Fluctuation properties of steady-state Langevin systems

Jeffrey B. Weiss[*]

Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado 80309, USA

(Received 24 August 2007; published 27 December 2007)

Motivated by stochastic models of climate phenomena, the steady state of a linear stochastic model with additive Gaussian white noise is studied. Fluctuation theorems for nonequilibrium steady states provide a constraint on the character of these fluctuations. The properties of the fluctuations which are unconstrained by the fluctuation theorem are investigated and related to the model parameters. The irreversibility of trajectory segments, which satisfies a fluctuation theorem, is used as a measure of nonequilibrium fluctuations. The moments of the irreversibility probability density function (pdf) are found and the pdf is seen to be non-Gaussian. The average irreversibility goes to zero for short and long trajectory segments and has a maximum for some finite segment length, which defines a characteristic time scale of the fluctuations. The initial average irreversibility growth rate is equal to the average entropy production and is related to noise amplification. For systems with a separation of deterministic time scales, modes with time scales much shorter than the trajectory time span and whose noise amplitudes are not asymptotically large, do not, to first order, contribute to the irreversibility statistics, providing a potential basis for dimensional reduction.

I. INTRODUCTION

Recent advances in nonequilibrium statistical mechanics have investigated fluctuation theorems in a variety of contexts [1]. The fluctuation theorem quantifies the probability of finding fluctuations in nonequilibrium systems that violate the second law of thermodynamics. Fluctuation theorems take many forms. The formulation we will focus on is in terms of the probability of observing finite time trajectory segments of a system [2,3]. In this context, the fluctuation theorem provides a constraint that such trajectory segments must satisfy. Here we investigate the fluctuations in a nonequilibrium steady state governed by Langevin dynamics. We go beyond the fluctuation theorem and study those properties of the fluctuations that are not constrained by the fluctuation theorem. These properties are not generic. They depend on the details of the specific dynamical system, and we investigate the relationship between the nonequilibrium fluctuations and the parameters defining the Langevin dynamics.

Our motivation for studying specific details of nonequilibrium fluctuations comes from work in theoretical climate dynamics. In recent years, linear stochastic dynamical systems have been successfully used to model many phenomena in the climate system such as El-Niño [4–6], the North Atlantic Gulf Stream [7], and a variety of atmospheric phenomena [8–12]. We shall refer to these phenomena as climate subsystems in that they are often considered to be dynamical systems that are separable from the larger climate system, at least on some set of spatial and temporal scales. Since these fluctuations have macroscopic time scales, it is important to investigate the character of individual fluctuations and the statistics of their properties.

In the work on climate subsystems, the focus has been on two considerations: the utility of linear stochastic systems in forecasting [4,6,13], and the potential for the deterministic

part of the dynamics to amplify the random noise [6,14–17]. It is often assumed that the large amplitudes of these phenomena requires them to be the result of dynamical instabilities. The recognition that deterministic dynamics can amplify small noise forcing, which in meteorology goes back to Lorenz [18], provides an alternative view of such phenomena. This amplification occurs when the deterministic matrix is non-normal [14,15]. One common critique of the noise-amplification view is that non-normality and the resulting amplification depends on the subjective choice of coordinate system, and can be removed by an appropriate coordinate transformation. Recently, this objection has been answered by noting that underlying the property of non-normality is the more fundamental, coordinate invariant property of detailed balance. Linear stochastic climate subsystem models share the property that they violate detailed balance, and this is what is responsible for the noise amplification [17]. Thus, a wide range of phenomena in the climate system can be interpreted as fluctuations in a nonequilibrium steady state. For climate fluctuations such as El-Niño, understanding the character of the fluctuations is extremely important. Further, due to global warming, the steady state is changing. Understanding how phenomena such as El-Niño will change as climate changes is a major uncertainty in climate change predictions [19]. Thus, improved understanding of how nonequilibrium fluctuations depend on the properties of the steady state could lead to improvements in climate change forecasts.

As a concrete example, we will focus on El-Niño. El-Niño is a coupled atmosphere-ocean phenomenon that is centered in the tropical Pacific Ocean and has global impacts. One key aspect of El-Niño is that the atmosphere evolves on a faster time scale than the ocean. The turbulent dynamics of the atmosphere has a predictability limit of about 2 weeks [20]. The ocean, on the other hand, has time scales of months. Thus, on the monthly time scale of El-Niño, the atmosphere is unpredictable and can be considered a random forcing [21]. While this parametrization of fast

───────
[*]jeffrey.weiss@colorado.edu

chaos as random noise is typically done empirically, there are some theoretical results [22–25].

The phenomenon of El-Niño is described in terms of the sea surface temperature (SST) of the tropical Pacific. Although the SST is a continuous field, both observations and models use a finite number N of SST values. Thus, the state vector x of the system is an N-dimensional vector of real numbers representing the discretized SST field. Often, the dimensionality is reduced by truncating to some number of leading modes. Typically, the mean SST is removed, so x represents the SST anomaly and can be positive or negative. A linear stochastic Langevin model for El-Niño is then

$$\frac{dx}{dt} = \mathbf{A}x + \mathbf{F}\xi, \tag{1}$$

where $\mathbf{A}$ is an $N \times N$ real matrix representing the linear deterministic dynamics, $\mathbf{F}$ is an $N \times N$ real matrix representing the noise forcing, ξ is N-dimensional Gaussian white noise, $\langle \xi(t)\xi^T(s)\rangle = \mathbf{I}\delta(t-s)$, where superscript T represents the transpose, $\mathbf{I}$ is the identity matrix, and the diffusion matrix is $\mathbf{D} = \mathbf{F}\mathbf{F}^T/2$. We require the system to have a steady state, which implies that the deterministic dynamics is stable, i.e., all eigenvalues of $\mathbf{A}$ have a negative real part. From a dynamical systems perspective, Eq. (1) describes a stable linear fixed point perturbed by additive Gaussian white noise. Equation (1) is the fundamental equation defining the dynamics, and our goal is to describe the nonequilibrium fluctuations in terms of $\mathbf{A}$, $\mathbf{D}$, and matrices derived from them.

