World Scientific
www.worldscientific.com

An analytical framework for consensus-based global optimization method

José A. Carrillo*

*Department of Mathematics, Imperial College London,
London SW7 2AZ, UK
carrillo@imperial.ac.uk*

Young-Pil Choi

*Department of Mathematics and
Institute of Applied Mathematics,
Inha University, Incheon 402-751, Republic of Korea
ypchoi@inha.ac.kr*

Claudia Totzeck

*Department of Mathematics,
Technische Universität Kaiserslautern,
Erwin-Schrödinger-Strasse, 67663 Kaiserslautern, Germany
totzeck@mathematik.uni-kl.de*

Oliver Tse

*Department of Mathematics and Computer Science,
Eindhoven University of Technology, P. O. Box 513,
5600MB Eindhoven, The Netherlands
o.t.c.tse@tue.nl*

In this paper, we provide an analytical framework for investigating the efficiency of a consensus-based model for tackling global optimization problems. This work justifies the optimization algorithm in the mean-field sense showing the convergence to the global minimizer for a large class of functions. Theoretical results on consensus estimates are

*Corresponding author

then illustrated by numerical simulations where variants of the method including non-linear diffusion are introduced.

Keywords: Global optimization; opinion dynamics; consensus formation; agent-based models; stochastic dynamics; mean-field limit.

AMS Subject Classification: 35Q83, 35Q91, 35Q93, 37N40, 60H10

1. Introduction

Over the last decades, individual-based models (IBMs) have been widely used in the investigation of complex systems that manifest self-organization or collective behavior. Examples of such complex systems include swarming behavior, crowd dynamics, opinion formation, synchronization, and many more, that are present in the field of mathematical biology, ecology and social dynamics, see for instance Refs. 5, 6, 13, 14, 18, 23, 28, 29, 34, and the references therein.

In the field of global optimization, IBMs may be found in a class of *metaheuristics* (e.g. evolutionary algorithms,[1,4,32] and swarm intelligence[20,26]). They play an increasing role in the design of fast algorithms to provide sufficiently good solutions in tackling hard optimization problems, which includes the traveling salesman problem that is known to be NP hard. Metaheuristics, in general, may be considered as high level concepts for exploring search spaces by using different strategies, chosen in such a way, that a dynamic balance is achieved between the exploitation of the accumulated search experience and the exploration of the search space.[9] Notable metaheuristics for global optimization include, for example, the Ant Colony Optimization, Genetic Algorithms, Particle Swarm Optimization and Simulated Annealing,[25,24] all of which are stochastic in nature.[7] Despite having to stand the test of time, a majority of metaheuristical methods lack the proper justification of its efficacy in the mathematical sense — the universal intent of research in the field is to ascertain whether a given metaheuristic is capable of finding an optimal solution when provided with sufficient information. Due to the stochastic nature of metaheuristics, answers to this question are nontrivial, and they are always probabilistic.

Recently, the use of opinion dynamics and consensus formation in global optimization has been introduced in Ref. 31, where the authors showed substantial numerical and partial analytical evidence of its applicability for solving multi-dimensional optimization problems of the form

$$\min_{x \in \Omega} f(x), \quad \Omega \subset \mathbb{R}^d \text{ a domain,}$$

for a given cost function $f \in \mathcal{C}(\Omega)$ that achieves its global minimum at a unique point in Ω. Without loss of generality, we may assume f to be positive and defined on the whole $\mathbb{R}^d$ by extending it outside Ω without changing its global minimum.

Throughout the paper, we will use the notations $\underline{f} = \inf f$, $\overline{f} = \sup f$, and

$$x_* = \arg\min f, \quad f_* = f(x_*).$$

As we assume to have a unique global minimizer, it holds $f_* = \underline{f}$.

The optimization algorithm involves the use of multiple agents located within the domain Ω to dynamically establish a consensual opinion amongst themselves in finding the global minimizer to the minimization problem, while taking into consideration the opinion of all active agents. First-order models for consensus have been studied within the mathematical community interested in granular materials and swarming that lead to aggregation-diffusion and kinetic equations, which have nontrivial stationary states or flock solutions (cf. Refs. 11, 16, 17, 12, and the references therein). They are also common tools in control engineering to establish consensus in graphs (cf. Refs. 30 and 37).

In order to achieve the goal of optimizing a given function $f(x)$, we consider an interacting stochastic system of $N \in \mathbb{N}$ agents with position $X_t^i \in \mathbb{R}^d$, described by the system of stochastic differential equations

$$dX_t^i = -\lambda(X_t^i - m_t)dt + \sigma|X_t^i - m_t|dW_t^i, \tag{1.1a}$$

$$m_t = \sum_{i=1}^N X_t^i \left(\frac{\omega_f^\alpha(X_t^i)}{\sum_{j=1}^N \omega_f^\alpha(X_t^j)} \right), \tag{1.1b}$$

with $\lambda, \sigma > 0$, where ω_f^α is a weight, which we take as $\omega_f^\alpha(x) = \exp(-\alpha f(x))$ for some appropriately chosen $\alpha > 0$. Notice that (1.1) resembles a geometric Brownian motion, which drifts toward $m_t \in \mathbb{R}^d$. This system is a simplified version of the algorithm introduced in Ref. 31, while keeping the essential ingredients and mathematical difficulties. The first term in (1.1a) imposes a global relaxation toward a position determined by the behavior of the normalized moment given by m_t, while the diffusion term tries to concentrate again around the behavior of m_t. In fact, agents with a position differing a lot from m_t are diffused more. Hence, they explore a larger portion of the landscape of the graph of $f(x)$, while the explorer agents closer to m_t diffuse much less. The normalized moment m_t is expected to dynamically approach the global minimum of the function f, at least when α is large enough, see below. This idea is also used in simulated annealing algorithms.[25,24] The well-posedness of this system will be thoroughly investigated in Sec. 2.

For the solution $X_t^1, \ldots, X_t^N$, $N \in \mathbb{N}$ of the particle system (1.1), we can consider its empirical measure given by

$$\rho_t^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i},$$

where δ_x is the Dirac measure at $x \in \mathbb{R}^d$. Observe that m_t may be re-written in terms of ρ_t^N, i.e.

$$m_t = \frac{1}{\|\omega_f^\alpha\|_{L^1(\rho_t^N)}} \int x\,\omega_f^\alpha d\rho_t^N =: m_f[\rho_t^N].$$

Therefore (1.1a) may be formulated as

$$dX_t^i = -\lambda(X_t^i - m_f[\mu_t^N])dt + \sigma|X_t^i - m_f[\mu_t^N]|dW_t^i,$$

for which we postulate the limiting $(N \to \infty)$ *nonlinear process* $\bar{X}_t$ to satisfy

$$d\bar{X}_t = -\lambda(\bar{X}_t - m_f[\rho_t])dt + \sigma|\bar{X}_t - m_f[\rho_t]|dW_t, \qquad (1.2a)$$

$$m_f[\rho_t] = \int x \, d\eta_t^\alpha, \quad \eta_t^\alpha = \omega_f^\alpha \rho_t / \|\omega_f^\alpha\|_{L^1(\rho_t)}, \quad \rho_t = \mathrm{law}(\bar{X}_t), \qquad (1.2b)$$

subject to the initial condition $\mathrm{law}(\bar{X}_0) = \rho_0$. We call η_t^α the α-*weighted measure*. The measure $\rho_t = \mathrm{law}(\bar{X}_t) \in \mathcal{P}(\mathbb{R}^d)$ is a Borel probability measure, which describes the evolution of a one-particle distribution resulting from the mean-field limit.

The (infinitesimal) generator corresponding to the nonlinear process (1.2a) is given by

$$L\varphi = \kappa\Delta\varphi - \mu \cdot \nabla\varphi, \quad \text{for } \varphi \in \mathcal{C}_c^2(\mathbb{R}^d),$$

with drift and diffusion coefficients

$$\mu_t = \lambda(x - m_f[\rho_t]), \quad \kappa_t = (\sigma^2/2)|x - m_f[\rho_t]|^2,$$

respectively. Therefore, the Fokker–Planck equation associated to (1.2) reads

$$\partial_t \rho_t = \Delta(\kappa_t \rho_t) + \nabla \cdot (\mu_t \rho_t), \quad \lim_{t \to 0} \rho_t = \rho_0, \qquad (1.3)$$

where $\rho_t \in \mathcal{P}(\mathbb{R}^d)$ for $t \geq 0$ satisfies (1.3) in the weak sense. Notice that the Fokker–Planck equation (1.3) is a nonlocal, nonlinear degenerate drift-diffusion equation, which makes its analysis a nontrivial task. Its well-posedness will be the topic of Sec. 3.

We recall from Ref. 31 (cf. Ref. 19), that for any $\rho \in \mathcal{P}(\mathbb{R}^d)$, $\omega_f^\alpha \rho$ satisfies the well-known *Laplace principle*:

$$\lim_{\alpha \to \infty} \left(-\frac{1}{\alpha} \log\left(\int e^{-\alpha f} d\rho \right) \right) = \underline{f}|_{\mathrm{supp}(\rho)} > 0.$$

Therefore, if f attains its minimum at a single point $x_* \in \mathrm{supp}(\rho)$, then the α-weighted measure $\eta_t^\alpha \in \mathcal{P}(\mathbb{R}^d)$ assigns most of its mass to a small region around x_* and hence it approximates a Dirac distribution δ_{x_*} at $x_* \in \mathbb{R}^d$ for large $\alpha \gg 1$. Consequently, the first moment of η_t^α, given by $m_f[\rho]$, provides a good estimate of $x_* = \arg\min f$. Using this fact, we proceed to give justifications for the applicability of the microscopic system (1.1) as a tool for solving global optimization problems, via its mean-field counterpart.

Our main results show in Sec. 4 that mild assumptions on the regularity of the objective function f, $f \in W^{2,\infty}(\mathbb{R}^d)$, one obtains a uniform consensus as the limiting measure $(t \to \infty)$ corresponding to (1.3), i.e.

$$\rho_t \to \delta_{\tilde{x}} \quad \text{as } t \to \infty,$$

for some $\tilde{x} \in \mathbb{R}^d$ possibly depending on the initial density ρ_0. It is also shown that this convergence happens exponentially in time. Moreover, under the same assumptions on f, the point of consensus $\tilde{x}$ may be made arbitrarily close to $x_* = \arg\min f$ by choosing $\alpha \gg 1$ sufficiently large, which is the main goal for global optimization. Our regularity assumptions allow for complicated landscapes of objective functions
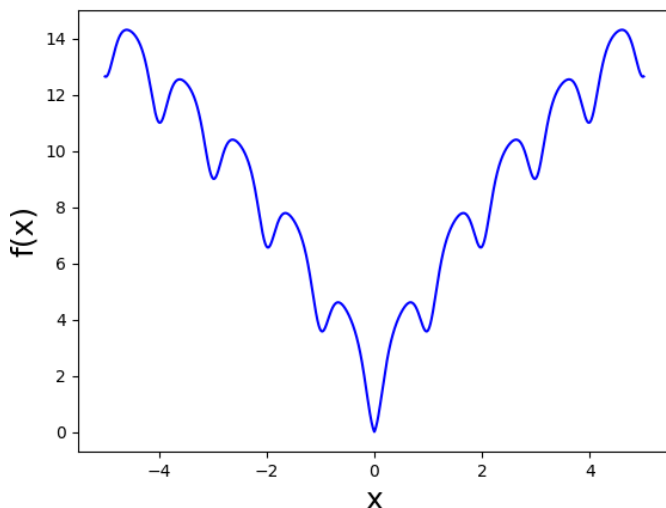
Fig. 1.   The Ackley function is a well-known benchmark for global optimization problems due to its various local minima and the unique global minimum.

with as many local minimizers as you want but with a well defined unique global minimum, see for instance the Ackley function — a well-known benchmark for global optimization problems[3] — used in Ref. 31 and depicted in Fig. 1. *Up to our knowledge, this work shows for the first time the convergence of an agent-based stochastic scheme for global optimization with mild assumptions on the regularity of the cost function.*

We conclude the paper with an extension of the Fokker–Planck equation (1.3) to include nonlinear diffusion of porous medium type and provide numerical evidence for consensus formation in the one-dimensional case. For this reason, we introduce an equivalent formulation of the mean-field equation in terms of the pseudo-inverse distribution $\chi_t(\eta) = \inf\{x \in \mathbb{R} \,|\, \rho_t((-\infty, x]) > \eta\}$. We also compare the microscopic approximation corresponding to the porous medium type Fokker–Planck equation with the original consensus-based microscopic system (1.1) and the proposed algorithm in Ref. 31, showcasing the exponential decay rate of the error in suitable transport distances toward the global minimizer.

