Chin. Ann. Math.
32B(3), 2011, 343–368
DOI: 10.1007/s11401-011-0647-2

Chinese Annals of Mathematics, Series B

Invariant Measures and Asymptotic Gaussian Bounds for Normal Forms of Stochastic Climate Model*

Yuan YUAN[1] Andrew J. MAJDA[2]

Abstract The systematic development of reduced low-dimensional stochastic climate models from observations or comprehensive high dimensional climate models is an important topic for atmospheric low-frequency variability, climate sensitivity, and improved extended range forecasting. Recently, techniques from applied mathematics have been utilized to systematically derive normal forms for reduced stochastic climate models for low-frequency variables. It was shown that dyad and multiplicative triad interactions combine with the climatological linear operator interactions to produce a normal form with both strong nonlinear cubic dissipation and Correlated Additive and Multiplicative (CAM) stochastic noise. The probability distribution functions (PDFs) of low frequency climate variables exhibit small but significant departure from Gaussianity but have asymptotic tails which decay at most like a Gaussian. Here, rigorous upper bounds with Gaussian decay are proved for the invariant measure of general normal form stochastic models. Asymptotic Gaussian lower bounds are also established under suitable hypotheses.

Keywords Reduced stochastic climate model, Invariant measure, Fokker-Planck equation, Comparison principle, Global estimates of probability density function

2000 MR Subject Classification 60H10, 60H30, 60E99, 35Q84

1 Introduction

There is a recent interest in deriving reduced stochastic models for climate and extended-range weather prediction. An attractive property of atmospheric low-frequency variability is that it can be efficiently described by just a few large-scale teleconnection patterns. These patterns exert a huge impact on surface climate and seasonal predictability. Thus, such reduced stochastic models are an attractive alternative for extended range prediction and climate sensitivity studies because they are computationally much more efficient than state-of-the-art climate models and were shown to have a comparable prediction skill (see [19]). Because such reduced models capture the essence of low-frequency processes they allow for a better understanding of the climate system. Reduced models can also be used for climate sensitivity and climate change studies via the fluctuation dissipation theorem (FDT) (see [21, 24, 16]). Thus,

Manuscript received March 4, 2010. Revised May 20, 2010. Published online April 19, 2011.
[1]Department of Mathematics, Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA. E-mail: yuan@cims.nyu.edu
[2]Department of Mathematics, and Center for Atmospheric Ocean Sciences, Courant Institute, New York University, 251 Mercer Street, New York, NY 10012, USA. E-mail: jonjon@cims.nyu.edu
*Project supported by the National Science Foundation Grant (No. DMS-0456713) and the Office of Naval Research Grant (No. N0014-05-1-1064).

systematic strategies are essential to successfully developing reduced models. The recently developed stochastic mode reduction strategy provides a systematic procedure for the derivation of reduced stochastic models (see [17–20, 22–23]). Using these techniques, a systematic normal form for reduced climate models which respects the important physical constraints of energy conservation in the original system has been developed recently (see [18]). These normal forms predict simultaneously cubic damping and correlated additive and multiplicative noise and this normal form is confirmed in data for low frequency teleconnection patterns (see [18]).

Much more naive adhoc analysis in [28] suggests a linear correlated additive and multiplicative (CAM) noise model for low frequency climate variables. The linear correlated additive and multiplicative noise model predicts a PDF whose tails decay according to a power-law. This certainly contradicts studies which examined a subspace spanned by the first few leading EOFs from data generated by long integrations of general circulation models in [1] and reanalysis data in [28, 29]. These studies find only small, though significant deviations from Gaussianity. These deviations occur most strongly in form of excess skewness and to a much lesser extent in form of excess kurtosis. Thus, there is not much empirical evidence for power-law decay of the tails of PDFs of the leading EOFs. However, this argument applies for low-frequency data (measured by the leading EOFs) and does not rule out the possible usefulness of the linear CAM model for other climate variables at high frequencies in station data in [30].

Thus, it is interesting to show rigorously that for the normal form in [18] for reduced stochastic climate models the joint PDF of its stationary distribution exists and has at most Gaussian decay thanks to the nonlinear cubic damping, and this is the main topic of this paper. Next, we make some informal remarks to connect this topic with recent mathematical developments.

Let us assume that the cubic nonlinearity gives strict energy dissipation. The existence and uniqueness of the pathwise solutions to our reduced climate model are known (see [15]). The existence of the invariant measure follows from the boundedness of the second moment of the pathwise solution with the standard tightness argument (see [6]). Moreover, the existence of the probability density function of this invariant measure was studied in [26].

If we further require that our climate model have additive dyads in all the low frequency modes such that it has uniformly non-degenerate diffusion, the uniqueness of the invariant measure follows from strong Feller property and irreducibility of the Markov transition semigroup in [6] and the smoothness of the density function was obtained in [3, 5]. Recent studies about global upper bounds of the probability density function in [25, 9] can be applied to our reduced climate model as well. However, uniformly non-degenerate diffusion is not necessary for the normal form for the stochastic climate modelling, for example, the 2-dimensional climate model (3.9) in Section 3.1.1. Moreover, global lower bound of the probability density function under some boundness assumption on the diffusion coefficients was studied in [4]. However, our climate model here can always have quadratic diffusion coefficients when there is CAM noise, so this earlier work cannot apply.

In this paper, we will mainly show that the density function of the stationary statistical solution (invariant measure) to the reduced stochastic model for N-climate variable has at most Gaussian decay as long as it belongs to C^2 and has power law decay. Moreover, we find sufficient conditions such that there exists a Gaussian lower bound which means that the global decaying rate of the density function can be exactly Gaussian. As a corollary, the global decaying behavior can be classified rigorously under the uniformly non-degenerate diffusion assumption. The idea of showing the global decay rate is to construct Gaussian comparison functions and use the comparison principle argument for the equilibrium Fokker-Planck equation to show that the tails of the stationary statistical solution t to the reduced stochastic climate model can be bounded above or bounded below by Gaussian comparison functions with different variances. Applying the comparison principle argument, we avoid the uniformly non-degenerate assumption, although we require that the density belong to C^2. Numerical evidence shows that C^2 is a reasonable assumption for our climate model, for example, the 2-dimensional climate model (3.9) in Section 3.1.1. Below, we present the general stochastic large scale geophysical flow model and its reduced system for the climate variables in Section 2. In Section 3, we explain our intuition through examples and the main results. Finally, we will present the proof in Section 4 step by step.

2 Normal Form Stochastic Models for N-Climate Variables

Here, we consider the dynamical core of comprehensive large scale geophysical flow models

$$\frac{\mathrm{d}u}{\mathrm{d}t} = F + Lu + B(u,u), \tag{2.1}$$

where F is a constant forcing, L is linear, B is a quadratically nonlinear operator and conserve energy (see [17]). An important example of quadratically nonlinear equations of the type as in (2.1) could be the barotropic flow on a beta plane with topography and mean flow (see [22]).

We can partition the variable $u = (u_1(t), u_2(t), \cdots, u_{N+M}(t))$ into low-frequency climate modes

$$x(t) = (u_1(t), u_2(t), \cdots, u_N(t)) \in \mathbb{R}^N$$

and high-frequency modes

$$y(t) = (u_{N+1}(t), u_{N+2}(t), \cdots, u_{N+M}(t)) \in \mathbb{R}^M.$$

In practical scenario, we could have millions of high-frequency modes, $M \gg N$. From the stochastic mode reduction procedure (see [18, 22]), we know that the triads between x_i, y_p, y_q ($p \neq q$) produce additional damping, forcing and uncorrelated additive noise which are not essential to the tail behavior. For simplicity and to concentrate on the phenomena involving the interactions of multiplicative dyads and triads, we will neglect the triads between x_i, y_p, y_q ($p \neq q$) in discussion. Then, as in [18, 22], we approximate the dyad interaction between high-frequency modes by using $-\frac{\gamma_p}{\varepsilon} y_p + \frac{\sigma_p}{\sqrt{\varepsilon}} \dot{W}_p$ so that (2.1) becomes

$$\begin{aligned} \frac{\mathrm{d}x_i}{\mathrm{d}t} &= F_i + \sum_j (L_{ij}x_j + I_{ij}x_ix_j - I_{ji}x_jx_j) \\ &\quad + \sum_p \Big(L_{ip}y_p + I^M_{ip}x_iy_p + I^A_{ip}y_py_p + \sum_{j\neq i} B_{ij,p}x_jy_p \Big), \\ \frac{\mathrm{d}y_p}{\mathrm{d}t} &= F_p + \sum_j \Big(L_{pj}x_j + I^M_{pj}x_jx_j + I^A_{pj}y_px_j + \sum_{k<j} B_{p,kj}x_kx_j \Big) - \frac{\gamma_p}{\varepsilon}y_p + \frac{\sigma_p}{\sqrt{\varepsilon}}\dot{W}_p. \end{aligned} \tag{2.2}$$

Energy conservation requires that dyad interactions satisfy $I_{ij}^M = -I_{ji}^M$, $I_{ij}^A = -I_{ji}^A$ and triad interactions satisfy $B_{ij,p}+B_{ji,p}+B_{p,ij}=0$. The working assumption of the stochastic modelling contains $\gamma_p > 0,\ \forall\, p$, so that the high frequency modes have physical regimes.

The functional form of general stochastic reduced models for climate with dyad interactions and triad interactions is systematically derived for (2.2) by the normal form stochastic mode reduction method (see [18]). The dyad interaction terms, I_{ip}^A, produce additional damping, forcing and uncorrelated additive noise; meanwhile, dyad interaction terms, I_{ip}^M, and triad interaction terms, $B_{ij,p}$, produce cubic nonlinearity, correlated additive and multiplicative noises (see [18, 22, 17]). The reduced model has the Itô form

$$\begin{aligned}\mathrm{d}x_i &= \widetilde{F}_i\mathrm{d}t + \widetilde{A}_i x\mathrm{d}t + \widetilde{B}_i(x,x)\mathrm{d}t\\&\quad+\sum_p \frac{1}{\gamma_p}\Big(I_{ip}^M x_i + \sum_{j\neq i} B_{ij,p}x_j\Big)\Big(\sum_j I_{pj}^M x_j x_j + \sum_{k<j} B_{p,kj}x_k x_j\Big)\mathrm{d}t\\&\quad+\sum_p\Big(L_{ip} + I_{ip}^M x_i + \sum_{j\neq i} B_{ij,p}x_j\Big)\frac{\sigma_p}{\gamma_p}\mathrm{d}W_p\\&\quad+\sigma_{A_i}\mathrm{d}W_{A_i}\end{aligned}\tag{2.3}$$

for $i=1,2,\cdots,N,\ p=1,2,\cdots,N_p$, where the dyad interactions satisfy $I_{ij}^M = -I_{ji}^M$ and triad interactions satisfy $B_{ij,p}+B_{ji,p}+B_{p,ij}=0$. It is clear that the reduced system contains cubic nonlinearity, CAM noises, and additive noises and the existence of CAM noise is due to the presence of the cubic terms. This is the normal form derived in [18]. We can also write it in a compact form

$$\mathrm{d}x = f\mathrm{d}t + \Sigma\mathrm{d}W,\tag{2.4}$$

where

$$\begin{aligned}f_i(x) &= \widetilde{F}_i + \widetilde{A}_i x + \widetilde{B}_i(x,x) + \sum_p \frac{1}{\gamma_p}\Big(I_{ip}^M x_i + \sum_{j\neq i} B_{ij,p}x_j\Big)\Big(\sum_j I_{pj}^M x_j x_j + \sum_{k<j} B_{p,kj}x_k x_j\Big),\\ g &= (g_{ij}) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}},\\ g_{ii}(x) &= \sum_p \frac{\sigma_p^2}{2\gamma_p^2}\Big(L_{ip} + I_{ip}^M x_i + \sum_{k\neq i} B_{ik,p}x_k\Big)^2 + \frac{\sigma_{A_i}^2}{2},\\ g_{ij}(x) &= \sum_p \frac{\sigma_p^2}{2\gamma_p^2}\Big(L_{ip} + I_{ip}^M x_i + \sum_{k\neq i} B_{ik,p}x_k\Big)\Big(L_{jp} + I_{jp}^M x_j + \sum_{k\neq j} B_{jk,p}x_k\Big),\quad i\neq j,\\ \mathrm{d}W &= (\mathrm{d}W_1,\mathrm{d}W_2,\cdots,\mathrm{d}W_{N_p},\mathrm{d}W_{A_1},\mathrm{d}W_{A_2},\cdots,\mathrm{d}W_{A_N})^{\mathrm{T}}.\end{aligned}$$