The Langevin dynamics, Eq. (1), describes both equilibrium and nonequilibrium steady states, depending on whether or not detailed balance is satisfied. For most matrices $\mathbf{A}$ and $\mathbf{D}$, detailed balance is violated and the system has a nonequilibrium steady state. This is the case for Langevin models of climate subsystems. Detailed balance requires $\mathbf{AD} - \mathbf{DA}^T = 0$, and then $\mathbf{\Omega} = \mathbf{A} + \mathbf{DQ}_0 = 0$, where $\mathbf{Q}_0$ is the inverse of the steady-state covariance, defined explicitly below, while both expressions are nonzero when detailed balance is violated. When detailed balance is violated, the steady-state distribution $p_0(x)$ is maintained by a nonzero divergence-free probability current $j(x) = \mathbf{\Omega}x p_0(x)$, where $\mathbf{\Omega}x$ is a phase space velocity and $\mathbf{\Omega}$ can be interpreted as a matrix of rotation frequencies. The probability current is divergence free, and thus $\text{Tr}(\mathbf{\Omega}) = 0$ The system satisfies detailed balance if and only if there exists a coordinate system where $\mathbf{A}$ and $\mathbf{D}$ are both diagonal. Thus, systems described by Eq. (1) in detailed balance can be transformed into a collection of uncoupled one-dimensional systems, while those violating detailed balance have an essential multidimensional character.

II. FLUCTUATION THEOREM

The fluctuation theorem can be written in terms of the probability of observing trajectory segments. Consider a long trajectory x and choose a time interval t. For any two states x_0 and x_1, we define the trajectory probability $p(x_0, x_1, t)$ as the probability of finding a trajectory segment within the long trajectory that begins at x_0 and ends at x_1 a time t later. This definition differs from the trajectory probability used in

some previous studies [3,26] in that it depends only on the end points of the trajectory segment and not on the full trajectory. By relying only on the end points of the trajectory segment, this definition may be better suited to discretely sampled data, such as that found in climate records.

The time-reversed trajectory segment, one starting at x_1 and ending at x_0, has a probability $p(x_1, x_0, t)$. For simplicity we shall drop the dependence on the time interval t when it causes no confusion. The irreversibility $r(x_0, x_1, t)$ of a trajectory segment with initial state x_0 and final state x_1 is defined by

$$r(x_0, x_1) = \ln \frac{p(x_0, x_1)}{p(x_1, x_0)}. \tag{2}$$

Thus, $r \equiv 0$ shows that one is equally likely to find forward and reverse trajectory segments and the system is reversible, while $r \neq 0$ shows that one can distinguish forward from reverse trajectories segments. The statistics of r over a long trajectory corresponds to the intuitive notion of irreversibility: they quantify how well one is able to distinguish a long trajectory from the same trajectory going backward in time.

The fluctuation theorem is now a direct result of the definition of r [3]. Equation (2) implies that $r(x_1, x_0) = -r(x_0, x_1)$ and $p(x_0, x_1) = \exp[r(x_0, x_1)]p(x_1, x_0)$. Then the probability $p_r(r)$ of finding a trajectory segment with irreversibility r is

$$p_r(r) = \int d^N x_0 d^N x_1 p(x_0, x_1)\delta[r(x_0, x_1) - r]$$

$$= \int d^N x_0 d^N x_1 e^{-r(x_1, x_0)} p(x_1, x_0)\delta[r(x_1, x_0) + r]$$

$$= e^r p_r(-r). \tag{3}$$

The final entry in Eq. (3) is the fluctuation theorem relating the probability of finding positive and negative irreversibilities. The fluctuation theorem thus puts a constraint on $p_r(r)$, but does not completely determine its functional form.

III. IRREVERSIBILITY OF TRAJECTORY SEGMENTS

For the linear stochastic Langevin dynamics, Eq. (1), the probability of finding a trajectory segment with a given irreversibility $p_r(r)$ can be written in terms of the basic parameters of the dynamics. We begin by defining two more traditional probability distributions: the steady-state probability $p_0(x)$ of finding the system in state x, and the transition probability $p(x_1, t|x_0)$ of finding the system in state x_1 conditioned on the system being in state x_0 a time t earlier. The trajectory probability is then

$$p(x_0, x_1, t) = p(x_1, t|x_0)p_0(x_0). \tag{4}$$

Because the system is linear with additive Gaussian white noise, these probabilities are also Gaussian and can be explicitly written in terms of the covariance of the dynamics.