2. Well-Posedness of the Microscopic Model

We begin by studying the existence of a unique process $\{\mathbf{X}_t^{(N)} \,|\, t \geq 0\}$ with $\mathbf{X}^{(N)} := (X^{(1,N)}, \ldots, X^{(N,N)})^\top$, satisfying the consensus-based optimization scheme (1.1), and write, for an arbitrary but fixed $N \in \mathbb{N}$, system (1.1) as

$$dX_t^{(N)} = -\lambda \mathbf{F}_N\big(\mathbf{X}_t^{(N)}\big)dt + \sigma \mathbf{M}_N\big(\mathbf{X}_t^{(N)}\big)d\mathbf{W}_t^{(N)}, \tag{2.1}$$

where $\mathbf{W} = (W^{(1,N)}, \ldots, W^{(N,N)})^\top$ is the standard Wiener process in $\mathbb{R}^{Nd}$, and

$$\mathbf{F}_N(\mathbf{X}) = \big(F_N^1(\mathbf{X}), \ldots, F_N^N(\mathbf{X})\big)^\top \in \mathbb{R}^{Nd},$$

where $F_N^i(\mathbf{X}) = \dfrac{\sum_{j \neq i}(X^i - X^j)\,\omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \in \mathbb{R}^d$,

$$\mathbf{M}_N(\mathbf{X}) = \operatorname{diag}\big(|F_N^1(\mathbf{X})|\mathbb{I}_d, \ldots, |F_N^N(\mathbf{X})|\mathbb{I}_d\big) \in \mathbb{R}^{Nd \times Nd}.$$

At this point, we make smoothness assumptions regarding our cost function f.

Assumption 2.1. The cost function $f \colon \mathbb{R}^d \to \mathbb{R}$ is locally Lipschitz continuous.

Under these conditions on f, we easily deduce that F_N^i, $1 \leq i \leq N$, is locally Lipschitz continuous and has linear growth. Consequently, $\mathbf{F}_N$ and $\mathbf{M}_N$ are locally Lipschitz continuous and have linear growth. To be more precise, we obtain the following result.

Lemma 2.1. *Let $N \in \mathbb{N}$, $\alpha, k > 0$ be arbitrary. Then for any $\mathbf{X}, \hat{\mathbf{X}} \in \mathbb{R}^{Nd}$ with $|\mathbf{X}|, |\hat{\mathbf{X}}| \leq k$ and all $i = 1, \ldots, N$ it holds*

$$|F_N^i(\mathbf{X}) - F_N^i(\hat{\mathbf{X}})| \leq |X^i - \hat{X}^i| + \left(1 + \frac{2c_k}{N}\sqrt{N|\hat{X}^i|^2 + |\hat{\mathbf{X}}|^2}\,\right)|\mathbf{X} - \hat{\mathbf{X}}|,$$

$$|F_N^i(\mathbf{X})| \leq |X^i| + |\mathbf{X}|,$$

where $c_k = \alpha\|\nabla f\|_{L^\infty(B_k)}\exp(\alpha\|f - \underline{f}\|_{L^\infty(B_k)})$ and $B_k = \{x \in \mathbb{R}^d \mid |x| \leq k\}$.

Proof. Let $\mathbf{X}, \hat{\mathbf{X}} \in \mathbb{R}^{Nd}$ with $|\mathbf{X}|, |\hat{\mathbf{X}}| \leq k$ for some $k \geq 0$ and $i \in \{1, \ldots, N\}$ be arbitrary. Then

$$F_N^i(\mathbf{X}) - F_N^i(\hat{\mathbf{X}}) = \frac{\sum_{j \neq i}(X^i - X^j)\,\omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} - \frac{\sum_{j \neq i}(\hat{X}^i - \hat{X}^j)\,\omega_f^\alpha(\hat{X}^j)}{\sum_j \omega_f^\alpha(\hat{X}^j)} = \sum_{\ell=1}^3 I_\ell,$$

where the terms I_ℓ, $\ell = 1, 2, 3$, are given by

$$I_1 = \frac{\sum_{j \neq i}(X^i - \hat{X}^i + \hat{X}^j - X^j)\,\omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)},$$

$$I_2 = \frac{\sum_{j \neq i}(\hat{X}^i - \hat{X}^j)\big(\omega_f^\alpha(X^j) - \omega_f^\alpha(\hat{X}^j)\big)}{\sum_j \omega_f^\alpha(X^j)},$$

$$I_3 = \sum_{j \neq i}(\hat{X}^i - \hat{X}^j)\,\omega_f^\alpha(\hat{X}^j)\frac{\sum_j\big(\omega_f^\alpha(\hat{X}^j) - \omega_f^\alpha(X^j)\big)}{\sum_j \omega_f^\alpha(X^j)\sum_j \omega_f^\alpha(\hat{X}^j)},$$

that may easily be estimated by

$$|I_1| \leq |X^i - \hat{X}^i| + |\mathbf{X} - \hat{\mathbf{X}}|,$$

$$|I_2| \leq \frac{\sqrt{2}c_k}{N}|\mathbf{X} - \hat{\mathbf{X}}|\sqrt{N|\hat{X}^i|^2 + |\hat{\mathbf{X}}|^2},$$

$$|I_3| \leq \frac{c_k}{N}|\mathbf{X} - \hat{\mathbf{X}}|\sqrt{N|\hat{X}^i|^2 + |\hat{\mathbf{X}}|^2}.$$

Putting all these terms together yields the required estimate.

As for the estimate of $|F_N^i(\mathbf{X})|$, we easily obtain

$$|F_N^i(\mathbf{X})| = \left| X^i - \frac{\sum_j X^j \, \omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \right| \le |X^i| + |\mathbf{X}|.$$

As i was chosen arbitrary, this concludes the result. $\qquad\square$

Due to Lemma 2.1, we may invoke standard existence results of strong solutions for (2.1).[21]

Theorem 2.1. *For each $N \in \mathbb{N}$, the stochastic differential equation (2.1) has a unique strong solution $\{\mathbf{X}_t^{(N)} \,|\, t \ge 0\}$ for any initial condition $\mathbf{X}_0^{(N)}$ satisfying $\mathbb{E}|\mathbf{X}_0^{(N)}|^2 < \infty$.*

Proof. As mentioned above, we make use of a standard result on existence of a unique strong solution. To this end, we show the existence of a constant $b_N > 0$, such that

$$-2\lambda\mathbf{X} \cdot \mathbf{F}_N(\mathbf{X}) + \sigma^2 \text{trace}(\mathbf{M}_N \mathbf{M}_N^\top)(\mathbf{X}) \le b_N |\mathbf{X}|^2. \qquad (2.2)$$

Indeed, since the following inequalities hold:

$$-X^i \cdot F_N^i(\mathbf{X}) = -X^i \cdot \frac{\sum_{j \ne i}(X^i - X^j)\omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \le -|X^i|^2 + |X^i||\mathbf{X}|,$$

$$|F_N^i(\mathbf{X})|^2 = \left| \frac{\sum_{j \ne i}(X^i - X^j)\omega_f^\alpha(X^j)}{\sum_j \omega_f^\alpha(X^j)} \right|^2 \le 2(|X^i|^2 + |\mathbf{X}|^2),$$

we conclude that

$$-2\lambda\mathbf{X} \cdot \mathbf{F}_N(\mathbf{X}) + \sigma^2 \text{trace}(\mathbf{M}_N \mathbf{M}_N^\top)(\mathbf{X})$$

$$= \sum_i \left(-2\lambda X^i \cdot F_N^i(\mathbf{X}) + d\sigma^2 |F_N^i(\mathbf{X})|^2 \right)$$

$$\le \sum_i 2\lambda \left(-|X^i|^2 + |X^i||\mathbf{X}| \right) + 2d\sigma^2 \left(|X^i|^2 + |\mathbf{X}|^2 \right)$$

$$\le 2(\lambda\sqrt{N} + 2d\sigma^2 N)|\mathbf{X}|^2 =: b_N |\mathbf{X}|^2.$$

Along with the local Lipschitz continuity and linear growth of $\mathbf{F}_N$ and $\mathbf{M}_N$, we obtain the assertion by applying Theorem 3.1 of Ref. 21. $\qquad\square$

Remark 2.1. In fact, the estimate (2.2) yields a uniform bound on the second moment of $\mathbf{X}_t$. Indeed, by application of the Itô formula, we obtain

$$\frac{d}{dt}\mathbb{E}|\mathbf{X}_t^{(N)}|^2 = -2\lambda\mathbb{E}\left[\mathbf{X}_t \cdot \mathbf{F}_N(\mathbf{X}_t^{(N)})\right] + \sigma^2\mathbb{E}\left[\text{trace}(\mathbf{M}_N \mathbf{M}_N^\top)(\mathbf{X}_t^{(N)})\right]$$

$$\le b_N \, \mathbb{E}|\mathbf{X}_t^{(N)}|^2.$$

Therefore, the Gronwall inequality yields

$$\mathbb{E}\big|\mathbf{X}_t^{(N)}\big|^2 \leq e^{b_N t}\mathbb{E}\big|\mathbf{X}_0^{(N)}\big|^2 \quad \text{for all } t \geq 0,$$

i.e. the solution exists globally in time for each fixed $N \in \mathbb{N}$.

Unfortunately, for the mean-field limit ($N \to \infty$) we lose control of the previous bound, since $b_N \to \infty$ as $N \to \infty$. Therefore, we will need a finer moment estimates on $\mathbf{X}^{(N)}$, which we establish at the end of the following section (cf. Lemma 3.4).

3. Well-Posedness of the Mean-Field Equation

In this section, we provide the well-posedness of the nonlocal, nonlinear Fokker–Planck equation (1.3). Since we will be working primarily with Borel probability measures on $\mathbb{R}^d$ with finite second moment, we provide its definition for the readers convenience.

We denote the space of Borel probability measures on $\mathbb{R}^d$ with finite second moment by

$$\mathcal{P}_2(\mathbb{R}^d) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) \text{ such that } \int_{\mathbb{R}^d} |z|^2 \mu(dz) < \infty \right\}.$$

that we equip with the 2-Wasserstein distance W_2 defined by

$$W_2^2(\mu, \hat{\mu}) = \inf\left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |z - \hat{z}|^2 \pi(dz, d\hat{z}), \ \pi \in \Pi(\mu, \hat{\mu}) \right\}, \quad \mu, \hat{\mu} \in \mathcal{P}_2(\mathbb{R}^d),$$

where $\Pi(\mu, \hat{\mu})$ denotes the collection of all Borel probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and $\hat{\mu}$ on the first and second factors, respectively. The set $\Pi(\mu, \hat{\mu})$ is also known as the set of all couplings of μ and $\hat{\mu}$. Equivalently, the Wasserstein distance may be defined by

$$W_2^2(\mu, \hat{\mu}) = \inf \mathbb{E}\big[|Z - \hat{Z}|^2\big],$$

where the infimum is taken over all joint distributions of the random variables Z and $\hat{Z}$ with marginals μ and $\hat{\mu}$, respectively. It is well-known that W_2 defines a metric on $\mathcal{P}_2(\mathbb{R}^d)$. Since $\mathbb{R}^d$ is a separable complete metric space and p a positive number, the metric space $(\mathcal{P}_p(\mathbb{R}^d), W_p)$ is separable and complete,[10] in particular this yields $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is Polish. We remind the reader that a Polish space is a separable completely metrizable topological space. With each point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and every $\epsilon > 0$ we associate an ϵ-ball $U_\epsilon(\mu) = \{\nu \in \mathcal{P}_2(\mathbb{R}^d) : W_2(\mu, \nu) < \epsilon\}$. The ϵ-balls form the basis of a topology, called the topology of the metric space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$. If this topology agrees with a given topology $\mathcal{O}$ on $\mathcal{P}_2(\mathbb{R}^d)$, we say that the metric W_2 is compatible with the topology $\mathcal{O}$. If for a topological space $(X, \mathcal{O})$ there exists a compatible metric, we say $(X, \mathcal{O})$ is metrizable or that the metric metrizes the topological space $(X, \mathcal{O})$. For $\mathbb{R}^d$ and $p \in [1, \infty)$ it is well-known that the Wasserstein distance W_p is compatible with the weak topology in $\mathcal{P}_p(\mathbb{R}^d)$. Moreover, convergence in W_p implies convergence of the first p moments (see Chap. 6 of Ref. 36 for more details). Altogether, W_2 metrizes the weak convergence in $\mathcal{P}_2(\mathbb{R}^d)$ and

convergence in W_2 implies convergence of the first two moments. Moreover, $\mathcal{P}_2(\mathbb{R}^d)$ with the weak topology is Polish.

