tilde{\chi}$, and it is then trivial from Lemma 3.27 that $\tilde{\chi}$ is measure consistent as well. $\square$

Putting Lemmas 3.27 and 3.29 together yields the following.

Definition 3.30 *Let $\mathcal{Q}'' \subseteq \mathcal{A}$ be a collection of sets satisfying*

1. *There exists a vector $\tilde{\chi}'' \in R^{\mathcal{Q}''}$ such that for any $\mathcal{P}$-measure χ (represented w.r.t. $\mathcal{A}$),*

$$\chi_{q''} = \tilde{\chi}''_{q''},\ \forall q'' \in \mathcal{Q}'' \tag{3.99}$$

2. *Every measure χ'' on $\mathcal{A}$ that coincides with $\tilde{\chi}''$ on $\mathcal{Q}''$ satisfies that*

$$\chi''_{P^c} = 0 \tag{3.100}$$

so that a measure χ on $\mathcal{A}$ is consistent with $\tilde{\chi}''$ iff χ is also a measure on $\mathcal{P}$ (represented w.r.t. $\mathcal{A}$). Or in other words, (since $\mathcal{P}$-measure consistency implies measure consistency,) a vector $\tilde{\chi} \in R^{\mathcal{Q}}$, $\mathcal{Q} \supseteq \mathcal{Q}''$, is $\mathcal{P}$-measure consistent iff it is measure consistent and it coincides with $\tilde{\chi}''$ on $\mathcal{Q}''$. The collection $\mathcal{Q}''$ will be said to be a "$\mathcal{P}$-measure test collection", and the vector $\tilde{\chi}''$ will be said to be a "$\mathcal{P}$-measure test vector" ("$\mathcal{P}$-test vector" for short).

We could also generalize the notion of a test vector to a case where a vector χ is $\mathcal{P}$-measure consistent iff it is measure consistent and its subvector on some $\mathcal{Q}'' \subseteq \mathcal{A}$ satisfies some constraint.

Corollary 3.31 *Let*

$$\Theta_{P^c}(Y_1, \ldots, Y_n) = P^c. \tag{3.101}$$

The vector $\tilde{\chi}$ defined on some collection of sets $\mathcal{Q} \subseteq \mathcal{A}$ defined by

$$\{\Theta_q(Y_, \ldots, Y_n) : q \in \mathcal{Q}\} \tag{3.102}$$

is $\mathcal{P}$-measure consistent iff there exist sets $W_1, \ldots, W_n$ contained in some Ω for which some measure X on the algebra generated by $W_1, \ldots, W_n$ satisfies

$$X(\Theta_q^{\Omega}(W_1, \ldots, W_n)) = \tilde{\chi}_q,\ \forall q \in \mathcal{Q} \tag{3.103}$$

and

$$X(\Theta_{P^c}^{\Omega}(W_1, \ldots, W_n)) = 0. \tag{3.104}$$

More generally, if $\tilde{\chi}''$ is a $\mathcal{P}$-test vector defined on a $\mathcal{P}$-test collection $\mathcal{Q}'' \subseteq \mathcal{A}$,

$$\{q'' = \Theta_{q''}(Y_1, \ldots, Y_n) : q'' \in \mathcal{Q}''\} \tag{3.105}$$

then the vector $\tilde{\chi}$ defined on the collection of sets $\mathcal{Q} \subseteq \mathcal{A}$ defined as above, is $\mathcal{P}$-measure consistent iff there exist sets $W_1, \ldots, W_n$ contained in some Ω for which some measure X satisfies

$$X(\Theta_q^{\Omega}(W_1, \ldots, W_n)) = \tilde{\chi}_q,\ \forall q \in \mathcal{Q} \tag{3.106}$$

and

$$X(\Theta_{q''}^{\Omega}(W_1, \ldots, W_n)) = \tilde{\chi}''_{q''},\ \forall q'' \in \mathcal{Q}''.\ \Box \tag{3.107}$$

Observe that given sets $\{W_1, \ldots, W_n\}$ and a universal set Ω, any measure X on any σ-algebra that includes the sets $\{W_i\}$ is also a measure on the algebra generated by the $\{W_i\}$. This corollary provides an interesting method for showing that a given description of a set $P \subseteq \{0,1\}^n$ by linear inequalities is convex hull defining.

Lemma 3.32 *Consider a set $P \subseteq \{0,1\}^n$ and a convex set T such that $P \subseteq T$, and let $\tilde{\chi}''$ be a $\mathcal{P}$-test vector defined on a $\mathcal{P}$-test collection $\mathcal{Q}'' \subseteq \mathcal{A}$. Then $T = Conv(P)$ iff for every $x \in T$ we can find sets $W_1, \ldots, W_n$ in some universal set Ω for which*

$$X(W_i) = x_i \tag{3.108}$$

where X is some probability measure on some σ-algebra including $\{W_i\}$, such that

$$X(\Theta_{q''}^{\Omega}(W_1, \ldots, W_n) = \tilde{\chi}''_{q''},\ \forall q'' \in \mathcal{Q}''. \tag{3.109}$$

Proof: A point $x \in R^n$ can be thought of as the set function χ defined on the collection

$$\mathcal{Q} = \{Y_1, \ldots, Y_n\} \tag{3.110}$$

and it belongs to the convex hull of P iff it is $\mathcal{P}$-probability-measure consistent ($\mathcal{P}$-measure consistency would guarantee membership in the cone). Thus by the corollary the set of points in the convex hull is the set of points x for which we can draw the sets described.

Thus the set T is contained in the convex hull iff we can draw such sets for every $x \in T$. But since we defined T to be convex and to contain P, we also have that $Conv(P) \subseteq T$, and this yields the lemma. □

So if, for example we let Ω be the unit interval on the real line and we let X be Lebesgue measure, then given a set of valid linear inequalities that hold on a set $P \subseteq \{0,1\}^n$, if those inequalities are convex hull defining, we have in principle a method of proving it "by picture".

Example: Consider the stable set problem with n nodes, and edge set E. Let the collection of stable sets be denoted $P \subseteq \{0,1\}^n$. Observe that the set of points in $\{0,1\}^n - P$ is exactly the set of points that have a 1 in positions i and j for some $\{i,j\} \in E$. Thus a measure on $\mathcal{A}$ assigns a measure of zero to the points in $\{0,1\}^n - P$ iff it assigns a measure of zero to all intersections $Y_i \cap Y_j,\ \{i,j\} \in E$. So the vector $\tilde{\chi}''$ defined on the collection

$$\mathcal{Q}'' = \{Y_i \cap Y_j : i,j \text{ s.t. } \{i,j\} \in E\} \tag{3.111}$$

with

$$\tilde{\chi}''_{Y_i \cap Y_j} = 0,\ \forall \text{ sets in } \mathcal{Q}'' \tag{3.112}$$

is a $\mathcal{P}$-test vector. Thus given a candidate convex set T such that

$$Conv(P) \subseteq T \subseteq [0,1]^n \tag{3.113}$$

if we can show that for any point $x \in T$ we can draw n sets W_i in the unit interval of length $x_i,\ i = 1,\ldots,n$, such that the Lebesgue measure of each $W_i \cap W_j$ is zero wherever $\{i,j\} \in E$, then we will have established that $Conv(P) = T$.[1] Observe that the conditions that must be satisfied therefore for $x \in R^n$ to belong to $Conv(P)$ are exactly the following:

$$\begin{gathered} x \in Conv(P) \text{ iff } \exists \text{ probability measure consistent } \tilde{\chi} \text{ defined on} \\ \mathcal{Q} = \{Y_1,\ldots,Y_n, Y_i \cap Y_j : \{i,j\} \in E\} \text{ with} \\ \tilde{\chi}_{Y_i} = x_i,\ i = 1,\ldots,n \text{ and} \\ \tilde{\chi}_{Y_i \cap Y_j} = 0,\ \forall \{i,j\} \in E \end{gathered} \tag{3.114}$$

Notice that beyond demanding that $\tilde{\chi}$ be probability measure consistent, the only additional requirements that $\tilde{\chi}$ must satisfy in order for its projection on the Y_i coordinates to belong to $Conv(P)$ are the $|E|$ equalities

$$\tilde{\chi}_{Y_i \cap Y_j} = 0,\ \forall \{i,j\} \in E. \tag{3.115}$$

[1] This fact is already implicit in [LS91] (putting together the "Remark" on page 186 with the characterization of $Cone(F)$ on page 187).

This establishes the following theorem.

Theorem 3.33 *If a polynomial time separation oracle exists for the set*

$$\{\tilde{\chi} \in R^{\mathcal{Q}} : \tilde{\chi} \textit{ is probability measure consistent}\} \tag{3.116}$$

where

$$\mathcal{Q} = \{Y_1, \ldots, Y_n, Y_i \cap Y_j : i, j = 1, \ldots, n\} \tag{3.117}$$

then P=NP.

It should also be noted that a polynomial time separation oracle exists for the set (3.116) iff one exists for the set

$$\{\tilde{\chi} \in R^{\bar{\mathcal{Q}}} : \tilde{\chi} \text{ is measure consistent}\} \tag{3.118}$$

where $\bar{\mathcal{Q}} = \mathcal{Q} \cup \{\{0,1\}^n\}$.

Proof: Let $P \subseteq \{0,1\}^n$ be the collection of (incidence vectors of) stable sets of a graph $G = (N, E)$ with $|N| = n$. If such an oracle existed, then we could test whether or not a vector $\tilde{\chi} \in R^{\mathcal{Q}}$ belongs to the set

$$V = \{\tilde{\chi} \in R^{\mathcal{Q}} : (\tilde{\chi}_{Y_1}, \ldots, \tilde{\chi}_{Y_n}) \in Conv(P)\} \tag{3.119}$$

by running the oracle on $\tilde{\chi}$ and then checking the $|E|$ equalities (3.115). Either both tests will pass or we will have a violated (in)equality. Thus this gives a polytime separation oracle for V. Since $Conv(P)$ is a projection of V a polytime separation oracle must then exist for $Conv(P)$, but this would imply that P=NP. □

Returning to the proofs by picture, and continuing with the assumption that P is the collection of stable sets of a graph, let us say now that the graph is complete. The linear inequalities

$$0 \leq x_i \leq 1, \ \forall i = 1, \ldots, n \tag{3.120}$$

$$\sum_{i=1}^{n} x_i \leq 1 \tag{3.121}$$

are clearly valid for $Conv(P)$. To prove that they are convex hull defining we need only show that for every $x \in R^n$ that satisfies these inequalities we can draw n sets W_i of length $x_i, i = 1, \ldots, n$ on the unit interval, such that no two of which intersect. But this is trivial: consider

$$W_1 = [0, x_1), \ W_2 = [x_1, x_1 + x_2), \ldots, W_n = (\sum_{i=1}^{n-1} x_i, \sum_{i=1}^{n} x_i] \tag{3.122}$$

As $\sum_{i=1}^{n} x_i \leq 1$, these sets clearly qualify. □

Example 2: Consider the set

$$P = \{y \in \{0,1\}^n : \sum_{i=1}^{n} y_i \geq \ 1\} \tag{3.123}$$

and consider the linear relaxation

$$0 \leq x_i \leq 1 \tag{3.124}$$

$$\sum_{i=1}^{n} x_i \geq 1. \tag{3.125}$$

By definition, a measure χ on $\mathcal{A}$ assigns a measure of zero to the points that do not belong to P iff it assigns a measure of zero to the points that violate the inequality $\sum_{i=1}^{n} y_i \geq \ 1$. The points that violate this inequality are exactly those that do not have any coordinates at value 1, i.e. the set $\bigcap_{i=1}^{n} Y_i^c$. So the test vector is

$$\tilde{\chi}''_{Y_1^c \cap \cdots \cap Y_n^c} = 0. \tag{3.126}$$

Thus a vector $x \in R^n$ belongs to $Conv(P)$ iff we can draw sets W_i of length x_i such that the length of $W_1^c \cap \cdots \cap W_n^c$ is zero, i.e. iff the sets W_i cover the whole interval. Clearly if the sum of the lengths is ≥ 1 then this can be done. □

Example 3: Consider the set

$$P = \{y \in \{0,1\}^n : \sum_{j \in A_i} y_j \geq 1,\ i = 1, \ldots, m\} \tag{3.127}$$

where each $A_i \subseteq \{1, \ldots, n\}$, or equivalently,

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} Y_j. \tag{3.128}$$

Theorem 3.34 *Suppose that the sets $\{A_i\}$ are pairwise disjoint, then the vector $x \in [0,1]^n$ belongs to $Conv(P)$ iff*

$$\sum_{j \in A_i} x_j \geq 1,\ i = 1, \ldots, m. \tag{3.129}$$

Proof: As above, by definition, a probability measure χ on $\mathcal{A}$ assigns a measure of zero to the points that do not belong to P iff it assigns a measure of zero to the points that violate the inequalities that define P. These are the points that have all zeroes in their A_i coordinates for some A_i. The test vector is therefore

$$(\tilde{\chi}''\left(\bigcap_{j \in A_1} Y_j^c\right), \ldots, \tilde{\chi}''(\bigcap_{j \in A_m} Y_j^c)) = 0. \tag{3.130}$$

So we are looking for a probability measure that assigns a value of zero to these sets. Equivalently, we are looking for a probability measure that assigns a value of 1 to each of the sets $\bigcup_{j\in A_i} Y_j$, i.e. a test vector can also be formulated as

$$(\tilde{\chi}''(\bigcup_{j\in A_1} Y_j),\ldots,\tilde{\chi}''(\bigcup_{j\in A_m} Y_j)) = (1,\ldots,1). \tag{3.131}$$

Let $\Omega = [0,1]$ and let X be the Lebesgue measure (which is a probability measure on Ω). Rename the elements of each A_i as $i_1,\ldots,i_{|A_i|}$ (with any ordering). Given $x \in [0,1]^n$ satisfying (3.129), for each $i = 1,\ldots,m$, draw sets $W_{i_1} = [0, x_{i_1}]$, and

$$W_{i_j} = \begin{cases} [\sum_{k=1}^{j-1} x_{i_k}, \sum_{k=1}^{j} x_{i_k}] : \sum_{k=1}^{j} x_{i_k} \leq 1 \\ [1 - x_{i_j}, 1] : \sum_{k=1}^{j} x_{i_k} > 1 \end{cases} \tag{3.132}$$

for $j \geq 1$. The fact that the $\{A_i\}$ are all disjoint implies that these definitions are consistent, $X(W_{i_j}) = x_{i_j}$ for all i_j, and by the constraints

$$\bigcup_{j\in A_i} W_j = \bigcup_{j=1}^{|A_i|} W_{i_j} = [0,1] \Rightarrow X\left(\bigcup_{j\in A_i} W_j\right) = 1,\ \forall i = 1,\ldots,m.\ \square \tag{3.133}$$

Example 4: Here we will use the set P of the previous example, but we will be looking only for simple measure consistency on the collection $\{Y_1,\ldots,Y_n,P^c\}$. The difference between this example and the previous is that here we will not require P^c to have a probability measure of 0. We will see that in this case, if we consider only constraints entailing variables exclusively from the collection $\chi(Y_1),\ldots,\chi(Y_n),\chi(P^c)$, then an exponential number of constraints is required to establish probability measure consistency. We will also see, however, that if we allow the constraints to entail some extra variables corresponding to some other sets, then the number of constraints can be reduced to a polynomial quantity.

Theorem 3.35 *Let χ be a set function on $\mathcal{A}$. Let s and $t_1,\ldots,t_s$ be nonnegative integers; let every $l_j^i \subseteq \{1,\ldots,n\}$; let every l_j^i be distinct, and denote the value*

$$\chi\left(\bigcup_{i=1}^{s}\bigcap_{j=1}^{t_i} Y_{l_j^i}\right) \tag{3.134}$$

as χ_q. Consider the collection of set function values

$$\{\chi(Y_1),\ldots,\chi(Y_n),\chi_q\}. \tag{3.135}$$

These values are probability measure consistent iff

1. $0 \le \chi(Y_i) \le 1,\ i = 1, \ldots, n$

2. $0 \le \chi_q \le 1,\ i = 1, \ldots, n$

3. $\chi_q \le \sum_{i=1}^{s} \chi(Y_{l^i_{j(i)}})$ *for all of the* $\prod_{i=1}^{s} t_i$ *possible choices of* $j(1), \ldots, j(s) : j(i) \in \{1, \ldots, t_i\}, i = 1, \ldots, s$

and the inequalities $\chi_q \le \sum_{i=1}^{s} \chi(Y_{l^i_{j(i)}})$ *are all facet defining for the polytope in* $[0,1]^{n+1}$ *that is defined by the system of constraints (1), (2) and (3). If, however, we append variables* $\{\chi_{q_i},\ i = 1, \ldots, s\}$ *(corresponding to the sets* $\{q_i = \bigcap_{j=1}^{t_i} Y_{l^i_j},\ i = 1, \ldots, s\}$*), then the collection* $\{\chi(Y_1), \ldots, \chi(Y_n), \chi_q\}$ *is probability measure consistent iff there is an expanded system of variables*

$$(\chi(Y_1), \ldots, \chi(Y_n), \chi_{q_1}, \ldots, \chi_{q_s}, \chi_q) \tag{3.136}$$

that satisfies

i. $0 \le \chi(Y_i),\ \chi_q,\ \chi_{q_j} \le 1,\ \forall i, j$

ii. $\chi_{q_i} \le \chi(Y_{l^i_j})$ *for each* $j \in \{1, \ldots, t_i\}$*.*

iii. $\chi_{q_i} \ge \sum_{j=1}^{t_i} \chi(Y_{l^i_j}) - t_i + 1$

iv. $\chi_q \ge \chi_{q_i}$ *for each* $i \in \{1, \ldots, s\}$*.*

v. $\chi_q \le \sum_{i=1}^{s} \chi_{q_i}$

Proof: We will consider the latter case first. We will prove the theorem by showing that we can find subsets $\{W_1, \ldots, W_n\}$ of $[0, 1]$ such that the length of each set W_i is $\chi(Y_i)$ and the length of

$$W_q := \bigcup_{i=1}^{s} \bigcap_{j=1}^{t_i} W_{l^i_j} \tag{3.137}$$

is χ_q iff constraints (i) – (v) hold. Define the sets $q_i,\ i = 1, \ldots, s$ by $q_i = \bigcap_{j=1}^{t_i} Y_{l^i_j}$. All the constraints stated are necessary since for any measure χ and any collection of sets $\{Q_1, \ldots, Q_m\}$ on which the measure is defined we have by the definition of measures

$$\chi\left(\bigcup_{i=1}^{m} Q_i\right) \le \sum_{i=1}^{m} \chi(Q_i) \tag{3.138}$$

and since χ is a probability measure

$$\chi\left(\bigcap_{i=1}^{m} Q_i\right) = 1 - \chi\left(\bigcup_{i=1}^{m} Q_i^c\right) \ge 1 - \sum_{i=1}^{m} \chi(Q_i^c) = \sum_{i=1}^{m} \chi(Q_i) - m + 1. \tag{3.139}$$

We will now prove sufficiency. By (i) and (iii) we have

$$\chi_{q_i} \geq \max\{0, \sum_{j=1}^{t_i} \chi(Y_{l^i_j}) - t_i + 1\} \tag{3.140}$$

and by (iv) we have

$$\chi_q \geq \max_{i=1,\ldots,s}\{\chi_{q_i}\} \tag{3.141}$$

and we therefore conclude that

$$\chi_q \geq \max_{i=1,\ldots,s}\left\{\max\{0, \sum_{j=1}^{t_i} \chi(Y_{l^i_j}) - t_i + 1\}\right\}. \tag{3.142}$$

Similarly by (i), (ii) and (v) we conclude that

$$\chi_q \leq \min\left\{1, \sum_{i=1}^{s} \min_{j=1,\ldots,t_i}\{\chi(Y_{l^i_j})\}\right\}. \tag{3.143}$$

These provide the upper and lower bounds on the values of χ_q allowable by the constraints. Since these constraints are necessary in any measure space, so where the Lebesgue measure of a set $W \subseteq [0,1]$ is denoted $X(W)$, we similarly have that

$$X(W_q) \geq \max_{i=1,\ldots,s}\left\{\max\{0, \sum_{j=1}^{t_i} X(W_{l^i_j}) - t_i + 1\}\right\} \tag{3.144}$$

and

$$X(W_q) \leq \min\left\{1, \sum_{i=1}^{s} \min_{j=1,\ldots,t_i}\{X(W_{l^i_j})\}\right\}. \tag{3.145}$$

We will show first that we can draw sets $W_1, \ldots, W_n$ in $[0,1]$ of lengths $\chi(Y_1), \ldots, \chi(Y_n)$ for which the length of $X(W_q)$ will meet these upper and lower bounds. We will show this first for the lower bound. Define the sets

$$W_{q_i} := \bigcap_{j=1}^{t_i} W_{l^i_j},\ i = 1, \ldots, s. \tag{3.146}$$

The idea here is that the length of W_q will be minimized by minimizing the length of each W_{q_i}, and then placing all of the sets $\{W_{q_i}\}$ on top of the longest one of them. Each W_{q_i} is minimized by distributing the sets $W_{l^i_j}$ as disparately as possible. Specifically, draw $W_{l^1_1}$ as the interval $[0, \chi(Y_{l^1_1}))$, then draw $W_{l^1_2}$ as the interval

$$[\chi(Y_{l^1_1}), \chi(Y_{l^1_1}) + \chi(Y_{l^1_2})) \tag{3.147}$$

where $\chi(Y_{l^1_1}) + \chi(Y_{l^1_2}) \leq 1$. If $\chi(Y_{l^1_1}) + \chi(Y_{l^1_2}) > 1$ then draw it on

$$[\chi(Y_{l^1_1}), 1] \cup [0, \chi(Y_{l^1_1}) + \chi(Y_{l^1_2}) - 1) \tag{3.148}$$

etc. Continuing in this manner for all t_1 sets $W_{l_j^1}$, we will make some number ($\leq t_1$) of passes over the interval $[0,1]$, and the number of sets $W_{l_j^1}$ to which any point in $[0,1]$ will belong will be the number of times that we will have passed over that point using this procedure (since no one set can pass over a single point twice). Thus the set of points that will belong to all t_1 sets is the set of points that are covered by the t_1'st pass over the interval $[0,1]$, and this set is empty unless we make more than $t_1 - 1$ passes over the interval. If there is, in fact, a t_1'st pass, then the size of the set covered in that pass is

$$\sum_{j=1}^{t_1} \chi(Y_{l_j^1}) - (t_1 - 1) \tag{3.149}$$

and therefore the size of W_{q_1} is

$$\max\{0, \sum_{j=1}^{t_1} \chi(Y_{l_j^1}) - t_1 + 1\}. \tag{3.150}$$

Moreover, the set W_{q_1} is the interval

$$[0, X(W_{q_1})). \tag{3.151}$$

Since the $\{l_j^i\}$ are all distinct we can repeat this for all $q_i,\ i = 1, \ldots, s$, and the length of W_q will thus be

$$\max_{i=1,\ldots,s} \left\{ \max\{0, \sum_{j=1}^{t_i} \chi(Y_{l_j^i}) - t_i + 1\} \right\}. \tag{3.152}$$

The largest length, on the other hand, that the set W_q could attain would be obtained by maximizing the length of each W_{q_i} by placing all of the $W_{l_j^i}$ for a given i on top of the longest one, and then by distributing the sets W_{q_i} as disparately as possible, as above. The length thus obtained for each W_{q_i} will be

$$\min_{j=1,\ldots,t_i} \{\chi(Y_{l_j^i})\} \tag{3.153}$$

and the length for W_q will be

$$\min \left\{ 1, \sum_{i=1}^{s} \min_{j=1,\ldots,t_i} \{\chi(Y_{l_j^i})\} \right\}. \tag{3.154}$$

It isn't hard to see that any length in between is possible as well, and thus all values for $\chi(Y_1), \ldots, \chi(Y_n), \chi_q$ allowed by the constraints correspond to the lengths of actual sets in $[0,1]$ and these values are therefore measure consistent.

We have now proved the second half of the theorem, but to make the argument more concrete, consider the following numerical example. Suppose we are given the collection

$$\{\chi(Y_1), \ldots, \chi(Y_9), \chi_q = \chi((Y_1 \cap Y_2 \cap Y_3) \cup (Y_4 \cap Y_5 \cap Y_6) \cup (Y_7 \cap Y_8 \cap Y_9))\} \tag{3.155}$$

with values

$\chi(Y_1)$	.7
$\chi(Y_2)$	.6
$\chi(Y_3)$	.1
$\chi(Y_4)$	.7
$\chi(Y_5)$	.8
$\chi(Y_6)$	.9
$\chi(Y_7)$	.1
$\chi(Y_8)$	.2
$\chi(Y_9)$	.5

The lower bound (3.142) for χ_q allowable under constraints (i) – (v) is

$$\chi_q \geq \max_{i=1,\ldots,3} \left\{ \max\{0, \sum_{j=1}^{3} \chi(Y_{l_j^i}) - 2\} \right\} = .4 \tag{3.156}$$

Draw sets

W_1	$[0, .7)$
W_2	$[.7, 1] \cup [0, .3)$
W_3	$[.3, .4)$
W_4	$[0, .7)$
W_5	$[.7, 1] \cup [0, .5)$
W_6	$[.5, 1] \cup [0, .4)$
W_7	$[0, .1)$
W_8	$[.1, .3)$
W_9	$[.3, .8)$

Thus

$$W_1 \cap W_2 \cap W_3 = \emptyset \tag{3.157}$$

$$W_4 \cap W_5 \cap W_6 = [0, .4) \tag{3.158}$$

$$W_7 \cap W_8 \cap W_9 = \emptyset \tag{3.159}$$

and

$$W_q = (W_1 \cap W_2 \cap W_3) \cup (W_4 \cap W_5 \cap W_6) \cup (W_7 \cap W_8 \cap W_9) = [0, .4) \tag{3.160}$$

which is indeed of Lebesgue measure .4.

The upper bound (3.143) for χ_q allowable under constraints (i) – (v) is

$$\chi_q \leq \min\left\{1, \sum_{i=1}^{3} \min_{j=1,\ldots,3}\{\chi(Y_{l_j^i})\}\right\} = .9 \tag{3.161}$$

Draw sets

W_1	$[0, .7)$
W_2	$[0, .6)$
W_3	$[0, .1)$
W_4	$[.1, .8)$
W_5	$[.1, .9)$
W_6	$[.1, 1)$
W_7	$[.8, .9)$
W_8	$[.7, 1)$
W_9	$[.5, 1)$

Thus

$$W_1 \cap W_2 \cap W_3 = [0, .1) \tag{3.162}$$

$$W_4 \cap W_5 \cap W_6 = [.1, .8) \tag{3.163}$$

$$W_7 \cap W_8 \cap W_9 = [.8, .9) \tag{3.164}$$

and

$$W_q = (W_1 \cap W_2 \cap W_3) \cup (W_4 \cap W_5 \cap W_6) \cup (W_7 \cap W_8 \cap W_9) = [0, .9) \tag{3.165}$$

which is indeed of Lebesgue measure .9. It is also to see that for any value of χ_q between .4 and .9 we can select sets $W_1, \ldots, W_9$ for which the Lebesgue measure of W_q will be χ_q. This concludes the numerical example. We will now prove the first half of the theorem.

If we tried to obtain the same result without adding the new variables $\{\chi_{q_i}\}$, the only way to ensure that

$$\chi_q \leq \min\left\{1, \sum_{i=1}^{s} \min_{j=1,\ldots,t_i}\{\chi(Y_{l_j^i})\}\right\} \tag{3.166}$$

would be to write the $\prod_{i=1}^{s} t_i$ inequalities

$$\chi_q \leq \sum_{i=1}^{s} \chi(Y_{l_{j(i)}^i}) \tag{3.167}$$

for all of the $\prod_{i=1}^{s} t_i$ possible choices of $j(1), \ldots, j(s) : j(i) \in \{1, \ldots, t_i\}, i = 1, \ldots, s$. In order to see this we will show that these constraints are all facet defining, as follows. Note that each choice of $j(1), \ldots, j(s)$ can be thought of as a function J. Represent the constraint corresponding to each such function J as $(a^J)^T\chi \geq \chi_q$. We will show that no vector $(0, a^J, -1) \in R^{n+2}$ can be written as a nonnegative linear combination of other vectors $(0, a^{J'}, -1)$, the unit vectors in R^{n+2}, $e_0, e_1, \ldots, e_{n+1}$ and the vectors $e_0 - e_i,\ i = 1, \ldots, n+1$ (these are the vectors that correspond to the conic form of the original constraint set). Suppose that one of the vectors $(0, \bar{a}, -1)$ of the form $(0, a^J, -1)$ is indeed such a nonnegative linear combination. Observe first that the combination could not possibly entail any of the vectors $e_0 - e_i$, so it must be of the form

$$(\bar{a}, -1) = \sum_{\text{functions } J} \lambda_J(a^J, -1) + \sum_{i=1}^{n+1} \gamma_i e_i, \quad \lambda, \gamma \geq 0. \tag{3.168}$$

Thus

$$-\sum_{\text{functions } J} \lambda_J + \gamma_{n+1} = -1 \Rightarrow \sum_{\text{functions } J} \lambda_J \geq 1. \tag{3.169}$$

Observe now that for all constraints of the form $(a^J)^T x \geq \chi_q$, where

$$H^i = \{l^i_j : j = 1, \ldots, t_i\} \tag{3.170}$$

then $a^J_h = 1$ for exactly one $h \in H^i$ and is zero for its remaining H^i coordinates, and that every positive coordinate $a^J_h = 1$ represents such a choice from some H^i. Say now that $\bar{a}_{h_l} = 1$ and that $h_l \in H^i$, so that $\bar{a}_{h_j} = 0,\ \forall h_j \in H^i, h_j \neq h_l$. Then for each positive λ_J, we must have

$$a^J_{h_j} = 0,\ \forall h_j \in H^i, h_j \neq h_l \Rightarrow a^J_{h_l} = 1. \tag{3.171}$$

Thus

$$1 = \bar{a}_{h_l} = \sum_J \lambda_J a^J_{h_l} + \gamma_{h_l} = \sum_J \lambda_J + \gamma_{h_l}. \tag{3.172}$$

But $\sum_J \lambda_J \geq 1$ and so we conclude that $\gamma_{h_l} = 0$ and $\sum_J \lambda_J = 1$. Thus for each positive coordinate $\bar{a}_h = 1$ we have $\gamma_h = 0$, and obviously for each zero coordinate $\bar{a}_h = 0$ we have $\gamma_h = 0$, so we conclude that $\gamma = 0$ and therefore $\bar{a}$ is a convex combination of a^J vectors. But this is impossible, as these are all distinct $0, 1$ vectors. □

Example 5: This is not an example of the proof by picture method, but it is an example of a classical problem for which regular probability measure consistency (i.e. $\mathcal{A}$-probability-measure consistency rather than $\mathcal{P}$-probability-measure consistency, recall that $\mathcal{A}$ is the subset algebra for $\{0,1\}^n$) is all that is needed, even without defining $\mathcal{P}$-test vectors. It also shows that the variables that most interest us need not always be $\chi(Y_i)$.

Consider the $0,1$ representation of the cut polytope. This is the convex hull of the set of all vectors y^C indexed by the edges of an n node undirected graph $G = (V, E)$, such that $y^C_e = 1$ if e crosses the cut C, and $y^C_e = 0$ otherwise. (A cut is a subset of the nodes of a graph, and an edge is said to "cross" the cut if exactly one of its endpoints belongs to that subset.) This can be modeled as follows. There is a cut C for each subset of the n nodes, so if we define the set P to be the set of all *node* incidence vectors of cuts, then $P = \{0, 1\}^n$. Observe now that for each cut C with incidence vector $y(C) \in P = \{0, 1\}^n$, the edge $\{i, j\}$ crosses the cut defined by y iff

$$y \in (Y_i \cap Y_j^c) \cup (Y_i^c \cap Y_j). \tag{3.173}$$

So letting $\zeta^{y(C)}$ be the zeta vector corresponding to $y(C)$ it follows that each $\{i, j\} \in E$ coordinate of the edge incidence vector of C is just

$$\zeta^{y(C)}((Y_i \cap Y_j^c) \cup (Y_i^c \cap Y_j)). \tag{3.174}$$

So the convex hull of edge incidence vectors for cuts is just the convex hull of the projections of the zeta vectors for $\mathcal{A}$ on the coordinates $(Y_i \cap Y_j^c) \cup (Y_i^c \cap Y_j),\ \{i, j\} \in E$, so by Corollary 3.19 a vector belongs to the convex hull of the edge incidence vectors of the cuts iff that vector, when it is construed to have a coordinate for each set $(Y_i \cap Y_j^c) \cup (Y_i^c \cap Y_j),\ \{i, j\} \in E$ (corresponding to what is usually construed as its $\{i, j\}$ coordinate), is probability measure consistent (i.e. $\mathcal{A}$-probability-measure consistent). (An alternative treatment of the cut polytope that deals with $+1, -1$ moment matrices is given in [Lau01], but this is the most natural approach for our framework.)

3.4 Delta Vectors

In the previous section we saw that a (signed) measure on $\mathcal{P}$ is completely determined by the values it places on the sets $\mathcal{S}^P$. We will show in this section that this is true for other collections of sets as well.

Definition 3.36 *Denote the zeta matrix of the algebra $\mathcal{P}$ as $\mathcal{Z}_{\mathcal{P}}$, and the zeta matrix of the algebra $\mathcal{A}$ as $\mathcal{Z}$. Denote the submatrix of $\mathcal{Z}_{\mathcal{P}}$ corresponding to the columns of the sets in $Q \subseteq \mathcal{P}$ as $\mathcal{Z}^Q_{\mathcal{P}}$, and denote the submatrix of $\mathcal{Z}^Q_{\mathcal{P}}$ corresponding to the rows of the sets in $G \subseteq \mathcal{P}$ as $\mathcal{Z}^Q_{\mathcal{P}}\{G\}$. (Individual rows of the matrix will also be denoted in the same fashion.) So the matrix made up of the zeta columns of the nonempty atoms is $\mathcal{Z}^{\mathcal{S}^P}_{\mathcal{P}}$. If $G \subseteq \mathcal{P}$ is such that the rows of $\mathcal{Z}^{\mathcal{S}^P}_{\mathcal{P}}\{G\}$ span all of the rows of $\mathcal{Z}^{\mathcal{S}^P}_{\mathcal{P}}$, then G will be said to be a* *spanning collection. If the rows of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ are linearly independent then G will be said to be a linearly independent collection.*

Observe that as $\mathcal{P} \subseteq \mathcal{A}$, the matrices $\mathcal{Z}_{\mathcal{P}}$ are submatrices of $\mathcal{Z}$.

Lemma 3.37 *The matrix $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}$ has full column rank. Thus where G is a linearly independent spanning collection, the matrix $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ is $|\mathcal{S}^P| \times |\mathcal{S}^P|$ and nonsingular.*

Proof: For each atom $r \in \mathcal{S}^P$, the only atom to contain r is r itself, thus the row $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{r\}$ has a 1 in its r position and zeroes elsewhere. The matrix $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\mathcal{S}^P\}$ is therefore just the identity matrix, which is of full rank. □

Definition 3.38 *Let $G \subseteq \mathcal{P}$ be a linearly independent spanning collection. For each $q \in \mathcal{P}$ define the vector*

$$(\mu^G(q))^T = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}(\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\})^{-1} \tag{3.175}$$

so that

$$(\mu^G(q))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} \tag{3.176}$$

i.e. $\mu^G(q)$ is the vector indexed by the rows of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ that describes how to obtain the q'th row of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}$ in terms of the G rows. In general, where G is an arbitrary collection, the set of vectors that describes the row $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ in terms of the rows of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ is the (possibly empty) affine set

$$M^G(q) = \{\mu^G(q) \in R^G : (\mu^G(q))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}\}. \tag{3.177}$$

Note that if G is spanning then for all $q \in \mathcal{P},\ M^G(q) \neq \emptyset$.

Observe that the q'th row of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}$ is just the list of points (nonempty atoms) that are contained in the set q. Thus the vectors $\mu^G(q)$ can be thought of as describing the set q in terms of the sets in G. Notice also that, where we denote the subcolumns of the zeta columns ζ^r corresponding to G as $\zeta^r\{G\}$, for any point (nonempty atom) $r \in \mathcal{S}^P$,

$$(\mu^G(q))^T \zeta^r\{G\} = \zeta^r_q = \begin{cases} 1 : r \subseteq q \\ 0 : \text{ otherwise} \end{cases} \tag{3.178}$$

Calculating the inner product of $\zeta^r\{G\}$ with $\mu^G(q)$ thus provides a single test for all points $r \in \mathcal{S}^P$ of whether or not the point r belongs to the set q. This test is the same for every column $\zeta^r\{G\}$, and this is responsible for the following lemma.

Lemma 3.39 *Given a set function χ on $\mathcal{P}$, let $\tilde{\chi} \in R^G$ be the vector that lists the function values on the sets in the linearly independent spanning collection $G \subseteq \mathcal{P}$. χ is a signed measure on $\mathcal{P}$ iff for every $q \in \mathcal{P} - G$,*

$$\chi_q = (\mu^G(q))^T \tilde{\chi} \tag{3.179}$$

or equivalently, iff for every $q \in \mathcal{P}$,

$$\chi_q = (\mu^G(q))^T \tilde{\chi}. \tag{3.180}$$

χ is a measure on $\mathcal{P}$ iff χ is a signed measure on $\mathcal{P}$, and $\tilde{\chi}$ is $\mathcal{P}$-measure consistent.

Proof: Let $\tilde{\tilde{\chi}}$ be the subvector of χ that lists its values on the sets in $\mathcal{P} - G$, so that $\chi = (\tilde{\chi}, \tilde{\tilde{\chi}})$. By Corollary 3.19, χ is a signed measure on $\mathcal{P}$ iff it is a linear combination of the columns of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}$, i.e. iff there exists an α with a coordinate for each of the nonempty atoms of $\mathcal{P}$ such that $\chi = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P} \alpha$, or equivalently, iff there exists an α such that

$$\tilde{\chi} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\alpha \text{ and } \tilde{\tilde{\chi}} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\mathcal{P} - G\}\alpha. \tag{3.181}$$

By definition this is equivalent to

$$\tilde{\chi} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\alpha \text{ and } \tilde{\tilde{\chi}}_q = (\mu^G(q))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\alpha = (\mu^G(q))^T \tilde{\chi}, \ \forall q \in \mathcal{P} - G. \tag{3.182}$$

Since $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ is nonsingular, an α that satisfies the first condition always exists and is uniquely determined. Observe also that the condition $\chi_q = (\mu^G(q))^T \tilde{\chi}$ for $q \in G$ is always satisfied as for $q \in G$, $\mu^G(q) = e_q$, so this proves that χ is a signed measure on $\mathcal{P}$ iff for every $q \in \mathcal{P}$,

$$\chi_q = (\mu^G(q))^T \tilde{\chi}. \tag{3.183}$$

The only additional condition required for χ to be a measure is that α (which is uniquely determined) is nonnegative. But this is exactly the condition for $\tilde{\chi}$ to be $\mathcal{P}$-measure consistent. □

Corollary 3.40 *Let G be a linearly independent collection, then every vector $\tilde{\chi} \in R^G$ is signed measure consistent.*

Proof: Let $G' \supseteq G$ be a linearly independent spanning collection. Append coordinates to $\tilde{\chi}$ for each set in $G' - G$ with arbitrary values. The resulting vector is signed measure consistent by Lemma 3.39. □

The vectors introduced in Definition 3.38 are quite useful, and we will be returning to them repeatedly later on. We now present a batch of results, all of which are fairly straightforward, that provide stretches and variations on the themes of Definition 3.38 and Lemma 3.39 that will prove helpful in later sections and chapters. The primary focus of these results will be on using these vectors to characterize set function values and to provide characterizations of measure and signed measure consistency.

Lemma 3.41 *Let $q \in \mathcal{P}$. Let $G \subseteq \mathcal{P}$ be a spanning collection, and let $\mathcal{Q} \subseteq G$. Then there exists a vector $\mu^{\mathcal{Q}}(q) \in M^{\mathcal{Q}}(q)$ iff the vector $(\mu^{\mathcal{Q}}(q), 0) \in R^G$ obtained by padding $\mu^{\mathcal{Q}}(q)$ with zeroes in its $G - \mathcal{Q}$ coordinates belongs to $M^G(q)$.* □

The following lemma lists some easy generalizations of the above results. The first statement of the lemma is obvious. The other statements concerning $\mathcal{P}$-signed-measure and $\mathcal{P}$-measure consistency can be proven in the same manner as Lemma 3.39.

Lemma 3.42 *Let $\mathcal{Q} \subseteq \mathcal{P}$ and let $\mathcal{Q}' \subseteq \mathcal{Q}$ be a linearly independent collection that is inclusion maximal with respect to $\mathcal{Q}$ (so its zeta rows span the zeta rows of $\mathcal{Q}$). For each set $q \in \mathcal{Q} - \mathcal{Q}'$ there exists a (unique) vector $\mu^{\mathcal{Q}'}(q) \in M^{\mathcal{Q}'}(q)$. Moreover, where $\tilde{\chi} \in R^{\mathcal{Q}}$, and $\tilde{\chi}'$ is the projection of $\tilde{\chi}$ on its $\mathcal{Q}'$ coordinates, then $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent iff*

$$(\mu^{\mathcal{Q}'}(q))^T \tilde{\chi}' = \tilde{\chi}_q, \ \forall q \in \mathcal{Q} - \mathcal{Q}', \tag{3.184}$$

and $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, and $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent. More generally, even if $\mathcal{Q}'$ is not linearly independent, so long as its zeta rows span the zeta rows of $\mathcal{Q}$ then $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent iff $\tilde{\chi}'$ is $\mathcal{P}$-signed-measure consistent and

$$(\mu^{\mathcal{Q}'}(q))^T \tilde{\chi}' = \tilde{\chi}_q, \ \forall q \in \mathcal{Q} - \mathcal{Q}'. \tag{3.185}$$

As above, $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, and $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent. □

Remark 3.43 *The μ vectors could also be adapted for the situation where the set function χ on $\mathcal{P}$ is represented as a vector $\chi' \in R^{\mathcal{A}}$ with coordinates corresponding to the sets in $\mathcal{A}$. Let $G \subseteq \mathcal{P}$ be a spanning collection for $\mathcal{P}$; let $G' \subseteq \mathcal{A}$ be such that $\{t' \cap P : t' \in G'\} = G$; let $\tilde{\chi}$ be the projection of χ on R^G, and let $\tilde{\chi}'$ be the projection of χ' on $R^{G'}$. The vector $\mu^G(q)$ is a list of multipliers corresponding to the sets in $G \subseteq \mathcal{P}$. Thus if χ is to be represented w.r.t. all of the sets in $\mathcal{A}$ then for each set $t \in \mathcal{P}$, the function value $\chi(t)$ will appear in all* *coordinates* $t' : t' \cap P = t$*, so any vector* $\bar{\mu}^{G'}(q) \in R^{G'}$ *for which*

$$\sum_{t' \in G' : t' \cap P = t} \bar{\mu}^{G'}(q)_{t'} = \mu^G(q)_t, \ \forall t \in G \tag{3.186}$$

will satisfy that

$$(\bar{\mu}^{G'}(q))^T \tilde{\chi}' = (\mu^G(q))^T \tilde{\chi} = \chi_q = \chi'_q = \chi'_{q'} \ \forall q' \in \mathcal{A} \ s.t. \ q' \cap P = q. \ \square \tag{3.187}$$

Thus for any (signed) measure, there is a single recipe for calculating the (signed) measure of every set in $\mathcal{P}$ so long as we know the (signed) measures of the sets in a linearly independent spanning collection. By putting together Lemmas 3.39 and 3.26 we therefore obtain the following characterization of $\mathcal{P}$-measure consistency for vectors defined with respect to spanning subcollections of $\mathcal{A}$.

Lemma 3.44 *Let* $G \subseteq \mathcal{P}$ *be a spanning collection for* $\mathcal{P}$*, and let* $G' \subseteq G$ *be a linearly independent spanning collection. Let* $\tilde{\chi} \in R^G$*, and let* $\tilde{\chi}'$ *be the subvector of* $\tilde{\chi}$ *with coordinates for each set in* G'*. Then* $\tilde{\chi}$ *is* $\mathcal{P}$*-measure consistent iff*

1. $\tilde{\chi}_q = (\mu^{G'}(q))^T \tilde{\chi}', \ \forall q \in G - G'$
2. $(\mu^{G'}(r))^T \tilde{\chi}' \geq 0, \ \forall r \in \mathcal{S}^P$

Recasting now for vectors expressed w.r.t. subcollections of $\mathcal{A}$*, let* $G \subseteq \mathcal{A}$ *be a spanning collection for* $\mathcal{A}$ *and let* $G' \subseteq G$ *be a linearly independent spanning collection. Let* $\tilde{\chi} \in R^G$*, and let* $\tilde{\chi}'$ *be the subvector of* $\tilde{\chi}$ *with coordinates for each set in* G'*. Then* $\tilde{\chi}$ *is* $\mathcal{P}$*-measure consistent iff*

1. $\tilde{\chi}_q = (\mu^{G'}(q))^T \tilde{\chi}', \ \forall q \in G - G'$
2. $(\mu^{G'}(r))^T \tilde{\chi}' \geq 0, \ \forall r \in \mathcal{S}$
3. $(\mu^{G'}(r))^T \tilde{\chi}' = 0, \ \forall r \in \mathcal{S} - \mathcal{S}^P \ \square$

We have seen already that the inner product of $\mu^G(q)$ with any column of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ is 0 or 1. More generally,

Lemma 3.45 *Given a collection* $G \subseteq \mathcal{P}$*, and denoting the subcolumns of the zeta columns* ζ^r *corresponding to* G *as* $\zeta^r\{G\}$*, a vector* $\mu \in R^G$ *satisfies*

$$\mu^T \zeta^r\{G\} \in \{0, 1\} \ \forall r \in \mathcal{S}^P \ \textit{iff} \ \ \mu = \mu^G(q) \tag{3.188}$$

for some $q \in \mathcal{P}$ *and for some* $\mu^G(q) \in M^G(q)$*.*

Proof: We have already shown sufficiency. Assume now that

$$\mu^T \zeta^r\{G\} \in \{0,1\}\ \forall r \in \mathcal{S}^P. \tag{3.189}$$

Let q be the set that is comprised of exactly those $r \in \mathcal{S}^P$ such that $\mu^T \zeta^r\{G\} = 1$. Then

$$\mu^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} \Rightarrow \tag{3.190}$$

$$\mu \in M^G(q).\ \square \tag{3.191}$$

Definition 3.46 *We will denote the vectors of the form $\mu^G(q)$ as "delta vectors" as these are the vectors whose inner product with the $\{\zeta^r\{G\} : r \in \mathcal{S}^P\}$ is always zero or one.*

Note that the delta vectors generalize the idempotents of $\bigvee$ described at the end of Chapter 2.

The following lemma shows that a type of additivity holds for the delta vectors.

Lemma 3.47 *Let $u, v \in \mathcal{P}$ be disjoint, then for any collection $G \subseteq \mathcal{P}$, for every $\mu^G(u) \in M^G(u)$ and every $\mu^G(v) \in M^G(v)$ there exists $\mu^G(u \cup v) \in M^G(u \cup v)$ such that*

$$\mu^G(u \cup v) = \mu^G(u) + \mu^G(v) \tag{3.192}$$

and conversely.

Proof: For every $r \in \mathcal{S}^P$, by Lemma 3.14

$$\zeta^r_{u \cup v} = \zeta^r_u + \zeta^r_v \tag{3.193}$$

and so

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{u \cup v\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{u\} + \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{v\} \Rightarrow \tag{3.194}$$

$$(\mu^G(u))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} + (\mu^G(v))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{u \cup v\} \Rightarrow \tag{3.195}$$

$$\mu^G(u) + \mu^G(v) \in M^G(u \cup v). \tag{3.196}$$

Conversely given $\mu^G(u \cup v) \in M^G(u \cup v)$, let $\mu^G(u)$ be any vector in $M^G(u)$, then it is easy to see by the same reasoning that

$$\mu^G(u \cup v) - \mu^G(u) \in M^G(v).\ \square \tag{3.197}$$

Corollary 3.48 *For any $q \in \mathcal{P}$, and any spanning collection G, for every $\mu^G(q) \in M^G(q)$, there exist $\{\mu^G(r) \in M^G(r) : r \in \mathcal{S}^P\}$ such that*

$$\mu^G(q) = \sum_{r \in \mathcal{S}^P : r \subseteq q} \mu^G(r) \tag{3.198}$$

and conversely. □

Observe that for $r \in \mathcal{S}^P$, where G is a linearly independent spanning collection,

$$\mu^G(r) = e_r^T (\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\})^{-1} \tag{3.199}$$

i.e. these are the rows of the inverse matrix to $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$.

We have seen already that the set $\mathcal{S}^P$ is a linearly independent spanning set. Here are some more examples of spanning sets.

Lemma 3.49 *Denote the sets $(Y_i^P)^c$ as N_i^P. The collections*

$$I^P = \left\{ \bigcap_{i \in V} Y_i^P : V \subseteq \{1, \ldots, n\} \right\} \tag{3.200}$$

$$I_N^P = \left\{ \bigcap_{i \in V} N_i^P : V \subseteq \{1, \ldots, n\} \right\} \tag{3.201}$$

$$U^P = \left\{ \bigcup_{i \in V \neq \emptyset} Y_i^P : V \subseteq \{1, \ldots, n\} \right\} \cup \{P\} \tag{3.202}$$

$$U_N^P = \left\{ \bigcup_{i \in V \neq \emptyset} N_i^P : V \subseteq \{1, \ldots, n\} \right\} \cup \{P\} \tag{3.203}$$

are all spanning, and are linearly independent if $\mathcal{P} = \mathcal{A}$. As usual, for the case $\mathcal{P} = \mathcal{A}$, these collections will be denoted without a superscript.

Note that P belongs to I^P and I_N^P as well (choose $V = \emptyset$.)

Proof: Given any set $q \subseteq P$, choose a representation

$$q = \Theta_q^P(Y_1^P, \ldots, Y_n^P). \tag{3.204}$$

Without loss of generality we can assume that this set theoretic expression entails no unions, since for any sets A and B

$$A \cup B = A^c \cap B^c \tag{3.205}$$

If the expression entails no complementation then $q \in I^P$. If it entails complementation, then consider the complementation that is performed last in evaluating the expression (according to some given order of operations). The complementation is performed on some expression A, and afterwards the only operation performed – if any further operations are indeed performed – is intersection. Thus the expression $\Theta_q^P(Y^P)$ is always either of the form

$$\Theta_q^P(Y^P) = A^c \tag{3.206}$$

or of the form

$$\Theta_q^P(Y^P) = A^c \cap B \tag{3.207}$$

for some expressions A and B (which themselves describe sets). By Lemma 3.14 we therefore have

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{A\} \tag{3.208}$$

in case (3.206), or

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{B\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{A \cap B\} \tag{3.209}$$

in case (3.207). For the latter case, both the expressions B and $A \cap B$ entail strictly fewer complementations than did $\Theta_q^P(Y^P) = A^c \cap B$. For the former case, $P \in I^P$ already, and A entails one fewer complementation than did $\Theta_q^P(Y^P) = A^c$. So we can repeat this procedure now for $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{B\}$ and for $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{A \cap B\}$ respectively (in the latter case, the former case is similar)), obtaining each of these as differences of rows of the form $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{A'\}$, again with strictly fewer complementations. Eventually we will reach a description of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ as a linear combination of rows $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{T\}$ where each T is an expression entailing no complementations, i.e. the set described by T belongs to I^P. (For example, consider

$$\Theta_q^P(Y^P) = \Big((Y_1^P Y_2^P)^c (Y_3^P)^c\Big)^c \tag{3.210}$$

where we have suppressed the intersection symbol $\cap$,

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{(Y_1^P Y_2^P)^c (Y_3^P)^c\} = \tag{3.211}$$

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} - \Big(\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{(Y_1^P Y_2^P)^c\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{(Y_1^P Y_2^P)^c Y_3^P\}\Big) = \tag{3.212}$$

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} - \Big((\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{Y_1^P Y_2^P\}) - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{(Y_1^P Y_2^P)^c Y_3^P\}\Big) = \tag{3.213}$$

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} - \Big((\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{Y_1^P Y_2^P\}) - (\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{Y_3^P\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{Y_1^P Y_2^P Y_3^P\})\Big) \tag{3.214}$$

which is a linear combination of rows corresponding to sets from I^P.) As for I_N^P, it is clear that we can recast any set theoretic expression in terms of Y_i^P to be in terms of N_i^P, and then we could follow the same procedure. For the unions, U^P, we could have started with an expression entailing no intersections, of the form

$$\Theta_q^P(Y^P) = A^c \cup B \tag{3.215}$$

and made use of the identity (also from Lemma 3.14)

$$\zeta^r_{A^c \cup B} = \zeta^r_P - \zeta^r_A + \zeta^r_B - \zeta^r_{A^c \cap B} = \tag{3.216}$$

$$\zeta^r_P - \zeta^r_A + \zeta^r_B - \zeta^r_B + \zeta^r_{A \cap B} = \tag{3.217}$$

$$\zeta^r_P - \zeta^r_A + \zeta^r_A + \zeta^r_B - \zeta^r_{A \cup B} = \tag{3.218}$$

$$\zeta^r_P + \zeta^r_B - \zeta^r_{A \cup B} \Rightarrow \tag{3.219}$$

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{P\} + \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{B\} - \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{A \cup B\} \tag{3.220}$$

repeatedly as above. Again, we could do something similar for U_N^P. Thus these collections are all spanning. For the case $P = \{0,1\}^n$ they are all also of size 2^n, which is $|\mathcal{S}|$ (the number of columns in $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}$), so in that case they are also linearly independent. □

Observe that we could also mix and match these collections to form different spanning collections. For example the collection (where all of the subscripts should be understood to range from 1 to n)

$$\{P, N_i^P, Y_i^P \cup Y_j^P, Y_i^P \cap Y_j^P \cap Y_k^P, N_i^P \cup N_j^P \cup N_k^P \cup, N_l^P, \ldots\} \tag{3.221}$$

is spanning.

Remark 3.50 *Observe that for the case $P = \{0,1\}^n$, the submatrix $\mathcal{Z}^{\mathcal{S}}(I)$ is exactly the zeta matrix Z for the lattice L described by Lovász and Schrijver and discussed in Chapters 1 and 2 (Definition 1.17). The delta vectors $\mu^I(r) : r \in \mathcal{S}$ are therefore the rows of the Möbius matrix, and the vectors $\mu^I(q) : q \in \mathcal{A}$ are the idempotents of $\bigvee$ described at the end of Chapter 2. From Lemma 2.23 we can also see that the vectors $m^{[u,v]}$ of Definition 1.50 are the vectors*

$$\mu^I(\bigcap_{i:s_i \in u} Y_i \cap \bigcap_{i:s_i \in v-u} N_i). \tag{3.222}$$

Observe also from the technique described in the proof of Lemma 3.49 that the vectors $m^{[u,v]}$ can have nonzeroes only in positions corresponding to sets of the form

$$\bigcap_{i \in W} Y_i,\ W \subseteq \{1,\ldots,n\},\ |W| \leq |v|. \tag{3.223}$$

Notice also that as I is linearly independent, every vector $\tilde{\chi} \in R^I$ is signed measure consistent. □

The technique described in the proof of Lemma 3.49 is a constructive method for obtaining delta vectors for any $q \in \mathcal{P}$ with respect to those spanning collections, but where $P \neq \{0,1\}^n$, the delta vectors thus obtained for a given $q \in \mathcal{P}$ are not not in general unique, as the collections are not in general linearly independent. Linear independence is significant as Lemma 3.39 will not hold without it, and its importance can be seen from Corollary 3.40 and from Lemma 3.44 as well. We will now find linearly independent subcollections of these collections and we will show how to obtain delta vectors w.r.t. those linearly independent subcollections.

Definition 3.51 *Define the following collections:*

$$\bar{I}^P = \left\{ \bigcap_{i \in V} Y_i^P : V \subseteq \{1,\dots,n\} \text{ such that } \exists y \in P \text{ for which } y_i = 1 \text{ iff } i \in V \right\} \tag{3.224}$$

$$\bar{I}_N^P = \left\{ \bigcap_{i \in V} N_i^P : V \subseteq \{1,\dots,n\} \text{ such that } \exists y \in P \text{ for which } y_i = 0 \text{ iff } i \in V \right\} \tag{3.225}$$

$$\bar{U}^P = \left\{ \bigcup_{i \in V} Y_i^P : V \subseteq \{1,\dots,n\} \text{ such that } \exists y \in P \text{ for which } y_i = 0 \text{ iff } i \in V \right\} \tag{3.226}$$

$$\bar{U}_N^P = \left\{ \bigcup_{i \in V} N_i^P : V \subseteq \{1,\dots,n\} \text{ such that } \exists y \in P \text{ for which } y_i = 1 \text{ iff } i \in V \right\} \tag{3.227}$$

The empty intersection, corresponding to $V = \emptyset$, should be understood to be the universal set P. Abusing convention, we will also say that the empty unions (for $\bar{U}^P$ and $\bar{U}_N^P$), corresponding to $V = \emptyset$, should be understood to be the universal set P.

Lemma 3.52

$$|\bar{I}^P| = |P| \tag{3.228}$$

Proof: Clearly the eligible V's are in one to one correspondence with the points of P. So all we need to show is that distinct V's yield distinct sets. Consider

$$q_V = \bigcap_{i \in V} Y_i^P \text{ and } \exists y \in P : y_i = 1 \text{ iff } i \in V \quad (\text{so } y \in q_V) \text{ and} \tag{3.229}$$

$$q_{V'} = \bigcap_{i \in V'} Y_i^P \text{ and } \exists y' \in P : y'_i = 1 \text{ iff } i \in V' \quad (\text{so } y' \in q_{V'}) \tag{3.230}$$

Suppose $y \in q^{V'}$ then it must be that $V \supseteq V'$. Similarly if $y' \in q_V$ then it must be that $V' \supseteq V$. Thus both sets could only share y and y' if $V = V'$. $\square$

So the collection $\bar{I}^P$ is the right size, all we need to show now is that it spans, after which we will be able to conclude that it is linearly independent (since the submatrix $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\bar{I}^P\}$

that it describes is square and of full column rank). Consider a set $q \in I^P - \bar{I}^P$. Such a set is an intersection of sets $Y_i^P : i \in V$ for some $V \subseteq \{1, \ldots, n\}$, but there is no point in that set that has a 1 only in its V positions and zeroes elsewhere. Thus every point in q has a 1 in some other non-V position, i.e. every point in q belongs to the set

$$\bigcup_{j \notin V} Y_j^P \tag{3.231}$$

and thus we have

$$q = \bigcap_{i \in V} Y_i^P \cap \left(\bigcup_{j \notin V} Y_j^P \right) = \bigcup_{j \notin V} \left(\bigcap_{i \in V} Y_i^P \cap Y_j^P \right). \tag{3.232}$$

Therefore by elementary measure theory (each coordinate of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ is just ζ_q^r for some $r \in \mathcal{S}^P$, and these are all measures),

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} = \sum_{j \notin V} \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P} \left\{ \bigcap_{i \in V} Y_i^P \cap Y_j^P \right\} - \tag{3.233}$$

$$\sum_{j_1, j_2 \notin V} \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P} \left\{ \bigcap_{i \in V} Y_i^P \cap Y_{j_1}^P \cap Y_{j_2}^P \right\} + \cdots - \cdots \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P} \left\{ \bigcap_{i=1}^{n} Y_i^P \right\}. \tag{3.234}$$

Thus we see that $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ can be written as a linear combination of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q'\}$ where the sets q' all belong to I^P and can all be written as intersections of strictly more that $|V|$ sets Y_i^P. Suppose that for one of the elements $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q'\}$ in this linear combination, q' also does not belong to $\bar{I}^P$, then by the same reasoning we could rewrite $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q'\}$ as a linear combination of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q''\}$ such that all $q'' \in I^P$, and all can be written as intersections of strictly more than $|V| + 1$ sets Y_i^P. Clearly this cannot repeat more than n times, so eventually we must conclude with a description of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ as a linear combination of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\bar{q}\}$ where all $\bar{q} \in \bar{I}^P$. Thus for any $u \in \mathcal{P}$ we could obtain $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{u\}$ as a linear combination of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\bar{q}\}$ where all $\bar{q} \in \bar{I}^P$ by first obtaining it as a linear combination of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ where all $\bar{q} \in I^P$, and then obtaining each $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ as a linear combination of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\bar{q}\}$ where all $\bar{q} \in \bar{I}^P$ as above. This proves the following theorem with regard to $\bar{I}^P$. The statements about the other collections follow from similar arguments.

Theorem 3.53 *$\bar{I}^P$, $\bar{I}_N^P$, $\bar{U}^P$ and $\bar{U}_N^P$ are all linearly independent spanning collections.* □

3.5 The Vectors $\nu^G(q)$

We have seen that the vectors $\mu^G(q)$ (where $G \subseteq \mathcal{P}$) describe the set q in terms of the sets in G in the sense that they describe linear combinations of the lists of points in each set of

G (i.e. the rows $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{g\} : g \in G$) that yield the list of points in q (i.e. the row $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$). This notion can be expanded to include linear combinations of lists of points in the sets of G that "overestimate" the list of points in q, i.e. they yield a list of points that may include points not in q, and that may count some points more than once. We will denote these vectors as $\nu^G(q)$, and we will define

$$N^G(q) = \{\nu^G(q) \in R^G : (\nu^G(q))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} \geq \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}\}. \tag{3.235}$$

Lemma 3.54 *Given $G \subseteq P$, a vector $\nu \in R^G$ is such that*

$$\nu^T \tilde{\chi} \geq \chi_q \tag{3.236}$$

for every $\mathcal{P}$-measure χ on $\mathcal{P}$ with subvector $\tilde{\chi}$ corresponding to the sets of G iff

$$\nu \in N^G(q). \tag{3.237}$$

Proof: If $\nu^T \tilde{\chi} \geq \chi_q$ for every $\mathcal{P}$-measure, then in particular this is true for every $\zeta^r : r \in \mathcal{S}^P$, and thus

$$\nu^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} \geq \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\} \tag{3.238}$$

and so $\nu \in N^G(q)$. Conversely if $\nu \in N^G(q)$, then for any $\mathcal{P}$-measure χ, there exists $\alpha \geq 0$ such that

$$\chi = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\alpha,\ \tilde{\chi} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\alpha \Rightarrow \tag{3.239}$$

$$\nu^T \tilde{\chi} = \nu^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\alpha \geq \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}\alpha = \chi_q. \ \square \tag{3.240}$$

The reason that we are interested in such vectors is that if the vector $\tilde{\chi} \in R^G$, where G is *not* a spanning set, then exact delta vectors for a set q may not exist. Specifically, if every linear combination describing $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{q\}$ in terms of other rows of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}$ entails a row for some set that does not belong to G, then there will be no delta vector $\mu^G(q)$. Vectors $\nu^G(q)$, however, are much easier to come by. The following theorem states that for any $G \subseteq \mathcal{P}$, the $\nu^G(q)$ vectors (the $\nu^G(r),\ r \in \mathcal{S}^P$ vectors in particular) can be used to characterize $\mathcal{P}$-measure consistency for vectors in R^G. The theorem is a generalization of Lemma 3.44.

Theorem 3.55 *Given $G \subseteq \mathcal{P}$ with $G' \subseteq G$ linearly independent and inclusion maximal subject to that property, and $\tilde{\chi} \in R^G$ with projection $\tilde{\chi}'$ on $R^{G'}$, the vector $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff for each $q \in G' - G$,*

$$(\mu^{G'}(q))^T \tilde{\chi}' = \tilde{\chi}_q \tag{3.241}$$

and for each $r \in \mathcal{S}^P$ *such that* $M^{G'}(r) \neq \emptyset$,

$$(\mu^{G'}(r))^T \tilde{\chi}' \geq 0 \tag{3.242}$$

and for each $r \in \mathcal{S}^P$ *such that* $M^{G'}(r) = \emptyset$,

$$(\nu^{G'}(r))^T \tilde{\chi}' \geq 0, \quad \forall \nu^G(r) \in N^G(r) \tag{3.243}$$

i.e. $\tilde{\chi}'$ *must be consistent with* $\tilde{\chi}$*, and every function value* χ_r, $r \in \mathcal{S}^P$ *that can be determined from the coordinates of* $\tilde{\chi}'$ *alone must be nonnegative, and for each* $r \in \mathcal{S}^P$ *for which the function value* χ_r *cannot be determined from the coordinates of* $\tilde{\chi}'$ *alone, then every* overestimation *of the function value* χ_r *that can be calculated using only the coordinates of* $\tilde{\chi}'$ *must be nonnegative.*

Proof: The vector $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff it is $\mathcal{P}$-signed-measure consistent and $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent (Lemma 3.42). Condition (3.241) establishes $\mathcal{P}$-signed measure consistency, so it suffices to prove that $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent. The vector $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent iff it is a projection of a vector that belongs to the cone of the $\zeta^r : r \in \mathcal{S}^P$, i.e. iff it belongs to the cone of the projected vectors $\zeta^r\{G'\} : r \in \mathcal{S}^P$. Thus $\tilde{\chi}'$ is measure consistent iff

$$a^T \tilde{\chi}' \geq 0, \quad \forall a \in \Big(Cone(\{\zeta^r\{G'\} : r \in \mathcal{S}^P\})\Big)^* \tag{3.244}$$

where the * symbol connotes the polar cone. Clearly if $\tilde{\chi}'$ is consistent with a measure χ on $\mathcal{P}$ then for each $r \in \mathcal{S}^P : M^{G'}(r) \neq \emptyset$ we will have $0 \leq \chi_r = (\mu^{G'}(r))^T \tilde{\chi}'$. Observe further that every $\nu^{G'}(r)$ must belong to $(Cone(\{\zeta^r\{G'\} : r \in \mathcal{S}^P\}))^*$. Thus if $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent so that $\tilde{\chi}' \in Cone(\{\zeta^r\{G'\} : r \in \mathcal{S}^P\})$, then it is also clear that for every $\nu^{G'}(r)$ we must have

$$(\nu^{G'}(r))^T \tilde{\chi}' \geq 0. \tag{3.245}$$

Conversely suppose that (3.242) and (3.243) hold as per the theorem. Observe first that any $a \neq 0$ in $(Cone(\{\zeta^r\{G'\} : r \in \mathcal{S}^P\}))^*$ must satisfy

$$a^T \zeta^r\{G'\} \geq 0, \quad \forall r \in \mathcal{S}^P \Rightarrow \tag{3.246}$$

$$a^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G'\} \geq 0. \tag{3.247}$$

Note moreover that by the linear independence of G' we must also have

$$a^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G'\} \neq 0. \tag{3.248}$$

Suppose now that there is some $r \in \mathcal{S}^P$ for which $a^T\zeta^r\{G'\} > 0$ and for which $M^{G'}(r) = \emptyset$, then there exists a scalar $\alpha > 0$ such that

$$\alpha a^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G'\} \geq \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{r\} \tag{3.249}$$

so that $\alpha a \in N^{G'}(r)$, which implies by (3.243) that

$$\alpha a^T \tilde{\chi}' \geq 0 \Rightarrow a^T \tilde{\chi}' \geq 0. \tag{3.250}$$

Suppose on the other hand that for each $r \in \mathcal{S}^P$ for which $a^T\zeta^r\{G'\} > 0$ we have $M^{G'}(r) \neq \emptyset$. Let us call the vector $a^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G'\}$ by the name $\bar{a}$, and define the vector

$$\mu := \sum_{r \in \mathcal{S}^P : \bar{a}_r > 0} \bar{a}_r \mu^{G'}(r). \tag{3.251}$$

Then

$$\mu^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G'\} = \sum_{r \in \mathcal{S}^P : \bar{a}_r > 0} \bar{a}_r e_r = \bar{a} \tag{3.252}$$

(where e_r is the r'th unit vector) which implies, by the linear independence of the rows of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G'\}$, that $\mu = a$. But by (3.242) we now have

$$a^T \tilde{\chi}' = \mu^T \tilde{\chi}' = \sum_{r \in \mathcal{S}^P : \bar{a}_r > 0} \bar{a}_r (\mu^{G'}(r))^T \tilde{\chi}' \geq 0. \ \square \tag{3.253}$$

It may also be worth noting that for any $G \subseteq \mathcal{P}$,

$$N^G(\emptyset) = (Cone(\{\zeta^r\{G\} : r \in \mathcal{S}^P\}))^*. \tag{3.254}$$

Here is a simple example of a ν^G vector for the case

$$G = \{Y_1, \ldots, Y_n, Y_i \cap Y_j, (i, j = 1, \ldots, n)\} \tag{3.255}$$

where we represent $\chi(Y_i)$ as χ_i and $\chi(Y_i \cap Y_j)$ as $\chi_{i,j}$, etc.

$$\chi(Y_1 \cap (Y_2 \cup Y_3)^c) = \chi_1 - \chi_{1,2} - \chi_{1,3} + \chi_{1,2,3} \leq \tag{3.256}$$

$$\chi_1 - \chi_{1,2} - \chi_{1,3} + \chi_{2,3} \tag{3.257}$$

for all measures χ, yielding the valid inequality

$$\chi_1 - \chi_{1,2} - \chi_{1,3} + \chi_{2,3} \geq 0. \tag{3.258}$$

Here is a somewhat more complicated example:

$$\chi((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)^c) = \tag{3.259}$$

$$\chi_{1\cup 2} - \chi_{1\cup 2,3} - \chi_{1\cup 2,4} + \chi_{1\cup 2,3,4} = \tag{3.260}$$

$$\chi_1 + \chi_2 - \chi_{1,2} - \chi_{1,3} - \chi_{2,3} + \chi_{1,2,3} - \chi_{1,4} - \chi_{2,4} + \chi_{1,2,4} + \chi_{1\cup 2,3,4} \leq \tag{3.261}$$

$$\chi_1 + \chi_2 + \chi_{1,2} - \chi_{1,3} - \chi_{1,4} - \chi_{2,3} - \chi_{2,4} + \chi_{3,4} \tag{3.262}$$

for all measures χ, yielding the valid inequality

$$\chi_1 + \chi_2 + \chi_{1,2} - \chi_{1,3} - \chi_{1,4} - \chi_{2,3} - \chi_{2,4} + \chi_{3,4} \geq 0. \tag{3.263}$$

3.6 Measure Preserving Operators

3.6.1 Characterization

Given a candidate vector $\tilde{\chi} \in R^{\mathcal{Q}}$ for some $\mathcal{Q} \subseteq \mathcal{P}$, and given some necessary condition $\nu^T \tilde{\chi} \geq 0$ valid for all measure consistent vectors, we may be able to enhance the power of this condition in the following way. Suppose an operator T exists such that wherever $\tilde{\chi}$ is measure consistent then so is $T\tilde{\chi}$, then we could enforce the condition $\nu^T T \tilde{\chi} \geq 0$ as well.

Definition 3.56 *Let $G \subseteq \mathcal{P}$ be a linearly independent spanning collection, and say $\tilde{\chi} \in R^G$. A function $T : R^G \to R^G$ is said to be $\mathcal{P}$-measure preserving if for every $\mathcal{P}$-measure consistent $\tilde{\chi}$ (recall that a measure is uniquely determined by $\tilde{\chi}$), $T\tilde{\chi}$ is also $\mathcal{P}$-measure consistent.*

Throughout this section we will assume that G is a linearly independent spanning collection.

Lemma 3.57 *The linear operator $T : R^G \to R^G$ is $\mathcal{P}$-measure preserving iff*

$$T\zeta^r\{G\} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\lambda, \ \lambda \geq 0 \tag{3.264}$$

for all $r \in \mathcal{S}^P$.

Proof: This follows from linearity since every $\mathcal{P}$-measure χ is in the cone of the $\zeta^r\{G\}$, $r \in \mathcal{S}^P$. □

Recall that the matrix $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ is invertible, and its inverse is the matrix to be denoted $\mathcal{M}^G$ whose rows are the (unique) vectors $\mu^G(r), r \in \mathcal{S}^P$.

Lemma 3.58 *The dual of the cone generated by the vectors $\zeta^r\{G\}$, $r \in \mathcal{S}^P$ is the cone generated by the rows of $\mathcal{M}^G$.*

Proof: This is always true for bases: let

$$\tilde{\chi} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\lambda,\ \lambda \geq 0 \text{ then} \tag{3.265}$$

$$\mathcal{M}^G\tilde{\chi} = \lambda \geq 0. \tag{3.266}$$

Conversely,

$$\mathcal{M}^G\tilde{\chi} = \lambda \geq 0 \Rightarrow \tag{3.267}$$

$$\tilde{\chi} = (\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\mathcal{M}^G)\tilde{\chi} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}(\mathcal{M}^G\tilde{\chi}) = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}\lambda.\ \Box \tag{3.268}$$

Putting together the previous two lemmas we get

Theorem 3.59 *If we represent the linear operator $T : R^G \to R^G$ as an $R^G \times R^G$ matrix, then T is $\mathcal{P}$-measure preserving iff there exists a matrix $F \geq 0$ such that*

$$\mathcal{M}^G T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} = F \tag{3.269}$$

or equivalently, iff

$$T = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}F\mathcal{M}^G,\ F \geq 0. \tag{3.270}$$

The set of $\mathcal{P}$-measure preserving linear operators is thus a cone whose extreme "rays" are the matrices $\zeta^r\{G\}(\mu^G(s))^T$ where $r, s \in \mathcal{S}^P$.

Proof: The first expression says that every $T\zeta^r\{G\}, r \in \mathcal{S}^P$ must belong to the cone of $\mathcal{P}$-measure consistent vectors. The second comes from multiplying both sides of the first expression by $(\mathcal{M}^G)^{-1} = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ on the left and $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}^{-1} = \mathcal{M}^G$ on the right. This establishes that the set of $\mathcal{P}$-measure preserving operators is a cone, and any such operator (represented as a matrix) is a nonnegative linear combination of the matrices $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}E_{r,s}\mathcal{M}^G$ where $E_{r,s}$ is the matrix with a 1 in position r, s and 0 everywhere else. But $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}E_{r,s}$ is the matrix whose s'th column is $\zeta^r\{G\}$, and is zero everywhere else, and that matrix times $\mathcal{M}^G$ has as its u, v entry the expression $\zeta^r\{G\}_u\mu^G(s)_v$ where $\mu^G(s)$ is the s'th row of $\mathcal{M}^G$. So we conclude that

$$\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}E_{r,s}\mathcal{M}^G = \zeta^r\{G\}(\mu^G(s))^T \tag{3.271}$$

and these are the extreme rays. $\Box$

In an efficient implementation of these ideas we will have only defined $\tilde{\chi}$ on some small subset of $\mathcal{P}$, and thus the only matrices T that will be useful to us are those whose rows have nonzero entries corresponding only to sets upon which we have defined $\tilde{\chi}$ (or else we will not be able to calculate any of the terms of the product $T\tilde{\chi}$). It does not matter, however, if T has too many rows, as we can just ignore the rows that do not correspond to the sets upon which $\tilde{\chi}$ is defined.

Lemma 3.60 *Let $G' \subseteq G$ be the collection of sets upon which the subvector $\tilde{\chi}'$ of $\tilde{\chi}$ is defined, and let T' be a linear operator on $R^{G'}$. Then T' preserves $\mathcal{P}$-measure consistency iff there exists a $\mathcal{P}$-measure preserving operator T such that the submatrix of T defined by its G' rows and columns is exactly T', and such that the $G - G'$ entries of the G' rows of T are all zero.*

Proof: If the operator T' is $\mathcal{P}$-measure consistency preserving then for every $r \in \mathcal{S}^P$ we have

$$T'\zeta^r\{G'\} = \sum_{p \in \mathcal{S}^P} \lambda_p(r)\zeta^p\{G'\} \tag{3.272}$$

where the $\lambda_p(r)$ are nonnegative numbers. Consider now the $\mathcal{P}$-measure preserving linear operator

$$\hat{T} = \sum_{q \in \mathcal{S}^P} \sum_{p \in \mathcal{S}^P} \lambda_p(q)\zeta^p\{G\}(\mu^G(q))^T. \tag{3.273}$$

For each $r \in \mathcal{S}^P$ we therefore have

$$\hat{T}\zeta^r\{G\} = \sum_{p \in \mathcal{S}^P} \lambda_p(r)\zeta^p\{G\}. \tag{3.274}$$

Thus the G' coordinates of any $\hat{T}\zeta^r\{G\}$ match exactly to $T'\zeta^r\{G'\}$. Replace now the G' rows of $\hat{T}$ with T', filling in zeroes for all the extra column positions, and refer to the matrix thus formed as T. By construction, T' is the submatrix of T corresponding to the sets G'. As the other rows remain unchanged we conclude that for any $\zeta^r\{G\}$,

$$\hat{T}\zeta^r\{G\} = T\zeta^r\{G\} \tag{3.275}$$

and as the $\zeta^r\{G\}$ form a basis of R^G this implies that $T = \hat{T}$. Conversely if there exists a $\mathcal{P}$-measure preserving operator T satisfying the conditions of the lemma then for each $r \in \mathcal{S}^P$, there exist nonnegative numbers $\lambda_p(r)$ such that

$$T\zeta^r\{G\} = \sum_{p \in \mathcal{S}^P} \lambda_p(r)\zeta^p\{G\} \tag{3.276}$$

which implies (since the nonzero entries of the G' rows of T all belong to the G' columns)

$$T'\zeta^r\{G'\} = \sum_{p \in \mathcal{S}^P} \lambda_p(r)\zeta^p\{G'\} \tag{3.277}$$

which establishes that T' preserves $\mathcal{P}$-measure consistency. □

So the $\mathcal{P}$-measure consistency preserving linear operators on G' are the submatrices (corresponding to G') of those $\mathcal{P}$-measure preserving operators for which the nonzero entries of the G' rows all belong to the G' columns. One easy class of matrices of this type is as follows.

Lemma 3.61 *Suppose $\tilde{\chi}' \in R^{G'}$ is $\mathcal{P}$-measure consistent, and $y' \in R^{G'}$ is also $\mathcal{P}$-measure consistent. For any delta vector $\mu^G(q)$ with nonzeroes only in coordinates corresponding to sets from G', $y'(\mu^{G'}(q))^T\tilde{\chi}'$ is $\mathcal{P}$-measure consistent, where $\mu^{G'}(q)$ is the restriction of $\mu^G(q)$ to its G' coordinates.*

Proof: Any vector $\mu^G(q)$ is the sum of the delta vectors of the nonempty atoms that comprise q, and any $\mathcal{P}$-measure consistent y' is the restriction of some nonnegative linear combination y of vectors $\zeta^r\{G\}$ to G' coordinates. Thus $y\mu^G(q)^T$ is $\mathcal{P}$-measure preserving. Considering now that $\mu^G(q)$ has only zeroes in its non-G' coordinates, we conclude that the G' rows of $y(\mu^G(q))^T$ have all non-G' column entries equal to zero, and their G' column entries are just $y'(\mu^{G'}(q))^T$. □

These "easy" matrices though are actually too easy to be of use. The fact that they are $\mathcal{P}$-measure preserving is also a consequence of the fact that for any delta vector $\mu^{G'}(q)$ expressed solely in terms of the sets of G' we must always have (if $\tilde{\chi}'$ is to be measure consistent)

$$(\mu^{G'}(q))^T\tilde{\chi}' = \chi_q \text{ (for some measure } \chi) \ \geq 0 \tag{3.278}$$

so that $y'(\mu^{G'}(q))^T\tilde{\chi}'$ is just a nonnegative multiple of the known $\mathcal{P}$-measure consistent vector y'.

3.6.2 Partial Summation

A more interesting example of a $\mathcal{P}$-measure preserving operator is partial summation. Stated loosely, partial summation corresponds to the situation where the matrix F of Theorem 3.59 is the identity matrix, but missing some of its 1's. Specifically, let G be a linearly independent spanning collection, and let the r'th column of F be

$$F^r = (\zeta^r\{G\})^T\mu^G(q)e_r, \ \forall r \in \mathcal{S}^P \tag{3.279}$$

where q is some set in $\mathcal{P}$, and where e_r is the r'th unit vector. Clearly $F \geq 0$ term for term so that the operator

$$T = \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}F\mathcal{M}^G \tag{3.280}$$

defined by this choice of F is $\mathcal{P}$-measure preserving. Naturally we could also consider the more general case of weighted summation where F is a nonnegative diagonal matrix (these are nonnegative linear combinations of the matrices F described above), i.e.

$$F^r = (\zeta^r\{G\})^T \nu^G(q)e_r, \ \forall r \in \mathcal{S}^P \tag{3.281}$$

(or a positive multiple thereof). Now for any

$$\tilde{\chi} = \sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{G\}, \tag{3.282}$$

$$T\tilde{\chi} = \sum_{r\in\mathcal{S}^P} \alpha_r \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}F\mathcal{M}^G\zeta^r\{G\} = \tag{3.283}$$

$$\sum_{r\in\mathcal{S}^P} \alpha_r \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}F^r = \sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{G\}(\zeta^r\{G\})^T \nu^G(q). \tag{3.284}$$

If $\nu^G(q) = \mu^G(q)$ this is just

$$\sum_{r\in\mathcal{S}^P:r\subseteq q} \alpha_r \zeta^r\{G\}. \tag{3.285}$$

(For general $\nu^G(q)$ it is a (positive multiple of) an upper bound on (3.285) if $\tilde{\chi}$ is $\mathcal{P}$-measure consistent.) Note moreover that for arbitrary $\mathcal{Q} \subseteq \mathcal{P}$ we also have that

$$\sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\}(\zeta^r\{\mathcal{Q}\})^T \mu^{\mathcal{Q}}(q) = \sum_{r\in\mathcal{S}^P:r\subseteq q} \alpha_r \zeta^r\{\mathcal{Q}\} \tag{3.286}$$

(and similarly, replacing $\mu^{\mathcal{Q}}(q)$ by $\nu^{\mathcal{Q}}(q)$ will yield an upper bound on $\sum_{r\in\mathcal{S}^P:r\subseteq q} \alpha_r \zeta^r\{\mathcal{Q}\}$ if $\sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\}$ is $\mathcal{P}$-measure consistent). Thus wherever we can write the expression

$$\sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\}(\zeta^r\{\mathcal{Q}\})^T \nu^{\mathcal{Q}}(q) \tag{3.287}$$

we can calculate a $\mathcal{P}$-measure consistency preserving transformation (which yields, where $\nu^{\mathcal{Q}}(q) = \mu^{\mathcal{Q}}(q)$, the partial sum over the nonempty atoms that belong to q, and which yields in general an upper bound on that partial sum if the full sum is $\mathcal{P}$-measure consistent).

Definition 3.62 *Given a collection $\mathcal{Q} \subseteq \mathcal{P}$, a vector $\tilde{\chi} \in R^{\mathcal{Q}}$, and given a collection $\mathcal{Q}' \subseteq \mathcal{P}$ such that for every pair of not necessarily distinct sets $u, v \in \mathcal{Q}'$, the intersection satisfies $u \cap v \in \mathcal{Q}$, define the matrix $U^{\tilde{\chi}}$ to be the $|\mathcal{Q}'| \times |\mathcal{Q}'|$ matrix whose u, v entry is $\tilde{\chi}_{u\cap v}$.*

Note that though $U^{\tilde{\chi}}$ is a function of $\mathcal{Q}'$, we will not use any dependence notation so long as the dependence is clear. Note also that $\mathcal{Q}' \subseteq \mathcal{Q}$ since for each $u \in \mathcal{Q}'$, $u = u \cap u \in \mathcal{Q}$.

Lemma 3.63 *If $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, so that there exists α such that*

$$\tilde{\chi} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\} \tag{3.288}$$

then for any such α,

$$U^{\tilde{\chi}} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T. \tag{3.289}$$

Proof:

$$\left(\sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T \right)_{u,v} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}'\}_u(\zeta^r\{\mathcal{Q}'\}_v)^T = \tag{3.290}$$

$$\sum_{r \in \mathcal{S}^P : r \subseteq u, r \subseteq v} \alpha_r = \sum_{r \in \mathcal{S}^P : r \subseteq u \cap v} \alpha_r = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\}_{u \cap v} = \tilde{\chi}_{u \cap v} \ \square \tag{3.291}$$

We therefore conclude

Lemma 3.64 *Given $\mathcal{P}$-signed-measure consistent $\tilde{\chi} \in R^{\mathcal{Q}}$, and given a delta vector $\mu^{\mathcal{Q}'}(q)$ for a set $q \in \mathcal{P}$, then for any α for which*

$$\tilde{\chi} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\}, \tag{3.292}$$

the partial sum of the projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$ of $\tilde{\chi}$ over the nonempty atoms that belong to q is

$$U^{\tilde{\chi}} \mu^{\mathcal{Q}'}(q). \ \square \tag{3.293}$$

Observe that the partial sum of $\tilde{\chi}'$ over the nonempty atoms that belong to q is the contribution to $\tilde{\chi}'$ from those points in P that satisfy the $\mathcal{P}$-logical condition $\theta_q^P(y)$.

The following lemma is an easy generalization of Lemma 3.64 that will prove useful shortly.

Lemma 3.65 *Let $\mathcal{Q}'' \subseteq \mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$ be such that for every $u \in \mathcal{Q}''$ and $v \in \mathcal{Q}'$ we have $u \cap v \in \mathcal{Q}$. Let $\tilde{\chi} \in R^{\mathcal{Q}}$ be $\mathcal{P}$-signed-measure consistent, with projections $\tilde{\chi}' \in R^{\mathcal{Q}'}$ and $\tilde{\chi}'' \in R^{\mathcal{Q}''}$. Define the $|\mathcal{Q}''| \times |\mathcal{Q}'|$ matrix $\hat{U}^{\tilde{\chi}}$ with each u, v entry equal to $\tilde{\chi}_{u \cap v}$. Let $q \in \mathcal{P}$ be such that a delta vector $\mu^{\mathcal{Q}'}(q)$ exists. Then for any α for which*

$$\tilde{\chi} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\}, \tag{3.294}$$

the partial sum of *$\tilde{\chi}''$ over the nonempty atoms that belong to q is*

$$\hat{U}^{\tilde{\chi}} \mu^{\mathcal{Q}'}(q). \tag{3.295}$$

Proof: Let $G = \{g \in \mathcal{P} : g = u \cap v,\ u, v \in \mathcal{Q}'\} \cup \mathcal{Q}$, and let $\bar{\chi}$ be any lifting of $\tilde{\chi}$ to R^G. The matrix $\hat{U}^{\tilde{\chi}}$ is just the $\mathcal{Q}''$ rows of the $|\mathcal{Q}'| \times |\mathcal{Q}'|$ matrix $U^{\bar{x}}$ with u, v entry of $\bar{\chi}_{u \cap v}$ for each $u, v \in \mathcal{Q}'$. Thus by Lemma 3.64 the partial sum of $\tilde{\chi}'$ over the nonempty atoms that belong to q is $U^{\bar{\chi}}\mu^{\mathcal{Q}'}(q)$, which implies that the partial sum of $\tilde{\chi}''$ over the nonempty atoms that belong to q is $\hat{U}^{\tilde{\chi}}\mu^{\mathcal{Q}'}(q)$. □

Obviously we can also conclude that $U^{\tilde{\chi}}\nu^{\mathcal{Q}'}(q)$ is a $\mathcal{P}$-measure preserving operator. Thus the matrices $U^{\tilde{\chi}}$, the partial sum operations $U^{\tilde{\chi}}\mu^{\mathcal{Q}}(q)$, and their generalizations $U^{\tilde{\chi}}\nu^{\mathcal{Q}'}(q)$ all arise naturally as a special case of $\mathcal{P}$-measure preserving operators.

Another way to approach partial summation, without explicit reference to the matrix $U^{\tilde{\chi}}$, is as follows. (The following lemma is an extension of Claim 3.16.)

Lemma 3.66 *Suppose $\tilde{\chi} \in R^{\mathcal{Q}}$ is $\mathcal{P}$-signed-measure consistent. If $\tilde{\chi}^q$ is any possible partial sum vector over some $q \in \mathcal{P}$,*

$$\tilde{\chi}^q_u = \tilde{\chi}_{q \cap u}. \tag{3.296}$$

(If q and u do not belong to $\mathcal{Q}$, then this should be understood to mean that the equality must hold for any signed measure χ with which $\tilde{\chi}$ is consistent.)

Proof: For any representation

$$\tilde{\chi} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\} \tag{3.297}$$

the partial sum vector over the nonempty atoms in $q \in \mathcal{P}$ for the expanded vector χ of $\tilde{\chi}$,

$$\chi^q = \sum_{r \in \mathcal{S}^P : r \subseteq q} \alpha_r \zeta^r \tag{3.298}$$

satisfies that for any $u \in \mathcal{P}$,

$$\chi^q_u = \sum_{r \in \mathcal{S}^P : r \subseteq q} \alpha_r \zeta^r_u = \sum_{r \in \mathcal{S}^P : r \subseteq q \cap u} \alpha_r \zeta^r = \chi_{q \cap u}. \ \square \tag{3.299}$$

Corollary 3.67 *Given any constraint*

$$\sum_{u \in U \subseteq \mathcal{Q}} \nu_u \tilde{\chi}_u \geq 0 \tag{3.300}$$

valid for all $\mathcal{P}$-measure consistent vectors, then the constraint

$$\sum_{u \in U \subseteq \mathcal{Q}} \nu_u \tilde{\chi}_{u \cap q} \geq 0 \tag{3.301}$$

is also valid for all $\mathcal{P}$-measure consistent vectors, and this is that same constraint applied to the partial sum $\tilde{\chi}^q$.

Proof: If $\tilde{\chi}$ is $\mathcal{P}$-measure consistent then it is $\mathcal{P}$-signed-measure consistent, so by the lemma the partial sum $\tilde{\chi}^q$ is such that

$$\tilde{\chi}^q_u = \tilde{\chi}_{q \cap u}, \ \forall q, u \tag{3.302}$$

so that

$$\sum_{u \in U \subseteq \mathcal{Q}} \nu_u \tilde{\chi}_{u \cap q} = \nu^T \tilde{\chi}^q \tag{3.303}$$

and this must be nonnegative as partial summation is $\mathcal{P}$-measure preserving. □

In this sense enforcing valid inequalities for the vector $\tilde{\chi} = (\tilde{\chi}_P, \tilde{\chi}_{Y_1}, \ldots, \tilde{\chi}_{Y_n})$ on the partial sum of $\tilde{\chi}$ over some q is an enforcement of those inequalities on the variables $(\tilde{\chi}_q, \tilde{\chi}_{q \cap Y_1^P}, \ldots, \tilde{\chi}_{q \cap Y_n^P})$.

The treatment we have given to partial summation here generalizes that given in the previous chapter. In particular, we will now show how to characterize the $\bar{N}^k$ operator of the first and second chapters in terms of the algebra $\mathcal{A}$. All of the statements in the following remark follow from the definition of $\bar{N}^k$ (Lemma 1.60), from Remark 3.50, and from Lemma 3.42.

Remark 3.68 *Let $P \subseteq \{0,1\}^n$, let $K = \{y \in \{0,1\}^{n+1} : y_0 = 1, (y_1, \ldots, y_n) \in P\}$, and let*

$$\bar{K} \subseteq Cone\Big(\{y \in \{0,1\}^{n+1} : y_0 = 1\}\Big) \tag{3.304}$$

satisfy that

$$\bar{K} \cap \{0,1\}^{n+1} = \{0\} \cup K. \tag{3.305}$$

Rename the coordinates $0, 1, \ldots, n$ as $\{0,1\}^n, Y_1, \ldots, Y_n$. Let k be a positive integer. Consider the following subcollections of $\mathcal{A}$:

$$\mathcal{Q} = \left\{ \bigcap_{i \in V} Y_i : V \subseteq \{1, \ldots, n\}, \ |V| \leq k+1 \right\} \tag{3.306}$$

$$\mathcal{Q}' = \left\{ \bigcap_{i \in V} Y_i : V \subseteq \{1, \ldots, n\}, \ |V| \leq k \right\} \tag{3.307}$$

$$\mathcal{Q}'' = \{\{0,1\}^n, Y_1, \ldots, Y_n\} \tag{3.308}$$

Observe that these subcollections are all linearly independent, and that for each set q of the form

$$q = \bigcap_{i \in V} Y_i \cap \bigcap_{i \in W} N_i, \ V, W \subseteq \{1, \ldots, n\}, \ |V| + |W| \leq k \tag{3.309}$$

(recall that $N_i = Y_i^c$) there exists exactly one vector $\mu^{\mathcal{Q}'}(q)$. The set $\bar{N}^k(\bar{K})$ is the set of points $\tilde{\chi}'' \in R^{\mathcal{Q}''}$ that have a lifting $\tilde{\chi} \in R^{\mathcal{Q}}$ such that the partial sum vector $\hat{U}^{\tilde{\chi}}\mu^{\mathcal{Q}'}(q) \in \bar{K}$, (where $\hat{U}^{\tilde{\chi}}$ is as in Lemma 3.65) for all sets q of form (3.309). Equivalently, if we let

$$\bar{\mathcal{Q}} = \left\{ T \cap \bigcap_{i \in V} Y_i \cap \bigcap_{i \in W} N_i : T \in \mathcal{Q}'',\ V, W \subseteq \{1, \dots, n\},\ |V| + |W| \leq k \right\} \tag{3.310}$$

then $\bar{N}^k(\bar{K})$ is the set of points $\tilde{\chi}'' \in R^{\mathcal{Q}''}$ that have a lifting $\bar{\chi} \in R^{\bar{\mathcal{Q}}}$ that is $\mathcal{A}$-signed-measure consistent, and satisfies

$$(\tilde{\chi}'')^q \in \bar{K} \tag{3.311}$$

for all q of the form (3.309), where for each $u \in \mathcal{Q}''$, $((\tilde{\chi}'')^q)_u = \bar{\chi}_{u \cap q}$. Stated another way, it is the set of points $\tilde{\chi}'' \in R^{\mathcal{Q}''}$ that have a lifting to a signed measure χ on $\mathcal{A}$ such that $(\tilde{\chi}'')^q \in \bar{K}$ for all q of the form (3.309), where for each $u \in \mathcal{Q}''$, $((\tilde{\chi}'')^q)_u = \chi_{u \cap q}$ (i.e. $(\tilde{\chi}'')^q$ is the projection of the partial sum signed measure χ^q on the $\mathcal{Q}''$ coordinates).[2]

Observe also that to ensure $\mathcal{A}$-signed-measure consistency on $\bar{\chi}$, by Lemma 3.42, we need only enforce the equations that describe the coordinates $q \in \bar{\mathcal{Q}} - \mathcal{Q}$ as linear combinations of $\mathcal{Q}$ coordinates. This can be easily done by following the constructive procedure outlined in Lemma 3.49. □

3.6.3 Term For Term Multiplication

Another measure preserving operator to which we will call attention is as follows.

