

Non-gaussian Test Models for Prediction and State Estimation with Model Errors*

Michał BRANICKI¹ Nan CHEN² Andrew J. MAJDA²

(In honor of the scientific heritage of Jacques-Louis Lions)

Abstract Turbulent dynamical systems involve dynamics with both a large dimensional phase space and a large number of positive Lyapunov exponents. Such systems are ubiquitous in applications in contemporary science and engineering where the statistical ensemble prediction and the real time filtering/state estimation are needed despite the underlying complexity of the system. Statistically exactly solvable test models have a crucial role to provide firm mathematical underpinning or new algorithms for vastly more complex scientific phenomena. Here, a class of statistically exactly solvable non-Gaussian test models is introduced, where a generalized Feynman-Kac formulation reduces the exact behavior of conditional statistical moments to the solution to inhomogeneous Fokker-Planck equations modified by linear lower order coupling and source terms. This procedure is applied to a test model with hidden instabilities and is combined with information theory to address two important issues in the contemporary statistical prediction of turbulent dynamical systems: the coarse-grained ensemble prediction in a perfect model and the improving long range forecasting in imperfect models. The models discussed here should be useful for many other applications and algorithms for the real time prediction and the state estimation.

Keywords Prediction, Model error, Information theory, Feynman-Kac framework, Fokker planck, Turbulent dynamical systems

2000 MR Subject Classification 60G25, 60H10, 60H30, 82C31, 94A15

1 Introduction

Turbulent dynamical systems involve dynamics with both a large dimensional phase space and a large number of positive Lyapunov exponents. Such extremely complex systems are ubiquitous in many disciplines of contemporary science and engineering such as climate-atmosphere-ocean science, neural science, material science, and engineering turbulence. Wide contemporary interest topics involve the statistical ensemble prediction (see [31]) and the real time state esti-

Manuscript received August 31, 2012.

¹Corresponding author. Department of Mathematics and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, USA.

E-mail: branicki@cims.nyu.edu

²Department of Mathematics and Center for Atmosphere Ocean Science, Courant Institute of Mathematical Sciences, New York University, New York, USA.

E-mail: jonjon@cims.nyu.edu chennan@cims.nyu.edu

*Project supported by the Office of Naval Research (ONR) Grants (No. ONR DRI N00014-10-1-0554) and the DOD-MURI award “Physics Constrained Stochastic-Statistical Models for Extended Range Environmental Prediction”.

mation/filtering (see [34]) for the extremely complex systems while coping with the fundamental limitations of model error and the curse of small ensemble size (see [22]).

An important role of mathematics in applied sciences is to develop simple and accurate (or easily solvable) test models with unambiguous mathematical features which nevertheless capture crucial features of vastly more complex systems in science and engineering. Such models provide the firm underpinning for both the advancing scientific understanding and the developing new numerical or statistical understanding. One of the authors developed this approach with various collaborators over the past few years for paradigm problems for turbulent dynamical systems. For example, simple statistically exactly solvable test models were developed for slow-fast systems (see [12–13]), turbulent tracers (see [2, 14, 21, 33]) and as stochastic parameterization algorithms for the real time filtering of turbulent dynamical systems with the judicious model error (see [8–9, 15, 34–35]). Such models were utilized as unambiguous test models for improving prediction with imperfect models in climate science through the empirical information theory (see [5, 10, 28–30]) and for testing algorithms for uncertainty quantification (see [4–5, 25]).

Here, we study non-Gaussian statistics in a class of test models which are statistically exactly solvable through a generalized Feynman-Kac formula (see [16, 21]) which reduces the exact behavior of conditional statistical moments to the solution to inhomogeneous Fokker-Planck equations modified by linear lower-order coupling terms and source terms. This exact procedure is developed in Section 2 below and involves only the marginal averaging and the integration by parts. In Section 3, elementary test models are introduced where the general procedure from Section 2 can be evaluated through elementary numerical solutions to the coupled generalized Fokker-Planck equations (CGFPE). Section 4 contains a brief introduction to the use of information theory to the quantify model error in a framework adapted to the present context. Section 5 contains two applications of the material in Sections 3–4 to the statistical ensemble forecasting: the first application involves the coarse-grained ensemble prediction in a perfect model with hidden instabilities; the second application involves the use of imperfect models for the long range forecasting.

2 Test Models with Exactly Solvable Conditional Moments

We consider a special class of test models and illustrate that the evolution of the exact conditional statistical moments can be calculated through the solution to coupled generalized Fokker-Planck equations (CGFPE). Our elementary derivation follows the philosophy of the generalized Feynman-Kac framework (see [16, 21]) although we do not know any specific reference for the general principle developed below.

Consider a vector $\mathbf{u} \in \mathbb{R}^M$ partitioned into components $\mathbf{u} = (\mathbf{u}_I, \mathbf{u}_{II})$ with $\mathbf{u}_I \in \mathbb{R}^{M_I}$, $\mathbf{u}_{II} \in \mathbb{R}^{M_{II}}$, and $M = M_I + M_{II}$. We focus on the special class of test models given by the system of (Itô) SDE's,

$$\begin{cases} d\mathbf{u}_I = F_I(\mathbf{u}_I, t) dt + \sigma_I(\mathbf{u}_I, t) dW_I(t), \\ d\mathbf{u}_{II} = (F_{II}(\mathbf{u}_I, t) + \Gamma(\mathbf{u}_I, t) \mathbf{u}_{II}) dt + \sigma_{II}(\mathbf{u}_I, t) dW_I(t) \\ \quad + \sigma_{II,A}(\mathbf{u}_I, t) dW_{II,A}(t) + (\sigma_{II,0} + \sigma_{II,M}(\mathbf{u}_I, t) \mathbf{u}_{II}) dW_{II,M}, \end{cases} \quad (2.1)$$

where W_I is an M_I -dimensional Wiener process, and $W_{II,A}, W_{II,0}, W_{II,M}$ are independent M_{II} -dimensional Wiener processes. Note that the dynamics of \mathbf{u}_I is arbitrary while the dynamics of \mathbf{u}_{II} is quasilinear, i.e., it is linear in \mathbf{u}_{II} in both the drift and the noise with general nonlinear coefficients depending on \mathbf{u}_I . Also, note that the noise for \mathbf{u}_I and \mathbf{u}_{II} can be correlated through W_I appearing in the equations for both \mathbf{u}_I and \mathbf{u}_{II} . All of the nonlinear test models for slow-fast systems (see [12–13]), turbulent tracers (see [2, 14, 21, 33]) and exactly solvable stochastic parameterized filters (see [8–9, 15, 34–35]) have the structural form as in (2.1). Such systems are known to have exactly solvable non-Gaussian statistics for filters, where \mathbf{u}_I is observed conditionally over a time interval [1, 20] [1,20]. Below, we derive explicit closed equations for the evolution of conditional moments of \mathbf{u}_2 through CGFPE.

The Fokker-Planck equation for the probability density $p(\mathbf{u}_I, \mathbf{u}_{II}, t)$ associated with (2.1) is given by [7, 36]

$$\begin{aligned} p_t = & -\nabla_I \cdot (F_I p) - \nabla_{II} \cdot ((F_{II} + \Gamma \mathbf{u}_{II})p) + \frac{1}{2} \nabla \cdot \nabla (Q p) \\ & + \frac{1}{2} \nabla_{II} \cdot \nabla_{II} (Q_A p) + \frac{1}{2} \nabla_{II} \cdot \nabla_{II} (Q_M p), \end{aligned} \quad (2.2)$$

where $\nabla = (\nabla_I, \nabla_{II})$, and

$$Q = (\sigma_I, \sigma_{II}) \otimes (\sigma_I^T, \sigma_{II}^T), \quad Q_A = \sigma_{II,A} \otimes \sigma_{II,A}^T, \quad Q_M = (\sigma_{II,0} + \sigma_{II,M} \mathbf{u}_{II}) \otimes (\sigma_{II,0}^T + \mathbf{u}_{II}^T \sigma_{II,M}^T). \quad (2.3)$$

We are interested in developing exact statistical approximations for $p(\mathbf{u}_I, \mathbf{u}_{II}, t)$ which, by Bayes theorem, can be written as

$$p(\mathbf{u}_I, \mathbf{u}_{II}, t) = p(\mathbf{u}_{II} | \mathbf{u}_I, t) \pi(\mathbf{u}_I, t), \quad (2.4)$$

where $\pi(\mathbf{u}_I, t)$ is the marginal distribution

$$\pi(\mathbf{u}_I, t) \equiv \int p(\mathbf{u}_I, \mathbf{u}_{II}, t) d\mathbf{u}_{II}. \quad (2.5)$$

We first integrate (2.2) with respect to \mathbf{u}_{II} and use the divergence theorem to verify that the marginal density $\pi(\mathbf{u}_I, t)$ satisfies the Fokker-Planck equation

$$\pi_t = \mathcal{L}_{FP,I} \pi \quad (2.6)$$

with

$$\mathcal{L}_{FP,I} \pi = -\nabla_I \cdot (F_I \pi) + \frac{1}{2} \nabla_I \cdot \nabla_I (Q_I \pi), \quad Q_I = \sigma_I \otimes \sigma_I^T. \quad (2.7)$$

Next, we derive the closed system of coupled generalized Fokker-Planck equations (CGFPE) for the conditional moments

$$\mathcal{M}_\alpha(\mathbf{u}_I, t) \equiv \int \mathbf{u}_{II}^\alpha p(\mathbf{u}_I, \mathbf{u}_{II}, t) d\mathbf{u}_{II} = \pi(\mathbf{u}_I, t) \int \mathbf{u}_{II}^\alpha p(\mathbf{u}_{II} | \mathbf{u}_I, t) d\mathbf{u}_{II}. \quad (2.8)$$

Note that $\mathcal{M}_0(\mathbf{u}_I, t) = \pi(\mathbf{u}_I, t)$ is just the marginal density of (2.1) in \mathbf{u}_I . Here and hereafter, we use the standard multi-index notation $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{M_{II}}) \in \mathbb{R}^{M_{II}}$ with

$$\mathbf{u}_{II}^\alpha \equiv (\mathbf{u}_{II})_1^{\alpha_1} (\mathbf{u}_{II})_2^{\alpha_2} \cdots (\mathbf{u}_{II})_{M_{II}}^{\alpha_{M_{II}}}. \quad (2.9)$$

We have the following general principles for computing the vector $\mathbf{M}_\alpha(\mathbf{u}_1, t) \equiv (\mathcal{M}_\alpha(\mathbf{u}_1, t))$ ($|\alpha| = N$) of conditional moments of order N .

Proposition 2.1 (Generalized Feynman-Kac Framework) *The vector $\mathbf{M}_N(\mathbf{u}_1, t)$ of conditional moments of order N associated with the probability density of (2.1) satisfies the system of coupled generalized Fokker-Planck equations (CGFPE)*

$$\begin{aligned} \frac{\partial \mathcal{M}_N(\mathbf{u}_1, t)}{\partial t} &= \mathcal{L}_{\text{FP}} \mathcal{M}_N(\mathbf{u}_1, t) + \mathcal{L}_N(\mathbf{u}_1, t) \mathcal{M}_N(\mathbf{u}_1, t) \\ &+ \mathcal{F}_N(\mathbf{u}_1, \mathcal{M}_{N-1}(\mathbf{u}_1, t), \nabla_{\mathbf{I}} \mathcal{M}_{N-1}(\mathbf{u}_1, t), \mathcal{M}_{N-2}(\mathbf{u}_1, t)) \end{aligned} \quad (2.10)$$

with the convention $\mathcal{M}_{-2} = \mathcal{M}_{-1} = 0$, where \mathcal{F}_N is an explicit linear function with coefficients depending on \mathbf{u}_1 of the lower order moments, \mathcal{L}_N is an $N \times N$ Feynman-Kac matrix potential which is an explicit linear function with coefficients depending on \mathbf{u}_1 of the quantities

$$\Gamma(\mathbf{u}_1, t), \quad Q_{\text{II}, \text{M}} = \sigma_{\text{II}, \text{M}} \otimes \sigma_{\text{II}, \text{M}}^{\text{T}}, \quad (2.11)$$

which vanishes when $\Gamma = 0$ and $Q_{\text{II}, \text{M}} = 0$.

The proof below immediately yields explicit formulas for \mathcal{L}_N and \mathcal{F}_N in any concrete application (see Section 3). But a general notation for these coefficients will be tedious and unnecessary to develop here. The advantage of CGFPE in (2.10) is that high resolution numerical integrators can be developed for (2.10) to find these statistics provided that M_I is low-dimensional or has the special algebraic structure (see Section 3).

The sketch of the proof below emphasizes the main contributions to the operator \mathcal{L}_N in (2.10). As in the derivation of (2.6), we first multiply the Fokker-Planck equation (2.2) by $\mathbf{u}_{\text{II}}^\alpha$ and integrate with respect to \mathbf{u}_{II} to obtain

$$\begin{aligned} \frac{\partial \mathcal{M}_N(\mathbf{u}_1, t)}{\partial t} &= \mathcal{L}_{\text{FP}} \mathcal{M}_N(\mathbf{u}_1, t) - \int \mathbf{u}_{\text{II}}^\alpha \cdot \nabla_{\text{II}} (\Gamma(\mathbf{u}_1, t) \mathbf{u}_{\text{II}} p) d\mathbf{u}_{\text{II}} \\ &+ \frac{1}{2} \int \mathbf{u}_{\text{II}}^\alpha \cdot \nabla_{\text{II}} \cdot \nabla_{\text{II}} (\sigma_{\text{II}, \text{M}} \mathbf{u}_{\text{II}} \otimes \mathbf{u}_{\text{II}}^{\text{T}} \sigma_{\text{II}, \text{M}}^{\text{T}} p) d\mathbf{u}_{\text{II}} + \dots, \end{aligned} \quad (2.12)$$

where “+...” denotes all the remaining terms which define the recursive source term \mathcal{F}_N . We simplify (2.12) by using the integration by parts of the last two terms on the right hand side, namely,

$$- \int \mathbf{u}_{\text{II}}^\alpha \cdot \nabla_{\text{II}} (\Gamma(\mathbf{u}_1, t) \mathbf{u}_{\text{II}} p) d\mathbf{u}_{\text{II}} = \int \nabla_{\text{II}} \mathbf{u}_{\text{II}}^\alpha \cdot (\Gamma(\mathbf{u}_1, t) \mathbf{u}_{\text{II}}) p d\mathbf{u}_{\text{II}} = \mathcal{L}_N^{(1,2)} \mathcal{M}_N(\mathbf{u}_1, t) \quad (2.13)$$

and

$$\begin{aligned} &\frac{1}{2} \int \mathbf{u}_{\text{II}}^\alpha \cdot \nabla_{\text{II}} \cdot \nabla_{\text{II}} (\sigma_{\text{II}, \text{M}} \mathbf{u}_{\text{II}} \otimes \mathbf{u}_{\text{II}}^{\text{T}} \sigma_{\text{II}, \text{M}}^{\text{T}} p) d\mathbf{u}_{\text{II}} \\ &= \frac{1}{2} \int \nabla_{\text{II}} \cdot \nabla_{\text{II}} \mathbf{u}_{\text{II}}^\alpha \cdot (\sigma_{\text{II}, \text{M}} \mathbf{u}_{\text{II}} \otimes \mathbf{u}_{\text{II}}^{\text{T}} \sigma_{\text{II}, \text{M}}^{\text{T}}) p d\mathbf{u}_{\text{II}} \\ &= \mathcal{L}_N^{(2,2)} \mathcal{M}_N(\mathbf{u}_1, t), \end{aligned} \quad (2.14)$$

so that $\mathcal{L}_N = \mathcal{L}_N^{(1,2)} + \mathcal{L}_N^{(2,2)}$ in (2.10). The remaining terms in “+...” are explicitly computed by a similar integration by parts to define \mathcal{F}_N . The correlated noise terms in (2.1) involving

W_I which defines the noise Q in (2.2) determine the dependence on $\nabla \mathcal{M}_{N-1}(\mathbf{u}_I, t)$ in \mathcal{F}_N since they have the typical form

$$-\int \mathbf{u}_{II}^\alpha \nabla_I \cdot \nabla_{II} (\sigma_I \sigma_{II}^T p) d\mathbf{u}_{II} = \nabla_I \cdot \int \nabla_{II} \mathbf{u}_{II}^\alpha (\sigma_I \sigma_{II}^T p) d\mathbf{u}_{II} = \mathcal{F}_N(\mathbf{u}_I, \nabla_I \mathcal{M}_{N-1}(\mathbf{u}_I, t)). \quad (2.15)$$

It is worth pointing out that \mathcal{F}_N depends only on the point-wise values of $\mathcal{M}_{N-1}(\mathbf{u}_I, t)$ and $\mathcal{M}_{N-2}(\mathbf{u}_I, t)$ if there are non-correlated noise interactions and $\sigma_{II} = 0$.

3 Application of the Conditional Moment PDE's to a Non-Gaussian Test Model

We develop the simplest non-Gaussian test model, where we can explicitly evaluate non-trivial statistical features utilizing the coupled system of PDE's in (2.10) from §2 for the conditional moments $\mathcal{M}_\alpha(\mathbf{u}_I, t)$. We then derive and validate a numerical procedure for the accurate numerical solution to the closed system of equations in (2.10) for the conditional moments in several stringent test problems. This explicit solution procedure is applied in §5 to understand the role of coarse-graining and non-Gaussian statistics with the model error in ensemble predictions.

