CrossMark

Random dynamical systems: addressing uncertainty, nonlinearity and predictability

Navaratnam Sri Namachchivaya (ORCID)

Abstract Nonlinearity and noise play a significant role in an enormous range of subjects across the entire spectrum of science and engineering. This paper considers several research topics that encompass the area of *random dynamical systems* (RDS). A general overview of the problems, the multidisciplinary methods required for their analysis, and relevant results achieved in RDS are given with particular emphasis on developments during the past 25 years. The first part of this paper focuses on developing methods to unravel complex interactions between noise and nonlinearities using a mix of multidisciplinary approaches from theory, modeling, and simulation. Practical applications of these research results are beginning to appear across the entire spectrum of mechanics; for example, vibration absorbers, panel flutter, variable speed machining processes, and mixing and transport phenomena in fluid mechanics. The second part of this paper focuses on developing new algorithms and tools for the collection, assimilation and harnessing of data by threading together ideas ranging from random dynamical systems to information theory. A new particle filtering algorithm that combines stochastic homogenization with filtering theory is presented. Importance sampling and control methods are then used as a basic and flexible tool for the construction of the proposal density inherent in particle filtering.

Keywords Stochastic differential equation · Lyapunov exponents · Invariant measures · Stochastic stability · Homogenization · Stochastic bifurcation · Nonlinear filtering · Particle filters · Chaos

1 Introduction

Mathematical models are valuable tools for understanding complex systems. However, these models are considerably uncertain. The uncertainty can be traced back to: the inability of models to capture important physical processes; inadequate use of observations to constrain and initialize the models; and uncertainties in forcing. An emerging challenge for inference and prediction of large-scale complex systems is to efficiently *analyze and assimilate* the ever-increasing high dimensional data produced by the vast number of engineered and natural systems. Tremendous strides have been made in recent years in uncertainty quantification, but the enormous complexity of the problems continues to pose challenges for predicting interactions among the physical and environmental

N. S. Namachchivaya (✉)
Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, 306 Talbot Laboratory, 104 S. Wright St, Urbana, IL 61801, USA
e-mail: navam@illinois.edu

systems. These unique challenges, due to the interactions among uncertainties, nonlinearities, and observations, will be the focus of this work.

This paper integrates the ideas from *random dynamical systems*, *homogenization methods*, *nonlinear filtering* and *Markov Chain Monte Carlo* methods to develop a general collection of new mathematical techniques that describe the behavior of complex dynamical systems, and dynamically assimilate new data from observations for prediction. As we shall show, the concept of combining random dynamical systems, homogenization methods and nonlinear filtering always sounds clear enough at a distance, but the nearer you get to the proofs the vaguer the subtleties of these interactions become. Key tools to explore these issues in a canonical way are explained in [2, 4, 19, 20, 34]. Novel results, contributions and their significance in *three* important areas of research within RDS are presented below.

The material in this paper is organized as follows. In Sect. 2, we consider the topic of random dynamical systems and briefly describe the basic concepts from invariant measures to Lyapunov exponents, particularly, the maximal Lyapunov exponent, which quantifies the rate of exponential growth or decay of the solution to a linear stochastic system. Section 3 summarizes the results on the stability in the almost-sure sense, which is determined by the sign of the maximal Lyapunov exponent. Since an invariant measure is the stochastic analogue of a fixed point, we present results on bifurcations of invariant measures; the main tool for this analysis is the approximation by amplitude equations. The reduction technique, examined in Sect. 4, entails an averaging result for the fast motion associated with the unperturbed dynamics. Then we state the well known results associated with the martingale problem. Section 5 addresses the effects of the multiscale signal and observation processes via the study of the Zakai equation [49]. We describe a lower dimensional stochastic PDE (Zakai type equation) that was constructed in a canonical way to addresses the effects of the multiscale signal and observation processes. Finally, we incorporate an optimal particle filtering algorithm that generates the best importance sampling density. This particle method consists of control terms in the "prognostic" equations that nudge the particles toward the observations. We apply the optimal

(optimal nudging) particle filtering algorithm to a chaotic system.

2 Random dynamical systems

Random dynamical systems are dynamical systems "corrupted" by random perturbations and are tailor-made to cover a vast number of engineered and natural systems, in particular those modeled as random and stochastic differential equations. To start with, define a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is a σ-algebra of measurable subsets of Ω called "events" and $\mathbb{P}$ is the probability measure. Here we restrict ourselves to the smooth (i.e. C^∞) case, two-sided continuous time $\mathbf{T} = \mathbb{R}$, and state space $\mathbb{R}^d$. A smooth random dynamical system consists consist of two "ingredients" (see Arnold [4]):

1. Model of the noise: By parametrizing the probability space by time, we are able to connect the state $\omega \in \Omega$ of the random environment at time $t = 0$, to it's state $\theta_t \omega \in \Omega$ after a passage of time t. More precisely, a *metric dynamical system* denoted as $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t \in \mathbb{R}})$ (for short: θ), i.e. a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a measurable flow of measure preserving transformations $\theta_t : \Omega \to \Omega$, i.e. $\theta_0 = \mathrm{id}$, $\theta_{t+s} = \theta_t \circ \theta_s$ for all $t, s \in \mathbb{R}$, $\theta_t \mathbb{P} = \mathbb{P}$, and $(t, \omega) \mapsto \theta_t \omega$ measurable.

2. Model of the system perturbed by noise: A system excited by an external stochastic process for a single realization ω can be interpreted via the path $\theta_t \omega$ in Ω. More precisely, a *cocycle* φ over θ of smooth mappings of $\mathbb{R}^d$, i.e. a measurable mapping

$$\varphi : \mathbb{R} \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d, \quad (t, \omega, x) \mapsto \varphi(t, \omega)x,$$

for which $(t, x) \mapsto \varphi(t, \omega)x$ is continuous in (t, x) and smooth in x, and φ satisfies the *cocycle property*

$$\varphi(0, \omega) = \mathrm{id}_{\mathbb{R}^d}, \quad \omega \in \Omega,$$
$$\varphi(t + s, \omega) = \varphi(t, \theta_s \omega) \circ \varphi(s, \omega) \quad \forall s, t \in \mathbb{R}.$$

The cocycle property implies that $\varphi(t, \omega)^{-1} = \varphi(-t, \theta_t \omega)$, i.e. the mapping $\varphi(t, \omega) : \mathbb{R}^d \to \mathbb{R}^d$ is a (smooth) diffeomorphism. This framework allows one to show that, for almost all realizations

$\omega \in \Omega$, the evolution in the state space $\mathbb{R}^d$ of a stochastic system from time $s < t$ to time t is described by a two-parameter family of transformations, starting at time $s = 0$.

The flow Θ_t on $\Omega \times \mathbb{R}^d$ given by $\Theta_t(\omega, x) := (\theta_t \omega, \varphi(t, \omega)x)$ is called the *skew product flow* corresponding to φ.

A *real noise* here means any stationary stochastic process $(\xi_t)_{t \in \mathbf{R}}$ that can be canonically modeled as a metric dynamical system with an appropriate state space. This paper examines a random dynamical system defined by

$$\dot{X}_t = f(X_t; \alpha) + g(X_t; \alpha)\xi_t, \quad X_0 \stackrel{\text{def}}{=} x \in \mathbb{R}^d, \qquad (2.1)$$

where $\xi_t = \xi(\omega, t)$ represents stationary stochastic processes (e.g. white noise, colored noise, etc.). In the case of additive noise, $g(x; \alpha)$ is a constant. The stability and nonlinear response of such stochastic systems have become research problems of increasing interest.

Dynamical systems driven by white noise are rigorously dealt with in stochastic analysis and are solutions of (Stratonovich) *stochastic differential equations*

$$dX_t = f(X_t; \alpha) \, dt + g(X_t; \alpha) \circ dW_t, \quad X_0 \stackrel{\text{def}}{=} x \in \mathbb{R}^d, \qquad (2.2)$$

where f and g are smooth vector fields in $\mathbb{R}^d$. Let us now consider (2.2), in the context of *random dynamical systems*. *White noise* (Wiener process) can be canonically modeled as a metric dynamical system as follows: Let $\Omega = \{\omega \in C(\mathbb{R}, \mathbb{R}^d) : \omega(0) = 0\}$, $\mathcal{F}$ the Borel σ-algebra of Ω, and $\mathbb{P}$ the Wiener measure, i.e. the measure generated by the Wiener process (Brownian motion) $(W_t)_{t \in \mathbb{R}}$ in $\mathbb{R}^m$. This process has stationary independent increments with $W_{t+h} - W_t \sim \mathcal{N}(0, |h|I)$, continuous trajectories, and satisfies $W_0 = 0$. The shift $\theta_t \omega(\cdot) := \omega(t + \cdot) - \omega(t)$ leaves $\mathbb{P}$ invariant since the increments are stationary. Then θ is an ergodic metric dynamical system on $(\Omega, \mathcal{F}, \mathbb{P})$ "driving" the stochastic differential equation (2.2) and $W_t = \omega(t)$.

Theorem 2.1 (Arnold and Scheutzow [8]) *Let $f, g \in C_b^\infty$. Then the stochastic differential equation (2.2) has a unique solution $x \mapsto \varphi(t, \omega)x$ which is a smooth random dynamical system. The Jacobian*

$D\varphi(t, \omega, x)$ *is a matrix cocycle over Θ and uniquely solves the* variational equation

$$dv = Df(\varphi(t, \cdot)x) \, v \, dt + Dg(\varphi(t, \cdot)x) \, v \circ dW_t. \quad (2.3)$$

In Eq. (2.3), the dependence of the vector fields f and g in the parameter α is suppressed.

2.1 Invariant measures

For all further steps we need the notion of an invariant measure for a random dynamical system. The invariant measures μ are defined on the product space $\Omega \times \mathbb{R}^d$ and the invariance is with respect to the skew product flow Θ so that $\Theta_t \mu = \mu$. The projection of μ on Ω is $\mathbb{P}$ while the invariance of μ in $\mathbb{R}^d$ corresponds to the use of random probability measures on $\mathbb{R}^d$. Let φ be a random dynamical system over θ. A random probability measure $\omega \mapsto \mu_\omega$ on $(\mathbb{R}^d, \mathbb{B}^d)$, where $\mathbb{B}^d$ represent the space of Borel sets in $\mathbb{R}^d$, is called *invariant* under φ, if

$$\varphi(t, \omega)\mu_\omega = \mu_{\theta_t \omega} \quad \mathbb{P} - a.s. \quad \text{for all} \quad t \in \mathbb{R}.$$

For random dynamical systems whose one-point motions $\mathbb{R}^+ \ni t \mapsto \varphi(t, \omega)x$ are Markov processes with transition probability $P(t, x, B) = \mathbb{P}\{\omega : \varphi(t, \omega)x \in B\}$ and generator G (for solutions of stochastic differential equation (2.2)), a measure ρ on $\mathbb{R}^d$ is called *stationary* if it satisfies for all $t \in \mathbb{R}^+$

$$\rho(\cdot) = \int_{\mathbb{R}^d} P(t, x, \cdot) \, \rho(dx),$$

equivalently, if it solves the *Fokker–Planck equation*

$$G^* \rho = 0, \quad G = f + \frac{1}{2}g^2. \qquad (2.4)$$

Here we have written G in the Hörmander form. There is a one-to-one correspondence between stationary ρ's and those invariant μ_ω's for φ which are measurable w.r.t. the past $\mathcal{F}_{-\infty}^0$ of the noise, via the "pullback"

$$\rho \mapsto \mu_\omega = \lim_{t \to \infty} \varphi(-t, \omega)^{-1} \rho, \quad \mu_\omega \mapsto \mathbb{E}\mu_. = \rho, \qquad (2.5)$$

(see Arnold [4], Sect. 1.7). The procedure of passing from a deterministic stationary measure ρ to a random invariant measure μ_ω described by (2.5) is called *disintegration* of ρ. However, there are in general,

more invariant measures μ_ω than those obtained from stationary measures.

