Data-driven structure-preserving model reduction for stochastic Hamiltonian systems

Tomasz M. Tyranowski*[1,2,3]

[1]University of Twente, Department of Applied Mathematics
PO Box 217, 7500AE Enschede, The Netherlands
[2]Max-Planck-Institut für Plasmaphysik
Boltzmannstraße 2, 85748 Garching, Germany
[3]Technische Universität München, Zentrum Mathematik
Boltzmannstraße 3, 85748 Garching, Germany

Abstract

In this work we demonstrate that SVD-based model reduction techniques known for ordinary differential equations, such as the proper orthogonal decomposition, can be extended to stochastic differential equations in order to reduce the computational cost arising from both the high dimension of the considered stochastic system and the large number of independent Monte Carlo runs. We also extend the proper symplectic decomposition method to stochastic Hamiltonian systems, both with and without external forcing, and argue that preserving the underlying symplectic or variational structures results in more accurate and stable solutions that conserve energy better than when the non-geometric approach is used. We validate our proposed techniques with numerical experiments for a semi-discretization of the stochastic nonlinear Schrödinger equation and the Kubo oscillator.

1 Introduction

The purpose of this work is twofold: to demonstrate that the conventional SVD-based model reduction methods for ordinary differential equations (ODEs) can be extended to stochastic differential equations (SDEs), and to show that in the case of stochastic Hamiltonian systems it is beneficial to maintain their symplectic (or variational) structure in the construction of the reduced spaces by extending the existing structure-preserving model reduction techniques for deterministic Hamiltonian systems.

Model reduction methods have been introduced in order to reduce the computational cost of solving high-dimensional dynamical systems. The goal of these techniques is to construct lower-dimensional models which are less expensive to solve numerically, but still capture the dominant features of the dynamics of the original system (see [11], [20], [21], [133] and the references therein). The Proper Orthogonal Decomposition (POD), first introduced in [109], is an SVD-based data-driven model reduction technique that employs an offline-online splitting. In the offline stage the

*tomasz.tyranowski@ipp.mpg.de

available empirical data about the solutions of the full system are used to identify an optimal subspace of the full configuration space. The equations governing the evolution of the system are then projected to that subspace, and in the online stage such a reduced model is solved numerically at a lower computational cost. The POD method has proved very successful and has been used in many scientific and engineering problems, such as fluid dynamics ([84], [89], [90], [99], [104], [127], [140]), electric circuit analysis ([121]), or structural dynamics ([4]); see also [30], [43], [65], [71], [135].

Consider a general stochastic differential equation

$$d_t u = a(u)\,dt + \sum_{\nu=1}^{m} b_\nu(u) \circ dW^\nu(t), \tag{1.1}$$

where $a : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the drift function, $[b_1, \ldots, b_m] : \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m}$ is the diffusion matrix, $W(t) = (W^1(t), \ldots, W^m(t))$ is the standard m-dimensional Wiener process, and $\circ$ denotes Stratonovich integration. We use d_t to denote the stochastic differential of stochastic processes (other than the Wiener process $W(t)$) to avoid confusion with the exterior derivative d of differential forms. We will assume that the drift function and the diffusion matrix are sufficiently smooth and satisfy all the necessary conditions for the existence and uniqueness of solutions to (1.1) (see [12], [87], [95], [100]). Much like in the deterministic case, numerical simulations of (1.1) become computationally expensive when the dimension n of the stochastic process $u(t)$ is large. Such a situation may occur, for instance, when (1.1) comes from a semi-discretization of a stochastic partial differential equation (see Section 5.1). Also, in order to calculate the statistical properties of the stochastic process $u(t)$, such as the probability density function or the expected value, one typically needs to simulate (1.1) for a very large number of sample paths. We will show that the computational cost of both these tasks can be alleviated by adapting the POD method to the stochastic setting. A number of model reduction methods have been applied to stochastic systems in various contexts. Examples include partial differential equations with random coefficients ([24], [25], [64], [69]), parametric closure models for the stochastic Burgers' equation ([107]), or a variance reduction method for SDEs ([26]). Stochastic reduced models were also used as noisy perturbations of deterministic reduced models in order to account for unresolved small-scale features ([16], [49], [108], [122]). However, to the best of our knowledge the adaptation of the POD method to SDEs driven by a Wiener process has been much less investigated. An application of the POD method which is similar in spirit to our approach appears in [88] and [164], but only in the specific context of solving the stochastic Burgers equation, empirical approximation of the nonlinear term is not addressed, and only low-order integration in time is used (see also [32]).

The POD method can, in principle, be applied to a Hamiltonian system. There is, however, no guarantee that the reduced system will maintain the Hamiltonian structure, nor that it will be stable, which may result in a blow-up of its solutions (see [132], [135]). A model reduction technique that retains the symplectic structure of Hamiltonian systems was introduced in [126]. In analogy to POD, this method is called the Proper Symplectic Decomposition (PSD). The PSD method has been proven to preserve the energy and stability of the system, and to yield better numerical solutions, especially when combined with symplectic integration in time. Therefore, the PSD method is better suited for model reduction of Hamiltonian systems than the classical POD approach, especially when long-time integration is required; see also [1], [2], [3], [40], [42], [74], [79], [93], [102], [124], [125], [131].

A stochastic Hamiltonian system is an SDE of the form

$$d_t q = \frac{\partial H}{\partial p}\, dt + \sum_{\nu=1}^{m} \frac{\partial h_\nu}{\partial p} \circ dW^\nu(t), \qquad d_t p = -\frac{\partial H}{\partial q}\, dt - \sum_{\nu=1}^{m} \frac{\partial h_\nu}{\partial q} \circ dW^\nu(t), \tag{1.2}$$

where $H : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ and $h_\nu : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ for $\nu = 1, \ldots, m$ are the Hamiltonian functions. Such systems can be used to model, e.g., mechanical systems with uncertainty, or error, assumed to arise from random forcing, limited precision of experimental measurements, or unresolved physical processes on which the Hamiltonian of the deterministic system might otherwise depend. Particular examples include modeling synchrotron oscillations of particles in particle storage rings (see [63], [143]) and stochastic dynamics of the interactions of singular solutions of the EPDiff basic fluids equation (see [81], [82]). More examples are discussed in Section 5; see also [83], [106], [112], [119], [141], [144], [146], [151]. Similar to their deterministic counterparts, stochastic Hamiltonian systems possess several important geometric features. In particular, their phase space flows (almost surely) preserve the canonical symplectic structure. We will argue that maintaining this property in model reduction is beneficial because the resulting reduced systems can then be integrated using stochastic symplectic methods (see [6], [7], [8], [9], [27], [36], [38], [47], [50], [52], [53], [62], [81], [83], [85], [86], [97], [110], [111], [115], [116], [117], [150], [160], [161], [163], [166]), and consequently more accurate numerical solutions can be obtained. This goal can be achieved by an appropriate adaptation of the PSD method in the stochastic setting. We are not aware of any previous studies on this topic.

Main content The main content of the remainder of this paper is, as follows.

In Section 2 we briefly review the POD method and present how it can be applied to the general SDE (1.1) in computations involving a single realization of the Wiener process.

In Section 3 we briefly review the PSD method and present how it can be applied to the stochastic Hamiltonian system (1.2) in computations involving a single realization of the Wiener process. We also further extend the PSD method to stochastic forced Hamiltonian systems.

In Section 4 we discuss how the problem of calculating the statistical properties of solutions to (1.1) or (1.2) can be recast as the problems presented in Section 2 and Section 3, respectively.

In Section 5 we present the results of our numerical experiments for the stochastic Nonlinear Schrödinger Equation and the Kubo oscillator, both with and without external forcing.

Section 6 contains the summary of our work.

2 Model reduction for SDEs

The Proper Orthogonal Decomposition is one of the standard model reduction techniques for ordinary differential equations (see [11], [20], [21], [133]). In this section we demonstrate how the POD method can be adapted in the context of stochastic differential equations.

2.1 Proper Orthogonal Decomposition

Suppose we would like to solve (1.1) for a single realization of the Wiener process $W(t)$. If the dimension n of the stochastic process $u(t)$ is a very high number, then the system (1.1) becomes

very expensive to solve numerically. The main idea of model reduction is to approximate such a high-dimensional stochastic dynamical system using a lower-dimensional one that can capture the dominant dynamic properties. Let Δ be an $n \times r$ matrix representing empirical data on the system (1.1). For instance, Δ can be a collection of snapshots of a solution of this system for the given realization of the Wiener process,

$$\Delta = [u(t_1)\, u(t_2)\, \dots\, u(t_r)], \tag{2.1}$$

at times $t_1, \dots, t_r$. These snapshots are calculated for a particular set of initial conditions or values of parameters that the system (1.1) depends on (see Section 5). A low-rank approximation of Δ can be done by performing the singular value decomposition (SVD) of Δ and truncating it after the first k largest singular values, that is,

$$\Delta = U\Sigma V^T \approx U_k \Sigma_k V_k^T, \tag{2.2}$$

where $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \dots)$ is the diagonal matrix of the singular values, U and V are orthogonal matrices, Σ_k is the diagonal matrix of the first k largest singular values, and U_k and V_k are orthogonal matrices constructed by taking the first k columns of U and V, respectively. Let ξ denote a vector in $\mathbb{R}^k$. Substituting $u = U_k \xi$ in (1.1) yields a reduced SDE for $\xi(t)$ as

$$d_t \xi = U_k^T a(U_k \xi)\, dt + \sum_{\nu=1}^{m} U_k^T b_\nu(U_k \xi) \circ dW^\nu(t). \tag{2.3}$$

The corresponding initial condition is calculated as $\xi_0 = U_k^T u_0$, where u_0 is an initial condition for (1.1). If the singular values of Δ decay sufficiently fast, then one can obtain a good approximation of Δ for k such that $k \ll n$. Equation (2.3) is then a low-dimensional approximation of (1.1) and can be solved more efficiently. The approximate solution of (1.1) is then reconstructed as $u(t) = U_k \xi(t)$. The process of constructing the matrix U_k from the empirical/simulation data ensemble Δ is typically called the *offline* stage of model reduction. Depending on the size of the data, this stage can be computationally expensive, but it is performed only once. Solving the low-dimensional system (2.3) is usually referred to as the *online* stage of model reduction, and is supposed to be faster and more efficient than solving the full system (1.1).

2.2 Discrete Empirical Interpolation Method

If the drift a or any of the diffusion terms b_ν is a complicated nonlinear function, then solving the reduced system (2.3) may not bring any computational savings, because one usually needs to compute the state variable $u = U_k \xi$ in the original coordinate system, evaluate the nonlinear drift or diffusion terms, and then project back to the column space of U_k. A technique called the Discrete Empirical Interpolation Method (DEIM) has been developed for ODEs in order to reduce the complexity in evaluating the nonlinear terms (see [43], [135]). This technique can be readily adapted also in the stochastic setting. For completeness and for the benefit of the reader, below we briefly outline the main ideas of DEIM. The further details can be found in, e.g., [43], [126]. Let the drift function be expressed as

$$a(u) = Lu + a_N(u), \tag{2.4}$$

where L is an $n \times n$ matrix, and a_N represents the nonlinear part of a. First, suppose that $a_N(u)$ lies approximately in the range of an $n \times \bar{k}$ matrix Ψ. Similar to U_k in (2.2), the matrix Ψ can be found by performing the SVD of another empirical data ensemble, namely

$$\bar{\Delta} = [a_N(u(t_1))\ a_N(u(t_2))\ \ldots\ a_N(u(t_r))], \tag{2.5}$$

and truncating it after the first $\bar{k}$ largest singular values. Next, calculate only $\bar{k}$ components of the vector $a_N(u)$ with the preselected indices $\beta_1, \ldots \beta_{\bar{k}}$. Those $\bar{k}$ components of $a_N(u)$ can be conveniently denoted by the expression $P^T a_N(u)$, where the $n \times \bar{k}$ matrix P is defined as

$$P = [\mathbb{e}_{\beta_1}, \ldots, \mathbb{e}_{\beta_{\bar{k}}}], \tag{2.6}$$

and $\mathbb{e}_{\beta_i}$ denotes the β_i-th column of the identity matrix $\mathbb{I}_n$. Given Ψ, the suitable set of indices $\beta_1, \ldots \beta_{\bar{k}}$ can be inductively constructed using a greedy algorithm (see Algorithm 1 in [43] for details). The DEIM approximation of the nonlinear term is then expressed as

$$\bar{a}_N(u) = \Psi(P^T\Psi)^{-1}P^T a_N(u), \tag{2.7}$$

and the approximation of the drift term in (2.3) becomes

$$U_k^T a(U_k\xi) \approx \bar{L}\xi + W g(\xi), \tag{2.8}$$

where

$$\bar{L} = U_k^T L U_k, \qquad W = U_k^T \Psi (P^T\Psi)^{-1}, \qquad g(\xi) = P^T a_N(U_k\xi). \tag{2.9}$$

Note that the matrices $\bar{L}$ and W are calculated only once at the offline stage, and at the online stage the function $g(\xi)$ evaluates only $\bar{k}$ components of $a_N(U_k\xi)$. A similar DEIM approximation can be applied to each of the diffusion terms $b_\nu(u)$, thus reducing the computational complexity of the reduced system (2.3).

2.3 Time integration

The SDEs (1.1) and (2.3) can be solved numerically using any general purpose stochastic numerical schemes (see [95], [114] and the references therein). In this work we will focus our attention on several stochastic Runge-Kutta methods, namely the explicit stochastic Heun and $R2$ methods (see [33], [34], [35], [37], [95]), and the implicit stochastic midpoint method (see [83], [97], [111], [116]). The stochastic mipoint method for (1.1) takes the form

$$u_{i+1} = u_i + a\left(\frac{u_i + u_{i+1}}{2}\right)\Delta t + \sum_{\nu=1}^{m} b_\nu\left(\frac{u_i + u_{i+1}}{2}\right)\Delta W^\nu, \tag{2.10}$$

where Δt denotes the time step and $\Delta W = (\Delta W^1, \ldots, \Delta W^m)$ are the increments of the Wiener process. All the mentioned methods are strongly convergent of order 1 for systems driven by a commutative noise, and of order $1/2$ in the non-commutative case.

3 Model reduction for stochastic Hamiltonian systems

The Proper Symplectic Decomposition is a model reduction technique that has been developed for deterministic Hamiltonian systems (see [1], [126]). In this section we demonstrate how the PSD method can be adapted in the context of stochastic Hamiltonian systems.

3.1 Proper Symplectic Decomposition

The stochastic Hamiltonian system (1.2) can be equivalently written as

$$d_t u = \mathbb{J}_{2n} \nabla_u H(u)\, dt + \sum_{\nu=1}^{m} \mathbb{J}_{2n} \nabla_u h_\nu(u) \circ dW^\nu(t), \tag{3.1}$$

where $u = (q, p)$ and $\mathbb{J}_{2n}$ is the canonical symplectic matrix defined as

$$\mathbb{J}_{2n} = \begin{pmatrix} 0 & \mathbb{I}_n \\ -\mathbb{I}_n & 0 \end{pmatrix}, \tag{3.2}$$

with $\mathbb{I}_n$ denoting the $n \times n$ identity matrix. In a manner similar to its deterministic counterpart (see, e.g., [75], [80], [113]), the stochastic Hamiltonian system (3.1) possesses several characteristic properties. Its stochastic flow F_t (almost surely) preserves the canonical symplectic form $\Omega = \sum_{i=1}^{n} dq^i \wedge dp^i$ on the phase space $\mathbb{R}^n \times \mathbb{R}^n$. This property expressed in terms of the standard basis for $\mathbb{R}^{2n}$ takes the form of the condition

$$(DF_t)^T \mathbb{J}_{2n} DF_t = \mathbb{J}_{2n}, \tag{3.3}$$

where DF_t denotes the Jacobi matrix of the flow map F_t. Moreover, if the Hamiltonian function H commutes with all the Hamiltonian functions h_ν, that is, if the canonical Poisson bracket satisfies

$$\{H, h_\nu\} = (\nabla_u H)^T \mathbb{J}_{2n} \nabla_u h_\nu = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial q^i} \frac{\partial h_\nu}{\partial p^i} - \frac{\partial H}{\partial p^i} \frac{\partial h_\nu}{\partial q^i} \right) = 0 \tag{3.4}$$

for all $\nu = 1, \ldots, m$, then the flow F_t also preserves (almost surely) the Hamiltonian function H, which can be easily verified by calculating the stochastic differential

$$d_t H(u(t)) = \sum_{\nu=1}^{m} \{H, h_\nu\} \circ dW^\nu(t), \tag{3.5}$$

where $u(t) = F_t(u_0)$ is the solution of (3.1) with the initial condition $u(0) = u_0$, and we used the rules of Stratonovich calculus (see [22], [83], [105], [116]).