A stochastic trajectory of Eq. (1) starting at $x(0)$ can be written as

$$x(t) = e^{\mathbf{A}t}x(0) + \int_0^t ds\, e^{\mathbf{A}(t-s)}\mathbf{F}\xi. \qquad (5)$$

Since we only consider additive noise, the above integral is the same in both Ito and Stratonovich interpretations. The time-dependent covariance matrix $\mathbf{C}_t$ for $\Delta x = x_1 - \exp(\mathbf{A}t)x_0$, the difference between the stochastic trajectory and its deterministic counterpart is

$$\mathbf{C}_t = \langle \Delta x \Delta x^T \rangle = 2 \int_0^t e^{\mathbf{A}(t-s)}\mathbf{D}e^{\mathbf{A}^T(t-s)}, \qquad (6)$$

and the steady-state covariance $\mathbf{C}_0 = \lim_{t \to \infty} \mathbf{C}_t$ satisfies the relation

$$\mathbf{A}\mathbf{C}_0 + \mathbf{C}_0\mathbf{A}^T + 2\mathbf{D} = \mathbf{0}. \qquad (7)$$

Writing $\mathbf{C}_0$ as

$$\mathbf{C}_0 = 2 \int_{-\infty}^0 ds\, e^{-\mathbf{A}s}\mathbf{D}e^{-\mathbf{A}^T s}, \qquad (8)$$

allows one to obtain the useful relation

$$e^{\mathbf{A}t}\mathbf{C}_0 e^{\mathbf{A}^T t} = \mathbf{C}_0 - \mathbf{C}_t. \qquad (9)$$

Define the inverse covariance matrices, sometimes called the concentration, by $\mathbf{Q}_t = \mathbf{C}_t^{-1}$, $\mathbf{Q}_0 = \mathbf{C}_0^{-1}$.

The steady-state probability is

$$p_0(x) = \mathcal{N}_0 \exp(-x^T \mathbf{Q}_0 x / 2), \qquad (10)$$

with normalization $\mathcal{N}_0 = 1/\sqrt{(2\pi)^N \mathrm{Det}(\mathbf{C}_0)}$. The transition probability is

$$p(x_1, t | x_0) = \mathcal{N}_t \exp(-\Delta x^T \mathbf{Q}_t \Delta x / 2), \qquad (11)$$

where the normalization is $\mathcal{N}_t = 1/\sqrt{(2\pi)^N \mathrm{Det}(\mathbf{C}_t)}$. The transition probability is thus a Gaussian distribution with covariance $\mathbf{C}_t$ distributed around the end point of the deterministic trajectory segment beginning at x_0.

The trajectory probability depends on both the initial and final states. We thus consider a $2N$-dimensional state space

$$z = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}. \qquad (12)$$

The forward trajectory probability is Gaussian and is given by

$$p(x_0, x_1) = \mathcal{N} \exp(-z^T \mathbf{R}_{01} z / 2), \qquad (13)$$

with normalization $\mathcal{N} = 1/[(2\pi)^N \sqrt{\mathrm{Det}(\mathbf{C}_0\mathbf{C}_t)}]$, and the concentration of the trajectory probability $\mathbf{R}_{01}$ is the $2N \times 2N$ matrix

$$\mathbf{R}_{01} = \begin{pmatrix} e^{\mathbf{A}^T t}\mathbf{Q}_t e^{\mathbf{A}t} + \mathbf{Q}_0 & -e^{\mathbf{A}^T t}\mathbf{Q}_t \\ -\mathbf{Q}_t e^{\mathbf{A}t} & \mathbf{Q}_t \end{pmatrix}. \qquad (14)$$

The reverse trajectory probability is

$$p(x_1, x_0) = \mathcal{N} \exp(-z^T \mathbf{R}_{10} z / 2), \qquad (15)$$

where the concentration of the reverse trajectory probability is related to that of the forward trajectory probability by swapping initial and final states,

$$\mathbf{R}_{10} = \mathbf{J}\mathbf{R}_{01}\mathbf{J} = \begin{pmatrix} \mathbf{Q}_t & -\mathbf{Q}_t e^{\mathbf{A}t} \\ -e^{\mathbf{A}^T t}\mathbf{Q}_t & e^{\mathbf{A}^T t}\mathbf{Q}_t e^{\mathbf{A}t} + \mathbf{Q}_0 \end{pmatrix}, \qquad (16)$$

with

$$\mathbf{J} = \begin{pmatrix} 0 & \mathbf{I} \\ \mathbf{I} & 0 \end{pmatrix}, \qquad (17)$$

and $\mathbf{I}$ is the $N \times N$ identity matrix.

The irreversibility r is now simply obtained from Eq. (2), $r(z) = z^T \mathbf{R} z / 2$ with the irreversibility concentration matrix $\mathbf{R} = \mathbf{R}_{10} - \mathbf{R}_{01}$. One can show that when detailed balance is satisfied, the irreversibility concentration matrix $\mathbf{R}$ is identically zero and the system is reversible. The irreversibility can be described by the eigenvalues and eigenvectors of $\mathbf{R}$. A $2N$-dimensional eigenvector v_n with eigenvalue λ_n can be written in terms of an N-dimensional initial end point v_{n0} and an N-dimensional final end point v_{n1}, $v_n^T = (v_{n0}^T, v_{n1}^T)$. Due to the structure of $\mathbf{R}$, each eigenvector has a corresponding time-reversed eigenvector with reversed end points, (v_{n1}^T, v_{n0}^T), with eigenvalue $-\lambda_n$.

