# CREATION OF CHAOS FOR INTERACTING BROWNIAN PARTICLES

ARMAND BERNOU, MITIA DUERINCKX, AND MATTHIEU MÉNARD

ABSTRACT. We consider a system of $N$ Brownian particles, with or without inertia, interacting in the mean-field regime via a weak, smooth, long-range potential, and starting initially from an arbitrary exchangeable $N$-particle distribution. In this model framework, we establish a fine version of the so-called creation-of-chaos phenomenon: in weak norms, the mean-field approximation for a typical particle is shown to hold with an accuracy $O(N^{-1})$ up to an error due solely to initial pair correlations, which is damped exponentially over time. Corresponding higher-order results are also derived in the form of higher-order correlation estimates. The approach is new and easily adaptable: we start from suboptimal correlation estimates obtained from an elementary use of Itô's calculus on moments of the empirical measure, together with ergodic properties of the mean-field dynamics, and these bounds are then made optimal after combination with PDE estimates on the BBKY hierarchy.

MSC-CLASS: Primary 60K35; Secondary 35Q70, 82C22, 82C31, 35Q84.

## CONTENTS

## 1. INTRODUCTION

1.1. **General overview.** We consider the Langevin dynamics for a system of $N$ Brownian particles with mean-field interactions, moving in a confining potential in $\mathbb{R}^d$, $d \geq 1$, as described by the following system of coupled SDEs: for $1 \leq i \leq N$,

$$\begin{cases} \mathrm{d}X_t^{i,N} = V_t^{i,N}\mathrm{d}t, \\ \mathrm{d}V_t^{i,N} = -\frac{\kappa}{N}\sum_{j=1}^{N}\nabla W(X_t^{i,N} - X_t^{j,N})\,\mathrm{d}t - \frac{\beta}{2}V_t^{i,N}\mathrm{d}t - \nabla A(X_t^{i,N})\,\mathrm{d}t + \mathrm{d}B_t^i, \qquad t \geq 0, \end{cases} \tag{1.1}$$

where $\{Z^{i,N} := (X^{i,N}, V^{i,N})\}_{1 \leq i \leq N}$ is the set of positions and velocities of the particles in the phase space $\mathbb{X} := \mathbb{R}^d \times \mathbb{R}^d$, where $W : \mathbb{R}^d \to \mathbb{R}$ is a long-range interaction potential with the action-reaction condition $W(x) = W(-x)$, where $A$ is a uniformly convex confining potential, where $\{B^i\}_i$ are i.i.d. $d$-dimensional Brownian motions, and where $\kappa, \beta > 0$ are given constants. For simplicity, we choose the confining potential $A$ to be quadratic,

$$A(x) := \tfrac{1}{2}a|x|^2, \qquad \text{for some } a > 0, \tag{1.2}$$

although perturbations can be considered as well (see Section 1.3). Turning to a statistical description of the system, in terms of a probability density $F^N$ on the $N$-particle phase space $\mathbb{X}^N$, the Langevin dynamics (1.1) is equivalent to the corresponding Liouville equation

$$\partial_t F^N + \sum_{i=1}^{N} v_i \cdot \nabla_{x_i} F^N = \tfrac{1}{2} \sum_{i=1}^{N} \mathrm{div}_{v_i}((\nabla_{v_i} + \beta v_i) F^N)$$
$$+ \tfrac{\kappa}{N} \sum_{i,j=1}^{N} \nabla W(x_i - x_j) \cdot \nabla_{v_i} F^N + \sum_{i=1}^{N} \nabla A(x_i) \cdot \nabla_{v_i} F^N, \qquad t \geq 0. \quad (1.3)$$

In this contribution, we investigate the behavior of the system in the regime of a large number $N \gg 1$ of particles in the general setting of initial data $\{Z^{i,N}|_{t=0}\}_{1 \leq i \leq N}$ that are only assumed to be exchangeable (that is, their law $F^N|_{t=0}$ is invariant under permutation of the particles), but that may be strongly correlated. Let us recall two standard types of results for the system:

— *Propagation of chaos:*

If particles have weak correlations initially, then propagation of chaos is expected to hold, which means that only little correlations can build over time. As a consequence, the distribution of a typical particle is approximated by the solution $\mu$ of the mean-field Vlasov–Fokker–Planck equation

$$\begin{cases} \partial_t \mu + v \cdot \nabla_x \mu = \tfrac{1}{2} \mathrm{div}_v((\nabla_v + \beta v)\mu) + (\nabla A + \kappa \nabla W * \mu) \cdot \nabla_v \mu, \qquad t \geq 0, \\ \mu|_{t=0} = \mu_\circ, \end{cases} \quad (1.4)$$

where $\mu_\circ$ stands for the initial single-particle distribution. Such a result has been established in various settings and we refer e.g. to the recent work [5, 4] in case of singular interactions. More precisely, for all $1 \leq m \leq N$, denoting by $F^{N,m}$ the distribution of $m$ typical particles in the system, as given by the $m$-th marginal of $F^N$,

$$F^{N,m}(z_1, \ldots, z_m) = \int_{\mathbb{X}^{N-m}} F^N(z_1, \ldots, z_N) \, \mathrm{d}z_{m+1} \ldots \mathrm{d}z_N, \qquad 1 \leq m \leq N,$$

we find that $F^{N,m}$ converges to $\mu^{\otimes m}$ as $N \uparrow \infty$ for any fixed $m$. Beyond this qualitative convergence, an optimal $O(N^{-1})$ error estimate for the convergence of marginals was first obtained in [10] for smooth interactions, and later in [15] for bounded non-smooth interactions in a better metric: if correlations vanish initially, then there holds

$$\|F_t^{N,m} - \mu_t^{\otimes m}\|_{TV} \leq C e^{Ct} m N^{-1}, \qquad \text{for all } 1 \leq m \leq N \text{ and } t \geq 0. \quad (1.5)$$

If in addition the intensity $\kappa$ of interaction forces is small enough, which ensures the uniqueness of the mean-field equilibrium, then this error estimate (1.5) holds uniformly in time without the exponential factor [16, 1] (see also previous work using coupling methods but only producing a suboptimal convergence rate [3, 13]).

— *Relaxation to equilibrium:*

If interaction forces are weak enough, even if particles are strongly correlated initially, their distribution is known to relax to Gibbs equilibrium on long times. For Lipschitz forces, the following uniform-in-$N$ exponential estimate was established in [3],

$$\mathcal{W}_2(F_t^{N,m}, M^{N,m}) \leq C_m e^{-\frac{1}{C} t}, \qquad \text{for all } 1 \leq m \leq N \text{ and } t \geq 0,$$

where $M^{N,m}$ stands for the $m$-th marginal of the $N$-particle Gibbs equilibrium $M^N$. As the latter exhibits weak correlations, this entails in particular that initial correlations between the particles must be exponentially damped over time.

Both uniform-in-time propagation of chaos and convergence to Gibbs equilibrium rely essentially on the same phenomenon: due to the diffusion, correlations cannot accumulate too much over time and must somehow be damped. This leads to the natural related question of *creation of chaos* as first raised by Del Moral and Tugaut [7] (see also "generation of chaos" in [17, 19]): *for general exchangeable initial data, is it possible to show that the mean-field approximation holds to order $O(N^{-1})$ up to an error coming from initial correlations, which is damped exponentially over time?*

Such a result can in fact be deduced immediately from the coupling argument of [3], but only with a suboptimal convergence rate $O(N^{-1/2})$: more precisely, for weak Lipschitz forces, it follows from [3,

Theorem 2 and Remark 12] that

$$\mathcal{W}_2(F_t^{N,m}, \mu_t^{\otimes m}) \leq C_m N^{-\frac{1}{2}} + C_m e^{-\frac{1}{C}t} \mathcal{W}_2(F_0^{N,m}, \mu_\circ^{\otimes m}), \quad \text{for all } 1 \leq m \leq N \text{ and } t \geq 0. \quad (1.6)$$

Our first objective is to achieve the optimal rate $O(N^{-1})$ in this creation-of-chaos result, which to our knowledge is not possible via the same coupling methods. More precisely, we look for bounds of the following form, in some suitable metric,

$$\|F_t^{N,m} - \mu_t^{\otimes m}\| \leq C_m N^{-1} + C_m e^{-\frac{1}{C}t}\|F_\circ^{N,m} - \mu_\circ^{\otimes m}\|, \quad \text{for all } 1 \leq m \leq N \text{ and } t \geq 0. \quad (1.7)$$

In the sequel, we establish estimates of this kind for weak smooth interactions, in negative Sobolev norms instead of Wasserstein distance; see Theorem 1.1 below.

While virtually any proof of uniform-in-time propagation of chaos comes with some associated creation-of-chaos result showing the damping of initial correlations, we emphasize that clean estimates of the form (1.7) are more difficult to establish. In particular, there has been a lot of recent progress based on hierarchical approaches to uniform-in-time propagation of chaos [15, 5, 16, 1, 4], starting from the BBGKY hierarchy of equations for marginals, but such methods require to deal with information on the *whole* hierarchy $\{F^{N,m} - \mu^{\otimes m}\}_{1 \leq m \leq N}$ at once and can thus a priori not be used to obtain estimates of the form (1.7) for any *fixed m*. For overdamped systems (see Section 1.3.1 below), modulated energy methods can also be used to establish creation of chaos, see [19], but they require much stronger initial information in the form of the relative entropy of the $N$-particle distribution with respect to a tensorized state, which again cannot lead to clean estimates of the form (1.7) for marginals.

Next to this question on the creation and propagation of chaos, we further refine our analysis to corresponding correlation estimates. Many-particle correlation functions provide finer information on the propagation of chaos in the system and we shall see that creation of chaos holds more generally at the level of arbitrary high-order correlations. Recall that the two-particle correlation function is defined as

$$G^{N,2} := F^{N,2} - (F^{N,1})^{\otimes 2},$$

which captures the defect to propagation of chaos (1.5) at the level of two-particle statistics. This quantity alone does not allow to reconstruct the full particle density $F^N$: we define higher-order correlation functions $\{G^{N,m}\}_{2 \leq m \leq N}$ as suitable polynomial combinations of marginals of $F^N$ in such a way that the full particle distribution $F^N$ is recovered in form of a cluster expansion,

$$F^N(z_1, \ldots, z_N) = \sum_{\pi \vdash [\![N]\!]} \prod_{B \in \pi} G^{N, \sharp B}(z_B), \quad (1.8)$$

where $\pi$ runs through the list of all partitions of the index set $[\![N]\!] := \{1, \ldots, N\}$, where $B$ runs through the list of blocks of the partition $\pi$, where $\sharp B$ is the cardinality of $B$, where for $B = \{i_1, \ldots, i_l\} \subset [\![N]\!]$ we write $z_B = (z_{i_1}, \ldots, z_{i_l})$, and where we have set $G^{N,1} := F^{N,1}$. As is easily checked, correlation functions are fully determined by prescribing (1.8) together with the requirement $\int_{\mathbb{X}} G^{N,m}(z_1, \ldots, z_m) \, dz_l = 0$ for $2 \leq m \leq N$ and $1 \leq l \leq m$. If initial correlations vanish, standard propagation of chaos leads to $G_t^{N,2} = O(N^{-1})$, and a formal analysis of the BBGKY hierarchy further leads to expect

$$G_t^{N,m} = O(N^{1-m}), \quad 2 \leq m \leq N. \quad (1.9)$$

This is referred to as *higher-order propagation of chaos* and is a key tool to describe deviations from mean-field theory, see e.g. [9, 10, 11] and in particular [2, Corollary 1.3] for the derivation of the so-called Bogolyubov correction to mean field. Such estimates have been obtained in several settings with an exponential time growth: non-Brownian particle systems with smooth interactions were first addressed in [10], while in [14] the overdamped dynamics was further covered for bounded non-smooth interactions. If interactions are weak enough, we proved in our previous work [2] that (1.9) actually holds uniformly in time, and the same was obtained later in [21] for the overdamped dynamics with bounded non-smooth interactions. All these results have been obtained only on the condition that initial correlations vanish. Yet, uniform-in-time versions of (1.9) hint towards a corresponding 'higher-order' creation-of-chaos phenomenon at the level of correlation functions: *for general exchangeable*

*initial data, is it possible to show that the m-particle correlation $G^{N,m}$ is of order $O(N^{1-m})$ only up to an error coming from initial m-particle correlations (or at least, from a finite number of high-order correlations), which is damped exponentially over time?* For two-particle correlations, it already follows from (1.7) that

$$\|G_t^{N,2}\| \ \leq \ CN^{-1} + Ce^{-\frac{1}{C}t}\|G_\circ^{N,2}\|. \tag{1.10}$$

In the sequel, we obtain similar estimates for all higher-order correlations, see Theorem 1.2 below. The only other result of this kind that we are aware of is the recent work of Lukkarinen and Vuoksenmaa [17] regarding energy cumulants for the stochastic Kac model. In that work, however, the authors appeal to hierarchical techniques, which require to deal with the whole hierarchy $\{G^{N,m}\}_{1 \leq m \leq N}$ at once and do not yield simpler estimates of the form (1.10) for correlations of a fixed order.

**Notation.**
— We denote by $C \geq 1$ any constant that only depends on the space dimension $d$, or possibly on other controlled quantities when specified. We use the notation $\lesssim$ for $\leq C\times$ up to such a multiplicative constant $C$. We write $a \simeq b$ when both $a \lesssim b$ and $b \lesssim a$ hold. We also use the notation $\ll$ (resp. $\gg$) for $\leq \frac{1}{C}\times$ (resp. $\geq C\times$) for some large enough constant $C$. We occasionally add subscripts to $C, \lesssim, \simeq, \ll, \gg$ to indicate dependence on other parameters.
— For any two integers $m \geq n \geq 0$, we use the short-hand notation $[\![m, n]\!] := \{m, m+1, \ldots, n\}$, and in addition for any integer $m \geq 1$ we set $[\![m]\!] := [\![1, m]\!]$.
— For all $z \in \mathbb{X}$, we use the notation $\langle z \rangle := (1 + |z|^2)^{\frac{1}{2}}$.
— For any measurable space $(E, \mathcal{E})$, we write $\mathcal{P}(E)$ for the space of probability measures on $E$. We endow $\mathbb{X}$ with the Borel $\sigma$-algebra.
— For a measure $\mu \in \mathcal{P}(\mathbb{X})$, we use the following short-hand notation for its moments,

$$Q_r(\mu) \ := \ \int_{\mathbb{X}} \langle z \rangle^r \mu(\mathrm{d}z), \qquad r \geq 0. \tag{1.11}$$

1.2. **Main results.** We start by defining the appropriate functional framework for our results: for $r \geq 0$, $1 < q < 2$, and $0 < p \leq 1$, we consider the following weighted negative Sobolev norms on functions $h : \mathbb{X} \to \mathbb{R}$,

$$\|h\|_{W_p^{-r,q}(\mathbb{X})} \ := \ \sup\left\{\Big|\int_{\mathbb{X}} \varphi h\Big| \ : \ \varphi \in C_c^\infty(\mathbb{X}), \ \ \|\langle\cdot\rangle^{-p}\varphi\|_{W^{r,q'}(\mathbb{X})} \leq 1\right\} \tag{1.12}$$

and for $m \geq 1$ and functions $h : \mathbb{X}^m \to \mathbb{R}$,

$$\|h\|_{W_p^{-r,q}(\mathbb{X})^{\otimes m}} \ := \ \sup\left\{\Big|\int_{\mathbb{X}^m} \varphi^{\otimes m} h\Big| \ : \ \varphi \in C_c^\infty(\mathbb{X}), \ \ \|\langle\cdot\rangle^{-p}\varphi\|_{W^{r,q'}(\mathbb{X})} \leq 1\right\}.[1] \tag{1.13}$$

The choice of these spaces is dictated by the available ergodic estimates on the linearized mean-field evolution, which are not known to hold on the simpler unweighted spaces $W^{-k,1}(\mathbb{X})$; see e.g. [2] and Lemma 2.1 below.

Our first main result justifies the desired estimates (1.7) on the creation and propagation of chaos in the above norms. The estimates appear essentially optimal, except for the dependence on $m$, for the strong regularity assumption, and perhaps for the loss of regularity in the norms.

