Eur. Phys. J. B **43**, 221–242 (2005)
DOI: 10.1140/epjb/e2005-00045-3

THE EUROPEAN
PHYSICAL JOURNAL B

Moment instabilities in multidimensional systems with noise

D.M. Wilkinson[a]

HP Labs, 1501 Page Mill Rd, Palo Alto, CA 94304, USA
and
Physics Department, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305-4060, USA

Received 28 July 2004 / Received in final form 23 November 2004
Published online 25 February 2005 –

Abstract. We present a systematic study of moment evolution in multidimensional stochastic difference systems, focusing on characterizing systems whose low-order moments diverge in the neighborhood of a stable fixed point. We consider systems with a simple, dominant eigenvalue and stationary, white noise. When the noise is small, we obtain general expressions for the approximate asymptotic distribution and moment Lyapunov exponents. In the case of larger noise, the second moment is calculated using a different approach, which gives an exact result for some types of noise. We analyze the dependence of the moments on the system's dimension, relevant system properties, the form of the noise, and the magnitude of the noise. We determine a critical value for noise strength, as a function of the unperturbed system's convergence rate, above which the second moment diverges and large fluctuations are likely. Analytical results are validated by numerical simulations. Finally, we present a short discussion of the extension of our results to the continuous time limit.

PACS. 02.50.Ey Stochastic processes – 02.50.Sk Multivariate analysis – 05.45.Ca Noise

1 Introduction

1.1 Motivation and previous work

The stability of fixed points in a multidimensional system is easily ascertained when the system is perfectly deterministic by using linear stability analysis [1]. Many real-world systems, however, are not perfectly deterministic because their interactions are subject to noise [2]. It is therefore of interest to consider the effect of a multiplicative noise term on a linearized system:

$$\mathbf{x}^t = (\mathbf{A} + \mathbf{B}^t)\mathbf{x}^{t-1}. \quad (1)$$

In this paper we analyze the effect of white, stationary mean zero noise in discrete systems. This type of noise has no effect on a system's stability in mean, because the expected value evolves exactly as if the system were unperturbed (Sect. 2.1). However, multiplicative noise processes cause fluctuations which can be large even if the fixed point is stable (Fig. 1), knocking the system out of the linear regime and coupling it to nonlinearities. Even for exact linear models, large fluctuations can cause long delays in convergence. An example of fluctuations in such a system is shown in Figure 1.

Fluctuations in a stochastic system are studied by way of the system's moments [2]. The pth moment of a multivariate system is simply the expected value of $|\mathbf{x}|^p$; large moments, especially the low order moments such as the second and third, indicate that a system attains large values with non-negligible probability [3]. For example, the system of Figure 1 has divergent moments for $p \geq 3$.

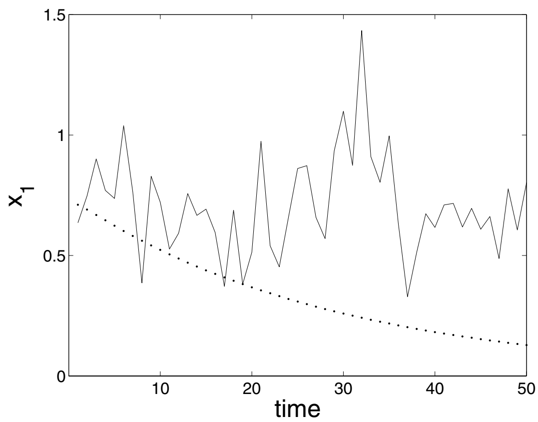

Fig. 1. Example of fluctuations in a linear system. The dotted line shows evolution of the first component of $\mathbf{x}$ without noise, and the solid line shows one instance of evolution with noise.

Multiplicative noise causes fluctuations because its effect is to cause the moments of a system to diverge, even when the system converges in mean [4,5]. In particular, divergent low-order moments in the neighborhood of a stable fixed point are likely to cause the large fluctuations

[a] e-mail: dennis.wilkinson@m4x.org

described above. The evolution of the moments is thus an important consideration in regards to fixed point stability in systems whose interactions are subject to noise.

Much of the previous work on stability in stochastic systems has focused on stability in mean. This subject is studied by way of a system's *Lyapunov exponents*, which are the asymptotic exponential rates of growth or decay of a system's eigenmodes, and in particular the *characteristic Lyapunov exponent*[1], which corresponds to the dominant mode. Calculation of the spectrum of Lyapunov exponents for multivariate systems is very difficult in general [6]. Stability in mean and calculation of the characteristic Lyapunov exponent or Lyapunov spectrum for discrete linear stochastic systems and random matrix products has been a major area of research in mathematics [7–12], control theory [13–16], engineering mechanics [17–19], and biology [5,20], among others. In physics, the Lyapunov spectrum is of great interest in stochastic [21] as well as chaotic [22] or disordered systems [23–25] which arise in many applications (e.g. [26,27]; see also [28]). The traditional approach to determining convergence in random systems is to use bounds, as in the above mathematics references; see also [29] for a physics approach and [30,31], for example, for continuous systems.

In the system (1) under consideration in this paper, the asymptotic behavior of the mean is easily determined, as we will see in Section 2.1, because only white noise is considered. However, the noise has a non-trivial effect on the moments of the distribution and it is of interest to determine their asymptotic behavior, as discussed above.

Moment evolution in stochastic systems is characterized by the *moment Lyapunov exponents*, which are the asymptotic exponential growth rates of the moments. There is relatively little previous work on moment Lyapunov exponents and moment stability in vector stochastic systems. [32] and [33] find small noise expansions for moment evolution in continuous 2-dimensional systems using a completely different approach from that of this paper. In the discrete case, research on the resistance of long, one-dimensional wires led to a study of moments of the trace of random matrix products [26,29,34]. The methods presented in these papers are specifically designed to calculate the trace of products of matrices subject to only one independent noise element, so that their application to the general problem presented in this paper would be quite difficult. In contrast, the methods presented here enable the calculation of the moment Lyapunov exponents for any unperturbed system subject to any arrangement of independent noises. Moreover, the analytic expressions found in this paper enable a general study of the dependence of the moments on system size, number of independent noise elements, and other parameters (see the overview of results, below).

[1] The terms *maximal Lyapunov exponent* or simply *Lyapunov exponent* are also commonly used to refer to the characteristic Lyapunov exponent.

1.2 Problem statement and notation

We are studying a system evolving according to the difference equation

$$\mathbf{x}^t = (\mathbf{A} + \mathbf{B}^t)\mathbf{x}^{t-1}, \tag{2}$$

or

$$\mathbf{x}^t = \left[\prod_{\tau=1}^{t} (\mathbf{A} + \mathbf{B}^\tau)\right] \mathbf{x}^0. \tag{3}$$

Here $\mathbf{x}$ is the system state, a vector of random variables and $\mathbf{B}^t$ is a matrix of white noise processes with mean 0. (That is, $\langle B_{ij}^t \rangle = 0$ and $\langle B_{ij}^t B_{ij}^{t'} \rangle \sim \delta_{tt'}$.) For simplicity, we consider only stationary processes wherein the B_{ij}^t are drawn from a distribution independent of t. The initial state $\mathbf{x}^0$ of the system is assumed to be fixed. The eigenvalues of the matrix $\mathbf{A}$ are λ_i; the largest eigenvalue λ_1 or simply λ is simple[2] and dominant, that is, $\lambda > |\lambda_i|$ for all $i \neq 1$. For simplicity we assume that $\lambda > 0$.

The system size is n. We define the mean of the A_{ij} to be a, and the variance to be σ_A^2. In the *mean value approximation*, $\mathbf{A} \approx a\mathbf{G}$ where $\mathbf{G}$ is the matrix whose elements are all 1. We will be diagonalizing $\mathbf{A}$ into the form $\mathbf{P\Lambda P}^{-1}$, where $\mathbf{\Lambda}$ is the diagonal matrix of eigenvalues of $\mathbf{A}$, and

$$\mathbf{u} = \mathbf{P}_{i1} = \text{normalized right eigenvector corresponding to } \lambda; \tag{4}$$

$$\mathbf{v} = (\mathbf{P}_{1i}^{-1})^T = \text{non-normalized left eigenvector corresponding to } \lambda, \tag{5}$$

so that $|\mathbf{u}| = 1$. Note that $\mathbf{v} \cdot \mathbf{u} = 1$, and therefore $|\mathbf{v}| \geq 1$.

A vector $\mathbf{x}^t$ will be said to converge in mean when $\langle x_i^t \rangle$ converges[3] to 0 for all i. As we will see in Section 2.1, convergence in mean corresponds to the case $\lambda < 1$ in agreement with the Oseledec theorem [22]. The system's fixed point is stable whenever the system converges in mean, because the initial state is irrelevant to convergence [35]. In the special case $\lambda = 1$, the system will converge in mean to a fixed non-zero vector, but every moment will diverge for any noise, however small, as we will see. Such a system is thus certain to eventually undergo large fluctuations and couple to nonlinearities. We will not devote any further discussion to this special case in particular.

We define the *pth moment* of the system to be $\langle |\mathbf{x}^t|^p \rangle$. Moment convergence can be elegantly expressed in terms of *moment Lyapunov exponents*, which are defined as

$$L_p = \lim_{t\to\infty} \frac{\log\langle |\mathbf{x}^t|^p \rangle}{t}. \tag{6}$$

[2] Simple eigenvalues have algebraic multiplicity 1 and thus only one associated eigenvector.

[3] The particular definition used for convergence is not of significant interest because the behavior of the mean in the equation considered in this paper is so simple. In keeping with the language of Lyapunov exponents and asymptotic exponential decay, we could say for example that $\langle x_i^t \rangle$ converges if there exists T, c_1, and $c_2 > 0$ such that for all $t > T$, $\langle x_i^t \rangle < c_1 e^{-c_2 t}$.

The existence of such a limit was demonstrated in [36]. The asymptotic behavior of the pth moment is then given by

$$\langle |\mathbf{x}^t|^p \rangle \sim e^{tL_p}$$

and the pth moment converges if

$$L_p < 0.$$

Finally, in the case that all the elements of the noise matrix have the same variance b^2, we define the *critical value* b_c^2 to be the level of noise above which the second moment diverges.

1.3 Overview of results

The central results of this paper are the approximations and exact expressions describing the evolution of moments of the system (2), in particular the second moment. The small noise case is treated first using a perturbation approach. This approach allows us to find an approximation for the system's moment Lyapunov exponents. For larger noises, an iteration technique is presented which gives both small and large noise results for the second moment. For certain types of noise, the iteration method allows us to calculate the second moment Lyapunov exponent exactly in any system. These are the first general analytic results for the moment Lyapunov exponents of discrete multivariate systems. They enable a study of the dependence of moment stability on parameters such as system size and magnitude and form of noise.

The analysis of this paper is valid in discrete systems with a simple, dominant eigenvalue. The eigenvalue requirement is satisfied by all nonnegative systems (see Appendix B) and many arbitrary systems. Nonnegative [37] and positive [38,39] discrete systems arise in Markov models, and the fields of biology, population models, economics (input-output models), finance, and cooperative problem solving, among others. Applications to arbitrary systems are too numerous to list.

Particular results of this paper are as follows. First, we show that in the small noise regime, the problem of finding the moment Lyapunov exponents of a multidimensional system reduces to the scalar case, which is trivial (Sects. 2.2, 4.1, 4.3). We thus obtain the expression

$$\langle |\mathbf{x}^t|^p \rangle = |\langle \mathbf{x}^t \rangle|^p e^{t\varepsilon^2 \frac{p(p-1)}{2} + O(\varepsilon^4)}, \tag{7}$$

where $\langle \mathbf{x}^t \rangle$ is the expected (unperturbed) value of the system at time t, and the parameter

$$\varepsilon^2 = \frac{\langle |\mathbf{v} \cdot \mathbf{B}\mathbf{u}|^2 \rangle}{\lambda^2} \tag{8}$$

depends on the noise as well as the dominant mode of the unperturbed system (recall the notation $\mathbf{u}$ and $\mathbf{v}$ for the right and left eigenvectors, respectively, of the dominant mode). The time superscript on the noise matrix $\mathbf{B}$ has been suppressed because the distribution its elements are drawn from does not depend on t. The value of ε^2 determines whether a given noise in a given system is small or large. Expressions for ε^2 are calculated for various forms of noise in Table 3. The above approximation is justified by simulation (Figs. 6–8) and its accuracy is discussed briefly in Section 4.5.

In the case of larger noise, the iteration technique of Section 5 presents a methodology for calculating the second moment Lyapunov exponent to any degree of accuracy in any system, provided the noise elements have the same variance. The exact value of the second moment Lyapunov exponent is expressed as the largest eigenvalue of a matrix and its accuracy is justified in the simulation of Figure 9.

It is shown that the results of the iteration technique agree with (7) for small noise, and with the $\lambda = 0$ limit for large noise. It is also shown that all results agree with the trivial scalar case discussed in Section 2.2 in the limit $n = 1$.

While the unperturbed value of the system depends only on the initial state and the dominant eigenvalue in the asymptotic limit, the moments depend on other properties of the system including the system size and the form of the noise. It is shown that

- the effect of a given level of noise can be magnified, in some cases greatly, if the dominant eigenvalue of the unperturbed system is ill-conditioned (Sects. 3, 4.3);
- the destabilizing effect of the noise is damped as the number of independent components of noise increases ("destructive interference" of independent noises) (Sect. 4.4, Fig. 7);
- the destructive interference of independent noises is maximized in the mean value limit (Sect. 1.2) and is mitigated by any deviation from this limit (Sect. 4.4, Fig. 8).
- large noise (Sect. 6.2), or small noise in systems with a very ill-conditioned dominant eigenvalue (Sect. 5.2, Fig. 9), almost certainly destabilizes the system.

We also present a discussion of the critical value for the noise variance above which the second moment diverges and fluctuations become a major consideration. For simplicity, this discussion is largely restricted to the case in which all the noise elements have the same variance. We obtain the following expression for the critical value:

$$b_c^2 = \frac{1}{n^k + \frac{f_v f_u \lambda^2}{1-\lambda^2}}, \tag{9}$$

where f_u and f_v are parameters related to $\mathbf{A}$ and to the type of noise considered, and where the factor k equals 1 if all the noise elements are independent and 2 if they are all correlated. The parameters f_u and f_v are very close to 1 in most systems. This expression is shown to be accurate in both the small and large noise cases, and is used to create a stability diagram for the system in Figures 11 and 12.

The dependence of the critical value on system parameters is discussed. We show that

- for small noise, the critical value depends weakly on the system size and type of noise considered (Sect. 6.4, Fig. 13);
- for large noise, the critical value depends strongly on the system size and type of noise (Sect. 6.4);

– the critical value provides a much more accurate indication of the level of noise below which the second moment converges than a simple bound on convergence (Sect. 6.5, Appendix D);
– for most convergent systems subject to small noise, the low-order moments diverge only if the unperturbed system converges slowly (Sect. 6.1).

This last statement is especially true for positive systems (Fig. 5). Note that systems with slow convergence may have other problems besides fluctuations due to noise, such as large transient behavior [40].

