Electron. J. Probab. **22** (2017), no. 8, 1–37.
ISSN: 1083-6489 DOI: 10.1214/17-EJP25

Uniform in time interacting particle approximations for nonlinear equations of Patlak-Keller-Segel type*

Amarjit Budhiraja Wai-Tong (Louis) Fan†

Abstract

We study a system of interacting diffusions that models chemotaxis of biological cells or microorganisms (referred to as particles) in a chemical field that is dynamically modified through the collective contributions from the particles. Such systems of reinforced diffusions have been widely studied and their hydrodynamic limits that are nonlinear non-local partial differential equations are usually referred to as Patlak-Keller-Segel (PKS) equations.

Solutions of the classical PKS equation may blow up in finite time and much of the PDE literature has been focused on understanding this blow-up phenomenon. In this work we study a modified form of the PKS equation that is natural for applications and for which global existence and uniqueness of solutions are easily seen to hold. Our focus here is instead on the study of the long time behavior through certain interacting particle systems.

Under the so-called "quasi-stationary hypothesis" on the chemical field, the limit PDE reduces to a parabolic-elliptic system that is closely related to granular media equations whose time asymptotic properties have been extensively studied probabilistically through certain Lyapunov functions [17, 4, 9]. The modified PKS equation studied in the current work is a parabolic-parabolic system for which analogous Lyapunov function constructions are not available. A key challenge in the analysis is that the associated interacting particle system is not a Markov process as the interaction term depends on the whole history of the empirical measure.

We establish, under suitable conditions, uniform in time convergence of the empirical measure of particle states to the solution of the PDE. We also provide uniform in time exponential concentration bounds for rate of the above convergence under additional integrability conditions. Finally, we introduce an Euler discretization scheme for the simulation of the interacting particle system and give error bounds that show that the scheme converges uniformly in time and in the size of the particle system as the discretization parameter approaches zero.

Keywords: weakly interacting particle systems; uniform propagation of chaos; McKean-Vlasov equations; kinetic equations; chemotaxis; reinforced diffusions; Patlak-Keller-Segel equations; granular media equations; uniform exponential concentration bounds; long time behavior; uniform in time Euler approximations.

*Research partially supported by National Science Foundation (DMS-1305120), the Army Research Office (W911NF-14-1-0331) and DARPA (W911NF-15-2-0122).

†E-mail: ariesfanhk@gmail.com

AMS MSC 2010: 60K35; 60H30; 60H35; 60K40; 60F05.
Submitted to EJP on July 24, 2016, final version accepted on January 8, 2017.

1 Introduction

Consider the following system of nonlinear nonlocal partial differential equations

$$\begin{cases} \partial_t u(t,x) = \dfrac{1}{2}\,\Delta u(t,x) - \nabla\cdot\big(u(t,x)[\chi\,\nabla h(t,x) - \nabla V(x)]\big) \\ \dfrac{1}{\gamma}\,\partial_t h(t,x) = \dfrac{1}{2}\,\Delta h(t,x) - \alpha h(t,x) + \beta\displaystyle\int_{\mathbb{R}^d} u(t,z)\,g(z-x)\,dz, \end{cases} \tag{1.1}$$

where $(t,x)\in(0,\infty)\times\mathbb{R}^d$ and α, β, γ, χ are positive constants. The symbols ∇, $\nabla\cdot$ and Δ denote the gradient operator, the divergence operator and the standard Laplacian respectively. Equations of the above form arise as reinforced diffusion models for chemotaxis of particles representing biological cells or microorganisms in which the particle diffusions are directed by the gradient of a chemical field which in turn is dynamically modified by the contributions of the particles themselves(cf. [20, 15]). The functions $u(t,x)$ and $h(t,x)$ represent, respectively, the continuum limits of the densities of the biological particles and the particles constituting the chemical field. The parameters α and β in the second equation model the decay rate of the chemical particles and the rate at which the biological particles contribute to the chemical field, respectively. The function g is the dispersal kernel which models the spread and amount of the chemical produced by the biological particles. A natural form for g is a Gaussian kernel $g(x,y)\doteq(2\pi\delta)^{-d/2}\exp\{-|y-x|^2/2\delta\}$, where δ is a small parameter. The first equation describes the collective motion of the biological particles. The dynamics of the individual particles is coupled through the gradient of the chemical field, i.e. ∇h, which defines their drift coefficient (up to a positive constant multiplier χ). Finally, the function V models a confinement potential for the particle motions. Thus the reinforcement mechanism is as follows. Particles are attracted to the chemical and they emit the chemical at a constant rate, resulting in a positive feedback: the more the cells are aggregated, the more concentrated in their vicinity is the chemical they produce which in turn attracts other cells.

The key feature of the model is the competition between the aggregation resulting from the above reinforcement mechanism and the diffusive effect which spreads out the biological and chemical particles in space. When $V=0$ and $g(z-x)dz$ is the Dirac delta measure δ_x, (1.1) becomes the classical Patlak-Keller-Segel (PKS) model [20, 15] which has been studied extensively. It is well known that for the 2-d PKS model (i.e. $d=2$), there is a critical mass M_c such that (i) the solution to (1.1) blows up in finite time if the initial mass $\int u(0,x)dx > M_c$, and (ii) a smooth solution exists for all time if $\int u(0,x)dx < M_c$. For $d\geq 3$, the blow up of the solution is related to the $L^{d/2}$ norm of the initial density but here the theory is less well developed. We refer the reader to the survey articles [13, 14] for references to the large literature on the PKS model and its variants.

One line of active research has focused on the prevention of finite time blowup of solutions via various modifications of the classical PKS equation that discourage mass concentration (cf. [10, 1, 7] and references therein.) The replacement of the Dirac delta measure by a smooth density $g(z-x)dz$, as is considered in the current work, can be regarded as one such natural modification of the PKS model in which chemicals are dispersed by cells over a region of positive area rather than over a single point. It is easy to see that there are global unique solutions for general initial conditions for the system (1.1) (cf. Proposition 2.3). The focus of this work is instead on the study of

the long time behavior of (1.1) and their particle approximations, for which very little is known. Long time behavior of weakly interacting particle systems of various types has been investigated in many recent works ([25, 8, 17, 9, 4]) and although our work uses many ideas similar to those in these works, one key distinguishing feature and challenge in the model considered here is that the associated weakly interacting particle system is not a Markov process. In particular, Lyapunov function constructions that have been extensively used in the proofs in the above works are not available for the model considered here.

Note that if (u, h) solve (1.1) and $\int u(0,x)dx = m$, then $(u/m, h)$ solves (1.1) with β replaced by βm. Thus we can (and will) assume without loss of generality that $\int u(0,x)dx = 1$. Our starting point is the following probabilistic representation for the solution of (1.1) in terms of a nonlinear diffusion of the McKean-Vlasov type.

$$\begin{cases} d\bar{X}_t = dB_t - \nabla V(\bar{X}_t)\, dt + \chi\, \nabla h(t,\, \bar{X}_t)\, dt, \\ \dfrac{1}{\gamma}\, \partial_t h(t,x) = \dfrac{1}{2}\, \Delta h(t,x) - \alpha h(t,x) + \beta \displaystyle\int_{\mathbb{R}^d} g(z-x)\, d\mu_t(z), \\ \mathcal{L}(\bar{X}_t) = \mu_t, \end{cases} \tag{1.2}$$

where $\{B_t\}$ is the standard Brownian motion in $\mathbb{R}^d$ and $\mathcal{L}(\bar{X}_t)$ denotes the probability law of $\bar{X}_t$. In Proposition 2.3 we will show that the above equation has a unique pathwise solution $(\bar{X}_t, h(t,\cdot))$ under natural conditions on the initial data and the kernel g. Furthermore, for $t > 0$ the measure μ_t admits a density $u(t,\cdot)$ with respect to the Lebesgue measure. The first equation in (1.1) can be regarded as the Kolmogorov's forward equation for the first equation in (1.2). In particular, it is easy to check that the pair $(u(t,\cdot), h(t,\cdot))$ is a solution of (1.1). Along with the nonlinear diffusion (1.2) we will also study a mesoscopic particle model for the chemotaxis phenomenon described above that is given through a stochastic system of weakly interacting particles of the following form.

$$\begin{cases} dX_t^{i,N} = dB_t^i - \nabla V(X_t^{i,N})\, dt + \chi\, \nabla h_N(t, X_t^{i,N})\, dt, \quad i = 1, \ldots, N \\ \dfrac{1}{\gamma}\, \partial_t h_N(t,x) = \dfrac{1}{2}\, \Delta h_N(t,x) - \alpha h_N(t,x) + \dfrac{\beta}{N} \displaystyle\sum_{i=1}^N g(X_t^{i,N} - x),\ x \in \mathbb{R}^d \end{cases} \tag{1.3}$$

where $\{B^i\}_{i=1}^N$ are independent standard Brownian motions in $\mathbb{R}^d$. Note that the second equation in (1.3) is the same as that in (1.2) with μ_t replaced by the empirical measure

$$\mu_t^N \doteq \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}}. \tag{1.4}$$

In this model a detailed evolution for biological particles is used whereas the chemical field is regarded as the continuum limit of much smaller chemical molecules. One can also consider a microscopic model where a detailed evolution equation of chemical particles replaces the second equation in (1.3). Such 'full particle system approximations' of (1.1) for the classical PKS model (i.e. when g is replaced by a Dirac probability measure) were studied in [22] where the convergence of the empirical measures of the biological and chemical particles to the solution of the limit PDE, up to the blow up time of the solutions, was established. Starting from the works of McKean and Vlasov [18], nonlinear diffusions and the associated weakly interacting particle models have been studied by many authors (See, for instance, [23, 19, 17, 16].) One important difference in (1.2) (and similarly (1.3)) from these classical papers is that the right side of the first equation depends not only on μ_t but rather on the full past trajectory of the laws, i.e. $\{\mu_s : 0 \le s \le t\}$.

The first main goal of this work is to rigorously establish that under suitable conditions, as N becomes large, the mesoscopic model (1.3) gives a good approximation for (1.2) (and thus also for (1.1)), uniformly in time. Specifically, our results will give, under conditions, uniform in time convergence of μ_t^N to μ_t, in a suitable sense. Such a result is important since it says in particular that the time asymptotic aggregation behavior of the particle system is well captured by the asymptotic density function $u(t,\cdot)$ as $t \to \infty$. In general one would also like to know how well μ^N approximates μ for a fixed value of N. In order to address such questions, in our second result, under stronger integrability conditions, we will provide uniform in time exponential concentration bounds that give estimates on rates of convergence of μ^N to μ.

A natural empirical approach for the study of long time properties of (1.3) is through numerical simulations. For example, under the Neumann boundary condition, a numerical simulation for (1.7) in a square demonstrates a separation of time scale: after an initial short time interval during which particles aggregates to form many crowded subpopulations, the subpopulations merge to form a stationary profile at a much slower time scale. See [11] and Figure 3.9 in [21] for such simulation results. Note however that the system cannot be simulated exactly and in practice one needs to do a suitable time discretization. For such simulations to form a reliable basis for mathematical intuition on the long time behavior, it is key that they approximate the system (1.3) or the PDE (1.1), uniformly in time. We will show that under suitable conditions a natural discretization scheme for (1.3) gives a uniform in time convergent approximation to the solution of (1.2) as $N \to \infty$ and as the discretization step size tends to zero. Our uniform in time numerical approximations offer qualitative insights for the long time dynamical behavior of such systems.

1.1 Existing results and some challenges

One of the key challenges in the study of (1.3) is that the Nd dimensional process $X^{(N)} = (X^{1,N}, \cdots, X^{N,N})$ is not a Markov process since the right side of the first equation in (1.3) depends on the full past history of the empirical measure, i.e. $\{\mu_s^N\}_{0\le s\le t}$. In order to get a Markovian descriptor one needs to consider the pair $(X^{(N)}, h_N)$ which is an infinite dimensional Markov process. Similar difficulties arise in the study of (1.1) where the form of the coupling between u and h makes the analysis challenging.

These difficulties do not occur for the reduced parabolic-elliptic system (1.5) obtained by formally letting $\gamma \to \infty$ in (1.1):

$$\begin{cases} \partial_t u(t,x) = \dfrac{1}{2}\,\Delta u(t,x) - \nabla \cdot \big(u(t,x)[\chi\,\nabla h(t,x) - \nabla V(x)]\big) \\ \qquad\qquad 0 = \dfrac{1}{2}\,\Delta h(t,x) - \alpha h(t,x) + \beta \displaystyle\int_{\mathbb{R}^d} u(t,z)\, g(z-x)\, dz. \end{cases} \tag{1.5}$$

In the context of chemotaxis, the model in (1.5) corresponds to a quasi-stationary hypothesis for the chemical h, that is, the chemical diffuses at a much faster time scale than the biological particles. Equation (1.5) is mathematically more tractable since here one can solve for h explicitly in terms of u and g as

$$h = \beta G_\alpha * u,$$

where $*$ denotes the standard convolution operator, $G_\alpha(z) = \int_0^\infty e^{-\alpha t} P_t g(z)\, dt$ and P_t is the standard heat semigroup, i.e. the semigroup generated by $\frac{1}{2}\Delta$. Using this expression for h, the system (1.5) can be expressed as a single equation of the form

$$\partial_t u = \frac{1}{2}\,\Delta u + \nabla \cdot (u\,[\nabla V - \chi\,\beta\,\nabla G_\alpha * u]). \tag{1.6}$$

Kinetic equations of the above form have been well studied in the literature [18, 25, 23, 8, 9, 17, 4] where they are sometimes referred to as granular media equations because of their use in the modeling of granular flows (cf. [2]). An interacting particle approximation for this equation takes the following simple form

$$dX_t^{(N)} = dB_t^{(N)} - \nabla\Phi_N(X_t^{(N)})\,dt, \tag{1.7}$$

where $X_t^{(N)} = (X_t^{1,N}, \ldots, X_t^{N,N})$, $\{B_t^{(N)}\}$ is a standard Brownian motion in $\mathbb{R}^{dN}$ and for $\mathbf{x} = (x_1, \cdots, x_N) \in \mathbb{R}^{dN}$,

$$\Phi_N(\mathbf{x}) \doteq \sum_{i=1}^N V(x_i) - \frac{\chi\beta}{2N}\sum_{i=1}^N\sum_{j=1}^N G_\alpha(x_i - x_j).$$

Note that in this case X^N is a Markov process given as a Nd dimensional diffusion with a gradient form drift. Law of large number results and propagation of chaos properties for such models over a finite time interval that rigorously connect the asymptotic behavior of (1.7) as $N \to \infty$ with the equation in (1.6) are classical and go back to the works of McKean [18] and Sznitman [23]. In recent years there has also been significant progress in the study of the time asymptotic behavior of (1.7) and (1.6). Under suitable growth and convexity assumptions on V and G_α, [25] studied the existence and local exponential stability of fixed points of (1.5) by a suitable construction of a Lyapunov function. Similar Lyapunov functions were used in [8, 17, 9] to establish a uniform in time propagation of chaos property and convergence of $\mu_t^N(dx)$ to $u(t,x)dx$ along with uniform in time exponential concentration bounds. In the context of the model in (1.7) and (1.6), simple modifications of arguments in proofs of [17, Theorem 1.3] and [9, Theorem 3.1] imply the following results. Suppose

$$\langle x - y,\, \nabla V(x) - \nabla V(y)\ \geq \lambda|x-y|^2 \quad \text{for all } x,\, y \in \mathbb{R}^d \tag{1.8}$$

and

$$\lambda > 2\,d\,\beta\chi\frac{\|\operatorname{Hess} g\|_\infty}{\alpha} \tag{1.9}$$

where $\|\operatorname{Hess} f\|_\infty \doteq \sup_{i,j}\sup_x |\partial_{x_i}\partial_{x_j} f(x)|$. Then as $N \to \infty$,

$$\sup_{t\geq 0}\mathcal{W}_2\Big(\mathcal{L}(X_t^{1,N},\, X_t^{2,N},\, \cdots,\, X_t^{k,N}),\ \big(u(t,x)dx\big)^{\otimes k}\Big) \longrightarrow 0 \tag{1.10}$$

for all positive integers k, where for $p \geq 1$, $\mathcal{W}_p$ is the Wasserstein-p distance (see Section 1.2) on the space of probability measures on $\mathbb{R}^{dk}$ and u is the solution to (1.6). Under the same assumptions (1.8) and (1.9), a uniform concentration bound of the form

$$\sup_{t\geq 0}\mathbb{P}\left(\mathcal{W}_1(\mu_t^N,\, u(t,x)dx) > \epsilon\right) \leq C_1(1+\epsilon^{-2})\ \exp\left(-C_2\,N\,\epsilon^2\right)$$

is obtained in [4, Theorem 2.12], where $C_1,\, C_2 \in (0,\infty)$ (the condition $\beta + 2\gamma > 0$ in [4, Theorem 2.12] is implied by (1.9).) The paper [17] proves uniform in time weak convergence of empirical measures constructed from an implicit Euler discretization scheme for the Markovian system (1.7) to the solution of (1.6). As remarked earlier, uniform in time numerical approximations are useful for obtaining qualitative insights for the long time dynamical behavior of such systems.

Much less is known for the model (1.1)–(1.3). For the classical parabolic-parabolic Patlak-Keller-Segel PDE a global existence in the subcritical case (i.e the initial mass is less than 8π) in $\mathbb{R}^2$ is established in [6] and the corresponding uniqueness result is established in [5]. We refer the reader to references in [5] for recent development of

the parabolic-parabolic Patlak-Keller-Segel PDE. None of these works consider particle approximations or long time behavior (however see [5] for recent stability results in the plane in a quasi parabolic-elliptic regime.) The goal of the current work is to develop the theory for the long time behavior of (1.1)–(1.3), analogous to the one for parabolic-elliptic model described above. As noted earlier, our approach is inspired by the ideas developed in [25, 8, 17, 9, 4]. Our main contributions are as follows.

1.2 Contributions of this work

In this work we identify conditions under which the particle system (1.3) converges to the nonlinear process (1.2) *uniformly* over the infinite time horizon and construct time stable numerical approximations for (1.3) and (1.2). More precisely, the main contributions of this paper are as follows.

1. Under suitable conditions, well-posedness of (1.2) and (1.3) is established in Propositions 2.2 and 2.3.
2. Sufficient conditions for a uniform in time propagation of chaos (POC) property for (1.3) are identified in Theorem 3.4. This implies, under the same conditions, a uniform in time law of large numbers (LLN) for the empirical measures μ^N (Corollary 3.5).
3. Under stronger integrability conditions, we establish uniform in time exponential concentration bounds for μ^N given in Theorem 3.8. These bounds say that the probability of observing deviation of μ_t^N from its LLN prediction μ_t is exponentially small, uniformly in t, as N increases.
4. An explicit Euler scheme for (1.3) is constructed and it is shown that it converges to the solution of (1.3) uniformly in time and in N(Theorem 3.10). Together with the POC result in 2 this shows that the Euler scheme gives a uniform in time convergent approximation for the nonlinear process as $N \to \infty$ and step size goes to 0 (Corollary 3.11).

Our main condition for uniform in time results in 2,3,4 is Assumption 2.4. This assumption can be regarded as the analog of condition (1.9) used in the study of (1.6)–(1.7).

The paper is organized as follows. In Section 2, we present the basic wellposedness results and introduce our main assumptions. Section 3 contains the main results of this work. Finally Section 4 is devoted to proofs.

