Estimating generalized Lyapunov exponents for products of random matrices

J. Vanneste*
School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom
(Received 12 November 2009; published 8 March 2010)

We discuss several techniques for the evaluation of the generalized Lyapunov exponents which characterize the growth of products of random matrices in the large-deviation regime. A Monte Carlo algorithm that performs importance sampling using a simple random resampling step is proposed as a general-purpose numerical method which is both efficient and easy to implement. Alternative techniques complementing this method are presented. These include the computation of the generalized Lyapunov exponents by solving numerically an eigenvalue problem, and some asymptotic results corresponding to high-order moments of the matrix products. Taken together, the techniques discussed in this paper provide a suite of methods which should prove useful for the evaluation of the generalized Lyapunov exponents in a broad range of applications. Their usefulness is demonstrated on particular products of random matrices arising in the study of scalar mixing by complex fluid flows.

DOI: 10.1103/PhysRevE.81.036701 PACS number(s): 02.70.Uu, 05.45.−a, 47.51.+a, 47.52.+j

I. INTRODUCTION

Products of random matrices arise in many physical models, of disordered media, of wave localization, and of chaotic dynamics in particular. The main quantity of interest is the largest Lyapunov exponent, which gives the rate of exponential growth of the products as the number N of factors increases to infinity. The free energy of random Ising models, for instance, is given by the largest Lyapunov exponent of a product of matrices, as is the localization length of some random Schrödinger operators. We refer the reader to the book by Crisanti, Paladin, and Vulpiani [1] for a discussion of these and other applications.

Often, it is necessary to go beyond the almost sure, infinite-N growth of the matrix product captured by the largest Lyapunov exponent, and examine finite-N fluctuations. These are characterized by the distribution of the so-called finite-time (or finite-N) Lyapunov exponents or, equivalently, by the generalized Lyapunov exponents $\ell(q)$, which give the growth rate of the q-th moment of the norm of the matrix product (e.g., [1–4]). At a mathematical level, the generalization involved entails the passage from the (mutliplicative, noncommutative) law of large numbers [5–7] to the corresponding theory of large deviations ([8], and references therein).

One area of application in which multiplicative large deviations and generalized Lyapunov exponents play a central role is the transport, mixing and reaction of constituents in complex fluid flows. In the last ten years or so, a number of results have related the macroscopic dynamics of scalars and fields in fluid flows to the large-deviation statistics of the stretching by these flows (see [9], for an early review). Specifically, the generalized Lyapunov exponents associated with the stretching have been found to control the decay rate of purely advected passive scalars [9–14], the spatial distribution of reacting scalars [15,16], the reaction rate of fast reactions [17], the distribution of vorticity in certain turbulent flows [18,19], the clustering of inertial particles [20], the magnetic field in kinematic dynamo models [21], etc. In most of these applications, the complex fluid flows are modeled by random processes which either are white in time (Kraichnan–Kazantsev flows), or consist of sequences of independent identically distributed (iid) processes (variously termed renewing, renovating, or innovating flows). In the latter case, the stretching is controlled by products of iid random matrices of the type considered in this paper.

In many of these applications, it is necessary to evaluate the generalized Lyapunov exponents for specific random matrices. Very few exact results are available, however. As is also the case for the usual Lyapunov exponent, given in fact by $\ell'(0)$, these are essentially limited to matrices satisfying an isotropy property that reduces the problem to scalar multiplication [1,22,23]. Thus approximations to $\ell(q)$, either perturbative or numerical, need to be obtained. Crisanti *et al.* [1] review several techniques including Cook and Derrida's asymptotic results for large sparse matrices [24], the weak-disorder expansion for near-identity matrices, the replica trick (applicable when q is even and positive), and the (heuristic) microcanonical estimate. Cycle expansions [25,26] provide yet another technique. However, these techniques are limited to special ensembles of matrices: the microcanonical and cycle-expansion estimates, for instance, are applicable to ensembles drawn from a small number of matrices. There is, therefore, a genuine need for numerical techniques that enable the estimation of $\ell(q)$ for a broad range of matrix ensembles. The main aim of the present paper is to develop one such numerical technique and to demonstrate its usefulness by applying it to a few examples.

Several of the papers on fluid mixing cited above contain numerical evaluations of the generalized Lyapunov exponents corresponding to simple renewing flows, and in particular to the alternating sine map [27] that has become a standard tool in the field. Most of these estimates are obtained using a straightforward Monte Carlo sampling of either the probability distribution of the finite-time Lyapunov exponents, or of the q-th moments of the norm of the matrix product. This approach, which we refer to as brute-force Monte Carlo in what follows, is highly inefficient unless $|q|$

*j.vanneste@ed.ac.uk

is small. This is because it attempts to sample events that have an exponentially small probability as $N \to \infty$. Clearly, what is needed is some form of importance sampling, which focuses the computational effort on the realizations dominating the estimate of $\ell(q)$. We propose and test a simple algorithms that has this property. This algorithm, which we call resampled Monte Carlo (RMC), falls in the category of sequential importance-sampling [28] or "go-with-the-winners" strategies [29] used extensively in statistical physics and elsewhere; it consists of a simple modification of the brute-force computation adding a (random) resampling step which drastically reduces the sample variance. As a result, it yields accurate estimates of $\ell(q)$ with ensembles that are orders of magnitude smaller than those required for the brute-force estimation. The algorithm is very close to the cloning/pruning algorithm recently developed to estimate large-deviation statistics for more general Markov chains [30,31]. However, our focus on products of matrices leads to an algorithm that is particularly simple to implement and to analyze.

Recently, Haynes and Vanneste [14] used an alternative approach to the brute-force Monte Carlo sampling for the evaluation of $\ell(q)$ for the alternating sine flow (see also [16]). This approach relates $\ell(q)$ to the eigenvalue of an (infinite dimensional) eigenvalue problem that can be discretized and solved numerically, at least for 2×2 and perhaps 3×3 matrices. We review this approach, first to compare its results with those of our RMC algorithm, but mostly because the eigenvalue problem can be used to derive interesting properties of $\ell(q)$. One such property relates the function $\ell(q)$ associated with an ensemble of matrices A to the corresponding function associated with the complementary ensemble of matrices $A^{-1}/|\det A|^{1/q}$. This relationship is of great practical interest since considering $A^{-1}/|\det A|^{1/q}$ instead of A can lead to more accurate estimates of $\ell(q)$ for some value of q. We demonstrate the usefulness of this observation in some examples.

For large $|q|$, $\ell(q)$ is controlled by exceedingly rare realizations of the matrix products, and hence it is difficult to estimate reliably using Monte Carlo numerical methods, even with importance sampling. An alternative, which we pursue in this paper, is to take advantage of the largeness of $|q|$ to derive asymptotic estimates. Starting with the eigenvalue problem and using a WKB ansatz, we obtain the asymptotics of $\ell(q)$ for ensembles of bounded matrices and for matrices with (not necessarily independent) Gaussian entries. These asymptotic estimates, together with the RMC method, the eigenvalue formulation, and the replica approach (which we briefly discuss) provide a suite of methods which should prove useful for the evaluation of the generalized Lyapunov exponents of products of random matrices arising in a broad range of applications.

The plan of the paper is as follows. In Sec. II, we review the definition of the generalized Lyapunov exponents $\ell(q)$ and their connection with the large-deviation distribution of the finite-time Lyapunov exponents. We also derive the eigenvalue problem from which $\ell(q)$ can be inferred, and we use it to relate $\ell(q)$ obtained for the matrices A to its counterpart obtained for the matrices $A^{-1}/|\det A|^{1/q}$. The RMC algorithm is presented and analyzed in Sec. III; there we show that the algorithm leads to an unbiased estimate for the q-th moment of the matrix product, and we derive an expression for the variance of this estimate. Section IV is devoted to alternative methods for the evaluation of $\ell(q)$, namely, the numerical solution of the eigenvalue problem, the replica method, and the large-$|q|$ asymptotic results. All the methods discussed in the paper are tested on three examples of random matrix ensembles in Sec. V. The paper concludes with a discussion in Sec. VI. A pseudocode implementing the RMC method, and some technical derivations are relegated to three appendices.

II. GENERALIZED LYAPUNOV EXPONENTS

A. Definitions and basic properties

We consider N successive products of a vector $X_0 \in \mathbb{R}^d$ by iid random matrices $A_n \in \mathbb{R}^{d \times d}$, $n=1,2,\cdots,N$. In other words, we consider the recurrence

$$X_n = A_n X_{n-1}, \quad n = 1,2,\cdots,N. \tag{1}$$

We assume that X_0 is deterministic and normalized: $X_0 = x_0$ with $\|x_0\| = 1$. The randomness of the matrices A_n implies the choice of a probability measure on $\mathbb{R}^{d \times d}$. We will not be specific as to the properties of this measure; what we have in mind, as illustrated by the examples of Sec. V, are random matrices defined by a number of random parameters taken from smooth distributions such as the normal or uniform distributions.

