Linear Response Theory for Statistical Ensembles in Complex Systems with Time-Periodic Forcing

Andrew J. Majda
Courant Institute, New York University
New York, NY 10012
and
Xiaoming Wang
Department of Mathematics, Florida State University
Tallahassee, FL 32306

Comm. Math. Sci., Vol. 8, issue 1 (March 2010), p.145-172.

Abstract

New linear response formulas for unperturbed chaotic (stochastic) complex dynamical systems with time periodic coefficients are developed here. Such time periodic systems arise naturally in climate change studies due to the seasonal cycle. These response formulas are developed through the mathematical interplay between statistical solutions for the time-periodic dynamical systems and the related skew-product system. This interplay is utilized to develop new systematic quasi-Gaussian and adjoint algorithms for calculating the climate response in such time-periodic systems. These new formulas are found in section 4. New linear response formulas are also developed here for general time-dependent statistical ensembles arising in ensemble prediction including the effects of deterministic model errors, initial ensembles, and model noise perturbations simultaneously . An information theoretic perspective is developed in calculating those model perturbations which yield the largest information deficit for the unperturbed system both for climate response and finite ensemble predictions.

Keywords: Linear response theory, fluctuation-dissipation theory, time periodic coefficients, quasi-Gaussian and Gaussian approximation, climate response, information content, relative entropy
AMS classification: 37N10, 82C31, 86A99, 60H10, 34C28, 94A15

1 Introduction

One of the cornerstones of modern statistical physics is the fluctuation-dissipation theorem (FDT) which roughly states that for systems of identical particles in statistical equilibrium, the average mean response to small external perturbations can be calculated through the

knowledge of suitable correlation functions of the unperturbed statistical systems with many practical applications [20, 26, 5]. The low frequency response to changes in external forcing for various components of the climate system is a central problem of contemporary climate change science. Leith [18] suggested that if the climate system satisfied a suitable FDT, then climate response to small external forcing could be calculated by estimating suitable statistics in the present climate. The climate system is a forced dissipative chaotic dynamical system with time periodic forcing due to the seasonal cycle which is physically quite far from the classical physicists' setting for FDT. Leith's sugestion has created a lot of recent activity in generating new theoretical formulations and approximate algorithms for FDT with applications to climate response [17, 15, 14, 16, 2, 1, 21, 3]. Thus, these approximate FDT algorithms have been applied to autonomous climate models of varying complexity and have ignored the important time periodic effect of the seasonal cycle [6, 12, 4]. One goal of the present paper is to develop a version of FDT for time-periodic systems which leads to new approximate algorithms for climate response with seasonal cycle. The new algorithms are presented in section 4 below.

Finite time ensemble predictions of turbulent chaotic dynamical systems with many degrees of freedoms is an important practical topic (see [27, 25] and Chapter 15 of [23]). The effect of model error, both deterministic and stochastic, in competition with initial ensemble error effecting the skill for ensemble prediction is a central topic [27, 25, 22]. Here linear response theory is extended to unperturbed prediction ensembles far from equilibrium to develop a theoretical framework to discuss the competition between various types of model error and initial ensemble perturbations.

Here a theoretical framework and new approximate algorithms to address both important scientific issues mentioned in the previous two paragraphs are developed in a unified fashion. In section 2, we develop an important mathematical interplay between time-periodic dynamical systems and the related skew-product system and their statistical solutions. In section 3, we develop systematic FDT theorems for the time-periodic equilibrium setting as well as finite time ensemble predictions together with various important approximations. In section 4, we utilize the theory developed earlier in section 2 and 3 to develop a suite of new systematic quasi-Gaussian [18, 21, 15, 17, 14, 16] and adjoint algorithms [2, 1, 3] as well as blended variants [1, 3] for climate response with a seasonal cycle (periodic forcing). The reader mainly interested in the proposed new algorithms for climate response can go to section 4 directly without difficulty. In section 5 an information theoretic perspective [21, 23] is developed in calculating those model perturbations which yield the largest information deficit for the unperturbed system both for climate response and finite ensemble predictions.

2 General Formulation and the Skew Product System

Consider a generic (Ito) SDE system which is assumed to be well-posed and describes the motion of some physical system

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X},t) + \sigma(\mathbf{X})\dot{W},\ \mathbf{X} \in \mathbb{R}^N,\ \mathbf{F} = (F_1, \cdots, F_N), \mathbf{f} = (f_1, \cdots, f_N), \tag{1}$$

where $\dot{W}$ denotes a standard M dimensional white (in time) noise, σ is an $N \times M$ matrix, and $\mathbf{F}(\mathbf{X}, t)$ has time periodic coefficients. For example,

$$\mathbf{F}(\mathbf{X}, t) = \mathbf{F}(\mathbf{X}) + \mathbf{f}(t) \tag{2}$$

where $\mathbf{f}(t)$ is a time-periodic forcing with period T_0. In the case of zero noise, i.e. $\sigma = 0$, the SDE reduces to an ODE system and we shall require that the system to be large ($N \gg 1$) and have the strong mixing property.

For a system with explicit time-dependent forcing which is not stationary in time, it is impossible for the system to reach any time-independent statistical equilibrium. Hence, we have to consider time dependent statistical solution of the system.

For an ensemble prediction, the unperturbed statistical solution $\bar{p}(\mathbf{X}, t)$ is the solution to the associated unperturbed time-dependent **Fokker-Planck equation**

$$\begin{aligned} \frac{\partial \bar{p}}{\partial t} &= -\nabla \cdot (\bar{p}\mathbf{F}(\mathbf{X}, t)) + \frac{1}{2}\nabla \cdot \nabla \cdot (Q\bar{p}) (\stackrel{def}{=} L_{FP}\bar{p}), && (3) \\ \bar{p}(\mathbf{X}, t)\Big|_{t=0} &= \bar{p}_0(\mathbf{X}), && (4) \end{aligned}$$

where

$$Q = \sigma\sigma^T \geq 0 \tag{5}$$

is an $N \times N$ matrix and

$$\nabla \cdot \nabla \cdot (Q\bar{p}) = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2 (Q_{ij}\bar{p})}{\partial x_i \partial x_j} \tag{6}$$

is the associated diffusion operator and the time dependent Fokker-Planck operator L_{FT} is defined as

$$L_{FP}(t)p = -\nabla \cdot (p\mathbf{F}(\mathbf{X}, t)) + \frac{1}{2}\nabla \cdot \nabla \cdot (Q\, p). \tag{7}$$

The Fokker-Planck equation reduces to the **Liouville equation** in the case of zero noise ($Q \equiv 0$). Our goal here is to develop a calculus useful for both climate response and system perturbation/model error in time dependent ensemble prediction.

2.1 Skew Product System

With the introduction of the time-periodic $\mathbf{F}(\mathbf{X}, t)$, it is unlikely that the system will reach any time-independent statistical equilibrium. Moreover, the classical linear response theory [26, 21] cannot be directly applied to this case since a straight forward calculation leads to an equation for linear response with time-dependent coefficients which makes it hard to be useful unless (infinite dimensional version) Floquet theory is invoked. In order to overcome this difficulty, we introduce a skew-product flow/system just as in the classical approach for non-autonomous systems. In the case of periodic $\mathbf{F}(\mathbf{X}, t)$ (with period T_0), the skew-product system is the following SDE on $\mathbb{R}^N \times \mathbb{S}$ where $\mathbb{S} = \mathbb{R}^1/\text{mod}T_0$ is the (one-dimensional) circle with circumference T_0

$$\begin{aligned} \frac{d\mathbf{X}}{dt} &= \mathbf{F}(\mathbf{X}, s) + \sigma(\mathbf{X})\dot{W}, \\ \frac{ds}{dt} &= 1, \end{aligned}$$

which can be written in the form of (1) for the extended (skew-product) variables

$$\hat{\mathbf{X}} = \begin{pmatrix} \mathbf{X} \\ s \end{pmatrix} \tag{8}$$

as

$$\frac{d\hat{\mathbf{X}}}{dt} = \hat{\mathbf{F}}(\hat{\mathbf{X}}) + \hat{\sigma}(\hat{\mathbf{X}})\dot{W}, \ \hat{\mathbf{X}} \in \mathbb{R}^N \times \mathbb{S}, \tag{9}$$

where

$$\hat{\mathbf{F}}(\hat{\mathbf{X}}) = \begin{pmatrix} \mathbf{F}(\mathbf{X}, s) \\ 1 \end{pmatrix}, \quad \hat{\sigma}(\hat{\mathbf{X}}) = \begin{pmatrix} \sigma(\mathbf{X}) \\ 0 \end{pmatrix}. \tag{10}$$

Introducing the gradient operator in the skew-product variable as $\hat{\nabla} = \begin{pmatrix} \nabla_{\mathbf{X}} \\ \frac{\partial}{\partial s} \end{pmatrix} = \begin{pmatrix} \nabla \\ \frac{\partial}{\partial s} \end{pmatrix}$, we may formally write down the Fokker-Planck equation for the skew-product system just as in (3) with the hat. Indeed, with the special form that we have assumed here, it can be written as

$$\frac{\partial \hat{p}}{\partial t} = -\nabla \cdot (\hat{p}\mathbf{F}(\mathbf{X}, s)) - \frac{\partial \hat{p}}{\partial s} + \frac{1}{2}\nabla \cdot \nabla \cdot (Q\hat{p}) (\stackrel{def}{=} \hat{L}_{FP}\hat{p}), \tag{11}$$

$$\hat{p}(\hat{\mathbf{X}}, t)\Big|_{t=0} = \bar{p}_0(\mathbf{X}) \times \delta_0(s) \tag{12}$$

where δ_0 is the Dirac delta function centered at zero and the skew-product Fokker-Planck operator $\hat{L}_{FT}$ is defined as

$$\hat{L}_{FP}p = -\nabla \cdot (p\mathbf{F}(\mathbf{X}, s)) - \frac{\partial p}{\partial s} + \frac{1}{2}\nabla \cdot \nabla \cdot (Q\, p). \tag{13}$$

We naturally inquire the relationship between the statistical solutions for the two formulations. In the case of a time-dependent statistical solution generated with initial data given by (4, 12), the relationship is simple. In fact we can easily verify the following result.

Lemma 1. *Let $\bar{p}(\mathbf{X}, t)$ be a solution of the time-dependent Fokker-Planck equation (3) with initial data (4). Then $\hat{p}(\hat{\mathbf{X}}, t)$ is a solution to the skew-product Fokker-Planck equation (11) with initial data (12) if and only if*

$$\hat{p}(\hat{\mathbf{X}}, t) = \bar{p}(\mathbf{X}, t) \times \delta_0(s - t). \tag{14}$$

Moreover, if $\hat{p}(\hat{\mathbf{x}}, t)$ is a smooth solution of the skew-product Fokker-Planck equation (11), then

$$p(\mathbf{x}, t) \stackrel{def}{=} T_0\hat{p}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t) \tag{15}$$

is a solution to the time-dependent Fokker-Planck equation (3) and is in fact a statistical solution to (1) with initial datum $\hat{p}(\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix}, 0)$.

Proof: We sketch the proof. The verification of (15) is a direct and simple calculation. The relationship (14) has to be understood in the sense of distribution. For this purpose we consider a smooth test functional $\hat{\phi}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t)$ with compact support, we have

$$
\begin{aligned}
& \int_0^\infty \int (\frac{\partial \hat{p}}{\partial t} - \hat{L}_{FT}\hat{p})\hat{\phi}\, d\hat{\mathbf{x}}dt \\
= & \int_0^\infty \int \bar{p}(\mathbf{x},t) \times \delta_0(s-t)(-\frac{\partial}{\partial t} - \hat{L}_{FT}^T)\hat{\phi}\, d\hat{\mathbf{x}}dt \\
= & -\int_0^\infty \int \bar{p}(\mathbf{x},t)(\frac{\partial}{\partial t}\hat{\phi}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t) + \mathbf{F}(\mathbf{x},t)\nabla\hat{\phi}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t) + \frac{\partial}{\partial s}\hat{\phi}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t) + \frac{1}{2}Q : \nabla\nabla\hat{\phi}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t))\, d\hat{\mathbf{x}} \\
= & -\int_0^\infty \int \bar{p}(\mathbf{x},t)(\frac{d}{dt} + L_{FP}^T)\hat{\phi}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t)\, d\hat{\mathbf{x}}dt \\
= & \int_0^\infty \int (\frac{\partial}{\partial t} - L_{FP})\bar{p}(\mathbf{x},t)\hat{\phi}(\begin{pmatrix} \mathbf{x} \\ t \end{pmatrix}, t)\, d\hat{\mathbf{x}}dt \\
= & \ 0.
\end{aligned}
$$

□

Next we ask about existence of statistical equilibrium of the skew-product system and possible long time asymptotic of the statistical behavior of the time dependent system as well as their relationship. Of course statistical equilibrium may not exist even for the skew-product system. However, if the skew-product system (or the original time-dependent system) is dissipative in certain appropriate sense, the existence of statistical equilibrium can be established via the classical Bogliubov-Krylov approach [11, 28]. These dissipative systems include the following type of special systems [21, 23] that are of great importance in geophysical fluid dynamics

$$
\frac{d\mathbf{X}}{dt} = B(\mathbf{X},\mathbf{X}) + L\mathbf{X} - \gamma\mathbf{X} + \mathbf{f}(t) + \sigma\dot{W} \tag{16}
$$

where B is a bilinear anti-symmetric operator with $\mathbf{X} \cdot B(\mathbf{X},\mathbf{X}) = 0$ (resulting from the quadratic advection term for instance), L is a linear anti-symmetric operator (derived from the Coriolis forcing for example), γ represents dissipation and $\mathbf{f}$ denotes external periodic (seasonal, annual) forcing. Both $L(t)$ and $\gamma(t)$ can also have time periodic components in applications centered about time periodic equilibrium states in a reformulation [12, 6, 4]. The global existence of solution to such kind of systems can be found in [9] among others under appropriate assumptions. Nevertheless, we do not know a priori if there is a unique statistical equilibrium since we have degenerate diffusion for the skew-product system (9) even if the noise for the time-dependent system is non-degenerate ($\text{rank}(Q) = N$). However, we will see below that utilizing the relative entropy for the time-dependent problem and a simple relationship between asymptotic behavior of statistical solutions of the time-dependent problem (1) and the skew-product problem (9) yields a somewhat surprising uniqueness of invariant measure (statistical equilibrium) for the skew-product system. This uniqueness of the invariant measure (statistical equilibrium) can be viewed as a manifestation of noise induced statistical stability in this time-dependent setting.

