International Congress
of Mathematicians 2026
Vol. 7, P. 129–149

Applied Random Matrix Theory

Joel A. Tropp†

Abstract. Random matrices now play a role in many parts of computational mathematics. To advance these applications, it is desirable to have tools that are flexible, easy to use, and powerful. Over the last 25 years, researchers have developed a remarkable family of results, called matrix concentration inequalities, that meet the criteria. This paper offers an invitation to the field of matrix concentration and its multifarious applications.

Key words. graph algorithms, high-dimensional statistics, numerical analysis, numerical linear algebra, quantum information, random matrix theory

MSC codes. 15B52, 60B20, 65-02

DOI. 10.1137/25M1806922

1 Motivation. Random matrix theory emerged from applications in *statistics* (sample covariance matrices [58]) and in *nuclear physics* (Hamiltonians of heavy nuclei [57]). The subject established itself within the mathematical firmament through deep connections to *number theory* (zeros of the Riemann zeta function [35]), *operator algebras* (free product factors [55]), *combinatorics* (longest increasing subsequence [3]), and beyond. Over the last generation, random matrices have attracted new attention in computational mathematics, starting with work in *algorithms* [33] and in *quantum information* [1], and expanding in waves that have touched the farthest shores.

The classic literature emphasizes random matrices that enjoy a harmonious mathematical structure, such as the independent and identically distributed (i.i.d.) entries of a Wigner matrix [57]. By now, we understand these models comprehensively [2, 38]. In contrast, contemporary applications often lead to random matrices of more baroque construction, such as a random set of columns drawn from a fixed data matrix [48, sect. 5.2]. The standard methods for studying random matrices have relatively little to say about these strange examples.

To advance into this frontier, researchers have developed new tools for applied random matrix theory, collectively called *matrix concentration inequalities* [48]. These results offer several key benefits:

1. *Flexibility.* They apply to a large family of random matrices.
2. *Ease of use.* They reduce the analysis to a calculation of simple summary statistics.
3. *Power.* They accurately describe features of the random matrix model.

Matrix concentration has roots in Banach space geometry [45, 34, 40]; it also owes a heavy debt to research on quantum information theory [1]. The core results crystallized in 2010 through the efforts of Oliveira [37] and the author [46, 47]. My monograph [48] collected many applications and popularized the theory. Matrix concentration soon became textbook material [54, 56], and it has now informed thousands of research papers.

The purpose of this memoir is to introduce the fundamental matrix concentration inequalities for independent sums and martingales. I will illustrate these tools through eight contemporary applications in uncertainty quantification, numerical linear algebra, spectral graph theory, high-dimensional statistics, and quantum information science. Parts of this work are adapted from my monograph [48] and lecture notes [51].

1.1 Concentration of random matrices. A *random matrix* is a matrix whose entries are random variables, not necessarily independent from each other. The distinctive concern of random matrix theory is the action of the random matrix as a random linear map between linear spaces. In that vein, we can investigate its geometric properties (reflected in operator norms or its action on sets) and its spectral features (singular values and singular vectors, eigenvalues and eigenvectors in the square case).

†Department of Computing and Mathematical Sciences, Caltech, Pasadena, CA 91125 USA (jtropp@caltech.edu, https://tropp.caltech.edu/).

Matrix concentration studies the deviation of a random matrix $\boldsymbol{S} \in \mathbb{C}^{d_1 \times d_2}$ from its expectation $\mathbb{E}[\boldsymbol{S}]$, as measured in the spectral norm $\|\cdot\|$. For levels $t \geq 0$, we would like to control the tail probability

$$\mathbb{P}\{\|\boldsymbol{S} - \mathbb{E}[\boldsymbol{S}]\| \geq t\} \quad \leq \quad \underline{\qquad ??? \qquad}. \tag{1.1}$$

The expectation of the random matrix is computed componentwise; we always assume that this expectation is defined and finite. The spectral norm is also called the ℓ_2 operator norm; it coincides with the largest singular value. If we control the deviation $\|\boldsymbol{S} - \mathbb{E}[\boldsymbol{S}]\|$, we also control other characteristics of the random matrix:

1. *Singular values.* The singular values of $\boldsymbol{S}$ and $\mathbb{E}[\boldsymbol{S}]$ are close.
2. *Singular vectors.* The singular vectors of $\boldsymbol{S}$ and $\mathbb{E}[\boldsymbol{S}]$ are close for isolated singular values.
3. *Linear functionals.* All linear functionals of $\boldsymbol{S}$ and $\mathbb{E}[\boldsymbol{S}]$ are comparable.

Many practical questions reduce to one of these considerations.

1.2 Random matrix models. We cannot hope to make progress on the question (1.1) without imposing some restrictions on the random matrix. The challenge is to identify models that capture a wide range of examples, yet offer enough scaffolding to support strong theoretical guarantees. This paper focuses on two models inspired by the most basic scalar stochastic processes: independent sums and martingales. Our attention to these templates will be rewarded by a host of applications. Section 6 outlines recent research and alternative models.

1.2.1 Independent sums. A random matrix $\boldsymbol{S} \in \mathbb{C}^{d_1 \times d_2}$ follows the *independent sum model* when it admits the decomposition

$$\boldsymbol{S} = \sum_{k=1}^{n} \boldsymbol{X}_k \quad \text{where the family } (\boldsymbol{X}_1, \dots, \boldsymbol{X}_n) \text{ is statistically independent.}$$

We use the term "statistical independence" to mark a contrast against linear independence. Let us emphasize that the entries within any particular matrix $\boldsymbol{X}_k$ may exhibit dependencies, but $\boldsymbol{X}_k$ cannot provide any information about events involving $(\boldsymbol{X}_j : j \neq k)$. Moreover, there may be many distinct decompositions of a random matrix as an independent sum. Inter alia, the independent sum model describes a random Monte Carlo approximation of a fixed matrix as an average of simple unbiased estimates [48, sect. 6.2].

We would like to capture information about the concentration of the independent sum $\boldsymbol{S}$ through summary statistics. In the scalar setting, we can achieve this goal with the Bernstein inequality [10, Thm. 2.10], which is arguably the most useful probability inequality for an independent sum of scalars. The matrix Bernstein inequality [47, Thm. 6.2] offers a perfect analogue in the matrix setting, where it may be the single most productive tool for studying random matrices. We present this result as Theorem 2.1. Section 3 showcases applications to active subspace methods, stochastic rounding, graph sparsification, and quantum state tomography.

1.2.2 Martingales. A *matrix martingale* is a sequence $(\boldsymbol{S}_0, \boldsymbol{S}_1, \boldsymbol{S}_2, \dots)$ of random matrices with common dimension $d_1 \times d_2$ that satisfies the "status quo" property:

$$\mathbb{E}[\boldsymbol{S}_{k+1} \,|\, \boldsymbol{S}_0, \dots, \boldsymbol{S}_k] = \boldsymbol{S}_k \quad \text{for each } k = 0, 1, 2, \dots.$$

Given what we know so far, our best estimate for the next matrix $\boldsymbol{S}_{k+1}$ is the current matrix $\boldsymbol{S}_k$. As a simple example, the partial sums of an independent family of centered random matrices compose a matrix martingale. More generally, matrix martingales can model complicated sequences of random matrices that are revealed one step at a time, such as the iterates of a randomized algorithm that performs a linear algebra computation.

One of the most powerful tools for studying scalar martingale sequences is the Freedman inequality [18], which is the martingale analogue of the Bernstein inequality. The matrix Freedman inequality [37, 46] generalizes the scalar result to matrix martingales. It allows us to treat sequences of random matrices that evolve adaptively, and it provides uniform control on the whole trajectory. We present this result as Theorem 4.4. Section 5 highlights applications to online covariance estimation, Cholesky decomposition with stochastic rounding, fast graph Laplacian solvers, and randomized approximation of quantum Hamiltonians.

1.3 Notation. The Pascal notation := or =: generates a definition. The symbols $\vee$ and $\wedge$ denote the infix maximum and minimum. We work over a scalar field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, which is real or complex. The star * delivers the (conjugate) transpose of a vector or matrix. The linear space $\mathbb{F}^d$ is equipped with the standard inner product $\langle \boldsymbol{b},\ \boldsymbol{a} \rangle := \boldsymbol{b}^* \boldsymbol{a}$ and the associated ℓ_2-norm $\|\cdot\|$.

In the space $\mathbb{F}^d$, the standard basis elements are $\mathbf{e}_1, \dots, \mathbf{e}_d$, and $\mathbf{1}$ is the vector of ones. The standard basis elements in a matrix space, such as $\mathbb{F}^{d_1 \times d_2}$, are $\mathbf{E}_{ij}$. The identity matrix is $\mathbf{I}$. Dimensions are determined by context. We write a_{ij} for the entries of a matrix $\boldsymbol{A}$, and we use the colon operator $\boldsymbol{A}(i,:)$ or $\boldsymbol{A}(:,j)$ to extract the ith row or jth column. For matrices, $\|\cdot\|$ is the spectral norm (also known as the Schatten ∞-norm), $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm (or Schatten 2-norm), and $\|\cdot\|_1$ refers to the trace norm (or Schatten 1-norm).

The linear space $\mathbb{M}_d(\mathbb{F})$ consists of $d \times d$ matrices with entries in $\mathbb{F}$, on which the operator Tr computes the trace. The real-linear subspace $\mathbb{H}_d(\mathbb{F})$ consists of the $d \times d$ Hermitian (i.e., conjugate-symmetric) matrices. The cone $\mathbb{H}_d^+(\mathbb{F})$ contains the $d \times d$ positive-semidefinite (psd) matrices, and $\mathbb{H}_d^{++}(\mathbb{F})$ is the cone of positive-definite matrices. For Hermitian matrices, the semidefinite partial order $\boldsymbol{A} \preccurlyeq \boldsymbol{H}$ signifies that $\boldsymbol{H} - \boldsymbol{A}$ is psd, while the maps $\lambda_{\max}$ and $\lambda_{\min}$ return the (algebraic) maximum and minimum eigenvalues.

The operator $\mathbb{E}$ computes the expectation of a random variable, and $\mathbb{P}$ reports the probability of an event. The symbol $\sim$ means "has the distribution." For a real random variable X, the function $\sup X$ specifies its least upper bound on the probability space. Within a master probability space, the function $\sigma(\cdot)$ returns the sigma-algebra generated by its arguments; the empty argument returns the trivial sigma-algebra.

The big-$\mathcal{O}$ notation is interpreted per the conventions of computer science.

2 The matrix Bernstein inequality. Among matrix concentration inequalities, the single most important result is the matrix extension of the scalar Bernstein inequality [10, Thm. 2.10]. The matrix Bernstein inequality was established independently by Oliveira [37] in late 2009 and the author [47] in early 2010. We state the version of this result from [48, Thm. 6.1.1].

THEOREM 2.1 (Matrix Bernstein). *Consider a statistically independent family $(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n)$ of $d_1 \times d_2$ random matrices, real or complex, that are centered and uniformly bounded: $\mathbb{E}[\boldsymbol{X}_k] = \mathbf{0}$ and $\|\boldsymbol{X}_k\| \le B$ for each index k. Form the sum, and calculate its matrix variance statistic:*

$$\boldsymbol{S} := \sum_{k=1}^{n} \boldsymbol{X}_k \quad \text{and} \quad v(\boldsymbol{S}) := \|\mathbb{E}[\boldsymbol{S}\boldsymbol{S}^*]\| \vee \|\mathbb{E}[\boldsymbol{S}^*\boldsymbol{S}]\|. \tag{2.1}$$

Then the spectral norm of the sum satisfies the probability inequalities

$$\mathbb{E}\,\|\boldsymbol{S}\| \le \sqrt{2v(\boldsymbol{S})\log(d_1+d_2)} + \frac{1}{3}B\log(d_1+d_2), \tag{2.2}$$

$$\mathbb{P}\{\|\boldsymbol{S}\| \ge t\} \le (d_1+d_2)\exp\left(\frac{-t^2/2}{v(\boldsymbol{S}) + Bt/3}\right) \quad \text{for } t \ge 0. \tag{2.3}$$

The proof of Theorem 2.1 parallels the proof of the scalar Bernstein inequality, but it requires sophisticated tools from matrix analysis. The argument will occupy the rest of this section. First, we elaborate on the meaning of the result.

The significant outcome of Theorem 2.1 is the expectation bound (2.2). In contrast, the tail bound (2.3) is usually somewhat loose; it can be improved by combining the expectation bound (2.2) with scalar concentration inequalities for the norm of an independent sum, such as [10, Prob. 13.33].

We remark that Theorem 2.1 reproduces the scalar Bernstein inequality when it is applied to random scalars (i.e., 1×1 random matrices), so the numerical constants are sharp. The dimensional factor $(d_1 + d_2)$ is a new feature in the matrix setting, and it is a necessary component of the bound; see subsection 2.6.

Like the scalar variance, the matrix variance $v(\boldsymbol{S})$ reflects the magnitude of the expected "square" of the centered random matrix, where we keep in mind that the non-Hermitian matrix $\boldsymbol{S}$ has two distinct squares $\boldsymbol{S}\boldsymbol{S}^*$ and $\boldsymbol{S}^*\boldsymbol{S}$. Both terms in the matrix variance are required, although they coincide when the sum is Hermitian. It is often convenient to express the matrix variance directly in terms of the summands:

$$v(\boldsymbol{S}) = \left\| \sum_{k=1}^{n} \mathbb{E}[\boldsymbol{X}_k \boldsymbol{X}_k^*] \right\| \vee \left\| \sum_{k=1}^{n} \mathbb{E}[\boldsymbol{X}_k^* \boldsymbol{X}_k] \right\|. \tag{2.4}$$

The last identity exploits independence and centering. For simplicity, Theorem 2.1 assumes that the matrix $\boldsymbol{S}$ is centered; otherwise, note that

$$\|\boldsymbol{S}\| \le \|\mathbb{E}[\boldsymbol{S}]\| + \|\boldsymbol{S} - \mathbb{E}[\boldsymbol{S}]\|.$$

The theorem now applies to the second term.