Our focus is on the large-N behavior of $\|X_N\|$. This can be characterized by considering the generalized Lyapunov exponents

$$\ell(q) = \lim_{N \to \infty} \frac{1}{N} \log \mathbf{E} \|X_N\|^q, \tag{2}$$

where $\mathbf{E}$ denotes the expectation over the random matrices. Note that these exponents are independent of x_0 for almost all x_0 and realizations of the matrices A_n (e.g., [1,4]). Correspondly, the large-N asymptotics of the moments of $\|X_N\|$ is given by

$$\mathbf{E} \|X_N\|^q \sim c_q e^{N\ell(q)}, \tag{3}$$

for some c_q. Note that in the commutative case $d=1$, Eq. (3) is exact with $c_q = 1$. An alternative to the definition (2) of $\ell(q)$ that makes the independence on x_0 obvious is

$$\ell(q) = \lim_{N \to \infty} \frac{1}{N} \log \mathbf{E} \|A_N \cdots A_1\|^q, \tag{4}$$

where the matrix norm is the 2-norm, so that $\|A_N \cdots A_1\|$ is the largest singular value of $A_N \cdots A_1$.

The generalized Lyapunov function $\ell(q)$, sometimes termed free energy, obviously satisfies $\ell(0)=0$ and can be shown to be convex. It is directly related to the statistics of $\|X_N\|$ for $N \gg 1$ (e.g., [1,4]). These are usually described in terms of the (largest) finite-N Lyapunov exponents

$$H_N = \frac{1}{N}\log\|A_N \cdots A_1\|. \tag{5}$$

The large-deviation theory asserts that the pdf p_N of H_N is approximately

$$p_N(h) \asymp e^{-Ng(h)}, \tag{6}$$

where $\asymp$ denotes rough asymptotic equivalence, that is, asymptotic equivalence of the logarithms as $N \to \infty$. The function g, variously termed rate function, Cramér function or entropy, is convex. It attains a minimum at the Lyapunov exponent $\bar{h}$, which satisfies

$$\bar{h} = \lim_{N\to\infty} H_N \tag{7}$$

for almost all realizations of the random matrices, and it can be taken such that $g(\bar{h})=g'(\bar{h})=0$. Note that g is in fact independent of the norm chosen for $A_N \cdots A_1$, and that the same g would be obtained if Eq. (5) was replaced by $H_N = N^{-1}\log\|X_N\|$. Using the latter point, Laplace's method can be applied to write

$$\mathbf{E}\|X_N\|^q \asymp \int e^{Nqh} e^{-Ng(h)} dh \asymp e^{N \sup_h [qh-g(h)]},$$

and conclude from Eq. (2) that ℓ and g are Legendre transforms of one another,

$$\ell(q) = \sup_h [qh - g(h)]. \tag{8}$$

(Rigorous conditions on the probability measure for the A_n that guarantee that Eqs. (6) and (8) hold are given in [8]) Since $g'(\bar{h})=0$, the Legendre relationship $\ell'(q)=h$ gives

$$\bar{h} = \ell'(0). \tag{9}$$

B. Eigenvalue problem

The generalized Lyapunov exponents $\ell(q)$ can be found by solving a family of eigenvalue problems parameterized by q. To see this, we consider

$$u_n(x) = \mathbf{E} f(A_n \cdots A_1 x), \tag{10}$$

for a given function $f: \mathbb{R}^d \to \mathbb{R}$. We now derive a backward equation for u_n by noting that

$$u_{n+1}(x) = \mathbf{E} f(A_{n+1} \cdots A_1 x) = \mathbf{E} f(A_n \cdots A_1 A x) = \mathbf{E} u_n(Ax),$$

where the last expectation involves the single matrix A only. Thus, for an arbitrary f, the u_n satisfy the recurrence relation

$$u_{n+1}(x) = \mathbf{E} u_n(Ax), \quad \text{with} \quad u_0(x) = f(x). \tag{11}$$

In the particular case where $f(x)=\|x\|^q$, so that $u_n(x_0) = \mathbf{E}\|X_n\|^q$, Eq. (11) admits solutions of the form

$$u_n(x) = \lambda^n \|x\|^q v(\hat{e}), \tag{12}$$

where $\hat{e}=x/\|x\| \in S^{d-1}$ is a unit vector. The scalar λ and function v are determined by introducing Eq. (12) into Eq. (11) to obtain

$$(\mathcal{L}_q v)(\hat{e}) = \lambda v(\hat{e}), \tag{13}$$

where we have introduced the linear operator $\mathcal{L}_q$ defined by

$$(\mathcal{L}_q v)(\hat{e}) = \mathbf{E}\|A\hat{e}\|^q v(A\hat{e}/\|A\hat{e}\|). \tag{14}$$

Comparing $\mathbf{E}\|X_n\|^q = u_n(x_0)$ with Eq. (2) gives the following:

Proposition 1. The generalized Lyapunov exponent $\ell(q)$ is the logarithm of the largest eigenvalue λ_1 of (13):

$$\ell(q) = \log \lambda_1. \tag{15}$$

Here we assume that the point of the spectrum with the largest modulus is an eigenvalue, λ_1. This can be guaranteed under certain assumptions. (See [8] where the eigenvalue problem (13) is studied in order to establish central limit and large-deviation results.) Note that since $\mathcal{L}_q$ maps positive functions to positive functions, $\lambda_1 > 0$.

The characterization (15) of the generalized Lyapunov exponents is useful for a number of purposes. First, it gives a deterministic method for finding $\ell(q)$ by solving an eigenvalue problem, analytically in simple cases and numerically in less simple cases. Second, the eigenvalue formulation can be used to examine the convergence of $\log \mathbf{E}\|X_N\|^q$ as $N \to \infty$ and conclude, for instance, that the convergence is typically exponential, with an error proportional to $|\lambda_2/\lambda_1|^N$, where λ_2 is the second largest eigenvalue of Eq. (13). Third, the eigenvalue formulation makes it possible to establish some useful properties of $\ell(q)$, which we now discuss.

In Appendix A, we show that the adjoint of $\mathcal{L}_q$ is the operator $\tilde{\mathcal{L}}_{-q-d}$, where $\tilde{\mathcal{L}}_q$ is defined as $\mathcal{L}_q$ in (13), but with the matrix A replaced by $A^{-1}/|\det A|^{1/q}$. We then have the following useful relationships between generalized Lyapunov exponents of the matrices A, $A^{-1}/|\det A|^{1/q}$ and A^{-1}.

Proposition 2. Let

$$\tilde{\ell}(q) = \lim_{N\to\infty} \frac{1}{N}\log \mathbf{E}\frac{\|A_N^{-1} \cdots A_1^{-1}\|^q}{|\det(A_N \cdots A_1)|} \quad \text{and}$$

$$\ell^-(q) = \lim_{N\to\infty} \frac{1}{N}\log \mathbf{E}\|A_N^{-1} \cdots A_1^{-1}\|^q. \tag{16}$$

Then,

(i) $\ell(q)=\tilde{\ell}(-q-d)$,
(ii) $\ell(q)=\ell^-(-q-d)$ if the matrices A_n satisfy $\det A_n = 1$,
(iii) $\ell(q)=\ell(-q-d)$ if the matrices A_n are symplectic.

Note that it follows from the first property that

$$\ell(-d) = \log \mathbf{E}|\det A|^{-1} \tag{17}$$

which extends the well-known observation that [6,32]

$$\ell(-d) = 0 \quad \text{if} \quad \det A = 1. \tag{18}$$

The properties in proposition 2 are established in Appendix A. They are useful in practice: because numerical methods for the estimation of the Lyapunov exponents are more accurate when $|q|$ is small, estimates for $\ell(q)$ with $q < -d$ can be obtained efficiently by evaluating $\tilde{\ell}(-q-d)$. Also, the replica method (described in Sec. IV B), which provides

estimates of $\ell(q)$ for q even and positive, can be used for some negative values of q when proposition 2 is exploited.

As a practical tool for the computation of generalized Lyapunov exponents, the eigenvalue problem (13) appears limited to small matrices with $d=2$ or $d=3$, because it requires the discretization of an operator acting on functions of $d-1$ variables. (See [14,16] and Sec. IV A below for some implementations with $d=2$.) In the next section we describe a Monte Carlo method that does not suffer from this limitation.

III. RESAMPLED MONTE CARLO

The simplest Monte Carlo method for the estimation of $\ell(q)$, which we term brute-force Monte Carlo, consists in computing the estimator

$$Z_N^{\mathrm{bf}} = \frac{1}{K}\sum_{k=1}^{K} \|A_N^{(k)} \cdots A_1^{(k)} x_0\|^q,$$

where the bracketed superscript indexes K independent realizations of the sequences of random matrices A_n. Clearly,

$$\mathbf{E} Z_N^{\mathrm{bf}} \asymp e^{N\ell(q)},$$

so $N^{-1}\log Z_N^{\mathrm{bf}}$ estimates $\ell(q)$. This method is hopelessly inefficient, however, unless $|q|$ is small. To see why, note that the variance of Z_N^{bf} is given by

$$\mathrm{var}\, Z_N^{\mathrm{bf}} = \frac{1}{K}\mathrm{var}\|X_N\|^q = \frac{1}{K}\{\mathbf{E}\|X_N\|^{2q} - (\mathbf{E}\|X_N\|^q)^2\} \sim \frac{c_{2q}e^{N\ell(2q)} - c_q e^{2N\ell(q)}}{K}.$$

The convexity of $\ell(q)$ then implies that $\exp[N\ell(2q)] \gg \exp[2N\ell(q)]$. So the variance of Z_N^{bf} is exponentially large in N, and a number of realizations $K \gg \exp\{N[\ell(2q) - 2\ell(q)]\}$ is in principle necessary for an accurate estimation of $\ell(q)$.