**Theorem 1.1** (Creation of chaos)**.** *Let the interaction potential $W$ be even, smooth and decaying in the sense of $W \in W^{d+3,\infty} \cap H^s(\mathbb{R}^d)$ for some $s > \frac{d}{2} + 5$, and let the confining potential $A$ be quadratic (1.2). Given an exchangeable $N$-particle distribution $F_\circ^N \in \mathcal{P}(\mathbb{X}^N)$ with all bounded moments, consider the weak solution $F^N \in C(\mathbb{R}^+; \mathcal{P}(\mathbb{X}^N))$ of the Liouville equation (1.3) with initial data $F_\circ^N$. Given $\mu_\circ \in \mathcal{P}(\mathbb{X})$ with bounded second moments, also consider the weak solution $\mu \in C(\mathbb{R}^+; \mathcal{P}(\mathbb{X}))$ of the mean-field equation (1.4) with initial data $\mu_\circ$. Then there is $\lambda_0 > 0$ such that the following holds:*

---

[1]This is the norm of the injective tensor product $W_p^{-r,q}(\mathbb{X})^{\otimes m}$, which of course differs from the norm of $W_p^{-r,q}(\mathbb{X}^m)$.

given $1 < q \le 2$ and $0 < p \le \frac{1}{6}$ with $pq' \gg 1$ large enough, and provided that $0 < \kappa \ll 1$ is small enough, we have for all $1 \le m \le N$ and $t \ge 0$,[2]

$$\|F_t^{N,m} - \mu_t^{\otimes m}\|_{W_p^{-3,q}(\mathbb{X})^{\otimes m}} \lesssim N^{-1} + \sum_{j=1}^{\tilde{m}} e^{-j\lambda_0 pt}\|F_\circ^{N,j} - \mu_\circ^{\otimes j}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes j}}, \tag{1.14}$$

where we have set for abbreviation $\tilde{m} = m$ for $m$ even and $\tilde{m} = m + 1$ for $m$ odd.

We turn to the creation of chaos at the level of correlation estimates. The precise combination of initial correlations that needs to be tracked quickly becomes quite cumbersome for high-order correlations. Although its expression is easily inferred from the proof, we stick here, for readability, to a simpler statement where we assume that correlations have a definite scaling initially. We emphasize that for estimates on $m$-particle correlations we only need information on initial $j$-particle correlations with $j \le 2m - 1$, cf. (1.16), in contrast to the much stronger requirements for hierarchical arguments. In the case of chaotic initial data, we recover in particular the uniform-in-time correlation estimates obtained in [2], cf. (1.18).

**Theorem 1.2** (Creation of chaos in correlation estimates). *Let $W, A, F^N, \lambda_0, p, q, \kappa$ be as in the statement of Theorem 1.1, and further assume $W \in C_b^\infty \cap H^\infty(\mathbb{R}^d)$. Then the two-particle correlation function satisfies for all $t \ge 0$,*[3]

$$\|G_t^{N,2}\|_{W_p^{-3,q}(\mathbb{X})^{\otimes 2}} \lesssim N^{-1} + e^{-\lambda_0 pt}\|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}}. \tag{1.15}$$

*In addition, given $2 \le m \le N$ and $\alpha \in [0, 1]$, if we have initially*

$$\|G_\circ^{N,j}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes j}} \le C_j^\circ N^{\alpha(1-j)}, \qquad \text{for all } 2 \le j \le 2m - 1, \tag{1.16}$$

*and if $\kappa$ is small enough (further depending on $m$), then we have for all $t \ge 0$,*

$$\|G_t^{N,m}\|_{W_{p_m}^{-r_m,q_m}(\mathbb{X})^{\otimes m}} \lesssim N^{1-m} + e^{-\lambda_0 p_m t}N^{\alpha(1-m)}, \tag{1.17}$$

*for $r_m = m + 1$, some $q_m' \simeq_m q'$, and $p_m \simeq_m p$ with $p_m q_m' = pq'$. In particular, if the initial data are chaotic, that is, if $F_\circ^N = \mu_\circ^{\otimes N}$ for some $\mu_\circ \in \mathcal{P}(\mathbb{X})$, then we have for all $2 \le m \le N$ and $t \ge 0$,*

$$\|G_t^{N,m}\|_{W_{p_m}^{-r_m,q_m}(\mathbb{X})^{\otimes m}} \lesssim N^{1-m}. \tag{1.18}$$

**Remark 1.3** (Smallness assumption). In the above results (1.17)–(1.18), the smallness assumption on $\kappa$ depends on the order $m$ of the correlation function of interest. This is necessary in our approach (except in the simpler setting of the overdamped dynamics, see Section 1.3.1 below), but it can be at least partially relaxed. More precisely, focusing for shortness on the particular case of chaotic initial data, $F_\circ^N = \mu_\circ^{\otimes N}$, we can prove the following suboptimal bounds: letting $W, A, F^N, p, q, \kappa$ be as in Theorem 1.1 (hence, with $\kappa$ independent of $m$), we have for all $2 \le m \le N$ and $t \ge 0$,

$$\|G_t^{N,m}\|_{W_{p_m}^{-2,q}(\mathbb{X})^{\otimes m}} \lesssim N^{-\frac{m}{2}}.$$

(To see this, we refer to Step 1 of the proof of Theorem 1.2 in Section 4, simply specializing it to the case of chaotic data, hence $\alpha = 1$.)

1.3. **Extensions.** We briefly discuss some simple extensions of the above results. The adaptation to the Kac model, based on ergodic estimates for the space-homogeneous Boltzmann equation and the use of a similar BBGKY hierarchy, is a bit more demanding and is postponed to a separate work, improving the recent result of Lukkarinen and Vuoksenmaa in [17].

---

[2]In this statement, the exponent $\lambda_0$ only depends on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$; the condition $pq' \gg 1$ only depends on $d, \beta, a$; the smallness condition $\kappa \ll 1$ further depends on $p, q, Q_2(\mu_\circ), Q_2(F_\circ^{N,1})$, and on controlled norms of $W$; and the multiplicative constant in (1.14) further depends on $m$ and $Q_{mp}(F_\circ^{N,1})$.

[3]The multiplicative constant in (1.15) only depends on $d, \beta, a, Q_2(F_\circ^{N,1})$, and on $\|W\|_{W^{d+3,\infty} \cap H^s(\mathbb{X})}$ for some $s > \frac{d}{2} + 5$; the additional smallness condition on $\kappa$ only depends on $p, q, Q_2(F_\circ^{N,1}), m$; and the multiplicative constants in (1.17) and (1.18) further depend on $Q_{mp}(F_\circ^{N,1}), \|W\|_{W^{d+m+2,\infty}(\mathbb{R}^d)}$, and on the constants $(C_j^\circ)_{2 \le j \le 2m-1}$ in (1.16).

1.3.1. *Overdamped dynamics.* Our results also apply to the overdamped Langevin dynamics, as described by the following system of coupled SDEs: for $1 \le i \le N$,

$$dY_t^{i,N} = \frac{\kappa}{N} \sum_{j=1}^N K(Y_t^{i,N} - Y_t^{j,N}) \, dt - \nabla A(Y_t^{i,N}) + dB_t^i, \qquad t \ge 0, \tag{1.19}$$

where $K : \mathbb{R}^d \to \mathbb{R}^d$ is a smooth force kernel with the action-reaction condition $K(x) = -K(-x)$ (we may take $K$ of the form $K = -\nabla W$, but this is not necessary). In terms of a probability density $\tilde{F}^N$ on the $N$-particle phase space $(\mathbb{R}^d)^N$, this particle dynamics is equivalent to the following Liouville equation,

$$\partial_t \tilde{F}^N = \tfrac{1}{2} \sum_{i=1}^N \triangle_{x_i} \tilde{F}_t^N - \frac{\kappa}{N} \sum_{i,j=1}^N \operatorname{div}_{x_i} \big( K(x_i - x_j) \tilde{F}^N \big) + \sum_{i=1}^N \operatorname{div}_{x_i} \big( \nabla A(x_i) \tilde{F}^N \big). \tag{1.20}$$

The associated mean-field equation is now the McKean–Vlasov equation

$$\partial_t \tilde{\mu} = \tfrac{1}{2} \triangle \tilde{\mu} - \kappa \operatorname{div}(\tilde{\mu}(K * \tilde{\mu})) + \operatorname{div}(\tilde{\mu} \nabla A). \tag{1.21}$$

In this setting, by standard parabolic theory, ergodic estimates for the linearized mean-field equation hold in the *unweighted* spaces $W^{-k,1}(\mathbb{R}^d)$ for $k \ge 0$, see [8, 2]. This simplifies our proof substantially: while working in *weighted* Sobolev spaces in the kinetic setting forces us to use fine concentration estimates for moments of the particle dynamics, this is no longer needed here. In addition, when adapting the proof to the present unweighted framework, we find that the smallness requirement on $\kappa$ in correlation estimates is no longer required to depend on the order of the correlation. More precisely, we obtain the following.

**Theorem 1.4** (Creation of chaos — overdamped setting)**.** *Let the interaction force $K$ be odd, smooth and decaying in the sense of $K \in C_b^\infty \cap H^\infty(\mathbb{R}^d)$, and let the confining potential $A$ be quadratic (1.2). Given an exchangeable $N$-particle distribution $\tilde{F}_\circ^N \in \mathcal{P}((\mathbb{R}^d)^N)$ with bounded second moments, we consider the solution $\tilde{F}^N \in C(\mathbb{R}^+; \mathcal{P}((\mathbb{R}^d)^N))$ of (1.20) with initial data $\tilde{F}_\circ^N$ and denote by $\{\tilde{G}^{N,m}\}_{2 \le m \le N}$ the corresponding correlation functions. Given $\tilde{\mu}_\circ \in \mathcal{P}(\mathbb{R}^d)$ with bounded second moments, also consider the weak solution $\tilde{\mu} \in C(\mathbb{R}^+; \mathcal{P}(\mathbb{R}^d))$ of the McKean–Vlasov equation (1.21) with initial data $\tilde{\mu}_\circ$. Then there is $\lambda_0 > 0$ such that the following hold for $0 < \kappa \ll 1$ small enough.*[4]*
— For all $1 \le m \le N$ and $t \ge 0$ we have*

$$\|\tilde{F}_t^{N,m} - \tilde{\mu}_t^{\otimes m}\|_{W^{-3,1}(\mathbb{R}^d)^{\otimes m}} \lesssim N^{-1} + \sum_{j=1}^{\tilde{m}} e^{-j\lambda_0 t} \|\tilde{F}_\circ^{N,j} - \tilde{\mu}_\circ^{\otimes j}\|_{W^{-2,1}(\mathbb{R}^d)^{\otimes j}}, \tag{1.22}$$

*where we recall the short-hand notation $\tilde{m} = m + \mathbb{1}_{m \ odd}$.*
— *Given $2 \le m \le N$ and $\alpha \in [0,1]$, if we have initially*

$$\|\tilde{G}_\circ^{N,j}\|_{W^{-2,1}(\mathbb{R}^d)^{\otimes j}} \le C_j^\circ N^{\alpha(1-j)}, \qquad \text{for all } 1 \le j \le 2m - 1, \tag{1.23}$$

*then we have for all $t \ge 0$,*

$$\|\tilde{G}_t^{N,m}\|_{W^{-m-1,1}(\mathbb{R}^d)^{\otimes m}} \lesssim N^{1-m} + e^{-\lambda_0 t} N^{\alpha(1-m)}. \tag{1.24}$$

*In addition, if the initial data are chaotic, that is, if $\tilde{F}_\circ^N = \tilde{\mu}_\circ^{\otimes N}$, then for all $2 \le m \le N$ and $t \ge 0$,*

$$\|\tilde{G}_t^{N,m}\|_{W^{-m-1,1}(\mathbb{R}^d)^{\otimes m}} \lesssim N^{1-m}.$$

---

[4]In this statement, the exponent $\lambda_0$ only depends on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$; the smallness condition on $\kappa$ further depends on $Q_2(\mu_\circ), Q_2(F_\circ^{N,1})$, and on some norms of $W$; the multiplicative constant in (1.22) further depends on $m$ and on some norms of $W$ also depending on $m$; the multiplicative constant in (1.24) further depends on the constants $(C_j^\circ)_{2 \le j \le 2m-1}$ in (1.23).

1.3.2. *Periodic setting.* Our analysis is easily adapted to particle systems on the torus $\mathbb{T}^d$, for instance in the spatially homogeneous setting $A \equiv 0$. However, in that case, for the Langevin dynamics (1.1), ergodic estimates for the linearized mean-field equation only yield a super-polynomial decay $t^{-\infty}$ instead of exponential; see indeed [20, Theorem 56] and [2, Remark 2.14]. As a result, Theorems 1.1 and 1.2 then take the same form as above up to replacing the exponential time decay $e^{-ct}$ by $t^{-\infty}$. For the overdamped dynamics (1.19), on the other hand, the exponential decay still holds in the periodic setting.

1.3.3. *Non-quadratic confinement.* Both for the Langevin dynamics (1.1) and the overdamped dynamics (1.19), our results would hold in the same form if the quadratic confinement (1.2) were replaced by $A(x) = a|x|^2 + A'(x)$ for some $a > 0$ and some smooth potential $A' \in C_c^\infty(\mathbb{R}^d)$, provided that $\|\nabla^2 A'\|_{L^\infty(\mathbb{R}^d)}$ is small enough (only depending on $\beta, W, a$). In that case, we can still appeal to [2] to ensure the validity of the ergodic estimates of Lemma 2.1 below, while the rest of our approach can be adapted directly without major difficulties.

1.3.4. *More general interactions.* While we focus for shortness on pairwise interactions deriving from a potential, the term $\frac{\kappa}{N} \sum_{j=1}^N \nabla W(X^{i,N} - X^{j,N})$ in (1.1) (or correspondingly in (1.19)) could be replaced by a more general interaction term of the form $\kappa\, b(X^{i,N}, \mu^N)$ for some smooth functional $b : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}$, where $\mu^N$ stands for the empirical measure

$$\mu_t^N := \frac{1}{N} \sum_{i=1}^N \delta_{Z_t^{i,N}} \in \mathcal{P}(\mathbb{X}). \tag{1.25}$$

Equation (1.1) corresponds to the particular choice $b(x, \mu) = -\nabla W * \mu(x)$, but our analysis is easily adapted to more general functionals $b$ under suitable regularity assumptions both on $b$ and on its derivatives with respect to the measure argument, together with a corresponding smallness condition on $\kappa$. We skip the details and refer to [8] for the needed adaptations (see in particular the regularity assumptions (Reg) and (Lip) in [8, Section 2.3.2]).

1.4. **Strategy and plan of the paper.** Our starting point is the derivation of new estimates for quantities of the form

$$A^m(t) := \sup_\varphi \mathbb{E}\left[ \left| \int_{\mathbb{X}} \varphi(\mu_t^N - \mu_t) \right|^m \right],$$

where the supremum runs over a suitable class of smooth test functions $\varphi$. Analyzing the dynamics of those quantities and appealing to the ergodic estimates of the linearized mean-field equation, we are led to proving roughly that for all $t \geq 0$,

$$A^m(t) \leq \left( \frac{Cm}{N} \right)^{\frac{m}{2}} + C^m e^{-\lambda_0 m t} A^m(0), \tag{1.26}$$

see Proposition 2.3. For the Langevin dynamics, as ergodic estimates only hold in weighted Sobolev spaces, we lose moments of the particle dynamics in the estimates, which we manage to control by means of suitable concentration estimates.

When reformulating $A^m$ in terms of marginals or correlations, this bound (1.26) leads only to suboptimal estimates. In particular, as $t \uparrow \infty$, we only deduce a bound $O(N^{-1/2})$ instead of $O(N^{-1})$ for the error in the propagation of chaos, and a bound $O(N^{-m/2})$ instead of $O(N^{1-m})$ for the $m$-particle correlation. In order to recover the optimal scalings, our next main ingredient is the BBGKY hierarchy: more precisely, appealing again to ergodic properties of the linearized mean-field equation, we show that the optimal scalings can be rigorously extracted from the BBGKY hierarchy, noting that the latter can be truncated using the suboptimal estimates obtained from (1.26).

