Higher-order propagation of chaos in L^2 for interacting diffusions

Elias Hess-Childs* Keefer Rowan†

10/14/2023

Abstract

In this paper, we study diffusions in the torus with bounded pairwise interaction. We show for the first time propagation of chaos on arbitrary time horizons in a stronger L^2-based distance, as opposed to the usual Wasserstein or relative entropy distances. The estimate is based on iterating inequalities derived from the BBGKY hierarchy and does not follow directly from bounds on the full N-particle density. This argument gives the optimal rate in N, showing the distance between the j-particle marginal density and the tensor product of the mean-field limit is $O(N^{-1})$. We use cluster expansions to give perturbative higher-order corrections to the mean-field limit. For an arbitrary order i, these provide "low-dimensional" approximations to the j-particle marginal density with error $O(N^{-(i+1)})$.

Contents

1. Introduction

In this paper we consider systems of N interacting particles in $\mathbb{T}^d$ of the form

$$\begin{cases} dX_{j,N}(t) = \frac{1}{N}\sum_{k=1}^N K(X_{j,N}(t), X_{k,N}(t))dt + \sqrt{2}dW_j(t), \quad j \in \{1, \cdots, N\} \\ X_{j,N}(0) = Y_j, \end{cases} \tag{1.1}$$

where the $W_j(t)$ are independent standard Brownian motions in $\mathbb{T}^d$, Y_j are i.i.d. random variables with probability density $f(x)$, and $K(x, y)$ denotes the drift on a particle at position y induces on a particle at position x. Particle systems of this form arise in many contexts such as vortices in

*Courant Institute of Mathematical Sciences, New York University. elias.hess-childs@courant.nyu.edu.

†Courant Institute of Mathematical Sciences, New York University. keefer.rowan@cims.nyu.edu.

viscous fluids [Ons49, MP94], the training of large neural networks [CB18, RVE22], and aggregation and collective motion of microscopic organisms [TBL06, Per07].

We recall that the law of the vector $(X_{1,N}, X_{2,N}, \ldots, X_{N,N})$ has a density $f_{N,N} : [0,\infty) \times \mathbb{T}^{Nd} \to \mathbb{R}$, which solves the Liouville equation,

$$\begin{cases} \partial_t f_{N,N} - \Delta f_{N,N} + \frac{1}{N}\sum_{k,\ell=1}^N \nabla_{x_k} \cdot (K(x_k, x_\ell) f_{N,N}) = 0, \\ f_{N,N}(0,x) = \prod_{k=1}^N f(x_k) = f^{\otimes N}(x). \end{cases} \tag{1.2}$$

By integrating the equation (1.2) over $x_{j+1}, \ldots, x_N$ one finds that the marginal densities $f_{j,N}$ satisfy the PDE hierarchy

$$\partial_t f_{j,N} - \Delta f_{j,N} + \frac{1}{N} \sum_{k,\ell=1}^{j} \nabla_{x_k} \cdot (K(x_k, x_\ell) f_{j,N}) = -\frac{N-j}{N} \sum_{k=1}^{j} \nabla_{x_k} \cdot \int K(x_k, x_*) f_{j+1,N}(x, x_*)\, dx_* \tag{1.3}$$

with initial data

$$f_{j,N}(0,\cdot) = f^{\otimes j}.$$

We note that $f_{N,N}$ is exchangeable, therefore the j-particle marginals $f_{j,N}$ are exchangeable and independent of which $N-j$ coordinates were integrated over.

We study the *propagation of chaos* of the system (1.1), that is for any fixed j, the convergence as $N \to \infty$ of the marginal density $f_{j,N} \to \rho^{\otimes j}$, where ρ solves the *McKean-Vlasov* equation

$$\begin{cases} \partial_t \rho(t,x) - \Delta \rho(t,x) + \nabla \cdot \left(\int K(x, x_*) \rho(t, x_*)\, dx_* \rho(t,x) \right) = 0, \\ \rho(t,\cdot) = f(\cdot). \end{cases} \tag{1.4}$$

Propagation of chaos has been shown under a wide range of conditions on f, ρ, and K and under various distances; for some recent results see [DEGZ20, LLX20, GBM23] and for a review of the vast literature see [CD22]. Recently, there has been lots of activity around quantitative propagation of chaos using relative entropy as a distance. In particular, global bounds—that is bounds on the relative entropy between $f_{N,N}$ and $\rho^{\otimes N}$—have been used to show quantitative propagation of chaos such as in [BAZ99, JW16, Lac18, Jab19] for non-singular interactions. Additionally, estimates of this kind have been used for a large class of singular interactions [JW18, BJW19, dCRS23, RS23].

Results based on global bounds at best show

$$\sqrt{H(f_{j,N} \mid \rho^{\otimes j})} = O\Big(\sqrt{\tfrac{j}{N}}\Big),$$

where $H(f \mid g)$ is the relative entropy of f with respect to g. This was widely believed to be optimal, but in [Lac23] it was shown that

$$\sqrt{H(f_{j,N} \mid \rho^{\otimes j})} = O\Big(\frac{j}{N}\Big),$$

for a class of interactions satisfying an exponential integrability condition. Further, this rate was shown to be optimal by constructing an example that saturates the bound. Instead of using global bounds, [Lac23] uses the BBGKY hierarchy (1.3) to get bounds on $H(f_{j,N} \mid \rho^{\otimes j})$ in terms of $H(f_{j+1,N} \mid \rho^{\otimes (j+1)})$. By iterating these bounds, one can show this optimal rate.

In this paper, we instead prove bounds in an L^2 norm. In particular, we show for initial conditions $f \in L^\infty$ and bounded interaction, that for any $j = o(N^{2/3})$,

$$D_{j,N} := \left(\int \Big| \frac{f_{j,N} - \rho^{\otimes j}}{\rho^{\otimes j}} \Big|^2 \rho^{\otimes j}\, dx \right)^{1/2} = O\Big(\frac{j}{N}\Big).$$

We note that $D^2_{j,N} = \chi^2(f_{j,N} \mid \rho^{\otimes j})$, where $\chi^2(\mu \mid \nu)$ is the *chi-squared divergence* of μ with respect to ν. Pinsker's inequality and [SV16, Theorem 2] respectively imply the inequalities

$$\|\mu - \nu\|^2_{TV} \leqslant \frac{1}{2} H(\mu \mid \nu) \leqslant \frac{1}{2} \chi^2(\mu \mid \nu),$$

for any probability measures μ and ν. The L^2-type convergence of $f_{j,N} \to \rho^{\otimes j}$ thus implies the relative entropy convergence at the same optimal rate as in [Lac23], which in turn implies TV convergence.

The L^2 bounds are shown using somewhat analogous techniques to [Lac23], controlling $D_{j,N}$ by $D_{j+1,N}$ and iterating these bounds. In contrast to relative entropy, no sufficiently strong global bounds on $D_{N,N}$ are available. In fact, the best bound we can show is

$$D_{N,N} \leqslant Ce^{CNt}.$$

This bound is far from sufficient to directly imply that $D_{j,N} \to 0$ for fixed j. As such, this L^2 distance is not amenable to global techniques and so necessitates analysis of the BBGKY hierarchy. We note in the recent preprint [BJS22], propagation of chaos is shown in certain L^p spaces and even for certain singular interactions, but only for sufficiently short times. In this paper, we show convergence on any time horizon.

One heuristic justification for propagation of chaos involves discarding terms of order N^{-1} in the BBGKY hierarchy (1.3) and noting that the $\rho^{\otimes j}$ are a solution to the resulting hierarchy of equations, as is explained in Subsection 1.2. This suggests that the tensor product of the McKean-Vlasov solution $\rho^{\otimes j}$ is the 0th-order term of a perturbative expansion of $f_{j,N}$ in powers of N^{-1}. This turns out to be in fact true, as we show in this paper by constructing this perturbative expansion and showing the appropriate bounds. Finding the correct perturbative approximation of order greater than 0 requires the introduction of the *cluster (or cumulant) expansion*, which rewrite the marginal densities $f_{j,N}$ in terms of certain sums of products of *cluster functions* $g_{j,N}$, made precise in (1.9). After the correct perturbative approximations are found through cluster expansion, proving that they approximate $f_{j,N}$ to the appropriate order follows by the same analysis as the 0th-order case discussed above.

The higher-order terms of the perturbative approximation are, unlike the 0th-order terms $\rho^{\otimes j}$, not positive. In fact, in order to preserve the mean of the perturbative approximation, the higher-order terms are mean-zero and hence take both positive and negative values. As such, without strong pointwise control on the higher-order terms, the positivity of the higher-order approximations is unclear. This makes analyzing the error between $f_{j,N}$ and its approximation less amenable to probabilistic techniques such as relative entropy. In contrast, there are no such issues in the L^2 analysis.

Cluster expansions—and related expansions using correlation errors or v-functions—have been used in a wide variety of contexts to study asymptotics of statistical particle systems, for example [DMP91, PS17, BGSRS20] and citations therein. Somewhat relevant to our current study, [Due21] uses Glauber calculus to estimate cluster functions in the kinetic setting without noise in the evolution. In that work, the author uses a non-hierarchical technique and requires strong bounds on the regularity of the interaction: to go to arbitrary order the interaction must be C^∞.

The pair of papers [PPS19, PP19] use a correlation error expansion and hierarchical techniques. The authors consider an abstract setting that covers both quantum mean-field models as well as stochastic jump processes such as the Kac model. Their analysis relies on iterating across the BBGKY hierarchy to pass estimates on the *correlation errors*, which in turn implies propagation of chaos. In [PP19], they take perturbative expansions of the correlation errors to construct higher order corrections to propagation of chaos. In their setting, the time evolution is unitary, allowing the use of techniques that are not clearly applicable to the setting we consider. Additionally, their perturbative expansion involves approximations which meaningfully depend on N, and the number of equations one must solve to construct the approximation of $f_{j,N}$ to order i depends on j. In contrast, in our approach neither of these properties appear.

