On the Largest Lyapunov Exponent for Products of Gaussian Matrices

Vladislav Kargin

Abstract The paper provides a new integral formula for the largest Lyapunov exponent of Gaussian matrices, which is valid in the real, complex and quaternion-valued cases. This formula is applied to derive asymptotic expressions for the largest Lyapunov exponent when the size of the matrix is large and compare the Lyapunov exponents in models with a spike and no spikes.

1 Introduction

Let A_k be random d-by-d matrices and let $P_n := A_n A_{n-1} \ldots A_1$. In a seminal paper Furstenberg and Kesten [6], showed that $\lim_{n\to\infty} n^{-1} \log \|P_n\|$ exists under some mild conditions on random matrices A_k. This limit is called *the largest Lyapunov exponent*. In fact, the limit of matrices $\left(P_n^* P_n\right)^{1/2n}$ is well defined (under some conditions on A_k), and the limit has d non-negative real eigenvalues [16,19]. The logarithms of these eigenvalues are called the *Lyapunov exponents* of matrices A_i. We denote them $\mu_1 \geq \cdots \geq \mu_d$.

As Lyapunov exponents are related to the behavior of complex dynamical systems, much research went into the study of their properties, such as their analyticity with respect to small matrix perturbations and to changes in the matrix random process [11,17,20], the multiplicity of the largest Lyapunov exponent [8], and others (see papers in [3]). Lyapunov exponents have been generalized to infinite-dimensional operators [10,21], and found important applications in the study of hydrodynamic stability [22] and random Schroedinger operators [2].

In Furstenberg [7], derived an integral formula for the largest Lyapunov exponent, which in the case of Gaussian matrices reduces to a multidimensional integral over a known density. However, there is only a limited number of cases in which this integral has been computed

V. Kargin (✉)
University of Cambridge, Palo Alto, CA, USA
e-mail: vladislav.kargin@gmail.com

in a closed form [4,13], although in some cases, it was shown that the exponents can be computed efficiently even if one does non have an explicit formula [18].

Recently, in Forrester [5], found a formula for all Lyapunov exponents of complex Gaussian matrices with a general covariance matrix Σ. His method is based on the Harish–Chandra–Itzykson–Zuber integration formula and cannot be applied to the case when matrices are real-valued.

In this paper we derive a new formula for the largest Lyapunov exponent of Gaussian matrices with general Σ, which is applicable in the real, complex and quaternion cases.

Recall that a d-by-d Gaussian matrix A with covariance matrix Σ has the probability density

$$P(A) = c_\beta \, (\det \Sigma)^{-d/2} \exp\left[-\frac{\beta}{2} \mathrm{Tr}\left(A^* \Sigma^{-1} A \right) \right],$$

where $\beta = 1, 2$, or 4 for real, complex or quaternion matrices, and c_β is a normalization constant. Equivalently, A can be defined as $\Sigma^{1/2} G$, where G has independent real, complex, or quaternion entries whose components are real Gaussian variables with variance $1/\beta$.

Theorem 1.1 *Let A_i be independent d-by-d Gaussian matrices with covariance matrix Σ, and let their entries be real, complex or quaternion, according to whether $\beta = 1, 2$, or 4. Assume that the eigenvalues of Σ are $\sigma_i^2 = 1/y_i$. Then, the following formula holds for the largest Lyapunov exponent of A_i,*

$$2\mu_1 = -\gamma + \log\left(\frac{2}{\beta}\right) + \int_0^\infty \left[\mathbf{1}_{[0,1]}(x) - \prod_{i=1}^d \left(1 + \frac{x}{y_i}\right)^{-\beta/2} \right] \frac{dx}{x}, \tag{1}$$

where $\gamma \approx 0.5772$ is the Euler constant.

As an application of this formula, we derive asymptotic expressions for the largest Lyapunov exponents of high-dimensional matrices. First, we compute the asymptotic behavior in the case when d is large and Σ does not have spikes so that all eigenvalues of Σ are of the same order. Next, we consider the case when the covariance matrix Σ does have a "spike". We consider the particular case when all except one eigenvalue of Σ equal $1/d$, and the exceptional eigenvalue does not depend on the dimension. We study the asymptotic behavior of the largest Lyapunov exponent when the dimension d is large, and find a significant difference relative to the case of no spikes.

More formally, we have the following result for the no-spike case.

Theorem 1.2 *Suppose that $A_i^{(d)}$ are independent d-by-d Gaussian matrices with covariance matrix Σ_d, and let the eigenvalues of Σ_d be $\theta_i^{(d)}/d$ where $i = 1, \ldots, d$. Assume that $\theta_i^{(d)}$ are bounded, $1 \leq \theta_i^{(d)} \leq L$, and that*

$$\lim_{d\to\infty} \mathrm{Tr}\,(\Sigma_d) \equiv \lim_{d\to\infty} \frac{1}{d} \sum_{i=1}^d \theta_i^{(d)} = \lambda.$$

Then, for the largest Lyapunov exponent of $A_i^{(d)}$, we have

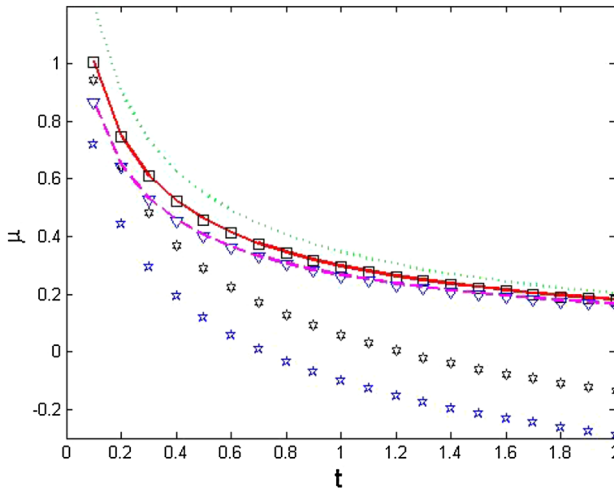$$\lim_{d\to\infty} \mu_1^{(d)} = \frac{1}{2} \log \lambda.$$