2.1 Reduction to the Hermitian case. Let us commence with the proof of Theorem 2.1. It suffices to consider the complex case. Introduce the *Hermitian dilation*:

$$\mathcal{H} : \mathbb{C}^{d_1 \times d_2} \to \mathbb{H}_{d_1+d_2}(\mathbb{C}) \quad \text{where} \quad \mathcal{H}(\boldsymbol{A}) := \begin{bmatrix} \mathbf{0} & \boldsymbol{A} \\ \boldsymbol{A}^* & \mathbf{0} \end{bmatrix}. \tag{2.5}$$

The map $\mathcal{H}$ is real-linear, and it preserves spectral data in the sense that $\lambda_{\max}(\mathcal{H}(\boldsymbol{A})) = -\lambda_{\min}(\mathcal{H}(\boldsymbol{A})) = \|\boldsymbol{A}\|$. As a consequence, to prove Theorem 2.1, we can pass to the random Hermitian matrix

$$\mathcal{H}(\boldsymbol{S}) = \sum_{k=1}^{n} \mathcal{H}(\boldsymbol{X}_k) \quad \text{where } \mathbb{E}[\mathcal{H}(\boldsymbol{X}_k)] = \mathbf{0} \text{ and } \|\mathcal{H}(\boldsymbol{X}_k)\| \le B.$$

Through this device, we can simply instate the assumption that the summands are Hermitian matrices.

2.2 The matrix Laplace transform method. To extract information about the distribution of a bounded random *Hermitian* matrix $\boldsymbol{S} \in \mathbb{H}_d(\mathbb{C})$, define the logarithm of the *matrix moment generating function* (briefly, the *matrix log-mgf*):

$$\boldsymbol{\Xi}_{\boldsymbol{S}}(\theta) := \log \mathbb{E}\, \mathrm{e}^{\theta \boldsymbol{S}} \in \mathbb{H}_d(\mathbb{C}) \quad \text{for a parameter } \theta \in \mathbb{R}.$$

We apply a scalar function to an Hermitian matrix by applying the function to each eigenvalue without modifying the associated eigenspace. Ahlswede and Winter [1] formulated the ideas in this section, while Oliveira [37] and the author [47, 48] crystallized the arguments.

Even in the matrix setting, we can pursue the same strategy that leads to exponential tail bounds in the scalar setting [10, Chap. 2]. When the parameter $\theta > 0$,

$$\begin{aligned} \mathbb{P}\{\lambda_{\max}(\boldsymbol{S}) \ge t\} &= \mathbb{P}\left\{\mathrm{e}^{\theta \lambda_{\max}(\boldsymbol{S})} \ge \mathrm{e}^{\theta t}\right\} \le \mathrm{e}^{-\theta t} \cdot \mathbb{E}\, \mathrm{e}^{\theta \lambda_{\max}(\boldsymbol{S})} \\ &= \mathrm{e}^{-\theta t} \cdot \mathbb{E}\, \lambda_{\max}(\mathrm{e}^{\theta \boldsymbol{S}}) \le \mathrm{e}^{-\theta t} \cdot \mathbb{E}\, \operatorname{Tr} \mathrm{e}^{\theta \boldsymbol{S}} = \mathrm{e}^{-\theta t} \cdot \operatorname{Tr} \exp(\boldsymbol{\Xi}_{\boldsymbol{S}}(\theta)). \end{aligned} \tag{2.6}$$

The first inequality is Markov's. Afterward, invoke the spectral mapping theorem to draw the eigenvalue map through the exponential, and note that the trace of a psd matrix dominates its maximum eigenvalue. The emergence of the trace exponential of the matrix log-mgf unlocks powerful tools for trace functions.

A similar calculation furnishes a bound for the expectation of the maximum eigenvalue. For each $\theta > 0$,

$$\mathbb{E}\, \lambda_{\max}(\boldsymbol{S}) = \theta^{-1} \log \exp(\mathbb{E}\, \lambda_{\max}(\theta \boldsymbol{S})) \le \theta^{-1} \log \mathbb{E}\, \mathrm{e}^{\lambda_{\max}(\theta \boldsymbol{S})} \le \theta^{-1} \log \mathbb{E} \operatorname{Tr} \mathrm{e}^{\theta \boldsymbol{S}} = \theta^{-1} \log \operatorname{Tr} \exp(\boldsymbol{\Xi}_{\boldsymbol{S}}(\theta)). \tag{2.7}$$

The first inequality is Jensen's. Once again, the matrix log-mgf controls the maximum eigenvalue.

2.3 Subadditivity of the matrix log-mgf. In the scalar setting, the log-mgf of a sum of independent real random variables agrees with the sum of the log-mgfs:

$$\log \mathbb{E}\, \mathrm{e}^{\theta \sum_{k=1}^{n} X_k} = \sum_{k=1}^{n} \log \mathbb{E}\, \mathrm{e}^{\theta X_k} \quad \text{where } (X_1, \dots, X_n) \in \mathbb{R}^n \text{ is an independent family.}$$

The latter formula depends on the fact $\mathrm{e}^{a+b} = \mathrm{e}^a \mathrm{e}^b$ for $a, b \in \mathbb{R}$, a property that catastrophically fails for matrices.

Remarkably, there is a substitute. I established a *subadditivity rule* for the matrix log-mgf [47, Lem. 3.4]:

$$\operatorname{Tr} \exp\big(\boldsymbol{\Xi}_{\boldsymbol{S}}(\theta)\big) \le \operatorname{Tr} \exp\left(\sum_{k=1}^{n} \boldsymbol{\Xi}_{\boldsymbol{X}_k}(\theta)\right) \quad \text{where } \boldsymbol{S} := \textstyle\sum_{k=1}^{n} \boldsymbol{X}_k. \tag{2.8}$$

In this expression, $(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n)$ is a statistically independent family of bounded, Hermitian random matrices of the same dimension. In the proof of (2.8), the key ingredient is a deep concavity theorem of Lieb [32, Thm. 6], which is one of the crown jewels of matrix analysis. See [48, Chap. 8] for a detailed proof of Lieb's result.

2.4 The matrix Bernstein log-mgf bound. The subadditivity rule (2.8) reduces the bound on the matrix log-mgf of a sum to individual bounds on the matrix log-mgfs of the summands. Classic scalar concentration inequalities [10, Chap. 2] provide inspiration about fruitful strategies for controlling the matrix log-mgfs [47, 48].

The matrix Bernstein inequality depends on the same type of log-mgf bound as the scalar Bernstein inequality [10, Thm. 2.10]. Indeed, the result [48, Lem. 6.1] states that the matrix log-mgf of a bounded, centered, random Hermitian matrix $\boldsymbol{X} \in \mathbb{H}_d(\mathbb{C})$ satisfies the relation

$$\boldsymbol{\Xi}_{\boldsymbol{X}}(\theta) \preccurlyeq \frac{\theta^2/2}{1 - B|\theta|/3} \cdot \mathbb{E}[\boldsymbol{X}^2] \quad \text{where } \mathbb{E}[\boldsymbol{X}] = \mathbf{0} \text{ and } \|\boldsymbol{X}\| \le B \text{ and } |\theta| < 3/B. \tag{2.9}$$

The semidefinite partial order $\boldsymbol{A} \preccurlyeq \boldsymbol{H}$ on Hermitian matrices means that $\boldsymbol{H} - \boldsymbol{A}$ is psd. The proof of (2.9) tracks the scalar argument. Expand the exponential function in the matrix log-mgf as a Taylor series, use the centering to remove the term with degree one, and apply the norm bound to control the terms with degree two and higher. We also employ the fact that the matrix logarithm respects the semidefinite partial order [48, Prop. 8.4.4].

2.5 Assembly line. We are prepared to establish Theorem 2.1. Consider the sum $\boldsymbol{S} := \sum_{k=1}^n \boldsymbol{X}_k$ of independent, random Hermitian matrices with dimension d that satisfy $\mathbb{E}[\boldsymbol{X}_k] = \mathbf{0}$ and $\|\boldsymbol{X}_k\| \le B$. Sequence the displays (2.6), (2.8), and (2.9) to arrive at the tail inequality

$$\begin{aligned} \mathbb{P}\{\lambda_{\max}(\boldsymbol{S}) \ge t\} &\le \mathrm{e}^{-\theta t} \cdot \operatorname{Tr} \exp\left(\frac{\theta^2/2}{1 - B|\theta|/3} \sum_{k=1}^n \mathbb{E}[\boldsymbol{X}_k^2] \right) \\ &= \mathrm{e}^{-\theta t} \cdot \operatorname{Tr} \exp\left(\frac{\theta^2/2}{1 - B|\theta|/3} \mathbb{E}[\boldsymbol{S}^2] \right) \le d \cdot \mathrm{e}^{-\theta t} \cdot \exp\left(\frac{v(\boldsymbol{S})\theta^2/2}{1 - B|\theta|/3} \right). \end{aligned} \tag{2.10}$$

The parameter is restricted to the interval $|\theta| < B/3$. The first inequality depends on the fact that the trace exponential respects the semidefinite partial order [48, Ex. 8.1]. The identity exploits the centering and independence of the summands. Afterward, bound the trace exponential by the dimension d times the maximum eigenvalue, and use spectral mapping to recognize the matrix variance $v(\boldsymbol{S}) = \lambda_{\max}(\mathbb{E}[\boldsymbol{S}^2])$.

Finally, in (2.10), set the parameter $\theta = t/(v(\boldsymbol{S}) + Bt/3)$ to reach the finished tail inequality

$$\mathbb{P}\{\lambda_{\max}(\boldsymbol{S}) \ge t\} \le d \cdot \exp\left(\frac{-t^2/2}{v(\boldsymbol{S}) + Bt/3} \right). \tag{2.11}$$

To establish the analogous tail bound (2.3) for a general non-Hermitian sum, apply the Hermitian dilation (2.5) and invoke (2.11). The bound (2.2) on the *expected* norm of the sum follows a similar argument starting from (2.7).

2.6 Optimality. As noted, the expectation bound (2.2) is the most significant outcome from Theorem 2.1. We can strengthen this bound, but not by very much. Introduce the *tail content* statistic:

$$B_2 := \left(\mathbb{E} \max_k \|\boldsymbol{X}_k\|^2 \right)^{1/2}.$$

The tail content satisfies $B_2 \le B$; the slackness of this inequality depends on the particular choice of summands. Using another style of argument based on symmetrization [49], we can obtain a two-sided bound of the form

$$\sqrt{v(\boldsymbol{S})} \;+\; B_2 \quad \lesssim \quad \mathbb{E}\|\boldsymbol{S}\| \quad \lesssim \quad \sqrt{v(\boldsymbol{S}) \log(d_1 + d_2)} \;+\; B_2 \log(d_1 + d_2). \tag{2.12}$$

The paper [49] provides examples to show that each term in (2.12) is necessary, so we cannot improve the bounds without a substantial amount of extra information (about the covariance structure of the random matrix). Section 6 mentions some situations where improvements are possible.

3 Independent sum model: Applications. In this section, we present some contemporary applications of the matrix Bernstein inequality in several areas of computational mathematics. As a caveat, these stylized applications may not reflect the intricacies of each problem. It is sometimes possible to find a simpler argument by invoking a more suitable matrix concentration inequality, and we can occasionally obtain sharper analyses using more complicated tools; see section 6.

3.1 Active subspace methods. In engineering design, researchers build computer models of complex physical systems that are governed by several input parameters. Key tasks include optimization of parameters, analysis of sensitivity to changes in parameters, and quantification of uncertainty about the system output. For example, an aeronautics engineer designs an airfoil by modifying its geometry to reach a target lift and drag coefficient. It is challenging to explore the parameter space of a complicated, nonlinear model, so it is common to seek reduced models. One basic methodology is to restrict our attention to the most salient directions in the input space, which we can identify through a random sampling procedure. The analysis of this approach depends on matrix concentration inequalities. This vignette is adapted from Constantine's book [16, Chap. 3]. The mathematics are similar with the classic problem of covariance estimation; see subsection 5.1.

Let $f : \mathbb{R}^d \to \mathbb{R}$ be an L-Lipschitz, differentiable function that models a complicated system, and assume that we can evaluate its gradient ∇f at each point in the parameter space—but the computational cost is significant. As part of the model, introduce a random vector $\boldsymbol{z} \in \mathbb{R}^d$ that describes a distribution over the parameter space; this distribution is often interpreted as a Bayesian prior. To capture the variability of the function, introduce the psd *sensitivity matrix*

$$\boldsymbol{\Sigma} := \mathbb{E}\left[(\nabla f(\boldsymbol{z}))(\nabla f(\boldsymbol{z}))^*\right] \in \mathbb{H}_d^+(\mathbb{R}). \tag{3.1}$$

The quadratic form in $\boldsymbol{\Sigma}$ reflects the sensitivity of the function to directional perturbations in the input space:

$$\boldsymbol{a}^*\boldsymbol{\Sigma}\boldsymbol{a} = \mathbb{E}\left[|\langle \nabla f(\boldsymbol{z}),\ \boldsymbol{a}\rangle|^2\right] \quad \text{for each } \boldsymbol{a} \in \mathbb{R}^d.$$

The subspace spanned by the leading eigenvectors of $\boldsymbol{\Sigma}$ is called an *active subspace* of the model, because it captures the most salient directions in the input space.

We cannot evaluate the sensitivity matrix $\boldsymbol{\Sigma}$ explicitly, but we can construct a Monte Carlo approximation:

$$\widehat{\boldsymbol{\Sigma}}_n := \frac{1}{n}\sum_{k=1}^{n} (\nabla f(\boldsymbol{z}_k))(\nabla f(\boldsymbol{z}_k))^* \quad \text{where } \boldsymbol{z}_k \sim \boldsymbol{z} \text{ i.i.d.} \tag{3.2}$$

How many gradient samples suffice to approximate the sensitivity matrix $\boldsymbol{\Sigma}$? Can we determine the directions in which the model varies substantially?

THEOREM 3.1 (Sampling active subspaces [16, Thm. 3.7]). *Let $f : \mathbb{R}^d \to \mathbb{R}$ be L-Lipschitz and differentiable. As in* (3.1) *and* (3.2)*, consider the sensitivity matrix $\boldsymbol{\Sigma} \in \mathbb{H}_d(\mathbb{R})$ and the empirical approximation $\widehat{\boldsymbol{\Sigma}}_n \in \mathbb{H}_d(\mathbb{R})$ determined by n gradient samples. The expectation of the relative approximation error satisfies*

$$\mathbb{E}\left[\frac{\|\widehat{\boldsymbol{\Sigma}}_n - \boldsymbol{\Sigma}\|}{\|\boldsymbol{\Sigma}\|}\right] \le \sqrt{2\beta_n} + \frac{1}{3}\beta_n \quad \textit{where} \quad \beta_n := \frac{L^2 \log(2d)}{\|\boldsymbol{\Sigma}\|} \cdot \frac{1}{n}.$$

For $\varepsilon \in (0, 1)$, the sample complexity $n \ge 4\varepsilon^{-2} L^2 \log(2d)/\|\boldsymbol{\Sigma}\|$ results in expected relative error at most ε.