The inefficiency of the brute-force Monte Carlo estimate stems from the fact that for finite q, $\mathbf{E}\|X_N\|^q$ is dominated by rare realizations which are undersampled unless K is exponentially large. To remedy this, we can resample at each iteration so that the dominant contributions to $\mathbf{E}\|X_N\|^q$ are represented by more realizations; this is the main idea behind sequential importance sampling or "go-with-the-winners" strategies [28,29]. We describe a particularly simple algorithm for such a resampling strategy which we term resampled Monte Carlo (RMC).

A. Algorithm

Like the brute-force Monte Carlo, the algorithm relies on N iterations and K realizations, calculating $X_n^{(k)}$ for $n=1,\ldots,N$ and $k=1,\ldots,K$. The difference is that the realizations are dependent. Rather than $X_n^{(k)}$, it is convenient to use the corresponding unit vector

$$E_n^{(k)} = \frac{X_n^{(k)}}{\|X_n^{(k)}\|}.$$

Starting with $E_0^{(k)} = x_0$ for $k=1,\ldots,K$, the algorithm proceeds iteratively with two steps at each iteration n:

(i) Draw K random matrices $A_n^{(k)}$, and compute

$$\hat{E}_n^{(k)} = \frac{A_n^{(k)} E_{n-1}^{(k)}}{\|A_n^{(k)} E_{n-1}^{(k)}\|} \quad \text{and} \quad \alpha_n^{(k)} = \|A_n^{(k)} E_{n-1}^{(k)}\|^q. \tag{19}$$

(ii) Resample by letting

$$E_n^{(k)} = \hat{E}_n^{(J_n^{(k)})}. \tag{20}$$

Here, the $J_n^{(k)}$ are independent random variables taking values in $\{1,\cdots,K\}$, with

$$\mathbf{P}(J_n^{(k)} = j) = \frac{\alpha_n^{(j)}}{\beta_n}, \quad \text{where} \quad \beta_n = \sum_{k=1}^{K} \alpha_n^{(k)}. \tag{21}$$

The estimate of $\mathbf{E}\|X_N\|^q$ is then given by

$$Z_N = \frac{1}{K^N}\beta_1\beta_2\cdots\beta_N. \tag{22}$$

Note that the resampling step ensures that, at each iteration n, the weight of each realization in the estimate of $\mathbf{E}\|X_n\|^q$ is the same. Note also that the resampling is tailored to a specific value of q. Unlike in the brute-force Monte Carlo, where the same ensemble can be used to estimate $\ell(q)$ for a range of values of q, the RMC approach requires a new sampling for each value of q (although it may be possible to use the same sampling for a narrow enough range of q). In several applications, though, $\ell(q)$ is only required for a single value of q (e.g., [10,13,14]).

In Appendix B, we give a pseudocode for the RMC algorithm. This illustrates the simplicity of the algorithm, and should be useful for readers wishing to implement it in a specific programming language.

B. Analysis

To analyze the algorithm further, we note that the NK random matrices $A_n^{(k)}$ involved in the computation form K independent paths consisting of the N matrices that are multiplied in succession to obtain each $E_N^{(k)}$. These paths are

$$A_N^{(I_N^{(k)})}, A_{N-1}^{(I_{N-1}^{(k)})}, \ldots, A_1^{(I_1^{(k)})},$$

where the random variables $I_n^{(k)}, n=1,\cdots,N$ are determined by k and by the random variables $J_n^{(k)}$ according to

$$I_N^{(k)} = J_N^{(k)}, \quad I_{N-1}^{(k)} = J_{N-1}^{(I_N^{(k)})}, \quad I_{N-2}^{(k)} = J_{N-2}^{(I_{N-1}^{(k)})}, \ldots.$$

The factors $\alpha_n^{(\cdot)}$ that are computed along the path that yields $E_N^{(k)}$ are then

$$\alpha_n^{(I_n^{(k)})} = \frac{\|A_n^{(I_n^{(k)})} A_{n-1}^{(I_{n-1}^{(k)})} \cdots A_1^{(I_1^{(k)})} x_0\|^q}{\|A_{n-1}^{(I_{n-1}^{(k)})} A_{n-2}^{(I_{n-2}^{(k)})} \cdots A_1^{(I_1^{(k)})} x_0\|^q}, \quad n=1,\ldots,N. \tag{23}$$

Note that the distribution of the $I_n^{(k)}$ is that same as that of the $J_n^{(k)}$, since the distribution of the latter is independent of k; thus,

$$\mathbf{P}(I_n^{(k)}=j)=\frac{\alpha_n^{(j)}}{\beta_n}.$$

For a given realization of the matrices $A_n^{(k)}$ for $n=1,\cdots,N$ and $k=1,\cdots,K$, the probability of a particular path

$$A_N^{(j_N)},A_{N-1}^{(j_{N-1})},\cdots A_1^{(j_1)}$$

is then

$$\mathbf{P}(j_1,\ldots,j_N|A_n^{(k)})=\frac{\alpha_1^{(j_1)}\alpha_2^{(j_2)}\cdots\alpha_N^{(j_N)}}{\beta_1\beta_2\cdots\beta_N}, \tag{24}$$

where

$$\alpha_n^{(j_n)}=\frac{\|A_n^{(j_n)}\cdots A_1^{(j_1)}x_0\|^q}{\|A_{n-1}^{(j_{n-1})}\cdots A_1^{(j_1)}x_0\|^q} \quad \text{and} \quad \beta_n=\sum_{j_n=1}^{K}\alpha_n^{(j_n)}. \tag{25}$$

Here we abuse notation slightly and use the same symbol α to denote, in Eq. (23), a random variable that depend on both the $A_n^{(k)}$ and the $J_n^{(k)}$, and in Eq. (25) one that depend only on the $A_n^{(k)}$; the same abuse of notation is made for β.

To compute the expected value of functions f of Z_N produced by the algorithm, we note that the corresponding expectation, $\mathbf{E}'$ is a combination of the expectation $\mathbf{E}$ over the random matrices $A_n^{(k)}$ and of the expectation over the random variables $J_n^{(k)}$. Using Eq. (24) to compute the latter expectation leads to

$$\mathbf{E}'f(Z_N)=\mathbf{E}\sum_{j_1,\ldots,j_N=1}^{K}\frac{\alpha_1^{(j_1)}\alpha_2^{(j_2)}\cdots\alpha_N^{(j_N)}}{\beta_1\beta_2\cdots\beta_N}f(z_N), \tag{26}$$

where $z_N=\beta_1\beta_2\cdots\beta_N/K^N$, and the $\alpha_n^{(j_n)}$ and β_n defined as in Eq. (25).

Using Eq. (26), it is immediate to establish

Proposition 3. Z_N is an unbiased estimator for $\mathbf{E}\|X_N\|^q$:

$$\mathbf{E}'Z_N=\mathbf{E}\|X_N\|^q. \tag{27}$$

This follows from the computation

$$\begin{aligned}\mathbf{E}'Z_N&=\frac{1}{K^N}\mathbf{E}\sum_{j_1,\cdots j_N=1}^{K}\alpha_1^{(j_1)}\alpha_2^{(j_2)}\cdots\alpha_N^{(j_N)}\\&=\frac{1}{K^N}\mathbf{E}\sum_{j_1,\ldots,j_N=1}^{K}\|A_N^{(j_N)}\cdots A_1^{(j_1)}x_0\|^q=\mathbf{E}\|A_N\cdots A_1x_0\|^q\\&=\mathbf{E}\|X_N\|^q,\end{aligned}$$

which uses Eqs. (22), (25), and (26).

In order to estimate the error of Z_N, we obtain the following expression:

Proposition 4. The expected value of Z_N^2 is

$$\mathbf{E}'Z_N^2=\frac{1}{K^{2N}}\mathbf{E}\sum_{j_1,\ldots,j_N=1}^{K}\sum_{j_1',\cdots,j_N'=1}^{K}\|A_N^{(j_N)}\cdots A_1^{(j_1)}x_0\|^q\|A_N^{(j_N')}\cdots A_1^{(j_1')}x_0\|^q. \tag{28}$$

This is obtained from

$$\mathbf{E}'Z_N^2=\frac{1}{K^{2N}}\mathbf{E}\sum_{j_1,\ldots,j_N=1}^{K}\alpha_1^{(j_1)}\alpha_2^{(j_2)}\cdots\alpha_N^{(j_N)}\beta_1\beta_2\cdots\beta_N$$

on using the definition (25) of the β_n.

