



# Linear response based parameter estimation in the presence of model error

He Zhang<sup>a,\*</sup>, John Harlim<sup>a,b,c</sup>, Xiantao Li<sup>a</sup>

<sup>a</sup> Department of Mathematics, the Pennsylvania State University, 109 McAllister Building, University Park, PA 16802-6400, USA

<sup>b</sup> Department of Meteorology and Atmospheric Science, the Pennsylvania State University, 503 Walker Building, University Park, PA 16802-5013, USA

<sup>c</sup> Institute for CyberScience, the Pennsylvania State University, 224B Computer Building, University Park, PA 16802-2101, USA

## ARTICLE INFO

### Article history:

Available online 8 January 2021

### Keywords:

Parameter estimation  
Linear response theory  
Missing dynamics  
Kernel embedding

## ABSTRACT

Recently, we proposed a method to estimate parameters of stochastic dynamics based on the linear response statistics. The method rests upon a nonlinear least-squares problem that takes into account the response properties that stem from the Fluctuation-Dissipation Theory. In this article, we address an important issue that arises in the presence of model error. In particular, when the equilibrium density function is high dimensional and non-Gaussian, and in some cases, is unknown, the linear response statistics are inaccessible. We show that this issue can be resolved by fitting the imperfect model to appropriate *marginal linear response statistics* that can be approximated using the available data and parametric or nonparametric models. The effectiveness of the parameter estimation approach is demonstrated in the context of molecular dynamical models (Langevin dynamics) with a non-uniform temperature profile, where the modeling error is due to coarse-graining, and a PDE (non-Langevin dynamics) that exhibits spatiotemporal chaos, where the model error arises from a severe spectral truncation. In these examples, we show how the imperfect models, the Langevin equation with parameters estimated using the proposed scheme, can predict the nonlinear response statistics of the underlying dynamics under *admissible external disturbances*.

© 2021 Elsevier Inc. All rights reserved.

## 1. Introduction

An important and routine task in scientific research is to determine parameters in a given mathematical model. Whether the model is formulated from direct empirical observations or deduced from an underlying dynamic that is much more comprehensive, model error is inevitable in general. Even in the context of first-principle approaches, e.g., quantum-mechanical descriptions [1,2], they are rarely implemented in its full form. Instead, various reductions are introduced so that the model fits into practical computations, e.g., tight-binding methods [3] and pseudo-potentials [4], which necessarily introduce model error, and the parameters have to be determined in the presence of the model error.

We have in mind a class of stochastic models, whose parameters can be classified as follows. We assume that there is a subspace of parameters directly responsible for the equilibrium statistics. This subspace of parameters can be determined by many standard statistical methods. Once the parameters in this subspace are determined, the full model parameters can be

\* Corresponding author.

E-mail addresses: hqz5159@psu.edu (H. Zhang), jharlim@psu.edu (J. Harlim), xxl12@psu.edu (X. Li).

estimated by fitting the parameters in the quotient space corresponding to the dynamics' response properties. In this paper, we consider fitting the dynamical behavior through an impulse/response approach by perturbing the system with a small external forcing. This approach is known as the linear response in statistical physics as a first step to study non-equilibrium phenomena [5]. A hallmark in linear response theory is the Fluctuation-Dissipation Theory (FDT). Basically, it states that for a dynamical system at equilibrium, the full response to small external perturbations can be characterized by first-order linear responses through appropriate time correlation functions of the *unperturbed* dynamical system. Motivated by the accessibility of FDT, we have developed a parameter estimation approach [6] as well as an efficient numerical scheme [7] under the perfect model assumption. By inferring the estimates from the linear response statistics, we are able to reproduce both the equilibrium statistics and the response properties under small perturbations.

This paper focuses primarily on the response-based method in the presence of model error. In general, the model error may be a result of an empirical assumption, truncation, coarse-graining, multiscale expansion, discounting the memory effect, etc. In practice, the highly desirable parameter estimation methods are those that are not aware of these effects. Such a property is particularly relevant when the only available information is a data set of some observables of interest, which poses several new challenges for parameter estimation methods. First, the external forcing, when applied to the underlying dynamics and then carried over to the imperfect model, may depend on variables other than the observed components. We clarify this issue by introducing a concept called *admissible external forcing*. Another emerging issue is that the linear response involves a variable conjugate to the external force [5] (also interpreted as a flux variable [8]), which, in general, is only accessible when the statistics of the full model are known. We circumvent this difficulty by introducing the marginal linear response statistics that are practically computable given the available data and parametric/nonparametric modeling.

These ideas are best explained with concrete examples. Our first example is motivated by an important problem in computational chemistry, i.e., coarse-graining (CG) molecular dynamics (MD). The need for CGMD is driven by the observation that full molecular models are limited by the small intrinsic time scale and it is difficult for direct simulations to reach the time scale of interest [9–11], where important biological processes occur [12]. CGMD circumvents the problem by projecting the dynamics to the CG variables, a much smaller set, often defined by the local averages of the atomic coordinates and momenta [13–15]. The interactions among the CG variables are represented via the free energy, which is also known as the potential of mean force (PMF). In the uniform temperature case, this characterization provides an explicit parametric form of the probability density function. Then traditional statistical methods can be used to determine the parameters in the PMF [16], which we refer to as the equilibrium parameters in the remainder of this article. In addition to this straightforward setup, we will also consider a more general case where the system temperature is non-uniform, which is more challenging since the functional form of the equilibrium density is unknown.

Meanwhile, there are parameters in the CG dynamics that do not appear in the PMF, a well-known example of which is the damping coefficient in the Langevin equation, and more generally, the friction kernel in the generalized Langevin equations. In the latter case, parameters have to be determined by following the statistical properties of the trajectories of the CG variables. Various techniques have been considered, including the Bayesian approach [17], relative entropy [18], as well as the Padé-type of algorithms that match the time correlation properties [19–21]. The application of our response-based approach for CGMD models has been motivated by the steered MD simulations [22,23], where forces are applied to proteins to facilitate the folding and unfolding processes. With a one-dimensional chain model as a simple test problem, we demonstrate how the response-based approach is formulated and implemented to determine the damping coefficient, which is assumed to be a more general band matrix.

Another important class of problems is PDEs with solutions that exhibit chaotic behavior. One well-known example is turbulence. These dynamics are often projected to Fourier modes to study the energy transfer among different modes. From a reduced-order modeling viewpoint, a challenging problem is to deduce a closure model that retains the statistical properties of a finite number of modes. Part of the challenge stems from the fact that the full model's equilibrium statistic is high-dimensional and unknown. We will consider the Kuramoto-Sivashinsky (KS) equation as an example and demonstrate how to circumvent this difficulty with a semi-parametric method using the kernel embedding method.

In principle, the linear response theory is a prediction tool by itself, and the imperfect models would seem to be of little value in this case. However, as we previously alluded to, the conjugate variables might not be computable. The semi-parametric approach addresses precisely this issue. Besides, we demonstrate with examples that our approach yields models that also predict the *full response* with reasonable accuracy.

The paper is organized as follows: We start with a short review of the linear response theory and the parameter estimation approach based on the linear response in Section 2. To motivate a general parameter estimation scheme in the presence of model error, a linear fast-slow model will be presented in Section 3 for which explicit calculations can be carried out. In Section 4, we elaborate on the general difficulties and outline a parameter estimation scheme to address those issues. The numerical scheme will be followed by two examples: a one-dimensional chain model with spatially varying temperature (Section 5) and the KS equation (Section 6). We close the paper with a summary and discussions in Section 7. We include three appendices, discussing the detailed computational method for verifying the results, the derivation of the parametric form of the equilibrium density in the CGMD example, and presenting additional numerical results of the isothermal CGMD.

## 2. The parameter estimation method for the perfect model via linear response statistics

Fluctuation-Dissipation Theory (FDT) is a mathematical framework for quantifying the linear response of a dynamical system subjected to small external forcing [24]. The linear response statistics, determined based on two-point equilibrium statistics of the unperturbed dynamics, provide estimates for the response properties. In statistical mechanics literature, FDT is known as a linear response approach [8], which serves as a foundation for defining transport coefficients, e.g., viscosity, diffusion constant, heat conductivity, etc.

We begin by reviewing the linear response theory and the concept of the essential statistics, which is a core component of the response-based estimation method introduced in [6]. This will be presented in the context of a stochastic dynamics, expressed in the form of an  $n$ -dimensional (time-homogeneous) stochastic differential equations (SDEs), also known as the Itô diffusion [25]. The SDEs, together with its perturbed dynamics, are written as follows,

$$\dot{X} = b(X; \theta) + \sigma(X; \theta) \dot{W}_t, \quad (1)$$

$$\dot{X}^\delta = b(X^\delta; \theta) + c(X^\delta) \delta f(t) + \sigma(X^\delta; \theta) \dot{U}_t, \quad (2)$$

respectively, where  $W_t$  and  $U_t$  are standard Wiener processes. In the unperturbed system (1), the vector field  $b(X; \theta)$  denotes the drift and  $\sigma(X; \theta)$  is the diffusion tensor; while in the perturbed system (2), an order- $\delta$  ( $|\delta| \ll 1$ ) external perturbation is introduced in the form of  $f(x, t) = c(x) \delta f(t)$ . In both equations,  $\theta \in D$  denotes the model parameters with parameter domain  $D \subset \mathbb{R}^N$ .

Throughout this section, which concerns with the *perfect model* case, our goal is to estimate the underlying true parameter value  $\theta^\dagger$  based on the time series of the unperturbed dynamics (1) at the equilibrium. In the case of *imperfect model*, which is the central problem concerned in this paper, there are no true parameter values since the parameters in the imperfect model are not consistent with the parameters in the true model. Before we introduce the concept of the response statistics, we pose the following assumptions for all  $\theta \in D$ :

1. The system governed by Eq. (1) is ergodic with equilibrium density  $p_{eq}(x; \theta)$ , that is,  $p_{eq}(x; \theta)$  is the unique steady-state solution of the corresponding Fokker-Planck equation. In particular, we denote  $p_{eq}^\dagger(x) := p_{eq}(x; \theta^\dagger)$ .
2. The statistical properties associated with Eq. (2) can be characterized by a perturbed density  $p^\delta(x, t; \theta)$ , which follows the corresponding Fokker-Planck equation under initial condition  $p^\delta(x, 0; \theta) = p_{eq}(x; \theta)$ .

There is an abuse of notation since for certain SDEs,  $p_{eq}$  can depend on a subspace of the parameter. We use the notation  $\mathbb{E}_{p_{eq}(\theta)}[\cdot]$  and  $\mathbb{E}_{p^\delta(\theta)}[\cdot]$  to emphasize that the corresponding statistics are computed from the sample generated by the corresponding model at parameter value  $\theta$  initially at equilibrium.

For any integrable observables  $A(\cdot)$ , one can define the difference

$$\Delta \mathbb{E}[A](t) := \mathbb{E}_{p^\delta(\theta)}[A(X^\delta)](t) - \mathbb{E}_{p_{eq}(\theta)}[A(X)] \quad (3)$$

as the full response statistics. The linear response theory allows us to estimate the order- $\delta$  term in (3) by a convolution integral, that is,

$$\Delta \mathbb{E}[A](t) = \int_0^t k_A(t-s; \theta) \delta f(s) ds + \mathcal{O}(\delta^2). \quad (4)$$

The FDT formulates the linear response operator,  $k_A(t)$  in (4), as the following two-point statistics:

$$k_A(t; \theta) := \mathbb{E}_{p_{eq}(\theta)}[A(X(t)) \otimes B(X(0); \theta)], \quad B_i(X; \theta) := -\frac{\partial X_i [c_i(X) p_{eq}(X; \theta)]}{p_{eq}(X; \theta)}, \quad (5)$$

where  $B_i$  and  $c_i$  denote the  $i$ th components of  $B$  and  $c$ , respectively. The variable  $B$  will be called the conjugate variable [8] to  $A$ . We will denote  $B(X; \theta^\dagger)$  as  $B^\dagger(X)$ .

It is worthwhile to point out that in statistical mechanics [26], the notion of equilibrium often refers to a statistical ensemble, e.g., the canonical distribution that describes a system at thermal equilibrium. But the approach of linear response can be naturally extended to problems modeled by stochastic differential equations [27], in which case the equilibrium corresponds to the stationary distribution. The framework of linear response is also applicable to systems at non-equilibrium steady-state [28].

The significance of FDT is that the response operator is defined without involving the perturbed density  $p^\delta(x, t; \theta)$ . Rather, it can be evaluated at the equilibrium of the unperturbed dynamics. To turn this into a practical tool, we first notice that for a given  $t \geq 0$ , the value of  $k_A(t; \theta^\dagger)$  can be computed using a Monte-Carlo sum based on the time series of the unperturbed system (1) at  $p_{eq}^\dagger$ . For example, let  $\{X_i = X(t_i)\}_{i=1}^M$  be the time series generated at  $p_{eq}^\dagger$  with step length  $\Delta t = t_{i+1} - t_i$ , then for  $t = k \Delta t$ , the Monte-Carlo approximation can be written as

$$k_A(t; \theta^\dagger) \approx \frac{1}{M-k} \sum_{i=1}^{M-k} A(X_{i+k}) \otimes B^\dagger(X_i). \quad (6)$$

In practice, the right-hand side of (6) can be computed directly from data. Meanwhile, the same response statistics can be computed for any  $\theta \in D$  by solving (1) over the corresponding parameter value. In particular,

$$\hat{k}_A(t; \theta) := \mathbb{E}_{p_{eq}(\theta)}[A(X(t)) \otimes B^\dagger(X(0))],$$

can still be approximated by a Monte-Carlo sum (6) using the time series generated at  $p_{eq}(x; \theta)$ .

Motivated by the accessibility of  $k_A(t; \theta^\dagger)$  as well as the finite-dimensional Padé approximation introduced in [19], we have proposed to infer the true parameter value  $\theta^\dagger$  from a discretization of  $k_A(t; \theta^\dagger)$ . In particular, the discrepancy of the linear response statistics is given by:

$$\left| \int_0^t (k_A^\dagger(t-s) - \hat{k}_A(t-s; \theta)) \delta f(s) ds \right| \leq C \left( \int_0^t (k_A^\dagger(s) - \hat{k}_A(s; \theta))^2 ds \right)^{\frac{1}{2}},$$

where we have used Cauchy-Schwartz inequality with constant  $C$  denotes the  $L^2$ -norm of the term  $\delta f(t)$  in the external forcing. Minimizing the  $L^2$ -error in the right-hand side on discrete time *essential statistics*  $\{k_A^\dagger(t_i) := k_A(t_i; \theta^\dagger)\}_{i=1}^K$ , leads to the nonlinear least-squares problem [6,7]:

$$\hat{\theta} := \arg \min_{\theta \in D} \sum_{i=1}^K f_i^2(\theta), \quad f_i(\theta) := k_A^\dagger(t_i) - \hat{k}_A(t_i; \theta), \quad i = 1, 2, \dots, K. \quad (7)$$

To ensure the solvability of (7), we assume that the total number of involved essential statistics is always strictly greater than the dimension of  $\theta$ , that is,  $K > N$ . In terms of choosing  $\{t_i\}$ , one can interpret the optimization problem (7) as a curve-fitting problem. Thus, the points  $\{t_i\}$ 's should be selected such that  $\{k_A^\dagger(t_i)\}$ 's represent the qualitative feature in  $k_A^\dagger(t)$ . For example, one may include more points around local extreme points of  $k_A^\dagger(t)$ . Also, we do not want  $t_i$  to be too large, since in that case the value of  $k_A^\dagger(t)$  would be too small and the Monte-Carlo error introduced by (6) will overwhelm the significance of the essential statistics.

**Remark 2.1.** The accessibility of  $k_A(t; \theta^\dagger)$  also allows us to use its time derivatives as the essential statistics to infer  $\theta^\dagger$ . For example, we have shown that in [6], under specific external forcing and observables  $A(\cdot)$ , the first-order time derivative of  $k_A$  at  $t = 0$  of a Langevin dynamics model provides direct estimates for the damping coefficients. It is worthwhile to mention that estimating higher-order derivatives of  $k_A(t; \theta^\dagger)$  requires time series with sufficient accuracy, which is not necessarily available, especially when the data are from experimental observations.