Fig. 1 The eigenvalues of the covariance matrix equal $1/d$ except one which equals $1/t$. The *blue pentagrams* show the largest Lyapunov exponents when the dimension $d = 2$ and $\beta = 1$; the *black hexagrams* are for $d = 2$ and $\beta = 2$; the *triangles* are for $d = 100$ and $\beta = 1$; the *squares* are for $d = 100$ and $\beta = 2$; the *red solid line* is the asymptotic prediction for $\beta = 2$ and the *magenta dashed line* is the prediction for $\beta = 1$; the *green dotted line* is the free probability prediction $\log(1 + 1/t)/2$ (Color figure online)

The result of Theorem 1.2 is in agreement with the free probability prediction (see for example, [9]).

For the model with a spike we have the following result illustrated in Fig. 1.

Theorem 1.3 *Suppose that A_i are independent d-by-d Gaussian matrices with the covariance matrix Σ_d, and that the eigenvalues of Σ_d are $\lambda_i^{(d)} = 1/d$ for $i = 1, \ldots, d - 1$, and $\lambda_d^{(d)} = 1/t$, where $0 < t < d$. In the complex case ($\beta = 2$), we have the following estimate,*

$$2\mu_1 = e^t \int_1^\infty e^{-tx} \frac{dx}{x} + O_t (1/d), \tag{2}$$
$$= -e^t \operatorname{Ei}(-t) + O_t (1/d),$$

where $\operatorname{Ei}(x)$ is the exponential integral function. In the real case ($\beta = 1$),

$$2\mu_1 = e^{t/2} \int_1^\infty e^{-tx/2} \frac{dx}{\sqrt{x}(\sqrt{x} + 1)} + O_t (1/d),$$

A comparison of the results for the spike and no-spike models shows that a spike can significantly influence the largest Lyapunov exponent. Indeed, in the situation of Theorem 1.3 with $\beta = 2$, the formula in Theorem 1.2 would incorrectly predict that for large d the first Lyapunov exponent is close to $\frac{1}{2} \log(1 + 1/t)$ (because $\lim_{d \to \infty} \frac{1}{d} \sum_{i=1}^d \theta_i^{(d)} = 1 + \frac{1}{t}$) instead of the correct $-\frac{1}{2} e^t \operatorname{Ei}(-t)$.

Theorem 1.3 concerns the case when the spike (i.e., the ratio of the exceptional eigenvalue to the bulk eigenvalues) grows at the same rate as the dimension. For other regimes, we have the following results.

Theorem 1.4 *Suppose that A_i are independent d-by-d complex Gaussian matrices with the covariance Σ, and that the eigenvalues of Σ_d are $\lambda_i^{(d)} = 1/d$ for $i = 1, \ldots, d - 1$, and*

$\lambda_d^{(d)} = \theta/d$, where $\theta > 1$. Then, for the largest Lyapunov exponent μ_1, we have

$$2\mu_1 = \frac{\theta - \frac{3}{2}}{d} + O\left(\left(\frac{\theta - 1}{d}\right)^2\right) + O\left(\frac{1}{d^2}\right).$$

Note that for $d \to \infty$ and θ fixed, this is essentially the no-spike model with $\lambda = \lim_{d\to\infty} \frac{1}{d}\sum_{i=1}^{d} \theta_i^{(d)} = 1$. Thus, Theorem 1.4 gives a second order correction to the result in Theorem 1.2.

Finally, consider the case of an exceptionally large eigenvalue which grows faster than the dimension d.

Theorem 1.5 *Suppose that A_i are independent d-by-d complex Gaussian matrices with the covariance Σ, and that the eigenvalues of Σ_d are $\lambda_i^{(d)} = 1$ for $i = 1, \ldots, d - 1$, and $\lambda_d^{(d)} = \theta > d$. Then, for the largest Lyapunov exponent μ_1, we have*

$$2\mu_1 = \log\theta - \gamma + O\left(\frac{d}{\theta}\log\theta\right).$$

In particular, if θ grows faster than d^α with $\alpha > 1$, then the last term is negligible for large θ. This happens, for example, if d is fixed and θ grows. It is remarkable that the Lyapunov exponent does not depend on the dimension d if the spike in covariance θ is much larger than d.

The rest of the paper is organized as follows. Section 2 collects necessary background information about Lyapunov exponents. Section 3 is devoted to the proof of the main formula. Section 4 derives asymptotic expressions for the largest Lyapunov exponent when the size of the matrices is large. And Sect. 5 is the conclusion.

2 Background Information About Lyapunov Exponents

One of the key facts about Lyapunov exponents is that they satisfy the following relation:

$$\mu_1 + \cdots + \mu_k = \lim_{n\to\infty} \frac{1}{n} \log \text{Vol}_k\left(y_1(n), \ldots, y_k(n)\right) \tag{3}$$

where $y_i(n) = P_n y_i(0)$ and $\{y_i(0)\}_{i=1}^{k}$ is an arbitrary orthonormal vector system.