1.1 Statement of main results

Before stating the results, we introduce some notation. The first three definitions are needed in order to express the perturbative expansions.

Definition. For any set A, we use the notation $\pi \vdash A$ to denote that π is a partition of A. When appearing in a sum

$$\sum_{\pi \vdash A}$$

we mean that the sum is taken over all possible partitions of A. We often take $\pi \vdash [j]$ for some $j \in \mathbb{N}$ where

$$[j] := \{1, \ldots, j\}.$$

Definition. Fix some finite set A and let $(h_j)_{1 \leqslant j \leqslant |A|}$ be a family of exchangeable functions such that $h_j : \mathbb{T}^{jd} \to \mathbb{R}$. Then for any partition $\pi \vdash A$, we denote

$$\prod_{P \in \pi} h_P : (\mathbb{T}^d)^A \to \mathbb{R}$$

such that

$$\prod_{P \in \pi} h_P(x) := \prod_{P \in \pi} h_{|P|}(x^P),$$

where for $x \in (\mathbb{T}^d)^A$

$$x^P := (x_k)_{k \in P}.$$

Note by exchangeability, the order of the x_k in x^P doesn't matter.

Definition. For any partition $\pi \vdash j$ where $\pi = \{P_1, ..., P_k\}$, by

$$\sum_{\substack{(i_P)_{P \in \pi} \\ \sum i_P = i}}$$

we denote that the sum is over all choices of $i_{P_1}, \ldots, i_{P_k} \in \mathbb{N}$ such that $\sum_{P \in \pi} i_P = i$.

Our first results are on the existence and representation of the perturbative approximation.

Definition 1.1. Let $T = \{(i, j) \in \mathbb{N}^2 : 1 \leqslant j \leqslant i + 1, j \geqslant 1\}$. We define an ordering on T by saying $(i, j) \leqslant (a, b)$ if

$$i < a \text{ or both } i = a \text{ and } j \geqslant b.$$

Proposition 1.2. *Suppose the initial distribution $f \in L^2(\mathbb{T}^d)$ and the interaction $K \in L^\infty(\mathbb{T}^{2d})$. Then there exists a family of functions $g_j^i \in L^\infty_{loc}([0, \infty), L^2(\mathbb{T}^{jd})) \cap L^2_{loc}([0, \infty), H^1(\mathbb{T}^{jd}))$, where $j \in \{1, 2, \ldots\}$ and $i \in \{0, 1, \ldots\}$ so that g_j^i solve the equations (2.4)[1] with initial data (2.5). More so, the g_j^i have the properties:*

[1]In the linked equation, the operators H_k and $S_{k,\ell}$ appear. These are defined below in Definition 2.1.

1. *For $(i,j) \notin T$, $g_j^i = 0$.*

2. *For $(i,j) \in T$, the equation for g_j^i depends only on the g_ℓ^k with $(k,\ell) \leqslant (i,j)$ under the ordering on T. More so, the equation is linear in g_j^i for $(i,j) > (0,1)$.*

3. *$g_1^0 = \rho$, the unique solution to the McKean-Vlasov equation (1.4).*

4. *Assuming Property 1, these solutions are unique for fixed f.*

The functions g_j^i at this stage are somewhat opaque but are the natural perturbative expansion of the cluster functions $g_{j,N}$, as will be made clear in Subsection 1.2.

Theorem 1.3. *Suppose the initial distribution $f \in L^2(\mathbb{T}^d)$ and the interaction $K \in L^\infty(\mathbb{T}^{2d})$. Then let*

$$f_j^i := \sum_{\pi \vdash [j]} \sum_{\substack{(i_P)_{P\in\pi} \\ \sum i_P = i}} \prod_{P\in\pi} g_P^{i_P} \tag{1.5}$$

where the g_j^i are as in Proposition 1.2. Then

$$\begin{aligned}\partial_t f_j^i - \Delta f_j^i = &- \sum_{k=1}^{j} \nabla_{x_k} \cdot \int K(x_k, x_*) f_{j+1}^i(x, x_*)\, dx_* - \sum_{k,\ell=1}^{j} \nabla_{x_k} \cdot (K(x_k, x_\ell) f_j^{i-1}) \\ &+ j \sum_{k=1}^{j} \nabla_{x_k} \cdot \int K(x_k, x_*) f_{j+1}^{i-1}(x, x_*)\, dx_*, \end{aligned} \tag{1.6}$$

where we take the convention $f_j^{-1} = 0$ and the f_j^i have initial data

$$f_j^i(0,\cdot) = \begin{cases} f^{\otimes j} & i = 0, \\ 0 & i \geqslant 1. \end{cases}$$

In particular, we have

$$f_j^0 = \rho^{\otimes j},$$

where ρ is the unique solution to the McKean-Vlasov equation (1.4).

The case of $i = 0$ in Proposition 1.2 and Theorem 1.3 is just the usual setting for propagation of chaos: $g_1^0 = \rho$ and $f_j^0 = \rho^{\otimes j}$. See Remark 1.13 for an explicit representation of the $i = 1$ case.

We note that the equation (1.6) is what one gets from formally expanding $f_{j,N} = \sum_{i=0}^\infty N^{-i} f_j^i$, plugging the right hand side into the BBGKY hierarchy (1.3), and collecting orders. Thus we expect any f_j^i solving (1.6) to be such that

$$f_{j,N} = \sum_{k=0}^{i} N^{-k} f_j^k + O(N^{-(i+1)}).$$

Theorem 1.3 gives an explicit representation (1.5) of solutions f_j^i to (1.6). Further, properties 1 and 2 of Proposition 1.2 ensure that the expression (1.5) for f_j^i is computable in terms of the finite collection $\{g_\ell^k : k \leqslant i, (k,\ell) \in T\}$ which depends only on i, not on j or N. That is, in order to compute f_j^i for any j, one only needs to solve $\frac{1}{2}(i+2)(i+1)$ equations.

The main result of this paper is then to show that the f_j^i as constructed in Theorem 1.3 appropriately approximate $f_{j,N}$.

Theorem 1.4. *Suppose $f \in L^\infty(\mathbb{T}^d), K \in L^\infty(\mathbb{T}^{2d})$, and there exists $m > 0$ such that that $f \geqslant m$. Then for each $i \in \mathbb{N}$, there exists $C(\|K\|_{L^\infty(\mathbb{T}^{2d})}, i) < \infty$ such that for any N and any j with*

$$j \leqslant C^{-1} e^{-Ct^2} N^{2/3},$$

we have the bound

$$\int \left| \frac{f_{j,N} - \sum_{k=0}^{i} N^{-k} f_j^k}{\rho^{\otimes j}} \right|^2 \rho^{\otimes j} \, dx \leqslant Ce^{Ct} \left(\frac{j}{N} \right)^{2(i+1)}, \tag{1.7}$$

with the f_j^k given as in Theorem 1.3.

Remark 1.5. We note that in the above theorem we require $f \in L^\infty(\mathbb{T}^d)$ and $f \geqslant m > 0$, but neither the L^∞ norm of f nor the strict lower bound m show up in the constant C. Thus these are only *qualitative assumptions.* We need these assumptions to make sense of the PDEs under study. For the well-posedness of the McKean-Vlasov equation as well as the Kolmogorov equation, the assumption that $f \in L^\infty(\mathbb{T}^d)$ is used (in that it implies $f \in L^2(\mathbb{T}^d)$). Additionally, in order to show the result, we must repeatedly deal with terms involving ρ^{-1}; for example taking derivatives of them or integrating by parts against them. We thus really would like ρ^{-1} to live in a reasonably nice space, e.g. $\rho^{-1} \in L^\infty$, so as not to cause issues in the computations. Since we are working on the torus and the McKean-Vlasov equation is diffusive, one can show for all positive times that ρ is strictly bounded away from 0, but if the initial data f does not have such a bound, then this breaks down as $t \to 0$. Thus we require the lower bound on the initial data $f \geqslant m > 0$, despite no constants depending on this m. Similarly, to get all of the integrals to be clearly finite, we need $f \in L^\infty(\mathbb{T}^d)$.

As these assumptions of $f \in L^\infty(\mathbb{T}^d)$ and f strictly bounded away from 0 are purely qualitative, they are very soft restrictions. One may suspect that they should be able to be easily removed, but doing so while preserving all of the estimates on the equations is somewhat non-obvious. In any case, such an argument would be highly technical and distract from the main point of this paper. The authors plan in forthcoming work to use more probabilistic techniques to prove L^2-based propagation of chaos without these assumptions—in particular allowing the domain to be $\mathbb{R}^d$ in order to cover the second-order in time case, in which case no uniform positivity of the initial data is possible.

Remark 1.6. For $i \geqslant 1$, the L^2-type distance between $f_{j,N}$ and $\sum_{k=0}^{i} N^{-k} f_j^k$ bounded in (1.7) is not a chi-squared divergence, hence does not bound the relative entropy. Nevertheless, an application of Hölder's inequality implies that under the same conditions of Theorem 1.4,

$$\left\| f_{j,N} - \sum_{k=0}^{i} N^{-k} f_j^k \right\|_{TV} \leqslant Ce^{Ct} \left(\frac{j}{N} \right)^{i+1}.$$

Remark 1.7. We note that in the $i = 0$ case, Theorem 1.4 gives the estimate

$$\sqrt{\chi^2(f_{j,N} \mid \rho^{\otimes j})} \leqslant Ce^{Ct} \frac{j}{N},$$

showing convergence in chi-squared divergence (and hence in relative entropy and total variation) with optimal rate in N^{-1}.

