RANDOM-BATCH METHOD FOR MULTI-SPECIES STOCHASTIC INTERACTING PARTICLE SYSTEMS

ESTHER S. DAUS, MARKUS FELLNER, AND ANSGAR JÜNGEL

ABSTRACT. A random-batch method for multi-species interacting particle systems is proposed, extending the method of S. Jin, L. Li, and J.-G. Liu [*J. Comput. Phys.* 400 (2020), 108877]. The idea of the algorithmus is to randomly divide, at each time step, the ensemble of particles into small batches and then to evolve the interaction of each particle within the batches until the next time step. This reduces the computational cost by one order of magnitude, while keeping a certain accuracy. It is proved that the L^2 error of the error process behaves like the square root of the time step size, uniformly in time, thus providing the convergence of the scheme. The numerical efficiency is tested for some examples, and numerical simulations of the opinion dynamics in a hierarchical company, consisting of workers, managers, and CEOs, are presented.

1. INTRODUCTION

The collective behavior of particles or agents of multiple species can be described by interacting particle systems, which are an important tool for modeling complex real-world phenomena with applications in physics, biology, and social sciences. The binary interaction between all particles makes numerical simulations very demanding when many agents need to be modeled, which explains the need for efficient algorithms. Averaged results can be obtained from the associated mean-field equations, while the individual dynamics is captured by direct simulations, using fast summation algorithms, like fast multipole methods [11], wavelet transforms [3], or variants of Monte–Carlo methods [5]. Recently, motivated by mini-batch gradient descent in machine learning (see, e.g., [23]), the authors of [16] suggested to use small random batches in interacting particle systems, which results in the reduction of the computational cost per time step from $O(N^2)$ to $O(N)$ (N being the number of particles or agents). Compared to other efficient sampling methods, like the Ewald summation or the fast multipole method, the random-batch method is easier to implement and more flexible to apply in complex systems. The results of [16] are valid in the single-species case. In this paper, we generalize their approach to multi-species systems.

In particular, we work out the dependence of the L^2 error with respect to the batch sizes of the different species and discuss the case of multiplicative noise.

1.1. **Setting.** The dynamics of the multi-species system is described by

$$
(1) \qquad \mathrm{d}X_i^k = -\nabla V_i(X_i^k)\mathrm{d}t + \sum_{j=1}^{n} \alpha_{ij} \sum_{\substack{\ell=1 \\ (i,k)\neq(j,\ell)}}^{N_j} K_{ij}(X_i^k - X_j^\ell)\mathrm{d}t + \sigma_i \mathrm{d}B_i^k(t),
$$

$$
(2) \qquad X_i^k(0) = X_{0,i}^k \quad \text{for } i = 1,\ldots,n, \ k = 1,\ldots,N_i,
$$

where

$$
(3) \qquad \alpha_{ij} = \frac{1}{N_j - \delta_{ij}}, \quad i,j = 1,\ldots,n.
$$

The stochastic process $X_i^k(t) \in \mathbb{R}^d$ $(d \geq 1)$ represents the position of the kth particle (or the features of the kth agent) of species i in a system of $N = \sum_{i=1}^n N_i$ particles. The function ∇V_i describes some (given) external force, K_{ii} and K_{ij} are the interaction kernels between particles of the same and of different species, respectively, $\sigma_i > 0$ are diffusion coefficients, and B_i^k are N independent standard Brownian motions. The initial data $X_{0,i}^1, \ldots, X_{0,i}^{N_i}$ are assumed to be independent and identically distributed.

Equations (1) can be used to model the information flow through social networks [2], the dynamics of opinions [8], the herding of sheep by dogs [26], or the segregation behavior of populations [6]. Stochastic gradient descent can be interpreted as the evolution of interacting particle systems governed by a potential related to the objective function used to train neural networks [25].

1.2. **Random-batch method.** The random-batch method is defined as follows. Let the number of particles $N_i \in \mathbb{N}$ of the ith species be an even number, where $i = 1,\ldots,n$. We introduce the time steps $t_m = m\tau$ with the time step size $\tau > 0$ and $m = 1,\ldots,M :=$ $\lceil T/\tau \rceil$, and $T > 0$ is the end time. For a given $m \in \{1,\ldots,M\}$, we divide the set $\{1,\ldots,N_i\}$ randomly into b_i batches $\mathcal{C}_{i,1},\ldots,\mathcal{C}_{i,b_i}$ of size p_i. This means that we choose $p_i \geq 2$ and $b_i \geq 1$ such that $N_i = b_i p_i$, and we consider not all interactions but only those in the same batch. Furthermore, we introduce the super-batches $\mathcal{C}_r = \{(i,k) : k \in \mathcal{C}_{i,r}\}$ for $1 \leq r \leq \max\{b_1,\ldots,b_n\}$ (see Figure 1). For any particle X_i^k, there exists exactly one super-batch such that $(i,k) \in \mathcal{C}_r$ for some $r \geq 0$.

We solve the particle system in the time interval $(t_{m-1}, t_m]$ with initial datum $\widetilde{X}_i^k(t_{m-1})$. The random-batch process $\widetilde{X}_i^k$ is defined for $t_{m-1} < t \leq t_m$ as the solution to

$$
(4) \qquad \mathrm{d}\widetilde{X}_i^k = -\nabla V_i(\widetilde{X}_i^k)\mathrm{d}t + \sum_{j=1}^{n} \beta_{ij} \sum_{\substack{\ell \in \mathcal{C}_{j,r} \\ (i,k)\neq(j,\ell)}} K_{ij}(\widetilde{X}_i^k - \widetilde{X}_j^\ell)\mathrm{d}t + \sigma_i \mathrm{d}B_i^k,
$$

where

$$
(5) \qquad \beta_{ij} = \frac{b_i}{(p_j - \delta_{ij})\min\{b_i, b_j\}}, \quad i,j = 1,\ldots,n.
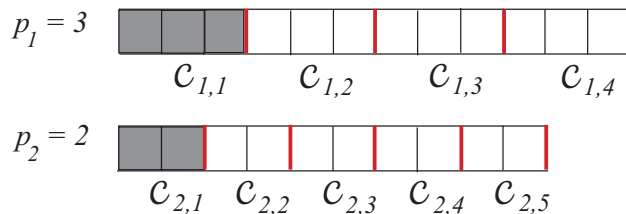$$

FIGURE 1. Batches $\mathcal{C}_{i,r}$ for a two-species system with $N = 22$ particles, four batches of size $p_1 = 3$, and five batches of size $p_2 = 2$. The particles in the super-batch $\mathcal{C}_1$ are marked in grey color.

Instead of summing over all interactions, the sum in (4) only accounts for the interactions in each small batch. Observe that we use the same Brownian motions as in (1). The sum over all $\ell \in \mathcal{C}_{j,r}$ means that we sum over all (j, ℓ) which are in the same super-batch as (i, k). The factor $b_i / \min\{b_i, b_j\}$ in (5) does not appear in [16]; it is necessary to achieve consistency and convergence of the scheme. The scaling results from the different number of nontrivial batches $\mathcal{C}_{i,r}$ of the different species. Indeed, let $b_j < b_i$. From the viewpoint of the particles of the ith species, they interact with the particles of the jth species only with the share b_j / b_i, since $\mathcal{C}_{i,r}$ is empty for $r > b_j$. This yields the factor b_i / b_j. The random-batch algorithm is summarized in Algorithm 1.2.

Algorithm 1 (Pseudo-code for the multi-species random-batch algorithm)

1: **for** $k = m, \ldots, M$ **do**
2:      **for** $i = 1, \ldots, n$ **do**
3:          Divide $\{1, \ldots, N_i\}$ randomly into b_i batches $\mathcal{C}_{i,1}, \ldots, \mathcal{C}_{i,b_i}$ with size p_i each.
4:          **for** $r = 1, \ldots, b_i$ **do**
5:              For every $(i, k) \in \mathcal{C}_r$, update $\widetilde{X}_i^k$ by solving (4) in the interval $(t_{m-1}, t_m]$ with initial datum $\widetilde{X}_i^k(t_{m-1})$.
6:          **end for**
7:      **end for**
8: **end for**

When we allow for pairwise interactions between *all* particles, the computational cost at each time step is of order $O(N^2)$. Since we have M time steps, the total cost of this naive algorithm is $O(MN^2)$. In the random-batch method, each particle ends up in exactly one super-batch $\mathcal{C}_r$ for some $r \geq 1$ and is chosen only once (i.e. without replacement). Then the total computatinal cost becomes $O(pMN)$, where $p = \sum_{i=1}^n p_i$. As p is typically a small number (often $p_i = 2$), the total cost has been reduced by approximately one order of magnitude. We show in this paper that, under suitable conditions on the external potentials and the kernel functions, the L^2 error of the error process $\widetilde{X}_i^k(t) - X_i^k(t)$ converges to zero as $\tau \to 0$ uniformly in time, and the convergence is, as expected, of order $O(\sqrt{\tau})$. The idea of the method is the fact that in time average, the random force is consistent

with the full interaction (see Proposition 2), and the convergence is like in the law of large numbers, but in time.

1.3. **Main result.** We start with some definitions and notation. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space, let $\xi_{m-1,i}$ denote the random division of batches of species i at t_{m-1}, and set $\xi_{m-1} = (\xi_{m-1,1}, \ldots, \xi_{m-1,n})$. We define the filtrations $(\mathcal{F}_m)_{m\geq 0}$ and $(\mathcal{G}_m)_{m\geq 0}$ by

$$\mathcal{F}_{m-1} = \sigma\big(X_{0,i}^k,\, B_i^k(t),\, \xi_{j,i} : t \leq t_{m-1},\, 1 \leq i \leq n,\, 1 \leq j \leq m-1\big),$$
$$\mathcal{G}_{m-1} = \sigma\big(X_{0,i}^k,\, B_i^k(t),\, \xi_{j,i} : t \leq t_{m-1},\, 1 \leq i \leq n,\, 1 \leq j \leq m-2\big).$$

The set $\mathcal{F}_{m-1}$ contains the information how the batches are constructed for $t \in [t_{k-1}, t_k)$. Denoting by $\sigma(\xi_{m-1,i})$ the σ-algebra generated by $\xi_{m-1,i}$, it holds that $\mathcal{F}_{m-1} = \sigma(\mathcal{G}_{m-1} \cup \sigma(\xi_{m-1,1}) \cup \cdots \cup \sigma(\xi_{m-1,n}))$. We write $\|\cdot\|_p = (\mathbb{E}|\cdot|^p)^{1/p}$ to denote the $L^p(\Omega)$ norm for $1 \leq p < \infty$ and set $\|\cdot\| = \|\cdot\|_2$. In the whole paper, $C > 0$, $C_i > 0$ denote generic constants whose values change from line to line. We set $X = (X_i^k)_{i=1,\ldots,n}^{k=1,\ldots,N_i}$ and $\widetilde{X} = (\widetilde{X}_i^k)_{i=1,\ldots,n}^{k=1,\ldots,N_i}$.

We impose the following assumptions:

(A1) Kernel functions: $K_{ij} \in C^2(\mathbb{R}^d)$ is bounded, Lipschitz continuous with Lipschitz constant $L_{ij} > 0$, and has a bounded second derivative.

(A2) Potential functions: $V_i \in C^2(\mathbb{R}^d)$, and there exist $C_V > 0$, $q_i > 0$ such that for all $x \in \mathbb{R}^d$,

$$|\nabla V_i(x)| + |D^2 V_i(x)| \leq C_V(1 + |x|^{q_i}), \quad i = 1, \ldots, n.$$

(A3) Strong convexity: The function $x \mapsto V_i(x) - r_i|x|^2/2$ is convex, where $r_i > 2\sum_{j=1}^n \max\{L_{ij}, L_{ji}\}$ and $i = 1, \ldots, n$.

(A4) Synchronous coupling: $X_i^k(0) = \widetilde{X}_i^k(0) = X_{0,i}^k$ for $i = 1, \ldots, n$, $k = 1, \ldots, N_i$, where $X_{0,i}^1, \ldots, X_{0,i}^{N_i}$ are independent and identically distributed, and $X_{0,i}^k$ is $\mathbb{F}_0$-measurable with $\mathbb{E}|X_{0,i}^k|^{2\max\{1,q_i\}} < \infty$.

Under these assumptions (in particular, the Lipschitz continuity), standard results for stochastic differential equations [19] guarantee that (1) and (4) have (up to $\mathbb{P}$-distinguishability) unique strong solutions. The polynomial growth conditions on ∇V_i and $D^2 V_i$ are needed to prove the stability; see Lemma 4. The smallness condition on the Lipschitz constants of the kernel functions ensures that the evolution group of the deterministic part of (1) is a contraction, thus yielding error bounds uniformly in time.

Our main result reads as follows.

Theorem 1 (Error estimate)**.** *Let Assumptions (A1)–(A4) hold. Then there exists a constant $C > 0$, which is independent of $(b_i, p_i)_{i=1,\ldots,n}$, m, and T, such that*

$$\sup_{0 < t < T} \sum_{i=1}^n \|(X_i^k - \widetilde{X}_i^k)(t)\| \leq C\sqrt{\tau}\bigg(\sum_{i=1}^n \Gamma_i\bigg)^{1/2} + C\tau(1 + \theta^\gamma), \quad t > 0,$$

where

$$(6) \qquad \theta = \frac{\max_{j=1,\dots,n} b_j}{\min_{j=1,\dots,n} b_j}, \quad \gamma = 3\big(\max\{1, q_1, \dots, q_n\} + 1\big),$$

$$(7) \qquad \Gamma_i = \sum_{\substack{j,j'=1 \\ j,j'\neq i, j\neq j'}}^{n} \left(\frac{\max\{b_i, b_j, b_{j'}\}}{\max\{b_j, b_{j'}\}} - 1 \right) + \sum_{\substack{j=1 \\ j\neq i}}^{n} \left(\frac{b_i - \min\{b_i, b_j\}}{\min\{b_i, b_j\}} - \frac{2 - \max\{b_i, b_j\}}{N_j} \right.$$

$$\left. + \frac{b_i}{p_j \min\{b_i, b_j\}} \right) + \left(\frac{1}{p_i - 1} - \frac{1}{N_i - 1} \right) \geq 0. \quad i = 1, \dots, n,$$

and q_i is introduced in Assumption (A2).

The theorem generalizes [16, Theorem 3.1] to the multi-species case. Indeed, if $n = 1$, Γ_1 reduces to $1/(p_1 - 1) - 1/(N_1 - 1)$ and $\theta = 1$. Then the error bound becomes $C\sqrt{\tau/(p_1 - 1)} + C\tau$, which corresponds to (3.9) in [16]. Compared to the result in [16], Theorem 1 shows the influence of the different batch sizes b_i of the species. Indeed, if the batch sizes are very different, θ is much larger than one, which increases the constant in the error estimate. This behavior is also observed in the numerical simulations; see Section 6.3.

The proof of Theorem 1 is based on estimates for the error process $Z_i^k := \widetilde{X}_i^k - X_i^k$. Since the noise terms are the same, Z_i^k solves

$$\mathrm{d}Z_i^k(t) = -(\nabla V_i(\widetilde{X}_i^k) - \nabla V_i(X_i^k))\mathrm{d}t + \sum_{j=1}^{n} \alpha_{ij} \sum_{\substack{\ell=1 \\ (j,k)\neq(j,\ell)}} \Delta K_{ij}^\ell \mathrm{d}t + \chi_i^k(\widetilde{X})\mathrm{d}t$$

for $t_{m-1} < t \leq t_m$, where $\Delta K_{ij}^\ell := K_{ij}(\widetilde{X}_i^k - \widetilde{X}_j^\ell) - K_{ij}(X_i^k - X_j^\ell)$ and $\chi_i^k(\widetilde{X})$ is a remainder term (defined in (9) below). An important ingredient of the proof is the computation of the variance of χ_i^k, which is more involved than in [16], since the multi-species case requires to distinguish several cases in the choice of indices (i, k) and (j, ℓ).

