Front. Math. China 2012, 7(2): 199–216
DOI 10.1007/s11464-012-0192-7

RESEARCH ARTICLE

Improved linear response for stochastically driven systems

Rafail V. ABRAMOV

Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA

Abstract The recently developed short-time linear response algorithm, which predicts the average response of a nonlinear chaotic system with forcing and dissipation to small external perturbation, generally yields high precision of the response prediction, although suffers from numerical instability for long response times due to positive Lyapunov exponents. However, in the case of stochastically driven dynamics, one typically resorts to the classical fluctuation-dissipation formula, which has the drawback of explicitly requiring the probability density of the statistical state together with its derivative for computation, which might not be available with sufficient precision in the case of complex dynamics (usually a Gaussian approximation is used). Here, we adapt the short-time linear response formula for stochastically driven dynamics, and observe that, for short and moderate response times before numerical instability develops, it is generally superior to the classical formula with Gaussian approximation for both the additive and multiplicative stochastic forcing. Additionally, a suitable blending with classical formula for longer response times eliminates numerical instability and provides an improved response prediction even for long response times.

Keywords Fluctuation-dissipation theorem, linear response, stochastic processes
MSC 37N10

1 Introduction

The fluctuation-dissipation theorem (FDT) is one of the cornerstones of modern statistical physics. Roughly speaking, the fluctuation-dissipation theorem states that for dynamical systems at statistical equilibrium the

Received February 1, 2011; accepted July 6, 2011
E-mail: abramov@math.uic.edu

average response to small external perturbations can be calculated through the knowledge of suitable correlation functions of the unperturbed dynamical system. The fluctuation-dissipation theorem has great practical use in a variety of settings involving statistical equilibrium of baths of identical gas or liquid molecules, Ornstein-Uhlenbeck Brownian motion, motion of electric charges, turbulence, quantum field theory, chemical physics, physical chemistry, and other areas. The general advantage provided by the fluctuation-dissipation theorem is that one can successfully predict the response of a dynamical system at statistical equilibrium to an arbitrary small external perturbation without ever observing the behavior of the perturbed system, which offers great versatility and insight in understanding behavior of dynamical processes near equilibrium in numerous scientific applications [10,16]. In particular, there has been a profound interest among the atmospheric/ocean science community to apply the fluctuation-dissipation theorem to predict global climate changes responding to variation of certain physical parameters [6–8,11–15,18,22], where the FDT has been used largely in its classical formulation [25]. A vivid demonstration of high predictive skill in low-frequency climate response despite structural instability of statistical states is given in [21].

Recently, Majda and the author [3–5] developed and tested a novel computational algorithm for predicting the mean response of nonlinear functions of states of a chaotic dynamical system to small change in external forcing based on the FDT. The major difficulty in this situation is that the probability measure in the limit as time approaches infinity in this case is typically a Sinai-Ruelle-Bowen probability measure which is supported on a large-dimensional (often fractal) set and is usually not absolutely continuous with respect to the Lebesgue measure [9,29]. In the context of Axiom A attractors, Ruelle [27,28] has adapted the classical calculations for FDT to this setting. The geometric algorithm (also called the short-time FDT, or ST-FDT algorithm in [3–5]) is based on the ideas of [26,28] and takes into account the fact that the dynamics of chaotic nonlinear forced-dissipative systems often reside on chaotic fractal attractors, where the classical FDT formula of the fluctuation-dissipation theorem often fails to produce satisfactory response prediction, especially in dynamical regimes with weak and moderate chaos and slower mixing. It has been discovered in [3–5] that the ST-FDT algorithm is an extremely precise response approximation for short response times, and can be blended with the classical FDT algorithm with Gaussian approximation of the state probability density (quasi-Gaussian FDT algorithm, or qG-FDT) for longer response times to alleviate undesirable effects of expanding Lyapunov directions (which cause numerical instability in ST-FDT for longer response times). Further developing the ST-FDT response algorithm for practical applications, in [1] the author designed a computationally inexpensive method for ST-FDT using the reduced-rank tangent map, and in [2] the ST-FDT algorithm is adapted for the response on slow variables of multi-scale dynamics, which improves its computational stability and simultaneously reduces computational expense.

However, dynamical systems describing real-world processes are often driven by a stochastic forcing. In this setting, the traditional approach is to use the classical FDT algorithm, which computes the linear response to small external forcing as a correlation function along a single long-term trajectory. Typically, it is assumed that the single long-term trajectory samples the statistical equilibrium state of the model, however, suitable generalizations for dynamics with time-periodic forcing can also be made [23,24]. A significant drawback of the classical FDT approach is that its computational algorithm requires the statistical state probability density together with its derivative to be explicitly computed, which is typically not possible for complex nonlinear systems. Usually, an approximation is used, such as the Gaussian approximation with suitable mean state and covariance matrix [3–5]. In this case, if the actual statistical state is far from the Gaussian, the predicted response is usually considerably different from what is observed by direct model perturbation (so called ideal response [3–5]). The reason why the Gaussian approximation is used, is because it is the simplest approximation to use for systems with many degrees of freedom, as only the mean state and covariance matrix should be known to define a Gaussian probability density. Besides, in the absence of any further information the Gaussian probability density maximizes Shannon entropy among all densities with the same mean and covariance, minimizing statistical bias [22].

On the other hand, the ST-FDT response algorithm is observed to be consistently superior to the classical FDT with Gaussian approximation for deterministic chaotic dynamical systems with strongly non-Gaussian statistical states for response times before the numerical instability occurs. In this work, we adapt the ST-FDT linear response algorithm to be used with stochastically forced dynamics (further called stochastic ST-FDT, or SST-FDT). Below, we observe that the SST-FDT response algorithm, adapted to stochastically driven dynamics and blended with the qG-FDT algorithm to avoid numerical instability, is also generally superior to the classical FDT with Gaussian approximation of the statistical state for both the additive and multiplicative noise, just as the ST-FDT algorithm in [3–5] for chaotic deterministic systems.

This paper is organized as follows. In Section 2, we develop the SST-FDT formula for general time-dependent stochastically forced dynamics, and design a practical computational algorithm for autonomous dynamics with invariant probability measure. In Section 3, we test the new algorithm for the stochastically driven Lorenz 96 model [19,20]. Section 4 summarizes the results of this work.

2 Fluctuation-dissipation theorem for stochastically driven systems

Here, we consider an Itô stochastic differential equation (SDE) of the form

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{f}_\alpha(\boldsymbol{x}, t)\mathrm{d}t + \boldsymbol{\sigma}(\boldsymbol{x}, t)\mathrm{d}\boldsymbol{W}_t, \tag{2.1}$$

where $\boldsymbol{x} = \boldsymbol{x}(t) \in \mathbb{R}^N$, $\boldsymbol{f}_\alpha\colon [\mathbb{R}^N \times T] \to \mathbb{R}^N$, $\boldsymbol{\sigma}\colon [\mathbb{R}^{N\times K} \times T] \to \mathbb{R}^N$ are smooth nonlinear functions, and $\boldsymbol{W}_t$ is the K-dimensional Wiener process. Additionally, $\boldsymbol{f}$ depends on a scalar parameter α. We say that the SDE in (2.1) is *unperturbed* if $\alpha = 0$, or *perturbed* otherwise. We also adopt the notation $\boldsymbol{f} \equiv \boldsymbol{f}_0$, with the assumption

$$\frac{\partial}{\partial\alpha}\,\boldsymbol{f}_\alpha(\boldsymbol{x},t)|_{\alpha=0} = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{\eta}(t), \tag{2.2}$$

where $\boldsymbol{B}(\boldsymbol{x})$ is an $N \times L$ matrix-valued function, and $\boldsymbol{\eta}(t)$ is an L-vector valued function. The practical meaning of the above assumption will become clear below.