Now if the time-dependent system is dissipative in certain sense, then the skew-product system is also dissipative and hence a generalized long time average defined through Banach limit would generate invariant measure (stationary statistical solution) of the system which we denote $\hat{p}^{eq}(\hat{\mathbf{x}})$. This statistical equilibrium of the skew-product system is related to the long time asymptotic statistical behavior of the original time-dependent dynamical system (1). Indeed, it is easy to verify from (3, 11) (see also (15)) that

$$p_{per}(\mathbf{x},t)=T_0\,\hat{p}^{eq}\begin{pmatrix}\mathbf{x}\\ t\end{pmatrix}, \tag{17}$$

is a solution to the original time-dependent Fokker-Planck equation (3). Since $\hat{p}^{eq}$ is a time independent solution to the extended (skew-product) Fokker-Planck equation (11), we see after integrating in $\mathbf{x}$ that

$$\frac{d}{dt}\int p_{per}(\mathbf{x},t)\,d\mathbf{x}=T_0\frac{d}{dt}\int \hat{p}^{eq}\begin{pmatrix}\mathbf{x}\\ t\end{pmatrix}d\mathbf{x}=T_0\int\frac{\partial}{\partial s}\hat{p}^{eq}\begin{pmatrix}\mathbf{x}\\ t\end{pmatrix}d\mathbf{x}=T_0\int L_{FP}\hat{p}^{eq}\begin{pmatrix}\mathbf{x}\\ t\end{pmatrix}d\mathbf{x}=0$$

and hence, since $\hat{p}^{eq}$ is a probability density function on $\mathbb{R}^N\times\mathbb{S}$,

$$\int p_{per}(\mathbf{x},t)\,d\mathbf{x}=1,\forall t.$$

Therefore, $p_{per}(\mathbf{x},t)$ is a time periodic (period T_0) statistical solution to (1). Conversely, if $p_{per}(\mathbf{x},t)$ is a statistical solution to (1) which is periodic in t with period T_0, we may define $\hat{p}^{eq}$ through (17) and it is easy to verify that $\hat{p}^{eq}$ is a stationary solution of the skew-product Fokker-Planck equation.

Now under generic noise ($\text{rank}(Q)=N$) it is natural to expect that all statistical solutions to (1), are in fact smooth and positive for all $\mathbf{x}\in\mathbb{R}^N$ and $t>0$ due to the non-degenerate diffusion in the time-dependent Fokker-Planck equation (11). Hence any two statistical solutions $p_j, j=1,2$ of (1) must be approaching each other under generic noise ($\text{rank}(Q)=N$) due to the following calculation on the relative entropy (information content) (This is related to the classical H-theorem, see [13, 26], or section 5 below)

$$\frac{d}{dt}H(p_1,p_2)=\frac{d}{dt}\int p_1(\mathbf{x},t)\ln\frac{p_1(\mathbf{x},t)}{p_2(\mathbf{x},t)}\,d\mathbf{x}=-\int\frac{p_1(\mathbf{x},t)}{2R^2(\mathbf{x},t)}\nabla R(\mathbf{x},t)\cdot Q\nabla R(\mathbf{x},t)\,d\mathbf{x}\le 0 \tag{18}$$

where $R=\frac{p_1}{p_2}$ and $H(p_1,p_2)$ denotes the relative entropy (an alternative notation is $\mathcal{P}(p_1,p_2)$, see section 5) which is defined as

$$H(p_1,p_2)=\int p_1(\mathbf{x})\ln\frac{p_1(\mathbf{x})}{p_2(\mathbf{x})}\,d\mathbf{x}. \tag{19}$$

The time derivative of the relative entropy is zero only if ∇R is independent of $\mathbf{x}$ which is possible only if $p_1\equiv p_2$. Therefore we may conclude that all statistical solutions of the time-dependent system (1) approaches the time-periodic asymptotic statistical solution p_{per}. This also implies that there is a unique statistical equilibrium for the skew-product system under the generic noise assumption in lieu of (17).

Of course the relative entropy calculation also applies to the skew-product system. For any two regular/smooth statistical solutions $\hat{p}_j(\hat{\mathbf{x}},t), j=1,2$ of the skew-product system, then the proof of the H-theorem dictates that

$$\frac{d}{dt}\int \hat{p}_1(\hat{\mathbf{x}},t)\ln\frac{\hat{p}_1(\hat{\mathbf{x}},t)}{\hat{p}_2(\hat{\mathbf{x}},t)}\,d\hat{\mathbf{x}} = -\int \frac{\hat{p}_1(\hat{\mathbf{x}},t)}{2\hat{R}^2(\hat{\mathbf{x}},t)}\nabla\hat{R}(\hat{\mathbf{x}},t)\cdot Q\nabla\hat{R}(\hat{\mathbf{x}},t)\,d\hat{\mathbf{x}} \leq 0 \tag{20}$$

where $\hat{R}=\frac{\hat{p}_1}{\hat{p}_2}$. The time derivative is zero only if $\nabla\hat{R}$ is independent of $\mathbf{x}$ which is possible only if $\hat{p}_1\equiv\hat{p}_2$ under the generic noise condition. This shows that the statistical equilibrium of the skew-product system (9) is unique under the generic noise assumption.

We summarize the results on long time statistical properties as following.

Theorem 1. *If the time-dependent system (1) is dissipative in appropriate sense, then it possesses at least one time-periodic (with period T_0) statistical solution p_{per} which is associated with a statistical equilibrium $\hat{p}^{eq}$ of the skew-product system (9) through (17). Moreover, under generic noise assumptions, i.e., rank$(Q)=N$, and appropriate decay property at infinity, the time-periodic statistical solution p_{per} captures all asymptotic statistical properties of the system (1) in the sense that for any statistical solution p*

$$\lim_{t\to\infty} H(p(t),p_{per}(t)) = \lim_{t\to\infty} H(p_{per}(t),p(t)) = 0. \tag{21}$$

In this case, the skew-product system (9) possesses a unique ergodic statistical equilibrium $\hat{p}^{eq}$ which is related to the asymptotic statistical solution p_{per} of (1) through (17). Moreover, in this case of generic noise, $p_{per}(t)$ is the pdf of the unique ergodic invariant measure of the Poincaré map of the system (1) with time period T_0 starting from t. Therefore, we have, for any bounded continuous functional $\varphi(\mathbf{x})$, and t_0,

$$\int \varphi(\mathbf{x})p_{per}(\mathbf{x},t_0)\,d\mathbf{x} = \lim_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K}\varphi(\mathbf{X}(t_0+kT_0)) \tag{22}$$

with the right hand naturally interpreted as the climatological mean.

Proof: The uniqueness of the statistical equilibrium of the skew-product system follows from the uniqueness of asymptotic statistical behavior of the original system (1) and the relationship (17). We point out that this uniqueness does not guarantee the convergence of all statistical solutions of the skew-product system (9) to this unique statistical equilibrium, only the time-average converges.

We only need to remark on the ergodicity (in the sense of equivalence between spatial and temporal averages) of the statistical equilibrium $\hat{p}^{eq}$ when it is unique. This follows from connection between generalized time averaging and invariant measure [11, 28, 29] but it seems that it is less known. Indeed, for any smooth test functional $\hat{\phi}$ with compact support, there exists an invariant measure $\hat{\mu}^{sup}$ induced by a special generalized Banach limit that agrees with the lim sup on this functional in the sense that

$$\limsup_{T\to\infty}\frac{1}{T}\int_0^T \hat{\phi}(\hat{\mathbf{x}}(t))\,dt = \int_H \hat{\phi}(\hat{\mathbf{x}})\,d\hat{\mu}^{sup}(\hat{\mathbf{x}}).$$

Similarly, there exists another invariant measure $\hat{\mu}_{inf}$ which agrees with the lim inf on this functional, i.e.,

$$\liminf_{T\to\infty} \frac{1}{T}\int_0^T \hat{\phi}(\hat{\mathbf{x}}(t))\,dt = \int_H \hat{\phi}(\hat{\mathbf{x}})\,d\hat{\mu}_{inf}(\hat{\mathbf{x}}).$$

Since there is a unique invariant measure $\hat{\mu}^{eq}$ we see that

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T \hat{\phi}(\hat{\mathbf{x}}(t))\,dt = \liminf_{T\to\infty} \frac{1}{T}\int_0^T \hat{\phi}(\hat{\mathbf{x}}(t))\,dt = \int_H \hat{\phi}(\hat{\mathbf{x}})\,d\hat{\mu}^{eq}(\hat{\mathbf{x}}) = \int_H \hat{\phi}(\hat{\mathbf{x}})\hat{p}^{eq}(\hat{\mathbf{x}})\,d\hat{\mathbf{x}}.$$

It is easy to see that the time T_0 Poincaré map of each of the phase shifted dynamical system

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}, t+t_0) + \sigma(\mathbf{X})\dot{W}, \quad \mathbf{X}(0) = \mathbf{x} \tag{23}$$

has a unique invariant measure which is exactly $p_{per}(\mathbf{x}, t_0)$ under the assumption. Since $p_{per}(\mathbf{x}, t_0) > 0, \forall \mathbf{x}$, it is ergodic and therefore the discrete long time average (22) in the statement of the theorem follows. This formula can be used to estimate the asymptotic pdf p_{per} from long time series of the system. □

It is also easy to see that a phase shift, i.e. replacing $\mathbf{f}(t)$ by $\mathbf{f}(t+\phi)$, does not alter the asymptotic statistical behavior. With this skew-product formulation, we may generalize the classical linear response calculation to this case with periodic in time forcing by systematically repeating the time independent formalism [26, 21]. This is done next.

3 The Linear Response Formula

It is well-known that we frequently encounter various uncertainties both in the deterministic forcing term $\mathbf{F}$, and in the noise term σ. Therefore it is natural to consider **system perturbation** induced by perturbation in the deterministic forcing term, and the noise term.

For the deterministic forcing term, we consider a space-time separable perturbation commonly used in climate studies [21] and the perturbation in noise is assumed to be of the same order. Therefore we have the following perturbed system

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}, t) + \mathbf{a}(\mathbf{X}) \bullet \delta\tilde{\mathbf{F}}(t) + (\sigma(\mathbf{X}) + \delta\tilde{\sigma}(\mathbf{X}))\dot{W}, \tag{24}$$

where $\mathbf{a} \bullet \mathbf{w}$ denotes the Hadamard (or Schur, or entrywise) product of $\mathbf{a}$ and $\mathbf{w}$, i.e. the jth component of it is the product of the jth component of $\mathbf{a}$ and $\mathbf{w}$, i.e.

$$(\mathbf{a} \bullet \mathbf{w})_j = a_j w_j. \tag{25}$$

More general perturbation in the noise term of the form of $\delta\tilde{\sigma}\dot{\tilde{W}}$ may be also considered in a similar fashion.

Since $\nabla \cdot (\mathbf{a}(\mathbf{X}) \bullet \delta\tilde{\mathbf{F}}(t)\hat{p}^{\delta}) = \nabla \bullet (\mathbf{a}(\mathbf{X})\hat{p}^{\delta}) \cdot \delta\tilde{\mathbf{F}}(t)$, the **perturbed Fokker-Planck equation** in the skew-product formulation then takes the form

$$\frac{\partial \hat{p}^{\delta}}{\partial t} = -\nabla \cdot (\hat{p}^{\delta}\mathbf{F}(\mathbf{X}, s)) - \frac{\partial \hat{p}^{\delta}}{\partial s} + \frac{1}{2}\nabla \cdot \nabla \cdot (Q\hat{p}^{\delta})$$

$$-\delta\nabla\bullet(\mathbf{a}(\mathbf{X})\hat{p}^\delta)\cdot\tilde{\mathbf{F}}(t)+\frac{\delta^2}{2}\nabla\cdot\nabla\cdot(\tilde{Q}\hat{p}^\delta)+\frac{\delta}{2}\nabla\cdot\nabla\cdot((\sigma\tilde{\sigma}^T+\tilde{\sigma}\sigma^T)\hat{p}^\delta), \tag{26}$$

$$\hat{p}^\delta(\hat{\mathbf{X}},t)\Big|_{t=0} = \hat{p}_0^\delta=\bar{p}_0(\mathbf{X})\times\delta_0(s)+\delta p_0'(\mathbf{X})\times\delta_0(s), \tag{27}$$

where $\tilde{Q}=\tilde{\sigma}^T\tilde{\sigma}$. Note that $\delta p_0'$ can incorporate initial errors in mean, variance, etc, in an ensemble prediction.

Remark: The perturbation in the initial pdf does not need to be of the order of δ of course. On a finite time interval, the leading order perturbation (at least formally) in the pdf will be that of the system perturbation if the perturbation in p_0 is of higher order, or that of the initial pdf if the perturbation in p_0 is of lower order. The perturbation in initial pdf may not play any role for long time behavior in the case the system (perturbed and unperturbed) is mixing and reaches a unique statistical equilibrium.

We now recall the *linear response calculation* [26, 21] applied to the skew-product system. For this purpose we assume

$$\hat{p}^\delta=\bar{\hat{p}}+\delta\hat{p}'+\mathcal{O}(\delta^2). \tag{28}$$

Inserting this into the perturbed Fokker-Plank equation (26) and dropping terms of the order of δ^2 in the perturbed Fokker-Planck equation, we arrive at the following

Approximate Linear Response Dynamics

$$\begin{aligned}\frac{\partial\hat{p}'}{\partial t} &= -\nabla\cdot(\hat{p}'\mathbf{F}(\mathbf{X},s))-\frac{\partial\hat{p}'}{\partial s}+\frac{1}{2}\nabla\cdot\nabla\cdot(Q\hat{p}')\\ &\quad -\nabla\bullet(\mathbf{a}(\mathbf{X})\bar{\hat{p}})\cdot\tilde{\mathbf{F}}(t)+\frac{1}{2}\nabla\cdot\nabla\cdot((\sigma\tilde{\sigma}^T+\tilde{\sigma}\sigma^T)\bar{\hat{p}})\\ &\stackrel{def}{=} \hat{L}_{FP}\hat{p}'+\mathbf{L_a}\bar{\hat{p}}\cdot\tilde{\mathbf{F}}+L_\sigma\bar{\hat{p}}, \end{aligned}\tag{29}$$

$$\hat{p}'(\hat{\mathbf{X}},t)\Big|_{t=0} = p_0'(\mathbf{X})\times\delta_0(s). \tag{30}$$

Here the external operators corresponding to the deterministic uncertainty ($\mathbf{L_a}$) and the uncertainty in noise (L_σ) are defined in an obvious way

$$\mathbf{L_a}p = -\nabla\bullet(\mathbf{a}p), \tag{31}$$

$$L_\sigma p = \frac{1}{2}\nabla\cdot\nabla\cdot((\sigma\tilde{\sigma}^T+\tilde{\sigma}\sigma^T)p). \tag{32}$$

This equation can be (formally) solved exactly to give the **perturbative pdf**

$$\hat{p}'(t)=e^{t\hat{L}_{FP}}\hat{p}_0'+\int_0^t[e^{(t-\tau)\hat{L}_{FP}}\mathbf{L_a}\bar{\hat{p}}(\tau)]\cdot\tilde{\mathbf{F}}(\tau)\,d\tau+\int_0^t e^{(t-\tau)\hat{L}_{FP}}L_\sigma\bar{\hat{p}}(\tau)\,d\tau. \tag{33}$$

Remark: Notice that $L_\sigma=0$ when $\sigma=0$, i.e. zero noise in the unperturbed case. Hence in the case of noise perturbation of an originally noiseless system, the perturbative noise level should be of the order of $\sqrt{\delta}$ in order to have non-trivial order δ perturbation to the pdf due to noise. This is in accordance with conventional wisdom [13].