Remark 1.8. A simple argument shows that the rate

$$\int \left| \frac{f_{j,N} - \sum_{k=0}^{i} N^{-k} f_j^k}{\rho^{\otimes j}} \right|^2 \rho^{\otimes j} \, dx = O(N^{-2(i+1)})$$

is optimal for some fixed j and i, provided the next order correction f_j^{i+1} is not identically zero. It is not completely straightforward to construct examples for which one can show that $f_j^{i+1} \neq 0$ for some j, but it would be extremely surprising if there were no such examples. If that were the case, there would be some i_* such that for any f and K, we would have $f_j^i = 0$ for all $i \geqslant i_*$ and all j. In particular, this would imply that $f_{j,N} - \sum_{k=0}^{i_*} N^{-k} f_j^k$ vanishes faster than any polynomial rate in N^{-1}.

Remark 1.9. Throughout this paper we assume that the initial data of $f_{j,N}$ is completely tensorized, that is $f_{j,N}(0,\cdot) = f^{\otimes j}$. As is usual, we don't strictly need this to be true; one can show the same bound (1.7) at order i, provided for all j, the initial data $f_{j,N}(0,\cdot) \in L^\infty(\mathbb{T}^{jd})$ and satisfies the quantitative bound

$$\int \left| \frac{f_{j,N}(0,\cdot) - f^{\otimes j}}{f^{\otimes j}} \right|^2 f^{\otimes j}\, dx \leqslant C_0 \Big(\frac{j}{N} \Big)^{2(i+1)},$$

for some C_0 independent of j. Of course then the constant C of the bound (1.7) would then depend on C_0. We omit this argument as it adds notational complexity without adding any real content.

Remark 1.10. We note the restriction $j = o(N^{2/3})$. This is not very constraining, and still shows strong bounds along a broad class of simultaneous limits of $(j,N) \to (\infty,\infty)$—these simultaneous limits are sometimes called *increasing propagation of chaos* [BAZ99, MM01]. The restriction is however worse than in [Lac23], which allows $j = O(N)$. This restriction originates from the prefactor $\frac{j^3}{N^2}$ that appears on a term in the fundamental energy-type estimate given in Proposition 3.3. We need the prefactor of this term to be $O(1)$ in order to not cause growth when the hierarchy of differential inequalities is iterated, thus we give the requirement that $j = O(N^{2/3})$. The time decay in the upper bound on j, that $j \leqslant C^{-1} e^{-Ct^2} N^{2/3}$—and hence that $j = o(N^{2/3})$—then comes from the iteration of a short time argument which requires us to restrict to a smaller set of j on each iteration.

Remark 1.11. Theorem 1.4 in particular implies that for fixed j

$$N(f_{j,N} - \rho^{\otimes j}) \xrightarrow{TV} f_j^1,$$

and similarly for the higher-order f_j^i. This justifies that the f_j^i are the natural next order corrections. We note that due to the N-dependence of the higher-order corrections in [PP19], no such result is available in their analysis.

Remark 1.12. The interaction K has not been assumed to be symmetric nor has $K(x,x)$ been assumed to be 0. In particular, we allow

$$K(x,y) := b(x) + \widehat{K}(x-y)$$

where b is a drift affecting all particles and $\widehat{K}$ is a translation-invariant pairwise interaction.

Remark 1.13. In order to make the general higher-order corrections more concrete, here we explicitly give the first-order corrections: the g_j^1 and f_j^1. All the $g_j^1 = 0$ except $j = 1, 2$. Letting ρ

be the unique solution to the McKean-Vlasov equation (1.4), g_2^1 solves the equation

$$\begin{aligned}\partial_t g_2^1 - \Delta g_2^1 &+ \nabla_x \cdot \int K(x,x_*)\big(\rho(x)g_2^1(y,x_*) + \rho(x_*)g_2^1(x,y)\big)\,dx_* \\ &+ \nabla_y \cdot \int K(y,x_*)\big(\rho(y)g_2^1(x,x_*) + \rho(x_*)g_2^1(x,y)\big)\,dx_* \\ = \nabla_x \cdot &\int K(x,x_*)\rho(x_*)\rho(x)\rho(y)\,dx_* + \nabla_y \cdot \int K(y,x_*)\rho(x_*)\rho(x)\rho(y)\,dx_* \\ &- \nabla_x \cdot \big(K(x,y)\rho(x)\rho(y)\big) - \nabla_y \cdot \big(K(y,x)\rho(x)\rho(y)\big).\end{aligned}$$

The equation for g_1^1 is

$$\begin{aligned}\partial_t g_1^1 - \Delta g_1^1 &+ \nabla \cdot \int K(x,x_*)\big(g_1^1(x_*)\rho(x) + \rho(x_*)g_1^1(x)\big)\,dx_* \\ &= \nabla \cdot \int K(x,x_*)\big(\rho(x_*)\rho(x) - g_2^1(x,x_*)\big)\,dx_* - \nabla \cdot \big(K(x,x)\rho(x)\big).\end{aligned}$$

Then for any j, f_j^1 is given by

$$f_j^1 = \sum_{k=1}^{j} g_1^1(x_k)\rho^{\otimes(j-1)}(x^{[j]-\{k\}}) + \sum_{1\leqslant k<\ell\leqslant j} g_2^1(x_k,x_\ell)\rho^{\otimes(j-2)}(x^{[j]-\{k,\ell\}}).$$

We note that the equation for g_2^1 only depends on ρ, the equation for g_1^1 only depends on ρ and g_2^1, and f_j^1 is computable for any j in terms of the three functions ρ, g_1^1, and g_2^1.

1.2 Overview of the argument

We first introduce the motivation and construction of the higher-order corrections f_j^i through the cluster expansion and perturbation theory. We then explain the L^2 analysis of the BBGKY hierarchy that allows us to prove the bound (1.7).

1.2.1. Higher-order corrections to propagation of chaos. One formal argument for propagation of chaos is given by discarding terms of order N^{-1} in the hierarchy (1.3), which gives the hierarchy

$$\begin{cases}\partial_t f_j^0 - \Delta f_j^0 = -\sum_{k=1}^{j} \nabla_{x_k} \cdot \int K(x_k,x_*) f_{j+1}^0(x,x_*)\,dx_*, \\ f_j^0(0,\cdot) = f^{\otimes j},\end{cases}$$

where the notation f_j^0 is due to the fact we are only keeping track of terms to 0th order in N^{-1}. One can then note that $f_j^0 := \rho^{\otimes j}$ is a solution to this system. Thus the tensor product $\rho^{\otimes j}$ is formally the 0th order term of a perturbative expansion of $f_{j,N}$. We are then interested in the higher-order terms of this expansion, so we formally suppose

$$f_{j,N} = \sum_{i=0}^{\infty} N^{-i} f_j^i. \tag{1.8}$$

Collecting orders of N^{-1}, we get that f_j^i solves the equation (1.6). We note that for each i, this is an infinite hierarchy of equations in j, with forcing depending on $\{f_j^{i-1} : j \in \mathbb{N}\}$. It is not at all clear how to directly construct solutions to these hierarchies.

To solve this problem, we introduce the *cluster (or cumulant) expansion*. That is, we express the $f_{j,N}$ in terms of a family of exchangeable functions $g_{1,N}, ..., g_{N,N}$, namely

$$f_{j,N} = \sum_{\pi \vdash [j]} \prod_{P \in \pi} g_{|P|,N}(x^P). \tag{1.9}$$

From this ansatz, one can deduce an inversion formula

$$g_{j,N} = \sum_{\pi \vdash [j]} (-1)^{|\pi|-1}(|\pi|-1)! \prod_{P \in \pi} f_{|P|,N}(x^P), \tag{1.10}$$

which defines the $g_{k,N}$ in terms of the $f_{j,N}$. The BBGKY hierarchy (1.3) then induces the hierarchy of equations (2.3) on the $g_{k,N}$. We can then take formal perturbative expansion of the $g_{k,N}$, writing

$$g_{k,N} = \sum_{i=0}^{\infty} N^{-i} g_k^i, \tag{1.11}$$

and collect orders in equation (2.3) to get equation (2.4) on the g_k^i. Unlike the equations (1.6), the equations for the g_k^i can be inductively solved. Then plugging the expansion (1.11) into the cluster expansion (1.9) and collecting orders of N^{-1}, we formally find representation of the f_j^i of (1.8) in terms of the g_j^i, as given in (1.5). Theorem 1.3 then gives that this expression for f_j^i in terms of the g_j^i actually solves the equation (1.6) we formally expect it to.

Remark 1.14. Note (1.7) gives that

$$f_{j,N} = \sum_{k=0}^{i} N^{-k} f_j^k + O(N^{-(i+1)}).$$

Inserting this approximation into (1.10) and using the definition of the f_j^k, (1.5), one can show that

$$g_{j,N} = \sum_{k=0}^{i} N^{-k} g_j^k + O(N^{-(i+1)}).$$

Since, as noted in Proposition 1.2 $g_j^k = 0$ for $k \leqslant j-2$, we see by letting $i = j-2$ that

$$g_{j,N} = O(N^{-(j-1)}).$$

This shows that that $g_{j,N}$ are small for all $j \geqslant 2$ and in particular allows estimates on the joint cumulants of observables on j particles. That is, for any $\varphi_1, \ldots, \varphi_j \in C(\mathbb{T}^d, \mathbb{R})$ we have that

$$\begin{aligned} \kappa\Big(\varphi_1\big(X_{1,N}(t)\big), \ldots, \varphi_j\big(X_{j,N}(t)\big)\Big) &= \int \prod_{k=1}^{j} \varphi_k(x_k)\, g_{j,N}(t, x_1, \ldots, x_j)\, dx \\ &\leqslant \prod_{k=1}^{j} \|\varphi_k\|_{C^0} \|g_{j,N}\|_{TV} \\ &= O(N^{-(j-1)}), \end{aligned}$$

where $\kappa(Z_1, .., Z_j)$ denotes the joint cumulant of $Z_1, ..., Z_j$. Thus the results of this paper in particular show the smallness of joint cumulants of observables of many particles, with a rate getting very small as the number of particles gets large. We note that these estimates on cumulants are related to the Bogolyubov corrections—a version of these bounds on the cumulants in the context of second-order in time interacting particle systems conjectured by physicists [Bog60].