Lemma 3.69 *Let ζ^q be the column of the zeta matrix of $\mathcal{A}$ corresponding to the set $q \in \mathcal{A}$. Then for any $v \in \mathcal{A}$,*

$$\zeta_v^q = \prod_{r \in \mathcal{S}^P : r \subseteq q} \zeta_v^r \quad \square \tag{3.312}$$

Let $\mathcal{Q} \subseteq I$, where, as earlier,

$$I = \left\{ \bigcap_{i \in V} Y_i : V \subseteq \{1, \dots, n\} \right\} \tag{3.313}$$

so that any $q \in \mathcal{Q}$ can be written

$$q = \bigcap_{i \in V_q} Y_i \tag{3.314}$$

[2] To see that $\bar{N}^n(\bar{K}) = Cone(K)$, observe that if $k = n$, then the sets of the form (3.309) include all of the atoms of $\mathcal{A}$, and by Claim 3.16, for each atom $r \in \mathcal{S}$, $(\tilde{\chi}'')^r = \chi_r(\tilde{\zeta}^r)''$. The constraint $(\tilde{\chi}'')^r \in \bar{K}$ now implies that $\chi_r \geq 0$ (since $\zeta^r \geq 0$, $(\tilde{\zeta}^r)'' \neq 0$ and $\bar{K} \subseteq R_+^{n+1}$), and therefore by the definition of $\bar{K}$, $\chi_r > 0$ iff the point $(\tilde{\zeta}^r)'' \in \{0,1\}^{n+1}$ belongs to K, or equivalently, iff $r \in \mathcal{S}^P$. Since this holds for each $r \in \mathcal{S}$, by Lemma 3.44 we conclude that χ defines a $\mathcal{P}$-measure, and that therefore $\tilde{\chi}'' \in Cone(K)$.

for some $V_q \subseteq \{1, \ldots, n\}$. Recall that $N_i := Y_i^c$, and consider the mapping

$$f(q) = \bigcap_{i \in V_q} Y_i \cap \bigcap_{i \in (V_q)^c} N_i \tag{3.315}$$

i.e. we map each $q \in \mathcal{Q}$ into the atom that belongs to exactly the $Y_i, i \in V_q$. Similarly every $r \in \mathcal{S}$ can be written

$$r = \bigcap_{i \in W_r} Y_i \cap \bigcap_{i \in (W_r)^c} N_i \tag{3.316}$$

for some $W_r \subseteq \{1, \ldots, n\}$. Consider the mapping from $\mathcal{S}$ into $\mathcal{A}$ defined by

$$g(r) = \bigcap_{i \in (W_r)^c} N_i \tag{3.317}$$

and consider also the mapping $h : \mathcal{S} \times \mathcal{S} \to \mathcal{S}$, defined by

$$h(r_1, r_2) = \bigcap_{i \in W_{r_1} \cap W_{r_2}} Y_i \cap \bigcap_{i \in (W_{r_1})^c \cup (W_{r_2})^c} N_i \tag{3.318}$$

and observe that $g(h(r_1, r_2)) = g(r_1) \cap g(r_2)$.

Lemma 3.70 *Given $u \in \mathcal{Q} \subseteq I$, $r \in \mathcal{S}$,*

$$\zeta_u^r = \zeta_{g(r)}^{f(u)} \tag{3.319}$$

Proof:

$$\zeta_u^r = 1 \Leftrightarrow r \subseteq u \Leftrightarrow V_u \subseteq W_r \Leftrightarrow f(u) \subseteq g(r) \ \Box \tag{3.320}$$

Thus the u'th element of the r'th column $\zeta^r\{\mathcal{Q}\}$ is the $f(u)$'th element of the $g(r)$'th row. So the column $\zeta^r\{\mathcal{Q}\}$ can also be thought of as a row, and term for term products of rows correspond to intersections.

Corollary 3.71 *Given $u \in \mathcal{Q} \subseteq I$, $r_1, r_2 \in \mathcal{S}$,*

$$\zeta_u^{r_1 \cup r_2} = \zeta_{g(r_1) \cap g(r_2)}^{f(u)} = \zeta_{g(h(r_1, r_2))}^{f(u)} \tag{3.321}$$

Proof:

$$\zeta_u^{r_1 \cup r_2} = 1 \Leftrightarrow \zeta_u^{r_1} = 1 \text{ and } \zeta_u^{r_2} = 1 \Leftrightarrow \tag{3.322}$$

$$\zeta_{g(r_1)}^{f(u)} = 1 \text{ and } \zeta_{g(r_2)}^{f(u)} = 1 \Leftrightarrow \zeta_{g(r_1) \cap g(r_2)}^{f(u)} = 1 \ \Box \tag{3.323}$$

In general, by the same reasoning,

Corollary 3.72 *Given $u, v \in \mathcal{Q} \subseteq I^P$,*

$$\zeta_u^v = \zeta^{f(u)}_{\bigcap_{r \subseteq v} g(r)} \quad \Box \tag{3.324}$$

But applying Lemma 3.70 now yields

Lemma 3.73

$$\zeta^{r_1 \cup r_2}\{\mathcal{Q}\} = \zeta^{h(r_1,r_2)}\{\mathcal{Q}\} \tag{3.325}$$

Proof: For each $u \in \mathcal{Q}$,

$$\zeta_u^{r_1 \cup r_2} = \zeta^{f(u)}_{g(r_1) \cap g(r_2)} = \zeta_u^{h(r_1,r_2)}. \ \Box \tag{3.326}$$

Thus where we restrict ourselves to $\mathcal{Q} \subseteq I$ coordinates, the columns $\zeta^{r_1 \cup r_2}$ are themselves zeta columns for the atom $h(r_1, r_2)$ whose "yeses" are the intersections of the "yeses" of the atoms r_1 and r_2.

Definition 3.74 *Let l be a positive integer and let x and y be vectors in R^l. Define $x*y \in R^l$ to be the vector*

$$(x * y)_i = x_i y_i. \tag{3.327}$$

Lemma 3.75 *Let $\mathcal{Q} \subseteq I$, and let $\tilde{\chi}, \xi \in R^{\mathcal{Q}}$ be measure consistent vectors. Then $\tilde{\chi} * \xi$ is measure consistent also.*

Proof: By Lemmas 3.73 and 3.69, for any ζ^{r_1} and ζ^{r_2},

$$\zeta^{r_1}\{\mathcal{Q}\} * \zeta^{r_2}\{\mathcal{Q}\} = \zeta^{h(r_1,r_2)}\{\mathcal{Q}\}. \tag{3.328}$$

Observe that

$$(\alpha_1\zeta^{r_1}\{\mathcal{Q}\} + \alpha_2\zeta^{r_2}\{\mathcal{Q}\}) * (\alpha_3\zeta^{r_3}\{\mathcal{Q}\} + \alpha_4\zeta^{r_4}\{\mathcal{Q}\}) = \tag{3.329}$$

$$\alpha_1\alpha_3(\zeta^{r_1}\{\mathcal{Q}\} * \zeta^{r_3}\{\mathcal{Q}\}) + \alpha_1\alpha_4(\zeta^{r_1}\{\mathcal{Q}\} * \zeta^{r_4}\{\mathcal{Q}\}) + \tag{3.330}$$

$$\alpha_2\alpha_3(\zeta^{r_2}\{\mathcal{Q}\} * \zeta^{r_3}\{\mathcal{Q}\}) + \alpha_2\alpha_4(\zeta^{r_2}\{\mathcal{Q}\} * \zeta^{r_4}\{\mathcal{Q}\}) \tag{3.331}$$

so $*$ is a linear operator. Thus if $\tilde{\chi}$ and ξ are both of the form

$$\sum_{r \in \mathcal{S}} \alpha_r \zeta^r\{\mathcal{Q}\}, \ \alpha \geq 0 \tag{3.332}$$

then so is $\tilde{\chi} * \xi$. $\Box$

Where $\mathcal{Q} \subseteq I$, the operator $\zeta^{r_1}\{\mathcal{Q}\} * \zeta^{r_2}\{\mathcal{Q}\}$ maps into $\zeta^{h(r_1,r_2)}\{\mathcal{Q}\}$ where $h(r_1, r_2)$ is the atom whose corresponding point has a 1 in each i'th coordinate iff r_1 and r_2 both have 1's in their i'th coordinate. Stated another way, it is the atom whose yeses are the intersection of the yeses of r_1 and r_2. We will now describe a similar operator that where

$$\mathcal{Q}' \subseteq U = \left\{ \bigcup_{i \in V \neq \emptyset} Y_i : V \subseteq \{1, \ldots, n\} \right\} \cup \{\{0,1\}^n\} \tag{3.333}$$

maps $\zeta^{r_1}\{\mathcal{Q}'\}, \zeta^{r_2}\{\mathcal{Q}'\}$ into $\zeta^{h'(r_1,r_2)}\{\mathcal{Q}'\}$ where $h'(r_1, r_2)$ is the atom whose yeses are the union of the yeses of r_1 and r_2.

We will first consider the collection

$$U' := \left\{ \bigcup_{i \in V} Y_i : V \subseteq \{1, \ldots, n\} \right\} \tag{3.334}$$

as this case will be easier to analyze, and we will adapt the result for U shortly. Let $\mathcal{Q}' \subseteq U'$, so that any $q \in \mathcal{Q}'$ can be written

$$q = \bigcup_{i \in V_q} Y_i \tag{3.335}$$

for some (possibly empty) $V_q \subseteq \{1, \ldots, n\}$. Consider the mapping

$$f'(q) = \bigcap_{i \in V_q} Y_i \cap \bigcap_{i \in (V_q)^c} N_i \tag{3.336}$$

i.e. we map each $q \in \mathcal{Q}'$ into the atom that belongs to exactly the $Y_i, i \in V_q$. Similarly every $r \in \mathcal{S}$ can be written

$$r = \bigcap_{i \in W_r} Y_i \cap \bigcap_{i \in (W_r)^c} N_i. \tag{3.337}$$

Consider the mapping from $\mathcal{S}$ into U' defined by

$$g'(r) = \bigcup_{i \in W_r} Y_i \tag{3.338}$$

and consider also the mapping $h' : \mathcal{S} \times \mathcal{S} \to \mathcal{S}$, defined by

$$h'(r_1, r_2) = \bigcap_{i \in W_{r_1} \cup W_{r_2}} Y_i \cap \bigcap_{i \in (W_{r_1})^c \cap (W_{r_2})^c} N_i \tag{3.339}$$

so that $g'(h'(r_1, r_2)) = g'(r_1) \cup g'(r_2)$.

Lemma 3.76 *Given $u \in \mathcal{Q}' \subseteq U'$, $r \in \mathcal{S}$,*

$$\zeta_u^r = \zeta_{g'(r)}^{f'(u)} \tag{3.340}$$

Proof:

$$\zeta_u^r = 1 \Leftrightarrow r \subseteq u \Leftrightarrow V_u \cap W_r \neq \emptyset \Leftrightarrow f'(u) \subseteq g'(r) \ \Box \tag{3.341}$$

Thus the u'th element of the r'th column $\zeta^r\{\mathcal{Q}'\}$ is the $f'(u)$'th element of the $g'(r)$'th row. So the column $\zeta^r\{\mathcal{Q}'\}$ can also be thought of as a row, and as above, term for term products of rows correspond to intersections.

Corollary 3.77 *Given $u \in \mathcal{Q}' \subseteq U'$, $r_1, r_2 \in \mathcal{S}$,*

$$\zeta_u^{r_1 \cup r_2} = \zeta_{g'(r_1) \cap g'(r_2)}^{f'(u)} \tag{3.342}$$

Proof:

$$\zeta_u^{r_1 \cup r_2} = 1 \Leftrightarrow \zeta_u^{r_1} = 1 \text{ and } \zeta_u^{r_2} = 1 \Leftrightarrow \tag{3.343}$$

$$\zeta_{g(r_1)}^{f(u)} = 1 \text{ and } \zeta_{g(r_2)}^{f(u)} = 1 \Leftrightarrow \zeta_{g(r_1) \cap g(r_2)}^{f(u)} = 1 \ \Box \tag{3.344}$$

Applying Lemma 3.76 therefore yields

Lemma 3.78 *Let $r_1, r_2 \in \mathcal{S}$ and let $\mathcal{Q}' \subseteq U \cup U'$. Then*

$$\zeta^{r_1 \cup r_2}\{\mathcal{Q}'\} = \zeta^{r_1}\{\mathcal{Q}'\} + \zeta^{r_2}\{\mathcal{Q}'\} - \zeta^{h'(r_1, r_2)}\{\mathcal{Q}'\}. \tag{3.345}$$

Proof: Note first that $U \cup U' = U' \cup \{\{0,1\}^n\}$, so each $u \in \mathcal{Q}'$ is either a member of U' or else $u = \{0,1\}^n$. If $u = \{0,1\}^n$ then since every set is a subset of $\{0,1\}^n$ we have

$$\zeta_u^{r_1 \cup r_2} = 1 = 1 + 1 - 1 = \zeta_u^{r_1} + \zeta_u^{r_2} - \zeta_u^{h'(r_1, r_2)}. \tag{3.346}$$

Otherwise we have $u \in U'$ and therefore

$$\zeta_u^{r_1 \cup r_2} = \zeta_{g'(r_1) \cap g'(r_2)}^{f'(u)} = \zeta_{g'(r_1)}^{f'(u)} + \zeta_{g'(r_2)}^{f'(u)} - \zeta_{g'(r_1) \cup g'(r_2)}^{f'(u)} = \tag{3.347}$$

$$\zeta_u^{r_1}(\mathcal{Q}') + \zeta_u^{r_2}(\mathcal{Q}') - \zeta_u^{h'(r_1, r_2)}. \ \Box \tag{3.348}$$

Definition 3.79 *Let l be a positive integer and let x and y be vectors in R^l. Define $x \circ y \in R^l$ to be the vector*

$$(x \circ y)_i = x_i + y_i - x_i y_i. \tag{3.349}$$

Lemma 3.80 *Let $\mathcal{Q}' \subseteq U$. Let $\tilde{\chi}, \xi \in R^{\mathcal{Q}'}$ be probability measure consistent vectors (i.e. $\tilde{\chi}_{\{0,1\}^n} = \xi_{\{0,1\}^n} = 1$). Then $\tilde{\chi} \circ \xi$ is probability measure consistent also.*

Proof: By Lemmas 3.78 and 3.69,

$$\zeta^{r_1}\{\mathcal{Q}'\} \circ \zeta^{r_2}\{\mathcal{Q}'\} = \zeta^{h'(r_1,r_2)}\{\mathcal{Q}'\}. \tag{3.350}$$

Consider now

$$\tilde{\chi} = \sum_{r \in \mathcal{S}} \alpha_r \zeta^r\{\mathcal{Q}'\},\ \xi = \sum_{t \in \mathcal{S}} \beta_t \zeta^t\{\mathcal{Q}'\},\ \alpha, \beta \geq 0,\ \sum_{r \in \mathcal{S}} \alpha_r = \sum_{t \in \mathcal{S}} \beta_t = 1. \tag{3.351}$$

Then for every u (where we write ζ^r instead of $\zeta^r\{\mathcal{Q}'\}$ to simplify notation)

$$(\tilde{\chi} \circ \xi)_u = \sum_{r \in \mathcal{S}} \alpha_r \zeta^r_u + \sum_{t \in \mathcal{S}} \beta_t \zeta^t_u - \sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t \zeta^r_u \zeta^t_u = \tag{3.352}$$

$$\left(\sum_{t \in \mathcal{S}} \beta_t\right) \sum_{r \in \mathcal{S}} \alpha_r \zeta^r_u + \left(\sum_{r \in \mathcal{S}} \alpha_r\right) \sum_{t \in \mathcal{S}} \beta_t \zeta^t_u - \sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t \zeta^r_u \zeta^t_u = \tag{3.353}$$

$$\sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t \zeta^r_u + \sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t \zeta^t_u - \sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t \zeta^r_u \zeta^t_u = \tag{3.354}$$

$$\sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t (\zeta^r \circ \zeta^t)_u \Rightarrow \tag{3.355}$$

$$\tilde{\chi} \circ \xi = \sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t (\zeta^r \circ \zeta^t) \tag{3.356}$$

and $\alpha_r \beta_t \geq 0,\ \forall r, t,\ \sum_{r \in \mathcal{S}} \sum_{t \in \mathcal{S}} \alpha_r \beta_t = (\sum_{r \in \mathcal{S}} \alpha_r)(\sum_{t \in \mathcal{S}} \beta_t) = 1$, so this is indeed probability measure consistent. □

Chapter 4

Positive Semidefiniteness

Let $P \subseteq \{0,1\}^n$, and let $\mathcal{P}$ be the algebra of subsets of P (the algebra generated by $Y_1^P, \ldots, Y_n^P$). Recall from Definition 3.62 that where $\tilde{\chi} \in R^{\mathcal{Q}}$ for some $\mathcal{Q} \subseteq \mathcal{P}$, and $\mathcal{Q}' \subseteq \mathcal{Q}$ is such that for all $u, v \in \mathcal{Q}'$ we have $u \cap v \in \mathcal{Q}$, then the matrices $U^{\tilde{\chi}}$ are defined as the matrices with rows and columns indexed by the elements of $\mathcal{Q}'$, with each u, v entry equal to $\tilde{\chi}(u \cap v)$. The focus of this chapter will be on the measure theoretic relevance of the condition that the matrices of the form $U^{\tilde{\chi}}$ be positive semidefinite. We will see shortly that where the vector $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent (Definitions 3.23 and 3.25), positive semidefiniteness of $U^{\tilde{\chi}}$ is a relaxation of the condition that $\tilde{\chi}$ be consistent with a measure on the algebra $\mathcal{P}$. (Recall that to establish that a point $x \in [0,1]^n$ belongs to $Conv(P)$ we need to show that x is consistent with a probability measure on $\mathcal{P}$.) In the first two sections of this chapter we will try to quantify in measure theoretic terms the nature of this approximation.

The first section focuses on the inequalities that are implied by $U^{\tilde{\chi}} \succeq 0$. The first two subsections are devoted to characterizing the inequalities implied by positive semidefiniteness in terms of the delta and ν vectors defined in Sections 3.4 and 3.5. We will see the crucial role played by the condition of $\mathcal{P}$-signed-measure consistency, and in this context we will also note a generalization of the $TH(G)$ operator of [GLS81] and [LS91].

In particular we will see that if $\tilde{\chi}$ is consistent with any signed measure χ on $\mathcal{P}$, and $U^{\tilde{\chi}} \succeq 0$, then where we denote the projection of $\tilde{\chi}$ on $\mathcal{Q}'$ by $\tilde{\chi}'$, for any set $q \in \mathcal{P}$ whose measure can be described in terms of the coordinates of $\tilde{\chi}'$ (i.e. there exists a delta vector $\mu^{\mathcal{Q}'}(q)$ as per Definition 3.38) we will be guaranteed to have $\chi(q) \geq 0$. We will see that this result provides a window toward understanding the effect of positive semidefiniteness in approximating $\mathcal{P}$-measure consistency. We will also show that in the special case where the collection $\mathcal{Q}'$ is an inclusion maximal linearly independent subcollection (Definition 3.36) of $\mathcal{Q}$ then if $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent and $U^{\tilde{\chi}} \succeq 0$, we will be guaranteed that $\tilde{\chi}$ is actually $\mathcal{P}$-measure consistent. We will use this result to show that where a set $P \subseteq \{0,1\}^n$ is such that for every point $y \in P$, for each $i, j \in \{1, \ldots, n\}$, the product $y_i \times y_j$ is a linear function of y, then a single semidefinite constraint constitutes a necessary and sufficient condition for a point $x \in \mathit{Affine}(P)$ to belong to $\mathit{Conv}(P)$. More generally, though the condition that $\mathcal{Q}'$ be a maximal linearly independent subcollection of $\mathcal{Q}$ is quite restrictive, this result can still be useful in establishing that particular *subvectors*, at least, of the lifted vector $\tilde{\chi}$ are $\mathcal{P}$-measure consistent. We will use such a methodology to prove the main theorem of Section 6.6.

Subsection 4.1.3 presents an application to the stable set polytope. In that subsection and the next we will also describe several possible methodologies for establishing that an inequality is implied by positive semidefiniteness in concert with other constraints. That is, given $\tilde{\chi} \in R^{\mathcal{Q}}$ as above, with the projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$, we will describe conditions under which the constraint $U^{\tilde{\chi}} \succeq 0$ will guarantee that, where $v \in R^{\mathcal{Q}'}$, we will have $v^T \tilde{\chi}' \geq 0$. For example, we will show that if v can be written (after possibly padding with zeroes) as a sum of delta vectors $\mu^G(q_1), \ldots, \mu^G(q_t)$ for some $\mathcal{P} \supseteq G \supseteq \mathcal{Q}'$, and if for some signed measure χ on $\mathcal{P}$ consistent with $\tilde{\chi}$ either the sum of the signed measures of the pairwise intersections $\chi(q_1 \cap q_j)$ is nonpositive, or the sum of the signed measures of the pairwise unions $\chi(q_i \cup q_j)$ is nonnegative, then the condition $U^{\tilde{\chi}} \succeq 0$ is sufficient to guarantee that $v^T \tilde{\chi}' \geq 0$. Thus if there are constraints on $\tilde{\chi}$ that can guarantee that there is in fact a signed measure χ consistent with $\tilde{\chi}$ such that either $\sum \chi(q_i \cap q_j) \leq 0$ or $\sum \chi(q_i \cup q_j) \geq 0$, then the additional constraint $U^{\tilde{\chi}} \succeq 0$ is sufficient to guarantee that $v^T \tilde{\chi}' \geq 0$. In particular, we will show examples of collections of linear constraints on the vector $\tilde{\chi}$ that guarantee that $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, and that all $\mathcal{P}$-signed measures consistent with $\tilde{\chi}$ are such that either $\sum \chi(q_i \cap q_j) \leq 0$ or $\sum \chi(q_i \cup q_j) \geq 0$. In Subsection 4.1.3 this methodology will be applied to the stable set problem, and in the following subsection it will be formalized and generalized. In that subsection we will also generalize the characterization of [LS91] of some of the situations in which positive semidefiniteness works in concert with other constraints.

The fact that positive semidefiniteness, in the form of the N^+ operator, does not always strengthen the N operator has already been noted in the literature ([GT01], see also [CD01] and [CL01]), Subsection 4.1.5 addresses a reverse question: How much is accomplished by the N constraints that could not be accomplished by positive semidefiniteness and $\mathcal{P}$-signed-measure consistency alone?

Section 4.2 gives a measure theoretic interpretation and generalization of the methodology of Lasserre's algorithm ([Las01], [Lau01]) in terms of the ν vectors and measure preserving operators.

As we noted, the question of when N^+ strengthens N has already been treated in the literature. In the case of the N-type operators (as per Remark 3.68), there is no guarantee of $\mathcal{P}$-signed-measure consistency, but $\mathcal{A}$-signed-measure consistency is indeed guaranteed. (See Definition 3.25, and Remarks 3.28 and 3.50, and recall from Definition 3.2 and Remark 3.7 that $\mathcal{A}$ is the algebra generated by $Y_1, \ldots, Y_n$, which is the algebra of subsets of $\{0,1\}^n$.) Thus positive semidefiniteness in the context of the N operator is actually a relaxation of $\mathcal{A}$-measure consistency, rather than $\mathcal{P}$-measure consistency. While $\mathcal{A}$-measure consistency is not the same thing as $\mathcal{P}$-measure consistency, it is a necessary condition for $\mathcal{P}$-measure consistency (Remark 3.28), and as we saw in the examples of Section 3.3, it can prove very useful in establishing $\mathcal{P}$-measure consistency. In Section 4.3 we will address the question of when the condition of $\mathcal{A}$-measure consistency itself, which is far stronger than positive semidefiniteness, helps in the context of the N operator. We will show that in a number of the cases where it has already been established that positive semidefiniteness will not help, measure consistency (i.e. $\mathcal{A}$-measure consistency) will actually not help either.

4.1 Inequalities Implied By Positive Semidefiniteness

4.1.1 Delta and ν Vectors

Recall that given $\mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$, with $u, v \in \mathcal{Q}' \Rightarrow u \cap v \in \mathcal{Q}$, and given $\tilde{\chi} \in R^{\mathcal{Q}}$ with projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$, the matrix $U^{\tilde{\chi}}$ is defined to be the $|\mathcal{Q}'| \times |\mathcal{Q}'|$ matrix whose u, v entry is $\tilde{\chi}_{u \cap v}$. Recall also that where $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, then for any α satisfying

$$\tilde{\chi} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\} \tag{4.1}$$

we have

$$U^{\tilde{\chi}} = \sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T. \tag{4.2}$$

Considering that the (sole) additional condition for $\tilde{\chi}$ to be $\mathcal{P}$-measure consistent is that $\alpha \geq 0$ we can immediately observe the following necessary condition.

Lemma 4.1 *If $\tilde{\chi}$ is $\mathcal{P}$-measure consistent then $U^{\tilde{\chi}} \succeq 0$.*

Proof: The matrices $\zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T$ are positive semidefinite, and nonnegative combinations of positive semidefinite matrices are positive semidefinite. $\square$

Moreover,

Lemma 4.2 *Let $U \in R^{m\times m}$ be a matrix belonging to the linear span of some collection of matrices $\{x^j(x^j)^T : j = 1,\ldots,k,\ k \le m\}$ where the vectors $\{x^j \in R^m\}$ are linearly independent. Then U belongs to the* cone *of the matrices $\{x^j(x^j)^T\}$ iff $U \succeq 0$.*

Proof: If U is in the cone then $U \succeq 0$ because each $x^j(x^j)^T \succeq 0$, and conic combinations of positive semidefinite matrices are positive semidefinite. Conversely if U is not in the cone, then complete $x^1,\ldots,x^k$ to a basis by choosing vectors $x^{k+1},\ldots,x^m$,and let X be the nonsingular matrix whose j'th column is $x^j,\ j = 1,\ldots,m$. By assumption,

$$U = \sum_{i=1}^{k} \alpha_i x^j(x^j)^T, \quad \alpha_h < 0 \tag{4.3}$$

for some $h \in \{1,\ldots,k\}$. Therefore, where we denote the h'th row of X^{-1} as (the column vector) X_h^{-1},

$$(X_h^{-1})^T U X_h^{-1} = \sum_{i=1}^{k} \alpha_i (X_h^{-1})^T x^j(x^j)^T X_h^{-1} = \alpha_h < 0 \Rightarrow \tag{4.4}$$

$$U \not\succeq 0. \ \square \tag{4.5}$$

Corollary 4.3 *Given a $\mathcal{P}$-signed-measure-consistent vector $\tilde{\chi} \in R^{\mathcal{Q}}$, where $\mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$ and*

1. *$\mathcal{Q}'$ is a spanning collection for $\mathcal{P}$, and*

2. *$u, v \in \mathcal{Q}' \Rightarrow u \cap v \in \mathcal{Q}$*

then $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff $U^{\tilde{\chi}} \succeq 0$.

Proof: Let $\bar{\mathcal{Q}}' \subseteq \mathcal{Q}'$ be a linearly independent spanning collection, and let us denote the $\bar{\mathcal{Q}}' \times \bar{\mathcal{Q}}'$ matrix whose u, v entry is $\tilde{\chi}_{u\cap v}$ as $\bar{U}^{\tilde{\chi}}$. By assumption, there are α for which

$$\tilde{\chi} = \sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}\} \text{ and} \tag{4.6}$$

$$U^{\tilde{\chi}} = \sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T \tag{4.7}$$

$$\bar{U}^{\tilde{\chi}} = \sum_{r\in\mathcal{S}^P} \alpha_r \zeta^r\{\bar{\mathcal{Q}}'\}(\zeta^r\{\bar{\mathcal{Q}}'\})^T. \tag{4.8}$$

By definition of linearly independent spanning collections, the columns $\zeta^r\{\bar{\mathcal{Q}}'\}$ are linearly independent, thus by the lemma, $\alpha \ge 0$ iff $\bar{U}^{\tilde{\chi}} \succeq 0$. Moreover,

$$\bar{U}^{\tilde{\chi}} \succeq 0 \Rightarrow \alpha \ge 0 \Rightarrow U^{\tilde{\chi}} \succeq 0 \tag{4.9}$$

and conversely

$$U^{\tilde{\chi}} \succeq 0 \Rightarrow \bar{U}^{\tilde{\chi}} \succeq 0 \tag{4.10}$$

as $\bar{U}^{\tilde{\chi}}$ is a principal minor of $U^{\tilde{\chi}}$, so $U^{\tilde{\chi}}$ is positive semidefinite iff $\bar{U}^{\tilde{\chi}}$ is positive semidefinite. □

Recall that if the spanning collection $\mathcal{Q}'$ of $\mathcal{P}$ is linearly independent, every vector $\tilde{\chi}'$ defined on $\mathcal{Q}'$ is $\mathcal{P}$-signed-measure consistent, but the expanded vector $\tilde{\chi} \in R^{\mathcal{Q}}$ is only $\mathcal{P}$-signed-measure consistent if we impose the $\mathcal{P}$-signed-measure consistency constraints (defined by the delta vectors) described in Lemma 3.39 on $\mathcal{Q}-\mathcal{Q}'$. Observe, however, that for the case $P = \{0,1\}^n$, the set I of all intersections of sets Y_i is a linearly independent spanning collection, and it is closed under intersections. Thus we could choose $\mathcal{Q} = \mathcal{Q}' = I$ and then we are always guaranteed signed measure consistency, and we have measure consistency as well iff $U^{\tilde{\chi}} \succeq 0$.

In general, where $\mathcal{Q}'$ is not a spanning collection for $\mathcal{P}$ we can still say the following.

Lemma 4.4 *Let $\mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$ be such that*

$$u, v \in \mathcal{Q}' \Rightarrow u \cap v \in \mathcal{Q}. \tag{4.11}$$

Let $\tilde{\chi} \in R^{\mathcal{Q}}$ be $\mathcal{P}$-signed-measure consistent, and let the projection of $\tilde{\chi}$ on $\mathcal{Q}'$ be denoted as $\tilde{\chi}'$. For any delta vector $\mu^{\mathcal{Q}'}(q) \in R^{\mathcal{Q}'}$, for every signed measure χ with which $\tilde{\chi}$ is consistent,

$$(\mu^{\mathcal{Q}'}(q))^T U^{\tilde{\chi}} \mu^{\mathcal{Q}'}(q) = (\mu^{\mathcal{Q}'}(q))^T \tilde{\chi}' = \chi_q. \tag{4.12}$$

Proof:

$$(\mu^{\mathcal{Q}'}(q))^T U^{\tilde{\chi}} \mu^{\mathcal{Q}'}(q) = (\mu^{\mathcal{Q}'}(q))^T (\sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T) \mu^{\mathcal{Q}'}(q) = \tag{4.13}$$

$$\sum_{r \in \mathcal{S}^P} \alpha_r (\mu^{\mathcal{Q}'}(q))^T \zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T) \mu^{\mathcal{Q}'}(q) = \tag{4.14}$$

$$\sum_{r \in \mathcal{S}^P} \alpha_r ((\mu^{\mathcal{Q}'}(q))^T \zeta^r\{\mathcal{Q}'\})^2 = \sum_{r \in \mathcal{S}^P} \alpha_r (\mu^{\mathcal{Q}'}(q))^T \zeta^r\{\mathcal{Q}'\} = \tag{4.15}$$

$$(\mu^{\mathcal{Q}'}(q))^T \tilde{\chi}' = \chi_q \ \square \tag{4.16}$$

Corollary 4.5 *Under the conditions of Lemma 4.4,*

$$U^{\tilde{\chi}} \succeq 0 \Rightarrow \chi_q \geq 0 \tag{4.17}$$

and if $P \in \mathcal{Q}'$ (so that there exists a vector $\mu^{\mathcal{Q}'}(q^c)$) then we also have

$$U^{\tilde{\chi}} \succeq 0 \Rightarrow \chi_q \leq \chi_P. \tag{4.18}$$

Proof: The first statement is clear from Lemma 4.4. As for the second statement, if $P \in \mathcal{Q}'$, then where e_P is the P'th unit vector in $R^{\mathcal{Q}'}$, the vector $e_P - \mu^{\mathcal{Q}'}(q) \in R^{\mathcal{Q}'}$ is a delta vector for the set q^c (i.e. it belongs to $M^{\mathcal{Q}'}(q^c)$). (To see this note that for any (signed) measure y on $\mathcal{P}$, with projection $\tilde{y}'$ on $\mathcal{Q}'$, we have $(e_P - \mu^{\mathcal{Q}'}(q))^T \tilde{y}' = y(P) - y(q) = y(q^c)$. In particular this is true for the zeta vectors, and so $e_P - \mu^{\mathcal{Q}'}(q) \in M^{\mathcal{Q}'}(q^c)$.) Thus Lemma 4.4 implies that $0 \leq \chi_{q^c} = \chi_P - \chi_q$. $\Box$

Definition 4.6 *We will refer to inequalities of the form*

$$(\mu^{\mathcal{Q}'}(q))^T \tilde{\chi}' \geq 0 \tag{4.19}$$

as delta vector inequalities.

Thus positive semidefiniteness implies that if $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, then for any set q that can be described in terms of the sets in $\mathcal{Q}'$ every signed measure that is consistent with $\tilde{\chi}$ assigns nonnegative measure to q, or in other words, every delta vector inequality is satisfied. Thus if all of the nonempty atoms can be described in terms of $\mathcal{Q}'$ then by Lemma 3.24, $\tilde{\chi}$ must be $\mathcal{P}$-measure consistent. Thus it is easy to see that Corollary 4.5 generalizes Corollary 4.3. Observe that the number of sets that can be described in terms of even a small collection $\mathcal{Q}'$ can be very large.

We will give here another corollary, this time of the *proof* of the lemma, that can be useful.

Corollary 4.7 *Under the conditions of the lemma, if $u \cap v = \emptyset$ then*

$$(\mu^{\mathcal{Q}'}(u))^T U^{\tilde{\chi}} \mu^{\mathcal{Q}'}(v) = 0. \tag{4.20}$$

Proof:

$$(\mu^{\mathcal{Q}'}(u))^T U^{\tilde{\chi}} \mu^{\mathcal{Q}'}(v) = \sum_{r \in \mathcal{S}^P} \alpha_r (\mu^{\mathcal{Q}'}(q))^T \zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T) \mu^{\mathcal{Q}'}(q) = \tag{4.21}$$

$$\sum_{r \in \mathcal{S}^P : r \subseteq u, r \subseteq v} \alpha_r = 0 \; \Box \tag{4.22}$$

Example: Consider $P \subseteq \{0,1\}^n$ where P is the set of incidence vectors of the stable sets of a graph G with n nodes $(v_1, \ldots, v_n)$, and the edge set E. (Edges in E will be identified by the indices of the nodes they lie between. Thus the edge between nodes v_i and v_j will be denoted $\{i,j\}$.) Let C be a clique in G. Then no stable set can have a 1 in the coordinates corresponding to any two nodes from C. In set theoretic terms,

$$Y_i^P \cap Y_j^P = \emptyset, \; \forall i,j \text{ such that } v_i, v_j \in C \Rightarrow \tag{4.23}$$

$$\bigcup_{i:v_i\in C} Y_i^P \text{ is a disjoint union} \Rightarrow \tag{4.24}$$

$$\chi\left(\bigcup_{i:v_i\in C} Y_i^P\right) = \sum_{i:v_i\in C} \chi(Y_i^P) \text{ and therefore} \tag{4.25}$$

$$\chi\left(\left(\bigcup_{i:v_i\in C} Y_i^P\right)^c\right) = \chi_P - \sum_{i:v_i\in C} \chi(Y_i^P) \tag{4.26}$$

for any signed measure χ on $\mathcal{P}$. Thus where we set

$$\mathcal{Q}' = \{P, Y_1^P, \ldots, Y_n^P\},\ \mathcal{Q} = \{P, Y_1^P, \ldots, Y_n^P, Y_i^P \cap Y_j^P, (i,j = 1, \ldots, n)\} \tag{4.27}$$

and we denote the set $(\bigcup_{i:v_i\in C} Y_i^P)^c$ as C^u, then we can write

$$\mu^{\mathcal{Q}'}(C^u) = e_P - \sum_{i:v_i\in C} e_i \tag{4.28}$$

where e_i is the unit vector that corresponds to the set Y_i^P, and e_P corresponds to the set P. By the same token, denote $\tilde{\chi}'_{Y_i^P}$ as $\tilde{\chi}'_i$, and $\tilde{\chi}_{Y_i^P\cap Y_j^P}$ as $\tilde{\chi}_{i,j}$. Thus so long as $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent then the constraint $U^{\tilde{\chi}} \succeq 0$ will guarantee that every signed measure with which $\tilde{\chi}$ is consistent must assign nonnegative value to C^u for every clique, i.e.

$$0 \le \chi_{c^u} = (\mu^{\mathcal{Q}'}(C^u))^T \tilde{\chi}' = \chi_P - \sum_{i:v_i\in C} \tilde{\chi}'_i \tag{4.29}$$

i.e. all of the clique constraints will be satisfied (and there can be exponentially many of them). Observe now that for any i, j such that $\{i,j\} \notin E$, the set $\{v_i, v_j\}$ is a stable set, and therefore the point in $\{0,1\}^n$ with a 1 in positions i and j and zeroes elsewhere belongs to P. Thus where

$$\bar{I}^P = \left\{\bigcap_{i\in V} Y_i^P : V \subseteq \{1, \ldots, n\} \text{ s.t. } \exists y \in P \text{ for which } y_i = 1 \text{ iff } i \in V\right\} \tag{4.30}$$

we conclude that $Y_i^P \cap Y_j^P \in \bar{I}^P$. Recall now that by Theorem 3.53, the collection $\bar{I}^P$ is linearly independent, thus the collection

$$\left\{P, Y_1^P, \ldots, Y_n^P, Y_i^P \cap Y_j^P : \{i,j\} \notin E\right\} \tag{4.31}$$

is linearly independent. Moreover, for any $i, j : \{i,j\} \in E$, we have

$$Y_i^P \cap Y_j^P = \emptyset \tag{4.32}$$

so all of these expressions describe a single set, namely $\emptyset$, and the delta vector $\mu^{\mathcal{Q}}(\emptyset) = 0$ describes this set. By Lemma 3.42 we conclude,

Lemma 4.8 *For the given example, all vectors $\tilde{\chi}$ indexed by*

$$\mathcal{Q} = \left\{P, Y_i^P(i = 1, \ldots, n), Y_i^P \cap Y_j^P(\{i,j\} \notin E), \emptyset\right\} \tag{4.33}$$

that satisfy $\tilde{\chi}_\emptyset = 0$ are $\mathcal{P}$-signed-measure consistent. □

Corollary 4.9 *If $\tilde{\chi}_\emptyset = 0$ (so $U^{\tilde{\chi}}_{i,j} = \tilde{\chi}_{Y_i^P \cap Y_j^P} = \tilde{\chi}_\emptyset = 0$ for all $\{i,j\} \in E$) and $U^{\tilde{\chi}} \succeq 0$, then $\tilde{\chi}'$ satisfies all of the clique inequalities.* □

It should be noted that the set of vectors $\tilde{\chi}'$ that are projections of some $\tilde{\chi}$ with $\tilde{\chi}_\emptyset = 0$, and for which $U^{\tilde{\chi}} \succeq 0$ (with the additional constraint $\tilde{\chi}'_P = 1$) is the same as the set $TH(G)$ introduced in [GLS81] and [LS91] and it was already noted there that all vectors in this set satisfy the clique inequalities. Thus for general $P \subseteq \{0,1\}^n$, the conditions that $\tilde{\chi}$ be $\mathcal{P}$-signed-measure consistent and $U^{\tilde{\chi}} \succeq 0$, can be viewed as a generalization of the idea of $TH(G)$. □

Here is another theorem, more specialized than Lemma 4.4, that also allows us to say something (sometimes) for the case where Q' is not a spanning set.

Theorem 4.10 *Let $\mathcal{Q}'$ be an inclusion maximal linearly independent subcollection of $\mathcal{Q} \subseteq \mathcal{P}$, where $u, v \in \mathcal{Q}' \Rightarrow u \cap v \in \mathcal{Q}$, and let $\tilde{\chi}$ (with a coordinate for each $q \in \mathcal{Q}$) be $\mathcal{P}$-signed-measure consistent. Then $\tilde{\chi}$ is $\mathcal{P}$-measure consistent iff $U^{\tilde{\chi}} \succeq 0$.*

Proof: It is clear that if $\tilde{\chi}$ is $\mathcal{P}$-measure consistent then we must have $U^{\tilde{\chi}} \succeq 0$, so we only need to prove the converse. By the linear independence of $\mathcal{Q}'$, there must exist some square nonsingular submatrix W of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\mathcal{Q}'\}$ (defined in Definition 3.36). The columns of this submatrix are indexed by a $|\mathcal{Q}'|$ size subcollection of the atoms $\mathcal{S}^P$ (or alternatively, by a cardinality $|\mathcal{Q}'|$ subset of P). Let us refer to the union of these atoms (i.e. to that cardinality $|\mathcal{Q}'|$ subset of P) as V, and let us rename the rows corresponding to sets $u \in \mathcal{Q}'$ as $u^V = u \cap V$, and let us denote the collection of all such sets as $\mathcal{Q}'^V$. (Note that the sets $u \cap V,\ u \in \mathcal{Q}$ are all distinct, since suppose $\exists u, w \in \mathcal{Q}, u \neq w$ and $u \cap V = w \cap V$. So the u row of $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}$ is not the same as the w row, but on the subrow corresponding to the atoms in V they match. Thus $\mu^{\mathcal{Q}'}(u) \neq \mu^{\mathcal{Q}'}(w)$, but $(\mu^{\mathcal{Q}'}(u))^T W = (\mu^{\mathcal{Q}'}(w))^T W$, which contradicts the nonsingularity of W.) Then the square submatrix W is exactly the submatrix $\mathcal{Z}_{\mathcal{V}}^{\mathcal{S}^V}\{\mathcal{Q}'^V\}$, of the zeta matrix for the subset algebra $\mathcal{V}$ of V, and by linear independence and the fact that $|V| = |\mathcal{Q}'^V|$ we conclude that the collection $\mathcal{Q}'^V$ is a linearly independent spanning collection for $\mathcal{V}$. Define the collection $\mathcal{Q}^V$ of sets $u \cap V,\ u \in \mathcal{Q}$, and observe that we still have $u, v \in \mathcal{Q}'^V \Rightarrow u \cap v \in \mathcal{Q}^V$. Observe also that where the vector $\tilde{\chi}^V$ is indexed by $\mathcal{Q}^V$ with

$\tilde{\chi}^V(u \cap V) = \tilde{\chi}(u), u \in \mathcal{Q}$ (recall that distinct $u \in \mathcal{Q}$ make for distinct $u \cap V \in \mathcal{Q}^V$), then $\tilde{\chi}^V$ is $\mathcal{V}$-signed-measure consistent (since the delta vectors describing rows in $\mathcal{Q}$ as linear combinations of the rows in $\mathcal{Q}'$ are unaffected by the elimination of some of the columns from the zeta matrix, so Lemma 3.42 continues to apply). Thus where the matrix $U^{\tilde{\chi}^V}$ has its rows and columns indexed by $\mathcal{Q}'^V$, or equivalently by $\mathcal{Q}'$, Corollary 4.3 implies that $\tilde{\chi}^V$ is $\mathcal{V}$-measure consistent iff

$$U^{\tilde{\chi}} = U^{\tilde{\chi}^V} \succeq 0 \tag{4.34}$$

where the equality follows from the fact that for each $u, v \in \mathcal{Q}'$,

$$U^{\tilde{\chi}^V}(u,v) = \tilde{\chi}^V(u \cap v \cap V) = \tilde{\chi}(u \cap v) = U^{\tilde{\chi}}(u,v). \tag{4.35}$$

Let χ^V be a $\mathcal{V}$-measure with which $\tilde{\chi}^V$ is consistent, and define the $\mathcal{P}$-measure χ by $\chi(u) = \chi^V(u \cap V),\ \forall u \subseteq P$. Thus for each $u \in \mathcal{Q}$ we have

$$\chi(u) = \chi^V(u \cap V) = \tilde{\chi}^V(u \cap V) = \tilde{\chi}(u) \tag{4.36}$$

which proves that $\tilde{\chi}$ is indeed $\mathcal{P}$-measure consistent. □

Corollary 4.11 *Let $P \subseteq \{0,1\}^n$ be such that for all $y \in P$ and for all $\{i,j\} \in \{1,\ldots,n\}$, $i \neq j$, the product*

$$y_i \times y_j = (\alpha^{\{i,j\}})^T y + \gamma^{\{i,j\}} \tag{4.37}$$

for some $\alpha^{i,j} \in R^n$ and some real number $\gamma^{\{i,j\}}$. Define

$$\alpha^{\{i,i\}} = \alpha^{\{0,i\}} = e_i, \quad \gamma^{\{i,i\}} = \gamma^{\{0,i\}} = 0 \tag{4.38}$$

where e_i is the i'th unit vector. Then a point $x \in R^n$ belongs to Conv(P) iff $x \in$ Affine(P), and $U^x \succeq 0$, where U^x is the square matrix with rows and columns indexed by $\{0,1,\ldots,n\}$, with $U^x(0,0) = 1$, and for all $\{i,j\} \neq \{0,0\}$,

$$U^x(i,j) = (\alpha^{\{i,j\}})^T x + \gamma^{\{i,j\}}. \tag{4.39}$$

Proof: It is easy to see that these conditions are all necessary, so we will only prove sufficiency. The vector $(1,x)$, which may be represented as

$$(x[P], x[Y_1^P], \ldots, x[Y_n^P]) \tag{4.40}$$

is $\mathcal{P}$-signed-measure consistent iff $x \in$ *Affine*(P) by Corollary 3.19. The conditions guarantee moreover that the expanded vector with coordinates for each pairwise intersection as per U^x is $\mathcal{P}$-signed-measure consistent as well (by Lemma 3.42). Since the conditions of the corollary also guarantee that some subcollection of $P, Y_1^P, \ldots, Y_n^P$ is an inclusion maximal linearly independent subcollection of $\{P, Y_i^P (i = 1, \ldots, n), Y_i^P \cap Y_j^P (i, j = 1, \ldots, n)\}$ we can apply Theorem 4.10 to conclude that $(1, x)$ is $\mathcal{P}$-measure consistent, which proves the corollary. □.

This is only one of the ways that positive semidefiniteness can be effective. One way to appreciate the power of positive semidefiniteness in general is as follows. Let $\bar{\mathcal{Q}}'$ be an inclusion maximal linearly independent subcollection of $\mathcal{Q}' \subseteq \mathcal{P}$. Recall from Lemma 3.42 that a vector $\tilde{\chi}' \in R^{\mathcal{Q}'}$ is $\mathcal{P}$-measure consistent iff it is $\mathcal{P}$-signed-measure consistent, and its projection $\bar{\tilde{\chi}}' \in R^{\bar{\mathcal{Q}}'}$ is $\mathcal{P}$-measure consistent. Recall from Section 3.5 that the vector $\bar{\tilde{\chi}}'$ is $\mathcal{P}$-measure consistent iff

$$(\nu^{\bar{\mathcal{Q}}'}(\emptyset))^T \bar{\tilde{\chi}}' \geq 0, \ \forall \nu^{\bar{\mathcal{Q}}'}(\emptyset) \in N^{\bar{\mathcal{Q}}'}(\emptyset) \tag{4.41}$$

since $N^{\bar{\mathcal{Q}}'}(\emptyset)$ is the polar cone of $\{\zeta^r\{\bar{\mathcal{Q}}'\} : r \in \mathcal{S}^P\}$, i.e. the vectors $\nu^{\bar{\mathcal{Q}}'}(\emptyset)$ are those that satisfy

$$(\nu^{\bar{\mathcal{Q}}'}(\emptyset))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\bar{\mathcal{Q}}'\} \geq 0. \tag{4.42}$$

For each $\nu^{\bar{\mathcal{Q}}'} \in N^{\bar{\mathcal{Q}}'}(\emptyset)$, define the row vector

$$w^\nu = (\nu^{\bar{\mathcal{Q}}'})^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\bar{\mathcal{Q}}'\}. \tag{4.43}$$

Since $w^\nu \geq 0$, denoting the unit vector in $R^{\mathcal{S}^P}$ corresponding to each $r \in \mathcal{S}^P$ by e_r, there exists a unique $\lambda^\nu \in R_+^{\mathcal{S}^P}$ such that

$$w^\nu = \sum_{r \in \mathcal{S}^P} \lambda_r^\nu e_r^T. \tag{4.44}$$

Let G be a linearly independent spanning collection such that $\bar{\mathcal{Q}}' \subseteq G$. For each $\nu^{\bar{\mathcal{Q}}'} \in N^{\bar{\mathcal{Q}}'}(\emptyset)$, define the vector ν to be the expansion of $\nu^{\bar{\mathcal{Q}}'}$ to $|G|$ dimensions, with a value of zero in all of the appended coordinates. Now

$$\nu^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} = (\nu^{\bar{\mathcal{Q}}'})^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{\bar{\mathcal{Q}}'\} = w^\nu = \tag{4.45}$$

$$\sum_{r \in \mathcal{S}^P} \lambda_r^\nu e_r^T = \sum_{r \in \mathcal{S}^P} \lambda_r^\nu (\mu^G(r))^T \mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\} \Rightarrow \tag{4.46}$$

$$\nu = \sum_{r \in \mathcal{S}^P} \lambda_r^\nu \mu^G(r) \tag{4.47}$$

since $\mathcal{Z}_{\mathcal{P}}^{\mathcal{S}^P}\{G\}$ is nonsingular. Thus for every expansion of $\bar{\tilde{\chi}}'$ to $\hat{\chi} \in R^G$ determining the signed measure χ on $\mathcal{P}$,

$$(\nu^{\bar{\mathcal{Q}}'})^T \bar{\tilde{\chi}}' = \nu^T \hat{\chi} = \sum_{r \in \mathcal{S}^P} \lambda_r^\nu (\mu^G(r))^T \hat{\chi} = \sum_{r \in \mathcal{S}^P} \lambda_r^\nu \chi_r. \tag{4.48}$$

The implications of these observations are as follows.

For the purposes of the next two lemmas, let $\bar{\mathcal{Q}}' \subseteq \mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$ where $\bar{\mathcal{Q}}'$ is an inclusion maximal linearly independent subcollection of $\mathcal{Q}'$. For the second lemma, assume in addition that for every $u, v \in \mathcal{Q}'$ we have $u \cap v \in \mathcal{Q}$. Let $\tilde{\chi}$ be in $R^{\mathcal{Q}}$ with projections $\tilde{\chi}' \in R^{\mathcal{Q}'}$ and $\bar{\tilde{\chi}}' \in R^{\bar{\mathcal{Q}}'}$ on $\mathcal{Q}'$ and $\bar{\mathcal{Q}}'$ respectively, and let $\tilde{\chi}$ be $\mathcal{P}$-signed-measure consistent (so that its projections are as well).

Lemma 4.12 *The vector $\tilde{\chi}'$ is $\mathcal{P}$-measure consistent iff $\bar{\tilde{\chi}}'$ satisfies that for some (and every) signed measure χ consistent with $\bar{\tilde{\chi}}'$, and every $\nu^{\bar{\mathcal{Q}}'} \in N^{\bar{\mathcal{Q}}'}(\emptyset)$ and corresponding $\lambda^{\nu} \in R^{\mathcal{S}^P}$,*

$$\sum_{r \in \mathcal{S}^P} \lambda_r^{\nu} \chi_r \geq 0. \tag{4.49}$$

(Note that $\lambda^{\nu} \geq 0$.) If $\bar{\mathcal{Q}}'$ is an inclusion maximal linearly independent subcollection of $\mathcal{Q}$ as well, then $\tilde{\chi}$ will also be $\mathcal{P}$-measure consistent under this condition.

Lemma 4.13 *If $U^{\tilde{\chi}} \succeq 0$, then*

$$\sum_{r \in \mathcal{S}^P} (\lambda_r^{\nu})^2 \chi_r \geq 0 \tag{4.50}$$

for every signed measure χ consistent with $\bar{\tilde{\chi}}'$, and every $\nu^{\bar{\mathcal{Q}}'} \in N^{\bar{\mathcal{Q}}'}(\emptyset)$ and corresponding $\lambda^{\nu} \in R^{\mathcal{S}^P}$.

Proof: Denote the $|\bar{\mathcal{Q}}'| \times |\bar{\mathcal{Q}}'|$ matrix with u, v entry equal to $\tilde{\chi}_{u \cap v}$ as $\bar{U}^{\tilde{\chi}}$. For any $\mathcal{P}$-signed-measure χ consistent with $\bar{\tilde{\chi}}'$, let $U^{\chi} \in R^{G \times G}$ be the matrix whose p, q entry is $\chi_{p \cap q}$ for all $p, q \in G$, where $G \supseteq \bar{\mathcal{Q}}'$ is a linearly independent spanning collection. Then by positive semidefiniteness,

$$0 \leq (\nu^{\bar{\mathcal{Q}}'})^T \bar{U}^{\tilde{\chi}} \nu^{\bar{\mathcal{Q}}'} = \nu^T U^{\chi} \nu = \tag{4.51}$$

$$\left(\sum_{r \in \mathcal{S}^P} \lambda_r^{\nu} \mu^G(r) \right)^T U^{\chi} \left(\sum_{r \in \mathcal{S}^P} \lambda_r^{\nu} \mu^G(r) \right) = \tag{4.52}$$

$$\sum_{r \in \mathcal{S}^P} (\lambda_r^{\nu})^2 (\mu^G(r))^T U^{\chi} \mu^G(r) + \sum_{u, v \in \mathcal{S}^P} \lambda_u^{\nu} \lambda_v^{\nu} (\mu^G(u))^T U^{\chi} \mu^G(v) = \tag{4.53}$$

$$\sum_{r \in \mathcal{S}^P} (\lambda_r^{\nu})^2 \chi_r \tag{4.54}$$

by Lemma 4.4 and Corollary 4.7 since all of the sets in $\mathcal{S}^P$ (the nonempty atoms) are mutually disjoint. □

One other specific example regarding the ν vectors that we will point out is as follows. Let

$$\mathcal{Q}' = \{P, Y_1^P, \ldots, Y_n^P\}, \quad \mathcal{Q} = \{P, Y_1^P, \ldots, Y_n^P, Y_i^P \cap Y_j^P, (i,j = 1, \ldots, n)\}. \tag{4.55}$$

Recall that the following "inclusion-exclusion" inequality is always valid for $\mathcal{P}$-measure consistent vectors $\tilde{\chi} \in R^{\mathcal{Q}}$, where we use the same notation as in the example above.

$$\sum_{i \in J \subseteq \{1,\ldots,n\}} \tilde{\chi}_i - \sum_{i,j \in J, i \neq j} \tilde{\chi}_{i,j} \leq \tilde{\chi}_P \tag{4.56}$$

If the matrix $U^{\tilde{\chi}}$ is positive semidefinite, then writing $v^T U^{\tilde{\chi}} v \geq 0$ where $v_P = 1$, $v_i = -\frac{2}{3}$ for $i \in J$, with the remaining coordinates at 0 yields

$$\sum_{i \in J} \tilde{\chi}_i - \sum_{i,j \in J, i \neq j} \tilde{\chi}_{i,j} \leq \frac{9}{8} \tilde{\chi}_P. \tag{4.57}$$

Moreover, where $v_i = -\frac{1}{2}$ for $i \in J$ the positive semidefineteness constraint gives

$$\frac{3}{2} \sum_{i \in J} \tilde{\chi}_i - \sum_{i,j \in J, i \neq j} \tilde{\chi}_{i,j} \leq 2\tilde{\chi}_P \tag{4.58}$$

so that wherever $\sum_{i \in J} \tilde{\chi}_i \geq 2\tilde{\chi}_P$ this implies that

$$\sum_{i \in J} \tilde{\chi}_i - \sum_{i,j \in J, i \neq j} \tilde{\chi}_{i,j} \leq \tilde{\chi}_P \tag{4.59}$$

and naturally (4.59) will also hold wherever $\sum_{i \in J} \tilde{\chi}_i \leq \tilde{\chi}_P$, and by (4.58) it will also hold wherever

$$\sum_{i,j \in J, i \neq j} \tilde{\chi}_{i,j} \geq \tilde{\chi}_P. \tag{4.60}$$

4.1.2 Combinations of Delta Vectors

One way to characterize the inequalities that are introduced by demanding positive semidefiniteness of $U^{\tilde{\chi}}$ is in terms of sums of delta vectors such that the sum has nonzero values only in its $\mathcal{Q}'$ coordinates. In what follows, wherever we make reference to the collections $\mathcal{Q}$ and $\mathcal{Q}'$ these should be understood to be subcollections of $\mathcal{P}$, with $\mathcal{Q}' \subseteq \mathcal{Q}$, and such that whenever $u, v \in \mathcal{Q}'$ we have $u \cap v \in \mathcal{Q}$.

Let $G \supseteq \mathcal{Q}$. For every vector $v \in R^{\mathcal{Q}'}$, define the vector $\bar{v} \in R^G$ to be the lifting of v to $|G|$ dimensions obtained by padding v with zeroes in all of the appended coordinates. Observe that there always exist sets $\{q_1, \ldots, q_t\} \subseteq \mathcal{P}$ for which

$$\bar{v} = \sum_{i=1,\ldots,t} \beta_i \mu^G(q_i). \tag{4.61}$$

Specifically, we could choose $\{q_1, \ldots, q_t\} = \mathcal{Q}'$ so that we can write $\mu^G(q_i) = e_{q_i}$ and we could choose $\beta_{q_i} = v_{q_i}$. This establishes the following lemma.

Lemma 4.14 *Let $G \supseteq \mathcal{Q}$. Let $\tilde{\chi} \in R^{\mathcal{Q}}$ be $\mathcal{P}$-signed-measure consistent. Let χ be any signed measure on $\mathcal{P}$ with which $\tilde{\chi}$ is consistent. For each $v \in R^{\mathcal{Q}'}$ there exists some collection of sets $\{q_1, \ldots, q_t\} \subseteq \mathcal{P}$ such that the vector $\bar{v} \in R^G$ obtained by padding v with zeroes is such that $\bar{v} = \sum_{i=1}^t \beta_i \mu^G(q_i)$ for some β, and such that*

$$v^T U^{\tilde{\chi}} v = \sum_{i=1}^t \beta_i^2 \chi(q_i) + 2 \sum_{i=1}^t \sum_{j=i+1}^t \beta_i \beta_j \chi(q_i \cap q_j). \tag{4.62}$$

Conversely, for every collection of sets $\{q_1, \ldots, q_t\} \subseteq \mathcal{P}$ for which there are scalars $\beta_1, \ldots, \beta_t$ such that $\sum_{i=1}^t \beta_i \mu^G(q_i)$ is nonzero only in $\mathcal{Q}'$ coordinates, then letting $\bar{v}$ denote the vector $\sum_{i=1}^t \beta_i \mu^G(q_i)$, and letting v denote the projection of $\bar{v}$ on $R^{\mathcal{Q}'}$, we also have

$$v^T U^{\tilde{\chi}} v = \sum_{i=1}^t \beta_i^2 \chi(q_i) + 2 \sum_{i=1}^t \sum_{j=i+1}^t \beta_i \beta_j \chi(q_i \cap q_j). \tag{4.63}$$

Proof: Let U^χ be the $|G| \times |G|$ matrix whose u, v entry is $\chi(u \cap v)$, for each $u, v \in G$. Then

$$v^T U^{\tilde{\chi}} v = \bar{v}^T U^\chi \bar{v} = \tag{4.64}$$

$$\left(\sum_{i=1}^t \beta_i \mu^G(q_i) \right)^T U^\chi \left(\sum_{i=1}^t \beta_i \mu^G(q_i) \right) = \tag{4.65}$$

$$\sum_{i=1}^t \beta_i^2 (\mu^G(q_i))^T U^\chi \mu^G(q_i) + 2 \sum_{i=1}^t \sum_{j=i+1}^t \beta_i \beta_j (\mu^G(q_i))^T U^\chi \mu^G(q_j) = \tag{4.66}$$

$$\sum_{i=1}^t \beta_i^2 (\mu^G(q_i))^T U^\chi \mu^G(q_i) + \tag{4.67}$$

$$2 \sum_{i=1}^t \sum_{j=i+1}^t \beta_i \beta_j (\mu^G(q_i))^T \left(\sum_{r \in \mathcal{S}^P} \alpha_r \zeta^r\{G\} (\zeta^r\{G\})^T \right) \mu^G(q_j) = \tag{4.68}$$

$$\sum_{i=1}^t \beta_i^2 \chi(q_i) + 2 \sum_{i=1}^t \sum_{j=i+1}^t \beta_i \beta_j \sum_{r \in \mathcal{S}^P : r \subseteq q_i \cap q_j} \alpha_r = \tag{4.69}$$

$$\sum_{i=1}^t \beta_i^2 \chi(q_i) + 2 \sum_{i=1}^t \sum_{j=i+1}^t \beta_i \beta_j \chi(q_i \cap q_j). \ \square \tag{4.70}$$

We can say a little bit more also. Observe that every $v \in R^{\mathcal{Q}'}$ can also be written as a linear combination of delta vectors of *disjoint* sets. Specifically, where $G \subseteq \mathcal{P}$ is a spanning collection and we write $\mathcal{Q}' = \{q_i : i = 1, \ldots, t\}$, then by Corollary 3.48 we can write the delta vector

$$e_{q_i} = \mu^G(q_i) = \sum_{r \in \mathcal{S}^P : r \subseteq q_i} \mu^G(r) \tag{4.71}$$

and the sets in $\mathcal{S}^P$ are all mutually disjoint. Thus

$$\bar{v} = \sum_{i=1\dots,t} v_i e_{q_i} \tag{4.72}$$

can be written as a linear combination of delta vectors of disjoint sets. Thus applying the previous lemma we conclude:

Lemma 4.15 *Given $\mathcal{P}$-signed-measure consistent $\tilde{\chi} \in R^{\mathcal{Q}'}$, then for any signed measure χ on $\mathcal{P}$ with which it is consistent, the inequalities implied by $U^{\tilde{\chi}} \succeq 0$ are all of the form*

$$0 \le v^T U^{\tilde{\chi}} v = \sum_{i=1}^{t} \beta_i^2 \chi(q_i) \tag{4.73}$$

for some collection of disjoint sets $\{q_1, \dots, q_t\} \subseteq \mathcal{P}$. Conversely, for every collection of disjoint sets $\{q_1, \dots, q_t\} \subseteq \mathcal{P}$ for which there exist scalars $\beta_1, \dots, \beta_t$ such that $\sum_{i=1}^t \beta_i \mu^G(q_i)$ is nonzero only in $\mathcal{Q}'$ coordinates, positive semidefiniteness implies that the inequality described above holds. $\square$

Thus for every mutually disjoint collection of sets such that some linear combination of delta vectors for those sets has nonzero entries only in positions corresponding to $\mathcal{Q}'$, the nonnegative linear combination of the signed measures of those sets obtained by squaring the coefficients one by one is guaranteed by positive semidefiniteness to be nonnegative, and every inequality generated by positive semidefiniteness is of this form. It is clear that Corollary 4.5 is a special case.

Thus each inequality generated by positive semidefiniteness says that a nonnegative linear combination of signed measures of sets will be nonnegative. Obviously this requirement is not as strong as a requirement that those sets themselves have nonnegative signed measure, but it is something nontrivial nonetheless.

Here is a simple example of a positive semidefiniteness inequality that draws on delta vectors that involve more than just $\mathcal{Q}'$.

Example: Consider

$$\mathcal{Q}' = \{\{0,1\}^n, Y_1, Y_2\}, \quad \mathcal{Q} = \{\{0,1\}^n, Y_1, Y_2, Y_1 \cap Y_2\} \tag{4.74}$$

(we have chosen $P = \{0,1\}^n$ so as to make notation cleaner) and consider the vector

$$v = (1, -\frac{1}{2}, -\frac{1}{2}). \tag{4.75}$$

This vector is a linear combination of the delta vectors corresponding to the disjoint sets

$$(Y_1 \cup Y_2)^c \text{ and } Y_1^c Y_2 \cup Y_1 Y_2^c \tag{4.76}$$

where we have suppressed the intersection signs to reduce clutter. Specifically, where G is any spanning set, and $\bar{v}$ is, as above, the lifting of v to R^G obtained by padding v with zeroes, we have

$$\bar{v} = \mu^G((Y_1 \cup Y_2)^c) + \frac{1}{2}\mu^G(Y_1^c Y_2 \cup Y_1 Y_2^c) \tag{4.77}$$

for some delta vectors $\mu^G((Y_1 \cup Y_2)^c)$ and $\mu^G(Y_1^c Y_2 \cup Y_1 Y_2^c)$. Here is the proof that (4.77) is indeed valid:

For any signed measure χ on $\mathcal{A}$, where we denote the signed measure χ of a set q as $\chi(q)$, and any delta vectors $\mu^G((Y_1 \cup Y_2)^c)$ and $\mu^G(Y_1^c Y_2 \cup Y_1 Y_2^c)$,

$$(\mu^G((Y_1 \cup Y_2)^c))^T \chi + \frac{1}{2}(\mu^G(Y_1^c Y_2 \cup Y_1 Y_2^c))^T \chi = \tag{4.78}$$

$$\chi((Y_1 \cup Y_2)^c) + \frac{1}{2}\chi(Y_1^c Y_2 \cup Y_1 Y_2^c) = \tag{4.79}$$

$$\chi(P) - \chi(Y_1) - \chi(Y_2) + \chi(Y_1 Y_2))+ \tag{4.80}$$

$$\frac{1}{2}\chi(Y_1 Y_2^c) + \frac{1}{2}\chi(Y_1^c Y_2) = \tag{4.81}$$

$$\chi(P) - \chi(Y_1) - \chi(Y_2) + \chi(Y_1 Y_2)+ \tag{4.82}$$

$$\frac{1}{2}(\chi(Y_1) - \chi(Y_1 Y_2)) + \frac{1}{2}(\chi(Y_2) - \chi(Y_1 Y_2)) = \tag{4.83}$$

$$\chi(P) - \frac{1}{2}\chi(Y_1) - \frac{1}{2}\chi(Y_2) = \bar{v}^T \chi \Rightarrow \tag{4.84}$$

for any delta vector $\mu^G((Y_1 \cup Y_2)^c)$,

$$\left(\bar{v} - \mu^G((Y_1 \cup Y_2)^c)\right)^T \chi = \frac{1}{2}\chi(Y_1^c Y_2 \cup Y_1 Y_2^c) \tag{4.85}$$

and thus, since this anlysis holds for any signed measure χ (and in particular it holds for the zeta vectors), $2(\bar{v} - \mu^G((Y_1 \cup Y_2)^c))$ is a delta vector for $Y_1^c Y_2 \cup Y_1 Y_2^c$, which establishes (4.77).

The positive semidefiniteness inequality generated by v is therefore

$$0 \le \chi((Y_1 \cup Y_2)^c) + \frac{1}{4}\chi(Y_1^c Y_2 \cup Y_1 Y_2^c) = \tag{4.86}$$

$$\chi(P) - \chi(Y_1) - \chi(Y_2) + \chi(Y_1 Y_2)+ \tag{4.87}$$

$$\frac{1}{4}(\chi(Y_1) - \chi(Y_1 Y_2)) + \frac{1}{4}(\chi(Y_2) - \chi(Y_1 Y_2)) = \tag{4.88}$$

$$\chi(P) - \frac{3}{4}\chi(Y_1) - \frac{3}{4}\chi(Y_2) + \frac{1}{2}x(Y_1 Y_2) \Rightarrow \tag{4.89}$$

$$\frac{3}{4}\chi(Y_1) + \frac{3}{4}\chi(Y_2) - \frac{1}{2}\chi(Y_1 Y_2) \le \chi(P) \tag{4.90}$$

This is not as strong as stating that

$$\chi(Y_1) + \chi(Y_2) - \chi(Y_1Y_2) \leq \chi(P) \tag{4.91}$$

$$\chi(Y_1) + \chi(Y_2) - 2\chi(Y_1Y_2) \leq \chi(P) \tag{4.92}$$

the first of which is implied by $\chi((Y_1 \cup Y_2)^c) \geq 0$ and the second of which is implied by $\chi(Y_1^c Y_2 \cup Y_1 Y_2^c) \geq 0$, both of which we would have learned from delta vector inequalities had $\mathcal{Q}'$ included the set Y_1Y_2 (as per Corollary 4.5). But nevertheless it is more than what we would know from the delta vectors that involve P, Y_1 and Y_2 alone.

4.1.3 An Example: Stable Set

This section will illustrate some of the concepts previously described in an application to the stable set problem. Most of the valid stable set inequalities that will occupy us here are the same as those that were treated in [LS91], but here we will approach them from within the framework that has been developed in the previous subsections.

Let $G = (N, E)$ be an undirected graph with n nodes, and let $P \subseteq \{0,1\}^n$ be the collection of incidence vectors of the stable sets of the graph. Formally, where the nodes are labeled $1, \ldots, n$, and for each $i, j \in \{1, \ldots, n\}$ for which there is an edge between the i'th and the j'th node, the edge is labeled $\{i, j\}$, then

$$P = \{y \in \{0,1\}^n : y_i + y_j \leq 1,\ \forall i, j : \{i, j\} \in E\}. \tag{4.93}$$

Remark 4.16 *To reduce notational clutter, throughout this section we will represent set functions on $\mathcal{P}$ with respect to collections of sets from $\mathcal{A}$ (as described in Definition 3.25). Thus instead of indexing vectors by set theoretic expressions involving sets of the form Y_i^P, we will be indexing them by expressions involving sets of the form Y_i. In particular we will be using*

$$\mathcal{Q}' = \{P, Y_1, \ldots, Y_n\},\ \mathcal{Q} = \{P, Y_1, \ldots, Y_n, Y_i \cap Y_j,\ (i, j = 1, \ldots, n)\}. \tag{4.94}$$

For any set function on $\mathcal{P}$, recall that where that set function is represented as a vector χ with coordinates corresponding to sets from $\mathcal{A}$, the value $\chi_q,\ q \in \mathcal{A}$ is the set function value of $q \cap P$, and that vectors that can be interpreted in this way as representations of set functions on $\mathcal{P}$ are said to be $\mathcal{P}$-set-function consistent. (Obviously such vectors also describe set functions on $\mathcal{A}$.) Recall also that a vector $\tilde{\chi}$ defined on a subcollection of $\mathcal{A}$ is said to be $\mathcal{P}$-signed-measure consistent iff it is $\mathcal{P}$-set-function consistent, and the set *function it induces on $\mathcal{P}$ is $\mathcal{P}$-signed-measure consistent. (Note also that if $\chi \in R^{\mathcal{A}}$ is $\mathcal{P}$-signed-measure consistent, then for any set $q \in \mathcal{A}$ such that $q \subseteq P^c$, we have $\chi_q = 0$.) Observe that where P is the collection of incidence vectors of stable sets, the sets*

$$P, Y_1^P, \ldots, Y_n^P, Y_i^P \cap Y_j^P, \ \{i,j\} \notin E \tag{4.95}$$

are all distinct (they form a linearly independent collection as shown above in Section 4.1.1), and the sets

$$\{Y_i^P \cap Y_j^P : \{i,j\} \in E\} \tag{4.96}$$

are all the same set, i.e. the empty set. So $\mathcal{P}$-set-function consistency is achieved by any $\tilde{\chi} \in R^{\mathcal{Q}}$ that assigns a common value to all $Y_i \cap Y_j : \{i,j\} \in E$, and as we saw in Lemma 4.8, $\mathcal{P}$-signed-measure consistency will be achieved by any $\tilde{\chi}$ that assigns all of those sets a value of zero.

Observe also that for any $\mathcal{P}$-set-function consistent χ, $\chi_{\{0,1\}^n} = \chi_P$, so for practical purposes the set P in $\mathcal{Q}$ and $\mathcal{Q}'$ can be considered to be interchangeable with the universal set $\{0,1\}^n$. More generally, since for any set $q \in \mathcal{A}$, $\chi(q) = \chi(q \cap P)$, for ease of presentation we will refer to set theoretic expressions of sets $Y_i \in \mathcal{A}$ as being disjoint (or equal) if the sets that they define are disjoint (or equal) when intersected with P. Thus we may say $A \cup B = C$ if $(A \cup B) \cap P = C \cap P$. □

We have already considered the situation for the clique inequalities. In this section we will consider the odd hole, odd antihole, and wheel inequalities. Assume that an odd sized collection $C \subseteq N$ of nodes is composed of nodes $v_1, \ldots, v_k$, $k \geq 3$, represented by $0,1$ variables $y_1, \ldots, y_k$. If C is a chordless cycle in the graph G, the valid inequalities

$$\sum_{i=1}^{k} y_i \leq \frac{k-1}{2} \tag{4.97}$$

are called the odd hole inequalities. If C is a chordless cycle in the graph (N, E^c), i.e. there are edges between every pair of nodes in C except for the sequence of pairs

$$\{v_1, v_2\}, \ldots, \{v_{k-1}, v_k\}, \{v_k, v_1\} \tag{4.98}$$

then the valid inequalities

$$\sum_{i=1}^{k} y_i \leq 2 \tag{4.99}$$

are called the odd antihole constraints. If C is a chordless cycle in the graph, and there exists some node $w \notin C$ (with incidence variable denoted y_w) such that there exists an edge between w and every $v \in C$, then the valid inequalities

$$\sum_{i=1}^{k} y_i + \frac{k-1}{2} y_w \leq \frac{k-1}{2} \tag{4.100}$$

are called the odd wheel constraints.

In what follows we will be assuming that $|C| = k \geq 3$ (odd), and that the nodes belonging to C are numbered $v_1, \ldots, v_k$.

We will now show how to obtain the odd hole, odd antihole and odd wheel inequalities from measure theoretic constraints, and from delta vector constraints (Definition 4.6) in particular. Observe first that for any $\mathcal{P}$-signed measure (represented with respect to $\mathcal{A}$), χ, the following equations are valid since for all $\{i, j\} \in E$ the sets Y_i and Y_j are disjoint, and $\{k, 1\}, \{i, i+1\},\ i = 1, \ldots, k-1$ are all edges in E.

$$\chi(Y_1^c Y_i^c Y_{i+1}^c) = \chi(Y_1^c) - \chi(Y_1^c Y_i) - \chi(Y_1^c Y_{i+1}) = \tag{4.101}$$

$$\chi(P) - \chi(Y_1) - \chi(Y_i) + \chi(Y_1 Y_i) - \chi(Y_{i+1}) + \chi(Y_1 Y_{i+1}) \tag{4.102}$$

and

$$\chi(Y_1 Y_i^c Y_{i+1}^c) = \chi(Y_1) - \chi(Y_1 Y_i) - \chi(Y_1 Y_{i+1}) \tag{4.103}$$

(where i is an integer in $\{2, \ldots, k-1\}$, and the notation $\chi(q)$ means the function value of χ on the set q). Rewriting χ_i for $\chi(Y_i)$, χ_P for $\chi(P)$, and $\chi_{i,j}$ for $\chi(Y_i Y_j)$, etc. to avoid clutter, we must therefore have for any $\mathcal{P}$-measure χ,

$$\chi(Y_1^c Y_i^c Y_{i+1}^c) = \chi_P - \chi_1 - \chi_i - \chi_{i+1} + \chi_{1,i} + \chi_{1,i+1} \geq 0 \tag{4.104}$$

and

$$\chi(Y_1 Y_i^c Y_{i+1}^c) = \chi_1 - \chi_{1,i} - \chi_{1,i+1} \geq 0. \tag{4.105}$$

These are delta vector constraints, as defined in Definition 4.6, of the form $\mu^{\mathcal{Q}}(q)^T \tilde{\chi} \geq 0$, where $\mathcal{Q}$ is as in (4.94), and $\tilde{\chi}$ is the projection of χ on its $\mathcal{Q}$ coordinates. Summing the inequalities of type (4.104) over all even $i = 2, 4, \ldots, k-1$ (recall that k is odd) and adding that sum to the sum of all inequalities of type (4.105) for odd $i = 3, 5, \ldots, k-2$, and noting that each $\chi_{i,i+1}$ (as well as $\chi_{k,1}$) is zero yields the (homogenized) odd hole inequality

$$\frac{k-1}{2} \chi_P - \sum_{i=1}^{k} \chi_i \geq 0. \tag{4.106}$$

The odd wheel inequalities are obtained the same way, but the inequalities of type (4.104) are strengthened to

$$\chi_P - \chi_1 - \chi_i - \chi_{i+1} - \chi_w + \chi_{1,i} + \chi_{1,i+1} \geq 0 \tag{4.107}$$

yielding the desired inequality. Inequality (4.107) is also a delta vector inequality since for all $i \in \{1, \ldots, k, w\}$, (where k mod k is construed as having the value k), Y_i, $Y_{i+1 \bmod k}$, and Y_w are all disjoint sets, so that for any $\mathcal{P}$-signed measure χ,

$$\chi(Y_1^c Y_i^c Y_{i+1}^c Y_w^c) = \chi_P - \chi_1 - \chi_i - \chi_{i+1} - \chi_w + \chi_{1,i} + \chi_{1,i+1}. \tag{4.108}$$

The odd antihole constraint is the sum of the following two delta vector constraints

$$\chi((Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c) = \chi_P - \sum_{\text{odd } i=1}^{k-2} \chi_i - \chi_{k-1,k} \geq 0 \tag{4.109}$$

and

$$\chi((Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c) = \chi_P - \sum_{\text{even } i=2}^{k-1} \chi_i - \chi_k + \chi_{k-1,k} \geq 0. \tag{4.110}$$

The first equality follows from the fact that the sets $Y_1, Y_3, \cdots, Y_{k-2}, Y_{k-1}Y_k$ are pairwise disjoint, and the second equality follows from the fact that the sets $Y_2, Y_4, \cdots, Y_{k-3}, Y_{k-1} \cup Y_k$ are pairwise disjoint, and $\chi(Y_{k-1} \cup Y_k) = \chi_{k-1} + \chi_k - \chi_{k-1,k}$.

Lemma 4.17 *Let P be as in (4.93). Define a vector $\tilde{\chi}$ with coordinates indexed by P and all intersections of up to four sets Y_i, and demand that $\chi_{i,j} = 0$ for all $\{i,j\} \in E$, and similarly for all $\chi_{i,j,k}$ and $\chi_{i,j,k,l}$. Let $\mathcal{Q}$ be as in (4.94). Define the matrix $U^{\tilde{\chi}}$ with rows and columns indexed by $\mathcal{Q}$ and with each u, v entry equal to $\tilde{\chi}(u \cap v)$, and let $U^{\tilde{\chi}} \succeq 0$. Then the projection of $\tilde{\chi}$ on its $P, Y_1, \ldots, Y_n$ coordinates satisfies all odd hole, odd wheel and odd antihole inequalities (homogenized).*

Proof: A trivial extension of Lemma 4.8 shows that $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, and the delta vectors for inequalities (4.104), (4.105), (4.107), (4.109), and (4.110) are all of the form $\mu^{\mathcal{Q}}(q)$. The lemma now follows from Corollary 4.5. □

At this point let us consider the behavior of the N operator of Chapter 1 on the stable set problem, as per its definition in Remark 3.68. The description we will give here may seem a little backhanded, considering that the behavior of N in this case is quite simple to describe in its original incarnation, but this methodology will prove useful repeatedly later on.

Let $\mathcal{Q}'$ be as in (4.94). Define $\bar{K}$ by

$$\bar{K} = \{\tilde{\chi}' \in R^{\mathcal{Q}'} : \tilde{\chi}'_i + \tilde{\chi}'_j \leq \tilde{\chi}'_P \ \forall \{i,j\} \in E, \ \tilde{\chi}' \geq 0, \ \tilde{\chi}'_i \leq \tilde{\chi}'_P, \ i = 1, \ldots, n\} \tag{4.111}$$

and recall from Remark 3.68 that $N(\bar{K}) = \bar{N}^1(\bar{K})$ is the set of vectors $\tilde{\chi}' \in R^{\mathcal{Q}'}$ for which there exists a lifting to a signed measure χ on $\mathcal{A}$ such that the projection of the partial sum signed measure χ^{M_i} on the $\mathcal{Q}'$ coordinates $(\tilde{\chi}')^{M_i} \in \bar{K}$ for each set M_i of the form Y_i or of the form $N_i := Y_i^c$. The constraints imposed by N are therefore

$$\chi_i^{Y_h} + \chi_j^{Y_h} \leq \chi_P^{Y_h},\ h = 1, \ldots, n,\ \{i,j\} \in E \tag{4.112}$$

$$0 \leq \chi_i^{Y_h} \leq \chi_P^{Y_h},\ i, h = 1, \ldots, n \tag{4.113}$$

$$\chi_i^{N_h} + \chi_j^{N_h} \leq \chi_P^{N_h},\ h = 1, \ldots, n,\ \{i,j\} \in E \tag{4.114}$$

$$0 \leq \chi_i^{N_h} \leq \chi_P^{N_h},\ i, h = 1, \ldots, n \tag{4.115}$$

or equivalently, where we represent $\chi(N_i)$ as χ_{i^c} and $\chi(N_i Y_j)$ as $\chi_{i^c,j}$, etc.

$$\chi_{h,i} + \chi_{h,j} \leq \chi_h,\ h = 1, \ldots, n,\ \{i,j\} \in E \tag{4.116}$$

$$0 \leq \chi_{h,i} \leq \chi_h,\ i, h = 1, \ldots, n \tag{4.117}$$

$$\chi_{h^c,i} + \chi_{h^c,j} \leq \chi_{h^c},\ h = 1, \ldots, n,\ \{i,j\} \in E \tag{4.118}$$

$$0 \leq \chi_{h^c,i} \leq \chi_{h^c},\ i, h = 1, \ldots, n. \tag{4.119}$$

Since χ is a signed measure we can restate constraints (4.118) and (4.119) as

$$\chi_i - \chi_{h,i} + \chi_j - \chi_{h,j} \leq \chi_P - \chi_h,\ h = 1, \ldots, n,\ \{i,j\} \in E \tag{4.120}$$

$$0 \leq \chi_i - \chi_{h,i} \leq \chi_P - \chi_h \tag{4.121}$$

What we have done here is to describe those coordinates that do not correspond to pure intersections of Y_i sets in terms of those that do. Constraints (4.116), (4.117), (4.120) and (4.121) now involve only coordinates from $\mathcal{Q}$, which is linearly independent (with respect to $\mathcal{A}$), so that $\mathcal{A}$-signed-measure consistency is assured, cf. the end of Remark 3.68. Thus constraints (4.116), (4.117), (4.120) and (4.121) completely describe $N(\bar{K})$. Observe also that for any $\{i,j\} \in E$, constraint (4.116) implies that $\chi_i + \chi_{i,j} \leq \chi_i$, which, together with constraint (4.117), implies that $\chi_{i,j} = 0$. Thus by Lemma 4.8, $\mathcal{P}$-signed-measure consistency is also assured.

The inequalities (4.104) and (4.105) are therefore enforced by the N operator, and thus the odd hole inequalities are all satisfied by $N(\bar{K})$, but for the strengthened inequalities (4.107), and therefore the odd wheel constraints, we have no guarantee. Similarly we have no guarantee that the odd antihole constraints will be satisfied by $N(\bar{K})$.

We saw in Lemma 4.17 that insisting on positive semidefiniteness for the matrix $U^{\tilde{\chi}}$ with rows and columns indexed by $\mathcal{Q}$ is enough to guarantee on its own that the odd hole, odd wheel and odd antihole inequalities are satisfied, so long as the entries of the matrix are $\mathcal{P}$-signed-measure consistent. But if we insist only that the matrix with rows and columns indexed by $\mathcal{Q}'$ be positive semidefinite (even continuing to assume $\mathcal{P}$-signed measure-consistency) then that argument will fail. We will see shortly, however, that positive semidefiniteness of the smaller matrix (as is demanded by $N^+(\bar{K})$), in concert with the $N(\bar{K})$ constraints, is in fact sufficient to guarantee that the odd antihole and odd wheel constraints are satisfied.

Let $\tilde{\chi} \in R^{\mathcal{Q}}$ be consistent with some $\mathcal{P}$-signed measure χ, and denote the projection of $\tilde{\chi}$ on $\mathcal{Q}'$ as $\tilde{\chi}'$. Observe first that if the odd antihole constraint $\sum_{i=1}^k \tilde{\chi}_i \leq 2\tilde{\chi}_P$[1] is represented as $v^T\tilde{\chi}' \geq 0$, and delta vector constraint (4.109) is represented as

$$\left(\mu^{\mathcal{Q}}((Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c)\right)^T \tilde{\chi} \geq 0, \tag{4.122}$$

and delta vector constraint (4.110) is represented as

$$\left(\mu^{\mathcal{Q}}((Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c)\right)^T \tilde{\chi} \geq 0 \tag{4.123}$$

then the vector $\bar{v} \in R^{\mathcal{Q}}$ obtained by padding v with zeroes satisfies that

$$\bar{v} = \mu^{\mathcal{Q}}((Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c) + \mu^{\mathcal{Q}}((Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c). \tag{4.124}$$

Observe furthermore that

$$(Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c \cap (Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c = Y_1^c Y_2^c \cdots Y_k^c. \tag{4.125}$$

Thus if the matrix $U^{\tilde{\chi}}$ with rows and columns indexed by $\mathcal{Q}'$ is positive semidefinite, then Lemma 4.14 implies that $v^T U^{\tilde{\chi}} v =$

$$\chi((Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c) + \chi((Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c) + 2\chi(Y_1^c Y_2^c \cdots Y_k^c) \geq 0. \tag{4.126}$$

By equation (4.124) and the definition of delta vectors,

$$\chi((Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c) + \chi((Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c) =$$

$$\bar{v}^T \tilde{\chi} = v^T \tilde{\chi}' = 2\tilde{\chi}_P - \sum_{i=1}^k \tilde{\chi}_i \tag{4.127}$$

[1] Note that since χ, $\tilde{\chi}$ and $\tilde{\chi}'$ are all consistent with one another we could have written this as $\sum_{i=1}^k \tilde{\chi}'_i \leq 2\tilde{\chi}'_P$, or as $\sum_{i=1}^k \chi_i \leq 2\chi_P$. We have chosen to write it here in the way that we have in order to emphasize that the vector $\tilde{\chi}$ is the one that is being constrained by $U^{\tilde{\chi}} \succeq 0$.

and we can therefore conclude from (4.126) and (4.127) that

$$\sum_{i=1}^{k} \tilde{\chi}_i - 2\chi(Y_1^c \cdots Y_k^c) \leq 2\chi_P. \tag{4.128}$$

Thus for any $\mathcal{P}$-signed measure consistent $\tilde{\chi}$ and any $\mathcal{P}$-signed measure χ with which it is consistent, $U^{\tilde{\chi}} \succeq 0$ will imply that (4.128) will hold. Thus if there is an χ consistent with $\tilde{\chi}$ for which $\chi(Y_1^c \cdots Y_k^c) \leq 0$, then (4.128) will imply $\sum_{i=1}^{k} \tilde{\chi}_i \leq 2\tilde{\chi}_P$, i.e. it will imply that $\tilde{\chi}'$ satisfies the odd antihole constraints. But we have not yet guaranteed that any such χ will exist. Perhaps every χ consistent with $\tilde{\chi}$ is such that $\chi(Y_1^c \cdots Y_k^c) > 0$ (the set $Y_1^c \cdots Y_k^c$ has a nonempty intersection with P, and can therefore have nonzero signed measure). This suggests several possible approaches that could now be taken toward guaranteeing that $\sum_{i=1}^{k} \tilde{\chi}_i \leq 2\tilde{\chi}_P$ will in fact be satisfied. One possible approach is to attempt to show that for each choice of $\tilde{\chi}$, after possibly imposing some valid constraints on $\tilde{\chi}$, we can actually construct a signed measure χ on $\mathcal{P}$ consistent with $\tilde{\chi}$ for which $\chi(Y_1^c \cdots Y_k^c) \leq 0$.

The approach that we will be taking in the proof of the following theorem, however, is to show that for any $\mathcal{P}$-signed measure χ consistent with $\tilde{\chi}$, the assumption that $\chi(Y_1^c \cdots Y_k^c)$ is *positive*, together with some valid constraints on $\tilde{\chi}$, is sufficient to guarantee that $\sum_{i=1}^{k} \tilde{\chi}_i \leq 2\tilde{\chi}_P$. That is, we will show that for any $\mathcal{P}$-signed-measure consistent $\tilde{\chi}$ that satisfies these additional valid constraints, if $\tilde{\chi}$ is consistent with a $\mathcal{P}$-measure χ on $\mathcal{P}$ for which $\chi(Y_1^c \cdots Y_k^c) > 0$, then $\sum_{i=1}^{k} \tilde{\chi}_i \leq 2\tilde{\chi}_P$. Thus for each $\mathcal{P}$-signed-measure consistent $\tilde{\chi}$ that satisfies $U^{\tilde{\chi}} \succeq 0$ as well as these additional constraints, if there is a $\mathcal{P}$-signed measure χ consistent with $\tilde{\chi}$ that satisfies $\chi(Y_1^c \cdots Y_k^c) \leq 0$, then positive semidefiniteness implies $\sum_{i=1}^{k} \tilde{\chi}_i \leq 2\tilde{\chi}_P$, and if there is no such χ, then since by $\mathcal{P}$-signed measure consistency there exists a $\mathcal{P}$-signed measure χ consistent with $\tilde{\chi}$, it must be that $\chi(Y_1^c \cdots Y_k^c) > 0$, which implies $\sum_{i=1}^{k} \tilde{\chi}_i \leq 2\tilde{\chi}_P$ by assumption. Thus in either case we are guaranteed that $\tilde{\chi}$ will indeed satisfy the odd antihole constraints $\sum_{i=1}^{k} \tilde{\chi}_i \leq 2\tilde{\chi}_P$.

For the purposes of the following theorem, let P be as in (4.93); let $\tilde{\chi} \in R^{\mathcal{Q}}$ with projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$ (where $\mathcal{Q}$ and $\mathcal{Q}'$ are as defined in (4.94)) satisfy $\tilde{\chi}_{i,j} = 0,\ \forall \{i,j\} \in E$, so that $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent by Lemma 4.8. Let χ be a signed measure on $\mathcal{P}$ consistent with $\tilde{\chi}$. The matrix $U^{\tilde{\chi}}$ is, as usual, the matrix with rows and columns indexed by $\mathcal{Q}'$, with each u, v entry equal to $\chi(u \cap v)$.

Theorem 4.18 *Assume the following inequalities hold for χ:*

$$\chi_{i,j} + \chi_{i,k} \leq \chi_i,\ \forall \{j,k\} \in E. \tag{4.129}$$

Then if $U^{\tilde{\chi}} \succeq 0$, the odd antihole constraints will all be satisfied by $\tilde{\chi}'$.