For Gaussian matrices this formula can be significantly simplified. Namely, let A_i be independent Gaussian matrices with covariance matrix Σ. The crucial observation is that the distribution of $A_i^* A_i$ is invariant relative to the transformation

$$A_i^* A_i \to Q^* A_i^* A_i Q,$$

where Q is an arbitrary orthogonal matrix. This implies that the changes in the volume of a k-dimensional element are independent from step to step and that their distribution is the same as if they were applied to the element spanned by the standard basis vectors $\mathbf{e}_i$. Hence, by the law of large numbers,

$$\mu_1 + \cdots + \mu_k = \mathbb{E}\log\text{Vol}_k\left(A_1\mathbf{e}_1, \ldots, A_1\mathbf{e}_k\right)$$
$$= \frac{1}{2}\mathbb{E}\log\det\left(G_k^*\Sigma G_k\right) = \frac{1}{2}\frac{d}{d\mu}\mathbb{E}\left[\det\left(G_k^*\Sigma G_k\right)^\mu\right]\Big|_{\mu=0}, \tag{4}$$

where G_k denotes a random d-by-k matrix with the identically distributed standard Gaussian entries. (For details of the argument see [14] and [15]).

While formula (4) allows one to compute all Lyapunov exponents, it is essentially a multidimensional integral which can be computationally demanding. For this reason, it is of interest to obtain a more explicit way to calculate the Lyapunov exponents.

For real Gaussian matrices and the simplest situation when $\Sigma = \sigma^2 I$ and I is the identity matrix, Newman showed in [15] that

$$\mu_i = \frac{1}{2}\left[\log\left(2\sigma^2\right) + \Psi\left(\frac{d-i+1}{2}\right)\right], \tag{5}$$

where $\Psi(x)$ is the digamma function, $\Psi(x) := (\log\Gamma(x))'$. (At the positive integer points, $\Psi(n) = \sum_{k=1}^{n-1}\frac{1}{k} - \gamma$, where $\gamma = 0.5772\ldots$ is the Euler constant. At half-integers, $\Psi(n+1/2) = \sum_{k=1}^{n}\frac{1}{k-1/2} - 2\log 2 - \gamma$. The asymptotic behavior of the digamma function is given by the formula $\Psi(z) = \log z - \frac{1}{2z} - \frac{1}{12z^2}\left(1 + O\left(\frac{1}{z^2}\right)\right)$).

In particular if we normalize $\sigma^2 = 1/d$, then for $d = 1$ the largest Lyapunov exponent $\mu_1 = [-\log 2 - \gamma]/2$ and for $d \to \infty$, $\mu_1 = -\frac{1}{2d} + O\left(\frac{1}{d^2}\right)$.

In Forrester [5], established an analogue of formulas (5) for the complex-valued Gaussian matrices. Namely, Forrester showed that in the case of complex-valued Gaussian matrices with $\Sigma = \sigma^2 I$,

$$2\mu_i = \log\sigma^2 + \Psi(d-i+1) \tag{6}$$

(see Proposition 1 in [5] and note that the absence of 1/2 before Ψ is a typo in the statement of this proposition). If $\sigma^2 = 1/d$, then for $d = 1$ the largest Lyapunov exponent $\mu_1 = -\gamma/2$ and for $d \to \infty$, $\mu_1 = -\frac{1}{4d} + O\left(\frac{1}{d^2}\right)$.

What is more, in the complex-valued case there is an explicit formula for all Lyapunov exponents even if Σ is general. Namely, it is shown in [5] that

$$\mu_k = \frac{1}{2}\Psi(k) + \frac{1}{2\prod_{i<j}(y_i - y_j)}\det\begin{bmatrix}\left[y_j^{i-1}\right]_{i=1,\ldots,k-1;\,j=1,\ldots,d} \\ \left[(\log y_j)\,y_j^{k-1}\right]_{j=1,\ldots,d} \\ \left[y_j^{i-1}\right]_{i=k+1,\ldots,d;\,j=1,\ldots,d}\end{bmatrix}, \tag{7}$$

where y_j are eigenvalues of Σ^{-1}. In particular for $k = 1$, one can re-write this as

$$\mu_1 = \frac{1}{2}\left[\Psi(1) - \sum_{j=1}^{d}\frac{\log y_j}{\prod_{l\neq j}(1 - y_j/y_l)}\right] \tag{8}$$

provided that all y_j are different.

The proof of formula (7) is based on the Harish–Chandra–Itzykson–Zuber integral and cannot be directly generalized to the case of real or quaternion Gaussian matrices.

In fact, it appears that for the real-valued case with general Σ, an explicit formula (due to Mannion [12]) is only known for products of 2-by-2 Gaussian matrices:

$$\mu_1 = \frac{1}{2}\left[\Psi(1) + \log\left(\frac{1}{2}\mathrm{Tr}\Sigma + \sqrt{\det\Sigma}\right)\right]. \tag{9}$$

(Some explicit formulas are also known for 2-by-2 random matrices with non-Gaussian entries, see [13]. In addition, there are methods which sometime allow one to compute Lyapunov exponents efficiently even when explicit formulas are not available, see [18]).