The ratio $r := L^2/\|\boldsymbol{\Sigma}\|$ reflects the number of energetic dimensions in the model. Given $n = \mathcal{O}(r \log d)$ gradient samples, the estimator $\widehat{\boldsymbol{\Sigma}}_n$ for the sensitivity matrix $\boldsymbol{\Sigma}$ allows us to find this active subspace. To reduce the relative error to a level ε, the estimator requires an additional factor of ε^{-2} samples, so the empirical approximation does not attain high accuracy. This is a familiar weakness of Monte Carlo methods.

Proof sketch. The error in the Monte Carlo approximation is an instance of the independent sum model. Abbreviate $\boldsymbol{y}_k := \nabla f(\boldsymbol{z}_k)$ for each index k. Since the function f is L-Lipschitz, the random vectors satisfy the uniform bound $\|\boldsymbol{y}_k\| \le L$. Introduce the matrix approximation error

$$\boldsymbol{S} := \widehat{\boldsymbol{\Sigma}}_n - \boldsymbol{\Sigma} = \sum_{k=1}^{n} \frac{1}{n}\left(\boldsymbol{y}_k\boldsymbol{y}_k^* - \mathbb{E}[\boldsymbol{y}_k\boldsymbol{y}_k^*]\right) =: \sum_{k=1}^{n} \boldsymbol{X}_k.$$

The summands $(\boldsymbol{X}_k)$ compose an independent family of bounded, centered, random Hermitian matrices with common dimension d. The statistics appearing in the matrix Bernstein inequality (Theorem 2.1) satisfy

$$B = \sup \|\boldsymbol{X}_k\| \le \frac{L^2}{n} \quad \text{and} \quad v(\boldsymbol{S}) = \|\mathbb{E}[\boldsymbol{S}^2]\| = n \cdot \|\mathbb{E}[\boldsymbol{X}_1^2]\| \le \frac{1}{n}\|\mathbb{E}[(\boldsymbol{y}_1\boldsymbol{y}_1^*)^2]\| \le \frac{L^2}{n} \cdot \|\boldsymbol{\Sigma}\|.$$

The stated result follows from (2.2). ∎

3.2 Stochastic rounding. Conventional computer architectures offer floating-point numbers with 16 decimal digits of precision or more. Driven by contemporary applications, computer engineers have started to develop new architectures with far lower precision—sometimes as few as 4 or 8 bits (i.e., 1–2 digits). At this extreme, large numerical errors can accumulate as we round successive calculations to the nearest machine-representable number. One way to mitigate these errors is to design computer systems that perform *stochastic rounding.* To analyze linear algebra computations that employ stochastic rounding, we can exploit matrix concentration tools. This section offers an idealized treatment of the simplest problem; see [17] for more texture.

A *floating-point number system* is a finite set of machine-representable numbers $\mathsf{F} \subset \mathbb{R}$. Ignoring underflow and overflow, a (deterministic) rounding rule approximates each real number $a \in \mathbb{R}$ by a floating-point number $\mathtt{float}(a) \in \mathsf{F}$ that admits the relative-error bound

$$|a - \mathtt{float}(a)| \le \mathtt{u} \cdot |a| \quad \text{where } \mathtt{u} \in \mathbb{R} \text{ is the unit-roundoff parameter.} \tag{3.3}$$

In IEEE double-precision arithmetic, the unit-roundoff parameter $\mathtt{u} = 2^{-53} \approx 10^{-16}$, and the rounding rule must satisfy additional requirements to meet the standard [17, Tables 1 and 2]. As an alternative, a *stochastic rounding rule* maps each real number $a \in \mathbb{R}$ to a *random* floating-point number $\mathtt{sr}(a) \in \mathsf{F}$ that satisfies

$$\mathbb{E}[\mathtt{sr}(a)] = a \quad \text{and} \quad |a - \mathtt{sr}(a)| \le \mathtt{u} \cdot |a|. \tag{3.4}$$

In other words, stochastic rounding is unbiased, and it commits a relative error on the order of the unit roundoff.

Suppose that we round each entry of a real-valued matrix stochastically to the nearest floating-point number. How much damage will we do? How does this bound compare with deterministic rounding?

THEOREM 3.2 (Stochastic rounding). *Consider a matrix $\boldsymbol{A} \in \mathbb{R}^{d_1 \times d_2}$. For $p \in \{1, 2\}$, define the norms*

$$\|\boldsymbol{A}\|_{\max} := \max_{ij} |a_{ij}| \quad \textit{and} \quad \|\boldsymbol{A}\|_{\mathrm{rc},p} := (\max_i \|\boldsymbol{A}(i,:)\|_p) \vee (\max_j \|\boldsymbol{A}(:,j)\|_p).$$

Form a random matrix $\mathtt{sr}(\boldsymbol{A}) \in \mathsf{F}^{d_1 \times d_2}$ by independently rounding each entry of $\boldsymbol{A}$, honoring the rule (3.4)*. Then*

$$\mathbb{E}\,\|\boldsymbol{A} - \mathtt{sr}(\boldsymbol{A})\| \le \mathtt{u} \cdot \left[\sqrt{2\|\boldsymbol{A}\|_{\mathrm{rc},2}^2 \log(d_1 + d_2)} + \frac{1}{3}\|\boldsymbol{A}\|_{\max} \log(d_1 + d_2)\right].$$

In contrast, the deterministic rule (3.3) *only guarantees that $\|\boldsymbol{A} - \mathtt{float}(\boldsymbol{A})\| \le \mathtt{u} \cdot \|\boldsymbol{A}\|_{\mathrm{rc},1}$.*

To interpret the result, it is helpful to consider a square matrix with dimension d whose entries all have similar magnitudes. In this case, the stochastic rounding bound is governed by the norm $\|\boldsymbol{A}\|_{\mathrm{rc},2}$, which is proportional to $\sqrt{d}$. Meanwhile, the deterministic rounding bound depends on $\|\boldsymbol{A}\|_{\mathrm{rc},1}$, which is proportional to the matrix dimension d. This contrast indicates that stochastic rounding confers a significant benefit.

Proof sketch. The matrix of rounding errors is an instance of the independent sum model:

$$\boldsymbol{S} := \boldsymbol{A} - \mathtt{sr}(\boldsymbol{A}) = \sum_{ij} \varepsilon_{ij} a_{ij} \mathbf{E}_{ij} \quad \text{where } \mathbb{E}[\varepsilon_{ij}] = 0 \text{ and } |\varepsilon_{ij}| \le \mathtt{u}.$$

The summands $\boldsymbol{X}_{ij} := \varepsilon_{ij} a_{ij} \mathbf{E}_{ij}$ are bounded, centered independent random matrices with dimension $d_1 \times d_2$. A short calculation confirms that the parameters in the matrix Bernstein inequality (Theorem 2.1) satisfy

$$B = \sup \|\boldsymbol{X}_{ij}\| \le \mathtt{u} \cdot \|\boldsymbol{A}\|_{\max} \quad \text{and} \quad v(\boldsymbol{S}) = \|\mathbb{E}[\boldsymbol{S}\boldsymbol{S}^*]\| \vee \|\mathbb{E}[\boldsymbol{S}^*\boldsymbol{S}]\| \le \mathtt{u}^2 \cdot \|\boldsymbol{A}\|_{\mathrm{rc},2}^2.$$

The bound on the expectation of the error follows directly from (2.2). The error bound for deterministic rounding holds because $\|\boldsymbol{M}\| \le \|\boldsymbol{M}\|_{\mathrm{rc},1}$ for any matrix $\boldsymbol{M}$; see [22, Prob. 5.6P21]. ∎

In fact, the error estimate for stochastic rounding in Theorem 3.2 admits a modest refinement:

$$\mathbb{E}\,\|\boldsymbol{A} - \mathtt{sr}(\boldsymbol{A})\| \le 4\mathtt{u} \cdot \left[\|\boldsymbol{A}\|_{\mathrm{rc},2} + \|\boldsymbol{A}\|_{\max}\sqrt{1 + \log(d_1 \wedge d_2)}\right].$$

This bound follows from a specialized result [12, Thm. 1.4] for random matrices with independent entries, plus a symmetrization argument [54, Lem. 6.4.2]. For most practical use cases, the improvement is insubstantial.

3.3 Graph sparsification. A combinatorial graph encodes the pairwise relationships among a family of objects. We may ask whether it is possible to find a simpler graph (with fewer edges) that preserves structural properties of the original graph, including the weights of vertex cuts and the mixing time of a random walk. Spielman and Srivastava [41] showed how to achieve this goal by randomly sampling edges from the graph. We can analyze the procedure with matrix concentration.

Consider a connected, weighted, undirected graph (V, w). The graph comprises a set $\mathsf{V} := \{1, \dots, n\}$ of vertices and a symmetric function $w : \mathsf{V} \times \mathsf{V} \to \mathbb{R}_+$ that assigns a nonnegative weight to each pair of vertices. We require that $w_{ii} = 0$ and $w_{ij} = w_{ji}$ for all $i, j \in \mathsf{V}$. The number of edges $m := \#\{i < j : w_{ij} > 0\}$.

We can also represent the graph by means of the *graph Laplacian* $\boldsymbol{L} \in \mathbb{H}_n^+(\mathbb{R})$, which is the psd matrix

$$\boldsymbol{L} := \sum_{i<j} w_{ij} \cdot (\mathbf{e}_i - \mathbf{e}_j)(\mathbf{e}_i - \mathbf{e}_j)^* =: \sum_{i<j} w_{ij} \cdot \boldsymbol{\Delta}_{ij}. \tag{3.5}$$

The sum involves m nonzero terms. Since the graph is connected, $\operatorname{null}(\boldsymbol{L}) = \operatorname{span}\{\mathbf{1}\}$. Without further notice, we restrict $\boldsymbol{L}$ to the $(n-1)$-dimensional subspace where it is nonsingular. Define the *effective resistance* of each vertex pair in the graph:

$$\varrho_{ij} := \operatorname{Tr}[\boldsymbol{L}^{-1/2} \boldsymbol{\Delta}_{ij} \boldsymbol{L}^{-1/2}] \quad \text{for each vertex pair } (i, j).$$

If the graph models an electrical network whose wires have conductances w_{ij}, then the effective resistance ϱ_{ij} is the voltage required to push one unit of current from vertex i to vertex j. Observe that $\sum_{i<j} w_{ij} \varrho_{ij} = n - 1$. We can approximate all m of the nonzero effective resistances in $\mathcal{O}(m \log n)$ arithmetic operations [41, Thm. 2].

Our goal is to produce a *sparse* graph whose Laplacian matrix $\widehat{\boldsymbol{L}} \in \mathbb{H}_n^+(\mathbb{R})$ is comparable with the original Laplacian $\boldsymbol{L}$ in the semidefinite partial order:

$$(1 - \varepsilon)\boldsymbol{L} \preccurlyeq \widehat{\boldsymbol{L}} \preccurlyeq (1 + \varepsilon)\boldsymbol{L} \quad \text{for a parameter } \varepsilon \in (0, 1). \tag{3.6}$$

The spectral equivalence (3.6) ensures that the sparse graph is structurally similar to the original graph. We plan to construct a spectral sparsifier $\widehat{\boldsymbol{L}}$ by randomly sampling edges according to a carefully chosen probability distribution. Introduce the random matrix $\boldsymbol{Y} \in \mathbb{H}_d^+(\mathbb{R})$ with distribution

$$\mathbb{P}\left\{ \boldsymbol{Y} = \frac{n-1}{\varrho_{ij}} \boldsymbol{\Delta}_{ij} \right\} = \frac{w_{ij} \varrho_{ij}}{n-1} \quad \text{for each } i < j \text{ where } \varrho_{ij} > 0.$$

In other words, $\boldsymbol{Y}$ is the Laplacian of a weighted edge between a pair (i, j) of vertices, chosen with probability proportional to the weight w_{ij} and the effective resistance ϱ_{ij}. It is easy to confirm that $\mathbb{E}[\boldsymbol{Y}] = \boldsymbol{L}$, so the random matrix $\boldsymbol{Y}$ is an unbiased estimator for the Laplacian $\boldsymbol{L}$ of the original graph. By averaging q independent copies of $\boldsymbol{Y}$, we obtain the Laplacian of a random sparse graph with at most q edges:

$$\widehat{\boldsymbol{L}} := \frac{1}{q} \sum_{k=1}^{q} \boldsymbol{Y}_k \quad \text{where } \boldsymbol{Y}_k \sim \boldsymbol{Y} \text{ i.i.d.} \tag{3.7}$$

How many edges q suffice to achieve the approximation guarantee (3.6)?

THEOREM 3.3 (Graph sparsification [41, Thm. 1]). *Let $\boldsymbol{L}$ be the Laplacian* (3.5) *of a weighted graph on n vertices, and let $\widehat{\boldsymbol{L}}$ be the Laplacian* (3.7) *of a random sparse graph with at most q edges. With probability at least $1 - \delta$, the random matrix $\widehat{\boldsymbol{L}}$ is an ε-approximation of $\boldsymbol{L}$ in the sense of* (3.6) *provided that $q \geq 3\varepsilon^{-2} n \log(2n/\delta)$.*

The result ensures that we can approximate a graph on n vertices by a sparse graph with $q = \mathcal{O}(n \log n)$ edges, regardless of the number m of edges in the original graph. We pay an extra factor ε^{-2} in the number of edges q to achieve accuracy ε. Although the randomized method is simple and efficient, there is a more expensive deterministic algorithm that finds an ε-spectral sparsifier with just $q = \mathcal{O}(n/\varepsilon^2)$ edges [8].