3 Main Results and Discussion

First, we give some intuition. The local change in energy is given through Itô's formula by

$$\mathrm{d}\frac{|x|^2}{2} = (x^{\mathrm{T}}\cdot f(x) + \mathrm{Tr}(g(x)))\mathrm{d}t + x^{\mathrm{T}}\Sigma(x)\mathrm{d}W.\tag{3.1}$$

The effective energy dissipation through the cubic term is $-x^{\mathrm{T}} \cdot f$ which has leading quartic part

$$Q_1(x) = \sum_p \frac{1}{\gamma_p}\Big(\sum_j I_{pj}^M x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j\Big)^2. \tag{3.2}$$

The random fluctuations in energy are given by $x^{\mathrm{T}}\Sigma(x)\mathrm{d}W$, so the quartic part of the variance of the random energy fluctuations is

$$Q_3(x) = \sum_p \frac{\sigma_p^2}{2\gamma_p^2}\Big(\sum_j I_{pj}^M x_j^2 + \sum_{k<j} B_{p,kj} x_k x_j\Big)^2. \tag{3.3}$$

For the upper bound, we need

$$\min_{|\vec{\omega}|=1} Q_1(\vec{\omega}) > 0, \tag{3.4}$$

i.e., strict energy dissipation from the cubic terms. For the lower bound, we need (3.4) and the variance of random fluctuation strictly positive

$$\min_{|\vec{\omega}|=1} Q_3(\vec{\omega}) > 0. \tag{3.5}$$

The following theorem discusses the existence of the density functions.

Theorem 3.1 (Existence of the Probability Density Function) *Consider the normal form climate model* (2.4) *with drift term $f(x)$ and diffusion coefficients $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. Given strict energy dissipation through the cubic terms as defined in* (3.4)*, the density function of the stationary statistical solution to* (2.4) *exists in a generalized function space.*

Our main results Theorems 3.2–3.3 find a Gaussian upper bound and a Gaussian lower bound under suitable assumptions respectively, by constructing Gaussian comparison functions.

Theorem 3.2 (Gaussian Upper Bound) *Consider the normal form reduced climate model* (2.4) *with drift term $f(x)$ and diffusion coefficients $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. Given the assumptions below, if $p(x)$ is a density function of the stationary statistical solution to* (2.4)*, then, it has at most Gaussian decay, i.e., there exists a positive real number r and a Gaussian measure $p^G(x)$ satisfying*

$$p^G(x) \geqslant p(x), \quad \forall\, |x| \geqslant r,$$

where

$$p^G(x) = \frac{1}{\sqrt{2\pi\sigma_G^2}} \mathrm{e}^{-\frac{|x|^2}{2\sigma_G^2}}, \quad x = (x_1, x_2, \cdots, x_N)^{\mathrm{T}} \in \mathbb{R}^N.$$

These assumptions are

(1) *Strict energy dissipation through the cubic terms as defined in* (3.4);

(2) *Weak polynomial decaying smooth stationary statistical solution: $p(x) \in C^2(\mathbb{R}^N)$, such that there exists an α with*

$$\alpha > \frac{\max\limits_{|\vec{\omega}|=1} Q_2(\vec{\omega})}{\min\limits_{|\vec{\omega}|=1} Q_1(\vec{\omega})}, \qquad \lim_{|x|\to+\infty} |x|^\alpha p(x) = 0,$$

where $Q_1(x)$ is defined in (3.2)*, and $Q_2(x)$ is the quartic part of $-|x|^2\nabla\cdot f(x)$, i.e.,*

$$\begin{aligned}Q_2(x) = \Big(&-\sum_i\sum_p I^M_{ip}\frac{1}{\gamma_p}\Big(\sum_j I^M_{pj}x_jx_j+\sum_{k<j}B_{p,kj}x_kx_j\Big)\\&-\sum_i\sum_p\frac{1}{\gamma_p}\Big(2I^M_{pi}x_i+\sum_{j\neq i}B_{p,ij}x_j\Big)\Big(I^M_{ip}x_i+\sum_{j\neq i}B_{ij,p}x_j\Big)\Big)|x|^2.\end{aligned}\tag{3.6}$$

Theorem 3.3 (Gaussian Lower Bound) *Consider the normal form reduced climate model* (2.4) *with drift term $f(x)$ and diffusion coefficients $g(x)=(g_{ij}(x))=\frac{1}{2}\Sigma\Sigma^{\rm T}$. Given the assumptions below, if $p(x)$ is a density function of the stationary statistical solution to* (2.4)*, then, it has at least Gaussian decay, i.e., there exist a positive real number r and a Gaussian measure $p^L(x)$ satisfying*

$$p(x)\geqslant p^L(x),\quad \forall\,|x|\geqslant r,$$

where

$$p^L(x)=\frac{1}{\sqrt{2\pi\sigma_L^2}}{\rm e}^{-\frac{|x|^2}{2\sigma_L^2}},\quad x=(x_1,x_2,\cdots,x_N)^{\rm T}\in\mathbb{R}^N.$$

These assumptions are

(1) *Strict energy dissipation through the cubic terms as defined in* (3.4)*;*

(2) *Strictly positive variance of the multiplicative energy fluctuation given in* (3.5)*;*

(3) *Weak polynomial decaying smooth stationary statistical solution: $p(x)\in C^2(\mathbb{R}^N)$, such that there exists an α with*

$$\alpha>\frac{\max\limits_{|\vec\omega|=1}Q_2(\vec\omega)}{\min\limits_{|\vec\omega|=1}Q_1(\vec\omega)},\quad \lim_{|x|\to+\infty}|x|^\alpha p(x)=0,$$

where $Q_1(x)$ is defined in (3.2)*, and $Q_2(x)$ is defined in* (3.6)*;*

(4) *$p(x)$ is positive everywhere*

$$\inf_{|x|\leq r}p(x)>0,\quad \forall\,r>0.$$

These conditions are readily satisfied in general. For example, both conditions (1)–(2) are satisfied for the 2-dimensional stochastic climate model (3.9) below in Sections 3.1.1 and 3.3 provided $I_1I_2\neq 0,\ \sigma_1\sigma_2\neq 0$, i.e., nontrivial dyad interactions.

If we further assume that the diffusion part is uniformly non-degenerate, we can classify the global decaying rate for the density function of the stationary statistical solution rigorously.

Corollary 3.1 *Consider the normal form reduced climate model* (2.4) *with drift term $f(x)$ and diffusion coefficients $g(x)=(g_{ij}(x))=\frac{1}{2}\Sigma\Sigma^{\rm T}$. There exist a unique density function of the stationary statistical solution to* (2.4)*, $p(x)\in C^\infty$, and $p(x)$ has at most Gaussian decay, i.e., there exist a positive real number r and Gaussian measures $p^G(x)$ satisfying*

$$p^G(x)\geqslant p(x),\quad \forall\,|x|\geqslant r,$$

where

$$p^G(x)=\frac{1}{\sqrt{2\pi\sigma_G^2}}{\rm e}^{-\frac{|x|^2}{2\sigma_G^2}},\quad x=(x_1,x_2,\cdots,x_N)^{\rm T}\in\mathbb{R}^N,\tag{3.7}$$

given the following assumptions:

(1) *Strict energy dissipation through the cubic terms as defined in* (3.4);

(2) *Uniformly non-degenerate diffusion: The diffusion coefficient matrix* $g(x)$ *is positive definite in* $\mathbb{R}^N$.

If we further require that the following condition hold:

(3) *Strictly positive variance of the multiplicative energy fluctuation given in* (3.5).

Then, the global decaying rate of $p(x)$ *is exactly Gaussian, i.e., besides* (3.7), *there exist a positive real number* r *and Gaussian measures* $p^L(x)$ *satisfying*

$$p^L(x) \leqslant p(x), \quad \forall\, |x| \geqslant r,$$

where

$$p^L(x) = \frac{1}{\sqrt{2\pi\sigma_L^2}} \mathrm{e}^{-\frac{|x|^2}{2\sigma_L^2}}, \quad x = (x_1, x_2, \cdots, x_N)^{\mathrm{T}} \in \mathbb{R}^N. \tag{3.8}$$

The explicit examples with the scalar normal form in Section 3.2 illustrate the necessity of the structured assumptions in Theorems 3.1–3.3 and the corollary for the smoothness and Gaussian upper and lower bounds for the invariant measure as discussed in Section 3.3.

3.1 Energy dissipation constraints through the cubic term

In practice the original system has well-behaved PDFs for the low-frequency patterns with rapidly decaying tails (see [1]). Thus, the nonlinear operator in (2.3) should induce an effective damping following [18]. This can be shown for the N-dimensional case by multiplying (2.3) by x_i, neglecting the noise terms and all other terms besides the cubic terms, and summing over i leading to the energy identity

$$\frac{1}{2}\frac{\mathrm{d}E}{\mathrm{d}t} = -Q_1(x),$$

where $Q_1(x) = \sum\limits_p \frac{1}{\gamma_p}\Big(\sum\limits_j I^M_{pj} x_j x_j + \sum\limits_{k<j} B_{p,kj} x_k x_j\Big)^2$ as defined in (3.2) and $E = \sum\limits_i x_i^2$.

Stability of the original system now requires that $\frac{1}{2}\frac{\mathrm{d}E}{\mathrm{d}t} < 0$. This is fulfilled if the homogeneous polynomial $Q_1(x)$ is positive definite which is given by our first assumption as defined in (3.4).

3.1.1 Energy dissipation constraints on a 2-dimensional stochastic climate model

There is wide interest in studying the normal form for a 2-dimensional stochastic climate model (see [18]) following [20, 17, 10] which can be systematically derived from a 4-mode climate model with triad interaction between x_1, x_2, y_2 and dyad interaction between x_1, y_2 and x_2, y_1. This toy model has statistics that are very close to being Gaussian and a strong nonlinear signature in the form of a double swirl in the mean phase space tendencies of its low-frequency variables, much like recently identified signatures of nonlinear planetary wave dynamics in prototype and comprehensive atmospheric general circulation models (GCMs) (see [10]).