Suppose we would like to solve (3.1) for a single realization of the Wiener process. Again, if the dimension $2n$ of the stochastic process $u(t)$ is very high, then the system (3.1) becomes expensive to solve. In principle, the POD method described in Section 2 could be applied to (3.1), but there is no guarantee that the reduced system (2.3) will retain the Hamiltonian structure and the geometric properties of (3.1). The PSD method, first proposed in [126] for deterministic systems, constructs a reduced model that is also a Hamiltonian system. A $2n \times 2k$ matrix is called symplectic if it satisfies the condition

$$A^T \mathbb{J}_{2n} A = \mathbb{J}_{2k}. \tag{3.6}$$

For a symplectic matrix A, we can define its symplectic inverse $A^+ = \mathbb{J}_{2k}^T A^T \mathbb{J}_{2n}$. It is an inverse in the sense that $A^+ A = \mathbb{I}_{2k}$. Let ξ be a vector in $\mathbb{R}^{2k}$. Substituting $u = A\xi$ in (3.1) yields a reduced equation

$$\begin{aligned} d_t \xi &= A^+ \mathbb{J}_{2n} \nabla_u H(u)\, dt + \sum_{\nu=1}^{m} A^+ \mathbb{J}_{2n} \nabla_u h_\nu(u) \circ dW^\nu(t) \\ &= \mathbb{J}_{2k} \nabla_\xi H(A\xi)\, dt + \sum_{\nu=1}^{m} \mathbb{J}_{2k} \nabla_\xi h_\nu(A\xi) \circ dW^\nu(t), \end{aligned} \tag{3.7}$$

which is a lower-dimensional stochastic Hamiltonian system with the Hamiltonian functions $\tilde{H}(\xi) = H(A\xi)$ and $\tilde{h}_\nu(\xi) = h_\nu(A\xi)$ for $\nu = 1, \dots, m$. The corresponding initial condition is calculated as $\xi_0 = A^+ u_0$, where u_0 is an initial condition for (3.1). Given a set of empirical data on a Hamiltonian system, the PSD method constructs a symplectic matrix A which best approximates that data in a lower-dimensional subspace. Several algorithms have been proposed to construct A, namely the cotangent lift algorithm, the complex SVD algorithm, the greedy algorithm, and the nonlinear programming algorithm (see [1], [126]). Any of these algorithms can be adapted in the stochastic setting, too, but in this work we will focus only on the cotangent lift algorithm, as it possesses two advantages discussed below. The cotangent lift algorithm constructs a symplectic matrix A which has the special block diagonal structure

$$A = \begin{pmatrix} \Phi & 0 \\ 0 & \Phi \end{pmatrix}, \tag{3.8}$$

where Φ is an $n \times k$ matrix with orthogonal columns, i.e., $\Phi^T \Phi = \mathbb{I}_k$. Suppose snapshots of a solution are given as an $n \times 2r$ matrix Δ of the form

$$\Delta = [q(t_1) \ \dots \ q(t_r) \ p(t_1) \ \dots \ p(t_r)]. \tag{3.9}$$

The SVD of Δ is truncated after the first k largest singular values, similar to (2.2). The matrix Φ is then chosen as $\Phi = U_k$. With the cotangent lift matrix A as in (3.8), the reduced stochastic Hamiltonian system (3.7) can be written as

$$\begin{aligned} d_t \eta = \ & \Phi^T \frac{\partial H}{\partial p}\Big(\Phi\eta, \Phi\chi\Big)\, dt + \sum_{\nu=1}^{m} \Phi^T \frac{\partial h_\nu}{\partial p}\Big(\Phi\eta, \Phi\chi\Big) \circ dW^\nu(t), \\ d_t \chi = & -\Phi^T \frac{\partial H}{\partial q}\Big(\Phi\eta, \Phi\chi\Big)\, dt - \sum_{\nu=1}^{m} \Phi^T \frac{\partial h_\nu}{\partial q}\Big(\Phi\eta, \Phi\chi\Big) \circ dW^\nu(t), \end{aligned} \tag{3.10}$$

where $\xi = (\eta, \chi)$.

The advantage of using the cotangent lift algorithm is twofold. First, if the original system (1.2) is separable, so is the reduced system (3.7): when $H(q,p) = T(p) + V(q)$, then $\tilde{H}(\eta, \chi) = H(\Phi\eta, \Phi\chi) = T(\Phi\chi) + V(\Phi\eta)$; analogously for the Hamiltonians h_ν. This is significant, because separable Hamiltonian systems appear often in practical applications, and many symplectic integrators, such as

the stochastic Störmer-Verlet method, become explicit in that case (see Section 3.4). Second, the PSD method with the cotangent lift symplectic matrix (3.8) preserves also the Lagrange-d'Alembert structure of stochastic forced Hamiltonian systems (see Section 3.3).

3.2 Symplectic Discrete Empirical Interpolation Method

Just like in the case of the POD method, if the drift and diffusion terms in the reduced model (3.7) are complicated nonlinear functions, the PSD method may not bring any computational savings. A technique called the Symplectic Discrete Empirical Interpolation Method (SDEIM), which applies DEIM to approximate the symplectic projection of the nonlinear terms, has been developed in [126]. We will argue that this technique can be adapted also in the stochastic setting. For completeness and for the benefit of the reader, below we briefly outline the main ideas of SDEIM. The further details can be found in [126]. Let the gradient of the Hamiltonian function H be split into the linear and nonlinear parts as

$$\nabla_u H(u) = Lu + a_N(u). \tag{3.11}$$

Similar to (2.7), one can use DEIM to approximate the nonlinear vector term $a_N(u)$. The SDEIM approximation of the drift term in (3.7) then takes the form

$$A^+ \mathbb{J}_{2n} \nabla_u H(u) \approx \mathbb{J}_{2k} \bar{L} \xi + \mathbb{J}_{2k} W g(\xi), \tag{3.12}$$

where

$$\bar{L} = A^T L A, \qquad W = A^T \Psi (P^T \Psi)^{-1}, \qquad g(\xi) = P^T a_N(A\xi). \tag{3.13}$$

Note that the matrices $\bar{L}$ and W are calculated only once at the offline stage, and at the online stage the function $g(\xi)$ evaluates only $\bar{k}$ components of $a_N(A\xi)$. A similar SDEIM approximation can be applied to each of the diffusion terms $A^+ \mathbb{J}_{2n} \nabla_u h_\nu(u)$, thus reducing the computational complexity of the reduced system (3.7). It should, however, be noted that the PSD reduced system with the SDEIM approximation of the drift and diffusion terms is not strictly Hamiltonian anymore. Similar to (3.11), let the gradient of the Hamiltonian functions h_ν be split into the linear and nonlinear parts as

$$\nabla_u h_\nu(u) = L_\nu u + a_{N,\nu}(u) \tag{3.14}$$

for $\nu = 1, \dots, m$. The PSD reduced system (3.7) with the SDEIM approximation takes the form

$$d_t \xi = \mathbb{J}_{2k} \big(\bar{L} \xi + A^T \bar{a}_N(A\xi) \big)\, dt + \sum_{\nu=1}^{m} \mathbb{J}_{2k} \big(\bar{L}_\nu \xi + A^T \bar{a}_{N,\nu}(A\xi) \big) \circ dW^\nu(t), \tag{3.15}$$

where $\bar{a}_N$ and $\bar{a}_{N,\nu}$ denote the DEIM approximations of a_N and $a_{N,\nu}$, respectively. In the following theorem, which is a stochastic generalization of Theorem 5.1 in [126], we show in what sense the SDEIM method approximates the properties of a PSD reduced stochastic Hamiltonian system.

Theorem 3.1. *Let the Hamiltonians of the reduced system* (3.7) *satisfy*

$$\{\tilde{H}, \tilde{h}_\nu\} = (\nabla_\xi \tilde{H})^T \mathbb{J}_{2k} \nabla_\xi \tilde{h}_\nu = 0 \tag{3.16}$$

for $\nu = 1, \ldots, m$, *and let* $\xi(t)$ *be the solution of* (3.15) *with the initial condition* $\xi(0) = \xi_0$. *Then the stochastic differential of the Hamiltonian* $\tilde{H}$ *along* $\xi(t)$, *that is* $E(t) = \tilde{H}(\xi(t))$, *takes the form*

$$d_t E(t) = \gamma(t)\, dt + \sum_{\nu=1}^{m} \lambda_\nu(t) \circ dW^\nu(t), \tag{3.17}$$

with the drift and diffusion terms given by

$$\begin{aligned} \gamma(t) &= \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} A^T \big(\bar{a}_N(A\xi) - a_N(A\xi)\big), \\ \lambda_\nu(t) &= \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} A^T \big(\bar{a}_{N,\nu}(A\xi) - a_{N,\nu}(A\xi)\big), \end{aligned} \tag{3.18}$$

for $\nu = 1, \ldots, m$. *Moreover, upper bounds on the drift and diffusion terms are given by*

$$\begin{aligned} \big|\gamma(t)\big| &\le C \cdot \big\|\nabla_\xi \tilde{H}(\xi)\big\| \cdot \big\|(\mathbb{I} - \Psi\Psi^T) a_N(A\xi)\big\|, \\ \big|\lambda_\nu(t)\big| &\le C_\nu \cdot \big\|\nabla_\xi \tilde{H}(\xi)\big\| \cdot \big\|(\mathbb{I} - \Psi_\nu \Psi_\nu^T) a_{N,\nu}(A\xi)\big\|, \end{aligned} \tag{3.19}$$

for $\nu = 1, \ldots, m$, *where* $C = \|(P^T\Psi)^{-1}\|$ *and* $C_\nu = \|(P_\nu^T \Psi_\nu)^{-1}\|$ *are constants, and the matrices* P, Ψ *and* P_ν, Ψ_ν *define the DEIM approximations of the nonlinear terms* a_N *and* $a_{N,\nu}$, *respectively.*

Proof. We calculate the stochastic differential $d_t E(t)$ as

$$\begin{aligned} d_t E(t) &= \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} \big(\bar{L}\xi + A^T \bar{a}_N(A\xi)\big)\, dt + \sum_{\nu=1}^{m} \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} \big(\bar{L}_\nu \xi + A^T \bar{a}_{N,\nu}(A\xi)\big) \circ dW^\nu(t) \\ &= \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} \big(\bar{L}\xi + A^T \bar{a}_N(A\xi) - \nabla_\xi \tilde{H}(\xi)\big)\, dt \\ &\qquad + \sum_{\nu=1}^{m} \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} \big(\bar{L}_\nu \xi + A^T \bar{a}_{N,\nu}(A\xi) - \nabla_\xi \tilde{h}_\nu(\xi)\big) \circ dW^\nu(t) \\ &= \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} A^T \big(\bar{a}_N(A\xi) - a_N(A\xi)\big)\, dt \\ &\qquad + \sum_{\nu=1}^{m} \big(\nabla_\xi \tilde{H}(\xi)\big)^T \mathbb{J}_{2k} A^T \big(\bar{a}_{N,\nu}(A\xi) - a_{N,\nu}(A\xi)\big) \circ dW^\nu(t) \\ &= \gamma(t)\, dt + \sum_{\nu=1}^{m} \lambda_\nu(t) \circ dW^\nu(t), \end{aligned} \tag{3.20}$$

where in the first equality we used the rules of Stratonovich calculus and substituted (3.15), and in the second equality we used $\{\tilde{H}, \tilde{H}\} = 0$ and $\{\tilde{H}, \tilde{h}_\nu\} = 0$. The upper bounds (3.19) are obtained like in the proof of Theorem 5.1 in [126].

□

Even though the system (3.15) is not necessarily Hamiltonian, Theorem 3.1 shows that $|\gamma(t)| \to 0$ and $|\lambda_\nu(t)| \to 0$ when $\big\|(\mathbb{I} - \Psi\Psi^T) a_N(A\xi)\big\| \to 0$ and $\big\|(\mathbb{I} - \Psi_\nu \Psi_\nu^T) a_{N,\nu}(A\xi)\big\| \to 0$.

3.3 Proper Symplectic Decomposition for stochastic forced Hamiltonian systems

The PSD method described in Section 3.1 can also be applied to Hamiltonian systems subject to external forcing. Stochastic forced Hamiltonian systems take the form

$$\begin{aligned} d_t q &= \frac{\partial H}{\partial p} dt + \sum_{\nu=1}^{m} \frac{\partial h_\nu}{\partial p} \circ dW^\nu(t), \\ d_t p &= \left[-\frac{\partial H}{\partial q} + F(q,p) \right] dt + \sum_{\nu=1}^{m} \left[-\frac{\partial h_\nu}{\partial q} + f_\nu(q,p) \right] \circ dW^\nu(t), \end{aligned} \tag{3.21}$$

where $H = H(q,p)$ and $h_\nu = h_\nu(q,p)$ for $\nu = 1, \dots, m$ are the Hamiltonian functions, $F = F(q,p)$ and $f_\nu = f_\nu(q,p)$ are the forcing terms, and $W(t) = (W^1(t), \dots, W^m(t))$ is the standard m-dimensional Wiener process. Applications of such systems arise in many models in physics, chemistry, and biology. Particular examples include molecular dynamics (see, e.g., [17], [91], [101], [145]), dissipative particle dynamics (see, e.g., [139]), investigations of the dispersion of passive tracers in turbulent flows (see, e.g., [142], [152]), energy localization in thermal equilibrium (see, e.g., [136]), lattice dynamics in strongly anharmonic crystals (see, e.g., [72]), description of noise induced transport in stochastic ratchets (see, e.g., [103]), and collisional kinetic plasmas ([94], [97], [147], [154]). While their stochastic flow is not symplectic in general, stochastic forced Hamiltonian systems have an underlying variational structure, that is, their solutions satisfy the stochastic Lagrange-d'Alembert principle (see [97]). It is therefore beneficial to preserve this variational structure both in time integration and in deriving reduced models. Stochastic Lagrange-d'Alembert schemes for time integration of (3.21), first proposed in [97], demonstrate better accuracy and stability properties in long-time simulations than non-geometric stochastic methods.

It was shown in [124] and [125] that the PSD method with the cotangent lift algorithm preserves the Lagrange-d'Alembert structure of deterministic forced Hamiltonian systems (see also [3], [79]). It is straightforward to verify that this also holds for stochastic forced Hamiltonian systems. Indeed, substituting $q = \Phi\eta$ and $p = \Phi\chi$ in (3.21) yields, similar to (3.10), a reduced system of the form

$$\begin{aligned} d_t \eta &= \frac{\partial \tilde{H}}{\partial \chi}\Big(\eta, \chi\Big)\, dt + \sum_{\nu=1}^{m} \frac{\partial \tilde{h}_\nu}{\partial \chi}\Big(\eta, \chi\Big) \circ dW^\nu(t), \\ d_t \chi &= \left[-\frac{\partial \tilde{H}}{\partial \eta}\Big(\eta, \chi\Big) + \tilde{F}(\eta, \chi) \right] dt + \sum_{\nu=1}^{m} \left[-\frac{\partial \tilde{h}_\nu}{\partial \eta}\Big(\eta, \chi\Big) + \tilde{f}_\nu(\eta, \chi) \right] \circ dW^\nu(t), \end{aligned} \tag{3.22}$$

with

$$\begin{aligned} \tilde{H}(\eta,\chi) &= H(\Phi\eta, \Phi\chi), & \tilde{F}(\eta,\chi) &= \Phi^T F(\Phi\eta, \Phi\chi), \\ \tilde{h}_\nu(\eta,\chi) &= h_\nu(\Phi\eta, \Phi\chi), & \tilde{f}_\nu(\eta,\chi) &= \Phi^T f_\nu(\Phi\eta, \Phi\chi), \end{aligned} \tag{3.23}$$

that is, also a stochastic forced Hamiltonian system. The matrix Φ can be constructed from empirical data as described in Section 3.1.

3.4 Time integration

The full (3.1) and reduced (3.7) models can be solved numerically using any general purpose stochastic scheme mentioned in Section 2.3. However, since these models are Hamiltonian, it is advisable to integrate them using structure-preserving methods. Stochastic symplectic integrators, similar to their deterministic counterparts, preserve the symplecticity of the Hamiltonian flow and demonstrate good energy behavior in long-time simulations (see [6], [7], [8], [9], [27], [36], [38], [47], [50], [52], [53], [62], [81], [83], [85], [86], [97], [110], [111], [115], [116], [117], [150], [160], [161], [163], [166] and the references therein). In this work we will focus on two stochastic symplectic Runge-Kutta methods, namely the stochastic midpoint method (2.10), which is symplectic when applied to a Hamiltonian system, and the stochastic Störmer-Verlet method ([83], [97], [111]). The latter method for (1.2) takes the form

$$\begin{aligned}
\Lambda &= p_i - \frac{1}{2}\frac{\partial H}{\partial q}(q_i,\Lambda)\Delta t - \sum_{\nu=1}^{m}\frac{1}{2}\frac{\partial h_\nu}{\partial q}(q_i,\Lambda)\Delta W^\nu,\\
q_{i+1} &= q_i + \frac{1}{2}\frac{\partial H}{\partial p}(q_i,\Lambda)\Delta t + \frac{1}{2}\frac{\partial H}{\partial p}(q_{i+1},\Lambda)\Delta t + \sum_{\nu=1}^{m}\frac{1}{2}\frac{\partial h_\nu}{\partial p}(q_i,\Lambda)\Delta W^\nu + \sum_{\nu=1}^{m}\frac{1}{2}\frac{\partial h_\nu}{\partial p}(q_{i+1},\Lambda)\Delta W^\nu,\\
p_{i+1} &= \Lambda - \frac{1}{2}\frac{\partial H}{\partial q}(q_{i+1},\Lambda)\Delta t - \sum_{\nu=1}^{m}\frac{1}{2}\frac{\partial h_\nu}{\partial q}(q_{i+1},\Lambda)\Delta W^\nu,
\end{aligned} \tag{3.24}$$

where Δt denotes the time step, $\Delta W = (\Delta W^1, \dots, \Delta W^m)$ are the increments of the Wiener process, and Λ is the internal momentum stage. The stochastic Störmer-Verlet method is strongly convergent of order 1 for systems driven by a commutative noise, and of order $1/2$ in the non-commutative case. In case the stochastic Hamiltonian system (1.2) is separable, the numerical scheme (3.24) becomes explicit (see [83]).