Let $A(\boldsymbol{x})$ be a nonlinear function of $\boldsymbol{x}$, and let $\mathbb{E}^{t_0,t}_{\boldsymbol{x},\alpha}[A]$, where $t > 0$ is the elapsed time after t_0, denote the expectation of A at time $t_0 + t$ over all realizations of the Wiener process in (2.1), under the condition that $\boldsymbol{x}(t_0) = \boldsymbol{x}$ (with the short notation $\mathbb{E}^{t_0,t}_{\boldsymbol{x},0}[A] = \mathbb{E}^{t_0,t}_{\boldsymbol{x}}[A]$). Let A at the time t_0 be distributed according to a probability measure ρ_{t_0}, that is, the average value of A at time t_0 is

$$\langle A\rangle(t_0) = \rho_{t_0}(A) = \int_{\mathbb{R}^N} A(\boldsymbol{x})\mathrm{d}\rho_{t_0}(\boldsymbol{x}), \tag{2.3}$$

where $\mathrm{d}\rho_{t_0}(\boldsymbol{x})$ denotes the measure of the infinitesimal Lebesgue volume $\mathrm{d}\boldsymbol{x}$ associated with $\boldsymbol{x}$. Then, for time $t_0 + t$, the average of A for the perturbed system in (2.1) is given by

$$\langle A\rangle_\alpha(t_0+t) = \rho_{t_0}(\mathbb{E}^{t_0,t}_{\boldsymbol{x},\alpha}[A]) = \int_{\mathbb{R}^N} \mathbb{E}^{t_0,t}_{\boldsymbol{x},\alpha}[A]\mathrm{d}\rho_{t_0}(\boldsymbol{x}). \tag{2.4}$$

In general, for the same initial distribution ρ_{t_0}, the average value $\langle A\rangle_\alpha(t_0+t)$ depends on the value of α. Here, we define the *average response* $\delta\langle A\rangle_\alpha(t_0+t)$ as

$$\delta\langle A\rangle_\alpha(t_0+t) = \int_{\mathbb{R}^N} (\mathbb{E}^{t_0,t}_{\boldsymbol{x},\alpha}[A] - \mathbb{E}^{t_0,t}_{\boldsymbol{x}}[A])\mathrm{d}\rho_{t_0}(\boldsymbol{x}). \tag{2.5}$$

The meaning of the average response in (2.5) is the following: for the same initial average value of A, it provides the difference between the future average values of A for the perturbed and unperturbed dynamics in (2.1).

If α is small, we can formally linearize (2.5) with respect to α by expanding in Taylor series around $\alpha = 0$ and truncating to the first order, obtaining the following general linear fluctuation-response formula:

$$\delta\langle A\rangle_\alpha(t_0+t) = \alpha\int_{\mathbb{R}^N} \partial_\alpha\mathbb{E}^{t_0,t}_{\boldsymbol{x}}[A]\mathrm{d}\rho_{t_0}(\boldsymbol{x}), \tag{2.6}$$

where we use the short notation

$$\partial_\alpha\bullet \equiv \frac{\partial\bullet}{\partial\alpha}\Big|_{\alpha=0}. \tag{2.7}$$

2.1 Stochastic short-time linear response

To compute the general linear fluctuation-response formula in (2.6), we need a suitable algorithm for $\partial_\alpha \mathbb{E}_{\boldsymbol{x}}^{t_0,t}[A]$. Let

$$x(t_0+t) = \phi_\alpha^{t_0,t}\boldsymbol{x}$$

be the trajectory of (2.1) starting at $\boldsymbol{x}$ at t_0 for a particular realization of the Wiener process $\boldsymbol{W}_{[t_0\dots t_0+t]}$. Then, the expectation $\mathbb{E}_{\boldsymbol{x},\alpha}^{t_0,t}[A]$ is given by

$$\mathbb{E}_{\boldsymbol{x},\alpha}^{t_0,t}[A] = \mathbb{E}[A(\phi_\alpha^{t_0,t}\boldsymbol{x})], \tag{2.8}$$

where the expectation in the right-hand side is taken with respect to all Wiener paths. Therefore,

$$\partial_\alpha \mathbb{E}_{\boldsymbol{x}}^{t_0,t}[A] = \mathbb{E}[\mathrm{D}A(\phi^{t_0,t}\boldsymbol{x})\partial_\alpha \phi^{t_0,t}\boldsymbol{x}], \tag{2.9}$$

where $\mathrm{D}A$ denotes the derivative of A with respect to its argument. For $\partial_\alpha\phi^{t_0,t}\boldsymbol{x}$, by taking the difference between the perturbed and unperturbed versions of (2.1) and linearizing with respect to α at $\alpha = 0$, we have

$$\begin{aligned}\mathrm{d}\partial_\alpha\phi^{t_0,t}\boldsymbol{x} &= (\mathrm{D}\boldsymbol{f}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}t + \mathrm{D}\boldsymbol{\sigma}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}\boldsymbol{W}_{t_0+t})\\ &\quad\cdot\partial_\alpha\phi^{t_0,t}\boldsymbol{x} + \partial_\alpha\boldsymbol{f}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}t,\end{aligned} \tag{2.10}$$

where $\mathrm{D}\boldsymbol{f}$ and $\mathrm{D}\boldsymbol{\sigma}$ are Jacobians of $\boldsymbol{f}$ and $\boldsymbol{\sigma}$, respectively. The above equation is a linear stochastic differential equation for $\partial_\alpha\phi^{t_0,t}\boldsymbol{x}$ with zero initial condition (as at t_0 both perturbed and unperturbed solutions start with the same $\boldsymbol{x}$). It can be solved as follows: let us first introduce the integrating factor $\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}$ (an $N\times N$ matrix) given by the solution of the equation

$$\mathrm{d}\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t} = (\mathrm{D}\boldsymbol{f}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}t + \mathrm{D}\boldsymbol{\sigma}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}\boldsymbol{W}_{t_0+t})\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}, \quad \boldsymbol{T}_{\boldsymbol{x}}^{t_0,0} = \boldsymbol{I}, \tag{2.11}$$

and represent $\partial_\alpha\phi^{t_0,t}\boldsymbol{x}$ as a product

$$\partial_\alpha\phi^{t_0,t}\boldsymbol{x} = \boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}, \tag{2.12}$$

where $\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}$ is an N-vector. Then, for the Itô differential of $\partial_\alpha\phi^{t_0,t}\boldsymbol{x}$ we obtain

$$\begin{aligned}\mathrm{d}\partial_\alpha\phi^{t_0,t}\boldsymbol{x} &= \mathrm{d}\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t} + \boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t} + \mathrm{d}\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}\\ &= (\mathrm{D}\boldsymbol{f}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}t + \mathrm{D}\boldsymbol{\sigma}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}\boldsymbol{W}_{t_0+t})\\ &\quad\cdot\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}(\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t} + \mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}) + \boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}\\ &= (\mathrm{D}\boldsymbol{f}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}t + \mathrm{D}\boldsymbol{\sigma}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}\boldsymbol{W}_{t_0+t})\\ &\quad\cdot\partial_\alpha\phi^{t_0,t}\boldsymbol{x} + (\boldsymbol{I} + \mathrm{D}\boldsymbol{\sigma}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}\boldsymbol{W}_{t_0+t})\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}.\end{aligned} \tag{2.13}$$

Comparing the right-hand sides of (2.10) and (2.13), we find that $\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}$ should satisfy the identity

$$(\boldsymbol{I} + \mathrm{D}\boldsymbol{\sigma}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}\boldsymbol{W}_{t_0+t})\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t} = \partial_\alpha\boldsymbol{f}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}t, \tag{2.14}$$

which holds only if $\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t}$ does not have a diffusion term. As a result, we obtain

$$\mathrm{d}\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t} = (\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t})^{-1}\partial_\alpha \boldsymbol{f}(\phi^{t_0,t}\boldsymbol{x}, t_0+t)\mathrm{d}t, \quad \boldsymbol{y}_{\boldsymbol{x}}^{t_0,0} = 0, \tag{2.15}$$

with the formal solution

$$\boldsymbol{y}_{\boldsymbol{x}}^{t_0,t} = \int_0^t (\boldsymbol{T}_{\boldsymbol{x}}^{t_0,\tau})^{-1}\partial_\alpha \boldsymbol{f}(\phi^{t_0,\tau}\boldsymbol{x}, t_0+\tau)\mathrm{d}\tau. \tag{2.16}$$

Therefore, $\partial_\alpha\phi^{t_0,t}\boldsymbol{x}$ is given by

$$\partial_\alpha\phi^{t_0,t}\boldsymbol{x} = \int_0^t \boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}(\boldsymbol{T}_{\boldsymbol{x}}^{t_0,\tau})^{-1}\partial_\alpha \boldsymbol{f}(\phi^{t_0,\tau}\boldsymbol{x}, t_0+\tau)\mathrm{d}\tau. \tag{2.17}$$