Finally, we consider the continuous limit and show that our results only immediately extend to this limit in special cases (Sects. 7.1, 7.2). The main difficulty in comparing the discrete and continuous cases is the difficulty in obtaining general analytical results for moment evolution in stochastic differential systems. This subject presents a very interesting area for future work.

1.4 Paper organization

The paper is organized as follows. In Section 2 we present simple preliminary results: asymptotic expressions for the system's expected value, moments in the scalar ($n = 1$) case, and second moment in the case that $\mathbf{A} = 0$. Section 3 discusses the important properties of the multivariate system. Moment evolution is calculated for multivariate systems in the small noise limit using a perturbation approach in Section 4, and the result is discussed. Section 5 uses an iteration approach to treat the second moment's evolution in the case of larger noise. The critical value of noise for second moment divergence is the subject of Section 6. The accuracy of the approximations is justified in numerical simulations throughout the paper. Finally, Section 7 presents a discussion of the continuous time limit.

2 Preliminaries

This section presents a calculation of the expected value of the system, as well as calculations of two limiting cases: a scalar stochastic system, and a noise-only multivariate system ($\mathbf{A} = \mathbf{0}$).

A discussion of the expected value of the system and its convergence properties is a necessary preliminary step in any study of the moments. The computation is trivial, as we show, because the noise is white with mean zero.

Calculation of the moments of a scalar system provides a framework which we will apply in the small noise limit of the multivariate case (Sect. 4). The scalar system also provides a demonstration of how multiplicative noise leads to moment divergence.

The noise only, $\mathbf{A} = \mathbf{0}$ multivariate system is a system in which the moments can be found exactly, yielding the zeroth order term for the large-noise limit (see Sect. 5.3). The calculations involved also provide a useful preview of those in Section 5.

All of the calculations in this section are quite straightforward and have very likely been presented, in whole or part, in some previous work. However, we did not find a specific reference with the exception of [16] which includes a cursory treatment of the scalar case.

2.1 Expected value and unperturbed system

The expected (average) state of the system and the state of the unperturbed system are equivalent since the noise is white with mean 0. White noise means that $\mathbf{x}^{t-1}$ and $\mathbf{B}^t$ are independent, so that

$$\langle \mathbf{x}^t \rangle = (\mathbf{A} + \langle \mathbf{B}^t \rangle)\langle \mathbf{x}^{t-1} \rangle = \mathbf{A}\langle \mathbf{x}^{t-1} \rangle,$$

since the mean of the B_{ij} is 0. Thus

$$\langle \mathbf{x}^t \rangle = \mathbf{A}^t \mathbf{x}^0 = \mathbf{x}^t_{\text{unperturbed}}.$$

In systems with a simple dominant eigenvalue λ, the asymptotic behavior of the unperturbed system is completely determined by λ [35]. For large t,

$$\mathbf{x}^t_{\text{unperturbed}} = \lambda^t (\mathbf{v} \cdot \mathbf{x}^0), \mathbf{u}, \tag{10}$$

where $\mathbf{u}$ and $\mathbf{v}$ are respectively the right and left eigenvector of the dominant mode. The moment Lyapunov exponents of the unperturbed system are thus simply

$$L^0_p = p \log \lambda. \tag{11}$$

The system will converge to 0 in mean for any initial conditions if $\lambda < 1$, and it will diverge if $\lambda_1 > 1$. We note that in the case of stochastic matrices with $\lambda = 1$, the above formula is accurate because λ is simple. We are not interested in the case in which $\mathbf{x}^0$ is orthogonal to $\mathbf{v}$.

2.2 Scalar stochastic system

In the case $n = 1$ it is not difficult to determine exact expressions for the low-order moments, and small-noise approximations for all moments. We go through a derivation here because this analysis will apply to the small noise multivariate case. For clarity we assume that the initial value x_0 of the system is positive. The system evolves according to

$$x_t = x_0 \prod_{\tau=1}^{t} (a + b_\tau),$$

where $\langle b_\tau \rangle = 0$ and $\langle b_\tau b_{\tau'} \rangle = b^2 \delta_{\tau\tau'}$. Notice that we express time as a subscript in this section, whereas in the multidimensional treatment time is a superscript.

2.2.1 Exact expressions

We have

$$\begin{aligned} \langle x^p_t \rangle &= x^p_0 \langle [a + b_\tau]^p \rangle^t \\ &= x^p_0 a^{pt} \left(\sum_{k=0}^{p} \binom{p}{k} \langle b_\tau / a \rangle^k \right)^t . \end{aligned} \tag{12}$$

In particular, $\langle x_t \rangle = x_0 a^t$ and $\langle x_t^2 \rangle = x_0^2(a^2 + b^2)^t$, so that

$$L_2 = \log\left(a^2 + b^2\right). \tag{13}$$

2.2.2 Approximate asymptotic distribution

In this subsection we assume small noise, that is, $|b_\tau/a| < 1$. This ensures that the system state is always nonnegative, allowing us to take logs, and also ensures that the moments of b_τ/a are well behaved. We have

$$\log\frac{x_t}{x_0} = t\log a + \sum_\tau s_\tau,$$

where

$$s_\tau = \log(1 + b_\tau/a).$$

The s_τ are i.i.d., so by the central limit theorem the sum is normal for large t with mean $t\mu_s$ and variance $t\sigma_s^2$, where μ_s and σ_s^2 are the mean and variance of the s_τ. The system is thus log-normally distributed in the asymptotic limit and its moments are given by

$$\langle x_t^p \rangle = x_0^p a^{pt} e^{pt\mu_s + p^2 t\sigma_s^2/2}. \tag{14}$$

Since we know that the first moment $\langle x_t \rangle = x_0 a^t$ is independent of the noise, we can conclude that $\mu_s = -\sigma_s^2/2$ and we have

$$\langle x_t^p \rangle = \langle x_t \rangle^p e^{-t\mu_s p(p-1)} \tag{15}$$

in the large t limit. Thus the pth moment Lyapunov exponent is given by

$$L_p = p\log a - \mu_s p(p-1),$$

or

$$L_p = L_p^0 - \mu_s p(p-1), \tag{16}$$

where $L_p^0 = p\log a$ is the pth moment Lyapunov exponent for the unperturbed system. Notice that $\mu_s < 0$ because the log function weights the negative values of b_t/a more heavily than the positive ones.

Expanding the log in the expression for $\mu_s = \langle \log(1 + b_\tau/a)\rangle$ we find:

$$\mu_s = -\sum_k \frac{\langle b_\tau/a\rangle^k}{k}(-1)^k.$$

The $\langle b_\tau/a\rangle^k$ term in the expansion must be $O(b/a)^k$ or smaller since b_τ/a can never exceed 1. Thus

$$L_p = L_p^0 + p(p-1)(b/a)^2/2 + O(b/a)^3. \tag{17}$$

The error is $O(b/a)^4$ if the noise is symmetric. In particular, for the second moment,

$$L_2 \approx L_p^0 + (b/a)^2 + O(b/a)^3, \tag{18}$$

in agreement with the exact value to second order.

The approximation and exact results for a scalar system are compared to simulation in Figure 2. This system converges in mean but has diverging moments for $p \geq 3$.

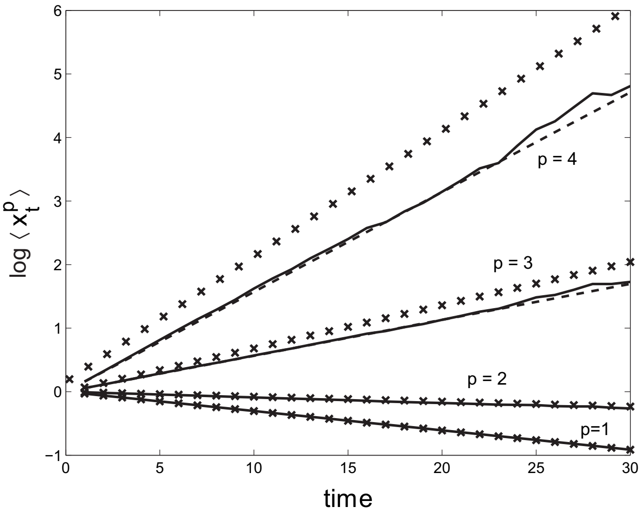

Fig. 2. Moment evolution in scalar system. The parameters for the simulation are $a = 0.97$ and "normal" noise with $b^2 = 0.05$ (noise values greater than a in absolute value were excluded). The solid lines show the average of the moments $\langle x_t^p \rangle$ for $p = 1, 2, 3$ and 4, over 10^6 runs. The dashed lines are the exact prediction (12) and are shown only for $p = 3$ and 4. The crosses are the approximation (17); the inaccuracy for $p = 4$ is due to the expansion of the log. The initial value was $x_0 = 1$ and noises larger than a were not allowed.

2.3 $\mathbf{A} = \mathbf{0}$ limit of multivariate case

Returning to the multivariate system, we consider as a first treatment the $\mathbf{A} = \mathbf{0}$ limit. The system's evolution in this limit is given by

$$\mathbf{x}^t = \left[\prod_{\tau=1}^{t} \mathbf{B}^\tau\right] \mathbf{x}^0. \tag{19}$$

The expected value of $\mathbf{x}$ is $\mathbf{0}$. The expression for moments contains two or more occurrences of each $\mathbf{B}^\tau$; the difficulty in its evaluation, and in general the difficulty of any multivariate system, is that the noise matrices do not commute. However, when $\mathbf{A} = \mathbf{0}$ and the noise is white, the sum may be evaluated explicitly. Its value depends on the type of noise considered. In this section we show the details of how to evaluate such a sum; in later sections such steps will be skipped.

In principle any moment of $\mathbf{x}$ could be calculated exactly, given a particular distribution for the noise elements. Here we restrict our calculation to the second moment for clarity and simplicity.

2.3.1 Independent noises

We first consider the case where all the noise elements vary independently. The matrices $\mathbf{B}^\tau$ do not commute so we must consider the full term by term expansion to evaluate the second moment:

$$\langle|\mathbf{x}^t|^2\rangle = \sum_i \sum_{j_1 j_2 \dots j_t} \sum_{k_1 k_2 \dots k_t} \langle B^1_{j_1 j_2} B^2_{j_2 j_3} \cdot \dots \cdot B^t_{j_t i} x^0_i B^1_{k_1 k_2} B^2_{k_2 k_3} \cdot \dots \cdot B^t_{k_t i} x^0_i \rangle, \quad (20)$$

where the expected value goes inside the sum because it is a linear function. All the elements of every $\mathbf{B}$ are independent, and we get a $\delta_{j_\tau k_\tau}$ for every $\tau = 1 \dots t$ when we sum on the k's. This gives

$$\langle|\mathbf{x}^t|^2\rangle = \sum_{j_1 j_2 \dots j_t i} \langle(B^1_{j_1 j_2})^2\rangle \langle(B^2_{j_2 j_3})^2\rangle \cdot \dots \cdot \langle(B^t_{j_t i})^2 (x^0_i)^2\rangle. \quad (21)$$

When all the noise elements have the same variance b^2, each of the t sums on $j_2, \dots, j_t$ and i simply gives a factor of nb^2. The remaining sum is just the norm squared of x^0, and we obtain

$$\langle|\mathbf{x}^t|^2\rangle = (nb^2)^t |\mathbf{x}^0|^2, \quad (22)$$

and

$$L_2^{\mathbf{A}=\mathbf{0}} = \log nb^2$$

for the second moment Lyapunov exponent.

If the noises did not all have the same variance, the result would be identical with b^2 replaced by an average variance

$$\bar{b}^2 = \frac{\sum_{ij} \langle(B_{ij})^2\rangle}{n^2}.$$

2.3.2 Correlated noises

When all the noises are correlated with the same variance the calculation is similar except that both the sum on the $\{j_\tau\}$ and the sum on the $\{k_\tau\}$ in equation (21) give a factor of n. We thus obtain

$$\langle|\mathbf{x}^t|^2\rangle = (n^2 b^2)^t |\mathbf{x}^0|^2. \quad (23)$$

We note that this result can be also found immediately by transforming to the eigenspace, since in this case $\mathbf{B}$ is proportional to the matrix $\mathbf{G}$ of all ones which has $\lambda = n$ and all other eigenvalues equal to 0. Second moment evolution in the noise-only case is shown in Figure 3.

The case of noises where only certain elements are correlated provides an intermediate case between independent and correlated noises. The expressions are complex and are left for future work.

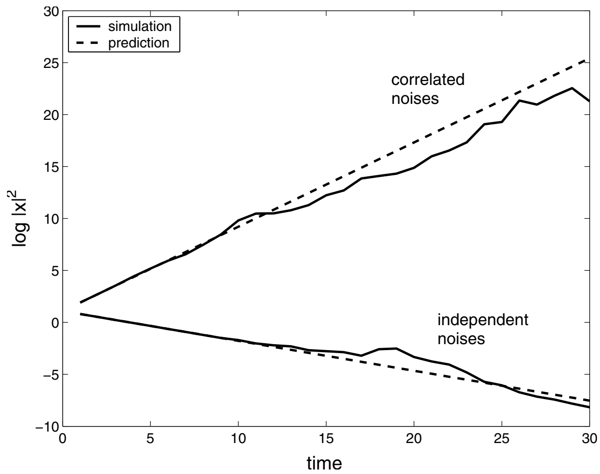

Fig. 3. Second moment evolution for $\mathbf{A} = \mathbf{0}$. The simulations show the value of $|\mathbf{x}|^2$ averaged over 1000 runs, in the independent noise case, and 100 000 runs for correlated noise. A large number of runs is necessary in the correlated noise case because of the divergent moments [4]. Here $n = 3$ and the elements of $\mathbf{B}$ were chosen from a normal distribution with variance 0.25. The simulations are compared to the predictions of (22) and (23).

3 Properties of multivariate stochastic systems

3.1 Heavy tail of distribution

As we saw in Section 2.2, scalar stochastic systems subjected to small multiplicative noise are log-normally distributed with parameters proportional to time, so the system moments evolve as $\exp[tp(\mu + p\sigma^2/2)]$. While μ is typically negative, for large p the positive $p\sigma^2/2$ term dominates and causes divergence. The effect of the multiplicative noise is thus to cause the system's pth moments to diverge for all p greater than some p_0.

The situation is more complicated in the multivariate case. It has been shown that any element of a product of t stationary random matrices is asymptotically log-normally distributed with parameters proportional to t [41,42]. Components of a multivariate stochastic difference system are thus linear combinations of log-normal variables with parameters proportional to t. This distribution may be very close to log-normal, as shown below, or not, as demonstrated in [29] where significant departure from the log-normal form was observed for certain parameter values in a particular system. This subject appears to be quite complex and would provide an interesting topic for future work. Nevertheless, as the results of this paper show, the moment Lyapunov exponents of a given system diverge as either the order of the moment increases or as the magnitude of the noise increases. This indicates that the actual distribution of the system states has a heavy tail even when it departs from the log-normal shape. In the particular case of small noise and simple dominant λ, the distribution of the elements of a multivariate system is very close to log-normal. This is demonstrated in the simulation of Figure 4.