Notation: For a Polish space (i.e. a complete separable metric space) S, $\mathcal{P}(S)$ denotes the space of all probability measures on S. This space is equipped with the topology of weak convergence. Distance on a metric space S will be denoted as $d_S(\cdot,\cdot)$ and if S is a normed linear space S the corresponding norm will be denoted as $\|\cdot\|_S$. If clear from the context S will be suppressed from the notation. The space of continuous functions from an interval $I \subset [0,\infty)$ to $\mathbb{R}^d$ is denoted by $\mathcal{C}(I:\mathbb{R}^d)$. The space $\mathcal{C}_T \doteq \mathcal{C}([0,T]:\mathbb{R}^d)$ will be equipped with the usual uniform norm and the Fréchet space $\mathcal{C} \doteq \mathcal{C}([0,\infty):\mathbb{R}^d)$ will be equipped with the distance

$$d_{\mathcal{C}}(x,y) \doteq \sum_{k=1}^{\infty} \frac{\|x-y\|_{\mathcal{C}_k} \wedge 1}{2^k}.$$

Given metric spaces S_i, $i = 1,\dots k$, the distance on the space $S_1 \times \cdots \times S_k$ is taken to be the sum of the k distances:

$$d_{S_1\times\cdots\times S_k}(x,y) \doteq \sum_{i=1}^{k} d_{S_i}(x_i,y_i),\ x=(x_1,\cdots x_k), y=(y_1,\cdots y_k).$$

The law of a S valued random variable X (an element of $\mathcal{P}(S)$), is denoted by $\mathcal{L}(X)$. A collection of S valued random variables $\{X_\alpha\}$ is said to be tight if their laws $\{\mathcal{L}(X_\alpha)\}$ are tight in $\mathcal{P}(S)$. For a signed measure μ on S and a μ-integrable function $f : S \to \mathbb{R}$, we write $\int f d\mu$ as $\langle f, \mu\rangle$. For a polish space S, the Wasserstein-p distance on $\mathcal{P}(S)$ is defined as

$$\mathcal{W}_p(\mu,\nu) \doteq \Big(\inf_{\pi} \int\int d_S(x,y)^p \, d\pi(x,y)\Big)^{1/p}, \tag{1.11}$$

where the infimum is taken over all probability measures $\pi \in \mathcal{P}(S \times S)$ with marginals μ and ν. Let $\mathcal{P}_p(S)$ be the set of $\mu \in \mathcal{P}(S)$ having finite p-th moments where $p \in [1, \infty)$. It is well known (cf. [26, Definition 6.8 and Theorem 6.9]) that $\mathcal{W}_p$ metrizes the weak convergence in $\mathcal{P}_p(S)$. For $p = 1$, the Kantorovich-Rubenstein duality (cf. [26, Remark 6.5]) says that for probability measures μ and ν which have finite first moments,

$$\mathcal{W}_1(\mu,\nu) = \sup_{f \in \mathrm{Lip}_1(S)} |\langle f, (\mu-\nu)\rangle|\,, \tag{1.12}$$

where $\mathrm{Lip}_1(S)$ is the space of Lipschitz functions on S whose Lipschitz constant is at most 1.

Throughout, $(\Omega, \mathcal{F}, \mathbb{P})$ will denote a probability space which is equipped with a filtration $(\mathcal{F}_t)$ satisfying the usual conditions. The symbol $\mathbb{E}$ denotes the expectation with respect to the probability measure $\mathbb{P}$. For a stochastic process X the notation X_t and $X(t)$ will be used interchangeably.

The space of all bounded continuous functions on S is denoted by $\mathcal{C}_b(S)$. The supremum of a function $f : S \to \mathbb{R}$ is denoted as $||f||_\infty \doteq \sup_{x\in S} |f(x)|$. Space of functions with k continuous (resp. continuous and bounded) derivatives will be denoted as $\mathcal{C}^k(\mathbb{R}^d)$ (resp. $\mathcal{C}_b^k(\mathbb{R}^d)$). For $f \in \mathcal{C}_b^1(\mathbb{R}^d)$ and $g \in \mathcal{C}_b^2(\mathbb{R}^d)$ we denote

$$\|\nabla f\|_\infty = \sup_x \Big(\sum_{i=1}^d (\partial_{x_i} f)^2\Big)^{1/2} \quad \text{and} \quad \|\operatorname{Hess} g\|_\infty = \sup_{i,j}\sup_x |\partial_{x_i}\partial_{x_j} g(x)|.$$

2 Preliminaries and well-posedness

Note that for a bounded function $g : \mathbb{R}^d \to \mathbb{R}$, a solution h to (1.2) by the variation of constants formula satisfies

$$h(t,x) = Q_t\, h_0\,(x) + \gamma\,\beta \int_0^t Q_{t-s}\Big(\int g(y-\cdot)\, d\mu_s(y)\Big)(x)\, ds, \tag{2.1}$$

where $\{Q_t\}_{t\geq 0}$ is the semigroup for $f \mapsto \frac{\gamma}{2}\Delta f - \gamma\alpha f$. That is for suitable $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$Q_t\phi(x) \doteq \mathbb{E}[\phi(x + B_{\gamma t})e^{-\gamma\alpha t}], \tag{2.2}$$

where $\{B_t\}$ is a standard d-dimensional Brownian motion. The equation in (2.1) can be rewritten as follows. For $m \in \mathcal{P}\big(\mathcal{C}([0,\infty) : \mathbb{R}^d)\big)$, let

$$\Theta_t^m(x) \doteq \int_0^t Q_{t-s}\Big(\int g(y-\cdot)\, dm_s(y)\Big)(x)\, ds,\ x \in \mathbb{R}^d \tag{2.3}$$

where $m_s \in \mathcal{P}(\mathbb{R}^d)$ is the marginal of m at time s, namely $m_s = m \circ (\pi_s)^{-1}$, where $\pi_s : \mathcal{C}([0,\infty) : \mathbb{R}^d) \to \mathbb{R}^d$ is the projection map, $\pi_s(w) \doteq w(s)$. Then (2.1) is same as

$$h(t,x) = Q_t\, h_0\,(x) + \gamma\,\beta\,\Theta_t^\mu(x) \tag{2.4}$$

where $\mu \in \mathcal{P}\Big(\mathcal{C}([0,\infty) :\ \mathbb{R}^d)\Big)$ is the probability law of $X = (X_t)_{t\geq 0}$ defined by the first equation in (1.2). Similarly, the solution h_N of (1.3) can be written as

$$h_N(t,x) = Q_t\, h_0\,(x) + \gamma\,\beta\,\Theta_t^{\mu^N}(x), \tag{2.5}$$

where $\mu^N \doteq \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ is the corresponding empirical measure.

2.1 Wellposedness

This section gives the basic wellposedness results for equations (1.2) and (1.3) under suitable conditions on the dispersal kernel g and initial chemical field h_0. Lemma 2.1 below gives a uniform boundedness and a uniform Lipschitz property for ∇h and ∇h_N. Its proof is straightforward but is included for completeness in Section 4.1.

Lemma 2.1. *Suppose g, $h_0 \in \mathcal{C}_b^2(\mathbb{R}^d)$. Define, for $m \in \mathcal{P}\big(\mathcal{C}([0,\infty) : \mathbb{R}^d)\big)$ and $(t,x) \in [0,\infty) \times \mathbb{R}^d$,*

$$h^m(t,x) = Q_t\, h_0\,(x) + \gamma\, \beta\, \Theta_t^m(x), \tag{2.6}$$

where Θ_t^m is given by (2.3)*. Then there exists $C \in (0,\infty)$ such that*

$$\begin{aligned} \sup_m \sup_{t\geq 0} \sup_{x\in\mathbb{R}^d} |\nabla h^m(t,x)| &\leq C \quad \text{and} \\ \sup_m \sup_{t\geq 0} |\nabla h^m(t,x) - \nabla h^m(t,y)| &\leq C\,|x-y|, \quad \text{for all } x,\, y \in \mathbb{R}^d, \end{aligned}$$

where the outside supremum is taken over all $m \in \mathcal{P}\big(\mathcal{C}([0,\infty) : \mathbb{R}^d)\big)$.

Denote by $\mathcal{C}^*([0,\infty)\times\mathbb{R}^d)$ the class of continuous functions $\zeta : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}$ such that for each $t \geq 0$, $\zeta(t,\cdot)$ is continuously differentiable and $\nabla\zeta(t,\cdot)$ is a bounded Lipschitz function. The space $\mathcal{C}^*([0,T]\times\mathbb{R}^d)$ is defined similarly. The above lemma shows that for an arbitrary $m \in \mathcal{P}\big(\mathcal{C}([0,\infty) : \mathbb{R}^d)\big)$, $h^m \in \mathcal{C}^*([0,\infty)\times\mathbb{R}^d)$. In Section 4.1, using Lemma 2.1 we prove the following proposition which gives the wellposedness of (1.3).

Proposition 2.2. *Suppose g, $h_0 \in \mathcal{C}_b^2(\mathbb{R}^d)$, $V \in \mathcal{C}^1(\mathbb{R}^d)$ and ∇V is Lipschitz. Let $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $\xi_0^N = (\xi_0^{1,N}, \cdots, \xi_0^{N,N})$ be a $\mathcal{F}_0$ measurable square integrable $\mathbb{R}^{dN}$ valued random variable with probability law $\mu_0^{\otimes N}$. Then the system of equations* (1.3) *has a unique pathwise solution $(X^N,\, h_N) \in \mathcal{C}([0,\infty) : \mathbb{R}^{Nd}) \times \mathcal{C}^*([0,\infty)\times\mathbb{R}^d)$ with $(X^N(0),\, h_N(0)) = (\xi_0^N,\, h_0)$.*

With another application of Lemma 2.1 and straightforward modifications of classical fixed point arguments (cf. [23]), we prove the following proposition in Section 4.1 as well.

Proposition 2.3. *Suppose g, $h_0 \in \mathcal{C}_b^2(\mathbb{R}^d)$, $V \in \mathcal{C}^1(\mathbb{R}^d)$ and ∇V is Lipschitz. Let ξ_0 be a $\mathcal{F}_0$ measurable $\mathbb{R}^d$ valued random variable with probability law $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$. Then equation* (1.2) *has a unique pathwise solution $(X,\, h) \in \mathcal{C}([0,\infty) : \mathbb{R}^d) \times \mathcal{C}^*([0,\infty)\times\mathbb{R}^d)$ with $(X(0),\, h(0)) = (\xi_0,\, h_0)$.*

2.2 Assumptions

The following will be our standing assumptions.

- $\gamma = 1$ and $\int_{\mathbb{R}^d} g(x)\,dx = 1$. The second assumption can be made without loss of generality by modifying the value of β whereas the first assumption is for notational convenience; the proofs for a general γ follows similarly.
- The functions $g,\, h_0 \in \mathcal{C}_b^2(\mathbb{R}^d)$.
- The confinement potential $V \in \mathcal{C}^1(\mathbb{R}^d)$, is symmetric, $V(0) = 0$ and ∇V is Lipschitz.
- The initial measure $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$.

The above assumptions will be used without further comment. Note that from the last assumption it follows that $\nabla V(0) = 0$.

In addition, for several results the following convexity assumption will be made. This condition plays an analogous role in the study of the long-time properties of the parabolic-parabolic system as condition (1.9) for the parabolic-elliptic system. Let

$$v_* \doteq \inf_{x\neq y} \frac{\langle x-y, \nabla V(x) - \nabla V(y)\rangle}{|x-y|^2}. \tag{2.7}$$

Note that since ∇V is Lipschitz, $|v_*| \leq L_{\nabla V}$ where latter is the the Lipschitz constant of ∇V. Let

$$\lambda \doteq \left(\| \operatorname{Hess} h_0 \|_\infty + \frac{2\beta \| \operatorname{Hess} g \|_\infty}{\alpha} \right) \chi d.$$

Assumption 2.4. *The confinement potential V is such that*

$$v_* > \lambda. \tag{2.8}$$

A prototypical example of a V that satisfies Assumption 2.4 is $V(x) = \langle x,\, Ax \rangle /2$ where A is a positive definite $d \times d$ matrix with spectrum bounded from below by λ.

3 Main results

3.1 Propagation of chaos

A standard approach to proving POC (see, for instance [23, 17, 9]) is by a coupling method. Let $\{B^i\}_{i=1}^N$ and $\{\xi_0^{i,N}\}_{i=1}^N$ be collections of independent standard d-dimensional $\{\mathcal{F}_t\}$-Brownian motions and $\mathcal{F}_0$ measurable i.i.d. square integrable random variables with probability law $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, respectively. Fix $h_0 \in \mathcal{C}_b^2(\mathbb{R}^d)$. We construct coupled systems $\{X^{i,N}\}_{i=1}^N$ and $\{\bar{X}^i\}_{i=1}^N$ of d-dimensional continuous stochastic processes in such a way that

- $\bar{X}_0^i = X_0^{i,N} = \xi_0^{i,N}$ for all $i = 1, \cdots N$.
- $X^N \doteq (X^{1,N}, \ldots, X^{N,N})$ is the solution to (1.3) with driving Brownian motions $\{B^i\}$ and for each $i = 1, \cdots N$, $\bar{X}^i$ is the solution to (1.2) driven by the Brownian motion B^i.

Using the above coupling we establish the following POC for any finite time horizon. Note this result does not require the convexity assumption (i.e. Assumption 2.4).

Theorem 3.1. *For each $T \geq 0$, there exists $C_T \in (0, \infty)$ such that*

$$\mathbb{E}\Big[\sup_{t \in [0,T]} |X_t^{i,N} - \bar{X}_t^i|^2 \Big] \leq \frac{C_T}{N}.$$

As an immediate consequence of this result we have the following result on asymptotic mutual independence of $(X^{1,N}, \ldots, X^{k,N})$ for each fixed k and the convergence of each $X^{i,N}$ to $\bar{X}^1$ (cf. [23]). The result in particular says that $\mathcal{L}(X^{1,N}, \cdots, X^{N,N})$ is $\mathcal{L}(\bar{X})$-chaotic in the terminology of [23].

Corollary 3.2. *As $N \to \infty$, we have*

$$\mathcal{W}_2\Big(\mathcal{L}(X^{1,N},\, X^{2,N},\, \cdots,\, X^{k,N}),\; \mathcal{L}(\bar{X}^1)^{\otimes k} \Big) \longrightarrow 0$$

for all $k \in \mathbb{N}$, where $\mathcal{W}_2$ is the Wasserstein-2 distance on $\mathcal{P}\left(\mathcal{C}^k\right)$.

Proof. From the definition of the Wasserstein-2 distance and using the fact that $(X^{i,N}, \bar{X}^i)$ has same distribution as $(X^{1,N}, \bar{X}^1)$, for $i = 1, \ldots, N$, we have

$$\mathcal{W}_2\Big(\mathcal{L}(X^{1,N},\, X^{2,N},\, \cdots,\, X^{k,N}),\; \mathcal{L}(\bar{X}^1,\, \bar{X}^2,\, \cdots,\, \bar{X}^k) \Big) \leq \sqrt{k} \sqrt{\mathbb{E}\, d_{\mathcal{C}}(X^{1,N}, \bar{X}^1)^2}.$$

The RHS tends to zero as $N \to \infty$ since $\mathbb{E}\|X^{1,N} - \bar{X}^1\|_{\mathcal{C}_T}^2 \to 0$ for all $T \geq 0$ by Theorem 3.1. The proof is complete since $\{\bar{X}^i\}$ are i.i.d. □

Since $\{X^{i,N}\}_{i=1}^N$ are exchangeable, by [23, Proposition 2.2], we have the following process level weak convergence of empirical distributions.

Corollary 3.3. *As $N \to \infty$, the random measures $\mu^N \doteq \frac{1}{N}\sum_{i=1}^N \delta_{X^{i,N}}$ converge to the deterministic measure $\mathcal{L}(\bar{X})$ in probability in $\mathcal{P}(\mathcal{C})$).*

Observe that Corollary 3.2 implies in particular that

$$\sup_{t\in[0,T]} \mathcal{W}_2\Big(\mathcal{L}(X_t^{1,N},\, X_t^{2,N},\, \cdots,\, X_t^{k,N}),\; \mathcal{L}(\bar{X}_t)^{\otimes k}\Big) \longrightarrow 0 \tag{3.1}$$

for all $T \geq 0$. However this result does not give uniform in time convergence of these multidimensional laws. To obtain a uniform in time result, we will make the stronger assumption in Assumption 2.4. The following is the analog of Theorem 3.1 over an infinite time horizon.

Theorem 3.4. *Suppose Assumption 2.4 is satisfied. Then there exists $C \in (0,\infty)$ such that*

$$\sup_{t\geq 0} \mathbb{E}\left[|X_t^{i,N} - \bar{X}_t^i|^2\right] \leq \frac{C}{N}.$$

As an immediate consequence of the theorem we have the following uniform in time propagation of chaos result and uniform in time convergence of the empirical measures $\mu^N(t)$.

Corollary 3.5. *Suppose Assumption 2.4 is satisfied. Then for all $N, k \in \mathbb{N}$, we have*

$$\sup_{t\geq 0} \mathcal{W}_2\Big(\mathcal{L}(X_t^{1,N},\, X_t^{2,N},\, \cdots,\, X_t^{k,N}),\; \mathcal{L}(\bar{X}_t^1)^{\otimes k}\Big) \leq \frac{C\sqrt{k}}{\sqrt{N}},$$

where $\mathcal{W}_2$ is the Wasserstein-2 distance on $\mathcal{P}(\mathbb{R}^{dk})$. Furthermore, if $\mu_0 \in \mathcal{P}_q(\mathbb{R}^d)$ for some $q > 2$, then

$$\sup_{t\geq 0} \mathbb{E}\left[\mathcal{W}_2^2(\mu_t^N, \mu_t)\right] \to 0$$

as $N \to \infty$, where $\mu^N(t) = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ and $\mu_t = \mathcal{L}(\bar{X}_t^1)$.

Proofs of Theorem 3.1, Theorem 3.4 and Corollary 3.5 are given in Section 4.2.

3.2 Concentration bounds

In this section, we present our concentration estimates for μ_t^N in $\mathcal{W}_1$-distance. As in the previous subsection, we first give a result for finite time horizons (this result will not use Assumption 2.4).

Theorem 3.6. *Suppose the initial distribution $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$ has a finite square-exponential moment, that is, there is $\theta_0 > 0$ such that*

$$\int_{\mathbb{R}^d} e^{\theta_0 |x|^2}\, d\mu_0(x) < \infty. \tag{3.2}$$

Fix $T \in (0,\infty)$. Then there is a $K \in (0,\infty)$ and, for any $d' \in (d,\infty)$, N_0 and C in $(0,\infty)$, such that

$$\mathbb{P}\Big(\sup_{0\leq t\leq T} \mathcal{W}_1(\mu_t^N, \mu_t) > \epsilon\Big) \leq C(1+\epsilon^{-2}) \exp\left(-K\, N\, \epsilon^2\right)$$

for all $N \geq N_0\, \max(\epsilon^{-(d'+2)}, 1)$ and $\epsilon > 0$.

We note that the constant K may depend on T but not on ϵ and d'; also C and N_0 may depend on T and d' but not on ϵ. The main idea in the proof is, as in [4], to (i) bound $\mathcal{W}_1(\mu_t^N, \mu_t)$ in terms of $\big(\mathcal{W}_1(\nu_s^N, \mu_s)\big)_{s\in[0,t]}$ where

$$\nu_s^N \doteq \frac{1}{N}\sum_{i=1}^N \delta_{\bar{X}_s^i}, \tag{3.3}$$

and $\{\bar{X}^i\}_{i=1}^N$ are the processes defined at the beginning of Section 3.1, and then; (ii) estimate $\left(\mathcal{W}_1(\nu_s^N, \mu_s)\right)_{s\in[0,t]}$ which is a quantity that concerns i.i.d. random variables $\{\bar{X}^i\}_{i=1}^N$. The first step is accomplished in Subsection 4.3.1 via a coupling argument similar to the one used in the proof of results in Section 3.1, while the second step relies on an estimate from [4] for the tail probabilities for empirical measures of i.i.d. random variables that is based on the equivalence between Talagrand's transportation inequalities (cf. [4]) and existence of a finite square-exponential moment. The precise result obtained in [4] is as follows. For $a, \alpha \in (0,\infty)$ we let

$$\mathcal{P}_{a,\alpha} \doteq \{\nu \in \mathcal{P}(\mathbb{R}^d) : \int_{\mathbb{R}^d} e^{\alpha\,|x|^2}\, d\nu(x) \le a\}.$$

Theorem 3.7. *[4, Theorem 2.1] Fix $a, \alpha \in (0,\infty)$. Then, there is a $\theta > 0$ such that for any $d' \in (d,\infty)$ there exists a positive integer N_0 such that*

$$\sup_{\nu\in\mathcal{P}_{a,\alpha}} \mathbb{P}\left(\mathcal{W}_1(\hat{\nu}^N,\, \nu) > \epsilon\right) \le e^{-\frac{\theta}{2}\,N\,\epsilon^2}.$$

for all $\epsilon > 0$ and $N \ge N_0 \max(\epsilon^{-(d'+2)},\, 1)$, where $\hat{\nu}^N \doteq \frac{1}{N}\sum_{i=1}^N \delta_{Z^i}$ and $(Z^i)_{i\in\mathbb{N}}$ are iid random variables with law ν.

