Spectral methods for multiscale stochastic differential equations*

A. Abdulle†, G.A. Pavliotis‡, and U. Vaes§

Abstract. This paper presents a new method for the solution of multiscale stochastic differential equations at the diffusive time scale. In contrast to averaging-based methods, e.g., the heterogeneous multiscale method (HMM) or the equation-free method, which rely on Monte Carlo simulations, in this paper we introduce a new numerical methodology that is based on a spectral method. In particular, we use an expansion in Hermite functions to approximate the solution of an appropriate Poisson equation, which is used in order to calculate the coefficients of the homogenized equation. Spectral convergence is proved under suitable assumptions. Numerical experiments corroborate the theory and illustrate the performance of the method. A comparison with the HMM and an application to singularly perturbed stochastic PDEs are also presented.

Key words. Spectral methods for differential equations, Hermite spectral methods, singularly perturbed stochastic differential equation, multiscale methods, homogenization theory, stochastic partial differential equations.

AMS subject classifications. 65N35, 65C30, 60H10, 60H15

1. Introduction.

Multiscale stochastic systems arise frequently in applications. Examples include atmosphere/ocean science [45] and materials science [19]. For systems with a clear scale separation it is possible, in principle, to obtain a closed – averaged or homogenized – equation for the slow variables [56]. The calculation of the drift and diffusion coefficients that appear in this effective (coarse-grained) equation requires appropriate averaging over the fast scales. Several numerical methods for multiscale stochastic systems that are based on scale separation and on the existence of a coarse-grained equation for the slow variables have been proposed in the literature. Examples include the heterogeneous multiscale method (HMM) [62, 21, 1] and the equation-free approach [36]. These techniques are based on evolving the coarse-grained dynamics, while calculating the drift and diffusion coefficients "on-the-fly" using short simulation bursts of the fast dynamics.

A prototype fast/slow system of stochastic differential equations (SDEs) for which the aforementioned techniques can be applied is [1]

$$dX_t^\varepsilon = \frac{1}{\varepsilon}\mathbf{f}(X_t^\varepsilon, Y_t^\varepsilon)\,dt + \sqrt{2}\,\boldsymbol{\sigma}_x\,dW_{xt}, \tag{1a}$$

*Submitted to the editors September 16, 2016.

Funding: A. Abdulle is supported by the Swiss national foundation. G.A. Pavliotis is supported by the Engineering and Physical Sciences Research Council of the UK through Grants Nos. EP/L020564, EP/L024926 and EP/L025159. U. Vaes is supported through a Roth PhD studentship by the Department of Mathematics, Imperial College London.

†Mathematics Section, École Polytechnique Fédérale de Lausanne (assyr.abdulle@epfl.ch).

‡Department of Mathematics, Imperial College London (g.pavliotis@imperial.ac.uk).

§Department of Mathematics, Imperial College London (u.vaes13@imperial.ac.uk).

[1] In this paper we will consider the fast/slow dynamics at the diffusive time scale, or, using the terminology of [56], the homogenization problem.

$$dY_t^\varepsilon = \frac{1}{\varepsilon^2}\mathbf{h}(X_t^\varepsilon, Y_t^\varepsilon)\,dt + \frac{\sqrt{2}}{\varepsilon}\boldsymbol{\sigma}_y\,dW_{yt}. \tag{1b}$$

where $X_t^\varepsilon \in \mathbf{R}^m$, $Y_t^\varepsilon \in \mathbf{R}^n$, $\varepsilon \ll 1$ is the parameter measuring scale separation, $\boldsymbol{\sigma}_x \in \mathbf{R}^{m\times d_1}$, $\boldsymbol{\sigma}_y \in \mathbf{R}^{n\times d_2}$ are constant matrices, and W_x, W_y are independent d_1 and d_2-dimensional Brownian motions, respectively.[2] For fast-slow systems of this form, a direct numerical approximation of the full dynamics would be prohibitively expensive, because resolving the fine scales would require a time step δt that scales as $\mathcal{O}(\varepsilon^2)$. Under appropriate assumptions on the coefficients and on the ergodic properties of the fast process Y_t^ε, it is well known that the slow process converges, in the limit as ε tends to 0, to a homogenized equation that is independent of the fast process and of ε [56, Ch. 11]:

$$dX_t = \mathbf{F}(X_t)\,dt + \mathbf{A}(X_t)\,dW_t. \tag{2}$$

The drift and diffusion coefficients in (2) can be calculated by solving a Poisson equation involving the generator of the fast process,[3]

$$-\mathcal{L}_y\boldsymbol{\phi} = \mathbf{f}, \tag{3}$$

where $\mathcal{L}_y = \mathbf{h}(x,y)\cdot\nabla_y + \boldsymbol{\sigma}_y\boldsymbol{\sigma}_y^T : \nabla_y\nabla_y$, together with appropriate boundary conditions, and calculating averages with respect to the invariant measure $\mu_x(dy)$ of Y_t^ε:

$$\mathbf{F}(x) = \int_{\mathbf{R}^n} \nabla_x\boldsymbol{\phi}(x,y)\,\mathbf{f}(x,y)\,\mu_x(dy), \tag{4a}$$

$$\mathbf{A}(x)\mathbf{A}(x)^T = \int_{\mathbf{R}^n} \left[\mathbf{f}(x,y)\otimes\boldsymbol{\phi}(x,y) + \boldsymbol{\phi}(x,y)\otimes\mathbf{f}(x,y)\right]\,\mu_x(dy). \tag{4b}$$

Once the drift and diffusion coefficients have been calculated, then it becomes computationally advantageous to solve the homogenized equations, in particular since we are usually interested in the evolution of observables of the slow process alone. The main computational task, thus, is to calculate the drift and diffusion coefficients that appear in the homogenized equation (2). When the state space of the fast process is high dimensional, the numerical solution of the Poisson equation and calculation of the integrals in (3) using deterministic methods become prohibitively expensive and Monte Carlo-based approaches have to be employed. In recent years different methodologies have been proposed for the numerical solution of the fast-slow system (1) that are based on the strategy outlined above, for example the Heterogeneous Multiscale Method (HMM) [62, 21, 1] and the equation-free approach [36]. In particular, the PDE-based formulas (4) are replaced by Green-Kubo type formulas [21, Sec. 1] that involve time averages and numerically calculated autocorrelation functions. The equivalence between the homogenization and the Green-Kubo formalism has been shown for a quite general class of fast/slow systems of SDEs [54]. See also [39, 41]. While offering several advantages, time and ensemble averages, on which these methods are based, imply that accurate solutions are

[2] It is straightforward to consider problems where the Brownian motions driving the fast and slow processes are correlated. This scenario might be relevant in applications to mathematical finance. See e.g. [16].

[3] We are assuming that the centering condition is satisfied, see (10) below.

computationally very expensive to obtain. Based on the analysis of [21], one deduces that the computational cost needed to obtain an error of order 2^{-p} scales as $\mathcal{O}(2^{p(2+1/l)})$, where l is the weak order of accuracy of the micro-solver used.

When the dimension of the state space of the fast process is relatively low, numerical approaches that are based on the accurate and efficient numerical solution of the Poisson equation (3) using "deterministic" techniques become preferable. Such an approach was taken in [9] for the study of the diffusion approximation of a kinetic model for swarming [12]. The computational methodology that was introduced and analyzed is based on the numerical calculation of the eigenvalues and eigenfunctions of an appropriate Schrödinger operator using a high-order finite element method. The authors showed rigorously and by means of numerical experiments that for sufficiently smooth interaction potentials the proposed numerical scheme performs extremely well; in particular, the numerical calculation of the first few eigenvalues and eigenfunctions are sufficient for the very accurate calculation of the drift and diffusion coefficients.

Methods that are based on the caclculation of the eigenvalues and eigenfunctions of the transfer (Koopman) operator have been introduced in [24, 25]. In addition diffusion maps have also been applied to multiscale stochastic systems [14, 13]. Techniques that are based on the transfer operator and diffusion maps are related to our approach since, unlike the HMM and the equation-free approach, they do not require the solution of the fast dynamics. On the other hand, it should be noted that these techniques are, in principle, more powerful than the methodology developed in this paper, since they enable us to identify the coarse-grained variables.

In this paper we develop further the methodology introduced in [9] and we apply it to the numerical solution of fast/slow systems of SDEs, including singularly perturbed stochastic partial differential equations (SPDEs) in bounded domains. Thus, we complement the work presented in [2], in which a hybrid HMM/spectral method for the numerical solution of singularly perturbed SPDEs with quadratic nonlinearities [7] at the diffusive time scale was developed.[4] The main difference between the methodology presented in [9] and the approach we take in this paper is that, rather than obtaining the orthonormal basis by solving the eigenvalue problem for an appropriate Schrödinger operator, we fix the orthonormal basis (Hermite functions) and expand the solution of the Poisson equation (3) (after the unitary transformation that maps it to an equation for a Schrödinger operator) in this basis. We show rigorously and by means of numerical experiments that our proposed methodology achieves spectral convergence for a wide class of fast processes in (1). Consequently, our method outperforms Monte Carlo-based methodologies such as the HMM and the equation-free method, at least for problems with low-dimensional fast processes. We discuss how our method can be modified so that it becomes efficient when the fast process has a high-dimensional state space in the conclusions section, Section 7.

In this paper we will consider fast/slow systems of SDEs for which the fast process is

[4] When the centering condition (see Equation (10)) is not satisfied, one needs to study the problem at a shorter time scale (called the advective time scale) before considering the diffusive limit. This problem is easier to study since it does not require the solution of a Poisson equation. The rigorous analysis of the HMM method for singularly perturbed SPDEs at the advective time scale was presented in [11].

reversible, i.e. it has a gradient structure [55, Sec. 4.8][5]

$$dX_t^\varepsilon = \frac{1}{\varepsilon}\mathbf{f}(X_t^\varepsilon, Y_t^\varepsilon)dt + \boldsymbol{\alpha}(X_t^\varepsilon, Y_t^\varepsilon)\, dW_{xt}, \quad X_0^\varepsilon = x_0, \tag{5a}$$

$$dY_t^\varepsilon = -\frac{1}{\varepsilon^2}\nabla V(Y_t^\varepsilon)dt + \frac{\sqrt{2}}{\varepsilon}dW_{yt}, \quad Y_0^\varepsilon = y_0, \tag{5b}$$

where $X_t^\varepsilon(t) \in \mathbf{R}^m$, $Y_t^\varepsilon(t) \in \mathbf{R}^n$, $\boldsymbol{\alpha}(\cdot,\cdot) \in \mathbf{R}^{m\times p}$, W_x and W_y are standard p and n-dimensional Brownian motions, and $V(\cdot)$ is a smooth confining potential.

Remark 1. *Our algorithm applies to the more general situation when the fast dynamics depends on the slow one in the following way*

$$dY_t^\varepsilon = -\frac{1}{\varepsilon^2}\nabla V(Y_t^\varepsilon)dt + \frac{1}{\varepsilon}\mathbf{h}(X_t^\varepsilon(t), Y_t^\varepsilon(t))\, dt + \frac{\sqrt{2}}{\varepsilon}dW_{yt}, \quad Y_0^\varepsilon = y_0. \tag{6}$$

The analysis presented in this paper also generalises with minor changes for such a fast process and the effective drift (4a) becomes

$$\mathbf{F}(x) = \int_{\mathbf{R}^n} \nabla_x \boldsymbol{\phi}(x,y)\, \mathbf{f}(x,y)\, \mu_x(dy) + \int_{\mathbf{R}^n} \nabla_y \boldsymbol{\phi}(x,y)\, \mathbf{h}(x,y)\, \mu_x(dy). \tag{7}$$

Hence the additional term appearing above involves the solution of a Poisson problem already computed.

SDEs of this form appear in several applications, e.g. in molecular dynamics [18, 40]. Furthermore, several interesting semilinear singularly perturbed SPDEs can be written in this form, see Section 6. It is well known [55, Sec. 4.9] that the generator of a reversible SDE is unitarily equivalent to an appropriate Schrödiner operator. Our approach is to first solve this Poisson equation for the Schrödinger operator via a spectral method using Hermite functions and then use this solution in order to calculate the integrals in (4). For smooth potentials that increase sufficiently fast at infinity our method has spectral accuracy, i.e. the error decreases faster than any negative power of the number of floating point operations performed. This, in turn, via a comparison for SDEs argument, implies that we can approximate very accurately the evolution of observables of the slow variable X_t^ε in (5) by solving an approximate homogenized equation in which the drift and diffusion coefficients are calculated using our spectral method. For relatively low dimensional fast-processes, this leads to a much more accurate and computationally efficient numerical method than any Monte Carlo-based methodology. We remark that our proposed numerical methodology becomes (analytically) exact when the fast process is, to leading order, an Ornstein-Uhlenbeck process, since in this case, for a suitable choice of the mean and the covariance matrix, the Hermite functions are the eigenfunctions of the corresponding Schrödinger operator.

The rest of the paper is organized as follows. In Section 2, we summarize the results from homogenization theory for the fast/slow system (5) that we will need in this work. In Section 3 we present our numerical method in an algorithmic manner. In Section 4, we summarize the

[5] We could, in principle, also consider reversible SDEs with a diffusion tensor that is not a multiple of the identity.

main theoretical results of this paper; in particular we show that our method, under appropriate assumptions on the coefficients of the fast/slow system, is spectrally accurate. The proofs of our main results are given in Section 5. In Section 6 we present details on the implementation of our numerical method, discuss the computational efficiency and present several numerical examples, including an example of the numerical solution of a singularly perturbed SPDE; for this example, we also present a brief qualitative comparison of our method with the HMM method. Section 7 is reserved for conclusions and discussion of further work. Finally in the appendices we present some results related to approximation theory in weighted Sobolev spaces that are needed in the proof of the main convergence theorem.

2. Diffusion Approximation and Homogenization. In this section, we summarize some of our working hypotheses and the results from the theory of homogenization used to derive the effective SDE for the system (5). Throughout this paper, the notation $|\cdot|$ denotes the Euclidian norm when applied to vectors or the 2-norm when applied to matrices. In addition, we denote the components of a vector $\mathbf{v} \in \mathbf{R}^d$ by $v_1, v_2 \cdots, v_d$. We start by assuming that $V(\cdot)$ is a smooth confining potential, [55, Definition 4.2]:

$$V \in C^\infty(\mathbf{R}^n), \qquad \lim_{|y|\to\infty} V(y) = \infty \quad \text{and} \quad e^{-V(\cdot)} \in L^1\left(\mathbf{R}^n\right). \tag{8}$$

These hypotheses guarantee that the fast process has a well defined solution for all positive times, with a unique invariant measure whose density is given by $\frac{1}{\mathcal{Z}} e^{-V(\cdot)}$, where $\mathcal{Z}$ is the normalization constant. Without loss of generality, we may assume that $\mathcal{Z} = 1$. To these assumptions, we add

$$\lim_{|y|\to\infty} \nabla V \cdot \left(\frac{y}{|y|}\right) = \infty \quad \text{and} \quad \lim_{|y|\to\infty} W(y) := \lim_{|y|\to\infty} \left(\frac{1}{4}|\nabla V(y)|^2 - \frac{1}{2}\Delta V(y)\right) = \infty, \tag{9}$$

which guarantee that the law of Y_t^ε converges to its invariant distribution e^{-V} exponentially fast (e.g. in relative entropy), see [47]. We assume furthermore that the drift coefficient in the slow equation of system (5) satisfies

$$\begin{aligned} &\mathbf{f}(x,y) \in \left(C^\infty(\mathbf{R}^m \times \mathbf{R}^n)\right)^m, \\ &\int_{\mathbf{R}^n} \mathbf{f}(x,y)\, e^{-V(y)}\, dy = \mathbf{0}, \text{ and} \\ &|\mathbf{f}(x,y)| \le K(1+|x|)(1+|y|^\beta) \quad \forall x \in \mathbf{R}^m \text{ and } \forall y \in \mathbf{R}^n, \end{aligned} \tag{10}$$

with β a positive integer and K a positive constant. Under Assumptions (8) and (9), we can use Proposition 17 and Lax-Milgram theorem to show that there exists for all $x \in \mathbf{R}^m$ a unique $\boldsymbol{\phi}(x,\cdot) \in H^1\left(\mathbf{R}^n, e^{-V}\right)^m$, where $H^1\left(\mathbf{R}^n, e^{-V}\right)$ is the weighted Sobolev space defined in Definition 15, such that

$$-\mathcal{L}_y \phi_i(x,y) := -\left(\Delta_y - \nabla_y V \cdot \nabla_y\right)\phi_i(x,y) = f_i(x,y) \quad \text{for } i = 1,\ldots,m \tag{11}$$

and

$$\int_{\mathbf{R}^n} \boldsymbol{\phi}(x,y)\, e^{-V(y)}\, dy = \mathbf{0}, \quad \forall x \in \mathbf{R}^m. \tag{12}$$

By the theory of elliptic regularity, the uniform ellipticity of $\mathcal{L}$ implies that $\boldsymbol{\phi}$ is smooth in y, and by [51, Theorem 1], $\boldsymbol{\phi}$ grows at most polynomially in y.

We also make the following assumption on the Lipschitz continuity with respect to x of the coefficients,

$$\left|\mathbf{f}(x,y)-\mathbf{f}(x',y)\right|+\left|\boldsymbol{\alpha}(x,y)-\boldsymbol{\alpha}(x',y)\right|\leq C(y)\left|x-x'\right|,\quad \forall x,x'\in\mathbf{R}^m, \tag{13}$$

with $C(y)$ a function bounded from above by a polynomial, and the following assumptions on the growth of the coefficients:

$$\begin{aligned}&|\nabla_x\mathbf{f}(x,y)|+\left|\nabla_x^2\mathbf{f}(x,y)\right|\leq K(1+|y|^{m_1}),\\ &|\boldsymbol{\alpha}(x,y)|\leq K(1+|x|^{1/2})(1+|y|^{m_2}),\end{aligned} \tag{14}$$

for positive integers m_1, m_2 and a positive constant K. It follows from this that $\boldsymbol{\phi}(\cdot,y)$ belongs to $\left(C^2(\mathbf{R}^m)\right)^m$ for all values of y. This can be shown by using the Feynman-Kac representation of the solution of (11) that was studied in [51]:

$$\phi_i(x,y)=\int_0^\infty \mathbf{E}_y f_i(x,z_t^y)\,dt,\qquad i=1,\dots,m, \tag{15}$$

where z_t^y is the solution of

$$dz_t^y=-\nabla_y V(z_t^y)\,dt+\sqrt{2}\,dW_t\quad\text{ with }\quad z_0^y=y.$$

Using the Feynman-Kac formula (15), one can show [51, p. 1073] that there exist $L,q>0$ such that:

$$\begin{aligned}&|\boldsymbol{\phi}(x,y)|+|\nabla_y\boldsymbol{\phi}(x,y)|\leq L(1+|x|)(1+|y|^q),\\ &|\nabla_x\boldsymbol{\phi}(x,y)|+|\nabla_y\nabla_x\boldsymbol{\phi}(x,y)|+|\nabla_x\nabla_x\boldsymbol{\phi}(x,y)|+|\nabla_y\nabla_x\nabla_x\boldsymbol{\phi}(x,y)|\leq L(1+|y|^q).\end{aligned} \tag{16}$$

Using the previous assumptions we can prove the following homogenization/diffusion approximation result [51, Theorem 3].