3.2 Relevant properties of A

3.2.1 Simple dominant eigenvalue

In this paper we only consider systems with simple, dominant λ. Geometrically, the effect of $\mathbf{A}$ repeatedly acting on a vector is to bring that vector into the direction of $\mathbf{u}$

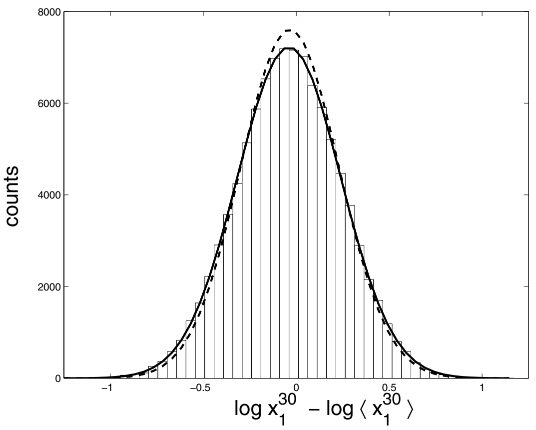

Fig. 4. Approximately log-normal distribution of the system state. The histogram plots the log of 100 000 instances of x_1^{30} in a positive system whose $\mathbf{A}$ is given in (64). The data were normalized by the expected value $\langle \mathbf{x}_1^{30} \rangle$. The solid line is a Matlab normal fit with $\mu = -0.0389$ and $\sigma = 0.2725$. The dashed line is the prediction of Section 4.3 and has $\mu = -0.0348$ and $\sigma = 0.2639$.

and to multiply its length repeatedly by λ. The behavior of unperturbed multivariate systems with a simple, dominant λ is thus equivalent to scalar systems in the asymptotic limit.

The requirement that λ be simple and dominant is met in all nonnegative systems of interest, as shown in Appendix B, so our treatment of nonnegative systems is comprehensive. Although many arbitrary systems meet this condition as well, some do not and we do not attempt to treat these cases. We also neglect systems with defective (non-diagonalizable) $\mathbf{A}$, which form a set of measure 0. This is reasonable because the nonzero elements of $\mathbf{A}$ are impossible to determine exactly in most applications.

3.2.2 Condition of λ

The effect of noise on a multivariate system, from a geometric perspective, is to perturb both the direction and length of the vector $\mathbf{x}$. Noise as a small perturbation means that a given noise matrix does not swing the $\mathbf{x}$ far from the direction of the eigenvector $\mathbf{u}$ or multiply $|\mathbf{x}|$ by a factor far from λ. In this regime, the dynamics are well approximated by the dynamics of a perturbed scalar system.

The regime of small noise, for multivariate systems, is determined not only by the size of the noise elements but also by the sensitivity of the system to perturbation. There exist matrices whose eigenvalues and eigenvectors are violently affected by even a small perturbation to the matrix elements [43,44]. To apply a perturbation treatment, we first need to know how much the dominant eigenvalue λ and its eigenvector $\mathbf{u}$ of the system are perturbed by a given level of noise.

The response of λ to noise is characterized by a quantity $\kappa(\lambda)$ called the *condition* of λ. When $\kappa(\lambda)$ is large, λ is said to be ill-conditioned, meaning that its response to a system perturbation is large with respect to the perturbation. Even a small noise can cause moment divergence in systems with an ill-conditioned λ. Conversely, when $\kappa(\lambda) = 1$, λ is said to be perfectly conditioned; its response to a perturbation of the matrix elements is the smallest possible and is on the order of the size of the perturbation. In systems with a well-conditioned λ, the perturbation approximation is applicable to relatively large noises.

The change in λ due to a small noise matrix $\mathbf{B}$ (small in the sense that $|\mathbf{B}| = \delta \ll 1$) is given by

$$\delta\lambda \approx \mathbf{v} \cdot (\mathbf{B}\mathbf{u}), \tag{24}$$

to first order in δ, where $\mathbf{u}$ and $\mathbf{v}$ are the right and left eigenvector corresponding to λ. Taking norms, we obtain the expression for the condition of λ in the case of normalized $\mathbf{u}$:

$$\kappa(\lambda) = |\mathbf{v}|. \tag{25}$$

It is clear that $\kappa \geq 1$ by the Schwartz inequality. The sensitivity of $\mathbf{u}$ to noise may also be calculated to first order [43] and depends on the condition of λ. It also depends on the gaps $\lambda - \lambda_i$ between the dominant eigenvalue and the others, and is therefore related to the accuracy of the approximation

$$\mathbf{A}^p \approx \lambda^p \mathbf{u}\mathbf{v}^T, \tag{26}$$

where T denotes transpose, obtained by neglecting all λ_i^p compared to λ^p. This approximation will be crucial in the manipulations of Section 5. In the limit that $\lambda_2 \to 0$, 26 is exact for all p and the sensitivity of $\mathbf{u}$ is minimized.

The level of noise which qualifies as a small perturbation must therefore depend on $\kappa(\lambda)$ and $\lambda - |\lambda_2|$, which it does, as we show is Section 4. Because the distance between the magnitudes of the dominant and second largest eigenvalues is important, we will introduce the terminology *eigenvalue gap* to refer to this quantity. The effect of a given level of noise is the smallest in *well-behaved* systems with a well-conditioned λ and small eigenvalue gap.

It is difficult to generally characterize the condition of λ and the eigenvalue gap in terms of more physical properties of the matrix $\mathbf{A}$. What we can say is summarized in Table 1. Systems close to the mean value approximation (recall $\mathbf{A} \approx a\mathbf{G}$ where $\mathbf{G}$ is a matrix of 1's in the mean value approximation; Sect. 1.2) are sure to be well-behaved; however, some systems far from the mean value approximation are also well-behaved, as shown in Figure 5 below as well as Figures 15 and 14.

The correlation between eigenvalue gap and condition number of λ is demonstrated in the scatter plots of Figure 5. As shown in the figure, the likelihood that a given system is well-behaved is larger for nonnegative matrices than for arbitrary matrices, and larger still for positive matrices. See Appendix C for further discussion.

3.2.3 Limits on average element size

Finally we note that in the case of nonnegative matrices, it is impossible to have a small λ if the elements of $\mathbf{A}$ are too

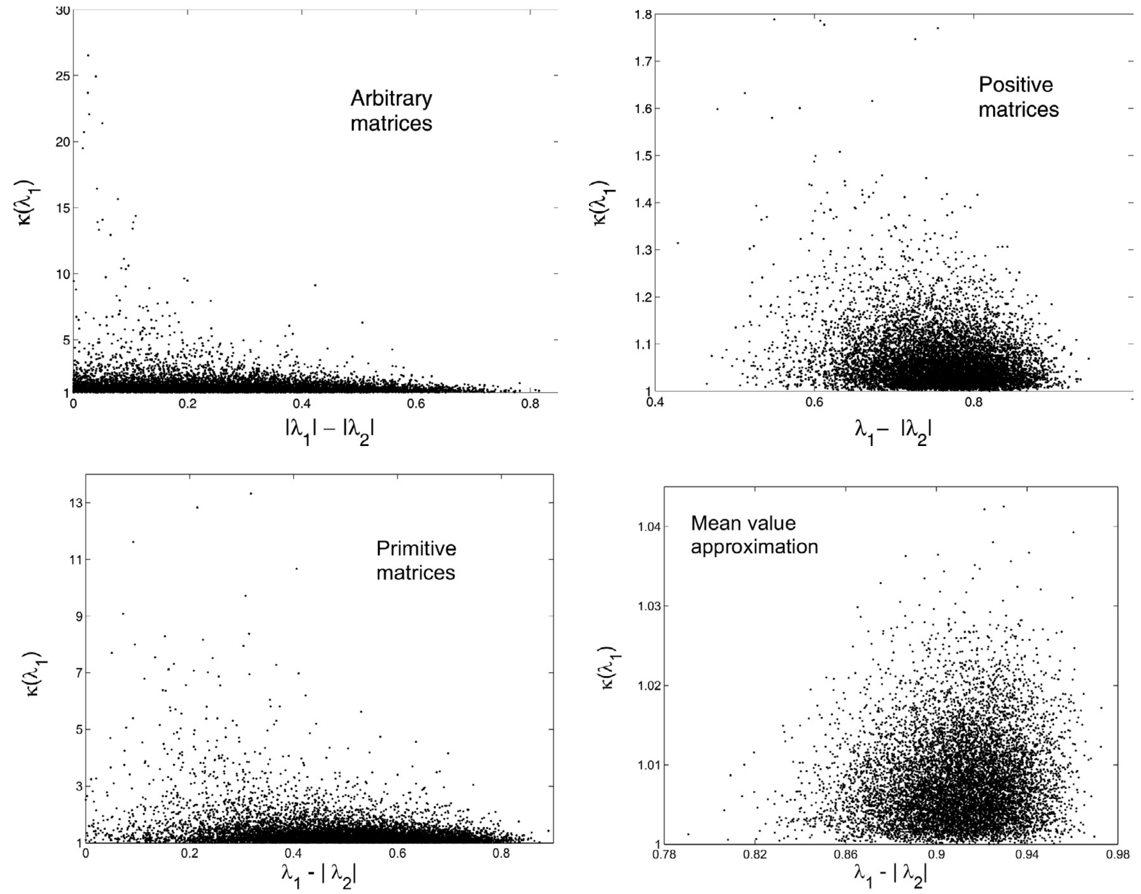

Fig. 5. Eigenvalue gap versus condition of λ in 10000 randomly generated 5×5 matrices. The matrices were generated from a normal distribution (top left), uniform distribution (top right), uniform distribution with probability 1/2 and 0 with probability 1/2 (bottom left) and uniform distribution with mean 0.2 and variance 0.02 (bottom right). For the arbitrary matrices, only those with real λ were accepted and the entries were normalized so that $\lambda = 1$. Note the difference in the regions plotted.

Table 1. Properties of some types of $\mathbf{A}$. Recall that $\mathbf{G}$ is the matrix of all 1's (mean value approximation) and σ_A^2 is the variance of the A_{ij}.

Matrix type	Eigenvalue gap $\lambda - \lvert\lambda_2\rvert$	Condition $\kappa(\lambda)$ of λ
$\mathbf{A} = a\mathbf{G}$	an	1
small σ_A^2	large	close to 1
normal $\mathbf{A}$	?	1
large σ_A^2	possibly small	possibly large

large. Many quite accurate bounds on the largest eigenvalue of nonnegative matrices exist (see [45] for a list); a relatively inaccurate but analytically tractable bound is the row sum bound, $\min_i(\sum_j A_{ij}) \le \lambda \le \max_i(\sum_j A_{ij})$. This estimate implies that on average we need to take

$$a < 1/n \tag{27}$$

to keep $\lambda < 1$ and ensure that the system converges in mean. This is exactly the asymptotic n result of [46], and the result we would obtain in the mean value approximation $\mathbf{A} \approx a\mathbf{G}$.

3.3 Types of noise for multivariate systems

For multivariate systems many different forms of noise are possible, distinguished by whether the elements are correlated and how large their relative variances are. In this paper we consider five cases which are analytically tractable and have relevance to physical systems. A summary of the types of noise considered is shown in Table 2.

For the correlation we consider three cases. *Uncorrelated* noise means that the elements of the noise matrix vary independently. *Totally correlated* noise means that all the noise elements vary in the same way at each time step. For symmetric systems, we consider *symmetrically correlated* noise.

For the variance we consider two possibilities. For *homogeneous* noise, the variance of every element is identical and equal to b^2. For *proportional* noise, the standard deviation of B_{ij} is proportional to A_{ij} by some factor q which we will take to be less than 1.

Table 2. Correlation rules for types of noise considered in this paper.

Noise type	Correlation rule
Uncorrelated homogeneous (UH)	$\langle B_{ij} B_{i'j'} \rangle = b^2 \delta_{ii'} \delta_{jj'}$
Symmetrically correlated homogeneous (SH)	$\langle B_{ij} B_{i'j'} \rangle = b^2 (\delta_{ii'} \delta_{jj'} + \delta_{ij'} \delta_{i'j})$
Totally correlated (T)	$\langle B_{ij} B_{i'j'} \rangle = b^2$
Uncorrelated proportional (UP)	$\langle B_{ij} B_{i'j'} \rangle = q^2 (A_{ij})^2 \delta_{ii'} \delta_{jj'}$
Symmetrically correlated proportional (SP)	$\langle B_{ij} B_{i'j'} \rangle = q^2 (A_{ij})^2 (\delta_{ii'} \delta_{jj'} + \delta_{ij'} \delta_{i'j})$

4 Small noise as a perturbation

In this section we determine approximate expressions for the moment Lyapunov exponents for multivariate systems subject to small noise using a perturbation treatment. We examine the dependence of the moment Lyapunov exponents on system properties, and discuss the accuracy of the approximation.

First let us reexpress the matrix product in (3):

$$\mathbf{x}^t = \left[\prod_{\tau=1}^{t} (\mathbf{A} + \mathbf{B}^\tau) \right] \mathbf{x}^0 \tag{28}$$

$$= \left[\sum_{\mathbf{Y}^t = \mathbf{A}, \mathbf{B}^t} \sum_{\mathbf{Y}^{t-1} = \mathbf{A}, \mathbf{B}^{t-1}} \sum_{\mathbf{Y}^1 = \mathbf{A}, \mathbf{B}^1} \mathbf{Y}^t \mathbf{Y}^{t-1} \cdots \mathbf{Y}^1 \right] \mathbf{x}^0$$

$$= \sum_{\text{each } \mathbf{Y}^\tau = \mathbf{A}, \mathbf{B}^\tau} \mathbf{Y}^t \mathbf{Y}^{t-1} \cdots \mathbf{Y}^1 \mathbf{x}^0, \tag{29}$$

meaning that each $\mathbf{Y}^\tau$ in the sum can be either $\mathbf{A}$ or $\mathbf{B}^\tau$, for $\tau = 1 \ldots t$. There are 2^t terms in the sum; each term is a vector.

4.1 Perturbation expansion

The perturbation expansion consists of considering only terms in (29) which have very few $\mathbf{B}$'s. For small noise, these terms make the only important contribution to the sum. Let us assume that this is so without justification, even before we define small noise.

The reason that this strategy simplifies the calculation is as follows. Consider the evolution in time of the length and direction of a single term of (29) with few $\mathbf{B}$'s. In the asymptotic limit, a typical term with few $\mathbf{B}$'s has long strings of consecutive $\mathbf{A}$'s broken by single occurrences of $\mathbf{B}$'s. As far as the direction of such a term, the long strings of $\mathbf{A}$'s act to bring it parallel to the dominant eigenvector $\mathbf{u}$ as previously mentioned (see (26)). When a $\mathbf{B}^\tau$ acts on the term, the term lies almost parallel to $\mathbf{u}$; even though the noise causes the term to point away from $\mathbf{u}$, the next string of $\mathbf{A}$'s brings it back to the direction of $\mathbf{u}$ before another noise term occurs. The action of $\mathbf{B}^\tau$ is thus independent of τ. As to the length, a string of p $\mathbf{A}$'s simply multiplies the term length by λ^p; and the $\mathbf{B}$'s multiply the length by some stationary random variable.