**Remark 2.2.** It is important to point out that, in general, the quantity  $B(X; \theta)$  in (5) depends on the unknown parameter  $\theta$ , which provides difficulties in computing  $k_A(t; \theta^\dagger)$  given only a time series of  $X$  at  $p_{eq}^\dagger$ . In [29–31], Qi and Majda addressed this issue by computing the “kicked response” based on a direct simulation of the underlying dynamics. However, since the underlying dynamics is unknown in our scenario and estimating  $p_{eq}^\dagger$  is a difficult computational task in high-dimensional problems, we will consider estimating a marginal invariant density  $p_R$  defined on the resolved (and identifiable unresolved) variables and define  $B$  using the estimated marginal density  $p_R$  in placed of  $p_{eq}^\dagger$ , as explained in Section 4.

One practical challenge in solving (7) is that, except for the very special cases, there is no explicit expressions for  $\hat{k}_A(t; \theta)$ . Therefore an iterative method, e.g., the Gauss-Newton method, would necessarily require solving the model (1) repeatedly in a sequential manner, which can be rather computationally demanding. We have mitigated such an issue by a polynomial based surrogate model in [7], where the cost functions  $f_i(\theta)$  in (7) are approximated by linear combinations of orthogonal polynomials, denoted as  $f_i^M(\theta)$ , based on the precomputed values  $\{f_i(\theta_j)\}$  over a set of grid points  $\{\theta_j\} \subset D$ , and such pre-computation can be proceed in parallel. Here,  $M$  stands for the order of the polynomials involved. In the subsequent approximations, we will replace  $f_i(\theta)$  in (7) by  $f_i^M(\theta)$ , and formulate a new least-squares problem for the order- $M$  polynomial based surrogate model. In [7], we have deduced the convergence analysis of such approximation. In particular, we proved that the minimizers  $\{\theta_M^*\}$  of the approximated least-squares problem converge to the minimizer of the true least-squares problem as  $M \rightarrow \infty$  under the perfect model assumption and appropriate regularity assumption.

To this end, we have reviewed a parameter estimation approach for Itô diffusions (1) as well as an efficient numerical scheme under the perfect model assumption. By inferring from the linear response statistics, our estimates can reproduce both the equilibrium statistics and the response under certain perturbations. A much more difficult problem, however, is when the available model is a truncated dynamics or only partially known. In this case, several interesting issues will emerge, which will be investigated in the remaining part of the paper. In the next section, we start with a simple two-scale system to demonstrate the importance of the response statistics in determining the model parameters.

### 3. A linear fast-slow model for insight

Consider the following two-dimensional linear fast-slow dynamics as the underlying model

$$\dot{x} = (a_{11}x + a_{12}y) + \sigma_x \dot{W}_x, \quad (8)$$

$$\dot{y} = \frac{1}{\epsilon}(a_{21}x + a_{22}y) + \frac{\sigma_y}{\sqrt{\epsilon}} \dot{W}_y, \quad (9)$$

where  $x$  and  $y$  are the slow and the fast variables, respectively. Here,  $W_x$  and  $W_y$  are two independent Wiener processes and  $\epsilon \ll 1$  characterizes the time-scale gap. Such a multi-scale Ornstein-Uhlenbeck process has been used as a linear test model in a variety of situations [32]. For the sake of non-degeneracy and ergodicity, we assume that

$$\sigma_x, \sigma_y > 0, \quad a_{11}a_{22} - a_{12}a_{21} > 0, \quad a_{11} + \frac{a_{22}}{\epsilon} < 0, \quad a_{22} < 0,$$

which are enough to ensure the drift coefficients

$$A_\epsilon := \begin{pmatrix} a_{11} & a_{12} \\ a_{21}/\epsilon & a_{22}/\epsilon \end{pmatrix}$$

to be negative definite as well as the existence of the averaged (effective) dynamics,

$$\dot{X} = \bar{a}X + \sigma_x \dot{W}_t, \quad \bar{a} := a_{11} - \frac{a_{12}a_{21}}{a_{22}} < 0. \quad (10)$$

Here, the solution,  $X(t)$ , with initial condition  $X(0) = x(0)$ , converges strongly up to order-one time  $t$  to the solution  $x(t)$  of the true dynamics in (8)-(9). See for example [33] for a detailed discussion.

We will keep the form of (10) by posing

$$\dot{X} = bX + \sigma \dot{W}_t, \quad b < 0, \quad \sigma > 0, \quad (11)$$

as our imperfect model of the fast-slow dynamics with  $\theta = (b, \sigma)$  as the unknown parameters to be determined. Applying the parameter estimation scheme reviewed in Section 2 to  $\theta$  by taking observable  $A(X) = X$  and a time-dependent external forcing  $\delta f(t)$ , we find that (see the proof of Theorem 1 in [6] for computational details),

$$\hat{k}_A(t; \theta) = \mathbb{E}_{p_{eq}(\theta)}[X(t)X(0)]/S^\dagger = e^{tb} \frac{S}{S^\dagger},$$

where  $S = -\sigma^2/2b$  is the equilibrium variance of  $X$ . In contrast,  $S^\dagger$  is the variance that corresponds to the “true parameter” value  $\theta^\dagger$ , when the underlying dynamics is exactly of the form (11) with  $\theta = \theta^\dagger$ . Fitting the essential statistics ( $k_A^\dagger(t) = \hat{k}_A(t; \theta)$ ) leads to

$$\mathbb{E}_{p_{eq}^\dagger}[X(t)X(0)] = e^{tb^\dagger} S^\dagger = e^{tb} S. \quad (12)$$

In sharp contrast to the perfect model situation, in our case, the partial observations are from an underlying two-dimensional dynamics (8)-(9), while the essential statistics in (12) are derived based on a one-dimensional imperfect model (11). Thus, the left-hand side of (12), legitimately, is not available, and will be replaced by the two-point statistics of the slow variable  $x$  in (8). Eq. (12) now becomes,

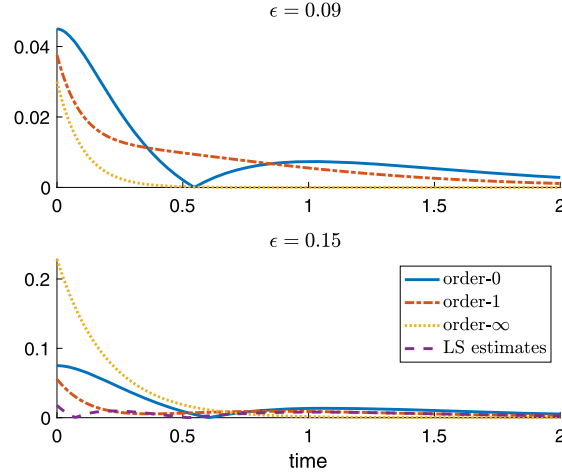
$$\mathbb{E}_{p_{eq}}[x(t)x(0)] = [1, 0]e^{tA_\epsilon} \Sigma [1, 0]^\top \approx e^{tb} S = -e^{tb} \frac{\sigma^2}{2b}, \quad (13)$$

where the covariance matrix  $\Sigma = (\Sigma_{ij})$  is determined by the Lyapunov equation [25]

$$A_\epsilon \Sigma + \Sigma A_\epsilon^\top = -Q, \quad Q = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2/\epsilon \end{pmatrix}. \quad (14)$$

One can interpret Eq. (13) as the consistency between the two time auto-covariance functions of the slow variable  $x$  and the variable  $X$  in the imperfect model. Obviously, seeking an exact fit for (13) is not feasible since its left-hand side has a total of 6 degrees of freedom ( $A_\epsilon$  and  $\Sigma$ ), which is much greater than that of the right-hand side. However, taking advantage of the multi-scale structure of the underlying dynamics, we can still find estimates for  $\theta$  when  $\epsilon \ll 1$  such that (13) is fulfilled up to an error of certain order. To see this, we first simplify the matrix exponential in (13) by computing the eigenvalues of  $A_\epsilon$ ,

$$\lambda_{1,2} = \frac{1}{2} \left( a_{11} + \frac{a_{22}}{\epsilon} \pm \sqrt{\Delta} \right), \quad \Delta = \left( a_{11} + \frac{a_{22}}{\epsilon} \right)^2 - \frac{4a_{22}\bar{a}}{\epsilon},$$



**Fig. 1.** The absolute error in reproducing  $\mathbb{E}_{p_{eq}}^\dagger [x(t)x(0)]$  under two different scale gaps  $\epsilon = 0.09$  (top) and  $\epsilon = 0.15$  (bottom). In terms of the estimates, order-0:  $\hat{b} = \tilde{a}$  and  $\hat{\sigma}^2 = \sigma_x^2$ ; order-1:  $\hat{b} = \tilde{a} - \epsilon a_{12} a_{21} \tilde{a} / a_{22}^2$  and  $\hat{\sigma}^2 = \sigma_x^2 - \epsilon (2\sigma_x^2 a_{12} a_{21} - \sigma_y^2 a_{12}^2) / a_{22}^2$ ; order- $\infty$ :  $\hat{b}$  and  $\hat{\sigma}^2$  are given by (16) without truncation, and LS: least-squares estimates.

where the  $\tilde{a}$  is the same as in Eq. (10). Assuming that the two eigenvalues are real, via a direct asymptotic expansions with respect to  $\epsilon$ , one can show that

$$\lambda_1 = \tilde{a} - \frac{a_{12} a_{21} \tilde{a}}{a_{22}^2} \epsilon + O(\epsilon^2), \quad \lambda_2 = \frac{a_{22}}{\epsilon} + \frac{a_{12} a_{21}}{a_{22}} + \frac{a_{12} a_{21} \tilde{a}}{a_{22}^2} \epsilon + O(\epsilon^2).$$

Notice that  $a_{22} < 0$  due to our assumption, and we are able to truncate the term containing  $e^{t\lambda_2}$  in (13) as long as  $t$  is large enough. Thus, up to a negligible error, (13) becomes

$$e^{\lambda_1 t} \frac{\Sigma_{11}(\lambda_2 - a_{11}) - \Sigma_{21} a_{12}}{\lambda_2 - \lambda_1} = -e^{tb} \frac{\sigma^2}{2b}, \quad (15)$$

which can be achieved by taking

$$\begin{aligned} \hat{b} = \lambda_1 &= \tilde{a} - \frac{a_{12} a_{21} \tilde{a}}{a_{22}^2} \epsilon + O(\epsilon^2), \\ \hat{\sigma}^2 &= -2\lambda_1 \frac{\Sigma_{11}(\lambda_2 - a_{11}) - \Sigma_{21} a_{12}}{\lambda_2 - \lambda_1} = \sigma_x^2 - \frac{2\sigma_x^2 a_{12} a_{21} - \sigma_y^2 a_{12}^2}{a_{22}^2} \epsilon + O(\epsilon^2), \end{aligned} \quad (16)$$

as our estimates. The same estimates have been obtained based on the Riccati equation in [34].

Compared to the averaged dynamics (10), our estimate in (16), based on matching the linear response statistics, exhausts the capability of the imperfect model (11) in the sense that the two-point statistics  $\mathbb{E}_{p_{eq}} [x(t)x(0)]$ , where  $t \gg \epsilon$ , of the slow variable  $x$  can be reproduced up to any order of  $\epsilon$  as  $\epsilon \rightarrow 0^+$ .

As a test, we set  $a_{11} = a_{21} = a_{22} = -1$ ,  $a_{12} = 1$ , and  $\sigma_x^2 = \sigma_y^2 = 2$  in (8) and (9). Fig. 1 shows the absolute error in reproducing the two-point statistics under two different choices of  $\epsilon$ , in comparison to various estimation methods of the imperfect model. In particular, we consider the estimate based on the classical averaging model in (10), which we denote as the order-0 model in Fig. 1 since it only keeps the leading order term in (16). In addition, we also consider the order-1 model, corresponding to truncating the order- $\epsilon^2$  in (16), and the order- $\infty$  model, corresponding to no truncation in (16), i.e., Eq. (15) is fulfilled. For small  $\epsilon$ , the order- $\infty$  model produces the most accurate estimation. In this regime, we empirically found that the estimation based on parameters obtained by solving Eq. (13) in the least-squares sense is nearly identical to the order- $\infty$  model (and thus, not reported here).

However, when  $\epsilon$  is relatively large, we are no longer allowed to truncate the term  $e^{t\lambda_2}$  in the simplification of (13), which in turn invalidates the expansion in (16). One can see the inaccurate estimate of the order- $\infty$  model in the second panel of Fig. 1. On the other hand, the solution of (13) in the least-squares sense is accurate and this justifies the importance of solving the nonlinear least-squares problem in (7) for parameter estimation in the presence of model error.

To summarize, we have elucidated the importance of the proposed parameter estimation scheme in the presence of model error on an idealistic problem. In particular, we have shown that the linear response-based parameter estimation scheme produces accurate linear mean response statistics, outperforming the conventional method, the averaging theory [33], even in a regime when such theory is valid.



#### 4. Parameter estimation in the presence of model error

Inspired by the result in Section 3, we consider extending the response-based parameter estimation scheme in the presence of model error arising from coarse-graining or missing dynamics. In the current section, we provide a formal discussion on the key issues in such a difficult scenario and propose a strategy to address them.

To elaborate the general idea, let us use the abstract notation,

$$\dot{u} = F(u), \quad X = P(u), \quad (17)$$

to denote the underlying full dynamics with partial observation of the variable of interest,  $X$ . Problems modeled by SDEs can be treated in a similar manner, and they will be addressed later. Here,  $P$  (not necessarily available) denotes a projection operator to the observable. We assume that  $F$  is unknown, instead, the available imperfect model is given as follows,

$$\dot{X} = R_1(X, Y; \theta), \quad \dot{Y} = R_2(X, Y; \theta). \quad (18)$$

Here, we have included an additional variable  $Y$  as a possible auxiliary variable and  $\theta \in D$  are the unknown model parameters. Our goal is to estimate the parameter  $\theta$  in (18) using the time series of  $X$  observed from (17) such that the resulting model in (18) can predict the full nonlinear response of the underlying dynamics (17) under *admissible external forcing*, as explained below. While the choice of model in (18) is central to the accuracy of the prediction, our focus is not on the specification of this model. Instead, we assume that the model in (18) is given and our focus, again, is to generalize the proposed parameter estimation scheme reviewed in Section 2 to this scenario.

In this case, the linear response operator (5) associated with the underlying dynamics (17) is no longer accessible since, in general, it is a two-point statistic that can be computed *only* if the full data set of  $u$  is available. Thus, to infer the parameter  $\theta$  in (18) from the linear response statistics of the underlying dynamics (17), we need to answer the following two questions:

1. How to find an alternative to the essential statistics  $k_A^\dagger(t_i)$  in (7), which is not computable given only the time series of  $X$ ?
2. How to formulate a computationally feasible nonlinear least-squares problem analogous to (7) for the imperfect model (18)?

To address these issues, we introduce the concepts of (i) *admissible external forcing*, and (ii) *marginal linear response (MLR)*. Using the notations in (17), similar to (1) and (2), the perturbed underlying dynamics can be written in the form

$$\dot{u}^\delta = F(u^\delta) + c(u^\delta)\delta f(t), \quad X^\delta = P(u^\delta). \quad (19)$$

Taking the time derivative of  $X$  in (17) and  $X^\delta$  in (19), we have

$$\dot{X} = \nabla P(u)F(u), \quad \dot{X}^\delta = \nabla P(u^\delta)F(u^\delta) + \delta f(t)\nabla P(u^\delta)c(u^\delta),$$

where  $\nabla P$  denotes the Jacobian matrix of  $P$ . Thus, the term  $\delta f(t)\nabla P(u^\delta)c(u^\delta)$  can be interpreted as the external force on the dynamics of  $X$ . In the case when the full dynamics (17) is a system of SDEs with an external force in the drift term, a similar observation can be made by using the Itô's formula. To this end, we define:

**Definition 4.1.** An external forcing  $f(u, t) = c(u)\delta f(t)$  to the full dynamics in (19) is *admissible* with respect to the dynamics (17) if  $\nabla P(u)c(u)$  can be written as a function of  $X$  denoted as  $c_R(X)$ .

