CONVERGENCE ANALYSIS OF A STOCHASTIC INTERACTING PARTICLE-FIELD ALGORITHM FOR 3D PARABOLIC-PARABOLIC KELLER-SEGEL SYSTEMS*

BOYI HU†, ZHONGJIAN WANG ‡, JACK XIN §, AND ZHIWEN ZHANG¶

Abstract. Chemotaxis models describe the movement of organisms in response to chemical gradients. In this paper, we present a stochastic interacting particle-field algorithm with random batch approximation (SIPF-r) for the three-dimensional (3D) parabolic-parabolic Keller-Segel (KS) system, also known as the fully parabolic KS system. The SIPF-r method approximates the KS system by coupling particle-based representations of density with a smooth field variable computed using spectral methods. By incorporating the random batch method (RBM), we bypass the mean-field limit and significantly reduce computational complexity. Under mild assumptions on the regularity of the original KS system and the boundedness of numerical approximations, we prove that, with high probability, the empirical measure of the SIPF-r particle system converges to the exact measure of the limiting McKean-Vlasov process in the 1-Wasserstein distance. Numerical experiments validate the theoretical convergence rates and demonstrate the robustness and accuracy of the SIPF-r method.

Key words. Fully parabolic Keller-Segel system, stochastic interacting particle-field (SIPF) algorithm, random batch method (RBM), three-dimensional (3D) simulations, convergence analysis.

MSC codes. 35K51, 65C05, 65M12, 65M75, 65T50.

1. Introduction.

Chemotaxis is a biological phenomenon concerning the movement of organisms (e.g. bacteria) in response to signals, typically chemical substances known as chemo-attractants, which can be produced by the organisms themselves. Theoretical and mathematical modeling was initiated by Patlak [29], Keller and Segel [19]. In this work, we focus on the fully parabolic KS system as follows:

$$\begin{aligned} \rho_t &= \nabla \cdot (\mu \nabla \rho - \chi \rho \nabla c), \\ \epsilon c_t &= \Delta c - \lambda^2 c + \rho, \\ \mathbf{x} &\in \Omega \subseteq \mathbb{R}^d, \quad t \in [0, T], \end{aligned} \tag{1.1}$$

where χ, μ (ϵ, λ) are positive (non-negative) constants. The model is called elliptic if $\epsilon = 0$, and parabolic if $\epsilon > 0$. Here ρ denotes the density of active particles (bacteria), and c represents the concentration of a chemical substance (chemo-attractant) emitted by the bacteria. KS partial differential equation (PDE) systems have diverse applications across disciplines. In biology, they help explain cell aggregation and migration behaviors, such as those of bacteria and cancer cells, driven by chemical gradients [30]. Ecologically, these models describe how organisms navigate environments using chemical cues [28]. In medicine, KS models are pivotal for studying cell migration in tissues, offering insights into wound healing, immune responses, and cancer spread [2, 34].

*Submitted to the editors DATE.

†Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, P.R.China. huby22@connect.hku.hk.

‡Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore 637371. zhongjian.wang@ntu.edu.sg

§Department of Mathematics, University of California at Irvine, Irvine, CA 92697, USA. jack.xin@uci.edu

¶Corresponding author. Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, P.R.China. Materials Innovation Institute for Life Sciences and Energy (MILES), HKU-SIRI, Shenzhen, P.R. China. zhangzw@hku.hk.

Due to the nonlinear and potentially singular behavior of the KS equations, particularly in the presence of blow-up phenomena [27, 15, 1, 10], numerical methods have become essential tools for studying their solutions. Mesh-based methods, such as finite difference [4, 33, 8], finite element [31, 9, 32], and finite volume schemes [11, 5, 40], are among the most widely used approaches for solving the KS system. Furthermore, Li, Shu, and Yang [22] introduced a local discontinuous Galerkin method with an optimal convergence rate for the two-dimensional (2D) KS model before blow-up occurs. Liu, Wang, and Zhou [23] proposed a semi-discrete scheme for 2D KS equations based on symmetrization reformation, which avoids nonlinear solvers, and asymptotically preserves the quasi-static limit. Despite their success, challenges remain in ensuring stability, convergence, and effective handling of singularities, making the numerical study of the KS model an active and evolving field of research.

In addition to mesh-based methods, particle-based approaches have also been developed to address the challenges posed by the KS system, offering a complementary perspective. Stevens [36] developed an N-particle system and established its convergence for the fully parabolic case. Haškovec and Schmeiser [14] proposed a convergent regularized particle system for the 2D parabolic-elliptic KS model. Moreover, Godinho and Quininao [12] showed well-posedness results and the propagation of chaos property in a subcritical KS equation. Craig and Bertozzi [6] proved the convergence of a blob method for the related aggregation equation. Liu and Yang [24] introduced a random particle blob method with a mollified kernel for the parabolic-elliptic case, proving its convergence when the macroscopic mean field equation possesses a global weak solution [25].

In [39], we proposed a novel stochastic interacting particle-field (SIPF) algorithm for the fully parabolic KS system (1.1) in 3D. The SIPF method approximates KS solutions ρ as empirical measures of particles (see Eq.(2.1)) coupled with a smoother field variable c computed using the spectral method (see Eq.(2.2)). Instead of relying on history-dependent heat kernels, the algorithm employs an implicit Euler discretization and a one-step recursion based on Green's function of an elliptic operator. Numerical experiments demonstrate that the algorithm efficiently studies finite-time blowup in 3D with only dozens of Fourier modes. It handles multi-modal initial data and tracks complex evolutions, such as particle cluster merging and singularity formation.

Despite that [39] introduced the efficient algorithm, numerically observed its convergence, and showed its uniform stability, a rigorous convergence analysis remains to be accomplished. In this paper, we fill this gap by establishing the convergence estimate of the algorithm and validating the estimates by numerical results. Our main result, presented in Theorem 3.3, shows the convergence of the solution of the SIPF-r method $(\widetilde{\rho}, \widetilde{c})$ to the exact solution (ρ, c) under mild assumptions. Specifically, the 1-Wasserstein distance between the SIPF-r and exact density distributions, denoted as $\mathcal{W}_1(\widetilde{\rho}_{t_n}, \rho_{t_n})$, depends on the time step δt, Fourier mode H, the number of particles P, and the batch size R, scales as $\mathcal{O}\left(\frac{1}{H^2} + \frac{H^2}{\sqrt{P}}\delta t + \frac{\delta t}{\sqrt{R}} + H\delta t\right)$, plus higher-order terms. Similarly, the maximum error in the truncated Fourier coefficients of the computed chemical concentration, $\max_{\mathbf{j}\in\mathcal{H}} \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\|$, where $\mathcal{H}$ denotes the finite set of Fourier modes retained in the SIPF-r method, is governed by $\mathcal{O}\left(\frac{1}{H} + \frac{H}{\sqrt{P}} + \frac{H}{\sqrt{R}}\delta t + \frac{H^3}{\sqrt{P}}\delta t\right)$. These results illustrate the dependence of the method's accuracy on the discretization parameters δt, H, P, and R, highlighting their interplay in determining the overall error behavior. The proof of this result relies on several key lemmas presented in Section 3 that carefully quantify the single-step update errors of ρ and c between the SIPF-r method and the exact solution.

These lemmas also analyze how the errors introduced at each time step propagate and accumulate over time.

The error between the chemical concentration $\widetilde{c}$ in the SIPF-r method and the exact solution c primarily arises from the truncation error in its Fourier series representation and the time discretization error introduced by the implicit Euler method. We assume c has a certain level of regularity, ensuring that the truncation error in the Fourier series approximation is controllable and tends to zero as the number of Fourier modes increases. A thorough analysis of this aspect is beyond the scope of this article and is not discussed in detail here. The time discretization error can be formulated using the differences between the Fourier coefficients of $\widetilde{c}$ and c. Specifically, the error after the one-time step is constructed based on the Fourier coefficient errors of $\widetilde{c}$ and c from the previous step, combined with the expected L^2 norm of the error between the SIPF-r trajectories $\widetilde{X}_t$ and the exact solution trajectories X_t, denoted as $\mathbb{E}(\|\widetilde{X}_t - X_t\|_{L^2})$, from the preceding step.

At the numerical discretization level, the RBM [18, 17, 16, 3] is incorporated into the SIPF-r algorithm. This ensures the assumption that the particles are fully independent and identically distributed (i.i.d.), thereby effectively circumventing the need to address the propagation of chaos [25]. At each time step, small random batches of particles are selected with replacement for particle interactions. In the error estimate between $\widetilde{c}$ and c, leveraging the i.i.d. property, applying a generalization of the mean value theorem to complex-valued functions [26], and using Bernstein's inequality [7], we bound the probability that the empirical mean of the particles deviates from the expected value of their trajectories. This deviation accounts for the uncertainty described in Theorem 3.3. Numerical experiments in Section 4 further demonstrate that, with the introduction of the RBM, the numerical examples maintain a high level of accuracy.

The error between X_t and $\widetilde{X}_t$ is influenced by the gradient of c and $\widetilde{c}$, reflecting the sensitivity of the particle trajectories to the interaction potential. With Parseval's identity [20], we can establish a relationship between the error measured in the L^2 norm of the gradient difference $\|\nabla\widetilde{c} - \nabla c\|_{L^2}$ and the error in the Fourier coefficients of $\widetilde{c}$ and c. During each update step in the SIPF-r method, we can establish two coupled recursive inequalities. The first inequality relates the $\|\nabla\widetilde{c} - \nabla c\|_{L^2}$ error to $\mathbb{E}(\|\widetilde{X}_t - X_t\|_{L^2})$. Conversely, the second inequality involves $\mathbb{E}(\|\widetilde{X}_t - X_t\|_{L^2})$ and includes a term related to $\|\nabla\widetilde{c} - \nabla c\|_{L^2}$ in its single-step update, as illustrated in Eqs.(3.44)-(3.45). By substituting and decoupling the recursive inequalities, we derive a general bound for $\mathbb{E}(\|\widetilde{X}_t - X_t\|_{L^2})$ that depends only on errors from previous time steps. By considering the natural coupling $\gamma_t = \mathrm{Law}(\widetilde{X}_t, X_t)$ induced by shared initial conditions and Brownian paths, we can relate $\mathbb{E}(\|\widetilde{X}_t - X_t\|_{L^2})$ to the 1-Wasserstein distance $\mathcal{W}_1(\widetilde{\rho}_t, \rho_t)$ between the distributions $\widetilde{\rho}_t$ and ρ_t of $\widetilde{X}_t$ and X_t respectively. This allows us to reformulate our bound in terms of this more analytically tractable metric. Applying the discrete Gronwall inequality [21] to the derived inequality and combining it with the error estimate for the Fourier coefficients of $\widetilde{c}$, we establish a global error bound for $(\widetilde{\rho}, \widetilde{c})$, ultimately leading to the result stated in Theorem 3.3.

The rest of the paper is organized as follows. In Section 2, we review the SIPF method for solving the fully parabolic KS system and present the derivation of the SIPF-r method, which incorporates the RBM to compute the particle interaction. In Section 3, under certain assumptions, we provide a detailed convergence analysis of the SIPF-r method by breaking the proof into several lemmas. In Section 4, we present numerical results to validate the necessity of the assumptions, demonstrate

the accuracy of the SIPF-r method, and confirm the theoretical convergence rate derived in our analysis. Finally, the paper is concluded in Section 5.

2. Derivation of SIPF-r Method. In this section, we present the SIPF-r algorithm for solving the fully parabolic KS model. It is viable that we restrict the system (1.1) in a large domain $\Omega = [-L/2, L/2]^3$ and assume Dirichlet boundary condition for particle density ρ and Neumann boundary condition for chemical concentration c.

Throughout this section, we use the standard notation ρ, c, etc., to represent the exact solutions of the fully parabolic KS model. For the variables computed or approximated using the SIPF-r algorithm, we instead use the notations $\widetilde{\rho}$, $\widetilde{c}$, etc.

As a discrete algorithm, we assume that the temporal domain $[0, T]$ is partitioned by $\{t_n\}_{n=0:n_T}$ with $t_0 = 0$ and $t_{n_T} = T$. We approximate the density $\widetilde{\rho}$ at $t = t_n$ by empirical particles $\{\widetilde{X}^p_{t_n}\}_{p=1:P}$, i.e.,

$$\widetilde{\rho}_{t_n} \approx \frac{M_0}{P} \sum_{p=1}^{P} \delta(x - \widetilde{X}^p_{t_n}), \quad P \gg 1, \tag{2.1}$$

where M_0 is the conserved total mass (integral of ρ). For chemical concentration $\widetilde{c}$, we approximate by Fourier basis, namely, $\widetilde{c}(\mathbf{x}, t)$ has a series representation

$$\sum_{\mathbf{j}\in\mathcal{H}} \widetilde{\alpha}_{t;\mathbf{j}} \exp(i2\pi j_1 x_1/L) \exp(i2\pi j_2 x_2/L) \exp(i2\pi j_3 x_3/L), \tag{2.2}$$

where $\mathcal{H}$ denotes index set

$$\{(\mathbf{j}) \in \mathbb{N}^3 : |j_1|, |j_2|, |j_3| \le \frac{H}{2}\}, \tag{2.3}$$

and $i = \sqrt{-1}$. The exact solution $c(\mathbf{x}, t)$ can also be approximated by a truncated spatial Fourier series expansion as follows:

$$\sum_{\mathbf{j}\in\mathcal{H}} \alpha_{t;\mathbf{j}} \exp(i2\pi j_1 x_1/L) \exp(i2\pi j_2 x_2/L) \exp(i2\pi j_3 x_3/L). \tag{2.4}$$

Remark 2.1. The choice of the Fourier basis over Hermite polynomials for approximating chemical concentration is based on the fact that since the blow-up phenomenon is localized near the domain center, periodic boundary conditions effectively emulate an infinite spatial domain in this configuration. When the spatial localization of the singularity remains distant from domain boundaries, its interaction with these artificial edges becomes negligible.

