COMMUN. MATH. SCI.
Vol. 16, No. 1, pp. 17–51

PREDICTING EXTREME EVENTS FOR PASSIVE SCALAR TURBULENCE IN TWO-LAYER BAROCLINIC FLOWS THROUGH REDUCED-ORDER STOCHASTIC MODELS*

DI QI† AND ANDREW J. MAJDA‡

Abstract. The capability of using imperfect stochastic reduced-order models to capture crucial passive tracer statistics is investigated. The passive scalar field is advected by a two-layer baroclinic turbulent flow which can generate various representative regimes in atmosphere and ocean. Much simpler and more tractable block-diagonal linear Gaussian stochastic models are proposed to approximate the complex and high-dimensional advection flow equations. The imperfect model prediction skill is improved through a judicious calibration of the model errors using leading order statistics of the background advection flow, while no additional prior information about the passive tracer field is required. A systematic framework of correcting model errors with empirical information theory is introduced, and optimal model parameters under this unbiased information measure can be achieved in a training phase before the prediction. It is demonstrated that crucial principal statistical quantities like the tracer spectrum and fat-tails in the tracer probability density functions in the most important large scales can be captured efficiently with accuracy using the reduced-order tracer model in various dynamical regimes of the flow field with distinct statistical structures. The skillful linear Gaussian stochastic modeling algorithm developed here should also be useful for other applications such as accurate forecast of mean responses and efficient algorithms for state estimation or data assimilation.

Keywords. passive tracer turbulence; intermittency; low-order stochastic model.

AMS subject classifications. 76F25; 86A32; 62B10.

1. Introduction

Passive scalar turbulence concerns about the advection and diffusion of a tracer field passively transported by a turbulent dynamical flow. The problem has many applications in atmosphere, ocean, and climate change sciences, as well as many other branches of environmental science, such as the behavior of anthropogenic and natural tracers and the spread of contaminants or hazardous plumes [8,9,11,19,24]. Important and representative characteristics in the passive tracer field that can be observed in atmospheric observations, laboratory experiments, and numerical simulations include the extreme events represented by non-Gaussian exponential tails in probability density functions of the tracer field [3–5, 10, 23, 24]. Another remarkable property of these systems is their complex multiscale structures in both space and time, thus intermittent energy back cascades along the spectrum add important small scale feedbacks to tracer large scale variability. Major difficulties in accurate prediction about the large-scale tracer statistics are due to the inadequate numerical resolution for small scale feedbacks and incomplete physical understanding about the energy transport mechanism of the system. One way to address such complications is to build instructive simplified models which maintain the skill in generating the key statistical features of the realistic passive tracer field to understand the major mechanism in the passive tracer turbulence, so that the essential dynamical structures can be identified [8,17,19,22].

The scalar field $T(\mathbf{x},t)$ describes the concentration of the passive tracer immersed in the two-dimensional fluid on $\mathbf{x}=(x,y)$ which is carried with the local fluid velocity $\mathbf{v}$ but

*Received: April 26, 2017. Accepted (in revised form): August 7, 2017. Communicated by Shi Jin.
†Department of Mathematics and Center for Atmosphere and Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, (qidi@cims.nyu.edu).
‡Department of Mathematics and Center for Atmosphere and Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, (jonjon@cims.nyu.edu).

does not itself significantly influence the dynamics of the fluid. We consider this issue in the context of the evolution of the scalar field through the joint effect of turbulent advection and diffusion [19,29,31]

$$\frac{\partial T}{\partial t}+\mathbf{v}\cdot\nabla T=\mathcal{D}T=-d_T T+\kappa\Delta T. \tag{1.1}$$

In general we assume the dissipation as a linear damping and diffusion, $\mathcal{D}T=-d_T T+\kappa\Delta T$, where d_T has a stronger effect on the large-scale modes while κ applies to the small-scale modes for the diffusion and unresolved effects together. The statistical description of tracer transport in turbulent flows is a prevalent concern in atmospheric and oceanic fluid dynamics when the advection flow $\mathbf{v}$ becomes turbulent [28,31]. Generally the statistical solution for equation (1.1) cannot be obtained in explicit form because of the nonlinear advection term from $\mathbf{v}\cdot\nabla T$. However with certain simplifications about a prescribed tracer mean gradient and fluctuations around the mean gradient, there exists well developed statistical-dynamical understanding for the above problem with much simpler master equations for the advection flow with realistic physical characteristics [19, 22]. The mechanism of the existence and behavior of intermittency is investigated under a rigorous mathematical framework. Fat-tailed probability density functions and random spikes in the time series of the tracer can also be generated by numerical simulations using the idealized model (1.1) with a wide range of scales in the advection flow field [3,17,34].

In this paper, the advection velocity field $\mathbf{v}(\mathbf{x},t)$ is modeled by the two-layer quasi-geostrophic (QG) system [29,33]. Despite its relatively simple structure, the two-layer QG model can generate a wide variety of dynamical regimes that capture many essential features of both the atmosphere and ocean [29,32,36]. Typically there are anisotropic, intermittent, and inhomogeneous structures, and strong non-Gaussian statistics may appear in the flow field especially among the dominant large scales. For example, in the atmosphere regime it would exhibit a representative flow shifting between a cross-sweep zonal jet in the east-west direction and the transverse Rossby waves in the north-south direction. Theories and parameterizations for the anisotropic transport of a tracer in this two-dimensional anisotropic flow with jets are developed in [2,31]. A multi-scale stochastic superparameterization strategy has also been developed for the system and has been successfully applied in predicting the passive tracer equilibrium spectra and in multi-scale filtering [15]. The complexity and large computational expense in resolving the highly turbulent true advection flow equations require the introduction of simpler and more tractable imperfect models which still maintain the ability in capturing the key intermittent features in the tracer field. Notice that the central quantity of interest here is the tracer statistics rather than the background flow field. Thus it is advantageous to develop simpler reduced-order stochastic models that can capture the most important features of the background advection flow $\mathbf{v}$, and check the performance and model errors in the resulting scalar tracer T due to the imperfect modeling of flow solutions.

In turbulence literature, additional dissipation is introduced for coarse resolution numerical modeling of the turbulent diffusion of passive tracers, which is usually called eddy diffusion [19, 25, 29]. Stochastically excited linear system has been designed for modeling the nonlinear eddy interactions in QG turbulence [6, 7], but the procedure there often reports limited skill in capturing the mean response of the true system accurately [18]. A class of physics-constrained multi-level nonlinear regression models was developed which involves both memory effects in time as well as physics-constrained energy conserving nonlinear interactions [12,18] with mathematical rigor. In this paper, we address the reduced-order modeling problem for obtaining crucial principal statistics

of the passive scalar field immersed in a turbulent high-dimensional fluid field with inhomogeneous statistics. The primary goal is to obtain a mathematically unambiguous reduced-order stochastic modeling strategy with high skill for the scientific investigation of complex turbulent diffusion problems arising in many applications for which no explicit solution is available. We begin with simple model approximation considering linear Gaussian dynamics in the advection flow equations, then calibrate the model errors and overcome the possible barriers due to using this crude approximation of the flow field $\mathbf{v}$. The linear Gaussian model for the flow equation is decomposed into 2×2 block-diagonal subsystems and only low order equilibrium statistics are required in the model calibration as a correction to the nonlinear small scale feedbacks, thus the computational cost is significantly reduced. The challenge here is to develop efficient reduced-order stochastic models for the flow velocity field that keep the skill in capturing the leading order moments, as well as the higher order statistics that characterize intermittency in the principal directions in the tracer system.

In designing the reduced-order models, we propose imperfect model parameters to quantify the unresolved small scale feedbacks [28] in the stochastic advection flow approximation, and then use the imperfect flow solution to predict the intermittent structures in the turbulent passive tracer field from reduced-order tracer equations of the model (1.1). Therefore a systematic procedure in calibrating the imperfect model errors due to the crude approximations is necessary. In a training phase before the prediction, an information-theoretic framework [13, 16, 26] is proposed to tune the imperfect model parameters in addition to the equilibrium consistency in the leading order moments, so that the model predicted stationary process can possess the least biased estimation in energy and autocorrelation functions compared with the truth. The major statistical quantities to check for the reduced-order model prediction skill will include the mean, variance of the leading dominant tracer spectral modes that determine the major structure in the two-dimensional turbulent tracer field, as well as the fat-tailed tracer PDFs. The reduced-order model is then tested in three typical dynamical regimes representing the ocean and atmosphere with distinct statistics. Despite strong non-Gaussianity and anisotropic structures in the original turbulent advection flow, the linear stochastic model displays promising predicting skill for the tracer statistics and intermittency in large scale leading modes with high accuracy and effective reduction in computational cost.

The remainder of this paper is organized as follows. Section 2 introduces the spectral formulations for both the background flow equations and the scalar passive tracer equations advected by the flow field. Then Section 3 describes the calibration strategies for the linear Gaussian stochastic advection flow model so that consistency in first two leading order moments and autocorrelation functions can be guaranteed. The reduced-order tracer equations can be applied using the solution from the optimized flow equations. The skill of the reduced-order models is then tested under the two-layer quasi-geostrophic system in Section 4. A summary discussion about the results is shown in the final section of this paper.

2. Spectral dynamical formulation for the turbulent diffusion models with mean gradient

First we consider the general formulation for a scalar tracer $T(\mathbf{x},t)$ which is advected by a velocity field $\mathbf{v}(\mathbf{x},t)$. The turbulent advection flow is described from the solution

of the two-layer QG equation [29,33]

$$\begin{aligned} &\frac{\partial q_j}{\partial t}+\mathbf{v}_j\cdot\nabla q_j+\left(\beta+k_d^2U_j\right)\frac{\partial \psi_j}{\partial x}=-\delta_{2j}r\Delta\psi_j-\nu\Delta^s q_j,\\ &q_j=\Delta\psi_j+\frac{k_d^2}{2}\left(\psi_{3-j}-\psi_j\right),\quad \mathbf{v}_j=(U_j,0)+\mathbf{v}_j'. \end{aligned} \tag{2.1}$$

Above the subindex $j=1,2$ is used to represent the upper and lower layer of the two-layer flow model. The two dimensional incompressible velocity field $\mathbf{v}_j$ is decomposed into a zonal mean cross sweep, $(U_j,0)$, and a fluctuating shear flow $\mathbf{v}_j'=\nabla^\perp\psi_j=(-\partial_y\psi_j,\partial_x\psi_j)$. Next we introduce the passive tracer equation. Assume the existence of a background mean gradient for the tracer field varying in x and y directions and a tracer fluctuation component

$$T_j(\mathbf{x},t)=\boldsymbol{\alpha}\cdot\mathbf{x}+T_j'(\mathbf{x},t), \tag{2.2}$$

where $\boldsymbol{\alpha}=(\alpha_x,\alpha_y)$ is the horizontal mean gradient in x and y directions. In the atmosphere analogy a tracer gas would include a north-south mean gradient and fluctuations restricted to a narrow latitude band. The existence of a background mean gradient for the tracer field has been shown responsible for many important phenomena [1,3,10].

With the tracer decomposition into mean gradient and fluctuations around the mean in equation (2.2), the fluctuation parts of the tracer in upper and lower layers satisfy the following dynamics for the *full fluctuation passive tracer model* of tracer field $T_j'(\mathbf{x},t), j=1,2$

$$\frac{\partial T_j'}{\epsilon^{-1}\partial t}+\mathbf{v}_j'(\mathbf{x},t)\cdot\nabla T_j'+U_j\frac{\partial T_j'}{\partial x}=-(\alpha_x u_j+\alpha_y v_j)(\mathbf{x},t)-d_T T_j'+\kappa\Delta T_j'. \tag{2.3}$$

Above $\mathbf{v}_j'=(u_j,v_j)$ is the fluctuating advection flow field from the solution of equations (2.1) together with a zonal mean flow U_j. We further introduce a scale separation in the tracer equation (2.3) with order ϵ, while keeping the solution of the advection flow (2.1) unchanged. The parameter ϵ is in a similar role of the Péclet number $\mathrm{Pe}=UL/\kappa$ in the nondimensionalized system

$$\frac{\partial T}{\partial t}+\mathrm{Pe}(\mathbf{v}\cdot\nabla T)=\Delta T.$$

It introduces a different time scale, $\tilde{t}=\epsilon^{-1}t$, in the tracer time as in various previous works [17, 19, 22]. As $\epsilon<1$, the velocity field is varying at a faster time scale than the passive tracer process, while on the other hand with $\epsilon>1$ the tracer evolves in a more rapid rate than the advection field. A long time rescaling limit with explicit analytic tracer solutions is derived in [22] and numerical simulations for varying values of ϵ among a wide range are investigated in [3] under a much simpler linear model. In general, different intermittent features will be generated from near Gaussian statistics to distributions with fat tails as the scale separation parameter value changes [3,19].

In this section, we first derive the explicit equations for the two-layer advection flow field as well as the passive tracer field in spectral space. They can offer a detailed illustration about the energy mechanism especially from the nonlinear effect between the tracer and underlying flow field. Through a thorough understanding about the true model mechanism, the basic ideas about constructing reduced-order models can be revealed. The reduced-order model is then constructed with a careful calibration according to the true model nonlinear interactions.

2.1. Exact passive tracer and advection flow dynamical equations in spectral space. Given periodic boundary condition in both the two-layer flow and the tracer field, we formulate the flow and tracer fields with Galerkin truncation to finite number of Fourier modes. Spatial Fourier decomposition in flow potential vorticity q_j and passive tracer disturbance T_j' can be written in the expansion under modes $\exp(i\mathbf{k}\cdot\mathbf{x})$ as

$$q_j = \sum_{\mathbf{k}} \hat{q}_{j,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \quad T_j' = \sum_{\mathbf{k}} \hat{T}_{j,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}.$$

By projecting the tracer and flow equations (2.3) and (2.1) to each Fourier spectral mode, equations for the spectral coefficients in each wavenumber of the two-layer tracer field $\vec{T}_\mathbf{k} = \left(\hat{T}_{1,\mathbf{k}}, \hat{T}_{2,\mathbf{k}}\right)^T$, and two-layer advection flow field $\vec{q}_\mathbf{k} = (\hat{q}_{1,\mathbf{k}}, \hat{q}_{2,\mathbf{k}})^T$, form the set of ODEs in the spectral domain as

$$\frac{d\vec{T}_\mathbf{k}}{dt} + \epsilon^{-1} \sum_{\mathbf{m}+\mathbf{n}=\mathbf{k}} \left(A_{\mathbf{km}} \vec{q}_\mathbf{m} \circ \vec{T}_\mathbf{n} + A_{\mathbf{kn}} \vec{q}_\mathbf{n} \circ \vec{T}_\mathbf{m} \right) = -\epsilon^{-1} (\gamma_{T,\mathbf{k}} + i\omega_{T,\mathbf{k}}) \vec{T}_\mathbf{k} + \epsilon^{-1} G_\mathbf{k} \vec{q}_\mathbf{k}, \tag{2.4a}$$

$$\frac{d\vec{q}_\mathbf{k}}{dt} + \sum_{\mathbf{m}+\mathbf{n}=\mathbf{k}} (A_{\mathbf{km}} \vec{q}_\mathbf{m} \circ \vec{q}_\mathbf{n} + A_{\mathbf{kn}} \vec{q}_\mathbf{n} \circ \vec{q}_\mathbf{m}) = -(\gamma_{q,\mathbf{k}} + i\omega_{q,\mathbf{k}}) \vec{q}_\mathbf{k}, \tag{2.4b}$$

where '$\circ$' is used to denote the pointwise produce, $\mathbf{a} \circ \mathbf{b} = (a_i b_i)$. The potential vorticities $\vec{q}_\mathbf{k}$ and stream functions $\vec{\psi}_\mathbf{k}$ in two layers are related by the transform matrix $H_\mathbf{k}$

$$\vec{q}_\mathbf{k} = H_\mathbf{k} \vec{\psi}_\mathbf{k} = -\begin{bmatrix} |\mathbf{k}|^2 + \frac{k_d^2}{2} & -\frac{k_d^2}{2} \\ -\frac{k_d^2}{2} & |\mathbf{k}|^2 + \frac{k_d^2}{2} \end{bmatrix} \vec{\psi}_\mathbf{k},$$

through the relation $q_j = \nabla^2 \psi_j + \frac{k_d^2}{2}(\psi_{3-j} - \psi_j)$ in the equations (2.1). Note the core nonlinearity are dominated by the quadratic exchange of energy among triad modes on the left hand sides of equations (2.4) with the same explicit form, $A_{\mathbf{km}} = \frac{1}{2}(k_x m_y - k_y m_x) H_\mathbf{m}^{-1}$, for both the tracer and flow equations. On the right hand side of the tracer dynamics (2.4a), shear flow from the velocity field $\boldsymbol{\alpha}\cdot\mathbf{v}$ introduces the operator $G_\mathbf{k}$ applied on the flow potential vorticity as an external forcing, and the tracer damping $\gamma_{T,\mathbf{k}}$ and phase shifting operators $\omega_{T,\mathbf{k}}$ are due to dissipation and the advection in large-scale zonal mean flow respectively, that is,

$$G_\mathbf{k} = -i\boldsymbol{\alpha}\cdot\mathbf{k}^\perp H_\mathbf{k}^{-1} = \Gamma_\mathbf{k} H_\mathbf{k}^{-1}, \quad \gamma_{T,\mathbf{k}} = d_T + \kappa|\mathbf{k}|^2, \quad \omega_{T,\mathbf{k}} = k_x \vec{U}.$$

In the advection flow field dynamics (2.4b), the operators get the explicit forms

$$\gamma_{q,\mathbf{k}} = (0,1)^T \circ r|\mathbf{k}|^2 H_\mathbf{k}^{-1} + \nu|\mathbf{k}|^{2s}, \quad \omega_{q,\mathbf{k}} = k_x \left(\vec{U} + \left(\beta + k_d^2 \vec{U} \right) H_\mathbf{k}^{-1} \right),$$

as linear dissipation and dispersion effects. $\gamma_{q,\mathbf{k}}$ is due to the Ekman friction only applied on the bottom layer and the hyperviscosity, and $\omega_{q,\mathbf{k}}$ is from the rotational β-effect as well as the background zonal mean flow advection from the original equation (2.1) applied on the vorticity modes.