Proof sketch. Define the linear map $\boldsymbol{\Phi}(\boldsymbol{A}) := \boldsymbol{L}^{-1/2} \boldsymbol{A} \boldsymbol{L}^{-1/2}$ on symmetric matrices that act on range$(\boldsymbol{L})$. The spectral equivalence (3.6) is the same as the requirement that $\|\boldsymbol{\Phi}(\widehat{\boldsymbol{L}} - \boldsymbol{L})\| \leq \varepsilon$. The matrix inside the norm can be written as an independent sum of centered random matrices acting on an $(n-1)$-dimensional space:

$$\boldsymbol{S} := \boldsymbol{\Phi}(\widehat{\boldsymbol{L}} - \boldsymbol{L}) = \sum_{k=1}^{q} \frac{1}{q} \boldsymbol{\Phi}(\boldsymbol{Y}_k - \mathbb{E}[\boldsymbol{Y}_k]) =: \sum_{k=1}^{q} \boldsymbol{X}_k.$$

To activate the matrix Bernstein inequality (Theorem 2.1), check that

$$B = \sup \|\boldsymbol{X}_k\| = \frac{n-1}{q} \quad \text{and} \quad v(\boldsymbol{S}) \le \frac{n-1}{q}.$$

The result now follows from the tail inequality (2.3). For details, see [51, sect. 5.2]. ∎

3.4 Quantum tomography. The state of a finite-dimensional quantum system, such as a register in a quantum computer, is characterized by a finite-dimensional psd matrix. We can probe the system by taking measurements, but the laws of quantum mechanics imply that each measurement returns a random number. As a consequence, methods for reconstructing the state of a quantum system lead to problems involving random matrices. Matrix concentration tools offer a quick way to analyze quantum state estimators. This section summarizes a lecture by Richard Kueng (see [51, Lecture 3]), which distills ideas from a paper of Guţă et al. [19]. Related ideas animate the theory of classical shadows, a major recent advance in quantum information science [27].

We model a quantum system by means of its *density matrix,* which is a complex psd matrix with trace one. The convex, compact set $\mathcal{D}_d := \{\boldsymbol{\rho} \in \mathbb{H}_d^+(\mathbb{C}) : \operatorname{Tr}[\boldsymbol{\rho}] = 1\}$ collects the density matrices of dimension d. Next, a *quantum measurement* is a family $\{\boldsymbol{H}_1, \dots, \boldsymbol{H}_m\} \subset \mathbb{H}_d^+(\mathbb{C})$ of psd matrices that partitions the identity: $\sum_{j=1}^m \boldsymbol{H}_j = \mathbf{I}$. When we apply the measurement to a quantum system with density matrix $\boldsymbol{\rho} \in \mathcal{D}_d$, two things happen. First, we sample a discrete random variable J with distribution

$$\mathbb{P}\{J = j \,|\, \boldsymbol{\rho}\} = \operatorname{Tr}[\boldsymbol{H}_j \boldsymbol{\rho}] \quad \text{for } j = 1, \dots, m. \tag{3.8}$$

Second, the quantum system "collapses." Therefore, to characterize a quantum system, we must prepare several copies in the same state and measure them repeatedly to obtain statistical evidence about the state.

For simplicity, we consider a special type of quantum measurement, called a *complex projective* 2*-design.* Here, each of the measurement matrices has rank one: $\boldsymbol{H}_j = (d/m)\boldsymbol{u}_j\boldsymbol{u}_j^*$ where $\boldsymbol{u}_j \in \mathbb{C}^d$ is a unit-norm vector. Moreover, these measurements support the reconstruction property

$$\frac{1}{m}\sum_{j=1}^m \operatorname{Tr}[\boldsymbol{u}_j\boldsymbol{u}_j^*\boldsymbol{A}] \cdot \boldsymbol{u}_j\boldsymbol{u}_j^* = \frac{1}{(d+1)d}(\boldsymbol{A} + \operatorname{Tr}[\boldsymbol{A}] \cdot \mathbf{I}) \quad \text{for each matrix } \boldsymbol{A} \in \mathbb{H}_d(\mathbb{C}). \tag{3.9}$$

The vectors $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ composing a complex projective 2-design are arranged very regularly over the complex sphere. Examples include systems of "mutually unbiased" orthonormal bases and systems of equiangular lines [51, Ex. 3.3]. This type of measurement system can stably distinguish any pair of density matrices.

To reconstruct a quantum system with density matrix $\boldsymbol{\rho} \in \mathcal{D}_d$, we perform a measurement to sample the random variable J. Then we build a random Hermitian matrix $\boldsymbol{Y} \in \mathbb{H}_d(\mathbb{C})$ according to the rule

$$\boldsymbol{Y} = (d+1)\boldsymbol{u}_J\boldsymbol{u}_J^* - \mathbf{I}.$$

The random matrix $\boldsymbol{Y}$ serves as an unbiased estimator for the density matrix:

$$\mathbb{E}[\boldsymbol{Y}] = \sum_{j=1}^m ((d+1)\boldsymbol{u}_j\boldsymbol{u}_j^* - \mathbf{I}) \cdot \mathbb{P}\{J = j \,|\, \boldsymbol{\rho}\} = \boldsymbol{\rho}.$$

This calculation follows from (3.8) and (3.9). By repeating the measurement on n independent (i.e., unentangled) copies of the quantum system, each with density matrix $\boldsymbol{\rho}$, we can obtain an estimator

$$\boldsymbol{S}_n := \frac{1}{n}\sum_{k=1}^n \boldsymbol{Y}_k \quad \text{where } \boldsymbol{Y}_k \sim \boldsymbol{Y} \text{ i.i.d.} \tag{3.10}$$

How many independent measurements suffice to approximate the underlying density matrix?

THEOREM 3.4 (Quantum tomography). *Let $\boldsymbol{\rho} \in \mathcal{D}_d$ be a density matrix with dimension d. Using a complex projective* 2*-design* (3.9)*, perform a quantum measurement on each of n independent quantum systems with common state $\boldsymbol{\rho}$. Form the sample average estimator $\boldsymbol{S}_n$, as in* (3.10)*. For $\varepsilon \in (0, 1)$,*

$$\mathbb{P}\{\|\boldsymbol{S}_n - \boldsymbol{\rho}\| \ge \varepsilon\} \le 2d \cdot \exp\left(\frac{-3n\varepsilon^2}{14d}\right).$$

In particular, the sample complexity $n \ge 5\varepsilon^{-2} d \log(2d/\delta)$ yields a failure probability no greater than $\delta \in (0, 1)$.

Proof sketch. To apply the matrix Bernstein inequality (Theorem 2.1) to the estimator (3.10), simply compute

$$B = \sup \|n^{-1}\boldsymbol{Y}\| = \frac{d}{n} \quad \text{and} \quad v(\boldsymbol{S}_n) = \frac{1}{n}\|\mathbb{E}[\boldsymbol{Y}^2]\| = \frac{1}{n}\|(d-1)\boldsymbol{\rho} + d\,\mathbf{I}\| \le \frac{2d-1}{n}.$$

The statement follows quickly from (2.3). ∎

While the sample average estimator (3.10) is unbiased, it generally does not produce a density matrix. Moreover, the trace norm yields a more interpretable distance between density matrices than the spectral norm. To address these concerns, define the *projected state estimator* using the Frobenius norm:

$$\widehat{\boldsymbol{\rho}}_n \in \arg\min_{\boldsymbol{\sigma}\in\mathcal{D}_d} \|\boldsymbol{\sigma} - \boldsymbol{S}_n\|_{\mathrm{F}}.$$

Lemma 3.9 of [51] ensures that the projected state estimator satisfies the trace-norm bound

$$\|\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}\|_1 \le 4r\|\boldsymbol{S}_n - \boldsymbol{\rho}\| \quad \text{where } r := \operatorname{rank}(\boldsymbol{\rho}).$$

As a consequence, we arrive at an upper bound for the sample complexity of the projected state estimator:

$$\mathbb{P}\{\|\widehat{\boldsymbol{\rho}}_n - \boldsymbol{\rho}\|_1 \ge \varepsilon\} \le \delta \quad \text{when} \quad n \ge 80\varepsilon^{-2}r^2 d\log(2d/\delta).$$

This bound implies that the projected state estimator $\widehat{\boldsymbol{\rho}}_n$ nearly achieves the *optimal* sample complexity that is possible in this setting [20].

4 Concentration for matrix martingales. The full majesty of the framework for exponential matrix concentration comes into view when we extend it to dynamically evolving sequences of matrices.

4.1 Matrix martingales. Fix a probability space. A *filtration* is an increasing sequence of sigma-algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$ inside the master sigma-algebra. A *matrix martingale* is a sequence $(\boldsymbol{S}_0, \boldsymbol{S}_1, \boldsymbol{S}_2, \dots)$ of complex random matrices with common dimension $d_1 \times d_2$ that satisfies three properties. For each index $k = 0, 1, 2, \dots$, the sequence is

1. *Adapted.* Each random matrix $\boldsymbol{S}_k$ is measurable with respect to $\mathcal{F}_k$.
2. *Integrable.* The expected norm is finite: $\mathbb{E}\,\|\boldsymbol{S}_k\| < +\infty$.
3. *Status quo.* The conditional expectation $\mathbb{E}[\boldsymbol{S}_{k+1} \,|\, \mathcal{F}_k] = \boldsymbol{S}_k$.

The *difference sequence* consists of the matrices $\boldsymbol{X}_k := \boldsymbol{S}_k - \boldsymbol{S}_{k-1}$ with $k \ge 1$. Each member of the difference sequence is conditionally centered: $\mathbb{E}[\boldsymbol{X}_k \,|\, \mathcal{F}_{k-1}] = \mathbf{0}$. Matrix martingales model a wide range of examples.

Example 4.1 (Adapted sums). Let $\boldsymbol{S}_0 := \boldsymbol{A}$ be a fixed matrix. For each index $k \ge 1$, suppose that $\boldsymbol{S}_k := \boldsymbol{S}_{k-1} + \boldsymbol{X}_k$ where the increments are conditionally centered: $\mathbb{E}[\boldsymbol{X}_k \,|\, \boldsymbol{X}_0, \dots, \boldsymbol{X}_{k-1}] = \mathbf{0}$. Then $(\boldsymbol{S}_0, \boldsymbol{S}_1, \boldsymbol{S}_2, \dots)$ is a matrix martingale with respect to the filtration $\mathcal{F}_k := \sigma(\boldsymbol{S}_0, \boldsymbol{X}_1, \dots, \boldsymbol{X}_k)$ defined for $k \ge 0$.

Example 4.2 (Adapted products). Let $\boldsymbol{S}_0 := \boldsymbol{A}$ be a fixed matrix. Suppose that $\boldsymbol{S}_k := \boldsymbol{Y}_k\boldsymbol{S}_{k-1}$ where the factors satisfy $\mathbb{E}[\boldsymbol{Y}_k \,|\, \boldsymbol{S}_0, \dots, \boldsymbol{S}_{k-1}] = \mathbf{I}$. Then $(\boldsymbol{S}_0, \boldsymbol{S}_1, \boldsymbol{S}_2, \dots)$ composes a matrix martingale with respect to the filtration $\mathcal{F}_k := \sigma(\boldsymbol{S}_0, \boldsymbol{Y}_1, \dots, \boldsymbol{Y}_k)$ defined for $k \ge 0$.

Example 4.3 (Lévy–Doob martingale). Fix a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$. Let $\boldsymbol{S}$ be a random $d_1 \times d_2$ matrix with $\mathbb{E}\,\|\boldsymbol{S}\| < +\infty$. Construct the sequence of conditional expectations:

$$\boldsymbol{S}_k := \mathbb{E}[\boldsymbol{S} \,|\, \mathcal{F}_k] \quad \text{for each } k = 0, 1, 2, \dots.$$

Then $(\boldsymbol{S}_0, \boldsymbol{S}_1, \boldsymbol{S}_2, \dots)$ composes a martingale with respect to the filtration.

To avoid technicalities, we only consider finite martingale sequences $(\boldsymbol{S}_0, \dots, \boldsymbol{S}_n)$. Furthermore, we require the elements of the martingale to be uniformly bounded in the sense that $\sup \|\boldsymbol{S}_k\| \le B_\infty < +\infty$ for each index $k = 0, \dots, n$. The filtration may be suppressed if it is determined by context.

4.2 Matrix Freedman. To analyze a matrix martingale, the most valuable theorem is the matrix Freedman inequality, originally due to Oliveira [37]. The author exploited the subadditivity (2.8) of the matrix log-mgf to refine Oliveira's result and to establish some other matrix martingale inequalities [46]. The version stated here follows from [46, Cor. 1.3, Thm. 3.1].

THEOREM 4.4 (Matrix Freedman). *Consider a finite matrix martingale sequence* $(\boldsymbol{S}_0, \dots, \boldsymbol{S}_n)$ *consisting of* $d_1 \times d_2$ *matrices, real or complex. Suppose that the elements of the difference sequence* $(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n)$ *are uniformly bounded:* $\|\boldsymbol{X}_k\| \le B$. *Introduce the conditional quadratic variation sequence:*

$$V_k := \left\| \sum_{j=1}^{k} \mathbb{E}\left[\boldsymbol{X}_j \boldsymbol{X}_j^* \,|\, \mathcal{F}_{j-1}\right] \right\| \vee \left\| \sum_{j=1}^{k} \mathbb{E}\left[\boldsymbol{X}_j^* \boldsymbol{X}_j \,|\, \mathcal{F}_{j-1}\right] \right\} \quad \text{for } k = 1, \dots, n.$$

For parameters $v, t \ge 0$, *we have the probability inequalities*

$$\mathbb{P}\left\{ \exists k : \|\boldsymbol{S}_k - \boldsymbol{S}_0\| \ge \sqrt{2vt} + Bt/3 \text{ and } V_k \le v \right\} \le (d_1 + d_2)\, \mathrm{e}^{-t}, \tag{4.1}$$

$$\mathbb{P}\left\{ \exists k : \|\boldsymbol{S}_k - \boldsymbol{S}_0\| \ge t \text{ and } V_k \le v \right\} \le (d_1 + d_2) \exp\left(\frac{-t^2/2}{v + Bt/3} \right). \tag{4.2}$$

The rest of this section presents a proof of Theorem 4.4 that combines ideas from my paper [46] and from the survey of Howard et al. [23]. First, we dilate on the meaning of the result.

The conditional quadratic variation is the martingale analogue of the matrix variance; cf. (2.4). At each step j of the process, we compute the expected "squares" of the element $\boldsymbol{X}_j$ in the difference sequence, given data $\mathcal{F}_{j-1}$ about the previous position of the martingale, and we combine the results. Thus, V_k is a random variable that reflects the total observed volatility of the matrix martingale up to time k. When the difference sequence $(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n)$ is statistically independent, each V_k reduces to the matrix variance $v(\boldsymbol{S}_k)$.

