

## MEAN FIELD LIMITS FOR INTERACTING DIFFUSIONS WITH COLORED NOISE: PHASE TRANSITIONS AND SPECTRAL NUMERICAL METHODS\*

S. N. GOMES<sup>†</sup>, G. A. PAVLIOTIS<sup>‡</sup>, AND U. VAES<sup>‡</sup>

**Abstract.** In this paper we consider systems of weakly interacting particles driven by colored noise in a bistable potential, and we study the effect of the correlation time of the noise on the bifurcation diagram for the equilibrium states. We accomplish this by solving the corresponding McKean–Vlasov equation using a Hermite spectral method, and we verify our findings using Monte Carlo simulations of the particle system. We consider both Gaussian and non-Gaussian noise processes, and for each model of the noise we also study the behavior of the system in the small correlation time regime using perturbation theory. The spectral method that we develop in this paper can be used for solving linear and nonlinear, local and nonlocal (mean field) Fokker–Planck equations, without requiring that they have a gradient structure.

**Key words.** McKean–Vlasov PDEs, nonlocal Fokker–Planck equations, interacting particles, Desai–Zwanzig model, colored noise, Hermite spectral methods, phase transitions

**AMS subject classifications.** 35Q70, 35Q83, 35Q84, 65N35, 65M70, 82B26

**DOI.** 10.1137/19M1258116

**1. Introduction.** Systems of interacting particles appear in a wide variety of applications, ranging from plasma physics and galactic dynamics [5] to mathematical biology [12, 31], the social sciences [16, 34], active media [4], dynamical density functional theory (DDFT) [18, 17], and machine learning [30, 39, 42]. They can also be used in models for cooperative behavior [8], opinion formation [16], and risk management [15], and also in algorithms for global optimization [38].

In most of the existing works on the topic, the particles are assumed to be subject to thermal additive noise that is modeled as a white noise process, i.e., a mean-zero Gaussian stationary process that is delta-correlated in time. There is extensive literature studying the behavior of these systems; we mention, for example, works on the rigorous passage to the mean field limit [35], the long-time behavior of solutions (see [8, 41] for a case of a ferromagnetic (quartic) potential, and [19] for more general potentials), multiscale analysis [20], and phase transitions [43].

In a more realistic scenario, the system has memory and the hypothesis of Markovianity does not hold [25, 26, 27]. This memory can be modeled by using *colored* noise, i.e., noise with a nonzero correlation time (or, more precisely, a nonsingular autocorrelation function), which is the approach we take in this paper. For simplicity, we will

---

\*Received by the editors April 24, 2019; accepted for publication (in revised form) April 28, 2020; published electronically September 2, 2020.

<https://doi.org/10.1137/19M1258116>

**Funding:** This work was supported by EPSRC through grants EP/P031587/1, EP/L024926/1, EP/L020564/1, and EP/K034154/1. The first author’s work was supported by the Leverhulme Trust through Early Career Fellowship ECF-2018-536. This research was funded in part by JPMorgan Chase & Co. Any views or opinions expressed herein are solely those of the authors listed, and may differ from the views and opinions expressed by JPMorgan Chase & Co. or its affiliates. This material is not a product of the Research Department of J.P. Morgan Securities LLC. This material does not constitute a solicitation or offer in any jurisdiction.

<sup>†</sup>Mathematics Institute, University of Warwick, Coventry, CV4 7AL, UK (susana.gomes@warwick.ac.uk).

<sup>‡</sup>Department of Mathematics, Imperial College London, London, SW7 2AZ, UK (g.pavliotis@imperial.ac.uk, u.vaes13@imperial.ac.uk).

assume that the noise is additive and that it can be represented by a finite-dimensional Markov process, as in the recent study [10] on mean field limits for non-Markovian interacting particles.

In this paper we will study the dynamics of a system of interacting particles of the Desai–Zwanzig type, interacting via a quadratic Curie–Weiss potential. The system of interacting particles is modeled by a system of stochastic differential equations (SDEs):

$$(1.1) \quad \frac{dX_t^i}{dt} = - \left( V'(X_t^i) + \theta \left( X_t^i - \frac{1}{N} \sum_{j=1}^N X_t^j \right) \right) + \sqrt{2\beta^{-1}} \xi_t^i, \quad i = 1, \dots, N,$$

where  $N$  is the number of particles,  $V(\cdot)$  is a confining potential,  $\theta$  is the interaction strength,  $\beta$  is the inverse temperature of the system, and  $\xi_t^i$  are independent, identically distributed (i.i.d.) noise processes.

Before discussing the Desai–Zwanzig model with colored noise, we present a brief overview of known results [8, 41] for the white noise problem. When  $\xi_t^i$  are white noise processes, we can pass to the mean field limit  $N \rightarrow \infty$  in (1.1) and obtain a nonlinear and nonlocal Fokker–Planck equation, known in the literature as a McKean–Vlasov equation, for the one-particle distribution function  $\rho(x, t)$ :

$$(1.2) \quad \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \left( V'(x) \rho + \theta \left( x - \int_{\mathbf{R}} x \rho(x, t) dx \right) \rho + \beta^{-1} \frac{\partial \rho}{\partial x} \right).$$

The McKean–Vlasov equation (1.2) is a gradient flow with respect to the quadratic Wasserstein metric for the free energy functional

$$(1.3) \quad \mathcal{F}[\rho] = \beta^{-1} \int_{\mathbf{R}} \rho(x) \ln \rho(x) dx + \int_{\mathbf{R}} V(x) \rho(x) dx + \frac{\theta}{2} \int_{\mathbf{R}} \int_{\mathbf{R}} F(x-y) \rho(x) \rho(y) dx dy,$$

where  $F(x) := x^2/2$  is the interaction potential. The long-time behavior of solutions depends on the number of local minima of the confining potential  $V$  [43]. It follows directly from (1.2) that any steady-state solution  $\rho_\infty(x)$  solves, together with its first moment, the following system of equations:

$$(1.4a) \quad \frac{\partial}{\partial x} \left( V'(x) \rho_\infty(x) + \theta (x - m) \rho_\infty(x) + \beta^{-1} \frac{\partial \rho_\infty}{\partial x}(x) \right) = 0,$$

$$(1.4b) \quad m = \int_{\mathbf{R}} x \rho_\infty(x) dx.$$

Since (1.4a) is, for  $m$  fixed, the stationary Fokker–Planck equation associated with the overdamped Langevin dynamics in the confining potential

$$(1.5) \quad V_{\text{eff}}(x; m, \theta) = V(x) + \frac{\theta}{2}(x - m)^2,$$

solutions can be expressed explicitly as

$$(1.6) \quad \rho_\infty(x; m, \beta, \theta) := \frac{1}{\mathcal{Z}(m, \beta, \theta)} e^{-\beta V_{\text{eff}}(x; m, \theta)},$$

where  $\mathcal{Z}(m, \beta, \theta)$  is the normalization constant (partition function); see [8, 19, 20] for more details. By substitution in (1.4b), a scalar fixed-point problem is obtained for

$m$ , the self-consistency equation:

$$(1.7) \quad m = \int_{\mathbf{R}} x \rho_{\infty}(x; m, \beta, \theta) dx =: R(m, \beta, \theta).$$

The stability of solutions to (1.4) depends on whether they correspond to a local minimum (stable) or to a local maximum/saddle point (unstable) of the free energy functional. The free energy along the one-parameter family (1.6), with parameter  $m$ , can be calculated explicitly [20],

$$\mathcal{F}[\rho_{\infty}(\cdot; m, \beta, \theta)] = -\beta^{-1} \ln \mathcal{Z}(m, \beta, \theta) - \frac{\theta}{2} (R(m, \beta, \theta) - m)^2,$$

from which we calculate that

$$\frac{\partial}{\partial m} \mathcal{F}[\rho_{\infty}(\cdot; m, \beta, \theta)] = -\beta\theta^2 (R(m, \beta, \theta) - m) \text{Var}(\rho_{\infty}(\cdot; m, \beta, \theta)),$$

where, for a probability density  $\psi$ ,

$$\text{Var}(\psi) := \int_{\mathbf{R}} \left( x - \int_{\mathbf{R}} \psi(x) dx \right)^2 \psi(x) dx.$$

Though incomplete, this informal argument suggests that the stability of a steady-state solution can also be inferred from the slope of  $R(m, \beta, \theta) - m$  at the corresponding value of  $m$ : If this slope is positive, the equilibrium is unstable, and conversely. The self-consistency map and the free energy of  $\rho_{\infty}(x; m, \beta, \theta)$ , for a range of values of  $m$ , are illustrated in Figure 1.1 for the bistable potential  $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$ . It is well-known that, when  $V(\cdot)$  is an even potential, (1.2) possesses a unique, mean-zero steady-state solution for sufficiently large temperatures (i.e., small  $\beta$ ). As the temperature decreases, this solution loses its stability and two new solutions of the self-consistency equation emerge, corresponding to a pitchfork bifurcation; see [8, 20] for details.

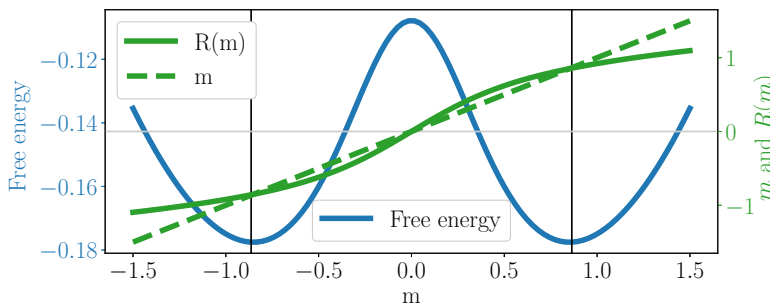


FIG. 1.1. Free energy (1.3) of the one-parameter family (1.6) of probability densities that solve (1.4a) for some value of  $m$  (in blue), and associated first moment  $R(m)$  (in green) for fixed  $\theta = 1$  and  $\beta = 5$ . Along the one-parameter family,  $m = 0$  is a local maximum of the free energy. Therefore, it corresponds to an unstable steady state of the McKean–Vlasov equation. (Figure in color online.)

As mentioned above, in this paper we focus on the case where the noise processes  $\xi_t^i$  in (1.1) have a nonzero correlation time. In particular, we assume that each noise process can be represented using a (possibly multidimensional) SDE, in which case

(1.1) leads to a Markovian system of SDEs in an extended phase space. The colored noise will be modeled by either an Ornstein–Uhlenbeck (OU) process, harmonic noise [36, Example 8.2], or a non-Gaussian reversible diffusion process.

Though more realistic, the use of colored noise presents us with some difficulties. First, the introduction of an extra SDE for the noise breaks the gradient structure of the problem; while we can still pass formally to the limit  $N \rightarrow \infty$  in (1.1) and obtain a McKean–Vlasov equation for the associated one-particle distribution function, it is no longer possible to write a free energy functional, such as (1.3), that is dissipated by this equation. Second, the McKean–Vlasov equation is now posed in an extended phase space, which increases the computational cost of its numerical solution via PDE methods. And third, it is no longer possible to obtain an explicit expression for the one-parameter family of (possible) stationary solutions to the mean field equation, as was possible in (1.6), which renders the calculation of steady states considerably more difficult.

When the correlation time of the noise is small, the latter difficulty can be somewhat circumvented by constructing an approximate one-parameter family of solutions through appropriate asymptotic expansions in terms of the correlation time, from which steady-state solutions of the McKean–Vlasov dynamics can be extracted by solving a self-consistency equation similar to (1.7); see (2.10). Outside of the small correlation time regime, however, finding the steady-states of the McKean–Vlasov equation requires a numerical method for PDEs in all but the simplest cases.

In this work, we propose a novel Hermite spectral method for the time-dependent and steady-state equations, applicable to the cases of both white and colored noise. Discretized in a basis of Hermite functions, the McKean–Vlasov equation becomes a system of ordinary differential equations with a quadratic nonlinearity originating from the interaction term. In contrast with other discretization methods for PDEs, the use of (possibly rescaled) Hermite functions for the problem under consideration leads to an efficient numerical method. This is because, first, Hermite functions have very good approximation properties in  $L^2$  and, second, all the differential operators appearing in the McKean–Vlasov equation lead to sparse matrices in Hermite space, with a small bandwidth related to the polynomial degree of  $V$  (provided that a suitable ordering of the multi-indices is employed). To solve the finite-dimensional system of equations obtained after discretization of the time-dependent equation, we employ either the Runge–Kutta 45 method (RK45) or a linear, semi-implicit time-stepping scheme.

We also verify that our results agree with known analytical solutions in simple settings, and with explicit asymptotic expansions in the small correlation time regime. We then use our spectral method, together with asymptotic expansions and Monte Carlo (MC) simulations of the particle system, to construct the bifurcation diagram of the first moment of the steady-state solutions as a function of the inverse temperature.

For the reader’s convenience, we summarize here the main results of this paper:

1. The systematic study of the effect of colored noise, both Gaussian and non-Gaussian, on the long-time behavior of the McKean–Vlasov mean field equation, including the effect of colored noise on the structure and properties of phase transitions.
2. The development and analysis of a spectral numerical method for the solution of linear or nonlinear, local or nonlocal Fokker–Planck-type equations. In particular, our method does not depend on an underlying gradient structure for the PDE.

