System Identification via Parabolic Relaxation

Adnan Nasir, *Student Member, IEEE,* Ramtin Madani, *Member, IEEE,* and Ali Davoudi, *Senior Member, IEEE*

***Abstract*—We offer a novel convex relaxation method to solve the equations governing the behavior of linear and nonlinear dynamical systems. This relaxation is used for the purpose of system identification. We demonstrate significant improvements in the overall scalability in comparison with the state-of-the-art convex relaxations. When the proposed relaxation is inexact, a sequential algorithm is provided which converges within a finite number of rounds. Theoretical guarantees and simulation results demonstrate the effectiveness of the proposed approach.**

***Index Terms*—Computational methods, Optimization, Identification.**

I. Introduction

SYSTEM identification methods aim to devise a dynamical model from system observations (e.g., input-output data) [1]. Linear systems employ statistical methods; e.g., maximum likelihood, Bayesian, cross validation, variance, regularization techniques, and least-squares methods [2]. More comprehensive methods [3]–[5], that handle both linear and nonlinear systems, could be prone to local minimums and have convergence issues. One could pursue convex optimization as a mean to identify dynamical systems [6], [7]. General-purpose convex relaxation methods; e.g., semidefinite programming (SDP) [8] and second-order cone programming (SOCP) relaxations [9], primarily rely on a process referred to as *lifting*. Lifting introduces new auxiliary variables by lift-and-project procedure performed to relax the problem [10]. The lifting mechanism severely limits the adoption of these algorithms due to the curse of dimensionality.

Alternatively, we formulate system identification as an optimization problem, and provide a scalable solution via a novel sequential parabolic relaxation method. The proposed approach relies on a modest number of auxiliary variables, and effectively addresses the computational issues caused by the lifting process in existing relaxation techniques. Generally, the state-of-the-art estimation methods require full input and state information, which may not be readily available. Here, we provide theoretical guarantees for an optimal solution within a short period of time using only a sample of input and state information. We validate our proposed method using linear and nonlinear benchmark systems, including the nonlinear dynamics of a permanent magnet synchronous machine (PMSM).

The remainder of the paper is organized as follows: section II, will elaborate the notations used in the paper. Section III defines the problem statement, and discusses the merits of parabolic relaxation against SDP and SOCP, for linear system identification. Section IV highlights the theoretical guarantees, and Section V extends the proposed approach to nonlinear systems. Section VI offers numerical experiments and, finally, Section VII concludes the paper.

This work is supported, in part, by the National Science Foundation under Grants 1509804 and 1809454. We also want to thank A. P. Yadav for his help in machine modeling. Authors are with the Department of Electrical Engineering, University of Texas at Arlington, TX, USA. (e-mail: adnan.nasir@mavs.uta.edu; ramtin.madani@uta.edu; davoudi@uta.edu).

II. Notations and Terminologies

The vectors and matrices are, respectively, shown by lower-case and upper-case bold letters. Symbols $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times m}$, respectively, denote the set of real scalars, real vectors of size n, and real matrices of size $n\times m$. The set of real $n\times n$ symmetric matrices and positive semidefinite matrices are shown with $\mathbb{S}_n$ and $\mathbb{S}_n^+$, respectively. Notation $\mathbf{A}\succeq 0$ ($\mathbf{A}\succ 0$) means that $\mathbf{A}$ is positive-semidefinite (positive definite). $\mathrm{tr}\{.\}$, $(.)^\top$, and $\mathrm{rank}\{.\}$, respectively, denote the trace, transpose, and rank operators. $\|.\|_p$ refers tonorm of a vector or a matrix. $|.|$ indicates the absolute value operator.

III. Identification of Linear Systems

A. Problem Formulation

Consider a linear dynamical system:

$$\mathbf{x}[t+1]=\mathbf{A}\mathbf{x}[t]+\mathbf{B}\mathbf{u}[t] \qquad t\in\mathcal{T} \tag{1a}$$
$$\mathbf{y}[t]=\mathbf{C}\mathbf{x}[t]+\mathbf{D}\mathbf{u}[t] \qquad t\in\mathcal{T}. \tag{1b}$$

$\mathcal{T}\triangleq\{1,2,\dots,\tau\}$ is a discrete time horizon, and $\{\mathbf{x}[t]\in\mathbb{R}^n\}_{t\in\mathcal{T}\cup\{\tau+1\}}$, $\{\mathbf{y}[t]\in\mathbb{R}^m\}_{t\in\mathcal{T}}$, and $\{\mathbf{u}[t]\in\mathbb{R}^o\}_{t\in\mathcal{T}}$ denote the state, observation, and system input vectors, respectively.

Assume that system matrices $(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D})$ and the state vectors at snapshots $\mathcal{T}'\triangleq\{t_1,t_2,\dots,t_{\tau'}\}\subseteq\mathcal{T}\cup\{\tau+1\}$ are unknown. We seek to determine the unknowns based on system observation, system input, and limited knowledge of state vectors at times $(\mathcal{T}\cup\{\tau+1\})\setminus\mathcal{T}'$. This problem can be formulated as follows:

$$\begin{aligned}\text{find}\quad & \mathbf{A}\in\mathbb{R}^{n\times n},\ \mathbf{B}\in\mathbb{R}^{n\times o},\ \mathbf{C}\in\mathbb{R}^{m\times n},\ \mathbf{D}\in\mathbb{R}^{m\times o},\\ & \{\mathbf{x}[t]\in\mathbb{R}^n\}_{t\in\mathcal{T}'} && \text{(2a)}\\ \text{subject to}\quad & \mathbf{x}[t+1]=\mathbf{A}\mathbf{x}[t]+\mathbf{B}\mathbf{u}[t] \qquad t\in\mathcal{T} && \text{(2b)}\\ & \mathbf{y}[t]=\mathbf{C}\mathbf{x}[t]+\mathbf{D}\mathbf{u}[t] \qquad t\in\mathcal{T} && \text{(2c)}\end{aligned}$$

Problem (2a) – (2c) is non-convex, due to the bi-linear terms $\mathbf{A}\mathbf{x}[t]$ and $\mathbf{C}\mathbf{x}[t]$. One approach for tackling the non-convexity of dynamical system equations is convex relaxation or approximation [11], which has been used in the parameter estimation [6], [12]. The number of feasible connected components induced by dynamical system equations can grow exponentially, which renders local search methods intractable [12]. Herein, our focus in on the scalability of convex relaxation by introducing a computationally-efficient parabolic relaxation to tackle constraints of the form (2b) and (2c).

B. State-of-the-Art Relaxations

We first cover two forms of the common-practice semidefinite programming (SDP) and second-order cone programming (SOCP) relaxations [13]–[15] and show that each suffer from major drawbacks compared to the proposed method.

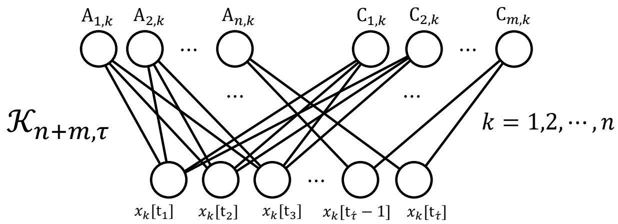

Figure 1. The bipartite sparsity pattern of the constraint (5).

1) Vector Formulation: Let the vector sets $\{\boldsymbol{a}_k \in \mathbb{R}^n\}_{k=1}^n$, $\{\boldsymbol{b}_k \in \mathbb{R}^o\}_{k=1}^n$, $\{\boldsymbol{c}_k \in \mathbb{R}^n\}_{k=1}^m$, and $\{\boldsymbol{d}_k \in \mathbb{R}^o\}_{k=1}^m$ represent the columns of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$, respectively, i.e.,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix}^\top \qquad \mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_n \end{bmatrix}^\top \tag{3a}$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_m \end{bmatrix}^\top \qquad \mathbf{D} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \dots & \mathbf{d}_m \end{bmatrix}^\top \tag{3b}$$

To formulate the standard conic relaxations of the problem (2a) – (2c), one can cast the identification problem with respect to the new vector and matrix variables:

$$\mathbf{h} \triangleq \begin{bmatrix}\mathbf{a}_1^\top & \mathbf{a}_2^\top & \dots & \mathbf{a}_n^\top & \mathbf{c}_1^\top & \mathbf{c}_2^\top & \dots & \mathbf{c}_n^\top \\ & & & & \mathbf{x}^\top[\mathsf{t}_1] & \dots & \mathbf{x}^\top[\mathsf{t}_{\tau'}]\end{bmatrix}^\top \in \mathbb{R}^{n(m+\tau)} \tag{4a}$$

$$\mathbf{H} \triangleq \mathbf{h}\mathbf{h}^\top \tag{4b}$$

using which the nonlinear constrains (2b) and (2c) can be cast linearly [16]. This approach is regarded as *lifting*, as we are transitioning to a higher-dimensional space with respect to the new matrix variable $\mathbf{H}$. The relations between the pair $(\mathbf{H}, \mathbf{h})$ can then be implicitly imposed as

$$\begin{bmatrix} \mathbf{H} & \mathbf{h} \\ \mathbf{h} & 1 \end{bmatrix} \in \mathcal{H}, \tag{5}$$

where $\mathcal{H} \in \mathbb{S}_{n(m+\tau)}$ is an appropriate convex set depending on the choice of relaxation / approximation [16].