At this point, observe that the solution $\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}$ of (2.11) can be represented as a product

$$\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t} = \boldsymbol{T}_{\phi^{t_0,\tau}\boldsymbol{x}}^{t_0+\tau,t-\tau}\boldsymbol{T}_{\boldsymbol{x}}^{t_0,\tau}, \quad \tau \leqslant t, \tag{2.18}$$

due to the fact that a solution of (2.11) can be multiplied by an arbitrary constant matrix on the right and still remains the solution. Then, (2.17) becomes

$$\partial_\alpha\phi^{t_0,t}\boldsymbol{x} = \int_0^t \boldsymbol{T}_{\phi^{t_0,\tau}\boldsymbol{x}}^{t_0+\tau,t-\tau}\partial_\alpha \boldsymbol{f}(\phi^{t_0,\tau}\boldsymbol{x}, t_0+\tau)\mathrm{d}\tau. \tag{2.19}$$

For smooth $\boldsymbol{f}_\alpha$ and $\boldsymbol{\sigma}$ in (2.1), $\phi^{t_0,t}\boldsymbol{x}$ smoothly depends on $\boldsymbol{x}$ [17], and the integrating factor $\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}$ is in fact the tangent map for the trajectory $\phi^{t_0,t}\boldsymbol{x}$:

$$\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t} = \frac{\partial}{\partial\boldsymbol{x}}\,\phi^{t_0,t}\boldsymbol{x}. \tag{2.20}$$

With (2.19), (2.9) becomes

$$\partial_\alpha\mathbb{E}_{\boldsymbol{x}}^{t_0,t}[A] = \int_0^t \mathbb{E}[\mathrm{D}A(\phi^{t_0,t}\boldsymbol{x})\boldsymbol{T}_{\phi^{t_0,\tau}\boldsymbol{x}}^{t_0+\tau,t-\tau}\partial_\alpha \boldsymbol{f}(\phi^{t_0,\tau}\boldsymbol{x}, t_0+\tau)]\mathrm{d}\tau. \tag{2.21}$$

Recalling (2.2), we write the above formula as

$$\partial_\alpha\mathbb{E}_{\boldsymbol{x}}^{t_0,t}[A] = \int_0^t \mathbb{E}[\mathrm{D}A(\phi^{t_0,t}\boldsymbol{x})\boldsymbol{T}_{\phi^{t_0,\tau}\boldsymbol{x}}^{t_0+\tau,t-\tau}\boldsymbol{B}(\phi^{t_0,\tau}\boldsymbol{x})]\boldsymbol{\eta}(t_0+\tau)\mathrm{d}\tau. \tag{2.22}$$

Then, the general linear response formula in (2.6) can be written as

$$\delta\langle A\rangle_\alpha(t_0+t) = \alpha\int_0^t \boldsymbol{R}_{\mathrm{SST}}(t_0,t,\tau)\boldsymbol{\eta}(t_0+\tau)\mathrm{d}\tau, \tag{2.23}$$

where the *linear response operator* $\boldsymbol{R}_{\mathrm{SST}}(t_0,t,\tau)$ is given by

$$\boldsymbol{R}_{\mathrm{SST}}(t_0,t,\tau) = \mathbb{E}\int_{\mathbb{R}^N} \mathrm{D}A(\phi^{t_0,t}\boldsymbol{x})\boldsymbol{T}_{\phi^{t_0,\tau}\boldsymbol{x}}^{t_0+\tau,t-\tau}\boldsymbol{B}(\phi^{t_0,\tau}\boldsymbol{x})\mathrm{d}\rho_{t_0}(\boldsymbol{x}). \tag{2.24}$$

Further, we refer to (2.24) as the *stochastic short-time fluctuation-dissipation theorem algorithm*, or SST-FDT algorithm. The reason is that in practice the computation of the tangent map in (2.11) for large t becomes numerically unstable because of exponential growth due to positive Lyapunov exponents (just as observed in [2–5] for deterministic chaotic dynamics). Note that if the stochastic forcing is removed from (2.1), the SST-FDT response operator becomes the usual ST-FDT from [2–5]. Apparently, (2.24) requires the average with respect to ρ_{t_0}. If ρ_{t_0} is not known explicitly, there are some opportunities to replace the ρ-average with time average, particularly for the autonomous dynamics with ρ_{t_0} being the invariant probability measure, and also for non-autonomous dynamical systems with explicit time-periodic dependence (as done in [23,24] for classical FDT response).

2.2 Classical linear response

The standard way to derive the classical linear response formula is through the Fokker-Planck equation (or, as it is also called, the forward Kolmogorov equation) for the perturbed system in (2.1) by neglecting the terms of higher order than the perturbation, as it is done in [3–5,22,24,25]. However, for the sake of clarity, here we show the derivation of the classical FDT directly from (2.6). Under the assumption of continuity of ρ_{t_0} with respect to the Lebesgue measure, that is,

$$\mathrm{d}\rho_{t_0}(\boldsymbol{x}) = p_{t_0}(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

where p_{t_0} is the probability density, we can also obtain a formal general expression for the classical fluctuation-response formula. Using the notations

$$\begin{aligned} L_{\mathrm{FP},\alpha}(\boldsymbol{x},t) &= -\frac{\partial}{\partial \boldsymbol{x}}\cdot(\boldsymbol{f}_\alpha(\boldsymbol{x},t)\bullet) + \Big(\frac{\partial}{\partial \boldsymbol{x}}\otimes\frac{\partial}{\partial \boldsymbol{x}}\Big)\cdot(\boldsymbol{\sigma}\boldsymbol{\sigma}^{\mathrm{T}}(\boldsymbol{x},t)\bullet), \\ \mathscr{L}_{\mathrm{FP},\alpha}(\boldsymbol{x},t_0,t) &= \mathscr{T}\exp\bigg(\int_0^t \mathrm{d}\tau L_{\mathrm{FP},\alpha}(\boldsymbol{x},t_0+\tau)\bigg), \end{aligned} \tag{2.25}$$

which are, respectively, the Fokker-Planck operator and its ordered exponential (which is the formal solution of the Fokker-Plank equation), we write the expectation $\mathbb{E}^{t_0,t}_{\boldsymbol{x},\alpha}[A]$ in the form

$$\begin{aligned} \mathbb{E}^{t_0,t}_{\boldsymbol{x},\alpha}[A] &= \int_{\mathbb{R}^N} A(\boldsymbol{y})\mathscr{L}_{\mathrm{FP},\alpha}(\boldsymbol{y},t_0,t)\delta(\boldsymbol{x}-\boldsymbol{y})\mathrm{d}\boldsymbol{y} \\ &= \int_{\mathbb{R}^N} \mathscr{L}^{\dagger}_{\mathrm{FP},\alpha}(\boldsymbol{y},t_0,t)A(\boldsymbol{y})\delta(\boldsymbol{x}-\boldsymbol{y})\mathrm{d}\boldsymbol{y} \\ &= \mathscr{L}^{\dagger}_{\mathrm{FP},\alpha}(\boldsymbol{x},t_0,t)A(\boldsymbol{x}), \end{aligned} \tag{2.26}$$

where $\delta(\boldsymbol{x})$ is the Dirac delta-function, and the adjoint is taken with respect to the standard inner product under the integral. Below, we use the notations

$$
\begin{aligned}
L_{\mathrm{FP}}(\boldsymbol{x},t) &\equiv L_{\mathrm{FP},0}(\boldsymbol{x},t), \quad \partial_\alpha L_{\mathrm{FP}}(\boldsymbol{x},t) \equiv \frac{\partial L_{\mathrm{FP},\alpha}(\boldsymbol{x},t)}{\partial\alpha}\Big|_{\alpha=0},\\
\mathscr{L}_{\mathrm{FP}}(\boldsymbol{x},t_0,t) &\equiv \mathscr{L}_{\mathrm{FP},0}(\boldsymbol{x},t_0,t), \quad \partial_\alpha \mathscr{L}_{\mathrm{FP}}(\boldsymbol{x},t_0,t) \equiv \frac{\partial \mathscr{L}_{\mathrm{FP},\alpha}(\boldsymbol{x},t_0,t)}{\partial\alpha}\Big|_{\alpha=0}.
\end{aligned} \tag{2.27}
$$