We split the results of this section into two parts, based on additional assumptions on f. We begin our investigation with the easier of the two, which also provides the means to prove the other. Throughout this section, we assume that f satisfies the following assumptions:

Assumption 3.1.

(1) The cost function $f\colon \mathbb{R}^d \to \mathbb{R}$ is bounded from below with $\underline{f} := \inf f$.
(2) There exist constants L_f and $c_u > 0$ such that

$$\begin{cases} |f(x) - f(y)| \leq L_f(|x| + |y|)|x - y| & \text{for all } x, y \in \mathbb{R}^d, \\ f(x) - \underline{f} \leq c_u(1 + |x|^2) & \text{for all } x \in \mathbb{R}^d. \end{cases} \tag{A1}$$

3.1. *Bounded cost functions*

In addition to Assumption 3.1 we consider cost functions f that are bounded from above. In particular, f has the upper bound $\overline{f} := \sup f$.

The main result of this section is provided by the following theorem.

Theorem 3.1. *Let f satisfy Assumption 3.1 and be bounded, and $\rho_0 \in \mathcal{P}_4(\mathbb{R}^d)$. Then there exists a unique nonlinear process $\bar{X} \in \mathcal{C}([0,T], \mathbb{R}^d)$, $T > 0$, satisfying*

$$d\bar{X}_t = -\lambda(\bar{X}_t - m_f[\rho_t])dt + \sigma|\bar{X}_t - m_f[\rho_t]|dW_t, \quad \rho_t = \text{law}(\bar{X}_t),$$

in the strong sense, and $\rho \in \mathcal{C}([0,T], \mathcal{P}_2(\mathbb{R}^d))$ satisfies the corresponding Fokker–Planck equation (1.3) *(in the weak sense) with $\lim_{t \to 0} \rho_t = \rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$.*

Before we prove the theorem, we discuss two results that not only facilitate the proof of Theorem 3.1, but are also interesting in their own right.

Lemma 3.1. *Let f satisfy Assumption 3.1 and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ with $\int |x|^2 d\mu \leq K$. Then*

$$\frac{e^{-\alpha \underline{f}}}{\|\omega_f^\alpha\|_{L^1(\mu)}} \leq \exp(\alpha c_u(1 + K)) =: c_K.$$

Proof. The proof follows from the Jensen inequality, which gives

$$e^{-\alpha \int (f - \underline{f}) d\mu} \leq \int e^{-\alpha(f - \underline{f})} d\mu.$$

A simple rearrangement of the previous inequality and using (A1) yields the required estimate. $\qquad \square$

Lemma 3.2. *Let f satisfy Assumption 3.1 and $\mu, \hat{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$ with*

$$\int |x|^4 d\mu, \quad \int |\hat{x}|^4 d\hat{\mu} \leq K.$$

Then the following stability estimate holds

$$|m_f[\mu] - m_f[\hat{\mu}]| \leq c_0 W_2(\mu, \hat{\mu}),$$

for a constant $c_0 > 0$ depending only on α, L_f and K.

Proof. Taking the difference, we obtain

$$m_f[\mu] - m_f[\hat{\mu}] = \iint \left[\frac{x \, \omega_f^\alpha(x)}{\|\omega_f^\alpha\|_{L^1(\mu)}} - \frac{\hat{x} \, \omega_f^\alpha(\hat{x})}{\|\omega_f^\alpha\|_{L^1(\hat{\mu})}} \right] d\pi =: \iint \big(h(x) - h(\hat{x}) \big) d\pi,$$

where $\pi \in \Pi(\mu, \hat{\mu})$ is an arbitrary coupling of μ and $\hat{\mu}$. Observe that the integrand on the right may be written as

$$h(x) - h(\hat{x}) = \frac{(x - \hat{x}) \, \omega_f^\alpha(\hat{x})}{\|\omega_f^\alpha\|_{L^1(\mu)}} + \frac{x(\omega_f^\alpha(x) - \omega_f^\alpha(\hat{x}))}{\|\omega_f^\alpha\|_{L^1(\mu)}}$$

$$+ \frac{\iint \big(\omega_f^\alpha(\hat{x}) - \omega_f^\alpha(x) \big) d\pi_t}{\|\omega_f^\alpha\|_{L^1(\mu)} \|\omega_f^\alpha\|_{L^1(\hat{\mu})}} \, \hat{x} \, \omega_f^\alpha(\hat{x}).$$

Under the assumption (A1) on f and Lemma 3.1, the terms may be estimated by

$$|h(x) - h(\hat{x})| \leq c_K |x - \hat{x}| + c_K \alpha L_f |x|(|x| + |\hat{x}|)|x - \hat{x}|$$

$$+ c_K^2 \alpha L_f |\hat{x}| \iint (|x| + |\hat{x}|)|x - \hat{x}| d\pi.$$

Using this estimate, we further obtain

$$|m_f[\mu] - m_f[\hat{\mu}]| \leq c_K \iint (1 + \alpha L_f |x|(|x| + |\hat{x}|))|x - \hat{x}| d\pi$$

$$+ c_K^2 \alpha L_f \left(\int |\hat{x}| d\hat{\mu} \right) \iint (|x| + |\hat{x}|)|x - \hat{x}| d\pi$$

$$\leq c_K (1 + \alpha L_f (1 + c_K) p_K) \left(\iint |x - \hat{x}|^2 d\pi \right)^{\frac{1}{2}},$$

where we applied the Hölder inequality in the second line and p_K is a polynomial in K. Finally, optimizing over all couplings π concludes the proof. $\square$

Proof of Theorem 3.1. Step 1: For some given $u \in \mathcal{C}([0, T], \mathbb{R}^d)$, we may uniquely solve the SDE

$$dY_t = -\lambda(Y_t - u_t)dt + \sigma|Y_t - u_t|dW_t, \quad \text{law}(Y_0) = \rho_0, \tag{3.1}$$

for some fixed initial measure $\rho_0 \in \mathcal{P}_4(\mathbb{R}^d)$, which induces $\nu_t = \text{law}(Y_t)$. Since $Y \in \mathcal{C}([0, T], \mathbb{R}^d)$, we obtain $\nu \in \mathcal{C}([0, T], \mathcal{P}_2(\mathbb{R}^d))$, which satisfies the following

Fokker–Planck equation

$$\frac{d}{dt}\int \varphi\, d\nu_t = \int \left((\sigma^2/2)|x-u_t|^2\Delta\varphi - \lambda(x-u_t)\cdot\nabla\varphi\right)d\nu_t, \tag{3.2}$$

for all $\varphi \in \mathcal{C}_b^2(\mathbb{R}^d)$. Setting $m_f[\nu] \in \mathcal{C}([0,T],\mathbb{R}^d)$ provides the self-mapping property of the map

$$\mathcal{T}\colon \mathcal{C}([0,T],\mathbb{R}^d) \to \mathcal{C}([0,T],\mathbb{R}^d); \quad u \mapsto \mathcal{T}u = m_f[\nu],$$

for which we show to be compact.

Step 2: Since $\rho_0 \in \mathcal{P}_4(\mathbb{R}^d)$, standard theory of SDEs (see e.g. Chap. 7 of Ref. 2) provides a fourth-order moment estimate for solutions to (3.1) of the form

$$\mathbb{E}|Y_t|^4 \leq (1+\mathbb{E}|Y_0|^4)e^{ct},$$

for some constant $c > 0$. In particular, $\sup_{t\in[0,T]}\int |x|^4 d\nu_t \leq K$ for some $K < \infty$. On the other hand, for any $t > s$, $t,s \in (0,T)$, the Itô isometry yields

$$\mathbb{E}|Y_t - Y_s|^2 \leq 2\lambda^2|t-s|\mathbb{E}\int_s^t |Y_\tau - u_\tau|^2 d\tau + 2\sigma^2\mathbb{E}\int_s^t |Y_\tau - u_\tau|^2 d\tau$$

$$\leq 4(\lambda^2 T + \sigma^2)(K + \|u\|_\infty^2)|t-s| =: c\,|t-s|,$$

and therefore, $W_2(\nu_t,\nu_s) \leq c\,|t-s|^{\frac{1}{2}}$, for some constant $c > 0$. Applying Lemma 3.2 with $\mu = \nu_t$ and $\hat{\mu} = \nu_s$, we obtain

$$|m_f[\nu_t] - m_f[\nu_s]| \leq c_0 W_2(\nu_t,\nu_s) \leq c_0\, c\,|t-s|^{\frac{1}{2}},$$

which provides the Hölder continuity of $t \mapsto m_f[\nu_t]$ with exponent $1/2$, and thereby the compactness of $\mathcal{T}$ due to the compact embedding $\mathcal{C}^{0,1/2}([0,T],\mathbb{R}^d) \hookrightarrow \mathcal{C}([0,T],\mathbb{R}^d)$.

Step 3: Now let $u \in \mathcal{C}([0,T],\mathbb{R}^d)$ satisfy $u = \tau\mathcal{T}u$ for $\tau \in [0,1]$. In particular, there exists $\rho \in \mathcal{C}([0,T],\mathcal{P}_2(\mathbb{R}^d))$ satisfying (3.2) such that $u = \tau m_f[\rho]$. Due to the boundedness assumption on f, we have for all $t \in (0,T)$ that

$$|u_t|^2 = \tau^2|m_f[\rho_t]|^2 \leq \tau^2 e^{\alpha(\overline{f}-\underline{f})}\int |x|^2 d\rho_t. \tag{3.3}$$

Therefore, a computation of the second moment gives

$$\frac{d}{dt}\int |x|^2 d\rho_t = \int (d\sigma^2|x-u_t|^2 - 2\lambda(x-u_t)\cdot x)d\rho_t$$

$$= \int ((d\sigma^2 - 2\lambda)|x|^2 + 2\gamma\,x\cdot u_t + d\sigma^2|u_t|^2)d\rho_t$$

$$\leq (d\sigma^2 - 2\lambda + |\gamma|)\int |x|^2 d\rho_t + (d\sigma^2 + |\gamma|)|u_t|^2 \leq c_\lambda \int |x|^2 d\rho_t,$$

where $\gamma = \lambda - d\sigma^2$ and $c_\lambda = (d\sigma^2 + |\gamma|)(1 + e^{\alpha(\overline{f}-\underline{f})})$. From Gronwall's inequality we easily deduce

$$\int |x|^2 d\rho_t \leq \left(\int |x|^2 d\rho_0\right)e^{c_\lambda t},$$

and consequently also an estimate for $\|u\|_\infty$ via (3.3). In particular, there is a constant $q > 0$ for which $\|u\|_\infty < q$. We conclude the proof by applying the Leray–Schauder fixed point theorem, [22, Chap. 11] which provides a fixed point u for the mapping $\mathcal{T}$ and thereby a solution of (1.2) (respectively (1.3)).