Clearly, the simplest models considered with the structure as in (2.1) have $M_I = M_{II} = 1$ so that the recursion formulas in (2.10) involve scalar fields and the CGFPE are integrated in a single spatial dimension. For \mathbf{u}_I , we choose the general nonlinear scalar Itô SDE

$$du_I = F_I(u_I, t)dt + \sigma_I(u_I, t)dW_I, \quad (3.1)$$

while for u_{II} , we utilize the quasi-linear equation

$$du_{II} = (-u_I u_{II} + f(t))dt + \sigma_{II} dW_{II}, \quad (3.2)$$

where $f(t)$ does not depend on u_I , and the noise σ_{II} is constant. Note that u_I enters in (3.2) as a multiplicative coefficient and fluctuations in u_I can show the growth and intermittent instabilities with highly non-Gaussian behavior even when u_I in (3.1) has a positive mean (see [3, 5, 25]). The stochastic models for u_I in (3.1) will vary from linear stochastic models (a special case of the SPEKF models for filtering (see [9, 25, 34–35])) to cubic nonlinear models with additive and multiplicative noise (see [25]). For the systems with dynamics as in (3.1)–(3.2), the closed equations for the conditional moments \mathcal{M}_α in (2.8) become

$$\begin{aligned} \frac{\partial}{\partial t} \mathcal{M}_N(u_I, t) &= \mathcal{L}_{FP} \mathcal{M}_N(u_I, t) - N u_I \mathcal{M}_N(u_I, t) + N f(t) \mathcal{M}_{N-1}(u_I, t) \\ &\quad + \frac{1}{2} N(N-1) \sigma_{II}^2 \mathcal{M}_{N-2}(u_I, t), \end{aligned} \quad (3.3)$$

where $N = 0, 1, \dots, N_{\max}$ and $\mathcal{M}_{-2} = \mathcal{M}_{-1} = 0$. Such models illustrate a wide range of intermittent non-Gaussian behavior mimicking one in vastly more complex systems (see [22]). These simple revealing models will be used in §5 to study various new aspects of model error in ensemble predictions for non-Gaussian turbulent systems.

3.1 Validation of a numerical method for solving the CGFPE

Determination of the time evolution of the conditional moments \mathcal{M}_α in (2.8) requires an accurate numerical procedure for solving the inhomogeneous system of coupled Fokker-Planck

equations (CGFPE) in (2.10). The algorithms discussed below is applied to the case $u_I \in \mathbb{R}$ (i.e., $M_I = 1$) which is sufficient for our purposes and leads to many new insights on the model error in imperfect ensemble predictions of turbulent systems with positive Lyapunov exponents, as discussed in §5. Similar to the case of the homogeneous Fokker-Planck equation, solving the inhomogeneous CGFPE system (2.10) for $M_I \geq 3$ poses a formidable challenge which, for convenience, is unnecessary here.

Here, the coupled system in (2.10) is solved by the third-order temporal discretization through the backward differentiation formulas (see [17]) and the second-order spatial discretization via the finite volume method (see [19] and see Appendix A in §7 for details). The performance of the numerical procedure for solving CGFPE in one spatial dimension (i.e., $u_I \in \mathbb{R}$ in (2.10)) is tested in the following widely varying dynamical configurations:

- (i) Dynamics with time-invariant statistics on the attractor/equilibrium with
 - (a) nearly Gaussian marginal equilibrium PDFs in u_{II} and linear Gaussian dynamics for u_I in (3.1),
 - (b) fat-tailed marginal equilibrium PDFs in u_{II} and linear Gaussian dynamics for u_I in (3.1),
 - (c) highly non-Gaussian marginal equilibrium PDFs in u_{II} and cubic dynamics for u_I in (3.1) with highly skewed equilibrium PDFs.
- (ii) Dynamics with the time-periodic statistics on the attractor with the time-periodic regime switching between nearly Gaussian and highly skewed regimes with cubic dynamics for u_I in (3.1) and highly non-Gaussian dynamics of u_{II} in (3.2).

Below, we introduce the relevant test models in §3.1.1 and provide the evidence for the good accuracy of the developed technique in §3.1.2, as well as its advantages over the direct Monte Carlo sampling.

3.1.1 Non-Gaussian test models for validating CGFPE

We consider two non-Gaussian models with intermittent instabilities and with the structure as in (3.1)–(3.2), where we adopt the following notations:

$$u_I = \gamma, \quad u_{II} = u.$$

The first model is a simplified version of the SPEKF model developed originally for filtering turbulent systems with stochastically parameterized unresolved variables (see [8–9, 15, 34–35]) and is given by

$$\begin{aligned} \text{(a)} \quad d\gamma &= (-d_\gamma(\gamma - \hat{\gamma}) + f_\gamma(t))dt + \sigma_\gamma dW_\gamma, \\ \text{(b)} \quad du &= (-\gamma u + f_u(t))dt + \sigma_u dW_u. \end{aligned} \tag{3.4}$$

Note that despite the Gaussian dynamics of the damping fluctuation γ , the dynamics of u in (3.4) can be highly non-Gaussian with intermittently positive Lyapunov exponents even when the equilibrium mean $\hat{\gamma}$ is positive (see [3–4, 25]). The system (3.4) possesses a wide range of turbulent dynamical regimes ranging from highly non-Gaussian dynamics with intermittency and fat-tailed marginal PDFs for u to laminar regimes with nearly Gaussian statistics. A detailed discussion of properties of this system can be found in [3, 5]. In the numerical tests discussed in the next section, we examine the accuracy of the numerical algorithm for solving CGFPE in the dynamical regime characterized by a highly intermittent marginal dynamics in

u associated with fat-tailed marginal equilibrium PDFs for u (see Figure 1 for examples of such dynamics).

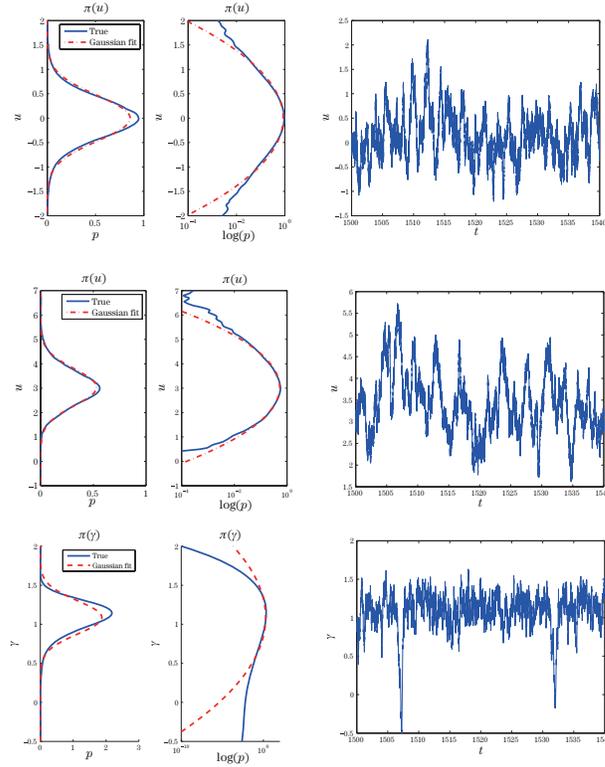


Figure 1 (Top) Marginal statistics $p_{\text{eq}}(u)$ and a path-wise solution $u(t)$ on the attractor of the system (3.4) in the non-Gaussian regime with invariant measure characterized by intermittent transient instabilities and fat-tailed marginal PDFs (dynamics of γ is Gaussian in this model). (Bottom) Marginal statistics $p_{\text{eq}}(\gamma)$ and $p_{\text{eq}}(u)$, and path-wise solutions $\gamma(t)$ and $u(t)$ on the attractor of the system (3.5) in the non-Gaussian regime with the regime switching in the path-wise dynamics despite a unimodal, skewed marginal PDF in γ .

The second model which we examine, has a cubic nonlinearity in the dynamics of the damping fluctuation γ , and is given by

$$\begin{aligned}
 \text{(a)} \quad d\gamma &= [-a\gamma + b\gamma^2 - c\gamma^3 + f_\gamma(t)]dt + (A - B\gamma)dW_C + \sigma_\gamma dW_\gamma, \\
 \text{(b)} \quad du &= (-\gamma u + f_u(t))dt + \sigma_u dW_u.
 \end{aligned} \tag{3.5}$$

The above nonlinear model for γ with correlated additive and multiplicative noise W_C and exactly solvable equilibrium statistics was first derived in [26] as a normal form for a single low-frequency variable in climate models, where the noise correlations arise through advection of the large scales by the small scales and simultaneously strong cubic damping. The nonlinear dynamics of γ has many interesting features which were studied in detail elsewhere (see [25]). Here, we consider a more complex problem with the dynamics of u in (a) of (3.5) coupled with γ through the quadratic nonlinearity. In the numerical tests below, we focus on the particularly interesting regime, where the damping fluctuation γ exhibits the regime switching despite the unimodality of the associated equilibrium statistics (see Figure 1). This configuration represents

the simplest possible test model for the analogous behavior occurring in comprehensive climate models (see [23, 27]). Another important configuration of (3.5) tested below with relevance to atmospheric/climate dynamics corresponds to time-periodic transitions in γ between a highly skewed and a nearly Gaussian phase in γ with the dynamics in u remaining highly non-Gaussian throughout the evolution (see Figure 2 for an illustration of such dynamics).

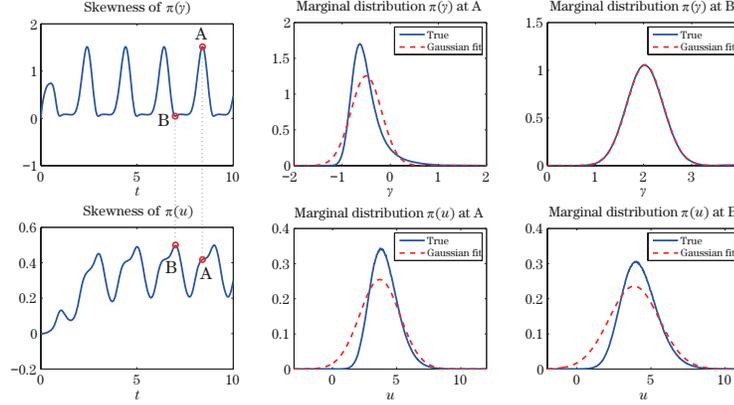


Figure 2 (Top) Time-periodic evolution of the skewness of the marginal dynamics of γ in the non-Gaussian system (3.5) with cubic nonlinearity in γ in the configuration, where γ cycles between a highly skewed (top middle) and a nearly Gaussian (top right) phase. The phases of high/low skewness in the marginal statistics of γ are correlated with those in the marginal statistics of u . However, note that the dynamics of u remains highly non-Gaussian throughout the evolution. The snapshots of the marginal PDFs in u on the bottom are shown for the times indicated on the top panel.

The above two non-Gaussian models are utilized below to validate the accuracy of our numerical method for solving the CGFPE system (2.10). This framework is then used to analyze the model error in imperfect predictions of turbulent non-Gaussian systems in §5.

3.1.2 Numerical tests

We use the test models introduced in the previous section to analyze the performance of the numerical scheme for solving the CGFPE system (2.10) in one-spatial dimension. In order to assess the accuracy of the algorithm, we consider the following two types of the relative error in the conditional moments: the point-wise relative error in the N -th conditional moment

$$\epsilon_N(\gamma, t) = \left| \frac{\mathcal{M}_N^{\text{CGFPE}}(\gamma, t) - \mathcal{M}_N^{\text{ref}}(\gamma, t)}{\mathcal{M}_N^{\text{ref}}(\gamma, t)} \right| \quad (3.6)$$

and the L^2 relative error for each fixed time

$$\epsilon_N(t) = \frac{\|\mathcal{M}_N^{\text{CGFPE}}(\gamma, t) - \mathcal{M}_N^{\text{ref}}(\gamma, t)\|_{L^2}}{\|\mathcal{M}_N^{\text{ref}}(\gamma, t)\|_{L^2}}. \quad (3.7)$$

The reference values for the conditional moments, $\mathcal{M}_N^{\text{ref}}$ in the above formulas, are obtained from either the analytical solutions (in the case of system (3.4) through the formulas derived in

[9]), or via the Monte Carlo estimates. The conditional moments are normalized in the standard fashion, with the conditional mean, variance, skewness and kurtosis given by

$$\widetilde{\mathcal{M}}_0(\gamma, t) = \mathcal{M}_0(\gamma, t), \quad \widetilde{\mathcal{M}}_1(\gamma, t) = \mathcal{M}_1(\gamma, t), \quad (3.8)$$

$$\widetilde{\mathcal{M}}_2(\gamma, t) = \int (u(t) - \mathcal{M}_1(\gamma, t))^2 p(u, \gamma, t) du = \mathcal{M}_2(\gamma, t) - \mathcal{M}_1^2(\gamma, t), \quad (3.9)$$

$$\begin{aligned} \widetilde{\mathcal{M}}_3(\gamma, t) &= \frac{1}{\widetilde{\mathcal{M}}_2^{\frac{3}{2}}(\gamma, t)} \int (u(t) - \mathcal{M}_1(\gamma, t))^3 p(u, \gamma, t) du \\ &= \frac{\mathcal{M}_3(\gamma, t) - 3\mathcal{M}_1(\gamma, t)\mathcal{M}_2(\gamma, t) + 2\mathcal{M}_1^3(\gamma, t)}{\widetilde{\mathcal{M}}_2^{\frac{3}{2}}(\gamma, t)}, \end{aligned} \quad (3.10)$$

$$\begin{aligned} \widetilde{\mathcal{M}}_4(\gamma, t) &= \frac{1}{\widetilde{\mathcal{M}}_2^2(\gamma, t)} \int (u(t) - \mathcal{M}_1(\gamma, t))^4 p(u, \gamma, t) du \\ &= \frac{\mathcal{M}_4(\gamma, t) - 4\mathcal{M}_1(\gamma, t)\mathcal{M}_3(\gamma, t) + 6\mathcal{M}_1^2(\gamma, t)\mathcal{M}_2(\gamma, t) - 3\mathcal{M}_1^4(\gamma, t)}{\widetilde{\mathcal{M}}_2^2(\gamma, t)}, \end{aligned} \quad (3.11)$$

respectively.

The L^2 errors for the two test models discussed in the previous section and parameters as specified below are listed in Tables 1–3. Note that the errors in the conditional moments do not exceed 6% for the wide range of dynamical regimes considered. Moreover, the comparison of the results in Tables 1–2 shows that the numerical algorithm developed here is more efficient and accurate than the Monte Carlo estimates, even when a relatively large sample size ($\sim 10^7$) is used in the MC simulations.

Table 1 Relative errors ϵ_N in (3.7), in the conditional moments \mathcal{M}_0 – \mathcal{M}_4 in (3.8)–(3.11) at equilibrium for the two test models (3.4)–(3.5) with the reference input obtained from Monte Carlo estimates from 10^7 runs.

	\mathcal{M}_0	\mathcal{M}_1	\mathcal{M}_2	\mathcal{M}_3	\mathcal{M}_4
System (3.4): Nearly Gaussian reg.	0.0031		0.0241		0.0494
System (3.4): Fat algebraic tail reg.	0.0225		0.0202		0.0593
System (3.5): High skewness reg.	0.0179	0.0181	0.0183	0.0196	0.0236

Table 2 Relative errors (3.7) in the conditional moments \mathcal{M}_0 , \mathcal{M}_1 and \mathcal{M}_3 at equilibrium for the two test models (3.4)–(3.5) with the reference input obtained from analytical solutions.

	\mathcal{M}_0	\mathcal{M}_1	\mathcal{M}_2	\mathcal{M}_3	\mathcal{M}_4
System (3.4): Nearly Gaussian reg.	2.1520×10^{-5}	0		0	
System (3.4): Fat algebraic tail reg.	3.6825×10^{-6}	0		0	
System (3.5): High skewness reg.	0.0018				

Table 3 Relative errors in time-periodic conditional moments \mathcal{M}_0 – \mathcal{M}_4 in (3.8)–(3.11) for the test model (3.5) in the regime with transitions (see Figure 1) between highly skewed and nearly Gaussian marginal densities $\pi_{att}(\gamma)$; the reference input obtained from Monte Carlo estimates from 10^7 runs.

	\mathcal{M}_0	\mathcal{M}_1	\mathcal{M}_2	\mathcal{M}_3	\mathcal{M}_4
$t_* = 7.00$	0.0185	0.0199	0.0218	0.0242	0.0271
$t_* = 8.40$	0.0299	0.0337	0.0400	0.0447	0.0561
$t_* = 7.70$	0.0309	0.0316	0.0321	0.0327	0.0332
$t_* = 9.00$	0.0182	0.0196	0.0229	0.0275	0.0330

In Figure 3, we illustrate the performance of the algorithm for computing the conditional moments in (2.10) associated with the conditional equilibrium density $p_{\text{eq}}(u | \gamma)$ for the system (3.4). The system parameters $(d_\gamma, \sigma_\gamma, \hat{\gamma}, \sigma_u)$ in (3.4) are chosen to represent the non-Gaussian dynamics in the regime with intermittent instabilities and a fat-tailed marginal equilibrium PDF in u . In particular, we choose

$$\sigma_\gamma = 10, \quad d_\gamma = 10, \quad \hat{\gamma} = 3, \quad f_u = f_\gamma = 0$$

(see Figure 1 for an example of the corresponding dynamics).