Then, the general response formula with

$$
\mathrm{d}\rho_{t_0}(\boldsymbol{x}) = p_{t_0}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}
$$

becomes

$$
\begin{aligned}
\delta\langle A\rangle_\alpha(t_0+t) &= \alpha\int_{\mathbb{R}^N} \partial_\alpha \mathbb{E}_{\boldsymbol{x}}^{t_0,t}[A]p_{t_0}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\
&= \alpha\int_{\mathbb{R}^N} \partial_\alpha \mathscr{L}_{\mathrm{FP}}^\dagger(\boldsymbol{x},t_0,t)A(\boldsymbol{x})p_{t_0}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\
&= \alpha\int_{\mathbb{R}^N} A(\boldsymbol{x})\partial_\alpha \mathscr{L}_{\mathrm{FP}}(\boldsymbol{x},t_0,t)p_{t_0}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{2.28}
$$

It is not difficult to show that the parametric derivative of an ordered exponential of a linear operator $L_\alpha(\boldsymbol{x},t)$ is computed as

$$
\begin{aligned}
&\frac{\partial}{\partial\alpha}\mathscr{T}\exp\left(\int_{t_0}^t \mathrm{d}\tau L_\alpha(\boldsymbol{x},\tau)\right)\\
&= \int_{t_0}^t \mathrm{d}\tau\mathscr{T}\exp\left(\int_\tau^t \mathrm{d}s L_\alpha(\boldsymbol{x},s)\right)\frac{\partial L_\alpha(\boldsymbol{x},\tau)}{\partial\alpha}\mathscr{T}\exp\left(\int_{t_0}^\tau \mathrm{d}s\, L_\alpha(\boldsymbol{x},s)\right).
\end{aligned} \tag{2.29}
$$

As a result, we obtain

$$
\begin{aligned}
\delta\langle A\rangle_\alpha(t_0+t) &= \alpha\int_0^t \mathrm{d}\tau\int_{\mathbb{R}^N} \mathscr{L}_{\mathrm{FP}}^\dagger(\boldsymbol{x},t_0+\tau,t-\tau)\\
&\qquad\cdot A(\boldsymbol{x})\partial_\alpha L_{\mathrm{FP}}(\boldsymbol{x},t_0+\tau)\mathscr{L}_{\mathrm{FP}}(\boldsymbol{x},t_0,\tau)p_{t_0}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\
&= \alpha\int_0^t \mathrm{d}\tau\int_{\mathbb{R}^N} \mathbb{E}_{\boldsymbol{x}}^{t_0+\tau,t-\tau}[A]\partial_\alpha L_{\mathrm{FP}}(\boldsymbol{x},t_0+\tau)p_{t_0+\tau}(\boldsymbol{x})\mathrm{d}\boldsymbol{x},
\end{aligned} \tag{2.30}
$$

where $p_{t_0+\tau}(\boldsymbol{x})$ is given by

$$
p_{t_0+\tau}(\boldsymbol{x}) = \mathscr{L}_{\mathrm{FP}}(\boldsymbol{x},t_0,\tau)p_{t_0}(\boldsymbol{x}). \tag{2.31}
$$

Recalling (2.2), we recover the classical linear fluctuation-response formula in the form

$$
\begin{aligned}
\delta\langle A\rangle_\alpha(t_0+t) &= \alpha\int_0^t \mathrm{d}\tau\int_{\mathbb{R}^N} \mathbb{E}_{\boldsymbol{x}}^{t_0+\tau,t-\tau}[A]\partial_\alpha L_{\mathrm{FP}}(\boldsymbol{x},t_0+\tau)p_{t_0+\tau}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}\\
&= \alpha\int_0^t \boldsymbol{R}_{\mathrm{class}}(t_0,t,\tau)\boldsymbol{\eta}(t_0+\tau)\mathrm{d}\tau,
\end{aligned} \tag{2.32}
$$

where the classical linear response operator $\boldsymbol{R}_{\mathrm{class}}$ is given by

$$
\boldsymbol{R}_{\mathrm{class}}(t_0,t,\tau) = -\mathbb{E}\int_{\mathbb{R}^N} A(\phi^{t_0+\tau,t-\tau}\boldsymbol{x})\frac{\partial}{\partial\boldsymbol{x}}\cdot(\boldsymbol{B}(\boldsymbol{x})p_{t_0+\tau}(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}. \tag{2.33}
$$

Observe that, unlike (2.24), in (2.33) one has to know $p_{t_0+\tau}(\boldsymbol{x})$ for all response times explicitly to perform differentiation with respect to $\boldsymbol{x}$. Usually, an approximation is used, such as the Gaussian approximation [3–5].

2.3 Special case for autonomous dynamics with ergodic invariant probability measure

Here, we consider the case where $\boldsymbol{f}$ and $\boldsymbol{\sigma}$ in (2.1) do not explicitly depend on t (although $\boldsymbol{f}_\alpha$ does with $\alpha \neq 0$), and we choose $\rho_{t_0} = \rho$ to be an ergodic invariant probability measure for (2.1). In this situation, one can replace the averaging with respect to the measure ρ with averaging over a single long-term trajectory which starts with an initial condition $\boldsymbol{x}$ in the support of ρ:

$$\begin{aligned}\boldsymbol{R}_{\mathrm{SST}}(t_0,t,\tau) = \mathbb{E}\lim_{r\to\infty}\frac{1}{r}\int_0^r &\mathrm{D}A(\phi^{t_0,t}\phi^{t_0-s,s}\boldsymbol{x})\\ &\cdot \boldsymbol{T}^{t_0+\tau,t-\tau}_{\phi^{t_0,\tau}\phi^{t_0-s,s}\boldsymbol{x}}\boldsymbol{B}(\phi^{t_0,\tau}\phi^{t_0-s,s}\boldsymbol{x})\mathrm{d}s, \end{aligned}\tag{2.34}$$

where, without loss of generality, the starting is time t_0-s, that is, the averaging occurs over the endpoints of $\phi^{t_0-s,s}\boldsymbol{x}$. Combining the solution operators, we obtain

$$\boldsymbol{R}_{\mathrm{SST}}(t_0,t,\tau) = \mathbb{E}\lim_{r\to\infty}\frac{1}{r}\int_0^r \mathrm{D}A(\phi^{t_0-s,s+t}\boldsymbol{x})\boldsymbol{T}^{s+\tau,t-\tau}_{\phi^{t_0-s,s+\tau}\boldsymbol{x}}\boldsymbol{B}(\phi^{t_0-s,s+\tau}\boldsymbol{x})\mathrm{d}s. \tag{2.35}$$

Since the averaging over all independent realizations of the Wiener process is needed, we can average over many statistically independent chunks of the Wiener path along a single long-time trajectory by setting $t_0 = s-\tau$:

$$\boldsymbol{R}_{\mathrm{SST}}(t,\tau) = \lim_{r\to\infty}\frac{1}{r}\int_0^r \mathrm{D}A(\phi^{-\tau,s+t}\boldsymbol{x})\boldsymbol{T}^{s,t-\tau}_{\phi^{-\tau,s+\tau}\boldsymbol{x}}\boldsymbol{B}(\phi^{-\tau,s+\tau}\boldsymbol{x})\mathrm{d}s. \tag{2.36}$$

Finally, replacing $\boldsymbol{x}$ with $\phi^{0,-\tau}\boldsymbol{x}$ (which for finite τ is also in the support of ρ), we find that

$$\boldsymbol{R}_{\mathrm{SST}}(t,\tau) = \lim_{r\to\infty}\frac{1}{r}\int_0^r \mathrm{D}A(\phi^{0,s+t-\tau}\boldsymbol{x})\boldsymbol{T}^{s,t-\tau}_{\phi^{0,s}\boldsymbol{x}}\boldsymbol{B}(\phi^{0,s}\boldsymbol{x})\mathrm{d}s, \tag{2.37}$$

or, denoting $\boldsymbol{x}(s) = \phi^{0,s}\boldsymbol{x}$,

$$\boldsymbol{R}_{\mathrm{SST}}(t,\tau) = \lim_{r\to\infty}\frac{1}{r}\int_0^r \mathrm{D}A(\boldsymbol{x}(s+t-\tau))\boldsymbol{T}^{s,t-\tau}_{\boldsymbol{x}(s)}\boldsymbol{B}(\boldsymbol{x}(s))\mathrm{d}s. \tag{2.38}$$