IV. IRREVERSIBILITY DISTRIBUTION

The probability density function of irreversibility $p_r(r)$ can be obtained in terms of its characteristic function $\hat{p}_r(k)$,

$$\hat{p}_r(k) = \langle e^{ikr} \rangle = \int dr\, e^{ikr} p_r(r),$$

$$p_r(r) = \frac{1}{2\pi} \int dk\, e^{-ikr} \hat{p}_r(k). \qquad (18)$$

Using the result from the preceding section, the irreversibility pdf can be written as

$$p_r(r) = \int d^{2N}z\, \delta(r - z^T \mathbf{R} z / 2) p(x_0, x_1). \qquad (19)$$

Writing the δ function in terms of its Fourier transform $\delta(r) = 1/(2\pi)\int dk \exp(ikr)$, changing the order of integration, and comparing with Eq. (18) gives the characteristic function

$$\hat{p}_r(k) = \mathcal{N} \int d^{2N}z\, \exp[-z^T(\mathbf{R}_{01} - ik\mathbf{R})z / 2]. \qquad (20)$$

Since $\mathbf{R}$ is symmetric and $\mathbf{R}_{01}$ is positive definite, the above expression can be simplified using the dual conjunctive diagonalization [27]. This diagonalizes the matrix in the exponential, transforming the expression into a product of one-dimensional integrals that can be easily carried out. Define the matrix $\mathbf{W} = \mathbf{R}_{01}^{-1}\mathbf{R}$, define the matrix $\mathbf{S}^{-1}$ to be the matrix of eigenvectors of $\mathbf{W}$, and define the diagonal matrix Λ to be the matrix whose elements are the $2N$ eigenvalues of $\mathbf{W}$, λ_n. Then $\mathbf{R}_{01} = \mathbf{S}^T\mathbf{S}$ and $\mathbf{R} = \mathbf{S}^T\Lambda\mathbf{S}$ and Eq. (20) can be integrated to obtain the characteristic function of the irreversibility pdf,

$$\hat{p}_r(k) = \frac{1}{\sqrt{\mathrm{Det}(\mathbf{I}_{2N} - ik\Lambda)}} = \frac{1}{\prod_{n=1}^{2N} \sqrt{1 - ik\lambda_n}}, \qquad (21)$$

where we have made use of the fact that $\mathrm{Det}(\mathbf{R}_{01})\mathrm{Det}(\mathbf{C}_0\mathbf{C}_t) = 1$. Thus, the characteristic function is

completely determined by the eigenvalues of **W**.

The covariance of the forward trajectory probability $\mathbf{R}_{01}^{-1}$, needed to compute **W**, can be written in a relatively simple form using Eq. (9),

$$\mathbf{R}_{01}^{-1} = \begin{pmatrix} \mathbf{C}_0 & \mathbf{C}_0 e^{\mathbf{A}^T t} \\ e^{\mathbf{A}t}\mathbf{C}_0 & \mathbf{C}_0 \end{pmatrix}. \qquad (22)$$

We have thus obtained a closed-form expression for $\hat{p}_r(k)$ in terms of the eigenvalues of **W**, which is an explicit function of the deterministic dynamics **A**, the steady-state and finite-time covariance matrices $\mathbf{C}_0$ and $\mathbf{C}_t$, and their inverses. The covariance matrices depend only on the deterministic dynamics and the diffusion matrix. The full pdf of irreversibility is then obtained by integrating Eq. (18) using Eq. (21). No closed form solution for the integral has been found.

The eigenvalues of **W** have an interesting interpretation. In the coordinate system $\hat{z} = \mathbf{S}z$, the trajectory probability $p(\hat{z}) \sim \exp(-|\hat{z}|^2/2)$, and the irreversibility is $r = \Sigma_i \lambda_i \hat{z}_i^2/2$. Thus, the eigenvalue λ_i of **W** gives the irreversibility weighting for coordinate i in a coordinate system where the trajectory probability is isotropic and has unit covariance.

The moments of the irreversibility pdf can be obtained by considering the cumulants of the pdf [28]. The cumulant function $g(k) = \ln \hat{p}(k)$ is

$$g(k) = -\frac{1}{2}\sum_{n=1}^{2N} \ln(1 - ik\lambda_n), \qquad (23)$$

and the cumulants are then

$$\kappa_m = (-i)^m \left.\frac{d^m g(k)}{dk^m}\right|_{k=0} = \frac{(m-1)!}{2}\mathrm{Tr}(\mathbf{W}^m). \qquad (24)$$

Using the relationship between cumulants and centered moments gives the moments of the pdf. The first four centered moments are

$$\langle r \rangle = \frac{1}{2}\mathrm{Tr}(\mathbf{W}),$$

$$\langle (r - \langle r \rangle)^2 \rangle = \frac{1}{2}\mathrm{Tr}(\mathbf{W}^2),$$

$$\langle (r - \langle r \rangle)^3 \rangle = \mathrm{Tr}(\mathbf{W}^3),$$

$$\langle (r - \langle r \rangle)^4 \rangle = 3\,\mathrm{Tr}(\mathbf{W}^4) + \frac{3}{4}\mathrm{Tr}^2(\mathbf{W}^2). \qquad (25)$$

