Statistical Science
2021, Vol. 36, No. 2, 256–263
https://doi.org/10.1214/20-STS780

Noncommutative Probability and Multiplicative Cascades

Ian W. McKeague

Abstract. Various aspects of standard model particle physics might be explained by a suitably rich algebra acting on itself, as suggested by Furey (2015). The present paper develops the asymptotics of large causal tree diagrams that combine freely independent elements in such an algebra. The Marčenko–Pastur law and Wigner's semicircle law are shown to emerge as limits of normalized sum-over-paths of nonnegative elements assigned to the edges of causal trees. These results are established in the setting of noncommutative probability. Trees with classically independent positive edge weights (random multiplicative cascades) were originally proposed by Mandelbrot as a model displaying the fractal features of turbulence. The novelty of the present work is the use of noncommutative (free) probability to allow the edge weights to take values in an algebra. An application to theoretical neuroscience is also discussed.

Key words and phrases: Mandelbrot cascade, Marčenko–Pastur law, martingale convergence, random matrices, Wigner's semicircle law.

1. INTRODUCTION

Initiated by Dan Voiculescu in around 1986, noncommutative probability has become a flourishing area of mathematics with close ties to random matrix theory (Anderson, Guionnet and Zeitouni, 2009) and quantum mechanics (Collins, Hayden and Nechita, 2017). Random multiplicative cascades were introduced by Mandelbrot in around 1974 as a model displaying the fractal features of turbulence. The present paper develops and discusses a noncommutative probability version of random multiplicative cascades. It is shown that some of the familiar limiting distributions in random matrix theory emerge in this setting, such as Wigner's semicircle law and the Marčenko–Pastur law (noncommutative probability analogues of the Gaussian and Poisson distributions, resp.). The discussion is essentially self-contained, and no previous knowledge of these areas is assumed.

The paper was initially motivated by the problem of understanding the behavior of large particle systems of the type recently proposed by Furey (2015) in order to explain aspects of the standard model of particle physics using "little more than an algebra." Physical concepts such as particles, causality and irreversible time are hypothesized by Furey to emerge from a sufficiently rich algebra $\mathcal{A}$ acting on itself. For instance, each of the standard model's Lorentz representations, scalars, spinors, four-vectors and the field strength tensor are shown to arise as certain invariant subspaces of the complex quaternions $\mathbb{C} \otimes \mathbb{H}$.

Furey (2015) further proposed that causality might be explained through the evaluation of algebraic expressions given by a sum-over-paths in rooted tree diagrams having edge weights in $\mathcal{A}$. In such diagrams, the sum-over-paths is defined as the sum over all self-avoiding paths between a leaf and the root of the tree, where the product of the edge weights along the path gives its contribution to the sum; see the example in Figure 1. In the sequel, we refer to this model as a *causal tree*. Various other concepts of causal trees and sets have been proposed in connection with models of quantum gravity (Bombelli et al., 1987, Markopoulou, 2000, Dowker, 2005). However, it is only in Furey's approach, which is also discussed in Cortês and Smolin (2014a, 2014b), that particles emerge at the fundamental level, rather than as artifacts of other phenomena.

Causal trees in which the number of children of each inner vertex is $N \geq 2$ and the edge weights are (classically) independent copies of a nondegenerate random variable $W \geq 0$ with $EW = 1$, are known as *multiplicative cascades*. These were introduced in the seminal work of Rosenblatt and Van Atta (1972, 1974a, 1974b) as a tractable way of explaining the fractal nature of turbulence. In the statistical physics literature, they were used by Derrida and Spohn (1988) to study Gibbs measures of directed polymers on disordered trees. For references

Ian W. McKeague is Professor of Biostatistics, Department of Biostatistics, Columbia University, New York, New York 10032, USA (e-mail: ian.mckeague@columbia.edu).

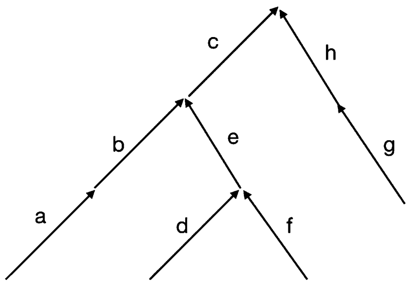

FIG. 1. *Example of a causal tree with the sum-over-paths given by the algebraic expression* $c \cdot (b \cdot a + e \cdot (d + f)) + h \cdot g$.

to the extensive literature on multiplicative cascades, including their connection to models of quantum gravity, see Barral et al. (2014), Barral (2014), Huang (2016).

Mandelbrot showed that for a multiplicative cascade on the N-ary tree of height n, the sum-over-paths (formally defined in Section 3) when normalized by N^n has an a.s. limit (X_N) as $n \to \infty$, where $EX_N \le 1$. This follows by application of the martingale convergence theorem. The distribution of X_N is characterized by a solution to the distributional fixed-point equation

$$X_N =_d \frac{1}{N} \sum_{i=1}^{N} W_i X_{i,N}, \tag{1.1}$$

where the $X_{i,N}$ are independent copies of X_N, and the W_i are independent copies of W that are independent of the $X_{i,N}$. The derivation is illustrated in Figure 2 below. From this equation, Kahane and Peyrière (1976) showed that $EX_N = 1$ if $EW \log W < \log N$, and otherwise $X_N = 0$ a.s.; when the condition is satisfied, $EX_N^2 < \infty$ if and only if $EW^2 < \infty$.

Assuming only that $EW^\gamma < \infty$ for some $\gamma > 0$, Durrett and Liggett (1983) gave a necessary and sufficient condition on W for (1.1) to have a nontrivial solution, and showed that the normalized sum-over-paths converges in distribution to X_N when the Laplace transforms of W and X_N have matching behavior at 0. It is difficult to find (nontrivial) explicit solutions to (1.1) except in some special cases; for instance, when W/N has a beta distribution with certain parameter values then X_N has a gamma

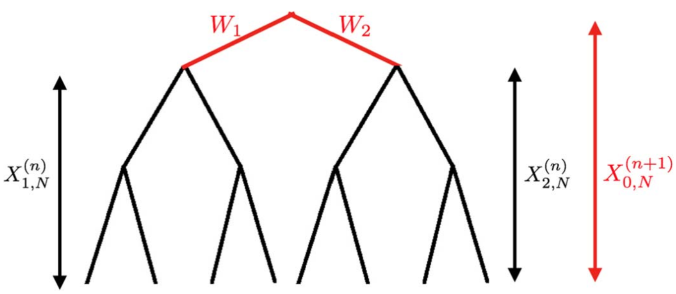

FIG. 2. *The recursion relation* (1.1) *is derived from the* N*-ary tree structure by noting that* $X_N^{(n+1)}$ *is a weighted sum of* N *i.i.d. copies of* $X_N^{(n)}$ *(where* $N = 2$ *in this case).*

distribution. This solution and a few other tractable solutions are discussed by Ossiander and Waymire (2002), who also study the problem of estimating the distribution of W given data on X_N. The asymptotic behavior of X_N as $N \to \infty$ has been studied by Liu and Rouault (2000), who showed using the Lindeberg–Feller theorem that if the distribution of W is fixed (i.e., not depending on N) and $EW^2 < \infty$, then $\sqrt{N}(X_N - 1)$ converges in distribution to normal with mean zero and variance $EW^2 - 1$.