Proof: By (4.128), it suffices to show that the inequality

$$\sum_{i=1}^{k} \chi_i \leq 2\chi_P \tag{4.130}$$

holds for all $\mathcal{P}$-signed measures χ that satisfy the conditions of the theorem, and for which

$$\chi(Y_1^c \cdots Y_k^c) \geq 0. \tag{4.131}$$

Since for any signed measure $\mathcal{X}$, and any pair of sets A, B, we always have $\mathcal{X}(A)+\mathcal{X}(B) = \mathcal{X}(A \cup B) + \mathcal{X}(A \cap B)$, and since χ is a signed measure we therefore have by (4.127) and (4.125),

$$2\chi_P - \sum_{i=1}^{k} \chi_i = \chi((Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c) + \chi((Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c) = \tag{4.132}$$

$$\chi((Y_1 \cup Y_3 \cup \cdots \cup Y_{k-2} \cup Y_{k-1}Y_k)^c \cup (Y_2 \cup Y_4 \cup \cdots \cup Y_{k-1} \cup Y_k)^c) + \chi(Y_1^c Y_2^c \cdots Y_k^c) = \tag{4.133}$$

$$\chi(Y_1^c Y_3^c \cdots Y_{k-2}^c (Y_{k-1}Y_k)^c \cup (Y_2^c Y_4^c \cdots Y_{k-1}^c Y_k^c)) + \chi(Y_1^c Y_2^c \cdots Y_k^c). \tag{4.134}$$

Now notice that a point $y \in P$ belongs to the set

$$Y_1^c Y_3^c \cdots Y_{k-2}^c (Y_{k-1}Y_k)^c \cup (Y_2^c Y_4^c \cdots Y_{k-1}^c Y_k^c) \tag{4.135}$$

if and only if it belongs either to none of the sets $Y_1, \dots, Y_k$, or if it belongs to exactly one of them. (Since vertices $v_1, \dots, v_k$ form an odd antihole, the only pair of vertices that can belong to a stable set are of the form v_i, v_{i+1} or v_k, v_1. Thus the only way for a point $y \in P$ to belong to two or more sets from among $\{Y_1, \dots, Y_k\}$ would be for it to belong to some Y_i and Y_{i+1}, $i \in \{1, \dots, k-1\}$, or to Y_k and Y_1, and no such point could belong to (4.135).) Thus since χ is a $\mathcal{P}$-signed measure (cf. the end of Remark 4.16), we obtain

$$\chi(Y_1^c Y_3^c \cdots Y_{k-2}^c (Y_{k-1}Y_k)^c \cup (Y_2^c Y_4^c \cdots Y_{k-1}^c Y_k^c)) + \chi(Y_1^c Y_2^c \cdots Y_k^c) = \tag{4.136}$$

$$\chi\left(\{y \in P : y \in \text{exactly one set } Y_i, i \in \{1, \dots, k\}\}\right) + 2\chi(Y_1^c Y_2^c \cdots Y_k^c). \tag{4.137}$$

By (4.131) this is at least as large as

$$\chi\left(\{y \in P : y \in \text{exactly one set } Y_i, i \in \{1, \dots, k\}\}\right) = \tag{4.138}$$

$$\sum_{i=1}^{k} \chi\left(\{y \in P : y \in Y_i \text{ only}\}\right) = \tag{4.139}$$

$$\sum_{i=1}^{k} (\chi_i - \chi_{i,i-1} - \chi_{i,i+1}) \geq 0 \tag{4.140}$$

where when $i = 1$ then we replace $i-1$ by k, and when $i = k$ we replace $i+1$ by 1. The last equality follows from the fact, noted above, that the only way for a point $y \in P$ to belong to more than one set from among $\{Y_1, \ldots, Y_k\}$ would be for it to belong to some Y_j and Y_{j+1}, so the set of points in P that belong to exactly the set Y_i from among $\{Y_1, \ldots, Y_k\}$ is $Y_i - Y_iY_{i-1} - Y_iY_{i+1}$ (cf. the end of Remark 4.16). The final inequality follows from (4.129). □

The situation for the wheel constraints is more complicated, but it points out one way to generalize the procedure of the above proof. Recall that the wheel constraint

$$\sum_{i=1}^{k} \tilde{\chi}_i + \frac{k-1}{2}\tilde{\chi}_w \leq \frac{k-1}{2}\tilde{\chi}_P \tag{4.141}$$

is the sum of the $\frac{k-1}{2}$ delta vector inequalities

$$\left(\mu^{\mathcal{Q}}(Y_1^cY_i^cY_{i+1}^cY_w^c)\right)^T \tilde{\chi} = \tilde{\chi}_P - \tilde{\chi}_1 - \tilde{\chi}_i - \tilde{\chi}_{i+1} - \tilde{\chi}_w + \tilde{\chi}_{1,i} + \tilde{\chi}_{1,i+1} \geq 0 \tag{4.142}$$

for all even $i = 2, 4, 6, \ldots, k-1$ (recall that k is odd), and the $\frac{k-1}{2} - 1$ delta vector inequalities

$$\left(\mu^{\mathcal{Q}}(Y_1Y_i^cY_{i+1}^c)\right)^T \tilde{\chi} = \tilde{\chi}_1 - \tilde{\chi}_{1,i} - \tilde{\chi}_{1,i+1} \geq 0 \tag{4.143}$$

for all odd $i = 3, 5, \ldots, k-2$.

Let the $\frac{k-1}{2}$ sets $Y_1^cY_i^cY_{i+1}^cY_w^c$ and the $\frac{k-1}{2} - 1$ sets $Y_1Y_i^cY_{i+1}^c$ be denoted as $q_1, \ldots, q_{k-2}$. Thus if $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent and $U^{\tilde{\chi}} \succeq 0$, by the same reasoning as we applied above in the case of the odd antihole inequalities, Lemma 4.14 implies that

$$\sum_{i=1}^{k} \tilde{\chi}_i + \frac{k-1}{2}\tilde{\chi}_w - 2\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) \leq \frac{k-1}{2}\tilde{\chi}_P. \tag{4.144}$$

Again this accomplishes half of the job for us, so that in order to be guaranteed that $\tilde{\chi}'$ will satisfy the odd wheel constraints it suffices, after imposing certain additional constraints on $\tilde{\chi}$, to establish that for each $\mathcal{P}$-signed-measure consistent $\tilde{\chi}$ that satisfies these additional constraints, if $\tilde{\chi}$ is consistent with a signed measure χ on $\mathcal{P}$ for which

$$\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) \geq 0, \tag{4.145}$$

then $\sum_{i=1}^{k} \tilde{\chi}_i + \frac{k-1}{2}\tilde{\chi}_w \leq \frac{k-1}{2}\tilde{\chi}_P$.

As above, let P be as in (4.93); let $\tilde{\chi} \in R^{\mathcal{Q}}$ with projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$ (where $\mathcal{Q}$ and $\mathcal{Q}'$ are as defined in (4.94)) satisfy $\tilde{\chi}_{i,j} = 0,\ \forall \{i,j\} \in E$, so that $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, and let χ be a signed measure on $\mathcal{P}$ consistent with $\tilde{\chi}$. The matrix $U^{\tilde{\chi}}$ is, as usual, the matrix with rows and columns indexed by $\mathcal{Q}'$, with each u, v entry equal to $\chi(u \cap v)$.

Theorem 4.19 *Assume that the following inequalities hold for χ:*

$$\chi_{i,j} + \chi_{i,k} \le \chi_i,\ \forall \{j,k\} \in E. \tag{4.146}$$

Then if $U^{\tilde{\chi}} \succeq 0$, the odd wheel constraints will all be satisfied by $\tilde{\chi}'$.

Proof: Denote the $k-2$ sets whose delta vector inequalities summed to give the odd wheel constraint, i.e. the $\frac{k-1}{2}$ sets

$$Y_1^c Y_i^c Y_{i+1}^c Y_w^c \tag{4.147}$$

and the $\frac{k-1}{2} - 1$ sets

$$Y_1 Y_i^c Y_{i+1}^c, \tag{4.148}$$

by $\{q_1, \ldots, q_{k-2}\}$, so the odd wheel constraint can be represented as

$$\sum_{i=1}^{k-2} \chi(q_i) \ge 0. \tag{4.149}$$

By (4.144), if

$$\sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) \le 0, \tag{4.150}$$

then the odd wheel constraints are satisfied. So let us assume

$$\sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) \ge 0. \tag{4.151}$$

Thus to prove the theorem it suffices to show that $\sum_{i=1}^{k-2} \chi(q_i) \ge 0$ for each $\mathcal{P}$-signed measure χ satisfying the conditions of the theorem for which the additional constraint $\sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) \ge 0$ is also assumed to hold.

Again, as in the case of the odd antihole constraints, there is a natural relationship between the intersections of the q_i sets and the odd wheel constraint. For the case of the odd antihole constraints we noted that the sum of the signed measures of the two sets whose delta vector inequalities sum to give the odd antihole constraints equals the signed measure of the union plus the signed measure of the intersection. Similarly here, if $k > 3$, then adding the signed measure of all pairwise unions of sets q_i to the signed measure of all pairwise intersections yields a multiple of $\sum_{i=1}^{k-2} \chi(q_i)$. Assume now that $k \ge 5$. (If $k = 3$ then the odd wheel is a clique and the odd wheel constraint is just a clique constraint, and it is therefore satisfied by virtue of positive semidefiniteness and $\mathcal{P}$-signed-measure consistency (Corollary 4.9).) Consider that for each q_i, adding the $k-3$ terms $\chi(q_i \cap q_j),\ j \ne i$ to the $k-3$ terms $\chi(q_i \cup q_j),\ j \ne i$ yields $k-3$ copies of $\chi(q_i)$. Thus

$$\sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cup q_j) + \sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) = \tag{4.152}$$

$$(k-3)\sum_{i=1}^{k-2}\chi(q_i). \tag{4.153}$$

Thus to establish that $\sum_{i=1}^{k-2}\chi(q_i)\geq 0$ it suffices to show that whenever we assume (4.151), we will have

$$\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-2}\chi(q_i\cup q_j)\geq 0. \tag{4.154}$$

It is also enough to show that (assuming (4.151),)

$$\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-2}\chi\left((q_i\cup q_j)-(q_i\cap q_j)\right)\geq 0 \tag{4.155}$$

since the pairwise intersections are a subset of the pairwise unions, and this is what we will be doing in particular. One way to think about the idea at work in these proofs is as follows. We are interested in establishing that $\sum_{i=1}^{k-2}\chi(q_i)\geq 0$, but positive semidefiniteness only establishes (4.144), which is an inequality of the form

$$\sum_{i=1}^{k-2}\chi(q_i)+z(\chi)\geq 0 \tag{4.156}$$

for some number $z(\chi)$. Thus if we can establish by other means that wherever $z(\chi)>0$ then

$$\sum_{i=1}^{k-2}\chi(q_i)-z(\chi)\geq 0 \tag{4.157}$$

as well then we will indeed be able to conclude that $\sum_{i=1}^{k-2}\chi(q_i)\geq 0$. The plan is therefore to show that $\sum_{i=1}^{k-2}\chi(q_i)$ minus (a multiple of) the sum of the intersections (which yields the sum of unions) is nonnegative.[2] (The demonstration is long and complicated and none of the later work will depend on it.)

[2] As we indicated, in our proof we will show that for every $\mathcal{P}$-signed-measure consistent $\tilde{\chi}$ that satisfies $\tilde{\chi}_{i,j}+\tilde{\chi}_{i,k}\leq\tilde{\chi}_i,\ \forall\{j,k\}\in E$, and every $\mathcal{P}$-signed measure χ consistent with $\tilde{\chi}$, either

$$\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-2}\chi(q_i\cap q_j)\leq 0 \tag{4.158}$$

or

$$\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-2}\chi(q_i\cup q_j)\geq 0. \tag{4.159}$$

It is worth noting, however, that strictly speaking it is not necessary to actually *prove* this for *every* χ consistent with $\tilde{\chi}$. In order to show that every $\tilde{\chi}\in R^{\mathcal{Q}}$ that satisfies the conditions of the theorem does in fact satisfy the odd wheel constraints, it is actually sufficient to show that for each choice of $\tilde{\chi}$ that satisfies the conditions of the theorem, there is either *some* χ consistent with $\tilde{\chi}$ such that

$$\sum_{i=1}^{k-2}\sum_{j=i+1}^{k-2}\chi(q_i\cap q_j)\leq 0 \tag{4.160}$$

Let us denote the sets q_i of the form

$$Y_1^c Y_i^c Y_{i+1}^c Y_w^c \tag{4.163}$$

as W_i, and let us denote the sets of the form

$$Y_1 Y_i^c Y_{i+1}^c \tag{4.164}$$

as T_i. Observe that for all i and j,

$$W_i \cap T_j = \emptyset \Rightarrow \chi(W_i \cup T_j) = \chi(W_i) + \chi(T_j). \tag{4.165}$$

Since there are $\frac{k-1}{2}$ sets W_i and $\frac{k-1}{2} - 1$ sets T_i, the signed measure of these unions will give us $\frac{k-1}{2} - 1$ copies of each $\chi(W_i)$, and $\frac{k-1}{2}$ copies of each $\chi(T_i)$. Thus, in sum, it gives us $\frac{k-1}{2} - 1$ copies of $\sum_{i=1}^{k-2} \chi(q_i)$ plus an additional copy of each $\chi(T_i)$. Subtract off those $\frac{k-1}{2} - 1$ copies of $\sum_{i=1}^{k-2} \chi(q_i)$ from both sides of inequality (4.153), and note that the conditions of the theorem already imply that each $\chi(T_i) \geq 0$. All we need to show, therefore, is that the sum of the signed measures of all the

$$W_i \cup W_j - W_i W_j \text{ and } T_i \cup T_j - T_i T_j \tag{4.166}$$

is nonnegative.

$$\chi(W_i \cup W_j) = \chi(Y_1^c Y_i^c Y_{i+1}^c Y_w^c \cup Y_1^c Y_j^c Y_{j+1}^c Y_w^c) = \tag{4.167}$$

$$\chi(Y_1^c Y_w^c (Y_i^c Y_{i+1}^c \cup Y_j^c Y_{j+1}^c)) \tag{4.168}$$

Thus (skipping a few steps)

$$\chi(W_i \cup W_j - W_i W_j) = \tag{4.169}$$

$$\chi(Y_1^c Y_w^c (Y_i^c Y_{i+1}^c (Y_j \cup Y_{j+1}))) + \chi(Y_1^c Y_w^c (Y_j^c Y_{j+1}^c (Y_i \cup Y_{i+1}))) = \tag{4.170}$$

(since $Y_w^c \supseteq Y_i \cup Y_{i+1},\ \forall i = 1, \ldots, k$, and since each Y_i and Y_{i+1} are disjoint)

$$\chi(Y_1^c Y_i^c Y_{i+1}^c Y_j) + \chi(Y_1^c Y_i^c Y_{i+1}^c Y_{j+1}) + \chi(Y_1^c Y_j^c Y_{j+1}^c Y_i) + \chi(Y_1^c Y_j^c Y_{j+1}^c Y_{i+1}). \tag{4.171}$$

---

or there is *some* χ consistent with $\tilde{\chi}$ such that

$$\sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cup q_j) \geq 0. \tag{4.161}$$

In the first case we have already seen from (4.144) that positive semidefiniteness would imply the odd wheel constraint, and if $\sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) > 0$, and $\sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cup q_j) \geq 0$ as well, then

$$0 \leq \sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cup q_j) + \sum_{i=1}^{k-2} \sum_{j=i+1}^{k-2} \chi(q_i \cap q_j) = (k-3) \sum_{i=1}^{k-2} \chi(q_i) \tag{4.162}$$

which implies that $\sum_{i=1}^{k-2} \chi(q_i) \geq 0$.

We get such terms for every even i and $j = 2, 4, \ldots, k-1$. Moreover each

$$\chi(Y_1^c Y_i^c Y_{i+1}^c Y_j) = \chi_{1^c,j} - \chi_{1^c,j,i} - \chi_{1^c,j,i+1} \tag{4.172}$$

since each Y_i and Y_j are disjoint. Since we get such terms for both j and $j+1$, we have such terms for all $j = 2, 3, \ldots, k$. By similar reasoning we get

$$\chi(T_i \cup T_j - T_i T_j) = \tag{4.173}$$

$$\chi(Y_1 Y_i^c Y_{i+1}^c Y_j) + \chi(Y_1 Y_i^c Y_{i+1}^c Y_{j+1}) + \chi(Y_1 Y_j^c Y_{j+1}^c Y_i) + \chi(Y_1 Y_j^c Y_{j+1}^c Y_{i+1}) \tag{4.174}$$

and we get such terms for every odd i and $j = 3, 5, \ldots, k-2$, where, as above, each

$$\chi(Y_1 Y_i^c Y_{i+1}^c Y_j) = \chi_{1,j} - \chi_{1,j,i} - \chi_{1,j,i+1} \tag{4.175}$$

and we get such terms for every $j = 3, 4, \ldots, k-1$. Now for $j = 2$ or k we have

$$\chi_{1^c,j} - \chi_{1^c,j,i} - \chi_{1^c,j,i+1} = \chi_j - \chi_{j,i} - \chi_{j,i+1} \geq 0 \tag{4.176}$$

by hypothesis. A term $\chi_{1^c,j}$ appears for every W_i other than the one for which i or $i+1 = j$, so it appears $\frac{k-1}{2} - 1$ times. A term $\chi_{1,j}$ appears for each T_i other than the one for which i or $i+1 = j$, so it appears $\frac{k-1}{2} - 2$ times. (Summing these will allow us to replace all of the $\chi_{1,j}$ and all but one of the $\chi_{1^c,j}$ with $\frac{k-1}{2} - 2$ terms χ_j.)

As for the terms with a minus sign, for the remaining $j = 3, \ldots, k-1$, for even j there is a term $\chi_{1^c,j,i}$ for each $i = 2, \ldots, k$, $i \neq j, j+1$ (we only took unions of distinct W_i and W_j). For $i = 2$, or k, however, $\chi_{1^c,j,i} = \chi_{j,i}$. For odd j there is a term $\chi_{1^c,j,i}$ for each $i = 2, \ldots, k$, $i \neq j, j-1$ with similar behavior. Each of the terms described appears exactly once for each ordered pair $\{j, i\}$. Similarly for every even j there is a term $\chi_{1,j,i}$ for each $i = 3, \ldots, k-1$, $i \neq j, j-1$, and for each odd j there is a term $\chi_{1,j,i}$ for each $i = 3, \ldots, k-1$, $i \neq j, j+1$. Again each such term appears exactly once (for each ordered pair $\{i, j\}$). Thus for all terms $\chi_{1^c,j,i}$ there is a term $\chi_{1,j,i}$, for all $i, j = 3, \ldots, k-1$, $i \neq j, |i-j| \neq 1$. As for the case of $i = 2$ or $i = k$ we found that we could anyway replace $\chi_{1^c,j,i}$ by $\chi_{j,i}$.

Consider at this point the part of the sum that was generated by the unions of W sets. This part of the sum is a sum of expressions of the form

$$\chi_{1^c,j} - \chi_{1^c,j,i} - \chi_{1^c,j,i+1}, \; i = 2, 4, \ldots, k-1, \; j = 3, 4, \ldots, k-1, \; i \neq j, j-1. \tag{4.177}$$

We have already dealt with those expressions corresponding to $j = 2$ or k. Consider all such expressions for a given $j \in \{3, \ldots, k-1\}$. All but one of the $\chi_{1^c,j}$ can be paired with a $\chi_{1,j}$. The single remaining $\chi_{1^c,j} = \chi_j - \chi_{j,1}$. All $\chi_{1^c,j,i}$, $i = 3, \ldots, k-1$ can be paired with a $\chi_{1,j,i}$. Notice that there is a term $\chi_{1^c,j,i}$ for each $i = 2, \ldots, k$ except where $i = j$ and except where $i = j - 1$ for odd j, and except where $i = j + 1$ for even j. But since in any case $\chi_{1^c,j,i}$ as well as $\chi_{1,j,i}$ are both zero wherever $|i - j| = 1$ we can ignore all such terms and deal only with $i \in \{2, \ldots, k\} - \{j - 1, j, j + 1\}$. This pairing exhausts all terms arising from the unions of T sets of the form

$$\chi_{1,j} - \chi_{1,j,i} - \chi_{1,j,i+1}, \; i = 3, 5, \ldots, k - 1, \; i \neq j, j - 1. \tag{4.178}$$

The remaining unpaired terms $\chi_{1^c,j,2}$ and $\chi_{1^c,j,k}$ are in any case equal to $\chi_{j,2}$ and $\chi_{j,k}$ respectively. Thus the sum of all these terms for a fixed j has $\frac{k-1}{2} - 1$ terms χ_j with a plus sign, and a term $-\chi_{j,i}$ for each $i = 1, \ldots, k, \; i \neq j, |i - j| \neq 1$. But we already know by hypothesis that

$$\chi_j - \chi_{j,j+2} - \chi_{j,j+3} \geq 0 \tag{4.179}$$

$$\chi_j - \chi_{j,j+4} - \chi_{j,j+5} \geq 0 \; \ldots \tag{4.180}$$

moving around the cycle until

$$\chi_j - \chi_{j,j-3} - \chi_{j,j-2} \geq 0. \tag{4.181}$$

This accounts for all the terms of the sum. Repeating over all $j = 3, \ldots, k - 1$ we conclude that the sum is nonnegative. □

4.1.4 Positive Semidefiniteness in Combination With Other Constraints

In this section we will carry over the methodology we applied to the stable set problem in the previous subsection to general sets $P \subseteq \{0, 1\}^n$.

Theorem 4.20 *Let $P \subseteq \{0, 1\}^n$, Let $\mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$ be such that for all sets $u, v \in \mathcal{Q}'$ we have $u \cap v \in \mathcal{Q}$. Let $\tilde{\chi} \in R^{\mathcal{Q}}$ with projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$ be $\mathcal{P}$-signed-measure consistent, and let $U^{\tilde{\chi}} \succeq 0$, where $U^{\tilde{\chi}}$ is the matrix with rows and columns indexed by $\mathcal{Q}'$, with each u, v entry equal to $\tilde{\chi}(u \cap v)$. Let $G \subseteq \mathcal{P}$ be such that $\mathcal{Q}' \subseteq G$, and assume that there exist sets $q_1, \ldots, q_k \in \mathcal{P}$ such that the vector $v \in R^{\mathcal{Q}'}$, when lifted to $\bar{v} \in R^G$ by appending coordinates for all $q \in G - \mathcal{Q}'$ all at value zero, can be written as a sum of delta vectors*

$$\bar{v} = \sum_{i=1}^{k} \mu^G(q_i). \tag{4.182}$$

If there exists a signed measure χ on $\mathcal{P}$ consistent with $\tilde{\chi}$ such that either

$$\sum_{i=1}^{k} \sum_{j=i+1}^{k} \chi(q_i \cup q_j) \geq 0 \tag{4.183}$$

or

$$\sum_{i=1}^{k}\sum_{j=i+1}^{k}\chi(q_i \cap q_j) \leq 0 \tag{4.184}$$

then

$$v^T\tilde{\chi}' = \sum_{i=1}^{k}\chi(q_i) \geq 0. \tag{4.185}$$

More generally, if $v \in R^{Q'}$ *is such that*

$$\bar{v} = \sum_{i=1}^{k}\beta_i\mu^G(q_i), \tag{4.186}$$

if $\sum_{i=1}^{k}\beta_i > 0$*, then if for some signed measure* χ *on* $\mathcal{P}$ *consistent with* $\tilde{\chi}$*,*

$$\sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_i \cup q_j) \geq 0 \tag{4.187}$$

then we can also conclude that

$$v^T\tilde{\chi}' = \sum_{i=1}^{k}\beta_i\chi(q_i) \geq 0. \tag{4.188}$$

Proof: If $k = 1$ then the theorem follows from Corollary 4.5 and the definition of delta vectors, so assume $k \geq 2$. The first part of the theorem is a direct consequence of the argument at the beginning of the proof of the Theorem 4.19. As for the second part, rewriting the statement of Lemma 4.14, by positive semidefiniteness we have

$$\sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_iq_j) \geq 0 \tag{4.189}$$

and therefore by hypothesis

$$0 \leq \sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_iq_j) + \sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_i \cup q_j) = \tag{4.190}$$

$$\sum_{i,j=1}^{k}\beta_i\beta_j(\chi(q_i) + \chi(q_j)) = \tag{4.191}$$

$$\sum_{i=1}^{k}\beta_i\sum_{j=1}^{k}\beta_j(\chi(q_i) + \chi(q_j)) = \tag{4.192}$$

$$\sum_{i=1}^{k}\beta_i(v^T\tilde{\chi}' + (\sum_{j=1}^{k}\beta_j)\chi(q_i)) = \tag{4.193}$$

$$\left(\sum_{i=1}^{k}\beta_i\right)v^T\tilde{\chi}' + \left(\sum_{j=1}^{k}\beta_j\right)v^T\tilde{\chi}' = \tag{4.194}$$

$$2\left(\sum_{i=1}^{k}\beta_i\right)v^T\tilde{\chi}' \Rightarrow \tag{4.195}$$

$$v^T\tilde{\chi}' \geq 0. \ \square \tag{4.196}$$

Naturally there are other ways to relate the expressions $\sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_iq_j)$, which are guaranteed by positive semidefiniteness to be nonnegative, to the sum of the signed measures $v^T\tilde{\chi}' = \sum_{i=1}^{k}\chi(q_i)$ as well. Note first that another way to look at the meaning of the constraint $\sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_iq_j) \geq 0$ is to observe that the relation

$$\sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_iq_j) \geq 0 \tag{4.197}$$

can be rewritten in terms of the vector $\bar{v} = \sum_{i=1}^{k}\beta_i\mu^G(q_i)$. Recall that the partial sum of χ with respect to the set $u \in \mathcal{P}$ is the set function χ^u that satisfies

$$\chi^u(q) = \chi(q \cap u). \tag{4.198}$$

Thus

$$\sum_{i=1}^{k}\beta_i\chi(q_iq_j) = \sum_{i=1}^{k}\beta_i\chi^{q_j}(q_i) \Rightarrow \tag{4.199}$$

$$\sum_{i,j=1}^{k}\beta_i\beta_j\chi(q_iq_j) = \tag{4.200}$$

$$\sum_{i=1}^{k}\beta_i\sum_{j=1}^{k}\beta_j\chi(q_iq_j) = \tag{4.201}$$

$$\sum_{i=1}^{k}\beta_i\bar{v}^T\chi^{q_i} = \sum_{i=1}^{k}\beta_i v^T(\tilde{\chi}')^{q_i} \tag{4.202}$$

where $(\tilde{\chi}')^{q_i}$ is the projection of χ^{q_i} on the $\mathcal{Q}'$ coordinates. So while positive semidefiniteness does not tell us anything conclusive about $v^T\tilde{\chi}'$, it does tell us something about the inner product of v with the partial sums, i.e. it tells us that

$$\sum_{i=1}^{k}\beta_i v^T(\tilde{\chi}')^{q_i} \geq 0. \tag{4.203}$$

(So in particular, where each $\beta_i = 1$, then this says that while positive semidefiniteness does not imply that $v^T\tilde{\chi}' \geq 0$, it does give the weaker result that the sum of the inner products of the partial sums χ^{q_i} with v is nonnegative.) A special case that gives rise to a simple relation between this sum and v^T itself is where one of the sets q_i is P. Note that $\chi^P = \chi$ and therefore, where we write $q_p = P$ we have

$$\beta_p v^T \tilde{\chi}' + \sum_{i=1,\ldots,k, i \neq p} \beta_i v^T (\tilde{\chi}')^{q_i} \geq 0. \tag{4.204}$$

This gives the following lemma.

Lemma 4.21 *Let $\mathcal{Q}, \mathcal{Q}', G, \tilde{\chi}, \tilde{\chi}'$, and $U^{\tilde{\chi}}$ all be as in Theorem 4.20, with $U^{\tilde{\chi}} \succeq 0$. Let $v \in R^{\mathcal{Q}'}$; let $\bar{v} \in R^G$ be obtained by padding v with zeroes, and assume that there exists some collection of sets $\{q_1, \ldots, q_t\} \subseteq \mathcal{P}$ that includes P, with P denoted as q_P, such that*

$$\bar{v} = \sum_{i=1,\ldots,t} \beta_i \mu^G(q_i). \tag{4.205}$$

Let χ be a signed measure on $\mathcal{P}$ that is consistent with $\tilde{\chi}$. Let $(\tilde{\chi}')^{q_i}$ be the projection of χ^{q_i} on the $\mathcal{Q}'$ coordinates. If, say $\beta_p > 0$ then if

$$\sum_{i=1,\ldots,k, i \neq p} \beta_i v^T (\tilde{\chi}')^{q_i} \leq 0 \tag{4.206}$$

then we can conclude that $v^T \tilde{\chi}' \geq 0$. $\square$

A special subcase (the easiest one) is where $\{q_1, \ldots, q_t\} = \mathcal{Q}'$. In this case the multipliers β_i are readily available - they are just the v_i, and $(\tilde{\chi}')^{q_i}$ is the vector in $R^{\mathcal{Q}'}$ whose q_j entry is

$$\chi_{q_i \cap q_j} = \tilde{\chi}_{q_i \cap q_j} = U^{\tilde{\chi}}_{q_i, q_j} \tag{4.207}$$

i.e. it is the q_i'th column of $U^{\tilde{\chi}}$. This also means that we do not need the assumption of $\mathcal{P}$-signed-measure consistency for this case, as we never need to make reference to any values of χ outside of $\mathcal{Q}$. In particular, if $v_p = \beta_p > 0$, all other $v_i = \beta_i \leq 0$ and the inner product of v with those columns B_i such that $v_i = \beta_i < 0$ is nonnegative then we will know that

$$\sum_{i=1,\ldots,k, i \neq p} v^T (\tilde{\chi}')^{q_i} \leq 0 \tag{4.208}$$

and therefore that $v^T \tilde{\chi}' \geq 0$. This is the case of Lovasz and Schrijver's Lemma 1.5 ([LS91]). This characterization can be used to show easily that positive semidefiniteness (in addition to the N constraints) implies that the clique, odd antihole, and wheel inequalities are satisfied, and they do so in their work.

4.1.5 Positive Semidefiniteness and the N Operator

We observed already that in the case of stable set, had the collection $\mathcal{Q}'$ indexing the rows and columns of the matrix $U^{\tilde{\chi}}$ included the intersections of pairs of sets Y_i, then so long as $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent, the odd hole, odd antihole, and odd wheel constraints would have been implied directly by $U^{\tilde{\chi}} \succeq 0$, without recourse to the N constraints (as is the case for the clique inequalities). This raises the question of how much in fact is accomplished by imposing the N constraints over and above what would be accomplished by $\mathcal{P}$-signed-measure consistency and positive semidefiniteness (which corresponds essentially to a generalization of $TH(G)$ as we noted above after Corollary 4.8) alone?

In the stable set case, all of the first iteration N constraints are actually themselves delta vector constraints involving intersections of up to two sets Y_i. Let $\mathcal{Q}$ and $\mathcal{Q}'$ be as in (4.94); let $\bar{\chi}$ with a coordinate for every intersection of up to four sets Y_i, with projections $\tilde{\chi} \in R^{\mathcal{Q}}$ and $\tilde{\chi}' \in R^{\mathcal{Q}'}$, be consistent with a $\mathcal{P}$-signed measure χ. The constraints underlying N for the stable set problem (i.e. those which define the polytope $\bar{K}$ defined in Section 4.1.3) are of the forms

$$\chi_i + \chi_j \leq \chi_P, \text{ and} \tag{4.209}$$

$$0 \leq \chi_i \leq \chi_P. \tag{4.210}$$

Both of these constraints only involve the signed measures of sets of the form Y_i. Constraint (4.209) is just a delta vector constraint

$$(\mu^{\mathcal{Q}'}(Y_i^c Y_j^c))^T \tilde{\chi}' \geq 0 \tag{4.211}$$

since

$$\chi(Y_i^c Y_j^c) = \chi_P - \chi_i - \chi_j + \chi_{i,j} = \chi_P - \chi_i - \chi_j \tag{4.212}$$

by $\mathcal{P}$-signed-measure consistency, and (4.210) represents the two delta vector constraints

$$(\mu^{\mathcal{Q}'}(Y_i))^T \tilde{\chi}' \geq 0 \text{ and } (\mu^{\mathcal{Q}'}(Y_i^c))^T \tilde{\chi}' \geq 0 \tag{4.213}$$

since $\chi(Y_i) = \chi_i$ and $\chi(Y_i^c) = \chi_P - \chi_i$. As we saw in Section 4.1.3 (and using the same notation), the constraints that are added in the first iteration of $N(\bar{K})$ are

$$\chi_{i,j} + \chi_{i,k} \leq \chi_i, \ \{j,k\} \in E \tag{4.214}$$

$$\chi_{i^c,j} + \chi_{i^c,k} \leq \chi_{i^c}, \ \{j,k\} \in E \tag{4.215}$$

$$0 \leq \chi_{i,j} \leq \chi_i, \text{ and } 0 \leq \chi_{i^c,j} \leq \chi_{i^c} \tag{4.216}$$

These are also just delta vector constraints

$$(\mu^{\mathcal{Q}}(Y_i Y_j^c Y_k^c))^T \tilde{\chi} \geq 0$$
$$(\mu^{\mathcal{Q}}(Y_i^c Y_j^c Y_k^c))^T \tilde{\chi} \geq 0$$
$$(\mu^{\mathcal{Q}}(Y_i Y_j))^T \tilde{\chi} \geq 0, \quad \text{and} \quad (\mu^{\mathcal{Q}}(Y_i Y_j^c))^T \tilde{\chi} \geq 0$$
$$(\mu^{\mathcal{Q}}(Y_i^c Y_j))^T \tilde{\chi} \geq 0, \quad \text{and} \quad (\mu^{\mathcal{Q}}(Y_i^c Y_j^c))^T \tilde{\chi} \geq 0$$

Inequalities (4.214), (4.215), and (4.216) all entail intersections of no more than two sets Y_i, and therefore by Corollary 4.5, if $\bar{\chi}$ is $\mathcal{P}$-signed-measure consistent and $U^{\bar{\chi}} \succeq 0$, where $U^{\bar{\chi}}$ is the matrix with rows and columns indexed by $\mathcal{Q}$, then (4.214), (4.215), and (4.216) are all satisfied by $\tilde{\chi}$. Thus we were to only enforce positive semidefiniteness and $\mathcal{P}$-signed-measure consistency and not bother with the N constraints, we would still obtain all valid inequalities on $N(\bar{K})$, but one "iteration later", in the sense that the matrix would need to be indexed by $\mathcal{Q}$ rather than by $\mathcal{Q}'$. Before we generalize this characterization we will first prove a claim.

Claim 4.22 *Given any $G \subseteq \mathcal{P}$, $q, t \in \mathcal{P}$ and any delta vector $\mu^G(q)$, if we define*

$$G' = \{g' \in \mathcal{P} : g' = g \cap t, \text{ for some } g \in G\} \tag{4.217}$$

then there exists a delta vector $\mu^{G'}(q \cap t)$.

Proof: The basic idea is that considering that $\mu^G(q)$ can be considered to be a collection of multipliers corresponding to the listing of the points of each set $g \in G$ and yielding a listing of the points in q, if the multipliers corresponding to each set $g \in G$ are assigned instead to $g \cap t$ then a listing of the points in $q \cap t$ will be obtained. Formally, for all $r \in \mathcal{S}^P$, where the expression $(\zeta^r)^t$ means the partial sum of ζ^r taken over t, and has value $\zeta^r_{v\cap t}$ in each v'th coordinate, the expression $(\zeta^r)^t\{G\}$ is the projection of $(\zeta^r)^t$ on its G coordinates, and where we refer to the vector $\mu^G(q)$ as μ for short,

$$\mu^T(\zeta^r)^t\{G\} = (\zeta^r)^t_q = \zeta^r_{q\cap t} \tag{4.218}$$

where the first equality follows from the fact that partial summation is $\mathcal{P}$-signed-measure preserving and the second follows from Lemma 3.66. But

$$\mu^T(\zeta^r)^t\{G\} = \sum_{g\in G} \mu_g \zeta^r_{g\cap t} = \sum_{g'\in G'} \left(\sum_{g\in G: g\cap t = g'} \mu_g \right) \zeta^r_{g'} = (\mu')^T \zeta^r\{G'\} \tag{4.219}$$

where $\mu' \in R^{G'}$ with each $\mu_{g'} = \sum_{g\in G: g\cap t=g'} \mu_g$. Since this holds for all $r \in \mathcal{S}^P$ we conclude that μ' is of the form $\mu^{G'}(q \cap t)$, and thus that such vectors exist. $\Box$

Lemma 4.23 *Let $P \subseteq \{0,1\}^n$. Given a collection of sets $\mathcal{Q}' = \{q_1, \ldots, q_h\} \subseteq \mathcal{P}$, define the collection $(\mathcal{Q}')^k$ by*

$$(\mathcal{Q}')^k = \{q \in \mathcal{P} : q = \bigcap_{j=1\ldots,k} q_j,\ q_j \in \mathcal{Q}' \text{ (not necessarily distinct)}\} \tag{4.220}$$

for all nonnegative integers k, i.e. (4.220) is the collection of all $\leq k$-fold intersections of sets from $\mathcal{Q}'$. Given a vector $\tilde{y}^{2k} \in R^{(\mathcal{Q}')^{2k}}$, let $U^{\tilde{y}^{2k}}$ denote the $|(\mathcal{Q}')^k| \times |(\mathcal{Q}')^k|$ matrix whose u, v entry is $\tilde{y}^{2k}_{u \cap v}$. Suppose now that there exists a vector $v \in R^{\mathcal{Q}'}$ such that for all $\mathcal{P}$-signed-measure consistent vectors $\tilde{\chi}^2 \in R^{(\mathcal{Q}')^2}$ that satisfy

$$U^{\tilde{\chi}^2} \succeq 0 \tag{4.221}$$

we have

$$v^T \tilde{\chi}^1 \geq 0 \tag{4.222}$$

where $\tilde{\chi}^1$ is the projection of $\tilde{\chi}^2$ on $R^{\mathcal{Q}'}$. Then for all $\mathcal{P}$-signed-measure consistent vectors $\tilde{\chi}^{2k+2} \in R^{(\mathcal{Q}')^{2k+2}}$ that satisfy

$$U^{\tilde{\chi}^{2k+2}} \succeq 0 \tag{4.223}$$

we must also have

$$v^T (\tilde{\chi}^1)^s \geq 0 \tag{4.224}$$

where $s \in \mathcal{P}$ is any set for which there exists a delta vector $\mu^{(\mathcal{Q}')^k}(s)$, and $(\tilde{\chi}^1)^s$ is the projection of the partial sum

$$(\tilde{\chi}^{k+1})^s = U^{\tilde{\chi}^{2k+2}} \mu^{(\mathcal{Q}')^{k+1}}(s) \tag{4.225}$$

on the $\mathcal{Q}'$ coordinates.

Observe that where $\mathcal{Q}' = \{\{0,1\}^n, Y_1, \ldots, Y_n\}$, and where s is a k-fold intersection of sets of the form $q : q \in \mathcal{Q}'$ or $q^c \in \mathcal{Q}'$ (this is an appropriate form for s by Remark 3.50), the constraints $v^T(\tilde{\chi}^1)^s \geq 0$ are the type of constraints that define the $\bar{N}^k$ operator (cf. Remark 3.68). Observe also that (4.225) follows from Lemma 3.64

Proof: If for all $\mathcal{P}$-signed-measure consistent $\tilde{\chi}^2$ for which $U^{\tilde{\chi}^2} \succeq 0$ we must also have $v^T \tilde{\chi}^1 \geq 0$, then there must exist some inequalities

$$0 \leq \sum_{i,j=1}^{h} \alpha_i^l \alpha_j^l \tilde{\chi}^2(q_i q_j) = (a^l)^T \tilde{\chi}^2 \tag{4.226}$$

and some equalities[3]

$$\left(\mu^{(\mathcal{Q}')^2}(q)\right)^T \tilde{\chi}^2 = \tilde{\chi}^2_q \text{ or stated more briefly, } (u^i)^T\tilde{\chi}^2 = 0 \tag{4.227}$$

such that where $\bar{v}^2$ is the lifting of v to $R^{(\mathcal{Q}')^2}$ obtained by padding with zeroes, there exist numbers $\lambda_l \geq 0$ and γ_i (unrestricted) such that

$$\sum_i \gamma_i u^i + \sum_l \lambda_l a^l = \bar{v}^2 \tag{4.228}$$

so that for all $\tilde{y}^2 \in R^{(\mathcal{Q}')^2}$ and projections $\tilde{y}^1 \in R^{\mathcal{Q}'}$,

$$\sum_i \gamma_i (u^i)^T\tilde{y}^2 + \sum_l \lambda_l (a^l)^T\tilde{y}^2 = (\bar{v}^2)^T\tilde{y}^2 = v^T\tilde{y}^1. \tag{4.229}$$

Consider now the vector

$$w^l = \sum_{i=1}^h \alpha_i^l \mu^{(\mathcal{Q}')^{k+1}}(q_i \cap s) \tag{4.230}$$

(note that $\mu^{(\mathcal{Q}')^{k+1}}(q_i \cap s)$ exists by Claim 4.22). Then for any $\mathcal{P}$-signed-measure χ consistent with $\tilde{\chi}^{2k+2}$, if $U^{\tilde{\chi}^{2k+2}} \succeq 0$,

$$0 \leq (w^l)^T U^{\tilde{\chi}^{2k+2}} w^l = \sum_{i,j=1}^h \alpha_i^l \alpha_j^l \chi(q_i q_j s) = (a^l)^T(\tilde{\chi}^2)^s \tag{4.231}$$

where $(\tilde{\chi}^2)^s$ is the projection of the partial sum χ^s of χ on $R^{(\mathcal{Q}')^2}$. Since partial summation is trivially $\mathcal{P}$-signed-measure preserving, every $\mathcal{P}$-signed measure consistency equality of the form $u^T\chi = 0$, where χ is a $\mathcal{P}$-signed measure, holds for each partial sum χ^s as well (compare to Lemma 3.66 and Corollary 3.67). Thus for each i we must also have $(u^i)^T(\tilde{\chi}^2)^s = 0$. Therefore

$$0 \leq \sum_i \gamma_i (u^i)^T(\tilde{\chi}^2)^s + \sum_l \lambda_l (a^l)^T(\tilde{\chi}^2)^s = (\bar{v}^2)^T(\tilde{\chi}^2)^s = v^T(\tilde{\chi}^1)^s. \tag{4.232}$$

(Observe that for any $\mathcal{P}$-signed-measure χ consistent with $\tilde{\chi}^{2k+2}$, the projection of the partial sum χ^s on $R^{(\mathcal{Q}')^{k+1}}$ is just $(\tilde{\chi}^{k+1})^s = U^{\tilde{\chi}^{2k+2}}\mu^{(\mathcal{Q}')^{k+1}}(s)$, and thus $(\tilde{\chi}^1)^s$ in expression (4.224) is just (as the notation implies) the projection of χ^s on $R^{\mathcal{Q}'}$, and so it is the projection of $(\tilde{\chi}^2)^s$ on $R^{\mathcal{Q}'}$ as well.) □

Thus if we have $\mathcal{P}$-signed-measure consistency and we are to enforce positive semidefiniteness, it will be most effective to introduce additional linear inequalities and then follow

[3] Observe from Lemma 3.42 that the constraints that establish $\mathcal{P}$-signed-measure consistency are all of this form. The fact that these constraints are all equalities set to zero also follows from the definition of the $\mathcal{P}$-signed-measure consistent vectors as a subspace.

an N type procedure where the inequality is successively applied to partial sums, if those inequalities have high positive semidefinite rank, in the sense that they are hard to derive using positive semidefiniteness and $\mathcal{P}$-signed-measure consistency alone. (Note that this result is not particular to $\mathcal{P}$-signed-measure consistency per se. The same methodology shows that if a constraint is implied by other constraints and positive semidefiniteness, then enforcing the constraints on the partial sums and positive semidefiniteness on the larger matrix also implies that the original constraint applied to the partial sums will be satisfied.)

We have seen that in the case of stable set, the covering inequalities

$$\chi_i + \chi_j \leq \chi_P, \; i, j : \{i, j\} \in E \tag{4.233}$$

were implied by positive semidefiniteness and $\mathcal{P}$-signed-measure consistency alone, i.e. these inequalities have "low positive semidefinite rank" in the manner described in the previous paragraph. We will soon give a general characterization of the positive semidefinite "rank" of covering constraints, but first we will prove a lemma.

Lemma 4.24 *Let $\{q_i : i = 1, \ldots, h\} \subseteq \mathcal{P}$. Let $\mathcal{Q}' \subseteq \mathcal{P}$ include P and all intersections of up to k sets q_i. Let $\mathcal{Q} \subseteq \mathcal{P}$ be such that for all $u, v \in \mathcal{Q}'$ we have $u \cap v \in \mathcal{Q}$. Let $\tilde{\chi} \in R^{\mathcal{Q}}$ with projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$, and let $U^{\tilde{\chi}}$ be the matrix with rows and columns indexed by $\mathcal{Q}'$ with u, v entry equal to $\tilde{\chi}(u \cap v)$. Then $U^{\tilde{\chi}} \succeq 0$ implies that the constraint*

$$\sum_{i=1}^{k} \tilde{\chi}'(q_i) \leq \tilde{\chi}'(q_1 q_2 \cdots q_k) + (k-1)\tilde{\chi}'_P \tag{4.234}$$

is satisfied (regardless of whether or not $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent).

Proof: By induction on k; let $k = 2$, consider the delta vector inequality

$$0 \leq \tilde{\chi}'(q_1^c q_2^c) = \tilde{\chi}'_P - \tilde{\chi}'(q_1) - \tilde{\chi}'(q_2) + \tilde{\chi}'(q_1 q_2) \Rightarrow \tag{4.235}$$

$$\tilde{\chi}'(q_1) + \tilde{\chi}'(q_2) \leq \tilde{\chi}'(q_1 q_2) + \tilde{\chi}'_P. \tag{4.236}$$

This constraint is implied by the positive semidefiniteness constraint $v^T U^{\tilde{\chi}} v \geq 0$ where v is the vector indexed by $\mathcal{Q}'$ with a 1 in its P position, -1 in its q_i and q_j positions, and zeroes elsewhere. Now assume that the lemma holds for arbitrary k, then

$$\sum_{i=1}^{k+1} \tilde{\chi}'(q_i) - k\tilde{\chi}'_P = \sum_{i=1}^{k} \tilde{\chi}'(q_i) - (k-1)\tilde{\chi}'_P + \tilde{\chi}'(q_{k+1}) - \tilde{\chi}'_P \leq \tag{4.237}$$

$$\tilde{\chi}'(q_1 q_2 \cdots q_k) + \tilde{\chi}'(q_{k+1}) - \tilde{\chi}'_P \leq \tag{4.238}$$

$$\tilde{\chi}'(q_1 q_2 \cdots q_k q_{k+1}) \tag{4.239}$$

where both inequalities are from the induction. □

Now we are ready for the generalization.

Theorem 4.25 *Let $\{q_i : i = 1, \ldots, h\} \subseteq \mathcal{P}$, and let $\mathcal{Q}' \subseteq \mathcal{P}$ include P and all intersections of up to $k-1$ sets q_i. Let $\tilde{\chi}$, $\tilde{\chi}'$ and $U^{\tilde{\chi}}$ be as in Lemma 4.24, with $U^{\tilde{\chi}} \succeq 0$. Then for any "forbidden configuration"*

$$q_1 q_2 \cdots q_k = \emptyset \tag{4.240}$$

if $\tilde{\chi}(\emptyset) = 0$, then the covering constraint

$$\sum_{i=1}^{k} \tilde{\chi}'(q_i) \leq (k-1)\tilde{\chi}'_P \tag{4.241}$$

is satisfied.

Proof: Note first that if we choose $u = q_1 \cdots q_{k-1} \in \mathcal{Q}'$ and $v = q_k \in \mathcal{Q}'$, then $u \cap v = q_1 q_2 \cdots q_k = \emptyset \in \mathcal{Q}$, so $\emptyset$ is one of the sets indexing $\tilde{\chi}$, and the expression $\tilde{\chi}(\emptyset)$ makes sense. Moreover we have $\tilde{\chi}(q_1 q_2 \cdots q_k) = \tilde{\chi}(\emptyset) = 0$ by assumption (observe that this would have been a consequence of $\mathcal{P}$-signed-measure consistency had we required it). As in the proof of Lemma 4.24, by positive semidefiniteness

$$0 \leq \tilde{\chi}'_P - \tilde{\chi}'(q_1 \cdots q_{k-1}) - \tilde{\chi}'(q_k) + \tilde{\chi}'(q_1 \cdots q_k) = \tilde{\chi}'_P - \tilde{\chi}'(q_1 \cdots q_{k-1}) - \tilde{\chi}'(q_k) \tag{4.242}$$

by hypothesis, so the delta vector inequality

$$\chi_P \geq \chi(q_1 q_2 \cdots q_{k-1} \cup q_k) = \chi(q_1 q_2 \cdots q_{k-1}) + \chi(q_k) \tag{4.243}$$

is satisfied by $\tilde{\chi}'$. Thus by Lemma 4.24

$$\sum_{i=1}^{k} \tilde{\chi}'(q_i) = \sum_{i=1}^{k-1} \tilde{\chi}'(q_i) + \tilde{\chi}'(q_k) \leq \tag{4.244}$$

$$\tilde{\chi}'(q_1 q_2 \cdots q_{k-1}) + (k-2)\tilde{\chi}'_P + \tilde{\chi}'(q_k) \leq (k-1)\tilde{\chi}'_P. \; \square \tag{4.245}$$

4.2 Positive Semidefiniteness and Measure Preserving Operators

One possible way to use positive semidefiniteness to greater effect is by way of measure preserving operators. If χ is meant to be measure consistent then we must have $U^{T\tilde{\chi}} \succeq 0$

for every measure preserving operator T. (Observe that this constraint depends only on measure consistency and remains valid even if T does not preserve $\mathcal{P}$-measure consistency.) Let $\mathcal{Q}'' \subseteq \mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$, with $\mathcal{Q}' \supseteq \{u \cap v : u, v \in \mathcal{Q}''\}$, $\mathcal{Q} \supseteq \{u \cap v : u, v \in \mathcal{Q}'\}$, and let $\tilde{\chi} \in R^{\mathcal{Q}}$. Consider, for example, the $\mathcal{P}$-measure preserving operator $T(\tilde{\chi}) = U^{\tilde{\chi}}\nu$ where $\nu \in R^{\mathcal{Q}'}$ and $\nu^T\tilde{\chi}' \geq 0$ is valid for all $\mathcal{P}$-measure consistent $\tilde{\chi}' \in R^{\mathcal{Q}'}$. Then where $U^{U^{\tilde{\chi}}\nu}$ is the matrix with rows and columns indexed by $\mathcal{Q}''$ with u, v entry $(U^{\tilde{\chi}}\nu)_{u\cap v}$, then for any $\mathcal{P}$-measure consistent $\tilde{\chi}$, by Lemma 4.1 we must have

$$U^{U^{\tilde{\chi}}\nu} \succeq 0. \tag{4.246}$$

This is essentially the Lasserre algorithm ([Las01], see also [Lau01]), generalized to our expanded framework. The following lemma shows that this constraint shares a similar relationship with the linear constraints

$$\nu^T(\tilde{\chi}')^q = \nu^T U^{\tilde{\chi}}\mu(q) \geq 0 \tag{4.247}$$

(where $\mu(q)$ is a delta vector using only $\mathcal{Q}''$ coordinates) as does the constraint $U^{\tilde{\chi}} \succeq 0$ with the constraints

$$\chi_q = (\mu(q))^T\tilde{\chi}' = (\mu(q))^T U^{\tilde{\chi}}\mu(q) \geq 0 \tag{4.248}$$

(where $\mu(q)$ is a delta vector using only $\mathcal{Q}'$ coordinates). It is essentially a "valid constraints" version of Lemma 4.4.

Lemma 4.26 *Let $\mathcal{Q}'' \subseteq \mathcal{Q}' \subseteq \mathcal{Q} \subseteq \mathcal{P}$, with $\mathcal{Q}' \supseteq \{u\cap v : u, v \in \mathcal{Q}''\}$, $\mathcal{Q} \supseteq \{u\cap v : u, v \in \mathcal{Q}'\}$; let $\nu \in R^{\mathcal{Q}'}$, and let $\tilde{\chi} \in R^{\mathcal{Q}}$ with projection $\tilde{\chi}' \in R^{\mathcal{Q}'}$ be $\mathcal{P}$-signed-measure consistent. Given any vector $\tilde{y} \in R^{\mathcal{Q}}$ let $U^{\tilde{y}}$ denote the matrix with rows and columns indexed by $\mathcal{Q}'$ with u, v entry $\tilde{y}_{u\cap v}$, and given any vector $\tilde{y}' \in R^{\mathcal{Q}'}$ let $U^{\tilde{y}'}$ denote the matrix with rows and columns indexed by $\mathcal{Q}''$ with u, v entry $\tilde{y}'_{u\cap v}$. For any delta vector $\mu^{\mathcal{Q}''}(q)$, if*

$$U^{U^{\tilde{\chi}}\nu} \succeq 0 \tag{4.249}$$

then

$$\nu^T(\tilde{\chi}')^q = \nu^T U^{\tilde{\chi}}\mu^{\mathcal{Q}'}(q) \geq 0. \tag{4.250}$$

Proof:

$$\tilde{\chi} = \sum_{r\in\mathcal{S}^P} \alpha_r\zeta^r\{\mathcal{Q}\} \Rightarrow U^{\tilde{\chi}} = \sum_{r\in\mathcal{S}^P} \alpha_r\zeta^r\{\mathcal{Q}'\}(\zeta^r\{\mathcal{Q}'\})^T \Rightarrow \tag{4.251}$$

$$U^{\tilde{\chi}}\nu = \sum_{r\in\mathcal{S}^P} \left(\alpha_r\nu^T\zeta^r\{\mathcal{Q}'\}\right)\zeta^r\{\mathcal{Q}'\} \Rightarrow \tag{4.252}$$

$$U^{U^{\tilde{\chi}}\nu} = \sum_{r\in\mathcal{S}^P} \left(\alpha_r \nu^T \zeta^r\{\mathcal{Q}'\}\right) \zeta^r\{\mathcal{Q}''\}(\zeta^r\{\mathcal{Q}''\})^T \Rightarrow \tag{4.253}$$

$$0 \leq \left(\mu^{\mathcal{Q}''}(q)\right)^T U^{U^{\tilde{\chi}}\nu} \mu^{\mathcal{Q}''}(q) = \sum_{r\in\mathcal{S}^P : r\subseteq q} \alpha_r \nu^T \zeta^r\{\mathcal{Q}'\} = \nu^T(\tilde{\chi}')^q \ \square \tag{4.254}$$

The next lemma shows that where $\mathcal{Q}, \mathcal{Q}', \mathcal{Q}''$ and the matrices $U^{\tilde{y}}$ and $U^{\tilde{y}'}$ are all defined as in the previous lemma, then if there is some collection $\bar{\mathcal{Q}}'' \subseteq \mathcal{Q}'$ for which $\bar{q}'' \cap q'' \in \mathcal{Q}'$ for every $\bar{q}'' \in \bar{\mathcal{Q}}''$, $q'' \in \mathcal{Q}''$, and if the constraint (4.246) is applied with $\nu = \mu(u)$ for a delta vector $\mu(u) \in R^{\mathcal{Q}'}$ (i.e. regular partial summation) such that $\mu(u)$ has nonzero entries only in its $\bar{\mathcal{Q}}''$ coordinates, then the positive semidefiniteness constraint (4.246) does not strengthen the condition $U^{\tilde{\chi}} \succeq 0$.

Lemma 4.27 *Let $\mathcal{Q}$, $\mathcal{Q}'$, $\mathcal{Q}''$, $\tilde{\chi}$, $\tilde{\chi}'$, $\tilde{\chi}''$ and the matrices of the form $U^{\tilde{y}}$ and $U^{\tilde{y}'}$ all be defined as in Lemma 4.26. Let $\bar{\mathcal{Q}}'' \subseteq \mathcal{Q}'$ satisfy $\bar{q}'' \cap q'' \in \mathcal{Q}'$ for every $\bar{q}'' \in \bar{\mathcal{Q}}''$, $q'' \in \mathcal{Q}''$, and let $\mu^{\mathcal{Q}'}(u)$ be a delta vector with all of its nonzeroes located in its $\bar{\mathcal{Q}}''$ coordinates. Then*

$$U^{\tilde{\chi}} \succeq 0 \Rightarrow U^{U^{\tilde{\chi}}\mu^{\mathcal{Q}'}(u)} \succeq 0. \tag{4.255}$$

Proof: By Claim 4.22, for each $q'' \in \mathcal{Q}''$ there is a vector $\mu^{\mathcal{Q}'}(u \cap q'')$ since by assumption there exists a vector $\mu^{\bar{\mathcal{Q}}''}(u)$ (namely the projection of $\mu^{\mathcal{Q}'}(u)$ on $R^{\bar{\mathcal{Q}}''}$), and

$$\{q \in \mathcal{P} : q = \bar{q}'' \cap q'' \text{ for some } \bar{q}'' \in \bar{\mathcal{Q}}''\} \subseteq \mathcal{Q}'. \tag{4.256}$$

Let v be any vector in $R^{\mathcal{Q}''}$. Observe that

$$\sum_{q''\in\mathcal{Q}''} v_{q''} \mu^{\mathcal{Q}'}(u \cap q'') \tag{4.257}$$

is a vector in $R^{\mathcal{Q}'}$. Where $(\tilde{\chi}')^u = U^{\tilde{\chi}}\mu^{\mathcal{Q}'}(u)$ has a coordinate for each $q' \in \mathcal{Q}'$ with value $\tilde{\chi}_{q'\cap u}$, and $U^{(\tilde{\chi}')^u} = U^{U^{\tilde{\chi}}\mu^{\mathcal{Q}'}(u)}$ is the matrix with rows and columns indexed by $\mathcal{Q}''$ with s, t entry $(\tilde{\chi}')^u_{s\cap t} = \tilde{\chi}_{s\cap t\cap u}$, we therefore have

$$0 \leq \left(\sum_{s\in\mathcal{Q}''} v_s \mu^{\mathcal{Q}'}(u \cap s)\right)^T U^{\tilde{\chi}} \sum_{s\in\mathcal{Q}''} v_s \mu^{\mathcal{Q}'}(u \cap s) = \tag{4.258}$$

$$\sum_{s\in\mathcal{Q}''} \sum_{t\in\mathcal{Q}''} v_s v_t \tilde{\chi}_{s\cap t\cap u} = v^T U^{(\tilde{\chi}')^u} v = v^T U^{U^{\tilde{\chi}}\mu^{\mathcal{Q}'}(u)} v. \ \square \tag{4.259}$$

This lemma proves Theorem 2.18, as in that case $\mathcal{Q}'$ is the collection of l-tuples of $\{0,1\}^n$ and Y_i, $\mathcal{Q}''$ is the collection of 1-tuples (i.e. $\{\{0,1\}^n, Y_1, \ldots, Y_n\}$), and $\bar{\mathcal{Q}}''$ is the collection of $(l-1)$-tuples.

4.3 When Does $\mathcal{A}$-Measure-Consistency Help?

The question of when the N^+ operator is stronger than N has been treated already by Goemans and Tunçel ([GT01]) (see also [CD01] and [CL01]). In this section we will shift the question to when does measure consistency (i.e. $\mathcal{A}$-measure consistency, see Definitions 3.2 and 3.23) help, and we will thereby broaden some of their results and give measure theoretic insight into why they hold. Our efforts here will focus primarily on a measure consistency supplemented $\bar{N}$-type paradigm (which is a strengthening of N^+), but we will also describe a theoretical situation for which a similar strengthening of the Lasserre operator would not help either.

Our first step will be to try to develop some geometric intuition into the nature of measure consistency within the framework of an N-type procedure. This intuition will be helpful in understanding where and why requiring measure consistency does not strengthen N.

4.3.1 The Geometry of Measure Consistency

Observe first that in using $U^{\tilde{\chi}} \succeq 0$ to imply delta vector constraints in the first sections of the chapter, we needed throughout the assumption that $\tilde{\chi}$ is $\mathcal{P}$-signed-measure consistent. In the absence of any such assumption, and given an arbitrary vector $\tilde{\chi}$ ordered by general set theoretic expressions, there is much less to be said. Even in the case where the expressions, when construed as being of sets of the form Y_i (see Remark 3.28), are known to define a linearly independent collection for $\mathcal{A}$ (such as in the case where they are all intersections of sets Y_i), positive semidefiniteness still only implies that for signed measures on $\mathcal{A}$ consistent with $\tilde{\chi}$, the signed measures of various sets in $\mathcal{A}$ are nonnegative. But this in any case only provides evidence that $\tilde{\chi}$ is consistent with an $\mathcal{A}$-measure, and not necessarily with a $\mathcal{P}$-measure. (Recall from Lemma 3.27 that for an $\mathcal{A}$-measure to correspond to a $\mathcal{P}$-measure requires $\chi(P^c) = 0$.) We saw in the previous chapter (Corollary 3.31) that if measure consistency is coupled with setting a "test vector" to an appropriate value, then measure consistency implies $\mathcal{P}$-measure consistency, but again, the test vector constraints there are crucial.

Before we go any further, we will illustrate the geometry of measure consistency with an example.

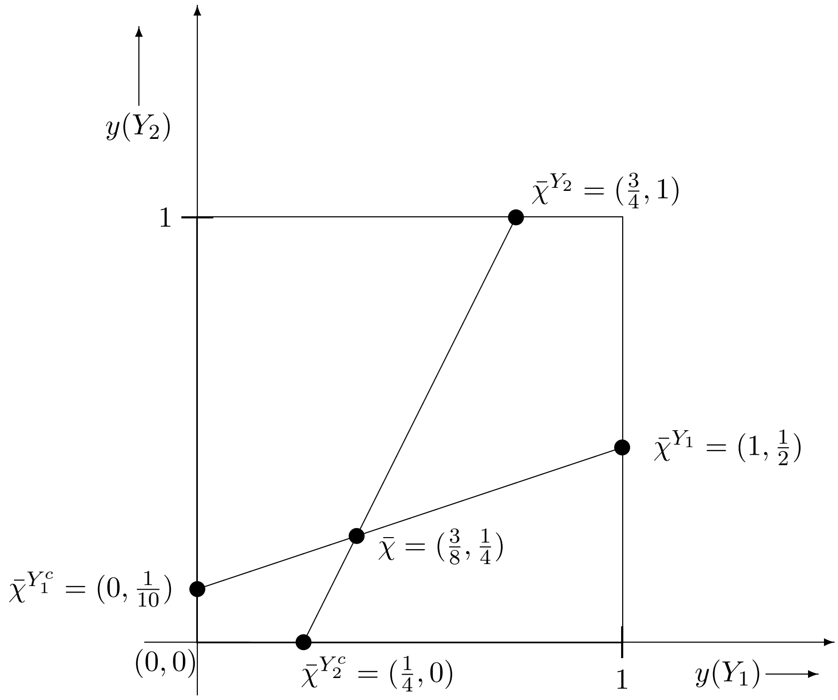

Figure 1

Aside from the differences in the labeling, this is Figure 3 of Chapter 1. In the diagram we have selected a point $\bar{\chi} \in [0,1]^2$. Considering that $\bar{\chi}$ belongs to the unit square, it is obvious that it may be written as a convex combination of the vertices $(1,1),\ (1,0),\ (0,1),\ (0,0)$ of the unit square, i.e. there is a choice of nonnegative numbers

$$\chi(Y_1 \cap Y_2),\ \chi(Y_1 \cap Y_2^c),\ \chi(Y_1^c \cap Y_2),\ \chi(Y_1^c \cap Y_2^c) \tag{4.260}$$

summing to the value 1 and such that

$$\bar{\chi} = \chi(Y_1 \cap Y_2)(1,1) + \chi(Y_1 \cap Y_2^c)(1,0) + \chi(Y_1^c \cap Y_2)(0,1) + \chi(Y_1^c \cap Y_2^c)(0,0). \tag{4.261}$$

Let us consider now these four numbers (4.260) to be the values of a probability measure χ on the four atomic sets $Y_1 \cap Y_2,\ Y_1 \cap Y_2^c,\ Y_1^c \cap Y_2,\ Y_1^c \cap Y_2^c$ respectively (of the algebra generated by Y_1 and Y_2). Then

$$\bar{\chi}(Y_1) = \chi(Y_1 \cap Y_2) + \chi(Y_1 \cap Y_2^c) = \chi(Y_1) \tag{4.262}$$

and

$$\bar{\chi}(Y_2) = \chi(Y_1 \cap Y_2) + \chi(Y_1^c \cap Y_2) = \chi(Y_2). \tag{4.263}$$

Recall that the partial sum χ^{Y_1} of the probability measure χ is a measure on the algebra generated by the sets Y_1 and Y_2, and if $\chi(Y_1) > 0$ then the normalized partial sum $\frac{\chi^{Y_1}}{\chi(Y_1)}$ is a probability measure. Before we continue, let us review a definition from probability theory (see Chapter 10 of [F99] for details).

Definition 4.28 *Let $\mathcal{X}$ be a probability measure defined on a σ-algebra $\mathcal{W}$ of subsets of a nonempty set Ω. Then given any set $Q \in \mathcal{W}$ with $\mathcal{X}(Q) > 0$, the conditional probability measure $\mathcal{X}|Q$ is defined by*

$$\mathcal{X}|Q(A) = \frac{\mathcal{X}(Q \cap A)}{\mathcal{X}(A)}, \quad \forall A \in \mathcal{W}. \tag{4.264}$$

Observe now that for any set q in the algebra generated by Y_1 and Y_2,

$$\frac{\chi^{Y_1}(q)}{\chi(Y_1)} = \frac{\chi(q \cap Y_1)}{\chi(Y_1)} \tag{4.265}$$

and so $\frac{\chi^{Y_1}}{\chi(Y_1)}$ is the conditional probability measure $\chi|Y_1$. Defining the vector

$$\bar{\chi}^{Y_1} = (\chi|Y_1(Y_1), \chi|Y_1(Y_2)) = \left(1, \frac{\chi(Y_1 \cap Y_2)}{\chi(Y_1)}\right), \tag{4.266}$$

observe that $\bar{\chi}^{Y_1}$ is just the normalized partial sum of the convex combination (4.261) taken over those vertices of the unit square that belong to Y_1, i.e.

$$\bar{\chi}^{Y_1} = \frac{1}{\chi(Y_1)}\left(\chi(Y_1 \cap Y_2)(1,1) + \chi(Y_1 \cap Y_2^c)(1,0)\right). \tag{4.267}$$

Similarly, where $\bar{\chi}^{Y_1^c}$ is defined by

$$\bar{\chi}^{Y_1^c} = (\chi|Y_1^c(Y_1), \chi|Y_1^c(Y_2)), \tag{4.268}$$

then $\bar{\chi}^{Y_1^c}$ is just the normalized partial sum of the convex combination (4.261) taken over those vertices of the unit square that belong to Y_1^c, and $\bar{\chi}$ is the convex combination

$$\bar{\chi} = \bar{\chi}(Y_1)\bar{\chi}^{Y_1} + (1 - \bar{\chi}(Y_1))\bar{\chi}^{Y_1^c}. \tag{4.269}$$

The diagram indicates a possible choice for $\bar{\chi}^{Y_1}$ and the consequent choice of $\bar{\chi}^{Y_1^c}$. Observe moreover that (where $\bar{\chi}^{Y_2}$ and $\bar{\chi}^{Y_2^c}$ are defined in the same manner as $\bar{\chi}^{Y_1}$ and $\bar{\chi}^{Y_1^c}$,) $\bar{\chi}^{Y_1}$ and $\bar{\chi}^{Y_2}$ are both determined by the choice of $\chi(Y_1 \cap Y_2)$ (and $\bar{\chi}$), so all four vectors $\bar{\chi}^{Y_1}$, $\bar{\chi}^{Y_1^c}$, $\bar{\chi}^{Y_2}$, $\bar{\chi}^{Y_2^c}$ are determined by $\bar{\chi}$ and the choice of $\chi(Y_1 \cap Y_2)$. The diagram shows the four vectors that would be determined by a (arbitrary) choice of $\chi(Y_1 \cap Y_2) = \frac{3}{16}$.

As we indicated, every choice of $\bar{\chi}$ in the unit square is consistent with some convex combination (with coefficients (4.260)) of the vertices of the unit square, and is thus consistent with the probability measure χ defined by (4.260). But obviously not every selection of

$\bar{\chi},\ \chi(Y_1 \cap Y_2)$ (and the four consequent conditional probability vectors $\bar{\chi}^{Y_1},\ \bar{\chi}^{Y_1^c},\ \bar{\chi}^{Y_2},\ \bar{\chi}^{Y_2^c}$) is compatible with a convex combination of the vertices (i.e. these vectors might not represent normalized partial sums of any convex combination of the vertices of the square) and therefore with a probability measure. For example had we chosen $\chi(Y_1 \cap Y_2) = \frac{9}{32}$ then we would have $\bar{\chi}^{Y_1}(Y_2) = \frac{3}{4}$ and it is easy to see from the diagram that this would imply that

$$\chi(Y_1^c Y_2) = \chi(Y_1^c)\bar{\chi}^{Y_1^c}(Y_2) = \frac{5}{8} \cdot \frac{-1}{20} = \frac{-1}{32} < 0. \tag{4.270}$$

The requirement that the choices of $\bar{\chi} = (\chi(Y_1), \chi(Y_2))$ and $\chi(Y_1, Y_2)$ be in fact probability measure consistent is thus the requirement that the consequent conditional probability vectors $\bar{\chi}^{Y_1},\ \bar{\chi}^{Y_1^c},\ \bar{\chi}^{Y_2},\ \bar{\chi}^{Y_2^c}$ (which are the convexifying vectors $v^1,\ w^1,\ v^2,\ w^2$ of Figure 3 from Chapter 1,) can actually correspond to some convex combination of the vertices of the cube that yields $\bar{\chi}$ (i.e. that they may represent normalized partial sums of that convex combination).

As a somewhat more instructive example, consider

$$\bar{\chi} = \left(\frac{5}{6}, \frac{1}{3}, \frac{3}{4}\right). \tag{4.271}$$

Again $\bar{\chi}$ can certainly be written as a convex combination of the vertices of the unit cube, and again any such convex combination would define a probability measure, and again the normalized partial sums

$$\bar{\chi}^{Y_1},\ \bar{\chi}^{Y_1^c},\ \bar{\chi}^{Y_2},\ \bar{\chi}^{Y_2^c},\ \bar{\chi}^{Y_3},\ \bar{\chi}^{Y_3^c}, \tag{4.272}$$

all of which are fixed by the values of $\chi(Y_1 \cap Y_2),\ \chi(Y_1 \cap Y_3)$ and $\chi(Y_2 \cap Y_3)$, represent conditional probabilities and are, in the notation of Chapter 1, the convexification vectors $v^1, w^1, v^2, w^2, v^3, w^3$. Here too, however, not every choice of $\bar{\chi}$ and $\chi(Y_1 \cap Y_2),\ \chi(Y_1 \cap Y_3)$ and $\chi(Y_2 \cap Y_3)$ (and the consequent choice of (4.272)) is compatible with a probability measure and therefore with a convex combination. Say, for example, that we have chosen

$$\chi(Y_1 \cap Y_2) = \frac{1}{6},\ \chi(Y_1 \cap Y_3) = \frac{3}{4},\ \chi(Y_2 \cap Y_3) = \frac{1}{4}. \tag{4.273}$$

This will fix

$$\bar{\chi}^{Y_1} = \left(1, \frac{1}{5}, \frac{9}{10}\right),\ \bar{\chi}^{Y_1^c} = (0, 1, 0) \tag{4.274}$$

$$\bar{\chi}^{Y_2} = \left(\frac{1}{2}, 1, \frac{3}{4}\right),\ \bar{\chi}^{Y_2^c} = \left(1, 0, \frac{3}{4}\right) \tag{4.275}$$

$$\bar{\chi}^{Y_3} = \left(1, \frac{1}{3}, 1\right),\ \bar{\chi}^{Y_3^c} = \left(\frac{1}{3}, \frac{1}{3}, 0\right). \tag{4.276}$$

Though all six of these points indeed belong to the unit cube, they are not consistent with any convex combination of the vertices of the cube yielding $\bar{\chi}$, i.e. they are not probability-measure consistent. To see this, let $\chi(Y_1 \cap Y_2 \cap Y_3)$ be denoted by θ. Then

$$\chi(Y_1 \cap Y_2 \cap Y_3^c) = \chi(Y_1 \cap Y_2) - \theta = \frac{1}{6} - \theta, \tag{4.277}$$

so that in order to have $\chi(Y_1 \cap Y_2 \cap Y_3^c) \geq 0$ we must have $\theta \leq \frac{1}{6}$, and

$$\chi(Y_1^c \cap Y_2^c \cap Y_3) = \chi(Y_1^c \cap Y_3) - \chi(Y_1^c \cap Y_2 \cap Y_3) = \chi(Y_3) - \chi(Y_1 \cap Y_3) - \chi(Y_2 \cap Y_3) + \theta = \theta - \frac{1}{4}, \tag{4.278}$$

so that in order to have $\chi(Y_1^c \cap Y_2^c \cap Y_3) \geq 0$ we must have $\theta \geq \frac{1}{4}$, which is a contradiction.

As we saw in Chapter 1, given some set $P \subseteq \{0,1\}^n$, and given some convex set $\bar{P} \subseteq [0,1]^n$ with $\bar{P} \cap \{0,1\}^n = P$, the convexification procedure applied to $\bar{P}$ requires there to be (for each $i \in \{1, \ldots, n\}$ for which $0 < \bar{\chi}(Y_i) < 1$) a choice of convexification vectors $\bar{\chi}^{Y_i}$ and $\bar{\chi}^{Y_i^c}$ lying on the hyperplanes $y(Y_i) = 1$ and $y(Y_i) = 0$ respectively, with $\bar{\chi}$ lying on each line connecting $\bar{\chi}^{Y_i}$ with $\bar{\chi}^{Y_i^c}$, and such that each vector $\bar{\chi}^{Y_i}$ and $\bar{\chi}^{Y_i^c}$ belongs to $\bar{P}$. Stated loosely (cf. Remark 3.68), the N operator (and more generally, the $\bar{N}$ operator) adds the requirement that these convexification vectors must be consistent with a choice of values $\chi(Y_i \cap Y_j)$. But as we have indicated, this is not in itself sufficient to ensure that the choice of convexification vectors is consistent with any convex combination of the vertices of even the hypercube, let alone with the subset of those vertices that constitutes P. We will now define a theoretical operator that is identical with the $\bar{N}$ operator, but which also requires that the choices of $\chi(Y_i \cap Y_j)$ (and more generally $\chi(\bigcap_i Y_i)$, for the case of $\bar{N}^k,\ k \geq 1$) must be measure consistent. (In line with all of the N type procedures, a coordinate $\chi(\{0,1\}^n)$ is also appended, and probability measure consistency can be ensured by the additional constraint $\chi(\{0,1\}^n) = 1$.)

Definition 4.29 *Define the operator $(N^{++})^k(\bar{K})$ to be the same as $\bar{N}^k(\bar{K})$, defined in Remark 3.68, but with the additional constraint that $\tilde{\chi}$, as defined in Remark 3.68, must be $\mathcal{A}$-measure consistent.*

The constraint that $\tilde{\chi}$ must be measure consistent can equivalently be recast as a constraint that the vector $\bar{\chi}$ defined in Remark 3.68 must be measure consistent, or as a constraint that the signed measure χ defined in Remark 3.68 must be a measure on $\mathcal{A}$.

Where K and $\bar{K}$ are as in Remark 3.68, it is clear that every vector $y \in K$ is $\mathcal{A}$-measure consistent (it is just a subvector of the zeta vector corresponding to the point $(y_1, \ldots, y_n)$), so since $Cone(K) \subseteq \bar{N}^k(\bar{K})$ it follows that $Cone(K) \subseteq (N^{++})^k(\bar{K})$ as well.

It is evident from Corollary 4.3 and the discussion following that result that the requirement that a matrix $U^{\tilde{\chi}}$ with rows and columns indexed by the sets

$$\bigcap_{i\in J} Y_i,\ J \subset \{1,\dots,n\}, |J| \le k,\ k < n \tag{4.279}$$

must be positive semidefinite is only a relaxation of the requirement that $\tilde{\chi}$ be $\mathcal{A}$-measure consistent. This should make it fairly evident that $(N^{++})^k$ refines $(N^+)^k$ (and this may be guessed from the discussion of N^+ in Chapter 2 as well). Formally, recalling Definition 2.17, we have

$$(N^{++})^k \subseteq (\bar{N}^*)^k \subseteq (N^+)^k. \tag{4.280}$$

The first inclusion of (4.280) can be seen as follows. Let $\mathcal{P} = \mathcal{A}$, and let $\mathcal{Q}'$ be as in Remark 3.68, and recall that the collection $\mathcal{Q}'$ is linearly independent. Let $\tilde{\chi}''$, as defined in Remark 3.68, belong to $(N^{++})^k(\bar{K}) \subseteq \bar{N}^k(\bar{K})$, and let χ, as defined in Remark 3.68, be the measure with which it is consistent. Thus the projection $\tilde{\chi}$ of χ (as in Remark 3.68) is a lifting of $\tilde{\chi}''$ satisfying all of the constraints of $\bar{N}^k(\bar{K})$. Let $U^{\hat{\chi}}$ be the matrix with rows and columns indexed by $\mathcal{Q}'$ with each u, v entry equal to $\chi(u \cap v)$ where $\hat{\chi}$ is the projection of χ on the appropriate space. Then since $\hat{\chi}$ is measure consistent, Lemma 4.1 implies that $U^{\hat{\chi}} \succeq 0$. Thus under the lifting $\hat{\chi}$ of $\tilde{\chi}''$ (which is also a lifting of $\tilde{\chi}$) the positive semidefiniteness constraint of $(\bar{N}^*)^k(\bar{K})$ is satisfied, so since $\tilde{\chi}$ satisfies all of the constraints of $\bar{N}^k(\bar{K})$, we can conclude that $\tilde{\chi}'' \in (\bar{N}^*)^k(\bar{K})$. The latter inclusion of (4.280) is from Theorem 2.18.

The operator N^{++} is actually vastly more powerful than N^+. For example, let $G = (V, E)$ be an undirected graph with vertex set V, $|V| = n$, and edge set E, and let $P \subseteq \{0,1\}^n$ be the collection of the incidence vectors of the stable sets of G. Define

$$K = \{y \in \{0,1\}^{n+1} : y_0 = 1,\ (y_1,\dots,y_n) \in P\} \tag{4.281}$$

and

$$\bar{K} = \{\chi \in R^{n+1} : \chi_i + \chi_j \le \chi_0,\ \forall i,j \in \{1,\dots,n\} : \{i,j\} \in E,\ 0 \le \chi_i \le \chi_0,\ 1 \le i \le n\} \tag{4.282}$$

so that $\bar{K} \cap \{0,1\}^{n+1} = \{0\} \cup K$. Rename the coordinates $0, 1, \dots, n$ as $\{0,1\}^n, Y_1, \dots, Y_n$, let $k = 1$ and let $\mathcal{Q}, \mathcal{Q}'$ and $\mathcal{Q}''$ all be as in Remark 3.68. Thus $N^{++}(\bar{K})$ is the set of points $\tilde{\chi}'' \in R^{\mathcal{Q}''}$ that have a lifting to a measure χ on $\mathcal{A}$ such that $(\tilde{\chi}'')^q \in \bar{K}$ for each q of the form (3.309), where each term $((\tilde{\chi}'')^q)_u$ denotes $\chi(u \cap q)$. So for each $\{i,j\} \in E$,

$$\chi(Y_i) + \chi(Y_i \cap Y_j) = ((\tilde{\chi}'')^{Y_i})_{Y_i} + ((\tilde{\chi}'')^{Y_i})_{Y_j} \le ((\tilde{\chi}'')^{Y_i})_{\{0,1\}^n} = \chi(Y_i) \Rightarrow \tag{4.283}$$

$$\chi(Y_i \cap Y_j) = 0. \tag{4.284}$$

Considering now that χ is a measure and

$$P^c = \bigcup_{\{i,j\}\in E} Y_i \cap Y_j, \tag{4.285}$$

it follows that $\chi(P^c) = 0$. Thus by Lemma 3.27, $\tilde{\chi}''$ is $\mathcal{P}$-measure consistent, and thus belongs to $Cone(K)$ by Corollary 3.19. We conclude that $N^{++}(\bar{K}) = Cone(K)$. Thus the $(N^{++})^k$ operator can characterize the homogenized version of the stable set polytope at level $k = 1$.

Nevertheless despite all of the additional power of the N^{++} operator, there will be cases of cones $\bar{K}$ for which N^{++} offers no improvement over $\bar{N}$, i.e.

$$(N^{++})^k(\bar{K}) = \bar{N}^k(\bar{K}) \neq Cone(\bar{K} \cap \{0,1\}^{n+1}). \tag{4.286}$$

We will see that some of the classes of problems for which it has been noted in the literature that positive semidefiniteness does not help, are actually of this type.

Returning to the diagram, recall that the original values $(\chi(Y_1), \chi(Y_2))$ are always measure consistent. Thus while not every choice of $(\chi(Y_1), \chi(Y_2), \chi(Y_1 \cap Y_2))$ is measure consistent, for each choice of $(\chi(Y_1), \chi(Y_2))$ there is always some choice of $\chi(Y_1 \cap Y_2)$, and consequently of vectors $\bar{\chi}^{Y_1}$, $\bar{\chi}^{Y_1^c}$, $\bar{\chi}^{Y_2}$, and $\bar{\chi}^{Y_2^c}$, that is indeed measure consistent. Thus measure consistency alone never eliminates any points from the hypercube. It is therefore clear already that measure consistency is only useful when coupled with other conditions.

In particular, the $\bar{N}$ conditions place restrictions on the conditional probability vectors $\bar{\chi}^{Y_i}$ and $\bar{\chi}^{Y_i^c}$ (i.e. the convexification vectors, which are the scaled partial sums) and thus on the choices of $\chi(Y_i \cap Y_j)$ that imply those vectors. But if the $\bar{N}$ conditions are such that for every point in the hypercube that they do not eliminate they leave available a choice of conditional probability (scaled partial sum) vectors that is measure consistent, then the measure consistency constraint, and therefore the positive semidefiniteness constraint in N^+ or $\bar{N}^*$, will not cut off any additional fractional points. We will describe two examples of where this happens. But first let us note that such a situation does not imply that the positive semidefinite Lasserre constraints will not help. Indeed, assume that $\bar{\chi} = (\chi(\{0,1\}^n), \chi(Y_1), \ldots, \chi(Y_n)) \in R^{n+1}$, and that the set $P \subseteq \{0,1\}^n$ is the set of integer solutions to a system of linear constraints, whose homogenized form is $k_i^T \bar{\chi} \geq 0,\ i = 1, \ldots, m$. Assume now that $\bar{\chi}$ is not $\mathcal{P}$-measure consistent, but assume that the lifted vector $\tilde{\chi}$ that is used by the Lasserre algorithm to construct its positive semidefinite matrices is nevertheless measure consistent. Then where we denote the restriction of the zeta vectors to their appropriate coordinates as $\tilde{\zeta}^r$,

$$\tilde{\chi} = \sum_{r\in\mathcal{S}} \alpha_r \tilde{\zeta}^r,\ \alpha \geq 0. \tag{4.287}$$

Thus since $\bar{\chi}$ is not $\mathcal{P}$-measure consistent, there must be some $s \in \mathcal{S}$ with $\alpha_s > 0$ such that the point $(\bar{\zeta}^s(Y_1), \ldots, \bar{\zeta}^s(Y_n)) \in \{0,1\}^n$ corresponding to the atom s does not belong to P. But this implies that there must be some k_i such that $k_i^T \bar{\zeta}^s < 0$ and therefore $\alpha_s k_i^T \bar{\zeta}^s < 0$ as well. Thus the vector

$$U^{\tilde{\chi}} k_i = \sum_{r \in \mathcal{S}} (\alpha_r k_i^T \bar{\zeta}^r) \tilde{\tilde{\zeta}}^r \tag{4.288}$$

(where the double tilde indicates some projection) cannot be guaranteed to be measure consistent, and therefore there is no guarantee that the matrix generated by that vector is positive semidefinite.

Thus given a vector that does not belong to the convex hull of P, the fact that we can expand the vector in a measure consistent fashion and such that various linear constraints will hold for the partial sums, is no guarantee that Lasserre's semidefinite constraints will hold. We could, however, guarantee the satisfaction of Lasserre's semidefinite constraints (for a point that does not belong to the convex hull) if we could show that the expanded point is measure consistent for the subset algebras of each of the m sets $P^i = \{y \in \{0,1\}^n : k_i^T y \geq 0\}$, via multiple representations. In particular, if for each $i = 1, \ldots, m$, there is a representation of the expanded vector $\tilde{\chi}$ as

$$\tilde{\chi} = \sum_{r \in \mathcal{S} : k_i^T \bar{\zeta}^r \geq 0} \alpha_r^i \tilde{\zeta}^r, \ \alpha^i \geq 0 \tag{4.289}$$

then the Lasserre constraints will be satisfied. Examples of this sort, however, are harder to construct, and for this reason it tends to be much more difficult to fix lower bounds for Lasserre rank than to do so for N^{++} rank.

4.3.2 Independent Sets

Definition 4.30 *Given a σ-algebra $\mathcal{W}$ of subsets of some universal set Ω, and given a probability measure $\mathcal{X}$ on $(\Omega, \mathcal{W})$, two sets $A, B \in \mathcal{W}$ are said to be independent with respect to the probability measure $\mathcal{X}$ if*

$$\mathcal{X}(A \cap B) = \mathcal{X}(A)\mathcal{X}(B). \tag{4.290}$$

See Chapter 10 of [F99] for details.

Recall that the set I (defined in Lemma 3.49) of all intersections of sets Y_i is a linearly independent spanning collection for $\mathcal{A}$. Thus every vector in R^I is $\mathcal{A}$-signed-measure consistent with a unique signed measure on $\mathcal{A}$.

Lemma 4.31 *Given a vector* $\bar{\chi} = (1, \chi(Y_1), \ldots, \chi(Y_n)) \in [0,1]^{n+1}$*, the (unique) signed measure* χ *on* $\mathcal{A}$ *such that for all collections* $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, n\}$*,*

$$\chi(Y_{i_1} \cap \cdots \cap Y_{i_k}) = \prod_{j=1}^{k} \chi(Y_{i_j}) \tag{4.291}$$

is a probability measure on $\mathcal{A}$*.*

This is just the probability measure χ for which the sets Y_i are (probability theoretically) independent with respect to χ. Recall that the number $\bar{\chi}^{Y_1}(Y_2)$ is the conditional probability of Y_2 given Y_1. Thus where the sets Y_1 and Y_2 are independent with respect to χ, we have $\bar{\chi}^{Y_1}(Y_2) = \chi(Y_2)$.

Proof:

One way to prove the lemma is to use the fact proven in the previous chapter (Lemma 3.29) that the numbers

$$\{\chi(Y_{i_1} \cap \cdots \cap Y_{i_k})\} \tag{4.292}$$

are probability measure consistent iff there exist sets $T_1, \ldots, T_n$ in some probability measure space $(\Omega, \mathcal{W}, \mathcal{X})$ such that the $\mathcal{X}$ measure of each $T(i_1) \cap \cdots \cap T(i_k)$ is $\chi(Y_{i_1} \cap \cdots \cap Y_{i_k})$. Thus all we need to find is some probability measure space $(\Omega, \mathcal{W}, \mathcal{X})$ in which there are some n independent events (i.e. n sets in $\mathcal{W}$ that are independent with respect to $\mathcal{X}$), with each i'th event being of probability $\chi(Y_i)$. From the standpoint of probability theory it is trivial that such spaces exist. Just consider, for example, n independent "Bernoulli" experiments (i.e. each experiment has two possible outcomes: "success" or "failure"), with each i'th experiment succeeding with probability $\chi(Y_i)$. The lemma can also be proven formally as follows. Let $\tilde{\chi}$ denote the projection of χ on R^I, and let $U^{\tilde{\chi}}$ be the matrix with rows and columns indexed by I with each u, v entry equal to $\tilde{\chi}(u \cap v)$ (note that $u \cap v \in I$ as well). Then considering that each $u, v \in I$ is an intersection of sets Y_i, then by definition of χ, we have

$$U^{\tilde{\chi}}_{u,v} = \chi(u \cap v) = \chi(u)\chi(v). \tag{4.293}$$

This implies that $U^{\tilde{\chi}} = \tilde{\chi}\tilde{\chi}^T \succeq 0$. Since I is a linearly independent spanning collection, $\tilde{\chi}$ is $\mathcal{A}$-signed-measure consistent and the lemma follows from Corollary 4.3 (letting $\mathcal{Q} = \mathcal{Q}' = I$). $\square$

Geometrically the case of Lemma 4.31 means that $\bar{\chi}^{Y_1}(Y_2) = \chi(Y_2)$, and $\bar{\chi}^{Y_2}(Y_1) = \chi(Y_1)$. In terms of the picture above in Figure 1 this would yield

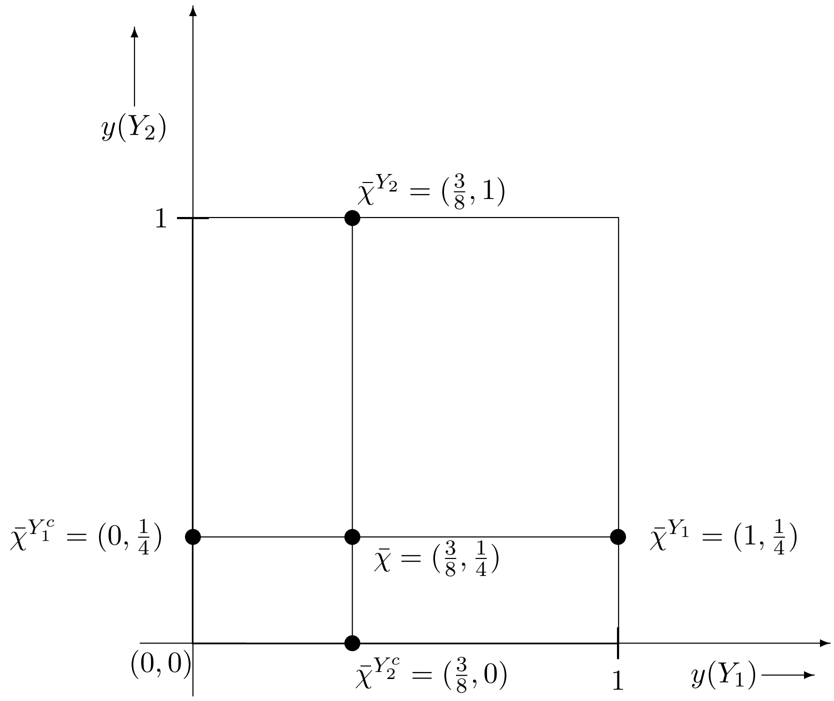

Figure 2

Let $S \subseteq [0,1]^n$, and let $\bar{K}(S)$ be the homogenized version of S in R^{n+1}, as in Definition 1.2. We claim that if $\bar{\chi} \in [0,1]^n$ is such that $\bar{\chi} - \bar{\chi}_i e_i$ and $\bar{\chi} + (1 - \bar{\chi}_i)e_i$ all belong to S, for all $i = 1, \ldots, n$, then the point $\bar{\chi}$ cannot be eliminated by $N^{++}(\bar{K}(S))$. This should be evident from the diagram, as the decomposition into partial sums depicted in the diagram is measure consistent and satisfies the N operator requirements by hypothesis, but to see this formally, note first that the lifting of $\bar{\chi}$ obtained by adding coordinates for each $Y_i \cap Y_j$ of value $\bar{\chi}(Y_i)\bar{\chi}(Y_j)$ satisfies the N operator constraints since the partial sum $\hat{\chi}^{Y_i}$ (where the hat indicates that there is a coordinate for the universal set - to be denoted by the subscript zero - as well) satisfies

$$\hat{\chi}_0^{Y_i} = \hat{\chi}^{Y_i}(Y_i) = \bar{\chi}(Y_i), \tag{4.294}$$

and for all $j \neq i$,

$$\hat{\chi}^{Y_i}(Y_j) = \bar{\chi}(Y_i)\bar{\chi}(Y_j). \tag{4.295}$$

Thus $\hat{\chi}^{Y_i} = \bar{\chi}(Y_i)$ times the vector in R^{n+1} with a 1 in its zero'th and in its i'th coodinates, and with $\bar{\chi}(Y_j)$ in each of its remaining j'th coordinates. By hypothesis this vector belongs to $\bar{K}(S)$. A similar situation holds for the partial sums $\hat{x}^{Y_i^c}$. The lifting moreover is measure consistent with the measure defined in Lemma 4.31. This gives us a stronger version of Goemans and Tuncel's Theorem 4.1 and Corollary 4.2 ([GT01]):

Definition 4.32 *Given $\bar{\chi} \in [0,1]^n$, define the vector $\bar{\chi}^{(j)}$ to be the same as $\bar{\chi}$ but with a 0 in the j'th position.*

For the purposes of the next theorem and corollary, let $S \subseteq [0,1]^n$ be convex, and let $N(S)$ denote the projection of $N(\bar{K}(S)) \cap \{\hat{\chi} \in R^{n+1} : \hat{\chi}_0 = 1\}$ on R^n, and similarly for N^0, N^+ and N^{++}.