In a term with few $\mathbf{B}$'s, therefore, the position of the matrices in the sum (29) is unrelated to their net effect on the term. Thus the matrices in the sum can be replaced by scalars, and the matrix product (28) becomes a product of scalars. To illustrate this, consider a typical term for $t = 10$ with a $\mathbf{B}^\tau$ only in the $\tau = 6$ spot:

$$\mathbf{AAAAB}^6\mathbf{AAAAAx}^0 \approx (\lambda^4 \mathbf{u}\mathbf{v}^T)(\mathbf{B}^6)(\lambda^5 \mathbf{u}\mathbf{v}^T)\mathbf{x}^0$$
$$\approx \lambda^{10} \varepsilon_6 \mathbf{u}(\mathbf{v} \cdot \mathbf{x}^0),$$

where (26) was applied to the strings of consecutive $\mathbf{A}$s and where we define the random variable

$$\varepsilon_\tau = \frac{\mathbf{v} \cdot \mathbf{B}^\tau \mathbf{u}}{\lambda}. \tag{30}$$

Recall that $\mathbf{u}$ and $\mathbf{v}$ are the right and left eigenvectors, respectively, of the dominant mode, In general, a term of the sum (29) that has long strings of $\mathbf{A}$'s and m isolated $\mathbf{B}$'s $\{\mathbf{B}^{\tau_1}, \ldots, \mathbf{B}^{\tau_m}\}$ points in the direction of $\mathbf{u}$ and has length $\lambda^t(\varepsilon_{\tau_1} \cdot \ldots \cdot \varepsilon_{\tau_m})\mathbf{u}(\mathbf{v} \cdot \mathbf{x}^0)$. Such terms dominate the sum (29) (see Sect. 4.5) and so the system state is given approximately by

$$\mathbf{x}^t \approx \mathbf{u}(\mathbf{v} \cdot \mathbf{x}^0) \lambda^t \prod_{\tau=1}^{t} (1 + \varepsilon_\tau). \tag{31}$$

The random variables ε_τ are i.i.d. and satisfy $\langle \varepsilon_\tau \rangle = 0$; the moments depend on the form of the noise. Notice that the numerator of ε_τ is exactly equal to the first order change in λ due to a small perturbation to $\mathbf{A}$ (Eq. (24)) and thus closely related to the condition $\kappa(\lambda)$ (Eq. (25)). The eigenvalue gap and thus the sensitivity of $\mathbf{u}$ is implicitly involved in this expression from the application of (26).

4.2 Criterion for small noise

The simplest small noise criterion is

$$\langle \varepsilon_\tau^p \rangle \ll 1$$

for all p. This is a rather complicated condition since the calculation of all the moments can be difficult for some forms of noise. Instead we choose a more restrictive (triple) condition,

$$P(|\varepsilon_\tau| > 1) = 0$$
$$\langle \varepsilon_\tau^p \rangle \sim \varepsilon^{p'}, \quad p' \le p \tag{32}$$
$$\varepsilon^2 \ll 1. \tag{33}$$

Note that this requirement is not trivial as in the scalar case because the condition of λ can be large. Less restrictive conditions are possible but this will enable us to better understand the dynamics by taking logs and expanding in a power series in ε.

4.3 Moment evolution

Using the perturbation expansion of Section 4.1, we can obtain approximate expressions for the moments of a multivariate stochastic system. We do so by calculating the approximate moment Lyapunov exponents, proceeding from (31) exactly as in the scalar case of Section 2.2.2 with ε_τ playing the role of b_τ and λ the multidimensional analog of a.

We thus find

$$L_p \approx p\log\lambda - p(p-1)\langle\ln(1+\varepsilon_\tau)\rangle \approx L_p^0 + p(p-1)\frac{\varepsilon^2}{2} + O(\varepsilon^3), \tag{34}$$

where

$$L_p^0 = p\log\lambda$$

is the pth moment Lyapunov exponent for the unperturbed system and the error is $O(\varepsilon^4)$ if the noise is symmetric. The system moments are

$$\langle|\mathbf{x}^t|^p\rangle \approx |\mathbf{v}\cdot\mathbf{x}^0|^p e^{tL_p} \approx |\langle\mathbf{x}^t\rangle|^p e^{t\varepsilon^2\frac{p(p-1)}{2}+O(\varepsilon^3)}. \tag{35}$$

In particular,

$$L_2 \approx L_2^0 + \varepsilon^2 \tag{36}$$

to second degree in ε, and

$$\langle|\mathbf{x}^t|^2\rangle \approx |\langle\mathbf{x}^t\rangle|^2 e^{t\varepsilon^2}. \tag{37}$$

Notice that to this level of approximation, first moment (norm) convergence is not distinguishable from convergence in mean.

To proceed beyond these expressions we must evaluate

$$\varepsilon^2 = \frac{\langle|\mathbf{v}\cdot\mathbf{B}\mathbf{u}|^2\rangle}{\lambda^2},$$

which we cannot do without specifying the form of the noise. The values of ε^2 for the noises described in Section 3.3 are easily calculated and presented in Table 3. The accuracy of the above approximations for the moment evolution is demonstrated in Figure 6.

4.4 Dependence on system size

We can now explore the n dependence of the moments. Because the small noise case is important for applications, we present a detailed discussion of the size dependence based only on the expressions developed thus far. A different discussion of the n dependence for larger noises is presented in Section 6.4. For simplicity, we consider only the second moment in this section.

As we show, independently varying noises "interfere" with each other and diminish the effect of the noise, compared to the unperturbed system. Thus, the effect of the noise decreases as n increases in the case of uncorrelated noise. There is no n dependence to second order, however, in the case of totally correlated noise. Symmetrically correlated noise provides an intermediate case.

Table 3. Values of ε^2 for the types of noise defined in Section 3.3. Here UP means uncorrelated proportional noise, UH uncorrelated homogeneous, SP symmetrically correlated proportional, SH symmetrically correlated homogeneous, and T totally correlated. For proportional noise, f is a factor which depends on the distribution chosen; for example, $f = 1$ for normal noise and $f = 1/3$ for uniform noise. In the case of symmetrically correlated noise, a symmetric $\mathbf{A}$ is assumed.

Noise Type	ε^2
UH	$\frac{v^2b^2}{\lambda^2}$
SH	$\frac{2b^2}{\lambda^2}$
T	$\frac{b^2}{\lambda^2}(\Sigma_i v_i)^2(\Sigma_i u_i)^2$
UP	$\frac{fq^2}{\lambda^2}\sum_{ij} v_i^2(A_{ij})^2u_j^2$
SP	$\frac{fq^2}{\lambda^2}\sum_{ij}(A_{ij})^2[v_i^2u_j^2 + v_iv_ju_iu_j]$

We also show that as the system deviates from the mean value approximation and in particular becomes closer to diagonal, the destructive interference is decreased and the noise has a greater effect.

4.4.1 Mean value approximation

The n dependence of the noise effect is most clearly seen in the mean value approximation where $\mathbf{A} \approx a\mathbf{G}$. In this case, $v_i \approx u_i \approx 1/\sqrt{n}$ for all i, and $\lambda \approx na$. We consider homogeneous noise (all the noise elements have the same variance) with variance $b^2 = q^2a^2$, $q < 1$. Because we are in the mean value approximation, homogeneous noise is essentially equivalent to proportional noise (whence the notation using the constant of proportionality q).

Using (36) and Table 3, the values of L_2 for three types of homogeneous noise are easily computed in the mean value approximation and are shown in Table 4. These expressions are to be compared to the scalar case $L_2 \approx L_2^0 + q^2$ (Eq. (18)).

Note in particular how the noise effect (the q^2 term) is divided by a factor related to the number of independent elements of the noise. This destructive interference is not surprising when we consider why multiplicative noise processes generate the anomalously large events which make up the heavy tail of the log-normal distribution. The anomalous events result from a long sequence of large, positive noises [47]. When there are n^2 independent noises per time step, as opposed to 1, anomalous events are rarer. However, when all the elements of noise vary identically, the effect of the noise is the same in scalar and multidimensional systems. Note that symmetric noise provides an intermediate calculable case; there are $n(n+1)/2$ independent components in a symmetric noise.

The dependence of the noise effect on the number of independent elements of the noise is demonstrated in the simulation of Figure 7.

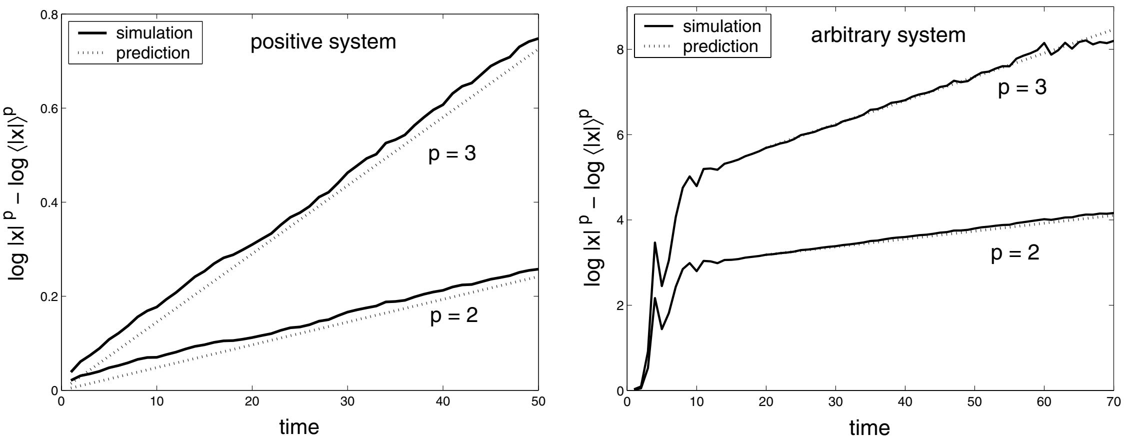

Fig. 6. Moment evolution in two systems. The log of the 2nd and 3rd moments of two randomly generated systems are shown. The plots are normalized by the expected value (unperturbed value) of the system and show only the noise part. The solid line is the average over 10 000 runs of the simulation, and the dotted line shows the analytic prediction of (35). At left is the positive system of (64) subjected to uniform UP noise with $q = 0.5$. Note that the asymptotic limit is reached almost immediately in this system. At right is an arbitrary system with simple dominant λ subjected to normal UH noise with $b = 0.1$. This system has large transient behavior before it settles in to its asymptotic limit around $t = 20$. The analytic prediction, which cannot account for the transient, has been artificially placed to demonstrate the asymptotic accuracy of the slope.

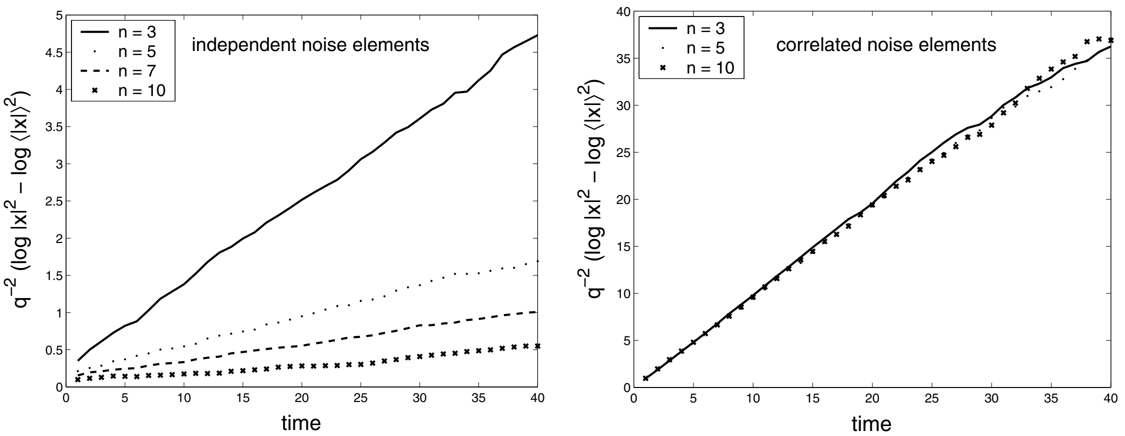

Fig. 7. Dependence of the noise effect on the number of independent noise elements. By Table 4 and equation (37), $\langle|\mathbf{x}|^2\rangle/|\langle\mathbf{x}\rangle|^2 = e^{tfq^2}$ in the asymptotic limit, where $f = 1/n^2$ for UH noise (all noise elements independent) and $f = 1$ for T noise (all noise elements correlated). Accordingly we plot $q^{-2}(\log|\mathbf{x}|^2 - \log|\langle\mathbf{x}\rangle|^2)$, averaged over 100 000 runs, for various values of n. On the left is the plot for UH noise, where the lines should have slope $1/n^2$; on the right is the plot for T noise where the slope should be 1 for any n. The agreement is excellent. The $\mathbf{A}$'s for these systems were randomly generated from uniform distributions with small variance. The initial state was a vector of 1's, and $q = 1/4$. Because we are in the mean value approximation, λ_2 is very small for these systems and the asymptotic limit for t begins almost immediately.

Table 4. Approximate value of L_2 for three types of homogeneous noise.

Noise Type	L_2
UH	$L_2 \approx L_2^0 + \frac{q^2}{n^2}$
SH	$L_2 \approx L_2^0 + \frac{q^2}{n^2/2}$
T	$L_2 \approx L_2^0 + q^2$

4.4.2 Deviation from the mean value approximation

We examine the effect of a deviation from the mean value approximation on L_2 in the case of uncorrelated proportional (UP) noise. The result is that the noise effect roughly increases the larger the deviation. This is reasonable because when the typical size of a noise element is proportional to the corresponding element of $\mathbf{A}$, the small entries contribute little to the interference effect of the independent noises.

In this subsection we assume that the approximation (26) is accurate for $p = 2$ so that $A_{ij}^2 \approx \lambda^2 u_i^2 v_j^2$. Recall that (26) is generally more accurate the closer $\mathbf{A}$ is to the mean value approximation, but it can be accurate even if the variance of the A_{ij} is large, as discussed above.