The natural framework for studying the behavior of normalized sum-over-paths in causal trees with algebra-valued edge weights is *noncommutative* (or free) probability. In this paper, we introduce free versions of the multiplicative cascades described above. We show that the fixed-point equation (1.1) in which the $X_{i,N}$ are freely independent copies of X_N, and the normalized weights $\bar{W}_i \equiv W_i/N$ are freely independent Bernoulli($1/N$) random variables, has a nontrivial solution, namely the Marčenko–Pastur distribution with parameter $\lambda = 1$, which implies the normalized sum-over-paths has a *nontrivial* weak limit. This is in marked contrast to the multiplicative cascade with the $\bar{W}_i$ as classically independent Bernoulli($1/N$), for which the weak limit of the normalized sum-over-paths is trivial (Dirac measure δ_0); this follows from the result of Kahane and Peyrière (1976) mentioned above (since $EW \log W = \log N$ in this case).

We also study the asymptotic distribution of general solutions X_N to the free version of the fixed-point equation (1.1) as $N \to \infty$, when the distribution of W is fixed with $EW^2 < \infty$. Specifically, we show that $\sqrt{N}(X_N - 1)$ converges in distribution to Wigner's semicircle law of radius $R = \sqrt{EW^2 - 1}$, amounting to a free analogue of Liu and Rouault's central limit theorem mentioned above.

The paper is organized as follows. Section 2 describes background on Voiculescu's theory of free probability that we need. Readers unfamiliar with this topic should refer for more detailed information to the many excellent surveys of the subject, including Nica and Speicher (2006), Anderson, Guionnet and Zeitouni (2009), Tao (2010), Kargin (2013) and Mingo and Speicher (2017). Section 3 introduces the notion of a free multiplicative cascade, which we call a free causal tree. Three results about the limiting distributional behavior of the sum-over-paths of large free causal trees are developed in Section 4, where we also discuss an open problem related to our main result and how it applies to free versions of Galton–Watson trees and critical percolation on Bethe lattices. Section 5 contains a discussion of how the idea of using suitably rich algebras acting on themselves to explain the behavior of complex causal tree structures may also be relevant in theoretical neuroscience, namely in connection with the quantum cognition hypothesis.

2. BACKGROUND ON FREE PROBABILITY

A noncommutative probability space is a pair $(\mathcal{A}, \mathbb{E})$, where $\mathcal{A}$ is a $*$-algebra over $\mathbb{C}$ and $\mathbb{E}\colon \mathcal{A} \to \mathbb{C}$ is a linear functional such that $\mathbb{E}(1_{\mathcal{A}}) = 1$, where $1_{\mathcal{A}}$ is the multiplicative identity. The elements $X \in \mathcal{A}$ are called (noncommutative) random variables, or "observables" in quantum probability language. $\mathbb{E}$ is a *state* and assumed to satisfy $\mathbb{E}(X^*X) \geq 0$, with equality only when $X = 0$. Random variables of the form X^*X will be said to be nonnegative (although "positive" is the standard terminology), and X is centered if $\mathbb{E}(X) = 0$. The operation $X \mapsto X - \mathbb{E}(X)1_{\mathcal{A}}$ is called centering. Classical probability theory is based on the commutative algebra of bounded measurable real-valued functions on some (standard) probability space, with $\mathbb{E}(X) = EX = \int X\, d\mathrm{P}$.

We further assume (essentially without loss of generality) that $\mathcal{A}$ is a von Neumann algebra: $\mathcal{A} \subset B(H)$, the space of bounded linear operators on a separable Hilbert space H, and is closed in the weak operator topology, with the adjoint being the involution. The state is assumed to be tracial, meaning $\mathbb{E}(XY) = \mathbb{E}(YX)$. The spectral distribution (or law) of X is the map $\mathcal{L}\colon \mathbb{C}[X] \to \mathbb{C}$ such that $\mathcal{L}(p) = \mathbb{E}(p(X))$, where $\mathbb{C}[X]$ is the set of polynomials in X having coefficients in $\mathbb{C}$.

When X is self-adjoint (i.e., $X^* = X$), which we assume for all random variables from now on, it can be shown (see Nica and Speicher, 2006, Proposition 3.13) that there is a unique probability measure μ on $\mathbb{R}$ such that $\mathcal{L}(p) = \int p(x)\, d\mu(x)$, denoted $X \sim \mu$. We say X is bounded if its spectral radius $\rho(X) = \lim_{k\to\infty} |\mathbb{E}X^{2k}|^{1/(2k)}$ is finite, and in that case μ is supported by $[-\rho(X), \rho(X)]$; see Tao (2010). A sequence of random variables $X_n \sim \mu_n$ is said to converge in distribution to X (denoted $X_n \to_d X$) if their laws converge; when the law of $X \sim \mu$ is determined by its moments, this is equivalent to weak convergence $\mu_n \to_d \mu$ in the usual sense.

A family $\{\mathcal{A}_k\}$ of subalgebras of $\mathcal{A}$, each containing $1_{\mathcal{A}}$, is called *freely independent* (or just free) if every finite product $X_1 \cdots X_n$ of centered $X_j \in \mathcal{A}_{k(j)}$ with alternating indices $k(j) \neq k(j+1)$ is also centered. A family of random variables $\{X_k\}$ is free if the algebras they generate are free. Although analogous, free independence is not a generalization of classical independence as it is based on a different factorization rule for calculating mixed moments. Two classically independent random variables X and Y are free only if one of them is a scalar, since classical independence implies $\mathbb{E}(X^2)\mathbb{E}(Y^2) = \mathbb{E}(XYXY)$, but the latter expression vanishes when X and Y are free and centered.