Now, the linear response formula in (2.23) and the response operator in (2.24) become, respectively,

$$\begin{aligned}\delta\langle A\rangle_\alpha(t_0+t) &= \alpha\int_0^t \boldsymbol{R}_{\mathrm{SST}}(t-\tau)\boldsymbol{\eta}(t_0+\tau)\mathrm{d}\tau,\\ \boldsymbol{R}_{\mathrm{SST}}(t) &= \lim_{r\to\infty}\frac{1}{r}\int_0^r \mathrm{D}A(\boldsymbol{x}(s+t))\boldsymbol{T}^{s,t}_{\boldsymbol{x}(s)}\boldsymbol{B}(\boldsymbol{x}(s))\mathrm{d}s.\end{aligned}\tag{2.39}$$

In a similar fashion, for the classical linear response in (2.32), we note that the Fokker-Planck operator L_{FP} does not depend on t, and both $\mathscr{L}_{\text{FP}}$ and its adjoint do not depend on t_0. Taking into account that $p_{t_0+\tau}(\boldsymbol{x}) = p(\boldsymbol{x})$, where $p(\boldsymbol{x})$ is the invariant probability density, we write

$$\boldsymbol{R}_{\text{class}}(t) = -\mathbb{E}\int_{\mathbb{R}^N} A(\phi^{t_0,t}\boldsymbol{x})\frac{\partial}{\partial \boldsymbol{x}}\cdot(\boldsymbol{B}(\boldsymbol{x})p(\boldsymbol{x}))\mathrm{d}\boldsymbol{x}, \tag{2.40}$$

or, after replacing the p-average with the average over the long-term trajectory,

$$\boldsymbol{R}_{\text{class}}(t) = -\mathbb{E}\lim_{r\to\infty}\frac{1}{r}\int_0^r A(\boldsymbol{x}(s+t))\frac{\frac{\partial}{\partial \boldsymbol{x}}\cdot(\boldsymbol{B}(\boldsymbol{x}(s))p(\boldsymbol{x}(s)))}{p(x(s))}\,\mathrm{d}s. \tag{2.41}$$

Here, the expectation can be removed since the averaging over different Wiener paths will automatically occur as the long time average is computed. As a result, we obtain

$$\boldsymbol{R}_{\text{class}}(t) = -\lim_{r\to\infty}\frac{1}{r}\int_0^r A(\boldsymbol{x}(s+t))\frac{\frac{\partial}{\partial \boldsymbol{x}}\cdot(\boldsymbol{B}(\boldsymbol{x}(s))p(\boldsymbol{x}(s)))}{p(\boldsymbol{x}(s))}\,\mathrm{d}s. \tag{2.42}$$

3 Application for stochastically driven Lorenz 96 model

The 40-mode deterministic Lorenz 96 model (L96) has been introduced by Lorenz and Emanuel [19,20] as a simple model with large scale features of complex nonlinear geophysical systems. The deterministic Lorenz 96 (L96) model is given by

$$\dot{X}_n = X_{n-1}(X_{n+1} - X_{n-2}) - X_n + F, \quad 1 \leqslant k \leqslant N, \tag{3.1}$$

with periodic boundary conditions given by $X_{n\pm N} = X_n$, where $N = 40$, and F being a constant forcing parameter. The model in (3.1) is designed to mimic midlatitude weather and climate behavior (in particular Rossby waves), so periodic boundary conditions are appropriate. It is demonstrated in [22, Chap. 2] that the dynamical regime of the L96 model varies with changing the value of constant forcing F: weakly chaotic dynamical regimes with $F = 5, 6$, strongly chaotic regime with $F = 8$, and turbulent regimes $F = 12, 16, 24$ with self-similar time autocorrelation decay.

Here, we apply the stochastic forcing to the L96 model as

$$\mathrm{d}X_k = [X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F]\mathrm{d}t + (\boldsymbol{\sigma}(\boldsymbol{X}))_k(\mathrm{d}\boldsymbol{W}_t)_k, \tag{3.2}$$

where $\boldsymbol{\sigma}\colon \mathbb{R}^N \to \mathbb{R}^N$ is a vector-valued function of $\boldsymbol{X}$, $\boldsymbol{W}$ is an N-dimensional Wiener process, and $(\mathrm{d}\boldsymbol{W}_t)_k$ is the k-th component of $\mathrm{d}\boldsymbol{W}$ (that is, effectively $\boldsymbol{\sigma}$ is a diagonal matrix multiplying the vector $\mathrm{d}\boldsymbol{W}$). As the stochastic Lorenz 96 (SL96) model above does not depend explicitly on time (except for the Wiener noise), we can assume that it has an invariant probability measure ρ.

In this work, we perturb the SL96 model in (3.2) by a small parameter α as

$$\mathrm{d}X_k = [X_{k-1}(X_{k+1} - X_{k-2}) - X_k + F + \alpha\eta_k]\mathrm{d}t + (\boldsymbol{\sigma}(\boldsymbol{X}))_k(\mathrm{d}\boldsymbol{W}_t)_k, \quad (3.3)$$

where $\boldsymbol{\eta} \in \mathbb{R}^N$ is a constant forcing vector perturbation, which is "turned on" at time $t_0 = 0$. With the invariant probability state ρ, and the perturbation given in (3.3), the general response formula in (2.6) becomes

$$\begin{aligned} \delta\langle A\rangle_\alpha(t) &= \alpha\mathscr{R}(t)\boldsymbol{\eta}, \\ \mathscr{R}(t) &= \int_0^t \boldsymbol{R}(\tau)\mathrm{d}\tau, \end{aligned} \quad (3.4)$$

where subscripts for $\boldsymbol{R}$ and $\mathscr{R}$ are omitted as both the SST-FDT and classical response operators apply. We also set the observable $A(\boldsymbol{x}) = \boldsymbol{x}$, that is, the response of the mean state is computed. As an approximation for the invariant probability density for the classical response, we choose the Gaussian distribution with the same mean and covariance as the actual invariant probability measure, which are determined by averaging along the long-term time series of unperturbed (3.2), and thus, further call it quasi-Gaussian FDT (qG-FDT) as in [3–5]. In this setting, the short-time and quasi-Gaussian linear response operators become

$$\begin{aligned} \boldsymbol{R}_{\mathrm{SST}}(t) &= \lim_{r\to\infty} \frac{1}{r}\int_0^r \boldsymbol{T}^{s,t}_{\boldsymbol{x}(s)}\mathrm{d}s, \\ \boldsymbol{R}_{\mathrm{qG}}(t) &= \lim_{r\to\infty} \frac{1}{r}\int_0^r \boldsymbol{x}(s+t)\boldsymbol{C}^{-1}(\boldsymbol{x}(s) - \overline{\boldsymbol{x}})\mathrm{d}s, \end{aligned} \quad (3.5)$$

where $\overline{\boldsymbol{x}}$ and $\boldsymbol{C}$ are the mean state and covariance matrix of the long-time series of unperturbed (3.2).

3.1 Blended SST/qG-FDT response

Following [3,5], we also compute the blended SST/qG-FDT response as

$$\boldsymbol{R}_{\mathrm{SST/qG}}(t) = [1 - H(t - t_{\mathrm{cutoff}})]\,\boldsymbol{R}_{\mathrm{SST}}(t) + H(t - t_{\mathrm{cutoff}})\,\boldsymbol{R}_{\mathrm{qG}}(t), \quad (3.6)$$

where the blending function H is the Heaviside step-function. The cut-off time t_{cutoff} is chosen as

$$t_{\mathrm{cutoff}} = \frac{3}{\lambda_1}, \quad (3.7)$$

where λ_1 is the largest Lyapunov exponent (for details, see [3,5]). This cut-off time allows to switch to the $\boldsymbol{R}_{\mathrm{qG}}$ just before the numerical instability occurs in $\boldsymbol{R}_{\mathrm{SST}}$, and thus, avoid the numerical instability. For constant external forcing and the Heaviside blending step-function, the blended response operators become

$$\mathscr{R}_{\mathrm{SST/qG}}(t) = \int_0^{t_{\mathrm{cutoff}}} \boldsymbol{R}_{\mathrm{SST}}(\tau)\mathrm{d}\tau + \int_{t_{\mathrm{cutoff}}}^t \boldsymbol{R}_{\mathrm{qG}}(\tau)\mathrm{d}\tau. \quad (3.8)$$

3.2 Computational experiments

Below, we perform computational experiments in the following setting:

