

A MODIFIED CONSENSUS-BASED OPTIMIZATION MODEL: CONSENSUS FORMATION AND UNIFORM-IN-TIME PROPAGATION OF CHAOS

YOUNG-PIL CHOI, SEUNGCHAN LEE, AND SIHYUN SONG

ABSTRACT. We introduce a modified Consensus-Based Optimization model that admits a fully unified and rigorous analysis of its finite-particle dynamics, the associated McKean–Vlasov equation, and their optimization behavior under a single set of structural framework. The key ingredient is a regularized Gibbs weight that stabilizes the consensus point and avoids degeneracies present in the classical formulation, eliminating the need for cutoffs, rescaling, or boundedness assumptions on the objective function. Our first main result establishes large-time consensus for the particle system: when the drift exceeds an explicit threshold, all particles converge exponentially to a common random limit that concentrates near the global minimizer. Our second result proves uniform-in-time propagation of chaos, providing quantitative and dimension-free convergence of the empirical measure to the McKean–Vlasov dynamics. Finally, we show that the mean-field system reaches deterministic consensus and that its consensus point approaches the global minimizer in the regime of highly concentrated Gibbs weights. Together, these results yield a unified and internally consistent theoretical framework for consensus-based optimization under substantially relaxed regularity assumptions on the objective function.

CONTENTS

1. Introduction	2
1.1. Main results	5
1.2. Organization of the paper	9
2. Well-posedness and large-time analysis of the particle system	10
2.1. Existence and uniqueness of strong solutions	10
2.2. Convergence analysis in $L^p(\Omega)$	10
2.3. Almost sure convergence	14
2.4. Consensus to the minimizer	18
3. Uniform-in-time propagation of chaos	20
3.1. Preliminary estimates and stability observations	20
3.2. Exponential decay and moment estimates for the synchronous coupling	22
3.3. Concentration estimates for the synchronous coupling	24

Date: November 25, 2025.

Key words and phrases. Consensus-Based Optimization, interacting particle systems, propagation of chaos, large-time behavior, McKean–Vlasov dynamics.

3.4. Proof of Theorem 1.2	30
4. Well-posedness and large-time analysis of the mean-field system	37
4.1. Well-posedness	37
4.2. Consensus formation and convergence toward minimizers	40
Acknowledgments	42
Appendix A. Exchangability of the solutions	42
Appendix B. Well-posedness of the modified CBO system	43
Appendix C. Uniform-in-time stability of particle system	45
References	49

1. INTRODUCTION

In the broader landscape of global optimization, metaheuristic algorithms such as random search [38], simulated annealing [35, 44], genetic algorithms [27], and particle swarm optimization [33, 34] have been widely used to tackle high-dimensional and nonconvex problems where gradient information is unavailable or unreliable. These approaches combine stochastic exploration and local exploitation mechanisms, often inspired by collective behavior or evolutionary processes, to overcome the limitations of deterministic optimization methods. However, while they have proven empirically successful, most of them lack rigorous mathematical guarantees of global convergence, primarily due to their algorithmic complexity and intrinsic stochasticity [3, 4, 5].

A major step toward a mathematically tractable metaheuristic was taken with the introduction of the *Consensus-Based Optimization (CBO)* framework in [36]. This model describes a system of interacting particles that collectively search for a global minimizer through attraction toward a consensus point, which is defined as a Gibbs-weighted average of their positions. The original formulation in [36] already incorporated stochastic effects, but the rigorous analysis was first carried out for a simplified deterministic setting, where the discontinuous drift switch was removed and the stochastic diffusion term was neglected for analytical tractability. The corresponding stochastic model was later analyzed at the mean-field level in [9], which established convergence to global minimizers under suitable assumptions and initiated the PDE-based theory of CBO dynamics. These foundational works laid the analytical basis for interpreting CBO as a stochastic interacting-particle algorithm performing an effective “convexification” of nonconvex landscapes in the mean-field limit. For an overview of recent advances in CBO, we refer the reader to [7, 22, 42] and the references therein.

In the following years, several studies further advanced the analytical understanding of the CBO model. The high-dimensional formulation introduced in [11] established convergence results in settings where each particle evolves under its own independent Brownian perturbation, providing a dimension-free analytical framework. Other works focused on a componentwise multiplicative noise structure based on the Hadamard product, in which the diffusion acts anisotropically along each coordinate

direction; see, for example, [24, 25, 45]. In this formulation, each particle evolves under independent stochastic perturbations, and the multiplicative nature of the noise allows for more flexible but technically involved dynamics. In parallel, a number of generalizations have been proposed to incorporate additional stochastic or geometric mechanisms, such as constrained CBO models [1, 6, 12], polarization-based models [8], and mirror-descent-type variants [7]. Further developments include formulations of CBO on manifolds [16, 17, 18, 26], consensus-based sampling methods [10], rescaled or truncated-noise systems [28, 29, 30, 20], and variants with Lévy noise or memory effects [32, 43].

The mean-field limit of CBO toward the corresponding McKean–Vlasov process has been rigorously studied in several works, though a major analytical challenge arises from the lack of regularity and nonlinearity of the consensus point. Quantitative finite-time mean-field limit results were obtained in [23, 16, 17, 18, 19, 32], while qualitative convergence over finite horizons was investigated in [31]. Uniform-in-time mean-field limit estimates have recently been proved for modified variants of CBO, including the rescaled formulation [29], the classical model under bounded objective functions [22], and cutoff-based systems ensuring bounded particle configurations [2]. These developments have considerably deepened the theoretical understanding of CBO, but most existing analyses rely on strong assumptions on the objective function f , such as global boundedness, Lipschitz regularity, or convexity, or on structural modifications of the dynamics, such as truncation, rescaling, or compactification of the noise.

In the present work, we propose a *modified CBO system* that addresses these limitations while preserving the essential consensus-seeking mechanism of the original model. Our formulation enables a fully dimension-free rigorous analysis without imposing upper bounds on f or introducing artificial cutoffs in the noise. The main novelty lies in the introduction of a *regularized Gibbs weight*, which stabilizes the consensus point and simplifies the stochastic structure of the system. Unlike previous formulations that rely on nonlinear or multiplicative diffusion terms, our model features a linear diffusion aligned with the consensus direction, yielding cleaner energy estimates and uniform-in-time propagation of chaos bounds independent of the space dimension.

Mathematically, we work on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ and consider the interacting particle system

$$\begin{cases} dX_t^i = -\lambda(X_t^i - m_t^h) dt + \sigma(X_t^i - m_t^h) dW_t^i, & i = 1, 2, \dots, N, \\ m_t^h := \frac{\sum_{i=1}^N X_t^i \psi_h(X_t^i)}{\sum_{i=1}^N \psi_h(X_t^i)}, \end{cases} \quad (1.1)$$

where $\lambda, \sigma > 0$ are positive parameters representing, respectively, the strength of drift toward consensus and the intensity of stochastic diffusion, and $\{W_t^i\}_{i=1}^N$ are independent one-dimensional Wiener processes. The term m_t^h represents the consensus point toward which all particles are attracted. It is constructed as a weighted average of the particle positions with respect to the regularized Gibbs weight

$$\psi_h(x) := \omega_f^\alpha(x) + h(\alpha), \quad \omega_f^\alpha(x) = e^{-\alpha f(x)}, \quad \alpha > 0,$$

where h is a strictly positive function of α . This additional term $h(\alpha)$ guarantees that the denominator in the definition of m_t^h remains uniformly bounded away from zero, thereby removing possible singularities and stabilizing the dynamics. The parameter λ will later be required to satisfy a quantitative

stability threshold ensuring exponential contraction of trajectories. By construction, the dynamics are symmetric with respect to particle indices. In particular, when the initial data $\{X_0^i\}_{i=1}^N$ are i.i.d. and the driving Wiener processes are independent, the unique strong solution $\{X_t^i\}_{i=1}^N$ remains exchangeable for all $t \geq 0$, since both the drift and diffusion coefficients depend only on the empirical measure. A short proof of this property is provided in Appendix A.

There is a key structural improvement behind our modification. In contrast to the classical CBO model, whose consensus point

$$m_t^0 := \frac{\sum_{i=1}^N X_t^i \omega_f^\alpha(X_t^i)}{\sum_{i=1}^N \omega_f^\alpha(X_t^i)}$$

can become unstable when the Gibbs weights $\omega_f^\alpha(X_t^i)$ are exceedingly small, our regularized formulation introduces a stabilizing term $h(\alpha)$ in the weight $\psi_h(x) = \omega_f^\alpha(x) + h(\alpha)$, ensuring that the denominator of m_t^h remains uniformly bounded away from zero. Equivalently, the new consensus point admits the interpolative representation

$$m_t^h = \beta_t m_t^0 + (1 - \beta_t) \left(\frac{1}{N} \sum_{i=1}^N X_t^i \right), \quad \beta_t := \frac{\sum_{i=1}^N \omega_f^\alpha(X_t^i)}{\sum_{i=1}^N \psi_h(X_t^i)}.$$

This structure provides a powerful self-regulating mechanism: when the Gibbs weights dominate ($\beta_t \approx 1$), the dynamics closely follow the classical CBO flow, whereas when the weights become negligible ($\beta_t \approx 0$), the system automatically shifts toward the empirical average, thereby maintaining stability and avoiding premature collapse to a poor local region. This interpolation between the weighted consensus and the empirical mean not only prevents numerical and analytical degeneracies but also enhances the exploration capability of the model, allowing particles to escape shallow traps in the optimization landscape. Consequently, our modification yields a model that is both analytically well-posed and algorithmically more robust than the classical CBO formulation, and it provides a unified framework in which the large-time behavior of the particle system, its mean-field limit, and the optimization properties of the McKean–Vlasov dynamics can all be rigorously analyzed under the same structural assumptions on f .

Throughout the paper, we assume that the objective function f is a locally Lipschitz function and attains a unique positive global minimizing point, and in particular

$$\exists! x_* \in \mathbb{R}^d \text{ for which } f(x_*) = \underline{f} := \inf_{x \in \mathbb{R}^d} f(x) > 0.$$

The local Lipschitz property of f is not explicitly used in this work, but we impose it as it guarantees the existence of a strong solution to (1.1) (see Appendix B). Further, we assume that the strictly positive function h satisfies

$$\limsup_{\alpha \rightarrow \infty} \frac{e^{-\alpha \underline{f}}}{h(\alpha)} < \infty,$$

which guarantees that the constant Λ_α which appears in the main results is finite. This imposes no significant restriction, as other than the above condition (and positivity) the choice of h is otherwise flexible.

1.1. Main results. The analysis developed in this work consists of three complementary components:

- the large-time behavior and optimization properties of the finite particle system,
- the uniform-in-time propagation of chaos linking the particle and mean-field descriptions, and
- the large-time behavior and optimization properties of the associated McKean–Vlasov equation.

Unlike much of the existing literature, where these aspects are treated separately and under different structural assumptions on the objective function f , our approach provides a unified and internally consistent framework in which all three results follow from the same core hypotheses. Within this framework, the particle dynamics, the mean-field limit, and the optimization behavior of f are derived under a common set of assumptions, ultimately leading to the convergence of the particle system toward the global minimizer of f .

We begin with the large-time analysis of the particle trajectories. This part relies on a stochastic stability framework that establishes both L^p and almost sure convergence of the system under suitable conditions on the interaction strength λ (which controls the drift toward the consensus point), the noise amplitude σ , and the modified weight function ψ_h . Our goal in this part is to characterize the asymptotic behavior of the system and identify conditions under which all particles converge toward a common limit, the consensus state, that subsequently concentrates near the global minimizer of f . The analysis of the large-time behavior requires only mild assumptions on the initial data and a lower bound on the interaction strength λ , quantified by the threshold condition $\lambda > \lambda_{p,\alpha,\sigma}$ (see Theorem 1.1 for its precise expression). This condition provides exponential decay of pairwise distances between particles and guarantees that all trajectories asymptotically collapse to a common random limit. To further connect this consensus state to the minimization of f , we impose additional regularity conditions on f , ensuring sufficient smoothness and growth control of its gradient and Hessian. These allow the application of Itô's formula to the weighted energy functional $\mathbb{E}\omega_f^\alpha(X_t^i)$, which plays a crucial role in linking the stochastic consensus dynamics to the optimization landscape of f .

Assumption 1.1. *The map $x \mapsto \omega_f^\alpha(x) := e^{-\alpha f(x)}$ is Lipschitz continuous, and its Lipschitz constant $L_{\omega_f^\alpha}$ satisfies $\lim_{\alpha \rightarrow \infty} L_{\omega_f^\alpha} = 0$.*

Assumption 1.2. *We assume $f \in \mathcal{C}^2(\mathbb{R}^d)$, and there exist constants $c_0, c_1 \geq 0$ such that*

$$\nabla^2 f(x) \preceq c_0 \mathbb{I}_d + c_1 \nabla f(x) [\nabla f(x)]^\top,$$

in other words

$$c_0 \mathbb{I}_d + c_1 \nabla f(x) [\nabla f(x)]^\top - \nabla^2 f(x) \text{ is positive semi-definite } \quad \forall x \in \mathbb{R}^d.$$

Prior to stating our main result, we leave some remarks on the diverse range of objective functions which satisfy the proposed assumptions above.

Remark 1.1. *Assumption 1.1 is, for instance, readily satisfied in the following cases:*

- f is any even-degree polynomial (with $\underline{f} > 0$). Indeed, in this case we can write $f(x) = p(x) + \underline{f}$ where $p \geq 0$ is an even polynomial. Note that the explicit expression of $L_{\omega_f^\alpha}$ is

$$L_{\omega_f^\alpha} = \alpha \cdot \sup_{x \in \mathbb{R}^d} |\nabla f(x) e^{-\alpha f(x)}| = \alpha e^{-\alpha \underline{f}} \sup_{x \in \mathbb{R}^d} |\nabla p(x) e^{-\alpha p(x)}|.$$

Since the supremum on the right-hand side is uniformly in α bounded for any p , we deduce that the right-hand side tends to zero as $\alpha \rightarrow \infty$.

- f is Lipschitz continuous. Indeed, the same expression above shows that

$$L_{\omega_f^\alpha} \leq \|\nabla f\|_{L^\infty} \alpha e^{-\alpha \underline{f}} \xrightarrow{\alpha \rightarrow \infty} 0.$$

Remark 1.2. Assumption 1.2 is automatically satisfied if, for instance, the objective function f has globally bounded second derivatives. Moreover, both Assumptions 1.1 and 1.2 are fulfilled by several standard benchmark functions used in global optimization. A representative example is the Rastrigin function (see, e.g., [36]), defined by

$$f_R(x) = \frac{1}{d} \sum_{i=1}^d [(x_i - B)^2 - 10 \cos(2\pi(x_i - B)) + 10] + C,$$

where $B, C \in \mathbb{R}$ are constant shifts. In this case, the function's linear gradient growth (resulting from its quadratic structure) and bounded Hessian ensure that Assumptions 1.1 and 1.2 are fulfilled.

The above structural conditions provide quantitative control of the energy evolution and form the basis for our first main result stated below. We emphasize that, in the setting of independent noises, to the best of our knowledge, our work provides the first rigorous proof of consensus in a CBO model without any cutoff in the noise or rescaling of the consensus point.

Theorem 1.1 (Large-time consensus and convergence to the minimizer). *Let $\{X_t^i\}_{i=1}^N$ be a global strong solution to the system (1.1) with i.i.d. initial data $\{X_0^i\}_{i=1}^N$, distributed as $X_{\text{in}} \in L^p(\Omega)$ for some $p \geq 2$. Define the consensus threshold*

$$\lambda_{p,\alpha,\sigma} := (p-1)\Lambda_\alpha^{\frac{2}{p}}\sigma^2 \quad \text{with } \Lambda_\alpha := \frac{e^{-\alpha \underline{f}} + h(\alpha)}{h(\alpha)}.$$

If $\lambda > \lambda_{p,\alpha,\sigma}$, then the following hold:

- (i) (Large-time convergence) *There exists a random vector $X_\infty = X_\infty^{N,\alpha} \in L^p(\Omega)$ such that*

$$X_t^i \xrightarrow[t \rightarrow \infty]{} X_\infty \quad \text{in } L^p(\Omega) \text{ for all } i = 1, \dots, N.$$

If, in addition, $p > 2$, the convergence holds almost surely.

- (ii) (Consensus to the minimizer) *Suppose furthermore that Assumptions 1.1 and 1.2 hold, and that $\lambda > \lambda_{p,\alpha,\sigma} \vee \lambda_{2,\alpha,\sigma}$. If there exists $0 < \epsilon \leq 1$ such that*

$$(1 - \epsilon)\mathbb{E}\omega_f^\alpha(X_{\text{in}}) \geq \frac{\lambda L_{\omega_f^\alpha}}{\lambda - \lambda_{2,\alpha,\sigma}} (2\Lambda_\alpha \text{Var}(X_{\text{in}}))^{\frac{1}{2}} + \frac{\alpha\sigma^2 c_0 \Lambda_\alpha e^{-\alpha \underline{f}}}{2(\lambda - \lambda_{2,\alpha,\sigma})} \text{Var}(X_{\text{in}})$$

then, the limit variable X_∞ satisfies

$$\text{ess inf}_{\omega \in \Omega} f(X_\infty(\omega)) \leq \text{ess inf}_{\omega \in \Omega} f(X_{\text{in}}(\omega)) + o(1) \quad (\alpha \rightarrow \infty).$$

Consequently, if the global minimizer x^ of f lies in $\text{supp}(\text{Law}(X_{\text{in}}))$, then*

$$\text{ess inf}_{\omega \in \Omega} f(X_\infty(\omega)) \leq \underline{f} + o(1) \quad (\alpha \rightarrow \infty),$$

showing that the consensus point X_∞ asymptotically approaches near the global minimizer of f .

We now turn to our second main result, which concerns the uniform-in-time propagation of chaos for the particle system (1.1). This result provides a quantitative and dimension-free description of the convergence of the N -particle dynamics toward its corresponding McKean–Vlasov equation as $N \rightarrow \infty$, thereby establishing a rigorous mean-field limit for the modified CBO model. It complements the large-time analysis developed in the previous section by connecting the microscopic particle description with its macroscopic mean-field counterpart.

Our approach is inspired by the framework of [22], which established uniform-in-time propagation of chaos for the original CBO equations through a novel interpretation of the weighted mean (see Section 3.1). Here, we extend that theory to our modified formulation under significantly weaker analytical assumptions. In particular, we do not require the objective function f to be upper bounded, and it suffices to assume that the Gibbs weight $\omega_f^\alpha = e^{-\alpha f}$ is Lipschitz continuous, instead of f itself. Moreover, our analysis relaxes the moment constraints on the initial data: when moments strictly higher than four are imposed, we obtain a quantitative decay rate for the 2-Wasserstein distance between the empirical and mean-field measures. Under higher-order moment conditions (e.g., eighth moments), we recover the $\frac{1}{N}$ convergence rate of [22], but under less restrictive assumptions on f .

We denote by $\bar{X}_t$ the solution to the following McKean–Vlasov equation, driven by the one-dimensional Wiener process $\{W_t\}_{t \geq 0}$:

$$\begin{cases} \bar{X}_t = \bar{X}_0 - \lambda \int_0^t (\bar{X}_s - m_f^h[\bar{\rho}_s]) ds + \sigma \int_0^t (\bar{X}_s - m_f^h[\bar{\rho}_s]) dW_s, \\ m_f^h[\rho] = \frac{\int_{\mathbb{R}^d} x \psi_h(x) \rho(dx)}{\int_{\mathbb{R}^d} \psi_h(x) \rho(dx)}, \quad \bar{\rho}_t := \text{Law}(\bar{X}_t). \end{cases} \quad (1.2)$$

The law $\bar{\rho}_t$ satisfies the nonlinear Fokker–Planck equation

$$\partial_t \bar{\rho}_t = \nabla \cdot (\lambda(x - m_f^h[\bar{\rho}_t]) \bar{\rho}_t) + \frac{\sigma^2}{2} \nabla^2 : [(x - m_f^h[\bar{\rho}_t]) \otimes (x - m_f^h[\bar{\rho}_t]) \bar{\rho}_t]. \quad (1.3)$$

Theorem 1.2 (Uniform-in-time propagation of chaos). *Let $p \geq 2$, $q > 2$, and $\alpha > 0$. Let $\{X_0^i\}_{i=1}^N$ be i.i.d. random variables with common law $\bar{\rho}_0 \in \mathcal{P}_{pq}(\mathbb{R}^d)$, and let $\{X_t^i\}_{i=1}^N$ be the strong solution to the particle system (1.1) with initial data $\{X_0^i\}_{i=1}^N$. Let $\bar{\rho}_t$ be the law of the McKean–Vlasov process $\bar{X}_t$ solving (1.2) with initial law $\bar{\rho}_0$. Assume that f satisfies Assumption 1.1. Then for any*

$$\lambda > \max \left\{ (pq - 1)(1 + \Lambda_\alpha^{\frac{2}{pq}}) \sigma^2 + 2(p - 1)(3 + \Lambda_\alpha^{\frac{2}{p}}) \sigma^2, (pq - 1) \Lambda_\alpha^{\frac{2}{pq}} \sigma^2 + 2(p - 1)(1 + \Lambda_\alpha) \sigma^2 \right\},$$

there exists a positive constant $C > 0$, depending only on $\lambda, \sigma, p, q, \Lambda_\alpha$ and $\bar{\rho}_0$, such that for all $t \geq 0$,

$$\begin{aligned} \mathbb{E} \mathcal{W}_p(\mu_t^N, \bar{\rho}_t) &\leq C \mathbb{E} \mathcal{W}_p(\mu_0^N, \bar{\rho}_0) + \frac{C}{N^{\frac{1}{2} \wedge \frac{q-2}{2p}}} \\ &\quad + C \begin{cases} N^{-\frac{1}{2}} + N^{-\frac{q-1}{q}} & \text{if } p > \frac{d}{2}, \\ N^{-\frac{1}{2}} \log(1 + N) + N^{-\frac{q-1}{q}} & \text{if } p = \frac{d}{2}, \\ N^{-\frac{p}{d}} + N^{-\frac{q-1}{q}} & \text{if } p \in (0, \frac{d}{2}) \text{ and } pq \neq \frac{d}{d-p}. \end{cases} \end{aligned} \quad (1.4)$$

Here $\mathcal{W}_p$ represents the p -Wasserstein distance, and $\mu_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^i}$ is the empirical measure associated with the particle system (1.1).

Remark 1.3. If the initial data are chosen synchronously, i.e., $\bar{X}_0^i = X_0^i$ almost surely for all $i = 1, \dots, N$, then the first term on the right-hand side of (1.4) vanishes. Consequently, the right-hand side decays to zero as $N \rightarrow \infty$, uniformly in time. This yields the convergence in probability

$$\sup_{t \geq 0} \lim_{N \rightarrow \infty} \mathbb{P}(\mathcal{W}_p(\mu_t^N, \bar{\rho}_t) \geq \epsilon) = 0 \quad \text{for any } \epsilon > 0,$$

which establishes the uniform-in-time propagation of chaos for the system. This argument parallels the classical mean-field limit framework (see [13, Lemma 3.19] or [41, Proposition 2.2]), with the essential improvement that the convergence obtained here is uniform in time, owing to the uniform-in-time stability estimate proved in Theorem 1.2.

We now complement the particle-level analysis by studying the corresponding mean-field dynamics. While the proof strategy parallels that of the particle system, relying on a contraction mechanism toward a weighted mean, the McKean–Vlasov equation enjoys two structural simplifications that lead to a sharper large-time description. First, the dynamics are driven by a single Wiener process and a self-consistent drift, so no fluctuation between particles remains at the mean-field level. Second, the contraction acts directly on the law $\bar{\rho}_t$, yielding a closed evolution for its variance. As a consequence, once λ exceeds an explicit threshold (the analogue of $\lambda_{p,\alpha,\sigma}$), the variance $\text{Var}(\bar{X}_t)$ decays uniformly in time, and $\bar{X}_t$ converges to a deterministic limit x_∞^α . Under the same structural assumptions on f used in the particle-level analysis (Assumptions 1.1–1.2), the weighted energy $\|\omega_f^\alpha\|_{L^1(\bar{\rho}_t)}$ again provides a quantitative link between the consensus point and the optimization landscape. This yields an explicit bound on $f(x_\infty^\alpha)$ and shows that, in the regime $\alpha \rightarrow \infty$, the mean-field consensus point approaches the global minimizer of f .

We summarize these conclusions in the following theorem.