The distributions of $X + Y$ and XY for free random variables $X \sim \mu$ and $Y \sim \nu$ are determined by μ and ν, and denoted $\mu \boxplus \nu$ and $\mu \boxtimes \nu$, respectively. The most tractable way of analyzing these free convolutions is through transforms, rather than directly with their moments. The Cauchy transform of a random variable $X \sim \mu$ is defined by analytically extending

$$G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z - x}\, d\mu(x) = \mathbb{E}(z - X)^{-1}$$

to the upper half of the complex plane $\{z\colon\ \mathrm{Im}(z) > 0\}$. G_μ uniquely determines μ by the Stieltjes inversion formula

$$\mu((a,b]) = -\lim_{\epsilon\to 0} \frac{1}{\pi} \int_a^b \mathrm{Im}\, G_\mu(y + i\epsilon)\, dy$$

for any $a < b$ such that $\mu\{a,b\} = 0$. For bounded X, the moment $\mathbb{E}X^k$ is the coefficient of $z^{-(k+1)}$ in the expansion of $G_\mu(z)$ obtained by applying $\mathbb{E}$ term-by-term to the formal Neumann series $(z - X)^{-1} = 1/z + X/z^2 + X^2/z^3 + \cdots$.

Two other transforms derived from the Cauchy transform play an important role. For compactly supported μ, the R-transform is $R_\mu(z) = zG_\mu^{-1}(z) - 1$, a variation on the original definition $G_\mu^{-1}(z) - 1/z$ introduced by Voiculescu. For compactly supported μ and ν, the addition formula $R_{\mu\boxplus\nu}(z) = R_\mu(z) + R_\nu(z)$ is analogous to the additivity of the log-characteristic function for classically independent random variables. An elegant heuristic proof can be found in Tao (2010). Further, the S-transform $S_\mu(z) = R_\mu^{-1}(z)/z$ uniquely determines μ, and for $\nu \neq \delta_0$ on $\mathbb{R}^+$ it can be shown that $\mu \boxtimes \nu \neq \delta_0$ and $S_{\mu\boxtimes\nu}(z) = S_\mu(z)S_\nu(z)$. See Chapter 5 of Anderson, Guionnet and Zeitouni (2009) for more details.

EXAMPLE 2.1 (Free central limit theorem). Let $Z_n = (X_1 + \cdots + X_n)/\sqrt{n}$, where $X_1, X_2, \ldots$ are freely independent copies of a centered, bounded random variable X with $\mathbb{E}X^2 = 1$. Then Z_n converges in distribution to the standard Wigner semicircle law μ with density $\sqrt{(4 - x^2)_+}/(2\pi)$. Voiculescu obtained this result early in the development of free probability. We sketch a proof based on finding the limit of the R-transform of Z_n. By direct integration, the Cauchy transform of μ is given by $G_\mu(z) = (z - \sqrt{z^2 - 4})/2$, so its R-transform is $R_\mu(z) = z$. Using the addition formula above, the R-transform of Z_n is seen to have the formal series expansion

$$R_{Z_n}(z) = \sqrt{n} R_X(z/\sqrt{n}) = z + z^2/\sqrt{n} + \cdots$$

in which only the leading term (namely $R_\mu(z)$) remains after $n \to \infty$. The moments of Z_n can be expressed as polynomials of the coefficients (the so-called free-cumulants of X, see Anderson, Guionnet and Zeitouni, 2009) in the above power series, so each moment of Z_n converges to the corresponding moment of μ, and we conclude that $Z_n \to_d \mu$.

3. FREE CAUSAL TREES

First, we recall the multiplicative cascade on the (full) N-ary tree, where $N \geq 2$ is fixed for now. Let $\Sigma = \bigcup_{n=1}^{\infty} \Sigma_n$, where each word $\sigma = \sigma_1 \cdots \sigma_n$ in $\Sigma_n = \{1, \ldots, N\}^n$ represents a self-avoiding path from the root to a vertex on level n. Let $\{W_\sigma, \sigma \in \Sigma\}$ be (classically) independent copies of a nondegenerate random variable $W \geq 0$ with $EW = 1$. As discussed in the Introduction, by the martingale convergence theorem the normalized sum-over-paths

$$X_N^{(n)} = \sum_{\sigma \in \Sigma_n} \prod_{i=1}^{n} W_{\sigma_1 \cdots \sigma_i} \Big/ N^n \to X_N \quad \text{a.s.}$$

as $n \to \infty$. Clearly, $X_N^{(n)}$ can be expressed as

(3.1)
$$X_N^{(n)} = \sum_{\sigma_1 \in \Sigma_1} W_{\sigma_1} \Bigg[\sum_{\sigma_2 \in \Sigma_1} W_{\sigma_1 \sigma_2} \Bigg[\cdots \Bigg[\sum_{\sigma_n \in \Sigma_1} W_{\sigma_1 \cdots \sigma_n} \Bigg] \cdots \Bigg] \Bigg] \Big/ N^n.$$

We introduce the free multiplicative cascade (that we call a *free causal tree*) by taking the edge-weights W_σ to be freely independent, identically distributed, nonnegative elements of $\mathcal{A}$ with $\mathbb{E}(W_\sigma) = 1$. In what follows, it is notationally convenient to work in terms of the law $\nu = \nu_N$ of the *normalized* edge weights $\bar{W}_\sigma \equiv W_\sigma / N$, which is a unique probability measure on $\mathbb{R}^+$ since nonnegative elements of $\mathcal{A}$ are normal. It is then easily seen from the nested form of (3.1) and an inductive argument that the distribution of $X_N^{(n)}$ can be expressed as $\mu^{(n)} = T_\nu^n(\delta_1)$, where T_ν^n is the nth iterate of the (free) Mandelbrot map $T_\nu : \mathcal{M}^+ \to \mathcal{M}^+$ defined by

$$T_\nu(\mu) = [\nu \boxtimes \mu]^{\boxplus N},$$

where $\mathcal{M}^+$ is the family of probability measures on $\mathbb{R}^+$. We say that $\mu \in \mathcal{M}^+$ is a *fixed point* of T_ν if $T_\nu(\mu) = \mu$.