The rest of this paper is organized as follows. In section 2, we present the models

for the colored noise and we derive formally the mean field McKean–Vlasov equation associated with the interacting particle system. In section 3, we present the numerical methods used to (a) solve the time-dependent and steady-state Fokker–Planck (or McKean–Vlasov) equations and (b) solve the finite-dimensional system of interacting diffusions (1.1). In section 4, we describe our methodology for constructing the bifurcation diagrams and we present the associated results. Section 5 is reserved for conclusions and perspectives for future work.

**2. The model.** We consider the following system of weakly interacting diffusions:

$$(2.1) \quad dX_t^i = - \left( V'(X_t^i) + \theta \left( X_t^i - \frac{1}{N} \sum_{j=1}^N X_t^j \right) \right) dt + \sqrt{2\beta^{-1}} \eta_t^i dt, \quad 1 \leq i \leq N,$$

where the noise processes  $\eta_t^i$  are independent, mean-zero, second-order stationary processes with almost surely continuous paths and autocorrelation function  $K(t)$ . In the rest of this paper, we will assume that the interaction strength  $\theta$  is fixed and equal to 1 and we will use the inverse temperature  $\beta^{-1}$  as the bifurcation parameter. We will consider two classes of models for the noise: Gaussian stationary noise processes with an exponential correlation function, and non-Gaussian noise processes that we construct by using the overdamped Langevin dynamics in a nonquadratic potential.

*Gaussian noise.* Stationary Gaussian processes in  $\mathbf{R}^n$  with continuous paths and an exponential autocorrelation function are solutions to an SDE of Ornstein–Uhlenbeck type:

$$(2.2) \quad d\mathbf{Y}_t^i = A \mathbf{Y}_t^i dt + \sqrt{2} D d\mathbf{W}_t^i, \quad i = 1, \dots, N,$$

where  $A, D$  are  $n \times n$  matrices satisfying Kalman’s rank condition [29, Chapter 9], and  $\mathbf{W}_t^i, 1 \leq i \leq n$ , are independent white noise processes in  $\mathbf{R}^n$ . We assume here that the noise is obtained by projection as  $\eta_t^i = \langle \mathbf{Y}_t^i, \mathbf{y}_\eta \rangle$ , where  $\langle \cdot, \cdot \rangle$  denotes the Euclidean inner product for some vector  $\mathbf{y}_\eta \in \mathbf{R}^n$ . Throughout this paper we will consider two particular examples, namely the scalar OU process and the harmonic noise [36, Chapter 8].

**(OU)** Scalar Ornstein–Uhlenbeck process:

$$d\eta_t^i = -\eta_t^i dt + \sqrt{2} dW_t^i.$$

The associated autocorrelation function is

$$K_{OU}(t) = e^{-|t|}.$$

**(H)** Harmonic noise:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & -\gamma \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{pmatrix}, \quad \mathbf{y}_\eta = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In this case the noise is the solution to the Langevin equation, with the first and second components of  $\mathbf{Y}$  corresponding to the position and velocity, respectively. Throughout this paper we will assume  $\gamma = 1$  for simplicity. The associated autocorrelation function of  $\eta^i$  is given by

$$K_H(t) = e^{-\frac{|t|}{2}} \left( \cos \left( \frac{\sqrt{3}}{2} t \right) + \frac{\sqrt{3}}{3} \sin \left( \frac{\sqrt{3}}{2} t \right) \right).$$

*Non-Gaussian noise.* In this case, instead of (2.2) we consider

$$d\eta_t^i = -V_\eta'(\eta_t^i) dt + \sqrt{2} dW_t^i,$$

where now  $V_\eta$  is a smooth nonquadratic confining potential satisfying the mean-zero condition:

$$(2.3) \quad \int_{\mathbf{R}} \eta e^{-V_\eta(\eta)} d\eta = 0.$$

We consider the following choices for  $V_\eta$ :

(B) The bistable potential  $V_\eta(\eta) = \eta^4/4 - \eta^2/2$ .

(NS) The shifted tilted bistable potential

$$(2.4) \quad V_\eta(\eta) = \frac{(\eta - \alpha)^4}{4} - \frac{(\eta - \alpha)^2}{2} + (\eta - \alpha),$$

with the constant  $\alpha \approx 0.885$  such that (2.3) is satisfied.

**2.1. Mean field limit.** For weakly interacting diffusions, the derivation of the mean field McKean–Vlasov PDE is a standard, well-known result [8, 41, 35]. When  $\xi_t^i$  in (1.1) are colored noise processes, it is also possible to pass to the mean field limit  $N \rightarrow \infty$  in (1.1) and to obtain a McKean–Vlasov equation for the one-particle distribution function  $\rho(x, \mathbf{y}, t)$ :

$$(2.5a) \quad \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \left( V' \rho + \theta (x - m(t)) \rho - \sqrt{2\beta^{-1}} \langle \mathbf{y}_i, \mathbf{y}_\eta \rangle \rho \right) + \mathcal{L}_{\mathbf{y}}^* \rho,$$

with the dynamic constraint

$$(2.5b) \quad m(t) = \int_{\mathbf{R}} \int_{\mathbf{R}^n} x \rho(x, \mathbf{y}, t) d\mathbf{y} dx.$$

Here  $\mathbf{y}$  are the noise variables, denoted by  $(\eta, \lambda)$  in the case of harmonic noise and just  $\eta$  otherwise, and

$$\mathcal{L}_{\mathbf{y}}^* \rho = \begin{cases} \partial_\eta (\eta \rho + \partial_\eta \rho) & \text{for scalar OU noise,} \\ \partial_\lambda (\lambda \rho + \partial_\lambda \rho) + (\eta \partial_\lambda \rho - \lambda \partial_\eta \rho) & \text{for harmonic noise,} \\ \partial_\eta (V_\eta' \rho + \partial_\eta \rho) & \text{for non-Gaussian noise.} \end{cases}$$

A formal derivation of the mean field limit is presented in [44], and this derivation can be justified rigorously using the results in [11, 33].

The main goal of this paper is the study of the effect of colored noise on the structure of the bifurcation diagram for the McKean–Vlasov equation with colored noise, (2.5a) and (2.5b). In other words, we want to gain insight into the number of solutions to the following stationary PDE and associated constraint (self-consistency equation):

$$(2.6a) \quad \frac{\partial}{\partial x} \left( V'(x) \rho + \theta (x - m) \rho - \sqrt{2\beta^{-1}} \langle \mathbf{y}_i, \mathbf{y}_\eta \rangle \rho \right) + \mathcal{L}_{\mathbf{y}}^* \rho = 0,$$

$$(2.6b) \quad m = \int_{\mathbf{R}} \int_{\mathbf{R}^n} x \rho(x, \mathbf{y}) d\mathbf{y} dx.$$

Although there still exists, for fixed  $\beta$  and fixed  $\theta$ , a one-parameter family of solutions to (2.6a) (with parameter  $m$ ), which we will denote by  $\{\rho_\infty(x, \mathbf{y}; m, \beta, \theta)\}_{m \in \mathbf{R}}$ , no closed form is available for these solutions. This is because the detailed balance condition no longer holds in the presence of colored noise, i.e., the probability flux at equilibrium does not vanish. Here, by probability flux, we mean the argument of the divergence in the Fokker–Planck operator; see [36, section 4.6].

**2.2. The white noise limit.** To study the limit of small correlation time, it will be convenient to rescale the noise as

$$\eta_t^i \rightarrow \zeta \eta_{t/\varepsilon^2}^i / \varepsilon,$$

where  $\varepsilon$  is a time scaling parameter, and  $\zeta$  is a model-dependent parameter ensuring that the autocorrelation function of the rescaled noise, given by  $\zeta^2 K(t/\varepsilon^2)/\varepsilon^2$ , satisfies

$$\int_0^\infty \zeta^2 K(t/\varepsilon^2)/\varepsilon^2 dt = \int_0^\infty \zeta^2 K(t) dt = \frac{1}{2}.$$

Then the autocorrelation of the noise converges to a Dirac delta when  $\varepsilon \rightarrow 0$ , and it can be shown that, in this limit, the solution of (2.1) converges to that of

$$dX_t^i = \left( -V'(X_t^i) - \theta \left( X_t^i - \frac{1}{N} \sum_{j=0}^N X_t^j \right) \right) dt + \sqrt{2\beta^{-1}} dW_t^i, \quad i = 1, \dots, N,$$

where  $W^i, i = 1, \dots, N$ , are independent Wiener processes; see [6] and [37, Chapter 11]. While not strictly necessary, including the parameter  $\zeta$  is convenient to obtain simpler formulas. The value of  $\zeta$  for each of the noise models considered in this paper is presented in Table 2.1. For the models **B** and **NS**,  $\zeta$  was calculated numerically and rounded to three significant figures in this table.

TABLE 2.1  
Value of  $\zeta$ .

Model	<b>OU</b>	<b>H</b>	<b>B</b>	<b>NS</b>
$\zeta$	$1/\sqrt{2}$	$1/\sqrt{2}$	0.624	0.944

In view of the convergence of the solution of the finite-dimensional particle system when  $\varepsilon \rightarrow 0$ , we expect that the  $x$ -marginals of the steady-state solutions to the McKean–Vlasov equation with colored noise, obtained by solving (2.6a) and (2.6b), should also converge to their white-noise counterparts as  $\varepsilon \rightarrow 0$ . It turns out that this is the case and, using asymptotic techniques from [25], it is possible to approximate the solutions  $\rho_\infty(x, \mathbf{y}; m, \beta, \theta)$  to (2.6a) by a power series expansion in  $\varepsilon$ ; using a superscript to emphasize the dependence on  $\varepsilon$ ,

$$(2.7) \quad \rho_\infty^\varepsilon(x, \mathbf{y}; m, \beta, \theta) = p_0(x, \mathbf{y}; m, \beta, \theta) + \varepsilon p_1(x, \mathbf{y}; m, \beta, \theta) + \varepsilon^2 p_2(x, \mathbf{y}; m, \beta, \theta) + \dots$$

From (2.7), we obtain a power series expansion for the  $x$ -marginal by integrating out the noise variable:

$$(2.8) \quad \begin{aligned} \rho_\infty^\varepsilon(x; m, \beta, \theta) &= \int_{\mathbf{R}^n} \rho_\infty^\varepsilon(x, \mathbf{y}; m, \beta, \theta) d\mathbf{y} \\ &=: \rho_\infty(x; m, \beta, \theta) + \varepsilon p_1(x; m, \beta, \theta) + \varepsilon^2 p_2(x; m, \beta, \theta) + \dots \end{aligned}$$

The methodology to obtain expressions for the terms works by substituting (2.7) in (2.6a) and grouping the terms in powers of  $\varepsilon$  in the resulting equation. This leads to a sequence of equations that can be studied using standard techniques. Details of the analysis leading to an explicit expression of the first nonzero correction in (2.8) can be found in [25, section 8] for the particular case of the **OU** noise, and in [44] for the other noise models we consider.

The order of the first nonzero correction in this expansion depends on the model: It is equal to 1 for model **NS**, to 2 for models **OU** and **B**, and to 4 for model **H**. In all cases, the first nontrivial term in the series expansion (2.8) can be calculated explicitly (possibly up to constant coefficients that have to be calculated numerically). For scalar Ornstein–Uhlenbeck noise, for example, we have, omitting the dependence of  $V_{\text{eff}}$  (the effective potential defined in (1.5)) on  $m$  and  $\theta$  for notational convenience,

$$(2.9) \quad \rho_{\infty}^{\varepsilon}(x; m, \beta, \theta) = \rho_{\infty}(x; m, \beta, \theta) \left[ 1 + \varepsilon^2 \left( C_{OU} - \frac{\beta}{2} (V'_{\text{eff}}(x))^2 + V''_{\text{eff}}(x) \right) \right] + \mathcal{O}(\varepsilon^4).$$

Here  $C_{OU}$  is a constant such that the correction integrates to 0. Similar expressions can be obtained for the other models; see [44].

Taking into account only the first nontrivial correction, the order of which we denote by  $\delta$ , the steady-state solutions to the McKean–Vlasov equation with colored noise can be approximated by solving the approximate self-consistency equation

$$(2.10) \quad \begin{aligned} m &= R_0(m, \beta, \theta) + \varepsilon^{\delta} R_{\delta}(m, \beta, \theta) \\ &:= \int_{\mathbf{R}} x \rho_{\infty}(x; m, \beta, \theta) dx + \varepsilon^{\delta} \int_{\mathbf{R}} x p_{\delta}(x; m, \beta, \theta) dx \\ &\approx R(m, \beta) := \int_{\mathbf{R}} x \rho_{\infty}^{\varepsilon}(x; m, \beta, \theta) dx. \end{aligned}$$

We show in Figure 2.1 that the equation  $R_0(m, \beta, \theta) + \varepsilon^2 R_2(m, \beta, \theta) = m$ , for fixed  $\beta = 10$ ,  $\theta = 1$  and  $\varepsilon = 0.1$ , admits three solutions in the case of **OU** noise, similarly to the case of white noise. This figure was generated using the asymptotic expansion (2.9).