The probability inequalities (4.1) and (4.2) state that the norm of the martingale element $\|\boldsymbol{S}_k\|$ is unlikely to be large when the conditional quadratic variation V_k is bounded above. The first inequality (4.1) formulates a clean bound on the tail probability, while the second inequality provides a direct bound on the level of $\|\boldsymbol{S}_k\|$. Let us emphasize that both estimates provide uniform control on the *entire trajectory* of the matrix martingale.

4.3 The log-mgf supermartingale. As in the proof of Theorem 2.1, we can and will employ the Hermitian dilation (2.5) to restrict our attention to Hermitian random matrices.

Consider a bounded Hermitian matrix martingale $(\boldsymbol{S}_0, \dots, \boldsymbol{S}_n)$ taking values in $\mathbb{H}_d(\mathbb{C})$, with initial condition $\boldsymbol{S}_0 := \mathbf{0}$ and with difference sequence $(\boldsymbol{X}_1, \dots, \boldsymbol{X}_n)$. For a parameter $\theta \in \mathbb{R}$, define the *conditional* log-mgfs and their partial sums:

$$\boldsymbol{\Xi}_{\boldsymbol{X}_k \,|\, \mathcal{F}_{k-1}}(\theta) := \log \mathbb{E}\left[\mathrm{e}^{\theta \boldsymbol{X}_k} \,|\, \mathcal{F}_{k-1} \right] \quad \text{and} \quad \boldsymbol{W}_k(\theta) := \sum_{j=1}^{k} \boldsymbol{\Xi}_{\boldsymbol{X}_j \,|\, \mathcal{F}_{j-1}}(\theta). \tag{4.3}$$

The element $\boldsymbol{W}_k(\theta) \in \mathbb{H}_d(\mathbb{C})$ of the log-mgf process reflects the total volatility of the martingale up to time k.

To track the evolution of the matrix martingale, let us pass to a real-valued random sequence:

$$M_0(\theta) := d \quad \text{and} \quad M_k(\theta) := \operatorname{Tr} \mathrm{e}^{\theta \boldsymbol{S}_k - \boldsymbol{W}_k(\theta)} \quad \text{where } k = 1, \dots, n. \tag{4.4}$$

Applying the subadditivity rule (2.8) for the matrix log-mgf conditionally, we quickly verify that $\mathbb{E}[M_k(\theta) \,|\, \mathcal{F}_{k-1}] \le M_{k-1}(\theta)$ for each $k = 1, \dots, n$. In other words, $(M_k(\theta) : k = 0, \dots, n)$ composes a positive supermartingale. We can study the trajectory of the matrix martingale by bounding the supermartingale [37, 46]. The approach here is adapted from Howard et al. [23].

PROPOSITION 4.5 (Matrix martingale: Eigenvalue bounds). *Consider a finite matrix martingale* $(\boldsymbol{S}_0, \dots, \boldsymbol{S}_n)$ *taking values in* $\mathbb{H}_d(\mathbb{C})$ *with log-mgf process* $(\boldsymbol{W}_1(\theta), \dots, \boldsymbol{W}_n(\theta))$, *as in* (4.3). *For parameters* $\theta, \alpha > 0$,

$$\mathbb{P}\left\{ \exists k : \lambda_{\max}(\boldsymbol{S}_k) \ge \theta^{-1}\big(\alpha + \lambda_{\max}(\boldsymbol{W}_k(\theta))\big) \right\} \le d \cdot \mathrm{e}^{-\alpha}.$$

Proof. The log-mgf supermartingale (4.4) admits a lower bound involving the eigenvalues of the sequences:

$$\log M_k(\theta) = \log \operatorname{Tr} \mathrm{e}^{\theta \boldsymbol{S}_k - \boldsymbol{W}_k(\theta)} \ge \theta \lambda_{\max}(\boldsymbol{S}_k) - \lambda_{\max}(\boldsymbol{W}_k(\theta)) \quad \text{for each } \theta > 0.$$

For the inequality, note that the trace dominates the maximum eigenvalue; then apply spectral mapping and Weyl's perturbation theorem [9, Cor. III.2.6]. Ville's inequality for positive supermartingales [23, Lem. 1] implies

$$\mathbb{P}\left\{ \exists k : \theta \lambda_{\max}(\boldsymbol{S}_k) - \lambda_{\max}(\boldsymbol{W}_k(\theta)) \ge \alpha \right\} \le \mathbb{P}\left\{ \exists k : \log M_k(\theta) \ge \alpha \right\} \le \mathrm{e}^{-\alpha} \cdot \mathbb{E}[M_0(\theta)] = d \cdot \mathrm{e}^{-\alpha}.$$

This is the advertised result. ∎

4.4 Proof of matrix Freedman. To invoke Proposition 4.5, we seek a bound for the maximum eigenvalue of the log-mgf process $(\boldsymbol{W}_k(\theta) : k = 0, \dots, n)$. In the setting of Theorem 4.4, each element of the difference sequence admits a bound on the conditional log-mgf of the Bernstein type (2.9):

$$\boldsymbol{\Xi}_{\boldsymbol{X}_j \,|\, \mathcal{F}_{j-1}}(\theta) \preccurlyeq g(\theta) \cdot \mathbb{E}\left[\boldsymbol{X}_j^2 \,|\, \mathcal{F}_{j-1}\right] \quad \text{where} \quad g(\theta) := \frac{\theta^2/2}{1 - B|\theta|/3} \quad \text{and} \quad |\theta| < 3/B.$$

By Weyl's monotonicity theorem [9, Cor. III.2.3],

$$\lambda_{\max}(\boldsymbol{W}_k(\theta)) = \lambda_{\max}\left(\sum_{j=1}^{k} \boldsymbol{\Xi}_{\boldsymbol{X}_j \,|\, \mathcal{F}_{j-1}}(\theta)\right) \le g(\theta) \cdot \lambda_{\max}\left(\sum_{j=1}^{k} \mathbb{E}\left[\boldsymbol{X}_j^2 \,|\, \mathcal{F}_{j-1}\right]\right) = g(\theta) \cdot V_k.$$

Proposition 4.5 yields

$$\begin{aligned} d \cdot \mathrm{e}^{-\alpha} &\ge \mathbb{P}\left\{\exists k : \lambda_{\max}(\boldsymbol{S}_k) \ge \theta^{-1}\big(\alpha + \lambda_{\max}(\boldsymbol{W}_k(\theta))\big)\right\} \ge \mathbb{P}\left\{\exists k : \lambda_{\max}(\boldsymbol{S}_k) \ge \theta^{-1}\big(\alpha + g(\theta)V_k\big)\right\} \\ &\ge \mathbb{P}\left\{\exists k : \lambda_{\max}(\boldsymbol{S}_k) \ge \theta^{-1}\big(\alpha + g(\theta)v\big) \text{ and } V_k \le v\right\}. \end{aligned}$$

In the Hermitian setting, we obtain an analogue of the first matrix Freedman tail bound (4.1) by selecting $\alpha = t$ and $\theta = \sqrt{t}/(\sqrt{v/2} + B\sqrt{t}/3)$. The second tail bound (4.2) follows from the first tail bound by inverting the function of t and making a simple bound; see [10, sects. 2.4 and 2.8]. Finally, we extend to a general (non-Hermitian) matrix martingale by applying these two bounds to the Hermitian dilation (2.5).

5 Matrix martingales: Applications. In this section, we outline some applications of the matrix Freedman inequality in computational mathematics. This technique is particularly valuable for handling sequences of matrices that evolve adaptively, and it provides uniform control over the entire trajectory of the process.

5.1 Uniform covariance estimation. Consider a centered random vector $\boldsymbol{y} \in \mathbb{R}^d$ that is uniformly bounded: $\mathbb{E}[\boldsymbol{y}] = \boldsymbol{0}$ and $\|\boldsymbol{y}\| \le L$. Its covariance matrix takes the form $\boldsymbol{\Sigma} := \mathbb{E}[\boldsymbol{y}\boldsymbol{y}^*]$. Given i.i.d. samples $\boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3, \dots$ from the distribution, we can construct a sequence of empirical covariance estimates:

$$\widehat{\boldsymbol{\Sigma}}_n := \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{y}_k \boldsymbol{y}_k^* \quad \text{for each } n = 1, 2, 3, \dots. \tag{5.1}$$

We would like to obtain a confidence region for the entire sequence of covariance estimates, ensuring simultaneous coverage at all times. The following statement is adapted from Howard et al. [24, sect. 4.3].

THEOREM 5.1 (Uniform covariance estimation [24, sect. 4.3]). *Let $\boldsymbol{y} \in \mathbb{R}^d$ be a bounded, centered random vector with $\|\boldsymbol{y}\| \le L$ and covariance matrix $\boldsymbol{\Sigma}$. As in* (5.1)*, construct the sequence $(\widehat{\boldsymbol{\Sigma}}_n : n \in \mathbb{N})$ of empirical covariance estimates. The following uniform confidence region attains confidence level $1 - \delta \in (0, 1)$:*

$$\mathbb{P}\left\{\forall n \in \mathbb{N} : \frac{\|\widehat{\boldsymbol{\Sigma}}_n - \boldsymbol{\Sigma}\|}{\|\boldsymbol{\Sigma}\|} \le \sqrt{8\beta_n} + \frac{2}{3}\beta_n\right\} \ge 1 - \delta \quad \textit{where} \quad \beta_n := \frac{L^2}{\|\boldsymbol{\Sigma}\|} \cdot \frac{\log(2d/\delta) + \log\log(\mathrm{e}n)}{n}.$$

The ratio $r := L^2/\|\boldsymbol{\Sigma}\|$ reflects the number of dimensions where the random vector has substantial fluctuation. We need $n = \mathcal{O}(r \log d)$ samples to reliably estimate the covariance in these directions. As the number n of samples continues to increase, the size of the confidence region decreases at a rate $\sqrt{n^{-1} \log\log n}$. The log-log factor is required to correct for occasional extreme fluctuations. More precise asymptotic results follow from the law of the iterated logarithm in a Banach space [31, Thm. 8.2].

Proof sketch. For a parameter $N \in \mathbb{N}$, introduce the finite martingale sequence

$$\boldsymbol{S}_0 := \boldsymbol{0} \quad \text{and} \quad \boldsymbol{S}_n := \sum_{k=1}^{n} (\boldsymbol{y}_k \boldsymbol{y}_k^* - \boldsymbol{\Sigma}) =: \sum_{k=1}^{n} \boldsymbol{X}_k \quad \text{for } n = 1, \dots, N.$$

The difference sequence $(\boldsymbol{X}_k)$ consists of i.i.d. centered random matrices with dimension d. A short calculation gives

$$B = \sup \|\boldsymbol{X}_k\| \le L^2 \quad \text{and} \quad V_n = \left\|\sum_{k=1}^{n} \mathbb{E}[\boldsymbol{X}_k^2]\right\| \le nL^2 \cdot \|\boldsymbol{\Sigma}\|.$$

The matrix Freedman inequality (4.1) ensures that

$$\mathbb{P}\left\{\exists n \le N : \|\boldsymbol{S}_n\| \ge \sqrt{2nL^2\|\boldsymbol{\Sigma}\|t} + L^2t/3\right\} \le 2d \cdot \mathrm{e}^{-t}.$$

Dividing each element $\boldsymbol{S}_n$ by the factor $n \cdot \|\boldsymbol{\Sigma}\|$ and limiting the index n to the range $[N/2, N]$, we find

$$\mathbb{P}\left\{\exists n \in [N/2, N] : \frac{\|\widehat{\boldsymbol{\Sigma}}_n - \boldsymbol{\Sigma}\|}{\|\boldsymbol{\Sigma}\|} \ge \sqrt{\frac{4L^2t}{n\|\boldsymbol{\Sigma}\|}} + \frac{L^2t}{3n\|\boldsymbol{\Sigma}\|}\right\} \le 2d \cdot \mathrm{e}^{-t}.$$

The key insight is to apply the last inequality on dyadic intervals: $N = 2^j$ for $j = 1, 2, 3, \dots$ at the level $t_j = \log(4j^2 d/\delta) \le 2\log(2d\log(en)/\delta)$. Combine the results with a union bound, then take the complement. ∎

5.2 Cholesky decomposition with stochastic rounding. Most numerical algorithms proceed in stages, and the numerical errors compound at each step of the process. Traditional analyses model these errors deterministically, and error bounds increase linearly with the number of steps in the algorithm. More recently, researchers have started to make a close accounting of error propagation under probabilistic models, where the errors accumulate more slowly because of cancellations. Martingale methods offer an elegant technique for tracking random errors that depend on the past history of the algorithm. Connolly and Higham [15] have employed this approach to give a probabilistic analysis of the Householder QR algorithm. In the same spirit, this section offers a stylized analysis of the Cholesky factorization algorithm under a stochastic rounding model (cf. subsection 3.2).

Let $\boldsymbol{A} \in \mathbb{H}_d^{++}(\mathbb{R})$ be a positive-definite matrix. For clarity of interpretation, we assume that $\operatorname{diag}(\boldsymbol{A}) = \mathbf{I}$. Cholesky's method iteratively produces a factorization $\boldsymbol{A} = \boldsymbol{C}\boldsymbol{C}^*$ where $\boldsymbol{C} \in \mathbb{M}_d(\mathbb{R})$ is lower triangular. Define the initial residual $\boldsymbol{R}_0 := \boldsymbol{A}$ and the initial triangular factor $\boldsymbol{C}_0 := \mathbf{0}$. At the kth step, reduce the residual $\boldsymbol{R}_{k-1}$ by performing a Schur complement with respect to its kth column:

$$\boldsymbol{c}_k := \frac{\boldsymbol{R}_{k-1}(:,k)}{\sqrt{\boldsymbol{R}_{k-1}(k,k)}} \quad \text{and} \quad \boldsymbol{C}_k := \boldsymbol{C}_{k-1} + \boldsymbol{c}_k\mathbf{e}_k^* \quad \text{and} \quad \boldsymbol{R}_k := \boldsymbol{R}_{k-1} - \boldsymbol{c}_k\boldsymbol{c}_k^*.$$

This procedure zeroes out the kth row and column from $\boldsymbol{R}_{k-1}$, while placing $\boldsymbol{c}_k$ in the kth column of $\boldsymbol{C}_k$. The matrix $\boldsymbol{C}_k$ remains lower triangular. At each step, the residual and the partial triangular factorization compose the original matrix: $\boldsymbol{A} = \boldsymbol{R}_k + \boldsymbol{C}_k\boldsymbol{C}_k^*$. After d steps, $\boldsymbol{R}_d = \mathbf{0}$ and $\boldsymbol{C}_d\boldsymbol{C}_d^* = \boldsymbol{A}$. Each step requires $\mathcal{O}(d^2)$ arithmetic operations, for a total cost of $\mathcal{O}(d^3)$.