Expression (28) makes clear why the variance of Z_N is much smaller than that of Z_N^{bf}. Only K^N terms of the K^{2N} terms in Eq. (28) lead to contributions term proportional to $\exp[N\ell(2q)]$ (those for which $j_1=j_1',\cdots j_N=j_N'$) with all the others leading to much smaller contributions with, in particular, $[K(K-1)]^N$ proportional to $\exp[2N\ell(q)]$ (those for which $j_1\neq j_1',\cdots j_N\neq j_N'$). In contrast, in $\mathbf{E}(Z_N^{\mathrm{bf}})^2$, all the terms are proportional to $\exp(N\ell(2q))$.

The improvement can be evaluated explicitly in the scalar case $d=1$. Admittedly, this is an uninteresting case as far as the numerical evaluation of $\ell(q)$ is concerned, since Eq. (2) holds exactly for finite N, but it is instructive nonetheless. For $d=1$, the asymptotic relation (3) holds exactly for all N and with $c_q=1$. It follows that the terms in Eq. (28) can be evaluated explicitly: if $j_k=j_k'$ for l values of k and $j_k\neq j_k'$ for the remaining $N-l$ values,

$$\mathbf{E}\|A_N^{(j_N)}\cdots A_1^{(j_1)}x_0\|^q\|A_N^{(j_N')}\cdots A_1^{(j_1')}x_0\|^q=e^{l\ell(2q)+2(N-l)\ell(q)}$$

Since there are $\binom{N}{l}K^N(K-1)^{N-l}$ such terms, Eq. (28) becomes

$$\begin{aligned}\mathbf{E}'Z_N^2&=\frac{1}{K^N}\sum_{l=0}^{N}\binom{N}{l}(K-1)^{N-l}e^{l\ell(2q)+2(N-l)\ell(q)}\\&=\frac{1}{K^N}[e^{\ell(2q)}+(K-1)e^{2\ell(q)}]^N\end{aligned}$$

The variance is then

$$\mathrm{var}\,Z_N=\frac{[e^{2\ell(q)}(K+\gamma_q)]^N-K^Ne^{2N\ell(q)}}{K^N}\sim\frac{N}{K}\gamma_qe^{2N\ell(q)},$$

where we have introduced

$$\gamma_q=e^{\ell(2q)-2\ell(q)}-1. \tag{29}$$

Thus the relative variance of Z_N is

$$\frac{\mathrm{var}\,Z_N}{(\mathbf{E}Z_N)^2}\sim\frac{N\gamma_q}{K}, \tag{30}$$

and the Monte Carlo estimation of $\mathbf{E}\|X_N\|^q$ by Z_N requires only that $K\gg N$ rather than $K\gg\exp\{N[\ell(2q)-2\ell(q)]\}$ as is the case for the brute-force Monte Carlo. This drastic gain in computational efficiency is expected to apply also for matrices with $d>1$: the noncommutativity is likely to modify Eq. (30) only through the introduction of an N- and K-independent factor on the right-hand side.

Although we have found that the RMC algorithm performs very well for a broad range of random-matrix products, it is useful to have alternative methods of evaluating $\ell(q)$ at one's disposal. This provides independent checks for the RMC results or, in the case of asymptotic approximation for $|q|\gg1$, makes it possible to estimate $\ell(q)$ when the RMC

approach becomes unreliable. Such alternative methods are discussed in the next section.

IV. OTHER ESTIMATES

A. Solving the eigenvalue problem (13)

For $d=2$ or 3, it is practical to compute $\ell(q)$ as the largest eigenvalue of the eigenvalue problem (13) for functions v on S^{d-1}. Here, we describe an implementation for $d=2$. In this case, the unit vector $\hat{e}$ can be parameterized by an angle θ, and v can be expanded in a Fourier series, which we write as

$$v(\theta) = \mathrm{Re} \sum_{n=0}^{M-1} v_n e^{in\theta},$$

and truncate at some M. A straighforward discretization of the eigenvalue problem (13) is then obtained by collocation at points $\theta_m = 2\pi m/M, m=0,\cdots,M-1$. This leads to the generalized matrix eigenvalue problem

$$P\boldsymbol{v} = \lambda Q\boldsymbol{v}, \tag{31}$$

where $\boldsymbol{v}=(v_0,\ldots,v_{M-1})^{\mathrm{T}}$, and the $M\times M$ matrices P and Q have entries given by

$$P_{mn} = \mathbf{E}\|A\hat{e}(\theta_m)\|^q e^{in\Theta(\theta_m)} \quad \text{and} \quad Q_{mn} = e^{in\theta_m}, \tag{32}$$

where $\hat{e}(\theta_m)=(\cos\theta_m, \sin\theta_m)^{\mathrm{T}}$ and $\Theta(\theta_m)$ is defined by

$$\begin{pmatrix} \cos\Theta(\theta_m) \\ \sin\Theta(\theta_m) \end{pmatrix} = \frac{A\hat{e}(\theta_m)}{\|A\hat{e}(\theta_m)\|}. \tag{33}$$

The expectation in the definition of P can be computed using a Monte Carlo approach, and the eigenvalue problem solved using a standard technique.

B. Replica method for positive even q

A useful method, known as the replica trick (see [1], and reference therein), makes it possible to compute $\ell(q)$ for q positive and even by finding the largest eigenvalue of a $qd\times qd$ (deterministic) matrix. To see how this can be achieved, observe that the q-fold tensor product X_n with itself satisfies

$$X_n^{\otimes q} = A_n^{\otimes q} X_{n-1}^{\otimes q}, \tag{34}$$

where $A_n^{\otimes q}$ is the q-fold Kronecker product of A_n with itself. Taking the expectation then leads to

$$\mathbf{E}X_n^{\otimes q} = \mathbf{E}A_n^{\otimes q}\mathbf{E}X_{n-1}^{\otimes q}. \tag{35}$$

Therefore

$$\mathbf{E}X_N^{\otimes q} \asymp e^{\mu_q N} y, \tag{36}$$

where μ_q is the largest eigenvalue of the $qd\times qd$ matrix $\mathbf{E}A_n^{\otimes q}$, and $y\in\mathbb{R}^{qd}$ is the corresponding eigenvector. Since for $q>0$ even, $\|X_N\|^q$ is obtained from $X_N^{\otimes q}$ by contraction,

$$\ell(q) = \mu_q \quad \text{for} \quad q>0 \quad \text{even}. \tag{37}$$

The result extends to the case of odd q when the matrices A have only non-negative entries.

C. Large-$|q|$ asymptotics

For large $|q|$, numerical methods that involve taking expectations by sampling become inefficient, and it is useful to develop analytic or semianalytic methods that take advantage of $|q|\gg 1$ to provide an asymptotic estimate for $\ell(q)$. The eigenvalue problem (13) is a good starting point. Since the expectation is an integral over the random parameters that define the matrix ensemble, we can attempt to approximate this integral for $|q|\gg 1$ using Laplace's method. A dominant-balance argument suggests that the eigenfunction $v(\hat{e})$, which depends implicitly on q, should have the asymptotic WKB form

$$v(\hat{e}) \sim z(\hat{e}) e^{qw(\hat{e})}, \tag{38}$$

where w and z are independent of q. Substituting this into (13) gives

$$\lambda z(\hat{e}) e^{qw(\hat{e})} = \mathbf{E}e^{q(\log\|A\hat{e}\| + w(A\hat{e}/\|A\hat{e}\|))} z(A\hat{e}/\|A\hat{e}\|). \tag{39}$$

When the values of $\|A\|$ are bounded, the expectation on the right-hand side is dominated by the matrices maximizing the argument of the exponential (assuming a nonzero probability density for the maximizing matrices). Concentrating on the case $q>0$, this gives

$$w(\hat{e}) = \sup_A(\log\|A\hat{e}\| + w(A\hat{e}/\|A\hat{e}\|)) - \kappa, \tag{40}$$

for some constant κ, where the supremum is over the support of the probability measure of the random matrices. Note that w is defined up to the addition of an arbitrary constant. Equation (40) can be interpreted as a nonlinear eigenvalue problem, with w as the eigenfunction and κ as the eigenvalue. If this eigenvalue problem has a solution, the largest value of κ governs the rough asymptotics of λ_1 and hence the asymptotics of $\ell(q)$, with the result

$$\ell(q) \sim \kappa q. \tag{41}$$

Note that this behavior implies that the rate function $g(h)$ of the finite-time Lyapunov exponents has a vertical asymptote for $h=\kappa$. Therefore κ is also given by the maximum possible (largest) finite-time Lyapunov exponent:

$$\kappa = \lim_{N\to\infty} \sup_{A_1\cdots A_N} \frac{1}{N}\log\|A_N\cdots A_1\|. \tag{42}$$

It would of course be difficult to attempt to determine κ by sampling the right-hand side of this expression. In general, $\kappa\le\sup_A \log\|A\|$, with the equality holding only in special cases; see Appendix C.