Then at $t_0 = 0$, we generate P empirical samples $\{\widetilde{X}^p_0\}_{p=1:P}$ according to the initial condition of $\widetilde{\rho}_0$ and set up $\widetilde{\alpha}_{0;\mathbf{j}}$ by the Fourier series of $\widetilde{c}_0$. For ease of presenting our algorithm, with a slight abuse of notation, we use $\widetilde{\rho}_n = \frac{M_0}{P} \sum_{p=1}^{P} \delta(x - \widetilde{X}^p_n)$, and

$$\widetilde{c}_n = \sum_{\mathbf{j}\in\mathcal{H}} \widetilde{\alpha}_{n;\mathbf{j}} \exp(i2\pi j_1 x_1/L) \exp(i2\pi j_2 x_2/L) \exp(i2\pi j_3 x_3/L) \tag{2.5}$$

to represent density $\widetilde{\rho}$ and chemical concentration $\widetilde{c}$ at time t_n.

Considering the time-stepping system (1.1) from t_n to t_{n+1}, with $\widetilde{\rho}_n$ and $\widetilde{c}_{n-1}$ known, our algorithm, inspired by the operator splitting technique, consists of two sub-steps: updating chemical concentration $\widetilde{c}$ and updating organism density $\widetilde{\rho}$.

Updating chemical concentration $\widetilde{c}$. Let $\delta t = t_{n+1} - t_n > 0$ be the time step. We discretize the $\widetilde{c}$ equation of (1.1) in time by an implicit Euler scheme:

$$\epsilon\,(\widetilde{c}_n - \widetilde{c}_{n-1})/\delta t = (\Delta - \lambda^2)\,\widetilde{c}_n + \widetilde{\rho}_n. \tag{2.6}$$

From Eq.(2.6), we obtain the explicit formula for $\widetilde{c}_n$ as:

$$(\Delta - \lambda^2 - \epsilon/\delta t)\,\widetilde{c}_n = -\epsilon\,\widetilde{c}_{n-1}/\delta t - \widetilde{\rho}_n. \tag{2.7}$$

It follows that:

$$\widetilde{c}_n = \widetilde{c}(\mathbf{x}, t_n) = -\mathcal{K}_{\epsilon,\delta t} * (\epsilon\,\widetilde{c}_{n-1}/\delta t + \widetilde{\rho}_n) = -\mathcal{K}_{\epsilon,\delta t} * (\epsilon\,\widetilde{c}(\mathbf{x}, t_{n-1})/\delta t + \widetilde{\rho}(\mathbf{x}, t_n)), \tag{2.8}$$

where $\mathcal{K}_{\epsilon,\delta t}$ is the Green's function of the operator $\Delta - \lambda^2 - \epsilon/\delta t$ and $*$ represents an approximation of spatial convolution, which is not exactly in the continuous setup, as $\widetilde{c}$ is computed using truncated Fourier basis functions and $\widetilde{\rho}$ is given by a discrete particle representation. Unless otherwise stated, all subsequent norms $\|\cdot\|$ will refer to the L^2 norms. In case of $\mathbb{R}^3$, the Green's function $\mathcal{K}_{\epsilon,\delta t}$ reads as follows

$$\mathcal{K}_{\epsilon,\delta t} = \mathcal{K}_{\epsilon,\delta t}(\mathbf{x}) = -\frac{\exp\{-\beta\|\mathbf{x}\|\}}{4\pi\|\mathbf{x}\|}, \quad \beta^2 = \lambda^2 + \epsilon/\delta t. \tag{2.9}$$

Green's function admits a closed-form Fourier transform,

$$\mathcal{F}\mathcal{K}_{\epsilon,\delta t}(\omega) = -\frac{1}{\|\omega\|^2 + \beta^2}. \tag{2.10}$$

For the term $-\mathcal{K}_{\epsilon,\delta t} * \widetilde{c}_{n-1}$ in Eq.(2.8), by Eq.(2.10) it is equivalent to modify Fourier coefficients $\widetilde{\alpha}_{\mathbf{j}}$ to $\widetilde{\alpha}_{\mathbf{j}}/(4\pi^2 j_1^2/L^2 + 4\pi^2 j_2^2/L^2 + 4\pi^2 j_3^2/L^2 + \beta^2)$.

For the second term $\mathcal{K}_{\epsilon,\delta t} * \widetilde{\rho}$, we first approximate $\mathcal{K}_{\epsilon,\delta t}$ with cos series expansion, then according to the particle representation of $\widetilde{\rho}$ in Eq.(2.1),

$$(\mathcal{K}_{\epsilon,\delta t} * \widetilde{\rho})_{\mathbf{j}} \approx \frac{M_0}{P}\sum_{p=1}^{P} \frac{\exp(-i2\pi j_1 \widetilde{X}^p_{n;1}/L - i2\pi j_2 \widetilde{X}^p_{n;2}/L - i2\pi j_3 \widetilde{X}^p_{n;3}/L)(-1)^{j_1+j_2+j_3}}{4\pi^2 j_1^2/L^2 + 4\pi^2 j_2^2/L^2 + 4\pi^2 j_3^2/L^2 + \beta^2}. \tag{2.11}$$

Finally, we summarize the one-step update of the Fourier coefficients of chemical concentration $\widetilde{c}$ in Alg.2.1, which follows the same procedure as in the original SIPF method.

Updating density of active particles $\widetilde{\rho}$. In the one-step update of density $\widetilde{\rho}_n$ represented by particles $\{\widetilde{X}^p_n\}_{p=1:P}$, we apply Euler-Maruyama scheme to solve the SDE

$$\widetilde{X}^p_{n+1} = \widetilde{X}^p_n + \chi\nabla_{\mathbf{x}}\widetilde{c}(\widetilde{X}^p_n, t_n)\delta t + \sqrt{2\,\mu\,\delta t}\,N^p_n, \tag{2.12}$$

where N^p_n's are i.i.d. standard normal distributions with respect to the Brownian paths in the SDE formulation. For $n > 1$, substituting Eq.(2.8) in Eq.(2.12) gives:

$$\widetilde{X}^p_{n+1} = \widetilde{X}^p_n - \chi\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t} * (\epsilon\,\widetilde{c}_{n-1}(\mathbf{x})/\delta t + \widetilde{\rho}_n(\mathbf{x}))|_{\mathbf{x}=\widetilde{X}^p_n}\delta t + \sqrt{2\,\mu\,\delta t}\,N^p_n, \tag{2.13}$$

from which $\widetilde{\rho}_{n+1}(\mathbf{x})$ is constructed via Eq.(2.1).

In this particle formulation, the computation of the spatial convolution differs slightly from that in the update of $\widetilde{c}$ (i.e., Eq.(2.8)).

Algorithm 2.1 One step update of chemical concentration in SIPF-r

Require: Distribution $\widetilde{\rho}_n$ represented by empirical samples $\widetilde{X}_n$,
initial concentration $\widetilde{c}_{n-1}$ represented by Fourier coefficients $\widetilde{\alpha}_{n-1}$.

for $(\mathbf{j}) \in \mathcal{H}$ **do**
  $\widetilde{\alpha}_{n;\mathbf{j}} \leftarrow \dfrac{\epsilon \widetilde{\alpha}_{n-1;\mathbf{j}}}{\delta t(4\pi^2 j_1^2/L^2 + 4\pi^2 j_2^2/L^2 + 4\pi^2 j_3^2/L^2 + \beta^2)}$
  $F_{\mathbf{j}} \leftarrow 0$
  for $p = 1$ **to** P **do**
    $F_{\mathbf{j}} \leftarrow F_{\mathbf{j}} + \exp(-i2\pi j_1 \widetilde{X}^p_{n;1}/L - i2\pi j_2 \widetilde{X}^p_{n;2}/L - i2\pi j_3 \widetilde{X}^p_{n;3}/L)$
  end for
  $F_{\mathbf{j}} \leftarrow F_{\mathbf{j}} \cdot \dfrac{(-1)^{j_1+j_2+j_3}}{4\pi^2 j_1^2/L^2 + 4\pi^2 j_2^2/L^2 + 4\pi^2 j_3^2/L^2 + \beta^2} \cdot \dfrac{M_0}{P}$
end for
$\widetilde{\alpha}_n \leftarrow \widetilde{\alpha}_n - F$

Ensure: Updated chemical concentration field from $\widetilde{c}_{n-1}$ to $\widetilde{c}_n$ via $\widetilde{\alpha}_n$.

For $\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{c}_{n-1}(\widetilde{X}^p_n)$, to avoid the singular points of $\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}$, we evaluate the integral with the quadrature points that are away from 0. Precisely, denote the standard quadrature point in Ω with

$$x_{\mathbf{j}} = (j_1\, L/H, j_2\, L/H, j_3\, L/H), \tag{2.14}$$

where j, m, l are integers ranging from $-H/2$ to $H/2-1$. When computing $\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{c}_{n-1}(\widetilde{X}^p_n)$, we evaluate $\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}$ at $\{\widetilde{X}^p_n + \bar{X}^p_n - x_{\mathbf{j}}\}_{\mathbf{j}}$ where a small spatial shift $\bar{X}^p_n = \frac{L}{2H} + \lfloor \frac{\widetilde{X}^p_n}{L/H} \rfloor \frac{L}{H} - \widetilde{X}^p_n$ and $\widetilde{c}$ at $\{x_{\mathbf{j}} - \bar{X}^p_n\}_{\mathbf{j}}$ correspondingly. The latter one is computed by inverse Fourier transform of the shifted coefficients, with $\widetilde{\alpha}_{\mathbf{j}}$ modified to $\widetilde{\alpha}_{\mathbf{j}} \exp(-i2\pi j_1 \bar{X}^p_{n;1}/L - i2\pi j_2 \bar{X}^p_{n;2}/L - i2\pi j_3 \bar{X}^p_{n;3}/L)$ where $(\bar{X}^p_{n;i})$ denotes the i-th component of $\bar{X}^p_n$.

Motivated by mini-batch sampling [13, 35, 37, 38] and random batch method (RBM) [18, 17, 3, 16], for each particle $\widetilde{X}^p_n$, we choose a small batch C_p with size R randomly with replacement. We just interact $\widetilde{X}^p_n$ with particles within this batch, i.e. approximate $\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{\rho}(\widetilde{X}^p_n, t_n)$ using $\sum_{s \in C_p, s \neq p} \frac{\chi M_0 \delta t}{R} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}^p_n - \widetilde{X}^s_n)$.

We summarize the one-step update (for $n > 1$) of the density in SIPF as in Alg.2.2.

Combining Eq.(2.8) and Eq.(2.13), we conclude that the recursion from $(\{\widetilde{X}^p_n\}_{p=1:P}, \widetilde{\rho}_n(\mathbf{x}), \widetilde{c}_{n-1}(\mathbf{x}))$ to $(\{\widetilde{X}^p_{n+1}\}_{p=1:P}, \widetilde{\rho}_{n+1}(\mathbf{x}), \widetilde{c}_n(\mathbf{x}))$ is complete. We summarize the SIPF-r method in the following Algorithm 2.3.

Particle-wise Independence due to RBM. In the above derivation, $\{\widetilde{X}^p_n\}_{p=1:P}$ are i.i.d. samples with distribution $\widetilde{\rho}_n$ and independent of $\widetilde{c}_{n-1}$. The one-step trajectories follow the discrete-time rule:

$$\widetilde{X}_{t_{n+1}} = \widetilde{X}_{t_n} + \chi \nabla \widetilde{c}(\widetilde{X}_{t_n}, t_n)\delta t + \int_{t_n}^{t_{n+1}} \sqrt{2\mu}\, dW_s, \tag{2.15}$$

where $\nabla \widetilde{c}$ is computed via Eq.(2.8), and W_s denotes the Brownian motion. It is worth noting that, for the updated position p-th particle $\widetilde{X}^p_{n+1}$ by Eq.(2.12), the interaction term, $\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{\rho}(\widetilde{X}^p_n, t_n)$ is computed by $\sum_{s \in C_p, s \neq p} \frac{\chi M_0 \delta t}{R} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}^p_n - \widetilde{X}^s_n)$, where

Algorithm 2.2 One step update of density in SIPF-r

Require: Distribution $\widetilde{\rho}_n$ represented by empirical samples $\widetilde{X}_n$,
Concentration $\widetilde{c}_{n-1}$ represented by Fourier coefficients $\widetilde{\alpha}_{n-1}$.

for $p = 1$ **to** P **do**
  $\widetilde{X}^p_{n+1} \leftarrow \widetilde{X}^p_{n+1} + \sqrt{2\mu\delta t}N$ {N is a standard normal random variable}
  $C_p \leftarrow$ random subset of $\{1, \dots, P\}$ with replacement, size R
  $\widetilde{X}^p_{n+1} \leftarrow \widetilde{X}^p_{n+1} - \sum_{s\in C_p} \frac{\chi M_0 \delta t}{R} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}^p_n - \widetilde{X}^s_n)$
  $\bar{X}^p_n \leftarrow \frac{L}{2H} + \lceil \frac{\widetilde{X}^p_n}{L/H} \rceil \frac{L}{H} - \widetilde{X}^p_n$
  for $(\mathbf{j}) \in \mathcal{H}$ **do**
    $F_{\mathbf{j}} \leftarrow \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}^p_n + \bar{X}^p_n - x_{\mathbf{j}})$ {$x_{\mathbf{j}}$ from Eq.(2.14)}
    $G_{\mathbf{j}} \leftarrow \alpha_{\mathbf{j}} \exp(-i2\pi j_1 \bar{X}^p_{n;1}/L - i2\pi j_2 \bar{X}^p_{n;2}/L - i2\pi j_3 \bar{X}^p_{n;3}/L)$
  end for
  $\check{G} \leftarrow \mathrm{iFFT}(G)$
  $\widetilde{X}^p_{n+1} \leftarrow \widetilde{X}^p_{n+1} - \epsilon\chi(F, \check{G}) \frac{L^3}{H^3}$ {$(\cdot,\cdot)\frac{L^3}{H^3}$ denotes an inner product corresponding to $L^2(\Omega)$ quadrature}
end for

Ensure: Updated distribution $\widetilde{\rho}_{n+1}$ represented by $\widetilde{X}_{n+1}$.