In actual practice, we commit floating-point errors when we reduce the residual—but we can extract a column from the residual without further error. If we employ stochastic rounding, then the residual update becomes

$$\boldsymbol{R}_k := \mathtt{sr}(\boldsymbol{R}_{k-1} - \boldsymbol{c}_k\boldsymbol{c}_k^*) = \boldsymbol{R}_{k-1} - \boldsymbol{c}_k\boldsymbol{c}_k^* + \boldsymbol{Y}_k.$$

The stochastic error $\boldsymbol{Y}_k$ is a random symmetric matrix that depends on the previous residual. The sigma-algebra $\mathcal{F}_{k-1} := \sigma(\boldsymbol{R}_0, \dots, \boldsymbol{R}_{k-1})$ captures the history of the algorithm through step $k-1$. As in subsection 3.2, each entry of $\boldsymbol{Y}_k$ satisfies the rounding properties in (3.4), conditional on $\mathcal{F}_{k-1}$, and each entry is rounded independently of the others (modulo symmetry). We frame the assumptions that

$$\mathbb{E}[\boldsymbol{Y}_k \,|\, \mathcal{F}_{k-1}] = \mathbf{0} \quad \text{and} \quad \big\|\mathbb{E}[\boldsymbol{Y}_k^2 \,|\, \mathcal{F}_{k-1}]\big\| \le 2\mathtt{u}^2 \cdot \|\boldsymbol{A}\| \quad \text{and} \quad \|\boldsymbol{Y}_k\|_{\max} \le \mathtt{u}. \tag{5.2}$$

The third property follows because $\operatorname{diag}(\boldsymbol{A}) = \mathbf{I}$ and $\operatorname{diag}(\boldsymbol{R}_{k-1})$ only decreases at each step. The second property holds if the computed residual $\boldsymbol{R}_{k-1}$ remains psd and satisfies $\|\boldsymbol{R}_{k-1}\| \le 2\|\boldsymbol{A}\|$. Under mild assumptions, these conditions can be justified with some additional effort. We also demand that no underflow or overflow occurs.

The computed residuals $\boldsymbol{R}_k$ and the partial triangular factorizations $\boldsymbol{C}_k\boldsymbol{C}_k^*$ induce a sequence of approximations of the original matrix:

$$\boldsymbol{S}_k := \boldsymbol{R}_k + \boldsymbol{C}_k\boldsymbol{C}_k^* = \boldsymbol{R}_k + \sum_{j=1}^{k} \boldsymbol{c}_j\boldsymbol{c}_j^* \quad \text{for } k = 0, 1, 2, \dots, d.$$

Using these approximations, we can calculate the error in the completed triangular factorization $\boldsymbol{C} := \boldsymbol{C}_d$.

$$\boldsymbol{C}\boldsymbol{C}^* - \boldsymbol{A} = \boldsymbol{S}_d - \boldsymbol{S}_0 = \sum_{k=1}^{d}(\boldsymbol{S}_k - \boldsymbol{S}_{k-1}) = \sum_{k=1}^{d}(\boldsymbol{R}_k - \boldsymbol{R}_{k-1} + \boldsymbol{c}_k\boldsymbol{c}_k^*) = \sum_{k=1}^{d}\boldsymbol{Y}_k.$$

In other words, the sequence of approximations composes a matrix martingale with difference sequence $(\boldsymbol{Y}_k : k = 1, \dots, d)$. We can use this martingale to assess the accumulated rounding error in the Cholesky decomposition.

THEOREM 5.2 (Cholesky decomposition with stochastic rounding). *Let $\boldsymbol{A} \in \mathbb{H}_d^{++}(\mathbb{R})$ be a $d \times d$ positive-definite matrix with $\operatorname{diag}(\boldsymbol{A}) = \mathbf{I}$. As described above, compute the Cholesky decomposition $\boldsymbol{C}\boldsymbol{C}^*$ with a stochastic rounding procedure that satisfies the assumption* (5.2). *The resulting factorization admits the error bound*

$$\mathbb{P}\left\{ \|\boldsymbol{C}\boldsymbol{C}^* - \boldsymbol{A}\| \geq \mathtt{u} \cdot \left(2\sqrt{d\,\|\boldsymbol{A}\|\, t} + \frac{1}{3} t \right) \right\} \leq 2d \cdot \mathrm{e}^{-t}.$$

For the Cholesky decomposition, the classic deterministic rounding error analysis [21, Chap. 10] gives a bound for the entrywise error $\|\boldsymbol{C}\boldsymbol{C}^* - \boldsymbol{A}\|_{\max}$ that is proportional to $\mathtt{u} \cdot d$. For stochastic rounding, the stylized analysis here indicates that the entrywise error is typically on the order of $\mathtt{u} \cdot \sqrt{d\,\|\boldsymbol{A}\|}$. The factor $\sqrt{d}$ improvement can be significant. A full comparison with deterministic rounding falls outside our ambit.

Proof sketch. We need to consider a more fine-grained martingale to exploit the fact that matrix entries are rounded independently. Decompose each error matrix as $\boldsymbol{Y}_k = \sum_{i \leq j} \boldsymbol{X}_{ijk}$, where $\boldsymbol{X}_{ijk}$ describes the rounding error in the (i,j) and (j,i) entries of $\boldsymbol{Y}_k$. For each k, the family $(\boldsymbol{X}_{ijk} : i \leq j)$ is conditionally independent, given $\mathcal{F}_{k-1}$. According to (5.2), the parameters in the matrix Freedman inequality (Theorem 4.4) satisfy

$$\begin{aligned} B &= \sup \|\boldsymbol{X}_{ijk}\| = \sup \|\boldsymbol{Y}_k\|_{\max} \leq \mathtt{u}, \\ V_d &= \left\| \sum_{k=1}^{d} \sum_{i \leq j} \mathbb{E}\left[\boldsymbol{X}_{ijk}^2 \,|\, \mathcal{F}_{k-1} \right] \right\| = \left\| \sum_{k=1}^{d} \mathbb{E}\left[\boldsymbol{Y}_k^2 \,|\, \mathcal{F}_{k-1} \right] \right\| \leq 2\mathtt{u}^2 d \cdot \|\boldsymbol{A}\|. \end{aligned}$$

The result follows from the matrix Freedman inequality (4.1). ∎

5.3 Fast Laplacian solvers. As in subsection 3.3, consider a weighted, connected, undirected graph on n vertices with Laplacian matrix $\boldsymbol{L} \in \mathbb{H}_n^+(\mathbb{R})$. A linear system in the Laplacian matrix takes the form

$$\boldsymbol{L}\boldsymbol{u} = \boldsymbol{f} \quad \text{where} \quad \mathbf{1}^*\boldsymbol{u} = \mathbf{1}^*\boldsymbol{f} = 0. \tag{5.3}$$

This type of linear system arises in a dizzying range of applications, including numerical discretizations of elliptic PDEs, clustering and partitioning of data, and network flow problems [44]. The basic algorithm for (5.3) starts by computing a Cholesky decomposition $\boldsymbol{L} = \boldsymbol{C}\boldsymbol{C}^*$, and it solves the linear system by two steps of triangular substitution. In general, this approach requires $\mathcal{O}(n^3)$ operations, which is often prohibitive. Can we do better?

Spielman and Teng [42] gave an affirmative answer by devising a theoretical Laplacian solver that runs in time linear in the number m of edges in the graph and polylogarithmic in the number n of vertices. After a train of refinements, Kyng and Sachdeva [29] designed a *practical* Laplacian solver that runs in time $\mathcal{O}(m \log^3 n)$. Their algorithm, `SparseCholesky`, performs an approximate Cholesky decomposition by randomly sparsifying the Schur complement at each step. The analysis relies on matrix martingales. See [51] for another account.

Here is the intuition. When we apply the Cholesky method to a graph Laplacian, each Schur complement step eliminates a vertex from the graph by adding a clique to the remaining vertices of the graph. The method is expensive because the clique involves up to $n - k$ vertices and up to $(n-k)^2$ edges at step k. Yet, as we saw in subsection 3.3, we can dramatically sparsify a graph by randomly sampling edges in proportion to their effective resistances. In this spirit, `SparseCholesky` starts with residual $\boldsymbol{R}_0 := \boldsymbol{L}$ and triangular factor $\boldsymbol{C}_0 := \mathbf{0}$. It performs the steps

$$\boldsymbol{c}_k := \frac{\boldsymbol{R}_{k-1}(:,k)}{\sqrt{\boldsymbol{R}_{k-1}(k,k)}} \quad \text{and} \quad \boldsymbol{C}_k := \boldsymbol{C}_{k-1} + \boldsymbol{c}_k \mathbf{e}_k^* \quad \text{and} \quad \boldsymbol{R}_k := \boldsymbol{R}_{k-1} - \mathtt{Sparsify}(\boldsymbol{c}_k \boldsymbol{c}_k^*).$$

The `Sparsify` method produces a random, unbiased approximation to the matrix $\boldsymbol{c}_k \boldsymbol{c}_k^*$, which contains the Laplacian of the clique. After sparsification, the remaining number of edges equals the number of neighbors of vertex k in the residual graph $\boldsymbol{R}_{k-1}$. The full algorithm requires several other innovations: it splits edges into smaller pieces to reduce their weights; it maintains estimates for the effective resistances to implement the sparsification step; and it eliminates vertices in a random order to limit the average number of neighbors. As in subsection 5.2, this procedure results in a matrix martingale, namely $\boldsymbol{S}_k := \boldsymbol{R}_k + \boldsymbol{C}_k \boldsymbol{C}_k^*$ for $k = 0, 1, 2, \ldots d$. We can analyze this martingale using the matrix Freedman inequality (Theorem 4.4). Here is the outcome.

THEOREM 5.3 (Sparse Cholesky approximation [29]). *Let $\boldsymbol{L} \in \mathbb{H}_d^+(\mathbb{R})$ be a weighted graph Laplacian, as in* (3.5)*. With probability at least $1 - d^{-1}$, the* `SparseCholesky` *algorithm produces an approximate Cholesky factorization $\boldsymbol{C}\boldsymbol{C}^*$ with the spectral guarantee*

$$0.5\boldsymbol{L} \preccurlyeq \boldsymbol{C}\boldsymbol{C}^* \preccurlyeq 1.5\boldsymbol{L}.$$

The lower-triangular matrix $\boldsymbol{C}$ has $\mathcal{O}(m \log^2 n)$ nonzero entries. The expected runtime is $\mathcal{O}(m \log^3 n)$ operations.

Once we have computed the approximate Cholesky decomposition $\boldsymbol{L} \approx \boldsymbol{C}\boldsymbol{C}^*$, we can solve the linear system (5.3) using preconditioned conjugate gradient (PCG) with the triangular preconditioner $\boldsymbol{C}^{-1}$.

PROPOSITION 5.4 (Fast Laplacian solvers [29]). *Given the triangular matrix $\boldsymbol{C}$ computed by the* `SparseCholesky` *algorithm, the PCG algorithm solves the consistent Laplacian system* (5.3) *to relative error ε in the Dirichlet energy norm after $\mathcal{O}(m \log^2(n) \log(1/\varepsilon))$ arithmetic operations.*

For a sparse graph with $m = \mathcal{O}(n)$ edges, the total cost of solving the Laplacian system (5.3) to fixed accuracy is just $\mathcal{O}(n \log^3 n)$ operations. For a dense graph with $m = \mathcal{O}(n^2)$ edges, the total cost is $\mathcal{O}(n^2 \log^3 n)$ operations. Both results compare favorably with the $\mathcal{O}(n^3)$ cost of the classic method based on a full Cholesky factorization.

5.4 Randomized Trotter formulas. The evolution of a quantum mechanical system is governed by a Hamiltonian matrix $\boldsymbol{H} \in \mathbb{H}_d(\mathbb{C})$. The state of the system at time $t \geq 0$ takes the form

$$\boldsymbol{\Phi}_t(\boldsymbol{\rho}) := \mathrm{e}^{-\mathrm{i}t\boldsymbol{H}} \cdot \boldsymbol{\rho} \cdot \mathrm{e}^{+\mathrm{i}t\boldsymbol{H}} \quad \text{where the initial state } \boldsymbol{\rho} \in \mathcal{D}_n.$$

A basic goal of quantum information science is to simulate the action $\boldsymbol{\Phi}_t$ induced by a complicated Hamiltonian through the application of simpler Hamiltonians. Randomized methods can provide effective algorithms for this task. This section summarizes a result of Chen et al. [14] that builds on work of Campbell [13].

Suppose that the Hamiltonian can be expressed as a sum of many "simple" Hermitian terms:

$$\boldsymbol{H} := \sum_{m=1}^{M} \boldsymbol{H}_m \quad \text{with} \quad L := \sum_{m=1}^{M} \|\boldsymbol{H}_m\| \quad \text{and each } \boldsymbol{H}_m \in \mathbb{H}_d(\mathbb{C}). \tag{5.4}$$

The statistic L measures the interaction strength within the Hamiltonian. This model encompasses applications in quantum chemistry and quantum physics. For clarity, fix the time horizon for the simulation at $t = 1$. Given a parameter $n \in \mathbb{N}$, construct a random unitary matrix by drawing a random term from the Hamiltonian according to an importance sampling distribution:

$$\boldsymbol{Y} := \mathrm{e}^{-\mathrm{i}\boldsymbol{Z}/n} \quad \text{where} \quad \mathbb{P}\left\{ \boldsymbol{Z} = \frac{L}{\|\boldsymbol{H}_j\|} \boldsymbol{H}_j \right\} = \frac{\|\boldsymbol{H}_j\|}{L} \quad \text{for } j = 1, \dots, M. \tag{5.5}$$

To approximate the unitary matrix $\boldsymbol{U} := \mathrm{e}^{-\mathrm{i}\boldsymbol{H}}$, form a product of n independent copies of $\boldsymbol{Y}$:

$$\boldsymbol{Q} := \boldsymbol{Y}_n \cdots \boldsymbol{Y}_3 \boldsymbol{Y}_2 \boldsymbol{Y}_1 \quad \text{where } \boldsymbol{Y}_k \sim \boldsymbol{Y} \text{ i.i.d. for } k = 1, \dots, n. \tag{5.6}$$

In other words, we simulate the action of the full Hamiltonian $\boldsymbol{H}$ by performing a series of short simulations with the simpler Hamiltonians $\boldsymbol{H}_m$. This is a randomized variant of the Lie–Trotter–Suzuki product formulas that are widely used to approximate the matrix exponential of a sum. We can analyze the quality of the simulation using matrix martingale methods.