Similarly, it is advisable that the full (3.21) and reduced (3.22) models are solved using structure-preserving methods. Stochastic Lagrange-d'Alembert integrators (see [97]) generalize the notion of symplectic integrators by preserving the underlying variational structure of forced Hamiltonian systems. As shown in [97], the stochastic Störmer-Verlet method, when applied to a forced system, is a Lagrange-d'Alembert integrator, and takes the form

$$\begin{aligned}
\Lambda &= p_i + \frac{1}{2}\Big[-\frac{\partial H}{\partial q}(q_i,\Lambda) + F(q_i,\Lambda)\Big]\Delta t + \frac{1}{2}\sum_{\nu=1}^{m}\Big[-\frac{\partial h_\nu}{\partial q}(q_i,\Lambda) + f_\nu(q_i,\Lambda)\Big]\Delta W^\nu,\\
q_{i+1} &= q_i + \frac{1}{2}\frac{\partial H}{\partial p}(q_i,\Lambda)\Delta t + \frac{1}{2}\frac{\partial H}{\partial p}(q_{i+1},\Lambda)\Delta t + \sum_{\nu=1}^{m}\frac{1}{2}\frac{\partial h_\nu}{\partial p}(q_i,\Lambda)\Delta W^\nu + \sum_{\nu=1}^{m}\frac{1}{2}\frac{\partial h_\nu}{\partial p}(q_{i+1},\Lambda)\Delta W^\nu,\\
p_{i+1} &= \Lambda + \frac{1}{2}\Big[-\frac{\partial H}{\partial q}(q_{i+1},\Lambda) + F(q_{i+1},\Lambda)\Big]\Delta t + \frac{1}{2}\sum_{\nu=1}^{m}\Big[-\frac{\partial h_\nu}{\partial q}(q_{i+1},\Lambda) + f_\nu(q_{i+1},\Lambda)\Big]\Delta W^\nu.
\end{aligned} \tag{3.25}$$

If the Hamiltonians in (3.21) are separable, the second equation in (3.25) becomes explicit. If in addition the forcing terms F and f_ν have special forms, then further improvements in efficiency are possible. For instance, if the forcing terms depend linearly on p, as is often the case in practical applications, then the first equation is a linear equation for Λ, and can be solved using linear solvers. In case the forcing terms are independent of p altogether, then the whole method becomes fully explicit.

4 Model reduction for a large number of Monte Carlo runs

In Sections 2 and 3 we have demonstrated how model reduction can be used to efficiently solve high dimensional SDEs for one given realization of the Wiener process. In this section we will argue that model reduction also allows an efficient approach to another typical computational issue arising in numerical simulations of SDEs. In order to compute the statistical properties of the solution of (1.1) or (1.2), one typically needs a very large number of Monte Carlo runs, that is, one needs to simulate the solution for a very large number of independent realizations of the Wiener process. We will show that when empirical data on the considered system is available, it can be used to construct a reduced model which requires less computational effort to yield the desired results.

4.1 Proper Orthogonal Decomposition

Suppose we are interested in solving the stochastic differential equation

$$d_t X = \Gamma(X)\, dt + B(X) \circ dW(t), \tag{4.1}$$

for an N-dimensional stochastic process $X(t)$, where $\Gamma : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is the drift function, $B : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ is the diffusion function, and $W(t)$ is the standard one-dimensional Wiener process. For convenience and clarity we restrict ourselves to a one-dimensional noise; the generalization to a multidimensional Wiener process is straightforward. Suppose we would like to compute the statistical properties of the stochastic process $X(t)$, e.g., its mean, variance, etc. For this we need to solve (4.1) for a large number M of independent realizations of the Wiener process. Unlike in the situation considered in Section 2.1, here we do not assume that the dimension N is very high. The main computational cost comes from the high value of M. Therefore, in order to apply model reduction, let us turn this problem into the problem discussed in Section 2. Let us consider M stochastic processes $X_1, \ldots, X_M$ satisfying the system of stochastic differential equations

$$\begin{aligned} d_t X_1 &= \Gamma(X_1)\, dt + B(X_1) \circ dW^1(t), \\ &\vdots \\ d_t X_M &= \Gamma(X_M)\, dt + B(X_M) \circ dW^M(t), \end{aligned} \tag{4.2}$$

where $W^1(t), \ldots, W^M(t)$ are the components of the standard M-dimensional Wiener process. Note that the equations in (4.2) are decoupled from each other, and each equation is driven by an independent Wiener process $W^\nu(t)$. Therefore, $X_1, \ldots, X_M$ are independent identically distributed (i.i.d.) stochastic processes, each with the same probability density function as the original stochastic process X. In this sense (4.2) is equivalent to (4.1). The advantage is that instead of considering M realizations of the N-dimensional stochastic process X, one can consider one realization of the NM-dimensional process $(X_1, \ldots, X_M)$. The value of any functional of X can then be approximated using the law of large numbers, e.g., the expected value $\mathbb{E}[X] \approx (X_1 + \ldots + X_M)/M$. Note that the system (4.2) has the form of (1.1) with $u = (X_1, \ldots, X_M)$, $n = NM$, $m = M$, and the drift and diffusion functions given by

$$a(u) = \begin{pmatrix} \Gamma(X_1) \\ \vdots \\ \vdots \\ \vdots \\ \Gamma(X_M) \end{pmatrix}, \qquad b_1(u) = \begin{pmatrix} B(X_1) \\ 0 \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \qquad b_2(u) = \begin{pmatrix} 0 \\ B(X_2) \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad b_M(u) = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ B(X_M) \end{pmatrix}. \tag{4.3}$$

With this setting the POD method described in Section 2 can now be directly applied, and the corresponding reduced model is given by (2.3). Since the vectors in (4.3) have a sparse structure, it pays off to split the matrix U_k into M blocks of size $N \times k$ each, that is,

$$U_k = \begin{pmatrix} U^{(1)} \\ \vdots \\ U^{(M)} \end{pmatrix}. \tag{4.4}$$

Then the drift and diffusion terms in (2.3) can be evaluated as

$$U_k^T a(U_k\xi) = \sum_{\nu=1}^{M} \left(U^{(\nu)}\right)^T \Gamma\left(U^{(\nu)}\xi\right) \quad \text{and} \quad U_k^T b_\nu(U_k\xi) = \left(U^{(\nu)}\right)^T B\left(U^{(\nu)}\xi\right) \quad \text{for } \nu = 1, \dots, M. \tag{4.5}$$

Remark. A related idea appears in [26], where a variance reduction method using the reduced basis paradigm is proposed. Variance reduction methods are a set of techniques to reduce the statistical error appearing in the Monte-Carlo estimation of the output expectation of a random variable. One of such techniques, the control variate method, involves introducing a correlated auxiliary variable (control variate) to reduce the variance of the estimator, leading to more accurate estimations of the expected value of a random variable. In [26] model reduction is used for the efficient calculation of control variates for a certain time-independent functional of the solution of a given parameter-dependent SDE, rather than for constructing a lower-dimensional stochastic system like (2.3). The reduced bases are constructed in the space of functionals of the solution of the underlying SDE, rather than in the space of solutions themselves; and the data used to construct those bases are the values of the functionals for a selected set of parameters, rather than the snapshots of the trajectories of the underlying stochastic system like (2.1). The approach outlined in this section is therefore conceptually different from the strategy employed in [26].

4.2 Proper Symplectic Decomposition

Let us now consider the problem of solving the stochastic Hamiltonian system

$$\begin{aligned} d_t Q &= \frac{\partial \bar{H}}{\partial P}(Q,P)\,dt + \frac{\partial \bar{h}}{\partial P}(Q,P) \circ dW(t), \\ d_t P &= -\frac{\partial \bar{H}}{\partial Q}(Q,P)\,dt - \frac{\partial \bar{h}}{\partial Q}(Q,P) \circ dW(t), \end{aligned} \tag{4.6}$$

for N-dimensional stochastic processes $Q(t)$ and $P(t)$, where $\bar{H} : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}$ and $\bar{h} : \mathbb{R}^N \times \mathbb{R}^N \longrightarrow \mathbb{R}$ are the Hamiltonian functions. Suppose we would like to solve (4.6) for a large number M of independent realizations of the Wiener process $W(t)$. In contrast to the scenario explored in Section 3.1, we do not assume that the dimension $2N$ of the system is very high. Rather, the main computational expense arises due to the large number M of Monte Carlo runs. Therefore, in a spirit similar to Section 4.1, let us consider $2M$ stochastic processes $Q_1, P_1, \ldots, Q_M, P_M$, with each pair (Q_ν, P_ν) satisfying the stochastic differential system

$$\begin{aligned} d_t Q_\nu &= \frac{\partial \bar{H}}{\partial P}(Q_\nu, P_\nu)\, dt + \frac{\partial \bar{h}}{\partial P}(Q_\nu, P_\nu) \circ dW^\nu(t), \\ d_t P_\nu &= -\frac{\partial \bar{H}}{\partial Q}(Q_\nu, P_\nu)\, dt - \frac{\partial \bar{h}}{\partial Q}(Q_\nu, P_\nu) \circ dW^\nu(t), \end{aligned} \tag{4.7}$$

for $\nu = 1, \ldots, M$, where $W^1(t), \ldots, W^M(t)$ are the components of the standard M-dimensional Wiener process. Note that the systems (4.7) are decoupled from each other for different values of ν, and each system is driven by an independent Wiener process $W^\nu(t)$. Therefore, the pairs (Q_ν, P_ν) for $\nu = 1, \ldots, M$ are independent identically distributed (i.i.d.) stochastic processes, each with the same probability density function as the original stochastic process (Q, P). In that sense the system (4.7) is equivalent to (4.6). Note that the system (4.7) has the form of (1.2) with $q = (Q_1, \ldots, Q_M)$, $p = (P_1, \ldots, P_M)$, $n = NM$, $m = M$, and the Hamiltonian functions given by

$$H(q,p) = \sum_{\nu=1}^{M} \bar{H}(Q_\nu, P_\nu) \qquad \text{and} \qquad h_\nu(q,p) = \bar{h}(Q_\nu, P_\nu) \qquad \text{for } \nu = 1, \ldots, M. \tag{4.8}$$

With this setting the PSD method described in Section 3 can now be applied, and the corresponding reduced stochastic Hamiltonian system is given by (3.10) for the cotangent lift algorithm. Similar to (4.3), the diffusion terms in (4.7) have a sparse structure, therefore it pays off to split the matrix Φ into M blocks of size $N \times k$

$$\Phi = \begin{pmatrix} \Phi^{(1)} \\ \vdots \\ \Phi^{(M)} \end{pmatrix}. \tag{4.9}$$

Then the drift and the diffusion terms of the reduced model (3.10) can be expressed as

$$\begin{aligned} \Phi^T \frac{\partial H}{\partial p}\Big(\Phi\eta, \Phi\chi\Big) &= \sum_{\nu=1}^{M} \big(\Phi^{(\nu)}\big)^T \frac{\partial \bar{H}}{\partial P}\Big(\Phi^{(\nu)}\eta, \Phi^{(\nu)}\chi\Big), \\ -\Phi^T \frac{\partial H}{\partial q}\Big(\Phi\eta, \Phi\chi\Big) &= -\sum_{\nu=1}^{M} \big(\Phi^{(\nu)}\big)^T \frac{\partial \bar{H}}{\partial Q}\Big(\Phi^{(\nu)}\eta, \Phi^{(\nu)}\chi\Big), \end{aligned} \tag{4.10}$$

and, for $\nu = 1, \ldots, M$,

$$\begin{aligned} \Phi^T \frac{\partial h_\nu}{\partial p}\Big(\Phi\eta, \Phi\chi\Big) &= \big(\Phi^{(\nu)}\big)^T \frac{\partial \bar{h}}{\partial P}\Big(\Phi^{(\nu)}\eta, \Phi^{(\nu)}\chi\Big), \\ -\Phi^T \frac{\partial h_\nu}{\partial q}\Big(\Phi\eta, \Phi\chi\Big) &= -\big(\Phi^{(\nu)}\big)^T \frac{\partial \bar{h}}{\partial Q}\Big(\Phi^{(\nu)}\eta, \Phi^{(\nu)}\chi\Big). \end{aligned} \tag{4.11}$$

Remark. An analogous strategy can be used for the stochastic forced system (3.21), where similar steps as above have to be applied also to the forcing terms. For brevity, we omit presenting detailed formulas.

5 Numerical experiments

We present the results of three numerical experiments that we have carried out to validate the methods proposed in Sections 2, 3, and 4. In the first experiment we have applied model reduction to the semi-discretization of a stochastic Nonlinear Schrödinger Equation, whereas in the second and third experiments model reduction has been used to reduce the computational cost of the simulations of the Kubo oscillator, unforced and forced, respectively. All computations have been performed in the Julia programming language with the help of the *GeometricIntegrators.jl* library (see [96]).

5.1 Stochastic Nonlinear Schrödinger Equation

The Nonlinear Schrödinger Equation (NLS) is a well-known nonlinear partial differential equation (PDE) with a broad spectrum of applications, ranging from wave propagation in nonlinear media to nonlinear optics, molecular biology, quantum physics, quantum chemistry, and plasma physics (see [149], [165] and the references therein). Model reduction for semi-discretizations of NLS is considered in, e.g., [1], [93]. Various stochastic perturbations of NLS have been proposed in order to, e.g., take into account inhomogeneities of the media or noisy sources (see [10], [14], [15], [51], [59], [60], [61], [66], [67], [134]). The stochastic NLS equation of the form

$$id_t\psi + \left(\frac{\partial^2\psi}{\partial x^2} + \epsilon|\psi|^2\psi\right)dt + \beta\psi \circ dW(t) = 0, \tag{5.1}$$

for a complex-valued function $\psi = \psi(x,t)$, where ϵ, β are real parameters, and i denotes the imaginary unit, has been proposed in [66] as a model for the propagation of optical pulses down a nonideal anomalously dispersive optical fiber, with the multiplicative noise term describing the effects of local density fluctuations in the fiber material. A similar equation has also been considered as a model of energy transfer in a monolayer molecular aggregate in the presence of thermal fluctuations (see [14], [15], [134]). Equation (5.1) reduces to the deterministic NLS equation for $\beta = 0$. It can be verified by a straightforward calculation that if $\psi_D(x,t)$ is a solution of the deterministic NLS equation, then $\psi(x,t) = \exp(i\beta W(t))\psi_D(x,t)$ is a solution of (5.1) (see [66]). This in particular means that Equation (5.1), similar to its deterministic counterpart, also possesses solitonic solutions. By considering the real and imaginary parts of ψ, Equation (5.1) can be rewritten as a system of coupled stochastic PDEs,

$$d_tq = -\left[\frac{\partial^2 p}{\partial x^2} + \epsilon(q^2+p^2)p\right]dt - \beta p \circ dW(t), \qquad d_tp = \left[\frac{\partial^2 q}{\partial x^2} + \epsilon(q^2+p^2)q\right]dt + \beta q \circ dW(t), \tag{5.2}$$

for real-valued functions $q = q(x,t)$ and $p = p(x,t)$, where $\psi = q + ip$. Equation (5.2) has the form of a stochastic Hamiltonian PDE, that is,

$$d_t q = \frac{\delta \mathcal{H}_0}{\delta p}\, dt + \frac{\delta \mathcal{H}_1}{\delta p} \circ dW(t), \qquad d_t p = -\frac{\delta \mathcal{H}_0}{\delta q}\, dt - \frac{\delta \mathcal{H}_1}{\delta q} \circ dW(t), \tag{5.3}$$

with the Hamiltonian functionals given by

$$\mathcal{H}_0[q,p] = \int \left[\frac{1}{2}\left(\frac{\partial q}{\partial x}\right)^2 + \frac{1}{2}\left(\frac{\partial p}{\partial x}\right)^2 - \frac{\epsilon}{4}(q^2+p^2)^2 \right] dx, \qquad \mathcal{H}_1[q,p] = -\frac{\beta}{2}\int (q^2+p^2)\, dx. \tag{5.4}$$

5.1.1 Semi-discretization

Suppose we would like to solve (5.2) on the bounded spatial domain $[0, X_{max}]$ with periodic boundary conditions. Let us introduce a uniform spatial mesh consisting of the N points $x_j = (j-1)\Delta x$ for $j = 1, \ldots, N$, where $\Delta x = X_{max}/N$ is the mesh size, and let us denote $q^j(t) = q(x_j, t)$ and $p^j(t) = p(x_j, t)$. Using central differences to approximate the second derivatives in (5.2), we obtain the system of $2N$ stochastic differential equations

$$\begin{aligned} d_t q^j &= -\left[\frac{p^{j+1} - 2p^j + p^{j-1}}{\Delta x^2} + \epsilon\Big((q^j)^2 + (p^j)^2\Big) p^j \right] dt - \beta p^j \circ dW(t), \\ d_t p^j &= \left[\frac{q^{j+1} - 2q^j + q^{j-1}}{\Delta x^2} + \epsilon\Big((q^j)^2 + (p^j)^2\Big) q^j \right] dt + \beta q^j \circ dW(t), \end{aligned} \tag{5.5}$$

for $j = 1, \ldots, N$, with $q^0 \equiv q^N$, $q^{N+1} \equiv q^1$, $p^0 \equiv p^N$, and $p^{N+1} \equiv p^1$. Equation (5.5) is a stochastic Hamiltonian system (1.2) with the Hamiltonians

$$H = \sum_{j=1}^{N} \left[\frac{1}{2}\left(\frac{q^{j+1} - q^j}{\Delta x}\right)^2 + \frac{1}{2}\left(\frac{p^{j+1} - p^j}{\Delta x}\right)^2 - \frac{\epsilon}{4}\Big((q^j)^2 + (p^j)^2\Big)^2 \right], \qquad h = -\frac{\beta}{2}\sum_{j=1}^{N}\Big((q^j)^2 + (p^j)^2\Big), \tag{5.6}$$

which are related to the discretizations of the Hamiltonian functionals (5.4). Since the noise is one-dimensional, for simplicity we write $h \equiv h_1$. One can check that the condition (3.4) is satisfied, therefore the Hamiltonian H is almost surely preserved on solutions to (5.5). The system (5.5) can be integrated in time using general purpose stochastic methods (see Section 2.3) or stochastic symplectic schemes (see Section 3.4).

5.1.2 Empirical data

Suppose we have the following computational problem: we would like to scan the domains of β and ϵ, that is, compute the numerical solution of (5.5) for a single realization of the Wiener process for a large number of values of β and ϵ. Given that in practical applications the system (5.5) is high-dimensional, this task is computationally intensive. Model reduction can alleviate this substantial computational cost. One can carry out full-scale computations only for a selected number of values of β and ϵ. These data can then be used to identify reduced models, as described in Sections 2 and 3.