In other words, an admissible external forcing  $c(u^\delta)\delta f(t)$  of the underlying dynamics (17) introduces a perturbation  $c_R(X^\delta)\delta f(t)$  to the imperfect model (18),

$$\dot{X}^\delta = R_1(X^\delta, Y^\delta; \theta) + c_R(X^\delta)\delta f(t), \quad \dot{Y}^\delta = R_2(X^\delta, Y^\delta; \theta). \quad (20)$$

Intuitively, this condition rules out other external forces that depend on  $u$ , which is not accessible since the imperfect model does not depend explicitly on  $u$ . We should point out that while the condition in Definition 4.1 provides a mathematical justification for the relevant perturbed imperfect model (18) that is consistent with the perturbed underlying dynamics in (19), this condition is not practically verifiable unless if  $P$ ,  $c$  and  $X$  are known. In our numerical examples, we will choose external forcings that are constant and nonzero on resolved components  $X$  such that the condition in Definition 4.1 is satisfied and they are admissible.

To simplify the discussion, let us define  $\theta_{eq} := \theta_{eq}(\theta)$  as the components of the parameters that directly appear in the equilibrium density  $\rho_{eq}(X, Y; \theta_{eq})$  of the imperfect model (18); they usually can be estimated directly by matching equilibrium statistics. For example, in the linear problem in (11), while the model parameters  $\theta$  consist of the damping coefficient  $b$  and noise amplitude  $\sigma$ , the parameter that appears in the equilibrium density,  $\theta_{eq}$ , is  $S = -\sigma^2/2b$ , which can be directly obtained by matching the equilibrium variances, i.e., considering  $t = 0$  in Eq. (13).

Once  $\theta_{eq}$  has been estimated (by  $\hat{\theta}_{eq}$ ), given an observable  $A(X)$ , the linear response operator (5) of the perturbed imperfect dynamics (20) is given as,

$$\hat{k}_A(t; \theta) = \mathbb{E}_{\rho_{eq}(\theta)} \left[ A(X(t)) \otimes B \left( X(0), Y(0); \hat{\theta}_{eq} \right) \right], \quad B_i \left( X, Y; \hat{\theta}_{eq} \right) = - \frac{\partial_{X_i} \left[ (c_R(X))_i \rho_{eq} \left( X, Y; \hat{\theta}_{eq} \right) \right]}{\rho_{eq} \left( X, Y; \hat{\theta}_{eq} \right)}, \quad (21)$$

where the expectation over  $\rho_{eq}(\theta)$  will be realized through Monte-Carlo averages with respect to the solutions of the unperturbed imperfect model in (18) for parameter  $\theta$  that satisfies the constrained of  $\theta_{eq}(\theta) = \hat{\theta}_{eq}$ .

Given an observable  $A(X)$  and an admissible external forcing  $f(u, t)$  in Definition 4.1, in analogous to the linear response operator (5) and the linear response statistics (4), we define:

**Definition 4.2.** Let  $p_R(X) := \int \delta(X - P(u)) p_{eq}^\dagger(u) du$  be the marginal invariant density of  $X$ , which can be estimated from the given data of  $X$  at the equilibrium. We define the marginal linear response (MLR) operator,

$$\hat{k}_A^\dagger(t) := \mathbb{E}_{p_{eq}^\dagger} \left[ A(X(t)) \otimes B_R^\dagger(X(0)) \right], \quad \left( B_R^\dagger \right)_i := - \frac{\partial_{X_i} \left[ (c_R(X))_i p_R(X) \right]}{p_R(X)}. \quad (22)$$

We call  $\left\{ \hat{k}_A^\dagger(t_i) \right\}_{i=1}^K$  the marginal essential statistics. Furthermore, we define,

$$\hat{\Delta} \mathbb{E}[A](t) = \int_0^t \hat{k}_A^\dagger(t-s) \delta f(s) ds \quad (23)$$

as the MLR statistics.

The two-point statistics in (22), defined with respect to the equilibrium distribution of the underlying dynamics (17), can be approximated via Monte-Carlo as in (6), based on the available time series of  $X$ . We require the external forcing to be admissible; otherwise,  $c_R$  in (22) would depend on  $u$  and the marginal essential statistics are not accessible when the data of  $u$  is not available. In practice, as we will show in the next two sections, the marginal density  $p_R$  can be learned from the available time series of  $X$  by using standard statistical methods. This way, the two-point statistics in (22) is computable as opposed to (6). We should also point out that, in principle,  $c_R$  can also depend on  $Y$ . In this case, computing  $k_A^\dagger$  in (22) requires samples of, both,  $X$  and  $Y$  for estimating the marginal density  $p_R$ , which is a function of  $X$  and  $Y$ . In Section 6.3, we will consider the case where the samples data of  $Y$  are not available.

To infer the parameter  $\theta$  in (18) from (22), compared with (7), we formulate the following new nonlinear least-squares problem

$$\hat{\theta} := \arg \min_{\left\{ \theta \in D \mid \theta_{eq}(\theta) = \hat{\theta}_{eq} \right\}} \sum_{i=1}^K f_i^2(\theta), \quad f_i(\theta) := \hat{k}_A^\dagger(t_i) - \hat{k}_A(t_i; \theta), \quad i = 1, 2, \dots, K, \quad (24)$$

where  $\hat{k}_A^\dagger(t_i)$  is the marginal essential statistics (22) and  $\hat{k}_A(t_i; \theta)$  is the linear response operator (21) of the imperfect model (18). Eq. (24) is a dynamic constrained nonlinear least-squares problem through the imperfect model in (18). We should point out the minimization is over a subspace of the parameter set, characterized by the constraint,  $\theta_{eq}(\theta) = \hat{\theta}_{eq}$ , where  $\hat{\theta}_{eq}$  is the estimate for the equilibrium parameters,  $\theta_{eq}$ , obtained from direct matching to the available equilibrium statistics.

In the next two sections, we validate the estimate  $\hat{\theta}$  by comparing the nonlinear response of the imperfect model at the estimated parameters with that of the underlying model under admissible external forcing. These statistics are computed via Monte-Carlo averages over solutions of the perturbed underlying dynamics (19) and perturbed imperfect model (20), where the former is only possible when the full underlying dynamic is known. See Appendix A for the detailed computational strategy to compute these statistics for stochastic dynamics.

**Remark 4.3.** For the fast-slow model discussed in Section 3, the underlying dynamics are governed by (8)-(9), while the imperfect model is given by (11). The time-dependent external forcing we proposed in the previous analysis is admissible since it acts only on the slow variable  $x$ .

## 5. A one-dimensional molecular model under a local temperature profile

We consider a MD model consisting of a chain of atoms with mass  $m$ . Let the equilibrium spacing between the particles be  $a_0$ , and the displacement of the  $i$ th particles from its equilibrium position  $R_i = ia_0$  be  $r_i$ . We take the Lennard-Jones (LJ) potential [35], expressed as



$$U_0(r) = \sum_i \sum_{j=i-2}^{i-1} \psi(r_i - r_j + (i-j)a_0), \quad \psi(r) = |r|^{-12} - |r|^{-6}. \quad (25)$$

### 5.1. The underlying dynamics

Our underlying dynamics, as a finite-dimensional approximation of the LJ lattice model, contains a total of  $N$  particles equipped with periodic boundary conditions. The size of the system is  $Na_0$ . Let  $\mathbf{r} = (r_1, r_2, \dots, r_N)^\top$  and  $\mathbf{v} = (v_1, v_2, \dots, v_N)^\top$  be the displacement and the velocities of the particles, respectively. The potential energy of the finite system, as a function of the relative displacement  $\mathbf{d} := (r_2 - r_1, r_3 - r_2, \dots, r_N - r_{N-1})^\top \in \mathbb{R}^{N-1}$ , can be written as,

$$U(\mathbf{d}) = \sum_{i=1}^{N+2} \sum_{j=i-2}^{i-1} \psi(r_i - r_j + (i-j)a_0). \quad (26)$$

The potential  $U$  determines the direct interactions among the particles. In the model (26), each particle is interacting with the nearest two neighbors from each side. We consider a Langevin dynamics of  $(\mathbf{r}, \mathbf{v})$ , driven by the potential energy  $U(\mathbf{d})$ , the frictional, and random forcing,

$$\begin{aligned} \dot{r}_i &= v_i, \\ \dot{v}_i &= -\nabla_{r_i} U(\mathbf{d}) - \gamma v_i + \sqrt{2k_B T_i \gamma} \dot{W}_t^{(i)}, \end{aligned} \quad (27)$$

for  $i = 1, 2, \dots, N$ , where the mass has been set to unity ( $m = 1$ ),  $\gamma$  denotes the friction constant, and  $W_t = (W_t^{(1)}, \dots, W_t^{(N)})$  is an  $N$ -dimensional Wiener process. The temperature in (27) is allowed to be non-uniform. As a specific example, we consider a sinusoidal profile

$$k_B T_i = k_B T_0 + k_B \Delta T \sin\left(\frac{2\pi i}{N}\right), \quad i = 1, 2, \dots, N. \quad (28)$$

The dynamic in (27) models coupled oscillators interacting with thermal reservoirs of different temperatures (e.g., [36]). Most of the effort in CGMD models has been focused on systems at a constant temperature ( $\Delta T \equiv 0$  in (28)), in which case, the equilibrium density is known (e.g., [15]). Unlike the isothermal case, the equilibrium density of (27) is not available in general. We will discuss how this issue is addressed in our approach.

In the numerical simulation, we set  $a_0 = 1.1196$ ,  $k_B T_0 = 0.25$ ,  $k_B \Delta T = 0.05$ ,  $m = 1.0$ ,  $\gamma = 0.5$  and  $N = 100$  in (27). A Verlet-type of integration algorithm [37] is applied using step length  $\Delta t = 0.05$ . The data is sub-sampled at every ten integration time steps ( $\delta t = 0.005$ ) at equilibrium with sample size  $5 \times 10^6$ . The parameter  $k_B T_0$  is set to be large enough, so that at the equilibrium, the marginal density of  $\mathbf{d}$  is non-Gaussian. Throughout the section, the underlying dynamics (27) will only be used for generating CG observations and computing the full response for verification of the parameter estimates.

### 5.2. The coarse-grained model

Motivated by the residue-based coarse-graining [9], we define the CG variables by dividing the system equally to consecutive blocks, each of which contains  $J$  atoms. In particular, the displacement and velocity of the CG particles are defined as

$$q_i := \frac{1}{J} \sum_{j=(i-1)J+1}^{iJ} r_j, \quad p_i := \dot{q}_i = \frac{1}{J} \sum_{j=(i-1)J+1}^{iJ} v_j, \quad i = 1, 2, \dots, n_b, \quad n_b := \frac{N}{J}, \quad (29)$$

respectively. Here,  $n_b$  denotes the number of CG particles. To construct a CG model, we propose the following function of the CG relative displacement  $\xi = (q_2 - q_1, q_3 - q_2, \dots, q_{n_b} - q_{n_b-1}) \in \mathbb{R}^{n_b-1}$  and CG velocity  $\mathbf{p}$

$$p_{eq}^\dagger(\xi, \mathbf{p}) \approx \rho(\xi, \mathbf{p}) \propto \prod_{i=1}^{n_b} \exp\left(-b_i \frac{p_i^2}{2} - \sum_{j=1}^4 a_{j,i} \xi_i^j - \sum_{k=1}^3 c_{k,i} p_i \xi_i^k\right), \quad a_{4,i}, b_i > 0, \quad i = 1, 2, \dots, n_b, \quad (30)$$

as the ansatz of the equilibrium distribution of the CG variables  $(\xi, \mathbf{p})$ . In the exponent of Eq. (30), the cross terms between  $p_i$  and  $\xi_i$  are introduced to incorporate the heat flux of the underlying dynamics (27) caused by the non-uniform temperature profile. When the temperature is uniform, there is no heat flux, and these cross terms would be zero (see Appendix B for details).

Formally, to capture the proposed equilibrium distribution (30), we define a pseudo-Hamiltonian

$$H := \sum_{i=1}^{n_b} \left( b_i \frac{p_i^2}{2} + \sum_{j=1}^4 a_{j,i} \xi_i^j + \sum_{k=1}^3 c_{k,i} p_i \xi_i^k \right),$$

such that the equilibrium density ansatz in (30) can be written as  $\rho \propto \exp(-H)$ . We consider the following imperfect model,

$$\begin{aligned} \dot{\mathbf{q}} &= D^{-1} \nabla_{\mathbf{p}} H, \\ \dot{\mathbf{p}} &= -D^{-1} \nabla_{\mathbf{q}} H - D^{-1} \Gamma \nabla_{\mathbf{p}} H + \sqrt{2} D^{-\frac{1}{2}} \sigma \dot{U}_t, \end{aligned} \quad (31)$$

where  $D = \text{diag}(b_1, b_2, \dots, b_{n_b})$ , and  $U_t$ , to distinguish from  $W_t$  in (27), is an  $n_b$ -dimensional Wiener process. In (31), the coefficient  $\Gamma$  is assumed to be a symmetric positive definite (SPD) matrix satisfying  $\Gamma = \sigma \sigma^\top$ . In practice,  $\sigma$  can be determined by a Cholesky decomposition of  $\Gamma$ . In particular,  $\rho$  in (30) solves the stationary Fokker-Planck equation of (31), which can be verified by direct computations.

To better understand the imperfect model (31), we rewrite (31) as

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p} + f(\mathbf{q}), \\ \dot{\mathbf{p}} &= F(\mathbf{q}) - \tilde{\Gamma} \mathbf{p} + g(\mathbf{q}, \mathbf{p}) + \Sigma \dot{W}_t, \end{aligned} \quad (32)$$

where

$$\begin{aligned} \tilde{\Gamma} &= D^{-1} \Gamma D, \quad \Sigma = \sqrt{2} D^{-\frac{1}{2}} \sigma, \\ f_i &= \sum_{k=1}^3 \frac{c_{k,i}}{b_i} \xi_i^k, \quad F_i = b_i^{-1} \sum_{j=1}^4 j \left( a_{j,i} \xi_i^{j-1} - a_{j,i-1} \xi_{i-1}^{j-1} \right), \\ g_i &= b_i^{-1} \sum_{k=1}^3 k \left( c_{k,i} p_i \xi_i^{k-1} - c_{k,i-1} p_{i-1} \xi_{i-1}^{k-1} \right), \quad i = 1, 2, \dots, n_b. \end{aligned}$$

Eq. (32) suggests that the imperfect model (31) is a perturbed dynamics of

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= F(\mathbf{q}) - \tilde{\Gamma} \mathbf{p} + \Sigma \dot{W}_t. \end{aligned} \quad (33)$$

Unlike the classical Langevin dynamics, the force  $F$  in (33) is nonconservative. In particular,

$$\frac{\partial F_i}{\partial q_{i+1}} = b_i^{-1} \sum_{j=2}^4 j(j-1) a_{j,i} \xi_i^{j-2} \neq \frac{\partial F_{i+1}}{\partial q_i} = b_{i+1}^{-1} \sum_{j=2}^4 j(j-1) a_{j,i} \xi_i^{j-2}.$$

By Theorem 3 in [38], we can verify that (33) is geometrically ergodic. Thus, such an imperfect model (31) can be interpreted as a perturbed dynamics of an ergodic Langevin dynamics with nonconservative forces. The ergodicity of the imperfect model (31) is left here as an open problem.