We shall apply this theorem to $\nu = \mu_s$. In order to do so, we need μ_s to have a finite squared-exponential moment. We will show in Section 4.3.2 that if μ_0 satisfies (3.2), then for every $T > 0$ there is a $\theta_T \in (0, \theta_0)$ such that

$$\sup_{s\in[0,T]} \int_{\mathbb{R}^d} e^{\theta_T\,|x|^2}\, d\mu_s(x) < \infty. \tag{3.4}$$

This will allow us to apply Theorem 3.7 in completing step (ii) in the proof of Theorem 3.6.

We next show in Theorem 3.8 below that when the convexity property in Assumption 2.4 is satisfied then a *uniform in time* concentration bound holds. The key step (see Section 4.3.2) is to argue (see Proposition 4.3) that under this assumption, for some $\theta_\infty > 0$, (3.4) holds with $[0,T]$ and θ_T replaced with $[0,\infty)$ and θ_∞ respectively. This together with another uniform bound established in Section 4.3.1 (Proposition 4.1) will imply the uniform in time concentration bound given in the theorem below.

Theorem 3.8. *Suppose that μ_0 satisfies* (3.2) *for some $\theta_0 > 0$. Suppose further that Assumption 2.4 is satisfied. Then there exists $K \in (0,\infty)$ such that for any $d' \in (d,\infty)$, there exist $C \in (0,\infty)$ and $N_0 \in (0,\infty)$ such that*

$$\sup_{t\ge 0} \mathbb{P}\left(\mathcal{W}_1(\mu_t^N, \mu_t) > \epsilon\right) \le C(1+\epsilon^{-2})\, \exp\left(-K\,N\,\epsilon^2\right)$$

for all $N \ge N_0 \max(\epsilon^{-(d'+2)}, 1)$ and $\epsilon > 0$.

We note that C and N_0 may depend on d' but not on ϵ.

Proofs of Theorems 3.6 and 3.8 will be given in Section 4.3.

3.3 Uniform convergence of Euler scheme

In this section we will introduce an Euler approximation for the collection of SDE in (1.3) which can be used for approximate simulation of the system. We show that the approximation error converges to 0 as the time discretization size ϵ converges to 0, *uniformly in time*. As a consequence it will follow that the empirical measure of the particle states in the approximate system converges to the law of the nonlinear process, uniformly in time, as $N \to \infty$ and $\epsilon \to 0$ (Corollary 3.11).

Note that Q_t has transition density $q(t,x,y) = e^{-\alpha t}p(t,x,y)$ with respect to Lebesgue measure, where $p(t,x,y)$ is the standard Gaussian kernel. Using (2.5), the system of equations governing the particle system $X^{(N)} = (X^{i,N})_{i=1}^N$ in (1.3) can be written as

$$dX_t^{i,N} = dB_t^i + \Big(\int_0^t G_{t-s}(X_s^{(N)}, X_t^{i,N})\, ds - \nabla V_t(X_t^{i,N}) \Big)\, dt \tag{3.5}$$

for $1 \le i \le N$, where

$$V_t = V - \chi Q_t h_0 \tag{3.6}$$

and for $\vec{x} = (x_1, \dots, x_N) \in \mathbb{R}^{dN}$ and $y \in \mathbb{R}^d$,

$$G_\theta(\vec{x}, y) = \frac{\chi\beta}{N} \sum_{i=1}^N \int_{\mathbb{R}^d} \nabla_y q(\theta, y, z) g(x_i - z)\, dz. \tag{3.7}$$

We now define an explicit Euler scheme for (3.5) with step size $\epsilon \in (0,1)$. Let $Y_0^{(N),\epsilon} = X_0^{(N)}$. Having defined $Y_n^{(N),\epsilon} = (Y_n^{i,N,\epsilon})_{i=1}^N$ for some $n \ge 0$, we define $Y_{n+1}^{(N),\epsilon}$ naturally as

$$Y_{n+1}^{i,N,\epsilon} \doteq Y_n^{i,N,\epsilon} + \Delta_n B^i + \epsilon \Big(\int_0^{n\epsilon} G_{n\epsilon - s}(\tilde{Y}_s^{(N),\epsilon},\, Y_n^{i,N,\epsilon})\, ds - \nabla V_{n\epsilon}(Y_n^{i,N,\epsilon}) \Big) \tag{3.8}$$

for $1 \le i \le N$, where $\Delta_n B^i \doteq B^i_{(n+1)\epsilon} - B^i_{n\epsilon}$ and

$$\tilde{Y}_s^{(N),\epsilon} \doteq Y_k^{(N),\epsilon} \quad \text{for } s \in [k\epsilon, (k+1)\epsilon).$$

Note that the integral on the right hand side of (3.8) can be written as

$$\int_0^{n\epsilon} G_{n\epsilon - s}(\tilde{Y}_s^{(N),\epsilon},\, Y_n^{i,N,\epsilon})\, ds = \sum_{k=0}^{n-1} \int_{k\epsilon}^{(k+1)\epsilon} G_{n\epsilon - s}(Y_k^{(N),\epsilon},\, Y_n^{i,N,\epsilon})\, ds.$$

Thus in order to evaluate a typical Euler step, one needs to compute terms of the form $\int_{[k\epsilon,(k+1)\epsilon]} G_\theta(\vec{x}, y) d\theta$ which can be done using numerical integration.

Our goal is to provide uniform in time estimates on the mean square error of the scheme, namely to estimate the quantity

$$\mathbb{E}\big| Y_n^{i,N,\epsilon} - X_{n\epsilon}^{i,N} \big|^2.$$

For that we begin by establishing moment bounds for the Euler scheme which are uniform in N, step size ϵ and time instant n. Recall v_* introduced in (2.7). Also recall that $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$.

Lemma 3.9. *Suppose $v_* > 0$. Then there exists $\epsilon_0 \in (0,1)$ such that*

$$\sup_{\epsilon \in (0,\, \epsilon_0)} \sup_{N \in \mathbb{N}_0} \sup_{1 \le i \le N} \sup_{n \ge 0} \mathbb{E}|Y_n^{i,N,\epsilon}|^2 < \infty.$$

We now present our main result on the uniform convergence of the Euler scheme. For this result we will make the stronger convexity assumption in Assumption 2.4.

Theorem 3.10. *Suppose Assumption 2.4 holds. Then there exists $\epsilon_0 \in (0,1)$ and $C \in (0,\infty)$ such that*

$$\sup_{N \ge 1} \sup_{1 \le i \le N} \sup_{n \in \mathbb{N}_0} \mathbb{E}\big| Y_n^{i,N,\epsilon} - X_{n\epsilon}^{i,N} \big|^2 \le C\, \epsilon \tag{3.9}$$

for all $\epsilon \in (0,\, \epsilon_0)$.

It is important that the estimate in (3.9) is uniform not only in time instant n but also in the size of the system N. As a consequence one has the desired property that in order to control the mean square error for larger systems one does not need smaller time discretization steps. This in particular implies that the Euler scheme provides a good numerical approximation to the nonlinear process, uniformly in time. Namely we have the following result.

Corollary 3.11. *Suppose Assumption 2.4 holds. There exists $\epsilon_0 \in (0,1)$ and $C \in (0,\infty)$ such that for any positive integer k,*

$$\sup_{n\geq 0} \mathcal{W}_2\Big(\mathcal{L}(Y_n^{1,N,\epsilon},\, Y_n^{2,N,\epsilon},\, \cdots,\, Y_n^{k,N,\epsilon}),\ \mathcal{L}(\bar{X}_{n\epsilon})^{\otimes k}\Big) \leq C\sqrt{k}\left(\sqrt{\epsilon} + \frac{1}{\sqrt{N}}\right)$$

for all $\epsilon \in (0,\epsilon_0)$. Furthermore if $\mu_0 \in \mathcal{P}_q(\mathbb{R}^d)$ for some $q \in (2,\infty)$. Then we have

$$\limsup_{N\to\infty} \sup_{n\geq 1} \mathbb{E}[\mathcal{W}_2^2(\mu_n^{N,\epsilon}, \mu_{n\epsilon})] \leq 2\,C\epsilon \tag{3.10}$$

for all $\epsilon \in (0,\, \epsilon_0)$, where $\mu_n^{N,\epsilon} \doteq \frac{1}{N}\sum_{i=1}^N \delta_{Y_n^{i,N,\epsilon}}$.

Proofs of Lemma 3.9, Theorem 3.10 and Corollary 3.11 will be given in Section 4.4.

4 Proofs

We will denote by $\kappa,\, \kappa_1,\, \kappa_2, \cdots$ the constants that appear in various estimates within a proof. These constants only depend on the model parameters or problem data, namely, $\alpha,\, \beta,\, \chi,\, g,\, V,\, h_0, d$ and μ_0. For estimates on a finite time horizon $[0,T]$, these constants may also depend on T and in that case we use $\kappa_T,\, \kappa_{1,T},\, \kappa_{2,T}, \cdots$ to denote such constants. The value of such constants may change from one proof to another.

4.1 Wellposedness

Proof of Lemma 2.1. From the definition of h^m in (2.6) and of the semigroup $\{Q_t\}$, we have for all $x, x_1, x_2 \in \mathbb{R}^d$ and $t \geq 0$, uniform bounds

$$\begin{aligned}
|h^m(t,x)| &\leq e^{-\alpha t}|P_t h_0(x)| + \frac{\beta(1-e^{-\alpha t})\|g\|_\infty}{\alpha} \\
&\leq e^{-\alpha t}\|h_0\|_\infty + \frac{\beta\,\|g\|_\infty}{\alpha}, \\
|\nabla h^m(t,x)| &\leq e^{-\alpha t}\|\nabla h_0\|_\infty + \frac{\beta\,\|\nabla g\|_\infty}{\alpha}, \\
|\nabla h^m(t,x_1) - \nabla h^m(t,x_2)| &\leq |x_1 - x_2|\, d\left(e^{-\alpha t}\|\operatorname{Hess} h_0\|_\infty + \frac{\beta\,\|\operatorname{Hess} g\|_\infty}{\alpha}\right).
\end{aligned}$$

The result is immediate from the above inequalities. □

Proof of Proposition 2.2. For notational simplicity, we suppress the index N and write $X^{N,i}$ as X_i.

Uniqueness. Suppose $(X,\, h)$ and $(\tilde{X},\, \tilde{h})$ are two solutions to (1.3) with $h(0,\cdot) = \tilde{h}(0,\cdot) = h_0$, and $X(0) = \tilde{X}(0)$, where $X = (X_1, \cdots,\, X_N)$ and $\tilde{X} = (\tilde{X}_1, \cdots,\, \tilde{X}_N)$. Letting $Y_i \doteq X_i - \tilde{X}_i$ and $H \doteq h - \tilde{h}$, we have

$$Y_i(t) = \int_0^t \Big(-\nabla V(X_i(s)) + \nabla V(\tilde{X}_i(s)) + \chi\nabla h(s, X_i(s)) - \chi\nabla\tilde{h}(s,\tilde{X}_i(s))\Big)\, ds, \tag{4.1}$$

$$H(t,x) = \frac{\beta}{N}\sum_{i=1}^N \int_0^t \Big(Q_{t-s}\big(g(X_i(s)-\cdot) - g(\tilde{X}_i(s)-\cdot)\big)(x)\Big)\, ds. \tag{4.2}$$

From (4.1) we have for all $t \geq 0$

$$\begin{aligned}\sup_{0\leq s\leq t}|Y_i(s)| &\leq d\int_0^t \Big(\|\operatorname{Hess} V\|_\infty \big|Y_i(s)\big| + \big|\nabla H(s,\tilde{X}_i(s))\big| \\ &\qquad + \chi\big|\nabla h(s,X_i(s)) - \nabla h(s,\tilde{X}_i(s))\big|\Big)\, ds \\ &\leq \kappa_1 \int_0^t \big(\big|Y_i(s)\big| + \|\nabla H(s)\|_\infty\big)\, ds, \end{aligned} \tag{4.3}$$

where $\|\nabla H(s)\|_\infty \doteq \sup_x |\nabla H(s,x)|$ and the last inequality follows from Lemma 2.1 on noting that h equals h^{μ^N}, where $\mu^N \doteq \frac{1}{N}\sum_{i=1}^N \delta_{X_i}$ is the path empirical measure (see (2.5)). From (4.2), the fact that $\nabla_x p(t,x,y) = -\nabla_y p(t,x,y)$ and integration by parts, we obtain

$$\nabla H(t,x) = \frac{\beta}{N}\sum_{i=1}^N \int_0^t e^{-\alpha(t-s)} \int_{\mathbb{R}^d} p((t-s),x,y)\nabla_y\big(g(X_i(s)-y) - g(\tilde{X}_i(s)-y)\big)\, dy\, ds.$$

Hence since $g \in \mathcal{C}_b^2(\mathbb{R}^d)$,

$$\|\nabla H(t)\|_\infty \leq \kappa_2 \int_0^t e^{-\alpha(t-s)} \max_{1\leq i\leq N}\big|Y_i(s)\big| ds. \tag{4.4}$$

Combining (4.3) and (4.4), and letting $Y(t) \doteq \max_{1\leq i\leq N}\big|Y_i(t)\big|$, we obtain

$$\begin{aligned}\sup_{0\leq s\leq t} Y(s) &\leq \kappa_3 \int_0^t \Big(Y(s) + \Big(\int_0^s e^{-\alpha(s-r)} Y(r)\, dr\Big)\Big)\, ds \\ &\leq \kappa_{4,t}\int_0^t \sup_{0\leq r\leq s} Y(r) ds. \end{aligned} \tag{4.5}$$

This implies $Y(t) = 0$ for all $t \geq 0$. Finally, from (4.2) we have $H(t,x) = 0$ for all $t \geq 0$ and $x \in \mathbb{R}^d$. This completes the proof of pathwise uniqueness.

Existence. This is argued by a minor modification of the standard Picard approximation method as follows. Define a sequence $\{(X^{(k)}, h^{(k)})\}_{k\geq 1}$, where $X^{(k)} = (X_1^{(k)}, \cdots, X_N^{(k)})$, of $\mathcal{C}([0,\infty) : \mathbb{R}^{Nd}) \times \mathcal{C}^*([0,\infty)\times\mathbb{R}^d)$ valued random variables as follows. Let $X^{(1)}(t) = (\xi_0^{1,N}, \cdots, \xi_0^{N,N})$ and $h^{(1)}(t,x) = h_0(x)$ for all t. We then define, for $k \geq 2$,

$$\begin{aligned} X_i^{(k+1)}(t) &\doteq \xi_0^{i,N} + \int_0^t \Big(-\nabla V(X_i^{(k)}(s)) + \nabla h^{(k)}(s, X_i^{(k)}(s))\Big)\, ds + B_t^i, \quad i = 1,\ldots,N, \\ h^{(k+1)}(t,x) &\doteq Q_t\, h_0\,(x) + \frac{\beta}{N}\sum_{i=1}^N \int_0^t Q_{t-s}\big(g(X_i^{(k+1)}(s) - \cdot)\big)(x)\, ds, \end{aligned}$$

Let $Y^{(k)}(t) \doteq \max_{1\leq i\leq N}\big|X_i^{(k+1)}(t) - X_i^{(k)}(t)\big|$. By similar estimates as that were used to obtain (4.5), we have

$$\sup_{0\leq s\leq t} Y^{(k)}(s) \leq \kappa_5 \int_0^t \Big(Y^{(k-1)}(s) + \Big(\int_0^s e^{-\alpha(s-r)} Y^{(k-1)}(r)\, dr\Big)\Big)\, ds.$$

Hence, for fixed $T > 0$ and $t \in [0,T]$,

$$\Big(\sup_{0\leq s\leq t} Y^{(k)}(s)\Big)^2 \leq \kappa_{6,T}\int_0^t (Y^{(k-1)}(s))^2\, ds.$$

A standard iteration argument then yields

$$\mathbb{E}\Big[\Big(\sup_{0\le s\le t} Y^{(k+1)}(s)\Big)^2\Big] \le C_T\,\frac{(\kappa_{7,T})^k}{k!};\quad 0\le t\le T,\, k\ge 1,$$

where $C_T = \max_{1\le i\le N}\sup_{0\le t\le T}\mathbb{E}[|X_i^{(2)}(t)-X_i^{(1)}(t)|^2]$ is finite from the uniform boundedness of ∇h_N proved in Lemma 2.1 and the Lipschitz property of ∇V. From this we conclude that $\{X^{(k)}(t)\}_{0\le t\le T}$ converges a.s in $\mathcal{C}([0,T]:\mathbb{R}^{Nd})$ to a continuous process $\{X(t)\}_{0\le t\le T}$. On other hand, using estimates similar to those used in obtaining (4.4) we have

$$\|h^{(k+1)}(t)-h^{(k)}(t)\|_\infty + \|\nabla h^{(k+1)}(t)-\nabla h^{(k)}(t)\|_\infty \le \kappa_8\int_0^t e^{-\alpha(t-s)}\sup_{0\le r\le s} Y^{(k)}(r)\,ds.$$

From this it follows that for every $t\in[0,T]$, $h^{(k)}(t,\cdot)$ converges uniformly to a continuously differentiable function $h(t,\cdot)$ and $\nabla h^{(k)}(t,\cdot)$ converges uniformly to $\nabla h(t,\cdot)$. Furthermore the convergence is uniform in $t\in[0,T]$, namely

$$\sup_{0\le t\le T}\Big(\|h^{(k)}(t)-h(t)\|_\infty + \|\nabla h^{(k)}(t)-\nabla h(t)\|_\infty\Big)\to 0$$

as $k\to\infty$. It is easy to verify that $h\in\mathcal{C}^*([0,\infty)\times\mathbb{R}^d)$ and

$$X_i(t) = \xi_0^{i,N} + \int_0^t \left(-\nabla V(X_i(s)) + \nabla h(s,X_i(s))\right)ds + B_t^i,\quad i=1,\dots,N,$$

$$h(t,x) = Q_t\,h_0\,(x) + \frac{\beta}{N}\sum_{i=1}^N\int_0^t Q_{t-s}\big(g(X_i(s)-\cdot)\big)(x)\,ds.$$

This establishes the desired existence of solutions. □

Proof of Proposition 2.3. The proof uses classical arguments from [23]. It suffices to show that for each $T>0$, equation (1.2) has a unique solution over the time horizon $[0,T]$ which belongs to $\mathcal{C}([0,T]:\mathbb{R}^d)\times\mathcal{C}^*([0,T]\times\mathbb{R}^d)$. Let $T>0$ be arbitrary. We note that a probability measure $m\in\mathcal{P}(\mathcal{C}^T)$ can be mapped naturally to a $\hat{m}\in\mathcal{P}(\mathcal{C}([0,\infty):\mathbb{R}^d))$ as $\hat{m}\doteq m\circ[\pi^T]^{-1}$ where $\pi^T:\mathcal{C}([0,T]:\mathbb{R}^d)\to\mathcal{C}([0,\infty):\mathbb{R}^d)$ is defined as $(\pi^T w)(s)\doteq w(s\wedge T)$ for $s\ge 0$, $w\in\mathcal{C}([0,T]:\mathbb{R}^d)$. Abusing notation we denote for $t\in[0,T]$ $\Theta_t^{\hat{m}}$ as Θ_t^m where Θ_t^m is defined in (2.3).