The lower-dimensional equations (2.3) or (3.10) can then be solved more efficiently for other values of β and ϵ, thus reducing the overall computational cost. Let us consider the initial conditions

$$q^j(0) = \sqrt{2}\,\mathrm{sech}(x_j - x_c)\cos\frac{c}{2}(x_j - x_c), \qquad p^j(0) = \sqrt{2}\,\mathrm{sech}(x_j - x_c)\sin\frac{c}{2}(x_j - x_c), \tag{5.7}$$

for $j = 1, \ldots, N$. These initial conditions correspond to a soliton for (5.1) centered at $x = x_c$ and propagating with the speed c in the case when $\epsilon = 1$ (see [149], [165]). For our experiment, we calculated the numerical solution to the full model (5.5) for 24 pairs of (arbitrarily selected) values of the parameters (β, ϵ), with β and ϵ taking the following values:

$$\beta = 0.12,\ 0.14,\ 0.16,\ 0.18, \qquad \epsilon = 0.95,\ 0.97,\ 0.99,\ 1.01,\ 1.03,\ 1.05. \tag{5.8}$$

Computations were carried out for $N = 256$ mesh points using the stochastic midpoint method (2.10) over the time interval $0 \le t \le 200$ with the time step $\Delta t = 0.01$. The remaining parameters were $X_{max} = 60$, $\Delta x \approx 0.2344$, $c = 1$, and $x_c = 30$. The same sample path of the Wiener process was used for all simulations (i.e., the same seed for the random number generator was used when calculating a sample path of $W(t)$). The generated data for all values of β and ϵ were put together and used to form the snapshot matrices (2.1) and (3.9). For instance, for the POD snapshot matrix (2.1) we used

$$\Delta = [u(t_1;\beta_1,\epsilon_1)\ u(t_2;\beta_1,\epsilon_1)\ u(t_3;\beta_1,\epsilon_1)\ \ldots\ u(t_1;\beta_1,\epsilon_2)\ u(t_2;\beta_1,\epsilon_2)\ u(t_3;\beta_1,\epsilon_2)\ \ldots]. \tag{5.9}$$

Then, following the description of the methods in Sections 2 and 3, reduced models (2.3) and (3.7) were derived. Note that the drift term in (5.5) has a nonlinear part, therefore its DEIM approximation (2.7) was also constructed. The decay of the singular values for the POD and PSD reductions, and for the DEIM approximation of the nonlinear term is depicted in Figure 5.1. In general, a sufficiently fast decay of the singular values indicates that it is possible to obtain an accurate low-rank approximation of the snapshot matrix (2.2) for a small value of k, which is a necessary condition for the success of reduced model simulations. Some problems are more amenable to model reduction than others (see, e.g., [31], [120], [130]). The decay in Figure 5.1 is not particularly fast (compare with Figure 5.10 and Figure 5.13), which means that perhaps one should not expect a very significant reduction of the dimension. This is typical behavior for wave-like phenomena and transport problems (see [73], [120]), for which more advanced model reduction techniques have been developed, such as online adaptive methods ([39], [123]), shifted PODs ([137]), or nonlinear manifold reduction methods ([138]). Nevertheless, the main purpose of our experiment is to compare the performance of POD and PSD reductions for stochastic systems, even if the level of reduction is not high. In addition, as described below in Section 5.1.3, in this particular case it is still possible to obtain computational advantage over full-model simulations. For other applications of POD and PSD to deterministic systems possessing wave-like solutions see, e.g., [1], [2], [3], [93], [126].

5.1.3 Reduced model simulations

The reduced models obtained in Section 5.1.2 from the empirical data can now be solved for any other desired values of the parameters β and ϵ. To test the accuracy of the considered model

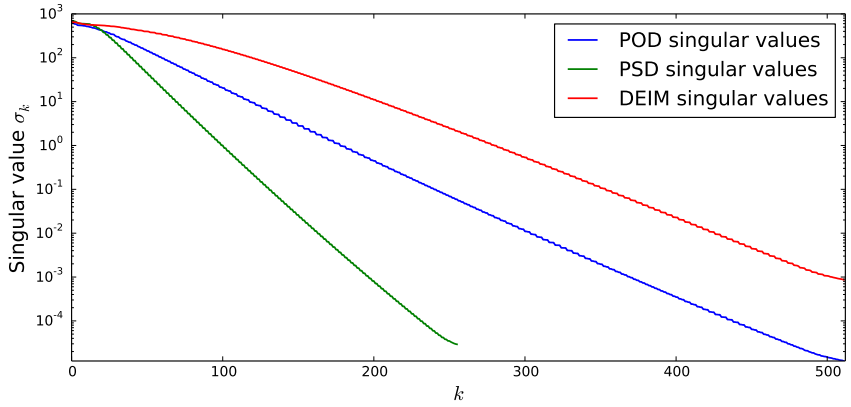

Figure 5.1: The decay of the singular values for the POD and PSD reductions, and for the DEIM approximation of the nonlinear term for the empirical data ensemble for the stochastic NLS equation.

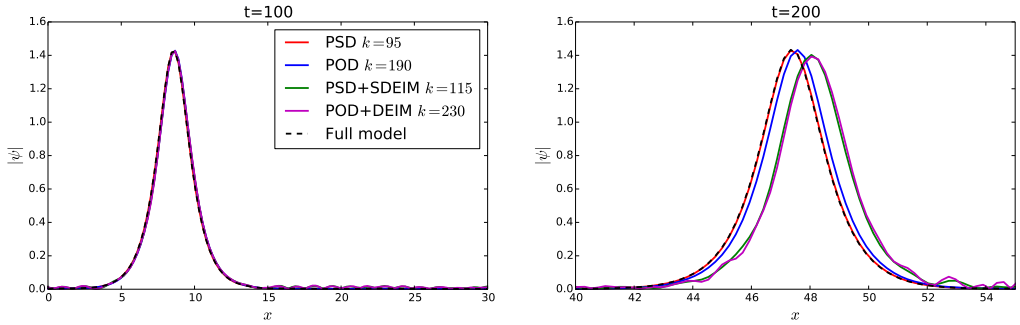

Figure 5.2: The solution $|\psi(x,t)| = |q(x,t) + ip(x,t)|$ of the stochastic NLS equation with the parameters $\beta = 0.15$ and $\epsilon = 1$ at times $t = 100$ (*Left*) and $t = 200$ (*Right*) obtained with the help of the reduced models integrated with the stochastic midpoint method using the time step $\Delta t = 0.01$. While at $t = 100$ all simulations resolve the soliton relatively well, at $t = 200$ the PSD models yield a more accurate solution than the POD models of the same dimension. The POD+DEIM and PSD+SDEIM simulations capture the propagation of the soliton, but create some spurious oscillations in its tail.

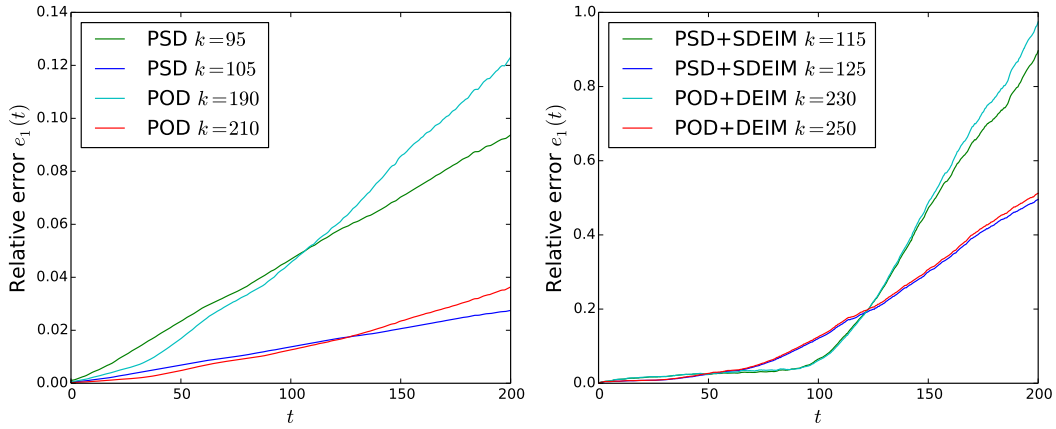

Figure 5.3: The relative error $e_1(t)$ as a function of time for several example reduced model simulations of the stochastic NLS equation using the stochastic midpoint method with the time step $\Delta t = 0.01$.

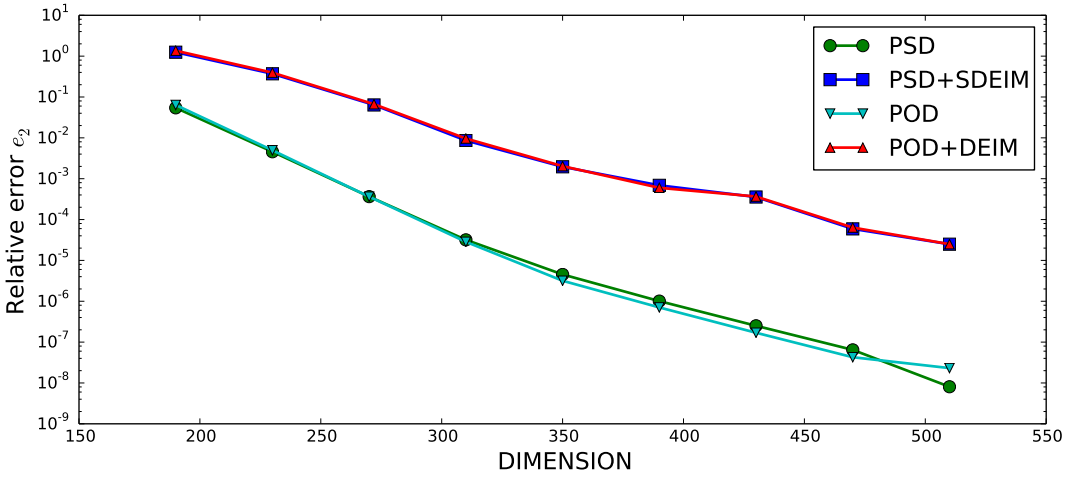

Figure 5.4: The relative error e_2 for the reduced model simulations of the stochastic NLS equation using the stochastic midpoint method with the time step $\Delta t = 0.01$ is depicted as a function of the dimension of the reduced system. Recall that the dimension of the reduced model is equal to k for the POD methods, and $2k$ for the PSD methods.

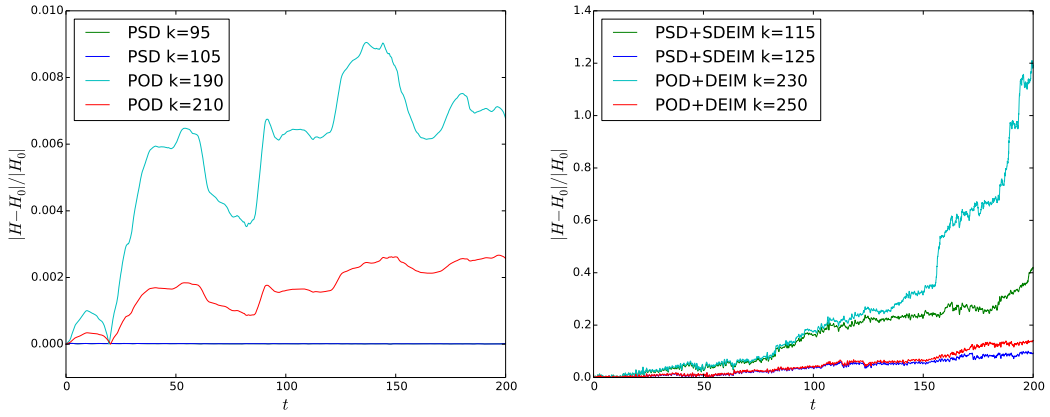

Figure 5.5: The relative error of the Hamiltonian as a function of time for several example reduced model simulations of the stochastic NLS equation using the stochastic midpoint method with the time step $\Delta t = 0.01$ is depicted, where H_0 denotes the initial value of the Hamiltonian. The PSD simulations preserve the Hamiltonian nearly exactly, in contrast to the POD simulations. Although the PSD reduced models in combination with the SDEIM approximation do not show such a good behavior, they still preserve the Hamiltonian better than the POD models of the same dimension in combination with the DEIM approximation. Note that the plots for the PSD method with $k = 95$ and $k = 105$ overlap very closely and are therefore indistinguishable.

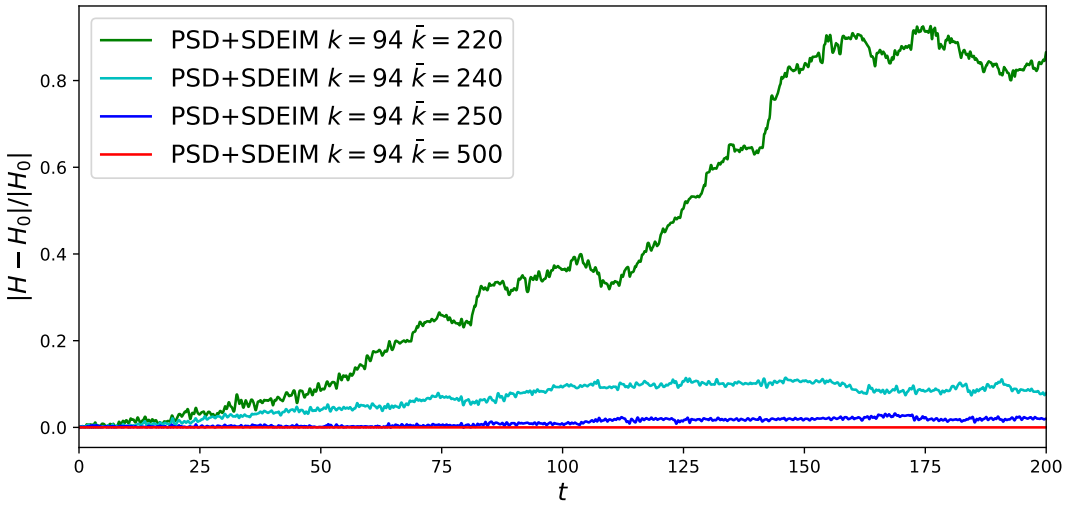

Figure 5.6: The relative error of the Hamiltonian as a function of time for several example PSD+SDEIM reduced model simulations of the stochastic NLS equation using the stochastic midpoint method with the time step $\Delta t = 0.01$ is depicted, where H_0 denotes the initial value of the Hamiltonian. A fixed dimension $2k = 188$ of the reduced system is used, but the number $\bar{k}$ of the calculated components of the nonlinear term in the SDEIM approximation is varied. We can see that the larger the number $\bar{k}$ is, the less the numerical evolution of the Hamiltonian deviates from the initial value, in line with the prediction based on Theorem 3.1.

reduction methods, we have compared the results of the reduced model simulations to a full-scale reference solution for a single arbitrary choice of the values of the parameters. The reference solution for $\beta = 0.15$ and $\epsilon = 1$ was calculated on the time interval $0 \le t \le 200$ in the same way as the empirical data in Section 5.1.2. Note that for this choice of ϵ the reference solution is a soliton. The reduced models were solved numerically on the same time interval using the stochastic midpoint, $R2$, and Heun methods. Note that when applied to a Hamiltonian system, the stochastic midpoint method is a symplectic integrator, while the $R2$ and Heun methods are not. The approximations of the nonlinear term were carried out with $\bar{k} = k$ for the POD+DEIM simulations (i.e., the simulations of the POD reduced models combined with the DEIM approximation), and with $\bar{k} = 2k$ for the PSD+SDEIM simulations (i.e., the simulations of the PSD reduced models combined with the SDEIM approximation), where k is the number of the modes kept in the SVD decompositions of (2.1) and (3.9), respectively. This way the number of the calculated components of the nonlinear term was commensurate with the dimension of the reduced system. The choice of k is a compromise between the speed and the accuracy: the smaller k the faster the computation, but also the larger the projection error. In practice, one may choose k based on the initial value of the error (5.10), i.e., the value of the projection error for the initial condition. For instance, in our experiment, the initial relative error for the POD simulations was equal to $5 \cdot 10^{-4}$ for $k = 190$, and $2.2 \cdot 10^{-4}$ for $k = 210$ (see also Figure 5.3).

The simulations using the stochastic midpoint method were carried out with the time step $\Delta t = 0.01$. The numerical solution for $|\psi(x,t)|$ obtained from the reduced models at times $t = 100$ and $t = 200$ is compared against the reference solution in Figure 5.2. We can see that while at $t = 100$ all simulations resolve the soliton relatively well, after a longer time the PSD models yield a more accurate solution than the POD models of the same dimension. We also see that both the POD+DEIM and PSD+SDEIM simulations capture the propagation of the soliton, but create some spurious oscillations in its tail. As a measure of accuracy of the reduced systems at time t we take the relative $L^2([0,60])$ error, that is,

$$e_1(t) = \frac{\|\psi(\cdot,t) - \psi_{\mathrm{ref}}(\cdot,t)\|_1}{\|\psi_{\mathrm{ref}}(\cdot,t)\|_1}, \tag{5.10}$$

where ψ_{ref} is the reference full-model solution, as described above, and $\|\psi(\cdot,t)\|_1 = \sqrt{\int_0^{60} |\psi(x,t)|^2\,dx}$. The relative error $e_1(t)$ for several example reduced model simulations is depicted in Figure 5.3. The accuracy of the reduced model simulations is improved, as the number of modes k is increased. The convergence of the solutions of the reduced systems to the full-model solution is depicted in Figure 5.4, where the relative error e_2 is defined as

$$e_2 = \frac{\|\psi - \psi_{\mathrm{ref}}\|_2}{\|\psi_{\mathrm{ref}}\|_2}, \tag{5.11}$$

and $\|\psi\|_2 = \sqrt{\int_0^{200} \int_0^{60} |\psi(x,t)|^2\,dxdt}$ is the $L^2([0,60] \times [0,200])$ norm. A clear advantage of using the PSD method in combination with symplectic integration in time is the preservation of the Hamiltonian H in long-time simulations. The relative error of the Hamiltonian for several example reduced model simulations is depicted in Figure 5.5. We can see that, in contrast to the POD simulations, the PSD simulations nearly exactly preserve the Hamiltonian. As pointed out in Section 3.2, the PSD method in combination with the SDEIM approximation does not result in a Hamiltonian system. Nevertheless, after a long integration time the PSD+SDEIM reduced models

dim	POD	PSD
190	7m:32.4s *(speedup 50.5%)*	2m:50s *(speedup 81.4%)*
210	9m:22.3s *(speedup 38.4%)*	3m:29.3s *(speedup 77.1%)*
230	11m:23.8s *(speedup 25.1%)*	4m:16.4s *(speedup 71.9%)*
250	13m:05.3s *(speedup 14%)*	5m:00.5s *(speedup 67.1%)*

Table 1: Runtimes of the POD and PSD reduced simulations of the stochastic NLS equation using the midpoint method (averaged over 5 runs) for different values of the dimension of the reduced system (the dimension is equal to k for POD reduced models, and $2k$ for PSD reduced models). The simulation of the full model (of dimension $2N = 512$) with the midpoint method took 15m:13.2s (averaged over 5 runs).