Our formula (1) in Theorem 1.1 provides an explicit formula applicable for real, complex and quaternion matrices with general Σ. It is limited, however, to the case of the largest Lyapunov exponent. For $\beta = 2$, Forrester's formula (8) can be derived from our formula by evaluating the integral in (1) using residues. Similarly, for $\beta = 4$, one can use residues and derive the following formula:

$$2\mu_1 = \Psi(1) - \log(2)$$
$$- \left(\prod_{i=1}^{d} y_i\right)^2 \sum_{i=1}^{d} \left\{ \frac{1}{y_i^2 \prod_{j\neq i} (y_i - y_j)^2} \left[1 - \log y_i \left(1 + \sum_{j\neq i} \frac{2y_i}{y_i - y_j}\right)\right]\right\}.$$
(10)

In the remaining case of $\beta = 1$, the integral can be easily evaluated numerically. In the asymptotic analysis of large-dimensional situations, we will use our formula (1) as most convenient even for $\beta = 2$ and 4.

3 Proof of the Formula for the Largest Lyapunov Exponent

We start the proof of Theorem 1.1 by interpreting the basic formula (4) for $k = 1$. Namely, by (4), μ_1 is the expected logarithm of $\left\|\Sigma^{1/2}G_1\right\|$. Since $\left\|\Sigma^{1/2}G_1\right\|^2$ is a weighted sum of the squares of independent Gaussian variables, its distribution is easy to calculate. Indeed, in the real case $\beta = 1$, the characteristic function of the sum is

$$f(t) = \prod_{j=1}^{d} \left(1 - 2i\sigma_j^2 t\right)^{-1/2}.$$

(In this formula, the function $z^{-1/2}$ is determined by making a cut along $z < 0$, and the selected branch of $\left(1 - 2i\sigma_j^2 t\right)^{-1/2}$ equals 1 at $t = 0$). Hence, the distribution density of the sum is

$$p(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\lambda t} \prod_{j=1}^{d} \left(1 - 2i\sigma_j^2 t\right)^{-1/2} dt$$
$$= \frac{1}{2\pi} \left(\Pi_{j=1}^{d} y_j\right)^{1/2} \int_{-\infty}^{\infty} e^{-i\lambda t} \prod_{j=1}^{d} (y_j - 2it)^{-1/2} dt$$
$$= c \int_{0+i\infty}^{0-i\infty} e^{-\lambda z/2} \prod_{i=1}^{d} (z - y_i)^{-1/2} dz,$$

where $y_i = 1/\sigma_i^2$ and the normalization constant c may depend on y_i and the choice of branch for functions $(z - y_i)^{-1/2}$. By changing the contour, we get the expression

$$p(\lambda) = \frac{1}{c} \int_{\mathcal{C}} e^{-\lambda z/2} \prod_{i=1}^{d} (z - y_i)^{-1/2} dz,$$

and the contour of integration goes in the counterclockwise direction around points y_i. In more detail, we can take the contour $\mathcal{C}$ that starts at a large $R > \max\{y_i\}$, goes along the upper edge of the real axis to $r > 0$, which is smaller than all of y_i, then crosses to the lower
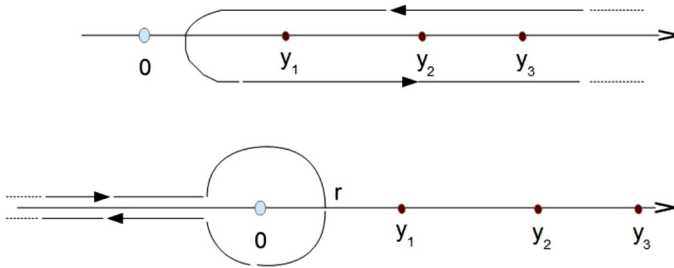
Fig. 2 Original and modified contours

edge of the real axis, and then returns along this lower edge to R. We then take the limit for $R \to \infty$.

For general β (i.e., $\beta = 1, 2,$ or 4), we can similarly obtain

$$p_\beta(\lambda) = \frac{1}{c_\beta} \int_C e^{-\beta \lambda z/2} \prod_{i=1}^{d} (z - y_i)^{-\beta/2} \, dz.$$

Next, we use formula (4) and find that

$$\mu_1 = \frac{1}{2c_\beta} \int_0^\infty \log \lambda \left[\int_C e^{-\beta \lambda z/2} \prod_{i=1}^{d} (z - y_i)^{-\beta/2} \, dz \right] d\lambda, \tag{11}$$

where

$$c_\beta = \int_0^\infty \int_C e^{-\beta \lambda z/2} \prod_{i=1}^{d} (z - y_i)^{-\beta/2} \, dz d\lambda. \tag{12}$$

By changing the order of integration in (11) and computing the inner integral, we find (with the help of the identity $\int_0^\infty (\log t) e^{-t} dt = -\gamma$) that

$$2\mu_1 = -\gamma + \frac{1}{\tilde{c}_\beta} \left[\frac{1}{2\pi i} \int_C \left(\frac{2}{\beta z} \right) \log \left(\frac{2}{\beta z} \right) \prod_{i=1}^{d} (z - y_i)^{-\beta/2} \, dz \right], \tag{13}$$

where

$$\tilde{c}_\beta = \frac{c_\beta}{2\pi i} = \frac{1}{2\pi i} \int_C \left(\frac{2}{\beta z} \right) \prod_{i=1}^{d} (z - y_i)^{-\beta/2} \, dz$$

$$= \frac{2}{\beta} \prod_{i=1}^{d} (-y_i)^{-\beta/2}.$$

This implies that

$$2\mu_1 = \Psi(1) + \log \left(\frac{2}{\beta} \right) + \frac{1}{2\pi i} \int_C \log(z) \prod_{i=1}^{d} \left(1 - \frac{z}{y_i} \right)^{-\beta/2} \frac{dz}{z}. \tag{14}$$

Now, let us change the contour. As before, let the contour start at a large $R > \max\{y_i\}$, goes along the upper edge of the real axis to $r > 0$, which is smaller than all of y_i, then crosses to the lower edge of the real axis, and then returns along this lower edge to R (Fig. 2).