1.2.2. L^2 hierarchy estimates. We now sketch the L^2-based estimates on the BBGKY hierarchy. Fundamentally we are concerned with estimating the size of solutions to the hierarchy

$$\partial_t \gamma_j - \Delta \gamma_j + \frac{1}{N} \sum_{k,\ell=1}^{j} \nabla_{x_k} \cdot (K(x_k, x_\ell)\gamma_j) + \frac{N-j}{N} \sum_{k=1}^{j} \nabla_{x_k} \cdot \int K(x_k, x_*)\gamma_{j+1}(x, x_*)\, dx_* = \nabla \cdot R_j, \tag{1.12}$$

where the γ_j have initial data $\gamma_j(0, \cdot) = 0$. In particular, for Theorem 1.4 we take for fixed i,

$$\gamma_j = f_{j,N} - \sum_{k=0}^{i} N^{-k} f_j^i.$$

By construction, this γ_j satisfies (1.12) with an error R_j such that $R_j = O(N^{-(i+1)})$. The goal then is to show that $\gamma_j = O(R_j)$. This is accomplished by noting

$$\begin{aligned} \frac{d}{dt} \int \Big|\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \leqslant 2j\|K\|_{L^\infty}^2 \bigg(\int \Big|\frac{\gamma_{j+1}}{\rho^{\otimes (j+1)}}\Big|^2 \rho^{\otimes (j+1)}\, dx_* dx - \int \Big|\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \bigg) \\ + 4\frac{j^3}{N^2}\|K\|_{L^\infty}^2 \int \Big|\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx + 2\int \Big|\frac{R_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx, \end{aligned} \tag{1.13}$$

which is shown (see Proposition 3.3 for details) by directly expanding the time derivative on the left hand side and using the equations solved by γ_j and $\rho^{\otimes j}$. This estimate is in many way analogous to [Lac23, Equation (1-17)]. The bound is also used similarly. In particular letting $\beta := 4\|K\|_{L^\infty}^2$ and

$$x_j := \int \Big|\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx, \quad r_j := 2\int \Big|\frac{R_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx,$$

(1.13) implies

$$\dot{x}_j \leqslant \beta j(x_{j+1} - x_j) + \beta\frac{j^3}{N^2}x_j + r_j. \tag{1.14}$$

In Proposition 3.6, we show that

$$r_j \leqslant Ce^{Ct}\Big(\frac{j}{N}\Big)^{2(i+1)},$$

where C does not depend on j. It is worth noting that a more naive bound on r_j would give a suboptimal rate in j—giving j^{3+4i} instead of $j^{2(i+1)}$—but by taking advantage of certain L^2 orthogonality, the above bound can be shown.

Now what remains to be shown is the r_j are the leading order contribution to the size of the x_j. To see this, x_j is controlled by iteratively applying Grönwall's inequality to (1.14), which gives an estimate of the form

$$x_j(t) \leqslant CI_j^\ell(t) \sup_{s\in[0,t]} x_{j+\ell}(s) + C\sum_{k=0}^{\ell-1} I_j^{k+1}(t)\frac{r_{j+k}}{j+k}, \tag{1.15}$$

provided $j+\ell \leqslant N^{2/3}$ and where the I_j^ℓ, defined in Definition 3.8, are certain iterated exponential integrals. The I_j^ℓ admit the estimates

$$I_j^\ell(t) \leqslant \Big(\frac{j+b}{j+\ell}\Big)^b e^{\beta bt}, \quad b \in \mathbb{N}. \tag{1.16}$$

By taking $b = 2i + 3$ and applying this bound we appropriately control the second term of (1.15), namely

$$C \sum_{k=0}^{\ell-1} I_j^{k+1}(t) \frac{r_{j+k}}{j+k} \leqslant Ce^{Ct} \Big(\frac{j}{N} \Big)^{2(i+1)}.$$

The remaining issue is to bound $I_j^\ell(t) \sup_{s \in [0,t]} x_{j+\ell}(s)$. In [Lac23], a simple *a priori* bound[2] on the analog to x_k was available, giving that $x_k \leqslant Ckt$. No such bound is available in our setting. The best *a priori* bound we have is given by (3.4) which implies that $x_k \leqslant Ce^{Ckt}$. Thus the best bound of the remaining term that we have available is

$$I_j^\ell(t) \sup_{s \in [0,t]} x_{j+\ell}(s) \leqslant Ce^{C(j+\ell)t} I_j^\ell(t).$$

For this, using (1.16) with fixed b is insufficient, as the exponential growth will always beat polynomial decay. Instead, by optimally choosing b in (1.16), one can deduce the exponential decay estimate

$$I_j^\ell(t) \leqslant \exp(-\tfrac{1}{3} e^{-\beta t - 1} \ell), \quad \text{for } j \leqslant \tfrac{1}{3} e^{-\beta t - 1} \ell.$$

By correctly choosing ℓ and constraining j, using this estimate one can show for sufficient small times,

$$I_j^\ell(t) \sup_{s \in [0,t]} x_{j+\ell}(s) \leqslant Ce^{C(j+\ell)t} I_j^\ell(t) \leqslant C \Big(\frac{j}{N} \Big)^{2(i+1)}.$$

From this, we then get for some t_* and for all $t \leqslant t_*$, $j \leqslant C^{-1} N^{2/3}$,

$$\int \Big| \frac{\gamma_j}{\rho^{\otimes j}} \Big|^2 \rho^{\otimes j} \, dx = x_j \leqslant C \Big(\frac{j}{N} \Big)^{2(i+1)}.$$

This is of course only a short time result and is essentially what is shown in Lemma 3.15. It turns out one can essentially iterate this argument to get the result for all times, though substantial care needs to be taken in propagating the correct estimates in time. See Lemma 3.15 for details.

1.2.3. Organization of the argument. In Section 2, we give all of the algebraic results of the cluster expansion and perturbation theory as well as the qualitative properties of the perturbative approximations. In Subsection 2.1, we introduce the cluster expansion and the hierarchy of equations solved by the terms of the cluster expansion. In Subsection 2.2, we perturbatively expand the terms of the cluster expansion and introduce the hierarchy of equations solved by these perturbative approximations. We then use the perturbative expansion of the terms of the cluster expansion to construct a perturbative expansion of the marginal densities. We then note the equations solved by the terms of the perturbative expansion of the marginal densities. In Subsection 2.3, we supply a proof of Proposition 1.2 as well as noting an additional important marginalization property of the functions g_j^i. Theorem 1.3 is a direct consequence of the results of Section 2, as will be made clear. Many of the proofs of the propositions stated in Section 2 will be deferred to Section 4, as they laborious, elementary, and unenlightening.

In Section 3, we proceed with the analytic work of proving Theorem 1.4. We start by proving a hierarchical "energy estimate" for the difference between $f_{j,N}$ and its perturbative approximation to finite order. The resulting bound can be viewed as a hierarchy of differential inequalities only involving time derivatives. We then note basic estimates of the terms involved, though the proofs

[2]By *a priori* we mean that the bound is in some way independent of the perturbation theory and is rather just an initial estimate of size.

of these estimates are deferred to the end of the section, Subsection 3.3, in order to not distract from the main analytic techniques for showing the L^2 bound. In Subsection 3.1, we prove estimates on hierarchies of differential inequalities. In Subsection 3.2, we use the estimates of Subsection 3.1 together with the "energy estimate" to prove Theorem 1.4.

1.3 Acknowledgments

The authors thank Daniel Lacker and Sylvia Serfaty for stimulating discussions about the presentation and context of the main results. EHC acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), [funding reference number PGSD3-557776-2021]. La recherche de EHC a été financée par le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), [numéro de référence PGSD3-557776-2021]

2. Cluster expansions and perturbation theory

For the remainder of the paper, we suppress the dependence of $f_{j,N}$ on N, simply writing f_j. In order to simplify the presentation of the algebra, we introduce abstract notation for the operators appearing the in BBGKY hierarchy.

Definition 2.1. Let $P \subseteq \mathbb{N} \cup \{*\}$[3] with $|P| < \infty$ and $h : (\mathbb{T}^d)^P \to \mathbb{R}$. Then for any $k, \ell \in P$ such that $k, \ell \neq *$, we define

$$S_{k,\ell} h : (\mathbb{T}^d)^P \to \mathbb{R}$$

by

$$S_{k,\ell} h(x) := \nabla_{x_k} \cdot (K(x_k, x_\ell) h(x)).$$

Then, provided $* \in P$, for any $k \in P$ such that $k \neq *$, we define

$$H_k h : (\mathbb{T}^d)^{P - \{*\}} \to \mathbb{R}$$

by

$$H_k h(x^{P-\{*\}}) := \nabla_{x_k} \cdot \int K(x_k, x_*) h(x) \, dx_*.$$

With this notation, we can rewrite the BBGKY hierarchy (1.3) abstractly as

$$\partial_t f_j - \Delta f_j + \frac{N-j}{N} \sum_{k \in [j]} H_k f_{[j] \cup \{*\}} + \frac{1}{N} \sum_{k,\ell \in [j]} S_{k,\ell} f_j = 0. \tag{2.1}$$

2.1 Cluster expansion

We now introduce the cluster expansion of the f_j.

Definition 2.2. Let $g_j : \mathbb{T}^{jd} \to \mathbb{R}$ be the exchangeable functions given by

$$g_j := \sum_{\pi \vdash [j]} (-1)^{|\pi|-1} (|\pi| - 1)! \prod_{P \in \pi} f_P. \tag{2.2}$$

We call the function g_j the *jth cluster function of the distribution.*

[3]Note here we are taking $*$ as some index distinct from $k \in \mathbb{N}$. It will always be used for the non-local operator H_k, for which it acts an index for the integration variable.

Remark 2.3. We note the expression of g_j in terms of the f_k is analogous to the expression the joint cumulant of a collection of j random variable in terms of their joint moments.