A straightforward computation, detailed in Section 4, shows that the error process satisfies

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}|Z_i^k(t)|^2 \leq - \left(r_i - 2\sum_{j=1}^{n} \max\{L_{ij}, L_{ji}\} \right) \mathbb{E}|Z_j^k(t)|^2 + \mathbb{E}\big(\chi_i^k(\widetilde{X}(t)) \cdot Z_i^k(t)\big).$$

The main difficulty is the estimate of the last term. The idea is to write it in terms of differences $Z_i^k(t) - Z_i^k(t_{m-1})$, $\chi_i^k(\widetilde{X}(t)) - \chi_i^k(\widetilde{X}(t_{m-1}))$, and $\chi_i^k(\widetilde{X}(t)) - \chi_i^k(X(t))$. These differences are estimated from the integral formulations of the differential equations satisfied by the corresponding processes, using Assumptions (A1)–(A4) and the stability results for X_i^k, $\widetilde{X}_i^k$, and Z_i^k. After some computations, we arrive at the differential inequality

$$\frac{\mathrm{d}u}{\mathrm{d}t} \leq - \min_{i=1,\dots,n} \left(r_i - 2\sum_{j=1}^{n} \max\{L_{ij}, L_{ji}\} \right) u + C'(\theta)\tau(u^{1/2} + \tau) + C''\tau \sum_{i=1}^{n} \Gamma_i,$$

where $u = \sum_{i=1}^{n} \|Z_i^k\|^2$ and the constants $C'(\theta) > 0$ and $C'' > 0$ do not depend on $(b_i, p_i)_{i=1,\dots,n}$, m, or T. In view of Assumption (A2), the first term on the right-hand side is nonpositive. The dependence of $C'(\theta)$ on θ arises from the terms involving $b_i / \min\{b_i, b_j\}$; see (5). It follows that $u(t)$ is bounded from above by $C(\theta)\tau + C\sqrt{\tau \sum_{i=1}^{n} \Gamma_i}$ for some other constants $C(\theta) > 0$ and $C > 0$.

1.4. **Link to related problems.** The random-batch scheme can be interpreted as a Monte–Carlo method to solve the mean-field equations associated to (1). In the mean-field limit $N \to \infty$, system (1) converges to

$$\mathrm{d}\overline{X}_i = -\nabla V_i(\overline{X}_i)\mathrm{d}t + \sum_{j=1}^{n}(K_{ij} * u_j)(\overline{X}_i)\mathrm{d}t + \sigma_i \mathrm{d}B_i, \quad i = 1, \dots, n,$$

where u_i is the probability density of $\overline{X}_i$ and solves the mean-field system

$$\partial_t u_i = \mathrm{div}(u_i \nabla V_i(x)) - \mathrm{div}\left(\sum_{j=1}^{n} u_i(K_{ij} * u_j)\right) + \frac{\sigma_i^2}{2}\Delta u_i \quad \text{in } \mathbb{R}^d, \ i = 1, \dots, n;$$

see, e.g., the review [13]. If $K_{ij} = \nabla k_{ij}$, it holds that $K_{ij} * u_j = k_{ij} * \nabla u_j$, and the density u_i solves a nonlocal cross-diffusion system. Moreover, if $k_{ij} = k_{ij}^\eta$ approximates the delta distribution δ according to $k_{ij}^\eta \to a_{ij}\delta$ in $\mathcal{D}'$ as $\eta \to 0$ for some numbers $a_{ij} \geq 0$, it was shown in [6] that the limit $N \to \infty$ and $\eta \to 0$ (in a certain sense) leads to the local cross-diffusion system

$$\partial_t u_i = \mathrm{div}(u_i \nabla V_i(x)) - \mathrm{div}\left(\sum_{j=1}^{n} a_{ij} u_i \nabla u_j\right) \quad \text{in } \mathbb{R}^d, \ i = 1, \dots, n.$$

The mean-field limit of the random-batch method was investigated in [14]. The authors showed that the (single-species) N-particle system is reduced to a p-particle system. This mean-field limit does not depend on the law of large numbers, and it is different from the standard mean-field limit, since the chaos is imposed at every time step, while in the standard limit, the chaos is propagated to later times.

The idea of choosing particles in a random way has been exploited in kinetic theory. For instance, subsampling was used in Monte–Carlo simulations [12] and for the symmetric Nabu algorithm, which relates to the random-batch method for $p_i = 2$ [1].

Random-batch methods can also be applied to second-order particle systems [18], many-particle Schrödinger equations [10], and kinetic equations [22]. They have been used to sample complicated or unknown probability distributions [21, 27], and they have been combined with model predictive control strategies to control the guiding problem for a herd of evaders [20]. In molecular dynamics, the interaction kernel is generally singular and given by, e.g., the Coulomb or Lennard–Jones potential. This situation is excluded in this paper because of Assumption (A1). However, one may split the kernel function into (singular) short-range and (smooth) long-range parts and apply the random-batch method only to the long-range part. This yields similar convergence results as above but

with constants depending on the end time [18]. We refer to the review [15] for further applications and references.

Theorem 1 provides the strong convergence with rate $O(\sqrt{\tau})$ of the error process. In [17], the weak convergence with rate $O(\tau)$ is proved for the single-species case. The proof makes use of the backward Kolmogorov equation and the contraction of the associated semigroup in $L^\infty(\mathbb{R}^d)$. In the multi-species situation, we obtain a system of equations for which contraction properties can be expected under Assumption (A2), but possibly in a weaker topology. A possible wayout is to use estimates in the space $H^s(\mathbb{R}^d) \subset L^\infty(\mathbb{R}^d)$ for $s > d/2$, derived for the mean-field limit [6]. We leave the details to future work.

Theorem 1 can be generalized to particle systems with multiplicative noise when the diffusion coefficients are Lipschitz continuous. We can only prove stability for particle systems with *interacting* diffusion coefficients like in [7], which lead in a mean-field-type limit to the Shigesada–Kawasaki–Teramoto population model. For details, we refer to Section 5.

The paper is organized as follows. The consistency of the scheme and stability of the stochastic processes X_i^k and $\widetilde{X}_i^k$ are proved in Section 2. Section 3 is concerned with the control of the error process $Z_i^k = \widetilde{X}_i^k - X_i^k$ and corresponding uniform estimates. Theorem 1 is proved in Section 4. We comment on the error estimate for particle systems with multiplicative noise in Section 5. Some numerical simulations, illustrating the convergence behavior and the influence of the batch sizes, are presented in Section 6. Finally, we collect some known results about the conditional expectation used in this paper in Appendix A.

2. Consistency and stability

We assume that Assumptions (A1)–(A4) hold. Let $i \in \{1, \ldots, n\}$, $k \in \{1, \ldots, N_i\}$ and let $\widetilde{X}_i^k$ with $k \in \mathcal{C}_{i,r}$ be a solution to (4). Then $\widetilde{X}_i^k$ solves

$$(8) \qquad \mathrm{d}\widetilde{X}_i^k = -\nabla V_i(\widetilde{X}_i^k)\mathrm{d}t + \sum_{j=1}^n \alpha_{ij} \sum_{\substack{\ell=1 \\ (i,k)\neq(j,\ell)}}^{N_j} K_{ij}(\widetilde{X}_i^k - \widetilde{X}_j^\ell)\mathrm{d}t + \sigma_i \mathrm{d}B_i^k + \chi_i^k(\widetilde{X})\mathrm{d}t,$$

where the remainder χ_i^k is defined for $x = (x_1^1, \ldots, x_n^{N_n}) \in \mathbb{R}^{dN_1 \times \cdots \times dN_n}$ by

$$(9) \qquad \chi_i^k(x) = \sum_{j=1}^n \beta_{ij} \sum_{\substack{\ell \in \mathcal{C}_{j,r} \\ (i,k)\neq(j,\ell)}} K_{ij}(x_i^k - x_j^\ell) - \sum_{j=1}^n \alpha_{ij} \sum_{\substack{\ell=1 \\ (i,k)\neq(j,\ell)}}^{N_j} K_{ij}(x_i^k - x_j^\ell)$$
$$=: f_i^k(x) - g_i^k(x).$$

The following proposition shows that the scheme is consistent.

Proposition 2 (Consistency)**.** *Let $p_i \geq 2$ for $i = 1, \ldots, n$ and $x = (x_1^1, \ldots, x_n^{N_n}) \in \mathbb{R}^{dN_1 \times \cdots \times dN_n}$. Then the expectation and variance of χ_i^k, defined in (9), are $\mathbb{E}(\chi_i^k(x)) = 0$*

and

$$\text{Var}(\chi_i^k(x)) = \sum_{\substack{j,j'=1 \\ j,j'\neq i, j\neq j'}}^{n} \left(\frac{\max\{b_i, b_j, b_{j'}\}}{\max\{b_i, b_{j'}\}} - 1\right) A_i^{jj'}(x)$$

$$(10) \quad + \sum_{\substack{j=1 \\ j\neq i}}^{n} \left(\frac{b_i - \min\{b_i, b_j\}}{\min\{b_i, b_j\}} - \frac{1}{N_j} + \frac{b_i}{p_j \min\{b_i, b_j\}}\right) A_i^j(x) + \frac{\max\{b_i, b_j\} - 1}{N_j} A_{i,1}^j(x)$$

$$+ \left(\frac{1}{p_i - 1} - \frac{1}{N_i - 1}\right) A_i(x),$$

where $i = 1, \ldots, n$, $k = 1, \ldots, N_i$, and

$$A_i^{jj'}(x) = \frac{1}{N_j N_{j'}} \sum_{\ell=1}^{N_j} \sum_{\ell'=1}^{N_{j'}} K_{ij}(x_i^k - x_j^\ell) K_{ij'}(x_i^k - x_{j'}^{\ell'}),$$

$$A_i^j(x) = \frac{1}{N_j(N_j - 1)} \sum_{\ell,\ell'=1, \ell\neq\ell'}^{N_j} K_{ij}(x_i^k - x_j^\ell) K_{ij}(x_i^k - x_j^{\ell'}),$$

$$A_{i,1}^j(x) = \frac{1}{N_j} \sum_{\ell=1}^{N_j} K_{ij}(x_i^k - x_j^\ell)^2,$$

$$A_i(x) = \frac{1}{N_i - 2} \sum_{\ell=1, \ell\neq k}^{N_i} \left(K_{ii}(x_i^k - x_i^\ell) - \frac{1}{N_i - 1} \sum_{\ell'=1, \ell'\neq k}^{N_i} K_{ii}(x_i^k - x_i^{\ell'})\right)^2.$$

Using definition (7), we can estimate the variance of $\chi_i^k(x)$ from above according to

$$\text{Var}(\chi_i^k(x)) \leq 8 \max_{i,j=1,\ldots,n} \|K_{ij}\|_\infty^2 \sum_{k=1}^{n} \Gamma_k.$$

As expected, for larger batch sizes p_i, the variance is smaller and the noise level is lower. In the single-species case, we recover [16, Lemma 3.1] since

$$\text{Var}(\chi_i^k(x)) = \left(\frac{1}{p_i - 1} - \frac{1}{N_i - 1}\right) A_i(x).$$

If the species numbers and batch sizes are the same, i.e. $N_i = N$ and $b_i = b$ for all $i = 1, \ldots, n$, it follows that

$$\text{Var}(\chi_i^k(x)) = \left(\frac{1}{p} - \frac{1}{N}\right) \sum_{j=1, j\neq i}^{n} \left(A_i^j(x) + A_{i,1}^j(x)\right) + \left(\frac{1}{p-1} - \frac{1}{N-1}\right) A_i(x).$$

We observe that the first term on the right-hand side of (10) vanishes. This means that, in case of different species numbers or batch sizes, the noise level is larger than in the uniform case.

Proof. The proof is similar to [16, Lemma 3.1], but since we have multiple species, the computations are more involved. Let $i \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, N_i\}$ be arbitrary but fixed. We write $I_i^k(j, \ell) = 1$ if (i, k) and (j, ℓ) are in the same batch, i.e., if there exists $r \geq 1$ such that $(i, k), (j, \ell) \in \mathcal{C}_r$. Otherwise, we set $I_i^k(j, \ell) = 0$. With this notation, we can write $f_i^k = f_i^k(x)$, defined in (9), as

$$f_i^k = \sum_{j=1}^n \beta_{ij} \sum_{\ell=1}^{N_j} K_{ij}(x_i^k - x_j^\ell) I_i^k(j, \ell).$$

Step 1: Computation of the expection. We claim that

(11) $$\mathbb{E} I_i^k(j, \ell) = \begin{cases} \frac{p_i - 1}{N_i - 1} & \text{if } i = j, \\ \frac{\min\{b_i, b_j\}}{b_i b_j} & \text{if } i \neq j. \end{cases}$$

The case $i = j$ is proved in [16, Lemma 3.1]. For $i \neq j$, we define $a(i, k)$ as the index of the super-batch $\mathcal{C}_r$ that contains (i, k), i.e. $a(i, k) = r$ if and only if $(i, k) \in \mathcal{C}_r$ or, equivalently, $k \in \mathcal{C}_{i,r}$. We have

$$\mathbb{P}(I_i^k(j, \ell) = 1) = \mathbb{P}((j, \ell) \in \mathcal{C}_{a(i,k)}) = \sum_{r=1}^{\min\{b_i, b_j\}} \mathbb{P}((j, \ell) \in \mathcal{C}_r | a(i, k) = r) \mathbb{P}(a(i, k) = r)$$

$$= \sum_{r=1}^{\min\{b_i, b_j\}} \mathbb{P}((j, \ell) \in \mathcal{C}_r) \mathbb{P}(a(i, k) = r).$$

The distribution of a particle of a certain species is uniform with respect to the species' batch in which it ends up, i.e. $\mathbb{P}(\ell \in \mathcal{C}_{j,r}) = \mathbb{P}(\ell \in \mathcal{C}_{j,s})$ for all $r, s = 1, \ldots, b_j$. Consequently, we have $\mathbb{P}(\ell \in \mathcal{C}_{j,r}) = 1/b_j$ for all $r = 1, \ldots, b_j$ and $\mathbb{P}(\ell \in \mathcal{C}_{j,r}) = 0$ otherwise, since $\mathcal{C}_{j,r} = \emptyset$ if $r > b_j$. This leads for $i \neq j$ to

$$\mathbb{E} I_i^k(j, \ell) = 1 \cdot \mathbb{P}(I_i^k(j, \ell) = 1) = \min\{b_i, b_j\} \frac{1}{b_j} \frac{1}{b_i}.$$

We infer from the definitions of α_{ij} and β_{ij} in (3) and (5), respectively, and from (11) that

$$\mathbb{E}(f_i^k) = \sum_{j=1}^n \beta_{ij} \sum_{\ell=1}^{N_j} K_{ij}(x_i^k - x_j^\ell) \mathbb{E} I_i^k(j, \ell)$$

$$= \frac{1}{N_i - 1} \sum_{\ell=1, \ell \neq k}^{N_i} K_{ii}(x_i^k - x_i^\ell) + \sum_{j=1, j \neq i}^n \frac{1}{N_j} \sum_{\ell=1}^{N_j} K_{ij}(x_i^k - x_j^\ell)$$

$$= \sum_{j=1}^n \alpha_{ij} \sum_{\substack{\ell=1 \\ (i,k) \neq (j,\ell)}} K_{ij}(x_i^k - x_j^\ell) = \mathbb{E}(g_i^k).$$

This shows that $\mathbb{E}(\chi_i^k(x)) = \mathbb{E}(f_i^k) - \mathbb{E}(g_i^k) = 0$.

Step 2: Preparation for the computation of the variance. We introduce the notation $G_j^\ell := K_{ij}(x_i^k - x_j^\ell)I_i^k(j,\ell)$ if $(i,k) \neq (j,\ell)$ and $G_j^\ell = 0$ if $(i,k) = (j,\ell)$. Then

$$\mathbb{E}(f_i^k)^2 = \sum_{j,j'=1}^n \beta_{ij}\beta_{ij'} \sum_{\ell=1}^{N_j} \sum_{\ell'=1}^{N_{j'}} \mathbb{E}(G_j^\ell G_{j'}^{\ell'}).$$

The expectation of $G_j^\ell G_{j'}^{\ell'}$ can be written as

$$\mathbb{E}(G_j^\ell G_{j'}^{\ell'}) = K_{ij}(x_i^k - x_j^\ell)K_{ij'}(x_i^k - x_{j'}^{\ell'})\mathbb{E}(I_i^k(j,\ell)I_i^k(j',\ell'))$$
$$= K_{ij}(x_i^k - x_j^\ell)K_{ij'}(x_i^k - x_{j'}^{\ell'})\mathbb{P}(I_i^k(j,\ell)I_i^k(j',\ell') = 1).$$

Thus, we need to calculate $\mathbb{P}(I_i^k(j,\ell)I_i^k(j',\ell') = 1)$. For this, we distinguish several cases.