Theorem 2. *Let (8)–(10), (13) and (14) be satisfied. Then for any $T>0$, the family of processes $\{X_t^\varepsilon, 0\leq t\leq T\}$ solving (5) is weakly relatively compact in $(C([0,T]))^m$. Any accumulation point X_t is a solution of the martingale problem associated to the operator:*

$$\mathcal{G}=\frac{1}{2}\mathbf{D}(x):\nabla_x\nabla_x+\mathbf{F}(x)\cdot\nabla_x$$

where

$$\mathbf{F}(x)=\int_{\mathbf{R}^n}\nabla_x\boldsymbol{\phi}(x,y)\,\mathbf{f}(x,y)\,e^{-V(y)}\,dy, \tag{17}$$

$$\mathbf{D}(x)=\int_{\mathbf{R}^n}\left(\boldsymbol{\alpha}(x,y)\boldsymbol{\alpha}(x,y)^T+\mathbf{f}(x,y)\otimes\boldsymbol{\phi}(x,y)+\boldsymbol{\phi}(x,y)\otimes\mathbf{f}(x,y)\right)\,e^{-V(y)}\,dy, \tag{18}$$

and $\boldsymbol{\phi}(x,y)$ is the centered solution of the Poisson equation (11). If, moreover, the martingale problem associated to $\mathcal{G}$ is well-posed, then $X_t^\varepsilon\Rightarrow X_t$ (convergence in law), where X_t is the unique diffusion process (in law) with generator $\mathcal{G}$.

The matrix $\mathbf{D}(x)$ is clearly symmetric, and one can show that it is also positive semi-definite, see [56, Section 11.5]. In view of Theorem 2, writing $\mathbf{D}(x) = \mathbf{A}(x)\mathbf{A}(x)^T$ we obtain the functions $\mathbf{F}(x)$, $\mathbf{A}(x)$ that appear in the homogenized SDE (2). Though the choice of $\mathbf{A}(x)$ is not unique, all choices lead to the same associated Fokker-Planck or backward Kolmogorov equations, hence the probability measures in path space are the same.

3. Numerical Method. In this section, we describe our method for the approximation of the effective dynamics, the analysis of which is postponed to Section 5. We start by introducing the necessary notation. We will denote by $L^2(\mathbf{R}^n)$ the space of square integrable functions on $\mathbf{R}^n$. The notation $L^2(\mathbf{R}^n, \rho)$, for a probability density ρ, will be used to denote the space of functions f such that $\sqrt{\rho} f \in L^2(\mathbf{R}^n)$. Weighted Sobolev spaces associated to a probability density are defined in Definition 15. whereas scales of Sobolev spaces, associated to an operator, are defined in Definition 16.

In addition to these function spaces, we will denote by $\mathbf{P}_d(\mathbf{R}^n)$ the space of polynomials in n variables of degree less than or equal to d, and by $H_\alpha(y; \mu, \Sigma)$ the Hermite polynomials on $\mathbf{R}^n$ defined in Appendix B:

$$H_\alpha(y; \mu, \Sigma) = H_\alpha(S^{-1}(y-\mu); 0, I), \text{ with } H_\alpha(z; 0, I) = \prod_{k=1}^n H_{\alpha_k}(z_k) \tag{19}$$

and $\alpha \in \mathbf{N}^n$. Here $H_{\alpha_k}(\cdot)$ denotes the one-dimensional Hermite polynomial of degree α_k, $\Sigma \in \mathbf{R}^{n\times n}$ is a symmetric positive definite matrix, D and Q are diagonal and orthogonal matrices such that $\Sigma = QDQ^T$, $S = QD^{1/2}$ and $\mu \in \mathbf{R}^n$. We recall from Appendix B that these polynomials form a complete orthonormal basis of $L^2(\mathbf{R}^n, G_{(\mu,\Sigma)})$, where $G_{\mu,\Sigma}$ denotes the Gaussian density on $\mathbf{R}^n$ with mean μ and covariance matrix Σ. Finally, we will use the notation $h_\alpha(y; \mu, \Sigma)$ to denote the Hermite functions corresponding to the Hermite polynomials (19), see Definition 21.

We recall from Section 2 that obtaining the drift and diffusion coefficients $\mathbf{F}(X)$ and $\mathbf{A}(X)$, respectively, of the homogenized equation

$$dX = \mathbf{F}(X)\, dt + \mathbf{A}(X)\, dW_t, \tag{20}$$

requires the solution of the Poisson equations (11). To emphasize the fact that x appears as a parameter in (11), we will use the notations $\boldsymbol{\phi}^x(\cdot) := \boldsymbol{\phi}(x, \cdot)$ and $\mathbf{f}^x(\cdot) := \mathbf{f}(x, \cdot)$. The weak formulation of the Poisson equation (11) is to find $\boldsymbol{\phi}^x \in H^1\left(\mathbf{R}^n, e^{-V}\right)^m$ such that for $i = 1, \ldots, m$,

$$a_V(\phi_i^x, v) := \int_{\mathbf{R}^n} \nabla\phi_i^x \cdot \nabla v\, e^{-V} dy = \int_{\mathbf{R}^n} f_i^x\, v\, e^{-V} dy \qquad \forall v \in H^1\left(\mathbf{R}^n, e^{-V}\right), \tag{21}$$

with the centering condition

$$\mathcal{M}(\boldsymbol{\phi}^x) := \int_{\mathbf{R}^n} \boldsymbol{\phi}^x\, e^{-V}\, dy = \mathbf{0}. \tag{22}$$

We recall that in order to be well-posed the condition $\mathcal{M}(\mathbf{f}^x) = 0$ must be satisfied.

We begin by performing the standard unitary transformation that maps the generator of a reversible Markov process to a Schrödinger operator: $e^{-V/2} : L^2\left(\mathbf{R}^n, e^{-V}\right) \to L^2\left(\mathbf{R}^n\right)$. Introducing

$$\mathcal{H} := e^{-V/2}\mathcal{L}_y\left(e^{V/2}\cdot\right) = \Delta - \left(\frac{1}{4}|\nabla V|^2 - \frac{1}{2}\Delta V\right) = \Delta - W(y), \tag{23}$$

and $\boldsymbol{\psi}^x = e^{-V/2}\boldsymbol{\phi}^x$, the Poisson equation (11) can be rewritten in terms of the operator (23) as:

$$-\mathcal{H}\boldsymbol{\psi}^x = e^{-V/2}\mathbf{f}^x. \tag{24}$$

Note that, because the superscript x already indicates that x plays the role of a parameter, we have have not appended the subscript y to $\mathcal{H}$. The weak formulation of this mapped problem reads: find $\boldsymbol{\psi}^x \in H^1\left(\mathbf{R}^n, \mathcal{H}\right)$ such that, for $i = 1, \ldots, m$,

$$a(\psi_i^x, v) := \int_{\mathbf{R}^n} \nabla\psi_i^x \cdot \nabla v + W(y)\,\psi_i^x\, v\, dy = \int_{\mathbf{R}^n} f_i^x\, v\, e^{-V/2} dy \qquad \forall v \in H^1\left(\mathbf{R}^n, \mathcal{H}\right), \tag{25}$$

where $H^1\left(\mathbf{R}^n, \mathcal{H}\right) = \left\{u \in H^1\left(\mathbf{R}^n\right) : \int_{\mathbf{R}^n} |W|\, u^2\, dy < \infty\right\}$, and such that the following centering condition is satisfied:

$$\hat{\mathcal{M}}(\boldsymbol{\psi}^x) := \int_{\mathbf{R}^n} \boldsymbol{\psi}^x e^{-V/2} dy = 0. \tag{26}$$

The formulas for the effective drift and diffusion coefficients can be written as

$$\mathbf{F}(x) = \int_{\mathbf{R}^n} \nabla_x \boldsymbol{\psi}^x \left(\mathbf{f}^x\, e^{-V/2}\right) dy, \tag{27a}$$

$$\mathbf{D}(x) = \int_{\mathbf{R}^n} \boldsymbol{\alpha}\boldsymbol{\alpha}^T(x,y)\,\mu^x(dy) + \mathbf{A}_0(x) + \mathbf{A}_0(x)^T, \tag{27b}$$

where

$$\mathbf{A}_0(x) = \int_{\mathbf{R}^n} \boldsymbol{\psi}^x \otimes \left(\mathbf{f}^x\, e^{-V/2}\right) dy. \tag{28}$$

The advantage of using the unitary transformation is that the solution of this new problem and its derivative lie in $L^2\left(\mathbf{R}^n\right)$, rather than in a weighted space.

To approximate numerically the coefficients of the effective SDE, we choose a finite-dimensional subspace $\hat{S}_d$ of $H^1\left(\mathbf{R}^n, \mathcal{H}\right)$, specified below, and consider the finite-dimensional approximation problem: find $\boldsymbol{\psi}_d^x \in (\hat{S}_d)^m$ such that, for $i = 1, \ldots, m$,

$$a(\psi_{di}^x, v_d) = \int_{\mathbf{R}^n} f_i^x\, v_d\, e^{-V/2} dy \qquad \forall v_d \in \hat{S}_d. \tag{29}$$

Existence and uniqueness, up to a function in the kernel of $\mathcal{H}$, if that space is contained in $\hat{S}_d$, of the solution to the finite-dimensional problem are inherited from the infinite-dimensional problem (25).

For a given basis $\{e_\alpha\}_{|\alpha|\le d}$ of $\hat{S}_d$, the finite-dimensional approximation of $\boldsymbol{\psi}^x$ can be expanded as $\boldsymbol{\psi}^x_d = \sum_{|\alpha|\le d} \boldsymbol{\psi}^x_\alpha\, e_\alpha$, and from the variational formulation (29) we obtain the following linear systems:

$$\sum_{|\beta|\le d} a(e_\alpha,\, e_\beta)\, \boldsymbol{\psi}^x_\beta = \mathbf{f}^x_\alpha \quad \text{with} \quad \mathbf{f}^x_\alpha = \int_{\mathbf{R}^n} \mathbf{f}^x\, e_\alpha\, e^{-V/2}\, dy. \tag{30}$$

We will use the notation $A_{\alpha\beta} = a(e_\alpha, e_\beta)$ for the stiffness matrix. In view of formula (27) we see that we also need an approximation of the gradient of the solution, which we denote by $(\nabla_x \boldsymbol{\psi}^x)_d$ and decompose in the basis as $(\nabla_x \boldsymbol{\psi}^x)_d = \sum_{|\alpha|\le d} (\nabla_x \boldsymbol{\psi}^x)_\alpha\, e_\alpha$. This approximation can be obtained by solving (30) with the right-hand side $(\nabla_x \mathbf{f}^x)_\alpha = \int_{\mathbf{R}^n} (\nabla_x \mathbf{f}^x)\, e_\alpha\, e^{-V/2}\, dy$. In practice, we solved (30) by using the function *solve* of the C++ linear algebra library Armadillo. As $d \to \infty$, the lowest eigenvalue of $A_{\alpha\beta}$ tends to zero. This could in principle induce numerical instabilities, because the system becomes ill-conditioned, but didn't cause any problem in the numerical experiments we present in Section 6.

Then, by substituting the approximations of $\boldsymbol{\psi}^x$ and $\nabla_x \boldsymbol{\psi}^x$ in (27), we calculate the approximate drift and diffusion as follows:

$$\mathbf{F}_d(x) = \int_{\mathbf{R}^n} (\nabla_x \boldsymbol{\psi}^x)_d \left(\mathbf{f}^x\, e^{-V/2}\right) dy = \sum_{|\alpha|\le d} (\nabla_x \boldsymbol{\psi}^x)_\alpha\, \mathbf{f}^x_\alpha, \tag{31a}$$

$$\mathbf{A}_{0d}(x) = \int_{\mathbf{R}^n} \boldsymbol{\psi}^x_d \otimes \left(\mathbf{f}^x\, e^{-V/2}\right) dy = \sum_{|\alpha|\le d} \boldsymbol{\psi}^x_\alpha \otimes \mathbf{f}^x_\alpha, \tag{31b}$$

$$\mathbf{D}_d(x) = \int_{\mathbf{R}^n} \boldsymbol{\alpha}\boldsymbol{\alpha}^T(x,y)\, e^{-V}\, dy + \mathbf{A}_{0d}(x) + \mathbf{A}_{0d}(x)^T, \quad \mathbf{A}_d(x)\mathbf{A}_d(x)^T = \mathbf{D}_d(x). \tag{31c}$$

For $\mathbf{A}_d(x)$ to be well defined, the symmetric matrix $\mathbf{D}_d$ must be positive semi-definite. Using the notations $\psi^x_\xi = \boldsymbol{\psi}^x_d \cdot \boldsymbol{\xi}$ and $f^x_\xi = (\mathbf{f}^x e^{-V/2}) \cdot \boldsymbol{\xi}$, we can show this by noticing that, for any $\boldsymbol{\xi} \in \mathbf{R}^m$, $\boldsymbol{\xi}^T\, (\mathbf{A}_{0d}(x) + \mathbf{A}_{0d}(x)^T)\boldsymbol{\xi} = 2 \int_{\mathbf{R}^n} \psi^x_\xi f^x_\xi\, dy = 2\, a(\psi^x_\xi, \psi^x_\xi)$, by (29).

Although $\boldsymbol{\psi}^x_d$ and $(\nabla_x \boldsymbol{\psi}^x)_d$ are not uniquely defined when $\hat{S}_d$ contains the kernel of $\mathcal{H}$, the coefficients defined in (31a) and (31b) are unique, because by the centering condition f is orthogonal in $L^2\,(\mathbf{R}^n)$ to the kernel of $\mathcal{H}$. The matrix $\mathbf{A}_d(x)$ is not, however, uniquely determined by (31c). Using these coefficients, we obtain the approximate homogenized SDE

$$dX_d \;=\; \mathbf{F}_d(X_d)dt + \mathbf{A}_d(X_d)dW_t. \tag{32}$$

This equation can now be easily solved using a standard numerical method, e.g. Euler-Maruyama.

Our numerical methodology is based on the expansion of the solution to (24) in Hermite functions:

$$\hat{S}_d = \operatorname{span}\{h_\alpha(y;\mu,\Sigma)\}_{|\alpha|\le d}. \tag{33}$$

A good choice of the mean and covariance, μ and Σ, respectively, is important for the efficiency of the algorithm. In our implementation we choose

$$\mu = \int_{\mathbf{R}^n} y\, e^{-V(y)} dy \quad \text{and} \quad \Sigma = \lambda \int_{\mathbf{R}^n} (y-\mu)(y-\mu)^T\, e^{-V(y)}\, dy, \tag{34}$$

where $\lambda > 0$ is a free parameter independent of the first two moments of e^{-V}. This choice for the mean and covariance guarantees that our method is invariant under the rescaling $\tilde{Y}_t^\varepsilon = \sigma(Y_t^\varepsilon - m)$. An example illustrating why this is desirable is when the mass of the probability density e^{-V} is concentrated far away from the origin. Using Hermite functions centered at 0 would provide a very poor approximation in this case, but choosing Hermite functions around the center of mass of e^{-V} leads to a much better approximation. Note that this is not the only choice that guarantees invariance under rescaling, but it is the most natural one.

Remark 3. *When the potential V is quadratic, say $V(y) = \frac{1}{2}(y-m)^T S(y-m)$, the eigenfunctions of the operator $\mathcal{H}$ (defined in (23)) are precisely the Hermite functions $h_\alpha(y; m, S)$. Hence choosing these as a basis, i.e. $e_\alpha = h_\alpha(y; m, S)$, leads to a diagonal matrix A in the linear systems (30), because $a(e_\alpha, e_\beta) = \lambda_\alpha \delta_{\alpha\beta}$, with λ_α defined in (87). This choice corresponds to $\lambda = 1$ in (34). The optimal choice for the parameters μ and Σ for a general density e^{-V} and function $\mathbf{f}$ has been partially studied. In particular, it was shown in [30] that $O(p^2)$ Hermite polynomials are necessary to resolve p wavelengths of a sine function, when keeping the scaling parameter fixed. This result carries over to the case of normalized Hermite functions, where the associated covariance matrix would play the role of the scaling parameter. In [60], it was shown that much better results could be obtained by choosing the scaling parameter as a function of the degree of approximation. In particular, the author shows that by choosing this parameter inversely proportional to the number of Hermite functions, only $O(p)$ functions are needed in order to resolve p wavelengths in one spatial dimension. More recently, the question of the optimal choice of the scaling parameter has also been studied in the framework of spectral methods for the Fokker-Planck equation, see [23].*

Summary of the Method. In short, the method can be summarized as follows.

For a given initial condition $X^\varepsilon(0) = X_0$, $n = 0, 1, 2, \ldots$, a given stochastic integrator $X_d^{n+1} = \Psi(X_d^n, \mathbf{F}_d, \mathbf{A}_d, \Delta t, \xi_n)$, and a chosen time step Δt, set $X_0^n = X_0$ and

1. Compute the solution $\boldsymbol{\psi}_d^{X_d^n}$ and $(\nabla_x \boldsymbol{\psi}^{X_d^n})_d$ of (30);
2. Evaluate $\mathbf{F}_d(X_d^n), \mathbf{A}_d(X_d^n)$ from (31);
3. Compute a time step $X_d^{n+1} = \Psi(X_d^n, \mathbf{F}_d, \mathbf{A}_d, \Delta t, \xi_n)$, and go back to 1.

4. Main Results.

In this section we present the main results on the analysis of our numerical method, the proof of which will be presented in Section 5. We first need to introduce some new notations. We will denote by $\langle \cdot, \cdot \rangle_{e^{-V}}$ the inner product of $L^2\left(\mathbf{R}^n, e^{-V}\right)$, defined by $\langle u, v \rangle_{e^{-V}} = \int_{\mathbf{R}^n} u\, v\, e^{-V}\, dy$, and by $\| \cdot \|_{e^{-V}}$ the associated norm. We will also use the notation $\| \cdot \|_{k,e^{-V}}$ for the norm of $H^k\left(\mathbf{R}^n, e^{-V}\right)$, and $\| \cdot \|_{k,\mathcal{O}}$, where $\mathcal{O}$ is a negative selfadjoint operator, for the norm of $H^k\left(\mathbf{R}^n, \mathcal{O}\right)$, see Appendix A. We will denote by $\pi(\cdot)$ the projection onto mean-zero functions of $L^2\left(\mathbf{R}^n, e^{-V}\right)$, defined by

$$\pi(v) = v - \langle v, 1 \rangle_{e^{-V}}, \quad v \in L^2\left(\mathbf{R}^n, e^{-V}\right). \tag{35}$$

We will work mostly with the Schrödinger formulation (24) of the Poisson equation. In that context, we will employ the $L^2\left(\mathbf{R}^n\right)$ projection operator on $\{\hat{v} \in L^2\left(\mathbf{R}^n\right) : \hat{M}(\hat{v}) = 0\}$, see (26),

which we denote by $\hat{\pi}(\cdot)$:

$$\hat{\pi}(\hat{v}) = \hat{v} - \langle \hat{v}, e^{-V/2} \rangle_0 \, e^{-V/2}, \quad \hat{v} \in L^2(\mathbf{R}^n), \tag{36}$$

where $\langle \cdot, \cdot \rangle_0$ denotes the $L^2(\mathbf{R}^n)$ inner product.

Finally, we will say that a function $g \in L^2(\mathbf{R}^n) \cap C^\infty(\mathbf{R}^n)$ decreases faster than any exponential function in the $L^2(\mathbf{R}^n)$ sense if

$$\int_{\mathbf{R}^n} g(x)^2 e^{\mu|y|} \, dy < \infty \quad \forall \mu \in \mathbf{R}, \tag{37}$$

and denote by $E(\mathbf{R}^n)$ the space of all such functions.

In addition to the hypotheses presented in Section 2, we will use the following assumptions.