EXAMPLE 3.1. The smallest concrete example of a noncommutative probability space $(\mathcal{A}, \mathbb{E})$ supporting a (nontrivial) countable family of edge-weights $\{W_\sigma\}$ for a free causal tree is the von Neumann algebra $B(H)$ on the free "baby" Fock space introduced by Attal and Nechita (2011):

$$H = \mathbb{C}\Omega \oplus \bigoplus_{n \geq 1} \bigoplus_{i_1 \neq \cdots \neq i_n} \mathbb{C}U_{i_1} \otimes \cdots \otimes \mathbb{C}U_{i_n}.$$

Here, $\mathbb{C}\Omega$ is the one-dimensional subspace spanned by the vacuum vector $\Omega = (1, 0)^\mathsf{T}$ and $U_i = (0, 1)^\mathsf{T} \in \mathbb{C}^2_{(i)}$, where $\mathbb{C}^2_{(i)}$ is an ith copy of $\mathbb{C}^2$. The ground state is given by $\mathbb{E}(X) = \langle X\Omega, \Omega \rangle$ for $X \in B(H)$. For $0 < p < 1$, the matrix

$$X_i = \begin{bmatrix} p & \sqrt{p(1-p)} \\ \sqrt{p(1-p)} & p \end{bmatrix},$$

viewed as a linear operator on $\mathbb{C}^2_{(i)}$, can be embedded as a linear operator on H, and as such furnishes a free sequence of (nonnegative) elements of $B(H)$ (Attal and Nechita, 2011). The kth moment of X_i for $k \geq 1$ is easily seen to be given by $\mathbb{E}(X_i^k) = \langle X_i^k \Omega, \Omega \rangle = p$, which agrees with the kth moment of $\mathrm{Ber}(p)$. This can be used to construct an explicit free causal tree with Bernoulli distributed edge-weights.

4. LIMITS OF FREE CAUSAL TREES

In the classical case, $X_N^{(n)}$ converges almost surely as $n \to \infty$ to the nontrivial limit $X_N \sim \mu$ provided μ is a unit-mean solution to the distributional fixed-point equation (1.1). Analogously, in our free probability setting, we can show that $X_N^{(n)} \sim \mu^{(n)}$ converges in distribution to a fixed point of T_ν. This is done using a noncommutative version of the martingale convergence theorem.

THEOREM 4.1. *If the Mandelbrot map T_ν has a unit-mean fixed point, $\mu \in \mathcal{M}^+$ that is determined by its moments, then μ is the unique fixed point of the Mandelbrot map and the normalized sum-over-paths $X_N^{(n)}$ converges in distribution to $X_N \sim \mu$.*

REMARK 4.2. In the classical case, there is a phase transition in the multiplicative cascade: Kahane and Peyrière (1976) gave a necessary and sufficient condition in terms of the edge-weight distribution (namely that $EW \log W < \log N$) for $X_N^{(n)}$ to have a nontrivial limit, representing a high-temperature phase. However, their proof relies on relating the existence of a nontrivial fixed-point solution of (1.1) to the left-derivative ($EW \log W$) of the function $h \mapsto EW^h$ at $h = 1$, and this technique has no parallel in the noncommutative setting. Later we give an example in which the conditions of Theorem 4.1 are satisfied, thus furnishing a nontrivial weak limit of $X_N^{(n)}$, but it would be of interest to establish a condition on the edge-weight distribution to characterize such behavior, as is possible in the classical case. Unfortunately, we have been unable to shed any light on this aspect of the free causal tree, so the existence of a phase transition remains an open question.

Pisier (2016) provides a summary (Chapter 14) of the various results on martingales in noncommutative L_p-spaces that we need to prove Theorem 4.1. For $p \geq 1$, define $L_p(\mathcal{A})$ as the completion of a von Neumann algebra $\mathcal{A}$ under the norm $\|X\|_p = (\mathbb{E}(|X|^p))^{1/p}$, where $|X| = (X^*X)^{1/2}$. It can be shown that $L_p(\mathcal{A})$ is a Banach space, and when $p = 2$ it is a Hilbert space with inner product $\langle X, Y \rangle = \mathbb{E}(Y^*X)$. A filtration in this setting is an increasing sequence of von Neumann subalgebras $\mathcal{A}_n \subset \mathcal{A}$. For a sequence $\{X_n\}$ with $X_n \in L_2(\mathcal{A}_n)$ for each $n \geq 1$, the role of the conditional expectation $\mathbb{E}(X_n | \mathcal{A}_{n-1})$ is played by the orthogonal projection of X_n

into $L_2(\mathcal{A}_{n-1})$. The martingale property is then defined in the usual fashion by $\mathbb{E}(X_n|\mathcal{A}_{n-1}) = X_{n-1}$.

We will need to appeal to the following version of the noncommutative martingale convergence theorem (the first part of Theorem 14.2 of Pisier, 2016).

THEOREM 4.3. *If the random variable $X \in \mathcal{A}$ has finite moments of all orders, then $\mathbb{E}(X|\mathcal{A}_n) \to \mathbb{E}(X|\mathcal{A}_\infty)$ in $L_p(\mathcal{A})$ for each $p \geq 1$, where $\mathcal{A}_\infty$ is the von Neumann algebra generated by the filtration $\{\mathcal{A}_n, n \geq 0\}$.*

Let $\mathcal{A}_n$ be the initial subalgebra generated by $\{\bar{W}_\sigma : \sigma \in \bigcup_{m=1}^n \Sigma_m\}$ for $n \geq 1$, and $\mathcal{A}_0$ the subalgebra generated by $1_\mathcal{A}$. By properties of noncommutative conditional expectation (see Pisier, 2016, page 527), we have for all $n \geq 1$

$$\begin{aligned}\mathbb{E}(X_N^{(n)}|\mathcal{A}_{n-1}) &= \sum_{\sigma_1 \in \Sigma_1} \bar{W}_{\sigma_1}\Bigg[\sum_{\sigma_2 \in \Sigma_1} \bar{W}_{\sigma_1\sigma_2}\Bigg[\cdots \\ &\quad \Bigg[\sum_{\sigma_n \in \Sigma_1} \mathbb{E}(\bar{W}_{\sigma_1\cdots\sigma_n}|\mathcal{A}_{n-1})\Bigg]\cdots\Bigg]\Bigg],\end{aligned}$$

and the martingale property of $X_N^{(n)}$ then follows from

$$\mathbb{E}(\bar{W}_{\sigma_1\cdots\sigma_n}|\mathcal{A}_{n-1}) = \mathbb{E}(\bar{W}_{\sigma_1\cdots\sigma_n})1_\mathcal{A} = \frac{1}{N}1_\mathcal{A}.$$