The 2-dimensional normal form stochastic climate model has cubic damping terms and CAM noises

$$\begin{aligned}
\mathrm{d}x_1 &= (\overline{B}_1(x,x) + \overline{A}_1 x + \overline{F}_1)\mathrm{d}t \\
&\quad + \frac{1}{\gamma_1} b_{123}(b_{312}x_1x_2^2 - I_2x_2^3)\mathrm{d}t - \frac{1}{\gamma_2} I_1^2 x_1^3 \mathrm{d}t \\
&\quad + \frac{\sigma_1}{\gamma_1}(L_{13} + b_{123}x_2)\mathrm{d}W_1 + \frac{\sigma_2}{\gamma_2} I_1 x_1 \mathrm{d}W_2, \\
\mathrm{d}x_2 &= (\overline{B}_2(x,x) + \overline{A}_2 x + \overline{F}_2)\mathrm{d}t \\
&\quad + \frac{1}{\gamma_1}(b_{312}x_1x_2 - I_2x_2^2)(b_{213}x_1 + I_2x_2)\mathrm{d}t \\
&\quad + \frac{\sigma_1}{\gamma_1}(b_{213}x_1 + I_2x_2)\mathrm{d}W_1 + \frac{\sigma_2}{\gamma_2} L_{24}\mathrm{d}W_2.
\end{aligned} \tag{3.9}$$

To mimic the main features of the GCM which has well-behaved PDFs for the low-frequency patterns with rapidly decaying tails, the nonlinear operator in (3.9) should also induce an effective damping. Multiplying (3.9) by x_i, neglecting the noise terms and all other terms besides the cubic terms, and summing over i leads to the energy identity

$$\frac{1}{2}\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{1}{\gamma_2} I_1^2 x_1^4 - \frac{1}{\gamma_1} x_2^2 (b_{312}x_1 - I_2x_2)^2.$$

The energy dissipation requires $\frac{1}{2}\frac{\mathrm{d}E}{\mathrm{d}t} < 0$ which is equivalent to $I_1 I_2 \neq 0$, so that there are nontrivial dyad interactions I_1 and I_2.

3.1.2 The stochastic model with only dyad interaction

The dissipation constraint on the quartic terms in $-x^{\mathrm{T}} \cdot f$ is also consistent with [18, Equation (14)] for the case with only dyads. If we take off the triad interaction in (2.4), the energy dissipation requirement becomes that $\sum_p \frac{1}{\gamma_p}\Big(\sum_j I_{pj}^M x_j x_j\Big)^2$ is positive definite in $\mathbb{R}^N$. That is

$$\sum_p \frac{1}{\gamma_p}\Big(\sum_j I_{pj}^M x_j x_j\Big)^2 = -\sum_i \Big(\sum_{j\neq i} \widetilde{A}_{ij} x_j^2 - \widetilde{I}_{ii} x_i^2\Big) x_i^2 = (x^2)^{\mathrm{T}}(\widetilde{I} - \widetilde{A})(x^2),$$

where

$$\widetilde{A}_{ij} = \sum_p \frac{I_{ip}^M I_{pj}^M}{\gamma_p}, \quad j \neq i, \quad \widetilde{A}_{ii} = 0; \quad \widetilde{I}_{ii} = \sum_p \frac{(I_{ip}^M)^2}{\gamma_p}; \quad \widetilde{I}_{ij} = 0, \quad j \neq i.$$

So, the energy dissipation requirement becomes that $\widetilde{I} - \widetilde{A}$ is positive definite on the positive cone in $\mathbb{R}^n$.

3.2 Why do smoothness and Gaussian decay happen? Examples through the scalar normal form climate model

What conditions make the Gaussian decay? Roughly, by looking at the Fokker-Planck equation of the scalar normal form climate model, we have the asymptotic behavior of the statistical solution $p(x)$

$$\ln p(x) \backsim \int \frac{\text{cubic dissipation}}{\text{quadratic or constant diffusion}}, \quad \text{as } |x| \text{ is large},$$

which gives us at most a Gaussian tail.

Let us discuss the scalar case from [18] in detail to get more intuition. Consider the reduced stochastic model for scalar low frequency climate variable

$$\mathrm{d}x = [F + ax + bx^2 - cx^3]\mathrm{d}t + \sum_p \frac{\sigma_p}{\gamma_p}(L_{1p} - I^M_{1p}x)\mathrm{d}W_p + \sigma \mathrm{d}W_A, \tag{3.10}$$

where $c = \sum\limits_p \frac{(I^M_{1p})^2}{\gamma_p}$. As shown in [18], this scalar case is exactly solvable. Since we can get the explicit decaying stationary statistical solution for this scalar case, the weak decaying requirement is automatically satisfied. The strict energy dissipation assumption through the cubic term, i.e., $I^M_{1p} \neq 0$ for some p, is equivalent to the cubic nonlinearity condition $c > 0$. The condition, strictly positive variance of the multiplicative energy fluctuation, i.e., $\sigma_p I^M_{1p} \neq 0$ for some p, is equivalent to the presence of the CAM noise $B \neq 0$. And the uniformly non-degenerate diffusion assumption is equivalent to the existence of the additional additive noise obtained from the additive dyads and triads, $\sigma \neq 0$.

The stationary statistical solution to (3.10), $p(x)$, satisfies Fokker Planck equation

$$-\frac{\partial}{\partial x}[(F + ax + bx^2 - cx^3)p(x)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[((Bx - A)^2 + \sigma^2)p(x)] = 0.$$

We could integrate the stationary Fokker Planck equation and get

$$\ln \frac{p(x)}{p(x_0)} = \int_{x_0}^x \frac{2B(Bx - A) + 2(F + ax + bx^2 - cx^3)}{(Bx - A)^2 + \sigma^2}\mathrm{d}x.$$

Let us denote

$$\begin{aligned}
a_1 &= 1 - \frac{-3A^2c + aB^2 + 2AbB + c\sigma^2}{B^4}, \quad b_1 = 2bB^2 - 4cAB, \quad c_1 = cB^2,\\
d &= \frac{d_1}{\sigma} + d_2\sigma, \quad d_1 = \frac{2A^2bB - 2A^3c + 2AaB^2 + 2B^3F}{B^4}, \quad d_2 = \frac{6cA - 2bB}{B^4},\\
A &= \frac{\sum_p \frac{\sigma_p}{\gamma_p} L_{1p} I^M_{1p}}{\Big(\sum\limits_p \Big(\frac{\sigma_p}{\gamma_p} I^M_{1p}\Big)^2\Big)^{\frac{1}{2}}}, \quad B = \Big(\sum_p \Big(\frac{\sigma_p}{\gamma_p} I^M_{1p}\Big)^2\Big)^{\frac{1}{2}},\\
\gamma &= -2a_{1_{\sigma=0}} = -2\Big(1 - \frac{-3A^2c + aB^2 + 2AbB}{B^4}\Big).
\end{aligned}$$

Assuming nontrivial cubic nonlinearity $c > 0$, we get that all the cases have at most Gaussian decay:

Case 1 CAM term exists $B \neq 0$, and the additive noise exists $\sigma \neq 0$. Then, $p(x)$ possibly has some fat tail decay in a finite range, but has Gaussian decay outside of the finite range,

$$p(x) = \frac{N_1}{((Bx - A)^2 + \sigma^2)^{a_1}} \mathrm{e}^{d \arctan(\frac{Bx - A}{\sigma})} \cdot \mathrm{e}^{\frac{-c_1x^2 + b_1x}{B^4}}. \tag{3.11}$$

Case 2 No CAM term $B = 0$. Then, $p(x)$ has even faster decay than Gaussian,

$$p(x) = N_2 \mathrm{e}^{\frac{2}{A^2 + \sigma^2}(Fx + \frac{ax^2}{2} + \frac{bx^3}{3} - \frac{cx^4}{4})}.$$

Case 3 CAM term exists $B \neq 0$, but there is no additional additive noise $\sigma = 0$. Without the additive noise, $p(x)$ exists only on half real line, or it has Dirac delta mass. So, this case is not in the physical regime we are interested in here. However, it is necessary to understand these different behaviors for the scalar case in order to consider reasonable hypotheses for the N-dimensional climate model with degenerate diffusions.

Here, the density function has bifurcations:

(1) $d_1 B > 0$, $\forall \gamma$,

$$p(x) = \begin{cases} 0, & x \leqslant \frac{A}{B}, \\ \dfrac{\mathrm{e}^{-\frac{|d_1|}{|B|x - \mathrm{sgn}(B)A} + \frac{-c_1 x^2 + b_1 x}{B^4}} \cdot |Bx - A|^{\gamma}}{\int_{\frac{A}{B}}^{+\infty} \mathrm{e}^{-\frac{|d_1|}{|B|y - \mathrm{sgn}(B)A} + \frac{-c_1 y^2 + b_1 y}{B^4}} \cdot |By - A|^{\gamma} \mathrm{d}y}, & x > \frac{A}{B}; \end{cases}$$

(2) $d_1 B < 0$, $\forall \gamma$,

$$p(x) = \begin{cases} 0, & x \geqslant \frac{A}{B}, \\ \dfrac{\mathrm{e}^{\frac{|d_1|}{|B|x - \mathrm{sgn}(B)A} + \frac{-c_1 x^2 + b_1 x}{B^4}} \cdot |Bx - A|^{\gamma}}{\int_{-\infty}^{\frac{A}{B}} \mathrm{e}^{\frac{|d_1|}{|B|y - \mathrm{sgn}(B)A} + \frac{-c_1 y^2 + b_1 y}{B^4}} \cdot |By - A|^{\gamma} \mathrm{d}y}, & x < \frac{A}{B}; \end{cases}$$

(3) $d_1 = 0$, $\gamma \geqslant 0$,

$$p(x) = \frac{\mathrm{e}^{\frac{-c_1 x^2 + b_1 x}{B^4}} \cdot |Bx - A|^{\gamma}}{\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{-c_1 y^2 + b_1 y}{B^4}} \cdot |Bx - A|^{\gamma} \mathrm{d}y}, \quad x \in (-\infty, +\infty);$$

(4) $d_1 = 0$, $-1 < \gamma < 0$,

$$p(x) = \frac{\mathrm{e}^{\frac{-c_1 x^2 + b_1 x}{B^4}} \cdot |Bx - A|^{\gamma}}{\int_{-\infty}^{+\infty} \mathrm{e}^{\frac{-c_1 y^2 + b_1 y}{B^4}} \cdot |Bx - A|^{\gamma} \mathrm{d}y}, \quad x \neq \frac{A}{B},$$

and there is a local integrable singularity such that

$$p\Big(\frac{A}{B}\Big) = +\infty;$$

(5) $d_1 = 0$, $\gamma \leqslant -1$,

$$p(x) = \delta\Big(x - \frac{A}{B}\Big), \quad -\infty < x < +\infty.$$

Let us check the hypoellipticity of the Fokker Planck operator by Hormander condition in [13] for this scalar case. We notice that the Hormander condition is equivalent to the bifurcation condition $d_1 \neq 0$.

We conclude that energy dissipation requirement through the cubic term is essential to the Gaussian decay and the presence of additive noises impacts the existence and regularity of the stationary statistical solution in the whole phase space.

3.3 Two conditions on the CAM noises: uniformly non-degenerate diffusion assumption and strictly positive variance of the multiplicative energy fluctuation

There are two possible assumptions on the CAM noises. One is the uniformly non-degenerate diffusion assumption which requires that the diffusion coefficient matrix $g(x)$ is positive definite in $\mathbb{R}^N$. The other is strictly positive variance of the multiplicative energy fluctuation as defined in (3.5).

Under the strict energy dissipation assumption through the cubic nonlinearity as defined in (3.4), a sufficient condition for the strictly positive variance of the multiplicative energy fluctuation can be the presence of the nontrivial dyad interaction between high frequency modes, $\sigma_p \neq 0,\ \forall p$, since the following inequality holds

$$Q_3(x) \geq \Big(\min_p \frac{\sigma_p^2}{2\gamma_p}\Big) Q_1(x).$$

Both these conditions on the CAM noises might not be held in a general reduced climate model, but they play important roles in a stochastic climate modeling scenario. For example, in the scalar climate model we discussed in Section 3.2, the uniformly non-degenerate diffusion assumption is equivalent to the existence of the additional additive noise $\sigma \neq 0$ obtained from the additive dyads and triads. It is essential to existence and regularity of the stationary statistical solution in the whole phase space. Strictly positive variance of the multiplicative energy fluctuation, i.e., $\sigma_p I_{1p}^M \neq 0$ for some p, implies the presence of the CAM noise $B \neq 0$ which is needed to find a Gaussian lower bound.