Step 4: As for uniqueness, we first note that a fixed point u of $\mathcal{T}$ satisfies $\|u\|_\infty < q$. Hence, the fourth-order moment estimate provided in **Step 2** holds and $\sup_{t \in [0,T]} \int |x|^4 d\rho_t \leq K < \infty$. Now suppose we have two fixed points u and $\hat{u}$ with

$$\|u\|_\infty, \quad \|\hat{u}\|_\infty < q, \quad \sup_{t \in [0,T]} \int |x|^4 d\rho_t, \quad \sup_{t \in [0,T]} \int |x|^4 d\hat{\rho}_t \leq K,$$

and their corresponding processes Y_t, $\hat{Y}_t$ satisfying (3.1), respectively. Then taking the difference $z_t := Y_t - \hat{Y}_t$ for the same Brownian path gives

$$z_t = z_0 - \lambda \int_0^t z_s \, ds + \lambda \int_0^t (u_s - \hat{u}_s) ds + \sigma \int_0^t (|Y_s - u_s| - |\hat{Y}_s - \hat{u}_s|) dW_s.$$

Squaring on both sides, taking the expectation and applying the Itô isometry yields

$$\mathbb{E}|z_t|^2 \leq 2\mathbb{E}|z_0|^2 + 8(\lambda^2 t + \sigma^2) \int_0^t \mathbb{E}|z_s|^2 ds + 4\lambda^2 t \int_0^t |m_f[\rho_t] - m_f[\hat{\rho}_t]|^2 ds.$$

Since Lemma 3.2 provides the estimate

$$|m_f[\rho_t] - m_f[\hat{\rho}_t]| \leq c_0 W_2(\rho_t, \hat{\rho}_t) \leq c_0 \sqrt{\mathbb{E}|z_t|^2}, \tag{3.4}$$

we further obtain

$$\mathbb{E}|z_t|^2 \leq 2\mathbb{E}|z_0|^2 + 4\big((2 + c_0^2)\lambda^2 t + 2\sigma^2\big) \int_0^t \mathbb{E}|z_s|^2 ds.$$

Therefore, applying Gronwall's inequality and using the fact that $\mathbb{E}|z_0|^2 = 0$ yields $\mathbb{E}|z_t|^2 = 0$ for all $t \in [0,T]$. In particular, $\|u - \hat{u}\|_\infty = 0$, i.e. $u \equiv \hat{u}$ due to (3.4). $\qquad\square$

3.2. *Cost functions with quadratic growth at infinity*

In this subsection, we allow for cost functions that have quadratic growth at infinity. More precisely, we suppose the following:

There exist constants $M > 0$ and $c_l > 0$ such that

$$f(x) - \underline{f} \geq c_l |x|^2 \quad \text{for } |x| \geq M. \tag{A2}$$

We provide a similar result as in the boundedness case under the assumption (A2).

Theorem 3.2. *Let f satisfy Assumption* 3.1 *and quadratic growth* (A2), *and $\rho_0 \in \mathcal{P}_4(\mathbb{R}^d)$. Then there exists a unique nonlinear process $\bar{X} \in \mathcal{C}([0,T], \mathbb{R}^d)$, $T > 0$, satisfying*

$$d\bar{X}_t = -\lambda(\bar{X}_t - m_f[\rho_t])dt + \sigma|\bar{X}_t - m_f[\rho_t]|dW_t, \quad \rho_t = \text{law}(\bar{X}_t),$$

in the strong sense, and $\rho \in \mathcal{C}([0,T], \mathcal{P}_2(\mathbb{R}^d))$ satisfies the corresponding Fokker–Planck equation (1.3) *(in the weak sense) with $\lim_{t \to 0} \rho_t = \rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$.*

Proof. The proof is a slight modification of **Step 3** in the proof of Theorem 3.1. Since **Steps 1**, **2** and **4** remain the same, we only show **Step 3**.

Step 3: Let $u \in \mathcal{C}([0,T], \mathbb{R}^d)$ satisfy $u = \tau \mathcal{T} u$ for $\tau \in [0,1]$, i.e. there exists $\rho \in \mathcal{C}([0,T], \mathcal{P}_2(\mathbb{R}^d))$ satisfying (3.2) such that $u = \tau m_f[\rho]$. Due to Lemma 3.3 below, we have that

$$|u_t|^2 = \tau^2 |m_f[\rho_t]|^2 \leq \tau^2 \int |x|^2 d\eta_t^\alpha \leq \tau^2 \left(b_1 + b_2 \int |x|^2 d\rho_t \right), \qquad (3.5)$$

for the constants b_1 and b_2 given in (3.6). Therefore, a similar computation of the second moment estimate as above gives

$$\frac{d}{dt} \int |x|^2 d\rho_t \leq (d\sigma^2 - 2\lambda + |\gamma|) \int |x|^2 d\rho_t + \tau(d\sigma^2 + |\gamma|) \left(b_1 + b_2 \int |x|^2 d\rho \right)$$

$$\leq (d\sigma^2 + |\gamma|)b_1 + (d\sigma^2 + |\gamma|)(1 + b_2) \int |x|^2 d\rho_t,$$

which by Gronwall's inequality yields

$$\int |x|^2 d\rho_t \leq e^{c_\lambda t} \int |x|^2 d\rho_0 + \frac{b_1}{1 + b_2} \left(e^{c_\lambda t} - 1 \right),$$

with $c_\lambda = (d\sigma^2 + |\gamma|)(1 + b_2)$, and consequently also an estimate for $\|u\|_\infty$ via (3.5). In particular, there is a constant $q > 0$ for which $\|u\|_\infty < q$. We conclude the proof by using the same argument as in **Step 3** in the proof of Theorem 3.1. $\qquad \square$

Lemma 3.3. *Let f satisfy Assumption 3.1 and (A2), and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Then*

$$\int |x|^2 d\eta^\alpha \leq b_1 + b_2 \int |x|^2 d\mu, \quad \eta^\alpha = \omega_f^\alpha \, \mu / \|\omega_f^\alpha\|_{L^1(\mu)} \qquad (3.6)$$

with constants

$$b_1 = M^2 + b_2, \quad b_2 = 2\frac{c_u}{c_l} \left(1 + \frac{1}{\alpha c_l} \frac{1}{M^2} \right),$$

depending only on M, c_u and c_l.

Proof. We begin by looking at $\eta^\alpha(\{f - f_* \geq k\})$ for some $k > 0$. A simple computation gives

$$\eta^\alpha(\{f - f_* \geq k\}) = \frac{1}{\|\omega_f^\alpha\|_{L^1(\mu)}} \int_{\{f - f_* \geq k\}} e^{-\alpha f} d\mu \leq \frac{e^{-\alpha(f_* + k)}}{\|\omega_f^\alpha\|_{L^1(\mu)}} \mu(\{f - f_* \geq k\}).$$

On the other hand,

$$\|\omega_f^\alpha\|_{L^1(\mu)} = \int e^{-\alpha f} \, d\mu \geq \mu(\{f - f_* < \ell\}) e^{-\alpha(f_* + \ell)},$$

for any $\ell > 0$. Consequently, we obtain our first estimate

$$\eta^\alpha(\{f - f_* \geq k\}) \leq e^{-\alpha(k - \ell)} \frac{\mu(\{f - f_* \geq k\})}{\mu(\{f - f_* < \ell\})}, \qquad (3.7)$$

which holds for any $k, \ell > 0$, as long as the terms on the right-hand side are finite.

Now let us estimate the second moment of η^α. For some $R_0 \geq M$ to be determined later,

$$\int |x|^2 d\eta^\alpha \leq R_0^2 \eta^\alpha(B_{R_0}) + \int_{B_{R_0}^c} |x|^2 d\eta^\alpha = R_0^2 + \int_{B_{R_0}^c} \left(|x|^2 - R_0^2\right) d\eta^\alpha$$

$$= R_0^2 + \sum_{i=1}^\infty \int_{A_i} \left(|x|^2 - R_0^2\right) d\eta^\alpha \leq R_0^2 + R_0^2 \sum_{i=1}^\infty \left(2^i - 1\right) \eta^\alpha(A_i),$$

where $A_i := \{2^{(i-1)/2} R_0 \leq |x| < 2^{i/2} R_0\}$, $i \in \mathbb{N}$. However, since (A2) tells us that $f - f_* \geq c_l |x|^2$ for all $|x| \geq R_0 \geq M$, we find that $A_i \subset \{f - f_* \geq c_l 2^{i-1} R_0^2\}$ and therefore $\eta^\alpha(A_i) \leq \eta^\alpha(\{f - f_* \geq c_l 2^{i-1} R_0^2\})$ for all $i \in \mathbb{N}$. From estimate (3.7) and the choice $\ell = c_l 2^{i-2} R_0^2$, we further obtain

$$\int |x|^2 d\eta^\alpha \leq R_0^2 + R_0^2 \sum_{i=1}^\infty \left(2^i - 1\right) \eta^\alpha\left(\{f - f_* \geq c_l 2^{i-1} R_0^2\}\right)$$

$$\leq R_0^2 + R_0^2 \sum_{i=1}^\infty \left(2^i - 1\right) e^{-\alpha c_l R_0^2 2^{i-2}} \frac{\mu\left(\{f - f_* \geq c_l 2^{i-1} R_0^2\}\right)}{\mu\left(\{f - f_* < c_l 2^{i-2} R_0^2\}\right)}. \tag{3.8}$$

We are now left with estimating the last term on the right.

The Markov inequality and (A1) yields

$$\mu(\{f - f_* \geq k\}) \leq \frac{1}{k} \int |f - f_*| d\mu \leq \frac{1}{k} c_u \left(1 + \int |x|^2 d\mu\right) =: \frac{c_u}{k} m,$$

for any $k > 0$. Furthermore, since

$$\mu(\{f - f_* < k\}) = 1 - \mu(\{f - f_* \geq k\}) \geq 1 - \frac{c_u}{k} m,$$

we have that

$$\frac{\mu(\{f - f_* \geq k\})}{\mu(\{f - f_* < k/2\})} \leq \frac{1}{2} \frac{m}{k/(2c_u) - m}.$$

Choosing $k = c_l 2^{i-1} R_0^2$, we obtain

$$\frac{\mu\left(\{f - f_* \geq c_l 2^{i-1} R_0^2\}\right)}{\mu\left(\{f - f_* < c_l 2^{i-2} R_0^2\}\right)} \leq \frac{1}{2} \frac{m}{(c_l/c_u) 2^{i-2} R_0^2 - m}.$$

At this point, we can choose

$$R_0^2 = M^2 + 2 \frac{c_u}{c_l} m > M^2,$$

to further obtain

$$\frac{\mu\left(\{f - f_* \geq c_l 2^{i-1} R_0^2\}\right)}{\mu\left(\{f - f_* < c_l 2^{i-2} R_0^2\}\right)} \leq \frac{1}{2} \frac{m}{(c_l/c_u) 2^{i-2} M^2 + (2^{i-1} - 1) m} \leq 2 \frac{c_u}{c_l} \frac{m}{2^i M^2}.$$

Inserting this into (3.8) yields

$$\int |x|^2 d\eta^\alpha \leq R_0^2 + 2 \frac{c_u}{c_l} \frac{m}{M^2} R_0^2 \sum_{i=1}^\infty \left(1 - 2^{-i}\right) e^{-\alpha c_l R_0^2 2^{i-2}}.$$

A coarse estimate of the summation on the right may be given by

$$\sum_{i=1}^{\infty}\left(1-2^{-i}\right)e^{-\alpha c_l R_0^2 2^{i-2}} \leq \sum_{i=1}^{\infty} e^{-\alpha c_l R_0^2 2^{i-2}} \leq \int_{\frac{1}{2}}^{\infty} e^{-\alpha c_l R_0^2 s}\,ds \leq \frac{1}{\alpha c_l R_0^2},$$

which finally gives an estimate of the form

$$\int |x|^2 d\eta^{\alpha} \leq M^2 + 2\frac{c_u}{c_l}\left(1+\frac{1}{\alpha c_l}\frac{1}{M^2}\right)\left(1+\int |x|^2 d\mu\right) =: b_1 + b_2 \int |x|^2 d\mu,$$

thereby concluding the proof. $\qquad\square$

Remark 3.1. Observe that the estimate (3.6) provided in Lemma 3.3 does not blow up as $\alpha \to \infty$, in contrast to estimate (3.3) used in the proof of the bounded case. In fact, the constant b_2 may be chosen to be independent of α for $\alpha \geq 1$.

3.3. *Moment estimates for the stochastic empirical measure*

For the solution $\mathbf{X}^{(N)} \in \mathcal{C}([0,T],\mathbb{R}^d)^N$, $N \in \mathbb{N}$ of the particle system (1.1), we denote by

$$\rho_t^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^{(i,N)}} \in \mathcal{P}_2(\mathbb{R}^d),$$

the empirical measure corresponding to $\mathbf{X}_t^{(N)}$ for every $t \in [0,T]$, where δ_x is the Dirac measure at $x \in \mathbb{R}^d$. Note that since $\mathbf{X}_t^{(N)}$ is a random variable, so is its empirical measure ρ_t^N.