The above spectral models (2.4) for tracer advected by two-layer baroclinic turbulent flow can be compared with the simpler master models discussed in detail in [17,19]

$$\begin{aligned} d\hat{T}_k + \epsilon^{-1} ikU_t \hat{T}_k dt &= -\epsilon^{-1}(\gamma_{T,k} + i\omega_{T,k})\hat{T}_k dt - \epsilon^{-1}\alpha \hat{v}_k dt, \\ d\hat{v}_k + ikU_t \hat{v}_k dt &= -(\gamma_{v,k} + i\omega_{v,k})\hat{v}_k dt + \sigma_{v,k} dW_{k,t}, \\ dU_t &= -\gamma_U U_t dt + \sigma_U dW_t. \end{aligned} \tag{2.5}$$

The simpler formulation in equations (2.5) models the nonlinear interaction through the zonal jet U_t from a mean zero Ornstein–Uhlenbeck process in the third equation for both the tracer and flow dynamics, and similar linear dissipation and dispersion terms are applied on the right hand sides of the tracer and flow equations. Explicit solutions for the mean, variance, and cross-covariance between tracer and flow modes can be derived [17] due to this simple structure, and intermittency in such system can be predicted through the random resonance between modes of the turbulent velocity and the passive tracer [22]. Motivated from the theoretical understanding achieved from the simplified system (2.5), the system (2.4) advected by the two-layer QG flow contains more realistic features with inherent internal instability and strong non-Gaussian statistics. Thus we are able to investigate the more complicated triad nonlinear interactions between the flow and tracer modes in different scales, and the effects of non-Gaussianity in the advection flow field to the final tracer distributions instead of the simple large-scale Gaussian advection in the system (2.5). More importantly, ideas about reduced-order modeling strategies can be proposed by comparing the more complicated system with strong non-Gaussianity (2.4) and the simpler master equations (2.5).

2.2. Reduced-order stochastic formulations for the passive tracer and turbulent flow field. The advection terms in the tracer and flow equations (2.4) involve interactions between modes of different scales along the entire spectrum in a large dimensional phase space, thus usually high computational cost is required in achieving accurate statistical results from direct numerical simulations. In general intermittency in tracer field is dominated by the variability in largest scales, thus we will concentrate on the large-scale modes with wavenumber $|\mathbf{k}| \le M \ll N$, where M is the number of resolved modes and N is the full dimensionality of the system. Usually we could choose M much smaller than N that only covers the essential most energetic directions in the flow system. Below we first propose the simple strategy with linear corrections to approximate the advection flow field in the leading modes. Then the calibration and improvement of the imperfect models due to model errors from this approximation will be discussed in Section 3.

2.2.1. Consistent linear Gaussian model in the advection flow. In approximating *the advection flow*, we begin with the simple Gaussian approximation by replacing the quadratic interactions $(\mathbf{v}\cdot\nabla q)_{\mathbf{k}}$ in the flow equations by additional linear damping and random Gaussian noise. From the exact solutions from the simpler system (2.5), it has been shown that important non-Gaussian and intermittent structures in the tracer field can be generated from a linear Gaussian advection flow. It is useful to check the optimal prediction skill for the nonlinear tracer statistics using only linear Gaussian approximation in the flow field.

Therefore the **reduced-order advection flow equations** are proposed as

$$\begin{aligned} d\vec{q}_{M,\mathbf{k}} &= -(\gamma_{q,\mathbf{k}} + i\omega_{q,\mathbf{k}})\vec{q}_{M,\mathbf{k}} dt - D^M_{q,\mathbf{k}} \vec{q}_{M,\mathbf{k}} dt + \Sigma^M_{q,\mathbf{k}} d\vec{W}_{q,\mathbf{k}}, \quad 1 \le |\mathbf{k}| \le M, \\ \mathbf{v}_M &= \nabla^\perp \vec{\psi}_M, \quad \vec{q}_{M,\mathbf{k}} = H_{\mathbf{k}} \vec{\psi}_{M,\mathbf{k}}, \end{aligned} \tag{2.6}$$

with only Gaussian statistics generated. Only the first M large-scale modes, $1 \le |\mathbf{k}| \le M$, are resolved in the reduced-order formulation in equations (2.6). $(\gamma_{q,\mathbf{k}}, \omega_{q,\mathbf{k}})$ are

the linear dissipation and dispersion operators as in equation (2.4b) that are easy to model. On the other hand, the expensive but crucial part in the original equations (2.4b) is the nonlinear interactions $A_{\mathbf{km}}\vec{q}_{\mathbf{m}}\circ\vec{q}_{\mathbf{n}}$ where modes with different scales are coupled. Especially the nonlinear term represents the energy transfers between different modes so that the unstable modes due to internal instability can be balanced. Thus additional damping and noise $\left(D_q^M,\Sigma_q^M\right)$ are introduced to correct model errors due to the neglected nonlinear interactions in the flow equations. In this way, the original flow equation (2.4b) is decomposed into a series of 2×2 block-diagonal subsystems for each fixed wavenumber $\mathbf{k}$.

One of the simplest and most direct way to estimate the undetermined coefficients $\left(D_q^M,\Sigma_q^M\right)$ is through the mean stochastic model (MSM) discussed in [12, 18]. The parameter values are estimated from the characteristic quantities of total energy and decorrelation time. Despite the simplicity in the linear autoregressive models, reasonably skillful prediction in uncertainty quantification as well as filtering can be reached under this strategy for turbulent systems [12, 18]. However MSM still suffers several shortcomings when strong nonlinearity and slow mixing take place in the system [26]. Thus more detailed calibration of the imperfect model parameters are required especially when non-Gaussianity becomes crucial in the system. The optimization strategy for the undetermined imperfect model parameters will be discussed in detail next in Section 3 using detailed calibration from equilibrium statistics.

2.2.2. Direct stochastic modeling in the passive tracer equations. In *the tracer dynamics*, we would not introduce additional model calibrations of the tracer field statistics in case of over fitting of data. The idea here is to improve the reduced-order model prediction skill by optimizing the background advection flow field, thus the **reduced-order passive tracer equations** can be modeled through a direct truncation

$$d\vec{T}_{M,\mathbf{k}}+\left(\tilde{\mathbf{v}}_M\cdot\nabla\vec{T}_M\right)_{\mathbf{k}}d\tilde{t}=\Gamma_{\mathbf{k}}\vec{\psi}_{M,\mathbf{k}}d\tilde{t}-(\gamma_{T,\mathbf{k}}+i\omega_{T,\mathbf{k}})\vec{T}_{M,\mathbf{k}}d\tilde{t},\quad 1\le|\mathbf{k}|\le M.\tag{2.7}$$
$$\tilde{\mathbf{v}}_M=\sum_{|\mathbf{k}|\le M_1}i\mathbf{k}^{\perp}\vec{\psi}_{M,\mathbf{k}}e^{i\mathbf{k}\cdot\mathbf{x}},\quad M_1\le M.$$

Consistent with the advection flow model (2.6), only the first leading modes $|\mathbf{k}|\le M\ll N$ are resolved in the tracer approximation model.

Again the major difficulty in modeling the tracer dynamics is from the accurate approximation of the tracer advection $A_{\mathbf{km}}\vec{q}_{\mathbf{m}}\circ\vec{T}_{\mathbf{n}}$ in equation (2.4a). Exact modeling about this nonlinear interaction term requires the flow mode solution $\vec{q}_{M,\mathbf{k}}$ along the entire spectrum $0<|\mathbf{k}|\le N$, while only the first M leading modes are available through the reduced-order model. One crude approximation idea could be to replace the nonlinear advection in the tracer field, $\mathbf{v}(\mathbf{x},t)\cdot\nabla\vec{T}(\mathbf{x},t)$, with linear damping and noise in a similar fashion as the flow approximation model (2.6). However, as discussed in previous works [19, 22], the nonlinear advection in the tracer equation is crucial in the generation of many important statistical features including the intermittency. Thus, the inclusion of nonlinear effects from the flow solution is essential, at least for the large scale modes. On the left-hand side of the equation (2.7), the nonlinear advection $\tilde{\mathbf{v}}_M\cdot\nabla\vec{T}_M$ is modeled explicitly, but only the first $M_1\le M$ largest scale flow modes in the model velocity solution $\tilde{\mathbf{v}}_M$ are used to calculate the imperfect model tracer advection. This nonlinear advection is essential in generating the accurate spectra in tracer statistics, while it is also not expensive to calculate since only leading modes are involved. The idea for this approximation is through the assumption that the dominant features in tracer statistics (such as intermittency and equilibrium spectrum in large scales) are due to the leading

advection flow modes with largest energy. The contribution from the leading dominant modes in tracer statistics will be further illustrated in Section 4.1 through explicit model simulations.

3. Calibration of the imperfect model (using equilibrium statistics and information theory)

In this section, we discuss the strategies for the optimal prediction models through calibration of low-order statistics in the advection flow system. In the stochastic reduced-order model proposed in equation (2.6) for the advection flow equation, efficient decoupled linear Gaussian dynamics are used, while large model errors are introduced due to the approximation in the turbulent flow field. Strategies for finding optimal imperfect model parameters that minimize the model error are proposed here according to the corresponding statistical equations of the flow equation. In this modeling process, we only consider the calibration of the advection flow field, and the flow equation solution is applied to the tracer equations for predicting tracer statistics directly.

3.1. Model calibration in the advection flow field.

3.1.1. Statistical equations for the advection flow fluctuation. In constructing reduced-order flow models, we need to make sure that the imperfect model calibration parameters $\left(D_q^M, \Sigma_q^M\right)$ can properly reflect the true nonlinear energy mechanism from the true system. The consistent imperfect model then can be proposed by consulting the model statistical dynamics. Therefore it is useful to investigate the statistical equations for the second-order moments from the fluctuation equations (2.4b). The dynamics for the covariance matrix $R_{\mathbf{k}}^q = \langle \vec{q}_{\mathbf{k}} \bar{q}_{\mathbf{k}}^* \rangle$ of flow potential vorticity can be derived as a 2×2 system for each wavenumber [21, 28]

$$\frac{dR_{\mathbf{k}}^q}{dt} + \mathcal{L}_{\mathbf{k}}(\bar{q})\, R_{\mathbf{k}}^q + R_{\mathbf{k}}^q \mathcal{L}_{\mathbf{k}}^*(\bar{q}) + Q_F^q = (\mathcal{L}_{\mathbf{k}}^q + \mathcal{D}_{\mathbf{k}}^q)\, R_{\mathbf{k}}^q + R_{\mathbf{k}}^q (\mathcal{L}_{\mathbf{k}}^q + \mathcal{D}_{\mathbf{k}}^q)^*, \quad |\mathbf{k}| \le N. \tag{3.1}$$

The linear operators $(\mathcal{L}^q, \mathcal{D}^q)$ represent the skew-symmetric dispersion and dissipation effects from the right hand side of equation (2.4b). The additional operator $\mathcal{L}_{\mathbf{k}}(\bar{q})$ represents the interactions with a non-zero statistical mean state, where internal instability occurs with positive growth rate. Most importantly, the nonlinear interactions between different spectral modes introduce the additional nonlinear flux term Q_F^q indicating higher-order interactions. The small and large scale modes are linked through third-order moments $\langle \vec{q}_{\mathbf{m}} \vec{q}_{\mathbf{n}} \bar{q}_{\mathbf{k}}^* \rangle$ in Q_F^q between the triad modes $\mathbf{m}+\mathbf{n}=\mathbf{k}$. Thorough discussions about the energy mechanism and construction of statistical closure models for system (3.1) are detailed in [21, 27, 28] for more generalized systems.

Here our focus is on the low-order stochastic realization of the statistical closure model of system (3.1), thus solving the statistical equation (3.1) directly is not favorable considering its complexity. Still the nonlinear flux Q_F^q corresponds to the unresolved nonlinear effects in the stochastic model (2.6). Thus it is useful to exploit the nonlinear flux Q_F^q so that the imperfect model parameters $\left(D_q^M, \Sigma_q^M\right)$ can be proposed according to the true model energy transfer mechanism. Especially in statistical equilibrium as $t \to \infty$ the nonlinear fluxes can be calculated easily from the localized lower-order moments

$$Q_{F,\mathrm{eq}}^q = (\mathcal{L}_{\mathbf{k}}^q + \mathcal{D}_{\mathbf{k}}^q - \mathcal{L}_{\mathbf{k}}(\bar{q}_{\mathrm{eq}}))\, R_{\mathbf{k},\mathrm{eq}}^q + R_{\mathbf{k},\mathrm{eq}}^q (\mathcal{L}_{\mathbf{k}}^q + \mathcal{D}_{\mathbf{k}}^q - \mathcal{L}_{\mathbf{k}}(\bar{q}_{\mathrm{eq}}))^*. \tag{3.2}$$

Above in equation (3.2) only steady state statistics about the first two moments $\left(\bar{q}_{\mathrm{eq}}, R_{\mathrm{eq}}^q\right)$ are required, thus we do not really need to calculate the entire spectrum to get the higher order feedbacks in the dominant low wavenumber modes.

3.1.2. Imperfect model correction from equilibrium statistics and additional correction terms. As illustrated in above sections, the expensive nonlinear flux Q_F^q in the statistical Equations (3.1) is replaced by linear damping and noise with parameters $\left(D_q^M, \Sigma_q^M\right)$ in the stochastic models (2.6) representing the unresolved interactions. In fact, we can decompose the matrix $Q_F^q = Q_F^{q,+} + Q_F^{q,-}$ by singular value decomposition into positive-definite and negative-definite components. The positive definite part $Q_F^{q,+}$ illustrates the additional energy that injected into this mode from other scales, while the negative definite part $Q_F^{q,-}$ shows the extraction of energy through nonlinear transfer to other scales.

In the first step of constructing the stochastic damping and noise parameters, the idea is to make use of the equilibrium statistics of the nonlinear flux $Q_{F,\mathrm{eq}}^q$ which can be achieved easily from the lower-order equilibrium statistics as in equation (3.2). In adopting the true equilibrium statistics from $Q_{F,\mathrm{eq}}^q$, the true model energy transfer mechanism is respected and least artificial effect is introduced into the imperfect approximation model. Considering all these aspects, the first proposal for the linear damping and Gaussian random noise correction can be introduced as

$$D_{q,\mathbf{k}}^{\mathrm{eq}} = -\frac{1}{2} Q_{F,\mathrm{eq},\mathbf{k}}^{q,-} \left(R_{\mathbf{k},\mathrm{eq}}^q\right)^{-1}, \quad \Sigma_{q,\mathbf{k}}^{\mathrm{eq}} = \left(Q_{F,\mathrm{eq},\mathbf{k}}^{q,+}\right)^{1/2}. \tag{3.3}$$

The additional damping is from the negative definite equilibrium flux $Q_{F,\mathrm{eq}}^{q,-}$; and the positive definite equilibrium flux $Q_{F,\mathrm{eq}}^{q,+}$ acts as additional noise to the system. The above additional damping and noise (3.3) offer a desirable quantification for the minimum amount of corrections to stabilize the system with consistent equilibrium statistics for the mean and variance. This is the same idea applied to the statistical modified quasi-linear Gaussian closures developed in [30].