The result (41) can be refined by noting that Laplace's method applied to Eq. (39) leads to the expectation of a term of the form $\exp(-q\langle A-A_*, A-A_*\rangle)$, where A_* is the maximizer in Eq. (40) and $\langle\cdot,\cdot\rangle$ is some scalar product (both A_* and $\langle\cdot,\cdot\rangle$ depend on $\hat{e}$). Carrying out the expectation

yields a factor $q^{-D/2}$, where D is the dimension of the support of the measure. It follows that

$$\ell(q) \sim \kappa q - \frac{D}{2}\log q + O(1). \tag{43}$$

This asymptotics implies that $g \sim -D\,\log(\kappa - h)/2$, which describes the manner in which $g(h)$ approaches the vertical asymptote at $h=\kappa$.

When $\|A\|$ is unbounded, the matrices A dominating the expectation in Eq. (39) are controlled by a balance between the argument of the exponential, which grows with $\|A\|$, and the probability density of A, which should decrease with $\|A\|$ if $\ell(q)$ is to be finite. This means that one needs to apply Laplace's method for movable maxima (e.g., [33]) and consider the q-dependent maximum of

$$\log\|A\hat{e}\| + w(A\hat{e}/\|A\hat{e}\|) + q^{-1}\,\log\,\pi(A), \tag{44}$$

where $\pi(A)$ is the probability density of A, and w depends on q. For instance, if $\pi(A)$ is Gaussian, this maximum corresponds to matrices A with $O(q^{1/2})$ entries, leading without further calculations to

$$\lambda \asymp e^{q\,\log\,q^{1/2}}, \quad \text{hence} \quad \ell(q) \sim \frac{q}{2}\log q. \tag{45}$$

Correspondingly, $g(h) \asymp e^h$ for $h \gg 1$.

V. EXAMPLES

A. Two-dimensional sine map

In studies of transport and mixing by complex fluid flows, numerous authors have used the random sine map proposed by Pierrehumbert [27] as a model of a completely chaotic, nondivergent flow. In two dimensions, this map is given by

$$x_{n+1} = x_n + a\,\sin(y_n + \phi_1), \quad y_{n+1} = y_n + b\,\sin(x_{n+1} + \phi_2), \tag{46}$$

where a and b are fixed parameters, and the random angles ϕ_1 and ϕ_2 are independent and uniformly distributed in $[0, 2\pi]$. The Jacobian matrix $\partial(x_{n+1}, y_{n+1})/\partial(x_n, y_n)$, whose statistics are independent of (x_n, y_n), is given at (0,0) by

$$A = \begin{bmatrix} 1 & a\,\cos\,\phi_1 \\ b\,\cos(\sin\,\phi_1 + \phi_2) & 1 + ab\,\cos\,\phi_1\,\cos(\sin\,\phi_1 + \phi_2) \end{bmatrix}. \tag{47}$$

It satisfies $\det A = 1$ and hence, since $d=2$, is symplectic.

The generalized Lyapunov exponents corresponding to the ensemble of matrices A generated by ϕ_1 and ϕ_2 characterize the separation of nearby particle in the sine flow. Remarkably, their knowledge makes it possible to predict, in some cases at least, the rate of decrease in the variance of a passive scalar released in the flow [10,11,14,19]. Specifically, this rate is given by $g(0)$ and, in view of the Legendre duality of $g(h)$ and $\ell(q)$, by $-\ell(q_*)$, where q_* is such that $\ell'(q_*) = 0$. Because of property (iii) of proposition 2, $q_* = -1$.

In the literature, $\ell(-1)$ has been evaluated using brute force Monte Carlo [12,13] and solving the eigenvalue problem [14,16]. Here we apply the algorithm of Sec. III to demonstrate its efficiency. In Fig. 1, we compare $\ell(q)$ obtained for $a=b=\pi$ using different numerical methods: brute force Monte Carlo, RMC, and numerical solution of the eigenvalue problem using 128 Fourier modes and 128 collocation points. For the latter two methods, we have used a relatively small ensemble, with $K=1000$, while for the brute force computation we have used the much larger $K=10^5$. The number of matrix multiplication N was taken as 100 for the RMC but only $N=50$ for the brute-force computation which is restricted to moderate values of N. Also shown are the very reliable estimates obtained for $q=2$, 4 and 6 using the replica method. The figure illustrates how impractical the brute force computation is to estimate $\ell(q)$ for, say, $q>2$ and $q<0$. The other methods, by contrast, provide good estimates for a wide range of q. Based on the comparison with the replica estimate, the RMC algorithm, which for the parameters chosen is the faster by a factor of about 5, appears to be the more accurate method. The inset in the figure zooms on the range $q \in [-2, 0]$ to emphasize the substantial differences in the estimates in that region leading, in particular, the inaccuracy in the estimates of $\ell(-1)$ needed for decay-rate predictions in the passive-scalar problem. In this regard, we note that a sequence of 500 RMC computations gives the average and standard deviation $\ell(-1) = -0.5916 \pm 0.0056$.

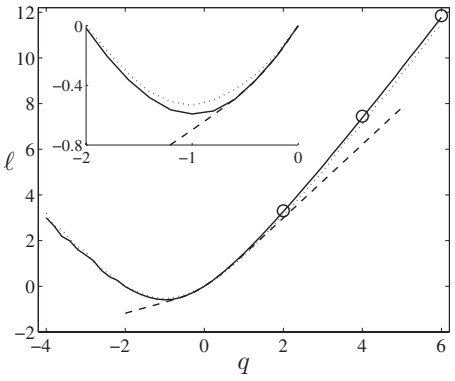

FIG. 1. Generalized Lyapunov exponents of the product of the random matrices (47) with $a=b=\pi$. The result of the RMC method (solid line) are compared with those of the brute-force Monte Carlo method (dashed line), and of the numerical solution of the eigenvalue problem (dotted line). The inset displays a close up of the region $-2 \le q \le 0$.

We have used the example of the two-dimensional sine flow with $q=-1$ to assess the dependence of the variance of the RMC estimate Z_N on the number of realizations K and on the number of iterations N. We have estimated this variance by performing 500 computations of Z_N for 9 combinations of the parameters K and N. The results are reported in Table I. Unsurprisingly, the sample variance scales roughly like $1/K$; more interestingly it also scales like N in agreement with the behavior (30) obtained in the scalar case. The behavior (30) can be tested further: since $\ell(-2)=0$, $\gamma_{-1} = \exp[-2\ell(-1)] - 1 \approx 2.26$, which compares reasonably well with the various estimates of $\mathrm{var}\,Z_N/(\mathbf{E}Z_N)^2$ that can be obtained from Table I.

Returning to Fig. 1, we note that the estimates of $\ell(q)$ appear less accurate for negative q unless $|q| \lesssim 1$; this can

TABLE I. Estimation of the normalized variance var $Z_N/(\mathbf{E}Z_N)^2$ for the RMC method applied to the matrices (47) for $q=-1$ and for different values of the number of realizations K and number of iterations N.

$K=$	500	1000	2000
$N=20$	0.12	0.052	0.027
$N=40$	0.24	0.12	0.059
$N=80$	0.62	0.23	0.10

easily be remedied, however, by using the third property in proposition 2, namely, $\ell(q)=\ell(-q-2)$, so that the only negative range that needs to be considered is $q\in[-1,0]$.

The estimation of $\ell(q)$ is truly challenging for large q. Here, we briefly consider it for the matrices (47) in order to assess both the reliability of the RMC method, and the asymptotic estimate (43). Figure 2 shows $\ell(q)$ for the matrices (47) with $a=b=\pi$, and with $a=\pi$ and $b=\pi/8$. In both cases, $\ell(q)$ can be approximated according to Eq. (43) with $D=2$ (since the matrices are defined by 2 random angles ϕ_1 and ϕ_2). The value of κ should be derived by solving Eq. (40). The case $a=b$ is special, however. It can be verified in this case that the maximum of $\|A\hat{e}\|$ is achieved for matrices A and unit vectors $\hat{e}$ such that $A\hat{e}/\|A\hat{e}\|=\hat{e}$. As a consequence, we have that

$$\kappa=\log\sup_A\|A\| \quad \text{for} \quad a=b. \tag{48}$$

This result, which holds for any matrix ensemble such that $A\hat{e}/\|A\hat{e}\|=\hat{e}$ for A and $\hat{e}$ maximizing $\|A\hat{e}\|$, is established in Appendix C. It enables a simple evaluation of κ when $a=b$, giving $\kappa=2.467$. The corresponding asymptotic estimate Eq. (43) is compared in Fig. 2 with the numerical estimates obtained using the RMC and replica methods. The $O(1)$ term in Eq. (43) is determined by matching the asymptotic and numerical results for the largest value of q on the figure. The figure demonstrates the validity of the asymptotic estimate; it also illustrates the reliability of the RMC method (used here

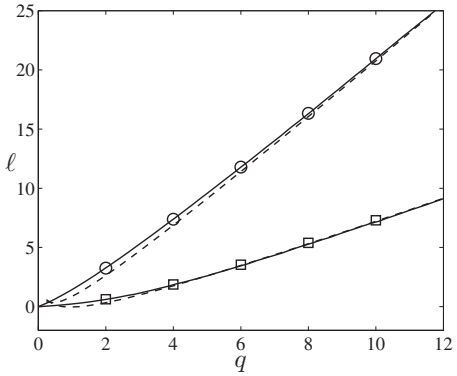

FIG. 2. Generalized Lyapunov exponents of the product of the random matrices (47) with $a=b=\pi$ (circles), and $a=\pi$ and $b=\pi/8$ (squares). The result of the RMC method (solid lines) are compared with the large-q asymptotic estimate (dashed lines), and the results of the replica method (symbols).