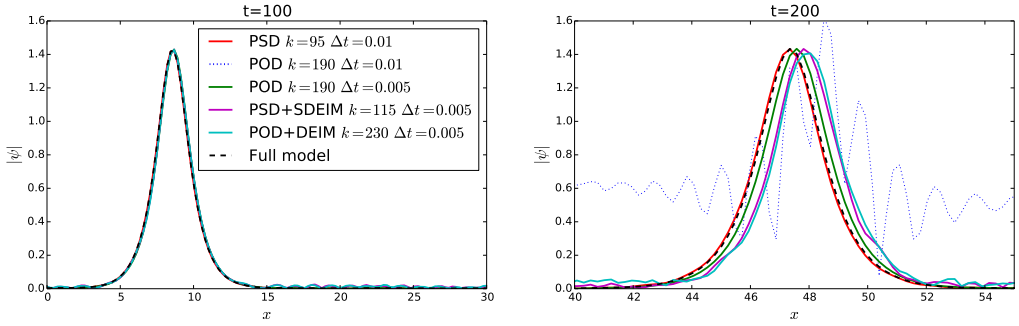

Figure 5.7: The solution $|\psi(x,t)| = |q(x,t) + ip(x,t)|$ of the stochastic NLS equation with the parameters $\beta = 0.15$ and $\epsilon = 1$ at times $t = 100$ (*Left*) and $t = 200$ (*Right*) obtained with the help of the reduced models integrated with the stochastic $R2$ method. While at $t = 100$ all simulations resolve the soliton relatively well, at $t = 200$ the PSD models yield a more accurate solution than the POD models of the same dimension and integrated with the same time step. The POD simulation with $k = 190$ and $\Delta t = 0.01$ eventually becomes unstable. The POD+DEIM and PSD+SDEIM simulations capture the propagation of the soliton, but create some spurious oscillations in its tail.

preserve the Hamiltonian better than the POD+DEIM reduced models of the same dimension. In order to further investigate the behavior of the SDEIM approximation in the stochastic setting, and to provide a numerical validation of Theorem 3.1, additional simulations for a fixed $k = 94$ and varying $\bar{k}$ were carried out. The results presented in Figure 5.6 indicate that the larger the number $\bar{k}$ is, the better the reduced system (3.15) preserves the Hamiltonian of the system (3.7). It is also worth noting that despite the slow decay of the singular values in Figure 5.1 and the relatively high dimension of the reduced system needed to obtain satisfactory accuracy (see Figure 5.4), it is still possible to achieve reasonable computational speedup compared to the full model simulations. The average runtimes of the POD and PSD reduced simulations using the stochastic midpoint method for several dimensions of the reduced system are summarized in Table 1.

The advantage of maintaining the Hamiltonian structure in constructing a reduced model is evident even when explicit non-symplectic schemes are used to integrate the reduced equations. The simulations using the stochastic $R2$ method were carried out with the time steps $\Delta t = 0.01$, $\Delta t = 0.005$, and $\Delta t = 0.001$. Generally, shorter time steps were needed in order to maintain the stability of the numerical solution over the integration time. The numerical solution for $|\psi(x,t)|$ obtained from the reduced models at times $t = 100$ and $t = 200$ is compared against the reference

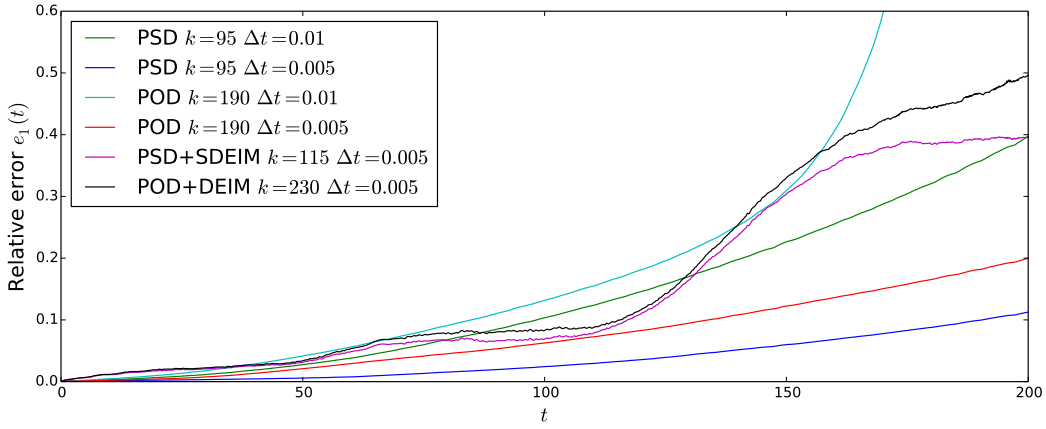

Figure 5.8: The relative error $e_1(t)$ as a function of time for several example reduced model simulations of the stochastic NLS equation using the stochastic $R2$ method. Note that the error for the POD method with $k = 190$ and $\Delta t = 0.01$ blows up.

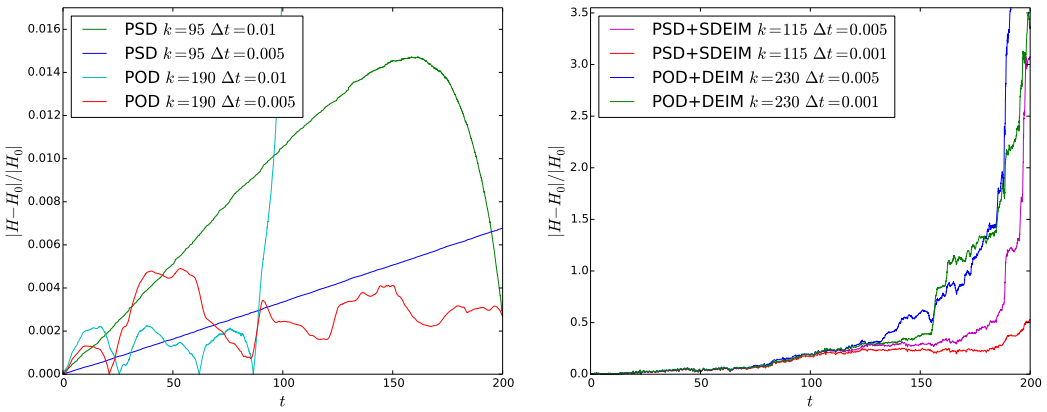

Figure 5.9: The relative error of the Hamiltonian as a function of time for several example reduced model simulations of the stochastic NLS equation using the stochastic $R2$ method is depicted, where H_0 denotes the initial value of the Hamiltonian. Note that the Hamiltonian for the POD method with $k = 190$ and $\Delta t = 0.01$ blows up.

solution in Figure 5.7. One can see that the PSD method with the time step $\Delta t = 0.01$ was able to produce an accurate solution up to the final time $t = 200$, while the POD method with the same time step eventually became unstable. The relative error $e_1(t)$ is depicted in Figure 5.8 and the relative error of the Hamiltonian is shown in Figure 5.9. The $R2$ method is not symplectic, therefore good preservation of the Hamiltonian is not expected, but also in this case we have observed that the PSD models produce more accurate solutions than the POD models of the same dimension and integrated with the same time step. The simulations using the stochastic Heun method yielded nearly identical results as the $R2$ method, therefore for brevity and clarity we skip presenting separate figures.

5.2 Kubo oscillator

The Kubo oscillator is a stochastic Hamiltonian system driven by a one-dimensional Wiener process with the Hamiltonians given by $H(q,p) = p^2/2 + q^2/2$ and $h(q,p) = \beta(p^2/2 + q^2/2)$, where β is the noise intensity (see [116]), and since the noise is one-dimensional, for simplicity we write $h \equiv h_1$. It is an example of an oscillator with a fluctuating frequency and it was first introduced in the context of the line-shape theory (see [5], [98]), but later also found many other applications in connection with mechanical systems, turbulence, laser theory, wave propagation (see [158] and the references therein), magnetic resonance spectroscopy, nonlinear spectroscopy (see [118] and the references therein), single molecule spectroscopy ([92]), and stochastic resonance ([44], [45], [46], [70]). The Kubo oscillator serves as a prototype for multiplicative stochastic processes, and since its solutions can be calculated analytically, it is often used for validation of numerical algorithms (see, e.g., [68], [111], [116], [150]). It is straightforward to verify that the exact solution is given by

$$q_e(t) = p_0 \sin(t + \beta W(t)) + q_0 \cos(t + \beta W(t)), \qquad p_e(t) = p_0 \cos(t + \beta W(t)) - q_0 \sin(t + \beta W(t)), \tag{5.12}$$

where q_0 and p_0 are the initial conditions. Note that (5.12) is the solution of the deterministic harmonic oscillator equation with the time argument shifted by $\beta W(t)$. It is also clearly evident that the Hamiltonian H is preserved along the solution (5.12). Since $W(t) \sim N(0,t)$ is normally distributed, one can explicitly calculate the mean position and momentum as, respectively,

$$\mathbb{E}[q_e(t)] = e^{-\frac{\beta^2}{2}t}(p_0 \sin t + q_0 \cos t), \qquad \mathbb{E}[p_e(t)] = e^{-\frac{\beta^2}{2}t}(p_0 \cos t - q_0 \sin t). \tag{5.13}$$

5.2.1 Empirical data

Suppose we have the following computational problem: we would like to calculate the expected value of the solution to (1.2) for a large number of parameters β. In order to accurately approximate the mean value of a stochastic process one typically needs thousands, or even millions Monte Carlo runs. This presents a computational challenge which can be alleviated by model reduction. One can carry out full-scale computations only for a selected number of values of β, and the resulting data can be used to identify a reduced model, as described in Section 4. The lower dimensional equations can then be solved for other values of β. In our experiment we have considered the initial conditions $q_0 = 0$ and $p_0 = 1$, and we have used the exact solution (5.12) to generate the empirical data for the following six (arbitrarily selected) values of the parameter β:

$$\beta = 0.00095,\ 0.00097,\ 0.00099,\ 0.00101,\ 0.00103,\ 0.00105. \tag{5.14}$$

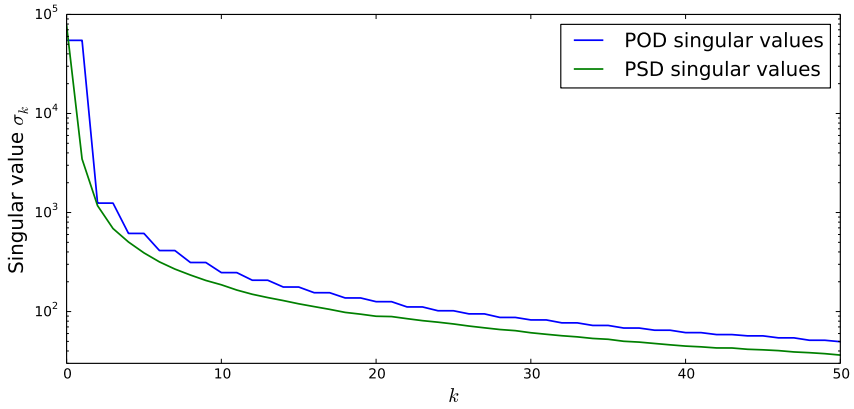

Figure 5.10: The decay of the singular values for the POD and PSD reductions for the empirical data ensemble for the Kubo oscillator. Because of their large number and rapid decay, for clarity only the first 50 singular values are depicted.

The exact solution (5.12) was sampled over the time interval $0 \le t \le 5000$ with the time step $\Delta t = 0.05$ for $M = 10000$ independent realizations of the Wiener process, $W(t;\omega_1),\ldots,W(t;\omega_M)$, where ω_ν denote elementary events in the probability space. Of course in the general situation the empirical data are generated using high-fidelity numerical methods applied to the full model, like in Section 5.1, but our experiment was meant as a validation of the model reduction method, therefore the exact solution was used for convenience. Then, following the description of each algorithm in Section 4, reduced models were derived. The single realization of the M-dimensional Wiener process in (4.2) and (4.7) was defined as $W^\nu(t) = W(t;\omega_\nu)$ for $\nu = 1,\ldots,M$. Note that for the Kubo oscillator the drift and diffusion terms are linear, therefore no DEIM approximations were necessary. The decay of the singular values for the POD and PSD methods is depicted in Figure 5.10.

5.2.2 Reduced model simulations

The reduced models obtained in Section 5.2.1 from the empirical data can now be solved for any other desired value of the parameter β. To test the accuracy of the considered model reduction methods, we have compared the results of the reduced model simulations to a full-scale reference solution for a single arbitrary choice of the parameter. The reference solution for $\beta = 0.001$ was generated on the time interval $0 \le t \le 5000$ in the same way as the empirical data in Section 5.2.1. The POD reduced model was solved using the stochastic R2 and Heun methods for $k = 42$ (thus reducing the dimensionality of the system (4.2) from $2M = 20000$ to 42). The PSD reduced model was solved using the stochastic Störmer-Verlet method (3.24) for $k = 21$ (thus reducing the dimensionality of the system (4.7) from $2M = 20000$ to $2k = 42$). Note that the Hamiltonians for the Kubo oscillator are separable, therefore the Störmer-Verlet method is in this case explicit.

As a measure of accuracy of the reduced models we considered three types of error. First, we investigated the relative error of a selected single sample path, that is,

$$E_1(t) = \frac{\|u(t) - u_e(t)\|}{\|u_e(t)\|}, \tag{5.15}$$

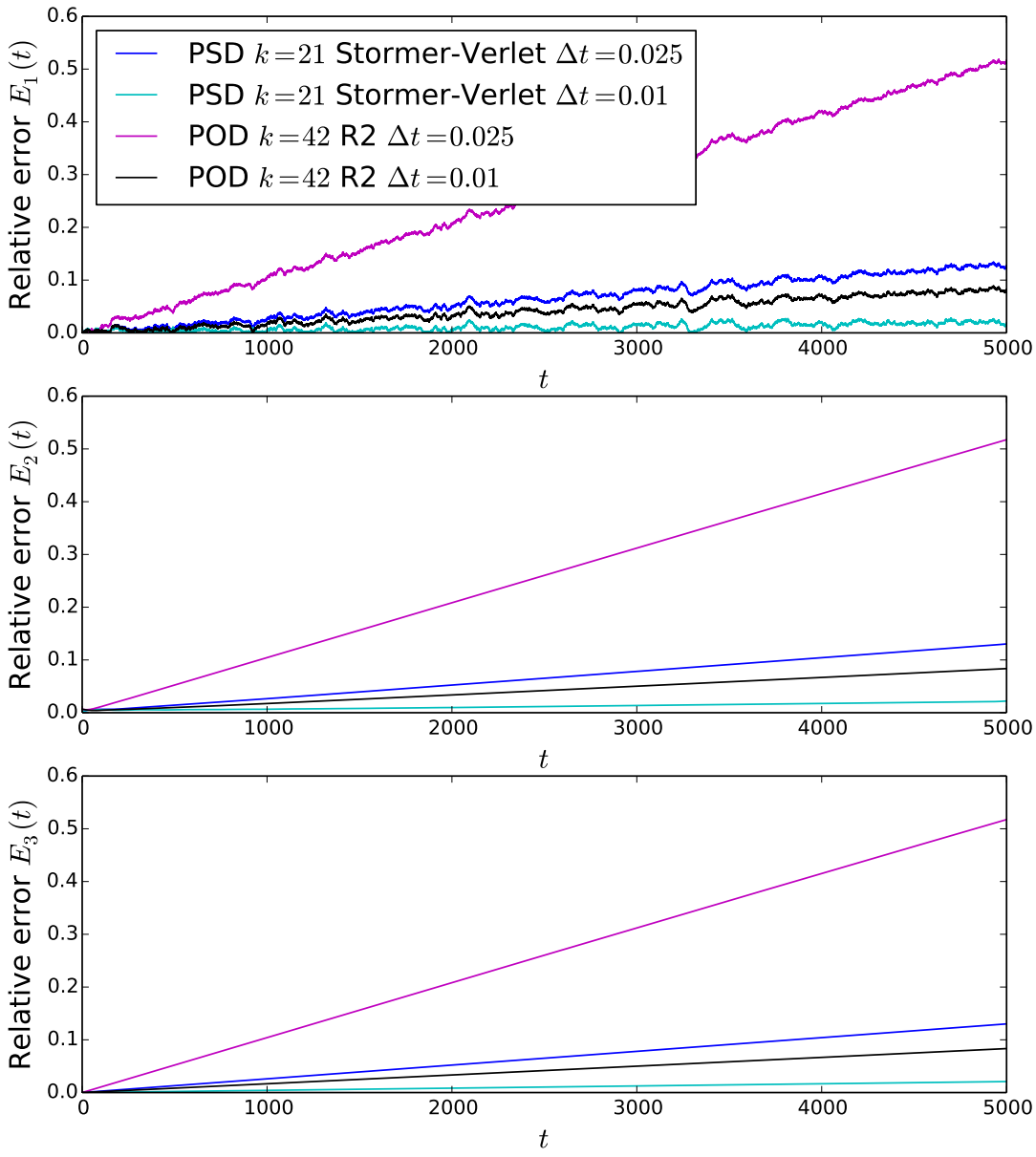

Figure 5.11: The relative errors $E_1(t)$ (*Top*), $E_2(t)$ (*Middle*), and $E_3(t)$ (*Bottom*) for the reduced model simulations of the Kubo oscillator are depicted as functions of time. The PSD method combined with the stochastic Störmer-Verlet scheme yields more accurate solutions than the POD method combined with the $R2$ integrator and using the same time step. The POD simulations with the Heun method yield nearly identical results as the $R2$ method, therefore separate plots are omitted for clarity.