Still the above estimation of parameters (3.3) may not be optimal for the reduced-order Gaussian model considering that: i) it only guarantees marginal stability in the unstable modes for equilibrium; and more importantly ii) the mixing time scale in each mode (represented by the autocorrelation functions) may still lack the accuracy in the approximation using only equilibrium information. The nonlinear energy transferring mechanism may change with large deviation from the equilibrium case when intermittent fluctuations are present. The shortcomings for purely using the approximation (3.3) only from equilibrium statistics can be observed from the various numerical simulations [26]. As a further correction, we propose additional terms on top of approximation (3.3) with a simple constant damping for all the spectral modes and an additional noise accordingly to make sure consistency in energy

$$Q_{M,\mathbf{k}}^{\mathrm{add}} = -D_M^{\mathrm{add}} R_{M,\mathbf{k}} - R_{M,\mathbf{k}} D_M^{\mathrm{add}*} + \left(\Sigma_{M,\mathbf{k}}^{\mathrm{add}}\right)^2, \quad D_M^{\mathrm{add}} = \mathrm{diag}\left\{d_M + i\omega_M, d_M - i\omega_M\right\}. \tag{3.4}$$

Specifically for the two-layer system, we introduce the additional damping parameter d_M with same rate in the upper and lower layer. The imaginary part of the parameter ω_M introduces additional dispersion effect for correcting the time decay scale among the two layer modes. This further correction term in equation (3.4) is aimed to offer stabilizing effects in the marginal stable equilibrium form (3.3), and to offer further corrections in modeling the autocorrelation function that is important for the mixing rate in each spectral mode.

Combining the ideas in equations (3.3) and (3.4), we propose the additional damping and noise corrections for the reduced-order flow potential vorticity model (2.6) in the

following form

$$D^M_{q,\mathbf{k}} = -\frac{1}{2} Q^{q,-}_{F,\mathrm{eq},\mathbf{k}} \left(R^q_{\mathbf{k},\mathrm{eq}}\right)^{-1} + D^{\mathrm{add}}_M, \quad \Sigma^M_{q,\mathbf{k}} = \left(Q^{q,+}_{F,\mathrm{eq},\mathbf{k}} + \left(\Sigma^{\mathrm{add}}_{M,\mathbf{k}}\right)^2\right)^{1/2}. \tag{3.5}$$

Above the parameters $\left(D^M_{q,\mathbf{k}}, \Sigma^M_{q,\mathbf{k}}\right)$ are 2×2 matrices applied on the two-layer spectral mode $\vec{q}_{\mathbf{k}}$. Comparing with the exact true system (2.4b), the reduced-order approximation is equivalent to replacing the nonlinear interaction terms with the judiciously calibrated damping and noise in consideration with both the equilibrium energy transfer mechanism and further sensitivity correction. More detailed explanation about the reduced-order closure strategy for equations (3.3) and (3.4) can be found in Appendix B and [21, 28].

Next the task left in the Gaussian linear flow model is a systematic strategy to estimate the optimal imperfect model parameters in equation (3.4) in a training phase making sure that i) statistical equilibrium consistency for the mean and variance can be maintained; and ii) the mixing time scale represented by the autocorrelation function in each mode can be modeled correctly with the truth.

3.2. Imperfect stochastic model consistency in equilibrium statistics and autocorrelation functions. From the above construction of reduced-order stochastic models, we still have the undetermined model parameters $\left(D^{\mathrm{add}}_M, \Sigma^{\mathrm{add}}_M\right)$ in the linear damping and noise corrections in equation (3.5). These additional parameters can offer the freedom to control the imperfect model consistency in the leading order moments, and agreement in autocorrelation functions. Now we describe the model calibrations for consistency in equilibrium moments, and model optimization using autocorrelation functions under a systematic information-theoretic framework. This information-theoretic framework based on the relative entropy offers an unbiased and invariant measure for model distributions [13, 14], and has been utilized to systematically improve model fidelity and sensitivity and to make an empirical link between model fidelity and forecasting skill [16].

3.2.1. Climate consistency of the stochastic models for the first two moments. First we calculate the constraints for the imperfect model parameters so that the resolved modes can converge to the same equilibrium statistics in the first and second order moments in statistical steady state. The linear parts form the same linear dissipation and dispersion operators as in system (3.1), thus we only need to consider the imperfect model approximation of the nonlinear flux compared with the truth in Q_M and Q^q_F. The contribution from the additional damping D^M_q in equation (3.5) adds the effective outflow of energy in each mode as

$$Q^-_{M,\mathbf{k}} = D^M_{q,\mathbf{k}} R_{M,\mathbf{k}} + R_{M,\mathbf{k}} D^{M*}_{q,\mathbf{k}} = -\frac{1}{2} Q^-_{F,\mathrm{eq},\mathbf{k}} R^{-1}_{\mathbf{k},\mathrm{eq}} R_{M,\mathbf{k}} + D^{\mathrm{add}}_M R_{M,\mathbf{k}} + c.c.$$

The additional noise Σ^M_q in equation (3.5) acts as the inflow of energy in each mode as

$$Q^+_{M,\mathbf{k}} = \left(\Sigma^M_{q,\mathbf{k}}\right)^2 = Q^{q,+}_{F,\mathrm{eq},\mathbf{k}} + \left(\Sigma^{\mathrm{add}}_{M,\mathbf{k}}\right)^2.$$

In the statistical steady state, we would like to make sure the imperfect model statistics converge to the truth as

$$R^q_M \to R^q_{\mathrm{eq}}, \quad \text{as} \quad t \to \infty.$$

From the negative and positive definite correction terms above in Q_M^- and Q_M^+, the first component acts as the nonlinear flux approximation from equilibrium statistics, thus it will converge to the truth $Q_{F,\mathrm{eq}}$ in the equilibrium. Therefore we get the *necessary condition* for the additional damping and noise to make sure climate consistency so that

$$\left(\Sigma_{M,\mathbf{k}}^{\mathrm{add}}\right)^2 = D_M^{\mathrm{add}} R_{\mathbf{k},\mathrm{eq}} + R_{\mathbf{k},\mathrm{eq}} D_M^{\mathrm{add}*}. \tag{3.6}$$

In the two-layer case, D_M^{add} is defined in equation (3.4) for the 2×2 uniform damping in each spectral mode, and the 2×2 matrix $\Sigma_{M,\mathbf{k}}^{\mathrm{add}}$ adds additional noise into each spectral mode accordingly. With equation (3.3) one can verify that at the equilibrium, the approximate flux through linear damping and noise is consistent with the nonlinear flux computed from equation (3.2), that is,

$$\left(-D_{q,\mathbf{k}}^M R_{\mathbf{k}}^q - R_{\mathbf{k}}^q D_{q,\mathbf{k}}^{M*} + \left(\Sigma_{q,\mathbf{k}}^M\right)^2\right)_{t\to\infty} = -D_{q,\mathbf{k}}^{\mathrm{eq}} R_{\mathbf{k},\mathrm{eq}}^q - R_{\mathbf{k},\mathrm{eq}}^q D_{q,\mathbf{k}}^{\mathrm{eq}*} + \left(\Sigma_{q,\mathbf{k}}^{\mathrm{eq}}\right)^2 = Q_{F,\mathrm{eq}}^q.$$

3.2.2. Training phase for optimal autocorrelation functions with information theory. By choosing parameters according to equation (3.6), the climate consistency for the imperfect reduced-order models in the unperturbed equilibrium is guaranteed. Still the imperfect stochastic model may not be optimized due to the incorrect modeling about time scales and mixing rate in each mode (See Figure 4.5 in Section 4.2.2 as an example). On the other hand, we still have the additional model parameters (d_M, ω_M) in equation (3.4) to control the performance in each mode. The mixing time scale in the stochastic process for $\vec{q}_{\mathbf{k}}(t)$ can be characterized by the autocorrelation functions defined as

$$\mathcal{R}_{\mathbf{k}}(t) \equiv \left\langle (\vec{q}_{\mathbf{k}}(\tau) - \langle \vec{q}_{\mathbf{k}} \rangle)(\vec{q}_{\mathbf{k}}(\tau+t) - \langle \vec{q}_{\mathbf{k}} \rangle)^* \right\rangle, \tag{3.7}$$

that forms a 2×2 matrix for each spectral mode. It measures the lagged-in-time covariance $\left\langle \vec{q}'(t)\vec{q}'(0)^* \right\rangle$ in statistical stationary state as 'memory' in remembering the previous state in each mode, thus characterizes the mixing rate. Next we follow the general strategy introduced in [26] for measuring autocorrelation functions with the help of information theory.

The problem is to find a proper metric to measure the errors in the imperfect model autocorrelation functions in the training phase in a balanced way. A preferable measure to offer an unbiased metric for the imperfect model probability distribution π_M from the truth π is to use the relative entropy [13, 14, 16]

$$\mathcal{P}(\pi, \pi_M) = \int \pi \log \frac{\pi}{\pi_M}. \tag{3.8}$$

However one additional difficulty in tuning the autocorrelation functions under the information metric is that it is improper to directly substitute the autocorrelation (3.7) into the relative entropy formula (3.8) to calibrate the model error. The autocorrelation function $\mathcal{R}(t)$ may oscillate with negative values, whereas the probability density function π must stay positive everywhere in value. The problem can be solved by instead considering the spectral representation of the random process $\vec{q}(t)$, from the *theory of spectral representation* of stationary random fields [35]. It is proved a positive-definite matrix $E(\lambda) > 0$ can be constructed so that the *spectral representation* of the autocorrelation $\mathcal{R}(t)$ and stationary process $\vec{q}$ become

$$\mathcal{R}(t) = \int_{-\infty}^{\infty} e^{i\lambda t} dF(\lambda), \quad \vec{q}(t) = (\hat{q}_1, \hat{q}_2)^T = \int_{-\infty}^{\infty} e^{i\lambda t} \hat{\vec{Z}}(d\lambda), \tag{3.9}$$

where the spectral random measure $\vec{\hat{Z}}(d\lambda) = \left(\hat{Z}_1(d\lambda), \hat{Z}_2(d\lambda)\right)^T$ has independent increments and energy spectrum measured by $E(\lambda)$ or $dF(\lambda)$

$$dF(\lambda) = E(\lambda)\, d\lambda = \mathbb{E}\left|\vec{\hat{Z}}(d\lambda)\, \vec{\hat{Z}}^*(d\lambda)\right|.$$

Here specifically for the application in the two layer flow system, we view the spectral density and distribution functions E and dF as 2×2 matrices (see Appendix A for more details about the spectral expansion). In this way, we can optimize the autocorrelation functions $\mathcal{R}(t)$ including upper and lower layer covariances at the same time since these two modes are always closely coupled together.

Applying the theory for spectral representation of stationary processes, we find the one-to-one correspondence between the autocorrelation function $\mathcal{R}(t)$ and positive-definite energy spectra $E(\lambda)$. Back to the comparisons of the true and model random fields $\vec{q}, \vec{q}_M$ constrained within first two moments, $\mathcal{R}(t)$ for the true process $\vec{q}$ can be achieved through the data from true model simulations in equation (2.4b), while the imperfect model autocorrelation $\mathcal{R}_M(t)$ can be solved explicitly through the Gaussian linear process in equations (2.6). $E(\lambda)$ and $E_M(\lambda)$ then are the Fourier transforms of the autocorrelation matrices $\mathcal{R}(t), \mathcal{R}_M(t)$ according to equation (3.9). Therefore we can construct the spectral measures for two Gaussian random fields as a product of increment independent normal distributions about the frequency λ

$$p_G(x;\lambda) = \prod \mathcal{N}(0, E(\lambda)), \quad p_G^M(x;\lambda) = \prod \mathcal{N}(0, E_M(\lambda)).$$

The normalized relative entropy (3.8) between the two processes then can be defined under the spectral densities

$$\mathcal{P}\left(p_G, p_G^M\right) = \mathcal{P}(E(\lambda), E_M(\lambda)) \triangleq \int_{-\infty}^{\infty} \mathcal{P}\left(p_G(x;\lambda), p_G^M(x;\lambda)\right) d\lambda. \tag{3.10}$$

See [26] and Appendix A for a more detailed derivation about the above formula. We abuse the notation above using the spectra $E(\lambda)$ to denote density functions. Since E and E_M are positive definite 2×2 matrices for the spectral random variables, it is well-defined of the last part of the above formula (3.10) using the information distance formula (3.8). Through measuring the information distance in the spectral coefficients $\vec{\hat{Z}}(d\lambda)$, we get the lack of information in the autocorrelation function $\mathcal{R}(t)$ from the model. Furthermore, we have shown in [26] that the error in autocorrelation functions $\|\mathcal{R}(t) - \mathcal{R}_M(t)\|$ of two stationary random processes $\vec{q}$ and $\vec{q}_M$ is bounded by the information distance of their energy spectra $\mathcal{P}(E, E_M)$.

To summarize, we can seek a spectral representation of the autocorrelation functions like

$$\mathcal{P}(E(\lambda), E_M(\lambda)) = \int \mathcal{D}\left(E(\lambda)\, E_M(\lambda)^{-1}\right) d\lambda, \tag{3.11}$$

where $\mathcal{D}(x) \equiv -\log\det x + \mathrm{tr}x - 2$ is the Gaussian relative entropy with a zero mean state [13]. And since we only concentrate on the leading order statistics (that is, mean and variance), thus this representation is enough. We can find the optimal model parameter $\boldsymbol{\theta}_* = (d_M, \omega_M)$ by minimizing the information metric defined in (3.11)

$$\mathcal{P}(E(\lambda), E_M(\lambda, \boldsymbol{\theta}_*)) = \min_{\boldsymbol{\theta}} \mathcal{P}(E(\lambda), E_M(\lambda, \boldsymbol{\theta})). \tag{3.12}$$

The same strategy has been tested successfully for the simpler L-96 system in [26].

3.3. Summary about the numerical algorithm for the reduced-order stochastic models. With the optimal reduced-order stochastic approximation for the background advection flow equations (3.5) in which consistent leading order statistics and autocorrelation functions are guaranteed, the prediction about tracer statistics can be carried out by applying the reduced-order tracer equations as in equations (2.7). As a summary, here we describe the general reduced-order stochastic modeling procedure for predicting passive tracer statistics advected by a two-layer turbulent flow field.

ALGORITHM 3.1 (Reduced-order stochastic model for passive scalar turbulence). *Determine the low-dimensional subspace spanned by orthonormal basis* $\{\mathbf{e_k}\}$ *covering the regime with largest variability in the flow spectrum. The stochastic dynamical equations for advection flow and passive tracer field (2.6) and (2.7) can be set up by Galerkin projecting the original equations to the resolved subspace for wavenumbers* $1 \le |\mathbf{k}| \le M$,

$$\begin{aligned} d\vec{T}_{M,\mathbf{k}} &= -\left(\mathbf{v}_M \cdot \nabla \vec{T}_M\right)_{\mathbf{k}} dt + \Gamma_{\mathbf{k}} \vec{\psi}_{M,\mathbf{k}} dt - (\gamma_{T,\mathbf{k}} + i\omega_{T,\mathbf{k}}) \vec{T}_{M,\mathbf{k}} dt, \\ d\vec{q}_{M,\mathbf{k}} &= -(\gamma_{q,\mathbf{k}} + i\omega_{q,\mathbf{k}}) \vec{q}_{M,\mathbf{k}} dt - D^M_{q,\mathbf{k}} \vec{q}_{M,\mathbf{k}} dt + \Sigma^M_{q,\mathbf{k}} d\vec{W}_{q,\mathbf{k}}, \\ \mathbf{v}_M &= \nabla^\perp \vec{\psi}_M = \sum_{|\mathbf{k}| \le M_1} i\mathbf{k}^\perp \vec{\psi}_{M,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, \quad \vec{q}_{M,\mathbf{k}} = H_{\mathbf{k}} \vec{\psi}_{M,\mathbf{k}}. \end{aligned}$$

The flow potential vorticity equation for $\vec{q}_M$ *needs to be calibrated in a training phase for optimal model parameters, and then the solution from the flow equation is used for the prediction of tracer statistics in* $\vec{T}_M$ *in the most energetic leading modes.*

- *Calibration in advection flow equations: in the potential vorticity equations about* $\vec{q}_M$*, additional linear damping and Gaussian noise* $\left(D^M_q, \Sigma^M_q\right)$ *are introduced according to the flow equilibrium statistics and further corrections in equation (3.5). Equilibrium consistency is guaranteed through the constraint in equation (3.6), and the optimal model parameter is achieved through tuning the imperfect model autocorrelation function with the truth under information metric using equations (3.11) and (3.12).*
- *Prediction in passive tracer equations: the advection flow solution* $\mathbf{v}_M$ *is calculated from reduced-order model with the optimized parameters. Using the advection flow solution, the truncated tracer model is applied to get the statistics in passive tracer field. Tracer energy spectrum can be achieved by averaging the time-series of the tracer solution, and the tracer intermittency can be checked by comparing the tails in tracer probability density functions.*

4. Numerical experiments for passive tracer statistics in various regimes

In this section, we test the reduced-order models described in the previous sections among different dynamical regimes of the two-layer QG flow with distinct statistical features. In numerical simulations, the true statistics are calculated by a pseudo-spectra code with $128 \times 128 \times 2$ grid points in total. We use this relatively small truncation since it already suffices to generate strong large-scale intermittency. The zonal mean flow $\vec{U} = (U, -U)$ is taken as the same strength with opposite directions in the two layers. In the tracer simulations, for simplicity we always consider the mean gradient along y direction, that is to assume, $T = T' + \alpha y$. This assumption is representative in many previous investigations [17, 22, 26]. We may change the scale separation parameter ϵ in

regime	N	β	F	U	r	ν	s	d_T	κ	α
ocean, high lat.	128	2	40	0.1	0.1	1×10^{-13}	4	0.5	0.001	1
atmosphere, high lat.	128	1	4	0.2	0.2	1×10^{-13}	4	0.1	0.001	1
atmosphere, mid lat.	128	2	4	0.2	0.1	1×10^{-13}	4	0.1	0.001	1

Table 4.1: *Model parameters representing test regimes in ocean and atmosphere. The ocean regime gets a larger deformation frequency F than the atmosphere regime. N is the model resolution, β is the rotation parameter, $F=f_0^2L^2/4\pi^2g'H=k_d^2/4$ is inversely proportional to the deformation radius, U is the background mean shear flow, r is the Ekman drag in the bottom layer, and the hyperviscosity is measured by the operator $-\nu\nabla^{2s}$. The last three columns display the parameters for the tracer equations for linear damping d_T, tracer diffusivity κ, and mean gradient α along y..*

equation (2.4a) to generate different tracer structures. In most of the test case here, we use $\epsilon^{-1}=5$ where intermittency is prominent. Parameter values for both ocean and atmosphere regimes are following [13, 20] and are shown in Table 4.1. In the reduced-order models, we only compute the modes $|\mathbf{k}|\le M=10$ in largest scales, compared with the true system resolution $N=128$.