- The number of variables (model size) $N = 40$.
- Constant forcing $F = 6$. The L96 model is observed to be weakly chaotic in this regime [3–5,22], and we would like to compare the responses for weakly chaotic deterministic dynamics and the stochastically driven dynamics.
- The tangent map $\boldsymbol{T}_{\boldsymbol{x}}^{t_0,t}$ in (2.11) is computed in the same fashion as in [2–5].
- Forward Euler numerical scheme with time step $\Delta t = 0.001$ for both (2.11) and (3.2).
- The linear response is tested for the following settings of the stochastic term σ:
  - $\sigma_k = 0$ (fully deterministic regime without stochastic forcing);
  - $\sigma_k = 1$ (additive noise);
  - $\sigma_k = 0.2X_k$, $\sigma_k = 0.5X_k$ (multiplicative noise).
- We compute the linear response operators $\mathscr{R}_{\mathrm{SST}}$, $\mathscr{R}_{\mathrm{qG}}$, and $\mathscr{R}_{\mathrm{SST/qG}}$, which are given by (3.4) and (3.8), and compare them with the ideal response operator $\mathscr{R}_{\mathrm{ideal}}$, which is computed through the direct model perturbations [2–5].
- The time-averaging is done along a time series of 10000 time units.
- The ideal response operator $\mathscr{R}_{\mathrm{ideal}}$ is computed via direct perturbations a 10000-member statistical ensemble.
- The comparison of the FDT response operators with the ideal response operator is carried out by evaluating the L_2 relative error

$$L_2\text{-error} = \frac{\|\mathscr{R}_{\mathrm{FDT}} - \mathscr{R}_{\mathrm{ideal}}\|}{\|\mathscr{R}_{\mathrm{ideal}}\|}, \tag{3.9}$$

and the correlation function

$$\mathrm{Corr} = \frac{(\mathscr{R}_{\mathrm{FDT}}, \mathscr{R}_{\mathrm{ideal}})}{\|\mathscr{R}_{\mathrm{FDT}}\|\,\|\mathscr{R}_{\mathrm{ideal}}\|}, \tag{3.10}$$

where $(\cdot,\cdot)$ denotes the standard Euclidean inner product. Observe that the L_2 error shows the general difference between the FDT and ideal responses, while the correlation function shows the extent to which the responses are collinear (that is, how well the location of the response is determined, without considering its magnitude).

Here, we would like to mention that, unlike the previous works [1–5], where systems of ordinary differential equations studied, here we sample the correlation functions from a stochastic Monte Carlo process, which must also include many realizations of a Wiener process. In practice, this leads to much longer time averaging windows than for systems of ordinary differential equations. However, as can be observed in the results below, the time averaging window of 10000 time units seems to be sufficient to provide adequate statistical sampling for the 40-mode Lorenz model.

In Fig. 1, we display the L_2 relative errors between the ideal response operator and the FDT response operators, together with the intrinsic error in the ideal response operator (which is the result of slight nonlinearity in the ideal response due to small but finite perturbations). Observe that in the fully deterministic regime ($F = 6$, $\sigma_k = 0$) the SST-FDT response provides a very precise prediction until the time $t \approx 3$, and then the errors in the SST-FDT grow exponentially rapidly, which is due to the positive Lyapunov exponents and numerical instability in the tangent map. On the other hand, the qG-FDT response is not precise (reaching about 80% by the time $t = 1.5$), due to the fact that the invariant probability measure associated with the deterministic regime is highly non-Gaussian, and most probably not continuous with respect to the Lebesgue measure (that is, it does not even possess a density). Remarkably, if we look at the stochastically driven regimes $\sigma_k = 1$ (additive noise) and $\sigma_k = 0.2X_k$, $\sigma_k = 0.5X_k$ (multiplicative noise), we see that the behavior of both the SST-FDT and qG-FDT responses is qualitatively the same as in the fully deterministic regime, even though the dynamics is qualitatively different. Apparently, the level of noise in the two stochastically driven regimes $\sigma_k = 1$ and $\sigma_k = 0.2X_k$ is insufficient to "smooth out" the invariant probability

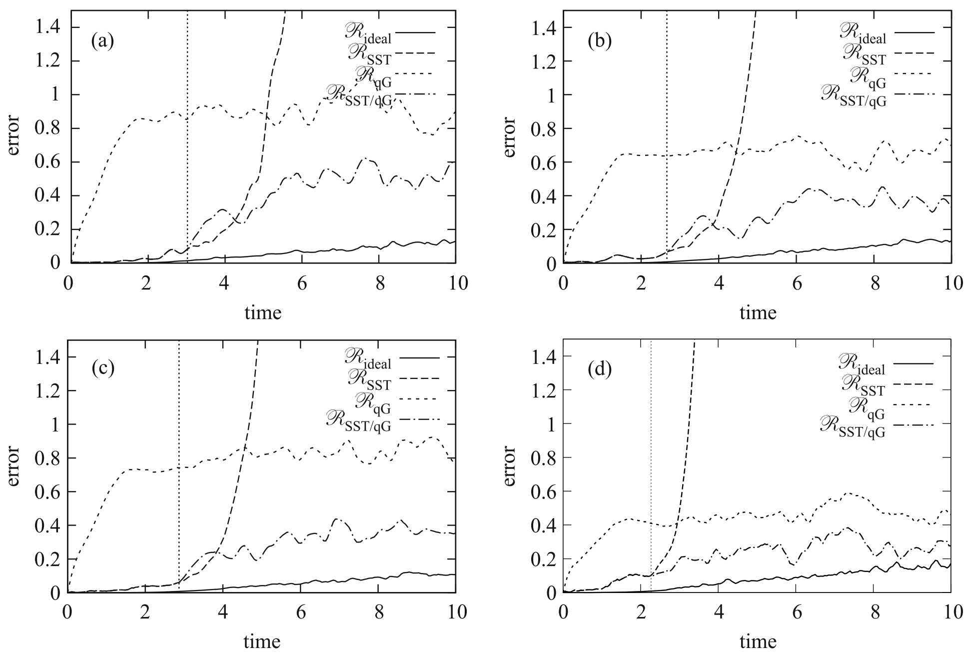

Fig. 1 L_2-errors of response operators for SL96 model, $N = 40$, $F = 6$, (a) $\sigma_k = 0$; (b) $\sigma_k = 1$; (c) $\sigma_k = 0.2X_k$; (d) $\sigma_k = 0.5X_k$. Straight dotted vertical line denotes blending cut-off time for SST/qG-FDT. $\mathscr{R}_{\text{ideal}}$ denotes intrinsic error in ideal response due to slight nonlinearity.

measure enough for it to resemble the Gaussian state and to destabilize the computation of the tangent map. However, in the $\sigma_k = 0.5X_k$ multiplicative noise regime, the errors in the initial qG-FDT response are reduced to about 40%, which is due to the fact that in this regime the invariant probability measure is closer to the Gaussian state because of strong noise. The blended SST/qG-FDT response yields the lowest errors in all cases, due to its explicit design to avoid numerical instability in the SST-FDT algorithm.

In Fig. 2, we show the correlation functions for the same simulations. Observe that, although significant L_2-errors were observed for the qG-FDT algorithm for the fully deterministic regime $\sigma_k = 0$, its correlations with the ideal response are generally on the level of around 0.7, which is remarkable. Also, the correlations of the SST-FDT response with the ideal response are roughly 1 (nearly perfect correlation) before the numerical instability manifests itself. As for the blended SST/qG-FDT response, the best correlations are achieved in the stochastically forced regimes $\sigma_k = 1$ (additive noise) and $\sigma_k = 0.2X_k$, $\sigma_k = 0.5X_k$ (multiplicative noise), were the correlations do not become lower than 0.95 for all response times. For the fully deterministic case $\sigma_k = 0$, the correlations of the blended SST/qG-FDT response are about 0.8.