Assumption 4.1. *The potential $W(y)$, introduced in (9), is bounded from above by a polynomial of degree $4k$, for some $k \in \mathbf{N}$. Furthermore, for every multi-index α, there exist constants $c_\alpha > 0$ and $\mu_\alpha \in \mathbf{R}$ such that*

$$\left| \partial_y^\alpha V \right| \leq c_\alpha \, e^{\mu_\alpha |y|},$$

where $V(\cdot)$ is the potential that appears in (5b).

Assumption 4.2. *The drift vector $\mathbf{f}(x, y)$ in (5a) is such that $e^{-V(\cdot)/2} \, \partial_y^\alpha \mathbf{f}(x, \cdot) \in (E(\mathbf{R}^n))^m$ and $e^{-V(\cdot)/2} \, \partial_y^\alpha \nabla_x \mathbf{f}(x, \cdot) \in (E(\mathbf{R}^n))^{m \times m}$ for all $\alpha \in \mathbf{N}^n$ and $x \in \mathbf{R}^m$.*

For the proof of our main theorem we will need to have control on higher order derivatives of the solution to the Poisson equation (11). To obtain such bounds we need to strengthen our assumptions on $\mathbf{f}(x, y)$ in (5a). In particular, in addition to (14), we assume the following:

Assumption 4.3. *For all $\alpha \in \mathbf{N}^n$, there exist constants $C_\alpha > 0$ and $\ell_\alpha \in \mathbf{N}$ such that*

$$\left| \partial_y^\alpha \mathbf{f} \right| + \left| \partial_y^\alpha \nabla_x \mathbf{f} \right| \leq C_\alpha \, (1 + |y|^{\ell_\alpha}). \tag{38}$$

In addition, the diffusion coefficient in the right-hand side of (5a) satisfies

$$|\boldsymbol{\alpha}(x, y)| \leq K(1 + |y|^{m_2}), \tag{39}$$

for constants K and m_2 independent of x.

From the Pardoux-Veretennikov bounds (16), a bootstrapping argument, Assumptions 4.1 and 4.3 and the integrability of monomials with respect to Gaussian weights we obtain the bounds

$$\| \boldsymbol{\phi}^x \|_{s, \mathcal{L}_{\mu, \Sigma}} \vee \| \nabla_x \boldsymbol{\phi}^x \|_{s, \mathcal{L}_{\mu, \Sigma}} \vee \| \mathbf{f}^x \|_{e^{-V}} \leq C(s), \tag{40}$$

for $s \in \mathbf{N}$ and a constant $C(s)$ independent of x, and where $a \vee b$ denotes the maximum between a and b.

Remark 4. *In Assumption 4.3 we assumed that the derivatives of the drift vector in (5a) with respect to y are bounded uniformly in x. This is a very strong assumption and it can be replaced by a linear growth bound as in (14). Under such an assumption the proof of Theorem 7 has to be modified using a localization argument that is based on the introduction of appropriate stopping times. Although tedious, this is a standard argument, see e.g. [32], and we will not present it in this paper. Details can be found in [61].*

Theorem 5 (Spectral convergence of the Hermite-Galerkin method). *Under Assumptions 4.1 to 4.3, for any $s \in \mathbf{N}$ there exists $C(s)$, independent of x, such that the approximate solutions $\boldsymbol{\psi}_d^x$ and $(\nabla_x \boldsymbol{\psi}^x)_d$ satisfy the following error estimate:*

$$\|\hat{\boldsymbol{\pi}}(\boldsymbol{\psi}_d^x) - \boldsymbol{\psi}^x\|_0 \vee \|\hat{\pi}((\nabla_x \boldsymbol{\psi}^x)_d) - \nabla_x \boldsymbol{\psi}^x\|_0 \leq C(s)\, d^{-s},$$

where $\|\cdot\|_0$ is the usual $L^2(\mathbf{R}^n)$ norm.

Using this result, we can prove spectral convergence for the calculation of the drift and diffusion coefficients.

Theorem 6 (Convergence of the drift and diffusion coefficients F_d and A_d). *Suppose that Assumptions 4.1, 4.2 and 4.3 hold. Then the error on the approximate drift and diffusion coefficients decreases faster than any negative power of d, uniformly in x, i.e. for all $s \in \mathbf{N}$ there exists $D(s)$ such that*

$$\sup_{x \in \mathbf{R}^m} |\mathbf{F}_d(x) - \mathbf{F}(x)| \vee \left|\mathbf{A}_d(x)\, \mathbf{A}_d(x)^T - \mathbf{A}(x)\, \mathbf{A}(x)^T\right| \leq D(s)\, d^{-s}.$$

Using the spectral convergence of the approximate calculation of the drift and diffusion coefficients, we can now control the distance between the solution of the homogenized SDE and its approximation (32).

As we have already mentioned, homogenization/diffusion approximation theorems are generally of the weak convergence type. Furthermore, the effective diffusion coefficient of the simplified equation is not uniquely defined – see Equation (18) and the fact that $\mathbf{D}(x) = \mathbf{A}(x)\mathbf{A}(x)^T$. Consequently, it is not possible to prove the strong convergence of the solution to the approximate SDE (32) to the solution to the homogenized SDE (20) without constraining the choice of the factorization. To establish the next result, we will therefore assume that $\mathbf{A}(x)$ and $\mathbf{A}_d(x)$ are the unique symmetric positive semi-definite square roots of $\mathbf{D}(x)$ and $\mathbf{D}_d(x)$.

Denoting by $X(t)$ the exact solution of the homogenized equation and by $X_d(t)$ the approximate solution, we use the following norm to measure the error:

$$|||X(t) - X_d(t)||| := \left(\mathbf{E}\left[\sup_{0 \leq t \leq T} |X(t) - X_d(t)|^2\right]\right)^{1/2}. \tag{41}$$

Theorem 7. *Let Assumptions 4.1 to 4.3 hold. Then the error between the approximate and exact solutions of the simplified equation satisfies*

$$|||X(t) - X_d(t)||| \leq \sqrt{4\,(T+4)\,\beta(s)\,T\,d^{-s}}\, \exp\left(2\,(T+4)\,C_L\,T\right), \tag{42}$$

for any $s \in \mathbf{N}$ and $T > 0$, and where $\beta(s)$ is a constant depending only on s.

Now we consider the fully discrete scheme. We need to consider an appropriate discretization of the approximate homogenized equation (32). For simplicity we present the convergence results for the case when we discretize the homogenized SDE using the Euler-Maruyama method:

$$X_d^{n+1} = X_d^n + \Delta t\, \mathbf{F}_d(X_d^n) + \mathbf{A}_d(X_d^n)\, \Delta W_n, \tag{43}$$

but we emphasize that any higher order integrator, e.g. the Milstein scheme, could be used [37, 49]. The following is a classical result on the convergence of X_d^n for which we refer to [37, 49, 32] for a proof.

Theorem 8 (Convergence of the SDE solver). *Assume that X_0 is a random variable such that $\mathbf{E}|X_0|^2 < \infty$ and that Assumptions 4.1 to 4.3 hold. Then*

$$\left(\mathbf{E}\left[\sup_{n\Delta t\in[0,T]} |X_d^n - X_d(t_n)|^2\right]\right)^{\frac{1}{2}} \leq C(T)\sqrt{\Delta t}. \tag{44}$$

for any choice of T, where X_d^n denotes the solution of (43).

Combined, Theorem 7 and Theorem 8 imply the weak convergence of the solution of (43) to the solution of the homogenized equation (20).

5. Proofs of the Main Results.

5.1. Convergence of the Spectral Method for the Poisson Equation.

In this section we present the proof of Theorem 5, establishing the convergence of the spectral method for the solution of the Poisson equation (21). Since the variable x only appears as a parameter in the Poisson equation, we will consider in this section that it takes an arbitrary value and will omit it from the notation. Additionally, to disencumber ourselves of vectorial notations, we will consider an arbitrary direction of $\mathbf{R}^n$, defined through a unit vector $\mathbf{e}$, and denote by f the projection $\mathbf{f}\cdot\mathbf{e}$.

We recall from [51, 52] that there exists a unique smooth mean-zero function of $\phi \in H^1\left(\mathbf{R}^n, e^{-V}\right)$ satisfying the variational formulation

$$a_V(\phi, v) := \langle \nabla\phi, \nabla v\rangle_{e^{-V}} = \langle f, v\rangle_{e^{-V}} \quad \forall v \in H^1\left(\mathbf{R}^n, e^{-V}\right). \tag{45}$$

Let S_d be the finite-dimensional subset of $H^1\left(\mathbf{R}^n, e^{-V}\right)$ defined by $S_d = e^{V/2}\hat{S}_d$, where $\hat{S}_d$ is the approximation space defined in eq. (33), and consider the following problem: find $\phi_d \in S_d$ satisfying:

$$a_V(\phi_d, v_d) = \langle f, v_d\rangle_{e^{-V}} \quad \forall v_d \in S_d. \tag{46}$$

Note that, by definition of f, $\phi = \boldsymbol{\phi}\cdot\mathbf{e}$ and $\phi_d = \boldsymbol{\phi}_d\cdot\mathbf{e}$. The convergence of ϕ_d to ϕ can be obtained using techniques from the theory of finite elements, in particular Céa's lemma and an approximation argument. We will use the notation that was introduced at the beginning Section 4.

Lemma 9 (Céa's lemma). *Let ϕ be the solution of (45) satisfying $\mathcal{M}(\phi) = 0$ and ϕ_d be a solution of (46). Then,*

$$\|\phi - \pi(\phi_d)\|_{1,e^{-V}} \leq C \inf_{v_d\in S_d} \|\phi - v_d\|_{1,e^{-V}}.$$

Proof. The proof is standard. It uses Poincaré inequality for the measure $e^{-V}dx$, recalled in Appendix A, Proposition 17. ∎

Since we will be working mostly with the Schrödinger formulation of Poisson equation, we need an analogue of Lemma 9 for the transformed PDE. We recall from Appendix A that the space $H^1(\mathbf{R}^n, \mathcal{H})$ is equipped with the norm

$$\|\psi\|_{1,\mathcal{H}}^2 = \|\psi\|_0^2 + \int_{\mathbf{R}^n} |\nabla\psi|^2 \, dy + \int_{\mathbf{R}^n} W\psi^2 \, dy.$$

Lemma 10. *Let $\boldsymbol{\psi}$ be the unique solution of (25) satisfying $\hat{\mathcal{M}}(\psi) = 0$ and $\boldsymbol{\psi}_d$ be a solution of (29). Then the projections $\psi = \boldsymbol{\psi} \cdot \mathbf{e}$ and $\psi_d = \boldsymbol{\psi}_d \cdot \mathbf{e}$ satisfy*

$$\|\psi - \hat{\pi}(\psi_d)\|_{1,\mathcal{H}} \leq C \inf_{v_d \in \hat{S}_d} \|\psi - v_d\|_{1,\mathcal{H}}. \tag{47}$$

Proof. The result follows directly by using the fact that $e^{-V/2}$ is also a unitary transformation from $H^1\left(\mathbf{R}^n, e^{-V}\right)$ to $H^1(\mathbf{R}^n, \mathcal{H})$. ∎

Next, we focus on establishing a result that will allow us to control the right-hand side of (47). In [26, Lemma 2.3] the authors show that any smooth square integrable function such that $(-\Delta + W)v = g$ lies in the space $E(\mathbf{R}^n)$ introduced in (37), provided that $g \in E(\mathbf{R}^n)$ and that Assumption (9) holds. Differentiating the equation with respect to y_i, we obtain:

$$(-\Delta + W)\, \partial_{y_i} v = \partial_{y_i} g - \partial_{y_i} W \, v,$$

so it is clear by Assumption 4.1 that $\partial_\alpha \psi \in E(\mathbf{R}^n)$ for all values of $\alpha \in \mathbf{N}^n$. This implies that ψ belongs to the Schwartz space $S(\mathbf{R}^n)$. We now generalize sligthly [26, Lemma 3.1]. This result will enable to control the norm $\|\cdot\|_{1,\mathcal{H}}$ on the right-hand side of (47) by a norm $\|\cdot\|_{k,\mathcal{H}_{\mu,\Sigma}}$, where $\mathcal{H}_{\mu,\Sigma}$ is an operator defined in Appendix A. From this appendix, we recall that the operator $\mathcal{H}_{\mu,\Sigma}$, with $\mu \in \mathbf{R}^n$ and Σ a symmetric positive definite matrix, is defined by $\mathcal{H}_{\mu,\Sigma} = -\Delta + W_{\mu,\Sigma}(y)$, where $W_{\mu,\Sigma}$ denotes the quadratic function $(y-\mu)^T\Sigma^{-2}(y-\mu)/4 - \operatorname{tr}\Sigma^{-1}/2$.

Lemma 11. *For every $k \in \mathbf{N}$ and $v \in S(\mathbf{R}^n)$,*

$$\int_{\mathbf{R}^n} |y|^{4k} \, v^2(y) \, dy \leq C(k, \mu, \Sigma) \|v\|_{2k,\mathcal{H}_{\mu,\Sigma}}^2,$$

where $C(k, \mu, \Sigma)$ is a constant independent of v.

Proof. We set $Q_{\mu,\Sigma} = (y-\mu)^T\Sigma^{-2}(y-\mu)/4$. Following the methodology used to prove lemma 3.1 in [26], we establish that:

$$\|Q_{\mu,\Sigma}(y)^{k+1} v\|_0^2 \leq \|Q_{\mu,\Sigma}(y)^k \left(\mathcal{H}_{\mu,\Sigma} + \operatorname{tr}\Sigma^{-1}/2\right) v\|_0^2 + C_1(k,\Sigma)\|Q_{\mu,\Sigma}(y)^k v\|_0^2,$$

for all $k \in \mathbf{N}$, and where $C_1(k,\Sigma) = (4k+2)(k\,\rho(\Sigma^{-2}) + \operatorname{tr}\Sigma^{-2}/4)$. Reasoning by recursion and applying the triangle inequality, this immediately implies

$$\begin{aligned}\|Q_{\mu,\Sigma}(y)^k v\|_0^2 &\leq \sum_{i=0}^{k} c_i(k,\Sigma) \, \| \left(\mathcal{H}_{\mu,\Sigma} + \operatorname{tr}\Sigma^{-1}/2\right)^i v\|_0^2 \\ &\leq C_2(k,\Sigma)\|v\|_{2k,\mathcal{H}_{\mu,\Sigma}}^2,\end{aligned}$$

To conclude, note that

$$|y|^{4k} \leq C_3 + C_4\, Q_{\mu,\Sigma}(y)^{2k},$$

for suitably chosen C_3 and C_4 depending on Σ and μ. ∎

A finer version of the previous inequality could be obtained by following the argumentation in [26, Theorem 3.2], but this will not be necessary for our purposes. Lemma 11 can be used to show the following result.

Lemma 12. *If $W(y)$ is bounded above by a polynomial of degree $4k$, there exists a constant C depending on k, μ, Σ, and W such that any $v \in S(\mathbf{R}^n)$ satisfies*

$$\|v\|_{1,\mathcal{H}} \leq C\, \|v\|_{2k,\mathcal{H}_{\mu,\Sigma}}.$$

Proof. This follows from the considerations of Appendix A. First we note that

$$\|v\|^2_{1,\mathcal{H}} = \|v\|^2_{1,\mathcal{H}_{\mu,\Sigma}} + \int_{\mathbf{R}^n} (W - W_{\mu,\Sigma}) v^2\, dy.$$

To bound the second term, we use Assumption 4.1 on W, together with Lemma 11:

$$\int_{\mathbf{R}^n} (W - W_{\mu,\Sigma}) v^2\, dy \leq \int_{\mathbf{R}^n} (C_1 + C_2\ |y|^{4k}) v^2\, dy \leq C_3 \|v\|^2_{2k,\mathcal{H}_{\mu,\Sigma}},$$

with C_1, C_2, C_3 depending on k, μ, Σ. ∎

Upon combining the results presented so far in this section, we can complete the proof of Theorem 5.

Proof of Theorem 5. In this proof, the constant C is independent of d and changes from line to line. By Lemmas 10 and 11, and the fact that the exact solution ψ and its derivatives are smooth and decrease faster than exponentials, we have:

$$\|\psi - \hat{\pi}(\psi_d)\|_{1,\mathcal{H}} \leq C \inf_{v_d \in \hat{S}_d} \|\psi - v_d\|_{1,\mathcal{H}} \leq C \inf_{v_d \in \hat{S}_d} \|\psi - v_d\|_{2k,\mathcal{H}_{\mu,\Sigma}}.$$

Using Corollary 23 on approximation by Hermite functions, we have for any $s > 2k$

$$\begin{aligned} \|\psi - \hat{\pi}(\psi_d)\|_{1,\mathcal{H}} &\leq C(d+1)^{-\frac{s-2k}{2}} \|\psi\|_{s,\mathcal{H}_{\mu,\Sigma}}, \\ &\leq C(d+1)^{-\frac{s-2k}{2}}, \end{aligned}$$

where we used the first estimate of (40) and the fact that $\|\psi\|_{s,\mathcal{H}_{\mu,\Sigma}} = \|\phi\|_{s,\mathcal{L}_{\mu,\Sigma}}$. The same reasoning can be applied to $\nabla_x \psi$. Since s was arbitrary, this proves the statement. ∎

5.2. Convergence of the Drift and Diffusion Coefficients. In this section we prove the convergence of the drift and diffusion coefficients obtained from the approximate solution of the Poisson equation.

Proof of Theorem 6. From the expressions of $\mathbf{F}$ and $\mathbf{F}_d$ we have:

$$\begin{aligned}\mathbf{F}(x) - \mathbf{F}_d(x) &= \int_{\mathbf{R}^n} (\nabla_x \boldsymbol{\psi}^x - (\nabla_x \boldsymbol{\psi}^x)_d) \cdot (\mathbf{f}^x \, e^{-V/2}) \, dy \\ &= \int_{\mathbf{R}^n} (\nabla_x \boldsymbol{\psi}^x - \hat{\pi}((\nabla_x \boldsymbol{\psi}^x)_d)) \cdot (\mathbf{f}^x \, e^{-V/2}) \, dy,\end{aligned}$$

where we used the fact that $\mathbf{f}$ is centered, thus orthogonal to $(\nabla_x \boldsymbol{\psi}^x)_d - \hat{\pi}((\nabla_x \boldsymbol{\psi}^x)_d)$. Using Theorem 5 and Cauchy-Schwarz inequality we deduce that there exists for any value of $s \in \mathbf{N}$ a constant $C(s)$ such that

$$\begin{aligned}|\mathbf{F}_d(x) - \mathbf{F}(x)| &\leq \|\nabla_x \boldsymbol{\psi}^x - \hat{\pi}((\nabla_x \boldsymbol{\psi}^x)_d)\|_0 \, \|\mathbf{f}^x \, e^{-V/2}\|_0 \\ &\leq C(s) \, d^{-s} \|\mathbf{f}^x\|_{e^{-V}}.\end{aligned}$$

The error on the diffusion term can be bounded similarly:

$$\begin{aligned}|\mathbf{A}_{0d}(x) - \mathbf{A}_0(x)| &= \int_{\mathbf{R}^n} (\boldsymbol{\psi}_d^x - \boldsymbol{\psi}^x) \otimes (\mathbf{f}^x \, e^{-V/2}) \, dy \\ &\leq C(s) \, d^{-s} \, \|\mathbf{f}^x\|_{e^{-V}}.\end{aligned}$$

The proof can then be concluded using the last bound from (40). ∎

5.3. Convergence of the Solution to the SDE. The proof of Theorem 7 relies on two results. First, it relies on the existence of a constant C_L, proved in Appendix C, such that the effective drift and diffusion coefficients, obtained using (17) and (18) and by calculating $\mathbf{A}(x)$ as the unique symmetric positive semi-definite square root of $\mathbf{D}(x)$, satisfy

$$|\mathbf{F}(a) - \mathbf{F}(b)|^2 \vee |\mathbf{A}(a) - \mathbf{A}(b)|^2 \leq C_L |a - b|^2, \tag{48}$$

for all $a, b \in \mathbf{R}^m$. Next, by Theorem 6 and Lemma 24, there exists for every $s \in \mathbf{N}$ a constant $\beta(s)$ independent of d and x such that

$$|\mathbf{F}_d(x) - \mathbf{F}(x)|^2 \vee |\mathbf{A}_d(x) - \mathbf{A}(x)|^2 \leq \beta(s) \, d^{-s}, \tag{49}$$

for any $x \in \mathbf{R}^m$. Upon combining (48) and (49), Theorem 7 can be proved.