We are interested in statistical quantities as usual. For a given functional (observable) $A(\hat{\mathbf{X}})$, the statistics under the perturbed dynamics is given by

$$\begin{aligned} E^{\delta}(A)(t) &= \int\int A(\hat{\mathbf{X}})\hat{p}^{\delta}(\hat{\mathbf{X}},t)\,d\hat{\mathbf{X}} \\ &= E^0 + \delta E' + \mathcal{O}(\delta^2) \end{aligned} \tag{34}$$

where

$$\delta E'(t) = \delta \int\int A(\hat{\mathbf{X}})\hat{p}'(\hat{\mathbf{X}},t)\,d\hat{\mathbf{X}} \tag{35}$$

is the leading order perturbation in the statistics which can be written as

$$\begin{aligned} E'(A)(t) &= \int e^{t\hat{L}_{FP}}\hat{p}_0'(\hat{\mathbf{x}})A(\hat{\mathbf{x}})\,d\hat{\mathbf{x}} + \int\int_0^t A(\hat{\mathbf{x}})[e^{(t-\tau)\hat{L}_{FP}}\mathbf{L_a}\bar{\hat{p}}(\hat{\mathbf{x}},\tau)]\cdot\tilde{\mathbf{F}}(\tau)\,d\tau\,d\hat{\mathbf{x}} \\ &\quad + \int\int_0^t A(\hat{\mathbf{x}})[e^{(t-\tau)\hat{L}_{FP}}L_\sigma\bar{\hat{p}}(\hat{\mathbf{x}},\tau)]\,d\tau\,d\hat{\mathbf{x}} \\ &= \int \hat{p}_0'(\hat{\mathbf{x}})e^{t\hat{L}_{FP}^T}A(\hat{\mathbf{x}})\,d\hat{\mathbf{x}} + \int_0^t\int[e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})][\mathbf{L_a}\bar{\hat{p}}(\hat{\mathbf{x}},\tau)]\cdot\tilde{\mathbf{F}}(\tau)\,d\hat{\mathbf{x}}\,d\tau \\ &\quad + \int\int_0^t[e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})][L_\sigma\bar{\hat{p}}(\hat{\mathbf{x}},\tau)]\,d\tau\,d\hat{\mathbf{x}} \qquad (36) \\ &= \int \hat{p}_0'(\hat{\mathbf{x}})e^{t\hat{L}_{FP}^T}A(\hat{\mathbf{x}})\,d\hat{\mathbf{x}} + \int_0^t\int\{\mathbf{L_a}^T[e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})]\}\cdot\tilde{\mathbf{F}}(\tau)\bar{\hat{p}}(\hat{\mathbf{x}},\tau)\,d\hat{\mathbf{x}}\,d\tau \\ &\quad + \int\int_0^t[L_\sigma^T e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})]\bar{\hat{p}}(\hat{\mathbf{x}},\tau)\,d\tau\,d\hat{\mathbf{x}} \qquad (37) \\ &\stackrel{def}{=} \int \hat{p}_0'(\hat{\mathbf{x}})e^{t\hat{L}_{FP}^T}A(\hat{\mathbf{x}})\,d\hat{\mathbf{x}} + \int_0^t \vec{R}_{\mathbf{a},A}(t,\tau)\tilde{\mathbf{F}}(\tau)\,d\tau + \int_0^t R_{\sigma,A}(t,\tau)\,d\tau \qquad (38) \end{aligned}$$

where $\hat{L}_{FP}^T, \mathbf{L_a}^T, L_\sigma^T$ are the adjoint operators of $\hat{L}_{FP}, \mathbf{L_a}, L_\sigma$ respectively. Here the **linear response operators** $\vec{R}_{\mathbf{a}}$ and R_σ which account for the system perturbations only are defined as

$$\vec{R}_{\mathbf{a},A}^T(t,\tau) = \int[e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})][\mathbf{L_a}\bar{\hat{p}}(\hat{\mathbf{x}},\tau)]\,d\hat{\mathbf{x}} = \int \mathbf{L_a}^T[e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})]\bar{\hat{p}}(\hat{\mathbf{x}},\tau)\,d\hat{\mathbf{x}}, \tag{39}$$

$$R_{\sigma,A}(t,\tau) = \int[e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})][L_\sigma\bar{\hat{p}}(\hat{\mathbf{x}},\tau)]\,d\hat{\mathbf{x}} = \int[L_\sigma^T e^{(t-\tau)\hat{L}_{FP}^T}A(\hat{\mathbf{x}})]\bar{\hat{p}}(\hat{\mathbf{x}},\tau)\,d\hat{\mathbf{x}}. \tag{40}$$

The first form may be suitable for quasi-Gaussian (and Gaussian) approximation [18, 21, 14, 16, 15, 17, 3] while the second one is more useful for direct adjoint approximation [2, 1, 3] as we shall see below.

Remark: This general linear response approximation is probably valid for finite time only due to the initial ensemble perturbation. For instance, if we have time-independent forcing perturbation, i.e. $\tilde{\mathbf{F}}(t) \equiv \tilde{\mathbf{F}}$, no time dependent external forcing, i.e. $\mathbf{F}(\mathbf{X},t) \equiv \mathbf{F}(\mathbf{X})$, and the system is mixing and hence all statistical solutions converge to the unique equilibrium state, the perturbation in the statistics should be independent of the perturbation in the initial pdf over a long time, i.e. $E(A)(t)$ should be essentially independent $\hat{p}_0'$ for large t.

But this is not consistent with the approximation formula in (38) as the last two terms (the two linear response operators) are independent of $\hat{p}_0'$ while the first depends on the initial perturbation in a linear fashion (unless there is magic cancelation). Therefore it might make sense to ignore perturbation in initial pdf, i.e. set $\hat{p}_0' = 0$, if we are interested in long time approximation.

In the case when the unperturbed pdf $\bar{\hat{p}}$ is smooth and non-vanishing, we may formally rewrite the linear response operator in the form of a statistical average which can be replaced by long time average if the unperturbed pdf is assumed to be ergodic [26, 21]. This is in the spirit of the fluctuation-dissipation theory in statistical physics [26, 21, 5].

Theorem 2 (FDT)**.** *Suppose that $\bar{\hat{p}}(\hat{\mathbf{X}},\tau) > 0, \forall \hat{\mathbf{X}}, \forall \tau > 0$ and it is smooth. Then the computation of the linear response operators (39, 40) can be reduced to the computation of the following statistical correlations*

$$\vec{R}^T_{\mathbf{a},A}(t,\tau) = < A(\hat{\mathbf{X}}(t))\hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{X}}(\tau)) >= \int [e^{(t-\tau)\hat{L}^T_{FP}} A(\hat{\mathbf{X}})]\hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{X}},\tau)\bar{\hat{p}}(\hat{\mathbf{X}},\tau)\, d\hat{\mathbf{X}}, \quad (41)$$

$$R_{\sigma,A}(t,\tau) = < A(\hat{\mathbf{X}}(t))\hat{B}_{\sigma}(\hat{\mathbf{X}}(\tau)) >, \quad (42)$$

with the special (vector) nonlinear functionals

$$\hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{X}},\tau) = \frac{\mathbf{L}_{\mathbf{a}}\bar{\hat{p}}(\hat{\mathbf{X}},\tau)}{\bar{\hat{p}}(\hat{\mathbf{X}},\tau)}, \quad \hat{B}_{\sigma}(\hat{\mathbf{X}},\tau) = \frac{L_{\sigma}\bar{\hat{p}}(\hat{\mathbf{X}},\tau)}{\bar{\hat{p}}(\hat{\mathbf{X}},\tau)}, \quad (43)$$

where the correlations are evaluated at the unperturbed pdf $\bar{\hat{p}}(\hat{\mathbf{X}},\tau)$. In the case when the unperturbed statistical solution to the skew-product system is related to the statistical solution $\bar{p}(\mathbf{X},\tau)$ of the original time-dependent system through (14) and under the assumption that $\bar{p}(\mathbf{X},\tau)$ is smooth and positive, then the computation of the linear response operators (39, 40) can be reduced to the computation of the following statistical correlations

$$\vec{R}^T_{\mathbf{a},A}(t,\tau) = < A(\hat{\mathbf{X}}(t))\mathbf{B}_{\mathbf{a}}(\hat{\mathbf{X}}(\tau)) >= \int [e^{(t-\tau)\hat{L}_{FP}} A(\hat{\mathbf{X}})]|_{s=\tau}\mathbf{B}_{\mathbf{a}}(\hat{\mathbf{X}},\tau)|_{s=\tau}\bar{p}(\mathbf{X},\tau)\, d\mathbf{X}, \quad (44)$$

$$R_{\sigma,A}(t,\tau) = < A(\hat{\mathbf{X}}(t))B_{\sigma}(\hat{\mathbf{X}}(\tau)) >, \quad (45)$$

with the special nonlinear functional (which are independent of s) computed through $\bar{p}(\mathbf{X},\tau)$ directly

$$\mathbf{B}_{\mathbf{a}}(\hat{\mathbf{X}},\tau) = \frac{L_{\mathbf{a}}\bar{p}(\mathbf{X},\tau)}{\bar{p}(\mathbf{X},\tau)}, \quad B_{\sigma}(\hat{\mathbf{X}},\tau) = \frac{L_{\sigma}\bar{p}(\mathbf{X},\tau)}{\bar{p}(\mathbf{X},\tau)}. \quad (46)$$

Proof: It is obvious that the correlation function of two functionals A and $\mathbf{B}$ with respect to $\bar{\hat{p}}$ can be written as

$$\begin{aligned} < A(\hat{\mathbf{X}}(t))\mathbf{B}(\hat{\mathbf{X}}(\tau)) > &= \int A(\hat{\mathbf{X}}) \int \mathbf{B}(\hat{\mathbf{Y}}) p(\hat{\mathbf{X}},t;\hat{\mathbf{Y}},\tau)\, d\hat{\mathbf{Y}} d\hat{\mathbf{X}} \\ &= \int A(\hat{\mathbf{X}}) \int \mathbf{B}(\hat{\mathbf{Y}}) p(\hat{\mathbf{X}},t\Big|\hat{\mathbf{Y}},\tau)\bar{\hat{p}}(\hat{\mathbf{Y}},\tau)\, d\hat{\mathbf{Y}} d\hat{\mathbf{X}} \\ &= \int A(\hat{\mathbf{X}}) \int \mathbf{B}(\hat{\mathbf{Y}}) e^{(t-\tau)\hat{L}_{FP}(\hat{\mathbf{X}})}\delta(\hat{\mathbf{X}}-\hat{\mathbf{Y}})\bar{\hat{p}}(\hat{\mathbf{Y}},\tau)\, d\hat{\mathbf{Y}} d\hat{\mathbf{X}} \end{aligned}$$

$$
\begin{aligned}
&= \int A(\hat{\mathbf{X}}) e^{(t-\tau)\hat{L}_{FP}(\hat{\mathbf{X}})} [\mathbf{B}(\hat{\mathbf{X}})\bar{\hat{p}}(\hat{\mathbf{X}},\tau)]\, d\hat{\mathbf{X}} \\
&= \int [e^{(t-\tau)\hat{L}_{FP}^T} A(\hat{\mathbf{X}})][\mathbf{B}(\hat{\mathbf{X}})\bar{\hat{p}}(\hat{\mathbf{X}},\tau)]\, d\hat{\mathbf{X}}. \qquad (47)
\end{aligned}
$$

This ends the proof of the theorem. □

Remark: The first formulation for the $\hat{\mathbf{B}}_{\mathbf{a}}, \hat{B}_\sigma$ may be suitable when the unperturbed pdf is the (presumed unique) stationary statistical solution of the skew-product system which is (assumed) to be smooth and positive while the second formulation is more suitable for ensemble prediction based on the original system and hence the unperturbed pdf given by (14) is neither smooth nor positive everywhere in the skew-product variable. On the other hand, it is quite natural to expect $\bar{p}(\mathbf{X},\tau)$ to be smooth and positive everywhere under the assumption of non-degenerate noise ($\text{rank}(Q) = N$) and some appropriate dissipative assumption.

Although the contribution from uncertainty in initial perturbation looks straightforward and there has been an abundant literature on this topic, the effect is not completely clarified, especially in the presence of system perturbation. Moreover, perturbation in initial data must be included in order to study ensemble perturbation which is of great importance in practice. Therefore, we have included both perturbation in initial data and system perturbation to account for model error [27, 24] in our study here for potential future applications.

In the special case when the unperturbed pdf is an equilibrium pdf of the unperturbed system, i.e., $\hat{L}_{FP}\bar{\hat{p}} \equiv 0$, and hence $B(\hat{\mathbf{X}},\tau)$ is independent of τ, the correlation can be written in a simpler form

$$
\begin{aligned}
R_{A,B}(t,\tau) &= < A(\hat{\mathbf{X}}(t))B(\hat{\mathbf{X}}(\tau)) > \\
&= R_{A,B}(t-\tau, 0) \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{T_*}^{T+T_*} A(\hat{\mathbf{X}}(t'+t-\tau))B(\hat{\mathbf{X}}(t'))\, dt' \qquad (48) \\
&= \lim_{T\to\infty} \frac{1}{T} \int_{T_*}^{T+T_*} A(\mathbf{X}(t'+t-\tau))B(\mathbf{X}(t'))\, dt' \qquad (49)
\end{aligned}
$$

where we have invoked the ergodicity assumption in the second to the last step, and we have assumed that A, B are independent of s in the last step. This last two formula may be particularly useful in the case of an ergodic statistical equilibrium since they do not involve the potentially unstable computation of tangent map although long time integration of the skew-product system is required. The issue of long time integrator which captures the equilibrium statistics is itself an interesting and challenging issue. See [10] for the case of finite dimensional Hamiltonian system using Andersen thermostat approach, and [7, 8, 30] for infinite dimensional dissipative systems.

3.1 Zero Noise Tangent Map Approach

Notice, as in [1, 2, 3], that in the case of *zero noise*, the adjoint skew Fokker-Planck equation can be solved explicitly via the characteristic method,

$$
e^{(t-\tau)\hat{L}_{FP}^T} A(\hat{\mathbf{x}}) = A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t-\tau)) \qquad (50)
$$

where $\hat{\mathbf{X}}(\hat{\mathbf{x}}, t-\tau)$ is the solution at time $t-\tau$ of the zero noise skew equation (9) which starts at $\hat{\mathbf{x}}$ at time zero.