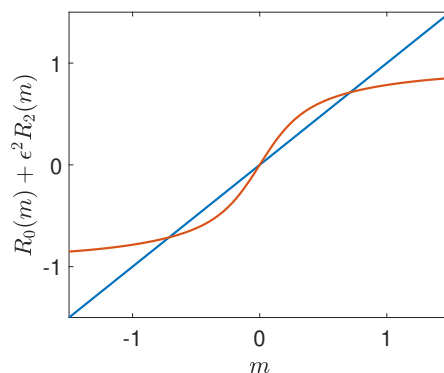


FIG. 2.1. Truncated asymptotic expansion of the self-consistency map,  $R_0 + \varepsilon^2 R_2$  as a function of  $m$  (red line) compared to  $y = m$  (blue line) for the scalar Ornstein–Uhlenbeck noise, with  $\beta = 10$ ,  $\theta = 1$ ,  $\varepsilon = 0.1$ . (Figure in color online.)



**3. The numerical method.** In this section, we describe the spectral numerical method that we will use in order to solve the time-dependent McKean–Vlasov equation, (2.5a) and (2.5b), as well as the steady-state equation, (2.6a) and (2.6b). Before looking at colored noise, we consider the case of white noise for which our method can be tested against the results in [20], which were obtained using the finite volume scheme developed in [7].

**3.1. Linear Fokker–Planck equation with white noise.** We start by presenting the methodology used in the absence of an interaction term, in which case (2.5a) reduces to a linear Fokker–Planck equation:

$$(3.1) \quad \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \left( V' \rho + \beta^{-1} \frac{\partial \rho}{\partial x} \right) =: \mathcal{L}_x^* \rho, \quad \rho(x, t = 0) = \rho_0(x).$$

We assume that  $V(\cdot)$  is a smooth confining potential and, consequently, the unique invariant distribution is given by  $\rho_s = \frac{1}{\mathcal{Z}} e^{-\beta V}$ , where  $\mathcal{Z}$  is the normalization constant [36, Proposition 4.2]. The Fokker–Planck operator in (3.1) is unitarily equivalent to a Schrödinger operator; see [1] and [36, section 4.9]. Defining  $u = \rho/\sqrt{\rho_s}$ , the function  $u$  satisfies

$$(3.2) \quad \begin{aligned} \frac{\partial u}{\partial t} &= \sqrt{\rho_s^{-1}} \mathcal{L}_x^* \left( \sqrt{\rho_s} u \right) \\ &= \beta^{-1} \frac{\partial^2 u}{\partial x^2} + \left( \frac{1}{2} V''(x) - \frac{\beta}{4} |V'(x)|^2 \right) u =: \mathcal{H}_x u, \end{aligned}$$

with the initial condition  $u(x, t = 0) = \rho_0/\sqrt{\rho_s} =: u_0$ . Several works made use of Hermite spectral methods to study equations of this type, e.g., [1, 13, 14]. The Schrödinger operator on the right-hand side of (3.1) is self-adjoint in  $L^2(\mathbf{R})$  and it has nonpositive eigenvalues. Under appropriate growth assumptions on the potential  $V(x)$  as  $x \rightarrow \infty$ , it can be shown that its eigenfunctions decrease more rapidly than any exponential function in the  $L^2(\mathbf{R})$  sense, in that they satisfy  $e^{\mu|x|} \varphi(x) \in L^2(\mathbf{R})$  for all  $\mu \in \mathbf{R}$ ; see [14] and also [2] for a detailed study. Under appropriate decay assumptions at infinity on the initial condition, we expect the solution to (3.1) to also decrease rapidly as  $|x| \rightarrow \infty$ .

We denote by  $\mathbf{P}(d)$  the space of polynomials of degree less than or equal to  $d$ , and by  $\langle \cdot, \cdot \rangle$  the usual  $L^2(\mathbf{R})$  inner product. For a quadratic potential  $V_q = \frac{1}{2} \left(\frac{x}{\sigma}\right)^2$ , with  $\sigma$  a scaling parameter, the Galerkin method we employ consists of finding  $u_d(t) \in e^{-V_q/2} \mathbf{P}(d)$  such that

$$(3.3a) \quad \left\langle \frac{\partial u_d}{\partial t}, w_d \right\rangle = \langle \mathcal{H}_x u_d, w_d \rangle_{\hat{d}} \quad \forall w_d \in e^{-V_q/2} \mathbf{P}(d), \quad \forall t > 0,$$

$$(3.3b) \quad \langle u_d(0), w_d \rangle = \langle u_0, w_d \rangle_{\hat{d}} \quad \forall w_d \in e^{-V_q/2} \mathbf{P}(d).$$

Here the subscript  $\hat{d} \geq d$  on the right-hand side of (3.3a) and (3.3b) indicates that the inner product is performed using a numerical quadrature with  $\hat{d} + 1$  points. With appropriately rescaled Gauss–Hermite points, inner products calculated using the quadrature are *exact* for functions in  $e^{-V_q/2} \mathbf{P}(\hat{d})$ ,

$$\langle v_d, w_d \rangle_{\hat{d}} = \langle v_d, w_d \rangle \quad \forall v_d, w_d \in e^{-V_q/2} \mathbf{P}(\hat{d}),$$

which is why we did not append the subscript  $\hat{d}$  to the inner products on the left-hand side of (3.3a) and (3.3b). When  $V$  is a polynomial, it is possible to show

using the recursion relations (A.1) and (A.2) in [21, Appendix A] that the inner product  $\langle \mathcal{H}_x u_d, w_d \rangle_{\hat{d}}$  on the right-hand side of (3.3a) is *exactly*  $\langle \mathcal{H}_x u_d, w_d \rangle$  when  $\hat{d} \geq d + \deg(|V'|^2)$ . This is the approach we take in all the numerical experiments presented in this paper. We will, therefore, omit the subscript  $\hat{d}$  in (3.3a) from now on.

The natural basis of  $\mathbf{P}(d)$  (from which a basis of  $e^{-V_q/2} \mathbf{P}(d)$  follows) to obtain a finite-dimensional system of differential equations from the variational formulation (3.3a) is composed of rescaled Hermite polynomials  $H_i^\sigma(x) := H_i(x/\sigma)$ ,  $0 \leq i \leq d$ , where  $H_i(x)$  are the Hermite polynomials orthonormal for the Gaussian weight  $\mathcal{N}(0, 1)$ ; the corresponding basis functions of  $e^{-V_q/2} \mathbf{P}(d)$  are then rescaled Hermite functions. The fundamental results on Hermite polynomials, Hermite functions, and the related approximation results that are used in this paper are summarized in [21, Appendix A].

It is possible to prove the convergence of the method presented above in the limit as  $d \rightarrow \infty$  given appropriate additional assumptions on the confining potential  $V(\cdot)$ . For simplicity we will make the following assumption, which is satisfied for the bistable potential that we consider in this work, but we note that less restrictive conditions would be sufficient.

*Assumption 3.1.* The confining potential  $V(\cdot)$  is a polynomial of (even) degree greater than or equal to 2. Consequently, it satisfies

$$C_1(1 + |x|^2) \leq C_2 + W := C_2 + \left( \frac{\beta}{4} |V'|^2 - \frac{1}{2} V'' \right) \leq C_3(1 + |x|^{2k})$$

for constants  $C_1, C_2, C_3 > 0$  and a natural number  $k \geq 1$ .

We will denote by  $H^m(\mathbf{R}; \mathcal{H}_x)$  the Hilbert space obtained by completion of  $C_c^\infty(\mathbf{R})$ , the space of smooth compactly supported functions, with the inner product

$$\langle u, v \rangle_{m, \mathcal{H}_x} := \langle (-\mathcal{H}_x + 1)^m u, v \rangle.$$

The norm associated with this Sobolev-like space will be denoted by  $\|\cdot\|_{m, \mathcal{H}_x}$ .

**THEOREM 3.1.** *Suppose that Assumption 3.1 holds and that the initial condition  $u_0$  is smooth and belongs to  $H^m(\mathbf{R}; \mathcal{H}_x)$  for some natural number  $m \geq 2k$ , where  $k$  is as in Assumption 3.1. Then for any  $d \geq m - 1$ , any final time  $T$ , and for all  $\alpha > 0$ , it holds that*

$$\sup_{t \in [0, T]} \|u(t) - u_d(t)\|^2 \leq C_\alpha e^{\alpha T} \frac{(d - m + 1)!}{(d - 2k + 1)!} \|u_0\|_{m, \mathcal{H}_x}$$

for a constant  $C_\alpha$  not depending on  $d$ ,  $u_0$ , or  $T$ , and where  $\|\cdot\|$  denotes the  $L^2(\mathbf{R})$  norm.

*Proof.* See Appendix A for the proof.  $\square$

*Remark 3.1.* Theorem 3.1 is not optimal. On one hand, it overestimates the error for large times: Both the numerical and exact solutions converge to stationary solutions as  $t \rightarrow \infty$ , so we expect the error  $\|u(t) - u_d(t)\|^2$  to tend to a finite limit when  $t \rightarrow \infty$ . Although the error between the stationary solutions can be bounded similarly to the transient error, (see Remark A.1), we have not obtained a result that combines both errors; we plan to return to this interesting question in future work. On the other hand, the bound on the transient error of Theorem 3.1 is probably

not sharp. Indeed, when the initial condition  $u_0$  is smooth and, together with all its derivatives, decreases exponentially as  $x \rightarrow \infty$ , Theorem 3.1 implies only that the error decreases faster than any negative power of  $d$ . In most practical examples, however, we observed numerically that the convergence is, in fact, exponential.

*Remark 3.2.* The condition that  $u_0 \in H^m(\mathbf{R}; m, \mathcal{H}_x)$  is quite restrictive. It requires, in particular, that  $u_0 \in L^2(\mathbf{R})$ , which is equivalent to requiring that  $\rho_0 \in L^2(\mathbf{R}; \rho_s^{-1})$ , because  $u_0 = \rho_0/\sqrt{\rho_s}$  by definition. Though natural from an  $L^2$ -theory perspective (see [36, sect. 4.5] and [32]), this condition excludes a large class of initial conditions. If  $V(x)$  behaves as  $|x|^4$  as  $|x| \rightarrow \infty$ , then it excludes Gaussian initial conditions, for example.

**3.2. McKean–Vlasov equation with white noise.** In the presence of an interaction term, the Fokker–Planck equation becomes nonlinear:

$$(3.4) \quad \begin{aligned} \frac{\partial \rho}{\partial t} &= \frac{\partial}{\partial x} \left( V' \rho + \theta(x - m(t)) \rho + \beta^{-1} \frac{\partial \rho}{\partial x} \right) \\ &=: (\mathcal{L}_x^m)^* \rho, \quad m(t) = \int_{\mathbf{R}} x \rho \, dx. \end{aligned}$$

For this equation the weighted  $L^2(\mathbf{R}; e^V)$  energy estimate of the linear case (A.7) does not hold, and there is, therefore, no longer a natural space for the Galerkin approximation. Because of this, and since we would like to employ the spectral numerical method with Gaussian initial conditions, which is not possible with a variational formulation of the type (3.3) in view of Remark 3.2, we will use Hermite functions to approximate the solution to (3.2) directly, i.e., we will look for an approximate solution in the space  $e^{-V_q/2} \mathbf{P}(d)$ . The variational formulation corresponding to the Galerkin approximation is then to find  $\rho \in e^{-V_q/2} \mathbf{P}(d)$  such that

$$(3.5a) \quad \left\langle \frac{\partial \rho_d}{\partial t}, w_d \right\rangle = \langle (\mathcal{L}_x^{m_d})^* \rho_d, w_d \rangle \quad \forall w_d \in e^{-V_q/2} \mathbf{P}(d),$$

$$(3.5b) \quad m_d = \frac{\langle x, \rho_d \rangle_{\hat{d}}}{\langle 1, \rho_d \rangle_{\hat{d}}} \approx \frac{\int_{\mathbf{R}} x \rho_d \, dx}{\int_{\mathbf{R}} \rho_d \, dx},$$

$$(3.5c) \quad \langle \rho_d(0), w_d \rangle = \langle \rho_0, w_d \rangle_{\hat{d}} \quad \forall w_d \in e^{-V_q/2} \mathbf{P}(d).$$

Dividing by  $\langle 1, \rho_d \rangle_{\hat{d}}$  in (3.5b) is useful to account for changes in the total mass of  $\rho_d$ , which can compromise the accuracy of the method when  $d$  is low, but doing so becomes unnecessary for large enough  $d$ . In contrast with the operator  $\mathcal{H}_x$  in (3.3a), the operator  $(\mathcal{L}_x^{m_d})^*$  is not self-adjoint in  $L^2(\mathbf{R})$ . Therefore, the associated stiffness matrix is not symmetric. In addition, the quadratic form  $\langle (\mathcal{L}_x^{m_d})^* \cdot, \cdot \rangle$  is not necessarily negative for the usual  $L^2(\mathbf{R})$  inner product, and indeed, we observe numerically that the eigenvalue with smallest real part of the discrete operator is often negative, although small when  $d$  is large enough.