Lemma 3.4. *Let f satisfy Assumption 3.1 and either (1) boundedness, or (2) quadratic growth at infinity (A2), and $\rho_0 \in \mathcal{P}_{2p}(\mathbb{R}^d)$, $p \geq 1$. Further, let $\mathbf{X}^{(N)}$, $N \in \mathbb{N}$ be the solution of the particle system (1.1) with $\rho_0^{\otimes N}$-distributed initial data $\mathbf{X}_0^{(N)}$ and ρ^N the corresponding empirical measure. Then there exists a constant $K > 0$, independent of N, such that*

$$\sup_{t\in[0,T]} \mathbb{E}\int |x|^{2p}d\rho_t^N, \quad \sup_{t\in[0,T]} \mathbb{E}|m_t^N|^{2p} \leq K, \tag{3.9}$$

and consequently also the estimates

$$\sup_{t\in[0,T]} \mathbb{E}\int |x|^2 d\eta_t^{\alpha,N}, \quad \sup_{t\in[0,T]} \mathbb{E}|X_t^{(i,N)}|^{2p} \leq K \quad \text{for } i = 1,\dots,N,$$

for the same constant K.

Proof. Let $\mathbf{X}^{(N)}$, $N \in \mathbb{N}$ be a solution of the particle system (1.1). Using the inequality $(a+b)^q \leq 2^{q-1}(a^q+b^q)$, $q \geq 1$ and the Itô isometry, it is easy to see that

the estimate

$$\mathbb{E}\big|X_t^{(i,N)}\big|^{2p} = 2^{2p-1}\mathbb{E}\big|X_0^{(i,N)}\big|^{2p}$$

$$+ 2^{3(2p-1)}(\lambda^{2p}T^p + \sigma^{2p})T^{p-1}\int_0^t \mathbb{E}\big|X_s^{(i,N)}\big|^{2p} + \mathbb{E}\big|m_s^N\big|^{2p}ds, \quad (3.10)$$

holds for each $i = 1, \ldots, N$. Summing the previous inequality over $i = 1, \ldots, N$, dividing by N and using the linearity of the expectation gives

$$\mathbb{E}\int |x|^{2p}d\rho_t^N \le 2^{2p-1}\mathbb{E}\int |x|^{2p}d\rho_0^N$$

$$+ 2^{3(2p-1)}(\lambda^{2p}T^p + \sigma^{2p})T^{p-1}\int_0^t \left[\mathbb{E}\int |x|^{2p}d\rho_s^N + \mathbb{E}\big|m_s^N\big|^{2p}\right]ds. \tag{3.11}$$

Following the strategy given in Sec. 3, we obtain for cost functions f satisfying Assumption 3.1 and either boundedness or (A2), the estimate

$$|m_s^N|^2 \le \int |x|^2 d\eta_s^{\alpha,N} \le c_1 + c_2 \int |x|^2 d\rho_s^N, \tag{3.12}$$

for appropriate constants c_1 and c_2, independent of N, where by construction $m_t^N = \int x \, d\eta_t^{\alpha,N}$ with $\eta_t^{\alpha,N} = \omega_f^\alpha \rho_t^N / \|\omega_f^\alpha\|_{L^1(\rho_t^N)} \in \mathcal{P}(\mathbb{R}^d)$. Therefore, we further obtain

$$|m_s^N|^{2p} \le \left(c_1 + c_2 \int |x|^2 d\rho_s^N\right)^p \le 2^{p-1}\left(c_1^p + c_2^p \int |x|^{2p}d\rho_s^N\right).$$

Inserting this into (3.11) and applying the Gronwall inequality provides a constant $K_p > 0$, independent of N, such that $\sup_{t\in[0,T]} \mathbb{E}\int |x|^{2p}d\rho_t^N \le K_p$ holds, and consequently also

$$\sup_{t\in[0,T]} \mathbb{E}|m_t^N|^{2p} \le 2^{p-1}(c_1^p + c_2^p K_p),$$

which concludes the proof of the estimates in (3.9) by choosing K sufficiently large. The other two estimates easily follow by (3.12) and by applying the Grownwall inequality on (3.10), respectively. $\qquad\square$

Remark 3.2. Despite having the estimates provided in Lemma 3.4, we are unable to prove a mean-field limit result for the interacting particle system (1.1) toward the nonlinear process (1.2) by means of classical tools (see e.g. Ref. 33). At this moment, we only postulate its corresponding mean-field equation (1.3), for which the numerical simulations have indicated to be true.

4. Large Time Behavior and Consensus Formation

We finally arrive at the most important part of the paper, i.e. we provide sufficient conditions on f such that *uniform consensus formation* happens. More precisely,

we say that uniform consensus occurs when

$$\rho_t \to \delta_{\tilde{x}} \quad \text{as } t \to \infty,$$

for some $\tilde{x} \in \mathbb{R}^d$ possibly depending on ρ_0.

In fact, in the framework of global optimization, we would like to further have that $\tilde{x} = x_* = \arg\min f$. In other words, we want that ρ_t concentrates at the global minimum of f. Unlike, the deterministic case, the formation of nonuniform consensus, i.e. multiple opinions in the limit $t \to \infty$, in the stochastic model cannot occur.[31] Hence, it is expected that uniform consensus is formed, whenever concentration happens. We will see that this is the case.

Assumption 4.1. Throughout this section we will assume that $f \in \mathcal{C}^2(\mathbb{R}^d)$ satisfies additionally

(1) $\inf f > 0$.
(2) $\|\nabla^2 f\|_\infty \le c_f$ and there exists constants $c_0, c_1 > 0$, such that

$$\Delta f \le c_0 + c_1 |\nabla f|^2 \quad \text{in } \mathbb{R}^d.$$

Let us point out that from the viewpoint of global optimization we can change the value of the function f outside a large ball in a way that condition (2) of Assumption 4.1 is satisfied as soon as it does not change the point where the global minimum is achieved. These cases include interesting functions in applications.

4.1. *Concentration estimate*

Let us begin by denoting the expectation and variance of ρ_t by

$$E(\rho_t) := \int x \, d\rho_t \quad \text{and} \quad V(\rho_t) := \frac{1}{2} \int |x - E(\rho_t)|^2 \, d\rho_t.$$

A simple computation of the evolution of its variance gives

$$\frac{d}{dt} V(\rho_t) = \frac{d}{dt} \frac{1}{2} \int |x - E(\rho_t)|^2 \, d\rho_t = -2\lambda V(\rho_t) + (d\sigma^2/2) \int |x - m_f[\rho_t]|^2 d\rho_t.$$

To estimate the last term on the right, we apply Jensen's inequality to obtain

$$\int |x - m_f[\rho_t]|^2 d\rho_t \le \frac{\iint |x - y|^2 \omega_f^\alpha d\rho_t(x) d\rho_t(y)}{\|\omega_f^\alpha\|_{L^1(\rho_t)}} \le 2 \frac{e^{-\alpha \underline{f}}}{\|\omega_f^\alpha\|_{L^1(\rho_t)}} V(\rho_t), \qquad (4.1)$$

and therefore

$$\frac{d}{dt} V(\rho_t) \le -\left(2\lambda - d\sigma^2 e^{-\alpha \underline{f}} / \|\omega_f^\alpha\|_{L^1(\rho_t)}\right) V(\rho_t). \qquad (4.2)$$

Remark 4.1. If f were bounded, then an obvious way to obtain concentration is to use the estimate $\|\omega_f^\alpha\|_{L^1(\rho_t)} \ge e^{-\alpha \overline{f}}$ to obtain

$$\frac{d}{dt} V(\rho_t) \le -\left(2\lambda - d\sigma^2 e^{\alpha(\overline{f} - \underline{f})}\right) V(\rho_t).$$

However, since $\overline{f} > \underline{f}$, we have that $e^{\alpha(\overline{f} - \underline{f})} \to \infty$ as $\alpha \to \infty$, and we would have to choose $\lambda \gg 1$ sufficiently large in order to obtain concentration. In the next

subsection, we will see that this is not desirable since α has to be chosen large to have a good approximation of the global minimizer.

In order to understand how $\|\omega_f^\alpha\|_{L^1(\rho_t)}$ evolves, we study its evolution given by

$$\frac{d}{dt}\|\omega_f^\alpha\|_{L^1(\rho_t)} = \alpha\lambda \int (\nabla f(x) - \nabla f(m_f[\rho_t])) \cdot (x - m_f[\rho_t])\omega_f^\alpha d\rho_t$$

$$+ \frac{\sigma^2}{2} \int \left(\alpha^2 |\nabla f(x)|^2 - \alpha\Delta f(x)\right)|x - m_f[\rho_t]|^2\omega_f^\alpha d\rho_t =: I_1 + I_2,$$

where we used the fact that

$$\int (x - m_f[\rho_t])\omega_f^\alpha d\rho_t = 0,$$

for the first term.

Lemma 4.1. *Under Assumption 4.1 for f and $\alpha \geq c_1$, we have*

$$\frac{d}{dt}\|\omega_f^\alpha\|_{L^1(\rho_t)}^2 \geq -b_1 V(\rho_t), \tag{4.3}$$

with $b_1 = b_1(\alpha, \lambda, \sigma) = 2\alpha e^{-2\alpha\underline{f}}(c_0\sigma^2 + 2\lambda c_f)$.
Note that from Assumption 4.1, we see that $b_1 \to 0$ as $\alpha \to \infty$.

Proof. From the assumptions on f, we obtain the following estimates

$$I_1 \geq -\alpha\lambda c_f e^{-\alpha\underline{f}} \int |x - m_f[\rho_t]|^2 d\rho_t$$

$$I_2 \geq \alpha(\alpha - c_1)\frac{\sigma^2}{2} \int |\nabla f(x)|^2 |x - m_f[\rho_t]|^2\omega_f^\alpha \, d\rho_t$$

$$- \alpha c_0 \frac{\sigma^2}{2} e^{-\alpha\underline{f}} \int |x - m_f[\rho_t]|^2 d\rho_t.$$

When $\alpha \geq c_1$, the first term in the estimate for I_2 is non-negative. Hence, we obtain

$$\frac{d}{dt}\|\omega_f^\alpha\|_{L^1(\rho_t)} \geq -\alpha e^{-\alpha\underline{f}}(c_0\sigma^2/2 + \lambda c_f) \int |x - m_f[\rho_t]|^2 d\rho_t$$

$$\geq -\alpha e^{-2\alpha\underline{f}}(c_0\sigma^2 + 2\lambda c_f)\frac{V(\rho_t)}{\|\omega_f^\alpha\|_{L^1(\rho_t)}},$$

where the second inequality follows from (4.1). We obtain (4.3) from the fact that

$$\frac{1}{2}\frac{d}{dt}\|\omega_f^\alpha\|_{L^1(\rho_t)}^2 = \|\omega_f^\alpha\|_{L^1(\rho_t)}\frac{d}{dt}\|\omega_f^\alpha\|_{L^1(\rho_t)} \geq -\alpha e^{-2\alpha\underline{f}}(c_0\sigma^2 + 2\lambda c_f)V(\rho_t),$$

thereby concluding the proof. □

We now have the ingredients to show the concentration of ρ_t. In particular, we show that the estimates (4.2) and (4.3) provide the means to identify assumptions

on the parameters α, λ and σ, for which we obtain the convergence $V(\rho_t) \to 0$ as $t \to \infty$ at an exponential rate.

Theorem 4.1. *Let f satisfy Assumption 4.1. Furthermore let the parameters α, λ and σ satisfy*

$$b_1 < \frac{3}{4}, \quad 2\lambda b_0^2 - K - 2d\sigma^2 b_0 e^{-\alpha \underline{f}} \geq 0,$$

with $K = V(\rho_0)$ and $b_0 = \|\omega_f^\alpha\|_{L^1(\rho_0)}$. Then $V(\rho_t) \leq V(\rho_0)e^{-qt}$ with

$$q = 2\big(\lambda - (d\sigma^2/b_0)e^{-\alpha \underline{f}}\big) \geq K/b_0^2.$$

In particular, there exists a point $\tilde{x} \in \mathbb{R}^d$ for which

$$E(\rho_t) \to \tilde{x} \quad and \quad m_f[\rho_t] \to \tilde{x} \quad as\ t \to \infty.$$

Remark 4.2. Since $b_1 \to 0$ and $e^{-\alpha \underline{f}} \to 0$ as $\alpha \to \infty$, the set of parameters for which the above conditions are satisfied is non-empty.