With the above approximation we have

$$L_2 \approx \ln\lambda^2 + q^2 w^4, \qquad (38)$$

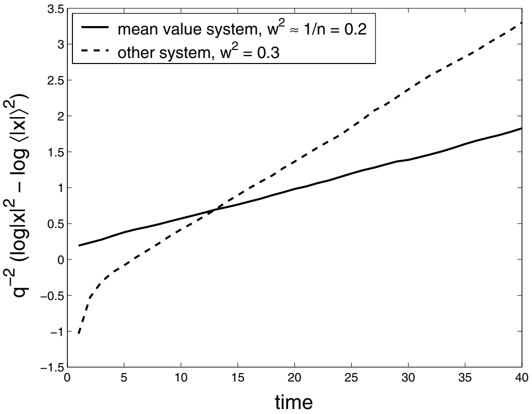

Fig. 8. Larger noise effect of a deviation from the mean value approximation. This figure is analogous to Figure 7 and compares the noise parts of the second moment $q^{-2}(\log|\mathbf{x}|^2 - \log|\langle\mathbf{x}\rangle|^2)$ of a mean value approximation (MVA) system and of a system with the same λ and $n = 5$ but in which the values of the entries of the matrix $\mathbf{A}$ are widely spread. UP noise with $q = 1/4$ was considered, so the slope is predicted to be w^4 by (38). The dashed line system has $w^2 = 0.2888$, $w^4 = 0.0834$ while the MVA system has $w^2 = 0.2002 \approx 1/n$ and $w^4 = 0.04$; a rough linear interpolation fit to the data shows an asymptotic slope 0.095 for the dashed system and 0.0385 for the MVA system, in agreement with (38). Note how the dashed system does not immediately reach its asymptotic limit; it has $|\lambda_2| \sim \lambda/2$ as opposed to $|\lambda_2| \ll 1$ for the MVA system.

where we define

$$w^2 = \sum_i v_i^2 u_i^2.$$

Comparing (38) to the mean value case, we see that the $1/n$ factor is replaced by w^2. This quantity satisfies

$$1/n \leq w^2 \leq 1$$

since $\mathbf{v} \cdot \mathbf{u} = 1$. The lower bound is achieved in the mean value case; the upper bound is achieved when $\mathbf{A}$ is diagonal. w^2 is thus a rough measure of the deviation of the A_{ij} from the mean value approximation; it generally increases as the variance of the A_{ij} increases.

Larger values of w^2 for deviations from the mean value approximation thus mean a larger noise effect. This effect is demonstrated in Figure 8.

4.4.3 Large n limit and homogeneous noise

When the noise is homogeneous we can apply results on spectral theory of matrices to study the n dependence in the large n limit without appealing to a mean value approximation. In the case of a symmetric matrix with entries drawn from a distribution with mean a and variance σ_a^2

$$\lambda = na + \sigma_a^2/a$$

on average [48]. For an arbitrary (asymmetric) matrix [46]

$$\lambda \sim na,$$

which is really just the mean value approximation.

Table 5. Large n dependence of arbitrary and symmetric systems with homogeneous noise.

Arbitrary system	$L_2 \sim L_2^0 + \frac{b^2/a^2}{n^2}$
Symmetric system	$L_2 \sim L_2^0 + \frac{b^2/a^2}{(n^2+2n\sigma_a^2/a^2)/2}$

We thus obtain Table 5 for the n dependence. Recall the notation b^2 for the variance of the homogeneous noise. Note again the n^2 damping in the arbitrary system, and a damping on the order of $n(n+1)/2$ in the symmetric case, in agreement with the previous analysis. We have assumed independently varying noise in the arbitrary system, and symmetrically varying noise in the symmetric system.

4.5 Approximation justification, accuracy, failure

The justification for equation (31) in the small noise approximation is as follows. Expand the product $\prod_\tau(1+\varepsilon_\tau)$ into a sum. The typical size of the random variable ε_τ is ε, and the largest contribution to the sum comes from terms with $k_{\max}$ ε_τ's, where

$$k_{\max} = [t\varepsilon]$$

is the binomial expected value. The brackets denote the closest integer. This means that the largest terms in the sum come from terms of (29) with $k_{\max}$ $\mathbf{B}$'s. From symmetry considerations it is clear that in the asymptotic limit, the average separation d between two $\mathbf{B}$'s in a term with $k_{\max}$ $\mathbf{B}$'s is

$$\langle d \rangle = \frac{t}{1+t\varepsilon},$$

which is large for small ε and asymptotically independent of t. Furthermore, in a term with $k_{\max} \approx t\varepsilon$ $\mathbf{B}$'s, the separation satisfies [3]

$$P(d < d_0) = 1 - \left(1 - \frac{d_0}{t}\right)^{t\varepsilon} \to \varepsilon d_0$$

in the asymptotic limit, which is small for small ε and independent of t. Therefore, the important terms of (29) for small ε are those with a few $\mathbf{B}$'s separated by long strings of $\mathbf{A}$'s for all t [4].

However, this analysis does not tell the entire story. The accuracy of the perturbation approximation is in fact much higher than one would expect from the above calculation. To understand this, consider a term of the sum (29) with many $\mathbf{B}$'s. This term's direction is impossible to determine in general because each noise matrix transforms it arbitrarily. There are many such terms and they are all affected by a different set of noise matrices. Their directions are thus widely distributed in $\mathbf{R}^n$ and mostly cancel out in the sum.

[4] Note that simulation of divergent moments in a convergent system for large t may not seem accurate, because as t increases the probability of an anomalous event becomes very small. A very large sample space is necessary to obtain an accurate simulation for large t; see the discussion in [4].

When ε is not small, terms in the sum (29) with many **B**'s become important. This causes the perturbation approximation to be inaccurate for two different reasons. First, when **B**'s are adjacent, the approximation of replacing **B** by ε_τ is poor; second, when there are many strings of only a few adjacent **A**'s, both replacing **A** by λ^2 and $\mathbf{B}^\tau$ by ε_τ can be inaccurate. The relative importance of these two inaccuracies can be different. For example, the accuracy of the **A** factor is independent of n while the accuracy of the **B** factor decreases as n increases.

It is difficult to determine a cut-off where ε becomes large. The overall error may be much smaller than the error of each term of the sum (29), because the deviations of the terms may lie in different directions and cancel out in the sum. It is clear that the cut-off depends on how quickly $\mathbf{A}^p$ brings a random vector into alignment with **u**, but even this is a complicated function of the eigenvalue gap and the condition of λ [43]. To account for large ε and handle the contribution from neighboring **B**'s accurately for large n, we develop a different approximation in the next section.

5 Arbitrary noise using iteration approximation

We now present a different method, the iteration technique, which can be applied to find an approximate value for the second moment for small or large noise in well- or ill-conditioned systems.

For homogeneous noise (that is, all the noise elements having the same variance), the approximation can be extended to any level of accuracy for any noise. Unfortunately, for other forms of noise including proportional noise, only the first approximation is applicable.

In addition to providing a way to treat systems where the noise effect is not small, this technique is able to detect the explicit n-dependence of the noise effect. This effect is very slight in the small noise case, but quite important for larger noises.

The general strategy of the method is to express $\langle|\mathbf{x}^{t+1}|^2\rangle$ as a time-independent function of $\{\langle|\mathbf{x}^t|^2\rangle, \langle|\mathbf{x}^{t-1}|^2\rangle, \ldots\}$ in the asymptotic limit. A similar technique was independently developed in [49] for other applications.

5.1 First approximation

The first approximation of this method consists of applying the relation

$$\sum_{ij}(A^r_{ij})^2 \approx \lambda^{2r}v^2, \tag{39}$$

for all r, even $r = 1$, where v is the length of the dominant left eigenvector after the dominant right eigenvector **u** has been normalized to have length 1. Using this approximation we can express $\langle|\mathbf{x}^{t+1}|^2\rangle$ as a t-independent function of $\langle|\mathbf{x}^t|^2\rangle$ alone, as we will see. We thus define $\mathbf{x}^t_A = \mathbf{A}\mathbf{x}^{t-1}$ and $\mathbf{x}^t_B = \mathbf{B}^t\mathbf{x}^{t-1}$, so that $\mathbf{x}^t = \mathbf{x}^t_A + \mathbf{x}^t_B$. We have

$$\langle|\mathbf{x}^t|^2\rangle = \langle|\mathbf{x}^t_A|^2\rangle + \langle|\mathbf{x}^t_B|^2\rangle; \tag{40}$$

the cross term is zero in expectation because there is one power of $\mathbf{B}^t$. We will establish a matrix recurrence relation

$$\begin{pmatrix} \langle|\mathbf{x}^{t+1}_A|^2\rangle \\ \langle|\mathbf{x}^{t+1}_B|^2\rangle \end{pmatrix} \approx \mathbf{M}_1 \begin{pmatrix} \langle|\mathbf{x}^{t}_A|^2\rangle \\ \langle|\mathbf{x}^{t}_B|^2\rangle \end{pmatrix}, \tag{41}$$

where the elements of $\mathbf{M}_1$ (subscript 1 for first approximation) are independent of time. The asymptotic behavior of the second moment is $\langle|\mathbf{x}^t|^2\rangle \sim \mu_1^t$, where μ_1 is the largest eigenvalue of $\mathbf{M}_1$, and

$$L_2 \approx \ln \mu_1$$

for the second moment Lyapunov exponent.

Using the new notation on the recurrence relation, we have:

$$\langle|\mathbf{x}^{t+1}|^2\rangle = \sum_{ijj'} \big\langle (A + B^{t+1})_{ij}(x^t_A + x^t_B)_j \times (A + B^{t+1})_{ij'}(x^t_A + x^t_B)_{j'}\big\rangle \tag{42}$$

or

$$\langle|\mathbf{x}^{t+1}|^2\rangle = \langle|\mathbf{A}\mathbf{x}^t_A|^2\rangle + \sum_{ijj'} A_{ij}A_{ij'}\langle(x^t_B)_j(x^t_B)_{j'}\rangle + \sum_{ijj'}\langle B_{ij}B_{ij'}\rangle\langle(x^t_A)_j(x^t_A)_{j'}\rangle + \sum_{ijj'}\langle B_{ij}B_{ij'}\rangle\langle(x^t_B)_j(x^t_B)_{j'}\rangle \tag{43}$$

because the noise is white with mean 0. This is the simplest form we can obtain without considering particular types of noise.

5.1.1 Homogeneous noise

Recall that homogeneous noise is a type of noise in which all the elements of **B** have the same variance. In this case, equation (43) becomes

$$\langle|\mathbf{x}^{t+1}|^2\rangle \approx \lambda^2\langle|\mathbf{x}^t_A|^2\rangle + (\lambda^2 f_v/n)\langle|\mathbf{x}^t_B|^2\rangle + (n^k b^2 f_u)\langle|\mathbf{x}^t_A|^2\rangle + (n^k b^2)\langle|\mathbf{x}^t_B|^2\rangle), \tag{44}$$

where we introduce the notation

$$f_v = \begin{cases} v^2, & \text{UH noise} \\ (\Sigma_i v_i)^2, & \text{T noise} \end{cases} \tag{45}$$

$$f_u = \begin{cases} 1, & \text{UH noise} \\ (\Sigma_i u_i)^2, & \text{T noise} \end{cases} \tag{46}$$

as well as the factor

$$k = \begin{cases} 1, & \text{UH noise} \\ 2, & \text{T noise} \end{cases} \tag{47}$$

to account for the difference between independent (UH) noise and correlated (T) noise. Multiple steps have been skipped in obtaining equation (44), including the use of (26) with $p = 2$ on the first term and (39) on the others.

We thus obtain

$$\mathbf{M}_1^{UH} = \begin{pmatrix} \lambda^2 & f_v\lambda^2/n \\ nf_u b^2 & n^k b^2 \end{pmatrix},$$

and the second moment evolves as μ_1^t, where μ_1 is the largest eigenvalue of $\mathbf{M}_1^{UH}$ and is given by

$$\mu_1 = \frac{\lambda^2 + n^k b^2 + \sqrt{(\lambda^2 - n^k b^2)^2 + 4\lambda^2 f_u f_v b^2}}{2}. \quad (48)$$

Recalling that for UH noise $v^2b^2/\lambda^2 = \varepsilon^2$, while for T noise $b^2(\Sigma_i u_i)^2(\Sigma_i v_i)^2/\lambda^2 = \varepsilon^2$, we thus have

$$L_2^{UH} \approx \ln\left[\lambda^2\left(\frac{1 + n^k b^2/\lambda^2 + \sqrt{(1 - n^k b^2/\lambda^2)^2 + 4\varepsilon^2}}{2}\right)\right]$$
$$\approx L_2^0 + \varepsilon^2 + \varepsilon^4\left(\frac{n^k}{v^2} - \frac{3}{2}\right) + O\left(\varepsilon^6\right), \quad (49)$$

where $L_2^0 = 2\log\lambda$ is the unperturbed second moment Lyapunov exponent, and the approximation in the second line is valid in the limit of small ε^2 *and* small nb^2/λ^2. The main difference between this expression and the perturbation expansion is that we have taken into account the effect of two neighboring **B**'s, which produces a factor of n. The n dependence enters only in the second and higher order terms; this expression agrees with the perturbation approximation (36) to first order.

5.1.2 Proportional noise

In the case where the noise elements satisfy $b_{ij} = qA_{ij}$ with $q < 1$, we apply (39) with $p = 1$ and proceed as above to find

$$L_2^{UP} \approx \ln\left[\lambda^2\left(\frac{1 + q^2w^2 + \sqrt{(1 - q^2w^2)^2 + 4q^2w^4}}{2}\right)\right]$$
$$\approx L_2^0 + \left(qw^2\right)^2 + \left(qw^2\right)^4\left(\frac{1}{w^2} - \frac{3}{2}\right) + O(qw^2)^6,$$

where w^2 was defined previously (Sect. 4.4) as $\sum_j v_j^2 u_j^2$, and the approximation in the second line is valid in the limit of small qw^2. As expected this agrees with the perturbation approximation result (38) for proportional noise to first order.

5.2 Further approximation

The above treatment is completely accurate in the way it handles the **B** for homogeneous noise. Any inaccuracy stems from using the approximation $\mathbf{A}^r \approx \lambda^r \mathbf{u}\mathbf{v}^T$ on the **A** for $r = 1$. We can improve on this inaccuracy to any desired degree, as explained below. Unfortunately, any approximation past the first order is only applicable to homogeneous noise (all variances the same) and not to proportional noise or any other form with different variances. For the remainder of this section, therefore, only homogeneous noise will be considered.

5.2.1 Second approximation

To illustrate the idea behind further approximation using the iteration technique, we begin with a second approximation wherein (39) is assumed to be accurate for $r = 2$ and higher, but not $r = 1$. In this second approximation, **A**'s which occur "alone" (surrounded by two B's) in an element contribute a factor $\alpha_1\lambda_1^2$ instead of just λ^2, so that α_1 accounts for the inaccuracy of (39).

We now break $\mathbf{x}^t$ into $\mathbf{x}^t = \mathbf{x}_{AA}^t + \mathbf{x}_{AB}^t + \mathbf{x}_B^t$ analogously to (40), where $\mathbf{x}_{AA}^t$ are the terms beginning with **AA**, etc. Proceeding just as above, we find that

$$\begin{pmatrix} \langle|\mathbf{x}_{AA}^{t+1}|^2\rangle \\ \langle|\mathbf{x}_{AB}^{t+1}|^2\rangle \\ \langle|\mathbf{x}_{B}^{t+1}|^2\rangle \end{pmatrix} \approx \mathbf{M}_2 \begin{pmatrix} \langle|\mathbf{x}_{AA}^{t}|^2\rangle \\ \langle|\mathbf{x}_{AB}^{t}|^2\rangle \\ \langle|\mathbf{x}_{B}^{t}|^2\rangle \end{pmatrix}$$

with

$$\mathbf{M}_2 = \begin{pmatrix} \lambda^2 & \lambda^2/\alpha_1 & 0 \\ 0 & 0 & f_v\alpha_1\lambda^2/n \\ nf_u b^2 & nf_u b^2 & n^k b^2 \end{pmatrix},$$

where f_u, f_v and k were previously defined in (46), (45) and (47) to account for the difference between UH noise and T noise.