Although this dependence is suppressed, the g_j depend on N through their dependence on the f_j. The g_j are defined exactly so that the following expansion for the f_j holds.

Proposition 2.4.

$$f_j = \sum_{\pi \vdash [j]} \prod_{P \in \pi} g_P.$$

This is a classical combinatorial fact frequently used to relate moments and cumulants of random variables. The proof is found in Section 4.

Remark 2.5. The g_j have the marginalization property that for any $j \geqslant 2$ and any $1 \leqslant \ell \leqslant j$,

$$\int g_j \, dx_\ell = 0.$$

We don't need this property, so we omit its proof. The argument uses the same elementary combinatorics as the rest of the proofs of the results of this section. We will however use the same marginalization property for the terms perturbative expansion of the g_j, g_j^i, which is noted in Proposition 2.10.

By taking the time derivative of (2.2) and using the BBGKY hiearchy (2.1), we see that the g_j themselves solve equations, which we now give.

Proposition 2.6. *For fixed N, the cluster functions g_j, $1 \leqslant j \leqslant N$, solve the hierarchy of equations*

$$\begin{aligned}
\partial_t g_j - \Delta g_j = &-\frac{N-j}{N} \sum_{k=1}^{j} H_k g_{[j]\cup\{*\}} + \sum_{k=1}^{j} \sum_{W \subseteq [j]-\{k\}} \frac{j-1-|W|}{N} H_k g_{W\cup\{k,*\}} g_{[j]-\{k\}-W} \\
&- \frac{N-j}{N} \sum_{k=1}^{j} \sum_{W \subseteq [j]-\{k\}} H_k g_{W\cup\{k\}} g_{[j]\cup\{*\}-W-\{k\}} \\
&+ \sum_{k=1}^{j} \sum_{W \subseteq [j]-\{k\}} \sum_{R \subseteq [j]-\{k\}-W} \frac{j-1-|W|-|R|}{N} H_k g_{W\cup\{k\}} g_{R\cup\{*\}} g_{[j]-R-W-\{k\}} \\
&- \frac{1}{N} \sum_{k,\ell=1}^{j} S_{k,\ell} g_j - \frac{1}{N} \sum_{\substack{k,\ell=1 \\ k\neq \ell}}^{j} \sum_{W \subseteq [j]-\{k,\ell\}} S_{k,\ell} g_{W\cup\{k\}} g_{[j]-\{k\}-W},
\end{aligned} \tag{2.3}$$

with initial conditions

$$g_j(0,\cdot) = \begin{cases} f & j = 0, \\ 0 & j \geqslant 1. \end{cases}$$

The proof of this proposition involves expanding out the g_j in terms of f_i, using the equations for f_i, and then re-expanding the f_i in terms of g_k. One must then carefully collect constant factors before identical terms. We defer the unenlightening proof to Section 4.

2.2 Perturbative expansion of the cluster functions

We note this subsection primarily consists of formal arguments, motivating the correct equations for g_j^i and f_j^i. The actual analytic content of this section isn't realized until we prove that this formal perturbation theory gives good approximations to the true marginal densities in Section 3.

We are interested in computing solutions to this hierarchy perturbatively in N^{-1}. Thus we now take the perturbative ansatz for g_j,

$$g_j = \sum_{i=0}^{\infty} N^{-i} g_j^i,$$

where the g_i^j are assumed to be N-independent. Plugging this into (2.3) and collecting orders of N^{-1}, we find that such g_j^i should be solutions to

$$\begin{aligned}
&\partial_t g_j^i - \Delta g_j^i + \sum_{k=1}^{j} H_k g_{\{k\}}^0 g_{[j]\cup\{*\}-\{k\}}^i + \sum_{k=1}^{j} H_k g_{[j]}^i g_{\{*\}}^0 \\
&\quad = -\sum_{k=1}^{j} H_k g_{[j]\cup\{*\}}^i - \sum_{k=1}^{j} \sum_{W\subseteq[j]-\{k\}} \sum_{m=1}^{i-1} H_k g_{W\cup\{k\}}^m g_{[j]\cup\{*\}-W-\{k\}}^{i-m} \\
&\qquad + j\sum_{k=1}^{j} H_k g_{[j]\cup\{*\}}^{i-1} + \sum_{k=1}^{j} \sum_{W\subseteq[j]-\{k\}} (j-1-|W|) \sum_{m=0}^{i-1} H_k g_{W\cup\{k,*\}}^m g_{[j]-\{k\}-W}^{i-1-m} \\
&\qquad + j\sum_{k=1}^{j} \sum_{W\subseteq[j]-\{k\}} \sum_{m=0}^{i-1} H_k g_{W\cup\{k\}}^m g_{[j]\cup\{*\}-W-\{k\}}^{i-1-m} \\
&\qquad + \sum_{k=1}^{j} \sum_{W\subseteq[j]-\{k\}} \sum_{R\subseteq[j]-\{k\}-W} (j-1-|W|-|R|) \sum_{m=0}^{i-1} \sum_{n=0}^{i-1-m} H_k g_{W\cup\{k\}}^m g_{R\cup\{*\}}^n g_{[j]-R-W-\{k\}}^{i-1-m-n} \\
&\qquad - \sum_{k,\ell=1}^{j} S_{k,\ell} g_j^{i-1} - \sum_{\substack{k,\ell=1\\k\neq\ell}}^{j} \sum_{W\subseteq[j]-\{k,\ell\}} \sum_{m=0}^{i-1} S_{k,\ell} g_{W\cup\{k\}}^m g_{[j]-\{k\}-W}^{i-1-m},
\end{aligned} \tag{2.4}$$

where we take the convention that $g_j^{-1} = 0$ for any j and $g_0^i = 0$ for any i. We also find that they should have initial conditions

$$g_j^i(0,\cdot) = \begin{cases} f & i=0, j=1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.5}$$

Now that we have a representation of the perturbative expansion for the cluster functions g_j, we turn our attention back to the marginals f_j. We seek a representation of their perturbative expansion. To that end we write the formal expansions

$$\sum_{i=0}^{\infty} N^{-i} f_j^i = f_j = \sum_{\pi\vdash[j]} \prod_{P\in\pi} g_P = \sum_{\pi\vdash[j]} \prod_{P\in\pi} \sum_{i_P=0}^{\infty} N^{-i_P} g_P^{i_P} = \sum_{i=0}^{\infty} N^{-i} \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P = i}} \prod_{P\in\pi} g_P^{i_P}.$$

Collecting terms by order, we get

$$f_j^i := \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P = i}} \prod_{P\in\pi} g_P^{i_P}.$$

Further, plugging the perturbative expansion for $f_j = \sum_{i=0}^{\infty} N^{-i} f_j^i$ into the BBGKY hierarchy (2.1) and collecting orders, we formally get

$$\partial_t f_j^i - \Delta f_j^i + \sum_{k=1}^{j} H_k f_{[j]\cup\{*\}}^i = j \sum_{k=1}^{j} H_k f_{[j]\cup\{*\}}^{i-1} - \sum_{k,\ell=1}^{j} S_{k,\ell} f_j^{i-1}.$$

Since the f_j^i are defined in terms of the g_j^i, which themselves solve equations, we need to check that under our definition of the f_j^i, this equation is in fact solved, as is given by the next proposition.

Proposition 2.7. *Let f_j^i be defined by (1.5) where the g_j^i solve the hierarchy (2.4). Then the f_j^i solve the hierarchy of equations (1.6).*

Remark 2.8. Theorem 1.3 is an immediate consequence of this proposition and Proposition 1.2, which will be proved in the next subsection.

The proof of Proposition 2.7 also proceeds by tedious but elementary algebraic manipulation, so has been deferred to Section 4.

2.3 Existence and basic properties of the g_j^i

As we will see in the proof of Proposition 1.2, g_0^1 will solve the equation

$$\begin{cases} \partial_t g_0^1 - \Delta g_0^1 + \nabla \cdot \int K(x,x_*) g_0^1(x_*) g_0^1(x)\, dx_* = 0 \\ g_0^1(0,\cdot) = f. \end{cases}$$

This makes g_0^1 special among the g_j^i in two ways, first it is the only g_j^i whose equation is nonlinear in g_j^i and second it is the only g_j^i with non-trivial initial data. We note that g_0^1 is the mean-field limit and its equation is the McKean-Vlasov equation. While existence theory for this equation is well known, it is mostly done from the probabilistic perspective, showing the existence of solutions to the associated McKean-Vlasov SDE, e.g. as in [MV21]. While the PDE existence can be deduced from the SDE existence, in order to make this presentation more self-contained, we give a purely PDE argument for the existence of solutions. The proof follows standard PDE arguments and so is moved to the end of the paper, Appendix A.

Proposition 2.9. *For $f \in L^2(\mathbb{T}^d)$ and $K \in L^\infty(\mathbb{T}^{2d})$, there exists a unique $\rho \in C([0,\infty), L^2(\mathbb{T}^d)) \cap L^2_{loc}([0,T], H^1(\mathbb{T}^d))$ such that*

$$\begin{cases} \partial_t \rho - \Delta \rho + \nabla \cdot \int K(x,x_*) \rho(x_*) \rho(x)\, dx_* \\ \rho(0,\cdot) = f. \end{cases}$$

For the remainder section we take ρ to be the unique solution to the McKean-Vlasov equation given by Proposition 2.9. Now that we have a solution to the McKean-Vlasov equation, we can prove that there actually is a solution to the hierarchy (2.4), which is the content of Proposition 1.2.

Proof of Proposition 1.2. The proof proceeds in two steps.