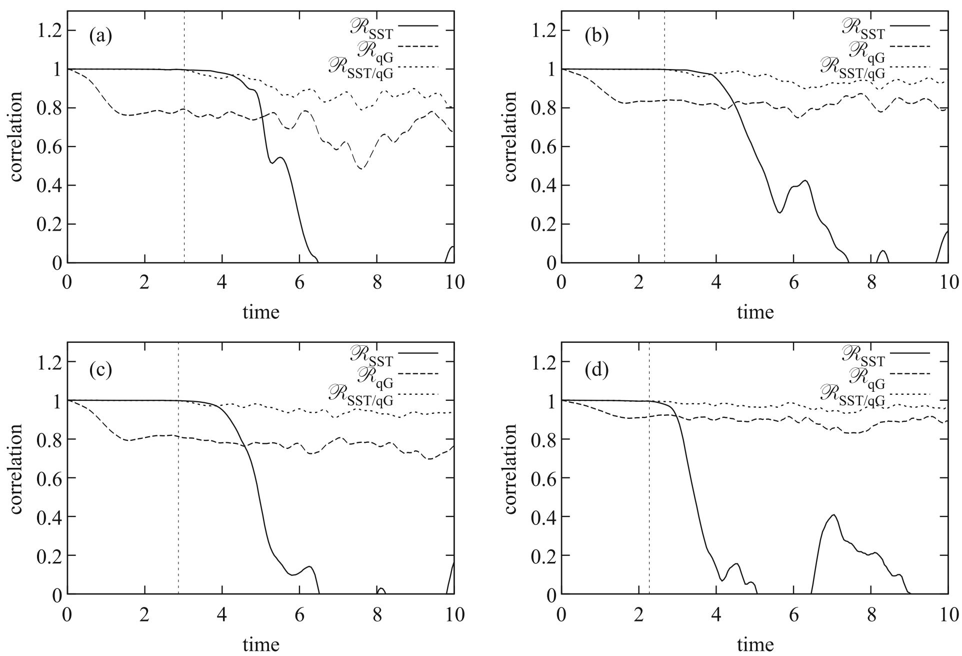

Fig. 2 Correlations of FDT response operators with ideal response operator for SL96 model, $N = 40$, $F = 6$, (a) $\sigma_k = 0$; (b) $\sigma_k = 1$; (c) $\sigma_k = 0.2X_k$; (d) $\sigma_k = 0.5X_k$. Straight dotted vertical line denotes blending cut-off time for SST/qG-FDT.

In addition to displaying the errors and correlations between the FDT response operators and the ideal response operator, in Figs. 3–5, we show the instantaneous snapshots of the linear response operators at times $T = 1$, $T = 2$ (which are before the SST/qG-FDT cutoff time), and $T = 5$ (which is after the

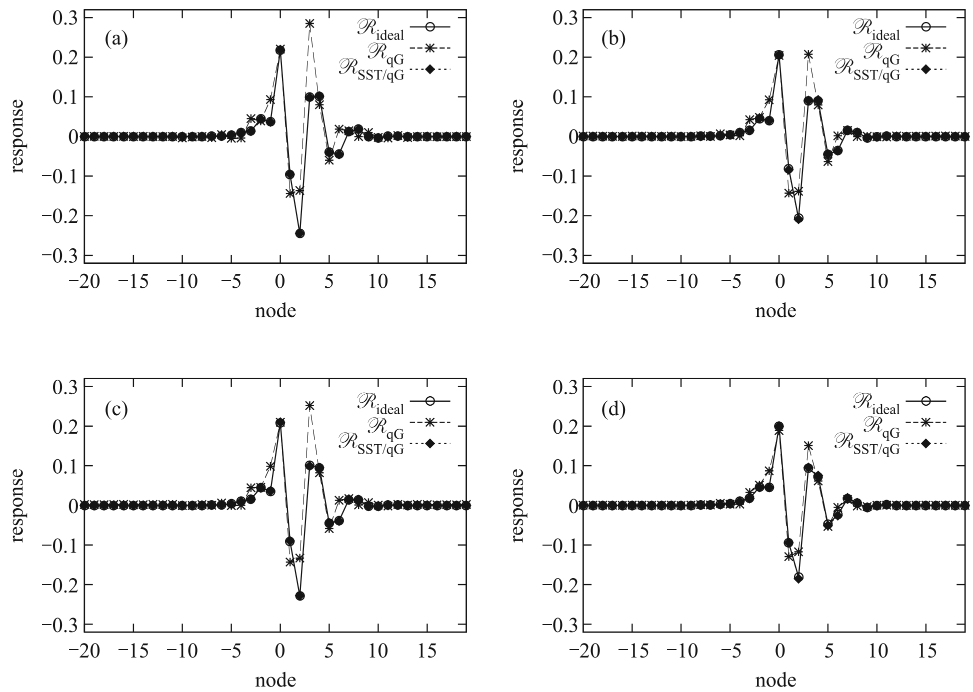

Fig. 3 Snapshots of response operators for SL96 model at $T = 1$, $N = 40$, $F = 6$, (a) $\sigma_k = 0$; (b) $\sigma_k = 1$; (c) $\sigma_k = 0.2X_k$; (d) $\sigma_k = 0.5X_k$.

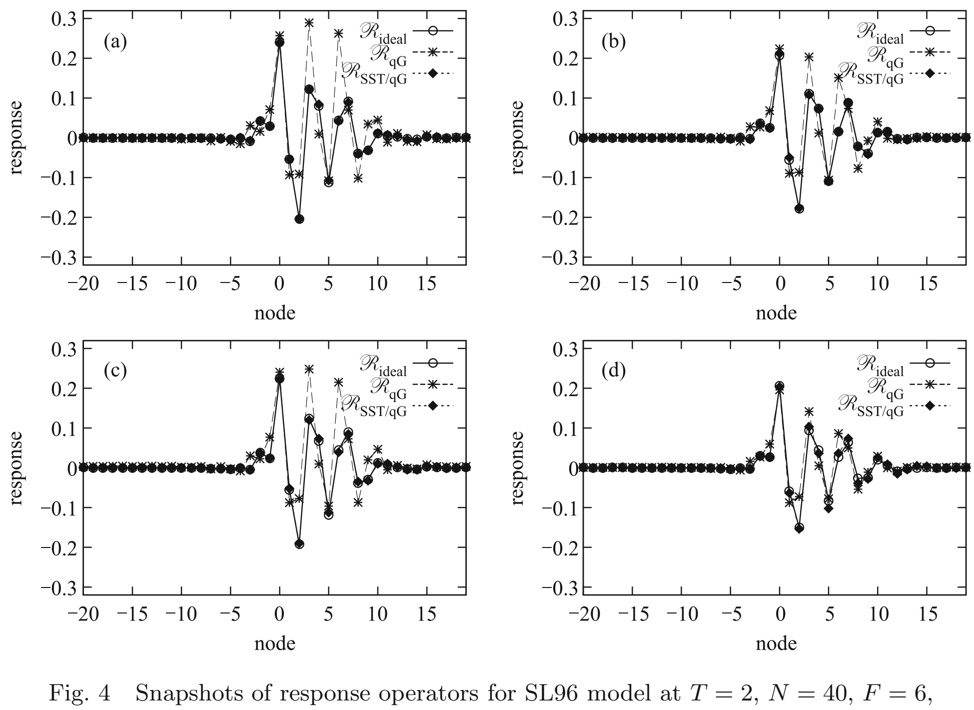

Fig. 4 Snapshots of response operators for SL96 model at $T = 2$, $N = 40$, $F = 6$, (a) $\sigma_k = 0$; (b) $\sigma_k = 1$; (c) $\sigma_k = 0.2X_k$; (d) $\sigma_k = 0.5X_k$.

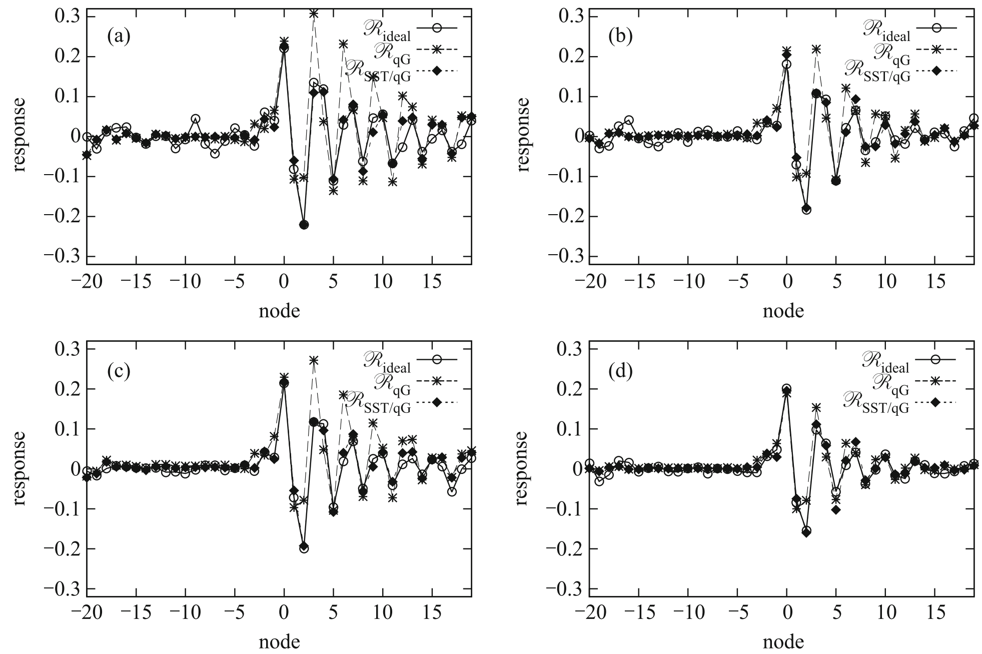

Fig. 5 Snapshots of response operators for SL96 model at $T = 5$, $N = 40$, $F = 6$, (a) $\sigma_k = 0$; (b) $\sigma_k = 1$; (c) $\sigma_k = 0.2X_k$; (d) $\sigma_k = 0.5X_k$.