Algorithm 2.3 Stochastic Interacting Particle-Field Method

Require: Initial distribution ρ_0, initial concentration c_0.
Generate P i.i.d. samples following distribution ρ_0: $X^1, X^2, \dots, X^P$.
for $p = 1$ **to** P **do**
  Compute $\widetilde{X}^p_1$ by Eq.(2.12), with $c_{-1} = c_0$.
end for
Compute $\widetilde{c}_1$ by Algorithm 2.1 with c_0 and $\widetilde{\rho}_1 = \sum_{p=1}^P \frac{M_0}{P} \delta_{\widetilde{X}^p_1}$.
for $n = 2$ **to** $N = T/\delta t$ **do**
  Compute $\widetilde{X}_n$ by Algorithm 2.2 with $\widetilde{\rho}_{n-1}$ and $\widetilde{c}_{n-2}$.
  Compute $\widetilde{c}_n$ by Algorithm 2.1 with $\widetilde{c}_{n-1}$ and $\widetilde{\rho}_n = \sum_{p=1}^P \frac{M_0}{P} \delta_{\widetilde{X}^p_n}$.
end for

Ensure: Final particle distribution $\widetilde{\rho}_N$ and concentration field $\widetilde{c}_N$.

the selection of C_p is independent of $\widetilde{X}^p_n$ and hence $\{\widetilde{X}^s_n\}_{s\in C_p}$ can be viewed as i.i.d samples of $\widetilde{\rho}_n$ independent of $\widetilde{c}_{n-1}$ and $\widetilde{X}^p_n$. Together with the independent Brownian motion term W^p_s, we can deduce the independency of $\{\widetilde{X}^p_{n+1}\}_{p=1:P}$.

Correspondingly, we denote the exact dynamics of the system by X_t, a $\rho(\cdot,t)$-distributed random variable evolving continuously in time:

$$X_t = X_{t_0} + \chi \int_{t_0}^t \nabla c(X_s, s)\, ds + \int_{t_0}^t \sqrt{2\mu}\, dW_s, \quad X_{t_0} = \widetilde{X}_{t_0}, \tag{2.16}$$

where $c(\cdot, s)$ is the exact concentration field, and the integral describes how the gradient evolves in continuous time. Both processes share the same Brownian motion W_s, indicating that both processes are driven by the same source of randomness.

3. L^2 Convergence of SIPF-r method to smooth solutions. We now prove the convergence of the SIPF-r method to classical solutions of the 3D parabolic-parabolic Keller-Segel equations. To ensure the validity of the following analysis, we

introduce a set of assumptions that impose structure on the concentration fields and their gradients.

ASSUMPTION 1. *We assume the approximation errors of particles and the gradient of the chemical concentration at any finite time t are bounded. Specifically, there exist constants $M_1, M_2 > 0$ such that for all $t \in [0,T]$ and $\mathbf{x} \in \mathbb{R}^3$,*

$$\|\widetilde{X}_t - X_t\| \le M_1, \tag{3.1}$$

$$\|\nabla\widetilde{c}(\mathbf{x},t) - \nabla c(\mathbf{x},t)\| \le M_2, \quad \forall \mathbf{x}, t. \tag{3.2}$$

Remark 3.1. The boundedness condition in Eq.(3.1) can be achieved by Eqs. (2.15)-(2.16) and Assumption 2(c). Eq.(3.2) follows immediately from the uniform bound in Assumption 2(c). It is important to note that this assumption only requires the errors to be bounded and does not demand them to converge to zero. The convergence of these errors to zero will be demonstrated later in subsequent theorem and proof.

ASSUMPTION 2. *Suppose both $\nabla\widetilde{c}$ and ∇c satisfy Lipschitz continuity conditions in space and time, along with regularity and boundedness properties as follows:*

(a) (Spatial Lipschitz Continuity) *There exists a constant $K > 0$, depending on the regularity of $\nabla\widetilde{c}$ and ∇c, as well as the parameters ϵ and λ in the system (1.1), such that for all $t \in [0,T]$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$,*

$$\left(\|\nabla\widetilde{c}(\mathbf{x},t) - \nabla\widetilde{c}(\mathbf{y},t)\|, \|\nabla c(\mathbf{x},t) - \nabla c(\mathbf{y},t)\|\right) \le K\|\mathbf{x} - \mathbf{y}\|.$$

This implies that the second derivatives (Hessian entries) $\nabla^2\widetilde{c}(\mathbf{x},t)$ exist almost everywhere and satisfy:

$$\sup_{\mathbf{x}\in\mathbb{R}^3, t\in[0,T]} \left(\|\nabla^2\widetilde{c}(\mathbf{x},t)\|\right) \le K.$$

(b) (Temporal Lipschitz Continuity) *There exists a constant $K_1 > 0$, depending on the regularity of ∇c and the parameters ϵ and λ in the system (1.1), such that for any $t_1, t_2 \in [0,T]$ and $\mathbf{x} \in \mathbb{R}^3$,*

$$\|\nabla c(\mathbf{x},t_1) - \nabla c(\mathbf{x},t_2)\| \le K_1|t_1 - t_2|.$$

(c) (Uniform Boundedness) *There exists a constant $M_3 > 0$, depending on the regularity of ∇c and the parameters ϵ and λ in the system (1.1), such that for all $t \in [0,T]$ and $\mathbf{x} \in \mathbb{R}^3$:*

$$\max\left(\|\nabla c(\mathbf{x},t)\|, \|\nabla\widetilde{c}(\mathbf{x},t)\|\right) \le M_3.$$

ASSUMPTION 3 (CFL-like Condition). *In the SIPF-r algorithm, we assume that the discrete time interval δt approaches 0 faster than the square of the Fourier mode H^2 diverges to positive infinity. Additionally, the number of particles P is assumed to grow sufficiently fast, ensuring that it outpaces the H^2 to infinity.*
More formally, as $\delta t \to 0$, $H \to \infty$, and $P \to \infty$, there exists parameters κ, ν such that:

$$\kappa := H \cdot \sqrt{\delta t},\ \kappa \to 0, \qquad \nu := \frac{H}{\sqrt{P}},\ \nu \to 0. \tag{3.3}$$

These assumptions guarantee that the gradients of the exact concentration field c and the approximated concentration field $\widetilde{c}$ exhibit sufficient regularity, boundedness,

and stability in both space and time. They establish the necessary framework to rigorously compare the SIPF-r approximation $\widetilde{X}_{t_n}$ with the exact solution X_{t_n} while ensuring the stability and convergence of the particle system.

Remark 3.2. The assumptions above are technical and provide the foundation for analyzing the convergence and stability of the SIPF-r algorithm. Their validity will be supported by numerical experiments. Specifically, we will provide detailed numerical results to demonstrate that the approximation errors, regularity, and boundedness conditions held in practice under realistic parameter settings. These experiments will confirm the assumptions held in practice, ensuring the robustness of our theoretical results.

We now state our main theorem, quantifying the convergence of the SIPF-r method.

THEOREM 3.3. *Suppose that the exact solutions and the solutions of the SIPF-r method satisfy Assumptions 1, 2, 3 in* $\mathbb{R}^3$*, consider the SIPF-r method with* H *being the Fourier mode,* P *being the number of particles,* R *being the batch size, and* δt *being the uniform time step. Then the quantities* $(\widetilde{\rho}, \widetilde{c})$*, which comprise the SIPF-r method, exist on discrete time steps* $t_n = n\delta t$ *for* $n = 0, 1, \ldots, \frac{T}{\delta t}$*, and satisfy the following with high probability:*
For $\forall n \in \{0, 1, \ldots, \frac{T}{\delta t}\}$*, the* 1*-Wasserstein distance (defined in Eq.(3.49)) between* $\widetilde{\rho}_{t_n}$ *and* ρ_{t_n} *satisfies:* $\mathcal{W}_1(\widetilde{\rho}_{t_n}, \rho_{t_n})$ *is* $\mathcal{O}\left(\frac{1}{H^2} + \frac{H^2}{\sqrt{P}}\delta t + \frac{\delta t}{\sqrt{R}} + H\delta t\right)$*, and the maximum error in the truncated Fourier coefficients of* $\widetilde{c}_{t_n}$ *and* c_{t_n} *satisfies:* $\max_{\mathbf{j}\in\mathcal{H}} \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\|$ *is* $\mathcal{O}\left(\frac{1}{H} + \frac{H}{\sqrt{P}} + \frac{H}{\sqrt{R}}\delta t + \frac{H^3}{\sqrt{P}}\delta t\right)$.
More specifically, for $\forall n \in \{0, 1, \ldots, \frac{T}{\delta t}\}$*, the errors are bounded with high probability by:*

$$\mathcal{W}_1(\widetilde{\rho}_{t_n}, \rho_{t_n}) \le \left(S_0(\frac{L}{H})^2 + S_1\delta t + \left(\frac{S_2 \cdot H^2}{\mathcal{O}(\sqrt{P})} + S_3 H + \frac{1}{\mathcal{O}(\sqrt{R})}\right)\delta t + \mathcal{O}(\delta t^2)\right) \cdot \exp(1 + S_4\delta t + S_2 H^2 \delta t),$$

$$\max_{\mathbf{j}\in\mathcal{H}} \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| \le \left(\frac{S_7}{H} + \left(S_8 H + S_9 H^2 + S_5\frac{H}{\mathcal{O}(\sqrt{R})} + S_{10}\frac{H^3}{\mathcal{O}(\sqrt{P})}\right)\delta t\right) \cdot \exp(1 + S_4\delta t + S_2 H^2\delta t) + S_5\frac{H}{\mathcal{O}(\sqrt{P})} + S_6 H\delta t, \tag{3.4}$$

where S_i*,* $i = 0, \ldots, 10$*, are constants specified in Eqs.(3.51)-(3.53), and* L *is the characteristic domain size.*

A direct consequence of the Theorem 3.3 reads, as $H, P \to \infty$ and $\delta t \to 0$, we have both $\mathcal{W}_1(\widetilde{\rho}_{t_n}, \rho_{t_n})$ and $\max_{\mathbf{j}\in\mathcal{H}} \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\|$ converge to 0 with high probability.

Combining the above theorem, to simplify the form, we obtain the following corollary:

COROLLARY 3.4. *Under the conditions of Theorem 3.3, assume the scaling relationships* $H = \Theta(P^{1/6})$ *and* $\delta t = \Theta(H^{-3}) = \Theta(P^{-1/2})$*. Then, the solutions* $(\widetilde{\rho}, \widetilde{c})$ *of the SIPF-r method satisfy the following simplified high-probability error estimates:*
For $\forall n \in \{0, 1, \ldots, \frac{T}{\delta t}\}$*,* $\mathcal{W}_1(\widetilde{\rho}_{t_n}, \rho_{t_n})$ *is* $\mathcal{O}\left(\delta t^{\frac{2}{3}}\right)$*,* $\max_{\mathbf{j}\in\mathcal{H}} \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\|$ *is* $\mathcal{O}\left(\delta t^{\frac{1}{3}}\right)$.

The result of Theorem 3.3 relies on the following lemmas concerning the change in single-step update error of the SIPF-r method and the complete proof of Theorem 3.3 is postponed to the end of this subsection.

The exact solution of chemical concentration c comes from solving a parabolic equation, which is no longer Markovian. At time $t > 0$, the solution of ρ in $[0, t]$ has to be involved in the representation of c, namely,

$$c(\cdot,t) = e^{-\frac{\lambda^2}{\epsilon}t}e^{t\Delta}c(\cdot,0) + \frac{1}{\epsilon}\int_0^t e^{\frac{\lambda^2}{\epsilon}(s-t)}e^{(t-s)\Delta}\rho(\cdot,s)\,ds, \tag{3.5}$$

where the heat semigroup operator $e^{t\Delta}$ is defined by

$$(e^{t\Delta}f)(\mathbf{x},t) := \int e^{-\frac{\epsilon\|\mathbf{x}-\mathbf{y}\|^2}{4t}}(\frac{\epsilon}{4\pi t})^{3/2}f(\mathbf{y})\,d\mathbf{y}.$$

From Eqs.(2.2)-(2.4), the error between $\widetilde{c}$ and c can be decomposed into two components: the error in their Fourier coefficients and the truncation error of c. As the Fourier mode H and domain size L tend to infinity, and due to the smoothness of c, the truncation error becomes negligible and can be omitted from the analysis. We now focus on the error analysis between the Fourier coefficients $\widetilde{\alpha}_{\mathbf{j}}$ and $\alpha_{\mathbf{j}}$ of $\widetilde{c}$ and c, as presented in the following lemma.

LEMMA 3.5. *For $\forall n \in \mathbb{N}_+$ and $\forall \mathbf{j} \in \mathcal{H}$ (the same index set as in Eq.(2.3)), under Assumption 1, the following inequality holds with high probability:*

$$\begin{aligned}\|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| \leq & \|\widetilde{\alpha}_{t_{n-1};\mathbf{j}} - \alpha_{t_{n-1};\mathbf{j}}\| + \frac{(\sqrt{2}M_0\|\omega_{\mathbf{j}}\|+1)}{\epsilon\cdot\mathcal{O}(\sqrt{P})}\delta t + \mathcal{O}(\frac{\delta t^2}{\|\omega_{\mathbf{j}}\|^4}) \\ & + C_1\|\omega_{\mathbf{j}}\|\delta t^2 + \frac{\sqrt{2}M_0\|\omega_{\mathbf{j}}\|}{\epsilon}\delta t\mathbb{E}[\|\widetilde{X}_{t_n} - X_{t_n}\|],\end{aligned} \tag{3.6}$$

where C_1 is a constant, and $t_n = n\delta t$.

Proof. We write the frequency $\omega_{\mathbf{j}} = \left(\frac{2\pi j_1}{L}, \frac{2\pi j_2}{L}, \frac{2\pi j_3}{L}\right)$. According to Section 2, in Eq.(2.8), the term $-\mathcal{K}_{\epsilon,\delta t} * \frac{\epsilon\widetilde{c}(\mathbf{x},t_{n-1})}{\delta t}$ modifies the Fourier coefficients $\widetilde{\alpha}_{t_{n-1};\mathbf{j}}$ to:

$$\frac{\epsilon\widetilde{\alpha}_{t_{n-1};\mathbf{j}}}{\delta t\left(\|\omega_{\mathbf{j}}\|^2+\beta^2\right)},$$

where $\beta^2 = \lambda^2 + \epsilon/\delta t$. Similarly, for the term $-\mathcal{K}_{\epsilon,\delta t} * \widetilde{\rho}(\mathbf{x}, t_n)$, it is modified as:

$$\frac{1}{1+Z_{\mathbf{j}}}\cdot\frac{\delta t}{\epsilon}\cdot\mathcal{F}_{\mathbf{j}}[\widetilde{\rho}(\mathbf{x},s)],$$

where $\mathcal{F}_{\mathbf{j}}[\widetilde{\rho}(\mathbf{x},t_n)] = \frac{M_0}{P}\sum_{p=1}^{P}\frac{e^{-i\omega_{\mathbf{j}}\cdot\widetilde{X}^p_{t_n}}}{1+Z_{\mathbf{j}}}$ represents the Fourier coefficient of $\widetilde{\rho}(\mathbf{x},t_n)$ at the frequency $\omega_{\mathbf{j}}$.