Theorem 1.3 (Mean-field consensus and convergence to the minimizer). *Let $\bar{X}_t$ denote a global solution to the McKean–Vlasov equation (1.2) with initial law $\bar{\rho}_0 \in \mathcal{P}_p(\mathbb{R}^d)$ for some $p \geq 2$. Define the mean-field consensus threshold*

$$\bar{\lambda}_{p,\alpha,\sigma} := \lambda_{p,\alpha,\sigma} + (p-1)\sigma^2 = (p-1)(1 + \Lambda_\alpha^{2/p})\sigma^2.$$

If $\lambda > \bar{\lambda}_{p,\alpha,\sigma}$, then the following hold:

- (i) (Deterministic consensus) *There exists a deterministic point $x_\infty^\alpha \in \mathbb{R}^d$ such that*

$$\bar{X}_t \xrightarrow[t \rightarrow \infty]{} x_\infty^\alpha \quad \text{in } L^p(\Omega),$$

or equivalently,

$$\mathcal{W}_p(\bar{\rho}_t, \delta_{x_\infty^\alpha}) \xrightarrow[t \rightarrow \infty]{} 0.$$

- (ii) (Consensus to minimizer) *Assume in addition that f satisfies Assumptions 1.1 and 1.2. If there exists $0 < \epsilon \leq 1$ such that*

$$(1 - \epsilon)\|\omega_f^\alpha\|_{L^1(\bar{\rho}_0)} \geq \frac{\lambda L_{\omega_f^\alpha} \Lambda_\alpha}{\lambda - \bar{\lambda}_{2,\alpha,\sigma}} \text{Var}(\bar{X}_0)^{\frac{1}{2}} + \frac{\alpha \sigma^2 c_0 e^{-\alpha f} \Lambda_\alpha}{2(\lambda - \bar{\lambda}_{2,\alpha,\sigma})} \text{Var}(\bar{X}_0),$$

then the limit point x_∞^α satisfies

$$f(x_\infty^\alpha) \leq \inf_{x \in \text{supp}(\bar{\rho}_0)} f(x) + o(1) \quad (\alpha \rightarrow \infty).$$

Consequently, if the global minimizer x^* of f lies in $\text{supp}(\bar{\rho}_0)$, then

$$f(x_\infty^\alpha) \leq \underline{f} + o(1) \quad (\alpha \rightarrow \infty).$$

Asymptotic convergence to the global minimizer. Finally, we describe how the three main theorems established above combine to yield the convergence of the N -particle system to the unique global minimizer $x^* \in \text{supp}(\bar{\rho}_0)$ in the joint limit $t \rightarrow \infty$, $N \rightarrow \infty$, and $\alpha \rightarrow \infty$.

Let μ_t^N denote the empirical distribution of the particle system (1.1), and let $\bar{\rho}_t$ be the solution to the McKean–Vlasov equation (1.2). For any $t \geq 0$, we decompose

$$\begin{aligned} \mathbb{E}\mathcal{W}_p(\mu_t^N, \delta_{x^*}) &\leq \mathbb{E}\mathcal{W}_p(\mu_t^N, \bar{\rho}_t) + \mathcal{W}_p(\bar{\rho}_t, \delta_{x_\infty^\alpha}) + \mathcal{W}_p(\delta_{x_\infty^\alpha}, \delta_{x^*}) \\ &= \mathbb{E}\mathcal{W}_p(\mu_t^N, \bar{\rho}_t) + \mathcal{W}_p(\bar{\rho}_t, \delta_{x_\infty^\alpha}) + |x_\infty^\alpha - x^*|, \end{aligned}$$

where the last equality follows from the identity $\mathcal{W}_p(\delta_a, \delta_b) = |a - b|$.

By Theorem 1.3 (i), the mean-field dynamics converge to the deterministic limit x_∞^α as $t \rightarrow \infty$. Similarly, by Theorem 1.1 (i), the particle system converges to a random consensus point $X_\infty^{N,\alpha}$. Letting $t \rightarrow \infty$ in the inequality above yields

$$\mathbb{E}|X_\infty^{N,\alpha} - x^*| = \mathbb{E}\mathcal{W}_p(\delta_{X_\infty^{N,\alpha}}, \delta_{x^*}) \leq \sup_{t \geq 0} \mathbb{E}\mathcal{W}_p(\mu_t^N, \bar{\rho}_t) + |x_\infty^\alpha - x^*|.$$

Assuming the initial data of the particle system are synchronously coupled, we may apply Theorem 1.2 and take $\limsup_{N \rightarrow \infty}$, obtaining

$$\limsup_{N \rightarrow \infty} \mathbb{E}|X_\infty^{N,\alpha} - x^*| \leq |x_\infty^\alpha - x^*|.$$

Finally, Theorem 1.3 (ii) shows that $x_\infty^\alpha \rightarrow x^*$ as $\alpha \rightarrow \infty$, provided $x^* \in \text{supp}(\bar{\rho}_0)$. Taking $\limsup_{\alpha \rightarrow \infty}$ then gives the full asymptotic convergence,

$$\limsup_{\alpha \rightarrow \infty} \limsup_{N \rightarrow \infty} \mathbb{E}|X_\infty^{N,\alpha} - x^*| = 0.$$

This establishes that, under appropriate scaling of N and α , the interacting particle system concentrates asymptotically near the global minimizer x^* of f .

1.2. Organization of the paper. The rest of the paper is organized as follows. In Section 2, we establish Theorem 1.1. We first prove global well-posedness of the finite-particle SDE system (the complete proof is given in the appendix), and then derive quantitative L^p -contraction and uniform moment bounds leading to the convergence stated in Theorem 1.1 (i). Under additional regularity assumptions, we obtain a weighted-energy estimate via Itô's formula and the Laplace principle to conclude Theorem 1.1 (ii). Section 3 is devoted to the uniform-in-time propagation of chaos. We develop a stability framework comparing the particle and mean-field systems, establish differential inequalities for fluctuation energies, and combine them with probabilistic concentration estimates to prove Theorem 1.2. Section 4 addresses the well-posedness and large-time analysis of the mean-field system. We first show that the McKean–Vlasov equation admits a unique strong solution. We then prove deterministic consensus formation using uniform decay estimates and finally establish quantitative convergence of the consensus point toward global minimizers, thereby proving Theorem 1.3. The appendices collect auxiliary results.

Appendix A verifies exchangeability of the finite-particle dynamics, Appendix B provides the detailed well-posedness proof and uniform moment bound, and Appendix C contains the concentration and stability estimates for the particle system.

2. WELL-POSEDNESS AND LARGE-TIME ANALYSIS OF THE PARTICLE SYSTEM

This section proves Theorem 1.1. We proceed in three steps: (i) well-posedness of the finite-particle SDE system; (ii) quantitative $L^p(\Omega)$ contraction of pairwise distances and uniform L^p bounds, yielding the L^p (and, for $p > 2$, almost sure) convergence asserted in Theorem 1.1 (i); (iii) a weighted-energy estimate based on Itô's formula under Assumptions 1.1 and 1.2, which, together with Laplace's principle, implies Theorem 1.1 (ii).

2.1. Existence and uniqueness of strong solutions. We begin by establishing the well-posedness of the finite-particle system (1.1), which serves as the foundation for all subsequent estimates. Since the analysis that follows relies on pathwise properties of $\{X_t^i\}_{i=1}^N$, we first confirm the global existence and uniqueness of strong solutions. Although the argument parallels that of [9, Theorem 2.1], we provide a complete proof in Appendix B for the sake of self-containment.

Theorem 2.1 (Global well-posedness). *Let $N \in \mathbb{N}$ and $d \geq 1$. For each $i = 1, \dots, N$, let $X_t^i : \Omega \rightarrow \mathbb{R}^d$ satisfy the SDE system (1.1) with initial data $X_0^i \in L^2(\Omega; \mathbb{R}^d)$. Then the system admits a unique strong solution $\{X_t^i\}_{t \geq 0}$ satisfying*

$$\mathbb{E}|X_t^i|^2 < \infty \quad \text{for all } t \geq 0.$$

2.2. Convergence analysis in $L^p(\Omega)$. In this subsection, we investigate the large-time behavior of the system in the $L^p(\Omega)$ sense. Our goal is to show that, under a suitable threshold condition on the interaction strength λ , the pairwise distances between particles decay exponentially in time. This contraction property forms the cornerstone of the proof of Theorem 1.1 (i), since it implies convergence of all trajectories to a common random limit, the consensus state.

A key simplification arises by expressing the dynamics in terms of relative positions. The next lemma quantifies the L^p -contraction of pairwise distances.

Lemma 2.2. *Let $p \geq 2$ and assume $\{X_0^i\}_{i=1}^N \subset L^p(\Omega)$. Then for each $1 \leq i, j \leq N$ and $t \geq 0$, we have $X_t^i - X_t^j \in L^p(\Omega)$ and*

$$\mathbb{E}|X_t^i - X_t^j|^p \leq \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p e^{-p(\lambda - (p-1)\Lambda_\alpha^{2/p}\sigma^2)t},$$

where

$$\Lambda_\alpha = \frac{e^{-\alpha f} + h(\alpha)}{h(\alpha)}.$$

Consequently, if $\lambda > \lambda_{p,\alpha,\sigma} = (p-1)\Lambda_\alpha^{2/p}\sigma^2$, then

$$X_t^i - X_t^j \xrightarrow[t \rightarrow \infty]{} 0 \quad \text{in } L^p(\Omega).$$

In particular, we have

$$\max_{1 \leq i \leq N} \mathbb{E}|X_t^i - m_t^h|^p \leq \Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p e^{-p(\lambda - \lambda_{p,\alpha,\sigma})t}. \quad (2.1)$$

Proof. We begin with the equation for $X_t^i - X_t^j$:

$$d(X_t^i - X_t^j) = -\lambda(X_t^i - X_t^j)dt + \sigma(X_t^i - m_t^h)dW_t^i - \sigma(X_t^j - m_t^h)dW_t^j.$$

Applying the Itô formula to $|X_t^i - X_t^j|^p$ yields

$$\begin{aligned} d|X_t^i - X_t^j|^p &= -p\lambda|X_t^i - X_t^j|^{p-1}dt + \frac{p\sigma^2}{2}|X_t^i - X_t^j|^{p-2} \left(|X_t^i - m_t^h|^2 + |X_t^j - m_t^h|^2 \right) dt \\ &\quad + \frac{p(p-2)\sigma^2}{2}|X_t^i - X_t^j|^{p-4} \langle X_t^i - X_t^j, X_t^i - m_t^h \rangle^2 dt \\ &\quad + \frac{p(p-2)\sigma^2}{2}|X_t^i - X_t^j|^{p-4} \langle X_t^i - X_t^j, X_t^j - m_t^h \rangle^2 dt \\ &\quad + p\sigma|X_t^i - X_t^j|^{p-2} \langle X_t^i - X_t^j, X_t^i - m_t^h \rangle dW_t^i \\ &\quad - p\sigma|X_t^i - X_t^j|^{p-2} \langle X_t^i - X_t^j, X_t^j - m_t^h \rangle dW_t^j. \end{aligned}$$

Taking expectations eliminates the martingale terms and gives

$$\begin{aligned} \frac{d}{dt}\mathbb{E}|X_t^i - X_t^j|^p &= -p\lambda\mathbb{E}|X_t^i - X_t^j|^{p-1} + \frac{p\sigma^2}{2}\mathbb{E} \left[|X_t^i - X_t^j|^{p-2} \left(|X_t^i - m_t^h|^2 + |X_t^j - m_t^h|^2 \right) \right] \\ &\quad + \frac{p(p-2)\sigma^2}{2}\mathbb{E} \left[|X_t^i - X_t^j|^{p-4} \langle X_t^i - X_t^j, X_t^i - m_t^h \rangle^2 \right] \\ &\quad + \frac{p(p-2)\sigma^2}{2}\mathbb{E} \left[|X_t^i - X_t^j|^{p-4} \langle X_t^i - X_t^j, X_t^j - m_t^h \rangle^2 \right]. \end{aligned}$$

Next, from the definition of m_t^h and the bounds $h(\alpha) \leq \psi_h \leq e^{-\alpha f} + h(\alpha)$, we obtain by Hölder's inequality

$$\begin{aligned} \mathbb{E}|X_t^i - m_t^h|^p &= \mathbb{E} \left| \frac{\sum_{k=1}^N (X_t^i - X_t^k) \psi_h(X_t^k)}{\sum_{k=1}^N \psi_h(X_t^k)} \right|^p \\ &\leq \mathbb{E} \left| \frac{(\sum_k |X_t^i - X_t^k|^p \psi_h(X_t^k))^{1/p} (\sum_k \psi_h(X_t^k))^{(p-1)/p}}{\sum_k \psi_h(X_t^k)} \right|^p \\ &\leq \frac{\Lambda_\alpha}{N} \sum_{k=1}^N \mathbb{E}|X_t^i - X_t^k|^p \\ &\leq \Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_t^i - X_t^j|^p. \end{aligned} \tag{2.2}$$

Using this bound and again applying Hölder's inequality, we have for $k = i, j$:

$$\begin{aligned} \mathbb{E} \left[|X_t^i - X_t^j|^{p-2} |X_t^k - m_t^h|^2 \right] &\leq \left(\mathbb{E}|X_t^i - X_t^j|^p \right)^{\frac{p-2}{p}} \left(\mathbb{E}|X_t^k - m_t^h|^p \right)^{\frac{2}{p}} \\ &\leq \left(\mathbb{E}|X_t^i - X_t^j|^p \right)^{\frac{p-2}{p}} \left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_t^i - X_t^j|^p \right)^{\frac{2}{p}} \\ &\leq \Lambda_\alpha^{\frac{2}{p}} \max_{1 \leq i, j \leq N} \mathbb{E}|X_t^i - X_t^j|^p. \end{aligned}$$

Substituting these estimates into the previous differential inequality, we arrive at

$$\frac{d}{dt}\mathbb{E}|X_t^i - X_t^j|^p \leq -p\lambda\mathbb{E}|X_t^i - X_t^j|^{p-1} + p(p-1)\Lambda_\alpha^{\frac{2}{p}}\sigma^2 \max_{1 \leq i, j \leq N} \mathbb{E}|X_t^i - X_t^j|^p.$$

Applying Grönwall's lemma to the maximal pair (i, j) gives

$$\max_{1 \leq i, j \leq N} \mathbb{E}|X_t^i - X_t^j|^p \leq \max_{1 \leq i, j \leq N} \mathbb{E}|X_t^i - X_t^j|^p e^{-p(\lambda - (p-1)\Lambda_\alpha^{2/p}\sigma^2)t}.$$

Finally, combining this with (2.2) deduces the desired estimate (2.1). $\square$

Next, we show that the particles themselves remain uniformly bounded in $L^p(\Omega)$ for all times. In particular, Lemma 2.2 already guarantees exponential decay of the inter-particle distances in $L^p(\Omega)$, and we now use this property to derive a uniform bound for each trajectory X_t^i .

Lemma 2.3. *Let $p \geq 2$. If $\lambda > \lambda_{p,\alpha,\sigma}$ and $\{X_0^i\}_{i=1}^N \subset L^p(\Omega)$, then we have*

$$X_t^i \in L^\infty((0, \infty); L^p(\Omega)) \quad \text{uniformly in } i.$$

More specifically, there is a constant C_* dependent only on $X_0^i, \lambda, p, \Lambda_\alpha$, and σ such that

$$\max_{1 \leq i \leq N} \sup_{t \geq 0} \mathbb{E}|X_t^i|^p \leq C_*.$$

Proof. Using the equation for each particle X_t^i and the convexity of $z \mapsto z^p$, we estimate

$$\begin{aligned} |X_t^i|^p &= \left| X_0^i - \lambda \int_0^t (X_s^i - m_s^h) ds + \sigma \int_0^t (X_s^i - m_s^h) dW_s^i \right|^p \\ &\leq 3^{p-1} |X_0^i|^p + 3^{p-1} \lambda^p \left(\int_0^t |X_s^i - m_s^h| ds \right)^p + 3^{p-1} \sigma^p \left| \int_0^t (X_s^i - m_s^h) dW_s^i \right|^p. \end{aligned}$$

Here, the stochastic integral

$$M_t := \int_0^t (X_s^i - m_s^h) dW_s^i$$

is a continuous $\mathbb{P}$ -martingale. Indeed, using Itô's isometry and Lemma 2.2, we obtain

$$\mathbb{E}|M_t|^2 = \mathbb{E} \int_0^t |X_s^i - m_s^h|^2 ds \leq \int_0^t (\mathbb{E}|X_s^i - m_s^h|^p)^{\frac{2}{p}} ds \leq \frac{1}{2(\lambda - \lambda_{p,\alpha,\sigma})} \left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p \right)^{\frac{2}{p}}.$$

The finiteness of this bound ensures that M_t belongs to $L^2(\Omega)$ for each $t \geq 0$, and hence M_t is a square-integrable continuous martingale.

We next control higher moments of M_t . By the Burkholder–Davis–Gundy inequality (see, e.g. [39, Theorem 19.20]), there exists a constant $C_{\text{BDG},p} > 0$ depending only on p such that

$$\mathbb{E} \left[\left| \int_0^t (X_s^i - m_s^h) dW_s^i \right|^p \right] \leq C_{\text{BDG},p} \mathbb{E} \left(\int_0^t |X_s^i - m_s^h|^2 ds \right)^{\frac{p}{2}}. \quad (2.3)$$

Taking expectations in the above inequality for $|X_t^i|^p$ and using (2.3), we get

$$\mathbb{E}|X_t^i|^p \leq 3^{p-1} \mathbb{E}|X_0^i|^p + 3^{p-1} \lambda^p I_t^i + 3^{p-1} \sigma^p C_{\text{BDG},p} II_t^i,$$

where

$$I_t^i := \mathbb{E} \left(\int_0^t |X_s^i - m_s^h| ds \right)^p, \quad II_t^i := \mathbb{E} \left(\int_0^t |X_s^i - m_s^h|^2 ds \right)^{\frac{p}{2}}.$$

We first estimate I_t^i . By Minkowski's integral inequality and Lemma 2.2, we have

$$\begin{aligned} I_t^i &\leq \left(\int_0^t (\mathbb{E}|X_s^i - m_s^h|^p)^{\frac{1}{p}} ds \right)^p \\ &\leq \left(\int_0^t \left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p \right)^{\frac{1}{p}} e^{-(\lambda - \lambda_{p, \alpha, \sigma})s} ds \right)^p \\ &\leq \frac{\Lambda_\alpha}{(\lambda - \lambda_{p, \alpha, \sigma})^p} \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p. \end{aligned}$$

A similar argument yields for II_t^i ,

$$\begin{aligned} II_t^i &\leq \left(\int_0^t \left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p \right)^{\frac{2}{p}} e^{-2(\lambda - \lambda_{p, \alpha, \sigma})s} ds \right)^{\frac{p}{2}} \\ &\leq \frac{\Lambda_\alpha}{(2(\lambda - \lambda_{p, \alpha, \sigma}))^{\frac{p}{2}}} \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p. \end{aligned} \tag{2.4}$$

Collecting the estimates for I_t^i and II_t^i above, we obtain the desired uniform bound

$$\max_{1 \leq i \leq N} \sup_{t \geq 0} \mathbb{E}|X_t^i|^p \leq C_*,$$

where $C_* > 0$ is constant depending only on $\lambda, \sigma, p, \Lambda_\alpha$, and the initial data. $\square$

In the previous lemma, we established that each process X_t^i remains uniformly bounded in $L^p(\Omega)$ under suitable assumptions on the initial data and the parameter λ . We now turn to the large-time behavior of the particle positions in $L^p(\Omega)$. The next result shows that all particle trajectories converge in $L^p(\Omega)$ to a common random limit.

Lemma 2.4. *Let $p \geq 2$, $\lambda > \lambda_{p, \alpha, \sigma}$, and assume $\{X_0^i\}_{i=1}^N \subset L^p(\Omega)$. Then there exists a random variable $X_\infty \in L^p(\Omega)$ such that*

$$\lim_{t \rightarrow \infty} X_t^i = X_\infty \quad \text{in } L^p(\Omega), \quad i = 1, 2, \dots, N.$$

Proof. We start from the integral representation of X_t^i :

$$X_t^i = X_0^i - \lambda \int_0^t (X_s^i - m_s^h) ds + \sigma \int_0^t (X_s^i - m_s^h) dW_s^i =: X_0^i - \lambda I_t^i + \sigma II_t^i.$$

We first analyze the martingale term II_t^i . From estimates (2.3) and (2.4), we recall that

$$\mathbb{E}[|II_t^i|^p] \leq \frac{C_{\text{BDG}, p} \Lambda_\alpha}{(2(\lambda - \lambda_{p, \alpha, \sigma}))^{\frac{p}{2}}} \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p,$$

which implies in particular that

$$\sup_{t \geq 0} \mathbb{E}[|II_t^i|^p] < \infty.$$

Hence the sequence $\{II_t^i\}_{t \geq 0}$ is uniformly bounded in $L^p(\Omega)$. Since II_t^i is the stochastic integral of a progressively measurable and square-integrable process, it defines an L^p -martingale. By the martingale convergence theorem, there exists a random variable $II_\infty^i \in L^p(\Omega)$ such that

$$II_t^i \xrightarrow{t \rightarrow \infty} II_\infty^i \quad \text{in } L^p(\Omega) \text{ and almost surely.}$$

We next examine the drift term I_t^i . Using Minkowski's inequality, for any $t \geq 0$ we estimate

$$\begin{aligned}
\mathbb{E} \left| \int_0^\infty (X_s^i - m_s^h) ds - I_t^i \right|^p &= \mathbb{E} \left| \int_t^\infty (X_s^i - m_s^h) ds \right|^p \\
&\leq \left(\int_t^\infty (\mathbb{E} |X_s^i - m_s^h|^p)^{\frac{1}{p}} ds \right)^p \\
&\leq \left(\int_t^\infty \left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E} |X_0^i - X_0^j|^p \right)^{\frac{1}{p}} e^{-(\lambda - \lambda_{p, \alpha, \sigma})s} ds \right)^p \\
&\leq \frac{\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E} |X_0^i - X_0^j|^p}{(\lambda - \lambda_{p, \alpha, \sigma})^p} e^{-p(\lambda - \lambda_{p, \alpha, \sigma})t} \\
&\xrightarrow[t \rightarrow \infty]{} 0.
\end{aligned}$$

Hence, we have

$$\int_0^t (X_s^i - m_s^h) ds \xrightarrow[t \rightarrow \infty]{} \int_0^\infty (X_s^i - m_s^h) ds \quad \text{in } L^p(\Omega).$$

Combining both limits, we obtain the existence of

$$X_\infty^i := X_0^i - \lambda \int_0^\infty (X_s^i - m_s^h) ds + \sigma \int_0^\infty (X_s^i - m_s^h) dW_s^i \in L^p(\Omega),$$

such that

$$\lim_{t \rightarrow \infty} X_t^i = X_\infty^i \quad \text{in } L^p(\Omega). \quad (2.5)$$

Finally, to identify the common limit, we note that Lemma 2.2 ensures

$$\mathbb{E} |X_t^i - X_t^j|^p \rightarrow 0 \quad \text{as } t \rightarrow \infty.$$

This and (2.5) imply $X_\infty^i = X_\infty^j$ for all $1 \leq i, j \leq N$, and the proof is complete. $\square$

2.3. Almost sure convergence. In this part, we strengthen the result of the previous subsection by establishing that the convergence of the particle system also holds almost surely, that is, pointwise in $\omega \in \Omega$. The motivation for this refinement is that the results of Section 2.2 ensure convergence only in the sense of expectations, which confirms stability of the system on average but does not guarantee that a single realization of the stochastic dynamics necessarily leads to the coalescence of particles.

To bridge this gap, we make use of the classical Kolmogorov continuity theorem (see, e.g., [39, Theorem 10.1]), which asserts that if a stochastic process $\{X_t\}_{t \geq 0}$ satisfies an estimate of the form

$$\mathbb{E} |X_t - X_s|^\alpha \leq c |t - s|^{1+\beta} \quad \text{for all } s, t \geq 0$$

for some $\alpha, \beta > 0$, then it admits a modification whose trajectories are γ -Hölder continuous for every $\gamma \in (0, \frac{\beta}{\alpha})$. We will verify below that this criterion is fulfilled by the particle trajectories of our system.

Lemma 2.5. *Let $p > 2$ and assume $\lambda > \lambda_{p, \alpha, \sigma}$. If $\{X_0^i\}_{i=1}^N \subset L^p(\Omega)$, then each process X_t^i admits a modification such that its sample paths $t \mapsto X_t^i$ are γ -Hölder continuous on $[0, \infty)$ for every $\gamma \in (0, \frac{p-2}{2p})$.*

Proof. For $t, s \geq 0$, we write the increment as

$$X_t^i - X_s^i = -\lambda \int_s^t (X_r^i - m_r^h) dr + \sigma \int_s^t (X_r^i - m_r^h) dW_r^i.$$

We estimate the p -th moment of this increment. Applying almost the same arguments as in the proofs of Lemmas 2.2 and 2.3, together with the Burkholder–Davis–Gundy inequality and Minkowski’s integral inequality, yields

$$\begin{aligned} \mathbb{E}|X_t^i - X_s^i|^p &\leq 2^{p-1}\lambda^p \mathbb{E} \left(\int_s^t |X_r^i - m_r^h| dr \right)^p + 2^{p-1}\sigma^p C_{\text{BDG},p} \mathbb{E} \left(\int_s^t |X_r^i - m_r^h|^2 dr \right)^{\frac{p}{2}} \\ &\leq 2^{p-1}\lambda^p \left(\int_s^t (\mathbb{E}|X_r^i - m_r^h|^p)^{\frac{1}{p}} dr \right)^p + 2^{p-1}\sigma^p C_{\text{BDG},p} \left(\int_s^t (\mathbb{E}|X_r^i - m_r^h|^p)^{\frac{2}{p}} dr \right)^{\frac{p}{2}} \\ &\leq 2^{p-1}\lambda^p \left(\int_s^t \left(\Lambda_\alpha \max_{1 \leq i,j \leq N} \mathbb{E}|X_0^i - X_0^j|^p \right)^{\frac{1}{p}} e^{-(\lambda - \lambda_{p,\alpha,\sigma})r} dr \right)^p \\ &\quad + 2^{p-1}\sigma^p C_{\text{BDG},p} \left(\int_s^t \left(\Lambda_\alpha \max_{1 \leq i,j \leq N} \mathbb{E}|X_0^i - X_0^j|^p \right)^{\frac{2}{p}} e^{-2(\lambda - \lambda_{p,\alpha,\sigma})r} dr \right)^{\frac{p}{2}}. \end{aligned}$$

Since $\lambda > \lambda_{p,\alpha,\sigma}$, the exponential factors in the integrals decay uniformly in time. Hence, both integrals on the right-hand side can be bounded proportionally to $|t - s|$, that is,

$$\int_s^t e^{-(\lambda - \lambda_{p,\alpha,\sigma})r} dr \leq C|t - s|$$

for some constant $C > 0$ independent of s, t . Substituting this estimate into the previous inequality yields

$$\mathbb{E}|X_t^i - X_s^i|^p \leq C \left(|t - s|^p + |t - s|^{\frac{p}{2}} \right),$$

where $C > 0$ is independent of t .