Case 1: $j, j' \neq i$ and $j \neq j'$. We compute, using the definition of the super-batches,

$$\{I_i^k(j,\ell)I_i^k(j',\ell') = 1\} = \big\{(j,\ell) \in \mathcal{C}_{a(i,k)}, \ (j',\ell') \in \mathcal{C}_{a(i,k)}\big\}$$
$$= \big\{\ell \in \mathcal{C}_{j,a(i,k)}, \ \ell' \in \mathcal{C}_{j',a(i,k)}\big\} = \bigcup_{r \in \mathbb{N}} \big\{\ell \in \mathcal{C}_{j,a(i,k)}, \ \ell' \in \mathcal{C}_{j',a(i,k)}, \ r = a(i,k)\big\}.$$

The random division $\xi_{m.1}$ of the batch $\mathcal{C}_{i,r}$ at time t_{m-1} is independent of the random devision of the batches $\mathcal{C}_{j,r}$ and $\mathcal{C}_{j',r}$. Thus, we can write

$$\mathbb{P}(I_i^k(j,\ell)I_i^k(j',\ell') = 1) = \sum_{r=1}^{\min\{b_i,b_j,b_{j'}\}} \mathbb{P}(k \in \mathcal{C}_{i,r})\mathbb{P}(\ell \in \mathcal{C}_{j,r})\mathbb{P}(\ell' \in \mathcal{C}_{j',r})$$
$$= \min\{b_i, b_j, b_{j'}\}\frac{1}{b_i}\frac{1}{b_j}\frac{1}{b_{j'}}.$$

Case 2: $j, j' \neq i$, $j = j'$ and $\ell \neq \ell'$. In this case, both ℓ and ℓ' are in the same batch such that

(12) $$\{I_i^k(j,\ell)I_i^k(j,\ell') = 1\} = \bigcup_{r \in \mathbb{N}} \big\{\ell, \ell' \in \mathcal{C}_{j,r}, \ r = a(i,k)\big\}$$
$$= \bigcup_{r \in \mathbb{N}} \{I_j^k(j,\ell') = 1\} \cap \{\ell \in \mathcal{C}_{j,r}\} \cap \{k \in \mathcal{C}_{i,r}\}.$$

Because of the uniformity of the random devision (as in Case 1), we have

$$\mathbb{P}(I_j^\ell(j,\ell') = 1, \ \ell \in \mathcal{C}_{j,r}) = \mathbb{P}(I_j^\ell(j,\ell') = 1, \ \ell \in \mathcal{C}_{j,s}) \quad \text{for all } 1 \leq r, s \leq b_j.$$

Since, by (11), $\mathbb{P}(I_j^\ell(j,\ell') = 1) = (p_j - 1)/(N_j - 1)$, we deduce from (12) that

$$\mathbb{P}(I_i^k(j,\ell)I_i^k(j',\ell') = 1) = \sum_{r=1}^{\min\{b_i,b_j\}} \mathbb{P}(I_j^\ell(j,\ell') = 1, \ \ell \in \mathcal{C}_{j,r})\mathbb{P}(k \in \mathcal{C}_{i,r})$$
$$= \frac{1}{b_ib_j} \sum_{r=1}^{\min\{b_i,b_j\}} \mathbb{P}(I_j^\ell(j,\ell') = 1) = \frac{\min\{b_i, b_j\}(p_j - 1)}{b_ib_j(N_j - 1)}.$$

Case 3: $j \neq i$, $j' = i$. If $\ell = k$, it follows from the definition of G_j^ℓ that $G_i^{\ell'} = G_i^k = 0$. If $\ell \neq k$, the definition of G_j^ℓ gives

$$\mathbb{E}(G_j^\ell G_{j'}^{\ell'}) = K_{ij}(x_i^k - x_j^\ell)K_{ii}(x_i^k - x_i^{\ell'})\mathbb{E}(I_i^k(j,\ell)I_i^k(i,\ell')),$$

and it remains the compute the expectation on the right-hand side. Proceeding as in the previous cases, we find that

$$\mathbb{P}(I_i^k(j,\ell)I_i^k(i,\ell') = 1) = \sum_{r=1}^{\min\{b_i,b_j\}} \mathbb{P}(I_i^k(i,\ell') = 1, k \in \mathcal{C}_{i,r})\mathbb{P}(\ell \in \mathcal{C}_{j,r})$$

$$= \frac{1}{b_i b_j} \sum_{r=1}^{\min\{b_i,b_j\}} \mathbb{P}(I_i^k(i,\ell') = 1) = \frac{\min\{b_i,b_j\}(p_i-1)}{b_i b_j(N_i-1)}.$$

Case 4: $j, j' = i$, $\ell \neq \ell' \neq k$. We need to compute the probability of $I_i^k(i,\ell)I_i^k(i,\ell') = 1$. This case happens exactly when the indices ℓ, ℓ', and k are in the same batch $\mathcal{C}_{i,a(i,k)}$. Similar arguments as for $\mathbb{P}(I_i^k(i,\ell) = 1)$ in the proof of Lemma 3.1 in [16] yield

$$\mathbb{P}(I_i^k(i,\ell)I_i^k(i,\ell') = 1) = \frac{(p_i-1)(p_i-2)}{(N_i-1)(N_i-2)}.$$

Case 5: $j, j' = i$, $\ell = \ell'$, $\ell \neq k$. We only need $\mathbb{E}(I_i^k(i,\ell))$, which we already computed:

$$\mathbb{P}(I_i^k(i,\ell) = 1) = \frac{p_i-1}{N_i-1}.$$

Summarizing these five cases, we obtain $\mathbb{E}(f_i^k)^2 = J_1 + \cdots + J_5$, where the term J_j corresponds to case j and

$$J_1 = \sum_{\substack{j,j'=1 \\ j,j' \neq i, j \neq j'}}^{n} \frac{\min\{b_i,b_j,b_{j'}\}b_i}{N_j N_{j'} \min\{b_i,b_j\} \min\{b_i,b_{j'}\}} \sum_{\ell,\ell'=1}^{N_j,N_{j'}} K_{ij}(x_i^k - x_j^\ell)K_{ij'}(x_i^k - x_{j'}^{\ell'}),$$

$$J_2 = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{(p_j-1)b_i}{(N_j-1)N_j \min\{b_i,b_j\}p_j} \sum_{\ell,\ell'=1, \ell \neq \ell'}^{N_j} K_{ij}(x_i^k - x_j^\ell)K_{ij}(x_i^k - x_j^{\ell'}),$$

$$J_2' = \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{b_i}{N_j \min\{b_i,b_j\}p_j} \sum_{\ell=1}^{N_j} K_{ij}(x_i^k - x_j^\ell)^2,$$

$$J_3 = 2\sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{(N_i-1)N_j} \sum_{\substack{\ell,\ell'=1 \\ \ell' \neq k}}^{N_j,N_i} K_{ij}(x_i^k - x_j^\ell)K_{ii}(x_i^k - x_i^{\ell'}),$$

$$J_4 = \frac{p_i - 2}{(p_i - 1)(N_i - 1)(N_i - 2)} \sum_{\substack{\ell,\ell'=1 \\ \ell \neq \ell'}}^{N_i} K_{ii}(x_i^k - x_i^\ell) K_{ii}(x_i^k - x_i^{\ell'}),$$

$$J_5 = \frac{1}{(p_i - 1)(N_i - 1)} \sum_{\ell=1}^{N_i} K_{ii}(x_i^k - x_i^\ell)^2.$$

For the term $(\mathbb{E}(f_i^k))^2 = (\mathbb{E}(g_i^k))^2$, we expand the square:

$$(\mathbb{E}(f_i^k))^2 = \left(\sum_{j=1}^n \alpha_{ij} \sum_{\substack{\ell=1 \\ (i,k) \neq (j,\ell)}}^n K_{ij}(x_i^k - x_j^\ell) \sum_{j'=1}^n \alpha_{ij'} \sum_{\substack{\ell'=1 \\ (i,k) \neq (j',\ell')}}^n K_{ij'}(x_i^k - x_{j'}^{\ell'}) \right)^2$$

$$= \widehat{J}_1 + \cdots + \widehat{J}_5, \quad \text{where}$$

$$\widehat{J}_1 = \sum_{\substack{j,j'=1 \\ j,j' \neq i,\, j \neq j'}}^n \frac{1}{N_j N_{j'}} \sum_{\ell,\ell'=1}^{N_j,N_{j'}} K_{ij}(x_i^k - x_j^\ell) K_{ij'}(x_i^k - x_{j'}^{\ell'}),$$

$$\widehat{J}_2 = \sum_{j=1,\, j\neq i}^n \frac{1}{N_j^2} \sum_{\ell,\ell'=1}^{N_j} K_{ij}(x_i^k - x_j^\ell) K_{ij}(x_i^k - x_j^{\ell'}),$$

$$\widehat{J}_3 = 2 \sum_{j=1,\, j\neq i}^n \frac{1}{(N_i - 1)N_j} \sum_{\ell=1}^{N_j} \sum_{\substack{\ell'=1 \\ \ell' \neq k}}^{N_i} K_{ij}(x_i^k - x_j^\ell) K_{ii}(x_i^k - x_i^{\ell'}),$$

$$\widehat{J}_4 = \frac{1}{(N_i - 1)^2)} \sum_{\substack{\ell,\ell'=1 \\ \ell \neq \ell'}}^{N_i} K_{i,i}(x_i^k - x_i^\ell) K_{ii}(x_i^k - x_i^{\ell'}),$$

$$\widehat{J}_5 = \frac{1}{(N_i - 1)^2} \sum_{\ell=1}^{N_i} K_{ii}(x_i^k - x_i^\ell)^2.$$

The variance of f_i^k is the difference $(J_1 + \cdots + J_5) - (\widehat{J}_1 + \cdots + \widehat{J}_5)$. We observe that $J_3 - \widehat{J}_3 = 0$ and that

$$\frac{\min\{b_i, b_j, b_{j'}\} b_i}{\min\{b_i, b_j\} \min\{b_i, b_{j'}\}} = \frac{\max\{b_i, b_j, b_{j'}\}}{\max\{b_j, b_{j'}\}},$$

A tedious but straightforward computation yields for the other terms:

$$\text{Var}(f_i^k) = \mathbb{E}(f_i^k)^2 - (\mathbb{E}f_i^k)^2 = (J_1 - \widehat{J}_1) + (J_2 + J_2' - \widehat{J}_2) + (J_4 + J_5 - \widehat{J}_4 - \widehat{J}_5)$$

$$= \sum_{\substack{j,j'=1 \\ j,j' \neq i,\, j \neq j'}}^n \left(\frac{\max\{b_i, b_j, b_{j'}\}}{\max\{b_j, b_{j'}\}} - 1 \right) \frac{1}{N_j N_{j'}} \sum_{\ell=1}^{N_j} \sum_{\ell'=1}^{N_{j'}} K_{ij}(x_i^k - x_j^\ell) K_{ij'}(x_i^k - x_{j'}^{\ell'})$$

$$+ \sum_{j=1, j\neq i}^{n} \left(\frac{b_i - \min\{b_i, b_j\}}{\min\{b_i, b_j\}} - \frac{1}{N_j} + \frac{b_i}{\min\{b_i, b_j\}p_j} \right) \frac{1}{N_j(N_j - 1)}$$

$$\times \sum_{\ell,\ell'=1, \ell\neq\ell'}^{N_j} K_{ij}(x_i^k - x_j^\ell)K_{ij}(x_i^k - x_j^{\ell'})$$

$$+ \sum_{j=1, j\neq i}^{n} \left(\frac{\max\{b_i, b_j\} - 1}{N_j} \right) \frac{1}{N_j} \sum_{\ell=1}^{N_j} K_{ij}(x_i^k - x_j^\ell)^2$$

$$+ \left(\frac{1}{p_i - 1} - \frac{1}{N_i - 1} \right) \frac{1}{N_i - 2} \sum_{\substack{\ell=1 \\ k\neq\ell}}^{N_i} \left(K_{ii}(x_i^k - x_j^\ell) - \frac{1}{N_i - 1} \sum_{\substack{\ell'=1 \\ \ell'\neq k}}^{N_i} K_{ii}(x_i^k - x_j^{\ell'}) \right)^2 .$$

The right-hand side equals (10), which finishes the proof. $\qquad\square$

For later use, we prove the following auxiliary result, which generalizes Lemma 3.2 in [16] to the multi-species case.

Lemma 3. *Let $i \in \{1, \ldots, n\}$, $k \in \{1, \ldots, N_i\}$, and $(i, k) \in \mathcal{C}_{i,r}$ for some $r = a(i, k) \leq b_i$. Let $S_j^\ell \in \mathbb{R}^d$ with $j, \ell \in \mathbb{N}$ be random variables which are independent of the partitioning random variable ξ_m. Then it holds*

$$\left\| \frac{1}{p_j} \sum_{\ell\in\mathcal{C}_{j,r}} S_j^\ell \right\| = \max_{\ell=1,\ldots,N_j} \|S_j^\ell\| \quad \text{if } i \neq j,$$

$$\left\| \frac{1}{p_j - 1} \sum_{\ell\in\mathcal{C}_{i,r}, \ell\neq k} S_i^\ell \right\| = \max_{\ell=1,\ldots,N_i} \|S_i^\ell\| \quad \text{if } i = j,$$

recalling that $\|\cdot\| = (\mathbb{E}(\cdot)^2)^{1/2}$.

Proof. The proof is similar to that one of [16, Lemma 3.2]. We present it for completeness. Let $i \neq j$ and set $I_i^k(j, \ell) = 1$ if (i, k) and (j, ℓ) are in same batch and $I_i^k(j, \ell) = 0$ otherwise. Due to the independency of S_j^ℓ and ξ_m, we have

$$\left\| \frac{1}{p_j} \sum_{\ell\in\mathcal{C}_{j,r}} S_j^\ell \right\|^2 = \frac{1}{p_j^2}\mathbb{E}\left(\sum_{\ell=1}^{N_j} I_i^k(j, \ell)S_j^\ell \right)^2 = \frac{1}{p_j^2} \sum_{\ell,\ell'=1}^{N_j} \mathbb{E}\left(I_i^k(j, \ell)I_i^k(j, \ell')S_j^\ell S_j^{\ell'} \right)$$

$$= \frac{1}{p_j^2} \sum_{\ell,\ell'=1}^{N_j} \mathbb{E}(I_i^k(j, \ell)I_i^k(j, \ell'))\mathbb{E}(S_j^\ell S_j^{\ell'}).$$

We know from (11) that $\mathbb{E}(I_i^k(j, \ell)I_i^k(j, \ell')) \leq 1/b_j$ in the case of $\ell = \ell'$ and from Case 2 of Proposition 2 that $\mathbb{E}(I_i^k(j, \ell)I_i^k(j, \ell')) \leq (p_j - 1)/(b_j(N_j - 1))$, if $\ell \neq \ell'$. Therefore, using

the Cauchy–Schwarz inequality and the fact that $N_j = b_j p_j$,

$$\left\| \frac{1}{p_j} \sum_{\ell \in \mathcal{C}_{j,r}} S_j^\ell \right\|^2 \leq \frac{1}{p_j^2} \left(\sum_{\ell, \ell'=1, \ell \neq \ell'}^{N_j} \frac{p_j - 1}{b_j(N_j - 1)} \|S_j^\ell\| \, \|S_j^{\ell'}\| + \sum_{\ell=1}^{N_j} \frac{1}{b_j} \|S_j^\ell\|^2 \right)$$

$$\leq \max_{\ell=1,\ldots,N_j} \|S_j^\ell\|^2 \left(\frac{(p_j - 1)(N_j - 1)N_j}{p_j N_j (N_j - 1)} + \frac{1}{p_j} \right) \leq \max_{\ell=1,\ldots,N_j} \|S_j^\ell\|^2.$$

The case $i = j$ is shown in a similar way. $\square$

The next result is concerned with the stability of X_i^k and $\widetilde{X}_i^k$.