The main drawback of the aforementioned approach is the curse of dimensionality caused by lifting, even if the present sparsity of the problem is fully leveraged. Let $\mathcal{K}$ denote the sparsity graph of the quadratic constraints (2b) – (2c), whose every vertex corresponds to an element of the vector $\mathbf{h}$, and every edge corresponds a bilinear term:

$$\mathsf{A}_{ik}\, \mathsf{x}_k[\mathsf{t}] \qquad \forall (i,k,\mathsf{t}) \in \mathcal{N} \times \mathcal{N} \times \mathcal{T}' \tag{6a}$$

$$\mathsf{C}_{jk}\, \mathsf{x}_k[\mathsf{t}] \qquad \forall (j,k,\mathsf{t}) \in \mathcal{M} \times \mathcal{N} \times \mathcal{T}' \tag{6b}$$

adding up to a total of $n(n+m)\tau'$ edges, where $\mathcal{N} = \{1, \dots, n\}$ and $\mathcal{M} = \{1, \dots, m\}$. More precisely each pair of elements in $\mathbf{h}$ are connected, if and only if their product appears in either (2b) or (2c). As demonstrated in Figure 1, it can be easily observed that

$$\mathcal{K} = \underbrace{\mathcal{K}_{n+m,\tau'} \cup \mathcal{K}_{n+m,\tau'} \cup \dots \cup \mathcal{K}_{n+m,\tau'}}_{n} \tag{7}$$

where $\mathcal{K}_{n+m,\tau'}$ denotes the complete bipartite graph with partitions of size $n+m$ and τ'. The complexity of solving cone programming relaxations is closely related to the sparsity graph of the problem [17], [18]. As shown in Table I, in the case of SOCP relaxation for problem (2a) – (2c), one can leverage the sparsity of $\mathcal{K}$ by incorporating only those elements of $\mathbf{H}$ that correspond to existing edges amounting to $n(n+m)\tau'$ auxiliary variables.

The curse of dimensionality caused by lifting is further pronounced in the case of SDP relaxation that requires the incorporation of every element in $\mathbf{H}$ corresponding to the edges of an arbitrary chordal extension (Table I). In the reminder of this section, we discuss an alternative conic relaxation for which a smaller number of variables are needed.

2) Matrix Formulation: Another approach for tackling problems of the form (2a) – (2c) through convex relaxation is the matrix formulation. To this end, constraints (2b) and (2c) can be cast in the following form:

$$\begin{bmatrix} \bar{\mathbf{A}} & \big[\mathbf{x}[\mathsf{t}_1{+}1]\dots\mathbf{x}[\mathsf{t}_{\tau'}{+}1]\big] - \mathbf{BU} & \mathbf{A} \\ * & \bar{\mathbf{X}} & \mathbf{X}^\top \\ * & * & \boldsymbol{I}_{n\times n} \end{bmatrix} \in \mathcal{C}_1 \tag{8a}$$

$$\begin{bmatrix} \bar{\mathbf{C}} & \big[\mathbf{y}[\mathsf{t}_1]\dots\mathbf{y}[\mathsf{t}_{\tau'}]\big] - \mathbf{DU} & \mathbf{C} \\ * & \bar{\mathbf{X}} & \mathbf{X}^\top \\ * & * & \boldsymbol{I}_{n\times n}. \end{bmatrix} \in \mathcal{C}_2 \tag{8b}$$

where $\mathbf{X} \triangleq \big[\mathbf{x}[\mathsf{t}_1]\dots\mathbf{x}[\mathsf{t}_{\tau'}]\big]$ and $\mathbf{U} = \big[\mathbf{u}[\mathsf{t}_1]\dots\mathbf{u}[\mathsf{t}_{\tau'}]\big]$, and $\bar{\mathbf{A}}$, $\bar{\mathbf{C}}$, and $\bar{\mathbf{X}}$ are auxiliary variables accounting for $\mathbf{A}\mathbf{A}^\top$, $\mathbf{C}\mathbf{C}^\top$ and $\mathbf{X}^\top\mathbf{X}$, respectively. Having

$$\mathcal{C}_1 = \{\boldsymbol{W} \in \mathbb{S}_{2n+\tau'} \mid \boldsymbol{W} \succeq 0, \quad \operatorname{rank}\{\boldsymbol{W}\} = n\} \tag{9a}$$

$$\mathcal{C}_2 = \{\boldsymbol{W} \in \mathbb{S}_{m+\tau'+n} \mid \boldsymbol{W} \succeq 0, \quad \operatorname{rank}\{\boldsymbol{W}\} = n\} \tag{9b}$$

makes the two formulations (2a) – (2c) and (8a) – (8b) equivalent. Additionally, various convex relaxations/approximation can be formulated via different choices for $\mathcal{C}_1 \subseteq \mathbb{S}_{2n+\tau'}$ and $\mathcal{C}_2 \subseteq \mathbb{S}_{m+\tau'+n}$, e.g., the common-practice SDP and SOCP relaxations. Next, we pursue an alternative approach with far less complexity, involving convex quadratic constraints only.

C. Parabolic Relaxation

To formulate parabolic relaxation of the problem (2a) – (2c), we need to introduce $n + m + \tau'$ auxiliary variables:

- Define $\bar{\mathbf{a}} \in \mathbb{R}^n$ as the variable whose k-th element represents $\|\mathbf{a}_k\|_2^2$,
- Define $\bar{\mathbf{c}} \in \mathbb{R}^m$ as the variable whose k-th element represents $\|\mathbf{c}_k\|_2^2$,
- Define $\bar{\mathbf{x}} \in \mathbb{R}^{\tau'}$ as the variable whose k-th element represents $\|\mathbf{x}[\mathsf{t}_k]\|_2^2$,

Then, problem (2a) – (2c) can be reformulated as:

$$\text{find} \quad \bar{\mathbf{a}} \in \mathbb{R}^n,\ \bar{\mathbf{c}} \in \mathbb{R}^m,\ \bar{\mathbf{x}} \in \mathbb{R}^{\tau'},\ \{\mathbf{x}[\mathsf{t}] \in \mathbb{R}^n\}_{\mathsf{t}\in\mathcal{T}'},\ \{\mathbf{a}_k \in \mathbb{R}^n\}_{k=1}^n, \{\mathbf{b}_k \in \mathbb{R}^o\}_{k=1}^n, \{\mathbf{c}_k \in \mathbb{R}^n\}_{k=1}^m, \{\mathbf{d}_k \in \mathbb{R}^o\}_{k=1}^m, \tag{10a}$$

subject to

$$\bar{\mathsf{a}}_k + \bar{\mathsf{x}}_\mathsf{t} + 2\mathsf{x}_k[\mathsf{t}+1] - 2\mathbf{b}_k^\top\mathbf{u}[\mathsf{t}] \geq \|\mathbf{a}_k + \mathbf{x}[\mathsf{t}]\|_2^2 \quad \mathsf{t}\in\mathcal{T}',\ k\in\mathcal{N} \tag{10b}$$

$$\bar{\mathsf{a}}_k + \bar{\mathsf{x}}_\mathsf{t} - 2\mathsf{x}_k[\mathsf{t}+1] + 2\mathbf{b}_k^\top\mathbf{u}[\mathsf{t}] \geq \|\mathbf{a}_k - \mathbf{x}[\mathsf{t}]\|_2^2 \quad \mathsf{t}\in\mathcal{T}',\ k\in\mathcal{N} \tag{10c}$$

$$\bar{\mathsf{c}}_k + \bar{\mathsf{x}}_\mathsf{t} + 2\mathsf{y}_k[\mathsf{t}] - 2\mathbf{d}_k^\top\mathbf{u}[\mathsf{t}] \geq \|\mathbf{c}_k + \mathbf{x}[\mathsf{t}]\|_2^2 \quad \mathsf{t}\in\mathcal{T}',\ k\in\mathcal{M} \tag{10d}$$