This verifies the assumption of the Kolmogorov continuity theorem with $\alpha = p$ and $\beta = \frac{p}{2} - 1$, so each X_t^i admits a modification whose trajectories are γ -Hölder continuous for every $\gamma \in (0, \frac{p-2}{2p})$. $\square$

By subadditivity of $\mathbb{P}$, without loss of generality, we may assume that all trajectories X_t^i possess γ -Hölder continuous sample paths. From this point on, we shall always work with such a modification.

Lemma 2.6. *Under the assumptions of Lemma 2.2, if $p > 2$, then we have*

$$X_t^i - X_t^j \xrightarrow[t \rightarrow \infty]{} 0 \quad \text{almost surely.}$$

Proof. Fix $\varepsilon > 0$. By Lemma 2.5 and the assumption of γ -Hölder continuity, the sample paths of $X_t^i - X_t^j$ are uniformly continuous. Hence, for each $\omega \in \Omega$, there exists $\delta = \delta(\omega) > 0$, independent of i and t , such that

$$|t - s| < \delta \implies |X_t^i - X_s^i| < \varepsilon, \quad \forall s, t \in [0, \infty).$$

Choose $n \in \mathbb{N}$ large enough that $\frac{1}{2^n} \leq \delta$, and define a uniform time partition $t_k := \frac{k}{2^n}$ for $k \in \mathbb{N}$. Let

$$A_k := \{\omega \in \Omega : |X_{t_k}^i - X_{t_k}^j| \geq \varepsilon\}.$$

Applying Markov’s inequality together with Lemma 2.2, we find

$$\mathbb{P}(A_k) \leq \frac{\mathbb{E}|X_{t_k}^i - X_{t_k}^j|^p}{\varepsilon^p} \leq \frac{\max_{1 \leq i,j \leq N} \mathbb{E}|X_0^i - X_0^j|^p}{\varepsilon^p} e^{-p(\lambda - \lambda_{p,\alpha,\sigma})t_k}.$$

The right-hand side forms a geometric series in k , and hence

$$\sum_{k=0}^{\infty} \mathbb{P}(A_k) \leq \sum_{k=0}^{\infty} \frac{\max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p}{\varepsilon^p} e^{-p(\lambda - \lambda_{p, \alpha, \sigma})t_k} < \infty.$$

By the Borel–Cantelli lemma, we thus have

$$\mathbb{P}(A_k \text{ infinitely often}) = 0,$$

which implies

$$X_{t_k}^i - X_{t_k}^j \xrightarrow[k \rightarrow \infty]{} 0 \quad \text{almost surely.} \quad (2.6)$$

To extend this convergence to all $t \geq 0$, fix ω outside a null set for which (2.6) holds. Then there exists $k_* = k_*(\omega)$ such that $|X_{t_k}^i - X_{t_k}^j| \leq \varepsilon$ for all $k \geq k_*$. For any $t \geq t_{k_*}$, there is a unique $t_k = \frac{k}{2^n}$ with $k \geq k_*$ such that $t_k \leq t < t_{k+1}$. Our assumption on n implies $|t_{k+1} - t_k| \leq \delta$, and therefore

$$|X_t^i - X_t^j| \leq |X_t^i - X_{t_k}^i| + |X_{t_k}^i - X_{t_k}^j| + |X_{t_k}^j - X_t^j| \leq 2\varepsilon, \quad \forall t \geq t_{k_*},$$

from which the result follows. $\square$

We can now establish the almost sure convergence in Theorem 1.1 (i) when $p > 2$.

Proof of Theorem 1.1 (i). We adopt the same notations as in the proof of Lemma 2.4. Since the (L^p) martingale convergence theorem ensures the almost sure convergence of II_t^i , it remains to verify that the deterministic integral term $I_t^i = \int_0^t (X_s^i - m_s^h) ds$ also converges almost surely as $t \rightarrow \infty$.

From Lemma 2.4, we already know that

$$I_t^i = \int_0^t (X_s^i - m_s^h) ds \longrightarrow I_\infty^i = \int_0^\infty (X_s^i - m_s^h) ds \quad \text{in } L^p(\Omega).$$

To strengthen this convergence to almost sure, we first observe that

$$\int_0^\infty \mathbb{E}|X_s^i - m_s^h| ds \leq \int_0^\infty (\mathbb{E}|X_s^i - m_s^h|^p)^{\frac{1}{p}} ds \leq \frac{\left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p\right)^{\frac{1}{p}}}{\lambda - \lambda_{p, \alpha, \sigma}} < +\infty.$$

Thus $X^i - m^h \in L^1((0, \infty) \times \Omega)$, and by Fubini's theorem we deduce that, almost surely, $X^i - m^h \in L^1(0, \infty)$. Consequently, the sample paths

$$(0, \infty) \ni t \mapsto I_t^i = \int_0^t (X_s^i - m_s^h) ds$$

are absolutely continuous almost surely, and in particular, uniformly continuous on $[0, \infty)$.

Now, we fix $\varepsilon > 0$ and $n \in \mathbb{N}$, and define the time partition $t_k := k/2^n$ and events

$$A_k := \{\omega \in \Omega : |I_{t_k}^i - I_\infty^i| \geq \varepsilon\}.$$

Applying Markov's inequality, Minkowski's integral inequality, and Lemma 2.2, we obtain

$$\mathbb{P}(A_k) \leq \frac{\mathbb{E}|I_{t_k}^i - I_\infty^i|^p}{\varepsilon^p} \leq \frac{1}{\varepsilon^p} \left(\int_{t_k}^\infty (\mathbb{E}|X_s^i - m_s^h|^p)^{\frac{1}{p}} ds \right)^p \leq \frac{\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p}{\varepsilon^p (\lambda - \lambda_{p, \alpha, \sigma})^p} e^{-p(\lambda - \lambda_{p, \alpha, \sigma})t_k}.$$

The right-hand side forms a geometric series in k , hence

$$\sum_{k=0}^{\infty} \mathbb{P}(A_k) < \infty.$$

Thus, by the Borel–Cantelli Lemma, we have

$$I_{t_k}^i \rightarrow I_{\infty}^i \quad \text{almost surely as } k \rightarrow \infty.$$

As in the proof of Lemma 2.6, the uniform continuity of the sample paths $t \mapsto I_t^i$ ensures that this convergence extends from the discrete sequence $\{t_k\}$ to all $t \geq 0$. Hence $I_t^i \rightarrow I_{\infty}^i$ almost surely, and therefore

$$X_t^i = X_0^i - \lambda I_t^i + \sigma I I_t^i \xrightarrow[t \rightarrow \infty]{\text{a.s.}} X_{\infty}^i := X_0^i - \lambda I_{\infty}^i + \sigma I I_{\infty}^i.$$

Finally, Lemma 2.6 implies that $X_t^i - X_t^j \rightarrow 0$ almost surely, and hence the limits X_{∞}^i must coincide for all i . We thus conclude that every particle converges almost surely to a common random vector X_{∞} , as claimed. $\square$

The preceding analysis has focused on a model where a d -dimensional state is driven by a one-dimensional Wiener process. It is worth noting that the core result of Lemma 2.2, namely the exponential contraction condition, can be adapted for different noise structures. The following remark summarizes how the consensus threshold $\lambda_{p,\alpha,\sigma}$ is adjusted for several common variations.

Remark 2.1. *The baseline consensus threshold established in Lemma 2.2,*

$$\lambda_{p,\alpha,\sigma} = (p-1)\Lambda_{\alpha}^{2/p}\sigma^2 \quad \text{with } \Lambda_{\alpha} = \frac{e^{-\alpha f} + h(\alpha)}{h(\alpha)}$$

corresponds to the case where each d -dimensional state vector is driven by an independent one-dimensional Wiener process. Depending on the dimensionality and structure of the noise, this threshold changes as follows. In all cases below, the same assumptions as in Lemma 2.2 still apply.

- **Model with one-dimensional common-direction noise.** Consider

$$dX_t^i = -\lambda(X_t^i - m_t^h)dt + \sigma|X_t^i - m_t^h|\mathbf{1}_d^{\top} dW_t^i,$$

where $\{W_t^i\}_{i=1}^N$ are one-dimensional independent Wiener processes. In this case, the effective diffusion acts identically along all d components, leading to a threshold increased by a factor of d :

$$\mathbb{E}|X_t^i - X_t^j|^p \leq \max_{1 \leq i,j \leq N} \mathbb{E}|X_0^i - X_0^j|^p e^{-p(\lambda - d\lambda_{p,\alpha,\sigma})t}.$$

- **Model with d -dimensional anisotropic noise.** Consider

$$dX_t^i = -\lambda(X_t^i - m_t^h)dt + \sigma(X_t^i - m_t^h) \odot dW_t^i,$$

where $\{W_t^i\}_{i=1}^N$ are d -dimensional independent Wiener processes and $\odot$ denotes the Hadamard (component) product. Since the stochastic perturbations are applied componentwise, the previous argument carries over directly, and the threshold remains unchanged:

$$\mathbb{E}|X_t^i - X_t^j|^p \leq \max_{1 \leq i,j \leq N} \mathbb{E}|X_0^i - X_0^j|^p e^{-p(\lambda - \lambda_{p,\alpha,\sigma})t}.$$

- **Model with d -dimensional isotropic noise.** Consider

$$dX_t^i = -\lambda(X_t^i - m_t^h)dt + \sigma|X_t^i - m_t^h|dW_t^i,$$

where $\{W_t^i\}_{i=1}^N$ are d -dimensional independent Wiener processes. Here, the magnitude of the stochastic forcing scales with the state discrepancy $|X_t^i - m_t^h|$, leading to a modified threshold

$$\mathbb{E}|X_t^i - X_t^j|^p \leq \max_{1 \leq i, j \leq N} \mathbb{E}|X_0^i - X_0^j|^p e^{-p(\lambda - \lambda_{p,d,\alpha,\sigma})t},$$

where $\lambda_{p,d,\alpha,\sigma} := \lambda_{p,\alpha,\sigma} + (d-1)\Lambda_\alpha^{2/p}\sigma^2$.

2.4. Consensus to the minimizer. In this subsection, we combine the large-time behavior established in Theorem 1.1 (i) with the Laplace principle to show that the particle trajectories of (1.1) indeed converge to the global minimizer of the objective function. This part provides the proof of Theorem 1.1 (ii). We start with an auxiliary estimate that will play a key role in linking the consensus dynamics to the decay of the objective functional.

Lemma 2.7. *Let $\{X_t^i\}_{1 \leq i \leq N}$ be the global strong solution to the system (1.1). Under the Assumptions 1.1 and 1.2, we assume that $X_{\text{in}} \in L^p(\Omega)$ and that $\lambda > \lambda_{p,\alpha,\sigma}$ and $\alpha \geq c_1$. Then we have*

$$\frac{d}{dt} \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E} \omega_f^\alpha(X_t^i) \right) \geq -\frac{\lambda L_{\omega_f^\alpha}}{N} \sum_{i=1}^N (\mathbb{E}|X_t^i - m_t^h|^2)^{\frac{1}{2}} - \frac{\alpha \sigma^2 c_0 e^{-\alpha f}}{2N} \sum_{i=1}^N \mathbb{E}|X_t^i - m_t^h|^2.$$

Proof. Applying the Itô chain rule to $\frac{1}{N} \sum_{i=1}^N \omega_f^\alpha(X_t^i)$ yields

$$d \left(\frac{1}{N} \sum_{i=1}^N \omega_f^\alpha(X_t^i) \right) = -\frac{\alpha}{N} \sum_{i=1}^N \omega_f^\alpha(X_t^i) \nabla f(X_t^i) \cdot dX_t^i + \frac{1}{2N} \sum_{i=1}^N \text{Tr}(dX_t^i [dX_t^i]^\top \nabla^2 \omega_f^\alpha(X_t^i)).$$

Substituting the dynamics (1.1) into the above, we obtain

$$\begin{aligned} d \left(\frac{1}{N} \sum_{i=1}^N \omega_f^\alpha(X_t^i) \right) &= \frac{\alpha \lambda}{N} \sum_{i=1}^N \omega_f^\alpha(X_t^i) \nabla f(X_t^i) \cdot (X_t^i - m_t^h) dt - \frac{\alpha \sigma}{N} \sum_{i=1}^N \omega_f^\alpha(X_t^i) \nabla f(X_t^i) \cdot (X_t^i - m_t^h) dW_t^i \\ &\quad + \frac{\sigma^2 \alpha}{2N} \sum_{i=1}^N \omega_f^\alpha(X_t^i) (X_t^i - m_t^h)^\top (\alpha \nabla f(X_t^i) \nabla f(X_t^i)^\top - \nabla^2 f(X_t^i)) (X_t^i - m_t^h) dt. \end{aligned}$$

Taking expectations eliminates the stochastic integral term and gives

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E} \omega_f^\alpha(X_t^i) \right) &= \frac{\alpha \lambda}{N} \sum_{i=1}^N \mathbb{E} \left[\omega_f^\alpha(X_t^i) \nabla f(X_t^i) \cdot (X_t^i - m_t^h) \right] \\ &\quad + \frac{\alpha \sigma^2}{2N} \sum_{i=1}^N \mathbb{E} \left[\omega_f^\alpha(X_t^i) (X_t^i - m_t^h)^\top (\alpha \nabla f(X_t^i) \nabla f(X_t^i)^\top - \nabla^2 f(X_t^i)) (X_t^i - m_t^h) \right]. \end{aligned}$$

For the first term on the right-hand side of the above, we utilize that $-\alpha \omega_f^\alpha \nabla f = \nabla w_f^\alpha$, and then Assumption 1.1 to estimate

$$\mathbb{E} \left[\alpha \omega_f^\alpha(X_t^i) \nabla f(X_t^i) \cdot [X_t^i - m_t^h] \right] \geq -L_{\omega_f^\alpha} \mathbb{E}|X_t^i - m_t^h|.$$

For the second term, by Assumption 1.2, we have

$$\begin{aligned} & \mathbb{E} [\omega_f^\alpha(X_t^i)(X_t^i - m_t^h)^\top \nabla^2 f(X_t^i)(X_t^i - m_t^h)] \\ & \leq c_0 e^{-\alpha f} \mathbb{E} |X_t^i - m_t^h|^2 + c_1 \mathbb{E} [\omega_f^\alpha(X_t^i)(\nabla f(X_t^i) \cdot (X_t^i - m_t^h))^2]. \end{aligned}$$

This deduces

$$\begin{aligned} & \mathbb{E} [\omega_f^\alpha(X_t^i)(X_t^i - m_t^h)^\top (\alpha \nabla f(X_t^i) \nabla f(X_t^i)^\top - \nabla^2 f(X_t^i))(X_t^i - m_t^h)] \\ & \geq -c_0 e^{-\alpha f} \mathbb{E} |X_t^i - m_t^h|^2 + (\alpha - c_1) \mathbb{E} [\omega_f^\alpha(X_t^i)(\nabla f(X_t^i) \cdot (X_t^i - m_t^h))^2] \\ & \geq -c_0 e^{-\alpha f} \mathbb{E} |X_t^i - m_t^h|^2, \end{aligned}$$

where the last inequality follows from $\alpha \geq c_1$.

Combining the two bounds above yields the desired differential inequality. $\square$

We are now in a position to provide the proof of Theorem 1.1 (ii).

Proof of Theorem 1.1 (ii). By Lemma 2.7, we find

$$\begin{aligned} \frac{d}{dt} \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E} \omega_f^\alpha(X_t^i) \right) & \geq -\frac{\lambda L_{\omega_f^\alpha}}{N} \sum_{i=1}^N (\mathbb{E} |X_t^i - m_t|^2)^{\frac{1}{2}} - \frac{\alpha \sigma^2 c_0 e^{-\alpha f}}{2N} \sum_{i=1}^N \mathbb{E} |X_t^i - m_t^h|^2 \\ & \geq -\lambda L_{\omega_f^\alpha} \left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E} |X_0^i - X_0^j|^2 \right)^{\frac{1}{2}} e^{-(\lambda - \lambda_{2, \alpha, \sigma})t} \\ & \quad - \frac{\alpha \sigma^2 c_0 e^{-\alpha f}}{2} \left(\Lambda_\alpha \max_{1 \leq i, j \leq N} \mathbb{E} |X_0^i - X_0^j|^2 \right) e^{-2(\lambda - \lambda_{2, \alpha, \sigma})t}, \end{aligned}$$

where the second inequality follows by Lemma 2.2 (applied with $p = 2$). Now integrating in t gives

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} \omega_f^\alpha(X_t^i) & \geq \mathbb{E} \omega_f^\alpha(X_{\text{in}}) - \lambda L_{\omega_f^\alpha} (2\Lambda_\alpha \text{Var}(X_{\text{in}}))^{\frac{1}{2}} \frac{1 - e^{-(\lambda - \lambda_{2, \alpha, \sigma})t}}{\lambda - \lambda_{2, \alpha, \sigma}} \\ & \quad - \frac{\alpha \sigma^2 c_0 e^{-\alpha f} \Lambda_\alpha (1 - e^{-2(\lambda - \lambda_{2, \alpha, \sigma})t})}{2(\lambda - \lambda_{2, \alpha, \sigma})} \text{Var}(X_{\text{in}}), \end{aligned}$$

where we used the identity

$$\mathbb{E} |X_0^i - X_0^j|^2 = 2(\mathbb{E} |X_{\text{in}}|^2 - |\mathbb{E} X_{\text{in}}|^2) = 2\text{Var}(X_{\text{in}}) \quad \forall i, j \in \{1, \dots, N\}.$$

Using that $X_t^i \rightarrow X_\infty$ in $L^1(\Omega)$, and that ω_f^α is Lipschitz, we can pass to the limit $t \rightarrow \infty$ to obtain

$$\begin{aligned} \mathbb{E} \omega_f^\alpha(X_\infty) & \geq \mathbb{E} \omega_f^\alpha(X_{\text{in}}) - \frac{\lambda L_{\omega_f^\alpha}}{\lambda - \lambda_{2, \alpha, \sigma}} (2\Lambda_\alpha \text{Var}(X_{\text{in}}))^{\frac{1}{2}} - \frac{\alpha \sigma^2 c_0 \Lambda_\alpha e^{-\alpha f}}{2(\lambda - \lambda_{2, \alpha, \sigma})} \text{Var}(X_{\text{in}}) \\ & \geq \epsilon \mathbb{E} \omega_f^\alpha(X_{\text{in}}). \end{aligned}$$

Therefore, taking logarithms, we have

$$-\frac{1}{\alpha} \log \mathbb{E} \omega_f^\alpha(X_\infty) \leq -\frac{1}{\alpha} \log \mathbb{E} \omega_f^\alpha(X_{\text{in}}) - \frac{1}{\alpha} \log \epsilon.$$

By Laplace's principle (see, e.g. [14, 37]), we conclude that

$$\text{ess inf}_{\omega \in \Omega} f(X_\infty(\omega)) \leq \text{ess inf}_{\omega \in \Omega} f(X_{\text{in}}(\omega)) + o(1) \quad (\alpha \rightarrow \infty),$$

which shows that the consensus point X_∞ asymptotically approaches to the global minimizer of f . $\square$

3. UNIFORM-IN-TIME PROPAGATION OF CHAOS

In this section, we establish our second main result, Theorem 1.2, which provides a quantitative mean-field limit for the modified CBO dynamics. More precisely, we show that the empirical distribution associated with the particle system (1.1) converges uniformly in time toward the law of its McKean–Vlasov counterpart, with a rate depending only on the initial data and model parameters. This result rigorously confirms that the collective dynamics of a large ensemble of particles can be faithfully described by the mean-field equation (1.2), even in the long-time regime when the system approaches consensus.

To prepare for the analysis, we first introduce basic notation that will be used throughout this section. For a probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ we denote

$$\mathcal{M}(\mu) := \int_{\mathbb{R}^d} x \mu(dx), \quad V_p(\mu) := \int_{\mathbb{R}^d} |x - \mathcal{M}(\mu)|^p \mu(dx), \quad \mathfrak{m}_p(\mu) := \int_{\mathbb{R}^d} |x|^p \mu(dx),$$

corresponding respectively to the mean, centered p -th moment, and absolute p -th moment of μ .

For the particle system (1.1), we introduce the empirical measure

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

so that the particle dynamics can be written concisely as

$$dX_t^i = -\lambda(X_t^i - m_f^h[\mu_t^N])dt + \sigma(X_t^i - m_f^h[\mu_t^N])dW_t^i. \quad (3.1)$$

3.1. Preliminary estimates and stability observations. We begin with several elementary but essential estimates that will serve as preliminary ingredients for the proof of Theorem 1.2. They provide quantitative control on the fluctuations of various averaged quantities, in particular on the sensitivity of the weighted mean $m_f^h[\mu]$ with respect to perturbations of the underlying measure μ .

The key difficulty in deriving a stability estimate for (3.1) lies in controlling the nonlinear dependence of the drift and diffusion terms on the empirical law through $m_f^h[\mu_t^N]$. While the usual mean operator $\mathcal{M}$ satisfies the optimal Lipschitz bound

$$|\mathcal{M}(\mu) - \mathcal{M}(\nu)| \leq \mathcal{W}_1(\mu, \nu), \quad \forall \mu, \nu \in \mathcal{P}_1(\mathbb{R}^d),$$

the same is not immediately true for the weighted mean $m_f^h[\mu]$, whose definition involves the nonlinearity of the (regularized) Gibbs weight. To overcome this difficulty, one can interpret $m_f^h[\mu]$ as a perturbation of the simple mean, i.e.,

$$m_f^h[\mu] - m_f^h[\nu] = (m_f^h[\mu] - \mathcal{M}(\mu)) + (\mathcal{M}(\mu) - \mathcal{M}(\nu)) + (\mathcal{M}(\nu) - m_f^h[\nu]).$$

This decomposition, exploited in [22] for the classical CBO system, allows one to treat the nonlinear weighted mean as a small deviation from $\mathcal{M}$ that can be quantified in terms of the Wasserstein distance and the centered moments of μ and ν .

In the sequel, we adapt this perturbative strategy to our modified model, which features the regularized Gibbs weight $\psi_h(x) = \omega_f^\alpha(x) + h(\alpha)$. Thanks to this regularization and the weaker assumption

that ω_f^α is Lipschitz continuous (see Assumption 1.1), the resulting estimates hold under milder conditions than those imposed in [22]. In particular, our analysis does not require f to be bounded from above, and only depends on the Lipschitz constant of the exponential weight ω_f^α . Moreover, while the approach of [22] was developed within an L^2 framework, we carry out the analysis in a more general L^p setting ($p \geq 2$), which allows for sharper quantitative estimates and a broader range of integrability assumptions on the initial data. We now make this precise in the following lemma, which captures the quantitative sensitivity of the weighted mean operator.

Lemma 3.1. *Let $p \geq 2$ and assume Assumption 1.1. Then for all $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$, we have*

$$|m_f^h[\mu] - \mathcal{M}(\mu) - m_f^h[\nu] + \mathcal{M}(\nu)| \leq \frac{L_{\omega_f^\alpha} \Lambda_\alpha}{h(\alpha)} \left(V_p^{\frac{1}{p}}(\mu) + V_p^{\frac{1}{p}}(\nu) \right) \mathcal{W}_p(\mu, \nu).$$

Proof. Although the structure of the proof is similar to that of [22], we provide it here for completeness and to highlight that, unlike [22, Assumption 1], our argument does not require any pointwise upper bound on the objective function f .

Set

$$Z_\mu := \int_{\mathbb{R}^d} \psi_h(x) \mu(dx), \quad g(x) := (x - \mathcal{M}(\nu)) (\psi_h(x) - Z_\mu),$$

and let $\pi \in \Pi(\mu, \nu)$ be any coupling. Then, by a simple rearrangement, we get

$$\begin{aligned} m_f^h[\mu] - \mathcal{M}(\mu) - m_f^h[\nu] + \mathcal{M}(\nu) &= \int_{\mathbb{R}^d} (x - \mathcal{M}(\nu)) (\psi_h(x) - Z_\mu) \left(\frac{\mu}{Z_\mu} - \frac{\nu}{Z_\nu} \right) \\ &= \frac{1}{Z_\mu} \iint_{\mathbb{R}^d \times \mathbb{R}^d} (g(x) - g(y)) \pi(dxdy) + \left(\frac{1}{Z_\mu} - \frac{1}{Z_\nu} \right) \int_{\mathbb{R}^d} g(x) \nu(dx) \\ &=: I + II. \end{aligned}$$

We first estimate I . Since $\psi_h(x) = \omega_f^\alpha(x) + h(\alpha)$ is $L_{\omega_f^\alpha}$ -Lipschitz by Assumption 1.1, the difference $g(x) - g(y)$ satisfies

$$|g(x) - g(y)| \leq |x - y| |\psi_h(x) - Z_\mu| + L_{\omega_f^\alpha} |y - \mathcal{M}(\nu)| |x - y|.$$

Applying Hölder's inequality and using that the marginals of π are μ and ν , we find

$$\begin{aligned} &\iint_{\mathbb{R}^d \times \mathbb{R}^d} |g(x) - g(y)| \pi(dxdy) \\ &\leq \left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dxdy) \right)^{\frac{1}{2}} \left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |\psi_h(x) - Z_\mu|^2 \pi(dxdy) \right)^{\frac{1}{2}} \\ &\quad + L_{\omega_f^\alpha} \left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dxdy) \right)^{\frac{1}{2}} \left(\iint_{\mathbb{R}^d \times \mathbb{R}^d} |y - \mathcal{M}(\nu)|^2 \pi(dxdy) \right)^{\frac{1}{2}}. \end{aligned}$$

Furthermore, since ψ_h is $L_{\omega_f^\alpha}$ -Lipschitz, the first integral satisfies

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} |\psi_h(x) - Z_\mu|^2 \pi(dxdy) = \int_{\mathbb{R}^d} |\psi_h(x) - Z_\mu|^2 \mu(dx) \leq L_{\omega_f^\alpha}^2 V_2(\mu).$$

Hence, we have

$$|I| \leq \frac{L_{\omega_f^\alpha}}{Z_\mu} (\sqrt{V_2(\mu)} + \sqrt{V_2(\nu)}) \mathcal{W}_2(\mu, \nu) \leq \frac{L_{\omega_f^\alpha}}{h(\alpha)} (\sqrt{V_2(\mu)} + \sqrt{V_2(\nu)}) \mathcal{W}_2(\mu, \nu),$$

due to $Z_\mu \geq h(\alpha)$.