4.1. Statistics in the true system with full resolution. In the first place, we display the true model solutions with full resolution as an illustration of the flow structures. Using the two-layer baroclinic model (2.1), many typical dynamical regimes with distinct statistical features can be generated by varying the parameter values. Here we focus on three representative regimes as shown in Table 4.1. In the high latitude regimes, the statistics are relatively homogeneous with a competition between the zonal modes and meridional modes. The typical structure is the exchange between the blocked regime with strong meridional heat transfer and unblocked regime with strong zonal jet (see Figure 4.3 and [28, 29]). The mid latitude atmosphere regime forms another interesting dynamical regime with representative anisotropic zonal jets (see Figure 4.1 and [29]). Thus stronger non-Gaussian statistics can be expected in this regime.

As a typical illustration about the advection flow field in the three test regimes in Table 4.1, we compare the zonal averaged flow time-series, $\bar{u}=-\oint\psi_y(x,y)\,dx$, in Figure 4.1. In the high latitude ocean regime, there is obvious competition between the unblocked zonal flow and blocked meridional heat flux (see Figure 4.3 for the snapshots of stream functions in these two regimes). In the high latitude atmosphere regime, similar structure can be observed with stronger turbulent mixing. In contrast, the mid latitude atmosphere regime generates a meandering zonal jet, which displays the highly anisotropic structure in this flow field. These three dynamical regimes get different statistics and share representative features with the typical flow fields that are observed in real applications.

4.1.1. Statistics in the advection flow field. In the statistics of the advection flow field, most importantly we consider the first two leading moments of the mean and variance in each mode (and they are also important in constructing the reduced-order model for the flow dynamics); but it is still useful to check the higher order statistics about how large the steady state statistics divert from the Gaussian distributions. Usually to measure the higher order moments, two characterizing statistical quantities are the *skewness* and *flatness* of the variables of interest, that is,

$$\text{skewness}=\frac{\left\langle (u-\bar{u})^3\right\rangle}{\left\langle (u-\bar{u})^2\right\rangle^{3/2}},\quad \text{flatness}=\frac{\left\langle (u-\bar{u})^4\right\rangle}{\left\langle (u-\bar{u})^2\right\rangle^{2}}. \tag{4.1}$$

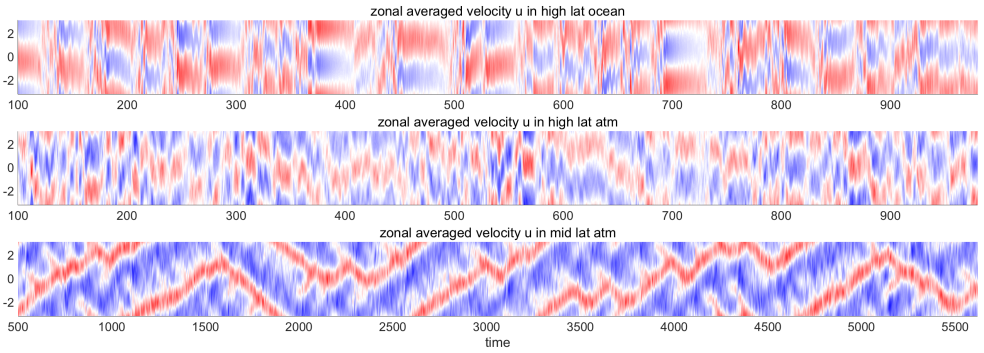


Fig. 4.1: *Time-series of the zonal averaged flow in the three test regimes with parameters in Table 4.1. The evolutions of the high latitude ocean (first row), the high latitude atmosphere (second row), and the mid latitude atmosphere (third row) are compared. The x-axis is time and y-axis is meridional coordinate.*

The skewness is used to measure the symmetry of the distributions of the random variables; flatness (kurtosis) indicates how the peak and tails of a distribution differ from the normal distribution. For a normal distribution, flatness gets the standard value of 3. A distribution with a flatness larger than 3 indicates that the distribution has tails heavier than the normal distribution (with same variance) and sharper peak. Likewise, a distribution with a flatness smaller than 3 indicates that the distribution has tails lighter than a normal distribution and flatter peak in the middle. One additional characterizing quantity is the autocorrelation function, $\mathcal{R}(u)$, which can be used to measure the mixing time scale of the process. The integrated correlation time of the process can be defined in two ways

$$M_{\text{corr}}^{\text{abs}} = \int |\mathcal{R}(t)|\, dt, \quad M_{\text{corr}} = \left| \int \mathcal{R}(t)\, dt \right|. \tag{4.2}$$

Both $M_{\text{corr}}^{\text{abs}}$ and M_{corr} are useful in measuring the mixing rate of the state variables for both the advection flow and the passive tracer modes [1, 19] especially when the autocorrelation function is highly oscillatory.

The equilibrium steady state statistics in the first four leading modes, $\hat{\psi}_{(1,0)}, \hat{\psi}_{(0,1)}, \hat{\psi}_{(1,1)}\hat{\psi}_{(-1,1)}$, of the flow stream functions are listed in Table 4.2 for the three test regimes. In the high latitude regimes, the first two modes (1,0) and (0,1) are dominant with much larger energy than the other modes, representing the Rossby waves and the zonal jets respectively. This can also be seen in Figure 4.2 with the energy spectra. For the variances in upper and lower layers of the flow, ocean regime has relatively the same amount of energy in both layers, and atmosphere regime accumulates more energy in the upper layer. In the mid latitude atmosphere regime, instead the zonal mode (0,1) gets much larger energy representing the meandering jet structure.

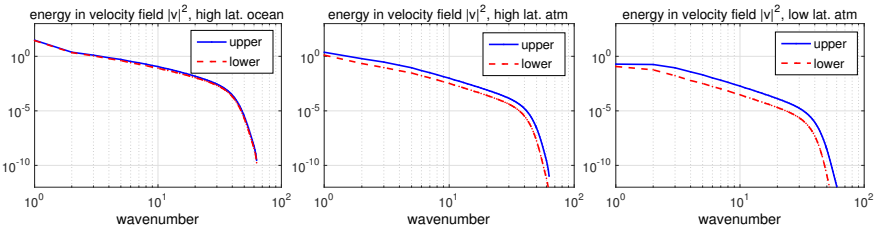
In higher order statistics, first note the skewness are all small within numerical errors for all the modes. This shows the symmetry in the distributions in the leading modes of interest. In high latitude regimes, for the fourth order moments for flatness, the first two modes (1,0) and (0,1) get values smaller than 3, representing sub-Gaussian distributions; while the next two modes (1,1) and (-1,1) have values larger than 3, representing fat tails. In mid latitude case, the zonal mode (0,1) becomes strongly sub-Gaussian while all the other modes are strongly super-Gaussian with fat tails. These

lower layer	mean	variance	skewness	flatness	correlation time
$\hat{\psi}_{(1,0)}$	0.0261	14.0134	-0.0136	2.4459	3.1190 (0.4673)
$\hat{\psi}_{(0,1)}$	0.1593	13.3364	0.0369	2.2103	6.2486 (6.1076)
$\hat{\psi}_{(1,1)}$	0.0050	0.4332	-0.0349	4.6093	1.3864 (0.2823)
$\hat{\psi}_{(-1,1)}$	0.0099	0.4379	0.0310	4.7116	1.4175 (0.2990)
high-latitude ocean regime $F=40, \beta=2$ (lower layer)					
lower layer	mean	variance	skewness	flatness	correlation time
$\hat{\psi}_{(1,0)}$	0.0046	0.5917	-0.0073	2.3438	4.1925 (1.1283)
$\hat{\psi}_{(0,1)}$	0.0215	0.5898	-0.0265	2.1646	3.5128 (2.8489)
$\hat{\psi}_{(1,1)}$	0.0018	0.0490	-0.0350	3.2605	1.7763 (1.0116)
$\hat{\psi}_{(-1,1)}$	0.0023	0.0500	-0.0421	3.1349	1.7144 (0.9326)
high-latitude atmosphere regime $F=4, \beta=1$ (lower layer)					
lower layer	mean	variance	skewness	flatness	correlation time
$\hat{\psi}_{(1,0)}$	0.0029	0.0111	0.0055	4.1219	17.7396 (0.7574)
$\hat{\psi}_{(0,1)}$	0.0002	0.1397	-0.0813	1.7995	97.6285 (73.672)
$\hat{\psi}_{(1,1)}$	0.0003	0.0176	-0.0557	3.6148	16.6485 (1.2648)
$\hat{\psi}_{(-1,1)}$	0.0005	0.0180	0.0082	3.5290	15.4606 (1.0893)
mid-latitude atmosphere regime $F=4, \beta=2$ (lower layer)					

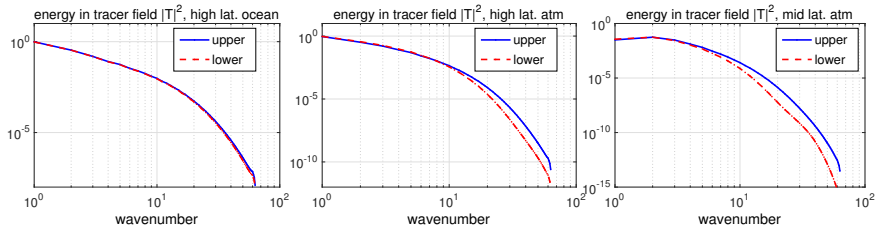
Table 4.2: *Equilibrium steady state statistics for the advection flow field stream functions $\hat{\psi}_{\mathbf{k}}$ in three test regimes. The (absolute) mean, variance, skewness, and flatness are compared in the first four most energetic modes. The original M_{corr} (in parenthesis) and absolute $M_{\text{corr}}^{\text{abs}}$ decorrelation time are shown in the last column, which illustrate the mixing scale for each mode in the advection flow.*

higher order statistics can be further observed in Figure 4.4 in the marginal PDFs. The correlation time M_{corr} and $M_{\text{corr}}^{\text{abs}}$ can be used to measure the mixing rate in each mode. Much longer mixing time can be observed also in the two leading modes. Another important feature to observe is the difference between M_{corr} and $M_{\text{corr}}^{\text{abs}}$. Mode $\hat{\psi}_{(0,1)}$ gets relatively similar values between M_{corr} and $M_{\text{corr}}^{\text{abs}}$, showing a slowly decaying mode (representing zonal flow); while mode $\hat{\psi}_{(1,0)}$ has much difference in M_{corr} and $M_{\text{corr}}^{\text{abs}}$, meaning strong oscillation (representing meridional transport mode).

4.1.2. Intermittency in the passive tracer field. Here we display the intermittent structure in the passive tracer field briefly in the true model. As a typical example we use the high latitude ocean regime with representative blocked and unblocked regimes as shown in Figure 4.3. Correspondingly we can observe the time-series of the meridional heat flux defined as $H_f = \int v\tau = \int \psi_x \tau$, with (ψ, τ) the barotropic and baroclinic mode. In the unblocked state with strong zonal flow, the meridional heat flux is suppressed; while in the blocked state strong meridional heat flux can be observed showing a strong meridional transport of the heat. The source of intermittency in passive scalar tracer is described qualitatively in [3] under an elementary model. The streamline topology for the flow field changes from unblocked behavior in the x direction (that is, parallel to the imposed tracer field mean scalar gradient αy along y), to very rapid heat transport in the y direction in the blocked state. The open streamlines in the meridional transport (like the left panel in first row in Figure 4.3), along with the y mean gradient, lead to large convective transport and large deformations of the isocontours for the passive scalar field, which promotes strong mixing by diffusion. On the other hand, in the regime with strong zonal flow along x direction (like the right panel



(a) energy spectra in advection flow field



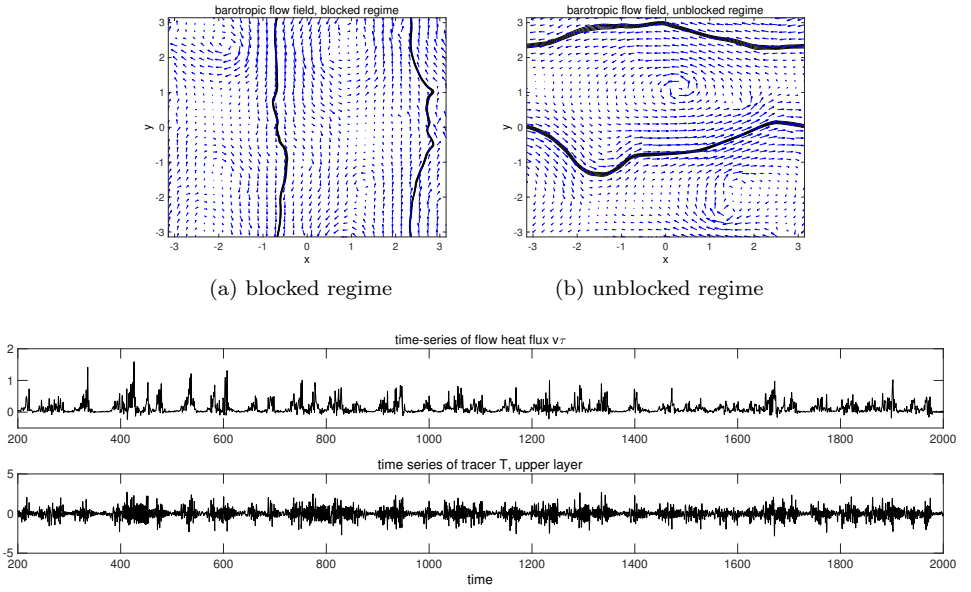
(b) energy spectra in passive tracer field

Fig. 4.2: *The energy spectra in the three typical test regimes. Both the energy spectra in flow velocity field,* $|\mathbf{k}|^2 \left\langle \left| \hat{\psi}_{\mathbf{k}} \right|^2 \right\rangle$*, and the tracer field,* $\left\langle \left| \hat{T}_{\mathbf{k}} \right|^2 \right\rangle$*, are shown in radial average.*

in first row in Figure 4.3), weak distortion of the scalar isocontours is caused by the almost entirely zonal convection of the flow field, and hence little opportunity for strong tracer mixing by diffusivity. Confirmed in the time-series of the scalar passive tracer trajectory in Figure 4.3, strong tracer intermittent bursts are always corresponding to the strong meridional heat transport in blocked regime leading to high mixing rate in tracer modes; the quite regime with little meridional heat flux in the unblocked regime also shows little intermittency of the tracer field in small amplitude.

The above on and off intermittent mechanism detailed in [3, 19] explains the turbulent mixing and intermittency via streamlines blocking and opening through simple elementary model with no instability and turbulence, while similar phenomena are also observed in the numerical simulations through much more complex two-layer baroclinic model in turbulence here. Even though simple and qualitative, the intuitive reasoning hints that the first two most energetic leading modes (0,1) and (1,0) which represent the competitive unblocked and blocked states might be crucial for modeling the intermittent structures in the tracer field. This also verifies the validity of only using small number of leading modes to model the tracer advection in equations (2.7) for the reduced-order tracer model.

4.2. Calibration in the advection flow field. Now we use the reduced-order stochastic model in equations (2.6) and (2.7) to predict scalar passive tracer statistics with only calibration of the background turbulent advection flow field. Only the leading modes $|\mathbf{k}| \leq 10$ are resolved for reducing model computational cost. As the model calibration strategy described in Section 3.1, we need to confirm equilibrium consistency in the energy spectra and time scale consistency in autocorrelation functions for the flow potential vorticity solutions.