2.2 Lyapunov exponents

The multiplicative ergodic theorem (MET) of Oseledec [37], which established the existence of finitely many deterministic exponential growth rates called *Lyapunov exponents*, has had a powerful influence upon the study of stochastic stability. Lyapunov exponents and Oseledec spaces provide us with the stochastic analogues of the real parts of the eigenvalues and eigenspaces of a deterministic constant matrix. The almost-sure stability of a solution to an RDS, is determined by the sign of the maximal Lyapunov exponent.

Let φ be a smooth random dynamical system, and let μ be an ergodic invariant measure. From Theorem 2.1, $D\varphi$ is a linear cocycle over Θ and uniquely solves the linear variational equation (2.3). Denote by

$$\lambda(\omega, x, v) := \lim_{t \to \infty} \frac{1}{t} \log \|D\varphi(t, \omega, x)v\|,$$

the *Lyapunov exponent* or the exponential growth rate of the solution $v_t(x, v)$, for the the initial condition v ($v \neq 0$) in (2.3). According to MET [37], λ takes on one of r fixed, non-random values $\lambda_1 > \cdots > \lambda_r$. Which λ_i is realized depends on the initial condition v. The multiplicities of the Lyapunov exponents sum to the dimension of the system, d. The maximum of these, λ_1, determines the almost-sure stability of the random dynamical system $\varphi(t, \omega)$ generated by (2.2) under the stationary measure ρ.

Rewriting the variational equation (2.3) in polar coordinates

$$s = \frac{v}{\|v\|} \in \mathbf{S}^{d-1}, \quad r = \|v\| \in (0, \infty)$$

yields

$$\begin{aligned} dr_t &= q_0(x_t, s_t) r_t dt + q_1(x_t, s_t) r_t \circ dW_t, \\ ds_t &= h_0(x_t, s_t) dt + h_1(x_t, s_t) \circ dW_t, \end{aligned} \quad (2.6)$$

where

$$h_0(x, s) \overset{\text{def}}{=} Df(x)s - q_0(x, s)s, \quad q_0(x, s) \overset{\text{def}}{=} \langle Df(x)s, s \rangle$$
$$h_1(x, s) \overset{\text{def}}{=} Dg(x)s - q_1(x, s)s, \quad q_1(x, s) \overset{\text{def}}{=} \langle Dg(x)s, s \rangle,$$
$$(2.7)$$

and $\langle x, y \rangle$ is the standard scalar product in $\mathbb{R}^{d-1}$. In (2.6), the equation for s_t is decoupled from the one

for r_t, so that the pair (x_t, s_t) forms a Markov process with state space $\mathbb{R}^d \times \mathbf{S}^{d-1}$, whose generator for the *additive noise* case simplifies to $\mathcal{L} \overset{\text{def}}{=} G + h_0(x, s)\frac{\partial}{\partial s}$. Integrating the equation for the radial process r_t in (2.6) and using the classical ergodic theorem yields the *Furstenberg–Khasminskii formula* ([4], Chap. 6) for the top Lyapunov exponent

$$\lambda = \int_{\mathbb{R}^d \times \mathbf{S}^{d-1}} Q(x, s) \, v(dx, ds), \quad (2.8)$$

where Q is some explicitly known function, which for the *additive noise* case simplifies to $Q(x, s) = q_0(x, s)$ and v is the (to be determined) joint stationary measure for the Markov process (x_t, s_t) on $\mathbb{R}^d \times \mathbf{S}^{d-1}$ with marginal ρ on $\mathbb{R}^d$. The sign of λ in (2.8) is of particular interest as it determines the stability of the variational equation (2.3) and in turn the stability of the original smooth nonlinear random dynamical system generated by (2.2). Formula (2.8) forms the basis of all asymptotic studies of Lyapunov exponents and particularly the presentation given in this paper.

2.3 Moment Lyapunov exponents

Although sample solutions may be stable with probability one, the mean square response of the system for the same parameter values may grow exponentially. It is well known that there are parameter values at which the top Lyapunov exponent λ is negative, indicating that the system is sample stable, while the pth moments grow exponentially for large p indicating the pth moment response is unstable. This implies that, although the system response $\|D\varphi(t, \omega, x)v\| \to 0$ as $t \to \infty$ with probability one at an exponential rate λ, there is a small probability that $\|D\varphi(t, \omega, x)v\|$ is large. This makes the expected value of this rare event large for large values of p and results in pth moment instability.

In a manner analogous to the above discussion, the concept of moment Lyapunov exponents was introduced for linear random dynamical systems by Arnold [4]. One can define the exponential growth rate of the pth moment of the solution i.e., $\mathbb{E}\|X_t(x, \omega)\|^p$, by the *moment Lyapunov exponent*

$$g(p; x) \overset{\text{def}}{=} \lim_{t \to \infty} \frac{1}{t} \log \mathbb{E}\|X_t(x, \omega)\|^p.$$

If $g(p; x) < 0$, then, by definition, $\mathbb{E}\|X_t(x, \omega)\|^p \to 0$ as $t \to \infty$ and this is referred to as pth moment stability.

The exponential growth rate of the pth moment has proven to be the key to all finer stability properties of a random dynamical system, in particular, $g(p)$ is a convex analytic function of $p \in \mathbb{R}$, with $g(0) = 0, \frac{g(p)}{p}$ increasing, and

$$g'(0) = \lim_{p \to 0} \frac{g(p)}{p} = \lambda.$$

The nonzero unique root γ (*the stability index*) (see Arnold and Khasminskii [7]) of the equation $g(p) = 0$ is connected with the asymptotic behavior. That is, the stability index controls the probability with which an almost surely stable system exceeds a threshold. If the top Lyapunov exponent $\lambda = g'(0) < 0$ and $g(p) = 0$ has a positive root $\gamma > 0$, then there exist a $K \geq 1$ such that for $\delta > 0$ and for all x with $||x|| < \delta$

$$\frac{1}{K} \left(\frac{||x||}{\delta} \right)^{\gamma} \leq \mathscr{P} \left\{ \sup_{t \geq 0} ||X_t(x, \omega)|| > \delta \right\} \leq K \left(\frac{||x||}{\delta} \right)^{\gamma}.$$

On the other hand, if the $\lambda = g'(0) > 0$ and $g(p) = 0$ has a negative root $\gamma < 0$, then there exist a $K \geq 1$ such that for $\delta > 0$ and for all $|x| > \delta$

$$\frac{1}{K} \left(\frac{||x||}{\delta} \right)^{\gamma} \leq \mathscr{P} \left\{ \inf_{t \geq 0} ||X_t(x, \omega)|| < \delta \right\} \leq K \left(\frac{||x||}{\delta} \right)^{\gamma}.$$

Over the past three decades the work on RDS raised hard and far reaching questions on stochastic stability of invariant measures of dynamical systems, stochastic bifurcations, stochastic flows and random attractors. Answers to these questions have made invaluable contributions to the modern theory of Lyapunov exponents based on the multiplicative ergodic theory of stochastic flows as well as developed a stochastic version of the center manifold theory and stochastic normal form theory [6, 31] for random dynamical systems. These are considered to be landmarks in the evolution of the field as amply explained in [4]. A primary goal of this section is to unravel core problems in the areas of almost sure stability [32, 46], stabilization by noise [35] and stochastic bifurcations [2, 29] in ways that are far more transparent. The practical results obtained have a wide range of applications and some of these results are even contrary to intuition, such as dissipation induced instability and stabilization by noise.

3 Stochastic stability and bifurcation

The scope of this section is to emphasize the challenges associated with the stochastic stability and stochastic bifurcation problems. As it turns out, this problem, apart from its obvious practical importance, requires answering several important asymptotic-theoretic questions as well. We first present a brief survey of the main stochastic stability and stabilization results for the case of multi-dimensional systems with critical and stable modes.

3.1 Stochastic stability and stabilization by noise

Let $\bar{x} \overset{\text{def}}{=} \tilde{x}(t, \omega; \alpha)$ be a stationary solution of Eq. (2.1). We have shown explicit dependence of ω since the solution is a random process. Then, one asks the following questions: • How do we find $\tilde{x}(t, \omega; \alpha)$? • Is $\tilde{x}(t, \omega; \alpha)$ stable in some sense? Both almost-sure stability and pth moment stability have been widely used in the study of stability of solutions of random dynamical systems. If a stationary solution $\tilde{x}(t, \omega; \alpha)$, or simply a fixed point $\bar{x}$ of dynamical system (2.1) is stable either almost-surely or in the pth moment, it is stable in distribution and in probability, as well. The challenge was to extend the stochastic techniques for the analysis of noisy nonlinear systems described in (2.1) in order to answer these questions. It is this need and challenge that was addressed by several researches over the past four decades.

The almost-sure stability is determined by the sign of the maximal Lyapunov exponent. For a linearized system perturbed by real noise,

$$\dot{X}_t = A(\xi(t))X_t, \quad X_0 \overset{\text{def}}{=} x \in \mathbb{R}^d$$

$$d\xi = \chi_0(\xi) + \sum_{i=1}^{r} \chi_i(\xi) \circ dW_i, \quad \xi \in M \tag{3.1}$$

in order to ensure that there is a unique smooth and positive invariant density v on the compact manifold M, we assume $\xi(t)$ is strongly elliptic in the sense that $\dim LA(X_1, \ldots, X_r)(\xi) = \dim M$ for all $\xi \in M$ where $LA(Z)$ denotes the Lie algebra generated by the set Z of vector fields. As before, reverting to polar coordinates in $\mathbb{R}^d$ through the Khas'minskii transformation we obtain

$$\| X_t(x) \| = \| x \| \exp \left\{ \int_0^t q(\tau) d\tau \right\}$$

$$\dot{s} = h(\xi(t), s) \quad \text{with} \quad h(\xi(t), s)$$
$$= (A(\xi(t)) - q(t)I)s,$$

where $q(t) \overset{\text{def}}{=} q(\xi(t), s) = s^T A(\xi(t))s$, and (ξ, s) together form a diffusion process on $M \times \mathbf{P}^{d-1}$ (obtained from S^{d-1} by identifying s and $-s$). The generator of this process is given by

$$\mathcal{L} = G + h \frac{\partial}{\partial s},$$

where $G = X_0 + \frac{1}{2} \sum_{i=1}^r X_i^2$ is the generator of ξ written in Hörmander form. For a fixed $\xi \in M, h(\xi, \cdot)$ is a vector field on the projective space. To avoid degenerate situations for $\mathcal{L}$, we impose the following ellipticity condition

$$(HR) \quad \dim LA \left(X_0 + h + \frac{\partial}{\partial t}, X_1, \ldots, X_r \right) (\xi, s, t)$$
$$= \dim M + d, \quad \forall (\xi, s, t) \in M \times \mathbf{P}^{d-1} \times \mathbb{R}.$$

Thus, combining the above results the following was proven by Arnold et al. [1] Assume (HR)

1. Let $\lambda = \int_M \int_{\mathbf{P}^{d-1}} q(\xi, s) d\mu$ where μ is the unique invariant probability measure of (ξ, s) on $M \times \mathbf{P}^{d-1}$. Then λ is the maximal Lyapunov exponent for (3.1), i.e. for $x_0 \neq 0$

$$\lim_{t \to \infty} \frac{1}{t} \log \| X_t(x) \| = \lambda \quad \text{almost-surely.}$$

2. For $p \in \mathbb{R}$, let $g(p)$ be the principal eigenvalue of $L(p) = \mathcal{L} + pq(\xi, s)$ acting on $C(M \times \mathbf{P}^{d-1})$. Then $g(p)$ is the pth moment Lyapunov exponent for (3.1), i.e. for $x \neq 0$

$$\lim_{t \to \infty} \frac{1}{t} \log \mathbb{E} \| X_t(x) \|^p = g(p).$$

Although the top Lyapunov exponent, λ, moment Lyapunov exponent, $g(p)$, and the stability index, γ, are very important characteristics for the analysis of linear RDS, in general it is impossible to find explicit expressions for them, except for some special linear oscillators under parametric white noise excitations: an exact formula for λ is given by Imkeller and Lederer [21].