For the exact solution c, when updating from t_{n-1} to t_n, the first term of Eq.(3.5) modifies the Fourier coefficients $\alpha_{t_{n-1};\mathbf{j}}$ as:

$$\alpha_{t_{n-1};\mathbf{j}}\cdot\exp\left(-\left(\|\omega_{\mathbf{j}}\|^2+\lambda^2\right)\cdot\frac{\delta t}{\epsilon}\right).$$

The second term modifies as:

$$\frac{1}{\epsilon}\int_{t_{n-1}}^{t_n} e^{-\left(\lambda^2+\|\omega_{\mathbf{j}}\|^2\right)\frac{t_n-s}{\epsilon}}\cdot\mathcal{F}_{\mathbf{j}}[\rho(\mathbf{x},s)]\,ds.$$

We write:

$$Z_{\mathbf{j}} = \left(\|\omega_{\mathbf{j}}\|^2 + \lambda^2\right) \cdot \frac{\delta t}{\epsilon}.$$

As $\delta t \to 0$, $Z_{\mathbf{j}} \to 0$. Using Taylor Expansion and the triangle inequality, we decompose the error into two terms:

$$I_1 := \|\alpha_{t_{n-1};\mathbf{j}} \cdot e^{-Z_{\mathbf{j}}} - \widetilde{\alpha}_{t_{n-1};\mathbf{j}} \cdot \frac{1}{Z_{\mathbf{j}}+1}\|,$$

$$I_2 := \left\| \frac{1}{\epsilon} \int_{t_{n-1}}^{t_n} e^{-\left(\lambda^2 + \|\omega_{\mathbf{j}}\|^2\right)\frac{t_n - s}{\epsilon}} \cdot \mathcal{F}_{\mathbf{j}}[\rho(\mathbf{x}, s)]\, ds - \frac{1}{1+Z_{\mathbf{j}}} \cdot \frac{\delta t}{\epsilon} \cdot \mathcal{F}_{\mathbf{j}}[\widetilde{\rho}(\mathbf{x}, s)] \right\|.$$

Applying Taylor Expansion and triangle inequality, we obtain

$$\begin{aligned} I_1 \leq & \|\alpha_{t_{n-1};\mathbf{j}} \cdot (e^{-Z_{\mathbf{j}}} - \frac{1}{Z_{\mathbf{j}}+1})\| + \|(\alpha_{t_{n-1};\mathbf{j}} - \widetilde{\alpha}_{t_{n-1};\mathbf{j}}) \cdot \frac{1}{Z_{\mathbf{j}}+1}\| \\ \leq & \frac{1}{2} Z_{\mathbf{j}}^2 \|\alpha_{t_{n-1};\mathbf{j}}\| + \|\alpha_{t_{n-1};\mathbf{j}} - \widetilde{\alpha}_{t_{n-1};\mathbf{j}}\|. \end{aligned} \tag{3.7}$$

Now we turn to I_2. To this end, we first list a generalization of the mean value theorem to complex-valued functions:

Let G be an open subset of $\mathbb{R}^n$, and let $f : G \to \mathbb{C}$ be a holomorphic function. Fix points $\mathbf{x}, \mathbf{y} \in G$ such that the line segment connecting $\mathbf{x}$ and $\mathbf{y}$ lies entirely within G. The there exists $c_1, c_2 \in (0,1)$ such that:

$$f(\mathbf{y}) - f(\mathbf{x}) = \mathrm{Re}\left(\nabla f((1-c_1)\mathbf{x} + c_1\mathbf{y})(\mathbf{y}-\mathbf{x})\right) + i\,\mathrm{Im}\left(\nabla f((1-c_2)\mathbf{x} + c_2\mathbf{y})(\mathbf{y}-\mathbf{x})\right). \tag{3.8}$$

The proof of (3.8) is direct. First, we define the function

$$g(t) = f((1-t)\mathbf{x} + t\mathbf{y}), \quad t \in [0,1].$$

Then g is also a holomorphic function. Then, by mean value theorem, there exist points $c_1, c_2 \in (0,1)$ such that,

$$\mathrm{Re}(g'(c_1)) = \mathrm{Re}(g(1) - g(0)),$$

$$\mathrm{Im}(g'(c_2)) = \mathrm{Im}(g(1) - g(0)),$$

which implies Eq.(3.8). Applying this result to to $f(\mathbf{x}) = e^{-i\omega_{\mathbf{j}}\mathbf{x}}$, we obtain:

$$\begin{aligned} & \|e^{-i\omega_{\mathbf{j}} \cdot \widetilde{X}^p_{t_n}} - e^{-i\omega_{\mathbf{j}} \cdot X^p_{t_n}}\| \\ \leq & \|\omega_{\mathbf{j}} \cdot \sin(\omega_{\mathbf{j}}((1-c_1)\widetilde{X}^p_{t_n} + c_1 X^p_{t_n})) \\ & + i\omega_{\mathbf{j}} \cdot \cos(\omega_{\mathbf{j}}((1-c_2)\widetilde{X}^p_{t_n} + c_2 X^p_{t_n}))\| \cdot \|\widetilde{X}^p_{t_n} - X^p_{t_n}\| \\ \leq & \sqrt{2}\|\omega_{\mathbf{j}}\| \cdot \|\widetilde{X}^p_{t_n} - X^p_{t_n}\|. \end{aligned} \tag{3.9}$$

Using the triangle inequality, we get:

$$
\begin{aligned}
I_2 \leq & \underbrace{\left\| \frac{1}{\epsilon} \int_{t_{n-1}}^{t_n} e^{-Z_{\mathbf{j}} \cdot (t_n - s)/\delta t} (\mathcal{F}_{\mathbf{j}}[\rho(\mathbf{x}, s)] - \mathcal{F}_{\mathbf{j}}[\rho_P(\mathbf{x}, s)])\, ds \right\|}_{I_{2,1}} \\
& + \underbrace{\left\| \frac{1}{\epsilon} \int_{t_{n-1}}^{t_n} e^{-Z_{\mathbf{j}} \cdot (t_n - s)/\delta t} (\mathcal{F}_{\mathbf{j}}[\rho_P(\mathbf{x}, s)] - \mathcal{F}_{\mathbf{j}}[\rho(\mathbf{x}, t_n)])\, ds \right\|}_{I_{2,2}} \\
& + \underbrace{\left\| \frac{\delta t}{\epsilon} \left(\frac{1 - e^{-Z_{\mathbf{j}}}}{Z_{\mathbf{j}}} - \frac{1}{1 + Z_{\mathbf{j}}} \right) \cdot \mathcal{F}_{\mathbf{j}}[\rho(\mathbf{x}, t_n)] \right\|}_{I_{2,3}} \\
& + \underbrace{\left\| \frac{\delta t}{\epsilon} \frac{1}{1 + Z_{\mathbf{j}}} \left(\mathcal{F}_{\mathbf{j}}[\rho(\mathbf{x}, t_n)] - \mathcal{F}_{\mathbf{j}}[\widetilde{\rho}(\mathbf{x}, t_n)] \right) \right\|}_{I_{2,4}}, \qquad (3.10)
\end{aligned}
$$

where $\rho_P(\mathbf{x}, t) = \frac{M_0}{P} \sum_{p=1}^{P} \delta(\mathbf{x} - X_t^p)$. By Glivenko-Cantelli's Theorem, ρ_P tends weakly to ρ as $P \to \infty$.

By the Central Limit Theorem, for the empirical measure ρ_P with i.i.d. samples X_t^p drawn from the distribution $\rho(\cdot, t)$, the difference in Fourier coefficients satisfies:

$$
\|\mathcal{F}_{\mathbf{j}}[\rho(\cdot, s)] - \mathcal{F}_{\mathbf{j}}[\rho_P(\cdot, s)]\| = M_0 \cdot \sqrt{\frac{\mathrm{Var}\left(e^{-i\omega_{\mathbf{j}} \cdot X_s^p}\right)}{P}} = \mathcal{O}\left(\frac{1}{\sqrt{P}}\right).
$$

Now, we can rewrite the expression:

$$
\begin{aligned}
I_{2,1} \leq & \left\| \frac{1}{\epsilon} \int_{t_{n-1}}^{t_n} e^{-Z_{\mathbf{j}} \cdot (t_n - s)/\delta t} \mathcal{O}\left(\frac{1}{\sqrt{P}}\right) ds \right\| \\
\leq & \frac{\delta t}{\epsilon} \mathcal{O}\left(\frac{1}{\sqrt{P}}\right). \qquad (3.11)
\end{aligned}
$$

According to Eqs.(2.16)-(3.8) and the uniform boundness property of ∇c in Assumption 2,

$$
\begin{aligned}
I_{2,2} &\leq \left\| \frac{1}{\epsilon} \frac{M_0}{P} \int_{t_{n-1}}^{t_n} e^{-Z_{\mathbf{j}} \cdot (t_n - s)/\delta t} \sum_{i=1}^{P} (\sqrt{2} \|\omega_{\mathbf{j}}\| \|X_{t_n}^i - X_s^i\|)\, ds \right\| \\
&\leq \left\| \frac{1}{\epsilon} \frac{M_0}{P} \int_{t_{n-1}}^{t_n} e^{-Z_{\mathbf{j}} \cdot (t_n - s)/\delta t} \sum_{i=1}^{P} \left(\sqrt{2} \|\omega_{\mathbf{j}}\| \| \int_s^{t_n} \nabla c(X_u^i, u)\, du \| \right) ds \right\| \\
&\leq \frac{\sqrt{2}}{\epsilon} M_0 M_3 \|\omega_{\mathbf{j}}\| \left| \int_{t_{n-1}}^{t_n} e^{-Z_{\mathbf{j}} \cdot (t_n - s)/\delta t} (t_n - s)\, ds \right| \\
&\leq \frac{\sqrt{2}}{\epsilon} M_0 M_3 \|\omega_{\mathbf{j}}\| \left| \frac{\delta t^2}{Z_{\mathbf{j}}^2} (1 - e^{-Z_{\mathbf{j}}} - Z_{\mathbf{j}} e^{-Z_{\mathbf{j}}}) \right| \\
&\leq \frac{\sqrt{2} M_0 M_3}{2\epsilon} \|\omega_{\mathbf{j}}\| \delta t^2. \qquad (3.12)
\end{aligned}
$$

Combined with Taylor Expansion, we obtain:

$$I_{2,3} \leq \frac{1}{2\epsilon} M_0 Z_{\mathbf{j}} \delta t. \tag{3.13}$$

Based on Eq.(3.9), we find that:

$$\begin{aligned} I_{2,4} &\leq \left\| \frac{\delta t}{\epsilon} \frac{1}{1+Z_{\mathbf{j}}} \frac{M_0}{P} \sum_{p=1}^{P} (e^{-i\omega_{\mathbf{j}} \cdot X_{t_n}^p} - e^{-i\omega_{\mathbf{j}} \cdot \widetilde{X}_{t_n}^p}) \right\| \\ &\leq \sqrt{2} M_0 \|\omega_{\mathbf{j}}\| \cdot \frac{\delta t}{\epsilon} \cdot \sum_{p=1}^{P} \frac{\|\widetilde{X}_{t_n}^p - X_{t_n}^p\|}{P}. \end{aligned} \tag{3.14}$$

Let $Y_p = \|\widetilde{X}_{t_n}^p - X_{t_n}^p\|$, where $\{Y_p\}_{p=1}^P$ are i.i.d. random variables. This follows from the fact that the particles $\{X_{t_n}^p\}_{p=1}^P$ and $\{\widetilde{X}_{t_n}^p\}_{p=1}^P$ are separately i.i.d.. Specifically, the i.i.d. property of $\{\widetilde{X}_{t_n}^p\}_{p=1}^P$ is ensured by the RBM described in Alg.2.2. Based on Assumption 1, Y_p is bounded. The empirical mean is defined as:

$$\bar{Y}_P = \frac{1}{P} \sum_{p=1}^{P} Y_p.$$

The expectation of Y_p is:

$$\mu = \mathbb{E}[Y_p] = \mathbb{E}[\|\widetilde{X}_{t_n} - X_{t_n}\|].$$

According to the Bernstein's inequality, for i.i.d. random variables $Y_1, Y_2, \ldots, Y_P$ with $|Y_p - \mu| \leq M_1$(from Assumption 1) almost surely, the probability that the empirical mean deviates from the expectation is bounded as:

$$\mathbb{P}\left(|\bar{Y}_P - \mu| \geq \eta\right) \leq 2\exp\left(-\frac{P\eta^2}{2\sigma^2 + \frac{2M_1\eta}{3}}\right),$$

where $\sigma^2 = \mathbb{E}[(Y_p - \mu)^2]$ is also bounded. With high probability (e.g., $1-\delta$ for very small $\delta > 0$), the following holds:

$$|\bar{Y}_P - \mu| \leq \sqrt{\frac{2\sigma^2 \ln(2/\delta)}{P}} + \frac{2M_1 \ln(2/\delta)}{3P}.$$

This implies that, with $1-\delta$ probability:

$$I_{2,4} \leq \sqrt{2} M_0 \|\omega_{\mathbf{j}}\| \cdot \frac{\delta t}{\epsilon} \cdot \left(\mathbb{E}[\|\widetilde{X}_{t_n} - X_{t_n}\|] + \sqrt{\frac{2\sigma^2 \ln(2/\delta)}{P}} + \frac{2M_1 \ln(2/\delta)}{3P} \right). \tag{3.15}$$