Theorem 4.33 *Let $\bar{\chi} \in S$ satisfy*

$$\bar{\chi}^{(j)} \text{ and } (\bar{\chi}^{(j)} + e_j) \in S \ \forall j : 0 < \bar{\chi}_j < 1. \tag{4.296}$$

Then $\bar{\chi} \in N^{++}(S)$. □

Corollary 4.34 *Let S be such that $(S \cap \{\bar{\chi} : \bar{\chi}_j = 0\}) + e_j = S \cap \{\bar{\chi} : \bar{\chi}_j = 1\}$ for all $j \in \{1, \ldots, n\}$ (see their diagram) then*

$$N^{++}(S) = N^+(S) = N(S) = N^0(S) = \bigcap_{j \in \{1,\ldots,n\}} \{\bar{\chi} : \bar{\chi}^{(j)} \in S\}. \ \square \tag{4.297}$$

4.3.3 Mutually Exclusive Sets

The other case where measure consistency does not help that we will discuss is in some ways the opposite of the first case. This case is more trivial, but it has some interesting behavior. Consider the vector

$$(\bar{\chi}_0, \bar{\chi}) \in R_+^{n+1}, \ \bar{\chi} \in [0, \bar{\chi}_0]^n \tag{4.298}$$

with the first coordinate corresponding to the universal set, and the subsequent n coordinates to $Y_1, \ldots, Y_n$ respectively. Write

$$\bar{\chi} = \sum_{i=1}^{n} \bar{\chi}(Y_i) e_i. \tag{4.299}$$

Where we define

$$N_j = \{y \in \{0,1\}^n : y_j = 0\} = Y_j^c, \tag{4.300}$$

the point $e_i \in R^n$ comprises the atom

$$r_i = Y_i \cap \bigcap_{j=1,\ldots,n, j\neq i} N_j \tag{4.301}$$

and it is the projection of ζ^{r_i} on its $Y_1, \ldots, Y_n$ coordinates. (Recall that ζ^{r_i} is the measure that assigns a value of 1 to every set that contains the atom r_i and zero to every other set. These are the "atomic measures".) Thus $\bar{\chi}$ is always consistent with the measure

$$\sum_{i=1}^{n} \bar{\chi}(Y_i)\zeta^{r_i}. \tag{4.302}$$

The measure defined by (4.302) assigns a measure of $\bar{\chi}(Y_i) \geq 0$ to each atom r_i, and zero measure to every other atom. But (4.302) may not be consistent with $(\bar{\chi}_0, \bar{\chi})$. To be consistent with $(\bar{\chi}_0, \bar{\chi})$, we have to also ensure that a measure of $\bar{\chi}_0$ is assigned to the universal set $\{0,1\}^n$. So consider the signed measure χ that, like (4.302), assigns $\bar{\chi}(Y_i)$ to each atom r^i, but which also assigns $\bar{\chi}_0 - \sum_{i=1}^{n} \bar{\chi}(Y_i)$ to the one atom that belongs to none of the sets Y_i, namely the atom

$$r_0 = \bigcap_{j=1,\ldots,n} N_j. \tag{4.303}$$

Since r_0 belongs to none of the Y_i, the signed measure of each set Y_i remains unchanged from what it was for (4.302), and therefore consistency with $\bar{\chi}$ continues to be maintained. The vector $(\bar{\chi}_0, \bar{\chi})$ is therefore consistent with the signed measure

$$(\bar{\chi}_0 - \sum_{i=1}^{n} \bar{\chi}(Y_i))\zeta^{r_0} + \sum_{i=1}^{n} \bar{\chi}(Y_i)\zeta^{r_i} \tag{4.304}$$

which is a measure iff

$$\bar{\chi}_0 - \sum_{i=1}^{n} \bar{\chi}(Y_i) \geq 0. \tag{4.305}$$

Assume that (4.305) holds, so that χ defined by (4.304) is in fact a measure. Observe that each set Y_i contains only one of these atoms (namely r_i) so the partial sums χ^{Y_i} are just the atomic measures ζ^{r_i} scaled by $\bar{\chi}(Y_i)$. The normalized partial sums (the conditional probability vectors) projected on their $Y_1, \ldots, Y_n$ coordinates, namely the vectors we have denoted $\bar{\chi}^{Y_i}$ in the diagrams, are just the vertices e_i, and the intersections of distinct Y_i are all of measure zero. (In probability terms, the sets Y_i are mutually exclusive, and thus the conditional probability of $Y_i|Y_j$, where Y_j is of positive probability, is one if $i = j$ and zero otherwise. The conditional probability given Y_i of every atom r_j comprised by the point $y^j \in \{0,1\}^n$, is thus zero unless it is contained in Y_i and in no other set Y_l, i.e. unless $y^j = e_i$.) Thus (4.304) assigns a measure of $\bar{\chi}(Y_i)$ to each set Y_i, and a measure of $\sum_{i=1}^n \bar{\chi}(Y_i)$ to their union, which is the maximum possible measure in general for unions. Equivalently, for any intersection

$$N_{i_1} \cap \cdots \cap N_{i_k} = (Y_{i_1} \cup \cdots \cup Y_{i_k})^c \tag{4.306}$$

we have

$$\chi(N_{i_1} \cap \cdots \cap N_{i_k}) = \chi_0 - \sum_{j=1}^k \chi(Y_{i_j}) \tag{4.307}$$

which is the minimum possible measure for intersections. Note also for every intersection q of sets N_i, the measure of the intersection of any Y_j (where Y_j is not one of the elements that intersected to give q) with q is just Y_j again. Thus this is the measure that gives the highest possible values for the measures of sets of the form $q \cap Y_j$.

In terms of our diagram,

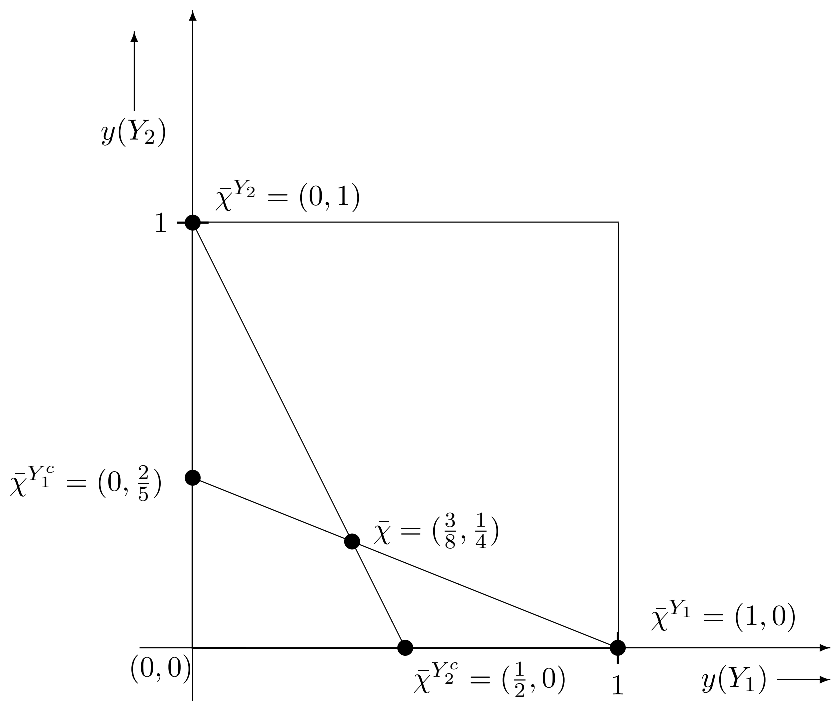

Figure 3

These facts are illustrated in the Cook and Dash example ([CD01])

$$S = \{\bar{\chi} \in [0,1]^n : \sum_{i=1}^{n} : \bar{\chi}_i \geq \frac{1}{2}\} \tag{4.308}$$

with homogenized form

$$\bar{K} = \{(\bar{\chi}_0, \bar{\chi}) \in R_+^{n+1} : \bar{\chi} \in [0, \bar{\chi}_0]^n, \sum_{i=1}^{n} \bar{\chi}_i \geq \frac{1}{2}\bar{\chi}_0\}. \tag{4.309}$$

Let $P = S \cap \{0,1\}^n$; let $K = \{y \in \{0,1\}^{n+1} : (y_1, \ldots, y_n) \in P\}$ (as per Definition 1.2), and note that

$$Cone(K) = \{(\bar{\chi}_0, \bar{\chi}) \in \bar{K} : \sum_{i=1}^{n} \chi_i \geq \bar{\chi}_0\}. \tag{4.310}$$

So the only candidates from $\bar{K}$ for being eliminated by the N operators are points

$$\{(\bar{\chi}_0, \bar{\chi}) \in \bar{K} : \sum_{i=1}^{n} \chi_i < \bar{\chi}_0\}. \tag{4.311}$$

But every such point is also a candidate for being represented as a measure (4.304) as described above. We will show that in fact none of the $\bar{N}^l$ constraints, for any l, eliminate this representation for any point that the $\bar{N}^l$ constraints do not eliminate altogether. Thus every point $\bar{\chi}$ that is not eliminated by $\bar{N}^l$ is already measure consistent and demanding measure consistency therefore adds nothing to $\bar{N}^l$.

By Remark 3.68, a point $(\bar{\chi}_0, \bar{\chi}) \in \bar{N}^l(\bar{K})$, iff it can be lifted to a signed measure, to be denoted $\tilde{\chi}$, on $\mathcal{A}$, such that for each set

$$q \in \mathcal{Q} := \left\{ \bigcap_{i \in V} Y_i \cap \bigcap_{i \in W} N_i : V, W \subseteq \{1, \ldots, n\}, |V| + |W| \leq l \right\}, \tag{4.312}$$

the following two constraints are satisfied:

$$\sum_{i=1}^{n} \tilde{\chi}(q \cap Y_i) \geq \frac{1}{2}\tilde{\chi}(q) \tag{4.313}$$

$$0 \leq \tilde{\chi}(q \cap Y_i) \leq \tilde{\chi}(q). \tag{4.314}$$

(These are the original constraints that defined $\bar{K}$, applied to the projection of the partial sum $\tilde{\chi}^q$ on the coordinates corresponding to $\{\{0,1\}^n, Y_1, \ldots, Y_n\}$, cf. Corollary 3.67.) So suppose that indeed $(\bar{\chi}_0, \bar{\chi}) \in \bar{N}^l(\bar{K})$, and that its lifting $\tilde{\chi}$ is a signed measure satisfying (4.313) and (4.314). Since $\tilde{\chi}$ is a signed measure we must also have for all $q \in \mathcal{Q}$,

$$\tilde{\chi}(q \cap Y_i) + \tilde{\chi}(q \cap N_i) = \tilde{\chi}(q). \tag{4.315}$$

Putting together (4.314) and (4.315) we obtain

$$0 \leq \tilde{\chi}(q \cap N_i) \leq \tilde{\chi}(q) \tag{4.316}$$

for each $q \in \mathcal{Q}$, and repeated application of (4.314) and (4.316) implies that for all $q \in \mathcal{Q}$,

$$\bar{\chi}(Y_i) \geq \tilde{\chi}(q \cap Y_i). \tag{4.317}$$

By (4.315) and (4.317) we now obtain, for each $k \leq l$,

$$\tilde{\chi}(N_{i_1} \cdots N_{i_k}) + \tilde{\chi}(Y_{i_k}) \geq \tilde{\chi}(N_{i_1} \cdots N_{i_{k-1}} N_{i_k}) + \tilde{\chi}(N_{i_1} \cdots N_{i_{k-1}} Y_{i_k}) = \tilde{\chi}(N_{i_1} \cdots N_{i_{k-1}}) \tag{4.318}$$

(where we have suppressed the intersection symbols) and by repeated application of (4.318)

$$\tilde{\chi}(N_{i_1} \cdots N_{i_k}) + \sum_{j=1}^{k} \bar{\chi}(Y_{i_j}) \geq \bar{\chi}_0. \tag{4.319}$$

Now consider what would happen had we expanded $\bar{\chi}$ into the measure (to be denoted χ) of the form (4.304). Clearly every measure is a signed measure, and every measure satisfies the constraints of the form (4.314). We will now show that χ also satisfies constraints (4.313), so that χ is a valid lifting for the purposes of $\bar{N}^l$, which establishes that $(\bar{\chi}_0, \bar{\chi}) \in (N^{++})^l(\bar{K})$ as well, since χ is a measure.

If $q \in \mathcal{Q}$ is the empty intersection, i.e. $q = \{0,1\}^n$, then

$$\sum_{i=1}^{n} \chi(q \cap Y_i) = \sum_{i=1}^{n} \chi(Y_i) = \sum_{i=1}^{n} \bar{\chi}(Y_i) \geq \frac{1}{2}\bar{\chi}_0 = \frac{1}{2}\chi(q) \tag{4.320}$$

where the second and third equalities follow from the definition of liftings, and the inequality follows from the fact that $(\bar{\chi}_0, \bar{\chi}) \in \bar{N}^l(\bar{K}) \subseteq \bar{K}$. Thus (4.313) is satisfied in this case. Consider now intersections q that entail one or more sets of the form Y_j. In this case we have

$$\sum_{i=1}^{n} \chi(q \cap Y_i) = \chi(q) \geq \frac{1}{2}\chi(q) \tag{4.321}$$

since the χ measure of any intersection of more than one set Y_j is zero, so (4.313) is still satisfied. Finally, if $q = N_{i_1} \cdots N_{i_k}$, then we already noted that each

$$\chi(q \cap Y_i) = \chi(Y_i) \tag{4.322}$$

(wherever $i \neq i_j,\ j = 1, \ldots, k$), and that

$$\chi(q) = \bar{\chi}_0 - \sum_{j=1}^{k} \bar{\chi}(Y_{i_j}). \tag{4.323}$$

By (4.322) and (4.317) we have

$$\sum_{i=1}^{n} \chi(q \cap Y_i) = \sum_{i=1,\ldots,n, i \neq i_1,\ldots,i_k} \bar{\chi}(Y_i) \geq \sum_{i=1}^{n} \tilde{\chi}(q \cap Y_i) \geq \frac{1}{2}\tilde{\chi}(q) \tag{4.324}$$

since the fact that $\tilde{\chi}$ is a signed measure implies that for all $j \in \{i_1, \ldots, i_k\}$, $\tilde{\chi}(q \cap Y_j) = \tilde{\chi}(\emptyset) = 0$, and the final inequality in the expression holds by hypothesis. Moreover by (4.319) and (4.323),

$$\tilde{\chi}(q) = \tilde{\chi}(N_{i_1} \cdots N_{i_k}) \geq \bar{\chi}_0 - \sum_{j=1}^{k} \bar{\chi}(Y_{i_j}) = \chi(q), \tag{4.325}$$

which together with (4.324) implies that

$$\sum_{i=1}^{n} \chi(q \cap Y_i) \geq \frac{1}{2}\chi(q), \tag{4.326}$$

and thus χ satisfies all constraints (4.313). We conclude that if any lifted vector $\tilde{\chi}$ satisfies the $\bar{N}$ constraints, then χ certainly does also. Thus in enforcing $\bar{N}$ conditions, for each $(\bar{\chi}_0, \bar{\chi})$, among the choices of expanded vectors that satisfy those conditions (if there are any) there is always a choice that corresponds to a measure (namely the measure (4.304)), and thus requiring measure consistency never eliminates any additional points at any level of $\bar{N}$. This thus strengthens the result of Cook and Dash.[4]

Geometrically, the polytope S is as follows.

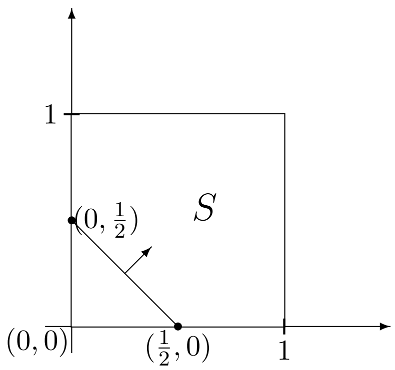

Figure 4

Though a two dimensional drawing is not really adequate, note how by choosing (in Figure 3) the vectors $\bar{\chi}^{Y_i}$ to be the vertices e_i of the square, the values of $\bar{\chi}^{Y_1^c}(Y_2)$ and $\bar{\chi}^{Y_2^c}(Y_1)$ are maximized, thus casting the vectors $\bar{\chi}^{Y_1^c}$ and $\bar{\chi}^{Y_2^c}$ as close as possible to the polytope S depicted in Figure 4. Contrast this to Figures 1 and 2, where $\bar{\chi}^{Y_i}$ were not

[4] By "strengthen" we mean that it shows that not only will positive semidefiniteness not help, as was shown by Cook and Dash, but measure consistency will not help either. It should be noted, however, that Cook and Dash addressed themselves to a slightly different problem. They showed that N^+ does not strengthen N or even N^0 (defined in the Definition 1.9) at any iteration. We have shown here that $(N^{++})^l$ does not strengthen $\bar{N}^l$ for any l.

chosen at the vertices e_i, and where the points $\bar{\chi}^{Y_i^c}$ are further from the polytope S. This illustrates that the choice of normalized partial sums $\bar{\chi}^{Y_i}$ at the vertices e_i is the optimal choice in the effort to ensure that the vectors $\bar{\chi}^{Y_i}$ and $\bar{\chi}^{Y_i^c}$ in fact belong to S.

One conclusion that we should reasonably draw from the results of this section is that in order to maximize the effectiveness of positive semidefiniteness we ought to try to enforce test vector conditions, or at least constraints to ensure $\mathcal{P}$-signed-measure consistency. The N^+ operator on stable set is an example where $\mathcal{P}$-signed-measure consistency and test vector consistency hold, and in that case N^+ is indeed much more powerful than N.

Chapter 5

Algorithms Driven by Set Theoretic Structure

5.1 Introduction

The previous chapters showed how lifting a set $P \subseteq \{0,1\}^n$ to the space with dimension indexed by P's subset algebra can capture the structure of P. In this chapter and the next we will turn our attention to the task of algorithmically exploiting this structure. The algorithms discussed in the first two chapters can all be understood to exploit this structure in one way or another, but there are several aspects of the structure exposed by the lifting that are not addressed by any of those algorithms. We have shown that all of those algorithms make either explicit or implicit use of what we called *partial summation* to successively approximate $Conv(P)$. They accomplish this by way of a gradual construction of a complete spanning set for the subset algebra $\mathcal{A}$ of $\{0,1\}^n$, which allows one to calculate every possible partial sum. Implicit or explicit constraints on the partial sums are then used to ensure $\mathcal{P}$-measure consistency (see Remark 3.68). Several points may be noted in this regard. One is that partial summation is an example of a measure-preserving operator. Lasserre's algorithm actually takes advantage of a more general measure-preserving operator, and in principle there may be other ways of utilizing measure-preserving operators to one's advantage as well. This is an area for further research, but it is one that we will not pursue here.

Secondly, none of the algorithms take specific notice of the measure-theoretic interpretation of the lifted vectors. Measure-consistency can be used as a source for generating almost limitless numbers of valid inequalities. While we do not actually want to enforce a limitless number of constraints, and we have seen that these constraints can be loosely approximated by positive semidefiniteness, these are nonetheless a largely untapped source of relationships that may be exploited among lifted variables.

Thirdly, these algorithms terminate only upon the construction of a complete spanning set for $\mathcal{A}$. This is effectively complete enumeration (we have hinted at this already at the beginning of Chapter 1), and these algorithms can in fact be viewed as merely a methodical process of complete enumeration. While arguably this ought to be expected of any algorithm that is meant to handle arbitrary integer programs, nevertheless the construction of a full spanning set is in some ways more than complete enumeration, as it completely determines the entire algebra $\mathcal{A}$, which may be far more information than we need.

But more importantly, it may be hoped that the *process and order* of the enumeration can be made to intelligently reflect the structure of the particular problem. All of the algorithms that have been considered so far, however, use effectively the same gradual construction of the spanning set of $\mathcal{A}$ regardless of P.

The algorithms that will be presented in this and the next chapter will also use the partial summation paradigm as a guide to the introduction of new variables. Partial summation is a sensible guide in that it introduces new variables with clear and known relationships amongst each other and the original variables. One particularly handy feature of partial sums is that if u and v are disjoint members of $\mathcal{P}$, and χ is a (signed) measure on $\mathcal{P}$, then the partial sums χ^u and χ^v satisfy

$$\chi^u + \chi^v = \chi^{u \cup v} \tag{5.1}$$

(since for each $q \in \mathcal{P}$, $\chi^u[q] + \chi^v[q] = \chi[u \cap q] + \chi[v \cap q] = \chi[(u \cup v) \cap q] = \chi^{u \cup v}[q]$). The algorithms of the first two chapters all made either explicit or implicit use of the following fact. Each pair of sets Y_i, N_i (with $Y_i = \{y \in \{0,1\}^n : y_i = 1\}$, $N_i = Y_i^c$) partitions $\{0,1\}^n$, and thus any (signed) measure χ on $\mathcal{A}$ can be decomposed as $\chi = \chi^{Y_i} + \chi^{N_i}$. This fact is useful because the (signed) measures χ^{Y_i} and χ^{N_i} are more highly structured than the (signed) measure χ. In particular, $\chi^{Y_i}[Y_i \cap q] = \chi^{Y_i}[q]$ for all $q \in \mathcal{A}$. Each χ^{Y_i} can similarly be decomposed as $\chi^{Y_i} = \chi^{Y_i \cap Y_j} + \chi^{Y_i \cap N_j}$, and so on. This progressive partitioning of $\{0,1\}^n$ and decomposition of χ is the principle that guides the selection of new variables in all of the algorithms of the first two chapters (regardless of P).

In this chapter and the next we will be considering partitioning schemes that focus on the partitioning of P rather than $\{0,1\}^n$, and which use the set theoretic structure of P itself as their guide. Thus if, for example, $P = (Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)$, then we might decompose a candidate measure χ on $\mathcal{P}$ (or a projection thereof) as

$$\chi = \chi^{(Y_1 \cup Y_2) \cap Y_3} + \chi^{(Y_1 \cup Y_2) \cap N_3 \cap Y_4}. \tag{5.2}$$

We will begin to see the details in the next section.

We will show that using such an approach, for certain classes of feasible regions P, most of the algorithms that we will present will produce sets that telescope to approximate $Conv(P)$ increasingly well in a quite concrete manner. Specifically, let us suggest the following definition.

Definition 5.1 *Given an inequality*

$$\alpha^T x \geq \beta,\ \alpha \geq 0,\ 0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{|support(\alpha)|} \tag{5.3}$$

We will say that the pitch *of the inequality, to be denoted* $\pi(\alpha, \beta)$ *is*

$$\pi(\alpha, \beta) = \min \left\{ k : \sum_{j=1}^{k} \alpha_j \geq \beta \right\}. \tag{5.4}$$

The pitch of an inequality may be thought of as a measure of how positive a $0, 1$ vector needs to be in order for the inequality to be satisfied. (To be completely precise, it is a measure of how positive those coordinates of the vector that are in the suppport of the inequality need to be in order for the inequality to be satisfied.)

Note that for any $P \subseteq \{0, 1\}^n$, every valid inequality $\alpha^T x \geq \beta,\ \alpha \geq 0$ has pitch $\leq n$. The notion of pitch can also be used to characterize inequalities $a^T x \geq \beta$ where $a \not\geq 0$ since we can always define

$$\alpha_{i'} = \begin{cases} a_i : a_i \geq 0 \\ 0 : \text{ otherwise} \end{cases} \tag{5.5}$$

$$\alpha_{i''} = \begin{cases} -a_i : a_i \leq 0 \\ 0 : \text{ otherwise} \end{cases} \tag{5.6}$$

$$x'_{i'} = x_i,\ \text{and}\ x'_{i''} = 1 - x_i, \quad i = 1, \ldots, n. \tag{5.7}$$

Thus $x', \alpha \in R^{2n}$ and $\alpha \geq 0$, and

$$\alpha^T x' \geq \beta + \sum_{i=1}^{n} \alpha_{i''} \ \text{ iff } \ a^T x \geq \beta, \tag{5.8}$$

and $\pi(\alpha, \beta + \sum_{i=1}^{n} \alpha_{i''}) \leq 2n$.

We will show that for certain classes of P, all constraints of pitch $\leq k$ that are valid for P (or for more general cases, the constraints that are valid for a particular relaxation of P) are valid for the approximation of $Conv(P)$ generated at the k'th "level" of most of the algorithms that we will present in this and the next chapter. The algorithms to be described can also terminate without having generated a spanning set. We will also make some use of measure theoretic inequalities, and we will see that these inequalities together with positive semidefiniteness and the P-driven choice of sets can generate interesting constraints that would be difficult to obtain in the absence of the positive semidefiniteness condition. (We have noted already in Section 4.2 that positive semidefiniteness in the absence of attention paid to the structure of P can be quite useless.)

In order to do any of this however, we will need to assume that P has a set theoretic structure that can be "nicely expressed" in some way. Where $A_i \subseteq \{1,\dots,n\},\ i = 1,\dots,m$, and P is the set

$$P = \{y \in \{0,1\}^n : \sum_{j \in A_i} y_j \geq 1,\ i = 1,\dots,m\} \tag{5.9}$$

(the points of P are the incidence vectors of the "set coverings" of the A_i) then we can write

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} Y_j \tag{5.10}$$

where, as usual, $Y_j = \{y \in \{0,1\}^n : y_j = 1\}$. This is a simple set theoretic structure, and we will see that it is easy to exploit. On the other hand,

$$P = \{y \in \{0,1\}^n : By \geq b\} \tag{5.11}$$

where B is an arbitrary $m \times n$ matrix and b is an arbitrary vector, does not necessarily have such a "nice" set theoretic description. We will say that a set-theoretic description for P is "nice" if it entails only sets Y_j, and arbitrary unions, intersections and complementations, (this is the defining characteristic of membership in the algebra generated by $\{Y_1,\dots,Y_n\}$), and is of manageable length. Equivalently, the sets P we will be interested in are those that can be described concisely by arbitrary logical constraints on the boolean variables $y_1,\dots,y_n$, entailing terms of the form "$y_i = 1$", "AND", "OR", and "NOT". Specifically, the sets P that we will be working with are those that have the form

$$P = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(\cdot)} \bigcap_{i_2=1}^{m_2(\cdot)} \bigcup_{j_2=1}^{t_2(\cdot)} \cdots \bigcap_{i_h=1}^{m_h(\cdot)} \bigcup_{j_h=1}^{t_h(\cdot)} M_{f(i_1,j_1,\cdots,i_h,j_h)} \tag{5.12}$$

where $M_{f(\cdot)}$ is a set either of the form Y_l or Y_l^c for some $l \in \{1,\dots,n\}$, each t_l is a function of $i_r,\ r \leq l$ and $j_r,\ r < l$, and each m_l is a function of i_r and $j_r,\ r < l$.

Typically we will say that f maps into the set $\{1', 1'', 2', 2'', \dots, n', n''\}$, and that for each $l \in \{1,\dots,n\}$, $M_{l'} = Y_l$, and $M_{l''} = Y_l^c = N_l$. For example, if $m_1 = 3$ and

$$t_1(1) = 2,\ t_1(2) = 3,\ t_1(3) = 2 \tag{5.13}$$

and

$$f(1,1) = 1',\ f(1,2) = 3'' \tag{5.14}$$

$$f(2,1) = 2'',\ f(2,2) = 1'',\ f(2,3) = 3' \tag{5.15}$$

$$f(3,1) = 2',\ f(3,2) = 1'' \tag{5.16}$$

then

$$\bigcap_{i_1=1}^{3} \bigcup_{j_1=1}^{t_1(i_1)} M_{f(i_1,j_1)} = \tag{5.17}$$

$$(Y_1 \cup N_3) \cap (N_2 \cup N_1 \cup Y_3) \cap (Y_2 \cup N_1). \tag{5.18}$$

It should be noted that any set theoretic expression composed of unions and/or intersections and/or complementations of some or all of the sets $Y_1, \ldots, Y_n$ can be put into the form (5.12) in time polynomial in the length of that expression. For example consider the expression,

$$\Theta = [(Y_1 \cup (Y_2 \cap Y_3)^c)^c \ \cup \ (Y_2^c \cap (Y_3 \cup Y_4)^c)^c]^c \tag{5.19}$$

The outermost complementation can be removed via the rule

$$(A \cup B)^c = A^c \cap B^c \tag{5.20}$$

(which can increase the length of the expression by no more than what it takes to represent one additional complementation) yielding

$$\Theta = (Y_1 \cup (Y_2 \cap Y_3)^c) \ \cap \ (Y_2^c \cap (Y_3 \cup Y_4)^c)\,. \tag{5.21}$$

By (5.20) and the rule

$$(A \cap B)^c = A^c \cup B^c \tag{5.22}$$

we can similarly conclude that

$$\Theta = (Y_1 \cup Y_2^c \cup Y_3^c) \ \cap \ (Y_2^c \cap Y_3^c \cap Y_4^c) = (Y_1 \cup N_2 \cup N_3) \cap N_2 \cap N_3 \cap N_4 \tag{5.23}$$

which is of the form

$$\Theta = \bigcap_{i_1=1}^{4} \bigcup_{j_1=1}^{t_1(i_1)} M_{f(i_1,j_1)} \tag{5.24}$$

where $t_1(1) = 3$ and $t_1(2) = t_1(3) = t_1(4) = 1$, and where

$$f(1,1) = 1',\ f(1,2) = 2'',\ f(1,3) = 3'',\ \text{and} \tag{5.25}$$

$$f(2,1) = 2'',\ f(3,1) = 3'',\ f(4,1) = 4''. \tag{5.26}$$

In general, for arbitrary set theoretic expressions, it is not hard to see that progressively removing the "outermost" complementations via (5.20) and (5.22) will always yield an expression of the form (5.12) in time polynomial in the length of the original expression, and of length no more than a constant multiple of the length of the original expression.

Sets P whose description is given in terms of linear constraints do not necessarily have nice set theoretic descriptions, i.e. their set theoretic descriptions may all be of a length that is exponentially larger than the length of their standard ILP descriptions. But this works both ways: a set may have a nice set theoretic description without having any concise (relative to the size of its set theoretic description) representation via linear inequalities, and this is in fact usually the case. For example, consider

$$P = \bigcap_{i=1}^{m} \bigcup_{j=1}^{m_i} \bigcap_{k=1}^{m_{i,j}} Y_{f(i,j,k)} \tag{5.27}$$

where each $f(i,j,k) \in \{1,\ldots,n\}$. This is the polynomial integer program

$$P = \{y \in \{0,1\}^n : \sum_{j=1}^{m_i} \prod_{k=1}^{m_{i,j}} y_{f(i,j,k)} \geq 1,\ i = 1,\ldots,m\}. \tag{5.28}$$

To express this as a linear integer program would require exponentially many linear constraints. Notice, however, that this is still a linear integer program of length polynomial in the size of the set theoretic representation of P, in the variables of the form

$$\prod_{j \in J \subseteq \{1,\ldots,n\}} y_j. \tag{5.29}$$

Recall from Subsection 2.1.4 that the $\bar{N}$ formalism could therefore still be applied to this problem. But consider now

$$P = \bigcap_{i=1}^{m} \bigcup_{j=1}^{m_i} \bigcap_{k=1}^{m_{i,j}} \bigcup_{l=1}^{m_{i,j,k}} Y_{f(i,j,k,l)}. \tag{5.30}$$

This will now require exponentially long constraints even if it is to be represented as a linear integer program in the variables $\prod y_j$, and for appropriate choices of f and $m_{(\cdot)}$, it will require exponentially long constraints even if it is to be represented as a linear integer program in the variables $\prod_{j \in J} y_j \prod_{j \in H}(1 - y_j)$. Nevertheless, this set still has a "nice" set theoretic description, and it can be handled by the algorithm that will be presented shortly. Thus while the approach that we will take here cannot be applied to every problem for which the previous algorithms may be applied, those approaches also cannot be applied to every problem for which this approach may be applied.

Let us also mention that in principle there may be other ways to "nicely describe" sets, and we will see hints of such features in the next chapter.

5.2 Feasible Space Partitioning Algorithms

5.2.1 Introduction

Notation: Throughout this chapter and the next, we will typically refer to general vectors, whether lifted or not, as x. The q coordinate of the vector (where q is a set) will generally be denoted $x[q]$, and the expression $x(\cdot)$ will usually have a different meaning. Partial sum vectors will be denoted x^q where the partial sum is taken over q. □

The algorithms that will be presented over the course of this chapter and the next are based primarily on a partial summation paradigm. We have noted already that partial summation can be thought of as an extended version of disjunctive programming. Indeed the first (and most basic) version of the algorithm that will be presented first can be interpreted almost solely in terms of disjunctive programming. By disjunctive programming we are referring here not to convexification but to the abstract formulation in which the feasible set is construed as a union of subsets, and potential points of the convex hull are decomposed into points that are meant to belong to the convex hull of these subsets. This is related to the idea that a measure can be decomposed into a sum of partial sum measures corresponding to a union of subsets of P that partitions P. The extra machinery of the broader algebraic interpretation, however, is required for defining the refinements of the algorithm described at the end of this chapter (and in the next), as well as for algorithms such as the breadth first partitioning algorithm at the end of this chapter, and for the semidefiniteness results of the next. It is also used in some of the details of the basic algorithm.

The following fundamental fact will be exploited repeatedly in what follows.

Lemma 5.2 *Given any collection of sets $Q_1, \dots, Q_t$, the union $\bigcup_{i=1}^t Q_i$ can be expressed as the* disjoint *union*

$$Q = \bigcup_{i=1}^{t} \left(\bigcap_{j=1}^{i-1} Q_j^c \right) Q_i = Q_1 \cup Q_1^c Q_2 \cup \cdots \cup Q_1^c Q_2^c \cdots Q_{t-1}^c Q_t \tag{5.31}$$

where we have dropped the intersection symbols to reduce clutter.

Proof: Obviously any point in Q belongs to $\bigcup_{i=1}^t Q_i$. Conversely, for any point y in $\bigcup_{i=1}^t Q_i$ there must be some smallest $i \leq t$ such that $y \in Q_i$ and $y \notin Q_j$ for all $j < i$. □

Before proceeding to the formal definition of the algorithms, let us give a high-level description of how the first of these algorithms will work in two simplified cases.

5.2.2 Example 1: Set Covering

Consider first the case where $A_i \subseteq \{1, \ldots, n\},\ i = 1, \ldots, m$, and

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} Y_j. \tag{5.32}$$

We noted already that (5.32) is equivalent to (5.9), which is the collection of set coverings of the A_i. In line (essentially) with the notation of (5.12), we represent (5.32) as,

$$P = \bigcap_{i=1}^{m} \bigcup_{j=1}^{t(i)} Y_{f(i,j)} \tag{5.33}$$

where $A_i = \{f(i,1), \ldots, f(i,t(i))\}$. The analysis will proceed in two steps.

Step 1:

Define

$$R_i = \bigcup_{j=1}^{t(i)} Y_{f(i,j)} \tag{5.34}$$

so that

$$P = \bigcap_{i=1}^{m} R_i. \tag{5.35}$$

By Lemma 5.2, we can partition R_i into the disjoint union

$$R_i = \bigcup_{j=1}^{t(i)} \left(\bigcap_{\bar{j}=1}^{j-1} N_{f(i,\bar{j})} \cap Y_{f(i,j)} \right) \tag{5.36}$$

and therefore

$$P = \bigcap_{\substack{\bar{i}=1 \\ \bar{i} \neq i}}^{m} R_{\bar{i}} \cap \bigcup_{j=1}^{t(i)} \left(\bigcap_{\bar{j}=1}^{j-1} N_{f(i,\bar{j})} \cap Y_{f(i,j)} \right). \tag{5.37}$$

Thus if we define

$$T(i,j) = \bigcap_{\substack{\bar{i}=1 \\ \bar{i} \neq i}}^{m} R_{\bar{i}} \cap \bigcap_{\bar{j}=1}^{j-1} N_{f(i,\bar{j})} \tag{5.38}$$

then P can be partitioned into the disjoint union

$$P = \bigcup_{j=1}^{t(i)} (T(i,j) \cap Y_{f(i,j)}). \tag{5.39}$$

Example: Let

$$P = (Y_1 \cup Y_3 \cup Y_4)(Y_2 \cup Y_5)(Y_1 \cup Y_5 \cup Y_6) \tag{5.40}$$

(where we have dropped the intersection symbols). Then

$$R_1 = Y_1 \cup Y_3 \cup Y_4 \tag{5.41}$$

$$R_2 = Y_2 \cup Y_5 \tag{5.42}$$

$$R_3 = Y_1 \cup Y_5 \cup Y_6 \tag{5.43}$$

$$T(1,1) = R_2 R_3 \tag{5.44}$$

$$T(1,2) = R_2 R_3 N_1 \tag{5.45}$$

$$T(1,3) = R_2 R_3 N_1 N_3 \tag{5.46}$$

$$T(2,1) = R_1 R_3 \tag{5.47}$$

$$T(2,2) = R_1 R_3 N_2 \tag{5.48}$$

$$T(3,1) = R_1 R_2 \tag{5.49}$$

$$T(3,2) = R_1 R_2 N_1 \tag{5.50}$$

$$T(3,3) = R_1 R_2 N_1 N_5 \tag{5.51}$$

and

$$P = R_1 R_2 R_3 = R_2 R_3 Y_1 \cup R_2 R_3 N_1 Y_3 \cup R_2 R_3 N_1 N_3 Y_4 = \tag{5.52}$$

$$T(1,1)Y_1 \cup T(1,2)Y_3 \cup T(1,3)Y_4 = \tag{5.53}$$

$$T(1,1)Y_{f(1,1)} \cup T(1,2)Y_{f(1,2)} \cup T(1,3)Y_{f(1,3)}. \tag{5.54}$$

Similarly,

$$P = R_1 R_2 R_3 = R_1 R_3 Y_2 \cup R_1 R_3 N_2 Y_5 = \tag{5.55}$$

$$T(2,1)Y_{f(2,1)} \cup T(2,2)Y_{f(2,2)} \tag{5.56}$$

and

$$P = R_1 R_2 R_3 = R_1 R_2 Y_1 \cup R_1 R_2 N_1 Y_5 \cup R_1 R_2 N_1 N_5 Y_6 = \tag{5.57}$$

$$T(3,1)Y_{f(3,1)} \cup T(3,2)Y_{f(3,2)} \cup T(3,3)Y_{f(3,3)}. \ \square \tag{5.58}$$

Thus for each $i = 1, \ldots, m$, P is a disjoint union of sets each of which is a subset of some $Y_{f(i,j)}$. Consider now that for any (signed) measure χ on $\mathcal{A}$ (the subset algebra of $\{0,1\}^n$), and any disjoint pair of sets $u, v \in \mathcal{A}$, the partial sum (signed) measures χ^u and χ^v satisfy

$$\chi^u + \chi^v = \chi^{u \cup v}. \tag{5.59}$$

Recall (Lemma 3.27 and Remark 3.28) that a measure χ on $\mathcal{A}$ defines a $\mathcal{P}$-measure (in the sense that there is a measure $\bar{\chi}$ on $\mathcal{P}$ with $\bar{\chi}[q \cap P] = \chi[q],\ \forall q \in \mathcal{A}$), iff $\chi(P^c) = 0$, i.e. iff $\chi^P = \chi$. Recall also that a vector $(x_1, \dots, x_n) \in R^n$ belongs to $Conv(P)$ iff there exists a measure χ on $\mathcal{A}$ that defines a $\mathcal{P}$-measure in this sense (i.e. $\chi^P = \chi$), such that $\chi[Y_i] = x_i,\ i = 1, \dots, n$, and such that $\chi[\{0,1\}^n] = \chi[P] = 1$. Thus by the disjointness of (5.39), any vector $(x_1, \dots, x_n) \in Conv(P)$ can be lifted to a probability measure χ on $\mathcal{A}$, with $\chi[Y_l] = x_l,\ l = 1, \dots, n$,[1] for which

$$\chi = \chi^P = \sum_{j=1}^{t(i)} \chi^{T(i,j) \cap Y_{f(i,j)}} \tag{5.60}$$

and for which $1 = \chi[\{0,1\}^n] = \chi[P]$. More generally, any $(x_0, x_1, \dots, x_n) \in R^{n+1}$ such that $(x_0, \dots, x_n) \in Cone(K(P))$ (Definition 1.2) can be lifted to a measure χ on $\mathcal{A}$ with $x_0 = \chi[\{0,1\}^n] = \chi[P],\ x_l = \chi[Y_l],\ l = 1, \dots, n$, for which (5.60) holds. Therefore any vector $x \in Cone(K(P))$ must be expressible as a sum of partial sum vectors (i.e. projections of partial sum measures)

$$x = \sum_{j=1}^{t(i)} x^{T(i,j) \cap Y_{f(i,j)}}. \tag{5.61}$$

(Recall that if $(x_0, \dots, x_n) \in Cone(K(P))$, then either $(x_0, \dots, x_n) = 0$ or $\frac{1}{x_0}(x_1, \dots, x_n) \in Conv(P)$.)

We will now show how to ensure that for every valid pitch k constraint, $\alpha^T x \geq \beta$ on P, the vector $(x_0, \dots, x_n)$, which will be construed as $(x[P], x[Y_1], \dots, x[Y_n])$, can be made to satisfy the homogenized constraint, $\alpha^T x \geq \beta x_0$. Our plan is to introduce new vectors of the same dimension to correspond with

$$x^{T(i,j) \cap Y_{f(i,j)}} = (x_0^{T(i,j) \cap Y_{f(i,j)}}, x_1^{T(i,j) \cap Y_{f(i,j)}}, \dots, x_n^{T(i,j) \cap Y_{f(i,j)}}) = \tag{5.62}$$

$$(x^{T(i,j) \cap Y_{f(i,j)}}[P], x^{T(i,j) \cap Y_{f(i,j)}}[Y_1], \dots, x^{T(i,j) \cap Y_{f(i,j)}}[Y_n]) \tag{5.63}$$

and to use the properties of partial summation to put valid constraints on *these* vectors that ensure that for each valid pitch $\leq k$ constraint, $\alpha^T x \geq \beta$, there will be some $i \in \{1, \dots, m\}$ such that every vector $x^{T(i,j) \cap Y_{f(i,j)}},\ j = 1, \dots, t(i)$, will satisfy $\alpha^T x \geq \beta \chi_0$. If this can be accomplished, then by enforcing (5.61), it will follow that for each valid $\alpha^T x \geq \beta$ of pitch $\leq k$, the vector $(x_0, \dots, x_n)$, as a sum of vectors each of which satisfy $\alpha^T x \geq \beta x_0$, must itself also satisfy $\alpha^T x \geq \beta x_0$.

[1] Note that $\chi[Y_l] = \chi[Y_l^P]$ since by assumption $\chi[P^c] = 0$. Nevertheless we will be describing sets throughout this chapter mostly by set theoretic expressions involving Y_l rather than Y_l^P, as it was felt that the presentation will be clearer this way. In the following chapter, however, it will be more convenient to describe sets by expressions involving Y_l^P.

To this end, observe first that if $(x_0, \ldots, x_n)$ can be lifted to a measure on $\mathcal{A}$ (with each new q'th coordinate denoted $x[q]$), then any partial sum vector $x^{T(i,j)\cap Y_{f(i,j)}}$ must satisfy

$$x^{T(i,j)\cap Y_{f(i,j)}}[Y_{f(i,j)}] = x[T(i,j) \cap Y_{f(i,j)} \cap Y_{f(i,j)}] = \tag{5.64}$$

$$x[T(i,j) \cap Y_{f(i,j)}] = x[T(i,j) \cap Y_{f(i,j)} \cap P] = x^{T(i,j)\cap Y_{f(i,j)}}[P]. \tag{5.65}$$

(Alternatively, this can be seen by noting that the partial sum vector for $T(i,j) \cap Y_{f(i,j)}$ is a nonnegative linear combination of the zeta vectors of the atoms that belong to the set $T(i,j) \cap Y_{f(i,j)}$, all of which satisfy $\zeta[P] = \zeta[T(i,j) \cap Y_{f(i,j)}]$.) Thus for each $i \in \{1, \ldots, m\}$, the vector $(x_0, \ldots, x_n)$ can be decomposed by (5.61) into a sum of vectors each of which can be validly constrained by $x_{f(i,j)} = x_0$ for some j.

Consider now that for any pitch k constraint, $k \geq 1$, $\alpha^T x \geq \beta$, that is valid for P, it must be that

$$\textit{support}(\alpha) \supseteq A_i \text{ for some } i \in \{1, \ldots, m\} \tag{5.66}$$

(we will prove this formally later). Consider also that for any $l \in \textit{support}(\alpha)$, the valid constraint

$$\bar{\alpha}^T x \geq \beta - \alpha_l \tag{5.67}$$

where $\bar{\alpha}$ is the same as α but with $\bar{\alpha}_l = 0$, has pitch strictly smaller than k (to be proven later). Observe moreover that if a vector x satisfies $\bar{\alpha}^T x \geq \beta - \alpha_l$ as well as $x_l = 1$ then it must also satisfy $\alpha^T x \geq \beta$, or more generally, if x satisfies $\bar{\alpha}^T x \geq (\beta - \alpha_l)x_0$, as well as $x_l = x_0$, then it must also satisfy $\alpha^T x \geq \beta x_0$.

Putting these facts together, we conclude that if each of the vectors $x^{T(i,j)\cap Y_{f(i,j)}}$ into which we have decomposed $(x_0, \ldots, x_n)$ can also be guaranteed to satisfy all of the valid constraints of pitch *less* than k, then for any valid constraint, $\alpha^T x \geq \beta$ of pitch $\leq k$, choosing $i \in \{1, \ldots, m\}$ such that $\textit{support}(\alpha) \supseteq A_i$, each vector $x^{T(i,j)\cap Y_{f(i,j)}}$ will satisfy $\alpha^T x \geq \beta x_0$ as well, since it satisfies $\bar{\alpha}^T x \geq (\beta - \alpha_{f(i,j)})x_0$ by assumption, and it is constrained to satisfy $x_{f(i,j)} = x_0$. Thus we will conclude that for each valid pitch $\leq k$ constraint, $\alpha^T x \geq \beta$, the vector $(x_0, \ldots, x_n)$ also satisfies $\alpha^T x \geq \beta x_0$, as it is a sum of vectors that each satisfy that constraint.

Example: Consider, for example, the set P defined by

$$P = (Y_1 \cup Y_2)(Y_1 \cup Y_3)(Y_2 \cup Y_3) \tag{5.68}$$

i.e.

$$P = \bigcap_{i=1}^{3} \bigcup_{j=1}^{t(i)} Y_{f(i,j)} \tag{5.69}$$

with $t(i) = 2$ for each $i = 1, \ldots, 3$, and with

$$f(1,1) = 1,\ f(1,2) = 2,\ f(2,1) = 1,\ f(2,2) = 3,\ f(3,1) = 2,\ f(3,2) = 3 \tag{5.70}$$

Stated another way, P is the set of points in $\{0,1\}^n$ that satisfy the system of constraints:

$$y_1 + y_2 \geq 1 \tag{5.71}$$

$$y_1 + y_3 \geq 1 \tag{5.72}$$

$$y_2 + y_3 \geq 1. \tag{5.73}$$

In this case we have

$$R_1 = Y_1 \cup Y_2 \tag{5.74}$$

$$R_2 = Y_1 \cup Y_3 \tag{5.75}$$

$$R_3 = Y_2 \cup Y_3 \tag{5.76}$$

and

$$T(1,1) = R_2 R_3, \quad T(1,2) = R_2 R_3 N_1 \tag{5.77}$$

$$T(2,1) = R_1 R_3, \quad T(2,2) = R_1 R_3 N_1 \tag{5.78}$$

$$T(3,1) = R_1 R_2, \quad T(3,2) = R_1 R_2 N_2. \tag{5.79}$$

Observe that the pitch 2 inequality

$$y_1 + y_2 + y_3 \geq 2 \tag{5.80}$$

is valid for P. In addition to the vector $(x_0, x_1, x_2, x_3) = (x[P], x[Y_1], x[Y_2], x[Y_3])$, the algorithm will define vectors $x^{T(i,j) \cap Y_{f(i,j)}}$ for each (i,j), all of which will have coordinates for P and for each of the sets Y_1, Y_2, Y_3, and will demand that

$$x = x^{T(1,1) \cap Y_1} + x^{T(1,2) \cap Y_2} \tag{5.81}$$

$$x = x^{T(2,1) \cap Y_1} + x^{T(2,2) \cap Y_3} \tag{5.82}$$

$$x = x^{T(3,1) \cap Y_2} + x^{T(3,2) \cap Y_3}. \tag{5.83}$$

The algorithm will also enforce

$$x^{T(i,j) \cap Y_{f(i,j)}}[Y_{f(i,j)}] = x^{T(i,j) \cap Y_{f(i,j)}}[P]. \tag{5.84}$$

Let us now assume that each of the vectors $x^{T(i,j)\cap Y_{f(i,j)}}$ also satisfies all valid pitch 1 constraints, so for example, we will have

$$x^{T(1,1)\cap Y_1}[Y_1] + x^{T(1,1)\cap Y_1}[Y_2] \geq x^{T(1,1)\cap Y_1}[P] \tag{5.85}$$

$$x^{T(1,1)\cap Y_1}[Y_1] + x^{T(1,1)\cap Y_1}[Y_3] \geq x^{T(1,1)\cap Y_1}[P] \tag{5.86}$$

$$x^{T(1,1)\cap Y_1}[Y_2] + x^{T(1,1)\cap Y_1}[Y_3] \geq x^{T(1,1)\cap Y_1}[P]. \tag{5.87}$$

Observe now that by (5.84), we have

$$x^{T(1,1)\cap Y_1}[Y_1] = x^{T(1,1)\cap Y_1}[P] \tag{5.88}$$

which together with (5.87) implies

$$x^{T(1,1)\cap Y_1}[Y_1] + x^{T(1,1)\cap Y_1}[Y_2] + x^{T(1,1)\cap Y_1}[Y_3] \geq 2x^{T(1,1)\cap Y_1}[P] \tag{5.89}$$

so we conclude that $x^{T(1,1)\cap Y_1}$ indeed satisfies the pitch 2 inequality (5.80). This result is not merely accidental. Considering that $\{1\}$ belongs to the support of inequality (5.80), and considering (5.88), in order to guarantee that $x^{T(1,1)\cap Y_1}$ indeed satisfies the pitch 2 inequality (5.80), it suffices to establish that $x^{T(1,1)\cap Y_1}$ satisfies the (homogenized) pitch 1 inequality (5.87) obtained by zeroing out the Y_1 coordinate in (5.80) and subtracting its coefficient from its right hand side .

But it is not only $\{1\}$ that belongs to the support of inequality (5.80), $\{2\}$ does as well, and in general the support of any valid pitch 2 constraint must always contain the support of some valid pitch 1 constraint. Thus since, by (5.84), we have

$$x^{T(1,2)\cap Y_2}[Y_2] = x^{T(1,2)\cap Y_2}[P] \tag{5.90}$$

we need only establish that $x^{T(1,2)\cap Y_2}$ satisfies the valid pitch 1 constraint

$$x^{T(1,2)\cap Y_2}[Y_1] + x^{T(1,2)\cap Y_2}[Y_3] \geq x^{T(1,2)\cap Y_2}[P]. \tag{5.91}$$

to guarantee

$$x^{T(1,2)\cap Y_2}[Y_1] + x^{T(1,21)\cap Y_2}[Y_2] + x^{T(1,2)\cap Y_2}[Y_3] \geq 2x^{T(1,2)\cap Y_2}[P]. \tag{5.92}$$

Thus if we maintain the assumption that each of the vectors $x^{T(i,j)\cap Y_{f(i,j)}}$ satisfies all valid pitch 1 constraints, then it will follow that $x^{T(1,2)\cap Y_2}$ also satisfies the pitch 2 inequality (5.80). Constraint (5.81) will now imply that x must satisfy (5.80) as well. □

Step 2:

We have thus established that if the vectors $x^{T(i,j)\cap Y_{f(i,j)}}$ can be guaranteed to satisfy all valid pitch $\leq k-1$ constraints then x will satisfy all valid pitch k constraints. We now need to establish that the vectors $x^{T(i,j)\cap Y_{f(i,j)}}$ can indeed be ensured to satisfy all valid pitch $\leq k-1$ constraints.

Define

$$T(\{i,j\}) = T(i,j) \cap Y_{f(i,j)}. \tag{5.93}$$

For the moment, let us assume that $m > 1$. Consider first that Lemma 5.2 implies that for each $i' \in \{1,\ldots,m\} - \{i\}$, the set $T(\{i,j\})$ can itself be partitioned as follows

$$T(\{i,j\}) = \bigcap_{\substack{\bar{i}=1 \\ \bar{i}\neq i}}^{m} R_{\bar{i}} \cap \bigcap_{\bar{j}=1}^{j-1} N_{f(i,\bar{j})} \cap Y_{f(i,j)} = \tag{5.94}$$

$$\left(\bigcup_{j'=1}^{t(i')} Y_{f(i',j')}\right) \cap \bigcap_{\substack{\bar{i}=1 \\ \bar{i}\neq i,i'}} R_{\bar{i}} \cap \bigcap_{\bar{j}=1}^{j-1} N_{f(i,\bar{j})} \cap Y_{f(i,j)} = \tag{5.95}$$

$$\bigcup_{j'=1}^{t(i')} \left(\bigcap_{\bar{j}'=1}^{j'-1} N_{f(i',\bar{j}')} \cap Y_{f(i',j')}\right) \cap \bigcap_{\substack{\bar{i}=1 \\ \bar{i}\neq i,i'}} R_{\bar{i}} \cap \bigcap_{\bar{j}=1}^{j-1} N_{f(i,\bar{j})} \cap Y_{f(i,j)} = \tag{5.96}$$

$$\bigcup_{j'=1}^{t(i')} \left(\bigcap_{\substack{\bar{i}=1 \\ \bar{i}\neq i,i'}} R_{\bar{i}} \cap \bigcap_{\bar{j}=1}^{j-1} N_{f(i,\bar{j})} \cap \bigcap_{\bar{j}'=1}^{j'-1} N_{f(i',\bar{j}')} \cap Y_{f(i,j)} \cap Y_{f(i',j')} \right) = \tag{5.97}$$

$$\bigcup_{j'=1}^{t(i')} \left(T(\{i,j\}) \cap T(\{i',j'\}\right) \tag{5.98}$$

where the last equality follows from the fact that $Y_{f(i,j)} \subseteq R_i$ and $Y_{f(i',j')} \subseteq R_{i'}$. Define now

$$T(\{i,j\},\{i',j'\}) = T(\{i,j\}) \cap T(\{i',j'\}). \tag{5.99}$$

Thus for each $i' \in \{1,\ldots,m\}$ other than i, each partial sum vector $x^{T(\{i,j\})}$ can itself be decomposed into the sum of partial sum vectors

$$x^{T(\{i,j\})} = \sum_{j'=1}^{t(i')} x^{T(\{i,j\},\{i',j'\})}. \tag{5.100}$$

Note moreover that we can also enforce

$$x^{T(\{i,j\},\{i',j'\})}[Y_{f(i,j)}] = x^{T(\{i,j\},\{i',j'\})}[Y_{f(i',j')}] = x^{T(\{i,j\},\{i',j'\})}[P]. \tag{5.101}$$

Example: For the example from Step 1, i.e. where P is as in (5.68), we would write

$$x^{T(1,1)\cap Y_1} = x^{T(\{1,1\})} = x^{T(\{1,1\},\{2,1\})} + x^{T(\{1,1\},\{2,2\})} = \tag{5.102}$$

$$x^{T(1,1)\cap T(2,1)\cap Y_1} + x^{T(1,1)\cap T(2,2)\cap Y_1\cap Y_3} \tag{5.103}$$

and

$$x^{T(\{1,1\})} = x^{T(\{1,1\},\{3,1\})} + x^{T(\{1,1\},\{3,2\})} = \tag{5.104}$$

$$x^{T(1,1)\cap T(3,1)\cap Y_1\cap Y_2} + x^{T(1,1)\cap T(3,2)\cap Y_1\cap Y_3} \tag{5.105}$$

and enforcing (5.101) amounts to writing

$$x^{T(\{1,1\},\{2,1\})}[Y_1] = x^{T(\{1,1\},\{2,1\})}[P] \tag{5.106}$$

$$x^{T(\{1,1\},\{2,2\})}[Y_1] = x^{T(\{1,1\},\{2,2\})}[Y_3] = x^{T(\{1,1\},\{2,2\})}[P] \tag{5.107}$$

$$x^{T(\{1,1\},\{3,1\})}[Y_1] = x^{T(\{1,1\},\{3,1\})}[Y_2] = x^{T(\{1,1\},\{3,1\})}[P] \tag{5.108}$$

$$x^{T(\{1,1\},\{3,2\})}[Y_1] = x^{T(\{1,1\},\{3,2\})}[Y_3] = x^{T(\{1,1\},\{3,2\})}[P]. \tag{5.109}$$

Note now that the valid pitch 1 inequalities where P is as in (5.68) are all dominated by the valid constraints

$$y_1 + y_2 \geq 1,\ y_1 + y_3 \geq 1,\ y_2 + y_3 \geq 1 \tag{5.110}$$

and that by (5.108) and (5.109) the vectors $x^{T(\{1,1\},\{3,1\})}$ and $x^{T(\{1,1\},\{3,2\})}$ all satisfy all three of these constraints (homogenized). Thus by (5.104) it follows that $x^{T(1,1)\cap Y_1} = x^{T(\{1,1\})}$ satisfies all of the pitch 1 constraints as well. Similar arguments apply for all of the vectors $x^{T(i,j)\cap Y_{f(i,j)}} = x^{T(\{i,j\})}$. □

Consider now a valid pitch $k-1$ constraint, $\alpha^T x \geq \beta$. As above, there must be some $A_l \subseteq$ *support*(α). Suppose first that $A_i \subseteq$ *support*(α). Thus for each $j = 1, \ldots, t(i)$, the valid constraint $(\bar{\alpha}^j)^T x \geq \beta - \alpha_j$ (where $\bar{\alpha}$ is the same as α but with $\alpha_j = 0$) is of pitch $\leq k-2$. Thus if we assume that $x^{T(\{i,j\})}$ satisfies all valid constraints of pitch $\leq k-2$, (and this will hold if all vectors $x^{T(\{i,j\},\ \{i',j'\})}$ satisfy all pitch $\leq k-2$ constraints), then considering that $x^{T(\{i,j\})}[Y_{f(i,j)}] = x^{T(\{i,j\})}[P]$, it will follow that $x^{T(\{i,j\})}$ satisfies the constraint $\alpha^T x \geq \beta x_0$ as well. If, on the other hand, $A_l \subseteq$ *support*(α), $l \neq i$, then a similar argument to the one used in Step 1 shows that if all of the vectors $x^{T(\{i,j\},\ \{i',j'\})}$ satisfy all valid pitch $k-2$ constraints, then $x^{T(\{i,j\})}$ satisfies the constraint $\alpha^T x \geq \beta x_0$ too.

We have been assuming so far that $m > 1$ so that there is in fact an i' in $\{1, \ldots, m\}$ other than i. If, however, $m = 1 = i$, then for every valid pitch $k-1$ constraint $\alpha^T x \geq \beta$, we must have $A_i \subseteq \mathit{support}(\alpha)$. Thus as above, so long as $x^{T(\{i,j\})}$ satisfies all pitch $k-2$ constraints then it will satisfy $\alpha^T x \geq \beta x_0$ as well. But the support of any pitch $k-2$ constraint $\hat{\alpha}^T x \geq \hat{\beta}$ similarly must contain A_i, and so the same reasoning implies that if $x^{T(\{i,j\})}$ satisfies all pitch $k-3$ constraints then it satisfies $\hat{\alpha}^T x \geq \hat{\beta}$ as well. Noting that the pitch 0 constraints are just the nonnegativity constraints, then it is easy to see by induction that so long as we impose nonnegativity, $x^{T(\{i,j\})}$ will satisfy all pitch $k-1$ constraints.

Returning now to the case $m > 1$, we need to show how to ensure that each of the vectors $x^{T(\{i,j\},\ \{i',j'\})}$ satisfies all valid constraints of pitch $\leq k-2$. If $m = 2$, then for any valid pitch $k-2$ constraint $\alpha^T x \geq \beta$, it must be that either $A_i \subseteq \mathit{support}(\alpha)$ or $A_{i'} \subseteq \mathit{support}(\alpha)$. Thus so long as $x^{T(\{i,j\},\{i',j'\})}$ satisfies all valid pitch $\leq k-3$ constraints then it must also satisfy $\alpha^T x \geq \beta$. As above, repeating the argument will show that $x^{T(\{i,j\},\{i',j'\})}$ will satisfy all valid pitch $k-2$ constraints.

If, however, $m > 2$, then the decomposition procedure we outlined can be again repeated to partition sets $T(\{i,j\},\{i',j'\})$ into disjoint unions of sets

$$T(\{i,j\},\{i',j'\},\{i'',j''\}) := T(\{i,j\}) \cap T(\{i',j'\}) \cap T(\{i'',j''\}) \tag{5.111}$$

for each $i'' \neq i, i'$, where the union is over $j'' = 1, \ldots, t(i'')$. It is easy to see that it suffices to establish that each of the partial sum vectors $x^{T(\{i,j\},\{i',j'\},\{i'',j''\})}$ satisfies the pitch $k-3$ constraints in order to guarantee that $x^{T(\{i,j\},\{i',j'\})}$ will satisfy the pitch $k-2$ constraints.

Recalling again that the pitch 0 constraints are all dominated by the nonnegativity constraints, it is easy to see that repeating the procedure until we have taken k-fold decompositions of $(x_0, \ldots, x_n)$ (or m-fold if $m < k$) will guarantee that $(x_0, \ldots, x_n)$ will satisfy all (homogenized) pitch k constraints.

5.2.3 Example 2: Covering Constraints

Here we consider the problem

$$P = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1} \cdots \bigcap_{i_h=1}^{m_h} \bigcup_{j_h=1}^{t_h} Y_{f(i_1,j_1,\cdots,i_h,j_h)}. \tag{5.112}$$

The valid pitch 1 constraints for this problem are dominated by the constraints (all of which are valid) of the form

$$\sum_{j_1=1}^{t_1} \sum_{j_2=1}^{t_2} \cdots \sum_{j_h=1}^{t_h} x_{f(i_1,j_1,i_2(j_1),j_2\cdots,i_h(j_1,\ldots,j_{h-1}),j_h)} \geq 1. \tag{5.113}$$

In each such constraint the value of i_1 is a constant in the range $\{1, \ldots, m_1\}$; the value of $i_2(j_1)$ varies in the range $\{1, \ldots, m_2\}$ as a function of j_1; the value of $i_3(j_1, j_2)$ varies in the range $\{1,\ldots,m_3\}$ as a function of j_1 and j_2, etcetera. In other words, the elements of a sum of the form (5.113) for which $j_1 = 1$ may have a different i_2 value than the elements of the sum for which $j_1 = 2$ (though all elements with $j_1 = 1$ will have the same i_2 value). Similarly, the elements of the sum with $j_1 = 1$ and $j_2 = 3$ can have a different i_3 value than those with $j_1 = 2$ and $j_2 = 3$. In general, for each term of the sum indexed by a given $j_1 = \bar{j}_1, \ldots, j_l = \bar{j}_l$ there can be a different choice of i_{l+1} from the range $\{1,\ldots,m_{l+1}\}$. There is a valid constraint of the form (5.113) for each of the exponentially many h-tuples of functions $(i_1, i_2(\cdot), i_3(\cdot), \ldots, i_h(\cdot))$. (We will formally prove all of this later.)

Example: Consider

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.114}$$

or in our notation

$$P = \bigcap_{i_1=1}^{1} \bigcup_{j_1=1}^{2} \bigcap_{i_2=1}^{2} \bigcup_{j_2=1}^{2} Y_{f(i_1,j_1,i_2,j_2)} \tag{5.115}$$

where f maps

$$(1,1,1,1) \to 1,\ (1,1,1,2) \to 2,\ (1,1,2,1) \to 3, \ldots, (1,2,2,2) \to 8. \tag{5.116}$$

This corresponds to the integer program

$$P = \{y \in \{0,1\}^8 : (y_1 + y_2)(y_3 + y_4) + (y_5 + y_6)(y_7 + y_8) \geq 1\}. \tag{5.117}$$

The constraints of the form (5.113) are

$$\sum_{j_1=1}^{2} \sum_{j_2=1}^{2} x_{f(i_1,j_1,i_2(j_1),j_2)} \geq 1 \tag{5.118}$$

and there is such a constraint for each function $i_2(j_1) : \{1,2\} \to \{1,2\}$ (here there is only one possible choice for i_1). Specifically these constraints are:

$$y_{f(1,1,1,1)} + y_{f(1,1,1,2)} + y_{f(1,2,1,1)} + y_{f(1,2,1,2)} = y_1 + y_2 + y_5 + y_6 \geq 1 \tag{5.119}$$

$$y_{f(1,1,1,1)} + y_{f(1,1,1,2)} + y_{f(1,2,2,1)} + y_{f(1,2,2,2)} = y_1 + y_2 + y_7 + y_8 \geq 1 \tag{5.120}$$

$$y_{f(1,1,2,1)} + y_{f(1,1,2,2)} + y_{f(1,2,1,1)} + y_{f(1,2,1,2)} = y_3 + y_4 + y_5 + y_6 \geq 1 \tag{5.121}$$

$$y_{f(1,1,2,1)} + y_{f(1,1,2,2)} + y_{f(1,2,2,1)} + y_{f(1,2,2,2)} = y_3 + y_4 + y_7 + y_8 \geq 1 \tag{5.122}$$

□

We will now show how to ensure that a vector $(x_0, \ldots, x_n)$, which will be denoted here simply as x, will satisfy all of the exponentially many pitch 1 constraints (5.113).

Let us first rewrite P as follows:

$$P = \bigcap_{i_1=1}^{m_1} R_{i_1} \quad \text{where} \tag{5.123}$$

$$R_{i_1} = \bigcup_{j_1=1}^{t_1} Q_{i_1,j_1} \quad \text{where} \tag{5.124}$$

$$Q_{i_1,j_1} = \bigcap_{i_2=1}^{m_2} R_{i_1,j_1,i_2} \quad \text{where} \tag{5.125}$$

$$R_{i_1,j_1,i_2} = \bigcup_{j_2=1}^{t_2} Q_{i_1,j_1,i_2,j_2} \quad \text{where} \tag{5.126}$$

$$\vdots \tag{5.127}$$

$$Q_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1}} = \bigcap_{i_h=1}^{m_h} R_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h} \quad \text{where} \tag{5.128}$$

$$R_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h} = \bigcup_{j_h=1}^{t_h} Q_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h,j_h} \quad \text{where} \tag{5.129}$$

$$Q_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h,j_h} = Y_{f(i_1,j_1,\ldots,i_h,j_h)} \tag{5.130}$$

and, as in the previous subsection, we will introduce new "partial sum" vectors with coordinates for P and for each Y_i. For each $i_1 = 1, \ldots, m_1$, we will then decompose x into the sum (over $j_1 = 1, \ldots, t_1$) of the partial sum vectors $x^{T(\{i_1,j_1\})}$, where

$$T(\{i_1, j_1\}) = T(i_1, j_1) \cap Q_{i_1,j_1} \quad \text{and} \tag{5.131}$$

$$T(i_1, j_1) = \bigcap_{\substack{\bar{i}_1=1 \\ \bar{i}_1 \neq i_1}}^{m_1} R_{\bar{i}_1} \cap \bigcap_{\bar{j}_1=1}^{j_1-1} (Q_{i_1,\bar{j}_1})^c. \tag{5.132}$$

Observe now that for each $i_2 \in \{1, \ldots, m_2\}$, Q_{i_1,j_1} can itself be partitioned as

$$Q_{i_1,j_1} = \bigcap_{\substack{\bar{i}_2=1 \\ \bar{i}_2 \neq i_2}}^{m_2} R_{i_1,j_1,\bar{i}_2} \cap \bigcup_{j_2=1}^{t_2} \left(\bigcap_{\bar{j}_2=1}^{j_2-1} (Q_{i_1,j_1,i_2,\bar{j}_2})^c \cap Q_{i_1,j_1,i_2,j_2} \right) = \tag{5.133}$$

$$\bigcup_{j_2=1}^{t_2} \left(\bigcap_{\substack{\bar{i}_2=1 \\ \bar{i}_2 \neq i_2}}^{m_2} R_{i_1,j_1,\bar{i}_2} \cap \bigcap_{\bar{j}_2=1}^{j_2-1} (Q_{i_1,j_1,i_2,\bar{j}_2})^c \cap Q_{i_1,j_1,i_2,j_2} \right). \tag{5.134}$$

Defining now

$$T(i_1, j_1, i_2, j_2) = \bigcap_{\substack{\bar{i}_2=1 \\ \bar{i}_2 \neq i_2}}^{m_2} R_{i_1,j_1,\bar{i}_2} \cap \bigcap_{\bar{j}_2=1}^{j_2-1} (Q_{i_1,j_1,i_2,\bar{j}_2})^c, \tag{5.135}$$

we can therefore write

$$Q_{i_1,j_1} = \bigcup_{j_2=1}^{t_2} \left(T(i_1, j_1, i_2, j_2) \cap Q_{i_1,j_1,i_2,j_2}\right). \tag{5.136}$$

Thus where we define

$$T(\{i_1, j_1, i_2, j_2\}) = T(i_1, j_1) \cap T(i_1, j_1, i_2, j_2) \cap Q_{i_1,j_1,i_2,j_2}, \tag{5.137}$$

we have the partition

$$T(\{i_1, j_1\}) = \bigcup_{j_2=1}^{t_2} T(\{i_1, j_1, i_2, j_2\}) \tag{5.138}$$

for each $i_2 = 1, \ldots, m_2$. It now follows that we can decompose, for each $i_2 = 1, \ldots, m_2$, the vector $x^{T(\{i_1,j_1\})}$ as

$$x^{T(\{i_1,j_1\})} = \sum_{j_2=1}^{t_2} x^{T(\{i_1,j_1,i_2,j_2\})} \tag{5.139}$$

which implies that

$$x = \sum_{j_1=1}^{t_1} \sum_{j_2=1}^{t_2} x^{T(\{i_1,j_1,i_2,j_2\})} \tag{5.140}$$

for all choices of i_1 and i_2. It is important to observe that in any such sum, the choice of i_2 needs not be constant. Each choice of values for j_1 can have a different corresponding i_2 value, and thus we may more precisely write

$$x = \sum_{j_1=1}^{t_1} \sum_{j_2=1}^{t_2} x^{T(\{i_1,j_1,i_2(j_1),j_2\})} \tag{5.141}$$

for all choices of i_1 and $i_2(j_1)$ in the appropriate ranges.

In general, for each $l \leq h$ we will define

$$T(i_1, j_1, \ldots, i_l, j_l) = \tag{5.142}$$

$$\bigcap_{\substack{\bar{i}_l=1 \\ \bar{i}_l \neq i_l}}^{m_l} R_{i_1,j_1,\ldots,i_{l-1},j_{l-1},\bar{i}_l} \cap \bigcap_{\bar{j}_l=1}^{j_l-1} (Q_{i_1,j_1,\ldots,i_{l-1},j_{l-1},i_l,\bar{j}_l})^c \tag{5.143}$$

and

$$T(\{i_1, j_1, \ldots, i_l, j_l\}) = T(i_1, j_1) \cap \cdots \cap T(i_1, j_1, \ldots, i_l, j_l) \cap Q_{i_1,j_1,\ldots,i_l,j_l} \tag{5.144}$$

and, for each $i_l = 1, \ldots, m_l$, we will have partitions

$$T(\{i_1, j_1, \ldots, i_{l-1}, j_{l-1}\}) = \bigcup_{j_l=1}^{t_l} T(\{i_1, j_1, \ldots, i_l, j_l\}). \tag{5.145}$$

The algorithm (after introducing the appropriate partial sum vectors) will therefore decompose each partial sum $x^{T(\{i_1,j_1,\ldots,i_{l-1},j_{l-1}\})}$, for each $i_l = 1,\ldots,m_l$, as

$$x^{T(\{i_1,j_1,\ldots,i_{l-1},j_{l-1}\})} = \sum_{j_l=1}^{t_l} x^{T(\{i_1,j_1,\ldots,i_l,j_l\})} \tag{5.146}$$

and by repeated application for all $l = 2,\ldots,h$, it will ultimately follow that

$$x = \sum_{j_1=1}^{t_1}\sum_{j_2=1}^{t_2}\cdots\sum_{j_h=1}^{t_h} x^{T(\{i_1,j_1,i_2(j_1),j_2,\ldots,i_h(j_1,\ldots,j_{h-1}),j_h\})}. \tag{5.147}$$

As above, in any such decomposition the value of each $i_l(\cdot)$ is a function of $j_1,\ldots,j_{l-1}$. The functions $\{i_l(\cdot)\}$ therefore determine both the decompositions of the form (5.147) and the constraints of the form (5.113), and the decompositions and the constraints are therefore in one to one correspondence.

The algorithm will now impose the valid constraint that each vector

$$x^{T(\{i_1,j_1,\ldots,i_h,j_h\})} = x^{T(i_1,j_1,\ldots,i_h,j_h)\cap T(i_1,j_1,\ldots,i_{h-1},j_{h-1})\cap\cdots\cap T(i_1,j_1)\cap Y_{f(i_1,j_1,\ldots,i_h,j_h)}} \tag{5.148}$$

must satisfy

$$x_{f(i_1,j_1,\ldots,i_h,j_h)} = x_0 \tag{5.149}$$

(i.e. $x^{T(\{i_1,j_1,\ldots,i_h,j_h\})}[Y_{f(i_1,j_1,\ldots,i_h,j_h)}] = x^{T(\{i_1,j_1,\ldots,i_h,j_h\})}[P]$), and it will also impose that all vectors must be nonnegative. It will now follow that for any given choice of functions $\{i_l(\cdot)\}$, to be denoted $\{i'_l(\cdot)\}$, every partial sum vector $x^{T(\{i'_1,j'_1,i'_2(\cdot),j'_2,\ldots,i'_h(\cdot),j'_h\})}$ that appears in the h-fold sum of the form (5.147) defined by $\{i'_l(\cdot)\}$ will satisfy

$$\sum_{j_1=1}^{t_1}\sum_{j_2=1}^{t_2}\cdots\sum_{j_h=1}^{t_h} x_{f(i'_1,j_1,i'_2(\cdot),j_2,\cdots,i'_h(\cdot),j_h)} \geq x_{f(i'_1,j'_1,i'_2(\cdot),j'_2,\ldots,i'_h(\cdot),j'_h)} = x_0 \tag{5.150}$$

and it will therefore follow that x will satisfy

$$\sum_{j_1=1}^{t_1}\sum_{j_2=1}^{t_2}\cdots\sum_{j_h=1}^{t_h} x_{f(i'_1,j_1,i'_2(\cdot),j_2,\cdots,i'_h(\cdot),j_h)} \geq x_0 \tag{5.151}$$

as well. Since the decompositions of the form (5.147) and the constraints of the form (5.113) are in one to one correspondence, it follows that x will satisfy all constraints of the form (5.113).

For example, say that P is as in (5.115), (the example at the beginning of the subsection) and that we have selected $i'_1 = 1,\ i'_2(1) = 2,\ i'_2(2) = 1$. Then the corresponding decomposition (5.147) is

$$x = \sum_{j_1=1}^{2}\sum_{j_2=1}^{2} x^{T(\{1,j_1,i_2(j_1),j_2\})} = x^{T(\{1,1,2,1\})}+x^{T(\{1,1,2,2\})}+x^{T(\{1,2,1,1\})}+x^{T(\{1,2,1,2\})}, \tag{5.152}$$

and the corresponding constraint (5.113) is

$$x_0 \le \sum_{j_1=1}^{2} \sum_{j_2=1}^{2} x_{f(1,j_1,i_2(j_1),j_2)} = x_{f(1,1,2,1)} + x_{f(1,1,2,2)} + x_{f(1,2,1,1)} + x_{f(1,2,1,2)}. \tag{5.153}$$

The algorithm imposes the constraint

$$x^{T(\{1,1,2,1\})}_{f(1,1,2,1)} = x^{T(\{1,1,2,1\})}_0 \tag{5.154}$$

which implies (together with nonnegativity) that

$$x^{T(\{1,1,2,1\})}_{f(1,1,2,1)} + x^{T(\{1,1,2,1\})}_{f(1,1,2,2)} + x^{T(\{1,1,2,1\})}_{f(1,2,1,1)} + x^{T(\{1,1,2,1\})}_{f(1,2,1,2)} \ge x^{T(\{1,1,2,1\})}_{f(1,1,2,1)} \ge x^{T(\{1,1,2,1\})}_0. \tag{5.155}$$

The constraints

$$x^{T(\{1,1,2,2\})}_{f(1,1,2,2)} = x^{T(\{1,1,2,2\})}_0 \tag{5.156}$$

$$x^{T(\{1,2,1,1\})}_{f(1,2,1,1)} = x^{T(\{1,2,1,1\})}_0 \tag{5.157}$$

$$x^{T(\{1,2,1,2\})}_{f(1,2,1,2)} = x^{T(\{1,2,1,2\})}_0 \tag{5.158}$$

similarly imply that $x^{T(\{1,1,2,2\})}, x^{T(\{1,2,1,1\})}$ and $x^{T(\{1,2,1,2\})}$ also satisfy (5.153), and thus by (5.152), x will satisfy (5.153) too.

Example: Here we consider again the case

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.159}$$

and we will show in detail how the procedure outlined will ensure that x will satisfy the pitch 1 constraints (5.119) - (5.122).

In this case $m_1 = 1,\ t_1 = m_2 = t_2 = 2,\ P = R_1$, and

$$R_1 = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.160}$$

$$Q_{1,1} = (Y_1 \cup Y_2) \cap (Y_3 \cup Y_4) \tag{5.161}$$

$$Q_{1,2} = (Y_5 \cup Y_6) \cap (Y_7 \cup Y_8) \tag{5.162}$$

$$R_{1,1,1} = Y_1 \cup Y_2 \tag{5.163}$$

$$R_{1,1,2} = Y_3 \cup Y_4 \tag{5.164}$$

$$R_{1,2,1} = Y_5 \cup Y_6 \tag{5.165}$$

$$R_{1,2,2} = Y_7 \cup Y_8 \tag{5.166}$$

$$T(1,1) = \{0,1\}^n \tag{5.167}$$

$$T(1,2) = Q^c_{1,1} \tag{5.168}$$

$$T(\{1,1\}) = Q_{1,1} \tag{5.169}$$

$$T(\{1,2\}) = Q^c_{1,1} \cap Q_{1,2}. \tag{5.170}$$

We now have the partition

$$P = Q_{1,1} \cup (Q^c_{1,1} \cap Q_{1,2}) = T(\{1,1\}) \cup T(\{1,2\}) \tag{5.171}$$

and we can therefore enforce

$$x = x^{Q_{1,1}} + x^{Q^c_{1,1} \cap Q_{1,2}} = x^{T(\{1,1\})} + x^{T(\{1,2\})}. \tag{5.172}$$

We also have

$$T(1,1,1,1) = R_{1,1,2} \tag{5.173}$$

$$T(1,1,1,2) = R_{1,1,2} \cap N_1 \tag{5.174}$$

$$T(1,1,2,1) = R_{1,1,1} \tag{5.175}$$

$$T(1,1,2,2) = R_{1,1,1} \cap N_3 \tag{5.176}$$

$$T(1,2,1,1) = R_{1,2,2} \tag{5.177}$$

$$T(1,2,1,2) = R_{1,2,2} \cap N_5 \tag{5.178}$$

$$T(1,2,2,1) = R_{1,2,1} \tag{5.179}$$

$$T(1,2,2,2) = R_{1,2,1} \cap N_7 \tag{5.180}$$

$$T(\{1,1,1,1\}) = R_{1,1,2} \cap Y_1 \tag{5.181}$$

$$T(\{1,1,1,2\}) = R_{1,1,2} \cap N_1 \cap Y_2 \tag{5.182}$$

$$T(\{1,1,2,1\}) = R_{1,1,1} \cap Y_3 \tag{5.183}$$

$$T(\{1,1,2,2\}) = R_{1,1,1} \cap N_3 \cap Y_4 \tag{5.184}$$

$$T(\{1,2,1,1\}) = Q^c_{1,1} \cap R_{1,2,2} \cap Y_5 \tag{5.185}$$

$$T(\{1,2,1,2\}) = Q^c_{1,1} \cap R_{1,2,2} \cap N_5 \cap Y_6 \tag{5.186}$$

$$T(\{1,2,2,1\}) = Q^c_{1,1} \cap R_{1,2,1} \cap Y_7 \tag{5.187}$$

$$T(\{1,2,2,2\}) = Q^c_{1,1} \cap R_{1,2,1} \cap N_7 \cap Y_8. \tag{5.188}$$

Observe now that we have the partitions

$$Q_{1,1} = (R_{1,1,1} \cap Y_3) \cup (R_{1,1,1} \cap N_3 \cap Y_4) \tag{5.189}$$

and

$$Q_{1,1} = (R_{1,1,2} \cap Y_1) \cup (R_{1,1,2} \cap N_1 \cap Y_2) \tag{5.190}$$

as well as the partitions

$$Q_{1,2} = (R_{1,2,2} \cap Y_5) \cup (R_{1,2,2} \cap N_5 \cap Y_6) \tag{5.191}$$

and

$$Q_{1,2} = (R_{1,2,1} \cap Y_7) \cup (R_{1,2,1} \cap N_7 \cap Y_8). \tag{5.192}$$

We can therefore enforce the decompositions

$$x^{T(\{1,1\}} = x^{Q_{1,1}} = x^{R_{1,1,2} \cap Y_1} + x^{R_{1,1,2} \cap N_1 \cap Y_2} = \tag{5.193}$$

$$x^{T(\{1,1,1,1\})} + x^{T(\{1,1,1,2\})} \tag{5.194}$$

and

$$x^{T(\{1,1\}} = x^{Q_{1,1}} = x^{R_{1,1,1} \cap Y_3} + x^{R_{1,1,1} \cap N_3 \cap Y_4} = \tag{5.195}$$

$$x^{T(\{1,1,2,1\})} + x^{T(\{1,1,2,2\})} \tag{5.196}$$

and

$$x^{T(\{1,2\})} = x^{Q_{1,1}^c \cap Q_{1,2}} = x^{Q_{1,1}^c \cap R_{1,2,2} \cap Y_5} + x^{Q_{1,1}^c \cap R_{1,2,2} \cap N_5 \cap Y_6} = \tag{5.197}$$

$$x^{T(\{1,2,1,1\})} + x^{T(\{1,2,1,2\})} \tag{5.198}$$

and

$$x^{T(\{1,2\})} = x^{Q_{1,1}^c \cap Q_{1,2}} = x^{Q_{1,1}^c \cap R_{1,2,1} \cap Y_7} + x^{Q_{1,1}^c \cap R_{1,2,1} \cap N_7 \cap Y_8} = \tag{5.199}$$

$$x^{T(\{1,2,2,1\})} + x^{T(\{1,2,2,2\})}. \tag{5.200}$$

Combining with (5.172) therefore implies the four decompositions:

$$x = x^{T(\{1,1,1,1\})} + x^{T(\{1,1,1,2\})} + x^{T(\{1,2,1,1\})} + x^{T(\{1,2,1,2\})} \tag{5.201}$$

$$x = x^{T(\{1,1,1,1\})} + x^{T(\{1,1,1,2\})} + x^{T(\{1,2,2,1\})} + x^{T(\{1,2,2,2\})} \tag{5.202}$$

$$x = x^{T(\{1,1,2,1\})} + x^{T(\{1,1,2,2\})} + x^{T(\{1,2,1,1\})} + x^{T(\{1,2,1,2\})} \tag{5.203}$$

$$x = x^{T(\{1,1,2,1\})} + x^{T(\{1,1,2,2\})} + x^{T(\{1,2,2,1\})} + x^{T(\{1,2,2,2\})} \tag{5.204}$$

Moreover we can enforce

$$x^{T(\{1,1,1,1\})}[Y_1] = x^{T(\{1,1,1,1\})}[P] \tag{5.205}$$

$$x^{T(\{1,1,1,2\})}[Y_2] = x^{T(\{1,1,1,2\})}[P] \tag{5.206}$$

$$x^{T(\{1,1,2,1\})}[Y_3] = x^{T(\{1,1,2,1\})}[P] \tag{5.207}$$

$$x^{T(\{1,1,2,2\})}[Y_4] = x^{T(\{1,1,2,2\})}[P] \tag{5.208}$$

$$x^{T(\{1,2,1,1\})}[Y_5] = x^{T(\{1,2,1,1\})}[P] \tag{5.209}$$

$$x^{T(\{1,2,1,2\})}[Y_6] = x^{T(\{1,2,1,2\})}[P] \tag{5.210}$$

$$x^{T(\{1,2,2,1\})}[Y_7] = x^{T(\{1,2,2,1\})}[P] \tag{5.211}$$

$$x^{T(\{1,2,2,2\})}[Y_8] = x^{T(\{1,2,2,2\})}[P] \tag{5.212}$$

From equations (5.205) - (5.212) it is clear that each term of the decomposition (5.201) satisfies

$$x_1 + x_2 + x_5 + x_6 \geq x_0 \tag{5.213}$$

as for each term of that decomposition at least one from among the first, second, fifth and sixth coordinates must be of value 1. It is similarly clear that each term of the decomposition (5.202) satisfies (5.120) (homogenized), that each term of the decomposition (5.203) satisfies (5.121) (homogenized), and that each term of the decomposition (5.204) satisfies (5.122) (homogenized). We are therefore ensured that x will satisfy all constraints (5.119) - (5.122) (homogenized). □

5.2.4 Pitch of Inequalities

Before going further, we will first prove some facts about the pitch of inequalities for sets of the form

$$P = \bigcap_{i=1}^{m} \bigcup_{j=1}^{t_i} M_{f(i,j)} \tag{5.214}$$

where, as above, f maps into $\{1', 1'', \ldots, n', n''\}$, and for each $l \in \{1, \ldots, n\}$, $M_{l'} = Y_l$, $M_{l''} = N_l$. We will begin by considering the easiest case, i.e. $P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} Y_j$, where $A_i \subseteq \{1, \ldots, n\}$, $i = 1, \ldots, m$.

Lemma 5.3 *If $P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} Y_j$, where $A_i \subseteq \{1, \ldots, n\}$, $i = 1, \ldots, m$, and $\alpha^T x \geq \beta$, $\alpha \geq 0$ is valid for P and of pitch 1, then $\alpha^T x \geq \beta$ is dominated by the inequalities $\sum_{j \in A_i} x_j \geq 1$.*

Proof: Let us assume that the variables are arranged so that

$$0 < \alpha_1 \leq \alpha_2 \leq \cdots < \alpha_{|support(\alpha)|} \tag{5.215}$$

We must have $\beta > 0$ or else the pitch would be zero. Thus by the definition of pitch $\alpha^T x \geq \beta$ is dominated by

$$\sum_{j=1}^{|support(\alpha)|} x_j \geq 1 \tag{5.216}$$

which must be valid as well (or else there would exist $y \in \{0,1\}^n$ for which $\alpha^T y = 0 < \beta$). Suppose now that there is no A_i such that $A_i \subseteq support(\alpha)$, then define $y \in \{0,1\}^n$ by

$$y_j = \begin{cases} 0 : j \in support(\alpha) \\ 1 : \text{ otherwise} \end{cases} \tag{5.217}$$

Then for all A_i, $\sum_{j \in A_i} y_j \geq 1 \Rightarrow y \in P$, but $\sum_{j=1}^{|support(\alpha)|} y_j = 0$, which is a contradiction. □

Lemma 5.4 *If $P = \bigcap_{i=1}^m \bigcup_{j \in A_i} Y_j$, where $A_i \subseteq \{1, \ldots, n\}$, $i = 1, \ldots, m$, and $a^T x \geq \beta$, a unrestricted, is valid for P, then $a^T x \geq \beta$ is dominated by inequalities of the form $\alpha^T x \geq \beta$, $\alpha \geq 0$.*

Proof: Consider a valid inequality $a^T x \geq \beta$ with $a_h < 0$. This inequality is dominated by $x_h \leq 1$ and $\bar{a}^T x \geq \beta - a_h$, where $\bar{a}$ is the same as a but with $\bar{a}_h = 0$. Suppose that $\bar{a}^T x \geq \beta - a_h$ were not valid for P, then there would be a $y \in P : \bar{a}^T y < \beta - a_h$, but since y must satisfy $a^T y \geq \beta$ we must have $y_h = 0$. Define now $\bar{y}$ to be the same as y but with $\bar{y}_h = 1$. Changing a coordinate from zero to one cannot violate any of the constraints $\sum_{j \in A_i} y_j \geq 1$ that define P, so $\bar{y} \in P$ as well, and

$$a^T \bar{y} = \bar{a}^T y + a_h < \beta \tag{5.218}$$

which is a contradiction. Repeating for all $a_h < 0$ proves the lemma. □

Lemma 5.5 *If $P = \bigcap_{i=1}^m \bigcup_{j \in A_i} Y_j$, where $A_i \subseteq \{1, \ldots, n\}$, $i = 1, \ldots, m$, and $\alpha^T x \geq \beta$, $\alpha \geq 0$ is valid for P and of pitch n, then $\alpha^T x \geq \beta$ is dominated by the valid pitch 1 inequalities.*

Proof: By the definition of pitch, the only way for a pitch n inequality to be valid is if the pitch 1 inequality $y_i \geq 1$ is valid for all $i = 1, \ldots, n$. It is easy to see that these inequalities will dominate any valid pitch n inequality. □

Taken together we conclude:

Lemma 5.6 *If $P = \bigcap_{i=1}^m \bigcup_{j \in A_i} Y_j$, where $A_i \subseteq \{1, \ldots, n\}$, $i = 1, \ldots, m$, then every valid inequality for P is dominated by a constraint with pitch $\leq n - 1$.* □

Unfortunately Lemma 5.6 does not hold in general for P of the form,

$$P = \bigcap_{i=1}^{m} \bigcup_{j=1}^{t_i} M_{f(i,j)} \tag{5.219}$$

where f maps into $\{1', 1'', \ldots, n', n''\}$, and $M_{l'} = Y_l,\ M_{l''} = N_l,\ l = 1, \ldots, n$. Nevertheless any P of this form can be relaxed to a set of the form $\bigcap \bigcup Y_j$. Consider first that where we define, for any $y \in \{0,1\}^n$, $y_{l'} = y_l,\ y_{l''} = 1 - y_l,\ l = 1, \ldots, n$, then P can be written as

$$P = \{y \in \{0,1\}^n : \sum_{j=1}^{t_i} y_{f(i,j)} \geq 1,\ i = 1, \ldots, m\}. \tag{5.220}$$

For example,

$$P = (Y_1 \cup N_3) \cap (N_2 \cup N_1 \cup Y_3) \cap (Y_2 \cup N_1) \tag{5.221}$$

can be expressed as

$$P = \{y \in \{0,1\}^n :$$

$$y_1 + (1 - y_3) \geq 1,\ (1 - y_2) + (1 - y_1) + y_3 \geq 1,\ y_2 + (1 - y_1) \geq 1\} = \tag{5.222}$$

$$\{y \in \{0,1\}^n : y_{1'} + y_{3''} \geq 1,\ y_{2''} + y_{1''} + y_{3'} \geq 1,\ y_{2'} + y_{1''} \geq 1\}. \tag{5.223}$$

Thus by reindexing the variables according to the $i',\ i''$, a set P as in (5.220) can be equivalently represented as the set

$$P = \{y' = (y'_{1'}, y'_{1''}, \ldots, y'_{n'}, y'_{n''}) \in \{0,1\}^{2n} : \sum_{j=1}^{t_i} y'_{f(i,j)} \geq 1,\ i = 1, \ldots, m,$$

$$y'_{l'} + y'_{l''} = 1,\ l = 1, \ldots, n\} \tag{5.224}$$

In this representation, for each $l = 1, \ldots, n$, $y'_{l'}$ replaces y_l, and the new variable $y'_{l''}$ is introduced with value fixed to $1 - y_l$. In set theoretic notation,

$$P = \left(\bigcap_{i=1}^{m} \bigcup_{j=1}^{t_i} Y'_{f(i,j)} \right) \cap \bigcap_{l=1}^{n} ((Y'_{l'} \cup Y'_{l''}) \cap (N'_{l'} \cup N'_{l''})) \tag{5.225}$$

where

$$Y'_j = \{y' \in \{0,1\}^{2n} : y'_j = 1\} \text{ and } N'_j = \{y' \in \{0,1\}^{2n} : y'_j = 0\}. \tag{5.226}$$

But (5.224) can be relaxed to the set

$$P' = \{y' \in \{0,1\}^{2n} : \sum_{j=1}^{t_i} y'_{f(i,j)} \geq 1,\ i = 1, \ldots, m,\ \ y'_{l'} + y'_{l''} \geq 1,\ l = 1, \ldots, n\} \tag{5.227}$$

or, in set theoretic notation,

$$P' = \left(\bigcap_{i=1}^{m} \bigcup_{j=1}^{t_i} Y'_{f(i,j)} \right) \cap \bigcap_{l=1}^{n} (Y'_{l'} \cup Y'_{l''}) \tag{5.228}$$

which is of the desired form. One nice feature of this relaxation is as follows.