PROOF OF THEOREM 4.1. The uniqueness of the fixed point obviously follows from the convergence part of the result. Let $Z \in \mathcal{A}$ have distribution $\mu \in \mathcal{M}^+$, a unit-mean fixed point of the Mandelbrot map. There exist freely independent elements $\{\bar{W}_\sigma, \sigma \in \Sigma\}$, having distribution ν, and freely independent elements $\{Z_\sigma, \sigma \in \Sigma\}$ having distribution μ that are free of $\{\bar{W}_\sigma, \sigma \in \Sigma\}$ such that for all $n \geq 1$

$$\begin{aligned}Z &= \sum_{\sigma_1 \in \Sigma_1} \bar{W}_{\sigma_1}\Bigg[\sum_{\sigma_2 \in \Sigma_1} \bar{W}_{\sigma_1\sigma_2}\Bigg[\cdots \\ &\quad \Bigg[\sum_{\sigma_n \in \Sigma_1} \bar{W}_{\sigma_1\cdots\sigma_n} Z_{\sigma_1\cdots\sigma_n}\Bigg]\cdots\Bigg]\Bigg].\end{aligned}$$

This is the noncommutative restatement of (7) in the proof of Theorem 1 of Kahane and Peyrière (1976). Since $\mathbb{E}(Z_{\sigma_1\cdots\sigma_n}|\mathcal{A}_n) = \mathbb{E}(Z_{\sigma_1\cdots\sigma_n})1_\mathcal{A} = 1_\mathcal{A}$, it follows that

$$\begin{aligned}&\mathbb{E}(Z|\mathcal{A}_n) \\ &= \sum_{\sigma_1 \in \Sigma_1} \bar{W}_{\sigma_1}\Bigg[\sum_{\sigma_2 \in \Sigma_1} \bar{W}_{\sigma_1\sigma_2}\Bigg[\cdots\Bigg[\sum_{\sigma_n \in \Sigma_1} \bar{W}_{\sigma_1\cdots\sigma_n}\Bigg]\cdots\Bigg]\Bigg] \\ &= X_N^{(n)}.\end{aligned}$$

Appealing to Theorem 4.3 and the assumption that μ has finite moments of all orders, we obtain $\mathbb{E}(Z|\mathcal{A}_n) \to \mathbb{E}(Z|\mathcal{A}_\infty)$ in $L_p(\mathcal{A})$ for each $p \geq 1$, where $\mathcal{A}_\infty$ is the von Neumann algebra generated by $\{\mathcal{A}_n, n \geq 0\}$.

Denote the distribution of $X_N = \mathbb{E}(Z|\mathcal{A}_\infty)$ by $\bar{\mu} \in \mathcal{M}^+$. Since $\mathbb{E}(\cdot|\mathcal{A}_\infty)$ is a contractive projection on $L_p(\mathcal{A})$, we have $\mathbb{E}([\mathbb{E}(Z|\mathcal{A}_\infty)]^p) \leq \mathbb{E}(Z^p)$ for $p \geq 2$, so $\bar{\mu}$ is determined by its moments (as μ is assumed to have this property). Also, since

$$\mathbb{E}([X_N^{(n)}]^p) = \mathbb{E}([\mathbb{E}(Z|\mathcal{A}_n)]^p) \to \mathbb{E}([\mathbb{E}(Z|\mathcal{A}_\infty)]^p)$$

for all integers $p \geq 1$, it follows that $X_N^{(n)} \sim \mu^{(n)} \to_d \bar{\mu}$. Recall that the notation $A \leq B$ for self-adjoint $A, B \in \mathcal{A}$ means that $B - A$ is nonnegative. Appealing to Corollary 3.3 of Bercovici and Voiculescu (1993), since $\mathbb{E}(Z|\mathcal{A}_\infty) \leq Z$ (and both are self-adjoint), we find that the distributions of $\mathbb{E}(Z|\mathcal{A}_\infty)$ and Z are stochastically ordered: $\bar{\mu}([t,\infty)) \leq \mu([t,\infty))$ for all $t \in \mathbb{R}$. Since $\bar{\mu}$ and μ have the same mean (namely 1), they must be equal (see, e.g., Theorem 1.A.8 of Shaked and Shanthikumar, 2007), and we conclude that $X_N^{(n)} \to_d X_N \sim \mu$. □

Our second result establishes the existence of a unit-mean fixed point of the Mandelbrot map. This fixed point is a particular instance ($\lambda = 1$) of the Marčenko–Pastur law, which for parameter values $\lambda \geq 1$ is defined by

$$d\mu_\lambda(x) = \frac{\sqrt{(b-x)(x-a)}}{2\pi x} 1_{[a,b]}(x)\,dx,$$

where $a = (1-\sqrt{\lambda})^2$ and $b = (1+\sqrt{\lambda})^2$.

THEOREM 4.4. *If the normalized edge-weights $\bar{W}_\sigma$ in the N-ary causal tree are freely independent Bernoulli random variables with distribution $\nu = (1-1/N)\delta_0 + (1/N)\delta_1$, then the Marčenko–Pastur law μ_1 is the unique fixed point of the Mandelbrot map T_ν, solving*

$$\mu_1 = [\nu \boxtimes \mu_1]^{\boxplus N}.$$

REMARK 4.5. It is easily checked that the Marčenko–Pastur law μ_1 has unit-mean, and of course it is determined by its moments (having compact support), so the conditions of Theorem 4.1 hold for T_ν with $\nu = (1 - 1/N)\delta_0 + (1/N)\delta_1$, and we obtain the nontrivial limit $X_N^{(n)} \to_d \mu_1$.

REMARK 4.6. The general Marčenko–Pastur law μ_λ is also known as the free-Poisson law (with parameter $\lambda > 0$) since it arises as the limit of a free additive convolution of Bernoulli distributions: $[(1-\lambda/n)\delta_0 + (\lambda/n)\delta_1]^{\boxplus n} \to_d \mu_\lambda$ as $n \to \infty$; see Nica and Speicher (2006). In our case with $\lambda = 1$, the Marčenko–Pastur law has density $\sqrt{(4-x)_+ x}/(2\pi x)$, which coincides with the density of the square of a standard Wigner semicircle random variable.