Define $\Phi:\mathcal{P}(\mathcal{C}^T)\to\mathcal{P}(\mathcal{C}^T)$ which maps m to the law $\mathcal{L}(Z)$, where $Z=(Z_t)_{t\in[0,T]}$ is the solution of

$$Z_t = \xi_0 + B_t + \int_0^t\left[-\nabla V(Z_s) + \chi\,\nabla h^m(s,Z_s)\right]ds,\quad t\in[0,T], \tag{4.6}$$

and h^m is as in (2.6). From Lemma 2.1 ∇h^m is a Lipschitz map and by assumption ∇V is Lipschitz as well, thus the equation in (4.6) has a unique solution and consequently the function Φ is well-defined. Observe that $(X,\,h)$ is a solution of (1.2) over $[0,T]$ if and only if $\mathcal{L}(X)\in\mathcal{P}(\mathcal{C}^T)$ is a fixed point of Φ and h is given by the right hand side of (2.1) with μ_t being the law of X_t. We will show that for all $m^1,\,m^2\in\mathcal{P}(\mathcal{C}^T)$, we have

$$D_t(\Phi(m^1),\Phi(m^2)) \le \kappa_T\int_0^t D_s(m^1,m^2)\,ds,\quad t\in[0,T], \tag{4.7}$$

for some $\kappa_T\in(0,\infty)$ where D_t is the Wasserstein-1 distance on $\mathcal{P}(\mathcal{C}^t)$, namely, for $m^1,m^2\in\mathcal{P}(\mathcal{C}^t)$, $D_t(m^1,m^2)$ is given by the right side of (1.11) with $p=1$ and $S=\mathcal{C}^t$.

Suppose for $i = 1, 2$, Z^i solves (4.6) with m replaced by m^i on the right side where $m^i \in \mathcal{P}(\mathcal{C}^T)$. Then, Z^i has law $\Phi(m^i)$ and

$$\begin{aligned}\sup_{s\le t}|Z_s^1 - Z_s^2| \quad \le \quad & \int_0^t \Big[|\nabla V(Z_s^2) - \nabla V(Z_s^1)| + \chi\,|\nabla Q_s h_0(Z_s^1) - \nabla Q_s h_0(Z_s^2)| \\ & +\beta\,\chi\,|\nabla \Theta_s^{m^1}(Z_s^1) - \nabla \Theta_s^{m^2}(Z_s^2)|\Big]\, ds \end{aligned} \tag{4.8}$$

From properties of the heat semigroup $|\nabla Q_s h_0(x) - \nabla Q_s h_0(y)| \le e^{-\alpha s}|x-y| d \|\operatorname{Hess} h_0\|_\infty$ and

$$\begin{aligned}
&|\nabla \Theta_t^{m^1}(x) - \nabla \Theta_t^{m^2}(y)| \\
&\le \Big| \int_0^t \int_{\mathbb{R}^d} q(t-s,x,z)\langle \nabla_z g(\cdot - z), m_s^1\ - q(t-s,y,z)\langle \nabla_z g(\cdot - z), m_s^2\ dz\, ds \Big| \\
&\le \Big| \int_0^t \int_{\mathbb{R}^d} q(t-s,x,z)\langle \nabla_z g(\cdot - z), m_s^1\ - q(t-s,y,z)\langle \nabla_z g(\cdot - z), m_s^1\ dz\, ds \Big| \\
&\quad + \int_0^t \int_{\mathbb{R}^d} q(t-s,y,z)\,\big|\langle \nabla_z g(\cdot - z),\ m_s^1 - m_s^2\,\big|\, dz\ ds \\
&\le d\|\operatorname{Hess} g\|_\infty |x-y| \int_0^t e^{-\alpha(t-s)}\, ds \\
&\quad + \int_0^t \int_{\mathbb{R}^d} q(t-s,y,z)\,\big|\langle \nabla_z g(\cdot - z),\ m_s^1 - m_s^2\,\big|\, dz\ ds \\
&\le \kappa\left(|x-y| + \int_0^t e^{-\alpha(t-s)} \int_{\mathbb{R}^d\times\mathbb{R}^d} |w_1(s) - w_2(s)|\, dM(w_1,w_2)\ ds\right),
\end{aligned}$$

for any $M \in \mathcal{P}(\mathcal{C}^t \times \mathcal{C}^t)$ with marginals m^1 and m^2, where the last step uses the fact that $y \mapsto \nabla_z g(y-z)$ is Lipschitz. Combining the above estimates with (4.8) and using the Lipschitz property of ∇V, we obtain

$$\sup_{s\le t}|Z_s^1 - Z_s^2| \le \kappa_1 \int_0^t \left[|Z_s^1 - Z_s^2| + \int_{\mathbb{R}^d\times\mathbb{R}^d} \sup_{r\le s}|w_1(r) - w_2(r)|\, dM(w_1,w_2)\right]\, ds$$

for any M as above. Hence

$$\sup_{s\le t}|Z_s^1 - Z_s^2| \le \kappa_2 \int_0^t \left[|Z_s^1 - Z_s^2| + D_s(m^1,m^2)\right]\, ds.$$

By Gronwall's Lemma, it now follows that

$$\sup_{s\le t}|Z_s^1 - Z_s^2| \le \kappa_{3,T} \int_0^t D_s(m^1,m^2) ds, \quad t \le T,$$

Taking expectations we obtain (4.7). Now by a standard fixed point argument, there exists a unique $m^* \in \mathcal{P}(\mathcal{C}^T)$ such that $m^* = \Phi(m^*)$. Let Z^* be the unique solution to (4.6) with m replaced by m^*. Then (1.2) has a unique pathwise solution $(Z^*,\, h^*)$ where h^* is given by the right hand side of (2.1) with $\mu_t(dy)$ being the law of Z_t^*. □

4.2 Propagation of chaos

The proofs of Theorems 3.1 and 3.4 are based on breaking the pathwise deviation

$$\Delta_s^i \doteq X_s^{i,N} - \bar{X}_s^i$$

into several manageable terms. For any N and i, integration by parts yields

$$d|\Delta_t^i|^2 = 2\,\Delta_t^i \cdot \Big[-\nabla V(X_t^{i,N}) + \nabla V(\bar{X}_t^i) + \chi\left(\nabla h^{\mu^N}(t, X_t^{i,N}) - \nabla h^{\mu}(t, \bar{X}_t^i)\right)\Big]\, dt$$

Note that, with v_* as in (2.7)

$$\Delta_t^i \cdot \Big[-\nabla V(X_t^{i,N}) + \nabla V(\bar{X}_t^i) \Big] \le -v_* |\Delta_t^i|^2.$$

Next, from (2.6)

$$\begin{aligned}
&\nabla h^{\mu^N}(t, X_t^{i,N}) - \nabla h^{\mu}(t, \bar{X}_t^i) \\
&\quad = \Big(\nabla Q_t h_0(X_t^{i,N}) - \nabla Q_t h_0(\bar{X}_t^i) \Big) \\
&\qquad + \beta \int_0^t \left[\nabla Q_{t-s} \left(\int g(y - \cdot)\, \mu_s^N(dy) \right) (X_t^{i,N}) - \nabla Q_{t-s} \left(\int g(y - \cdot)\, \mu_s(dy) \right) (\bar{X}_t^i) \right] ds.
\end{aligned} \tag{4.9}$$

Since $h_0 \in \mathcal{C}_b^2(\mathbb{R}^d)$, for the first term on the right side of (4.9) we have

$$\Delta_t^i \cdot \Big(\nabla Q_t h_0(X_t^{i,N}) - \nabla Q_t h_0(\bar{X}_t^i) \Big) \le e^{-\alpha t} d \| \operatorname{Hess} h_0 \|_\infty |\Delta_t^i|^2.$$

Next note that for any $m \in \mathcal{P}(\mathbb{R}^d)$,

$$\nabla Q_{t-s} \Big(\int g(y - \cdot)\, m(dy) \Big)(x) = e^{-\alpha(t-s)} \int \nabla P_{t-s} g(y - x)\, m(dy).$$

For the second term on the right side in (4.9) we will use the decomposition

$$\begin{aligned}
&\int \nabla P_{t-s} g(y - X_t^{i,N})\, \mu_s^N(dy) - \int \nabla P_{t-s} g(y - \bar{X}_t^i)\, \mu_s(dy) \\
&\quad = \int \nabla P_{t-s} g(y - X_t^{i,N})\, \mu_s^N(dy) - \int \nabla P_{t-s} g(y - \bar{X}_t^i)\, \nu_s^N(dy) \\
&\qquad + \int \nabla P_{t-s} g(y - \bar{X}_t^i)\, \nu_s^N(dy) - \int \nabla P_{t-s} g(y - \bar{X}_t^i)\, \mu_s(dy),
\end{aligned}$$

where ν^N is the empirical measure of $\{\bar{X}^i\}_{i=1}^N$, i.e. $\nu^N = \frac{1}{N} \sum_{i=1}^N \delta_{\bar{X}^i}$, and ν_t^N is the marginal at time instant t.

From the above observations we have

$$\begin{aligned}
d|\Delta_t^i|^2 \le\ & 2 \Big[-v_* + \chi\, e^{-\alpha t} d \, \| \operatorname{Hess}(h_0) \|_\infty \Big] |\Delta_t^i|^2\, dt \\
&+ 2\chi\beta \int_0^t \Big(e^{-\alpha(t-s)} \Delta_t^i \cdot \Big(\int \nabla P_{t-s} g(y - X_t^{i,N})\, \mu_s^N(dy) \\
&- \int \nabla P_{t-s} g(y - \bar{X}_t^i)\, \nu_s^N(dy) \Big) \Big) ds\, dt \\
&+ 2\chi\beta \int_0^t e^{-\alpha(t-s)} \Delta_t^i \cdot \Big(\int \nabla P_{t-s} g(y - \bar{X}_t^i)\, (\nu_s^N - \mu_s)(dy) \Big) ds\, dt,
\end{aligned} \tag{4.10}$$

where the above inequality is interpreted in the integral sense. That is, $d\phi_t \le \psi_t\, dt$ is interpreted as $\phi_b - \phi_a \le \int_a^b \psi_s\, ds$ for all $0 \le a \le b$. In second term on the right of (4.10), the integrand has absolute value

$$\begin{aligned}
&2\chi\beta \frac{e^{-\alpha(t-s)}}{N} \Big| \sum_j \Delta_t^i \cdot \big[\nabla P_{t-s} g(X_s^{j,N} - X_t^{i,N}) - \nabla P_{t-s} g(\bar{X}_s^j - \bar{X}_t^i) \big] \Big| \\
&\le\ 2\chi\beta \frac{e^{-\alpha(t-s)}\, d \| \operatorname{Hess} g \|_\infty}{N} \sum_j |\Delta_t^i| (|\Delta_t^i| + |\Delta_s^j|).
\end{aligned} \tag{4.11}$$

For the third term on the right of (4.10), we let

$$a^{i,j}(s,t) \doteq \big(\nabla P_{t-s} g(\bar{X}^j_s - \bar{X}^i_t) - \int \nabla P_{t-s} g(y - \bar{X}^i_t)\, \mu_s(dy) \big).$$

Then $\mathbb{E}[a^{i,j}(s,t)\, a^{i,k}(s,t)] = 0$ for all $0 \le s \le t$ whenever $j \neq k$ and hence

$$\begin{aligned}
\mathbb{E}\Big[\Big| \int \nabla P_{t-s} g(y - \bar{X}^i_t)\, (\nu^N_s - \mu_s)(dy) \Big|^2 \Big] &= \frac{1}{N^2} \mathbb{E}\Big[\Big| \sum_j a^{i,j}(s,t) \Big|^2 \Big] \\
&= \frac{1}{N^2} \mathbb{E}\Big[\sum_{j,k} a^{i,j}(s,t) \cdot a^{i,k}(s,t) \Big] \\
&= \frac{1}{N^2} \mathbb{E}\Big[\sum_j |a^{i,j}(s,t)|^2 \Big] \le \frac{2\, \|\nabla g\|^2_\infty}{N}. \qquad (4.12)
\end{aligned}$$

The proofs of Theorems 3.1 and 3.4 will make use of the above calculations.

Proof of Theorem 3.1. Fix $T \in (0,\infty)$. Letting $f^i(t) \doteq \sup_{s\in[0,t]} |\Delta^i_t|^2$, we have from (4.10) on noting that $|v_*| \le L_{\nabla V}$, for all $T_1 \in [0,T]$,

$$\begin{aligned}
f^i(T_1) &\le 2(\chi d \| \operatorname{Hess} h_0\|_\infty + L_{\nabla V}) \int_0^{T_1} f^i(t) dt \\
&\quad + \frac{2\chi\beta d \| \operatorname{Hess} g\|_\infty}{N} \sum_j \int_0^{T_1} \sqrt{f^i(t)} \int_0^t (\sqrt{f^i(t)} + \sqrt{f^j(s)}) ds\, dt \\
&\quad + \frac{2\chi\beta}{N} \int_0^{T_1} \sqrt{f^i(t)} \int_0^t |\sum_j a^{i,j}(s,t)| ds\, dt \\
&\le \kappa \int_0^{T_1} \left[f^i(t) + t \left(f^i(t) + \sqrt{f^i(t)} \frac{\sum_j \sqrt{f^j(t)}}{N} \right) + \sqrt{f^i(t)} \int_0^t \frac{|\sum_j a^{i,j}(s,t)|}{N}\, ds \right] dt.
\end{aligned}$$

Taking expectation, using Cauchy-Schwarz inequality, (4.12) and the fact that

$$\vartheta(t) \doteq \frac{\sum_i \mathbb{E} f^i(t)}{N} \ge \Big(\frac{1}{N} \sum_i \sqrt{\mathbb{E} f^i(t)} \Big)^2,$$

we obtain

$$\vartheta(T_1) \le \kappa_1 \int_0^{T_1} \left[(1+2t)\, \vartheta(t) + \frac{t\sqrt{\vartheta(t)}}{\sqrt{N}} \right] dt$$

and hence

$$\vartheta(T_1) \le \kappa_{2,T} \int_0^{T_1} \left[\vartheta(t) + \frac{\sqrt{\vartheta(t)}}{\sqrt{N}} \right] dt$$

for all $T_1 \in [0,T]$. Note that $\Phi(z) \doteq z + \frac{\sqrt{z}}{\sqrt{N}}$, $z \in \mathbb{R}_+$, is an increasing function. From Bihari's generalization of Gronwall's lemma (see Section 3 of [3])

$$\vartheta(t) \le A^{-1}\left(A(0) + \kappa_{2,T} t \right), \text{ for all } t \in [0,T],$$

where

$$A(u) = \int_0^u \frac{ds}{\Phi(s)}, \ u \ge 0.$$

Observing that

$$A(u) = 2\log\left(\frac{N^{-1/2}+u^{1/2}}{N^{-1/2}}\right) \text{ and } A^{-1}(v) = \Big((N^{-1/2})(e^{v/2}-1)\Big)^2,$$

we see that for all $t \in [0,T]$,

$$\vartheta(t) \le \left(\frac{(e^{\kappa_{2,T}\, t/2}-1)}{\sqrt{N}}\right)^2 \le \frac{\kappa_{3,T}}{N} \quad \text{for } t \in [0,T].$$

The proof is complete. □

Proof of Theorem 3.4. For $i,j = 1,\cdots N$ and $0 \le s \le t$

$$\mathbb{E}[\,|\Delta_t^i|(|\Delta_t^i| + |\Delta_s^j|)\,] \le \mathbb{E}|\Delta_t^i|^2 + \sqrt{\mathbb{E}|\Delta_t^i|^2\,\mathbb{E}|\Delta_s^j|^2}. \tag{4.13}$$

Combining (4.10), (4.11), (4.12) and (4.13), we see that $f(t) \doteq \mathbb{E}|\Delta_t^i|^2$ satisfies

$$\begin{aligned} df(t) \;&\le\; 2\,\big(-v_* \,+\, \chi d\,\|\operatorname{Hess} h_0\|_\infty\big)\, f(t)\, dt \\ &\quad + 2\,\chi\,\beta \int_0^t e^{-\alpha(t-s)} d\|\operatorname{Hess} g\|_\infty \big(f(t) + \sqrt{f(t)\,f(s)}\,\big)\, ds\, dt \\ &\quad + \frac{2\,\chi\,\beta}{N} \int_0^t e^{-\alpha(t-s)}\, \sqrt{f(t)\, 2\,\|\nabla g\|_\infty^2\, N}\, ds\, dt \\ &\le\; 2\,\big(-v_* \,+\chi C_1\big)\, f(t)\, dt \\ &\quad + 2\,C_2\,\chi\,\beta \int_0^t e^{-\alpha(t-s)} \big(f(t) + \sqrt{f(t)\,f(s)}\,\big)\, ds\, dt \\ &\quad + \frac{4\,C_3\,\chi\,\beta}{\sqrt{N}}\,\sqrt{f(t)} \int_0^t e^{-\alpha(t-s)}\, ds\, dt, \end{aligned} \tag{4.14}$$

where $C_1 = d\|\operatorname{Hess} h_0\|_\infty$, $C_2 = d\|\operatorname{Hess} g\|_\infty$ and $C_3 = \|\nabla g\|_\infty$. Thus

$$df(t) \;\le\; \left(-2\,\tilde\lambda\, f(t) + 2\,\tilde C_2 \sqrt{f(t)} \int_0^t e^{-\alpha(t-s)} \sqrt{f(s)}\, ds + \frac{2\,\tilde C_3}{\sqrt{N}}\,\sqrt{f(t)}\right) dt, \tag{4.15}$$

where $-\tilde\lambda \doteq -v_* + \chi C_1 + C_2\chi\beta/\alpha$, $\tilde C_2 \doteq C_2\,\chi\,\beta$ and $\tilde C_3 \doteq 2\,C_3\,\chi\,\beta/\alpha$. Under Assumption 2.4 $\tilde\lambda \ge \tilde C_2/\alpha$. Note that $f(0) = 0$ since $X_0^{i,N} = \bar X_0^i$. Multiplying both sides of (4.15) by $e^{2\,\tilde\lambda\, t}$ and letting $\vartheta(t) \doteq e^{2\,\tilde\lambda\, t}\, f(t)$, we obtain

$$d\vartheta(t) \le 2\sqrt{\vartheta(t)}\Big(\int_0^t \tilde C_2\, e^{(\tilde\lambda-\alpha)\,(t-s)}\, \sqrt{\vartheta(s)}\, ds + \frac{\tilde C_3 e^{\tilde\lambda\, t}}{\sqrt{N}}\Big). \tag{4.16}$$

In rest of the proof we estimate $\vartheta(t)$ using the above inequality. For this, heuristically, one can set $\zeta = \sqrt{\vartheta}$ to obtain a simplification

$$\zeta'(t) \le \tilde C_2 \int_0^t e^{(\tilde\lambda-\alpha)\,(t-s)} \zeta(s)\, ds + \frac{\tilde C_3 e^{\tilde\lambda\, t}}{\sqrt{N}} \quad \text{on } \{\zeta \ne 0\}.$$

Since we do not have any control for $\zeta'(t)$ on $\{\zeta = 0\}$, we instead consider $\zeta_\epsilon(t) = \sqrt{\vartheta(t)+\epsilon^2}$ where $\epsilon > 0$. Then $\vartheta' = 2\zeta_\epsilon\zeta_\epsilon'$ and $\sqrt{\vartheta} = \sqrt{\zeta_\epsilon^2 - \epsilon^2} \le \zeta_\epsilon$. Hence (4.16) implies that

$$d(\zeta_\epsilon(t))^2 \le 2\zeta_\epsilon(t)\Big(\int_0^t \tilde C_2\, e^{(\tilde\lambda-\alpha)\,(t-s)} \zeta_\epsilon(s)\, ds + \frac{\tilde C_3 e^{\tilde\lambda\, t}}{\sqrt{N}}\Big).$$