Proof of Theorem 4.1. Let $\gamma = K/b_0^2$. Choose any $\varepsilon > 0$ and set

$$\mathcal{T}^\varepsilon := \{t > 0 : V(\rho_s)e^{\gamma s} < K_\varepsilon \text{ for } s \in [0, t)\},$$

where $K_\varepsilon := K + \varepsilon$. Then, by continuity of $V(\rho_t)$, we get $\mathcal{T}^\varepsilon \neq \emptyset$. Set $\mathcal{T}_*^\varepsilon := \sup \mathcal{T}^\varepsilon$, and we claim $\mathcal{T}_*^\varepsilon = \infty$. Suppose not, i.e. $\mathcal{T}_*^\varepsilon < \infty$. Then this yields

$$\lim_{t \nearrow \mathcal{T}_*^\varepsilon} V(\rho_t)e^{\gamma t} = K_\varepsilon.$$

On the other hand, it follows from (4.3) that for $t < \mathcal{T}_*^\varepsilon$, we have

$$\|\omega_f^\alpha\|_{L^1(\rho_t)}^2 \geq b_0^2 - b_1 \int_0^t V(\rho_s)ds > b_0^2 - b_1 K_\varepsilon \int_0^t e^{-\gamma s}\, ds$$

$$= b_0^2 \left(1 - b_1\frac{K_\varepsilon}{K}\right) + \frac{b_1 K_\varepsilon}{\gamma}e^{-\gamma \mathcal{T}_*^\varepsilon}.$$

Since $b_1 < 3/4$ and $K_\varepsilon \to K$ as $\varepsilon \to 0$, there exists $\varepsilon_0 > 0$ such that $1 - b_1 K_\varepsilon/K \geq 1/4$ for $0 < \varepsilon \leq \epsilon_0$. Thus we obtain that for $t < \mathcal{T}_*^\varepsilon$ with $0 < \varepsilon \leq \varepsilon_0$

$$\|\omega_f^\alpha\|_{L^1(\rho_t)} \geq \frac{b_0}{2} + \xi_\varepsilon \quad \text{for some } \xi_\varepsilon > 0.$$

Inserting this into (4.2) gives

$$\frac{d}{dt}V(\rho_t) \leq -2\left(\lambda - \frac{d\sigma^2}{b_0 + 2\xi_\epsilon}e^{-\alpha \underline{f}}\right)V(\rho_t).$$

Applying Gronwall's inequality yields

$$V(\rho_t)e^{q_\varepsilon t} \leq K \quad \text{where } q_\varepsilon = 2\left(\lambda - \frac{d\sigma^2}{b_0 + 2\xi_\varepsilon}e^{-\alpha \underline{f}}\right),$$

for $t < \mathcal{T}_*^\varepsilon$ with $0 < \varepsilon \leq \varepsilon_0$. On the other hand, taking the limit $t \nearrow \mathcal{T}_*^\varepsilon$ to the above inequality gives

$$K_\varepsilon = \lim_{t \to \mathcal{T}_*^\varepsilon -} V(\rho_t)e^{\gamma t} < \lim_{t \to \mathcal{T}_*^\varepsilon -} V(\rho_t)e^{q_\varepsilon t} = K < K_\epsilon,$$

where we used the second condition, which gives $q_\varepsilon > \gamma$ for all $0 < \varepsilon \leq \varepsilon_0$. This is a contradiction and implies that $\mathcal{T}_*^\varepsilon = \infty$. Finally, by taking the limit $\varepsilon \to 0$, we complete the proof.

For the second part of the statement, we first observe that the expectation of ρ_t satisfies

$$\frac{d}{dt}E(\rho_t) = \frac{d}{dt}\int x \, d\rho_t = -\lambda \int (x - m_f[\rho_t])d\rho_t. \tag{4.4}$$

Taking the absolute value of the equation above and then integrating in time yields

$$\int_0^t \left|\frac{d}{ds}E(\rho_s)\right| ds \leq \lambda \int_0^t \int |x - m_f[\rho_s]|d\rho_s \, ds \leq 4(\lambda/b_0)e^{-\alpha \underline{f}}V(\rho_0)\int_0^t e^{-qs}ds$$

$$= \frac{4\lambda}{qb_0}e^{-\alpha \underline{f}}V(\rho_0)(1 - e^{-qt}) \leq \frac{4\lambda}{qb_0}e^{-\alpha \underline{f}}V(\rho_0),$$

where we used the fact that $\|\omega_f^\alpha\|_{L^1(\rho_t)}^2 \geq b_0/2$ and the concentration estimate $V(\rho_t) \leq V(\rho_0)e^{-qt}$. The previous estimate tells us that $dE(\rho_t)/dt \in L^1(0, \infty)$ and thus, provides the existence of a point $\tilde{x} \in \mathbb{R}^d$, possibly depending on ρ_0, such that

$$\hat{x} = E(\rho_0) + \int_0^\infty \frac{d}{dt}E(\rho_t)dt = \lim_{t \to \infty} E(\rho_t).$$

Moreover, since $|E(\rho_t) - m_f[\rho_t]| \to 0$ for $t \to \infty$, we obtain $\lim_{t \to \infty} m_f[\rho_t] = \tilde{x}$. □

Remark 4.3. A very important takeaway from the conditions provided in Theorem 4.1 is the fact that λ and σ may be chosen independently of α. Indeed, if

$$2\lambda b_0^2 - K - 2d\sigma^2 b_0 \geq 0, \tag{4.5}$$

then the second condition of Theorem 4.1 is trivially satisfied since $f \geq 0$ by assumption. Therefore, when (4.5) is satisfied, consensus is achieved for arbitrarily large α satisfying $b_1 \leq 3/4$.

4.2. *Approximate global minimizer*

While the previous results provided a sufficient condition for uniform consensus to occur, we will argue further that the point of consensus $\tilde{x} \in \mathbb{R}^d$ may be made arbitrarily close to the global minimizer x_* of f, for $f \in \mathcal{C}^2(\mathbb{R}^d)$ satisfying Assumption 4.1.

In the following, we assume further that f attains a unique global minimum at $x_* \in \text{supp}(\rho_0)$.

Theorem 4.2. *Let f satisfy Assumption* 4.1. *For any given $0 < \epsilon_0 \ll 1$ arbitrarily small, there exist some $\alpha_0 \gg 1$ and appropriate parameters (λ, σ) such that uniform*

consensus is obtained at a point $\tilde{x} \in B_{\epsilon_0}(x_*)$. More precisely, we have that $\rho_t \to \delta_{\tilde{x}}$ for $t \to \infty$, with $\tilde{x} \in B_{\epsilon_0}(x_*)$.

For the proof of Theorem 4.2, we will make use of the following auxiliary result.

Lemma 4.2. *Let f satisfy Assumption 4.1 and ρ_t satisfy*

$$(1 - b(\alpha))\|\omega_f^\alpha\|_{L^1(\rho_0)}^2 \leq \|\omega_f^\alpha\|_{L^1(\rho_t)}^2 \quad \text{for all } t \geq 0,$$

with $0 \leq b(\alpha) \leq 1$ and $b(\alpha) \to 0$ as $\alpha \to \infty$. Furthermore, assume that $V(\rho_t) \to 0$ and $E(\rho_t) \to \tilde{x}$ as $t \to \infty$. Then for any given $0 < \epsilon_0 \ll 1$ arbitrarily small, there exist some $\alpha_0 \gg 1$ such that $\tilde{x} \in B_{\epsilon_0}(x_*)$.

Proof. From the assumptions $V(\rho_t) \to 0$ and $E(\rho_t) \to \tilde{x}$ as $t \to \infty$, we deduce

$$\left| \int e^{-\alpha f} d\rho_t - \int_{B_\delta(\tilde{x})} e^{-\alpha f} d\rho_t \right| = \int_{\{|x - \tilde{x}| \geq k\}} e^{-\alpha f} d\rho_t \leq \rho_t(\{|x - \tilde{x}| \geq k\})$$

$$\leq \frac{1}{k^2} \int_{\{|x - \tilde{x}| \geq k\}} |x - \tilde{x}|^2 \, d\rho_t$$

$$\leq \frac{2}{k^2}(V(\rho_t) + |E(\rho_t) - \tilde{x}|^2),$$

for any $k > 0$. Furthermore, from Chebyshev's inequality, we know that $\rho_t \rightharpoonup \delta_{\tilde{x}}$ narrowly as $t \to \infty$. Thus, since f is continuous and bounded on $B_\delta(\tilde{x})$, we may pass to the limit in t to obtain

$$\lim_{t \to \infty} \int_{B_\delta(\tilde{x})} e^{-\alpha f} d\rho_t = e^{-\alpha f(\tilde{x})}.$$

Altogether, we obtain the convergence

$$\|\omega_f^\alpha\|_{L^1(\rho_t)} \to e^{-\alpha f(\tilde{x})} \quad \text{as } t \to \infty. \tag{4.6}$$

On the other hand, the Laplace principle (see e.g. Ref. 31) gives us the possibility to choose $\alpha_1 \gg 1$ large enough for any given $\epsilon > 0$, such that

$$-\frac{1}{\alpha} \log \|\omega_f^\alpha\|_{L^1(\rho_0)} - f_* < \frac{\epsilon}{2}$$

for any $\alpha \geq \alpha_1$. By assumption, we have that

$$-\frac{1}{\alpha} \log \|\omega_f^\alpha\|_{L^1(\rho_t)} \leq -\frac{1}{\alpha} \log \|\omega_f^\alpha\|_{L^1(\rho_0)} - \frac{1}{2\alpha} \log(1 - b(\alpha)).$$

From the continuity of the logarithm, we can pass to the limit $t \to \infty$ to obtain

$$f(\tilde{x}) \leq -\frac{1}{\alpha} \log \|\omega_f^\alpha\|_{L^1(\rho_0)} - \frac{1}{2\alpha} \log(1 - b(\alpha)).$$

This, together with using the assumption that $b(\alpha) \to 0$ as $\alpha \to \infty$, yields

$$0 \leq f(\tilde{x}) - f_* < \frac{\epsilon}{2} - \frac{1}{2\alpha} \log(1 - b(\alpha)) < \epsilon.$$

for $\alpha \geq \alpha_2$ with $\alpha_2 \gg 1$ sufficiently large. Finally, using the continuity of f and the uniqueness of the global minimum, we can find a suitable small $\epsilon_0 > 0$ (correspondingly $\alpha_0 \geq \max\{\alpha_1, \alpha_2\}$) such that $\tilde{x} \in B_{\epsilon_0}(x_*)$ for all $\alpha \geq \alpha_0$. $\qquad\square$

Proof of Theorem 4.2. It follows from Lemma 4.1 that

$$\|\omega_f^\alpha\|_{L^1(\rho_t)}^2 \geq \|\omega_f^\alpha\|_{L^1(\rho_0)}^2 - b_1(\alpha) \int_0^t V(\rho_s) ds,$$

where $b_1(\alpha) \to 0$ as $\alpha \to \infty$. Furthermore, by choosing λ and σ according to (4.5), we obtain from Theorem 4.1 that

$$\int_0^t V(\rho_s) ds \leq \int_0^t V(\rho_0) e^{-qs}\, ds = \frac{V(\rho_0)}{q}(1 - e^{-qt}) \leq (1 - e^{-qt})\|\omega_f^\alpha\|_{L^1(\rho_0)}^2,$$

and consequently,

$$\|\omega_f^\alpha\|_{L^1(\rho_t)}^2 \geq \|\omega_f^\alpha\|_{L^1(\rho_0)}^2 (1 - b_1(\alpha)).$$

We conclude the proof by choosing $\alpha \geq \alpha_0$ with α_0 obtained from Lemma 4.2.

□

5. Pseudo-Inverse Distribution, Extended Models and Numerical Results

In this section, we consider the Fokker–Planck equation (1.3) in one dimension and derive an equivalent formulation of the equation in terms of the pseudo-inverse distribution function. Then, we introduce an extension of the current model to replace the diffusion term with nonlinear diffusions of porous media type, which would guarantee compact support of the probability measure ρ_t. In this case the pseudo-inverse distribution function allows us to investigate concentration by only considering the evolution of the boundary points. To investigate this numerically, we introduce schemes for the porous media type equation and the evolution equation for the pseudo-inverse distribution function. We conclude with numerical results that illustrate the convergence results shown above.