We can also check these conditions for the 2-dimensional stochastic climate model (3.9) which we discussed in Section 3.1.1. We mentioned that this toy model has statistics that are very close to being Gaussian and a strong nonlinear signature in the form of a double swirl in the mean phase space tendencies of its low-frequency variables, much like recently identified signatures of nonlinear planetary wave dynamics in prototype and comprehensive atmospheric general circulation models (GCMs) (see [10]). The CAM noise coefficients are given by

$$\Sigma(x) = \begin{pmatrix} \frac{\sigma_1}{\gamma_1}(L_{13} + b_{123}x_2) & \frac{\sigma_2}{\gamma_2} I_1 x_1 \\ \frac{\sigma_1}{\gamma_1}(b_{213}x_1 + I_2 x_2) & \frac{\sigma_2}{\gamma_2} L_{24} \end{pmatrix}.$$

The diffusion coefficient matrix has the form $g(x) = \frac{1}{2}\Sigma(x)\Sigma(x)^{\mathrm{T}}$ which clearly does not satisfy the uniformly non-degenerate condition. Moreover, on the hyperbolic curve defined by $\det(\Sigma(x)) = 0$, this diffusion matrix is degenerate. Furthermore, the quartic part of the variance of the random energy fluctuation $x^{\mathrm{T}} g(x) x$ is given by

$$Q_3^{2D} = \frac{\sigma_1^2}{2\gamma_2^2} I_1^2 x_1^4 + \frac{\sigma_2^2}{2\gamma_2^2} x_2^2 (b_{312}x_1 - I_2 x_2)^2.$$

The strict positivity of Q_3^{2D} holds as long as there exists nontrivial the dyad interaction, i.e., $I_1 I_2 \neq 0,\ \sigma_1\sigma_2 \neq 0$. Since we have learned in Section 3.1.1 that the strict energy dissipation assumption through the cubic nonlinearity as defined in (3.4) is equivalent to $I_1 I_2 \neq 0$, the positivity of Q_3^{2D} can be obtained by assuming the strict energy dissipation and $\sigma_1\sigma_2 \neq 0$.

Although the 2-dimensional stochastic climate model (3.9) does not have a uniform non-degenerate diffusion, (3.9) has nice and fast decaying invariant probability density functions under various parameter sets as observed through a rich family of numerical simulations. We can also check the hypoellipticity of the Fokker Planck operator by Hormander condition (see [13]). We notice that the Hormander condition holds unless the parameter sets satisfy the following equation

$$\det\begin{pmatrix} \frac{\sigma_1}{\gamma_1}(L_{13}+b_{123}x_2) & \frac{\sigma_2}{\gamma_2}I_1x_1 \\ D_1(x) & D_2(x) \end{pmatrix}=0,$$

where $x=(x_1,x_2)$ is on the hyperbolic curve where the diffusion part is degenerate

$$\det\Sigma(x)=\det\begin{pmatrix} \frac{\sigma_1}{\gamma_1}(L_{13}+b_{123}x_2) & \frac{\sigma_2}{\gamma_2}I_1x_1 \\ \frac{\sigma_1}{\gamma_1}(b_{213}x_1+I_2x_2) & \frac{\sigma_2}{\gamma_2}L_{24} \end{pmatrix}=0$$

and $D_1(x)$, $D_2(x)$ are cubic drift terms of the 2D model

$$\begin{aligned} D_1(x)&=(\overline{B}_1(x,x)+\overline{A}_1x+\overline{F}_1)+\frac{1}{\gamma_1}b_{123}(b_{312}x_1x_2^2-I_2x_2^3)\mathrm{d}t-\frac{1}{\gamma_2}I_1^2x_1^3,\\ D_2(x)&=(\overline{B}_2(x,x)+\overline{A}_2x+\overline{F}_2)+\frac{1}{\gamma_1}(b_{312}x_1x_2-I_2x_2^2)(b_{213}x_1+I_2x_2). \end{aligned}$$

In this paper, we will make reasonable assumptions for the degenerate diffusion case that the density function is C^2 and has power law decay.

4 Proof of the Theorems

The normal form reduced stochastic model for N-climate variables (2.3) is rewritten here for emphasis

$$\begin{aligned} \mathrm{d}x_i&=\widetilde{F}_i\mathrm{d}t+\widetilde{A}_ix\mathrm{d}t+\widetilde{B}_i(x,x)\mathrm{d}t\\ &\quad+\sum_p\frac{1}{\gamma_p}\Big(I_{ip}^Mx_i+\sum_{j\neq i}B_{ij,p}x_j\Big)\Big(\sum_j I_{pj}^Mx_jx_j+\sum_{k<j}B_{p,kj}x_kx_j\Big)\mathrm{d}t\\ &\quad+\sum_p\Big(L_{ip}+I_{ip}^Mx_i+\sum_{j\neq i}B_{ij,p}x_j\Big)\frac{\sigma_p}{\gamma_p}\mathrm{d}W_p+\sigma_{A_i}\mathrm{d}W_{A_i} \end{aligned}$$

for $i=1,2,\cdots,N,\ p=1,2,\cdots,N_p$. The compact form (2.4) is rewritten here as well

$$\mathrm{d}x=f\mathrm{d}t+\Sigma\mathrm{d}W,$$

where

$$\begin{aligned} f_i(x)&=\widetilde{F}_i+\widetilde{A}_ix+\widetilde{B}_i(x,x)\\ &\quad+\sum_p\frac{1}{\gamma_p}\Big(I_{ip}^Mx_i+\sum_{j\neq i}B_{ij,p}x_j\Big)\Big(\sum_j I_{pj}^Mx_jx_j+\sum_{k<j}B_{p,kj}x_kx_j\Big),\\ g&=(g_{ij})=\frac{1}{2}\Sigma\Sigma^{\mathrm{T}},\\ g_{ii}(x)&=\sum_p\frac{\sigma_p^2}{2\gamma_p^2}\Big(L_{ip}+I_{ip}^Mx_i+\sum_{k\neq i}B_{ik,p}x_k\Big)^2+\frac{\sigma_{A_i}^2}{2},\\ g_{ij}(x)&=\sum_p\frac{\sigma_p^2}{2\gamma_p^2}\Big(L_{ip}+I_{ip}^Mx_i+\sum_{k\neq i}B_{ik,p}x_k\Big)\Big(L_{jp}+I_{jp}^Mx_j+\sum_{k\neq j}B_{jk,p}x_k\Big),\quad i\neq j,\\ \mathrm{d}W&=\Big(\mathrm{d}W_1,\mathrm{d}W_2,\cdots,\mathrm{d}W_{N_p},\mathrm{d}W_{A_1},\mathrm{d}W_{A_2},\cdots,\mathrm{d}W_{A_N}\Big)^{\mathrm{T}}. \end{aligned}$$

The drift term satisfies the strict energy dissipation assumption as defined in (3.4) that the leading order part of $-x^{\mathrm{T}}f$, $Q_1(x) = \sum_p \frac{1}{\gamma_p}\Big(\sum_j I^M_{pj} x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j\Big)^2$, is positive definite in $\mathbb{R}^N$.

We are going to discuss its moment estimations of the pathwise solutions, stationary statistical solutions (invariant measure), the asymptotic behavior of the probability density function of the invariant measure, and finally conclude the theorems respectively.

4.1 Moment estimations

The existence and uniqueness of pathwise solutions to our normal form reduced climate model are known by Krylov [15]. Krylov considered a larger class of equations which includes our reduced climate model. By introducing the notion of Euler solvability for stochastic differential problems, Krylov proved some useful estimates for the p-th moments of the solutions. Cerrai applied Krylov's idea to SDEs with polynomial drifts in [6]. Theorem 1.2.5 in [6] gave us a moment estimation with time dependent bound.

Proposition 4.1 *Consider the normal form reduced climate model* (2.4) *with drift term* $f(x)$ *and diffusion coefficients* $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. *Given the strict energy dissipation through the cubic terms as defined in* (3.4), (2.4) *has a unique* N*-dimensional pathwise solution* $x(t)$, *having continuous trajectories and satisfying*

$$E \sup_{0\le s\le t} |x(s)|^n \le a_n(t)(|x(0)|^n + 1), \quad \forall\, n \ge 1, \tag{4.1}$$

where $a_n(t)$ *is a suitable increasing function.*

Here, based on Proposition 4.1, we find a uniform bound for all the moments wrt time evolution.

Proposition 4.2 *Consider the normal form reduced climate model* (2.4) *with drift term* $f(x)$ *and diffusion coefficients* $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. *Given the strict energy dissipation through the cubic terms as defined in* (3.4), *the pathwise solution of* (2.4), $x(t)$, *has uniformly bounded* n*-th moment in the following sense:*

$$E|x(t)|^n \le C(n, x(0)), \quad n \ge 0, \tag{4.2}$$

where the bound $C(n, x(0))$ *depends on the initial position* $x(0)$ *and the order of the moment* n, *but does not depend on time* t.

Proof $x(t)$ satisfies SDE model $\mathrm{d}x = f\mathrm{d}t + \Sigma\mathrm{d}W$. The local energy change is given by (3.1) which is rewritten below for emphasis

$$\mathrm{d}\frac{|x|^2}{2} = (x^{\mathrm{T}}f(x) + \mathrm{Tr}(g(x)))\mathrm{d}t + x^{\mathrm{T}}\Sigma(x)\mathrm{d}W.$$

Applying Itô's lemma one more time, we get

$$\begin{aligned}
\mathrm{d}|x|^{2m} = {}& 2m|x|^{2(m-1)}x^{\mathrm{T}}f(x)\mathrm{d}t \\
&+ [2m|x|^{2(m-1)}\mathrm{Tr}(g(x)) + 4m(m-1)|x|^{2(m-2)}(x^{\mathrm{T}}g(x)x)]\mathrm{d}t \\
&+ 2m|x|^{2(m-1)}x^{\mathrm{T}}\Sigma(x)\mathrm{d}W, \quad \forall\, m > 2.
\end{aligned}$$

For $r \geq 0$, we define the stopping time τ_r by setting

$$\tau_r = \inf\{t \geq 0,\ |x(t)| \geq r\}.$$

Then, we have the pathwise solution

$$\begin{aligned}
&|x(t\wedge\tau_r)|^{2m} \\
&= |x(0)|^{2m} + 2m\int_0^{t\wedge\tau_r} |x(s)|^{2(m-1)} x(s)^{\mathrm{T}} f(x(s))\mathrm{d}s \\
&\quad + \int_0^{t\wedge\tau_r} [2m|x(s)|^{2(m-1)}\mathrm{Tr}(g(x(s))) + 4m(m-1)|x(s)|^{2(m-2)}(x(s)^{\mathrm{T}} g(x(s))x(s))]\mathrm{d}s \\
&\quad + \int_0^{t\wedge\tau_r} 2m|x(s)|^{2(m-1)} x(s)^{\mathrm{T}}\Sigma(x(s))\mathrm{d}W(s), \quad \forall\, m > 2.
\end{aligned}$$

The $2m$-th moment satisfies

$$\begin{aligned}
E|x(t\wedge\tau_r)|^{2m} &= |x(0)|^{2m} + 2m\int_0^t E\{\chi_{\{\tau_r\geq s\}}|x(s)|^{2(m-1)} x(s)^{\mathrm{T}} f(x(s))\}\mathrm{d}s \\
&\quad + \int_0^t E\{\chi_{\{\tau_r\geq s\}} 2m|x(s)|^{2(m-1)}\mathrm{Tr}(g(x(s)))\}\mathrm{d}s \\
&\quad + 4m(m-1)\int_0^t E\{\chi_{\{\tau_r\geq s\}}|x(s)|^{2(m-2)}(x(s)^{\mathrm{T}} g(x(s))x(s))\}\mathrm{d}s. \qquad (4.3)
\end{aligned}$$

The strict energy dissipation assumption implies that there exists $\lambda > 0$ such that the leading order part of $-x^{\mathrm{T}} f$, $Q_1(x)$, satisfies $Q_1(x) \geq \lambda|x|^4$.