Lemma 4 (Stability). *Let* $q \geq 2$, *and* $X_{0,i}^k \in L^q(\Omega)$, *where* $i \in \{1, \ldots, n\}$ *and* $k \in \{1, \ldots, N_i\}$. *Then there exist constants* $C(q)$, $C_1 > 0$, *independent of* $(p_i, b_i)_{i=1,\ldots,n}$, m, *and* T, *such that*

(13) $$\sup_{t>0} \mathbb{E}|X_i^k(t)|^q \leq C(q), \quad \sup_{t>0} \mathbb{E}|\widetilde{X}_i^k(t)|^q \leq C(q)(1 + \theta^q),$$

(14) $$\sup_{t_{m-1} < t < t_m} \mathbb{E}\big(|\widetilde{X}_i^k(t)|^q \big| \mathcal{F}_{m-1}\big) \leq |\widetilde{X}_i^k(t_{m-1})|^q + C(q)(1 + \theta^q),$$

where θ *is defined in* (6). *Furthermore, it holds that*

(15) $$\left| \mathbb{E}\big(\widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1}) \big| \mathcal{F}_{m-1}\big) \right| \leq C_V \tau |\widetilde{X}_i^k(t_{m-1})|^{\widetilde{q}_i} + C_1 \tau (1 + \theta^{\widetilde{q}_i}),$$

where $\widetilde{q}_i = \max\{2, q_i\}$ *and* $C_V > 0$ *is introduced in Assumption (A2).*

Proof. Let $i \in \{1, \ldots, n\}$ and $k \in \{1, \ldots, N_i\}$ be arbitrary but fixed. The proof is similar to [16, Lemma 3.3] with the exception that we work out the dependence on the number of batches b_i in terms of the quotient θ.

Step 1: Stability for $X_i^k(t)$. Let $d \geq 2$. We use Itô's calculus for the process $|X_i^k|^q$ and apply the expectation as in [16, Lemma 3.3], which yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathbb{E}|X_i^k(t)|^q = -q\mathbb{E}\big(|X_i^k(t)|^{q-2} X_i^k(t) \cdot \nabla V_i(X_i^k(t))\big)$$

$$+ \frac{q}{N_i - 1} \mathbb{E}\bigg(|X_i^k(t)|^{q-2} X_i^k(t) \cdot \sum_{\ell=1, \ell \neq k}^{N_i} K_{ii}(X_i^k(s) - X_i^\ell(s)) \bigg)$$

$$+ \sum_{j=1, j \neq i}^{n} \frac{q}{N_j} \mathbb{E}\bigg(|X_i^k(t)|^{q-2} X_i^k(t) \cdot \sum_{\ell=1}^{N_j} K_{ij}(X_i^k(s) - X_j^\ell(s)) \bigg)$$

$$+ \frac{\sigma_i^2}{2} q(q + d - 2)\mathbb{E}|X_i^k(t)|^{q-2}.$$

The mean-value theorem with intermediate value $\zeta \in \mathbb{R}^d$ and the convexity of $x \mapsto V_i(x) - r_i|x|^2/2$ (Assumption (A2)) imply that for all $x \in \mathbb{R}^d$,

$$x \cdot \nabla V_i(x) = x^T D^2 V_i(\zeta) x + x \cdot \nabla V_i(0) \geq r_i |x|^2 + x \cdot \nabla V_i(0).$$

Together with Fubini's theorem, the boundedness of the kernels K_{ij} (Assumption (A1)), and Young's inequality, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}|X_i^k(t)|^q \le -qr_i\mathbb{E}|X_i^k(t)|^q + q\bigg(|\nabla V_i(0)|\sum_{j=1}^n \|K_{ij}\|_\infty\bigg)\mathbb{E}|X_i^k(t)|^{q-1}$$

$$+ \frac{\sigma_i^2}{2}q(q+d-2)\mathbb{E}|X_i^k(t)|^{q-2} \le -\frac{qr_i}{2}\mathbb{E}|X_i^k(t)|^q + C_2,$$

where $C_2 > 0$ depends on ∇V_i, K_{ij}, σ_i, d, and q. Gronwall's lemma implies that $\mathbb{E}|X_i^k(t)|^q$ is bounded by a constant depending on q (and not depending on T).

Step 2: Stability for $\widetilde{X}_i^k$. Let $t \in (t_{m-1}, t_m]$ and let $(i,k) \in \mathcal{C}_r$ for some $r \in \mathbb{N}$. Similarly as in the previous step, we use Itô's calculus and apply the conditional expectation with respect to $\mathcal{F}_{m-1}$, observing that $|\widetilde{X}_i^k(t_{m-1})|^q$ is $\mathcal{F}_{m-1}$-measurable. Then, applying Lemmas 9 and 10 in the appendix,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big(|\widetilde{X}_i^k(t)|^q\big|\mathcal{F}_{m-1}\big) = -q\mathbb{E}\big(|\widetilde{X}_i^k(t)|^{q-2}\widetilde{X}_i^k \cdot \nabla V_i(\widetilde{X}_i^k(t))\big|\mathcal{F}_{m-1}\big)$$

$$+ \frac{q}{p_i-1}\sum_{\ell\in\mathcal{C}_{i,r},\,\ell\neq k}\mathbb{E}\big(|\widetilde{X}_i^k(t)|^{q-2}\widetilde{X}_i^k(t)\cdot K_{ii}(\widetilde{X}_i^k - \widetilde{X}_i^\ell)\big|\mathcal{F}_{m-1}\big)$$

$$+ \sum_{j=1,\,j\neq i}^n \frac{qb_i}{p_j\min\{b_i,b_j\}}\sum_{\ell\in\mathcal{C}_{j,r}}\mathbb{E}\big(|\widetilde{X}_i^k(t)|^{q-2}\widetilde{X}_i^k(t)\cdot K_{ij}(\widetilde{X}_i^k(t) - \widetilde{X}_j^\ell(t))\big|\mathcal{F}_{m-1}\big)$$

$$+ \frac{\sigma_i^2}{2}q(q+d-2)\mathbb{E}\big(|\widetilde{X}_i^k(t)|^{q-2}\big|\mathcal{F}_{m-1}\big).$$

Proceeding as in the previous step and using $b_i/\min\{b_i,b_j\} \le \theta$, we infer that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big(|\widetilde{X}_i^k(t)|^q\big|\mathcal{F}_{m-1}\big) \le -\frac{qr_i}{2}\mathbb{E}\big(|\widetilde{X}_i^k(t)|^q\big|\mathcal{F}_{m-1}\big) + C_3(1+\theta)^q,$$

and Gronwall's lemma on $(t_{m-1}, t_m]$ implies (14). Finally, the second estimate in (13) is proved in a similar way, using the Gronwall lemma on $[0, t]$ and taking into account that $\mathbb{E}|X_{0,i}^k|^q$ is bounded by assumption.

Step 3: Proof of estimate (15). We apply Itô's lemma, take the conditional expectation of $\widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1})$, and use the polynomial growth condition for ∇V_i in Assumption (A2) as well as the boundedness of K_{ij}:

$$\mathbb{E}\big(\widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1})\big|\mathcal{F}_{m-1}\big) = -\int_{t_{m-1}}^t \mathbb{E}\big(\nabla V_i(\widetilde{X}_i^k(s))\big|\mathcal{F}_{m-1}\big)\mathrm{d}s$$

$$+ \frac{1}{p_i-1}\int_{t_{m-1}}^t \mathbb{E}\bigg(\sum_{\ell\in\mathcal{C}_{i,r},\,\ell\neq k}K_{ii}(\widetilde{X}_i^k(s) - \widetilde{X}_i^\ell(s))\big|\mathcal{F}_{m-1}\bigg)\mathrm{d}s$$

$$+ \sum_{j=1,\,j\neq i}^n \frac{b_i}{p_j\min\{b_i,b_j\}}\int_{t_{m-1}}^t \mathbb{E}\bigg(\sum_{\ell\in\mathcal{C}_{j,r}}K_{ij}(\widetilde{X}_i^k(s) - \widetilde{X}_j^\ell(s))\big|\mathcal{F}_{m-1}\bigg)\mathrm{d}s$$

$$\leq C_V \tau + C_V \int_{t_{m-1}}^{t} \mathbb{E}\big(|\widetilde{X}_i^k(s)|^{q_i}\big|\mathcal{F}_{m-1}\big)\mathrm{d}s + \tau \sum_{j=1}^{n} \frac{\|K_{ij}\|_{\infty} b_i}{\min\{b_i, b_j\}}.$$

It follows from (14) with $q = \widetilde{q}_i := \max\{2, q_i\}$ that

$$\mathbb{E}\big(\widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1})\big|\mathcal{F}_{m-1}\big)$$

$$\leq C_V \tau |\widetilde{X}_i^k(t_{m-1})|^{\widetilde{q}_i} + \tau\bigg(C_V + C_V C(\widetilde{q}_i)(1 + \theta^{\widetilde{q}_i}) + \sum_{j=1}^{n} \frac{\|K_{ij}\|_{\infty} b_i}{\min\{b_i, b_j\}}\bigg)$$

$$\leq C_V \tau |\widetilde{X}_i^k(t_{m-1})|^{\widetilde{q}_i} + C_4(\widetilde{q}_i)(1 + \theta^{\widetilde{q}_i}).$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3. Control of the error process

We prove first a bound for the difference $\widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1})$.

Lemma 5. *Let $t \in (t_{m-1}, t_m]$, let $\widetilde{X}$ be the stochastic process defined in (4), and let $i \in \{1, \dots, n\}$. Set $q_i' = 2\max\{1, q_i\}$, where q_i is defined in Assumption (A2). Then, for any $(i, k) \in \mathcal{C}_r$ for some $r \leq b_i$ such that $X_{0,i}^k \in L^{q_i'}(\Omega)$, there exists a constant $C > 0$, independent of $(p_i, b_i)_{i=1,\dots,n}$, and ξ_m, such that*

$$\mathbb{E}\big(|\widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1})|^2\big|\mathcal{F}_{m-1}\big) \leq C\tau(1 + \theta^{q_i'/2+1})\big(1 + |\widetilde{X}_i^k(t_{m-1})|^{q_i'/2+1}\big).$$

Proof. Again, the proof is similar to [16, Lemma 3.3] and based on Itô's calculus. Let $t \in (t_{m-1}, t_m]$ and $(i, k) \in \mathcal{C}_r$ for some $r \leq b_i$, satisfying the assumptions of the lemma. Set $S(t) := \widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1})$. We apply Itô's lemma to $|S(t)|^2$ and the conditional expectation and use Lemmas 9 and 10:

$$(16) \quad \mathbb{E}(|S(t)|^2|\mathcal{F}_{m-1}) \leq 2\int_{t_{m-1}}^{t} \big|\mathbb{E}\big(S(s) \cdot \nabla V_i(\widetilde{X}_i^k(s))\big|\mathcal{F}_{m-1}\big)\big|\mathrm{d}s + d\int_{t_{m-1}}^{t} \sigma_i^2 \mathrm{d}s$$

$$+ \frac{2}{p_i - 1} \int_{t_{m-1}}^{t} \bigg|\mathbb{E}\bigg(\sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} K_{ii}(\widetilde{X}_i^k(s) - \widetilde{X}_i^\ell(t)) \cdot S(s)\bigg|\mathcal{F}_{m-1}\bigg)\bigg|\mathrm{d}s$$

$$+ \sum_{j=1, j \neq i}^{} \frac{2b_i}{p_j \min\{b_i, b_j\}} \int_{t_{m-1}}^{t} \bigg|\mathbb{E}\bigg(\sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} K_{ij}(\widetilde{X}_i^k(s) - \widetilde{X}_j^\ell(t)) \cdot S(s)\bigg|\mathcal{F}_{m-1}\bigg)\bigg|\mathrm{d}s$$

$$=: J_6 + \cdots + J_9.$$

By the Cauchy–Schwarz inequality, the polynomial growth condition on ∇V_i (Assumption (A2)), and stability estimate (13) with $q = q_i'$, we have

$$J_6 \leq 2C_V^{1/2} \int_{t_{m-1}}^{t} \big(\mathbb{E}(|S(s)|^2|\mathcal{F}_{m-1})\big)^{1/2}\big(\mathbb{E}(1 + |\widetilde{X}_i^k(s)|^{2q_i}|\mathcal{F}_{m-1})\big)^{1/2}\mathrm{d}s$$

$$\leq 2C_V^{1/2}\big(1 + C(q)(1 + \theta^q) + |\widetilde{X}_i^k(t_{m-1})|^q\big)^{1/2} \int_{t_{m-1}}^{t} \big(\mathbb{E}(|S(s)|^2|\mathcal{F}_{m-1})\big)^{1/2}\mathrm{d}s.$$

Next, using the boundedness of K_{ii}, Lemma 8, and Hölder's inequality,

$$J_8 \leq C\|K_{ii}\|_\infty \int_{t_{m-1}}^t \big(\mathbb{E}(|S(s)|^2|\mathcal{F}_{m-1})\big)^{1/2}\mathrm{d}s,$$

$$J_9 \leq C\sum_{j=1,\,j\neq i}^n \frac{b_i}{\min\{b_i,b_j\}}\|K_{ii}\|_\infty \int_{t_{m-1}}^t \big(\mathbb{E}(|S(s)|^2|\mathcal{F}_{m-1})\big)^{1/2}\mathrm{d}s$$

$$\leq C\theta \int_{t_{m-1}}^t \big(\mathbb{E}(|S(s)|^2|\mathcal{F}_{m-1})\big)^{1/2}\mathrm{d}s.$$

Hence, we infer from (16) that

$$(17)\qquad \mathbb{E}(|S(t)|^2|\mathcal{F}_{m-1}) \leq C_5 \int_{t_{m-1}}^t \big(\mathbb{E}(|S(s)|^2|\mathcal{F}_{m-1})\big)^{1/2}\mathrm{d}s + d\sigma_i^2(t-t_{m-1}),$$

where $C_5 := 2C_V^{1/2}(1+C(q)(1+\theta^q)+|\widetilde{X}_i^k(t_{m-1})|^q)^{1/2}+C\theta$. We deduce from estimate (13) that the integrand on the right-hand side can be estimated according to

$$\mathbb{E}(|S(s)|^2|\mathcal{F}_{m-1}) \leq \frac{1}{2}\mathbb{E}\big(|\widetilde{X}_i^k(s)|^2\big|\mathcal{F}_{m-1}\big) + \frac{1}{2}|\widetilde{X}_i^k(t_{m-1})|^2$$

$$\leq \frac{C(2)}{2}(1+\theta^2) + \frac{1}{2}|\widetilde{X}_i^k(t_{m-1})|^2.$$

Inserting this estimate into (17), we conclude that

$$\mathbb{E}(|S(t)|^2|\mathcal{F}_{m-1}) \leq C_6\tau(1+\theta^{q/2+1})\big(1+|\widetilde{X}_i^k(t_{m-1})|^{q/2+1}\big),$$

where $C_6 > 0$ does not depend on b_i, p_i, or ξ_m. $\qquad\square$

We define the error process $Z_i^k(t) := \widetilde{X}_i^k(t) - X_i^k(t)$ and prove some estimates for $Z_i^k(t)$, generalizing [16, Lemma 3.4].