Lemma 5.7 *Let P be as in (5.224), and let P' be as in (5.227). Suppose $x' \in R^{2n}$ belongs to $Conv(P')$, then $x' \in Conv(P)$ as well iff $x'_{l'} + x'_{l''} = 1, \ \forall l = 1, \ldots, n$.*

Proof: If the condition is violated then obviously $x' \notin P$. Conversely, if $x' \in Conv(P)$ and $x'_{l'} + x'_{l''} = 1$ for all l, then we can write

$$x' = \sum_{y' \in P'} \lambda_{y'} y', \ \lambda \geq 0, \ \sum_{y' \in P'} \lambda_{y'} = 1. \tag{5.229}$$

Suppose now that $\lambda_{y'} > 0$ for some $y' \in P' - P$. Since $y' \in P' - P$ we must have $y'_{l'} + y'_{l''} > 1$ for some l, but since for all $y' \in P'$ we have $y'_{l'} + y'_{l''} \geq 1$, we must have

$$x'_{l'} + x'_{l''} = \sum_{y' \in P'} \lambda_{y'} (y'_{l'} + y'_{l''}) > \sum_{y' \in P'} \lambda_{y'} = 1 \tag{5.230}$$

which is a contradiction. □

Corollary 5.8 *Let P be as in (5.224), and let P' be as in (5.227). Every valid constraint for P (in the R^{2n} representation) is dominated by valid constraints for P' of the form $\alpha^T x' \geq \beta, \ \alpha \geq 0$, of pitch $\leq 2n - 1$, and the constraints $x'_{l'} + x'_{l''} = 1, \ l = 1, \ldots, n$.* □

5.2.5 Pitch 2 Inequalities

We will now show that even in the simplest (nontrivial) case, namely $P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} Y_j$, which as has been noted, corresponds to set covering problems, it is no simple matter to obtain even the valid pitch 2 inequalities. Note that the set of valid pitch 2 inequalities for set covering problems can be equivalently cast as the inequalities with all coefficients (including the right hand side) in $\{0, 1, 2\}$. (It is clear that any such inequality is of pitch 2 or less, and it is not hard to show that every valid pitch 2 inequality can be dominated by valid $0, 1, 2$ inequalities.) This class of inequalities has previously been analyzed by Balas and Ng ([BN89]), and they showed that these inequalities can be characterized by a certain type of rank 1 Chvátal-Gomory cut. But an explicit construction of all of these inequalities via their characterization would still require exponentially many such cuts.

Consider first that there may be exponentially many facet defining pitch 2 inequalities. For example, consider the following system.

Let $A \subset \{1, \ldots, n\}, \ |A| \geq 2$, let $A_i = A - \{i\}$ for each $i \in A$, and let $\{B_i : i \in A\}$ be $|A|$ disjoint subsets of $\{1, \ldots, n\}$, with $A \cap B_i = \emptyset$ for all i. Define $x(A_i) = \sum_{j \in A_i} x_j$. Let P be the set of $0, 1$ points that satisfy

$$x(A_i) + x_j \geq 1, \ \forall j \in B_i, \ \forall i \in A. \tag{5.231}$$

For every set $Q \subset \{1, \ldots, n\}, |Q| = |A|$, consisting of exactly one element drawn from each set $B_i, i \in A$, it is not hard to prove that the constraint

$$x(A) + x(Q) \geq 2 \tag{5.232}$$

is valid and facet defining for P. There is such a constraint for each of the exponentially many choices of Q, and these constraints are all of pitch 2.

We will now show that the N^{++} procedure, defined in Definition 4.29, (recall that this is a vastly more powerful operator than N^+), can also perform poorly in obtaining pitch 2 inequalities for set covering problems. For the purposes of the following theorem, given $P \subseteq \{0,1\}^n$, given $\bar{P} \subseteq [0,1]^n$ with $\bar{P} \cap \{0,1\}^n = P$, and recalling (Definition 1.2) that $\bar{K}(\bar{P})$ is the homogenized version of $\bar{P}$, the "N^{++}" rank of a valid inequality, $\alpha^T x \geq \beta$, for $P \subseteq \{0,1\}^n$ will refer to the smallest integer k such that all points of $(N^{++})^k(\bar{K}(\bar{P}))$ satisfy the homogenized inequality $\alpha^T x \geq \beta x_0$.

Theorem 5.9 *Define*

$$A_i = \{1, \ldots, n\} - \{i\}, \; i = 1, \ldots, n, \quad n \geq 3. \tag{5.233}$$

Let

$$P = \{y \in \{0,1\}^n : y(A_i) \geq 1, \; i = 1, \ldots, n\} \tag{5.234}$$

$$\bar{P} = \{y \in [0,1]^n : y(A_i) \geq 1, \; i = 1, \ldots, n\}. \tag{5.235}$$

The pitch 2 *inequality*

$$\sum_{j=1}^{n} y_j \geq 2 \tag{5.236}$$

is valid for P, and its N^{++} rank is $\geq n - 2$.

Proof: We will construct a measure χ on $\mathcal{A}$ for which the vector

$$(\chi[\{0,1\}^n], \chi[Y_i], \ldots, \chi[Y_n]) \tag{5.237}$$

violates the constraint

$$\sum_{j=1}^{n} \chi[Y_j] \geq 2\chi[\{0,1\}^n] \tag{5.238}$$

while having every partial sum χ^Q satisfy every constraint

$$\sum_{j \in A_i} \chi^Q[Y_j] \geq \chi^Q[\{0,1\}^n], \; i = 1, \ldots, n \tag{5.239}$$

for every Q of the form

$$Q = \bigcap_{h=1}^{k} M_h \tag{5.240}$$

where each $M_h \in \{Y_1, \ldots, Y_n, N_1, \ldots, N_n\}$, and $k < n-2$. By Remark 3.68 and Definition 4.29 it will then follow that for every $k < n-2$, the vector $(\chi[\{0,1\}^n], \chi[Y_i], \ldots, \chi[Y_n])$, which does not belong to $Cone(K(P))$ (Definition 1.2), nevertheless belongs to $(N^{++})^k(K(\bar{P}))$, which proves the theorem.

Before we proceed to the construction of the measure, recall that a partial sum χ^Q is the measure on $\mathcal{A}$ that matches the value of χ on every atom in Q, and assigns a measure of zero elsewhere, and recall also that the measure of any set is the sum of the measures of the atoms that are contained in that set. Recall also that a measure χ on $\mathcal{A}$ defines a measure on $\mathcal{P}$, in the sense that there is a measure $\bar{\chi}$ on $\mathcal{P}$ with $\bar{\chi}[q \cap P] = \chi[q]$ for all $q \in \mathcal{A}$, iff it assigns a measure of zero to all atoms in P^c. Recall finally that if a measure χ on $\mathcal{A}$ defines a measure on $\mathcal{P}$ in this sense, then $(\chi[\{0,1\}^n], \chi[Y_1], \ldots, \chi[Y_n]) \in Cone(K(P))$. The construction is as follows. For the atom

$$r = \bigcap_{j=1}^{n} N_j \tag{5.241}$$

assign $\chi[r] = 1$, and for each $J \subset \{1, \ldots, n\}$, $|J| = n-2$, for each atom

$$s^J = \bigcap_{j \in J} N_j \cap \bigcap_{j \notin J} Y_j \tag{5.242}$$

assign $\chi[s^J] = 1$. Assign all remaining atoms a measure of zero. Each s^J atom contributes 1 unit of measure to $\chi[\{0,1\}^n]$, and one unit of measure to each of the two $\chi[Y_j]$, $j \notin J$. Thus each s^J atom contributes two units of measure to each side of expression (5.238). But r contributes two units to the right side and nothing to the left, so χ indeed violates (5.238).

Consider now that for each set Q of the form (5.240) that entails a "yes" (i.e. some element M_h of the intersection (5.240) is of the form Y_j), then $r \not\subseteq Q$, so the measure χ^Q assigns zero measure to r, and nonzero measure only to (some of the) s^J atoms. Thus since all s^J atoms are in P, χ^Q defines a $\mathcal{P}$-measure, so $(\chi^Q[\{0,1\}^n], \chi^Q[Y_1], \ldots, \chi^Q[Y_n]) \in Cone(K(P))$ and χ^Q therefore certainly satisfies all constraints (5.239). So suppose that Q entails only "no's", i.e. it is of the form

$$Q = \bigcap_{j \in q} N_j \tag{5.243}$$

where $q \subset \{1, \ldots, n\}$ and suppose that $|q| < n-2$. Then for any $i \in \{1, \ldots, n\}$, r contributes one unit of measure to the right side of (5.239) and zero to the left, as it belongs to no set Y_j. Each $s^J \subset Q$ atom contributes one unit to the right side and at least one unit to the left (as each J^c overlaps each A_i in at least one location), and each $s^J \subset Q$ for which $|J^c \cap A_i| = 2$ (i.e. the two "yeses" of s^J both overlap A_i), contributes 2 units to the left side. Thus if we can establish that there is some $s^J \subset Q$ for which indeed $|J^c \cap A_i| = 2$, then we will be guaranteed that (5.239) will be satisfied. Observe now that for any $|q| < n-2$, $|A_i - q| \geq 2$, so where S is any size 2 subset of $A_i - q$, and we define

$$J(i) = \{1, \ldots, n\} - S \tag{5.244}$$

then $s^{J(i)} \subset Q$, and $J(i)^c \cap A_i = 2$ (i.e. the indices of the two "yeses" of $s^{J(i)}$ both belong to A_i, but neither belongs to q), so all constraints of the form (5.239) will be satisfied. □

This is particularly interesting considering that it is easy to see that where P is as in Theorem 5.9, the "Common Factor Algorithm" at level 2, to be defined in the next chapter, is dominated by $(\bar{N})^{n-2}$, which is itself dominated by $(N^{++})^{n-2}$. Thus since the common factor algorithm, as we will see, obtains all pitch 2 constraints by level 2, it follows from Theorem 5.9 that the $\bar{N}$ rank, as well as the N^{++} rank, of the constraint (5.236) is exactly $n-2$. Thus the measure consistency requirement that distinguishes N^{++} from $\bar{N}$ did not help in this case, and the N^{++} algorithm did not guarantee (5.236) until the "last minute", i.e. until it dominated the common factor algorithm.

It is also worth pointing out that in the last stage of the proof of Theorem 5.9, if $|q| = n-3$ then there are A_i for which $|A_i - q| = 2$, so that there is exactly one choice of a pair of indices in $A_i - q$, and there is exactly one $J(i)$ for which $s^{J(i)} \subset Q$ and such that $J(i)^c$ overlaps A_i twice. But if $|q| = n-4$, then $|A_i - q| \geq 3$, so there are at least $\binom{3}{2} = 3$ appropriate size 2 sets S and there are therefore at least 3 sets J for which $s^J \subset Q$ and such that J^c overlaps A_i twice. Thus even had we assigned a measure of 3 to the atom r, all of the constraints (5.239) would continue to be satisfied by every partial sum χ^Q for which Q is an intersection of no more than $n-4$ sets M_i. In general, if $|q| = k$, $(k \leq n-3)$, then there would be at least $\binom{n-k-1}{2}$ sets J for which $s^J \subset Q$ and such that J^c overlaps A_i twice, and therefore even had we assigned a measure of $\binom{n-k-1}{2}$ to the atom r, all of the constraints (5.239) would continue to be satisfied by every partial sum χ^Q for which Q is an intersection of no more than k sets M_i. Denoting this measure with $\chi[r] = \binom{n-k-1}{2}$ as $\hat{\chi}$, this means that the vector $(\hat{\chi}[\{0,1\}^n], \hat{\chi}[Y_1], \ldots, \hat{\chi}[Y_n])$ belongs to $(N^{++})^k(\bar{K}(\bar{P}))$. Observe now that considering that there are $\binom{n}{2}$ atoms s^J in total, each of which belongs to exactly two sets Y_i, assigning a measure of $\binom{n-k-1}{2}$ to the atom r would imply that

$$\hat{\chi}[\{0,1\}^n] = \binom{n}{2} + \binom{n-k-1}{2} = \frac{n(n-1) + (n-k-1)(n-k-2)}{2}, \tag{5.245}$$

so that

$$\sum_{j=1}^{n} \hat{\chi}[Y_j] = 2\binom{n}{2} = \frac{2n(n-1)}{n(n-1)+(n-k-1)(n-k-2)}\hat{\chi}[\{0,1\}^n]. \tag{5.246}$$

Thus, for $(N^{++})^k(\bar{K}(\bar{P}))$ to satisfy even the constraint, say,

$$\sum_{j=1}^{n} x_j \geq 1.8x_0 \tag{5.247}$$

requires the level k to be such that

$$\frac{2n(n-1)}{n(n-1)+(n-k-1)(n-k-2)} \geq 1.8. \tag{5.248}$$

Observe, however, that there is no fixed k for which (5.248) will hold for all n. Thus where P and $\bar{P}$ are as in Theorem 5.9, and we define

$$(N^{++})^k(\bar{P}) := (N^{++})^k(\bar{K}(\bar{P})) \cap \{x \in R^{n+1} : x_0 = 1\}, \tag{5.249}$$

and we write

$$c^* := \min\left\{\sum_{i=1}^{n} x_i \ : \ x \in P\right\} \tag{5.250}$$

and

$$c^k = \min\left\{\sum_{i=1}^{n} x_i \ : \ x \in (N^{++})^k(\bar{P})\right\} \tag{5.251}$$

then considering that $c^* = 2$, it follows that there is no fixed k for which we can always be guaranteed that $c^k \geq .9c^*$. The choice, moreover, of 1.8 for the right hand side of (5.247) was arbitrary. Any number greater than 1 would have yielded the same result. Note now that the inequality $\sum_{i=1}^{n} x_i \geq 2$ is a rank 1 Chvátal-Gomory cut on $\bar{P}$. We therefore conclude as follows.

Theorem 5.10 *Given a set covering problem, denoted (SC),*

$$\min\left\{c^T x \ : \ x(A_i) \geq 1, \ i = 1, \ldots, m, \ x \in \{0,1\}^n\right\} \tag{5.252}$$

with feasible region denoted P, define

$$\bar{P} = x \in [0,1]^n : x(A_i) \geq 1, \ i = 1, \ldots, m; \tag{5.253}$$

define $\bar{P}^{C-G}$ to be the rank 1 Chvátal-Gomory closure of $\bar{P}$, and define

$$c^k(SC) = \min\left\{c^T x \ : \ x \in (N^{++})^k(\bar{P})\right\}, \quad \textit{and} \quad \bar{c}(SC) = \min\left\{c^T x \ : \ x \in \bar{P}^{C-G}\right\}. \tag{5.254}$$

Then for every number $\epsilon < \frac{1}{2}$, there is no fixed integer k for which

$$c^k(SC) \geq (1-\epsilon)\bar{c}(SC) \tag{5.255}$$

for all set covering problems (SC). □

This is particularly noteworthy considering the following fact. It can be shown ([BZ03]) that for any fixed positive integer r and any fixed number $\epsilon > 0$, there exists a fixed integer k such that any set covering problem (SC) of the form (5.252), with feasible region denoted P, for which we define

$$P^k = \{x \in [0,1]^n : \alpha^T x \geq \beta, \text{ for all valid constraints on } P \text{ with } \pi(\alpha, \beta) \leq k\} \quad (5.256)$$

and $\bar{P}^{C-G(r)}$ to be the rank r Chvátal-Gomory closure of $\bar{P}$, and for which we define $\hat{c}^k(SC) := \min\{c^T x \ : \ x \in P^k\}$, and $\vec{c}^r(SC) := \min\{c^T x \ : \ x \in \bar{P}^{C-G(r)}\}$, satisfies

$$\hat{c}^k(SC) \geq (1 - \epsilon)\vec{c}^r(SC). \quad (5.257)$$

This will imply that for any set covering problem (SC) of the form (5.252), the "Depth-First Algorithm" to be introduced later in this chapter, as well as the "Common Factor Algorithms" of the following chapter, all of which are capable of generating in polynomial time a relaxation of $Conv(P)$ that satisfies all valid pitch $\leq k$ constraints on P, can always ϵ-approximate $\vec{c}^r(SC)$ in polynomial time for any fixed ϵ and r.

5.3 Preliminaries

Sets of the form (5.12) can be complicated objects, so before we can describe the algorithms we will need to understand the basics of the structure of such sets. The first result that we will show here is that sets of the form (5.12) have two other representations, each of which will prove useful to us in due course. For example, consider once more the case

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)). \quad (5.258)$$

As we saw earlier, this can be represented as

$$P = \bigcap_{i_1=1}^{1} \bigcup_{j_1=1}^{2} \bigcap_{i_2=1}^{2} \bigcup_{j_2=1}^{2} Y_{f(i_1,j_1,i_2,j_2)} \quad (5.259)$$

where f maps

$$(1,1,1,1) \to 1, \ (1,1,1,2) \to 2, \ (1,1,2,1) \to 3, \ldots, (1,2,2,2) \to 8. \quad (5.260)$$

Expanding (5.258) by distributing the intersections over the unions, i.e. via the rule

$$(A \cup B) \cap (C \cup D) = (A \cap C) \cup (A \cap D) \cup (B \cap C) \cup (B \cap D) \quad (5.261)$$

yields

$$P = (Y_1 \cap Y_3) \cup (Y_1 \cap Y_4) \cup (Y_2 \cap Y_3) \cup (Y_2 \cap Y_4) \cup$$

$$(Y_5 \cap Y_7) \cup (Y_5 \cap Y_8) \cup (Y_6 \cap Y_7) \cup (Y_6 \cap Y_8) = \tag{5.262}$$

$$\begin{aligned}
&(Y_{f(1,1,1,1)} \cap Y_{f(1,1,2,1)}) \ \cup \ (Y_{f(1,1,1,1)} \cap Y_{f(1,1,2,2)}) \cup \\
&(Y_{f(1,1,1,2)} \cap Y_{f(1,1,2,1)}) \ \cup \ (Y_{f(1,1,1,2)} \cap Y_{f(1,1,2,2)}) \cup \\
&(Y_{f(1,2,1,1)} \cap Y_{f(1,2,2,1)}) \ \cup \ (Y_{f(1,2,1,1)} \cap Y_{f(1,2,2,2)}) \cup \\
&(Y_{f(1,2,1,2)} \cap Y_{f(1,2,2,1)}) \ \cup \ (Y_{f(1,2,1,2)} \cap Y_{f(1,2,2,2)}).
\end{aligned} \tag{5.263}$$

There is a pattern in (5.263). Each element in the union is of the form

$$Y_{f(1,j_1(1),1,j_2(1,1))} \cap Y_{f(1,j_1(1),2,j_2(1,2))} \tag{5.264}$$

for some pair of functions $(j_1(1), j_2(\cdot))$ such that $j_1(1)$ is a function whose domain is the single element $\{1\}$ (which is the only possible value for the first subscript i_1), and thus has a constant value in the range $\{1, 2\}$, and in which $j_2(\cdot)$ is a function of (i_1, i_2) (i.e. the first and third subscripts) with range $\{1, 2\}$. There are obviously two possible choices for $j_1(1)$, and there are four choices for $j_2(1, i_2)$, namely

1. $(1, 1) \to 1, \ (1, 2) \to 1$
2. $(1, 1) \to 1, \ (1, 2) \to 2$
3. $(1, 1) \to 2, \ (1, 2) \to 1$
4. $(1, 1) \to 2, \ (1, 2) \to 2.$

There are thus eight possible pairs of functions, and there are eight corresponding elements in the union. In the first element of the union, for example, j_1 and j_2 both have constant value 1. In the second element of the union j_1 takes the value 1, and j_2 take the value 1 when $i_2 = 1$ and the value 2 where $i_2 = 2$. In the third, j_1 takes the value 1, and j_2 takes the value 2 when $i_2 = 1$, etcetera. Note also that each element (5.264) of the union is the intersection of the sets $Y_{f(i_1,j_1(i_1),i_2,j_2(i_1,i_2))}$ over all possible choices of i_1, i_2 in the domain of the pair $(j_1(i_1), j_2(i_1, i_2))$.

Observe moreover that for a point $y \in \{0, 1\}^n$ to belong to P, it must belong to either Y_1 or Y_2 or Y_5 or Y_6. Similarly it must belong to either Y_1 or Y_2 or Y_7 or Y_8, as well as to Y_3 or Y_4 or Y_5 or Y_6, and to Y_3 or Y_4 or Y_7 or Y_8. Conversely, if y belongs to Y_1 or Y_2 or Y_5 or Y_6, and to Y_1 or Y_2 or Y_7 or Y_8, and to Y_3 or Y_4 or Y_5 or Y_6, and to Y_3 or Y_4 or Y_7 or Y_8, then it must belong to P. Thus we can restate P as

$$P = (Y_{f(1,1,1,1)} \cup Y_{f(1,1,1,2)} \cup Y_{f(1,2,1,1)} \cup Y_{f(1,2,1,2)}) \cap$$

$$(Y_{f(1,1,1,1)} \cup Y_{f(1,1,1,2)} \cup Y_{f(1,2,2,1)} \cup Y_{f(1,2,2,2)}) \cap$$
$$(Y_{f(1,1,2,1)} \cup Y_{f(1,1,2,2)} \cup Y_{f(1,2,1,1)} \cup Y_{f(1,2,1,2)}) \cap$$
$$(Y_{f(1,1,2,1)} \cup Y_{f(1,1,2,2)} \cup Y_{f(1,2,2,1)} \cup Y_{f(1,2,2,2)}) \tag{5.265}$$

Again there is a pattern here. Each element of the intersection corresponds to a different function describing how to choose i_2 for each j_1. In the first element, $i_2 = 1$ regardless of j_1. In the second $i_2 = 1$ when $j_1 = 1$, and $i_2 = 2$ when $j_1 = 2$. In the third $i_2 = 2$ when $j_1 = 1$, and $i_2 = 1$ when $j_1 = 2$, and in the fourth $i_2 = 2$ regardless of j_1. In parallel to the representation (5.263), each element of the intersection (5.265) is the union over all possible choices of j_1, j_2 for the given rule. Before we formalize and generalize these alternate representations for P, we will first pose a definition that makes the characterization of these "indexing" functions precise.

Definition 5.11 *Let*

$$P = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} \bigcap_{i_2=1}^{m_2(i_1,j_1)} \bigcup_{j_2=1}^{t_2(i_1,j_1,i_2)} \cdots \bigcap_{i_h=1}^{m_h(i_1,\ldots,j_{h-1})} \bigcup_{j_h=1}^{t_h(i_1,\ldots,j_{h-1},i_h)} M_{f(i_1,j_1,\cdots,i_h,j_h)} \tag{5.266}$$

where f maps into the set $\{1', 2', \ldots, n', 1'', 2'', \ldots, n''\}$ and where when $l \in \{1, 2, \ldots, n\}$, then $M_{l'} = Y_l$ and $M_{l''} = N_l$. Given h integer valued functions of integers,

$$I_1,\ I_2(j_1),\ I_3(j_1, j_2), \ldots, I_h(j_1, \ldots, j_{h-1}), \tag{5.267}$$

we will say that these h functions comprise an ordered indexing family $\mathcal{I}$ *for P if these functions are as follows:*

I_1 is a constant $\in \{1, \ldots, m_1\}$.

$I_2(j_1)$ has domain $\{1, \ldots, t_1(I_1)\}$ and each $I_2(j_1) \in \{1, \ldots, m_2(I_1, j_1)\}$.

$I_3(j_1, j_2)$ has domain

$$\{(j_1, j_2) : j_1 \in \{1, \ldots, t_1(I_1)\},\ j_2 \in \{1, \ldots, t_2(I_1, j_1, I_2(j_1))\}\} \tag{5.268}$$

and each $I_3(j_1, j_2) \in \{1, \ldots, m_3(I_1, j_1, I_2(j_1), j_2)\}$.
In general, the elements of the domain of $I_l(j_1, j_2, \ldots, j_{l-1})$ are the tuples

$$\{(j_1, \ldots, j_{l-1}) : j_k \in \{1, \ldots, t_k(I_1, j_1, \ldots, I_{k-1}(\cdot), j_{k-1}, I_k(\cdot))\},\ k = 1, \ldots, l-1\} \tag{5.269}$$

and each

$$I_l(j_1, j_2, \ldots, j_{l-1}) \in \{1, \ldots, m_l(I_1, j_1, I_2(j_1), j_2, \ldots, I_{l-1}(\cdot), j_{l-1})\}. \tag{5.270}$$

Similarly we will say that h functions

$$J_1(i_1), J_2(i_1, i_2), \ldots, J_h(i_1, \ldots, i_h) \tag{5.271}$$

*comprise a j-*indexing family *for P if these functions are as follows:*

For each $l = 1, \ldots, h$, the elements of the domain of $J_l(i_1, i_2, \ldots, i_l)$ are the tuples

$$\{(i_1, i_2, \ldots, i_l) : i_k \in \{1, \ldots, m_k(i_1, J_1(i_1), \ldots, i_{k-1}, J_{k-1}(\cdot)\},\ k = 1, \ldots, l\} \tag{5.272}$$

and each

$$J_l(i_1, i_2, \ldots, i_l) \in \{1, \ldots, t_l(i_1, J_1(i_1), \ldots, i_{l-1}, J_{l-1}(\cdot), i_l)\}. \tag{5.273}$$

Given an indexing family $\mathcal{I}$, we will write

$$\mathcal{I}(j_1, \ldots, j_h) = (I_1, j_1, I_2(j_1), j_2, \ldots, I_h(j_1, \ldots, j_{h-1}), j_h) \tag{5.274}$$

and we will refer to the domain of the function $\mathcal{I}(j_1, \ldots, j_h)$, i.e. the h-tuples $(j_1, \ldots, j_h)$ for which the numbers $(I_1, j_1, \ldots, I_h, j_h)$ are defined and within the appropriate bounds, as $j(\mathcal{I})$. The function $\mathcal{J}(i_1, \ldots, i_h)$ and the set $i(\mathcal{J})$ are defined similarly.

Observe that I_2 is not technically a function of I_1, but I_2 can only serve as the second member of an indexing family for P if there is an appropriate I_1. In general, I_l is not technically a function of $I_1, \ldots, I_{l-1}$, though for I_l to be the l'th member of an indexing family for P requires that appropriate $I_1, \ldots, I_{l-1}$ exist. Note also that to reduce clutter, we will often refer to the functions $I_l(\cdot)$ and $J_l(\cdot)$ merely as I_l and J_l.

Lemma 5.12 *For any set of the form*

$$W = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} \cdots \bigcap_{i_h=1}^{m_h(i_1,\ldots,j_{h-1})} \bigcup_{j_h=1}^{t_h(i_1,\ldots,j_{h-1},i_h)} Z(i_1, j_1, \cdots, i_h, j_h) \tag{5.275}$$

we have

$$W = W^I := \bigcap_{\mathcal{I}} \bigcup_{j(\mathcal{I})} Z(I_1, j_1, \ldots, I_h, j_h) \tag{5.276}$$

where the intersection is taken over all indexing families of functions $\mathcal{I}$ for W, and

$$W = W^J := \bigcup_{\mathcal{J}} \bigcap_{i(\mathcal{J})} Z(i_1, J_1, \ldots, i_h, J_h) \tag{5.277}$$

where the union is taken over all j-indexing families of functions $\mathcal{J}$ for W. Moreover, if we expand the expression (5.275) defining W by repeatedly distributing the intersections over the unions then this will also yield the expression W^J.

Proof: The proof will be by induction on h. Let us first prove $W = W^I$. Where $h = 1$ this equality holds by definition. Assume now that $W = W^I$ for all $h \leq r$, and consider $h = r+1$. Suppose $y \in W$ but $y \notin W^I$, so for some indexing family of functions $I_1, \ldots, I_{r+1}$,

$$y \notin \bigcup_{j_1=1}^{t_1(I_1)} \bigcup_{j_2=1}^{t_2(I_1,j_1,I_2)} \cdots \bigcup_{j_{r+1}=1}^{t_{r+1}(I_1,j_1,\ldots,I_{r+1})} Z(I_1, j_1, I_2, j_2 \cdots, I_{r+1}, j_{r+1}). \tag{5.278}$$

Then for every $j_1 = 1, \ldots, t_1(I_1)$ we also have

$$y \notin \bigcup_{j_2=1}^{t_2(I_1,j_1,I_2)} \cdots \bigcup_{j_{r+1}=1}^{t_{r+1}(I_1,j_1,\ldots,I_{r+1})} Z(I_1, j_1, I_2, j_2 \cdots, I_{r+1}, j_{r+1}). \tag{5.279}$$

Note that I_1, j_1 and therefore I_2 are all constant in this expression. But since $y \in W$, for any value of $I_1 \in \{1, \ldots, m_1\}$, there must be some $j_1 \in \{1, \ldots, t(I_1)\}$ such that

$$y \in W_{I_1,j_1} := \bigcap_{i_2=1}^{m_2(I_1,j_1)} \bigcup_{j_2=1}^{t_2(I_1,j_1,i_2)} \cdots \bigcap_{i_{r+1}=1}^{m_{r+1}(I_1,j_1,i_2,\ldots,j_r)} \bigcup_{j_r=1}^{t_r(I_1,j_1,i_2,\ldots,j_r,i_{r+1})} Z(I_1, j_1, i_2, \cdots, j_{r+1}). \tag{5.280}$$

Thus by induction for every indexing family of functions for W_{I_1,j_1},

$$\{\bar{I}_2, \bar{I}_3(j_2), \ldots, \bar{I}_{r+1}(j_2, j_3, \ldots, j_r)\} \tag{5.281}$$

we must have

$$y \in \bigcup_{j_2=1}^{t_2(I_1,j_1,\bar{I}_2)} \cdots \bigcup_{j_{r+1}=1}^{t_{r+1}(I_1,j_1,\bar{I}_2,j_2,\ldots,\bar{I}_{r+1})} Z(I_1, j_1, \bar{I}_2, j_2 \cdots, \bar{I}_{r+1}, j_{r+1}). \tag{5.282}$$

But where I_1, j_1 and therefore I_2 are all constant then $\{I_2, \ldots, I_{r+1}\}$ is itself an indexing family of the form $\{\bar{I}_2, \bar{I}_3(j_2), \ldots, \bar{I}_{r+1}(j_2, j_3, \ldots, j_r)\}$ for Q_{I_1,j_1}, and we therefore obtain a contradiction.

Suppose now that $y \in W^I$, but that $y \notin W$. Since $y \notin W$, for some $i'_1 \in \{1, \ldots, m_1\}$, we must have $y \notin W_{i'_1,j_1}$ for any $j_1 \in 1, \ldots, t_1(i'_1)$. Thus by induction, for each $j_1 \in \{1, \ldots, t_1(i'_1)\}$ there must be some indexing family of functions for $W_{i'_1,j_1}$

$$\{\bar{I}_2^{j_1}, \bar{I}_3^{j_1}(j_2), \ldots, \bar{I}_{r+1}^{j_1}(j_2, j_3, \ldots, j_r)\} \tag{5.283}$$

for which

$$y \notin \bigcup_{j_2=1}^{t_2(i'_1,j_1,\bar{I}_2^{j_1})} \cdots \bigcup_{j_{r+1}=1}^{t_{r+1}(i'_1,j_1,\bar{I}_2^{j_1},j_2,\ldots,\bar{I}_{r+1}^{j_1})} Z(i'_1, j_1, \bar{I}_2^{j_1}, j_2 \cdots, \bar{I}_{r+1}^{j_1}, j_{r+1}) \Rightarrow \tag{5.284}$$

$$y \notin \bigcup_{j_1=1}^{t_1(i'_1)} \bigcup_{j_2=1}^{t_2(i'_1,j_1,\bar{I}_2^{j_1})} \cdots \bigcup_{j_{r+1}=1}^{t_{r+1}(i'_1,j_1,\bar{I}_2^{j_1},j_2,\ldots,\bar{I}_{r+1}^{j_1})} Z(i'_1, j_1, \bar{I}_2^{j_1}, j_2 \cdots, \bar{I}_{r+1}^{j_1}, j_{r+1}). \tag{5.285}$$

Define now the following indexing family of functions for W

$$I_1 = i'_1, \quad I_r(j_1, j_2, \ldots, j_{r-1}) = \bar{I}_r^{j_1}(j_2, \ldots, j_{r-1}), \quad r = 2, \ldots, r+1. \tag{5.286}$$

We therefore have

$$y \in \bigcup_{j_1=1}^{t_1(i'_1)} \bigcup_{j_2=1}^{t_2(i'_1,j_1,I_2)} \cdots \bigcup_{j_{r+1}=1}^{t_{r+1}(i'_1,j_1,\ldots,I_{r+1})} Z(i'_1, j_1, I_2, j_2 \cdots, I_{r+1}, j_{r+1}) = \tag{5.287}$$

$$\bigcup_{j_1=1}^{t_1(i'_1)} \bigcup_{j_2=1}^{t_2(i'_1,j_1,\bar{I}_2^{j_1})} \cdots \bigcup_{j_{r+1}=1}^{t_{r+1}(i'_1,j_1,\bar{I}_2^{j_1},j_2,\ldots,\bar{I}_{r+1}^{j_1})} Z(i'_1, j_1, \bar{I}_2^{j_1}, j_2 \cdots, \bar{I}_{r+1}^{j_1}, j_{r+1}) \tag{5.288}$$

which is a contradiction.

We will now show that the expression W can be expanded via distibution of the intersections into the expression W^J. Where $h = 1$, the general term of the union obtained by distributing the intersections over the unions in the expression

$$\bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} Z(i_1, j_1) \tag{5.289}$$

is an intersection with one element drawn from each union $\bigcup_{j_1=1}^{t_1(i_1)} Z(i_1, j_1),\ i_1 = 1, \ldots, m_1$, and there is such a term for each set of choices of elements to draw. So for each $i_1 = 1, \ldots, m_1$, a choice is made from among the $t_1(i_1)$ elements in the i_1'th union. Thus there is a term for each function of i_1 mapping i_1 into $\{1, \ldots, t_1(i_1)\}$, and this term is the intersection of these choices over all i_1. We therefore conclude that indeed

$$\bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} Z(i_1, j_1) = \bigcup_{\mathcal{J}^1} \bigcap_{i_1=1}^{m_1} Z(i_1, J_1) = \bigcup_{\mathcal{J}^1} \bigcap_{i(\mathcal{J}^1)} Z(i_1, J_1). \tag{5.290}$$

Assume now that the lemma holds for all $h \leq r$ and consider the case where $h = r+1$. Then by induction,

$$W = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} W_{i_1,j_1} = \tag{5.291}$$

$$\bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} \left(\bigcup_{\mathcal{J}^r(i_1,j_1)} \bigcap_{i(\mathcal{J}^r(i_1,j_1))} Z(i_1, j_1, i_2, J_2^{i_1,j_1}, \ldots, i_{r+1}, J_{r+1}^{i_1,j_1}) \right) \tag{5.292}$$

where each $\mathcal{J}^r(i_1, j_1)$ is a j-indexing family of r functions

$$\{J_2^{i_1,j_1}(i_2), \ldots, J_{r+1}^{i_1,j_1}(i_2, \ldots, i_{r+1})\} \tag{5.293}$$

for W_{i_1,j_1}. Thus

$$W = \bigcap_{i_1=1}^{m_1} \bigcup_{\substack{\mathcal{J}^r(i_1,j_1), \\ j_1=1,\ldots,t_1(i_1)}} \left(\bigcap_{i(\mathcal{J}^r(i_1,j_1))} Z(i_1, j_1, i_2, J_2^{i_1,j_1}, \ldots, i_{r+1}, J_{r+1}^{i_1,j_1}) \right). \tag{5.294}$$

If we distribute the left-most $\bigcap$ over the $\bigcup$, then as above, we obtain a union whose general term is an intersection over $i_1 = 1, \ldots, m_1$ of elements

$$\bigcap_{i(\mathcal{J}^r(i_1,j_1))} Z(\cdot) = \bigcap_{i_2=1}^{m_2(i_1,j_1)} \bigcap_{i_3=1}^{m_3(i_1,j_1,i_2,J_2^{i_1,j_1})} \cdots \bigcap_{i_{r+1}=1}^{m_{r+1}(\cdot)} Z(i_1, j_1, i_2, J_2^{i_1,j_1}, \ldots, i_{r+1}, J_{r+1}^{i_1,j_1}) \tag{5.295}$$

where for each $i_1 = 1, \ldots, m_1$, the element (5.295) is determined by some choice of a pair $j_1, \mathcal{J}^r(i_1, j_1)$. Thus we obtain a union whose general term is of the form

$$\bigcap_{i_1=1}^{m_1} \bigcap_{i_2=1}^{m_2(i_1,J_1(i_1))} \bigcap_{i_3=1}^{m_3(i_1,J_1(i_1),i_2,J_2^{i_1,J_1(i_1)})} \cdots \bigcap_{i_{r+1}=1}^{m_{r+1}(\cdot)} Z(i_1, J_1(i_1), i_2, J_2^{i_1,J_1(i_1)}, \ldots, i_{r+1}, J_{r+1}^{i_1,J_1(i_1)}) \tag{5.296}$$

where J_1 is a function of i_1, and the j-indexing family of functions for $W_{i_1,J(i_1)}$,

$$\{J_2^{i_1,J_1(i_1)}(i_2), \ldots, J_{r+1}^{i_1,J_1(i_1)}(i_2, \ldots, i_{r+1})\} \tag{5.297}$$

is also a function of i_1. The union moreover contains a term (5.296) for each selection of functions

$$J_1(i_1), J_2^{i_1,J_1(i_1)}(i_2), \ldots, J_{r+1}^{i_1,J_1(i_1)}(i_2, \ldots, i_{r+1}). \tag{5.298}$$

Now observe that each selection of functions (5.298) is itself a j-indexing family of functions $\mathcal{J} = \{J_1, \ldots, J_{r+1}\}$ for W and vice-versa. Thus

$$W = \bigcap_{i_1=1}^{m_1} \bigcup_{\substack{\mathcal{J}^r(i_1,j_1), \\ j_1=1,\ldots,t_1(i_1)}} \left(\bigcap_{i(\mathcal{J}^r(i_1,j_1))} Z(i_1, j_1, i_2, J_2^{i_1,j_1}, \ldots, i_{r+1}, J_{r+1}^{i_1,j_1}) \right) = \tag{5.299}$$

$$\bigcup_{\mathcal{J}} \left(\bigcap_{i_1=1}^{m_1} \bigcap_{i_2=1}^{m_2(i_1,J_1)} \bigcap_{i_3=1}^{m_3(i_1,J_1,i_2,J_2)} \cdots \bigcap_{i_{r+1}=1}^{m_{r+1}(\cdot)} Z(i_1, J_1, i_2, J_2, \ldots, i_{r+1}, J_{r+1}) \right) = \tag{5.300}$$

$$\bigcup_{\mathcal{J}} \bigcap_{i(\mathcal{J})} Z(\cdot). \ \Box \tag{5.301}$$

The following definitions formalize and generalize the notation introduced in Subsections 5.2.2 and 5.2.3.

Definition 5.13 *Given*

$$P = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} \bigcap_{i_2=1}^{m_2(i_1,j_1)} \bigcup_{j_2=1}^{t_2(i_1,j_1,i_2)} \cdots \bigcap_{i_h=1}^{m_h(i_1,\ldots,j_{h-1})} \bigcup_{j_h=1}^{t_h(i_1,\ldots,j_{h-1},i_h)} M_{f(i_1,j_1,\cdots,i_h,j_h)} \tag{5.302}$$

where f maps into the set $\{1',2',\ldots,n',1'',2'',\ldots,n''\}$ and where when $l \in \{1,2,\ldots,n\}$, then $M_{l'} = Y_l$ and $M_{l''} = N_l$, we will write

$$P = \bigcap_{i_1=1}^{m_1} R_{i_1} \tag{5.303}$$

$$R_{i_1} = \bigcup_{j_1=1}^{t_1(i_1)} Q_{i_1,j_1} \tag{5.304}$$

$$Q_{i_1,j_1} = \bigcap_{i_2=1}^{m_2(i_1,j_1)} R_{i_1,j_1,i_2} \tag{5.305}$$

$$R_{i_1,j_1,i_2} = \bigcup_{j_2=1}^{t_2(i_1,j_1,i_2)} Q_{i_1,j_1,i_2,j_2} \tag{5.306}$$

$$\vdots \tag{5.307}$$

$$Q_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1}} = \bigcap_{i_h=1}^{m_h(i_1,\ldots,j_{h-1})} R_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h} \tag{5.308}$$

$$R_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h} = \bigcup_{j_h=1}^{t_h(i_1,\ldots,j_{h-1},i_h)} Q_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h,j_h} \tag{5.309}$$

$$Q_{i_1,j_1,i_2,j_2,\ldots,i_{h-1},j_{h-1},i_h,j_h} = M_{f(i_1,j_1,\ldots,i_h,j_h)}. \tag{5.310}$$

As we noted in the previous section, unions and intersections of sets of the form Y_i are much easier to deal with than sets defined by unions and intersections of sets of the form M_i (where M_i may be either of the form Y_i or of the form N_i). The following definition suggests a relaxation of the general P, as defined in Definition 5.13, to a set that can be defined by unions and intersections of "yes" sets exclusively.

Definition 5.14 *Given P as in Definition 5.13, define the set P' by*

$$P' = \left(\bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} \bigcap_{i_2=1}^{m_2(i_1,j_1)} \bigcup_{j_2=1}^{t_2(i_1,j_1,i_2)} \cdots \bigcap_{i_h=1}^{m_h(i_1,\ldots,j_{h-1})} \bigcup_{j_h=1}^{t_h(i_1,\ldots,j_{h-1},i_h)} Y'_{f(i_1,j_1,\cdots,i_h,j_h)} \right) \cap \bigcap_{l=1}^{n} (Y'_{l'} \cup Y'_{l''}) \tag{5.311}$$

where for each $j \in \{1',1'',\ldots,n',n''\}$,

$$Y'_j = \{y' = (y'_{1'}, y'_{1''}, \ldots, y'_{n'}, y'_{n''}) \in \{0,1\}^{2n} : y'_j = 1\}. \tag{5.312}$$

For example, given $n = 3$ and

$$P = (Y_1 \cup N_3) \cap (N_2 \cup N_1 \cup Y_3) \cap (Y_2 \cup N_1) \tag{5.313}$$

which can be equivalently recast as

$$P = (M_{1'} \cup M_{3''}) \cap (M_{2''} \cup M_{1''} \cup M_{3'}) \cap (M_{2'} \cup M_{1''}) \tag{5.314}$$

we would have

$$P' = (Y'_{1'} \cup Y'_{3''}) \cap (Y'_{2''} \cup Y'_{1''} \cup Y'_{3'}) \cap (Y'_{2'} \cup Y'_{1''}) \cap \bigcap_{l=1}^{3} (Y'_{l'} \cup Y'_{l''}) \tag{5.315}$$

where, for example,

$$Y'_{2''} = \{(y'_{1'}, y'_{1''}, y'_{2'}, y'_{2''}, y'_{3'}, y'_{3''}) \in R^6 : y'_{2''} = 1\}. \tag{5.316}$$

Definition 5.15 *Where P is as in Definition 5.13, define*

$$T(i_1, j_1, \ldots, i_l, j_l) = \tag{5.317}$$

$$\bigcap_{\substack{\bar{i}_l=1 \\ \bar{i}_l \neq i_l}}^{m_l(i_1,\ldots,j_{l-1})} R_{i_1,j_1,\ldots,i_{l-1},j_{l-1},\bar{i}_l} \cap \bigcap_{\bar{j}_l=1}^{j_l-1} (Q_{i_1,j_1,\ldots,i_{l-1},j_{l-1},i_l,\bar{j}_l})^c \tag{5.318}$$

and

$$T(\{i_1, j_1, \ldots, i_l, j_l\}) = T(i_1, j_1) \cap \cdots \cap T(i_1, j_1, \ldots, i_l, j_l) \cap Q_{i_1,j_1,\ldots,i_l,j_l} \tag{5.319}$$

and

$$T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) = \tag{5.320}$$

$$\bigcap_{r=1}^{s} T(\{i_1^r, j_1^r, \ldots, i_{l^r}^r, j_{l^r}^r\}). \tag{5.321}$$

We will refer to the sets $\{i_1^r, \ldots, j_{l^r}^r\}$ as "ordered index sets".

The T notation has already been introduced in Subsection 5.2.3, and we refer the reader to that subsection for examples. Note that for a set of the form $v = T(\{\cdot\}, \ldots, \{\cdot\})$, changing the order within one of the ordered index sets will typically change the set v, i.e. $T(\{1,2\}) \neq T(\{2,1\})$ in general. It does not make a difference, however, in what order the ordered index sets sets themselves are listed in the definition of v, i.e. $T(\{1,2\},\{3,4\}) = T(\{3,4\},\{1,2\})$. Thus such a set v is defined by an *unordered* collection of ordered index sets.

For the purposes of the following definition, note that a "lexicographical" ordering for ordered index sets is an ordering in which an ordered index set $\{i_1, j_1, \ldots, i_l, j_l\}$ is listed before a different ordered index set $\{i'_1, j'_1, \ldots, i'_{l'}, j'_{l'}\}$ iff there exists $k \in \{1, \ldots, l\}$ such that $i_r \leq i'_r$, $j_r \leq j'_r$ for all $r \leq k-1$ and either $i_k < i'_k$ or $i_k = i'_k$ and $j_k < j'_k$, or $i_k = i'_k$ and $j_k = j'_k$ and $k = l < l'$. (This is the same principle as "alphabetical ordering" but applied to numbers.) Thus for example, $\{1,2,1,5\}$ is listed prior to $\{1,3\}$ and to $\{1,2,2,4\}$ and to $\{1,2,1,5,1,1\}$.

Definition 5.16 *Given*

$$v = T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.322}$$

assume that no two of the ordered index sets are identical, and that the ordered index sets are arranged in lexicographical order. Where $r \in \{1, \ldots, s\}$, define

$$v(r, i_{l^r+1}^r, j_{l^r+1}^r) = T(\{i_1^1, \ldots, j_{l^1}^1\}, \ldots \{i_1^{r-1}, \ldots, j_{l^{r-1}}^{r-1}\}, \{i_1^r, \ldots, j_{l^r}^r, i_{l^r+1}^r, j_{l^r+1}^r\},$$

$$\{i_1^{r+1}, \ldots, j_{l^{r+1}}^{r+1}\}, \ldots, \{i_1^s, \ldots, j_{l^s}^s\}) \tag{5.323}$$

i.e. append $i_{l^r+1}^r, j_{l^r+1}^r$ onto the r'th ordered index set.

Similarly for any v and any positive integer $l^{s+1} \leq h$, define

$$v(s+1, \{i_1^{s+1}, j_1^{s+1}, \ldots, i_{l^{s+1}}^{s+1}, j_{l^{s+1}}^{s+1}\}) = \tag{5.324}$$

$$T(\{i_1^1, \ldots, j_{l^1}^1\}, \ldots \{i_1^s, \ldots, j_{l^s}^s\}, \{i_1^{s+1}, \ldots, j_{l^{s+1}}^{s+1}\}) \tag{5.325}$$

i.e. append the $s+1$'st ordered index set, $\{i_1^{s+1}, \ldots, j_{l^{s+1}}^{s+1}\}$, to v.

Recall that the sets v of the form (5.322) are defined by unordered collections of ordered index sets. The reason for introducing the lexicographical order in the definition of the sets $v(\cdot)$ is only as a means of identifying to which ordered index set we intend to append coordinates.

As an example of the $v(\cdot)$ notation, say that

$$v = T(\{1,2,1,4,2,4\}, \{1,1,2,3\}) \tag{5.326}$$

then

$$v(2,3,6) = T(\{1,1,2,3\}, \{1,2,1,4,2,4,3,6\}). \tag{5.327}$$

(The "2" in $v(2,3,6)$ means that the indices $3,6$ should be appended to the ordered index set that is second in the lexicographical order, namely $\{1,2,1,4,2,4\}$.) Similarly,

$$v(3,1,3,2,3,1,5) = T(\{1,1,2,3\}, \{1,2,1,4,2,4\}, \{1,3,2,3,1,5\}). \tag{5.328}$$

Stated loosely (and this should be evident from the discussion in Subsection 5.2.3), the sets

$$T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.329}$$

are the elements of the partition of P obtained by partitioning (as per Lemma 5.2) each set $R_{i_1^r, j_1^r, \ldots, i_{n^r}^r, j_{n^r}^r}$ for all $r = 1, \ldots, s$ and all $n^r = 1, \ldots, l^r$. The next few lemmas will quantify formally and precisely in what ways these $T(\cdot)$ sets can be used to partition P, but first we will outline the general idea.

There are essentially two ways in which we will be partitioning sets. Sets of the form $T(\{i_1^1, j_1^1\})$ are obtained from P by partitioning the set $R_{i_1^1}$ as

$$R_{i_1^1} = \bigcup_{j_1^1=1}^{t_1(i_1^1)} \left(\bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1, \bar{j}_1})^c \cap Q_{i_1^1, j_1^1} \right) \tag{5.330}$$

which implies that P may be partitioned as

$$P = \bigcap_{i_1=1}^{m_1} R_{i_1} = \bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1}}^{m_1} R_{i_1} \cap \bigcup_{j_1^1=1}^{t_1(i_1^1)} \left(\bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1, \bar{j}_1})^c \cap Q_{i_1^1, j_1^1} \right) = \tag{5.331}$$

$$\bigcup_{j_1^1=1}^{t_1(i_1^1)} \left(\bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1}}^{m_1} R_{i_i} \cap \bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1, \bar{j}_1})^c \cap Q_{i_1^1, j_1^1} \right) = \tag{5.332}$$

$$\bigcup_{j_1^1=1}^{t_1(i_1^1)} T(\{i_1^1, j_1^1\}). \tag{5.333}$$

Considering that $Q_{i_1^1, j_1^1}$ is itself an intersection of sets of the form $R_{i_1^1, j_1^1, i_2^1}$ it follows that the sets $T(\{i_1^1, j_1^1\})$ are intersections of sets of the form $R_{i_1},\ i_1 \neq i_1^1$, sets of the form $R_{i_1^1, j_1^1, i_2^1}$, and sets of the form $(Q_{i_1^1, \bar{j}_1})^c$. The lattermost sets we will be ignoring for the moment, so the methodological question that presents itself is whether to further partition $T(\{i_1^1, j_1^1\})$ by going back now to partition one of the other $R_{i_1},\ i_1 \neq i_1^1$, or to partition $T(\{i_1^1, j_1^1\})$ by partitioning one of the $R_{i_1^1, j_1^1, i_2^1}$ (which can be thought of as a further partitioning of $R_{i_1^1}$). The particulars of these two strategies are as follows. The first possible strategy is to partition some $R_{i_1^2},\ i_1^2 \neq i_1^1$ as

$$R_{i_1^2} = \bigcup_{j_1^2=1}^{t_1(i_1^2)} \left(\bigcap_{\bar{j}_1=1}^{j_1^2-1} (Q_{i_1^2, \bar{j}_1})^c \cap Q_{i_1^2, j_1^2} \right) \tag{5.334}$$

which yields a partition of $T(\{i_1^1, j_1^1\})$ as

$$T(\{i_1^1, j_1^1\}) = \bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1}}^{m_1} R_{i_1} \cap \bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1,\bar{j}_1})^c \cap Q_{i_1^1,j_1^1} = \tag{5.335}$$

$$\bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1, i_1^2}}^{m_1} R_{i_1} \cap \bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1,\bar{j}_1})^c \cap Q_{i_1^1,j_1^1} \cap \bigcup_{j_1^2=1}^{t_1(i_1^2)} \left(\bigcap_{\bar{j}_1=1}^{j_1^2-1} (Q_{i_1^2,\bar{j}_1})^c \cap Q_{i_1^2,j_1^2} \right) = \tag{5.336}$$

$$\bigcup_{j_1^2=1}^{t_1(i_1^2)} \left(\bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1, i_1^2}}^{m_1} R_{i_1} \cap \bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1,\bar{j}_1})^c \cap Q_{i_1^1,j_1^1} \cap \bigcap_{\bar{j}_1=1}^{j_1^2-1} (Q_{i_1^2,\bar{j}_1})^c \cap Q_{i_1^2,j_1^2} \right) = \tag{5.337}$$

$$\bigcup_{j_1^2=1}^{t_1(i_1^2)} (T(\{i_1^1, j_1^1\}) \cap T(\{i_1^2, j_1^2\})) = \tag{5.338}$$

$$\bigcup_{j_1^2=1}^{t_1(i_1^2)} (T(\{i_1^1, j_1^1\}, \{i_1^2, j_1^2\}). \tag{5.339}$$

The second possible strategy is that rather than partitioning some other R_{i_1}, we may partition the sets $R_{i_1^1,j_1^1,i_2^1}$ as

$$R_{i_1^1,j_1^1,i_2^1} = \bigcup_{j_2^1=1}^{t_2(i_1^1,j_1^1,i_2^1)} \left(\bigcap_{\bar{j}_2=1}^{j_2^1-1} (Q_{i_1^1,j_1^1,i_2^1,\bar{j}_2})^c \cap Q_{i_1^1,j_1^1,i_2^1,j_2^1} \right), \tag{5.340}$$

yielding a partition of $T(\{i_1^1, j_1^1\})$ as

$$T(\{i_1^1, j_1^1\}) = \bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1}}^{m_1} R_{i_1} \cap \bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1,\bar{j}_1})^c \cap Q_{i_1^1,j_1^1} = \tag{5.341}$$

$$\bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1}}^{m_1} R_{i_1} \cap \bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1,\bar{j}_1})^c \cap \bigcap_{\substack{i_2=1 \\ i_2 \neq i_2^1}}^{m_2(i_1^1,j_1^1)} R_{i_1^1,j_1^1,i_2} \cap \bigcup_{j_2^1=1}^{t_2(i_1^1,j_1^1,i_2^1)} \left(\bigcap_{\bar{j}_2=1}^{j_2^1-1} (Q_{i_1^1,j_1^1,i_2^1,\bar{j}_2})^c \cap Q_{i_1^1,j_1^1,i_2^1,j_2^1} \right) = \tag{5.342}$$

$$\bigcup_{j_2^1=1}^{t_2(i_1^1,j_1^1,i_2^1)} \left(\bigcap_{\substack{i_1=1 \\ i_1 \neq i_1^1}}^{m_1} R_{i_1} \cap \bigcap_{\bar{j}_1=1}^{j_1^1-1} (Q_{i_1^1,\bar{j}_1})^c \cap \bigcap_{\substack{i_2=1 \\ i_2 \neq i_2^1}}^{m_2(i_1^1,j_1^1)} R_{i_1^1,j_1^1,i_2} \cap \bigcap_{\bar{j}_2=1}^{j_2^1-1} (Q_{i_1^1,j_1^1,i_2^1,\bar{j}_2})^c \cap Q_{i_1^1,j_1^1,i_2^1,j_2^1} \right) = \tag{5.343}$$

$$\bigcup_{j_2^1=1}^{t_2(i_1^1,j_1^1,i_2^1)} \Big(T(i_1^1, j_1^1) \cap T(i_1^1, j_1^1, i_2^1, j_2^1) \cap Q_{i_1^1,j_1^1,i_2^1,j_2^1} \Big) = \tag{5.344}$$

$$\bigcup_{j_2^1=1}^{t_2(i_1^1,j_1^1,i_2^1)} T(\{i_1^1, j_1^1, i_2^1, j_2^1\}). \tag{5.345}$$

The "depth first" methodology is to first partition $T(\{i_1^1, j_1^1\})$ by way of a partitioning of one of the $R_{i_1^1,j_1^1,i_2^1}$ to form sets of the form $T(\{i_1^1, j_1^1, i_2^1, j_2^1\})$, as we have seen. We will soon see that each of these sets can then in turn be partitioned by way of a partitioning of one of the $R_{i_1^1,j_1^1,i_2^1,j_2^1,i_3^1}$ to yield sets of the form $T(\{i_1^1, j_1^1, i_2^1, j_2^1, i_3^1, j_3^1\})$. These sets are similarly partitioned until eventually by way of a partitioning of sets of the form $R_{i_1^1,j_1^1,\ldots,i_h^1}$ into intersections of sets of the form M_l we will obtain finally a partition of P into sets of the form $T(\{i_1^1, j_1^1, \ldots, i_h^1, j_h^1\})$. When we reach this point and can no longer continue to partition in this manner then we return to one of the other R_{i_1}, and then repeat the methodology from there.

Explicitly, the depth first methodology is to first partition P as

$$P = \bigcup_{j_1^1=1}^{t_1(i_1^1)} T(\{i_1^1, j_1^1\}) \tag{5.346}$$

for each $i_1^1 = 1, \ldots, m_1$, and then to partition eah set $T(\{i_1^1, j_1^1\})$ as

$$T(\{i_1^1, j_1^1\}) = \bigcup_{j_2^1=1}^{t_2(i_1^1,j_1^1,i_2^1)} T(\{i_1^1, j_1^1, i_2^1, j_2^1\}) \tag{5.347}$$

for each $i_2^1 = 1, \ldots, m_2(i_1^1, j_1^1)$, and then to partition each $T(\{i_1^1, j_1^1, i_2^1, j_2^1\})$ as

$$T(\{i_1^1, j_1^1, i_2^1, j_2^1\}) = \bigcup_{j_3^1=1}^{t_3(i_1^1,j_1^1,i_2^1,j_2^1,i_3^1)} T(\{i_1^1, j_1^1, i_2^1, j_2^1, i_3^1, j_3^1\}) \tag{5.348}$$

for each $i_3^1 = 1, \ldots, m_3(i_1^1, j_1^1, i_2^1, j_2^1)$, until eventually we partition each $T(\{i_1^1, \ldots, j_{h-1}^1\})$ as

$$T(\{i_1^1, \ldots, j_{h-1}^1\}) = \bigcup_{j_h^1=1}^{t_h(\cdot)} T(\{i_1^1, \ldots, j_{h-1}^1, i_h^1, j_h^1\}) \tag{5.349}$$

for each $i_h^1 = 1, \ldots, m_h(\cdot)$. Each $T(\{i_1^1, \ldots, j_h^1\})$ is then partitioned as

$$T(\{i_1^1, \ldots, j_h^1\}) = \bigcup_{j_1^2=1}^{t_1(i_1^2)} T(\{i_1^1, \ldots, j_h^1\}, \{i_1^2, j_1^2\}) \tag{5.350}$$

for each $i_1^2 = 1, \ldots, m_1$. (Technically we do not need to consider the case where $i_1^2 = i_1^1$, but the statement (5.350) remains true in this case also, and the presentation is made easier by assuming that we allow $i_1^2 = i_1^1$.) Each $T(\{i_1^1, \ldots, j_h^1\}, \{i_1^2, j_1^2\})$ is then partitioned as

$$T(\{i_1^1, \ldots, j_h^1\}, \{i_1^2, j_1^2\}) = \bigcup_{j_2^2=1}^{t_2(i_1^2,j_1^2,i_2^2)} T(\{i_1^1, \ldots, j_h^1\}, \{i_1^2, j_1^2, i_2^2, j_2^2\}) \tag{5.351}$$

for each $i_2^2 = 1, \ldots, m_2(i_1^2, j_1^2)$, and so on.

The alternative methodology, to which we will refer as "breadth first", is to partition $T(\{i_1^1, j_1^1\}$ into sets of the form $T(\{i_1^1, j_1^1\}, \{i_1^2, j_1^2\})$ by way of partitioning some $R_{i_1^2},\ i_1^2 \neq i_1^1$, as we have seen. We will see soon that the $T(\{i_1^1, j_1^1\}, \{i_1^2, j_1^2\})$ sets can themselves be partitioned into sets of the form $T(\{i_1^1, j_1^1\}, \{i_1^2, j_1^2\}, \{i_1^3, j_1^3\})$ by partitioning some $R_{i_1^3},\ i_1^3 \neq i_1^1, i_1^2$, until eventually after partitioning every $R_{i_1},\ i_1 = 1, \ldots, m_1$ we will obtain a partition of P into sets of the form $T(\{i_1^1, j_1^1\}, \ldots, \{i_1^{m_1}, j_1^{m_1}\})$. At this point we will first begin to partition sets of the form $R_{i_1^1, j_1^1, i_2^1}$ in order to partition the sets of the form $T(\{i_1^1, j_1^1\}, \ldots, \{i_1^{m_1}, j_1^{m_1}\})$ into sets of the form

$$T(\{i_1^1, j_1^1, i_2^1, j_2^1\}, \{i_1^2, j_1^2\}, \ldots, \{i_1^{m_1}, j_1^{m_1}\}). \tag{5.352}$$

In like manner we will partition sets of the form (5.352) into sets of the form

$$T(\{i_1^1, j_1^1, i_2^1, j_2^1\}, \{i_1^2, j_1^2, i_2^2, j_2^2\}, \{i_1^3, j_1^3\} \ldots, \{i_1^{m_1}, j_1^{m_1}\}), \tag{5.353}$$

and so on.

Lemma 5.17 will formally describe depth first partitioning, and Lemma 5.18 will formally describe breadth first partitioning.

Lemma 5.17 *For each size k collection of indexing families of functions for P,*

$$\{\mathcal{I}^r = (I_1^r, I_2^r(j_1), \ldots, I_h^r(j_1, \ldots, j_{h-1})) : \ r = 1, \ldots, k\} \tag{5.354}$$

P can be partitioned into the disjoint union

$$P = \bigcup_{j_1^1=1}^{t_1(I_1^1)} \cdots \bigcup_{j_h^1=1}^{t_h(I_1^1, j_1^1 \ldots, I_h^1)} \cdots \bigcup_{j_1^k=1}^{t_1(I_1^k)} \cdots \bigcup_{j_h^k=1}^{t_h(I_1^k, j_1^k \ldots, I_h^k)} T(\{I_1^1, \ldots, j_h^1\}, \ldots, \{I_1^k, \ldots, j_h^k\}) \tag{5.355}$$

(Though we have suppressed the dependence notation, expressions such as $T(\{I_1^r, j_1^r, I_2^r, j_2^r\})$ should be understood to mean $T(\{I_1^r, j_1^r, I_2^r(j_1^r), j_2^r\})$, etc.)

Proof: Observe first that by Lemma 5.2, for any integer $1 \leq s \leq h-1$, and any $Q_{i_1, j_1, \ldots, i_s, j_s}$, for each $i_{s+1} \in \{1, \ldots, m_{s+1}(i_1, \ldots, j_s)\}$, we have

$$Q_{i_1, j_1, \ldots, i_s, j_s} = \bigcap_{\bar{i}_{s+1}=1}^{m_{s+1}(i_1, j_1, \ldots, i_s, j_s)} R_{i_1, j_1, \ldots, i_s, j_s, \bar{i}_{s+1}} = \tag{5.356}$$

$$\bigcap_{\substack{\bar{i}_{s+1}=1 \\ \bar{i} \neq i_{s+1}}}^{m_{s+1}(i_1, \ldots, j_s)} R_{i_1, \ldots, j_s, \bar{i}_{s+1}} \cap \bigcup_{j_{s+1}=1}^{t_{s+1}(i_1, \ldots, j_s, i_{s+1})} \left(\bigcap_{\bar{j}_{s+1}=1}^{j_{s+1}-1} (Q_{i_1, \ldots, j_s, i_{s+1}, \bar{j}_{s+1}})^c \cap Q_{i_1, \ldots, j_s, i_{s+1}, j_{s+1}} \right) = \tag{5.357}$$

$$\bigcup_{j_{s+1}=1}^{t_{s+1}(i_1,\ldots,j_s,i_{s+1})} (T(i_1,j_1,\ldots,i_{s+1},j_{s+1}) \cap Q_{i_1,j_1,\ldots,i_{s+1},j_{s+1}}). \tag{5.358}$$

This also holds for the case $s = 0$, if we think of P as being of the form $Q_{i_1,j_1,\ldots,i_s,j_s}$ with $s = 0$.

Thus applying Lemma 5.2 repeatedly, for each $r = 1,\ldots,k$, we have

$$P = \bigcup_{j_1^r=1}^{t_1(I_1^r)} (T(I_1^r,j_1^r) \cap Q_{I_1^r,j_1^r}) = \tag{5.359}$$

$$\bigcup_{j_1^r=1}^{t_1(I_1^r)} \left(T(I_1^r,j_1^r) \cap \left(\bigcup_{j_2^r=1}^{t_2(I_1^r,j_1^r,I_2^r)} (T(I_1^r,j_1^r,I_2^r,j_2^r) \cap Q_{I_1^r,j_1^r,I_2^r,j_2^r}) \right) \right) = \tag{5.360}$$

$$\bigcup_{j_1^r=1}^{t_1(I_1^r)} \bigcup_{j_2^r=1}^{t_2(I_1^r,j_1^r,I_2^r)} (T(I_1^r,j_1^r) \cap T(I_1^r,j_1^r,I_2^r,j_2^r) \cap Q_{I_1^r,j_1^r,I_2^r,j_2^r}) = \cdots = \tag{5.361}$$

$$\bigcup_{j_1^r=1}^{t_1(I_1^r)} \cdots \bigcup_{j_h^r=1}^{t_h(I_1^r,\ldots,j_h^r)} (T(I_1^r,j_1^r) \cap \cdots \cap T(I_1^r,j_1^r,\ldots,I_h^r,j_h^r) \cap Q_{I_1^r,j_1^r,\ldots,I_h^r,j_h^r}) = \tag{5.362}$$

$$\bigcup_{j_1^r=1}^{t_1(I_1^r)} \cdots \bigcup_{j_h^r=1}^{t_h(I_1^r,\ldots,j_h^r)} T(\{I_1^r,j_1^r,\ldots,I_h^r,j_h^r\}) \tag{5.363}$$

where all unions are disjoint. But since

$$T(\{I_1^1,j_1^1,\ldots,I_h^1,j_h^1\}) = T(\{I_1^1,j_1^1,\ldots,I_h^1,j_h^1\}) \cap P \tag{5.364}$$

this implies that

$$P = \bigcup_{j_1^1=1}^{t_1(I_1^1)} \cdots \bigcup_{j_h^1=1}^{t_h(I_1^1,\ldots,j_h^1)} T(\{I_1^1,j_1^1,\ldots,I_h^1,j_h^1\}) = \tag{5.365}$$

$$\bigcup_{j_1^1=1}^{t_1(I_1^1)} \cdots \bigcup_{j_h^1=1}^{t_h(I_1^1,\ldots,j_h^1)} \bigcup_{j_1^2=1}^{t_1(I_1^2)} \cdots \bigcup_{j_h^2=1}^{t_h(I_1^2,\ldots,j_h^2)} (T(\{I_1^1,\ldots,j_h^1\}) \cap T(\{I_1^2,\ldots,j_h^2\})). \tag{5.366}$$

Repeating the argument, we conclude that

$$P = \bigcup_{j_1^1=1}^{t_1(I_1^1)} \cdots \bigcup_{j_h^1=1}^{t_h(I_1^1,\ldots,j_h^1)} \cdots \bigcup_{j_1^k=1}^{t_1(I_1^k)} \cdots \bigcup_{j_h^k=1}^{t_h(I_1^k,\ldots,j_h^k)} T(\{I_1^1,\ldots,j_h^1\},\ldots,\{I_1^k,\ldots,j_h^k\}). \;\Box \tag{5.367}$$

Thus with depth first partitioning, P is partitioned into sets $T(\{i_1^1,j_1^1,\ldots,i_h^1,j_h^1\})$, and then by compounding the procedure it is partitioned into sets of the form

$$T(\{i_1^1,j_1^1,\ldots,i_h^1,j_h^1\},\{i_1^2,j_1^2,\ldots,i_h^2,j_h^2\}), \tag{5.368}$$

etc. with the ordered index sets always of maximum length (i.e. length $2h$).

Example: Consider again

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.369}$$

which, as above, can be represented as

$$P = \bigcap_{i_1=1}^{1} \bigcup_{j_1=1}^{2} \bigcap_{i_2=1}^{2} \bigcup_{j_2=1}^{2} Y_{f(i_1,j_1,i_2,j_2)} \tag{5.370}$$

where f maps

$$(1,1,1,1) \to 1,\ (1,1,1,2) \to 2,\ (1,1,2,1) \to 3, \ldots, (1,2,2,2) \to 8. \tag{5.371}$$

Let $\mathcal{I}^1$ be defined by

$$I_1^1 = 1,\ I_2^1(1) = 1,\ I_2^1(2) = 2 \tag{5.372}$$

and let $\mathcal{I}^2$ be defined by

$$I_1^2 = 1,\ I_2^2(1) = 2,\ I_2^2(2) = 1. \tag{5.373}$$

Lemma 5.17 now says that P can be partitioned as

$$P = \bigcup_{j_1^1=1}^{2} \bigcup_{j_2^1=1}^{2} \bigcup_{j_1^2=1}^{2} \bigcup_{j_2^2=1}^{2} T(\{1, j_1^1, I_2^1(j_1^1), j_1^1\}, \{1, j_1^2, I_2^2(j_1^2), j_1^2\}) = \tag{5.374}$$

$$\begin{aligned}
&T(\{1,1,1,1\},\{1,1,2,1\}) \cup T(\{1,1,1,1\},\{1,1,2,2\}) \cup \\
&T(\{1,1,1,1\},\{1,2,1,1\}) \cup T(\{1,1,1,1\},\{1,2,1,2\}) \cup \\
&T(\{1,1,1,2\},\{1,1,2,1\}) \cup T(\{1,1,1,2\},\{1,1,2,2\}) \cup \\
&T(\{1,1,1,2\},\{1,2,1,1\}) \cup T(\{1,1,1,2\},\{1,2,1,2\}) \cup \\
&T(\{1,2,2,1\},\{1,1,2,1\}) \cup T(\{1,2,2,1\},\{1,1,2,2\}) \cup \\
&T(\{1,2,2,1\},\{1,2,1,1\}) \cup T(\{1,2,2,1\},\{1,2,1,2\}) \cup \\
&T(\{1,2,2,2\},\{1,1,2,1\}) \cup T(\{1,2,2,2\},\{1,1,2,2\}) \cup \\
&T(\{1,2,2,2\},\{1,2,1,1\}) \cup T(\{1,2,2,2\},\{1,2,1,2\}).
\end{aligned} \tag{5.375}$$

It is worth noting that in simply compounding the procedure we will end up with some sets in the partition that are empty. The set $T(\{1,1\})$, for example, is an element in the partition of P obtained by partitioning R_1, and the set $T(\{1,2\})$ is a different element of that same partition. Thus $T(\{1,1\},\{1,2\}) = T(\{1,1\}) \cap T(\{1,2\}) = \emptyset$. Specifically, in the decomposition (5.375), the second, fourth, fifth, and seventh lines of the expression are comprised exclusively of empty sets. In general, the set

$$T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.376}$$

is empty if there exist $r, r' \in \{1, \ldots, s\}$ for which there is a $k \geq 1$ such that $i_l^r = i_l^{r'}$, $j_l^r = j_l^{r'}$, $l = 1, \ldots, k-1$, $i_k^r = i_k^{r'}$, and $j_{k+1}^r \neq j_{k+1}^{r'}$. $\square$

Lemma 5.18 *Define the j-indexing family of functions*

$$\mathcal{J}^r = \{J_1(i_1), J_2(i_1, i_2), \ldots, J_r(i_1, \ldots, i_r)\} \tag{5.377}$$

define the function

$$\mathcal{J}^r(i_1, \ldots, i_r) = (i_1, J_1(i_1), \ldots, i_r, J_r(i_1, \ldots, i_r)) \tag{5.378}$$

and let $i(\mathcal{J}^r)$ be the domain of that function. Then for each $r = 1, \ldots, h$, P can be partitioned into the disjoint union

$$P = \bigcup_{\mathcal{J}^r} \bigcap_{i(\mathcal{J}^r)} T(\{i_1, J_1, \ldots, i_r, J_r\}). \tag{5.379}$$

Proof: For each $i_1 = 1, \ldots, m_1$, by Lemma 5.2 we have

$$P = \bigcup_{j_1=1}^{t_1(i_1)} T(\{i_1, j_1\}) \Rightarrow \tag{5.380}$$

$$P = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} T(\{i_1, j_1\}). \tag{5.381}$$

Moreover, for any $T(\{i_1, j_1, \ldots, i_s, j_s\})$, and any $i_{s+1} = 1, \ldots, m_{s+1}(i_1, \ldots, j_s)$,

$$T(\{i_1, j_1, \ldots, i_s, j_s\}) = \bigcup_{j_{s+1}=1}^{t_{s+1}(i_1,\ldots,j_s)} T(\{i_1, j_1, \ldots, i_{s+1}, j_{s+1}\}) \Rightarrow \tag{5.382}$$

$$T(\{i_1, j_1, \ldots, i_s, j_s\}) = \bigcap_{i_{s+1}=1}^{m_{s+1}(i_1,\ldots,j_s)} \bigcup_{j_{s+1}=1}^{t_{s+1}(i_1,\ldots,j_s)} T(\{i_1, j_1, \ldots, i_{s+1}, j_{s+1}\}). \tag{5.383}$$

Repeating this argument for $s = 1, \ldots, r$ we conclude that

$$P = \bigcap_{i_1=1}^{m_1} \bigcup_{j_1=1}^{t_1(i_1)} \cdots \bigcap_{i_r=1}^{m_r(\cdot)} \bigcup_{j_r=1}^{t_r(\cdot)} T(\{i_1, \ldots, j_r\}) \Rightarrow \tag{5.384}$$

$$P = \bigcup_{\mathcal{J}^r} \bigcap_{i(\mathcal{J}^r)} T(\{i_1, J_1, \ldots, i_r, J_r\}) \tag{5.385}$$

by Lemma 5.12. To show that the union is disjoint, consider two distinct j-indexing families $\mathcal{J}^r$ and $(\mathcal{J}^r)'$. As these are distinct, there must be some $(i_1, \ldots, i_r)$ such that $\mathcal{J}^r(i_1, \ldots, i_r) \neq (\mathcal{J}^r)'(i_1, \ldots, i_r)$. Let l be minimum subject to $J_l(i_1, \ldots, i_l) \neq J_l'(i_1, \ldots, i_l)$, and assume without loss of generality that $J_l(i_1, \ldots, i_l) < J_l'(i_1, \ldots, i_l)$, then

$$T(\{i_1, J_1, \ldots, i_{l-1}, J_{l-1}, i_l, J_l\}) \subseteq Q_{i_1,J_1,\ldots,i_{l-1},J_{l-1},i_l,J_l} \tag{5.386}$$

and

$$T(\{i_1, J_1', \ldots, i_{l-1}, J_{l-1}', i_l, J_l'\}) = \tag{5.387}$$

$$T(\{i_1, J_1, \ldots, i_{l-1}, J_{l-1}, i_l, J_l'\}) \subseteq (Q_{i_1,J_1,\ldots,i_{l-1},J_{l-1},i_l,J_l})^c \tag{5.388}$$

by definition of the set $T(\{i_1, J_1, \ldots, i_{l-1}, J_{l-1}, i_l, J_l'\})$. □

Example: Consider again

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.389}$$

which, as above, can be represented as

$$P = \bigcap_{i_1=1}^{1} \bigcup_{j_1=1}^{2} \bigcap_{i_2=1}^{2} \bigcup_{j_2=1}^{2} Y_{f(i_1,j_1,i_2,j_2)} \tag{5.390}$$

where f maps

$$(1,1,1,1) \to 1,\ (1,1,1,2) \to 2,\ (1,1,2,1) \to 3, \ldots, (1,2,2,2) \to 8. \tag{5.391}$$

Let $r = h = 2$. Then the possible functions $J_1(i_1)$ are the constants $J_1 = 1$ and $J_1 = 2$. Denote these functions as J_1' and J_1'' respectively. The possible functions J_2 are as follows

$$J_2' : \quad J_2(1,1) = 1,\ J_2(1,2) = 1 \tag{5.392}$$

$$J_2'' : \quad J_2(1,1) = 1,\ J_2(1,2) = 2 \tag{5.393}$$

$$J_2''' : \quad J_2(1,1) = 2,\ J_2(1,2) = 1 \tag{5.394}$$

$$J_2'''' : \quad J_2(1,1) = 2,\ J_2(1,2) = 2 \tag{5.395}$$

There are thus eight possible families of functions $\mathcal{J}$, to be denoted $\mathcal{J}^1, \ldots, \mathcal{J}^8$, and these are

$$(J_1', J_2'),\ (J_1', J_2''),\ (J_1', J_2'''),\ (J_1', J_2''''),\ (J_1'', J_2'),\ (J_1'', J_2''),\ (J_1'', J_2'''),\ (J_1'', J_2''''). \tag{5.396}$$

Lemma 5.18 therefore says that P can be partitioned as

$$P = \bigcup_{r=1}^{8} T(\{1, J_1^r, 1, J_2^r(1,1)\}\{1, J_1^r, 2, J_2^r(1,2)\}) = \tag{5.397}$$

$$T(\{1,1,1,1\},\{1,1,2,1\}) \cup T(\{1,1,1,1\},\{1,1,2,2\}) \cup$$
$$T(\{1,1,1,2\},\{1,1,2,1\}) \cup T(\{1,1,1,2\},\{1,1,2,2\}) \cup$$
$$T(\{1,2,1,1\},\{1,2,2,1\}) \cup T(\{1,2,1,1\},\{1,2,2,2\}) \cup$$
$$T(\{1,2,1,2\},\{1,2,2,1\}) \cup T(\{1,2,1,2\},\{1,2,2,2\}) \tag{5.398}$$

The reader can check that expression (5.398) is the same as the partition we yielded for this example by applying Lemma 5.17 (expression 5.375) above (ignoring the empty elements in 5.375 as described there). □

Either via breadth first partitioning or via depth first partitioning or via some combination of the two, the partition of P can be made increasingly fine by partitioning each $R_{\{\cdot\}}$ set until eventually $T(\{\cdot\},\ldots,\{\cdot\})$ becomes an intersection of sets of the form $Q^c_{\{\cdot\}}$ and $M_{f(\cdot)}$ alone.

Lemma 5.19 *For any j-indexing family of functions $\mathcal{J}$,*

$$\bigcap_{i(\mathcal{J})} T(\{i_1, J_1, \ldots, i_h, J_h\}) = \tag{5.399}$$

$$\bigcap_{i(\mathcal{J})} \left(\left(\bigcap_{l=1}^{h} \bigcap_{\bar{j}_l=1}^{J_l-1} (Q_{i_1,J_1,\ldots,i_{l-1},J_{l-1},i_l,\bar{j}_l})^c \right) \cap M_{f(i_1,J_1,\ldots,i_h,J_h)} \right). \tag{5.400}$$

Proof: Observe that

$$R_{i_1,J_1,\ldots,i_l} \supseteq Q_{i_1,J_1,\ldots,i_l,J_l}.$$

Observe moreover that

$$Q_{i_1,J_1,\ldots,i_l,J_l} = \bigcap_{i_{l+1}=1}^{m_{l+1}(\cdot)} R_{i_1,J_1,\ldots,i_l,J_l,i_{l+1}} =$$
$$\bigcap_{i_{l+1}=1}^{m_{l+1}(\cdot)} \bigcup_{j_{l+1}=1}^{t_{l+1}(\cdot)} Q_{i_1,J_1,\ldots,i_{l+1},j_{l+1}} \supseteq \bigcap_{i_{l+1}=1}^{m_{l+1}(\cdot)} Q_{i_1,J_1,\ldots,i_{l+1},J_{l+1}}. \tag{5.401}$$

Putting these results together and applying (5.401) repeatedly we conclude that

$$R_{i_1,J_1,\ldots,i_l} \supseteq \bigcap_{i_{l+1}=1}^{m_{l+1}(\cdot)} \cdots \bigcap_{i_h=1}^{m_h(\cdot)} M_{f(i_1,J_1,\ldots,i_h,J_h)}. \tag{5.402}$$

Now

$$\bigcap_{i(\mathcal{J})} T(\{i_1, J_1, \ldots, i_h, J_h\}) = \tag{5.403}$$

$$\bigcap_{i(\mathcal{J})}\left(\bigcap_{l=1}^{h} T(i_1, J_1, \ldots, i_l, J_l) \cap M_{f(i_1,J_1,\ldots,i_h,J_h)}\right) = \tag{5.404}$$

$$\bigcap_{i(\mathcal{J})}\left(\bigcap_{l=1}^{h}\left(\bigcap_{\substack{\bar{i}_l=1\\ \bar{i}_l\neq i_l}}^{m_l(i_1,\ldots,J_{l-1})} R_{i_1,J_1,\ldots,i_{l-1},J_{l-1},\bar{i}_l} \cap \bigcap_{\bar{j}_l=1}^{J_l-1}(Q_{i_1,J_1,\ldots,i_{l-1},J_{l-1},i_l,\bar{j}_l})^c\right) \cap M_{f(i_1,J_1,\ldots,i_h,J_h)}\right) = \tag{5.405}$$

$$\bigcap_{i(\mathcal{J})}\left(\left(\bigcap_{l=1}^{h}\bigcap_{\bar{j}_l=1}^{J_l-1}(Q_{i_1,J_1,\ldots,i_{l-1},J_{l-1},i_l,\bar{j}_l})^c\right) \cap M_{f(i_1,J_1,\ldots,i_h,J_h)}\right) \tag{5.406}$$

by (5.402). □

Corollary 5.20

$$P = \bigcup_{\mathcal{J}}\bigcap_{i(\mathcal{J})}\left(\left(\bigcap_{l=1}^{h}\bigcap_{\bar{j}_l=1}^{J_l-1}(Q_{i_1,J_1,\ldots,i_{l-1},J_{l-1},i_l,\bar{j}_l})^c\right) \cap M_{f(i_1,J_1,\ldots,i_h,J_h)}\right) \tag{5.407}$$

and the union is disjoint. □

In the case of the example

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.408}$$

discussed above, we can observe that the term $T(\{1,2,1,1\},\{1,2,2,1\})$ from expressions (5.375) and (5.398) is (by applying (5.185) and (5.187))

$$T(\{1,2,1,1\},\{1,2,2,1\}) = T(\{1,2,1,1\}) \cap T(\{1,2,2,1\}) = \tag{5.409}$$

$$Q_{1,1}^c \cap R_{1,2,1} \cap R_{1,2,2} \cap M_{f(1,2,1,1)} \cap M_{f(1,2,2,1)} = \tag{5.410}$$

$$Q_{1,1}^c \cap M_{f(1,2,1,1)} \cap M_{f(1,2,2,1)} \tag{5.411}$$

and a similar situation holds for the other terms of (5.375) and (5.398).

5.4 Depth-First Partitioning Algorithm

Recall from Corollary 3.19 that a vector $(x_1,\ldots,x_n) \in R^n$, construed as $(x[Y_1^P],\ldots,x[Y_n^P])$, belongs to $Conv(P)$ iff there is a probability measure $\bar{\chi}$ on $\mathcal{P}$ consistent with the values $x[Y_1^P],\ldots,x[Y_n^P]$. Note now that there is a probability measure $\bar{\chi}$ on $\mathcal{P}$ consistent with $x[Y_1^P],\ldots,x[Y_n^P]$ iff there is a measure χ on $\mathcal{A}$ consistent with $x[Y_1^P],\ldots,x[Y_n^P]$ such that $\chi[P] = 1$. (If an appropriate probability measure $\bar{\chi}$ exists then define χ by, say, $\chi[q] = \chi[q \cap P]$, $\forall q \in \mathcal{A}$, and if an appropriate χ exists then define $\bar{\chi}$ by $\bar{\chi}[q] = \chi[q]$, $\forall q \subseteq P$.) Note also that there is a measure χ on $\mathcal{A}$ consistent with $x[Y_1^P], \ldots, x[Y_n^P]$ such that $\chi[P] = 1$ iff there is a probability measure $\hat{\chi}$ on $\mathcal{A}$ with $\hat{\chi}[Y_l] = x_l,\ l = 1, \ldots, n$, and such that $\hat{\chi} = \hat{\chi}^P$. (If an appropriate $\hat{\chi}$ exists, then define $\chi = \hat{\chi}$, and if an appropriate χ exists then define $\hat{\chi} = \chi^P$.)

There are thus a number of equivalent ways in which we can describe the liftings that we will be performing. For the sake of maximal transparency, though this is a slight departure from the presentation in the earlier sections, the original vector $(x_1, \ldots, x_n) \in R^n$ that we seek to ensure belongs to $Conv(P)$, will be construed as $(x[Y_1^P], \ldots, x[Y_n^P])$, and the vector $(x_1, \ldots, x_n) = (x[Y_1^P], \ldots, x[Y_n^P])$ will belong to $Conv(P)$ if and only if there is a measure χ on $\mathcal{A}$ consistent with $x[Y_1^P], \ldots, x[Y_n^P]$, and such that $\chi[P] = 1$. Or equivalently, the vector $(x_1, \ldots, x_n) = (x[Y_1^P], \ldots, x[Y_n^P])$ will belong to $Conv(P)$ if and only if there is a probability measure $\bar{\chi}$ on $\mathcal{P}$ consistent with $x[Y_1^P], \ldots, x[Y_n^P]$.

We will lift the original vector $(x[Y_1^P], \ldots, x[Y_n^P])$ by creating new variables $x[q]$ corresponding to the set function values $\chi[q]$ on additional sets $q \in \mathcal{A}$, and we will place constraints on these new values arising from the requirements that the set function χ be a measure on $\mathcal{A}$. Note that since the partial sum χ^V, where $V \in \mathcal{A}$, is the set function on $\mathcal{A}$ defined by $\chi^V[q] = \chi[V \cap q]$ for each $q \in \mathcal{A}$, defining appropriate variables $x[q \cap V] = \chi[q \cap V]$ will allow us to describe (projections of) the partial sum vector χ^V as well.

Subsections 5.2.2 and 5.2.3 already outlined the basic structure of the algorithm, and in this section we will present it formally. What follows is a basic implementation; many refinements are possible, some of which will be described in the course of this and the next chapter.

Algorithm at Level $k \geq 1$

Step 1 : Form the Matrix

Where P is as in Definition 5.13, form a matrix U with rows indexed by the sets

$$P, Y_1, \ldots, Y_n, N_1, \ldots, N_n. \tag{5.412}$$

Form a column for each of the sets

$$v = \bigcap_{u=1}^{s} T(\{i_1^u, j_1^u, \ldots, i_{l^u}^u, j_{l^u}^u\}) =$$

$$T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots, \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.413}$$

(defined in Definition 5.15) for all unordered collections of s ordered $2l^r$-tuples of positive integers ($r = 1, \ldots, s$),

$$\{i_1^1, \ldots, j_{l^1}^1\}, \ldots, \{i_1^s, \ldots, j_{l^s}^s\} \tag{5.414}$$

with all $i_u^r \leq m_u(i_1^r, \ldots, j_{u-1}^r)$, $j_u^r \leq t_u(i_1^r, \ldots, j_{u-1}^r, i_u^r)$ for all $1 \leq l^r \leq h$ and $0 \leq s \leq k$ for which the following conditions hold:

1. No ordered set $\{i_1^r, j_1^r, \ldots, i_{l^r}^r, j_{l^r}^r\}, 1 \leq r \leq s$, is equal to any other ordered set $\{i_1^u, j_1^u, \ldots, i_{l^r}^u, j_{l^r}^u\},\ 1 \leq u \leq s,\ u \neq r$.

2. For each $r, r' \leq s,\ r \neq r'$,

$$\{i_1^r, j_1^r, \ldots, i_u^r\} = \{i_1^{r'}, j_1^{r'}, \ldots, i_u^{r'}\} \Rightarrow j_u^r = j_u^{r'}. \tag{5.415}$$

(Technically, the columns are indexed by the tuples (5.414.)

Where v is of the form (5.413) and $s = 0$, we will say that $v = P$, and we will refer to the corresponding column as x^P.

Step 2 : Impose Constraints

Step 2(A) : General Measure Theoretic Constraints

Enforce:

$$x^P[P] = 1. \tag{5.416}$$

Where v is of the form (5.413), for each column U^v, we will denote U^v by x^v and we will denote the entries of the column by:

$$U^v[P] \leftrightarrow x_0^v \tag{5.417}$$

$$U^v[Y_i] \leftrightarrow x_{i'}^v \tag{5.418}$$

$$U^v[N_i] \leftrightarrow x_{i''}^v \tag{5.419}$$

For each v'th column, x^v, with v of the form (5.413), impose the constraints:

$$x^v \geq 0 \tag{5.420}$$

$$x^v[q] \leq x^v[P] \quad \text{for every row } q \tag{5.421}$$

$$x^v[Y_i] = x^v[P] - x^v[N_i], \quad i = 1, \ldots, n. \tag{5.422}$$

For each v'th column, with v of the form (5.413), for which $l^r = h$ for some $r \in \{1, \ldots, s\}$, impose the constraint

$$x^v_{f(i^r_1, j^r_1, \ldots, i^r_h, j^r_h)} = x^v_0. \tag{5.423}$$

Step 2(B) : Partitioning Constraints

Recalling the notation introduced in Definition 5.16, for each expression,

$$v = T(\{i^1_1, j^1_1, \ldots, i^1_{l^1}, j^1_{l^1}\}, \{i^2_1, j^2_1, \ldots, i^2_{l^2}, j^2_{l^2}\}, \ldots \{i^s_1, j^s_1, \ldots, i^s_{l^s}, j^s_{l^s}\}) \tag{5.424}$$

(assuming that the superscripts reflect a lexicographic ordering of the ordered index sets) and every

$$\{i^r_1, j^r_1, \ldots, i^r_{l^r}, j^r_{l^r}, i^r_{l^r+1}\}, \quad r \in \{1, \ldots, s\}, \ l^r < h \tag{5.425}$$

such that there is a column $v(r, i^r_{l^r+1}, j^r_{l^r+1})$ for each $j^r_{l^r+1} = 1, \ldots, t_{l^r+1}(\cdot)$, define $\bar{v}$ to be the expression obtained by discarding from the expression v all ordered index sets that are subsets of other ordered index sets, and impose

$$x^{\bar{v}} = \sum_{j^r_{l^r+1}=1}^{t_{l^r+1}(i^r_1, j^r_1, \ldots, i^r_{l^r}, j^r_{l^r}, i^r_{l^r+1})} x^{v(r, i^r_{l^r+1}, j^r_{l^r+1})}. \tag{5.426}$$

Finally, for each v'th column, where v is of the form (5.413), $s < k$, and each ordered set

$$\{l^{s+1}, i^{s+1}_1, j^{s+1}_1, \ldots, i^{s+1}_{l^{s+1}-1}, j^{s+1}_{l^{s+1}-1}, i^{s+1}_{l^{s+1}}\} \tag{5.427}$$

such that $h \geq l^{s+1} \geq 1$ and such that, if $l^{s+1} \geq 2$,

i. The ordered subset

$$\{i^{s+1}_1, j^{s+1}_1, \ldots, i^{s+1}_{l^{s+1}-1}, j^{s+1}_{l^{s+1}-1}\} \tag{5.428}$$

is equal to some ordered set

$$\{i^r_1, j^r_1, \ldots, i^r_{l^{s+1}-1}, j^r_{l^{s+1}-1}\}, \quad r \in \{1, \ldots, s\}, \ l^r > l^{s+1} - 1 \tag{5.429}$$

ii. The ordered set

$$\{i^{s+1}_1, j^{s+1}_1, \ldots, i^{s+1}_{l^{s+1}-1}, j^{s+1}_{l^{s+1}-1}, i^{s+1}_{l^{s+1}}\} \tag{5.430}$$

is not equal to any ordered set

$$\{i^r_1, j^r_1, \ldots, i^r_{l^{s+1}-1}, j^r_{l^{s+1}-1}, i^r_{l^{s+1}}\}, \quad r \in \{1, \ldots, s\} \tag{5.431}$$

we impose the constraint

$$x^v = \sum_{j^{s+1}_{l^{s+1}}=1}^{t_{l^{s+1}}(i^{s+1}_1, j^{s+1}_1, \ldots, i^{s+1}_{l^{s+1}-1}, j^{s+1}_{l^{s+1}-1}, i^{s+1}_{l^{s+1}})} x^{v(s+1,\{i^{s+1}_1, j^{s+1}_1, \ldots, i^{s+1}_{l^{s+1}}, j^{s+1}_{l^{s+1}}\})}. \square \qquad (5.432)$$

Comments on the Depth First Algorithm:

- Each entry $x^v[q]$ of the matrix is construed by the algorithm to be the value $x[v \cap q]$ of a lifted vector x consistent with some set function χ on $\mathcal{A}$. Each column x^v is thus a projection of the partial sum χ^v. For any set function χ on $\mathcal{A}$, we have $\chi^P[Y_l] = \chi[Y_l^P]$, so as we indicated at the beginning of the section, the vector $(x^P[Y_1], \ldots, x^P[Y_n]) = (\chi[Y_1^P], \ldots, \chi[Y_n^P])$ belongs to $Conv(P)$ iff χ can be chosen to be a measure on $\mathcal{A}$ with $\chi[P] = 1$. The constraints imposed by the algorithm are all necessity conditions for this to in fact be the case. (Equivalently, we may think of the lifted vector x as being consistent with a set function $\bar{\chi}$ on $\mathcal{P}$ (since for each q, v entry of the matrix, $q \cap v \subseteq P$), and $(x^P[Y_1], \ldots, x^P[Y_n]) = (\bar{\chi}[Y_1^P], \ldots, \bar{\chi}[Y_n^P])$ belongs to $Conv(P)$ iff $\bar{\chi}$ can be chosen to be a probability measure on $\mathcal{P}$.) The relaxation of $Conv(P)$ that is produced by the algorithm at level k is thus the set of vectors $\{x \in R^n : x = (U^P[Y_1], \ldots, U^P[Y_n])\}$ for some matrix U satisfying the algorithm constraints at level k.