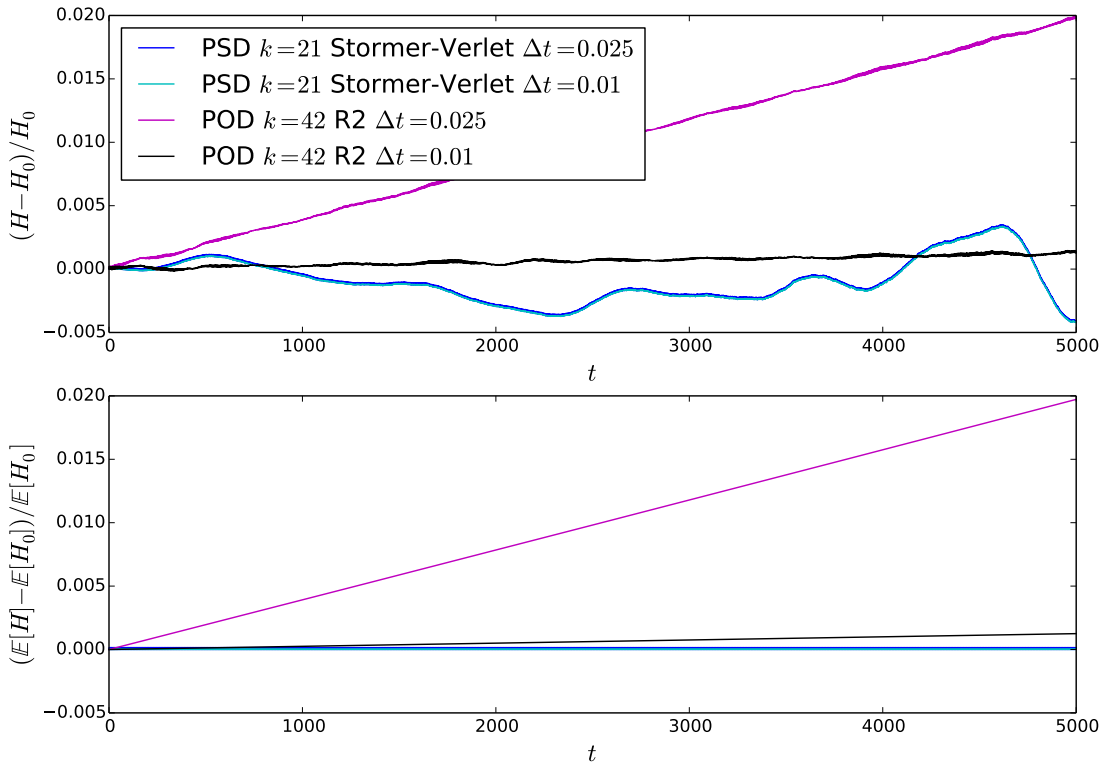

Figure 5.12: The relative errors of the Hamiltonian H for a selected sample path (*Top*) and of the mean Hamiltonian $\mathbb{E}[H]$ (*Bottom*) for the reduced model simulations of the Kubo oscillator are depicted as functions of time, where H_0 denotes the initial value of the Hamiltonian. The PSD simulations using the Störmer-Verlet method nearly exactly preserve the mean Hamiltonian, while the POD simulations using the $R2$ and Heun methods demonstrate a linear growth trend. Note that the plots for the PSD simulations with the time steps Δt = 0.025 and Δt = 0.01 overlap very closely and are therefore barely distinguishable. The POD simulations with the Heun method yield nearly identical results as the $R2$ method, therefore separate plots are omitted for clarity.

where $u_e = (q_e, p_e)$ and $u = (q, p)$ denote single sample paths of the exact reference solution and of the solution reconstructed from the reduced model, respectively, and $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^2$. Second, we studied the relative mean-square error, that is,

$$E_2(t) = \sqrt{\frac{\mathbb{E}\big[\|u(t) - u_e(t)\|^2\big]}{\mathbb{E}\big[\|u_e(t)\|^2\big]}}. \tag{5.16}$$

Finally, we used the relative error of the mean (also known as the weak error), that is,

$$E_3(t) = \frac{\Big\|\mathbb{E}[u(t)] - \mathbb{E}[u_e(t)]\Big\|}{\Big\|\mathbb{E}[u_e(t)]\Big\|}. \tag{5.17}$$

All errors are depicted in Figure 5.11. We can see that the reduced model simulations indeed yield good approximations of the exact solution. We also see that the PSD method combined with the stochastic Störmer-Verlet scheme yields more accurate solutions than the POD method combined with the $R2$ and Heun integrators. The numerical values of the Hamiltonian H for a selected single sample path and of the mean Hamiltonian $\mathbb{E}[H]$ as functions of time are depicted in Figure 5.12. The PSD simulations using the Störmer-Verlet method nearly exactly preserve the mean Hamiltonian, while the POD simulations using the $R2$ and Heun methods demonstrate a linear growth trend.

5.3 Forced Kubo oscillator

In this experiment we consider the Kubo oscillator described, as in Section 5.2, by the Hamiltonians $H(q,p) = p^2/2 + q^2/2$ and $h(q,p) = \beta(p^2/2 + q^2/2)$, subject to external damping, with the forcing terms given by $F(q,p) = -\nu p$ and $f(q,p) = -\beta\nu p$, where ν is the damping coefficient. Since the noise is one-dimensional, for simplicity we write $h \equiv h_1$ and $f \equiv f_1$. It is straightforward to verify that the exact solution is given by

$$\begin{aligned}
q_e(t) &= q_0 e^{-\frac{\nu}{2}(t+\beta W(t))}\cos\omega(t+\beta W(t)) + \frac{1}{\omega}(p_0 + \frac{\nu}{2}q_0)e^{-\frac{\nu}{2}(t+\beta W(t))}\sin\omega(t+\beta W(t)),\\
p_e(t) &= p_0 e^{-\frac{\nu}{2}(t+\beta W(t))}\cos\omega(t+\beta W(t)) - \frac{1}{\omega}(q_0 + \frac{\nu}{2}p_0)e^{-\frac{\nu}{2}(t+\beta W(t))}\sin\omega(t+\beta W(t)),
\end{aligned} \tag{5.18}$$

where q_0 and p_0 are the initial conditions, the angular frequency is $\omega = \frac{1}{2}\sqrt{4-\nu^2}$, and we have assumed the underdamped case $0 \le \nu < 2$. Note that (5.18) is the solution of the deterministic damped harmonic oscillator with the time argument shifted by $\beta W(t)$. Given that $W(t) \sim N(0,t)$ is normally distributed, one can explicitly calculate the expected value of the Hamiltonian H as a function of time as

$$E\Big(H\big(q_e(t), p_e(t)\big)\Big) = ae^{-\frac{\nu(2-\beta^2\nu)}{2}t} + e^{-((2-\nu^2)\beta^2+\nu)t}\Big[b\cos\big(2(1-\beta^2\nu)\omega t\big) + c\sin\big(2(1-\beta^2\nu)\omega t\big)\Big], \tag{5.19}$$

where

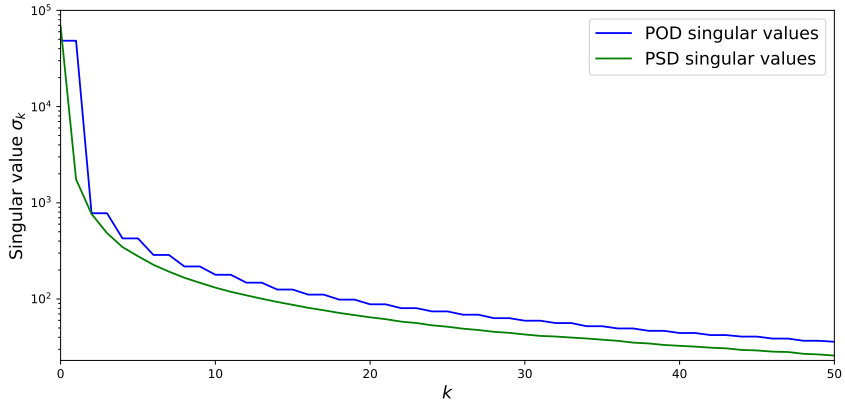

Figure 5.13: The decay of the singular values for the POD and PSD reductions for the empirical data ensemble for the forced Kubo oscillator. Because of their large number and rapid decay, for clarity only the first 50 singular values are depicted.

$$a = \frac{2(p_0^2 + q_0^2 + \nu p_0 q_0)}{4 - \nu^2}, \qquad b = -\frac{\nu^2(p_0^2 + q_0^2) + 4\nu p_0 q_0}{2(4 - \nu^2)}, \qquad c = \frac{\nu(q_0^2 - p_0^2)}{2\sqrt{4 - \nu^2}}. \tag{5.20}$$

5.3.1 Empirical data

Suppose that, similar to Section 5.2.1, we have the following computational problem: a solution (or its statistical properties) to (3.21) is needed for a large number of parameters β and ν. In order to efficiently calculate, e.g., the expected value of the solution, we use full-model data for a selected number of values of β and ν to construct a reduced model which is less expensive to solve. In our experiment we have considered the initial conditions $q_0 = 2$ and $p_0 = 0$, and we have used the exact solution (5.18) to generate the empirical data for $\beta = 0.001$ and for the following six (arbitrarily selected) values of the damping coefficient ν:

$$\nu = 0.00097,\ 0.00099,\ 0.00101,\ 0.00103,\ 0.00105,\ 0.00107. \tag{5.21}$$

The exact solution (5.18) was sampled over the time interval $0 \le t \le 5000$ with the time step $\Delta t = 0.05$ for $M = 10000$ independent realizations of the Wiener process, $W(t;\omega_1), \ldots, W(t;\omega_M)$, where ω_ν denote elementary events in the probability space. Then, following the description of each algorithm in Section 4, reduced models were derived. The single realization of the M-dimensional Wiener process in (4.2) and (4.7) was defined as $W^\nu(t) = W(t;\omega_\nu)$ for $\nu = 1, \ldots, M$. Note that for the forced Kubo oscillator with the linear forcing terms, the drift and diffusion terms are linear, therefore no DEIM approximations were necessary. The decay of the singular values for the POD and PSD methods is depicted in Figure 5.13.

5.3.2 Reduced model simulations

The reduced models obtained in Section 5.3.1 can now be solved for any other desired value of the parameter ν. To test the accuracy of the considered model reduction methods, we have compared

the results of the reduced model simulations to a full-scale reference solution for a single arbitrary choice of the parameter. The reference solution for $\beta = 0.001$ and $\nu = 0.001$ was generated on the time interval $0 \leq t \leq 5000$ in the same way as the empirical data in Section 5.3.1. Both the POD and PSD reduced models were integrated using the stochastic Störmer-Verlet method. Note that in the case of the forced Kubo oscillator, the Störmer-Verlet method is implicit for both the POD and PSD models. As explained in Section 3.4, when applied to the PSD reduced model, the Störmer-Verlet method is a structure-preserving Lagrange-d'Alembert integrator, whereas it does not have this property in the case of the POD reduced model. The POD reduced model was solved for $k = 42$ (thus reducing the dimensionality of the full system from $2M = 20000$ to 42). The PSD reduced model was solved for $k = 21$ (thus reducing the dimensionality of the full system from $2M = 20000$ to $2k = 42$).

The errors $E_1(t)$, $E_2(t)$, and $E_3(t)$ (see (5.15), (5.16), and (5.17), respectively) are depicted in Figure 5.14. We can see that the reduced model simulations indeed yield good approximations of the exact solution. We also see that the PSD method yields more accurate solutions than the POD method with the same time step. The theoretical time evolution (5.19) of the mean Hamiltonian $\mathbb{E}[H_e]$ for the chosen initial conditions and parameters is depicted in Figure 5.15, whereas the errors of the corresponding numerical values $\mathbb{E}[H]$ are depicted in Figure 5.16. It is evident that the PSD simulations capture the evolution of the Hamiltonian more accurately than the POD simulations with the same time step.

6 Summary

We have successfully demonstrated that SVD-based model reduction methods known for ordinary differential equations can be extended to stochastic differential equations, and can be used to reduce the computational cost arising from both the high dimension of the considered stochastic system and the large number of independent Monte Carlo runs. We have also argued that in model reduction of stochastic Hamiltonian systems it is advisable to maintain their symplectic or variational structures, to which end we have adapted to the stochastic setting the proper symplectic decomposition method known for deterministic Hamiltonian systems. We have further applied our proposed techniques to a semi-discretization of the stochastic Nonlinear Schrödinger equation and to the Kubo oscillator, providing numerical evidence that model reduction is a viable tool for obtaining accurate numerical approximations of stochastic systems, and that preserving the geometric structures of stochastic Hamiltonian systems results in more accurate and stable solutions that conserve energy better than when the non-geometric approach is used.

Our work can be extended in many ways. A natural follow-up study would be the application of the presented model reduction techniques to semi-discretizations of other stochastic PDEs with underlying geometric structures, such as the stochastic Camassa-Holm equation ([19], [81], [82]), the stochastic quasi-geostrophic equation ([18], [54], [55], [56], [57]), the stochastic Korteweg-de Vries equation ([48], [58], [77], [82]), or the stochastic Sine-Gordon equation ([76], [153], [155], [162]) to name just a few. Particle discretizations of collisional Vlasov equations have been recently proved to have the structure of stochastic forced Hamiltonian systems ([97], [154], [157]), therefore kinetic plasma theory is yet another area where structure-preserving model reduction techniques could be applied. Furthermore, it would be interesting to extend the methods discussed in our work to stochastic non-canonical systems. For instance, one could consider structure-preserving model reduction techniques for stochastic Poisson systems by combining the method developed in [78] with

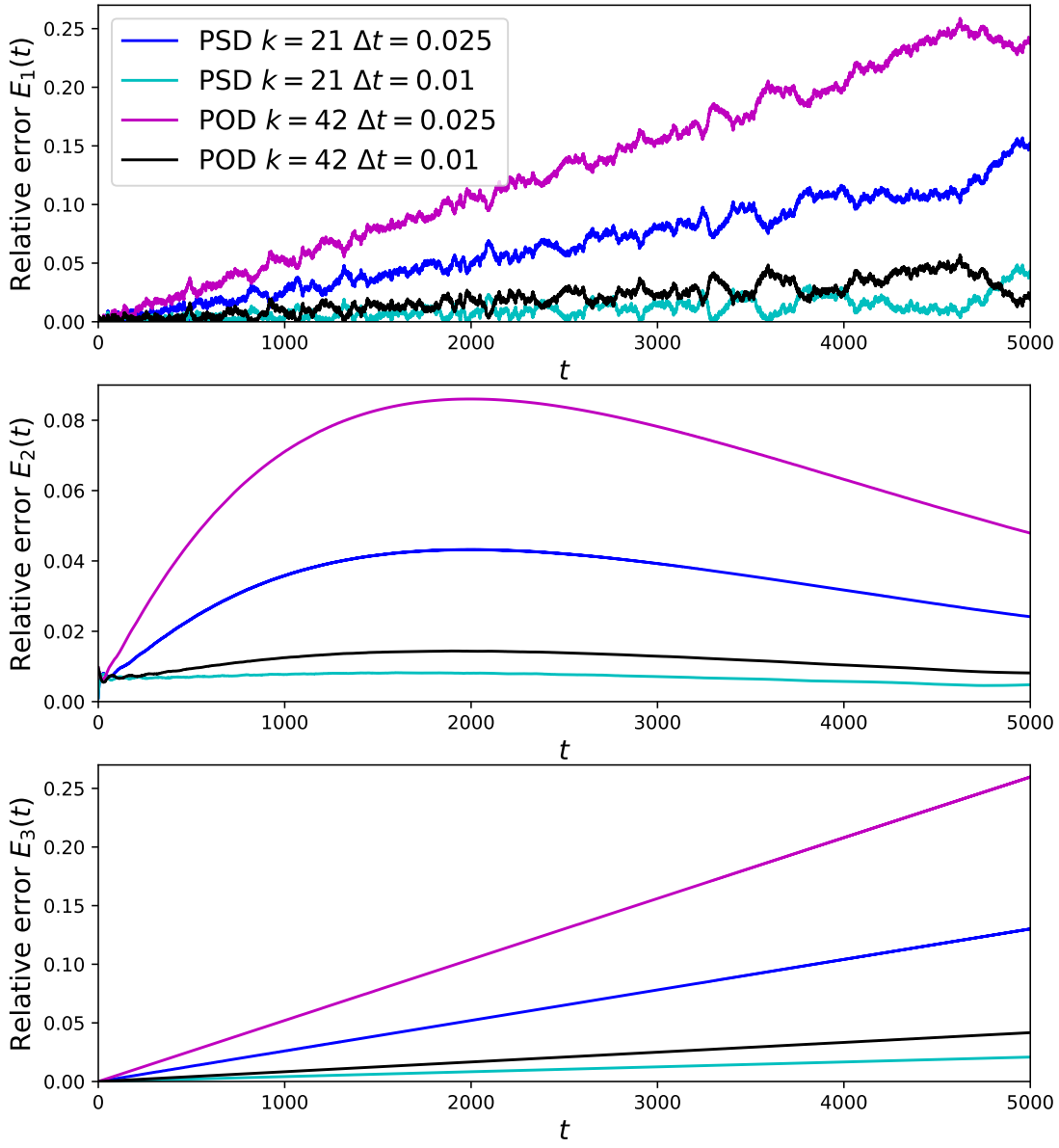

Figure 5.14: The relative errors $E_1(t)$ (*Top*), $E_2(t)$ (*Middle*), and $E_3(t)$ (*Bottom*) for the reduced model simulations of the forced Kubo oscillator using the stochastic Störmer-Verlet scheme are depicted as functions of time. The PSD method yields more accurate solutions than the POD method using the same time step.