We emphasize that relatively few ingredients are used in this approach: the basic estimate (1.26) relies solely on ergodic estimates for the linearized mean-field equation, on concentration properties, and on some regularity of the interaction kernel. The strategy appears versatile and should be applicable

to other mean-field systems, including the Kac model mentioned above, which is postponed to a future work.

*Plan of the paper.* Section 2 is devoted to the statement of ergodic estimates, the proof of concentration bounds for moments of the particle dynamics, and most importantly the proof of (1.26), cf. Proposition 2.3. Next, Sections 3 and 4 are devoted to the proof of Theorems 1.1 and 1.2 as a consequence of (1.26) together with BBGKY estimates.

## 2. Preliminary results

As correlations are driven by the linearized mean-field operator, a key ingredient to understand the damping of initial correlations is naturally given by ergodic estimates for linearized mean field. Recall that $\mu$ stands for the solution of the mean-field equation (1.4) with initial data $\mu_\circ$. This equation admits a unique steady state $M$ for $0 < \kappa \ll 1$ small enough, which can be characterized as the unique solution to the fixed-point Gibbs equation

$$M(x,v) = c_M \exp\Big[ - \beta\Big(\tfrac{1}{2}|v|^2 + A(x) + \kappa W * M(x)\Big)\Big], \qquad (x,v) \in \mathbb{X},$$

where $c_M$ is the normalizing constant such that $\int_{\mathbb{X}} M = 1$. This fixed-point equation has indeed a unique solution when $\kappa\beta\|W\|_{L^\infty(\mathbb{R}^d)} < 1$. For $h \in C_c^\infty(\mathbb{X})$ and $t \geq 0$, we consider the (non-autonomous) backward dual linearized mean-field evolution,

$$\begin{cases} \partial_s U_{s,t} h + v \cdot \nabla_x U_{s,t} h = \Big( - \tfrac{1}{2}\triangle_v + \tfrac{1}{2}\beta v \cdot \nabla_v + (\nabla A + \kappa\nabla W * \mu_s) \cdot \nabla_v\Big) U_{s,t} h, \qquad s \leq t, \\ U_{s,t} h|_{s=t} = h. \end{cases} \quad (2.1)$$

In [2, Proposition 7.1], we obtained the following ergodic estimates in weighted Sobolev spaces. This is a consequence of hypocoercivity techniques together with a suitable use of the enlargement theory of Mischler and Mouhot [18]. (In [2], an estimate is obtained instead for the forward primal linearized evolution, but it is equivalent to this one by duality.)

**Lemma 2.1** (Dual ergodic estimates, [2]). *There exist $\kappa_0, \lambda_0 > 0$ (depending on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$) such that given $\kappa \in [0, \kappa_0]$ we have for all $r \geq 0$, $1 < q \leq 2$, $0 < p \leq 1$ with $pq' \gg 1$ large enough (only depending on $d, \beta, a$), $t \geq s \geq 0$, and $h \in C_c^\infty(\mathbb{X})$,*

$$\Big\|\langle z\rangle^{-p}\Big(U_{s,t}h - \int_{\mathbb{X}} hM\Big)\Big\|_{W^{r,q'}(\mathbb{X})} \lesssim e^{-\lambda_0 p(t-s)}\|\langle z\rangle^{-p}h\|_{W^{r,q'}(\mathbb{X})},$$

*where the multiplicative constant only depends on $d, \beta, p, q, r, a, \|W\|_{W^{r+d+1,\infty}(\mathbb{R}^d)}$, and $Q_2(\mu_\circ)$.*

We emphasize that the above ergodic estimates are not known to hold in $W^{r,\infty}(\mathbb{X})$, but only in weighted spaces $\langle z\rangle^p W^{r,q'}(\mathbb{X})$ with $pq' \gg 1$. As we work in such weighted spaces, we will need to control moments of the particle dynamics. More precisely, we will appeal to the following concentration bounds for moments of the empirical measure $\mu^N$ given in (1.25).

**Lemma 2.2** (Moment estimates). *Let $\kappa_0 > 0$ be as in Lemma 2.1. There exist $\lambda_0 > 0$ (only depending on $\beta, a$) and $C_0 > 0$ (further depending on $d, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$), such that for all $2 \leq m \leq N$, $0 < p \leq 1$, $L \geq C_0 Q_{2p}(\mu_\circ)$, and $t \geq 0$ we have*

$$\mathbb{E}\big[(Q_{2p}(\mu_t^N) - L)_+^m\big] \leq \Big(\frac{C_0 L m}{N}\Big)^{\frac{m}{2}} + C_0^m e^{-\lambda_0 pmt}\mathbb{E}\Big[\Big|\int_{\mathbb{X}} \langle z\rangle^{2p}(\mu_0^N - \mu_\circ)\Big|^m\Big]. \quad (2.2)$$

*In addition, for all $r \geq 0$,*

$$\mathbb{E}[Q_r(\mu_t^N)] = Q_r(F_t^{N,1}) \leq C_0(C_0 r)^{\frac{r}{2}} + C_0^r e^{-\lambda_0 rt} Q_r(F_\circ^{N,1}). \quad (2.3)$$

*Proof.* We split the proof into two steps, starting with the proof of the concentration estimate (2.2), before turning to the more classical proof of the moment bound (2.3).

*Step* 1. Proof of (2.2).

Let $0 < p \leq 1$ be fixed. In the spirit of [3], we consider the random process

$$G_t^N := \frac{1}{N} \sum_{i=1}^{N} (R_t^{i,N})^p, \qquad R_t^{i,N} := 1 + a|X_t^{i,N}|^2 + |V_t^{i,N}|^2 + \eta X_t^{i,N} \cdot V_t^{i,N},$$

for some $\eta \in (0, 2\sqrt{a})$ to be chosen later on. Note that this range of $\eta$ ensures

$$R_t^{i,N} \simeq_{a,\eta} \langle Z_t^{i,N} \rangle^2, \qquad G_t^N \simeq_{a,\eta} Q_{2p}(\mu_t^N). \tag{2.4}$$

The particle dynamics (1.1) and Itô's formula yield

$$\mathrm{d}R_t^{i,N} = -\Big(a\eta|X_t^{i,N}|^2 + (\beta - \eta)|V_t^{i,N}|^2 + \frac{\eta\beta}{2}X_t^{i,N} \cdot V_t^{i,N} - 1\Big)\mathrm{d}t$$

$$- \frac{\kappa}{N} \sum_{j=1}^{N} (2V_t^{i,N} + \eta X_t^{i,N}) \cdot \nabla W(X_t^{i,N} - X_t^{j,N})\mathrm{d}t + (2V_t^{i,N} + \eta X_t^{i,N}) \cdot \mathrm{d}B_t^i.$$

Provided that $0 < \eta \ll 1$ is chosen small enough (only depending on $\beta, a$), we can find $\lambda_0 > 0$ only depending on $\beta, a$ such that

$$\mathrm{d}R_t^{i,N} \leq -2\lambda_0 R_t^{i,N}\mathrm{d}t + C_0\mathrm{d}t + (2V_t^{i,N} + \eta X_t^{i,N}) \cdot \mathrm{d}B_t^i, \tag{2.5}$$

where henceforth $C_0$ stands for a positive constant only depending on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$, the value of which may change from line to line. By Itô's formula, using $p(p-1) \leq 0$ and $(R_t^{i,N})^{p-1} \leq 1$, we deduce

$$\begin{aligned}\mathrm{d}(R_t^{i,N})^p &\leq -2\lambda_0 p(R_t^{i,N})^p\mathrm{d}t + C_0 p(R_t^{i,N})^{p-1}\mathrm{d}t + \tfrac{1}{2}p(p-1)(R_t^{i,N})^{p-2}|2V_t^{i,N} + \eta X_t^{i,N}|^2\mathrm{d}t \\ &\quad + p(R_t^{i,N})^{p-1}(2V_t^{i,N} + \eta X_t^{i,N}) \cdot \mathrm{d}B_t^i \\ &\leq -2\lambda_0 p(R_t^{i,N})^p\mathrm{d}t + C_0 p\,\mathrm{d}t + p(R_t^{i,N})^{p-1}(2V_t^{i,N} + \eta X_t^{i,N}) \cdot \mathrm{d}B_t^i,\end{aligned}$$

and thus, summing over $i$ and recalling the definition of $G_t^N$,

$$\mathrm{d}G_t^N \leq -2\lambda_0 p G_t^N\mathrm{d}t + C_0 p\,\mathrm{d}t + \frac{p}{N}\sum_{i=1}^{N}(R_t^{i,N})^{p-1}(2V_t^{i,N} + \eta X_t^{i,N}) \cdot \mathrm{d}B_t^i.$$

Further using Itô's generalized formula for convex functions, bounding classically the expectation of the local time using the quadratic variation, we get for all $L \geq 0$ and $2 \leq m \leq N$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[(G_t^N - L)_+^m\big] &\leq \mathbb{E}\Big[ -2\lambda_0 pm(G_t^N - L)_+^{m-1}G_t^N + C_0 pm(G_t^N - L)_+^{m-1} \\ &\qquad + \frac{p^2 m(m-1)}{2N^2}(G_t^N - L)_+^{m-2}\sum_{i=1}^{N}(R_t^{i,N})^{2p-2}|2V_t^{i,N} + \eta X_t^{i,N}|^2\Big] \\ &\leq \mathbb{E}\Big[ -2\lambda_0 pm(G_t^N - L)_+^{m-1}G_t^N + C_0 pm(G_t^N - L)_+^{m-1} + \frac{C_0 pm^2}{N}(G_t^N - L)_+^{m-2}G_t^N\Big],\end{aligned}$$

hence, decomposing $G_t^N = G_t^N - L + L$, we get for $L \geq C_0/\lambda_0$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[(G_t^N - L)_+^m\big] \leq \mathbb{E}\Big[ -2\lambda_0 pm(G_t^N - L)_+^m + \frac{C_0 Lpm^2}{N}(G_t^N - L)_+^{m-2}\Big].$$

By Young's inequality, this entails

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[(G_t^N - L)_+^m\big] \leq -\lambda_0 pm\mathbb{E}[(G_t^N - L)_+^m] + 2\lambda_0 p\Big(\frac{C_0 Lm}{\lambda_0 N}\Big)^{\frac{m}{2}},$$

and thus, by Grönwall's inequality,

$$\mathbb{E}\big[(G_t^N - L)_+^m\big] \leq \Big(\frac{C_0 L m}{\lambda_0 N}\Big)^{\frac{m}{2}} + e^{-\lambda_0 pmt}\,\mathbb{E}\big[(G_0^N - L)_+^m\big].$$

Noting that

$$\mathbb{E}\big[(G_0^N - L)_+^m\big] \leq \mathbb{E}\Big[\Big(C_0 \int_{\mathbb{X}} \langle z\rangle^{2p}\mu_0^N - L\Big)_+^m\Big] = C_0^m\,\mathbb{E}\Big[\Big(\int_{\mathbb{X}} \langle z\rangle^{2p}(\mu_0^N - \mu_\circ) + Q_{2p}(\mu_\circ) - C_0^{-1}L\Big)_+^m\Big],$$

and further imposing $L \geq C_0 Q_{2p}(\mu_\circ)$, we conclude

$$\mathbb{E}\big[(G_t^N - L)_+^m\big] \leq \Big(\frac{C_0 L m}{\lambda_0 N}\Big)^{\frac{m}{2}} + C_0^m e^{-\lambda_0 pmt}\,\mathbb{E}\Big[\Big|\int_{\mathbb{X}} \langle z\rangle^{2p}(\mu_0^N - \mu_\circ)\Big|^m\Big],$$

and the conclusion (2.2) follows.

*Step* 2. Proof of (2.3).
Using the same notation as in Step 1, and starting from (2.5), we find by Itô's formula, for any $r > 0$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[(R_t^{1,N})^r\big] \leq -2\lambda_0 r\mathbb{E}\big[(R_t^{1,N})^r\big] + C_0 r\mathbb{E}\big[(R_t^{1,N})^{r-1}\big] + \tfrac{1}{2}r(r-1)\mathbb{E}\big[(R_t^{1,N})^{r-2}|2V_t^{1,N}+\eta X_t^{1,N}|^2\big]. \quad (2.6)$$

In case $r \geq 1$, using $|2V_t^{1,N} + \eta X_t^{1,N}|^2 \lesssim R_t^{1,N}$ and Young's inequality, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[(R_t^{1,N})^r\big] \leq -2\lambda_0 r\mathbb{E}\big[(R_t^{1,N})^r\big] + C_0 r^2\mathbb{E}\big[(R_t^{1,N})^{r-1}\big] \leq -\lambda_0 r\mathbb{E}\big[(R_t^{1,N})^r\big] + \lambda_0\Big(\frac{C_0 r}{\lambda_0}\Big)^r.$$

Hence, by Grönwall's inequality, recalling $r \geq 1$

$$\mathbb{E}\big[(R_t^{1,N})^r\big] \leq \Big(\frac{C_0 r}{\lambda_0}\Big)^r + e^{-\lambda_0 rt}\,\mathbb{E}\big[(R_0^{1,N})^r\big],$$

and the conclusion (2.3) then follows (up to renaming the constants), using

$$\mathbb{E}\big[(R_t^{1,N})^r\big] \simeq_r \mathbb{E}\big[Q_{2r}(\mu_t^N)\big] = \mathbb{E}\Big[\int_{\mathbb{X}} \langle z\rangle^{2r}\mu_t^N\Big] = Q_{2r}(F_t^{N,1}).$$

In case $0 < r \leq 1$, rather using $r(r-1) \leq 0$ and $(R_t^{1,N})^{r-1} \leq 1$ in (2.6), we find

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\big[(R_t^{1,N})^r\big] \leq -2\lambda_0 r\mathbb{E}\big[(R_t^{1,N})^r\big] + C_0 r,$$

and the conclusion follows similarly. $\qquad\square$

With the above two preliminary lemmas at hand, we may now turn to the proof of the following moment estimates on the mean-field approximation error, which we deduce from a suitable application of Itô's formula together with a Grönwall argument. The proof can be compared with some related computations in [6].