Proof of Theorem 7. The error $e_d(t) = X(t) - X_d(t)$ satisfies

$$e_d(t) = \int_0^t \mathbf{F}(X(\tau)) - \mathbf{F}_d(X_d(\tau)) \, d\tau + \int_0^t \mathbf{A}(X(\tau)) - \mathbf{A}_d(X_d(\tau)) \, \mathrm{d}W_\tau.$$

Using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ and Cauchy-Schwarz, we have

$$\begin{aligned}\mathbf{E}\left[\sup_{0 \leq t \leq T} |e_d(t)|^2\right] &\leq 2\,T\,\mathbf{E}\left[\int_0^T |\mathbf{F}(X(\tau)) - \mathbf{F}_d(X_d(\tau))|^2 \, \mathrm{d}\tau\right] \\ &\quad + 2\,\mathbf{E}\left[\sup_{0 \leq t \leq T} \left|\int_0^t \mathbf{A}(X(\tau)) - \mathbf{A}_d(X_d(\tau)) \, \mathrm{d}W_\tau\right|^2\right].\end{aligned} \tag{50}$$

The first term in the right-hand side can be bounded by using the triangle inequality with the decomposition $\mathbf{F}(X(\tau)) - \mathbf{F}_d(X_d(\tau)) = (\mathbf{F}(X(\tau)) - \mathbf{F}(X_d(\tau))) + (\mathbf{F}(X_d(\tau)) - \mathbf{F}_d(X_d(\tau)))$, the Lipschitz continuity of $\mathbf{F}(\cdot)$ and the convergence of $\mathbf{F}_d$ to $\mathbf{F}$:

$$\begin{aligned} &\mathbf{E}\left[\int_0^T |\mathbf{F}(X(\tau)) - \mathbf{F}_d(X_d(\tau))|^2 \,\mathrm{d}\tau\right] \\ &\qquad \le \mathbf{E}\left[2\,\beta(s)\,T\,d^{-s} + 2\,C_L \int_0^T |X(\tau) - X_d(\tau)|^2 \,d\tau\right] \\ &\qquad \le 2\,\beta(s)\,T\,d^{-s} + 2\,C_L \int_0^T \mathbf{E}\left[\sup_{0\le t\le \tau} |e_d(t)|^2\right] d\tau \end{aligned} \tag{51}$$

The second term can be bounded in a similar manner by using Burkholder–Davis–Gundy inequality, see for example [35, Theorem 3.28], and Itô isometry :

$$\begin{aligned} &\mathbf{E}\left[\sup_{0\le t\le T}\left|\int_0^t \mathbf{A}(X(\tau)) - \mathbf{A}_d(X_d(\tau))\,\mathrm{d}W_\tau\right|^2\right] \\ &\qquad \le 4\,\mathbf{E}\left|\int_0^T \mathbf{A}\,(X(\tau)) - \mathbf{A}_d(X_d(\tau))\,\mathrm{d}W_\tau\right|^2 \\ &\qquad \le 8\,\beta(s)\,T\,d^{-s} + 8\,C_L \int_0^T \mathbf{E}\left[\sup_{0\le t\le \tau} |e_d(t)|^2\right] d\tau. \end{aligned} \tag{52}$$

Using (51) and (52) in (50), we obtain:

$$\mathbf{E}\left[\sup_{0\le t\le T} |e_d(t)|^2\right] \le 4\,(T+4)\left(\beta(s)\,T\,d^{-s} + C_L \int_0^T \mathbf{E}\left[\sup_{0\le t\le \tau} |e_d(t)|^2\right] \mathrm{d}\tau\right).$$

By Gronwall's inequality, this implies:

$$\mathbf{E}\left[\sup_{0\le t\le T} |e_d(t)|^2\right] \le 4\,(T+4)\,\beta(s)\,T\,d^{-s}\,\exp\left(4\,(T+4)\,C_L\,T\right), \tag{53}$$

which finishes the proof. ∎

Remark 13. *Note that, as mentioned in Section 4, the convergence of the solution can still be proved if we assume that the Lipschitz continuity and convergence of the coefficients holds only locally, provided there exists $p > 2$ and a constant K independent of d such that the solutions of the equations*

$$\mathrm{d}X = \mathbf{F}(X)\,\mathrm{d}t + \mathbf{A}(X)\,\mathrm{d}W_t, \qquad X(0) = X_0,$$

and

$$\mathrm{d}X_d = \mathbf{F}_d(X_d)\,\mathrm{d}t + \mathbf{A}_d(X_d)\,\mathrm{d}W_t, \qquad X_d(0) = X_0,$$

satisfy the moment bounds

$$\mathbf{E}\left[\sup_{0\le t\le T} |X(t)|^p\right] \vee \mathbf{E}\left[\sup_{0\le t\le T} |X_d(t)|^p\right] \le K.$$

With these alternative assumptions, we can show that

$$\mathbf{E}\left[\sup_{0\le t\le T}|X(t)-X_d(t)|^2\right] \le 4\,(T+4)\,D_R(s)\,T\,d^{-s}\exp\left(4\,(T+4)\,C_R\,T\right) + 2\,K\left(\frac{2^p\,\delta}{p}+\frac{p-2}{R^p\,p\,\delta^{\frac{2}{p-2}}}\right)$$

for any $\delta > 0$ and $R > |X_0|$, and where C_R and D_R are the local constants for the assumptions. The proof of this estimate is very similar to the one of the strong convergence of Euler-Maruyama scheme in [32, Theorem 2.2], and will thus not be presented here. From this estimate, we deduce that the solution of the approximate homogenized equation converges to the exact solution when $d \to \infty$.

6. Implementation of the Algorithm and Numerical Experiments. In this section, we discuss the implementation of the algorithm and present some numerical experiments to validate the method and illustrate our theoretical findings.

6.1. Implementation details. Below we discuss the quadrature rules used and the approach taken for the calculation of the matrix and right-hand side of the linear system of equations (30).

The algorithm requires the calculation of Gaussian integrals of the type:

$$I = \int_{\mathbf{R}^n} f(y)\,G_{(\mu,\Sigma)}(y)\,dy. \tag{54}$$

Several approaches, either Monte Carlo-based or deterministic, can be used for the calculation of such integrals. Probabilistic methods offer an advantage when the dimension n of the state space of the fast process is large, but since the HMM is more efficient than our approach in that case, in practice we don't use them. Instead, we use a multi-dimensional quadrature rule obtained by tensorization of one-dimensional Gauss-Hermite quadrature rules.

For the calculation of the stiffness matrix, we can take advantage of the diagonality of A when the potential is equal to $V_{\mu,\Sigma} := \frac{1}{2}(y-\mu)\Sigma^{-1}(y-\mu) + \log(\sqrt{(2\pi)^n \det\Sigma})$.[6] Using the notation $\mathcal{H}_{\mu,\Sigma}$ to denote the same operator as in Lemma 11, and the shorthand notations H_α and h_α, for $\alpha \in \mathbf{N}^n$, in place of $H_\alpha(y;\mu,\Sigma)$ and $h_\alpha(y;\mu,\Sigma)$, respectively, we have:

$$A_{\alpha\beta} = -\int_{\mathbf{R}^n}(\mathcal{H}-\mathcal{H}_{\mu,\Sigma})\,h_\alpha\,h_\beta\,dy - \int_{\mathbf{R}^n}\mathcal{H}_{\mu,\Sigma}\,h_\alpha\,h_\beta\,dy =: A^\delta_{\alpha\beta} + D_{\alpha\beta}, \tag{55}$$

where D is a diagonal matrix whose entries can be computed explicitly and

$$A^\delta_{\alpha\beta} = \int_{\mathbf{R}^n}(W-W_{\mu,\Sigma})\,f_\alpha f_\beta\,dy = \int_{\mathbf{R}^n}(W-W_{\mu,\Sigma})\,G_{(\mu,\Sigma)}\,H_\alpha H_\beta\,dy, \tag{56}$$

where $W_{\mu,\Sigma}$ is the potential obtained from $V_{\mu,\Sigma}$ according to (9). To simplify the calculation of these coefficients, we expand the Hermite polynomials in terms of monomials:

$$H_\alpha(y;\mu,\Sigma) = \sum_{|\beta|\le d} c_{\alpha\beta}\,y^\beta. \tag{57}$$

[6] The constant $\log(\sqrt{(2\pi)^n \det\Sigma})$ in $V(\mu,\Sigma)$ is chosen so that $\int_{\mathbf{R}^n} e^{-V}\,dy = 1$.

With this notation, we write:

$$A^{\delta}_{\alpha\beta} = \sum_{|\rho|\le d}\sum_{|\sigma|\le d} c_{\alpha\rho}\, c_{\beta\sigma} \int_{\mathbf{R}^n} (W - W_{\mu,\Sigma})\, G_{(\mu,\Sigma)}\, y^{\rho+\sigma}\, dy =: \sum_{|\rho|\le d}\sum_{|\sigma|\le d} c_{\alpha\rho}\, c_{\beta\sigma} I_{\rho+\sigma}, \tag{58}$$

The integrals I_α are computed by numerical quadrature. Denoting by w_i and q_i the weights and nodes of the Gauss-Hermite quadrature, respectively, I_α is approximated as

$$I_\alpha \approx \sum_{i=1}^{N_q} w_i \left(W(q_i) - W_{\mu,\Sigma}(q_i)\right)\, G_{(\mu,\Sigma)}(q_i)\, q_i^{\alpha}, \qquad |\alpha| \le 2d, \tag{59}$$

where N_q denotes the number of points in the quadrature. Only the last factor of the previous expression depends on α, so the numerical calculation of these integrals can be performed by evaluating for each grid point the value of $w_i \left(W(q_i) - W_{\mu,\Sigma}(q_i)\right)\, G_{(\mu,\Sigma)}(q_i)$ and the values of q_i^α for $|\alpha| \le 2d$. For the numerical experiments presented in this paper, we used a quadrature rule with 100 points in each dimension, so $N_q = 100^n$. This quadrature rule was used also to approximate the exact effective coefficients from (27).

A similar method can be applied for the calculation of the right-hand side, whose elements are expressed as:

$$b_\alpha = \int_{\mathbf{R}^n} e^{-V/2} f\, e_\alpha\, dy. \tag{60}$$

By expanding the Hermite functions in terms of Hermite polynomials multiplying $G_{(\mu,\Sigma)}^{1/2}$, the previous equation can be rewritten as

$$b_\alpha = \sum_{|\beta|\le d} c_{\alpha\beta} \int_{\mathbf{R}^n} \left(\frac{e^{-V}}{G_{(\mu,\Sigma)}}\right)^{\frac{1}{2}} f(x,y)\, y^\beta\, G_{(\mu,\Sigma)}\, dy, \tag{61}$$

which is a Gaussian integral that can also be calculated using a multi-dimensional Gauss-Hermite quadrature.

6.2. Numerical experiments. Now we present the results of some numerical experiments.

The Euler-Maruyama scheme is used to approximate both $X(t)$ and $X_d(t)$ with a time step of 0.01 for $T = 1$, and $N_r = 100$ replicas of the driving Brownian motion are used for the numerical computation of expectations. The ith replica of the discretized approximations of $X(t)$ and $X_d(t)$ are noted $X^{n,i}$ and $X_d^{n,i}$ respectively. In most of the numerical experiments below, the error is measured by:

$$E(d) = \left(\frac{1}{N_r}\sum_{i=1}^{N_r} \max_{0\le n\,\Delta t\le 1} |X^{n,i} - X_d^{n,i}|^2\right)^{\frac{1}{2}}, \tag{62}$$

which is an approximation of the norm $|||\cdot|||$ used in Theorem 7.

In the numerical experiments presented in this paper, we have chosen the scaling parameter λ in (34) by trial-and-error. A natural extension of the work presented in this paper is to develop a systematic methodology for identifying the optimal scaling parameter, see also the discussion in Remark 3.

6.2.1. Test of the method for single well potentials. For the two problems in this section, the scaling parameter is chosen as $\lambda = 0.5$ for all degrees of approximation. We start by considering the following problem.

$$
\begin{cases}
dx_{0t} = -\dfrac{1}{\varepsilon}\,\mathcal{L}_y\left[\cos\left(x_{0t}+y_{0t}+y_{1t}\right)\right]\,dt,\\
dx_{1t} = -\dfrac{1}{\varepsilon}\,\mathcal{L}_y\left[\sin\left(x_{1t}\right)\sin\left(y_{0t}+y_{1t}\right)\right]\,dt,\\
dy_{0t} = -\dfrac{1}{\varepsilon^2}\,\partial_{y_0}V(y)\,dt + \dfrac{1}{\varepsilon}\left[\cos\left(x_{0t}\right)\cos\left(y_{0t}\right)\cos\left(y_{1t}\right)\right]\,dt + \dfrac{\sqrt{2}}{\varepsilon}dW_{0t},\\
dy_{1t} = -\dfrac{1}{\varepsilon^2}\,\partial_{y_1}V(y)\,dt + \dfrac{1}{\varepsilon}\left[\cos\left(x_{0t}\right)\cos\left(y_{0t}+y_{1t}\right)\right]\,dt + \dfrac{\sqrt{2}}{\varepsilon}dW_{1t},
\end{cases}
\tag{63}
$$

with

$$
V(y) = y_0^2 + y_1^2 + 0.5\left(y_0^2+y_1^2\right)^2, \tag{64}
$$

and where $\mathcal{L}_y = -\nabla V\cdot\nabla + \Delta$. We have written the right-hand side of the equations for the slow processes x_{0t} and x_{1t} in this form to ensure that the centering condition is satisfied. The convergence of the approximate solution of the effective equation for this problem is illustrated in Figure 1. Here the potential is localized, so Hermite functions are well suited for the approximation of the solution, which is reflected in the very good convergence observed.

In the next example, the state space of the fast process has dimension 3:

$$
\begin{cases}
dx_{0t} = -\dfrac{1}{\varepsilon}\,\mathcal{L}_y\left[\cos\left(x_{0t}+y_{0t}+y_{1t}\right)\right]\,dt,\\
dx_{1t} = -\dfrac{1}{\varepsilon}\,\mathcal{L}_y\left[\sin\left(x_{1t}\right)\sin\left(y_{0t}+y_{1t}+2y_{2t}\right)\right]\,dt,\\
dy_{0t} = -\dfrac{1}{\varepsilon^2}\,\partial_{y_0}V(y)\,dt + \dfrac{1}{\varepsilon}\left[\cos\left(x_{0t}\right)\cos\left(y_{1t}\right)\cos\left(y_{0t}+y_{2t}\right)\right]\,dt + \dfrac{\sqrt{2}}{\varepsilon}\,dW_{0t},\\
dy_{1t} = -\dfrac{1}{\varepsilon^2}\,\partial_{y_1}V(y)\,dt + \dfrac{1}{\varepsilon}\left[\cos\left(x_{0t}\right)\cos\left(y_{0t}+y_{1t}\right)\right]\,dt + \dfrac{\sqrt{2}}{\varepsilon}\,dW_{1t},\\
dy_{2t} = -\dfrac{1}{\varepsilon^2}\,\partial_{y_2}V(y)\,dt + \dfrac{\sqrt{2}}{\varepsilon}\,dW_{2t},
\end{cases}
\tag{65}
$$

with

$$
V(y) = y_0^4 + 2y_1^4 + 3y_2^4. \tag{66}
$$

Because computing the effective coefficients is much more expensive computationally than in the previous case, we measure the error for a given value of the slow variables, by

$$
e(d,x) = \frac{|\mathbf{F}(x)-\mathbf{F}_d(x)|}{|\mathbf{F}(x)|} + \frac{|\mathbf{A}(x)-\mathbf{A}_d(x)|}{|\mathbf{A}(x)|}. \tag{67}
$$

The value we chose for the comparison is $x = (0.2, 0.2)$, for which the denominators in the previous equation are non-zero. The relative error on the homogenized coefficients is illustrated in Figure 2. In this case, the method also performs very well, although it is slightly less accurate than in the previous example.

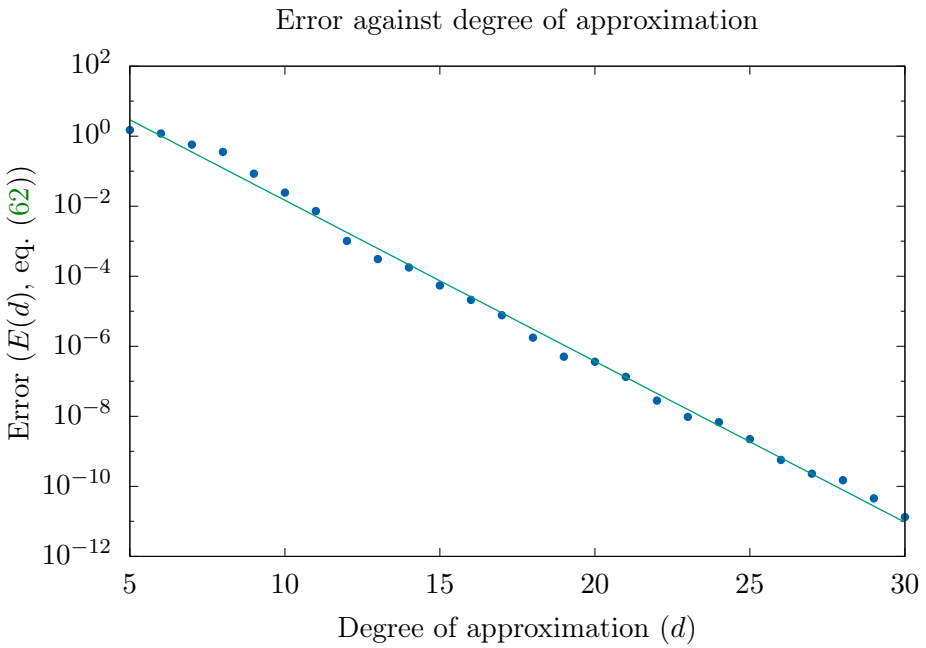

Figure 1. *Error $E(d)$, see (62), for the fast-slow SDE (63). The blue dots were obtained by numerical experiments, and the straight line is a function of the type $C\,e^{-\alpha\,d}$ obtained by linear regression in the $d - log_{10}(E(d))$ plane. A super-algebraic convergence is observed.*

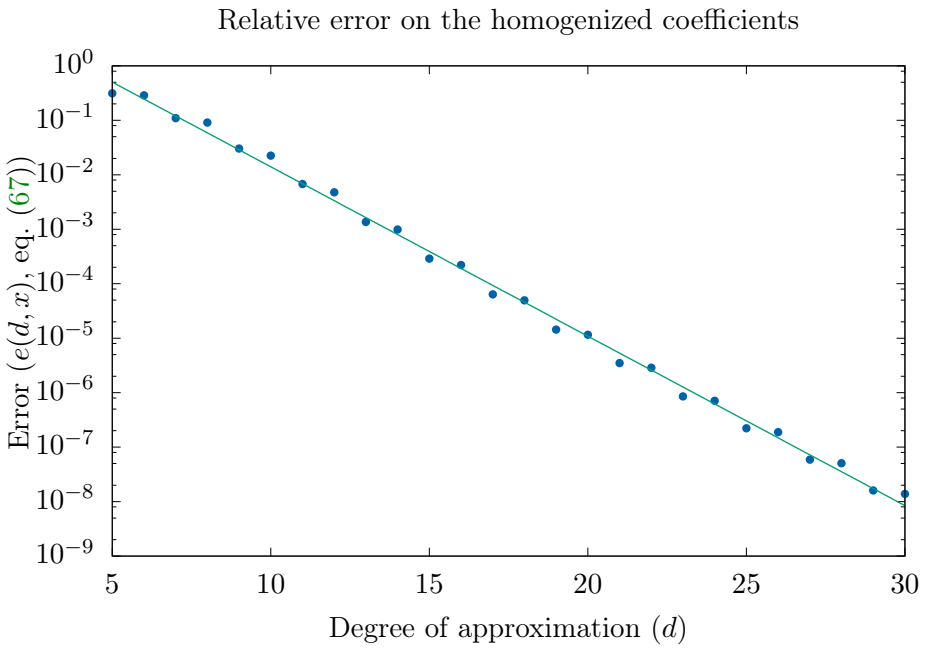

Figure 2. *Relative error of the homogenized coefficients, $e(d,x)$, see (67), for the fast/slow SDE (65) at $x = (0.2, 0.2)$. In this case, the convergence is also super-algebraic.*

6.2.2. Test of the method for potentials with multiple wells. Now we consider multiple-well potentials that lead to multi-modal distributions. Potentials with multiple wells pose challenges to our method because of the localized nature of Hermite functions. In the examples below, we consider potentials with wells that are close to each other, so the solutions can still be approximated well by Hermite functions. The first potential that we analyze is the standard bistable potential,

$$V(y) = y^4/4 - y^2/2. \tag{68}$$

We consider the fast/slow SDE system:

$$\begin{cases} dx_t = -\dfrac{1}{\varepsilon}\mathcal{L}_y\left(x_t \sin(y_t)\right) dt, \\ dy_t = -\dfrac{1}{\varepsilon^2}\partial_y V(y_t)\, dt + \dfrac{\sqrt{2}}{\varepsilon}\, dW_t. \end{cases} \tag{69}$$

We choose the parameter λ in (34) to be $\lambda = 0.5$. The convergence of the method is illustrated in Figure 3. Although the method is less accurate than in the previous cases, a super-algebraic convergence can still be observed, and a good accuracy can be reached by choosing a high enough value for the degree of approximation. Note that the computational cost in this case is low – the numerical solution can be calculated in a matter of seconds on a personal computer.