- We could also view the rows as indexed by P, Y_i^P, N_i^P, as in any case we only defined columns corresponding to sets $v \subseteq P$. Equivalently we could see them as being indexed by the $2n$ dimensional representation of P and $Y'_{i'}, N'_{i'}$ or $Y'_{i'}, Y'_{i''}$. Note also that strictly speaking we didn't need to define rows for the N_i, but defining these rows will make certain aspects of the analysis cleaner.

- If a set $T(\{\cdot\}, \cdots, \{\cdot\})$ is such that condition (1) fails to hold, then we could discard the smaller index set and remain with the same T set. If $T(\{\cdot\}, \cdots, \{\cdot\})$ is such that condition (2) fails to hold, then it is empty.

- Constraints (5.420), (5.421) and (5.422) are justified by the facts that measures must be nonnegative, each set $v \cap q$ is a subset of P (since for each v'th column, $v \subseteq P$), and $Y_i \cap v$ and $N_i \cap v$ partition $P \cap v$.

- If a set v of the form (5.413) fails to satisfy condition (2), then any set $v(r, i^r_{l^r+1}, j^r_{l^r+1})$ must violate condition (2) as well. Thus if the matrix U indeed has a column corresponding to $v(r, i^r_{l^r+1}, j^r_{l^r+1})$ then v could not violate condition (2). Thus if the matrix

U has a column corresponding to $v(r, i^r_{l^r+1}, j^r_{l^r+1})$, then where $\bar{v}$ is the expression obtained by discarding from the expression v all ordered index sets that are subsets of other ordered index sets, then $\bar{v}$ satisfies condition (1) and there must therefore be a column in U for $\bar{v}$ as well.

- The partitioning constraints (5.426) and (5.432) are justified as follows: Observe that where $l^r < h$, and

$$v = T(\{i^1_1, j^1_1, \ldots, i^1_{l^1}, j^1_{l^1}\}, \{i^2_1, j^2_1, \ldots, i^2_{l^2}, j^2_{l^2}\}, \ldots \{i^s_1, j^s_1, \ldots, i^s_{l^s}, j^s_{l^s}\}) = \bigcap_{w=1}^{s} T(\{i^w_1, j^w_1, \ldots, i^w_{l^w}, j^w_{l^w}\}) \tag{5.433}$$

then

$$T(\{i^r_1, \ldots, j^r_{l^r}\}) = T(i^r_1, j^r_1) \cap \cdots \cap T(i^r_1, j^r_1, \cdots, i^r_{l^r}, j^r_{l^r}) \cap Q_{i^r_1, \ldots, j^r_{l^r}} \tag{5.434}$$

and

$$Q_{i^r_1, \ldots, j^r_{l^r}} = \bigcap_{i^r_{l^r+1}=1}^{m_{l^r+1}(\cdot)} R_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}}. \tag{5.435}$$

Thus for each $i^r_{l^r+1} = 1, \ldots, m_{r+1}(\cdot)$, we may apply Lemma 5.2 to partition the set

$$R_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}} = \bigcup_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} Q_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, j^r_{l^r+1}} \tag{5.436}$$

as

$$R_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}} = \bigcup_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} \left(\bigcap_{\bar{j}^r_{l^r+1}=1}^{j^r_{l^r+1}-1} (Q_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, \bar{j}^r_{l^r+1}})^c \cap Q_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, j^r_{l^r+1}} \right). \tag{5.437}$$

This now yields the partition, for each $i^r_{l^r+1} = 1, \ldots, m_{r+1}(\cdot)$,

$$Q_{i^r_1, \ldots, j^r_{l^r}} =$$

$$\bigcap_{\substack{\bar{i}^r_{l^r+1}=1 \\ \bar{i}^r_{l^r+1} \neq i^r_{l^r+1}}}^{m_{l^r+1}(\cdot)} R_{i^r_1, \ldots, j^r_{l^r}, \bar{i}^r_{l^r+1}} \cap \bigcup_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} \left(\bigcap_{\bar{j}^r_{l^r+1}=1}^{j^r_{l^r+1}-1} (Q_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, \bar{j}^r_{l^r+1}})^c \cap Q_{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, j^r_{l^r+1}} \right) = \tag{5.438}$$

$$\bigcup_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} T(i^r_1, \ldots, j^r_{l^r+1}) \cap Q_{i^r_1, \ldots, j^r_{l^r+1}}, \tag{5.439}$$

which yields the partition

$$T(\{i^r_1, \ldots, j^r_{l^r}\}) = \bigcup_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} T(\{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, j^r_{l^r+1}\}) \tag{5.440}$$

which yields the partition

$$v = \bigcup_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} v(r, i^r_{l^r+1}, j^r_{l^r+1}). \tag{5.441}$$

Note now that if some u'th ordered index set is a subset of, say, the r'th ordered index set, $r \neq u$, then

$$T(\{i^r_1, \ldots, j^r_{l^r}\}) = T(i^r_1, j^r_1) \cap \cdots \cap T(i^r_1, \ldots, j^r_{l^u}) \cap \cdots \cap T(i^r_1, j^r_1, \cdots, i^r_{l^r}, j^r_{l^r}) \cap Q_{i^r_1,\ldots,j^r_{l^r}} = \tag{5.442}$$

$$T(i^u_1, j^u_1) \cap \cdots \cap T(i^u_1, \ldots, j^u_{l^u}) \cap T(i^r_1, \ldots, j^r_{l^u+1}) \cap \cdots \cap T(i^r_1, j^r_1, \cdots, i^r_{l^r}, j^r_{l^r}) \cap Q_{i^r_1,\ldots,j^r_{l^r}}. \tag{5.443}$$

But

$$Q_{i^u_1,\ldots,j^u_{l^u}} = Q_{i^r_1,\ldots,j^r_{l^u}} \supseteq \tag{5.444}$$

$$T(i^r_1, \ldots, j^r_{l^u+1}) \cap \cdots \cap T(i^r_1, j^r_1, \cdots, i^r_{l^r}, j^r_{l^r}) \cap Q_{i^r_1,\ldots,j^r_{l^r}} \tag{5.445}$$

by the argument at the beginning of the proof of Lemma 5.17. This now implies that

$$T(\{i^r_1, \ldots, j^r_{l^r}\}) \subseteq T(i^u_1, j^u_1) \cap \cdots \cap T(i^u_1, \ldots, j^u_{l^u}) \cap Q_{i^u_1,\ldots,j^u_{l^u}} =$$

$$T(\{i^u_1, \ldots, j^u_{l^u}\}) \tag{5.446}$$

which implies that

$$v = \bigcap_{\substack{w=1 \\ w \neq u}}^{s} T(\{i^w_1, j^w_1, \ldots, i^w_{l^w}, j^w_{l^w}\}). \tag{5.447}$$

Repeating the argument we conclude that $v = \bar{v}$, so that $\bar{v}$ may also be partitioned as

$$\bar{v} = \bigcup_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} v(r, i^r_{l^r+1}, j^r_{l^r+1}) \tag{5.448}$$

which justifies constraint (5.426).

Similarly where v is as in (5.433), and some $s+1$'st ordered index set $\{i^{s+1}_1, \ldots, j^{s+1}_{l^{s+1}}\}$, $l^{s+1} \leq h$, matches in its first $2(l^{s+1}-1)$ coordinates to, say, the r'th ordered index set ($r \in \{1, \ldots, s\}$), then by the reasoning used in (5.446),

$$T(\{i^r_1, \ldots, j^r_{l^r}\}) \subseteq T(\{i^{s+1}_1, \ldots, j^{s+1}_{l^{s+1}-1}\}) \tag{5.449}$$

which implies that

$$v = T(\{i^1_1, \ldots, j^1_{l^1}\}, \ldots \{i^s_1, \ldots, j^s_{l^s}\}\{i^{s+1}_1, \ldots, j^{s+1}_{l^{s+1}-1}\}). \tag{5.450}$$

Applying now the rule (5.441) to the $s+1$'st ordered index set of the expression (5.450), we thus obtain that for each $i^{s+1}_{l^{s+1}}$, we can partition v as

$$v = \bigcup_{j^{s+1}=1}^{t_{l^{s+1}}(\cdot)} T(\{i^1_1,\ldots,j^1_{l^1}\},\ldots\{i^s_1,\ldots,j^s_{l^s}\}\{i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}-1},i^{s+1}_{l^{s+1}},j^{s+1}_{l^{s+1}}\}) = \quad (5.451)$$

$$\bigcup_{j^{s+1}_{l^{s+1}}=1}^{t_{l^{s+1}}(\cdot)} v(s+1,\{i^{s+1}_1,j^{s+1}_1,\ldots,i^{s+1}_{l^{s+1}},j^{s+1}_{l^{s+1}}\}), \quad (5.452)$$

and this justifies constraint (5.432). The additional requirements, $l^r > l^{s+1}-1$, and condition (ii), are required to ensure that the sets

$$v(s+1,\{i^{s+1}_1,j^{s+1}_1,\ldots,i^{s+1}_{l^{s+1}},j^{s+1}_{l^{s+1}}\}) \quad (5.453)$$

do not violate conditions (1) or (2), so that there will indeed be such columns in the matrix.

- Constraint (5.423) is justified by the fact that under the stated circumstances, $v \cap M_{f(i^r_1,j^r_1,\ldots,i^r_h,j^r_h)} = v$.
- For the case of $v = P$, we considered P to be of the form

$$v = T(\{i^1_1,j^1_1,\ldots,i^1_{l^1},j^1_{l^1}\},\{i^2_1,j^2_1,\ldots,i^2_{l^2},j^2_{l^2}\},\ldots\{i^s_1,j^s_1,\ldots,i^s_{l^s},j^s_{l^s}\}) \quad (5.454)$$

but with $s=0$. This identification is sensible as the meaning of "$s=0$" is that v is the element of the partition of P obtained where none of the R_i are partitioned, and this means that $v=P$. Moreover for each $i^1_1 = 1,\ldots,m_1$, P can be partitioned as

$$P = \bigcup_{j^1_1=1}^{t_1(i^1_1)} T(i^1_1,j^1_1) \quad (5.455)$$

so the identification of the case $s=0$ with P is consistent with the definition for the case $s>0$, and constraint (5.432), which enforces that for each $i^1_1 = 1,\ldots,m_1$, we must have $l^1 = 1$, and

$$x^P = \sum_{j^1_1=1}^{t_1(i^1_1)} x^{T(\{i^1_1,j^1_1\})} \quad (5.456)$$

is valid. □

Example: Let us consider once more the case

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \quad (5.457)$$

which, as above, can be represented as

$$P = \bigcap_{i_1=1}^{1} \bigcup_{j_1=1}^{2} \bigcap_{i_2=1}^{2} \bigcup_{j_2=1}^{2} Y_{f(i_1,j_1,i_2,j_2)} \tag{5.458}$$

where f maps

$$(1,1,1,1) \to 1,\ (1,1,1,2) \to 2,\ (1,1,2,1) \to 3, \ldots, (1,2,2,2) \to 8. \tag{5.459}$$

The depth first partitioning algorithm at level 2 forms a matrix with rows indexed by $P, Y_1, \ldots, Y_8$, $N_1, \ldots, N_8$, and with columns indexed by P (the empty tuple), and the sets

$$T(\{1,1\}), T(\{1,2\}), \tag{5.460}$$

$$T(\{1,1,1,1\}), T(\{1,1,1,2\}), T(\{1,1,2,1\}), T(\{1,1,2,2\}), \tag{5.461}$$

$$T(\{1,2,1,1\}), T(\{1,2,1,2\}), T(\{1,2,2,1\}), T(\{1,2,2,2\}), \tag{5.462}$$

$$T(\{1,1,1,1\},\{1,1,2,1\}), T(\{1,1,1,1\},\{1,1,2,2\}), \tag{5.463}$$

$$T(\{1,1,1,2\},\{1,1,2,1\}), T(\{1,1,1,2\},\{1,1,2,2\}), \tag{5.464}$$

$$T(\{1,2,1,1\},\{1,2,2,1\}), T(\{1,2,1,1\},\{1,2,2,2\}), \tag{5.465}$$

$$T(\{1,2,1,2\},\{1,2,2,1\}), T(\{1,2,1,2\},\{1,2,2,2\}). \tag{5.466}$$

The subcolumn of the P column indexed by $Y_1, \ldots, Y_8$ is the vector that we are seeking to ensure belongs to $Conv(P)$. The (5.426) constraints are

$$x^{T(\{1,1\})} = x^{T(\{1,1,1,1\})} + x^{T(\{1,1,1,2\})} \tag{5.467}$$

$$x^{T(\{1,1\})} = x^{T(\{1,1,2,1\})} + x^{T(\{1,1,2,2\})} \tag{5.468}$$

$$x^{T(\{1,2\})} = x^{T(\{1,2,1,1\})} + x^{T(\{1,2,1,2\})} \tag{5.469}$$

$$x^{T(\{1,2\})} = x^{T(\{1,2,2,1\})} + x^{T(\{1,2,2,2\})}. \tag{5.470}$$

The (5.432) constraints are

$$x^P = x^{T(\{1,1\})} + x^{T(\{1,2\})} \tag{5.471}$$

$$x^{T(\{1,1,1,1\})} = x^{T(\{1,1,1,1\},\{1,1,2,1\})} + x^{T(\{1,1,1,1\},\{1,1,2,2\})} \tag{5.472}$$

$$x^{T(\{1,1,1,2\})} = x^{T(\{1,1,1,2\},\{1,1,2,1\})} + x^{T(\{1,1,1,2\},\{1,1,2,2\})} \tag{5.473}$$

$$x^{T(\{1,1,2,1\})} = x^{T(\{1,1,2,1\},\{1,1,1,1\})} + x^{T(\{1,1,2,1\},\{1,1,1,2\})} \tag{5.474}$$

$$x^{T(\{1,1,2,2\})} = x^{T(\{1,1,2,2\},\{1,1,1,1\})} + x^{T(\{1,1,2,2\},\{1,1,1,2\})} \tag{5.475}$$

$$x^{T(\{1,2,1,1\})} = x^{T(\{1,2,1,1\},\{1,2,2,1\})} + x^{T(\{1,2,1,1\},\{1,2,2,2\})} \tag{5.476}$$

$$x^{T(\{1,2,1,2\})} = x^{T(\{1,2,1,2\},\{1,2,2,1\})} + x^{T(\{1,2,1,2\},\{1,2,2,2\})} \tag{5.477}$$

$$x^{T(\{1,2,2,1\})} = x^{T(\{1,2,2,1\},\{1,2,1,1\})} + x^{T(\{1,2,2,1\},\{1,2,1,2\})} \tag{5.478}$$

$$x^{T(\{1,2,2,2\})} = x^{T(\{1,2,2,2\},\{1,2,1,1\})} + x^{T(\{1,2,2,2\},\{1,2,1,2\})}. \tag{5.479}$$

It should be clear how to apply the other constraints. □

Lemma 5.21 *Denote the length of the set theoretic description of P, as in Definition 5.13, by Φ. For each fixed level k, the matrices U that satisfy the constraints of the depth first partitioning algorithm at level k are described by the solution set of a linear system with a number of variables and constraints that is bounded by a polynomial in Φ. Moreover, the depth first partitioning algorithm constructs this linear system in time polynomial in Φ.*

Proof: Let Θ be the number of distinct choices of

$$(i_1, j_1, i_2, j_2, \dots, i_h, j_h) \tag{5.480}$$

within the given bounds. Clearly $\Theta \leq O(\Phi)$, as the set theoretic desciption of P entails a string representing $M_{f(\cdot)}$ for each such choice. We can also assume that the length of the description is at least n (or else we could have shrunk the dimension), and clearly we can also assume that $\Phi \geq h$. The number of ordered index sets is no more than $h\Theta$, so the number of columns is $\leq (h\Theta)^k$. This bound on the number of columns would hold even if we would not enforce conditions (1) and (2), and so the number of comparisons performed in determining which sets v of the form (5.413) will have a corresponding column in U is certainly $\leq O(h^{k+1}\Theta^k)$. The number of rows as well as the number of constraints that apply within each column is $O(n)$, and the number of constraints that describe any given column as a sum of other columns is clearly no more than $O(\Theta)$. □

In analyzing the behavior of the algorithm, the following lemma frees us from the obligation to check whether or not expressions of the form

$$v = \bigcap_{u=1}^{s} T(\{i_1^u, j_1^u, \dots, i_{l^u}^u, j_{l^u}^u\}) \tag{5.481}$$

satisfy conditions (1) and (2).

Lemma 5.22 *Given an arbitrary expression*

$$v = \bigcap_{u=1}^{s} T(\{i_1^u, j_1^u, \dots, i_{l^u}^u, j_{l^u}^u\}), \tag{5.482}$$

or equivalently, given an unordered collection of s ordered $2l^r$-tuples of positive integers ($r = 1, \dots, s$),

$$\{i_1^1, \dots, j_{l^1}^1\}, \dots, \{i_1^s, \dots, j_{l^s}^s\} \tag{5.483}$$

with all $l^r \leq h$, $i^r_u \leq m_u(i^r_1, \ldots, j^r_{u-1})$, $j^r_u \leq t_u(i^r_1, \ldots, j^r_{u-1}, i^r_u)$, and given a matrix U satisfying the algorithm constraints at level $k \geq s$, define the vector x^v by

$$x^v = \begin{cases} U^v : \text{ there is a } v \text{ column in } U \\ U^{\bar{v}} : v \text{ violates condition (1) alone} \\ 0 : v \text{ violates condition (2)} \end{cases} \tag{5.484}$$

where $\bar{v}$ is the expression obtained by discarding from the expression v all ordered index sets that are subsets of other ordered index sets. Then constraints (5.420), (5.421), (5.422), (5.426) (for r such that $l^r < h$), and (5.423) all hold for all expressions x^v, regardless of whether or not they correspond to columns of U. Moreover, where $k > s$ and

$$\{i^{s+1}_1, \ldots, j^{s+1}_{l^{s+1}-1}\} \tag{5.485}$$

matches with the corresponding elements of one of the ordered index sets that defines v, then for any

$$\{i^{s+1}_1, \ldots, j^{s+1}_{l^{s+1}-1}, i^{s+1}_{l^{s+1}}\}, \tag{5.486}$$

constraint (5.432) holds for all expressions x^v, regardless of whether or not they correspond to columns of U.

Proof: The statement regarding constraints (5.420), (5.421), (5.422), and (5.423) is clear from the definitions. Let v be as in the statement of the lemma, and let r be such that $l^r < h$. We will show that (5.426) holds for every vector x^v defined in the statement of the lemma, regardless of whether or not U has columns for the sets $v(r, i^r_{l^r+1}, j^r_{l^r+1})$. If v violates condition (2), then both sides of inequality (5.426) are zero and (5.426) is satisfied trivially. So we will assume that v does not violate condition (2). Suppose for now that no $v(r, i^r_{l^r+1}, j^r_{l^r+1})$ violates condition (2) either, then $\bar{v}$ and $\overline{v(r, i^r_{l^r+1}, j^r_{l^r+1})}$ are both columns of U. If we assume additionally that the r'th ordered index set of v is not redundant (i.e. it is not a subset of any other ordered index set), then assuming that the r'th ordered index set of v is the $\bar{r}$'th ordered index set of $\bar{v}$, then there is a column in U for the set $\bar{v}(\bar{r}, i^r_{l^r+1}, j^r_{l^r+1})$ as well (since considering that $\bar{v}$ does not violate condition (1), appending two new terms to its $\bar{r}$'th ordered index set can create no new violation of conditions (1) or (2)). Thus if we assume that the r'th ordered index set of v is indeed not redundant, then

$$\bar{v}(\bar{r}, i^r_{l^r+1}, j^r_{l^r+1}) = \overline{v(r, i^r_{l^r+1}, j^r_{l^r+1})}, \tag{5.487}$$

so by (5.426),

$$x^{\bar{v}} = \sum_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} x^{\bar{v}(\bar{r}, i^r_{l^r+1}, j^r_{l^r+1})} = \sum_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} x^{\overline{v(r, i^r_{l^r+1}, j^r_{l^r+1})}} = x^{v(r, i^r_{l^r+1}, j^r_{l^r+1})} \tag{5.488}$$

so (5.426) is satisfied in this case, regardless of whether or not there are columns in U for the sets $v(r, i^r_{l^r+1}, j^r_{l^r+1})$.

Suppose now that the r'th ordered index set is in fact redundant, i.e. there is some $r' \neq r,\ r' \leq s$ for which $i^r_1, \ldots, j^r_{l^r} = i^{r'}_1, \ldots, j^{r'}_{l^r}$, then $\bar{v}$ is defined by $\bar{s} < s$ ordered index sets, and where $i^r_{l^r+1} \neq i^{r'}_{l^r+1}$ (note that by the assumption that v does not violate (2), all ordered index sets dominating the r'th (and not equal to the r'th) must have the same value for i_{l^r+1}), the set

$$\bar{v}(\bar{s}+1, \{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, j^r_{l^r+1}\}) \tag{5.489}$$

is of the appropriate form for the application of (5.432), it violates neither (1) nor (2) (so it corresponds to a column of U), and is equal to $\overline{v(r, i^r_{l^r+1}, j^r_{l^r+1})}$. Thus applying (5.432) to $\bar{v}$ we obtain

$$x^{\bar{v}} = \sum_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} x^{\bar{v}(\bar{s}+1,\{i^r_1,\ldots,j^r_{l^r},i^r_{l^r+1},j^r_{l^r+1}\})} = \sum_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} x^{\overline{v(r,i^r_{l^r+1},j^r_{l^r+1})}} = \tag{5.490}$$

$$\sum_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} x^{v(r,i^r_{l^r+1},j^r_{l^r+1})} \tag{5.491}$$

and so (5.426) holds here as well.

Suppose now that $i^r_{l^r+1} = i^{r'}_{l^r+1}$, then $v(r, i^r_{l^r+1}, j^r_{l^r+1})$ violates condition (2) for every $j^r_{l^r+1}$ in the range of summation $1, \ldots, t_{l^r+1}(\cdot)$, except for $j^r_{l^r+1} = j^{r'}_{l^r+1}$. But this implies that

$$\sum_{j^r_{l^r+1}=1}^{t_{l^r+1}(\cdot)} x^{v(r,i^r_{l^r+1},j^r_{l^r+1})} = x^{v(r,i^r_{l^r+1},j^{r'}_{l^r+1})} = x^{\overline{v(r,i^r_{l^r+1},j^{r'}_{l^r+1})}} = x^{\bar{v}} \tag{5.492}$$

since

$$\{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}, j^{r'}_{l^r+1}\} = \{i^{r'}_1, \ldots, j^{r'}_{l^r}, i^{r'}_{l^r+1}, j^{r'}_{l^r+1}\}, \tag{5.493}$$

so in this case (5.426) is trivially satisfied by identity.

Suppose finally that some $v(r, i^r_{l^r+1}, j^r_{l^r+1})$ does violate condition (2), then there must be some $r' \leq s,\ r \neq r'$ such that

$$\{i^{r'}_1, \ldots, j^{r'}_{l^r}, i^{r'}_{l^r+1}\} = \{i^r_1, \ldots, j^r_{l^r}, i^r_{l^r+1}\}. \tag{5.494}$$

In this case constraint (5.426) again becomes an identity, as above, and is satisfied trivially.

We have thus proven that (5.426) holds for all expressions v. We will now show that (5.432) also holds for all v and all $\{i^{s+1}_1, \ldots, j^{s+1}_{l^{s+1}-1}, i^{s+1}_{l^{s+1}}\}$ that satisfy the conditions of the lemma. We will continue to assume that v does not violate condition (2), as if it does then (5.432) is satisfied trivially. Suppose first that the index set $\{i^{s+1}_1, \ldots, j^{s+1}_{l^{s+1}-1}, i^{s+1}_{l^{s+1}}\}$

matches to the first $2(l^{s+1}-1)+1$ elements of some other, say the r'th, ordered index set of v. As above, in this case we have

$$\sum_{j^{s+1}_{l^{s+1}}=1}^{t_{l^{s+1}}(\cdot)} x^{v(s+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})} = \sum_{j^{s+1}_{l^{s+1}}=1}^{t_{l^{s+1}}(\cdot)} x^{v(s+1,i^r_1,\ldots,j^r_{l^{s+1}-1},i^r_{l^{s+1}},j^{s+1}_{l^{s+1}})} = \tag{5.495}$$

$$x^{v(s+1,i^r_1,\ldots,j^r_{l^{s+1}-1},i^r_{l^{s+1}},j^r_{l^{s+1}})} = x^{\overline{v(s+1,i^r_1,\ldots,j^r_{l^{s+1}-1},i^r_{l^{s+1}},j^r_{l^{s+1}})}} = x^{\bar{v}} = x^v \tag{5.496}$$

and so (5.432) is trivially satisfied by identity.

So suppose that the index set $\{i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}-1},i^{s+1}_{l^{s+1}}\}$ does not match to the first $2(l^{s+1}-1)+1$ elements of any other ordered index set of v. Thus no set $v(s+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})$ violates condition (2), and moreover the index set $\{i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}-1},i^{s+1}_{l^{s+1}},j^{s+1}_{l^{s+1}}\}$ is not redundant for any set $v(s+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})$. So there is a column in U corresponding to the set

$$\overline{v(s+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})} \tag{5.497}$$

and the set (5.497) has an ordered index set $\{i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}}\}$. Let us say that the set $\bar{v}$ obtained by discarding the redundant ordered index sets of v is defined by $\bar{s}\leq s$ of the ordered index sets of v. Then there is a column in U for $\bar{v}$, and

$$\bar{v}(\bar{s}+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}}) = \overline{v(s+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})}. \tag{5.498}$$

Thus by 5.432,

$$x^v = x^{\bar{v}} = \sum_{j^{s+1}_{l^{s+1}}=1}^{t_{l^{s+1}}(\cdot)} x^{\bar{v}(\bar{s}+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})} = \sum_{j^{s+1}_{l^{s+1}}=1}^{t_{l^{s+1}}(\cdot)} x^{\overline{v(s+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})}} = \tag{5.499}$$

$$\sum_{j^{s+1}_{l^{s+1}}=1}^{t_{l^{s+1}}(\cdot)} x^{v(s+1,i^{s+1}_1,\ldots,j^{s+1}_{l^{s+1}})}, \tag{5.500}$$

so (5.432) holds in this case as well. □

Before we move on to address the pitch k constraints, we will prove a batch of results that describe how and why the algorithm works. Our first two results will show that the algorithm guarantees that the column x^P of U can be decomposed as the sum of vectors x^v. The first result parallels the decomposition of P discussed in Lemma 5.17, and it shows that there are an exponentially large number of such decompositions. The second result, which parallels the decomposition of P discussed in Lemma 5.18, only yields one decomposition of x^P, but we will see that it is a useful decomposition. We will then show that a sufficiently detailed decomposition of x^P into vectors x^v, together with some very obvious constraints on the vectors x^v (arising from the fact that these vectors are meant to represent measures of sets $v \cap M_{f(\cdot)}$), is sufficient to guarantee that $(x^P[Y_1], \ldots, x^P[Y_n]) \in Conv(P)$. That result will then show that once the algorithm reaches a level at which it implies such a detailed decomposition, then we will be guaranteed that indeed $(x^P[Y_1], \ldots, x^P[Y_n]) \in Conv(P)$. This will thus give us our first termination criterion for this algorithm. In the next section we will derive another termination criterion.

Lemma 5.23 *Given a matrix U satsfying the algorithm constraints at level k, the P'th column, x^P, of U satisfies*

$$x^P = \sum_{j_1^1=1}^{t_1(I_1^1)} \cdots \sum_{j_h^1=1}^{t_h(I_1^1,\ldots,j_{h-1}^1,I_h^1)} \cdots \sum_{j_1^k=1}^{t_1(I_1^k)} \cdots \sum_{j_h^k=1}^{t_h(I_1^k,\ldots,j_{h-1}^k,I_h^k)} x^{T(\{I_1^1,\ldots,j_h^1\},\ldots,\{I_1^k,\ldots,j_h^k\})} \tag{5.501}$$

for all choices of k indexing families for P, $\mathcal{I}^r = \{I_1^r, \ldots, I_h^r\}$, $r = 1, \ldots, k$, where we define the vectors $x^{T(\{I_1^1,\ldots,j_h^1\},\ldots,\{I_1^k,\ldots,j_h^k\})}$ as in Lemma 5.22. □

Lemma 5.23, which parallels Lemma 5.17, follows directly from the definition of the algorithm and repeated application of (5.619) (which will be shown later). Lemma 5.23 represents a kind of depth first decomposition of x^P. Of greater interest for our present purposes, however, will be the following breadth first decomposition of x^P, paralleling Lemma 5.18.

Lemma 5.24 *Let $r \leq h$; let the j-indexing families of functions $\mathcal{J}^r = \{J_1(\cdot), \ldots, J_r(\cdot)\}$ be defined as in Lemma 5.18, and let $i(\mathcal{J}^r)$ also be as defined in Lemma 5.18. Then where the level of the algorithm is sufficiently high so that all of the terms in the expression (5.502) are defined (as per Lemma 5.22), then the column vectors x^P satisfy*

$$x^P = \sum_{\mathcal{J}^r} x^{\bigcap_{i(\mathcal{J}^r)} T(\{i_1,\ldots,J_r\})} \tag{5.502}$$

where the sum is taken over all familes $\mathcal{J}^r$, and the intersection is taken over all elements of $i(\mathcal{J}^r)$.

Proof: Define the binary operator $\odot$ on the vectors x^v defined in Lemma 5.22 by

$$x^w \odot x^v = x^{w \cap v} \tag{5.503}$$

wherever these expressions are defined as per Lemma 5.22. Let v be of the form

$$v = \bigcap_{u=1}^{s} T(\{i_1^u, j_1^u, \ldots, i_{l^u}^u, j_{l^u}^u\}) \tag{5.504}$$

and let w be of the form

$$w = \bigcap_{u=s+1}^{\sigma} T(\{i_1^u, j_1^u, \ldots, i_{l^u}^u, j_{l^u}^u\}) \tag{5.505}$$

and say that the r'th ordered index set of v, according to the lexicographical ordering, is the r''th ordered index set of $v \cap w = \bigcap_{u=1}^{\sigma} T(\{i_1^u, j_1^u, \ldots, i_{l^u}^u, j_{l^u}^u\})$ according to the lexicographical ordering. Observe now that by (5.426),

$$x^w \odot x^v = x^w \odot \left(\sum_{j_{l^r+1}^r=1}^{t_{l^r+1}(\cdot)} x^{v(r,i_{l^r+1}^r,j_{l^r+1}^r)} \right) \tag{5.506}$$

and also by (5.426),

$$x^w \odot x^v = x^{\bigcap_{u=1}^{\sigma} T(\{i_1^u,j_1^u,\ldots,i_{l^u}^u,j_{l^u}^u\})} = \sum_{j_{l^r+1}^r=1}^{t_{l^r+1}(\cdot)} x^{(w\cap v)(r',i_{l^r+1}^r,j_{l^r+1}^r)} = \tag{5.507}$$

$$\sum_{j_{l^r+1}^r=1}^{t_{l^r+1}(\cdot)} x^{w\cap(v(r,i_{l^r+1}^r,j_{l^r+1}^r))} = \sum_{j_{l^r+1}^r=1}^{t_{l^r+1}(\cdot)} x^w \odot x^{v(r,i_{l^r+1}^r,j_{l^r+1}^r)} \tag{5.508}$$

and so we conclude that

$$x^w \odot \left(\sum_{j_{l^r+1}^r=1}^{t_{l^r+1}(\cdot)} x^{v(r,i_{l^r+1}^r,j_{l^r+1}^r)} \right) = \sum_{j_{l^r+1}^r=1}^{t_{l^r+1}(\cdot)} x^w \odot x^{v(r,i_{l^r+1}^r,j_{l^r+1}^r)}. \tag{5.509}$$

By similar reasoning we also obtain

$$x^w \odot \left(\sum_{j_{l^{s+1}}^{s+1}=1}^{t_{l^{s+1}}(\cdot)} x^{v(s+1,i_1^{s+1},\ldots,j_{l^{s+1}}^{s+1})} \right) = \sum_{j_{l^{s+1}}^{s+1}=1}^{t_{l^{s+1}}(\cdot)} x^w \odot x^{v(s+1,i_1^{s+1},\ldots,j_{l^{s+1}}^{s+1})}. \tag{5.510}$$

Thus the operator $\odot$ is distributive over all sums of the form (5.426) and (5.432). Thus

$$x^P = x^P \odot x^P = \left(\sum_{j_1(1)=1}^{t_1(1)} x^{T(\{1,j_1(1)\})} \right) \odot \left(\sum_{j_1(2)=1}^{t_1(2)} x^{T(\{2,j_1(2)\})} \right) = \tag{5.511}$$

$$\sum_{j_1(1)=1}^{t_1(1)} \left(\left(\sum_{j_1(2)=1}^{t_1(2)} x^{T(\{2,j_1(2)\})} \right) \odot x^{T(\{1,j_1(1)\})} \right) = \tag{5.512}$$

$$\sum_{j_1(1)=1}^{t_1(1)} \sum_{j_1(2)=1}^{t_1(2)} x^{T(\{1,j_1(1)\})}) \odot x^{T(\{2,j_1(2)\})}. \tag{5.513}$$

Repeating, we obtain,

$$x^P = \bigodot_{i_1=1}^{m_1} x^P = \bigodot_{i_1=1}^{m_1} \sum_{j_1=1}^{t_1(i_1)} x^{T(\{i_1,j_1\})} = \tag{5.514}$$

$$\sum_{j_1(1)=1}^{t_1(1)} \cdots \sum_{j_1(m_1)=1}^{t_1(m_1)} \left(x^{T(\{1,j_1(1)\})} \odot \cdots \odot x^{T(\{m_1,j_1(m_1)\})} \right) = \tag{5.515}$$

$$\sum_{J_1(i_i)} \bigodot_{i(J_1)} x^{T(\{i_1,J_1(i_1)\})}. \tag{5.516}$$

where the sum is taken over all functions $J_1(i_1)$ such that $J_1(i_1) \in \{1, \ldots, t_1(i_1)\}$ for each $i_1 = 1, \ldots, m_1$, and where $i(J_1)$ is the set of numbers i_1 that belong to the domain of the function J_1 (so in this case $i(J_1) = \{1, \ldots, m_1\}$).

Note now that similarly

$$x^{T(\{i_1,j_1\})} = \bigodot_{i_2=1}^{m_2(i_1,j_1)} x^{T(\{i_1,j_1\})} = \bigodot_{i_2=1}^{m_2(i_1,j_1)} \sum_{j_2=1}^{t_2(i_1,j_1,i_2)} x^{T(\{i_1,j_1,i_2,j_2\})} \tag{5.517}$$

and repeating the argument we obtain

$$x^P = \bigodot_{i_1=1}^{m_1} \sum_{j_1=1}^{t_1(i_1)} \cdots \bigodot_{i_r=1}^{m_r(\cdot)} \sum_{j_r=1}^{t_r(\cdot)} x^{T(\{i_1,\ldots,j_r\})}. \tag{5.518}$$

By the distributivity of $\odot$ as per (5.509) and (5.510) and by adapting the argument used in the proof of the statement in Lemma 5.12 that $W = W^{\mathcal{J}}$, it is not hard to see that the expression (5.518) is equal to

$$\sum_{\mathcal{J}^r} \bigodot_{i(\mathcal{J}^r)} x^{T(\{i_1,\ldots,J_r\})} = \tag{5.519}$$

$$\sum_{\mathcal{J}^r} x^{\bigcap_{i(\mathcal{J}^r)} T(\{i_1,\ldots,J_r\})}. \ \square \tag{5.520}$$

Example: Consider again the example

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.521}$$

represented as

$$P = \bigcap_{i_1=1}^{1} \bigcup_{j_1=1}^{2} \bigcap_{i_2=1}^{2} \bigcup_{j_2=1}^{2} Y_{f(i_1,j_1,i_2,j_2)} \tag{5.522}$$

where f maps

$$(1,1,1,1) \to 1, \ (1,1,1,2) \to 2, \ (1,1,2,1) \to 3, \ldots, (1,2,2,2) \to 8. \tag{5.523}$$

Let $r = h = 2$. Then as we saw earlier in our discussion of this example (after the proof of Lemma 5.18), there are eight possible families of functions $\mathcal{J}^r$, and thus Lemma 5.24 claims that

$$x^P = \sum_{\mathcal{J}^r} x^{\bigcap_{i(\mathcal{J}^r)} T(i_1,\ldots,J_r)} = \tag{5.524}$$

$$x^{T(\{1,1,1,1\},\{1,1,2,1\})} + x^{T(\{1,1,1,1\},\{1,1,2,2\})} +$$

$$x^{T(\{1,1,1,2\},\{1,1,2,1\})} + x^{T(\{1,1,1,2\},\{1,1,2,2\})} +$$

$$x^{T(\{1,2,1,1\},\{1,2,2,1\})} + x^{T(\{1,2,1,1\},\{1,2,2,2\})} +$$

$$x^{T(\{1,2,1,2\},\{1,2,2,1\})} + x^{T(\{1,2,1,2\},\{1,2,2,2\})}. \tag{5.525}$$

Applying constraints (5.467) - (5.479) we have

$$x^P = x^{T(\{1,1\})} + x^{T(\{1,2\})} = \tag{5.526}$$

$$x^{T(\{1,1,2,1\})} + x^{T(\{1,1,2,2\})} + x^{T(\{1,2,1,1\})} + x^{T(\{1,2,1,2\})} = \tag{5.527}$$

$$x^{T(\{1,1,2,1\},\{1,1,1,1\})} + x^{T(\{1,1,2,1\},\{1,1,1,2\})} +$$

$$x^{T(\{1,1,2,2\},\{1,1,1,1\})} + x^{T(\{1,1,2,2\},\{1,1,1,2\})} +$$

$$x^{T(\{1,2,1,1\},\{1,2,2,1\})} + x^{T(\{1,2,1,1\},\{1,2,2,2\})} +$$

$$x^{T(\{1,2,1,2\},\{1,2,2,1\})} + x^{T(\{1,2,1,2\},\{1,2,2,2\})}. \tag{5.528}$$

Recalling that the order in which the ordered index sets appear is irrelevant, so that, for example, $x^{T(\{1,1,2,1\},\{1,1,1,1\})} = x^{T(\{1,1,1,1\},\{1,1,2,1\})}$, the reader can check that expressions (5.525) and (5.528) are indeed equal. □

We will now establish the conditions under which the decomposibility of x^P will imply that the vector $(x^P[Y_1], \ldots, x^P[Y_n]) \in Conv(P)$.

Lemma 5.25 *If $P \subseteq \{0,1\}^n$ can be written as a* disjoint *union*

$$P = \bigcup_{j=1}^{t} \bar{Q}_j \tag{5.529}$$

where

$$\bar{Q}_j = \bigcap_{i \in V_j} M_i \tag{5.530}$$

where each $V_j \subseteq \{1', 1'', \ldots, n', n''\}$, then where for each vector $(x_0, \ldots, x_n) \in R^{n+1}$, we write $x_{l'} = x_l,\ x_{l''} = x_0 - x_l,\ l = 1, \ldots, n$, we have

$$Conv(P) = \{x \in R^n : \textit{ for each } j = 1, \ldots, t, \textit{ there exist vectors}$$

$$x^j \in R^n \textit{ and real numbers } x_0^j \textit{ with}$$

1. $x_0^j \geq x_i^j \geq 0,\ i = 1', 1'', \ldots, n', n''$

2. $x_i^j = x_0^j, \ \forall i \in V_j$

3. $\sum_{j=1}^t (x_0^j, x^j) = (1, x)\}$

Proof: Any point $x \in Conv(P)$ can be decomposed in this manner since, by our comments at the beginning of the section, x, when construed as $(x[Y_1^P], \ldots, x[Y_n^P])$, is consistent with a measure χ on $\mathcal{P}$ satisfying $\chi[P] = 1$. Thus (by the disjointness of the union), $(1, x) = (\chi[P], \chi[Y_1^P], \ldots, \chi[Y_n^P])$ can be decomposed into the sum of the partial sum vectors

$$(\chi^{\bar{Q}_j}[P], \chi^{\bar{Q}_j}[Y_1^P], \ldots, \chi^{\bar{Q}_j}[Y_n^P]), \tag{5.531}$$

and setting each $(x_0^j, x^j) = (\chi^{\bar{Q}_j}[P], \chi^{\bar{Q}_j}[Y_1^P], \ldots, \chi^{\bar{Q}_j}[Y_n^P])$, it is clear that conditions (1) - (3) of the lemma will all be satisfied. Conversely, note that since

$$\bar{Q}_j = \{x \in \{0,1\}^n : x_i = 1, \ \forall i \in V_j\}, \tag{5.532}$$

the convex hull of $\bar{Q}_j$ is just its linear relaxation

$$Conv(\bar{Q}_j) = \{x \in [0,1]^n : x_i = 1, \ \forall i \in V_j\}. \tag{5.533}$$

Thus any (x_0^j, x^j) that satisfies the first two conditions satisfies that either $x_0^j = 0$, or

$$x^j / x_0^j \in Conv(\bar{Q}_j) \subseteq Conv(P). \tag{5.534}$$

Thus if vectors (x_0^j, x^j) exist that satisfy all three conditions, then

$$x = \sum_{x_0^j \neq 0} x^j = \sum_{x_0^j \neq 0} (x_0^j (x^j / x_0^j)) \in Conv(P) \tag{5.535}$$

since $\sum_{x_0^j \neq 0} x_0^j = 1$ by (3). □

As is evident from its proof, one way to interpret Lemma 5.25 is as an application of the fact that the convex hull of P can be characterized as those vectors that can be written as the scaled sum of vectors each of which belong to the convex hull of some element of a partition of P. The significance of this fact is obviously that the convex hull of the elements of the partition may be easy to characterize, as is the case in the lemma. In this way, this can be seen as a disjunctive programming result, and in fact the more general version of this lemma, which will be presented as Theorem 5.28, is very similar to a theorem proved by Balas ([B74], [B79]) in the context of disjunctive programming. Nevertheless, as was noted earlier, partial summation embraces more general variables and measure theoretic constraints, and this is what will allow us to refine the procedure and to extend it to a breadth first version (as will be described later), as well as to make use of positive semidefiniteness.

Lemma 5.25 shows that by repeatedly partitioning P until eventually each set in the partition is of the form $\bigcap M_i$, we can obtain the convex hull of P. Our algorithm, however, only partitions sets of the form $R_{\{\cdot\}}$, and it never partitions the sets of the form $Q^c_{\{\cdot\}}$, and so it will never, in general (where $h > 2$) *completely* partition P in the manner of Lemma 5.25.

It should be noted however, that we could have partitioned the sets $Q^c_{\{\cdot\}}$ as well by the rule

$$Q^c = \left(\bigcap_i R_i\right)^c = \bigcup_i^{\text{disjoint}} \left(\bigcap_{j=1}^{i-1} R_j \cap R_i^c\right), \tag{5.536}$$

and since we already know how to partition R_i, and $R_i^c = \bigcap_j Q^c_{i,j}$, we could have repeated the procedure until we eventually would have achieved a decomposition as in Lemma 5.25. The variant of the algorithm that this suggests will be referred to as a *"Complete Partitioning Algorithm"*. (It should be noted, however, that even the complete partitioning variant of the algorithm does not partition completely in the manner of the Sherali Adams operators, as it does not, in general, partition down to the atomic level.)

The following theorem will show that a similar result to Lemma 5.25 will hold even for the original version of the algorithm. It would be interesting nevertheless to compare the actual performance of the original algorithm with that of its complete partitioning variant. The original algorithm, as stated, actually completely ignores the fact that the sets it works with are subsets of sets of the form $Q^c_{\{\cdot\}}$. We will show later, however, that even if we do not attempt to partition completely by partitioning the $Q^c_{\{\cdot\}}$, there are sensible constraints that can be imposed to approximate the effect of partitioning the $Q^c_{\{\cdot\}}$ sets.

For the purposes of the following theorem, as above in Lemma 5.25, given any vector $(x_0, \ldots, x_n) \in R^{n+1}$, for each $l = 1, \ldots, n$, we will write $x_{l'} = x_l$ and $x_{l''} = x_0 - x_l$.

Theorem 5.26 *Let P be as in Definition 5.13 and for each j-indexing family of functions $\mathcal{J}$ for P, define*

$$V_{\mathcal{J}} = \{i \in \{1', 1'', \ldots, n', n''\} : \exists (i_1, \ldots, i_h) \in i(\mathcal{J}) \text{ s.t. } f(i_1, \ldots, J_h) = i\}. \tag{5.537}$$

The convex hull of P is the set of $x \in R^n$ such that for each $\mathcal{J}$, there exist vectors $(x_0^{\mathcal{J}}, x^{\mathcal{J}}) \in R^{n+1}$ for which

1. $x_0^{\mathcal{J}} \geq x_i^{\mathcal{J}} \geq 0,\ i = 1', 1'', \ldots, n', n''$

2. $x_i^{\mathcal{J}} = x_0^{\mathcal{J}}, \ \forall i \in V_{\mathcal{J}}$

3. $\sum_{\mathcal{J}} (x_0^{\mathcal{J}}, x^{\mathcal{J}}) = (1, x)$.

Proof: Recall first that by Corollary 5.20, P can be written as the disjoint union

$$P = \bigcup_{\mathcal{J}} \left(\bigcap_{i(\mathcal{J})} \left(\bigcap_{l=1}^{h} \bigcap_{\bar{j}_l=1}^{J_l-1} (Q_{i_1,J_1,\ldots,i_{l-1},J_{l-1},i_l,\bar{j}_l})^c \right) \cap \bar{Q}_{\mathcal{J}} \right) \tag{5.538}$$

where

$$\bar{Q}_{\mathcal{J}} = \bigcap_{i(\mathcal{J})} M_{f(i_1,J_1,\ldots,i_h,J_h)}). \tag{5.539}$$

Thus, as in the proof of Lemma 5.25, it is easy to see that for any $x \in Conv(P)$ there must be a decomposition as described in the theorem. Conversely, suppose that vectors $x^{\mathcal{J}}$ exist satisfying the conditions of the theorem. We will show that each $x^{\mathcal{J}}$ must be in $Conv(P)$, which will imply that $x \in Conv(P)$, proving the theorem (as in the proof of Lemma 5.25). From the proof of Lemma 5.25 we already know that for any such $x^{\mathcal{J}}$, either $(x_0^{\mathcal{J}}, x^{\mathcal{J}}) = 0$, or

$$x^{\mathcal{J}}/x_0^{\mathcal{J}} \in Conv(\bar{Q}_{\mathcal{J}}) \tag{5.540}$$

so if we can show that $\bar{Q}_{\mathcal{J}} \subseteq P$ then, (as in the proof of Lemma 5.25,) the theorem will be proven. By Lemma 5.12, we have

$$P = \bigcup_{\mathcal{J}} \bigcap_{i(\mathcal{J})} M_{f(\cdot)} = \bigcup_{\mathcal{J}} \bar{Q}_{\mathcal{J}} \Rightarrow \tag{5.541}$$

$$\bar{Q}_{\mathcal{J}} \subseteq P. \ \Box \tag{5.542}$$

Corollary 5.27 *Let*

$$\bar{\Theta} = \max_{\mathcal{J}} |i(\mathcal{J})| \tag{5.543}$$

and observe that $\bar{\Theta}$ is bounded from above by the number of distinct tuples $(i_1, i_2, \ldots, i_h)$ that can be chosen within the given ranges. Then the subcolumn $(x^P[Y_1], \ldots, x^P[Y_n])$ of the column x^P of any matrix U that satisfies the algorithm constraints as any level $k \geq \bar{\Theta}$, will belong to $Conv(P)$.

Proof: At level $k \geq \bar{\Theta}$, vectors $x^{\bigcap_{i(\mathcal{J})} T(\{i_1,j_1,\ldots,i_h,j_h\})}$ are defined (as per Lemma 5.22) for every j-indexing family of functions $\mathcal{J}$ for P. For each $\mathcal{J}$, define $(x_0^{\mathcal{J}}, x^{\mathcal{J}}) \in R^{n+1}$ by

$$x_0^{\mathcal{J}} = x^{\bigcap_{i(\mathcal{J})} T(\{i_1,j_1,\ldots,i_h,j_h\})}[P] \tag{5.544}$$

$$x_l^{\mathcal{J}} = x^{\bigcap_{i(\mathcal{J})} T(\{i_1,j_1,\ldots,i_h,j_h\})}[Y_l], \ l = 1, \ldots, n. \tag{5.545}$$

Recall the notation $x^v[P] = x^v_0$, $x^v[Y_l] = x^v_{l'}$, $x^v[N_l] = x^v_{l''}$, and recall that algorithm constraint (5.423) requires that given $\mathcal{J} = \{J_1(\cdot), \ldots, J_h(\cdot)\}$, for each $r \in \{1', 1'', \ldots, n', n''\}$ such that $f(i_1, J_1, \ldots, i_h, J_h) = r$ for some $(i_1, \ldots, i_h) \in i(\mathcal{J})$, we must have

$$x_r^{\bigcap_{i(\mathcal{J})} T(\{i_1, j_1, \ldots, i_h, j_h\})} = x_0^{\bigcap_{i(\mathcal{J})} T(\{i_1, j_1, \ldots, i_h, j_h\})}. \tag{5.546}$$

Thus by algorithm constraints (5.420), (5.421), (5.422), and (5.423), conditions (1) and (2) of Theorem 5.26 are satisfied by all vectors $(x_0^{\mathcal{J}}, x^{\mathcal{J}})$. Moreover by Lemma 5.24,

$$x^P = \sum_{\mathcal{J}} x^{\bigcap_{i(\mathcal{J})} T(\{i_1, j_1, \ldots, i_h, j_h\})}, \tag{5.547}$$

so (since $x^P[P] = 1$) condition (3) is met as well. Thus Theorem 5.26 implies that $(x^P[Y_1], \ldots, x^P[Y_n]) \in Conv(P)$. □

The idea at work in Theorem 5.26 is that it is not actually necessary to form a partition of P in order to obtain the result of Lemma 5.25. All that is needed is for the sets $\bar{Q}_j$ to each be a subset of P, and for their union to cover P. We will state this as a separate theorem. (As was noted earlier, this theorem, in a slightly different form, was proven in [B74].)

Theorem 5.28 *Given any set $P \subseteq \{0, 1\}^n$ that can be written as a (not necessarily disjoint) union $P = \bigcup_{j=1}^t W_j$, we have*

$$Conv(P) = \{x : \textit{ for each } j = 1, \ldots, t, \textit{ there exist vectors } (x_0^j, x^j) \textit{ for which}$$

1. $0 \le x_l^j \le x_0^j, \ l = 1, \ldots, n$
2. *either* $x_0^j = 0$, *or* $x^j / x_0^j \in Conv(W_j)$
3. $\sum_{j=1}^t (x_0^j, x^j) = (1, x)\}$

Proof: Form the collection

$$Q_1 = W_1, \ Q_2 = W_2 - W_1, \ Q_3 = W_3 - (W_1 \cup W_2), \ \ldots, Q_t = W_t - \bigcup_{i=1}^{t-1} W_i \tag{5.548}$$

Clearly

$$P = \bigcup_{j=1}^t W_j = \bigcup_{j=1}^t Q_j \tag{5.549}$$

and the latter union is disjoint. So for each $x \in Conv(P)$, we can construe x as the vector $(x[Y_1^P], \ldots, x[Y_n^P])$ and (by our comment at the beginning of the section) there exists a measure χ on $\mathcal{A}$ with $\chi[P] = 1$ that is consistent with $(x[Y_1^P], \ldots, x[Y_n^P])$, and naturally the partial sum measure χ^P of χ is also consistent with $(x[Y_1^P], \ldots, x[Y_n^P])$ and also satisfies $\chi^P[P] = 1$. Consider now the partial sum vectors

$$(x_0^j, x^j) := (\chi^{Q_j}[P], \chi^{Q_j}[Y_1^P], \ldots, \chi^{Q_j}[Y_n^P]) = (\chi[Q_j], \chi[Y_1^{Q_j}], \ldots, \chi[Y_n^{Q_j}]). \tag{5.550}$$

Either $x_0^j = 0$, in which case $x^j = 0$ also, or if $x^j > 0$ then x^j / x_0^j is consistent with the measure $\frac{\chi}{\chi[Q_j]}$ on $\mathcal{A}$ (defined by $\frac{\chi}{\chi[Q_j]}[q] = \frac{\chi[q]}{\chi[Q_j]}$, $\forall q \in \mathcal{A}$). So since $\frac{\chi}{\chi[Q_j]}[Q_j] = 1$, it follows from our comment at the beginning of the section that

$$x^j / x_0^j \in Conv(Q_j) \subseteq Conv(W_j). \tag{5.551}$$

The vectors (x_0^j, x^j) thus satisfy condition (2) of the theorem, and they satisfy the other two conditions by the fact that the sets Q_j are disjoint, by the definition of measures and of partial sums, and by the fact that $\chi^P[P] = 1$. As for the converse, the proof given in Lemma 5.25 suffices, as it does not assume that the union is disjoint. □

5.5 The Depth First Algorithm and the Pitch k Constraints

We will now set out to show that where P is as in Definition 5.13 and P' is as in Definition 5.14, at level k, the P'th column of U will satisfy every (homogenized) pitch $\leq k$ constraint that is valid for P'. Recall the notation $x^v[P] = x_0^v$, $x^v[Y_l] = x_{l'}^v$, $x^v[N_l] = x_{l''}^v$, and recall that the points $y \in P'$ are of the form $y = (y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''})$. We will thus show that where we denote the projection of x^P on its Y_l and N_l coordinates as $\hat{x}^P$, for any valid constraint $\alpha^T x \geq \beta$ on P' with $(\alpha, \beta) \geq 0$, and of pitch $\leq k$, the column x^P of any matrix U satisfying the algorithm constraints at level $\geq k$ will satisfy

$$\beta x_0^P \leq \alpha^T \hat{x}^P = \sum_{l=1}^{n} \alpha_{l'} x_{l'}^P + \sum_{l=1}^{n} \alpha_{l''} x_{l''}^P. \tag{5.552}$$

Note that if all $M_{f(\cdot)}$ are of the form Y_j, then it is easy to see that P is just the projection of P' on the $1', \ldots, n'$ coordinates. Thus every constraint that is valid for P will be valid for P' as well (where the constraint is applied to its $1', \ldots, n'$ coordinates). So in that case we will be guaranteed that at level k, the vector x^P will satisfy every (homogenized) pitch $\leq k$ constraint that is valid for P. Stated precisely, where we denote the projection of x^P on its Y_l coordinates as $\bar{x}^P$, and where we denote the Y_l coordinates of x^P as $x^P[Y_l] = x_l^P$, then for any valid constraint $\alpha^T x \geq \beta$ on P with $(\alpha, \beta) \geq 0$, and of pitch $\leq k$, the column

x^P of any matrix U satisfying the algorithm constraints at level $\geq k$ will satisfy in this case

$$\beta x_0^P \leq \alpha^T \bar{x}^P = \sum_{l=1}^{n} \alpha_l x_l^P. \tag{5.553}$$

The next few lemmas will characterize P and P' in standard integer programming formulation, i.e. as the set of points in $\{0,1\}^n$ that satisfy a (exponentially large) collection of linear constraints. We will see that P' is the feasible region of a set covering problem, and that it can be thought of as a relaxation of P.

Lemma 5.29 *Where P is as in Definition 5.13, then*

$$P = \{y \in \{0,1\}^n :$$

$$\sum_{j_1=1}^{t_1(I_1)} \sum_{j_2=1}^{t_2(I_1,j_1,I_2)} \cdots \sum_{j_h=1}^{t_h(I_1,j_1,I_2,j_2,\ldots,I_{h-1}j_{h-1},I_h)} y_{f(I_1,j_1,I_2,j_2\cdots,I_h,j_h)} \geq 1$$

$$\forall \text{ indexing families } \{I_1,\ldots,I_h\}\}$$

where we define $y_{j'} = y_j$, and $y_{j''} = 1 - y_{j'}$, $j = 1,\ldots,n$

Proof: This follows from Lemma 5.12 and (5.220). □

For our standard example

$$P = ((Y_1 \cup Y_2) \cap (Y_3 \cup Y_4)) \quad \cup \quad ((Y_5 \cup Y_6) \cap (Y_7 \cup Y_8)) \tag{5.554}$$

the indexing function I_1 must be $I_1 = 1$, and there are four possible functions I_2:

$$I_2(1) = 1,\ I_2(2) = 1 \tag{5.555}$$

$$I_2(1) = 1,\ I_2(2) = 2 \tag{5.556}$$

$$I_2(1) = 2,\ I_2(2) = 1 \tag{5.557}$$

$$I_2(1) = 2,\ I_2(2) = 2. \tag{5.558}$$

There are thus four possible families $\mathcal{I}$, and thus by Lemma 5.29, P is the set of points $y \in \{0,1\}^n$ that satisfy the four constraints

$$y_{f(1,1,1,1)} + y_{f(1,1,1,2)} + y_{f(1,2,1,1)} + y_{f(1,2,1,2)} = y_1 + y_2 + y_5 + y_6 \geq 1 \tag{5.559}$$

$$y_{f(1,1,1,1)} + y_{f(1,1,1,2)} + y_{f(1,2,2,1)} + y_{f(1,2,2,2)} = y_1 + y_2 + y_7 + y_8 \geq 1 \tag{5.560}$$

$$y_{f(1,1,2,1)} + y_{f(1,1,2,2)} + y_{f(1,2,1,1)} + y_{f(1,2,1,2)} = y_3 + y_4 + y_5 + y_6 \geq 1 \tag{5.561}$$

$$y_{f(1,1,2,1)} + y_{f(1,1,2,2)} + y_{f(1,2,2,1)} + y_{f(1,2,2,2)} = y_3 + y_4 + y_7 + y_8 \geq 1. \tag{5.562}$$

Lemma 5.30 *Where P and P' are as in Definitions 5.13 and 5.14 respectively, then*

$$P' = \{y \in \{0,1\}^{2n} :$$

$$\sum_{j_1=1}^{t_1(I_1)} \sum_{j_2=1}^{t_2(I_1,j_1,I_2)} \cdots \sum_{j_h=1}^{t_h(I_1,j_1,I_2,j_2,\ldots,I_{h-1}j_{h-1},I_h)} y_{f(I_1,j_1,I_2,j_2\cdots,I_h,j_h)} \geq 1 \tag{5.563}$$

$$\forall \text{ indexing families } \{I_1,\ldots,I_h\}\} \text{ for } P$$

$$y_{l'} \geq 1 - y_{l''},\ l = 1,\ldots,n\}.$$

Proof: Observe first that P' can be written as

$$P' = \bigcap_{i_1=1}^{m_1+n} \bigcup_{j_1=1}^{t_1(i_1)} \bigcap_{i_2=1}^{m_2(i_1,j_1)} \bigcup_{j_2=1}^{t_2(i_1,j_1,i_2)} \cdots \bigcap_{i_h=1}^{m_h(i_1,\ldots,j_{h-1})} \bigcup_{j_h=1}^{t_h(i_1,\ldots,j_{h-1},i_h)} Y'_{f(i_1,j_1,\cdots,i_h,j_h)} \tag{5.564}$$

where for each $l \in \{1,\ldots,n\}$, the value $i_1 = m_1 + l$ satisfies $t_2(m_1+l) = 2$, and $t_r(m_1 + l,\ldots) = m_r(m_1+l,\ldots) = 1$ for all $r = 2,\ldots,h$, and $f(m_1+l,1,1,\ldots,1) = l'$ and $f(m_1 + l,2,1,\ldots,1) = l''$. The indexing families for which $I_1 = 1,\ldots,m_1$ are exactly the same for P' as for P, and for any indexing family for which $I_1 \in \{m_1+1,\ldots,m_1+n\}$, the functions $I_2,\ldots,I_h$ are all constants with value 1. Thus there is only one indexing family $\mathcal{I} = \{I_1,\ldots,I_h\}$ for each of the n values $I_1 = m_1 + l$. Applying Lemma 5.29 to these n families, we obtain the n additional constraints

$$1 \leq \sum_{j_1=1}^{2} \sum_{j_2=1}^{1} \cdots \sum_{j_h=1}^{1} y_{f(m_1+l,j_1,1,\ldots,1)} = y_{l'} + y_{l''}.\ \square \tag{5.565}$$

Definition 5.31 *Where P is as in Definition 5.13, and where $\mathcal{I} = \{I_1,\ldots,I_h\}$ is an indexing family for P, then let*

$$A(\mathcal{I}) = \{q \in \{1',\ldots,n',1'',\ldots,n''\} :$$

$$f(I_1,j_1,I_2(j_1),j_2,\ldots,I_h(j_1,\ldots,j_{h-1}),j_h) = q \quad \text{for some } (j_1,\ldots,j_h)\} \tag{5.566}$$

and define

$$y(A(\mathcal{I})) = \sum_{q\in A(\mathcal{I})} y_q. \tag{5.567}$$

The purpose of introducing the sets $A(\mathcal{I})$ is to reduce all positive coefficients in constraints of the form (5.563) to 1. For example, if

$$P = ((Y_1 \cap N_2) \cup (Y_1 \cap Y_4)) \quad \cap \quad ((Y_3 \cap N_1) \cup (Y_2 \cap Y_4)) \tag{5.568}$$

or

$$P = \bigcap_{i_1=1}^{2} \bigcup_{j_1=1}^{2} \bigcap_{i_1=1}^{2} \bigcup_{j_2=1}^{1} M_{f(i_1,j_1,i_2,j_2)} \tag{5.569}$$

where

$$f(1,1,1,1) = 1',\ f(1,1,2,1) = 2'',\ f(1,2,1,1) = 1',\ f(1,2,2,1) = 4' \tag{5.570}$$

$$f(2,1,1,1) = 3',\ f(2,1,2,1) = 1'',\ f(2,2,1,1) = 2',\ f(2,2,2,1) = 4'. \tag{5.571}$$

Then say $I_1 = 1$ and $I_2 = 1$ as well. The equation of the form (5.563) corresponding to this indexing family is

$$1 \leq y_{f(1,1,1,1)} + y_{f(1,2,1,1)} = y_1 + y_1 = 2y_1 \tag{5.572}$$

(where we understand $y_{l'}$ to be y_l and $y_{l''}$ to be $1 - y_l$ for each $l = 1, \ldots, n$). Obviously the coefficient, 2, can be reduced to 1 while maintaining the validity of the constraint. In this case $A(I_1, I_2) = 1'$, and

$$y(A(I_1, I_2)) = y_1. \tag{5.573}$$

We thus obtain from Lemmas 5.29 and 5.30 the following lemma.

Lemma 5.32 *Where P and P' are as in Definitions 5.13 and 5.14 respectively, then*

$$P = \{y \in \{0,1\}^n :$$

$$y(A(\mathcal{I})) \geq 1,\ \forall \textit{ indexing families } \mathcal{I} \textit{ for } P\} \tag{5.574}$$

where we take $y_{l'}$ to be y_l and $y_{l''}$ to be $1 - y_l$ for each $l = 1, \ldots, n$. Equivalently,

$$P = \{y = (y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''}) \in \{0,1\}^{2n} :$$

$$y(A(\mathcal{I})) \geq 1,\ \forall \textit{ indexing families } \mathcal{I} \textit{ for } P,$$

$$y_{j'} + y_{j''} = 1,\ j = 1, \ldots, n\} \tag{5.575}$$

and

$$P' = \{y = (y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''}) \in \{0,1\}^{2n} :$$

$$y(A(\mathcal{I})) \geq 1,\ \forall \textit{ indexing families } \mathcal{I} \textit{ for } P$$

$$y_{j'} + y_{j''} \geq 1,\ j = 1, \ldots, n\}. \ \square \tag{5.576}$$

Consider again the example $P \subseteq \{0,1\}^4$ defined by

$$P = ((Y_1 \cap N_2) \cup (Y_1 \cap Y_4)) \quad \cap \quad ((Y_3 \cap N_1) \cup (Y_2 \cap Y_4)) \tag{5.577}$$

(introduced after Definition 5.31). There are eight possible families $\mathcal{I}$ corresponding to the two possible choices for I_1 (i.e. $I_1 = 1$ and $I_1 = 2$), and the four possible choices for I_2 (i.e. $\{I_2(1), I_2(2)\} = \{1,1\}$ or $\{1,2\}$ or $\{2,1\}$ or $\{2,2\}$). The eight corresponding sets $A(\mathcal{I})$ are

1. $\{1'\}$
2. $\{1', 4'\}$
3. $\{2'', 1'\}$
4. $\{2'', 4'\}$
5. $\{3', 2'\}$
6. $\{3', 4'\}$
7. $\{1'', 2'\}$
8. $\{1'', 4'\}$

(There is considerable redundancy here, but that needn't concern us.) Thus P is the set of points $y \in \{0, 1\}^4$ that satisfy

$$y_1 \geq 1,\ y_1 + y_4 \geq 1,\ (1 - y_2) + y_1 \geq 1,\ (1 - y_2) + y_4 \geq 1, \tag{5.578}$$

$$y_3 + y_2 \geq 1,\ y_3 + y_4 \geq 1,\ (1 - y_1) + y_2 \geq 1,\ (1 - y_1) + y_4 \geq 1 \tag{5.579}$$

Equivalently, P is the set of points $(y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''}) \in \{0, 1\}^8$ that satisfy

$$y_{1'} \geq 1,\ y_{1'} + y_{4'} \geq 1,\ y_{2''} + y_{1'} \geq 1,\ y_{2''} + y_{4'} \geq 1, \tag{5.580}$$

$$y_{3'} + y_{2'} \geq 1,\ y_{3'} + y_{4'} \geq 1,\ y_{1''} + y_{2'} \geq 1,\ y_{1''} + y_{4'} \geq 1 \tag{5.581}$$

$$y_{l'} + y_{l''} = 1,\ l = 1, \ldots, 4 \tag{5.582}$$

and P' is the set of points $(y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''}) \in \{0, 1\}^8$ that satisfy

$$y_{1'} \geq 1,\ y_{1'} + y_{4'} \geq 1,\ y_{2''} + y_{1'} \geq 1,\ y_{2''} + y_{4'} \geq 1, \tag{5.583}$$

$$y_{3'} + y_{2'} \geq 1,\ y_{3'} + y_{4'} \geq 1,\ y_{1''} + y_{2'} \geq 1,\ y_{1''} + y_{4'} \geq 1 \tag{5.584}$$

$$yl' + y_{l''} \geq 1,\ l = 1, \ldots, 4. \tag{5.585}$$

We have thus established that all P defined as in Definition 5.13 have a natural relaxation as the feasible region of a set covering problem. The methodology that we used therefore to prove the pitch k result for set covering problems in Subsection 5.2.2 can therefore be applied here as well. Before we come to a formal proof, however, we need two more preliminary results.

Lemma 5.33 *Consider an inequality $\alpha^T x \geq \beta$, $\alpha \geq 0$, with pitch $\pi(\alpha, \beta) = k > 0$. Let $v \in \{1, \dots, n\}$ be such that $\alpha_v > 0$, and define $\bar{\alpha}$ to be the same as α, but with $\bar{\alpha}_v$ set to zero, then*

$$\pi(\bar{\alpha}, \beta - \alpha_v) \leq k - 1. \tag{5.586}$$

Proof: Consider the following listing of the positive coordinates of α:

$$0 < \alpha_{s(1)} \leq \alpha_{s(2)} \leq \cdots \leq \alpha_{s(|support(\alpha)|)}. \tag{5.587}$$

Let us say that α_v is the u'th smallest positive coefficient of α, i.e. $v = s(u)$, then

$$\sum_{i=1}^{k-1} \bar{\alpha}_{s(i)} = \begin{cases} \sum_{i=1}^{k-1} \alpha_{s(i)} \geq \beta - \alpha_{s(k)} \geq \beta - \alpha_{s(u)} = \beta - \alpha_v : u \geq k \\ \sum_{i=1,\dots,k,\ i \neq u} \alpha_{s(i)} \geq \beta - \alpha_{s(u)} = \beta - \alpha_v : u < k \end{cases} \quad \square \tag{5.588}$$

Lemma 5.34 *Let P and P' be as in Definitions 5.13 and 5.14 respectively. Given an inequality $\alpha^T x \geq \beta$, $\alpha \geq 0$ that is valid for P', then either there exists some indexing family, $\mathcal{I} = \{I_1, \dots, I_h\}$, for P, such that*

$$A(\mathcal{I}) \subseteq support(\alpha) \tag{5.589}$$

or $\alpha^T \geq \beta$ is dominated by the inequalities

$$x_j \geq 0, \ j = 1', 1'', \dots, n', n'' \text{ and} \tag{5.590}$$

$$x_{i'} + x_{i''} \geq 1, \ i = 1, \dots, n. \tag{5.591}$$

Proof: We will show that if there is no indexing family $\mathcal{I} = \{I_1, \dots, I_h\}$ for P, such that

$$A(\mathcal{I}) \subseteq support(\alpha) \tag{5.592}$$

then either $\beta \leq 0$ (and so $\alpha^T x \geq \beta$ is dominated by $x \geq 0$), or

$$\sum_{i \in \{1,\dots,n\} : i', i'' \in support(\alpha)} \min(\alpha_{i'}, \alpha_{i''}) \geq \beta \tag{5.593}$$

in which case $x \geq 0$ and $x_{i'} + x_{i''} \geq 1, \ i = 1, \dots, n$, imply that $\alpha^T x \geq \beta$. The proof is by induction on $\pi(\alpha, \beta)$. Where $\pi(\alpha, \beta) = 0$, then the result is simple. Assume now that the lemma holds for all constraints with pitch $\leq k$, where $k \geq 0$, and consider a valid inequality (for P'), $\alpha^T x \geq \beta$, with $\pi(\alpha, \beta) = k + 1$. Suppose that there is no $A(\mathcal{I}) \subseteq support(\alpha)$. Suppose further that there is no $i \in \{1, \dots, n\}$ such that i' and i'' are both $\in support(\alpha)$. Then consider $y \in \{0, 1\}^{2n}$ defined by

$$y_j = 1 \text{ iff } j \notin support(\alpha), \ j = 1', 1'', \dots, n', n'' \tag{5.594}$$

then y satisfies all inequalities $y(A(\mathcal{I})) \geq 1$ as well as all inequalities $y_{i'} + y_{i''} \geq 1$ and nevertheless $\alpha^T y = 0$. Thus since $\alpha^T x \geq \beta$ was assumed to be valid, it must be that $\beta \leq 0$ (note that this also contradicts the assumption that $\pi(\alpha, \beta) = k+1 > 0$). Assume now that that there is some $i \in \{1, \ldots, n\}$ such that i' and i'' both belong to *support*(α), (w.l.o.g.) with $\alpha_{i'} \leq \alpha_{i''}$, and define $\bar{\alpha}$ to be the same as α but with $\bar{\alpha}_{i'} = 0$. Then

$$\bar{\alpha}^T x \geq \beta - \alpha_{i'} \tag{5.595}$$

is also valid for P', and by Lemma 5.33, it is of pitch $\leq k$. If $\beta \leq \alpha_{i'}$ then (5.593) is satisfied trivially, so assume that $\beta - \alpha_{i'} > 0$. Thus by induction, since there is no $A(\mathcal{I}) \subseteq$ *support*$(\bar{\alpha})$ either, it must be that

$$\beta - \alpha_{i'} \leq \sum_{j \in \{1,\ldots,n\}: j', j'' \in support(\bar{\alpha})} \min(\bar{\alpha}_{j'}, \bar{\alpha}_{j''}) = \tag{5.596}$$

$$\sum_{j \in \{1,\ldots,n\}: j', j'' \in support(\bar{\alpha})} \min(\alpha_{j'}, \alpha_{j''}). \tag{5.597}$$

Observe now that $i', i'' \in$ *support*(α), but they are not both in *support*$(\bar{\alpha})$, and therefore

$$\sum_{j \in \{1,\ldots,n\}: j', j'' \in support(\alpha)} \min(\alpha_{j'}, \alpha_{j''}) = \tag{5.598}$$

$$\sum_{j \in \{1,\ldots,n\}: j', j'' \in support(\bar{\alpha})} \min(\alpha_{j'}, \alpha_{j''}) + \min(\alpha_{i'}, \alpha_{i''}) = \tag{5.599}$$

$$\sum_{j \in \{1,\ldots,n\}: j', j'' \in support(\bar{\alpha})} \min(\alpha_{j'}, \alpha_{j''}) + \alpha_{i'} \geq \beta \tag{5.600}$$

which proves the lemma. □

Now we are ready for our main result. Recall that by Lemma 5.22, at level w of the depth first partitioning algorithm, for every

$$v = T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots, \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.601}$$

with $s \leq w$, i.e. for every unordered collection of ordered sets of integers

$$\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots, \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\} \tag{5.602}$$

with $s \leq w$, and with every i_u^r and j_u^r within their appropriate bounds, the column vector $x^v \in R^{2n+1}$ is defined and satisfies all of the algorithm constraints regardless of whether or not v satisfies algorithm conditions (1) and (2). Note also that for the purposes of the following theorem, we will be representing vectors x^v as $x^v = (x_0^v, \hat{x}^v)$.

Theorem 5.35 *Let P and P' be as in Definitions 5.13 and 5.14 respectively. Then each vector x^v defined in Lemma 5.22 with respect to level w of the depth first partitioning algorithm, where*

$$v = T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots, \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.603}$$

and $s \leq w$, satisfies

$$\alpha^T \hat{x}^v \geq \beta x_0^v \tag{5.604}$$

for every constraint $\alpha^T x \geq \beta$ that is valid for P', with $\pi(\alpha, \beta) \leq w - s$. In particular, where $v = P$, so that $s = 0$ and $x^v = x^P = (1, \hat{x}^P)$, the column x^P satisfies that for every valid constraint $\alpha^T x \geq \beta$ on P' of pitch $\leq w$, we have $\alpha^T \hat{x}^P \geq \beta$.

Note: For ease of expression, given a valid constraint $\alpha^T x \geq \beta$ for P' (so that $\alpha \in R^{2n}$), we will allow ourselves to say that a vector x^v, which is in R^{2n+1}, "satisfies" $\alpha^T x \geq \beta$, if $\alpha^T \hat{x}^v \geq \beta x_0^v$.

Proof: The proof will be by induction on the pitch of the valid constraints. The pitch 0 constraints are all dominated by the inequalities $x^v \geq 0$ which hold for every v, so every column x^v satisfies all pitch 0 constraints. If a valid constraint $\alpha^T x \geq \beta$ is dominated by the inequalities $x \geq 0$ and $x_{j'} + x_{j''} \geq 1$ then it is also trivial that every column x^v satisfies $\alpha^T x \geq \beta$, as the $x_{j'} + x_{j''} \geq x_0$ inequalities are themselves dominated by the inequalities $x_{j'} + x_{j''} = x_0$, which hold for every column v. Assume now that $w - 1 \geq k \geq 0$ and every valid constraint $\alpha^T \geq \beta$ with $\pi(\alpha, \beta) \leq k$ holds for every vector x^v, for each

$$v = T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots, \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.605}$$

with $s \leq w$, and for which $w - s \geq k$. That is, assume that every valid constraint of pitch $\leq k$ holds for each vector for which the theorem asserts that it will hold. Consider now an arbitrary valid inequality on P', $\alpha^T x \geq \beta$ of pitch $k+1$, where $\alpha^T x \geq \beta$ is not dominated by the inequalities $x \geq 0$ and $x_{j'} + x_{j''} \geq 1$, and consider an arbitrary x^v, where

$$v = T(\{i_1^1, j_1^1, \ldots, i_{l^1}^1, j_{l^1}^1\}, \{i_1^2, j_1^2, \ldots, i_{l^2}^2, j_{l^2}^2\}, \ldots, \{i_1^s, j_1^s, \ldots, i_{l^s}^s, j_{l^s}^s\}) \tag{5.606}$$

with $w - s \geq k+1$. If we can prove that x^v must satisfy $\alpha^T x \geq \beta$ then the theorem will be proven.

By Lemma 5.34, there must be some indexing family, $\mathcal{I} = \{I_1, \ldots, I_h\}$, for P, such that $A(\mathcal{I}) \subseteq \textit{support}(\alpha)$. Assume first that for some $r \in \{1, \ldots, s\}$, $l^r = h$, $i_1^r = I_1$, and

$$i_l^r = I_l(j_1^r, \ldots, j_{l-1}^r), \ \forall l = 2, \ldots, h. \tag{5.607}$$

Then $f(i_1^r, \ldots, j_h^r) \in A(\mathcal{I}) \subseteq \mathit{support}(\alpha)$, and by 5.423, we also have

$$x^v_{f(i_1^r,\ldots,j_h^r)} = x_0^v. \tag{5.608}$$

Observe now that the constraint $\bar{\alpha}^T x \geq \beta - \alpha_{f(i_1^r,\ldots,j_h^r)}$, where $\bar{\alpha}$ is the same as α but with $\bar{\alpha}_{f(i_1^r,\ldots,j_h^r)} = 0$, is also valid for P'. Moreover, by Lemma 5.33, its pitch is strictly less than $\pi(\alpha, \beta) = k + 1$, so by induction x^v satisfies

$$\bar{\alpha}^T \hat{x}^v \geq (\beta - \alpha_{f(i_1^r,\ldots,j_h^r)}) x_0^v \Rightarrow \tag{5.609}$$

$$\alpha^T \hat{x}^v = \bar{\alpha}^T \hat{x}^v + \alpha_{f(i_1^r,\ldots,j_h^r)} x^v_{f(i_1^r,\ldots,j_h^r)} = \tag{5.610}$$

$$\bar{\alpha}^T \hat{x}^v + \alpha_{f(i_1^r,\ldots,j_h^r)} x_0^v \geq \beta x_0^v \tag{5.611}$$

and so x^v satisfies $\alpha^T x \geq \beta$. So assume that there is no $r \in \{1, \ldots, s\}$ for which $i_1^r = I_1$, and $i_l^r = I_l(j_1^r, \ldots, j_{l-1}^r)$, $\forall l = 2, \ldots, h$, and let $r \in \{1, \ldots, s\}$ be such that $i_l^r = I_l(j_1^r, \ldots, j_{l-1}^r)$, $\forall l = 1, \ldots, u$ for maximal $u < h$. (If there is no $r \in \{1, \ldots, s\}$ with $i_1^r = I_1$ then let $u = 0$.) Let $l^{s+1} = u + 1$. Define

$$j_l^{s+1} = j_l^r, \ l = 1, \ldots, u. \tag{5.612}$$

Thus by construction, the ordered set

$$\{I_1, j_1^{s+1}, \ldots, I_u(j_1^{s+1}, \ldots, j_{u-1}^{s+1}), j_u^{s+1}\} \tag{5.613}$$

coincides with $\{i_1^r, j_1^r, \ldots, i_u^r, j_u^r\}$, but the ordered set

$$\{I_1, j_1^{s+1}, \ldots, I_u(j_1^{s+1}, \ldots, j_{u-1}^{s+1}), j_u^{s+1}, I_{u+1}(j_1^{s+1}, \ldots, j_u^{s+1})\} \tag{5.614}$$

does not coicide with any $\{i_1^r, j_1^r, \ldots, i_u^r, j_u^r, i_{u+1}^r\}$. We can thus apply (5.432) to obtain

$$x^v = \sum_{j_{u+1}^{s+1}=1}^{t_{u+1}(I_1, j_1^{s+1}, \ldots, I_u, j_u^{s+1}, I_{u+1})} x^{v(s+1, \{I_1, j_1^{s+1}, \ldots, I_{u+1}, j_{u+1}^{s+1}\})}. \tag{5.615}$$

We can now further decompose x^v by applying (5.426) to each element of the sum (5.615), yielding $x^v =$

$$\sum_{j_{u+1}^{s+1}=1}^{t_{u+1}(\cdot)} \sum_{j_{u+2}^{s+1}=1}^{t_{u+2}(I_1, j_1^{s+1}, \ldots, I_{u+1}, j_{u+1}^{s+1}, I_{u+2})} x^{v(s+1, \{I_1, j_1^{s+1}, \ldots, I_{u+1}, j_{u+1}^{s+1}, I_{u+2}, j_{u+2}^{s+1}\})} \tag{5.616}$$

since

$$[v(s+1, \{I_1, \ldots, j_{u+1}^{s+1}\})](s+1, I_{u+2}, j_{u+2}^{s+1}) = \tag{5.617}$$

$$v(s+1,\{I_1,\ldots,j_{u+1}^{s+1},I_{u+2},j_{u+2}^{s+1}\}). \tag{5.618}$$

Applying (5.426) repeatedly in the same manner, we eventually obtain

$$x^v = \sum_{j_{u+1}^{s+1}=1}^{t_{u+1}(\cdot)} \cdots \sum_{j_h^{s+1}=1}^{t_h(\cdot)} x^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})}. \tag{5.619}$$

By (5.423), each column $x^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})}$ satisfies

$$x_{f(I_1,\ldots,j_h^{s+1})}^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})} = x_0^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})}. \tag{5.620}$$

Moreover $f(I_1,\ldots,j_h^{s+1}) \in A(\mathcal{I}) \subseteq support(\alpha)$, and therefore where $\bar{\alpha}$ is the same as α but with $\bar{\alpha}_{f(I_1,\ldots,j_h^{s+1})} = 0$, then

$$\bar{\alpha}^T x \geq \beta - \alpha_{f(I_1,\ldots,j_h^{s+1})} \tag{5.621}$$

is valid for P' and of pitch $\leq k$ as above. Thus, by induction, each vector $x^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})}$ must satisfy the corresponding constraint (5.621), and therefore by (5.620), it must satisfy $\alpha^T x \geq \beta$ as well, as above, i.e.

$$\alpha^T \hat{x}^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})} \geq \beta x_0^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})}. \tag{5.622}$$

Since, by (5.619), x^v is a sum of vectors $x^{v(s+1,\{I_1,\ldots,j_h^{s+1}\})}$, we conclude that $\alpha^T \hat{x}^v \geq \beta x_0^v$ as well. □

The following corollary now follows from the theorem and from Lemma 5.6 and Corollary 5.8.

Corollary 5.36 *The convex hull of P is obtained by the algorithm no later than level $2n-1$, and where each $f(\cdot)$ maps into $\{1',\ldots,n'\}$ then it is obtained no later than level $n-1$.*

It is also worth mentioning that there is a particular significance in the fact that the partial sum vectors x^v are also guaranteed to satisfy all constraints of up to a certain pitch. It is often the case that a high pitch constraint $\alpha^T x \geq \beta$ is not implied by valid lower pitch constraints of the form $(\alpha^i)^T x \geq \beta^i$, but is nevertheless implied by constraints of the form $(\alpha^i)^T x^{v^i} \geq \beta^i$ for some collection of sets $\{v^i\}$ (where the x^{v_i} are partial sum vectors) and low pitch valid constraints $(\alpha^i)^T x \geq \beta^i$. Recall that such was the case for the odd hole, odd antihole and odd wheel constraints for the stable set problem (i.e. imposing valid low pitch constraints for P on partial sums of x was sufficient to guarantee that x will satisfy the odd hole, odd antihole and odd wheel constraints), as was shown in chapter 4.

5.6 Breadth First Partitioning

The breadth first version of the algorithm represents essentially the same idea as the depth first algorithm, but the expansions will be done in a different order. An important difference between them, however, is that at a given level k, the breadth first algorithm may never define the set P in terms of explicit intersections of sets Y or N, but only in terms of sets Q and R. Thus it will be necessary to append constraints that relate the sets Q and R to the sets Y and N. In any case this is something that ought to be done for the depth first algorithm as well, as was noted earlier, and the methodology to be presented here can be applied there as well. One other point to note is that the breadth first algorithm at level k will only run in polynomial time if the terms $m_1, \ldots, t_k(\cdot)$ are all constants.

Breadth First Partitioning Algorithm, Level $k \geq 1$

Step 1 : Form the Matrix

Where P is as in Definition 5.13, form a matrix U with rows indexed by P and each of the sets $R_{\{\cdot\}}$ and $Q_{\{\cdot\}}$ (including the sets $Q_{i_1,\ldots,j_h} = M_{f(i_1,\ldots,j_h)}$), and append a row for each $Q^c_{\{\cdot\}}$ as well. Form a column for each of the sets

$$v = T(\{i^1_1, j^1_1, \ldots, i^1_{l^1}, j^1_{l^1}\}, \{i^2_1, j^2_1, \ldots, i^2_{l^2}, j^2_{l^2}\}, \ldots, \{i^s_1, j^s_1, \ldots, i^s_{l^s}, j^s_{l^s}\}) \tag{5.623}$$

for all unordered collections of s ordered $2l^r$-tuples of positive integers ($r = 1, \ldots, s$),

$$\{i^1_1, \ldots, j^1_{l^1}\}, \ldots, \{i^s_1, \ldots, j^s_{l^s}\} \tag{5.624}$$

with all $i^r_u \leq m_u(i^r_1, \ldots, j^r_{u-1})$, $j^r_u \leq t_u(i^r_1, \ldots, j^r_{u-1}, i^r_u)$ for all $1 \leq l^r \leq k$, and all $s \geq 0$ for which the following conditions hold:

1. No ordered set $\{i^r_1, j^r_1, \ldots, i^r_{l^r}, j^r_{l^r}\}, 1 \leq r \leq s$, is equal to any other ordered set $\{i^u_1, j^u_1, \ldots, i^u_{l^r}, j^u_{l^r}\},\ 1 \leq u \leq s,\ u \neq r$.

2. For each $r, r' \leq s,\ r \neq r'$,

$$\{i^r_1, j^r_1, \ldots, i^r_u\} = \{i^{r'}_1, j^{r'}_1, \ldots, i^{r'}_u\} \Rightarrow j^r_u = j^{r'}_u. \tag{5.625}$$

If $s = 0$ then the set defined by (5.623) is defined to be P, and the associated column is denoted x^P. In general, the v'th column, where v is of the form (5.623), will be denoted x^v.

Step 2 : Impose Constraints

Each column thus corresponds to either P or to the pure intersection of Q sets, Q^c sets and R sets that describes the sets of the form (5.623), as per Definition 5.15. For each column x^v of the matrix, we will write

$$\delta^Q(v) = \{\text{all indices } \{i_1, j_1, \ldots i_l, j_l\} : Q_{i_1,j_1,\ldots i_l,j_l}$$

$$\text{is an element of the intersection that describes } v \text{ in Definition 5.15}\} \tag{5.626}$$

and for $v = P$, define $\delta^Q(v) = \emptyset$. Define $\delta^{Q^c}(v)$ in a similar fashion. The set $\delta^R(v)$, however, will be defined as follows:

$$\delta^R(v) = \{\text{all indices } \{i_1, j_1, \ldots i_{l-1}, j_{l-1}, i_l\} :$$

$$\text{either } R_{i_1,j_1,\ldots i_{l-1},j_{l-1},i_l} \text{ or some } Q_{i_1,j_1,\ldots i_l,j_l}$$

$$\text{is an element of the intersection that describes } v \text{ in Definition 5.15}\} \tag{5.627}$$

and these three sets together uniquely identify v. Similarly for each row $Q_{i_1,\ldots,j_l}$ of the matrix we will write

$$\delta^Q(Q_{i_1,\ldots,j_l}) = \{i_1, \ldots, j_l\} \tag{5.628}$$

$$\delta^R(Q_{i_1,\ldots,j_l}) = \{i_1, j_1 \ldots, i_{l-1}\} \tag{5.629}$$

$$\delta^{Q^c}(Q_{i_1,\ldots,j_l}) = \emptyset \tag{5.630}$$

and for each row $(Q_{i_1,\ldots,j_l})^c$ we will write

$$\delta^{Q^c}((Q_{i_1,\ldots,j_l})^c) = \{i_1, \ldots, j_l\} \tag{5.631}$$

$$\delta^Q((Q_{i_1,\ldots,j_l})^c) = \emptyset = \delta^R((Q_{i_1,\ldots,j_l})^c) \tag{5.632}$$

and for each row $R_{1_1,j_1\ldots,i_l}$ we will write

$$\delta^R((R_{i_1,j_1\ldots,i_l}) = \{i_1, j_1 \ldots, i_l\} \tag{5.633}$$

$$\delta^Q((R_{i_1,j_1\ldots,i_l}) = \emptyset = \delta^{Q^c}((R_{i_1,j_1\ldots,i_l}). \tag{5.634}$$

For each matrix entry w, v we associate the triple

$$(\delta^Q(w,v), \delta^{Q^c}(w,v), \delta^R(w,v)) =$$

$$(\delta^Q(w) \cup \delta^Q(v), \delta^{Q^c}(w) \cup \delta^{Q^c}(v), \delta^R(w) \cup \delta^R(v). \tag{5.635}$$

For any pair of matrix entries (w, v) and (w', v') such that

$$\delta^Q(w,v) \subseteq \delta^Q(w',v'),\ \delta^{Q^c}(w,v) \subseteq \delta^{Q^c}(w',v'), \delta^R(w,v) \subseteq \delta^R(w',v') \tag{5.636}$$

write the constraint

$$x^v[w] \geq x^{v'}[w']. \tag{5.637}$$

Impose the constraint

$$x^P[P] = 1. \tag{5.638}$$

For each column x^v of U impose the constraints:

$$x^v \geq 0 \tag{5.639}$$

$$x^v[Q_{\{\cdot\}}] + x^v[Q^c_{\{\cdot\}}] = x^v[P], \ \forall Q_{\{\cdot\}} \tag{5.640}$$

$$\text{if } f(i_1, \ldots, j_h) = f(\bar{i}_1, \ldots, \bar{j}_h) \text{ then}$$

$$x^v[Q_{i_1,\ldots,j_h}] = x^v[Q_{\bar{i}_1,\ldots,\bar{j}_h}] \tag{5.641}$$

$$\text{if } f(i_1, \ldots, j_h) = j' \text{ and } f(\bar{i}_1, \ldots, \bar{j}_h) = j'' \text{ then}$$

$$x^v[Q_{i_1,\ldots,j_h}] = x^v[(Q_{\bar{i}_1,\ldots,\bar{j}_h})^c]. \tag{5.642}$$

We will relate the Q and R rows to each other via the following constraints. For each column x^v of U impose:

$$x^v[P] \leq x^v[R_{i_1}], \ i_1 = 1, \ldots, m_1 \tag{5.643}$$

$$x^v[P] \geq \sum_{i_1=1}^{m_1} x^v[R_{i_1}] - (m_1 - 1)x^v[P] \tag{5.644}$$

$$x^v[R_{i_1}] \geq x^v[Q_{i_1,j_1}], \ \forall (i_1, j_1) \tag{5.645}$$

$$x^v[R_{i_1}] \leq \sum_{j_1=1}^{t_1(i_1)} x^v[Q_{i_1,j_1}], \ \forall (i_1, j_1) \tag{5.646}$$

$$x^v[Q_{i_1,j_1}] \leq x^v[R_{i_1,j_1,i_2}], \ \forall (i_1, j_1, i_2) \tag{5.647}$$

$$x^v[Q_{i_1,j_1}] \geq \sum_{i_2=1}^{m_2(i_1,j_1)} x^v[R_{i_1,j_1,i_2}] - (m_2(i_1, j_1) - 1)x^v[P], \ \forall (i_1, j_1) \tag{5.648}$$

$$\vdots \tag{5.649}$$

$$x^v[R_{i_1,j_1,\ldots,i_h}] \geq x^v[Q_{i_1,\ldots,j_h}], \ \forall (i_1, \ldots, j_h) \tag{5.650}$$

$$x^v[R_{i_1,j_1,\ldots,i_h}] \leq \sum_{j_h=1}^{t_h(\cdot)} x^v[Q_{i_1,\ldots,j_h}], \ \forall (i_1, j_1, \ldots, i_h) \tag{5.651}$$

The partitioning constraints are the same as they were for the depth first algorithm. □

Comments on the Breadth First Algorithm:

- Observe that for each $l = 1, \ldots, n$, there must be some Y_l or N_l in the set theoretic expression that defines P (i.e. some $f(i_1, \ldots, j_h) \in \{l', l''\}$ so that $M_{f(i_1,\ldots,j_h)}$ is either Y_l or N_l) or else P can be equivalently recast in a lower dimension. Thus considering that for each $Q_{\{\cdot\}}$ set (including $M_{f(\cdot)}$), there is a row for both $Q_{\{\cdot\}}$ and $Q^c_{\{\cdot\}}$, without loss of generality we can assume that there is a row for each Y_l, $l = 1, \ldots, n$. There is thus an entry $x^P[Y_l]$ for each $l = 1, \ldots, n$. As usual, each entry $x^v[q]$ of the matrix is construed by the algorithm to be the value $x[v \cap q]$ of a lifted vector x consistent with some set function χ on $\mathcal{A}$, and the vector $(x^P[Y_1], \ldots, x^P[Y_n]) = (\chi[Y_1^P], \ldots, \chi[Y_n^P])$ belongs to $Conv(P)$ iff χ can be chosen to be a measure on $\mathcal{A}$ with $\chi[P] = 1$. The constraints imposed by the algorithm are all necessity conditions for this to in fact be the case. (As with the depth first algorithm, we may equivalently think of the lifted vector x as being consistent with a set function $\bar{\chi}$ on $\mathcal{P}$ and $(x^P[Y_1], \ldots, x^P[Y_n]) = (\bar{\chi}[Y_1^P], \ldots, \bar{\chi}[Y_n^P])$ belongs to $Conv(P)$ iff $\bar{\chi}$ can be chosen to be a probability measure on $\mathcal{P}$.)