Step 1: We check that if we take $g_j^i = 0$ for all $(i,j) \notin T$, then this does not contradict the equations (2.4). That is to say, we just need to verify that if $(i,j) \notin T$, then all terms in the right hand side of the equation for g_j^i involve g_ℓ^k for some $(k,\ell) \notin T$. This is easy to check for all terms which do not involve products of the g_ℓ^k. There are 5 terms which do involve products. We check the first term (in the order they appear in (2.4)), which already sharply generates the constraint $j \geqslant i+2$, the others follow similarly. Consider

$$H_k g^m_{W\cup\{k\}} g^{i-m}_{[j]\cup\{*\}-W-\{k\}},$$

for $k \in [j], W \subseteq [j]-\{k\}$. Then the only way $(m, |W|+1)$ and $(i-m, j-|W|)$ are both in T is if

$$|W|+1 \leqslant m+1 \text{ and } j-|W| \leqslant i-m+1.$$

Simplifying and combining these constraints we find that $j \leqslant i+1$, which contradicts the assumption that $(i,j) \notin T$. The analysis of all other terms follows directly analogously.[4]

Step 2: Using Step 1, we will now make the *a priori* assumption that $g^i_j = 0$ for all $(i,j) \notin T$. We will inductively show unique existence for function g^i_j with $(i,j) \in T$ using the ordering defined above.

First, in the base case $(0,1)$, the equation for g^0_1 reduces to the McKean-Vlasov equation (1.4). Proposition 2.9 implies that there is a unique solution to g^0_1, namely ρ.

Now, assuming that g^k_ℓ have been shown to uniquely exist for $(k,\ell) < (i,j) \in T$, we consider the equation for g^i_j. We note that all the terms on the right hand side of the equation only involve terms which are zero or satisfy $(k,\ell) < (i,j)$, while the terms on the left hand side are linear in g^i_j. Standard parabolic existence theory (for example [LM72]) gives unique existence of a solution to (2.4) in $L^2_{loc}([0,\infty), H^1(\mathbb{T}^d)) \cap L^\infty_{loc}([0,\infty), L^2(\mathbb{T}^d))$. This completes the induction. $\square$

Now that we have constructed a solution to the hierarchy (2.4), we wish to show that the g^i_j have the same marginalization properties as the g_j. This is shown by inductively using the Grönwall inequality, where the induction is done in the ordering on T.

Proposition 2.10. *For $f \in L^2(\mathbb{T}^d)$ and $K \in L^\infty(\mathbb{T}^{2d})$, if g^i_j are as given by Proposition 1.2, then for any i,j and any $1 \leqslant \ell \leqslant j$,*

$$\int g^i_j \, dx_\ell = \begin{cases} 1 & i=0, j=1, \\ 0 & \textit{otherwise.} \end{cases}$$

Thus

$$\int f^i_{j+1} \, dx_{j+1} = f^i_j \textit{ and } \int \varphi^i_{j+1} \, dx_{j+1} = \varphi^i_j.$$

Proof. We will show this inductively using the order given in Definition 1.1. The base case holds trivially since $g^0_1 = \rho$ is a probability density.

Fixing (i,j) such that $(i,j) \geqslant (0,1)$, suppose now that the marginalization holds for all g^k_ℓ with $(k,\ell) < (i,j)$. We define

$$\psi(x_1, ..., x_{j-1}) := \int g^i_j(x) \, dx_j.$$

Then integrating the equation (2.4) over x_j we get

[4] In particular the first, fourth, and fifth terms sharply give the constraint $j \geqslant i+2$, while the second and third terms are not sharp

$$
\begin{aligned}
&\partial_t\psi - \Delta\psi + \sum_{k=1}^{j-1} H_k g^0_{\{k\}}\psi(x^{[j-1]-\{k\}\cup\{*\}}) + \sum_{k=1}^{j-1} H_k \psi(x) g^0_{\{*\}} \\
&\quad = -\sum_{k=1}^{j-1} H_k \int g^i_{[j]\cup\{*\}}\, dx_j - \sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}}\sum_{m=1}^{i-1} H_k \int g^m_{W\cup\{k\}} g^{i-m}_{[j]\cup\{*\}-W-\{k\}}\, dx_j && (2.6)\\
&\qquad + j\sum_{k=1}^{j-1} H_k \int g^{i-1}_{[j]\cup\{*\}}\, dx_j + \sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}} (j-1-|W|)\sum_{m=0}^{i-1} H_k \int g^m_{W\cup\{k,*\}} g^{i-1-m}_{[j]-\{k\}-W}\, dx_j && (2.7)\\
&\qquad + j\sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}}\sum_{m=0}^{i-1} H_k \int g^m_{W\cup\{k\}} g^{i-1-m}_{[j]\cup\{*\}-W-\{k\}}\, dx_j && (2.8)\\
&\qquad + \sum_{k=1}^{j-1}\sum_{\substack{W\subseteq[j]-\{k\}\\ R\subseteq[j]-\{k\}-W}} (j-1-|W|-|R|)\sum_{m=0}^{i-1}\sum_{n=0}^{i-1-m} H_k \int g^m_{W\cup\{k\}} g^n_{R\cup\{*\}} g^{i-1-m-n}_{[j]-R-W-\{k\}}\, dx_j \\
&\qquad - \sum_{k=1}^{j-1}\sum_{\ell=1}^{j}\int S_{k,\ell} g^{i-1}_j\, dx_j - \sum_{k=1}^{j-1}\sum_{\substack{\ell=1\\ k\neq\ell}}^{j}\sum_{W\subseteq[j]-\{k,\ell\}}\sum_{m=0}^{i-1}\int S_{k,\ell} g^m_{W\cup\{k\}} g^{i-1-m}_{[j]-\{k\}-W}\, dx_j. && (2.9)
\end{aligned}
$$

where we've used that

$$\int \nabla_{x_j}\cdot h\, dx_j = 0$$

for any function h. Both sums on line (2.6) are equal to zero by the induction hypothesis as all the superscripts are larger than 1. The induction assumption also implies that the first sum on line (2.7) equals 0 when $i \geqslant 2$ as then the superscript $i-1 \geqslant 1$. When $i = 1$, it also equals 0, but instead because $g^0_{[j]\cup\{*\}} = 0$ as $|[j]\cup\{*\}| \geqslant 2$.

The second sum on line (2.7) will be shown later to cancel with the first sum on line (2.9), so we skip it for now.

For line (2.8), we note that all terms in the sum corresponding to $0 < m < i-1$ equal zero by the induction hypothesis. We are thus left with

$$j\sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}} H_k \int g^0_{W\cup\{k\}} g^{i-1}_{[j]\cup\{*\}-W-\{k\}}\, dx_j + H_k \int g^{i-1}_{W\cup\{k\}} g^0_{[j]\cup\{*\}-W-\{k\}}\, dx_j.$$

The terms in this sum can be broken into two cases, either $j \in W$ or $j \notin W$. If $j \in W$ then $|W \cup \{k\}| \geqslant 2$, thus $g^0_{W\cup\{k\}} = 0$ and

$$\int g^{i-1}_{W\cup\{k\}}\, dx_j = 0,$$

thus all these terms equal to 0. When $j \notin W$, then $|[j]\cup\{*\} - W - \{k\}| \geqslant 2$, thus by an analogous argument all the corresponding terms equal zero as well. This shows that line (2.8) equals zero as well.

We have so far simplified the entire equation to

$$\partial_t\psi - \Delta\psi + \sum_{k=1}^{j-1} H_k g^0_{\{k\}}\psi(x^{[j-1]-\{k\}\cup\{*\}}) + \sum_{k=1}^{j-1} H_k\psi(x)g^0_{\{*\}}$$

$$= \sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}}(j-1-|W|)\sum_{m=0}^{i-1} H_k\int g^m_{W\cup\{k,*\}}g^{i-1-m}_{[j]-\{k\}-W}\,dx_j \tag{2.10}$$

$$+ \sum_{k=1}^{j-1}\sum_{\substack{W\subseteq[j]-\{k\}\\ R\subseteq[j]-\{k\}-W}}(j-1-|W|-|R|)\sum_{m=0}^{i-1}\sum_{n=0}^{i-1-m} H_k\int g^m_{W\cup\{k\}}g^n_{R\cup\{*\}}g^{i-1-m-n}_{[j]-R-W-\{k\}}\,dx_j \tag{2.11}$$

$$- \sum_{k=1}^{j-1}\sum_{\ell=1}^{j}\int S_{k,\ell}g^{i-1}_j\,dx_j \tag{2.12}$$

$$- \sum_{k=1}^{j-1}\sum_{\substack{\ell=1\\ k\neq\ell}}^{j}\sum_{W\subseteq[j]-\{k,\ell\}}\sum_{m=0}^{i-1}\int S_{k,\ell}g^m_{W\cup\{k\}}g^{i-1-m}_{[j]-\{k\}-W}\,dx_j. \tag{2.13}$$

First we claim that the sum (2.10) can be reduced to

$$\sum_{k=1}^{j-1} H_k g^{i-1}_{[j-1]\cup\{*\}}.$$

This is clear when $j = 1$. Using the induction hypothesis when $0 < m < i-1$ we reduce (2.10) to

$$\sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}}(j-1-|W|)\left(H_k\int g^0_{W\cup\{k,*\}}g^{i-1}_{[j]-\{k\}-W}\,dx_j + H_k\int g^{i-1}_{W\cup\{k,*\}}g^0_{[j]-\{k\}-W}\,dx_j\right)$$

Since $|W\cup\{k,*\}| \geqslant 2$, $g^{i-1}_{W\cup\{k,*\}} = 0$ hence this further reduces to

$$\sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}}(j-1-|W|)H_k\int g^{i-1}_{W\cup\{k,*\}}g^0_{[j]-\{k\}-W}\,dx_j.$$