THEOREM 5.5 (Random product formulas [14, Thm. 1]). *Consider a Hamiltonian of the form* (5.4) *with interaction strength L, and form the unitary matrix $\boldsymbol{U} := \mathrm{e}^{-\mathrm{i}\boldsymbol{H}}$. Fix parameters $\varepsilon, \delta \in (0, 1)$ where $\varepsilon \leq L$. For $n \geq 40\varepsilon^{-2} L^2 \log(2d/\delta)$, construct a random unitary matrix $\boldsymbol{Q}$ with n factors via the random product formula* (5.6)*. Then*

$$\mathbb{P}\left\{ \sup\nolimits_{\boldsymbol{\rho} \in \mathcal{D}_n} \|\boldsymbol{Q} \cdot \boldsymbol{\rho} \cdot \boldsymbol{Q}^* - \boldsymbol{U} \cdot \boldsymbol{\rho} \cdot \boldsymbol{U}^*\|_1 \geq 2\varepsilon \right\} \leq \delta.$$

This result states that, with high probability, the random product formula (5.6) approximates the evolution of the quantum Hamiltonian for *all* initial states. The number n of factors in the product does not depend on the number M of summands in the Hamiltonian (5.4), but it grows quadratically with the interaction strength L. For comparison, the number of factors in a deterministic product formula depends linearly on the number M of summands in the Hamiltonian, but more weakly on the interaction strength L. As a consequence, random product formulas are most effective for short-time simulations of Hamiltonians with many summands.

Proof sketch. The quantity of interest depends on the spectral-norm distance between the two unitaries:

$$E := \sup_{\boldsymbol{\rho}\in\mathcal{D}_n} \|\boldsymbol{Q}\cdot\boldsymbol{\rho}\cdot\boldsymbol{Q}^* - \boldsymbol{U}\cdot\boldsymbol{\rho}\cdot\boldsymbol{U}^*\|_1 \le 2\,\|\boldsymbol{Q}-\boldsymbol{U}\|.$$

This point follows from the triangle inequality, unitary invariance, and the operator ideal property of the trace norm. The right-hand side consists of a random fluctuation term and a deterministic bias:

$$\|\boldsymbol{Q}-\boldsymbol{U}\| \le \|\boldsymbol{Q}-\mathbb{E}[\boldsymbol{Q}]\| + \|\mathbb{E}[\boldsymbol{Q}]-\boldsymbol{U}\| =: \text{flux} + \text{bias}.$$

By statistical independence of the factors in (5.6), the expectation $\mathbb{E}[\boldsymbol{Q}] = (\mathbb{E}[\boldsymbol{Y}])^n$.

The bias term admits a bound that depends on the interaction strength L of the Hamiltonian:

$$\text{bias} = \|(\mathbb{E}[\boldsymbol{Y}])^n - (\mathrm{e}^{-\mathrm{i}\boldsymbol{H}/n})^n\| \le n\cdot\|\mathbb{E}[\boldsymbol{Y}] - \mathrm{e}^{-\mathrm{i}\boldsymbol{H}/n}\| \le n\cdot\left(\frac{L}{n}\right)^2 = \frac{L^2}{n}.$$

The first inequality follows from repeated application of the triangle inequality and unitary invariance, while the second inequality requires a second-order approximation of the matrix exponential. See [14, Lems. 3.5 and 3.6] for the details. We insist that $L^2/n \le \varepsilon/2$ to control the bias term.

For the fluctuation term, we introduce a Lévy–Doob martingale. Let $\mathcal{F}_k := \sigma(\boldsymbol{Y}_1,\dots,\boldsymbol{Y}_k)$, and define

$$\boldsymbol{S}_k := \mathbb{E}[\boldsymbol{Q}\,|\,\mathcal{F}_k] = (\mathbb{E}\,\boldsymbol{Y})^{n-k}\boldsymbol{Y}_k\dots\boldsymbol{Y}_1 \quad \text{for } k=0,\dots,n.$$

Observe that $\boldsymbol{S}_n = \boldsymbol{Q}$ and $\boldsymbol{S}_0 = \mathbb{E}[\boldsymbol{Q}]$. Construct the difference sequence $\boldsymbol{X}_k := \boldsymbol{S}_k - \boldsymbol{S}_{k-1}$ for $k=1,\dots,n$. By first-order approximation of the matrix exponential, the difference sequence satisfies uniform bounds

$$\|\boldsymbol{X}_k\| \le \|\boldsymbol{Y}_k - \mathbb{E}[\boldsymbol{Y}_k]\| \le \frac{2L}{n} \quad\text{and}\quad \left\|\mathbb{E}\left[\boldsymbol{X}_k\boldsymbol{X}_k^*\,|\,\mathcal{F}_{k-1}\right]\right\| \vee \left\|\mathbb{E}\left[\boldsymbol{X}_k^*\boldsymbol{X}_k\,|\,\mathcal{F}_{k-1}\right]\right\| \le \|\boldsymbol{Y}_k - \mathbb{E}[\boldsymbol{Y}_k]\|^2 \le \frac{4L^2}{n^2}.$$

See [14, Prop. 3.3] for the details. The matrix Freedman inequality (4.2) now implies a probability inequality for $\text{flux} = \|\boldsymbol{S}_n - \boldsymbol{S}_0\|$ at level $t = \varepsilon/2$. Combine the bias and fluctuation to get a tail bound for the quantity E. ∎

6 Matrix concentration: Extensions. Over the last few years, researchers have made significant advances in understanding the behavior of the independent sum model. These results require a detour into the study of Gaussian random matrices, which serve as ideal models for independent sums. First, we outline some improvements of matrix concentration for Gaussian random matrices, and then we explain how Gaussian models can capture the spectral features of an independent sum of random matrices. Last, we mention several other random matrix models that remain under active study.

6.1 Variance statistics. We can describe the behavior of random matrix models more fully by employing a broader family of variance statistics. Introduce the real inner product

$$\langle \boldsymbol{B},\ \boldsymbol{A}\rangle_{\mathrm{Re}} := \mathrm{Re}\,\mathrm{Tr}[\boldsymbol{B}^*\boldsymbol{A}] \quad \text{for } \boldsymbol{A},\boldsymbol{B}\in\mathbb{F}^{d_1\times d_2}.$$

The variance function of a random matrix $\boldsymbol{S}\in\mathbb{F}^{d_1\times d_2}$ is defined as

$$\mathsf{Var}[\boldsymbol{S}] : \mathbb{F}^{d_1\times d_2} \to \mathbb{R}_+ \quad\text{where}\quad \mathsf{Var}[\boldsymbol{S}](\boldsymbol{A}) := \mathrm{Var}[\langle \boldsymbol{S},\ \boldsymbol{A}\rangle_{\mathrm{Re}}] \quad \text{for } \boldsymbol{A}\in\mathbb{F}^{d_1\times d_2}. \tag{6.1}$$

Variance functions are in correspondence with psd quadratic forms on $\mathbb{F}^{d_1\times d_2}$, treated as a real linear space.

The variance function $\mathsf{Var}[\boldsymbol{S}]$ packs up all the second-order statistics of the random matrix. In particular, it completely determines the matrix variance $v(\boldsymbol{S})$, defined in (2.1). We can also introduce the *weak variance,* which is related to the tail behavior of the sum: $v_*(\boldsymbol{S}) := \sup_{\|\boldsymbol{A}\|_1\le 1}\mathsf{Var}[\boldsymbol{S}](\boldsymbol{A})$. The weak variance $v_*(\boldsymbol{S})$ is always smaller than the matrix variance $v(\boldsymbol{S})$, and often much smaller.

6.2 Gaussian matrix models. A random matrix is *Gaussian* when each of its linear marginals is Gaussian. More precisely, $\boldsymbol{Z}\in\mathbb{F}^{d_1\times d_2}$ is Gaussian if and only if the random variable $\langle \boldsymbol{Z},\ \boldsymbol{A}\rangle_{\mathrm{Re}}$ follows a real Gaussian distribution for each matrix $\boldsymbol{A}\in\mathbb{F}^{d_1\times d_2}$. A Gaussian random matrix is fully characterized by its expectation $\mathbb{E}[\boldsymbol{Z}]$ and by its variance function $\mathsf{Var}[\boldsymbol{Z}]$. For a specified expectation matrix $\boldsymbol{M}\in\mathbb{F}^{d_1\times d_2}$ and a variance function $\mathsf{V} : \mathbb{F}^{d_1\times d_2}\to\mathbb{R}_+$, we denote the (unique) Gaussian distribution with these statistics by $\textsc{normal}(\boldsymbol{M},\mathsf{V})$.

6.3 Second-order matrix Khinchin inequalities. The norm of a Gaussian matrix $\boldsymbol{Z}$ satisfies an elegant matrix concentration bound, called the *matrix Khinchin inequality,* originally due to Lust-Piquard [34]:

$$\sqrt{(2/\pi)\, v(\boldsymbol{Z})} \;\le\; \mathbb{E}\,\|\boldsymbol{Z} - \mathbb{E}[\boldsymbol{Z}]\| \;\le\; \sqrt{2v(\boldsymbol{Z})\log(d_1+d_2)}. \tag{6.2}$$

The lower bound involves Gaussian isoperimetry [30, Cor. 3]; the upper bound relies on exponential matrix concentration arguments [48, Thm. 4.1.1]. Both inequalities are saturated.

To improve on the matrix Khinchin inequality (6.2), we need extra information about the variance function of the Gaussian matrix. Define the *interaction energy* to be the maximum eigenvalue of the quadratic form:

$$w(\boldsymbol{Z}) := 2\sup\nolimits_{\|\boldsymbol{A}\|_{\mathrm{F}}\le 1} \mathsf{Var}[\boldsymbol{Z}](\boldsymbol{A}).$$

The interaction energy $w(\boldsymbol{Z})$ can be much smaller than the matrix variance $v(\boldsymbol{Z})$, but they are incomparable. Heuristically, the interaction energy is small when $\boldsymbol{Z}$ is "highly noncommutative." This statistic supports a second-order refinement of the matrix Khinchin inequality [5, Cor. 2.2 and Lem. 2.5]:

$$\mathbb{E}\,\|\boldsymbol{Z} - \mathbb{E}[\boldsymbol{Z}]\| \le 2\sqrt{v(\boldsymbol{Z})} + \mathrm{Const}\cdot\big[v(\boldsymbol{Z})w(\boldsymbol{Z})\log^3(d_1+d_2)\big]^{1/4}. \tag{6.3}$$

When $w(\boldsymbol{Z}) \ll v(\boldsymbol{Z})$, then the result (6.3) improves over (6.2). As a simple example, consider a real Ginibre matrix $\boldsymbol{G} \in \mathbb{M}_d(\mathbb{R})$, which is a $d\times d$ matrix with i.i.d. standardized Gaussian entries. The matrix variance $v(\boldsymbol{G}) = d$, while the interaction energy $w(\boldsymbol{G}) = 2$. Thus,

$$\begin{aligned} &(6.2) \quad \text{implies} \quad \mathbb{E}\,\|\boldsymbol{G}\| \le \sqrt{2d\log(2d)};\\ &(6.3) \quad \text{implies} \quad \mathbb{E}\,\|\boldsymbol{G}\| \le 2\sqrt{d} + \mathrm{Const}\cdot\big[2d\log^3(2d)\big]^{1/4}. \end{aligned}$$

The latter inequality is the stronger one, and its first term is numerically sharp.

Roughly speaking, the logarithmic factor in (6.2) arises when the Hermitian dilations of two independent copies $\mathcal{H}(\boldsymbol{Z}), \mathcal{H}(\boldsymbol{Z}')$ of the Gaussian matrix "almost commute" with each other, in an appropriate sense. To quantify this insight, inspired by free probability, I introduced a subtle statistic called the *matrix alignment* [50, Def. 3.1]. For a special class of Gaussian matrices, I also established a preliminary version of the second-order Khinchin inequality [50, Thm. 3.1] that reveals how the matrix alignment arises.

The finished result (6.3) was obtained through additional insights of Bandeira and colleagues (see [4, 5]). The latter papers use matrix alignment statistics to compare the spectral norm of a Gaussian matrix with the spectral norm of a free probability model (namely, an operator semicircle) that admits explicit formulas. The role of the interaction energy $w(\boldsymbol{Z})$ is to provide a computable bound for the matrix alignment [4, Rem. 3.3]. Similar results hold for other spectral properties, such as the support of the spectrum and the mean singular value distribution.

6.4 Independent sums: Gaussian comparison. Let us return to a more general setting. Consider an independent sum of bounded, centered random matrices with dimension $d_1 \times d_2$:

$$\boldsymbol{S} := \sum_{k=1}^{n} \boldsymbol{X}_k \quad \text{where } \mathbb{E}[\boldsymbol{X}_k] = \boldsymbol{0} \text{ and } \|\boldsymbol{X}_k\| \le B.$$

Let $\mathsf{V} := \mathsf{Var}[\boldsymbol{S}]$ be the variance function of the sum. The multivariate central limit theorem suggests that we should compare $\boldsymbol{S}$ with the centered Gaussian random matrix $\boldsymbol{Z} \sim \textsc{normal}(\boldsymbol{0}, \mathsf{V})$, but it is not clear how to quantify the difference between their distributions.

Under weak assumptions, we can obtain detailed nonasymptotic comparisons for *spectral properties* of the independent sum $\boldsymbol{S}$ and the Gaussian model $\boldsymbol{Z}$. Brailovskaya and van Handel [11, Cor. 2.7] established that

$$\big|\,\mathbb{E}\,\|\boldsymbol{S}\| - \mathbb{E}\,\|\boldsymbol{Z}\|\big| \;\lesssim\; \big[v(\boldsymbol{S})B\log^2(d)\big]^{1/3} + B\log(d) + \sqrt{v_*(\boldsymbol{S})\log(d)} \quad \text{where } d := d_1 + d_2. \tag{6.4}$$

In view of the matrix Khinchin inequality (6.2), the Gaussian matrix satisfies $\mathbb{E}\,\|\boldsymbol{Z}\| \approx \sqrt{v(\boldsymbol{Z})} = \sqrt{v(\boldsymbol{S})}$. The weak variance term $\sqrt{v_*(\boldsymbol{S})}$ is often negligible. Therefore, in case $B^2 \ll v(\boldsymbol{S})$, the discrepancy between the expected norms is much smaller than the expected norm of the Gaussian matrix. In other words, when the bound B on the summands is sufficiently small, we can accurately approximate $\mathbb{E}\,\|\boldsymbol{S}\|$ by way of estimates for $\mathbb{E}\,\|\boldsymbol{Z}\|$. The second-order matrix Khinchin inequality (6.3) provides one such estimate.