For the integration in time, we used either the RK45 method (using the `solve_ivp` method from the *SciPy integrate* module), or a linear semi-implicit method obtained by treating  $m_d$  explicitly and the other terms implicitly at each time step. The former is most useful when an accurate time-dependent solution is required, while the latter enables the use of larger time steps and is, therefore, more convenient when only the steady-state solution is sought, as will be the case for the construction of bifurcation diagrams. Denoting the time step by  $\Delta t$  and the Galerkin approximation of  $\rho_d(n \Delta t)$

by  $\rho_d^n$ , the semi-implicit method is based on obtaining  $\rho_d^{n+1}$  by solving

$$(3.6a) \quad \langle \rho_d^{n+1} - \rho_d^n, w_d \rangle = \Delta t \left\langle (\mathcal{L}_x^{m_d})^* \rho_d^{n+1}, w_d \right\rangle \quad \forall w_d \in e^{-V_q/2} \mathbf{P}(d),$$

$$(3.6b) \quad m_d^{n+1} = \frac{\langle x, \rho_d^{n+1} \rangle_{\hat{d}}}{\langle 1, \rho_d^{n+1} \rangle_{\hat{d}}}.$$

*Remark 3.3* (computational considerations). Discretizing the operators appearing in the Galerkin approximations (3.3a) and (3.5) requires the calculation of multiple matrices corresponding to operators of the type  $\Pi_d (f \partial_x) \Pi_d$ , where  $f$  is a polynomial and  $\Pi_d$  is the  $L^2(\mathbf{R}; e^{-V_q})$  projection operator onto  $\mathbf{P}(d)$ . These calculations can be carried out by noticing that

$$\Pi_d \left( f \frac{d^m}{dx^m} \right) \Pi_d = (\Pi_d f \Pi_d) \left( \Pi_d \frac{d^m}{dx^m} \Pi_d \right).$$

The matrix representation of the first operator on the right-hand side, in a basis of Hermite polynomials, can be obtained from the Hermite transform of  $f$ . The matrix representation of the second operator, on the other hand, is a matrix with zero entries everywhere except on the  $m$ th superdiagonal, in view of the recursion relation (A.1) in [21].

**3.3. Linear Fokker–Planck equation with colored noise.** In this section, we turn our attention to the case of Gaussian or non-Gaussian colored noise given in terms of overdamped Langevin dynamics. The case of harmonic noise can be treated in a similar fashion, and for conciseness we do not present the associated Galerkin formulation explicitly here. We start by considering the linear (without the interaction term) Fokker–Planck equation with colored noise:

$$(3.7) \quad \begin{aligned} \frac{\partial \rho}{\partial t} &= \frac{\partial}{\partial x} \left( \frac{\partial V}{\partial x} \rho - \frac{\zeta}{\varepsilon} \sqrt{2\beta^{-1}\eta} \rho \right) \\ &+ \frac{1}{\varepsilon^2} \frac{\partial}{\partial \eta} \left( V'_\eta \rho + \frac{\partial \rho}{\partial \eta} \right) =: \mathcal{L}_\varepsilon^* \rho. \end{aligned}$$

We recall that  $\varepsilon^2$  controls the correlation time of the colored noise and  $\zeta$  is a parameter such that the white noise limit is recovered (with inverse temperature  $\beta$ ) when  $\varepsilon \rightarrow 0$ . We include  $\varepsilon$  in (3.3a) and (3.3b) because, although we do not consider the white noise limit in this section, large values of  $\varepsilon$  are, in general, more difficult to tackle numerically, and it will be, therefore, convenient to use smaller correlation times in the numerical experiments below. The problem is now two-dimensional and the operator on the right-hand side of (3.3a) and (3.3b) is no longer elliptic. In contrast with the white noise case, there does not exist an explicit formula for the steady-state solution for (3.3a) and (3.3b).

The procedure for obtaining a Galerkin formulation is the same as in (3.1), except that we now use tensorized Hermite polynomials/functions. To retain some generality, we will consider that the Galerkin approximation space is of the form  $S_d = e^{-U(x,\eta)/2} e^{-V_q(x,\eta)/2} \mathbf{P}(\mathcal{I}_d)$  for some function  $U : \mathbf{R}^2 \mapsto \mathbf{R}$ , a nondegenerate quadratic potential  $V_q$  to be determined, and where  $\mathbf{P}(\mathcal{I}_d) := \text{span} \{ x^{\alpha_x} \eta^{\alpha_\eta} : (\alpha_x, \alpha_\eta) \in \mathcal{I}_d \}$  for some index set  $\mathcal{I}_d \subset \mathbf{N}^2$  that grows with  $d \in \mathbf{N}$ . Compared to the one-dimensional case, there are now two scaling parameters,  $V_q := x^2/2\sigma_x^2 + \eta^2/2\sigma_\eta^2$ . The

Galerkin approximation we propose consists of finding  $\rho_d \in S_d$  such that

$$(3.8) \quad \left\langle \frac{\partial \rho_d}{\partial t}, w_d \right\rangle_{e^U} = \langle \mathcal{L}_\varepsilon^* \rho_d, w_d \rangle_{e^U} \quad \forall w_d \in S_d, \quad \forall t > 0,$$

with appropriate initial conditions. The choice of the weight  $e^U$  in the inner products of (3.8) is motivated by the fact that differential operators admit sparse representations in the Hermite-type basis naturally associated with  $S_d$ , and we note that  $e^{-U(x,\eta)/2} e^{-V_q(x,\eta)/2} \mathbf{P}(\mathbf{N}^2)$ , where  $\mathbf{P}(\mathbf{N}^2)$  is the space of polynomials in two dimensions, is dense in  $L^2(\mathbf{R}^2; e^U)$ . In practice, we obtain  $\rho_d$  as  $e^{-U(x,\eta)/2} e^{-V_q(x,\eta)/2} v_d$ , where  $v_d$  is obtained by solving

$$(3.9) \quad \left\langle \frac{\partial v_d}{\partial t}, w_d \right\rangle_{e^{-V_q}} = \langle \mathcal{H}_\varepsilon v_d, w_d \rangle_{e^{-V_q}} \quad \forall w_d \in \mathbf{P}(\mathcal{I}_d), \quad \forall t > 0,$$

where, for a test function  $\varphi$ ,  $\mathcal{H}_\varepsilon \varphi := (e^{U/2} e^{V_q/2}) \mathcal{L}_\varepsilon^* (e^{-U/2} e^{-V_q/2} \varphi)$ , and the basis functions used for (3.9) are Hermite polynomials orthonormal with respect to the Gaussian weight  $e^{-V_q}$ . Regarding the index set, several choices are possible, with the simplest ones being the triangle  $\{\alpha \in \mathbf{N}^2 : |\alpha|_1 \leq d\}$  and the square  $\{\alpha \in \mathbf{N}^2 : |\alpha|_\infty \leq d\}$ ; see Figures B.1a and B.2a. It was demonstrated in [44] that, in order to study the limit  $\varepsilon \rightarrow 0$ , a rectangle-shaped index set is usually the only suitable choice. When studying the behavior as  $d$  increases, however, we observed spectral convergence irrespectively of the index set utilized.

Clearly, it is necessary that  $\rho \in L^2(\mathbf{R}^2; e^U)$  for the Galerkin discretization (3.8) to produce good results. Since the  $1/\varepsilon^2$  part of the operator on the right-hand side of (3.3a) and (3.3b),  $\mathcal{L}_0^* = \partial_\eta (V_\eta'(\eta) \cdot + \partial_\eta \cdot)$ , is self-adjoint in  $L^2(\mathbf{R}; e^{-U_x(x)/2 - V_\eta(\eta)/2})$  for any choice of  $U_x$ , it is natural to choose  $e^{-U(x,\eta)/2} = e^{-U_x(x)/2 - V_\eta(\eta)/2}$  for some one-dimensional potential  $U_x$ . This guarantees that the matrix representation of  $\mathcal{L}_0^*$  is symmetric and negative semidefinite, but this is not a requirement.

The performance of the Galerkin approximation (3.8) is investigated through numerical experiments in Appendix B. An asymptotic analysis of the numerical method in the limit as  $\varepsilon \rightarrow 0$  is presented in [21], which is a longer version of this paper.

**3.4. McKean–Vlasov equation with colored noise.** We consider now the nonlinear McKean–Vlasov initial value problem with OU noise: Recalling that  $\zeta = 1/\sqrt{2}$  in this case,

$$(3.10a) \quad \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \left( \frac{\partial V}{\partial x} \rho + \theta (x - m(t)) \rho - \frac{1}{\varepsilon} \sqrt{\beta^{-1}} \eta \rho \right) + \frac{1}{\varepsilon^2} \frac{\partial}{\partial \eta} \left( \eta \rho + \frac{\partial \rho}{\partial \eta} \right),$$

$$(3.10b) \quad m(t) = \int_{\mathbf{R}} \int_{\mathbf{R}} x \rho(x, \eta, t) \, d\eta \, dx,$$

$$(3.10c) \quad \rho(x, \eta, t = 0) = \rho_0(x, \eta)$$

for some initial distribution  $\rho_0(x, \eta)$  such that the noise is not necessarily started at stationarity. The method that we use in this case, which applies *mutatis mutandis* to the other noise models, is the same as in (3.8), with the addition of the interaction term, and we use the same time-stepping schemes as in (3.1).

Numerical experiments, testing the convergence of the method for the two time-stepping schemes, are presented in Appendix B.

**3.5. Monte Carlo simulations.** We will compare the bifurcation diagrams obtained using the spectral method described above to those obtained by direct MC simulations of the system of interacting particles (2.1). We use the Euler–Maruyama method:

$$X_{k+1}^i = X_k^i - V'(X_k^i) \Delta t - \theta \left( X_k^i - \frac{1}{N} \sum_{j=1}^N X_k^j \right) \Delta t + \frac{\zeta}{\varepsilon} \sqrt{2\beta^{-1}} \eta_k^i \Delta t,$$

where  $\eta_k^i$  is the appropriate projection of the stochastic process  $\mathbf{Y}_t$ . In the case of Gaussian noise, this is discretized as follows:

$$\mathbf{Y}_{k+1}^i = \mathbf{Y}_k^i + \frac{1}{\varepsilon^2} A \mathbf{Y}_k^i \Delta t + \frac{1}{\varepsilon} \sqrt{2 \Delta t} D \xi,$$

where  $\xi \sim N(0, 1)$ , and  $X_k$ ,  $\mathbf{Y}_k$  and  $\eta_k$  are the approximations to  $X(k\Delta t)$ ,  $\mathbf{Y}(k\Delta t)$  and  $\eta(k\Delta t)$ , respectively. The time step used was always  $\mathcal{O}(\varepsilon^2)$ , to ensure the accurate solution of the equation. This scheme has weak order of convergence one (see [23, 24]), and we find that we capture the correct behavior as long as the time step is sufficiently small.

**4. Results: Effect of colored noise on bifurcations.** In this section we present the bifurcation diagrams corresponding to the four models of the noise introduced in section 2. We begin with the case of Gaussian noise, and later move to the case of non-Gaussian noise.

**4.1. Construction of the bifurcation diagrams for the mean field equation.** We constructed the bifurcation diagrams using three different approaches.

*Monte Carlo simulations.* We solved the system of interacting particles (2.1) with a sufficiently large number of particles, and we approximated the first moment by ergodic average over an interval  $(T, T + \Delta T)$ , where  $T$  is sufficiently large to guarantee that the system has reached its stationary state and  $\Delta T$  is sufficiently large to ensure that the ergodic averages are accurate. By applying this procedure for a range of inverse temperatures,  $\beta = 0.1, 0.15, 0.2, \dots, 10$ , we obtained the desired bifurcation diagram.

*Perturbation expansions.* This approach, which we already outlined in section 2, relies on the fact that the self-consistency map can be approximated as  $R(m, \beta) \approx R_0(m, \beta) + \varepsilon^\delta R_\delta(m, \beta)$ , with good accuracy when  $\varepsilon \ll 1$ . Here we used the same notation as in section 2, and, in particular,  $\delta$  denotes the order of the first nontrivial correction in (2.8). Using arclength continuation<sup>1</sup> for the resulting approximate self-consistency equation,  $m = R_0(m, \beta) + \varepsilon^\delta R_\delta(m, \beta)$ , we can plot the first moment  $m$  as a function of  $\beta$  for a fixed value of  $\varepsilon$ . We note that, in view of the typical shape of the self-consistency map, depicted in a particular case in Figure 2.1, a standard root finding algorithm can be employed to initiate the arclength continuation at some initial inverse temperature  $\beta_0$ .

*The spectral method.* Finally, we employed the Galerkin method presented in (3.3a) and (3.3b). We considered two different methodologies: On the one hand, by calculating numerically an approximation  $\rho_{d,\infty}(x, \eta; \beta, m)$  of the steady-state solution

<sup>1</sup>We do this using the Moore–Penrose quasi-arclength continuation algorithm. The rigorous mathematical construction of the arclength continuation methodology can be found, e.g., in [28, 3]. Some useful practical aspects of implementing arclength continuation are also given in [9]. See also [19].

of the linear Fokker–Planck equation (2.6a) with fixed  $m$  and  $\beta$ , we approximated the self-consistency map as  $R(m, \beta) \approx \int_{\mathbf{R}^n} \int_{\mathbf{R}^n} x \rho_{d,\infty}(x, \eta; \beta, m) dx d\mathbf{y}$ , after which a bifurcation diagram can be constructed by using the same method as in the previous paragraph. Each evaluation of the self-consistency map requires the computation of the eigenvector associated with the eigenvalue of smallest magnitude of the discretized operator, which can be performed efficiently for sufficiently small systems using the *SciPy* toolbox. On the other hand, the time-dependent (nonlinear) McKean–Vlasov equation can be integrated directly using our spectral method. Since only the final solution is of interest to us, the semi-implicit time-stepping scheme (3.6) can be used with a large time step, which enables a quick and accurate approximation of the steady-state solutions. While both methodologies work well in the two-dimensional case, in three dimensions (harmonic noise) solving the McKean–Vlasov equation directly proved more efficient, so this is the approach we employed for all the tests presented in this section.