We now turn to II . A straightforward computation yields

$$\begin{aligned}
|II| &\leq \frac{1}{h(\alpha)^2} |Z_\mu - Z_\nu| \int_{\mathbb{R}^d} |x - \mathcal{M}(\nu)| e^{-\alpha f(x)} \nu(dx) \\
&\leq \frac{e^{-\alpha \underline{f}}}{h(\alpha)^2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |e^{-\alpha f(x)} - e^{-\alpha f(y)}| \pi(dxdy) \int_{\mathbb{R}^d} |x - \mathcal{M}(\nu)| \nu(dx) \\
&\leq \frac{L_{\omega_f^\alpha} e^{-\alpha \underline{f}}}{h(\alpha)^2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - y| \pi(dxdy) \int_{\mathbb{R}^d} |x - \mathcal{M}(\nu)| \nu(dx) \\
&\leq \frac{L_{\omega_f^\alpha} e^{-\alpha \underline{f}}}{h(\alpha)^2} \mathcal{W}_1(\mu, \nu) V_1(\nu).
\end{aligned}$$

Combining the two estimates gives

$$|m_f^h[\mu] - \mathcal{M}(\mu) - m_f^h[\nu] + \mathcal{M}(\nu)| \leq \frac{L_{\omega_f^\alpha}}{h(\alpha)} \left(\sqrt{V_2(\mu)} + \sqrt{V_2(\nu)} \right) \mathcal{W}_2(\mu, \nu) + \frac{L_{\omega_f^\alpha} e^{-\alpha \underline{f}}}{h(\alpha)^2} \mathcal{W}_1(\mu, \nu) V_1(\nu).$$

Finally, we use the relations $\mathcal{W}_1 \leq \mathcal{W}_2 \leq \mathcal{W}_p$, $V_1 \leq V_2^{1/2} \leq V_p^{1/p}$ for $p \geq 2$, to conclude the desired result. $\square$

3.2. Exponential decay and moment estimates for the synchronous coupling. In this subsection, we derive quantitative estimates for the mean-field process $\bar{X}_t$, which solves the limiting SDE (1.2). We note that (1.2) admits a unique strong solution under Assumption 1.1, together with suitable moment bounds on the initial data. This justifies all computations carried out in this section. For readability, however, we postpone the discussion of well-posedness until Section 4.1. Since (1.2) is well-posed, the associated law $\bar{\rho}_t = \text{Law}(\bar{X}_t)$ defines a weak solution to (1.3).

The following lemma establishes that $\bar{X}_t$ converges exponentially fast towards its mean and that its fluctuations around the weighted average $m_f^h[\bar{\rho}_t]$ also decay exponentially in time.

Lemma 3.2. *Let $p \geq 2$ and assume $\bar{\rho}_0 \in \mathcal{P}_p(\mathbb{R}^d)$. Then for $t \geq 0$, we have*

$$\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p \leq C_p \mathfrak{M}_p(\bar{\rho}_0) e^{-p(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t}$$

and

$$\mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p \leq C_{p,\Lambda_\alpha} \mathfrak{M}_p(\bar{\rho}_0) e^{-p(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t},$$

where

$$\bar{\lambda}_{p,\alpha,\sigma} := (p-1)(1 + \Lambda_\alpha^{2/p})\sigma^2 = \lambda_{p,\alpha,\sigma} + (p-1)\sigma^2.$$

Proof. We start by analyzing the evolution of the mean of $\bar{X}_t$. Taking the expectation in (1.2) gives

$$\frac{d}{dt} \mathbb{E}\bar{X}_t = -\lambda(\mathbb{E}\bar{X}_t - m_f^h[\bar{\rho}_t]).$$

Substituting this into the Itô's formula for $|\bar{X}_t - \mathbb{E}\bar{X}_t|^p$ yields

$$\begin{aligned} d|\bar{X}_t - \mathbb{E}\bar{X}_t|^p &= -p\lambda|\bar{X}_t - \mathbb{E}\bar{X}_t|^p dt + \frac{p\sigma^2}{2}|\bar{X}_t - \mathbb{E}\bar{X}_t|^{p-2}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^2 dt \\ &\quad + \frac{p(p-2)\sigma^2}{2}|\bar{X}_t - \mathbb{E}\bar{X}_t|^{p-4}\langle \bar{X}_t - \mathbb{E}\bar{X}_t, \bar{X}_t - m_f^h[\bar{\rho}_t] \rangle^2 dt \\ &\quad + p\sigma|\bar{X}_t - \mathbb{E}\bar{X}_t|^{p-2}\langle \bar{X}_t - \mathbb{E}\bar{X}_t, \bar{X}_t - m_f^h[\bar{\rho}_t] \rangle dW_t. \end{aligned} \quad (3.2)$$

Similarly as in the proof of Lemma 2.2, using the Cauchy-Schwarz and Hölder's inequalities, then taking expectations, we find

$$\frac{d}{dt}\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p \leq -p\lambda\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p + \frac{p(p-1)\sigma^2}{2} \left[(\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p)^{\frac{p-2}{p}} (\mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p)^{\frac{2}{p}} \right]. \quad (3.3)$$

Next, we control the distance between the mean $\mathbb{E}\bar{X}_t$ and the weighted mean $m_f^h[\bar{\rho}_t]$. By direct calculation,

$$\begin{aligned} |\mathbb{E}\bar{X}_t - m_f^h[\bar{\rho}_t]|^p &= \left| \frac{\int_{\mathbb{R}^d} (x - \mathcal{M}(\bar{\rho}_t))\psi_h(x)\bar{\rho}_t(dx)}{\int_{\mathbb{R}^d} \psi_h(x)\bar{\rho}_t(dx)} \right|^p \\ &\leq \left| \frac{(\int_{\mathbb{R}^d} |x - \mathcal{M}(\bar{\rho}_t)|^p \psi_h(x)\bar{\rho}_t(dx))^{1/p} (\int_{\mathbb{R}^d} \psi_h(x)\bar{\rho}_t(dx))^{(p-1)/p}}{\int_{\mathbb{R}^d} \psi_h(x)\bar{\rho}_t(dx)} \right|^p \\ &\leq \Lambda_\alpha \int_{\mathbb{R}^d} |x - \mathcal{M}(\bar{\rho}_t)|^p \bar{\rho}_t(dx) \\ &= \Lambda_\alpha \mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p, \end{aligned} \quad (3.4)$$

where we used $h(\alpha) \leq \psi_h \leq e^{-\alpha f} + h(\alpha)$.

This deduces

$$(\mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p)^{\frac{2}{p}} \leq \left[(\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p)^{\frac{1}{p}} + (\mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p)^{\frac{1}{p}} \right]^2 \leq 2(1 + \Lambda_\alpha^{\frac{2}{p}}) (\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p)^{\frac{2}{p}}.$$

Plugging this estimate into (3.3) yields

$$\frac{d}{dt}\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p \leq -\left(p\lambda - p(p-1)(1 + \Lambda_\alpha^{\frac{2}{p}})\sigma^2\right) \mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p.$$

Then by Grönwall's inequality, we have

$$\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p \leq \mathbb{E}|\bar{X}_0 - \mathbb{E}\bar{X}_0|^p e^{-p(\lambda - (p-1)(1 + \Lambda_\alpha^{\frac{2}{p}})\sigma^2)t} \leq 2^p \mathfrak{m}_p(\bar{\rho}_0) e^{-p(\lambda - (p-1)(1 + \Lambda_\alpha^{\frac{2}{p}})\sigma^2)t}.$$

To estimate $\mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p$, we use the decomposition

$$|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p \leq 2^{p-1}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p + 2^{p-1}|\mathbb{E}\bar{X}_t - m_f^h[\bar{\rho}_t]|^p,$$

together with (3.4). Taking expectations yields

$$\mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p \leq 2^{p-1}(1 + \Lambda_\alpha)\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p,$$

and substituting the previous decay estimate concludes the proof. $\square$

The exponential decay established above immediately implies that the mean-field trajectories remain uniformly concentrated around their average. We now show that this concentration yields a uniform-in-time L^p bound for $\bar{X}_t$.

Lemma 3.3. *Let $p \geq 2$. If $\lambda > \bar{\lambda}_{p,\alpha,\sigma}$ and $\bar{\rho}_0 \in \mathcal{P}_p(\mathbb{R}^d)$, then there exists a constant $\bar{C}_* = \bar{C}_*(\lambda, \sigma, p, \Lambda_\alpha, \bar{\rho}_0) > 0$ such that*

$$\sup_{t \geq 0} \mathbb{E} |\bar{X}_t|^p \leq \bar{C}_*.$$

Proof. The proof follows the same strategy as Lemma 2.3, now exploiting the explicit exponential decay established in Lemma 3.2. From (1.2) and the triangle inequality, we first estimate

$$\begin{aligned} \mathbb{E} |\bar{X}_t|^p &\leq 3^{p-1} \mathbb{E} |\bar{X}_0|^p + 3^{p-1} \mathbb{E} \left[\int_0^t |\bar{X}_s - m_f^h[\bar{\rho}_s]| ds \right]^p + 3^{p-1} C_{\text{BDG},p} \mathbb{E} \left[\int_0^t |\bar{X}_s - m_f^h[\bar{\rho}_s]|^2 ds \right]^{\frac{p}{2}} \\ &=: 3^{p-1} \mathbb{E} |\bar{X}_0|^p + 3^{p-1} I_t + 3^{p-1} C_{\text{BDG},p} II_t. \end{aligned}$$

where the last term comes from the Burkholder–Davis–Gundy inequality.

To control I , we apply Minkowski's integral inequality together with Lemma 3.2, obtaining

$$\begin{aligned} I_t &\leq \left(\int_0^t (\mathbb{E} |\bar{X}_s - m_f^h[\bar{\rho}_s]|^p)^{\frac{1}{p}} ds \right)^p \\ &\leq 2^{2p-1} (1 + \Lambda_\alpha) \mathfrak{M}_p(\bar{\rho}_0) \left(\int_0^t e^{-(\lambda - \bar{\lambda}_{p,\alpha,\sigma})s} ds \right)^p \\ &\leq \frac{2^{2p-1} (1 + \Lambda_\alpha) \mathfrak{M}_p(\bar{\rho}_0)}{(\lambda - \bar{\lambda}_{p,\alpha,\sigma})^p}. \end{aligned}$$

Similarly, for II we have

$$\begin{aligned} II_t &\leq \left(\int_0^t (\mathbb{E} |\bar{X}_s - m_f^h[\bar{\rho}_s]|^p)^{\frac{2}{p}} ds \right)^{\frac{p}{2}} \\ &\leq 2^{2p-1} (1 + \Lambda_\alpha) \mathfrak{M}_p(\bar{\rho}_0) \left(\int_0^t e^{-2(\lambda - \bar{\lambda}_{p,\alpha,\sigma})s} ds \right)^{\frac{p}{2}} \\ &\leq \frac{2^{2p-1} (1 + \Lambda_\alpha) \mathfrak{M}_p(\bar{\rho}_0)}{(2(\lambda - \bar{\lambda}_{p,\alpha,\sigma}))^{\frac{p}{2}}}. \end{aligned}$$

Combining all bounds, we deduce that there exists a constant $\bar{C}_* > 0$, depending only on the parameters $(\lambda, \sigma, p, \Lambda_\alpha, \bar{\rho}_0)$, such that

$$\sup_{t \geq 0} \mathbb{E} |\bar{X}_t|^p \leq \bar{C}_*.$$

This completes the proof. $\square$

3.3. Concentration estimates for the synchronous coupling. We now quantify how tightly the empirical weighted mean concentrates around its population counterpart in the synchronous coupling. This provides the probabilistic ingredient required later to pass from the moment estimates in L^p to the uniform-in-time propagation of chaos.

For clarity, we recall the relevant notation. Let $\{W_t^i\}_{i=1}^N$ be a family of independent one-dimensional Wiener processes and let $\bar{X}_t^i$, $i = 1, \dots, N$, denote the corresponding i.i.d. solutions of the mean-field

SDE (1.2) with common initial law $\bar{\rho}_0$. We denote by $\bar{\rho}_t = \text{Law}(\bar{X}_t^i)$ their common distribution, and by

$$\bar{\mu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{\bar{X}_t^i}$$

the associated empirical measure of the system.

The next lemma quantifies the deviation between $m_f^h[\bar{\mu}_t^N]$ and its population value $m_f^h[\bar{\rho}_t]$, showing that it is of order $N^{-1/2}$ in L^p , with an exponential decay in time as obtained from Lemma 3.2.

Lemma 3.4. *Let $p \geq 2$ and assume $\bar{\rho}_0 \in \mathcal{P}_p(\mathbb{R}^d)$. Then there exists a constant $C = C(p, \Lambda_\alpha, \bar{\rho}_0) > 0$ such that for all $t \geq 0$,*

$$\mathbb{E}|m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^p \leq \frac{C}{N^{\frac{p}{2}}} e^{-p(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t}.$$

Proof. By definition of $m_f^h[\cdot]$ and using that $\psi_h \geq h(\alpha)$, we find

$$\mathbb{E}|m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^p = \mathbb{E} \left| \frac{\sum_{i=1}^N (\bar{X}_t^i - m_f^h[\bar{\rho}_t]) \psi_h(\bar{X}_t^i)}{\sum_{i=1}^N \psi_h(\bar{X}_t^i)} \right|^p \leq \frac{1}{h(\alpha)^p} \mathbb{E} \left| \frac{1}{N} \sum_{i=1}^N (\bar{X}_t^i - m_f^h[\bar{\rho}_t]) \psi_h(\bar{X}_t^i) \right|^p.$$

The random variables

$$Y_i := (\bar{X}_t^i - m_f^h[\bar{\rho}_t]) \psi_h(\bar{X}_t^i)$$

are independent and centered (since $\mathbb{E}Y_i = 0$). Thus, by applying the Marcinkiewicz–Zygmund inequality (see, e.g., [40, Theorem 2, §10.3]), Jensen’s inequality, and using Lemma 3.2, we deduce

$$\begin{aligned} \mathbb{E}|m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^p &\leq \frac{C_{\text{MZ},p}}{N^{\frac{p}{2}} h(\alpha)^p} \mathbb{E} \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^2 \psi_h(\bar{X}_t^i)^2 \right)^{\frac{p}{2}} \\ &\leq \frac{C_{\text{MZ},p} \Lambda_\alpha^p}{N^{\frac{p}{2}}} \left[\frac{1}{N} \sum_{i=1}^N \mathbb{E} |\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^p \right] \\ &\leq \frac{C_{\text{MZ},p} \Lambda_\alpha^p}{N^{\frac{p}{2}}} \mathbb{E} |\bar{X}_t^1 - m_f^h[\bar{\rho}_t]|^p \\ &\leq \frac{C_{p,\Lambda_\alpha}}{N^{\frac{p}{2}}} \mathfrak{m}_p(\bar{\rho}_0) e^{-p(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t}. \end{aligned}$$

This completes the proof. $\square$

We now extend the L^p control to a uniform-in-time concentration inequality for the synchronous coupling. Under a quantitative contraction on λ , an exponentially weighted energy functional yields a drift-martingale estimate leading to a sharp-in- N tail bound for $\sup_{t \geq 0} e^{\bar{\kappa}t} V_p(\bar{\mu}_t^N)$, consistent with the exponential decay above.

Lemma 3.5 (Concentration for the synchronous coupling). *Let $p, q \geq 2$ and assume $\bar{\rho}_0 \in \mathcal{P}_{pq}(\mathbb{R}^d)$. For the particle trajectories, we suppose that $\bar{X}_0^i$ are i.i.d. with $\bar{X}_0^i \sim \bar{\rho}_0$ for all $i = 1, \dots, N$. Denoting*

$$\bar{c}_{\text{con},p} := 2\bar{\lambda}_{p,\alpha,\sigma} + 4(p-1)\sigma^2,$$

further assume

$$\lambda > \bar{\lambda}'_{p,q,\alpha,\sigma} := \bar{\lambda}_{pq,\alpha,\sigma} + \bar{c}_{\text{con},p}. \quad (3.5)$$

Then, for all small enough $\bar{\kappa} > 0$ and every $A > 0$, the following concentration estimate holds:

$$\mathbb{P} \left[\sup_{t \geq 0} e^{\bar{\kappa} t} V_p(\bar{\mu}_t^N) \geq 2^p V_p(\bar{\rho}_0) + 2^p A \right] \leq \frac{C}{A^q N^{\frac{q}{2}}}, \quad (3.6)$$

where $C > 0$ depends only on $\lambda, p, q, \sigma, \Lambda_\alpha$ and $\bar{\rho}_0$. In particular, (3.6) holds for all fixed $0 < \bar{\kappa} < p(\lambda - (\bar{c}_{\text{con}, p} \vee \bar{\lambda}_{pq, \alpha, \sigma}))$.

Proof. Step 1: Reduction to centered p -moments and a differential inequality. We first observe

$$\begin{aligned} V_p(\bar{\mu}_t^N) &:= \frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathcal{M}(\bar{\mu}_t^N)|^p \\ &\leq \frac{2^{p-1}}{N} \sum_{i=1}^N \left(|\bar{X}_t^i - \mathcal{M}(\bar{\rho}_t)|^p + \left| \mathcal{M}(\bar{\rho}_t) - \frac{1}{N} \sum_{i=1}^N \bar{X}_t^i \right|^p \right) \\ &\leq \frac{2^p}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathcal{M}(\bar{\rho}_t)|^p, \end{aligned} \quad (3.7)$$

and thus the evolution of $V_p(\bar{\mu}_t^N)$ can be estimated by that of $\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p$.

From the identity in (3.2) and Jensen's inequality, we have

$$\begin{aligned} d \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right) &\leq -\frac{p\lambda}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p dt + \frac{p(p-1)\sigma^2}{2N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^{p-2} |\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^2 dt \\ &\quad + \frac{p\sigma}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^{p-2} \langle \bar{X}_t^i - \mathbb{E} \bar{X}_t^i, \bar{X}_t^i - m_f^h[\bar{\rho}_t] \rangle dW_t^i. \end{aligned}$$

Here we estimate the second term on the right-hand side as follows. By Hölder's inequality, we first get

$$\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^{p-2} |\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^2 \leq \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right)^{\frac{p-2}{p}} \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^p \right)^{\frac{2}{p}}, \quad (3.8)$$

where the last term can be estimated further, using (3.4), Minkowski and Jensen's inequalities:

$$\begin{aligned} \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^p \right)^{\frac{2}{p}} &\leq 4 \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right)^{\frac{2}{p}} + 4 |\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^2 \\ &\quad + 4 |\mathcal{M}(\bar{\mu}_t^N) - m_f^h[\bar{\mu}_t^N]|^2 + 4 |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^2 \\ &\leq 4(1 + \Lambda_\alpha^{\frac{2}{p}}) \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right)^{\frac{2}{p}} + 4 |\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^2 \\ &\quad + 4 |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^2. \end{aligned}$$

Moreover, by using Young's inequality, we obtain

$$4 \left[\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right]^{\frac{p-2}{p}} |\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^2 \leq \frac{4(p-2)}{p} \left[\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right] + \frac{8}{p} |\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^p$$

$$\leq 4 \left[\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right] + \frac{8}{p} |\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^p,$$

and similarly,

$$4 \left[\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right]^{\frac{p-2}{p}} |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^2 \leq 4 \left[\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right] + \frac{8}{p} |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^p.$$

These combined with (3.8) yield

$$\begin{aligned} d \left(\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right) &\leq -p(\lambda - \bar{c}_{\text{con},p}) \left[\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right] dt \\ &\quad + 4(p-1) [|\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^p + |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^p] dt \\ &\quad + \frac{p\sigma}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^{p-2} \langle \bar{X}_t^i - \mathbb{E} \bar{X}_t^i, \bar{X}_t^i - m_f^h[\bar{\rho}_t] \rangle dW_t^i. \end{aligned} \quad (3.9)$$

By the hypothesis (3.5), we have $p(\lambda - \bar{c}_{\text{con},p}) > p\lambda_{pq,\alpha,\sigma} > 0$. Thus we fix

$$0 < \bar{\kappa} < p(\lambda - \bar{c}_{\text{con},p}),$$

to be further restricted later, and proceed to the weighted energy estimate in Step 2.

Step 2: Exponential weight and decomposition into drift + martingale. Motivated from [22], we define

$$\bar{\mathcal{E}}_t := \left[\frac{1}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^p \right] e^{\bar{\kappa}t}, \quad \bar{Z}_t := 4(p-1)e^{\bar{\kappa}t} [|\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^p + |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^p],$$

and

$$d\bar{M}_t := \left[\frac{p\sigma}{N} \sum_{i=1}^N |\bar{X}_t^i - \mathbb{E} \bar{X}_t^i|^{p-2} \langle \bar{X}_t^i - \mathbb{E} \bar{X}_t^i, \bar{X}_t^i - m_f^h[\bar{\rho}_t] \rangle dW_t^i \right] e^{\bar{\kappa}t}.$$

Multiplying (3.9) by $e^{\bar{\kappa}t}$ and using $\bar{\kappa} < p(\lambda - \bar{c}_{\text{con},p})$, we derive

$$d\bar{\mathcal{E}}_t \leq \bar{Z}_t dt + d\bar{M}_t \implies \bar{\mathcal{E}}_t \leq \bar{\mathcal{E}}_0 + \int_0^t \bar{Z}_s ds + \int_0^t d\bar{M}_s.$$

Using Markov's inequality, we can split and estimate the event of concentration as

$$\begin{aligned} &\mathbb{P} \left[\sup_{s \in [0,t]} \bar{\mathcal{E}}_s \geq \mathbb{E} \bar{\mathcal{E}}_0 + A \right] \\ &\leq \mathbb{P} \left[\sup_{s \in [0,t]} \left(\bar{\mathcal{E}}_0 + \int_0^s \bar{Z}_r dr + \int_0^s d\bar{M}_r \right) \geq \mathbb{E} \bar{\mathcal{E}}_0 + A \right] \\ &\leq \mathbb{P} \left[\bar{\mathcal{E}}_0 - \mathbb{E} \bar{\mathcal{E}}_0 \geq \frac{A}{3} \right] + \mathbb{P} \left[\sup_{s \in [0,t]} \left| \int_0^s \bar{Z}_r dr \right| \geq \frac{A}{3} \right] + \mathbb{P} \left[\sup_{s \in [0,t]} \left| \int_0^s d\bar{M}_r \right| \geq \frac{A}{3} \right] \\ &\leq \frac{3^q}{A^q} \mathbb{E} |\bar{\mathcal{E}}_0 - \mathbb{E} \bar{\mathcal{E}}_0|^q + \frac{3^q}{A^q} \mathbb{E} \left[\sup_{s \in [0,t]} \left| \int_0^s \bar{Z}_r dr \right|^q \right] + \frac{3^q}{A^q} \mathbb{E} \left[\sup_{s \in [0,t]} \left| \int_0^s d\bar{M}_r \right|^q \right] \\ &=: \frac{3^q}{A^q} (I + II + III). \end{aligned} \quad (3.10)$$

Step 3: Bounds for I , II , and III . First, by the Marcinkiewicz–Zygmund inequality and independence of $\{\bar{X}_0^i\}$, we estimate

$$\begin{aligned}
I &= \frac{1}{N^q} \mathbb{E} \left| \sum_{i=1}^N (|\bar{X}_0^i - \mathcal{M}(\bar{\rho}_0)|^p - \mathbb{E}|\bar{X}_0^1 - \mathcal{M}(\bar{\rho}_0)|^p) \right|^q \\
&\leq \frac{C_{\text{MZ},q}}{N^{\frac{q}{2}}} \mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N [|\bar{X}_0^i - \mathcal{M}(\bar{\rho}_0)|^p - \mathbb{E}|\bar{X}_0^1 - \mathcal{M}(\bar{\rho}_0)|^p]^2 \right]^{\frac{q}{2}} \\
&\leq \frac{C_{\text{MZ},q}}{N^{\frac{q}{2}}} \left[\frac{1}{N} \sum_{i=1}^N \mathbb{E} [|\bar{X}_0^i - \mathcal{M}(\bar{\rho}_0)|^p - \mathbb{E}|\bar{X}_0^1 - \mathcal{M}(\bar{\rho}_0)|^p]^q \right] \\
&\leq \frac{2^q C_{\text{MZ},q}}{N^{\frac{q}{2}}} \left[\frac{1}{N} \sum_{i=1}^N \mathbb{E} |\bar{X}_0^i - \mathcal{M}(\bar{\rho}_0)|^{pq} \right] \\
&\leq \frac{C}{N^{\frac{q}{2}}},
\end{aligned}$$

where $C > 0$ depends on p, q , and $\mathfrak{M}_{pq}(\bar{\rho}_0)$, and the last line follows from Lemma 3.2.