The contribution to the leading order statistics from the perturbation of the initial pdf can be handled easily

$$\int \hat{p}_0'(\hat{\mathbf{x}}) e^{t\hat{L}_{FP}^T} A(\hat{\mathbf{x}})\, d\hat{\mathbf{x}} = \int \hat{p}_0'(\hat{\mathbf{x}}) A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t))\, d\hat{\mathbf{x}} = \int p_0'(\mathbf{x}) A(\hat{\mathbf{X}}(\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix}, t))\, d\mathbf{x}. \tag{51}$$

As for the contribution to the perturbation of the leading order statistics due to perturbation in external forcing, we can rewrite the **linear response operator** in this zero noise case as

$$\begin{aligned}
\vec{R}_{\mathbf{a},A}^T(t,\tau) = \vec{R}^T(t,\tau) &= \int \mathbf{L}_{\mathbf{a}}^T [e^{(t-\tau)\hat{L}_{FP}^T} A(\hat{\mathbf{x}})] \bar{\hat{p}}(\hat{\mathbf{x}}, \tau)\, d\hat{\mathbf{x}} \\
&= \int A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t-\tau)) [\mathbf{L}_{\mathbf{a}} \bar{\hat{p}}(\hat{\mathbf{x}}, \tau)]\, d\hat{\mathbf{x}} \\
&= \int A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t-\tau)) (-\frac{\partial (a_j(\mathbf{x}) \bar{\hat{p}}(\hat{\mathbf{x}}, \tau))}{\partial x_j})\, d\hat{\mathbf{x}} \\
&= \int (\frac{\partial A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t-\tau))}{\partial x_j} a_j(\mathbf{x}) \bar{\hat{p}}(\hat{\mathbf{x}}, \tau))\, d\hat{\mathbf{x}} \\
&= \int \nabla_{\mathbf{x}} A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t-\tau)) \bullet \mathbf{a}(\mathbf{x})\, \bar{\hat{p}}(\hat{\mathbf{x}}, \tau)\, d\hat{\mathbf{x}} \\
&= \int_{\mathbb{R}^N} \int_{\mathbb{T}} \nabla_{\mathbf{x}} A(\hat{\mathbf{X}}(\begin{pmatrix} \mathbf{x} \\ s \end{pmatrix}, t-\tau)) \bullet \mathbf{a}(\mathbf{x})\, \bar{\hat{p}}(\begin{pmatrix} \mathbf{x} \\ s \end{pmatrix}, \tau)\, ds\, d\mathbf{x}.
\end{aligned} \tag{52}$$

This is a simple generalization of the classical linear response formula [26, 21] to the current environment of perturbation away from time-dependent statistical state. In the case of perturbation near an equilibrium, the average with respect to the equilibrium pdf $\bar{\hat{p}}$ can be replaced by long time average after invoking the ergodicity assumption. As developed in [1, 2, 3], the current form allows us to compute the short time linear response without explicit knowledge of the unperturbed state (only the statistics are needed which could be obtained via Monte-Carlo simulation or observation).

The derivative $\nabla_{\mathbf{x}} A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t))$ may be calculated via solving the linearized equation (tangent map) and is related to the finite time Lyapunov exponents of the skew-product dynamical system (9) since

$$\frac{\partial A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t))}{\partial x_j} = \nabla_{\mathbf{X}} A(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t)) \cdot \frac{\partial \mathbf{X}(\hat{\mathbf{x}}, t)}{\partial x_j}, \tag{53}$$

and

$$\frac{d \nabla_{\mathbf{x}} \mathbf{X}(\hat{\mathbf{x}}, t)}{dt} = \nabla_{\mathbf{X}} \mathbf{F}(\mathbf{X}) \Big|_{\mathbf{X} = \mathbf{X}(\hat{\mathbf{x}}, t)} \nabla_{\mathbf{x}} \mathbf{X}(\hat{\mathbf{x}}, t), \tag{54}$$

$$\nabla_{\mathbf{x}} \mathbf{X}(\hat{\mathbf{x}}, t) \Big|_{t=0} = I. \tag{55}$$

Therefore

$$\nabla_{\mathbf{x}} \mathbf{X}(\hat{\mathbf{x}}, t'+t) = \exp(\int_{t'}^{t'+t} \nabla_{\mathbf{X}} \mathbf{F}(\mathbf{X}) \Big|_{\mathbf{X} = \mathbf{X}(\hat{\mathbf{x}}, \tau)} d\tau) \nabla_{\mathbf{x}} \mathbf{X}(\hat{\mathbf{x}}, t') (\stackrel{def}{=} T_{(\hat{\mathbf{x}}, t')}^t \nabla_{\mathbf{x}} \mathbf{X}(\hat{\mathbf{x}}, t')). \tag{56}$$

At the first glance, it seems unlikely for the linear response theory to be valid for long time for system with at least one positive Lyapunov exponent as the linear response operator will grow exponentially in time. On the other hand, we are considering statistical averages here which makes those worst scenario argument inapplicable unless we consider a degenerate ensemble of a single trajectory. Ample evidence of the practical skill of the adjoint tangent map approach in the case of time-independent deterministic forcing and perturbation away from statistical equilibiurm may be found in [1, 2, 3].

There are two special cases that merit elaboration. The first is the case when the unperturbed statistical solution $\bar{\hat{p}}$ of the skew-product system (9) is related to the statistical solution $\bar{p}$ of the time dependent system (1) through (14). We further assume that the observable A depends on the spatial location $\mathbf{X}$ only, i.e. $A(\hat{\mathbf{X}}) = A(\mathbf{X})$ (see section 4 below for general observable), we have

$$\begin{aligned} \vec{R}^T_{\mathbf{a},A}(t,\tau) &= \int A(\hat{\mathbf{X}}(\begin{pmatrix} \mathbf{x} \\ \tau \end{pmatrix}, t-\tau))[\mathbf{L_a}\bar{p}(\mathbf{x},\tau)]\, d\mathbf{x} \\ &= \int \nabla_{\mathbf{x}} A(\mathbf{X}_\tau(\mathbf{x}, t-\tau)) \bullet \mathbf{a}(\mathbf{x})\bar{p}(\mathbf{x},\tau)]\, d\mathbf{x} \end{aligned} \tag{57}$$

where $\mathbf{X}_\tau(\mathbf{x}, t-\tau)$ denotes the $\mathbf{X}$ component of $\hat{\mathbf{X}}(\begin{pmatrix} \mathbf{x} \\ \tau \end{pmatrix}, t-\tau)$.

Now in the special subcase of the unperturbed pdf $\bar{p}$ is generated via finite ensemble prediction

$$\bar{p}(\mathbf{x},t) = \sum_{j=1}^{R} p_j \delta_0(\mathbf{x} - \mathbf{x}_j(t)), \quad \sum_{j=1}^{R} p_j = 1, \tag{58}$$

the linear response operator can be calculated with the help of the characteristic method and the tangent map as

$$\vec{R}^T_{\mathbf{a},A}(t,\tau) = \sum_{j=1}^{R} p_j \nabla_{\mathbf{x}} A(\mathbf{X}_\tau(\mathbf{x}, t-\tau)) \bullet \mathbf{a}(\mathbf{x})\Big|_{\mathbf{x}=\mathbf{x}_j(\tau)}. \tag{59}$$

The term involving p_0' can be handled similarly.

The other case is when the unperturbed pdf $\bar{\hat{p}}$ is the statistical equilibrium $\hat{p}^{eq}(\hat{\mathbf{x}}) = p_{per}(\mathbf{x}, s)$ of the skew-product system. In this case the correlation is a function of $t-\tau$ only and

$$\vec{R}^T_{\mathbf{a},A}(t) = \int_{\mathbb{R}^N} \left(\int_{\mathbb{S}} \nabla_{\mathbf{x}} A(\mathbf{X}(\begin{pmatrix} \mathbf{x} \\ s \end{pmatrix}, t)) \bullet \mathbf{a}(\mathbf{x})\, \hat{p}^{eq} \begin{pmatrix} \mathbf{x} \\ s \end{pmatrix} ds \right) d\mathbf{x} \tag{60}$$

$$= \int_{\mathbb{R}^N} \left(\int_{\mathbb{S}} \nabla_{\mathbf{x}} A(\mathbf{X}(\begin{pmatrix} \mathbf{x} \\ s \end{pmatrix}, t)) \bullet \mathbf{a}(\mathbf{x})\, p_{as}(\mathbf{x}, s)\, ds \right) d\mathbf{x} \tag{61}$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{T^*}^{T+T^*} \nabla_{\mathbf{x}} A(\mathbf{X}(\hat{\mathbf{x}}, t+\tau)) \bullet \mathbf{a}(\mathbf{X}(\hat{\mathbf{x}}, \tau))\, d\tau \tag{62}$$

where in the last step we invoked the ergodicity assumption and the inner integral can be viewed as an average over the phases. Note that for applications of linear response theory to

climate change with seasonal forcing, it is very interesting to have more general functionals, $A(\mathbf{X}, s)$, in applications. See section 4 below.

It seems that whether $\tilde{\mathbf{F}}$ is periodic in t with period T_0 or not does not affect the calculation above. The additional periodic assumption does not provide further simplification unless the unperturbed system is assumed to be at statistical equilibrium.

3.2 Quasi-Gaussian Approximation, Gaussian Approximation and Ensemble Prediction

As an alternative approximation to the direct approach discussed in the previous subsection, we may utilize the assumption that the unperturbed pdf $\bar{p}$ or the statistical equilibrium $\hat{p}^{eq}$ is close to a Gaussian in many applications and hence we may replace it by a Gaussian with the same mean and variance in some appropriate fashion. Depending on the manner on how the equivalent Gaussians are utilized, we may end-up with the so-called quasi-Gaussian approximation [21] or the simple direct Gaussian approximation. A crucial advantage of quasi-Gaussian approximation is that *we do not need linear tangent model to assess response behavior approximately.*

Due to the presence of time-dependent deterministic forcing, there are two ways to introduce the quasi-Gaussian or Gaussian approximation. The first is closer to finite ensemble approach and utilizes the relationship (14) between the statistical solutions of the time-dependent system and the skew product system. The second approach utilizes appropriate Gaussian approximation of the statistical equilibrium $\hat{p}^{eq}$ of the skew-product system.

3.2.1 The Quasi-Gaussian Approximation

We first consider the case of statistical ensembles generated by the time-dependent system since this is the one most useful in practice. More specifically, for each fixed time t, we define a Gaussian pdf $p^G(\mathbf{X}, t)$ with the same mean and second moments as the unperturbed pdf $\bar{p}(\mathbf{X}, t)$, i.e.

$$\int \mathbf{X} p^G(\mathbf{X}, t)\, d\mathbf{X} = \int \mathbf{X} \bar{p}(\mathbf{X}, t)\, d\mathbf{X} \stackrel{def}{=} \bar{\mathbf{X}}(t), \tag{63}$$

$$\int (\mathbf{X} - \bar{\mathbf{X}}(t)) \otimes (\mathbf{X} - \bar{\mathbf{X}}(t)) p^G(\mathbf{X}, t)\, d\mathbf{X} = \int (\mathbf{X} - \bar{\mathbf{X}}(t)) \otimes (\mathbf{X} - \bar{\mathbf{X}}(t)) \bar{p}(\mathbf{X}, t)\, d\mathbf{X} \stackrel{def}{=} \mathcal{C}(t, t) \tag{64}$$

where $\mathcal{C}(t, t)$ is the covariance matrix of $\mathbf{X}(t)$, and hence

$$p^G(\mathbf{X}, t) = \frac{1}{(2\pi)^{\frac{N}{2}} \det \mathcal{C}^{\frac{1}{2}}} \exp(-\frac{(\mathbf{X} - \bar{\mathbf{X}}(t)) \cdot \mathcal{C}^{-1} (\mathbf{X} - \bar{\mathbf{X}}(t))}{2}). \tag{65}$$

Now we replace $\bar{p}(\mathbf{X}, t)$ by $p^G(\mathbf{X}, t)$ in the calculation of the $\mathbf{B_a}$ and B_σ in the correlation formulation (46) in the previous theorem, i.e., we propose

$$(\vec{R}^G_{\mathbf{a},A})^T(t, \tau) =< A(\hat{\mathbf{X}}(t)) \mathbf{B}^G_{\mathbf{a}}(\mathbf{X}(\tau)) >, \quad R^G_{\sigma,A}(t, \tau) =< A(\hat{\mathbf{X}}(t)) B^G_\sigma(\mathbf{X}(\tau)) >, \tag{66}$$

with the approximate special (vector) nonlinear functionals

$$\mathbf{B}^G_{\mathbf{a}}(\mathbf{X}, \tau) = \frac{\mathbf{L_a} p^G(\mathbf{X}, \tau)}{p^G(\mathbf{X}, \tau)}, \quad B^G_\sigma(\mathbf{X}, \tau) = \frac{L_\sigma p^G(\mathbf{X}, \tau)}{p^G(\mathbf{X}, \tau)}. \tag{67}$$

Therefore if the unperturbed pdf is generated via finite ensemble prediction given in (58), we have, in the zero noise case

$$(\vec{R}^G_{\mathbf{a},A})^T(t,\tau) = \sum_{j=1}^{R} p_j\, A(\mathbf{x}_j(t))\mathbf{B}^G_{\mathbf{a}}(\mathbf{x}_j(\tau)), \tag{68}$$

Notice that no tangent map is needed here and this is a huge advantage for this approximation for finite ensemble prediction. This should be contrasted to the tangent map adjoint approach presented in the previous subsection.

In the special case of perturbation in the external forcing only, i.e. $\mathbf{a}(\mathbf{X}) \equiv \mathbf{a}$ is a constant vector, we may assume, without loss of generality,

$$\mathbf{L}_{\mathbf{a}} = -\nabla_{\mathbf{X}}. \tag{69}$$

Hence we have

$$\mathbf{B}^G_{\mathbf{a}}(\mathbf{X},\tau) \;=\; \mathcal{C}^{-1}(t,t)(\mathbf{X} - \bar{\mathbf{X}}(\tau)). \tag{70}$$

Thus if A is a linear function in $\mathbf{X}$, the linear response operator $\vec{R}^G_{\mathbf{a},A}(t,\tau)$ is essentially the auto-correlation of $\mathbf{X}$.

In the case when the unperturbed statistical solution $\bar{\hat{p}}$ of the skew-product system is the statistical equilibrium $\hat{p}^{eq}$ of the system, the quasi-Gaussian approximation must be developed in a slightly different way. First we notice that $\hat{p}^{eq}$ cannot be Gaussian since the variable s lives on a circle. What we can expect is that each slice is close to a Gaussian for fixed s. Of course this concept needs to be tested on some simple models. Thanks to (17), we see that for each fixed s, $T_0\hat{p}^{eq}(\mathbf{x},s)$ is a pdf and we may approximate it by a Gaussian denoted $p^{G,s}(\mathbf{x})$. We then propose the following quasi-Gaussian approximation

$$\mathbf{B}^{G,eq}_{\mathbf{a}}(\mathbf{X},s) \;=\; \frac{\mathbf{L}_{\mathbf{a}}p^{G,s}(\mathbf{X})}{p^{G,s}(\mathbf{X})}, \tag{71}$$

$$B^{G,eq}_{\sigma}(\mathbf{X},s) \;=\; \frac{L_{\sigma}p^{G,s}(\mathbf{X})}{p^{G,s}(\mathbf{X})}, \tag{72}$$

$$(\vec{R}^{G,eq}_{\mathbf{a},A})^T(t) \;=\; < A(\hat{\mathbf{X}}(t))\mathbf{B}^{G,eq}_{\mathbf{a}}(\mathbf{X},s) >, \tag{73}$$

$$R^{G,eq}_{\sigma,A}(t) \;=\; < A(\hat{\mathbf{X}}(t))B^{G,eq}_{\sigma}(\mathbf{X},s) > . \tag{74}$$

There is a similar version of the Gaussian approximation at equilibrium with this approximation.