$$\bar{\mathsf{c}}_k + \bar{\mathsf{x}}_\mathsf{t} - 2\mathsf{y}_k[\mathsf{t}] + 2\,\mathbf{d}_k^\top\mathbf{u}[\mathsf{t}] \geq \|\mathbf{c}_k - \mathbf{x}[\mathsf{t}]\|_2^2 \quad \mathsf{t}\in\mathcal{T}',\ k\in\mathcal{M} \tag{10e}$$

$$\bar{\mathsf{x}}_t = \|\mathbf{x}[\mathsf{t}]\|_2^2 \qquad \mathsf{t} \in \mathcal{T}' \tag{10f}$$

$$\bar{\mathsf{a}}_k = \|\mathbf{a}_k\|_2^2 \qquad k \in \mathcal{N} \tag{10g}$$

$$\bar{\mathsf{c}}_k = \|\mathbf{c}_k\|_2^2 \qquad k \in \mathcal{M} \tag{10h}$$

$$\mathsf{x}_k[\mathsf{t}+1] = \mathbf{a}_k^\top\mathbf{x}[\mathsf{t}] + \mathbf{b}_k^\top\mathbf{u}[\mathsf{t}] \qquad \mathsf{t} \in \mathcal{T}\backslash\mathcal{T}', \quad k \in \mathcal{N} \tag{10i}$$

$$\mathsf{y}_k[\mathsf{t}] = \mathbf{c}_k^\top\mathbf{x}[\mathsf{t}] + \mathbf{d}_k^\top\mathbf{u}[\mathsf{t}] \qquad \mathsf{t} \in \mathcal{T}\backslash\mathcal{T}', \quad k \in \mathcal{M} \tag{10j}$$

Table I
COMPLEXITY OF STATE-OF-THE-ART CONVEX RELAXATIONS VERSUS PARABOLIC RELAXATION.

	Max size of SDP constraints	**Number of SDP constraints**	**Number of new variables**
SDP relaxation of (4b)*	$(\max_{V\in\mathcal{V}} \lvert V\rvert) \times (\max_{V\in\mathcal{V}} \lvert V\rvert)$	$n\lvert\mathcal{V}\rvert$	$n(n+m)\tau'$
SOCP relaxation of (4b)	2×2	$n(n+m)\tau'$	$n(n+m)\tau'$
SDP relaxation of (8)	$(n+\max\{m,n\}+\tau')\times(n+\max\{m,n\}+\tau')$	2	$\binom{n+1}{2}+\binom{m+1}{2}+\binom{\tau'+1}{2}$
SOCP relaxation of (8)	2×2	$\binom{n+m+\tau'}{2}+\binom{2n+\tau'}{2}$	$\binom{n+1}{2}+\binom{m+1}{2}+\binom{\tau'+1}{2}$
Parabolic relaxation	Convex Quadratic	$2(n+m)\tau'+n+m+\tau'$	$n+m+\tau'$

* $(\mathcal{V},\mathcal{E})$ is an arbitrary tree decomposition of $\mathcal{K}_{n+m,\tau'}$

Problems (2a) – (2c) and (10a) – (10j) are equivalent since the pair of constraints (10b) and (10c) are equivalent to

$$\mathsf{x}_k[\mathsf{t}+1] \geq \mathbf{a}_k^\top \mathbf{x}[\mathsf{t}] + \mathbf{b}_k^\top \mathbf{u}[\mathsf{t}] - (\bar{\mathsf{x}}_t - \|\mathbf{x}[\mathsf{t}]\|_2^2) - (\bar{\mathsf{a}}_k - \|\mathbf{a}_k\|_2^2) \quad (11a)$$

$$\mathsf{x}_k[\mathsf{t}+1] \leq \mathbf{a}_k^\top \mathbf{x}[\mathsf{t}] + \mathbf{b}_k^\top \mathbf{u}[\mathsf{t}] + (\bar{\mathsf{x}}_t - \|\mathbf{x}[\mathsf{t}]\|_2^2) + (\bar{\mathsf{a}}_k - \|\mathbf{a}_k\|_2^2) \quad (11b)$$

and the pair of constraints (10d) and (10e) are equivalent to:

$$\mathsf{y}_k[\mathsf{t}] \geq \mathbf{c}_k^\top \mathbf{x}[\mathsf{t}] + \mathbf{d}_k^\top \mathbf{u}[\mathsf{t}] - (\bar{\mathsf{x}}_t - \|\mathbf{x}[\mathsf{t}]\|_2^2) - (\bar{\mathsf{c}}_k - \|\mathbf{c}_k\|_2^2) \quad (12a)$$

$$\mathsf{y}_k[\mathsf{t}] \leq \mathbf{c}_k^\top \mathbf{x}[\mathsf{t}] + \mathbf{d}_k^\top \mathbf{u}[\mathsf{t}] + (\bar{\mathsf{x}}_t - \|\mathbf{x}[\mathsf{t}]\|_2^2) + (\bar{\mathsf{c}}_k - \|\mathbf{c}_k\|_2^2). \quad (12b)$$

According to (10f), (10g) and (10h), the terms $(\bar{\mathsf{x}}_t - \|\mathbf{x}[\mathsf{t}]\|_2^2)$, $(\bar{\mathsf{a}}_k - \|\mathbf{a}_k\|_2^2)$, and $(\bar{\mathsf{c}}_k - \|\mathbf{c}_k\|_2^2)$ are zero, which means that:

$$\mathsf{x}_k[\mathsf{t}+1] = \mathbf{a}_k^\top \mathbf{x}[\mathsf{t}] + \mathbf{b}_k^\top \mathbf{u}[\mathsf{t}] \qquad \mathsf{t}\in\mathcal{T}', \quad k\in\mathcal{N} \quad (4a)$$

$$\mathsf{y}_k[\mathsf{t}] = \mathbf{c}_k^\top \mathbf{x}[\mathsf{t}] + \mathbf{d}_k^\top \mathbf{u}[\mathsf{t}] \qquad \mathsf{t}\in\mathcal{T}', \quad k\in\mathcal{M} \quad (4b)$$

that are equivalent to (2b) and (2c).

The primary motivation behind the use of formulation (10a) – (10h) is that the only non-convex constraints in this new formulation are (10a), (10b), and (10c). These non-convex constraints can be readily relaxed to

$$\bar{\mathsf{x}}_t \geq \|\mathbf{x}[\mathsf{t}]\|_2^2 \qquad \bar{\mathsf{a}}_k \geq \|\mathbf{a}_k\|_2^2 \qquad \bar{\mathsf{c}}_k \geq \|\mathbf{c}_k\|_2^2 \quad (5)$$

We regard this as the *parabolic relaxation* of the problem (2a) – (2c). Unlike the common-practice SDP and SOCP relaxations that require $O(n^2)$ number of auxiliary variables (even when the sparsity of the problem is leveraged), the proposed approach relies on a modest number of auxiliary variables, which is its primary strength.

If the relaxed problem has a unique solution for which the non-convex equalities (10b) – (10h) hold true, i.e., if

$$\bar{\mathsf{x}}_t^{\text{sol}} = \|\mathbf{x}^{\text{sol}}[\mathsf{t}]\|_2^2 \quad \bar{\mathsf{a}}_k^{\text{sol}} = \|\mathbf{a}_k^{\text{sol}}\|_2^2 \quad \bar{\mathsf{c}}_k^{\text{sol}} = \|\mathbf{c}_k^{\text{sol}}\|_2^2, \quad (6)$$

then we refer to the relaxation as *exact*. However, like any other convex relaxation, we may encounter cases where the equalities in (6) are not true and the relaxation is inexact. To remedy this issue, we introduce a family of penalty functions whose minimization can help address this issue.

D. Sequential Penalization

Motivated by [19], we minimize a penalty function to obtain feasible points for the problem (10a) – (10j). The *penalized parabolic relaxation* of the bi-linear problem (10a) – (10j) is

$$\underset{\substack{\mathbf{A},\,\mathbf{B},\,\mathbf{C},\,\mathbf{D},\\ \{\mathbf{x}[\mathsf{t}]\}_{\mathsf{t}\in\mathcal{T}'},\\ \bar{\mathbf{a}},\,\bar{\mathbf{c}},\,\bar{\mathbf{x}}}}{\text{minimize}} \quad \eta_1\big(\mathbf{1}_n^\top \bar{\mathbf{a}} - 2\,\mathrm{tr}\big\{[\mathbf{a}_1\ldots\mathbf{a}_n]^\top \check{\mathbf{A}}\big\} + \|\check{\mathbf{A}}\|_{\mathrm{F}}^2\big) + \eta_2\big(\mathbf{1}_n^\top \bar{\mathbf{c}} - 2\,\mathrm{tr}\big\{[\mathbf{c}_1\ldots\mathbf{c}_n]^\top \check{\mathbf{C}}\big\} + \|\check{\mathbf{C}}\|_{\mathrm{F}}^2\big) + \eta_3\sum\nolimits_{\mathsf{t}\in\mathcal{T}}\big(\bar{\mathsf{x}}_\mathsf{t} - 2\,\check{\mathbf{x}}^\top[\mathsf{t}]\,\mathbf{x}[\mathsf{t}] + \|\check{\mathbf{x}}[\mathsf{t}]\|_2^2\big) \quad (14a)$$

$$\text{subject to} \quad (10b),(10c),(10d),(10e),(5),(10i),(10j) \quad (14b)$$

$\eta_1,\eta_2,\eta_3 > 0$ are fixed regularization parameters, and $\check{\mathbf{A}} \in \mathbb{R}^{n\times n}$, $\check{\mathbf{C}} \in \mathbb{R}^{m\times n}$. $\{\check{\mathbf{x}}[\mathsf{t}] \in \mathbb{R}^n\}_{\mathsf{t}\in\mathcal{T}'}$ represent an arbitrary initial guess for the solution. We propose Algorithm 1 which starts from an initial point and solves a sequence of penalized relaxations until the ground-truth system is identified.