Applying the strict energy dissipation assumption and Hölder's inequality, we can estimate the derivative of the $2m$-th moment

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}E|x(t\wedge\tau_r)|^{2m} &= 2mE\{\chi_{\{\tau_r\geq t\}}|x(t)|^{2(m-1)} x(t)^{\mathrm{T}} f(x(t))\} \\
&\quad + E\{\chi_{\{\tau_r\geq t\}} 2m|x(t)|^{2(m-1)}\mathrm{Tr}(g(x(t)))\} \\
&\quad + 4m(m-1)E\{\chi_{\{\tau_r\geq t\}}|x(s)|^{2(m-2)}(x(t)^{\mathrm{T}} g(x(t))x(t))\} \\
&\leq E\{\chi_{\{\tau_r\geq t\}}(-\widetilde{\lambda}|x(t)|^{2(m+1)} + c_1(m))\} \quad (0 < \widetilde{\lambda} < \lambda) \\
&\leq E(-\widetilde{\lambda}|x(t\wedge\tau_r)|^{2(m+1)} + \widetilde{\lambda} r^{2(m+1)}\chi_{\{\tau_r<t\}}) + c_1(m) \\
&= -\widetilde{\lambda}E|x(t\wedge\tau_r)|^{2(m+1)} + \widetilde{\lambda} r^{2(m+1)}\mathcal{P}(\tau_r < t) + c_1(m) \\
&\leq -\widetilde{\lambda}(E|x(t\wedge\tau_r)|^{2m})^{1+\frac{1}{m}} + \widetilde{\lambda} r^{2(m+1)}\mathcal{P}(\tau_r < t) + c_1(m) \\
&\leq -\widetilde{\lambda}E|x(t\wedge\tau_r)|^{2m} + \widetilde{\lambda} r^{2(m+1)}\mathcal{P}(\tau_r < t) + c_2(m), \qquad (4.4)
\end{aligned}$$

where the last step is due to the inequality $y - y^{1+a} \leq \frac{a}{(1+a)^{1+\frac{1}{a}}},\ \forall a > 0,\ y \in R$.

This simply implies

$$\begin{aligned}
E|x(t\wedge\tau_r)|^{2m} &\leq \mathrm{e}^{-\widetilde{\lambda}t}|x(0)|^{2m} + \int_0^t \mathrm{e}^{-\widetilde{\lambda}(t-s)}(\widetilde{\lambda} r^{2(m+1)}\mathcal{P}(\tau_r < s) + c_2(m))\mathrm{d}s \\
&\leq \mathrm{e}^{-\widetilde{\lambda}t}|x(0)|^{2m} + \widetilde{\lambda}\int_0^t \mathrm{e}^{-\widetilde{\lambda}(t-s)} r^{2(m+1)}\mathcal{P}(\tau_r < s)\mathrm{d}s + \frac{1}{\widetilde{\lambda}}c_2(m). \qquad (4.5)
\end{aligned}$$

Thanks to the Chebyschev inequality and (4.1), we get

$$
\begin{aligned}
r^{2(m+1)}\mathcal{P}(\tau_r < s) &\leq r^{2(m+1)}\mathcal{P}\Big(\sup_{0\leq\varrho\leq s}|x(\varrho)|^{2(m+2)} \geq r^{2(m+2)}\Big) \\
&\leq \frac{1}{r^2}E\sup_{0\leq\varrho\leq s}|x(\varrho)|^{2(m+2)} \leq \frac{1}{r^2}a_{2m}(s)(|x(0)|^{2m}+1). \qquad (4.6)
\end{aligned}
$$

Let r goes to $+\infty$ in (4.5). Due to Fatou's lemma, we obtain

$$E|x(t)|^{2m} \leq \mathrm{e}^{-\widetilde{\lambda}t}|x(0)|^{2m} + \frac{1}{\widetilde{\lambda}}c_2(m),$$

which implies $E|x(t)|^{2m} \leq C(m, x(0)),\ \forall\, m > 2$ for some constant $C(m, x(0))$ that does not depend on time t.

Finally, we use Hölder's inequality to conclude that $E|x(t)|^n \leq C(n, x(0))$ holds for all $n \geq 0$.

4.2 Existence and uniqueness of the invariant measure

Existence of the invariant measure of our reduced climate model follows from the uniform boundness of the second moment of the pathwise solution with the standard tightness argument (see [6]).

Proposition 4.3 (Existence) *Consider the normal form reduced climate model* (2.4) *with drift term $f(x)$ and diffusion coefficients $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. Given the strict energy dissipation assumption through the cubic terms as defined in* (3.4)*,* (2.4) *has an invariant measure.*

And the uniqueness follows from strong Feller property and irreducibility of the Markov transition semigroup which require nondegenerate CAM noises (see [6]).

Proposition 4.4 (Uniqueness) *Consider the normal form reduced climate model* (2.4) *with drift term $f(x)$ and diffusion coefficients $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. Given the strict energy dissipation assumption through the cubic terms as defined in* (3.4) *and uniformly non-degenerate diffusion that g is uniformly positive definite in $\mathbb{R}^N$, the invariant measure of* (2.4) *is unique, ergodic and strongly mixing.*

We present the main idea of the proof below for completeness (see [6]).

Consider stochastic problem

$$\mathrm{d}\xi(t) = f(\xi(t))\mathrm{d}t + \Sigma(\xi(t))\mathrm{d}W(t), \quad \xi(0) = x \in \mathbb{R}^N, \qquad (4.7)$$

where f and g satisfy sufficient assumptions ensuring the existence and uniqueness of the solution $\xi(t, x)$ and the continuity of the trajectories. For any Borel bounded function $\varphi \in B_b(\mathbb{R}^N)$, the Markov transition semigroup $\mathcal{P}_t$ corresponding to problem (4.7) is defined by

$$\mathcal{P}_t\varphi(x) = E\varphi(\xi(t, x)).$$

We say that a semigroup $\mathcal{P}_t$ acting on $B_b(\mathbb{R}^N)$ enjoys the Feller property if

$$\varphi \in C_b(\mathbb{R}^N) \Rightarrow \mathcal{P}_t\varphi \in C_b(\mathbb{R}^N), \quad t \geq 0.$$

Instead, we say that $\mathcal{P}_t$ enjoys the strong Feller property if

$$\varphi \in B_b(\mathbb{R}^N) \Rightarrow \mathcal{P}_t\varphi \in C_b(\mathbb{R}^N), \quad t > 0.$$

A measure μ is invariant for the semigroup $\mathcal{P}_t$, if for any $t \geq 0$, $\mathcal{P}_t^*\mu = \mu$ holds, or equivalently,

$$\int_{\mathbb{R}^N} \mathcal{P}_t\varphi(y)\mu(\mathrm{d}y) = \int_{\mathbb{R}^N} \varphi(y)\mu(\mathrm{d}y).$$

Markovian transition probability, the law of $\xi(t,x)$, is defined for $x \in \mathbb{R}^N$ and $t \geq 0$ by

$$\mathcal{P}_t(x, A) = \mathcal{P}_t\chi_A(x),$$

where A is any Borel set in $\mathbb{R}^N$. Moreover, we say that $\mathcal{P}_t$ is irreducible if for every non empty open set $A \subseteq \mathbb{R}^N$ and $x \in \mathbb{R}^N$, we have $\mathcal{P}_t(x, A) > 0,\ \forall\, t$.

We recall that a set of probability measures Λ is said to be tight if for any $\varepsilon > 0$, there exists a compact set $K_\varepsilon \in \mathbb{R}^N$, such that for any $\mu \in \Lambda$,

$$\mu(K_\varepsilon) \geq 1 - \varepsilon.$$

The following theorem gives a property of tight subsets.

Theorem 4.1 (Prokhorov) *If $\{\mu_n\}$ is a tight sequence in $P(\mathbb{R}^N)$ the collection of probability measures on $\mathbb{R}^N$, then there exists a subsequence $\{\mu_{n_i}\}$ and probability measure μ in $P(\mathbb{R}^N)$, such that $\{\mu_{n_i}\}$ converges weakly to μ.*

For any $\nu \in P(\mathbb{R}^N)$, we define

$$\mathcal{R}_t^*\nu = \frac{1}{t}\int_0^t \mathcal{P}_s^*\nu \mathrm{d}s,$$

which means that for any Borel bounded function φ,

$$\int_{\mathbb{R}^N} \varphi(y)\mathcal{R}_t^*\nu(\mathrm{d}y) = \frac{1}{t}\int_0^t \int_{\mathbb{R}^N} \mathcal{P}_s\varphi(y)\nu(\mathrm{d}y)\mathrm{d}s.$$

The following theorem describes a method of constructing an invariant measure by using $\mathcal{R}_t^*\nu$.

Theorem 4.2 (Krylov-Bogoliubov) *Assume that $\mathcal{P}_t$ is a Feller semigroup. If for some $\nu \in P(\mathbb{R}^N)$ and some sequence $\{t_n\}$ increasing to $+\infty$, we have that $\mathcal{R}_{t_n}^*\nu$ converges weakly to some $\mu \in P(\mathbb{R}^N)$, as $n \to +\infty$. Then μ is an invariant measure for $\mathcal{P}_t$.*

Furthermore, Khas'minskii and Doob give us important tools to study the uniqueness and asymptotic behavior of the invariant measure.

Theorem 4.3 (see [14]) *If the semigroup $\mathcal{P}_t$ is a strong Feller and irreducible, then all probability measures $\mathcal{P}_t(x,\cdot)$, with $x \in \mathbb{R}^N$ and $t > 0$, are equivalent.*

Theorem 4.4 (see [7]) *Let μ be the invariant measure of the semigroup $\mathcal{P}_t$. Assume that there exists t_0 such that all probability measures $\mathcal{P}_t(x,\cdot)$, with $x \in \mathbb{R}^N$ and $t > t_0$, are mutually equivalent. Then the following statements hold:*

(1) *μ is the unique invariant measure for $\mathcal{P}_t$ and in particular is ergodic;*
(2) *μ is equivalent to all the probability measures $\mathcal{P}_t(x,\cdot)$ for any $x \in \mathbb{R}^N$ and $t > t_0$;*
(3) *μ is strongly mixing.*

Let us now consider our normal form reduced climate model. Under the strict energy dissipation assumption, one can show that the semigroup $\mathcal{P}_t$ for our model has the Feller property (see [6]). The uniform boundness of the second moment of the pathwise solution is known by Proposition 4.2. Thus, by Chebyschev inequality, the family of measures $\{\mathcal{P}_t(x_0,\,\cdot\,)\}_{t\geq 0}$ is tight for any fix $x_0 \in \mathbb{R}^N$. This means that for some $\nu \in P(\mathbb{R}^N)$, the family of measures $\{\mathcal{R}_t^*\nu\}_{t\geq 0}$ is tight. By Prokhorov theorem, there exists a sequence $\{\mathcal{R}_{t_n}^*\nu\}$ converging weakly to some $\mu \in P(\mathbb{R}^N)$, as $n \to +\infty$, for some sequence $\{t_n\}$ increasing to $+\infty$. Then, by Krylov-Bogoliubov theorem, μ is an invariant measure for $\mathcal{P}_t$. This completes the proof of Proposition 4.3 on the existence of the invariant measure of our climate model.