The following results discuss the challenges of evaluating the top and the moment Lyapunov exponents of the stationary measures of noisy dynamical systems asymptotically, when the noise is weak. The results on stochastic stability of higher dimensional problems with one critical mode, broke new ground by providing a possible mechanism for stabilization by noise [35]. Naturally, problems of this type frequently appear in many multi-degree-of-freedom systems exhibiting Hopf bifurcation and have received considerable attention. A four-dimensional version of this is

$$\dot{X}_t = AX_t + \sqrt{\varepsilon v}\, \xi(t) BX_t, \quad X_0 = x \in \mathbb{R}^4, \qquad (3.2)$$

where v represents the intensity of the noise ξ_t,

$$A = \begin{bmatrix} \varepsilon\delta_1 & \omega_1 & 0 & 0 \\ -\omega_1 & \varepsilon\delta_1 & 0 & 0 \\ 0 & 0 & -\delta_2 & \omega_2 \\ 0 & 0 & -\omega_2 & -\delta_2 \end{bmatrix},$$

$$B = \begin{bmatrix} K_{11} & K_{12} & M_{11} & M_{12} \\ K_{21} & K_{22} & M_{21} & M_{22} \\ N_{11} & N_{12} & L_{11} & L_{12} \\ N_{21} & N_{22} & L_{21} & L_{22} \end{bmatrix},$$

and the quantity δ_2 corresponds to the real part of the eigenvalue of the stable mode, δ_1 is the unfolding of the critical eigenvalue which may represent the rate at which the real part crosses the imaginary axis of the complex plane or a *small* detuning damping parameter. The B matrix is conveniently represented in terms of four sub-matrices. In the case of white noise the maximal Lyapunov exponent is

$$\lambda = \varepsilon\,\delta_1 + \varepsilon\, \frac{v}{8} \Big[(K_{12} + K_{21})^2 + (K_{11} - K_{22})^2$$
$$+ 2(M_{11}N_{11} + M_{22}N_{22} + M_{12}N_{21} + M_{21}N_{12}) \Big].$$
$$(3.3)$$

This formula contains terms of the type, $M_{ij}N_{ji}$, which represent the contribution from the stochastic components in the stable "heavily damped" modes to that of the "critical" modes. Unlike the two dimensional linear system, stabilization is possible under stochastic excitations if at least one of the pairs $(M_{11}, N_{11}), (M_{22}, N_{22}), (M_{12}, N_{21}), (M_{21}, N_{12})$ contains elements of opposite signs and the maximal Lyapunov exponent is negative. These results on

stabilization by noise showed how to model and design engineering systems that take advantage of unavoidable turbulence or noise. These results were then used to explain on a rigorous theoretical basis the stabilization by grid generated turbulence of a smooth circular cylinder reported in Popp and Romberg [42] by modeling the immersed cylinder as a two-degree of freedom oscillator and the turbulence as a stochastic process.

An asymptotic expansion for the maximal Lyapunov exponent and the rotation numbers for a general four dimensional dynamical system (3.2) with two critical modes, (i.e, $\delta_2 \to \varepsilon\delta_2$) driven by a small intensity real noise process were constructed in [16, 32]. Results are presented only for the case of the coupled oscillators [32] described by equations of motion of the form

$$\ddot{q}_i + \omega_i^2 q_i + 2\varepsilon\zeta\omega_i\dot{q}_i + \sqrt{\varepsilon}\sum_{j=1}^{2} k_{ij}q_j f(\xi(t)) = 0, \ i = 1, 2$$

where the q_i's are generalized coordinates, ω_i is the ith natural frequency, and $\varepsilon\zeta$ represents a small viscous damping coefficient. It is assumed that the natural frequencies are *noncommensurable*. The stochastic term $\xi(t)$ is a small-intensity, real-noise process on a smooth connected Riemannian manifold M (with or without boundary), as in (3.1). In order to make the problem tractable, the associated infinitesimal generator G will be assumed to have an *isolated simple zero eigenvalue*. Hence, the only solution of $Gu = 0$ is $u \equiv$ constant. It follows that the associated adjoint operator G^* also has zero as a simple, isolated eigenvalue and the normalized invariant measure satisfies $G^*v(\xi) = 0$.

Let's define $p_{ij} \stackrel{\text{def}}{=} \frac{k_{ij}}{\omega_i}$ and define λ_i, $i = 1, 2, a$ and b in terms of cosine spectra at $2\omega_i$ and $\Omega^\pm = \omega_1 \pm \omega_2$ as,

$$\lambda_i = -\delta_i + \frac{1}{8}p_{ii}^2 S(2\omega_i), \quad i = 1, 2$$

$$a = \frac{1}{16}\sum_{i=1}^{2} p_{ii}^2 S(2\omega_i) - \frac{1}{4}p_{21}p_{12}S(\Omega^+) \quad (3.4)$$

$$b = \frac{1}{16}\sum_{i=1}^{2} p_{ii}^2 S(2\omega_i) + \frac{1}{4}p_{21}p_{12}S(\Omega^-),$$

with $S(\omega) = 2\int_0^\infty \mathcal{R}(\tau)\cos\omega\tau\,d\tau$, where $\mathcal{R}(\tau)$ is the correlation function. Then the maximal Lyapunov

exponent in terms of coefficients of the stochastic terms is

$$\lambda^\varepsilon = \frac{\varepsilon}{2}\left\{(\lambda_1 + \lambda_2) + (\lambda_1 - \lambda_2)\coth\left[\frac{(\lambda_1 - \lambda_2)}{2}\beta(\frac{\pi}{2})\right]\right.$$
$$\left.\frac{1}{2}p_{12}p_{21}[S(\Omega^+) - S(\Omega^-)]\right\},$$
$$(3.5)$$

where

$$\beta\left(\frac{\pi}{2}\right) = \begin{cases} \dfrac{1}{\sqrt{ab}}\ln\left|\dfrac{\sqrt{a} + \sqrt{b}}{\sqrt{a} - \sqrt{b}}\right| & \text{if } ab > 0; \\[3mm] \dfrac{2}{a + b} & \text{if } ab = 0; \\[3mm] \dfrac{1}{\sqrt{-ab}}\tan^{-1}\left(\dfrac{2\sqrt{-ab}}{a + b}\right) & \text{if } ab < 0. \end{cases}$$

The extension of these almost-sure asymptotic stability results to infinite dimensional systems [25] was achieved for delay differential equations (DDEs), where the time derivative can depend on both past and present values of the variable. Application of these results to cutting-tool chatter in turning and milling processes demonstrated rigorously the potential effectiveness of spindle speed variation (fluctuating) to eliminate chatter. The model of interest is a second-order delay differential equation for the position q of a point on a machine tool which is cutting material from a shaft which is rotating with a time period of revolution r. The effects of small random perturbations in structural parameters of a delay differential oscillator is of interest, that is,

$$\ddot{q}(t) + 2\zeta\dot{q}(t) + q(t) = -\kappa[q(t) - q(t - r)]$$
$$+ \varepsilon\sigma(\xi(t))[\kappa q(t - r)]. \quad (3.6)$$

The term $\sigma(\xi(t))$ represents small random perturbations, for example, in the natural frequency, the term $\kappa[q(t) - q(t - r)]$ in (3.6) represents the assumption that the force acting on the tool is proportional to width of the chip being cut, and the width is the difference between the present position $q(t)$ of the tool and its position one revolution earlier $q(t - r)$. It is known that, for a fixed r, there exists a critical κ_c such that the amplitude q of the oscillator decreases exponentially if $\kappa < \kappa_c$ and increases exponentially if $\kappa > \kappa_c$. When

$\kappa = \kappa_c$ oscillations of constant amplitude persist with frequency ω_c. In machining, this oscillatory behavior is called chatter. Chatter—the self-excited relative vibration between workpiece and cutting tool—is a common problem in the machining process. Chatter occurs with greater rate of cut or with larger depth of cut, resulting in poor surface finish. Our results showed that stabilization is possible by the noise in (3.6). That is, by suitably selecting the parameters in the equation (3.6), it is possible to have λ negative. For such a selection of parameters, it might be possible to increase κ slightly from the critical value without undergoing chatter.

Now we present some results of asymptotic stability of nonlinear systems (forces derivable from a potential) with noise, which describe a host of physically interesting problems in random vibrations, from simple oscillators [5] to noisy autoparametric systems [30]. It is well known that such systems, under the assumptions that $\xi_i(t)$'s are uncorrelated Gaussian processes, and the ratio of the spectral density of each excitation, $\xi_i(t)$, to the corresponding damping, β_i, is the same, have stationary probability densities. However, there are no concrete results on the sign of the top Lyapunov exponents corresponding to these stationary measures. Hence, their stability is not known.

It is well-known that the two point motion of a one dimensional nonlinear stochastic system is unique. More precisely if a noisy one dimensional equation,

$$\dot{X}_t = f(X_t) + \sum_{i=1}^{2} g_i(X_t)\xi_i(t), \quad X_0 = x \in \mathbb{R}, \quad (3.7)$$

with $g^2(x) \stackrel{\text{def}}{=} (g_1^2 + g_2^2(x)) > 0$ for all x, has a stationary invariant measure with density

$$g(x) \stackrel{\text{def}}{=} \sqrt{(g_1^2 + g_2^2(x))}$$

provided $p(x)$ is normalizable, then the Lyapunov exponent is

$$\lambda = -2 \int_0^\infty \frac{\left(2f^2(x) + \frac{1}{2}g_2'^2(x)g_1^2\right)}{(g_1^2 + g_2^2(x))} p(x)\,dx. \quad (3.8)$$

The Lyapunov exponent is always negative provided $f(x) \neq 0$ [23].

The challenge was to explicitly evaluate the top Lyapunov exponent of higher ($d \geq 2$) dimensional

nonlinear systems with noise, and in particular additive white noise. The analytical study of asymptotic stability of nonlinear systems with noise, initiated in Arnold et al. [5], opened the door to a host of physically interesting problems in random vibrations, from simple oscillators to noisy auto-parametric systems. It was shown analytically in [5] that the top Lyapunov exponent of a nonlinear oscillator under additive noise is positive, for small noise and small dissipation. Their work was devoted to the effect of noise on the Duffing oscillator

$$\ddot{x}_t + \varepsilon \dot{x}_t + U'(x_t) = \sqrt{2\varepsilon}\,\xi(t), \quad (3.9)$$

where $U(x) = \frac{a}{2}x^2 + \frac{b}{4}x^4$, $a, b > 0$. It was shown that the top Lyapunov exponent is positive, i.e.,

$$\lambda(\varepsilon) > 0 \quad \text{for } \varepsilon \text{ not too large.}$$

Schimansky-Geier and Herzel [43] were the first to consider numerically the Lyapunov exponents of a two dimensional nonlinear double-well potential with $U(x) = -\frac{a}{2}x^2 + \frac{b}{4}x^4$, $a, b > 0$, which was studied by Kramers in his celebrated work (e.g., [19, 43]). It was shown that the top Lyapunov exponent is positive for small ε. The top Lyapunov exponent is determined by the simultaneous behavior of two neighboring orbits, or the two-point motion of the noisy nonlinear oscillators. The positivity of the top Lyapunov exponent is significant, because it implies that while for each initial condition the solution trajectory asymptotically approaches the volume element in the state space giving rise to a nontrivial stationary measure, the distance between any two initial conditions will grow at an exponentially fast rate. Hence, an additive noise in (3.9) induces an unstable stationary measure. A result of Baxendale and Goukasian [11] showed that the top Lyapunov exponent is positive for small ε for a multiplicative noise in (3.9). Namachchivaya et al. [30] extended these results to study the single mode solution of two degree-of-freedom noisy auto-parametric systems. The unique properties of auto-parametric systems, and Diliberto transformations were used to reduce the problem to two uncoupled two dimensional nonlinear systems. As the noise intensity was increased, these questions were examined by developing a sequence of averaging and computational procedures that were uniquely adapted to study noisy auto-parametric systems.

3.2 Stochastic bifurcation

Stochastic bifurcations are those where small changes in noisy parameters play an important role, and they result in large uncertainty about the system's dynamical behavior. Stochastic bifurcation theory studies the *qualitative changes* in parameterized families of random dynamical systems, e.g. those generated by a family of stochastic differential equations. We study such bifurcations for the reduced systems, but also investigate how bifurcation interacts with a separation of scales. How are the qualitative changes of probability densities of the reduced model connected to the phenomenological changes of the original noisy systems?