Algorithm 1 Sequential Penalized Parabolic Relaxation.

Require: $\check{\mathbf{A}} \in \mathbb{R}^{n\times n}$, $\check{\mathbf{C}} \in \mathbb{R}^{m\times n}$, $\{\check{\mathbf{x}}[\mathsf{t}] \in \mathbb{R}^n\}_{\mathsf{t}\in\mathcal{T}'}$, and $\eta_1,\eta_2,\eta_3 > 0$

repeat
  solve the penalized convex problem (14a) – (14b) to obtain $(\mathbf{A}^{\text{sol}},\mathbf{B}^{\text{sol}},\mathbf{C}^{\text{sol}},\mathbf{D}^{\text{sol}},\ \{\mathbf{x}^{\text{sol}}[\mathsf{t}]\}_{\mathsf{t}\in\mathcal{T}'},\ \bar{\mathbf{a}}^{\text{sol}},\bar{\mathbf{c}}^{\text{sol}},\bar{\mathbf{x}}^{\text{sol}})$.
  set $\check{\mathbf{A}} := \mathbf{A}^{\text{sol}}$, $\check{\mathbf{C}} := \mathbf{C}^{\text{sol}}$, and $\check{\mathbf{x}}[\mathsf{t}] := \mathbf{x}^{\text{sol}}[\mathsf{t}]$ for all $\mathsf{t}\in\mathcal{T}'$.
until stopping criterion is met.
return $(\mathbf{A}^{\text{sol}},\mathbf{B}^{\text{sol}},\mathbf{C}^{\text{sol}},\mathbf{D}^{\text{sol}})$

IV. THEORETICAL GUARANTEES

We first offer some preliminary notations and definition and, then, state the main theoretical result of the paper.

Definition 1. *For every $\mathbf{A}\in\mathbb{R}^{n\times n}$, $\mathbf{C}\in\mathbb{R}^{m\times n}$, and $\{\mathbf{x}[\mathsf{t}]\in\mathbb{R}^n\}_{\mathsf{t}\in\mathcal{T}'}$, define the Jacobian matrix*

$$\mathbf{J}\big(\mathbf{A},\mathbf{C},\mathbf{x}[\mathsf{t}_1],\ldots,\mathbf{x}[\mathsf{t}_{\tau'}]\big) \triangleq \begin{bmatrix} \frac{\partial\mathrm{vec}\{\mathbf{Q}\}}{\partial\mathrm{vec}\{\mathbf{A}\}} & \frac{\partial\mathrm{vec}\{\mathbf{Q}\}}{\partial\mathrm{vec}\{\mathbf{B}\}} & \frac{\partial\mathrm{vec}\{\mathbf{Q}\}}{\partial\mathrm{vec}\{\mathbf{C}\}} & \frac{\partial\mathrm{vec}\{\mathbf{Q}\}}{\partial\mathrm{vec}\{\mathbf{D}\}} & \frac{\partial\mathrm{vec}\{\mathbf{Q}\}}{\partial\mathbf{x}[\mathsf{t}_1]} \cdots \frac{\partial\mathrm{vec}\{\mathbf{Q}\}}{\partial\mathbf{x}[\mathsf{t}_{\tau'}]} \\ \frac{\partial\mathrm{vec}\{\mathbf{R}\}}{\partial\mathrm{vec}\{\mathbf{A}\}} & \frac{\partial\mathrm{vec}\{\mathbf{R}\}}{\partial\mathrm{vec}\{\mathbf{B}\}} & \frac{\partial\mathrm{vec}\{\mathbf{R}\}}{\partial\mathrm{vec}\{\mathbf{C}\}} & \frac{\partial\mathrm{vec}\{\mathbf{R}\}}{\partial\mathrm{vec}\{\mathbf{D}\}} & \frac{\partial\mathrm{vec}\{\mathbf{R}\}}{\partial\mathbf{x}[\mathsf{t}_1]} \cdots \frac{\partial\mathrm{vec}\{\mathbf{R}\}}{\partial\mathbf{x}[\mathsf{t}_{\tau'}]} \end{bmatrix} \quad (15)$$

where the matrix functions

$$\mathbf{Q}: \mathbb{R}^{n\times n}\times\mathbb{R}^{n\times o}\times\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times o}\times\mathbb{R}^{n}\times\ldots\times\mathbb{R}^{n} \quad\rightarrow\quad \mathbb{R}^{n\times\tau}$$

$$\mathbf{R}: \mathbb{R}^{n\times n}\times\mathbb{R}^{n\times o}\times\mathbb{R}^{m\times n}\times\mathbb{R}^{m\times o}\times\mathbb{R}^{n}\times\ldots\times\mathbb{R}^{n} \quad\rightarrow\quad \mathbb{R}^{m\times\tau}$$

are defined as:

$$\mathbf{Q}\big(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{x}[\mathsf{t}_1],\ldots,\mathbf{x}[\mathsf{t}_{\tau'}]\big) \triangleq \big[\mathbf{x}[2]\ldots\mathbf{x}[\tau+1]\big] - \mathbf{A}\big[\mathbf{x}[1]\ldots\mathbf{x}[\tau]\big] - \mathbf{B}\big[\mathbf{u}[1]\ldots\mathbf{u}[\tau]\big] \quad (16a)$$

$$\mathbf{R}\big(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{x}[\mathsf{t}_1],\ldots,\mathbf{x}[\mathsf{t}_{\tau'}]\big) \triangleq \big[\mathbf{y}[1]\ldots\mathbf{y}[\tau]\big] - \mathbf{C}\big[\mathbf{x}[1]\ldots\mathbf{x}[\tau]\big] - \mathbf{D}\big[\mathbf{u}[1]\ldots\mathbf{u}[\tau]\big] \quad (16b)$$

The point $(\mathbf{A},\mathbf{C},\{\mathbf{x}[\mathsf{t}]\}_{\mathsf{t}\in\mathcal{T}'})$ is said to satisfy the linear independence constraint qualification (LICQ) condition if the columns of $\mathbf{J}\big(\mathbf{A},\mathbf{C},\mathbf{x}[\mathsf{t}_1],\ldots,\mathbf{x}[\mathsf{t}_{\tau'}]\big)$ are linearly independent.

$$\mathbf{K}\big(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{x}[t_1],\dots,\mathbf{x}[t_{\tau'}],\bar{\mathbf{a}},\bar{\mathbf{c}},\bar{\mathbf{x}}\big)\triangleq$$
$$\begin{bmatrix}\mathrm{diag}\{\bar{\mathbf{a}}-\mathrm{diag}\{\mathbf{A}\mathbf{A}^\top\}\} & \mathbf{0}_{n\times m} & [\mathbf{x}[t_1+1]\dots\mathbf{x}[t_{\tau'}+1]]-\mathbf{A}[\mathbf{x}[t_1]\dots\mathbf{x}[t_{\tau'}]]-\mathbf{B}[\mathbf{u}[t_1]\dots\mathbf{u}[t_{\tau'}]]\\ * & \mathrm{diag}\{\bar{\mathbf{c}}-\mathrm{diag}\{\mathbf{C}\mathbf{C}^\top\}\} & [\mathbf{y}[t_1]\dots\mathbf{y}[t_{\tau'}]]-\mathbf{C}[\mathbf{x}[t_1]\dots\mathbf{x}[t_{\tau'}]]-\mathbf{D}[\mathbf{u}[t_1]\dots\mathbf{u}[t_{\tau'}]]\\ * & * & \mathrm{diag}\Big\{\bar{\mathbf{x}}-\mathrm{diag}\{[\mathbf{x}[t_1]\dots\mathbf{x}[t_{\tau'}]]^\top[\mathbf{x}[t_1]\dots\mathbf{x}[t_{\tau'}]]\}\Big\}\end{bmatrix} \tag{18}$$

Moreover, the singularity function $s:\mathbb{R}^{n\times n}\times\mathbb{R}^{m\times n}\times\mathbb{R}^n\times\dots\times\mathbb{R}^n\to\mathbb{R}$ *is defined as*

$$s\big(\mathbf{A},\mathbf{C},\mathbf{x}[t_1],\dots,\mathbf{x}[t_{\tau'}]\big)\triangleq\begin{cases}\sigma_{\min}\{\mathbf{J}(\mathbf{A},\mathbf{C},\mathbf{x}[t_1],\dots,\mathbf{x}[t_{\tau'}])\} & \mathbf{J}\ \textit{is full row rank}\\ 0 & \textit{otherwise,}\end{cases}$$

where $\sigma_{\min}$ *denotes the smallest singular value operator.*

The next theorem states that if the initial guess of the penalized convex relaxation (14a) – (14b) is sufficiently close to the ground truth, then the relaxation is exact and the solution can be recovered using the penalized convex relaxation.