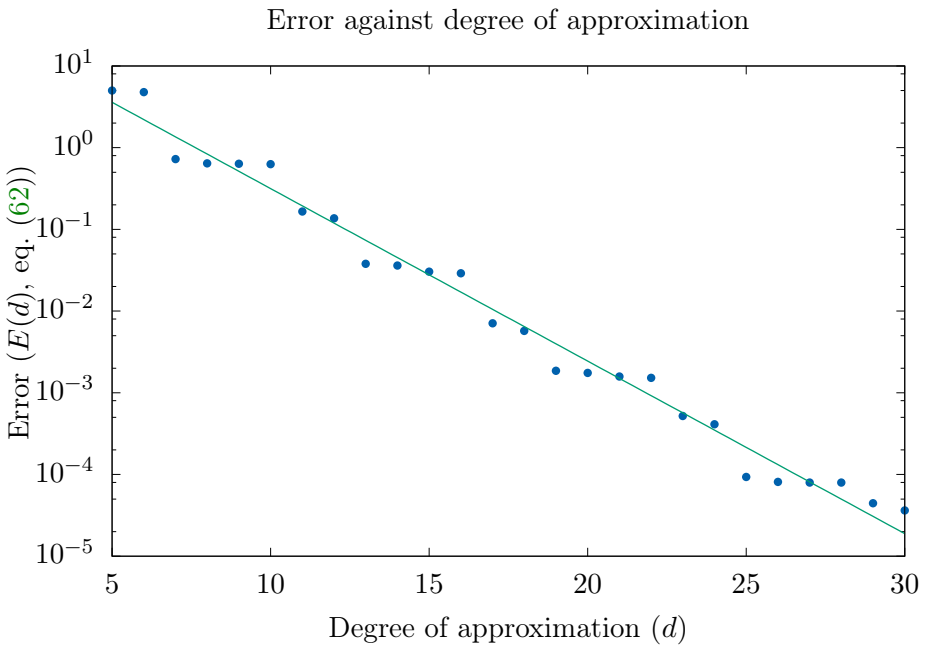

Figure 3. *Error $E(d)$, see (62), for the fast/slow SDE (69).*

Next we consider the tilted bistable potential

$$V(y) = y^4/4 - y^2/2 + 10y, \tag{70}$$

which corresponds to the case $\gamma = 1$, $\delta = 10$ in the examples considered in [9], and the fast/slow SDE

$$\begin{cases} dx_t = -\dfrac{1}{\varepsilon}\mathcal{L}_y\left(x_t\,\sin(y_t) + y_t^2\right)\,dt, \\ dy_t = -\dfrac{1}{\varepsilon^2}\partial_y V(x_t, y_t)\,dt + \dfrac{\sqrt{2}}{\varepsilon}\,dW_t. \end{cases} \tag{71}$$

The convergence of the solution in this case is presented in Figure 5, for the scaling parameter $\lambda = 1$. Due to the presence of a strong linear term, the potential is localized, see Figure 4. As a result, the convergence of the spectral method is good, though it does not appear to be super-algebraic.

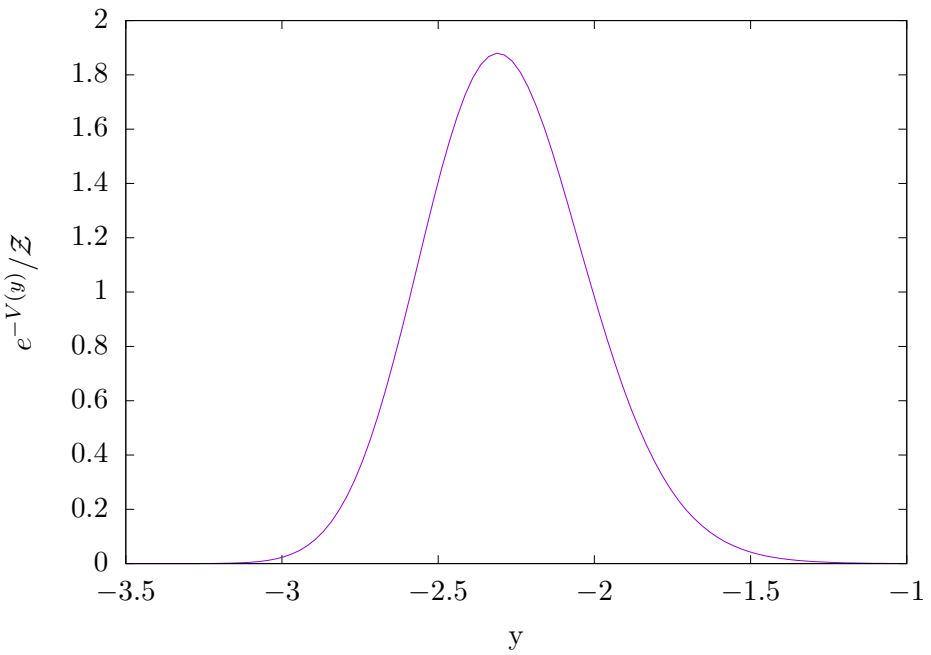

Figure 4. *Probability density $e^{-V(\cdot)}/\mathcal{Z}$ associated to the potential (70).*

Finally, we consider a three-well potential in $\mathbf{R}^2$,

$$V(y) = \left((y_0 - 1)^2 + y_1^2\right)\left(\left(y_0 + \frac{1}{2}\right)^2 + \left(y_1 - \frac{\sqrt{3}}{2}\right)^2\right)\left(\left(y_0 + \frac{1}{2}\right)^2 + \left(y_1 + \frac{\sqrt{3}}{2}\right)^2\right), \tag{72}$$

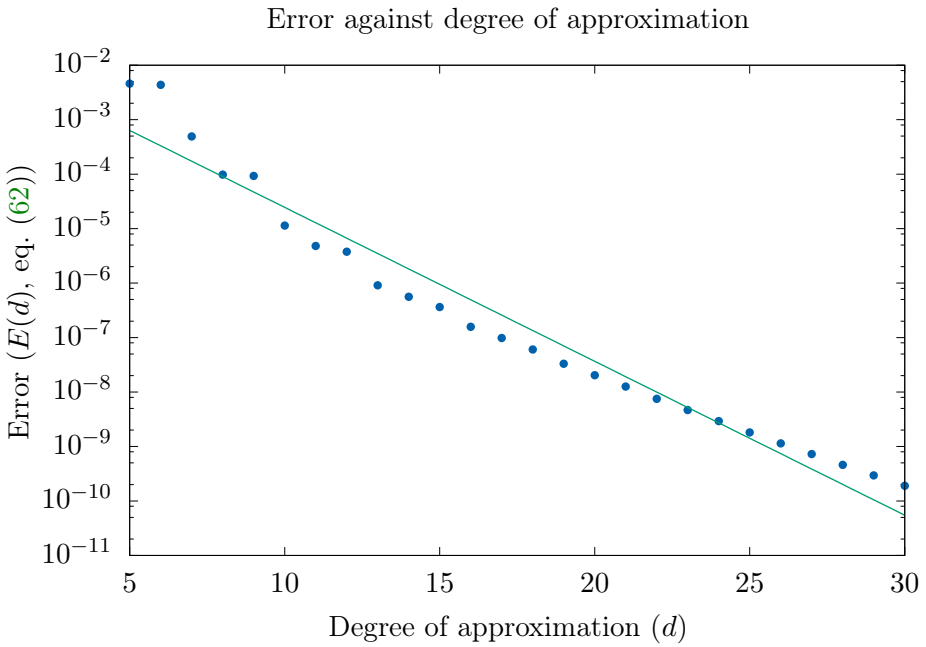

Figure 5. *Error $E(d)$, see (62), for the fast/slow system (71).*

and the following fast/slow SDE:

$$
\begin{cases}
dx_{0t} = -\dfrac{1}{\varepsilon}\,\mathcal{L}_y\left[\cos\left(x_{0t}+y_{0t}+y_{1t}\right)\right]\,dt,\\
dx_{1t} = -\dfrac{1}{\varepsilon}\,\mathcal{L}_y\left[\sin\left(x_{1t}\right)\sin\left(y_{0t}+y_{1t}\right)\right]\,dt,\\
dy_{0t} = -\dfrac{1}{\varepsilon^2}\,\partial_{y_0}V(y)\,dt + \dfrac{1}{\varepsilon}\left[\cos\left(x_{0t}\right)\cos\left(y_{0t}\right)\cos\left(y_{1t}\right)\right]\,dt + \dfrac{\sqrt{2}}{\varepsilon}\,dW_{0t},\\
dy_{1t} = -\dfrac{1}{\varepsilon^2}\,\partial_{y_1}V(y)\,dt + \dfrac{1}{\varepsilon}\left[\cos\left(x_{0t}\right)\cos\left(y_{0t}+y_{1t}\right)\right]\,dt + \dfrac{\sqrt{2}}{\varepsilon}\,dW_{1t}.
\end{cases}
\tag{73}
$$

For this fast/slow SDE, we choose $\lambda = 0.35$. A contour plot of the potential is shown in Figure 6, and the convergence graph is presented in Figure 7. In this case the error is very large for degrees of approximation lower than 10, beyond which the convergence is clear and super-algebraic. The accuracy reached with a degree of approximation equal to 30 is of the order of 1×10^{-4}, which is good in comparison with the accuracy that can be achieved using Monte Carlo-based methods.

6.2.3. Discretization of a multiscale stochastic PDE. As mentioned in the introduction, our numerical method is particularly well-suited for the solution of singularly perturbed stochastic PDEs (SPDEs), and constitutes a very good complement to the method proposed in [2]. Let us recall how the method introduced in [2] works for a singularly perturbed SPDE

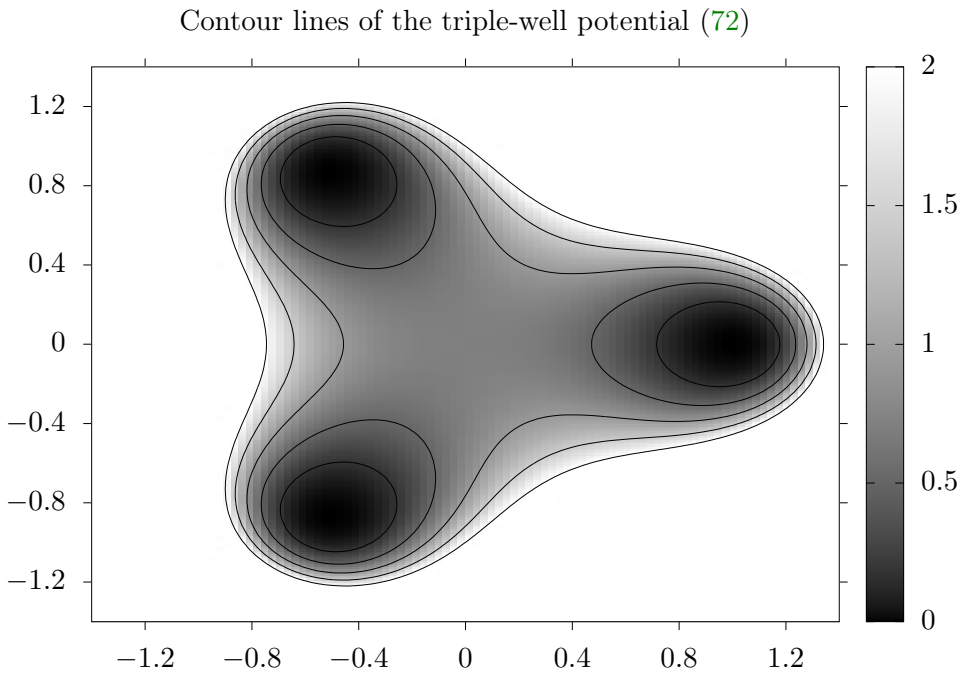

Figure 6. *Potential* (72)*, used in equation* (73)*.*

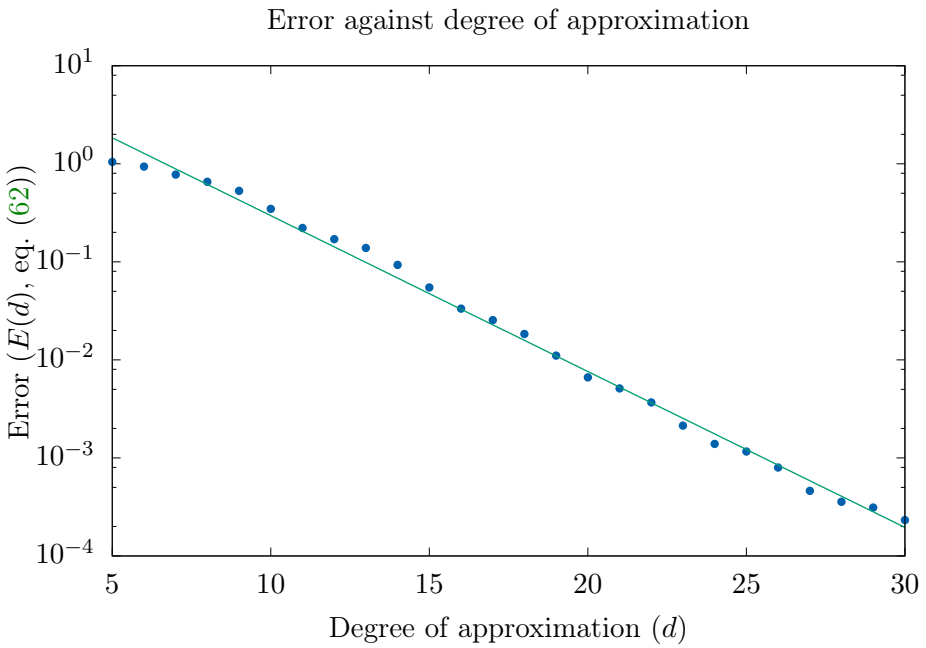

Figure 7. *Error* $E(d)$*, see* (62)*, for the fast/slow system* (73)*.*

of the following form

$$\frac{\partial u}{\partial t} = \frac{1}{\varepsilon^2}\mathcal{A}u + \frac{1}{\varepsilon}F(u) + \frac{1}{\varepsilon}Q\dot{W}, \tag{74}$$

posed in a bounded domain of $\mathbf{R}^m$ with suitable boundary conditions. We consider the set-up from [7]. In particular, the operator $\mathcal{A}$ in (74) is a differential operator, assumed to be negative and selfadjoint in a Hilbert space $\mathcal{H}$, and with compact resolvent. It is furthermore assumed that $\mathcal{A}$ has a finite dimensional kernel, denoted by $\mathcal{M}$. The term W denotes a cylindrical Wiener process on $\mathcal{H}$ and Q denotes the covariance operator of the noise, which is assumed to be positive, selfadjoint, and bounded. It is assumed that Q and $\mathcal{A}$ commute, and that the noise acts only on the orthogonal complement of $\mathcal{M}$, denoted by $\mathcal{M}^\perp$, see [7, Assumption 2.4]. The function $F(\cdot)$ is a polynomial function representing a nonlinearity that has to be such that the above scaling makes sense.[7]

Since $\mathcal{A}$ is selfadjoint with compact resolvent, there exists an orthonormal basis of $\mathcal{H}$ consisting of eigenfunctions of $\mathcal{A}$. We denote by $\{\lambda_k, e_k\}$ the eigenvalues and corresponding eigenfunctions of $\mathcal{A}$. We arrange the eigenpairs by increasing absolute value of the eigenvalues, so the m first eigenfunctions are in the kernel of the differential operator, $\mathcal{M} = \operatorname{span}\{e_1, \ldots, e_m\}$. Formally, the cylindrical Brownian motion can be expanded in the basis as $W(t) = \sum_{i=1}^{\infty} e_i\, w_i(t)$, where $\{w_i\}_{i=1}^{\infty}$ are independent Brownian motions. The assumption that the covariance operator Q commutes with the differential operator $\mathcal{A}$ means that this operator satisfies $Q\, e_i = q_i\, e_i$, while the assumption that the noise only acts on $\mathcal{M}^\perp$ implies that $q_i = 0$ for $i = 1, 2, \ldots, m$.

We now summarize how the dynamics of the slow modes in (74) can be approximated by solving a multiscale system of SDEs using the methodology developed in [2].

First, we write the solution of (74) as

$$u = x + y, \quad \text{with } x = \sum_{k=1}^{m} x_k\, e_k \text{ and } y = \sum_{k=m+1}^{\infty} y_k\, e_k.$$

Note that $x = \mathcal{P}u$, and $y = (I - \mathcal{P})u$, where $\mathcal{P}$ is the projection operator from $\mathcal{H}$ onto $\mathcal{M}$. By assumption, the noise term can be expanded in the same way, as $\sum_{k=1}^{\infty} q_k\, e_k\, \dot{w}_k(t)$. Substitution of these expansions in the SPDE gives:

$$\frac{d}{dt}\left(\sum_{k=1}^{m} x_k\, e_k + \sum_{k=m+1}^{\infty} y_k\, e_k\right) = -\frac{1}{\varepsilon^2}\sum_{k=m+1}^{\infty} \lambda_k\, y_k\, e_k + \frac{1}{\varepsilon}F(u) + \frac{1}{\varepsilon}\sum_{k=m+1}^{\infty} q_k\, e_k\, \dot{w}_k(t).$$

The equations that govern the evolution of the coefficients x_k and y_k can be obtained by taking the inner product (of $\mathcal{H}$) of both sides of the above equation by each of the eigenfunctions of the operator, and using orthonormality :

$$\begin{cases} \dot{x}_i = \dfrac{1}{\varepsilon}\langle F(u), e_i\rangle & i = 1, \ldots, m; \\ \dot{y}_i = -\dfrac{1}{\varepsilon^2}\lambda_i\, y_i + \dfrac{1}{\varepsilon}\langle F(u), e_i\rangle + \dfrac{1}{\varepsilon}q_i\, \dot{w}_i & i = m+1, m+2, \ldots \end{cases} \tag{75}$$

[7] i.e., the centering condition is satisfied.