For the second term, we apply Minkowski's integral inequality and Lemmas 3.2, 3.4, to find

$$\begin{aligned}
II &\leq \mathbb{E} \left[\left| \int_0^\infty \bar{Z}_r dr \right|^q \right] \\
&\leq \left(\int_0^\infty (\mathbb{E} |\bar{Z}_r|^q)^{\frac{1}{q}} dr \right)^q \\
&\leq 4^q (p-1)^q \left(\int_0^\infty \left[(\mathbb{E} |\mathcal{M}(\bar{\rho}_t) - \mathcal{M}(\bar{\mu}_t^N)|^{pq})^{\frac{1}{q}} + (\mathbb{E} |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^{pq})^{\frac{1}{q}} \right] e^{\bar{\kappa}r} dr \right)^q \\
&\leq \frac{4^q (p-1)^q}{N^{\frac{pq}{2}}} \left(\int_0^\infty C e^{-(p(\lambda - \bar{\lambda}_{pq,\alpha,\sigma}) - \bar{\kappa})r} dr \right)^q \\
&\leq \frac{C}{N^{\frac{pq}{2}} (p(\lambda - \bar{\lambda}_{pq,\alpha,\sigma}) - \bar{\kappa})^q},
\end{aligned}$$

where $C > 0$ depends on $p, q, \Lambda_\alpha, \mathfrak{M}_{pq}(\bar{\rho}_0)$, and the last line is valid provided that $\bar{\kappa} < p(\lambda - \bar{\lambda}_{pq,\alpha,\sigma})$. Since $\lambda - \bar{\lambda}_{pq,\alpha,\sigma} > \bar{c}_{\text{con},p}$, it is clear that we can assume $\bar{\kappa}$ is small enough so that it lies in this range.

We finally estimate III . Under the given condition of λ , and owing to Lemma 3.2, we note that the process $\{\bar{M}_s\}_{s \geq 0}$ is a $\mathbb{P}$ -martingale. Then we have

$$\begin{aligned}
\mathbb{E} \left[\sup_{s \leq t} |\bar{M}_t|^q \right] &\leq C_{\text{BDG},q} \mathbb{E} \left[\langle \bar{M} \rangle_t^{\frac{q}{2}} \right] \\
&\leq \frac{C_{\text{BDG},q} p^q \sigma^q}{N^{\frac{q}{2}}} \mathbb{E} \left[\frac{1}{N} \sum_{i=1}^N \left(\int_0^t e^{2\bar{\kappa}s} |\bar{X}_s^i - \mathbb{E} \bar{X}_s^i|^{2p-2} |\bar{X}_s^i - m_f^h[\bar{\rho}_s]|^2 ds \right)^{\frac{q}{2}} \right] \\
&= \frac{C_{\text{BDG},q} p^q \sigma^q}{N^{\frac{q}{2}}} \mathbb{E} \left[\int_0^t e^{2\bar{\kappa}s} |\bar{X}_s^1 - \mathbb{E} \bar{X}_s^1|^{2p-2} |\bar{X}_s^1 - m_f^h[\bar{\rho}_s]|^2 ds \right]^{\frac{q}{2}} \\
&\leq \frac{C_{\text{BDG},q} p^q \sigma^q}{N^{\frac{q}{2}}} \left[\int_0^t e^{2\bar{\kappa}s} \left(\mathbb{E} \left[|\bar{X}_s^1 - \mathbb{E} \bar{X}_s^1|^{(p-1)q} |\bar{X}_s^1 - m_f^h[\bar{\rho}_s]|^q \right] \right)^{\frac{2}{q}} ds \right]^{\frac{q}{2}}
\end{aligned}$$

$$\begin{aligned}
&\leq \frac{C_{\text{BDG},q} p^q \sigma^q}{N^{\frac{q}{2}}} \left[\int_0^t e^{2\bar{\kappa}s} \left[\left(\mathbb{E}|\bar{X}_s^1 - \mathbb{E}\bar{X}_s^1|^{pq} \right)^{\frac{p-1}{p}} \left(\mathbb{E}|\bar{X}_s^1 - m_f^h[\bar{\rho}_s]|^{pq} \right)^{\frac{1}{p}} \right]^{\frac{2}{q}} ds \right]^{\frac{q}{2}} \\
&\leq \frac{C(p, q, \sigma, \mathfrak{M}_{pq}(\bar{\rho}_0), \Lambda_\alpha)}{N^{\frac{q}{2}}} \left[\int_0^t e^{(2\bar{\kappa} - 2p(\lambda - \bar{\lambda}_{pq, \alpha, \sigma}))s} ds \right]^{\frac{q}{2}} \\
&\leq \frac{C}{N^{\frac{q}{2}} (p(\lambda - \bar{\lambda}_{pq, \alpha, \sigma}) - \bar{\kappa})^{\frac{q}{2}}}.
\end{aligned}$$

In the above, the first line is due to the BDG inequality; the second line is by definition of the quadratic variation, the Cauchy–Schwarz inequality, and the following estimate

$$\left(\frac{1}{N^2} \sum_{i=1}^N \int_0^t |\sigma_i(s)|^2 ds \right)^{\frac{q}{2}} = \frac{1}{N^{\frac{q}{2}}} \left(\frac{1}{N} \sum_{i=1}^N \int_0^t |\sigma_i(s)|^2 ds \right)^{\frac{q}{2}} \leq \frac{1}{N^{\frac{q}{2}}} \frac{1}{N} \sum_{i=1}^N \left(\int_0^t |\sigma_i(s)|^2 ds \right)^{\frac{q}{2}},$$

which holds by Jensen for general sufficiently integrable processes σ_i . The third line then follows as the $\bar{X}_t^i$ are i.i.d. Next, the fourth line is by Minkowski’s integral inequality, and the fifth is due to Hölder. Finally, the second-to-last line is due to Lemma 3.2 and the last line follows from direct computation.

Step 4: Conclusion and choice of $\bar{\kappa}$. The bounds for I, II, III obtained in Step 3 are independent of t . Hence, by the bounded convergence theorem we can pass to the limit $t \rightarrow \infty$ in (3.10) and deduce

$$\mathbb{P} \left[\sup_{t \geq 0} \bar{\mathcal{E}}_t \geq \mathbb{E}\bar{\mathcal{E}}_0 + A \right] \leq \frac{C}{N^{\frac{q}{2}} A^q} \left[1 + \frac{1}{N^{\frac{q}{2}(p-1)} (p(\lambda - \bar{\lambda}_{pq, \alpha, \sigma}) - \bar{\kappa})^q} + \frac{1}{(p(\lambda - \bar{\lambda}_{pq, \alpha, \sigma}) - \bar{\kappa})^{\frac{q}{2}}} \right].$$

Recalling (3.7), we have the relations

$$2^p \bar{\mathcal{E}}_t \geq e^{\bar{\kappa}t} V_p(\bar{\mu}_t^N), \quad \mathbb{E}\bar{\mathcal{E}}_0 = V_p(\bar{\rho}_0).$$

Consequently,

$$\mathbb{P} \left[\sup_{t \geq 0} e^{\bar{\kappa}t} V_p(\bar{\mu}_t^N) \geq 2^p \mathbb{E}V_p(\bar{\rho}_0) + 2^p A \right] \leq \mathbb{P} \left[\sup_{t \geq 0} \bar{\mathcal{E}}_t \geq \mathbb{E}V_p(\bar{\rho}_0) + A \right].$$

Combining the two estimates yields the claimed concentration inequality

$$\mathbb{P} \left[\sup_{t \geq 0} e^{\bar{\kappa}t} V_p(\bar{\mu}_t^N) \geq 2^p V_p(\bar{\rho}_0) + 2^p A \right] \leq \frac{C}{A^q N^{\frac{q}{2}}},$$

where $C > 0$ depends only on $\lambda, p, q, \sigma, \Lambda_\alpha$, and $\bar{\rho}_0$. Finally, $\bar{\kappa}$ can be chosen in the admissible range

$$0 < \bar{\kappa} < p(\lambda - (\bar{c}_{\text{con}, p} \vee \bar{\lambda}_{pq, \alpha, \sigma})),$$

which ensures all exponential integrals above are finite. $\square$

An analogous argument applies to the particle trajectories $\{X_t^i\}_{i=1}^N$, yielding a concentration estimate of the same form as in Lemma 3.5. The proof closely follows the reasoning of [22, Remark 4.10], with modifications analogous to those used in Lemma 3.5. For completeness, we include a detailed proof in Appendix C.

Lemma 3.6 (Concentration for the particle system). *Let $p, q \geq 2$ and assume $X_{\text{in}} \in L^{pq}(\Omega)$. For the particle trajectories, we suppose that X_0^i are i.i.d. with $X_0^i \sim X_{\text{in}}$ for all $i = 1, \dots, N$. Denoting*

$$c_{\text{con},p} := 2\Lambda_\alpha^{1-\frac{2}{p}} \lambda_{p,\alpha,\sigma} + 2(p-1)\sigma^2,$$

further assume

$$\lambda > \lambda'_{p,q,\alpha,\sigma} := \lambda_{pq,\alpha,\sigma} + c_{\text{con},p}. \quad (3.11)$$

Then for all small enough $\kappa > 0$ and every $A > 0$, the following concentration estimate holds:

$$\mathbb{P} \left[\sup_{t \geq 0} e^{\kappa t} V_p(\mu_t^N) \geq \mathbb{E} [V_p(\mu_0^N)] + A \right] \leq \frac{C}{A^q N^{\frac{q}{2}}}, \quad (3.12)$$

for a constant $C > 0$ independent of A and N . In particular, (3.12) holds for all fixed $0 < \kappa < p(\lambda - (c_{\text{con},p} \vee \lambda_{pq,\alpha,\sigma}))$.

3.4. Proof of Theorem 1.2. We now establish the uniform-in-time propagation of chaos. Throughout this subsection, we assume that

$$\lambda > \tilde{\lambda}_{p,q,\alpha,\sigma} := \max\{\bar{\lambda}'_{p,q,\alpha,\sigma}, \lambda'_{p,q,\alpha,\sigma}\} \quad (3.13)$$

where we recall that $\bar{\lambda}'_{p,q,\alpha,\sigma}$ is given in (3.5), and $\lambda'_{p,q,\alpha,\sigma}$ in (3.11). This condition ensures that both Lemmas 3.5 and 3.6 are applicable. Note that we can choose $\tilde{\kappa} > 0$ satisfying

$$\tilde{\kappa} < p \left[\left(\lambda - (\bar{c}_{\text{con},p} \vee \bar{\lambda}_{pq,\alpha,\sigma}) \right) \wedge \left(\lambda - (c_{\text{con},p} \vee \lambda_{pq,\alpha,\sigma}) \right) \right], \quad (3.14)$$

so that both concentration estimates (3.6) and (3.12) hold simultaneously with $\tilde{\kappa}$ in place of $\bar{\kappa}$ and κ , respectively.

Moreover, the inequalities (3.13) and (3.5) together imply

$$\lambda \geq \bar{\lambda}'_{p,q,\alpha,\sigma} > \bar{c}_{\text{con},p} > 2\bar{\lambda}_{p,\alpha,\sigma} > 2\lambda_{p,\alpha,\sigma}.$$

This guarantees that all preceding results, such as Lemmas 2.2 and 3.2, can be employed in the forthcoming analysis.

Following the classical strategy, we introduce the fluctuation variable

$$Z_t^i := X_t^i - \bar{X}_t^i.$$

For convenience, we also set

$$\begin{aligned} \Delta_t^{\mathcal{M}} &= m_f^h[\mu_t^N] - m_f^h[\bar{\mu}_t^N], \quad \Delta_t^{\mathcal{M}} = \mathcal{M}(\mu_t^N) - \mathcal{M}(\bar{\mu}_t^N), \\ \bar{\mathcal{G}}_{t,p} &:= \frac{1}{N} \sum_{i=1}^N |Z_t^i|^p, \quad \bar{\mathcal{D}}_{t,p} := \frac{1}{N} \sum_{i=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^p, \quad \bar{\mathcal{O}}_{t,p} := |\Delta_t^{\mathcal{M}}|^p. \end{aligned} \quad (3.15)$$

Then each Z_t^i satisfies

$$dZ_t^i = -\lambda \left(Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \right) dt + \sigma \left(Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \right) dW_t^i.$$

3.4.1. Estimates for the fluctuation functionals. As discussed in Section 3.1, we regard the right-hand side as a perturbation of the mean-field dynamics. It is therefore crucial to control the discrepancy $\Delta_t^{\mathcal{M}} - \Delta_t^m$, which quantifies the deviation between the empirical means of the interacting particle system and the mean-field process. The following lemma, relying on the concentration estimates from Lemmas 3.5 and 3.6, provides an L^p -estimate for this quantity.

Lemma 3.7 (Control of mean discrepancies). *Let $p \geq 2$ and $q > 2$. Let the law $\bar{\rho}_0$ (corresponding to $\bar{X}_0$) have finite pq -moments. Then whenever $\lambda > \tilde{\lambda}_{p,q,\alpha,\sigma}$, and $\tilde{\kappa}$ satisfies (3.14), there exists $0 < \theta < \tilde{\kappa}$ for which*

$$\mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^p \leq C \left[\frac{1}{N^{q-2}} + \mathbb{E}\bar{\mathcal{G}}_{t,p} \right] e^{-\theta t},$$

where $C > 0$ depends only on $\lambda, \alpha, \sigma, p, q$ and $\bar{\rho}_0$.

Proof. We recall from Lemma 3.1 that for $p \geq 2$:

$$|\Delta_t^{\mathcal{M}} - \Delta_t^m|^p \leq \frac{1}{2} \left(\frac{2L_{\omega_f} \Lambda_\alpha}{h(\alpha)} \right)^p (V_p(\mu_t^N) + V_p(\bar{\mu}_t^N)) \mathcal{W}_p^p(\mu_t^N, \bar{\mu}_t^N). \quad (3.16)$$

Thus, it suffices to estimate the expectation of

$$(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N)) \mathcal{W}_p^p(\mu_t^N, \bar{\mu}_t^N). \quad (3.17)$$

We will compute the expectation of (3.17) by writing $\Omega = \Omega_{\tilde{\kappa}} \cup (\Omega \setminus \Omega_{\tilde{\kappa}})$, where

$$\Omega_{\tilde{\kappa}} = \left\{ \omega \in \Omega : \sup_{t \geq 0} e^{\tilde{\kappa}t} V_p(\mu_t^N) \geq V_p(\bar{\rho}_0) + 1 \right\} \cup \left\{ \omega \in \Omega : \sup_{t \geq 0} e^{\tilde{\kappa}t} V_p(\bar{\mu}_t^N) \geq 2^p V_p(\bar{\rho}_0) + 2^p \right\}.$$

By definition of $\Omega_{\tilde{\kappa}}$, we have

$$\forall t \geq 0, \quad (V_p(\mu_t^N) + V_p(\bar{\mu}_t^N)) \mathbf{1}_{\Omega \setminus \Omega_{\tilde{\kappa}}} \leq (2^p + 1)(V_p(\bar{\rho}_0) + 1) e^{-\tilde{\kappa}t}.$$

On the other hand, by subadditivity of $\mathbb{P}$ and Lemmas 3.5, 3.6, we find

$$\mathbb{P}[\Omega_{\tilde{\kappa}}] \leq \frac{C}{N^{\frac{q}{2}}}.$$

Thus, the expectation of (3.17) is computed as

$$\begin{aligned} & \mathbb{E}[(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N)) \mathcal{W}_p^p(\mu_t^N, \bar{\mu}_t^N)] \\ &= \mathbb{E}[(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N)) \mathcal{W}_p^p(\mu_t^N, \bar{\mu}_t^N) \mathbf{1}_{\Omega_{\tilde{\kappa}}}] + \mathbb{E}[(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N)) \mathcal{W}_p^p(\mu_t^N, \bar{\mu}_t^N) \mathbf{1}_{\Omega \setminus \Omega_{\tilde{\kappa}}}] \\ &\leq \mathbb{P}[\Omega_{\tilde{\kappa}}]^{\frac{q-2}{q}} \left(\mathbb{E}[(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N))^{q/2} \mathcal{W}_p^{pq/2}(\mu_t^N, \bar{\mu}_t^N)] \right)^{\frac{2}{q}} + e^{-\tilde{\kappa}t} C_1 \mathbb{E}\bar{\mathcal{G}}_{t,p}, \end{aligned} \quad (3.18)$$

where $C_1 := (2^p + 1)(V_p(\bar{\rho}_0) + 1)$. It remains to compute the term in parentheses. Applying the Cauchy-Schwarz and Jensen's inequalities, and using (3.7), we obtain

$$\begin{aligned} & \mathbb{E}[(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N))^{q/2} \mathcal{W}_p^{pq/2}(\mu_t^N, \bar{\mu}_t^N)] \\ &\leq \left(\mathbb{E}[(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N))^q] \right)^{\frac{1}{2}} \left(\mathbb{E}\mathcal{W}_p^{pq}(\mu_t^N, \bar{\mu}_t^N) \right)^{\frac{1}{2}} \\ &\leq (2^{q-1} \mathbb{E}V_{pq}(\mu_t^N) + 2^{q-1} \mathbb{E}V_{pq}(\bar{\mu}_t^N))^{\frac{1}{2}} (\mathbb{E}\bar{\mathcal{G}}_{t,pq})^{\frac{1}{2}} \end{aligned}$$

$$\leq \left(2^{q-1} \mathbb{E} V_{pq}(\mu_t^N) + \frac{2^{(p+1)q-1}}{N} \sum_{i=1}^N \mathbb{E} |\bar{X}_t^i - \mathcal{M}(\bar{\rho}_t)|^{pq} \right)^{\frac{1}{2}} (\mathbb{E} \bar{\mathcal{G}}_{t,pq})^{\frac{1}{2}}.$$

The right-hand side can be estimated further as follows. We use Lemmas 2.2 and 3.2 to have

$$\begin{aligned} 2^{q-1} \mathbb{E} V_{pq}(\mu_t^N) &= \frac{2^{q-1}}{N} \sum_{i=1}^N \mathbb{E} |X_t^i - \mathcal{M}(\mu_t^N)|^{pq} \\ &\leq 2^{q-1} \mathbb{E} |X_0^1 - X_0^2|^{pq} e^{-pq(\lambda - \bar{\lambda}_{pq,\alpha,\sigma})t} \\ &\leq C \mathfrak{m}_{pq}(\bar{\rho}_0) e^{-pq(\lambda - \bar{\lambda}_{pq,\alpha,\sigma})t} \end{aligned}$$

and

$$\frac{2^{(p+1)q-1}}{N} \sum_{i=1}^N \mathbb{E} |\bar{X}_t^i - \mathcal{M}(\bar{\rho}_t)|^{pq} \leq C \mathfrak{m}_{pq}(\bar{\rho}_0) e^{-pq(\lambda - \bar{\lambda}_{pq,\alpha,\sigma})t}.$$

Using Lemmas 2.3 and 3.3, we also find

$$\mathbb{E} \mathcal{G}_{t,pq} = \frac{1}{N} \sum_{i=1}^N \mathbb{E} |X_t^i - \bar{X}_t^i|^{pq} \leq \frac{2^{pq-1}}{N} \sum_{i=1}^N (\mathbb{E} |X_t^i|^{pq} + \mathbb{E} |\bar{X}_t^i|^{pq}) \leq C,$$

since we assume $\bar{\rho}_0$ has finite pq -th moments. Thus, combining the above estimates yields

$$\mathbb{E} \left[(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N))^{q/2} \mathcal{W}_p^{pq/2}(\mu_t^N, \bar{\mu}_t^N) \right] \leq C e^{-\frac{pq}{2}(\lambda - \bar{\lambda}_{pq,\alpha,\sigma})t}.$$

where $C > 0$ depends on $p, q, \sigma, \Lambda_\alpha, \bar{\rho}_0$ but is independent of N and t . Subsequently, inserting this into (3.18) gives

$$\mathbb{E} [(V_p(\mu_t^N) + V_p(\bar{\mu}_t^N)) \mathcal{W}_p^p(\mu_t^N, \bar{\mu}_t^N)] \leq \frac{C}{N^{\frac{q-2}{2}}} e^{-p(\lambda - \bar{\lambda}_{pq,\alpha,\sigma})t} + C_1 \mathbb{E} \bar{\mathcal{G}}_{t,p} e^{-\tilde{\kappa}t}.$$

Finally, taking expectations in (3.16), then inserting the above bound, we conclude the desired result. $\square$

At this point we remind the reader of the notations in (3.15). Our goal is to control the evolution of $\bar{\mathcal{G}}_{t,p}$ uniformly in time and with explicit N -dependence. Using the elementary inequality

$$\bar{\mathcal{G}}_{t,p} \leq 2^{p-1} (\bar{\mathcal{D}}_{t,p} + \bar{\mathcal{O}}_{t,p}) \leq (2^{2p-1} + 2^{p-1}) \bar{\mathcal{G}}_{t,p}, \quad (3.19)$$

it suffices to estimate $\bar{\mathcal{D}}_{t,p}$ and $\bar{\mathcal{O}}_{t,p}$. We begin with $\bar{\mathcal{D}}_{t,p}$.

Lemma 3.8 (Estimate for $\bar{\mathcal{D}}_{t,p}$). *Assume that the assumptions of Theorem 1.2 hold. Then for any*

$$\tilde{\theta} \in \left(0, \frac{\theta}{p} \wedge (\lambda - \bar{\lambda}_{p,\alpha,\sigma}) \right)$$

(where θ is the constant from Lemma 3.7), there exists a positive constant $C > 0$, dependent only on $\lambda, \sigma, p, q, \Lambda_\alpha$, and $\bar{\rho}_0$, such that

$$\frac{d}{dt} \mathbb{E} \bar{\mathcal{D}}_{t,p} \leq C \left(\mathbb{E} \bar{\mathcal{D}}_{t,p} e^{-\tilde{\theta}t} + \mathbb{E} \bar{\mathcal{O}}_{t,p} e^{-\tilde{\theta}t} + \frac{1}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} e^{-\tilde{\theta}t} \right). \quad (3.20)$$

Proof. Applying the Itô formula to $\bar{\mathcal{D}}_{t,p} = \frac{1}{N} \sum_{i=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^p$ and using the SDE for Z_t^i yields

$$d\bar{\mathcal{D}}_{t,p} = -p\lambda \bar{\mathcal{D}}_{t,p} dt$$

$$+ \frac{p\sigma^2}{2N} \left(1 - \frac{2}{N}\right) \sum_{i=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^2 dt \quad (3.21)$$

$$+ \frac{p\sigma^2}{2N^3} \sum_{i=1}^N \sum_{j=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |Z_t^j - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^2 dt \quad (3.22)$$

$$+ \frac{p(p-2)\sigma^2}{2N} \left(1 - \frac{2}{N}\right) \sum_{i=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-4} \langle Z_t^i - \Delta_t^{\mathcal{M}}, Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \rangle^2 dt \quad (3.23)$$

$$+ \frac{p(p-2)\sigma^2}{2N^3} \sum_{i=1}^N \sum_{j=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-4} \langle Z_t^i - \Delta_t^{\mathcal{M}}, Z_t^j - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \rangle^2 dt \quad (3.24)$$

$$+ \frac{p\sigma}{N} \sum_{i=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} \langle Z_t^i - \Delta_t^{\mathcal{M}}, Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \rangle dW_t^i \\ - \frac{p\sigma}{N^2} \sum_{i=1}^N \sum_{j=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} \langle Z_t^i - \Delta_t^{\mathcal{M}}, Z_t^j - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \rangle dW_t^j.$$

Taking expectations, the stochastic integrals vanish. Regarding the expectation of (3.21), we use simple algebraic manipulations such as

$$\begin{aligned} & \sum_{i=1}^N \mathbb{E} |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^2 \\ &= \sum_{i=1}^N \mathbb{E} |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |Z_t^i - \Delta_t^{\mathcal{M}} + \Delta_t^{\mathcal{M}} - \Delta_t^m - m_f^h[\bar{\mu}_t^N] + m_f^h[\bar{\rho}_t]|^2 \\ &\leq 3\mathbb{E}\mathcal{D}_{t,p} + 3 \sum_{i=1}^N \mathbb{E} |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |\Delta_t^{\mathcal{M}} - \Delta_t^m|^2 + 3 \sum_{i=1}^N \mathbb{E} |Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]|^2. \end{aligned}$$

An analogous estimate holds for (3.23). Grouping (3.22) and (3.24) together, we then have at this point

$$\begin{aligned} \frac{d}{dt} \mathbb{E}\bar{\mathcal{D}}_{t,p} &\leq -p \left(\lambda - \frac{3(p-1)\sigma^2}{2} \right) \mathbb{E}\bar{\mathcal{D}}_{t,p} + \frac{3p(p-1)\sigma^2}{2N} \sum_{i=1}^N \mathbb{E} [|Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |\Delta_t^{\mathcal{M}} - \Delta_t^m|^2] \\ &\quad + \frac{3p(p-1)\sigma^2}{2N} \sum_{i=1}^N \mathbb{E} [|Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^2] \\ &\quad + \frac{p(p-1)\sigma^2}{2N^3} \sum_{i=1}^N \sum_{j=1}^N \mathbb{E} [|Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |Z_t^j - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^2] \\ &=: -p \left(\lambda - \frac{3(p-1)\sigma^2}{2} \right) \mathbb{E}\bar{\mathcal{D}}_{t,p} + I + II + III. \end{aligned}$$

By our assumption on λ , the first term in the inequality above is negative, thus it suffices to estimate the remaining terms, I , II , and III .