Lemma 6 (Control of the error process). *Let $i \in \{1,\ldots,n\}$, $k \in \{1,\ldots,N_i\}$, and $X_{0,i}^k \in L^{q_i'}(\Omega)$, where $q_i' = 2\max\{1,q_i\}$ and q_i is introduced in Assumption (A2). Then there exists a constant $C > 0$, which is independent of $(b_i,p_i)_{i=1,\ldots,n}$, and m such that for all $t \in (t_{m-1},t_m]$,*

$$(18)\quad \|Z_i^k(t) - Z_i^k(t_{m-1})\| \leq C\tau(1+\theta^{q_i'/2}), \quad |Z_i^k(t)| \leq C\tau\theta + |Z_i^k(t_{m-1})|,$$

$$(19)\quad \big|\mathbb{E}\big((Z_i^k(t) - Z_i^k(t_{m-1}))\chi_i^k(\widetilde{X}(t_{m-1}))\big)\big|$$

$$\leq C\tau\bigg((1+\theta^{3q_i'/2})\tau + (1+\theta^{q_i'})\|Z_i^k(t)\| + \sum_{j=1}^n \|Z_j^1(t)\|\bigg) + 8\tau\max_{j=1,\ldots,n}\|K_{ij}\|_\infty^2\Gamma_i,$$

where Γ_i and χ_i^k are defined in (7) and (9), respectively.

Proof. Since the Brownian motions are the same for X_i^k and $\widetilde{X}_i^k$, the process $Z_i^k(t)$ solves for $t \in (t_{m-1}, t_m]$ the deterministic equation

$$(20) \quad dZ_i^k(t) = -\big(\nabla V_i(\widetilde{X}_i^k(t)) - \nabla V_i(X_i^k(t))\big)dt + \sum_{j=1}^{n} \beta_{ij} \sum_{\ell \in \mathcal{C}_{j,r},\,(i,k)\neq(j,\ell)} K_{ij}(\widetilde{X}_i^k - \widetilde{X}_j^\ell)dt$$

$$- \sum_{j=1}^{n} \alpha_{ij} \sum_{\ell=1,\,(i,k)\neq(j,\ell)}^{N_j} K_{ij}(X_i^k - X_j^\ell)dt.$$

Step 1: Proof of (18). Let $(i,k) \in \mathcal{C}_{i,r}$ for some $r \leq b_i$. We take the expectation of the difference of the equations (20) solved by $Z_i^k(t)$ and $Z_i^k(t_{m-1})$, respectively, and distinguish the cases $j = i$ and $j \neq i$, leading to

$$(21) \qquad \|Z_i^k(t) - Z_i^k(t_{m-1})\| \leq J_{10} + \cdots + J_{14}, \quad \text{where}$$

$$J_{10} = \left\| \int_{t_{m-1}}^{t} \big(\nabla V_i(\widetilde{X}_i^k(s)) - \nabla V_i(X_i^k(s))\big)ds \right\|,$$

$$J_{11} = \frac{1}{p_i - 1} \left\| \int_{t_{m-1}}^{t} \sum_{\ell \in \mathcal{C}_{j,r},\,(i,k)\neq(j,\ell)} K_{ii}(\widetilde{X}_i^k(s) - \widetilde{X}_i^\ell(s))ds \right\|,$$

$$J_{12} = \sum_{j=1,\,j\neq i}^{n} \frac{b_i}{p_j \min\{b_i, b_j\}} \left\| \int_{t_{m-1}}^{t} \sum_{\ell \in \mathcal{C}_{j,r}} K_{ij}(\widetilde{X}_i^k(s) - \widetilde{X}_j^\ell(s))ds \right\|,$$

$$J_{13} = \frac{1}{N_i - 1} \sum_{\ell=1,\,\ell\neq k}^{N_j} \left\| \int_{t_{m-1}}^{t} K_{ii}(X_i^k(s) - X_i^\ell(s))ds \right\|,$$

$$J_{14} = \sum_{j=1,\,j\neq i}^{n} \frac{1}{N_j} \sum_{\ell=1}^{N_j} \left\| \int_{t_{m-1}}^{t} K_{ij}(X_i^k(s) - X_j^\ell(s))ds \right\|.$$

For the first term, we use the Cauchy–Schwarz inequality, the growth condition of ∇V_i, and stability estimate (13) with $q = q_i'$:

$$J_{10} \leq \sqrt{\tau} \left(\mathbb{E} \int_{t_{m-1}}^{t} \big|\nabla V_i(\widetilde{X}_i^k(s)) - \nabla V_i(X_i^k(s))\big|^2 ds \right)^{1/2}$$

$$\leq C_V \sqrt{\tau} \left(\mathbb{E} \int_{t_{m-1}}^{t} \big(1 + \mathbb{E}|\widetilde{X}_i^k(s)|^q + \mathbb{E}|X_i^k(s)|^q\big)ds \right)^{1/2} \leq C\tau(1 + \theta^{q/2}).$$

For the remaining terms, we exploit the boundedness of K_{ij}, yielding

$$J_{11} + \cdots + J_{14} \leq C\tau \sum_{j=1}^{n} \left(1 + \frac{b_i}{\min\{b_i, b_j\}} \right) \|K_{ij}\|_\infty \leq C\tau(1 + \theta).$$

Thus, we deduce from (21) that

$$\|Z_i^k(t) - Z_i^k(t_{m-1})\| \leq C\tau(1 + \theta + \theta^{q/2}),$$

which proves the first inequality in (18).

We estimate similarly as in the proof of Lemma 4, using the strong convexity of V_i and the boundedness of K_{ij}:

$$\frac{\mathrm{d}}{\mathrm{d}t}|Z_i^k(t)|^2 \le -r_i|Z_i^k(t)|^2 + \sum_{j=1}^n \frac{b_i\|K_{ij}\|_\infty}{\min\{b_i, b_j\}}|Z_i^k(t)| \le C\theta|Z_i^k(t)|.$$

This implies after integration with respect to time that $|Z_i^k(t)| \le C\theta\tau + |Z_i^k(t_{m-1})|$, showing the second inequality in (18).

Step 2: Proof of (19). Set $\Delta K_{ij}^\ell := K_{ij}(\widetilde{X}_i^k - \widetilde{X}_j^\ell) - K_{ij}(X_i^k - X_j^\ell)$. Using the formulation (8) for $\widetilde{X}_i^k$, we find that

(22)
$$\left|\mathbb{E}\big((Z_i^k(t) - Z_i^k(t_{m-1}))\chi_i^k(X(t))\big)\right| \le J_{15} + \cdots + J_{18}, \quad \text{where}$$

$$J_{15} = \mathbb{E}\bigg(\int_{t_{m-1}}^t \big|\nabla V_i(\widetilde{X}_i^k(s)) - \nabla V_i(X_i^k(s))\big|\mathrm{d}s|\chi_i^k(X(t))|\bigg),$$

$$J_{16} = \mathbb{E}\bigg(\frac{1}{p_i - 1}\int_{t_{m-1}}^t \sum_{\ell \in \mathcal{C}_{i,r}, \ell \ne k} |\Delta K_{ii}^\ell(s)|\mathrm{d}s|\chi_i^k(X(t))|\bigg),$$

$$J_{17} = \mathbb{E}\bigg(\sum_{j=1, j\ne i}^n \frac{b_i}{p_j \min\{b_i, b_j\}}\int_{t_{m-1}}^t \sum_{\ell \in \mathcal{C}_{j,r}} |\Delta K_{ij}^\ell(s)|\mathrm{d}s|\chi_i^k(X(t))|\bigg),$$

$$J_{18} = \mathbb{E}\bigg|\int_{t_{m-1}}^t \chi_i^k(X(s))\mathrm{d}s \cdot \chi_i^k(X(t))\bigg|.$$

For the term J_{15}, we use the mean-value theorem and the growth condition for D^2V_i (Assumption (A2)):

$$\big|\nabla V_i(\widetilde{X}_i^k - \nabla V_i(X_i^k)\big| \le |\widetilde{X}_i^k - X_i^k|\int_0^1 |D^2V_i(\widetilde{X}_i^k - \eta(\widetilde{X}_i^k - X_i^k))|\mathrm{d}\eta$$

$$\le C_V|Z_i^k|\int_0^1 \big(1 + |\widetilde{X}_i^k - \eta(\widetilde{X}_i^k - X_i^k)|^{q_i}\big)d\eta \le C|Z_i^k|(1 + |\widetilde{X}_i^k|^{q_i} + |X_i^k|^{q_i}).$$

Since $|\chi_i^k| \le 2\sum_{j=1}^n \|K_{ij}\|_\infty$, the Cauchy–Schwarz inequality and stability estimate (13) lead to

$$J_{15} \le C\sum_{j=1}^n \|K_{ij}\|_\infty \int_{t_{m-1}}^t \mathbb{E}\big(|Z_i^k(s)|(1 + |\widetilde{X}_i^k(s)|^{q_i} + |X_i^k(s)|^{q_i})\big)\mathrm{d}s$$

$$\le C\sum_{j=1}^n \|K_{ij}\|_\infty \int_{t_{m-1}}^t \|Z_i^k(s)\|\|1 + |\widetilde{X}_i^k(s)|^{q_i} + |X_i^k(s)|^{q_i}\|\mathrm{d}s$$

$$\le C(q)\tau(1 + \theta^{q_i})\big(\tau(1 + \theta^{q_i/2}) + \|Z_i^k(t)\|\big).$$

The last inequality follows from

(23)
$$\|Z_i^k(s)\| \le \|Z_i^k(s) - Z_i^k(t_{m-1})\| + \|Z_i^k(t_{m-1}) - Z_i^k(t)\| + \|Z_i^k(t)\|$$

$$\leq 2C\tau(1 + \theta^{q_i/2}) + \|Z_i^k(t)\|,$$

which in turn is a consequence of estimate (18). We conclude that

$$J_{15} \leq C\tau^2(1 + \theta^{3q_i/2}) + C\tau(1 + \theta^{q_i})\|Z_i^k(t)\|.$$

We use the Lipschitz continuity of K_{ij} (Assumption (A1)) to obtain

$$J_{16} \leq \frac{2}{p_i - 1} \sum_{j=1}^{n} \|K_{ij}\|_\infty \int_{t_{m-1}}^{t} \mathbb{E} \sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} |\Delta K_{ij}^\ell(s)| \mathrm{d}s$$

$$\leq \frac{CL_{ii}}{p_i - 1} \int_{t_{m-1}}^{t} \mathbb{E} \sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} \left(|\widetilde{X}_i^k(s) - X_i^k(s)| + |X_i^\ell(s) - \widetilde{X}_i^\ell(s)|\right) \mathrm{d}s$$

$$\leq \frac{CL_{ii}}{p_i - 1} \int_{t_{m-1}}^{t} \left(\left\| \sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} Z_i^k(s) \right\| + \left\| \sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} Z_i^\ell(s) \right\|\right) \mathrm{d}s.$$

It follows from the second estimate in (18), i.e. $|Z_i^k(t)| \leq C\tau\theta + |Z_i^k(t_{m-1})|$, that

$$J_{16} \leq \frac{CL_{ii}}{p_i - 1} \int_{t_{m-1}}^{t} \left((p_i - 1)C\tau\theta + (p_i - 1)\|Z_i^k(s)\| + \left\| \sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} Z_i^\ell(t_{m-1}) \right\|\right) \mathrm{d}s.$$

The variable $Z_i^\ell(s)$ is $\mathcal{G}_{m-1}$-measurable for all $t_{m-1} < s < t$ and hence it is independent of ξ_{m-1}. Therefore, we can apply Lemma 3 to the last term of the integrand to find that

$$(24) \qquad J_{16} \leq CL_{ii} \int_{t_{m-1}}^{t} \left(C\tau\theta + \|Z_i^k(s)\| + \|Z_i^k(t_{m-1})\|\right) \mathrm{d}s.$$

Here, we have taken into account the fact that $\|Z_i^k(t)\| = \|Z_i^\ell(t)\|$ for every $k, \ell = 1, \ldots, N_i$. The last two terms of the integrand can be estimated, by estimate (18), according to (23) and

$$\|Z_i^k(t_{m-1})\| \leq \|Z_i^k(t_{m-1}) - Z_i^k(t)\| + \|Z_i^k(t)\| \leq C\tau(1 + \theta^{q_i/2}) + \|Z_i^k(t)\|.$$

Hence, we conclude from (24) that

$$J_{16} \leq C\tau\left(\tau(1 + \theta^{q/2}) + \|Z_i^k(t)\|\right),$$

where $C > 0$ does not depend on b_i, p_i, or m and recalling that we have chosen $q = 2\max\{1, q_i\}$. Similar arguments lead to

$$J_{17} \leq 2\tau \sum_{j=1, j\neq i}^{n} \|K_{ij}\|_\infty\left(\tau(1 + \theta^{q/2}) + \|Z_i^k(t)\| + \|Z_j^1(t)\|\right)$$

$$\leq C\tau\left(\tau(1 + \theta^{q/2}) + \|Z_i^k(t)\| + \sum_{j=1, j\neq i}^{n} \|Z_j^1(t)\|\right).$$

Finally, we estimate the remaining term. By the Cauchy–Schwarz inequality,

$$(25) \qquad J_{18} = \int_{t_{m-1}}^{t} \mathbb{E}\left|\chi_i^k(X(s)) \cdot \chi_i^k(X(t))\right| \mathrm{d}s \leq \int_{t_{m-1}}^{t} \|\chi_i^k(X(s))\| \|\chi_i^k(X(t))\| \mathrm{d}s.$$

By Lemma 8 in the appendix,

$$\|\chi_i^k(X(s))\|^2 = \mathbb{E}|\chi_i^k(X(s))|^2 = \mathbb{E}\big[\mathbb{E}\big(|\chi_i^k(X(s))|^2\big|\sigma(X(s))\big)\big],$$

where $\sigma(X(s))$ is the σ-algebra generated by $X(s)$. Proposition 2 states that $\mathbb{E}\chi_i^k(X(s)) = 0$ and $\mathrm{Var}(\chi_i^k(X(s))) \leq 8\max_{j=1,\ldots,n}\|K_{ij}\|_\infty^2\Gamma_i$. Therefore,

$$\|\chi_i^k(X(s))\|^2 = \mathrm{Var}_{\sigma(X(s))}(\chi_i^k(X(s))) \leq 8\max_{j=1,\ldots,n}\|K_{ij}\|_\infty^2\Gamma_i.$$

Inserting this estimate into (25) leads to

$$J_{18} \leq 8\tau\max_{j=1,\ldots,n}\|K_{ij}\|_\infty^2\Gamma_i.$$

Summarizing, we obtain from (22)

$$\big|\mathbb{E}\big((Z_i^k(t) - Z_i^k(t_{m-1}))\chi_i^k(X(t))\big)\big|$$

$$\leq C(q)\tau\bigg(\tau(1 + \theta^{3q/2}) + (1 + \theta^q)\|Z_i^k(t)\| + \sum_{j=1}^n\|Z_j^1(t)\|\bigg) + 8\max_{j=1,\ldots,n}\|K_{ij}\|_\infty^2\Gamma_i,$$

which finishes the proof. $\qquad\square$

4. Proof of Theorem 1

Let $i \in \{1,\ldots,n\}$ and $k \in \{1,\ldots,N_i\}$ be such that $(i,k) \in \mathcal{C}_{i,r}$ for some $r \leq b_i$. As in the last section, we set $\Delta K_{ij}^\ell := K_{ij}(\widetilde{X}_i^k - \widetilde{X}_j^\ell) - K_{ij}(X_i^k - X_j^\ell)$. The process Z_i^k satisfies

$$\mathrm{d}Z_i^k(t) = -(\nabla V_i(\widetilde{X}_i^k(t)) - \nabla V_i(X_i^k(t)))\mathrm{d}t + \frac{1}{N_i - 1}\sum_{\ell=1,\,\ell\neq k}^{N_i}\Delta K_{ii}^\ell(t)\mathrm{d}t$$

$$+ \sum_{j=1,\,j\neq i}^n\frac{1}{N_j}\sum_{\ell=1}^{N_j}\Delta K_{ij}^\ell(t)\mathrm{d}t + \chi_i^k(\widetilde{X}(t))\mathrm{d}t.$$

In particular, Z_i^k is pathwise a.e. differentiable in time.