**4.2. Gaussian case.** The one-dimensional Ornstein–Uhlenbeck noise provides an ideal testbed for the three methods we use to construct bifurcation diagrams. Figure 4.1 below plots the bifurcation diagram of the first moment  $m$  as a function of  $\beta$  for  $\varepsilon = 0.1, 0.2, \dots, 0.5$ . Three different initial conditions ( $\mathbf{X}_0 \sim N(0, 0.1)$ ,  $\mathbf{X}_0 \sim N(0.1, 0.1)$ , and  $\mathbf{X}_0 \sim N(-0.1, 0.1)$ ) were used for the MC simulations. Although we observe that the results of MC simulations tend to be less precise around the bifurcation point, the agreement between the three methods overall is excellent for  $\varepsilon = 0.1, 0.2$ . For the other values of  $\varepsilon$ , while the results of MC simulations and of our spectral method continue to agree, those obtained from the asymptotic expansion are significantly less accurate, which is consistent with the observations presented in [44].

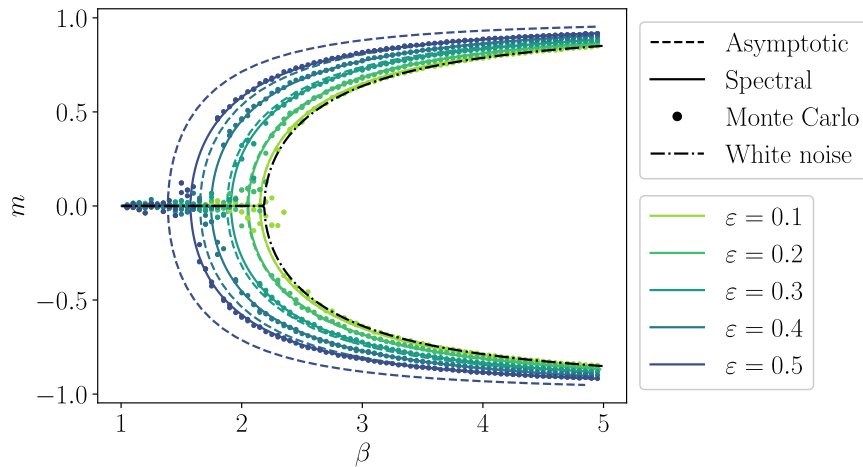


FIG. 4.1. Bifurcation diagram of  $m$  against  $\beta$  for Ornstein–Uhlenbeck noise, obtained via MC simulation, the spectral method, and the asymptotic expansion (2.9).

The case of harmonic noise, corresponding to a three-dimensional McKean–Vlasov equation, is more challenging to tackle using our spectral method. When using 40 basis functions in each direction, the CPU time required to construct the full bifurcation diagram was of the order of a week. As a consequence of the lower number of basis functions used in this case, we observe a small discrepancy between the results of the spectral method and those of MC simulations for large  $\beta$  in the case  $\varepsilon = 0.4$ .

Nevertheless, as can be seen in Figure 4.2, for small  $\varepsilon$  the overall agreement between the three methods is excellent. We note, in particular, that, as suggested by the asymptotic expansions, the use of harmonic noise produces results much closer to the white noise limit than scalar OU noise.

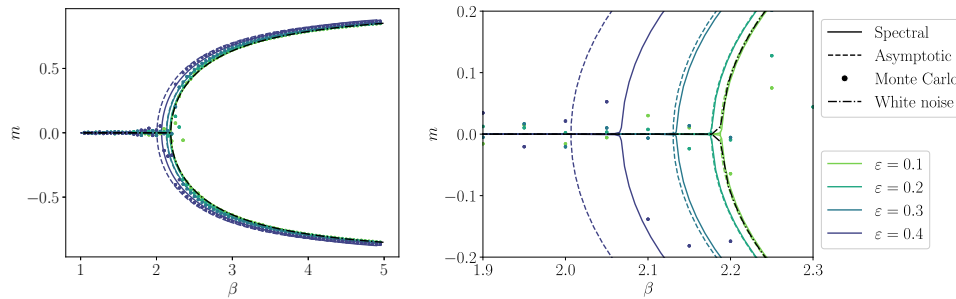


FIG. 4.2. Bifurcation diagram of  $m$  against  $\beta$  for harmonic noise (model **H**), obtained via MC simulation, the Hermite spectral method, and the asymptotic expansion (2.9).

**4.3. Non-Gaussian noise.** For the non-Gaussian noise processes considered in this paper, the  $x^4$  asymptotic growth of the confining potentials in both directions causes the McKean–Vlasov equation to be stiffer than in the cases of OU and harmonic noise, especially for large values of  $\varepsilon$ . Consequently, we were not able to consider as wide a range of  $\varepsilon$  as in the previous subsection using the spectral method. Since, on the other hand, MC simulations become overly computationally expensive for small  $\varepsilon$ , the comparisons in this section comprise only results obtained using our spectral method and asymptotic expansions. Results of simulations for the bistable noise (model **B**) are presented in Figure 4.3, in which a very good agreement can be observed.

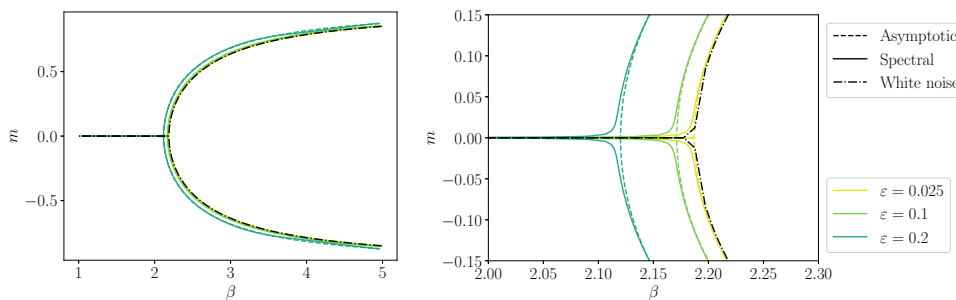


FIG. 4.3. Bifurcation diagram of  $m$  against  $\beta$  for the bistable noise (model **B**), using the spectral method and a truncated asymptotic expansion including the first nonzero correction. We see that, overall, the agreement between the two methods is excellent.

For nonsymmetric noise (model **NS**), the two branches in the bifurcation diagram are separate, as illustrated in Figure 4.4. Here too, the agreement between the spectral method and the asymptotic expansion is excellent. In contrast with the other models considered, the first nonzero term in the asymptotic expansion is of order  $\varepsilon$ , which is reflected by the manifestly higher sensitivity to the correlation time of the noise. In the right panel of Figure 4.4, we present the graph of  $R_0(m; \beta) + \varepsilon R_1(m; \beta)$  for a value of  $\beta$  close to the point at which new branches (one stable and one unstable) emerge.



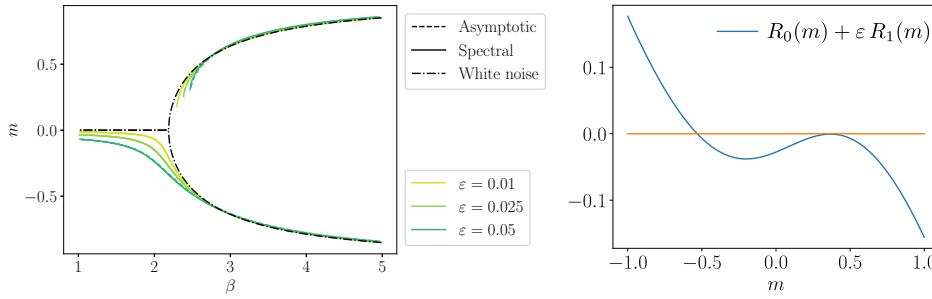


FIG. 4.4. Left: bifurcation diagram of  $m$  against  $\beta$  for the nonsymmetric noise (model **NS**), using the spectral method and a truncated asymptotic expansion including the first nonzero correction. Right:  $R_0 + \varepsilon R_1(m) - m$  against  $m$  for  $\varepsilon = 0.1$  and  $\beta = 2.6$ .

**4.4. Dependence of the critical temperature on  $\varepsilon$ .** For the noise models **OU**, **H**, and **B**, the effect of colored noise on the dynamics is a shift of the critical temperature: The pitchfork bifurcation occurs for *smaller* values of  $\beta$  (i.e., larger temperatures) as the correlation time increases. In order to further investigate the effect of the correlation time on the long time behavior of the system of interacting particles, we will compute the critical temperature as a function of  $\varepsilon$  based on the asymptotic expansions and compare with the results of spectral and MC simulations; see Figure 4.5. Rather than finding the critical inverse temperature  $\beta_C$  for a range of values of  $\varepsilon$  (and for a fixed  $\theta$ ), it is convenient to fix  $\beta_C$  and find the corresponding  $\varepsilon$ , satisfying

$$(4.1) \quad \frac{d}{dm} \left( \int_{\mathbf{R}} x p_0(x; \beta_C, m) dx \right)_{m=0} + \varepsilon^\delta \frac{d}{dm} \left( \int_{\mathbf{R}} p_\delta(x; \beta_C, m) dx \right)_{m=0} = 1,$$

which is merely a polynomial equation in  $\varepsilon$ , the coefficient of which can be calculated by numerical differentiation. With this procedure, the dependence of the critical  $\beta$  upon  $\varepsilon$  can be calculated on a fine mesh. In the case of **OU** noise, for example, both coefficients on the left-hand side of (4.1) are positive, implying that the equation has a solution (in fact, two, but one of them negative) only if  $\beta_C$  is lower than the inverse critical temperature in the white noise case.

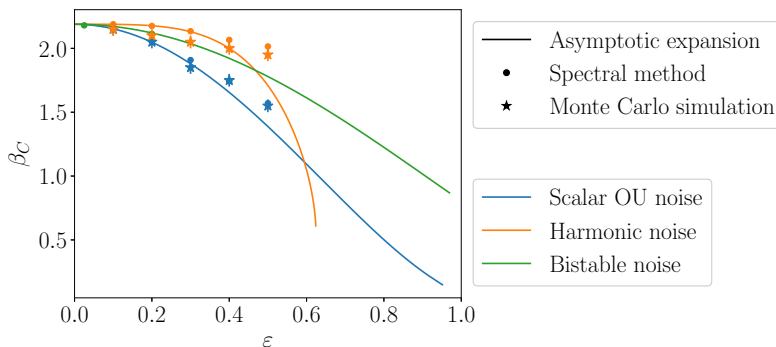


FIG. 4.5. Critical  $\beta$  against  $\varepsilon$ .

Of the three methods employed in Figure 4.5, the approach based on the asymp-

otic expansions has the lowest computational cost: Calculating all the solid curves took only about a couple of minutes on a personal computer with an Intel i7-3770 processor. The data points associated with the spectral method and the MC simulations were obtained from the bifurcation diagrams presented above.

**5. Conclusions.** In this paper, we introduced a robust spectral method for the numerical solution of linear and nonlinear, local and nonlocal Fokker–Planck-type PDEs that does not require that the PDE is a gradient flow. We then used our method to construct the bifurcation diagram for the stationary solutions of the mean field limit of a system of weakly interacting particles driven by colored noise.

To verify our results, we also constructed the bifurcation diagrams by using two other independent approaches, namely by MC simulation of the  $N$ -particle system and by using explicit asymptotic expansions with respect to correlation time of the noise. In the small correlation time regime, we observed a very good agreement between all three methods. For larger values of the correlation time, the asymptotic expansions become inaccurate, but the results obtained via the spectral method and MC simulations continue to be in good agreement.

It appeared from our study that, unless the potential in which the noise process is confined is asymmetric, the correlation structure of the noise does not influence the topology of the bifurcation diagram: The mean-zero steady-state solution, which is stable for sufficiently large temperatures, becomes unstable as the temperature decreases below a critical value, at which point two new stable branches emerge in the same manner as reported in [8, 41]. The correlation structure does, however, influence the temperature at which bifurcation occurs, and, in general, this temperature increases as the correlation time of the noise increases. In the presence of an asymmetry in the confining potential of the noise, on the other hand, the two stable branches in the bifurcation diagram are separate, indicating that the system always reaches the same equilibrium upon slowly decreasing the temperature. This behavior is similar to what has been observed previously in the white noise case when a tilt is introduced in the confining potential  $V(\cdot)$ ; see [19, 20].

Several problems remain open for future work. On the theoretical front, we believe that the analysis we presented in (3.1) and Appendix A for the linear Fokker–Planck equation can be extended to both the linear Fokker–Planck equation with colored noise and the nonlinear McKean–Vlasov equation. Another direction for future research could be the rigorous study of bifurcations and, more specifically, of fluctuations and critical slowing down near the bifurcation point. On the modeling front, it would be interesting to consider more general evolution equations for the interacting particles, such as the generalized Langevin equation, and also to study systems of interacting particles subject to colored noise that is multiplicative.