Brailovskaya and van Handel [11] developed Gaussian comparisons, similar to (6.4), for other spectral statistics of the independent sum, including the support of the spectrum and the mean distribution of the singular values. Their arguments rely on Stein's method, cumulant expansions, Möbius inversion, and a selection of matrix inequalities. My paper [53] offers an alternative proof of their results, based on the method of exchangeable pairs. I recently obtained another type of Gaussian comparison [52] for the *minimum* eigenvalue of an independent sum of *psd* matrices; the proof exploits a deep result from matrix analysis, called Stahl's theorem [43].

We conclude with a critical observation. For independent sums, the latest matrix concentration results, such as (6.3) and (6.4), offer new insights. Nevertheless, it may be difficult to deploy these inequalities in applications, such as the ones in section 3, because we often lack fine-grained information about the variance function of the random matrix model. As a consequence, classic matrix concentration tools (e.g., Theorem 2.1) and the more recent refinements play complementary roles.

6.5 Other models. While the matrix concentration theory for the independent sum model is rather complete, other types of random matrix models remain mysterious and have continued to attract attention. Citations are limited to a few typical papers, as there is not enough space to do justice to this rich literature.

The nonlinear random matrix model concerns a matrix-valued function of underlying random variables that are usually (but not always) independent. Results for this model include matrix Efron–Stein inequalities [39] and (local) matrix Poincaré inequalities [25, 26]. There are also specialized results for matrix-valued polynomial chaos [7]. A related challenge arises from random matrix models with dependencies, such as the sum of fixed matrices modulated by negatively correlated scalar random variables; some recent progress on these questions appears in [28]. In addition to matrix martingales, researchers have also treated other types of sequential processes, such as matrix-valued Markov chains; for example, see [36]. Another line of work [6] pursues extensions of matrix concentration to random tensors.

Many of the models in the last paragraph emerged from problems in combinatorics and algorithms, and the theoretical results have led to satisfying progress. Thus, research on matrix concentration tools remains active, and it continues to make a profound impact on applications.

Acknowledgments. JAT gratefully acknowledges support under ONR Award N00014-24-1-2223 and the Caltech Carver Mead New Adventures Fund. Many thanks are due to Ethan Epperly for his valuable insights on a preliminary draft of this article.

References

[1] R. Ahlswede and A. Winter, *Strong converse for identification via quantum channels*, IEEE Trans. Inform. Theory, 48 (2002), pp. 569–579, https://doi.org/10.1109/18.985947.

[2] Z. Bai and J. W. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices*, 2nd ed., Springer, New York, 2010, https://doi.org/10.1007/978-1-4419-0661-8.

[3] J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Amer. Math. Soc., 12 (1999), pp. 1119–1178, https://doi.org/10.1090/S0894-0347-99-00307-0.

[4] A. S. Bandeira and M. T. Boedihardjo, *The Spectral Norm of Gaussian Matrices with Correlated Entries*, preprint, https://arxiv.org/abs/2104.02662, 2021.

[5] A. S. Bandeira, M. T. Boedihardjo, and R. van Handel, *Matrix concentration inequalities and free probability*, Invent. Math., 234 (2023), pp. 419–487, https://doi.org/10.1007/s00222-023-01204-6.

[6] A. S. Bandeira, S. Gopi, H. Jiang, K. Lucca, and T. Rothvoss, *Tensor concentration inequalities: A geometric approach*, in Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC '25, Association for Computing Machinery, New York, 2025, pp. 822–832, https://doi.org/10.1145/3717823.3718188.

[7] A. S. Bandeira, K. Lucca, P. Nizic-Nikolac, and R. van Handel, *Matrix chaos inequalities and chaos of combinatorial type*, in Proceedings of the 57th Annual ACM Symposium on Theory of Computing, STOC '25, Association for Computing Machinery, New York, 2025, pp. 795–805, https://doi.org/10.1145/3717823.3718238.

[8] J. Batson, D. A. Spielman, and N. Srivastava, *Twice-Ramanujan sparsifiers*, SIAM Rev., 56 (2014), pp. 315–334, https://doi.org/10.1137/130949117.

[9] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1997, https://doi.org/10.1007/978-1-4612-0653-8.

[10] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*, Oxford University Press, Oxford, 2013, https://doi.org/10.1093/acprof:oso/9780199535255.001.0001.

[11] T. Brailovskaya and R. van Handel, *Universality and sharp matrix concentration inequalities*, Geom. Funct. Anal., 34 (2024), pp. 1734–1838, https://doi.org/10.1007/s00039-024-00692-9.

[12] T. Brailovskaya and R. van Handel, *Extremal random matrices with independent entries and matrix superconcentration inequalities*, Ann. Probab., to appear.

[13] E. Campbell, *Random compiler for fast Hamiltonian simulation*, Phys. Rev. Lett., 123 (2019), 070503, https://doi.org/10.1103/PhysRevLett.123.070503.

[14] C.-F. Chen, H.-Y. Huang, R. Kueng, and J. A. Tropp, *Concentration for random product formulas*, Phys. Rev. X Quantum, 2 (2021), 040305, https://doi.org/10.1103/PRXQuantum.2.040305.

[15] M. P. Connolly and N. J. Higham, *Probabilistic rounding error analysis of Householder QR factorization*, SIAM J. Matrix Anal. Appl., 44 (2023), pp. 1146–1163, https://doi.org/10.1137/22M1514817.

[16] P. G. Constantine, *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, SIAM, Philadelphia, 2015, https://doi.org/10.1137/1.9781611973860.

[17] M. Croci, M. Fasi, N. J. Higham, T. Mary, and M. Mikaitis, *Stochastic rounding: Implementation, error analysis and applications*, Roy. Soc. Open Sci., 9 (2022), 211631, https://doi.org/10.1098/rsos.211631.

[18] D. A. Freedman, *On tail probabilities for martingales*, Ann. Probab., 3 (1975), pp. 100–118, https://doi.org/10.1214/aop/1176996452.

[19] M. Guţă, J. Kahn, R. Kueng, and J. A. Tropp, *Fast state tomography with optimal error bounds*, J. Phys. A, 53 (2020), 204001, https://doi.org/10.1088/1751-8121/ab8111.

[20] J. Haah, A. W. Harrow, Z. Ji, X. Wu, and N. Yu, *Sample-optimal tomography of quantum states*, IEEE Trans. Inform. Theory, 63 (2017), pp. 5628–5641, https://doi.org/10.1109/TIT.2017.2719044.

[21] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002, https://doi.org/10.1137/1.9780898718027.

[22] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, 2013.

[23] S. R. Howard, A. Ramdas, J. McAuliffe, and J. Sekhon, *Time-uniform Chernoff bounds via nonnegative supermartingales*, Probab. Surv., 17 (2020), pp. 257–317, https://doi.org/10.1214/18-PS321.

[24] S. R. Howard, A. Ramdas, J. McAuliffe, and J. Sekhon, *Time-uniform, nonparametric, nonasymptotic confidence sequences*, Ann. Statist., 49 (2021), pp. 1055–1080, https://doi.org/10.1214/20-AOS1991.

[25] D. Huang and J. A. Tropp, *From Poincaré inequalities to nonlinear matrix concentration*, Bernoulli, 27 (2021), pp. 1724–1744, https://doi.org/10.3150/20-bej1289.

[26] D. Huang and J. A. Tropp, *Nonlinear matrix concentration via semigroup methods*, Electron. J. Probab., 26 (2021), 8, https://doi.org/10.1214/20-EJP578.

[27] H.-Y. Huang, R. Kueng, and J. Preskill, *Predicting many properties of a quantum system from very few measurements*, Nature Phys., 16 (2020), pp. 1050–1057, https://doi.org/10.1038/s41567-020-0932-7.

[28] T. Kaufman, R. Kyng, and F. Soldá, *Scalar and matrix Chernoff bounds from ℓ_∞-independence*, in Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA), SIAM, Philadelphia; ACM, New York, 2022, pp. 3732–3753, https://doi.org/10.1137/1.9781611977073.147.

[29] R. Kyng and S. Sachdeva, *Approximate Gaussian elimination for Laplacians—fast, sparse, and simple*, in 57th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2016, IEEE Computer Society, Los Alamitos, CA, 2016, pp. 573–582, https://doi.org/10.1109/FOCS.2016.68.

[30] R. Latała and K. Oleszkiewicz, *Gaussian measures of dilatations of convex symmetric sets*, Ann. Probab., 27 (1999), pp. 1922–1938, https://doi.org/10.1214/aop/1022677554.

[31] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: Isoperimetry and Processes*, Springer-Verlag, Berlin, 1991, https://doi.org/10.1007/978-3-642-20212-4.

[32] E. H. Lieb, *Convex trace functions and the Wigner–Yanase–Dyson conjecture*, Advances in Math., 11 (1973), pp. 267–288, https://doi.org/10.1016/0001-8708(73)90011-X.

[33] N. Linial, E. London, and Y. Rabinovich, *The geometry of graphs and some of its algorithmic applications*, Combinatorica, 15 (1995), pp. 215–245, https://doi.org/10.1007/BF01200757.

[34] F. Lust-Piquard, *Inégalités de Khintchine dans C_p $(1 < p < \infty)$*, C. R. Acad. Sci. Paris Sér. I Math., 303 (1986), pp. 289–292.

[35] H. L. Montgomery, *The pair correlation of zeros of the zeta function*, in Analytic Number Theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), American Mathematical Society, Providence, RI, 1973, pp. 181–193.

[36] J. Neeman, B. Shi, and R. Ward, *Concentration inequalities for sums of Markov-dependent random matrices*, Inf. Inference, 13 (2024), iaae032, https://doi.org/10.1093/imaiai/iaae032.

[37] R. I. Oliveira, *Concentration of the Adjacency Matrix and of the Laplacian in Random Graphs with Independent Edges*, preprint, https://arxiv.org/abs/0911.0600, 2010.

[38] L. Pastur and M. Shcherbina, *Eigenvalue Distribution of Large Random Matrices*, American Mathematical Society, Providence, RI, 2011, https://doi.org/10.1090/surv/171.

[39] D. Paulin, L. Mackey, and J. A. Tropp, *Efron-Stein inequalities for random matrices*, Ann. Probab., 44 (2016), pp. 3431–3473, https://doi.org/10.1214/15-AOP1054.

[40] M. Rudelson, *Random vectors in the isotropic position*, J. Funct. Anal., 164 (1999), pp. 60–72, https://doi.org/10.1006/jfan.1998.3384.

[41] D. A. Spielman and N. Srivastava, *Graph sparsification by effective resistances*, SIAM J. Comput., 40 (2011), pp. 1913–1926, https://doi.org/10.1137/080734029.

[42] D. A. Spielman and S.-H. Teng, *Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems*, in Proceedings of the 36th Annual ACM Symposium on Theory of Computing (STOC), ACM Press, New York, 2004, pp. 81–90, https://doi.org/10.1145/1007352.1007372.

[43] H. R. Stahl, *Proof of the BMV conjecture*, Acta Math., 211 (2013), pp. 255–290, https://doi.org/10.1007/s11511-013-0104-z.

[44] S.-H. Teng, *The Laplacian paradigm: Emerging algorithms for massive graphs*, in Theory and Applications of Models of Computation, Springer, Berlin, 2010, pp. 2–14, https://doi.org/10.1007/978-3-642-13562-0_2.

[45] N. Tomczak-Jaegermann, *The moduli of smoothness and convexity and the Rademacher averages of trace classes* S_p $(1 \leq p < \infty)$, Studia Math., 50 (1974), pp. 163–182, http://eudml.org/doc/217886.

[46] J. A. Tropp, *Freedman's inequality for matrix martingales*, Electron. Commun. Probab., 16 (2011), pp. 262–270, https://doi.org/10.1214/ECP.v16-1624.

[47] J. A. Tropp, *User-friendly tail bounds for sums of random matrices*, Found. Comput. Math., 12 (2012), pp. 389–434, https://doi.org/10.1007/s10208-011-9099-z.

[48] J. A. Tropp, *An introduction to matrix concentration inequalities*, Found. Trends Mach. Learn., 8 (2015), pp. 1–230, https://doi.org/10.1561/2200000048.

[49] J. A. Tropp, *The expected norm of a sum of independent random matrices: An elementary approach*, in High Dimensional Probability VII, Springer, Cham, 2016, pp. 173–202, https://doi.org/10.1007/978-3-319-40519-3_8.

[50] J. A. Tropp, *Second-order matrix concentration inequalities*, Appl. Comput. Harmon. Anal., 44 (2018), pp. 700–736, https://doi.org/10.1016/j.acha.2016.07.005.

[51] J. A. Tropp, *Matrix Concentration & Computational Linear Algebra*, CMS Lecture Notes 2019-01, Caltech, Pasadena, CA, 2019, https://doi.org/10.7907/nbx6-vm05.

[52] J. A. Tropp, *Comparison Theorems for the Minimum Eigenvalue of a Random Positive-Semidefinite Matrix*, preprint, https://doi.org/10.48550/arXiv.2501.16578, 2025.

[53] J. A. Tropp, *Universality Laws for Random Matrices via Exchangeable Pairs*, manuscript, 2025.

[54] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*, Cambridge University Press, Cambridge, 2018, https://doi.org/10.1017/9781108231596.

[55] D. V. Voiculescu, K. J. Dykema, and A. Nica, *Free Random Variables*, American Mathematical Society, Providence, RI, 1992, https://doi.org/10.1090/crmm/001.

[56] M. J. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, Cambridge University Press, Cambridge, 2019, https://doi.org/10.1017/9781108627771.

[57] E. P. Wigner, *Characteristic vectors of bordered matrices with infinite dimensions*, Ann. of Math. (2), 62 (1955), pp. 548–564, https://doi.org/10.2307/1970079.

[58] J. Wishart, *The generalised product moment distribution in samples from a normal multivariate population*, Biometrika, 20A (1928), pp. 32–52, https://doi.org/10.1093/BIOMET/20A.1-2.32.