**Proposition 2.3.** *Consider the following quantities measuring the mean-field approximation error, for any $1 < q \leq 2$, $0 < p \leq 1$, $m \geq 1$, and $t \geq 0$,*

$$A_{p,q}^{N,m}(t) := \sup\Big\{\mathbb{E}\Big[\Big|\int_{\mathbb{X}} \varphi(\mu_t^N - \mu_t)\Big|^m\Big] \ : \ \varphi \in C_c^\infty(\mathbb{X}), \ \|\langle z\rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})} \leq 1\Big\}. \quad (2.7)$$

*There exists $\lambda_0 > 0$ (only depending on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$) such that the following holds: given $1 < q \leq 2$ and $0 < p \leq 1$ with $pq' \gg 1$ large enough (only depending on $d, \beta, a$), there exist $\kappa_0, C_0 > 0$ (further depending on $p, q$, $Q_2(\mu_\circ)$, and $\|W\|_{W^{d+3,\infty}\cap H^s(\mathbb{R}^d)}$ for $s > \frac{d}{2} + 3$) such that given $\kappa \in [0, \kappa_0]$ we have for all $2 \leq m \leq N$ and $t \geq 0$,*

$$A_{p,q}^{N,m}(t) \leq \Big(\frac{C_0 m}{N}\Big)^{\frac{m}{2}} + C_0^m e^{-\lambda_0 pmt} A_{3p,q}^{N,m}(0).$$

*Proof.* Let $\lambda_0$ stand for the minimum between the corresponding exponents in Lemmas 2.1 and 2.2. Let $1 < q \le 2$ and $0 < p \le 1$ be fixed with $pq' \gg_{\beta,a} 1$. Instead of controlling directly the evolution of $A_{p,q}^{N,m}(t)$, let us consider

$$D_{p,q}^{N,m}(t) := \sup\left\{ e^{\gamma m(t-s)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{s,t}\varphi)(\mu_s^N - \mu_s)\right|^m\right] \ : \ 0 \le s \le t, \ \varphi \in C_c^\infty(\mathbb{X}), \ \|\langle z \rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})} \le 1 \right\},$$

for some exponent $\gamma > 0$ to be chosen appropriately later, where we recall that $\{U_{s,t}\}_{0 \le s \le t}$ is the backward linearized mean-field evolution defined in (2.1). Given a test function $\varphi \in C_c^\infty(\overline{\mathbb{X}})$ with $\|\langle z \rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})} \le 1$, we find by Itô's formula, for all $m \ge 2$ and $0 \le s \le t$,

$$e^{\gamma m(t-s)}\mathbb{E}\left[\left|\int_{\mathbb{X}}(U_{s,t}\varphi)(\mu_s^N - \mu_s)\right|^m\right] \le e^{\gamma mt}\mathbb{E}\left[\left|\int_{\mathbb{X}}(U_{0,t}\varphi)(\mu_0^N - \mu_\circ)\right|^m\right]$$
$$- \gamma m \int_0^s e^{\gamma m(t-\tau)}\mathbb{E}\left[\left|\int_{\mathbb{X}}(U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^m\right]\mathrm{d}\tau$$
$$+ \kappa m \int_0^s e^{\gamma m(t-\tau)}\mathbb{E}\left[\left|\int_{\mathbb{X}}(U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-1}\left|\int_{\mathbb{X}}\mu_\tau^N \nabla_v(U_{\tau,t}\varphi)\cdot\nabla W * (\mu_\tau^N - \mu_\tau)\right|\right]\mathrm{d}\tau$$
$$+ \frac{m(m-1)}{N}\int_0^s e^{\gamma m(t-\tau)}\mathbb{E}\left[\left|\int_{\mathbb{X}}(U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-2}\int_{\mathbb{X}}|\nabla_v(U_{\tau,t}\varphi)|^2 \mu_\tau^N\right]\mathrm{d}\tau. \quad (2.8)$$

We start by examining the penultimate right-hand side term. Similarly as e.g. in [6], it is convenient to decompose $\nabla W * (\mu_s^N - \mu_s)$ as a superposition of Fourier modes,

$$\int_{\mathbb{X}}\mu_\tau^N \nabla_v(U_{\tau,t}\varphi)\cdot\nabla W * (\mu_\tau^N - \mu_\tau) = C\int_{\mathbb{R}^d} i\xi\widehat{W}(\xi)\cdot\left(\int_{\mathbb{X}}\nabla_v(U_{\tau,t}\varphi)e_\xi\mu_\tau^N\right)\left(\int_{\mathbb{X}}e_{-\xi}(\mu_\tau^N - \mu_\tau)\right)\mathrm{d}\xi,$$

where we have set $e_\xi(x) := e^{ix\cdot\xi}$. By the ergodic estimates of Lemma 2.1, together with the Sobolev embedding with $\ell > 2d/q'$, we find

$$\|\langle z \rangle^{-p}\nabla_v(U_{\tau,t}\varphi)e_\xi\|_{L^\infty(\mathbb{X})} \lesssim \|\langle z \rangle^{-p}\nabla_v(U_{\tau,t}\varphi)\|_{W^{\ell,q'}(\mathbb{X})} \lesssim e^{-\lambda_0 p(t-\tau)}\|\langle z \rangle^{-p}\varphi\|_{W^{\ell+1,q'}(\mathbb{X})}. \quad (2.9)$$

Choosing $q' \ge pq' \gg 1$ large enough, we can choose $\ell > 2d/q'$ with $\ell + 1 \le 2$. Hence, recalling the choice of $\varphi$ with $\|\langle z \rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})} \le 1$, we get

$$\left|\int_{\mathbb{X}}\mu_\tau^N \nabla_v(U_{\tau,t}\varphi)\cdot\nabla W * (\mu_\tau^N - \mu_\tau)\right| \lesssim e^{-\lambda_0 p(t-\tau)}Q_p(\mu_\tau^N)\int_{\mathbb{R}^d}|\xi\widehat{W}(\xi)|\left|\int_{\mathbb{X}}e_{-\xi}(\mu_\tau^N - \mu_\tau)\right|\mathrm{d}\xi,$$

where the multiplicative constant only depends on $d, \beta, p, q, a, Q_2(\mu_\circ)$, and $\|W\|_{W^{d+3,\infty}(\mathbb{R}^d)}$. Given a parameter $L \ge 1$ to be chosen appropriately later, decomposing $Q_p(\mu_\tau^N) \le L + (Q_p(\mu_\tau^N) - L)_+$, noting that $|\int_{\mathbb{X}}e_{-\xi}(\mu_\tau^N - \mu_\tau)| \le 2$ and that the assumptions on $W$ ensure $\int_{\mathbb{R}^d}|\xi\widehat{W}(\xi)|\mathrm{d}\xi \lesssim 1$, we can bound

$$\left|\int_{\mathbb{X}}\mu_\tau^N \nabla_v(U_{\tau,t}\varphi)\cdot\nabla W * (\mu_\tau^N - \mu_\tau)\right| \lesssim Le^{-\lambda_0 p(t-\tau)}\int_{\mathbb{R}^d}|\xi\widehat{W}(\xi)|\left|\int_{\mathbb{X}}e_{-\xi}(\mu_\tau^N - \mu_\tau)\right|\mathrm{d}\xi$$
$$+ e^{-\lambda_0 p(t-\tau)}(Q_p(\mu_\tau^N) - L)_+.$$

Now using this to estimate the penultimate term in (2.8), we find

$$\int_0^s e^{\gamma m(t-\tau)}\mathbb{E}\left[\left|\int_{\mathbb{X}}(U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-1}\left|\int_{\mathbb{X}}\mu_\tau^N \nabla_v(U_{\tau,t}\varphi)\cdot\nabla W * (\mu_\tau^N - \mu_\tau)\right|\right]\mathrm{d}\tau$$
$$\lesssim L\int_{\mathbb{R}^d}|\xi\widehat{W}(\xi)|\int_0^s e^{\gamma m(t-\tau)}e^{-\lambda_0 p(t-\tau)}\mathbb{E}\left[\left|\int_{\mathbb{X}}e_{-\xi}(\mu_\tau^N - \mu_\tau)\right|\left|\int_{\mathbb{X}}(U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-1}\right]\mathrm{d}\tau\mathrm{d}\xi$$
$$+ \int_0^s e^{\gamma m(t-\tau)}e^{-\lambda_0 p(t-\tau)}\mathbb{E}\left[(Q_p(\mu_\tau^N) - L)_+\left|\int_{\mathbb{X}}(U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-1}\right]\mathrm{d}\tau. \quad (2.10)$$

Let us start by examining the first right-hand side term in this estimate. Noting that the definition of $D_{p,q}^{N,m}(t)$ entails in particular

$$\mathbb{E}\left[\left|\int_{\mathbb{X}} e_{-\xi}(\mu_\tau^N - \mu_\tau)\right|^m\right]^{\frac{1}{m}} \le \|\langle z \rangle^{-p} e_{-\xi}\|_{W^{2,q'}(\mathbb{X})} D_{p,q}^{N,m}(\tau)^{\frac{1}{m}} \lesssim \langle \xi \rangle^2 D_{p,q}^{N,m}(\tau)^{\frac{1}{m}},$$

and noting that the assumptions on $W$ allow to estimate as follows the remaining integral with respect to $\xi$, for some $\delta > 0$,

$$\int_{\mathbb{R}^d} \langle \xi \rangle^2 |\xi \widehat{W}(\xi)| d\xi \lesssim_\delta \left(\int_{\mathbb{R}^d} \langle \xi \rangle^{2(\frac{d+\delta}{2}+3)} |\widehat{W}(\xi)|^2 d\xi\right)^{\frac{1}{2}} \simeq \|W\|_{H^{\frac{d+\delta}{2}+3}(\mathbb{R}^d)} < \infty,$$

we can use Hölder's inequality to estimate

$$L\int_{\mathbb{R}^d} |\xi \widehat{W}(\xi)| \int_0^s e^{\gamma m(t-\tau)} e^{-\lambda_0 p(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} e_{-\xi}(\mu_\tau^N - \mu_\tau)\right| \left|\int_{\mathbb{X}} (U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-1}\right] d\tau d\xi$$

$$\lesssim L\int_0^s e^{\gamma m(t-\tau)} e^{-\lambda_0 p(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^m\right]^{1-\frac{1}{m}} D_{p,q}^{N,m}(\tau)^{\frac{1}{m}} d\tau.$$

Inserting this into (2.10) and further using Young's inequality, we get

$$\int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-1} \left|\int_{\mathbb{X}} \mu_\tau^N \nabla_v(U_{\tau,t}\varphi) \cdot \nabla W * (\mu_\tau^N - \mu_\tau)\right|\right] d\tau$$

$$\lesssim L\int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^m\right] d\tau$$

$$+ \frac{L}{m}\int_0^s e^{(\gamma-\lambda_0 p)m(t-\tau)} D_{p,q}^{N,m}(\tau) d\tau + \frac{1}{m}\int_0^s e^{(\gamma-\lambda_0 p)m(t-\tau)} \mathbb{E}\left[(Q_p(\mu_\tau^N) - L)_+^m\right] d\tau. \quad (2.11)$$

We turn to the last term in (2.8). Applying again Young's inequality, we can bound

$$\frac{m(m-1)}{N}\int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-2} \int_{\mathbb{X}} |\nabla_v(U_{\tau,t}\varphi)|^2 \mu_\tau^N\right] d\tau$$

$$\le \frac{\gamma m}{2}\int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} U_{\tau,t}\varphi(\mu_\tau^N - \mu_\tau)\right|^m\right] d\tau + \gamma\left(\frac{2m}{\gamma N}\right)^{\frac{m}{2}} \int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left(\int_{\mathbb{X}} |\nabla_v(U_{\tau,t}\varphi)|^2 \mu_\tau^N\right)^{\frac{m}{2}}\right] d\tau,$$

and thus, using again the ergodic estimates of Lemma 2.1 as in (2.9),

$$\frac{m(m-1)}{N}\int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^{m-2} \int_{\mathbb{X}} |\nabla_v(U_{\tau,t}\varphi)|^2 \mu_\tau^N\right] d\tau$$

$$\le \frac{\gamma m}{2}\int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} U_{\tau,t}\varphi(\mu_\tau^N - \mu_\tau)\right|^m\right] d\tau + \gamma\left(\frac{Cm}{\gamma N}\right)^{\frac{m}{2}} \int_0^s e^{(\gamma-\lambda_0 p)m(t-\tau)} \mathbb{E}\left[Q_{2p}(\mu_\tau^N)^{\frac{m}{2}}\right] d\tau.$$

Inserting this together with (2.11) into (2.8), we are led to

$$e^{\gamma m(t-s)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{s,t}\varphi)(\mu_s^N - \mu_s)\right|^m\right] \le e^{\gamma mt} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{0,t}\varphi)(\mu_0^N - \mu_\circ)\right|^m\right]$$

$$+ (C\kappa L - \tfrac{1}{2}\gamma)m \int_0^s e^{\gamma m(t-\tau)} \mathbb{E}\left[\left|\int_{\mathbb{X}} (U_{\tau,t}\varphi)(\mu_\tau^N - \mu_\tau)\right|^m\right] d\tau$$

$$+ C\kappa L \int_0^s e^{(\gamma-\lambda_0 p)m(t-\tau)} D_{p,q}^{N,m}(\tau) d\tau + C\kappa \int_0^s e^{(\gamma-\lambda_0 p)m(t-\tau)} \mathbb{E}\left[(Q_p(\mu_\tau^N) - L)_+^m\right] d\tau$$

$$+ \gamma\left(\frac{Cm}{\gamma N}\right)^{\frac{m}{2}} \int_0^s e^{(\gamma-\lambda_0 p)m(t-\tau)} \mathbb{E}\left[Q_{2p}(\mu_\tau^N)^{\frac{m}{2}}\right] d\tau.$$

Note that for the first right-hand side term the ergodic estimates of Lemma 2.1 and the definition of $D_{p,q}^{N,m}(0)$ yield

$$\mathbb{E}\Big[\Big|\int_{\mathbb{X}}(U_{0,t}\varphi)(\mu_0^N-\mu_\circ)\Big|^m\Big] \leq \Big\|\langle z\rangle^{-p}\Big(U_{0,t}\varphi - \int_{\mathbb{X}}\varphi M\Big)\Big\|_{W^{2,q'}(\mathbb{X})}^m D_{p,q}^{N,m}(0) \leq C^m e^{-\lambda_0 pmt}D_{p,q}^{N,m}(0).$$

Provided that $\gamma \geq 2C\kappa L$, taking the supremum over $\varphi$ and $0 \leq s \leq t$, we deduce

$$D_{p,q}^{N,m}(t) \leq C^m e^{(\gamma-\lambda_0 p)mt}D_{p,q}^{N,m}(0) + C\kappa L\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}D_{p,q}^{N,m}(\tau)\,\mathrm{d}\tau$$

$$+ C\kappa\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathbb{E}\big[(Q_{2p}(\mu_\tau^N)-L)_+^m\big]\mathrm{d}\tau$$

$$+ \gamma\Big(\frac{Cm}{\gamma N}\Big)^{\frac{m}{2}}\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathbb{E}\big[Q_{2p}(\mu_\tau^N)^{\frac{m}{2}}\big]\mathrm{d}\tau. \quad (2.12)$$

Noting that for $pq' \gg 1$ we have

$$\mathbb{E}\Big[\Big|\int_{\mathbb{X}}\langle z\rangle^{2p}(\mu_0^N-\mu_\circ)\Big|^m\Big] \leq C^m D_{3p,q}^{N,m}(0),$$

and choosing $L \gg 1$ large enough (only depending on $d, \beta, a, Q_2(\mu_\circ), \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$), the concentration bound of Lemma 2.2 yields, for $2 \leq m \leq N$,

$$\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathbb{E}\big[(Q_{2p}(\mu_\tau^N)-L)_+^m\big]\mathrm{d}\tau$$

$$\leq \Big(\frac{CLm}{N}\Big)^{\frac{m}{2}}\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathrm{d}\tau + C^m D_{3p,q}^{N,m}(0)\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}e^{-\lambda_0 pm\tau}\mathrm{d}\tau,$$

and thus, provided that $\gamma \leq \frac{1}{2}\lambda_0 p$,

$$\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathbb{E}\big[(Q_{2p}(\mu_\tau^N)-L)_+^m\big]\mathrm{d}\tau \leq \Big(\frac{CLm}{N}\Big)^{\frac{m}{2}} + C^m e^{(\gamma-\lambda_0 p)mt}D_{3p,q}^{N,m}(0).$$

Similarly, we can also bound

$$\Big(\frac{Cm}{N}\Big)^{\frac{m}{2}}\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathbb{E}\big[Q_{2p}(\mu_\tau^N)^{\frac{m}{2}}\big]\mathrm{d}\tau$$

$$\leq \Big(\frac{CLm}{N}\Big)^{\frac{m}{2}}\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathrm{d}\tau + C^m\int_0^t e^{(\gamma-\lambda_0 p)m(t-\tau)}\mathbb{E}\big[(Q_{2p}(\mu_\tau^N)-L)_+^m\big]\mathrm{d}\tau$$

$$\leq \Big(\frac{CLm}{N}\Big)^{\frac{m}{2}} + C^m e^{(\gamma-\lambda_0 p)mt}D_{3p,q}^{N,m}(0).$$

Inserting these bounds into (2.12), choosing $\gamma := \frac{1}{2}\lambda_0 p$, and assuming that $\kappa$ is small enough in the sense of $2C\kappa L \leq \frac{1}{2}\lambda_0 p$, we find

$$D_{p,q}^{N,m}(t) \leq \Big(\frac{CLm}{N}\Big)^{\frac{m}{2}} + C^m e^{-\frac{1}{2}\lambda_0 pmt}D_{3p,q}^{N,m}(0) + C\kappa L\int_0^t e^{-\frac{1}{2}\lambda_0 pm(t-\tau)}D_{p,q}^{N,m}(\tau)\,\mathrm{d}\tau,$$

and thus, by Grönwall's inequality,

$$D_{p,q}^{N,m}(t) \leq \Big(\frac{CLm}{N}\Big)^{\frac{m}{2}} + C^m e^{-\frac{1}{2}\lambda_0 pmt}D_{3p,q}^{N,m}(0)$$

$$+ C\kappa Le^{-\frac{1}{2}\lambda_0 pmt}\int_0^t e^{C\kappa L(t-\tau)}\Big(e^{\frac{1}{2}\lambda_0 pm\tau}\Big(\frac{CLm}{N}\Big)^{\frac{m}{2}} + C^m D_{3p,q}^{N,m}(0)\Big)\mathrm{d}\tau.$$

As the smallness requirement on $\kappa$ precisely ensures $C\kappa L \leq \frac{1}{4}\lambda_0 p$, we get after computing the remaining time integrals,

$$D_{p,q}^{N,m}(t) \leq \Big(\frac{CLm}{N}\Big)^{\frac{m}{2}} + C^m e^{-\frac{1}{4}\lambda_0 pmt}D_{3p,q}^{N,m}(0).$$

As by definition we have $A_{p,q}^{N,m}(t) \le D_{p,q}^{N,m}(t)$ and $A_{p,q}^{N,m}(0) = D_{p,q}^{N,m}(0)$, the conclusion follows (up to renaming $\lambda_0$). $\qquad\square$

## 3. Creation of chaos

This section is devoted to the proof of Theorem 1.1 using the preliminary results above. Our argument shows how the estimates of Proposition 2.3 can be turned into suboptimal mean-field estimates for marginals, which in turn can be made optimal in $N$ after combination with direct estimates on the BBGKY hierarchy. A similar idea will be used again for the proof of Theorem 1.2 in the next section.