Thus for a.e. $t \geq 0$,

$$\zeta_\epsilon'(t) \leq \int_0^t \tilde{C}_2\, e^{(\tilde{\lambda}-\alpha)\,(t-s)} \zeta_\epsilon(s)\, ds + \frac{\tilde{C}_3 e^{\tilde{\lambda}\, t}}{\sqrt{N}}. \tag{4.17}$$

We will now use a comparison result for ordinary differential equations (ODE). Let k_ϵ be the solution of the ODE

$$k_\epsilon''(t) = (\tilde{\lambda}-\alpha)\, k_\epsilon'(t) + \tilde{C}_2\, k_\epsilon(t) + \frac{\alpha\, \tilde{C}_3}{\sqrt{N}}\, e^{\tilde{\lambda}\, t},\ k_\epsilon(0) = \epsilon,\ k_\epsilon'(0) = \frac{\tilde{C}_3}{\sqrt{N}}. \tag{4.18}$$

Note that the solution k_ϵ solves the integral equation

$$k_\epsilon'(t) = \int_0^t \tilde{C}_2\, e^{(\tilde{\lambda}-\alpha)\,(t-s)} k_\epsilon(s)\, ds + \frac{\tilde{C}_3 e^{\tilde{\lambda}\, t}}{\sqrt{N}}, \quad k_\epsilon(0) = \epsilon. \tag{4.19}$$

It is straightforward to verify that the unique solution of (4.18) converges to

$$k(t) \doteq \frac{\tilde{C}_3}{(r_1-r_2)\sqrt{N}} \Big[e^{r_1 t} - e^{r_2 t} + \alpha \Big(\frac{e^{\tilde{\lambda} t} - e^{r_1 t}}{\tilde{\lambda} - r_1} - \frac{e^{\tilde{\lambda} t} - e^{r_2 t}}{\tilde{\lambda} - r_2} \Big) \Big] \tag{4.20}$$

uniformly on compacts when $\epsilon \to 0$, where $r_2 < 0 < r_1$ are the zeros of the characteristic polynomial

$$\theta(r) = r^2 - (\tilde{\lambda}-\alpha)\, r - \tilde{C}_2. \tag{4.21}$$

In particular,

$$\sup_{t \leq T} \sup_{\epsilon \in (0,1)} |k_\epsilon(t)| < \infty \quad \text{for any } T \geq 0. \tag{4.22}$$

Subtracting (4.19) from (4.17), we see that $\phi_\epsilon = \zeta_\epsilon - k_\epsilon$ satisfies

$$\phi_\epsilon(t) \leq \tilde{C}_2 \int_0^t \int_0^{t_1} e^{(\tilde{\lambda}-\alpha)\,(t_1-t_2)} \phi_\epsilon(t_2)\, dt_2\, dt_1, \quad t \geq 0.$$

Upon iterating M times, one has

$$\phi_\epsilon(t) \leq (\tilde{C}_2)^M \int_{\{0 \leq t_{2M} \leq t_{2M-1} \leq \cdots \leq t_1 \leq t\}} \prod_{i=1}^{M} e^{(\tilde{\lambda}-\alpha)\,(t_{2i-1}-t_{2i})} \phi_\epsilon(t_{2M})\, dt_{2M} \cdots dt_2\, dt_1. \tag{4.23}$$

For any $T > 0$, $C_T = \sup_{t \leq T} \sup_{\epsilon \in (0,1)} |\phi_\epsilon(t)| < \infty$ by (4.22) and Theorem 3.1. Hence the integrand of (4.23) is at most $C_T\, e^{(\tilde{\lambda}-\alpha)\, T\, M}$. Thus (4.23) implies that for every $T \in (0,\infty)$ there exists $\tilde{C}_T \in (0,\infty)$ such that for all $t \in [0,T]$,

$$\begin{aligned} \phi_\epsilon(t) &\leq \tilde{C}_T^M \,\text{Volume}\, \{0 \leq t_{2M} \leq t_{2M-1} \leq \cdots \leq t_1 \leq t\} \\ &\leq (\tilde{C}_T T^2)^M/(2M)! \end{aligned}$$

Letting $M \to \infty$, we see that $\sup_{t \in [0,T]} \phi_\epsilon(t) \leq 0$. Since $T > 0$ is arbitrary, we have $\phi_\epsilon(t) \leq 0$ for all $t \geq 0$. Thus $\zeta_\epsilon(t) \leq k_\epsilon(t)$ for all $t \geq 0$. Letting $\epsilon \to 0$, we obtain that $\sqrt{\vartheta}$ is bounded by k defined by (4.20). Recalling that $\vartheta(t) \doteq e^{2\,\tilde{\lambda}\, t}\, f(t)$, we have

$$\sqrt{f(t)} \leq \frac{\tilde{C}_3}{(r_1-r_2)\sqrt{N}} \Big[e^{(r_1-\tilde{\lambda})t} - e^{(r_2-\tilde{\lambda})t} + \alpha \Big(\frac{1-e^{(r_1-\tilde{\lambda})t}}{\tilde{\lambda}-r_1} - \frac{1-e^{(r_2-\tilde{\lambda})t}}{\tilde{\lambda}-r_2} \Big) \Big]. \tag{4.24}$$

Observe that Assumption 2.4 implies that $\tilde{C}_2 < \tilde{\lambda}\alpha$ from which it follows that

$$\sqrt{(\tilde{\lambda}-\alpha)^2 + 4\tilde{C}_2} < \tilde{\lambda} + \alpha.$$

Recalling that r_1, r_2 are the positive and negative roots of (4.21) we now see that under Assumption 2.4, $\tilde{\lambda} > r_1 > 0 > r_2$. Thus the right hand side of (4.24) is at most

$$\frac{\tilde{C}_3}{(r_1 - r_2)\sqrt{N}} \Big[1 + \frac{\alpha}{\tilde{\lambda} - r_1}\Big]$$

The proof is complete. □

We can now complete the proof of Corollary 3.5.

Proof of Corollary 3.5. The first statement in the corollary is immediate from Theorem 3.4 on noting that

$$\mathcal{W}_2\Big(\mathcal{L}(X_t^{1,N}, X_t^{2,N}, \cdots, X_t^{k,N}),\ \mathcal{L}(\bar{X}_t^1)^{\otimes k}\Big) \le \left(\mathbb{E}\sum_{i=1}^{k} |X_t^{i,N} - \bar{X}_t^i|^2\right)^{1/2}.$$

For the second statement note that since $\mu_0 \in \mathcal{P}_q(\mathbb{R}^d)$, $\sup_{t\ge 0}\int_{\mathbb{R}^d} |x|^{\tilde{q}}\mu_t(dx) < \infty$ for some $\tilde{q} > 2$(see Remark 4.4). Hence from Theorem 1.1 of [12], we have

$$\limsup_{N\to\infty} \sup_{t\ge 0} \mathbb{E}[\mathcal{W}_2^2(\nu_t^N, \mu_t)] = 0.$$

Also, from Theorem 3.4, as $N \to \infty$

$$\sup_{t\ge 0} \mathbb{E}[\mathcal{W}_2^2(\mu_t^N, \nu_t^N)] \le \sup_{t\ge 0} \mathbb{E}\frac{1}{N}\sum_{i=1}^{N} |X_t^{i,N} - \bar{X}_t^i|^2 \to 0.$$

The result now follows on combining the above two displays and using the triangle inequality

$$\mathcal{W}_2(\mu_t^N, \mu_t) \le \mathcal{W}_2(\mu_t^N, \nu_t^N) + \mathcal{W}_2(\nu_t^N, \mu_t).$$ □

4.3 Concentration bounds

In this section we will first provide exponential concentration bounds that are uniform over compact time intervals. Under the stronger property in Assumption 2.4 we will then show that these bounds can be strengthened to be uniform over the infinite time horizon. We begin with an upper bound for $\mathcal{W}_1(\mu_t^N, \mu_t)$ in terms of $\big(\mathcal{W}_1(\nu_s^N, \mu_s)\big)_{s\in[0,t]}$ where ν^N is as introduced in (3.3).

4.3.1 Bounds in terms of empirical measures of independent variables

The following proposition is a generalization of Proposition 5.1 in [4]. Let $\tilde{\lambda} \doteq v_* - C_1\chi - C_2\chi\beta/\alpha$ be as above (4.16).

Proposition 4.1. *For all $t \ge 0$*

$$\mathcal{W}_1(\mu_t^N, \mu_t) \le \mathcal{W}_1(\nu_t^N, \mu_t) + \frac{\sqrt{C_2\,\chi\,\beta}}{2}\int_0^t \Big(e^{(r_1-\tilde{\lambda})(t-s)} - e^{(r_2-\tilde{\lambda})(t-s)}\Big)\mathcal{W}_1(\nu_s^N, \mu_s)\,ds,$$

where, as before, $r_2 < 0 < r_1$ are the solutions of $r^2 - (\tilde{\lambda} - \alpha)\,r - C_2\,\chi\,\beta = 0$.

Proof. From (4.10) and (4.11) we have

$$\begin{aligned} d|\Delta_t^i|^2 \le\ & 2\Big[-v_* + e^{-\alpha t}\,\chi\,d\|\operatorname{Hess} h_0\|_\infty\Big]\,|\Delta_t^i|^2\,dt \\ & + 2\,\chi\,\beta\int_0^t \left(\frac{e^{-\alpha(t-s)}\,d\|\operatorname{Hess} g\|_\infty}{N}\sum_j |\Delta_t^i|(|\Delta_t^i| + |\Delta_s^j|)\right) ds\ dt \\ & + 2\,\chi\,\beta\int_0^t \left(e^{-\alpha(t-s)}\Delta_t^i \cdot \Big(\int \nabla P_{t-s}g(y - \bar{X}_t^i)\,(\nu_s^N - \mu_s)(dy)\Big)\right) ds\ dt. \end{aligned} \tag{4.25}$$

Instead of taking expectations as in Section 4.2, we now bound the third term on the right hand side of (4.25) using the inequality

$$\Big| \int \nabla P_{t-s} g(\bar{X}^i_t - y)\, (\nu^N_s - \mu_s)(dy) \Big| \le d \| \operatorname{Hess} g \|_\infty \, \mathcal{W}_1(\nu^N_s,\, \mu_s) \tag{4.26}$$

which follows from the Kantorovich-Rubenstein duality (1.12) and the fact that the Lipschitz norm of the function $x \mapsto \nabla_j P_{t-s} g(\bar{X}^i_t - x)$ is bounded by $d \| \operatorname{Hess} g \|_\infty$ for each $j = 1, \cdots, d$.

Applying (4.26) to (4.25), then summing over i and using the inequality $\sum_i |\Delta^i_t| \le \sqrt{N \sum_i |\Delta^i_t|^2}$, we see that with $F(t) \doteq \sum_{i=1}^N |\Delta^i_t|^2/N$

$$\begin{aligned}
dF(t) \le & 2 \Big[- v_* \; + \; \chi\, d \| \operatorname{Hess} h_0 \|_\infty \Big] \, F(t)\, dt \\
& + 2\, \chi\, \beta \int_0^t \Big(e^{-\alpha(t-s)}\, d \| \operatorname{Hess} g \|_\infty \, (F(t) + \sqrt{F(t)\, F(s)}) \Big)\, ds\, dt \\
& + 2\, \chi\, \beta \int_0^t \Big(e^{-\alpha(t-s)} \sqrt{F(t)}\, d \| \operatorname{Hess} g \|_\infty \, \mathcal{W}_1(\nu^N_s,\, \mu_s) \Big)\, ds\, dt \\
\le & 2 \Big[- v_* \; + \; C_1\, \chi \Big] \, F(t)\, dt \\
& + 2\, \chi\, \beta\, C_2 \int_0^t e^{-\alpha(t-s)} \, (F(t) + \sqrt{F(t)\, F(s)})\, ds\, dt \\
& + 2\, \chi\, \beta\, C_2 \sqrt{F(t)} \int_0^t e^{-\alpha(t-s)} \, \mathcal{W}_1(\nu^N_s,\, \mu_s)\, ds\, dt \\
= & 2 \Big[- v_* \; + \; C_1\, \chi \; + \; \frac{\chi\, \beta\, C_2}{\alpha} \Big] \, F(t)\, dt \\
& + 2\, \chi\, \beta\, C_2 \sqrt{F(t)} \int_0^t e^{-\alpha(t-s)} \, \Big(\sqrt{F_s} + \mathcal{W}_1(\nu^N_s,\, \mu_s) \Big)\, ds\, dt
\end{aligned}$$

Recalling that $\tilde{\lambda} = v_* - C_1 \chi - C_2 \chi \beta / \alpha$, we obtain

$$dF(t) \; \le \; \Big(-2\, \tilde{\lambda}\, F(t) + 2\, \tilde{C}_2 \sqrt{F(t)} \int_0^t e^{-\alpha(t-s)} \, \Big(\sqrt{F(s)} + \mathcal{W}(s) \Big) \, ds \Big)\, dt, \tag{4.27}$$

where $\mathcal{W}(s) \doteq \mathcal{W}_1(\nu^N_s,\, \mu_s)$ and $\tilde{C}_2$ is as introduced below (4.15). Recall that $F_0 = 0$. We now use (4.27) to obtain an upper bound for $F(t)$ in terms of $\{\mathcal{W}(s)\}_{s \in [0,t]}$.

As in (4.16), $G(t) \doteq e^{2\, \tilde{\lambda}\, t}\, F(t)$ satisfies

$$G'(t) \le 2\, \tilde{C}_2 \sqrt{G(t)} \Big(\int_0^t e^{(\tilde{\lambda} - \alpha)\, (t-s)} \sqrt{G(s)}\, ds \; + \; e^{(\tilde{\lambda} - \alpha)\, t} \int_0^t e^{\alpha s} \mathcal{W}(s)\, ds \Big).$$

Following the same comparison argument as was used to obtain the bound for $\sqrt{\vartheta}$ (ϑ was introduced above (4.16)), we let $H_\epsilon(t) = \sqrt{G(t) + \epsilon^2}$ where $\epsilon > 0$ and obtain for a.e. $t \ge 0$

$$H'_\epsilon(t) \le \tilde{C}_2 \Big(\int_0^t e^{(\tilde{\lambda} - \alpha)\, (t-s)} H_\epsilon(s)\, ds \; + \; e^{(\tilde{\lambda} - \alpha)\, t} \int_0^t e^{\alpha s} \mathcal{W}(s)\, ds \Big). \tag{4.28}$$

This time we need to solve the inhomogeneous second order ODE

$$K''_\epsilon(t) - (\tilde{\lambda} - \alpha)\, K'_\epsilon(t) - \tilde{C}_2\, K_\epsilon(t) = \tilde{C}_2\, e^{\tilde{\lambda}\, t} \mathcal{W}(t)$$

with initial conditions $K_\epsilon(0) = \epsilon$ and $K'_\epsilon(0) = 0$. On solving this ODE, we obtain, as in (4.20) and (4.24), that K_ϵ converges uniformly on compacts as $\epsilon \to 0$ to K defined as

$$K(t) \doteq \frac{\tilde{C}_2}{(r_1 - r_2)} \Big(e^{r_1\, t} \int_0^t e^{(\tilde{\lambda} - r_1) s} \mathcal{W}(s)\, ds - e^{r_2\, t} \int_0^t e^{(\tilde{\lambda} - r_2) s} \mathcal{W}(s)\, ds \Big) \tag{4.29}$$

Similar to the argument below (4.22) we have that for all $\epsilon > 0$ and $t \geq 0$

$$\sqrt{G(t) + \epsilon^2} \leq H_\epsilon(t) \leq K_\epsilon(t).$$

Sending $\epsilon \to 0$,

$$\sqrt{F(t)} \leq e^{-\tilde{\lambda}t}\sqrt{G(t)} \leq e^{-\tilde{\lambda}t}K(t) = \frac{\tilde{C}_2}{(r_1 - r_2)} \int_0^t \left(e^{(r_1 - \tilde{\lambda})(t-s)}\mathcal{W}(s) - e^{(r_2 - \tilde{\lambda})(t-s)}\mathcal{W}(s) \right) ds. \tag{4.30}$$

On the other hand,

$$\mathcal{W}_1(\mu_t^N, \nu_t^N) \leq \mathcal{W}_2(\mu_t^N, \nu_t^N) \leq \Big(\sum_{i=1}^N |\Delta_t^i|^2/N \Big)^{1/2} = \sqrt{F(t)}. \tag{4.31}$$

The desired equality now follows from the triangle inequality $\mathcal{W}_1(\mu_t^N, \mu_t) \leq \mathcal{W}_1(\nu_t^N, \mu_t) + \mathcal{W}_1(\mu_t^N, \nu_t^N)$ and the observation that $r_1 - r_2 \geq 2\sqrt{\tilde{C}_2}$. □

The following corollary is an immediate consequence of the above proposition.

Corollary 4.2. *For every $T \in (0, \infty)$, there exists $C_T \in (0, \infty)$ such that*

$$\mathcal{W}_1(\mu_t^N, \mu_t) \leq \mathcal{W}_1(\nu_t^N, \mu_t) + C_T \int_0^t \mathcal{W}_1(\nu_s^N, \mu_s)\, ds \quad \text{for } t \in [0, T].$$

Also, for all $t \geq 0$

$$\mathcal{W}_1(\mu_t^N, \mu_t) \leq \mathcal{W}_1(\nu_t^N, \mu_t) + \frac{\sqrt{C_2\, \chi\, \beta}}{2} \int_0^t e^{(r_1 - \tilde{\lambda})(t-s)} \mathcal{W}_1(\nu_s^N, \mu_s)\, ds.$$

Recall from Section 4.2 that under Assumption 2.4 $(r_1 - \tilde{\lambda}) < 0$. This will be key in obtaining a uniform in time bound from the last inequality in the corollary above.

4.3.2 Moment bounds

Let $(\bar{X}_t)_{t \geq 0}$ be the nonlinear process solving (1.2) and μ_t be its law at t. In Proposition 4.3 below we will give bounds on the square exponential moments of $\bar{X}_t$ under appropriate conditions. The first part of the proposition holds under our standing assumptions along with a suitable integrability condition. For the second part we will make in addition an assumption that is weaker than Assumption 2.4, namely $v_* > 0$, where v_* was defined in (2.7). Let, for $t, \theta \geq 0$,

$$S_\theta(t) \doteq \int_{\mathbb{R}^d} e^{\theta\,|x|^2}\, d\mu_t(x) = \mathbb{E}\big[e^{\theta\,|\bar{X}_t|^2}\big].$$

Proposition 4.3. *Suppose μ_0 is as in Theorem 3.8, namely for some $\theta_0 > 0$, $S_{\theta_0}(0) < \infty$. Then*

(i) For any $T \in (0, \infty)$, there exists $\theta_T > 0$ such that $\sup_{t \in [0,T]} S_{\theta_T}(t) < \infty$.

(ii) Suppose that $v_* > 0$. Then for any $\theta \in (0,\, \theta_0/4 \wedge v_*/8)$, we have $\sup_{t \geq 0} S_\theta(t) < \infty$.