5.1. *Evolution of the inverse distribution function*

We first define the well-known cumulative distribution F_t of a probability measure ρ_t and its pseudo-inverse χ_t (see e.g. Ref. 35) by

$$F_t(x) = \int_{-\infty}^x d\rho_t = \rho_t((-\infty, x]).$$

Then, the pseudo-inverse of F_t on the interval $[0, 1]$ is defined by

$$\chi_t(\eta) := F_t^{-1}(\eta) := \inf\{x \in \mathbb{R} \,|\, F_t(x) > \eta\}.$$

Both, F_t and χ_t are by definition right-continuous. To derive the evolution equation for χ_t we use the properties of ρ_t, F_t and χ_t collected in the following corollary. These properties can be obtained by basic calculus using the above definitions.

Corollary 5.1. *Let ρ_t be a probability measure, F_t the corresponding cumulative distribution and χ_t the pseudo-inverse distribution of F_t as defined above. Then,*

the following equations hold:

$$\chi(t, F(t,x)) = x, \quad F(t, \chi(t,\eta)) = \eta, \quad \partial_t F = -\rho \partial_t \chi, \quad \partial_\eta \chi = \frac{1}{\rho},$$

$$\partial_\eta \chi \, \partial_x F = 1, \quad \frac{\partial_x \rho}{\rho} = \frac{\partial_{xx} F}{\partial_x F} = -\frac{\partial_{\eta\eta}\chi}{(\partial_\eta \chi)^2}$$

restricted to $x = \chi(t, \eta)$, $\eta = F(t, x)$ *and* $x \in \mathrm{supp}(\rho)$, *respectively.*

From these properties we may now derive an integro-differential equation for the pseudo-inverse χ_t, $t \geq 0$. Indeed, let us consider the solution ρ_t to (1.3) satisfying $\rho_t \in \mathcal{C}([0,\infty), \mathcal{P}_2(\mathbb{R}))$. From the definition of F_t, we deduce that

$$\partial_t F_t(x) - \mu_t(x)\rho_t(x) = \rho_t(x)\partial_x \kappa_t(x) + \kappa_t(x)\partial_x \rho_t(x).$$

We then use the relations between χ_t and ρ_t provided in Corollary 5.1 to obtain

$$-\rho_t(x)\partial_t \chi_t(\eta) - \mu_t(x)\rho_t(x) = \rho_t(x)\partial_x \kappa_t(x) + \kappa_t(x)\partial_x \rho_t(x),$$

which consequently yields

$$\partial_t \chi_t(\eta) + \mu_t(x) = -\partial_x \kappa_t(x) - \kappa_t(x)\frac{\partial_x \rho_t(x)}{\rho_t(x)} = -\frac{\partial_\eta \kappa_t(\eta)}{\partial_\eta \chi_t(\eta)} + \kappa_t(\eta)\frac{\partial_{\eta\eta}\chi_t(\eta)}{(\partial_\eta \chi_t(\eta))^2}.$$

On the other hand, Corollary 5.1 implies

$$\mu_t = \lambda(\chi_t - m_f[\chi_t]) \quad \text{and} \quad \kappa_t = (\sigma^2/2)|\chi_t - m_f[\chi_t]|^2,$$

respectively, where $m_f[\chi_t]$ is given by

$$m_f[\chi_t] = \frac{\int_0^1 \chi_t \exp(-\alpha f(\chi_t(\eta)))d\eta}{\int_0^1 \exp(-\alpha f(\chi_t(\eta)))d\eta}.$$

Hence, the pseudo-inverse distribution χ_t satisfies the following integro-differential equation:

$$\partial_t \chi_t + \mu_t = -\partial_\eta(\kappa_t(\partial_\eta \chi_t)^{-1}).$$

Clearly, the results of Sec. 4 concerning the concentration and approximation of the global minimizer hold in the equivalent formulation in terms of χ_t as well.

5.2. *Porous media version of the evolution equation*

One very common application of the pseudo-inverse distribution χ_t is to study the behavior of the support $\mathrm{supp}(\rho_t)$ of the corresponding probability measure ρ_t. This is especially interesting when ρ_t has compact support. Unfortunately, we do not have that in the present case due to the diffusion, which causes ρ_t to have full support in $\mathbb{R}$. This naturally leads to the idea of increasing the power of ρ_t in the

diffusion term, inspired by the porous media equation.[15] The evolution equation for ρ_t then becomes

$$\partial_t \rho_t + \partial_x(\mu_t \rho_t) = \partial_{xx}(\kappa_t \rho_t^p), \tag{5.1}$$

with porous media coefficient $p \geq 1$. Notice that the previous model is included here for $p = 1$. The derivation of the evolution equation for χ_t corresponding to this equation may be analogously done, which leads to

$$\partial_t \chi_t + \mu_t = -\partial_\eta(\kappa_t(\partial_\eta \chi_t)^{-p}). \tag{5.2}$$

Further investigation of the diffusion term results in

$$\partial_\eta(\kappa_t(\partial_\eta \chi_t)^{-p}) = \partial_\eta(\kappa_t \, \rho_t^p) = \rho_t^p \partial_\eta \kappa_t + \kappa_t \partial_\eta(\rho_t^p) = \rho_t^p \partial_\eta \kappa_t + p\kappa_t \rho_t^{p-1} \partial_\eta \rho_t$$

in (η, t) variables. For $p > 1$ we can do the following formal computations. Due to mass conservation of ρ_t we assume a no flux condition for (5.1) which in (x, t) variables reads

$$\mu_t \rho_t = \partial_x(\kappa_t \rho_t^p) = \rho_t^p \partial_x \kappa_t + p\kappa_t \rho_t^{p-1} \partial_x \rho_t,$$

on the boundary points of $\text{supp}(\rho_t)$. Consequently, we obtain

$$\partial_\eta(\kappa_t(\partial_\eta \chi_t)^{-p})(F_t(x)) = [\rho_t^p \partial_x \kappa_t + p\kappa_t \rho_t^{p-1} \partial_x \rho_t]\partial_\eta \chi_t(F_t(x))$$

$$= \mu_t \rho_t \partial_\eta \chi_t(F_t(x)) = \mu_t,$$

on the boundary points of $\text{supp}(\rho_t)$. Therefore, restricting (5.2) onto the boundary points yields

$$\partial_t \chi_t(\eta) = -2\mu_t(\eta) = -2\lambda(\chi_t(\eta) - m_f[\chi_t]) \quad \text{for } \eta \in \{0, 1\}. \tag{5.3}$$

As $m_f[\chi_t]$ is contained in the interior of $\text{supp}(\rho_t)$ by definition, μ_t is negative at the left boundary point $\eta = 0$ and positive at the right boundary point $\eta = 1$. Hence, (5.3) implies the shrinking of $\text{supp}(\rho_t)$. In particular, one expects the concentration of χ_t at m_f as $t \to \infty$. We will numerically check this behavior in the next subsection.

5.3. *Discretization of the evolution equation for χ_t*

To investigate the behavior of the pseudo-inverse χ_t numerically, we use an implicit finite difference scheme. Following the ideas in Ref. 8 we denote the discretized version of χ_t by χ_k^i, where the spatial discretization is indexed by k and the temporal discretization by i. The spatial and temporal step sizes are denoted by h and τ, respectively. A straightforward discretization of the general equation (5.2) yields

$$\frac{\chi_k^{i+1} - \chi_k^i}{\tau} = -\left(\frac{\kappa(\chi_{k+1}^{i+1}, m_f^i)}{(\chi_{k+1}^{i+1} - \chi_k^{i+1})^p} - \frac{\kappa(\chi_k^{i+1}, m_f^i)}{(\chi_k^{i+1} - \chi_{k-1}^{i+1})^p} \right) h^{p-1} + \lambda(m_f^i - \chi_k^{i+1}),$$

$$\tag{5.4}$$

for $\eta \in (0,1)$, where $m_f^i = m_f[\chi^i]$. At the boundary points $\eta = 0, 1$ the expressions

$$\left(\chi_k^i - \chi_{k-1}^i\right)^{-p} \quad \text{and} \quad \left(\chi_{k+1}^i - \chi_k^i\right)^{-p}$$

are set to zero, respectively. As stopping criterion for the iteration procedure we use

$$\|\chi^{i+1} - \chi^i\|_{L^2(0,1)} < \texttt{tol}.$$

Since we expect the density ρ_t to concentrate close to the minimizer $x_* \in \mathbb{R}^d$ of the cost function f, the pseudo-inverse χ_t should converge toward the constant function with value $x_* \in \mathbb{R}^d$. This causes problems in the computation of the fractions appearing in (5.4). Our workaround is to evaluate the fractions up to a tolerance and set them artificially to zero if the denominator is too small. The scheme is tested with the well-known Ackley benchmark function for global optimization problems (see e.g. Ref. 3) shown in Fig. 1.

5.4. *Particle approximation*

To compare the results of the extension $p > 1$ to the scheme in Ref. 31, we are interested in a particle scheme corresponding to the evolution equation for $p = 2$. Note that in contrast to the pseudo-inverse distribution case, we are not restricted to one dimension here. We derive a numerical scheme by rewriting (5.1) as

$$\partial_t \rho_t = -\nabla_x(\mu_t \rho_t) + \Delta(\kappa_t \rho_t^2) = \nabla_x[-\mu_t \rho_t + \rho_t(\nabla_x(\kappa_t \rho_t) + \kappa_t \nabla_x \rho_t)].$$

The terms $\nabla_x(\kappa_t \rho_t)$ and $\nabla_x \rho_t$ are mollified in the spirit of Ref. 27 with the help of a mollifier φ_ϵ,

$$\nabla_x(\kappa_t \rho_t) \approx \nabla_x \varphi_\epsilon * (\kappa_t \rho_t) \quad \text{and} \quad \nabla_x \rho_t \approx \nabla_x \varphi_\epsilon * \rho_t.$$

Altogether this yields the approximate deterministic microscopic system

$$\dot{X}_t^i = -\lambda(X_t^i - m_t) + \frac{\sigma}{N} \sum_{j=1}^N \nabla_x \varphi_\epsilon(X_t^i - X_t^j)\left[|X_t^j - m_t|^{2p} + |X_t^i - m_t|^{2p}\right], \quad (5.5)$$

for $i = 1, \ldots N$, using the notation in Introduction.

Remark 5.1. Note that the scheme (5.5) is deterministic in contrast to the scheme (1.1) for $p = 1$. Unfortunately, it is not trivial to extend the particle scheme to $p > 2$.

5.5. *Numerical results*

In the following, numerical results corresponding to the above discretizations are reported. We use 200 grid points for the spatial discretization of χ_t and 500 particles
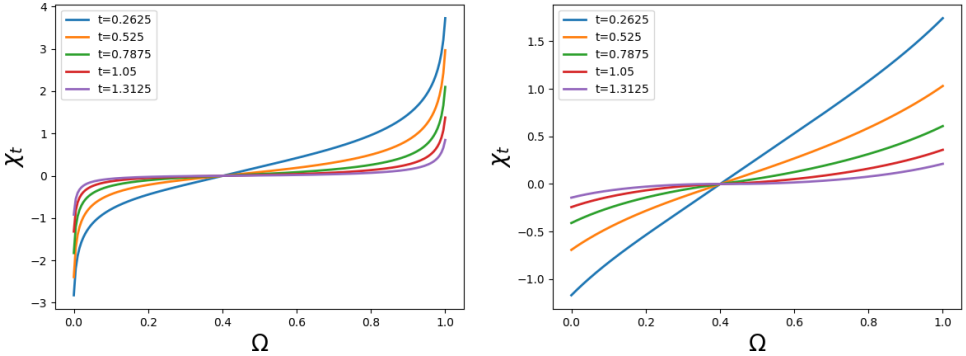
Fig. 2.   Ackley benchmark in one-dimensional. Progression of the inverse distribution function over time for the three benchmarks. Left: Diffusion with $p = 1$. Right: Diffusion with $p = 2$.

for the particle approximation schemes. Further parameters are fixed as

$$\tau = 2.5 \cdot 10^{-3}, \quad \alpha = 30, \quad \sigma = 0.8, \quad \texttt{tol} = 10^{-6}.$$

The mollifier is chosen to be $\varphi_\epsilon = \epsilon^{-d}\varphi(x/\epsilon)$, where

$$\varphi(x) = \frac{1}{Z_d} \begin{cases} \exp\left(\frac{1}{|x|^2 - 1}\right) & \text{if } |x| < 1, \\ 0 & \text{else,} \end{cases}$$

with normalizing constant Z_d.