*Proof of Theorem 1.1.* We split the proof into three steps. Let $\lambda_0, p, q, \kappa$ be as in Proposition 2.3, and assume $p \le \frac{1}{6}$. In the sequel of this proof, all multiplicative constants are allowed to depend on $d, \beta, a, p, q, Q_2(\mu_\circ), Q_2(F_\circ^{N,1})$, and $\|W\|_{W^{d+3,\infty} \cap H^s(\mathbb{R}^d)}$ for some $s > \frac{d}{2}+5$. A subscript '$m$' to constants is used to indicate further dependence on $m$ and on $Q_{mp}(F_\circ^{N,1})$.

**Step 1:** Reformulation of Proposition 2.3: proof of the suboptimal mean-field estimate

$$\|F_t^{N,1} - \mu_t\|_{W_p^{-2,q}(\mathbb{X})}^2 + \|G_t^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}}$$
$$\lesssim N^{-1} + e^{-2\lambda_0 pt}\Big(\|F_\circ^{N,1} - \mu_\circ\|_{W_{3p}^{-2,q}(\mathbb{X})}^2 + \|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}}\Big). \quad (3.1)$$

Noting that

$$\mathbb{E}\Big[\Big(\int_\mathbb{X} \varphi\,(\mu_t^N - \mu_t)\Big)\Big] = \int_\mathbb{X} \varphi(F_t^{N,1} - \mu_t),$$

and that

$$A_{p,q}^{N,2}(t) \ge \sup\Big\{\Big(\int_\mathbb{X} \varphi(F_t^{N,1} - \mu_t)\Big)^2 \ : \ \varphi \in C_c^\infty(\mathbb{X}),\ \|\langle\cdot\rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})} \le 1\Big\}$$
$$= \|F_t^{N,1} - \mu_t\|_{W_p^{-2,q}(\mathbb{X})}^2,$$

Proposition 2.3 implies for $m = 2$,

$$\|F_t^{N,1} - \mu_t\|_{W_p^{-2,q}(\mathbb{X})}^2 \lesssim N^{-1} + e^{-2\lambda_0 pt} A_{3p,q}^{N,2}(0). \quad (3.2)$$

Also note that

$$\mathbb{E}\Big[\Big(\int_\mathbb{X} \varphi\,(\mu_t^N - \mu_t)\Big)^2\Big] = \Big(\int_\mathbb{X} \varphi(F_t^{N,1} - \mu_t)\Big)^2 + \int_\mathbb{X} \varphi^{\otimes 2} G_t^{N,2} + \frac{1}{N}\int_\mathbb{X} \varphi^2 F_t^{N,1} - \frac{1}{N}\int_\mathbb{X} \varphi^{\otimes 2} F_t^{N,2},$$

and thus, taking the supremum over $\varphi$, and using the Sobolev embedding together with the moment bounds of Lemma 2.2, which yields in particular

$$\|F_t^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}} \lesssim \int_{\mathbb{X}^2} \langle z\rangle^p \langle z_*\rangle^p F_t^{N,2}(z, z_*)\,\mathrm{d}z\mathrm{d}z_* \le \int_\mathbb{X} \langle z\rangle^{2p} F_t^{N,1} = Q_{2p}(F_t^{N,1}), \quad (3.3)$$

we get

$$\Big|A_{p,q}^{N,2}(t) - \|G_t^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}}\Big| \le \|F_t^{N,1} - \mu_t\|_{W_p^{-2,q}(\mathbb{X})}^2 + CN^{-1} Q_{2p}(F_t^{N,1}).$$

Appealing to Lemma 2.2 in form of $Q_{2p}(F_t^{N,1}) \lesssim 1 + Q_{2p}(F_\circ^{N,1}) \lesssim 1$, this means

$$\Big|A_{p,q}^{N,2}(t) - \|G_t^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}}\Big| \lesssim \|F_t^{N,1} - \mu_t\|_{W_p^{-2,q}(\mathbb{X})}^2 + CN^{-1}. \quad (3.4)$$

Combining with the result of Proposition 2.3 to get a control on $G_t^{N,2}$, we obtain

$$\|G_t^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}} \lesssim \|F_t^{N,1} - \mu_t\|_{W_p^{-2,q}(\mathbb{X})}^2 + N^{-1} + e^{-2\lambda_0 pt} A_{3p,q}^{N,2}(0),$$

which thus yields, together with (3.2),

$$\|F_t^{N,1} - \mu_t\|_{W_p^{-2,q}(\mathbb{X})}^2 + \|G_t^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}} \lesssim N^{-1} + e^{-2\lambda_0 pt} A_{3p,q}^{N,2}(0). \quad (3.5)$$

Now applying (3.4) at $t = 0$ with $p$ replaced by $3p$, we find

$$A_{3p,q}^{N,2}(0) \leq \|F_\circ^{N,1} - \mu_\circ\|_{W_{3p}^{-2,q}(\mathbb{X})}^2 + \|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}} + CN^{-1},$$

and the claim (3.1) follows.

**Step 2:** Proof of the improved mean-field bound

$$\|F_t^{N,1} - \mu_t\|_{W_p^{-3,q}(\mathbb{X})}$$
$$\lesssim N^{-1} + e^{-\frac{1}{2}\lambda_0 pt}\Big(\|F_\circ^{N,1} - \mu_\circ\|_{W_p^{-2,q}(\mathbb{X})} + \|F_\circ^{N,1} - \mu_\circ\|_{W_{3p}^{-2,q}(\mathbb{X})}^2 + \|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}}\Big)$$
$$\lesssim N^{-1} + e^{-\frac{1}{2}\lambda_0 pt}\Big(\|F_\circ^{N,1} - \mu_\circ\|_{W_{3p}^{-2,q}(\mathbb{X})} + \|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}}\Big). \tag{3.6}$$

Integrating the Liouville equation (1.3) with respect to its last $N - 1$ coordinates, and decomposing the second marginal as $F^{N,2} = F^{N,1} \otimes F^{N,1} + G^{N,2}$, we find that the first marginal $F^{N,1}$ satisfies the following BBGKY equation,

$$\partial_t(F^{N,1} - \mu) = R_\mu(F^{N,1} - \mu) + \kappa \nabla W * (F^{N,1} - \mu) \cdot \nabla_v F^{N,1}$$
$$+ \kappa \int_{\mathbb{X}} \nabla W(\cdot - x_*) \cdot \nabla_v(G^{N,2} - \tfrac{1}{N} F^{N,2})(\cdot, z_*) \, \mathrm{d}z_*,$$

where $R_\mu$ stands for the following (truncated) linearized mean-field operator (linearized at the mean-field solution $\mu$),

$$R_\mu h := \tfrac{1}{2}\mathrm{div}_v((\nabla_v + \beta v)h) - v \cdot \nabla_x h + (\nabla A + \kappa \nabla W * \mu) \cdot \nabla_v h.$$

Given $\varphi \in C_c^\infty(\mathbb{X})$, recalling the definition (2.1) of the backward linearized evolution $\{U_{s,t}\}_{0 \leq s \leq t}$, we may then compute

$$\int_{\mathbb{X}} \varphi(F_t^{N,1} - \mu_t) = \int_{\mathbb{X}} (U_{0,t}\varphi)(F_\circ^{N,1} - \mu_\circ) - \kappa \int_0^t \Big(\int_{\mathbb{X}} \nabla_v(U_{s,t}\varphi) \cdot \nabla W * (F_s^{N,1} - \mu_s) F_s^{N,1}\Big) \mathrm{d}s$$
$$- \kappa \int_0^t \Big(\int_{\mathbb{X}^2} \nabla_v(U_{s,t}\varphi)(z) \cdot \nabla W(x - x_*)(G_s^{N,2} - \tfrac{1}{N} F_s^{N,2})(z, z_*) \, \mathrm{d}z\mathrm{d}z_*\Big) \mathrm{d}s. \tag{3.7}$$

Let us examine the three right-hand side terms separately. Using the ergodic estimates of Lemma 2.1, the first one is bounded by

$$\Big|\int_{\mathbb{X}} (U_{0,t}\varphi)(F_\circ^{N,1} - \mu_\circ)\Big| \lesssim e^{-\lambda_0 pt}\|\langle z\rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})}\|F_\circ^{N,1} - \mu_\circ\|_{W_p^{-2,q}(\mathbb{X})}.$$

For the second term, using again the ergodic estimates of Lemma 2.1 along with the Sobolev embedding, we can bound

$$\Big|\int_{\mathbb{X}} \nabla_v(U_{s,t}\varphi) \cdot \nabla W * (F_s^{N,1} - \mu_s) F_s^{N,1}\Big|$$
$$\lesssim Q_p(F_s^{N,1}) \|\langle z\rangle^{-p}\nabla_v(U_{s,t}\varphi)\|_{\mathrm{L}^\infty(\mathbb{X})} \sup_{x \in \mathbb{R}^d} \Big|\int_{\mathbb{X}} \nabla W(x - \cdot)(F_s^{N,1} - \mu_s)\Big|$$
$$\lesssim Q_p(F_s^{N,1}) e^{-\lambda_0 p(t-s)} \|\langle z\rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})} \|F_s^{N,1} - \mu_s\|_{W_p^{-3,q}(\mathbb{X})},$$

and thus, further using the moment bounds of Lemma 2.2,

$$\Big|\int_{\mathbb{X}} \nabla_v(U_{s,t}\varphi) \cdot \nabla W * (F_s^{N,1} - \mu_s) F_s^{N,1}\Big| \lesssim e^{-\lambda_0 p(t-s)} \|\langle z\rangle^{-p}\varphi\|_{W^{2,q'}(\mathbb{X})} \|F_s^{N,1} - \mu_s\|_{W_p^{-3,q}(\mathbb{X})}.$$

For the last term in (3.7), we aim to appeal to the bound (3.1) of Step 1 for $G^{N,2}$. Note, however, that the map $(z, z_*) \mapsto \nabla_v(U_{s,t}\varphi)(z) \cdot \nabla W(x - x_*)$ is not an admissible test function for such estimates in $W_p^{-2,q}(\mathbb{X})^{\otimes 2}$, cf. (1.13). To remedy this, we decompose $W$ as a superposition of Fourier modes, which indeed allows to reduce to a tensorized test function. More precisely, by polarization, using that

$G_s^{N,2} - \frac{1}{N}F_s^{N,2}$ is symmetric in its two variables, and recalling the short-hand notation $e_\xi(x) = e^{ix\cdot\xi}$, we can bound

$$\left|\int_{\mathbb{X}^2} \nabla_v(U_{s,t}\varphi)(z) \cdot \nabla W(x - x_*)\,(G_s^{N,2} - \tfrac{1}{N}F_s^{N,2})(z, z_*)\,\mathrm{d}z\mathrm{d}z_*\right|$$

$$\leq \int_{\mathbb{R}^d} |\xi\widehat{W}(\xi)| \left|\int_{\mathbb{X}^2} e_\xi(x_*)e_{-\xi}(x)\nabla_v(U_{s,t}\varphi)(z)\,(G_s^{N,2} - \tfrac{1}{N}F_s^{N,2})(z, z_*)\,\mathrm{d}z\mathrm{d}z_*\right|\mathrm{d}\xi \qquad (3.8)$$

$$\leq 4\|G_s^{N,2} - \tfrac{1}{N}F_s^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}} \int_{\mathbb{R}^d} |\xi\widehat{W}(\xi)|\|\langle z\rangle^{-p}e_\xi\|_{W^{2,q'}(\mathbb{X})}\|\langle z\rangle^{-p}e_\xi\nabla_v(U_{s,t}\varphi)\|_{W^{2,q'}(\mathbb{X})}\,\mathrm{d}\xi.$$

Noting that

$$\|\langle z\rangle^{-p}e_\xi\|_{W^{2,q'}(\mathbb{X})} \lesssim \langle\xi\rangle^2, \qquad \|\langle z\rangle^{-p}e_\xi\nabla_v(U_{s,t}\varphi)\|_{W^{2,q'}(\mathbb{X})} \lesssim \langle\xi\rangle^2 e^{-\lambda_0 p(t-s)}\|\langle z\rangle^{-p}\varphi\|_{W^{3,q'}(\mathbb{X})},$$

we deduce

$$\left|\int_{\mathbb{X}^2} \nabla_v(U_{s,t}\varphi)(z) \cdot \nabla W(x - x_*)\,(G_s^{N,2} - \tfrac{1}{N}F_s^{N,2})(z, z_*)\,\mathrm{d}z\mathrm{d}z_*\right|$$

$$\lesssim e^{-\lambda_0 p(t-s)}\|\langle z\rangle^{-p}\varphi\|_{W^{3,q'}(\mathbb{X})}\|G_s^{N,2} - \tfrac{1}{N}F_s^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}} \int_{\mathbb{R}^d} \langle\xi\rangle^4|\xi\widehat{W}(\xi)|\,\mathrm{d}\xi,$$

where the last factor can be controlled by $\|W\|_{H^s(\mathbb{R}^d)}$ for $s > \frac{d}{2} + 5$. Combining these estimates for the different terms in (3.7), and taking the supremum over $\varphi$, we get

$$\|F_t^{N,1} - \mu_t\|_{W_p^{-3,q}(\mathbb{X})} \lesssim e^{-\lambda_0 pt}\|F_\circ^{N,1} - \mu_\circ\|_{W_p^{-2,q}(\mathbb{X})}$$

$$+ \kappa \int_0^t e^{-\lambda_0 p(t-s)}\|F_s^{N,1} - \mu_s\|_{W_p^{-3,q}(\mathbb{X})}\,\mathrm{d}s + \kappa \int_0^t e^{-\lambda_0 p(t-s)}\|G_s^{N,2} - \tfrac{1}{N}F_s^{N,2}\|_{W_p^{-2,q}(\mathbb{X})^{\otimes 2}}\,\mathrm{d}s.$$

Now inserting the bound (3.1) of Step 1 for $G^{N,2}$, using (3.3) together with $Q_{2p}(F_t^{N,1}) \lesssim 1$, we obtain

$$\|F_t^{N,1} - \mu_t\|_{W_p^{-3,q}(\mathbb{X})} \leq CN^{-1} + C\kappa \int_0^t e^{-\lambda_0 p(t-s)}\|F_s^{N,1} - \mu_s\|_{W_p^{-3,q}(\mathbb{X})}\,\mathrm{d}s$$

$$+ Ce^{-\lambda_0 pt}\left(\|F_\circ^{N,1} - \mu_\circ\|_{W_p^{-2,q}(\mathbb{X})} + \|F_\circ^{N,1} - \mu_\circ\|_{W_{3p}^{-2,q}(\mathbb{X})}^2 + \|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}}\right).$$

Provided that $\kappa$ is chosen small enough so that $C\kappa < \frac{1}{2}\lambda_0 p$, the claim (3.6) follows from Grönwall's inequality.