Step 1: Differential inequality for $|Z_i^k|^2$. Together with the strong convexity of V_i (Assumption (A2)) and the Lipschitz continuity of K_{ij} (Assumption (A1)), we find that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|Z_i^k|^2 = -(\nabla V_i(\widetilde{X}_i^k) - \nabla V_i(X_i^k))\cdot Z_i^k + \frac{1}{N_i - 1}\sum_{\ell=1,\,\ell\neq k}^{N_i}\Delta K_{ii}^\ell\cdot Z_i^k$$

$$+ \sum_{j=1,\,j\neq i}^n\frac{1}{N_j}\sum_{\ell=1}^{N_j}\Delta K_{ij}^\ell\cdot Z_i^k + \chi_i^k(\widetilde{X})\cdot Z_i^k$$

$$\leq -r_i|Z_i^k|^2 + \frac{L_{ii}}{N_i - 1}\sum_{\ell=1,\,\ell\neq k}^{N_i}(|Z_i^k| + |Z_i^\ell|)|Z_i^k|$$

$$+ \sum_{j=1,\, j\neq i}^{n} \frac{L_{ij}}{N_j} \sum_{\ell=1}^{N_j} (|Z_i^k| + |Z_j^\ell|)|Z_i^k| + \chi_i^k(\widetilde{X}) \cdot Z_i^k.$$

By taking the expectation and using Young's inequality, it follows after a standard computation that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}|Z_i^k|^2 \leq -r_i\mathbb{E}|Z_i^k|^2 + \frac{3}{2}\sum_{j=1}^{n} L_{ij}\mathbb{E}|Z_i^k|^2 + \frac{1}{2}\sum_{j=1}^{n} L_{ij}|Z_j^k|^2 + \mathbb{E}(\chi_i^k(\widetilde{X}) \cdot Z_i^k).$$

Without loss of generality, we may take $k = 1$ (since the distributions coincide). A summation over $i = 1, \ldots, n$ and exchanging the summation indices in the third term of the right-hand side leads to

$$(26) \quad \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{n}\mathbb{E}|Z_i^1|^2 \leq -\sum_{i=1}^{n} r_i\mathbb{E}|Z_i^1|^2 + \frac{3}{2}\sum_{i,j=1}^{n} L_{ij}\mathbb{E}|Z_i^1|^2 + \frac{1}{2}\sum_{i,j=1}^{n} L_{ji}\mathbb{E}|Z_i^1|^2$$

$$+ \sum_{i=1}^{n}\mathbb{E}(\chi_i^1(\widetilde{X}) \cdot Z_i^1)$$

$$\leq -\min_{i=1,\ldots,n}\left(r_i - 2\sum_{j=1}^{n}\max\{L_{ij}, L_{ji}\}\right)\sum_{i=1}^{n}\|Z_i^1\|^2 + \mathbb{E}(\chi_i^1(\widetilde{X}) \cdot Z_i^1).$$

It remains to estimate the last term $\mathbb{E}(\chi(\widetilde{X}) \cdot Z_i^1)$. To this end, we write

$$(27) \qquad \mathbb{E}(\chi(\widetilde{X}(t)) \cdot Z_i^1(t)) = J_{19} + \cdots + J_{22}, \quad \text{where}$$

$$J_{19} = \mathbb{E}\big(Z_i^1(t_{m-1}) \cdot \chi_i^1(\widetilde{X}(t_{m-1}))\big),$$

$$J_{20} = \mathbb{E}\big((Z_i^1(t) - Z_i^1(t_{m-1})) \cdot \chi_i^1(X(t))\big),$$

$$J_{21} = \mathbb{E}\big(Z_i^1(t_{m-1}) \cdot (\chi_i^1(\widetilde{X}(t)) - \chi_i^1(\widetilde{X}(t_{m-1})))\big),$$

$$J_{22} = \mathbb{E}\big((Z_i^1(t) - Z_i^1(t_{m-1})) \cdot (\chi_i^1(\widetilde{X}(t)) - \chi_i^1(X(t)))\big).$$

Step 2: Estimate of J_{19} and J_{20}. Since $\xi_{m,i}$ is independent of $\mathcal{G}_{m-1}$ and $Z_i^1(t_{m-1})$ is $\mathcal{G}_{m-1}$-measurable, we obtain from Lemma 8 in the appendix that

$$\mathbb{E}\big(Z_i^1(t_{m-1})\big|\mathcal{G}_{m-1}\big) = Z_i^1(t_{m-1}),$$

$$\mathbb{E}\big(Z_i^1(t_{m-1}) \cdot \chi(\widetilde{X}(t_{m-1}))\big|\mathcal{G}_{m-1}\big) = Z_i^1(t_{m-1}) \cdot \mathbb{E}\big(\chi(\widetilde{X}(t_{m-1}))\big|\mathcal{G}_{m-1}\big).$$

This shows that, using Proposition 2,

$$J_{19} = \mathbb{E}\big[\mathbb{E}\big(Z_i^1(t_{m-1}) \cdot \chi(\widetilde{X}(t_{m-1}))\big|\mathcal{G}_{m-1}\big)\big] = \mathbb{E}\big[Z_i^1(t_{m-1}) \cdot \mathbb{E}\big(\chi(\widetilde{X}(t_{m-1}))\big|\mathcal{G}_{m-1}\big)\big] = 0.$$

The term J_{20} can be directly estimated from (19):

$$J_{20} \leq C\tau\left((1 + \theta^{3q_i'/2})\tau + (1 + \theta^{q_i'})\|Z_i^1(t)\| + \sum_{j=1}^{n}\|Z_j^1(t)\|\right) + 8\tau\max_{j=1,\ldots,n}\|K_{ij}\|_\infty^2\Gamma_i.$$

Step 3: Estimate of J_{21}. We observe that $Z_i^1(t_{m-1})$ is $\mathcal{F}_{m-1}$-measurable. By the law of total expectation (Lemma 8) and the Cauchy–Schwarz inequality,

$$(28) \qquad J_{21} = \mathbb{E}\big[Z_i^1(t_{m-1})\mathbb{E}\big(\chi_i^1(\widetilde{X}_i^1(t)) - \chi_i^1(\widetilde{X}(t_{m-1}))\big|\mathcal{F}_{m-1}\big)\big]$$
$$\leq \|Z_i^1(t_{m-1})\|\,\big\|\mathbb{E}\big(\chi_i^1(\widetilde{X}_i^1(t)) - \chi_i^1(\widetilde{X}(t_{m-1}))\big|\mathcal{F}_{m-1}\big)\big\|.$$

We deduce from (23) that the first factor on the right-hand side is bounded from above by

$$(29) \qquad \|Z_i^1(t_{m-1})\| \leq C\tau(1 + \theta^{q_i'/2}) + \|Z_i^k(t)\|.$$

For the second factor, we introduce the notation

$$\Delta\widetilde{K}_{ij}^\ell := K_{ij}(\widetilde{X}_i^1(t) - \widetilde{X}_j^\ell(t)) - K_{ij}(\widetilde{X}_i^1(t_{m-1}) - \widetilde{X}_j^\ell(t_{m-1})),$$
$$\Delta\widetilde{X}_{ij}^\ell := (\widetilde{X}_i^1(t) - \widetilde{X}_j^\ell(t)) - (\widetilde{X}_i^1(t_{m-1}) - \widetilde{X}_j^\ell(t_{m-1})).$$

Since ξ_m is $\mathcal{F}_{m-1}$-measurable, we can write the second factor on the right-hand side of (28) as follows:

$$(30) \quad \mathbb{E}\big(\chi_i^1(\widetilde{X}(t)) - \chi_i^1(\widetilde{X}(t_{m-1}))\big|\mathcal{F}_{m-1}\big)$$

$$= \frac{1}{p_i - 1}\sum_{\ell \in \mathcal{C}_{i,r},\,\ell \neq 1}\mathbb{E}(\Delta\widetilde{K}_{ii}^\ell|\mathcal{F}_{m-1}) - \frac{1}{N_i - 1}\sum_{\ell=1,\,\ell\neq i}^{N_i}\mathbb{E}(\Delta\widetilde{K}_{ii}^\ell|\mathcal{F}_{m-1})$$

$$+ \sum_{j=1,\,j\neq i}^n \frac{b_i}{p_j\min\{b_i,b_j\}}\sum_{\ell\in\mathcal{C}_{j,r}}\mathbb{E}(\Delta\widetilde{K}_{ij}^\ell|\mathcal{F}_{m-1}) - \sum_{j=1,\,j\neq i}^n\frac{1}{N_j}\sum_{\ell=1}^{N_j}\mathbb{E}(\Delta\widetilde{K}_{ij}^\ell|\mathcal{F}_{m-1}).$$

We perform a Taylor expansion of K_{ij} at $\widetilde{X}_i^1(t_{m-1}) - \widetilde{X}_j^\ell(t_{m-1})$ and use the fact that K_{ij} is Lipschitz continuous with constant L_{ij}, such that DK_{ij} can be bounded from above by L_{ij}:

$$\big|\mathbb{E}(\Delta\widetilde{K}_{ii}^\ell|\mathcal{F}_{m-1})\big| \leq L_{ij}\big|\mathbb{E}(\Delta\widetilde{X}_{ij}^\ell|\mathcal{F}_{m-1})\big| + \frac{d}{2}\|D^2 K_{ij}\|_\infty\mathbb{E}\big(|\Delta\widetilde{X}_{ij}^\ell|^2\big|\mathcal{F}_{m-1}\big).$$

Inserting

$$\Delta\widetilde{X}_{ij}^\ell = \big(\widetilde{X}_i^1(t) - \widetilde{X}_i^1(t_{m-1})\big) + \big(\widetilde{X}_j^\ell(t) - \widetilde{X}_j^\ell(t_{m-1})\big)$$

into the previous estimate and taking into account the stability estimates of Lemmas 4 and 5, we infer that

$$\big\|\mathbb{E}(\Delta\widetilde{K}_{ii}^\ell|\mathcal{F}_{m-1})\big\| \leq C\tau L_{ij}\big(1 + \theta^{\widetilde{q}_i} + \theta^{\widetilde{q}_j}\big)$$
$$+ C\tau\|D^2 K_{ij}\|_\infty(1 + \theta^{q_i'/2+1})\big(1 + \||\widetilde{X}_i^1(t_{m-1})|^{q_i'/2+1}\|\big)$$
$$\leq C\tau\big(1 + \theta^{q_i'+2}\big),$$

where the constant $C > 0$ does not depend on b_i, p_i, or m. We use this estimate in (30) and observe that $b_i/\min\{b_i,b_j\} \leq \theta$, yielding

$$(31) \qquad \big\|\mathbb{E}\big(\chi_i^1(\widetilde{X}_i^1(t)) - \chi_i^1(\widetilde{X}(t_{m-1}))\big|\mathcal{F}_{m-1}\big)\big\| \leq C\tau\big(1 + \theta^{q_i'+3}\big).$$

Finally, we combine estimates (29) and (31) to conclude from (28) that

$$J_{21} \leq C\tau(1 + \theta^{q_i'+3})\|Z_i^1(t)\| + C\tau^2(1 + \theta^{3q_i'/2+3}).$$

Step 4: Estimate of J_{22}. Set $\Delta K_{ij}^\ell := K_{ij}(\widetilde{X}_i^1(t) - \widetilde{X}_j^\ell(t)) - K_{ij}(X_i^1(t) - X_j^\ell(t))$. We use the Cauchy–Schwarz inequality and (18) to obtain

$$
\begin{aligned}
(32) \qquad J_{22} &\leq \|Z_i^1(t) - Z_i^1(t_{m-1})\| \, \|\chi_i^1(\widetilde{X}(t)) - \chi_i^1(X(t))\| \\
&\leq C\tau(1 + \theta^{q_i'/2})\|\chi_i^1(\widetilde{X}(t)) - \chi_i^1(X(t))\| \\
&\leq C\tau(1 + \theta^{q_i'/2})\Bigg(\frac{1}{p_i - 1}\bigg\| \sum_{\ell \in \mathcal{C}_{i,r},\,\ell \neq k} \Delta K_{ii}^\ell \bigg\| + \frac{1}{N_i - 1} \sum_{\ell=1,\,\ell\neq k}^{N_i} \|\Delta K_{ii}^\ell\| \\
&\qquad + \sum_{j=1,\,j\neq i}^{n} \frac{b_i}{p_j \min\{b_i, b_j\}} \bigg\| \sum_{\ell \in \mathcal{C}_{j,r}} \Delta K_{ij}^\ell \bigg\| + \sum_{j=1,\,j\neq i}^{n} \frac{1}{N_j} \sum_{\ell=1}^{N_j} \|\Delta K_{ij}^\ell\| \Bigg).
\end{aligned}
$$

The difference ΔK_{ij}^ℓ can be estimated according to (see the second inequality in (18))

$$
|\Delta K_{ij}^\ell| \leq L_{ij}\big(|Z_i^1(t)| + |Z_j^\ell(t)|\big) \leq C\big(\tau\theta + |Z_i^1(t_{m-1})| + |Z_j^\ell(t_{m-1})|\big).
$$

Then, with the help of the auxiliary Lemma 3,

$$
\begin{aligned}
\frac{1}{p_i - 1}\bigg\| \sum_{\ell \in \mathcal{C}_{i,r},\,\ell\neq k} \Delta K_{ii}^\ell \bigg\| &\leq \frac{C}{p_i - 1}\bigg\| \sum_{\ell \in \mathcal{C}_{i,r},\,\ell\neq k} \big(\tau\theta + |Z_i^1(t_{m-1})| + |Z_i^\ell(t_{m-1})|\big) \bigg\| \\
&\leq C\tau\theta + C\|Z_i^1(t_{m-1})\| \leq C\tau(1 + \theta^{q_i'/2}) + C\|Z_i^1(t)\|, \\
\frac{1}{p_j}\bigg\| \sum_{\ell \in \mathcal{C}_{j,r},\,\ell\neq k} \Delta K_{ij}^\ell \bigg\| &\leq C\tau(1 + \theta^{\gamma/2}) + C\|Z_i^1(t)\| + C\|Z_j^1(t)\|,
\end{aligned}
$$

where $\gamma = \max_{j=1,\ldots,n} q_j'$. Therefore, because of $b_i / \min\{b_i, b_j\} \leq \theta$, (32) becomes

$$
J_{22} \leq C\tau(1 + \theta^{q_i'/2})(1 + \theta)\bigg(\tau(1 + \theta^{\gamma/2}) + \sum_{j=1}^{n} \|Z_j^1(t)\| \bigg).
$$

We deduce from (27) and the previous estimates for $J_{19}, \ldots, J_{22}$ that

$$
\begin{aligned}
(33) \quad \sum_{i=1}^{n} \mathbb{E}\big(\chi_i^1(\widetilde{X}(t)) \cdot Z_i^1\big) &\leq C\tau^2(1 + \theta^{3\gamma/2+3}) + C\tau(1 + \theta^{\gamma+3}) \sum_{i=1}^{n} \|Z_i^1(t)\| + C\tau \sum_{i=1}^{n} \Gamma_i \\
&\leq C_7(\theta)\tau^2 + C_8(\theta)\tau \bigg(\sum_{i=1}^{n} \|Z_i^1(t)\|^2 \bigg)^{1/2} + C_9\tau \sum_{i=1}^{n} \Gamma_i.
\end{aligned}
$$

Step 5: End of the proof. Let

$$
u(t) = \sum_{i=1}^{n} \|Z_i^1(t)\|^2, \quad r = \min_{i=1,\ldots,n}\bigg(r_i - 2 \sum_{j=1}^{n} \max\{L_{ij}, L_{ji}\} \bigg) > 0.
$$

We infer from (26) and (33) that

$$
\frac{\mathrm{d}u}{\mathrm{d}t} \leq -ru + C_7\tau^2 + C_8\tau u^{1/2} + C_9\tau \sum_{i=1}^{n} \Gamma_i.
$$

The positive solution z_+ of the quadratic equation $-rz^2 + C_7\tau^2 + C_8\tau z + C_9\tau \sum_{i=1}^n \Gamma_i = 0$ gives us an upper bound for $u(t)^{1/2}$, since $\mathrm{d}u/\mathrm{d}t \le 0$ otherwise. Consequently,

$$u(t)^{1/2} \le z_+ = \frac{C_9}{2r}\tau + \frac{\sqrt{\tau}}{2r}\left(C_8^2\tau + 4C_7 r\tau + 4C_9 r \sum_{i=1}^n \Gamma_i\right)^{1/2}$$

$$\le C\tau(1 + \theta^{3\gamma/2+3}) + C\sqrt{\tau}\left(\sum_{i=1}^n \Gamma_i\right)^{1/2}.$$

This ends the proof of Theorem 1.