Combining all the equations above and merging the first and second terms, we conclude that, with high probability:

$$\begin{aligned} \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| \leq & \|\widetilde{\alpha}_{t_{n-1};\mathbf{j}} - \alpha_{t_{n-1};\mathbf{j}}\| + \frac{(\sqrt{2}M_0\|\omega_{\mathbf{j}}\| + 1)}{\epsilon \cdot \mathcal{O}(\sqrt{P})} \delta t + \mathcal{O}(\frac{\delta t^2}{\|\omega_{\mathbf{j}}\|^4}) \\ & + C_1 \|\omega_{\mathbf{j}}\| \delta t^2 + \frac{\sqrt{2}M_0\|\omega_{\mathbf{j}}\|}{\epsilon} \delta t \mathbb{E}[\|\widetilde{X}_{t_n} - X_{t_n}\|], \end{aligned} \tag{3.16}$$

where $C_1 = \frac{\sqrt{2}M_0 M_3}{2\epsilon}$ is a constant. ∎

The error estimate between ∇c and $\nabla \widetilde{c}$ is more complex than that between c and $\widetilde{c}$. To analyze this, we introduce an intermediate quantity $\nabla \widetilde{\widetilde{c}}$. Using the frequency notation $\omega_{\mathbf{j}}$ from Lemma 3.5 that $\omega_{\mathbf{j}} = \left(\frac{2\pi j_1}{L}, \frac{2\pi j_2}{L}, \frac{2\pi j_3}{L}\right)$, we define

$$\begin{aligned}
\nabla \widetilde{\widetilde{c}}(\mathbf{x}, t_n) :&= \sum_{\mathbf{j}\in\mathcal{H}} i\omega_{\mathbf{j}} \widetilde{\alpha}_{n;\mathbf{j}} \exp(i\omega_{\mathbf{j}}\mathbf{x}) \\
&= -\frac{\epsilon}{\delta t} \int \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\mathbf{x}-\mathbf{y}) \widetilde{c}_{n-1}(\mathbf{y})\, d\mathbf{y} - \sum_{q=1}^{P} \frac{M_0}{P} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\mathbf{x} - \widetilde{X}_n^q) \\
&= -\frac{\epsilon}{\delta t} \underbrace{\int \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\mathbf{x}+\bar{\mathbf{x}}-\mathbf{y}) \widetilde{c}_{n-1}(\mathbf{y}-\bar{\mathbf{x}})\, d\mathbf{y}}_{I_4} - \underbrace{\sum_{q=1}^{P} \frac{M_0}{P} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\mathbf{x} - \widetilde{X}_n^q)}_{I_5}, \qquad (3.17)
\end{aligned}$$

where $\bar{\mathbf{x}} = \frac{L}{2H} + \lfloor \frac{\mathbf{x}}{L/H} \rfloor \frac{L}{H} - \mathbf{x}$. From Alg.2.2, we have:

$$\begin{aligned}
\nabla \widetilde{c}(\mathbf{x}, t_n) &= -\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * (\epsilon\, \widetilde{c}_{n-1}(\mathbf{x})/\delta t + \widetilde{\rho}_n(\mathbf{x})) \\
&= -\underbrace{\frac{\epsilon}{\delta t} \frac{L^3}{H^3} \sum_{\mathbf{j}\in\mathcal{H}} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\mathbf{x}+\bar{\mathbf{x}}-x_{\mathbf{j}}) \widetilde{c}_{n-1}(x_{\mathbf{j}}-\bar{\mathbf{x}})}_{I_6} \\
&\quad - \underbrace{\sum_{s\in C_p, s\neq p} \frac{M_0}{R} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\mathbf{x} - \widetilde{X}_n^s)}_{I_7}, \qquad (3.18)
\end{aligned}$$

where $x_{\mathbf{j}}$ is the same notation in Eq.(2.14). The error between ∇c and $\nabla \widetilde{c}$ can be estimated by:

$$\|\nabla c(\mathbf{x}, t_n) - \nabla \widetilde{c}(\mathbf{x}, t_n)\| \leq \|\nabla c(\mathbf{x}, t_n) - \nabla \widetilde{\widetilde{c}}(\mathbf{x}, t_n)\| + \|\nabla \widetilde{\widetilde{c}}(\mathbf{x}, t_n) - \nabla \widetilde{c}(\mathbf{x}, t_n)\|. \qquad (3.19)$$

To estimate the error between $\nabla \widetilde{c}$ and $\nabla \widetilde{\widetilde{c}}$, we divide the analysis into two parts:

$$\|\nabla \widetilde{c}(\mathbf{x}, t_n) - \nabla \widetilde{\widetilde{c}}(\mathbf{x}, t_n)\| \leq \frac{\epsilon}{\delta t} \|I_4 - I_6\| + \|I_5 - f_7\|. \qquad (3.20)$$

The first part, involving I_4 and I_6, focuses on the different methods of approximating $(\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{c})$ in $\nabla \widetilde{c}$ and $\nabla \widetilde{\widetilde{c}}$, while the second part, involving I_5 and I_7, examines the differences in the approximations of $(\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{\rho})$ between $\nabla \widetilde{c}$ and $\nabla \widetilde{\widetilde{c}}$. Specifically, I_4 represents the continuous integral, while I_6 is constructed as a discrete Riemann sum that approximates this integral, excluding the interval $[-\frac{L}{2H}, \frac{L}{2H}]^3$.

To analyze the error introduced by the approximation of $(\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{c})$, we rely on the following lemma:

LEMMA 3.6. *For $\forall n \in \mathbb{N}_+$, based on the definitions of I_4 and I_6 in Eq.(3.17) and Eq.(3.18), the following error bound holds:*

$$\|I_4 - I_6\| \leq C_2 (\frac{L}{H})^2 + \frac{1}{\lambda^2 + \frac{\epsilon}{\delta t}} \left(M_3 + \frac{K}{2}\right), \qquad (3.21)$$

where C_2 is a constant that depends on the norm of the second derivative of $\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \widetilde{c}_{n-1}$, M_3 is the uniform bound of $\nabla \widetilde{c}$, K is the spatial Lipschitz constant for $\nabla \widetilde{c}$, L is the characteristic domain size, and H is the grid spacing. Moreover, λ and ϵ are parameters in the system (1.1), $t_n = n\delta t$.

Proof. We rewrite the integral as follows to facilitate the computation of the error between I_4 and I_6. Specifically, we have

$$I_4 = \underbrace{\int_{\mathbf{x}+\bar{\mathbf{x}}-\mathbf{y}\in[-\frac{L}{2H},\frac{L}{2H}]^3} \nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\mathbf{x}+\bar{\mathbf{x}}-\mathbf{y})\widetilde{c}_{n-1}(\mathbf{y}-\bar{\mathbf{x}})\,d\mathbf{y}}_{I_{4,1}}$$

$$+ \underbrace{\int_{\mathbf{x}+\bar{\mathbf{x}}-\mathbf{y}\in[-\frac{L}{2},\frac{L}{2}]^3\setminus[-\frac{L}{2H},\frac{L}{2H}]^3} \nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\mathbf{x}+\bar{\mathbf{x}}-\mathbf{y})\widetilde{c}_{n-1}(\mathbf{y}-\bar{\mathbf{x}})\,d\mathbf{y}}_{I_{4,2}}. \tag{3.22}$$

The leading-order term of the error $\|I_6 - I_{4,2}\|$ depends on the smoothness of the integrand, specifically the second derivatives of the product of functions $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}$ and $\widetilde{c}_{n-1}$. Under Assumption 2, $\widetilde{c}_{n-1}$ is assumed to be twice continuously differentiable with uniformly bounded derivatives. Furthermore, with the inclusion of the shift term, $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}$ can be regarded as smooth, similar to $I_{4,2}$. The smoothness of $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}$ and $\widetilde{c}_{n-1}$ in $[-\frac{L}{2},\frac{L}{2}]^3 \setminus [-\frac{L}{2H},\frac{L}{2H}]^3$ ensures that the integrand is twice differentiable, and its second derivatives are uniformly bounded. As a result, the error $\|I_6 - I_{4,2}\|$ can be bounded by:

$$\|I_6 - I_{4,2}\| \le C_2(\frac{L}{H})^2, \tag{3.23}$$

where the constant C_2 satisfies that

$$C_2 = \mathcal{O}(\|\nabla^2\big(\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t} * \widetilde{c}_{n-1}\big)\|).$$

The boundedness of C_2 is a combined outcome of the derivation, as it relies on the uniform bounds of the second derivatives of the integrand, which are guaranteed by both the smoothness of $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}$ in $[-\frac{L}{2},\frac{L}{2}]^3 \setminus [-\frac{L}{2H},\frac{L}{2H}]^3$ and Assumption 2 on $\widetilde{c}_{n-1}$.

The integral $I_{4,1}$ is defined as:

$$\begin{aligned} I_{4,1} &= \int_{\mathbf{z}\in\left[-\frac{L}{2H},\frac{L}{2H}\right]^3} \nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\mathbf{z})\widetilde{c}_{n-1}(\mathbf{x}-\mathbf{z})\,d\mathbf{z} \\ &= \int_{\mathbf{z}\in\left[-\frac{L}{2H},\frac{L}{2H}\right]^3} \nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\mathbf{z})\left(\widetilde{c}_{n-1}(\mathbf{x}) - \nabla\widetilde{c}_{n-1}(\mathbf{x})\cdot\mathbf{z} + \frac{1}{2}\mathbf{z}^\top H(\widetilde{c}_{n-1}(\xi))\mathbf{z}\right)d\mathbf{z} \\ &:= I_{4,1}^{(0)} + I_{4,1}^{(1)} + R_1, \end{aligned} \tag{3.24}$$

where $H(\widetilde{c}_{n-1}(\xi))$ is the Hessian matrix of $\widetilde{c}_{n-1}$ (composed of second-order partial derivatives at some point ξ between $\mathbf{z}$ and $\mathbf{x}$).

Since $\int_{\mathbf{z}} \mathbf{z}\,d\mathbf{z} = 0$ over a symmetric domain, the zeroth-order term $I_{4,1}^{(0)}$ vanishes

$$I_{4,1}^{(0)} = \widetilde{c}_{n-1}(\mathbf{x})\int_{\mathbf{z}} \frac{\exp(-\beta\|\mathbf{z}\|)}{4\pi\|\mathbf{z}\|^3}(1+\beta\|\mathbf{z}\|)\mathbf{z}\,d\mathbf{z} = 0, \tag{3.25}$$

where $\beta = \sqrt{\lambda^2 + \epsilon/\delta t}$ is the same notation in Eq.(2.9).

For the first term $I_{4,1}^{(1)}$, switching to spherical coordinates: let $\|\mathbf{z}\| = r$, $\mathbf{z} = r\hat{\mathbf{z}}$, where $\hat{\mathbf{z}}$ is the unit vector $(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Substituting these, the integral

becomes:

$$\begin{aligned}
\|I_{4,1}^{(1)}\| &\le \| -\int_0^{\frac{L}{2H}} \int_0^{\pi} \int_0^{2\pi} \frac{\exp(-\beta r)}{4\pi r^3}(1+\beta r) r^4 \hat{\mathbf{z}}(\hat{\mathbf{z}} \cdot \nabla \widetilde{c}_{n-1}(\mathbf{x})) \sin\theta \, d\phi \, d\theta \, dr \| \\
&= \| \frac{1}{3} \nabla \widetilde{c}_{n-1}(\mathbf{x}) \int_0^{\frac{L}{2H}} r \exp(-\beta r)(1+\beta r) \, dr \| \\
&= \left\| -\frac{\nabla \widetilde{c}_{n-1}(\mathbf{x})}{\beta^2} \left[1 - \left(1 + \beta \frac{L}{2H} + \frac{1}{3}(\beta \frac{L}{2H})^2 \right) \exp\left(-\beta \frac{L}{2H} \right) \right] \right\| \\
&\le \left\| \frac{\nabla \widetilde{c}_{n-1}(\mathbf{x})}{\beta^2} \right\| \\
&\le \left| \frac{M_3}{\beta^2} \right|, \qquad (3.26)
\end{aligned}$$

where, according to Assumption 3, β diverges to positive infinity at a faster rate than H, and according to Assumption 2, M_3 is the uniform bound of $\nabla \widetilde{c}_{n-1}$.

Under Assumption 2, $H(\widetilde{c}_{n-1}(\xi))$ is bounded, then we can get the inequality for the remainder term R_1.

$$\begin{aligned}
|R_1| &\le |\frac{1}{2} \int_0^{\frac{L}{2H}} \int_0^{\pi} \int_0^{2\pi} \frac{\exp(-\beta r)}{4\pi r^3}(1+\beta r) \cdot r^2 (\hat{\mathbf{z}}^\top H(\widetilde{c}_{n-1}(\xi)) \hat{\mathbf{z}}) \cdot r^2 \sin\theta \, d\phi \, d\theta \, dr| \\
&\le |\frac{K}{6} \int_0^{\frac{L}{2H}} r \exp(-\beta r)(1+\beta r) \, dr| \\
&\le \frac{K}{2\beta^2}, \qquad (3.27)
\end{aligned}$$

where K is the spatial Lipschitz constant for $\nabla \widetilde{c}$.