It is important to study how these invariant measures can bifurcate as the system parameters are varied. The invariant measures, in other words, the solutions of the *Fokker–Planck equation*

$$\mathscr{L}_\alpha^* p_\alpha = 0 \tag{3.10}$$

arising from a nonlinear system depending upon a parameter α may, for example, exhibit transitions from one-peak to two-peak or crater-like densities. These have been observed experimentally, numerically and analytically. The number and locations of the extrema of the stationary densities have been carefully studied.

This concept can be formalized based on the ideas of Zeeman [50], according to which two probability densities p, q are called equivalent, $p \sim q$, if there are two diffeomorphisms α, β such that $p = \alpha \circ q \circ \beta$. Then the family p_α is structurally unstable at $\alpha = \alpha_0$ since, in each neighborhood of α_0, there are non-equivalent densities. Hence, $\alpha = \alpha_0$ can rightly be called a "phenomenological", or "P-bifurcation point", and we shall call a phenomenon like this a "P-bifurcation". Even though there are many drawbacks to such a phenomenological approach, they are clearly of engineering importance since they are what one would observe experimentally or numerically in the original system.

On the other hand, a dynamical approach studies bifurcations of invariant measures which represent the stochastic analogues of fixed points. The parameter values for which the maximal Lyapunov exponent (corresponding to a known invariant measure) vanishes are known as D-bifurcation points, provided there is a nontrivial invariant measure branching from the known measure (see Baxendale [10]).

These new techniques captured the essence of the stochastic Hopf bifurcation phenomenon that contains several novel features, including phenomenological and dynamical stochastic bifurcation scenarios. These theoretical results were applied to several practical engineering systems, including the dynamics of a flat panel in a supersonic flow with boundary-layer turbulence, vibro-impact systems, and noisy auto-parametric systems. It is convenient to transfer these problems into a first order vector form:

$$\frac{dZ_t}{dt} = A(\mu)Z_t + F(Z_t; \mu) + vG(\mu)Z_t\,\xi_t,$$
$$Z_0 = z \in \mathbb{R}^n, \quad \mu \in \mathbb{R}, \tag{3.11}$$

where A and G are $n \times n$ matrices depending on a system parameter μ, v represents the intensity of the noise ξ_t, which we shall assume to be white, and F satisfies: $F(0; \mu) \equiv D_z F(0; \mu) \equiv 0$. Equation (3.11) can be written as the following Itô equation:

$$dZ_t = \left[A(\mu)Z_t + F(Z_t, \mu) + \frac{v^2}{2}G^2(\mu)Z_t \right] dt$$
$$+ vG(\mu)Z_t\,dW_t. \tag{3.12}$$

We assume that at the critical system parameter $\mu = \mu_c$, the matrix A has the following block diagonal form:

$$A(\mu_c) = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}, \quad \text{where} \quad \mathrm{B} = \begin{bmatrix} 0 & -\omega_0 \\ \omega_0 & 0 \end{bmatrix},$$
$$\omega_0 \in \mathbb{R}^+$$

and C is an $(n-2) \times (n-2)$ matrix whose eigenvalues, λ_i, $i = 1, 2, \ldots, n-2$, have negative real parts. Due to the block diagonal form of $A(\mu_c)$, it is convenient to write the matrices $A'(\mu_c)$ and $G(\mu_c)$ as

$$A'(\mu_c) = \begin{bmatrix} D & E \\ H & J \end{bmatrix}, \quad G(\mu_c) = \begin{bmatrix} K & M \\ N & L \end{bmatrix},$$

with B, D, K being 2×2 matrices; E, M being $2 \times (n-2)$ matrices; H, N being $(n-2) \times 2$ matrices and C, J, L being $(n-2) \times (n-2)$ matrices.

Putting $Z_t = (X_t, Y_t)$, $z = (x, y) \in \mathbb{R}^2 \times \mathbb{R}^{n-2}$, and $F(z) = (f^3(x,y), g^3(x,y)) \in \mathbb{R}^2 \times \mathbb{R}^{n-2}$, where $f^{(3)}$ and $g^{(3)}$ are homogeneous polynomials of degree 3. Results for more general homogeneous polynomials of

degree j, are given in [33]. The essential dynamic behavior of the multi-dimensional system (3.11) is determined by the evolution of the "critical" modes. When the rest of the modes are "heavily damped", trajectories are rapidly attracted to some low-dimensional invariant manifold, which may be parameterized by the amplitudes of the critical modes. Hence defining the amplitude and phase of the critical mode as

$$r = \mathcal{R}(z) \stackrel{\text{def}}{=} \|x\|_{\mathbb{R}^2}, \quad \theta = \Theta(z) \stackrel{\text{def}}{=} \arctan\left(\frac{x_2}{x_1}\right),$$
(3.13)

the goal is to study the behavior of $\mathcal{R}(Z^\varepsilon)$ in the small vicinity of μ_c and for small values of noise. It was shown in [33], that in the small vicinity of μ_c for small values of noise, the law of $\{\mathcal{R}(Z^\varepsilon_t); t \geq 0\}$ converges to the law of $\{\check{r}_t; t \geq 0\}$, where $\check{r}$ is the solution of the SDE

$$d\check{r}_t = b_{\mathcal{R}}(\check{r}_t)\, dt + \sigma_{\mathcal{R}}(\check{r}_t)\, dW_t, \quad \text{with} \quad \check{r}_0 = \mathcal{R}(z).$$
(3.14)

As shown in [33], a straightforward evaluation of terms in (3.14) reveals that the averaged diffusion is given by

$$\sigma^2_{\mathcal{R}}(r) = \frac{v^2}{8}\Big[2(K_{11} + K_{22})^2 + (K_{11} - K_{22})^2$$
$$+ (K_{12} + K_{21})^2\Big] r^2,$$
(3.15)

which contains only the stochastic effects from the "critical" modes. The averaged drift coefficient in (3.14) is given by

$$b_{\mathcal{R}}(r) = \frac{v^2}{8}\bigg\{ (K_{11} + K_{22})^2 + \frac{3}{2}\Big[(K_{11} - K_{22})^2$$
$$+ (K_{12} + K_{21})^2\Big]\bigg\} r + \frac{v^2}{4}\big[M_{IJ}N_{JI}\big]\, r$$
$$+ \frac{\alpha}{2}(D_{11} + D_{22})r + \frac{3}{8}\Big\{\widehat{f}^{(3)}_{1111} + \widehat{f}^{(3)}_{1122}$$
$$+ \widehat{f}^{(3)}_{2112} + \widehat{f}^{(3)}_{2222}\Big\} r^3, \quad \alpha \stackrel{\text{def}}{=} \mu - \mu_c.$$

The averaged drift coefficient contains two distinct components: (1) terms of the type, $M_{ij}N_{ji}$, which represent the contribution from the stochastic components in the stable "heavily damped" modes to that of

the "critical" modes, as well as (2) the nonlinear terms. In [33], the difficult task of including the quadratic nonlinearities as well as the stochastic parametric terms was rigorously implemented in the stochastic averaging context—derived using the martingale approach in a consistent formulation that proves limit theorems for stochastic averaging for systems with rapidly oscillating and decaying components. Now that we have identified the appropriate asymptotic law of $\mathcal{R}(Z_t)$ (see for example [44] for the real noise case), we can use this information to study its qualitative properties for computing standard statistical measures of stability, exit time laws, and stationary solutions. The work [27] extended the stochastic averaging results for delay differential equations using the powerful martingale methods of Papanicolaou et al. [38]. When the trivial stable fixed point 0 becomes unstable, and a new (stable) random invariant measure appears, it is important to know whether or not this new measure is stable. To this end, in [12] the stochastic averaging was carried out for the n-point motions of the stochastic differential equations (stochastic averaging for stochastic flows), as opposed to the standard stochastic averaging which is most commonly done just for the one-point motion. In particular the law of the linearized process of the non-trivial measure and it's top Lyapunov exponent along trajectories (top Lyapunov exponent of invariant measures under stochastic averaging) was uniquely determined. The work [32] presented the results for one specific second order equation (known as the *noisy Duffing–van der Pol oscillator*), which were generalized in [44].

4 Dimensional reduction and homogenization

The analytical core of this part of the work is *dimensional reduction*. In large complex systems, non-linearities of the governing physical processes allow energy transfer between different scales, and many aspects of this complex behavior can be represented by stochastic models. In such problems with scale separation, one of most studied models of random perturbations is represented by a diffusive Markov process X^ε_t whose semigroup of transition operators T^ε is generated by $\mathcal{L}^\varepsilon$, which is a second

order elliptic (partial) differential operator as explained in [19]. High dimensional multi-scale stochastic systems often behave like a smaller, reduced-order model; however, the reduced-order model is not known a priori. In these problems, extracting coarse-grained dynamics is at heart a problem of weak convergence of stochastic processes, or more exactly weak convergence of the laws of Markov processes. One of the preeminent modern frameworks for considering convergence of the laws of Markov processes is that of the *martingale problem* [17, 38, 45], which was used in [29, 34, 44] to develop the reduced models.

Efficient utilization of the low-dimensional models is a necessary component of simulations in large-scale settings. The goal of stochastic dimensional (or model) reduction is to extract the essential dynamics of large, complex, and noisy systems as accurately as possible; this is done by reducing the number of state variables under consideration. In order to understand their effects, one should look for a reduced model which encodes the structure of the unperturbed dynamical system but which allows one to look at the quantities of interest on an appropriate time scale. Homogenization yields a lower-dimensional model by averaging out the fast stochastic dynamics. The lower-dimensional model is strictly valid only in the limit of infinitesimally small noise. Nonetheless, the stochastically averaged models provide qualitatively useful results and are helpful in developing inexpensive lower-dimensional computational models as shown in [36]. These reduced models can be used in place of the original complex models, either for state estimation and prediction or real-time control as described below.

The starting point for this work will be the stochastic version of a multi-scale dynamical system, where Z^ϵ and X^ϵ represent the fast and slow variables, respectively.

$$dX_t^\epsilon = b(X_t^\epsilon, Z_t^\epsilon)dt + \sigma(X_t^\epsilon, Z_t^\epsilon)dB_t,$$
$$dZ_t^\epsilon = \frac{1}{\epsilon}f(X_t^\epsilon, Z_t^\epsilon)dt + \frac{1}{\sqrt{\epsilon}}g(X_t^\epsilon, Z_t^\epsilon)dW_t, \quad (4.1)$$

where $X_t^\epsilon \in \mathbb{R}^m, Z_t^\epsilon \in \mathbb{R}^n, W_t \in \mathbb{R}^l$ and $B_t \in \mathbb{R}^k$ are independent standard Brownian motions, $b : \mathbb{R}^{m+n} \to \mathbb{R}^m, \sigma : \mathbb{R}^{m+n} \to \mathbb{R}^{m\times k}, f : \mathbb{R}^{m+n} \to \mathbb{R}^n, g : \mathbb{R}^{m+n} \to \mathbb{R}^{n\times l}$. All the functions above are assumed to be Borel-measurable. For fixed $x \in \mathbb{R}^m$, define

$$dZ_t^x = f(x, Z_t^x)dt + g(x, Z_t^x)dW_t. \quad (4.2)$$

Then the solution Z^x of (4.2) with $X^\epsilon = x$ fixed is ergodic and converges rapidly to its unique stationary distribution $p_\infty(x : \cdot)$ For simplicity, let's consider a Markov process $\zeta^\epsilon \stackrel{\text{def}}{=} \{(X_t^\epsilon, Z_t^\epsilon); t \geq 0\}$ with generator

$$\mathscr{L}^\epsilon = \frac{1}{\epsilon}\mathscr{L}_F + \mathscr{L}_S,$$

where $\mathscr{L}_F$ and $\mathscr{L}_S$ represent generators of fast and slow variables. The primary objective is to derive the self-contained description of the coarse-grained dynamics without fully resolving the dynamics described in fast scales. We show that the limiting process as ϵ (scaling parameter) tends to zero is simply a Markov process $\bar{X}$ with the generator $\mathscr{L}^\dagger$:

$$\mathscr{L}^\dagger = \sum_{i=1}^m \bar{b}_i(x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^m \bar{a}_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j}, \quad (4.3)$$

where

$$\bar{b}(x) = \int b(x,z)p_\infty(x, dz), \ \bar{a}(x) = \int (\sigma\sigma^T)(x,z)p_\infty(x, dz).$$

In general, the low-dimensional models will take their values in a reduced space $\mathfrak{M}$. The geometry of this space is found from the the coarse-grained dynamics as shown in a series of papers [29, 36] on stochastic dimensional reduction.