Motivated by the symmetry of the system, the damping matrix  $\Gamma$  is assumed to be a symmetric circulant matrix:

$$\Gamma = \begin{pmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_2 & \gamma_1 \\ \gamma_1 & \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_2 \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ \gamma_2 & \cdots & \gamma_2 & \gamma_1 & \gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_2 & \cdots & \gamma_2 & \gamma_1 & \gamma_0 \end{pmatrix} \in \mathbb{R}^{n_b \times n_b}, \quad (34)$$

which is fully determined by the first  $\lfloor \frac{n_b}{2} \rfloor + 1$  elements  $(\gamma_0, \dots, \gamma_{\lfloor \frac{n_b}{2} \rfloor})$  in the first row. To guarantee that  $\Gamma$  is positive definite, we pose a diagonal dominant constraint:

$$\gamma_0 > \begin{cases} 2 \sum_{i=1}^{\frac{n_b}{2}-1} |\gamma_i| + \left| \gamma_{\frac{n_b}{2}} \right|, & n_b \text{ even,} \\ 2 \sum_{i=1}^{\frac{n_b-1}{2}} |\gamma_i|, & n_b \text{ odd.} \end{cases}$$

**Table 1**  
Part of the estimates of  $\theta_{eq}$  in (30).

Parameter ( $\theta_{eq}$ )	$b_1^{-1}$	$a_{1,1}$	$a_{2,1}$	$a_{3,1}$	$a_{4,1}$	$c_{1,1}$	$c_{2,1}$	$c_{3,1}$
Estimates ( $\hat{\theta}_{eq}$ )	0.02631	-0.2582	5.951	-2.781	0.6885	-0.4949	0.1446	0.04487

For simplicity, we will set  $\gamma_3 = \gamma_4 = \dots = \gamma_{\lfloor \frac{n_b}{2} \rfloor} = 0$  in (34), and the parameter space of the imperfect model (31) can be summarized as

$$\theta := \{b_i, a_{1,i}, \dots, a_{4,i}, c_{1,i}, \dots, c_{3,i}, \gamma_0, \gamma_1, \gamma_2\}_{i=1}^{n_b}, \quad a_{4,i}, b_i > 0, \quad i = 1, 2, \dots, n_b, \quad \gamma_0 > 2(|\gamma_1| + |\gamma_2|). \quad (35)$$

To be consistent with the notation in Section 4, the parameters in (35) include  $\theta_{eq} := \{b_i, a_{j,i}, c_{k,i}\}$ , which are the coefficients in the density function (30), and  $\gamma_i$  are the component of the damping matrix (34).

There are many other approaches in the model reduction of Langevin dynamics, e.g., the partitioning technique [39], the Petrov-Galerkin projection [20], and the maximum likelihood estimator [17]. Here, as we have emphasized, we consider the imperfect model (31) as a mean to introduce the model error, and use the one-dimensional system to demonstrate the parameter estimation procedure, instead of focusing on the model reduction technique.

### 5.3. Parameter estimation and numerical results

To this end, we have introduced our underlying dynamics (27) and the corresponding imperfect model (31) with unknown parameters  $\theta$  in (35). In the numerical experiment, the observation of the CG model (31) will be acquired from Eq. (29) based on the time series of  $(\mathbf{d}, \mathbf{v})$  generated by the underlying dynamics. For instance, the CG relative displacement  $\xi$  can be determined by:

$$\xi_i = q_{i+1} - q_i = \frac{1}{J} \sum_{j=(i-1)J+1}^{iJ} (r_{j+J} - r_j) = \frac{1}{J} \sum_{j=(i-1)J+1}^{iJ} \sum_{k=j}^{j+J-1} (r_{k+1} - r_k), \quad i = 1, 2, \dots, n_b - 1. \quad (36)$$

Here, Eq. (36), together with the second equation of (29) correspond to the function  $P(\cdot)$  in our general setup (17) with the variable of interest  $X = (\xi, \mathbf{p})$ . Thus, we are able to obtain the time series of  $(\xi, \mathbf{p})$ , containing model error, for the imperfect model (31), which will be the only information used in estimating  $\theta$ .

For parameters  $\theta_{eq}$  that appear in the density ansatz (30), we will apply the maximum entropy method [40] using the equation-by-equation solver introduced in [41]. In our implementation, we solve for  $(b_i, a_{1,i}, \dots, a_{4,i}, c_{1,i}, \dots, c_{3,i})$  for each  $i$  independently since

$$\rho_i(\xi_i, p_i) \propto \exp \left( -b_i \frac{p_i^2}{2} - \sum_{j=1}^4 a_{j,i} \xi_i^j - \sum_{k=1}^3 c_{k,i} p_i \xi_i^k \right) \quad (37)$$

fully determines the marginal distributions of  $(\xi_i, p_i)$ . Thus, the parameter estimation problem of  $\theta_{eq}$  can be decomposed into a total of  $n_b$  two-dimensional maximum entropy problems. By maximizing the Shannon's entropy under the moment constraints, we formulate the following maximum entropy estimates for  $i = 1, 2, \dots, n_b$ ,

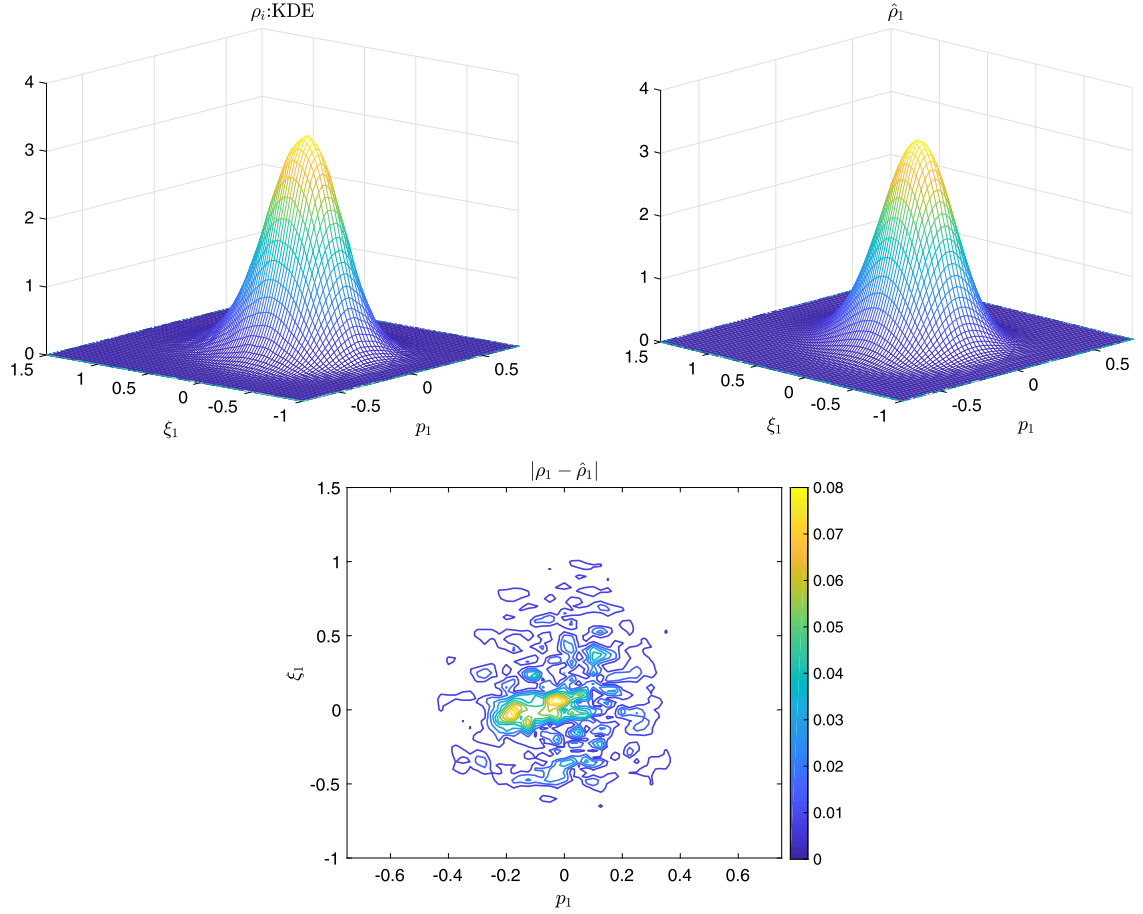
$$(\hat{b}_i, \hat{a}_{1,i}, \dots, \hat{a}_{4,i}, \hat{c}_{1,i}, \dots, \hat{c}_{3,i}) = \arg \min \int -\rho_i(\xi_i, p_i) \log \rho_i(\xi_i, p_i) d\xi_i dp_i, \quad (38)$$

$$\text{s.t. } \mathbb{E}_{\rho_i}[1] = 1, \quad \mathbb{E}_{\rho_i}[p_i^2] = \mathbb{E}_{p_{eq}^\dagger}[p_i^2], \quad \mathbb{E}_{\rho_i}[\xi_i^j] = \mathbb{E}_{p_{eq}^\dagger}[\xi_i^j], \quad \mathbb{E}_{\rho_i}[p_i \xi_i^k] = \mathbb{E}_{p_{eq}^\dagger}[p_i \xi_i^k], \quad j = 1, 2, 3, 4, \quad k = 1, 2, 3, \quad (39)$$

where  $\mathbb{E}_{p_{eq}^\dagger}[\cdot]$  will be approximated using Monte-Carlo average over the available time series of  $(\xi, \mathbf{p})$ . Thus, such constrained minimization problem (38) and (39) is equivalent to the reverse Monte Carlo method [42]. Table 1 shows the value of the estimates of  $\theta_{eq}$  for  $i = 1$ . Fig. 2 compares  $\rho_1$  determined by a two-dimensional kernel density estimates based on the CG observations  $(\xi_1, p_1)$  and  $\hat{\rho}_1$  using the parameter estimates in Table 1. We can see that the maximum entropy estimates provide a good fit towards the marginal distribution of  $(\xi_1, p_1)$  at the equilibrium of the underlying dynamics.

Up to this point, we have estimated  $\theta_{eq}$  in the density ansatz (30) of the imperfect model (31). We now apply our parameter estimation approach to estimate  $\{\gamma_i\}$  in the damping matrix  $\Gamma$ . These parameters determine the average dynamical behavior at equilibrium. To satisfy the assumptions posed in Section 4, the external forcing of the underlying dynamics is set to be a constant external forcing, exerted equally on the first block of atoms, and the observables  $A(X)$  are the velocities of the first two CG particles. As a result, following (20), the corresponding external forces  $c \delta f$  and  $c_R \delta f$  for the underlying dynamics (acting on variable  $\mathbf{v}$ ) and imperfect model (acting on CG variable  $\mathbf{p}$ ), respectively, satisfy

$$\mathbf{c} = (\underbrace{1, \dots, 1}_{J\text{-terms}}, 0, \dots, 0) \in \mathbb{R}^N, \quad \mathbf{c}_R = (1, 0, \dots, 0)^\top \in \mathbb{R}^{n_b}, \quad \delta f = \delta. \quad (40)$$



**Fig. 2.** The density function  $\rho_1$  (up-left, plot using the “mvksdensity” syntax in MATLAB), and  $\hat{\rho}_1$  (up-right, plot using the parametric form in (37) and the parameter estimates in Table 1) the density estimates generated via maximum entropy. The absolute error between  $\rho_1$  and  $\hat{\rho}_1$  is presented in the lower panel. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

According to the criterion postulated in Definition 4.1, this forcing is admissible. With the observables  $A = (p_1, p_2)^\top$ , following (21)–(22), we write

$$\begin{aligned}\hat{k}_A^\dagger(t) &= \mathbb{E}_{p_{eq}^\dagger} \left[ (p_1(t), p_2(t))^\top \otimes \hat{B}(p_1(0), \xi_1(0)) \right], \\ \hat{k}_A(t; \theta) &= \mathbb{E}_{\rho(\theta)} \left[ (p_1(t), p_2(t))^\top \otimes \hat{B}(p_1(0), \xi_1(0)) \right],\end{aligned}\quad (41)$$

where

$$\hat{B}(p_1, \xi_1) = -\frac{\partial}{\partial p_1} \log \left( \rho(\xi, \mathbf{p}; \hat{\theta}_{eq}) \right) = \hat{b}_1 p_1 + \sum_{k=1}^3 \hat{c}_{k,1} \xi_1^k,$$

and  $\hat{b}_1$  and  $\{\hat{c}_{k,1}\}$  are the estimates obtained from solving the maximum entropy problem (38)–(39). Subsequently, we formulate the following constrained nonlinear least-squares problem:

$$\min_{\theta} \sum_{i=1}^K \sum_{j=1}^2 \left( g_j^\dagger(t_i) - \hat{g}_j(t_i; \theta) \right)^2, \quad \text{s.t. } \theta_{eq} = \hat{\theta}_{eq}, \quad \gamma_0 > 2(|\gamma_1| + |\gamma_2|), \quad (42)$$

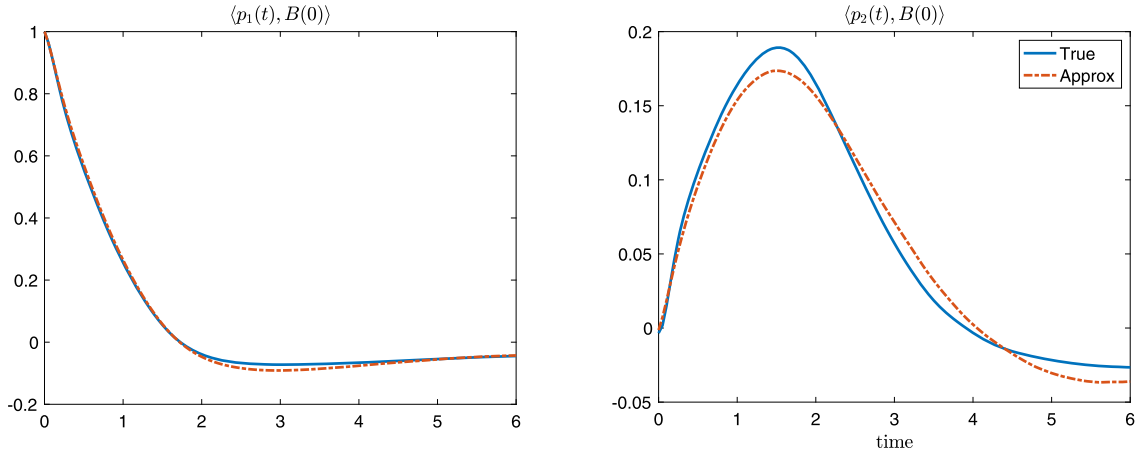
where

$$g_j^\dagger(t) = \mathbb{E}_{p_{eq}^\dagger} \left[ p_j(t) \hat{B}(p_1(0), \xi_1(0)) \right], \quad \hat{g}_j(t; \theta) = \mathbb{E}_{\rho(\theta)} \left[ p_j(t) \hat{B}(p_1(0), \xi_1(0)) \right], \quad j = 1, 2.$$

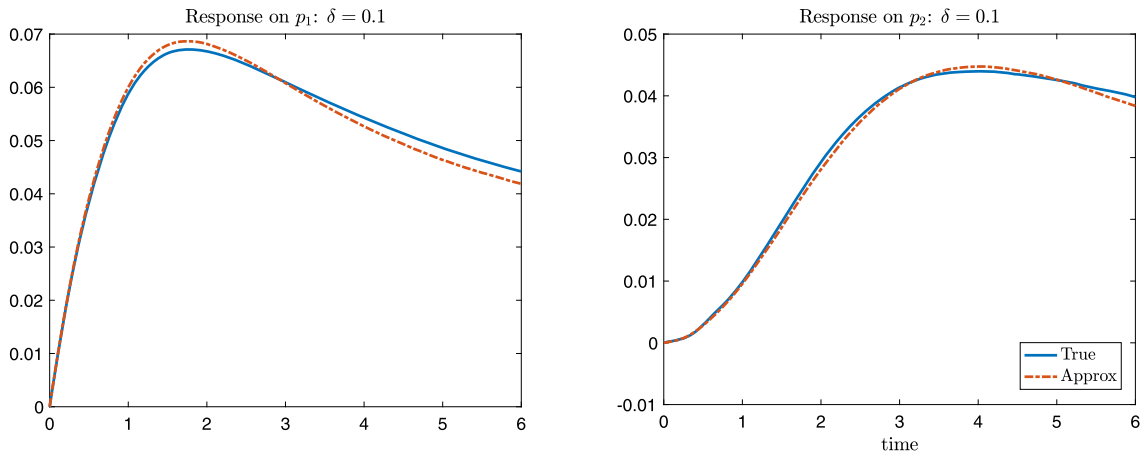
Recall that both  $\{g_j^\dagger(t_i)\}$  and  $\{\hat{g}_j(t_i; \theta)\}$  will be approximated by Monte-Carlo of the form (6) based on, respectively, the available CG observations and the time series generated by simulating the imperfect model (31) at the corresponding parameter

**Table 2**  
Estimates of the parameter in the damping matrix  $\Gamma$  (34).

Parameter ( $\gamma_i$ )	$\gamma_0$	$\gamma_1$	$\gamma_2$
Estimates ( $\hat{\gamma}_i$ )	0.8409	-0.1758	$7.616 \times 10^{-3}$



**Fig. 3.** Comparison of the time correlation functions involved in the nonlinear least-squares problem (42). The blue solids curves are computed from the available CG observations, while the red dot-dash curves are the approximation of the imperfect model based on the estimates. (Sample size:  $1 \times 10^6$ .)