REMARK 4.7. Theorem 4.4 is in marked contrast to the analogous result in the classical case in which the normalized edge weights $\bar{W}_\sigma$ are independent Ber$(1/N)$ random variables. Then (1.1) has only the trivial fixed-point solution δ_0 and $X_N^{(n)} \to_d \delta_0$. This follows from Theorem 1 of Kahane and Peyrière (1976), or even more directly as $X_N^{(n)}$ can be viewed as the size Z_n of the nth generation of a Galton–Watson branching process with offspring distribution $\mathrm{Bin}(N, 1/N)$; this is the critical case because the mean number of offspring is 1. As is well known, in a critical Galton–Watson tree the population eventually becomes extinct with probability 1, and the number of generations until extinction, $\tau = \min\{n \geq 1 \colon Z_n = 0\}$, satisfies $P\{\tau > n\} \sim 2/(n\sigma^2)$, where σ^2 is the variance of the offspring distribution ($\sigma^2 = 1 - 1/N$ in this case).

REMARK 4.8. More generally, consider the classical Galton–Watson branching process with offspring distribution $\mathrm{Bin}(N, p)$ where $0 \leq p \leq 1$, having mean $m = Np$. Note that $0 \leq m < 1$ is the subcritical case, $m > 1$ the supercritical case, and $m = 1$ is the critical case discussed above. We define the free version of this process by taking the "size" Z_n of the nth generation as the sum-over-paths of length n in the N-ary tree with freely independent Bernoulli(p) edge weights. This is equivalent to defining $Z_n = m^n X_N^{(n)}$, where $X_N^{(n)}$ is the previously defined normalized sum-over-paths for the free causal tree with (unit-mean) edge weights $W_\sigma \sim \mathrm{Ber}(p)/p$. We conjecture that $X_N^{(n)} = Z_n/m^n$ converges in distribution to a nondegenerate limit for all $m \geq 1$, although we have not been able to prove this (except in the critical case $m = 1$), because finding a fixed point μ of the Mandelbrot map appears difficult unless $m = 1$. Of course, in a trivial sense the phase transition behavior is still evident: $\mathbb{E}Z_n \to 0$ in the subcritical case and $\mathbb{E}Z_n \to \infty$ in the supercritical case.

REMARK 4.9. In the late 1930s, the physicist, Leo Szilard, reinvented the theory of Galton–Watson processes to help explain how the production of free neutrons in nuclear fission can lead to a sustained chain reaction. An implication of Theorem 4.4 is that the quantum probability treatment of such processes has a nontrivial limit precisely at criticality, in contrast to the classical Galton–Watson process which dies out at criticality. Similarly, the result furnishes an example of quantum stability for a large particle system of the type suggested by Furey (2015), as mentioned in the Introduction, in reference to her proposal that irreversible time in the standard model can be explained through the evaluation of algebraic expressions arising from causal trees.

REMARK 4.10. In the classical case, the sum-over-paths Z_n in Remark 4.8 is the number of paths connecting the root to the nth level of the tree. The occurrence at least one infinite path with positive probability is known as *bond percolation*, as defined originally by Broadbent and Hammersley (1957). Braga, Sanchis and Schieber (2005) give an accessible introduction to percolation theory focusing on Bethe lattices, which are infinite unrooted trees in which each vertex has a fixed number of neighbors ($K \geq 2$, the *coordination* number). Bethe lattices are of particular interest in statistical physics because probabilistic models based on them are often exactly solvable. In the case of classically independent Bernoulli(p) edge weights on the Bethe lattice, the critical probability for percolation is $p = 1/(K-1)$, but percolation does not occur unless $p > 1/(K-1)$. In the *freely* independent case, however, as we have seen previously, percolation occurs even at the critical probability. Further, at criticality, the limiting distribution of the sum-over-paths Z_n from a fixed central vertex of the Bethe lattice to the nth shell around it is Marčenko–Pastur μ_λ with parameter $\lambda = K/(K-1)$. This can be seen by noting that any fixed central vertex of the free Bethe lattice is connected to K disjoint N-ary causal trees (each having $N = K - 1$ children at each vertex). Thus, by Theorem 4.4, the limit distribution of Z_n is the K-fold free additive convolution $[\nu \boxtimes \mu_1]^{\boxplus K}$, where $\nu = \mathrm{Ber}(1/N)$, which coincides with μ_λ from similar arguments in the proof of Theorem 4.4.

PROOF OF THEOREM 4.4. By direct integration, the Cauchy transform of the Marčenko–Pastur law μ_λ is

$$G_{\mu_\lambda}(z) = \frac{z + 1 - \lambda - \sqrt{(z - 1 - \lambda)^2 - 4z}}{2z},$$

and it follows that the R- and S-transforms of μ_λ are given by

$$R_{\mu_\lambda}(z) = zG_{\mu_\lambda}^{-1}(z) - 1 = \frac{\lambda z}{1 - z},$$

$$S_{\mu_\lambda}(z) = \frac{1}{z} R_{\mu_\lambda}^{-1}(z) = \frac{1}{\lambda + z}.$$

The S-transform of ν is

$$S_\nu(z) = \frac{1 + z}{1/N + z}$$

so the S-transform of $\nu \boxtimes \mu_1$ is

$$S_{\nu\boxtimes\mu_1}(z) = S_\nu(z) S_{\mu_1}(z) = \frac{1 + z}{1/N + z} \cdot \frac{1}{1 + z} = \frac{1}{1/N + z}$$

which is the S-transform of $\mu_{1/N}$, so $\nu \boxtimes \mu_1 = \mu_{1/N}$. The R-transform of $[\nu \boxtimes \mu_1]^{\boxplus N}$ is therefore

$$R_{[\nu\boxtimes\mu_1]^{\boxplus N}}(z) = N R_{\mu_{1/N}}(z) = N \cdot \frac{z/N}{1 - z}$$

$$= \frac{z}{1 - z} = R_{\mu_1}(z),$$

and we conclude that $\mu_1 = [\nu \boxtimes \mu_1]^{\boxplus N}$, as required. □

Our last result shows that Wigner's semicircle law is the asymptotic distribution of the centered and normalized sum-over-paths in the N-ary tree as $N \to \infty$.