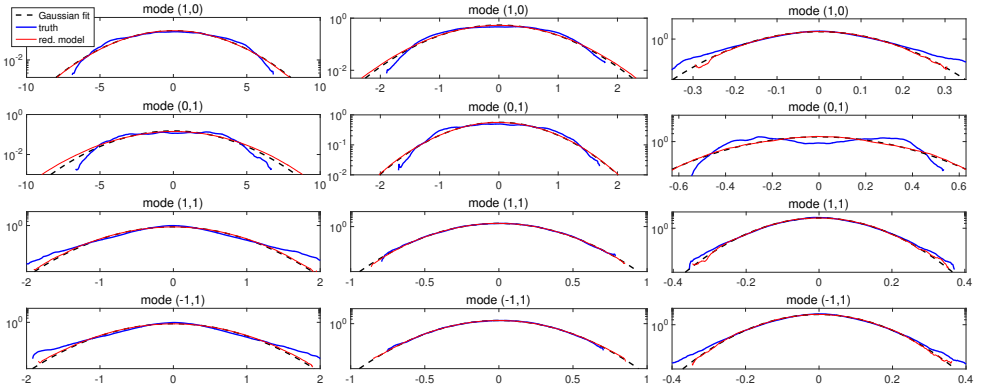
(a) blocked regime (b) unblocked regime

(c) time-series of heat flux and scalar tracer

Fig. 4.3: *Illustration about tracer intermittency with heat flux. The first row is the snapshots of the flow field with blocked and unblocked regime. The following parts show typical time-series for both the heat flux and scalar passive tracer.*

4.2.1. Consistency in energy spectra. As described in the modeling strategy in the previous section, the linear Gaussian flow approximation equations (2.6) should recover the true statistical energy in equilibrium in the first place. Choosing the parameters according to the constraint (3.6), naturally the same statistics can be reached since the true equilibrium first two moments are used to help construct the reduced model approximation for additional damping and noise corrections. To further confirm this convergence and make sure the validity of the numerical schemes, Figure 4.4 displays the equilibrium PDFs from the first 4 leading modes, $\hat{\psi}_{(1,0)}, \hat{\psi}_{(0,1)}, \hat{\psi}_{(1,1)} \hat{\psi}_{(-1,1)}$, of flow stream functions in truth and reduced-order model predictions. Consistency in variances is confirmed in these marginal distributions for the reduced-order model predictions overlapping the Gaussian fit of the truth with the same variance. Besides, consistent with the numbers in Table 4.2, the true leading modes all display some degree of non-Gaussian statistics with sub-Gaussian or super-Gaussian structures in the tails. Especially for the first two leading modes, that is, (1,0) for the Rossby waves and (0,1) for the large-scale zonal flow, strongly non-Gaussian features can be observed in all three regimes. On the other hand, the reduced-order model only captures the Gaussian statistics since only linear Gaussian model is used.

4.2.2. Consistency in autocorrelations. Still the climate consistency in the leading order moments is not sufficient for representing the low-order imperfect approximations in an optimal way. It is observed that large model error could still exist in the autocorrelation functions of each mode. As illustrated in Section 4.1.2 and Figure 4.3, the tracer intermittency is related to the competition between the zonal mode (0,1) and meridional mode (1,0). Thus the correct modeling about the mixing time scales in these



(a) high lat. ocean (b) high lat. atmosphere (c) mid lat. atmosphere

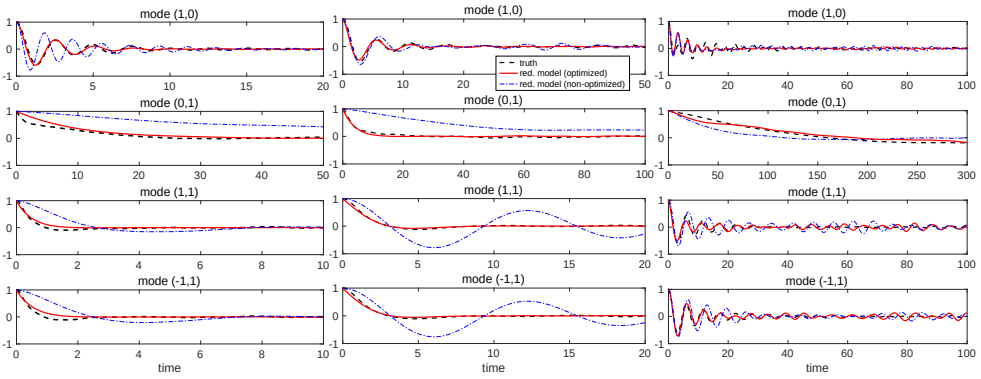
Fig. 4.4: *The marginal PDFs of the first four leading modes in stream functions in all three test regimes. The truth is shown in blue and red is the reduced model result. The Gaussian distributions with the same equilibrium variance are in dashed black lines. Since we use the equilibrium statistics to correct the linear Gaussian approximation models, exact recovery about the Gaussian fit can be achieved through this reduced-order model.*

modes is directly linked with the occurrence of intermittency, then with the accurate prediction of the fat-tails in tracer distributions.

The autocorrelation functions of the first four leading modes in the flow stream functions are plotted in Figure 4.5. Again we observe that the zonal mode (0,1) has a long decaying time representing the persistent zonal jet especially in mid latitude regime; while the Rossby mode (1,0) is highly oscillating responsible for the intermittent heat transport. Comparing the reduced-order model results with the optimal parameters from equation (3.12) and the non-optimized model with no additional corrections $d_M = 0, \sigma_M = 0$, it can be seen that in all three test regimes, the tuning process in the training phase can effective improve the accuracy in the predicted autocorrelation functions, while large error may still exist for the non-optimal case without additional correction. Important time scales in these crucial modes will be missed if no additional model calibration is considered. As we can see in the predictions for tracer statistics in the next section, this is also important for the reduced model skill for tracer results.

4.3. Predictions of turbulent tracer statistics in reduced-order models. In this section, we test the prediction skill of the reduced-order tracer models in equations (2.7) using the solution from the optimized flow equations as achieved above. As we have discussed, the nonlinear advection in the tracer equation $\mathbf{v}_M \cdot \nabla T_M$ is important for the final tracer statistical structure, while this part is expensive to calculate explicitly since it is nonlocal requiring the modes in different scales. So the strategy is to consider only principal modes with largest variance in $\mathbf{v}_M$ in calculating the nonlinear term and check the imperfect model prediction skill as in equations (2.7). Besides, we still have the scale separation parameter ϵ and linear damping d_T to control the tracer structures in the tracer model. This enables us to investigate the reduced model skill with different statistics by changing these parameter values.

4.3.1. Prediction in tracer equilibrium spectra and autocorrelation functions. In the reduced-order tracer equation (2.7), we would not like to introduce further calibrations of the tracer field. Thus in the first place, we should check the im-

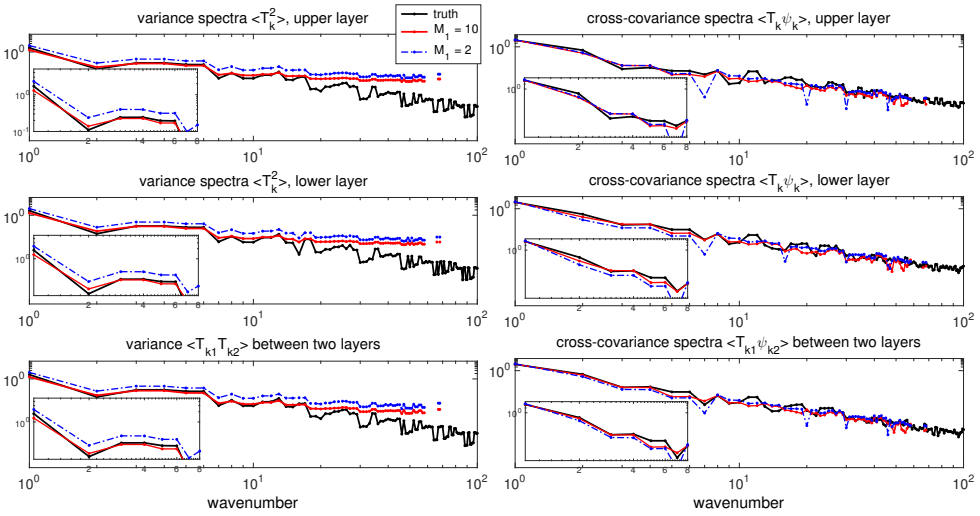


(a) high lat. ocean (b) high lat. atmosphere (c) mid lat. atmosphere

Fig. 4.5: *Autocorrelation functions in the first four most energetic modes in the flow field stream functions in all three test regimes. The truth in dashed blacked lines are compared with the optimized reduced-order model estimations in red lines. The optimal parameters in the reduced-order model is achieved by minimizing the information metric between the autocorrelations with the truth. For comparison, we also give the non-optimized results without additional corrections* $(d_M, \sigma_M) = 0$ *in dotted-dashed blue lines.*

perfect model skill in recovering the energy spectra and autocorrelation functions in the resolved tracer modes. As from the original tracer model (2.4a), the nonlinear advection in the tracer dynamics is nonlocal, while the reduced-order model only uses the leading modes in the flow field solution $\mathbf{v}_M = \sum \hat{\mathbf{v}}_{M,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{x}}, |\mathbf{k}| \le M_1$, to calculate the nonlinear part. By changing the size of truncated modes M_1, we can check the contribution of modes in different scales in tracer advection. In Figure 4.6-4.8, the equilibrium spectra in tracer modes and cross-covariance between tracer modes and flow stream functions are shown. To clearly compare the two-dimensional modes along one axis, the tracer modes are ordered all together with descending energy according to the advection flow modes. In the reduced-order model only the leading modes $|\mathbf{k}| \le 10$ are computed explicitly in the equations. Furthermore, we consider two different truncation sizes $M_1 = 10$ and $M_1 = 2$ in calculating the nonlinear advection in $\mathbf{v}_M$. Thus with $M_1 = 2$, only the first two dominant modes $(1,0)$ and $(0,1)$ are used in the nonlinear advection velocity, $\mathbf{v}_M \cdot \nabla T_M$. In general, the energy structure in the largest scales can be captured with desirable accuracy, while large errors appear with smaller size of advection modes. The cross-covariances between the tracer mode and stream function can also be captured through the reduced-order formulation. Comparing in more details in the first few dominant modes, using only two modes $M_1 = 2$ leads to larger errors due to the inaccurate modeling about the unresolved small scale feedbacks to the largest scales. Especially in the mid latitude regime with persistent zonal jet, using only two modes $M_1 = 2$ becomes insufficient with large errors. This shows the importance of small scale contributions in this regime. Overall, using the total $M_1 = 10$ resolved modes in the advection, the leading mode energy can be approximated with good agreement with the truth.

Next we check the reduced model prediction of the autocorrelation functions in the leading tracer modes. Results in all regimes in ocean and atmosphere are displayed in Figure 4.9. Overall the tracer autocorrelations show shorter decorrelation time than the corresponding advection flow modes in Figure 4.5. For the mid latitude case, the most important modes become the slow varying zonal modes, where the extremely long mixing



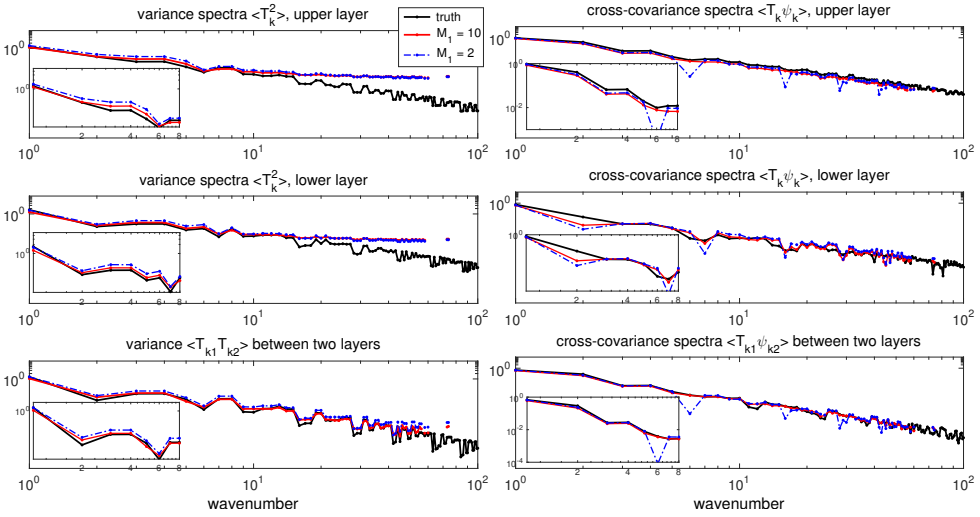
(a) tracer energy spectra (b) tracer-flow covariance

Fig. 4.6: *Tracer variance spectra predicted with the reduced-order model in high latitude ocean regime with tracer parameters* $(\epsilon^{-1}, d_T) = (5, 0.5)$. *The modes are ordered in descending order of the corresponding advection flow energy. The truth is shown in black, and we consider different number of resolved modes* $M_1 = 10$ *and* $M_1 = 2$ *in the reduced model advection.*

time is still captured in the model. Among all three test cases, the tracer autocorrelation functions can be captured with good accuracy in the reduced-order models.

4.3.2. Prediction in tracer intermittency and fat-tails in PDFs. Another important issue in the reduced-order models is the prediction of intermittency in the leading tracer modes. It has been observed from simulations in the true models that the tracers undergo a transition from Gaussian behavior to a probability distribution with approximately exponential fat-tails over a wide range of its variability with the change of model parameters. In the formulation in the tracer equations (2.4a), we can change the parameter value ϵ for scale separation between the tracer field and turbulent flow. The change of values in ϵ leads to changes in the Péclet number $\mathrm{Pe} = UL/\kappa$, where U is the characteristic scale in the advection flow $\mathbf{v}$. It is shown [3, 17, 22] that the passive tracer produces stronger intermittency as the tracer field approaches the long time slow varying limit $\epsilon \to 0$. Still here we would not like to push the system to the extremes but consider two typical cases with $\epsilon = 1/5$ (where strong fat-tailed distributions are developed) and $\epsilon = 1$ (where the tracer statistics become near Gaussian).

In Figure 4.10 and 4.11, we compare the high latitude atmosphere regime with two different scale separation parameter value $\epsilon = 1/5$ and $\epsilon = 1$. First observe that in the true signals from the perfect system with high resolution, larger spikes are produced in the tracer trajectories with smaller value $\epsilon = 1/5$, and obvious fat tails can be observed in the leading modes representing strong intermittent structures. On the other hand with larger value $\epsilon = 1$, the tracer field reduces to near Gaussian statistics where all the modes have the distributions close to the Gaussian fit. Similar case can be seen in the ocean regime case in Figure 4.12 and 4.13. Intermittency becomes weaker as ϵ is in larger value. Both the near Gaussian and highly non-Gaussian regimes can be representative showing a wide range of variability in the passive tracer field.



(a) tracer energy spectra (b) tracer-flow covariance

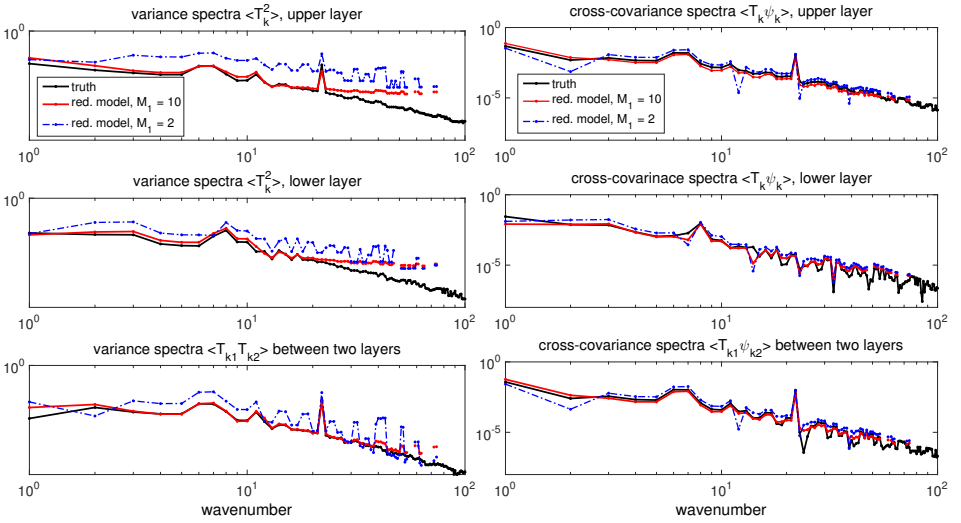
Fig. 4.7: *Tracer variance spectra predicted with the reduced-order model in high latitude atmosphere regime with tracer parameters* $(\epsilon^{-1}, d_T) = (5, 0.1)$. *The modes are ordered in descending order of the corresponding advection flow energy. The truth is shown in black, and we consider different number of resolved modes* $M_1 = 10$ *and* $M_1 = 2$ *in the reduced advection.*

Then we observe the reduced-order model skill in capturing these central non-Gaussian fat-tails. In Figure 4.10-4.13, we compare the imperfect model prediction in the high latitude regime in both ocean and atmosphere. Since the tracer intermittency is usually most important in the first few leading modes, we compare the representative time-series and tracer PDFs in statistical steady state. Among all the regimes, the fat-tails in the distribution functions can be captured, and similar characteristic structures can be seen in the truth and reduced model time-series. On the other hand in the nearly Gaussian regimes, still similar structures can be seen by comparing the truth and imperfect model results, and the reduced-order model can recover the tracer PDFs again with accuracy.