Equation (75) can be written in the form

$$\begin{cases} \dot{x} = \dfrac{1}{\varepsilon} a(x,y), \\ \dot{y} = \dfrac{1}{\varepsilon^2} \mathcal{A}\, y + \dfrac{1}{\varepsilon} b(x,y) + \dfrac{1}{\varepsilon} Q\, \dot{W}, \end{cases} \tag{76}$$

where $a(x,y)$ and $b(x,y)$ are the projections of $F(u)$ on $\mathcal{M}$ and $\mathcal{M}^{\perp}$, respectively:

$$a(x,y) = \sum_{i=1}^{m} a^i(x,y)\, e_i \quad \text{with} \quad a^i(x,y) = \langle F(x+y), e_i \rangle,$$

and

$$b(x,y) = \sum_{i=m+1}^{\infty} b^i(x,y)\, e_i \quad \text{with} \quad b^i(x,y) = \langle F(x+y), e_i \rangle.$$

The scale separation now appears clearly. We now truncate the fast process in (76) as $y \approx \sum_{i=m+1}^{m+n} y_i\, e_i$ to derive the following finite dimensional system:

$$\begin{cases} \dot{x}_i = \dfrac{1}{\varepsilon} a^i(x,y) & i = 1, \ldots, m; \\ \dot{y}_i = -\dfrac{1}{\varepsilon^2} \lambda_i y_i + \dfrac{1}{\varepsilon}\, b^i(x,y) + \dfrac{1}{\varepsilon} q_i\, \dot{w}_i & i = m+1, \ldots m+n, \end{cases} \tag{77}$$

In [43], the authors investigate the use of the heterogeneous multiscale method (HMM) for solving the problem (77), and show that a good approximation can be obtained using this method. However, when the nonlinearity is a polynomial function of u, the function a in the system above, which also appears on the right-hand side of the Poisson equation, is polynomial in x and y. In addition, the generator of this system of stochastic differential equations is of Ornstein-Uhlenbeck type to leading order, and so its eigenfunctions are Hermite polynomials. This means that the right-hand side can be expanded exactly in Hermite polynomials, and so the exact effective coefficients can be computed. Note that although equivalent, applying the unitary transformation is not necessary in this case, as we can work directly with Hermite polynomials in the appropriate weighted L^2 space.

We consider the SPDE (74), with $\mathcal{A} = \frac{\partial^2}{\partial x^2} + 1$ and $F(u) = u^2 \frac{\partial u^2}{\partial x}$, posed on $[-\pi, \pi]$ with periodic boundary conditions:

$$\frac{\partial u}{\partial t} = \frac{1}{\varepsilon^2} \left(\frac{\partial^2}{\partial x^2} + 1 \right) u + \frac{1}{\varepsilon}\, u^2\, \frac{\partial u^2}{\partial x} + \frac{1}{\varepsilon}\, Q \dot{W}. \tag{78}$$

The eigenfunctions of $\mathcal{A}$ on $[-\pi, \pi]$ with periodic B.C. are

$$e_i = \begin{cases} \dfrac{1}{\sqrt{\pi}} \sin\left(\dfrac{i+1}{2}\, x \right) & \text{if } i \text{ is odd}, \\ \dfrac{1}{\sqrt{\pi}} \cos\left(\dfrac{i}{2}\, x \right) & \text{if } i \text{ is even}, \end{cases}$$

and the corresponding eigenvalues are $\lambda_i = 1 - \frac{(i+1)^2}{4}$ if i is odd and $\lambda_i = 1 - \frac{i^2}{4}$ if i is even. For the numerical experiments, we chose $q_i = 1/i$ if $3 \leq i \leq 5$, $q_i = 0$ otherwise, and we used 7 modes for the truncation of the Fourier series. In this case the null space of $\mathcal{A}$ is two-dimensional. We consider a noise process of the form:

$$Q\dot{W} = \sum_{i=3}^{\infty} q_i \, \dot{w}_i. \tag{79}$$

Following the methodology outlined above, we approximate the solution by a truncated Fourier series:

$$u \;=\; x_1 \, e_1 \;+\; x_2 \, e_2 \;+\; \sum_{i=3}^{n+2} y_i \, e_i. \tag{80}$$

Substituting in the nonlinearity and taking the inner product with each of the eigenfunctions, a system of equation of the type (77) is obtained. The operator $\mathcal{A}$ and the nonlinearity were chosen so that the centering condition is satisfied. The homogenized equation for the slow variables (x_1, x_2) reads

$$dX_t \;=\; \mathbf{F}(X_t)\, dt + \mathbf{A}(X_t)\, dW_t, \tag{81}$$

where $\mathbf{F}(\cdot)$ and $\mathbf{A}(\cdot)$ are given by equations (4a) and (4b), respectively, and W is a standard Wiener process in $\mathbf{R}^2$. The Euler-Maruyama solver was used for both the macro and micro solvers, and the parameters of the HMM were chosen as

$$(\delta t/\varepsilon^2, N_T, M, N, N') = (2^{-p}, 16, 1, 10{\times}2^{3p}, 2^p p). \tag{82}$$

Here δt is the time step of the micro-solver, N_T is the number of steps that are omitted in the time-averaging process to reduce transient effects, M is the number of samples used for ensemble averages, and N, N' are the number of time steps employed for the calculation of time averages and the discretization of integrals originating from Feynman-Kac representation formula (15), respectively. See [21, 62] for a more detailed description of the method and a detailed explanation of the parameters in (82). In Figures 8 and 9, we compare the solutions obtained using the HMM method with the one obtained using our approach, using the same macro-solver and the same replica of the driving Brownian motion for both, and with the initial condition $x_{i0} = 1.2$ for $i = 1, \ldots, m$. The former is denoted by $\hat{X}^n$ and the latter by X^n. Notice that when the value of the parameter p increases, the solution obtained using the HMM converges to the exact solution obtained using the Hermite spectral method.

We now investigate the dependence on the precision parameter p of the error between the homogenized coefficients. The same error measure as in [21] is used to compare the two methods:

$$E_p \;=\; \frac{\Delta t}{T} \left(\sum_{n \leq T/\Delta t} |\mathbf{F}^p_{HMM}(X^n) - \mathbf{F}_{Sp}(X^n)| \;+\; |\mathbf{A}^p_{HMM}(X^n) \;-\; \mathbf{A}_{Sp}(X^n)| \right). \tag{83}$$

Here $\mathbf{F}^p_{HMM}$ and $\mathbf{A}^p_{HMM}$ are the drift and diffusion coefficients obtained using the HMM with the precision parameter equal to p, while $\mathbf{F}_{Sp}$ and $\mathbf{A}_{Sp}$ are the coefficients given by the Hermite spectral method developed in this paper. Given the choice of parameters (82), the theory developed in [21] predicts that the error should decrease as $\mathcal{O}(2^{-p})$. This error is presented in Figure 10 as a function of the precision parameter p, showing a good agreement with the theory developed in [2, 21].

For the SPDE described above, solving the Poisson equation associated with (77) using Hermite polynomials does recover exactly the corresponding effective parameters, and the only source of error is the macroscopic discretization scheme. This is in sharp contrast with the HMM-based method developed in [1], for which the micro-averaging process to recover the effective coefficients represents a non-negligible computational cost.

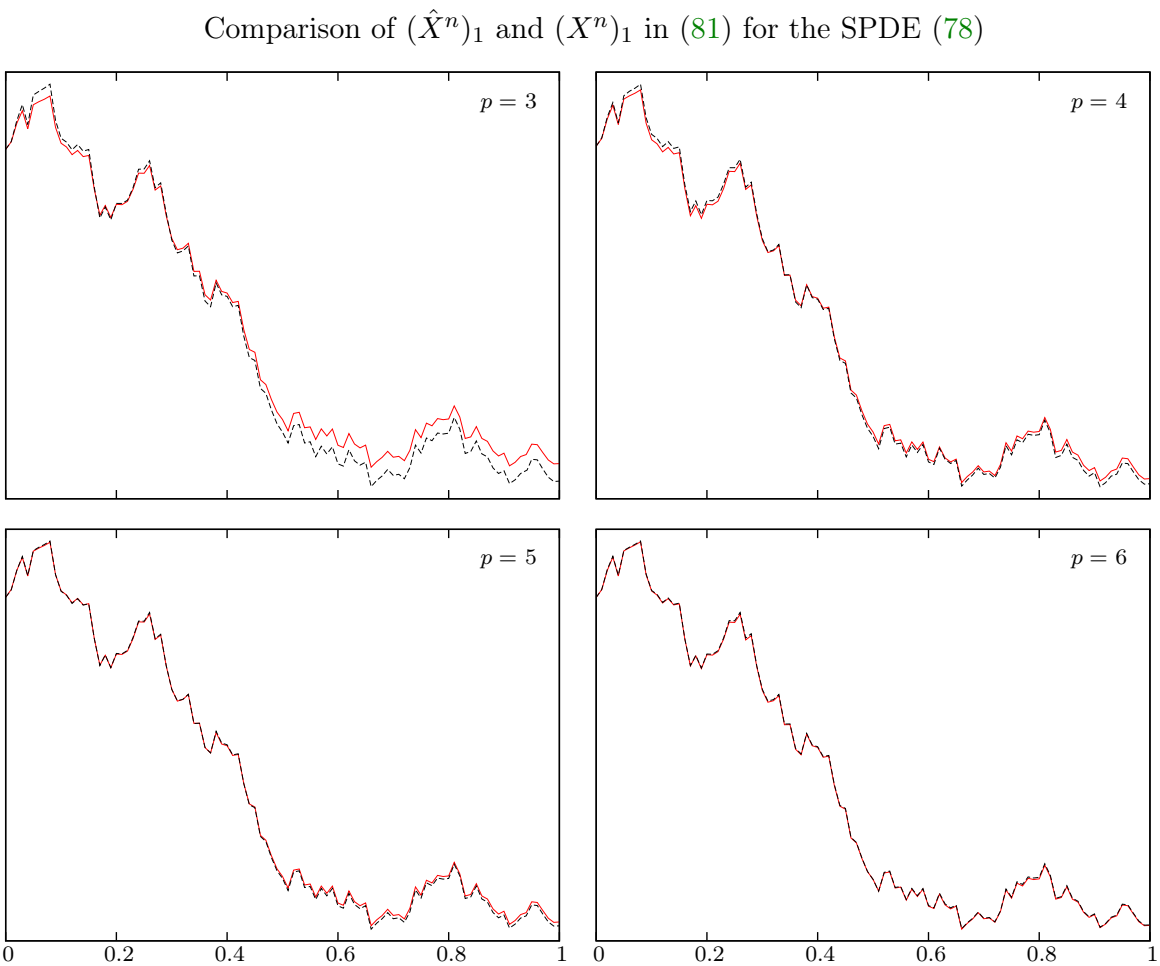

Figure 8. *Evolution of the coefficient x_1 of the first term in the Fourier expansion (80) of the the solution to the SPDE (78), obtained numerically by the HMM (black) and the Hermite spectral method (red), for one sample of the driving Brownian motion.*

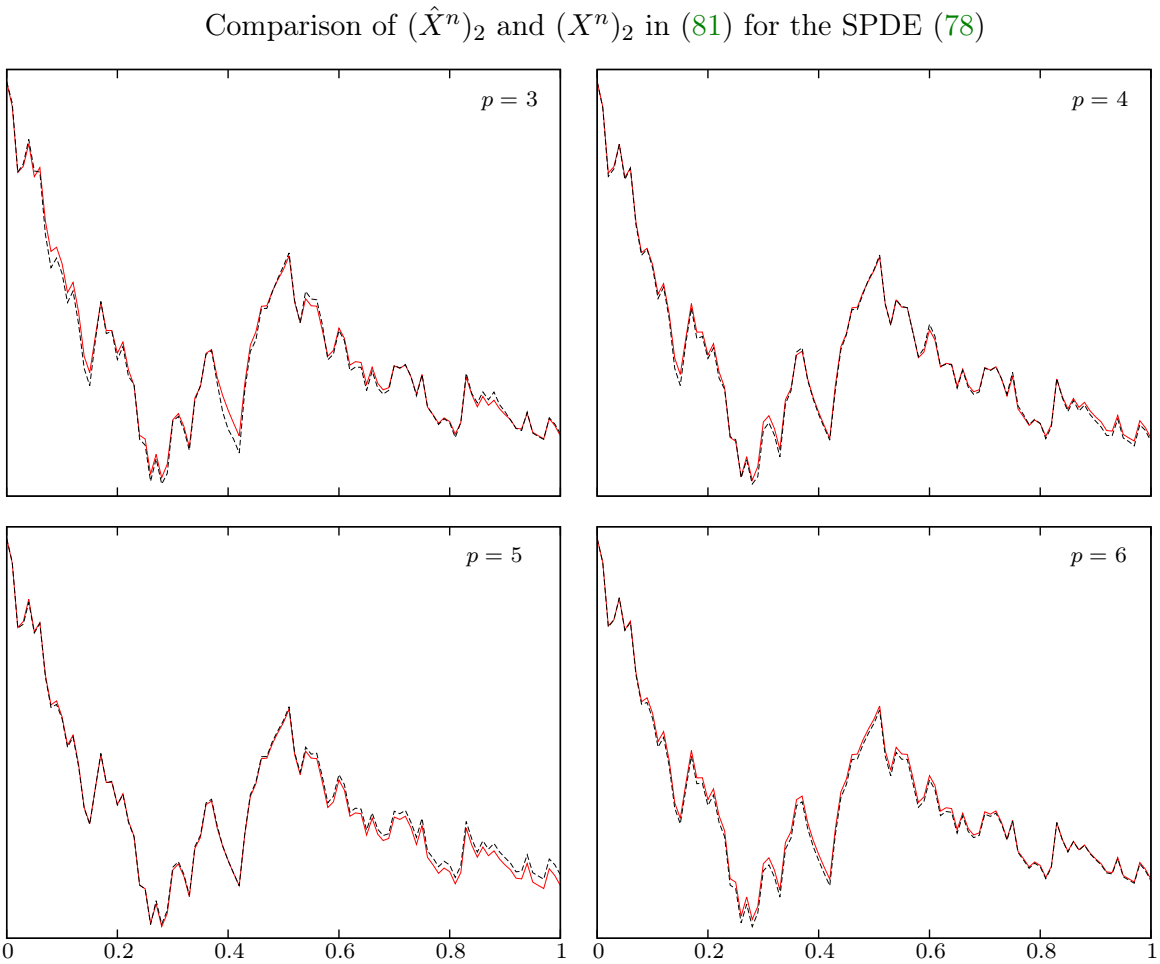

Figure 9. *Evolution of the coefficient x_2 of the second term in the Fourier expansion (80) of the the solution to the SPDE (78), obtained numerically by the HMM (black) and the Hermite spectral method (red), for one sample of the driving Brownian motion.*

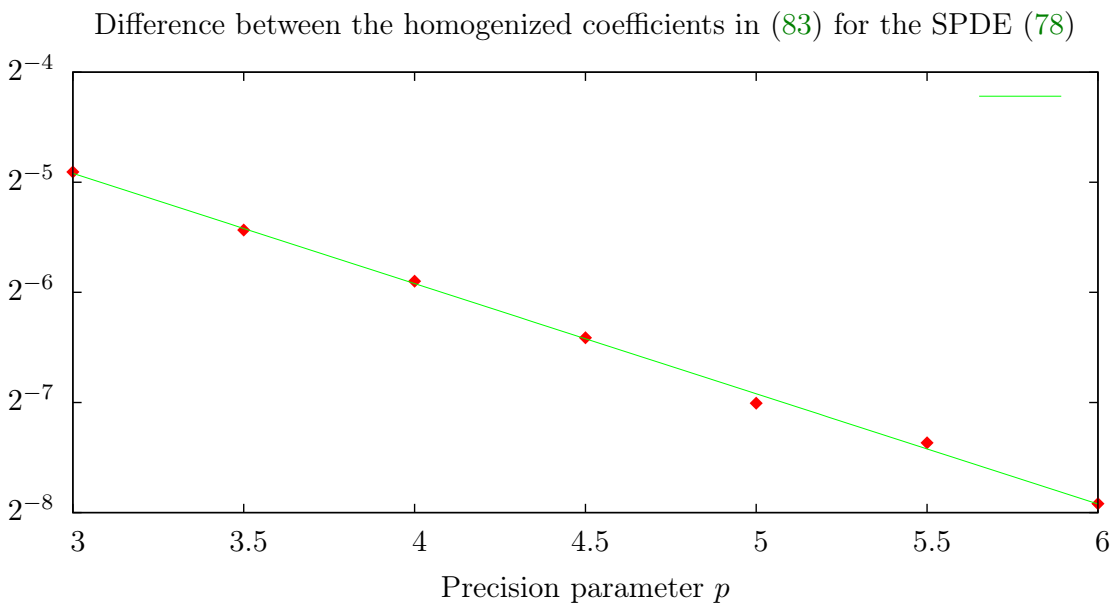

Figure 10. *Error between the homogenized coefficients (see (83)) for the SPDE (78), as a function of the precision parameter p. The green line, obtained by polynomial fitting, has slope* -1.01 *in the* $p - \log_2(E_p)$ *plane, which is close to the theoretical value of -1, showing a perfect agreement with the theory.*

7. Conclusion and Further Work. In this paper, we proposed a new approach for the numerical approximation of the slow dynamics of fast/slow SDEs for which a homogenized equation exists. Starting from the appropriate Poisson equation, the same unitary transformation as in [9] was utilized to obtain formulas for the drift and diffusion coefficients in terms of the solution to a Schrödinger equation. This equation is solved at each discrete time by means of a spectral method using Hermite functions, from which approximations of the homogenized drift and diffusion coefficients were calculated. A stochastic integrator was then used to evolve the slow variables by one time step, and the procedure is repeated.

Building on the work of [26], spectral convergence of the homogenized coefficients was rigorously established, from which weak convergence of the discrete approximation in time to the exact homogenized solution was derived. In the final section, the accuracy and efficiency of the proposed methodology were examined through numerical experiments.

The method presented, although not as general as the HMM, has proven more precise and more efficient for a broad class of problems. It performs particularly well for singularly perturbed SPDEs, and constitutes in this case a good complement to the HMM-based method presented in [2]. It also works comparatively very well when the fast dynamics is of relatively low dimension – typically less than or equal to 3 – and especially so when the potential is localized, since fewer Hermite functions are required to accurately resolve the Poisson equations in this situation. Our method also has several advantages compared to the approach taken in [9]: it does not require truncation of the domain, does not require the calculation of the eigenvalues and eigenfunctions of the Schrödinger operator, and has better asymptotic convergence properties.

The limitations of the method are two-fold; its generality is limited by the requirement of the gradient structure for fast dynamics, and its efficiency is limited by the curse of dimensionality, which causes the computation time to become prohibitive when the dimension of the state space of the fast process increases.