If we further assume that the diffusion part is uniformly non-degenerate, i.e., $g = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$ is uniformly positive definite in $\mathbb{R}^N$, one can show that the semigroup $\mathcal{P}_t$ for our reduced climate model is strong Feller and irreducible (see [6]). By Khas'minskii's theorem and Doob's theorem, we conclude that the invariant measure of our reduced climate model is unique, ergodic and strongly mixing. This completes the proof of Proposition 4.4.

4.3 Absolute continuity of the invariant measure

Our aim here is to discuss absolute continuity of the invariant measure with respect to the Lebesgue measure, in other words, the existence of the density.

The following result is proved in [3, 26].

Proposition 4.5 *Consider the normal form reduced climate model* (2.4) *with drift term* $f(x)$ *and diffusion coefficients* $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. *Given the strict energy dissipation through the cubic terms as defined in* (3.4) *and uniformly non-degenerate diffusion that* g *is uniformly positive definite in* $\mathbb{R}^N$, *the invariant measure of* (2.4) μ *is absolutely continuous with respect to the Lebesgue measure* η *and the density function*

$$p(x) := \frac{\mathrm{d}\mu}{\mathrm{d}\eta} > 0, \quad \eta\text{-a.e.},\ p(x) \in L^q_{\mathrm{loc}}(\mathbb{R}^N),\ \forall\, q \in \Big[1, \frac{N}{N-1}\Big).$$

The absolute continuity of the invariant measure follows from the absolute continuity of the transition probabilities $\mathcal{P}_t(x,\,\cdot\,)$. Let us consider the semigroup $\mathcal{P}_t^r$ defined by

$$\begin{aligned}\mathcal{P}_t^r\varphi(x) &= E\varphi(\xi(t\wedge\tau_r^x, x)), \quad x \in B(0,r),\ \varphi \in B_r(\mathbb{R}^N),\\ \tau_r^x &= \inf\{t \geq 0 : |\xi(t,x)| \geq r\}.\end{aligned}$$

The absolute continuity of the transition probabilities $\mathcal{P}_t(x,\,\cdot\,)$ can be obtained by utilizing the absolute continuity of the transition probabilities $\mathcal{P}_t^r(x,\,\cdot\,)$ and taking the limit as r goes to $+\infty$.

However, if we do not have the uniformly non-degenerate diffusion part, the absolute continuity of the invariant measure of our normal form reduced climate model is not clear in general. For example, in the 1-D case (3.10) in Section 3.2, if we do not have the additive noise, we can have some cases with smooth and integrable densities, we can also find some special parameters such that the invariant measure has Dirac delta mass. Actually, the following result has been shown in [3].

Proposition 4.6 *Consider the normal form reduced climate model* (2.4) *with drift term* $f(x)$ *and diffusion coefficients* $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. *Given the strict energy dissipation through the cubic terms as defined in* (3.4), *if* μ *is an invariant measure, then* $(\det(g))^{\frac{1}{N}}\mu$ *has a density function which belongs to* $L_{\mathrm{loc}}^{\frac{N}{N-1}}(\mathbb{R}^N)$.

4.4 Smoothness and Gaussian upper bound assuming uniformly non-degenerate diffusion

It is important to check the regularity of the probability density function here, since we finally want to find the global bounds by comparison principle arguments.

Here, we assume that the diffusion part is uniformly non-degenerate. The results in [3] establish that the density function is positive everywhere and belongs to $W_{\mathrm{loc}}^{1,q}(\mathbb{R}^N)$, $\forall\, q > N$, where $W^{k,q}(\mathbb{R}^N)$ is the standard Sobolev space of functions on $\mathbb{R}^N$ whose generalized derivatives up to order k are in L^q equipped with its natural norm. Standard local regularity theorems as in [5] tell us that the weak solution to the elliptic equation with smooth coefficients is smooth. Thus, our density function is smooth since it satisfies Fokker Planck equation which is a second order linear elliptic equation. Moreover, we can use [25, 9] to get a Gaussian upper bound due to the strict energy dissipation through the cubic terms. Let us present the regularity and global behavior results by applying [3, 5, 25, 9] as the following.

Proposition 4.7 *Consider the normal form reduced climate model* (2.4) *with drift term* $f(x)$ *and diffusion coefficients* $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. *Given the strict energy dissipation through the cubic terms as defined in* (3.4) *and uniformly non-degenerate diffusion that* g *is uniformly positive definite in* $\mathbb{R}^N$, *the density function of the stationary statistical solution to* (2.4), $p(x)$, *has the following properties:*

(1) $p(x) \in C^\infty(\mathbb{R}^N)$;

(2) $p(x)$ *is positive everywhere*

$$\inf_{|x|\le r} p(x) > 0, \quad \forall\, r > 0;$$

(3) $p(x)$ *has at most Gaussian decay, i.e., there exists a positive real number* r *and a Gaussian measure* $p^G(x)$ *satisfying*

$$p^G(x) \geqslant p(x), \quad \forall\, |x| \geqslant r,$$

where

$$p^G(x) = \frac{1}{\sqrt{2\pi\sigma_G^2}}\mathrm{e}^{-\frac{|x|^2}{2\sigma_G^2}}, \quad x = (x_1, x_2, \cdots, x_N)^{\mathrm{T}} \in \mathbb{R}^N. \tag{4.8}$$

4.5 Gaussian upper bound by comparison principle argument

Here, we find a Gaussian upper bound for the density function of the stationary statistical solution of the normal form reduced climate model (2.4), given the strict energy dissipation through the cubic terms as defined in (3.4) and sufficient smoothness of the density function.

The smoothness and Gaussian upper bound are known by Proposition 4.7 assuming uniformly non-degenerate diffusion. When it comes to the case with degenerate diffusion, the regularity of the density is not clear in general, since the density can have delta mass sometimes (see 1-D case). However, in some important climate models, for example the 2-D climate model (3.9) in Section 3.1.1, the diffusion part is not uniformly non-degenerate either. But a rich family of numerical simulations show that the probability density functions of their invariant measures exist, and have smoothness and fast decay under various parameter sets. We also learn from Section 3.3 that for a rich family of parameter sets, the Fokker Planck operator is hypoelliptic by Hormander condition. So, it will be reasonable to assume that $p(x) \in C^2(\mathbb{R}^N)$ and has at most power law decay.

Proposition 4.8 *Consider the normal form reduced climate model* (2.4) *with drift term* $f(x)$ *and diffusion coefficients* $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$. *Given the assumptions below, if* $p(x)$ *is a density function of the stationary statistical solution to* (2.4), *then it has at most Gaussian decay, i.e., there exists a positive real number* r *and a Gaussian measure* $p^G(x)$ *satisfying*

$$p^G(x) \geqslant p(x), \quad \forall\, |x| \geqslant r,$$

where

$$p^G(x) = \frac{1}{\sqrt{2\pi\sigma_G^2}} \mathrm{e}^{-\frac{|x|^2}{2\sigma_G^2}}, \quad x = (x_1, x_2, \cdots, x_N)^{\mathrm{T}} \in \mathbb{R}^N.$$

These assumptions are

(1) *Strict energy dissipation through the cubic terms as defined in* (3.4);

(2) *Weak polynomial decaying smooth stationary statistical solution:* $p(x) \in C^2(\mathbb{R}^N)$, *such that there exists an* α *with*

$$\alpha > \frac{\max\limits_{|\vec{\omega}|=1} Q_2(\vec{\omega})}{\min\limits_{|\vec{\omega}|=1} Q_1(\vec{\omega})}, \qquad \lim_{|x| \to +\infty} |x|^{\alpha} p(x) = 0,$$

where $Q_1(x)$ *is defined in* (3.2), *and* $Q_2(x)$ *is defined in* (3.6).

Proof Assume that $p(x)$ is the probability density function of the stationary statistical solution (invariant measure) of the reduced stochastic model for N-climate variables which is rewritten here for emphasis

$$\begin{aligned} \mathrm{d}x_i = {} & \widetilde{F}_i \mathrm{d}t + \widetilde{A}_i x \mathrm{d}t + \widetilde{B}_i(x,x)\mathrm{d}t \\ & + \sum_p \frac{1}{\gamma_p}\Big(I_{ip}^M x_i + \sum_{j \neq i} B_{ij,p} x_j\Big)\Big(\sum_j I_{pj}^M x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j\Big)\mathrm{d}t \\ & + \sum_p \Big(L_{ip} + I_{ip}^M x_i + \sum_{j \neq i} B_{ij,p} x_j\Big)\frac{\sigma_p}{\gamma_p}\mathrm{d}W_p + \sigma_{A_i}\mathrm{d}W_{A_i}. \end{aligned}$$

The density $p(x)$ satisfies Fokker-Planck equation in the strong sense assuming $p(x) \in C^2(\mathbb{R}^N)$ which is satisfied under strict energy dissipation assumption,

$$0 = L_{\mathrm{FP}} p = -\nabla \cdot [f(x)p(x)] + \nabla^2[g(x)p(x)] = \sum_i \sum_j g_{ij} p_{x_i x_j} + \sum_i b_i p_{x_i} + hp, \tag{4.9}$$

where $(\cdot)_{x_i}$ denotes the first derivative and $(\cdot)_{x_i x_j}$ denotes the second derivative, and

$$\begin{aligned}
f_i(x) &= \widetilde{F}_i + \widetilde{A}_i x + \widetilde{B}_i(x,x) \\
&\quad + \sum_p \frac{1}{\gamma_p}\Big(I_{ip}^M x_i + \sum_{j\neq i} B_{ij,p} x_j\Big)\Big(\sum_j I_{pj}^M x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j\Big), \\
g_{ii}(x) &= \sum_p \frac{\sigma_p^2}{2\gamma_p^2}\Big(L_{ip} + I_{ip}^M x_i + \sum_{k\neq i} B_{ik,p} x_k\Big)^2 + \frac{\sigma_{A_i}^2}{2}, \\
g_{ij}(x) &= \sum_p \frac{\sigma_p^2}{2\gamma_p^2}\Big(L_{ip} + I_{ip}^M x_i + \sum_{k\neq i} B_{ik,p} x_k\Big)\Big(L_{jp} + I_{jp}^M x_j + \sum_{k\neq j} B_{jk,p} x_k\Big), \quad i\neq j, \\
b_i(x) &= \sum_j (g_{ij} + g_{ji})_{x_j} - f_i(x), \\
h(x) &= \sum_i \sum_j (g_{ij})_{x_i x_j} - \sum_i (f_i)_{x_i}.
\end{aligned}$$

Our idea is to construct a Gaussian comparison function $p^G(x) = \frac{1}{\sqrt{2\pi\sigma_G^2}} \mathrm{e}^{-\frac{|x|^2}{2\sigma_G^2}}$ such that it can control the tails of the stationary statistical solution of the reduced stochastic climate model. In terms of mathematics, we want to use the comparison principle procedure for the above equilibrium Fokker-Planck equation, the second order elliptic partial differential equations, to show that $p^G(x) - p(x) > 0$ for $|x|$ large enough.

However, since $p^G(x)$ does not monotonically depend on σ_G, it is not easy to directly pick the variance parameter σ_G for the Gaussian comparison function $p^G(x)$. Instead, we use an intermediate function $u = \mathrm{e}^{-\frac{|x|^2}{2\sigma_0^2}}$ which has Gaussian shape but is not normalized. It will be easier to show that there exists a large enough constant C_0 such that $C_0 u(x) - p(x) > 0$ for $|x|$ large enough, since we can adjust C_0 and σ_G separately and monotonically.

Another difficulty is that the comparison principle procedure requires $h(x)$ in (4.9) to be non-positive for all $|x|$ large enough, but it is not the case in our problem. Thus, we apply the comparison principle procedure to a new function $v(x) = \frac{C_0 u(x) - p(x)}{\omega(x)}$, where $\omega = |x|^{-\alpha}$. Moreover, we are going to utilize the fact that $v(x)$ decays to 0 as $|x|$ goes to $+\infty$, which can be obtained by using the power law upper bound of $p(x)$.