Finally, in Figure 4.14 and 4.15, the mid latitude regime with persistent zonal jet is tested. In this case, the representative feature is the slow-varying leading zonal modes that become dominant. In this case, the streamlines will be open with strong zonal flow for most of the time in the general situation, thus the contour lines of the tracer mean gradient along y is less frequent to be broken. Thus tracer turbulent mixing in this case is weaker and fat-tails in the PDFs is less obvious. The reduced-order model maintains the skill in capturing tracer statistics in leading modes among this regime. Notice here in all the reduced-order simulations, only the largest scale modes $|\mathbf{k}| \leq 10$ are resolved compared with the full model with resolution of 128 grid points along each horizontal direction. The model efficiency can be effectively improved by solving the low-order model.

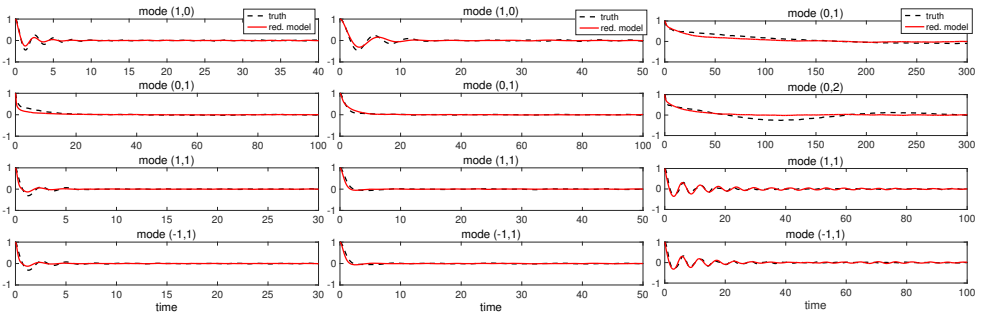
Reduced model error due to truncated modes in tracer advection term

In the advection term $\mathbf{v}_M \cdot \nabla T_M$ in modeling the passive tracer field (2.7) in the reduced-order approximation, only the resolved large-scale modes in the flow equation solution $\mathbf{v}_M$ are used in calculating the nonlinear tracer advection. That is, the tracer advection in the model is calculated through the large scale truncation. In above simu-



(a) tracer energy spectra (b) tracer-flow covariance

Fig. 4.8: *Tracer variance spectra predicted with the reduced-order model in mid latitude atmosphere regime with tracer parameters* $(\epsilon^{-1}, d_T) = (5, 0.1)$. *The modes are ordered in descending order of the corresponding advection flow energy. The truth is shown in black, and we consider different number of resolved modes* $M_1 = 10$ *and* $M_1 = 2$ *in the reduced advection.*



(a) tracer high lat ocean (b) tracer high lat atmosphere (c) tracer mid lat atmosphere

Fig. 4.9: *Prediction of the reduced-order autocorrelation functions in the tracer field in the first four most energetic modes. In the ocean case* $F = 40$, *the tracer parameters are* $(\epsilon^{-1}, d_T) = (5, 0.5)$; *and in the atmosphere case* $F = 4$, *the tracer parameters are* $(\epsilon^{-1}, d_T) = (5, 0.1)$. *In the autocorrelation functions, the truth is shown in black dashed lines and the reduced-order model prediction in red lines.*

lation results for tracer PDFs, we always use the entire number of resolved flow modes $M_1 = M$ from the advection velocity solution $\mathbf{v}_M$. Here as a further test for model errors through the high wavenumber truncation, we can even consider less modes $M_1 < M$ in computing the nonlinear advection in the tracer dynamics. In Figure 4.16, PDFs in the two representative atmosphere and ocean regimes are compared with different truncation sizes $M_1 = 10, 5, 2$ in computing the nonlinear tracer advection term with $\mathbf{v}_M$. With $M_1 = M = 10$, all the resolved modes from linear Gaussian flow model (2.6) are used in the tracer advection. Using half of the resolved modes, $M_1 = 5$, still we observe that

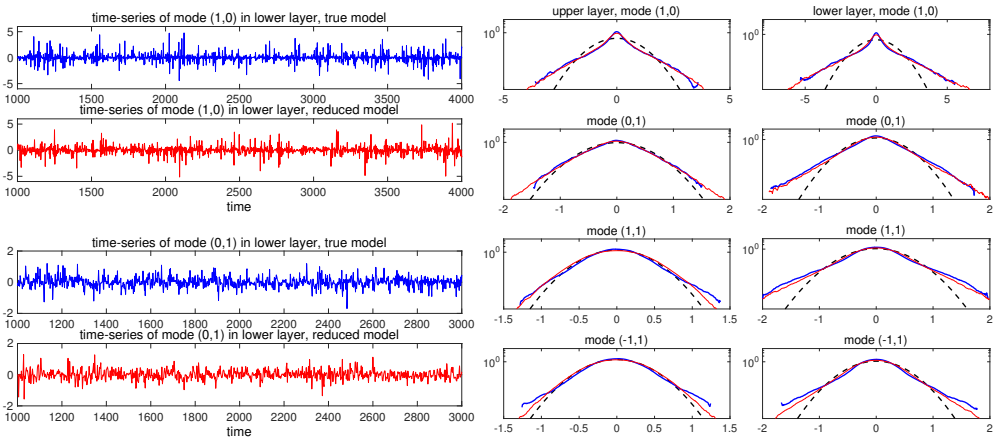


Fig. 4.10: *Prediction of tracer intermittency in high latitude atmosphere regime with parameters* $(\epsilon^{-1}, d_T) = (5, 0.1)$. *The left panel is the time-series for the first two leading modes (1,0) and (0,1) between true model and reduced model results; the right panel compares the PDFs in the first four modes between the truth in blue and reduced model prediction in red with the Gaussian fit in dashed black lines.*

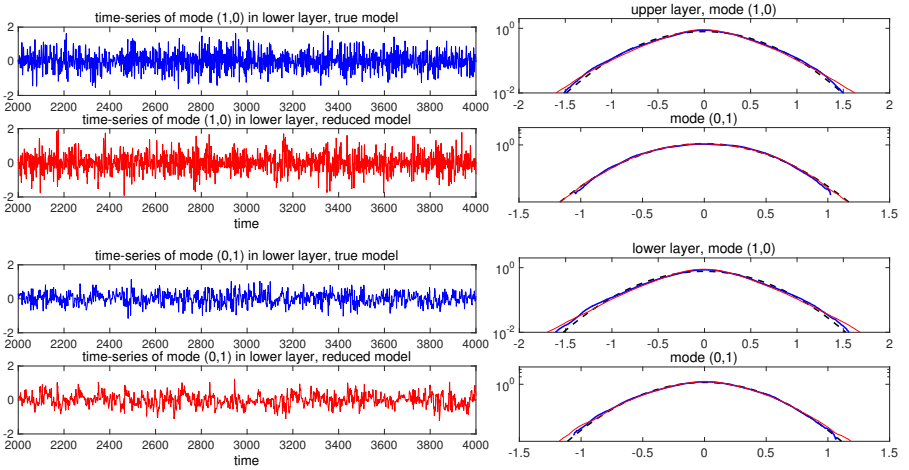


Fig. 4.11: *Prediction of tracer intermittency in high latitude atmosphere regime with parameters* $(\epsilon^{-1}, d_T) = (1, 0.1)$. *The left panel is the time-series for the first two leading modes (1,0) and (0,1) between true model and reduced model results; the right panel compares the PDFs in the first four modes between the truth in blue and reduced model prediction in red with the Gaussian fit in dashed black lines.*

the major structure with fat-tails can be captured in the leading PDFs with only a little larger error in the tails of the distributions. In the extreme case using the severe truncation $M_1 = 2$, only the first two modes are included in calculating the nonlinear advection. As a result, larger errors appear with this crude approximation due to the insufficient quantification for flow advection modes.

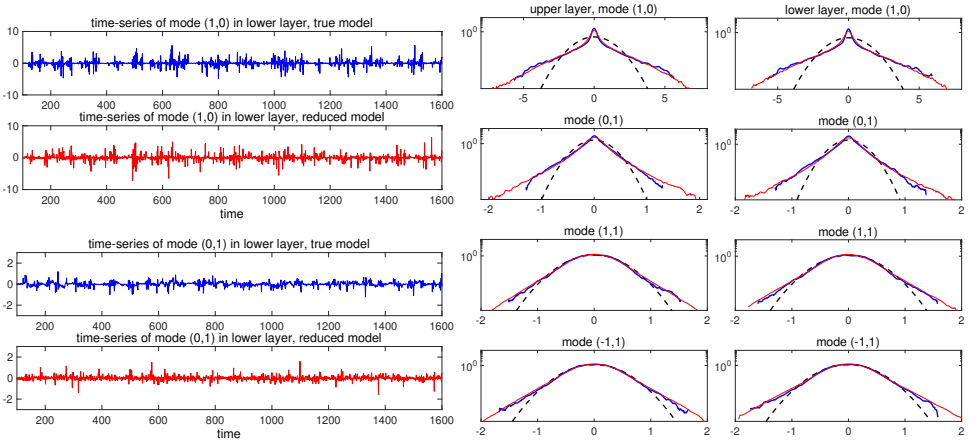


Fig. 4.12: *Prediction of tracer intermittency in high latitude ocean regime with parameters* $(\epsilon^{-1}, d_T) = (5, 0.5)$. *The left panel is the time-series for the first two leading modes (1,0) and (0,1) between true model and reduced model results; the right panel compares the PDFs in the first four modes between the truth in blue and reduced model prediction in red with the Gaussian fit in dashed black lines.*

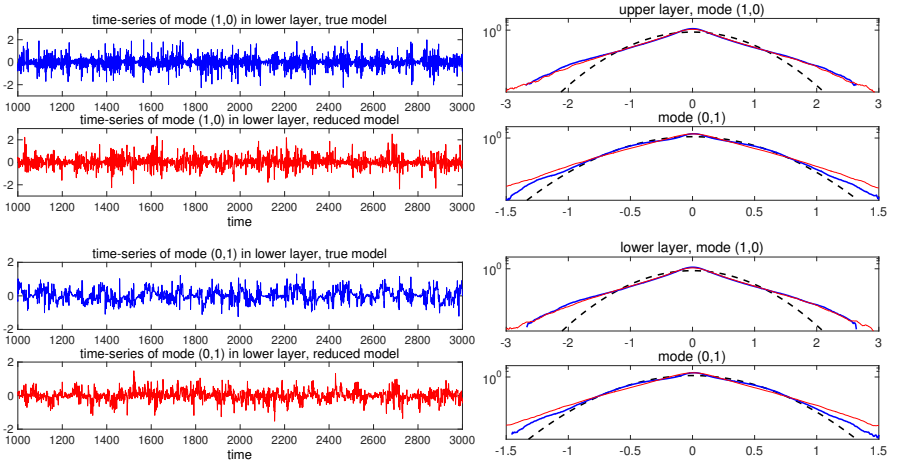


Fig. 4.13: *Prediction of tracer intermittency in high latitude ocean regime with parameters* $(\epsilon^{-1}, d_T) = (1, 0.5)$. *The left panel is the time-series for the first two leading modes (1,0) and (0,1) between true model and reduced model results; the right panel compares the PDFs in the first four modes between the truth in blue and reduced model prediction in red with the Gaussian fit in dashed black lines.*

4.3.3. Eddy diffusivity approximations for tracers. Parameterization by eddy diffusivity in the tracer system is often used in practice to account for unresolved small scales [17, 19]. Here we demonstrate the accuracy of the imperfect model in recovering the tracer eddy diffusivity. Consider the mean statistical dynamics about $\bar{T}_{\mathbf{k}} = \langle T_{\mathbf{k}} \rangle$ of the tracer equations by taking ensemble average about the original model (2.4a)

$$d\bar{T}_{\mathbf{k}} + \left(\bar{\mathbf{v}} \cdot \nabla \bar{T}\right)_{\mathbf{k}} dt + \langle \mathbf{v}' \cdot \nabla T' \rangle_{\mathbf{k}} dt = \Gamma_{\mathbf{k}} \bar{\psi}_{\mathbf{k}} dt - (\gamma_{T,\mathbf{k}} + i\omega_{T,\mathbf{k}}) \bar{T}_{\mathbf{k}} dt.$$

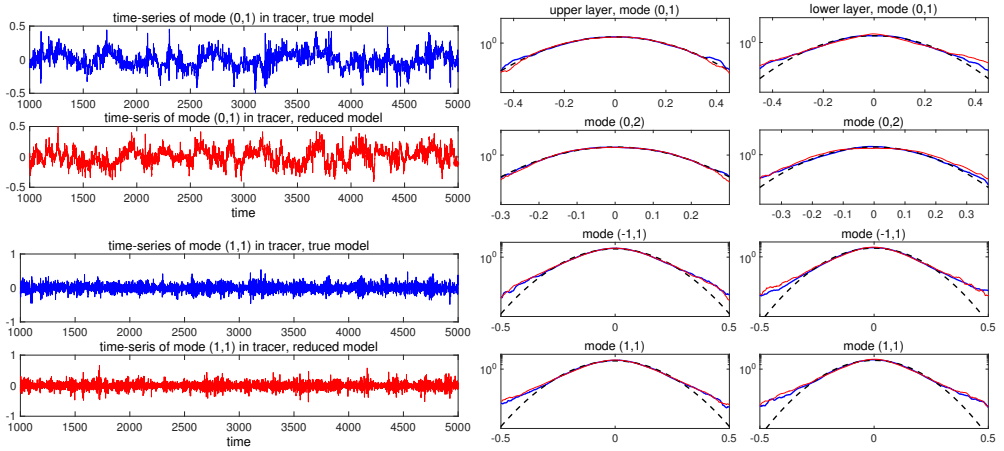


Fig. 4.14: *Prediction of tracer intermittency in mid latitude atmosphere regime with parameters* $(\epsilon^{-1}, d_T) = (5, 0.1)$. *The left panel is the time-series for two leading modes (0,1) and (1,1) between true model and reduced model results; the right panel compares the PDFs in the first four modes between the truth in blue and reduced model prediction in red with the Gaussian fit in dashed black lines.*

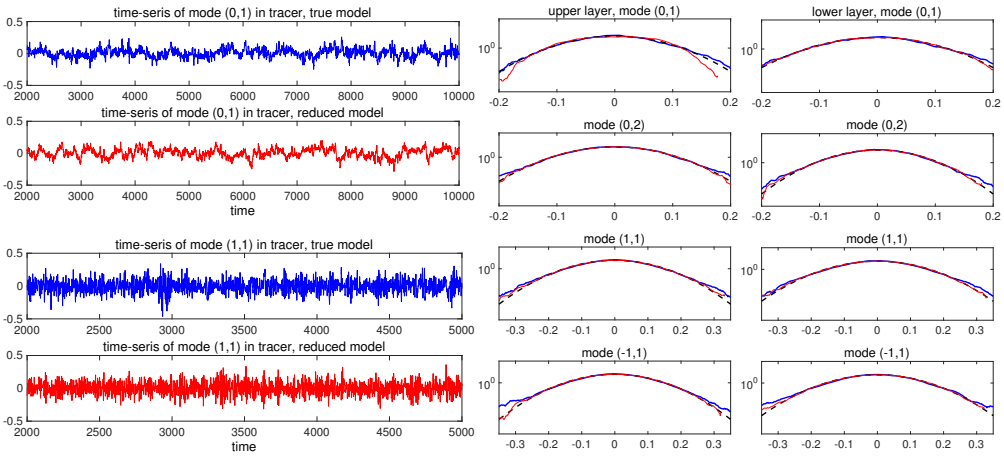
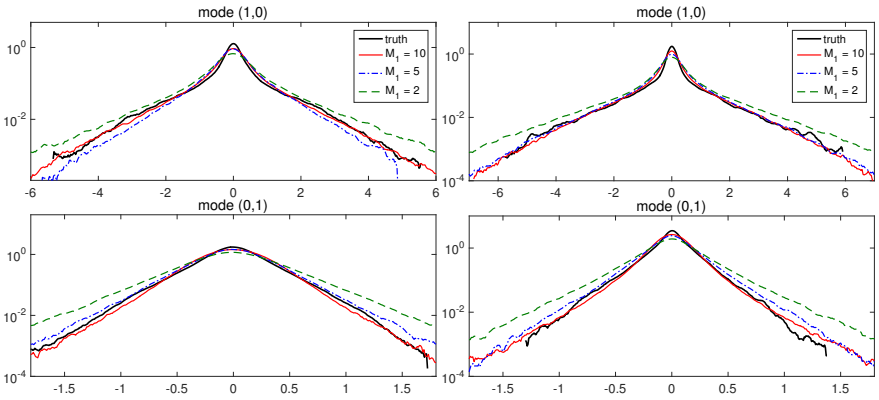


Fig. 4.15: *Prediction of tracer intermittency in mid latitude atmosphere regime with parameters* $(\epsilon^{-1}, d_T) = (1, 0.1)$. *The left panel is the time-series for two leading modes (0,1) and (1,1) between true model and reduced model results; the right panel compares the PDFs in the first four modes between the truth in blue and reduced model prediction in red with the Gaussian fit in dashed black lines.*

Above $\langle \mathbf{v}' \cdot \nabla T' \rangle_{\mathbf{k}}$ is the *eddy diffusivity* for the tracer dynamics. This is a spatially non-local term involving the interactions between all the spectral modes among all scales. In the turbulent-viscosity hypothesis [19, 29], the tracer turbulent advection is represented by a diffusion term as

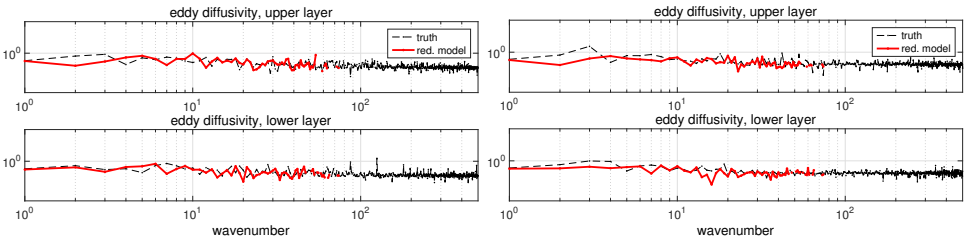
$$\langle \mathbf{v}' \cdot \nabla T' \rangle = -\nabla \cdot \left(\kappa_T \nabla \bar{T} \right), \tag{4.3}$$

where the coefficient κ_T is a constant "eddy diffusivity". The turbulent-viscosity hypothesis implies that the scalar flux vector is aligned with the mean scalar gradient vector. It is introduced through a homogenization about the turbulent system.