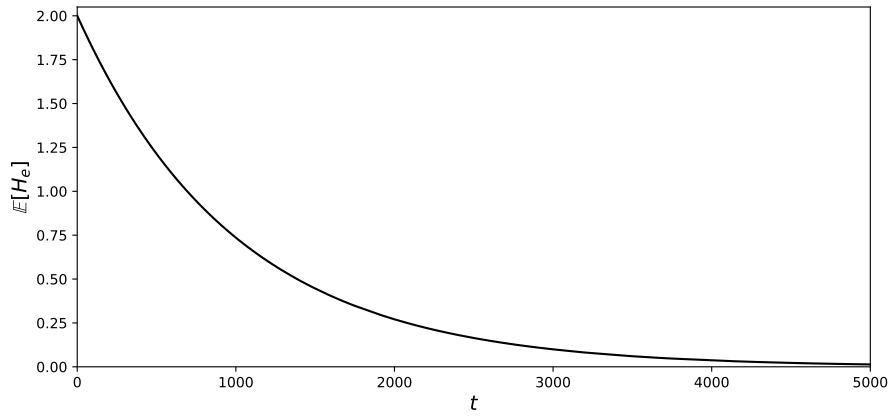

Figure 5.15: The time dependence of the expected value of the Hamiltonian $\mathbb{E}[H_e]$ for the exact solution (5.18) of the forced Kubo oscillator with the initial conditions $q_0 = 2$ and $p_0 = 0$, and the parameters $\beta = 0.001$ and $\nu = 0.001$.

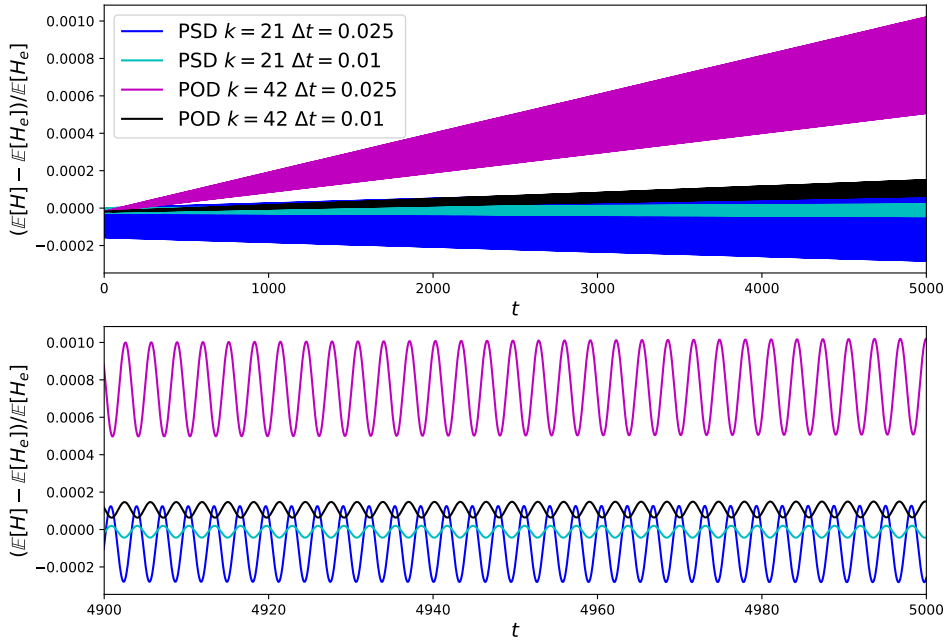

Figure 5.16: *Top:* The relative error of the mean Hamiltonian $\mathbb{E}[H]$ for the reduced model simulations of the forced Kubo oscillator using the stochastic Störmer-Verlet method is depicted as a function of time over the whole integration interval, where $H_e = H(q_e(t), p_e(t))$ denotes the value of the Hamiltonian for the exact solution (5.18). It is evident that the PSD simulations capture the evolution of Hamiltonian more accurately than the POD simulations with the same time step. *Bottom:* The same plot zoomed in on the time interval [4900,5000] for a more detailed depiction of the oscillatory character of the time dependence of the error.

the stochastic geometric integrators presented in [50] and [52]. This would be of great interest for systems appearing in gyrokinetic and guiding-center theories (see [13], [23], [28], [29], [41], [128], [129], [148], [156], [159]).

Acknowledgements

We would like to thank Tobias Blickhan, Michael Kraus, Paul Skerritt and Udo von Toussaint for useful comments and references. The study is a contribution to the Reduced Complexity Models grant number ZT-I-0010 funded by the Helmholtz Association of German Research Centers.

References

[1] B. Afkham and J. Hesthaven. Structure preserving model reduction of parametric Hamiltonian systems. *SIAM Journal on Scientific Computing*, 39(6):A2616–A2644, 2017.

[2] B. M. Afkham, A. Bhatt, B. Haasdonk, and J. S. Hesthaven. Symplectic model-reduction with a weighted inner product. Unpublished, arXiv:1803.07799, 2018.

[3] B. M. Afkham and J. S. Hesthaven. Structure-preserving model-reduction of dissipative Hamiltonian systems. *Journal of Scientific Computing*, 81:3–21, 2019.

[4] M. Amabili, A. Sarkar, and M. Païdoussis. Reduced-order models for nonlinear vibrations of cylindrical shells via the proper orthogonal decomposition method. *Journal of Fluids and Structures*, 18(2):227–250, 2003. Axial and Internal Flow Fluid-Structure Interactions.

[5] P. Anderson. A mathematical model for the narrowing of spectral lines by exchange or motion. *Journal of the Physical Society of Japan*, 9(3):316–339, 1954.

[6] S. Anmarkrud and A. Kværnø. Order conditions for stochastic Runge-Kutta methods preserving quadratic invariants of Stratonovich SDEs. *Journal of Computational and Applied Mathematics*, 316:40 – 46, 2017.

[7] C. Anton, J. Deng, and Y. S. Wong. Weak symplectic schemes for stochastic Hamiltonian equations. *Electronic Transactions on Numerical Analysis*, 43:1–20, 2014.

[8] C. Anton, Y. S. Wong, and J. Deng. On global error of symplectic schemes for stochastic Hamiltonian systems. *International Journal of Numerical Analysis and Modeling, Series B*, 4(1):80–93, 2013.

[9] C. Anton, Y. S. Wong, and J. Deng. Symplectic numerical schemes for stochastic systems preserving Hamiltonian functions. In I. Dimov, I. Faragó, and L. Vulkov, editors, *Numerical Analysis and Its Applications: 5th International Conference, NAA 2012, Lozenetz, Bulgaria, June 15-20, 2012, Revised Selected Papers*, pages 166–173. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[10] R. Anton and D. Cohen. Exponential integrators for stochastic Schrödinger equations driven by Itô noise. *Journal of Computational Mathematics*, 36(2):276–309, 2018.

[11] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. A survey of model reduction methods for large-scale systems. *Contemporary Mathematics*, 280:193–219, 2001.

[12] L. Arnold. *Stochastic Differential Equations: Theory and Applications*. Dover Books on Mathematics. Dover Publications, 2013.

[13] R. Balescu, H. Wang, and J. H. Misguich. Langevin equation versus kinetic equation: Subdiffusive behavior of charged particles in a stochastic magnetic field. *Physics of Plasmas*, 1(12):3826–3842, 1994.

[14] O. Bang, P. Christiansen, F. If, K. Rasmussen, and Y. Gaididei. Temperature effects in a nonlinear model of monolayer scheibe aggregates. *Phys. Rev. E*, 49:4627–4636, May 1994.

[15] O. Bang, P. Christiansen, F. If, K. Rasmussen, and Y. Gaididei. White noise in the two-dimensional nonlinear Schrödinger equation. *Applicable Analysis*, 57(1-2):3–15, 1995.

[16] D. Bastine, L. Vollmer, M. Wächter, and J. Peinke. Stochastic wake modelling based on POD analysis. *Energies*, 11(3), 2018.

[17] D. A. Beard and T. Schlick. Inertial stochastic dynamics. I. Long-time-step methods for Langevin dynamics. *The Journal of Chemical Physics*, 112(17):7313–7322, 2000.

[18] T. M. Bendall and C. J. Cotter. Statistical properties of an enstrophy conserving finite element discretisation for the stochastic quasi-geostrophic equation. *Geophysical & Astrophysical Fluid Dynamics*, 113(5-6):491–504, 2019.

[19] T. M. Bendall, C. J. Cotter, and D. D. Holm. Perspectives on the formation of peakons in the stochastic Camassa-Holm equation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477(2250):20210224, 2021.

[20] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Review*, 57(4):483–531, 2015.

[21] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox. *Model Reduction and Approximation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2017.

[22] J. Bismut. Mecanique aleatoire. In P. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour X - 1980*, volume 929 of *Lecture Notes in Mathematics*, pages 1–100. Springer Berlin Heidelberg, 1982.

[23] A. Bottino and E. Sonnendrücker. Monte Carlo particle-in-cell methods for the simulation of the Vlasov-Maxwell gyrokinetic equations. *Journal of Plasma Physics*, 81(5):435810501, 2015.

[24] S. Boyaval, C. Bris, T. Lelièvre, Y. Maday, N. Nguyen, and A. Patera. Reduced basis techniques for stochastic problems. *Archives of Computational Methods in Engineering*, 17:435–454, 12 2010.

[25] S. Boyaval, C. L. Bris, Y. Maday, N. C. Nguyen, and A. T. Patera. A reduced basis approach for variational problems with stochastic parameters: Application to heat conduction with variable Robin coefficient. *Computer Methods in Applied Mechanics and Engineering*, 198(41):3187–3206, 2009.

[26] S. Boyaval and T. Lelièvre. A variance reduction method for parametrized stochastic differential equations using the reduced basis paradigm. *Communications in Mathematical Sciences*, 8(3):735–762, 2010.

[27] C.-E. Bréhier, D. Cohen, and T. Jahnke. Splitting integrators for stochastic Lie-Poisson systems. Preprint arXiv:2111.07387, 2021.

[28] A. J. Brizard. Variational principle for nonlinear gyrokinetic Vlasov-Maxwell equations. *Physics of Plasmas*, 7(12):4816–4822, 2000.

[29] A. J. Brizard and C. Tronci. Variational formulations of guiding-center Vlasov-Maxwell theory. *Physics of Plasmas*, 23(6):062107, 2016.

[30] P. Buchfink, A. Bhatt, and B. Haasdonk. Symplectic model order reduction with non-orthonormal bases. *Mathematical and Computational Applications*, 24:43, 2019.

[31] A. Buffa, Y. Maday, A. T. Patera, C. Prud´homme, and G. Turinici. A priori convergence of the greedy algorithm for the parametrized reduced basis method. *ESAIM: M2AN*, 46(3):595–603, 2012.

[32] J. Burkardt, M. D. Gunzburger, and C. Webster. Reduced order modeling of some nonlinear stochastic partial differential equations. *International Journal of Numerical Analysis and Modeling*, 4(3-4):368–391, 2007.

[33] K. Burrage and P. M. Burrage. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Appl. Numer. Math.*, 22:81–101, 1996.

[34] K. Burrage and P. M. Burrage. General order conditions for stochastic Runge-Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Appl. Numer. Math.*, 28:161–177, 1998.

[35] K. Burrage and P. M. Burrage. Order Conditions of Stochastic Runge–Kutta Methods by B-Series. *SIAM Journal on Numerical Analysis*, 38(5):1626–1646, 2000.

[36] K. Burrage and P. M. Burrage. Low rank Runge-Kutta methods, symplecticity and stochastic Hamiltonian problems with additive noise. *Journal of Computational and Applied Mathematics*, 236(16):3920 – 3930, 2012.

[37] P. Burrage. *Runge-Kutta methods for stochastic differential equations*. PhD thesis, University of Queensland, 1999.

[38] P. M. Burrage and K. Burrage. Structure-preserving Runge-Kutta methods for stochastic Hamiltonian equations with additive noise. *Numerical Algorithms*, 65(3):519–532, 2014.

[39] K. Carlberg. Adaptive h-refinement for reduced-order models. *International Journal for Numerical Methods in Engineering*, 102(5):1192–1210, 2015.

[40] K. Carlberg, R. Tuminaro, and P. Boggs. Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics. *SIAM Journal on Scientific Computing*, 37(2):B153–B184, 2015.

[41] J. R. Cary and A. J. Brizard. Hamiltonian theory of guiding-center motion. *Rev. Mod. Phys.*, 81:693–738, 2009.

[42] S. Chaturantabut, C. Beattie, and S. Gugercin. Structure-preserving model reduction for nonlinear port-Hamiltonian systems. *SIAM Journal on Scientific Computing*, 38(5):B837–B865, 2016.

[43] S. Chaturantabut and D. C. Sorensen. Nonlinear model reduction via discrete empirical interpolation. *SIAM Journal on Scientific Computing*, 32(5):2737–2764, 2010.

[44] J. R. Chaudhuri and S. Chattopadhyay. Microscopic realization of Kubo oscillator. *Chemical Physics Letters*, 480(1):140 – 143, 2009.

[45] J. R. Chaudhuri and S. Chattopadhyay. Kubo oscillator and its application to stochastic resonance: A microscopic realization. In R. K. Chaudhuri, M. Mekkaden, A. V. Raveendran, and A. Satya Narayanan, editors, *Recent Advances in Spectroscopy*, pages 75–83, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[46] J. R. Chaudhuri, P. Chaudhury, and S. Chattopadhyay. Harmonic oscillator in presence of nonequilibrium environment. *The Journal of Chemical Physics*, 130(23):234109, 2009.

[47] C. Chen, D. Cohen, R. D'Ambrosio, and A. Lang. Drift-preserving numerical integrators for stochastic hamiltonian systems. *Adv. Comput. Math.*, 46(2):27, 2020.

[48] Y. Chen, Q. Wang, and B. Li. The stochastic soliton-like solutions of stochastic KdV equations. *Chaos, Solitons & Fractals*, 23(4):1465–1473, 2005.

[49] A. J. Chorin and F. Lu. Discrete approach to stochastic parametrization and dimension reduction in nonlinear dynamics. *Proceedings of the National Academy of Sciences*, 112(32):9804–9809, 2015.

[50] D. Cohen and G. Dujardin. Energy-preserving integrators for stochastic Poisson systems. *Communications in Mathematical Sciences*, 12:1523–1539, 01 2014.

[51] D. Cohen and G. Dujardin. Exponential integrators for nonlinear Schrödinger equations with white noise dispersion. *Stochastics and Partial Differential Equations: Analysis and Computations*, 5(4):592–613, 2017.

[52] D. Cohen and G. Vilmart. Drift-preserving numerical integrators for stochastic Poisson systems. *International Journal of Computer Mathematics*, 0(0):1–17, 2021.

[53] F. G. Cordoni, L. Di Persio, and R. Muradore. Discrete stochastic port-Hamiltonian systems. *Automatica*, 137:110122, 2022.

[54] C. Cotter, D. Crisan, D. Holm, W. Pan, and I. Shevchenko. Data assimilation for a quasi-geostrophic model with circulation-preserving stochastic transport noise. *Journal of Statistical Physics*, 179:1186–1221, 06 2020.

[55] C. Cotter, D. Crisan, D. Holm, W. Pan, and I. Shevchenko. Modelling uncertainty using stochastic transport noise in a 2-layer quasi-geostrophic model. *Foundations of Data Science*, 2(2):173–205, 2020.

[56] C. Cotter, D. Crisan, D. D. Holm, W. Pan, and I. Shevchenko. Numerically modeling stochastic lie transport in fluid dynamics. *Multiscale Modeling & Simulation*, 17(1):192–232, 2019.

[57] C. J. Cotter, G. A. Gottwald, and D. D. Holm. Stochastic partial differential fluid equations as a diffusive limit of deterministic Lagrangian multi-time dynamics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 473(2205), 2017.

[58] A. de Bouard and A. Debussche. On the stochastic Korteweg-de Vries equation. *Journal of Functional Analysis*, 154(1):215–251, 1998.

[59] A. de Bouard, A. Debussche, and L. Di Menza. Theoretical and numerical aspects of stochastic nonlinear Schrödinger equations. *Journées équations aux dérivées partielles*, 2001.

[60] A. Debussche and L. Di Menza. Numerical resolution of stochastic focusing nls equations. *Appl. Math. Lett.*, 15:661–669, 2002.

[61] A. Debussche and L. Di Menza. Numerical simulation of focusing stochastic nonlinear Schrödinger equations. *Physica D: Nonlinear Phenomena*, 162(3):131–154, 2002.

[62] J. Deng, C. Anton, and Y. S. Wong. High-order symplectic schemes for stochastic Hamiltonian systems. *Communications in Computational Physics*, 16:169–200, 2014.

[63] G. Dôme. Theory of RF acceleration. In S. Turner, editor, *Proceedings of CERN Accelerator School, Oxford, September 1985*, volume 1, pages 110–158. CERN European Organization for Nuclear Research, 1987.

[64] A. Doostan, R. G. Ghanem, and J. Red-Horse. Stochastic model reduction for chaos representations. *Computer Methods in Applied Mechanics and Engineering*, 196(37):3951–3966, 2007.