In Figures 5–6, we illustrate the performance of our algorithm for computing the conditional moments $\mathcal{M}_\alpha(\gamma, t)$ of u in the system (3.5) with the cubic nonlinearity in γ which is coupled multiplicatively to the dynamics in u . Here, we consider two distinct configurations. For the constant forcing, we choose the parameters in (3.5) in such a way that γ displays regime switching with the unimodal, highly skewed marginal equilibrium PDF for γ , while the marginal dynamics of u is highly non-Gaussian and second-order stable. This dynamical configuration can be achieved by setting, for example,

$$\begin{aligned} a &= 1, & b &= 1, & c &= 1, \\ A &= 0.5, & B &= -2, \\ \sigma &= 1, & f_u &= 1, & f_\gamma &= 3 \end{aligned}$$

(see Figure 1 for an illustration of such dynamics). For the time-periodic forcing, when the dynamics in γ cycles between a highly skewed and a nearly Gaussian phase while u remains highly non-Gaussian, we set

$$\begin{aligned} a &= 1, & b &= 1, & c &= 1, \\ A &= 0.5, & B &= -0.5, \\ \sigma &= 0.5, & f_u &= -0.5 \end{aligned}$$

with the time-periodic forcing in γ given by

$$f_\gamma(t) = 6.5 \sin\left(\pi t - \frac{\pi}{2}\right) + 2.5.$$

Based on the results summarized in Figures 3–6 and Tables 1–3, we make the following points:

- (1) The numerical algorithm for solving the coupled system (2.10) in the CGFPE framework with $u_i \in \mathbb{R}$ provides robust and accurate estimates for the conditional moments (2.8).
- (2) The discrepancies between the estimates obtained from (2.10) and the direct Monte Carlo estimates with a large sample size ($\sim 10^7$) are below 6% for both time-periodic and time-invariant attractor statistics.
- (3) The largest discrepancies in the normalized conditional moments obtained from CGFPE and Monte Carlo estimates in the normalized moments occur in tail regions, where the corresponding probability densities are very small.
- (4) The developed algorithm for solving the CGFPE system (2.10) is more efficient and more accurate than the Monte Carlo estimates with relatively large sample sizes ($\sim 10^7$).

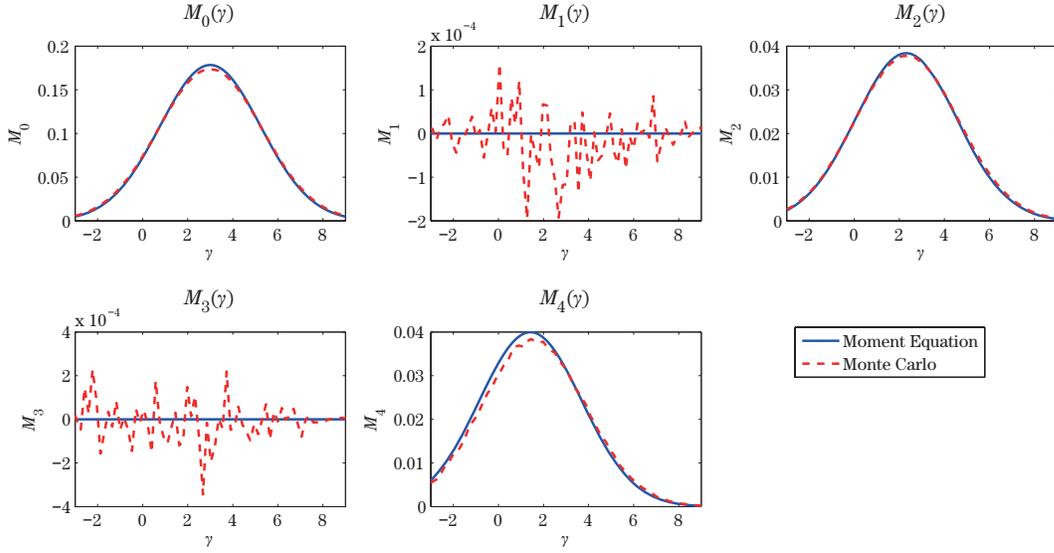


Figure 3 Equilibrium conditional statistics of the system (3.4) with Gaussian damping fluctuations, intermittent instabilities and fat-tailed marginal PDFs in u . Unnormalized conditional moments, $\mathcal{M}_0(\gamma) - \mathcal{M}_4(\gamma)$, (2.8) of u at the equilibrium of the two-dimensional non-Gaussian turbulent system (3.4) with intermittent instabilities due to Gaussian damping fluctuations; the results of CGFPE (2.10) and Monte Carlo estimates from 10^7 runs are compared. In the dynamical regime shown in the marginal equilibrium PDF, $p_{\text{eq}}(u)$, is symmetric and fat-tailed due to these intermittent instabilities (see Figure 1). Note the errors in the Monte Carlo estimates in the odd moments.

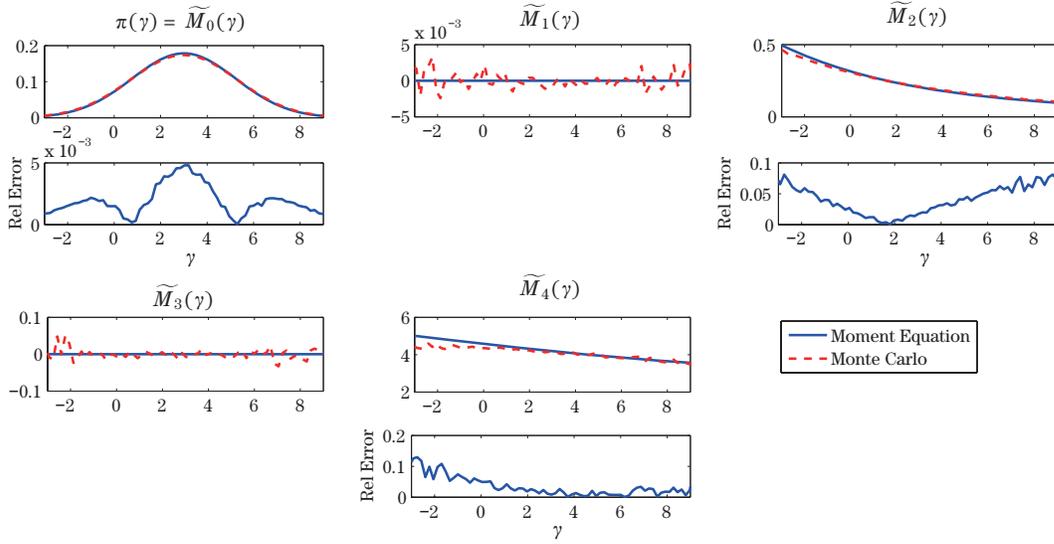


Figure 4 Teh same as that in Figure 3 above but for centered, normalized conditional moments $\widetilde{\mathcal{M}}_0(\gamma) - \widetilde{\mathcal{M}}_4(\gamma)$, which correspond to marginal density $\pi(\gamma)$, the conditional mean, variance, skewness and kurtosis given by (3.8)–(3.11), respectively.

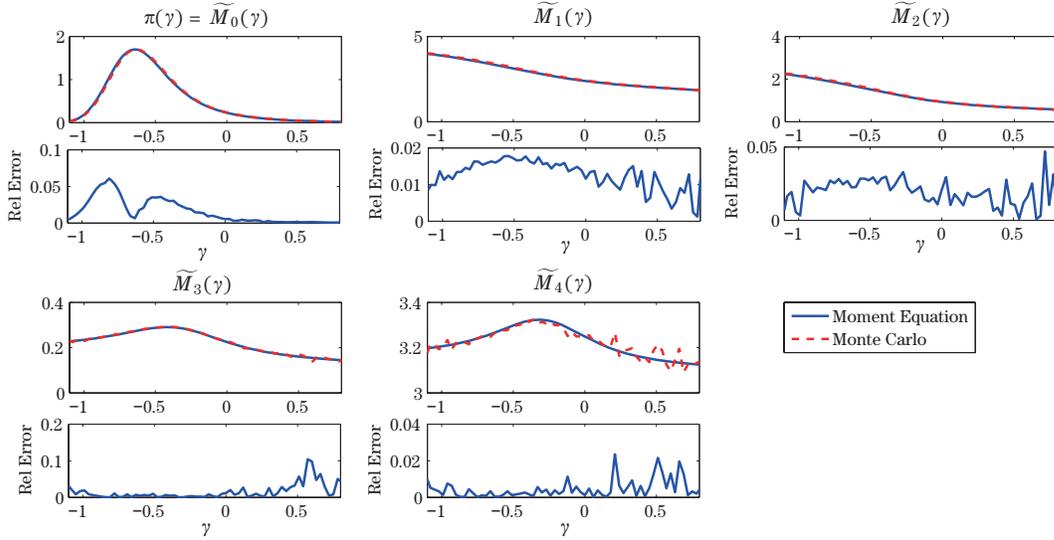


Figure 5 Snapshots of time-periodic conditional statistics on the attractor of the system (3.5); the cubic nonlinearity in damping fluctuations, the highly skewed PDF phase. Normalized conditional moments $\widetilde{\mathcal{M}}_0(\gamma, t_*) - \widetilde{\mathcal{M}}_4(\gamma, t_*)$ in (3.8)–(3.11) of u on the time-periodic attractor of the two-dimensional non-Gaussian turbulent system (3.5) with cubic dynamics of damping fluctuations γ ; the results obtained via CGFPE (2.10) and Monte Carlo simulations with 10^7 sample runs are compared at time $t_* = 8.4$ which corresponds to the highly non-Gaussian phase with highly skewed marginal PDFs, $\pi_{att}(u, t_*)$, $\pi_{att}(\gamma, t_*)$ (see Figure 2). The normalized conditional moments are the conditional mean, variance skewness and kurtosis.

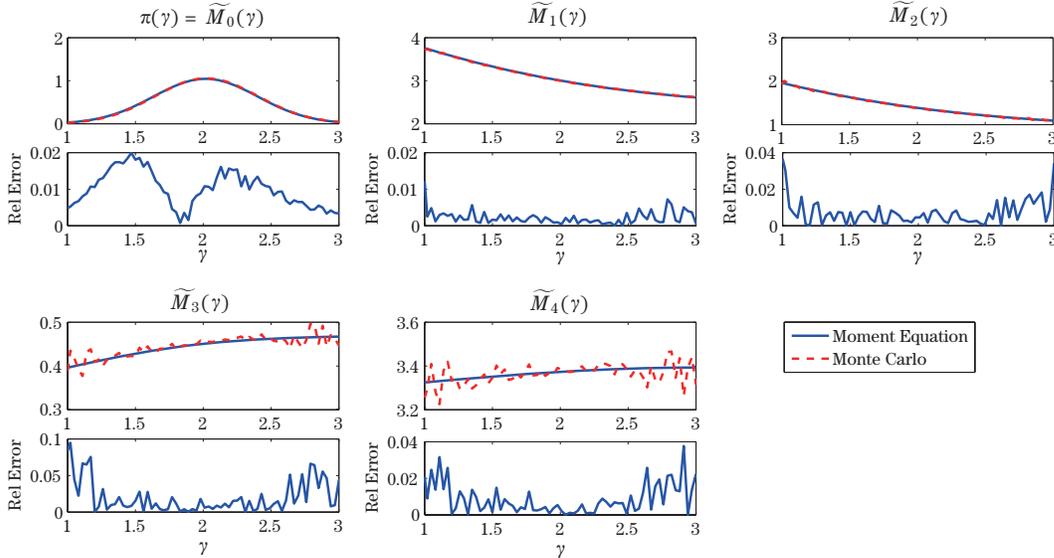


Figure 6 The same as that in Figure 5 above, but showing the normalized conditional moments $\widetilde{\mathcal{M}}_1(\gamma, t_*) - \widetilde{\mathcal{M}}_4(\gamma, t_*)$ (i.e., the conditional mean, variance, skewness and kurtosis), at $t_* = 7$ which corresponds to the nearly Gaussian phase in γ , but has a highly skewed marginal $\pi_{att}(u, t_*)$ (see also Figure 2).

4 Quantifying Model Error Through Empirical Information Theory

As discussed extensively recently (see [5, 11, 25, 28–30]), a very natural way to quantify the model error in statistical solutions to complex systems is through the relative entropy $\mathcal{P}(p, q) \geq 0$ for two probability measures p and q given by

$$\mathcal{P}(p, q) = \int p \ln \frac{p}{q} = -\mathcal{S}(p) - \int p \ln q, \quad (4.1)$$

where

$$\mathcal{S}(p) = - \int p \ln p \quad (4.2)$$

is the Shannon entropy of the probability measure p . The relative entropy $\mathcal{P}(p, q)$ measures the lack of information in q about the probability measure p . If p is the perfect density and p_M , $M \in \mathcal{M}$ is a class of probability densities, then M_1 is a better model than M_2 provided that

$$\mathcal{P}(p, p_{M_1}) < \mathcal{P}(p, p_{M_2}), \quad (4.3)$$

and the best model $M^* \in \mathcal{M}$ satisfies

$$\mathcal{P}(p, p_{M^*}) = \min_{M \in \mathcal{M}} \mathcal{P}(p, p_M). \quad (4.4)$$

There are extensive applications of information theory to improve imperfect models in climate science developed recently (see [5, 11, 25, 28–30]); the interested reader can refer to these references. The goal here is to develop and illustrate this perspective of the information theory on the model error for direct application to the estimate of model error for the setup developed above in Sections 2–3. These formulas are utilized in §5 below.

We consider a probability density for the perfect model $p(\mathbf{u}_I, \mathbf{u}_{II})$ which can be written by Bayes theorem as

$$p(\mathbf{u}_I, \mathbf{u}_{II}) = p(\mathbf{u}_{II} | \mathbf{u}_I) \pi(\mathbf{u}_I), \quad (4.5)$$

here and hereafter, $\pi(\mathbf{u}_I)$ is the marginal

$$\pi(\mathbf{u}_I) = \int p(\mathbf{u}_I, \mathbf{u}_{II}) d\mathbf{u}_{II}. \quad (4.6)$$

From the CGFPE procedure developed in Sections 2–3, we have exact expressions for the conditional moments up to some order L for $p(\mathbf{u}_{II} | \mathbf{u}_I)$ evolving in time already, this is a source of information loss through the coarse graining of $p(\mathbf{u}_I, \mathbf{u}_{II})$. To quantify this information loss by measuring only the conditional moments up to order L , let

$$p_L(\mathbf{u}_I, \mathbf{u}_{II}) = p_L(\mathbf{u}_{II} | \mathbf{u}_I) \pi(\mathbf{u}_I), \quad (4.7)$$

where for each value \mathbf{u}_I , the conditional density $p_L(\mathbf{u}_{II} | \mathbf{u}_I)$ satisfies the maximum entropy (least biased) criterion (see [24, 31–32])

$$\mathcal{S}(p_L(\mathbf{u}_{II} | \mathbf{u}_I)) = \max_{\pi_L \in \mathcal{L}} \mathcal{S}(\pi_L(\mathbf{u}_{II})), \quad (4.8)$$

where \mathfrak{L} is a class of marginal densities π_L with identical moments up to order L , i.e.,

$$\int \mathbf{u}_{\text{II}}^\alpha \pi_L(\mathbf{u}_{\text{II}}) d\mathbf{u}_{\text{II}} = \int \mathbf{u}_{\text{II}}^\alpha p_L(\mathbf{u}_I, \mathbf{u}_{\text{II}}) d\mathbf{u}_{\text{II}} = \int \mathbf{u}_{\text{II}}^\alpha p(\mathbf{u}_I, \mathbf{u}_{\text{II}}) d\mathbf{u}_{\text{II}}, \quad |\alpha| \leq L. \quad (4.9)$$

Below and in Section 5, we will always apply the variational problem in (4.8) for $L = 2$ which guarantees that $p_L(\mathbf{u}_{\text{II}} | \mathbf{u}_I)$ is a Gaussian density with the specified conditional mean and variance. In general, for L even and $L > 2$, it is a subtle issue as to whether the solution to the variational problem (4.8) exists (see [34]), but here we tacitly assume this. We remark here that highly non-Gaussian densities can have Gaussian conditional densities like $p_L(\mathbf{u}_{\text{II}} | \mathbf{u}_I)$ as discussed in §5.

Natural imperfect densities with the model error have the form

$$p_L^M(\mathbf{u}_I, \mathbf{u}_{\text{II}}) = p_L^M(\mathbf{u}_{\text{II}} | \mathbf{u}_I) \pi^M(\mathbf{u}_I). \quad (4.10)$$

The simplest model with the model error is a Gaussian density $p_G(\mathbf{u}_I, \mathbf{u}_{\text{II}})$ which is defined by its mean and variance; the standard regression formula for Gaussian densities (see [7]) automatically guarantees that the form in (4.10) is applied with $L = 2$ in this important case.

Another important way of generating an imperfect model with the form (4.10) is to have a different model (see [22, 25]) for the stochastic dynamics of \mathbf{u}_I rather than that in (2.1) and to compute the conditional moments up to order L in the approximate model through CGFPE so that the model approximations automatically have the form (4.10) (see Section 5 below).

Here we have a precise way to quantify the model error in an imperfect model in the present setup.