The irreversibility pdf is not, in general, Gaussian. The distribution is skewed and has kurtosis different from the Gaussian value. Such non-Gaussian distributions have been previously seen in a variety of nonequilibrium systems [29–33]. Considering the definition of r, Eq. (2), and writing moments of the irreversibility as

$$\langle r^n \rangle = \int d^N\boldsymbol{x}_0 d^N\boldsymbol{x}_1 r^n(\boldsymbol{x}_0,\boldsymbol{x}_1)p(\boldsymbol{x}_0,\boldsymbol{x}_1), \qquad (26)$$

one sees that all (noncentered) moments are non-negative. The average irreversibility can be simplified to

$$\langle r \rangle = \mathrm{Tr}[\mathbf{C}_0\mathbf{Q}_t(\mathbf{I} - e^{2\mathbf{A}t}) - \mathbf{I}]. \qquad (27)$$

It is interesting to consider how the irreversibility scales with the strength of the noise. Consider taking a system and increasing the diffusion matrix by a constant factor α. Then the covariance matrices increase by α, and the concentration matrices are scaled by $1/\alpha$. The irreversibility for a segment with fixed end points scales by $1/\alpha$, but since the covariance scales by α, the probability of finding those end points scales. As a result, the irreversibility pdf is unchanged by the scaling. Note that this invariance is only valid for multiplication by a scalar. A matrix transformation of the noise will, in general, change the dynamics and the irreversibility. The irreversibility is coordinate invariant, so we are free to consider any convenient coordinate system. If one considers a coordinate system where **D** is diagonal, obtainable by an orthogonal transformation, then the system can be considered to be coupled to N thermal reservoirs with temperatures given by the eigenvalues of **D**. Then, the irreversibility statistics are unchanged by changing the temperature of all heat baths by a constant factor. Further, if one considers the coordinate system where **D**=**I**, obtainable by a nonorthogonal transformation, then the system can be considered to have all N degrees of freedom coupled to a single thermal reservoir with unit temperature. This coordinate system makes explicit the fact that all coordinate-invariant properties, including the irreversibility, are independent of the temperature of this single reservoir. Thus, irreversibility is neither a measure of the temperature of the reservoir, nor a measure of the amplitude of the fluctuations, but rather is related to the amplification of the noise above that seen in equilibrium.

V. TIME DEPENDENCE OF AVERAGE IRREVERSIBILITY

The time dependence of the average irreversibility gives information about the time dependence of the fluctuations. Because the dynamics is stable, the following properties hold: the eigenvalues of **A** all have negative real part, $\lim_{t\to\infty} e^{\mathbf{A}t} = 0$, $\lim_{t\to\infty}\mathbf{C}_t = \mathbf{C}_0$, and $\lim_{t\to\infty}\mathbf{Q}_t = \mathbf{Q}_0$. Thus, Eqs. (14) and (16) give

$$\lim_{t\to\infty}\mathbf{R}_{01} = \lim_{t\to\infty}\mathbf{R}_{10} = \begin{pmatrix} \mathbf{Q}_0 & 0 \\ 0 & \mathbf{Q}_0 \end{pmatrix}. \qquad (28)$$

This is easily understood in that as the time becomes large, the initial and final states become uncorrelated and the trajectory probability is merely the product of the steady-state probabilities of the initial and final states. As a result, $\lim_{t\to\infty}\mathbf{R} = 0$, and the irreversibility of long trajectory segments goes to zero.

For short times $t = \tau \ll 1$, one can use the asymptotic expansion for $e^{\mathbf{A}\tau}$ to obtain expressions for $\mathbf{C}_\tau$ and $\mathbf{Q}_\tau$. This allows one to write the average irreversibility as

$$\langle r \rangle = \frac{\tau}{2}\mathrm{Tr}[\mathbf{A}\mathbf{C}_0(\mathbf{A}^T\mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{A})] + O(\tau^2). \qquad (29)$$

Thus, as $t \to 0$, $\langle r \rangle \to 0$. Further, for short times the dependence of $\langle r \rangle$ on the violation of detailed balance, which is equivalent to $\mathbf{A}^T\mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{A} \neq 0$, is manifest.

Thus, the average irreversibility is positive semidefinite for all time and goes to zero as $t \to 0$ and $t \to \infty$. This means that either $\langle r \rangle(t) = 0$ for all t and the system is in equilibrium, or there is a time t^* where the average irreversibility reaches a global maximum. This defines a characteristic time scale of the irreversible fluctuations.

Fluctuation theorems are often formulated in terms of the entropy production [1]. Chernyak *et al.* [3] speculate that for multidimensional systems such as Eq. (1), the time average entropy increase in the thermal reservoirs coupled to the system over a time t is, in our notation,

$$\Sigma_t = \frac{1}{t} \int_0^t d\mathbf{x}^T \mathbf{D}^{-1} \mathbf{\Omega} \mathbf{x}, \tag{30}$$

where the integral is interpreted in the Stratonovich sense. Using the equation of motion (1) to write $d\mathbf{x}$ in terms of dt and taking the average one obtains

$$\langle \Sigma_t \rangle = \frac{1}{t} \int_0^t ds \langle \mathbf{x}^T(s) \mathbf{A}^T \mathbf{D}^{-1} \mathbf{\Omega} \mathbf{x}(s) \rangle$$
$$+ \frac{1}{t} \int_0^t ds \langle \boldsymbol{\xi}^T(s) \mathbf{F}^T \mathbf{D}^{-1} \mathbf{\Omega} \mathbf{x}(s) \rangle. \tag{31}$$