The extent to which some of these constraints can be lifted constitutes an interesting topic for future work. We believe that it is possible to generalize our method to a broader class of problems while retaining its efficiency and accuracy. When the fast process is not reversible, then tools from the spectral theory for non-selfadjoint problems could be used [31]. In the case of the underdamped Langevin dynamics, expansion of the solution in Hermite polynomials in both position and momentum can be performed [38]. On the other hand, high-dimensional integrals could be computed more efficiently. For example, an alternative to the tensorized quadrature approach taken in this work is to use a sparse grid method; such a method can in principle offer the same degree of polynomial exactness with a significantly lower number of nodes, see e.g. [27, 34]. Furthermore, efficient preconditioning techniques can be used in order to reduce the computational cost of the numerical solution. Preconditioning can be naturally implemented for nonreversible fast processes that have a well-defined structure, such as the underdamped Langevin dynamics, [57], or the Markovian approximation of the generalized Langevin equation, [50].

Acknowledgments. The authors thank Andrew Duncan, Gabriel Stoltz and Julien Roussel for useful discussions, and the anonymous referees for their valuable input.

Appendix A. Weighted Sobolev Spaces. In this section, we recall a few results about weighted Sobolev spaces needed for the analysis presented in Section 5. For more details on this topic, see [26, 64, 8, 44]. Throughout the appendix, V denotes a smooth confining potential whose derivatives are all bounded above by a polynomial and such that $\rho := e^{-V}$ is normalized.

Definition 14. *The weighted space $L^2(\mathbf{R}^n, \rho)$ is defined as*

$$L^2(\mathbf{R}^n, \rho) = \left\{ u \text{ measurable } : \int_{\mathbf{R}^n} u^2 \, \rho \, dy < \infty \right\}.$$

It is a Hilbert space for the inner product given by:

$$\langle u, v \rangle_\rho = \int_{\mathbf{R}^n} u \, v \, \rho \, dy.$$

Definition 15. *The weighted Sobolev space $H^s(\mathbf{R}^n, \rho)$, with $s \in \mathbf{N}$, is defined as*

$$H^s(\mathbf{R}^n, \rho) = \left\{ u \in L^2(\mathbf{R}^n, \rho) : \partial^\alpha u \in L^2(\mathbf{R}^n, \rho) \ \forall \, |\alpha| \leq s \right\}.$$

It is a Hilbert space for the inner product given by:

$$\langle u, v \rangle_{s,\rho} = \sum_{|\alpha| \leq s} \langle \partial^\alpha u, \partial^\alpha v \rangle_\rho$$

We also define the following spaces.

Definition 16. *Given $s \in \mathbf{N}$ and a positive selfadjoint operator $-\mathcal{L}$ on a Hilbert space H of functions on $\mathbf{R}^n$, we define $H^s(\mathbf{R}^n, \mathcal{L})$ as the space obtained by completion of $C_c^\infty(\mathbf{R}^n)$ for the inner product:*

$$\langle u, v \rangle_{s,\mathcal{L}} = \sum_{i=0}^{s} \langle (-\mathcal{L})^i u, v \rangle_H.$$

We will denote the associated norm by $\| \cdot \|_{s,\mathcal{L}}$.

It can be shown that $C_c^\infty(\mathbf{R}^n)$ is dense in $H^1(\mathbf{R}^n, \rho)$, see [63]. By integration by parts, this implies that $H^1(\mathbf{R}^n, \rho) = H^1(\mathbf{R}^n, \mathcal{L})$, where $-\mathcal{L}$ is the positive selfadjoint operator on $L^2(\mathbf{R}^n, \rho)$ defined by $\mathcal{L} = \Delta - \nabla V \cdot \nabla$. For the rest of this appendix, we make the additional assumption that the potential V satisfies

$$\lim_{|y| \to \infty} \left(\frac{1}{4} |\nabla V|^2 - \frac{1}{2} \Delta V \right) = \infty \quad \text{and} \quad \lim_{|y| \to \infty} |\nabla V| = \infty. \tag{84}$$

With this, the following compactness result holds.

Proposition 17. *Assume that (84) holds. Then the embedding $H^1(\mathbf{R}^n, \rho) \subset L^2(\mathbf{R}^n, \rho)$ is compact, and the measure ρ satisfies Poincaré inequality:*

$$\int_{\mathbf{R}^n} (u - \bar{u})^2 \, \rho \, dy \leq C \int_{\mathbf{R}^n} |\nabla u|^2 \, \rho \, dy \quad \forall u \in H^1(\mathbf{R}^n, \rho),$$

where $\bar{u} = \int_{\mathbf{R}^n} u \, \rho \, dy$.

Proof. See [44], sec. 8.5, p. 216. ∎

Remark 18. *Alternative conditions on the potential that ensure that the corresponding Gibbs measure satisfies a Poincaré inequality are presented in [42, Theorem 2.5].*

Now we consider the unitary transformation $e^{-V/2} : L^2(\mathbf{R}^n, \rho) \to L^2(\mathbf{R}^n)$, and characterize the spaces obtained by applying this mapping to the weighted Sobolev spaces.

Proposition 19. *The multiplication operator $e^{-V/2}$ is a unitary transformation from $H^s(\mathbf{R}^n, \mathcal{L})$ to $H^s(\mathbf{R}^n, \mathcal{H})$, where $-\mathcal{H}$ is the positive selfadjoint operator on $L^2(\mathbf{R}^n)$ defined by*

$$-\mathcal{H} = e^{-V/2}\,\mathcal{L}\,e^{V/2} = -\Delta + \left(\frac{|\nabla V|^2}{4} - \frac{\Delta V}{2}\right) =: -\Delta + W.$$

Proof. Since $(-\mathcal{H})^i = e^{-V/2}(-\mathcal{L})^i e^{V/2}$, $\langle u, v\rangle_{s,\mathcal{L}} = \langle e^{-V/2}u, e^{-V/2}v\rangle_{s,\mathcal{H}}$ for any $u, v \in C_c^\infty(\mathbf{R}^n)$ and any exponent $i \in \mathbf{N}$, from which the result follows by density. ∎

The space $H^1(\mathbf{R}^n, \mathcal{H})$, for $\mathcal{H}$ defined as above, is of particular relevance to this paper. This space can be equivalently defined by

$$H^1(\mathbf{R}^n, \mathcal{H}) = \left\{ u \in H^1(\mathbf{R}^n) : \int_{\mathbf{R}^n} |W| u^2\, dy < \infty \right\},$$

and, for $u \in H^1(\mathbf{R}^n, \mathcal{H})$,

$$\|u\|_{1,\mathcal{H}}^2 = \int_{\mathbf{R}^n} u^2 + |\nabla u|^2\, dy + \int_{\mathbf{R}^n} W u^2\, dy.$$

Appendix B. Hermite Polynomials and Hermite Functions.

In this appendix, we recall the results about Hermite polynomials and Hermite functions that are essential for the analysis presented in this paper.

Hermite polynomials. In one dimension, it is well-known that the polynomials

$$H_r(s) = \frac{(-1)^r}{\sqrt{r!}} \exp\left(\frac{s^2}{2}\right) \frac{\mathrm{d}^r}{\mathrm{d}s^r}\left(\exp\left(\frac{-s^2}{2}\right)\right) \qquad r = 0, 1, 2, \ldots \tag{85}$$

form a complete orthonormal basis of $L^2\left(\mathbf{R}, G_{(0,1)}\right)$, where $G_{(0,1)}$ is the Gaussian density of mean 0 and variance 1. These polynomials can be naturally extended to the multidimensional case. For $\mu \in \mathbf{R}^n$ and a symmetric positive definite matrix $\Sigma \in \mathbf{R}^{n\times n}$, consider the Gaussian density $G_{(\mu,\Sigma)}$ of mean μ and covariance matrix Σ. Let D and Q be diagonal and orthogonal matrices such that $\Sigma = QDQ^T$, and note $S = QD^{1/2}$, such that $\Sigma = SS^T$. The polynomials defined for $\alpha \in \mathbf{N}^n$ by

$$H_\alpha(y; \mu, \Sigma) = H_\alpha(S^{-1}(y-\mu); 0, I), \text{ with } H_\alpha(z; 0, I) = \prod_{k=1}^n H_{\alpha_k}(z_k), \tag{86}$$

form a complete orthonormal basis of $L^2(\mathbf{R}^n, G_{(\mu,\Sigma)})$. Note that the Hermite polynomial corresponding to a multi-index α depends on the orthogonal matrix Q chosen. When μ and Σ are clear from the context, we will sometimes omit them to simplify the notation.

In addition to forming a complete orthonormal basis, the Hermite polynomials are the eigenfunctions of the Ornstein-Ulhenbeck operator

$$-\mathcal{L}_{\mu,\Sigma} = \Sigma^{-1}(y-\mu)\cdot\nabla - \Delta,$$

which is the operator $\mathcal{L}$ from the previous appendix for $\rho = G_{(\mu,\Sigma)}$. The eigenvalue associated to $H_\alpha(y;\mu,\Sigma)$ is given by

$$\lambda_\alpha = \sum_{i=1}^{n} \alpha_i \lambda_i, \tag{87}$$

where $\{\lambda_i\}_{i=1}^n$ are the diagonal elements of D^{-1}.

Proof. Let us introduce, for an orthogonal matrix $P \in \mathbf{R}^{n\times n}$ and a vector $u \in \mathbf{R}^n$, the rotation operator $\mathcal{R}_P$ and the translation operator $\mathcal{T}_u$ defined respectively by $\mathcal{R}_P f(\cdot) = f(P^T\cdot)$ and $\mathcal{T}_u f(\cdot) = f(\cdot - u)$. With these notations, note that $\mathcal{L}_{\mu,\Sigma} = \mathcal{T}_\mu \circ \mathcal{R}_Q \circ \mathcal{L}_{0,D} \circ \mathcal{R}_{Q^T} \circ \mathcal{T}_{-\mu}$ and $H_\alpha(\cdot;\mu,\Sigma) = \mathcal{T}_\mu \circ \mathcal{R}_Q H_\alpha(\cdot;0,D)$, so it is sufficient to show that $H_\alpha(\cdot;0,D)$ is an eigenfunction of $\mathcal{L}_{0,D}$. Writing $\mathcal{L}_{0,D}$ as a sum of one-dimensional operators, we see that it is enough to prove this in one dimension, which follows from a simple recursion. ∎

It follows that Hermite polynomials are also orthogonal in $H^s(\mathbf{R}^n, \mathcal{L}_{\mu,\Sigma})$, and that the norm of $H_\alpha(y;\mu,\Sigma)$ in that space is given by $(1+\lambda_\alpha+\lambda_\alpha^2+\cdots+\lambda_\alpha^s)^{1/2}$. Noting that $1+r+r^2+\cdots+r^s \le e^{\frac{1}{r}} r^s$ for $r>0$ and that $\lambda_\alpha \to \infty$ when $|\alpha|\to\infty$, this implies that, for a function $u \in H^s(\mathbf{R}^n, \mathcal{L}_{\mu,\Sigma})$ and $c_\alpha = \langle u, H_\alpha(\cdot;\mu,\Sigma)\rangle_{G_{(\mu,\Sigma)}}$,

$$\left(c_0^2 + \sum_{|\alpha|>0} \lambda_\alpha^s c_\alpha^2\right) \le \|u\|^2_{s,\mathcal{L}_{\mu,\Sigma}} \le L\left(c_0^2 + \sum_{|\alpha|>0} \lambda_\alpha^s c_\alpha^2\right), \quad \text{where } L = \max_{|\alpha|>0} e^{\frac{1}{\lambda_\alpha}}. \tag{88}$$

Hermite polynomials have very good approximation properties for smooth functions in $L^2(\mathbf{R}^n, G_{(\mu,\Sigma)})$. In Proposition 20, we note $\pi(\cdot, \mathbf{P}_d) : L^2(\mathbf{R}^n, G_{(\mu,\Sigma)}) \to \mathbf{P}_d$ the $L^2(\mathbf{R}^n, G_{(\mu,\Sigma)})$ projection operator on the space of polynomials of degree less than or equal to d.

Proposition 20 (Approximation by polynomials in weighted spaces). *Let Σ be a symmetric positive definite matrix, and suppose that $f \in H^s(\mathbf{R}^n, \mathcal{L}_{\mu,\Sigma})$. Then*

$$\|f - \pi(f, \mathbf{P}_d)\|_{r,\mathcal{L}_{\mu,\Sigma}} \le C(\Sigma, r, s)\,(d+1)^{-\frac{(s-r)}{2}}\,\|f\|_{s,\mathcal{L}_{\mu,\Sigma}},$$

for $r \in \mathbf{N}$ such that $0 \le r \le s$.

Proof. See [26, Theorem 3.1]. From (88), we have that

$$\|f - \pi(f, \mathbf{P}_d)\|^2_{r,\mathcal{L}_{\mu,\Sigma}} \le L \sum_{|\alpha|>d} \lambda_\alpha^r c_\alpha^2 \le L\, M^{r-s} \sum_{\alpha\in\mathbf{N}^n} \lambda_\alpha^s c_\alpha^2 \le LM^{r-s}\|f\|^2_{s,\mathcal{L}_{\mu,\Sigma}},$$

with $c_\alpha = \langle f, H_\alpha(\cdot;\mu,\Sigma)\rangle_{G_{(\mu,\Sigma)}}$ and $M = \min_{|\alpha|>d} \lambda_\alpha$. Since $\lambda_\alpha > C(\Sigma)\,|\alpha|$, the conclusion follows. ∎

Hermite Functions. Hermite functions can be defined from Hermite polynomials as follows:

Definition 21. *Given $\mu \in \mathbf{R}^n$, and $\Sigma \in \mathbf{R}^{n\times n}$ positive definite, we define the Hermite functions $h_\alpha(y;\mu,\Sigma)$ by:*

$$h_\alpha(y;\mu,\Sigma) = \sqrt{G_{(\mu,\Sigma)}}\, H_\alpha(y;\mu,\Sigma) \qquad \text{for } \alpha \in \mathbf{N}^n.$$

These functions form a complete orthonormal basis of $L^2(\mathbf{R}^n)$. Since they are obtained from Hermite polynomials by a multiplication by $\sqrt{G_{(\mu,\Sigma)}}$, they satisfy:

Proposition 22. *Given $\mu \in \mathbf{R}^n$ and $\Sigma \in \mathbf{R}^{n\times n}$ positive definite, the Hermite functions $h_\alpha(y;\mu,\Sigma)$ are the eigenfunctions of the operator:*

$$-\mathcal{H}_{\mu,\Sigma} = (G_{(\mu,\Sigma)})^{\frac{1}{2}} (-\mathcal{L}_{\mu,\Sigma}) (G_{(\mu,\Sigma)})^{-\frac{1}{2}} = -\Delta + \left(\frac{(y-\mu)^T\Sigma^{-2}(y-\mu)}{4} - \frac{\operatorname{tr}\Sigma^{-1}}{2} \right),$$

with the same eigenvalues as in (87).

Hermite functions inherit the good approximation properties of Hermite polynomials expressed in Proposition 20. In the following result, π refers to the $L^2(\mathbf{R}^n)$ projection operator, i.e.

$$\pi\left(f, \sqrt{G_{(\mu,\Sigma)}}\mathbf{P}_d\right) = \sum_{|\alpha|\le d} \langle f, h_\alpha(\cdot;\mu,\Sigma)\rangle\, h_\alpha(\cdot;\mu,\Sigma). \tag{89}$$

Corollary 23 (Approximation by Hermite functions in flat space). *Let $\mu \in \mathbf{R}^n$ and $\Sigma \in \mathbf{R}^{n\times n}$ be a symmetric positive definite matrix, and suppose that $f \in H^s(\mathbf{R}^n, \mathcal{H}_{\mu,\Sigma})$. Then*

$$\|f - \pi\left(f, \sqrt{G_{(\mu,\Sigma)}}\mathbf{P}_d\right)\|_{r,\mathcal{H}_{\mu,\Sigma}} \le C(\Sigma,r,s)\,(d+1)^{-\frac{(s-r)}{2}}\, \|f\|_{s,\mathcal{H}_{\mu,\Sigma}},$$

for any $r \in \mathbf{N}$ such that $0 \le r \le s$.

Appendix C. Proof of (48).

In this appendix, we show the Lipschitz continuity of the effective coefficients. The proof relies on the following technical lemma.

Lemma 24. *For any $m \in \mathbf{N}$ and any symmetric positive semi-definite matrices $M_1, M_2 \in \mathbf{R}^{m\times m}$, there exists $c(m) > 0$ such that*

$$\sup_{|x|=1} \left| \sqrt{x^T M_2^2 x} - \sqrt{x^T M_1^2 x} \right| \ge c(m)\,|M_2 - M_1|\,,$$

where $|\cdot|$ is the matrix 2-norm.