Proposition 4.1 *Given the perfect model distribution $p(\mathbf{u}_I, \mathbf{u}_{\text{II}})$ with its conditional approximation $p_L(\mathbf{u}_I, \mathbf{u}_{\text{II}})$ in (4.7) and the imperfect model density $p_L^M(\mathbf{u}_I, \mathbf{u}_{\text{II}})$ defined in (4.10), we have*

$$\mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{\text{II}}), p_L^M(\mathbf{u}_I, \mathbf{u}_{\text{II}})) = \mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{\text{II}}), p_L(\mathbf{u}_I, \mathbf{u}_{\text{II}})) + \mathcal{P}(p_L(\mathbf{u}_I, \mathbf{u}_{\text{II}}), p_L^M(\mathbf{u}_I, \mathbf{u}_{\text{II}})), \quad (4.11)$$

where

$$\begin{aligned} 0 &\leq \mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{\text{II}}), p_L(\mathbf{u}_I, \mathbf{u}_{\text{II}})) \\ &= \int \pi(\mathbf{u}_I) [\mathcal{S}(p_L(\mathbf{u}_{\text{II}} | \mathbf{u}_I)) - \mathcal{S}(p(\mathbf{u}_{\text{II}} | \mathbf{u}_I))] d\mathbf{u}_I \\ &= \int \pi(\mathbf{u}_I) \mathcal{P}(p(\mathbf{u}_{\text{II}} | \mathbf{u}_I), p_L(\mathbf{u}_{\text{II}} | \mathbf{u}_I)) d\mathbf{u}_I \end{aligned} \quad (4.12)$$

and

$$\begin{aligned} 0 &\leq \mathcal{P}(p_L(\mathbf{u}_I, \mathbf{u}_{\text{II}}), p_L^M(\mathbf{u}_I, \mathbf{u}_{\text{II}})) \\ &= \mathcal{P}(\pi(\mathbf{u}_I), \pi^M(\mathbf{u}_I)) + \int \pi(\mathbf{u}_I) [\mathcal{S}(p_L^M(\mathbf{u}_{\text{II}} | \mathbf{u}_I)) - \mathcal{S}(p_L(\mathbf{u}_{\text{II}} | \mathbf{u}_I))] d\mathbf{u}_I \\ &= \mathcal{P}(\pi(\mathbf{u}_I), \pi^M(\mathbf{u}_I)) + \int \pi(\mathbf{u}_I) \mathcal{P}(p(\mathbf{u}_{\text{II}} | \mathbf{u}_I), p_L^M(\mathbf{u}_{\text{II}} | \mathbf{u}_I)) d\mathbf{u}_I. \end{aligned} \quad (4.13)$$

In particular, $\mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{\text{II}}), p_L(\mathbf{u}_I, \mathbf{u}_{\text{II}}))$ quantifies an intrinsic information barrier (see [5, 11, 25, 29–30]) for all imperfect model densities with the form as in (4.10).

To prove Proposition 4.1, first, utilize the general identity (see [6]) to calculate

$$\mathcal{P}(p_L(\mathbf{u}_I, \mathbf{u}_{II}), p_L^M(\mathbf{u}_I, \mathbf{u}_{II})) = \mathcal{P}(\pi(\mathbf{u}_I), \pi^M(\mathbf{u}_I)) + \int \pi(\mathbf{u}_I) \mathcal{P}(p(\mathbf{u}_{II} | \mathbf{u}_I), p_L^M(\mathbf{u}_{II} | \mathbf{u}_I)) d\mathbf{u}_I, \quad (4.14)$$

which is easily verified by the reader. Next, for each \mathbf{u}_I , use the general identity for least biased densities, which follows from the max-entropy principle in (4.8) (see [24, Chapter 2])

$$\mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{II}), p_L^M(\mathbf{u}_I, \mathbf{u}_{II})) = \mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{II}), p_L^M(\mathbf{u}_I, \mathbf{u}_{II})) + \mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{II}), p_L^M(\mathbf{u}_I, \mathbf{u}_{II})), \quad (4.15)$$

and insert this in (4.14). Finally, computing $\mathcal{P}(p(\mathbf{u}_I, \mathbf{u}_{II}), p_L(\mathbf{u}_I, \mathbf{u}_{II}))$ and $\mathcal{P}(p_L(\mathbf{u}_I, \mathbf{u}_{II}), p_L^M(\mathbf{u}_I, \mathbf{u}_{II}))$ by the formula in (4.14) once again, with simple algebra, we arrive at the required formulas in (4.11)–(4.13).

5 Non-Gaussian Test Models for Statistical Prediction with Model Errors

We apply the material developed in Sections 3–4 with $L = 2$ to gain new insight into statistical predictions with the effects of coarse-graining and model errors in the non-Gaussian setting. In the first part of this section, we consider the effect of model errors through coarse-graining the statistics in a perfect model setting (see [18]) for short, medium, and long range forecasting. In the second part of this section, we consider the effect of model errors in the dynamics of \mathbf{u}_I (see [25]) on the long range forecasting skill. The errors in both the full probability density and the marginal densities in \mathbf{u}_{II} are considered.

5.1 Choice of initial statistical conditions

As already mentioned in §4, we are particularly interested in assessing the model error due to various coarse-grainings of the perfect statistics. These model errors arise naturally either when deriving the approximate least-biased conditional densities through estimating the conditional moments in the CGFPE framework of §2, or when deriving the Gaussian estimators of non-Gaussian densities. The effects of initial conditions are clearly important in the short and medium range predictions, for both the perfect and the coarse-grained statistics, and the choice of a representative set of statistical initial conditions requires some care.

In the following sections, we consider the least-biased conditionally Gaussian estimators (i.e., $L = 2$ in §4) of the true statistics $p(u, \gamma, t)$, leading to the non-Gaussian densities $p_2(u, \gamma, t)$, as well as fully Gaussian approximations $p_G(u, \gamma, t)$ of the true non-Gaussian statistics $p(u, \gamma, t)$. Therefore, in order to compare the effects of coarse-graining the structure of the PDFs in a standardized setting, we consider the initial joint densities with identical second-order moments, i.e., any two initial densities, \tilde{p}_i and \tilde{p}_j , satisfy

$$\int u^\alpha \gamma^\beta \tilde{p}_i(u, \gamma) du d\gamma = \int u^\alpha \gamma^\beta \tilde{p}_j(u, \gamma) du d\gamma, \quad 0 \leq \alpha + \beta \leq 2. \quad (5.1)$$

For simplicity, we choose the initial densities with uncorrelated variables,

$$\tilde{p}_i(u, \gamma) = \tilde{\pi}_i(u) \tilde{\pi}_i(\gamma),$$

where the marginal densities $\tilde{\pi}_i(u)$ and $\tilde{\pi}_i(\gamma)$ are given by the mixtures of simple densities (see Appendix B in §7 for more details). This procedure is sufficient for the present purposes and reduces the complexity of exposition. An analogous procedure can be used to generate PDFs with correlated variables by, for example, changing the coordinate frame. Such a step might be necessary when studying the model error in filtering problems.

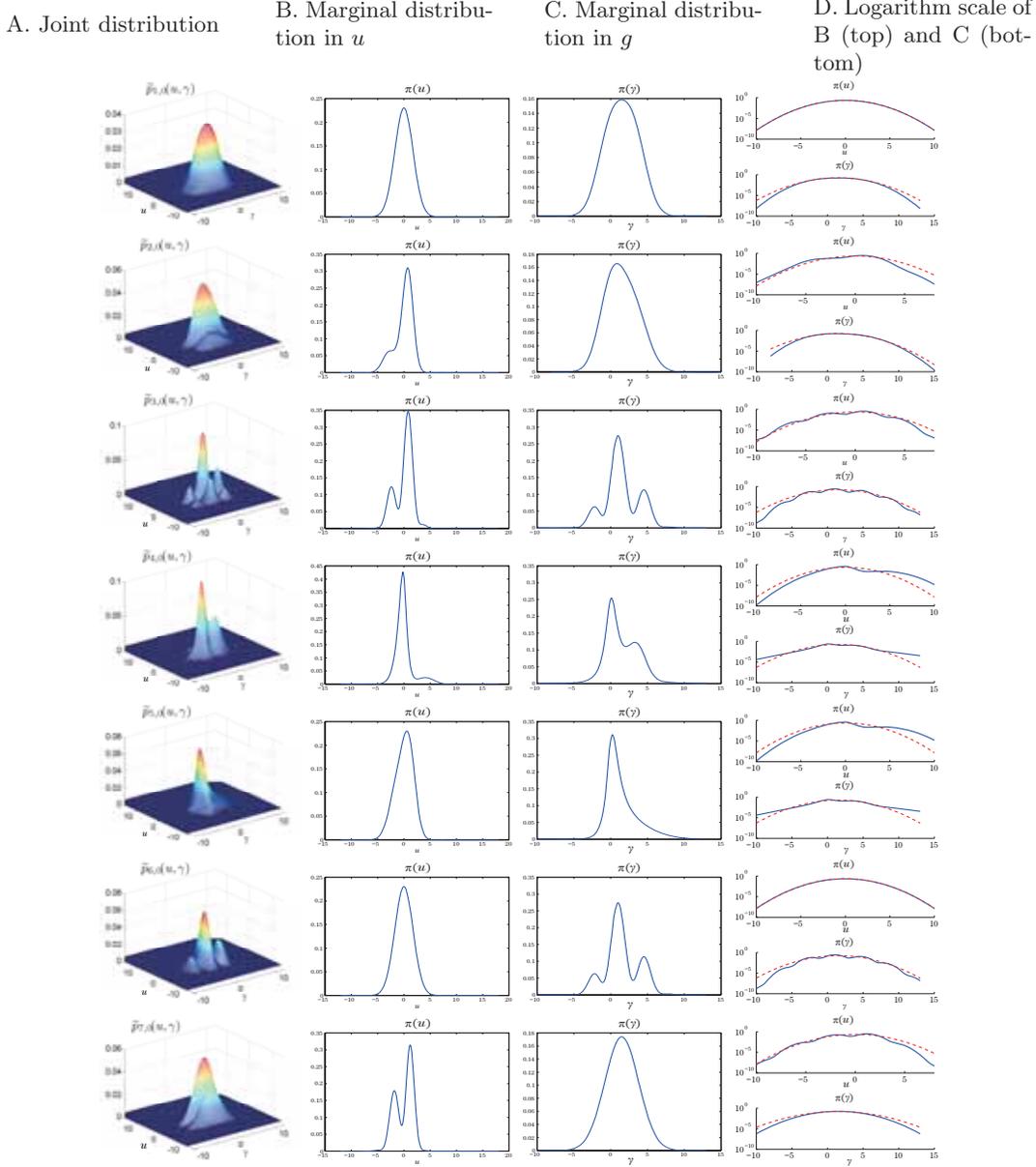


Figure 7 The set of seven non-Gaussian initial conditions with identical second-order statistics used in the tests in Figures 9–20 (see §5.1 and Appendix B in §7 for more details).

The following sets of non-Gaussian initial conditions, shown in Figure 7 and constructed in the way described above, are used in the suite of tests discussed next (see also Appendix B in

§7):

(1) $\tilde{p}_1(u, \gamma)$: Nearly Gaussian PDF with the Gaussian marginal in u and a weakly sub-Gaussian marginal in γ .

(2) $\tilde{p}_2(u, \gamma)$: PDF with a bimodal marginal in u and a weakly skewed marginal in γ .

(3) $\tilde{p}_3(u, \gamma)$: Multimodal PDF with a bimodal marginal in u and a tri-modal marginal in γ .

(4) $\tilde{p}_4(u, \gamma)$: PDF with a highly skewed marginal in u and a bimodal marginal in γ .

(5) $\tilde{p}_5(u, \gamma)$: PDF with a weakly skewed marginal in u and a highly skewed marginal in γ .

(6) $\tilde{p}_6(u, \gamma)$: Multimodal PDF with a Gaussian marginal in u and a tri-modal marginal in γ .

γ .

(7) $\tilde{p}_7(u, \gamma)$: Multimodal PDF with a bimodal marginal in u and a Gaussian marginal γ .

5.2 Ensemble prediction with model error due to coarse-graining the perfect dynamics

We consider the dynamics of the same non-Gaussian system (3.4) with intermittent instabilities as in §3.1.1 which has the general structure as in (3.1)–(3.2). The wide range of interesting turbulent dynamical regimes (see [3–5, 25]) makes this statistically exactly solvable system an unambiguous tested for studying the effects of model errors introduced through various coarse-grainings of the perfect density $p(u, \gamma, t)$ as discussed in §4. In this section, following the methodology introduced in §4, we focus on the model error arising from two particular coarse-grainings of the perfect model density $p(u, \gamma, t)$:

(1) $p_2(u, \gamma, t)$: Non-Gaussian density obtained through the least-biased conditionally Gaussian approximation of the true conditional densities such that the true density $p(u, \gamma, t)$ and the coarse-grained density $p_2(u, \gamma, t)$ have the same first two conditional moments, i.e., for each fixed γ and t , we set

$$\mathcal{S}(p_2(u | \gamma, t)) = \max_{\mathcal{M}_{N,2} = \mathcal{M}_N} \mathcal{S}(q(u)),$$

where

$$\mathcal{M}_N = \int u^N p(u | \gamma, t) du, \quad \mathcal{M}_{N,2} = \int u^N q(u) du, \quad 0 \leq n \leq 2.$$

Note that, despite the Gaussian approximations for the conditional densities $p_2(u | \gamma, t)$, the coarse-grained joint and marginal densities

$$p_2(u, \gamma, t) = p_2(u | \gamma, t) \pi(\gamma, t), \quad \pi_2(u, t) = \int p_2(u, \gamma, t) d\gamma$$

can be highly non-Gaussian.

(2) $p_G(u, \gamma, t)$: Gaussian approximation of the joint density $p(u, \gamma, t)$. The error in the Gaussian estimators $p_G(u, \gamma, t)$ and $\pi_G(u, t) = \int p_G(u, \gamma, t) d\gamma$, arises from the least-biased approximation of the true non-Gaussian density $p(u, \gamma, t)$, which for each fixed t , maximizes the entropy

$$\mathcal{S}(p_G(u, \gamma, t)) = \max_{\mathcal{M}_{ij,G} = \mathcal{M}_{ij}} \mathcal{S}(q(u, \gamma)),$$

subject to the following moment constraints:

$$\mathcal{M}_{i,j} = \int u^i \gamma^j p(u, \gamma, t) du d\gamma, \quad \mathcal{M}_{i,j,G} = \int u^i \gamma^j q(u, \gamma) du d\gamma, \quad 0 \leq i + j \leq 2.$$

In the above set-up, the conditional approximations p_2 and π_2 represent the best possible (least-biased) estimates for the true joint and marginal densities, given the first two conditional moments. Thus, the errors $\mathcal{P}(p, p_2)$ and $\mathcal{P}(\pi, \pi_2)$ represent the intrinsic information barriers which can not be overcome by models based on utilizing two-moment approximations of the true densities (see Proposition 4.1 in §4).

In Figures 9–11, we show the evolution of the model error (4.11) due to different coarse-grainings in p_2 and p_G in the following three dynamical regimes of the system (3.4) with Gaussian damping fluctuations (see also Figure 8):

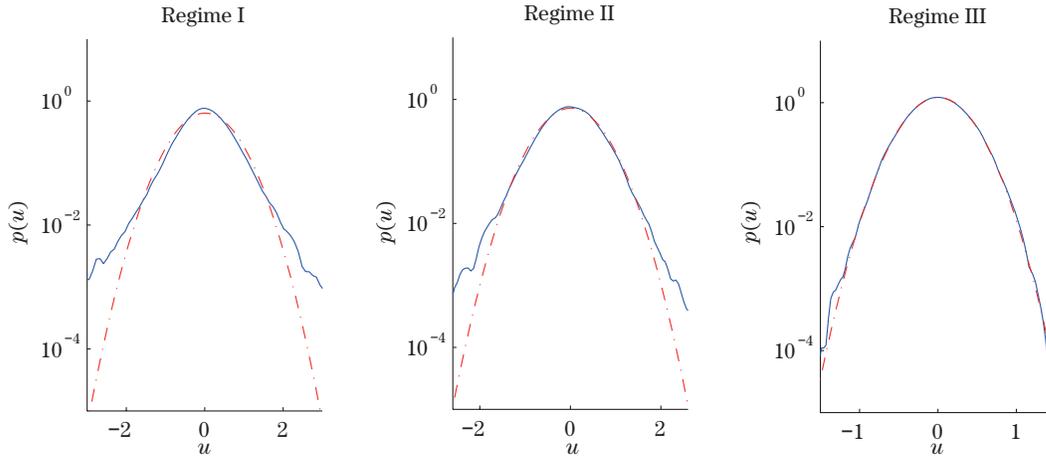


Figure 8 Three dynamical regimes of the non-Gaussian system (3.4) characterized by different equilibrium marginal densities $\pi_{\text{eq}}(u)$ used for studying the model error in coarse-grained densities in §5.2 (see Figures 9–11). Regimes I–II of (3.4) are characterized by intermittent dynamics of u due to transient instabilities induced by the damping fluctuation γ .