**Step 3:** Conclusion.
Let $\varphi \in C_c^\infty(\mathbb{X})$ be momentarily fixed with $\|\langle z\rangle^{-p}\varphi\|_{W^{3,q'}(\mathbb{X})} \leq 1$. Moments of the empirical measure can be computed as follows, for all $m \geq 1$,

$$\mathbb{E}\left[\left(\int_{\mathbb{X}} \varphi(\mu_t^N - \mu_t)\right)^m\right] = N^{-m} \sum_{i_1,\dots,i_m=1}^N \mathbb{E}\left[\prod_{l=1}^m \left(\varphi(Z_t^{i_l,N}) - \int_{\mathbb{X}}\varphi\mu_t\right)\right]$$

$$= N^{-m} \sum_{\pi\vdash[\![m]\!]} N(N-1)\dots(N-\sharp\pi+1) \int_{\mathbb{X}^{\sharp\pi}} \left(\bigotimes_{B\in\pi} \left(\varphi - \int_{\mathbb{X}}\varphi\mu_t\right)^{\sharp B}\right) F_t^{N,\sharp\pi},$$

where we use the same notation as in (1.8) for sums over partitions. Note that the moment bounds of Lemma 2.2 together with the Sobolev embedding and the choice of $\varphi$ yield, for all $\pi \vdash [\![m]\!]$, mimicking (3.3),

$$\left|\int_{\mathbb{X}^{\sharp\pi}} \left(\bigotimes_{B\in\pi} \left(\varphi - \int_{\mathbb{X}}\varphi\mu_t\right)^{\sharp B}\right) F_t^{N,\sharp\pi}\right| \lesssim_m \int_{\mathbb{X}} \langle z\rangle^{mp}F_t^{N,1} \leq (Cm)^{\frac{mp}{2}} + C^m e^{-\lambda_0 pmt}Q_{mp}(F_\circ^{N,1}) \lesssim_m 1.$$

Hence, neglecting all $O(N^{-1})$ terms in the above, we get

$$\left| \mathbb{E}\left[ \left( \int_{\mathbb{X}} \varphi(\mu_t^N - \mu_t) \right)^m \right] - \int_{\mathbb{X}^m} \left( \varphi - \int_{\mathbb{X}} \varphi \mu_t \right)^{\otimes m} F_t^{N,m} \right| \lesssim_m N^{-1}.$$

Noting that

$$\int_{\mathbb{X}^m} \left( \varphi - \int_{\mathbb{X}} \varphi \mu_t \right)^{\otimes m} F_t^{N,m} = \sum_{j=0}^m \binom{m}{j} (-1)^{m-j} \left( \int_{\mathbb{X}} \varphi \mu_t \right)^{m-j} \int_{\mathbb{X}^j} \varphi^{\otimes j} F_t^{N,j}$$

$$= \sum_{j=1}^m \binom{m}{j} (-1)^{m-j} \left( \int_{\mathbb{X}} \varphi \mu_t \right)^{m-j} \int_{\mathbb{X}^j} \varphi^{\otimes j} (F_t^{N,j} - \mu_t^{\otimes j}),$$

and focusing on the term corresponding to $j = m$ in the sum, we deduce

$$\left| \mathbb{E}\left[ \left( \int_{\mathbb{X}} \varphi(\mu_t^N - \mu_t) \right)^m \right] - \int_{\mathbb{X}^m} \varphi^{\otimes m}(F_t^{N,m} - \mu_t^{\otimes m}) \right| \lesssim_m N^{-1} + \sum_{j=1}^{m-1} \left| \int_{\mathbb{X}^j} \varphi^{\otimes j}(F_t^{N,j} - \mu_t^{\otimes j}) \right|. \quad (3.9)$$

Taking the supremum over $\varphi \in C_c^\infty(\mathbb{X})$ with $\|\langle z \rangle^{-p} \varphi\|_{W^{3,q'}(\mathbb{X})} \le 1$, we deduce in particular for all $m \ge 1$,

$$\|F_t^{N,m} - \mu_t^{\otimes m}\|_{W_p^{-3,q}(\mathbb{X})^{\otimes m}} \lesssim_m N^{-1} + A_{p,q}^{N,m}(t) + \sum_{j=1}^{m-1} \|F_t^{N,j} - \mu_t^{\otimes j}\|_{W_p^{-3,q}(\mathbb{X})^{\otimes j}}, \quad (3.10)$$

and thus, after a direct iteration,

$$\|F_t^{N,m} - \mu_t^{\otimes m}\|_{W_p^{-3,q}(\mathbb{X})^{\otimes m}} \lesssim_m N^{-1} + \sum_{j=2}^m A_{p,q}^{N,j}(t) + \|F_t^{N,1} - \mu_t\|_{W_p^{-3,q}(\mathbb{X})}.$$

Inserting the result of Proposition 2.3 and the result (3.6) of Step 2, we then get

$$\|F_t^{N,m} - \mu_t^{\otimes m}\|_{W_p^{-3,q}(\mathbb{X})^{\otimes m}} \lesssim_m N^{-1} + \sum_{j=2}^m e^{-\lambda_0 pjt} A_{3p,q}^{N,j}(0)$$

$$+ e^{-\frac{1}{2}\lambda_0 pt} \left( \|F_\circ^{N,1} - \mu_\circ\|_{W_{3p}^{-2,q}(\mathbb{X})} + \|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}} \right). \quad (3.11)$$

From here, it remains to express the $A_{3p,q}^{N,j}(0)$'s in terms of initial correlations. Noting that

$$\mathbb{E}\left[ \left| \int_{\mathbb{X}} \varphi(\mu^N - \mu) \right|^j \right] \le \begin{cases} \mathbb{E}\left[ \left( \int_{\mathbb{X}} \varphi(\mu^N - \mu) \right)^j \right] & : \ j \text{ even}, \\ \mathbb{E}\left[ \left( \int_{\mathbb{X}} \varphi(\mu^N - \mu) \right)^{j-1} \right]^{\frac{1}{2}} \mathbb{E}\left[ \left( \int_{\mathbb{X}} \varphi(\mu^N - \mu) \right)^{j+1} \right]^{\frac{1}{2}} & : \ j \text{ odd}, \end{cases} \quad (3.12)$$

using (3.9) at $t = 0$, and taking the supremum over $\varphi \in C_c^\infty(\mathbb{X})$ with $\|\langle z \rangle^{-3p} \varphi\|_{W^{2,q'}(\mathbb{X})} \le 1$, we find for all $j \ge 2$,

$$A_{3p,q}^{N,j}(0) \lesssim_j N^{-1} + \sum_{l=1}^{\tilde{j}} \|F_\circ^{N,l} - \mu_\circ^{\otimes l}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes l}}, \qquad \tilde{j} := j + \mathbb{1}_{j \text{ odd}}.$$

Inserting this into (3.11), we get

$$\|F_t^{N,m} - \mu_t^{\otimes m}\|_{W_p^{-3,q}(\mathbb{X})^{\otimes m}} \lesssim_m N^{-1} + e^{-\frac{1}{2}\lambda_0 pt} \|G_\circ^{N,2}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes 2}} + \sum_{j=1}^{\tilde{m}} e^{-\frac{1}{2}\lambda_0 pjt} \|F_\circ^{N,j} - \mu_\circ^{\otimes j}\|_{W_{3p}^{-2,q}(\mathbb{X})^{\otimes j}},$$

with $\tilde{m} = m + \mathbb{1}_{m \text{ odd}}$. Decomposing

$$G_\circ^{N,2} = (F_\circ^{N,2} - \mu_\circ^{\otimes 2}) - (F_\circ^{N,1} - \mu_\circ) \otimes F_\circ^{N,1} - \mu_\circ \otimes (F_\circ^{N,1} - \mu_\circ), \quad (3.13)$$

the conclusion (1.14) follows for all $m \ge 1$ (up to renaming $\lambda_0$). $\qquad \square$

## 4. Correlation estimates

This section is devoted to the proof of Theorem 1.2. Similarly as for Theorem 1.1, our argument starts from the estimates of Proposition 2.3: we show how these can be used to deduce suboptimal estimates on correlations, which in turn can be made optimal in $N$ after combination with direct estimates on the BBGKY hierarchy. Note that this time we will need to iterate BBGKY estimates multiple times to achieve optimality. We split the proof into the next three subsections.

While in the introduction we defined correlation functions so as to satisfy the cluster expansion (1.8), we recall that they can be defined more explicitly as polynomial combinations of marginals:

$$
\begin{aligned}
G^{N,1} &= F^{N,1}, \\
G^{N,2} &= F^{N,2} - F^{N,1} \otimes F^{N,1}, \\
G^{N,3} &= \mathrm{sym}\big(F^{N,3} - 3F^{N,2} \otimes F^{N,1} + 2(F^{N,1})^{\otimes 3}\big), \\
G^{N,4} &= \mathrm{sym}\big(F^{N,4} - 4F^{N,3} \otimes F^{N,1} - 3F^{N,2} \otimes F^{N,2} + 12F^{N,2} \otimes (F^{N,1})^{\otimes 2} - 6(F^{N,1})^{\otimes 4}\big),
\end{aligned}
$$

and so on, where the symbol 'sym' stands for the symmetrization of coordinates. More generally, we can write for all $2 \le m \le N$,

$$
G^{N,m}(z_1, \ldots, z_m) = \sum_{\pi \vdash \llbracket m \rrbracket} (\sharp\pi - 1)!(-1)^{\sharp\pi - 1} \prod_{B \in \pi} F^{N,\sharp B}(z_B), \tag{4.1}
$$

where we use a similar notation as in (1.8) for sums over partitions and where $\sharp\pi$ stands for the number of blocks in a partition $\pi$. We also use the following abbreviation for functional spaces: for all $r, p, q, m$,

$$
W_{p,m}^{-r,q} := W_p^{-r,q}(\mathbb{X})^{\otimes m}.
$$

4.1. **From moment mean-field estimates to correlations.** From now on, we let $\mu$ be the weak solution of the mean-field equation (1.4) with initial condition

$$
\mu_\circ := F_\circ^{N,1}.
$$

In this section, we examine how the correlation functions $\{G^{N,m}\}_{2 \le m \le N}$ can be related to the moments of $\mu_t^N - \mu_t$ in terms of the $A_{p,q}^{N,m}$'s defined in (2.7), and we show that Proposition 2.3 implies the following hierarchy of estimates on correlations with suboptimal $N$-dependence ($N^{-m/2}$ instead of $N^{1-m}$, cf. (1.9)).

**Proposition 4.1.** Let $\lambda_0, p, q, \kappa$ be as in Proposition 2.3 with $\mu_\circ = F_\circ^{N,1}$. Then we have for all $2 \le m \le N$, $r \ge 2$, $0 < p_0 \le p$, and $t \ge 0$,

$$
\|G_t^{N,m}\|_{W_{p_0,m}^{-r,q}} \lesssim_m N^{-\frac{m}{2}} + \sum_{\substack{1 \le k \le l \le m-1}} \sum_{\substack{j_1, \ldots, j_k \ge 1 \\ j_1 + \ldots + j_k = l}} N^{l-k-m+1} \prod_{i=1}^k \|G_t^{N,j_i}\|_{W_{mp_0,j_i}^{-r,q}}
$$

$$
+ e^{-\lambda_0 pmt} \times
\begin{cases}
\sum_{\ell=1}^{m/2} \sum_{\ell \le K \le L \le m} \sum_{\substack{j_1, \ldots, j_K \ge 1 \\ j_1 + \ldots j_K = L}} N^{L-K-m+\ell} \prod_{i=1}^K \|G_\circ^{N,j_i}\|_{W_{3mp,j_i}^{-2,q}} & : \ m \ \text{even}, \\[2ex]
\sum_{\ell=1}^m \sum_{\ell \le K \le L \le 2m} \sum_{\substack{1 \le j_1, \ldots, j_K \le m+1 \\ j_1 + \ldots + j_K = L}} \left( N^{L-K-2m+\ell} \prod_{i=1}^K \|G_\circ^{N,j_i}\|_{W_{3mp,j_i}^{-2,q}} \right)^{\frac{1}{2}} & : \ m \ \text{odd},
\end{cases}
$$

where the multiplicative constant only depends on $d, \beta, a, p_0, p, q, r, m, Q_2(F_\circ^{N,1})$, and $\|W\|_{W^{d+3,\infty} \cap H^s(\mathbb{R}^d)}$ for some $s > \frac{d}{2} + 3$.

*Proof.* As a starting point, we recall the standard link between cumulants of the empirical measure and correlation functions. We first used this link in [10, Section 4] and we further included a self-contained statement and proof in [2, Lemma 2.6]. It can be stated as follows: for all $\varphi \in C_c^\infty(\mathbb{X})$, $1 \le m \le N$, and $t \ge 0$,

$$
\left| \kappa^m \Big[ \int_{\mathbb{X}} \varphi \mu_t^N \Big] - \int_{\mathbb{X}^m} \varphi^{\otimes m} G_t^{N,m} \right|
$$

$$\lesssim_m \sum_{\substack{\pi \vdash \llbracket m \rrbracket \\ \sharp\pi < m}} \sum_{\rho \vdash \pi} N^{\sharp\pi - \sharp\rho - m + 1} \bigg| \int_{\mathbb{X}^{\sharp\pi}} \Big( \bigotimes_{B \in \pi} \varphi^{\sharp B} \Big) \Big( \bigotimes_{D \in \rho} G_t^{N, \sharp D}(z_D) \Big) \, \mathrm{d}z_\pi \bigg|, \quad (4.2)$$

where $\kappa^m[\cdot]$ stands for the $m$-th cumulant. For $m \geq 2$, as cumulants are invariant under translation, we can write $\kappa^m[\int_{\mathbb{X}} \varphi \mu_t^N] = \kappa^m[\int_{\mathbb{X}} \varphi(\mu_t^N - \mu_t)]$. By the definition of cumulants in terms of moments, we also recall for any random variable $X$,

$$\kappa^m[X] = \sum_{\pi \vdash \llbracket m \rrbracket} (-1)^{\sharp\pi - 1} (\sharp\pi - 1)! \prod_{B \in \pi} \mathbb{E}[X^{\sharp B}] \lesssim_m \mathbb{E}[|X|^m]. \quad (4.3)$$

Hence, from the above, we deduce in particular for all $\varphi \in C_c^\infty(\mathbb{X})$, $2 \leq m \leq N$, and $t \geq 0$,

$$\bigg| \int_{\mathbb{X}^m} \varphi^{\otimes m} G_t^{N,m} \bigg| \lesssim_m \mathbb{E}\bigg[ \Big| \int_{\mathbb{X}} \varphi(\mu_t^N - \mu_t) \Big|^m \bigg]$$
$$+ \sum_{\substack{\pi \vdash \llbracket m \rrbracket \\ \sharp\pi < m}} \sum_{\rho \vdash \pi} N^{\sharp\pi - \sharp\rho - m + 1} \bigg| \int_{\mathbb{X}^{\sharp\pi}} \Big( \bigotimes_{B \in \pi} \varphi^{\sharp B} \Big) \Big( \bigotimes_{D \in \rho} G_t^{N, \sharp D}(z_D) \Big) \, \mathrm{d}z_\pi \bigg|.$$

For $r \geq 2$, $1 < q \leq 2$ with $q' > d$, and $0 < p_0 \leq p \leq 1$, taking the supremum over the test function $\varphi$, recognizing the definition (2.7) of $A_{p_0,q}^{N,m}(t)$, noting that for $p_0 \leq p$ we have $A_{p_0,q}^{N,m} \leq A_{p,q}^{N,m}$, and also noting that the Sobolev embedding implies

$$\|\langle z \rangle^{-p_0 \ell} \varphi^\ell\|_{W^{r,q'}(\mathbb{X})} \lesssim \|\langle z \rangle^{-p_0} \varphi\|_{W^{r,q'}(\mathbb{X})}^\ell, \qquad \text{for } \ell \geq 1, \quad (4.4)$$

we deduce for all $2 \leq m \leq N$ and $t \geq 0$,

$$\|G_t^{N,m}\|_{W_{p_0,m}^{-r,q}} \lesssim_m A_{p,q}^{N,m}(t) + \sum_{\substack{1 \leq k \leq l \leq m-1}} \sum_{\substack{j_1,\ldots,j_k \geq 1 \\ j_1 + \ldots + j_k = l}} N^{l-k-m+1} \prod_{i=1}^k \|G_t^{N,j_i}\|_{W_{mp_0,j_i}^{-r,q}}.$$