5. PARTICLE SYSTEMS WITH MULTIPLICATIVE NOISE

The technique of the proof of Theorem 1 can be applied to particle systems with multiplicative noise,

$$\mathrm{d}X_i^k = -\nabla V_i(X_i^k)\mathrm{d}t + \sum_{j=1}^n \alpha_{ij} \sum_{\substack{\ell=1 \\ (i,k)\ne(j,\ell)}}^{N_j} K_{ij}(X_i^k - X_j^\ell)\mathrm{d}t + \sigma_i(X_i^k)\mathrm{d}B_i^k(t),$$

with initial conditions (2), and $\alpha_{ij} = 1/(N_j - \delta_{ij})$, $i, j = 1, \ldots, n$, $k = 1, \ldots, N_i$. The random-batch process $\widetilde{X}_i^k$ is defined as in (4) but with $\sigma_i(\widetilde{X}_i^k)$ instead of σ_i. In addition to Assumptions (A1)–(A4), we suppose the following conditions:

(B1) Diffusion: $\sigma_i \in C^0(\mathbb{R}^d)$ is bounded and Lipschitz continuous with Lipschitz constant $L_i > 0$.

(B2) Strong convexity: The function $x \mapsto V_i(x) - r_i|x|^2/2$ is convex, where $r_i > 2\sum_{j=1}^n \max\{L_{ij}, L_{ji}\} + L_i^2 d$ and $r_i > 2L_i^2(2\max\{1, q_i\} + d - 2)$, $i = 1, \ldots, n$.

Theorem 7. *Let Assumptions (A1)–(A2), (A4), (B1)–(B2) hold. Then there exists a constant $C > 0$, which is independent of $(b_i, p_i)_{i=1,\ldots,n}$, m, and T, such that*

$$\sup_{0<t<T} \sum_{i=1}^n \|(X_i^k - \widetilde{X}_i^k)(t)\| \le C\sqrt{\tau}\left(1 + \sum_{i=1}^n \Gamma_i\right)^{1/2} + C\tau(1 + \theta^\gamma),$$

and θ, γ, Γ_i are defined in (6)–(7).

Sketch of the proof. The proof is similar to that one for Theorem 1 except for some additional estimates for the multiplicative noise term. In particular, Proposition 2 keeps unchanged since it is concerned with the shuffling process only. For the stability (Lemma 4), we need the condition $2 \le q \le q_i' = 2\max\{1, q_i\}$. The proof is essentially the same, except for the estimate of the term $\frac{1}{2}q(q + d - 2)\mathbb{E}(\sigma_i^2|X_i^k|^{q-2})$. Here, we use the Lipschitz continuity of σ_i and the stricter condition on r_i in Assumption (B2). In the estimate for $\widetilde{X}_i^k(t) - \widetilde{X}_i^k(t_{m-1})$ (Lemma 5), the diffusion σ_i is controlled by the Lipschitz continuity, $\sigma_i(\widetilde{X}_i^k)^2 \le 2L_i^2|\widetilde{X}_i^k|^2 + 2\sigma_i(0)^2$, and Lemma 4. Finally, for the control of the error process (Lemma 6), estimates (18)–(19) need to be changed to

$$(34) \qquad \|Z_i^k(t) - Z_i^k(t_{m-1})\| \le C\tau(1 + \theta^{q_i'/2}) + C\sqrt{\theta},$$

$$(35) \quad \left|\mathbb{E}\big((Z_i^k(t) - Z_i^k(t_{m-1}))\chi_i^k(\widetilde{X}(t_{m-1}))\big)\right| \leq C\tau^2(1 + \theta^{3q_i'/2}) + 8\tau \max_{j=1,\dots,n} \|K_{ij}\|_\infty^2 \Gamma_i$$

$$+ \sqrt{\tau}\bigg((1 + \sqrt{\tau})(1 + \theta^{q_i'})\|Z_i^k(t)\| + \sqrt{\tau}(1 + \sqrt{\tau})\sum_{j=1}^n \|Z_j^1(t)\|\bigg).$$

For the proof of estimate (34), the right-hand side of (21) contains the additional term

$$\widetilde{J}_{14} = \left\|\int_{t_{m-1}}^t (\sigma_i(\widetilde{X}_i^k) - \sigma_i(X_i^k))\mathrm{d}B_i^k\right\|.$$

The square of $\widetilde{J}_{14}$ is estimated by using the Itô isometry and the Lipschitz continuity of σ_i. Integrating and taking the square root then leads to the additional $C\sqrt{\tau}$ term.

The proof of (35) is very similar to (19), except that we need the inequality

$$\left\|\frac{1}{p_i - 1} \sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} |Z_i^k|\right\| \leq C\theta(\tau + \|Z_i^1(t_{m-1})\|).$$

The square of the left-hand side is formulated as

$$\left\|\frac{1}{p_i - 1} \sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} |Z_i^k|\right\|^2 = \mathbb{E}\bigg\{\mathbb{E}\bigg[\bigg(\sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} |Z_i^k|\bigg)^2 \Big|\mathcal{F}_{m-1}\bigg]\bigg\}.$$

Since ξ_{m-1} is $\mathcal{F}_{m-1}$ measurable, the inner expectation becomes

$$\mathbb{E}\bigg[\bigg(\sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} |Z_i^k|\bigg)^2 \Big|\mathcal{F}_{m-1}\bigg] = \sum_{\ell,\ell' \in \mathcal{C}_{i,r}, \ell,\ell' \neq k} \mathbb{E}\big(|Z_i^\ell||Z_i^{\ell'}|\big|\mathcal{F}_{m-1}\big)$$

$$\leq \sum_{\ell,\ell' \in \mathcal{C}_{i,r}, \ell,\ell' \neq k} \sqrt{\mathbb{E}(|Z_i^\ell|^2|\mathcal{F}_{m-1})}\sqrt{\mathbb{E}(|Z_i^{\ell'}|^2|\mathcal{F}_{m-1})},$$

using the Cauchy–Schwarz inequality for the conditional expectation. A straightforward computation leads to

$$\mathbb{E}(|Z_i^\ell|^2|\mathcal{F}_{m-1}) \leq C\theta^2(\tau + |Z_i^k(t_{m-1})|)^2,$$

from which we infer that

$$\mathbb{E}\bigg[\bigg(\sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} |Z_i^k|\bigg)^2 \Big|\mathcal{F}_{m-1}\bigg] \leq C\theta^2\bigg(\sum_{\ell \in \mathcal{C}_{i,r}, \ell \neq k} (\tau + |Z_i^k(t_{m-1})|)\bigg)^2.$$

As $Z_i^\ell(t_{m-1})$ is independent of ξ_{m-1}, the proof finishes after applying Lemma 3. $\qquad\square$

A more complicated particle system with multiplicative noise was considered in [7], which leads in a mean-field-type limit to the Shigesada–Kawasaki–Teramoto population model:

$$\mathrm{d}X_i^k = -\nabla V_i(X_i^k)\mathrm{d}t + \bigg(\sigma_i^2 + \sum_{j=1}^n f\bigg(\alpha_{ij} \sum_{\substack{\ell=1 \\ (i,k)\neq(j,\ell)}}^{N_j} K_{ij}(X_i^k - X_j^\ell)\bigg)\bigg)^{1/2}\mathrm{d}B_i^k(t),$$

with initial conditions (2), $i = 1, \ldots, n$, $k = 1, \ldots, N_i$, and the function f is globally Lipschitz continuous. Again, the random-batch process $\widetilde{X}_i^k$ is similar to (4). For this system, we have been not able to prove an error estimate of order $\sqrt{\tau}$, but only a stability estimate of the form

$$\sum_{i=1}^{n} \|(X_i^k - \widetilde{X}_i^k)(t)\| \leq C(t)\sqrt{t}\left(\sqrt{\tau}h(t, \tau, \theta) + \sum_{i=1}^{n}\Gamma_i\right), \quad t > 0,$$

where $h(t, \tau, \theta)$ is a smooth function. Compared to the error estimates of Theorems 1 and 7, the bound $\sum_{i=1}^{n}\Gamma_i$ for the variance of the remainder (9) is not multiplied by $\sqrt{\tau}$. Numerical simulations (not shown) reveal a saturation effect when τ becomes very small, indicating that the previous estimate cannot be improved.

6. NUMERICAL SIMULATIONS

We present numerical results for a test example, a population system, and an opinion-formation model. The algorithm is implemented in Matlab. The random shuffling is realized using the command `randperm`, and the stochastic differential equations are discretized by the standard Euler–Maruyama scheme.

6.1. **Discrete L^2 error for a test example.** We generalize the test example of [16, Section 4.1]. For this, we consider system (1) with $n = 3$ species in $d = 2$ dimensions and specify the functions

$$\nabla V_i(x) = r_i(x - m^{(i)}), \quad K_{ij}(x) = \frac{Q_i Q_j x}{1 + |x|^2}, \quad x \in \mathbb{R}^2, \ i, j = 1, 2, 3,$$

where the model parameters are $(Q_1, Q_2, Q_3) = (-1, 2, -2)$, $(r_1, r_2, r_3) = (1, 4, 2)$, and $m^{(1)} = (1, 0)^T$, $m^{(2)} = -(1, 1)^T$, $m^{(3)} = (1, 1)^T$. This choice incorporates different repulsive and attracting effects. The initial data are centered Gaussian distributions with the variances $(v_1, v_2, v_3) = (2, 2, 1)$, where the index signifies the number of the species.

For the first experiment, we choose the diffusion coefficients $\sigma_i = 0.5$ for $i = 1, 2, 3$ and the time step sizes $\tau = 2^{-2}, \ldots, 2^{-6}$. The end time is $T = 1$, the batch sizes are $p_i = 2$ for $i = 1, 2, 3$, and the numbers N_i of particles of the ith species are $(N_1, N_2, N_3) = (100, 100, 200)$, $(1000, 1000, 2000)$, or $(2500, 2500, 5000)$. Thus the total number of particles is $N = 400$, 4000, or 10000. We compare the random-batch solution with a reference solution, obtained by solving the fully coupled system using the time step size $2^{-4}, \ldots, 2^{-8}$. Figure 2 (left) shows the discrete $L^2(\Omega)$ error for the different time step sizes, defined by

$$E = \left(\sum_{i=1}^{n}\frac{1}{N_i}\sum_{k=1}^{N_i}|\widetilde{X}_i^k(T) - X_i^k(T)|^2\right)^{1/2}.$$

The reference line has the slope $1/2$. The results clearly show that the convergence rate is of order $O(\sqrt{\tau})$ as predicted by Theorem 1.

Figure 2 (right) illustrates the $L^2(\Omega)$ error as a function of the computational time, represented by the number of FLOPs (floating-point operations). We choose $\sigma_i = 0$ for all $i = 1, \ldots, n$ to allow for the comparison of the random-batch solution with a reference
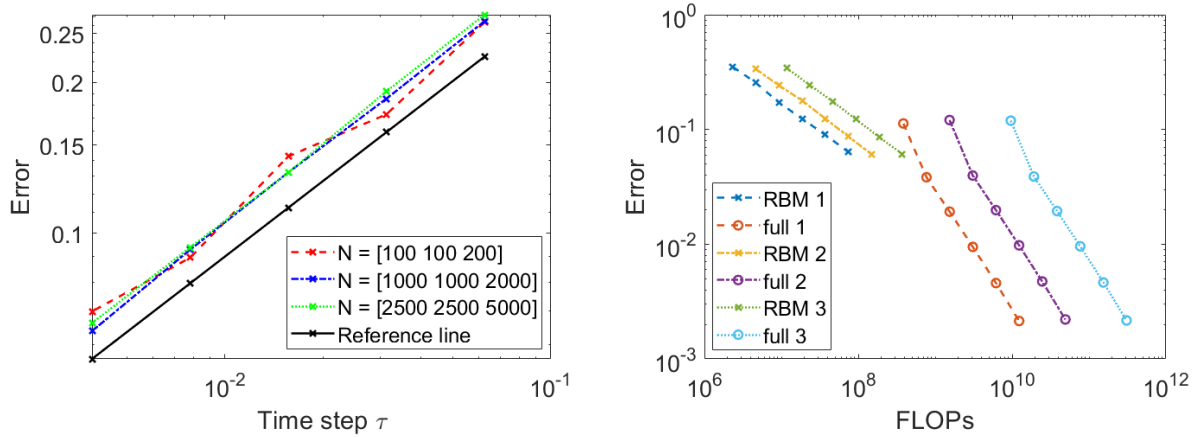
FIGURE 2. Left: Discrete $L^2(\Omega)$ error E versus time step size τ for various total particle numbers N. Right: Discrete $L^2(\Omega)$ error versus number of FLOPs for various random-batch simulations (RBM) and the corresponding reference solutions (ref).

solution that is calculated beforehand. The parameters for the random-batch algorithm are $T = 1$, $n = 2$, $d = 2$, $(p_1, p_2) = (2, 2)$, $\tau = 2^{-3}, \ldots, 2^{-7}$, and $(N_1, N_2) = (1250, 1250)$ (RBM1, full 1), $(2500, 2500)$ (RBM2, full 2), or $(5000, 5000)$ (RBM3, full 3). The reference solution is calculated from an explicit Euler scheme with the time step size $\tau = 2^{-1}, \ldots, 2^{-5}$. The number of FLOPs needed for the Matlab-internal functions are determined by the lightspeed toolbox of Tom Minka (https://github.com/tminka/lightspeed). The total numbers of FLOPs are then calculated by adding all needed operations manually.

Figure 2 (right) shows that the random-batch algorithm needs almost three orders of magnitude less FLOPs than the reference algorithm. As expected, the discrete $L^2(\Omega)$ error of the random-batch scheme is larger than that one of the reference scheme for a given time step. However, for a given error, the number of FLOPs of the random-batch algorithm is still much smaller compared to the reference algorithm, namely by about two orders of magnitude.

6.2. **A population system.** We consider the population system derived in [6] without external potentials using the following parameters: $n = 3$, $d = 1$, $T = 2$, $N_i = 5000$ for $i = 1, 2, 3$, and $(\sigma_1, \sigma_2, \sigma_3) = (1, 2, 3)$. The interaction kernels are given by $K_{ij} = \nabla B_{ij}^\eta$, where $B_{ij}^\eta(x) = \eta^{-1} B_{ij}(x/\eta)$, $B_{ij}(x) = D_{ij} \exp(1 - 1/(1 - |x|^2))1_{\{|x|<1\}}(x)$ for $x \in \mathbb{R}$, $\eta = 2$, and

$$(D_{ij}) = \begin{pmatrix} 0 & 355 & 355 \\ 25 & 0 & 25 \\ 355 & 0 & 0 \end{pmatrix}.$$

The initial data are Gaussian normal distributions with means $(m_1, m_2, m_3) = (-1, 2, 3)$ and variances $(v_1, v_2, v_3) = (2, 2, 2)$.

Figure 3 (left) illustrates the approximate probability densities at time $T = 2$ obtained by simulating the particle system 1000 times with the batch sizes $p_i = 20$ for $i = 1, 2, 3$ and the time step size $\tau = 10^{-2}$. We observe that the species segregate and avoid each other. Each of the simulation requires about $2 \cdot 10^{10}$ FLOPs, which needs to be compared to about $5 \cdot 10^{12}$ FLOPs required when using full interactions. This is a reduction of the numerical effort of more than two orders of magnitude.