Consider the following simple signal model [41] to illustrate the effect of homogenization (advective time scale)
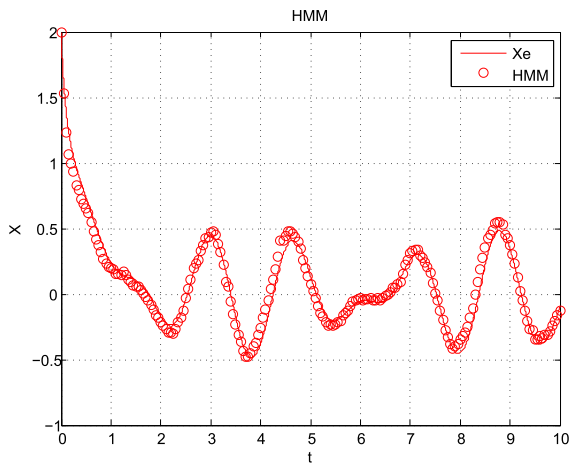
$$\dot{Z}_t^\epsilon = -\frac{1}{\epsilon}(Z_t^\epsilon - X_t^\epsilon) + \frac{1}{\sqrt{\epsilon}}\dot{W}_t, \ Z_0^\epsilon = z_0 \quad (4.4a)$$

$$\dot{X}_t^\epsilon = -(Z_t^\epsilon)^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t), \ X_0^\epsilon = x_0 \quad (4.4b)$$

For a fixed $X_t^\epsilon = x$, (4.4a) becomes the Ornstein–Uhlenbeck process in which stationary density is

$$\mu(z|x) = \frac{1}{\sqrt{\pi}}\exp\{-(z - x)^2\}.$$

As $\epsilon \to 0$, one can show that [47] $X_t^\epsilon \to X_t^0$ strongly, where X_t^0 satisfies

Fig. 1 HMM solution and the original signal

$$\dot{X}_t^0 = -(X_t^0)^3 - \frac{3}{2}X_t^0 + \sin(\pi t) + \cos(\sqrt{2}\pi t), \quad (4.5)$$

with $X_0^0 = x_0$. Figure 1 compares the Homogeneous Multi-scale Method (HMM) solution [47] with the analytical solution (4.5).

4.1 Nonstandard reduction

In the previous section, physical reasoning helped identify the time scales present in the dynamical system (4.1). Within this framework, multiple time scales constituted another type of special structure that specified a class of models whose generic properties were determined.

In general, the models are described in a high dimensional state space, say $\mathbb{R}^n$, without explicit time scales. The dynamics of interest $\zeta^\varepsilon \equiv \{\zeta_t^\varepsilon, t \geq 0\}$ of the time evolving system take place in a subspace, say $\mathfrak{M} \subset \mathbb{R}^n$. The primary objective of this section is to predict the self-contained description of this dynamics in a *stratified space* $\mathfrak{M}$ (manifolds that are required to fit together in a certain way).

For simplicity, let's consider a four dimensional Markov process $\zeta^\varepsilon \stackrel{\text{def}}{=} \{\zeta_t^\varepsilon; t \geq 0\}$ with generator $\mathscr{L}^\varepsilon = \mathscr{L}_F + \mathscr{L}_S$. The operator $\mathscr{L}_F \varphi(\zeta) = (\overline{\nabla} H, \nabla \varphi)(\zeta)$ is the vector field generating an integrable Hamiltonian and $\mathscr{L}_S$ is a second-order operator corresponding to the diffusive and dissipative perturbations. In the absence of dissipation and random

perturbations, system is integrable. The interesting part of the analysis is near bifurcations, where the structure of the fast orbits changes. To perform such an analysis, the original state space is decomposed into a collection of open subsets $\{A_1, A_2 \ldots A_n\}$ of $\mathbb{R}^4$ or $\mathbb{R}^2 \times \mathbb{R}^+ \times \mathbb{S}^1$ which are separated by hyperbolic invariant manifolds. Defining $\pi : \mathbb{R}^4 \to \mathfrak{M}$ to be the mapping, we have $\pi(A_i) \stackrel{\text{def}}{=} \Gamma_i \in \mathfrak{M}$. Inside any of the A_i's, the structure of the fast orbits are diffeomorphic and $\pi(A_i) = \Gamma_i$ is a smooth manifold with a boundary. In the absence of resonance and away from a homoclinic orbit γ, the invariant measure on the orbit can be written as a Lebesgue measure on a torus. In the A_i's, in the absence of resonance, standard stochastic averaging (see [34]) should lead to asymptotic dynamics of the law of $\pi(\zeta^\varepsilon)$; which should be diffusive as long as they remain on one of the planes of the arrowhead, Γ_i in Fig. 2.

At the line $\pi(\gamma)$ (mapping of the homoclinic manifold) where the planes of the arrowhead meet, *gluing conditions* define the behavior of the process. The glueing conditions roughly mean that when the process hits $\pi(\gamma)$, it flips a coin to decide where to make the next excursion. When ζ^ε leaves one of the A_i's, the gluing conditions essentially give a probabilistic way to select an "adjacent" A_j into which $\pi(\zeta^\varepsilon)$ should (in an asymptotic sense) make an excursion. These gluing conditions in some sense define a (possibly unfair) three-sided "coin". Informally, each time the process hits the line, this coin is flipped to decide on which plane the next excursion will occur.

To better understand the nature of the glueing conditions, let's assume that ζ^ε starts at z in the normally hyperbolic invariant manifold, γ. Assuming that the generator $\mathscr{L}_S$ is sufficiently nondegenerate, ζ^ε should make infinitely many excursions into different A_i's based upon the diffusivity of $\mathscr{L}_S$ in the directions normal to the invariant manifold γ. The effective behavior of the asymptotic reduced model should involve a combination of the normal diffusivity and the fast tangential motion. Noting that $\mathscr{L}^\varepsilon$ has a large drift component of order $1/\varepsilon$ and a small noisy part, we see that the stochastic dimensional reduction is a *singular perturbation* problem. The general theory of singular perturbations suggests a rescaling in the direction normal to the hyperbolic manifold, to make the normal diffusion of the same order as the fast motion. This corresponds to a boundary-layer
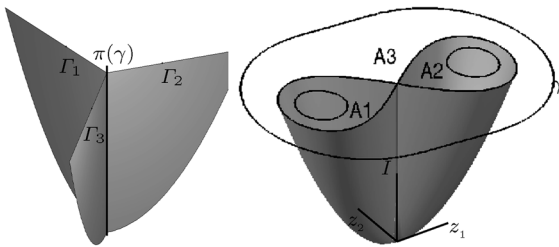
Fig. 2 Two-dimensional reduced space "arrowhead" $\mathfrak{M}$ and the unperturbed phase space
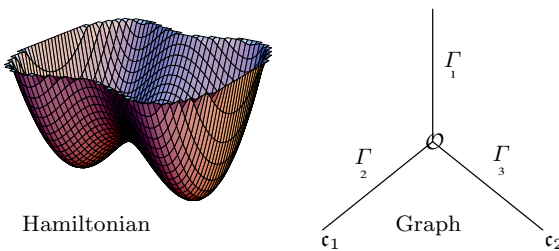
Fig. 3 Homoclinic orbit at $H^{-1}(0)$ which divides the phase space of the Hamiltonian flow (viz. $\mathbb{R}^2$) into three regions, which are mapped to the "Y" graph [18, 29]. Any two points x and y in $\mathbb{R}^2$ are equivalent, i.e., $x \sim y$, if $H(x) = H(y)$ and they are in the same connected component of $H^{-1}(H(x)) = H^{-1}(H(y))$

expansion in which the excursions into the different A_j's become comparable to the fast motion.

The averaged system will take values in a reduced space $\mathfrak{M}$. The geometry of this space is found from the unperturbed system or the coarse-grained dynamics. In a typical four dimensional example, the reduced space $\mathfrak{M}$ looks like an "arrowhead" where the phase-space regions separated by hyperbolic invariant manifolds are mapped to distinct planes (leaves).

Then, under small perturbations, the integrals of motion evolves slowly and stochastic averaging makes use of the integrable structure to identify a reduced diffusive model on a space which encodes the structure of the fixed points and have dimensional singularities. At these singularities, *glueing conditions* were derived, thereby completing the specification of the dynamics of the reduced model. It was shown in [36] that the law of $\{\pi(\zeta_t^\varepsilon); t \geq 0\}$ tends to that of a $\mathfrak{M}$-valued Markov process with a two-dimensional generator $\mathscr{L}^\dagger$

$$(\mathscr{L}^\dagger f)(x) \stackrel{\text{def}}{=} \sum_j^2 b_j^i(x) \frac{\partial f_i}{\partial x_j}(x) + \frac{1}{2} \sum_{jk}^2 a_{jk}^i(x) \frac{\partial^2 f_i}{\partial x_j \partial x_k}(x),$$

$$\text{for all } x \stackrel{\text{def}}{=} (h, I) \in \Gamma_i. \tag{4.6}$$

Here b_j^i is a drift vector and a_{jk}^i a diffusion matrix and the (three) Γ_i's denote individual leaf of the "arrowhead" $\mathfrak{M}$ (see Fig. 2). The formal derivation of the drift and diffusion coefficients is performed using the martingale problem [17, 45].

4.1.1 Example in $\mathbb{R}^2$

To explain, let's consider these results to a *simpler* $\mathbb{R}^2$-valued system. Circumventing the cumbersome algebra of four dimensional problems discussed above, consider a noisy Duffing–van der Pol oscillator

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = \alpha x_1(t) - x_1^3(t) + \epsilon(\beta + b x_1^2(t)) x_2(t) \tag{4.7}$$
$$+ \sqrt{\epsilon}(x_1(t) v_2 \xi_2(t) + v_1 \xi_1(t)),$$

with

$$H(x_1, x_2) = \frac{1}{2} x_2^2 - \alpha \frac{1}{2} x_1^2 + \frac{1}{4} x_1^4.$$

We achieve the model-reduction through *non-standard stochastic averaging*, where the reduced Markov process takes its values on a graph $\Gamma = \bigcup_{i=0}^3 [\mathfrak{c}_i] \cup \bigcup_{i=1}^3 \Gamma_i$, Fig. 3, where

$$\Gamma_1 \stackrel{\text{def}}{=} \cup_{\substack{x = (x_1, x_2) \in \mathcal{G} \\ H(x) < 0 \\ x \neq \mathfrak{c}_1 \\ x_1 < 0}} [x], \quad \Gamma_2 \stackrel{\text{def}}{=} \cup_{\substack{x = (x_1, x_2) \in \mathcal{G} \\ H(x) < 0 \\ x \neq \mathfrak{c}_2 \\ x_1 > 0}} [x],$$

$$\Gamma_3 \stackrel{\text{def}}{=} \cup_{\substack{x \in \mathcal{G} \\ H(x) > 0}} [x], \quad \text{and} \quad \mathfrak{c}_i's \text{ are the critical points,}$$

with certain glueing conditions at the vertex of the graph.