Next, let us move this contour of integration so that the new contour $\mathcal{C}'$ starts at $-R$, goes along the upper edge of the real axis to $-r$, then goes along the circle of radius r around the 0 in the clockwise direction, and then returns to $-R$ along the lower edge of the real axis. The change in the integral corresponding to the change from contour $\mathcal{C}$ to contour $\mathcal{C}'$ is small for large R.

Indeed, consider the contour that starts at $-R$, then goes to $-r$ along the upper edge of the real axis, then to r clockwise along the circle centered at zero, and then to R along the upper edge of the real axis. Finally its returns back to $-R$ conterclocwise along the circle centered at zero. The integral over this closed contour is zero because the integrand is holomorphic inside the contour. In addition, the integral over the semicircle with radius R is small for large R. Hence, the integral over the first portion of this contour is also small. This shows that the difference in integrals is small for those portions of the contours $\mathcal{C}$ and $\mathcal{C}'$ that are above the real axis. Similar we can treat the portions of the contours $\mathcal{C}$ and $\mathcal{C}'$ that are below the real axis.

Then by computing the integrals over the two rays and the circle, and by letting $r \to 0$, we find that

$$
\lim_{\substack{R \to \infty, \\ r \to 0}} \frac{-1}{2\pi i} \int_{\mathcal{C}'} \log(z) \prod_{i=1}^{d} \left(1 - \frac{z}{y_i}\right)^{-\beta/2} \frac{dz}{z} = \lim_{r \to 0} \left\{ \int_r^\infty \prod_{i=1}^{d} \left(1 + \frac{x}{y_i}\right)^{-\beta/2} \frac{dx}{x} + \log r \right\}
$$

$$
= \int_0^1 \left(\prod_{i=1}^{d} \left(1 + \frac{x}{y_i}\right)^{-\beta/2} - 1 \right) \frac{dx}{x}
$$

$$
+ \int_1^\infty \prod_{i=1}^{d} \left(1 + \frac{x}{y_i}\right)^{-\beta/2} \frac{dx}{x}.
$$

(In the first integral we used the fact that the branches of the logarithm on the upper and the lower edges of the real axis differ by $2\pi i$). This proves formula (1). □

4 Asymptotic Behavior for Large Matrix Size

4.1 A Model Without Spikes

We start the proof of Theorem 1.2 by noting that a straightforward calculation gives the identity

$$
\Psi(1) + \log\left(\frac{2}{\beta}\right) + \int_0^\infty \left[1_{[0,1]}(x) - e^{-(\beta/2)\lambda x} \right] \frac{dx}{x} = \log \lambda.
$$

If we compare this with the formula (1), then we find that it is sufficient to show that

$$
\int_0^\infty \left[\prod_{i=1}^{d} \left(1 + \frac{\theta_i}{d} x\right)^{-\beta/2} - e^{-(\beta/2)\lambda x} \right] \frac{dx}{x} \to 0 \tag{15}
$$

as $d \to \infty$. First, consider

$$
I_1 := \int_0^\infty \left[\prod_{i=1}^{d} \left(1 + \frac{\theta_i}{d} x\right)^{-\beta/2} - \prod_{i=1}^{d} \exp\left(-\frac{\beta}{2} \frac{\theta_i}{d} x\right) \right] \frac{dx}{x}.
$$

We split I_1 in two integrals, $I_1' + I_1''$, the first one is over the interval from 0 to M, and the other is over the interval from M to infinity. For the second integral, we have

$$\int_M^\infty \exp\left(-\frac{\beta}{2}\frac{\sum \theta_i}{d}x\right)\frac{dx}{x} = \int_{M\left(\frac{\beta}{2}\right)\frac{1}{d}\sum\theta_i}^\infty \frac{e^{-t}dt}{t} \le e^{-M(\beta/2)} \to 0$$

as $M \to \infty$ and the convergence is uniform in d. In addition,

$$\int_M^\infty \prod_{i=1}^d \left(1 + \frac{\theta_i}{d}x\right)^{-\beta/2}\frac{dx}{x} \le \int_{M/d}^\infty (1+t)^{-d\beta/2}\frac{dt}{t}$$

$$\le \frac{d}{M}\int_{M/d}^\infty (1+t)^{-d\beta/2}\,dt$$

$$= \frac{d}{M}\frac{1}{d\,(\beta/2)-1}\left(1+\frac{M}{d}\right)^{-d\beta/2+1} \to 0$$

as $M \to \infty$, again uniformly in $d > 1$. We conclude that for every $\varepsilon > 0$, we can find M_0 such that $\left|I_1''\right| \le \varepsilon$ for all $M \ge M_0$ and all $d > 1$. In words, we can make I_1'' arbitrarily small uniformly in d by taking M sufficiently large.

For the integral I_1', we estimate the integrand by using the fact that if $|z_i| \le 1$ and $|w_i| \le 1$, then

$$\left|\prod_{i=1}^d z_i - \prod_{i=1}^d w_i\right| \le \sum_{i=1}^d |z_i - w_i|,$$

(see Lemma 1 of Section 27 in Billingsley [1]). Since

$$\left|\left(1+\frac{\theta_i}{d}x\right)^{-\beta/2} - \exp\left(-\frac{\beta}{2}\frac{\theta_i}{d}x\right)\right| \le C\left(\frac{\theta_i}{d}x\right)^2$$

for all $x \le d/L$, therefore (for $d \ge LM$), we estimate

$$I_1' \le \int_0^M C\frac{L^2}{d}x\,dx = \frac{CL^2M^2}{2d}.$$

For a fixed M, this can be made arbitrarily small by choosing d sufficiently large. Hence, $I_1 \to 0$ as $d \to \infty$.