**Fig. 4.** Reconstruction of full response statistics: underlying dynamics (blue solid) v.s. the imperfect model (red dot-dash). We take  $\delta = 0.1$  in (40) as the external forcing. The full response statistics (3) are computed via Monte-Carlo. Check Appendix A for the details.

value. In the numerical simulation of the imperfect model (31), we use a second-order weak trapezoidal method introduced in [43].

As an example, we set  $J = 10$  such that  $n_b = 10$  since  $N = 100$ . Setting  $t_i = 0.2i$  with  $i = 1, 2, \dots, 20$ , we solve (42) using the numerical scheme reviewed in Section 2. Specifically, a polynomial chaos expansion with order-6 Legendre polynomials is used to approximate the cost function (42) to avoid the excessive computational cost in evaluating  $\hat{g}_j$  on new parameter  $\theta$  (see [7] for details). Table 2 shows the value of the estimates. We also show in Figs. 3 and 4 the results from fitting the essential statistics and recovering the full response of the underlying dynamics, respectively. One can observe from the figures that the estimates provide an excellent fit on the linear response operator  $\langle p_1(t), B(0) \rangle$ . For the response operator  $\langle p_2(t), B(0) \rangle$ , the poor performance at the longer time might be caused by the Monte-Carlo error involved in (42) and possibly the polynomial chaos approximation on the cost function. Interestingly, the imperfect model produces reasonably accurate estimates of the full nonlinear response statistics of the underlying dynamics. We also tested the isothermal case, as a special case of (27), and the results also show good accuracy in terms of the estimation of the full response statistics. These results are reported in Appendix C for interested readers.

In the next section, we will present an example from solutions of a PDE. At this point, it is worthwhile to emphasize several points.

**Remark 5.1.** While in general the choice of the imperfect model in (18) depends on the specific applications, in current and next section, we will consider Langevin type stochastic dynamics, regardless of whether the underlying dynamics are Langevin or not. In the example above, the imperfect Langevin model (31), as a result of coarse-graining, is a closed equation of  $X = (\xi, \mathbf{p})$ , with no dependence on any auxiliary variables  $Y$ . In this case, we propose a natural parametric form as an ansatz of the associated equilibrium density and use the maximum entropy principle to estimate  $\theta_{eq}$ . In the next section, we shall see that the imperfect model (also Langevin dynamics) is not a closed equation of  $X$ .

**Remark 5.2.** In the example above, we considered  $p_R(X) = \rho(X; \theta_{eq})$  in (30). We should point out that choosing  $p_R$  to depend only on  $X$  in (22) is only sensible (in the linear response theory) if  $X$  is independent to the orthogonal component,  $X^\perp := u - P(u)$ , at equilibrium state, in that  $p_{eq}(X, X^\perp) = p_R(X)p_{R^\perp}(X^\perp)$ . In this case, we have

$$\frac{\partial_{X_i} [c_i(X) p_{eq}(X, X^\perp)]}{p_{eq}(X, X^\perp)} = \frac{\partial_{X_i} [c_i(X) p_R(X) p_{R^\perp}(X^\perp)]}{p_R(X) p_{R^\perp}(X^\perp)} = \frac{\partial_{X_i} [c_i(X) p_R(X)]}{p_R(X)},$$

which suggests that the MLR operator is identical to the linear response operator of the underlying dynamics,  $\hat{k}_A^\dagger(t) = k_A^\dagger(t)$ . In general, however, the distributions of  $X$  and  $X^\perp$  are not independent and fitting to the marginal statistics defined through  $p_R(X)$  may not be the optimal strategy. In the next section, we shall see that additional information is needed to achieve a reasonable linear response estimate.

## 6. A severely truncated spatio-temporal system

In this section, we consider a parameter estimation problem where the underlying equilibrium distribution is high-dimensional and unknown. With this constraint, there is a practical issue in the estimation of the linear response statistics  $k_A^\dagger$  for solving the least-squares problem in (7), since the FDT approximation to these statistics requires an explicit form of the underlying density. To overcome this difficulty, we will consider the kernel embedding method [44–47] as a non-parametric estimator of the equilibrium density. We shall see that the resulting linear response statistics obtained from the kernel embedding estimate is more accurate than the Quasi-Gaussian FDT (QG-FDT) linear response [48,49]. To test the performance of the parameter estimation method in recovering the full response statistics, we consider estimating parameters of an arbitrary closure model of a severely truncated Kuramoto-Shivashinsky (KS) equation, which full solutions exhibit spatiotemporal chaotic patterns observed in many applications, such as the trapped ion modes in plasma [50] and phase dynamics in reaction-diffusion systems [51].

### 6.1. The underlying dynamics

The underlying dynamics of the KS equation, defined on a one-dimensional periodic domain  $[0, L]$ , can be approximated on uniformly distributed spatial nodes (e.g., [52–54]) with discrete Fourier modes that satisfy,

$$u_0 = 0, \quad \dot{u}_k = (q_k^2 - q_k^4)u_k - \frac{iq_k}{2} \sum_{1 \leq |\ell|, |k-\ell| \leq \frac{N}{2}-1} u_\ell u_{k-\ell}, \quad |k| \leq \frac{N}{2} - 1. \quad (43)$$

Here,  $q_k = \frac{2\pi k}{L}$  and  $u_k(t)$  denote the Fourier coefficients that satisfy  $u_{-k} = u_k^*$ .

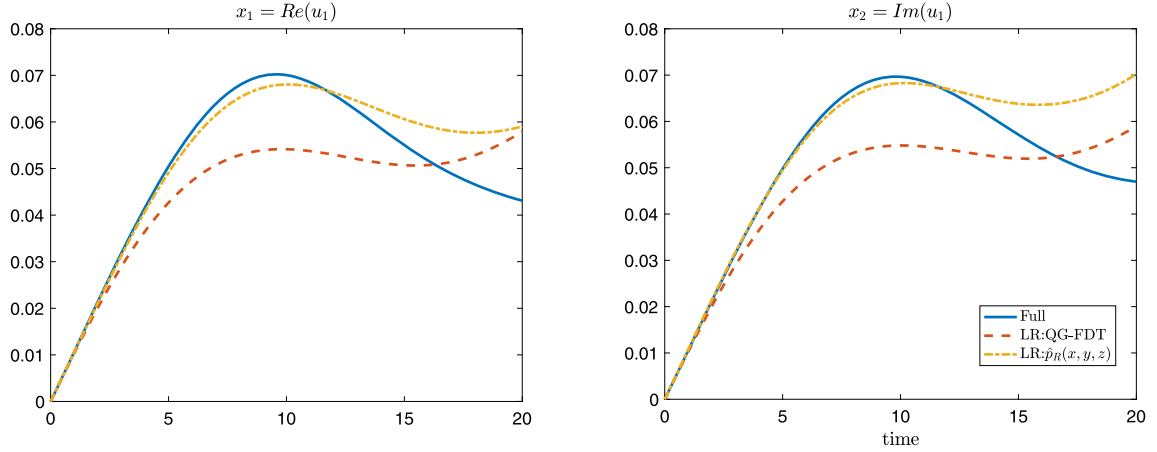
We take the numerical solution to the system (43) with  $N = 48$  as the solutions to the underlying dynamics; this choice of  $N$  is sufficient to recover the inertial manifold of the system [54]. The numerical approximate solution is generated using an exponential time difference fourth order Runge-Kutta (ETDRK4) method introduced in [53] with the discrete Fourier transform of  $(1 + \sin(x))\cos(x)$  on the grid points  $\{x_i = iL/N\}_{i=0}^{N-1}$  as the initial condition. Here, the ETDRK4 method is applied with  $L = 2\pi/\sqrt{0.085}$ . The value of  $L$  is large enough to yield chaotic solutions [52]. We drop the first  $5 \times 10^4$  time units to allow the system to relax, and take  $u_1$  and  $u_2$  in the next  $5 \times 10^5$  time units with step length  $\Delta t = 0.1$  as the training data set. We denote this training data set as  $\{u_{1,t}, u_{2,t}\}_{t=1}^T$  with  $T = 5 \times 10^6$ .

We should point out that the chaotic behavior of the problem is characterized by positive Lyapunov exponents, the exponentially decaying time correlations, and other metrics [55] that are empirically estimated from the numerical solutions of (43). Following many other works that used this model as a benchmark for recovery of the invariant statistics (e.g., [52, 54, 56]), we assume that the numerical solutions of the KS system in (43) constitute a time series that is stationary in time and homogeneous in space. Effectively, this is the ergodicity assumption that is motivated by the empirical characterization of the chaotic behavior of the numerical solutions of (43) as discussed above.

### 6.2. FDT response based on kernel embedding estimates

Our interest is in predicting the full response of the first Fourier mode,  $u_1$ , under small external forcing, without knowing the full underlying dynamics in (43). Instead, we are only given the dynamical equation associated to the first Fourier mode,





**Fig. 5.** The response on  $u_1$  component of the underlying dynamics (43) to a constant external forcing  $\delta f = \frac{1}{N}(0.5 + 0.5i)$ . The blue solid curves show the full nonlinear response of  $u_1$  approximated by Monte-Carlo based on  $5 \times 10^5$  times of simulations; the red dash shows the performance of QG-FDT based on the first six modes  $(u_1, u_2, \dots, u_6)$ ; yellow dot-dash curves correspond to the MLR statistics produced by (52) based on  $(u_1, u_2, g)$  as the resolved variable.

$$\dot{u}_1 = (q_1^2 - q_1^4)u_1 - iq_1u_1u_2 + g, \quad g := -\frac{iq_k}{2} \sum_{1 \leq |\ell|, |k-\ell| \leq \frac{N}{2}-1} u_\ell u_{k-\ell} + iq_1u_1u_2, \quad (44)$$

where  $g$  is the *identifiable unresolved component* in the following sense. Given the partial dynamics (44) and the training data set  $\{u_{1,t}, u_{2,t}\}_{t=1}^T$ , one can extract  $\{g_t\}_{t=1}^T$  by fitting these data set into the first equation in (44), where the time derivative is replaced by a finite difference approximation [54].

As we previously alluded to, the key issue here is that we have no access to the explicit form of the equilibrium distribution of the full dynamics, which is very high-dimensional. Since we only have data set  $\{u_{1,t}, u_{2,t}, g_t\}_{t=1}^T$ , we propose to consider the linear response operator in (5), computed by averaging over  $B$  that is defined with respect to a marginal density,  $p_R$ , of these three variables  $(u_1, u_2, g)$ . Beyond this practical consideration, we should allude that the inclusion of the identifiable unresolved variable  $g$  is essential. To understand this, we first point out that if we follow the notation in (17), then  $X$  corresponds to  $u_1$ . If we only consider the MLR operator with  $p_R$  that depends only on  $u_1$  (or even on  $u_1, u_2$ ), we found that the corresponding linear response estimates do not approximate well the full response statistics. This is not surprising since  $X = u_1$  is not independent to the conjugate variables,  $X^\perp = (u_2, \dots, u_N)$ , as discussed in Remark 5.2. Likewise,  $(u_1, u_2)$  is not independent to its conjugate variables  $(u_3, \dots, u_N)$ . As we shall see below (see Fig. 5), the resulting MLR statistics corresponds to  $p_R$  that depends on  $(u_1, u_2, g)$ , is an accurate estimator of the full nonlinear response statistic of the underlying dynamics under admissible external disturbances up to a finite time. The numerical results suggest that the inclusion of  $(u_2, g)$  in  $p_R$  is important since it carries information in  $X^\perp$  that is relevant for accurate linear response estimations.

To facilitate the following discussion, we define

$$x := (\text{Re}(u_1), \text{Im}(u_1))^\top, \quad y := (\text{Re}(u_2), \text{Im}(u_2))^\top, \quad \text{and} \quad z := (\text{Re}(g), \text{Im}(g))^\top. \quad (45)$$

Under a constant external forcing on  $u_1$ , which is admissible according to the criteria in Definition 4.1, Eq. (22) suggests that, using the real representation (45), the MLR operator of (43) is given by

$$\hat{k}_A^\dagger(t) = \mathbb{E}_{p_{eq}} [A(x(t)) \otimes B(x(0), y(0), z(0))], \quad B(x, y, z) := -\nabla_x \log(p_R(x, y, z)), \quad (46)$$

where the unknown  $p_R$  corresponds to the marginal distribution of  $(x, y, z)$  at the equilibrium of the underlying dynamics (43). While the expectation over  $p_{eq}^\dagger$  can be realized using Monte-Carlo average over the available data, one needs an actual expression of  $p_R$  which appears in  $B$ . This way, we have replaced the problem of estimating a very high-dimensional density of  $(u_1, \dots, u_N)$  to a moderately low-dimensional (in this case, six-dimensional real-valued) function. In the remainder of this section, we will discuss how to approximate  $p_R$  with a kernel embedding estimator,  $\hat{p}_R$ , and check the validity of the MLR in predicting the full response statistics of the underlying dynamics under small perturbations.

Recall that kernel embedding, as a nonparametric approach to learn a distribution function from the sample, can be formulated with an arbitrary kernel that uniquely determines the reproducing kernel Hilbert space (RKHS). We consider kernels defined by the orthogonal basis of a weighted  $L^2$ -space. Specifically, let the target density function  $f \in \mathcal{H}_\rho \subset L^2(\mathbb{R}^d, q^{-1})$ , where  $\mathcal{H}_\rho$  denotes the RKHS induced by a set of orthonormal basis  $\{\Psi_{\vec{m}} = \psi_{\vec{m}} q\}$  of  $L^2(\mathbb{R}^d, q^{-1})$  and  $q$  denotes a positive weight function on  $\mathbb{R}^d$ , (see [45,46] for details) where  $\vec{m} = (m_1, m_2, \dots, m_d)$  is a multi-index notation with

nonnegative components. In particular, we take  $q$  to be the standard  $d$ -dimensional Gaussian distribution, and define the kernel

$$k_\rho(\mathbf{x}, \mathbf{y}) := \sum_{\vec{m} \geq 0} \lambda_{\vec{m}} \Psi_{\vec{m}}(\mathbf{x}) \Psi_{\vec{m}}(\mathbf{y}), \quad \lambda_{\vec{m}} := \rho^{\|\vec{m}\|_1}, \quad \|\vec{m}\|_1 := \sum_{k=1}^d m_k, \quad (47)$$

for  $\rho \in (0, 1)$ . In [47], we have shown that equipped with the inner product

$$\langle f, g \rangle_{\mathcal{H}_\rho} := \sum_{\vec{m} \geq 0} \frac{\hat{f}_{\vec{m}} \hat{g}_{\vec{m}}}{\rho^{\|\vec{m}\|_1}}, \quad f = \sum_{\vec{m} \geq 0} \hat{f}_{\vec{m}} \Psi_{\vec{m}}, \quad g = \sum_{\vec{m} \geq 0} \hat{g}_{\vec{m}} \Psi_{\vec{m}}, \quad (48)$$

$\mathcal{H}_\rho$  is an RKHS with a bounded kernel  $k_\rho$  satisfying

$$k_\rho(\mathbf{x}, \mathbf{y}) = (2\pi)^{-d} (1 - \rho^2)^{-\frac{d}{2}} \exp \left[ -\frac{1}{2(1 - \rho^2)} \left( \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\rho \sum_{i=1}^d x_i y_i \right) \right].$$

Here,  $k_\rho$  (47) can be interpreted as a generalization of Mehler kernel [57]. Then the order- $M$  kernel embedding estimate of  $f$ , denoted by  $f_M$ , is given by

$$f_M := \sum_{\|\vec{m}\|_1 \leq M} \hat{f}_{\vec{m}} \Psi_{\vec{m}}, \quad \hat{f}_{\vec{m}} = \int_{\mathbb{R}^d} f \Psi_{\vec{m}} q^{-1} d\mathbf{x} = \int_{\mathbb{R}^d} f \psi_{\vec{m}} d\mathbf{x}. \quad (49)$$

Since  $\{\psi_{\vec{m}}\}$  forms an orthonormal basis for  $L^2(q)$ , we take

$$\psi_{\vec{m}}(\mathbf{x}) = \prod_{k=1}^d \psi_{m_k}(x_k), \quad \mathbf{x} = (x_1, x_2, \dots, x_d),$$

where  $\psi_{m_k}$  is the order- $m_k$  normalized Hermite polynomial. We should point out with this choice of basis representation, we basically arrive at a polynomial chaos approximation [58] of a density function  $f$ . In practice, the integral in (49) can be approximated by Monte-Carlo,

$$\hat{f}_{\vec{m}} \approx \hat{f}_{\vec{m}, N} = \frac{1}{N} \sum_{n=1}^N \psi_{\vec{m}}(X_n), \quad (50)$$

where  $\{X_n\}_{n=1}^N$  are sampled from the target density function  $f$ . One can define the order- $M$  empirical kernel embedding estimate of  $f$  as

$$f_{M, N} := \sum_{\|\vec{m}\|_1 \leq M} \hat{f}_{\vec{m}, N} \Psi_{\vec{m}}. \quad (51)$$

In our application, the sample  $\{X_n\}$ , as a time series of  $(x, y, z)$  (45), can be interpreted as a stationary process with the target distribution function  $f = p_R$ .