Proof. (i) Note that for all $t \in [0, T]$

$$|\bar{X}_t| \leq |\bar{X}_0| + \chi C T + \sup_{0 \leq t \leq T} |B_t| + L_{\nabla V} \int_0^t |\bar{X}_s|\, ds, \quad t \geq 0,$$

where C is as in Lemma 2.1. By Gronwall's lemma

$$\sup_{0\le t\le T}|\bar{X}_t| \le a(T)e^{L_{\nabla V}T}$$

where

$$a(T) = |\bar{X}_0| + \chi CT + \sup_{0\le t\le T}|B_t|.$$

Thus for $\theta > 0$

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\le t\le T} e^{\theta|\bar{X}_t|^2}\Big] &\le \kappa_{1,T}\mathbb{E}\Big[e^{2\theta|\bar{X}_0|^2}e^{2\theta(\chi CT+\sup_{0\le t\le T}|B_t|)^2}\Big]\\ &\le \kappa_{1,T}\left(\mathbb{E}\left[e^{4\theta|\bar{X}_0|^2}\right]\right)^{1/2}\left(\mathbb{E}\left[e^{4\theta(\chi CT+\sup_{0\le t\le T}|B_t|)^2}\right]\right)^{1/2}.\end{aligned}$$

Choose $\theta \le \frac{\theta_0}{4}$ such that $\mathbb{E}\left[e^{8\theta|B_T|^2)}\right] < \infty$. Then for all such θ, $\sup_{0\le t\le T} S_\theta(t) < \infty$.

(ii) Fix $\theta > 0$. We apply Itô's formula to $\phi(x) = e^{\theta|x|^2}$. Note that $\nabla\phi(x) = 2\theta e^{\theta|x|^2}x$ and $\Delta\phi(x) = 2\theta e^{\theta|x|^2}\big(d + 2\theta|x|^2\big)$. Thus for $t \ge 0$

$$d\phi(\bar{X}_t) = e^{\theta|\bar{X}_t|^2}\Big(2\theta\bar{X}_t\cdot(dB_t - \nabla V(\bar{X}_t)\,dt + \chi\nabla h(t,\bar{X}_t)\,dt) + \theta\big(d + 2\theta|\bar{X}_t|^2\big)\,dt\Big). \quad (4.32)$$

We claim that with $\sigma \doteq \theta_0/4 \wedge v_*/8$, for all $\theta \in (0,\sigma)$,

$$M_\theta(t) \doteq \int_0^t e^{\theta|\bar{X}_s|^2}\bar{X}_s\cdot dB_s$$

is a square integrable martingale. Suppose for now that the claim is true. Then $\mathbb{E}[M_\theta(t)] = 0$ for all $t \ge 0$ and $\theta \in (0,\sigma)$. By our assumption $-x\cdot\nabla V(x) \le -v_*|x|^2$ for all $x \in \mathbb{R}^d$. Also, for any $\eta > 0$, $|x| \le \eta + \frac{|x|^2}{4\eta}$. Combining these observations with Lemma 2.1, we obtain that

$$dS_\theta(t) \le \mathbb{E}\Big[e^{\theta|\bar{X}_t|^2}\big(A + B\,|\bar{X}_t|^2\big)\Big]\,dt, \quad (4.33)$$

for all $\eta > 0$, where

$$\begin{aligned}A &= \theta\big(2\chi\,\|\nabla h\|_\infty\eta + d\big),\\ B &= 2\,\theta^2 - 2b\,\theta \quad \text{and } b = v_* - \frac{\chi}{4\eta}\|\nabla h\|_\infty.\end{aligned}$$

Since $v_* > 0$, we can find $\eta \in (0,\infty)$ so that $b = v_*/2$ and consequently, since $\sigma < v_*/2$, with this choice of η, $B < 0$ for all $\theta \in (0,\sigma)$. Therefore for such θ

$$dS_\theta(t) \le \kappa\int_{\mathbb{R}^d}(R^2 - |x|^2)\,e^{\theta|x|^2}\,d\mu_t(x), \quad (4.34)$$

for some R, $\kappa \in (0,\infty)$ depending only on A and B. Decomposing the integrand on the right of (4.34) according to the size of $|x|$, one obtains

$$dS_\theta(t) \le \big(\kappa_1 - \kappa_2\,S_\theta(t)\big)\,dt$$

where κ_1, $\kappa_2 \in (0,\infty)$. Since $\theta < \sigma \le \theta_0$, $S_\theta(0) < \infty$. A standard estimate now shows that $\sup_{t\ge0} S_\theta(t) < \infty$ for all $\theta \in (0,\sigma)$.

Finally we verify the claim. Using the estimates $-x\cdot\nabla V(x) \le -v_*|x|^2$ and $|x| \le \eta + \frac{|x|^2}{4\eta}$ once again, and choosing η such that $b = v_*/2$ as before, we have by an application of Itô's formula

$$|\bar{X}_t|^2 \le |\bar{X}_0|^2 + \int_0^t 2\bar{X}_s\cdot dB_s - \int_0^t v_*|\bar{X}_s|^2\,ds + \kappa_3\,t \quad (4.35)$$

where $\kappa_3 = 2\chi\|\nabla h\|_\infty \eta + d$. Provided that $\sigma_1 \leq v_*/4 \wedge \theta_0/2$, we can bound $S_{\sigma_1}(t)$ above by

$$\begin{aligned}
&e^{\sigma_1\,\kappa_3\,t}\,\mathbb{E}\Big[\exp\Big(\sigma_1|\bar{X}_0|^2\Big)\cdot\exp\Big(\int_0^t 2\sigma_1\bar{X}_s\cdot dB_s - \int_0^t v_*\sigma_1|\bar{X}_s|^2\,ds\Big)\Big]\\
&\quad\leq e^{\sigma_1\,\kappa_3\,t}\,\mathbb{E}\Big[\exp\Big(\sigma_1|\bar{X}_0|^2\Big)\cdot\exp\Big(\int_0^t 2\sigma_1\bar{X}_s\cdot dB_s - \int_0^t 4\sigma_1^2|\bar{X}_s|^2\,ds\Big)\Big]\\
&\quad\leq e^{\sigma_1\,\kappa_3\,t}\,\sqrt{\mathbb{E}\Big[\exp\Big(2\sigma_1|\bar{X}_0|^2\Big)\Big]\cdot\mathbb{E}\Big[\exp\Big(\int_0^t 4\sigma_1\bar{X}_s\cdot dB_s - \int_0^t 8\sigma_1^2|\bar{X}_s|^2\,ds\Big)\Big]}\\
&\quad\leq e^{\sigma_1\,\kappa_3\,t}\,\sqrt{\mathbb{E}\Big[\exp\Big(2\sigma_1|\bar{X}_0|^2\Big)\Big]}\\
&\quad\leq \kappa_4 e^{\sigma_1\,\kappa_3\,t},
\end{aligned}\tag{4.36}$$

where the next to last inequality follows on noting that

$$\exp\Big(\int_0^t 4\sigma_1\bar{X}_s\cdot dB_s - 8\sigma_1^2\int_0^t|\bar{X}_s|^2\,ds\Big)$$

is a supermartingale and in the last inequality we have used the fact that since $\sigma_1 \leq \theta_0/2$, $\mathbb{E}\exp(2\sigma_1|\bar{X}_0|^2) < \infty$. Finally for $\theta < \sigma_1/2$ and $t < \infty$

$$\begin{aligned}
\int_0^t \mathbb{E}\left[e^{2\theta|X_s|^2}|X_s|^2\right]ds &\leq \frac{1}{\sigma_1 - 2\theta}\int_0^t \mathbb{E}e^{\sigma_1|X_s|^2}\\
&\leq \frac{\kappa_4}{\sigma_1 - 2\theta}\int_0^t e^{\sigma_1\,\kappa_3\,s}\,ds\\
&< \infty.
\end{aligned}$$

This completes the proof of the claim and the result follows. □

Remark 4.4. (i) In a similar manner it can be shown that if μ_0 admits a finite square exponential moment of order $\theta_0 > 0$, then for every $T > 0$ there is a $\theta_T > 0$ such that

$$\sup_{N,i}\sup_{0\leq t\leq T}\mathbb{E}\big[e^{\theta_T\,|X_t^{N,i}|^2}\big] < \infty.$$

Furthermore if $v_* > 0$, we have for all θ as in part (ii) of the above proposition

$$\sup_{N,i}\sup_{t\geq 0}\mathbb{E}\big[e^{\theta\,|X_t^{N,i}|^2}\big] < \infty.$$

(ii) Suppose that instead of assuming that μ_0 has a finite squared exponential moment, we assume that $\mu_0 \in \mathcal{P}_{q_0}(\mathbb{R}^d)$ for some $q_0 \geq 2$. Then it follows easily that for any fixed $T < \infty$

$$\sup_{0\leq t\leq T}\mathbb{E}|\bar{X}_t|^{q_0} < \infty.$$

Furthermore, by applying Itô formula to $|x|^q$ instead of $e^{\theta|x|^2}$ one can check that, if in addition $v_* > 0$, with $q = \frac{q_0}{2} + 1$,

$$\sup_{0\leq t<\infty}\mathbb{E}|\bar{X}_t|^q < \infty.$$

Also, analogous statements as in (i) hold with the squared exponential moment replaced by the q-th moment.

4.3.3 Time-regularity

In this section we give some estimates on the moments of the increments of the nonlinear process $\bar{X}$. These estimates are needed to appeal to results in [4] for proofs of our concentration bounds.

We start with moment estimates for $|\bar{X}_t - \bar{X}_s|$. By Lemma 2.1 and our assumption on ∇V we have

$$|\bar{X}_t - \bar{X}_s| \leq |B_t - B_t| + L_{\nabla V} \int_s^t |\bar{X}_r| \, dr + C\,(t-s), \quad 0 \leq s \leq t < \infty,$$

where C is as in Lemma 2.1.

Throughout this section we will assume that $S_{\theta_0}(0) < \infty$ for some $\theta_0 > 0$. Taking powers and using Proposition 4.3, we obtain the following result.

Lemma 4.5. *For all $p \geq 1$ and $T > 0$, there exists $C_{T,p} \in (0, \infty)$ such that*

$$\mathbb{E}[|\bar{X}_t - \bar{X}_s|^p] \leq C_{T,p}\,|t-s|^{p/2} \quad \textit{for } s,\, t \in [0,T]. \tag{4.37}$$

Moreover, if Assumption 2.4 is satisfied, then for some $C_p \in (0, \infty)$

$$\mathbb{E}[|\bar{X}_t - \bar{X}_s|^p] \leq C_p(|t-s|^p + |t-s|^{p/2}) \quad \textit{for } s,\, t \geq 0.$$

Lemma 4.5 immediately implies

$$\mathcal{W}_p(\mu_t,\, \mu_s) \leq \tilde{C}_{T,p}\,|t-s|^{1/2} \quad \text{for } s,\, t \in [0,T] \tag{4.38}$$

and under Assumption 2.4, for some $\tilde{C}_p > 0$

$$\mathcal{W}_p(\mu_t,\, \mu_s) \leq \tilde{C}_p(|t-s| + |t-s|^{1/2}) \quad \text{for } s,\, t \geq 0.$$

Next we give an estimate on the exponential moments of the increments.

Lemma 4.6. *For all $T > 0$, there exist $\theta_T, C_T \in (0, \infty)$ such that*

$$\mathbb{E}\Big[\sup_{a \leq s, t \leq a+b} \exp(\theta_T |\bar{X}_t - \bar{X}_s|^2)\Big] \leq 1 + C_T\, b$$

for all $a,\, b \in [0,T]$.

Proof. By an application of Cauchy-Schwarz inequality we see that it suffices to show that for some $\theta_T, C_T \in (0, \infty)$

$$\mathbb{E}\Big[\sup_{a \leq t \leq a+b} \exp(\theta_T |\bar{X}_t - \bar{X}_a|^2)\Big] \leq 1 + C_T\, b \tag{4.39}$$

for all $a, b \in [0,T]$. Let C be as in Lemma 2.1 and $\mathcal{H}_C$ be the class of all functions $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\sup_{t \geq 0} |v(t,x) - v(t,y)| \leq C|x-y|,\ \sup_{t \geq 0} |v(t,x)| \leq C \text{ for all } x, y \in \mathbb{R}^d.$$

Note that from Lemma 2.1, for every $m \in \mathcal{P}\big(\mathcal{C}([0,\infty) : \mathbb{R}^d)\big)$, $\nabla h^m \in \mathcal{H}_C$.

Given $v \in \mathcal{H}_C$ and $z \in \mathbb{R}^d$, let $Y^{v,z}$ be the solution of the stochastic differential equation

$$dY_t^{v,z} = dB_t - \nabla V(Y_t^{v,z}) + \chi v(t, Y_t^{v,z}) dt,\ Y_0^{v,z} = z.$$

By a standard conditioning argument it suffices to argue that for some $\theta_T, C_T \in (0, \infty)$

$$\sup_{z \in \mathbb{R}^d, v \in \mathcal{H}_C} \mathbb{E}\left[\sup_{0 \leq t \leq b} \exp(\theta_T |Y_t^{v,z} - z|^2)\right] \leq 1 + C_T\, b. \tag{4.40}$$

Fix $(z,v) \in \mathbb{R}^d \times \mathcal{H}_C$ and suppress it in the notation (i.e. write $Y^{v,z}$ as Y). Let $\theta : [0,T] \to \mathbb{R}$ be a non-negative continuously differentiable function and write for $t \in [0,b]$

$$Z_t \doteq e^{\theta(t)\,|Y_t - z|^2}.$$

Using Itô's formula we obtain,

$$\begin{aligned} dZ_t = Z_t \Big[& 2\theta(t)(Y_t - z) \cdot (dB_t - \nabla V(Y_t)\,dt + \chi v(t, Y_t)\,dt) \\ & + \theta(t)\big(d + 2\theta(t)|Y_t - z|^2\big)\,dt + \theta'(t)|Y_t - z|^2\,dt \Big]. \end{aligned} \tag{4.41}$$

Integrating, we obtain

$$\begin{aligned} Z_t - 1 = M_t + \int_0^t Z_r \Big[& 2\theta(r)(Y_r - z) \cdot \big(- \nabla V(Y_r) + \chi v(r, Y_r)\big) \\ & + \theta(r)\big(d + 2\theta(r)|Y_r - z|^2\big) + \theta'(r)|Y_r - z|^2 \Big]\,dr, \end{aligned} \tag{4.42}$$

where $M_t \doteq 2\int_0^t Z_r \theta(r)(Y_r - z) \cdot dB_r$. From a similar argument as for the proof of Proposition 4.3(i), there is a $\varsigma > 0$ such that

$$\sup_{(z,v)\in\mathbb{R}^d\times\mathcal{H}_C} \sup_{0\le s\le T} \mathbb{E} e^{\varsigma |Y_s^{z,v} - z|^2} < \infty. \tag{4.43}$$

One of the properties of $\{\theta(t)\}_{0\le t\le T}$ chosen below will be that $\sup_{0\le t\le T}\theta(t) < \varsigma/2$. With such a choice of $\{\theta(t)\}$, $\{M_t\}$ is a martingale and consequently $\mathbb{E}(M_t) = 0$ for all $t \ge 0$.

Applying Young's inequality we see that for every $\eta > 0$

$$2(Y_r - z) \cdot \big(- \nabla V(Y_r) + \chi v(r, Y_r)\big) \le \eta |Y_r - z|^2 + \frac{|-\nabla V(Y_r) + \chi v(r, Y_r)|^2}{\eta},$$

and thus

$$Z_t \le 1 + M_t + \int_0^t Z_r \big(A_r + B_r\,|Y_r - z|^2\big)\,dr, \tag{4.44}$$

where

$$\begin{aligned} A_r &= \theta(r)\left(d + \frac{L_{\nabla V}|Y_r| + \chi C}{\eta}\right) \\ B_r &= \eta\,\theta(r) + 2\theta^2(r) + \theta'(r). \end{aligned}$$

Rest of the argument is similar to [4, Section 5.1] and so we only give a sketch. Choose $\theta(r)$ to be the solution of the ODE

$$\eta\,\theta(r) + 2\theta^2(r) + \theta'(r) = 0$$

with $\theta(0)$ to be a strictly positive and smaller than $\varsigma/2$. It is easy to see that the solution is decreasing and strictly positive. Thus B_r is identically zero and $0 < \theta(T) \le \theta(t) \le \varsigma/2$ for every $t \in [0,T]$. As a consequence

$$\mathbb{E} \sup_{0\le t\le b} Z_t \le 1 + \mathbb{E} \sup_{0\le t\le b} M_t + \int_0^b \mathbb{E} Z_r A_r dr.$$

Using the bound in (4.43) it is now checked exactly as in [4] that

$$\sup_{(z,v)\in\mathbb{R}^d\times\mathcal{H}_C} \sup_{0\le s\le T} \mathbb{E} Z_r A_r < \infty \text{ and } \sup_{(z,v)\in\mathbb{R}^d\times\mathcal{H}_C} \mathbb{E} \sup_{0\le s\le b} M_s \le \tilde{C}_T b,$$

for some $\tilde{C}_T < \infty$. This proves (4.40) and thus the result follows. □

With Lemma 4.6, standard estimates for Brownian motion and the Chebyshev's exponential inequality then yields the following time-regularity of the empirical measures ν^N. The proof is contained in [4, Section 5.2, page 577–578] and is omitted here.

Proposition 4.7. *For any $T > 0$, there exist C_1, $C_2 \in (0, \infty)$ such that*

$$\mathbb{P}\Big(\sup_{a \le s, t \le a+b} \mathcal{W}_1(\nu_s^N, \nu_t^N) > \epsilon \Big) \le \exp\big(- N(C_1\, \epsilon^2 - C_2\, b)\big)$$

for all a, $b \in [0, T]$ and $\epsilon > 0$.

Remark 4.8. Using Lemma 2.1, it is immediate that Lemma 4.6 and Proposition 4.7 remain valid with the same constants if we replace $\bar{X}_t$ by $X_t^{N,i}$ and ν^N by μ^N respectively.

4.3.4 Proofs for the concentration bounds

Equipped with the results Sections 4.3.1, 4.3.2 and 4.3.3 in the previous subsections, the proofs of Theorem 3.6 and Theorem 3.8 can be completed as in [4, Section 7.1] and [4, Section 7.2] respectively. We only provide a sketch.

For Theorem 3.6, note that first bound in Corollary 4.2 implies

$$\sup_{t \in [0,T]} \mathcal{W}_1(\mu_t^N, \mu_t) \le \kappa_T \sup_{t \in [0,T]} \mathcal{W}_1(\nu_t^N, \mu_t).$$

This in turn implies

$$\mathbb{P}\Big(\sup_{t \in [0,T]} \mathcal{W}_1(\mu_t^N, \mu_t) > \epsilon \Big) \le \mathbb{P}\Big(\sup_{t \in [0,T]} \mathcal{W}_1(\nu_t^N, \mu_t) > \tilde{\epsilon} \Big),$$

where $\tilde{\epsilon} = \epsilon / \kappa_T$. which is analogous to equation (75) in [4]. Part (i) of Proposition 4.3 guarantees that we can apply Theorem 3.7 to assert that for any $d' \in (d, \infty)$ and $\theta' \in (0, \theta)$, there exists a positive integer N_0 such that

$$\sup_{t \in [0,T]} \mathbb{P}\Big(\mathcal{W}_1(\nu_t^N, \mu_t) > \tilde{\epsilon} \Big) \le \exp\big(- \kappa_{2,T}\, \theta'\, N\, \epsilon^2 \big)$$

for all $\epsilon > 0$ and $N \ge N_0 \max(\epsilon^{-(d'+2)}, 1)$.

To complete the proof of Theorem 3.6, it remains to improve this estimate by interchanging $\sup_{t \in [0,T]}$ and $\mathbb{P}$. This "exchange" can be achieved using the continuity estimates on ν_t^N and μ_t in Subsection 4.3.3. Details can be verified as in [4, Section 7.1]. Specifically, Corollary 4.2, Proposition 4.3 and Proposition 5.1 in [4] are replaced by, respectively, part (i) of Proposition 4.3, (4.38) with $p = 1$, and Proposition 4.7.

For Theorem 3.8, we start from the second bound in Corollary 4.2 and argue as in [4, Section 7.2]. The key ingredient is Proposition 4.1 in [4]. This result is replaced by part (ii) of Proposition 4.3 which gives uniform in time estimate for the square exponential moment for μ_t. We omit the details.

4.4 Uniform convergence of Euler scheme

In the proofs of Lemma 3.9 and Theorem 3.10, we need to solve difference inequalities which are harder to handle than similar differential inequalities that appeared in the proofs of Theorem 3.4 and Proposition 4.1.