Theorem 1. *Let* $(\mathring{\mathbf{A}},\mathring{\mathbf{B}},\mathring{\mathbf{C}},\mathring{\mathbf{D}},\{\mathring{\mathbf{x}}[t]\}_{t\in\mathcal{T}'})$ *denote a solution for the problem (2a) – (2c) that satisfies the LICQ condition. If*

$$\sqrt{\eta_1^2\|\mathring{\mathbf{A}}-\check{\mathbf{A}}\|_{\mathrm{F}}^2+\eta_2^2\|\mathring{\mathbf{C}}-\check{\mathbf{C}}\|_{\mathrm{F}}^2+\eta_3^2\textstyle\sum_{t\in\mathcal{T}'}\|\mathring{\mathbf{x}}[t]-\check{\mathbf{x}}[t]\|_2^2}\ <$$
$$\min\left\{\frac{\eta_1}{\sqrt{\tau}},\frac{\eta_2}{\sqrt{\tau}},\frac{\eta_3}{\sqrt{m+n}}\right\}\frac{s\big(\mathbf{A},\mathbf{C},\mathbf{x}[t_1],\dots,\mathbf{x}[t_{\tau'}]\big)}{2} \tag{17}$$

then $(\mathring{\mathbf{A}},\mathring{\mathbf{B}},\mathring{\mathbf{C}},\mathring{\mathbf{D}},\{\mathring{\mathbf{x}}[t]\}_{t\in\mathcal{T}'})$ *is the unique solution of the penalized convex relaxation (14a) – (14b).*

Corollary 1. *According to Theorem 1, it can be observed that if Algorithm 1 converges to a ground truth solution, then convergence occurs within a finite number of rounds.*

Proof. Define the matrix function $\mathbf{K}$ as shown in (18). Constraint (14b) is true if and only if

$$\mathbf{K}\big(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{x}[t_1],\dots,\mathbf{x}[t_{\tau'}],\bar{\mathbf{a}},\bar{\mathbf{c}},\bar{\mathbf{x}}\big)\in\mathbb{F}, \tag{19}$$

where

$$\mathbb{F}\triangleq\big\{\mathbf{F}\in\mathbb{S}_{n+m+\tau'}\mid \mathsf{F}_{ii}+\mathsf{F}_{jj}\geq 2\mathsf{F}_{ij}\quad \forall i,j\in\{1,\dots,n+m+\tau'\}\big\}. \tag{20}$$

In order to prove the optimality of $(\mathring{\mathbf{A}},\mathring{\mathbf{B}},\mathring{\mathbf{C}},\mathring{\mathbf{D}},\{\mathring{\mathbf{x}}[t]\}_{t\in\mathcal{T}'})$, it suffices to construct a dual certificate. To this end, define $\boldsymbol{\Lambda}\in\mathbb{R}^{n\times\tau'}$ and $\boldsymbol{\Pi}\in\mathbb{R}^{m\times\tau'}$ as the pair of matrices satisfying:

$$\begin{bmatrix}\mathrm{vec}\{\boldsymbol{\Lambda}\}\\ \mathrm{vec}\{\boldsymbol{\Pi}\}\end{bmatrix}=2\boldsymbol{J}^+\big(\mathring{\mathbf{A}},\mathring{\mathbf{C}},\mathring{\mathbf{x}}[t_1],\dots,\mathring{\mathbf{x}}[t_{\tau'}]\big)\begin{bmatrix}\eta_1\mathrm{vec}\{\mathring{\mathbf{A}}-\check{\mathbf{A}}\}\\ \eta_2\mathrm{vec}\{\mathring{\mathbf{C}}-\check{\mathbf{C}}\}\\ \eta_3(\mathring{\mathbf{x}}[t_1]-\check{\mathbf{x}}[t_1])\\ \vdots\\ \eta_3(\mathring{\mathbf{x}}[t_{\tau'}]-\check{\mathbf{x}}[t_{\tau'}])\end{bmatrix}. \tag{21}$$

We claim that the matrix

$$\boldsymbol{\Gamma}\triangleq\begin{bmatrix}\eta_1\boldsymbol{I}_{n\times n} & \mathbf{0}_{n\times m} & \boldsymbol{\Lambda}\\ * & \eta_2\boldsymbol{I}_{m\times m} & \boldsymbol{\Pi}\\ * & * & \eta_3\boldsymbol{I}_{\tau'\times\tau'}\end{bmatrix} \tag{22}$$

qualifies as a Lagrange multiplier associated with (19).

Stationarity with respect to primal variables is an immediate consequence of (21). Additionally, according to (21), we have

$$\sqrt{\|\boldsymbol{\Lambda}\|_{\mathrm{F}}^2+\|\boldsymbol{\Pi}\|_{\mathrm{F}}^2}\leq \frac{2\sqrt{\eta_1^2\|\mathring{\mathbf{A}}-\check{\mathbf{A}}\|_{\mathrm{F}}^2+\eta_2^2\|\mathring{\mathbf{C}}-\check{\mathbf{C}}\|_{\mathrm{F}}^2+\eta_3^2\sum_{t\in\mathcal{T}'}\|\mathring{\mathbf{x}}[t]-\check{\mathbf{x}}[t]\|_2^2}}{s\big(\mathbf{A},\mathbf{C},\mathbf{x}[t_1],\dots,\mathbf{x}[t_{\tau'}]\big)}, \tag{23}$$

which concludes that

$$\sqrt{\|\boldsymbol{\Lambda}\|_{\mathrm{F}}^2+\|\boldsymbol{\Pi}\|_{\mathrm{F}}^2}<\min\left\{\frac{\eta_1}{\sqrt{\tau}},\frac{\eta_2}{\sqrt{\tau}},\frac{\eta_3}{\sqrt{m+n}}\right\}, \tag{24}$$

meaning that $\boldsymbol{\Gamma}$ is strictly diagonally-dominant. Since, the set of strictly diagonally-dominant matrices is the dual cone of $\mathbb{F}$, $\boldsymbol{\Gamma}$ is dual feasible and qualifies as a dual certificate. □

V. Extension to Non-linear Systems

In this section, we extend the proposed methodology to non-linear systems. Consider the following dynamical equations:

$$\mathbf{x}[t+1]=\mathbf{A}\,\mathbf{x}[t]+\mathbf{B}\,\mathbf{u}[t]+\mathbf{E}\,\mathbf{s}[t]+\mathbf{F}\,\mathbf{r}[t]\qquad t\in\mathcal{T} \tag{25a}$$
$$\mathbf{y}[t]=\mathbf{C}\,\mathbf{x}[t]+\mathbf{D}\,\mathbf{u}[t]\qquad t\in\mathcal{T} \tag{25b}$$
$$\mathbf{r}[t]=(\mathbf{G}_1\,\mathbf{x}[t])\circ(\mathbf{G}_2\,\mathbf{x}[t])\qquad t\in\mathcal{T} \tag{25c}$$
$$\mathbf{s}[t]=\mathbf{x}[t]\circ\mathbf{x}[t]\qquad t\in\mathcal{T} \tag{25d}$$

where the matrices $(\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D},\mathbf{E},\mathbf{F})$ are unknown, and $(\mathbf{G}_1,\mathbf{G}_2)$ are known binary incidence matrices that are selected in such a way that $\mathbf{r}[t]$ contains all possible monomials of type $\mathsf{x}_i[t]\mathsf{x}_j[t]$ $(i\neq j)$ that may appear in the dynamics, i.e.,

- $\mathbf{s}[t]$ encapsulates all monomials of the type $\mathsf{x}_i[t]^2$,
- $\mathbf{r}[t]$ encapsulates a known subset of monomials of the type $\mathsf{x}_i[t]\mathsf{x}_j[t]$, where $i\neq j$.