Then according to Theorem 1 in [29], the limiting process on the graph is defined by the generator
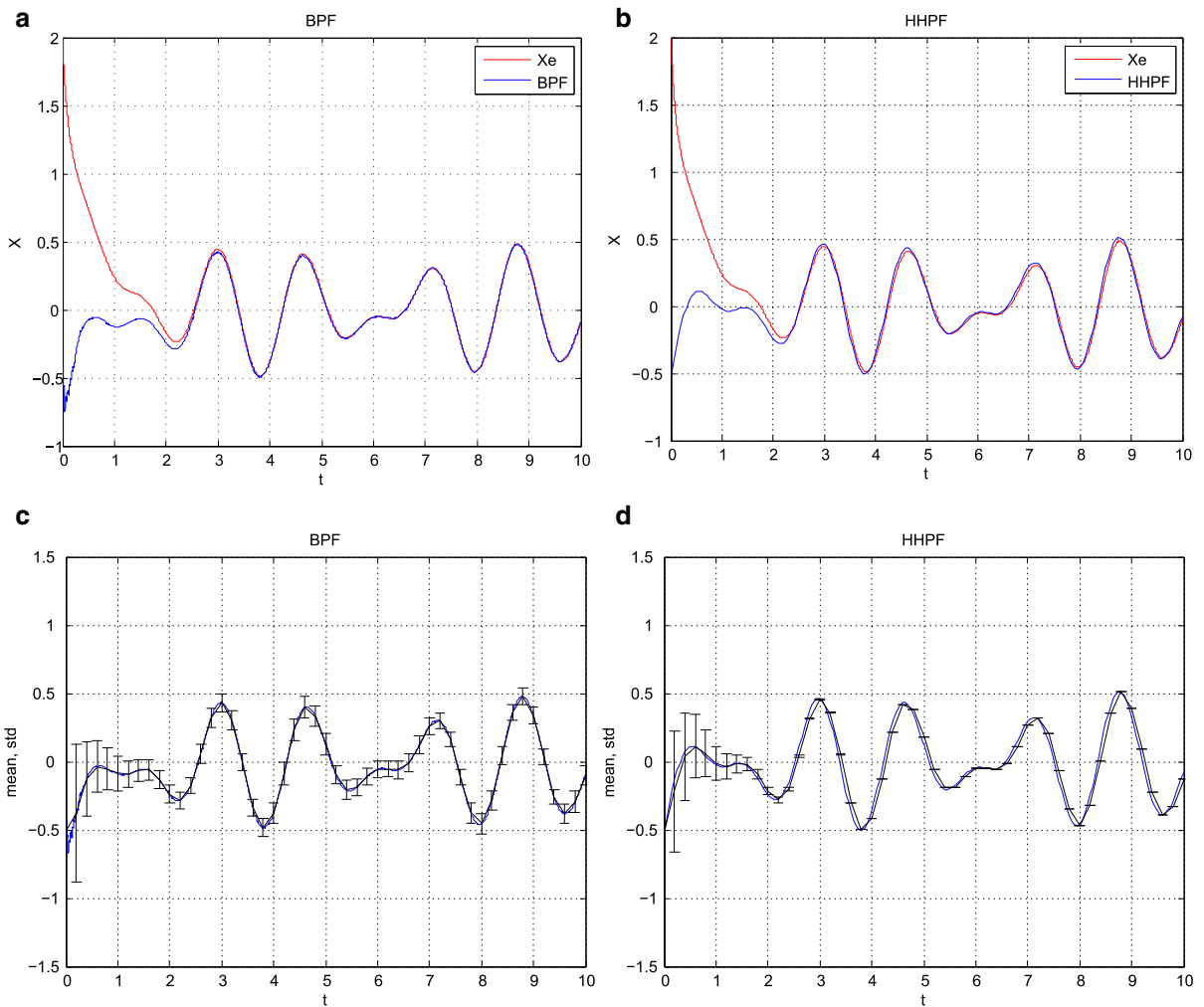
Fig. 4 a Particle filter (PF), **b** HHPF, **c** PF μ and σ and **d** HHPF μ and σ

$$\mathscr{L}_i^\dagger f_i(H) = \bar{A}_i(H)f_i' + \frac{1}{2}\bar{\sigma}_i^2(H)f_i'', \quad [x] \in \Gamma_i \qquad (4.8)$$

on the three edges of the graph, where H is only a local coordinate in each edge and it can take the same value for different trajectories. The domain of the averaged generator (4.8) is given by

$$\mathscr{D}(\mathscr{L}^\dagger) \stackrel{\text{def}}{=} \{f^\dagger \in C(\Gamma) \ : \ f^\dagger \in C^2(\cup_{i=1}^3 \Gamma_i),$$

$$\lim_{H \to H(\mathcal{O})} \mathscr{L}_i^\dagger \text{ exist, } \sum_{i=1}^3 (\pm)\sigma_i^2(\mathcal{O})f_i'(\mathcal{O}) = 0 \Bigg\}$$

$$(4.9)$$

where $f_i'(\mathcal{O}) = \lim_{H \to H(\mathcal{O})} f_i'(H)$ for $(H, i) \in I_i$ and the $\pm$sign denotes whether the coordinate H on the edge I_i

is *greater than or less than* $H(\mathcal{O})$. In Eq. (4.9), the glueing condition for the vertex $\mathcal{O}$

$$\sum_{i=1}^3 (\pm)\sigma_i^2(\mathcal{O})f_i'(\mathcal{O}) = 0 \qquad (4.10)$$

that corresponds to the saddle point, roughly means the following. Define

$$\Delta \stackrel{\text{def}}{=} \sigma_1^2(\mathcal{O}) + \sigma_2^2(\mathcal{O}) + \sigma_3^2(\mathcal{O}).$$

If the limiting process starts in edge 1 of the graph Γ, it evolves according to Eq. (4.8) with $i = 1$. Upon reaching the vertex, it flips a three-sided coin to decide where to go next. It will go back to edge 1 with likelihood$\sigma_1^2(\mathcal{O})/\Delta$, to edge 2 with likelihood
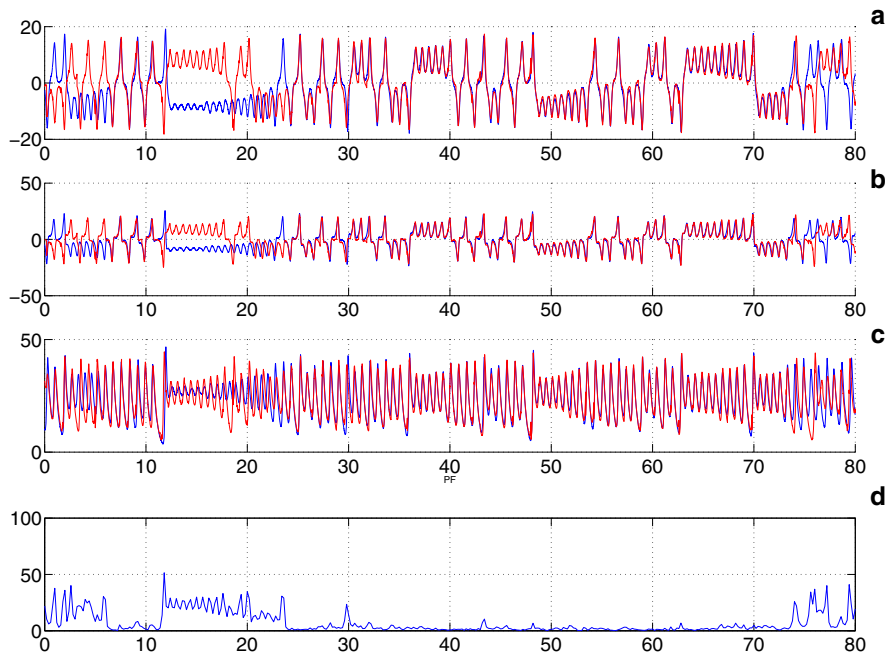
Fig. 5 Particle filter without control: **a** X^1, **b** X^2, **c** X^3 and **d** total error $e_t = ||(X_t^1, X_t^2, X_t^3)_{true} - (X_t^1, X_t^2, X_t^3)_{esti}||_2$

$\sigma_2^2(\mathcal{O})/\Delta$, and to edge 3 with likelihood $\sigma_3^2(\mathcal{O})/\Delta$. Once it is in any of these edges, it will evolve according to Eq. (4.8) with σ_1 and B_1 replaced by the appropriate σ_i and B_i. When it again hits the vertex, the coin-flipping procedure is repeated (with a new coin).

In evaluating the drift and diffusion coefficients for each edge, we change the time integral to the path integral with respect to the fast variable x_t^ε while averaging over one period of the fast motion of x_t^ε. For different values of H, we have different path integrals (oscillations or rotations) and thus different drift $\bar{A}(H)$ and diffusion coefficients $\bar{\sigma}(H)$. They are evaluated in [29, 34]. For practical problems it was crucial to consider real noise excitations. The above results were extended in [15] for non-white stationary stochastic processes for the treatment of large amplitude ship rolling in random seas.

5 Data assimilation in multi-scale systems

In a large number of applications it is essential to not only track the state of the system but also to understand if the system state has entered or is approaching a new dynamic mode in real time. This requires data and the common features of much of current data are: their complex structures—complexity in relations between different parts of the data and the nature of the data itself; noisiness—most of the measurement processes are inherently subject to random fluctuations; indirect observation—the desired state is not directly observed. This section deals with the assimilation of such data into evolving complex systems. Data assimilation or filtering involves blending information from observations of the actual system states with information from dynamical models to estimate the current system states or certain model parameters. The filtering problem relies on three fundamental ingredients, namely (1) sensor placement: where the sensors are placed in order to obtain the most useful information, (2) sensor fusion: how to combine the measurements from different sensors, and (3) estimation: how to use the measurements to obtain the best possible state estimates. Continuous time state estimation for linear stochastic systems is based on a single unifying theme, namely that state estimation is equivalent to projection onto a closed linear subspace generated by an observation process in a Hilbert space of random variables. This formulation of state estimation leads linear estimation and prediction, such as the Kalman–Bucy filtering formulae, which are much used for
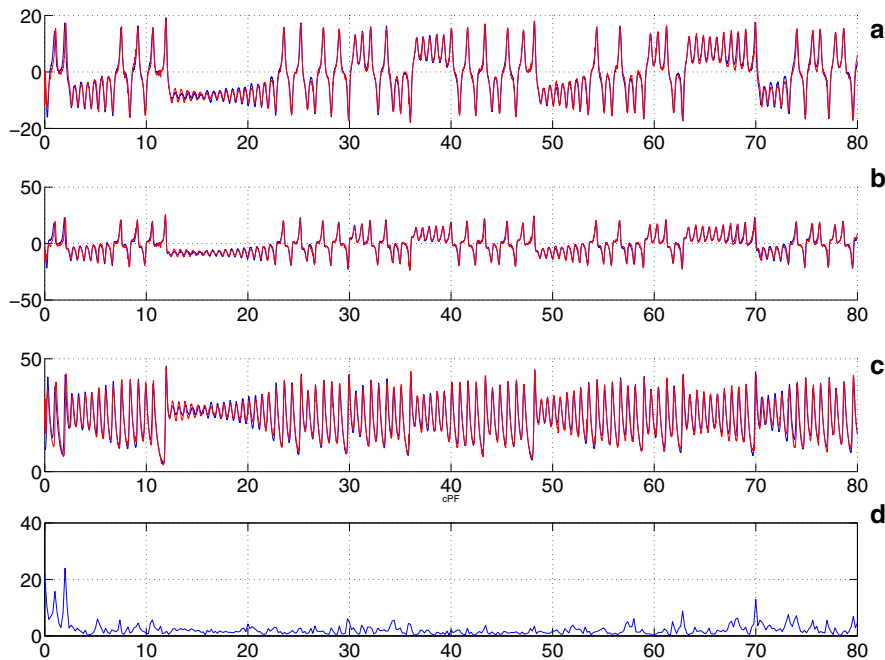
Fig. 6 Particle filter with control: **a** X^1, **b** X^2, **c** X^3 and **d** total error $e_t = ||(X_t^1, X_t^2, X_t^3)_{true} - (X_t^1, X_t^2, X_t^3)_{esti}||_2$

example in problems of inertial guidance and control in aerospace and in stochastic optimal control.