**Appendix A. Proof of Theorem 3.1.** Using the same notation as in [21, Appendix A], we let  $\Pi_d$  be the  $L^2(\mathbf{R}; e^{-V_q})$  projection operator on  $\mathbf{P}(d)$  and  $\hat{\Pi}_d := e^{-V_q/2} \Pi_d e^{V_q/2}$ . The solution  $u_d$  of (3.3a) satisfies  $\partial_t u_d = \hat{\Pi}_d \mathcal{H}_x \hat{\Pi}_d u_d =: \mathcal{H}_d u_d$ . Clearly, the operator  $\mathcal{H}_d$  is self-adjoint on  $e^{-V_q/2} \mathbf{P}(d)$  with the  $L^2(\mathbf{R})$  inner product, and it is also negative, by negativity of  $\mathcal{H}_x$ :

$$(A.1) \quad \langle \mathcal{H}_d w_d, w_d \rangle = \langle \mathcal{H}_x w_d, w_d \rangle \leq 0 \quad \forall w_d \in e^{-V_q/2} \mathbf{P}(d).$$

To prove the convergence of  $u_d$  when  $d \rightarrow \infty$ , we will rely on the following lemma.

**LEMMA A.1.** *Let  $\hat{\partial}_x := \partial_x + x/2$ , and assume that  $\hat{\partial}_x^n u \in L^2(\mathbf{R})$  for  $n = 0, \dots, m$ . Then for all natural numbers  $m_1, m_2$  such that  $m_1 + m_2 \leq m$ , it holds that  $x^{m_1} u^{(m_2)} \in$*

$L^2(\mathbf{R})$  and

$$(A.2) \quad K_1(m) \max_{m_1+m_2 \leq m} \|x^{m_1} u^{(m_2)}\| \leq \max_{0 \leq i \leq m} \|\hat{\partial}_x^i u\| \\ \leq K_2(m) \max_{m_1+m_2 \leq m} \|x^{m_1} u^{(m_2)}\|,$$

where  $K_1(m), K_2(m)$  are positive constants depending only on  $m$  and  $\|\cdot\|$  is the usual  $L^2(\mathbf{R})$  norm.

*Proof.* We denote by  $H^m(\mathbf{R}; e^{-x^2/2})$  the Sobolev space weighted by  $e^{-x^2/2}$ ,

$$H^m(\mathbf{R}; e^{-x^2/2}) = \{v : v^{(i)} \in L^2(\mathbf{R}; e^{-x^2/2}) \text{ for } i = 0, \dots, m\},$$

and by  $\|\cdot\|_{m, e^{-x^2/2}}$  the associated norm:  $\|v\|_{m, e^{-x^2/2}}^2 = \sum_{i=0}^m \|v^{(i)}\|_{e^{-x^2/2}}^2$ . For the first inequality, we know from [40, Lemma B.6] that

$$\|xv\|_{e^{-x^2/2}} \leq 4 \|v\|_{1, e^{-x^2/2}} \quad \forall v \in H^1(\mathbf{R}; e^{-x^2/2}).$$

Applying this inequality repeatedly, we obtain

$$(A.3) \quad \|x^{m_1} v\|_{e^{-x^2/2}} \leq C(m) \|v\|_{m, e^{-x^2/2}}, \quad m_1 = 0, \dots, m, \quad \forall v \in H^m(\mathbf{R}; e^{-x^2/2})$$

for a constant  $C(m)$  depending only on  $m$ . By definition,  $\hat{\partial}_x u = e^{-x^2/4} \partial_x(e^{x^2/4} u)$ , so the assumption implies that  $e^{x^2/4} u \in H^m(\mathbf{R}; e^{-x^2/2})$ , from which we obtain using (A.3) that, for  $0 \leq m_1 \leq m$ ,

$$\|x^{m_1} u\| = \|x^{m_1} u e^{x^2/4}\|_{e^{-x^2/2}} \leq C(m) \|u e^{x^2/4}\|_{m, e^{-x^2/2}} = C(m) \sqrt{\sum_{i=0}^m \|\hat{\partial}_x^i u\|^2}.$$

This proves the first inequality of (A.2) in the case  $m_2 = 0$ . We assume now that the statement is proved up to  $m_2 - 1$ , and we show that it is valid for  $m_2$ . Using the triangle inequality we obtain

$$\|x^{m_1} u^{(m_2)}\| \leq \|x^{m_1} (u^{(m_2)} - \hat{\partial}_x^{m_2} u)\| + \|x^{m_1} \hat{\partial}_x^{m_2} u\|.$$

The derivatives in the first term are of order strictly lower than  $m_2$ . Therefore, this term can be bounded by the induction assumption. The second term is bounded by applying the base case to  $\hat{\partial}_x^{m_2} u$ : Introducing  $v := \hat{\partial}_x^{m_2} u$ , we notice that  $\hat{\partial}_x^{m-m_2} v = \hat{\partial}_x^m u \in L^2(\mathbf{R})$  by assumption, so we can apply the first inequality in (A.2), without any derivative of  $v$  in the left-hand side, to deduce

$$\|x^{m-m_2} v\| \leq \max_{0 \leq i \leq m-m_2} \|\hat{\partial}_x^i v\| \leq \max_{0 \leq i \leq m} \|\hat{\partial}_x^i u\|.$$

The second inequality in (A.2) then holds trivially by expanding  $\hat{\partial}_x$  and applying a triangle inequality.  $\square$

With Assumption 3.1, we can show that the two norms in Lemma A.1 can be bounded from above by the norm  $\sqrt{\langle (-\mathcal{H}_x + 1)^m u, u \rangle}$  for appropriate  $m$ .

LEMMA A.2 (bound by alternative norm). *If Assumption 3.1 holds, then*

$$\sum_{i=0}^m \|\hat{\partial}_x^m u\|^2 \leq C \langle (-\mathcal{H}_x + 1)^m u, u \rangle$$

for any smooth  $u$  for which the right-hand side is well-defined. Here  $C$  is a positive constant that depends on  $\beta$ ,  $m$ , and on the particular expression of the potential  $W$  defined in Assumption 3.1.

*Proof.* Below  $C_1$  and  $C_2$  denote the same constants as in Assumption 3.1. First, we notice that, for any constant  $K > 1$ ,

$$(A.4) \quad \langle (-\mathcal{H}_x + K)^m u, u \rangle \leq K^m \langle (-\mathcal{H}_x + 1)^m u, u \rangle,$$

because  $\mathcal{H}_x$  is a negative operator. Since  $W$  is a polynomial, its derivatives grow asymptotically more slowly than  $W$  itself, and so it is possible for any  $\varepsilon > 0$  to find  $K \geq C_2$  large enough that

$$(A.5) \quad |W^{(i)}(x)| \leq \varepsilon (W(x) + K) \quad \forall x \in \mathbf{R}, \quad i = 1, 2, \dots$$

For this proof to go through, it is, in fact, sufficient that this inequality be satisfied for  $i = 1, \dots, m$ . We decompose  $-\mathcal{H}_x + K$  as  $(-\beta^{-1}\partial_x^2) + (W(x) + K)$ . The two operators in this sum are positive because  $K \geq C_2$  and by assumption  $W(x) + C_2 \geq C_1(1 + |x|^2)$ . Expanding the inner product in the left-hand side of (A.4) and using integration by parts,

$$(A.6) \quad \langle (-\mathcal{H}_x + K)^m u, u \rangle = \left( \sum_{i=0}^m \beta^{-i} \binom{m}{i} \int_{\mathbf{R}} (W(x) + K)^i (u^{(m-i)}(x))^2 dx \right) + \dots,$$

where the remainder terms originate from the fact that the operators  $\partial_x$  and  $(W(x) + K)$  do not commute. By (A.5), these terms can be bounded for sufficiently large  $K$  by half the leading term in (A.6). To conclude, we further expand this leading term:

$$\begin{aligned} \langle (-\mathcal{H}_x + K)^m u, u \rangle &\geq \frac{1}{2} \sum_{i=0}^m \binom{m}{i} \beta^{-i} \int_{\mathbf{R}} (W(x) + K)^i (u^{(m-i)}(x))^2 dx \\ &\geq \frac{1}{2} \sum_{i=0}^m \binom{m}{i} C_1^i \beta^{-i} \int_{\mathbf{R}} (1 + x^2)^i (u^{(m-i)}(x))^2 dx \\ &\geq \frac{1}{2} \sum_{i=0}^m \sum_{j=0}^i \binom{m}{i} \binom{i}{j} C_1^i \beta^{-i} \int_{\mathbf{R}} x^{2j} (u^{(m-i)}(x))^2 dx \\ &\geq C(m, \beta, C_1) \sum_{m_1+m_2 \leq m} \|x^{m_1} u^{(m_2)}\|^2, \end{aligned}$$

from which Lemma A.1 allows us to conclude.  $\square$

*Proof of Theorem 3.1.* We assume for simplicity that  $\sigma = 1$ , and we begin by splitting the error as  $u_d - u = (u_d - \hat{\Pi}_d u) + (\hat{\Pi}_d u - u) =: e_d + \delta_d$ . The first term is related to the so-called consistency error, and the second to the approximation error. We obtain from (3.1) and (3.3a)

$$\partial_t e_d = \hat{\Pi}_d \mathcal{H}_x \hat{\Pi}_d e_d + (\hat{\Pi}_d \mathcal{H}_x \hat{\Pi}_d - \hat{\Pi}_d \mathcal{H}_x) u.$$

Taking the inner product with  $e_d$  and using (A.1), this implies

$$\begin{aligned} \langle \partial_t e_d, e_d \rangle &\leq \langle \mathcal{H}_x(\hat{\Pi}_d u - u), e_d \rangle \\ &\leq \frac{\alpha}{2} \langle e_d, e_d \rangle + \frac{1}{2\alpha} \langle \mathcal{H}_x^2(u - \hat{\Pi}_d u), (u - \hat{\Pi}_d u) \rangle \quad \forall \alpha > 0, \end{aligned}$$

where we used Young’s inequality. We see from this equation that  $e_d$  can be controlled if one can bound the second inner product on the right-hand side. For this we use arguments similar to the ones employed in [1, 14]. Since  $\mathcal{H}_x$  is negative and self-adjoint, we notice that

$$\begin{aligned} \langle (-\mathcal{H}_x)^i u(t), u(t) \rangle &= \langle (-\mathcal{H}_x)^i u_0, u_0 \rangle + \int_0^t \frac{d}{ds} \langle (-\mathcal{H}_x)^i u(s), u(s) \rangle ds \\ &= \langle (-\mathcal{H}_x)^i u_0, u_0 \rangle - 2 \int_0^t \langle (-\mathcal{H}_x)^{i+1} u(s), u(s) \rangle ds \\ \text{(A.7)} \quad &\leq \langle (-\mathcal{H}_x)^i u_0, u_0 \rangle \end{aligned}$$

for  $i = 1, 2, \dots$ , which implies that the inner products  $\langle (-\mathcal{H}_x)^i u, u \rangle$  remain bounded for all positive times. We can now apply [21, Corollary A.2] to obtain, using Lemmas A.1 and A.2 and Assumption 3.1,

$$\begin{aligned} \langle \mathcal{H}_x^2(u - \hat{\Pi}_d u), (u - \hat{\Pi}_d u) \rangle &\leq C \sum_{i=0}^{2k} \|\hat{\partial}_x^i (u - \hat{\Pi}_d u)\|^2 \\ &\leq C \frac{(d - m + 1)!}{(d - 2k + 1)!} \|\hat{\partial}_x^m u\|^2 \\ &\leq C \frac{(d - m + 1)!}{(d - 2k + 1)!} \langle (-\mathcal{H}_x + 1)^m u, u \rangle \\ &\leq C \frac{(d - m + 1)!}{(d - 2k + 1)!} \langle (-\mathcal{H}_x + 1)^m u_0, u_0 \rangle. \end{aligned}$$

We note that when  $V$  is quadratic,  $k = 1$  is a valid choice in Assumption 3.1, and the bound above can be obtained by simply expanding  $u$  in terms of the eigenfunctions of  $\mathcal{H}_x$ , which in that case are just rescaled Hermite functions. Using Grönwall’s inequality, we finally obtain

$$\begin{aligned} \|e_d(t)\|^2 &\leq e^{\alpha t} \|e_d(0)\|^2 + \int_0^t e^{\alpha(t-s)} \langle \mathcal{H}_x^2(u - \hat{\Pi}_d u), (u - \hat{\Pi}_d u) \rangle ds, \\ \text{(A.8)} \quad &\leq e^{\alpha t} \left( \|e_d(0)\|^2 + C_\alpha \frac{(d - m + 1)!}{(d - 2k + 1)!} \right). \end{aligned}$$

The first term, proportional to  $\|e_d(0)\|^2$ , depends only on the interpolation error of the initial condition, which is nonzero when using a Gauss–Hermite quadrature. It was proved that this error term also decreases spectrally (see, e.g., [40, Theorems 7.17, 7.18]), and in our case faster than the second error term. For the approximation error  $\delta_d$ , similar inequalities to the ones used above can be used to obtain a bound of the type (A.8), which leads to the conclusion.  $\square$