Step 1 We want to choose an α large enough such that $L_{\mathrm{FP}}\omega < 0$ for all $|x| > r$, where r is fixed and large.

We compute the derivatives of ω for $x \neq 0$,

$$\begin{aligned}
\omega_{x_i} &= \frac{-\alpha x_i}{|x|^2}\omega, \\
\omega_{x_i x_i} &= (-\alpha|x|^{-2} + \alpha(\alpha+2)x_i^2|x|^{-4})\omega, \\
\omega_{x_i x_j} &= (\alpha(\alpha+2)|x|^{-4} x_i x_j)\omega, \quad j \neq i.
\end{aligned}$$

Then, we observe that the cubic dissipation term dominates the asymptotic behavior of $L_{\mathrm{FP}}\omega$ as x large,

$$L_{\mathrm{FP}}\omega = \sum_i \sum_j g_{ij}\omega_{x_i x_j} + \sum_i b_i \omega_{x_i} + h\omega$$

$$= \Big(-\sum_i (f_i)_{x_i} |x|^2 - \alpha\Big(-\sum_i f_i x_i \Big) + O(|x|^3) \Big) \omega |x|^{-2}$$
$$= (Q_2(x) - \alpha Q_1(x) + O(|x|^3)) \omega |x|^{-2},$$

where $\frac{O(|x|^3)}{|x|^3}$ is bounded as x goes to ∞; $Q_1(x)$ is the leading order part of $-\sum_i f_i x_i$, as defined in (3.2),

$$\begin{aligned} Q_1(x) &= -\sum_i x_i \sum_p \frac{1}{\gamma_p} \Big(I^M_{ip} x_i + \sum_{j\neq i} B_{ij,p} x_j \Big) \Big(\sum_j I^M_{pj} x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j \Big) \\ &= -\sum_p \frac{1}{\gamma_p} \sum_i \Big(I^M_{ip} x_i x_i + \sum_{j\neq i} B_{ij,p} x_j x_i \Big) \Big(\sum_j I^M_{pj} x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j \Big) \\ &= \sum_p \frac{1}{\gamma_p} \Big(\sum_i I^M_{pi} x_i x_i - \sum_i \sum_{j<i} (B_{ij,p} + B_{ji,p}) x_j x_i \Big) \Big(\sum_j I^M_{pj} x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j \Big) \\ &= \sum_p \frac{1}{\gamma_p} \Big(\sum_j I^M_{pj} x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j \Big)^2; \end{aligned} \tag{4.10}$$

$Q_2(x)$ is the leading order part of $-\sum_i (f_i)_{x_i} |x|^2$, as defined in (3.6),

$$\begin{aligned} Q_2(x) = \Big(&-\sum_i \sum_p I^M_{ip} \frac{1}{\gamma_p} \Big(\sum_j I^M_{pj} x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j \Big) \\ &- \sum_i \sum_p \frac{1}{\gamma_p} \Big(2 I^M_{pi} x_i + \sum_{j\neq i} B_{p,ij} x_j \Big) \Big(I^M_{ip} x_i + \sum_{j\neq i} B_{ij,p} x_j \Big) \Big) |x|^2. \end{aligned}$$

By energy dissipation requirement, $Q_1(x) \geq \lambda |x|^4$. Moreover, there exists a positive constant μ, such that $Q_2(x) \leq \mu |x|^4$. Now, let us choose α big enough such that $\alpha > \frac{\mu}{\lambda}$. Then we get

$$L_{\mathrm{FP}} \omega \leq ((\mu - \alpha\lambda)|x|^4 + O(|x|^3)) \omega |x|^{-2} < 0,$$

as $|x|$ large enough.

Step 2 We want to choose a σ_0 large enough such that $L_{FP} u < 0$ $(\forall |x| > r)$, where r is fixed and large, $u(x) = \mathrm{e}^{-\frac{|x|^2}{2\sigma_0^2}}$.

We compute the derivatives of u

$$\nabla u = -\frac{x}{\sigma_0^2} u, \quad u_{x_i x_i} = \Big(\frac{x_i^2}{\sigma_0^4} - \frac{1}{\sigma_0^2} \Big) u, \quad u_{x_i x_j} = \frac{x_i x_j}{\sigma_0^4} u, \quad j \neq i.$$

Now, the cubic dissipation term and the quadratic diffusion term both play important roles in the asymptotic behavior of $L_{\mathrm{FP}} u$,

$$\begin{aligned} L_{\mathrm{FP}} u &= \sum_i \sum_j g_{ij} u_{x_i x_j} + \sum_i b_i u_{x_i} + hu \\ &= \Big(\sum_i \sum_j g_{ij} \frac{x_i x_j}{\sigma_0^4} - \frac{-\sum_i f_i x_i}{\sigma_0^2} + O(|x|^3) \Big) u \\ &= \Big(\frac{Q_3(x)}{\sigma_0^4} - \frac{Q_1(x)}{\sigma_0^2} + O(|x|^3) \Big) u, \end{aligned}$$

where $\frac{O(|x|^3)}{|x|^3}$ is bounded as x goes to ∞. Here, $Q_1(x)$ is the leading order part of $-\sum_i f_i x_i$, as defined in (3.2); $Q_3(x)$ is the leading order part of $\sum_i \sum_j g_{ij} x_i x_j$, as defined in (3.3),

$$\begin{aligned}
Q_3(x) &= \sum_i \sum_j \sum_p \frac{\sigma_p^2}{2\gamma_p^2} \Big(I_{ip}^M x_i + \sum_{k\neq i} B_{ik,p} x_k \Big) \Big(I_{jp}^M x_j + \sum_{k\neq j} B_{jk,p} x_k \Big) x_i x_j \\
&= \sum_i \sum_j \sum_p \frac{\sigma_p^2}{2\gamma_p^2} \Big(I_{ip}^M x_i x_i + \sum_{k\neq i} B_{ik,p} x_k x_i \Big) \Big(I_{jp}^M x_j x_j + \sum_{k\neq j} B_{jk,p} x_k x_j \Big) \\
&= \sum_p \frac{\sigma_p^2}{2\gamma_p^2} \sum_i \sum_j \Big(I_{ip}^M x_i x_i + \sum_{k\neq i} B_{ik,p} x_k x_i \Big) \Big(I_{jp}^M x_j x_j + \sum_{k\neq j} B_{jk,p} x_k x_j \Big) \\
&= \sum_p \frac{\sigma_p^2}{2\gamma_p^2} \Big(\sum_i \Big(I_{ip}^M x_i^2 + \sum_{j\neq i} B_{ij,p} x_i x_j \Big) \Big)^2 \\
&= \sum_p \frac{\sigma_p^2}{2\gamma_p^2} \Big(\sum_j I_{pj}^M x_j x_j + \sum_{k<j} B_{p,kj} x_k x_j \Big)^2 .
\end{aligned}$$

By strict energy dissipation assumption, we know that there exists a $\lambda > 0$, such that $Q_1(x) \geqslant \lambda |x|^4$ for all $x \in \mathbb{R}^n$. Moreover, $Q_3(x)$ is a semi-positive definite homogeneous polynomial with degree 4, there exists a $\mu' > 0$, such that $Q_3(x) \leq \mu' |x|^4$ for all $x \in \mathbb{R}^n$.

Now, we can choose $\sigma_0^2 > \frac{\mu'}{\lambda}$ such that $L_{\mathrm{FP}} u < 0$, $\forall\, |x| > r$, as r is large enough.

Step 3 In this step, we show that there is a C_0 so that $C_0 u(x) \geq p(x)$.

Consider the domain $D = \{|x| > r\}$, r is large enough such that $L_{\mathrm{FP}} u < 0$, $L_{\mathrm{FP}} \omega < 0$, $L_{\mathrm{FP}} p = 0$ in D.

Choose C_0 big enough such that $C_0 \mathrm{e}^{-\frac{r^2}{2\sigma_0^2}} > \max_{|x|=r} p(x)$. Set $v(x) = \frac{C_0 u(x) - p(x)}{\omega(x)}$, we have

$$L_{\mathrm{FP}}(v\omega) < 0, \quad x \in\ D \quad \text{and} \quad v(x) > 0, \quad |x| = r.$$

Moreover, weak power law decay of $p(x)$ says $|x|^\alpha p(x) \to 0$ as $|x| \to \infty$. So, $|v(x)| \to 0$ as $|x| \to \infty$. Thus, $v(x)$ is bounded in D.

Assume that there exists an $x_0 \in D$, such that $v(x_0) < 0$. There must be some point $y_0 \in D$ such that

$$v(y_0) = \min_{\overline{D}} v(x) < 0$$

by continuity of v.

Thus, we should have $\sum_i \sum_j g_{ij} v_{x_i x_j}(y_0) \geqslant 0$ (since (g_{ij}) is a semi-positive definite matrix), $\nabla v(y_0) = 0$, $v(y_0) < 0$. Therefore, at y_0,

$$L_{\mathrm{FP}}(v\omega) = (L_{\mathrm{FP}}\omega) v + \sum_i \Big(b_i \omega + 2 \sum_j g_{ij} \omega_{x_j} \Big) v_{x_i} + \omega \sum_i \sum_j g_{ij} v_{x_i x_j} \geqslant 0.$$

Since we know $L_{\mathrm{FP}}(v\omega) < 0$, this gives a contradiction.

Therefore, we must have $v(x) \geqslant 0$ in D, that is

$$p(x) \leqslant C_0 \mathrm{e}^{-\frac{|x|^2}{2\sigma_0^2}}, \quad \forall\, |x| \geqslant r.$$

Step 4 Now we use Step 3 to dominate $p(x)$ by the Gaussian comparison function $p^G(x)$. Choose $\sigma_G > \sigma_0$. Then for $|x|$ large, we have

$$p^G(x) = \frac{1}{\sqrt{2\pi\sigma_G^2}}\mathrm{e}^{-\frac{|x|^2}{2\sigma_G^2}} > C_0\mathrm{e}^{-\frac{|x|^2}{2\sigma_0^2}} = C_0 u \geq p(x).$$

4.6 Gaussian lower bound by comparison principle argument

Here, we find sufficient conditions such that there exists a Gaussian lower bound for the density function of the stationary statistical solution to the normal form reduced climate model (2.4).

Proposition 4.9 *Consider the normal form reduced climate model* (2.4) *with drift term* $f(x)$ *and diffusion coefficients* $g(x) = (g_{ij}(x)) = \frac{1}{2}\Sigma\Sigma^{\mathrm{T}}$*. Given the assumptions below, if* $p(x)$ *is a density function of the stationary statistical solution of* (2.4)*, then, it has at least Gaussian decay, i.e., there exists a positive real number* r *and a Gaussian measure* $p^L(x)$ *satisfying*

$$p(x) \geqslant p^L(x), \quad \forall\, |x| \geqslant r,$$

where

$$p^L(x) = \frac{1}{\sqrt{2\pi\sigma_L^2}}\mathrm{e}^{-\frac{|x|^2}{2\sigma_L^2}}, \quad x = (x_1, x_2, \cdots, x_N)^{\mathrm{T}} \in \mathbb{R}^N.$$

These assumptions are

(1) *Strict energy dissipation through the cubic terms as defined in* (3.4);

(2) *Strictly positive variance of the multiplicative energy fluctuation given in* (3.5);

(3) *Weak polynomial decaying smooth stationary statistical solution:* $p(x) \in C^2(\mathbb{R}^N)$*, such that there exists an* α *with*

$$\alpha > \frac{\max\limits_{|\vec{\omega}|=1} Q_2(\vec{\omega})}{\min\limits_{|\vec{\omega}|=1} Q_1(\vec{\omega})}, \quad \lim_{|x|\to+\infty} |x|^\alpha p(x) = 0,$$

where $Q_1(x)$ *is defined in* (3.2)*, and* $Q_2(x)$ *is defined in* (3.6);

(4) $p(x)$ *is positive everywhere*

$$\inf_{|x|\leq r} p(x) > 0, \quad \forall\, r > 0.$$

The proof will be similar to the comparison principle argument for Gaussian upper bound in Section 4.5. Let us use the same notation as in Section 4.5.