Example 1. *This example presents the model of a Permanent Magnet Synchronous Machine (PMSM) in the above form. The machine dynamical equations are formulated in qd reference frame [20]. The internal parameters are* r_{s} *as the stator resistance,* l_{q} *as the q-axis inductance,* l_{d} *as the d-axis inductance,* λ_{m} *as the permanent magnet flux,* p *as pole pairs,* j_{m} *as the rotor inertia, and* d_{m} *as the viscous friction coefficient. The time-domain system states are* $i_{\mathrm{qs}}[t]$ *as the q-axis stator current,* $i_{\mathrm{ds}}[t]$ *as the d-axis stator current,* $w_{\mathrm{r}}[t]$ *as the rotor electrical speed, and* $\theta_{\mathrm{r}}[t]$ *as the rotor electrical position. The time-domain inputs are* $v_{\mathrm{qs}}[t]$ *as the q-axis stator voltage,* $v_{\mathrm{ds}}[t]$ *as the d-axis stator voltage, and* $t_{\mathrm{l}}[t]$ *as the load torque.*

The dynamics of the PMSM can be formulated as

$$\mathbf{x}[t+1]=\mathbf{A}\mathbf{x}[t]+\mathbf{B}\mathbf{u}[t]+\mathbf{F}\mathbf{r}[t] \tag{26a}$$
$$\mathbf{r}[t]=(\mathbf{G}_1\mathbf{x}[t])\circ(\mathbf{G}_2\mathbf{x}[t]) \tag{26b}$$

with respect to the time-domain state and input vectors

$$\mathbf{x}[t]=\big[i_{\mathrm{qs}}[t]\ \ i_{\mathrm{ds}}[t]\ \ w_{\mathrm{r}}[t]\ \ \theta_{\mathrm{r}}[t]\big]^\top \tag{27a}$$
$$\mathbf{u}[t]=\big[v_{\mathrm{qs}}[t]\ \ v_{\mathrm{ds}}[t]\ \ t_{\mathrm{l}}[t]\big]^\top \tag{27b}$$

where

$$\mathbf{A}=\begin{bmatrix}1-\frac{r_\mathrm{s}\delta}{l_\mathrm{q}} & 0 & -\frac{\lambda_\mathrm{m}\delta}{l_\mathrm{q}} & 0\\ 0 & 1-\frac{r_\mathrm{s}\delta}{l_\mathrm{d}} & 0 & 0\\ \frac{3\lambda_\mathrm{m}p^2\delta}{2j_\mathrm{m}} & 0 & 1-\frac{d_\mathrm{m}\delta}{j_\mathrm{m}} & 0\\ 0 & 0 & 1 & \delta\end{bmatrix}\ \mathbf{B}=\begin{bmatrix}\frac{\delta}{l_\mathrm{q}} & 0 & 0\\ 0 & \frac{\delta}{l_\mathrm{d}} & 0\\ 0 & 0 & -\frac{p\delta}{j_\mathrm{m}}\\ 0 & 0 & 0\end{bmatrix}$$

$$\mathbf{F}=\begin{bmatrix}-\frac{l_\mathrm{d}\delta}{l_\mathrm{q}} & 0 & 0\\ 0 & \frac{l_\mathrm{q}\delta}{l_\mathrm{d}} & 0\\ 0 & 0 & \frac{3(l_\mathrm{d}-l_\mathrm{q})p^2\delta}{2j_\mathrm{m}}\\ 0 & 0 & 0\end{bmatrix} \tag{28}$$

and δ is the time step, and

$$\mathbf{G}_1=\begin{bmatrix}0&0&1&0\\0&0&1&0\\1&0&0&0\end{bmatrix}\qquad \mathbf{G}_2=\begin{bmatrix}0&1&0&0\\1&0&0&0\\0&1&0&0\end{bmatrix}. \tag{29}$$

In Section VI, we demonstrate how the proposed method recovers the unknown time-domain parameters of this machine.

In order to convexify the dynamical equations (25), we first introduce the auxiliary vectors $\{\mathbf{x}^\mathbf{A}[\mathrm{t}]=\mathbf{A}\mathbf{x}[\mathrm{t}]\in\mathbb{R}^n\}_{\tau\in\mathcal{T}}$, $\{\mathbf{s}^\mathbf{E}[\mathrm{t}]=\mathbf{E}\,\mathbf{s}[\mathrm{t}]\in\mathbb{R}^n\}_{\tau\in\mathcal{T}}$, $\{\mathbf{r}^\mathbf{F}[\mathrm{t}]=\mathbf{F}\,\mathbf{r}[\mathrm{t}]\in\mathbb{R}^n\}_{\tau\in\mathcal{T}}$, and $\{\mathbf{x}^\mathbf{C}[\mathrm{t}]=\mathbf{C}\,\mathbf{x}[\mathrm{t}]\in\mathbb{R}^n\}_{\tau\in\mathcal{T}}$. Using these auxiliary variables, the dynamical equations (25) can be cast linearly in form of:

$$\mathbf{x}[\mathrm{t}+1]=\mathbf{x}^\mathbf{A}[\mathrm{t}]+\mathbf{B}\mathbf{u}[\mathrm{t}]+\mathbf{s}^\mathbf{E}[\mathrm{t}]+\mathbf{r}^\mathbf{F}[\mathrm{t}]\qquad \mathrm{t}\in\mathcal{T} \tag{30a}$$
$$\mathbf{y}[\mathrm{t}]=\mathbf{x}^\mathbf{C}[\mathrm{t}]+\mathbf{D}\mathbf{u}[\mathrm{t}]\qquad \mathrm{t}\in\mathcal{T} \tag{30b}$$

Additionally, let $\mathbf{e}_k$ and $\mathbf{f}_k$ be the k-th column of $\mathbf{E}^\top$ and $\mathbf{F}^\top$, respectively. We then impose the following constraints to relate $\mathbf{x}^\mathbf{A}[\mathrm{t}]$, $\mathbf{s}^\mathbf{E}[\mathrm{t}]$, $\mathbf{r}^\mathbf{F}[\mathrm{t}]$, and $\mathbf{x}^\mathbf{C}[\mathrm{t}]$ and their corresponding values:

$$\bar{\mathrm{a}}_k+\bar{\mathrm{x}}_\mathrm{t}+2\mathrm{x}^\mathbf{A}_k[\mathrm{t}]\geq\|\mathbf{a}_k+\mathbf{x}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{N} \tag{31a}$$
$$\bar{\mathrm{a}}_k+\bar{\mathrm{x}}_\mathrm{t}-2\mathrm{x}^\mathbf{A}_k[\mathrm{t}]\geq\|\mathbf{a}_k-\mathbf{x}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{N} \tag{31b}$$
$$\bar{\mathrm{e}}_k+\bar{\mathrm{s}}_\mathrm{t}+2\mathrm{s}^\mathbf{E}_k[\mathrm{t}]\geq\|\mathbf{e}_k+\mathbf{s}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{N} \tag{31c}$$
$$\bar{\mathrm{e}}_k+\bar{\mathrm{s}}_\mathrm{t}-2\mathrm{s}^\mathbf{E}_k[\mathrm{t}]\geq\|\mathbf{e}_k-\mathbf{s}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{N} \tag{31d}$$
$$\bar{\mathrm{f}}_k+\bar{\mathrm{r}}_\mathrm{t}+2\mathrm{r}^\mathbf{F}_k[\mathrm{t}]\geq\|\mathbf{f}_k+\mathbf{r}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{N} \tag{31e}$$
$$\bar{\mathrm{f}}_k+\bar{\mathrm{r}}_\mathrm{t}-2\mathrm{r}^\mathbf{F}_k[\mathrm{t}]\geq\|\mathbf{f}_k-\mathbf{r}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{N} \tag{31f}$$
$$\bar{\mathrm{c}}_k+\bar{\mathrm{x}}_\mathrm{t}+2\mathrm{x}^\mathbf{C}_k[\mathrm{t}]\geq\|\mathbf{a}_k+\mathbf{x}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{M} \tag{31g}$$
$$\bar{\mathrm{c}}_k+\bar{\mathrm{x}}_\mathrm{t}-2\mathrm{x}^\mathbf{C}_k[\mathrm{t}]\geq\|\mathbf{a}_k-\mathbf{x}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T},\quad k\in\mathcal{M} \tag{31h}$$