For the estimate of I , by using Hölder's inequality, Lemma 3.7, and (3.19), we obtain

$$I \leq \frac{3p(p-1)\sigma^2}{2} (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} (\mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^p)^{\frac{2}{p}}$$

$$\begin{aligned}
&\leq C (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} \left[\frac{1}{N^{\frac{q-2}{2}}} + \mathbb{E}\bar{\mathcal{G}}_{t,p} \right]^{\frac{2}{p}} e^{-\frac{2}{p}\theta t} \\
&\leq C (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} \left[\frac{1}{N^{\frac{q-2}{2}}} + 2^{p-1}(\mathbb{E}\bar{\mathcal{D}}_{t,p} + \mathbb{E}\bar{\mathcal{O}}_{t,p}) \right]^{\frac{2}{p}} e^{-\frac{2}{p}\theta t} \\
&\leq C \left[\frac{1}{N^{\frac{q-2}{p}}} (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} + 4\mathbb{E}\bar{\mathcal{D}}_{t,p} + 4 (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} (\mathbb{E}\bar{\mathcal{O}}_{t,p})^{\frac{2}{p}} \right] e^{-\frac{2}{p}\theta t} \\
&\leq C \left[\frac{1}{N^{\frac{q-2}{2}}} + \mathbb{E}\bar{\mathcal{D}}_{t,p} + \mathbb{E}\bar{\mathcal{O}}_{t,p} \right] e^{-\frac{2}{p}\theta t},
\end{aligned}$$

where the last line follows from Young's inequality.

In a similar manner, by using Lemma 3.4 and Young's inequality, we estimate

$$\begin{aligned}
II &\leq \frac{3p(p-1)\sigma^2}{2N} (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} (\mathbb{E}|m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]|^p)^{\frac{2}{p}} \\
&\leq \frac{C}{N^2} (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} e^{-2(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t} \\
&\leq C \left(\frac{1}{N^p} + \mathbb{E}\bar{\mathcal{D}}_{t,p} \right) e^{-2(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t}.
\end{aligned}$$

For III, we recall (see the beginning of Section 3.4) that both Lemmas 2.2 and 3.2 are available due to λ being large enough. Hence, we have

$$\begin{aligned}
III &\leq \frac{p(p-1)\sigma^2}{2N^3} \sum_{i=1}^N \sum_{j=1}^N (\mathbb{E}|Z_t^i - \Delta_t^{\mathcal{M}}|^p)^{\frac{p-2}{p}} \left(\mathbb{E}|Z_t^j - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^p \right)^{\frac{2}{p}} \\
&\leq \frac{p(p-1)\sigma^2}{N} (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} \left(\frac{1}{N} \sum_{i=1}^N (\mathbb{E}|X_t^i - m_f^h[\mu_t^N]|^p)^{\frac{2}{p}} + (\mathbb{E}|\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^p)^{\frac{2}{p}} \right) \\
&\leq \frac{C}{N} (\mathbb{E}\bar{\mathcal{D}}_{t,p})^{\frac{p-2}{p}} e^{-2(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t} \\
&\leq C \left[\frac{1}{N^{\frac{p}{2}}} + \mathbb{E}\bar{\mathcal{D}}_{t,p} \right] e^{-2(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t}.
\end{aligned}$$

Collecting the bounds for I, II, III and choosing $\tilde{\theta} > 0$ sufficiently small ($\tilde{\theta} < \frac{\theta}{p}$ and $\tilde{\theta} < \lambda - \bar{\lambda}_{p,\alpha,\sigma}$), we absorb the exponentially decaying factors into $e^{-\tilde{\theta}t}$ and obtain (3.20) with a suitable constant $C > 0$ that is independent of t and N . $\square$

Lemma 3.9 (Estimate for $\bar{\mathcal{O}}_{t,p}$). *Assume that the assumptions of Theorem 1.2 hold. Then there exists a constant $C > 0$, depending only on $\lambda, \sigma, p, q, \Lambda_\alpha$ and $\bar{\rho}_0$, such that for the same $\tilde{\theta} > 0$ as in Lemma 3.8:*

$$\frac{d}{dt} \mathbb{E}\bar{\mathcal{O}}_{t,p} \leq C \left(\mathbb{E}\bar{\mathcal{O}}_{t,p} e^{-\tilde{\theta}t} + \mathbb{E}\bar{\mathcal{D}}_{t,p} e^{-\tilde{\theta}t} + \frac{1}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} e^{-\tilde{\theta}t} \right). \quad (3.25)$$

Proof. Applying the Itô formula to $\bar{\mathcal{O}}_{t,p}$ gives

$$\begin{aligned}
d\bar{\mathcal{O}}_{t,p} &= -p\lambda |\Delta_t^{\mathcal{M}}|^{p-2} \langle \Delta_t^{\mathcal{M}}, \Delta_t^{\mathcal{M}} - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \rangle dt \\
&\quad + \frac{p\sigma^2}{2N^2} |\Delta_t^{\mathcal{M}}|^{p-2} \sum_{i=1}^N |Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^2 dt
\end{aligned}$$

$$\begin{aligned}
& + \frac{p(p-2)\sigma^2}{2N^2} |\Delta_t^{\mathcal{M}}|^{p-4} \sum_{i=1}^N \langle \Delta_t^{\mathcal{M}}, Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \rangle^2 dt \\
& + \frac{p\sigma}{N} |\Delta_t^{\mathcal{M}}|^{p-2} \sum_{i=1}^N \langle \Delta_t^{\mathcal{M}}, Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t]) \rangle dW_t^i.
\end{aligned}$$

Taking expectations and using the Cauchy-Schwarz inequality, we get

$$\begin{aligned}
\frac{d}{dt} \mathbb{E} \bar{\mathcal{O}}_{t,p} & \leq -p\lambda \mathbb{E} [|\Delta_t^{\mathcal{M}}|^{p-2} \langle \Delta_t^{\mathcal{M}}, \Delta_t^{\mathcal{M}} - \Delta_t^m \rangle] + p\lambda \mathbb{E} [|\Delta_t^{\mathcal{M}}|^{p-2} \langle \Delta_t^{\mathcal{M}}, m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t] \rangle] \\
& + \frac{p(p-1)\sigma^2}{2N^2} \mathbb{E} \left[|\Delta_t^{\mathcal{M}}|^{p-2} \sum_{i=1}^N |Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^2 \right] \\
& =: I + II + III,
\end{aligned}$$

where by using a similar argument as in the proof of Lemma 3.8, we estimate

$$\begin{aligned}
I & \leq p\lambda \mathbb{E} [|\Delta_t^{\mathcal{M}}|^{p-1} |\Delta_t^{\mathcal{M}} - \Delta_t^m|] \\
& \leq p\lambda (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-1}{p}} (\mathbb{E} |\Delta_t^{\mathcal{M}} - \Delta_t^m|^p)^{\frac{1}{p}} \\
& \leq C (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-1}{p}} \left[\frac{1}{N^{\frac{q-2}{2}}} + \mathbb{E} \bar{\mathcal{G}}_{t,p} \right]^{\frac{1}{p}} e^{-\frac{1}{p}\theta t} \\
& \leq C (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-1}{p}} \left[\frac{1}{N^{\frac{q-2}{2}}} + 2^{p-1} (\mathbb{E} \bar{\mathcal{D}}_{t,p} + \mathbb{E} \bar{\mathcal{O}}_{t,p}) \right]^{\frac{1}{p}} e^{-\frac{1}{p}\theta t} \\
& \leq C \left[\frac{1}{N^{\frac{q-2}{2p}}} (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-1}{p}} + 2\mathbb{E} \bar{\mathcal{O}}_{t,p} + 2 (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-1}{p}} (\mathbb{E} \bar{\mathcal{D}}_{t,p})^{\frac{1}{p}} \right] e^{-\frac{1}{p}\theta t} \\
& \leq C \left[\frac{1}{N^{\frac{q-2}{2}}} + \mathbb{E} \bar{\mathcal{O}}_{t,p} + \mathbb{E} \bar{\mathcal{D}}_{t,p} \right] e^{-\frac{1}{p}\theta t}, \\
II & \leq p\lambda \mathbb{E} [|\Delta_t^{\mathcal{M}}|^{p-1} |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|] \\
& \leq p\lambda (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-1}{p}} (\mathbb{E} |m_f^h[\bar{\mu}_t^N] - m_f^h[\bar{\rho}_t]|^p)^{\frac{1}{p}} \\
& \leq \frac{C}{N^{\frac{1}{2}}} (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-1}{p}} e^{-(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t} \\
& \leq C \left(\frac{1}{N^{\frac{p}{2}}} + \mathbb{E} \bar{\mathcal{O}}_{t,p} \right) e^{-(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t},
\end{aligned}$$

and

$$\begin{aligned}
III & \leq \frac{p(p-1)\sigma^2}{2N} \left[\frac{1}{N} \sum_{i=1}^N (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-2}{p}} (\mathbb{E} |Z_t^i - (m_f^h[\mu_t^N] - m_f^h[\bar{\rho}_t])|^p)^{\frac{2}{p}} \right] \\
& \leq \frac{p(p-1)\sigma^2}{N} (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-2}{p}} \left(\frac{1}{N} \sum_{i=1}^N (\mathbb{E} |X_t^i - m_f^h[\mu_t^N]|^p)^{\frac{2}{p}} + (\mathbb{E} |\bar{X}_t^i - m_f^h[\bar{\rho}_t]|^p)^{\frac{2}{p}} \right) \\
& \leq \frac{C}{N} (\mathbb{E} \bar{\mathcal{O}}_{t,p})^{\frac{p-2}{p}} e^{-2(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t} \\
& \leq C \left[\frac{1}{N^{\frac{p}{2}}} + \mathbb{E} \bar{\mathcal{O}}_{t,p} \right] e^{-2(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t}.
\end{aligned}$$

Collecting the bounds and choosing $\tilde{\theta} \in (0, \frac{\theta}{p} \wedge (\lambda - \bar{\lambda}_{p,\alpha,\sigma}))$, we see that the decay rate of I, II, III is given by $e^{-\tilde{\theta}t}$, and we thus obtain (3.25) up to constructive constants. $\square$

3.4.2. *Uniform-in-time bound for the fluctuation energy.* Combining (3.20) and (3.25) yields

$$\frac{d}{dt} (\mathbb{E}\bar{\mathcal{D}}_{t,p} + \mathbb{E}\bar{\mathcal{O}}_{t,p}) \leq C (\mathbb{E}\bar{\mathcal{D}}_{t,p} + \mathbb{E}\bar{\mathcal{O}}_{t,p}) e^{-\tilde{\theta}t} + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} e^{-\tilde{\theta}t}$$

for some constant $C > 0$ depending only on $\lambda, \sigma, p, q, \Lambda_\alpha, \bar{\rho}_0$. Thus, applying Grönwall's inequality yields

$$\mathbb{E}\bar{\mathcal{D}}_{t,p} + \mathbb{E}\bar{\mathcal{O}}_{t,p} \leq \left[\mathbb{E}\bar{\mathcal{D}}_{0,p} + \mathbb{E}\bar{\mathcal{O}}_{0,p} + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} \right] e^{C/\tilde{\theta}},$$

and by (3.19), we deduce

$$\mathbb{E}\bar{\mathcal{G}}_{t,p} \leq \left[(2^{2p-1} + 2^{p-1}) \mathbb{E}\bar{\mathcal{G}}_{0,p} + \frac{2^{p-1}C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} \right] e^{C/\tilde{\theta}}.$$

Hence, for $\lambda > \tilde{\lambda}_{p,q,\alpha,\sigma}$, there exist a positive constant $C > 0$ dependent only on $\lambda, \sigma, p, q, \Lambda_\alpha$ and $\bar{\rho}_0$ such that

$$\sup_{t \geq 0} \frac{1}{N} \sum_{i=1}^N \mathbb{E}|X_t^i - \bar{X}_t^i|^p \leq \frac{C}{N} \sum_{i=1}^N \mathbb{E}|X_0^i - \bar{X}_0^i|^p + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}}.$$

3.4.3. *Mean-field convergence and propagation of chaos.* We now rigorously deduce the uniform-in-time propagation of chaos from the quantitative estimates obtained above. The key quantity of interest is the p -Wasserstein distance between the empirical distribution of the particle system and the limiting mean-field distribution. By the triangle inequality, we can decompose it as

$$\mathcal{W}_p(\mu_t^N, \bar{\rho}_t) \leq \mathcal{W}_p(\mu_t^N, \bar{\mu}_t^N) + \mathcal{W}_p(\bar{\mu}_t^N, \bar{\rho}_t), \quad (3.26)$$

where $\bar{\mu}_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{\bar{X}_t^i}$ denotes the empirical measure associated with the nonlinear (mean-field) processes $\{\bar{X}_t^i\}_{i=1}^N$. Each term in (3.26) has a distinct meaning: the first term quantifies the average discrepancy between the interacting particle trajectories and their mean-field counterparts, while the second term measures the deviation of the mean-field empirical distribution from its law $\bar{\rho}_t$.

To estimate the second term, we recall a classical result from probability theory concerning the convergence rate of empirical measures in the Wasserstein distance. The following statement, adapted from Fournier and Guillin [21, Theorem 1], provides a quantitative bound that depends only on the number of particles N , the dimension d , and the available moment bounds of the limiting distribution.

Lemma 3.10. *For $\mu \in \mathcal{P}(\mathbb{R}^d)$, consider an i.i.d. sequence $(X_k)_{k \geq 1}$ of μ -distributed random variables and, for $N \geq 1$, the empirical measure*

$$\mu_N := \frac{1}{N} \sum_{k=1}^N \delta_{X_k}.$$

Assume that $p > 1$ and $\mathfrak{m}_r(\mu) < \infty$ for some $r > p$. There exists a constant C depending only on p, d, r such that, for all $N \geq 1$,

$$\mathbb{E}\mathcal{W}_p(\mu_N, \mu) \leq C \mathfrak{m}_r^{\frac{p}{r}}(\mu) \begin{cases} N^{-\frac{1}{2}} + N^{-\frac{r-p}{r}} & \text{if } p > \frac{d}{2} \text{ and } r \neq 2p, \\ N^{-\frac{1}{2}} \log(1+N) + N^{-\frac{r-p}{r}} & \text{if } p = \frac{d}{2} \text{ and } r \neq 2p, \\ N^{-\frac{p}{d}} + N^{-\frac{r-p}{r}} & \text{if } p \in (0, \frac{d}{2}) \text{ and } r \neq \frac{d}{d-p}. \end{cases}$$

We first analyze the term $\mathcal{W}_p(\mu_t^N, \bar{\mu}_t^N)$, which compares the particle system with the mean-field system. Under the assumptions of Theorem 1.2, the coupling argument used in the proof of Theorem 1.2 yields

$$\mathbb{E}\mathcal{W}_p(\mu_t^N, \bar{\mu}_t^N) \leq \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E}|X_t^i - \bar{X}_t^i|^p \right)^{\frac{1}{p}} \leq \left(\frac{1}{N} \sum_{i=1}^N \mathbb{E}|X_0^i - \bar{X}_0^i|^p \right)^{\frac{1}{p}} + \frac{C}{N^{\frac{1}{2} \wedge \frac{q-2}{2p}}}, \quad t \geq 0,$$

where $C > 0$ is independent of t and N . This bound is both uniform in time and quantitative with respect to N , providing a precise rate of convergence for the particle system towards its mean-field approximation.

We next address the second term $\mathcal{W}_p(\bar{\mu}_t^N, \bar{\rho}_t)$, which describes the discrepancy between the empirical distribution of the nonlinear processes and the law of a single McKean–Vlasov particle. This purely probabilistic term is independent of the interaction structure and can thus be controlled by Lemma 3.10. By setting $r = pq$ and recalling that $\sup_{t \geq 0} \mathfrak{m}_{pq}(\bar{\rho}_t) < \infty$ (cf. Lemma 3.3), we deduce that

$$\mathbb{E}\mathcal{W}_p(\bar{\mu}_t^N, \bar{\rho}_t) \leq C \mathfrak{m}_{pq}^{\frac{1}{q}}(\bar{\rho}_t) \begin{cases} N^{-\frac{1}{2}} + N^{-\frac{q-1}{q}} & \text{if } p > \frac{d}{2}, \\ N^{-\frac{1}{2}} \log(1+N) + N^{-\frac{q-1}{q}} & \text{if } p = \frac{d}{2}, \\ N^{-\frac{p}{d}} + N^{-\frac{q-1}{q}} & \text{if } p \in (0, \frac{d}{2}) \text{ and } pq \neq \frac{d}{d-p}, \end{cases}$$

uniformly for all $t \geq 0$. The crucial point here is that the pq -moment of $\bar{\rho}_t$ remains bounded in time, guaranteeing that the probabilistic convergence rate does not depend on t .

Combining the two bounds above with (3.26), we obtain the quantitative mean-field limit

$$\begin{aligned} \mathbb{E}\mathcal{W}_p(\mu_t^N, \bar{\rho}_t) &\leq C \mathbb{E}\mathcal{W}_p(\mu_0^N, \bar{\rho}_0) + \frac{C}{N^{\frac{1}{2} \wedge \frac{q-2}{2p}}} \\ &\quad + C \begin{cases} N^{-\frac{1}{2}} + N^{-\frac{q-1}{q}} & \text{if } p > \frac{d}{2}, \\ N^{-\frac{1}{2}} \log(1+N) + N^{-\frac{q-1}{q}} & \text{if } p = \frac{d}{2}, \\ N^{-\frac{p}{d}} + N^{-\frac{q-1}{q}} & \text{if } p \in (0, \frac{d}{2}) \text{ and } pq \neq \frac{d}{d-p}. \end{cases} \end{aligned}$$

This completes the proof.

4. WELL-POSEDNESS AND LARGE-TIME ANALYSIS OF THE MEAN-FIELD SYSTEM

This section is devoted to the rigorous study of the mean-field equation (1.2). We first establish the existence and uniqueness of strong solutions, thereby justifying all subsequent computations. We then obtain uniform decay estimates that yield deterministic consensus formation in the large-time limit. Finally, using the structural assumptions on the objective function f , we quantify the asymptotic closeness between the consensus point and the global minimizer.

4.1. Well-posedness. The mean-field equation (1.2) differs fundamentally from the finite-particle system since its drift and diffusion coefficients depend on the evolving law $\bar{\rho}_t$. While each linearized equation with a prescribed path can be handled by classical SDE theory, the full McKean–Vlasov dynamics require closing a nonlinear self-consistency condition between the process and its own distribution. For

this reason, the well-posedness of (1.2) does not follow automatically from standard existence results and must be established separately.

The theorem below confirms that the McKean–Vlasov equation admits a unique strong solution on every finite time interval.

Theorem 4.1. *Let f satisfy Assumption 1.1, and suppose $\bar{\rho}_0 \in \mathcal{P}_2(\mathbb{R}^d)$. Then the following nonlinear stochastic differential equation*

$$d\bar{X}_t = -\lambda(\bar{X}_t - m_f^h[\bar{\rho}_t])dt + \sigma(\bar{X}_t - m_f^h[\bar{\rho}_t])dW_t, \quad \bar{\rho}_t := \text{Law}(\bar{X}_t),$$

has a unique continuous strong solution $\bar{X} \in \mathcal{C}([0, T]; \mathbb{R}^d)$ for each time horizon $T > 0$.

Remark 4.1. *According to the estimate in Lemma 3.3, the second moments of $\bar{X}_t$ remains bounded uniformly in time, and thus we can deduce that $\bar{X}_t \in \mathcal{C}([0, \infty); \mathbb{R}^d)$ by a simple continuation argument.*

Proof of Theorem 4.1. For each $v \in \mathcal{C}([0, T]; \mathbb{R}^d)$, let $Y_t = Y_t(v)$ be the unique strong solution to the linear SDE:

$$dY_t = -\lambda(Y_t - v_t)dt + \sigma(Y_t - v_t)dW_t, \quad Y_0 \sim \bar{\rho}_0 \in \mathcal{P}_2(\mathbb{R}^d).$$

We refer to [15] as a classical reference for the well-posedness of the above equation. Then, we denote $\eta_t := \text{law}(Y_t)$, and define the operator $\mathcal{T} : \mathcal{C}([0, T]; \mathbb{R}^d) \rightarrow \mathcal{C}([0, T]; \mathbb{R}^d)$ with $\mathcal{T} : v \mapsto m_f^h[\eta]$. Note that the solution to (1.2) is realized by a fixed point of $\mathcal{T}$. We will thus show in the steps to follow that $\mathcal{T}$ satisfies all assumptions of the Leray–Schauder–Schaefter fixed point theorem.

(Compactness) We denote $B_R := \{v \in \mathcal{C}([0, T]; \mathbb{R}^d) : \|v\|_\infty \leq R\}$, with $\|v\|_\infty := \sup_{t \in [0, T]} |v_t|$. Through elementary calculations, we can first check that $\sup_{t \in [0, T]} \mathbb{E}|Y_t|^2 \leq K$ for some constant $K = K(R, T, \bar{\rho}_0) > 0$. Thus, using Itô's isometry:

$$\begin{aligned} \mathbb{E}|Y_t - Y_s|^2 &\leq 2\lambda^2 \mathbb{E} \left| \int_s^t (Y_r - v_r) dr \right|^2 + 2\sigma^2 \int_s^t \mathbb{E}|Y_r - v_r|^2 dr \\ &\leq 2\lambda^2 |t - s| \int_s^t \mathbb{E}|Y_r - v_r|^2 dr + 2\sigma^2 \int_s^t \mathbb{E}|Y_r - v_r|^2 dr \\ &\leq 4(\lambda^2 T + \sigma^2)(K + \|v\|_\infty^2)|t - s|. \end{aligned}$$

By definition of the Wasserstein distance we have $\mathcal{W}_2(\eta_t, \eta_s) \leq (\mathbb{E}|Y_t - Y_s|^2)^{\frac{1}{2}}$, from which we deduce $\mathcal{W}_2(\eta_t, \eta_s) \leq C|t - s|^{\frac{1}{2}}$ for some time-independent constant $C > 0$. Utilizing Lemma 3.1, with $\mu = \eta_t$, $\nu = \eta_s$, and also using that $\sup_{[0, T]} \int_{\mathbb{R}^d} |x|^2 d\eta_t \leq C$, we have

$$\begin{aligned} |m_f^h[\eta_t] - m_f^h[\eta_s]| &\leq |m_f^h[\eta_t] - \mathcal{M}(\eta_t) - m_f^h[\eta_s] + \mathcal{M}(\eta_s)| + |\mathcal{M}(\eta_t) - \mathcal{M}(\eta_s)| \\ &\leq C\mathcal{W}_2(\eta_t, \eta_s) \\ &\leq C|t - s|^{\frac{1}{2}}, \end{aligned} \tag{4.1}$$

thanks to the elementary estimate $|\mathcal{M}(\mu) - \mathcal{M}(\nu)| \leq \mathcal{W}_1(\mu, \nu) \leq \mathcal{W}_2(\mu, \nu)$. In particular, we note that $t \mapsto m_f^h[\eta_t]$ is Hölder continuous with exponent $\frac{1}{2}$. Thus $\mathcal{T}(B_R) \subset C^{0, \frac{1}{2}}([0, T]; \mathbb{R}^d)$, and since the latter is compact in $\mathcal{C}([0, T]; \mathbb{R}^d)$ by Ascoli's theorem, we conclude that $\mathcal{T}$ is a compact operator.

(Continuity) We consider two dynamical equations with a common Wiener process W_t and initial law η_0 :

$$\begin{aligned} dY_t &= -\lambda(Y_t - v_t)dt + \sigma(Y_t - v_t)dW_t, \\ d\hat{Y}_t &= -\lambda(\hat{Y}_t - \hat{v}_t)dt + \sigma(\hat{Y}_t - \hat{v}_t)dW_t. \end{aligned}$$

Applying Itô's formula and Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \frac{d}{dt} \mathbb{E}|Y_t - \hat{Y}_t|^2 &= -2\lambda \mathbb{E}\langle Y_t - \hat{Y}_t, Y_t - \hat{Y}_t - (v_t - \hat{v}_t) \rangle + \sigma^2 \mathbb{E}|Y_t - \hat{Y}_t - (v_t - \hat{v}_t)|^2 \\ &\leq (-\lambda + 2\sigma^2) \mathbb{E}|Y_t - \hat{Y}_t|^2 + (\lambda + 2\sigma^2) |v_t - \hat{v}_t|^2. \end{aligned}$$

By Grönwall's inequality, we have

$$\sup_{t \in [0, T]} \mathbb{E}|Y_t - \hat{Y}_t|^2 \leq C \|v - \hat{v}\|_\infty^2$$

for some $C > 0$ depending only on T, λ, σ , and hence, $\sup_{t \in [0, T]} \mathcal{W}_2(\eta_t, \hat{\eta}_t) \leq C \|v - \hat{v}\|_\infty$. Now Lemma 3.1 shows, similarly as before, that

$$\sup_{t \in [0, T]} |m_f^h[\eta_t] - m_f^h[\hat{\eta}_t]| \leq C \|v - \hat{v}\|_\infty,$$

which implies $\mathcal{T}$ is continuous on $\mathcal{C}([0, T]; \mathbb{R}^d)$.