From the above inequalities, we can conclude that:

$$\|I_4 - I_6\| \le C_2 (\frac{L}{H})^2 + \frac{1}{\beta^2} \left(M_3 + \frac{K}{2} \right). \qquad (3.28)$$ ∎

We now proceed to estimate $(\nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t} * \rho_n)(\widetilde{X}_n^p)$ in $\nabla \widetilde{c}$ and $\nabla \widetilde{\widetilde{c}}$. Using the RBM in Alg.2.2, we replace

$$\sum_{q=1, q\neq p}^{P} \frac{M_0}{P} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)$$

with

$$\sum_{s \in C_p, s \neq p} \frac{M_0}{R} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^s).$$

We write

$$\zeta_{n,p} := \sum_{q=1, q\neq p}^{P} \frac{M_0}{P} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q) - \sum_{s \in C_p, s \neq p} \frac{M_0}{R} \nabla_{\mathbf{x}} \mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^s).$$

LEMMA 3.7. *For $\forall n \in \mathbb{N}_+$, $p \in 1, 2, ..., P$,*

$$\mathbb{E}(\|\zeta_{n,p}\|) \le M_0 M_4 \sqrt{\left(\frac{1}{R} - \frac{1}{P} \right)}, \qquad (3.29)$$

where $M_4 = \max_{q\neq p} \|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\|$, M_0 *is the conserved total mass,* P *is the total number of particles,* R *is the batch size.*

Proof. Similar to Lemma 3.1 in [18], we rewrite

$$f_p = \sum_{s\in C_p, s\neq p} \frac{M_0}{R}\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^s) = \sum_{q=1,q\neq p}^{P} \frac{M_0}{R}\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)I(p,q), \tag{3.30}$$

where $I(p,q)$ means q is in the batch C_p. Here we have that $I(p,q)$ is a Bernoulli random variable with $\mathbb{E}[I(p,q)] = \frac{R}{P}$, which indicates that $\mathbb{E}[\zeta_{n,p}] = 0$.

$$\begin{aligned}
\mathbb{E}|f_p|^2 =& \frac{M_0^2}{R^2} \sum_{\substack{q,r:\\ q\neq r, q\neq p,\\ r\neq p}} \|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\cdot\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^r)\|^2 P(I(p,q)I(p,r)=1) \\
&+ \frac{M_0^2}{R^2}\sum_{q=1,q\neq p}^{P} \|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\|^2 P(I(p,q)=1) \\
=& \frac{M_0^2}{RP}\sum_{q=1,q\neq p}^{P} \|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\|^2 \\
&+ \frac{M_0^2}{P^2}\sum_{q,r:q\neq r,q\neq p,r\neq p} \|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\cdot\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^r)\|^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
\mathrm{Var}(\zeta_{n,p}) &= \mathbb{E}|f_p|^2 - (\mathbb{E}|f_p|)^2 \\
&= M_0^2(\frac{1}{R}-\frac{1}{P})\frac{1}{P}\sum_{q=1,q\neq p}^{P} \|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\|^2.
\end{aligned} \tag{3.31}$$

According to Jensen's Inequality, we obtain:

$$\mathbb{E}(\|\zeta_{n,p}\|) \le \sqrt{\mathbb{E}(\|\zeta_{n,p}\|^2)} = \sqrt{\mathrm{Var}(\zeta_{n,p})} \le M_0 M_4\sqrt{\left(\frac{1}{R}-\frac{1}{P}\right)}, \tag{3.32}$$

where $M_4 = \max_{q\neq p} \|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\|$. Since all particles are located at distinct positions in the SIPF-r algorithm ($\widetilde{X}_n^p \neq \widetilde{X}_n^q$ for $p \neq q$), there exists a minimum separation distance $d_{\min} > 0$ between any two particles. Consequently, $\|\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}(\widetilde{X}_n^p - \widetilde{X}_n^q)\|$ is bounded for all pairs of particles. This ensures that M_4, which is the maximum of these kernel gradient norms, is finite. ∎

Now we quantify the error between $\nabla\widetilde{\widetilde{c}}$ and ∇c as follows.

LEMMA 3.8. *For* $\forall n \in \mathbb{N}_+$*, with high probability:*

$$\begin{aligned}
&\|\nabla\widetilde{\widetilde{c}}(\mathbf{x},t_n) - \nabla c(\mathbf{x},t_n)\| \\
\le& L^{3/2}\max_{\mathbf{j}\in\mathcal{H}}\|\omega_{\mathbf{j}}\|\|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| \\
\le& \left(C_3\cdot\frac{H^2}{\mathcal{O}(\sqrt{P})} + C_4\cdot H\right)\delta t + \mathcal{O}(\frac{\delta t^2}{H^3}) + C_5\cdot H^2\cdot\delta t^2 \\
&+ C_3\cdot H^2\delta t\cdot\mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|) + L^{\frac{3}{2}}\max_{\mathbf{j}\in\mathcal{H}}\|\omega_{\mathbf{j}}\||\widetilde{\alpha}_{t_{n-1};\mathbf{j}} - \alpha_{t_{n-1};\mathbf{j}}|,
\end{aligned} \tag{3.33}$$

where C_3, C_4, C_5 are constants, $t_n = n\delta t$.

Proof. We begin with the Fourier transform of the difference between $\widetilde{\widetilde{c}}$ and c:

$$\mathcal{F}\{\widetilde{\widetilde{c}} - c\}(\mathbf{k}) = \hat{\widetilde{\widetilde{c}}}(\mathbf{k}) - \hat{c}(\mathbf{k}),$$

where $\mathbf{k} = (k_1, k_2, k_3)$ is the Fourier dual variable.

Next, we consider the Fourier transform of the gradient difference $\nabla\widetilde{\widetilde{c}} - \nabla c$. Using the properties of the Fourier transform, we have:

$$\mathcal{F}\{\nabla\widetilde{\widetilde{c}} - \nabla c\}(\mathbf{k}) = i\mathbf{k} \cdot \big(\hat{\widetilde{\widetilde{c}}}(\mathbf{k}) - \hat{c}(\mathbf{k})\big).$$

Here, the operation $i\mathbf{k}$ corresponds to multiplication in the Fourier domain, which is the Fourier representation of the gradient operator in real space. By the Parseval's theorem,

$$\|\nabla c(\mathbf{x}, t_n) - \nabla\widetilde{\widetilde{c}}(\mathbf{x}, t_n)\| = \sqrt{\frac{L^3}{H^3}\sum_{\mathbf{j}\in\mathcal{H}} \|\omega_{\mathbf{j}}\|^2 |\alpha_{t_n;\mathbf{j}} - \widetilde{\alpha}_{t_n;\mathbf{j}}|^2}. \tag{3.34}$$

Using the conclusion of Lemma 3.5 and inequality $\sum_{j=1}^{n}(a_j b_j) \le n \cdot \max a_j b_j$, we obtain that with high probability:

$$\begin{aligned}
&\|\nabla c(\mathbf{x}, t_n) - \nabla\widetilde{\widetilde{c}}(\mathbf{x}, t_n)\| \\
\le& L^{\frac{3}{2}} \max_{\mathbf{j}\in\mathcal{H}} \|\omega_{\mathbf{j}}\| \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| \\
\le& \left(C_3 \cdot \mathcal{O}\left(\frac{H^2}{\sqrt{P}}\right) + C_4 \cdot H\right)\delta t + \mathcal{O}(\frac{\delta t^2}{H^3}) + C_5 \cdot H^2 \cdot \delta t^2 \\
&+ C_3 \cdot H^2 \delta t \cdot \mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|) + L^{\frac{3}{2}} \max_{\mathbf{j}\in\mathcal{H}} \|\omega_{\mathbf{j}}\| |\widetilde{\alpha}_{t_{n-1};\mathbf{j}} - \alpha_{t_{n-1};\mathbf{j}}|,
\end{aligned} \tag{3.35}$$

where the constants C_3, C_4, C_5 are given by:

$$C_3 = \frac{3\sqrt{2}\pi^2 M_0}{\epsilon\sqrt{L}}, \quad C_4 = \frac{2\pi\sqrt{3L}M_0}{\epsilon}, \quad C_5 = \frac{3\pi^2 M_0 M_3}{\epsilon\sqrt{2L}}. \tag{3.36}$$

∎

Hence, combining Lemma 3.6, 3.7, 3.8 and Eq.(3.19), we get that with high probability:

$$\begin{aligned}
&\mathbb{E}(\|\nabla c(\widetilde{X}_{t_n}, t_n) - \nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n)\|) \\
\le& \mathbb{E}(\|\nabla c(\widetilde{X}_{t_n}, t_n) - \nabla\widetilde{\widetilde{c}}(\widetilde{X}_{t_n}, t_n)\|) + \mathbb{E}(\|\nabla\widetilde{\widetilde{c}}(\widetilde{X}_{t_n}, t_n) - \nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n)\|) \\
\le& \frac{\epsilon}{\delta t}\left(C_2(\frac{L}{H})^2 + \frac{1}{\lambda^2 + \frac{\epsilon}{\delta t}}(M_3 + \frac{K}{2})\right) + M_0 M_4\sqrt{\left(\frac{1}{R} - \frac{1}{P}\right)} \\
&+ L^{3/2} \max_{\mathbf{j}\in\mathcal{H}} \|\omega_{\mathbf{j}}\| \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| \\
\le& \frac{\epsilon}{\delta t}\left(C_2(\frac{L}{H})^2 + \frac{1}{\lambda^2 + \frac{\epsilon}{\delta t}}(M_3 + \frac{K}{2})\right) + M_0 M_4\sqrt{\left(\frac{1}{R} - \frac{1}{P}\right)} \\
&+ \left(C_3 \cdot \mathcal{O}\left(\frac{H^2}{\sqrt{P}}\right) + C_4 \cdot H\right)\delta t + \mathcal{O}(\frac{\delta t^2}{H^3}) + C_5 \cdot H^2 \cdot \delta t^2 \\
&+ C_3 \cdot H^2 \delta t \cdot \mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|) + L^{3/2} \max_{\mathbf{j}\in\mathcal{H}} \|\omega_{\mathbf{j}}\| |\widetilde{\alpha}_{t_{n-1};\mathbf{j}} - \alpha_{t_{n-1};\mathbf{j}}|.
\end{aligned} \tag{3.37}$$

For simplicity of notation in the proof below, we define:

$$a_n := \mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|), \tag{3.38}$$

$$b_n := L^{3/2} \max_{\mathbf{j}\in\mathcal{H}} \|\omega_{\mathbf{j}}\| \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\|. \tag{3.39}$$

The following provides a bound on the error between $\widetilde{X}_{t_{n+1}}$ and $X_{t_{n+1}}$.

LEMMA 3.9. *For $\forall n \in \mathbb{N}_+$,*

$$\begin{aligned}\mathbb{E}(\|\widetilde{X}_{t_{n+1}} - X_{t_{n+1}}\|) \leq & (1+\chi K\delta t)\mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|) + \chi\epsilon C_2(\frac{L}{H})^2 \\ & + \chi\delta t\left(L^{3/2}\max_{\mathbf{j}\in\mathcal{H}}\|\omega_{\mathbf{j}}\|\|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| + 3M_3 + \frac{K}{2} + \frac{1}{\mathcal{O}(\sqrt{R})}\right),\end{aligned} \tag{3.40}$$

where K is the Lipschitz constant, M_3 is the uniform bound of ∇c, C_2 is a constant that depends on the norm of the second derivative of $\nabla_{\mathbf{x}}\mathcal{K}_{\epsilon,\delta t}$, $t_n = n\delta t$.

Proof. According to Eqs.(2.15)-(2.16),

$$\begin{aligned}&\mathbb{E}(\|\widetilde{X}_{t_{n+1}} - X_{t_{n+1}}\|) \\ \leq & \mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|) + \chi\mathbb{E}(\int_{t_n}^{t_{n+1}} \|\nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n) - \nabla c(X_s, s)\|\, ds) \\ = & \mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|) + \chi\int_{t_n}^{t_{n+1}} \mathbb{E}(\|\nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n) - \nabla c(X_s, s)\|)\, ds,\end{aligned} \tag{3.41}$$

by the triangle inequality and Tonelli's theorem. According to the Assumption 2,

$$\begin{aligned}&\|\nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n) - \nabla c(X_s, s)\| \\ \leq & \|\nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n) - \nabla c(\widetilde{X}_{t_n}, t_n)\| + \|\nabla c(\widetilde{X}_{t_n}, t_n) - \nabla c(X_{t_n}, t_n)\| \\ & + \|\nabla c(X_{t_n}, t_n) - \nabla c(X_s, s)\| \\ \leq & \|\nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n) - \nabla c(\widetilde{X}_{t_n}, t_n)\| + K\|X_{t_n} - \widetilde{X}_{t_n}\| \\ & + \|\nabla c(X_{t_n}, t_n) - \nabla c(X_s, s)\|,\end{aligned} \tag{3.42}$$

where K is the Lipschitz constant.

Using the notations for a_n and b_n in Eqs.(3.38)-(3.39), we have:

$$\begin{aligned}
a_{n+1} =&\mathbb{E}(\|\widetilde{X}_{t_{n+1}} - X_{t_{n+1}}\|)\\
\leq&\mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|) + \chi\int_{t_n}^{t_{n+1}} \mathbb{E}(\|\nabla\widetilde{c}(\widetilde{X}_{t_n}, t_n) - \nabla c(\widetilde{X}_{t_n}, t_n)\|)\, ds\\
&+ \chi K\int_{t_n}^{t_{n+1}} \mathbb{E}(\|\widetilde{X}_{t_n} - X_{t_n}\|)\, ds + \chi\int_{t_n}^{t_{n+1}} \mathbb{E}(\|\nabla c(X_{t_n}, t_n) - \nabla c(X_s, s)\|)\, ds\\
\leq&(1+\chi K\delta t)a_n + \chi\int_{t_n}^{t_{n+1}} \mathbb{E}(\|\nabla c(X_{t_n}, t_n) - \nabla c(X_s, s)\|)\, ds\\
&+ \chi\delta t\left(b_n + \frac{\epsilon}{\delta t}\left(C_2(\frac{L}{H})^2 + \frac{1}{\lambda^2 + \frac{\epsilon}{\delta t}}(M_3 + \frac{K}{2})\right) + M_0M_4\sqrt{\left(\frac{1}{R} - \frac{1}{P}\right)}\right)\\
\leq&(1+\chi K\delta t)a_n + \chi\delta t b_n\\
&+ \chi\delta t\left(2M_3 + \frac{\epsilon}{\delta t}\left(C_2(\frac{L}{H})^2 + \frac{1}{\lambda^2 + \frac{\epsilon}{\delta t}}(M_3 + \frac{K}{2})\right) + M_0M_4\sqrt{\left(\frac{1}{R} - \frac{1}{P}\right)}\right)
\end{aligned} \tag{3.43}$$

$$\leq(1+\chi K\delta t)a_n + \chi\delta t\left(b_n + 3M_3 + \frac{K}{2} + \mathcal{O}(\frac{1}{\sqrt[4]{P}})\right) + \chi\epsilon C_2(\frac{L}{H})^2,$$

where K is the Lipschitz constant, M_3 is the uniform bound of ∇c, thereby concluding the proof. ∎

Now, we are ready to prove Theorem 3.3.