Hence, by Proposition 2.3, for suitable $\lambda_0, p, q, \kappa$,

$$\|G_t^{N,m}\|_{W_{p_0,m}^{-r,q}} \lesssim_m N^{-\frac{m}{2}} + e^{-\lambda_0 pmt} A_{3p,q}^{N,m}(0) + \sum_{\substack{1 \leq k \leq l \leq m-1}} \sum_{\substack{j_1,\ldots,j_k \geq 1 \\ j_1 + \ldots + j_k = l}} N^{l-k-m+1} \prod_{i=1}^k \|G_t^{N,j_i}\|_{W_{mp_0,j_i}^{-r,q}}.$$

It remains to express $A_{3p,q}^{N,m}(0)$ in the right-hand side in terms of initial correlations. For that purpose, we first appeal to the cluster expansion of moments in terms of cumulants: given $\varphi \in C_c^\infty(\mathbb{X})$, by definition of cumulants, we can expand for all $m \geq 1$,

$$\mathbb{E}\bigg[ \Big( \int_{\mathbb{X}} \varphi(\mu_0^N - \mu_\circ) \Big)^m \bigg] = \sum_{\pi \vdash \llbracket m \rrbracket} \prod_{B \in \pi} \kappa^{\sharp B} \bigg[ \int_{\mathbb{X}} \varphi(\mu_0^N - \mu_\circ) \bigg].$$

With the choice $\mu_\circ = F_\circ^{N,1}$, we note that $\kappa^1[\int_{\mathbb{X}}(\mu_0^N - \mu_\circ)] = \mathbb{E}[\int_{\mathbb{X}}(\mu_0^N - \mu_\circ)] = 0$, and we may thus restrict the sum to partitions $\pi$ such that $\sharp B > 1$ for all $B \in \pi$. Appealing to (4.2) at $t = 0$ in order to control cumulants back in terms of correlation functions, and taking the supremum over $\varphi$, we deduce

$$\sup \bigg\{ \mathbb{E}\bigg[ \Big( \int_{\mathbb{X}} \varphi(\mu_0^N - \mu_\circ) \Big)^m \bigg] \, : \, \varphi \in C_c^\infty(\mathbb{X}), \ \|\langle z \rangle^{-3p} \varphi\|_{W^{2,q'}(\mathbb{X})} \leq 1 \bigg\}$$
$$\lesssim \sum_{\ell=1}^m \sum_{\substack{j_1,\ldots,j_\ell \geq 2 \\ j_1 + \ldots + j_\ell = m}} \prod_{i=1}^\ell \bigg( \sum_{\substack{1 \leq k \leq l \leq j_i}} \sum_{\substack{n_1,\ldots,n_k \geq 1 \\ n_1 + \ldots + n_k = l}} N^{l-k-j_i+1} \prod_{s=1}^k \|G_\circ^{N,n_s}\|_{W_{3mp,n_s}^{-2,q}} \bigg).$$

In the right-hand side, we note that the sum over $j_1,\ldots,j_\ell \geq 2$ with $j_1 + \ldots + j_\ell = m$ requires $2\ell \leq m$, so the first sum can be restricted accordingly. Reorganizing the sums, we are then led to

$$\sup\left\{\mathbb{E}\left[\left(\int_{\mathbb{X}}\varphi(\mu_0^N-\mu_\circ)\right)^m\right] \: : \: \varphi\in C_c^\infty(\mathbb{X}),\ \|\langle z\rangle^{-3p}\varphi\|_{W^{2,q'}(\mathbb{X})}\le 1\right\}$$
$$\lesssim \sum_{\ell=1}^{m/2}\sum_{\ell\le K\le L\le m}\sum_{\substack{j_1,\ldots,j_K\ge 1\\ j_1+\ldots+j_K=L}} N^{L-K-m+\ell}\prod_{i=1}^{K}\|G_\circ^{N,j_i}\|_{W_{3mp,j_i}^{-2,q}}.$$

For $m$ even, this means

$$A_{3p,q}^{N,m}(0)\lesssim \sum_{\ell=1}^{m/2}\sum_{\ell\le K\le L\le m}\sum_{\substack{j_1,\ldots,j_K\ge 1\\ j_1+\ldots j_K=L}} N^{L-K-m+\ell}\prod_{i=1}^{K}\|G_\circ^{N,j_i}\|_{W_{3mp,j_i}^{-2,q}},$$

and for $m$ odd, after combination with (3.12),

$$A_{3p,q}^{N,m}(0)\lesssim \sum_{\ell=1}^{m}\sum_{\ell\le K\le L\le 2m}\sum_{\substack{1\le j_1,\ldots,j_K\le m+1\\ j_1+\ldots+j_K=L}} \left(N^{L-K-2m+\ell}\prod_{i=1}^{K}\|G_\circ^{N,j_i}\|_{W_{3mp,j_i}^{-2,q}}\right)^{\frac{1}{2}}.$$

This concludes the proof.  □

4.2. **BBGKY hierarchical estimates.** Starting from the BBGKY hierarchy for marginals of $F^N$, and recalling how marginals can be expanded in terms of correlations (and vice versa), we can derive a corresponding BBGKY hierarchy of equations for correlations. Before stating it, we start by introducing some useful notation.

**Definition 4.2.** Consider a collection $\{h^m\}_{1\le m\le N}$ of functions $h^m : \mathbb{X}^m\to\mathbb{R}$ such that for all $m$ the function $h^m$ is symmetric in its $m$ entries (such as $\{F^{N,m}\}_{1\le m\le N}$ or $\{G^{N,m}\}_{1\le m\le N}$).

— For $P\subset \llbracket m\rrbracket$ with $P\ne\varnothing$, we define $h^P : \mathbb{X}^m\to\mathbb{R}$ as
$$h^P(z_{\llbracket m\rrbracket}) \: := \: h^{\sharp P}(z_P),$$
and for $P=\varnothing$ we set $h^\varnothing := 0$.

— For $P\subset \llbracket m\rrbracket$ with $P\ne \llbracket N\rrbracket$, we define $h^{P\cup\{*\}} : \mathbb{X}^{m+1}\to\mathbb{R}$ as
$$h^{P\cup\{*\}}(z_{\llbracket m\rrbracket},z_*) \: := \: h^{\sharp P+1}(z_P,z_*),$$
and for $P=\llbracket N\rrbracket$ we set $h^{\llbracket N\rrbracket\cup\{*\}}=0$.

— For $P\subset \llbracket m\rrbracket$ and $k,\ell\in P$, we define
$$S_{k,\ell}h^P \: := \: \nabla W(x_k-x_\ell)\cdot\nabla_{v_k}h^P.$$

— For $P\subset \llbracket m\rrbracket$ and $k\in P$, we define
$$H_k h^{P\cup\{*\}}(z_{\llbracket m\rrbracket}) \: := \: \int_{\mathbb{X}}\nabla W(x_k-x_*)\cdot\nabla_{v_k}h^{P\cup\{*\}}(z_{\llbracket m\rrbracket},z_*)\,\mathrm{d}x_*.$$

With this notation, we may now formulate the BBGKY hierarchy of equations satisfied by correlation functions. To the best of our knowledge, such hierarchies for correlations were first written down by Ernst and Cohen [12] in the context of Boltzmann-type systems. We refer to [14, Section 4] for the derivation in the case of the overdamped dynamics (1.19): lengthy but straightforward algebraic manipulations are needed to collect the different factors in the correlation hierarchy. For convenience, we shall write $A-B=A\setminus B$ for set difference, with for instance $A-B-C=A\setminus(B\cup C)$ and $A\cup B-C=(A\cup B)\setminus C$.

**Lemma 4.3** (Correlation hierarchy, e.g. [14]). *For fixed $N$, correlation functions $\{G^{N,m}\}_{1\le m\le N}$ satisfy the following hierarchy of equations: for all $1\le m\le N$,*

$$\partial_t G^{N,m}=L_{N,m}G^{N,m}+\kappa\frac{N-m}{N}\sum_{k=1}^{m}H_k G^{N,\llbracket m\rrbracket\cup\{*\}}$$

$$- \kappa \sum_{k=1}^{m} \sum_{A \subset \llbracket m \rrbracket - \{k\}} \frac{m-1-\sharp A}{N} H_k \big( G^{N, A \cup \{k, *\}} G^{N, \llbracket m \rrbracket - \{k\} - A} \big)$$

$$+ \kappa \frac{N-m}{N} \sum_{k=1}^{m} \sum_{\substack{A \subsetneq \llbracket m \rrbracket - \{k\} \\ A \neq \varnothing}} H_k \big( G^{N, A \cup \{k\}} G^{N, \llbracket m \rrbracket \cup \{*\} - A - \{k\}} \big)$$

$$- \kappa \sum_{k=1}^{m} \sum_{A \subset \llbracket m \rrbracket - \{k\}} \sum_{B \subset \llbracket m \rrbracket - \{k\} - A} \frac{m-1-\sharp A - \sharp B}{N} H_k \big( G^{N, A \cup \{k\}} G^{N, B \cup \{*\}} G^{N, \llbracket m \rrbracket - A - B - \{k\}} \big)$$

$$+ \frac{\kappa}{N} \sum_{k, \ell=1}^{m} S_{k, \ell} G^{N, \llbracket m \rrbracket} + \frac{\kappa}{N} \sum_{k \neq \ell}^{m} \sum_{A \subset \llbracket m \rrbracket - \{k, \ell\}} S_{k, \ell} \big( G^{N, A \cup \{k\}} G^{N, \llbracket m \rrbracket - A - \{k\}} \big)$$

$$+ \kappa \frac{N-m}{m} \sum_{k=1}^{m} H_k \Big( (F^{N, \{*\}} - \mu(z_*)) G^{N, \llbracket m \rrbracket} + (F^{N, \{k\}} - \mu(z_k)) G^{N, \llbracket m \rrbracket \cup \{*\} - \{k\}} \Big),$$

where $L_{N,m}$ stands for the following (non-autonomous) m-particle linearized mean-field operator,

$$L_{N,m} \quad := \quad \sum_{k=1}^{m} \mathrm{Id}^{\otimes k-1} \otimes L_{N,m}^{\circ} \otimes \mathrm{Id}^{\otimes m-k}, \tag{4.5}$$

$$L_{N,m}^{\circ} h \quad := \quad \tfrac{1}{2} \mathrm{div}_v ((\nabla_v + \beta v) h) - v \cdot \nabla_x h + \Big( \nabla A + \kappa \frac{N-m}{N} \nabla W * \mu \Big) \cdot \nabla_v h$$

$$+ \kappa \frac{N-m}{N} \nabla_v \mu \cdot \int_{\mathbb{X}} \nabla W (x - x_*) h(z_*) \, dz_*.$$

Starting from this hierarchy of equations for correlations, combined with ergodic estimates for the linearized mean-field evolution, we establish the following a priori estimates.

**Proposition 4.4.** *There is $\lambda_0 > 0$ (only depending on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$) such that the following holds: given $1 < q \le 2$ and $0 < p \le \frac{1}{6}$ with $pq' \gg 1$ large enough (only depending on $d, \beta, a$), provided that $0 < \kappa \ll 1$ is small enough (further depending on $p, q, Q_2(F_\circ^{N,1})$, and on $\|W\|_{W^{d+3,\infty} \cap H^s(\mathbb{R}^d)}$ for some $s > \frac{d}{2} + 5$), we have for all $2 \le m \le N$, $r \ge 4$, and $t \ge 0$,*

$$\|G_t^{N,m}\|_{W_{p,m}^{-r,q}} \lesssim e^{-\lambda_0 pmt} \|G_\circ^{N,m}\|_{W_{p,m}^{-r,q}} + \int_0^t e^{-\lambda_0 pm(t-s)} \bigg( \|G_s^{N,m+1}\|_{W_{p,m+1}^{1-r,q}}$$

$$+ \sum_{\substack{i_1+i_2=m+1 \\ 1 \le i_1, i_2 \le m-1}} \|G_s^{N,i_1}\|_{W_{p,i_1}^{1-r,q}} \|G_s^{N,i_2}\|_{W_{p,i_2}^{1-r,q}} + N^{-1} \sum_{\substack{i_1+i_2=m \\ 1 \le i_1, i_2 \le m-1}} \|G_s^{N,i_1}\|_{W_{p,i_1}^{1-r,q}} \|G_s^{N,i_2}\|_{W_{p,i_2}^{1-r,q}}$$

$$+ N^{-1} \sum_{\substack{i_1+i_2+i_3=m+1 \\ 1 \le i_1, i_2, i_3 \le m-1}} \|G_s^{N,i_1}\|_{W_{p,i_1}^{1-r,q}} \|G_s^{N,i_2}\|_{W_{p,i_2}^{1-r,q}} \|G_s^{N,i_3}\|_{W_{p,i_3}^{1-r,q}}$$

$$+ \Big( N^{-1} + e^{-\lambda_0 ps} \|G_\circ^{N,2}\|_{W_{3p,2}^{-2,q}} \Big) \|G_s^{N,m}\|_{W_{p,m}^{1-r,q}} \bigg) \, ds, \tag{4.6}$$

where the multiplicative constant further depends on $m, r, Q_{mp}(F_\circ^{N,1})$, and on $\|W\|_{W^{d+r+1,\infty} \cap H^s(\mathbb{R}^d)}$ for some $s > 2r + \frac{d}{2} - 1$.

*Proof.* We start by recalling the available ergodic estimates for the linearized mean-field operator $L_{N,m}$. Given $m \ge 1$ and $u_\circ^m \in C_c^\infty(\mathbb{X}^m)$ with $\int_{\mathbb{X}} u_\circ^m(z_{\llbracket m \rrbracket}) \, dz_j = 0$ for all $1 \le j \le m$, let us consider the solution $u^m$ of

$$\begin{cases} \partial_t u^m = L_{N,m} u^m, & t \ge 0, \\ u^m|_{t=0} = u_\circ^m, \end{cases}$$

where in the definition (4.5) of $L_{N,m}$ we recall that we have chosen $\mu$ as the solution of the mean-field equation (1.4) with initial condition $\mu_\circ = F_\circ^{N,1}$. By definition of $L_{N,m}$ as a Kronecker sum, we can write

$$u_t^m = (V_{0,t}^{N,m})^{\otimes m} u_\circ^m,$$

where $\{V_{s,t}^{N,m}\}_{0 \le s \le t}$ stands for the one-particle evolution defined by

$$\begin{cases} \partial_t V_{s,t}^{N,m} h = L_{N,m}^\circ V_{s,t}^{N,m}, & t \ge s, \\ V_{s,t}^{N,m} h|_{t=s} = h. & s \end{cases}$$