[65] M. Drohmann, B. Haasdonk, and M. Ohlberger. Reduced basis approximation for nonlinear parametrized evolution equations based on empirical operator interpolation. *SIAM Journal on Scientific Computing*, 34(2):A937–A969, 2012.

[66] J. Elgin. Stochastic perturbations of optical solitons. *Physics Letters A*, 181(1):54–60, 1993.

[67] G. E. Falkovich, I. Kolokolov, V. Lebedev, and S. K. Turitsyn. Statistics of soliton-bearing systems with additive noise. *Phys. Rev. E*, 63:025601, Jan 2001.

[68] R. Fox, R. Roy, and A. Yu. Tests of numerical simulation algorithms for the Kubo oscillator. *Journal of Statistical Physics*, 47:477–487, 1987.

[69] R. G. Ghanem, G. Saad, and A. Doostan. Efficient solution of stochastic systems: Application to the embankment dam problem. *Structural Safety*, 29:238–251, 2007.

[70] M. Gitterman. Harmonic oscillator with fluctuating damping parameter. *Phys. Rev. E*, 69:041101, Apr 2004.

[71] S. Glavaski, J. Marsden, and R. Murray. Model reduction, centering, and the Karhunen-Loeve expansion. In *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No.98CH36171)*, volume 2, pages 2071–2076 vol.2, 1998.

[72] Y. N. Gornostyrev, M. I. Katsnelson, A. V. Trefilov, and S. V. Tret'jakov. Stochastic approach to simulation of lattice vibrations in strongly anharmonic crystals: Anomalous frequency dependence of the dynamic structure factor. *Phys. Rev. B*, 54:3286–3294, Aug 1996.

[73] C. Greif and K. Urban. Decay of the Kolmogorov N-width for wave problems. *Applied Mathematics Letters*, 96:216–222, 2019.

[74] S. Gugercin, R. V. Polyuga, C. Beattie, and A. van der Schaft. Structure-preserving tangential interpolation for model reduction of port-Hamiltonian systems. *Automatica*, 48(9):1963–1974, 2012.

[75] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics. Springer, New York, 2002.

[76] M. Hairer and H. Shen. The dynamical Sine-Gordon model. *Communications in Mathematical Physics*, 341:933–989, 02 2016.

[77] R. L. Herman and A. Rose. Numerical realizations of solutions of the stochastic KdV equation. *Math. Comput. Simul.*, 80(1):164–172, sep 2009.

[78] J. Hesthaven and C. Pagliantini. Structure-preserving reduced basis methods for Hamiltonian systems with a state-dependent Poisson structure. *Mathematics of Computation*, 90:1701–1740, 11 2020.

[79] J. Hesthaven, C. Pagliantini, and N. Ripamonti. Structure-preserving model order reduction of Hamiltonian systems. Unpublished, arXiv:2109.12367, 2021.

[80] D. D. Holm, T. Schmah, and C. Stoica. *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Oxford Texts in Applied and Engineering Mathematics. Oxford University Press, Oxford, 2009.

[81] D. D. Holm and T. M. Tyranowski. Variational principles for stochastic soliton dynamics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 472(2187), 2016.

[82] D. D. Holm and T. M. Tyranowski. New variational and multisymplectic formulations of the Euler–Poincaré equation on the Virasoro–Bott group using the inverse map. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 474(2213), 2018.

[83] D. D. Holm and T. M. Tyranowski. Stochastic discrete Hamiltonian variational integrators. *BIT Numerical Mathematics*, 58(4):1009–1048, 2018.

[84] P. Holmes, J. Lumley, G. Berkooz, and C. Rowley. *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge Monographs on Mechanics. Cambridge University Press, 2012.

[85] J. Hong and L. Sun. *Symplectic Integration of Stochastic Hamiltonian Systems*. Lecture Notes in Mathematics. Springer Singapore, 1st edition, February 2023.

[86] J. Hong, D. Xu, and P. Wang. Preservation of quadratic invariants of stochastic differential equations via Runge-Kutta methods. *Applied Numerical Mathematics*, 87:38 – 52, 2015.

[87] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. Kodansha scientific books. North-Holland, 1989.

[88] T. Iliescu, H. Liu, and X. Xie. Regularized reduced order models for a stochastic Burgers equation. *International Journal of Numerical Analysis and Modeling*, 15(4-5):594–607, 2018.

[89] K. Ito and S. Ravindran. A reduced-order method for simulation and control of fluid flows. *J. Comput. Phys.*, 143(2):403–425, jul 1998.

[90] K. Ito and S. Ravindran. Reduced basis method for optimal control of unsteady viscous flows. *International Journal of Computational Fluid Dynamics*, 15(2):97–113, 2001.

[91] J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel. Langevin stabilization of molecular dynamics. *The Journal of Chemical Physics*, 114(5):2090–2098, 2001.

[92] Y. Jung, E. Barkai, and R. J. Silbey. *A Stochastic Theory of Single Molecule Spectroscopy*, chapter 4, pages 199–266. John Wiley & Sons, Ltd, 2003.

[93] B. Karasözen and M. Uzunca. Energy preserving model order reduction of the nonlinear Schrödinger equation. *Advances in Computational Mathematics*, 44(6):1769–1796, 2018.

[94] R. Kleiber, R. Hatzky, A. Könies, K. Kauffmann, and P. Helander. An improved control-variate scheme for particle-in-cell simulations with collisions. *Computer Physics Communications*, 182(4):1005 – 1012, 2011.

[95] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics : Stochastic Modelling and Applied Probability. Springer, 1995.

[96] M. Kraus, T. Tyranowski, C. Albert, and C. Rackauckas. DDMGNI/GeometricIntegrators.jl: v0.2.0, Feb. 2020. https://doi.org/10.5281/zenodo.3648326.

[97] M. Kraus and T. M. Tyranowski. Variational integrators for stochastic dissipative Hamiltonian systems. *IMA Journal of Numerical Analysis*, 41(2):1318–1367, 2020.

[98] R. Kubo. Note on the stochastic theory of resonance absorption. *Journal of the Physical Society of Japan*, 9(6):935–944, 1954.

[99] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM Journal on Numerical Analysis*, 40(2):492–515, 2002.

[100] H. Kunita. *Stochastic Flows and Stochastic Differential Equations*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1997.

[101] A. M. Lacasta, J. M. Sancho, A. H. Romero, I. M. Sokolov, and K. Lindenberg. From subdiffusion to superdiffusion of particles on solid surfaces. *Phys. Rev. E*, 70:051104, Nov 2004.

[102] S. Lall, P. Krysl, and J. E. Marsden. Structure-preserving model reduction for mechanical systems. *Physica D: Nonlinear Phenomena*, 184(1):304–318, 2003. Complexity and Nonlinearity in Physical Systems – A Special Issue to Honor Alan Newell.

[103] P. S. Landa. Noise-induced transport of Brownian particles with consideration for their mass. *Phys. Rev. E*, 58:1325–1333, Aug 1998.

[104] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. *Model Order Reduction in Fluid Dynamics: Challenges and Perspectives*, pages 235–273. Springer International Publishing, 2014.

[105] J. A. Lázaro-Camí and J. P. Ortega. Stochastic Hamiltonian Dynamical Systems. *Reports on Mathematical Physics*, 61(1):65 – 122, 2008.

[106] T. Lelièvre and G. Stoltz. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numerica*, 25:681–880, 2016.

[107] F. Lu. Data-driven model reduction for stochastic Burgers equations. *Entropy*, 22(12), 2020.

[108] F. Lu, K. K. Lin, and A. J. Chorin. Data-based stochastic model reduction for the Kuramoto-Sivashinsky equation. *Physica D: Nonlinear Phenomena*, 340:46–57, 2017.

[109] J. L. Lumley. *The structure of inhomogeneous turbulent flows*, pages 166–178. Nauka, Moscow, 1967.

[110] Q. Ma, D. Ding, and X. Ding. Symplectic conditions and stochastic generating functions of stochastic Runge-Kutta methods for stochastic Hamiltonian systems with multiplicative noise. *Applied Mathematics and Computation*, 219(2):635–643, 2012.

[111] Q. Ma and X. Ding. Stochastic symplectic partitioned Runge-Kutta methods for stochastic Hamiltonian systems with multiplicative noise. *Appl. Math. Comput.*, 252(C):520–534, Feb. 2015.

[112] X. Mao. *Stochastic Differential Equations and Applications*. Elsevier Science, 2007.

[113] J. Marsden and T. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer Verlag, 1994.

[114] G. Milstein. *Numerical Integration of Stochastic Differential Equations*. Mathematics and Its Applications. Springer Netherlands, 1995.

[115] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov. Symplectic integration of Hamiltonian systems with additive noise. *SIAM J. Numer. Anal.*, 39(6):2066–2088, June 2001.

[116] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov. Numerical methods for stochastic systems preserving symplectic structures. *SIAM J. Numer. Anal.*, 40(4):1583 – 1604, 2002.

[117] T. Misawa. Symplectic integrators to stochastic Hamiltonian dynamical systems derived from composition methods. *Mathematical Problems in Engineering*, 2010. Vol. 2010, Article ID 384937, 12 pages.

[118] S. Mukamel. *Principles of nonlinear optical spectroscopy*. Oxford series in optical and imaging sciences. Oxford University Press, 1995.

[119] E. Nelson. Stochastic mechanics and random fields. In P.-L. Hennequin, editor, *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, pages 427–459. Springer Berlin Heidelberg, Berlin, Heidelberg, 1988.

[120] M. Ohlberger and S. Rave. Reduced basis methods: Success, limitations and future challenges. *Proceedings of the Conference Algoritmy*, pages 1–12, 2016.

[121] P. Parrilo, S. Lall, F. Paganini, G. Verghese, B. Lesieutre, and J. Marsden. Model reduction for analysis of cascading failures in power systems. In *Proceedings of the 1999 American Control Conference (Cat. No. 99CH36251)*, volume 6, pages 4208–4212 vol.6, 1999.

[122] M. B. Paskyabi, M. Krutova, F. G. Nielsen, J. Reuder, and O. E. Guernaoui. On stochastic reduced-order and LES-based models of offshore wind turbine wakes. *Journal of Physics: Conference Series*, 1669(1):012018, oct 2020.

[123] B. Peherstorfer and K. Willcox. Online adaptive model reduction for nonlinear systems via low-rank updates. *SIAM Journal on Scientific Computing*, 37(4):A2123–A2150, 2015.

[124] L. Peng and K. Mohseni. Geometric model reduction of forced and dissipative Hamiltonian systems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 7465–7470, 2016.

[125] L. Peng and K. Mohseni. Structure-preserving model reduction of forced Hamiltonian systems. Unpublished, arXiv:1603.03514, 2016.

[126] L. Peng and K. Mohseni. Symplectic model reduction of Hamiltonian systems. *SIAM Journal on Scientific Computing*, 38(1):A1–A27, 2016.

[127] J. S. Peterson. The reduced basis method for incompressible viscous flow calculations. *SIAM Journal on Scientific and Statistical Computing*, 10(4):777–786, 1989.

[128] D. Pfirsch. New variational formulation of Maxwell-Vlasov and guiding center theories local charge and energy conservation laws. *Zeitschrift für Naturforschung A*, 39(1):1–8, 1984.

[129] D. Pfirsch and P. J. Morrison. Local conservation laws for the Maxwell-Vlasov and collisionless kinetic guiding-center theories. *Phys. Rev. A*, 32:1714–1721, Sep 1985.

[130] A. Pinkus. *N-widths in Approximation Theory*. Ergebnisse der Mathematik und ihrer Grenzgebiete : a series of modern surveys in mathematics. Folge 3. Springer-Verlag, 1985.

[131] R. V. Polyuga and A. van der Schaft. Structure preserving model reduction of port-Hamiltonian systems by moment matching at infinity. *Automatica*, 46(4):665–672, 2010.

[132] S. Prajna. POD model reduction with stability guarantee. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, volume 5, pages 5254–5258 Vol.5, 2003.

[133] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations: An Introduction*. UNITEXT. Springer International Publishing, 2015.

[134] K. Rasmussen, Y. Gaididei, O. Bang, and P. Christiansen. The influence of noise on critical collapse in the nonlinear Schrödinger equation. *Physics Letters A*, 204(2):121–127, 1995.

[135] M. Rathinam and L. R. Petzold. A new look at proper orthogonal decomposition. *SIAM Journal on Numerical Analysis*, 41(5):1893–1925, 2003.

[136] R. Reigada, A. H. Romero, A. Sarmiento, and K. Lindenberg. One-dimensional arrays of oscillators: Energy localization in thermal equilibrium. *The Journal of Chemical Physics*, 111(4):1373–1384, 1999.

[137] J. Reiss, P. Schulze, J. Sesterhenn, and V. Mehrmann. The shifted proper orthogonal decomposition: A mode decomposition for multiple transport phenomena. *SIAM Journal on Scientific Computing*, 40(3):A1322–A1344, 2018.

[138] D. Rim, B. Peherstorfer, and K. T. Mandli. Manifold approximations via transported subspaces: Model reduction for transport-dominated problems. *SIAM Journal on Scientific Computing*, 45(1):A170–A199, 2023.

[139] M. Ripoll, M. H. Ernst, and P. Español. Large scale and mesoscopic hydrodynamics for dissipative particle dynamics. *The Journal of Chemical Physics*, 115(15):7271–7284, 2001.

[140] C. W. Rowley, T. Colonius, and R. M. Murray. Model reduction for compressible flows using POD and Galerkin projection. *Physica D: Nonlinear Phenomena*, 189(1):115–129, 2004.

[141] J. Sanz-Serna and A. Stuart. Ergodicity of dissipative differential equations subject to random impulses. *Journal of Differential Equations*, 155(2):262 – 284, 1999.

[142] B. L. Sawford. Turbulent relative dispersion. *Annu. Rev. Fluid Mech.*, 33:289–317, 2001.

[143] M. Seeßelberg, H. P. Breuer, H. Mais, F. Petruccione, and J. Honerkamp. Simulation of one-dimensional noisy Hamiltonian systems and their application to particle storage rings. *Zeitschrift für Physik C Particles and Fields*, 62(1):63–73, 1994.

[144] T. Shardlow. Splitting for dissipative particle dynamics. *SIAM Journal on Scientific Computing*, 24(4):1267–1282, 2003.

[145] R. D. Skeel. Integration schemes for molecular dynamics and related applications. In M. Ainsworth, J. Levesley, and M. Marletta, editors, *The Graduate Student's Guide to Numerical Analysis '98: Lecture Notes from the VIII EPSRC Summer School in Numerical Analysis*, pages 119–176. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[146] C. Soize. *The Fokker-Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solutions*. Advanced Series on Fluid Mechanics. World Scientific, 1994.

[147] E. Sonnendrücker, A. Wacher, R. Hatzky, and R. Kleiber. A split control variate scheme for PIC simulations with collisions. *Journal of Computational Physics*, 295:402 – 419, 2015.

[148] H. Sugama. Gyrokinetic field theory. *Physics of Plasmas*, 7(2):466–480, 2000.

[149] C. Sulem and P. Sulem. *The Nonlinear Schrödinger Equation: Self-Focusing and Wave Collapse*. Applied Mathematical Sciences. Springer New York, 2007.

[150] L. Sun and L. Wang. Stochastic symplectic methods based on the Padé approximations for linear stochastic Hamiltonian systems. *Journal of Computational and Applied Mathematics*, 2016. http://dx.doi.org/10.1016/j.cam.2016.08.011.

[151] D. Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8(2):163–198, 2002.

[152] D. J. Thomson. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, 180:529–556, 1987.

[153] H. C. Tuckwell. Numerical solutions of some stochastic hyperbolic wave equations including sine-Gordon equation. *Wave Motion*, 65:130–146, 2016.

[154] T. M. Tyranowski. Stochastic variational principles for the collisional Vlasov-Maxwell and Vlasov-Poisson equations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477(2252):20210167, 2021.

[155] T. M. Tyranowski and M. Desbrun. R-adaptive multisymplectic and variational integrators. *Mathematics*, 7(7), 2019.

[156] T. M. Tyranowski and M. Desbrun. Variational partitioned Runge-Kutta methods for Lagrangians linear in velocities. *Mathematics*, 7(9), 2019.

[157] T. M. Tyranowski and M. Kraus. Symplectic model reduction methods for the Vlasov equation. *Contributions to Plasma Physics*, page e202200046, 2022.

[158] N. Van Kampen. Stochastic differential equations. *Physics Reports*, 24(3):171 – 228, 1976.

[159] E. vanden Eijnden and A. Grecos. Stochastic modelling of turbulence and anomalous transport in plasmas. *Journal of Plasma Physics*, 59(4):683–694, June 1998.

[160] L. Wang. *Variational Integrators and Generating Functions for Stochastic Hamiltonian Systems*. PhD thesis, Karlsruhe Institute of Technology, 2007.

[161] L. Wang and J. Hong. Generating functions for stochastic symplectic methods. *Discrete and Continuous Dynamical Systems*, 34(3):1211–1228, 2014.

[162] P. Wang and G. Chen. Effective approximation of stochastic sine-Gordon equation with a fast oscillation. *Journal of Mathematical Physics*, 62(3):032702, 2021.

[163] P. Wang, J. Hong, and D. Xu. Construction of symplectic Runge-Kutta methods for stochastic Hamiltonian systems. *Communications in Computational Physics*, 21(1):237–270, 2017.

[164] X. Xie, F. Bao, and C. G. Webster. Evolve filter stabilization reduced-order model for stochastic Burgers equation. *Fluids*, 3(4), 2018.

[165] V. E. Zakharov and S. Manakov. On the complete integrability of a nonlinear Schrödinger equation. *Theoretical and Mathematical Physics*, 19:551–559, 1974.

[166] W. Zhou, J. Zhang, J. Hong, and S. Song. Stochastic symplectic Runge-Kutta methods for the strong approximation of Hamiltonian systems with additive noise. *Journal of Computational and Applied Mathematics*, 325:134 – 148, 2017.