Proof of Lemma 3.9. From integration by parts in (3.7) we see that

$$G_\theta(\vec{x}, y) = \frac{-\chi \beta e^{-\alpha\theta}}{N} \sum_{j=1}^{N} \int_{\mathbb{R}^d} p(\theta, y, z) \nabla g(x_j - z)\, dz$$

and thus

$$\sup_{N\geq 1}\sup_{\vec{x}\in\mathbb{R}^{dN}}\sup_{y\in\mathbb{R}^d}|G_\theta(\vec{x},y)|\leq \chi\beta\, e^{-\alpha\theta}\|\nabla g\|_\infty. \tag{4.45}$$

Also recalling that $V_t = V - \chi Q_t h_0$,

$$\sup_{t\geq 0}|\nabla V_t(x)|\leq |\nabla V(x)| + \chi\sup_{t\geq 0}|\nabla Q_t h_0(x)|\leq \kappa_1\,(|x|+1). \tag{4.46}$$

Recall from (3.8) that

$$Y^i_{n+1} = Y^i_n + \Delta_n B^i + \epsilon\Big(\int_0^{n\epsilon} G_{n\epsilon-s}(\tilde{Y}^{(N),\epsilon}_s,\, Y^i_n)\,ds - \nabla V_{n\epsilon}(Y^i_n)\Big) \tag{4.47}$$

for $1\leq i\leq N$, where we have suppressed N,ϵ in the notation. Also note that $|\nabla(Q_t h_0)(y)|\leq\|\nabla h_0\|_\infty$ for every $y\in\mathbb{R}^d$. Hence by (4.45) and (4.46),

$$\begin{aligned}|Y^i_{n+1}|^2 \;\leq\; & -2\epsilon Y^i_n\cdot\nabla V_{n\epsilon}(Y^i_n) + |Y^i_n|^2 + |\Delta_n B^i|^2\\ & + \kappa_2\,\epsilon^2(|Y^i_n|^2+1) + \xi^i_n\cdot\Delta_n B^i + \kappa_3\,\epsilon|Y^i_n|\end{aligned}$$

where $\kappa_4 = \kappa_3 + 2\chi\|\nabla h_0\|_\infty$ and ξ^i_n is measurable with respect to $\mathcal{F}^{B^i}_{n\epsilon}\doteq\sigma\{B^i_s : 0\leq s\leq n\epsilon\}$. Note that $\mathbb{E}[\xi_n\Delta_n B^i]=0$ for every n,i. Our assumption on V then gives

$$\mathbb{E}|Y^i_{n+1}|^2\leq -2\epsilon v_*\mathbb{E}|Y^i_n|^2 + \mathbb{E}|Y^i_n|^2 + \epsilon + \kappa_2\epsilon^2\mathbb{E}|Y^i_n|^2 + \kappa_2\epsilon^2 + \kappa_4\epsilon\mathbb{E}|Y^i_n|.$$

Let $a_n = \mathbb{E}|Y^i_n|^2$. Then for ϵ small enough, we have the nonlinear difference inequality

$$a_{n+1}\leq (1-\epsilon v_*)a_n + \kappa\,\epsilon\sqrt{a_n} + 2\epsilon \tag{4.48}$$

Note that by our assumption $v_*>0$. By Young's inequlity $\kappa\sqrt{a_n}\leq\eta\,a_n + \kappa^2/(4\eta)$ for all $\eta>0$. Taking $\eta = v_*/2$ we obtain

$$a_{n+1} - \delta\,a_n\leq A \tag{4.49}$$

where $\delta = 1-\frac{v_*\epsilon}{2}$ and $A=\epsilon\Big(\frac{\kappa^2}{2v_*}+2\Big)$. Note that $\delta\in(0,1)$ for $\epsilon>0$ small enough. Multiplying both sides of (4.49) by δ^{-n}, we obtain

$$b_{n+1} - b_n\leq A\,\delta^{-n} \tag{4.50}$$

where $b_n = a_n\,\delta^{-(n-1)}$. Summing over n then gives

$$b_{n+1} - b_1\leq A\sum_{i=1}^n\delta^{-i}.$$

Since $a_0=0$ (giving $b_1\leq A$), we obtain

$$a_{n+1}\leq A\Big(\frac{1-\delta^{n+1}}{1-\delta}\Big) = \frac{2}{v_*}\Big(1-\big(1-v_*\epsilon/2\big)^{n+1}\Big)\Big(\frac{\kappa^2}{2v_*}+2\Big).$$

Thus we have

$$\sup_{n\geq 0}a_n\leq\frac{2}{v_*}\Big(\frac{\kappa^2}{2v_*}+2\Big). \tag{4.51}$$

The proof is complete. □

Remark 4.9. By (4.45) and (4.46), we also have (suppressing N and ϵ in notation)

$$\begin{aligned}|Y^i_n - Y^i_{n-1}| &\le |\Delta_{n-1}B^i| + \epsilon\Big|\nabla V_{(n-1)\epsilon}(Y^i_{n-1}) - \int_0^{(n-1)\epsilon} G_{(n-1)\epsilon - s}(\tilde{Y}_s, Y^i_{n-1})\,ds\Big| \\ &\le |\Delta_{n-1}B^i| + \epsilon\Big(\kappa_1\,(|Y^i_{n-1}|+1) + \frac{\chi\beta\|\nabla g\|_\infty}{\alpha}\Big) \\ &\le |\Delta_{n-1}B^i| + \epsilon\,\kappa_2\,(|Y^i_{n-1}|+1).\end{aligned}$$

From this and the uniform bounds in Lemma 3.9 we have that if $v_* > 0$,

$$\sup_{i,n}\mathbb{E}|Y^i_n - Y^i_{n-1}|^2 \le \kappa_3\,\epsilon \tag{4.52}$$

Similarly, with $X^{i,N}_t$ as in (3.5), for $t \in [(n-1)\epsilon, n\epsilon]$,

$$|X^{i,N}_t - X^{i,N}_{n\epsilon}| \le |B^i_t - B^i_{n\epsilon}| + \kappa_2\int_t^{n\epsilon}(|X^{i,N}_r|+1)\,dr$$

and by the uniform bound for $\sup_{t\ge 0}\mathbb{E}|X^{i,N}_t|$ (see Remark 4.4),

$$\sup_{i,N}\mathbb{E}|X^{i,N}_t - X^{i,N}_{n\epsilon}|^2 \le \kappa_3\,|t - n\epsilon| \quad \text{for all } n\ge 1,\ t\in[(n-1)\epsilon, n\epsilon]. \tag{4.53}$$

Proof of Theorem 3.10. To simplify notations, we supress N and ϵ to write Y^i_n for $Y^{i,N,\epsilon}_n$ and X^i_t for $X^{i,N}_t$. Denote $Z^i_n \triangleq Y^i_n - X^i_{n\epsilon}$ to be the error of the scheme. From (3.8) and (3.5) we obtain

$$Z^i_n = Z^i_{n-1} + a^i_n + b^i_n \quad \text{for } n\ge 1,\ 1\le i\le N, \tag{4.54}$$

where

$$\begin{aligned}a^i_n &= \int_{(n-1)\epsilon}^{n\epsilon}\Big(\int_0^{(n-1)\epsilon} G_{(n-1)\epsilon-s}(\tilde{Y}_s, Y^i_{n-1})\,ds - \int_0^t G_{t-s}(X_s, X^i_t)\,ds\Big)\,dt \\ b^i_n &= \int_{(n-1)\epsilon}^{n\epsilon}\big[-\nabla V_{(n-1)\epsilon}(Y^i_{n-1}) + \nabla V_t(X^i_t)\big]\,dt.\end{aligned}$$

From (4.54) we have $Z^i_n = \sum_{k=1}^n (a^i_k + b^i_k)$. Hence

$$|Z^i_n|^2 - |Z^i_{n-1}|^2 = (a^i_n + b^i_n)\,\big(Z^i_n + Z^i_{n-1}\big). \tag{4.55}$$

Step (i) We first estimate $|b^i_n\,Z^i_n + b^i_n\,Z^i_{n-1}|$. For this, we shall use the estimates

$$|\nabla Q_t h_0(y) - \nabla Q_t h_0(x)| \le e^{-\alpha t} d\|\operatorname{Hess} h_0\|_\infty |x-y|. \tag{4.56}$$

$$\begin{aligned}|\nabla Q_t h_0(x) - \nabla Q_s h_0(x)| &\le \Big|e^{-\alpha t}\big(\nabla P_t h_0(x) - \nabla P_s h_0(x)\big) + (e^{-\alpha t} - e^{-\alpha s})\nabla P_s h_0(x)\Big| \\ &\le d\,\|\operatorname{Hess} h_0\|_\infty e^{-\alpha t}\sqrt{t-s} \;+\; \|\nabla h_0\|_\infty \alpha\, e^{-\alpha\,(t\wedge s)}|t-s|.\end{aligned} \tag{4.57}$$

By (4.56), (4.57) and our convexity assumption on V,

$$\begin{aligned}b^i_n\,Z^i_n &= \int_{(n-1)\epsilon}^{n\epsilon}\Big[-\nabla V_{(n-1)\epsilon}(Y^i_{n-1}) + \nabla V_{n\epsilon}(Y^i_n) \\ &\qquad + \nabla V_t(X^i_t) - \nabla V_{n\epsilon}(X^i_{n\epsilon}) - \nabla V_{n\epsilon}(Y^i_n) + \nabla V_{n\epsilon}(X^i_{n\epsilon})\Big]\,dt\cdot(Y^i_n - X^i_{n\epsilon}) \\ &\le -v_*\,\epsilon\,|Z^i_n|^2 + |Z^i_n|\int_{(n-1)\epsilon}^{n\epsilon}\Big(\chi d\,\|\operatorname{Hess} h_0\|_\infty\,|Z^i_n| \\ &\qquad + |-\nabla V(Y^i_{n-1}) + \nabla V(Y^i_n)| + \kappa_1\Big[e^{-\alpha n\epsilon}|Y^i_n - Y^i_{n-1}| + \sqrt{\epsilon}\Big]\end{aligned}$$

$$
\begin{aligned}
&\qquad + |\nabla V(X_t^i) - \nabla V(X_{n\epsilon}^i)| + \kappa_1 \left[e^{-\alpha n\epsilon} |X_t^i - X_{n\epsilon}^i| + \sqrt{\epsilon} \right] \Big)\, dt \\
&\le -v_*\, \epsilon\, |Z_n^i|^2 \, + |Z_n^i| \int_{(n-1)\epsilon}^{n\epsilon} \Big(\chi d \,\| \operatorname{Hess} h_0 \|_\infty \, |Z_n^i| \\
&\qquad + L_{\nabla V} |Y_n^i - Y_{n-1}^i| + \kappa_1 \left[e^{-\alpha n\epsilon} |Y_n^i - Y_{n-1}^i| + \sqrt{\epsilon} \right] \\
&\qquad + L_{\nabla V} |X_t^i - X_{n\epsilon}^i| + \kappa_1 \left[e^{-\alpha n\epsilon} |X_t^i - X_{n\epsilon}^i| + \sqrt{\epsilon} \right] \Big)\, dt \\
&\le \big(-v_* + \chi d \,\| \operatorname{Hess} h_0 \|_\infty \big)\, |Z_n^i|^2 \epsilon \, + 2\kappa_1\, \epsilon^{3/2} |Z_n^i| \\
&\quad + \epsilon \Big(L_{\nabla V} + \kappa_1\, e^{-\alpha n\epsilon} \Big) |Z_n^i|\, |Y_n^i - Y_{n-1}^i| \\
&\quad + \Big(L_{\nabla V} + \kappa_1\, e^{-\alpha n\epsilon} \Big) |Z_n^i| \int_{(n-1)\epsilon}^{n\epsilon} |X_t^i - X_{n\epsilon}^i|\, dt. \qquad (4.58)
\end{aligned}
$$

Taking expectations in (4.58) and using (4.52) and (4.53) we obtain

$$
\mathbb{E}[b_n^i\, Z_n^i] \le \big(-v_* + \chi d \,\| \operatorname{Hess} h_0 \|_\infty \big)\, \epsilon\, \mathbb{E}|Z_n^i|^2 + \kappa_2\, \epsilon^{3/2} \sqrt{\mathbb{E}|Z_n^i|^2}. \qquad (4.59)
$$

The same argument gives

$$
\mathbb{E}[b_n^i\, Z_{n-1}^i] \le \big(-v_* + \chi d \,\| \operatorname{Hess} h_0 \|_\infty \big)\, \epsilon\, \mathbb{E}|Z_{n-1}^i|^2 + \kappa_2\, \epsilon^{3/2} \sqrt{\mathbb{E}|Z_{n-1}^i|^2}. \qquad (4.60)
$$

Step (ii) Next we estimate $|a_n^i|$. By (4.45),

$$
|a_1^i| = \Big| \int_0^\epsilon \int_0^t G_{t-s}(X_s, X_t^i)\, ds\, dt \Big| \le \chi\beta \,\|\nabla g\|_\infty \frac{\epsilon^2}{2}.
$$

For $n \ge 2$, by (4.45) again,

$$
\begin{aligned}
|a_n^i| &= \Big| \int_{(n-1)\epsilon}^{n\epsilon} \Big(\int_0^{(n-1)\epsilon} \Big(G_{(n-1)\epsilon - s}(\tilde{Y}_s, Y_{n-1}^i) - G_{t-s}(X_s, X_t^i) \Big)\, ds \\
&\qquad - \int_{(n-1)\epsilon}^t G_{t-s}(X_s, X_t^i)\, ds \Big)\, dt \Big| \\
&\le \Big| \int_{(n-1)\epsilon}^{n\epsilon} A_n^{(1,i)}(t) + A_n^{(2,i)}(t)\, dt \Big| + \chi\beta \,\|\nabla g\|_\infty \frac{\epsilon^2}{2}, \qquad (4.61)
\end{aligned}
$$

where

$$
\begin{aligned}
A_n^{(1,i)}(t) &= \int_0^{(n-1)\epsilon} \Big(G_{(n-1)\epsilon - s}(X_s, X_t^i) - G_{t-s}(X_s, X_t^i) \Big)\, ds \\
A_n^{(2,i)}(t) &= \int_0^{(n-1)\epsilon} \Big(G_{(n-1)\epsilon - s}(\tilde{Y}_s, Y_{n-1}^i) - G_{(n-1)\epsilon - s}(X_s, X_t^i) \Big)\, ds.
\end{aligned}
$$

For $A_n^{(1,i)}(t)$, using that $g \in \mathcal{C}_b^2(\mathbb{R}^d)$, we have for $\theta_1 \le \theta_2$, $|\theta_1 - \theta_2| \le \epsilon$

$$
\begin{aligned}
&|G_{\theta_1}(\vec{x}, y) - G_{\theta_2}(\vec{x}, y)| \\
&\le \frac{\chi\beta\, e^{-\alpha\theta_1}}{N} \sum_{j=1}^N \int_{\mathbb{R}^d} \big(p(\theta_1, y, z) - p(\theta_2, y, z) \big) \nabla g(x_j - z)\, dz + \chi\beta \,\|\nabla g\|_\infty |e^{-\alpha\theta_1} - e^{-\alpha\theta_2}| \\
&\le \kappa_3 \sqrt{\epsilon}\, e^{-\alpha\theta_1}.
\end{aligned}
$$

Putting $\theta_1 = (n-1)\epsilon - s$ and $\theta_2 = t - s$, we obtain

$$
|A_n^{(1,i)}(t)| \le \kappa_3 \sqrt{\epsilon} \int_0^{(n-1)\epsilon} e^{-\alpha((n-1)\epsilon - s)}\, ds \le \frac{\kappa_3}{\alpha} \sqrt{\epsilon}. \qquad (4.62)
$$

For $A_n^{(2,i)}(t)$, note that for $s \in [0,(n-1)\epsilon]$, $t \in [(n-1)\epsilon, n\epsilon]$

$$
\begin{aligned}
&|G_\theta(\tilde{Y}_s^j, Y_{n-1}^i) - G_\theta(X_s, X_t^i)| \\
&= \frac{\chi\beta\, e^{-\alpha\theta}}{N} \sum_{j=1}^N \int_{\mathbb{R}^d} \Big(p(\theta, Y_{n-1}^i - z)\nabla g(\tilde{Y}_s^j - z) - p(\theta, X_t^i - z)\nabla g(X_s^j - z) \Big)\, dz \\
&= \frac{\chi\beta\, e^{-\alpha\theta}}{N} \sum_{j=1}^N \int p(\theta, X_t^i - z)\Big[\nabla g(\tilde{Y}_s^j - Y_{n-1}^i + X_t^i - z) - \nabla g(X_s^j - z) \Big]\, dz \\
&\le \frac{\chi\beta\, e^{-\alpha\theta}}{N} \sum_{j=1}^N d\| \operatorname{Hess} g\|_\infty (|\tilde{Y}_s^j - X_s^j + X_t^i - Y_{n-1}^i|)
\end{aligned}
$$

where we have used the substitution $z \mapsto Y_{n-1}^i - X_t^i + w$ in the first integral in the first equality. When $s \in [(k-1)\epsilon, k\epsilon)$, we have $\tilde{Y}_s = Y_{k-1}$ and

$$
|\tilde{Y}_s^j - X_s^j + X_t^i - Y_{n-1}^i| \le |Z_{k-1}^j| + |Z_n^i| + \mathcal{R}_{s,t}(i,j),
$$

where $\mathcal{R}_{s,t}(i,j) \doteq |X_{(k-1)\epsilon}^j - X_s^j| + |X_t^i - X_{n\epsilon}^i| + |Y_n^i - Y_{n-1}^i|$ is an error term which, in view of (4.52) and (4.53) satisfies,

$$
\mathbb{E}[\mathcal{R}_{s,t}(i,j)^2] \le \kappa_4\, \epsilon. \tag{4.63}
$$

From the above calculations and recalling that $C_2 = d\| \operatorname{Hess} g\|_\infty$

$$
\begin{aligned}
|A_n^{(2,i)}(t)| &= \Big| \sum_{k=1}^{n-1} \int_{(k-1)\epsilon}^{k\epsilon} G_{(n-1)\epsilon - s}(Y_{k-1}, Y_{n-1}^i) - G_{(n-1)\epsilon - s}(X_s, X_t^i)\, ds \Big| \\
&\le \frac{\chi\beta\, C_2}{N} \sum_{j=1}^N \sum_{k=1}^{n-1} \int_{(k-1)\epsilon}^{k\epsilon} e^{-\alpha((n-1)\epsilon - s)} \left(|Z_{k-1}^j| + |Z_n^i| + \mathcal{R}_{s,t}(i,j) \right) ds. \qquad (4.64)
\end{aligned}
$$

Using (4.62) and (4.64) in (4.61),

$$
\begin{aligned}
|a_n^i| \le \kappa_5\, \epsilon^{3/2} \;+\; \frac{\chi\beta\, C_2}{N} \int_{(n-1)\epsilon}^{n\epsilon} \Big(\sum_{k=1}^{n-1} \int_{(k-1)\epsilon}^{k\epsilon} e^{-\alpha((n-1)\epsilon - s)} \\
\sum_{j=1}^N \left(|Z_{k-1}^j| + |Z_n^i| + \mathcal{R}_{s,t}(i,j) \right) ds \Big) dt. \qquad (4.65)
\end{aligned}
$$