SST/qG-FDT cutoff time). Although the linear response operator at a given time is a 40×40 matrix, it has the property of translational invariance (just like the L96 model itself), and thus, can be averaged along the main diagonal with wrap-around aliasing of rows (or columns) into a single vector. These averaged vectors are displayed in Figs. 3–5. Observe that for the early times of the response $T = 1, 2$, the SST/qG-FDT response is virtually indistinguishable from the ideal response. As for the qG-FDT response, its best performance is observed in the case of strong multiplicative noise $\sigma_k = 0.5X_k$, where the discrepancies between the qG-FDT and ideal response are not much larger than those between the SST/qG-FDT response and the ideal response. This is probably the consequence of the fact that the strong multiplicative noise changes the invariant probability density of the SL96 model to the point where it is relatively close to the Gaussian. For other regimes, by the response time $T = 2$, significant errors develop in the qG-FDT response to the right of the main response diagonal. For the longer response time $T = 5$ and all regimes, the blended SST/qG-FDT response is very similar to the ideal response, while the qG-FDT response again develops large discrepancies to the right of the main response diagonal for $\sigma_k = 0,\ 1$, and $0.2X_k$. For the strong multiplicative noise regime, $\sigma_k = 0.5X_k$, and response time $T = 5$, the qG-FDT yields lower errors than in the other regimes, but is still less precise than the SST/qG-FDT response.

4 Summary

The classical fluctuation-dissipation theorem, by its design, is suitable for computing the linear response for stochastically driven systems, as it assumes the continuity of the probability measure of the statistical ensemble distribution with respect to the Lebesgue measure (which is guaranteed in many stochastically driven systems). However, the drawback of the classical fluctuation-response formula is that it requires the probability density together with its derivative (or their suitable approximations) explicitly in the response formula. Unfortunately, for complex systems with many variables such an approximation might not be necessarily available with required precision.

In this work, we develop the stochastic short-time fluctuation-dissipation formula (SST-FDT) for stochastically driven systems which does not require the probability measure of the statistical state of the system to be known explicitly. This formula is the analog of the general linear response formula [3–5,9,28] for chaotic (but not stochastically driven) nonlinear systems. We demonstrate that, before the numerical instability due to positive Lyapunov exponents occurs, the SST-FDT for the stochastically driven Lorenz 96 model is generally superior to the classical FDT formula where the probability density of the statistical state is approximated by the Gaussian density with the same mean and covariance (qG-FDT). We test the new SST-FDT formula for the L96 model with stochastic forcing for both the additive and multiplicative noise, and observe that the SST-FDT response formula is generally better than the qG-FDT in both the error and correlation comparison, before the numerical instability develops in the SST-FDT response. Additionally, the blended SST/qG-FDT response with a simple Heaviside blending function clearly performs on top of both the qG-FDT and SST-FDT in all studied regimes. The results of this work suggest that the SST/qG-FDT algorithm can be used in practical applications with stochastic parameterization, such as the climate change prediction.

Acknowledgements The author thanks Ibrahim Fatkullin for helpful comments and remarks. This work was supported by the NSF CAREER grant DMS-0845760 and the ONR grant N000140610286.

References

1. Abramov R. Short-time linear response with reduced-rank tangent map. Chin Ann Math, 2009, 30B(5): 447–462
2. Abramov R. Approximate linear response for slow variables of dynamics with explicit time scale separation. J Comput Phys, 2010, 229(20): 7739–7746
3. Abramov R, Majda A. Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems. Nonlinearity, 2007, 20: 2793–2821
4. Abramov R, Majda A. New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems. J Nonlin Sci, 2008, 18(3):

303–341
5. Abramov R, Majda A. New algorithms for low frequency climate response. J Atmos Sci, 2009, 66: 286–309
6. Bell T. Climate sensitivity from fluctuation dissipation: Some simple model tests. J Atmos Sci, 1980, 37(8): 1700–1708
7. Carnevale G, Falcioni M, Isola S, Purini R, Vulpiani A. Fluctuation-response in systems with chaotic behavior. Phys Fluids A, 1991, 3(9): 2247–2254
8. Cohen B, Craig G. The response time of a convective cloud ensemble to a change in forcing. Quart J Roy Met Soc, 2004, 130(598): 933–944
9. Eckmann J, Ruelle D. Ergodic theory of chaos and strange attractors. Rev Mod Phys, 1985, 57(3): 617–656
10. Evans D, Morriss G. Statistical Mechanics of Nonequilibrium Liquids. New York: Academic Press, 1990
11. Gritsun A. Fluctuation-dissipation theorem on attractors of atmospheric models. Russ J Numer Math Modeling, 2001, 16(2): 115–133
12. Gritsun A, Branstator G. Climate response using a three-dimensional operator based on the fluctuation-dissipation theorem. J Atmos Sci, 2007, 64: 2558–2575
13. Gritsun A, Branstator G, Dymnikov V. Construction of the linear response operator of an atmospheric general circulation model to small external forcing. Num Anal Math Modeling, 2002, 17: 399–416
14. Gritsun A, Branstator G, Majda A. Climate response of linear and quadratic functionals using the fluctuation dissipation theorem. J Atmos Sci, 2008, 65: 2824–2841
15. Gritsun A, Dymnikov V. Barotropic atmosphere response to small external actions, theory and numerical experiments. Atmos Ocean Phys, 1999, 35(5): 511–525
16. Kubo R, Toda M, Hashitsume N. Statistical Physics II: Nonequilibrium Statistical Mechanics. New York: Springer-Verlag, 1985
17. Kunita H. Stochastic Flows and Stochastic Differential Equations. Cambridge: Cambridge University Press, 1997
18. Leith C. Climate response and fluctuation-dissipation. J Atmos Sci, 1975, 32: 2022–2025
19. Lorenz E. Predictability—a problem partly solved. In: Palmer T, Hagedorn R, eds. Predictability of Weather and Climate. Cambridge: Cambridge University Press, 2006
20. Lorenz E, Emanuel K. Optimal sites for supplementary weather observations. J Atmos Sci, 1998, 55: 399–414
21. Majda A, Abramov R, Gershgorin B. High skill in low frequency climate response through fluctuation dissipation theorems despite structural instability. Proc Natl Acad Sci, 2010, 107(2): 581–586
22. Majda A, Abramov R, Grote M. Information Theory and Stochastics for Multiscale Nonlinear Systems. CRM Monograph Series of Centre de Recherches Mathématiques, Université de Montréal, Vol 25. Providence: American Mathematical Society, 2005
23. Majda A, Gershgorin B. A test model for fluctuation-dissipation theorems with time-periodic statistics. Physica D, 2010, 239(17): 1741–1757
24. Majda A, Wang X. Linear response theory for statistical ensembles in complex systems with time-periodic forcing. Commun Math Sci, 2010, 8(1): 145–172
25. Risken F. The Fokker-Planck Equation. 2nd ed. New York: Springer-Verlag, 1989
26. Ruelle D. Chaotic Evolution and Strange Attractors. Cambridge: Cambridge University Press, 1989
27. Ruelle D. Differentiation of SRB states. Comm Math Phys, 1997, 187: 227–241
28. Ruelle D. General linear response formula in statistical mechanics, and the fluctuation-dissipation theorem far from equilibrium. Phys Lett A, 1998, 245: 220–224
29. Young L -S. What are SRB measures, and which dynamical systems have them? J Stat Phys, 2002, 108(5-6): 733–754