The integral

$$\int g^{i-1}_{W\cup\{k,*\}}g^0_{[j]-\{k\}-W}\,dx_j = 0$$

unless $W = [j-1]-\{k\}$ hence

$$\sum_{k=1}^{j-1}\sum_{W\subseteq[j]-\{k\}}(j-1-|W|)H_k\int g^{i-1}_{W\cup\{k,*\}}g^0_{[j]-\{k\}-W}\,dx_j = \sum_{k=1}^{j-1}H_k\int g^{i-1}_{[j-1]\cup\{*\}}g^0_{\{j\}}\,dx_j$$
$$= \sum_{k=1}^{j-1}H_k g^{i-1}_{[j-1]\cup\{*\}},$$

as claimed. This cancels exactly with (2.12) since if $\ell \neq j$ then

$$\int S_{k,\ell}g^{i-1}_j\,dx_j = 0$$

while when $\ell = j$ exchangeability implies

$$\int S_{k,j} g_j^{i-1} \, dx_j = H_k g_{[j-1] \cup \{*\}}^{i-1}.$$

Similarly, we can reduce (2.11) to

$$\sum_{k=1}^{j-1} \sum_{W \subseteq [j-1]-\{k\}} \sum_{m=0}^{i-1} H_k g_{W \cup \{k\}}^m g_{[j-1]-\{k\}-W \cup \{*\}}^{i-1-m}.$$

Indeed, if $j \in W$ or $j \in R$ then either

$$\int g_{W \cup \{k\}}^m \, dx_j = 0 \text{ or } \int g_{R \cup \{*\}}^n \, dx_j = 0,$$

respectively. When $j \notin R \cup W$ the integral

$$\int g_{[j]-R-W-\{k\}}^{i-1-m-n} \, dx_j = 0$$

unless $i-1-m-n = 0$ and $R \cup W = [j-1] - \{k\}$. Thus

$$\begin{aligned} &\sum_{k=1}^{j-1} \sum_{W \subseteq [j]-\{k\}} \sum_{R \subseteq [j]-\{k\}-W} (j-1-|W|-|R|) \sum_{m=0}^{i-1} \sum_{n=0}^{i-1-m} H_k \int g_{W \cup \{k\}}^m g_{R \cup \{*\}}^n g_{[j]-R-W-\{k\}}^{i-1-m-n} \, dx_j \\ &\quad = \sum_{k=1}^{j-1} \sum_{W \subseteq [j-1]-\{k\}} (j-1-(j-2)) \sum_{m=0}^{i-1} H_k \int g_{W \cup \{k\}}^m g_{[j-1]-W-\{k\} \cup \{*\}}^{i-1-m} g_{\{j\}}^0 \, dx_j \\ &\quad = \sum_{k=1}^{j-1} \sum_{W \subseteq [j-1]-\{k\}} \sum_{m=0}^{i-1} H_k g_{W \cup \{k\}}^m g_{[j-1]-W-\{k\} \cup \{*\}}^{i-1-m}. \end{aligned} \tag{2.14}$$

This then will cancel with (2.13). To see this not that when $\ell \neq j$,

$$\int S_{k,\ell} g_{W \cup \{k\}}^m g_{[j]-\{k\}-W}^{i-1-m} \, dx_j = 0$$

since if $j \in W$ then $|W \cup \{k\}| \geqslant 2$ and if $j \in [j] - \{k\} - W$ then $|[j] - \{k\} - W| \geqslant 2$. The sum (2.13) thus reduces to

$$\sum_{k=1}^{j-1} \sum_{W \subseteq [j]-\{k,\ell\}} \sum_{m=0}^{i-1} \int S_{k,j} g_{W \cup \{k\}}^m g_{[j]-\{k\}-W}^{i-1-m} \, dx_j$$

which is then equal to (2.14) by exchangeability, so the terms cancel exactly.

We have thus shown that

$$\partial_t \psi - \Delta \psi + \sum_{k=1}^{j-1} H_k g_{\{k\}}^0 \psi(x^{[j-1]-\{k\} \cup \{*\}}) + \sum_{k=1}^{j-1} H_k \psi g_{\{*\}}^0 = 0.$$

The claim is then completed by a Grönwall argument on $\|\psi\|_{L^2}^2(t)$ using that $\psi(0, \cdot) = 0$. Using the marginalization of the g_j^i to give the marginalization of the f_j^i and the φ_j^i is direct from the definition (1.5) of f_j^i and then the definition (3.1) of φ_j^i. $\square$

3. Hierarchy bounds

For this section, we let g_j^i be the unique family of functions solving (2.4) given by Proposition 1.2. We will also need the stronger assumptions on f throughout, namely that $f \in L^\infty$ and that there exists $m > 0$ such that $f \geqslant m$. Given these assumptions, we note that by basic parabolic theory applied to the McKean-Vlasov equation, since the initial data is upper and lower bounded, we get that $\rho + \rho^{-1} \in L^\infty_{loc}([0,\infty), L^\infty(\mathbb{T}^d))$. Additionally, using the Liouville equation and marginalization, we see that $f_j \in L^\infty_{loc}([0,\infty), L^\infty(\mathbb{T}^d))$. Similarly, using the equations for the g_j^i, we see that $g_j^i \in L^\infty_{loc}([0,\infty), L^\infty(\mathbb{T}^d))$. We omit these arguments as they are standard, and we only need these bounds qualitatively to ensure all the integrals are finite.

Having constructed the g_j^i and f_j^i and shown basic properties of them, we are now prepared to show the main result, which appropriately controls the error between f_j and its approximation to order i. First let's introduce some more notation.

Definition 3.1. Letting f_j^i be defined by (1.5), we let

$$\varphi_j^i := \sum_{k=0}^{i} N^{-k} f_j^k, \tag{3.1}$$

$$R_j^i := \frac{1}{N^{i+1}} \sum_{k=1}^{j} e_k \otimes \sum_{\ell=1}^{j} \int K(x_k, x_*) f^i_{[j]\cup\{*\}} dx_* - K(x_k, x_\ell) f_j^i.$$

Remark 3.2. The tensor product notation used in the definition of R_j^i is given such that

$$\nabla \cdot R_j^i = \frac{1}{N^{i+1}} \sum_{k=1}^{j} \nabla_{x_k} \cdot \sum_{\ell=1}^{j} \int K(x_k, x_*) f^i_{[j]\cup\{*\}} dx_* - K(x_k, x_\ell) f_j^i,$$

where $\nabla \cdot$ denotes the divergence on $\mathbb{T}^{jd}$.

One can readily check using the equations the f_j^i solve that φ_j^i solves the following equation.

$$\partial_t \varphi_j^i - \Delta \varphi_j^i + \frac{N-j}{N} \sum_k \nabla_{x_k} \cdot \int K(x_k, x_*) \varphi^i_{j+1}(x^{[j]\cup\{*\}})\, dx_* + \frac{1}{N} \sum_{k,\ell=1}^{j} \nabla_{x_k} \cdot (K(x_k, x_\ell) \varphi_j^i) = \nabla \cdot R_j^i. \tag{3.2}$$

We now show the essential L^2 energy-type estimate for difference $\varphi_j^i - f_j$. We note that at $t = 0$, $\varphi_j^i = f_j$, so this estimate allows us to control the size of $\varphi_j^i - f_j$ for $t > 0$ by a Grönwall-type argument. We also give a somewhat brutal bound on the growth of f_j that doesn't depend on φ_j^i. We will use this brutal bound to "close" the hierarchy.

Proposition 3.3. *Suppose $f \in L^\infty(\mathbb{T}^d), K \in L^\infty(\mathbb{T}^{2d})$, and there exists $m > 0$ such that $f \geqslant m$. Then letting*

$$\gamma_j^i := \varphi_j^i - f_j,$$

we have that

$$\frac{d}{dt} \int \left|\frac{\gamma_j^i}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\, dx \leqslant 2j \|K\|^2_{L^\infty} \left(\int \left|\frac{\gamma^i_{j+1}}{\rho^{\otimes(j+1)}}\right|^2 \rho^{\otimes(j+1)}\, dx_* dx - \int \left|\frac{\gamma_j^i}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\, dx \right)$$
$$+ 4\frac{j^3}{N^2} \|K\|^2_{L^\infty} \int \left|\frac{\gamma_j^i}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\, dx + 2 \int \left|\frac{R_j^i}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\, dx. \tag{3.3}$$

We also have that
$$\frac{d}{dt}\int\Big|\frac{f_j}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx \leqslant 12j\|K\|_{L^\infty}^2\int\Big|\frac{f_j}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx. \tag{3.4}$$

Remark 3.4. Note the interesting—and essential—property of this estimate that all constants are independent of ρ.

Proof. For notational simplicity, let us fix $i \in \mathbb{N}$ and drop the i dependence, writing γ_j for γ_j^i. We note that γ_j solves the equation

$$\partial_t\gamma_j - \Delta\gamma_j + \frac{N-j}{N}\sum_{k=1}^{j}\nabla\cdot\int K(x_k,x_*)\gamma_{j+1}(x^{[j]\cup\{*\}})\,dx_* + \frac{1}{N}\sum_{k,\ell=1}^{j}\nabla_{x_k}\cdot(K(x_k,x_\ell)\gamma_j) = \nabla\cdot R_j,$$

where $R_j = R_j^i$. We also have that

$$\partial_t\rho^{\otimes j} - \Delta\rho^{\otimes j} + \sum_{k=1}^{j}\nabla\cdot\int K(x_k,x_*)\rho(x_*)\,dx_*\rho^{\otimes j} = 0.$$

We then compute

$$\begin{aligned}
\frac{d}{dt}\int\frac{\gamma_j^2}{\rho^{\otimes j}}\,dx &= \int 2\frac{\gamma_j}{\rho^{\otimes j}}\partial_t\gamma_j - \frac{\gamma_j^2}{(\rho^{\otimes j})^2}\partial_t\rho^{\otimes j}\,dx\\
&= \int -2\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\nabla\gamma_j + 2\frac{N-j}{N}\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\sum_{k=1}^{j}e_k\int K(x_k,x_*)\gamma_{j+1}(x^{[j]\cup\{*\}})\,dx_*\\
&\quad + \frac{2}{N}\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\sum_{k,\ell=1}^{j}e_kK(x_k,x_\ell)\gamma_j - 2\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot R_j\\
&\quad + 2\frac{\gamma_j}{\rho^{\otimes j}}\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\nabla\rho^{\otimes j} - 2\frac{\gamma_j}{\rho^{\otimes j}}\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\sum_{k=1}^{j}e_k\int K(x_k,x_*)\rho(x_*)\,dx_*\rho^{\otimes j}\,dx.
\end{aligned}$$