(Eigenvector bounds) We now prove that the set

$$\mathcal{A} := \{v \in \mathcal{C}([0, T], \mathbb{R}^d) : \exists \tau \in [0, 1] \text{ such that } v = \tau \mathcal{T}v\}$$

is bounded. Let $v \in \mathcal{A}$, so that by definition there exists $\rho \in \mathcal{C}([0, T]; \mathcal{P}_2(\mathbb{R}^d))$ such that $v = \tau m_f^h[\rho]$. The definition of m_f^h gives that for all $t \in (0, T)$:

$$|v_t|^2 = \tau^2 |m_f^h[\bar{\rho}_t]|^2 \leq \tau^2 \Lambda_\alpha \int_{\mathbb{R}^d} |x|^2 d\bar{\rho}_t. \quad (4.2)$$

Further, using (4.2), we can compute

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^d} |x|^2 d\bar{\rho}_t &= \int_{\mathbb{R}^d} [-2\lambda \langle x - v_t, x \rangle + \sigma^2 |x - v_t|^2] d\bar{\rho}_t \\ &\leq (-2\lambda + \sigma^2 + |\lambda - \sigma^2|) \int_{\mathbb{R}^d} |x|^2 d\bar{\rho}_t + (\sigma^2 + |\lambda - \sigma^2|) |v_t|^2 \\ &\leq C \int_{\mathbb{R}^d} |x|^2 d\bar{\rho}_t, \end{aligned}$$

where $C = -2\lambda + (1 + \tau^2 \Lambda_\alpha) \sigma^2 + (1 + \tau^2 \Lambda_\alpha) |\lambda - \sigma^2|$. From Grönwall's inequality, we deduce that

$$\int_{\mathbb{R}^d} |x|^2 d\bar{\rho}_t \leq \left(\int_{\mathbb{R}^d} |x|^2 d\bar{\rho}_0 \right) e^{Ct},$$

and therefore, $\|v\|_\infty \leq \Lambda_\alpha K e^{CT}$ which guarantees the boundedness of $\mathcal{A}$.

(Uniqueness) Up to this point, we have seen that all conditions of the Leray-Schauder-Schaefer fixed point theorem are satisfied. Therefore, the existence of a fixed point $v_t = m_f^h[\bar{\rho}_t]$, and thus a solution $\bar{X}_t$ to (1.2), is guaranteed. We now verify this solution is unique. Let $\bar{X}_t$ and $\hat{X}_t$ be two solutions to (1.2) with

same initial data $\bar{X}_0$ and driven by the same Wiener process W_t . Let $\bar{\rho}_t := \text{Law}(\bar{X}_t)$ and $\hat{\rho}_t := \text{Law}(\hat{X}_t)$. Denote the difference $Z_t = \bar{X}_t - \hat{X}_t$, then Z_t satisfies:

$$dZ_t = -\lambda(Z_t - \Delta m_t)dt + \sigma(Z_t - \Delta m_t)dW_t,$$

where $\Delta m_t := m_f^h[\bar{\rho}_t] - m_f^h[\hat{\rho}_t]$. Applying the Itô formula and Cauchy–Schwarz inequality yields:

$$\begin{aligned} \frac{d}{dt} \mathbb{E}|Z_t|^2 &= -2\lambda \mathbb{E}\langle Z_t, Z_t - \Delta m_t \rangle + \sigma^2 \mathbb{E}|Z_t - \Delta m_t|^2 \\ &\leq (-2\lambda + \sigma^2 + |\lambda - \sigma^2|) \mathbb{E}|Z_t|^2 + (\sigma^2 + |\lambda - \sigma^2|) |\Delta m_t|^2. \end{aligned}$$

Utilizing Lemma 3.1, in the same manner as in (4.1), we have $|\Delta m_t|^2 \leq C\mathcal{W}_2(\bar{\rho}_t, \hat{\rho}_t)^2 \leq C\mathbb{E}|Z_t|^2$. Substituting this into the differential inequality yields:

$$\frac{d}{dt} \mathbb{E}|Z_t|^2 \leq C\mathbb{E}|Z_t|^2$$

for some time independent constant $C > 0$. Applying Grönwall's inequality then gives $\mathbb{E}|Z_t|^2 = 0$ for all $t \in [0, T]$. $\square$

4.2. Consensus formation and convergence toward minimizers. We now turn to the long-time behavior of the mean-field dynamics (1.2), providing the proof of Theorem 1.3. In contrast with the finite-particle system, the evolution of $\bar{X}_t$ is entirely determined by its own distribution, and the resulting limit point turns out to be deterministic. This feature plays an important role: the self-consistent structure of the McKean–Vlasov equation allows one to derive sharper decay estimates than those available for the particle trajectories, where the interaction with all other particles creates additional fluctuations.

Our first objective is to show that the solution $\bar{X}_t$ indeed collapses to a single point in the large-time limit.

Proof of Theorem 1.3 (i). We first recall that $\mathcal{M}(\bar{\rho}_t) = \mathbb{E}[\bar{X}_t]$ satisfies (by taking expectations in (1.2))

$$\frac{d}{dt} \mathcal{M}(\bar{\rho}_t) = -\lambda \mathbb{E}[\bar{X}_t - m_f^h[\bar{\rho}_t]].$$

Hence, by (3.4) and Lemma 3.2, we find

$$\begin{aligned} \int_0^\infty \left| \frac{d}{dt} \mathcal{M}(\bar{\rho}_t) \right| dt &\leq C \int_0^\infty \mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]| dt \\ &\leq C \int_0^\infty (\mathbb{E}|\bar{X}_t - m_f^h[\bar{\rho}_t]|^p)^{\frac{1}{p}} dt \\ &\leq C \int_0^\infty (\mathbb{E}|\bar{X}_t - \mathbb{E}\bar{X}_t|^p)^{\frac{1}{p}} dt \\ &\leq C \int_0^\infty e^{-(\lambda - \bar{\lambda}_{p,\alpha,\sigma})t} dt < \infty. \end{aligned}$$

Therefore, $\frac{d\mathcal{M}(\bar{\rho}_t)}{dt} \in L^1(0, \infty)$, and this provides the existence of

$$x_\infty := \mathcal{M}(\bar{\rho}_0) + \int_0^\infty \frac{d}{dt} \mathcal{M}(\bar{\rho}_t) dt = \lim_{t \rightarrow \infty} \mathcal{M}(\bar{\rho}_t).$$

Using Jensen's inequality, we then write

$$\mathbb{E}|\bar{X}_t - x_\infty|^p \leq 2^{p-1}\mathbb{E}|\bar{X}_t - \mathcal{M}(\bar{\rho}_t)|^p + 2^{p-1}|\mathcal{M}(\bar{\rho}_t) - x_\infty|^p.$$

As $t \rightarrow \infty$, the first term on the right-hand side tends to 0 by applying Lemma 3.2, and the second term tends to zero by definition of x_∞ . $\square$

Theorem 1.3 (i) shows that the mean-field dynamics converge to a deterministic point x_∞^α .

The next result addresses the central question of optimization: whether this consensus point is close to a global minimizer of the objective function f in the regime of large α .

Proof of Theorem 1.3 (ii). We follow the strategy used for the particle system in Lemma 2.7 and Theorem 1.1 (ii). Applying Itô's formula to $\omega_f^\alpha(\bar{X}_t)$ using the McKean-Vlasov SDE (1.2), and then taking expectations, yields:

$$\begin{aligned} \frac{d}{dt} \|\omega_f^\alpha\|_{L^1(\bar{\rho}_t)} &= -\lambda \int_{\mathbb{R}^d} \alpha e^{-\alpha f(x)} \nabla f(x) \cdot (x - m_f^h[\bar{\rho}_t]) d\bar{\rho}_t \\ &\quad + \frac{1}{2} \alpha \sigma^2 \int_{\mathbb{R}^d} (x - m_f^h[\bar{\rho}_t])^\top (\alpha \nabla f(x) \nabla f(x)^\top - \nabla^2 f(x)) (x - m_f^h[\bar{\rho}_t]) d\bar{\rho}_t. \end{aligned}$$

Using Assumptions 1.1 and 1.2, for $\alpha \geq c_1$, we have

$$\begin{aligned} \frac{d}{dt} \|\omega_f^\alpha\|_{L^1(\bar{\rho}_t)} &\geq -\lambda L_{\omega_f^\alpha} \int_{\mathbb{R}^d} |x - m_f^h[\bar{\rho}_t]| d\bar{\rho}_t - \frac{\alpha \sigma^2 c_0 e^{-\alpha f}}{2} \int_{\mathbb{R}^d} |x - m_f^h[\bar{\rho}_t]|^2 d\bar{\rho}_t \\ &\geq -\lambda L_{\omega_f^\alpha} \Lambda_\alpha \left(\int_{\mathbb{R}^d} |x - \mathcal{M}(\bar{\rho}_t)|^2 d\bar{\rho}_t \right)^{\frac{1}{2}} - \alpha \sigma^2 c_0 e^{-\alpha f} \Lambda_\alpha \int_{\mathbb{R}^d} |x - \mathcal{M}(\bar{\rho}_t)|^2 d\bar{\rho}_t, \end{aligned} \quad (4.3)$$

where we used the fact that

$$\int_{\mathbb{R}^d} |x - m_f^h[\bar{\rho}_t]|^2 d\bar{\rho}_t \leq \frac{\int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \psi_h(y) d\bar{\rho}_t d\bar{\rho}_t}{\int_{\mathbb{R}^d} \psi_h(y) d\bar{\rho}_t} \leq 2\Lambda_\alpha \int_{\mathbb{R}^d} |x - \mathcal{M}(\bar{\rho}_t)|^2 d\bar{\rho}_t.$$

Therefore, integrating both sides of (4.3) then applying Lemma 3.2 yields

$$\begin{aligned} \|\omega_f^\alpha\|_{L^1(\bar{\rho}_t)} &\geq \|\omega_f^\alpha\|_{L^1(\bar{\rho}_0)} - \frac{\lambda L_{\omega_f^\alpha} \Lambda_\alpha}{\lambda - \bar{\Lambda}_{2,\alpha,\sigma}} \text{Var}(\bar{X}_0)^{\frac{1}{2}} (1 - e^{-(\lambda - \bar{\Lambda}_{2,\alpha,\sigma})t}) \\ &\quad - \frac{\alpha \sigma^2 c_0 e^{-\alpha f} \Lambda_\alpha}{2(\lambda - \bar{\Lambda}_{2,\alpha,\sigma})} \text{Var}(\bar{X}_0) (1 - e^{-2(\lambda - \bar{\Lambda}_{2,\alpha,\sigma})t}). \end{aligned}$$

Letting $t \rightarrow \infty$ and applying the dominated convergence theorem yields

$$\begin{aligned} e^{-\alpha f(x_\infty)} &\geq \|\omega_f^\alpha\|_{L^1(\bar{\rho}_0)} - \frac{\lambda L_{\omega_f^\alpha} \Lambda_\alpha}{\lambda - \bar{\Lambda}_{2,\alpha,\sigma}} \text{Var}(\bar{X}_0)^{\frac{1}{2}} - \frac{\alpha \sigma^2 c_0 e^{-\alpha f} \Lambda_\alpha}{2(\lambda - \bar{\Lambda}_{2,\alpha,\sigma})} \text{Var}(\bar{X}_0) \\ &\geq \epsilon \|\omega_f^\alpha\|_{L^1(\bar{\rho}_0)}. \end{aligned}$$

Therefore, taking logarithms on both sides, we have

$$f(x_\infty) \leq -\frac{1}{\alpha} \log \|\omega_f^\alpha\|_{L^1(\bar{\rho}_0)} - \frac{1}{\alpha} \log \epsilon.$$

By Laplace's principle, we conclude that

$$f(x_\infty) \leq \inf_{x \in \text{supp}(\bar{\rho}_0)} f(x) + o(1) \quad (\alpha \rightarrow \infty).$$

This completes the proof. $\square$

ACKNOWLEDGMENTS

This work is supported by NRF grant no. RS-2024-00406821.

APPENDIX A. EXCHANGABILITY OF THE SOLUTIONS

In this appendix, we verify that the finite-particle system (1.1) preserves exchangeability over time.

Lemma A.1. *Let $\{X_t^i\}_{i=1}^N$ be the unique strong solution to the particle system (1.1) with i.i.d. initial data $\{X_0^i\}_{i=1}^N$ and independent Wiener processes $\{W_t^i\}_{i=1}^N$. Then the family $\{X_t^i\}_{i=1}^N$ is exchangeable for all $t \geq 0$.*

Proof. Let $\pi \in \text{Perm}(N)$ be any permutation of $\{1, 2, \dots, N\}$. We define a new process $\{\hat{\mathbf{X}}_t\}_{t \geq 0}$ by permuting the components of $\mathbf{X}_t$, that is,

$$\hat{X}_t^i := X_t^{\pi(i)}, \quad i = 1, 2, \dots, N.$$

First, we rewrite the dynamics (1.1) for each particle using the empirical measure

$$\mu_{\mathbf{X}_t^N} = \frac{1}{N} \sum_{j=1}^N \delta_{X_t^j}$$

as

$$dX_t^i = F(X_t^i, \mu_{\mathbf{X}_t^N}) dt + G(X_t^i, \mu_{\mathbf{X}_t^N}) dW_t^i,$$

where the drift and diffusion coefficients F and G are given, for any $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d)$, by

$$F(x, \mu) = -\lambda(x - m_f^h[\mu]), \quad G(x, \mu) = \sigma(x - m_f^h[\mu]).$$

Since the empirical measure is invariant under permutation, we have

$$\mu_{\hat{\mathbf{X}}_t^N} = \frac{1}{N} \sum_{i=1}^N \delta_{\hat{X}_t^i} = \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{\pi(i)}} = \mu_{\mathbf{X}_t^N}.$$

This invariance implies that each component $\hat{X}_t^i = X_t^{\pi(i)}$ satisfies the dynamics

$$d\hat{X}_t^i = dX_t^{\pi(i)} = F(X_t^{\pi(i)}, \mu_{\mathbf{X}_t^N}) dt + G(X_t^{\pi(i)}, \mu_{\mathbf{X}_t^N}) dW_t^{\pi(i)} = F(\hat{X}_t^i, \mu_{\hat{\mathbf{X}}_t^N}) dt + G(\hat{X}_t^i, \mu_{\hat{\mathbf{X}}_t^N}) d\hat{W}_t^i,$$

where we introduced the permuted Wiener processes $\hat{W}_t^i := W_t^{\pi(i)}$.

Next, we check the initial and noise distributions:

- (i) Since the initial data $\{X_0^i\}_{i=1}^N$ are i.i.d., the permuted vector $\hat{\mathbf{X}}_0 = (X_0^{\pi(1)}, \dots, X_0^{\pi(N)})^\top$ is identically distributed to $\mathbf{X}_0$.
- (ii) Since $\{W_t^i\}_{i=1}^N$ are independent one-dimensional Wiener processes, the permuted family $\{\hat{W}_t^i\}_{i=1}^N$ also consists of independent one-dimensional Wiener processes. Hence, the vector processes $\{\hat{\mathbf{W}}_t\}_{t \geq 0}$ and $\{\mathbf{W}_t\}_{t \geq 0}$ have the same law.

Therefore, $\{\mathbf{X}_t\}_{t \geq 0}$ and $\{\hat{\mathbf{X}}_t\}_{t \geq 0}$ satisfy the same SDE, driven by initial data and Wiener processes that are equal in law. By pathwise uniqueness of strong solutions (Theorem 2.1), the two processes coincide in law:

$$\text{Law}(\{\hat{\mathbf{X}}_t\}_{t \geq 0}) = \text{Law}(\{\mathbf{X}_t\}_{t \geq 0}).$$

Since this holds for any permutation $\pi \in \text{Perm}(N)$, we conclude that the family $\{X_t^i\}_{i=1}^N$ is exchangeable. $\square$

APPENDIX B. WELL-POSEDNESS OF THE MODIFIED CBO SYSTEM

In this appendix, we provide a detailed proof of Theorem 2.1. Throughout this section, we fix $N \in \mathbb{N}$ and denote

$$\mathbf{X}_t := (X_t^1, X_t^2, \dots, X_t^N)^\top \in \mathbb{R}^{Nd}.$$

The system (1.1) can then be rewritten in the vector form

$$d\mathbf{X}_t = -\lambda \mathbf{F}_N(\mathbf{X}_t) dt + \sigma \mathbf{M}_N(\mathbf{X}_t) d\mathbf{W}_t,$$

where $\mathbf{W}_t = (W_t^1 \mathbf{1}_d, W_t^2 \mathbf{1}_d, \dots, W_t^N \mathbf{1}_d)^\top \in \mathbb{R}^{Nd}$, with $\{W^i\}_{i=1}^N$ being a collection of independent one-dimensional Wiener processes and $\mathbf{1}_d = (1, 1, \dots, 1) \in \mathbb{R}^d$. The drift term

$$\mathbf{F}_N(\mathbf{X}) = (F_N^1(\mathbf{X}), F_N^2(\mathbf{X}), \dots, F_N^N(\mathbf{X}))^\top \in \mathbb{R}^{Nd}$$

has components

$$F_N^i(\mathbf{X}) = \frac{\sum_{j=1}^N (X^i - X^j) \psi_h(X^j)}{\sum_{j=1}^N \psi_h(X^j)} \in \mathbb{R}^d.$$

The diffusion matrix $\mathbf{M}_N(\mathbf{X})$ is block-diagonal:

$$\mathbf{M}_N(\mathbf{X}) = \begin{pmatrix} \mathbf{M}_N^1(\mathbf{X}) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_N^2(\mathbf{X}) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{M}_N^3(\mathbf{X}) & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{M}_N^N(\mathbf{X}) \end{pmatrix},$$

where each $\mathbf{M}_N^i(\mathbf{X}) \in \mathbb{R}^{d \times d}$ has entries

$$(\mathbf{M}_N^i(\mathbf{X}))_{nm} = \begin{cases} (F_N^i)_n & n = m, \\ 0 & n \neq m. \end{cases}$$

Under the local Lipschitz regularity assumption on f and the positivity of h , one verifies that each F_N^i is locally Lipschitz and has at most linear growth, hence $\mathbf{F}_N$ and $\mathbf{M}_N$ inherit these properties.

Lemma B.1. *Let $N \in \mathbb{N}$, $k, \alpha > 0$ be arbitrary and assume h is strictly positive function. Then for any $\mathbf{X}, \tilde{\mathbf{X}} \in \mathbb{R}^{Nd}$ and all $i = 1, 2, \dots, N$ with $|\mathbf{X}|, |\tilde{\mathbf{X}}| \leq k$ it holds,*

$$\begin{aligned} |F_N^i(\mathbf{X}) - F_N^i(\tilde{\mathbf{X}})| &\leq |X^i - \tilde{X}^i| + \left(\frac{\Lambda_\alpha}{\sqrt{N}} + \frac{c_k(1 + \Lambda_\alpha)}{Nh(\alpha)} \sqrt{N|\tilde{X}^i| + |\tilde{\mathbf{X}}|^2} \right) |\mathbf{X} - \tilde{\mathbf{X}}|, \\ |F_N^i(\mathbf{X})| &\leq |X^i| + \frac{\Lambda_\alpha}{\sqrt{N}} |\mathbf{X}|, \end{aligned}$$

where $c_k = \sqrt{2}\alpha \|\nabla f \cdot \omega_f^\alpha\|_{L^\infty(B(0,k))}$.

Proof. Let $\mathbf{X}, \tilde{\mathbf{X}} \in \mathbb{R}^{Nd}$ with $|\mathbf{X}|, |\tilde{\mathbf{X}}| \leq k$ for some $k \geq 0$ and $i \in \{1, 2, \dots, N\}$ be arbitrary. We can easily notice that

$$F_N^i(\mathbf{X}) - F_N^i(\tilde{\mathbf{X}}) = \frac{\sum_{j \neq i} (X^i - X^j) \psi_h(X^j)}{\sum_{j=1}^N \psi_h(X^j)} - \frac{\sum_{j \neq i} (\tilde{X}^i - \tilde{X}^j) \psi_h(\tilde{X}^j)}{\sum_{j=1}^N \psi_h(\tilde{X}^j)} =: I^i + II^i + III^i,$$

where

$$\begin{aligned} |I^i| &= \left| \frac{\sum_{j \neq i} (X^i - \tilde{X}^i + \tilde{X}^j - X^j) \psi_h(X^j)}{\sum_{j=1}^N \psi_h(X^j)} \right| \leq |X^i - \tilde{X}^i| + \frac{\Lambda_\alpha}{\sqrt{N}} |\mathbf{X} - \tilde{\mathbf{X}}|, \\ |II^i| &= \left| \frac{\sum_{j \neq i} (\tilde{X}^i - \tilde{X}^j) (\psi_h(X^j) - \psi_h(\tilde{X}^j))}{\sum_{j=1}^N \psi_h(X^j)} \right| \leq \frac{\sqrt{2}\alpha \|\nabla f \cdot \omega_f^\alpha\|_{L^\infty(B(0,k))}}{Nh(\alpha)} \sqrt{N|\tilde{X}^i|^2 + |\tilde{\mathbf{X}}|^2} |\mathbf{X} - \tilde{\mathbf{X}}|, \\ |III^i| &= \left| \sum_{j \neq i} (\tilde{X}^i - \tilde{X}^j) \psi_h(\tilde{X}^j) \frac{\sum_{j=1}^N (\psi_h(\tilde{X}^j) - \psi_h(X^j))}{\sum_{j=1}^N \psi_h(X^j) \sum_{j=1}^N \psi_h(\tilde{X}^j)} \right| \\ &\leq \frac{\sqrt{2}\alpha \Lambda_\alpha \|\nabla f \cdot \omega_f^\alpha\|_{L^\infty(B(0,k))}}{Nh(\alpha)} \sqrt{N|\tilde{X}^i|^2 + |\tilde{\mathbf{X}}|^2} |\mathbf{X} - \tilde{\mathbf{X}}|. \end{aligned}$$

Summing all these together yields the required estimate. As for the estimate of $|F_N^i(\mathbf{X})|$, we easily obtain,

$$|F_N^i(\mathbf{X})| \leq |X^i| + \frac{\Lambda_\alpha}{\sqrt{N}} |\mathbf{X}|. \quad (\text{B.1})$$

This completes the proof. $\square$

Lemma B.1 implies that both $\mathbf{F}_N$ and $\mathbf{M}_N$ satisfy the local Lipschitz and linear growth conditions. We can therefore invoke standard SDE theory to establish global existence and uniqueness, as stated in Theorem 2.1.

Proof of Theorem 2.1. Applying Itô's formula to $|\mathbf{X}_t|^2$, we obtain

$$d|\mathbf{X}_t|^2 = -2\lambda \langle \mathbf{X}_t, \mathbf{F}_N(\mathbf{X}_t) \rangle dt + 2\sigma \langle \mathbf{X}_t, \mathbf{M}_N(\mathbf{X}_t) d\mathbf{W}_t \rangle + \sigma^2 \|\mathbf{M}_N(\mathbf{X}_t)\|_{\mathbb{F}}^2 dt.$$

Using the estimate

$$-\langle X_t^i, F_N^i(\mathbf{X}_t) \rangle = - \left\langle X_t^i, \frac{\sum_{j=1}^N (X_t^i - X_t^j) \psi_h(X_t^j)}{\sum_{j=1}^N \psi_h(X_t^j)} \right\rangle \leq -|X_t^i|^2 + \frac{\Lambda_\alpha}{\sqrt{N}} |X_t^i| |\mathbf{X}_t|,$$

together with (B.1) and Hölder's inequality, we deduce

$$\begin{aligned} -2\lambda \langle \mathbf{X}_t, \mathbf{F}_N(\mathbf{X}_t) \rangle + \sigma^2 \|\mathbf{M}_N(\mathbf{X}_t)\|_{\mathbb{F}}^2 &= \sum_{i=1}^N [-2\lambda \langle X_t^i, F_N^i(\mathbf{X}_t) \rangle + \sigma^2 |F_N^i(\mathbf{X}_t)|^2] \\ &\leq -2\lambda \sum_{i=1}^N \left[|X_t^i|^2 - \frac{\Lambda_\alpha}{\sqrt{N}} |X_t^i| |\mathbf{X}_t| \right] + 2\sigma^2 (1 + \Lambda_\alpha) |\mathbf{X}_t|^2 \\ &\leq (-2\lambda + 2\lambda \Lambda_\alpha + 2\sigma^2 (1 + \Lambda_\alpha)) |\mathbf{X}_t|^2. \end{aligned} \quad (\text{B.2})$$

The local Lipschitz continuity and linear growth of $\mathbf{F}_N$ and $\mathbf{M}_N$, established in Lemma B.1, allow us to apply the standard theory of stochastic differential equations (see, e.g. [15, Theorem 3.1]), which guarantees the existence and uniqueness of a local-in-time strong solution $\mathbf{X}_t$ to (1.1).

To show that the solution is in fact global, we derive a uniform second-moment bound. It follows from (B.2) that

$$\frac{d}{dt} \mathbb{E}|\mathbf{X}_t|^2 \leq b_\alpha \mathbb{E}|\mathbf{X}_t|^2, \quad b_\alpha := -2\lambda + 2\lambda\Lambda_\alpha + 2\sigma^2(1 + \Lambda_\alpha).$$

Applying Grönwall's inequality then yields:

$$\mathbb{E}|\mathbf{X}_t|^2 \leq \mathbb{E}|\mathbf{X}_0|^2 e^{b_\alpha t} \quad \text{for all } t \geq 0.$$

This bound ensures that $\mathbb{E}|\mathbf{X}_t|^2$ remains finite for every finite time t , so no explosion can occur before $t = \infty$. Hence, the local strong solution extends globally in time.