Regime I (see Figure 11) A regime with plentiful, short-lasting transient instabilities in the resolved component $u(t)$ with fat-tailed marginal equilibrium densities $\pi(u)$, where, the parameters used in (3.4) are

$$\hat{\gamma} = 2, \quad \sigma_\gamma = d_\gamma = 10, \quad \sigma_u = 1, \quad f_u = 0.$$

Regime II (see Figure 10) A regime with intermittent large-amplitude bursts of the instability in $u(t)$ with fat-tailed marginal equilibrium densities $\pi(u)$, where, the parameters used in (3.4) are

$$\hat{\gamma} = 2, \quad \sigma_\gamma = d_\gamma = 2, \quad \sigma_u = 1, \quad f_u = 0.$$

Regime III (see Figure 9) A regime with nearly Gaussian marginal equilibrium density $\pi(u)$, where, the parameters used in (3.4) are

$$\hat{\gamma} = 7, \quad \sigma_\gamma = d_\gamma = 1, \quad \sigma_u = 1, \quad f_u = 0.$$

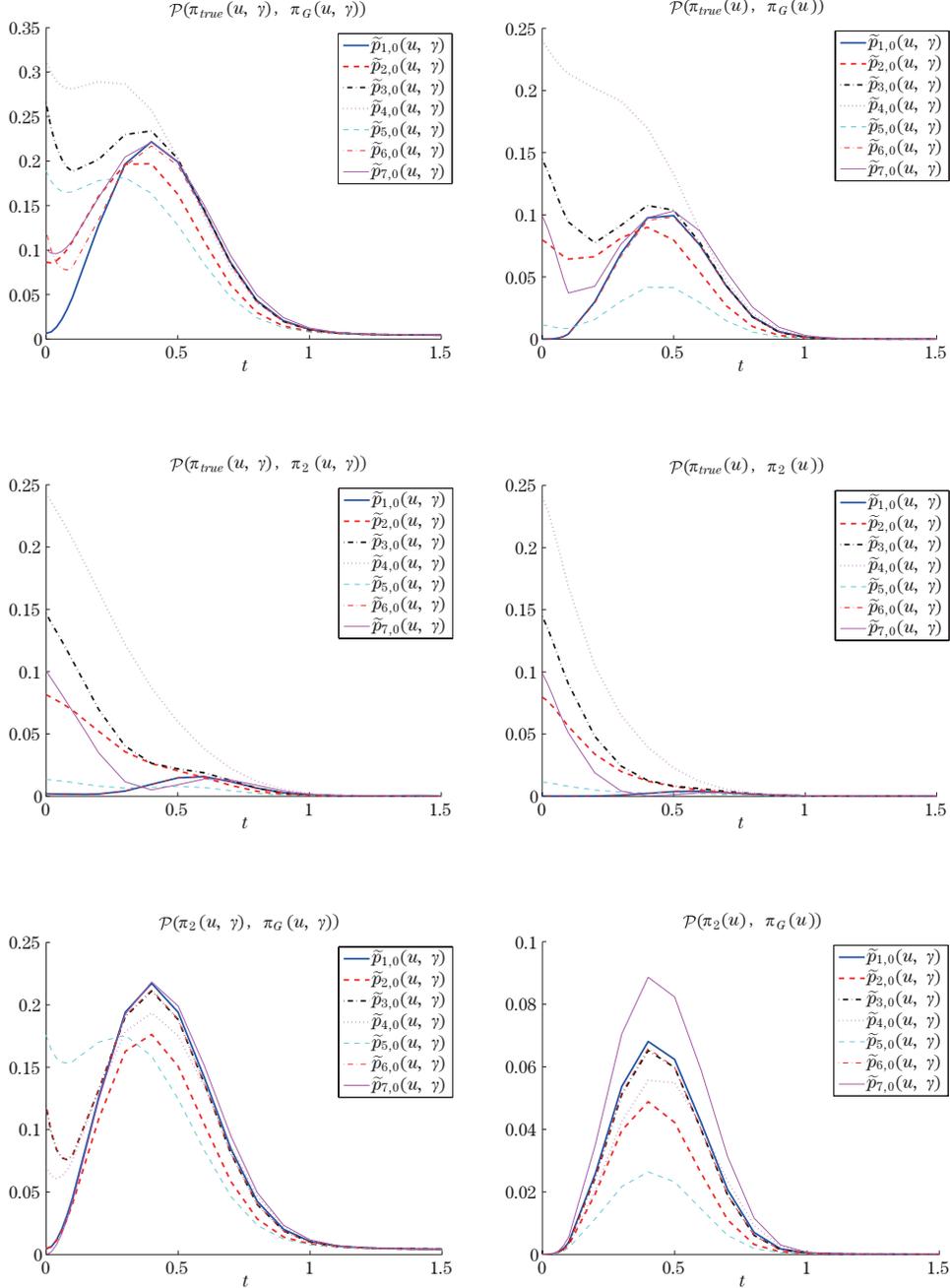


Figure 9 Model errors due to coarse-graining the perfect dynamics of the system (3.4) in the nearly Gaussian regime (Regime III in Figure 8). (Top two rows) Evolution of the model errors (4.11) due to different coarse-grainings of the perfect dynamics in the system (3.4) with Gaussian damping fluctuations. The non-Gaussian joint and marginal densities, p_2 and π_2 , are obtained through the Gaussian coarse-graining of the conditional statistics $p(u | \gamma)$ (see §3–4), while p_G and π_G are the joint and the marginal densities of the Gaussian estimators (see §4). The information barrier (bottom row) equals $\mathcal{P}(p, p_G) - \mathcal{P}(p_2, p_G)$ (see (4.11)). The respective statistical initial conditions, all with the same second-order moments, are described in §5.1 and shown in Figure 7.

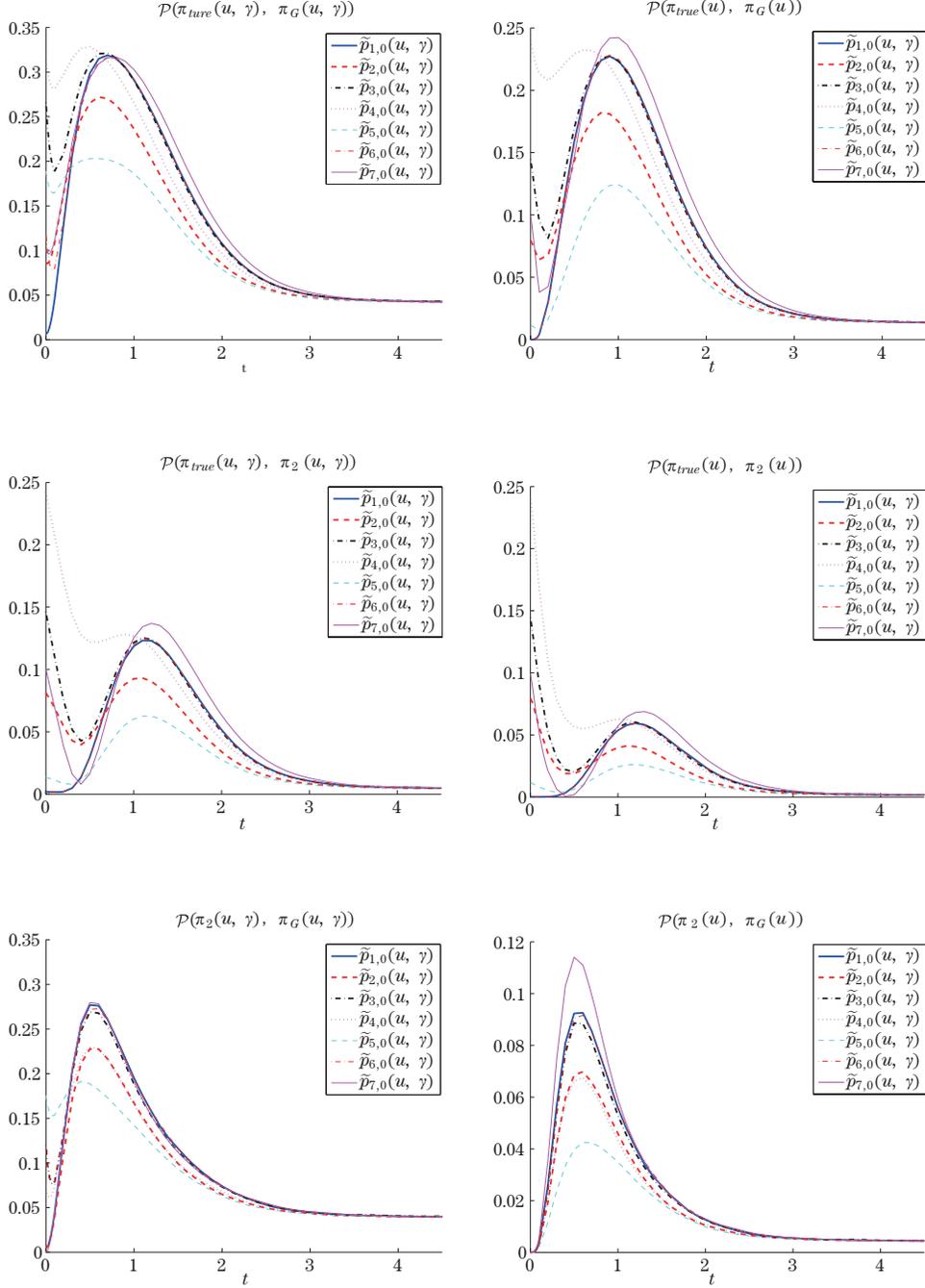


Figure 10 Model errors due to coarse-graining perfect dynamics; the system (3.4) in the regime with intermittent large amplitude instabilities. (Top two rows) Evolution of the model errors (4.11) due to different coarse-grainings of the perfect dynamics in the system (3.4) with Gaussian damping fluctuations. The non-Gaussian joint and marginal densities p_2 and π_2 , are obtained through the Gaussian coarse-graining of the conditional statistics $p(u|\gamma)$ (see §3–4), while p_G and π_G are the joint and the marginal densities of the Gaussian estimators (see §4). The information barrier (bottom row) equals $\mathcal{P}(p, p_G) - \mathcal{P}(p_2, p_G)$ (see (4.11)). The respective statistical initial conditions, all with the same second-order moments, are described in §5.1 and shown in Figure 7.

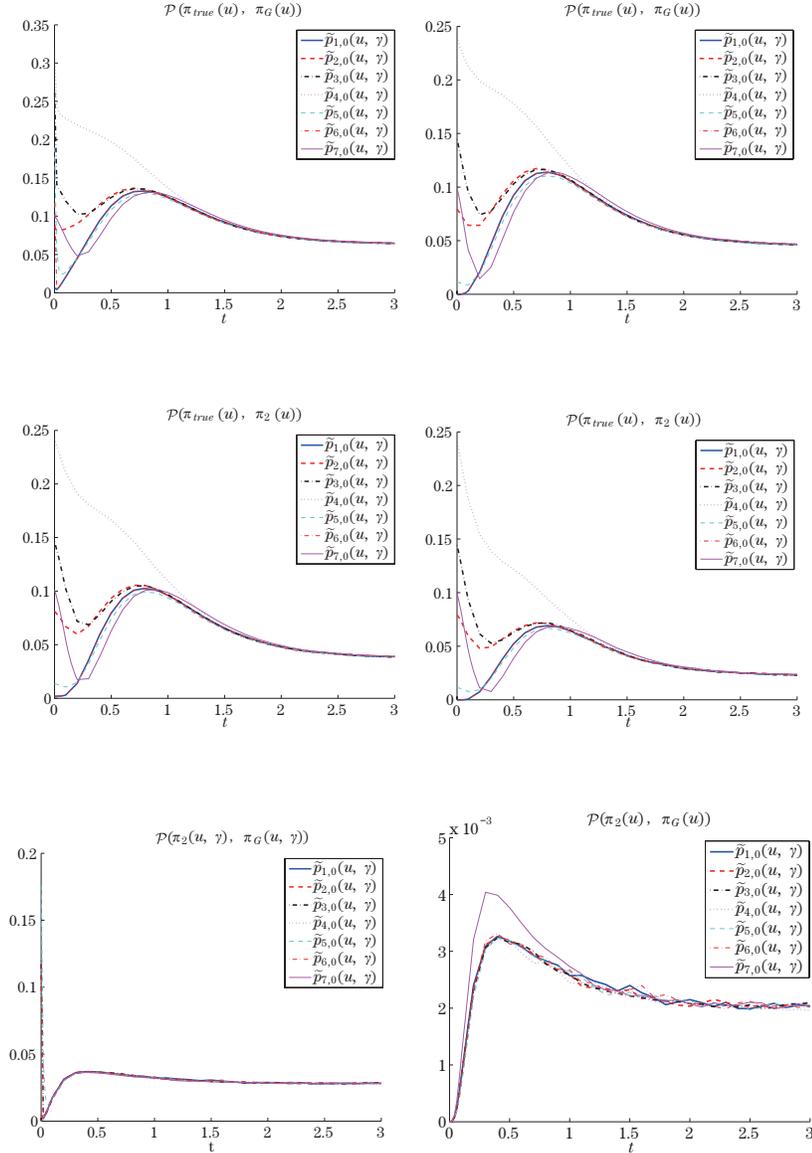


Figure 11 Model errors due to coarse-graining perfect dynamics; the system (3.4) in the regime with abundant transient instabilities. (Top two rows) Evolution of the model errors (4.11) due to different coarse-grainings of the perfect dynamics in the system (3.4) with Gaussian damping fluctuations. The non-Gaussian joint and marginal densities p_2 and π_2 , are obtained through the Gaussian coarse-graining of the conditional statistics $p(u | \gamma)$ (see §3–4), while p_G and π_G are the joint and the marginal densities of the Gaussian estimators (see §4). The information barrier (bottom row) equals $\mathcal{P}(p, p_G) - \mathcal{P}(p_2, p_G)$ (see (4.11)). The respective statistical initial conditions, all with the same second-order moments, are described in §5.1 and shown in Figure 7.

In each regime, the model error in the ensemble predictions is examined for the set of seven different initial densities introduced in §5.1 and Figure 7 with identical second-order statistics. The evolution of the true density $p(u, \gamma, t)$ is estimated via Monte Carlo simulations with 10^7 samples, while the coarse-grained joint densities p_2, p_G , and their marginals π_2, π_G are computed according to the moment-constrained maximum entropy principle in (4.8) using

the conditional moments computed from the CGFPE procedure (2.10).

The top row in Figures 9–11 shows the evolution of the model error in the Gaussian estimators $p_G(u, \gamma, t)$ and $\pi_G(u, t)$ of the true density. The intrinsic information barrier in the Gaussian approximation (see Proposition 4.1), represented by the lack of information in the least-biased density p_2 , based on two conditional moment constraints, is shown for each regime in the middle row. It can be seen in Figures 9–11 that the common feature of the model error evolution in all the examined regimes of (3.4) is the presence of a large error at the intermediate lead times. The source of this phenomenon is illustrated in Figure 12 in Regime III of (3.4) with nearly Gaussian attractor statistics. The large error arises from the presence of a robust transient phase of fat-tailed dynamics in the system (3.4) which is poorly captured by the coarse grained statistics.

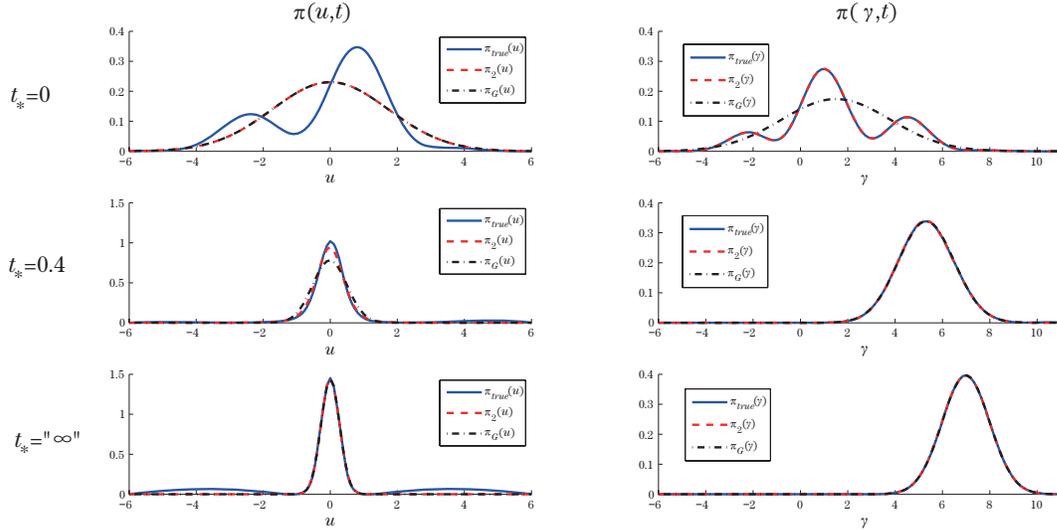


Figure 12 Three distinct stages in the statistical evolution of the system (3.4) illustrated for the regime with nearly Gaussian dynamics and highly non-Gaussian multimodal initial statistical conditions $\tilde{p}_3(u, \gamma)$ (see Figure 7). These three stages exist regardless of the dynamical regime of (3.4) and the form of the initial conditions (not shown). (Top) The initial configuration projected on the marginal densities at $t_* = 0$. (Middle) The fat-tailed phase in the marginal $\pi(u, t)$ corresponding to the large error phase in the coarse grained models (see Figures 9–11), (Bottom) Equilibrium marginal statistics on the attractor in the regime with nearly Gaussian statistics (see Regime III in Figure 8).

Below, we summarize the results illustrated in Figures 9–12 with the focus on the model error in the Gaussian approximations $p_G(u, \gamma, t)$ and $\pi_G(u, t)$:

(1) For both the Gaussian estimators $p_G(u, \gamma)$, $\pi_G(u)$ and the conditionally Gaussian estimators $p_2(u, \gamma)$, $\pi_2(u)$, there exists a phase of large model errors at intermediate lead times. This phase exists in all the examined regimes of (3.4) irrespective of the initial conditions, and it arises due to a transient highly non-Gaussian fat-tailed dynamical phase in (3.4) which the Gaussian estimators fail to capture.

(2) The trends in the model error evolution for the joint and the marginal densities are similar. This is to be expected based on Proposition 4.1.

(3) The contributions to the model error in the Gaussian estimators $p_G(u, \gamma)$ and $\pi_G(u)$ from the intrinsic information barrier $\mathcal{P}(p, p_2)$ (see Proposition 4.1), and from the error $\mathcal{P}(p_2, p_G)$ due to the fully Gaussian vs conditionally Gaussian approximations depend on the dynamical regime.