Quasi-Gaussian approximations are quite successful in many geophysical applications, especially for estimating linear climate response in the mean ($A(\mathbf{X}) = \mathbf{X}$) and variance ($A(\mathbf{X}) = \mathbf{X} \otimes \mathbf{X}$) [14, 16, 21]. The success can be partially explained through the following short time asymptotic expansion.

Theorem 3 (Short time validity of quasi-Gaussian approximation)**.** *Assume that the unperturbed pdf $\bar{p}(\mathbf{X},t)$ of the system (1) is smooth and non-vanishing. Furthermore, we assume that the deterministic perturbation is defined by external forcing only, i.e. $\mathbf{a}(\mathbf{X}) \equiv \mathbf{a}$. Then*

for the special linear functionals, $A(\mathbf{X}) = X_j$, the quasi-Gaussian approximation defined in (66) satisfies

$$\vec{R}^G_{\mathbf{a},A}(t,\tau) \;=\; \vec{R}_{\mathbf{a},A}(t,\tau) + \mathcal{O}(t-\tau). \tag{75}$$

Proof: Under the assumption, we may assume $\mathbf{L_a} = -\nabla_{\mathbf{X}}$ so that

$$\vec{R}^T_{\mathbf{a},A}(t,\tau) \;=\; -\int [e^{(t-\tau)\hat{L}^T_{FP}} A(\hat{\mathbf{x}})]\Big|_{s=\tau} [\nabla_{\mathbf{x}} \bar{p}(\mathbf{x},\tau)]\, d\mathbf{x} = \int \nabla_{\mathbf{x}} [e^{(t-\tau)\hat{L}^T_{FP}} A(\hat{\mathbf{x}})]\Big|_{s=\tau} \bar{p}(\mathbf{x},\tau)\, d\mathbf{x}, \tag{76}$$

$$(\vec{R}^G_{\mathbf{a},A})^T(t,\tau) \;=\; -\int [e^{(t-\tau)\hat{L}^T_{FP}} A(\hat{\mathbf{x}})]\Big|_{s=\tau} [\frac{\nabla_{\mathbf{x}} p^G(\mathbf{x},\tau)}{p^G(\mathbf{x},\tau)}] \bar{p}(\mathbf{x},\tau)\, d\mathbf{x} = \int [e^{(t-\tau)L^T_{FP}} A(\mathbf{x})]\Big|_{s=\tau} \mathbf{B}^G_{\mathbf{a}}(\mathbf{x},\tau)\bar{p}(\mathbf{x},\tau)\, d\mathbf{x} \tag{77}$$

It is easy to see that for linear test functional A we have

$$\begin{aligned}
\vec{R}^T_{\mathbf{a},A}(t,t) &= \int \nabla A(\mathbf{x}) \bar{p}(\mathbf{x},t)\, d\mathbf{x} \\
&= \int \nabla A(\mathbf{x}) p^G(\mathbf{x},t)\, d\mathbf{x} \\
&= -\int A(\mathbf{x}) \nabla p^G(\mathbf{x},t)\, d\mathbf{x} \\
&= \int A(\mathbf{x}) \mathbf{B}^G_{\mathbf{a}}(\mathbf{x},t) p^G(\mathbf{x},t)\, d\mathbf{x} \\
&= \int A(\mathbf{x}) \mathbf{B}^G_{\mathbf{a}}(\mathbf{x},t) \bar{p}(\mathbf{x},t)\, d\mathbf{x} \\
&= (\vec{R}^G_{\mathbf{a},A})^T(t,t)
\end{aligned}$$

where in the second to the last step we have utilized the fact that $\mathbf{B}^G_{\mathbf{a}}(\mathbf{x},t)$ is a polynomial of first degree (and hence $A(\mathbf{x})\mathbf{B}^G_{\mathbf{a}}(\mathbf{x},t)$ is a polynomial of second degree) and the first two moments of $p^G(\mathbf{x},t)$ and $\bar{p}(\mathbf{x},t)$ are the same. This leads to the conclusion for general unperturbed pdf $\bar{p}$. A similar argument applies for the quasi-Gaussian approximation in statistical equilibrium for the skew system. □

3.2.2 Gaussian approximation

For higher order accuracy estimates, we may formally differentiate $\vec{R}_{\mathbf{a},A}(t,\tau)$ and $\vec{R}^G_{\mathbf{a},A}(t,\tau)$ in τ and evaluate at $\tau = t$ and hope that the derivatives match. Unfortunately this is not the case. Nevertheless, if we consider Gaussian approximation instead of quasi-Gaussian approximation, i.e. replacing $\bar{p}$ by p^G directly in the formula for the linear response operators in (39, 40), this is possible at least under appropriate noise. Hence we introduce the following Gaussian approximation linear response operators

$$\begin{aligned}
& R^g_{\sigma,A}(t,\tau) \\
=\; & -\int [e^{(t-\tau)\hat{L}^T_{FP}} A(\hat{\mathbf{x}})][L_\sigma p^G(\mathbf{x},\tau)]\delta_0(s-\tau)\, d\hat{\mathbf{x}} = -\int [L^T_\sigma e^{(t-\tau)\hat{L}^T_{FP}} A(\hat{\mathbf{x}})] p^G(\mathbf{x},\tau)\delta_0(s-\tau)\, d\hat{\mathbf{x}}, \qquad (78) \\
& (\vec{R}^g_{\mathbf{a},A})^T(t,\tau) \\
=\; & -\int [e^{(t-\tau)\hat{L}^T_{FP}} A(\hat{\mathbf{x}})][\nabla_{\mathbf{x}} p^G(\mathbf{x},\tau)]\delta_0(s-\tau)\, d\hat{\mathbf{x}} = \int [\nabla_{\mathbf{x}} e^{(t-\tau)L^T_{FP}} A(\mathbf{x})] p^G(\mathbf{x},\tau)\delta_0(s-\tau)\, d\mathbf{x}. \qquad (79)
\end{aligned}$$

Further more, we assume Q is either a constant matrix (corresponding to additive noise only) or a matrix with each entry a polynomial of degree no more than 2 in $\mathbf{X}$ (corresponding to simple multiplicative noise, i.e. with $\sigma(\mathbf{X})$ being a linear function in $\mathbf{X}$. This is consistent with stochastic mode reduction procedure [21], and hence we have the following properties when L_σ^T applied onto a polynomial $q(\mathbf{x})$.

simple multiplicative noise and perturbation	$\deg(L_\sigma^T q) \le \deg(q)$
additive noise and perturbation	$\deg(L_\sigma^T q) \le \max(\deg(q)-2, 0)$
simple multiplicative noise and additive perturbation	$\deg(L_\sigma^T q) \le \max(\deg(q)-1, 0)$
additive noise and simple multiplicative perturbation	$\deg(L_\sigma^T q) \le \max(\deg(q)-1, 0)$

We now restrict to the special case of quadratic nonlinearity $\mathbf{F}$ which is physically relevant such as quadratic advection term in fluid problems (16) [21]. In this case we can easily see the effect of L_{FP}^T applied to a polynomial q under simple multiplicative noise and perturbation assumption.

$$\deg(\hat{L}_{FP}^T q) \le \deg(q) + 1. \tag{80}$$

It is then easy to see that

$$\begin{aligned}
\vec{R}_{\mathbf{a},A}(t,t) &= \vec{R}^g_{\mathbf{a},A}(t,t) \quad A \text{ cubic} \\
R_{\sigma,A}(t,t) &= R^g_{\sigma,A}(t,t) \quad A \text{ quadratic under simple multiplicative noise and perturbation} \\
R_{\sigma,A}(t,t) &= R^g_{\sigma,A}(t,t) \quad A \text{ cubic under mixed additive/multiplicative noise and perturbation} \\
R_{\sigma,A}(t,t) &= R^g_{\sigma,A}(t,t) \quad A \text{ quadratic under additive noise and perturbation}
\end{aligned}$$

Simple calculation leads to

$$\begin{aligned}
\frac{d^k}{d\tau^k}\vec{R}^T_{\mathbf{a},A}(t,t)|_{\tau=t} &= \int \sum_{j=0}^{k}\sum_{i=0}^{j} C_k^j C_j^i [\nabla_{\mathbf{x}}(-\hat{L}_{FP}^T)^{k-j} A(\hat{\mathbf{x}})] \frac{\partial^{j-i}}{\partial t^{j-i}} \bar{p}(\mathbf{x},t)\delta_0^{(i)}(s-t)\, d\hat{\mathbf{x}}, \\
\frac{d^k}{d\tau^k}(\vec{R}^g_{\mathbf{a},A})^T(t,t)|_{\tau=t} &= \int \sum_{j=0}^{k}\sum_{i=0}^{j} C_k^j C_j^i [\nabla_{\mathbf{x}}(-\hat{L}_{FP}^T)^{k-j} A(\hat{\mathbf{x}})] \frac{\partial^{j-i}}{\partial t^{j-i}} p^G(\mathbf{x},t)\delta_0^{(i)}(s-t)\, d\hat{\mathbf{x}}, \\
\frac{d^k}{d\tau^k}R_{\sigma,A}(t,t)|_{\tau=t} &= -\int \sum_{j=0}^{k}\sum_{i=0}^{j} C_k^j C_j^i [L_\sigma^T(-\hat{L}_{FP}^T)^{k-j} A(\mathbf{x})] \frac{\partial^{j-i}}{\partial t^{j-i}} \bar{p}(\mathbf{x},t)\delta_0^{(i)}(s-t)\, d\hat{\mathbf{x}}, \\
\frac{d^k}{d\tau^k}R^g_{\sigma,A}(t,t)|_{\tau=t} &= -\int \sum_{j=0}^{k}\sum_{i=0}^{j} C_k^j C_j^i [L_\sigma^T(-\hat{L}_{FP}^T)^{k-j} A(\mathbf{x})] \frac{\partial^{j-i}}{\partial t^{j-i}} p^G(\mathbf{x},t)\delta_0^{(i)}(s-t)\, d\hat{\mathbf{x}}.
\end{aligned}$$

Combining the above and the fact that the first and second moments of $\frac{\partial^j}{\partial t^j}p(\mathbf{x},t)$ and $\frac{\partial^j}{\partial t^j}p^G(\mathbf{x},t)$ match, we have the following result on the validity of Gaussian approximation on short time interval.

Theorem 4 (Short time validity of Gaussian approximation)**.** *Suppose that the deterministic forcing term* $\mathbf{F}$ *in (1) is quadratic in* $\mathbf{X}$*, and we have at most simple multiplicative noise and noise perturbation, i.e.* $\sigma, \tilde{\sigma}$ *are linear in* $\mathbf{X}$*. Furthermore, we assume that the deterministic perturbation is defined by external forcing only, i.e.* $\mathbf{a}(\mathbf{X}) \equiv \mathbf{a}$*. Then*

- *for a linear functional $A(\mathbf{X})$, the Gaussian approximations defined in (78, 79) satisfy*

$$\vec{R}^g_{\mathbf{a},A}(t,\tau) = \vec{R}_{\mathbf{a},A}(t,\tau) + \mathcal{O}((t-\tau)^3). \tag{81}$$
$$R^g_{\sigma,A}(t,\tau) = R_{\sigma,A}(t,\tau) + \mathcal{O}((t-\tau)^2), \quad \textit{simple multiplicative noise and perturbation} \tag{82}$$
$$R^g_{\sigma,A}(t,\tau) = R_{\sigma,A}(t,\tau) + \mathcal{O}((t-\tau)^3), \quad \textit{mixed additive/multiplicative noise and perturbation} \tag{83}$$
$$R^g_{\sigma,A}(t,\tau) = R_{\sigma,A}(t,\tau) + \mathcal{O}((t-\tau)^4), \quad \textit{additive noise and perturbation} \tag{84}$$

- *For a quadratic functional $A(\mathbf{X})$, the Gaussian approximations defined in (78, 79) satisfy*

$$\vec{R}^g_{\mathbf{a},A}(t,\tau) = \vec{R}_{\mathbf{a},A}(t,\tau) + \mathcal{O}((t-\tau)^2), \tag{85}$$
$$R^g_{\sigma,A}(t,\tau) = R_{\sigma,A}(t,\tau) + \mathcal{O}(t-\tau), \quad \textit{simple multiplicative noise and perturbation} \tag{86}$$
$$R^g_{\sigma,A}(t,\tau) = R_{\sigma,A}(t,\tau) + \mathcal{O}((t-\tau)^2), \quad \textit{mixed additive/multiplicative noise and perturbation} \tag{87}$$
$$R^g_{\sigma,A}(t,\tau) = R_{\sigma,A}(t,\tau) + \mathcal{O}((t-\tau)^3), \quad \textit{additive noise and perturbation} \tag{88}$$

A disadvantage of the Gaussian approximations presented above with higher order accuracy is the need to build suitable efficient approximations to the backward operator, $e^{t\hat{L}^T_{FP}}A(\hat{\mathbf{X}})$, directly. We leave this topic for future research.

4 Computational Algorithms for Climate Response with Periodic Forcing

The mathematical framework developed in section 2 together with the zero noise adjoint form and the quasi-Gaussian approximations developed in sections 3.1 and 3.2.1 lead to new algorithms for computing the equilibrium response in a periodic system via the FDT theorem. The most important new practical application is computation of the changes in the equilibrium response to models for the climate system with time periodic forcing coefficients reflecting the seasonal cycle [12, 6, 4] with the prototype structure presented in (16). The goal here is to present the form of such algorithms for future applications to climate response; there are natural generalizations of the quasi-Gaussian FDT algorithms [18, 21, 15, 17, 14, 16, 1, 2, 3] following 3.2.1, the short time FDT algorithms [1, 2, 3] and the blended response algorithms [1, 3] to the situation with a seasonal cycle. The full structure of the skew-system formulation in section 2 together with the theory in section 2 will be utilized below.

The typical functionals $\hat{A}\begin{pmatrix}\mathbf{x}\\ s\end{pmatrix}$ for climate response have the separable form

$$\hat{A}\begin{pmatrix}\mathbf{x}\\ s\end{pmatrix} = \tilde{A}(\mathbf{x})\chi(|s-s_0|<P) \tag{89}$$

where $\chi(\mathcal{S})$ is the characteristic function of the set $\mathcal{S}$. For example, for the seasonal cycle so that the period T_0 is one year, we might be interested in the change in the low frequency teleconnection patterns [12, 6, 3] during each month or season so that $\tilde{A}(\mathbf{x})$ is $\mathbf{x}$ (corresponding to the mean) and/or $\mathbf{x}\otimes\mathbf{x}$ (corresponding to the second moments or variance), s_0 is the

fifteenth day of each month, and P is fifteen days. Similarly, we might be interested in the mean temperature change of its variance in each month at a specific location rather than just the annual changes [14, 16].