THEOREM 4.11. *In free causal trees for $N \geq 2$, suppose that all the edge-weights W_σ are identically distributed with finite second moment and that $X_N \sim \mu_N$, where μ_N is a unit-mean fixed point of the Mandelbrot map T_ν for the N-ary tree. Then $X_N \to 1_\mathcal{A}$ in $L_2(\mathcal{A})$ and $\sqrt{N}(X_N - 1_\mathcal{A})$ converges in distribution as $N \to \infty$ to Wigner's semicircle law with density $2\sqrt{(R^2 - x^2)_+}/(\pi R^2)$ and radius $R = \sqrt{\mathbb{E}(W_\sigma^2) - 1}$.*

PROOF OF THEOREM 4.11. From the definition of a fixed point of the Mandelbrot map T_ν, there exist freely independent elements $\{W_i, i = 1, \dots, N\}$ that are equal in distribution to $W_\sigma \sim \nu$, and freely independent elements $\{X_{i,N}, i = 1, \dots, N\}$ having the same distribution as $X_N \sim \mu_N$ that are free of $\{W_i, i = 1, \dots, N\}$, such that (1.1) holds. This gives

$$\mathbb{E}(X_N^2) = \frac{1}{N^2}\left[\sum_{i \neq j} \mathbb{E}(W_i X_{i,N} W_j X_{j,N}) + \sum_{i=1}^{N} \mathbb{E}(W_i X_{i,N} W_i X_{i,N})\right]$$
$$= \frac{1}{N^2}\left[(N^2 - N) + N(\mathbb{E}(W_\sigma^2) + \mathbb{E}(X_N^2) - 1)\right],$$

where the last step follows from a standard free probability calculation of mixed moments; see, for example, (1.14) of Mingo and Speicher (2017). Solving for $\mathbb{E}(X_N^2)$, we obtain

$$\mathbb{E}(X_N^2) = \frac{\mathbb{E}(W_\sigma^2) + N - 2}{N - 1} = \frac{R^2}{(N-1)} + 1,$$

so

$$\mathbb{E}[(X_N - 1_\mathcal{A})^2] = \mathbb{E}(X_N^2) - 1 = R^2/(N-1) \to 0$$

as $N \to \infty$, proving the first part of the theorem. From (1.1),

$$\sqrt{N}(X_N - 1_\mathcal{A}) =_d N^{-1/2} \sum_{i=1}^{N} (W_i X_{i,N} - 1_\mathcal{A}).$$

The terms $(W_i X_i - 1_\mathcal{A})$ in the above sum are identically distributed, freely independent, have mean zero and variance

$$\mathbb{E}((W_i X_{i,N})^2) - 1 = \mathbb{E}(W_\sigma^2) + \mathbb{E}((X_N)^2) - 2 \to R^2$$

as $N \to \infty$. The second part of the theorem then follows from a special case of the free central limit theorem for triangular arrays; see Corollary 2.3 of Chistyakov and Götze (2008). □

5. DISCUSSION

The initial motivation for this paper came from a question related to particle physics, but the idea of using suitably rich algebras acting on themselves to explain the behavior of complex causal tree structures may also be useful in theoretical neuroscience. A longstanding yet still highly controversial idea in that field has been that biological neuronal networks effectively operate near phase transitions in order to enhance information processing (the critical brain hypothesis); see Brochini et al. (2016) for a recent study of the emergence of self-organized criticality in neuronal networks, along with numerous references to the literature (of over 500 papers) on this topic.

The critical brain yypothesis was initiated by some remarks of Alan Turing in his famous "Imitation Game" paper (Turing, 1950):

> "*Another simile would be an atomic pile of less than critical size: an injected idea is to correspond to a neutron entering the pile from without. Each such neutron will cause a certain disturbance which eventually dies away. If, however, the size of the pile is sufficiently increased, the disturbance caused by such an incoming neutron will very likely go on and on increasing until the whole pile is destroyed. Is there a corresponding phenomenon for minds, and is there one for machines? There does seem to be one for the human mind. The majority of them seems to be subcritical, that is, to correspond in this analogy to piles of subcritical size. An idea presented to such a mind will on average give rise to less than one idea in reply. A smallish proportion are supercritical. An idea presented to such a mind may give rise to a whole "theory" consisting of secondary, tertiary and more remote ideas ...*"

Turing's use of the "atomic pile" simile to speculate on criticality in cognition has an interesting connection with our discussion of the free Galton–Watson branching process (Remark 4.9). There is a further connection with the (highly speculative) idea that the brain may function like a quantum computer (the quantum cognition hypothesis). The physicist Matthew Fisher recently made an intriguing proposal in support of this hypothesis: the nuclear spins of phosphorus atoms could serve as neural qubits (Fisher, 2015). Indeed, if the hypothesis is true, since neuronal networks are typically organized as complex tree-like or forest-like structures (Ascoli, 2015), it is conceivable that their quantum behavior is reflected in the behavior of the free Galton–Watson process, or percolation on the free Bethe lattice, as we have described. Our finding that these processes have a nontrivial limit precisely at criticality may even be interpreted as *circumstantial* evidence that quantum effects are present in the brain, since classical versions of these processes die out at criticality.

An important problem left open by this work is to determine whether phase transitions exist in free multiplicative cascades. Indeed, despite showing that a stable high-temperature phase exists for the free Galton–Watson tree, we have been unable to find any nontrivial examples of free multiplicative cascades that have a low-temperature phase (with δ_0 as the limit distribution), even though we think it highly likely that such examples exist.

ACKNOWLEDGMENTS

The author thanks Ivan Corwin and Sunder Sethuraman for helpful discussions concerning multiplicative cascades, and he also thanks Cohl Furey for many helpful discussions on causal trees in physics.

REFERENCES

ANDERSON, G. W., GUIONNET, A. and ZEITOUNI, O. (2009). *An Introduction to Random Matrices*, 1st ed. *Cambridge Studies in Advanced Mathematics* **118**. Cambridge Univ. Press, Cambridge. MR2760897

ASCOLI, G. A. (2015). *Trees of the Brain*, *Roots of the Mind*, 1st ed. MIT Press, Cambridge.