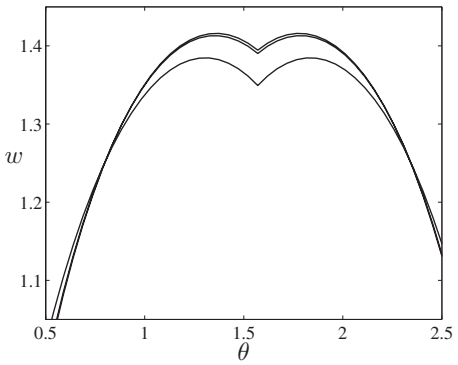

FIG. 3. Iterates $w^{(k)}(\theta)$ for $k=1,2,3$ in the numerical solution of the problem (40) determining κ for the matrices (47) with $a=\pi$ and $b=\pi/8$.

with an ensemble of $K=1000$ matrices), which provides accurate estimates of $\ell(q)$ for q as large as 12, at least for matrices considered here.

The simple result (48) is very special. In general, when $a\neq b$, the right-hand side of Eq. (48) is a strict upper bound for κ. There is then no explicit expression for κ, and the problem (40) must be solved for both κ and $w(\hat{e})$. We have implemented a numerical solution of this problem for the matrices (47). The implementation relies on an iteration: successive iterates $w^{(k)}, k=1,2,\cdots$, regarded as functions of the angle θ parameterizing $\hat{e}$, are represented using the truncated Fourier series

$$w^{(k)}(\theta)=\mathrm{Re}\sum_{n=1}^{M-1}w_n^{(k)}e^{in\theta},$$

from which the average ($n=0$) term is omitted in order to fix the arbitrary constant in the definition of w. The iteration scheme

$$w^{(k+1)}(\theta_m)+\kappa^{(k+1)}=\sup_A\left(\log\|A\hat{e}(\theta_m)\|+\sum_{n=1}^{M-1}w_n^{(k)}e^{in\Theta(\theta_m)}\right), \tag{49}$$

where $\Theta(\theta_m)$ is defined in Eq. (33), determines $w^{(k+1)}$ on the grid points $\theta_m=2m\pi/M$, with $\kappa^{(k+1)}$ fixed using the condition of zero average for $w^{(k+1)}$. The supremum is evaluated numerically by finding the maximum over a finite number of matrices A obtained for values of ϕ_1 and ϕ_2 on a grid. An inverse Fourier transform then gives $w_n^{(k+1)}$, and the iteration can continue. Figure 3 shows the first three iterates of this method applied in the case $a=\pi$ and $b=\pi/8$. The functions $w^{(k)}(\theta)$ are defined for $\theta=[0,2\pi]$ and π-periodic; here we show an interval of θ around the maxima of these functions. The first iterate, corresponding to the lowest curve, is simply $w^{(1)}(\theta)=\sup_A\log\|A(\hat{e}(\theta))\|$. The next two iterates illustrate the rapid convergence of the method; after 4 iterations, convergence is achieved, and the estimate $\kappa=1.061$ is obtained; this is substantially less than $\log\sup_A\|A\|=1.385$. The validity of our asymptotic formula and evaluation of κ are confirmed by Fig. 2, which shows an excellent match between the asymptotic and numerical estimates of $\ell(q)$. A similar

match was found for other values of a and b.

B. Three-dimensional sine map

In order to explore matrices that are not symplectic but have determinant 1, we consider the stretching by the volume-preserving map of $\mathbb{R}^3$,

$$x_{n+1} = x_n + a\,\sin(y_n + \phi_1), \quad y_{n+1} = y_n + b\,\sin(z_n + \phi_2),$$
$$z_{n+1} = z_n + c\,\sin(x_{n+1} + \phi_3), \tag{50}$$

where the $\phi_j, j=1,2,3$ are independent uniformly distributed in $[0,2\pi]$. This map generalizes to three dimensions the two-dimensional alternating sine map (46). The corresponding Jacobian matrix at the origin is

$$A = \begin{bmatrix} 1 & a\cos\phi_1 & 0 \\ 0 & 1 & b\cos\phi_2 \\ c\cos(a\sin\phi_1 + \phi_3) & ac\cos\phi_1\cos(a\sin\phi_1 + \phi_3) & 1 \end{bmatrix}. \tag{51}$$

The results of several numerical computations with these matrices are displayed in Fig. 4. In the main panel, we show three different estimates of $\ell(q)$, all obtained using the RMC method with $N=100$. The first (solid line) applies the RMC method to the matrices A with an ensemble size $K=1000$; the second (dotted line) also uses the RMC method but with the much smaller ensemble size $K=100$. The results illustrate the difficulties that arise when evaluating numerically $\ell(q)$ for $q<0$: for $q\lesssim -2$ in this case, the numerical estimates appear to be very unreliable, and the situation does not improve much when the number of realizations is increased from $K=100$ to $K=1000$. The problem is easily remedied, however, using property (ii) of proposition 2: by applying the RMC algorithm to A^{-1} rather than to A, we estimate $\ell^-(q)$; this estimate, which proves accurate for $q\gtrsim -2$, then provides a reliable approximation for $\ell(q)$ with $q\lesssim -1$ since $\ell(q) = \ell^-(-q-3)$. The curve of $\ell^-(-q-3)$ is shown by the dashed curve in Fig. 4. The best estimate of $\ell(q)$ should be read as the dashed curved for $q\lesssim -1$ and the solid curve for $q\gtrsim -2$. For definiteness, one could choose the point $q=-d/2 = -3/2$ for the transition between the two approximations.

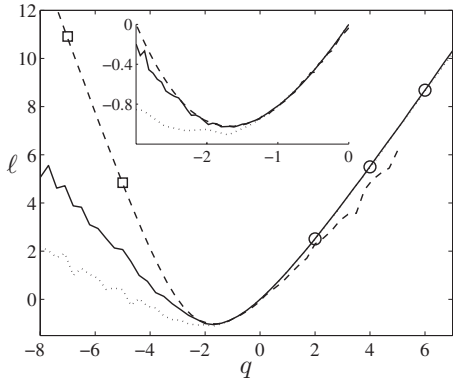

FIG. 4. Generalized Lyapunov exponents of the product of the random matrices (51) with $a=b=c=\pi$. The RMC estimates of $\ell(q)$ obtained for $N=100$ with an ensemble size $K=1000$ (solid line) and $K=100$ (dotted line) are compared with an estimate of $\ell^-(-q-3)$ obtained by applying the RMC method to A^{-1} with $K=1000$ (dashed line). The results of the replica method, applied to A (circles) and A^{-1} (squares) are also indicated. The solid and dashed curves in the inset, which displays a close up of the region $-3\leq q \leq 0$, have been computed using $K=5000$.

C. Gaussian matrices

As a last example, we consider the case of Gaussian matrices. When all the entries are independent $N(0,\sigma^2)$ variables, the statistics of $\|A\hat{e}\|$ are independent of $\hat{e}$, which leads to the explicit expression

$$\ell(q) = \log \mathbf{E}\|A\hat{e}\|^q = \frac{q}{2}\log(2\sigma^2) + \log\Gamma\left(\frac{q+d}{2}\right) - \log\Gamma\left(\frac{d}{2}\right), \tag{52}$$

for $q>-d$, with $\ell(q)=\infty$ for $q\leq -d$ [1]. No such explicit expressions are available when the entries are correlated, however, and $\ell(q)$ needs to be estimated numerically. Here, we examine the case of symmetric matrices with iid $N(0,\sigma^2)$

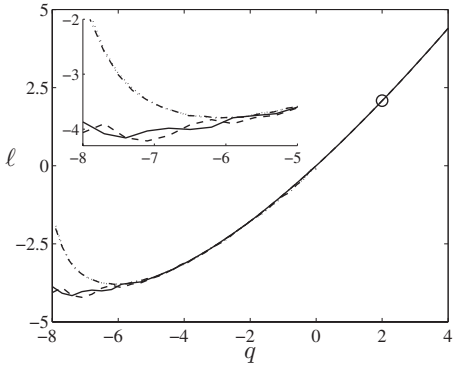

FIG. 5. Generalized Lyapunov exponents of the product of 8×8 symmetric matrices with independent, zero mean and variance-one Gaussian upper-triangular entries. The RMC estimates for $\ell(q)$ obtained with $K=2000$ (dashed line) and $K=4000$ (solid line) realizations are compared with the RMC estimates for $\tilde{\ell}(-q-8)$ obtained with $K=2000$ (dash-dotted line) and $K=4000$ (dotted line, almost indistinguishable from the dash-dotted line). The replica estimate for $q=2$ is indicated by the circle.

upper-triangular entries. As in the case of independent entries, $\ell(q)=\infty$ for $q<-d$, and so we can expect difficulties in estimating $\ell(q)$ for values of q slightly larger than $-d$, say for $q\lesssim -d/2$. It is indeed the case, as Fig. 5 demonstrates for $d=8$: the figure shows the direct estimates for $\ell(q)$ obtained using the RMC algorithm with 2000 and 4000 realizations. The differences between the results for $q\lesssim -d/2$ hints at their inaccuracy, as does an examination of the variance of these estimates. More obviously, the estimates fail to capture the rapid growth of $\ell(q)$ as $q\to -d$. Once again, we can invoke proposition 2 to remedy this problem, at least partially. Applying the RMC algorithm to the matrices $A^{-1}/|\det A|^{1/q}$ to estimate $\tilde{\ell}(q)$, then use the equality $\ell(q)=\tilde{\ell}(-q-d)$ gives a much better approximation for $\ell(q)$ in the range $-d\leq q\lesssim -d/2$. Thus, the approximation for $\ell(q)$ obtained in this manner with $K=2000$ and $K=4000$ are very close to one another and provide a satisfactory estimate for q close to $q=-d=-8$, though the divergence at $-d$ remains difficult to capture. Note that the large-q asymptotics (45) has been verified to apply to the symmetric Gaussian matrices with $d=8$ considered here; it is easy to check directly from Eq. (52) that it is satisfied for matrices with iid Gaussian entries.