To develop approximations utilizing Theorem 2 for FDT or the quasi-Gaussian approximations sketched in 3.2.1, we need to first gather accurate statistics for the pdf $\hat{p}^{eq}\begin{pmatrix}\mathbf{x}\\ s\end{pmatrix}$. First from (17), we have the formula

$$\hat{p}^{eq}\begin{pmatrix}\mathbf{x}\\ s\end{pmatrix} = T_0^{-1} p_{per}(\mathbf{x}, s) \tag{90}$$

where according to theorem 1, $p_{per}(\mathbf{x}, s)$ is periodic in s with period T_0 and arises from a long time integration of the periodic dynamical systems in (1) with or without noise assuming ergodicity and strong mixing. It can be estimated from long time series of the system (see Theorem 1). Take the period interval T_0 and divide it into L equal intervals centered at $s_j, 1 \leq j \leq L$ with width $\Delta s = \frac{T_0}{L}$. Then, using the long time series, one can calculate the appropriate statistics of approximate pdf's denoted here by $p_{per}(\mathbf{x}, s_j)$ by doing conditional statistics of the time series of the trajectory to the sets, $\{t, |t - ks_j| < \frac{\Delta s}{2}, k > K_0\}$. There are two important points:

(A) Direct FDT Algorithm: If $x \in \mathbb{R}^N$ where N is low dimensional (roughly, $N \leq 4$), and (1) is a low order stochastic model [22, 4], then the entire pdf $p_{per}(\mathbf{x}, s_j)$ can be found with reasonably high accuracy [22]. In this situation, Theorem 2 on FDT can be applied directly with the functional

$$A(\hat{\mathbf{x}}) = \hat{A}(\mathbf{x})\hat{B}^{eq}(\mathbf{x}, s_j), \quad \hat{B}^{eq}_{\mathbf{a}}(\mathbf{x}, s_j) = \frac{\mathbf{L_a} p_{per}(\mathbf{x}, s_j)}{p_{per}(\mathbf{x}, s_j)}. \tag{91}$$

(B) Quasi-Gaussian FDT Algorithm: On the other hand, if $x \in \mathbb{R}^N$ with$N \gg 1$ as occurs in contemporary climate models [14, 16], one can calculate the low order statistics of $p_{per}(\mathbf{x}, s_j)$ involving the mean and covariance with reasonable precision and build the quasi-Gaussian approximation as suggested in section 3.2.1

$$\hat{B}^{G,eq}_{\mathbf{a}}(\mathbf{x}, s_j) = \frac{\mathbf{L_a} p^G_{per}(\mathbf{x}, s_j)}{p^G_{per}(\mathbf{x}, s_j)}. \tag{92}$$

Both the direct algorithm and the approximated algorithm require evaluation of the response operator, $\vec{R}_{\mathbf{a},A}(t)$, through the correlation with a suitable $\hat{B}$ in either (91) or (92), i.e. the (approximate) response operator, $\vec{R}_{\mathbf{a},A}(t)$, is given by

$$\begin{aligned}\vec{R}^T_{\mathbf{a},A}(t) &\approx < A(\hat{\mathbf{X}}(t)\hat{B}_{\mathbf{a}}(\hat{\mathbf{X}}(0)) > \\ &= \frac{1}{T_0}\int_0^{T_0}\int_{\mathbb{R}^N} \mathbb{E}A(\mathbf{X}(\mathbf{x}, t, s), t+s)\hat{B}_{\mathbf{a}}(\mathbf{x}, s) p_{per}(\mathbf{x}, s)\, d\mathbf{x} ds. \end{aligned} \tag{93}$$

In (93), $\mathbb{E}$ is the expectation, $\mathbf{X}(\mathbf{x}, t, s)$ is the trajectory of (1) satisfying the phase shifted dynamical systems,

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X}, t+s) + \sigma(\mathbf{X})\dot{W}, \quad \mathbf{X}(0) = \mathbf{x}. \tag{94}$$

Using ergodicity with respect to $\mathbf{X}$, the discrete approximation described in the paragraph below (90), and the special form for the functionals in (89), we have

$$\vec{R}^T_{\mathbf{a},A}(t) \approx \frac{1}{LK}\sum_{j=1}^{L}\sum_{k=1}^{K}\mathbb{E}\tilde{A}(\mathbf{X}(t+s_j+kT_0,s_j))\chi(|s_j+t+kT_0-s_0|\le P)\hat{B}(\mathbf{X}(s_j+kT_0,s_j)) \tag{95}$$

With the two approximate formulas for $\hat{B}$ in (91, 92), this leads to the **direct FDT** and **Quasi-Gaussian FDT algorithms** for systems with periodic forcing. Note that there is non-trivial phase averaging for general functionals like those in (89); if $\chi = [0, T_0]$ so that we are interested in only mean averaged statistics, then we can use $L \equiv 1$ in (95) coupled with (91, 92).

For the case of (1) with zero noise, the exact adjoint formula described in section 3.1 and similar considerations as above in (93, 95) leads to a general response algorithm for functionals of the form in (89)

$$\begin{aligned}&\vec{R}^T_{\mathbf{a},A}(t)\\ \approx\ &\frac{1}{LK}\sum_{j=1}^{L}\sum_{k=1}^{K}\nabla_{\mathbf{X}(s_j+kT_0,s_j)}\tilde{A}(\mathbf{X}(t+s_j+kT_0,s_j))\bullet\mathbf{a}(\mathbf{X}(s_j+kT_0,s_j))\chi(|s_j+t+kT_0-s_0|\le P)\hat{B}(\mathbf{X}(s_j\end{aligned}$$

for $K \gg 1$. We call this algorithm, the **short-time FDT algorithm** after [1, 2, 3]; no explicit knowledge of the time-periodic equilibrium measure, p_{per}, is needed. Clearly, these algorithms presented here can be combined in time to create blended response algorithms following [1, 3]. The accuracy of these proposed algorithms depends on the sampling width P in (89) and the number, L, of trajectories (the width Δs), as well as the functional $\tilde{A}(\mathbf{X})$, and the length of the time series available.

5 The Information Content in Linear Response

The linear response operator that we derived above can be used to calculate perturbation effect on **information content** which can be further utilized to determine the most sensitive direction under spatial-temporal separable perturbation using information content as the criterion. This is potentially quite useful in climate response studies [21] and for ensemble predictions (see Chapter 15 of [23] and the references therein).

The information content of a pdf $\hat{p}^\delta$ over another pdf $\bar{\hat{p}}$ is defined through the **relative entropy** [21] or the Lyapunov function [26]

$$\mathcal{P}(\hat{p}^\delta,\bar{\hat{p}}) = \int \hat{p}^\delta \ln(\frac{\hat{p}^\delta}{\bar{\hat{p}}})\left(= H(\hat{p}^\delta,\bar{\hat{p}})\right). \tag{98}$$

It measures the lack of information in $\bar{\hat{p}}$ compared with $\hat{p}^\delta$. Clearly in both climate response and ensemble prediction, the perturbations with the largest information deficit for $\bar{\hat{p}}$ are the most significant ones.

It is easy to see that the relative entropy is semi positive definite utilizing Jensen's inequality or the elementary inequality $\ln x \le x - 1$ for instance [23, 26]. Notice that it is not symmetric nor satisfying triangle inequality.

For our problem of complex system under periodic in time external forcing, there are two related concepts of statistical solutions: one associated with the skew-product system and the other linked to the original time-dependent system. The approach that we provide below is general enough to handle both situations. We will focus on the case of statistical solution to the skew-product system since statistical solution to the time-dependent system can be lifted to a (singular) statistical solution to the skew-product system via (14). We will remark on the case of information content in terms of the time-dependent system at the end of this section.

Recall that the perturbed Fokker-Planck equation with space-time separable deterministic perturbation $\mathbf{a} \bullet \tilde{\mathbf{F}}$ and same order noise perturbation takes the form of (26). We also recall that the *Fisher information I* associated with $\hat{p}^\delta$ is defined as

$$I(\hat{p}^\delta(t)) = \frac{1}{2}\int \frac{(\frac{d\hat{p}^\delta}{d\delta})^2\Big|_{\delta=0}}{\bar{\hat{p}}}\, d\hat{\mathbf{X}} = \frac{1}{2}\int \frac{\hat{p}'(\hat{\mathbf{X}},t)^2}{\bar{p}(\hat{\mathbf{X}},t)}\, d\hat{\mathbf{X}}. \tag{99}$$

Hence the relative entropy $\mathcal{P}$ is related to the Fisher information in the following fashion [21] after a simple manipulation based on the formal expansion for $\hat{p}^\delta = \hat{p}_0 + \delta\hat{p}' + \mathcal{O}(\delta^2)$.

$$\mathcal{P}(\hat{p}^\delta, \bar{\hat{p}}) = \delta^2 \tilde{P}(\hat{p}^\delta, \bar{\hat{p}}) + \mathcal{O}(\delta^3) \tag{100}$$

where

$$\tilde{P}(\hat{p}^\delta, \bar{\hat{p}}) = \frac{\delta^2}{2}\int \frac{(\hat{p}'(\hat{\mathbf{X}},t))^2}{\bar{\hat{p}}(\hat{\mathbf{X}},t)}\, d\hat{\mathbf{X}} = \delta^2 I(\hat{p}^\delta(t)). \tag{101}$$

In order to see the effect of system perturbation on relative entropy, we take the time derivative of the relative entropy and follow a classical argument on H-theorem [26]. Denoting $\hat{R}^\delta = \frac{\hat{p}^\delta}{\bar{\hat{p}}}$ and noticing that $\hat{R}^\delta = 1 + \delta\frac{\hat{p}'}{\bar{\hat{p}}} + \mathcal{O}(\delta^2)$, we have

$$\begin{aligned}
&\frac{d\mathcal{P}}{dt} \\
=\ & \int \{(1+\ln(\frac{\hat{p}^\delta}{\bar{\hat{p}}}))\frac{\partial \hat{p}^\delta}{\partial t} - \frac{\hat{p}^\delta}{\bar{\hat{p}}}\frac{\partial \bar{\hat{p}}}{\partial t}\} \\
=\ & \int \{\ln \hat{R}^\delta \frac{\partial \hat{p}^\delta}{\partial t} - \hat{R}^\delta \frac{\partial \bar{\hat{p}}}{\partial t}\} \\
=\ & \int \{\ln \hat{R}^\delta (\hat{L}_{FP}\hat{p}^\delta + \delta \mathbf{L_a}\hat{p}^\delta \cdot \tilde{\mathbf{F}} + \delta L_\sigma \hat{p}^\delta) - \hat{R}^\delta \hat{L}_{FP}\bar{\hat{p}}\} + \mathcal{O}(\delta^3) \\
=\ & \int \{(\hat{L}_{FP}^T \ln \hat{R}^\delta)\hat{p}^\delta + \delta^2 \frac{\hat{p}'}{\bar{\hat{p}}}(\mathbf{L_a}\bar{\hat{p}} \cdot \tilde{\mathbf{F}} + L_\sigma \bar{\hat{p}}) - \hat{R}^\delta \hat{L}_{FP}\bar{\hat{p}}\} + \mathcal{O}(\delta^3) \\
=\ & \int \{((\mathbf{F}(\mathbf{x},s)) \cdot \frac{\nabla \hat{R}^\delta}{\hat{R}^\delta} + \frac{1}{\hat{R}^\delta}\frac{\partial \hat{R}^\delta}{\partial s} + \frac{1}{2}Q : \nabla(\frac{\nabla \hat{R}^\delta}{\hat{R}^\delta}))\hat{p}^\delta + \delta^2 \frac{\hat{p}'}{\bar{\hat{p}}}(\mathbf{L_a}\bar{\hat{p}} \cdot \tilde{\mathbf{F}} + L_\sigma \bar{\hat{p}}) - \hat{R}^\delta \hat{L}_{FP}\bar{\hat{p}}\} + \mathcal{O}(\delta^3) \\
=\ & \int \{\bar{\hat{p}}(\mathbf{F}(\mathbf{x},s)) \cdot \nabla \hat{R}^\delta + \bar{\hat{p}}\frac{\partial \hat{R}^\delta}{\partial s} + \frac{1}{2}\bar{\hat{p}}Q : \nabla(\nabla \hat{R}^\delta) - \hat{R}^\delta \hat{L}_{FP}\bar{\hat{p}} - \frac{\hat{p}^\delta}{2(\hat{R}^\delta)^2}\nabla \hat{R}^\delta \cdot Q \nabla \hat{R}^\delta \\
& + \delta^2 \frac{\hat{p}'}{\bar{\hat{p}}}(\mathbf{L_a}\bar{\hat{p}} \cdot \tilde{\mathbf{F}} + L_\sigma \bar{\hat{p}})\} + \mathcal{O}(\delta^3)
\end{aligned}$$

$$
\begin{aligned}
&= \int \{\bar{\hat{p}}\hat{L}_{FP}^T\hat{R}^\delta - \hat{R}^\delta\hat{L}_{FP}\bar{\hat{p}} - \frac{\hat{p}^\delta}{2(\hat{R}^\delta)^2}\nabla\hat{R}^\delta\cdot Q\nabla\hat{R}^\delta + \delta^2\frac{\hat{p}'}{\bar{\hat{p}}}(\mathbf{L_a}\bar{\hat{p}}\cdot\tilde{\mathbf{F}} + L_\sigma\bar{\hat{p}})\} + \mathcal{O}(\delta^3)\\
&= -\int \frac{\bar{\hat{p}}^2}{2\hat{p}^\delta}\nabla\hat{R}^\delta\cdot Q\nabla\hat{R}^\delta + \delta^2\int\frac{\hat{p}'}{\bar{\hat{p}}}(\mathbf{L_a}\bar{\hat{p}}\cdot\tilde{\mathbf{F}} + L_\sigma\bar{\hat{p}}) + \mathcal{O}(\delta^3)\\
&= -\frac{\delta^2}{2}\int\bar{\hat{p}}\nabla\frac{\hat{p}'}{\bar{\hat{p}}}\cdot Q\nabla\frac{\hat{p}'}{\bar{\hat{p}}} + \delta^2\int\frac{\hat{p}'}{\bar{\hat{p}}}\mathbf{L_a}\bar{\hat{p}}\cdot\tilde{\mathbf{F}} + \delta^2\int\frac{\hat{p}'}{\bar{\hat{p}}}L_\sigma\bar{\hat{p}} + \mathcal{O}(\delta^3). \qquad (102)
\end{aligned}
$$

Therefore we have the following result

Proposition 1. *The information content represented by the relative entropy $\mathcal{P}$ can be estimated via the following general formula,*

$$
\begin{aligned}
&\mathcal{P}(\hat{p}^\delta(T),\bar{\hat{p}}(T))\\
&= -\frac{\delta^2}{2}\int_0^T\int\bar{\hat{p}}\nabla\frac{\hat{p}'}{\bar{\hat{p}}}\cdot Q\nabla\frac{\hat{p}'}{\bar{\hat{p}}} + \delta^2\int_0^T\int\frac{\hat{p}'}{\bar{\hat{p}}}\mathbf{L_a}\bar{\hat{p}}\cdot\tilde{\mathbf{F}} + \delta^2\int_0^T\int\frac{\hat{p}'}{\bar{\hat{p}}}L_\sigma\bar{\hat{p}} + \frac{\delta^2}{2}\int\frac{\hat{p}_0'(\hat{\mathbf{x}})^2}{\bar{\hat{p}}_0(\hat{\mathbf{x}})} + \mathcal{O}(\delta^3) \qquad (103)
\end{aligned}
$$

The first term correspond to the contribution of noise, the second term is the direct contribution of the perturbation in deterministic forcing, the third term is the result of perturbation in noise while the fourth (last) term is the leading order contribution from the perturbation in initial pdf.