Figure 2 shows the progression of χ_t over time. On the left side the case $p = 1$ is depicted. The tails mentioned in the discussion of (5.1) can be seen near the boundary. On the right side the diffusion coefficient is $p = 2$, in this case no tails occur as expected.

In Ref. 31, the following scheme with an approximate Heaviside function was proposed:

$$dX_t^i = -\lambda(X_t^i - m_t)H_\epsilon(f(X_t^i) - f(m_t))dt + \sqrt{2}\sigma|X_t^i - m_f|dW_t^i,$$

where m_t is as given in (1.1b). Initially, the Heaviside function was included to assure that the particles do not concentrate abruptly. This is essential in cases where the weight parameter $\alpha > 0$ is chosen too small, thereby yielding a rough approximation of the minimizer at the start of the simulation. In fact, the presence of the Heaviside function prevents particles that attain function values smaller than the function values at the average, i.e. $f(X^i) < f(m_t)$, from drifting to m_t. In those cases, only the diffusion part is active.

An analogous particle scheme for the porous media equation with $p = 2$ reads

$$\dot{X}_t^i = -\lambda(X_t^i - m_t)H_\epsilon(f(X_t^i) - f(m_f))$$

$$+ \frac{\sigma}{N}\sum_{j=1}^{N}\nabla_x\varphi(X_t^i - X_t^j)\big[|X_t^j - m_t|^{2p} + |X_t^i - m_t|^{2p}\big].$$
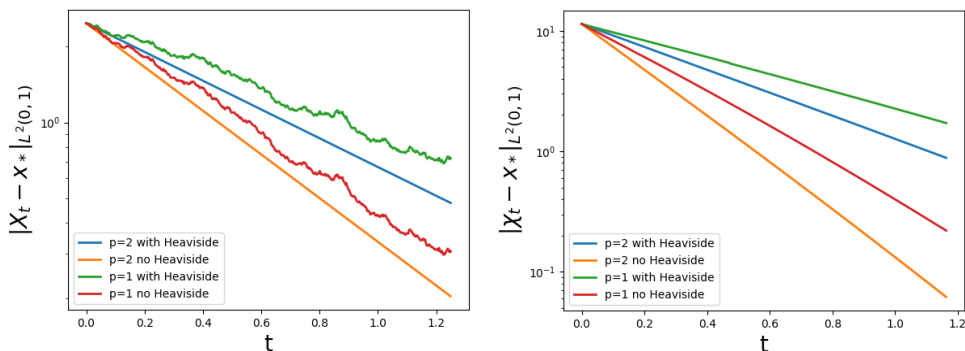
Fig. 3.   Ackley benchmark in one-dimension. L_2 error of the solution with respect to the minimizer x_* or, equivalently, the 2-Wasserstein distance between the solution and δ_{x^*} for the different benchmarks. Left: Particle scheme. The stochastic influence is visible for $p = 1$, as the graphs rely on data of one realization. Right: Pseudo-inverse distribution.

For both schemes a smooth approximation of the Heaviside function of the form,

$$H_\epsilon = (1 + \operatorname{erf}(x/\epsilon))/2$$

is used. We compare the results with and without the Heaviside function in Fig. 3. In these simulations, we see the damping effect of the Heaviside function. The simulations without Heaviside converge faster. Due to the large value of α, the minimizer is approximated well by m_f, thus the concentration happens very close to the actual minimum of the objective functions.

The graphs show the L_2-distance of $\mathbf{X}_t$ (left) and χ_t (right) to the known minimizer x_* or equivalently the 2-Wasserstein distance between the solutions of the mean-field equation and the particle scheme to the global consensus at δ_{x_*}. The schemes with nonlinear diffusion $p = 2$ converge faster than their corresponding schemes with linear diffusion. Nevertheless, for practical applications with large number of particles, the scheme with linear diffusion is more reasonable due to shorter computation times. In fact, in each iteration of the scheme (5.5) the convolution of all particles has to be computed. The error of the simulation for χ_t is smaller than the one for $\mathbf{X}_t$ at equal times. The linear graphs with respect to the logarithmic scaling of the y-axis in Fig. 3 indicate the exponential convergence shown in the theoretical section (cf. Theorem 4.1). The stochasticity influencing the schemes for $p = 1$ can be observed in the graphs in Fig. 3 (left).

In contrast to the pseudo-inverse distribution function, which is only available in one dimension, the particle scheme can be easily generalized to higher dimensions. We conclude the manuscript with some numerical results obtained with the particle scheme in two dimensions applied to the Ackley benchmark. In Fig. 4 (left), we see the surface and contour plot of the Ackley function in two dimensions. In Fig. 4 (right), the convergence of the different particle schemes is illustrated. For the stochastic scheme with $p = 1$ the data is averaged over 100 Monte Carlo simulations. The graphs are in good agreement with the corresponding graphs of the
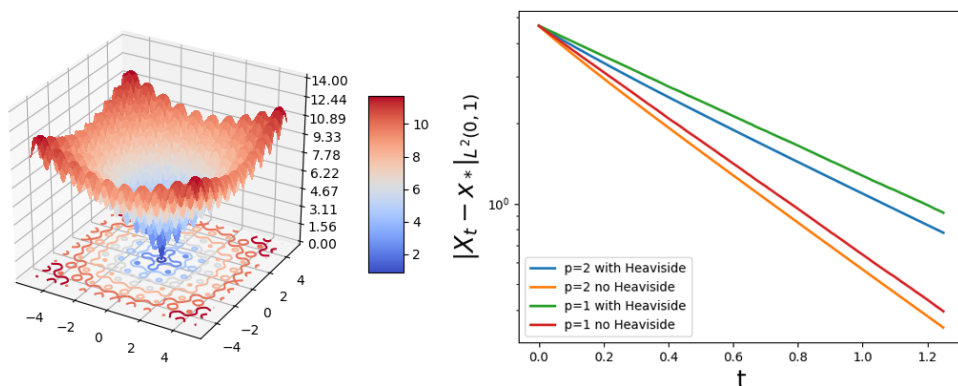
Fig. 4.   Ackley benchmark function in two dimensions. Left: Surface and contour plot of the function in two dimensions. Right: Convergence of the particle schemes. For the stochastic cases with $p = 1$ the average of 100 Monte Carlo simulations is shown.

pseudo-inverse distribution function in one dimension. Due to the averaging the stochastic influence which can be seen in Fig. 3 (left) disappears.

Acknowledgments

References

1. E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines* (John Wiley and Sons Inc., 1988).
2. L. Arnold, *Stochastic Differential Equations* (1974).
3. O. Askari-Sichani and M. Jalili, Large-scale global optimization through consensus of opinions over complex networks, *Compl. Adapt. Syst. Model.* **1** (2013) 1–18.
4. T. Back, D. B. Fogel and Z. Michalewicz, *Handbook of Evolutionary Computation* (IOP Publishing Ltd., 1997).
5. N. Bellomo, A. Bellouquid and D. Knopoff, From the microscale to collective crowd dynamics, *Multiscale Model. Simul.* **11** (2013) 943–963.
6. A. L. Bertozzi, J. Rosado, M. B. Short and L. Wang, Contagion shocks in one dimension, *J. Stat. Phys.* **158** (2015) 647–664.
7. L. Bianchi, M. Dorigo, L. M. Gambardella and W. J. Gutjahr, A survey on metaheuristics for stochastic combinatorial optimization, *Natur. Comput. Int. J.* **8** (2009) 239–287.
8. A. Blanchet, V. Calvez and J. A. Carrillo, Convergence of the mass-transport steepest descent scheme for the subcritical Patlak–Keller–Segel model, *SIAM J. Numer. Anal.* **46** (2008) 691–721.

9. C. Blum and A. Roli, Metaheuristics in combinatorial optimization: Overview and conceptual comparison, *ACM Comput. Surv.* **35** (2003) 268–308.

10. F. Bolley, Separability and completeness for the wasserstein distance, in *Séminaire de Probabilités XLI* (Springer, 2008), pp. 371–377.

11. F. Bolley, I. Gentil and A. Guillin, Uniform convergence to equilibrium for granular media, *Arch. Ration. Mech. Anal.* **208** (2013) 429–445.

12. J. A. Carrillo, Y.-P. Choi and M. Hauray, The derivation of swarming models: Mean-field limit and Wasserstein distances, in *Collective Dynamics from Bacteria to Crowds*, CISM Courses and Lectures, Vol. 553 (Springer, 2014), pp. 1–46.

13. J. A. Carrillo, M. Fornasier, J. Rosado and G. Toscani, Asymptotic flocking dynamics for the kinetic Cucker–Smale model, *SIAM J. Math. Anal.* **42** (2010) 218–236.

14. J. A. Carrillo, M. Fornasier, G. Toscani and F. Vecil, Particle, kinetic, and hydro-dynamic models of swarming, in *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences*, Modeling and Simulation in Science, Engineering and Technology (Birkhäuser, 2010), pp. 297–336.

15. J. A. Carrillo, M. P. Gualdani and G. Toscani, Finite speed of propagation in porous media by mass transportation methods, *C. R. Math.* **338** (2004) 815–818.

16. J. A. Carrillo, Y. Huang and S. Martin, Explicit flock solutions for quasi-morse potentials, *Europ. J. Appl. Math.* **25** (2014) 553–578.

17. J. A. Carrillo, Y. Huang and S. Martin, Nonlinear stability of flock solutions in second-order swarming models, *Nonlinear Anal. Real World Appl.* **17** (2014) 332–343.

18. F. Cucker and S. Smale, On the mathematics of emergence, *Jpn. J. Math.* **2** (2007) 197–227.

19. A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Vol. 38 (Springer Science & Business Media, 2009).

20. M. Dorigo and C. Blum, Ant colony optimization theory: A survey, *Theor. Comput. Sci.* **344** (2005) 243–278.

21. R. Durrett, *Stochastic Calculus*: *A Practical Introduction*, Vol. 6 (CRC press, 1996).

22. D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order* (Springer, 2015).

23. S.-Y. Ha and E. Tadmor, From particle to kinetic and hydrodynamic descriptions of flocking, *Kinet. Relat. Models* **1** (2008) 415–435.

24. R. A. Holley, S. Kusuoka and D. W. Stroock, Asymptotics of the spectral gap with applications to the theory of simulated annealing, *J. Funct. Anal.* **83** (1989) 333–347.

25. R. Holley and D. Stroock, Simulated annealing via sobolev inequalities, *Comm. Math. Phys.* **115** (1988) 553–569.

26. J. Kennedy, Particle swarm optimization, in *Encyclopedia of Machine Learning* (Springer, 2010), pp. 760–766.

27. A. Klar and S. Tiwari, A multiscale meshfree method for macroscopic approximations of interacting particle systems, *Multiscale Model. Simul.* **12** (2014) 1167–1192.

28. T. Kolokolnikov, J. A. Carrillo, A. Bertozzi, R. Fetecau and M. Lewis, Emergent behavior in multi-particle systems with non-local interactions [Editorial], *Phys. D* **260** (2013) 1–4.

29. S. Motsch and E. Tadmor, Heterophilious dynamics enhances consensus, *SIAM Rev.* **56** (2014) 577–621.

30. R. Olfati-Saber and R. M. Murray, Consensus problems in networks of agents with switching topology and time-delays, *IEEE Trans. Automat. Control* **49** (2004) 1520–1533.

31. R. Pinnau, C. Totzeck, O. Tse and S. Martin, A consensus-based model for global optimization and its mean-field limit, *Math. Mod. Methods Appl. Sci.* **27** (2017) 183–204.

32. C. Reeves, *Genetic Algorithms* (Springer, 2003).
33. A.-S. Sznitman, Topics in propagation of chaos, in *Ecole d'été de Probabilités de Saint-Flour, XIX1989* (Springer, 1991), pp. 165–251.
34. G. Toscani, Kinetic models of opinion formation, *Commun. Math. Sci.* **4** (2006) 481–496.
35. C. Villani, *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58 (Amer. Math. Soc., 2003).
36. C. Villani, *Optimal Transport*: *Old and New*, Vol. 338 (Springer Science & Business Media, 2008).
37. J. von Brecht, T. Kolokolnikov, A. L. Bertozzi and H. Sun, Swarming on random graphs, *J. Stat. Phys.* **151** (2013) 150–173.