Proof of Theorem 3.3. From Lemmas 3.9 and Eq.(3.37), we obtain the system of inequalities that couples a_n and b_n defined in Eqs.(3.38)-(3.39):

$$a_{n+1} \leq(1+\chi K\delta t)a_n + \chi\delta t\left(b_n + 3M_3 + \frac{K}{2} + \mathcal{O}(\frac{1}{\sqrt{R}})\right) + \chi\epsilon C_2(\frac{L}{H})^2, \tag{3.44}$$

$$b_{n+1} \leq \left(C_3\cdot\mathcal{O}\left(\frac{H^2}{\sqrt{P}}\right) + C_4\cdot H\right)\delta t + \mathcal{O}(\frac{\delta t^2}{H^3}) + C_5\cdot H^2\delta t^2 + C_3\cdot H^2\delta t a_{n+1} + b_n. \tag{3.45}$$

From this coupled system, we can derive a general bound for a_n. Substituting Eq.(3.45) into Eq.(3.44), we iteratively propagate and simplify the inequality to derive:

$$\begin{aligned}
a_{n+1} \leq&(1+\chi(K\delta t + C_3\cdot H^2\delta t^2))a_n + \chi\epsilon C_2(\frac{L}{H})^2 + \sum_{j=1}^{n-1}\chi C_3\cdot H^2\delta t^2 a_j\\
&+ \chi\delta t\left(3M_3 + \frac{K}{2} + \mathcal{O}(\frac{1}{\sqrt{R}}) + T\cdot\left(C_3\cdot\mathcal{O}\left(\frac{H^2}{\sqrt{P}}\right) + C_4\cdot H\right)\right) + \mathcal{O}(\delta t^2).
\end{aligned} \tag{3.46}$$

By the discrete Gronwall inequality, if (u_n) and (w_n) be nonnegative sequences satisfying

$$u_n \leq \alpha + \sum_{k=0}^{n-1} u_k w_k \quad \forall n \geq 1,$$

for some constant $\alpha \geq 0$. Then for all $n \geq 1$, the sequence (u_n) satisfies the bound:

$$u_n \leq \alpha \exp\left(\sum_{k=0}^{n-1} w_k\right).$$

Applying this result to the recursive inequality (3.46), we obtain the following bound with high probability:

$$a_{n+1} \leq \left(N_0(\frac{L}{H})^2 + N_1\delta t + \left(N_2 \cdot \mathcal{O}\left(\frac{H^2}{\sqrt{P}}\right) + N_3 \cdot H + \frac{1}{\mathcal{O}(\sqrt{R})}\right) \cdot \delta t + \mathcal{O}(\delta t^2)\right)$$

$$\cdot \exp(1 + N_4\delta t + N_2 H^2 \delta t), \tag{3.47}$$

for $\forall n \geq 0$, where

$$N_0 = \chi\epsilon C_2, \quad N_1 = \chi(3M_3 + \frac{K}{2}), \quad N_2 = \chi T C_3, \quad N_3 = \chi C_4, \quad N_4 = \chi K. \tag{3.48}$$

Here, C_2 is a constant that defined in Eq.(3.23) in Lemma 3.6, C_3, C_4 are constants that defined in Eq.(3.36) in Lemma 3.8.
According to the discrete and continuous dynamics defined in Eqs.(2.15)-(2.16), the 1-Wasserstein distance between the approximate and exact distributions at time t_{n+1} is given by:

$$\mathcal{W}_1(\widetilde{\rho}_{t_{n+1}}, \rho_{t_{n+1}}) = \inf_{\gamma \in \Pi(\widetilde{\rho}_{t_{n+1}}, \rho_{t_{n+1}})} \left(\int_{\mathbb{R}^3 \times \mathbb{R}^3} \|\mathbf{x} - \mathbf{y}\|_{L^1} \, d\gamma(\mathbf{x}, \mathbf{y})\right), \tag{3.49}$$

where the infimum is taken over all possible couplings of the two distributions. Under the natural coupling induced by shared initial conditions and Brownian motion paths (i.e., $\widetilde{X}_{t_n}$ and X_{t_n} evolve via the same Wiener process W_s), we explicitly construct a joint distribution $\gamma_n = \text{Law}(\widetilde{X}_{t_n}, X_{t_n})$. This coupling allows us to bound the Wasserstein distance as:

$$\begin{aligned}
\mathcal{W}_1(\widetilde{\rho}_{t_{n+1}}, \rho_{t_{n+1}}) \leq & \mathbb{E}(\|\widetilde{X}_{t_{n+1}} - X_{t_{n+1}}\|_{L^1}) \\
\leq & \sqrt{3}\mathbb{E}(\|\widetilde{X}_{t_{n+1}} - X_{t_{n+1}}\|_{L^2}) \\
\leq & \left(S_0(\frac{L}{H})^2 + S_1\delta t + \left(S_2\mathcal{O}\left(\frac{H^2}{\sqrt{P}}\right) + S_3 H + \frac{1}{\mathcal{O}(\sqrt{R})}\right)\delta t + \mathcal{O}(\delta t^2)\right) \\
& \cdot \exp(1 + S_4\delta t + S_2 H^2 \delta t),
\end{aligned} \tag{3.50}$$

where for $\forall n \geq 0$

$$S_i = \sqrt{3}N_i, \quad \text{for all } i = 0, \ldots, 4. \tag{3.51}$$

The inequality follows from the fact that the Wasserstein distance is defined as the infimum over all possible couplings, and our construction provides one such coupling. The transition is obtained through the elementary norm inequality $\|\mathbf{x}\|_{L^1} \leq \sqrt{3}\|\mathbf{x}\|_{L^2}$ for vectors in $\mathbb{R}^3$, which follows from the Cauchy-Schwarz inequality.
Combining Eq.(3.47) with Lemma 3.5, we have:

$$\max_{\mathbf{j} \in \mathcal{H}} \|\widetilde{\alpha}_{t_n;\mathbf{j}} - \alpha_{t_n;\mathbf{j}}\| \leq \left(\frac{S_7}{H} + \left(S_8 H + S_9 H^2 + S_5 \frac{H}{\mathcal{O}(\sqrt{R})} + S_{10}\frac{H^3}{\mathcal{O}(\sqrt{P})}\right)\delta t\right)$$

$$\cdot \exp(1 + S_4\delta t + S_2 H^2 \delta t) + S_5 \frac{H}{\mathcal{O}(\sqrt{P})} + S_6 H \delta t, \tag{3.52}$$

for $\forall n \geq 0$, where

$$S_5 = \frac{\sqrt{6}M_0T}{2\epsilon}, \quad S_6 = \frac{\sqrt{3}\pi TC_1}{2}, \quad S_7 = \frac{\sqrt{6}L^2M_0TS_0}{2\epsilon}, \tag{3.53}$$
$$S_8 = \frac{\sqrt{6}M_0TS_1}{2\epsilon}, \quad S_9 = \frac{\sqrt{6}L^2M_0TS_3}{2\epsilon}, \quad S_{10} = \frac{\sqrt{6}L^2M_0TS_2}{2\epsilon},$$

where C_1 is a constant defined in Eq.(3.16) in Lemma 3.5. This completes the proof of Theorem 3.3.

4. Numerical Experiments. The numerical experiments are divided into two main subsections: (1) validation of the assumptions and (2) validation of the convergence rate of the SIPF-r method. These experiments aim to empirically verify the theoretical foundations and practical performance of the algorithm. The interested reader is referred to [39] for demonstrations that our algorithm can handle multi-modal initial data and resolve complex evolution processes, including merging of particle clusters and finite-time singularity formation in the 3D fully parabolic KS systems.

4.1. Validation of Convergence Rate.

4.1.1. Accuracy of SIPF-r Method. Because some adjustments have been made to the original SIPF algorithm [39], and the RBM [18] has been introduced, we verify the accuracy of the SIPF-r method in Section 3. In the radially symmetric case, the fully parabolic KS system (1.1) can be expressed as $\rho(x,y,z,t) = \rho(r,t)$ and $c(x,y,z,t) = c(r,t)$, where $r = \sqrt{x^2+y^2+z^2}$. The system is then rewritten as follows:

$$\begin{cases} \rho_t = \mu\left(\dfrac{\partial^2\rho}{\partial r^2} + \dfrac{2}{r}\dfrac{\partial\rho}{\partial r}\right) - \chi\left(\dfrac{\partial\rho}{\partial r}\dfrac{\partial f}{\partial r} + \rho\cdot(\dfrac{\partial^2 f}{\partial r^2} + \dfrac{2}{r}\dfrac{\partial f}{\partial r})\right), \\ \epsilon c_t = \left(\dfrac{\partial^2 c}{\partial r^2} + \dfrac{2}{r}\dfrac{\partial c}{\partial r}\right) - \lambda^2 c + \rho. \end{cases} \tag{4.1}$$

To quantify the accuracy of the SIPF-r method, we compute a reference solution using a very fine mesh for the radial system, which serves as a benchmark for comparison. We define the relative error between the cumulative distribution functions (CDFs) obtained from the radial finite difference method (FDM) and the SIPF-r method as

$$\text{Relative Error} = \frac{1}{N}\sum_{i=1}^{N}\begin{cases} 0, & \text{if } F_{\text{FDM}}(s_i) = 0, \\ \frac{|F_{\text{SIPF-}r}(s_i) - F_{\text{FDM}}(s_i)|}{F_{\text{FDM}}(s_i)}, & \text{otherwise,} \end{cases} \tag{4.2}$$

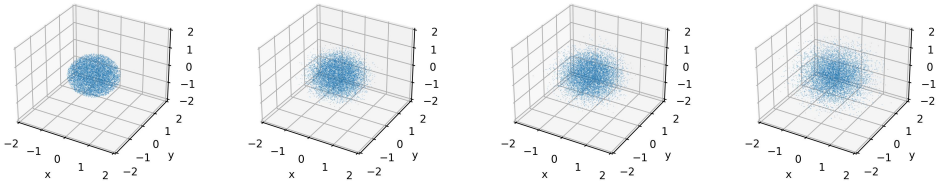
where $F_{\text{SIPF-}r}(s_i)$ and $F_{\text{FDM}}(s_i)$ represent the CDFs of ρ computed via the SIPF-r and FDM methods respectively, and s_i denotes the i-th radial mesh point in the FDM, which are the discrete points along the radial direction starting from the origin. To ensure the relative error is well-defined, we set it to zero wherever $F_{\text{FDM}}(s_i) = 0$.

Here the initial distribution ρ_0 is assumed to be a uniform distribution over a ball centered at $(0,0,0)^T$ with radius 1. The model parameters are chosen as follows:

$$\mu = \chi = 1, \quad \epsilon = 10^{-4}, \quad \lambda = 10^{-1}. \tag{4.3}$$

For the numerical computation, we use $H = 24$ Fourier basis in each spatial dimension to discretize the chemical concentration c and use $P = 10000$ particles to represent the approximated distribution ρ, where the batch size in Alg.2.2 is $R = \lfloor \sqrt{P} \rfloor = 100$. The computational domain is $\Omega = [-L/2, L/2]^3$, where $L = 8$, and the total mass is chosen to be $M_0 = 20$. The evolution of c and ρ is computed using Alg.2.3 with a time step size $\delta t = 10^{-4}$, up to the final simulation time $T = 0.1$.

In Fig.1, we present the evolution of particles over time, showing the dynamic behavior of ρ. Additionally, in Fig.2, we compare the cumulative probability curves of ρ obtained from the radial FDM and the SIPF-r method at $T = 0.1$, with a mean relative error of 0.05512 as defined in Eq.(4.2). This comparison demonstrates that the SIPF-r algorithm achieves high accuracy in approximating the true solution. These results validate the effectiveness of the SIPF-r algorithm in capturing the behavior of the particle distribution.



(a) t=0 (b) t=0.025 (c) t=0.05 (d) t=0.1

Fig. 1: Scattering plot of particles with $M_0 = 20$.

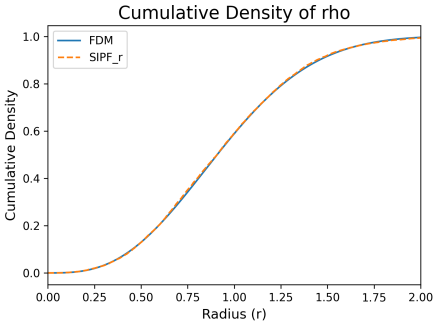
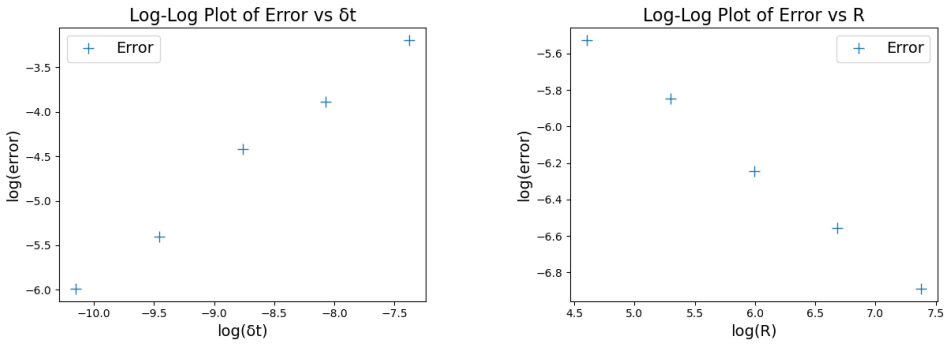


Fig. 2: Cumlative distribution of ρ computed by SIPF-r and radial FDM

4.1.2. Convergence of the SIPF-r Method. In this subsection, we validate the convergence of the SIPF-r numerically. Based on Eq.(3.52), the error between $\widetilde{c}$ and c can be quantified by the L^2 error between their Fourier coefficients $\widetilde{\alpha}$ and α. We adopt the same initial conditions in Subsection 4.1.1. To eliminate the uncertainty introduced by the RBM, the reference solution is computed using the original SIPF method [39] with parameters $\delta t = 10^{-6}$, $H = 24$, and $P = 10000$. Additionally, we set $M_0 = 20, T = 0.01$ to ensure that the system remains free of singularities,

as verified in Fig.3 of [39]. To investigate the convergence with respect to the time step δt, we vary δt from $2^{-8}T$ to $2^{-4}T$. Since Theorem 3.3 holds with high probability, we perform 100 independent experiments for each δt to empirically validate the algorithm's accuracy. The mean L^2 error of the Fourier coefficients is computed over these 100 trials. As shown in Fig.3a, the slope of the mean L^2 error versus δt on a logarithmic scale indicates an approximate first-order convergence rate, with $e(\delta t) = \mathcal{O}(\delta t^{1.023})$. This result aligns with the theoretical bound given in Eq.(3.4) of Theorem 3.3. Furthermore, we examine the mean L^2 error of $\widetilde{c}(\cdot, T)$ for varying batch sizes $R = 100, 200, 400, 800, 1600$, while keeping $P = 10000$. From Eq.(3.4), with other parameters unchanged, the theoretical L^2 error of $\widetilde{c}$ with respect to the batch size R should scale as $\mathcal{O}(R^{-\frac{1}{2}})$. This is empirically verified in Fig.3b, where the fitted convergence rate is $e(R) = \mathcal{O}(R^{-0.495})$, closely matching the theoretical prediction.