For the quadratic in $\hat{p}'$ terms, we could use the explicit formula for $\hat{p}'$, (33), to reduce them to linear terms in $\hat{p}'$ and hence again linear response terms (see below). However such manipulation may lead to the usage of (high) derivatives of $\bar{\hat{p}}$ which may be hard to compute accurately in practice (unless with quasi-Gaussian approximation or its generalization, see below). Now if we are interested in statistical solutions generated by ensemble of the time-dependent system (9), i.e., $\bar{\hat{p}} = \bar{p}(\mathbf{x},t)\times\delta_0(s-t), \hat{p}^\delta = p^\delta(\mathbf{x},t)\times\delta_0(s-t)$, we see that the formula remain the same with the hat $\hat{\cdot}$ on the pdf's removed and integration with respect to $\hat{\mathbf{x}}$ replaced by $\mathbf{x}$.

5.1 Zero Noise Case

In the special case of zero noise, i.e. $Q\equiv 0$, the formula degenerates into one that is very similar to the perturbation away from equilibrium case [21]. Indeed, we have

$$
\begin{aligned}
&\mathcal{P}(\hat{p}^\delta(T),\bar{\hat{p}}(T))\\
&= \delta^2\int_0^T\int\hat{p}'(t)\frac{\mathbf{L_a}\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}\cdot\tilde{\mathbf{F}}(t)\,dtd\hat{\mathbf{x}} + \frac{\delta^2}{2}\int\frac{\hat{p}_0'(\hat{\mathbf{x}})^2}{\bar{\hat{p}}_0(\hat{\mathbf{x}})}\,d\hat{\mathbf{x}} + \mathcal{O}(\delta^3)\\
&= \delta^2\int\int_0^T\left(e^{t\hat{L}_{FP}}\hat{p}_0'(\hat{\mathbf{x}}) + \int_0^t[e^{(t-\tau)\hat{L}_{FP}}\mathbf{L_a}\bar{\hat{p}}(\tau)]\cdot\tilde{\mathbf{F}}(\tau)\,d\tau\right)\frac{\mathbf{L_a}\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}\cdot\tilde{\mathbf{F}}(t)\,dtd\hat{\mathbf{x}}\\
&\quad + \frac{\delta^2}{2}\int\frac{\hat{p}_0'(\mathbf{x})^2}{\bar{\hat{p}}_0(\hat{\mathbf{x}})}\,d\hat{\mathbf{x}} + \mathcal{O}(\delta^3)\\
&= \delta^2\int\int_0^T[e^{t\hat{L}_{FP}}\hat{p}_0'(\hat{\mathbf{x}})][\frac{\mathbf{L_a}\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}\cdot\tilde{\mathbf{F}}(t)]\,dtd\hat{\mathbf{x}} + \delta^2\int_0^t(\mathcal{R}_a(t,\tau)\cdot\tilde{\mathbf{F}}(\tau))\cdot\tilde{\mathbf{F}}(t)\,d\tau\,dt\\
&\quad + \frac{\delta^2}{2}\int\frac{\hat{p}_0'(\hat{\mathbf{x}})^2}{\bar{\hat{p}}_0(\hat{\mathbf{x}})}\,d\hat{\mathbf{x}} + \mathcal{O}(\delta^3) \qquad (104)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{R}_{\mathbf{a}}(t,\tau) &= \int \mathbf{L}_{\mathbf{a}}^T [e^{(t-\tau)\hat{L}_{FP}^T} \frac{\mathbf{L}_{\mathbf{a}}\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}]\bar{\hat{p}}(\tau)\, d\hat{\mathbf{x}} \\
&= \int [e^{(t-\tau)\hat{L}_{FP}^T} \frac{\mathbf{L}_{\mathbf{a}}\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}] \otimes [\frac{\mathbf{L}_{\mathbf{a}}\bar{\hat{p}}(\tau)}{\bar{\hat{p}}(\tau)}]\bar{\hat{p}}(\tau)\, d\hat{\mathbf{x}} \\
&= \int \hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{X}}(\hat{\mathbf{x}}, t-\tau), t) \otimes \hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{x}},\tau)\bar{\hat{p}}(\tau)\, d\hat{\mathbf{x}} \qquad (105) \\
&= \int_{\mathbb{R}^N}\int_{\mathbb{T}} \hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{X}}(\begin{pmatrix}\mathbf{x}\\ s\end{pmatrix}, t-\tau), t) \otimes \hat{\mathbf{B}}_{\mathbf{a}}(\begin{pmatrix}\mathbf{x}\\ s\end{pmatrix},\tau)\bar{\hat{p}}(\begin{pmatrix}\mathbf{x}\\ s\end{pmatrix},\tau)\, ds\, d\mathbf{x} \qquad (106) \\
&= < \hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{X}}(t), t) \otimes \hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{X}}(\tau),\tau) > \qquad (107)
\end{aligned}
$$

is an auto-correlation matrix of $\hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{x}},\tau) = \frac{\mathbf{L}_{\mathbf{a}}\bar{\hat{p}}(\hat{\mathbf{x}},\tau)}{\bar{\hat{p}}(\hat{\mathbf{x}},\tau)}$ (see Theorem 2.2 in [21]). The inner integral may be considered as average over the phases.

Proposition 2. *In the case when the external perturbation $\tilde{\mathbf{F}}(t) = \tilde{\mathbf{F}}$ is time-independent, only the symmetric part of $\mathcal{R}_{\mathbf{a}}$, namely $\mathcal{R}_{\mathbf{a},sym}(t,\tau) = \frac{1}{2}(\mathcal{R}_{\mathbf{a}} + \mathcal{R}_{\mathbf{a}}^T)$ is relevant in the information content formula above, i.e., we have*

$$
\begin{aligned}
& \mathcal{P}(\hat{p}^\delta(T), \bar{\hat{p}}(T)) \\
= \;& \delta^2 \tilde{\mathbf{F}} \cdot \int_0^T \int_0^t \mathcal{R}_{\mathbf{a},sym}(t,\tau)\, d\tau dt \tilde{\mathbf{F}} + \delta^2 \int \int_0^T [e^{t\hat{L}_{FP}^T} \hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{x}},t)]\, dt \cdot \tilde{\mathbf{F}} \hat{p}_0'(\hat{\mathbf{x}})\, d\hat{\mathbf{x}} + \frac{\delta^2}{2} \int \frac{\hat{p}_0'(\hat{\mathbf{x}})^2}{\bar{\hat{p}}_0(\hat{\mathbf{x}})}\, d\hat{\mathbf{x}} + \mathcal{O}(\delta^3) \\
\stackrel{def}{=} \;& \delta^2 T\, \tilde{\mathbf{F}} \cdot \mathcal{M}_{\mathbf{a},T} \tilde{\mathbf{F}} + \delta^2 \int \int_0^T [e^{t\hat{L}_{FP}^T} \hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{x}},t)]\, dt \cdot \tilde{\mathbf{F}} \hat{p}_0'(\hat{\mathbf{x}})\, d\hat{\mathbf{x}} + \frac{\delta^2}{2} \int \frac{\hat{p}_0'(\hat{\mathbf{x}})^2}{\bar{\hat{p}}_0(\hat{\mathbf{x}})}\, d\hat{\mathbf{x}} + \mathcal{O}(\delta^3) \qquad (108)
\end{aligned}
$$

where

$$
\mathcal{M}_{\mathbf{a},T} \stackrel{def}{=} \frac{1}{T} \int_0^T \int_0^t \mathcal{R}_{\mathbf{a},sym}(t,\tau)\, d\tau dt. \qquad (109)
$$

The first term represent contribution from pure (independent of perturbation in initial pdf, i.e. $\hat{p}_0'$) system perturbation in the space-time separable fashion considered in this manuscript, the second term is the cross contribution to information content of perturbation in forcing and perturbation in initial pdf, the third term is the contribution to the information content due to the perturbation in initial pdf.

Furthermore, it is easy to see the following:

- The contribution to the information content due purely to the perturbation in the initial pdf is positive, i.e. $\int \frac{\hat{p}_0'(\mathbf{x})^2}{\bar{\hat{p}}_0(\mathbf{x})}\, d\hat{\mathbf{x}} \geq 0$. In the absence of perturbation in forcing ($\hat{\mathbf{B}}_{\mathbf{a}} \equiv 0$), the relative entropy (information content) is a constant in time in the absence of noise [13, 26] which is consistent with the relation above to leading order.

- The contribution to the information content due to perturbation in forcing alone must be positive, i.e.

$$
\mathcal{M}_{\mathbf{a},T} \geq 0. \qquad (110)
$$

  Indeed, for the special case of no perturbation in initial pdf, i.e. $\hat{p}_0' \equiv 0$, $\tilde{\mathbf{F}} \cdot \mathcal{M}_{\mathbf{a},T} \tilde{\mathbf{F}}$ is the leading order term of the non-negative function $\mathcal{P}(\hat{p}^\delta(T), \bar{\hat{p}}(T))$ and hence must

be non-negative definite. It then implies that $\mathcal{M}_{\mathbf{a},T} \geq 0$ since the matrix is symmetric and $\tilde{\mathbf{F}}$ is arbitrary. Notice that the definition of $\mathcal{M}_{\mathbf{a},T}$ is completely independent of the perturbation $\hat{p}_0'$ to the initial pdf. Hence the conclusion remains valid for general non-zero $\hat{p}_0'$ case. A very useful corollary to this observation is that *the direction of the eigenvector associated with the largest eigenvalue of $\mathcal{M}_{\mathbf{a},T}$ is the direction with the largest information content in response at time T in the absence of perturbation in initial pdf.* This is a generalization of a similar result valid for equilibrium unperturbed pdf [21]. In practical computation, one may approximate $\mathcal{R}_{\mathbf{a}}$ via the quasi-Gaussian approach described in section 3.2.1 above. Namely we approximate $\mathcal{R}_{\mathbf{a}}$ by

$$\begin{aligned}\mathcal{R}_{\mathbf{a}}(t,\tau) &\sim \mathcal{R}_{\mathbf{a}}^G(t,\tau) \\ &= \int [e^{(t-\tau)\hat{L}_{FP}^T}\hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{x}},t)] \otimes \hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{x}},\tau)\bar{\hat{p}}(\hat{\mathbf{x}},\tau)\, d\hat{\mathbf{x}} \\ &= \int \hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{X}}(\hat{\mathbf{x}},t-\tau),t)] \otimes \hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{x}},\tau)\bar{\hat{p}}(\hat{\mathbf{x}},\tau)\, d\hat{\mathbf{x}} \qquad (111)\end{aligned}$$

where $\hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{x}},\tau) = \frac{\mathbf{L}_{\mathbf{a}}\hat{p}^G(\mathbf{x},s,\tau)}{\hat{p}^G(\mathbf{x},s,\tau)}$ with $\hat{p}^G(\cdot,s,\tau)$ being the Gaussian that has the same first and second moments (in $\mathbf{x}$) as $\bar{\hat{p}}(\begin{pmatrix} \cdot \\ s \end{pmatrix},\tau)$. In the case of zero noise and the unperturbed pdf $\bar{\hat{p}}$ is given by a finite ensemble like (58), this approximation can be represented as

$$\mathcal{R}_{\mathbf{a}}^G(t,\tau) = \sum_{j=1}^{R} p_j \int \hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{x}}_j(t),t) \otimes \hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{x}}_j(\tau),\tau). \qquad (112)$$

- The contribution from the cross term has no definite sign. However, the overall contribution of the three terms must be non-negative since it is the leading order expansion of the relative entropy which is non-negative. The computation of this cross term can be handled using the quasi-Gaussian approximation idea as well.

$$\begin{aligned}\int\int_0^T [e^{t\hat{L}_{FP}^T}\hat{\mathbf{B}}_{\mathbf{a}}(\hat{\mathbf{x}},t)]\, dt \cdot \tilde{\mathbf{F}}\hat{p}_0'(\hat{\mathbf{x}})\, d\hat{\mathbf{x}} &\sim \int\int_0^T [e^{t\hat{L}_{FP}^T}\hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{x}},t)]\, dt \cdot \tilde{\mathbf{F}}\hat{p}_0'(\hat{\mathbf{x}})\, d\hat{\mathbf{x}} \\ &= \int\int_0^T \hat{\mathbf{B}}_{\mathbf{a}}^G(\hat{\mathbf{X}}(\hat{\mathbf{x}},t),t)]\, dt \cdot \tilde{\mathbf{F}}\hat{p}_0'(\hat{\mathbf{x}})\, d\hat{\mathbf{x}}. \qquad (113)\end{aligned}$$

  This approximate formula together with the approximate formula for the auto-correlation (111) can be utilized to investigate the relationships on $\tilde{\mathbf{F}}, \mathbf{a}, \hat{p}_0'$ (the perturbations) that maximize (or minimize) the information content.

In the special case of $\bar{\hat{p}} = \hat{p}^{eq}$ being an equilibrium pdf of the system, i.e., $\hat{L}_{FP}\bar{\hat{p}} = 0$, we recover a formula which is almost the same as the one from [21, 17]. Indeed, we have

$$\mathcal{R}_{\mathbf{a}}(t,\tau) = \mathcal{R}_{\mathbf{a}}(t-\tau,0) \stackrel{def}{=} \mathcal{R}_{\mathbf{a}}(t-\tau) \qquad (114)$$

and hence assuming the auto-correlation matrix decay fast enough

$$\mathcal{M}_{\mathbf{a},T} = \frac{1}{T}\int_0^T (T-t)\mathcal{R}_{\mathbf{a},sym}(t)\, dt \stackrel{T\to\infty}{\longrightarrow} \mathcal{M}_{\mathbf{a}}(\infty) = \int_0^\infty \mathcal{R}_{\mathbf{a},sym}(t)\, dt \qquad (115)$$

where

$$\mathcal{R}_{\mathbf{a},sym}(t,\tau) = \frac{1}{2}(\mathcal{R}_{\mathbf{a}} + \mathcal{R}_{\mathbf{a}}^T) \tag{116}$$

$$\mathcal{R}_{\mathbf{a}}(t) = \int_{\mathbb{R}^N}\int_{\mathbb{T}} \hat{\mathbf{B}}_{\mathbf{a}}^{eq}(\hat{\mathbf{X}}((\begin{smallmatrix}\mathbf{x}\\ s\end{smallmatrix}),t)) \otimes \hat{\mathbf{B}}_{\mathbf{a}}^{eq}((\begin{smallmatrix}\mathbf{x}\\ s\end{smallmatrix}))\hat{p}^{eq}((\begin{smallmatrix}\mathbf{x}\\ s\end{smallmatrix}))\, ds\, d\mathbf{x} \tag{117}$$

$$\hat{\mathbf{B}}_{\mathbf{a}}^{eq}(\hat{\mathbf{x}}) = \frac{\mathbf{L}_{\mathbf{a}}\hat{p}^{eq}(\hat{\mathbf{x}})}{\hat{p}^{eq}(\hat{\mathbf{x}})}. \tag{118}$$

Thus, *the natural low frequency basis for long times under constant external forcing perturbation, zero perturbation in initial data, is the one which diagonalizes the non-negative symmetric matrix* $\mathcal{M}_{\mathbf{a}}(\infty)$.