The second moment will diverge when the largest eigenvalue of M_2 is greater than 1. This eigenvalue is the largest root of the equation

$$\mu^3 - \mu^2(\lambda_1^2 + n^k b^2) + \mu\lambda_1^2 b^2(n^k - f_u f_v \alpha_1)$$
$$+ b^2\lambda_1^4 f_u f_v(1 - \alpha_1) = 0.$$

Notice that in the limit $\alpha_1 = 1$, that is, the limit that the first approximation is accurate, we recover the characteristic equation for the first approximation (48).

5.2.2 Higher order approximation

We can extend the above procedure to any level of accuracy. Define a vector α by

$$\alpha_r = \frac{1}{f_u f_v \lambda^{2r}} \sum_{ab} (A_{ab}^r)^2.$$

The elements of α are the successive corrections to (39). As r increases, α_r tends to 1 because $\lambda^p \gg |\lambda_i|^p$ becomes very accurate for large p. The second moment Lyapunov exponent of the system is given by the log of the largest

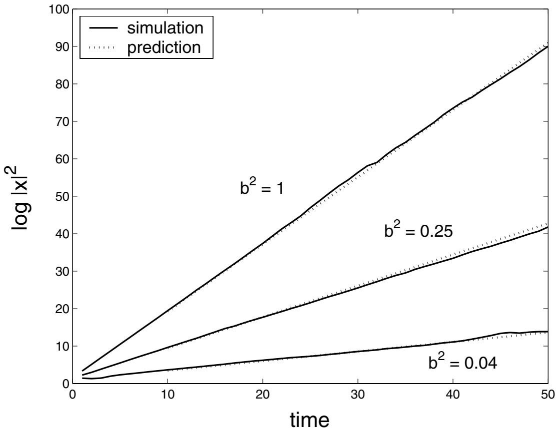

Fig. 9. Accuracy of approximation for very ill-conditioned system with large noise. The second moment of a system with a very poorly-behaved **A** (given in (65)), subject to normal UH noise with various b^2, is simulated. The solid lines are the average over 10 000 runs of the simulation, and the dotted lines are the analytical prediction of Section 5.2.2 for $r = 6$. The **A** in this system has $v^2 = 170.51$ and so virtually any noise is not treatable using the perturbation approximation. The slopes of the dotted lines are 0.255, 0.833, and 1.794 for $b^2 = 0.04, 0.25,$ and 1 respectively. Compare to the perturbation approximation which estimates 2.02, 3.85 and 5.24. This system is convergent in mean with $\lambda = 0.95$; notice how quickly the second moment diverges even for small noise because v^2 is large.

eigenvalue of

$$\mathbf{M}_r = \begin{pmatrix} \lambda^2 & \lambda^2 \frac{1}{\alpha_r} & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda^2 \frac{\alpha_r}{\alpha_{r-1}} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^2 \frac{\alpha_2}{\alpha_1} & 0 \\ 0 & 0 & 0 & \dots & 0 & \lambda^2 \alpha_1 f_v / n \\ n f_u b^2 & n f_u b^2 & n f_u b^2 & \dots & n f_u b^2 & n^k b^2 \end{pmatrix}. \quad (50)$$

in the large p limit. The characteristic equation for this matrix can be expressed iteratively as in Section 5.3 below, but the largest eigenvalue must be computed numerically. This method is exact for any noise and any **A** with a simple, dominant eigenvalue, however ill-conditioned λ may be. The accuracy of the higher-order approximation method is demonstrated in the simulation of Figure 9.

5.3 Large noise limit

Using the results of the higher order approximation, we can obtain an approximation for the second moment Lyapunov exponent for second moment evolution in the large noise limit.

Note that situations where a small noise has a large effect because the system's dominant eigenvalue is ill-conditioned are *not* treatable using the formalism of this section, for the reasons given below.

5.3.1 Criterion for large noise

A good estimate for the onset of the large noise regime can be obtained by comparing the largest eigenvalue of **A** to the average value of the largest eigenvalue of **B**.

For independent (UH) noises with variance b^2 chosen from a normal distribution, the magnitude of the largest eigenvalue of **B** is given on average by [50]

$$\overline{\lambda_B} = b\sqrt{n}, \quad \text{independent noises.} \quad (51)$$

For correlated (T) noises chosen from a normal distribution, the matrix **B** is simply a normal random multiple of the matrix **G** of all ones. **G** has largest eigenvalue n and so the largest eigenvalue of **B** is on average

$$\overline{\lambda_B} = bn, \quad \text{correlated noises.} \quad (52)$$

The large noise case corresponds to $\lambda_B \gg \lambda$, that is,

$$n^k b^2 \gg \lambda^2, \quad (53)$$

where $k = 1$ for independent (UH) noise and $k = 2$ for correlated (T) noises.

5.3.2 Large noise second moment Lyapunov exponent

To find the Lyapunov exponent for second moment evolution in the large noise limit that $n^k b^2 \gg \lambda^2$, we introduce the small parameter

$$\delta = \frac{\lambda^2}{n^k b^2} \ll 1. \quad (54)$$

We will appeal to the iteration treatment of Section 5.2.2 which was accurate for any noise. Recall that the second moment Lyapunov exponent is given in this treatment by the log of the largest eigenvalue μ_r of the matrix M_r (Eq. (50)), where r is any integer. The approximation is more accurate the larger r is, but as we will see, in the large noise limit there is no need to consider large r.

Note that the parameters $\{\alpha_i\}$, f_u, and f_v enter into the calculation of μ_r (Eq. (50)). If these parameters are large, they can ruin our expansion since they multiply δ. We therefore assume that they are $O(1)$. This amounts to assuming the system is well-behaved, and is why, as noted above, this expansion is not applicable to ill-conditioned systems.

The characteristic equation for $\mathbf{M}_{r+1}$ can be written

$$0 = (\lambda^2 - \mu)[\mu D_r + \alpha_r \lambda^{2r} b^2 f_u f_v] + \lambda^{2r} b^2 f_u f_v$$

where μ are the eigenvalues and the D_r are defined recursively by

$$D_{r+1} = -\mu D_r + \alpha_r \lambda^{2r} b^2 f_u f_v$$

with

$$D_2 = \det \begin{pmatrix} -\mu & f_v \alpha_1 \lambda^2 / n \\ n f_u b^2 & n^k b^2 - \mu \end{pmatrix}.$$

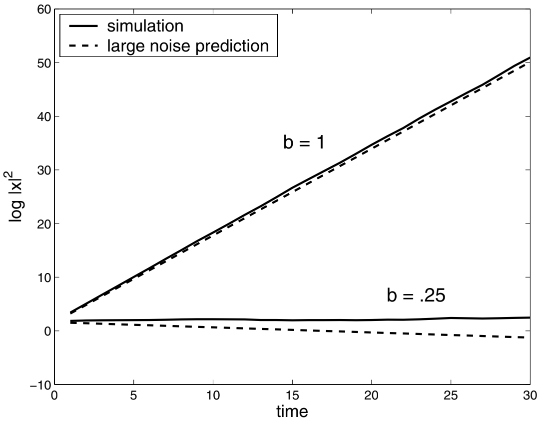

Fig. 10. Accuracy (and inaccuracy) of large noise prediction. The second moment is plotted against the prediction of (56). The $\mathbf{A}$ of (64) is used, which has $n = 5$ and $\lambda = 0.96$. Thus $b = 1$ is well within the large noise regime, as shown in the figure. However, $b = 0.25$ is not within the large noise regime, and the prediction is not very accurate in this case.

Keeping only terms to first order in δ in the characteristic equation above, we find that the largest eigenvalue μ_{r+1} is given approximately by

$$\mu_{r+1} = n^k b^2 \left(\frac{1 + \delta + \sqrt{(\delta - 1)^2 + \delta \frac{4\alpha_1 f_u f_v}{n}}}{2} \right) \quad (55)$$

independent of r, so that

$$L_2 \approx \log n^k b^2 + \delta \frac{\alpha_1 f_u f_v}{n^k} + O(\delta^2) \quad (56)$$

$$\approx \log n^k b^2 + \lambda^2 \frac{\alpha_1 f_u f_v}{n^{2k} b^2} + O(\lambda^4). \quad (57)$$

Note that the zeroth order term corresponds to that found in the $\mathbf{A} = 0$ limit by different means in Section 2.3.

The range of applicability of the large noise approximation is demonstrated in Figure 10.

6 Critical value and stability diagram

As a general rule, large deviations from the average become reasonably likely when the noise is large enough that the second moment diverges. The onset of second moment divergence therefore marks a threshold between two types of behavior and defines a critical value of the size of the noise.

In the case of a homogeneous noise, in which all the noise elements have the same variance b^2, the critical value can be simply expressed as the value of the variance where the second moment Lyapunov exponent L_2 equals 0. For a proportional noise, the critical value is the value q_c of the constant of proportionality (see Tab. 2) for which $L_2 = 1$. The approaches used in this paper allow for a detailed treatment of the critical value b_c^2 in the case of homogeneous noise, but unfortunately not for proportional or other forms of noise, which is left as a topic for future work. Proportional noise is only discussed in the mean value case where it takes the same form as homogeneous noise.

Throughout this section we will assume that the system is well-behaved, that is, f_u, f_v and the $\{\alpha_i\}$ are close to 1. Recall that f_u and f_v account for the difference between independent and correlated noises, and the $\{\alpha_i\}$ measure the accuracy of the approximation (39) for successive powers $\mathbf{A}^i$.

6.1 $\lambda^2 \to 1$ limit

A limit of particular interest when considering the critical value is $\lambda^2 \to 1$, where, as we will see, the critical value drops sharply to 0. Since only small noise is required to cause divergence in this limit, we can apply the first approximation of the iteration treatment, in particular equation (48), to find the critical value

$$(b_c^2)_{\text{small noise}} \approx \frac{1}{n^k + \frac{f_u f_v \lambda^2}{1-\lambda^2}}, \quad (58)$$

where $k = 1$ for independent noises and $k = 2$ for correlated noises. The sharp dropoff to 0 of the critical value is evident from this expression and demonstrated in figure 11 below. Note that when $n = 1$, in which case $f_u = f_v = 1$ and $\lambda = a$ we retrieve the scalar result $b_c^2 = 1 - a^2$.

Our above expression for b_c^2 should coincide with that implied by the perturbation approximation of Section 4. Using equation (36) combined with Table 3 and definitions (46) and (45) for f_u and f_v, we obtain the perturbation approximation result

$$(b_c^2)_{\text{pert}} = \frac{1 - \lambda^2}{f_u f_v},$$

which is equivalent to (58) to first order in $(1 - \lambda^2)$.

6.2 $\lambda \longrightarrow 0$ limit

When λ^2 is small and the system well-behaved, the second moment will only diverge if the noise is large. In this case the limit can be obtained from the large noise treatment of Section 5.3, in particular equation (55) whence we find

$$(b_c^2)_{\text{large noise}} \approx \frac{1}{n^k + \frac{\alpha_1 f_v f_u \lambda^2}{1-\lambda^2}}. \quad (59)$$

For well behaved systems with $\alpha_1 \approx 1$, this expression is almost equivalent to the small noise expression (58)! The scalar result is again retrieved from this expression. We also note that when $\lambda = 0$, we obtain the critical value

$$(b_c^2)_{\lambda=0} = 1/n^k,$$

which was obtained in a different way in Section 2.3.

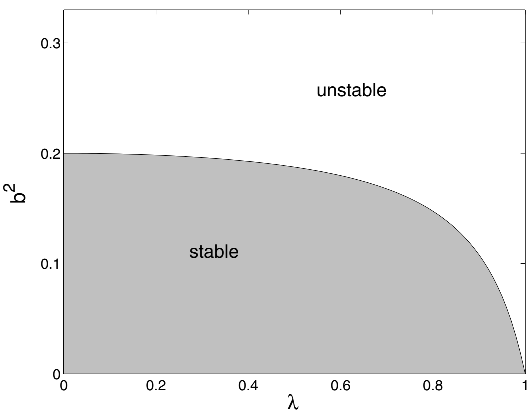

Fig. 11. Second moment stability diagram for $n = 5$ and independent noise elements. The critical value (60) is plotted against the size of the dominant eigenvalue of the unperturbed system.

6.3 Stability diagram

In the case of a well-behaved system with $\alpha_1 \approx 1$, the functional form of the critical value is the same for small and large values of λ^2. We thus propose the expression

$$b_c^2 = \frac{1}{n^k + \frac{f_v f_u \lambda^2}{1-\lambda^2}} \qquad (60)$$

for the critical value for all ranges of homogeneous noise in well-behaved systems. Using this expression we can make a phase plot for the stability regions of the system, as shown in Figure 11 for $n = 5$ and independent (UH) noise.

As for proportional noise, we can produce a stability diagram to compare with the homogeneous noise case in the mean value limit. This diagram indicates how large, as a fraction of the size of the unperturbed elements, the noise must be to cause divergence.

Recall that the size of a proportional noise was defined by the factor q, the constant of proportionality between the typical noise size and the average element of $\mathbf{A}$ (see Tab. 2). In the mean value approximation, the largest eigenvalue is given approximately by $\lambda \approx an$ and we thus obtain the critical value

$$q_c = \frac{n}{\lambda} b_c.$$

This relation produces the phase plot of Figure 12 for the constant of proportionality q_c as a function of λ.

6.4 n dependence of critical value

Expression (60) can be used to study the n dependence of the critical value. The n dependence is weak for large λ, but it is strong when λ is small and a large noise is required to cause divergence.

In the limit $\lambda \longrightarrow 1$, where only a small noise is needed to create second moment divergence, it is seen from expression (60) that the n dependence is quite weak. Indeed,

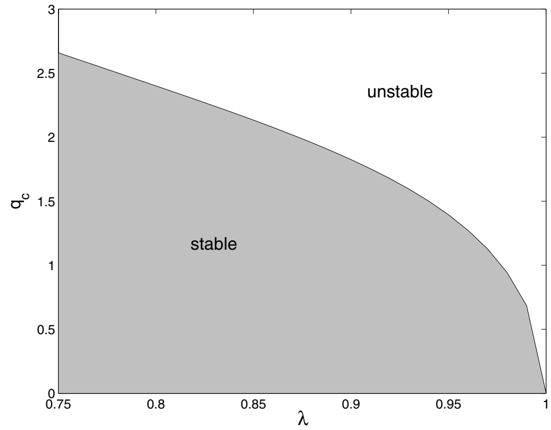

Fig. 12. Stability diagram for proportional noise in mean value approximation, $n = 5$, analogous to Figure 11.