Take the observable  $A(x, y, z) = x$ , and let  $\hat{p}_{R, M}$  be the order- $M$  kernel embedding estimator of  $p_R$ . Then, Eq. (46) is approximated by an order- $M$  MLR estimator,

$$\hat{k}_A^\dagger(t) \approx -\mathbb{E}_{p_{eq}^\dagger} [x(t) \otimes \nabla_x \log(\hat{p}_{R, M}(x(0), y(0), z(0)))]. \quad (52)$$

Recall that the expectation is still over  $p_{eq}^\dagger$  since we are using the observed time series from the underlying dynamics to compute the statistics (52). We should point out that  $\nabla_x \log(\hat{p}_{R, M})$  in Eq. (52) is not an accurate *pointwise* estimator for  $\nabla_x \log(p_R)$  since approximating  $\nabla_x \log(p_R)$  is numerically ill-posed when  $p_R$  is close to zero. However, estimating (46) with the estimator (52) is not an ill-posed problem since the MLR operator  $\hat{k}_A^\dagger$ , as a statistical quantity, is defined as an expectation with respect to  $p_R$ , which cancels out the pointwise error.

**Remark 6.1.** Theoretically, the two-point statistics on the right-hand side of (52) may not be well-defined, since the basis function  $\Psi_{\vec{m}}$  used in the order- $M$  kernel embedding estimate (49) can be negative. We resolve this issue by restricting  $\hat{p}_{R, M}$  in (52) to  $D_{M, \delta} = \{\hat{p}_{R, M} \geq \delta\}$  for  $\delta > 0$  such that the two-point statistics in (52) is well-defined and the error is negligible [47]. In practice, fix a  $\delta > 0$ , for large enough  $M$ , the sample points almost surely stay inside  $D_{M, \delta}$ , and the MLR estimator can be approximated by a Monte-Carlo integral over the entire data set.

Using the approximate response operator in (52), we now compare the resulting approximate linear response statistics with the corresponding full response statistics. The numerical result (Fig. 5) suggests that the full response is well captured by the linear response that uses  $\hat{p}_{R,M}(x, y, z)$  up to finite time. This result empirically validates the MLR response in this model. In fact, the MLR statistic is more accurate compared to the QG-FDT response [48,49], which is obtained from the response operator (6) with  $B^\dagger$  replaced by the one defined with a Gaussian approximation of the invariant density.

### 6.3. The truncated model

The focus in this section is to devise a simple dynamical model that can improve the estimation of the full response, yet retaining the accuracy of the equilibrium distribution of variable  $u_1$  obtained through the kernel embedding approach discussed in Section 6.2. We should point out that while many reduced-order models [52,59,54,56] (just to name a few) have been proposed for the KS model, their constructions typically require the knowledge of more than two Fourier modes. In our case, we only work with the data set of the first two Fourier modes and the resulting identifiable irrelevant variables,  $g$ , as defined in (44). While the dynamical equation for the first component in (44) is given, the fact that the linear term is unstable, since  $q_1^2 - q_1^4 > 0$ , makes it difficult to devise a closure model that is stable and ergodic with accurate invariant density of the variable  $x := (\text{Re}(u_1), \text{Im}(u_1))^\top$ .

Given such constraints, we propose the following “semi-parametric” extended Langevin equation as our imperfect model

$$\begin{aligned}\dot{x} &= v, \\ \dot{v} &= \Lambda \nabla_x \log(\hat{\rho}(x)) - \Gamma v + \sigma \dot{W}_t,\end{aligned}\tag{53}$$

where  $x \in \mathbb{R}^2$  is the real representation of  $u_1$  as in (45) and  $W_t$  denotes the standard two-dimensional Wiener process. By “semi-parametric”, we refer to the combination of linear parametric equation in the right-hand-side with a “nonparametric” term that involves  $\hat{\rho}(x)$  that is estimated by the kernel embedding approximation of the marginal distribution of  $x$  at equilibrium. We should point out that since the density  $\hat{\rho}$  is approximated using a specific choice of kernel with Hermite polynomials, as explained in Section 6.2, the resulting model in (53) is effectively parametric.

**Remark 6.2.** Assume the extended Langevin equation (53) satisfies the following conditions.

1.  $\Lambda$  is SPD,  $\Gamma$  is positive definite, and  $\sigma$  satisfies the Lyapunov equation

$$\Gamma \Lambda + \Lambda \Gamma^\top = \sigma \sigma^\top.\tag{54}$$

2.  $\hat{\rho}(x)$  is a bounded smooth density function defined on  $\mathbb{R}^2$ , and there exist constants  $D, E > 0$  and  $F$ , such that

$$-\langle \nabla_x \log(\hat{\rho}(x)), x \rangle \geq -D \log(\hat{\rho}(x)) + E \|x\|_2^2 + F, \quad \forall x \in \mathbb{R}^2.$$

Then the geometric ergodicity of the imperfect model (53) follows from Theorem 3.2 in [60]. In particular, consider a change of variables

$$q := \Lambda^{-\frac{1}{2}} x, \quad p := \Lambda^{-\frac{1}{2}} v,$$

then  $(q, p)$  satisfies

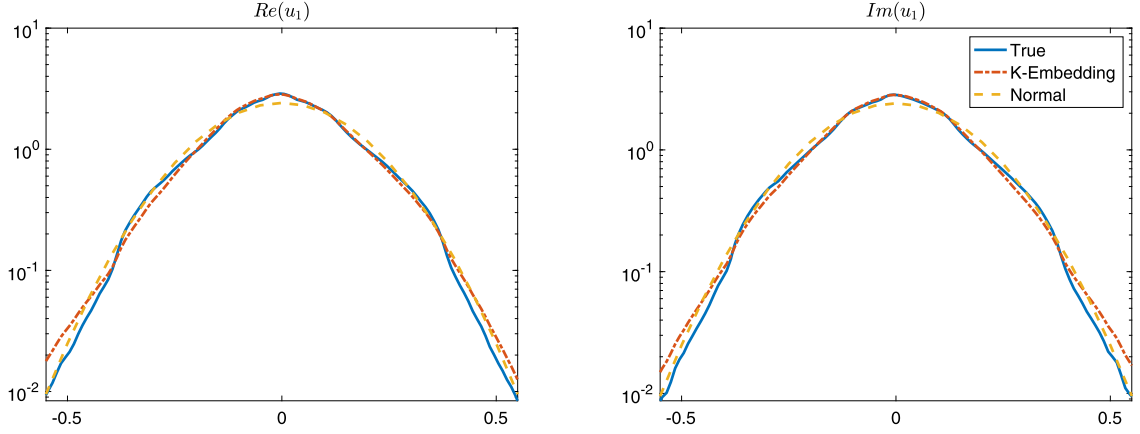
$$\begin{aligned}\dot{q} &= p, \\ \dot{p} &= \nabla_q \log(\hat{\rho}(\Lambda^{\frac{1}{2}} q)) - \tilde{\Gamma} p + \tilde{\sigma} \dot{W}_t,\end{aligned}\tag{55}$$

with  $\tilde{\Gamma} = \Lambda^{-\frac{1}{2}} \Gamma \Lambda^{\frac{1}{2}}$ ,  $\tilde{\sigma} = \Lambda^{-\frac{1}{2}} \sigma$ , which suggests that

$$\rho_{eq}(x, v) \propto \hat{\rho}(x) \exp\left(-\frac{1}{2} v^\top \Gamma^{-1} v\right)\tag{56}$$

is the unique equilibrium distribution of the imperfect model (53).

In this example, choosing Langevin dynamic as the imperfect model is motivated largely by the separable structure of its equilibrium distribution as shown in (56). Following the notation in (21),  $X = x$  and  $Y = v$ . In this case, the auxiliary variable  $v$  is Gaussian with parameters that can be estimated directly from its empirical variance. As we have mentioned at the beginning of this section, while other choices of imperfect models that involve variables  $(u_2, g)$  with accurate equilibrium density estimate  $\hat{\rho}(x)$  are possible, they tend to be more complicated and required higher wavenumbers  $(u_3, u_4, u_5, u_6)$  as reported in [54,56].



**Fig. 6.** The kernel embedding approximation of the marginal distribution of  $x$  in (45). The blue solid curves are the PDFs obtained from the observed time series of  $u_1^N$  using kernel density estimates for smoothing. The red dot-dash curves are the order-5 kernel embedding estimates. The yellow dash curves are the density function of the normal distribution of the same mean and variance.

Fig. 6 shows the performance of the kernel embedding in estimating  $\hat{\rho}(x)$ . One can see that, compared with the Gaussian approximation, the kernel embedding estimator has a better performance near the peaks of the target density functions. For the corresponding parameter estimation problem, with the Lyapunov equation (54), the parameter space of (53) can be reduced to

$$\theta := (\Lambda, \Gamma), \quad (57)$$

where  $\Lambda$ , as a covariance matrix of the auxiliary variable  $v$ , is SPD and  $\Gamma$  is positive definite. Connecting to Section 4, components of  $\Lambda$  are the equilibrium parameters,  $\theta_{eq}$ .

#### 6.4. Parameter estimation and numerical results

Next, we will estimate the parameters  $\theta = (\Lambda, \Gamma)$  using the scheme introduced in Section 4 and check whether the imperfect model in (53) with the estimated parameters can produce a better estimate of the full response statistics compared to the linear responses as shown in Fig. 5. First, notice that the imperfect model (53) yields an equilibrium distribution (56), which suggests the auxiliary variable  $v$  is Gaussian with covariance matrix  $\Lambda$  at the equilibrium. Thus, we can use the sampled covariance of  $\dot{x}$ , approximated through a finite-difference, to directly estimate  $\Lambda$ . We will denote the estimate for  $\Lambda$  as  $\hat{\Lambda}$ .

For the damping coefficient  $\Gamma$ , we take the diagonal elements of the approximated linear response operator (52) as the essential statistics (they are shown as the yellow stars in Fig. 7), and we are able to set up the following nonlinear least square problems to estimate  $\Gamma$ :

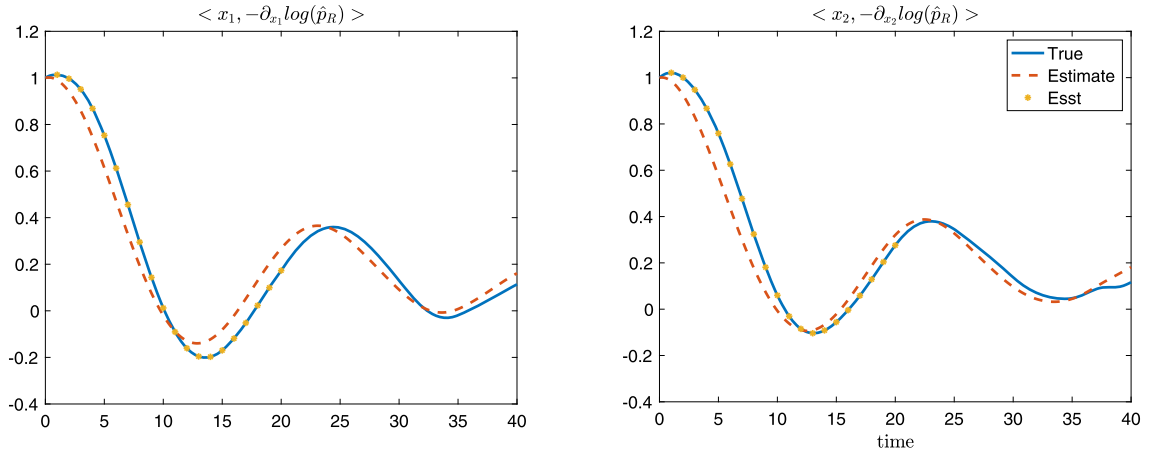
$$\min_{\Gamma} \sum_{i=1}^K \sum_{j=1}^2 \left( g_j^\dagger(t_i) - \hat{g}_j(t_i; \Gamma, \hat{\Lambda}) \right)^2, \quad (58)$$

where

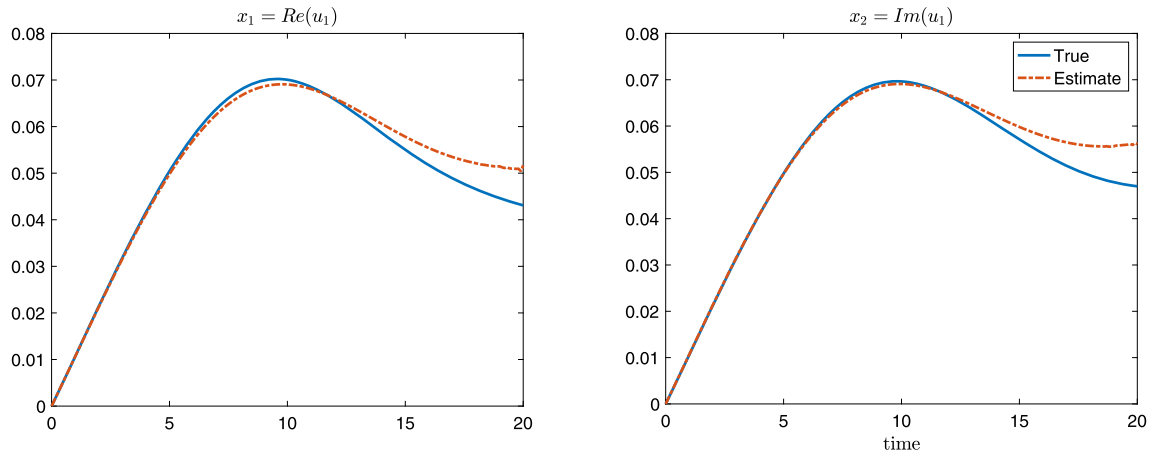
$$g_j^\dagger(t) = \mathbb{E}_{p_{eq}} [x_j(t) \partial_{x_j} \log(\hat{p}_{R,M})], \quad \hat{g}_j(t; \Gamma, \hat{\Lambda}) = \mathbb{E}_{\rho_{eq}(\Gamma, \hat{\Lambda})} [x_j(t) \partial_{x_j} \log(\hat{\rho})], \quad j = 1, 2.$$

In particular we take  $t_i = i$  for  $i = 1, 2, \dots, 20$ . Here, the step length of the marginal essential statistics is larger than that of the MD example in (42) since the correlation times of the linear response operator are different, as suggested by Fig. 3 and Fig. 7. We should point out that the  $\mathbb{E}_{p_{eq}}^\dagger$  will be approximated by Monte-Carlo over the true time series of  $\{u_{1,t}, u_{2,t}, g_t\}_{t=1, \dots, T}$ . We use Euler-Maruyama method to generate time series from the imperfect dynamics at  $\Gamma$  and  $\hat{\Lambda}$ . Similar to the setup of the numerical scheme of the underlying dynamics (43), we drop the first  $5 \times 10^4$  time units to relax the system, and sub-sample  $x$  in the next  $5 \times 10^5$  time units with step length  $\Delta t = 0.1$ . Table 3 records the estimated parameter values.

In Fig. 7, one can see that the estimated linear response statistics (red dash curves) from the imperfect model reproduces the qualitative feature of the true linear response statistics. Fig. 8 shows the recovery of the full response using the imperfect model in (53) based on the estimated parameters in Table 3. One can see that the real and imaginary components of  $u_1$  share similar behavior in both the two-point statistics and the full response statistics. We note that the imperfect model is able to capture such symmetry. Compared with the approximated linear response statistics showed in Fig. 5, the full response statistics of the imperfect model produce a more accurate estimation for larger values of  $t$ .