(a) vs. time step δt (log-scale)

(b) vs. batch size R (log-scale)

Fig. 3: L^2 error of $\widetilde{c}$ in SIPF-r

4.2. Validation of Theoretical Assumptions.

4.2.1. Spatial Lipschitz Continuity. To verify the spatial Lipschitz continuity in Assumption 2, we change the spatial discretization, varying H from 6 to 24. At the final time $T = 0.1$, we randomly select 1000 pairs of particle points from a total of 10,000 particles in each calculation. The Spatial Lipschitz Constant $L(H)$ for $\nabla\widetilde{c}$ is defined as the maximum ratio of the gradient difference to the spatial distance over all pairs of particle points $\{\mathbf{x}, \mathbf{y}\}$:

$$L(H) := \max_{\{\mathbf{x},\mathbf{y}\}} \frac{\|\nabla\widetilde{c}(\mathbf{x}, T) - \nabla\widetilde{c}(\mathbf{y}, T)\|}{\|\mathbf{x} - \mathbf{y}\|}. \tag{4.4}$$

The results, shown in Table 1, list the computed Lipschitz constant $L(H)$ for each value of H. The variation in these values is relatively small, confirming that the spatial Lipschitz continuity holds for $\nabla\widetilde{c}$ computed by the SIPF-r algorithm.

4.2.2. CFL-like Condition. To validate Assumption 3, we conduct experiments by selecting several pairs of $(\delta t, H)$ and (P, H) that violate the conditions outlined in the assumption. Specifically, we choose the following pairs:
- For $(\delta t, H)$:

$$(8 \times 10^{-3}, 4),\ (4 \times 10^{-3}, 8),\ (2 \times 10^{-3}, 16),\ (1 \times 10^{-3}, 32).$$

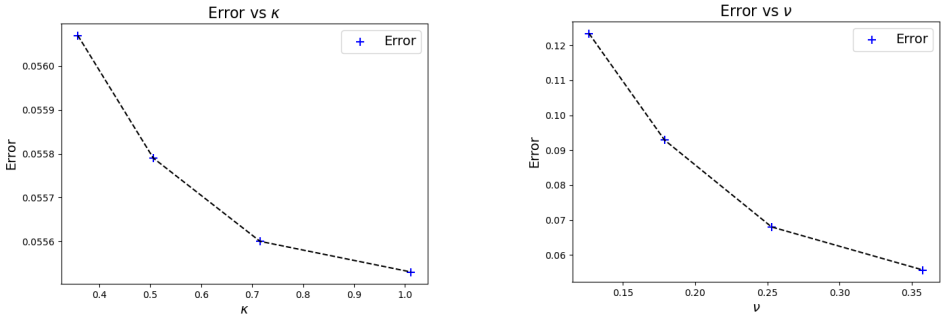
Fourier Modes(H)	**Spatial Lipschitz Constant($L(H)$)**
6	0.002085
12	0.002106
18	0.002036
24	0.001957

Table 1: Spatial Lipschitz Constant of $\nabla\widetilde{c}$ vs. H.

- For (P, H):

$$(1000, 4),\ (2000, 8),\ (4000, 16),\ (8000, 32).$$

For each pair, we repeat the experiment 100 times and compute the mean of the relative error defined in Eq.(4.2). Following the notations in Assumption 3, we define $\kappa = H\sqrt{\delta t}$ and $\nu = \frac{H}{\sqrt{P}}$. In Fig.4, we plot the error versus κ and ν, corresponding to the above pairs of $(\delta t, H), (P, H)$. As δt decreases and H increases, it is evident that the error decreases. However, since H increases at a faster rate than δt decreases, which violates the condition $\kappa = H\sqrt{\delta t} \to 0$ in Assumption 3. Similarly, H increases at a faster rate than $\sqrt{P}$, which violates the condition $\nu = \frac{H}{\sqrt{P}} \to 0$. As a result, the error reduction slows down, and convergence cannot be achieved under these conditions. This demonstrates that convergence cannot be achieved under these conditions, thereby validating the necessity of Assumption 3.



(a) vs. $\kappa = H\sqrt{\delta t}$ (b) vs. $\nu = \frac{H}{\sqrt{P}}$

Fig. 4: Relative Error

5. Conclusions. In this paper, we introduced a random batch variant [18] of the original SIPF method [39] to approximate the 3D fully parabolic KS system. This modification leverages the randomness in batch sampling to bypass the mean-field limit, reducing computational complexity without sacrificing accuracy. We established the L^2 convergence of the SIPF-r method for the 3D fully parabolic KS system. Specifically, we prove the convergence with high probability for both the density $\widetilde{\rho}(\mathbf{x}, t)$ and the concentration field $\widetilde{c}(\mathbf{x}, t)$ to their respective exact solutions $\rho(\mathbf{x}, t)$ and $c(\mathbf{x}, t)$. The error bounds reveal a dependence on δt, H, P, and R, with the density and concentration field exhibiting distinct but interrelated convergence behaviors.

Computational results further validated the effectiveness of the SIPF-r method which maintains accuracy while supporting our assumptions on the regularity of the original KS system and the boundedness of the numerical approximation. The observed convergence rates for both the time step δt and the batch size R align closely with the theoretical predictions derived in Theorem 3.3. Our error estimates can be seen as a theoretical and computational advancement over the prior work [39], as we justify SIPF-r by providing a convergence analysis supported by numerical experiment. Future work will focus on improving the efficiency of the algorithm, particularly in high-dimensional settings, and refining error estimates, particularly the overestimated bounds for the Fourier mode H. Additionally, extending the SIPF-r method to other related systems, such as models with more complex chemo-attractant dynamics or systems involving anisotropic interactions, offers an exciting direction for future research.

Acknowledgements. ZW was partially supported by NTU SUG-023162-00001 and MOE AcRF Tier 1 Grant RG17/24. JX was partially supported by NSF grant DMS-2309520. ZZ was partially supported by the National Natural Science Foundation of China (Projects 92470103 and 12171406), the Hong Kong RGC grant (Projects 17307921 and 17304324), the Seed Funding Programme for Basic Research (HKU), the Outstanding Young Researcher Award of HKU (2020-21), and Seed Funding for Strategic Interdisciplinary Research Scheme 2021/22 (HKU).

REFERENCES

[1] N. Bellomo, A. Bellouquid, Y. Tao, and M. Winkler, *Toward a mathematical theory of Keller-Segel models of pattern formation in biological tissues*, Mathematical Models and Methods in Applied Sciences, 25 (2015), pp. 1663–1763.
[2] F. Bubba, C. Pouchol, N. Ferrand, G. Vidal, L. Almeida, B. Perthame, and M. Sabbah, *A chemotaxis-based explanation of spheroid formation in 3d cultures of breast cancer cells*, Journal of Theoretical Biology, 479 (2019), pp. 73–80.
[3] Z. Cai, J.-G. Liu, and Y. Wang, *Convergence of random batch method with replacement for interacting particle systems*, arXiv preprint arXiv:2407.19315, (2024).
[4] A. Chertock, Y. Epshteyn, H. Hu, and A. Kurganov, *High-order positivity-preserving hybrid finite-volume-finite-difference methods for chemotaxis systems*, Advances in Computational Mathematics, 44 (2018), pp. 327–350.
[5] A. Chertock and A. Kurganov, *A second-order positivity preserving central-upwind scheme for chemotaxis and haptotaxis models*, Numerische Mathematik, 111 (2008), pp. 169–205.
[6] K. Craig and A. Bertozzi, *A blob method for the aggregation equation*, Mathematics of Computation, 85 (2016), pp. 1681–1717.
[7] M. H. DeGroot, *Probability and statistics*, (1986).
[8] Y. Epshteyn, *Upwind-difference potentials method for Patlak-Keller-Segel chemotaxis model*, Journal of Scientific Computing, 53 (2012), pp. 689–713.
[9] Y. Epshteyn and A. Izmirlioglu, *Fully discrete analysis of a discontinuous finite element method for the Keller-Segel chemotaxis model*, Journal of Scientific Computing, 40 (2009), pp. 211–256.
[10] I. Fatkullin, *A study of blow-ups in the Keller-Segel model of chemotaxis*, Nonlinearity, 26 (2012), p. 81.
[11] F. Filbet, *A finite volume scheme for the Patlak-Keller-Segel chemotaxis model*, Numerische Mathematik, 104 (2006), pp. 457–488.
[12] D. Godinho and C. Quininao, *Propagation of chaos for a subcritical Keller-Segel model*, in Annales de l'IHP Probabilités et Statistiques, vol. 51, 2015, pp. 965–992.
[13] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, MIT press, 2016.
[14] J. Haškovec and C. Schmeiser, *Stochastic particle approximation for measure valued solutions of the 2d Keller-Segel system*, Journal of Statistical Physics, 135 (2009), pp. 133–151.
[15] M. A. Herrero, E. Medina, and J. J. Velázquez, *Self-similar blow-up for a reaction-diffusion system*, Journal of Computational and Applied Mathematics, 97 (1998), pp. 99–119.

[16] Z. Huang, S. Jin, and L. Li, *Mean field error estimate of the random batch method for large interacting particle system*, ESAIM: Mathematical Modelling and Numerical Analysis, 59 (2025), pp. 265–289.
[17] S. Jin and L. Li, *Random batch methods for classical and quantum interacting particle systems and statistical samplings*, in Active Particles, Volume 3: Advances in Theory, Models, and Applications, Springer, 2021, pp. 153–200.
[18] S. Jin, L. Li, and J.-G. Liu, *Random batch methods (RBM) for interacting particle systems*, Journal of Computational Physics, 400 (2020), p. 108877.
[19] E. F. Keller and L. A. Segel, *Initiation of slime mold aggregation viewed as an instability*, Journal of Theoretical Biology, 26 (1970), pp. 399–415.
[20] T. W. Körner, *Fourier analysis*, Cambridge University Press, 2022.
[21] S. Larsson and V. Thomée, *Partial differential equations with numerical methods*, vol. 45, Springer, 2003.
[22] X. H. Li, C.-W. Shu, and Y. Yang, *Local discontinuous Galerkin method for the Keller-Segel chemotaxis model*, Journal of Scientific Computing, 73 (2017), pp. 943–967.
[23] J.-G. Liu, L. Wang, and Z. Zhou, *Positivity-preserving and asymptotic preserving method for 2d Keller-Segel equations*, Mathematics of Computation, 87 (2018), pp. 1165–1189.
[24] J.-G. Liu and R. Yang, *A random particle blob method for the Keller-Segel equation and convergence analysis*, Mathematics of Computation, 86 (2017), pp. 725–745.
[25] J.-G. Liu and R. Yang, *Propagation of chaos for the Keller-Segel equation with a logarithmic cut-off*, Methods and Applications of Analysis, 26 (2019), pp. 319–348.
[26] A. I. Markushevich, *Theory of functions of a complex variable*, American Mathematical Soc., 2013.
[27] T. Nagai, *Blow-up of radially symmetric solutions to a chemotaxis system*, Adv. Math. Sci. Appl., 5 (1995), p. 581.
[28] K. J. Painter, *Mathematical models for chemotaxis and their applications in self-organisation phenomena*, Journal of Theoretical Biology, 481 (2019), pp. 162–182.
[29] C. S. Patlak, *Random walk with persistence and external bias*, The Bulletin of Mathematical Biophysics, 15 (1953), pp. 311–338.
[30] B. Perthame, *Transport equations in biology*, Springer Science & Business Media, 2006.
[31] N. Saito, *Conservative upwind finite-element method for a simplified Keller-Segel system modelling chemotaxis*, IMA Journal of Numerical Analysis, 27 (2007), pp. 332–365.
[32] N. Saito, *Error analysis of a conservative finite-element approximation for the Keller-Segel system of chemotaxis*, Communications on Pure and applied Analysis, 11 (2011), pp. 339–364.
[33] N. Saito and T. Suzuki, *Notes on finite difference schemes to a parabolic-elliptic system modelling chemotaxis*, Applied Mathematics and Computation, 171 (2005), pp. 72–90.
[34] J. A. Sherratt, *Chemotaxis and chemokinesis in eukaryotic cells: The Keller-Segel equations as an approximation to a detailed model*, Bulletin of Mathematical Biology, 56 (1994), pp. 129–146.
[35] M. Sommerfeld, J. Schrieber, Y. Zemel, and A. Munk, *Optimal transport: Fast probabilistic approximation with exact solvers*, Journal of Machine Learning Research, 20 (2019), pp. 1–23.
[36] A. Stevens, *The derivation of chemotaxis equations as limit dynamics of moderately interacting stochastic many-particle systems*, SIAM Journal on Applied Mathematics, 61 (2000), pp. 183–212.
[37] Z. Wang, J. Xin, and Z. Zhang, *DeepParticle: learning invariant measure by a deep neural network minimizing Wasserstein distance on data generated by an interacting particle method*, Journal of Computational Physics, 464 (2022), p. 111309.
[38] Z. Wang, J. Xin, and Z. Zhang, *A DeepParticle method for learning and generating aggregation patterns in multi-dimensional Keller-Segel chemotaxis systems*, Physica D: Nonlinear Phenomena, 460 (2024), p. 134082.
[39] Z. Wang, J. Xin, and Z. Zhang, *A novel stochastic interacting particle-field algorithm for 3d parabolic-parabolic Keller-Segel chemotaxis system*, Journal of Scientific Computing, 102 (2025), pp. 1–23.
[40] G. Zhou and N. Saito, *Finite volume methods for a Keller-Segel system: discrete energy, error estimates and numerical blow-up analysis*, Numerische Mathematik, 135 (2017), pp. 265–311.