Similarly,

$$I_2'' := \int_M^\infty \left(e^{-(\beta/2)\lambda x} - e^{-(\beta/2)d^{-1}(\sum\theta_i)x}\right)\frac{dx}{x} \to 0$$

as $M \to \infty$ uniformly in d, and for

$$I_2' := \int_0^M \left(e^{-(\beta/2)\lambda x} - e^{-(\beta/2)d^{-1}(\sum\theta_i)x}\right)\frac{dx}{x},$$

we estimate

$$|I_2'| \le \int_0^M C\left|\lambda - d^{-1}\left(\sum \theta_i\right)\right|dx$$

$$= CM\left|\lambda - d^{-1}\left(\sum \theta_i\right)\right| \to 0$$

as $d \to \infty$ for a fixed M. Altogether, the convergence of I_1 and $I_2 := I_2' + I_2''$ to zero proves (15) and completes the proof.   □

4.2 A Model with a Spike

In order to prove Theorem 1.3, consider a slightly modified model in which the eigenvalues of Σ_d are $\lambda_i^{(d)} = 1$ for $i = 1, \ldots, d-1$, and $\lambda_d^{(d)} = \theta = d/t > 1$. This model is obtained from the model in Theorem 1.3 by multiplying Σ_d by d. Therefore, the results in the theorem can be recovered by subtracting $\frac{1}{2} \log d$ from the largest Lyapunov exponent of the modified model.

For the complex case, we have

$$
\begin{aligned}
2\mu_1 &= \Psi(d) + \int_0^\infty \frac{1}{(1+x)^{d-1}} \left[\frac{1}{1+x} - \frac{1}{1+\theta x} \right] \frac{dx}{x} \\
&= \Psi(d) + \int_0^\infty \frac{1}{(1+u/d)^d} \left[\frac{1}{1+u/d} - \frac{1}{1+u/t} \right] \frac{du}{u} \\
&= \log d + \int_0^\infty e^{-u} \left[1 - \frac{1}{1+u/t} \right] \frac{du}{u} + O_t(1/d) \\
&= \log d + \int_0^\infty e^{-xt} \frac{dx}{1+x} + O_t(1/d) \\
&= \log d + e^t \int_1^\infty e^{-xt} \frac{dx}{x} + O_t(1/d).
\end{aligned}
$$

For the real case, we assume $d = 2k$ (the other case is similar) and write:

$$
\begin{aligned}
2\mu_1 &= \log 2 + \Psi(k) + \int_0^\infty \frac{1}{(1+x)^{k-1/2}} \left[\frac{1}{\sqrt{1+x}} - \frac{1}{\sqrt{1+\theta x}} \right] \frac{dx}{x} \\
&= \log 2 + \Psi(k) + \int_0^\infty \frac{1}{\left(1+\frac{u}{k}\right)^{k-1/2}} \left[\frac{1}{\sqrt{1+u/k}} - \frac{1}{\sqrt{1+2u/t}} \right] \frac{du}{u} \\
&= \log d + \int_0^\infty e^{-u} \left[1 - \frac{1}{\sqrt{1+2u/t}} \right] \frac{du}{u} + O_t(1/d) \\
&= \log d + e^{t/2} \int_1^\infty e^{-tx/2} \frac{dx}{\sqrt{x}\left(\sqrt{x}+1\right)} + O_t(1/d).
\end{aligned}
$$

(The last step uses the change of variable $x = 1 + 2u/t$.) □

Proof of Theorem 1.4 Again, it is more convenient to use the model in which the eigenvalues of Σ_d are $\lambda_i^{(d)} = 1$ for $i = 1, \ldots, d-1$, and $\lambda_d^{(d)} = \theta > 1$. It is elementary to compute that

$$
\int_0^\infty \left[\mathbf{1}_{[0,1]}(x) - (1+x)^{-d} \right] \frac{dx}{x} = 1 + \frac{1}{2} + \cdots + \frac{1}{d-1} = \Psi(d) - \Psi(1).
$$

(For example, one can use the identity $\frac{1}{(1+x)^d x} = \frac{1}{x} - \frac{1}{1+x} - \cdots - \frac{1}{(1+x)^d}$). Hence, by using formula (1), we can write

$$
\begin{aligned}
2\mu_1 &= \Psi(1) + \int_0^\infty \left[\mathbf{1}_{[0,1]}(x) - \frac{1}{(1+x)^{d-1}(1+\theta x)} \right] \frac{dx}{x} \\
&= \Psi(d) + \int_0^\infty \left[\frac{1}{(1+x)^{d-1}} \left(\frac{1}{1+x} - \frac{1}{1+\theta x} \right) \right] \frac{dx}{x}, \\
&= \Psi(d) + f_d, \hspace{4cm} (16)
\end{aligned}
$$

where

$$f_d := (\theta - 1) \int_0^\infty \frac{1}{(1+x)^d (1+\theta x)} dx \le \frac{\theta - 1}{d} \qquad (17)$$

One can check that for $d \ge 1$, the additional term f_d satisfies the recursion equation:

$$f_d = \left(\frac{\theta - 1}{\theta} \right) \left(\frac{1}{d} + f_{d+1} \right). \qquad (18)$$