where $\bar{\mathrm{e}}_k=\|\mathbf{e}_k\|^2$, $\bar{\mathrm{f}}_k=\|\mathbf{f}_k\|^2$, $\bar{\mathrm{s}}_\mathrm{t}=\|\mathbf{s}[\mathrm{t}]\|^2$, and $\bar{\mathrm{r}}_\mathrm{t}=\|\mathbf{r}[\mathrm{t}]\|^2$. The following constraints impose the relationship between $\mathbf{r}[\mathrm{t}]$ and the corresponding monomials that it contains:

$$\mathbf{G}_1\,\mathbf{s}[\mathrm{t}]+\mathbf{G}_2\,\mathbf{s}[\mathrm{t}]+2\,\mathbf{r}[\mathrm{t}]\geq$$
$$(\mathbf{G}_1\,\mathbf{x}[\mathrm{t}]+\mathbf{G}_2\,\mathbf{x}[\mathrm{t}])\circ(\mathbf{G}_1\,\mathbf{x}[\mathrm{t}]+\mathbf{G}_2\,\mathbf{x}[\mathrm{t}])\qquad \mathrm{t}\in\mathcal{T} \tag{32a}$$
$$\mathbf{G}_1\,\mathbf{s}[\mathrm{t}]+\mathbf{G}_2\,\mathbf{s}[\mathrm{t}]-2\,\mathbf{r}[\mathrm{t}]\geq$$
$$(\mathbf{G}_1\,\mathbf{x}[\mathrm{t}]-\mathbf{G}_2\,\mathbf{x}[\mathrm{t}])\circ(\mathbf{G}_1\,\mathbf{x}[\mathrm{t}]-\mathbf{G}_2\,\mathbf{x}[\mathrm{t}])\qquad \mathrm{t}\in\mathcal{T} \tag{32b}$$

Finally, the following constraints complete the proposed parabolic relaxation for nonlinear systems:

$$\bar{\mathrm{x}}_t=\mathbf{1}^\top\mathbf{s}[\mathrm{t}],\qquad \bar{\mathrm{r}}_t\geq\|\mathbf{r}[\mathrm{t}]\|_2^2\qquad \mathrm{t}\in\mathcal{T} \tag{33a}$$
$$\bar{\mathrm{s}}_t\geq\|\mathbf{s}[\mathrm{t}]\|_2^2,\qquad \mathbf{s}[\mathrm{t}]\geq\mathbf{x}[\mathrm{t}]\circ\mathbf{x}[\mathrm{t}]\qquad \mathrm{t}\in\mathcal{T} \tag{33b}$$
$$\bar{\mathrm{a}}_k\geq\|\mathbf{a}_k\|_2^2,\quad \bar{\mathrm{e}}_k\geq\|\mathbf{e}_k\|_2^2,\quad \bar{\mathrm{f}}_k\geq\|\mathbf{f}_k\|_2^2\qquad k\in\mathcal{N} \tag{33c}$$
$$\bar{\mathrm{c}}_k\geq\|\mathbf{c}_k\|_2^2\qquad k\in\mathcal{M}. \tag{33d}$$

If equality holds for all of the constraints in (33), then the relaxation is exact. The penalized parabolic relaxation for identification of nonlinear systems can cast as follows:

$$\begin{aligned}\text{minimize}\quad &\eta_1\big(\mathbf{1}_n^\top\,\bar{\mathbf{a}}-2\,\mathrm{tr}\big\{[\mathbf{a}_1\ldots\mathbf{a}_n]^\top\check{\mathbf{A}}\big\}+\|\check{\mathbf{A}}\|_\mathrm{F}^2\big)+\\ &\eta_2\big(\mathbf{1}_n^\top\,\bar{\mathbf{c}}-2\,\mathrm{tr}\big\{[\mathbf{c}_1\ldots\mathbf{c}_n]^\top\check{\mathbf{C}}\big\}+\|\check{\mathbf{C}}\|_\mathrm{F}^2\big)+\\ &\eta_3\big(\mathbf{1}_n^\top\,\bar{\mathbf{e}}-2\,\mathrm{tr}\big\{[\mathbf{e}_1\ldots\mathbf{e}_n]^\top\check{\mathbf{E}}\big\}+\|\check{\mathbf{E}}\|_\mathrm{F}^2\big)+\\ &\eta_4\big(\mathbf{1}_n^\top\,\bar{\mathbf{f}}-2\,\mathrm{tr}\big\{[\mathbf{f}_1\ldots\mathbf{f}_n]^\top\check{\mathbf{F}}\big\}+\|\check{\mathbf{F}}\|_\mathrm{F}^2\big)+\\ &\eta_5\sum\nolimits_{\mathrm{t}\in\mathcal{T}}\big(\bar{\mathrm{r}}_\mathrm{t}-2\,\check{\mathbf{r}}^\top[\mathrm{t}]\,\mathbf{r}[\mathrm{t}]+\|\check{\mathbf{r}}[\mathrm{t}]\|_2^2\big)+\\ &\eta_6\sum\nolimits_{\mathrm{t}\in\mathcal{T}}\big(\bar{\mathrm{s}}_\mathrm{t}-2\,\check{\mathbf{s}}^\top[\mathrm{t}]\,\mathbf{s}[\mathrm{t}]+\|\check{\mathbf{s}}[\mathrm{t}]\|_2^2\big)+\\ &\eta_7\sum\nolimits_{\mathrm{t}\in\mathcal{T}}\big(\bar{\mathrm{x}}_\mathrm{t}-2\,\check{\mathbf{x}}^\top[\mathrm{t}]\,\mathbf{x}[\mathrm{t}]+\|\check{\mathbf{x}}[\mathrm{t}]\|_2^2\big)\end{aligned} \tag{34a}$$

$$\text{subject to}\quad (30),(31),(32),(33). \tag{34b}$$

This problem can then be solved sequentially to obtain a feasible solution for the set of dynamical equations (25).

VI. Numerical Results

A. Linear System Identification

We consider system identification problems with $n=16$, $m=12$, $o=10$, $\tau=250$, and $\hat{\mathcal{T}}=\mathcal{T}\setminus\{1,6,11,\ldots,246\}$. The resulting problem is 3928-dimensional accounting for the elements of $\mathbf{A}\in\mathbb{R}^{16\times16}$, $\mathbf{B}\in\mathbb{R}^{16\times10}$, $\mathbf{C}\in\mathbb{R}^{12\times16}$, $\mathbf{D}\in\mathbb{R}^{12\times10}$, and $\{\mathbf{x}[\mathrm{t}]\in\mathbb{R}^{16}\}_{\mathrm{t}\in\mathcal{T}'}$. The ground truth matrices are

- Every element of $\mathbf{A}-\boldsymbol{I}_{n\times n}$ is uniformly chosen from the interval $[-0.25,+0.25]$.
- The elements of $\mathbf{B}\in\mathbb{R}^{16\times10}$, $\mathbf{C}\in\mathbb{R}^{12\times16}$, and $\mathbf{D}\in\mathbb{R}^{12\times10}$ have zero-mean standard normal distribution.
- The elements of $\mathbf{x}[1]$ are uniformly chosen from the interval $[0.5;1.5]$.
- For every $\mathrm{t}\in\mathcal{T}$, we have $\mathbf{u}[\mathrm{t}]=\mathbf{F}\mathbf{x}[\mathrm{t}]+\mathbf{w}[\mathrm{t}]$, where the elements of $\mathbf{w}[\mathrm{t}]$ have independent Gaussian distribution with zero mean and standard deviation 0.1. Additionally, $\mathbf{F}=-(\boldsymbol{I}_{m\times m}+\mathbf{B}^\top\mathbf{P})^{-1}\mathbf{B}^\top\mathbf{P}\mathbf{A}$ is an optimal LQR controller with $\mathbf{P}$ representing the unique positive-definite solution to the Riccati equation:

$$\mathbf{A}^\top\mathbf{P}\mathbf{A}+\boldsymbol{I}_{n\times n}-\mathbf{P}=\mathbf{A}^\top\mathbf{P}\mathbf{B}(\boldsymbol{I}_{n\times n}+\mathbf{B}^\top\mathbf{P}\mathbf{B})^{-1}\mathbf{B}^\top\mathbf{P}\mathbf{A} \tag{35}$$

We generate 15 random systems and solve the resulting problem using the sequential Algorithm 1 with the initial point $\check{\mathbf{A}}=\boldsymbol{I}_{n\times n}$ and $\check{\mathbf{x}}[\mathrm{t}]=\mathbf{0}_n$ for every $t\in\hat{\mathcal{T}}$. Figure 2 illustrates the convergence of the error function

$$\left(\sum_{\mathrm{t}=1}^{\tau}\|\mathbf{x}[\mathrm{t}]-\mathbf{x}^{\mathrm{true}}[\mathrm{t}]\|_2^2\right)^{1/2}, \tag{36}$$

for all 15 experiments. Each round of the penalized parabolic is solved in less than 15 seconds.