To begin with, we consider the data assimilation problem for multi-timescale nonlinear systems. An understanding of how scales interact with information can lead to the development of rigorous reduced-order data assimilation techniques for these high-dimensional problems. The nonlinear filtering problem is generically framed by augmenting the dynamics (4.1) of the state by an observation process Y_t^ε (see for example, [9, 22]). We will consider the case where the information about the state is available only indirectly through sensors (partial observation), $h(x, z)$, corrupted by sensor noise V_t, that is, a d-dimensional observation is given by

$$Y_t^\epsilon = \int_0^t h(X_s^\epsilon, Z_s^\epsilon)ds + V_t \qquad (5.1)$$

with Borel-measurable $h : \mathbb{R}^{m+n} \to \mathbb{R}^d$. V_t is assumed to be a d-dimensional standard Brownian motion that is independent of the signal (4.1) noise W_t and B_t. The only available information about the signal/state of the system is contained in the observation σ-field $\mathscr{Y}_t^\varepsilon \overset{\text{def}}{=} \sigma \{Y_s^\varepsilon : 0 \le s \le t\}$. The main objective of filtering theory is to estimate the statistics of the signal $(X_t^\varepsilon, Z_t^\varepsilon)$ at time

t based on the information $\mathscr{Y}_t^\varepsilon$ in the observation process up to time t, more precisely, for each $t \ge 0$ find the conditional *law* of $(X_t^\varepsilon, Z_t^\varepsilon)$ given $\mathscr{Y}_t^\varepsilon$.

The main objective of this section is to describe some recent results for the best estimate of the slow state X_t^ε at time t based on the information $\mathscr{Y}_t^\varepsilon$ up to time t of the observation process Y_t^ε which depends also on the fast process Z_t^ε. Since the fast variable Z_t^ε rapidly attains its invariant measure, and standard averaging techniques suggest that as $\varepsilon \searrow 0$, we should replace the dynamics of the slow variables by $\bar{X}_t$. Hence, we can average out the effects of the fast variable Z^ε, regard $\{\bar{X}_t, t \ge 0\}$ as the reduced dynamical model. The recent work by Imkeller et al. [20] showed that the marginal of the original conditional density, that is, for each $t \ge 0$ and $A \in \mathscr{B}(\mathbb{R}^n)$, $\pi_t^{\varepsilon,x}(A, \mathbf{Y}_{[0,t]}^\varepsilon) = \mathbb{P}\{X_t^\varepsilon \in A | \mathscr{Y}_t^\varepsilon\}$, converges to a $\mathcal{P}(\bar{\mathbb{X}})$-valued process $\bar{p}$ (or $\bar{u}$), as $\varepsilon \to 0$. More precisely, it was shown how the equations of filtering interact with the reduced dynamics described by the low-dimensional generator $\mathscr{L}^\dagger$ (4.3).

The convergence of $(X^\varepsilon, Y^\varepsilon)$ itself does not guarantee the convergence of filters. In a series of papers, Namachchivaya and coworkers [20, 39, 40] showed how the *scaling* interacts with *filtering*. This issue

naturally appears when one has to replace the sensor observations by their effective quantities. Both dimension reduction methods and nonlinear filtering techniques were used in [20, 40] to create new capabilities for the analysis and prediction of large-scale complex systems.

The homogenized process generated by $\mathscr{L}^{\dagger}$ is combined with the actual observation Y^{ε} in defining the recursive stochastic PDE conditional density $\bar{p}_t(\cdot, \mathscr{F}_t^{Y^{\varepsilon}})$. Also, define $\bar{h}(x) = \int h(x, z) p_{\infty}(x, dz)$. It was shown that the marginal of the original conditional density $p_t^{\varepsilon, x}(\cdot, \mathscr{F}_t^{Y^{\varepsilon}})$ is close to the homogenized conditional density $\bar{p}_t(\cdot, \mathscr{F}_t^{Y^{\varepsilon}})$. For $A \in \mathscr{B}(\mathbb{R}^n)$, the conditional law of the coarse-grained dynamics is

$$\bar{\pi}_t(A, \mathbf{Y}_{[0,t]}^{\varepsilon}) \stackrel{\text{def}}{=} \int_{x \in A} \bar{p}^{\varepsilon}(t, x) dx = \frac{\int_{x \in A} \bar{u}^{\varepsilon}(t, x) dx}{\int_{x \in \mathbb{R}^n} \bar{u}^{\varepsilon}(t, x) dx},$$

where $\bar{u}^{\varepsilon}(t, x)$ is governed by the following stochastic PDE

$$d\bar{u}^{\varepsilon}(t, x) = \mathscr{L}^{\dagger *} \bar{u}^{\varepsilon}(t, x) dt + \bar{h}(x) \bar{u}^{\varepsilon}(t, x) dY_t^{\varepsilon}, \quad (5.2)$$

with $\bar{u}^{\varepsilon}(0, x) = p_x$. The recursive computability of the lower-dimensional nonlinear filter of the coarse-grained dynamics is brought out much more explicitly through the stochastic PDE (5.2). Note that the homogenized filter is still driven by the *real observation* Y_t^{ϵ} and not by a "homogenized observation", which is practical for implementation of the homogenized filter in applications since such averaged observation is usually not available. However, even if such homogenized observation is available, using it would lead to loss of information for estimating the signal compared to using the actual observation. This theory is enabled using an efficient class of filtering methods called particle methods, which include sequential Monte Carlo and interacting particle filters (e.g., [24, 41]). Particle filters represent the posterior conditional distribution of the state variables by a collection of particles, which evolves and adapts recursively as new information becomes available. This method involves the simulation of a sample of independent particles of the signal according to the signal's stochastic law, and the resampling of these particles to incorporate information from the observations. This combined method of sampling and averaging to solve the filtering equation in a multiscale

setting is called the Homogenized Hybrid Particle Filter (HHPF).

We now apply the HHPF to an example to illustrate its potential for high-dimensional complex problems. Consider the signal model given by (4.4a) and (4.4b) with the observation

$$Y_t^{\varepsilon} = \frac{1}{2}(X_s^{\varepsilon})^2 + B_t.$$

Note that the above observation can be written in a differential form

$$dY_t^{\varepsilon} = X_t^{\varepsilon} \left\{ -(Z_t^{\varepsilon})^3 + \sin(\pi t) + \cos(\sqrt{2}\pi t) \right\} dt + dB_t,$$

which is a proper form (5.1) of HHPF. We obtain the averaged observation with the following sensor function

$$\bar{h}(X_t^0) = -(X_t^0)^4 - \frac{3}{2}(X_t^0)^2 + \{\sin(\pi t) + \cos(\sqrt{2}\pi t)\} X_t^0.$$

We then applied the algorithms described in [41] to the above simple problem using Matlab R2007a. The results from the branching particle filter and the HHPF are given in Fig. 4a, b, respectively. Both results are also compared with the analytical solution (4.5). The sample mean μ and standard deviation σ for each method are shown in Fig. 4c, d, respectively, with the error bars representing the standard deviations. The time taken for these simulations are 448 and 15 s respectively with a 2 GHZ Intel Core 2 Duo MacBook, see [41] for simulation parameters.

5.1 Optimal nudging in particle filters

Based on the above results, an efficient particle filtering algorithm for multi-scale systems was constructed. Particle methods that are adapted for dynamical systems which are inherently chaotic are used for approximating the solution to the SPDE (5.2). Importance sampling and control methods are then used as a basic and flexible tool for the construction of the proposal density inherent in particle filtering. We superimpose a control on the particle dynamics which aims to drive the particles to locations most representative of the observations, while still trying to remain faithful to the original signal dynamics. To this end, we evolve the particle i according to

$$d\widehat{X}_t^i = \bar{b}(t, \widehat{X}_t^i) dt + u^i(t) dt + \bar{\sigma}(t, \widehat{X}_t^i) dW_t,$$
$$t_k \leq t \leq t_{k+1}, \quad \widehat{X}_{t_k}^i = x_{t_k}^i, \quad (5.3)$$

where $\bar{b}$ and $\bar{\sigma}$ are defined in (4.3). Note the difference between the above equation and the homogenized SDE associated with (4.3). The $\widehat{X}$ is used to indicate that *the particles* are evolved according to the controlled dynamics which is different from the system dynamics. This control is obtained by minimizing a specific cost functional which consists of a running cost due to input energy and a terminal cost that penalizes for being away from observation. The control can be interpreted as steering particles gradually toward locations indicated by the next observation. The optimization also results in minimal weight variance for each particle. The measure change, needed to compensate for the addition of control in the "prognostic" equations, corresponds to that involved in optimal importance sampling. The filtering algorithm presented in [26] utilizes the next available observation to steer particles in the time interval in between observations to construct better posterior densities at the observation times.

In practice, evaluation of the optimal control based on the Feynman–Kac representation and Malliavin derivative can become computationally overwhelming for nonlinear signals. However, the optimal control solution can be obtained explicitly for linearized systems:

$$u(t,x) = Q(e^{A(t_{k+1}-t)})^*[I + H^*R^{-1}H\Sigma]^{-1}$$
$$\cdot H^*R^{-1}[(Y_{t_{k+1}} - H\mu)] \tag{5.4}$$

where A is the linear operator, $\mu := e^{A(t_{k+1}-t)}x$, and $\Sigma := \int_t^{t_{k+1}} e^{A(t_{k+1}-s)}Q(e^{A(t_{k+1}-s)})^*ds$ are the mean and variance of the linearized system at time t_{k+1} when it starts at time t at x.

As shown in [26], a linear control strategy was implemented as a suboptimal control solution on particle filtering for the 3-dimensional Lorenz '63 system. The signal and observation processes are given as:

$$d\begin{bmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{bmatrix} = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho & -1 & 0 \\ 0 & 0 & -\beta \end{bmatrix} \begin{bmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{bmatrix} dt$$
$$+ \begin{bmatrix} 0 \\ -X_t^1 X_t^3 \\ X_t^1 X_t^2 \end{bmatrix} dt + \alpha\, dW_t$$
$$dY_{t_k} = hX_{t_k} + g\,dB_{t_k}, \quad k = 1, 2, \dots$$

We use the standard parameters $\sigma = 10$, $\rho = 8/3$, and $\beta = 28$ in the signal equations. The signal and sensor noise is simulated as a vector of Gaussian random numbers premultiplied by the following correlation matrices,

$$Q \overset{\text{def}}{=} \alpha\alpha^T = \begin{bmatrix} 1 & 0.5 & 0.25 \\ 0.5 & 1 & 0.5 \\ 0.25 & 0.5 & 1 \end{bmatrix} \text{ and}$$
$$R \overset{\text{def}}{=} gg^T = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

The sensor function is a 3×3 identity matrix, $h \overset{\text{def}}{=} I_{3\times3}$. Observations are recorded every $\Delta t = 0.2$, which corresponds to roughly $1/4$ of the error doubling time for the deterministic Lorenz '63 system.

The filtering results are shown in Figs. 5 and 6. Implementation of the suboptimal control was sufficient to ensure consistent tracking of the signal, even when the time interval between observations is large. The difference between the estimated state and the true state of the system constitutes the error

$$e_t = \sqrt{(\mathbf{X}_t - \hat{\mathbf{X}}_t)(\mathbf{X}_t - \hat{\mathbf{X}}_t)^T},$$

where $\mathbf{X}_t$ represent true signal and $\hat{\mathbf{X}}_t$ be the estimate, in specifying or forecasting the state, which is amplified in chaotic systems that have a number of positive Lyapunov exponents. Figure 5 shows the results of the particle filter without any control. A total of 20 particles are used, and resampling is done if the effective number of particles falls below 5. The time step used for integration is $dt = 0.004$ and the observation is received every $50\,dt$. This is a difficult filtering problem as discussed in [26] and the particle filter without any control misses the transitions from one wing of the Lorenz butterfly to the other, as can be seen from Fig. 5. For the same parameters and with the same initial sample as the above filter, Fig. 6 shows the results of the particle filter with the approximate linear control (5.4) where the trajectories of the controlled particle filter tracks "true" signal. Resampling is done (if the effective number of particles falls below 5) only at the time when observation is recorded.