VI. DISCUSSION

Motivated by the key role played by the large-deviation statistics of Lagrangian stretching in controlling several aspects of fluid mixing, this paper examines the generalized Lyapunov exponents of products of independent random matrices. Such products appear in this context when renewing flows, that is, sequences of simple iid steady flows, are used to model complex fluid motion. Products of random matrices appear of course in many other areas such as disordered media and wave localization.

The main aim of the paper is to present and test a reliable numerical procedure for the evaluation of the generalized Lyapunov exponents. The procedure proposed remedies the undersampling problem that affects the straightforward, brute-force Monte Carlo estimation by introducing a resampling step, which ensures that the variance of the estimate scales linearly with N, the (large) number of matrix multiplications, rather than exponentially. The algorithm chosen, which we term resampled Monte Carlo, is a particularly simple example of sequential importance sampling; its efficiency could be improved, e.g., by resampling every few iterations only, or by modifying the resampling method (see [28] for alternative approaches).

In particular, resampling methods can be devised on the model of the prune-enriched Rosenbluth method (PERM) method used in the simulation of polymer chains ([29,34,35] and references therein). In this method, the resampling is carried out only for realizations whose weight (i.e., contribution to the estimate of $\mathbf{E}\|X_n\|^q$ in our context) exceeds or falls below two chosen thresholds. If a weight exceeds the upper threshold, the realization is cloned a number of times, with the weight of each clone divided accordingly; if a weight falls below the lower threshold, the realization is either pruned with probability 1/2 or has its weight doubled. We have implemented a method of this type, using also a random pruning to keep the number of realizations constant. The results are similar to those obtained with the RMC method, but the PERM-like method proved somewhat slower in the examples we considered. However, we have made no attempt at optimizing the choice of the parameters that appear in the method (threshold values and number of clones). The PERM method has the advantage of potentially alleviating the problem of sample impoverishment, which occurs for large N when most of the realizations share the same early history. This problem does not appear to be serious for the computations of the generalized Lyapunov exponents of the matrix ensembles we treat in this paper, because convergence is achieved at moderately large N. Perhaps a more significant advantage of the PERM method in our context is that it can be implemented in a depth-first version, where the successive matrix multiplications are performed for a single realization at a time. The drastically reduced memory requirements of depth-first approaches make them suitable for the computations of the generalized Lyapunov exponents of very large matrices.

We have emphasized that the RMC method, and indeed all methods based on "go-with-the-winners" strategies have resampling strategies that are tailored to a particular value of q. When estimates of $\ell(q)$ are desired over a range of values of q, the computational efficiency could be improved by using the same ensemble, and hence the same resampling, for several values of q within a narrow interval, rather than a separate ensemble for each value of q. We do not pursue these improvements here, preferring to focus on the simple version of the algorithm, which can be easily analyzed and already provides a dramatic improvement compared with the brute-force Monte Carlo used by many authors.

In addition to providing a numerical method for the evaluation of the generalized Lyapunov exponents, the paper discusses some of their properties and, in particular, the relationship between the exponents associated with an ensemble of matrices A and those associated with the corresponding ensemble of matrices $A^{-1}/|\det A|^{1/q}$. This relationship is useful in practice to estimate $\ell(q)$ for negative q, when a direct application of our algorithm to the matrices A can be inaccurate. We also examine the asymptotic form of $\ell(q)$ for $|q|\gg 1$ and illustrate, on a specific example, how this form can be obtained by semianalytical means. Asymptotic results of this type usefully complement the direct numerical estimates of $\ell(q)$ which require very large samples as $|q|$ increases.

We conclude this paper by indicating a few possible extensions of the work reported. While the paper focuses on the largest generalized Lyapunov exponents, which encode the large-deviation statistics of the largest finite-N Lyapunov exponents, analogous statistics for lower Lyapunov exponents (proportional to the logarithm of smaller singular values of $A_N\cdots A_1$) are of interest. It would therefore be useful to develop an efficient numerical method to evaluate the corresponding generalized Lyapunov exponents. Work along these lines is currently in progress. Another useful extension concerns product of correlated matrices. The algorithm presented in Sec. III uses the independence of the matrices only to take a sequential approach, and hence it can be employed

for dependent matrices provided that the dependence is on the past only, that is, that A_n remain independent of A_k for $k>n$. More involved dependence would require a rethink of the algorithm. Since the literature on fluid mixing makes extensive use of white-in-time velocity fields as an alternative to renewing flows, it would also be desirable to develop methods for the efficient evaluation of generalized Lyapunov exponents in the context of linear stochastic differential equations (see [36] for recent analytical results). Finally, we note that the methods discussed in this paper apply to large matrices ($d \gg 1$) and so could be employed to study the large-deviation statistics of discretised infinite-dimensional systems as arise, for instance, in the problem of passive scalar decay [37].

ACKNOWLEDGMENTS

The author acknowledges the support of the Leverhulme Trust and thanks A. Tzella for useful discussions. The anonymous referees are thanked for their useful suggestions.

APPENDIX A: PROOF OF PROPOSITION 2

We first obtain the adjoint in $L_2(S^{d-1})$ of $\mathcal{L}_q$. Denoting by $\mathrm{d}\hat{e}$ the volume element on S^{d-1}, we consider two arbitrary functions $v(\hat{e})$ and $w(\hat{e})$ and compute

$$\int_{S^{d-1}} w(\hat{e})\mathcal{L}_q v(\hat{e})\mathrm{d}\hat{e}$$

$$= \mathbf{E}\int_{S^{d-1}} \|A\hat{e}\|^q w(\hat{e}) v(A\hat{e}/\|A\hat{e}\|)\mathrm{d}\hat{e}$$

$$= \mathbf{E}\int_{S^{d-1}} \|A^{-1}\hat{e}'\|^{-q-d} w(A^{-1}\hat{e}'/\|A^{-1}\hat{e}\|) v(\hat{e}')\mathrm{d}\hat{e}'/|\det A|,$$

where we have changed integration variable from $\hat{e}$ to $\hat{e}' = A\hat{e}/\|A\hat{e}\|$ and used that $\mathrm{d}\hat{e} = \|A\hat{e}\|^d \mathrm{d}\hat{e}'/|\det A| = \|A^{-1}\hat{e}'\|^{-d}\mathrm{d}\hat{e}'/|\det A|$ (cf. [32]). This gives the adjoint of $\mathcal{L}_q$ as

$$\mathcal{L}_q^\dagger = \widetilde{\mathcal{L}}_{-q-d}, \tag{A1}$$

where the operator $\widetilde{\mathcal{L}}_q$ is defined by

$$(\widetilde{\mathcal{L}}_q v)(\hat{e}) = \mathbf{E}\frac{\|A^{-1}\hat{e}\|^q}{|\det A|} v(A^{-1}\hat{e}/\|A^{-1}\hat{e}\|). \tag{A2}$$

Note that $\widetilde{\mathcal{L}}_q$ is equivalent to $\mathcal{L}_q$, with the matrices A replaced by $A^{-1}/|\det A|^{1/q}$. Thus, according to proposition 1, $\widetilde{\ell}(q)$ defined in Eq. (16) is the logarithm of the largest eigenvalue of $\widetilde{\mathcal{L}}_q$. The first part of proposition 2 follows from the fact that $\mathcal{L}_q$ and $\mathcal{L}_q^\dagger = \widetilde{\mathcal{L}}_{-q-d}$ have the same spectrum. The second part is a particular case of the first for $\det A = 1$.