Proof. To alleviate the notations, let $M = M_1$ and $\Delta = M_2 - M_1$. The inequality is trivially satisfied when $\Delta = 0$; we assume now that $\Delta \neq 0$. Without loss of generality, we assume that $|\Delta| = 1$, that Δ has an eigenvalue equal to $+1$, and that M is diagonal with its

eigenvalues occuring in descending order on the diagonal. For $x \in \mathbf{R}^m$ arbitrary,

$$\begin{aligned}\sqrt{x^T(M+\Delta)^2x} &= \sqrt{x^TM^2x + x^T(M\Delta+\Delta M)x + x^T\Delta^2x} \\ &= \sqrt{\left(\sqrt{x^TM^2x} + \sqrt{x^T\Delta^2x}\right)^2 - 2\left(\sqrt{x^TM^2x}\sqrt{x^T\Delta^2x} - x^TM\Delta x\right)}\end{aligned}$$

By Cauchy-Schwarz inequality, the expression in the second brackets is non-negative, so assuming that x is such that the expression in first brackets is positive,

$$\begin{aligned}\left|\sqrt{x^T(M+\Delta)^2x} - \sqrt{x^TM^2x}\right| &\geq \frac{\left|2\,x^TM\Delta x + x^T\Delta^2x\right|}{2(\sqrt{x^TM^2x} + \sqrt{x^T\Delta^2x})} \\ &\geq \frac{2\,x^TM\Delta x + x^T\Delta^2x}{2(\sqrt{x^TM^2x}+1)} =: R_{M,\Delta}(x)\end{aligned}$$

To conclude the proof, we show that for any m, M, Δ as above, there exists $x \in \mathbf{R}^m$ such that $R_{M,\Delta}(x)$ is bounded from below by a constant depending only on m. Assume that there exist $\{M_i\}_{i=1}^\infty$ and $\{\Delta_i\}_{i=1}^\infty$ such that $\sup_{|x|=1} R_{M_i,\Delta_i}(x) \to 0$, and let x_i denote an eigenvector corresponding to the eigenvalue 1 of Δ_i. Passing to a subsequence if necessary, we can assume that $x_i \to x_\infty$. By assumption,

$$R_{M_i,\Delta_i}(x_i) = \frac{2\,x_i^TM_ix_i + 1}{2\sqrt{x_i^TM_i^2x_i} + 2} \to 0 \Rightarrow \frac{x_i^TM_ix_i}{\sqrt{x_i^TM_i^2x_i}} \to 0 \Rightarrow x_i^T\left(\frac{M_i}{|M_i|}\right)x_i \to 0,$$

so in particular the first component of x_∞ is equal to 0. Let now $\{y_i\}_{i=1}^\infty$ be the sequence obtained from $\{x_i\}_{i=1}^\infty$ by replacing the first component by 0. By assumption,

$$R_{M_i,\Delta_i}(y_i) = \frac{(M_i\,y_i)^T(\Delta_iy_i)}{|M_iy_i|+1} + \frac{|\Delta_iy_i|^2}{2\,|M_iy_i|+2} \to 0.$$

Since $(\Delta_iy_i - y_i) \to 0$, both terms must tend to 0. This implies that $y_i^TM_iy_i/\,|M_iy_i| \to 0$, so the second component of x_∞ is 0 too. Continuing this way, we obtain $x_\infty = 0$, which is a contradiction because $x_i \to x_\infty$ and $|x_i| = 1$. ∎

Proof of (48). For the drift coefficient, we use (16) and Assumption 4.3:

$$\begin{aligned}|\mathbf{F}(a) - \mathbf{F}(b)| \leq &\int_{\mathbf{R}^n} |\nabla_x\boldsymbol{\phi}(a,y) - \nabla_x\boldsymbol{\phi}(b,y)|\;|\mathbf{f}(a,y)|\;e^{-V} \\ &+ |\nabla_x\boldsymbol{\phi}(b,y)|\;|\mathbf{f}(a,y) - \mathbf{f}(b,y)|\;e^{-V}\,dy \\ \leq |a-b|&\int_{\mathbf{R}^n} \sup_{x\in(a,b)} |\nabla_x\nabla_x\boldsymbol{\phi}(x,y)\,\mathbf{f}(a,y)|\;e^{-V} \\ &+ \sup_{x\in(a,b)} |\nabla_x\boldsymbol{\phi}(b,y)\,\nabla_x\mathbf{f}(x,y)|\;e^{-V}\,dy \\ \leq |a-b|&\int_{\mathbf{R}^n} 2\,C_0\,L\,(1+|y|^{\ell_0})(1+|y|^q)\,e^{-V}\,dy.\end{aligned}$$

The proof of the Lipschitz continuity of $A(\cdot)$ relies on Lemma 24. By virtue of this result, it is sufficient to show that, for all $\boldsymbol{\xi} \in \mathbf{R}^m$ with $\boldsymbol{\xi}^T\boldsymbol{\xi} = 1$, $(\boldsymbol{\xi}^T\mathbf{D}(x)\boldsymbol{\xi})^{1/2}$ is Lipschitz continuous with a Lipschitz constant independent of $\boldsymbol{\xi}$. Let $\boldsymbol{\xi} \in \mathbf{R}^m$ be given, and let us introduce the notations $f_\xi = \mathbf{f}^T\boldsymbol{\xi}$, $\phi_\xi = \boldsymbol{\phi}^T\boldsymbol{\xi}$, and $\boldsymbol{\alpha}_\xi = \boldsymbol{\alpha}^T\boldsymbol{\xi}$. Clearly, ϕ_ξ is the centered solution of $-\mathcal{L}\phi_\xi = f_\xi$, so

$$\boldsymbol{\xi}^T\mathbf{D}(x)\boldsymbol{\xi} = \int_{\mathbf{R}^n} \left(|\boldsymbol{\alpha}_\xi|^2 - \mathcal{L}\phi_\xi\,\phi_\xi\right) e^{-V}\,dy =: D_1(x) + D_2(x).$$

From the triangle inequality and (13),

$$\begin{aligned}\left|\sqrt{D_1(x+\Delta x)} - \sqrt{D_1(x)}\right| &\le \left(\int_{\mathbf{R}^n} |\mathbf{a}_\xi(x+\Delta x, y) - \mathbf{a}_\xi(x,y)|^2\, e^{-V(y)}\,dy\right)^{\frac{1}{2}} \\ &\le |\Delta x| \left(\int_{\mathbf{R}^n} C(y)^2 e^{-V(y)}\,dy\right)^{\frac{1}{2}}.\end{aligned}$$

Likewise, using (14) and (16),

$$\begin{aligned}\left|\sqrt{D_2(x+\Delta x)} - \sqrt{D_2(x)}\right| &\le \left(\int_{\mathbf{R}^n} |\phi_\xi(x+\Delta x, y) - \phi_\xi(x,y)|\;|f_\xi(x+\Delta x, y) - f_\xi(x,y)|\; e^{-V(y)} dy\right)^{\frac{1}{2}} \\ &\le |\Delta x| \left(\int_{\mathbf{R}^n} KL(1+|y|^q)(1+|y|^{m_1})\, e^{-V(y)}\,dy\right)^{\frac{1}{2}}.\end{aligned}$$

Thus $\sqrt{\boldsymbol{\xi}^T\mathbf{D}(x)\boldsymbol{\xi}}$ is Lipschitz continuous with a constant independent of $\boldsymbol{\xi}$. ∎

REFERENCES

[1] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden, *The heterogeneous multiscale method*, Acta Numer., 21 (2012), pp. 1–87.

[2] A. Abdulle and G. A. Pavliotis, *Numerical methods for stochastic partial differential equations with multiple scales*, J. Comput. Phys., 231 (2012), pp. 2482–2497, doi:10.1016/j.jcp.2011.11.039, http://dx.doi.org/10.1016/j.jcp.2011.11.039.

[3] S. Agmon, *Lectures on exponential decay of solutions of second-order elliptic equations: bounds on eigenfunctions of N-body Schrödinger operators*, vol. 29 of Mathematical Notes, Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo, 1982.

[4] G. E. Andrews, R. Askey, and R. Roy, *Special functions*, vol. 71 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1999, doi:10.1017/CBO9781107325937, http://dx.doi.org/10.1017/CBO9781107325937.

[5] A. Arnold, J.-P. Bartier, and J. Dolbeault, *Interpolation between logarithmic Sobolev and Poincaré inequalities*, Commun. Math. Sci., 5 (2007), pp. 971–979, http://projecteuclid.org/euclid.cms/1199377560.

[6] W. Beckner, *A generalized Poincaré inequality for Gaussian measures*, Proc. Amer. Math. Soc., 105 (1989), pp. 397–400, doi:10.2307/2046956, http://dx.doi.org/10.2307/2046956.

[7] D. Blömker, M. Hairer, and G. A. Pavliotis, *Multiscale analysis for stochastic partial differential equations with quadratic nonlinearities*, Nonlinearity, 20 (2007), pp. 1721–1744, doi:10.1088/0951-7715/20/7/009, http://dx.doi.org/10.1088/0951-7715/20/7/009.

[8] V. I. Bogachev, N. V. Krylov, M. Röckner, and S. V. Shaposhnikov, *Fokker-Planck-Kolmogorov equations*, vol. 207 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2015, doi:10.1090/surv/207, http://dx.doi.org/10.1090/surv/207.

[9] V. Bonnaillie-Noël, J. A. Carrillo, T. Goudon, and G. A. Pavliotis, *Efficient numerical calculation of drift and diffusion coefficients in the diffusion approximation of kinetic equations*, IMA Journal of Numerical Analysis, (2016), doi:10.1093/imanum/drv066, http://imajna.oxfordjournals.org/content/early/2016/03/04/imanum.drv066.abstract.

[10] J. P. Boyd, *Asymptotic coefficients of Hermite function series*, J. Comput. Phys., 54 (1984), pp. 382–410, doi:10.1016/0021-9991(84)90124-4, http://dx.doi.org/10.1016/0021-9991(84)90124-4.

[11] C.-E. Bréhier, *Analysis of an HMM time-discretization scheme for a system of stochastic PDEs*, SIAM J. Numer. Anal., 51 (2013), pp. 1185–1210, doi:10.1137/110853078, http://dx.doi.org/10.1137/110853078.

[12] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil, *Particle, kinetic, and hydrodynamic models of swarming*, in Mathematical modeling of collective behavior in socio-economic and life sciences, Model. Simul. Sci. Eng. Technol., Birkhäuser Boston, Inc., Boston, MA, 2010, pp. 297–336, doi:doi:10.1007/978-0-8176-4946-3_12, http://dx.doi.org/10.1007/978-0-8176-4946-3_12.

[13] R. R. Coifman, I. G. Kevrekidis, S. Lafon, M. Maggioni, and B. Nadler, *Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems*, Multiscale Model. Simul., 7 (2008), pp. 842–864, doi:10.1137/070696325, http://dx.doi.org/10.1137/070696325.

[14] R. R. Coifman and S. Lafon, *Diffusion maps*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 5–30, doi:10.1016/j.acha.2006.04.006, http://dx.doi.org/10.1016/j.acha.2006.04.006.

[15] R. R. Coifman and S. Lafon, *Geometric harmonics: a novel tool for multiscale out-of-sample extension of empirical functions*, Appl. Comput. Harmon. Anal., 21 (2006), pp. 31–52, doi:10.1016/j.acha.2005.07.005, http://dx.doi.org/10.1016/j.acha.2005.07.005.

[16] C. Cuthbertson, G. A. Pavliotis, A. Rafailidis, and P. Wiberg, *Asymptotic analysis for foreign exchange derivatives with stochastic volatility*, Int. J. Theor. Appl. Finance, 13 (2010), pp. 1131–1147, doi:10.1142/S0219024910006145, http://dx.doi.org/10.1142/S0219024910006145.

[17] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, vol. 152 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, second ed., 2014, doi:10.1017/CBO9781107295513, http://dx.doi.org/10.1017/CBO9781107295513.

[18] A. B. Duncan, S. Kalliadasis, G. A. Pavliotis, and M. Pradas, *Noise-induced transitions in rugged energy landscapes*, Phys. Rev. E, 94 (2016), p. 032107, doi:10.1103/PhysRevE.94.032107, http://link.aps.org/doi/10.1103/PhysRevE.94.032107.

[19] W. E, *Principles of multiscale modeling*, Cambridge University Press, Cambridge, 2011.

[20] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden, *Heterogeneous multiscale methods: a review*, Commun. Comput. Phys., 2 (2007), pp. 367–450.

[21] W. E, D. Liu, and E. Vanden-Eijnden, *Analysis of multiscale methods for stochastic differential equations*, Comm. Pure Appl. Math., 58 (2005), pp. 1544–1585, doi:10.1002/cpa.20088, http://dx.doi.org/10.1002/cpa.20088.

[22] L. C. Evans, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010, doi:10.1090/gsm/019, http://dx.doi.org/10.1090/gsm/019.

[23] J. C. M. Fok, B. Guo, and T. Tang, *Combined Hermite spectral-finite difference method for the Fokker-Planck equation*, Math. Comp., 71 (2002), pp. 1497–1528, doi:10.1090/S0025-5718-01-01365-5, http://dx.doi.org/10.1090/S0025-5718-01-01365-5.

[24] G. Froyland, G. A. Gottwald, and A. Hammerlindl, *A computational method to extract macroscopic variables and their dynamics in multiscale systems*, SIAM J. Appl. Dyn. Syst., 13 (2014), pp. 1816–1846, doi:10.1137/130943637, http://dx.doi.org/10.1137/130943637.

[25] G. Froyland, G. A. Gottwald, and A. Hammerlindl, *A trajectory-free framework for analysing multiscale systems*, Phys. D, 328/329 (2016), pp. 34–43, doi:10.1016/j.physd.2016.04.010, http://dx.doi.org/10.1016/j.physd.2016.04.010.

[26] J. Gagelman and H. Yserentant, *A spectral method for Schrödinger equations with smooth confinement potentials*, Numer. Math., 122 (2012), pp. 383–398, doi:10.1007/s00211-012-0458-8, http://dx.doi.org/10.1007/s00211-012-0458-8.

[27] T. Gerstner and M. Griebel, *Numerical integration using sparse grids*, Numer. Algorithms, 18 (1998), pp. 209–232, doi:10.1023/A:1019129717644, http://dx.doi.org/10.1023/A:1019129717644.

[28] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Classics in

Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[29] V. Gol′dshtein and A. Ukhlov, *Weighted Sobolev spaces and embedding theorems*, Trans. Amer. Math. Soc., 361 (2009), pp. 3829–3850, doi:10.1090/S0002-9947-09-04615-7, http://dx.doi.org/10.1090/S0002-9947-09-04615-7.

[30] D. Gottlieb and S. A. Orszag, *Numerical analysis of spectral methods: theory and applications*, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1977. CBMS-NSF Regional Conference Series in Applied Mathematics, No. 26.

[31] B. Helffer and F. Nier, *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, vol. 1862 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2005, doi:10.1007/b104762, http://dx.doi.org/10.1007/b104762.

[32] D. J. Higham, X. Mao, and A. M. Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equations*, SIAM J. Numer. Anal., 40 (2002), pp. 1041–1063 (electronic), doi:10.1137/S0036142901389530, http://dx.doi.org/10.1137/S0036142901389530.

[33] P. D. Hislop and I. M. Sigal, *Introduction to spectral theory*, vol. 113 of Applied Mathematical Sciences, Springer-Verlag, New York, 1996, doi:10.1007/978-1-4612-0741-2, http://dx.doi.org/10.1007/978-1-4612-0741-2. With applications to Schrödinger operators.

[34] V. Kaarnioja, *Smolyak quadrature*, (2013).

[35] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, vol. 113 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1991, doi:10.1007/978-1-4612-0949-2, http://dx.doi.org/10.1007/978-1-4612-0949-2.

[36] I. G. Kevrekidis, C. W. Gear, J. M. Hyman, P. Kevrekidis, O. Runborg, and C. Theodoropoulos, *Equation-free, coarse-grained multiscale computation: enabling microscopic simulators to perform system-level analysis*, Commun. Math. Sci., 1 (2003), pp. 715–762, http://projecteuclid.org/euclid.cms/1119655353.

[37] P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, vol. 23 of Applications of Mathematics (New York), Springer-Verlag, Berlin, 1992, doi:10.1007/978-3-662-12616-5, http://dx.doi.org/10.1007/978-3-662-12616-5.

[38] J. C. Latorre, G. A. Pavliotis, and P. R. Kramer, *Corrections to Einstein's relation for Brownian motion in a tilted periodic potential*, J. Stat. Phys., 150 (2013), pp. 776–803, doi:10.1007/s10955-013-0692-1, http://dx.doi.org/10.1007/s10955-013-0692-1.

[39] B. Leimkuhler, C. Matthews, and G. Stoltz, *The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics*, IMA J. Numer. Anal., 36 (2016), pp. 13–79, doi:10.1093/imanum/dru056, http://dx.doi.org/10.1093/imanum/dru056.

[40] T. Lelièvre, M. Rousset, and G. Stoltz, *Free energy computations*, Imperial College Press, London, 2010, doi:10.1142/9781848162488, http://dx.doi.org/10.1142/9781848162488. A mathematical perspective.

[41] T. Lelièvre and G. Stoltz, *Partial differential equations and stochastic methods in molecular dynamics*, Acta Numer., 25 (2016), pp. 681–880, doi:10.1017/S0962492916000039, http://dx.doi.org/10.1017/S0962492916000039.

[42] T. Lelièvre and G. Stoltz, *Partial differential equations and stochastic methods in molecular dynamics*, Acta Numer., 25 (2016), pp. 681–880, doi:10.1017/S0962492916000039, http://dx.doi.org/10.1017/S0962492916000039.

[43] T. Li, A. Abdulle, and W. E, *Effectiveness of implicit methods for stiff stochastic differential equations*, Commun. Comput. Phys., 3 (2008), pp. 295–307.

[44] L. Lorenzi and M. Bertoldi, *Analytical methods for Markov semigroups*, vol. 283 of Pure and Applied Mathematics (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, 2007.

[45] A. Majda, I. Timofeyev, and E. V. Eijnden, *A mathematical framework for stochastic climate models*, Comm. Pure Appl. Math., 54 (2001), pp. 891–974.

[46] X. Mao, *Stochastic differential equations and applications*, Horwood Publishing Limited, Chichester, second ed., 2008, doi:10.1533/9780857099402, http://dx.doi.org/10.1533/9780857099402.

[47] P. A. Markowich and C. Villani, *On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis*, Mat. Contemp., 19 (2000), pp. 1–29. VI Workshop on Partial Differential Equations, Part II (Rio de Janeiro, 1999).

[48] G. Metafune, D. Pallara, and E. Priola, *Spectrum of Ornstein-Uhlenbeck operators in L^p spaces*

with respect to invariant measures, J. Funct. Anal., 196 (2002), pp. 40–60, doi:10.1006/jfan.2002.3978, http://dx.doi.org/10.1006/jfan.2002.3978.

[49] G. N. Milstein and M. V. Tretyakov, *Stochastic numerics for mathematical physics*, Scientific Computation, Springer-Verlag, Berlin, 2004, doi:10.1007/978-3-662-10063-9, http://dx.doi.org/10.1007/978-3-662-10063-9.

[50] M. Ottobre and G. A. Pavliotis, *Asymptotic analysis for the generalized Langevin equation*, Nonlinearity, 24 (2011), pp. 1629–1653, doi:10.1088/0951-7715/24/5/013, http://dx.doi.org/10.1088/0951-7715/24/5/013.

[51] E. Pardoux and A. Y. Veretennikov, *On the Poisson equation and diffusion approximation. I*, Ann. Probab., 29 (2001), pp. 1061–1085, doi:10.1214/aop/1015345596, http://dx.doi.org/10.1214/aop/1015345596.

[52] E. Pardoux and A. Y. Veretennikov, *On Poisson equation and diffusion approximation. II*, Ann. Probab., 31 (2003), pp. 1166–1192, doi:10.1214/aop/1055425774, http://dx.doi.org/10.1214/aop/1055425774.

[53] E. Pardoux and A. Y. Veretennikov, *On the Poisson equation and diffusion approximation. III*, Ann. Probab., 33 (2005), pp. 1111–1133, doi:10.1214/009117905000000062, http://dx.doi.org/10.1214/009117905000000062.

[54] G. A. Pavliotis, *Asymptotic analysis of the Green-Kubo formula*, IMA J. Appl. Math., 75 (2010), pp. 951–967, doi:10.1093/imamat/hxq039, http://dx.doi.org/10.1093/imamat/hxq039.

[55] G. A. Pavliotis, *Stochastic processes and applications*, vol. 60 of Texts in Applied Mathematics, Springer, New York, 2014, doi:10.1007/978-1-4939-1323-7, http://dx.doi.org/10.1007/978-1-4939-1323-7. Diffusion processes, the Fokker-Planck and Langevin equations.

[56] G. A. Pavliotis and A. M. Stuart, *Multiscale methods*, vol. 53 of Texts in Applied Mathematics, Springer, New York, 2008. Averaging and homogenization.

[57] G. A. Pavliotis and A. Vogiannou, *Diffusive transport in periodic potentials: Underdamped dynamics*, Fluct. Noise Lett., 8 (2008), pp. L155–173.

[58] M. Reed and B. Simon, *Methods of modern mathematical physics. IV. Analysis of operators*, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.

[59] J. Shen and L.-L. Wang, *Some recent advances on spectral methods for unbounded domains*, Commun. Comput. Phys., 5 (2009), pp. 195–241.

[60] T. Tang, *The Hermite spectral method for Gaussian-type functions*, SIAM J. Sci. Comput., 14 (1993), pp. 594–606, doi:10.1137/0914038, http://dx.doi.org/10.1137/0914038.

[61] U. Vaes, *Numerical solution of multiscale stochastic equations*, master's thesis, Department of Mathematics, Imperial College London, 2014.

[62] E. Vanden-Eijnden, *Numerical techniques for multi-scale dynamical systems with stochastic effects*, Commun. Math. Sci., 1 (2003), pp. 385–391.

[63] C. Villani, *Hypocoercivity*, Mem. Amer. Math. Soc., 202 (2009), pp. iv+141, doi:10.1090/S0065-9266-09-00567-5, http://dx.doi.org/10.1090/S0065-9266-09-00567-5.

[64] H. Yserentant, *Regularity and approximability of electronic wave functions*, vol. 2000 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2010, doi:10.1007/978-3-642-12248-4, http://dx.doi.org/10.1007/978-3-642-12248-4.