For comparison, we have performed system identification on the same randomly generated systems using the methods in [21]. We found that subspace system identification works well when all the instances of the state matrix $\mathbf{x}[\mathrm{t}]$ are well known in advance. However, when only a portion of the input states $\mathbf{x}[\mathrm{t}]$ were known, that method failed. By contrast, our proposed method converges quite quickly with high accuracy.

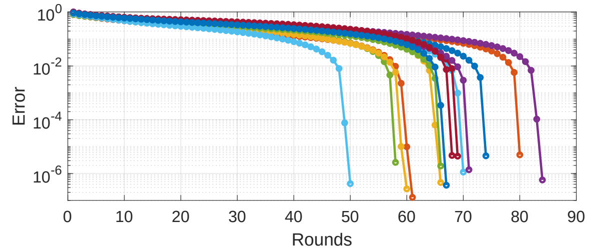

Figure 2. The performance of the sequential Algorithm 1 for linear system identification.

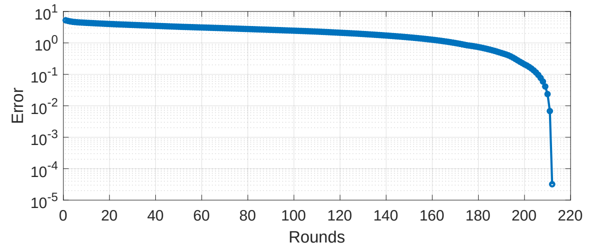

Figure 3. The performance of the proposed sequential approach on the nonlinear model of the PMSM example.

B. Nonlinear System Identification

We consider a PMSM machine with the parameters $r_{\rm s} = 0.1465$, $l_{\rm d} = l_{\rm q} = 0.0001395$, $d_{\rm m} = 0.009$, $\lambda_{\rm m} = 0.005$, $p = 7$, and $j_{\rm m} = 0.0000835$ [22]. We have generated time-domain signals with time steps of size $\delta = 5 \times 10^{-6}$ seconds and attempted to recover the unknown parameters using the approach proposed in Section V. We have considered a time horizon of length 20, with the following input signals:

- Each $v_{\rm qs}[\mathrm{t}]$ random Gaussian distribution with mean 12 and standard deviation 1.
- Each $v_{\rm ds}[\mathrm{t}]$ random Gaussian distribution with mean 0.25 and standard deviation 1.
- $t_{\rm l}[\mathrm{t}] = \begin{cases} 0 & \mathrm{t} < 10 \\ 0.25 & \mathrm{t} \geq 10 \end{cases}$

It should be noted that above inputs are not physical input, rather excitements considered for the sole purpose of system identification. The following parameters are assumed known:

- The matrices $\mathbf{G}_1$ and $\mathbf{G}_2$.
- The last row of matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{F}$, due to the trivial relation between the rotor position and rotor speed.
- The state vector $\mathbf{x}[\mathrm{t}]$ at times $\mathrm{t} \in \{1, 5, 15, 20\}$.

The convergence of the error function (36) is shown in Fig. 3.

VII. Conclusion

System identification is the process of finding optimized parameters of an unknown system using the input/output data. This paper presents a novel parabolic relaxation to identify linear and nonlinear dynamical systems. The proposed convex relaxation relies on a significantly smaller number of auxiliary variables leading to a higher efficiency compared to the state-of-the-art methods. We have provided theoretical guarantees on the method's feasibility, and used numerical simulation to validate this method for both linear and nonlinear systems.

References

[1] L. Fu and P. Li, "The research survey of system identification method," *Proceedings - 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2013*, vol. 2, pp. 397 – 401, Aug. 2013.

[2] L. Ljung, T. Chen, and B. Mu, "A shift in paradigm for system identification," *International Journal of Control*, vol. 93, no. 2, pp. 173–180, Feb. 2020.

[3] B. Mavkov, M. Forgione, and D. Piga, "Integrated neural networks for nonlinear continuous-time system identification," *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 851 – 856, Oct. 2020.

[4] K. Asan, G. Saravanakumar, K. Valarmathi, D. Devaraj, and T. Radhakrishnan, "Real-coded genetic algorithm for system identification and tuning of a modified model reference adaptive controller for a hybrid tank system," *Applied Mathematical Modelling*, vol. 37, no. 6, pp. 3829 – 3847, March 2013.

[5] Z. Wang, L. Zhao, and X. Luo, "Levy-particle swarm optimization intelligent search-based iterative identification for nonparametric models of bilinear systems with gaussian mixture noises," *Transactions of the Institute of Measurement and Control*, vol. 41, no. 14, pp. 3970 – 3978, Oct. 2019.

[6] L. Vandenberghe, "Convex optimization techniques in system identification," *IFAC Proceedings Volumes*, vol. 16, no. PART 1, pp. 71 – 76, July 2012.

[7] M. M. Tobenkin, I. R. Manchester, J. Wang, A. Megretski, and R. Tedrake, "Convex optimization in identification of stable non-linear state space models," *Proceedings of the IEEE Conference on Decision and Control*, pp. 7232 – 7237, Dec. 2010.

[8] S. Burer, "Semidefinite programming in the space of partial positive semidefinite matrices," *SIAM Journal on Optimization*, vol. 14, no. 1, pp. 139 – 172, 2003.

[9] F. Alizadeh and D. Goldfarb, "Second-order cone programming," *Mathematical Programming, Series B*, vol. 95, no. 1, pp. 3 – 51, Jan. 2003.

[10] M. Goemans and L. Tuncel, "When does the positive semidefiniteness constraint help in lifting procedures?" *Mathematics of Operations Research*, vol. 26, no. 4, pp. 796 – 815, Nov. 2001.

[11] A. A. Ahmadi and A. Majumdar, "DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization," *2014 48th Annual Conference on Information Sciences and Systems, CISS, Princeton, NJ, USA*, pp. 1–5, March 2014.

[12] H. Feng and J. Lavaei, "On the exponential number of connected components for the feasible set of optimal decentralized control problems," *2019 American Control Conference (ACC)*, pp. 1430 – 1437, July 2019.

[13] A. Barzegar, D. K. Molzahn, and R. Su, "A method for quickly bounding the optimal objective value of an OPF problem using a semidefinite relaxation and a local solution," *Electric Power Systems Research*, vol. 177, pp. 41 – 49, Dec. 2019.

[14] Z. Tian and W. Wu, "Recover feasible solutions for SOCP relaxation of optimal power flow problems in mesh networks," *IET Generation, Transmission and Distribution*, vol. 13, no. 7, pp. 1078 – 1087, April 2019.

[15] D. Molzahn and I. Hiskens, "Mixed SDP/SOCP moment relaxations of the optimal power flow problem," *2015 IEEE Eindhoven PowerTech*, pp. 1–6, June 2015.

[16] R. Madani, S. Sojoudi, G. Fazelnia, and J. Lavaei, "Finding low-rank solutions of sparse linear matrix inequalities using convex optimization," *SIAM Journal on Optimization*, vol. 27, no. 2, pp. 725 – 758, 2017.

[17] M. Fukuda, M. Kojima, K. Murota, and K. Nakata, "Exploiting sparsity in semidefinite programming via matrix completion I: General framework," *SIAM Journal on Optimization*, vol. 11, no. 3, pp. 647–674, 2000.

[18] K. Nakata, K. Fujisawa, M. Fukuda, M. Kojima, and K. Murota, "Exploiting sparsity in semidefinite programming via matrix completion II: implementation and numerical results," *Mathematical Programming*, vol. 95, no. 2, pp. 303–327, Feb. 2003.

[19] M. Kheirandishfard, F. Zohrizadch, M. Adil, and R. Madani, "Convex relaxation of bilinear matrix inequalities part II: Applications to optimal control synthesis," *Proceedings of the IEEE Conference on Decision and Control (CDC)*, pp. 75 – 82, Dec. 2018.

[20] S. Carpiuc and C. Villegas, "Fpga-based rapid control prototyping of permanent magnet synchronous motor servo drives," *PCIM Europe 2019; International Exhibition and Conference for Power Electronics, Intelligent Motion, Renewable Energy and Energy Management*, pp. 1581 – 1586, May 2019.

[21] W. Favoreel, B. De Moor, and P. Van Overschee, "Subspace state space system identification for industrial processes," *Journal of Process Control*, vol. 10, no. 2, pp. 149 – 155, April 2000.

[22] Maxon Group, "Ec 60 flat ϕ 60 mm, brushless, 100 watt," last visited Feb. 2021. [Online]. Available: https://www.maxongroup.com/medias/sys_master/root/8841185656862/EN-294.pdf