(a) high lat. atmosphere (b) high lat. ocean

Fig. 4.16: *Tracer PDFs in the first two leading modes (1,0) and (0,1) with different truncation* $M_1 = 10,5,2$ *in computing the advection velocity* $\mathbf{v}_M$*. The truth in black is compared with the reduced-order model predictions with different reduction in the nonlinear advection term.*



(a) $(\epsilon^{-1}, d_T) = (5, 0.1)$ (b) $(\epsilon^{-1}, d_T) = (1, 0.1)$

Fig. 4.17: *Eddy diffusivity calculated from equation (4.4) for the mid latitude atmosphere regime with different parameters. The modes are ordered according to the descending order of the advection flow energy. The true model results are shown in dashed black lines, while the reduced-order model predictions are in red.*

Here we want to compare the eddy diffusivity from the truth with the reduced-order approximations. From equation (4.3) we can calculate the eddy diffusivity coefficients for both the true system and reduced-order models through each spectral mode solution

$$\kappa_T = \frac{\langle \mathbf{v}' \cdot \nabla T' \rangle_{\mathbf{k}}}{|\mathbf{k}|^2 \bar{T}_{\mathbf{k}}}, \quad \kappa_T^M = \frac{\langle \mathbf{v}'_M \cdot \nabla T'_M \rangle_{\mathbf{k}}}{|\mathbf{k}|^2 \bar{T}_{M,\mathbf{k}}}. \tag{4.4}$$

The statistics are achieved through time averaging along solution trajectories in both true system and reduced stochastic model. Note that in the reduced-order models, only the first 10 leading modes are resolved, thus $\langle \mathbf{v}'_M \cdot \nabla T'_M \rangle_{\mathbf{k}}$ are calculated based on these available resolved modes only. Also, especially in homogeneous case, the statistical mean states of the tracer modes become small. Thus equation (4.4) can only qualitatively describe the mixing rate in each mode.

Figure 4.17 shows the results for eddy diffusivity estimated from equation (4.4) for mid latitude atmosphere $F = 4$ with jets. The modes are ordered in the same descending

order according to the equilibrium statistical energy spectra of the vorticity as before. From the black lines for the true model eddy diffusivity, beginning from the intermediate scales from wavenumber around 10, the numerical estimation about eddy diffusivity κ_T stays in nearly constant values with small fluctuations. This illustrates that turbulent-viscosity hypothesis is valid in tracer intermediate to small scales, and can actually offer good approximation in the eddy-diffusivity homogenization as in equation (4.3). In the reduced-order predictions with model errors, eddy diffusivity approximation κ_T^M has agreement in amplitude along the spectrum except the largest scales, showing that the most important fluctuation interactions in the reduced-order approximations are modeled with accuracy. In the largest scale modes, larger errors appear and the eddy diffusivity approximation gets violated. This is related with the highly anisotropic structure in the largest scale representing the zonal jets and the large errors in the approximation for the mean state with large fluctuations. The important practical issue of model error from eddy diffusivity approximation in the tracer statistics is an important research area worth further investigation in the future.

5. Concluding discussion

In this paper we developed reduced-order stochastic models for predicting extreme events and intermittency in passive scalar turbulence. The turbulent flow field is generated by a two-layer quasi-geostrophic model containing non-Gaussian and anistropic structures that can be observed in many realistic natural flows. The passive tracer is transported by this advection flow and forced by a mean tracer gradient. Various representative statistical features in the passive tracer field can be generated through this system from near Gaussian distributions to highly intermittent structures with fat-tailed PDFs in the leading tracer modes. This passive tracer turbulence advected by the complex high dimensional flow can be compared with the much simpler model using the L-96 system as the advection flow investigated in [26], which shares many similarities in statistical structures with the QG flow but gets much simpler dynamical equations (in fact, the $F = 5$ case in L-96 model has strong non-Gaussian statistics with leading dominant mode as the atmosphere regime with jets, and the $F = 8$ case in L-96 model can be viewed as a simplification of the high latitude ocean with near-Gaussian homogeneous statistics). Therefore the model reduction idea in [26] can be used as a guideline for the model development here under more complex setup.

The complex strongly turbulent dynamical system requires proper reduced-order modeling strategy of adopting simple advection flow models especially in situations when realistic high dimensional applications apply, but still maintaining the skill in recovering the most important statistical features in the tracer field. Linear Gaussian stochastic model that is block-diagonal is then introduced to approximate the background flow field, and the tracer equations with advection terms using only the resolved large-scale modes are applied for efficient prediction. Calibration of the imperfect models with error is then required to achieve the optimal performance. First linear damping and random noise corrections are introduced in the reduced-order model to replace the expensive higher order nonlinear interactions in the original flow equations. This is from a stochastic realization of the statistical closure models in [21,28]. High order feedbacks from small scales can be approximated by low order moments in equilibrium, thus the efficiency in calculating the imperfect model parameters is guaranteed. Second the model prediction skill is further improved through a systematic framework by measuring the model autocorrelation functions in the spectral representation with the information metric [26]. The mixing time scales in the leading modes can be modeled with accuracy. The skill of imperfect models in capturing the crucial large-scale structures in tracer

statistics is tested among a wide variety of dynamical regimes ranging from strongly non-Gaussian to near Gaussian statistics with strong mixing. The reduced-order stochastic model displays uniform skill in capturing both tracer energy spectra in equilibrium and fat-tails in the leading resolved modes, as well as a good approximation for the tracer eddy diffusivity in the atmosphere regime with jets. Besides, the skillful linear Gaussian stochastic modeling algorithm developed here should also be useful for other applications such as accurate forecast of mean responses to changes in forcing [18] and efficient algorithms for state estimation or data assimilation [10].

Acknowledgments. This research of Andrew Majda is partially supported by the Office of Naval Research through MURI N00014-16-1-2161 and DARPA through W911NF-15-1-0636. Di Qi is supported as a graduate research assistant on these grants.

Appendix A. Information criterion for measuring autocorrelation functions of the stationary random fields.

A.1. Spectral representation of general stationary random fields. For a scalar stationary process $u(t)$, its autocorrelation function $\mathcal{R}(t)=\langle u(0)u(t)^*\rangle$ is positive-definite in time, that is,

$$\sum_{i,j=1}^{N}\mathcal{R}(t_i-t_j)a_i a_j^*\geq 0,\quad \forall\{a_1,\cdots a_N\}\in\mathbb{C}^N.$$

Thus it is proved by Khinchin's formula [35] that the autocorrelation function has a spectral representation

$$\mathcal{R}(t)=\int_{-\infty}^{\infty}e^{i\lambda t}dF(\lambda)=\int_{-\infty}^{\infty}e^{i\lambda t}E(\lambda)d\lambda, \tag{A.1}$$

where $F(\lambda)$ is the non-decreasing *spectral distribution function* of the stationary process $u(t)$, and $E(\lambda)>0$ is the *spectral density function* of the stationary process $u(t)$. According to the formula (A.1), the spectral distribution function $F(\lambda)$ of the process $u(t)$ can be determined from its autocorrelation function $\mathcal{R}(t)$ via Fourier transform. Conversely, the autocorrelation function can be calculated from the corresponding spectral distribution function. Therefore we can construct the spectral representation of the stationary random field $u(t)$ as

$$u(t)=\int_{-\infty}^{\infty}e^{i\lambda t}\hat{Z}(d\lambda), \tag{A.2}$$

with $\hat{Z}(d\lambda)$ a random measure which can be defined in a similar form of Fourier transform as

$$\hat{Z}(d\lambda)=\lim_{T\to\infty}\frac{d\lambda}{2T}\int_{-T}^{T}e^{-i\lambda t}u(t)dt.$$

By combining formulas (A.1) and (A.2) we can find that $E(\lambda)$ or $dF(\lambda)$ can be viewed as the energy spectrum of the random field u which measures the variance in $\hat{Z}(d\lambda)$

$$dF(\lambda)=E(\lambda)d\lambda=\mathbb{E}\left|\hat{Z}(d\lambda)\right|^2. \tag{A.3}$$

Note here that in general $\hat{Z}(d\lambda)$, which depends on the statistics of the stationary process $u(t)$, is not necessarily a Gaussian random variable and may include higher

order statistical information. But in the main text we only concern fitting a linear Gaussian process with consistent statistics up to second order, so a simple Gaussian approximation $u_G(t)$ with statistics up to second order moments is sufficient. The true model autocorrelation function $\mathcal{R}(t)$ and corresponding spectral density $E(\lambda)$ can be achieved directly through the Fourier transform (A.1) based on true observation of the data. As the discussion in the main text, we use an Ornstein–Uhlenbeck process to approximate the flow model, that is,

$$du_M = -(\gamma + i\omega)\, u_M dt + \sigma dW.$$

The autocorrelation function and corresponding spectral density function can be calculated in exact forms,

$$\mathcal{R}_M(t) = \exp(-(\gamma+i\omega)t), \quad E_M(\lambda) = \int_{-\infty}^{\infty} e^{-i\lambda t}\mathcal{R}(t)\,dt = \frac{2\gamma}{\gamma^2 + (\lambda+\omega)^2}. \tag{A.4}$$

A.2. Spectral information criterion for measuring the autocorrelation functions in two-layer model. In the linear stochastic model (2.6) of advection flow potential vorticity $\vec{q}_M(t)$, one additional difference is that we are considering a 2-vector field. Still this can be easily reduced to the previous scalar case. In fact in the stochastic approximation model, model (2.6) becomes a two-dimensional Ornstein–Uhlenbeck process for each vorticity spectral mode $\vec{q}_M = (\hat{q}_1, \hat{q}_2)^T$ including the upper and lower layer modes

$$d\vec{q}_M = -(\Gamma + i\Omega)\,\vec{q}_M + \Sigma d\vec{W},$$

where Γ, Ω, Σ are 2×2 matrices and we neglect the subscripts in the state variables. Explicitly in the two-layer QG model we have

$$\Gamma + i\Omega = \begin{bmatrix} 0 & \\ & r|\mathbf{k}|^2 \end{bmatrix} H_{\mathbf{k}}^{-1} + \nu|\mathbf{k}|^{2s} I_2 + D_{q,\mathbf{k}}^M + ik_x \begin{bmatrix} U & \\ & -U \end{bmatrix} + ik_x \begin{bmatrix} \beta + k_d^2 U & \\ & \beta - k_d^2 U \end{bmatrix} H_{\mathbf{k}}^{-1}, \Sigma = \Sigma_{q,\mathbf{k}}^M,$$

with the original linear operators (2.6) and additional imperfect model corrections (3.5) combined together. The autocorrelation function about the above system becomes the 2×2 matrix

$$\mathcal{R}_M(t,t') \equiv \left\langle \vec{q}_M(t)\, \vec{q}_M(t')^* \right\rangle = \int_{t_0}^{t\wedge t'} e^{-(t-s)(\Gamma+i\Omega)} \Sigma^2 e^{-(\Gamma - i\Omega)(t'-s)}\, ds. \tag{A.5}$$

In each entry of the above matrix, the integrant is still an exponential about time; then in stationary state the autocorrelation function is only dependent on the time difference $|t'-t|$. Thus each entry of the above autocorrelation matrix can be reduced to the exact form of the scalar case (A.4). The corresponding spectral density $E_M(\lambda)$ is the 2×2 matrix that can be defined componentwise as in equation (A.4).

The true process constraint in Gaussian statistics, $\vec{q}_G$, and the imperfect linear model approximation, $\vec{q}_M(t)$, can be then naturally expressed under an independent 2-dimensional Wiener random measure $d\vec{\hat{W}}(\lambda) = \left(d\hat{W}_1, d\hat{W}_2\right)$. That is,

$$\vec{q}_G(t) = \int_{-\infty}^{\infty} e^{i\lambda t} E^{1/2}(\lambda)\, d\vec{\hat{W}}(\lambda), \tag{A.6}$$

$$\vec{q}_M(t) = \int_{-\infty}^{\infty} e^{i\lambda t} E_M^{1/2}(\lambda)\, d\vec{\hat{W}}_M(\lambda). \tag{A.7}$$

Above $E(\lambda)$ is the true spectral energy, while $E_M(\lambda)$ is the explicit form for the density of autocorrelation function for the linear model from equation (A.5). Instead of measuring the information distance between the stationary random fields $\vec{q}_G(t)$ and $\vec{q}_M(t)$ at fixed time t, a more favorable way is to compare the distributions between the true spectral random measure and the imperfect model approximation field which can include the information about the autocorrelation functions. With the help of the spectral decomposition of stationary random process, we decompose the state variables into the spectral forms in equations (A.6) and (A.7) respectively under each Fourier basis $\{e^{i\lambda t}\}$ in time and measure the information distance between one single spectral increment at frequency λ. Here consider the full spectra of random measures

$$p_G = \prod_{\lambda} p_G(x;\lambda) \sim \left\{ E^{1/2}(\lambda)\, d\vec{\hat{W}}(\lambda) \right\}_{\lambda=-\infty}^{\infty}$$

and

$$p_M = \prod_{\lambda} p_M(x;\lambda) \sim \left\{ E_M^{1/2}(\lambda)\, d\vec{\hat{W}}_M(\lambda) \right\}_{\lambda=-\infty}^{\infty}$$

as two random processes about $\lambda \in \mathbb{R}$. Above '$\sim$' means that the probability distribution $p(x;\lambda)$ has the random process representation $E^{1/2}(\lambda)\, d\vec{\hat{W}}(\lambda)$ at each frequency λ. Due to the independent increments between $\vec{\hat{W}}(d\lambda_1)$ and $\vec{\hat{W}}(d\lambda_2)$, the information distance between these two random processes can be expressed as the integration between all the spectral modes [26,35]

$$\begin{aligned} \mathcal{P}(p_G, p_M) &= \mathcal{P}\left(\prod_{\lambda} p_G(x;\lambda), \prod_{\lambda} p_M(x;\lambda) \right) \\ &= \int_{-\infty}^{\infty} d\lambda \mathcal{P}(p_G(x;\lambda), p_M(x;\lambda)) \\ &= \int d\lambda \mathcal{D}\left(E(\lambda) E_M^{-1}(\lambda) \right), \end{aligned} \tag{A.8}$$

where $\mathcal{D}(x) = -\log\det x + \mathrm{tr}x - N$ is the Gaussian relative entropy with a zero mean state.

Appendix B. Reduced-order statistical closure model for advection flow equation.