The first term is simplified by considering $\mathbf{x}(s)$ to be in the steady state so $\langle x_i(s) x_j(s) \rangle = C_{0_{ij}}$. The second term can be simplified to $\mathrm{Tr}(\mathbf{\Omega}) = 0$. Manipulating the result using the definition of $\mathbf{\Omega}$ and Eq. (7) one finds that the average steady-state rate of entropy production is identical to the zero-time growth rate of the average trajectory irreversibility given by Eq. (29),

$$\langle \Sigma_t \rangle = \left. \frac{d\langle r \rangle(t)}{dt} \right|_{t=0}. \tag{32}$$

The entropy production can be related to the noise amplification that occurs when detailed balance is violated. First consider one-dimensional dynamics governed by deterministic dynamics $A < 0$, diffusion constant $D > 0$, and steady-state covariance $C_0 > 0$. Then Eq. (7) gives $-(AC_0/D + 1) = 0$. The negative sign is chosen so that amplification will correspond to a positive number. If, through some additional forcing, the covariance were increased to $C' > C_0$, then one measure of that additional amplitude is $-(AC'/D + 1) > 0$. In multidimensional systems, increased variance occurs not through some additional forcing, but through violation of detailed balance [17]. One analogous measure for noise amplification is the nondimensional gain matrix $\mathbf{G} = -(\mathbf{A} C_0 \mathbf{D}^{-1} + \mathbf{I})$. Using Eq. (7) allows one to write Σ_t as

$$\langle \Sigma_t \rangle = \mathrm{Tr}(\mathbf{A} \mathbf{G}). \tag{33}$$

The entropy production rate has units of 1/time, and the time scale of the system is set by the deterministic dynamics $\mathbf{A}$, Thus, $\mathbf{A}\mathbf{G}$ is a matrix measure of the noise amplification per unit time, whose trace gives the steady-state rate of entropy production.

VI. REDUCTION TO SLOW MODES

Complex spatiotemporal systems are typically high-dimensional systems with a wide range of time scales. One often would like to reduce the system down to a more manageable number of degrees of freedom. There are several common truncations but none are entirely satisfactory. Here we explore dimensional reduction based on a separation of time scales and examine the effect on irreversibility.

We have already seen that for times longer than the longest deterministic time scale the irreversibility becomes zero. It is thus not unreasonable to hypothesize that those degrees of freedom whose time scale is much shorter than the time span of a trajectory segment will have no affect on its irreversibility. If this were the case, then in considering the irreversibility of trajectory segments of a particular time span, one could reduce the dimensionality to those degrees of freedom whose time scales are similar to or longer than the segment time span. We now show that when there is a separation of time scales in the dynamics this is indeed the case and the fast modes have a small effect on the irreversibility statistics.

Climate subsystems often have slow modes which comprise a reduced dimensional subspace. Here we shall investigate the consequences of the deterministic dynamics having a separation of time scales. Note that by modeling the system as an N-dimensional stochastic dynamical system we have already assumed a separation into N slow modes, considered to be deterministic, and unresolved fast modes which are parametrized as random noise. Now we decompose the slow deterministic modes into slower modes and faster modes. The faster modes are fast compared to the slower modes, but still slow compared to the random noise. More specifically, assume that the N eigenvalues λ_i' of $\mathbf{A}$ can be divided into two groups, N_s slow modes and $N_f = N - N_s$ fast modes, where $\mathrm{Re}(\lambda_i') \sim O(1)$ for $1 \leq i \leq N_s$, and $\mathrm{Re}(\lambda_i') \sim O(1/\epsilon)$ for $N_s < i \leq N$, and $\epsilon \ll 1$. Further, we restrict ourselves to times t that are of the order of 1, so that $\mathrm{Re}(\lambda_{\mathrm{fast}}')t \sim O(1/\epsilon)$. We shall rescale the fast eigenvalues so that the small parameter is explicit, $\lambda_{\mathrm{slow}} = \lambda_{\mathrm{slow}}'$, $\lambda_{\mathrm{fast}} = \lambda_{\mathrm{fast}}' \epsilon$, and the real parts are all $O(1)$.

We will work in a coordinate system where $\mathbf{A}$ is diagonal,

$$\mathbf{A} = \begin{pmatrix} \Lambda_s & 0 \\ 0 & \Lambda_f/\epsilon \end{pmatrix}, \tag{34}$$

where Λ_s is a diagonal matrix of slow eigenvalues and Λ_f is a diagonal matrix of scaled fast eigenvalues. Since the eigenvalues will typically be complex, the coordinate transformation is also complex. The various equations presented above must then be modified by changing transpose operators to adjoint operators, indicated with a superscript dagger. For more details on coordinate transformations in stochastic linear systems see [17]. In these coordinates, the state space variable $\mathbf{x}$ can be decomposed into slow variables $\mathbf{y}$ and fast variables $\mathbf{w}$, $\mathbf{x}^\dagger = (\mathbf{y}^\dagger, \mathbf{w}^\dagger)$.