Hence, we obtain a convergent series for f_d,

$$f_d = s \sum_{k=0}^\infty \frac{s^k}{d+k},$$

where $s := (\theta - 1) /\theta < 1$. Hence,

$$\left| df_d - \frac{s}{s-1} \right| = s \left| \sum_{k=0}^\infty \left(\frac{s^k}{1+k/d} - s^k \right) \right|$$

$$= \frac{s}{d} \sum_{k=0}^\infty \frac{ks^k}{1+k/d} \le \frac{1}{d} \left(\frac{s}{s-1} \right)^2.$$

That is,

$$\left| f_d - \frac{\theta - 1}{d} \right| \le \left(\frac{\theta - 1}{d} \right)^2.$$

Together with the asymptotic expansion for the digamma function, $\Psi(d) = \log d - \frac{1}{2d} + O\left(\frac{1}{d^2} \right)$, this limit implies the statement of the theorem.   □

Proof of Theorem 1.5 We can use the recursion in (18) and the initial condition $f_1 = \log \theta$ in order to obtain

$$f_d = \left(\frac{\theta}{\theta - 1} \right)^{d-1} \left(\log \theta - \sum_{k=1}^{d-1} \frac{1}{k} \left(1 - \frac{1}{\theta} \right)^k \right). \qquad (19)$$

Note that

$$\left| \left(\frac{\theta}{\theta - 1} \right)^{d-1} - 1 \right| = \left| \exp \left[(d-1) \log \left(1 + \frac{1}{\theta - 1} \right) \right] - 1 \right|$$

$$= \left| \exp \left[\frac{d-1}{\theta - 1} + O\left(\frac{d-1}{(\theta - 1)^2} \right) \right] - 1 \right|$$

$$= O\left(\frac{d-1}{\theta - 1} \right).$$

Similarly, we estimate

$$\sum_{k=1}^{d-1} \frac{1}{k} \left(1 - \frac{1}{\theta} \right)^k = \sum_{k=1}^{d-1} \frac{1}{k} \left[1 + O\left(\frac{d}{\theta} \right) \right].$$

Since $\sum_{k=1}^{d-1} \frac{1}{k} = \Psi(d) - \Psi(1) \sim \log d$, we find that

$$f_d = \log \theta - \Psi(d) + \Psi(1) + O\left(\frac{d}{\theta} \log \theta\right).$$

Therefore by (16), $2\mu_1 = \log \theta - \gamma + O\left(\frac{d}{\theta} \log \theta\right)$.   □

5 Conclusion

We derived a new formula for the largest Lyapunov exponent of Gaussian matrices and the asymptotic expressions for this exponent when the matrices are large. The asymptotic expressions are derived for two distinct cases, one is when the covariance matrix has no spikes and another one is when it has a single spike.

An interesting question is whether the asymptotic expressions derived in this paper remain valid for non-Gaussian random matrices with independent entries. Some numerical evidence concerning this question is presented in Figs. 3 and 4.

In order to create plots in these figures, we estimate the Lyapunov exponent using simulations and then plot its difference from the theoretical prediction in the formula (1). The difference is shown against the dimension of the matrix in order to check whether it declines with the increase in dimension as the concept of universality would predict.

Figure 3 presents this plot for the matrices GA, where G is the diagonal matrix with half of the diagonal entries equal to 1 and the remaining half equal to 2. The dashed line is for a d-by-d Gaussian matrix A where each entry is independent and have variance d^{-1}. It is very close to zero even for relatively small d.

The solid line is for a d-by-d matrix A where each entry is independent and equals $\pm d^{-1/2}$ with probability $1/2$. The solid line approaches zero as d grows, which suggests that
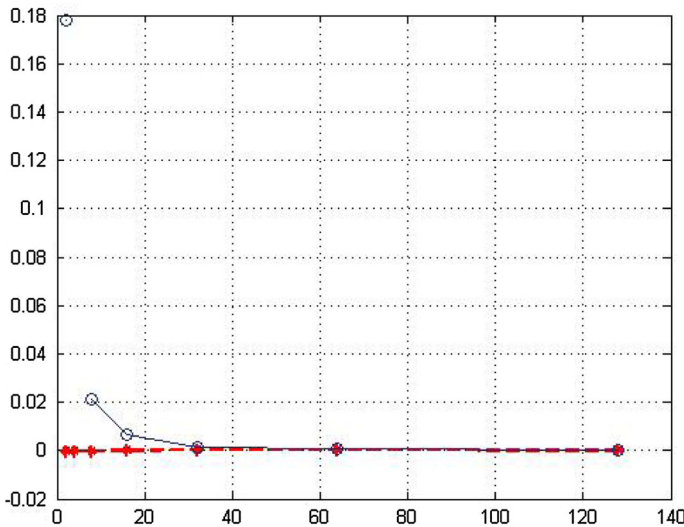

Fig. 3 The deviation of the largest Lyapunov exponent from the theoretical prediction. The no-spike case. The *dashed line* is for Gaussian and the *solid line* is for Bernoulli matrices. The *horizontal axis* shows the size of the matrix
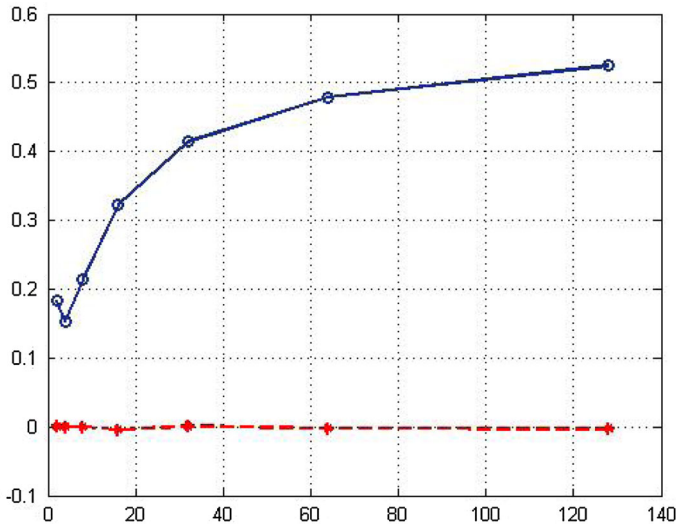
Fig. 4 The deviation of the largest Lyapunov exponent from the theoretical prediction. The spike case

the theoretical approximation gives a good approximation for the Lyapunov exponent of the Bernoulli matrix A.