We provide more details here about the derivation of the statistical closure model used in equations (3.3) and (3.4), which is developed in [21,27] for general systems with energy conservation principle. The statistical energy mechanism of the system including high-order nonlinear interactions can be illustrated by the dynamical equation of covariance matrix of flow potential vorticity by taking ensemble average of the fluctuation equations (2.4b), so that,

$$\frac{dR_{\mathbf{k}}^q}{dt} + \mathcal{L}_{\mathbf{k}}(\bar{q}) R_{\mathbf{k}}^q + R_{\mathbf{k}}^q \mathcal{L}_{\mathbf{k}}^*(\bar{q}) + Q_F^q = (\mathcal{L}_{\mathbf{k}}^q + \mathcal{D}_{\mathbf{k}}^q) R_{\mathbf{k}}^q + R_{\mathbf{k}}^q (\mathcal{L}_{\mathbf{k}}^q + \mathcal{D}_{\mathbf{k}}^q)^*, \quad |\mathbf{k}| \le N. \tag{B.1}$$

The 2×2 covariance matrix $R_{\mathbf{k}}^q$ in one wavenumber

$$R_{\mathbf{k}}^q = \langle \vec{q}_{\mathbf{k}} \vec{q}_{\mathbf{k}}^* \rangle = \begin{bmatrix} \left\langle |q_{1,\mathbf{k}}|^2 \right\rangle & \left\langle q_{1,\mathbf{k}}^* q_{2,\mathbf{k}} \right\rangle \\ \left\langle q_{1,\mathbf{k}} q_{2,\mathbf{k}}^* \right\rangle & \left\langle |q_{2,\mathbf{k}}|^2 \right\rangle \end{bmatrix}, \tag{B.2}$$

contains the variances of potential vorticity in upper and lower layers in diagonal entries, and the exchange of energy between layers in the off-diagonal components. On the right hand side of the covariance equation (B.1), linear operators $(\mathcal{L}^q, \mathcal{D}^q)$ are due to the skew-symmetric dispersion and dissipation effects. The dispersion operator $\mathcal{L}^q_{\mathbf{k}}$ is due to the β-effect from the rotation and the large-scale zonal mean flow $\vec{U}$, and the dissipation operator $\mathcal{D}^q_{\mathbf{k}}$ is from the Ekman bottom friction together with the hyperviscosity from the right hand side of the original flow equation (2.4b). $\mathcal{L}_{\mathbf{k}}(\bar{q})$ represents the interactions with a non-zero statistical mean state $\bar{q}$, implying internal instability with positive growth rate. Q^q_F represents the nonlinear interactions between different spectral modes, which play the crucial role as a nonlinear flux balancing the internal instabilities, so that higher order moments are introduced

$$Q^q_F(\vec{q}_{\mathbf{k}}) = \frac{1}{2}\sum_{\mathbf{m}+\mathbf{n}=\mathbf{k}} \langle (A_{\mathbf{km}}\vec{q}_{\mathbf{m}} \circ \vec{q}_{\mathbf{n}} + A_{\mathbf{kn}}\vec{q}_{\mathbf{n}} \circ \vec{q}_{\mathbf{m}})\, \vec{q}^*_{\mathbf{k}} \rangle .$$

Third-order moments $\langle \vec{q}_{\mathbf{m}} \vec{q}_{\mathbf{n}} \vec{q}^*_{\mathbf{k}} \rangle$ between the triad modes $\mathbf{m}+\mathbf{n}=\mathbf{k}$ are introduced between small and large scale modes in the second-order dynamics. First of all, it is important to notice the conservation of the total nonlinear flux $\mathrm{tr} Q^q_F \equiv 0$. This equality implies that the nonlinear interactions will not introduce additional energy source or sink into the system. Second, the nonlinear flux Q^q_F is crucial in the energy mechanism that balances the unstable directions due to internal instability from the linear operators.

In general in the statistical dynamics, the nonlinear flux Q^q_F describes the energy transfer from the unstable subspace to the stable one through higher-order interactions. The low-order approximation of this nonlinear flux is through additional damping and noise by splitting this operator into two separate components, $Q^q_M = Q^{q,+}_M + Q^{q,-}_M$. The low-order correction made in Q^q_M is only constructed from the first two order of statistics. The task is to propose proper forms to calibrate the nonlinear flux forms using additional damping in $Q^{q,-}_M$ and additional noise in $Q^{q,+}_M$.

Comparing with the true statistical equation (B.1), the reduced-order approximation is equivalent to replacing the nonlinear flux Q^q_F with the judiciously calibrated damping and noise $\left(D^M_{q,\mathbf{k}}, \Sigma^M_{q,\mathbf{k}}\right)$ in consideration with both the equilibrium energy transfer mechanism and further sensitivity correction, explicitly that is,

$$-\frac{1}{2}\sum_{\mathbf{m}+\mathbf{n}=\mathbf{k}} (A_{\mathbf{km}}\vec{q}_{\mathbf{m}} \circ \vec{q}_{\mathbf{n}} + A_{\mathbf{kn}}\vec{q}_{\mathbf{n}} \circ \vec{q}_{\mathbf{m}}) \to -D^M_{q,\mathbf{k}}\vec{q}_{M,\mathbf{k}}dt + \Sigma^M_{q,\mathbf{k}}d\vec{W}_{q,\mathbf{k}}.$$

Therefore the statistical nonlinear flux is approximated by the corresponding statistical damping and noise in the statistical equation

$$Q^q_{F,\mathbf{k}} \to -D^M_{q,\mathbf{k}}R^q_{\mathbf{k}} - R^q_{\mathbf{k}}D^{M*}_{q,\mathbf{k}} + \left(\Sigma^M_{q,\mathbf{k}}\right)^2 .$$

The reduced-order closure form consists of two indispensable components:

i) *higher-order corrections from equilibrium statistics*: in the first part of the correction, unperturbed equilibrium statistics in the nonlinear flux are used to calibrate the higher-order moments as additional energy sink and source. The true equilibrium higher-order flux can be calculated without error from equilibrium first and second order moments $\left(\bar{q}_{\mathrm{eq}}, R^q_{\mathrm{eq}}\right)$ from the true statistical dynamics (B.1) with $dR^q_{\mathbf{k}}/dt \equiv 0$ in steady state

$$Q^q_{F,\mathrm{eq}} = (\mathcal{L}^q_{\mathbf{k}} + \mathcal{D}^q_{\mathbf{k}} - \mathcal{L}_{\mathbf{k}}(\bar{q}_{\mathrm{eq}}))\, R^q_{\mathbf{k},\mathrm{eq}} + R^q_{\mathbf{k},\mathrm{eq}} (\mathcal{L}^q_{\mathbf{k}} + \mathcal{D}^q_{\mathbf{k}} - \mathcal{L}_{\mathbf{k}}(\bar{q}_{\mathrm{eq}}))^* .$$

The linear damping and Gaussian random noise correction can be first introduced from the equilibrium nonlinear flux $Q_{F,\mathrm{eq}}$

$$D^{\mathrm{eq}}_{q,\mathbf{k}} = -\frac{1}{2} Q^{q,-}_{F,\mathrm{eq},\mathbf{k}} \left(R^{q}_{\mathbf{k},\mathrm{eq}}\right)^{-1}, \quad \Sigma^{\mathrm{eq}}_{q,\mathbf{k}} = \left(Q^{q,+}_{F,\mathrm{eq},\mathbf{k}}\right)^{1/2}, \tag{B.3}$$

where the equilibrium high-order nonlinear flux is decomposed into a positive and a negative definite component by

$$Q^{q}_{F,\mathrm{eq}} = P^{*}\Lambda P, \qquad Q^{q,-}_{F,\mathrm{eq}} = P^{*}\Lambda^{-} P, \quad Q^{q,+}_{F,\mathrm{eq}} = P^{*}\Lambda^{+} P,$$

with $\Lambda^{+} = \max\{0,\Lambda\}$ and $\Lambda^{-} = \min\{0,\Lambda\}$. The damping operator $D^{\mathrm{eq}}_{q,\mathbf{k}}$ is from the negative component normalized by the equilibrium covariance $R^{q}_{\mathbf{k},\mathrm{eq}}$, and the noise is directly from the positive part as a Gaussian process. In this way, only the minimum amount of corrections to stabilize the system with consistent equilibrium statistics for the mean and variance is introduced in the above terms.

ii) *additional damping and noise to model nonlinear flux*: the above closure by using equilibrium information for nonlinear flux is not sufficient for accurate prediction in the reduced-order methods since the scheme is only marginally stable and the energy transferring mechanism may change with large deviation from the equilibrium case when external perturbations are applied. As a further correction for the noise and damping, we propose additional terms on top of the form (B.3) with a simple constant damping for all the spectral modes and an additional noise accordingly to make sure consistency in energy

$$Q^{\mathrm{add}}_{M,\mathbf{k}} = -D^{\mathrm{add}}_{M} R_{M,\mathbf{k}} + \left(\Sigma^{\mathrm{add}}_{M,\mathbf{k}}\right)^{2}, \quad D^{\mathrm{add}}_{M} = \begin{bmatrix} d_M + i\omega_M & \\ & d_M - i\omega_M \end{bmatrix}. \tag{B.4}$$

This further additional damping term in equation (B.4) is aimed to offer stabilizing effects in the marginal stable equilibrium form (B.3), and the additional noise to offer further corrections in modeling the transient state and autocorrelation function that is important for the mixing rate in each spectral mode. Notice that the nonlinear energy transferring mechanism may change with large deviation from the equilibrium case when intermittent fluctuations are present.

Combining the ideas in equations (B.3) and (B.4), The additional damping and noise corrections for the reduced-order flow potential vorticity model become the following form

$$D^{M}_{q,\mathbf{k}} = -\frac{1}{2} Q^{q,-}_{F,\mathrm{eq},\mathbf{k}} \left(R^{q}_{\mathbf{k},\mathrm{eq}}\right)^{-1} + D^{\mathrm{add}}_{M}, \quad \Sigma^{M}_{q,\mathbf{k}} = \left(Q^{q,+}_{F,\mathrm{eq},\mathbf{k}} + \left(\Sigma^{\mathrm{add}}_{M,\mathbf{k}}\right)^{2}\right)^{1/2}. \tag{B.5}$$

Above the parameters $\left(D^{M}_{q,\mathbf{k}}, \Sigma^{M}_{q,\mathbf{k}}\right)$ are 2×2 matrices applied on the two-layer spectral mode $\vec{q}_{\mathbf{k}} = (q_{1,\mathbf{k}}, q_{2,\mathbf{k}})^{T}$.

REFERENCES

[1] M. Avellaneda and A.J. Majda, *Simple examples with features of renormalization for turbulent transport*, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 346:205–233, 1994.

[2] P. Bartello and G. Holloway, *Passive scalar transport in β-plane turbulence*, Journal of Fluid Mechanics, 223:521–536, 1991.

[3] A. Bourlioux and A. Majda, *Elementary models with probability distribution function intermittency for passive scalars with a mean gradient*, Physics of Fluids (1994-present), 14:881–897, 2002.
[4] J.C. Bronski and R.M. McLaughlin, *The problem of moments and the Majda model for scalar intermittency*, Physics Letters A, 265:257–263, 2000.
[5] R. Camassa, N. Martinsen-Burrell, and R.M. McLaughlin, *Dynamics of probability density functions for decaying passive scalars in periodic velocity fields*, Physics of Fluids, 19:117104, 2007.
[6] T. DelSole, *Stochastic models of quasigeostrophic turbulence*, Surveys in Geophysics, 25:107–149, 2004.
[7] T. DelSole and B.F. Farrell, *A stochastically excited linear system as a model for quasigeostrophic turbulence: Analytic results for one-and two-layer fluids*, Journal of the Atmospheric Sciences, 52:2531–2547, 1995.
[8] G. Falkovich, K. Gawędzki, and M. Vergassola, *Particles and fields in fluid turbulence*, Reviews of modern Physics, 73:913–975, 2001.
[9] D.M. Frierson, *Robust increases in midlatitude static stability in simulations of global warming*, Geophysical Research Letters, 33(24), 2006.
[10] B. Gershgorin and A.J. Majda, *Filtering a statistically exactly solvable test model for turbulent tracers from partial observations*, J Comput. Phys., 230:1602–1638, 2011.
[11] J. Gould, D. Roemmich, S. Wijffels, H. Freeland, M. Ignaszewsky, X. Jianping, S. Pouliquen, Y. Desaubies, U. Send, K. Radhakrishnan, et al., *Argo profiling floats bring new era of in situ ocean observations*, Eos, 85(19):185–190, 2004.
[12] J. Harlim, A. Mahdi, and A.J. Majda, *An ensemble kalman filter for statistical estimation of physics constrained nonlinear regression models*, J Comput. Phys., 257:782–812, 2014.
[13] R. Kleeman and A. Majda, *Predictability in a model of geophysical turbulence*, Journal of the Atmospheric Sciences, 62:2864–2879, 2005.
[14] S. Kullback and R.A. Leibler, *On information and sufficiency*, The Annals of Mathematical Statistics, 22:79–86, 1951.
[15] Y. Lee, A.J. Majda, and D. Qi, *Stochastic superparameterization and multiscale filtering of turbulent tracers*, Multiscale Modeling & Simulation, 15:215–234, 2017.
[16] A. Majda, R.V. Abramov, and M.J. Grote, *Information Theory and Stochastics for Multiscale Nonlinear Systems*, Amer. Math. Soc., 25, 2005.
[17] A.J. Majda and B. Gershgorin, *Elementary models for turbulent diffusion with complex physical features: eddy diffusivity, spectrum and intermittency*, Phil. Trans. R. Soc. A, 371:2012.0184, 2013.
[18] A.J. Majda, B. Gershgorin, and Y. Yuan, *Low-frequency climate response and fluctuation-dissipation theorems: theory and practice*, Journal of the Atmospheric Sciences, 67:1186–1201, 2010.
[19] A.J. Majda and P.R. Kramer, *Simplified models for turbulent diffusion: theory, numerical modelling, and physical phenomena*, Physics Reports, 314:237–574, 1999.
[20] A.J. Majda and D. Qi, *Improving prediction skill of imperfect turbulent models through statistical response and information theory*, Journal of Nonlinear Science, 26:233–285, 2016.
[21] A.J. Majda and D. Qi, *Strategies for reduced-order models for predicting the statistical responses and uncertainty quantification in complex turbulent dynamical systems*, submitted to SIAM Review, 2016.
[22] A.J. Majda and X.T. Tong, *Intermittency in turbulent diffusion models with a mean gradient*, Nonlinearity, 28:4171–4208, 2015.
[23] M.A. Mohamad and T.P. Sapsis, *Probabilistic description of extreme events in intermittently unstable dynamical systems excited by correlated stochastic processes*, SIAM/ASA Journal on Uncertainty Quantification, 3:709–736, 2015.
[24] J.D. Neelin, B.R. Lintner, B. Tian, Q. Li, L. Zhang, P.K. Patra, M.T. Chahine, and S.N. Stechmann, *Long tails in deep columns of natural and anthropogenic tropospheric tracers*, Geophysical Research Letters, 37(5), 2010.
[25] S.B. Pope, *Turbulent Flows*, IOP Publishing, Bristol, 2001.
[26] D. Qi and A.J. Majda, *Predicting fat-tailed intermittent probability distributions in passive scalar turbulence with imperfect models through empirical information theory*, Comm. Math. Sci, 14(6):1687–1722, 2016.
[27] D. Qi and A.J. Majda, *Low-dimensional reduced-order models for statistical response and uncertainty quantification: barotropic turbulence with topography*, Physica D: Nonlinear Phenomena, 343(15):7–27, 2017.
[28] D. Qi and A.J. Majda, *Low-dimensional reduced-order models for statistical response and uncertainty quantification: two-layer baroclinic turbulence*, Journal of the Atmospheric Sciences,

73:4609–4639, 2016.
[29] R. Salmon, *Lectures on Geophysical Fluid Dynamics*, Oxford University Press, 1998.
[30] T.P. Sapsis and A.J. Majda, *Statistically accurate low-order models for uncertainty quantification in turbulent dynamical systems*, Proceedings of the National Academy of Sciences, 110:13705–13710, 2013.
[31] K.S. Smith, *Tracer transport along and across coherent jets in two-dimensional turbulent flow*, Journal of Fluid Mechanics, 544:133–142, 2005.
[32] A.F. Thompson and W.R. Young, *Two-layer baroclinic eddy heat fluxes: Zonal flows and energy balance*, Journal of the Atmospheric Sciences, 64:3214–3231, 2007.
[33] G.K. Vallis, *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*, Cambridge University Press, 2006.
[34] E. Vanden Eijnden, *Non-gaussian invariant measures for the Majda model of decaying turbulent transport*, Communications on Pure and Applied Mathematics, 54:1146–1167, 2001.
[35] A.M. Yaglom, *An introduction to the theory of stationary random functions*, Courier Corporation, 2004.
[36] P. Zurita-Gotor, J. Blanco-Fuentes, and E.P. Gerber, *The impact of baroclinic eddy feedback on the persistence of jet variability in the two-layer model*, Journal of the Atmospheric Sciences, 71:410–429, 2014.