In these diagonal-$\mathbf{A}$ coordinates, the covariance matrix can be obtained by direct integration of Eq. (6) to obtain

$$C_{t_{ij}} = \frac{2D_{ij}}{\lambda_i' + \lambda_j'^*}\{\exp[(\lambda_i' + \lambda_j'^*)t] - 1\}, \qquad (35)$$

where the asterisk denotes complex conjugate and the eigenvalues are unscaled. For the order of 1 times, $\exp(\lambda_{\text{fast}}'t) \sim \exp(-1/\epsilon) \approx 0$, and the deterministic Green function is

$$e^{\mathbf{A}t} = \begin{pmatrix} e^{\Lambda_s t} & 0 \\ 0 & 0 \end{pmatrix}. \qquad (36)$$

Then, the covariance matrices take the form

$$\mathbf{C}_t = \begin{pmatrix} \mathbf{C}_{tyy} & \epsilon\mathbf{C}_{0yw} \\ \epsilon\mathbf{C}_{0wy} & \epsilon\mathbf{C}_{0ww} \end{pmatrix}, \quad \mathbf{C}_0 = \begin{pmatrix} \mathbf{C}_{0yy} & \epsilon\mathbf{C}_{0yw} \\ \epsilon\mathbf{C}_{0wy} & \epsilon\mathbf{C}_{0ww} \end{pmatrix}, \quad (37)$$

where the subscripts y or w indicate fast or slow modes, the subscripts t or 0 indicate whether the covariance submatrix is time dependent or steady state, and all submatrices are $O(1)$. Thus, the time-dependent covariance matrix is decomposed into an $O(1)$ slow-slow covariance which depends on time, and its complement which is small and equal to the steady-state value. Note that this analysis assumes that all components of the diffusion matrix $\mathbf{D}$ are $O(1)$. It may be the case that the fast modes have larger noise than the slow modes. If the noise in the fast modes is $O(1/\epsilon)$ then this asymptotic expansion breaks down.

Using this decomposition in the equations leading to $\mathbf{R}$ and $\mathbf{W}$, and making frequent use of matrix identities for block matrices, leads to the following results. For any given trajectory segment with end points $x_0^\dagger = (y_0^\dagger, w_0^\dagger)$, $x_1^\dagger = (y_1^\dagger, w_1^\dagger)$, the irreversibility $z^\dagger \mathbf{R} z$ depends on both the fast and slow variables. On the other hand, the characteristic function for the irreversibility pdf depends to first order only on the slow variables. Thus, the statistics of the irreversibility are, to lowest order, unaffected by truncating the fast variables. These two results may seem contradictory. However, the fast eigenvalues have two effects: the correlations decay rapidly and, as seen in Eq. (37), the covariance of the fast modes is small. Thus, the typical size of the fast variables is small and its contribution to the irreversibility is a higher-order effect. While large rare fluctuations in the fast modes do effect the irreversibility of isolated trajectory segments, they do not, to lowest order, impact the statistics. Thus, in considering irreversibility, and provided the diffusion matrix is $O(1)$, one can safely neglect fast modes and reduce the dimensionality of the dynamics to just the slow modes of the system.

VII. DISCUSSION

Fluctuation theorems in nonequilibrium systems have focused attention on the irreversibility of trajectory segments, and provide a constraint for their distributions. Theories of climate subsystems provide motivation for analyzing the nonequilibrium fluctuations of linear stochastic dynamical systems in more detail. Here we combined these two perspectives and obtained information about the trajectory irreversibility that is not constrained by the fluctuation theorem.

For linear stochastic dynamics with additive Gaussian white noise, the irreversibility was shown to be governed by an irreversibility concentration matrix $\mathbf{R}$ which is expressed in terms of the fundamental matrices governing the dynamics. The moments of the irreversibility pdf can be written explicitly and the pdf is seen to be non-Gaussian. For nonequilibrium steady states, the average irreversibility grows from zero as the trajectory time span increases, with the initial growth rate being equal to the average entropy production rate, which is related to the noise amplification. In nonequilibrium steady states, there is a finite time where the average irreversibility is maximal, and it decays back to zero as time goes to infinity. For a system with a separation of time scales in the deterministic dynamics and without asymptotically large noise in the fast modes, only the slow modes contribute to the irreversibility statistics, while modes faster than the trajectory time span contribute higher order corrections. There is, however, evidence that in at least some climate subsystems the noise in fast modes is indeed large, and thus this approach to dimensional reduction may not be applicable [5,34].

It is important to note that the irreversibility depends on the multivariate nature of both the deterministic dynamics and the random noise. Nonequilibrium, the violation of detailed balance, and the noncommutivity relation $\mathbf{AD} - \mathbf{DA}^T \neq 0$ are all equivalent, and determining the irreversibility requires knowledge of both parts of the dynamics.

These results were obtained in the idealized context of a linear dynamical system forced by additive Gaussian white noise. However, the definition of the irreversibility and the constraint provided by the fluctuation theorem are general. Thus, one can ask similar questions about the distribution of irreversibility in more complex systems. It remains to be seen which of the above results generalize. In the context of climate, the fact that many climate subsystems can be modeled by such simple stochastic dynamical systems indicates that the properties obtained here are at least approximately valid for these climate subsystems. Further, even for climate phenomena that are not well approximated by such simple models, the behavior of these models provides a null hypothesis for the phenomenon.

Fluctuations in climate subsystems are complex, multidimensional phenomena with many characteristics. It is often convenient to reduce this behavior down to a single scalar, referred to as an index. These indices are usually based on subjectively chosen phenomenological features of the fluctuations. For example, one common index used in El-Niño studies is the NINO3 index, defined as the mean sea surface temperature anomaly from climatology over the region 5°N-5°S, 90°W-150°W [35]. Different indices capture different aspects of a phenomenon, and if the character of the phenomenon changes as climate changes, then any particular index may lose its utility.