In each component, we may then appeal to the ergodic estimate derived in [2, Theorem 2.13]: there exist $\kappa_0, \lambda_0 > 0$ (only depending on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$) such that, given $\kappa \in [0, \kappa_0]$, $1 < q \le 2$, and $0 < p \le 1$ with $pq' \gg 1$ (only depending on $d, \beta, a$), we have for all $r \ge 2$ and $t \ge 0$,

$$\|u_t^m\|_{W_{p,m}^{-r,q}} \lesssim e^{-\lambda_0 pmt} \|u_\circ^m\|_{W_{p,m}^{-r,q}},$$

where the multiplicative constant only depends on $d, \beta, a, p, q, r, m, Q_2(F_\circ^{N,1}), \|W\|_{W^{d+r+1,\infty}(\mathbb{R}^d)}$. Testing the definition of $S_{k,\ell} h^P$ with a tensorized test function, decomposing $W$ as a superposition of Fourier modes and using polarization similarly as in (3.8) to rewrite the resulting expression as a superposition of integrals of $h^P$ tested with a tensorized test function (recall the definition (1.13) of functional spaces), we are led to

$$\|S_{k,\ell} h^P\|_{W_{p,\sharp P}^{-r,q}} \lesssim \|h^P\|_{W_{p,\sharp P}^{1-r,q}},$$

and similarly,

$$\|H_k h^{P \cup \{*\}}\|_{W_{p,\sharp P}^{-r,q}} \lesssim \|h^{P \cup \{*\}}\|_{W_{p,\sharp P+1}^{1-r,q}},$$

where the multiplicative constants only depend on $d, r, \sharp P$, and $\|W\|_{H^s(\mathbb{R}^d)}$ for some $s > 2r + \frac{d}{2} - 1$. Now applying this to the equation for $G^{N,m}$ in Lemma 4.3, we get for all $2 \le m \le N$, $r \ge 2$, and $t \ge 0$,

$$\|G_t^{N,m}\|_{W_{p,m}^{-r,q}} \lesssim e^{-\lambda_0 pmt} \|G_\circ^{N,m}\|_{W_{p,m}^{-r,q}}$$

$$+ \int_0^t e^{-\lambda_0 pm(t-s)} \Bigg( \|G_s^{N,m+1}\|_{W_{p,m+1}^{1-r,q}} + N^{-1} \sum_{\substack{i_1+i_2=m+1 \\ 1 \le i_1, i_2 \le m}} \|G_s^{N,i_1}\|_{W_{p,i_1}^{1-r,q}} \|G_s^{N,i_2}\|_{W_{p,i_2}^{1-r,q}}$$

$$+ \sum_{\substack{i_1+i_2=m+1 \\ 1 \le i_1, i_2 \le m-1}} \|G_s^{N,i_1}\|_{W_{p,i_1}^{1-r,q}} \|G_s^{N,i_2}\|_{W_{p,i_2}^{1-r,q}} + N^{-1} \sum_{\substack{i_1+i_2=m \\ 1 \le i_1, i_2 \le m-1}} \|G_s^{N,i_1}\|_{W_{p,i_1}^{1-r,q}} \|G_s^{N,i_2}\|_{W_{p,i_2}^{1-r,q}}$$

$$+ N^{-1} \sum_{\substack{i_1+i_2+i_3=m+1 \\ 1 \le i_1, i_2, i_3 \le m-1}} \|G_s^{N,i_1}\|_{W_{p,i_1}^{1-r,q}} \|G_s^{N,i_2}\|_{W_{p,i_2}^{1-r,q}} \|G_s^{N,i_3}\|_{W_{p,i_3}^{1-r,q}}$$

$$+ \left( N^{-1} + \|F_s^{N,1} - \mu_s\|_{W_{p,1}^{1-r,q}} \right) \|G_s^{N,m}\|_{W_{p,m}^{1-r,q}} \Bigg) ds.$$

Note that in the first sum in the right-hand side the terms with $(i_1, i_2) = (m,1)$ or $(1,m)$ can be bounded by $N^{-1} \|G_s^{N,m}\|_{W_{p,m}^{1-r,q}}$, which already appears in the last term. Hence, this first sum can be restricted to $1 \le i_1, i_2 \le m-1$, which is then bounded by the second sum. For the last right-hand side term, involving $F^{N,1} - \mu$, we appeal to Theorem 1.1 with initial condition $\mu_\circ = F_\circ^{N,1}$: provided that $\lambda_0$ is chosen small enough (only depending on $d, \beta, a, \|W\|_{W^{1,\infty}(\mathbb{R}^d)}$), for all $1 < q \le 2$ and $0 < p \le \frac{1}{6}$ with $pq' \gg 1$ large enough (only depending on $d, \beta, a$), provided that $\kappa \ll 1$ is small enough (further depending on $p, q, Q_2(F_\circ^{N,1}), \|W\|_{W^{d+3,\infty} \cap H^s(\mathbb{R}^d)}$ for some $s > \frac{d}{2} + 5$), we have

$$\|F_t^{N,1} - \mu_t\|_{W_{p,1}^{-3,q}} \lesssim N^{-1} + e^{-2\lambda_0 pt} \|G_\circ^{N,2}\|_{W_{3p,2}^{-2,q}},$$

which then yields the conclusion. $\qquad\square$

4.3. **Proof of Theorem 1.2.** The claim (1.15) already follows from (3.1) in the proof of Theorem 1.1 with the choice $\mu_\circ = F_\circ^{N,1}$. It remains to prove (1.17). Let $1 < q \le 2$ and $0 < p \le \frac{1}{6}$ with $pq' \gg 1$ large enough (only depending on $d, \beta, a$), let $2 \le m \le N$ and $\alpha \in [0,1]$ be fixed, and assume that initially

$$\|G_\circ^{N,j}\|_{W_{3p,j}^{-2,q}} \le C_j^\circ N^{-\alpha(j-1)}, \qquad \text{for all } 1 \le j \le 2m-1. \tag{4.7}$$

We split the proof into two main steps.

**Step 1:** Suboptimal estimates.
Let $\lambda_0, \kappa$ be as in Proposition 2.3. By a straightforward computation, under the initial assumption (4.7), the result of Proposition 4.1 with $p_0 = p/n!$ yields for all $2 \le n \le 2m-1$, $r \ge 2$, and $t \ge 0$,

$$\|G_t^{N,n}\|_{W_{p/n!,n}^{-r,q}} \lesssim N^{-\frac{n}{2}} + e^{-\frac{\lambda_0 p}{n!}t} N^{-\alpha\frac{n}{2}} + \sum_{1 \le k \le l \le n-1} \sum_{\substack{j_1,\ldots,j_k \ge 1 \\ j_1 + \ldots + j_k = l}} N^{l-k-n+1} \prod_{i=1}^k \|G_t^{N,j_i}\|_{W_{p/(n-1)!,j_i}^{-r,q}}.$$

By a direct iteration of these estimates, inductively eliminating correlations in the right-hand side, and using $\|G_t^{N,1}\|_{W_{p,1}^{-r,q}} \lesssim 1$, we deduce for all $2 \le n \le 2m-1$, $r \ge 2$, and $t \ge 0$,

$$\|G_t^{N,n}\|_{W_{p/n!,n}^{-r,q}} \lesssim N^{-\frac{n}{2}} + e^{-\frac{\lambda_0 p}{n!}t} N^{-\alpha\frac{n}{2}}. \tag{4.8}$$

**Step 2:** Conclusion via BBGKY analysis.
We leverage the suboptimal correlation estimates of Step 1, using the BBGKY estimates of Proposition 4.4 and arguing by induction. As the desired estimates (1.17) are already known to hold for two-particle correlations, cf. (1.15), we can assume that there is some $3 \le n \le m$ such that the following holds: provided that $\kappa$ is small enough (also depending on $n$), we have for all $t \ge 0$,

$$\|G_t^{N,j}\|_{W_{p_j,j}^{-r_j,q_j}} \lesssim N^{1-j} + e^{-\lambda_0 p_j t} N^{\alpha(1-j)}, \qquad \text{for all } 2 \le j \le n-1, \tag{4.9}$$

where we have set

$$r_j := j+1, \qquad p_j := \frac{p}{(2j-2)!}, \qquad q_j := \frac{(2j-2)!q'}{(2j-2)!q'-1}.$$

Under this induction assumption, we shall show that the same estimate (4.9) is also automatically valid for $j = n$ up to further restricting the range of $\kappa$, which then concludes the proof. We split the argument into two further substeps.

**Substep 2.1:** Proof that for $\kappa$ small enough (also depending on $n$) we have for all $0 \le k \le n-2$, $r \ge 4$, and $t \ge 0$,

$$\|G_t^{N,n+k}\|_{W_{p_n,n+k}^{-r,q_n}} \lesssim N^{1-n-k} + e^{-\lambda_0 p_n t} N^{\alpha(1-n-k)}$$
$$+ \int_0^t e^{-\lambda_0 p_n (n+k)(t-s)} \bigg( \|G_s^{N,n+k+1}\|_{W_{p_n,n+k+1}^{1-r,q_n}} + \Big( N^{-1} + e^{-\lambda_0 p_n s} N^{-\alpha} \Big) \|G_s^{N,n+k}\|_{W_{p_n,n+k}^{1-r,q_n}}$$
$$+ \sum_{i=2}^{k+1} \Big( N^{1-i} + e^{-\lambda_0 p_n s} N^{\alpha(1-i)} \Big) \|G_s^{N,n+k+1-i}\|_{W_{p_n,n+k+1-i}^{1-r,q_n}} \bigg) \, \mathrm{d}s. \tag{4.10}$$

We apply Proposition 4.4 with $m = n + k$ and with exponents $p_n, q_n$ satisfying $p_n q_n' = pq' \gg 1$. Provided that $\kappa$ is small enough (depending on $n$), we get from Proposition 4.4, for all $0 \le k \le n-2$, $r \ge 4$, and $t \ge 0$,

$$\|G_t^{N,n+k}\|_{W_{p_n,n+k}^{-r,q_n}} \lesssim e^{-\lambda_0 p_n t} \|G_\circ^{N,n+k}\|_{W_{p_n,n+k}^{-r,q_n}} + \int_0^t e^{-\lambda_0 p_n (n+k)(t-s)} \bigg( \|G_s^{N,n+k+1}\|_{W_{p_n,n+k+1}^{1-r,q_n}}$$
$$+ \sum_{\substack{i_1+i_2=n+k+1 \\ 1 \le i_1,i_2 \le n+k-1}} \|G_s^{N,i_1}\|_{W_{p_n,i_1}^{1-r,q_n}} \|G_s^{N,i_2}\|_{W_{p_n,i_2}^{1-r,q_n}} + N^{-1} \sum_{\substack{i_1+i_2=n+k \\ 1 \le i_1,i_2 \le n+k-1}} \|G_s^{N,i_1}\|_{W_{p_n,i_1}^{1-r,q_n}} \|G_s^{N,i_2}\|_{W_{p_n,i_2}^{1-r,q_n}}$$

$$+ N^{-1} \sum_{\substack{i_1+i_2+i_3=n+k+1 \\ 1 \le i_1,i_2,i_3 \le n+k-1}} \|G_s^{N,i_1}\|_{W_{p_n,i_1}^{1-r,q_n}} \|G_s^{N,i_2}\|_{W_{p_n,i_2}^{1-r,q_n}} \|G_s^{N,i_3}\|_{W_{p_n,i_3}^{1-r,q_n}}$$

$$+ \left( N^{-1} + e^{-\lambda_0 p_n s} \|G_\circ^{N,2}\|_{W_{3p_n,2}^{-2,q_n}} \right) \|G_s^{N,n+k}\|_{W_{p_n,n+k}^{1-r,q_n}} \right) ds.$$

Note that the conditions $p_j q_j' = pq' > 2d$ and $r_j = j + 1$ entail by Jensen's inequality, for $j \le n-1$,

$$\|G^{N,j}\|_{W_{p_n,j}^{1-r_n,q_n}} \lesssim \|G^{N,j}\|_{W_{p_j,j}^{-r_j,q_j}}.$$

Using the initial assumption (4.7), as well as the induction assumption (4.9) for $G^{N,j}$ with $j \le n-1$, the claim (4.10) follows from the above estimate after straightforward simplifications.

**Substep 2.2:** Proof that for $\kappa$ small enough (depending on $n$) we have for all $0 \le \ell \le n-2$ and $t \ge 0$,

$$\|G_t^{N,n+k}\|_{W_{p_n,n+k}^{-3-\ell,q_n}} \lesssim N^{-\frac{n+k+\ell}{2}} + e^{-\lambda_0 p_n t} N^{-\alpha \frac{n+k+\ell}{2}}, \qquad \text{for all } 0 \le k \le n-2-\ell. \tag{4.11}$$

This will conclude the proof: indeed, choosing $\ell = n-2$ and $k = 0$, the above estimate precisely yields (4.9) with $j = n$ as desired. We turn to the proof of (4.11) and argue by induction on $\ell$. As for $\ell = 0$ the claim (4.11) follows from (4.8), we can assume that there is some $1 \le L < n-1$ such that (4.11) holds with $\ell = L - 1$, that is,

$$\|G_t^{N,n+k}\|_{W_{p_n,n+k}^{-2-L,q_n}} \lesssim N^{-\frac{n+k+L-1}{2}} + e^{-\lambda_0 p_n t} N^{-\alpha \frac{n+k+L-1}{2}}, \qquad \text{for all } 0 \le k \le n-1-L, \tag{4.12}$$

and it remains to show that (4.11) then automatically also holds for $\ell = L$. This immediately follows from the result (4.10) of Step 2.1 combined with the induction assumption (4.12). $\qquad \square$

## Acknowledgements

## References

[1] A. Béjar-López, A. Blaustein, P.-E. Jabin, and J. Soler. Longtime and chaotic dynamics in microscopic systems with singular interactions. *Preprint arXiv:2411.08614*, 2024.

[2] A. Bernou and M. Duerinckx. Uniform-in-time estimates on corrections to mean-field for interacting Brownian particles. Preprint, arXiv:2405.19306, 2024.

[3] F. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *M2AN Math. Model. Numer. Anal.*, 44(5):867–884, 2010.

[4] D. Bresch, M. Duerinckx, and P.-E. Jabin. A duality method for mean-field limits with singular interactions. Preprint, arXiv:2402.04695.

[5] D. Bresch, P.-E. Jabin, and J. Soler. A new approach to the mean-field limit of Vlasov-Fokker-Planck equations. Preprint, arXiv:2203.15747.

[6] F. Cornalba, J. Fischer, J. Ingmanns, and C. Raithel. Density fluctuations in weakly interacting particle systems via the Dean-Kawasaki equation. Preprint, arXiv:2303.00429.

[7] P. Del Moral and J. Tugaut. Uniform Propagation of Chaos and Creation of Chaos for a Class of Nonlinear Diffusions. *Stoch. Anal. Appl.*, 37(6):909–935, 2019.

[8] F. Delarue and A. Tse. Uniform in time weak propagation of chaos on the torus. Preprint, arXiv:2104.14973.

[9] M. Duerinckx. Lenard-Balescu correction to mean-field theory. *Probab. Math. Phys.*, 2(1):27–69, 2021.

[10] M. Duerinckx. On the Size of Chaos via Glauber Calculus in the Classical Mean-Field Dynamics. *Commun. Math. Phys.*, 382:613–653, 2021.

[11] M. Duerinckx and P.-E. Jabin. Dynamics of point-vortex systems near thermal equilibrium: relaxation or not? Preprint, arXiv:2401.01940.

[12] M. H. Ernst and E. G. D. Cohen. Nonequilibrium fluctuations in mu space. *J. Stat. Phys.*, 25(1):153–180, 1981.

[13] A. Guillin, P. Le Bris, and P. Monmarché. Convergence rates for the Vlasov-Fokker-Planck equation and uniform in time propagation of chaos in non convex cases. *Electronic Journal of Probability*, 27(none), January 2022.

[14] E. Hess-Childs and K. Rowan. Higher-order propagation of chaos in $L^2$ for interacting diffusions. Preprint, arXiv:2310.09654.

[15] D. Lacker. Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probab. Math. Phys.*, 4(2):377–432, 2023.

[16] D. Lacker and L. Le Flem. Sharp uniform-in-time propagation of chaos. *Probab. Theory Related Fields*, 187:443–480, 2023.

[17] J. Lukkarinen and A. Vuoksenmaa. Generation of chaos in the cumulant hierarchy of the stochastic Kac model. Preprint, arXiv:2407.17068.

[18] S. Mischler and C. Mouhot. Exponential stability of slowly decaying solutions to the kinetic-Fokker-Planck equation. *Arch. Ration. Mech. Anal.*, 221(2):677–723, 2016.

[19] M. Rosenzweig and S. Serfaty. Modulated logarithmic sobolev inequalities and generation of chaos. Preprint, arXiv:2307.07587.

[20] C. Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009.

[21] P. Xie. Uniform-in-Time Estimates on the Size of Chaos for Interacting Particle Systems. Preprint, arXiv:2411.15406, 2024.

(Armand Bernou) Université de Lyon, Université Claude Bernard Lyon 1, Laboratoire de Sciences Actuarielle et Financière, 50 Avenue Tony Garnier, F-69007 Lyon, France
*Email address*: armand.bernou@univ-lyon1.fr

(Mitia Duerinckx) Université Libre de Bruxelles, Département de Mathématique, 1050 Brussels, Belgium
*Email address*: mitia.duerinckx@ulb.be

(Matthieu Ménard) Université Libre de Bruxelles, Département de Mathématique, 1050 Brussels, Belgium
*Email address*: matthieu.menard@ulb.be