Step (iii) We now combine steps (i) and (ii) to obtain an inequality for

$$
f_n \doteq \frac{1}{N} \sum_{i=1}^N \mathbb{E}|Z_n^i|^2.
$$

This inequality is the discrete analogue of a differential inequality similar to (4.15). By exchangeability we have

$$
f_n = \mathbb{E}|Z_n^i|^2,\ i = 1, \cdots, N. \tag{4.66}
$$

By Cauchy-Schwartz inequality, the fact $\frac{1}{N}\sum_{i=1}^N \sqrt{\mathbb{E}|Z_n^i|^2} \le \sqrt{f_n}$ and the bound (4.63),

$$
\begin{aligned}
&\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \mathbb{E}\Big[|Z_n^i| \left(|Z_{k-1}^j| + |Z_n^i| + \mathcal{R}_{s,t}(i,j) \right) \Big] \\
&\le f_n + \Big(\frac{1}{N} \sum_{i=1}^N \sqrt{\mathbb{E}|Z_n^i|^2} \Big) \Big(\frac{1}{N} \sum_{i=1}^N \sqrt{\mathbb{E}|Z_{k-1}^i|^2} \Big) + \frac{1}{N} \sum_{i=1}^N \sqrt{\mathbb{E}|Z_n^i|^2} \left(\frac{1}{N} \sum_{j=1}^N \mathbb{E}|\mathcal{R}_{s,t}(i,j)|^2 \right)^{1/2} \\
&\le f_n + \sqrt{f_n\, f_{k-1}} + \kappa_6\, \sqrt{\epsilon}\, \sqrt{f_n}. \qquad (4.67)
\end{aligned}
$$

Let

$$\sigma_{n,k} \doteq \int_{(k-1)\epsilon}^{k\epsilon} e^{-\alpha((n-1)\epsilon - s)}\, ds = \frac{e^{-\alpha\epsilon\, n}}{\alpha}\big(e^{\alpha\epsilon(k+1)} - e^{\alpha\epsilon\, k}\big) \tag{4.68}$$

Then $\sum_{k=1}^{n-1} \sigma_{n,k} = (1 - e^{-\alpha(n-1)\epsilon})/\alpha \le 1/\alpha$ and from (4.65) and (4.67) we have

$$\frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[|a_n^i\, Z_n^i|] \le \frac{\chi\beta\, C_2}{\alpha}\,\epsilon\, f_n + \chi\beta\, C_2\,\epsilon\,\sqrt{f_n}\sum_{k=1}^{n-1}\sigma_{n,k}\,\sqrt{f_{k-1}} + \kappa_7\,\epsilon^{3/2}\sqrt{f_n}. \tag{4.69}$$

Similarly,

$$\frac{1}{N}\sum_{i=1}^{N} \mathbb{E}[|a_n^i\, Z_{n-1}^i|] \le \frac{\chi\beta\, C_2}{\alpha}\,\epsilon\, f_{n-1} + \chi\beta\, C_2\,\epsilon\,\sqrt{f_{n-1}}\sum_{k=1}^{n-1}\sigma_{n,k}\,\sqrt{f_{k-1}} + \kappa_7\,\epsilon^{3/2}\sqrt{f_{n-1}}. \tag{4.70}$$

Therefore, summing over i in (4.55) and then using (4.59), (4.60) (4.69) and (4.70), we obtain a nonlinear "difference-summation" inequality

$$\begin{aligned} f_n - f_{n-1} \le\ & -\tilde{\lambda}\,\epsilon\big(f_n + f_{n-1}\big) + \tilde{C}_2\,\epsilon\big(\sqrt{f_n} + \sqrt{f_{n-1}}\big)\sum_{k=1}^{n-1}\sigma_{n,k}\,\sqrt{f_{k-1}} \\ & + \kappa_8\,\epsilon^{3/2}\big(\sqrt{f_n} + \sqrt{f_{n-1}}\big), \end{aligned} \tag{4.71}$$

where $\sigma_{n,k}$ is defined in (4.68), and as before, $-\tilde{\lambda} \doteq -v_* + \chi C_1 + C_2\chi\beta/\alpha$ and $\tilde{C}_2 \doteq C_2\,\chi\,\beta$. We can rewrite (4.71) as

$$\frac{f_n - f_{n-1}}{\epsilon} \le -\,\tilde{\lambda}\,\big(f_n + f_{n-1}\big) + \big(\sqrt{f_n} + \sqrt{f_{n-1}}\big)\Big(\tilde{C}_2\sum_{k=1}^{n-1}\sigma_{n,k}\,\sqrt{f_{k-1}} + \kappa_8\,\sqrt{\epsilon}\Big) \tag{4.72}$$

which is the discrete analogue of a differential inequality similar to (4.15).

Step (iv) Finally, we use (4.71) to obtain a uniform (in N, n, ϵ) upper bound for f_n, under Assumption 2.4. Note that under this assumption $\tilde{\lambda} > 0$. Let

$$\delta \doteq \frac{1 - \tilde{\lambda}\,\epsilon}{1 + \tilde{\lambda}\,\epsilon}.$$

Note that $\delta \in (0,1)$ for ϵ small enough. Moving the negative term $-\tilde{\lambda}\,\epsilon\big(f_n + f_{n-1}\big)$ to the other side of the inequality in (4.72) and then multiplying both sides by $\delta^{-(n-1)}$ and letting $g_n = f_n/\delta^{n-1}$ we obtain

$$\begin{aligned} g_n - g_{n-1} &\le \frac{\epsilon}{(1+\tilde{\lambda}\,\epsilon)\,\delta^{n-1}}\,\big(\sqrt{f_n} + \sqrt{f_{n-1}}\big)\Big(\tilde{C}_2\sum_{k=1}^{n-1}\sigma_{n,k}\,\sqrt{f_{k-1}} + \kappa_8\,\sqrt{\epsilon}\Big) \\ &\le \frac{\epsilon}{(1+\tilde{\lambda}\,\epsilon)\,\delta^{n/2}}\,\big(\sqrt{g_n} + \sqrt{g_{n-1}}\big)\Big(\tilde{C}_2\sum_{k=1}^{n-1}\sigma_{n,k}\,\delta^{(k-2)/2}\,\sqrt{g_{k-1}} + \kappa_8\,\sqrt{\epsilon}\Big) \end{aligned} \tag{4.73}$$

where the second inequality follows on noting that $\sqrt{f_n} + \sqrt{f_{n-1}} \le \delta^{(n-2)/2}(\sqrt{g_n} + \sqrt{g_{n-1}})$ for $n \ge 2$. Similar to the proof of Theorem 3.4 we consider a small positive perturbation of g_n and let $h_\theta(n) \doteq \sqrt{g_n + \theta^2}$ where $\theta > 0$. The inequality in (4.73) then implies

$$\begin{aligned} & h_\theta^2(n) - h_\theta^2(n-1) \\ & \quad \le \frac{\epsilon}{(1+\tilde{\lambda}\,\epsilon)\,\delta^{n/2}}\,\big(h_\theta(n) + h_\theta(n-1)\big)\Big(\tilde{C}_2\sum_{k=1}^{n-1}\sigma_{n,k}\,\delta^{(k-2)/2}\,h_\theta(k-1) + \kappa_8\,\sqrt{\epsilon}\Big). \end{aligned}$$

Since h_θ is strictly positive, we obtain

$$h_\theta(n) - h_\theta(n-1) \leq \frac{\epsilon}{(1+\tilde{\lambda}\,\epsilon)\,\delta^{n/2}} \Big(\tilde{C}_2 \sum_{k=1}^{n-1} \sigma_{n,k}\, \delta^{(k-2)/2}\, h_\theta(k-1) + \kappa_8 \sqrt{\epsilon} \Big). \tag{4.74}$$

Consider the recursion equation obtained by replacing the inequality in (4.74) by equality, namely

$$k_n - k_{n-1} = A \Big(\sum_{i=1}^{n-1} \sigma_{n,i}\, \delta^{(i-2-n)/2}\, k_{i-1} \Big) + \frac{B}{\delta^{n/2}}, \tag{4.75}$$

and $k_0 = h_\theta(0) = \theta$, where

$$A = \frac{\epsilon\, \tilde{C}_2}{1+\tilde{\lambda}\,\epsilon} = O(\epsilon) \quad \text{and} \quad B = \frac{\epsilon^{3/2}\kappa_8}{1+\tilde{\lambda}\,\epsilon} = O(\epsilon^{3/2}). \tag{4.76}$$

By evaluating $\Delta_n - \frac{e^{-\alpha\epsilon}}{\delta^{1/2}}\Delta_{n-1}$ where $\Delta_{n-1} \doteq k_n - k_{n-1}$, we can convert the above equation (4.75) to the following second order linear difference equation:

$$\begin{cases} k_{n+1} - p\, k_n + q\, k_{n-1} = A_n,\ n \geq 2, \\ \quad k_0 = \theta, \quad k_1 = \theta + \frac{B}{\delta^{1/2}}, \end{cases} \tag{4.77}$$

where

$$p = 1 + \frac{e^{-\alpha\epsilon}}{\delta^{1/2}} \to 2,\ q = \frac{e^{-\alpha\epsilon}}{\delta^{1/2}} - \frac{A\,(1-e^{-\alpha\epsilon})}{\alpha\,\delta^{3/2}} \to 1 \text{ as } \epsilon \to 0 \text{ and } A_n = \frac{B(1-e^{-\alpha\epsilon})}{\delta^{(n+1)/2}}.$$

Solution of (4.77) can be explicitly given as

$$k_n = c_1^{(\theta)}\, r_1^n + c_2^{(\theta)} r_2^n + c_3\, \delta^{-n/2} \tag{4.78}$$

where $r_1 > r_2 > 0$ are the distinct positive real roots of $r^2 - pr + q = 0$, and

$$c_2^{(\theta)} = \frac{r_1(\theta - c_3) - \theta - \delta^{-1/2}(B - c_3)}{r_1 - r_2},\ c_1^{(\theta)} = \theta - c_3 - c_2^{(\theta)} \text{ and } c_3 = \frac{B(1-e^{-\alpha\epsilon})}{1 - p\,\delta^{1/2} + q\,\delta}. \tag{4.79}$$

On other hand, induction easily gives $h_\theta(n) \leq k_n$ for all $n \geq 0$ and $\theta > 0$.

Thus we have for all $\theta > 0$

$$\sqrt{f_n} = \delta^{(n-1)/2}\sqrt{g_n} = \delta^{(n-1)/2}\sqrt{h_\theta^2(n) - \theta^2} \leq \delta^{(n-1)/2}\, k_n.$$

Sending $\theta \to 0$ in (4.79) we obtain

$$\sqrt{f_n} \leq \frac{c_1}{\delta^{1/2}}\,(\delta^{1/2}\, r_1)^n + \frac{c_2}{\delta^{1/2}}\,(\delta^{1/2}\, r_2)^n + \frac{c_3}{\delta^{1/2}}$$

where

$$c_2 = \frac{(\delta^{-1/2} - r_1)c_3 - \delta^{-1/2}B}{r_1 - r_2} \quad \text{and} \quad c_1 = -c_3 - c_2. \tag{4.80}$$

We claim that for some $\kappa_9, \kappa_{10} \in (0, \infty)$

$$0 < c_3 < \kappa_9\sqrt{\epsilon},\ r_1 - r_2 > \kappa_9\epsilon \quad \text{and} \quad 0 < 1 - \delta^{1/2}\, r_1 < \kappa_9\epsilon \tag{4.81}$$

for $\epsilon \in (0,\, \kappa_{10})$. This will imply from (4.76) that both $|c_3|$ and $|c_2|$ are of order $\sqrt{\epsilon}$ and hence by (4.80), $|c_1|$ is also of order $\sqrt{\epsilon}$. Also, $0 < 1 - \delta^{1/2}\, r_1$ implies $(\delta^{1/2}\, r_1)^n \to 0$ as $n \to \infty$. Therefore we obtain the desired bound

$$\sup_{n \geq 0} \sqrt{f_n} \leq \kappa_{11}\, \sqrt{\epsilon},$$

for ϵ sufficiently small. The claim (4.81) is established in the Appendix. The proof is now complete in view of (4.66). □

We now complete the proof of Corollary 3.11.

Proof of Corollary 3.11. The first statement in the corollary is immediate from Corollary 3.5 and Theorem 3.10. For the second statement, we have from triangle inequality

$$\mathcal{W}_2(\mu_n^{N,\epsilon}, \mu_{n\epsilon}) \leq \mathcal{W}_2(\mu_n^{N,\epsilon}, \mu_{n\epsilon}^N) + \mathcal{W}_2(\mu_{n\epsilon}^N, \mu_{n\epsilon}).$$

Also, from Theorem 3.10,

$$\limsup_{N\to\infty} \sup_{n\geq 1} \mathbb{E}\mathcal{W}_2^2(\mu_n^{N,\epsilon}, \mu_{n\epsilon}^N) \leq \limsup_{N\to\infty} \sup_{n\geq 1} \Big(\frac{1}{N}\sum_{i=1}^N \mathbb{E}|Y_n^{i,N,\epsilon} - X_{n\epsilon}^{i,N}|^2\Big) \leq C\epsilon.$$

The result now follows on combining the above two displays with Corollary 3.5. □

A Proof of (4.81)

To see the first inequality in the claim (4.81), note that

$$\frac{c_3}{\sqrt{\epsilon}} = \frac{\delta^{1/2}\alpha\kappa_8}{\alpha\left(\frac{\delta^{1/2}-\delta}{\epsilon}\right)(1+\tilde{\lambda}\,\epsilon) - \tilde{C}_2}. \tag{A.1}$$

The inequality is now a consequence of the observation that $\delta^{1/2} \to 1$ as $\epsilon \to 0$ and

$$\lim_{\epsilon\to 0} \alpha\left(\frac{\delta^{1/2}-\delta}{\epsilon}\right)(1+\tilde{\lambda}\,\epsilon) - \tilde{C}_2 = \alpha\,\tilde{\lambda} - \tilde{C}_2 > 0,$$

where the last inequality is from Assumption 2.4. Hence the first estimate holds.

The second inequality in the claim (4.81) follows on observing that, as $\epsilon \to 0$,

$$\frac{(r_1-r_2)^2}{\epsilon^2} = \frac{p^2-4q}{\epsilon^2} = \frac{1}{\epsilon^2}\Big(1-\frac{e^{-\alpha\epsilon}}{\delta^{1/2}}\Big)^2 + \frac{4A\,(1-e^{-\alpha\epsilon})}{\epsilon^2\alpha\,\delta^{3/2}} \geq \frac{4A\,(1-e^{-\alpha\epsilon})}{\epsilon^2\alpha\,\delta^{3/2}} \to 4\tilde{C}_2.$$

For the third inequality, we need to show that $1-\delta^{1/2}\,r_1$ is of order at most $\sqrt{\epsilon}$. Clearly

$$\begin{aligned} 1-\sqrt{\delta}r_1 &= 1 - \frac{1}{2}\Big[\sqrt{(p^2-4q)\delta} + p\sqrt{\delta}\Big] \\ &= \frac{1}{2}\Big[(2-p\sqrt{\delta}) - \sqrt{(p^2-4q)\delta}\Big]. \end{aligned}$$

Regarding p, q and δ as functions of ϵ, we see that

$$\begin{aligned} p &= p(\epsilon) = 2 + \epsilon p'(0) + O(\epsilon^2), \\ q &= q(\epsilon) = 1 + \epsilon q'(0) + O(\epsilon^2), \\ \sqrt{\delta} &= \sqrt{\delta(\epsilon)} = 1 + \frac{1}{2}\epsilon\delta'(0) + O(\epsilon^2) \end{aligned}$$

Thus

$$(p^2-4q) = 4\epsilon(p'(0)-q'(0)) + O(\epsilon^2) = O(\epsilon^2)$$

where the last equality follows on checking that $p'(0) = q'(0)$. This shows that $\sqrt{(p^2-4q)\delta} = O(\epsilon)$. Also, clearly $2-p\sqrt{\delta} = O(\epsilon)$ and so

$$(2-p\sqrt{\delta}) - \sqrt{(p^2-4q)\delta} = O(\epsilon).$$

The desired inequality follows. □

References

[1] J. Bedrossian, N. Rodríguez and A. L. Bertozzi. Local and global well-posedness for aggregation equations and Patlak-Keller-Segel models with degenerate diffusion. *Nonlinearity*, 24(6), 1683, 2011.

[2] D. Benedetto, E. Caglioti, J. A. Carrillo and M. Pulvirenti. A non-Maxwellian steady distribution for one-dimensional granular media. *J. Statist. Phys.*, 91, 979–990, 1998.

[3] I. Bihari, A generalization of a lemma of Bellman and its application to uniqueness problems of differential equations. *Acta Mathematica Hungarica*, 7(1), 81–94, 1956.

[4] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Relat. Fields*, 137(3–4), 541–593, 2007.

[5] K. Carrapatoso and S. Mischler. Uniqueness and long time asymptotic for the parabolic-parabolic Keller-Segel equation. *arXiv preprint*, arXiv:1406.6006.

[6] V. Calvez and L. Corrias. The parabolic-parabolic Keller-Segel model in R2. *Communications in Mathematical Sciences*, 6(2), 417–447, 2008.

[7] J. A. Carrillo, S. Hittmeir and A. Jüngel. Cross diffusion and nonlinear diffusion preventing blow up in the Keller–Segel model. *Mathematical Models and Methods in Applied Sciences*, 22(12), 1250041, 2012.

[8] J. Carrillo, R. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19, 971–1018, 2003.

[9] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Probab. Theory Relat. Fields*, 140(1–2), 19–40, 2008.

[10] Di Francesco, M. and Rosado, J. Local and global well-posedness for aggregation equations and Patlak-Keller-Segel models with degenerate diffusion. *Nonlinearity*, 21(11), 2715, 2008.

[11] I. Fatkullin, A study of blow-ups in the Keller-Segel model of chemotaxis. *Nonlinearity*, 26 (2013), 81–94.

[12] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Relat. Fields*, 162(3–4), 707–738, 2015.

[13] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences I. *Jahresbericht der Deutschen Mathematiker Vereinigung*, 106(2), 51–70. 2003.

[14] D. Horstmann. From 1970 until present: the Keller-Segel model in chemotaxis and its consequences II. *Jahresbericht der Deutschen Mathematiker Vereinigung*, 106(2), 51–70. 2004.

[15] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3), 399–415, 1970.

[16] Peter M. Kotelenez and Thomas G. Kurtz. Macroscopic limits for stochastic partial differential equations of McKeanVlasov type. *Probab. Theory Relat. Fields*, 146, 189–222, 2010.

[17] F. Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Prob.*, 13(2), 540–560, 2003.

[18] H. P. McKean Jr. Propagation of chaos for a class of non-linear parabolic equations. *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ.*, 41–57, 1967.

[19] S. Méléard. Asymptotic behavior of some interacting particle systems; McKean-Vlasov and Boltzmann models. *Lecture Notes in Mathematics*, 1627, 42–95, Springer, Berlin, 1996.

[20] C. S. Patlak. Random walk with persistence and external bias. *Bulletin of Mathematical Biophysics*, 15(3), 311–338. 1953.

[21] F. Schweitzer, and J. D. Farmer, *Brownian agents and active particles: collective dynamics in the natural and social sciences.* Springer Science and Business Media. 2007.

[22] A. Stevens. The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems. *SIAM Journal on Appl. Math.*, 61(1), 183–212, 2000.

[23] A.-S. Sznitman. Topics in propagation of chaos. *Lecture Notes in Mathematics*, 1464, 165–251, 1991.

[24] D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Appl*, 8(2), 163–198, 2002.

[25] Y. Tamura. Free energy and the convergence of distributions of diffusion processes of McKean type. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 34(2), 443–484, 1987.

[26] C. Villani. *Optimal transport: old and new*, Vol. 338. Springer Science and Business Media, 2008.

Acknowledgments. We would like to thank Tom Kurtz for many helpful discussions related to this work. We also thank the referee for a careful review.

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1]LOCKSS: Lots of Copies Keep Stuff Safe `http://www.lockss.org/`
[2]EJMS: Electronic Journal Management System `http://www.vtex.lt/en/ejms.html`
[3]IMS: Institute of Mathematical Statistics `http://www.imstat.org/`
[4]BS: Bernoulli Society `http://www.bernoulli-society.org/`
[5]Project Euclid: `https://projecteuclid.org/`
[6]IMS Open Access Fund: `http://www.imstat.org/publications/open.htm`