*Remark A.1.* As mentioned in Remark 3.1, Theorem 3.1 is not optimal. It leaves open, in particular, the question of precisely how the error behaves as  $t \rightarrow \infty$ . In

this remark, we give a partial answer to the question: We show how a bound on the stationary error can be obtained, under the assumption that the solution to the discretized equation is rescaled in time in such a way that the integral of  $\rho_d := \sqrt{\rho_s} u_d$  remains equal to 1. With this rescaling and with the notation  $S_d := \hat{\Pi}_d(L^2(\mathbf{R}))$ , the stationary solution of the discretized (in space) equation is

$$\hat{u}_s := \arg \max_{\psi_d \in S_d, \int_{\mathbf{R}} \sqrt{\rho_s} \psi_d = 1} \langle \mathcal{H}_x \psi_d, \psi_d \rangle.$$

Taking

$$\psi_d = \frac{\hat{\Pi} \sqrt{\rho_s}}{\int_{\mathbf{R}} \sqrt{\rho_s} \hat{\Pi} \sqrt{\rho_s}},$$

we deduce

$$-\langle \mathcal{H}_x \hat{u}_s, \hat{u}_s \rangle \leq - \left| \frac{1}{\int_{\mathbf{R}} \sqrt{\rho_s} \hat{\Pi} \sqrt{\rho_s}} \right|^2 \langle \mathcal{H}_x(\hat{\Pi} \sqrt{\rho_s}), \hat{\Pi} \sqrt{\rho_s} \rangle.$$

Since  $\mathcal{H}_x$  is self-adjoint in  $L^2(\mathbf{R})$  and  $\mathcal{H}_x \sqrt{\rho_s} = 0$ , it follows that

$$-\langle \mathcal{H}_x(\hat{u}_s - \sqrt{\rho_s}), \hat{u}_s - \sqrt{\rho_s} \rangle \leq - \left| \frac{1}{\int_{\mathbf{R}} \sqrt{\rho_s} \hat{\Pi} \sqrt{\rho_s}} \right|^2 \langle \mathcal{H}_x(\hat{\Pi} \sqrt{\rho_s} - \sqrt{\rho_s}), \hat{\Pi} \sqrt{\rho_s} - \sqrt{\rho_s} \rangle.$$

An approximation argument similar to the one above can be employed to show that the right-hand side, and therefore also the left-hand side, decrease to zero as  $d \rightarrow \infty$  faster than  $d^{-n}$  for any  $n > 0$ . To conclude, a Poincaré-type inequality can be invoked, which is justified because

$$\int_{\mathbf{R}} (\hat{u}_s - \sqrt{\rho_s}) \sqrt{\rho_s} = 1 - 1 = 0$$

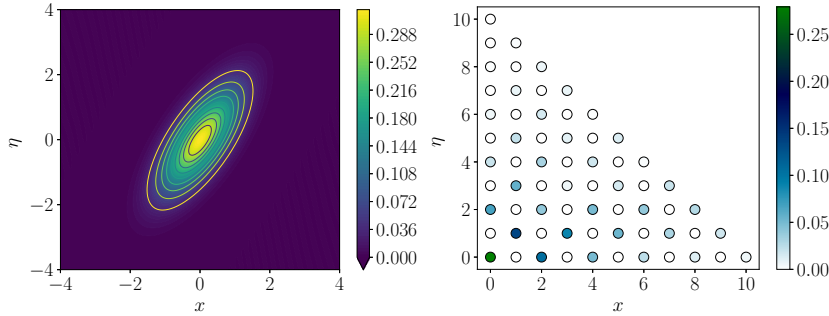
to obtain a bound of the type

$$\|\hat{u}_s - \sqrt{\rho_s}\| \leq -C \langle \mathcal{H}_x(\hat{u}_s - \sqrt{\rho_s}), \hat{u}_s - \sqrt{\rho_s} \rangle.$$

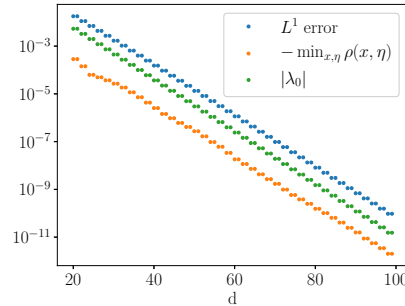
**Appendix B. Benchmark tests for the spectral numerical method.** In this section, we investigate the performance of the spectral method through numerical experiments.

**B.1. Linear Fokker–Planck equation with colored noise.** We focus first on the Galerkin approximation (3.8). Here we consider only the cases where  $V(\cdot)$  is a quadratic or a bistable potential and where the noise is described by an OU process, but results of additional numerical experiments, corresponding to harmonic noise and non-Gaussian noise, are presented in [44].

We start with the case  $V(x) = x^2/2$ , for which the exact solution to the Fokker–Planck equation (3.7) can be calculated explicitly by substitution of a Gaussian ansatz; see [36, section 3.7]. We study the convergence of the steady-state solution, obtained by calculating the eigenfunction associated with the eigenvalue of lowest magnitude of  $\hat{\Pi}_d \mathcal{L}_\varepsilon^* \hat{\Pi}_d$ , where  $\hat{\Pi}_d$  is the  $L^2(\mathbf{R}^2; e^U)$  projection operator on  $S_d$ , directly using the method `eigs` from the `SciPy` toolbox. The parameters used for this simulation are the following:  $\beta = \varepsilon = 1$ ,  $\sigma_x^2 = \frac{1}{10}$ ,  $\sigma_\eta^2 = 1$ ,  $e^{-U(x,\eta)/2} = e^{-V_\eta(\eta)/2} = e^{-\eta^2/4}$ . With these parameters, the steady-state solution to (3.3a) and (3.3b) is equal to



(a) Steady-state solution of the Fokker–Planck equation (3.8) with the associated field lines of the probability flux (left) and absolute value of the coefficients of degree less than equal to 10 in the Hermite expansion (right).



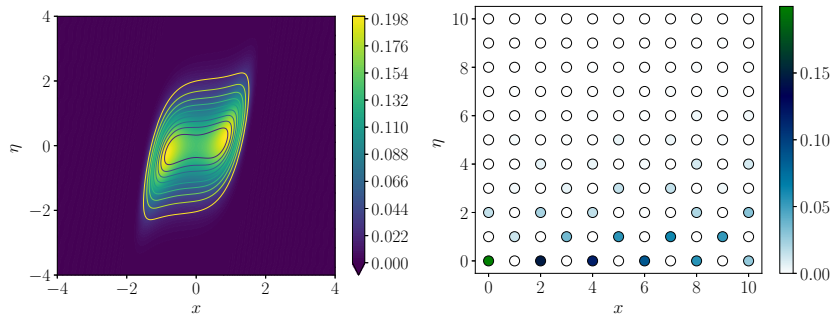
(b) Convergence of the method, using three different metrics for the error: The  $L^1$  norm of the error between the numerical and exact solutions, the negative of the minimum of the numerical solution, and the absolute value of the eigenvalue with smallest real part.

FIG. B.1. Simulation data when  $V(x) = x^2/2$ .

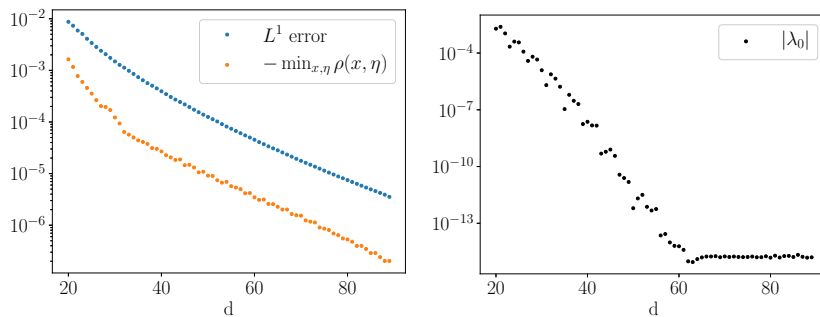
$\rho_\infty(x, \eta) = e^{-2x^2+2x\eta-\eta^2} / \pi$ , and clearly  $\rho_\infty \in L^2(\mathbf{R}^2; e^U)$ . Figure B.1a presents the steady-state solution, obtained using the spectral method with Hermite polynomials up to degree 100 ( $d = 100$ ) and a triangular index set, and Figure B.1b presents the convergence of the method. Since the solution satisfies  $\rho_\infty(x, \eta) = \rho_\infty(-x, -\eta)$ , the Hermite coefficients corresponding to even values of  $i + j$  are zero, where  $i$  and  $j$  are the indices in the  $x$  and  $\eta$  directions, respectively.

Now we consider that  $V$  is the bistable potential  $x^4/4 - x^2/2$ , which was solved numerically in [22] using generalized Hermite functions and a variation of the matrix continued fraction technique. For this case an explicit analytical solution is not available. The parameters we use are the following:  $\beta = 1$ ,  $\varepsilon = \frac{1}{2}$ ,  $\sigma_x^2 = \frac{1}{20}$ ,  $\sigma_\eta^2 = 1$ . Through numerical exploration, we noticed that a good convergence could be obtained by using the multiplier function  $e^{-U(x,\eta)/2} = e^{-\beta V(x)/2 - \eta^2/4}$ , rather than just  $e^{-\eta^2/4}$  in the previous paragraph. We note that this would have been the natural choice if the noise in the  $x$  direction had been white noise. The solution obtained using a square-shaped index set and  $d = 100$ , as well as the corresponding Hermite coefficients up to degree 10, is illustrated in Figure B.2a. We observe that the Hermite coefficients corresponding to the degree 0 in the  $\eta$  direction (i.e., to the basis function  $e^{-\eta^2/2}$ ) are significantly larger than the other coefficients, which is consistent with the fact

that, as  $\varepsilon \rightarrow 0$ , the steady-state solution approaches  $e^{-\beta V(x)} e^{-\eta^2/2}$  (up to a constant factor). The associated convergence curves are presented in Figure B.2b. We observe that the convergence is exponential, which is better than the rate of convergence predicted in Theorem 3.1.



(a) Steady-state solution to (3.8) and associated field lines of the probability flux (left), and absolute value of the coefficients of degree less than or equal to 10 in the Hermite expansion (right).



(b) Convergence of the method using the same measures of the error as in Figure B.1a, except that the  $L^1$  error is calculated by comparison with the numerical solution obtained when  $d = 100$ .

FIG. B.2. Simulation data when  $V(x) = x^4/4 - x^2/2$ .

**B.2. McKean–Vlasov equation with colored noise.** We focus now on the Galerkin approximation to (3.10). When the potential  $V(\cdot)$  is quadratic and the initial condition is Gaussian, it is well-known that the McKean–Vlasov equation has an explicit solution and that this solution is Gaussian. We assume that  $V(x) = x^2/2$  and we rewrite (3.10) in the formalism of [10] as

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \left( B \mathbf{x} \rho + \int_{\mathbf{R}^2} K(\mathbf{x} - \mathbf{x}') \rho(\mathbf{x}', t) d\mathbf{x}' \rho - D \nabla \rho \right),$$

where  $\mathbf{x} = (x, \eta)^T$  and

$$B = \begin{pmatrix} -1 & \varepsilon^{-1} \beta^{-1/2} \\ 0 & -\varepsilon^{-2} \end{pmatrix}, \quad K = \begin{pmatrix} -\theta & 0 \\ 0 & 0 \end{pmatrix}, \quad D = \begin{pmatrix} 0 & 0 \\ 0 & \varepsilon^{-2} \end{pmatrix}.$$



Adapting [10, Proposition 2.3] to our case, we deduce that the solution is of the type

$$\rho(\mathbf{x}, t) = \frac{1}{(2\pi)^d |\Sigma(t)|} \exp\left(-\frac{1}{2} (\mathbf{x} - \mu(t))^T \Sigma^{-1}(t) (\mathbf{x} - \mu(t))\right),$$

where  $\mu(t)$  and  $\Sigma(t)$  are given by (B.1)

$$\mu(t) = e^{Bt} \mu(0), \quad \Sigma(t) = e^{t(B+K)} \Sigma(0) e^{t(B+K)^T} + 2 \int_0^t e^{s(B+K)} D e^{s(B+K)^T} ds.$$

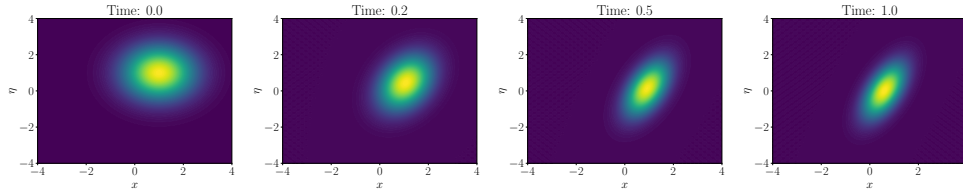


FIG. B.3. Probability density solution of (3.10) (obtained using the spectral method) at times 0, 0.2, 0.5, 1.

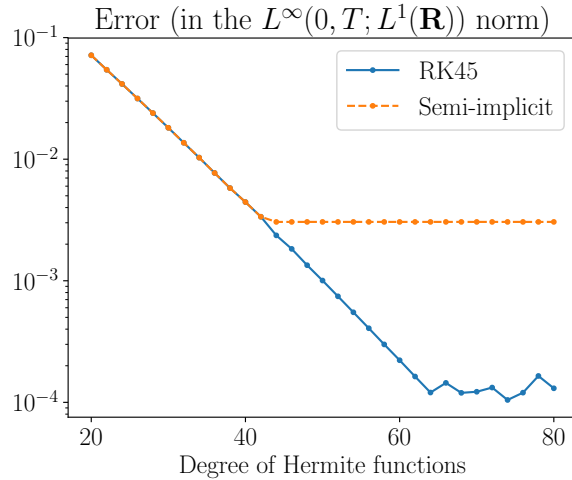


FIG. B.4. Convergence of the Hermite spectral method for the nonlinear McKean–Vlasov equation with colored noise. The error decreases exponentially for low enough values of  $d$ , and then reaches a plateau when it becomes dominated by the error induced by the time discretization.