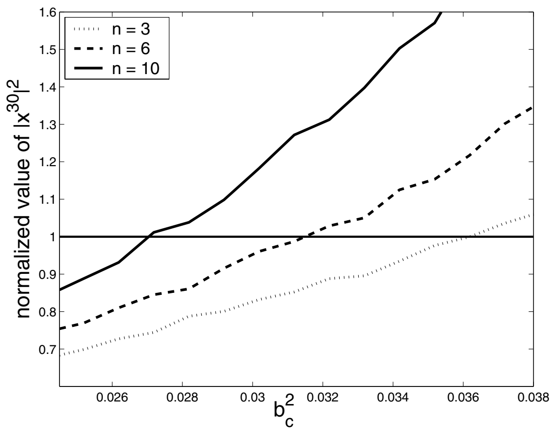

Fig. 13. n dependence of critical value. The figure plots the value of the second moment at $t = 30$ for systems with various b, normalized and averaged over 300 000 runs. Because of the normalization, the initial value (at $t = 0$) of the system's second moment was 1. The critical value b_c^2 is thus indicated for each n by x-coordinate for which the plotted curve's y-coordinate begins to exceed 1. Compare the values of b_c^2 given by the simulation to the analytic estimates $b_c^2 = 0.0365, 0.0326, 0.0288$ for $n = 3, 6, 10$, respectively from equation (60). The values of the elements of the $\mathbf{A}$s used in this plot were generated randomly from a normal distribution and normalized to set $\lambda = 0.98$; matrices without simple, dominant eigenvalues were not accepted.

the expansion (49) for small noise showed that n dependence enters only in the second order term in this limit. The effects of different forms of noise in this case are quite similar.

The n dependence of the critical value for small, independent (UH) noises is demonstrated in Figure 13.

In the opposite limit of $\lambda \longrightarrow 0$ and large noise, the n dependence is quite strong, as shown in (56). It is in this limit that the difference between independent and correlated noises is quite marked, due to the "destructive interference" phenomenon of independently varying noises discussed previously in Section 4.4.

6.5 Comparison to convergence bounds

The critical value (60) provides a much more accurate estimate of the "safe" level of noise for which the second moment does not diverge than do the convergence bounds of Appendix D. We bring this point up because the traditional mathematical approach to stochastic stability is to use bounds.

These bounds are $b^2 < \frac{(1-\lambda_1^2)}{n}$ for the second moment (73) and $b^2 < \frac{(1-\lambda_1)^2}{4n}$ for any moment (72) in the large n limit. For typical well–conditioned systems with v^2 relatively close to 1 (Fig. 5, Appendix C), it is clear that the critical value is much less restrictive than either of the bounds. That is, the bounds stipulate that we must take a very small noise to guarantee convergence of the second moment; but the critical value indicates that the second moment will converge for a much larger range of noise. When v^2 is large, λ and thus the matrix $\mathbf{A}$ are ill-conditioned and the norm of $\mathbf{A}$ is typically much greater than λ, so the bounds (73, 72) are not accurate.

7 Continuous limit

As a last subject, we discuss the continuous time limit of our stochastic system (2). For multidimensional systems, the continuous limit of (2) is the stochastic differential equation (in the Ito sense)

$$d\mathbf{x} = [(\mathbf{A} - \mathbf{I})dt + d\mathbf{B}]\mathbf{x} \quad (61)$$

where $d\mathbf{B} = bd\mathbf{W}$ is a matrix of Wiener processes with mean 0 and standard deviation proportional to $b\sqrt{dt}$, and $\mathbf{I}$ is the identity matrix. Solution or even simulation of this differential system is a difficult problem, and there are few existing analytical results. Two cases which do yield to analysis are the mean value limit where $\mathbf{A} = a\mathbf{G}$, $\mathbf{G}$ the matrix of all ones, and the limit of small noise in two dimensional systems which has been treated in [32] and [33].

There are several fundamental differences between stochastic difference and differential systems. First, there is no such thing as small noise, in the sense we have used, in the continuous time limit. This is because the correct limit of a white noise process has standard deviation proportional to $\sqrt{dt}$ [51], so that the noise dominates as $dt \to 0$. To illustrate this point, consider a particle moving in a one dimensional diffusion process

$$dx = ((a - 1)dt + bdw)x$$

where dw is a Wiener process. When we consider the system's average motion on a large time scale, the particle generally progresses along the curve $x_0 e^{(a-1)t}$. However, on very small time scale, the motion is completely erratic because it is dominated by the noise. Similarly, in the system (61) above, the motion is completely dominated by the noise and the vector $\mathbf{x}$ is transformed erratically around in $\mathbf{R}^n$. The system can never become aligned with $\mathbf{u}$ because the noise causes it to couple with the other modes of $\mathbf{A}$. Only when $\lambda_2 \to 0$ does the system become aligned with $\mathbf{u}$ and behave similarly to the perturbation approximation, above.

A second difference between stochastic difference and differential system is that the updating process in discrete time evolution is synchronous, while that of continuous time evolution is asynchronous. This asynchronicity is a source of difficulty in simulation of stochastic differential systems. Numerical treatment has shown that markedly different behavior occurs in complex, multidimensional systems depending on whether the updating is performed synchronously or asynchronously [52]. Only when the system is completely dominated by one eigenmode, i.e. the mean value limit, should this concern be irrelevant, as the system is effectively one-dimensional.

Both of these concerns thus suggest that we should only expect our discrete results to correspond to those from the stochastic differential system (61) in the mean value limit. This conjecture is borne out by the two examples treated in the remainder of this section. However, a full treatment of this question, in particular for medium and large noise cases, is beyond the scope of this paper and presents a very interesting subject for future research.

7.1 Correspondence between continuous and discrete results in the mean value approximation

For correspondence between the discrete and continuous cases we consider a system in which $\lambda_2 = 0$: the mean value limit that $\mathbf{A} = a\mathbf{G}$. For this $\mathbf{A}$ and totally correlated small noise, an analytic solution to (61) is possible because $\mathbf{A}$ and $d\mathbf{W} = \mathbf{G}dw$ commute [31,53]. Here dw is a one-dimensional Wiener process and $\lambda = na$ is the only nonzero eigenvalue of $\mathbf{A}$. The solution to (61) is

$$\mathbf{x}(t) = e^{(\mathbf{A}-\mathbf{I}-b^2\mathbf{G}^2/2)t+b\mathbf{G}w}\mathbf{x}(0).$$

where $w = \int_0^t dw$ is normal with variance t. From this it is straightforward to calculate that

$$\langle|\mathbf{x}(t)|^p\rangle = e^{pt[(na-1)-\frac{n^2b^2}{2}]}e^{\frac{b^2p^2n^2}{2}t}\left|\frac{1}{n}\mathbf{G}\mathbf{x}(0)\right|^p$$

in the asymptotic limit, and the pth moment Lyapunov exponent is

$$\ell_p = -p\delta + p(p-1)\frac{n^2b^2}{2}$$

where we have taken $\lambda = na = 1-\delta$. This can be compared with the discrete result for small, totally correlated noise in the mean value limit:

$$L_p \approx -p\delta + p(p-1)\frac{n^2b^2}{2}(1-2\delta) + O(\delta^2) + O(b^4),$$

where we have applied $v_i = u_i = n^{-1/2}$. In the limit of small time step the expressions are equivalent to lowest order. This same analysis can also be performed for a scalar system where there are no other modes to couple to.

7.2 Discrepancy between discrete and continuous results

When $\mathbf{A}$ has modes other than the dominant which have nonzero eigenvalues, the discrete result should not, and does not, correspond to the continuous limit. This can be verified by comparison to the result of [32] for small noise moment Lyapunov exponents of arbitrary two-dimensional linear stochastic differential equations. This result, for white noise, is

$$\ell_p = -p\delta + p\gamma_1 \frac{b^2}{2} + p^2\gamma_2 \frac{b^2}{2} + pO(b^2) + O(p^2) \quad (62)$$

where we take $\lambda = 1 - \delta$. The γ factors depend on the form of noise considered. γ_2 depends only on the dominant eigenmode, while γ_1 depends on both eigenmodes. To proceed we assume UH noise for definiteness, wherein one can show that $\gamma_2 = v^2$. The discrete result (36) for UH noise is

$$L_p \approx -p\delta + p(p-1)\frac{v^2 b^2}{2}(1-2\delta) + O(\delta^2) + O(b^4). \quad (63)$$

Comparing this expression to the continuous version (62), we see that the $1-2\delta$ factor on the noise term accounts for the difference between discrete and continuous evolution, as in Section 7.1, and that the $p^2\gamma_2\frac{b^2}{2} + pO(b^2)$ in (62) probably corresponds to the $p(p-1)\frac{v^2b^2}{2}$ term in (63). Note that the $p(p-1)$ form is present in both continuous scalar and T noise cases and is typical of log-normal distributions.

However, the term in (62) proportional to γ_1 is completely absent in the discrete result; moreover, it depends on λ_2 and its eigenvector which have no effect on the small noise discrete system. This term shows how the solution is coupled to all modes, not just the dominant one, in the continuous limit. In fact, for UH noise, one can show (see (68)) that $\gamma_1 = 1 - v^2$; in the mean value limit $v^2 = 1$ and the contribution of the second mode is 0, in agreement with the analysis of Section 7.1.

The author would like to thank David Luenberger, Gene Golub, and especially Rob Schreiber and Bernardo Huberman for helpful discussions and suggestions. This work was partially supported by the National Science Foundation under Grant No. 9986651.

Appendix A: Matrices used to generate figures

All of the unperturbed matrices $\mathbf{A}$ used to generate the figures of this paper were randomly generated. Two particular $\mathbf{A}$s are presented here. The others were generated as described in the text from distributions with very small variance and it is not important to show the exact value of their entries (see discussion in Sect. 3.2).

The matrix which was used to generate Figure 1 and many others as noted in the text is

$$\mathbf{A} = \begin{pmatrix} 0.1795 & 0.0861 & 0.1860 & 0.0924 & 0.1661 \\ 0.1429 & 0.1680 & 0.0517 & 0.2626 & 0.3272 \\ 0.3558 & 0.0127 & 0.2797 & 0.0221 & 0.3227 \\ 0.2766 & 0.2654 & 0.1611 & 0.0408 & 0.0745 \\ 0.3539 & 0.3059 & 0.0596 & 0.2933 & 0.3147 \end{pmatrix}. \quad (64)$$

This matrix has largest eigenvalue $\lambda = 0.966$, second largest eigenvalue $|\lambda_2| = 0.228$, and $v^2 = 1.10$ as computed by Matlab. This matrix is thus quite well-behaved as defined in Section 3.2.

The matrix with ill-conditioned λ used to generate the plot of Figure 9 is

$$\mathbf{A} = \begin{pmatrix} 0.5086 & 0.3496 & 0.0795 & -0.2044 & -0.3530 \\ -0.6168 & 0.1553 & 0.5224 & -0.0293 & 0.0137 \\ -0.5526 & 0.0069 & 0.0008 & -0.3189 & 0.4345 \\ 0.4805 & 0.8053 & -0.5502 & 0.6173 & -0.3041 \\ -0.4307 & 0.8960 & 0.0255 & 0.1454 & 0.6965 \end{pmatrix}, \quad (65)$$

with largest eigenvalue $\lambda = 0.950$, second largest eigenvalue $|\lambda_2| = 0.888$, and $v^2 = 170.3$ as computed by Matlab.

Appendix B: Reduction of nonnegative stability analysis to primitive systems

The reason that λ is simple and dominant in all nonnegative systems of interest is that we need only consider systems with primitive $\mathbf{A}$, and primitive matrices have the above property by the Perron-Frobenius theorem. Stability analysis of any nonnegative system whose matrix is not primitive reduces to analysis of primitive subsystems.

More precisely, nonnegative matrices which are not primitive may be either reducible or irreducible imprimitive. Reducible matrices are those which can be written in the form

$$\begin{pmatrix} \mathbf{C} & \mathbf{X} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}, \quad (66)$$

where $\mathbf{C}$ and $\mathbf{D}$ are square, by renaming the indices [54]. Stability analysis reduces to analysis of the subsystems $\mathbf{C}$ and $\mathbf{D}$, because $\begin{pmatrix} \mathbf{C} & \mathbf{X} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}^n = \begin{pmatrix} \mathbf{C}^n & \mathbf{Y} \\ \mathbf{0} & \mathbf{D}^n \end{pmatrix}$. A similar reduction occurs on the subsystems unless they are irreducible. Irreducible imprimitive matrices can be written as

$$\begin{pmatrix} \mathbf{0} & \mathbf{C}_{12} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_{23} & \dots & \mathbf{0} \\ \vdots & & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{C}_{h-1h} \\ \mathbf{C}_{h1} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}, \quad (67)$$

where the 0 blocks along the diagonal are square (second part of Perron Frobenius theorem). The hth power of such a matrix is block diagonal and the blocks are primitive [54], so the stability analysis is again reduced.

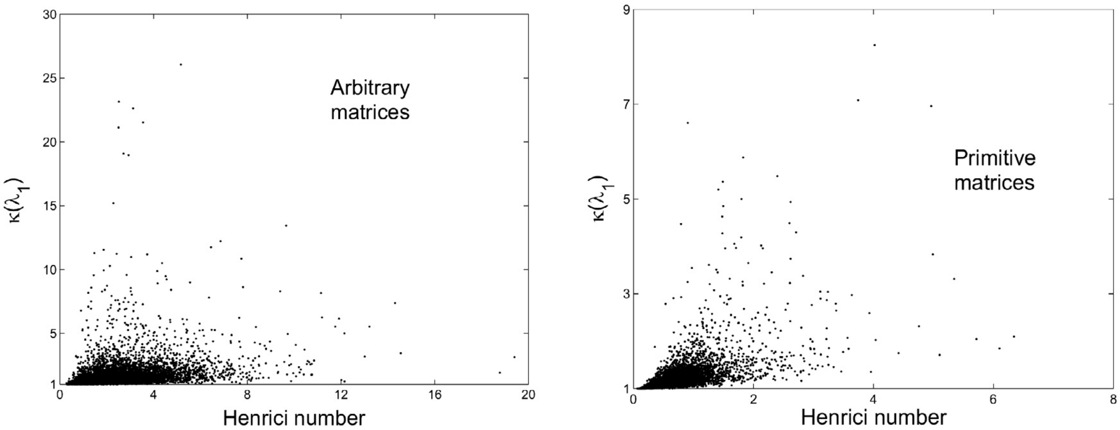

Fig. 14. Scatter plots of $|\mathbf{v}|$ versus the Henrici number $|\mathbf{A}\mathbf{A}^T - \mathbf{A}^T\mathbf{A}|$, a measure of non-normality, for 10,000 randomly generated 5×5 matrices normalized so that $\lambda = 1$. At left the elements were chosen from a normal distribution and only plotted if λ was real. At right, the matrices are nonnegative primitive; a random number of elements were 0, and the nonzero elements were chosen from a uniform distribution.

Physically, primitive matrices have the property that their powers are positive[5] (have no 0 elements). From a physical perspective, primitive systems are thus "fully interacting". This is in contrast to other nonnegative matrices which have zero blocks when raised to any power.