The irreversibility provides an interesting alternative to traditional indices. Like an index, irreversibility is a scalar defined in terms of the time series of the system. However, unlike traditional indices, the irreversibility reflects fundamental properties of the nonequilibrium dynamics. The time scale of climate phenomena is often understood in terms of the specific physical properties of the system of interest. El-Niño time scales, for example, involve the time for waves to propagate across the tropical Pacific Ocean. Other time scales such as the growth rates of perturbations based on

singular vector analysis, depend on the subjective coordinate system used. By virtue of its coordinate independence, the time of maximum average irreversibility provides an objective choice for a fluctuation time scale.

A major obstacle to the practical use of irreversibility is that its direct measurement from Eq. (2) will require extremely long time series. In order to compute r directly from a time series, one must quantify the probability of finding rare events, which means one must have a time series long enough to contain those rare events. Even with the long time series produced by numerical models, direct computation of the irreversibility may be prohibitive for all but the very simplest models. There are techniques that have been used to accelerate nonequilibrium computations, but it remains to be seen if these techniques can be applied to climate models. The most promising avenue may be to use climate time series to estimate the parameters of the linear stochastic model in Eq. (1). This is the technique used, for example, in constructing linear inverse models of El-Niño [4]. Then, the irreversibility could be computed using the expressions obtained here.

In many climate subsystems the phenomena modeled here by stationary random noise have a strong seasonal compo-

nent. A better stochastic model for these phenomena would then be cyclostationary noise. Additional complexities to be considered are red-noise processes and multiplicative noise processes. All of these modifications will probably impact the results obtained here.

As climate changes, the nonequilibrium steady-state changes. Many climate subsystems have time scales shorter than the time scale for climate change, and the change in the steady state can be considered to be adiabatic. By obtaining relations between the parameters defining the steady state and the fluctuations, we have solved part of the question of how climate subsystem fluctuations will be affected by climate change. The question of how the steady state itself evolves under climate change remains. However, it may be possible that climate models do a better job of capturing the evolving steady state than the fluctuations. If so, this work may lead to improved climate change forecasts.

ACKNOWLEDGMENT

[1] D. J. Evans and D. J. Searles, Adv. Phys. **51**, 1529 (2002).

[2] D. J. Searles and D. J. Evans, Phys. Rev. E **60**, 159 (1999).

[3] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, J. Stat. Mech.: Theory Exp. (2006) P08001.

[4] C. Penland and T. Magorian, J. Clim. **6**, 1067 (1993).

[5] C. Penland and P. D. Sardeshmukh, J. Clim. **8**, 1999 (1995).

[6] A. M. Moore and R. Kleeman, Q. J. R. Meteorol. Soc. **122**, 1405 (1996).

[7] A. M. Moore and B. F. Farrell, J. Phys. Oceanogr. **23**, 1682 (1993).

[8] B. F. Farrell and P. J. Ioannou, J. Atmos. Sci. **50**, 4044 (1993).

[9] B. F. Farrell and P. J. Ioannou, J. Atmos. Sci. **51**, 2685 (1994).

[10] M. Newman, P. D. Sardeshmukh, and C. Penland, J. Atmos. Sci. **54**, 435 (1997).

[11] J. S. Whitaker and P. D. Sardeshmukh, J. Atmos. Sci. **55**, 237 (1998).

[12] K. M. Weickmann, W. A. Robinson, and C. Penland, J. Geophys. Res. **105**, 15543 (2000).

[13] C. Penland and L. Matrosova, Mon. Weather Rev. **129**, 1740 (2001).

[14] B. F. Farrell and P. J. Ioannou, J. Atmos. Sci. **53**, 2025 (1996).

[15] B. F. Farrell and P. J. Ioannou, J. Atmos. Sci. **53**, 2041 (1996).

[16] P. J. Ioannou, J. Atmos. Sci. **52**, 1155 (1995).

[17] J. B. Weiss, Tellus, Ser. A **55A**, 208 (2003).

[18] E. N. Lorenz, Tellus **17**, 321 (1965).

[19] G. J. van Oldenborgh, S. S. Y. Philip, and M. Collins, Ocean Science **1**, 81 (2005).

[20] G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics* (Cam-

bridge University Press, Cambridge, 2006).

[21] C. Penland, Physica D **98**, 534 (1996).

[22] C. E. Leith, Physica D **98**, 481 (1996).

[23] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden, Proc. Natl. Acad. Sci. U.S.A. **96**, 14687 (1999).

[24] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden, Commun. Pure Appl. Math. **54**, 891 (2001).

[25] A. J. Majda, I. Timofeyev, and E. Vanden-Eijnden, Nonlinearity **19**, 769 (2006).

[26] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).

[27] M. Brooks, http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/intro.html

[28] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Academic, New York, 1975).

[29] D. M. Carberry, J. C. Reid, G. M. Wang, E. M. Sevick, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **92**, 140601 (2004).

[30] F. Ritort, J. Stat. Mech.: Theory Exp. (2004) P10016.

[31] V. Blickle, T. Speck, L. Helden, U. Seifert, and C. Bechinger, Phys. Rev. Lett. **96**, 070603 (2006).

[32] G. E. Crooks and C. Jarzynski, Phys. Rev. E **75**, 021116 (2007).

[33] A. Gomez-Marin and P. Pagonabarraga, Phys. Rev. E **74**, 061113 (2006).

[34] G. P. Compo, P. D. Sardeshmukh, and C. Penland, J. Clim. **14**, 3356 (2001).

[35] K. E. Trenberth, Bull. Am. Meteorol. Soc. **78**, 2771 (1997).