This solution can be obtained by introducing  $g = -\ln \rho$ , rewriting (3.10) as an equation for  $g$ , and using a quadratic ansatz for  $g$ . The eigenvalue decomposition of  $B + K$  is

$$(B + K) \begin{pmatrix} 1 & -\varepsilon \\ 0 & \sqrt{\beta}(1 - \varepsilon^2(1 + \theta)) \end{pmatrix} = \begin{pmatrix} 1 & -\varepsilon \\ 0 & \sqrt{\beta}(1 - \varepsilon^2(1 + \theta)) \end{pmatrix} \begin{pmatrix} -1 - \theta & 0 \\ 0 & -\varepsilon^{-2} \end{pmatrix},$$

which enables the explicit calculation of the integral in the expression of  $\Sigma(t)$ . From (B.1) and the structure of  $B$  and  $K$ , it holds that  $\mu(t) \rightarrow 0$  and

$$\Sigma(t) \rightarrow \Sigma_\infty = 2 \int_0^\infty e^{s(B+K)} D e^{s(B+K)^T} ds,$$

in the limit as  $t \rightarrow \infty$ . The limiting values for  $\mu(t)$  and  $\Sigma(t)$  coincide with the mean and covariance matrix, respectively, of the solution to the steady state linear Fokker–Planck equation corresponding to (3.10) with  $m(t) = 0$ . For this test case, we use the following parameters:  $\beta = \theta = 1$ ,  $\varepsilon = 1/2$ ,  $\sigma_x^2 = \sigma_\eta^2 = 1/5$ ,  $e^{-U(x,\eta)/2} = e^{-V(x)/2} e^{-V_\eta(\eta)/2}$ . The initial condition is taken to be the Gaussian density  $\mathcal{N}((1, 1)^T, I_{2 \times 2})$ . The evolution of the probability density is illustrated in Figure B.3, and the convergence of the method, in the  $L^\infty(0, T; L^1(\mathbf{R}^2))$  norm, is illustrated in Figure B.4.

**Acknowledgments.** The authors are grateful to J.A. Carrillo for making available the finite volume code we employed for validating our spectral method for the McKean–Vlasov dynamics with white noise; the associated numerical results are presented in [44]. The authors are also grateful to B. Goddard and D. Papageorgiou for useful comments and suggestions.

## REFERENCES

- [1] A. ABDULLE, G. A. PAVLIOTIS, AND U. VAES, *Spectral methods for multiscale stochastic differential equations*, SIAM/ASA J. Uncertain. Quantif., 5 (2017), pp. 720–761, <https://doi.org/10.1137/16M1094117>.
- [2] S. AGMON, *Lectures on Exponential Decay of Solutions of Second-order Elliptic Equations: Bounds on Eigenfunctions of N-body Schrödinger Operators*, Math. Notes 29, Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo, 1982.
- [3] E. L. ALLGOWER AND K. GEORG, *Introduction to numerical continuation methods*, Classics Appl. Math. 45, SIAM, Philadelphia, 2003, <https://doi.org/10.1137/1.9780898719154>.
- [4] N. BAIN AND D. BARTOLO, *Critical mingling and universal correlations in model binary active liquids*, Nat. Commun., 8 (2017), 15969.
- [5] J. BINNEY AND S. TREMAINE, *Galactic Dynamics*, Princeton Series in Astrophysics 20, Princeton University Press, Princeton, NJ, 2011.
- [6] G. BLANKENSHIP AND G. C. PAPANICOLAOU, *Stability and control of stochastic systems with wide-band noise disturbances*, I, SIAM J. Appl. Math., 34 (1978), pp. 437–476, <https://doi.org/10.1137/0134036>.
- [7] J. A. CARRILLO, A. CHERTOCK, AND Y. HUANG, *A finite-volume method for nonlinear nonlocal equations with a gradient flow structure*, Commun. Comput. Phys., 17 (2015), pp. 233–258, <https://doi.org/10.4208/cicp.160214.010814a>.
- [8] D. A. DAWSON, *Critical dynamics and fluctuations for a mean-field model of cooperative behavior*, J. Statist. Phys., 31 (1983), pp. 29–85, <https://doi.org/10.1007/BF01010922>.
- [9] A. DHOOGHE, W. GOVAERTS, Y. A. KUZNETSOV, W. MESTROM, A. M. RIET, AND B. SAUTOIS, *MATCONT and CL MATCONT: Continuation toolboxes in MATLAB*, Universiteit Gent, Belgium and Utrecht University, The Netherlands, 2006.
- [10] M. H. DUONG AND G. A. PAVLIOTIS, *Mean field limits for non-Markovian interacting particles: convergence to equilibrium, GENERIC formalism, asymptotic limits and phase transitions*, Commun. Math. Sci., 16 (2018), pp. 2199–2230, <https://doi.org/10.4310/CMS.2018.v16.n8.a7>.
- [11] A. DURMUS, A. EBERLE, A. GUILLIN, AND R. ZIMMER, *An Elementary Approach to Uniform in Time Propagation of Chaos*, preprint, <https://arxiv.org/abs/1805.11387>, 2018.
- [12] F. FARKHOUI AND W. STANNAT, *A Complete Mean-field Theory for Dynamics of Binary Recurrent Neural Networks*, preprint, <https://arxiv.org/abs/1701.07128>, 2017.
- [13] 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, <https://doi.org/10.1090/S0025-5718-01-01365-5>.
- [14] J. GAGELMAN AND H. YSERENTANT, *A spectral method for Schrödinger equations with smooth confinement potentials*, Numer. Math., 122 (2012), pp. 383–398, <https://doi.org/10.1007/s00211-012-0458-8>.
- [15] J. GARNIER, G. PAPANICOLAOU, AND T.-W. YANG, *Large deviations for a mean field model of systemic risk*, SIAM J. Finan. Math., 4 (2013), pp. 151–184, <https://doi.org/10.1137/12087387X>.
- [16] J. GARNIER, G. PAPANICOLAOU, AND T.-W. YANG, *Consensus convergence with stochastic effects*, Vietnam J. Math., 45 (2017), pp. 51–75, <https://doi.org/10.1007/s10013-016-0190-2>.

- [17] B. D. GODDARD, A. NOLD, N. SAVVA, G. A. PAVLIOTIS, AND S. KALLIADASIS, *General dynamical density functional theory for classical fluids*, Phys. Rev. Lett., 109 (2012), 120603, <https://doi.org/10.1103/PhysRevLett.109.120603>.
- [18] B. D. GODDARD, G. A. PAVLIOTIS, AND S. KALLIADASIS, *The overdamped limit of dynamic density functional theory: rigorous results*, Multiscale Model. Simul., 10 (2012), pp. 633–663, <https://doi.org/10.1137/110844659>.
- [19] S. N. GOMES, S. KALLIADASIS, G. A. PAVLIOTIS, AND P. YATSYSHIN, *Dynamics of the Desai-Zwanzig model in multiwell and random energy landscapes*, Phys. Rev. E, 99 (2019), 032109, <https://doi.org/10.1103/PhysRevE.99.032109>.
- [20] S. N. GOMES AND G. A. PAVLIOTIS, *Mean field limits for interacting diffusions in a two-scale potential*, J. Nonlinear Sci., 28 (2018), pp. 905–941, <https://doi.org/10.1007/s00332-017-9433-y>.
- [21] S. N. GOMES, G. A. PAVLIOTIS, AND U. VAES, *Mean-field Limits for Interacting Diffusions with Colored Noise: Phase Transitions and Spectral Numerical Methods*, preprint, <https://arxiv.org/abs/1904.05973>, 2020.
- [22] P. HANGGI AND P. JUNG, *Colored noise in dynamical systems*, Adv. in Chem. Phys., 89 (1995), pp. 239–326.
- [23] D. J. HIGHAM, *An algorithmic introduction to numerical simulation of stochastic differential equations*, SIAM Rev., 43 (2001), pp. 525–546, <https://doi.org/10.1137/S0036144500378302>.
- [24] 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, <https://doi.org/10.1137/S0036142901389530>.
- [25] W. HORSTHEMKE AND R. LEFEVER, *Noise-induced Transitions: Theory and Application in Physics, Chemistry, and Biology*, Springer Ser. Synergetics 15, Springer-Verlag, Berlin, Heidelberg, 1984.
- [26] A. IGARASHI, P. V. E. MCCLINTOCK, AND N. G. STOCKS, *Velocity spectrum for non-Markovian Brownian motion in a periodic potential*, J. Statist. Phys., 66 (1992), pp. 1059–1070, <https://doi.org/10.1007/BF01055716>.
- [27] A. IGARASHI AND T. MUNAKATA, *Non-Markovian Brownian motion in a periodic potential*, J. Phys. Soc. Jpn., 57 (1988), pp. 2439–2447.
- [28] B. KRAUSKOPF, H. M. OSINGA, AND J. GALÁN-VIOQUE, EDS., *Numerical Continuation Methods for Dynamical Systems*, Understanding Complex Systems, Springer, Dordrecht, 2007, <https://doi.org/10.1007/978-1-4020-6356-5>.
- [29] L. LORENZI AND M. BERTOLDI, *Analytical Methods for Markov Semigroups*, in Pure Appl. Math. (Boca Raton) 283, Chapman & Hall/CRC, Boca Raton, FL, 2007.
- [30] J. LU, Y. LU, AND J. NOLEN, *Scaling Limit of the Stein Variational Gradient Descent Part I: The Mean Field Regime*, preprint, <https://arxiv.org/abs/1805.04035>, 2018.
- [31] E. LUÇON AND W. STANNAT, *Transition from Gaussian to non-Gaussian fluctuations for mean-field diffusions in spatial interaction*, Ann. Appl. Probab., 26 (2016), pp. 3840–3909, <https://doi.org/10.1214/16-AAP1194>.
- [32] P. A. MARKOWICH AND C. VILLANI, *On the trend to equilibrium for the Fokker-Planck equation: an interplay between physics and functional analysis*, VI Workshop on Partial Differential Equations, Part II (Rio de Janeiro, 1999), Mat. Contemp., 19 (2000), pp. 1–29.
- [33] P. MONMARCHÉ, *Long-time behaviour and propagation of chaos for mean field kinetic particles*, Stochastic Process. Appl., 127 (2017), pp. 1721–1737, <https://doi.org/10.1016/j.spa.2016.10.003>.
- [34] S. MOTSCH AND E. TADMOR, *Heterophilous dynamics enhances consensus*, SIAM Rev., 56 (2014), pp. 577–621, <https://doi.org/10.1137/120901866>.
- [35] K. OELSCHLÄGER, *A martingale approach to the law of large numbers for weakly interacting stochastic processes*, Ann. Probab., 12 (1984), pp. 458–479, [www.jstor.org/stable/2243483](http://www.jstor.org/stable/2243483).
- [36] G. A. PAVLIOTIS, *Stochastic Processes and Applications, Diffusion Processes, the Fokker-Planck and Langevin Equations*, Texts Appl. Math. 60, Springer, New York, 2014, <https://doi.org/10.1007/978-1-4939-1323-7>.
- [37] G. A. PAVLIOTIS AND A. M. STUART, *Multiscale Methods, Averaging and Homogenization*, Texts Appl. Math. 53, Springer, New York, 2008.
- [38] R. PINNAU, C. TOTZECK, O. TSE, AND S. MARTIN, *A consensus-based model for global optimization and its mean-field limit*, Math. Models Methods Appl. Sci., 27 (2017), pp. 183–204, <https://doi.org/10.1142/S0218202517400061>.
- [39] G. M. ROTSKOFF AND E. VANDEN-ELJNDEN, *Trainability and Accuracy of Neural Networks: An Interacting Particle System Approach*, preprint, <https://arxiv.org/abs/1805.00915>, 2018.

- [40] J. SHEN, T. TANG, AND L.-L. WANG, *Spectral Methods: Algorithms, Analysis and Applications*, Springer Ser. Comput. Math. 41, Springer, Heidelberg, 2011, <https://doi.org/10.1007/978-3-540-71041-7>.
- [41] M. SHIINO, *Dynamical behavior of stochastic systems of infinitely many coupled nonlinear oscillators exhibiting phase transitions of mean-field type: H theorem on asymptotic approach to equilibrium and critical slowing down of order-parameter fluctuations*, Phys. Rev. A, 36 (1987), pp. 2393–2412, <https://doi.org/10.1103/PhysRevA.36.2393>.
- [42] J. SIRIGNANO AND K. SPILIOPOULOS, *Mean Field Analysis of Neural Networks: A Law of Large Numbers*, preprint, <https://arxiv.org/abs/1805.01053>, 2018.
- [43] J. TUGAUT, *Phase transitions of McKean-Vlasov processes in double-wells landscape*, Stochastics, 86 (2014), pp. 257–284, <https://doi.org/10.1080/17442508.2013.775287>.
- [44] U. VAES, *Topics in Multiscale Modelling: Numerical Analysis and Applications*, Ph.D. thesis, Department of Mathematics, Imperial College London, London, UK, 2019.