In the special case of $\mathbf{a}$ being a constant vector, i.e., perturbation in external forcing only, and the equilibrium state is Gaussian in $\mathbf{x}$, the phase averaged auto-correlation matrix (117) is essentially the auto-correlation of $\mathbf{x}$ (it is, if the mean is zero).

The inclusion of the effect of initial condition is important conceptually. For instance, this is potentially useful for us to address the question of for a given external perturbation class, which kind of perturbation generate largest additional information content in an ensemble prediction.

5.2 The Effect of Noise

In the presence of noise, we first observe that *the effect of the noise term is to decrease relative entropy* as is clear from the semi-positive definiteness of $\bar{\hat{p}}\nabla\frac{\hat{p}'}{\bar{\hat{p}}}\cdot Q\nabla\frac{\hat{p}'}{\bar{\hat{p}}}$. This is in accordance with the general result that noise reduces relative entropy (H-theorem, [13, 26]) although the two statistical solutions under investigation here are not for the same system (different parameter) since we are studying system perturbation.

The contribution from perturbation in noise, $\int_0^T\int \frac{\hat{p}'}{\bar{\hat{p}}}L_\sigma\bar{\hat{p}}$, does not seem to have a definite sign in general. In the case without perturbation in deterministic forcing ($\mathbf{L}_{\mathbf{a}} = 0$) and no perturbation in initial pdf ($\hat{p}_0' = 0$), this term can be represented as

$$\begin{aligned}
\int_0^T\int \frac{\hat{p}'(t)}{\bar{\hat{p}}(t)}L_\sigma\bar{\hat{p}}(t)\, d\hat{\mathbf{x}}dt &= \int_0^T\int\int_0^t [e^{(t-\tau)\hat{L}_{FP}}L_\sigma\bar{\hat{p}}(\tau)]\frac{L_\sigma\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}\, d\tau d\hat{\mathbf{x}}dt \\
&= \int_0^T\int\int_0^t [L_\sigma\bar{\hat{p}}(\tau)][e^{(t-\tau)\hat{L}_{FP}^T}\frac{L_\sigma\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}]\, d\tau d\hat{\mathbf{x}}dt \\
&= \int_0^T\int\int_0^t [e^{(t-\tau)\hat{L}_{FP}^T}\frac{L_\sigma\bar{\hat{p}}(t)}{\bar{\hat{p}}(t)}][\frac{L_\sigma\bar{\hat{p}}(\tau)}{\bar{\hat{p}}(\tau)}]\bar{\hat{p}}(\tau)\, d\tau d\hat{\mathbf{x}}dt \\
&= \int_0^T\int_0^t\int [e^{(t-\tau)\hat{L}_{FP}^T}\hat{B}_\sigma(t)][\hat{B}_\sigma(\tau)]\bar{\hat{p}}(\tau)\, d\hat{\mathbf{x}}d\tau dt
\end{aligned} \tag{119}$$

which is the time integral of the auto-correlation of $\hat{B}_\sigma(\hat{\mathbf{x}},\tau) = \frac{L_\sigma\bar{\hat{p}}(\hat{\mathbf{x}},\tau)}{\bar{\hat{p}}(\hat{\mathbf{x}},\tau)}$. This can be approximated via quasi-Gaussian approximation among others in practice.

Our second observation is that contribution to the information content of the noise term through $\tilde{\mathbf{F}}$ can be described via a semi-negative definite matrix so long as there is no perturbation to the initial pdf, i.e. $\hat{p}_0' = 0$, and there is no perturbation in noise, i.e. $L_\sigma = 0$.

Indeed, let

$$V(t) = \int_0^t \nabla \frac{[e^{(t-\tau)\hat{L}_{FP}} \mathbf{L_a} \bar{\hat{p}}(\tau)]}{\bar{\hat{p}}(t)} \, d\tau$$

we have, after utilizing the formula for perturbative pdf (33), and assuming no perturbation in initial pdf ($\hat{p}_0' = 0$), no noise perturbation ($L_\sigma = 0$) and constant external perturbation ($\tilde{\mathbf{F}}(t) \equiv \tilde{\mathbf{F}}$)

$$\begin{aligned} \frac{1}{2}\int_0^T \int \bar{\hat{p}}(t) \nabla \frac{\hat{p}'(t)}{\bar{\hat{p}}(t)} \cdot Q \nabla \frac{\hat{p}'(t)}{\bar{\hat{p}}(t)} \, d\hat{\mathbf{x}} dt &= \frac{1}{2}\int_0^T \int \bar{\hat{p}}(t)(V(t)\tilde{\mathbf{F}}) \cdot Q(V(t)\tilde{\mathbf{F}}) \, d\hat{\mathbf{x}} dt \quad (120) \\ &= \tilde{\mathbf{F}} \cdot \int_0^T \int \frac{\bar{\hat{p}}(t)}{2} V^T(t) Q V(t) \, d\hat{\mathbf{x}} dt \tilde{\mathbf{F}} \quad (121) \\ &= T\tilde{\mathbf{F}} \cdot \mathcal{V}_T \tilde{\mathbf{F}} \quad (122) \end{aligned}$$

where

$$\mathcal{V}_T = \frac{1}{2T}\int_0^T \int \bar{\hat{p}}(t) V^T(t) Q V(t) \, d\hat{\mathbf{x}} dt \quad (123)$$

is a symmetric semi-positive definite matrix.

Proposition 3. *In the special case of no perturbation in initial pdf ($\hat{p}_0' = 0$), no perturbation in noise ($L_\sigma = 0$), and constant perturbation in external forcing ($\tilde{\mathbf{F}}(t) = \tilde{\mathbf{F}}$), we have*

$$\begin{aligned} \mathcal{P}(\hat{p}^\delta(T), \bar{\hat{p}}(T)) &= -\frac{\delta^2}{2}\int_0^T \int \bar{\hat{p}} \nabla \frac{\hat{p}'}{\bar{\hat{p}}} \cdot Q \nabla \frac{\hat{p}'}{\bar{\hat{p}}} + \delta^2 \int_0^T \int \frac{\hat{p}'}{\bar{\hat{p}}} \mathbf{L_a} \bar{\hat{p}} \cdot \tilde{\mathbf{F}} + \mathcal{O}(\delta^3) \\ &= T\delta^2 \tilde{\mathbf{F}} \cdot (-\mathcal{V}_{T,Q} + \mathcal{M}_{\mathbf{a},T})\tilde{\mathbf{F}} + \mathcal{O}(\delta^3) \quad (124) \end{aligned}$$

where $\mathcal{V}_{T,Q}$ is semi-positive definite while $\mathcal{M}_{\mathbf{a},T}$ is semi-positive definite. Hence the information content is a tug war between noise and system perturbation in constant forcing. The first tends to diminish information content while the latter tends to increase the information content.

The positivity of $\mathcal{M}_{\mathbf{a},T}$ follows from the positivity of $\mathcal{P}$ and the semi-positivity of $\mathcal{V}_{T,Q}$. However, the difference of the two matrices must be semi-positive definite since $\mathcal{P}(T) \geq 0$. Therefore we can still conclude that the direction which is most sensitive to information flow at time T is the direction of the eigenvector associated with the largest eigenvalue of the semi-positive definite symmetric matrix $-\mathcal{V}_{T,Q} + \mathcal{M}_{\mathbf{a},T}$. This may be viewed as a generalization of the case without noise discussed in the previous subsection. Quasi-Gaussian type approximation may be used to approximate the matrices for practical purposes in studies of model error.

6 Concluding Discussion

Several generalizations of linear response theory have been developed here along with proposed numerical algorithms with potentially significant applications to both finite time ensemble predictions and climate response with time periodic forcing. For ensemble predictions,

new formulas to evaluate model error combined with initial ensemble perturbation [27] have been developed including variation in model noise [25]; an information theoretic perspective has been used in section 5 to assess these perturbations where the current model has the largest information deficit. When applied to the equilibrium distribution, this allows for assessing the most important perturbations with the largest response for a given functional (see Theorem 2.2 of [21]).

A general framework has been developed in sections 2 and 3 for the general fluctuation response of a (stochastic) chaotic dynamical system with periodic coefficients as arises with the seasonal cycle for climate response experiments or the diurnal cycle for moist convection. This framework leads to new algorithms presented in section 4 for calculating the low frequency climate response in periodic systems like those with a seasonal cycle.

All of the algorithms and concepts developed here in a theoretical framework for both climate response and ensemble prediction require further extensive testing and development to assess their performance. The suite of test models should range from low order stochastic models [22], to versions of the L-96 and geophysical equilibrium statistical mechanics models [23] to intermediate [3] and comprehensive climate models [14, 16, 6, 12] with realistic dissipation and forcing. The authors intend to do this in the near future with various collaborators.

Acknowledgments

This work is supported in part by grants from NSF DMS0456713(for AJM) and DMS0606671(for XW), and ONR grant N00014-05-0164(for AJM).

References

[1] Abramov, R.V. and Majda, A.J.; *Blended Response Algorithms for Linear Fluctuation-Dissipation for Complex Nonlinear Dynamical Systems*, Nonlinearity, 20, Issue 12, 2007, 2793 2821.

[2] Abramov, R.V. and Majda, A.J.; *New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems*, J. Nonlinear Science, 2007, published electronically.

[3] Abramov, R.V. and Majda, A.J.; *New Algorithms for Low Frequency Climate Response*, J. Atmos. Sci., accepted and in press, 2008.

[4] Achatz, U. and Opsteegh, J.D.; *Primitive-equation-based low order models with seasonal cycle.* J. Atmos. Sci., vol. 60, 649-690, 2002.

[5] Balescu, R.; *Statistical Dynamics.* London, Imperial College Press, 1997.

[6] Branstator, G. and Frederiksen J.; *The Seasonal Cycle of Interannual Variability and the Dynamical Imprint of the Seasonally Varying Mean State*, Journal of the Atmospheric Sciences Volume 60, Issue 13 (July 2003) pp. 1577-1592, DOI: 10.1175/3002.1

[7] Cheng, W.; Wang, X.; *A Uniformly Dissipative Scheme for Stationary Statistical Properties of the Infinite Prandtl Number Model*, Applied Mathematics Letters, in press. doi:10.1016/j.aml.2007.07.036

[8] Cheng, W. and Wang, X.; *A semi-implicit scheme for stationary statistical properties of the infinite Prandtl number model*, SIAM J. Num. Anal., vol.47, no.1, 250-270, 2008.. DOI. 10.1137/080713501

[9] Durrett, R.; *Stochastic Calculus, a practical introduction*, CRC Press, Boca Raton, Florida, 1996.

[10] E, W; D. Li, *The Andersen thermostat in molecular dynamics*, Preprint, 2007.

[11] Foias, C., Manley, O., Rosa, R. and Temam, R.; *Navier-Stokes Equations and Turbulence*, Cambridge University Press, Cambridge, UK, 2001.

[12] Frederiksen, J.S. and Branstator, G.; *Seasonal variability of teleconnection patterns.* J. Atmos. Sci., vol. 62, 1346-1365, 2005.

[13] Gardiner, C.W.; *Handbook of Stochastic Methods, 3rd ed*, Springer-Verlag, New York, 2004.

[14] Gritsun, A.S.; Branstator, G.; *Climate Response Using a Three-Dimensional Operator Based on the FluctuationCDissipation Theorem*, Journal of the Atmospheric Sciences Volume 64, Issue 7 (July 2007) pp. 2558C2575 DOI: 10.1175/JAS3943.1

[15] Gritsoun, A.S., Branstator, G., and Dymnikov, V.P.; *Construction of the linear response operator of an atmospheric general circulation model to small external forcing*, Russian J. Numer. Anal. Math. Moddeling, **17** (2002), 399-416.

[16] Gritsun, A.S.; Branstator, G.; and Majda A.J.; *Climate response of linear and quadratic functionals using the fluctuation-dissipation theorem*, J. Atmos. Sci., accepted and in press, 2008.

[17] Gritsoun, A.S. and Dymnikov, V.P.; *Barotropic atmosphere response to small external actions: theory and numerical experiments*, zv. Russ. Acad. Sci. Atmosphere Ocean Phys. **35** (1999), no.5, 511-525.

[18] Leith, C.; *Climate response and fluctuation-dissipation.* J. Atmos. Sci., vol. 32, 2022-2025, 1975.

[19] Kalnay, E., *Atmospheric modeling, data assimilation and predictability*, Cambridge University Press, New York, 2003.

[20] Kubo, R., Toda, M., and Hashitsume, N., *Statistical physics II: nonequilibrium statistical mechanics*, Springer-Verlag, New York, 1985.

[21] Majda, A.J., Abramov, R., and Grote, M.; *Information theory and stochastics for multiscale nonlinear systems*, CRM monograph series, American Mathematical Society, 2005.

[22] Majda, A.J., Franzke, C., and Khouider, B.; *An applied mathematics perspective on stochastic modelling for climate*. Phil. Trans. R. Soc. A., vol. 366, 2429-2455, 2008.

[23] Majda, A.J. and Wang, X.; *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*. Cambridge University Press, 2006.

[24] Nicolis, C.; *Dynamics of Model Error: Some Generic Features*, J. Atmos. Sci., Vol. 60, no. 17 (Sept. 2003) pp. 2208-2218

[25] Palmer, T.N.; *A nonlinear dynamical perspective on model error: a proposal for non-local stochastic-dynamic parameterizations in weather and climate prediction models*. Q. J. R. Meteorol. Soc., vol. 127, 279-304, 2001.

[26] Risken, H.; *The Fokker-Planck Equations*, Springer-Verlag, Berlin, 1984.

[27] Tribbia, J.J. and Baumhefner, D.P.; *The reliability of improvements in deterministic short-range forecasts in the presence of initial state and modelling deficiencies*. Mon. Weather Rev., vol. 116, 2276-2288, 1988.

[28] Vishik, M.I.; Fursikov, A.V., *Mathematical Problems of Statistical Hydromechanics*. Kluwer Acad. Publishers, Dordrecht/Boston/London, 1988.

[29] Wang, X.; *Upper Semi-Continuity of Stationary Statistical Properties of Dissipative Systems*. Dedicated to Prof. Li Ta-Tsien on the occasion of his 70th birthday. Discrete and Continuous Dynamical Systems, A. Vol. 23, no.1/2, pp.521-540, 2009. doi:10.3934/dcds.2009.23.521

[30] Wang, X.; *Approximation of stationary statistical properties of dissipative dynamical systems: time discretization*, Preprint, 2008.