Finally, note that the estimate above does not depend on the number of particles N . Therefore, the family of solutions $\{\mathbf{X}_t\}_{N \in \mathbb{N}}$ enjoys a uniform second-moment bound, which implies a strong stability of the system with respect to N . $\square$

APPENDIX C. UNIFORM-IN-TIME STABILITY OF PARTICLE SYSTEM

In this appendix, we provide the statements and proofs of the concentration and stability estimates for the particle system (1.1).

The first result below is a concentration estimate, which can be viewed as an adaptation of [22, Lemma 4.9] to our setting. It also extends that result in the sense that we work under weaker assumptions on the objective function. For the reader's convenience, we restate the concentration estimate here and include a sketch of the proof.

Lemma C.1 (Concentration estimate). *Let $p, q \geq 2$ and assume $X_{\text{in}} \in L^{pq}(\Omega)$. For the particle trajectories, we suppose that X_0^i are i.i.d. with $X_0^i \sim X_{\text{in}}$ for all $i = 1, \dots, N$. Denoting*

$$c_{\text{con},p} = 2\Lambda_\alpha^{1-\frac{2}{p}} \lambda_{p,\alpha,\sigma} + 2(p-1)\sigma^2,$$

further assume

$$\lambda > \lambda'_{p,q,\alpha,\sigma} := \lambda_{pq,\alpha,\sigma} + c_{\text{con},p}.$$

Then for all small enough $\kappa > 0$ and every $A > 0$, the following concentration estimate holds:

$$\mathbb{P} \left[\sup_{t \geq 0} e^{\kappa t} V_p(\mu_t^N) \geq \mathbb{E} [V_p(\mu_0^N)] + A \right] \leq \frac{C}{A^q N^{\frac{q}{2}}} \quad (\text{C.1})$$

for some constant $C > 0$ independent of A and N . In particular, (C.1) holds for all fixed $0 < \kappa < p(\lambda - (c_{\text{con},p} \vee \lambda_{pq,\alpha,\sigma}))$.

Proof. Since the proof is similar to that of Lemma 3.5, we give a sketch of proof. We apply Itô's formula to $V_p(\mu_t^N) = \frac{1}{N} \sum_{i=1}^N |\mathcal{K}_t^i|^p$, with $\mathcal{K}_t^i := X_t^i - \mathcal{M}(\mu_t^N)$, and then use the Cauchy-Schwarz inequality to

find

$$\begin{aligned}
dV_p(\mu_t^N) &\leq -\lambda p V_p(\mu_t^N) dt + \frac{p(p-1)\sigma^2}{2N} \sum_{i=1}^N |\mathcal{K}_t^i|^{p-2} |X_t^i - m_f^h[\mu_t^N]|^2 dt \\
&\quad + \frac{p(p-1)\sigma^2}{2N^3} \sum_{i=1}^N \sum_{j=1}^N |\mathcal{K}_t^i|^{p-2} |X_t^j - m_f^h[\mu_t^N]|^2 dt \\
&\quad + \frac{p\sigma}{N} \sum_{i=1}^N \left[|\mathcal{K}_t^i|^{p-2} \langle \mathcal{K}_t^i, X_t^i - m_f^h[\mu_t^N] \rangle - \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N |\mathcal{K}_t^j|^{p-2} \langle \mathcal{K}_t^j, X_t^i - m_f^h[\mu_t^N] \rangle \right] dW_t^i.
\end{aligned}$$

Next, by Jensen's inequality, for any $m \geq 1$ we have

$$|X_t^i - m_f^h[\mu_t^N]|^m \leq 2^{m-1} |\mathcal{K}_t^i|^m + \frac{2^{m-1} \Lambda_\alpha}{N} \sum_{i=1}^N |\mathcal{K}_t^i|^m. \quad (\text{C.2})$$

Applying (C.2) together with Hölder's inequality yields the differential inequality

$$\begin{aligned}
dV_p(\mu_t^N) &\leq -\left(p\lambda - p(p-1)(1 + \Lambda_\alpha) \left(1 + \frac{1}{N}\right) \sigma^2\right) V_p(\mu_t^N) dt \\
&\quad + \frac{p\sigma}{N} \sum_{i=1}^N \left[|\mathcal{K}_t^i|^{p-2} \langle \mathcal{K}_t^i, X_t^i - m_f^h[\mu_t^N] \rangle - \frac{1}{N} \sum_{j=1}^N |\mathcal{K}_t^j|^{p-2} \langle \mathcal{K}_t^j, X_t^i - m_f^h[\mu_t^N] \rangle \right] dW_t^i \\
&\leq -p(\lambda - c_{\text{con},p}) V_p(\mu_t^N) dt + dM_t.
\end{aligned}$$

Since $\lambda - c_{\text{con},p} > \lambda_{pq,\alpha,\sigma}$, we may choose $\kappa \in (0, p(\lambda - c_{\text{con},p}))$. Then, we define the auxiliary process $\mathcal{E}_t = e^{\kappa t} V_p(\mu_t^N)$. Applying Markov's inequality, along with Marcinkiewicz–Zygmund inequality and the Burkholder–Davis–Gundy inequality, we obtain in the same way as in the proof of Lemma 3.5:

$$\mathbb{P} \left[\sup_{t \geq 0} \mathcal{E}_t \geq \mathbb{E} \mathcal{E}_0 + A \right] \leq \frac{C}{A^q N^{\frac{q}{2}}} \left(1 + \frac{1}{(p(\lambda - \lambda_{pq,\alpha,\sigma}) - \kappa)^{\frac{q}{2}}} \right),$$

where $C > 0$ is dependent on $p, q, \mathbb{E}|X_{\text{in}}|^{pq}, \Lambda_\alpha$, and σ but independent of A or N . We remark that the above estimate is valid for

$$\kappa < p(\lambda - \lambda_{pq,\alpha,\sigma}). \quad (\text{C.3})$$

Since

$$\lambda - \lambda_{pq,\alpha,\sigma} > c_{\text{con},p} \geq \Lambda_\alpha^{1-\frac{2}{p}} \lambda_{p,\alpha,\sigma} + (p-1)\sigma^2,$$

we can choose $\kappa > 0$ small enough so that (C.3) holds. We conclude the proof by proceeding in the same way as at the end of the proof of Lemma 3.5. $\square$

We now proceed to prove the stability estimates for $p \geq 2$ in Theorem C.2. The proof relies on Lemma 3.1, which allows us to decompose the difference between the consensus points and the empirical means of the particles. This decomposition introduces terms involving the empirical p -th variances, which are subsequently controlled through the concentration estimate obtained in Lemma C.1. The latter ensures that these variance terms remain uniformly bounded with high probability.

Theorem C.2. Let $p \geq 2$, $q > 2$, and assume that $\lambda > \lambda'_{p,q,\alpha,\sigma}$. Suppose that the objective function f satisfies Assumptions 1.1. Consider two copies $\{X_t^i\}_{i=1}^N$ and $\{Y_t^i\}_{i=1}^N$ of the particle system (1.1), driven by the same one-dimensional Wiener processes $\{W_t^i\}_{i=1}^N$, but with different i.i.d. initial data: for each $1 \leq i \leq N$, assume $X_0^i \sim X_{\text{in}}$ and $Y_0^i \sim Y_{\text{in}}$, where $X_{\text{in}}, Y_{\text{in}} \in L^{pq}(\Omega)$. Then the following stability estimates hold:

(1) if $p = 2$,

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E}|X_t^i - Y_t^i|^2 \leq \frac{C}{N} \sum_{i=1}^N \mathbb{E}|X_0^i - Y_0^i|^2 + \frac{C}{N^{\frac{q-2}{2}}},$$

(2) if $p > 2$,

$$\frac{1}{N} \sum_{i=1}^N \mathbb{E}|X_t^i - Y_t^i|^p \leq \frac{C}{N} \sum_{i=1}^N \mathbb{E}|X_0^i - Y_0^i|^p + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}}$$

holds, where $C > 0$ is a constant depending only on $\lambda, p, q, \sigma, \Lambda_\alpha, X_{\text{in}}$, and Y_{in} .

Proof. We adopt the same notations as in the proof of Theorem C.2, replacing $\bar{X}_t^i$ there by Y_t^i :

$$Z_t^i = X_t^i - Y_t^i, \quad \nu_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{Y_t^i}, \quad \Delta_t^m = m_f^h[\mu_t^N] - m_f^h[\nu_t^N], \quad \Delta_t^{\mathcal{M}} = \mathcal{M}(\mu_t^N) - \mathcal{M}(\nu_t^N).$$

We also set

$$\mathcal{G}_{t,p} := \frac{1}{N} \sum_{i=1}^N |Z_t^i|^p, \quad \mathcal{D}_{t,p} := \frac{1}{N} \sum_{i=1}^N |Z_t^i - \Delta_t^{\mathcal{M}}|^p, \quad \mathcal{O}_{t,p} := |\Delta_t^{\mathcal{M}}|^p.$$

The overall strategy is similar to that of Theorem 1.2. Our goal is to control $\mathcal{G}_{t,p}$ by estimating $\mathcal{D}_{t,p}$ and $\mathcal{O}_{t,p}$. The first crucial step is to derive a bound for $\mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^p$. From Lemma 3.1, we recall

$$|\Delta_t^{\mathcal{M}} - \Delta_t^m|^p \leq \frac{1}{2} \left(\frac{2L(\alpha)C_\alpha}{h(\alpha)} \right)^p (V_p(\mu_t^N) + V_p(\nu_t^N)) \mathcal{W}_p^p(\mu_t^N, \nu_t^N).$$

Using the concentration estimate from Lemma C.1, we can proceed as in Lemma 3.7 and deduce that for some $0 < \eta < \kappa$,

$$\mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^p \leq \left[\frac{C}{N^{\frac{q-2}{2}}} + C\mathbb{E}\mathcal{G}_{t,p} \right] e^{-\eta t}, \quad (\text{C.4})$$

where $C > 0$ is a constant depending only on $\lambda, \sigma, p, q, \Lambda_\alpha, X_{\text{in}}$, and Y_{in} .

We now consider the two cases $p = 2$ and $p > 2$ separately.

• (The case $p = 2$) Applying the Itô formula to $\mathcal{G}_{t,2}$ leads to

$$d\mathcal{G}_{t,2} = - \left[\frac{2\lambda}{N} \sum_{i=1}^N |Z_t^i|^2 - 2\lambda \langle \Delta_t^{\mathcal{M}}, \Delta_t^m \rangle - \frac{\sigma^2}{N} \sum_{i=1}^N |Z_t^i - \Delta_t^m|^2 \right] dt + \frac{2\sigma}{N} \sum_{i=1}^N \langle Z_t^i, Z_t^i - \Delta_t^m \rangle dW_t^i.$$

By using the elementary identity

$$\frac{1}{N} \sum_{i=1}^N |z_i|^2 = \frac{1}{N} \sum_{i=1}^N |z_i - z^{\text{avg}}|^2 + |z^{\text{avg}}|^2, \quad z^{\text{avg}} := \frac{1}{N} \sum_{i=1}^N z_i,$$

and taking the expectation of both sides, we obtain for some $0 < \tilde{\eta} < \frac{\eta}{2}$

$$\frac{d}{dt} \mathbb{E}\mathcal{G}_{t,2} = - \frac{2\lambda - \sigma^2}{N} \sum_{i=1}^N \mathbb{E}|Z_t^i - \Delta_t^{\mathcal{M}}|^2 - 2\lambda \mathbb{E} \langle \Delta_t^{\mathcal{M}}, \Delta_t^{\mathcal{M}} - \Delta_t^m \rangle + \sigma^2 \mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^2$$

$$\begin{aligned}
&\leq 2\lambda (\mathbb{E}\mathcal{G}_{t,2})^{\frac{1}{2}} (\mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^2)^{\frac{1}{2}} + \sigma^2 \mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^2 \\
&\leq \lambda \mathbb{E}\mathcal{G}_{t,2} e^{-\tilde{\eta}t} + (\lambda e^{\tilde{\eta}t} + \sigma^2) \mathbb{E}|\Delta_t^{\mathcal{M}} - \Delta_t^m|^2.
\end{aligned}$$

The first term on the right-hand side is negative by the assumption on λ , while the remaining terms can be controlled by Cauchy–Schwarz and Young’s inequalities. Substituting the estimate (C.4) into the above yields

$$\frac{d}{dt} \mathbb{E}\mathcal{G}_{t,2} \leq C \mathbb{E}\mathcal{G}_{t,2} e^{-\tilde{\eta}t} + \frac{C}{N^{\frac{q-2}{2}}} e^{-\tilde{\eta}t}.$$

Applying Grönwall’s inequality yields the desired estimate for the $p = 2$ case:

$$\mathbb{E}\mathcal{G}_{t,2} \leq \left(\mathbb{E}\mathcal{G}_{0,2} + \frac{C}{N^{\frac{q-2}{2}}} \right) e^{C/\tilde{\eta}}.$$

• (The case $p > 2$) In the same way as in (3.19), we have

$$\mathcal{G}_{t,p} \leq 2^{p-1} (\mathcal{D}_{t,p} + \mathcal{O}_{t,p}) \leq (2^{2p-1} + 2^{p-1}) \mathcal{G}_{t,p}. \quad (\text{C.5})$$

Applying Itô’s formula to $\mathcal{D}_{t,p}$ and using Cauchy–Schwarz gives

$$\begin{aligned}
\frac{d}{dt} \mathbb{E}\mathcal{D}_{t,p} &\leq -p(\lambda - (p-1)\sigma^2) \mathbb{E}\mathcal{D}_{t,p} + \frac{p(p-1)\sigma^2}{N} \sum_{i=1}^N \mathbb{E}[|Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |\Delta_t^{\mathcal{M}} - \Delta_t^m|^2] \\
&\quad + \frac{p(p-1)\sigma^2}{2N^3} \sum_{i,j=1}^N \mathbb{E}[|Z_t^i - \Delta_t^{\mathcal{M}}|^{p-2} |Z_t^j - \Delta_t^m|^2].
\end{aligned}$$

Since the leading term is strictly negative, the remaining ones are handled as in Lemma 3.8, using Lemmas 2.2 and C.1, together with Young’s inequality and (C.4). Thus, there exist a constant $C > 0$, depending only on $\lambda, p, q, \sigma, \Lambda_\alpha, X_{\text{in}}$ and Y_{in} , and also $\hat{\eta} > 0$ such that

$$\frac{d}{dt} \mathbb{E}\mathcal{D}_{t,p} \leq C \mathbb{E}\mathcal{D}_{t,p} e^{-\hat{\eta}t} + C \mathbb{E}\mathcal{O}_{t,p} e^{-\hat{\eta}t} + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} e^{-\hat{\eta}t}. \quad (\text{C.6})$$

Next, applying Itô’s formula to $\mathcal{O}_{t,p}$ and applying the Cauchy–Schwarz inequality shows

$$\frac{d}{dt} \mathbb{E}\mathcal{O}_{t,p} \leq -p\lambda \mathbb{E}|\Delta_t^{\mathcal{M}}|^{p-2} \langle \Delta_t^{\mathcal{M}}, \Delta_t^{\mathcal{M}} - \Delta_t^m \rangle + \frac{p(p-1)\sigma^2}{2N^2} \mathbb{E} \left[|\Delta_t^{\mathcal{M}}|^{p-2} \sum_{i=1}^N |Z_t^i - \Delta_t^m|^2 \right].$$

Then, by using a similar argument as in Lemma 3.9, we obtain

$$\frac{d}{dt} \mathbb{E}\mathcal{O}_{t,p} \leq C \mathbb{E}\mathcal{O}_{t,p} e^{-\hat{\eta}t} + C \mathbb{E}\mathcal{D}_{t,p} e^{-\hat{\eta}t} + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} e^{-\hat{\eta}t} \quad (\text{C.7})$$

for some $C > 0$, which depends only on $\lambda, p, q, \sigma, \Lambda_\alpha, X_{\text{in}}$, and Y_{in} .

Combining (C.6) and (C.7), we have

$$\frac{d}{dt} (\mathbb{E}\mathcal{D}_{t,p} + \mathbb{E}\mathcal{O}_{t,p}) \leq C (\mathbb{E}\mathcal{D}_{t,p} + \mathbb{E}\mathcal{O}_{t,p}) e^{-\hat{\eta}t} + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} e^{-\hat{\eta}t}$$

for some positive constant C , depending only on $\lambda, p, q, \sigma, \Lambda_\alpha, X_{\text{in}}$ and Y_{in} . Applying Grönwall’s inequality gives

$$\mathbb{E}\mathcal{D}_{t,p} + \mathbb{E}\mathcal{O}_{t,p} \leq \left[\mathbb{E}\mathcal{D}_{0,p} + \mathbb{E}\mathcal{O}_{0,p} + \frac{C}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} \right] e^{C/\hat{\eta}}.$$

Finally, using (C.5), we deduce

$$\mathbb{E}\mathcal{G}_{t,p} \leq C \left[\mathbb{E}\mathcal{G}_{0,p} + \frac{1}{N^{\frac{p}{2} \wedge \frac{q-2}{2}}} \right] e^{C/\hat{\eta}},$$

completing the proof of Theorem C.2. $\square$

REFERENCES

- [1] H.-O. Bae, S.-Y. Ha, M. Kang, H. Lim, C. Min, and J. Yoo. A constrained consensus based optimization algorithm and its application to finance. *Appl. Math. Comput.*, 416:Paper No. 126726, 10, 2022.
- [2] E. Bayraktar, I. Ekren, and H. Zhou. Uniform-in-time weak propagation of chaos for consensus-based optimization, arXiv:2502.00582.
- [3] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar. *Convex analysis and optimization*. Athena Scientific, Belmont, MA, 2003.
- [4] L. Bianchi, M. Dorigo, L. M. Gambardella, and W. J. Gutjahr. A survey on metaheuristics for stochastic combinatorial optimization. *Nat. Comput.*, 8(2):239–287, 2009.
- [5] E. Bonabeau, M. Dorigo, and G. Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, 10 1999.
- [6] G. Borghi, M. Herty, and L. Pareschi. Constrained consensus-based optimization. *SIAM J. Optim.*, 33(1):211–236, 2023.
- [7] L. Bungert, F. Hoffmann, D. Y. Kim, and T. Roith. Mirrorco: A consensus-based optimization method in the spirit of mirror descent. *Math. Models Methods Appl. Sci.*, 35(14):3083–3170, 2025.
- [8] L. Bungert, T. Roith, and P. Wacker. Polarized consensus-based dynamics for optimization and sampling. *Math. Program.*, 211(1-2):125–155, 2025.
- [9] J. A. Carrillo, Y.-P. Choi, C. Totzeck, and O. Tse. An analytical framework for consensus-based global optimization method. *Math. Models Methods Appl. Sci.*, 28(6):1037–1066, 2018.
- [10] J. A. Carrillo, F. Hoffmann, A. M. Stuart, and U. Vaes. Consensus-based sampling. *Stud. Appl. Math.*, 148(3):1069–1140, 2022.
- [11] J. A. Carrillo, S. Jin, L. Li, and Y. Zhu. A consensus-based global optimization method for high dimensional machine learning problems. *ESAIM Control Optim. Calc. Var.*, 27:Paper No. S5, 22, 2021.
- [12] J. A. Carrillo, C. Totzeck, and U. Vaes. Consensus-based optimization and ensemble Kalman inversion for global optimization problems with constraints. In *Modeling and simulation for collective dynamics*, pages 195–230. World Scientific, 2023.
- [13] L.-P. Chaintron and A. Diez. Propagation of chaos: a review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models*, 15(6):895–1015, 2022.
- [14] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [15] R. Durrett. *Stochastic Calculus: A Practical Introduction*. CRC Press, Boca Raton, 1st edition, 1996. eBook Published 29 March 2018.

- [16] M. Fornasier, H. Huang, L. Pareschi, and P. Sünnen. Consensus-based optimization on hypersurfaces: well-posedness and mean-field limit. *Math. Models Methods Appl. Sci.*, 30(14):2725–2751, 2020.
- [17] M. Fornasier, H. Huang, L. Pareschi, and P. Sünnen. Consensus-based optimization on the sphere: convergence to global minimizers and machine learning. *J. Mach. Learn. Res.*, 22:Paper No. 237, 55, 2021.
- [18] M. Fornasier, H. Huang, L. Pareschi, and P. Sünnen. Anisotropic diffusion in consensus-based optimization on the sphere. *SIAM J. Optim.*, 32(3):1984–2012, 2022.
- [19] M. Fornasier, T. Klock, and K. Riedl. Consensus-based optimization methods converge globally. *SIAM J. Optim.*, 34(3):2973–3004, 2024.
- [20] M. Fornasier, P. Richtárik, K. Riedl, and L. Sun. Consensus-based optimisation with truncated noise. *European J. Appl. Math.*, 36(2):292–315, 2025.
- [21] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields*, 162(3-4):707–738, 2015.
- [22] N. J. Gerber, F. Hoffmann, D. Kim, and U. Vaes. Uniform-in-time propagation of chaos for consensus-based optimization, arXiv:2505.08669.
- [23] N. J. Gerber, F. Hoffmann, and U. Vaes. Mean-field limits for consensus-based optimization and sampling. *ESAIM Control Optim. Calc. Var.*, 31:Paper No. 74, 44, 2025.
- [24] S.-Y. Ha, S. Jin, and D. Kim. Convergence of a first-order consensus-based global optimization algorithm. *Math. Models Methods Appl. Sci.*, 30(12):2417–2444, 2020.
- [25] S.-Y. Ha, S. Jin, and D. Kim. Convergence and error estimates for time-discrete consensus-based optimization algorithms. *Numer. Math.*, 147(2):255–282, 2021.
- [26] S.-Y. Ha, M. Kang, D. Kim, J. Kim, and I. Yang. Stochastic consensus dynamics for nonconvex optimization on the Stiefel manifold: mean-field limit and convergence. *Math. Models Methods Appl. Sci.*, 32(3):533–617, 2022.
- [27] J. H. Holland. Genetic algorithms. *Scientific American*, 267(1):66–73, 1992.
- [28] H. Huang and H. Kouhkouh. Self-interacting CBO: existence, uniqueness, and long-time convergence. *Appl. Math. Lett.*, 161:Paper No. 109372, 6, 2025.
- [29] H. Huang and H. Kouhkouh. Uniform-in-time mean-field limit estimate for the Consensus-Based Optimization. *ESAIM Control Optim. Calc. Var.*, 31:Paper No. 69, 2025.
- [30] H. Huang, H. Kouhkouh, and L. Sun. Faithful global convergence for the rescaled consensus-based optimization, arXiv:2503.08578.
- [31] H. Huang and J. Qiu. On the mean-field limit for the consensus-based optimization. *Math. Methods Appl. Sci.*, 45(12):7814–7831, 2022.
- [32] D. Kalise, A. Sharma, and M. V. Tretyakov. Consensus-based optimization via jump-diffusion stochastic differential equations. *Math. Models Methods Appl. Sci.*, 33(2):289–339, 2023.
- [33] J. Kennedy. *Particle Swarm Optimization*, pages 760–766. Springer US, Boston, MA, 2010.
- [34] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of ICNN’95 - International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4, 1995.
- [35] S. Kirkpatrick, C. D. Gelatt, Jr., and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

- [36] R. Pinnau, C. Totzeck, O. Tse, and S. Martin. A consensus-based model for global optimization and its mean-field limit. *Math. Models Methods Appl. Sci.*, 27(1):183–204, 2017.
- [37] F. Rassoul-Agha and T. Seppäläinen. *A course on large deviations with an introduction to Gibbs measures*, volume 162 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2015.
- [38] L. Rastrigin. The convergence of the random search method in the extremal control of a many parameter system. *Automaton & Remote Control*, 24:1337–1342, 1963.
- [39] R. L. Schilling. *Brownian Motion: A Guide to Random Processes and Stochastic Calculus*. De Gruyter, Berlin and Boston, 3rd revised and extended edition, 2021.
- [40] Y. C. Shih and H. Teicher. *Probability Theory*. Springer Texts in Statistics. Springer New York, NY, 3 edition, 1997.
- [41] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [42] C. Totzeck. Trends in consensus-based optimization. In *Active particles. Vol. 3. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 201–226. Birkhäuser/Springer, Cham, [2022] ©2022.
- [43] C. Totzeck and M.-T. Wolfram. Consensus-based global optimization with personal best. *Math. Biosci. Eng.*, 17(5):6026–6044, 2020.
- [44] P. J. M. van Laarhoven and E. H. L. Aarts. *Simulated annealing: theory and applications*, volume 37 of *Mathematics and its Applications*. D. Reidel Publishing Co., Dordrecht, 1987.
- [45] J. Wei and W. Bian. Smoothing iterative consensus-based optimization algorithm for nonsmooth nonconvex optimization problems with global optimality. *J. Sci. Comput.*, 104(1):Paper No. 22, 33, 2025.

(Young-Pil Choi)

DEPARTMENT OF MATHEMATICS, YONSEI UNIVERSITY, SEOUL 03722, REPUBLIC OF KOREA

Email address: ypchoi@yonsei.ac.kr

(Seungchan Lee)

DEPARTMENT OF MATHEMATICS, YONSEI UNIVERSITY, SEOUL 03722, REPUBLIC OF KOREA

Email address: seungchan2718@yonsei.ac.kr

(Sihyun Song)

DEPARTMENT OF MATHEMATICS, YONSEI UNIVERSITY, SEOUL 03722, REPUBLIC OF KOREA

Email address: ssong@yonsei.ac.kr