**Fig. 7.** Results of the nonlinear least squares fitting (58). The values of the essential statistics  $\{g_j^\dagger(t_i)\}$ , taken from the blue curves, are highlighted by yellow stars. The red dash curves are the corresponding two-time statistics of the imperfect model (53) the estimated parameters reported in Table 3.



**Fig. 8.** The prediction of the full response. The blue solid curves indicate the same full response in Fig. 5; The red dot-dash curves are the full response from the imperfect model (53) based on the parameter estimates reported in Table 3.

**Table 3**  
Estimates of the parameter in (57).

$\Lambda_{11}$	$\Lambda_{12}$	$\Lambda_{22}$	$\Gamma_{11}$	$\Gamma_{12}$	$\Gamma_{21}$	$\Gamma_{22}$
$1.3675 \times 10^{-3}$	$-1.01 \times 10^{-5}$	$1.3544 \times 10^{-3}$	0.5327	-0.5038	-0.4953	0.6176

## 7. Summary and discussion

Motivated by FDT, we have demonstrated a response-based estimation method to determine parameters that are responsible for the response statistics in stochastic dynamics. Unlike the prior works [6,7], which were focused on the perfect model case, the emphasis of this paper is on a more practical scenario where the model error is present. We started with a conventional two-scale model and showed that the response-based approach can outperform a classical averaging method that is based on the equilibrium statistics, even in the regime where the averaging theory is valid.

More importantly, we demonstrated that in general, the parameter estimation problems are far more complicated by the fact that only partial observations are available and the response statistics might not be accessible. To overcome these difficult issues, we have outlined a general strategy that can be efficiently implemented in practice. In particular, we introduced the concepts of admissible external forcing, for which the forces on the variable of interest can be identified and computed, as well as the marginal linear response (MLR) statistics, as computable statistics in place of the linear response statistics of the underlying full dynamics that are inaccessible in general.

To illustrate this general strategy in a specific context, we considered the Langevin dynamics model, a widely used stochastic model in practice. We studied two examples, including a one-dimensional chain model, coarse-grained from a

full MD model; and an extended Langevin dynamics, obtained from a severely truncated KS equation in the Fourier space. In both cases, we demonstrate how our general strategy is implemented and validated. More specifically, after the model parameters are determined from the least-squares approach, we compared the *nonlinear* responses of the imperfect model with those of the underlying dynamics under admissible forcing. A quantitative agreement has been found.

For the second example, we also employed a nonparametric kernel embedding formulation to estimate the marginal density, and more importantly, the marginal essential statistics, to enable the estimation using the response statistics. We report the convergence analysis of this estimator in [47].

### CRedit authorship contribution statement

**He Zhang:** Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft. **John Harlim:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Visualization, Writing – review & editing. **Xiantao Li:** Funding acquisition, Methodology, Resources, Software, Writing – review & editing.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgement

This research was supported by NSF under grant DMS-1619661 and DMS-1819011. JH also acknowledges the support from NSF grant DMS-1854299.

### Appendix A. Approximation of the full response statistics via Monte-Carlo

To test the validity of the parameter estimates, we compare the full response statistics of the underlying dynamics and those of the imperfect model. For example, in Section 5, to generate Fig. 4, one needs to compute the full response statistics of the underlying dynamics (27) and the imperfect dynamics (31) under a constant external forcing on the first block of particles and the first CG particle, respectively. In this Appendix, we review the Monte-Carlo (MC) approach used in approximating the full response statistics  $\Delta\mathbb{E}[A](t)$  (3) under a general Itô diffusion setting.

Recall that in Section 2, we have used (1) and (2) to represent the unperturbed and perturbed dynamics respectively. In the verification, the unknown parameter  $\theta$  in the two dynamics will be replaced by the estimate  $\hat{\theta}$ . To compute  $\Delta\mathbb{E}[A](t)$ , since  $\mathbb{E}_{p_{eq}}[A(X)]$  in (3) is an equilibrium statistics of the unperturbed dynamics, it can be directly approximated by an ensemble average

$$\mathbb{E}_{p_{eq}}[A(X)] \approx \frac{1}{N} \sum_{i=1}^N A(x_i), \quad (\text{A.1})$$

where  $\{x_i = X(t_i)\}$  denotes a time series of the unperturbed dynamics (1) generated at the equilibrium state. Thus, we are able to reduce our problem to approximating the response statistics  $\mathbb{E}_{p^\delta}[A(X^\delta(t))]$ .

Introducing the transition kernel  $\rho^\delta(x, t|x_0, 0)$  of the perturbed dynamics (2),  $\mathbb{E}_{p^\delta}[A(X^\delta(t))]$  can be calculated by a double integral

$$\mathbb{E}_{p^\delta}[A(X^\delta(t))] = \int \int A(x) \rho^\delta(x, t|x_0, 0) p_{eq}(x_0) dx dx_0, \quad (\text{A.2})$$

where we have used the fact that the perturbed dynamics (2) is also initiated at the equilibrium state of the unperturbed dynamics (1) with equilibrium distribution  $p_{eq}$ . Further notice that the transition kernel satisfies the Fokker-Planck equation

$$\frac{\partial \rho^\delta}{\partial t} = (\mathcal{L}^\delta)^* \rho^\delta, \quad \rho^\delta(x, 0|x_0, 0) = \delta(x - x_0),$$

where  $\mathcal{L}^\delta$  denotes the generator of the perturbed dynamics (2). One can rewrite the transition kernel using the semi-group notation:  $\rho^\delta(x, t|x_0, 0) = e^{t(\mathcal{L}^\delta)^*} \delta(x - x_0)$ . Then, the double integral (A.2) becomes

$$\begin{aligned} \mathbb{E}_{p^\delta}[A(X^\delta(t))] &= \int \int A(x) e^{t(\mathcal{L}^\delta)^*} \delta(x - x_0) p_{eq}(x_0) dx dx_0 \\ &= \int e^{t\mathcal{L}^\delta} A(x) \int \delta(x - x_0) p_{eq}(x_0) dx_0 dx \\ &= \int e^{t\mathcal{L}^\delta} A(x) p_{eq}(x) dx. \end{aligned} \quad (\text{A.3})$$



Here,  $u^\delta(x, t) := e^{t\mathcal{L}^\delta} A(x)$  solves the Kolmogorov's backward equation (dual form of the Fokker-Planck equation)

$$\frac{\partial u^\delta}{\partial t} = \mathcal{L}^\delta u^\delta, \quad u^\delta(x, 0) = A(x),$$

which suggests that  $u^\delta(x, t)$  corresponds to the conditional expectation  $\mathbb{E}[A(X^\delta(t)) | X^\delta(0) = x]$ . Together with Eq. (A.3), we derive the following two-layer Monte-Carlo approximation,

$$\mathbb{E}_{p^\delta}[A(X^\delta(t))] \approx \frac{1}{N} \sum_{i=1}^N u^\delta(x_i, t) \approx \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M X^{\delta, x_i}(t; U_j), \quad (\text{A.4})$$

where  $X^{\delta, x_i}(t; U_j)$  denotes the realization of (2) with initial condition  $X^\delta(0) = x_i$ , given a sample path  $U_j$  of the Wiener process  $U_s$  ( $0 \leq s \leq t$ ) in (2).

In the numerical implementation, after generating the time series  $\{x_i\}$  at  $p_{eq}$ , one can duplicate  $x_i$  up to  $M$  times and solve (2) in parallel based on the samples from the time series as initial conditions. In particular, we take  $N = 10^5$  and  $M = 10^3$  in estimating the full response statistics of the underlying dynamics and imperfect model shown in Fig. 4.

## Appendix B. Heat flux and the density ansatz

The main goal of this Appendix is to deduce a parametric form of the equilibrium density (30) of the coarse-grained system. A system with non-uniform temperatures usually exhibits heat conduction, in which case the induced heat flux becomes relevant in our parameter estimation procedure (e.g., [36,61]). In Section 5, the local heat flux  $j_i$  corresponding to the definition of the CG variables  $(\xi, \mathbf{p})$  (29,36) is given by,

$$j_i \propto (p_{i+1} + p_i) V_i'(\xi_i),$$

where  $V_i$  denotes the local potential (e.g., [36]).

Let  $\rho$  be the equilibrium distribution of the CG variables in (31). It is important to note that this notion of equilibrium density is more general due to the non-zero current. Such a density function  $\rho$  corresponds to the steady-state solution of the Fokker-Planck equation of the CG model. Except for very special cases, the explicit form of  $\rho$  is unknown. Due to the presence of heat conduction, it is necessary to incorporate the heat flux in approximating the density function  $\rho$ .

Motivated by the maximum entropy approach, we consider the following two set of constraints on the local Hamiltonian  $\mathcal{H}_i := \frac{p_i^2}{2} + V_i$  and  $j_i$ ,

$$\mathbb{E}_\rho[\mathcal{H}_i] = \mathbb{E}_{p_{eq}}^\dagger[\mathcal{H}_i], \quad \mathbb{E}_\rho[j_i] = \mathbb{E}_{p_{eq}}^\dagger[j_i], \quad i = 1, 2, \dots, n_b, \quad (\text{B.1})$$

which can be interpreted as a system of integral equations to  $\rho$ . Each heat flux constraint in (B.1) involves a three-dimensional integral with respect to  $\xi_i$ ,  $p_i$ , and  $p_{i+1}$ , which causes issues in our later applications of maximum entropy. We resolve such an issue by exploiting the equality using the stationarity,

$$\frac{d}{dt} \mathbb{E}_{p_{eq}}^\dagger[V_i(\xi_i)] = 0, \quad i = 1, 2, \dots, n_b,$$

which implies that

$$\mathbb{E}_{p_{eq}}^\dagger[p_i V_i'(\xi_i)] = \mathbb{E}_{p_{eq}}^\dagger[p_{i+1} V_i'(\xi_i)].$$

As a result, the constraints in (B.1) can be equivalently rewritten as,

$$\mathbb{E}_\rho[\mathcal{H}_i] = \mathbb{E}_{p_{eq}}^\dagger[\mathcal{H}_i], \quad \mathbb{E}_\rho[p_i V_i'(\xi_i)] = \mathbb{E}_{p_{eq}}^\dagger[p_i V_i'(\xi_i)], \quad i = 1, 2, \dots, n_b, \quad (\text{B.2})$$

such that each constraint is a two-dimensional integral equation. With constraints in (B.2), the corresponding maximum entropy solution is of the form

$$\rho(\xi, \mathbf{p}) \propto \exp \left[ - \sum_{i=1}^{n_b} b_i \left( \frac{p_i^2}{2} + V_i(\xi_i) \right) - \sum_{i=1}^{n_b} c_i p_i V_i'(\xi_i) \right], \quad (\text{B.3})$$

together with the periodic boundary condition. Such an ansatz in (B.3) can be interpreted as a perturbed Gibbs measure [26].

The potential function  $V_i$  in  $\rho$  (B.3) is assumed to be a quartic function of the following form,

$$V_i(\xi_i) = \sum_{j=1}^4 A_{j,i} \xi_i^j, \quad (\text{B.4})$$

with unknown parameters  $\{A_{j,i}\}$ . Substituting (B.4) to (B.3), we obtain (30) with parameters:  $a_{j,i} = A_{j,i}b_i$  for  $j = 1, \dots, 4$ , and  $c_{j,i} = (j+1)A_{j+1,i}C_i$  for  $j = 1, \dots, 3$ . The restriction of parameters makes the maximum entropy formulation to be inconvenient in finding  $b_i$  and  $c_i$ . To overcome this issue, we keep the same ansatz for the equilibrium density in (30) and solve the maximum entropy problem in (38)-(39) for each coarse grained component.

### Appendix C. The isothermal one-dimensional molecular model

In Section 5, we emphasized on the parameter estimation problem of the imperfect model of a coarse-grained molecular model (27) with non-uniform temperature. We also tested the isothermal case, where  $k_B T_i \equiv k_B T = 0.25$  in (27) for  $i = 1, 2, \dots, N$ . We included the results in this section for interested readers. As a simplification of (27),  $(\mathbf{r}, \mathbf{v})$  obeys the Langevin equation of motion,

$$\begin{aligned}\dot{\mathbf{r}} &= \mathbf{v}, \\ \dot{\mathbf{v}} &= -\nabla_{\mathbf{r}} U(\mathbf{d}) - \gamma \mathbf{v} + \sqrt{2k_B T \gamma} \dot{W}_t.\end{aligned}\quad (\text{C.1})$$

The Langevin dynamics is a mean to sample canonical ensemble [62], or the Gibbs measure. The equilibrium distribution of the relative displacement  $\mathbf{d}$  and the velocity  $\mathbf{v}$ , is of the form

$$p_{eq}(\mathbf{d}, \mathbf{v}) \propto \exp \left[ -\frac{1}{k_B T} \left( U(\mathbf{d}) + \frac{1}{2} \mathbf{v}^2 \right) \right], \quad (\text{C.2})$$

under the constraint induced by the periodic boundary condition.

In this isothermal case, the CG model in (31) reduces to

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{q}} V(\xi; \mathbf{a}) - \Gamma \mathbf{p} + \sqrt{2\beta^{-1} \sigma} \dot{U}_t,\end{aligned}\quad (\text{C.3})$$

where the damping matrix  $\Gamma$  satisfies  $\sigma \sigma^\top = \Gamma$ , and the PMF  $V(\xi; \mathbf{a})$ , as a function of the CG relatively displacement  $\xi$  is of the form

$$V(\xi, \mathbf{a}) = \sum_{i=1}^J \sum_{j=1}^4 a_j \xi_i^j, \quad q_{k+n_b} = q_k, \quad a_4 > 0, \quad (\text{C.4})$$

with unknown parameters  $\mathbf{a} = (a_1, a_2, a_3, a_4)$ . Adopting the form of the damping matrix  $\Gamma$  in (34), the parameter set of the imperfect model (C.3) is

$$\theta = (\mathbf{a}, \beta, \gamma_0, \gamma_1, \gamma_2),$$

with  $\theta_{eq} := (\mathbf{a}, \beta)$ . To determine  $\theta_{eq}$ , consider the Gibbs measure of the imperfect model (C.3) (following the notation in (21)),

$$\rho_{eq}(\xi, \mathbf{p}; \theta_{eq}) \propto \exp \left[ -\beta \left( V(\xi; \mathbf{a}) + \frac{1}{2} \mathbf{p}^2 \right) \right], \quad (\text{C.5})$$

which suggests that the temperature  $\beta^{-1}$  can be directly estimated from the sample variance of the CG velocities. In practice, since  $\{p_i\}$  are independent identical distributed at the equilibrium, we take

$$\hat{\beta}^{-1} = \mathbb{E}_{p_{eq}^\dagger} [p_1^2], \quad (\text{C.6})$$

as the estimate for  $\beta^{-1}$ . For the parameters in the potential (C.4), we apply the maximum entropy method to the marginal distribution of  $q_{i+1} - q_i$  with respect to  $\rho_{eq}$  (C.5),

$$\rho(x; \mathbf{a}, \beta) \propto \exp \left( -\beta \sum_{j=1}^4 a_j x^j \right),$$