- We defined a separate row for each set $Q_{i_1,\ldots,j_h} = M_{f(i_1,\ldots,j_h)}$ and then asserted that these rows must be equal for each $i_1, \ldots, j_h$, and $\bar{i}_1, \ldots, \bar{j}_h$ for which $f(i_1, \ldots, j_h) = f(\bar{i}_1, \ldots, \bar{j}_h)$. It would have been easier to simply define one row for each of the elements $l \in \{1', 1'', \ldots, n', n''\}$ for which some $f(i_1, \ldots, j_h) = l$. With such a construction (done carefully), constraint (5.637) would have also implied constraints (5.641) and (5.642) directly. The algorithm is easier to describe, however, in the way that we have written it.

- For all row entries w, the valid constraint $x^v[w] \leq x_0^v$, which follows from the fact that for every w, v entry of the matrix, $v \cap w \subseteq v = v \cap P$, is implied by (5.637).

- Constraint (5.637) generalizes constraint (5.423)

- We have used the indices of the intersection described in Definition 5.15 to form the δ sets, but we have not made the most efficient use of these indices. In particular we have included indices for sets $R_{i_1,j_1,\ldots i_{l-1},j_{l-1}}$ that are elements of the intersection even where $Q_{i_1,j_1,\ldots i_l,j_l}$ is also an element of the intersection, despite the fact that $R_{i_1,j_1,\ldots i_{l-1},j_{l-1}} \supseteq Q_{i_1,j_1,\ldots i_l,j_l}$. For example, where $P = R_1 \cap R_2 \cap R_3$, the description as per Definition 5.15 of $T(\{1,1\},\{2,1\})$ is

$$R_2 \cap R_3 \cap Q_{1,1} \cap R_1 \cap R_3 \cap Q_{2,1} \tag{5.652}$$

which would yield

$$\delta^R(T(\{1,1\},\{2,1\}) = \{1\},\{2\},\{3\}. \tag{5.653}$$

But $R_1 \supseteq Q_{1,1}$ and $R_2 \supseteq Q_{2,1}$ so $T(\{1,1\},\{2,1\})$ can actually be written as

$$R_3 \cap Q_{1,1} \cap Q_{2,1} \tag{5.654}$$

which would yield

$$\delta^R(T(\{1,1\},\{2,1\}) = \{3\}. \tag{5.655}$$

(We could also observe that the set $T(\{1,1\},\{2,1\})$ is an element of the partition obtained by partitioning both R_1 and R_2 as per Lemma 5.2, so in principle neither R_1 nor R_2 needs to be included in the intersection.) We have moreover included indices even of sets $R_{i_1,j_1,\ldots i_{l-1},j_{l-1}}$ that do not appear in the intersection at all, so long as some set $Q_{i_1,j_1,\ldots i_l,j_l}$ does. Our reason for doing this is to make explicit that every set $Q_{i_1,j_1,\ldots i_l,j_l}$ can also be written as $Q_{i_1,j_1,\ldots i_l,j_l} \cap R_{i_1,j_1,\ldots i_{l-1},j_{l-1}}$.

- The sequence of constraints (5.643) - (5.651) follows from the fact that for every w, v entry of the matrix, $v \cap w \subseteq P$, and from the following rules, which hold for all measures χ and all measurable subsets $\{S_1, \ldots, S_l\}$ of a measurable set Ω with $\chi[\Omega] < \infty$:

$$\chi\left[\bigcup_{i=1}^{l} S_i\right] \geq \chi[S_h],\ \forall h = 1, \ldots, l \tag{5.656}$$

$$\chi\left[\bigcap_{i=1}^{l} S_i\right] \leq \chi[S_h],\ \forall h = 1, \ldots, l \tag{5.657}$$

$$\chi\left[\bigcup_{i=1}^{l} S_i\right] \leq \sum_{i=1}^{l} \chi[S_i] \tag{5.658}$$

$$\chi\left[\bigcap_{i=1}^{l} S_i\right] \geq \sum_{i=1}^{l} \chi[S_i] - (l-1)\chi[\Omega].\ \Box \tag{5.659}$$

It is easy to see that all of the constraints imposed by the algorithm are valid. We will not be discussing this algorithm in detail, but here is an example to help concretize the ideas.

Example: Consider

$$P = \bigcap_{i_1=1}^{2} \bigcup_{j_1=1}^{2} \bigcap_{i_2=1}^{2} \bigcup_{j_2=1}^{2} M_{f(i_1,j_1,i_2,j_2)} \tag{5.660}$$

with

$$f(1,1,1,1) = 1',\ f(1,1,1,2) = 3'',\ f(1,1,2,1) = 2',\ f(1,1,2,2) = 5'', \tag{5.661}$$

$$f(1,2,1,1) = 1'',\ f(1,2,1,2) = 5',\ f(1,2,2,1) = 2'',\ f(1,2,2,2) = 3'', \tag{5.662}$$

$$f(2,1,1,1) = 4',\ f(2,1,1,2) = 1'',\ f(2,1,2,1) = 3'',\ f(2,1,2,2) = 5', \tag{5.663}$$

$$f(2,2,1,1) = 1',\ f(2,2,1,2) = 5',\ f(2,2,2,1) = 4',\ f(2,2,2,2) = 3'' \tag{5.664}$$

i.e.

$$P = [((Y_1 \cup N_3) \cap (Y_2 \cup N_5)) \quad \cup \quad ((N_1 \cup Y_5) \cap (N_2 \cup N_3))] \quad \cap \tag{5.665}$$

$$[((Y_4 \cup N_1) \cap (N_3 \cup Y_5)) \quad \cup \quad ((Y_1 \cup Y_5) \cap (Y_4 \cup N_3))]. \tag{5.666}$$

The $Q_{\{\cdot\}}$ and $R_{\{\cdot\}}$ sets are thus as follows:

$$R_1 = ((Y_1 \cup N_3) \cap (Y_2 \cup N_5)) \quad \cup \quad ((N_1 \cup Y_5) \cap (N_2 \cup N_3)) \tag{5.667}$$

$$R_2 = ((Y_4 \cup N_1) \cap (N_3 \cup Y_5)) \quad \cup \quad ((Y_1 \cup Y_5) \cap (Y_4 \cup N_3)) \tag{5.668}$$

$$Q_{1,1} = (Y_1 \cup N_3) \cap (Y_2 \cup N_5) \tag{5.669}$$

$$Q_{1,2} = (N_1 \cup Y_5) \cap (N_2 \cup N_3) \tag{5.670}$$

$$Q_{2,1} = (Y_4 \cup N_1) \cap (N_3 \cup Y_5) \tag{5.671}$$

$$Q_{2,2} = (Y_1 \cup Y_5) \cap (Y_4 \cup N_3) \tag{5.672}$$

$$R_{1,1,1} = Y_1 \cup N_3 \tag{5.673}$$

$$R_{1,1,2} = Y_2 \cup N_5 \tag{5.674}$$

$$R_{1,2,1} = N_1 \cup Y_5 \tag{5.675}$$

$$R_{1,2,2} = N_2 \cup N_3 \tag{5.676}$$

$$R_{2,1,1} = Y_4 \cup N_1 \tag{5.677}$$

$$R_{2,1,2} = N_3 \cup Y_5 \tag{5.678}$$

$$R_{2,2,1} = Y_1 \cup Y_5 \tag{5.679}$$

$$R_{2,2,2} = Y_4 \cup N_3 \tag{5.680}$$

$$Y_1 = Q_{1,1,1,1} = Q_{2,2,1,1} = (Q_{1,2,1,1})^c = (Q_{2,1,1,2})^c \tag{5.681}$$

$$Y_2 = Q_{1,1,2,1} = (Q_{1,2,2,1})^c \tag{5.682}$$

$$N_3 = Q_{1,1,1,2} = Q_{1,2,2,2} = Q_{2,1,2,1} = Q_{2,2,2,2} \tag{5.683}$$

$$Y_4 = Q_{2,1,1,1} = Q_{2,2,2,1} \tag{5.684}$$

$$Y_5 = Q_{1,2,1,2} = Q_{2,1,2,2} = Q_{2,2,1,2} = (Q_{1,1,2,2})^c \tag{5.685}$$

There is a row for each of the $R_{\{\cdot\}}$ sets, and for each $Q_{\{\cdot\}}$, there is a row for $Q_{\{\cdot\}}$ and for $Q^c_{\{\cdot\}}$. At level 1 of the algorithm there will be columns for each of the sets:

$$P,\ T(\{1,1\}),\ T(\{1,2\}),\ T(\{2,1\}),\ T(\{2,2\}), \tag{5.686}$$

$$T(\{1,1\},\{2,1\}),\ T(\{1,1\},\{2,2\}),\ T(\{1,2\},\{2,1\}),\ T(\{1,2\},\{2,2\}). \tag{5.687}$$

Here is an example of the δ^Q notation:

$$\delta^Q(T(\{1,1\})) = \{1,1\} \tag{5.688}$$

since $T(\{1,1\}) = R_2 \cap Q_{1,1}$. Another example is

$$\delta^Q(T(\{1,2\},\{2,2\})) = \{1,1\},\{2,2\} \tag{5.689}$$

since $T(\{1,2\},\{2,2\}) = R_2 \cap (Q_{1,1})^c \cap Q_{1,2} \cap R_1 \cap (Q_{2,1})^c \cap Q_{2,2}$. Note that the R_1 and R_2 can actually be dropped from this intersection as $Q_{1,1} \subseteq R_1$ and $Q_{2,2} \subseteq R_2$. As we indicated in the comments, a different (more efficient) implementation of the algorithm would recognize that $T(\{1,2\},\{2,2\})$ is an element of the partition formed when both R_1 and R_2 are partitioned as per Lemma 5.2, and thus neither R_1 nor R_2 would be included in the representation of $T(\{1,2\},\{2,2\})$ that is used to generate δ^R. The approach that we have taken in our implementation, however, is the reverse, and thus for example,

$$\delta^R(Q_{1,1}) = \{1\} \tag{5.690}$$

and therefore

$$\delta^R(T(\{1,1\})) = \{1\},\{2\}. \tag{5.691}$$

Here are some examples of the δ^{Q^c} notation:

$$\delta^{Q^c}(Q_{2,1}, T(\{1,1\})) = \emptyset \tag{5.692}$$

$$\delta^{Q^c}(T(\{1,2\},\{2,2\})) = \{1,1\},\{2,1\} \tag{5.693}$$

We will now give two examples that illustrate how the various constraints work together.

Illustration 1: Note first that

$$x^{T(\{1,2\})}[Q_{1,1}] = \chi[T(\{1,2\}) \cap Q_{1,1}] \tag{5.694}$$

for a set function χ consistent with x. But $T(\{1,2\}) \cap Q_{1,1} = \emptyset$, so we would like to have $x^{T(\{1,2\})}[Q_{1,1}] = 0$. We will now show how the algorithm constraints guarantee that this will in fact be the case. Note that

$$\delta^Q(Q_{1,1}^c, T(\{1,2\})) = \delta^Q(P, T(\{1,2\})) \tag{5.695}$$

and

$$\delta^R(Q_{1,1}^c, T(\{1,2\})) = \delta^R(P, T(\{1,2\})) \tag{5.696}$$

and since $Q^c_{1,1}$ is an element of the intersection that defines $x^{T(\{1,2\})}$, we also have

$$\delta^{Q^c}(Q^c_{1,1}, T(\{1,2\})) = \delta^{Q^c}(P, T(\{1,2\})). \tag{5.697}$$

Thus (5.637) implies that

$$x^{T(\{1,2\})}[Q^c_{1,1}] = x^{T(\{1,2\})}[P] \tag{5.698}$$

and this implies, by (5.640), that

$$x^{T(\{1,2\})}[Q_{1,1}] = 0. \tag{5.699}$$

Illustration 2: As another example of how the constraints enforce set theoretic relationships, consider

$$x^{T(\{1,1\})}[Q_{2,1}] = \chi[R_2 \cap Q_{1,1} \cap Q_{2,1}] = \chi[Q_{1,1} \cap Q_{2,1}] \tag{5.700}$$

and consider

$$x^{T(\{2,1\})}[R_2] = \chi[R_1 \cap Q_{2,1} \cap R_2] = \chi[R_1 \cap Q_{2,1}]. \tag{5.701}$$

Thus since $Q_{1,1} \subseteq R_1$, if the set function χ is to be measure, then we must have $\chi[Q_{1,1} \cap Q_{2,1}] \leq \chi[R_1 \cap Q_{2,1}]$, and thus

$$x^{T(\{1,1\})}[Q_{2,1}] \leq x^{T(\{2,1\})}[R_2]. \tag{5.702}$$

The algorithm constraints ensure that this relationship holds since

$$\delta^Q(Q_{2,1}, T(\{1,1\})) = \{1,1\}, \{2,1\} \supseteq \delta^Q(R_2, T(\{2,1\})) \tag{5.703}$$

$$\delta^R(Q_{2,1}, T(\{1,1\})) = \{1\}, \{2\} = \delta^R(R_2, T(\{2,1\})) \tag{5.704}$$

$$\delta^{Q^c}(Q_{2,1}, T(\{1,1\})) = \emptyset = \delta^{Q^c}(R_2, T(\{2,1\})) \tag{5.705}$$

and thus inequality (5.702) follows from (5.637). It is interesting to note that even had we defined $\delta^R(v)$ in the same fashion as $\delta^Q(v)$ and $\delta^{Q^c}(v)$, i.e. only counting the indices $\{i_1.j_1, \ldots, i_l\}$ of sets $R_{i_1.j_1,\ldots,i_l}$ that actually appear in the intersection that defined v in Definition 5.15, the algorithm constraints would still be sufficient to establish (5.702). The partitioning constraint

$$x^{T(\{1,1\})} = x^{T(\{1,1\},\{2,1\})} + x^{T(\{1,1\},\{2,2\})} \tag{5.706}$$

implies that

$$x^{T(\{1,1\})}[Q_{2,1}] = x^{T(\{1,1\},\{2,1\})}[Q_{2,1}] + x^{T(\{1,1\},\{2,2\})}[Q_{2,1}] \tag{5.707}$$

and

$$x^{T(\{1,1\},\{2,2\})}[Q_{2,1}] = 0 \tag{5.708}$$

by the reasoning used in Illustration 1. Thus

$$x^{T(\{1,1\})}[Q_{2,1}] = x^{T(\{1,1\},\{2,1\})}[Q_{2,1}]. \tag{5.709}$$

Now even if $\delta^R(v)$ only includes the indices of the R sets that actually appear in the intersection that defined v, we still have

$$\delta^R(Q_{2,1}, T(\{1,1\},\{2,1\})) = \{1\},\{2\} = \delta^R(R_2, T(\{2,1\})) \tag{5.710}$$

and, as above,

$$\delta^{Q^c}(Q_{2,1}, T(\{1,1\},\{2,1\})) = \emptyset = \delta^{Q^c}(R_2, T(\{2,1\})) \tag{5.711}$$

and

$$\delta^Q(Q_{2,1}, T(\{1,1\},\{2,1\})) = \{1,1\},\{2,1\} \supseteq \delta^Q(R_2, T(\{2,1\})) \tag{5.712}$$

and thus, as above, inequality (5.702) follows from (5.637). □

Again, as with the depth first partitioning algorithm, we could also describe complete partitioning variants, but in this implementation we have in any case already taken care not to ignore the Q^c sets.

Our main objective in introducing this second algorithm is to show that there are potentially many ways to partition, and we need not always partition down to intersections of sets Y_i and N_i. Rather we can partition down to more complicated sets and then use measure theortic constraints to relate these to measures of the sets Y_i. The following chapter will take this observation further by introducing some completely different partitioning strategies.

Chapter 6

Common Factor Algorithms

6.1 Introduction

The fundamental idea that makes partial summation (or disjunctive programming, for that matter) useful is that subsets of the feasible region that are smaller and more highly structured may be easier to characterize than the feasible region as a whole. Thus if the entire feasible region can be covered by such subsets then this better characterization of the individual subsets may translate into a better characterization of the feasible region as a whole. The algorithms described in the previous chapter partition P in a methodical manner that will eventually characterize $Conv(P)$ completely. The algorithm to be described in this chapter also accomplishes this goal (though it is applicable only to a much narrower class of sets P), but in a more interesting way.

In this chapter we will be dealing with sets $P \subseteq \{0,1\}^n$ of the form

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} M_j \tag{6.1}$$

where $A_i \subseteq \{1', 1'', \ldots, n', n''\}$, and for each $l = 1, \ldots, n$, $M_{l'} = Y_l = \{y \in \{0,1\}^n : y_l = 1\}$ and $M_{l''} = N_l = Y_l^c$. As was noted in the previous chapter, P can also be represented as

$$P = \{y \in \{0,1\}^n : y(A_i) \geq 1,\ i = 1, \ldots, n\} \tag{6.2}$$

where we define $y(A_i) = \sum_{j \in A_i} y_j$ and where for each $l = 1, \ldots, n$ we define $y_{l'} = y_l$ and $y_{l''} = 1 - y_l$.

The basic idea underlying the algorithm to follow is to partition P into parts that make the specific linear constraints that we are given most effective. Specifically, note that the constraints of the form

$$y(A_i) \geq 1,\ i = 1, \ldots, m \tag{6.3}$$

become maximally effective, in the sense that they are convex hull defining, when there is no overlap between the index sets A_i (we will prove this formally later). Speaking loosely, one way to eliminate overlapping variables from a system of inequalities is to assign them values. The algorithm will therefore consider the subsets of P defined by assigning particular $0,1$ values to all overlapping variables for various subsets of the constraints.

Example: If P is the set of $0,1$ solutions to the system of constraints

$$x_1 + x_2 + x_3 + x_6 \geq 1 \tag{6.4}$$

$$x_2 + x_3 + x_4 + x_7 \geq 1 \tag{6.5}$$

$$x_3 + x_4 + x_5 + x_6 \geq 1 \tag{6.6}$$

then the overlapping variables for the first and second constraints are x_2 and x_3. If we assign the values, say, $x_2 = 0$ and $x_3 = 0$ then the system of constraints becomes

$$x_1 + x_6 \geq 1 \tag{6.7}$$

$$x_4 + x_7 \geq 1 \tag{6.8}$$

$$x_5 + x_6 \geq 1 \tag{6.9}$$

The first and second constraints now have no overlapping variables, and obviously the set of $0,1$ integer solutions $y \in \{0,1\}^n$ to this system (with $y_2 = 0 = y_3$) is a subset of P. □

Note that the set of $0,1$ solutions to the modified system of constraints is a subset of P. Thus if we can cover P with such sets, and if we can characterize the convex hulls of those sets (a job that has become simpler due to the elimination of the overlapping variables of at least some of the constraints), then by Theorem 5.28, we will obtain a characterization of $Conv(P)$ as well.

Obviously we cannot efficiently consider all possible $0,1$ values that may be assigned to the overlapping variables, but we will see that nice results can be obtained even if we consider only the case where values of zero are assigned to the overlapping variables, and then a limited number of other cases. We will identify two ways in which the remaining cases can be handled efficiently. Perhaps the more interesting of the two arises from the observation that for any pitch k inequality (Definition 5.1), $\alpha^T x \geq \beta$, the subset $T \subseteq P$ made up of all points for which k or more of the coordinates indexed by $support(\alpha)$ have value 1, always satisfies that $\alpha^T x \geq \beta$ for every $y \in T$. This is a complicated set to describe, and the description that we gave is not "nice", in the sense of "niceness" defined in the first section of the previous chapter, but this will nevertheless prove to be a useful subset of P.

This algorithm, like the depth first partitioning algorithm of the previous chapter, will also generate in polynomial time a relaxation of $Conv(P)$ whose feasible points satisfy all valid pitch k constraints for each fixed k. But this is not actually its most interesting feature, as the depth first partitioning algorithm will accomplish this goal (generally) faster, and for a much broader range of problems. This algorithm is interesting in that it takes advantage of the specific behavior of the constraints so as to partition in a less obvious, somewhat asymmetric fashion, and, in one of its versions, over more unusual sets. A consequence of the new methodology will be a new termination criterion that is independent of the number of variables and the number of constraints. Thus this algorithm can in principle terminate with $Conv(P)$ very quickly (in terms of m and n), which is something that cannot be said for any of the other algorithms described until now. (Those other algorithms may obtain the convex hull quickly, but they have no means of recognizing this. The Sherali-Adams type algorithms will actually never terminate until they have described a complete spanning set for $\mathcal{A}$.) We will also see that this extra structure leads to some nice results if positive semidefiniteness is to be enforced.

6.2 The Set Covering Case

We will consider first the set covering case, i.e.

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in \bar{A}_i \subseteq \{1,1' \ldots, n', n''\}} M_j \tag{6.10}$$

where there is no $l \in \{1, \ldots, n\}$ for which there are numbers $i, h \in \{1, \ldots, m\}$ such that $l' \in \bar{A}_i$ and $l'' \in \bar{A}_h$. By way of some changes of variables, we can equivalently express P in this case as

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i \subseteq \{1, \ldots, n\}} Y_j. \tag{6.11}$$

The general case

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in \bar{A}_i \subseteq \{1,1' \ldots, n', n''\}} M_j \tag{6.12}$$

for which we can have $l \in \{1, \ldots, n\}$ for which there are numbers $i, h \in \{1, \ldots, m\}$ such that $l' \in \bar{A}_i$ and $l'' \in \bar{A}_h$, will be considered later.

The following lemma, which states that where the sets $A_i \subseteq \{1, \ldots, n\}$ are mutually disjoint, then the system of inequalities $\sum_{j \in A_i} x_j \geq 1, \ i = 1, \ldots, m$ is convex hull defining, was proven earlier (Theorem 3.34), but is repeated here for convenience.

Lemma 6.1 *Consider*

$$H = \{y \in \{0,1\}^n : y(B_i) \geq 1,\ \forall i = 1,\dots,m\} \tag{6.13}$$

where $B_1,\dots,B_m$ *are disjoint subsets of the index set* $\{1,\dots,n\}$*, and* $y(B_i) := \sum_{j \in B_i} y_j$*, i.e. there are no overlapping variables. Then where*

$$\bar{H} = \{x \in [0,1]^n : x(B_i) \geq 1,\ \forall i = 1,\dots,m\} \tag{6.14}$$

we have

$$Conv(H) = \bar{H}.\ \square \tag{6.15}$$

The following statement is a direct consequence of Lemma 6.1, but we will state it explicitly for clarity.

Corollary 6.2 *Let* $P \subseteq \{0,1\}^n$ *be defined by*

$$P = \{y \in \{0,1\}^n : y(A_i) \geq 1,\ i = 1,\dots,m\} \tag{6.16}$$

where A_i *is a subset of the index set* $\{1,\dots,n\}$*, and consider the* strengthened subsystem *that is obtained by removing all overlapping variables from a particular size* k *subset of the constraints indexed by some* $r = \{r(1),\dots,r(k)\} \subseteq \{1,\dots,m\}$*,*

$$\bar{P}^r = \{y \in \{0,1\}^n : y(B_{r(i)}) \geq 1,\ i = 1,\dots,k\} \tag{6.17}$$

where

$$B_{r(i)} = A_{r(i)} - \bigcup_{j=1,\dots,k,\ j \neq i} A_{r(j)} \tag{6.18}$$

and assume that all $B_{r(i)} \neq \emptyset$*. Let* $\alpha^T x \geq \beta$ *be any inequality that is valid for* $\bar{P}^r$*. Then any* $x \in [0,1]^n$ *for which*

1. $x_j = 0,\ \forall j \in A_{r(h)} \cap A_{r(l)}$*, for any* $r, l \in \{1,\dots,k\},\ r \neq l$*, i.e. all of the overlapping variables are set to zero*

2. $x(A_{r(i)}) \geq 1,\ i = 1,\dots,k$

also satisfies $\alpha^T x \geq \beta$*.* $\square$

(Note that the expression "strengthened subsystem" as a description of $\bar{P}^r$ refers to the fact that $\bar{P}^r$ is a strengthening of the subsystem described by the k constraints $y(A_{r(i)}) \geq 1,\ i = 1,\dots,k$.)

Thus any point x that belongs to the subset of $[0,1]^n$ in which all of the overlapping variables are set to zero, and that satisfies the constraints $x(A_i) \geq 1,\ i = 1, \ldots, m$, will also satisfy all constraints that are valid for the strengthened subsystem. Recall that if χ is a (signed) measure and v is a set, then the partial sum (signed) measure χ^v is the (signed) measure for which $\chi^v[q] = \chi[v \cap q]$ for all sets q on which the (signed) measure χ is defined. Observe now that for any $x = (x[Y_1^P], \ldots, x[Y_n^P]) \in Conv(P)$, and any lifting of that x to a (signed) measure, the partial sum of the lifted x defined with respect to the set

$$C(r) = \bigcap_{j:j \text{ belongs to 2 distinct } A_{r(i)}} N_j^P \tag{6.19}$$

must satisfy $x^{C(r)}[Y_j^P] = 0$ for every overlapping coordinate j (since $x^{C(r)}[Y_j^P] = x[C(r) \cap Y_j^P] = x[\emptyset] = 0$). Thus the constraint

$$x^{C(r)}[Y_j^P] = 0 \tag{6.20}$$

is valid for each j such that j belongs to 2 distinct $A_{r(i)}$. Note also that the constraint

$$x^{C(r)}(A_i) \geq x_0^{C(r)} \tag{6.21}$$

(where we have represented $x^{C(r)}[P]$ as $x_0^{C(r)}$) is valid as well for each $i = 1, \ldots, m$ as all constraints that are valid for x are valid for all partial sums of x too (Corollary 3.67). Imposing (6.20) and (6.21) thus guarantees that for any valid constraint $\alpha^T x \geq \beta$ on the strengthened subsystem we will have $\alpha^T x^{C(r)} \geq \beta x_0^{C(r)}$.

We will now show that for any pitch k constraint, $\alpha^T x \geq \beta$, that is valid for P, there is some k or smaller sized subset of the constraints $y(A_i) \geq 1$, indexed by some $r(1), \ldots, r(s),\ s \leq k$, such that $\alpha^T x \geq \beta$ is indeed valid for $\bar{P}^r$. Imposing (6.20) and (6.21) will thus guarantee that $x^{C(r)}$ will in fact satisfy the (homogenized) pitch k constraint $\alpha^T x \geq \beta x_0$.

Theorem 6.3 *Let $P \subseteq \{0,1\}^n$ be defined as in Corollary 6.2, let $k \geq 0$ and let $\alpha^T x \geq \beta$, $\alpha \geq 0$, with $0 \leq \pi(\alpha, \beta) \leq k$, be an inequality that holds for all $y \in P$. Then there exists some subcollection*

$$\{A_{r(1)}, \ldots, A_{r(\lambda)}\},\ 0 \leq \lambda \leq k \tag{6.22}$$

such that where we define

$$B_{r(i)} = A_{r(i)} - \bigcup_{j=1,\ldots,\lambda,\ j \neq i} A_{r(j)} \tag{6.23}$$

and

$$P^\alpha = \{y \in \{0,1\}^n : y(B_{r(i)}) \geq 1,\ i = 1, \ldots, \lambda\} \tag{6.24}$$

we have

1. $A_{r(i)} \subseteq support(\alpha),\ i = 1, \ldots, \lambda$

2. $B_{r(i)} \neq \emptyset,\ i = 1, \ldots, \lambda$

3. $\alpha^T x \geq \beta$ *is valid for* P^α

Proof: Consider first the case $\alpha^T x \geq \beta$ where $\pi(\alpha, \beta) = 0$. Let $\lambda = 0$ and therefore $P^\alpha = \{0,1\}^n$. As a pitch zero constraint we must have $\beta \leq 0$, so since $\alpha \geq 0$, $\alpha^T x \geq \beta$ is indeed valid for $\{0,1\}^n$.

Assume now that the theorem holds for all valid constraints of pitch $j,\ j \leq k-1 \geq 0$, and consider a valid constraint $\alpha^T x \geq \beta > 0$ of pitch k. Note first that there must be some $A_v \subseteq support(\alpha)$, or else we could set $y_j = 0$ for all $j \in support(\alpha)$, and $y_j = 1$ everywhere else, and thereby satisfy every constraint and nevertheless have $\alpha^T y = 0$. Choose $A_v \subseteq support(\alpha)$ such that no $A_i,\ i \in \{1, \ldots, m\}$, is a proper subset of A_v. Let $v(1) \in A_v$ be the index of the minimum coefficient $\alpha_j : j \in A_v$. We will construct our strengthened subsystem in three steps. First consider the collection

$$A^\alpha = \{A_j : A_j \subseteq support(\alpha)\} \tag{6.25}$$

and note that $A_v \in A^\alpha$.

Observe that $\alpha^T x \geq \beta$ is valid for the system

$$\{y \in \{0,1\}^n : y(A_j) \geq 1, \forall A_j \in A^\alpha\} \tag{6.26}$$

(Otherwise there would be a $y \in \{0,1\}^n$ that satisfies all constraints $y(A_j) \geq 1$ for which $A_j \in A^\alpha$, but for which nevertheless $\alpha^T y < \beta$. Resetting all $y_j,\ j \notin support(\alpha)$, to 1 will maintain $\alpha^T y < \beta$ and will guarantee that y satisfies the rest of the constraints as well, which is a contradiction.)

First we will eliminate only those overlapping variables that are indexed by $A_v - \{v(1)\}$. For all $A_j \in A^\alpha - \{A_v\}$ define

$$\bar{B}_j = A_j - (A_v - \{v(1)\}) \tag{6.27}$$

Observe that $\bar{B}_j \neq \emptyset$ for any j since, by assumption, no $A_j \subset A_v$. Clearly we must still have that $\alpha^T x \geq \beta$ is valid for the system

$$\bar{P}^\alpha = \{y \in \{0,1\}^n : y(A_v) \geq 1, y(\bar{B}_j) \geq 1,\ \forall j \text{ s.t. } A_j \in A^\alpha - \{A_v\}\} \tag{6.28}$$

as this is just a strengthening of the system (6.26) defined above.

Consider now the valid constraint

$$\bar{\alpha}^T x \geq \beta - \alpha_{v(1)} \tag{6.29}$$

where $\bar{\alpha}$ is the same as α but with $\alpha_{v(1)}$ reset to zero. By Lemma 5.33 we have $\pi(\bar{\alpha}, \beta - \alpha_{v(1)}) \leq k-1$. By induction there must therefore be a subcollection of $\{A_v, \bar{B}_j : j \text{ s.t. } A_j \in A^{\alpha} - \{A_v\}\}$ that satisfies the three conditions of the theorem. Thus there must be

$$\bar{B}_{r(1)}, \ldots, \bar{B}_{r(\lambda)},\ \lambda \leq k-1 \tag{6.30}$$

each of which is in *support*$(\bar{\alpha})$ (so this excludes A_v), such that when we define

$$\hat{B}_{r(i)} = \bar{B}_{r(i)} - \bigcup_{j=1,\ldots,\lambda,\ j\neq i} \bar{B}_{r(j)} \tag{6.31}$$

then no $\hat{B}_{r(i)}$ is empty, and the constraint $\bar{\alpha}^T x \geq \beta - \alpha_{v(1)}$ is valid for the system defined by the $\hat{B}_{r(i)}$. This completes the second step; the third and final step, which is to append on A_v with all of its overlapping indices removed, follows now.

Consider the collection

$$A_{r(1)}, \ldots, A_{r(\lambda)}, A_v \tag{6.32}$$

and define $B_{r(i)}$ and B_v as per the statement of the theorem. Condition (1) is satisfied for this collection by construction, and clearly

$$B_{r(i)} = \hat{B}_{r(i)},\ i = 1, \ldots, \lambda \tag{6.33}$$

as the $\hat{B}_{r(i)}$ have already had their indices that overlap with $A_v - \{v(1)\}$ removed, and they never overlapped $v(1)$. Moreover, $v(1) \in B_v$, so $B_v \neq \emptyset$, so condition (2) is satisfied as well. Suppose now that we are given an arbitrary $y \in \{0,1\}^n$ that satisfies $y(B_v) \geq 1$ and all $y(B_{r(i)}) \geq 1$. Consider that we must have $y_j = 1$ for some $j \in B_v$, that $\alpha_j \geq \alpha_{v(1)}$, and that, since B_v and all of the $B_{r(i)}$ are disjoint, if we define $\bar{y}$ to be the same as y but with $y_j = 0$, then $\bar{y}$ still satisfies all $B_{r(i)} \geq 1$. Thus, by induction,

$$\bar{\alpha}^T \bar{y} \geq \beta - \alpha_{v(1)} \Rightarrow \tag{6.34}$$

$$\alpha^T y = \bar{\alpha}^T \bar{y} + \alpha_j y_j = \bar{\alpha}^T \bar{y} + \alpha_j \geq \beta - \alpha_{v(1)} + \alpha_j \geq \beta.\ \square \tag{6.35}$$

Definition 6.4 *Let $k \geq 0$. For every collection $\mathcal{F}_k$ of k of the A_i, define*

$$C(\mathcal{F}_k) = \bigcap_{j:j \text{ belongs to two distinct } A_i \in \mathcal{F}_k} N_j^P. \tag{6.36}$$

Thus $C(\mathcal{F}_k)$ is the subset of P in which all of the overlapping variables from the collection of k constraints defined by $\mathcal{F}_k$ are set to zero. If $k < 2$, then

$$C(\mathcal{F}_k) = P \tag{6.37}$$

i.e. the empty intersection is construed as P.

Observe that each set A_i can be thought of as representing a *forbidden configuration* in the sense that no point of P can have all of its A_i coordinates set to zero. In set theoretic notation,

$$\bigcap_{j \in A_i} N_j^P = \emptyset. \tag{6.38}$$

The set $C(\mathcal{F}_k)$ can be thought of as a kind of common factor of the forbidden configurations $\bigcap_{j \in A_i} N_j^P,\ A_i \in \mathcal{F}_k$, in the sense that for each $A_i \in \mathcal{F}_k$ we have

$$C(\mathcal{F}_k) \cap \bigcap_{j \in B_i} N_j^P = \emptyset \tag{6.39}$$

where $B_i = A_i - \bigcup_{j: A_j \in \mathcal{F}_k,\ j \neq i} A_j$.

Definition 6.5 *Let $k \geq 0$. Define $\bar{\mathcal{C}}_k$ to be the collection of all expressions $C(\mathcal{F}_j),\ 0 \leq j \leq k$, for which $\mathcal{F}_j = \{A_{r(1)}, \ldots, A_{r(j)}\}$ is such that where we define*

$$B_{r(i)} = A_{r(i)} - \bigcup_{h=1,\ldots,j, h \neq i} A_{r(h)} \tag{6.40}$$

then for all $i = 1, \ldots, j$,

$$B_{r(i)} \neq \emptyset. \tag{6.41}$$

Observe that technically this is not a collection of sets but rather of set theoretic expressions, which can be represented by the index sets of the intersections. Thus one would not double list sets with identical index sets for their intersections. For example $C(\emptyset) = P$, and $C(A_1) = P$ as well, but a listing of the members of, say, $\bar{\mathcal{C}}_2$ would not list P twice.

Example: Suppose that P is the set of points in $\{0,1\}^n$ that satisfies

$$x_1 + x_2 + x_3 \geq 1 \tag{6.42}$$

$$x_1 + x_2 + x_4 \geq 1 \tag{6.43}$$

$$x_1 + x_3 + x_4 \geq 1 \tag{6.44}$$

$$x_2 + x_3 + x_4 \geq 1. \tag{6.45}$$

Then

$$A_1 = \{1,2,3\},\ A_2 = \{1,2,4\},\ A_3 = \{1,3,4\},\ A_4 = \{2,3,4\} \tag{6.46}$$

and

$$C(\mathcal{F}_k) = P, \quad \forall k < 2 \tag{6.47}$$

$$C(\{A_1, A_2\}) = N_1^P \cap N_2^P \tag{6.48}$$

$$C(\{A_1, A_3\}) = N_1^P \cap N_3^P \tag{6.49}$$

$$C(\{A_1, A_4\}) = N_2^P \cap N_3^P \tag{6.50}$$

$$C(\{A_2, A_3\}) = N_1^P \cap N_4^P \tag{6.51}$$

$$C(\{A_2, A_4\}) = N_2^P \cap N_4^P \tag{6.52}$$

$$C(\{A_3, A_4\}) = N_3^P \cap N_4^P \tag{6.53}$$

and

$$\bar{\mathcal{C}}_2 = \{P,\ C(\{A_1, A_2\}),\ C(\{A_1, A_3\}),\ C(\{A_1, A_4\}), \tag{6.54}$$

$$C(\{A_2, A_3\}),\ C(\{A_2, A_4\}),\ C(\{A_3, A_4\})\}. \tag{6.55}$$

But

$$C(\{A_1, A_2, A_3\}) = C(\{A_1, A_2, A_4\}) = C(\{A_1, A_3, A_4\}) = C(\{A_2, A_3, A_4\}) = \tag{6.56}$$

$$N_1^P \cap N_2^P \cap N_3^P \cap N_4^P \tag{6.57}$$

and

$$A_1 - (A_2 \cup A_3) = \emptyset = A_1 - (A_2 \cup A_4) = \tag{6.58}$$

$$A_1 - (A_3 \cup A_4) = A_2 - (A_3 \cup A_4) \tag{6.59}$$

and therefore

$$\bar{\mathcal{C}}_3 = \bar{\mathcal{C}}_2 \ \Box \tag{6.60}$$

Definition 6.6 *For any expression of the form*

$$C = \bigcap_{j \in J \subseteq \{1,\ldots,n\}} N_j^P \tag{6.61}$$

define $\bar{\delta}(C) \subseteq \{1, \ldots, n\}$ *to be the index set of the intersection, i.e.*

$$\bar{\delta}(C) = J. \tag{6.62}$$

We will also define $\bar{\delta}(P)$*, corresponding to the empty intersection, as*

$$\bar{\delta}(P) = \emptyset. \tag{6.63}$$

(Note that $\bar{\delta}$ should not be confused with the δ^Q, δ^R, or δ^{Q^c} sets that were described in the previous chapter. Those sets were also index sets, but they are not related to $\bar{\delta}$.)

The following statement now follows from Theorem 6.3

Corollary 6.7 *For any valid constraint, $\alpha^T x \geq \beta$ of pitch no more than k, there exists some $C_k \in \bar{\mathcal{C}}_k$, with $\bar{\delta}(C_k) \subseteq support(\alpha)$, such that any $(x_0, x) \in [0, x_0]^n$, ($x_0 \geq 0$), for which*

1. $x_j = 0, \ \forall j \in \bar{\delta}(C_k)$
2. $x(A_i) \geq x_0, \ i = 1, \ldots, m$

also satisfies $\alpha^T x \geq \beta x_0$.

Proof: By Theorem 6.3, there exists a collection of $\lambda \leq k$ constraints, with index sets $A_{r(i)}, \ i = 1, \ldots, \lambda$ such that where we define

$$B_{r(i)} = A_{r(i)} - \bigcup_{j=1,\ldots,\lambda, \ j \neq i} A_{r(j)} \tag{6.64}$$

and

$$P^\alpha = \{y \in \{0,1\}^n : y(B_{r(i)}) \geq 1, \ i = 1, \ldots, \lambda\} \tag{6.65}$$

we have

1. $A_{r(i)} \subseteq support(\alpha), \ i = 1, \ldots, \lambda$
2. $B_{r(i)} \neq \emptyset, \ i = 1, \ldots, \lambda$
3. $\alpha^T x \geq \beta$ is valid for P^α.

Thus where we write $C = C(\{A_{r(1)}, \ldots, A_{r(\lambda)}\})$, then $C \in \bar{\mathcal{C}}_k$ and $\bar{\delta}(C) \subseteq support(\alpha)$. Thus since there are no overlapping variables in the system of constraints that defines P^α, any $x \in [0,1]^n$ for which

1. $x_j = 0, \ \forall j \in \bar{\delta}(C)$
2. $x(A_{r(i)}) \geq 1, \ i = 1, \ldots, \lambda$

must satisfy $\alpha^T x \geq \beta$. Thus if $(x_0, x) \in [0, x_0]^n$, ($x_0 \geq 0$) satisfies the conditions of the corollary, then either $x_0 = 0$, in which case the corollary clearly holds, or if $x_0 > 0$, then we have $\alpha^T (x/x_0) \geq \beta$, which implies that $\alpha^T x \geq \beta x_0$. □

Thus so long as we impose the valid constraints

$$x_j^{C^k} = 0, \quad \forall j \in \bar{\delta}(C_k) \tag{6.66}$$

and

$$x^{C^k}(A_i) \geq x_0^{C^k}, \ i = 1, \ldots, m \tag{6.67}$$

for each $C_k \in \bar{\mathcal{C}}_k$, then we will be guaranteed that for every constraint $\alpha^T x \geq \beta$ with pitch $j \leq k$, some x^{C_k} satisfies $\alpha^T x^{C_k} \geq \beta x_0^{C^k}$.

Definition 6.8 *For each $C \in \bar{\mathcal{C}}_k$, defined by*

$$C = \bigcap_{j \in \bar{\delta}(C)} N_j^P, \tag{6.68}$$

for each $r \leq |\bar{\delta}(C)|$, define $\bar{\mathcal{C}}^{-r}(C)$ to be the collection of all expressions

$$\bigcap_{j \in J'} N_j^P \cap \bigcap_{j \in J''} Y_j^P \tag{6.69}$$

for which $J' \cup J'' = \bar{\delta}(C)$, $J' \cap J'' = \emptyset$, and for which $|J''| = r$, i.e. it is the collection of sets that arise from negating (complementing) r of the elements of the intersection (6.68) defining C.

Define $\bar{\mathcal{C}}_k^{-r}$ as

$$\bar{\mathcal{C}}_k^{-r} = \bigcup_{C \in \mathcal{C}_k} \bar{\mathcal{C}}^{-r}(C). \tag{6.70}$$

Obviously, $\bar{\mathcal{C}}_k^{-0} = \bar{\mathcal{C}}_k$.

Let $r \geq 0$. For each $C \in \bar{\mathcal{C}}_k$ of the form (6.68), with $|\bar{\delta}(C)| \geq r+1$, define $C^{->r}$ to be the set of all points in P that fail to belong to $r+1$ or more of the terms of the intersection (6.68). Define $\bar{\mathcal{C}}_k^{->r}$ to be the collection $\{C^{->r} : C \in \bar{\mathcal{C}}_k\}$.

As above, the collections $\bar{\mathcal{C}}_k^{-r}$ and $\bar{\mathcal{C}}_k^{->r}$ would not be defined technically as collections of sets, but rather as collections of the indices that define those sets. Observe also that $\mathcal{C}^{-r}(P) = \emptyset$ for any $r > 0$.

Example: Where P is as in the example following Definition 6.5, so $N_1^P \cap N_2^P \in \bar{\mathcal{C}}_2$, then

$$\bar{\mathcal{C}}^{-1}(N_1^P \cap N_2^P) = \{Y_1^P \cap N_2^P, \;\; N_1^P \cap Y_2^P\} \tag{6.71}$$

and

$$(N_1^P \cap N_2^P)^{->1} = \{y \in P : y_1 = 1 = y_2\} = Y_1^P \cap Y_2^P. \tag{6.72}$$

For a more instructive example of the $C_k^{->r}$ sets, consider P defined as the set of $0,1$ solutions to the system

$$x_1 + x_2 + x_3 + x_4 + x_5 \geq 1 \tag{6.73}$$

$$x_1 + x_2 + x_3 + x_4 + x_6 \geq 1 \tag{6.74}$$

$$x_1 + x_2 + x_3 + x_5 + x_6 \geq 1 \tag{6.75}$$

$$x_1 + x_2 + x_4 + x_5 + x_6 \geq 1 \tag{6.76}$$

$$x_1 + x_3 + x_4 + x_5 + x_6 \geq 1 \tag{6.77}$$

$$x_2 + x_3 + x_4 + x_5 + x_6 \geq 1 \tag{6.78}$$

Here the set $C_2 := N_1^P \cap N_2^P \cap N_3^P \cap N_4^P$ belongs to $\bar{\mathcal{C}}_2$, and

$$(C_2)^{->1} = \{y \in P : 2 \text{ or more from among } \{y_1, y_2, y_3, y_4\} \text{ have value } 1\} = \tag{6.79}$$

$$(Y_1^P \cap Y_2^P) \ \cup \ (Y_1^P \cap Y_3^P) \ \cup \ (Y_1^P \cap Y_4^P) \cup \tag{6.80}$$

$$(Y_2^P \cap Y_3^P) \ \cup \ (Y_2^P \cap Y_4^P) \ \cup \ (Y_3^P \cap Y_4^P). \ \Box \tag{6.81}$$

We have now established that for an arbitrary pitch j constraint, $\alpha^T x \geq \beta$, $j \leq k$, so long as we introduce partial sum vectors x^{C_k} for each $C_k \in \bar{\mathcal{C}}_k$, and we constrain them by (6.66) and (6.67), there will be some x^{C_k} that satisfies $\alpha^T x^{C_k} \geq \beta x_0^{C_k}$. We would like to ensure that the full vector x also satisfies $\alpha^T x \geq \beta x_0$. But C_k is only a small subset of P, and recall that in order to ensure that x itself satisfies the constraint, we need to cover P by subsets of P *all* of whose partial sum vectors satisfy the constraint. That is, in order to guarantee that x satisfies $\alpha^T x \geq \beta x_0$, we need a *sequence* of partial sum vectors $x^{q_1}, \ldots, x^{q_t}$ such that the sets $q_1, \ldots, q_t$ partition P, so that

$$x = \sum_{i=1}^{t} x^{q_i}, \tag{6.82}$$

and such that every x^{q_i} satisfies $\alpha^T x \geq \beta x_0$.

Note now that P can indeed be partitioned by the subsets defined by all possible assignments of values to the overlapping variables. For example if the overlapping variables are y_1, y_2 and y_5, then P is partitioned by the eight subsets

$$Y_1^P \cap Y_2^P \cap Y_5^P \tag{6.83}$$

$$Y_1^P \cap Y_2^P \cap N_5^P \tag{6.84}$$

$$\vdots \tag{6.85}$$

$$N_1^P \cap N_2^P \cap N_5^P \tag{6.86}$$

corresponding to the eight possible choices of values for $y_1,\ y_2$ and y_5. We have shown that for any valid pitch j constraint, $\alpha^T x \geq \beta$, with $j \leq k$, there exists some subcollection of $\lambda \leq j$ of the $\{A_i\}$ such that the partial sum vector corresponding to the subset $C_k \in \bar{\mathcal{C}}_k$ of P defined by a choice of 0 for each overlapping variable must satisfy $\alpha^T x \geq \beta x_0$. If we can guarantee that every assignment of values to the overlapping variables will also yield a partial sum vector that obeys the constraint, then we will indeed be able to conclude that x also satisfies the constraint. But there are two problems. Firstly, if some overlapping variables are assigned a value of 1, then this assignment yields a subset C of P of the form

$$\bigcap_{j \in \bar{\delta}(C)} N_j^P \cap \bigcap_{j \in \bar{\Delta}(C)} Y_j^P. \tag{6.87}$$

But instead of constraints of the form (6.66), this assignment will yield constraints

$$x_j^C = 0, \quad j \in \bar{\delta}(C) \tag{6.88}$$

$$x_j^C = x_0^C, \quad j \in \bar{\Delta}(C), \tag{6.89}$$

and there is no guarantee that constraints (6.67) together with constraints (6.88) and (6.89) will also be sufficient to ensure that $\alpha^T x^C \geq \beta x_0^C$. A second problem is that there may be too many assignments of values to consider.

Considering the second problem first, observe that we need not consider every assignment explicitly. For any subset C of P defined by an assignment of k or more values of 1 to overlapping variables, the constraint

$$x^C(\bar{\delta}(C_k)) \geq k x_0^C \tag{6.90}$$

(where C_k is as above) is valid. One way to see this is to note that if x can be lifted to a measure, the partial sum x^C is a nonnegative linear combination of the (projected) zeta vectors of the points (the atoms, to be precise) that belong to C, all of which have k or more 1's among their $\bar{\delta}(C_j)$ coordinates. (Another way to see the validity of (6.90) is to note that a partial sum of a $\mathcal{P}$-measure with respect to a set $C \subseteq P$ is itself a measure on the subset algebra of C. Thus if x is to be $\mathcal{P}$-measure consistent, then the partial sum vector x^C (for which each q'th coordinate has value $x[C \cap q]$ as per the lifted vector x), must be consistent with a measure on the subset algebra of C. Thus where x^C refers to the vector $(x^C[Y_1^P], \ldots, x^C[Y_n^P])$, and $x_0^C = x^C[P]$, it must be that either $x_0^C = 0$ or else the vector

$$\frac{1}{x_0^C} x^C = \frac{1}{x_0^C}(x^C[Y_1^P], \ldots, x^C[Y_n^P]) = \frac{1}{x[C]}(x[Y_1^C], \ldots, x[Y_n^C]) \tag{6.91}$$

is consistent with a probability measure on the subset algebra of C, and must therefore belong to $Conv(C)$. In either case, for all constraints $\gamma^T x \geq \omega,\ \gamma \geq 0$ that are valid for C

we must have $\gamma^T x^C \geq \omega x_0^C$, and clearly as every point in C has k or more 1's among its $\bar{\delta}(C_j)$ coordinates, the constraint $x(\bar{\delta}(C_k)) \geq k$ is valid for C.)

As (6.90) holds for every subset C defined by an assignment of k or more values of 1 to overlapping variables, it must also hold for the union of these sets, namely $C_k^{>k-1}$, i.e.

$$x^{C_k^{>k-1}}(\bar{\delta}(C_k)) \geq k x_0^{C_k^{>k-1}} \tag{6.92}$$

and the following lemma will show that (6.92) implies that

$$\alpha^T x^{C_k^{>k-1}} \geq \beta x_0^{C_k^{>k-1}}. \tag{6.93}$$

Lemma 6.9 *Given any inequality $\alpha^T x \geq \beta$, with $\pi(\alpha,\beta) \leq k$, then any vector $(x_0, x) \in [0, x_0]^{n+1}$ that satisfies*

$$x(\mathit{support}(\alpha)) \geq k x_0 \tag{6.94}$$

also satisfies $\alpha^T x \geq \beta x_0$. In particular, for any intersection C of terms of the form N_j, such that $\bar{\delta}(C) \subseteq \mathit{support}(\alpha)$, any vector $(x_0, x) \in [0, x_0]^{n+1}$ that satisfies $x(\bar{\delta}(C)) \geq k x_0$ also satisfies $\alpha^T x \geq \beta x_0$.

Proof: Arrange the coordinates of α so that

$$0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{|\mathit{support}(\alpha)|},\ \alpha_j = 0,\ \forall j > |\mathit{support}(\alpha)| \tag{6.95}$$

If $x_0 = 0$ then the lemma is obvious, so suppose $x_0 > 0$. Since $x(\mathit{support}(\alpha)) \geq k x_0$ (so that $|\mathit{support}(\alpha)| \geq k$), and since the α_j are increasing ($j \leq |\mathit{support}(\alpha)|$), where we define the vector $\hat{x}$ by $\hat{x}_i = x_0,\ i = 1, \ldots, k$, and $\hat{x}_i = 0,\ i > k$, it must be that

$$\alpha^T x \geq \alpha^T \hat{x}. \tag{6.96}$$

But

$$\alpha^T \hat{x} = \sum_{j=1}^{k} \alpha_j x_0 \geq \beta x_0 \tag{6.97}$$

by the definition of pitch k constraints. □

It is therefore not necessary to consider all of the exponentially many possible assignment of values to the overlapping variables, we need only consider assignments with fewer than k values of 1, and the subset $C_k^{>k-1}$. The plan therefore is to decompose the vector x into the sum of partial sum vectors corresponding to the subsets defined by assignments of fewer than k values of 1 to the overlapping variables, and the vector $x^{C_k^{>k-1}}$. We have now seen how to ensure that the partial sum vector x^{C_k}, corresponding to the assignment of all zeroes, as well as the partial sum vector $x^{C_k^{>k-1}}$ can both be guaranteed to satisfy $\alpha^T x \geq \beta x_0$. It remains to show how to guarantee that partial sum vectors x^C, where C is the subset defined by an assignment of more than zero but less than k 1's to the overlapping variables, will also satisfy the constraint.

Note first that the case for which all of the overlapping variables are assigned value zero is in a sense the hardest case scenario for satisfying $\alpha^T x \geq \beta$. If some of the overlapping variables are known to have value 1, then since all of the overlapping variables are in $support(\alpha)$, to establish $\alpha^T x \geq \beta$ we need only establish a weaker inequality. For example if $x_1 = 1$ and $\alpha_1 > 0$ then we need only establish $\bar{\alpha}^T x \geq \beta - \alpha_1$, where $\bar{\alpha}$ is α, but with the first coordinate replaced by zero, and we have seen this already (Lemma 5.33) to be a constraint of lower pitch. Thus where, say C is the subset of P defined by an assignment of value 1 to exactly one of the overlapping variables, say y_3, then after we impose the valid constraint

$$x^C[Y_3] = x_0^C \tag{6.98}$$

if we assume that x^C satisfies all valid pitch $k-1$ constraints on P, it will follow that x^C satisfies $\alpha^T x \geq \beta x_0$ as well.

In particular, since the pitch 1 inequalities are all dominated by the m constraints $x(A_i) \geq 1,\ i = 1, \ldots, m$, if $k = 2$ then we only need require the partial sum vectors x^C, where C is the subset defined by an assignment of exactly one value of 1 to an overlapping variable, to satisfy the constraints $x^C(A_i) \geq x_0^C$ (in addition to the constraints of the form (6.98)) in order to guarantee that x^C satisfies the pitch 2 constraint $\alpha^T x \geq \beta x_0$.

Example: Consider P defined as the set of $0, 1$ solutions to the system

$$x_1 + x_2 + x_3 + x_4 + x_5 \geq 1 \tag{6.99}$$

$$x_1 + x_2 + x_3 + x_4 + x_6 \geq 1 \tag{6.100}$$

$$x_1 + x_2 + x_3 + x_5 + x_6 \geq 1 \tag{6.101}$$

$$x_1 + x_2 + x_4 + x_5 + x_6 \geq 1 \tag{6.102}$$

$$x_1 + x_3 + x_4 + x_5 + x_6 \geq 1 \tag{6.103}$$

$$x_2 + x_3 + x_4 + x_5 + x_6 \geq 1 \tag{6.104}$$

and observe that the pitch 2 inequality

$$\sum_{j=1}^{6} x_i \geq 2 \tag{6.105}$$

is valid for P. Consider now the following subset $C_2 \in \bar{\mathcal{C}}_2$ obtained by taking the "common factor" of the "forbidden configurations" defined by the first two constraints:

$$N_1^P \cap N_2^P \cap N_3^P \cap N_4^P. \tag{6.106}$$

Form partial sum vectors for C_2 and for $C_2^{->1}$ and for each of the sets

$$C_2^{-1}(1) := Y_1^P \cap N_2^P \cap N_3^P \cap N_4^P \tag{6.107}$$

$$C_2^{-1}(2) := N_1^P \cap Y_2^P \cap N_3^P \cap N_4^P \tag{6.108}$$

$$C_2^{-1}(3) := N_1^P \cap N_2^P \cap Y_3^P \cap N_4^P \tag{6.109}$$

$$C_2^{-1}(4) := N_1^P \cap N_2^P \cap N_3^P \cap Y_4^P. \tag{6.110}$$

Impose the partitioning constraints

$$x = x^{C_2} + x^{C_2^{-1}(1)} + x^{C_2^{-1}(2)} + x^{C_2^{-1}(3)} + x^{C_2^{-1}(4)} + x^{C_2^{->1}} \tag{6.111}$$

$$x_0 = x_0^{C_2} + x_0^{C_2^{-1}(1)} + x_0^{C_2^{-1}(2)} + x_0^{C_2^{-1}(3)} + x_0^{C_2^{-1}(4)} + x_0^{C_2^{->1}} \tag{6.112}$$

and the valid constraints (where the coordinate x_j is understood to mean $x[Y_j]$)

$$x_1^{C_2} = x_2^{C_2} = x_3^{C_2} = x_4^{C_2} = 0 \tag{6.113}$$

$$x_1^{C_2^{-1}(1)} = x_0^{C_2^{-1}(1)}, \text{ and } x_2^{C_2^{-1}(1)} = x_3^{C_2^{-1}(1)} = x_4^{C_2^{-1}(1)} = 0 \tag{6.114}$$

$$x_2^{C_2^{-1}(2)} = x_0^{C_2^{-1}(2)}, \text{ and } x_1^{C_2^{-1}(2)} = x_3^{C_2^{-1}(2)} = x_4^{C_2^{-1}(2)} = 0 \tag{6.115}$$

$$x_3^{C_2^{-1}(3)} = x_0^{C_2^{-1}(3)}, \text{ and } x_1^{C_2^{-1}(3)} = x_2^{C_2^{-1}(3)} = x_4^{C_2^{-1}(3)} = 0 \tag{6.116}$$

$$x_4^{C_2^{-1}(4)} = x_0^{C_2^{-1}(4)}, \text{ and } x_1^{C_2^{-1}(4)} = x_2^{C_2^{-1}(4)} = x_3^{C_2^{-1}(4)} = 0 \tag{6.117}$$

$$\sum_{j=1}^{4} x_j^{C_2^{->1}} \geq 2x_0^{C_2^{->1}} \tag{6.118}$$

and impose the valid constraints

$$x_1 + x_2 + x_3 + x_4 + x_5 \geq x_0 \tag{6.119}$$

$$x_1 + x_2 + x_3 + x_4 + x_6 \geq x_0 \tag{6.120}$$

$$x_1 + x_2 + x_3 + x_5 + x_6 \geq x_0 \tag{6.121}$$

$$x_1 + x_2 + x_4 + x_5 + x_6 \geq x_0 \tag{6.122}$$

$$x_1 + x_3 + x_4 + x_5 + x_6 \geq x_0 \tag{6.123}$$

$$x_2 + x_3 + x_4 + x_5 + x_6 \geq x_0 \tag{6.124}$$

on all partial sum vectors.

Putting together (6.119), (6.120) and (6.113) we can see that x^{C_2} satisfies

$$\sum_{j=1}^{6} x_j \geq 2x_0. \tag{6.125}$$

Putting together (6.124) and (6.114) we can see that $x^{C_2^{-1}(1)}$ satisfies (6.125). Similarly, putting together (6.123) and (6.115) we can see that $x^{C_2^{-1}(2)}$ satisfies (6.125), and similarly (6.122) and (6.116) imply that $x^{C_2^{-1}(3)}$ satisfies (6.125), and (6.121) and (6.117) imply that $x^{C_2^{-1}(4)}$ satisfies (6.125). By (6.118), the partial sum $x^{C_2^{->1}}$ also satisfies (6.125), and thus the partitioning constraints (6.111) and (6.112) imply that x satisfies (6.125) as well. □

If $k = 3$, however, then where C is the subset defined by an assignment of exactly one value of 1 to the overlapping variables, we need a way to guarantee that x^C will satisfy the pitch 2 constraints. But we now know how this may be accomplished: for each $C_2 \in \mathcal{C}_2$, partition C itself into sets of the form $C \cap C_2$, $C \cap C_2^{-1}$, and $C \cap C_2^{->1}$ (where the sets $C_2^{-1} \in \mathcal{C}^{-1}(C_2)$), and follow the procedure outlined above. Clearly this methodology can be repeated in polynomial time for all fixed integers k, and this methodology forms the essence of the algorithm to which we will refer as "Version 1" of the "Common Factor Algorithm".

Another methodology for partitioning $P - C_k$, is to do so in the same manner as the algorithms of the previous chapter, i.e. where we write

$$C_k = \bigcap_{j=1}^{t} N_{v(j)}^P, \quad \text{then} \tag{6.126}$$

$$P - C_k = \bigcup_{j=1,\ldots,t}^{\text{disjoint}} \left(Y_{v(j)}^P \cap \bigcap_{i=1}^{j-1} N_{v(i)}^P \right). \tag{6.127}$$

Again the inequality $\alpha^T x \geq \beta$ reduces to a lower pitch inequality on each set $Y_{v(j)}^P \cap \bigcap_{i=1}^{j-1} N_{v(i)}^P$, implying an inductive technique. The algorithm that arises from this methodology will be referred to as "Version 2" of the "Common Factor Algorithm".

Before we describe the algorithms formally, we will first generalize the results of this section, which dealt exclusively with set-covering type problems, to the general case.

6.3 The General Case

Definition 6.10 *In line with the notation of the previous chapter, for each $l = 1, \ldots, n$ we will write $M_{l'}^P = Y_l^P$, $M_{l''}^P = N_l^P$. We will also define (somewhat asymmetrically)*

$$N_j^P = (M_j^P)^c, \quad \textit{for each } j \in \{1', 1'', \ldots, n', n''\} \tag{6.128}$$

so that $N_{l'}^P = N_l^P$, *and* $N_{l''}^P = Y_l^P$, *for all* $l = 1, \ldots, n$.

The results of the previous section depended heavily on the fact that where P is defined as

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} Y_j, \quad A_i \subseteq \{1, \ldots, n\},\ i = 1, \ldots, m \tag{6.129}$$

the valid pitch 1 constraints on P are all dominated by the constraints $x(A_i) \geq 1$. For the general case

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} M_j \tag{6.130}$$

where

$$A_i \subseteq \{1', 1'', \ldots, n', n''\}, \quad \text{and} \quad M_{l'} = Y_l,\ M_{l''} = N_l,\ l = 1, \ldots, n \tag{6.131}$$

this does not necessarily hold. Nevertheless there are two fairly straightforward ways to apply the methodology used for set covering in the previous section to the general problem. Recall first that for the general problem, P can be equivalently defined as,

$$P = \{y \in \{0,1\}^n : y(A_i) \geq 1,\ i = 1, \ldots, n\} \tag{6.132}$$

where $y(A_i) = \sum_{j \in A_i} y_j$, and $y_{l'} = y_l$, and $y_{l''} = 1 - y_l$. The forbidden configurations that define this problem are

$$\bigcap_{l : l' \in A_i} N_l^P \cap \bigcap_{l : l'' \in A_i} Y_l^P = \bigcap_{j \in A_i} N_j^P \tag{6.133}$$

(i.e. for each $i = 1 \ldots, m$, the set of points that violates $y(A_i) \geq 1$ is $\bigcap_{j \in A_i} N_j$). Recall also that P can be equivalently expressed as

$$P = \{(y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''}) \in \{0,1\}^{2n} :$$

$$y(A_i) \geq 1,\ i = 1, \ldots, m,\ y_{l'} + y_{l''} = 1,\ l = 1, \ldots, n\} \tag{6.134}$$

which, where for each $j \in \{1', 1'', \ldots, n', n''\}$ we write

$$Y_j' = \{y \in P \subseteq \{0,1\}^{2n} : y_j = 1\}, \quad N_j' = \{y \in P \subseteq \{0,1\}^{2n} : y_j = 0\}, \tag{6.135}$$

is defined by the forbidden configurations

$$\bigcap_{j \in A_i} N_j',\ i = 1, \ldots, m, \quad \text{and} \quad Y_{l'}' \cap Y_{l''}',\ N_{l'}' \cap N_{l''}',\ l = 1, \ldots, n. \tag{6.136}$$

The most obvious option is to apply the set covering technique to the relaxation

$$P' = \{y = (y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''}) \in \{0,1\}^{2n} :$$

$$y(A_i) \geq 1,\ i = 1, \ldots, m,\ y_{l'} + y_{l''} \geq 1,\ l = 1, \ldots, n \tag{6.137}$$

instead. (We could then strengthen the relaxation by demanding $x_{l'} + x_{l''} = 1$ from each of the vectors produced by applying the technique; we will see the details later.) This relaxation is of the set covering form, and is defined by the forbidden configurations

$$\bigcap_{j \in A_i} N_j^{P'} \text{ (paralleling the original f.c.'s), and } N_{l'}^{P'} \cap N_{l''}^{P'},\ l = 1, \ldots, n. \tag{6.138}$$

These $m + n$ forbidden configurations will then be used to produce the collections $\mathcal{C}_k$ of "common factors", and the algorithm will proceed as in the previous section. (Observe that the essential characterization of the set covering case is the fact that its forbidden configurations are all intersections of "no's", i.e. sets of the form $N_l^P,\ l \in \{1, \ldots, n\}$. This is why we have chosen to recast the "yeses" in the general problem as "no's".)

The methodology that we will be using, however, is slightly different. We will see that it suffices to make use of just the original m forbidden configurations $\bigcap_{j \in A_i} N_j^{P'}$ in forming the collections of "common factors" $\mathcal{C}_k$. We will see, however, that it will be necessary to relax slightly the definition of the collections $\bar{\mathcal{C}}_k$.

Note first that the fundamental observation that the removal of overlapping variables leaves convex hull defining constraints applies to the general problem as well. We will give here a direct proof of this result for P. The statement will hold for P' as well, but that will not directly concern us.

Lemma 6.11 *Consider*

$$H = \{y \in \{0,1\}^{2n} :$$

$$y(B_i) \geq 1,\ \forall i = 1, \ldots, m,\ y_{l'} + y_{l''} = 1,\ \forall l = 1, \ldots, n\} \tag{6.139}$$

where the index sets B_i are mutually disjoint, and $|B_i \cap \{l', l''\}| \leq 1$ for each $1 \leq i \leq m$ and $1 \leq l \leq n$. Then where

$$\bar{H} = \{x \in [0,1]^{2n} :$$

$$x(B_i) \geq 1,\ \forall i = 1, \ldots, m,\ x_{l'} + x_{l''} = 1,\ \forall l = 1, \ldots, n\} \tag{6.140}$$

we have

$$Conv(H) = \bar{H}. \tag{6.141}$$

Proof: Let x^* be any extreme point of $\bar{H}$. Form an undirected bipartite graph G with vertex set

$$\{u_j :\ j = 1, \ldots, n\} \quad \cup \quad \{v_i :\ i = 1, \ldots, m\} \tag{6.142}$$

and with an edge $\{u_j, v_i\}$ for each j, i such that either

$$j' \in B_i \quad \text{and} \quad x^*_{j'} > 0 \tag{6.143}$$

or

$$j'' \in B_i \quad \text{and} \quad x^*_{j''} > 0. \tag{6.144}$$

Note that, by assumption, we can never have both $j' \in B_i$ and $j'' \in B_i$. Note also that each node u_j has degree of either 0, 1 or 2, and each node v_i has positive degree. Note also that the degree of a node u_j is 2 if and only if $0 < x^*_{j'} < 1$ and $0 < x^*_{j''} < 1$, and that the degree of a node v_i is 1, with a single edge $\{u_j, v_i\}$ incident if and only if either $j' \in B_i$ and $x^*_{j'} = 1$ or $j'' \in B_i$ and $x^*_{j''} = 1$.

Consider now any subpath in G of the form

$$u_1, v_1, u_2, v_2, u_3 \tag{6.145}$$

and assume without loss of generality that $1'' \in B_1$, $2' \in B_1$, $2'' \in B_2$, and that $3' \in B_2$. It must be that $0 < x^*_{2'} < 1$ and $0 < x^*_{2''} < 1$. Assume first that $x^*_{1''} = 1 = x^*_{3'}$. Then where we define the vectors x^1 and x^2 to be the same as x^* but with

$$x^1_{2'} = x^*_{2'} - \epsilon \text{ and } x^1_{2''} = x^*_{2''} + \epsilon \tag{6.146}$$

and

$$x^2_{2'} = x^*_{2'} + \epsilon \text{ and } x^2_{2''} = x^*_{2''} - \epsilon \tag{6.147}$$

for some sufficiently small $\epsilon > 0$, it is clear that x^1 and x^2 will both belong to $\bar{H}$. But $x^* = \frac{1}{2}x^1 + \frac{1}{2}x^2$, which contradicts the assumption that x^* is an extreme point of $\bar{H}$. Assume now that $x_{1''} < 1$ and $x_{3'} = 1$, then where x^1 and x^2 are as above, but with $x^1_{1''} = x^*_{1''} + \epsilon$ and $x^2_{1''} = x^*_{1''} - \epsilon$, then we still have $x^1, x^2 \in \bar{H}$ and $x^* = \frac{1}{2}x^1 + \frac{1}{2}x^2$. The case for $x_{1''} = 1$ and $x_{3'} < 1$ is similar. For the case $x_{1''} < 1$ and $x_{3'} < 1$ we will also let $x^1_{3'} = x^*_{3'} - \epsilon$ and $x^2_{3'} = x^*_{3'} + \epsilon$ to again obtain the same contradiction.

Thus G can contain no subpath of the form (6.145). It is easy to see that the same contradiction is obtained if u_3 were replaced by u_1, so this implies that G can contain no cycles either. Note also that G cannot contain any *path*

$$v_1, u_2, v_2 \quad \text{or} \quad v_1, u_2, v_2, u_3 \tag{6.148}$$

since we would have to have in that case (again assuming $2' \in B_1,\ 2'' \in B_2$) both $x_{2'} = 1$ (since v_1 is a leaf node) and $x_{2'} < 1$ (since $degree(u_2) = 2$). Thus the only paths that can exist in G are those that either consist of a single edge of the form $\{u_1, v_1\}$, or which consist of exactly two edges, and are of the form u_1, v_1, u_2. In the former case we must have $x^*_{1''} = 1$. For the latter case it is easy to see that we must have $x^*_{1''} = x^*_{2'} = 1$ or else we could again break up x^* into a sum $\frac{1}{2}x^1 + \frac{1}{2}x^2$ where $x^1, x^2 \in \bar{H}$.

We conclude that every extreme point x^* of $\bar{H}$ is integral, and thus actually belongs to H. This proves the lemma. □

The following statement, which is an analog of Corollary 6.2, follows directly from Lemma 6.11, but is stated explicitly for clarity.

Corollary 6.12 *Let $P \subseteq \{0,1\}^{2n}$ be defined by*

$$P = \{y \in \{0,1\}^{2n} : y(A_i) \geq 1,\ i = 1, \ldots, m,\ y_{l'} + y_{l''} = 1,\ l = 1, \ldots, n\} \tag{6.149}$$

and consider the strengthened subsystem *that is obtained by removing all overlapping variables from a particular size k subset of the constraints*

$$\bar{P} = \{y \in \{0,1\}^{2n} : y(B_{r(i)}) \geq 1,\ i = 1, \ldots, k,\ y_{l'} + y_{l''} = 1,\ l = 1, \ldots, n\}\} \tag{6.150}$$

where

$$B_{r(i)} = A_{r(i)} - \bigcup_{j=1,\ldots,k,\ j \neq i} A_{r(j)} \tag{6.151}$$

and where we define $y(\emptyset) = 0$. Let $\alpha^T x \geq \beta$ be any inequality that is valid for $\bar{P}$, then any $x \in [0,1]^{2n}$ for which

1. $x_{l'} + x_{l''} = 1,\ l = 1, \ldots, n$
2. $x_j = 0$, *for all j that belong to two distinct $A_{r(i)}$*
3. $\sum_{j \in A_{r(i)}} x_j \geq 1,\ i = 1, \ldots, k$

also satisfies $\alpha^T x \geq \beta$. □

Note that, as in the previous section, where C is the common factor of the forbidden configurations defined by $A_{r(1)}, \ldots, A_{r(k)}$, i.e.

$$C = \bigcap_{j:j \text{ belongs to 2 distinct } A_{r(i)}} N^P_j \tag{6.152}$$

then x^C can be validly constrained by all three of the enumerated conditions of Corollary 6.12. Thus by Corollary 6.12, so long as we have imposed these constraints on x^C, we will guaranteed that x^C will satisfy $\alpha^T x \geq \beta x_0$ for every inequality $\alpha^T x \geq \beta$ that is valid for

$\bar{P}$. Now we will not be able to show that every valid pitch k constraint for P, $\alpha^T x \geq \beta$, is valid for some strengthened subsystem $\bar{P}$ defined by no more than k of the constraints (which would prove that x^C, where C is the common factor induced by those constraints, satisfies $\alpha^T x^C \geq \beta x_0^C$). But we will show that every valid pitch k constraint for P' (which is also valid for P), as defined in (6.137), is indeed valid for some strengthened subsystem $\bar{P}$ defined by no more than k of the constraints.

The following theorem is an analog of a weaker version of Theorem 6.3.

Theorem 6.13 *Let P be as in (6.134); let P' be as in (6.137), and let $\alpha^T x \geq \beta$, $\alpha \geq 0$, with $\pi(\alpha, \beta) = k \geq 0$, be an inequality that holds for all $y \in P'$. Then there exists some (possibly empty) subcollection*

$$\{A_{r(1)}, \ldots, A_{r(\lambda)}\},\ 0 \leq \lambda \leq k \tag{6.153}$$

such that where we define

$$B_{r(i)} = A_{r(i)} - \bigcup_{j=1,\ldots,\lambda,\ j\neq i} A_{r(j)} \tag{6.154}$$

and

$$P^\alpha = \{y \in \{0,1\}^{2n} : y(B_{r(i)}) \geq 1,\ i = 1, \ldots, \lambda, \quad y_{l'} + y_{l''} = 1,\ l = 1, \ldots, n\} \tag{6.155}$$

$$(P')^\alpha = \{y \in \{0,1\}^{2n} : y(B_{r(i)}) \geq 1,\ i = 1, \ldots, \lambda, \quad y_{l'} + y_{l''} \geq 1,\ l = 1, \ldots, n\} \tag{6.156}$$

we have

1. *$A_{r(i)} \subseteq support(\alpha)$, $i = 1, \ldots, \lambda$*
2. *$\alpha^T x \geq \beta$ is valid for $(P')^\alpha$, and therefore for P^α as well.*

Note that $y(\emptyset) = 0$, so that if for some $i \in \{1, \ldots, \lambda\}$, $B_{r(i)} = \emptyset$, then $(P')^\alpha = \emptyset$.

Proof: First consider the collection

$$A^\alpha = \{i : A_i \subseteq support(\alpha)\}. \tag{6.157}$$

Observe that $\alpha^T x \geq \beta$ is valid for the system

$$(\bar{P'})^\alpha = \{y \in \{0,1\}^{2n} : y(A_i) \geq 1, \forall i \in A^\alpha, \quad y_{l'} + y_{l''} \geq 1,\ l = 1, \ldots, n\}. \tag{6.158}$$

(Otherwise there would be a $y \in \{0,1\}^{2n}$ that satisfies all constraints $y_{l'} + y_{l''} \geq 1$ and all constraints $y(A_i) \geq 1$ for which A_i has elements only from *support*(α), but for which nevertheless $\alpha^T y < \beta$. Resetting all y_j, $j \notin support(\alpha)$, to 1 will maintain $\alpha^T y < \beta$ and will guarantee that y satisfies the rest of the constraints as well, which is a contradiction.)

Suppose first that $|A^\alpha| \leq k$. Then for all $i \in A^\alpha$ define

$$B_i = A_i - \bigcup_{j \in A^\alpha - \{i\}} A_j \tag{6.159}$$

and

$$(P')^\alpha = \{y \in \{0,1\}^{2n} : y(B_i) \geq 1, \ \forall i \in A^\alpha, \ \ y_{l'} + y_{l''} \geq 1, \ l = 1, \dots, n\}. \tag{6.160}$$

Note that $(P')^\alpha$ is just a strengthening of $(\bar{P'})^\alpha$ and therefore every point of $(P')^\alpha$ must satisfy $\alpha^T x \geq \beta$ as well. (If for some $i \in A^\alpha$, $B_i = \emptyset$ then $(P')^\alpha = \emptyset$, and so the constraint $\alpha^T x \geq \beta$ is certainly valid for $(P')^\alpha$.)

Suppose now that $|A^\alpha| > k$. Consider any size k subset $\hat{A}^\alpha \subset A^\alpha$, and for all $i \in \hat{A}^\alpha$ define

$$B_i = A_i - \bigcup_{j \in \hat{A}^\alpha - \{i\}} A_j \tag{6.161}$$

and

$$(P')^\alpha = \{y \in \{0,1\}^{2n} : y(B_i) \geq 1, \ \forall i \in \hat{A}^\alpha, \ \ y_{l'} + y_{l''} \geq 1, \ l = 1, \dots, n\}. \tag{6.162}$$

Again, if there is an $i \in \hat{A}^\alpha$ for which $B_i = \emptyset$ then $(P')^\alpha = \emptyset$, and the constraint $\alpha^T x \geq \beta$ is valid for $(P')^\alpha$. So assume that there is no such i. Thus every $y \in (P')^\alpha$ satisfies

$$y(B_i) \geq 1, \ \forall i \in \hat{A}^\alpha. \tag{6.163}$$

But since there are k such constraints, and all B_i are disjoint and nonempty, there must be at least k coordinates j for which $y_j = 1$ and so we must have $\alpha^T y \geq \beta$ by the definition of pitch k constraints. □

The following definition is an analog of a relaxed version of the definition of the common factors given in Definition 6.5.

Definition 6.14 *Where $\mathcal{F}_k$ is a collection of k sets $A_i \subseteq \{1', 1'', \dots, n', n''\}$, for some $k \geq 0$, define $C(\mathcal{F}_k)$ by*

$$C(\mathcal{F}_k) = \bigcap_{j : j \text{ belongs to two distinct } A_i \in \mathcal{F}_k} N_j^P. \tag{6.164}$$

Define $\mathcal{C}_k$ to be the collection of all expressions $C(\mathcal{F}_j)$, $j \leq k$, where the collection $\mathcal{F}_j = \{A_{r(1)}, \dots, A_{r(j)}\}$ is a size j subcollection of $\{A_1, \dots, A_m\}$.

The definition of the sets $\mathcal{C}^{-r}(C)$, $C^{->r}$, $\mathcal{C}_k^{-r}$ and $\mathcal{C}_k^{->r}$ is parallel to that given in Definition 6.8.

As in Definition 6.4, if $\mathcal{F}_k$ is such that $\{j : j \text{ belongs to two distinct } A_i \in \mathcal{F}_k\} = \emptyset$, then $C(\mathcal{F}_k) = P$.

In the discussions to follow it will usually be most convenient to describe intersections of sets Y_l^P and N_l^P as intersections of sets N_j^P, $j \in \{1', 1'', \ldots, n', n''\}$, and this is what motivates the following definition:

Definition 6.15 *For any expression of the form*

$$C = \bigcap_{j \in J \subseteq \{1',1'',\ldots,n',n''\}} N_j^P \tag{6.165}$$

define $\delta(C) \subseteq \{1', 1'', \ldots, n', n''\}$ to be the index set of the intersection, i.e.

$$\delta(C) = J \tag{6.166}$$

and define

$$\delta(P) = \emptyset. \tag{6.167}$$

For any $j \in \{1', 1'', \ldots, n', n''\}$, define

$$\hat{j} = \begin{cases} l'' : & \text{if } j = l' \\ l' : & \text{if } j = l'' \end{cases}, \quad l = 1, \ldots, n. \tag{6.168}$$

If the expression C is of the form

$$C = \bigcap_{j \in J} N_j^P \cap \bigcap_{j \in \bar{J}} Y_j^P \tag{6.169}$$

then we will define

$$\delta(C) = J \cup \{j : \hat{j} \in \bar{J}\} \tag{6.170}$$

reflecting the fact that

$$C = \bigcap_{j \in J} N_j^P \cap \bigcap_{j : \hat{j} \in \bar{J}} N_j^P. \tag{6.171}$$

For example

$$\delta(N_{1'}^P \cap N_{2''}^P \cap Y_{6'}^P \cap Y_{9''}^P) = \{1', 2'', 6'', 9'\}. \tag{6.172}$$

It now follows from Theorem 6.13 that for any pitch k constraint, $\alpha^T x \geq \beta$, that is valid for P', there is in fact a strengthened subsystem P^α defined by $\lambda \leq k$ of the A_i, for which the constraint $\alpha^T x \geq \beta$ is valid as well, and moreover each of those A_i is a subset of $support(\alpha)$. Thus where C is the common factor of those A_i, then $C \in \mathcal{C}_k$, $\delta(C) \subseteq support(\alpha)$, and by Corollary 6.12, any vector $x \in \{0,1\}^{2n}$ for which $x_{l'} + x_{l''} = 1$, $l = 1, \ldots, n$, $x_j = 0$ for all $j \in \delta(C)$, and for which $x(A_i) \geq 1$, $i = 1, \ldots, m$ (this last condition is actually more than we need) must also satisfy $\alpha^T x \geq \beta$. We will now state this formally as a corollary to Theorem 6.13.

Corollary 6.16 *For any pitch k constraint, $\alpha^T x \geq \beta$, that is valid for P', there exists some $C_k \in \mathcal{C}_k$, with $\delta(C_k) \subseteq support(\alpha)$, such that any $(x_0, x) \in [0, x_0]^{2n+1}$, ($x_0 \geq 0$), for which*

1. $x_{l'} + x_{l''} = x_0, \ l = 1, \ldots, n$
2. $x_j = 0, \ \forall j \in \delta(C_k)$
3. $x(A_i) \geq x_0, \ i = 1, \ldots, m$

also satisfies $\alpha^T x \geq \beta x_0$. $\square$

As we noted above, where α, β and C_k are as in Corollary 6.16, then x_k^C can be validly constrained by all three of the corollary's conditions, and this will then guarantee that $\alpha^T x^{C_k} \geq \beta x_0^{C_k}$.

We are now in a position to apply the algorithms outlined in the previous section to the general problem, and we will do this formally in the following section.

6.4 The Algorithms

6.4.1 Version 1

As with the algorithms of the previous chapter, the original vector $(x_1, \ldots, x_n) \in R^n$ that we seek to ensure belongs to $Conv(P)$ will be construed as the n values $(x[Y_1^P], \ldots, x[Y_n^P])$ for some set function x on $\mathcal{P}$, and $(x_1, \ldots, x_n) \in Conv(P)$ if and only if that set function can be chosen to be a probability measure. We will lift the original vector $(x[Y_1^P], \ldots, x[Y_n^P])$ by creating new variables for set function values on additional sets $x[q]$, $q \in \mathcal{P}$, and we will place constraints on these new values arising from the requirement that the set function x be a probablity measure on $\mathcal{P}$. Recall also that where $V \in \mathcal{P}$, the partial sum x^V is the set function on $\mathcal{P}$ such that $x^V[q] = x[V \cap q]$ for each $q \in \mathcal{P}$, so that defining appropriate variables $x[q \cap V]$ can allow us to describe (projections of) the partial sum vector x^V as well.

The basic theme of this algorithm in particular is to partition sets $V \subseteq P$ into sets of the form $V \cap C$ where C is a set from $\mathcal{C}_k^{-r}$ or $\mathcal{C}_k^{->r}$, for some k and r. Given such partitions we will write the partial sum vector x^V, which will have coordinates for each of the sets $\{P, M_j^P\}$ (among others), as the sum of partial sum vectors $x^{V \cap C}$.

We will want to ensure that where V is an intersection of sets of the form N_j^P, $(j \in \{1', 1'', \ldots, n', n''\})$, and we write $x^V[M_j^P]$ as x_j^V and $x^V[P]$ as x_0^V, that

1. $x_j^V = 0, \ \forall j \in \delta(V)$ (since $V \cap N_j^P = V \Rightarrow V \cap M_j^P = \emptyset$ for those j)
2. $x^V(A_i) \geq x_0^V, \ i = 1, \ldots, m$

as these relationships are essential to the theorems of the previous sections. These equalities can be simply enforced directly, but we will obtain them instead from more general relationships.

The algorithm will successively partition and subpartition P. As is suggested by the general description of the algorithm in Section 6.2, in order to ensure that pitch k constraints are satisfied, it will be necessary for elements of these partitions to satisfy certain lower pitch constraints. Given a partition element of the form

$$C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \text{ or } C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}} \cap C_j^{->r_j} \tag{6.173}$$

the following function will be used to determine what pitch constraints the associated partial sum will be required to satisfy.

Definition 6.17 *Let $k \geq 1$. For each $j \leq k+1$, and every $k-j+1$–tuple, $(r_k, r_{k-1}, \ldots, r_j)$ of positive integers, define*

$$f(k, r_k, \ldots, r_j) = k - \sum_{t=j}^{k} r_t. \tag{6.174}$$

Where $j = k+1$, then we write

$$f(k) = k. \tag{6.175}$$

Given $k-j$ positive integers, $r_k, \ldots, r_{j+1}$, $(j \leq k)$, such that $f(k, r_k, \ldots, r_{j+1}) \geq 2$, define

$$f(k, r_k, \ldots, r_{j+1}, 0) = 1. \tag{6.176}$$

At the k'th "level" of the algorithm, given $j \leq k+1$, an ordered collection of numbers $r = \{r_j, \ldots, r_k\}$, and a set v of the form,

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \tag{6.177}$$

with each $C_h^{-r_h} \in \mathcal{C}_h^{-r_h}$, the quantity $f(k, r_k, \ldots, r_j)$ will identify the pitch of the valid constraints that the partial sum vector x^v will be guaranteed by the algorithm to satisfy. Where $j = k+1$, so that $r = \emptyset$, then the set v is the empty intersection, which we will construe to be P. The vector x^P will therefore be guaranteed by the algorithm at level k to satisfy all valid constraints of pitch $\leq k$. The algorithm will also define vectors of the form x^v, for

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}} \cap C_j^{->r_j} \tag{6.178}$$

but these vectors will never be required to satisfy any more than the pitch 1 constraints.

In the course of describing and analyzing the algorithm we will want to make reference to the pitch of the constraints that partial sums x^v, where v is of the form (6.177) or (6.178), will be required to satisfy. In other words, we would like a terminology that will refer to the "f value" of *sets*. We therefore suggest the following definition.

Definition 6.18 *Let $k \geq 1$. Given $j \leq k+1$, given $r_j, \ldots, r_k$ for which $f(k, r_k, \ldots, r_j)$ is defined, and given a collection of $k-j+1$ sets $C_i^{-r_i} \in \mathcal{C}_i^{-r_i}$, $i = j, \ldots, k$, then where*

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j}, \tag{6.179}$$

we will write

$$f(k, v) := f(k, r_k, \ldots, r_j) \tag{6.180}$$

as a shorthand. Similarly, if $C_j \in \mathcal{C}_j$, $C_i^{-r_i} \in \mathcal{C}_i^{-r_i}$, $i = j+1, \ldots, k$, and

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}} \cap C_j^{->r_j} \tag{6.181}$$

then we will write

$$f(k, v) := f(k, r_k, \ldots, r_j) \tag{6.182}$$

as a shorthand.

Observe that the intersection $C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j}$ is itself an intersection of sets of the form Y_j^P and N_j^P. So

$$\delta(C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j}) \tag{6.183}$$

is the index set of that intersection (when that intersection is reexpressed as an intersection of sets N_j^P alone, as per Definition 6.15). It should be noticed, however, that a set of the form

$$C = \bigcap_{j \in \delta(C)} N_j^P \tag{6.184}$$

might have more than one representation of the form (6.177). In this case the algorithm will form a vector for each representation and it may pose different requirements for the pitch of the constraints that the vectors need to satisfy. (A more efficient implementation of the algorithm would avoid this duplication, but for ease of presentation we will allow it.) Thus – and this is implicit in Definition 6.18 – the expression $f(k, v)$ is not well-defined if v is given merely as $\bigcap_{j \in \delta(v)} N_j^P$. It is only well-defined where we are given the values $r_j, \ldots, r_k$ in the expression (6.177). Similarly, where v is of the form (6.178), the expression $f(k, v)$ is only well-defined where we are given the values $r_j, \ldots, r_k$ in the expression (6.178).

Observe also that for every set of the form

$$C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}} \cap C_j^{->r_j}, \tag{6.185}$$

while this is not an intersection of sets of the form Y_j^P and N_j^P, every point y that belongs to this set must satisfy

$$y \in N_j^P, \ \forall j \in \delta(C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}}). \tag{6.186}$$

This suggests the following definition.

Definition 6.19

$$\delta(C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}} \cap C_j^{->r_j}) := \delta(C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}}) \tag{6.187}$$

Algorithm at Level $k \geq 1$:

Step 1 : Form the Matrix

Let P be as in (6.130). Form a matrix U as follows. Let the rows of U be indexed by the sets $P, Y_1^P, \ldots, Y_n^P, N_1^P, \ldots, N_n^P$, the elements of $\mathcal{C}_2$, and the forbidden configurations

$$\bigcap_{j \in A_i} N_j^P, \ i = 1, \ldots, m. \tag{6.188}$$

The columns of U will be of the following two types.

1. U has a column for each collection of $k-j+1$ sets, $C_i^{-r_i},\ i=j,\ldots,k$ such that

   (a) $j \le k+1$

   (b) $f(k,r_k,\ldots,r_j)$ is defined, and $1 \le f(k,r_k,\ldots,r_j) \le k$

   (c) for each $i=j,\ldots,k,\quad C_i^{-r_i} \in \mathcal{C}^{-r_i}(C_i)$, for some $C_i \in \mathcal{C}_i$, satisfying

$$\delta(C_j) \neq \delta(C_{j+1}) \neq \cdots \neq \delta(C_k) \ \text{ and } \ \delta(C_j) \neq \emptyset. \tag{6.189}$$

2. U has a column for each collection of $k-j+1$ sets, $C_j^{->r_j},\ C_i^{-r_i},\ i=j+1,\ldots,k$ such that

   (a) $j \le k$

   (b) $r_j,\ldots,r_k > 0$ and $f(k,r_k,\ldots,r_j)=1$

   (c) $C_j \in \mathcal{C}_j,\ |\delta(C_j)| > r_j$, and for each $i=j+1,\ldots,k,\quad C_i^{-r_i} \in \mathcal{C}^{-r_i}(C_i)$, for some $C_i \in \mathcal{C}_i$ such that the sets $C_j,\ldots,C_k$ satisfy (6.189).

The columns of type (1) will correspond to the sets

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \tag{6.190}$$

and the columns of type (2) will correspond to the sets

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_j+1} \cap C_j^{->r_j}. \tag{6.191}$$

As a shorthand, a column of type (1) may be identified as the "v'th column" of the matrix, where v is the set of the form (6.190) obtained by intersecting the elements of the $k-j+1$ member collection of sets that corresponds to that column, i.e. we may refer to the $C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j}$ column of the matrix. When we refer, however, to the $v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j}$ column, it should be understood that we are identifying the column not merely by the set $\bigcap_{j\in\delta(v)} N_j^P$ (which may not identify the column uniquely), but by the stated selection of $r_j,\ldots,r_k$, and of sets $C_j^{r_j},\ldots,C_k^{r_k}$ (and the implied selection of j). Thus as an additional shorthand (as per Definition 6.18), we will use the notation

$$f(k,v) \tag{6.192}$$

with reference to the "v'th column of the matrix" to refer to the quantity $f(k,r_k,\ldots,r_j)$, for the selection of integers $r_j,\ldots,r_k$ corresponding to the representation (6.190) of v.

Similarly a column of type (2) may be identified as the "v'th column" of the matrix, where v is the set of the form (6.191) obtained by intersecting the elements of the

$k-j+1$ member collection of sets that corresponds to that column, i.e. we may refer to the $C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_j+1} \cap C_j^{->r_j}$ column of the matrix. Here too, when we refer to the v'th column, it should be understood that we are referring to a particular choice of $j, r_j, \ldots, r_k, C_j, C_{j+1}^{r_j+1}, \ldots, C_k^{r_k}$, and here too we will allow ourselves to use the notation $f(k,v)$ to refer to $f(k, r_k, \ldots, r_j)$.

The column of type (1) that is obtained by the choice of $j = k+1$ corresponds to the empty intersection of sets of the form $C_i^{-r_i}$. We will construe this column as corresponding to the set P, and we will refer to it as U^P or as x^P.