Appendix C: Further discussion of properties of A

There is a correlation between an ill-conditioned λ and a small eigenvalue gap. This is so because a matrix with a large $\kappa(\lambda)$ is close to a matrix where λ is repeated. In particular [43], there exists a matrix $\mathbf{E}$ such that λ is a repeated eigenvalue of $\mathbf{A} + \mathbf{E}$ and

$$|\mathbf{E}| \leq \frac{|\mathbf{A}|}{\sqrt{(\kappa(\lambda))^2 - 1}}.$$

However, $\kappa(\lambda)$ may be small even if the gap is small. The relation between $\kappa(\lambda)$ and the eigenvalue gap is shown in Figure 5, above.

There is also a correlation between normality of $\mathbf{A}$ and a small $\kappa(\lambda)$. When $\mathbf{A}$ is normal, that is, $\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A}$, all of its eigenvectors are orthogonal and all the eigenvalues are perfectly conditioned. However, $\kappa(\lambda)$ may be small in matrices which are far from normal. The relation between $\kappa(\lambda)$ and the normality of $\mathbf{A}$ is shown in Figure 14.

Another way to characterize $\kappa = |\mathbf{v}|$ is the relation

$$v^2 = 1 - \sum_{i \neq 1} (\mathbf{u} \cdot \mathbf{e}_i^R)(\mathbf{v} \cdot \mathbf{e}_i^L), \tag{68}$$

where $\mathbf{e}_i^R$ is the ith column of $\mathbf{P}$ (the right eigenvector corresponding to λ_i), $\mathbf{e}_i^L$ is the ith row of $\mathbf{P}^{-1}$ (the left eigenvector corresponding to λ_i), and $\mathbf{u}$ and $\mathbf{v}$ are the right and left eigenvectors corresponding to λ. This relation is established by noting that $\sum_{ij} \mathbf{e}_i^R \mathbf{e}_j^R \mathbf{e}_i^L \mathbf{e}_j^L = 1$. It shows how v^2 is related to the angles between the eigenvectors. In particular, we see that for a normal matrix where the eigenvectors are orthogonal, $v^2 = 1$; but in general, the angular distribution of the eigenvalues is complicated.

Finally, there is a correlation between $|\lambda_2| \to 0$ and the variance σ_A^2 of the elements of $\mathbf{A}$. Bounds for the second largest eigenvalue can be found in the case of row (or column) stochastic matrices, for example [54]: $|\lambda_2| \leq \min\left(1 - \sum_i \min_j A_{ij}, \sum_i \max_j A_{ij} - 1\right)$. This shows that, at least for stochastic matrices, a small variance σ_A^2 corresponds to a large eigenvalue gap. This is shown to be true for all matrices in Figure 15. Of course, the converse is not true; matrices with large σ_A^2 can also have a large eigenvalue gap, as also is shown in Figure 15.

Appendix D: Bounds on convergence of $\langle |\mathbf{x}|^2 \rangle$

In this section we apply the matrix 2-norm to determine two different bounds on the variance of the noise which, if satisfied, ensure the convergence of $\langle |\mathbf{x}|^2 \rangle$. These conditions are sufficient but by no means necessary. The second moment will of course never converge in a system of the form (2) if the system does not converge in mean. We therefore take $\lambda < 1$ in this section.

The norm of a matrix is any function satisfying the regular properties of a vector norm and additionally the inequality $|\mathbf{A}\mathbf{B}| \leq |\mathbf{A}||\mathbf{B}|$. The matrix 2-norm corresponding to the usual Euclidean vector norm is

$$|\mathbf{A}|_2 = (\rho(\mathbf{A}\mathbf{A}^*))^{1/2}, \tag{69}$$

where ρ is the spectral radius. Note that for any norm, $|A| \geq |\mathbf{x}\mathbf{A}|/|\mathbf{x}| = |\lambda|$ for any eigenvalue λ, so that in particular,

$$\lambda \leq |\mathbf{A}|. \tag{70}$$

[5] More exactly, the pth power of a nonnegative $n \times n$ primitive matrix $\mathbf{A}$ has no zero elements for all $p \geq \gamma(\mathbf{A})$, where $\gamma(\mathbf{A})$ (the index of primitivity) is at most $n^2 - 2n + 2$, and usually much less [54]).

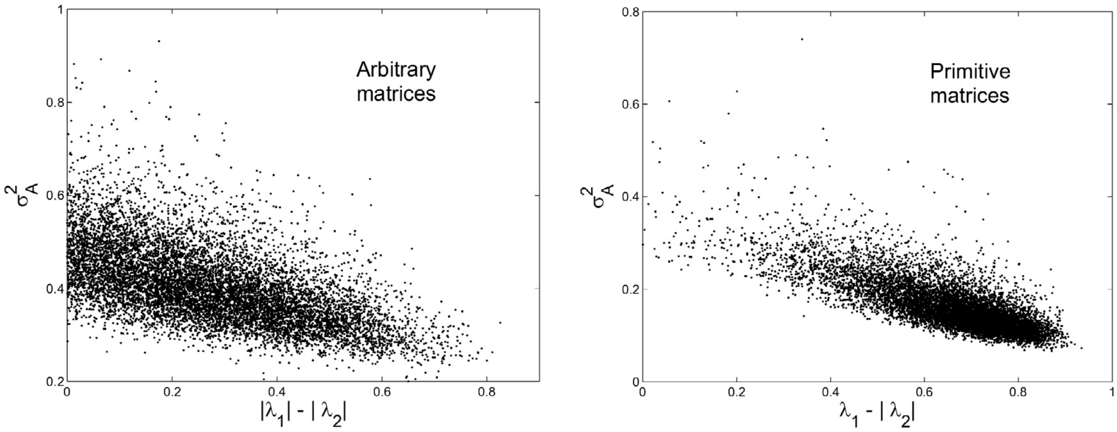

Fig. 15. Scatter plots of eigenvalue gap $\lambda - |\lambda_2|$ versus standard deviation of the $\mathbf{A}_{ij}$ for 10,000 randomly generated 5×5 matrices, normalized so that $\lambda = 1$. At left the elements were chosen from a normal distribution and only plotted if λ is real. At right, the matrices are nonnegative primitive; a random number of elements were 0, and the nonzero elements were chosen from a uniform distribution.

For ill-conditioned matrices, which includes those with ill-conditioned λ, $|\mathbf{A}|$ is typically much larger than λ [43].

D.1 Bound on convergence of any moment

We have $|\mathbf{x}^{t+1}|^p \le |\mathbf{A} + \mathbf{B}^t|^p |\mathbf{x}^t|^p$, so that

$$\langle |\mathbf{x}^t|^p \rangle \le \left[\prod_{\tau=1}^{t-1} \langle |\mathbf{A} + \mathbf{B}^\tau|\rangle^p \right] |\mathbf{x}^0|^p,$$

where the expected value goes inside the product because the noise is white noise. $\langle |\mathbf{x}|^p \rangle$ will thus converge for any p provided that $\langle |\mathbf{A} + \mathbf{B}| \rangle < 1$ (we neglect the time superscript because the noise is stationary), or more usefully

$$\langle |\mathbf{B}| \rangle < 1 - |\mathbf{A}| \tag{71}$$

using $|\mathbf{A} + \mathbf{B}| \le |\mathbf{A}| + |\mathbf{B}|$. Since convergence of every moment is a much stronger condition than convergence of just the second moment, this bound is typically poor when applied to the second moment.

We may estimate a lower limit for this bound for well-conditioned systems in the large n limit when the noise is UH (uncorrelated B_{ij} all with the same variance b^2). We do so by using $|\mathbf{A}| > \lambda$ (Eq. (70)) and a result of [55] that $\liminf |\mathbf{B}| \ge 2b\sqrt{n}$ almost surely in the large n limit, provided that the elements of $\mathbf{B}$ are mean 0 i.i.d. and their moments do not grow too fast (which is satisfied for any reasonable noise). Thus

$$\lambda_1 + 2b\sqrt{n} \le |\mathbf{A}| + \langle |\mathbf{B}| \rangle$$

and the condition (71) on b for convergence is at least weaker than

$$b^2 < \frac{(1-\lambda_1)^2}{4n} \tag{72}$$

in the large n limit. That is to say, (71) is more restrictive on b than (72). For ill-conditioned systems, (72) may not be accurate because $|\mathbf{A}|$ may be much larger than λ.

D.2 Second moment bound

A different bound on the convergence of the second moment in the case of UH noise can be found by applying the expected value before taking norms. We have

$$\begin{aligned}\langle |\mathbf{x}^t|^2 \rangle &= \langle |(\mathbf{A} + \mathbf{B}^t)\mathbf{x}^{t-1}|^2 \rangle \\ &\le (|\mathbf{A}|^2 + nb^2)\langle |\mathbf{x}^{t-1}|^2 \rangle,\end{aligned}$$

where we have used the properties of the norm and the fact that the noise is UH, white and has mean 0. We thus have the convergence condition $|\mathbf{A}|^2 + nb^2 < 1$, or

$$b^2 < \frac{(1 - |\mathbf{A}|^2)}{n}$$

for the convergence of $\langle |x|^2 \rangle$. Note that this condition is at least weaker than the condition

$$b^2 < \frac{(1-\lambda_1^2)}{n} \tag{73}$$

because of (70). Again, (73) may not be accurate for ill-conditioned $\mathbf{A}$ because $|\mathbf{A}|$ may be much larger than λ.

References

1. J. Guckenheimer, Lect. Applied Math. **17**, 187 (1979)
2. C.W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 2nd edn. (Springer-Verlag, Berlin, 1985)
3. J.A. Rice, *Mathematical Statistics and Data Analysis*, 2nd edn. (Duxbury Press, Belmont, California, 1995)
4. S. Redner, Am. J. Phys. **53**, 267 (1990)
5. R.C. Lewontin, D. Cohen, Proc. Nat. Ac. Sci. **62**, 1056 (1969)
6. J.N. Tsitsiklis, V.D. Blondel, Mathematics of Control, Signals, and Systems **10**, 31 (1997)
7. P. Bougerol, J. Lacroix, *Products of Random Matrices with Applications to Schrödinger Operators* (Birkhäuser, Boston, 1985)

8. G.W. Stewart, University of Maryland Institute for Advanced Computer Studies, Report No. 91, 2001 (unpublished)
9. J.E. Cohen, C.M. Newman, The Annals of Probability **12**, 283 (1984)
10. L. Gurvits, Linear Algebra and its Applications **231**, 47 (1995)
11. N.E. Barabanov, Automation and Remote Control **49**, 152 (1988)
12. D. Hinrichsen, N.K. Son, in *Proceedings of the Third International Conference on Approximation and Optimization*, edited by B. Bank et al. (The Electronic Library of Mathematics, 1995)
13. P.R. Kumar, P. Varaiya, *Stochastic Systems: Estimation, Identification and Adaptive Control* (Prentice-Hall, Englewood Cliffs, New Jersey, 1986)
14. C. Xiao, D.J. Hill, P. Agathoklis, IEEE Transactions on Circuits and Systems **44**, 614 (1997)
15. G.V. Moustakidis, Int. J. Adaptive Control and Signal Processing **12**, 579 (1998)
16. M. Milisavljević, E.I. Verriest, in *Proceedings of the European Control Conference*, 1997
17. F. Ma, T.K. Caughey, Int. J. Nonlinear Mech. **16**, 139 (1981)
18. F. Kozin, Automatica **5**, 95 (1969)
19. W.M. Haddad, V. Chellaboina, S.G. Nersesov, Mathematical Problems in Engineering **8**, 493 (2002)
20. T. Hogg, B.A. Huberman, J.M. McGlade, Proceedings of the Royal Society of London B **237**, 43 (1989)
21. N.G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992)
22. J.-P. Eckmann, D. Ruelle, Rev. Mod. Phys. **57**, 617 (1985)
23. B. Derrida, E. Gardner, J. Phys. France **45**, 1283 (1984)
24. B. Derrida, K. Mecheri, J. Pichard, J. Phys. France **48**, 733 (1987)
25. N. Zanon, B. Derrida, J. Stat. Phys. **50**, 509 (1988)
26. J.B. Pendry, J. Phys. C **15**, 4821 (1982)
27. G. Parisi, J. Phys. A **19**, L675 (1986)
28. A. Crisanti, G. Paladin, A. Vulpiani, *Products of Random Matrices in Statistical Physics* (Springer-Verlag, Berlin, 1993)
29. J.P. Bouchaud, A. Georges, D. Hansel, P.L. Doussel, J.M. Maillard, J. Phys. A **19**, L1145 (1986)
30. R. Khasminskii, Prikladnaya Matematika i Mekhanika (Journal of Applied Mathematics and Mechanics) **31**, 1021 (1967)
31. X. Mao, *Stochastic Differential Equations and Applications* (Horwood, Chichester, 1997)
32. L. Arnold, M.M. Doyle, N.S. Namachchivaya, Dynamics and Stability of Systems **12**, 187 (1997)
33. R. Khasminskii, N. Moschuk, SIAM J. Appl. Math. **58**, 245 (1998)
34. A. Abrahams, M.J. Stephen, J. Phys. C **13**, L377 (1980)
35. G. Strang, *Linear Algebra and Its Applications* (Academic Press, New York, 1976)
36. *Lyapunov Exponents*, Lecture Notes in Mathematics Vol. 1186, edited by L. Arnold, V. Wihstutz (Springer, New York, 1986)
37. A. Berman, M. Neumann, R.J. Stern, *Nonnegative Matrices in Dynamic Systems* (John Wiley and Sons, New York, 1989)
38. L. Farina, S. Rinaldi, *Positive Linear Systems: Theory and Applications* (John Wiley and Sons, New York, 2000)
39. *Positive Systems*, Lecture Notes in Control and Information Sciences Vol. 294, edited by L. Benvenuti, A. de Santis, L. Farina (Berlin, 2003, Springer)
40. L.N. Trefethen, A.E. Trefethen, S.C. Reddy, T.A. Driscoll, Science **261**, 578 (1993)
41. R. Bellman, Duke Mathematical J. **21**, 491 (1954)
42. H. Furstenberg, H. Kesten, Annals of Mathematical Statistics **31**, 457 (1960)
43. G.H. Golub, C.F. van Loan, *Matrix Computations*, 2nd edn. (The Johns Hopkins University Press, Baltimore, 1989)
44. A. Edelman, Ph.D. dissertation, MIT, 1989
45. H. Minc, *Nonnegative Matrices* (John Wiley and Sons, New York, 1988)
46. F. Juhasz, Discrete Mathematics **41**, 161 (1982)
47. Y.B. Zel'dovich, S.A. Molchanov, A.A. Ruzmaikin, D.D. Sokolov, Uspekhi Fizicheskikh Nauk **152**, 3 (1987)
48. Z. Füredi, J. Komlós, Combinatorica **1**, 233 (1981)
49. S. Roy, Ph.D. dissertation, MIT, 2003
50. S. Geman, Ann. Prob. **14**, 1318 (1986)
51. P.E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer-Verlag, Berlin, 1992)
52. B.A. Huberman, N.S. Glance, Proc. Natl. Acad. Sci. **90**, 7716 (1993)
53. B.K. Øksendal, *Stochastic Differential Equations: an Introduction with Applications*, 6th edn. (Springer, Berlin, 2003)
54. A. Berman, R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences* (Academic Press, New York, 1979)
55. S. Geman, Ann. Prob. **8**, 252 (1980)