The filtering algorithm can be supplemented by a scheme that extracts more accurate observations (that contain the maximum amount of information), which has the potential to further reduce the error in the

analysis of the initial state for the forecast. For chaotic systems, solution settles near a subset of the state space, called an attractor. The state trajectories are sensitive to initial conditions, i.e. trajectories starting from initial conditions that are close can deviate far apart in the future. This sensitivity to initial conditions is characterized by finite time Lyapunov exponents. Observations can potentially be improved by constructing a sensor function that is more sensitive in unstable directions (directions corresponding to positive Lyapunov exponents).

The filtering algorithm described so far can be supplemented by a sensor design scheme that extracts better observations (that contain the maximum amount of information), which has the potential to further reduce the error in the analysis of the initial state for the forecast. Using information-theoretic formulations and information flow, adaptive sensing can be designed to extract more information content at a given observation time. Finally, [48] considers sensor selection with the goal of improving the analysis at observation times, by looking at mutual information between the signal and observation. Additionally, in chaotic systems, error growth and uncertainty can be characterized by Lyapunov exponents. Specifically, for a chaotic system, solutions settle near a subset of the state space, called an attractor. However, state trajectories are sensitive to initial conditions, i.e. trajectories starting from initial conditions that are close can deviate far apart in the future. This sensitivity to initial conditions is characterized by (finite time) Lyapunov exponents. Observations can potentially be improved by constructing a sensor function that is more sensitive in unstable directions (directions corresponding to positive Lyapunov exponents).

6 Conclusions

Summarizing the results of this paper, the first section presented result on the maximal Lyapunov exponent of the response of multi-degree of freedom bilinear [32, 35, 46] and single-well nonlinear oscillators [5, 30] driven by either an additive or multiplicative white noise. Then it was shown that stochastic homogenization of the nonlinear systems yielded a set of equations which, together with their variational equations, were explicitly solved and hence their

bifurcation behavior completely analyzed [2, 29, 44]. Nonstandard reduction methods [29] with glueing conditions (boundary conditions at the "tipping points" for the reduced diffusion) were applied to study the behavior of noisy, strongly nonlinear mechanical systems with bifurcations in their fast deterministic dynamics. The relationship between stochastic bifurcations and how they are related to Lyapunov and moment Lyapunov exponents [3, 28] were examined. These results make the design of mechanical systems robust and will have an important impact on the design of advanced engineering systems and their reliability.

In the context of filtering applications, the reduced-order nonlinear filtering equations [20, 24] provided the estimation of coarse-grained dynamics (slow dynamics) without requiring explicit knowledge of the fast dynamics, hence reducing computational complexities and information storage requirements. Finally, we presented new particle filters that represent the posterior conditional distribution of the state variables by a collection of particles, which evolves and adapts recursively as new information becomes available. Lorenz '63 system was an excellent test-bed for the optimal importance sampling scheme that was developed in [26]. In most real applications, for example weather prediction models, the observation process is within the same environment as the signal; hence the signal and sensor noise are correlated. Hence the filter convergence results of [20, 41] need to be extended to the correlated noise setting.

The change in scale and complexity of the types of data and phenomena being studied in high dimensional, multi-scale complex systems poses new mathematical and conceptual challenges. In this data-centric multi-scale environment, model-based and data-driven methods are at the center of development of a large number of key fields in the sciences and engineering. The scope of data-centric methods spans a wide range of areas, from technological (electric power grid) and geophysical (climate and weather) to environmental (chemical and biological) and social systems (crime and terrorism). The bottleneck is the lack availability of good statistical techniques that quickly and effectively extract useful information and assimilate them for cooperative health monitoring and vulnerability assessment in real time. It is this need and challenge that was addressed in this paper.

There are several topics in RDS that are not covered in this paper due to lack of space. They include *large deviations* [14, 19], which is concerned with asymptotic estimates of probability of rare events associated with stochastic processes, *information theoretic methods* [48] to quantify the time evolution of uncertainty about the signal and mutual information between signal and observation, and *random vibration* [13] which provides a more applied aspect of RDS, to name a few. The same information theoretic concepts as in the discrete-time sensor selection problem are used, but now we will revert to the continuous-time setting and obtain equations for uncertainty and information flow.

References

1. Arnold L, Oeljeklaus E, Pardoux E (1986) Almost sure and moment stability for linear itô equations. Lyapunov exponents, lecture notes in mathematics, vol 1186. Springer, New York

2. Arnold L, Namachchivaya NS, Schenk KL (1996) Toward an understanding of stochastic Hopf bifurcations: a case study. J Bifurc Chaos 6:1947–1975

3. Arnold L, Doyle M, Namachchivaya NS (1997) Small noise expansion of moment Lyapunov exponents for two-dimensional systems. Dyn Stab Syst 12:187–211

4. Arnold L (1998) Random dynamical systems. Springer monographs in mathematics. Springer, Berlin

5. Arnold L, Imkeller P, Namachchivaya NS (2004) The asymptotic stability of a noisy non-linear oscillator. J Sound Vib 269(3–5):1003–1029

6. Arnold L, Imkeller P (1998) Normal forms for stochastic differential equations. Probab Theory Relat Fields 110(4):559–588

7. Arnold L, Khasminskii R (1995) Stability index for nonlinear stochastic differential equations. Proc Symp Pure Math 57:543–551

8. Arnold L, Scheutzow M (1995) Perfect cocycles through stochastic differential equations. Probab Theory Relat Fields 101:65–88

9. Bain A, Crisan D (2009) Fundamentals of stochastic filtering. Springer monographs in mathematics. Springer, Berlin

10. Baxendale PH (1994) A stochastic Hopf bifurcation. Probab Theory Relat Fields 99(2):581–616

11. Baxendale PH (2002) Lyapunov exponents for small random perturbations of Hamiltonian systems. Ann Probab 30(1):101–134

12. Baxendale PH (2004) Stochastic averaging and asymptotic behavior of the stochastic Duffing–van der Pol equation. Stoch Process Appl 113(2):235–272

13. Crandall S, Zhu WQ (1983) Random vibration: a survey of recent developments. J Appl Mech 50(4b):953–962

14. Dembo A, Zeitouni O (2010) Large deviations techniques and applications. Stochastic modelling and applied probability. Springer, Berlin

15. Dostal L, Kreuzer E, Namachchivaya N Sri (2012) Nonstandard stochastic averaging of large-amplitude ship rolling in random seas. Proc R Soc A 468(2148):4146–4173

16. Doyle M, Namachchivaya NS (1994) Almost-sure asymptotic stability of a general four dimensional dynamical system driven by real noise. J Stat Phys 75(3/4):525–552

17. Ethier SN, Kurtz TG (1986) Markov processes: characterization and convergence. Wiley, New York

18. Freidlin MI, Weber M (1998) Random perturbations of nonlinear oscillators. Ann Probab 26(3):925–967

19. Freidlin MI, Wentzell AD (1998) Random perturbations of dynamical systems. Grundlehren der Mathematischen Wissenschaften, vol 260. Springer, Berlin

20. Imkeller P, Namachchivaya NS, Perkowski N, Yeong HC (2013) Dimensional reduction in nonlinear filtering: a homogenization approach. Ann Appl Probab 23(6):2161–2603

21. Imkeller P, Lederer C (1999) An explicit description of the Lyapunov exponents of the noisy damped harmonic oscillator. Dyn Stab Syst 14(4):385–405

22. Kushner HJ (1990) Weak convergence methods and singularly perturbed stochastic control and filtering problems. Birkhäuser, Boston

23. Leng G, Namachchivaya NS, Talwar S (1992) Robustness of nonlinear systems perturbed by external random excitation. ASME J Appl Mech 59(4):1015–1022

24. Lingala N, Namachchivaya NS, Perkowski N, Yeong HC (2012) Particle filtering in high-dimensional chaotic systems. Chaos 22:047509

25. Lingala N, Namachchivaya NS, O'Reilly OM, Wihstutz V (2013) Almost sure asymptotic stability of an oscillator with delay feedback excited by finite state Markov noise. Probab Eng Mech 32:21–30

26. Lingala N, Namachchivaya NS, Perkowski N, Yeong HC (2014) Optimal nudging in particle filters. Probab Eng Mech 32:21–30

27. Lingala N, Namachchivaya NS (2016) Perturbations of linear delay differential equations at the verge of instability. Phys Rev E 93:21–30

28. Moshchuk N, Khasminskii R (1998) Moment Lyapunov exponent and stability index for linear conservative system with small random perturbation. SIAM J Appl Math 58(1):245–256

29. Namachchivaya NS, Sowers RB, Vedula L (2001) Non-standard reduction of noisy Duffing—van der Pol equation. J Dyn Syst 16(3):223–245
30. Namachchivaya NS, Kok D, Ariaratnam ST (2007) Stability of noisy nonlinear auto-parametric systems. Nonlinear Dyn 47(3):143–165
31. Namachchivaya NS, Lin YK (1991) Method of stochastic normal forms. Int J Nonlinear Mech 26(6):931–943
32. Namachchivaya NS, Van Roessel HJ (1993) Maximal Lyapunov exponent and rotation numbers for two coupled oscillators driven by real noise. J Stat Phys 71(3/4):549–567
33. Namachchivaya NS, Van Roessel HJ (2004) Averaging of noisy nonlinear systems with rapidly oscillating and decaying components. Nonlinear Dyn 36:329–347
34. Namachchivaya N Sri, Sowers R (2002) Rigorous stochastic averaging at a center with additive noise. Meccanica 37(2):85–114
35. Namachchivaya NS, Vedula L (2000) Stabilization of linear systems by noise: application to flow induced oscillations. Dyn Stab Syst 15(2):185–208
36. Onu K, Namachchivaya N Sri (2010) Stochastic averaging of surface waves. Proc R Soc A 466(3):2363–2381
37. Oseledec VI (1968) A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. Trans Mosc Math Soc 19:197–231
38. Papanicolaou GC, Stroock D, Varadhan SRS (1977) Martingale approach to some limit theorem. In: Statistical mechanics and dynamical systems, and papers from 1976 turbulence conference. Mathematics Department, Duke University
39. Park JH, Namachchivaya NS, Sowers RB (2008) A problem in stochastic averaging of nonlinear filters. Stoch Dyn 8(3):543–560
40. Park JH, Namachchivaya NS, Sowers RB (2010) Dimensional reduction in nonlinear filtering. Nonlinearity 23(2):305–325
41. Park JH, Namachchivaya NS, Yeong HC (2011) Particle filters in a multiscale environment: homogenized hybrid particle filter. J Appl Mech 78(6):1–10
42. Popp K, Romberg O (2001) Influence of stochastic effects on flow induced vibrations in tube bundles. In: Narayanan S, Iyengar RN (eds) IUTAM symposium on nonlinearity and stochastic structural dynamics. Solid mechanics and its applications, vol 85. Kluwer Academic Publishers, pp 197–208
43. Schimansky-Geier L, Herzel H (1993) Positive Lyapunov exponents in the Kramers oscillator. J Stat Phys 782:141–147
44. Singh P, Yeong HC, Zhang H, Rapti Z, Sri Namachchivaya N (2016) Stochastic stability and dynamics of a two-dimensional structurally nonlinear airfoil in turbulent flow. Meccanica. doi:10.1007/s11012-016-0445-8
45. Stroock DW, Varadhan SRS (1979) Multidimensional diffusion processes. Grundlehren der Mathematischen Wissenschaften, vol 233. Springer, Berlin
46. Wedig WV (1990) Invariant measures and Lyapunov exponents for generalized parameter fluctuations. Struct Saf 8(1–4):13–25
47. Weinan E, Liu D, Vanden-Eijnden E (2005) Analysis of multiscale methods for stochastic differential equations. Commun Pure Appl Math 58:1544–1585
48. Yeong HC, Beeson R, Namachchivaya NS, Perkowski N, Sauer PW (2017) Dynamic data-driven adaptive observations in data assimilation for multi-scale systems. In: Ravela S, Blasch E, Aved A (eds) Dynamic data driven applications systems. Springer, New York (to appear)
49. Zakai M (1969) On the optimal filtering of diffusion processes. Z Wahrscheinlichkeitstheorie Verwandte Geb 11:230–243
50. Zeeman E (1988) On the classification of dynamical systems. Bull Lond Math Soc 20:545–557