For each column U^v, we will denote U^v by x^v and we will denote the entries of the column by

$$U^v[P] \leftrightarrow x_0^v \tag{6.193}$$

$$U^v[Y_i^P] \leftrightarrow x_{i'}^v \tag{6.194}$$

$$U^v[N_i^P] \leftrightarrow x_{i''}^v \tag{6.195}$$

Step 2 : Impose Constraints

Step 2(A) : General Measure Theoretic Constraints

Enforce the following constraints:

$$x^P[P] = 1 \tag{6.196}$$

$$x^v[u] \geq 0, \ \forall u, v \tag{6.197}$$

$$x^v[Y_l^P] + x^v[N_l^P] = x^v[P], \ j = 1, \ldots, n, \ \forall v \tag{6.198}$$

Where u is a forbidden configuration, then

$$x^v[u] = 0, \ \forall v. \tag{6.199}$$

For each u, v entry of the matrix U, where v is of the form (6.190) (i.e. v is a pure intersection), if there exists some h, l entry of U for which

$$\delta(u) \cup \delta(v) \subseteq \delta(h) \cup \delta(l) \tag{6.200}$$

then enforce

$$U(u,v) \geq U(h,l). \tag{6.201}$$

If v is of the form (6.191) then (6.201) will also be enforced if l is of the form $\tilde{C}_k^{-\tilde{r}_k} \cap \tilde{C}_{k-1}^{-\tilde{r}_{k-1}} \cap \cdots \cap \tilde{C}_{t+1}^{-\tilde{r}_{t+1}} \cap \tilde{C}_t^{->\tilde{r}_t}$, with

$$\delta(u) \cup \delta(v) \subseteq \delta(h) \cup \delta(l),\ \delta(\tilde{C}_t) \subseteq \delta(C_j),\ \text{and}\ \tilde{r}_t \geq r_j. \tag{6.202}$$

For each row u other than the rows corresponding to the sets P, Y_l^P and N_l^P, enforce on each v'th column the constraint

$$x^v[u] \geq \sum_{j \in \delta(u)} x^v[N_j^P] - (|\delta(u)| - 1)x_0^v. \tag{6.203}$$

For each column corresponding to a set of the form (6.191) enforce the following inequality:

$$\sum_{j \in \delta(C_j)} x^v[M_j^P] \geq (r_j + 1)x^v[P]. \tag{6.204}$$

Step 2(B) : Partitioning Constraints

For each type (1) column x^v of the matrix with

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \tag{6.205}$$

and $f(k, v) := f(k, r_k, \ldots, r_j) = 2$, for each $C_{j-1} \in \mathcal{C}_{j-1}$ with

$$\delta(C_{j-1}) \notin \{\emptyset, \delta(C_k), \ldots, \delta(C_j)\}, \tag{6.206}$$

(and recalling from Definitions 6.5 and 6.14 that $\mathcal{C}^{-1}(C_{j-1})$ is the collection of sets obtained by negating exactly one element of the intersection that defined C_{j-1},) enforce

$$x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j}} = x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}} +$$

$$\sum_{C_{j-1}^{-1} \in \mathcal{C}^{-1}(C_{j-1})} x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}^{-1}} + x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}^{->1}}. \tag{6.207}$$

Thus for each v'th column for which $f(k, v) = 2$, x^v is identified as a sum of columns $x^{\bar{v}}$ for which $f(\bar{v}) = 1$. In general, for each type (1) column x^v of the matrix with

$$v = C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j}, \tag{6.208}$$

$f(k, v) = t$ and $2 \leq t \leq k$, for each $C_{j-1} \in \mathcal{C}_{j-1}$ with

$$\delta(C_{j-1}) \notin \{\emptyset, \delta(C_k), \ldots, \delta(C_j)\}, \tag{6.209}$$

enforce

$$x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}} = x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}} + \sum_{C_{j-1}^{-1}\in\mathcal{C}^{-1}(C_{j-1})} x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{-1}} + \cdots + \sum_{C_{j-1}^{-(t-1)}\in\mathcal{C}^{-(t-1)}(C_{j-1})} x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{-(t-1)}} + x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{->(t-1)}} \tag{6.210}$$

where if C_{j-1} is an intersection of $u < t$ sets N_j^P, we say

$$x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{->(t-1)}} = x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{-(u+h)}} = 0 \tag{6.211}$$

for all $h > 0$, and if C_{j-1} is an intersection of exactly t sets N_j^P, we say

$$x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{->(t-1)}} = x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{-t}}. \tag{6.212}$$

Thus for each v'th column for which $f(k,v) = t \geq 2$, x^v is identified as a sum of columns $x^{\bar{v}}$ for which $1 \leq f(\bar{v}) < t$. □

Comments on the Algorithm

- Each entry $x^v[q]$ of the matrix is the value of $x[v \cap q]$, i.e. the $v \cap q$ coordinate of the lifted vector x, and each column x^v of the matrix is a projection of the partial sum of this lifted x taken over the set v. Thus x^P is a projection of the lifted vector x itself. The constraints are all necessity conditions for the lifted vector x (and therefore its projections) to be $\mathcal{P}$-probability measure consistent. (This is clear in the case of constraints (6.196) - (6.199); we will deal with the other constraints in the later comments.) The projection $(x^P[Y_1^P], \ldots, x^P[Y_n^P]) = (x[Y_1^P], \ldots, x[Y_n^P])$ of x^P will belong to $Conv(P)$ if x is indeed $\mathcal{P}$-probability measure consistent.

- Note that we could define rows indexed by other sets as well, and this can make the algorithm stronger, but none of the results to be proven here will depend on the presence of more than these rows alone.

- We noted in the definition of the algorithm that the type (1) column for which $j = k+1$ coresponds to the empty intersection P. Observe that for all other choices of $j \leq k$, there is no type (1) column x^v, with v of the form (6.190), such that $\delta(v) = \emptyset$. This can be seen from the fact that P is not a member of any collection $\mathcal{C}_i^{-r}$ for any $r > 0$, and while we can have $r_j = 0$, and $P \in \mathcal{C}_j$, the restriction (6.189) requires $\delta(C_j) \neq \emptyset$. Observe also that (again by the fact that $\delta(C_j) \neq \emptyset$) we must always have $j \geq 2$. Thus where $k = 1$, there can be no columns of type (2) (this also follows from condition (b) of type (2) columns), and the only column of type (1) arises from the choice of $j = k + 1$, i.e. the only column is x^P.

- The idea behind the restriction (6.189) is that if v is of the form (6.190), and for some q and s we have $\delta(C_q) = \delta(C_s)$, then either $C_q^{-r_q} = C_s^{-r_s}$, in which case $C_q^{-r_q}$ can be removed from the expression without changing the set v, or else $C_q^{-r_q} \cap C_s^{-r_s} = \emptyset$, and v is empty. Similarly if $\delta(C_j) = \emptyset$ (so that $r_j = 0$), then C_j can be removed from the expression without changing the set v. A similar argument holds if v is of the form (6.191).

- With regard to constraint (6.201), it is clear that if v is of the form (6.190) (u is always a pure intersection), and (6.200) holds, that $u \cap v \supseteq h \cap l$. If v is of the form (6.191) and l is $\tilde{C}_k^{-\tilde{r}_k} \cap \tilde{C}_{k-1}^{-\tilde{r}_{k-1}} \cap \cdots \cap \tilde{C}_{t+1}^{-\tilde{r}_{t+1}} \cap \tilde{C}_t^{->\tilde{r}_t}$, then $u \cap v$ is an intersection of the form $Q \cap C_j^{->r_j}$, where Q is a pure intersection, and $h \cap l$ is an intersection of the form $Q' \cap \tilde{C}_t^{->\tilde{r}_t}$, where Q' is a pure intersection. If, additionally, (6.202) holds, then we can note immediately that $Q' \subseteq Q$. Recall now that the set $C_j^{->r_j}$ is defined as the set of points $y \in P$ for which more than r_j of the coordinates $N_j^P,\ j \in \delta(C_j)$ have value 0 (recall that for the N_j^P coordinate to have value 0, where $j \in \{1', 1'', \ldots, n', n''\}$ means that $y_l = 1$ if $j = l'$, and $y_l = 0$ if $j = l''$). The set $\tilde{C}_t^{->\tilde{r}_t}$ is similarly defined as the set of points $y \in P$ for which more than $\tilde{r}_t$ of the coordinates $N_j^P,\ j \in \delta(\tilde{C}_t)$ have value 0. Thus since $\delta(\tilde{C}_t) \subseteq \delta(C_j)$ and $\tilde{r}_t \geq r_j$, every point in $\tilde{C}_t^{->\tilde{r}_t}$ must have at least r_j of its coordinates $N_j^P,\ j \in \delta(C_j)$ at value 0, and must therefore belong to $C_j^{->r_j}$ as well. Thus $\tilde{C}_t^{->\tilde{r}_t} \subseteq C_j^{->r_j}$, which implies that $h \cap l \subseteq u \cap v$.

- With regard to constraint (6.203), recall that for each $l \in \{1, \ldots, n\}$, $N_{l'}^P$ is defined as N_l^P, and $N_{l''}^P$ is defined as Y_l^P. The constraint is justified by noting that for any measure X, and any collection of measurable subsets $\{T_1, \ldots, T_h\}$ of a measurable set Ω with $X[\Omega] < \infty$, the measure-theoretic inequality

$$X\left[\bigcap_{j=1}^{h} T_j\right] \geq \sum_{j=1}^{h} X[T_j] - (h-1)X[\Omega] \tag{6.213}$$

is always valid.

- Note that the terms

$$x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}^{->(t-1)}} \tag{6.214}$$

and

$$x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}^{-(u+h)}}, \quad h > 0 \tag{6.215}$$

in expression (6.211), as well as the expression

$$x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}^{->(t-1)}} \tag{6.216}$$

in (6.212) have no associated columns, as in this case $C_{j-1}^{->(t-1)}$ and $\mathcal{C}^{-(u+h)}(C_{j-1})$ are undefined.

- The arguments we gave on page 285 to justify (6.90) will justify (6.204) as well, but we will reiterate one of those arguments here for convenience. Observe that the partial sum vector

$$x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}} \cap C_j^{->r_j}} \tag{6.217}$$

is the sum of the partial sum vectors for each of the atoms of $\mathcal{P}$ that belong to the set $C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_{j+1}^{-r_{j+1}} \cap C_j^{->r_j}$, and each of these is a nonnegative multiple of the (projected) zeta vector for that atom. But for any atom $q \subseteq C_j^{->r_j}$ the zeta vector ζ^q must satisfy

$$\sum_{j \in \delta(C_j)} \zeta^q[M_j^P] \geq r_j + 1 = (r_j + 1)\zeta^q[P] \tag{6.218}$$

by definition of $C_j^{->r_j}$, and therefore the partial sum vector x^q must satisfy

$$\sum_{j \in \delta(C_j)} x^q[M_j^P] + \geq (r_j + 1)x^q[P]. \tag{6.219}$$

- Note that as the partitioning constraints ensure that each v'th column for which $f(k, v) > 1$ can be written as a sum of w'th columns for which $f(k, w) = 1$, constraints (6.197), (6.198), (6.199) and (6.203) only actually need to be enforced on the w columns for which $f(k, w) = 1$.

- For the case $v = P$ (i.e. $j = k + 1$), $f(k, P) = f(k) = k$, and thus applying (6.210) to x^P, we obtain that for each $C_k \in \mathcal{C}_k$,

$$x^P = x^{C_k} + \sum_{\mathcal{C}^{-1}(C_k)} x^{C_k^{-1}} + \cdots + \sum_{\mathcal{C}^{-(k-1)}(C_k)} x^{C_k^{-(k-1)}} + x^{C_k^{->(k-1)}} \tag{6.220}$$

- It is easy to see that for each fixed k, the collections $\mathcal{C}_j^{-r_j}$ with $j, r_j \leq k$, are bounded in size by a polynomial in m (the number of constraints defining the original integer programming formulation). Thus the algorithm at level k runs in polynomial time, and produces a linear system with a number of variables and constraints that is polynomially bounded in n and m. □

Example: Let P be the set of $y \in \{0,1\}^n$ that satisfy the following system of constraints:

$$y_1 + y_2 + (1 - y_6) \geq 1 \tag{6.221}$$

$$y_2 + (1 - y_3) + (1 - y_5) + y_6 \geq 1 \tag{6.222}$$

$$y_1 + (1 - y_3) + (1 - y_4) + (1 - y_5) \geq 1 \tag{6.223}$$

$$y_2 + y_3 + (1 - y_4) + y_6 \geq 1 \tag{6.224}$$

The forbidden configurations are therefore:

$$N_1^P \cap N_2^P \cap Y_6^P = N_{1'}^P \cap N_{2'}^P \cap N_{6''}^P \tag{6.225}$$

$$N_2^P \cap Y_3^P \cap Y_5^P \cap N_6^P = N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P \cap N_{6'}^P \tag{6.226}$$

$$N_1^P \cap Y_3^P \cap Y_4^P \cap Y_5^P = N_{1'}^P \cap N_{3''}^P \cap N_{4''}^P \cap N_{5''}^P \tag{6.227}$$

$$N_2^P \cap N_3^P \cap Y_4^P \cap N_6^P = N_{2'}^P \cap N_{3'}^P \cap N_{4''}^P \cap N_{6'}^P \tag{6.228}$$

The elements of the collection $\mathcal{C}_2$ (with distinct index sets $\delta(C)$) are:

$$P,\ N_{1'}^P,\ N_{2'}^P,\ N_{3''}^P \cap N_{5''}^P,\ N_{2'}^P \cap N_{6'}^P,\ N_{4''}^P \tag{6.229}$$

The collection $\mathcal{C}_3$ is comprised of the sets that comprise $\mathcal{C}_2$, and the additional sets:

$$N_{1'}^P \cap N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P, \tag{6.230}$$

$$N_{2'}^P \cap N_{6'}^P, \tag{6.231}$$

$$N_{1'}^P \cap N_{2'}^P \cap N_{4''}^P, \tag{6.232}$$

$$N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P \cap N_{6'}^P \cap N_{4''}^P \tag{6.233}$$

An example of a set in $\mathcal{C}^{-1}(N_{3''}^P \cap N_{5''}^P)$ is $N_{3'}^P \cap N_{5''}^P$. An example of a set in $\mathcal{C}^{-1}(N_{1'}^P \cap N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P)$ is

$$N_{1''}^P \cap N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P. \tag{6.234}$$

An example of a set in $\mathcal{C}^{-2}(N_{1'}^P \cap N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P)$ is

$$N_{1''}^P \cap N_{2''}^P \cap N_{3''}^P \cap N_{5''}^P. \tag{6.235}$$

At level 3 of the algorithm there will be a row for P, for each of $Y_1^P, \ldots, Y_6$ and $N_1^P, \ldots, N_6^P$, for each of the elements of $\mathcal{C}_2$, and for each of the forbidden configurations, and a column:

- for P (with f value 3)

- for each $C_3 \in \mathcal{C}_3 - \{P\}$ (with f value 1),
- for each $C_3^{-1} \in \mathcal{C}_3^{-1}$ (with f value 2),
- for each $C_3^{-2} \in \mathcal{C}_3^{-2}$ (with f value 1),
- for each $C_3^{-1} \cap C_2$ with $C_3^{-1} \in \mathcal{C}_3^{-1},\ C_2 \in \mathcal{C}_2$, subject to (6.189) (with f value 1),
- for each $C_3^{-1} \cap C_2^{-1}$ with $C_3^{-1} \in \mathcal{C}_3^{-1},\ C_2^{-1} \in \mathcal{C}_2^{-1}$, subject to (6.189) (with f value 1),
- for each $C_3^{->2} \in \mathcal{C}_3^{->2}$ (with f value 1), and
- for each $C_3^{-1} \cap C_2^{->1}$ with $C_3^{-1} \in \mathcal{C}_3^{-1},\ C_2^{->1} \in \mathcal{C}_2^{->1}$, subject to (6.189) (with f value 1).

Observe that there will be a column for the set

$$v = N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P \cap N_{6'}^P \cap N_{4''}^P \in \mathcal{C}_3 \tag{6.236}$$

but for each u'th entry of that column,

$$\delta(u) \cup \delta(v) \supseteq \delta(v) = \{2', 3'', 5'', 6', 4''\} = \delta(N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P \cap N_{6'}^P) \cup \delta(v) \tag{6.237}$$

and thus by constraints (6.199) and (6.201) every entry of that column has value zero.

An example of constraint (6.203) together with (6.196) is:

$$x^P[N_{3''}^P \cap N_{5''}^P] \geq x^P[N_{3''}^P] + x^P[N_{5''}^P] - x^P[P] = x^P[Y_3^P] + x^P[Y_5^P] - 1. \tag{6.238}$$

Another example, combining with constraint (6.199) is

$$0 = x^P[N_{2'}^P \cap N_{3''}^P \cap N_{5''}^P \cap N_{6'}^P] \geq x^P[N_2^P] + x^P[Y_3^P] + x^P[Y_5^P] + x^P[N_6^P] - 3 \tag{6.239}$$

which, together with (6.198), implies that

$$(1 - x^P[Y_2^P]) + x^P[Y_3^P] + x^P[Y_5^P] + (1 - x^P[Y_6^P]) \leq 3 \Rightarrow \tag{6.240}$$

$$x^P[Y_2^P] + (1 - x^P[Y_3^P]) + (1 - x^P[Y_5^P]) + x^P[Y_6^P] \geq 1 \tag{6.241}$$

which is the second of the constraints that defined P. Constraints (6.198), (6.203) and (6.199) actually combine to imply that all columns satisfy all four of the initial constraints (homogenized) that defined P, and we will return to this point soon.

Choosing $N_{1'}^P \cap N_{2'}^P \cap N_{4''}^P \in \mathcal{C}_3$, we have the following example of a partitioning constraint:

$$x^P = x^{N_{1'}^P \cap N_{2'}^P \cap N_{4''}^P} + x^{N_{1''}^P \cap N_{2'}^P \cap N_{4''}^P} + x^{N_{1'}^P \cap N_{2''}^P \cap N_{4''}^P} + x^{N_{1'}^P \cap N_{2'}^P \cap N_{4'}^P} +$$

$$x^{N^P_{1''}\cap N^P_{2''}\cap N^P_{4''}} + x^{N^P_{1''}\cap N^P_{2'}\cap N^P_{4'}} + x^{N^P_{1'}\cap N^P_{2''}\cap N^P_{4'}} +$$

$$x^{C^{->2}(N^P_{1'}\cap N^P_{2'}\cap N^P_{4''})}. \tag{6.242}$$

Observe also that the set $C^{->2}(N^P_{1'} \cap N^P_{2'} \cap N^P_{4''})$ is the set

$$N^P_{1''} \cap N^P_{2''} \cap N^P_{4'} \tag{6.243}$$

and that (6.204) and (6.201) imply that

$$x^{C^{->2}(N^P_{1'}\cap N^P_{2'}\cap N^P_{4''})}[Y^P_1] = \tag{6.244}$$

$$x^{C^{->2}(N^P_{1'}\cap N^P_{2'}\cap N^P_{4''})}[Y^P_2] = x^{C^{->2}(N^P_{1'}\cap N^P_{2'}\cap N^P_{4''})}[N^P_4] = \tag{6.245}$$

$$x^{C^{->2}(N^P_{1'}\cap N^P_{2'}\cap N^P_{4''})}[P]. \ \Box \tag{6.246}$$

Lemma 6.20 *Each v'th column x^v of U satisfies*

1. $x^v[M^P_j] = 0, \ \forall j \in \delta(v)$
2. $x^v[N^P_j] = x^v[P], \ \forall j \in \delta(v)$
3. $\sum_{j\in A_i} x^v[M^P_j] \geq x^v[P], \ \forall i = 1, \ldots, m$

Proof: If $j \in \delta(v)$ then $\delta(v) = \delta(v) \cup \delta(P) = \delta(v) \cup \delta(N^P_j)$, so we conclude by (6.201) that $x^v[P] = x^v[N^P_j]$, which implies that $x^v[M^P_j] = 0$ by (6.198). The third relationship follows from (6.198), (6.199) and (6.203). $\Box$

Before we state the next theorem, recall the notation,

$$x_{l'} = x[M^P_{l'}] = x[N^P_{l''}] = x[Y^P_l] \tag{6.247}$$

$$x_{l''} = x[M^P_{l''}] = x[N^P_{l'}] = x[N^P_l] \tag{6.248}$$

for each $l = 1, \ldots, n$. Thus, for example, where $l, h \in \{1, \ldots, n\}$,

$$\alpha_{l'} x_{l'} + \alpha_{h''} x_{h''} = \alpha_{l'} x[Y^P_l] + \alpha_{h''} x[N^P_h]. \tag{6.249}$$

Theorem 6.21 *Let $A_i \subseteq \{1', 1'', \ldots, n', n''\}, \ i = 1, \ldots, m$. Let*

$$P = \{y \in \{0,1\}^n : y(A_i) \geq 1, \ i = 1, \ldots, m\} \tag{6.250}$$

where $y_{l'} = y_l$ and $y_{l''} = 1 - y_l$, and let

$$P' = \{y \in \{0,1\}^{2n}) : y(A_i) \geq 1, \ \forall i = 1, \ldots, m, \ y_{l'} + y_{l''} \geq 1, \ l = 1, \ldots, n\}. \tag{6.251}$$

Denote the subvector of x^v indexed by $Y_1^P,\ldots,Y_n^P,N_1^P,\ldots,N_n^P$ as $\bar{x}^v$.

The algorithm at level k will satisfy that for every column x^v of U, where v is of the form (6.190) or (6.191), for which $f(k,v)=t\le k$, we have $\alpha^T\bar{x}^v\ge\beta x_0^v$, for every constraint $\alpha^Tx\ge\beta$ that is valid for P' such that $(\alpha,\beta)\ge 0$ and $\pi(\alpha,\beta)\le t$. In particular, $\alpha^T\bar{x}^P\ge\beta x_0^P$ will hold for every constraint $\alpha^Tx\ge\beta$ that is valid for P' for which $\pi(\alpha,\beta)\le k$.

Proof: The valid pitch 1 constraints for P' are all dominated by the constraints $y(A_i)\ge 1$, and by Lemma 6.20,

$$x^v(A_i)\ge x_0^v,\ i=1,\ldots,m \tag{6.252}$$

for every v. This also implies that the theorem holds for each column x^v of type (2), as for each such v, $f(k,v)=1$. Assume now by induction that for some $1\le t\le k-1$, the theorem holds for every valid constraint of pitch $\le t$, and consider an arbitrary valid constraint for P', $\alpha^Tx\ge\beta$ for which $\pi(\alpha,\beta)=t+1$. Consider now an arbitrary type (1) column corresponding to v of the form

$$v=C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j} \tag{6.253}$$

for which $f(k,r_k,\ldots,r_j)\ge t+1\ge 2$ (and so $r_j>0$ by construction). We will show that x^v satisfies $\alpha^T\bar{x}\ge\beta x_0$ by showing that for some $C_{j-1}\in\mathcal{C}_{j-1}$, every term in the sum

$$\begin{aligned} x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}} &= x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}}+\\ &\sum_{C_{j-1}^{-1}\in\mathcal{C}^{-1}(C_{j-1})} x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{-1}}+\cdots+\\ &\sum_{C_{j-1}^{-t}\in\mathcal{C}^{-t}(C_{j-1})} x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{-t}}+\\ &x^{C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1}^{->t}} \end{aligned} \tag{6.254}$$

(if it exists) satisfies $\alpha^T\bar{x}\ge\beta x_0$.

Note first that $t+1\le f(k,r_k,\ldots,r_j)\le k-(k-j+1)=j-1$, so by Cor 6.16 and Lemma 6.20, there exists some $C_{j-1}\in\mathcal{C}_{j-1}$ with $\delta(C_{j-1})\subseteq \mathit{support}(\alpha)$, such that if U has a column x^w for the expression

$$w=C_k^{-r_k}\cap C_{k-1}^{-r_{k-1}}\cap\cdots\cap C_j^{-r_j}\cap C_{j-1} \tag{6.255}$$

then x^w satisfies $\alpha^T\bar{x}\ge\beta x_0$. If $\delta(C_{j-1})=\emptyset$, then Lemma 6.20 implies that x^v already satisfies $\alpha^T\bar{x}\ge\beta x_0$, and we are done. So suppose that $\delta(C_{j-1})\ne\emptyset$, and let us also suppose, for the moment, that $\delta(C_{j-1}) \neq \delta(C_s)$ for any $s = j \ldots, k$, so there is indeed a column x^w in U.

Consider now the vector

$$x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}^{-u}}, \tag{6.256}$$

where $t \geq u \geq 1$. (Our assumption that $\delta(C_{j-1}) \neq \delta(C_s)$ implies that this vector does not violate (6.189) either.) If C_{j-1} is an intersection of fewer than u sets of the form N_i^P, then (6.256) is zero, which certainly satisfies $\alpha^T \bar{x} \geq \beta x_0$, so assume that this is not the case. Let $\Delta(C_{j-1}^{-u})$ be the index set of those terms of the intersection C_{j-1} that were negated from the form N_i^P to M_i^P in forming C_{j-1}^{-u}. If (6.256) satisfies the valid constraint

$$\sum_{i \in support(\alpha) - \Delta(C_{j-1}^{-u})} \alpha_i x_i \geq \beta x_0 - \sum_{i \in \Delta(C_{j-1}^{-u})} \alpha_i x_0 \tag{6.257}$$

then it also satisfies $\alpha^T \bar{x} \geq \beta x_0$ since for (6.256), each x_i coordinate, $i \in \Delta(C_{j-1}^{-u})$, has value x_0 by Lemma 6.20. Define $\bar{\alpha}$ by

$$\bar{\alpha}_i = \begin{cases} \alpha_i : i \notin \Delta(C_{j-1}^{-u}) \\ 0 : i \in \Delta(C_{j-1}^{-u}) \end{cases} \tag{6.258}$$

The constraint (6.257) is therefore

$$\bar{\alpha}^T \bar{x} \geq (\beta - \sum_{i \in \Delta(C_{j-1}^{-u})} \alpha_i) x_0. \tag{6.259}$$

But by repeated application of Lemma 5.33,

$$\pi(\bar{\alpha}, \beta - \sum_{i \in \Delta(C_{j-1}^{-u})} \alpha_i) \leq \tag{6.260}$$

$$\pi(\alpha, \beta) - |\Delta(C_{j-1}^{-u})| = \pi(\alpha, \beta) - u = t + 1 - u. \tag{6.261}$$

But $f(k, r_k, r_{k-1}, \cdots, r_j, u) \geq t + 1 - u$, so by induction (since $t + 1 - u \leq t$) inequality (6.257), and therefore also the inequality $\alpha^T \bar{x} \geq \beta x_0$, are indeed satisfied by (6.256).

Finally, consider the vector

$$x^{C_k^{-r_k} \cap C_{k-1}^{-r_{k-1}} \cap \cdots \cap C_j^{-r_j} \cap C_{j-1}^{->t}}. \tag{6.262}$$

Again, if C_{j-1} is an intersection of fewer than $t+1$ sets of the form N_i^P, then this vector is zero, which certainly satisfies $\alpha^T \bar{x} \geq \beta x_0$. Otherwise, by (6.204) this vector satisfies

$$\sum_{i \in \delta(C_{j-1})} x_i \geq (t+1) x_0 \Rightarrow \alpha^T \bar{x} \geq \beta x_0 \tag{6.263}$$

by Lemma 6.9 (since $\delta(C_{j-1}) \subseteq support(\alpha)$, and $\pi(\alpha,\beta) = t+1$). We thus conclude from equation (6.254) that x^v satisfies $\alpha^T\bar{x} \geq \beta x_0$ as well.

Until this point we have been assuming that $\delta(C_{j-1}) \neq \delta(C_s)$ for any $s = j,\ldots,k$. Assume now that there is an $s \in \{j,\ldots,k\}$ for which $\delta(C_{j-1}) = \delta(C_s)$. Thus we have $\delta(C_s) \subseteq support(\alpha)$, and since $r_s > 0$, Lemma 6.20 implies that for some $i \in support(\alpha)$ we have $x^v_i = x^v_0$. Thus x^v will satisfy $\alpha^T\bar{x} \geq \beta x_0$ so long as it satisfies $\bar{\alpha}^T\bar{x} \geq (\beta - \alpha_i)x_0$ (where $\bar{\alpha}$ is the same as α but with $\bar{\alpha}_i = 0$). As above, $\pi(\bar{\alpha}, \beta - \alpha_i) \leq \pi(\alpha,\beta) - 1 = t$, and so the therorem follows by induction. □

Remark 6.22 *If P itself is of the form*

$$\{y \in \{0,1\}^n : \sum_{l \in A_i} y_l \geq 1,\ i = 1,\ldots,m\} \tag{6.264}$$

where each $A_i \subseteq \{1,\ldots,n\}$, then we could strengthen the algorithm by replacing the collections $\mathcal{C}^{-r}$ and $\mathcal{C}^{->r}$ by $\bar{\mathcal{C}}^{-r}$ and $\bar{\mathcal{C}}^{->r}$ respectively, and at level k of the algorithm, for each $t \leq k$, the valid (homogenized) pitch $\leq t$ constraints for P would all be satisifed by each subcolumn $\bar{x}^v$ for which $f(k,v) \geq t$.

6.4.2 Version 2

Definition 6.23 *For every $C \in \mathcal{C}_j$, where we represent C as*

$$C = \bigcap_{r=1}^{|\delta(C)|} N^P_{v(r)} \tag{6.265}$$

define the collection of sets

$$\mathcal{C}^{Pt}(C) = \{C,\ M^P_{v(1)},\ N^P_{v(1)} \cap M^P_{v(2)},\ N^P_{v(1)} \cap N^P_{v(2)} \cap M^P_{v(3)}, \cdots, \left(\bigcap_{r=1}^{|\delta(C)|-1} N^P_{v(r)}\right) \cap M^P_{v(|\delta(C)|)}\}. \tag{6.266}$$

The collection of all sets that belong to some $\mathcal{C}^{Pt}(C),\ C \in \mathcal{C}_j$, will be denoted $\mathcal{C}^{Pt}_j$.

Lemma 6.24 *For any j and any $C \in \mathcal{C}_j$,*

$$\bigcup_{C^{Pt} \in \mathcal{C}^{Pt}(C)} C^{Pt} = P \tag{6.267}$$

and the union is disjoint.

Proof: It is clear that the union is disjoint. Represent C as in (6.265). Any point y in the universal set P must either belong to every $N^P_{v(r)}$, in which case $y \in C$, or fail to belong to some $N^P_{v(r)}$. If it fails to belong to some $N^P_{v(r)}$, then let

$$u = \min w : y \notin N^P_{v(w)} \tag{6.268}$$

(obviously $u \leq |\delta(C_j)|$). Then

$$y \in \bigcap_{r=1}^{u-1} N^P_{v(r)} \cap M^P_{v(u)}. \ \Box \tag{6.269}$$

Definition 6.25 *Given j, k, $(2 \leq j \leq k+1)$, and any set v represented as*

$$v = C^{Pt}_k \cap C^{Pt}_{k-1} \cap \cdots \cap C^{Pt}_j \tag{6.270}$$

where each $C^{Pt}_t \in \mathcal{C}^{Pt}_t$, define

$$g(k, v) = j - 1. \tag{6.271}$$

The empty intersection, $v = P$, will be said to be of the form (6.270) with $j = k+1$, and thus

$$g(k, P) = k. \tag{6.272}$$

Again, as with Definition 6.18, $g(k, v)$ is not well-defined if v is given only by $v = \bigcap_{j \in \delta(v)} N^P_j$. The definition requires that we be given $k-j+1$ sets $C^{Pt}_i \in \mathcal{C}^{Pt}_i$ such that v is as in (6.270).

Algorithm Version 2, Level $k \geq 1$

Step 1 : Form the Matrix

Form the matrix U whose rows are indexed by $P, Y^P_1, \ldots, Y^P_n$, $N^P_1, \ldots, N^P_n$, the elements of $\mathcal{C}_2$, and the forbidden configurations

$$\bigcap_{j \in A_i} N^P_j, \ i = 1, \ldots, m \tag{6.273}$$

and whose columns are indexed by all collections of $k-j+1$ sets C^{Pt}_i, $i = j, \ldots, k$ such that

1. $2 \leq j \leq k+1$
2. for each $i = j, \ldots, k$, each $C^{Pt}_i \in \mathcal{C}^{Pt}_i(C_i)$ for some $C_i \in \mathcal{C}_i$ such that the sets $C_j, \ldots, C_k$ satisfy

$$\delta(C_j) \neq \delta(C_{j+1}) \neq \cdots \neq \delta(C_k) \ \text{ and } \ \delta(C_i) \neq \emptyset, \ i = j, \ldots, k. \tag{6.274}$$

Each such column will be said to correspond to the set

$$v = C_k^{Pt} \cap C_{k-1}^{Pt} \cap \cdots \cap C_j^{Pt}. \tag{6.275}$$

Where $j = k+1$, the column corresponds to the empty intersection, and we will refer to this column as x^P.

Step 2 : Enforce Constraints

Enforce constraints (6.196) through (6.199), (6.201) and (6.203) as in the first version of the algorithm (constraint (6.204) is not relevant here).

Here are the partitioning constraints:

For each column v of U for which $g(k,v) \geq 2$, so that v is of the form (6.275), v satisfies restriction (6.274), and $j \geq 3$, impose the following constraint: For each $C_{j-1} \in \mathcal{C}_{j-1}$ such that $\delta(C_{j-1}) \neq \emptyset$ and such that $\delta(C_{j-1})$ is distinct from each $\delta(C_t)$, $t = j,\ldots,k$, enforce

$$x^{C_k^{Pt} \cap C_{k-1}^{Pt} \cap \cdots \cap C_j^{Pt}} = \sum_{C_{j-1}^{Pt} \in \mathcal{C}^{Pt}(C_{j-1})} x^{C_k^{Pt} \cap C_{k-1}^{Pt} \cap \cdots \cap C_j^{Pt} \cap C_{j-1}^{Pt}}. \ \square \tag{6.276}$$

Comments on Version 2:

- For each x^w in the sum on the right hand side of (6.276), $g(k,w) \geq 1$, and restriction (6.274) is satisfied. Thus there is actually a column for each such x^w, and the constraint is well defined. Observe also that $g(k,w) < g(k,v)$.
- It is clear from Lemma 6.24 that constraint (6.276) is valid, and thus all constraints imposed by the algorithm are valid. It is also easy to see that for each fixed k the algorithm runs in polynomial time.
- Applying (6.276) to the empty intersection $v = P$, we have $j = k+1$, and thus for each $C_k \in \mathcal{C}_k$ we obtain

$$x^P = \sum_{C^{Pt} \in \mathcal{C}^{Pt}(C_k)} x^{C^{Pt}}. \tag{6.277}$$

- The idea behind (6.274) is that if v is of the form (6.275) and some $\delta(C_r) = \delta(C_{r'})$, then either $C_r^{Pt} = C_{r'}^{Pt}$, in which case C_r^{Pt} can be removed from the intersection without altering the set v, or if $C_r^{Pt} \neq C_{r'}^{Pt}$ then $v = \emptyset$, since the elements of $\mathcal{C}^{Pt}(C_r)$ are mutually disjoint. Similarly if any $\delta(C_r) = \emptyset$, so that $C_r^{Pt} = C_r = P$, then C_r^{Pt} can be removed from the intersection without altering the set v.

- Note that Lemma 6.20 holds for Version 2 as well.
- In contradistinction to Version 1, in Version 2 we have included columns for intersections of the form $C_k \cap \cdots \cap C_j$, where each $C_t \in \mathcal{C}_t$ and $j < k$. In Version 2 we have also not fixed the f values of intersections $C_k \cap \cdots \cap C_j$ with $C_j \in \mathcal{C}_j$ at 1, i.e. Version 2 may further partition such sets. It should be noted, however, that Version 2 could have also been defined in the absence of both of these features without jeopardizing Theorem 6.26 (the pitch k result). We have defined it as we have in order to easily obtain the termination bound for the algorithm that will be described in the next section. □

Theorem 6.26 *Let P and P' be as in Theorem 6.21. The algorithm at level k will satisfy that for any subcolumn $\bar{x}^v$ for which $g(k, v) = t \leq k$, we have $\alpha^T \bar{x}^v \geq \beta x_0^v$ for every constraint $\alpha^T x \geq \beta$ that is valid for P' for which $\pi(\alpha, \beta) \leq t$. In particular $\alpha^T \bar{x}^P \geq \beta x_0^P$ for every constraint $\alpha^T x \geq \beta$ that is valid for P' for which $\pi(\alpha, \beta) \leq k$.*

Proof: The proof is similar to the proof for Version 1. As in the proof of Theorem 6.21, the result certainly holds for all valid constraints of pitch ≤ 1. Let $2 \leq t \leq k$. Assume now by induction that for all valid constraints $\alpha^T x \geq \beta$ on P' of pitch no more than $t-1$, the constraint $\alpha^T \bar{x} \geq \beta x_0$ holds for every column x^v for which $g(k, v) \geq t-1$. Consider now an arbitrary valid constraint $\alpha^T x \geq \beta$ of pitch t on P', and consider an arbitrary column x^v for which $k \geq g(k, v) = h \geq t$. Thus v is of the form

$$v = C_k^{Pt} \cap C_{k-1}^{Pt} \cap \cdots \cap C_{h+1}^{Pt} \tag{6.278}$$

with each C_r^{Pt} belonging to some $\mathcal{C}^{Pt}(C_r)$ for some $C_r \in \mathcal{C}_r$. (If $h = k$ then this is the empty intersection, i.e. $v = P$.) If we can show that x^v satisfies

$$\alpha^T \bar{x}^v \geq \beta x_0^v \tag{6.279}$$

then the theorem will be proven. By Corollary 6.16 there exists some $C_k \in \mathcal{C}_k$, with $\delta(C_k) \subseteq support(\alpha)$, such that any $(x_0, x) \in [0, x_0]^{2n+1}$, $(x_0 \geq 0)$, for which

1. $x_{l'} + x_{l''} = x_0,\ l = 1, \ldots, n$
2. $x_j = 0,\ \forall j \in \delta(C_k)$

3. $x(A_i) \geq x_0,\ i = 1, \ldots, m$

also satisfies $\alpha^T x \geq \beta x_0$. Thus if there is a column

$$x^{C_k^{Pt} \cap C_{k-1}^{Pt} \cap \cdots \cap C_{h+1}^{Pt} \cap C_h} \tag{6.280}$$

in U, then as in the proof of Theorem 6.21, by Lemma 6.20 the algorithm constraints will guarantee that (6.280) satisfies (6.279). As in the proof of Theorem 6.21, it can also be shown easily by induction that for every $C_h^{Pt} \in \mathcal{C}^{Pt}(C_h) - \{C\}$, the column

$$x^{C_k^{Pt} \cap C_{k-1}^{Pt} \cap \cdots \cap C_{h+1}^{Pt} \cap C_h^{Pt}} \tag{6.281}$$

will satisfy (6.279) as well. Thus if there is in fact a column (6.280), then x^v will indeed satisfy (6.279). If, however, there is no column (6.280), that could only be either because $\delta(C_h) = \emptyset$, or because for some $r \in \{h+1, \ldots, k\}$, $\delta(C_r) = \delta(C_h)$. But if $\delta(C_h) = \emptyset$, then Corollary 6.16 implies that x^v already satisfies (6.279). Similarly if $\delta(C_r) = \delta(C_h)$ and C_r^{Pt} is C_r then $\delta(C_r^{Pt}) = \delta(C_h)$ and Corollary 6.16 implies that x^v already satisfies (6.279). Finally, if $\delta(C_r) = \delta(C_h)$, and $C_r^{Pt} \in \mathcal{C}^{Pt}(C_r) - \{C_r\}$, then for some $j \in \delta(C_r)$, (6.201) implies that $x_j^v = x_0^v$. Thus since $\delta(C_r) = \delta(C_h) \subseteq \textit{support}(\alpha)$, it follows (again as in the proof of Theorem 6.21) that so long as x^v satisfies all valid pitch $t-1$ constraints it will also satisfy (6.279). By induction, the theorem is now proven. □

6.5 Termination Criteria

As with the depth-first partitioning algorithm of the previous chapter, by Lemma 5.6 and Corollary 5.8 we know that the convex hull is obtained by both versions of the algorithm by level $n-1$ in the set covering case, and by level $2n-1$ in the general case. One of the interesting features of these algorithms, however, is that there will be other criteria as well, potentially independent of n and m, that will guarantee that the convex hull has been obtained. For example if the index sets A_i are disjoint, so there are no common factors other than the "empty" factor P, so $\mathcal{C}_2 = \{P\}$, then the algorithm "terminates" after level 1 with the convex hull (c.f. Lemma 6.11). By "termination" we mean that the matrices produced at all levels $k \geq 1$ are all of exactly the same size and are defined by exactly the same constraints, as in this case, for all levels $k \geq 1$ the matrix U will have no columns aside from P. The way that we have defined the algorithms, it is always possible to define additional levels, but eventually the subsequent levels will all be identical and they will do no new work. We will describe here two simple criteria, one for Version 1 and one for Version 2, that will guarantee that the convex hull has already been obtained.

Theorem 6.27 *Let P and P' be as in Theorem 6.21. Let L be the set of indices $l \in \{1,\ldots,n\}$ such that either l' belongs to two distinct A_i, or l'' belongs to two distinct A_i. Let*

$$t = \min\{k : \exists C \in \mathcal{C}_k \textit{ satisfying } l' \in \delta(C) \textit{ or } l'' \in \delta(C),\ \forall l \in L,\ \textit{and } |\delta(C)| \leq k\} \quad (6.282)$$

Then by level t of Version 1 of the algorithm, the vector $(x^P[Y_1^P],\ldots,x^P[Y_n^P])$ will be guaranteed to belong to $Conv(P)$.

Proof: Where $C \in \mathcal{C}_k$ is as in the statement of the theorem, for each $r \geq 0$, each set $C^{-r} \in \mathcal{C}^{-r}(C)$ is a (possibly empty) subset of P for which every $x_{l'}, x_{l''},\ l \in L$, which includes every overlapping variable (from the constraints that define P), has been assigned given $0,1$ values. Thus the set C^{-r} is the set of $y \in \{0,1\}^n$ that satisfies this assignment and that satisfies the system of nonoverlapping constraints that is obtained by plugging that assignment into the original system of constraints. Therefore, denoting the projection on the $Y_1,\ldots,Y_n$ coordinates with the hat symbol, by Lemma 6.11, the algorithm constraints, and Lemma 6.20, for each vector $x^{C^{-r}}$, either $(x_0^{C^{-r}}, \hat{x}^{C^{-r}}) = 0$, or

$$\hat{x}^{C^{-r}}/x_0^{C^{-r}} \in Conv(C^{-r}) \subseteq Conv(P) \quad (6.283)$$

Thus by (6.220), since $|\delta(C)| \leq t$,

$$\hat{x}^P = \hat{x}^C + \sum_{C^{-1} \in \mathcal{C}^{-1}(C)} \hat{x}^{C^{-1}} + \cdots + \sum_{C^{-|\delta(C)|} \in \mathcal{C}^{-|\delta(C)|}(C)} \hat{x}^{C^{-|\delta(C)|}} \quad (6.284)$$

Since the sets, $C^{-r}, 0 \leq r \leq |\delta(C)|$, cover P (actually thay partition P), we conclude by Theorem 5.28 that $\hat{x}^P \in Conv(P)$. $\square$

Theorem 6.28 *Version 2 of the algorithm always obtains the convex hull of P by level $|\mathcal{C}_2|$.*

Proof: Consider the following simplified implementation of the algorithm. Redefine

$$\mathcal{C}_j := \mathcal{C}_2,\ \forall j > 2 \quad (6.285)$$

so that at any level $t \geq 1$ of the algorithm, the columns are indexed by the expressions

$$v = C_1^{Pt} \cap C_2^{Pt} \cap \cdots C_h^{Pt} \quad (6.286)$$

where $h \leq t-1$, and $C_1^{Pt} \in \mathcal{C}^{Pt}(C_1),\ldots,C_h^{Pt} \in \mathcal{C}^{Pt}(C_h)$, for some h distinct members $C_1,\ldots,C_h$ of $\mathcal{C}_2 - \{P\}$ (by (6.274)). (Recall that $\mathcal{C}_2$ is technically defined as the collection of *index sets* of the common factors, and thus common factors C with distinct index sets $\delta(C)$ are distinct members of $\mathcal{C}_2$.) Thus for all levels $k \geq |\mathcal{C}_2|$ of Version 2, the algorithm constructs the same size matrix with the same constraints. Thus if the simplified algorithm does in fact guarantee that

$$(x^P[Y_1], \ldots, x^P[Y_n]) \in Conv(P) \tag{6.287}$$

for some finite level k, then (6.287) must hold at level $|\mathcal{C}_2|$ as well. As the simplified algorithm at level k is just a weakening of the original algorithm at level k (since $\mathcal{C}_2 \subseteq \mathcal{C}_j,\ j \geq 2$), this will complete the proof of the theorem. Considering that every $C \in \mathcal{C}_k$ is an intersection of no more than $\binom{k}{2}$ sets from $\mathcal{C}_2$, this simplified algorithm is actually very similar to the original Version 2 algorithm. This will allow us to prove an analog of Theorem 6.26 that will show that the simplified algorithm also guarantees that for any k, x^P will satisfy all valid pitch $\leq k$ constraints at some finite level t of the algorithm. This will then prove (6.287), and thus the theorem as well. The proof is essentially identical to that of Theorem 6.26, but it is somewhat more complicated, and we will therefore describe it explicitly.

By Lemma 6.20, the valid pitch 1 constraints are all satisfied for each column of the matrix generated at every level of the algorithm. We will show, by induction, that for any $k \geq 2$, at any level $t \geq \sum_{r=2}^{k} \binom{r}{2}$ of the algorithm, every column x^v, where v is of the form

$$v = C_1^{Pt} \cap C_2^{Pt} \cap \cdots \cap C_h^{Pt} \tag{6.288}$$

and each $C_j^{Pt} \in \mathcal{C}^{Pt}(C_j)$ for some $C_j \in \mathcal{C}_2$, satisfies that for any integer $s,\ 2 \leq s \leq k+1$ such that

$$h \leq \sum_{r=s}^{k} \binom{r}{2}, \tag{6.289}$$

x^v satisfies all valid constraints for P' of pitch $\leq s-1$. We will say that $h = 0$ when $v = P$ (the empty intersection), so considering that $h = 0$ satisfies (6.289) for each $s \leq k+1$ (where $s = k+1$, the sum on the right hand side of (6.289) has value 0), this will mean that at level $t \geq \sum_{r=2}^{k} \binom{r}{2}$, the column x^P satisfies all valid constraints of pitch $\leq k$. For example if $k = 5$, then at level 20 of the algorithm, all intersections of ≤ 20 sets C^{Pt} satisfy all pitch 1 constraints. All intersections of ≤ 19 sets C^{Pt} satisfy all pitch 2 constraints. All intersections of ≤ 16 sets C^{Pt} satisfy all pitch 3 constraints. All intersections of ≤ 10 sets C^{Pt} satisfy all pitch 4 constraints, and x^P satisfies all pitch 5 constraints.

The case $s = 2$ is trivial, as all columns satisfy all pitch 1 constraints. Assume now that the hypothesis holds for all $s \leq \phi$, for some $\phi \leq k$, and consider now a valid constraint $\alpha^T x \geq \beta$ for P' of pitch $\phi + 1$, and an arbitrary column x^v with v of the form (6.288), with

$$h \leq \sum_{r=\phi+2}^{k} \binom{r}{2}. \tag{6.290}$$

By Corollary 6.16, recalling that every $C \in \mathcal{C}_k$ (where $\mathcal{C}_k$ is as it was originally defined in Definition 6.14) is an intersection of no more than $\binom{k}{2}$ sets from $\mathcal{C}_2$, there must be some collection of sets

$$\{C_{h+1}, C_{h+2}, \ldots, C_{h+w}\} \subseteq \mathcal{C}_2, \quad w \leq \binom{\phi+1}{2} \tag{6.291}$$

with

$$\bigcup_{r=h+1}^{h+w} \delta(C_r) \subseteq support(\alpha) \tag{6.292}$$

such that any $(x_0, x) \in [0, x_0]^{2n+1}$, $(x_0 \geq 0)$, for which

1. $x_{l'} + x_{l''} = x_0,\ l = 1, \ldots, n$

2. $x_j = 0,\ \forall j \in \bigcup_{r=h+1}^{h+w} \delta(C_r)$

3. $x(A_i) \geq x_0,\ i = 1, \ldots, m$

also satisfies $\alpha^T x \geq \beta x_0$. Without loss of generality, assume that the index sets $\delta(C_{h+r}),\ r = 1, \ldots, w$ are all distinct. Consider now the (possibly empty) collection of sets

$$\mathcal{W} = \{C_{h+r} : r \in \{1, \ldots, w\},\ \delta(C_{h+r}) \notin \{\emptyset, \delta(C_1^{Pt}), \ldots, \delta(C_h^{Pt})\}\} \tag{6.293}$$

and rename the elements of $\mathcal{W}$ as

$$\mathcal{W} = \{C_{h+1}, \ldots, C_{h+w'}\} \tag{6.294}$$

where $0 \leq w' = |\mathcal{W}| \leq w$. Consider now the set

$$u = C_1^{Pt} \cap \cdots C_h^{Pt} \cap C_{h+1} \cap \cdots \cap C_{h+w'}. \tag{6.295}$$

If there is a column x^u, then as in the proof of Theorem 6.26, the algorithm constraints guarantee that x^u will satisfy $\alpha^T x \geq \beta x_0$. If $w' = 0$ (in which case there is certainly a column x^u), then we are done. Otherwise, if there is indeed such a column x^u, then there is also a column $x^{u'}$ for each

$$u' = C_1^{Pt} \cap \cdots C_h^{Pt} \cap C_{h+1} \cap \cdots \cap C_{h+w'}^{Pt} \tag{6.296}$$

where $C_{h+w'}^{Pt} \in \mathcal{C}^{Pt}(C_{h+w'}) - C_{h+w'}$, and moreover since

$$h + w' \leq h + w \leq h + \binom{\phi+1}{2} \leq \sum_{r=\phi+1}^{k} \binom{r}{2}, \tag{6.297}$$

each $x^{u'}$ satisfies all valid pitch ϕ constraints for P'. But this will guarantee that each $x^{u'}$ also satisfies $\alpha^T x \geq \beta x_0$ since $\delta(C_{h+w}) \subseteq support(\alpha)$ and for some $j \in \delta(C_{h+w})$ the algorithm constraints imply that $x_j^{u'} = x_0^{u'}$ (as in the proof of Theorem 6.26). Thus by partitioning constraint (6.276) we obtain that where

$$\bar{u} = C_1^{Pt} \cap \cdots C_h^{Pt} \cap C_{h+1} \cap \cdots \cap C_{h+w'-1} \tag{6.298}$$

the column $x^{\bar{u}}$ satisfies $\alpha^T x \geq \beta x_0$. Repeating the argument, we conclude that x^v also satisfies $\alpha^T x \geq \beta x_0$.

If, however, there is no column x^u, then, by construction, this could only be because $\delta(C_q) = \delta(C_{q'})$ for some $q \in \{1, \ldots, h\}$, $q' \in \{h+1, \ldots, h+w'\}$. By construction $\delta(C_{q'}) \neq \delta(C_q^{Pt})$, so this implies that $C_q^{Pt} \in \mathcal{C}^{Pt}(C_q) - \{C_q\}$. But this implies that for some $j \in \delta(C_q)$, $x_j^v = x_0^v$. Thus since $\delta(C_q) = \delta(C_{q'}) \subseteq \mathit{support}(\alpha)$, it follows that so long as x^v satisfies all valid pitch ϕ constraints, then it will also satisfy $\alpha^T x \geq \beta x_0$. The theorem now follows from induction. □

6.6 A Positive Semidefiniteness Result

Recall that for a stable set problem on a graph $G = (N, E)$, if $C \subseteq N$ is a clique, then the "clique constraint" that corresponds to C is

$$\sum_{\{i,j\} \in C \times C} y_{\{i,j\}} \leq 1. \tag{6.299}$$

This constraint reflects the fact that if $y \in \{0,1\}^n$ is an incidence vector of a stable set, and no pair of coordinates y_i, y_j can simultaneously have value 1 for any distinct $i, j \in C$, then there cannot be more than one coordinate with value 1 among all of the $\{y_i : i \in C\}$. In this section we will deal with a large generalization of clique constraints applied to general problems with feasible regions

$$P = \bigcap_{i=1}^{m} \bigcup_{j \in A_i} M_j = \tag{6.300}$$

$$\{y \in \{0,1\}^n : y(A_i) \geq 1,\ i = 1, \ldots, m\} \tag{6.301}$$

with each $A_i \subseteq \{1', 1'', \ldots, n', n''\}$, and $y_{l'} = y_l$, $y_{l''} = 1 - y_l$, $l = 1, \ldots, n$. We will be considering cases where no k *blocks* of variables from among some t blocks of variables can simultaneously hold a particular assignment of $0, 1$ values. For example, if $n = 12$ and the blocks of variables are

1. $\{y_1, y_2, y_3\}$

2. $\{y_4, y_5, y_6\}$

3. $\{y_7, y_8, y_9\}$

4. $\{y_{10}, y_{11}, y_{12}\}$

then we will be concerned with situations in which, say, no more than 2 of the following assignments of values can hold simultaneously (so $k = 3$ in this case):

1. $y_1 = 1,\ y_2 = 1,\ y_3 = 0$
2. $y_4 = 0,\ y_5 = 0,\ y_6 = 0$
3. $y_7 = 1,\ y_8 = 1,\ y_9 = 1$
4. $y_{10} = 1,\ y_{11} = 0,\ y_{12} = 1.$

This would be the case if there were the following constraints $y(A_i) \geq 1$ in the initial definition of P:

$$y_{1''} + y_{2''} + y_{3'} + y_{4'} + y_{5'} + y_{6'} + y_{7''} + y_{8''} + y_{9''} \geq 1 \tag{6.302}$$

$$y_{1''} + y_{2''} + y_{3'} + y_{4'} + y_{5'} + y_{6'} + y_{10''} + y_{11'} + y_{12''} \geq 1 \tag{6.303}$$

$$y_{1''} + y_{2''} + y_{3'} + y_{7''} + y_{8''} + y_{9''} + y_{10''} + y_{11'} + y_{12''} \geq 1 \tag{6.304}$$

$$y_{4'} + y_{5'} + y_{6'} + y_{7''} + y_{8''} + y_{9''} + y_{10''} + y_{11'} + y_{12''} \geq 1. \tag{6.305}$$

For the assignment, say, $y_1 = 1,\ y_2 = 1,\ y_3 = 0$, to fail to hold means that $y_{1''} + y_{2''} + y_{3'} \geq 1$. Thus if at least two of the given assignments must fail to hold, then at least two of the inequalities

$$y_{1''} + y_{2''} + y_{3'} \geq 1 \tag{6.306}$$

$$y_{4'} + y_{5'} + y_{6'} \geq 1 \tag{6.307}$$

$$y_{7''} + y_{8''} + y_{9''} \geq 1 \tag{6.308}$$

$$y_{10''} + y_{11'} + y_{12''} \geq 1 \tag{6.309}$$

must hold, which means that

$$y_{1''} + y_{2''} + y_{3'} + y_{4'} + y_{5'} + y_{6'} + y_{7''} + y_{8''} + y_{9''} + y_{10''} + y_{11'} + y_{12''} \geq 2 \tag{6.310}$$

is a valid constraint on P. In general, an assignment of values to a block of variables will be represented by an index set $S \subseteq \{1', 1'', \ldots, n', n''\}$. For example the assignment $y_1 = 1,\ y_2 = 1,\ y_3 = 0$ would be represented by the index set $\{1'', 2'', 3'\}$ so that the inequality that will hold iff the assignment fails to hold is $y(S) \geq 1$. Under this terminology, the generalization of the clique constraint is the inequality

$$\sum_{i=1}^{t} x(S_i) \geq t - k + 1. \tag{6.311}$$

Observe that the standard clique inequality corresponds to the special case where the index sets S_i are the singletons $\{h''\},\ h \in C,\ k = 2$ and $t = |C|$. We will now show that for either version of the common factor algorithm, if positive semidefiniteness is imposed on a particular submatrix of the matrix U generated by the algorithm, then the vector x^P will be guaranteed to satisfy these generalized clique constraints at level 2 if $t \geq 2k - 1$, and at level $k - 1$ otherwise. We will then show that though the relaxation produced by the N^+ operator indeed satisfies the standard clique constraints (as shown in Chapter 4), even the stronger N^{++} operator defined in Definition 4.29 will require exponential time to satisfy the generalized clique constraints.

Theorem 6.29 *Let $A_i \subseteq \{1', 1'', \ldots, n', n''\},\ i = 1, \ldots, m,$ satisfy $|A_i \cap \{l', l''\}| \leq 1$ for all $i = 1, \ldots, m$ and $l = 1, \ldots, n,$ and let*

$$P = \{y \in \{0, 1\}^n : y(A_i) \geq 1,\ i = 1, \ldots, m\} \tag{6.312}$$

where $y_{l'} = y_l$ and $y_{l''} = 1 - y_l$. Let $\bar{U}$ be the square submatrix of U whose rows and columns are indexed by P and the other elements of $\mathcal{C}_2$. In addition to the constraints imposed by the algorithm (either version), let us enforce

$$\bar{U} \succeq 0 \tag{6.313}$$

as well. Assume that there exist disjoint subsets $S_1, \ldots, S_t$ of the indices $\{1', 1'', \ldots, n', n''\}$, such that $|S_i \cap \{l', l''\}| \leq 1$ for each $1 \leq i \leq t$ and each $1 \leq l \leq n$, and a positive integer $k \leq t$ such that every k-fold union

$$\bigcup_{j=1}^{k} S_{i_j} = A_h \tag{6.314}$$

(where $i_1, \ldots, i_k$ are all distinct elements of $\{1, \ldots, t\}$) for some $h \in \{1, \ldots, m\}$. Thus for every k-tuple of distinct elements, $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, t\}$, for each $y \in P$,

$$\sum_{j=1}^{k} y(S_{i_j}) \geq 1 \tag{6.315}$$

is one of the defining inequalities, $y(A_h) \geq 1$, of P.

Then the following constraint will hold for the column vector x^P at the 2 level of either version of the algorithm if $t \geq 2k - 1$,

$$\sum_{i=1}^{t} x(S_i) \geq t - k + 1 \tag{6.316}$$

and it will hold regardless of t at the $k - 1$ level.

Proof: Observe that each S_i represents a block of variables, the $l' \in S_i$ variables set to zero and the $l'' \in S_i$ variables set to one, such that for every k blocks there is a constraint in the original definition of P that specifically disallows those assignments of values from holding simultaneously.

As usual, we define the relaxation P' of P by

$$P' = \{y = (y_{1'}, y_{1''}, \ldots, y_{n'}, y_{n''}) \in \{0,1\}^{2n} :$$

$$y(A_i) \geq 1, \; i = 1, \ldots, m, \; y_{l'} + y_{l''} \geq 1, \; l = 1, \ldots, n\}. \tag{6.317}$$

Note that the constraint $\sum_{i=1}^{t} x(S_i) \geq t - k + 1$ is valid for P' under the conditions of the theorem. Moreover, if $t \leq 2k - 2$, then $t - k + 1 \leq k - 1$ which implies that the constraint is of pitch less or equal to $k - 1$. Thus by Theorem 6.21 and Theorem 6.26, this constraint must be satisfied by the column x^P at level $k - 1$.

So suppose now that $t \geq 2k - 1$. Recall that for each $j \in \{1', 1'', \ldots, n', n''\}$, the set $N_j^{P'}$ is defined as

$$N_j^{P'} = \{y \in P' : y_j = 0\}, \tag{6.318}$$

and recall that each row of the matrix U is indexed by an intersection C of the form $\bigcap_{j \in \delta(C)} N_j^P$. Our proof method will be to show that under the conditions of the theorem, if we rename the coordinates of the column x^P from the form $C = \bigcap_{j \in \delta(C)} N_j^P$ to the form $C' = \bigcap_{j \in \delta(C)} N_j^{P'}$, then a particular subvector of the column x^P will be $\mathcal{P}'$-probability measure consistent. In other words, we will be showing that for some subvector of x^P indexed by some collection of sets $C_1, \ldots, C_\phi \in \mathcal{P}$, there exists a probability measure χ on the susbset algebra $\mathcal{P}'$ of P' such that

$$\chi \left[\bigcap_{j \in \delta(C_i)} N_j^{P'} \right] = x^P \left[\bigcap_{j \in \delta(C_i)} N_j^P \right], \quad \forall i = 1, \ldots, \phi. \tag{6.319}$$

We will then use the properties of probability measures to obtain relationships between the quantities of the form $\chi[\bigcap_{j \in \delta(C_i)} N_j^{P'}] = x^P[C_i]$ and to thereby prove the theorem.

Let us say that the block of variables S_i is *violated* by $y \in P$ if $y(S_i) = 0$, i.e. if $y_j = 0$ for all $j \in S_i$. Thus the set of points in P that violates the i'th block is

$$B_i = \bigcap_{j \in S_i} N_j^P \quad (= \bigcap_{l' \in S_i} N_l^P \cap \bigcap_{l'' \in S_i} Y_l^P). \tag{6.320}$$

More generally, where $0 \leq r \leq t$, let $g = \{g_1, \ldots, g_r\} \subseteq \{1, \ldots, t\}$ index a collection of distinct blocks of variables. Then the set of points in P that violates all blocks $g_1, \ldots, g_r$, which will be denoted $T(g)$, is

$$T(g) = \bigcap_{i=1}^{r} B_{g_i} = \bigcap_{j \in \bigcup_{i=1}^{r} S_{g_i}} N_j^P, \tag{6.321}$$

and if $r = 0$ then $T(g) = P$. We will show first that for any $T(g),\ |g| = r,\ 0 \leq r < k$, there exists some $C(g) \in \mathcal{C}_2$ such that $C(g) = T(g)$. More specifically, we will show that there exists a unique element $C(g) \in \mathcal{C}_2$ whose index set $\delta(C(g))$ is also $\bigcup_{i=1}^{r} S_{g_i}$.[1] The proof is as follows. The case $r = 0$ is trivial, as the empty intersection (namely P) is a member of $\mathcal{C}_2$, so assume $r \geq 1$. Select subsets J_i and $\bar{J}_i$ of $\{1, \ldots, t\}$ such that

$$|J_i| = |\bar{J}_i| = k - r,\ J_i \cap \bar{J}_i = \emptyset,\ g_j \notin J_i \cup \bar{J}_i,\ j = 1, \ldots, r. \tag{6.322}$$

(This construction requires only that there be $2k - r$ distinct blocks, and by assumption $t \geq 2k - 1 \geq 2k - r$.) Thus, by assumption, there exist distinct u and v for which

$$A_u = \bigcup_{l=1}^{r} S_{g_l} \cup \bigcup_{h \in J_i} S_h,\quad A_v = \bigcup_{l=1}^{r} S_{g_l} \cup \bigcup_{h \in \bar{J}_i} S_h \tag{6.323}$$

so by the disjointness of the blocks, the common factor of A_u and A_v is

$$C(\{A_u, A_v\}) = \bigcap_{j \in \bigcup_{i=1}^{r} S_{g_i}} N_j^P = T(g). \tag{6.324}$$

There is thus a unique row and column in the matrix $\bar{U}$ for each $g,\ |g| < k$.

For each $i = 1, \ldots, t$, define the set B_i' by

$$B_i' = \bigcap_{j \in S_i} N_j^{P'} \tag{6.325}$$

and for each g with $|g| \leq t$, define the set $T'(g)$ as

$$T'(g) = \bigcap_{i=1}^{r} B'_{g_i} = \bigcap_{j \in \bigcup_{i=1}^{r} S_{g_i}} N_j^{P'} \tag{6.326}$$

and where $g = \emptyset$, we construe $T'(g)$ as

$$T'(g) = P' \tag{6.327}$$

[1] Recall that the "common factors" $C(\mathcal{F}) \in \mathcal{C}_2$ are identified by their index sets $\delta(C(\mathcal{F}))$, which are the sets of indices shared by the elements of $\mathcal{F}$. By our definitions, common factors of collections $\mathcal{F}$ with identical index sets are only listed once in $\mathcal{C}_2$, so the index set $\bigcup_{i=1}^{r} S_{g_i}$ uniquely identifies an element of $\mathcal{C}_2$. Even if, however, we had neglected to enforce such a rule and listed elements of $\mathcal{C}_2$ for every family $\mathcal{F}$ of ≤ 2 distinct A_i, it still follows from algorthm constraint 6.201 that the rows of U corresponding to elements of $\mathcal{C}_2$ with identical index sets are themselves identical.

(as P' is the universal set with respect to sets of the form $N_j^{P'}$). Observe that for each $g = \{g_1, \ldots, g_r\}$ with $r \geq k$, there exists some $h \in \{1, \ldots, m\}$ such that

$$T'(g) = \bigcap_{i=1}^{r} B'_{g_i} = \bigcap_{j \in \bigcup_{i=1}^{r} S_{g_i}} N_j^{P'} \subseteq \bigcap_{j \in A_h} N_j^{P'} = \emptyset \tag{6.328}$$

and $T(g) = \emptyset$ as well for the same reason.

Define now the vector X with a coordinate for each $T'(g),\ |g| < k$ (technically, X should be construed as having a coordinate for each $g,\ |g| < k$, but we will be referring to each g'th coordinate as $X[T'(g)]$) with value $x^P[T(g)]$. (The quantity $x^P[T(g)]$ is more precisely referred to as $x^P[C(g)]$, where $C(g)$ is the element of $\mathcal{C}_2$ that is defined by the index set $\delta(C(g)) = \bigcup_{i=1}^{r} S_{g_i}$, but we will refer to this quantity as well using "T" notation.) Note that X has a coordinate for $T'(\emptyset) = P'$ with value $x^P[T(\emptyset)] = x^P[P] = 1$ by (6.196). Consider the subvector $\hat{X}$ of X with coordinates for only those g such that $T'(g) \neq \emptyset$. Observe now that $T'(g) = \emptyset$ means that there are no points $y \in P'$ with a 0 in each j coordinate for every $j \in \delta(T(g))$.[2] But we claim that this implies that either:

- Indices l' and l'' both belong to $\delta(T(g))$ for some $l \in \{1, \ldots, n\}$, and therefore $X[T'(g)] = x^P[T(g)] = 0$ by algorithm constraints (6.198) and (6.201). Or:
- There must be some $A_i,\ i \in \{1, \ldots, m\}$ such that $A_i \subseteq \delta(T(g))$, in which case $X[T'(g)] = x^P[T(g)] = 0$ by algorithm constraints (6.199) and (6.201).

To see this, suppose that there is no $l \in \{1, \ldots, n\}$ with $l', l'' \in \delta(T(g))$, and that there is also no A_i with $A_i \subseteq \delta(T(g))$. Then the point $\bar{y} \in \{0,1\}^{2n}$ with zeroes in exactly the $\delta(T(g))$ coordinates would satisfy $y(A_i) \geq 1,\ \forall i = 1, \ldots, m$, and $y_{l'} + y_{l''} \geq 1,\ \forall l = 1, \ldots, n$, which implies that $\bar{y} \in P'$, which is a contradiction. We therefore conclude that

$$X = (\hat{X}, 0). \tag{6.329}$$

Observe now that for each $T'(g)$ such that $T'(g) \neq \emptyset$ then where $y \in \{0,1\}^{2n}$ is such that $y_j = 0$ iff $j \in \delta(T(g))$, then $y \in P'$ (or else $T'(g)$ would have been empty). Thus the collection, to be denoted $\mathcal{T} \subseteq \mathcal{P}'$, of the nonempty sets $T'(g),\ |g| < k$ is a subcollection of the linearly independent spanning collection $\bar{I}_N^{P'}$ (defined in Definition 3.51). By Theorem 3.53 and Corollary 3.40, there therefore exists a $\mathcal{P}'$-signed-measure χ' that agrees with $\hat{X}$ in the sense that for each $T'(g),\ |g| < k$,

$$\hat{X}[T'(g)] = \chi'[T'(g)]. \tag{6.330}$$

[2] By this we mean $\bigcup_{i=1}^{r} S_{g_i}$. This set is more accurately referred to, however, as $\delta(C(g))$ as δ is technically a function of *set theoretic expressions* such as $C(g)$ rather than of *sets* such as $T(g)$.

Define the collection of sets $\mathcal{T}' \subseteq \mathcal{P}'$ by

$$\mathcal{T}' = \{u \cap v : u, v \in \mathcal{T}\} \tag{6.331}$$

and define $\tilde{\chi}$ to be the projection of the signed measure χ' on $R^{\mathcal{T}'}$. Define now the matrix $U^{\tilde{\chi}}$ with rows and columns indexed by $\mathcal{T}$, with each $(T'(g), T'(g^*))$ entry denoted as $U^{\tilde{\chi}}(g, g^*)$, and of value $\tilde{\chi}[T'(g) \cap T'(g^*)] = \tilde{\chi}[T'(g \cup g^*)]$. Thus by definition, $\tilde{\chi}$ is $\mathcal{P}'$-signed-measure consistent. Observe moreover that

$$\mathcal{T}' = \mathcal{T} \cup \{\emptyset\} \tag{6.332}$$

since for any $T'(g), T'(g^*) \in \mathcal{T}$, if $T'(g \cup g^*) \notin \mathcal{T}$ then $T'(g \cup g^*) = \emptyset$ by definition if $|g \cup g^*| < k$, and if $|g \cup g^*| \geq k$, then we also have $T'(g \cup g^*) = \emptyset$ by (6.328). Thus $\mathcal{T}$ is an inclusion maximal linearly independent subcollection of $\mathcal{T}'$. It now follows from Theorem 4.10 that if additionally, $U^{\tilde{\chi}} \succeq 0$, then $\tilde{\chi}$ must actually be consistent with a *measure* on $\mathcal{P}'$. Considering moreover that $\tilde{\chi}[P'] = \tilde{\chi}[T'(\emptyset)] = \chi'[T'(\emptyset)] = \hat{X}[T'(\emptyset)] = X[T'(\emptyset)] = 1$ (as shown above), we would conclude that $\tilde{\chi}$ is $\mathcal{P}'$-probability-measure consistent. We will now show that the algorithm constraints, together with the constraint $\bar{U} \succeq 0$, do in fact guarantee that $U^{\tilde{\chi}} \succeq 0$.

Consider the submatrix $\hat{U}$ of $\bar{U}$ with rows for each $T(g)$ such that $T'(g) \neq \emptyset$. Each entry of the matrix $\hat{U}$ satisfies

$$\hat{U}(g, g^*) = \begin{cases} x^P[T(g \cup g^*)] = X[T'(g \cup g^*)] \;:\; |g \cup g^*| < k \\ 0 : |g \cup g^*| \geq k \end{cases} \tag{6.333}$$

by (6.201) and (6.199). Moreover, for each g, g^* with $|g \cup g^*| < k$ such that $T'(g \cup g^*) \neq \emptyset$, we have

$$X[T'(g \cup g^*)] = \hat{X}[T'(g \cup g^*)] = \tilde{\chi}[T'(g \cup g^*)] = U^{\tilde{\chi}}(g, g^*). \tag{6.334}$$

For each g, g^* such that $|g \cup g^*| < k$, but $T'(g \cup g^*) = \emptyset$ we have by (6.329),

$$X[T'(g \cup g^*)] = 0 = \chi[\emptyset] = \tilde{\chi}[\emptyset] = U^{\tilde{\chi}}(g, g^*). \tag{6.335}$$

(The second equality follows from the additivity property of signed measures). Finally, for each g, g^* such that $|g \cup g^*| \geq k$ we also have $T'(g \cup g^*) = \emptyset$ by (6.328), and therefore

$$\hat{U}(g, g^*) = 0 = \chi[\emptyset] = \tilde{\chi}[\emptyset] = U^{\tilde{\chi}}(g, g^*). \tag{6.336}$$

We conclude that $\hat{U} = U^{\tilde{\chi}}$. Thus since $\hat{U}$ is a submatrix of the positive semidefinite matrix $\bar{U}$, it follows that $U^{\tilde{\chi}} \succeq 0$ as well, and we now conclude that $\tilde{\chi}$ is consistent with some probability measure χ on $\mathcal{P}'$.

Note that for any finite measure $\mathcal{X}$ defined on $\mathcal{P}'$, and any collection of sets $\{W_1, \ldots, W_r\}$ in $\mathcal{P}'$,

$$\sum_{i=1}^{r} \mathcal{X}[W_i] =$$

$$\sum_{i=1}^{r} \mathcal{X}\left[\{y \in P' : \ y \text{ belongs to} \geq i \text{ sets from among } \{W_1, \ldots, W_r\}\}\right]. \tag{6.337}$$

This is true in greater generality as well, but to see why this is true in this case we need only note that for each zeta vector corresponding to a point $y \in P'$, which we will denote ζ^y, $\zeta^y[W_i] = 1$ iff $y \in W_i$. Thus if y belongs to j of the sets $\{W_1, \ldots, W_r\}$, then for exactly those j sets $\zeta^y[W_i] = 1$, and for each $i \leq j$ (and for no $i > j$),

$$\zeta^y\left[\{y \in P' : \ y \text{ belongs to} \geq i \text{ sets from among } \{W_1, \ldots, W_r\}\}\right] = 1. \tag{6.338}$$

So clearly ζ^y satisfies (6.337), and therefore each finite measure, which is a nonnegative linear combination of the ζ^y, must satisfy (6.337) as well.

Consider now that as a probability measure on $\mathcal{P}'$, χ must satisfy that for each $i \geq 0$,

$$\chi\left[\{y \in P' : \ y \text{ belongs to at least } i \text{ sets of the form } B'_j\}\right] \leq 1 \tag{6.339}$$

and by (6.328), where $i \geq k$ we have

$$\chi\left[\{y \in P' : \ y \text{ belongs to at least } i \text{ sets of the form } B'_j\}\right] = \chi[\emptyset] = 0. \tag{6.340}$$

Thus by (6.337), χ must satisfy

$$\sum_{i=1}^{t} \chi[B'_i] =$$

$$\sum_{i=1}^{t} \chi\left[\{y \in P' : \ y \text{ belongs to at least } i \text{ sets of the form } B'_j\}\right] =$$

$$\sum_{i=1}^{k-1} \chi\left[\{y \in P' : \ y \text{ belongs to at least } i \text{ sets of the form } B'_j\}\right] \leq$$

$$k - 1. \tag{6.341}$$

(A simpler, though arguably less instructive proof of (6.341) would be obtained by noting that each zeta vector ζ^y, $y \in P'$ must satisfy $\sum_{i=1}^{t} \zeta^r[B'_i] \leq k - 1$, as no point in P' can violate more than $k-1$ blocks. Thus χ, as a convex combination of those zeta vectors, must also satisfy this constraint.)

Note now that for each $i = 1, \ldots, t$, $B'_i = T'(g)$ where $g = \{i\}$, so X has a coordinate for each B'_i. If $B'_i \neq \emptyset$, then

$$X[B'_i] = \hat{X}[B'_i] = \tilde{\chi}[B'_i] = \chi[B'_i] \tag{6.342}$$

and if $B'_i = \emptyset$, so that $\hat{X}$ has no B'_i coordinate, we still have $X[B'_i] = 0 = \chi(B'_i)$ by (6.329). Thus by (6.341),

$$\sum_{i=1}^{t} X[B'_i] = \sum_{i=1}^{t} \chi[B'_i] \leq k - 1. \tag{6.343}$$

Now applying algorithm constraints (6.196) and (6.203), we have

$$X[B'_i] = x^P[B_i] \geq \sum_{j \in S_i} x[N_j^P] - |S_i| + 1 = \tag{6.344}$$

$$\sum_{j \in S_i} (1 - x[M_j^P]) - |S_i| + 1 = \tag{6.345}$$

$$|S_i| - \sum_{j \in S_i} x[M_j^P] - |S_i| + 1 = \tag{6.346}$$

$$1 - x(S_i). \tag{6.347}$$

So by (6.343),

$$k - 1 \geq \sum_{i=1}^{t} X[B'_i] \geq t - \sum_{i=1}^{t} x(S_i) \Rightarrow \tag{6.348}$$

$$\sum_{i=1}^{t} x(S_i) \geq t - k + 1. \ \square \tag{6.349}$$

As we noted at the beginning of the section, the constraints of the form (6.316) described by Theorem 6.29 are quite a considerable generalization of the clique constraints of the stable set problem. It is the fact that the common factors capture a great deal of the problem's structure that is largely responsible for this result. Similar "block of variables" versions of the odd antihole, odd hole and odd wheel constraints can also be shown to hold, and for the latter two the constraints can be obtained quickly even without enforcing positive semidefiniteness.

We will show now that even for the case of $k = 2$, the N^{++} rank of the constraint (6.316) rises with n. In the worst case its rank will be $\geq \lfloor \frac{n}{3} \rfloor$. To make the presentation simpler, we will also assume that P is of the form $P = \bigcap_{i=1}^{m} \bigcup_{l \in A_i} Y_l$ (with each $A_i \subseteq \{1, \ldots, n\}$), but this assumption does not change anything substantial.

For the purposes of the following theorem, recall from Definition 1.2 that where $P \subseteq \{0,1\}^n$, then $K(P)$ is defined as

$$\{y \in \{0,1\}^{n+1} : y_0 = 1, (y_1, \ldots, y_n) \in P\} \tag{6.350}$$

and that where $Q \subseteq [0,1]^n$, then $\bar{K}(Q)$ is the homogenized version of Q, i.e.

$$\bar{K}(Q) = Cone(\{x \in [0,1]^{n+1} : x_0 = 1, (x_1, \ldots, x_n) \in Q\}). \tag{6.351}$$

Recall also that

$$\bar{K}(Conv(P)) = Cone(K(P)). \tag{6.352}$$

Theorem 6.30 *Let $S_1, \ldots, S_t$ be disjoint subsets of $\{1, \ldots, n\}$. Let*

$$P = \{y \in \{0,1\}^n : y(S_i) + y(S_j) \geq 1, \ \forall i = 1, \ldots, n, \ j = 1, \ldots, n, \ i \neq j\} \tag{6.353}$$

and let

$$\bar{P} = \{y \in [0,1]^n : y(S_i) + y(S_j) \geq 1, \ \forall i = 1, \ldots, n, \ j = 1, \ldots, n, \ i \neq j\}. \tag{6.354}$$

Let γ be the cardinality of the second-largest set S_i. Then for all positive integers $l < \min\{t-2, \gamma\}$, the valid constraint

$$\sum_{i=1}^{t} y(S_i) \geq (t-1)y_0 \tag{6.355}$$

on $K(P)$ is not valid for the set $(N^{++})^l(\bar{K}(\bar{P}))$.

Proof: We will construct a measure χ on $\mathcal{A}$ (the subset algebra of $\{0,1\}^n$), for which the vector

$$(\chi[\{0,1\}^n], \chi[Y_1], \ldots, \chi[Y_n]) \tag{6.356}$$

violates the constraint

$$\sum_{i=1}^{t} \left(\sum_{j \in S_i} \chi[Y_j] \right) \geq (t-1)\chi[\{0,1\}^n] \tag{6.357}$$

but for which each partial sum measure χ^Q satisfies

$$\sum_{u \in S_i \cup S_j} \chi^Q[Y_u] \geq \chi^Q[\{0,1\}^n], \ \forall i \neq j \tag{6.358}$$

for each Q of the form

$$Q = \bigcap_{h=1}^{w} M_h \tag{6.359}$$

where each $M_h \in \{Y_1, \ldots, Y_n, N_1, \ldots, N_n\}$, and $w < \min\{t-2, \gamma\}$. By Remark 3.68 and Definition 4.29 it will then follow that for every $w < \min\{t-2, \gamma\}$, the vector defined by $(\chi[\{0,1\}^n], \chi[Y_i], \ldots, \chi[Y_n])$, which violates (6.357) and therefore does not belong to $Cone(K(P))$, nevertheless belongs to $(N^{++})^w(K(\bar{P}))$, which proves the theorem.

The construction is as follows. Without loss of generality, let us assume that $S_1, \ldots, S_t$ are arranged in order of decreasing cardinality. For each $i = 1, \ldots, t$, let j_i be a particular element in S_i. Define the atom

$$r = \bigcap_{j \in S_1 \cup S_2} N_j \cap \bigcap_{i=3}^{t} Y_{j_i} \cap \bigcap_{j \in (S_1 \cup S_2 \cup \bigcup_{i=3}^{t} \{j_i\})^c} N_j = \bigcap_{i=3}^{t} Y_{j_i} \cap \bigcap_{j \in (\bigcup_{i=3}^{t} \{j_i\})^c} N_j \tag{6.360}$$

and assign $\chi(r) = 1$. For each t-tuple,

$$H = (l, h_1, h_2, \ldots, h_{l-1}, h_{l+1}, \ldots, h_t) \tag{6.361}$$

where each $h_i \in S_i$, and $l \in \{3, \ldots, t\}$, define the atom

$$s^H = \bigcap_{\substack{i=1 \\ i \neq l}}^{t} Y_{h_i} \cap \bigcap_{j \in (\bigcup_{\substack{i=1 \\ i \neq l}}^{t} \{h_i\})^c} N_j \tag{6.362}$$

and assign $\chi(s^H) = 1$, and assign all other atoms measure zero. Recall that the measure of a set is the sum of the measures of the atoms that the set contains. For each atom s^H, for each $i = 1, \ldots, t,\ i \neq l$, there is exactly one $j \in S_i$ such that $s^H \subseteq Y_j$. Thus each atom s^H contributes $t-1$ units to either side of (6.357). The atom r, however, is contained in only $t-2$ of the sets Y_j, and so r contributes $t-1$ to the right side and only $t-2$ to the left side, so χ indeed violates (6.357). For every $\{i, j\} \neq \{1, 2\}$, every s^H such that $s^H \subseteq Q$, as well as r (if $r \subseteq Q$), contribute no more to the right side of (6.358) than to the left, so all partial sums satisfy all of these constraints. (Recall that atoms that are not subsets of Q do not contribute to (6.358) at all.) Furthermore, even where $\{i, j\} = \{1, 2\}$, for any Q such that some $s^H \subseteq Q$, s^H contributes 2 units to the left side of (6.358) and one unit to the right. So even if $r \subset Q$ as well, and therefore r contributes one unit to the right side and none to the left, the partial sum χ^Q continues to satisfy the constraint (6.358). Thus all we need to show is that for all

$$Q = \bigcap_{h=1}^{w} M_h,\ w < \min\{t-2, \gamma\},\ r \subseteq Q \tag{6.363}$$

there is some $s^H \subset Q$. Given such a Q, $r \subseteq Q$ implies that there exist subsets

$$J \subseteq \bigcup_{i=3}^{t}\{j_i\} \text{ and } L \subseteq \left(\bigcup_{i=3}^{t}\{j_i\}\right)^c, \quad |J|+|L|=w \tag{6.364}$$

such that

$$Q = \bigcap_{j\in L} N_j \cap \bigcap_{j\in J} Y_j. \tag{6.365}$$

Since $w < \gamma$ there must be some $h_1 \in S_1$ and $h_2 \in S_2$ such that neither belong to L, and since $w < t-2$ there must be some $l \in \{3,\ldots,t\}$ such that $l \notin J$. Define

$$H = (l, h_1, h_2, j_3, \ldots, j_{l-1}, j_{l+1}, \ldots, j_t). \tag{6.366}$$

Then $s^H \subset Y_j$ for each $j \in J$, and $h_1, h_2 \notin L$, together with $L \cap \bigcup_{i=3}^{t}\{j_i\} = \emptyset$ implies moreover that $s^H \subset N_j$ for each $j \in L$. Thus $s^H \subset Q$ and the theorem is proven. □

For the case

$$t = n - 2\lfloor\frac{n}{3}\rfloor + 2, \quad |S_1| = |S_2| = \lfloor\frac{n}{3}\rfloor, \ |S_i| = 1, \ i = 3,\ldots,t \tag{6.367}$$

the theorem gives us a lower bound of $\lfloor\frac{n}{3}\rfloor$ for the N^{++} rank of (6.355).

6.7 All Configurations Forbidden

In this section we will consider the set P given by

$$P = \{y \in \{0,1\}^n : \sum_{j\in J} y_j + \sum_{j\in J^c}(1-y_j) \geq \frac{1}{2}, \ \forall J \subseteq \{1,\ldots,n\}\}. \tag{6.368}$$

This set was first introduced in [CCH89], and was analyzed in [CD01], [GT01] and in [Lau01]. The set is clearly empty, as every possible configuration is forbidden. Nevertheless, the N^+ rank of the linear relaxation of this problem is n, as shown in [CD01] (and it is not hard to show that its N^{++} rank is n as well), and it is conjectured in [Lau01] that its Lasserre rank is $n-1$. Changing the right hand side to ≥ 1 reduces the bound for N^{++} only by 1, but it makes the problem suitable for the application of our algorithms. We will show that the first version of the common factor algorithm determines the set to be empty by level $k = 3$.

Theorem 6.31 *Let*

$$P = \{y \in \{0,1\}^n : \sum_{j\in J} y_j + \sum_{j\in J^c}(1-y_j) \geq 1, \ \forall J \subseteq \{1,\ldots,n\}. \tag{6.369}$$

Then the system of constraints enforced by Version 1 of the algorithm at level 3*, applied to* P*, is infeasible.*

Proof: We will show that the algorithm constraints at level 3 will require in this case that the column $x^P = 0$. But constraint (6.196) demands $x^P[P] = 1$, and so there can be no feasible solutions.

Observe first that every intersection

$$\bigcap_{l\in J} Y_l^P \cap \bigcap_{l\in \bar{J}} N_l^P \tag{6.370}$$

for which J and $\bar{J}$ partition $\{1,\ldots,n\}$ is a forbidden configuration. Thus every intersection

$$\bigcap_{l\in J} Y_l^P \cap \bigcap_{l\in \bar{J}} N_l^P \tag{6.371}$$

for which J and $\bar{J}$ are disjoint subsets of $\{1,\ldots,n\}$ with $|J| + |\bar{J}| \leq n-1$, belongs to $\mathcal{C}_2$. In particular, for each $l = 1,\ldots,n$, the expression $N_l^P \in \mathcal{C}_2$, and for every disjoint $J, \bar{J} \subseteq \{1,\ldots,n\}$ such that $|J| + |\bar{J}| \leq n-1$, every intersection

$$C = \bigcap_{l\in J} N_l^P \cap \bigcap_{l\in \bar{J}} Y_l^P \in \mathcal{C}_2^{-1} \subseteq \mathcal{C}_3^{-1}. \tag{6.372}$$

Consider now the case $|J|+|\bar{J}| = n-1$, $J\cap\bar{J} = \emptyset$, and $\{1,\ldots,n\}-(J\cup\bar{J}) = \{h\}$. Applying the algorithm constraints (as per Lemma 6.20),

$$\sum_{l\in J} x_{l'} + \sum_{l\in \bar{J}} x_{l''} + x_{h'} \geq x_0 \text{ and} \tag{6.373}$$

$$\sum_{l\in J} x_{l'} + \sum_{l\in \bar{J}} x_{l''} + x_{h''} \geq x_0 \tag{6.374}$$

to the column x^C yields $x_{h'}^C \geq x_0^C$ and $x_{h''}^C \geq x_0^C$. But $x_{h'}^C + x_{h''}^C = x_0^C$ (by (6.198)), which implies (by (6.197) and (6.201)) that

$$x_0^C \geq 2x_0^C \Rightarrow x_0^C = 0 \Rightarrow x^C = 0. \tag{6.375}$$

We will now prove by a backwards induction that for each $r = 0,\ldots,n-1$, for every intersection

$$C = \bigcap_{l\in J} N_l^P \cap \bigcap_{l\in \bar{J}} Y_l^P,\ |J| + |\bar{J}| = r \tag{6.376}$$

(with J and $\bar{J}$ disjoint), each type (1) column x^v of the matrix U (i.e. v is of the form (6.190)) for which $\delta(v) = \delta(C)$ satisfies $x^v = 0$. Recalling that by algorithm contraint (6.201), all matrix entries in type (1) columns associated with a common intersection of sets Y and N have the same value, the base case, $r = n-1$, has already been established. Assume now that the hypothesis holds for each $r \in \{t,\ldots,n-1\}$ for some $1 \leq t \leq n-1$, and consider the set

$$C_3^{-1} = \bigcap_{l\in J} N_l^P \cap \bigcap_{l\in \bar{J}} Y_l^P,\ |J| + |\bar{J}| = t-1 \tag{6.377}$$

(with J and $\bar{J}$ disjoint). Observe that $C_3^{-1} \in \mathcal{C}_3^{-1}$. Let $h \in \{1, \ldots, n\}$ be such that $h \notin J \cup \bar{J}$. Choose $N_h^P \in \mathcal{C}_2$, and apply algorithm constraint (6.210) to obtain that for each type (1) column x^v of the matrix for which $\delta(v) = \delta(C_3^{-1})$,

$$x^v = x^{C_3^{-1}} = x^{C_3^{-1} \cap N_h^P} + x^{C_3^{-1} \cap Y_h^P} = 0 + 0 = 0 \tag{6.378}$$

by hypothesis, which proves the induction. Thus where $r = 0$ then the set C defined by (6.376) is P, and we conclude that $x^P = 0$. □

6.8 Further Work

There are a number of avenues that call for further study. The first and most obvious is the question of what other types of partitioning schemes can be used, and can the choice of partitioning scheme be tailored to the problem? What other results can be obtained by partitioning (or covering - we noted already that strict partitions are not necessary) over clever choices of sets?

Is there a way to partition effectively using sets that are "not nice", in the terminology used at the beginning of Chapter 5, to develop algorithms to handle arbitrary feasible sets of the form $P = \{y \in \{0,1\}^n : Ay \geq b\}$? Our use of the sets $C^{->r}$ perhaps indicates that some use can be made even out of sets that have no "nice" characterization.

We indicated in Chapter 3 that the relationship between measures and convex hulls can be generalized to countably large feasible sets. This generalization can be pusued further.

Another point to consider is that the algorithms of the final two chapters did not make use of all of the machinery developed in Chapters 3 and 4. In particular they made only light use of positive semidefiniteness and measure theoretic constraints. These could be used to greater advantage perhaps in the context of an attempt to ensure $\mathcal{P}$-signed measure consistency, and a choice of sets that maximizes the effectiveness of results such as Lemma 4.4 and Theorem 4.10. Measure preserving operators also seem to be interesting and potentially useful objects.

Bibliography

[B74] E. Balas, Disjunctive Programs: Properties of the convex hull of feasible points, MSRR No. 348, Carnegie Mellon University (Pittsburgh, PA, 1974).

[B79] E. Balas, Disjunctive Programming. *Annals of Discrete Math.* **5** (1979) 3–51.

[BCC93] E. Balas, S. Ceria and G. Cornuéjols, A lift-and-project cutting plane algorithm for mixed 0-1 programs, *Mathematical Programming* **58** (1993), 295 – 324.

[BN89] E. Balas and S.M. Ng, On the set covering polytope: I. All the facets with coefficients in $\{0, 1, 2\}$, *Mathematical Programming* **45** (1989), 1 – 20.

[BH01] E. Boros and P.L. Hammer, Pseudo-Boolean Optimization, *RUTCOR Research Report* 48-2001, Rutgers University (2001).

[BZ02] D. Bienstock and M. Zuckerberg, Subset Algebra Lifting Methods for 0-1 Integer Programming, *CORC Technical Report 2002-01*, to appear in *SIAM J. on Optimization.*

[BZ03] D. Bienstock and M. Zuckerberg, Set Covering Problems and Chvátal-Gomory Cuts, *CORC Technical Report 2003-01.*

[CCH89] V. Chvátal, W. Cook and M. Hartmann, On cutting-plane proofs in combinatorial optimization, *Linear Algebra and its Applications* **114** (1989), 455 – 499.

[CD01] W. Cook and S. Dash, On the matrix-cut rank of polyhedra, *Mathematics of Operations Research* **26** (2001), 19 – 30.

[CL01] G. Cornuéjols and Y. Li, On the rank of mixed 0-1 polyhedra. In K. Aardal and A.M.H. Gerards (eds.), IPCO 2001, *Lecture Notes in Computer Science* **2081** (2001), 71 – 77.

[F99] G. B. Folland, *Real Analysis*, Wiley (1999).

[GLS81] M. Grötschel, L. Lovász and A. Schrijver, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag (1988)

[GT01] M.X. Goemans and L. Tunçel, When does the positive semidefiniteness constraint help in lifting procedures, *Mathematics of Operations Research* **26** (2001), 796-815.

[H00] J. Hooker, Logic, Optimization and Constraint Programming, *Informs J. on Computing* (2002).

[Las01] J. B. Lasserre, An explicit exact SDP relaxation for nonlinear 0-1 programs, in *Lecture Notes in Computer Science* (K. Aardal and A.M.H. Gerards, eds.) (2001), 293-303.

[LS91] L. Lovász and A. Schrijver, Cones of matrices and set-functions and 0-1 optimization, *SIAM J. on Optimization* **1** (1991), 166-190.

[Lau01] M. Laurent, A Comparison of the Sherali-Adams, Lovász-Schrijver and Lasserre Relaxations for 0-1 Programming, Technical Report PNA-R0108, CWI (2001).

[Ro64] G.-C. Rota, On the foundations of combinatorial theory I. Theory of Möbius functions, *Z. Wahrsch. Verw. Gebiete* **2** (1964), 340-368.

[Ru64] W. Rudin, *Principles of Mathematical Analysis*, McGraw Hill (1964).

[SA90] S. Sherali and W. Adams, A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems, *SIAM J. on Discrete Mathematics* **3** (1990), 411-430.

[S86] A. Schrijver, *Theory of Linear and Integer Programming*, Wiley (1986).

[W98] L.A. Wolsey, *Integer Programming*, Wiley (1998).