(a) The effects of the intrinsic information barrier are the most pronounced in the non-Gaussian Regime I of (3.4) with abundant transient instabilities in u (see Figures 8 and 11). In this regime, the information barrier dominates the total model errors. In the nearly Gaussian regime, the intrinsic information barrier is negligible except at short times due to the errors in coarse-graining the highly-non-Gaussian initial conditions (see Figures 7 and 9).

(b) In the highly non-Gaussian Regime I with abundant instabilities and the fat-tailed equilibrium PDFs (see Figure 8), the differences in the model error between different initial conditions quickly become irrelevant. The intrinsic information barrier dominates the model error, and there is a significant error for long range predictions in both the joint and the marginal coarse-grained densities.

(c) In the non-Gaussian Regime II of (3.4) with large amplitude intermittent instabilities, the intrinsic information barrier dominates the error in the Gaussian estimators at short ranges. At intermediate lead times, the error due to the fully Gaussian vs conditionally Gaussian approximations exceeds the intrinsic barrier. The error at long lead times is significantly smaller than those in Regime I with comparable contributions from $\mathcal{P}(p, p_2)$ and $\mathcal{P}(p_2, p_G)$.

(d) In the nearly Gaussian Regime III of (3.4), the intrinsic information barrier in the Gaussian estimators is small and dominated by the errors in coarse-graining the non-Gaussian initial conditions.

(4) The intrinsic information barriers in the joint density $\mathcal{P}(p, p_2)$ and in the marginal density $\mathcal{P}(\pi, \pi_2)$, are comparable throughout the evolution and almost identical at short lead times.

5.3 Ensemble prediction with model errors due to imperfect dynamics

We focus on the model error which arises through common approximations associated with the ensemble prediction: (i) Errors due to imperfect/simplified dynamics, and (ii) errors due to coarse-graining the statistics of the perfect system which is used for tuning the imperfect models. While the above two approximations are often simultaneously present in applications and are generally difficult to disentangle, it is important to understand the effects of these two contributions in a controlled environment which is developed below.

Similar to the framework used in the previous sections, we consider the dynamics with the structure as in the test models (3.1)–(3.2), where the non-Gaussian perfect system, as in (3.5), is given by

$$\begin{aligned}
 \text{(a)} \quad d\gamma &= [-a\gamma + b\gamma^2 - c\gamma^3 + f_\gamma(t)]dt + (A - B\gamma)dW_C + \sigma_\gamma dW_\gamma, \\
 \text{(b)} \quad du &= (-\gamma u + f_u(t))dt + \sigma_u dW_u,
 \end{aligned}
 \tag{5.2}$$

with cubic nonlinearity in the damping fluctuations γ . The imperfect non-Gaussian model

introduces errors by assuming Gaussian dynamics in the damping fluctuations, as in (3.4),

$$\begin{aligned} \text{(a)} \quad d\gamma^M &= (-d_\gamma^M(\gamma^M - \hat{\gamma}^M) + f_\gamma^M(t))dt + \sigma_\gamma^M dW_\gamma^M, \\ \text{(b)} \quad du^M &= (-\gamma^M u^M + f_u^M(t))dt + \sigma_u^M dW_u^M. \end{aligned} \quad (5.3)$$

The imperfect model (5.3) is optimized by tuning its marginal attractor statistics, in either u^M or γ^M depending on the context, to reproduce the respective true marginal statistics. This is a prototype problem for a number of important issues. Two topical examples are:

(1) Reduced models with a subset of unresolved variables (here γ^M) whose statistics is tuned for statistical fidelity in the resolved variables (here u^M).

(2) Simplification of parts of the dynamics in complex multi-component models such as the coupled atmosphere-ocean-land models in climate science; in the present toy-model setting γ can be regarded as the atmospheric forcing of the ocean dynamics u .

In order to illustrate the framework developed in §2–4, we compare the model error arising in the optimized imperfect statistics, $p^{M*}(u, \gamma, t)$ or $\pi^{M*}(u, t)$, associated with (5.3) with the model error in $p_2(u, \gamma, t)$ or $\pi_2(u, t)$ due to the Gaussian coarse-graining of the conditional density $p(u | \gamma, t)$ of the perfect system (5.2) using the CGFPE framework of §2.

In particular, we show that a small model error can be achieved at medium and long lead times for imperfect predictions of the marginal dynamics $\pi^{M*}(u)$ using models with tuned unresolved dynamics γ despite a large model error in the joint density $p^{M*}(u, \gamma)$.

5.3.1 Ensemble predictions with imperfect dynamics and time-independent statistics on the attractor

We consider the perfect system (5.2) and its model (5.3) with invariant measures at their respective equilibria. This configuration is achieved by assuming constant forcing $f_\gamma=0.8220$, $f_u=-0.5$, $f_\gamma^M=0$, $f_u^M=-0.5$ in both (5.2) and (5.3). We first examine the effects of model errors associated with two distinct ways of optimizing the imperfect model (5.3):

(I) Tuning the marginal equilibrium statistics of the damping fluctuations γ^M in (5.3) for fidelity to the true statistics of γ in (5.2).

In order to tune the mean and variance of γ^M to coincide with the true moments, we simply set

$$\hat{\gamma}^M = \langle \gamma \rangle_{\text{eq}}, \quad \frac{\sigma_{\gamma^M}^2}{2d_\gamma^M} = \text{Var}_{\text{eq}}(\gamma), \quad (5.4)$$

which leads to a one-parameter family of models in (5.3) with a correct marginal equilibrium density in γ^M . Below, we choose the damping d_γ^M in (5.3) as the free parameter and study the dependence of model errors in the class of models satisfying (5.4) and parameterized by the damping/decorrelation dime in γ^M (see Figure 13). Note that only one model in this family can match both the equilibrium density $\pi(\gamma)$ and the decorrelation time $\tau_\gamma = \int \text{Corr}_\gamma(\tau) d\tau$, of the true damping fluctuations in (5.2). For such a model we have, in addition to (5.4),

$$\tau_{\gamma^M}^M = \frac{1}{d_{\gamma^M}^M} = \tau_\gamma. \quad (5.5)$$

Examples of the prediction error in models of (5.3) optimized for equilibrium fidelity in γ^M but different dampings d_γ^{M*} are shown in Figure 13 for the two-state unimodal regime of (5.2) (see Figure 1). We highlight two important observations here:

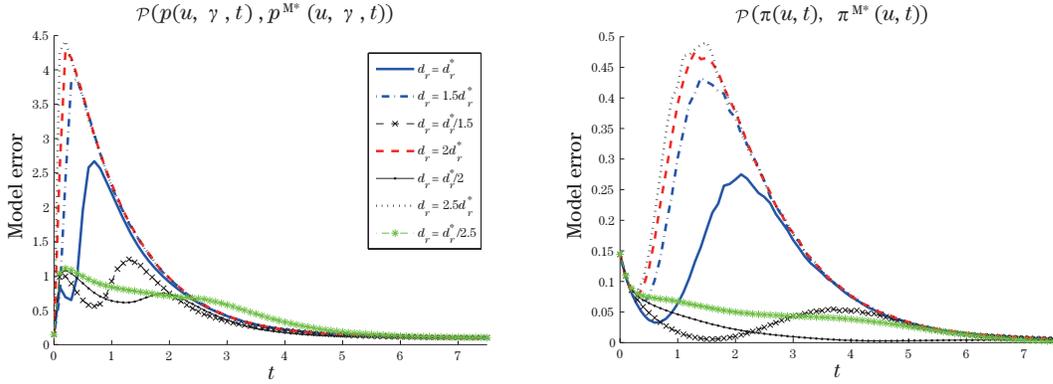


Figure 13 the ensemble prediction of (5.2) with imperfect models in (5.3); dependence of model error on the decorrelation time in the imperfect model. The model error (4.11) via the relative entropy for the imperfect prediction of the system (5.2) using imperfect models in (5.3) with the correct climatology in γ^M but different decorrelation times of the damping fluctuations. Note that in this case underdamped imperfect models have the best medium range prediction skill. The results shown are obtained in the skewed two-state unimodal regime (see Figure 1) of (5.2), starting from the statistical initial condition $\tilde{p}_1(u, \gamma)$ (see §5.1 and Figure 7).

(1) Underdamped models of (3.4) optimized for equilibrium fidelity in the damping fluctuations γ^M have the smallest error for medium range forecasts (all models are comparable for long range forecasts). These results are similar to those reported recently in [25], where the short and medium range predictive skills of linear models with optimized marginal statistics of the unresolved dynamics were shown to often exceed the skills of models with correct marginal statistics and decorrelation time.

(2) Despite the striking reduction in the model error at intermediate lead times achieved through underdamping the unresolved dynamics in (3.4), caution is needed when tuning imperfect models for short range forecasts or forced response predictions, where the damping, in both the resolved and unresolved dynamics, is relevant for correct system responses (see [25]).

(II) Tuning the marginal equilibrium statistics of the damping fluctuations γ^M in (5.3) for fidelity to the true statistics of u in (5.2).

This case corresponds to the situation in which we construct a simplified model of a system with unresolved degrees of freedom (here γ); these stochastically ‘superparameterized’ unresolved dynamics are then tuned to correctly reproduce the statistical features of the resolved dynamics (here u).

We consider this optimization in the Gaussian framework and optimize the imperfect model (5.3) by tuning the dynamics of the damping fluctuations γ^M in order to minimize the lack of information in the imperfect marginal density for the resolved variable, i.e., the optimal imperfect model satisfies

$$\mathcal{P}(\pi_G(u), \pi_G^{M*}(u)) = \min_{d_\gamma^M, \sigma_\gamma^M, \tilde{\gamma}^M} \mathcal{P}(\pi_G(u), \pi_G^M(u)), \quad (5.6)$$

where π_G and π_G^M are the Gaussian estimators of the respective marginal densities associated with (5.2) and (5.3), respectively. With the conditional moments of u in the perfect system (5.2), $M_1(\gamma)$ and $M_2(\gamma)$ obtained by solving (2.10) in the CGFPE framework in §2, the mean and the variance of $p_G(u)$ are given by

$$\bar{u} = \int \mathcal{M}_1(\gamma) d\gamma, \quad R_u = \int \mathcal{M}_2(\gamma) d\gamma - \bar{u}^2, \quad (5.7)$$

respectively. Analogous expressions hold for the mean and the variance of $p_G^M(u)$ which are used in the optimization (5.6).

The two types of model optimization are compared in Figure 14 for the two-state unimodal regime of (5.2) (see Figure 1). Both procedures yield comparably good results at long lead times when the model error in the marginal densities in $\pi^{M^*}(u, t)$ is considered. Unsurprisingly, optimizing the marginal dynamics of u^M by tuning the dynamics of γ^M generally leads to a smaller model error for short and medium range predictions. But the type of the optimization largely depends on the applications.

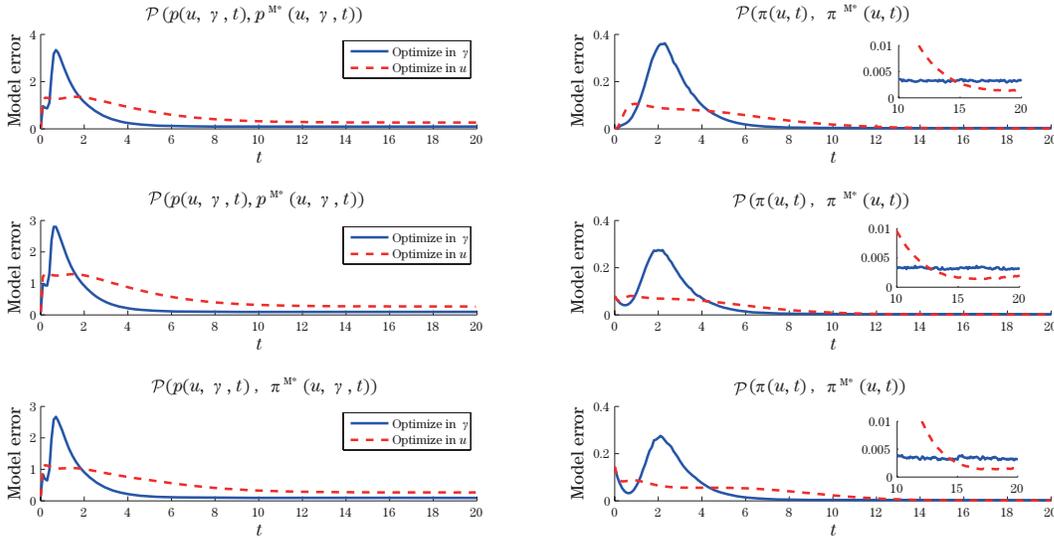


Figure 14 The ensemble prediction of (5.2) with imperfect models in (5.3); comparison of model errors for different types of model optimization. Evolution of the model error (4.11) via the relative entropy for imperfect models in (5.3) where the imperfect dynamics of the damping fluctuations γ^M , is either (I) tuned to correctly reproduce the marginal equilibrium statistics of γ , or (II) tuned to correctly reproduce the marginal equilibrium statistics of u in (5.2). The results shown are obtained for the perfect dynamics in (5.2) in the regime with skewed unimodal statistics and the two-state switching in the path-wise dynamics (see Figure 1), and for three different statistical initial conditions: (top) the initial density $\tilde{p}_1(u, \gamma)$, (middle) the initial density $\tilde{p}_2(u, \gamma)$, (bottom) the initial density $\tilde{p}_3(u, \gamma)$ (see also Figure 7 and §5.1).

In Figures 15–18, we illustrate the evolution of the model error in the imperfect statistical prediction of (5.2) which is optimized according to the procedure (I) above. Two non-Gaussian regimes of the true system (5.2) illustrated in Figure 1 are used to analyze the error in imperfect predictions with optimized models in (5.3).

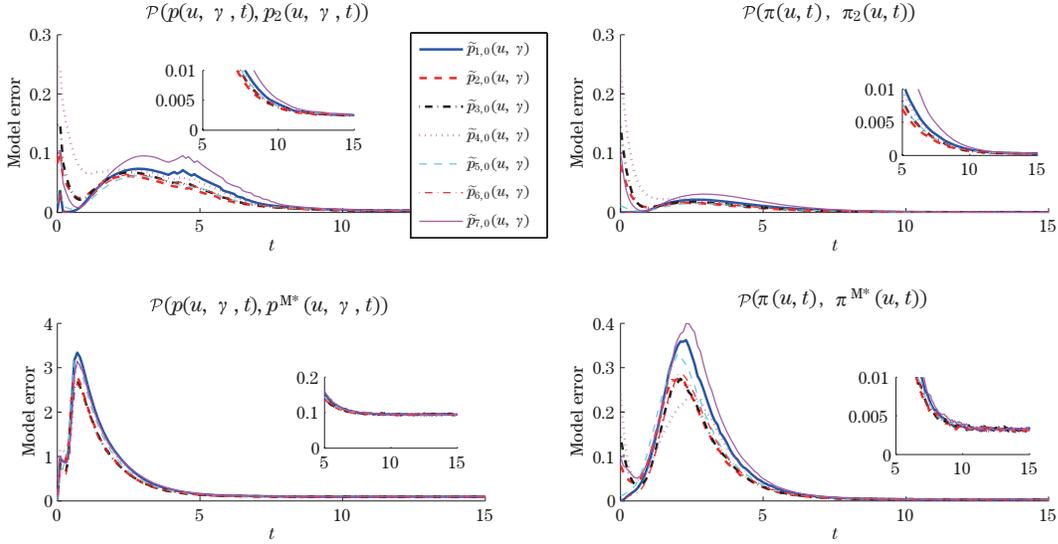


Figure 15 The ensemble prediction with optimized imperfect dynamics; the perfect system (5.2) with skewed unimodal statistics and the regime switching, imperfect model given by (5.3). Comparison of two types of model errors in ensemble predictions: (top row) the model error (4.11) due to coarse-graining the perfect conditional statistics (see §4), and (bottom row) the model error due to imperfect dynamics in (5.3) where γ^M is tuned for the correct marginal equilibrium statistics and the correlation time of the damping fluctuations γ in (5.2). The model error via the relative entropy 4.11 is shown for the joint densities (left column) and the marginal densities in u (right column). The respective initial conditions are shown in Figure 7.

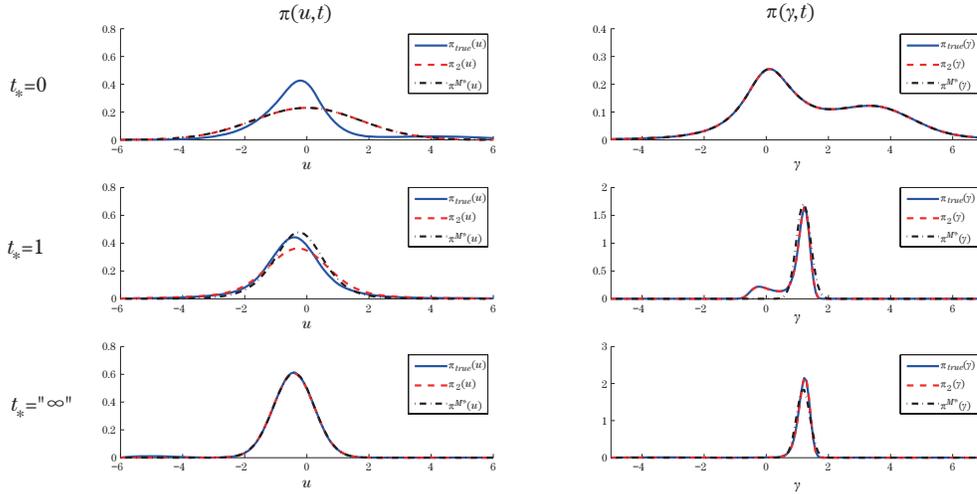


Figure 16 Three distinct stages in the statistical evolution of the system (5.2) and its imperfect models (5.3) with different contributions to model errors; the example shown corresponds to the evolution from the initial condition \tilde{p}_3 (see §5.1) in the regime with time-invariant statistics at the equilibrium with unimodal PDFs and the regime switching (see Figure 1). (Top) The initial configuration at $t_* = 0$. (Middle) The fat-tailed phase in the true marginal $\pi(u, t)$ corresponding to the large error phase in the coarse-grained and the Gaussian models (see Figure 15). (Bottom) Equilibrium marginal statistics on the attractor with the skewed marginals $\pi(\gamma)$ and $\pi(\gamma^M)$ of the damping fluctuations.