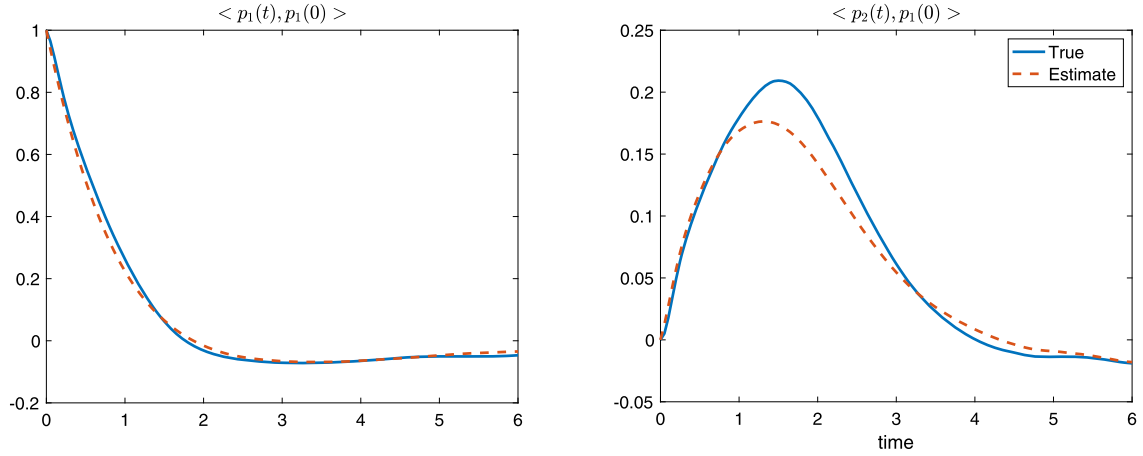
for  $i = 1, 2, \dots, n_b - 1$ . By maximizing the Shannon's entropy under the moment constraints, we formulate the following maximum entropy estimates for  $\mathbf{a}$ ,

$$\begin{aligned}\hat{\mathbf{a}} &= \arg \min \int -\rho(x; \mathbf{a}, \hat{\beta}) \log \rho(x; \mathbf{a}, \hat{\beta}) dx, \\ \text{s.t. } \int x^j \rho(x; \mathbf{a}, \hat{\beta}) dx &= \frac{1}{n_b - 1} \sum_{i=1}^{n_b-1} \mathbb{E}_{p_{eq}^\dagger} [\xi_i^j], \quad j = 1, 2, 3, 4,\end{aligned}$$

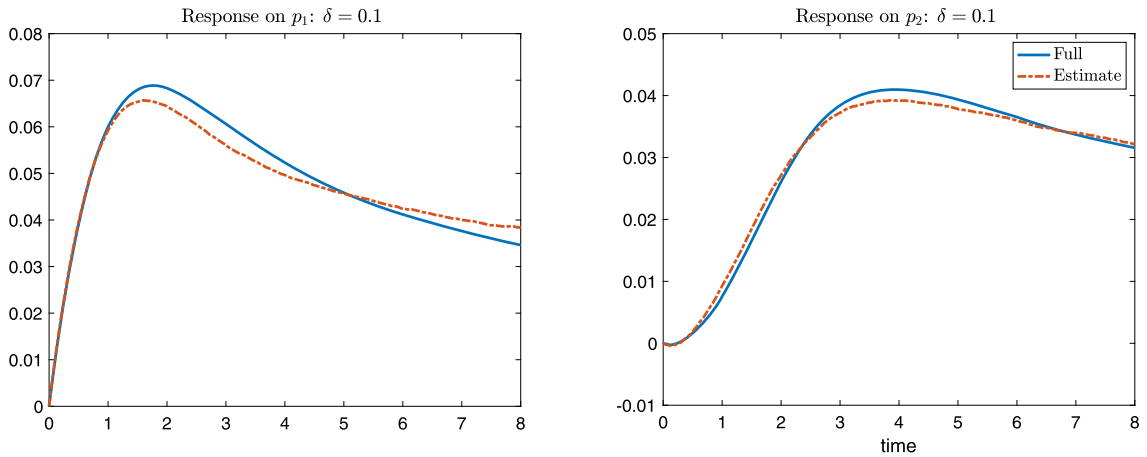
**Table C.4**

Estimates of the parameter in (C.3). The scale of  $a_j$  is different from the scale of  $a_{j,i}$  in Table 2 since  $\beta$  is a prefactor of potential function  $V$  in (C.5).

Parameter ( $\theta$ )	$\beta^{-1}$	$a_1$	$a_2$	$a_3$	$a_4$	$\gamma_0$	$\gamma_1$	$\gamma_2$
Estimates ( $\hat{\theta}$ )	0.0247	0.0193	0.148	-0.0830	0.0331	1.23	-0.371	-0.0420



**Fig. C.9.** Comparison of the time correlation functions involved in the nonlinear least-squares problem. The blue solids curves are computed from the available CG observations, while the red dash curves are the approximation of the imperfect model based on the estimates.



**Fig. C.10.** Reconstruction of full response statistics: underlying dynamics (blue solid) v.s. the imperfect model (red dot-dash).

where  $\hat{\beta}$  is given by (C.6).

For parameters  $\{\gamma_i\}$ , we take the same approach as in Section 5. Table C.4 shows the value of the estimates. Fig. C.9 and Fig. C.10 provide the results in fitting the essential statistics and recovering the full response of the underlying dynamics, respectively.

As in Section 5, we consider the damping matrix  $\Gamma$  in the form of symmetric and circulant matrix with non-zero elements  $(\gamma_0, \gamma_1, \gamma_2)$ . In our numerical experiments of the isothermal molecular model, we also tried damping matrix  $\Gamma$  with smaller bandwidth. We found that using  $\Gamma = \gamma_0 I$  produces a much less accurate recovery of the time correlation (results are not reported) compared to those obtained using non-diagonal  $\Gamma$  as presented in Fig. C.9. We also did not continue with testing the damping matrix with larger bandwidth since the estimates of  $\gamma_2$  are already small.

## References

- [1] W. Kohn, L.J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.* 140 (4A) (1965) A1133–A1138, <https://doi.org/10.1103/PhysRev.140.A1133>.
- [2] E. Runge, E.K. Gross, Density-functional theory for time-dependent systems, *Phys. Rev. Lett.* 52 (12) (1984) 997.

- [3] R. Haydock, V. Heine, M. Kelly, Electronic structure based on the local atomic environment for tight-binding bands, *J. Phys. C, Solid State Phys.* 5 (20) (1972) 2845.
- [4] D. Hamann, M. Schlüter, C. Chiang, Norm-conserving pseudopotentials, *Phys. Rev. Lett.* 43 (20) (1979) 1494.
- [5] M. Toda, R. Kubo, N. Hashitsume, *Statistical Physics II. Nonequilibrium Statistical Mechanics*, Springer, 1983.
- [6] J. Harlim, X. Li, H. Zhang, A parameter estimation method using linear response statistics, *J. Stat. Phys.* 168 (2017) 146–170.
- [7] H. Zhang, X. Li, J. Harlim, A parameter estimation method using linear response statistics: numerical scheme, *Chaos* 29 (3) (2019) 033101.
- [8] D.J. Evans, G. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, Cambridge University Press, 2008.
- [9] M. Baaden, S.J. Marrink, Coarse-grain modelling of protein-protein interactions, *Curr. Opin. Struct. Biol.* 23 (2013) 878–886, <https://doi.org/10.1016/j.sbi.2013.09.004>, <http://linkinghub.elsevier.com/retrieve/pii/S0959440X13001735>.
- [10] A.R. Leach, *Molecular Modelling: Principles and Applications*, Pearson Education, 2001.
- [11] G.A. Voth, *Coarse-Graining of Condensed Phase and Biomolecular Systems*, CRC Press, 2008.
- [12] T. Schlick, *Molecular Modeling and Simulation: An Interdisciplinary Guide*, Springer-Verlag, 2002.
- [13] Z. Zhang, L. Lu, W.G. Noid, V. Krishna, J. Pfandtner, G.A. Voth, A systematic methodology for defining coarse-grained sites in large biomolecules, *Biophys. J.* 95 (11) (2008) 5073–5083.
- [14] W. Noid, Perspective: coarse-grained models for biomolecular systems, *J. Chem. Phys.* 139 (9) (2013) 09B201\_1.
- [15] W. Noid, J.-W. Chu, G.S. Ayton, V. Krishna, S. Izvekov, G.A. Voth, A. Das, H.C. Andersen, The multiscale coarse-graining method. I. A rigorous bridge between atomistic and coarse-grained models, *J. Chem. Phys.* 128 (24) (2008) 244114.
- [16] C. Chipot, A. Pohorille, *Free Energy Calculations*, Springer, 2007.
- [17] A. Dequidt, J.G. Solano Canchaya, Bayesian parametrization of coarse-grain dissipative dynamics models, *J. Chem. Phys.* 143 (8) (2015) 084122.
- [18] M.A. Katsoulakis, P. Plecháč, Information-theoretic tools for parametrized coarse-graining of non-equilibrium extended systems, *J. Chem. Phys.* 139 (7) (2013) 074115.
- [19] H. Lei, N.A. Baker, X. Li, Data-driven parameterization of the generalized Langevin equation, *Proc. Natl. Acad. Sci.* 113 (50) (2016) 14183–14188.
- [20] L. Ma, X. Li, C. Liu, Coarse-graining Langevin dynamics using reduced-order techniques, *J. Comput. Phys.* 380 (2019) 170–190.
- [21] L. Ma, X. Li, C. Liu, The derivation and approximation of coarse-grained dynamics from Langevin dynamics, *J. Chem. Phys.* 145 (20) (2016) 204117.
- [22] R.B. Best, B. Li, A. Steward, V. Daggett, J. Clarke, Can non-mechanical proteins withstand force? Stretching barnase by atomic force microscopy and molecular dynamics simulation, *Biophys. J.* 81 (4) (2001) 2344–2356.
- [23] H. Lu, K. Schulten, Steered molecular dynamics simulations of force-induced protein domain unfolding, *Proteins, Struct. Funct. Bioinform.* 35 (4) (1999) 453–463.
- [24] C. Leith, Climate response and fluctuation dissipation, *J. Atmos. Sci.* 32 (10) (1975) 2022–2026.
- [25] G.A. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*, vol. 60, Springer, 2014.
- [26] R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford University Press, 2001.
- [27] M. Hairer, A.J. Majda, A simple framework to justify linear response theory, *Nonlinearity* 23 (4) (2010) 909.
- [28] M. Baiesi, C. Maes, B. Wynants, Fluctuations and response of nonequilibrium states, *Phys. Rev. Lett.* 103 (1) (2009) 010602.
- [29] A.J. Majda, D. Qi, Improving prediction skill of imperfect turbulent models through statistical response and information theory, *J. Nonlinear Sci.* 26 (1) (2016) 233–285.
- [30] D. Qi, A.J. Majda, Predicting fat-tailed intermittent probability distributions in passive scalar turbulence with imperfect models through empirical information theory, *Commun. Math. Sci.* 14 (6) (2016) 1687–1722.
- [31] A.J. Majda, D. Qi, Strategies for reduced-order models for predicting the statistical responses and uncertainty quantification in complex turbulent dynamical systems, *SIAM Rev.* 60 (3) (2018) 491–549.
- [32] G.A. Gottwald, J. Harlim, The role of additive and multiplicative noise in filtering complex dynamical systems, *Proc. R. Soc. A, Math. Phys. Eng. Sci.* 469 (2155) (2013) 20130096.
- [33] G. Pavliotis, A. Stuart, *Multiscale Methods: Averaging and Homogenization*, Springer Science & Business Media, 2008.
- [34] T. Berry, J. Harlim, Linear theory for filtering nonlinear multiscale systems with model error, *Proc. R. Soc. A, Math. Phys. Eng. Sci.* 470 (2167) (2014) 20140168.
- [35] Y. Ishimori, Solitons in a one-dimensional Lennard-Jones lattice, *Prog. Theor. Phys.* 68 (2) (1982) 402–410.
- [36] S. Lepri, R. Livi, A. Politi, Thermal conduction in classical low-dimensional lattices, *Phys. Rep.* 377 (1) (2003) 1–80.
- [37] C. Dellago, P.G. Bolhuis, D. Chandler, Efficient transition path sampling: application to Lennard-Jones cluster rearrangements, *J. Chem. Phys.* 108 (22) (1998) 9236–9245.
- [38] M. Sachs, B. Leimkuhler, V. Danos, Langevin dynamics with variable coefficients and nonconservative forces: from stationary states to numerical methods, *Entropy* 19 (12) (2017) 647.
- [39] C.R. Sweet, P. Petrone, V.S. Pande, J.A. Izaguirre, Normal mode partitioning of Langevin dynamics for biomolecules, *J. Chem. Phys.* 128 (14) (2008) 04B606.
- [40] E. Jaynes, Information theory and statistical mechanics, *Phys. Rev.* 106 (4) (1957) 620.
- [41] W. Hao, J. Harlim, An equation-by-equation method for solving the multidimensional moment constrained maximum entropy problem, *Commun. Appl. Math. Comput. Sci.* 13 (2) (2018) 189–214.
- [42] A.P. Lyubartsev, A. Laaksonen, Calculation of effective interaction potentials from radial distribution functions: a reverse Monte Carlo approach, *Phys. Rev. E* 52 (4) (1995) 3730.
- [43] D.F. Anderson, J.C. Mattingly, A weak trapezoidal method for a class of stochastic differential equations, *Commun. Math. Sci.* 9 (1) (2011) 301–318.
- [44] K. Muandet, K. Fukumizu, B. Sriperumbudur, B. Schölkopf, et al., Kernel mean embedding of distributions: a review and beyond, *Found. Trends Mach. Learn.* 10 (1–2) (2017) 1–141.
- [45] S.W. Jiang, J. Harlim, Parameter estimation with data-driven nonparametric likelihood functions, *Entropy* 21 (6) (2019) 559.
- [46] S.W. Jiang, J. Harlim, Modeling of missing dynamical systems: Deriving parametric models using a nonparametric framework, *Res. Math. Sci.* 7 (3) (2020) 1–25, Springer.
- [47] H. Zhang, J. Harlim, X. Li, Estimating linear response statistics using orthogonal polynomials: An RKHS formulation, *Found. Data Sci.* 2 (4) (2020) 443–485, <https://doi.org/10.3934/fods.2020021>.
- [48] A. Majda, R. Abramov, M. Grote, *Information Theory and Stochastics for Multiscale Nonlinear Systems*, CRM Monograph Series, vol. 25, American Mathematical Society, Providence, Rhode Island, USA, 2005.
- [49] A. Gritsun, G. Branstator, Climate response using a three-dimensional operator based on the fluctuation–dissipation theorem, *J. Atmos. Sci.* 64 (7) (2007) 2558–2575, <https://doi.org/10.1175/JAS3943.1>.
- [50] R.E. LaQuey, S. Mahajan, P. Rutherford, W. Tang, Nonlinear saturation of the trapped-ion mode, *Phys. Rev. Lett.* 34 (7) (1975) 391.
- [51] Y. Kuramoto, T. Tsuzuki, Persistent propagation of concentration waves in dissipative media far from thermal equilibrium, *Prog. Theor. Phys.* 55 (2) (1976) 356–369.
- [52] P. Stinis, Stochastic optimal prediction for the Kuramoto–Sivashinsky equation, *Multiscale Model. Simul.* 2 (4) (2004) 580–612.
- [53] A.-K. Kassam, L.N. Trefethen, Fourth-order time-stepping for stiff pdes, *SIAM J. Sci. Comput.* 26 (4) (2005) 1214–1233.

- [54] F. Lu, K.K. Lin, A.J. Chorin, Data-based stochastic model reduction for the Kuramoto–Sivashinsky equation, *Physica D* 340 (2017) 46–57.
- [55] S. Toh, Statistical model with localized structures describing the spatio-temporal chaos of Kuramoto–Sivashinsky equation, *J. Phys. Soc. Jpn.* 56 (3) (1987) 949–962.
- [56] J. Harlim, S.W. Jiang, S. Liang, H. Yang, Machine learning for prediction with missing dynamics, *J. Comput. Phys.* 428 (2021) 109922, <https://doi.org/10.1016/j.jcp.2020.109922>.
- [57] F.G. Mehler, Ueber die entwicklung einer function von beliebig vielen variablen nach laplaceschen functionen höherer ordnung, *J. Reine Angew. Math.* 66 (1866) 161–176.
- [58] D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University Press, 2010.
- [59] M. Schmuck, M. Pradas, G.A. Pavliotis, S. Kalliadasis, A new mode reduction strategy for the generalized Kuramoto–Sivashinsky equation, *IMA J. Appl. Math.* 80 (2) (2013) 273–301.
- [60] J.C. Mattingly, A.M. Stuart, D.J. Higham, Ergodicity for sdes and approximations: locally Lipschitz vector fields and degenerate noise, *Stoch. Process. Appl.* 101 (2) (2002) 185–232.
- [61] X. Wu, X. Li, On consistent definitions of momentum and energy fluxes for molecular dynamics models with multi-body interatomic potentials, *Model. Simul. Mater. Sci. Eng.* 23 (1) (2015) 015003.
- [62] B. Leimkuhler, C. Matthews, *Molecular Dynamics*, Springer, 2016.