We establish the third part of the proposition by showing that

$$\ell(q) = \ell^-(q) \quad \text{for symplectic matrices.}$$

Recall first that the matrices A are symplectic if d is even, and

$$A^{\mathrm{T}}\mathbb{J}A = \mathbb{J}, \quad \text{where} \quad \mathbb{J} = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix},$$

with $\mathbb{I}$ the $d/2 \times d/2$ identity matrix. Now, we define the operator $\mathcal{J}$ acting on functions on S^{d-1} according to

$$(\mathcal{J}v)(\hat{e}) = v(\mathbb{J}\hat{e})$$

and compute

$$\begin{aligned}(\mathcal{J}\mathcal{L}_q v)(\hat{e}) &= \mathbf{E}\|A\mathbb{J}\hat{e}\|^q v(A\mathbb{J}\hat{e}/\|A\mathbb{J}\hat{e}\|) \\ &= \mathbf{E}\|\mathbb{J}A^{-\mathrm{T}}\hat{e}\|^q v(\mathbb{J}A^{-\mathrm{T}}\hat{e}/\|\mathbb{J}A^{-\mathrm{T}}\hat{e}\|) \\ &= \mathbf{E}\|A^{-\mathrm{T}}\hat{e}\|^q v(\mathbb{J}A^{-\mathrm{T}}\hat{e}/\|A^{-\mathrm{T}}\hat{e}\|) \\ &= \mathbf{E}\|A^{-\mathrm{T}}\hat{e}\|^q (\mathcal{J}v)(A^{-\mathrm{T}}\hat{e}/\|A^{-\mathrm{T}}\hat{e}\|) = (\mathcal{L}_q^{-\mathrm{T}}\mathcal{J}v)(\hat{e}),\end{aligned}$$

where $\mathcal{L}_q^{-\mathrm{T}}$ is the analog of $\mathcal{L}_q$ with $A^{-\mathrm{T}}$ replacing A. This computation shows that $\mathcal{L}_q$ and $\mathcal{L}_q^{-\mathrm{T}}$ have the same spectrum, hence $\ell(q) = \ell^{-\mathrm{T}}(q)$. The results follows from observing that $\ell^{-\mathrm{T}}(q) = \ell^-(q)$ since they can be expressed as expectation of the largest singular values of $A_N^{-1}\cdots A_1^{-1}$ and $A_N^{-\mathrm{T}}\cdots A_1^{-\mathrm{T}}$, which coincide.

APPENDIX B: PSEUDOCODE FOR THE RMC METHOD

We give below a pseudocode for the RMC algorithm. The notation is as in Sec. III A except for the omission of the superscripts (k) and subscripts n when these are unnecessary for the numerical implementation. The variables $\gamma^{(k)}$ are introduced to perform the random resampling (20) and (21) using the uniformly distributed random variables ϵ.

```
fix q
E^(k) = (1,0,...,0)^T,  k = 1,...,K        (unit vectors in R^d)
for n = 1 to N                              (loop over iterations)
  for k = 1 to K                            (loop over realizations)
    draw random matrix A
    Ê^(k) = AE^(k)/‖AE^(k)‖
    α^(k) = ‖AE^(k)‖^q
  end
  γ^(k) = Σ_{l=1}^{k} α^(l),  k = 1,...,K
  β_n = γ^(K)
  for k = 1 to K                            (resampling)
    draw ε uniformly in [0, β_n]
    j = min{l ∈ {1,...,K} | γ^(l) − ε ≥ 0}
    E^(k) = Ê^(j)
  end
end
ℓ = (Σ_{n=1}^{N} log β_n)/N − log K         [estimate of ℓ(q)]
end
```

APPENDIX C: BOUNDS ON κ

Starting from Eq. (40), we write κ as

$$\kappa = \sup_A [\log\|A\hat{e}\| + w(A\hat{e}/\|A\hat{e}\|)] - w(\hat{e}), \tag{C1}$$

which holds for any $\hat{e}$. Taking $\hat{e}=\hat{e}_w$, where $\hat{e}_w$ maximizes w gives

$$\kappa \leq \sup_A \log\|A\hat{e}_w\|,$$

and in particular

$$\kappa \leq \sup_{\hat{e},A} \log\|A\hat{e}\| = \log \sup_{\hat{e},A} \|A\hat{e}\| = \log \sup_A \|A\|.$$

This is also obvious from the fact that $\lambda_1 \leq \sup_{\hat{e},A}\|A\hat{e}\|^q$. On the other hand, denoting by $\hat{e}_*$ and A_* the maximizers of $\|A\hat{e}\|$, and evaluating Eq. (C1) at $\hat{e}=\hat{e}_*$, we have that

$$\begin{aligned}\kappa &\geq \log\|A_*\hat{e}_*\| + w(A_*\hat{e}_*/\|A_*\hat{e}_*\|) - w(\hat{e}_*) \\ &= \log \sup_A \|A\| + w(A_*\hat{e}_*/\|A_*\hat{e}_*\|) - w(\hat{e}_*).\end{aligned}$$

A consequence is that $w(A_*\hat{e}_*/\|A_*\hat{e}_*\|) \leq w(\hat{e}_*)$. In the special case where

$$\frac{A_*\hat{e}_*}{\|A_*\hat{e}_*\|} = \hat{e}_*,$$

the two inequalities obtained imply that

$$\kappa = \log \sup_A \|A\|.$$

[1] A. Crisanti, G. Paladin, and A. Vulpiani, *Products of Random Matrices in Statistical Physics* (Springer, New York, 1993), p. 166.
[2] A. Crisanti, G. Paladin, and A. Vulpiani, J. Stat. Phys. **53**, 583 (1988).
[3] A. Crisanti, S. Nicolis, G. Paladin, and A. Vulpiani, J. Phys. A **23**, 3083 (1990).
[4] E. Ott, *Chaos in Dynamical Systems*, 2nd ed. (Cambridge University Press, Cambridge, England, 2002), p. 478.
[5] H. Furstenberg and H. Kesten, Ann. Math. Stat. **31**, 457 (1960).
[6] H. Furstenberg, Trans. Am. Math. Soc. **108**, 377 (1963).
[7] V. I. Oseledets, Trans. Moscow Math. Soc. **19**, 197 (1968).
[8] P. Bougerol and J. Lacroix, *Products of Random Matrices with Applications to Schrödinger Operators* (Birkhäuser, Boston, 1985), p. 283.
[9] G. Falkovich, Rev. Mod. Phys. **73**, 913 (2001).
[10] T. M. Antonsen, Z. Fan, E. Ott, and E. Garcia-Lopez, Phys. Fluids **8**, 3094 (1996).
[11] E. Balkovsky and A. Fouxon, Phys. Rev. E **60**, 4164 (1999).
[12] D. R. Fereday and P. H. Haynes, Phys. Fluids **16**, 4359 (2004).
[13] Y.-K. Tsang, T. M. Antonsen, and E. Ott, Phys. Rev. E **71**, 066301 (2005).
[14] P. H. Haynes and J. Vanneste, Phys. Fluids **17**, 097103 (2005).
[15] Z. Neufeld, C. López, E. Hernández-Garcia, and T. Tél, Phys. Rev. E **61**, 3857 (2000).
[16] A. Tzella and P. H. Haynes, Phys. Rev. E **81**, 016322 (2010).
[17] Y.-K. Tsang, Phys. Rev. E **80**, 026305 (2009).
[18] C. Reyl, T. M. Antonsen, and E. Ott, Physica D **111**, 202 (1998).
[19] Y.-K. Tsang, E. Ott, T. M. Antonsen, and P. N. Guzdar, Phys. Rev. E **71**, 066313 (2005).
[20] E. Balkovsky, G. Falkovich, and A. Fouxon, Phys. Rev. Lett. **86**, 2790 (2001).
[21] M. Chertkov, G. Falkovich, I. Kolokolov, and M. Vergassola, Phys. Rev. Lett. **83**, 4065 (1999).
[22] J. E. Cohen and C. M. Newman, Ann. Probab. **12**, 283 (1984).
[23] C. M. Newman, Commun. Math. Phys. **103**, 121 (1986).
[24] J. Cook and B. Derrida, J. Stat. Phys. **61**, 961 (1990).
[25] R. Mainieri, Phys. Rev. Lett. **68**, 1965 (1992).
[26] Z.-Q. Bai, J. Phys. A: Math. Theor. **40**, 8315 (2007).
[27] R. T. Pierrehumbert, Chaos, Solitons Fractals **4**, 1091 (1994).
[28] J. S. Liu, *Monte Carlo Strategies in Scientific Computing* (Springer–Verlag, Berlin, 2001).
[29] P. Grassberger, Comput. Phys. Commun. **147**, 64 (2002).
[30] C. Giardinà, J. Kurchan, and L. Peliti, Phys. Rev. Lett. **96**, 120603 (2006).
[31] V. Lecomte and J. Tailleur, J. Stat. Mech.: Theory Exp. (2007) 03004.
[32] Ya. B. Zel'dovich, A. A. Ruzmaikin, S. A. Molchanov, and D. D. Sokolov, J. Fluid Mech. **144**, 1 (1984).
[33] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (Springer, New York, 1999), p. 593.
[34] P. Grassberger, Phys. Rev. E **56**, 3682 (1997).
[35] H.-P. Hsu, V. Mehra, W. Nadler, and P. Grassberger, J. Chem. Phys. **118**, 444 (2003).
[36] R. Chetrite, J.-Y. Delannoy, and K. Gawędzki, J. Stat. Phys. **126**, 1165 (2007).
[37] J. Vanneste, Phys. Fluids **18**, 087108 (2006).