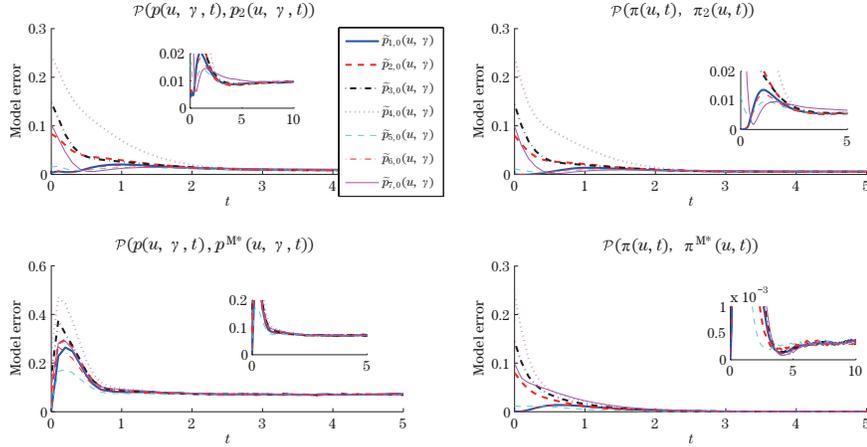


Figure 17 The ensemble prediction with optimized imperfect dynamics; the perfect system (5.2) with fat-tailed statistic, imperfect model given by (5.3). Comparison of two types of model errors in ensemble predictions: (Top row) The model error due to coarse-graining the perfect dynamics (5.2), and (bottom row) the model error due to imperfect dynamics (5.3), where γ^M is tuned for the correct marginal equilibrium statistics and teh correlation time of the damping fluctuations γ in (5.2). The model error via the relative entropy (4.11) is shown for the joint densities (left column) and the marginal densities in u (right column). The respective initial conditions are shown in Figure 7.

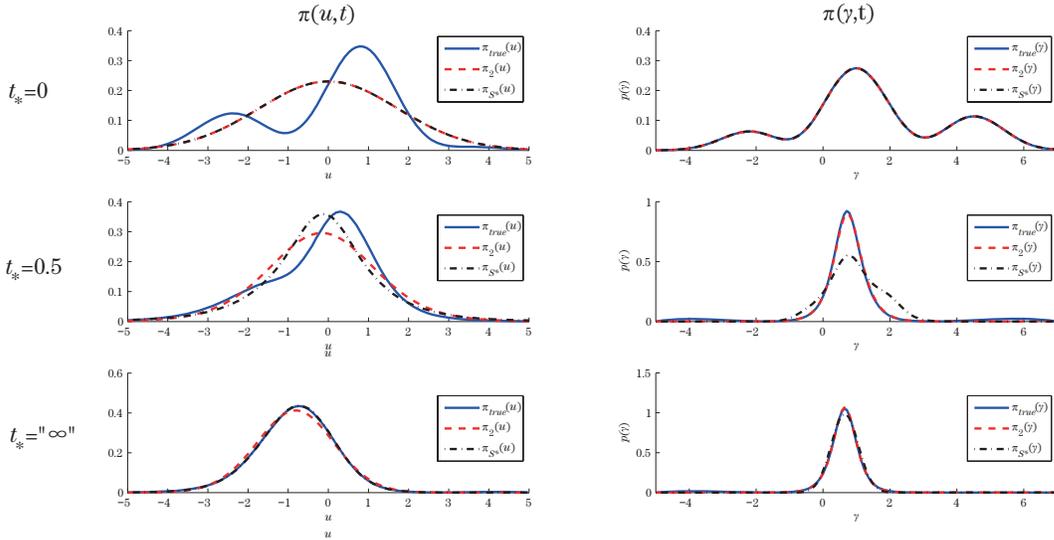


Figure 18 Three distinct stages in the statistical evolution of the perfect system (5.2) and its imperfect models in (5.3) with different contributions to the model error; the example shown corresponds to the evolution from the initial condition \tilde{p}_3 (see §5.1) in the regime with time-invariant statistics at the equilibrium and fat-tailed PDFs (see Figure 1). (Top) The initial configuration at $t_* = 0$. (Middle) The fat-tailed phase in the true marginal $\pi(u, t)$ corresponding to the large error phase in the coarse-grained and the Gaussian models (see Figure 15). (Bottom) Equilibrium marginal statistics on the attractor with the fat-tailed marginals $\pi(\gamma)$ and $\pi(\gamma^M)$ of the damping fluctuations.

5.3.2 Ensemble predictions with imperfect dynamics and time-periodic statistics on the attractor

We finish the analysis by considering the dynamics of the perfect system (5.2) and its model (5.3) with time-periodic statistics on the attractor. We focus on the highly non-Gaussian regime of the perfect system (5.2) with the cubic nonlinearity in the damping fluctuations periodic transitions between the nearly Gaussian and highly skewed marginal densities in the damping fluctuations γ which are induced by the simple time-periodic forcing

$$f_\gamma(t) = f_{\gamma,0} + f_{\gamma,1} \sin(\omega t + \phi).$$

This regime was previously used in §3.1.1 to validate the CGFPE framework (see Figure 2). Similar to the configurations studied with time-independent equilibrium statistics in the previous section, we are interested in the differences between the model error arising in the optimized imperfect dynamics $p^{M*}(u, \gamma, t)$ and $\pi^{M*}(u)$, and the error due to coarse-graining the perfect statistics in the densities $p_2(u, \gamma, t), \pi_2(u)$ obtained through the Gaussian approximations of the conditionals $p(u | \gamma, t)$.

The issue of tuning the marginal attractor statistics of the damping fluctuations γ^M in the imperfect model (5.3) requires more care than in the case with time-independent equilibrium statistics; this is due to the presence of an intrinsic information barrier (see §4 or [5, 25]) when tuning the statistics of the Gaussian damping fluctuations γ^M in (5.3) to the true statistics of (5.2) in γ . Similar to the time-independent case, we aim at tuning the marginal attractor statistics in γ^M for best fidelity to the true marginal statistics in γ . However, there exists an information barrier associated with the fact that the attractor variance of the Gaussian fluctuations γ^M is always constant regardless of the forcing $f_\gamma^M(t)$. One way to optimize the imperfect statistics of γ is to tune its decorrelation time, and time-averaged mean and variance on the attractor to reproduce the true time-averaged quantities. However, such an approach is clearly insensitive to phase variations of the respective statistical moments. Here, instead, we optimize the imperfect model by first tuning the decorrelation times of γ^M and γ and then minimizing the period-averaged relative entropy between the marginal densities for the damping fluctuations, i.e., the optimized model (5.3) satisfies

$$\overline{\mathcal{P}(\pi_G(\gamma, t), \pi_G^{M*}(\gamma, t))} = \min_{\sigma_\gamma^M, \{f_\gamma^M\}} \overline{\mathcal{P}(\pi_G(\gamma, t), \pi_G^M(\gamma, t))}, \quad (5.8)$$

where the overbar denotes the temporal average over one period, and $\{f_\gamma^M\}$ denotes a set of parameters in the forcing f_γ^M in (5.3). In the examples below we assume that the form of the forcing f_γ^M with the same time dependence on the true one, i.e.,

$$f_\gamma^M(t) = f_{\gamma,0}^M + f_{\gamma,1}^M \sin(\omega^M t + \phi^M) \quad \text{with } \omega^M = \omega, \quad \phi^M = \phi,$$

so that the optimization in (5.8) is carried out over a three-parameter space $\{\sigma_\gamma^M, f_{\gamma,0}, f_{\gamma,1}\}$ (the optimization in the phase and the frequency are often crucial and interesting, but we skip the discussions for the sake of brevity).

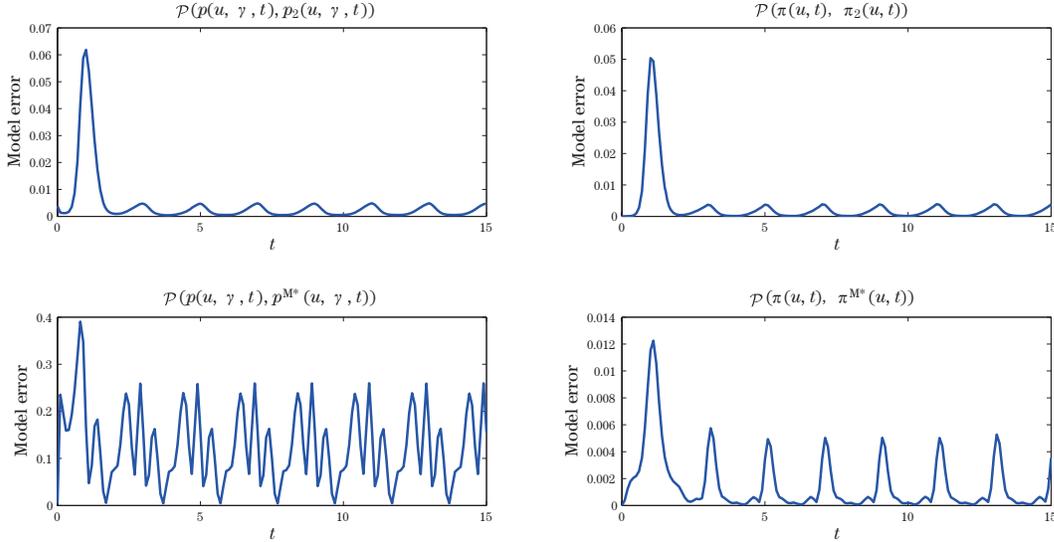


Figure 19 The model error in imperfect optimized ensemble predictions of non-Gaussian systems with time-periodic statistics; the perfect model (5.2) with attractor statistics nearly Gaussian \longleftrightarrow high skewness in γ (see Figure 2). Evolution of the model error (4.11) associated with the statistical prediction of (5.2) in the highly non-Gaussian regime with time-periodic statistics using two non-Gaussian models: (Top row) Models with a coarse-grained perfect conditional density $p_2(u | \gamma)$ (see §4), and (bottom row) models with imperfect dynamics of the damping fluctuations, γ^M in (5.3) which are optimized by matching the decorrelation time of γ and minimizing the period-averaged relative entropy (see §5.3.2 and §4 for details).

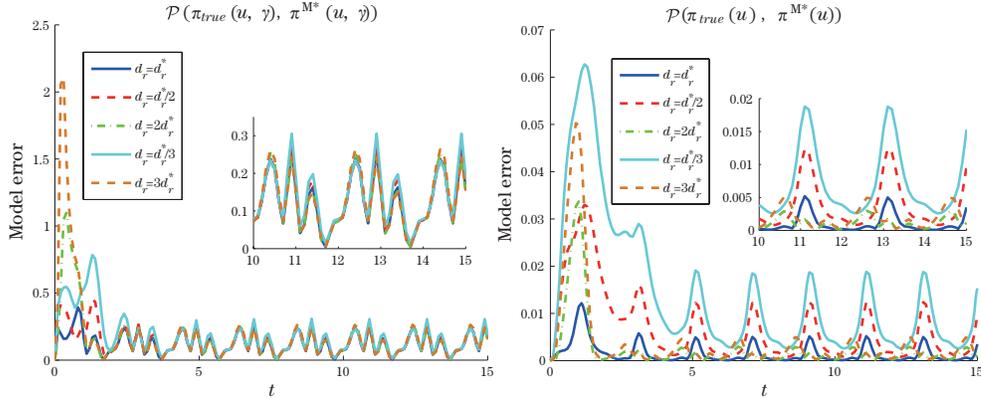


Figure 20 Dependence on the model error on decorrelation time in imperfect optimized ensemble predictions of non-Gaussian systems with time-periodic statistics; the perfect model (5.2) and its attractor statistics as in Figure 19. The evolution of the model error (4.11) for imperfect predictions of the true dynamics (5.2) using the models in (5.3) with different decorrelation times of damping fluctuations γ^M . $\tau_\gamma = \frac{1}{d_\gamma}$ denotes the decorrelation time of γ in the true dynamics (5.2). For a given decorrelation time $\frac{1}{d_\gamma^*}$, the model (5.3) is optimized in the remaining parameters by minimizing the period-averaged relative entropy $\overline{\mathcal{P}(p(u, \gamma, t), p^{M^*}(u, \gamma, t))}$ (see §5.3.2 and §4 for details).

In Figures 19–20, we show the model error for the coarse-grained joint and marginal densities p_2, π_2 , and compare them with the model error in the joint and marginal densities associated

with the optimized imperfect model (5.3). Here, the parameters used in (5.2) are

$$\begin{aligned} a = 1, \quad b = 1, \quad c = 1, \quad A = 0.5, \quad B = -0.5, \quad \sigma = 0.5, \quad \sigma_u = 1, \\ f_u = -0.5, \quad f_{\gamma,0} = 2.5, \quad f_{\gamma,1} = 6.5, \quad \omega = \pi, \quad \phi = -\frac{\pi}{2}. \end{aligned} \quad (5.9)$$

In Figure 19, the decorrelation time $\tau^{\text{M}^*} = \frac{1}{d_{\gamma}^{\text{M}^*}}$ of the damping fluctuations γ^{M} is the same as the one in the true dynamics while the results shown in Figure 20 illustrate the dependence of the model error in the optimized imperfect model on the decorrelation time (see also Figure 13 for the configuration with time-independent equilibrium statistics).

The following points summarize the results of §5.3.1 and §5.3.2:

(1) A small model error can be achieved at medium and long lead times for imperfect predictions of the marginal dynamics $\pi^{\text{M}^*}(u)$ using models with tuned unresolved dynamics γ despite a large model error in the joint density $p^{\text{M}^*}(u, \gamma)$ (see Figures 13–15, 17, 19–20).

(2) The error in the coarse-grained densities $p_2(u, \gamma, t)$, $\pi_2(u, t)$ is much smaller than that in the optimized models with imperfect dynamics with $p^{\text{M}^*}(u, \gamma, t)$, $\pi^{\text{M}^*}(u, t)$ (see Figures 15–18).

(3) The largest error in the optimized models (3.4) is associated with the presence of transient multimodal phases which can not be captured by the imperfect models in the class (3.4) (see Figures 15–19).

(4) At long lead times, the model error in the joint density $\mathcal{P}(p(u, \gamma, t), p^{\text{M}^*}(u, \gamma, t))$, is largely insensitive to the variation of the damping $d_{\gamma}^{\text{M}^*}$ (see Figure 20).

(5) The model error in the marginal densities $\pi^{\text{M}^*}(u, t)$ of the optimized models has non-trivial dependence on the decorrelation time $\frac{1}{d_{\gamma}^{\text{M}^*}}$ of the damping fluctuations. The overall trend is that underdamped imperfect models have smaller errors in the marginals $\pi^{\text{M}^*}(u, t)$ for the constant or slow forcing, while the overdamped imperfect models are better for the strongly varying forcing (see Figures 13 and 20 for two extreme cases).

6 Concluding Discussion

We consider a class of statistically exactly solvable non-Gaussian test models where the generalized Feynman-Kac formulation developed here reduces the exact behavior of conditional statistical moments to the solution to inhomogeneous Fokker-Planck equations modified by linear lower order coupling and source terms. This procedure is applied to test models with hidden instabilities and is combined with information theory to address two important issues in contemporary statistical predictions of turbulent dynamical systems: The coarse-grained ensemble prediction in a perfect model and the improving long range forecasting in imperfect models. Here, the focus is on studying these model errors in conditionally Gaussian approximations of the highly non-Gaussian test models. In particular, we show that in many turbulent non-Gaussian dynamical regimes, a small model error can be achieved for imperfect medium and long range forecasts of the resolved variables using models with appropriately tuned statistics of the unresolved dynamics. The framework developed here, combining the generalized Feynman-Kac approach with information theory, also allows for identifying dynamical regimes with information barriers and/or transient phases in the non-Gaussian dynamics, where the imperfect models fail to capture the characteristics of the true dynamics. The techniques and

models developed here should be useful for quantifying and mitigating the model error in filtering and prediction in a variety of other contexts. These applications will be developed by the authors in the near future.

7 Appendix

Appendix A: The numerical scheme for solving the CGFPE system (2.10)

Here, we outline the numerical method for solving the CGFPE system in (2.10) in one spatial dimension. This is achieved by combining the third-order backward differentiation formulas [17] with the method of (see [19]) and the second-order, finite-volume representation for (2.10).