Step 1 By Step 1 in Section 4.5, we have the fact that there exists an α, such that $L_{\mathrm{FP}}\omega < 0$ for all $|x| > r$, where $\omega = |x|^{-\alpha}$, r is fixed and large.

Step 2 We want to choose a small enough $\sigma_0 > 0$ such that $L_{\mathrm{FP}}u > 0$, $\forall |x| > r$, where r is fixed and large, $u(x) = \mathrm{e}^{-\frac{|x|^2}{2\sigma_0^2}}$.

Applying the same calculation in Step 2, Section 4.5, we get

$$\begin{aligned}L_{\mathrm{FP}}u &= \sum_i\sum_j g_{ij}u_{x_ix_j} + \sum_i b_iu_{x_i} + hu\\ &= \Big(\sum_i\sum_j g_{ij}\frac{x_ix_j}{\sigma_0^4} - \frac{-\sum_i f_ix_i}{\sigma_0^2} + O(|x|^3)\Big)u\\ &= \Big(\frac{Q_3(x)}{\sigma_0^4} - \frac{Q_1(x)}{\sigma_0^2} + O(|x|^3)\Big)u,\end{aligned}$$

where $\frac{O(|x|^3)}{|x|^3}$ is bounded as x goes to ∞. Here, $Q_1(x)$ is the leading order part of $-\sum\limits_i f_ix_i$, as defined in (3.2); $Q_3(x)$ is the leading order part of $\sum\limits_i\sum\limits_j g_{ij}x_ix_j$, as defined in (3.3).

By conditions (1)–(2), $Q_1(x)$, $Q_3(x)$ are positive definite homogeneous polynomials with degree 4 in $\mathbb{R}^N$, so there exists an $\eta > 0$, such that $Q_3(x) \geq \eta|x|^4$ for all $x \in \mathbb{R}^N$, and there exists a $\lambda' > 0$, such that $Q_1(x) \leqslant \lambda'|x|^4$ for all $x \in \mathbb{R}^N$.

Now, we can choose $0 < \sigma_0^2 < \frac{\eta}{\lambda'}$ such that $L_{\mathrm{FP}}u > 0$, $\forall |x| > r$ as r is large enough.

Step 3 In this step, we show that there is a $C_0 > 0$, so that $C_0u(x) \leq p(x)$.

Consider the domain $D = \{|x| > r\}$, r is large enough such that $L_{\mathrm{FP}}u > 0$, $L_{\mathrm{FP}}\omega < 0$, $L_{\mathrm{FP}}p = 0$ in D.

Choose $C_0 > 0$ small enough such that $C_0\mathrm{e}^{-\frac{r^2}{2\sigma_0^2}} < \min\limits_{|x|=r} p(x)$. The minimum value on $\{|x| = r\}$ is positive due to the positivity requirement of the density $p(x)$. Set $v(x) = \frac{p(x)-C_0u(x)}{\omega(x)}$, we have

$$L_{\mathrm{FP}}(v\omega) < 0, \quad x \in D \quad \text{and} \quad v(x) > 0, \quad |x| = r.$$

Moreover, power law decay of $p(x)$ says $|x|^\alpha p(x) \to 0$ as $|x| \to \infty$. So, $|v(x)| \to 0$ as $|x| \to \infty$. Thus, $v(x)$ is bounded in D.

Assume that there exists an $x_0 \in D$, such that $v(x_0) < 0$. There must be some point $y_0 \in D$ such that

$$v(y_0) = \min_{\overline{D}} v(x) < 0$$

by continuity of v.

Thus, we should have $\sum\limits_i\sum\limits_j g_{ij}v_{x_ix_j}(y_0) \geqslant 0$ (since (g_{ij}) is a semi-positive definite matrix), $\nabla v(y_0) = 0$, $v(y_0) < 0$. Therefore, at y_0

$$L_{\mathrm{FP}}(v\omega) = (L_{\mathrm{FP}}\omega)v + \sum_i \Big(b_i\omega + 2\sum_j g_{ij}\omega_{x_j}\Big)v_{x_i} + \omega\sum_i\sum_j g_{ij}v_{x_ix_j} \geqslant 0.$$

Since we know $L_{\mathrm{FP}}(v\omega) < 0$, this gives a contradiction.

Therefore, we must have $v(x) \geqslant 0$ in D, that is

$$p(x) \geqslant C_0e^{-\frac{|x|^2}{2\sigma_0^2}}, \quad \forall\,|x| \geqslant r.$$

Step 4 Now we use Step 3 to find a Gaussian lower bound for $p(x)$.

Choose $0 < \sigma_L < \sigma_0$. Then for $|x|$ large, we have

$$p^L(x) = \frac{1}{\sqrt{2\pi\sigma_L^2}}\mathrm{e}^{-\frac{|x|^2}{2\sigma_L^2}} < C_0\mathrm{e}^{-\frac{|x|^2}{2\sigma_0^2}} = C_0u \leq p(x).$$

4.7 Conclusion of proofs for Theorems 3.1–3.3 and Corollary 3.1

Here, let us complete our proofs for the results in Section 3 by applying the propositions we got from previous sections.

Theorem 3.1, the existence of the density function, can be concluded from Proposition 4.6 (degenerate diffusion case) and Proposition 4.5 (uniformly non-degenerate diffusion case) in Section 4.3.

Theorem 3.2, Gaussian upper bound, is the direct result from Proposition 4.8 in Section 4.5.

Theorem 3.3, Gaussian lower bound, is the direct result from Proposition 4.9 in Section 4.6.

In Corollary 3.1, we further assume that the diffusion part is uniformly non-degenerate. So, the uniqueness of the density function can be obtained from Proposition 4.4, the smoothness and the Gaussian upper bound can be found in Proposition 4.7, and the Gaussian lower bound can be obtained by applying Propositions 4.7 and 4.9.

Acknowledgement The authors thank Professor Martin Hairer for helpful discussions.

References

[1] Berner, J. and Branstator, G., Linear and nonlinear signatures in planetary wave dynamics of an AGCM: probability density functions, *J. Atmos. Sci.*, **64**, 2007, 117–136.

[2] Bogachev, V. I., Krylov, N. V. and Rockner, M., Elliptic equations for measures: regularity and global bounds of densities, *J. Math. Pures Appl.*, **85**(6), 2006, 743–757.

[3] Bogachev, V. I., Rockner, M. and Shaposhnikov, S. V., On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions, *Comm. Part. Diff. Eqs.*, **26**(1112), 2001, 2037–2080.

[4] Bogachev, V. I., Rockner, M. and Shaposhnikov, S. V., Lower estimates of densities of solutions of elliptic equations for measures, *Doklady Mathematics*, **426**(2), 2009, 156–161.

[5] Browder, F. E., Regularity theorems for solutions of partial differential equations with variable coefficients, *Proc. Natl. Acad. Sci. USA*, **43**(2), 1957, 234–236.

[6] Cerrai, S., Second Order PDE's in Finite and Infinite Dimension: A Probabilistic Approach, Lecture Notes in Mathematics, **1762**, Springer-Verlag, New York, 2001.

[7] Doob, J. L., Asymptotic properties of markov transition probabilities, *Tran. Amer. Math. Soc.*, **3**, 1948, 393–421.

[8] Evans, L. C., Partial Differential Equations, Graduate Studies in Mathematics, **19**, A. M. S., Providence, RI, 1998.

[9] Fornaro, S., Fusco, N., Metafune, G. and Pallara, D., Sharp upper bounds for the density of some invariant measures, *Proc. Roy. Soc. Edinburgh, Sect. A Math.*, **139**, 2009, 1145–1161.

[10] Franzke, C., Majda A. J. and Branstator, G., The origin of nonlinear signatures of planetary wave dynamics: mean phase space tendencies and their information, *J. Atmos. Sci.*, **64**(11), 2007, 3987–4003.

[11] Gritsun, A. and Branstator, G., Climate response using a three-dimensional operator based on the fluctuation-dissipation theorem, *J. Atmos. Sci.*, **64**, 2007, 2558–2575.

[12] Gritsun, A., Branstator, G. and Majda, A. J., Climate response of linear and quadratic functionals using the fluctuation-dissipation theorem, *J. Atmos. Sci.*, **65**, 2008, 2824–2841.

[13] Hormander, L., Hypoelliptic second order differential equations, *Acta Math.*, **119**, 1967, 147–171.

[14] Khas'minskii, R. Z., Ergodic properties of recurrent diffusion processes and stabilization of the solution to the Cauchy problem for parabolic equations, *Theory of Prob. and Its Appl.*, **9**, 1960, 179–196.

[15] Krylov, N. V., Introduction to the Theory of Diffusion Processes, Translations of Mathematical Monographs, **142**, A. M. S., Providence, RI, 1995,

[16] Majda, A. J., Abramov R. and Gershgorin, B., High skill in low frequency climate response through fluctuation dissipation theorems despite structural instability, *PNAS*, **107**(2), 2010, 581–586.

[17] Majda, A. J., Abramov, R. and Grote, M., Information Theory and Stochastics for Multiscale Nonlinear Systems, CRM Monograph Series, **25**, A. M. S., Providence, RI, 2005.

[18] Majda, A. J., Franzke, C. and Crommelin, D., Normal forms for reduced stochastic climate models, *PNAS*, **16**(10), 2009, 3649–3653.

[19] Majda, A. J., Franzke C., Fischer A. and Crommelin, D., Distinct metastable atmospheric regimes despite nearly Gaussian statistics: A paradigm model, *PNAS*, **103**(22), 2006, 8309–8314.

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

[21] Majda, A. J., Gershgorin B. and Yuan, Y., Low frequency response and fluctuation-dissipation theorems: theory and practice, *J. Atmos. Sci.*, 2010, in press. DOI: 10.1175/2009JAS3264.1

[22] Majda, A. J., Timofeyev, I. and Vanden-Eijnden, E., A mathematical framework for stochastic climate models, *Commun. Pure Appl. Math.*, **54**, 2001, 891–974.

[23] Majda, A. J., Timofeyev, I. and Vanden-Eijnden, E., Models for stochastic climate prediction, *PNAS*, **96**(26), 1999, 14687–14691.

[24] Majda, A. J. and Wang, X. M., Linear response theory for statistical ensembles in complex systems with time-periodic forcing, *Comm. Math. Sci.*, **8**(1), 2010, 142–172.

[25] Metafune, G., Pallara, D. and Rhandi, A., Global regularity of invariant measures, *J. Funct. Anal.*, **223**, 2005, 396–424.

[26] Prato, G. D. and Goldys, B., Elliptic operators on $\mathbb{R}^d$ with unbounded coefficients, *J. Diff. Eqs.*, **172**(2), 2001, 333–358.

[27] Protter, M. H. and Weinberger, H. F., Maximum Principles in Differential Equations, Springer-Verlag, New York, 1984.

[28] Sardeshmukh, P. D. and Sura, P., Reconciling non-Gaussian climate statistics with linear dynamics, *J. Climate*, **22**(5), 2009, 1193–1207.

[29] Stephenson, D. B., Hannachi, A. and Oneill, A., On the existence of multiple climate regimes, *Q. J. R. Meteorol. Soc.*, **130**, 2004, 583–605.

[30] Sura, P. and Sardeshmukh, P., A global view of non-Gaussian SST variability, *J. Phys. Oceonogr.*, **38**, 2008, 639–647.