Clearly, the reduction of computational cost comes at the price of an increased error. Figure 3 (right) presents the discrete $L^2(\Omega)$ error versus the number of FLOPs for various configurations of the batch sizes and various time step sizes. The end time is $T = 1$, and we used batch sizes $p_i = 2, 10, 100, 1000$ and time step sizes $\tau = 2^{-1}, \ldots, 2^{-7}$. The different points per line correspond to different values of τ. The reference solution is computed from the Euler–Maruyama scheme with the step size $\tau = 2^{-9}$; this simulation needed about 10^{13} FLOPs. We see that the error decreases with the time step size and larger batch sizes. The red dot in the figure indicates the number of FLOPs needed to compute a numerical solution with full interactions and step size $\tau = 10^{-2}$, to give a more practical point of reference. This simulation required about $2.5 \cdot 10^{12}$ FLOPs, while the random-batch algorithm with $\tau = 2^{-7}$ was about four times faster.
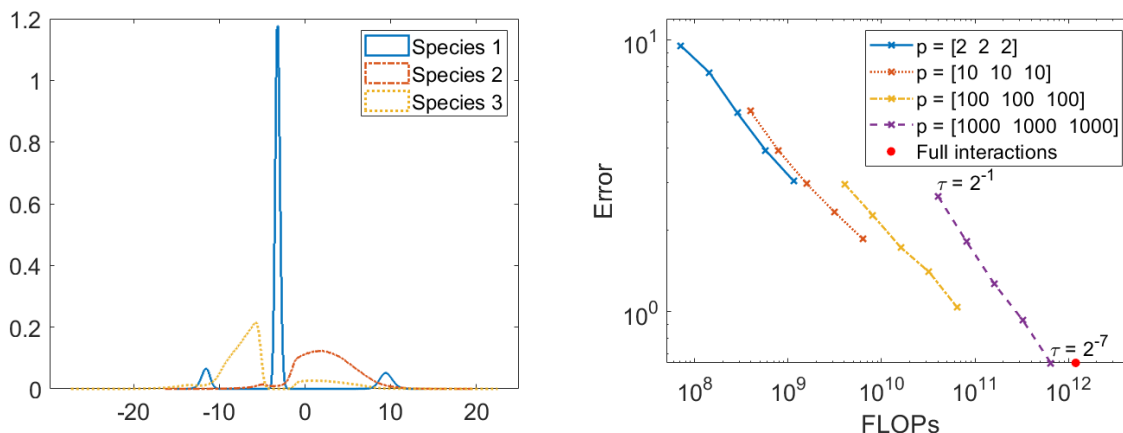

FIGURE 3. Left: Histogram of the population model derived in [6] for three species at time $T = 2$. Right: Discrete $L^2(\Omega)$ error versus number of FLOPs for various batch sizes p and time step sizes τ.

6.3. **Opinion dynamics model.** We model a company whose internal hierarchy regulates the communication between three different types of agents: workers (species 1), managers (species 2), and CEOs (species 3). The agents obey the following rules:

- CEOs can be only influenced by other CEOs. They influence managers (but not vice versa) and they do not interact with workers.
- Managers can influence workers but not other managers or CEOs.
- Workers can only influence each other.

The dynamics of opinions is described by the system

$$dX_i^k(t) = \sum_{j=1}^{3} \frac{1}{N_i - \delta_{ij}} \sum_{\ell=1,\,(i,k)\neq(j,\ell)}^{N_j} K_{ij}(X_i^k(t) - X_j^\ell(t))dt + \sigma dt,$$

$$X_i^k(0) = X_{0,i}^k, \quad i, j = 1, 2, 3, \ k = 1, \dots, N_i, \ 0 < t \leq T,$$

which is a generalization of a model discussed in [24]. The interaction is modeled by $K_{ij}(x) = -D_{ij}\phi(x/R_j)x$ for $x \in \mathbb{R}$, where $\phi(x) = \exp(1 - 1/(1 - |x|^{10}))1_{(-1,1)}$ is a smooth approximation of the characteristic function $1_{(-1,1)}$. The value D_{ij} is a measure of the influence that an agent of species j has over an agent of species i. According to the above-mentioned interaction rules, the matrix $D = (D_{ij})$ has the structure

$$D = \begin{pmatrix} D_{11} & D_{12} & 0 \\ 0 & 0 & D_{23} \\ 0 & 0 & D_{33} \end{pmatrix}.$$

As the only way for CEOs to communicate with the workers happens indirectly via the managers, we wish to explore the influence of the managers to achieve a consensus. In particular, we consider managers that are very submissive to authority ($D_{23} \gg 1$) or that are less obedient ($D_{23} \leq 1$). For the simulations, we use 5000 workers, 10 managers and 2 CEOs. The parameters are $\sigma = 0.1$, $T = 5$, $\tau = 10^{-5}$, and $(p_1, p_2, p_3) = (20, 2, 2)$. The initial conditions are drawn from a uniform distribution on the interval $[0, 10]$. The interaction radii are $(R_1, R_2, R_3) = (1, 2.5, 5)$.

In the first case (submissive managers), we choose the influence values

$$D_{11} = 5, \quad D_{12} = 10, \quad D_{23} = 25, \quad D_{33} = 0.1.$$

Figure 4 (left) shows one simulation of the particle system. We observe that the managers are very eager to find a compromise between the opinions of the two CEOs. This change of the opinion occurs too fast for the workers with more extreme opinions, as they are not as susceptible as the managers (since $D_{12} < D_{23}$). Therefore, they leave quickly the range of interaction of the managers and form their own clusters. Only those workers who have an opinion already close to that one of the CEOs, agree with the company policy and change their opinion accordingly.

In the second case (less obedient managers), we choose the same values of D_{ij} as before except $D_{23} = 1$. This means that the influence of the CEOs over the managers is rather small. Figure 4 (right) shows that the managers change their opinion slowly enough for the workers to adapt their opinion, as they stay within their range of interaction. Eventually, this leads to a consensus of opinion.

The simulations suggest that small changes over time are more likely to lead in an adjustment of the opinion and eventually to a consensus. In this picture, managers should not impose their opinion too quickly, but they should introduce the changes sufficiently slowly such that the workers can adjust in time.

Finally, we explore the influence of the batch size on the running time and the error. We consider 10000 workers, 100 managers, and 10 CEOs and choose the parameters $\tau =$
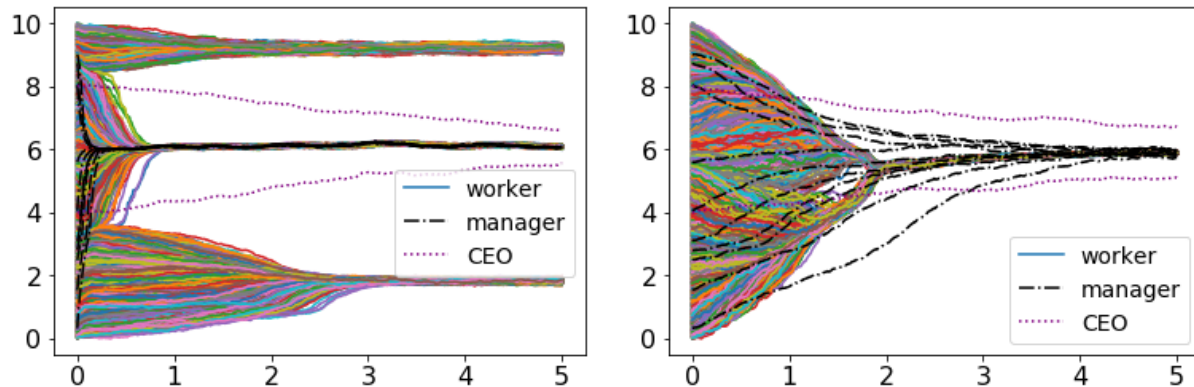
FIGURE 4. Opinion versus time of the CEOs and managers in the case of very submissive (left) or less obedient (right) managers.

$2^{-3}, \ldots, 2^{-7}$, $T = 4$, and $\sigma = 0.1$. The batch sizes are $(p_1, p_2, p_3) = (2, 2, 2)$, $(20, 5, 2)$, $(200, 20, 2)$, and $(2000, 20, 2)$. Figure 5 shows that the discrete $L^2(\Omega)$ error decreases with larger batch sizes (since this involves more interactions), smaller time step sizes, or θ closer to one, which is consistent with our error estimate. Clearly, the number of FLOPs increases with larger batch sizes.
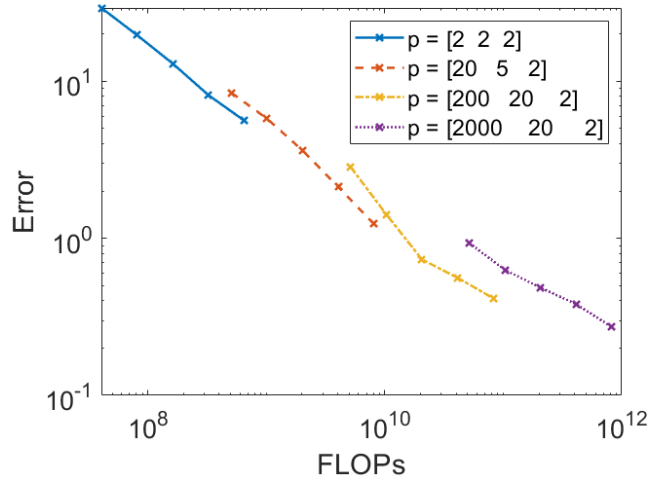

FIGURE 5. $L^2(\Omega)$ error versus number of FLOPs for different batch sizes and time step sizes $\tau = 2^{-3}, \ldots, 2^{-7}$.

APPENDIX A. AUXILIARY RESULTS

We recall some results involving the conditional expectation; see [9, Chapter 5]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Lemma 8. *Let $\mathcal{H}$ be a sub-σ-algebra of $\mathcal{F}$ and let X, $Y : \Omega \to \mathbb{R}^d$ be random variables such that X is $\mathcal{H}$-measurable. Then*

$$\mathbb{E}(X|\mathcal{H}) = X, \quad \mathbb{E}(XY|\mathcal{H}) = X\mathbb{E}(Y|\mathcal{H}).$$

In particular, the law of total expectation holds: $\mathbb{E}[\mathbb{E}(X|\mathcal{H})] = \mathbb{E}(X)$.

Lemma 9. *Let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra, and $(X(t))_{t\geq0}$ be an integrable stochastic process. Then, for any $t > 0$,*

$$\mathbb{E}\left(\int_0^t X(s)\mathrm{d}s \bigg| \mathcal{G} \right) = \int_0^t \mathbb{E}(X(s)|\mathcal{G})\mathrm{d}s.$$

The lemma is a consequence of Fubini's theorem [4, Lemma 2.3].

Lemma 10. *Let $T > 0$, $(B(t))_{t\geq0}$ be a d-dimensional Brownian motion, and $\mathcal{F}_t = \sigma(B(s), s \leq t)$ for $t \leq T$. Furthermore, let $X(t) \in \mathbb{R}^d$ be a square integrable, progressively measurable process with respect to $\mathcal{F}_t$. Then, for any $0 \leq s_1 \leq s_2 \leq T$,*

$$\mathbb{E}\left(\int_{s_1}^{s_2} X(t)\mathrm{d}B(t) \bigg| \mathcal{F}_{s_1} \right) = 0.$$

This lemma follows from the fact that $S(t) := \int_0^t X(s)\mathrm{d}B(s)$ is a martingale and consequently, $\mathbb{E}(S(s_1) - S(s_2)) = 0$ a.s. for $0 \leq s_1 \leq s_2 \leq T$.

References

[1] G. Albi and L. Pareschi. Binary interaction algorithms for the simulation of flocking and swarming dynamics. *Multiscale Model. Simul.* 11 (2013), 1–29.

[2] D. Aldous. Interacting particle systems as stochastic social dynamics. *Bernoulli* 19 (2013), 1122–1149.

[3] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms. I. *Commun. Pure Appl. Math.* 44 (1991), 141–183.

[4] R. A. Brooks. Conditional expectations associated with stochastic processes. *Pacific J. Math.* 41 (1972), 33–42.

[5] R. E. Caflisch. Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.* 1998 (1998), 1–49.

[6] L. Chen, E. S. Daus, and A. Jüngel. Rigorous mean-field limits and cross diffusion. *Z. Angew. Math. Phys.* 70 (2019), 122, 21 pages.

[7] L. Chen, E. S. Daus, A. Holzinger, and A. Jüngel. Rigorous derivation of population cross-diffusion systems from moderately interacting particle systems. Submitted for publication, 2020. arXiv:2010.12389.

[8] S. Fagioli and E. Radici. Opinion formation systems via deterministic particles approximation. *Kinetic Related Models* 14 (2021), 45–76.

[9] W. Feller. *An Introduction to Probability Theory and Its Applications.* John Wiley & Sons, New York, 1970.

[10] F. Golse, S. Jin, and T. Paul. The random batch method for N-body quantum dynamics. To appear in *J. Comput. Math.*, 2021. arXiv:1912.07424.

[11] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.* 73 (1987), 325–348.

[12] A.-L. Haji-Ali and R. Tempone. Multilevel and multi-index Monte Carlo methods for the McKean–Vlasov equation. *Stat. Comput.* 28 (2018), 923–935.

[13] P.-E. Jabin and Z. Wang. Mean field limit for stochastic particle systems. In: *Active Particles*, Vol. 1, pp. 379–402. Springer, Boston, 2017.

[14] S. Jin and L. Li. On the mean field limit of the Random Batch Method for interacting particle systems. Submitted for publication, 2020. arXiv:2005.11740.

[15] S. Jin and L. Li. Random Batch Methods for classical and quantum interacting particle systems and statistical samplings. Submitted for publication, 2021. arXiv:2104.04337.

[16] S. Jin, L. Li, and J.-G. Liu. Random batch methods (RBM) for interacting particle systems. *J. Comput. Phys.* 400 (2020), 108877, 30 pp.

[17] S. Jin, L. Li, and J.-G. Liu. Convergence of Random Batch Method for interacting particles with disparate species and weights. *SIAM J. Numer. Anal.* 59 (2021), 746–768.

[18] S. Jin, L. Li, and Y. Sun. On the Random Batch Method for second order interacting particle systems. Submitted for publication, 2020. arXiv:2011.10778.

[19] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, Berlin, 1992.

[20] D. Ko and E. Zuazua. Model predictive control with random batch methods for a guiding problem. Submitted for publication, 2020. arXiv:2004.14834.

[21] L. Li, Y. Li, J.-G. Liu, Z. Liu, and J. Lu. A stochastic version of Stein variational gradient descent for efficient sampling. *Commun. Appl. Math. Comput. Sci.* 15 (2020), 37–63.

[22] L. Li, J.-G. Liu, and Y. Tang. A direct simulation approach for the Poisson–Boltzmann equation using the Random Batch Method. Submitted for publication, 2020. arXiv:2004.05614.

[23] M. Li, T. Zhang, Y. Chen, and A. Smola. Efficient mini-batch training for stochastic optimization. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, pp. 661–670.

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

[25] G. Rotskoff and E. Vanden-Eijnden. Trainability and accuracy of neural networks: An interacting particle system approach. Preprint, 2018. arXiv:1805.00915.

[26] D. Strömbom, R. Mann, A. Wilson, S. Hailes, A. Morton, D. J. T. Sumpter, and A. King. Solving the shepherding problem: heuristics for herding autonomous, interacting agents. *J. Roy. Soc. Interface* 11 (2014), 20140719, 9 pages.

[27] X. Ye and Z. Zhou. Efficient sampling of thermal averages of interacting quantum particle systems with random batches. *J. Chem. Phys.* 154 (2021), 204106, 24 pages.

INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING, VIENNA UNIVERSITY OF TECHNOLOGY, WIEDNER HAUPTSTRASSE 8–10, 1040 WIEN, AUSTRIA
E-mail address: esther.daus@tuwien.ac.at

INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING, VIENNA UNIVERSITY OF TECHNOLOGY, WIEDNER HAUPTSTRASSE 8–10, 1040 WIEN, AUSTRIA
E-mail address: markus.fellner@tuwien.ac.at

INSTITUTE FOR ANALYSIS AND SCIENTIFIC COMPUTING, VIENNA UNIVERSITY OF TECHNOLOGY, WIEDNER HAUPTSTRASSE 8–10, 1040 WIEN, AUSTRIA
E-mail address: juengel@tuwien.ac.at