Recall that the CGFPE system consists of a hierarchy of inhomogeneous Fokker-Planck equations for the conditional moments $\mathcal{M}_N(\gamma, t)$ with the forcing terms depending linearly on $\mathcal{M}_N(\gamma, t)$ and inhomogeneities depending linearly on $\mathcal{M}_{N-i}(\gamma, t)$, $i > 1$. Thus, due to the form of (2.10), the linearity of the forcing and inhomogeneities, we outline here the present algorithm applied to the homogeneous Fokker-Planck part of (2.10), written in the conservative form

$$\frac{\partial \pi}{\partial t} = -\frac{\partial}{\partial \gamma} \left[\left(F - \frac{1}{2} G \gamma \right) \pi - \frac{1}{2} G \pi \gamma \right], \quad (7.1)$$

where $\pi(\gamma, t) = \int p(u, \gamma, t) du$ and $G(\gamma, t) = \tilde{\sigma}^2(\gamma, t)$. Given the spatial grid with nodes γ_i , $i = 1, \dots, N$, the uniform spacing $\Delta\gamma$, and the approximation

$$Q_i(t) \equiv \frac{1}{\Delta\gamma} \int_{\gamma_{i-\frac{1}{2}}}^{\gamma_{i+\frac{1}{2}}} \pi(\gamma, t) d\gamma, \quad (7.2)$$

we discretize (7.1) in space through the second-order finite volume formula as

$$\begin{aligned} \frac{dQ_i}{dt} = & -\frac{1}{\Delta\gamma} \left[\left(F - \frac{1}{2} G \gamma \right)_{i+\frac{1}{2}} \left(\frac{9}{16} Q_i + \frac{9}{16} Q_{i+1} - \frac{1}{16} Q_{i-1} - \frac{1}{16} Q_{i+2} \right) \right. \\ & \left. - \left(F - \frac{1}{2} G \gamma \right)_{i-\frac{1}{2}} \left(\frac{9}{16} Q_{i-1} + \frac{9}{16} Q_i - \frac{1}{16} Q_{i-2} - \frac{1}{16} Q_{i+1} \right) \right] \\ & + \frac{1}{2} \frac{1}{\Delta\gamma} \left[G_{i+\frac{1}{2}} \left(-\frac{9}{8} Q_i + \frac{9}{8} Q_{i+1} + \frac{1}{24} Q_{i-1} - \frac{1}{24} Q_{i+2} \right) \right. \\ & \left. - G_{i-\frac{1}{2}} \left(-\frac{9}{8} Q_{i-1} + \frac{9}{8} Q_i + \frac{1}{24} Q_{i-2} - \frac{1}{24} Q_{i+1} \right) \right]. \end{aligned} \quad (7.3)$$

The above expression is obtained by seeking higher order interpolants for $Q_{i+\frac{1}{2}}^{n+1}$ in the standard finite-volume formulation

$$\begin{aligned} \frac{dQ_i}{dt} = & -\frac{1}{\Delta\gamma} \left[\left(F - \frac{1}{2} G \gamma \right)_{i+\frac{1}{2}} Q_{i+\frac{1}{2}}^{n+1} - \left(F - \frac{1}{2} G \gamma \right)_{i-\frac{1}{2}} Q_{i-\frac{1}{2}}^{n+1} \right] \\ & + \frac{1}{2\Delta\gamma} [G_{i+\frac{1}{2}} Q_{i+\frac{1}{2}}^{n+1} - G_{i-\frac{1}{2}} Q_{i-\frac{1}{2}}^{n+1}]. \end{aligned} \quad (7.4)$$

The second order approximations for $Q_{i+\frac{1}{2}}^{n+1}$ are obtained by determining the coefficients a, b, c, d in the expansion

$$\tilde{Q}_{i+\frac{1}{2}} = a Q_i + b Q_{i+1} + c Q_{i-1} + d Q_{i+2},$$

such that $\tilde{Q}_{i+\frac{1}{2}} - Q_{i+\frac{1}{2}}$ is of order $O((\Delta\gamma)^3)$.

The time discretization of (7.1) or (2.10) is obtained by using the three-step backward differentiation formula (BDF3) (see [17]), which belongs to the family of linear multistep methods. In particular, (7.1) is discretized in time as follows

$$Q^{n+3} - \frac{18}{11}Q^{n+2} + \frac{9}{11}Q^{n+1} - \frac{2}{11}Q^n = \frac{6}{11}\Delta t f(Q^{n+3}). \quad (7.5)$$

The above implicit formulation can be solved explicitly due to the linearity of (7.1), where

$$f(Q^{n+3}) = \begin{cases} MQ^{n+3} & \text{for solving } \mathcal{M}_0, \\ MQ^{n+3} + f_{Q_3} & \text{for solving } \mathcal{M}_i \text{ with } i > 1. \end{cases}$$

Thus, (7.5) can be rewritten as

$$Q^{n+3} = \begin{cases} \left(I - \frac{6}{11}\Delta t M\right)^{-1} \left(\frac{18}{11}Q^{n+2} - \frac{9}{11}Q^{n+1} + \frac{2}{11}Q^n\right) & \text{for solving } \mathcal{M}_0, \\ \left(I - \frac{6}{11}\Delta t M\right)^{-1} \left(\frac{18}{11}Q^{n+2} - \frac{9}{11}Q^{n+1} + \frac{2}{11}Q^n + \frac{6}{11}\Delta t f_{Q_3}\right) & \text{for solving } \mathcal{M}_i \text{ with } i > 1. \end{cases}$$

The (local) accuracy of the temporal discretization is $\mathcal{O}((\Delta t)^3)$. Analogous discretization is implemented for solving the inhomogeneous system (2.10).

Appendix B: Expressions for the initial densities

Here, we list the formulas used for generating the initial densities $\tilde{p}_i(u, \gamma)$ introduced in §5.1. Recall that we chose the initial densities with uncorrelated variables,

$$\tilde{p}_i(\gamma, u) = \tilde{\pi}_i(\gamma) \tilde{\pi}_i(u),$$

where the marginal densities $\tilde{\pi}_i(\gamma)$ and $\tilde{\pi}_i(u)$ are given by the mixtures

$$\tilde{\pi}_i(\gamma) \propto \sum_n R_n(\gamma), \quad \tilde{\pi}_i(u) \propto \sum_n Q_n(u)$$

with the identical first and second moments chosen as

$$\langle u \rangle = 0, \quad \langle u^2 \rangle = 3, \quad \langle \gamma \rangle = 1.5, \quad \langle \gamma^2 \rangle = 7.5, \quad \langle \gamma u \rangle = 0.$$

In particular, the seven initial densities in §5.1 with the same joint second-order statistics are obtained as follows:

(1) Joint density

$$\tilde{p}_1(u, \gamma) = \frac{1}{2}(R_1(\gamma) + R_2(\gamma))Q_1(u),$$

where

$$R_i(\gamma) \propto \exp\left(-\frac{(\gamma - \bar{\gamma}_i)^2}{2\sigma_i^\gamma}\right), \quad Q_1(u) \propto \exp\left(-\frac{(u - \bar{u}_1)^2}{2\sigma_1^u}\right).$$

(2) Joint density

$$\tilde{p}_2(u, \gamma) = \frac{1}{4}(R_1(\gamma) + R_2(\gamma))(Q_1(u) + Q_2(u)),$$

where

$$R_i(\gamma) \propto \exp\left(-\frac{(\gamma - \bar{\gamma}_i)^2}{2\sigma_i^\gamma}\right), \quad Q_i(u) \propto \exp\left(-\frac{(u - \bar{u}_i)^2}{2\sigma_i^u}\right)(2 + \sin(u)).$$

(3) Joint density

$$\tilde{p}_3(u, \gamma) = \frac{1}{4}(R_1(\gamma) + R_2(\gamma))(Q_1(u) + Q_2(u)),$$

where

$$R_i(\gamma) \propto \exp\left(-\frac{(\gamma - \bar{\gamma}_i)^2}{2\sigma_i^\gamma}\right)\left(\frac{3}{2} + \sin\left(\frac{\pi\gamma}{2}\right)\right), \quad Q_i(u) \propto \exp\left(-\frac{(u - \bar{u}_i)^2}{2\sigma_i^u}\right)\left(\frac{3}{2} + \sin\left(\frac{\pi u}{2}\right)\right).$$

(4) Joint density

$$\tilde{p}_4(u, \gamma) = \frac{1}{4}(R_1(\gamma) + R_2(\gamma))(Q_1(u) + Q_2(u)),$$

where

$$R_i(\gamma) \propto \exp\left(-\frac{(\gamma - \bar{\gamma}_i)^2}{2\sigma_i^\gamma}\right)\frac{1}{\gamma^2 + 1}, \quad Q_i(u) \propto \exp\left(-\frac{(u - \bar{u}_i)^2}{2\sigma_i^u}\right)\frac{1}{u^2 + 1}.$$

(5) Joint density

$$\tilde{p}_5(u, \gamma) = \frac{1}{2}R_1(\gamma)(Q_1(u) + Q_2(u)),$$

where

$$R_i(\gamma) \propto \exp\left(-\frac{(\gamma - \bar{\gamma}_i)^2}{2\sigma_i^\gamma}\right)\frac{1}{\gamma^2 + 1}, \quad Q_i(u) \propto \exp\left(-\frac{(u - \bar{u}_i)^2}{2\sigma_i^u}\right).$$

(6) Joint density

$$\tilde{p}_6(u, \gamma) = \frac{1}{2}(R_1(\gamma) + R_2(\gamma))Q_1(u),$$

where

$$R_i(\gamma) \propto \exp\left(-\frac{(\gamma - \bar{\gamma}_i)^2}{2\sigma_i^\gamma}\right)\left(\frac{3}{2} + \sin\left(\frac{\pi\gamma}{2}\right)\right), \quad Q_1(u) \propto \exp\left(-\frac{(u - \bar{u}_1)^2}{2\sigma_1^u}\right).$$

(7) Joint density

$$\tilde{p}_7(u, \gamma) = \frac{1}{2}R_1(\gamma)(Q_1(u) + Q_2(u)),$$

where

$$R_1(\gamma) \propto \exp\left(-\frac{(\gamma - \bar{\gamma}_1)^2}{2\sigma_1^\gamma}\right), \quad Q_i(u) \propto \exp\left(-\frac{(u - \bar{u}_i)^2}{2\sigma_i^u}\right)\left(\frac{3}{2} + \sin\left(\frac{\pi u}{2} - 1\right)\right).$$

Table 4 The parameters used in (1)–(7)

	$\bar{\gamma}_1$	$\bar{\gamma}_2$	σ_1^γ	σ_2^γ	\bar{u}_1	\bar{u}_2	σ_1^u	σ_2^u
(1)	0.0000	3.0000	3.0000	3.0000	0.0000		3.0000	
(2)	0.0506	2.9494	2.6492	3.6492	-1.1667	0.5291	2.7234	1.3088
(3)	0.0167	2.9055	2.7649	3.6316	-0.9653	0.8210	2.7216	1.7763
(4)	4.0209	4.1964	21.9235	1.2482	-1.0970	5.0522	2.2703	2.0612
(5)	5.2632		11.1937		1.0204	-1.0203	1.4575	2.4603
(6)	0.0163	2.9064	2.7691	3.6204	0.0000	5.0000	3.0000	2.0000
(7)	1.5000		5.2500		-0.7417	-0.9372	1.3784	2.4465

References

- [1] Bensoussan, A., Stochastic Control of Partially Observable Systems, Cambridge University Press, Cambridge, 1992.
- [2] Bourlioux, A. and Majda, A. J., An elementary model for the validation of flamelet approximations in non-premixed turbulent combustion, *Combust. Theory Modell.*, **4**(2), 2000, 189–210.
- [3] Branicki, M., Gershgorin, B. and Majda, A. J., Filtering skill for turbulent signals for a suite of nonlinear and linear Kalman filters, *J. Comp. Phys.*, **231**, 2012, 1462–1498.
- [4] Branicki, M. and Majda, A. J., Fundamental limitations of polynomial chaos for uncertainty quantification in systems with intermittent instabilities, *Comm. Math. Sci.*, **11**(1), 2012, in press.
- [5] Branicki, M. and Majda, A. J., Quantifying uncertainty for predictions with model error in non-Gaussian models with intermittency, *Nonlinearity*, **25**, 2012, 2543–2578.
- [6] Cover, T. A. and Thomas, J. A., Elements of Information Theory, 2nd ed., Wiley-Interscience, Hoboken, 2006.
- [7] Gardiner, C., Stochastic Methods: A Handbook for the Natural and Social Sciences, 4th ed., Springer Series in Synergetics, Springer, Berlin, 2010.
- [8] Gershgorin, B., Harlim, J. and Majda, A. J., Improving filtering and prediction of spatially extended turbulent systems with model errors through stochastic parameter estimation, *J. Comp. Phys.*, **229**, 2010, 32–57.
- [9] Gershgorin, B., Harlim, J. and Majda, A. J., Test models for improving filtering with model errors through stochastic parameter estimation, *J. Comp. Phys.*, **229**, 2010, 1–31.
- [10] Gershgorin, B. and Majda, A. J., A test model for fluctuation-dissipation theorems with time-periodic statistics, *Physica D*, **239**, 2010, 1741–1757.
- [11] Gershgorin, B. and Majda, A. J., Quantifying uncertainty for climate change and long range forecasting scenarios with model errors. Part I: Gaussian models, *J. Climate*, **25**, 2012, 4523–4548.
- [12] Gershgorin, B. and Majda, A. J., A nonlinear test model for filtering slow-fast systems, *Comm. Math. Sci.*, **6**, 2008, 611–649.
- [13] Gershgorin, B. and Majda, A. J., Filtering a nonlinear slow-fast system with strong fast forcing, *Comm. Math. Sci.*, **8**, 2009, 67–92.
- [14] Gershgorin, B. and Majda, A. J., Filtering a statistically exactly solvable test model for turbulent tracers from partial observations, *J. Comp. Phys.*, **230**, 2011, 1602–1638.
- [15] Harlim, J. and Majda, A. J., Filtering turbulent sparsely observed geophysical flows, *Mon. Wea. Rev.*, **138**(4), 2010, 1050–1083.
- [16] Hersh, R., Random evolutions: a survey of results and problems, *Rocky Mountain J. Math.*, **4**(3), 1974, 443–477.
- [17] Iserles, A., A First Course in the Numerical Analysis of Differential Equations, Cambridge University Press, Cambridge, 1996.
- [18] Kleeman, R., Information theory and dynamical system predictability, *Entropy*, **13**, 2011, 612–649.
- [19] LeVeque, R., Numerical Methods for Conservation Laws, ETH Lectures in Mathematics Series, Birkhauser-Verlag, 1990.
- [20] Liptser, R. S. and Shiryaev, A. N., Statistics of random process, 2nd ed., Springer-Verlag, New York, 2001.
- [21] Majda, A. and Kramer, P., Simplified models for turbulent diffusion: Theory, numerical modeling, and physical phenomena, *Phys. Reports*, **314**(4), 1999, 237–257.
- [22] Majda, A. J., Challenges in climate science and contemporary applied mathematics, *Comm. Pure Appl. Math.*, **65**(7), 2012, 920–948.
- [23] Majda, A. J., Abramov, R. and Gershgorin, B., High skill in low frequency climate response through fluctuation dissipation theorems despite structural instability, *Proc. Natl. Acad. Sci. USA*, **107**(2), 2010, 581–586.
- [24] Majda, A. J., Abramov, R. V. and Grote, M. J., Information theory and stochastics for multiscale nonlinear systems, CRM Monograph Series, vol. 25, Americal Mathematical Society, Providence, 2005.
- [25] Majda, A. J. and Branicki, M., Lessons in uncertainty quantification for turbulent dynamical systems, *Discrete Cont. Dyn. Systems*, **32**(9), 2012, 3133–3231.
- [26] Majda, A. J., Franzke, C. and Crommelin, D., Normal forms for reduced stochastic climate models, *Proc. Natl. Acad. Sci.*, **106**(10), 2009, 3649–3653.

- [27] Majda, A. J., Franzke, C., Fischer, A., Crommelin, D. T., Distinct metastable atmospheric regimes despite nearly Gaussian statistics: A paradigm model, *Proc. Natl. Acad. Sci.*, **103**(22), 2006, 8309–8314.
- [28] Majda, A. J. and Gershgorin, B., Quantifying uncertainty in climate change science through empirical information theory, *Proc. Natl. Acad. Sci.*, **107**(34), 2010, 14958–14963.
- [29] Majda, A. J. and Gershgorin, B., Improving model fidelity and sensitivity for complex systems through empirical information theory, *Proc. Natl. Acad. Sci.*, **108**(5), 2011, 10044–10049.
- [30] Majda, A. J. and Gershgorin, B., Link between statistical equilibrium fidelity and forecasting skill for complex systems with model error, *Proc. Natl. Acad. Sci.*, **108**(31), 2011, 12599–12604.
- [31] Majda, A. J., Kleeman, R. and Cai, D., A mathematical framework for predictability through relative entropy, *Methods Appl. Anal.*, **9**(3), 2002, 425–444.
- [32] Majda, A. J. and Wang, X., *Nonlinear Dynamics and Statistical Theories for Basic Geophysical Flows*, Cambridge University Press, Cambridge, 2006.
- [33] Majda, A. J. and Gershgorin, B., Elementary models for turbulent diffusion with complex physical features: eddy diffusivity, spectrum, and intermittency, *Phil. Trans. Roy. Soc.*, 2011, In press.
- [34] Majda, A. J. and Harlim, J., *Filtering Complex Turbulent Systems*, Cambridge University Press, Cambridge, 2012.
- [35] Majda, A. J., Harlim, J. and Gershgorin, B., Mathematical strategies for filtering turbulent dynamical systems, *Discrete Cont. Dyn. Systems*, **27**, 2010, 441–486.
- [36] Risken, H., *The Fokker-Planck Equation: Methods of Solutions and Applications*, 2nd ed., Series in Synergetics, Springer-Verlag, Berlin, 1989.