ATTAL, S. and NECHITA, I. (2011). Discrete approximation of the free Fock space. In *Séminaire de Probabilités XLIII*. *Lecture Notes in Math*. **2006** 379–394. Springer, Berlin. MR2790382 https://doi.org/10.1007/978-3-642-15217-7_16

BARRAL, J. (2014). Mandelbrot cascades and related topics. In *Geometry and Analysis of Fractals*. *Springer Proc*. *Math*. *Stat*. **88** 1–45. Springer, Heidelberg. MR3275997 https://doi.org/10.1007/978-3-662-43920-3_1

BARRAL, J., KUPIAINEN, A., NIKULA, M., SAKSMAN, E. and WEBB, C. (2014). Critical Mandelbrot cascades. *Comm*. *Math*. *Phys*. **325** 685–711. MR3148099 https://doi.org/10.1007/s00220-013-1829-4

BERCOVICI, H. and VOICULESCU, D. (1993). Free convolution of measures with unbounded support. *Indiana Univ*. *Math*. *J*. **42** 733–773. MR1254116 https://doi.org/10.1512/iumj.1993.42.42033

BOMBELLI, L., LEE, J., MEYER, D. and SORKIN, R. D. (1987). Space-time as a causal set. *Phys*. *Rev*. *Lett*. **59** 521–524. MR0899046 https://doi.org/10.1103/PhysRevLett.59.521

BRAGA, G. A., SANCHIS, R. and SCHIEBER, T. A. (2005). Critical percolation on a Bethe lattice revisited. *SIAM Rev*. **47** 349–365. MR2179901 https://doi.org/10.1137/S0036144503433685

BROADBENT, S. R. and HAMMERSLEY, J. M. (1957). Percolation processes. I. Crystals and mazes. *Proc*. *Camb*. *Philos*. *Soc*. **53** 629–641. MR0091567 https://doi.org/10.1017/s0305004100032680

BROCHINI, L., DE ANDRADE COSTA, A., ABADI, M., ROQUE, A. C., STOL, J. and KINOUCHI, O. (2016). Phase transitions and self-organized criticality in networks of stochastic spiking neurons. *Sci*. *Rep*. **6** 35831.

CHISTYAKOV, G. P. and GÖTZE, F. (2008). Limit theorems in free probability theory. I. *Ann*. *Probab*. **36** 54–90. MR2370598 https://doi.org/10.1214/009117907000000051

COLLINS, B., HAYDEN, P. and NECHITA, I. (2017). Random and free positive maps with applications to entanglement detection. *Int*. *Math*. *Res*. *Not*. *IMRN* **2017** 869–894. MR3658153 https://doi.org/10.1093/imrn/rnw054

CORTÊS, M. and SMOLIN, L. (2014a). Quantum energetic causal sets. *Phys*. *Rev*. *D* **90** 044035.

CORTÊS, M. and SMOLIN, L. (2014b). The universe as a process of unique events. *Phys*. *Rev*. *D* **90** 084007.

DERRIDA, B. and SPOHN, H. (1988). Polymers on disordered trees, spin glasses, and traveling waves **51** 817–840. MR0971033 https://doi.org/10.1007/BF01014886

DOWKER, F. (2005). Causal sets and the deep structure of spacetime. In 100 *Years of Relativity* 445–464. World Sci. Publ., Hackensack, NJ. MR2223124 https://doi.org/10.1142/9789812700988_0016

DURRETT, R. and LIGGETT, T. M. (1983). Fixed points of the smoothing transformation. *Z*. *Wahrsch*. *Verw*. *Gebiete* **64** 275–301. MR0716487 https://doi.org/10.1007/BF00532962

FISHER, M. P. A. (2015). Quantum cognition: The possibility of processing with nuclear spins in the brain. *Ann*. *Physics* **362** 593–602. MR3411643 https://doi.org/10.1016/j.aop.2015.08.020

FUREY, C. (2015). Standard model physics from an algebra? Ph.D. thesis, University of Waterloo. Available at arXiv:1611.09182 [hep-th].

HUANG, C. (2016). Moments for multidimensional Mandelbrot cascades. *J*. *Appl*. *Probab*. **53** 1–21. MR3540776 https://doi.org/10.1017/jpr.2015.4

KAHANE, J.-P. and PEYRIÈRE, J. (1976). Sur certaines martingales de Benoit Mandelbrot. *Adv*. *Math*. **22** 131–145. MR0431355 https://doi.org/10.1016/0001-8708(76)90151-1

KARGIN, V. (2013). Lecture notes on free probability. Available at arXiv:1305.2611 [math.PR].

LIU, Q. and ROUAULT, A. (2000). Limit theorems for Mandelbrot's multiplicative cascades. *Ann*. *Appl*. *Probab*. **10** 218–239. MR1765209 https://doi.org/10.1214/aoap/1019737670

MANDELBROT, B. (1974a). Multiplications aléatoires itérées et distributions invariantes par moyenne pondérée aléatoire. *C*. *R*. *Acad*. *Sci*. *Paris*, *Sér*. *A* **278** 289–292. MR0431351

MANDELBROT, B. B. (1974b). Intermittent turbulence in self-similar cascades, divergence of high moments and dimension of the carrier. *J*. *Fluid Mech*. **62** 331–358.

MARKOPOULOU, F. (2000). The internal description of a causal set: What the universe looks like from the inside. *Comm*. *Math*. *Phys*. **211** 559–583. MR1773808 https://doi.org/10.1007/s002200050826

MINGO, J. A. and SPEICHER, R. (2017). *Free Probability and Random Matrices*. *Fields Institute Monographs* **35**. Springer, New York. MR3585560 https://doi.org/10.1007/978-1-4939-6942-5

NICA, A. and SPEICHER, R. (2006). *Lectures on the Combinatorics of Free Probability*. *London Mathematical Society Lecture Note Series* **335**. Cambridge Univ. Press, Cambridge. MR2266879 https://doi.org/10.1017/CBO9780511735127

OSSIANDER, M. and WAYMIRE, E. C. (2002). On estimation theory for multiplicative cascades. *Sankhyā Ser*. *A* **64** 323–343. MR1981762

PISIER, G. (2016). *Martingales in Banach Spaces*, 1st ed. *Cambridge Studies in Advanced Mathematics* **155**. Cambridge Univ. Press, Cambridge. MR3617459

ROSENBLATT, M. and VAN ATTA, C., eds. (1972). Statistical models and turbulence. In *Proceedings of a Symposium Held at the University of California*, *San Diego* (*La Jolla*), *Calif*., *July* 15–21, 1971. *Lecture Notes in Physics* **12**. Springer, Berlin. MR0438885

SHAKED, M. and SHANTHIKUMAR, J. G. (2007). *Stochastic Orders*. *Springer Series in Statistics*. Springer, New York. MR2265633 https://doi.org/10.1007/978-0-387-34675-5

TAO, T. (2010). Free probability. Available at terrytao.wordpress.com, 254A, Notes 5.

TURING, A. M. (1950). Computing machinery and intelligence. *Mind* **59** 433–460. MR0037064 https://doi.org/10.1093/mind/LIX.236.433