(This numerical evidence should be taken with a grain of salt. While it might indicate the universality of the theoretical approximation for a large class of matrices, in fact this approximation is not valid for Bernoulli matrices (at least in the strict sense). Indeed, with positive probability the Bernoulli matrix A will have all entries equal to $d^{-1/2}$ and therefore the starting vector will be mapped to a vector proportional to $(1, \ldots, 1, 2, \ldots, 2)$. For $d \geq 4$, this vector has a positive probability to be mapped to 0 by an appropriate Bernoulli matrix A. Hence, with probability 1 the product of independent matrices A_k will eventually maps an arbitrary starting vector to zero. It can be seen that this happens for $d = 4$ in our plot where the data point is missing. However, for large d the simulations would need a prohibitively large time to detect this phenomenon).

Figure 4 shows an analogous plot for the case when the matrix G has a spike. Namely, G is assumed to be diagonal with all diagonal entries except one equal to 1 and the exceptional diagonal entry equal to d. This plot shows that the theoretical approximation from Theorem 1.1 is valid for Gaussian but invalid for Bernoulli matrices. This raises a natural question of how the approximation should be modified for large non-Gaussian matrices with a spike in the covariance matrix.

References

1. Billingsley, P.: Probability and Measure, 3rd edn. Wiley, New York (1995)
2. Bougerol, P., Lacroix, J.: Products of Random Matrices with Applications to Schrodinger Operators, Volume 8 of Progress in Probability and Statistics. Birkhauser, Boston (1985)
3. Cohen J.E., Kesten, H., Newman, C.M. (eds): Random Matrices and Their Applications, Volume 50 of Contemporary Mathematics. American Mathematical Society (1986)
4. Cohen, J.E., Newman, C.M.: The stability of large random matrices and their products. Ann. Probab. **12**, 283–310 (1984)

5. Forrester, P.J.: Lyapunov exponents for products of complex Gaussian random matrices. J. Stat. Phys. **151**, 796–808 (2013). arXiv:1206.2001
6. Furstenberg, H., Kesten, H.: Products of random matrices. Ann. Math. Stat. **31**, 457–469 (1960)
7. Furstenberg, H.: Noncommuting random products. Trans. Am. Math. Soc. **108**, 377–428 (1963)
8. Guivarch, Y., Raugi, A.: Frontiere de Furstenberg, proprietes de contraction et theoremes de convergence. Zeit. Fur Wahrscheinlichkeitstheorie und Verw. Gebiete **67**, 265–278 (1985)
9. Kargin, V.: The norm of products of free random variables. Probab. Theory Relat. Fields **139**, 397–413 (2007). arXiv:math/0611593
10. Kargin, V.: Lyapunov exponents of free operators. J. Funct. Anal. **255,** 1874–1888 (2008). arXiv:0712.1378
11. Le Page, E.: Régularité de plus grand exposant charactéristique de produit de matrices aléatoires indépendantes et applications. Annales de l'Institut Henri Poincaré. Probabilités et Statistiques **25**, 109–142 (1989)
12. Mannion, D.: Products of 2×2 random matrices. Ann. Appl. Probab. **3**, 1189–1218 (1993)
13. Marklof, J., Tourigny, Y., Wolowski, L.: Explicit invariant measures for products of random matrices. Trans. Am. Math. Soc. **360**, 3391–3427 (2008)
14. Newman, C.M.: Lyapunov exponents for some products of random matrices: exact expressions and asymptotic distributions. In: Cohen, J.E., Kesten, H., Newman, C.M. (eds) Random Matrices and Their Applications, Volume 50 of Contemporary Mathematics, pp. 183–195. American Mathematical Society (1986)
15. Newman, C.M.: The distribution of Lyapunov exponents: exact results for random matrices. Commun. Math. Phys. **103**, 121–126 (1986)
16. Oseledec, V.I.: A multiplicative ergodic theorem. Ljapunov characteristic numbers for dynamical systems. Trans. Mosc. Math. Soc. **19**, 197–231 (1968)
17. Peres, Y.: Domains of analytic continuation for the top Lyapunov exponent. Ann. I. H. Poincare **28**, 131–148 (1992)
18. Pollicott, M.: Maximal Lyapunov exponents for random matrix products. Invent. Math. **181**, 209–226 (2010)
19. Raghunathan, M.S.: A proof of Oseledec's multiplicative ergodic theorem. Isr. J. Math. **32**, 356–362 (1979)
20. Ruelle, D.: Analyticity properties of the characterisic exponents of random matrix products. Adv. Math. **32**, 68–80 (1979)
21. Ruelle, D.: Characteristic exponents and invariant manifolds in Hilbert space. Ann. Math. **115**, 243–290 (1982)
22. Ruelle, D.: Characteristic exponents for a viscous fluid subjected to time dependent forces. Commun. Math. Phys. **93**, 285–300 (1984)