We note that
$$\rho^{\otimes j}\nabla\frac{\gamma_j}{\rho^{\otimes j}} = \nabla\gamma_j - \frac{\gamma_j}{\rho^{\otimes j}}\nabla\rho^{\otimes j}.$$

Thus
$$\int -2\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\nabla\gamma_j + 2\frac{\gamma_j}{\rho^{\otimes j}}\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\nabla\rho^{\otimes j}\,dx = -2\int\Big|\nabla\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx.$$

We then group terms,

$$\begin{aligned}
\frac{d}{dt}\int\frac{\gamma_j^2}{\rho^{\otimes j}}\,dx &= -2\int\Big|\nabla\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx\\
&\quad + 2\frac{N-j}{N}\int\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\sum_{k=1}^{j}e_k\int K(x_k,x_*)\Big(\frac{\gamma_{j+1}}{\rho^{\otimes(j+1)}} - \frac{\gamma_j}{\rho^{\otimes j}}\Big)\rho(x_*)\,dx_*\ \rho^{\otimes j}\,dx\\
&\quad - 2\frac{j}{N}\int\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\sum_{k=1}^{j}e_k\int K(x_k,x_*)\rho(x_*)\,dx_*\frac{\gamma_j}{\rho^{\otimes j}}\rho^{\otimes j}\,dx\\
&\quad + \frac{2}{N}\int\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\sum_{k,\ell=1}^{j}e_kK(x_k,x_\ell)\frac{\gamma_j}{\rho^{\otimes j}}\rho^{\otimes j}\,dx - 2\int\nabla\frac{\gamma_j}{\rho^{\otimes j}}\cdot\frac{R_j}{\rho^{\otimes j}}\rho^{\otimes j}\,dx.
\end{aligned}$$

Thus applying Young's inequality, we see that

$$\frac{d}{dt}\int \frac{\gamma_j^2}{\rho^{\otimes j}}\,dx \leqslant 2j\int\Big|\int K(x_1,x_*)\Big(\frac{\gamma_{j+1}}{\rho^{\otimes(j+1)}}-\frac{\gamma_j}{\rho^{\otimes j}}\Big)\rho(x_*)\,dx_*\Big|^2\rho^{\otimes j}\,dx$$
$$+\,4\frac{j^3}{N^2}\|K\|_{L^\infty}^2\int\frac{\gamma_j^2}{\rho^{\otimes j}}\,dx+2\int\Big|\frac{R_j}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx.$$

We then note that by Hölder's inequality,

$$\begin{aligned}
&\Big|\int K(x_1,x_*)\Big(\frac{\gamma_{j+1}}{\rho^{\otimes(j+1)}}-\frac{\gamma_j}{\rho^{\otimes j}}\Big)\rho(x_*)\,dx_*\Big|^2\\
&\quad\leqslant \int K(x_1,x_*)^2\rho(x_*)\,dx_*\int\Big(\frac{\gamma_{j+1}}{\rho^{\otimes(j+1)}}-\frac{\gamma_j}{\rho^{\otimes j}}\Big)^2\rho(x_*)\,dx_*\\
&\quad\leqslant \|K\|_{L^\infty}^2\left(\int\Big|\frac{\gamma_{j+1}}{\rho^{\otimes(j+1)}}\Big|^2\rho(x_*)\,dx_*-2\frac{\gamma_j}{(\rho^{\otimes j})^2}\int\gamma_{j+1}\,dx_*+\Big|\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2\int\rho(x_*)\,dx_*\right)\\
&\quad= \|K\|_{L^\infty}^2\left(\int\Big|\frac{\gamma_{j+1}}{\rho^{\otimes(j+1)}}\Big|^2\rho(x_*)\,dx_*-\Big|\frac{\gamma_j}{\rho^{\otimes j}}\Big|^2\right),
\end{aligned}$$

where we use Proposition 2.10 for the last line. Combining this with the previous inequality, we get (3.3). Turning our attention (3.4), we note that repeating that above computations with f_j in place of γ_j, we get that

$$\frac{d}{dt}\int\frac{f_j^2}{\rho^{\otimes j}}\,dx\leqslant 2j\int\Big|\int K(x_1,x_*)\Big(\frac{f_{j+1}}{\rho^{\otimes(j+1)}}-\frac{f_j}{\rho^{\otimes j}}\Big)\rho(x_*)\,dx_*\Big|^2\rho^{\otimes j}\,dx+4\frac{j^3}{N^2}\|K\|_{L^\infty}^2\int\frac{f_j^2}{\rho^{\otimes j}}\,dx.$$

Then we note that

$$\Big|\int K(x_1,x_*)\Big(\frac{f_{j+1}}{\rho^{\otimes(j+1)}}-\frac{f_j}{\rho^{\otimes j}}\Big)\rho(x_*)\,dx_*\Big|\leqslant\|K\|_{L^\infty}\Big(\frac{f_j}{\rho^{\otimes j}}+\frac{1}{\rho^{\otimes j}}\int f_{j+1}\,dx_*\Big)=2\|K\|_{L^\infty}\frac{f_j}{\rho^{\otimes j}}.$$

Thus

$$\frac{d}{dt}\int\frac{f_j^2}{\rho^{\otimes j}}\,dx\leqslant 8j\|K\|_{L^\infty}^2\int\Big|\frac{f_j}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx+4\frac{j^3}{N^2}\|K\|_{L^\infty}^2\int\frac{f_j^2}{\rho^{\otimes j}}\,dx\leqslant 12j\|K\|_{L^\infty}^2\int\Big|\frac{f_j}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx,$$

giving (3.4). ☐

With the energy estimate (3.3) in hand, we now need to understand how to bound hierarchies of differential inequalities of the above sort. This is the focus of Subsection 3.1. Before that though, we need to note bounds on the terms involved. The bound given by Proposition 3.6 is essential for estimating the contribution of the remainder terms of (3.3); the bounds of Proposition 3.5 will turn out to be useful as well for somewhat subtler reasons. We defer the combinatorial proofs of these bounds to Subsection 3.3 so as not to distract from the heart of the argument.

Proposition 3.5. *Suppose $f\in L^\infty(\mathbb{T}^d), K\in L^\infty(\mathbb{T}^{2d})$, and there exists $m>0$ such that $f\geqslant m$, then there exists $C(\|K\|_{L^\infty},i)<\infty$ such that*

$$\int\Big|\frac{f_j^i}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx\leqslant Ce^{Ct}j^{2i},$$

and so

$$\int\Big|\frac{\varphi_j^i}{\rho^{\otimes j}}\Big|^2\rho^{\otimes j}\,dx\leqslant Ce^{Ct}.$$

Proposition 3.6. *Suppose $f \in L^\infty(\mathbb{T}^d), K \in L^\infty(\mathbb{T}^{2d})$, and there exists $m > 0$ such that $f \geqslant m$, then there exists $C(\|K\|_{L^\infty}, i) < \infty$ such that*

$$\int \left|\frac{R_j^i}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\, dx \leqslant Ce^{Ct}\Big(\frac{j}{N}\Big)^{2(i+1)}.$$

Remark 3.7. We note that the j dependence of the above bounds is a consequence of certain L^2 orthogonality implicit in the definitions of the f_j^i, which in turn is a consequence of the 0 marginalization property of the g_j^i noted in Proposition 2.10. Without exploiting this L^2 orthogonality, one can derive similar bounds but with worse rates in j. This then propagates to worse rates in j in Theorem 1.4. Thus the marginalization property of Proposition 2.10 is essential to getting good bounds in j.

3.1 ODE hierarchy estimates

We now consider passing estimates on hierarchies of differential inequalities of the form (3.3). By repeatedly applying the Grönwall inequality to the hierarchy, iterated exponential integrals appear. We introduce the following notation for these integrals.

Definition 3.8. For $\beta := 4\|K\|_{L^\infty}^2$, let

$$I_j^\ell(t) := \beta^\ell \frac{(j+\ell-1)!}{(j-1)!} e^{-\beta j t}\int_0^t \int_0^{s_\ell} \cdots \int_0^{s_2} e^{-\beta \sum_{k=2}^\ell s_k} e^{\beta(j+\ell-1)s_1}\, ds_1 \cdots ds_\ell,$$

where $I_j^0(t) := 1$ by convention.

The I_j^ℓ are related in the following way.

Proposition 3.9.

$$\beta j e^{-\beta j t}\int_0^t e^{\beta j s} I_{j+1}^\ell(s)\, ds = I_j^{\ell+1}(t).$$

Proof. If $\ell = 0$, we see that

$$\beta j e^{-\beta j t}\int_0^t e^{\beta j s} I_{j+1}^0(s)\, ds = \beta j e^{-\beta j t}\int_0^t e^{\beta(j+1-1)s_1}\, ds_1 = I_j^1.$$

Otherwise, we compute

$$\begin{aligned}
&\beta j e^{-\beta j t}\int_0^t e^{\beta j s_{\ell+1}} I_{j+1}^\ell(s_{\ell+1})\, ds_{\ell+1} \\
&\quad= \beta^{\ell+1}\frac{(j+\ell)!}{(j-1)!} e^{-\beta j t}\int_0^t e^{\beta j s_{\ell+1}} e^{-\beta(j+1)s_{\ell+1}} \int_0^{s_{\ell+1}} \cdots \int_0^{s_2} e^{-\beta\sum_{k=2}^\ell s_k} e^{\beta(j+1+\ell-1)s_1}\, ds_1 \cdots ds_\ell ds_{\ell+1} \\
&\quad= \beta^{\ell+1}\frac{(j+\ell+1-1)!}{(j-1)!} e^{-\beta j t}\int_0^t \int_0^{s_{\ell+1}} \cdots \int_0^{s_2} e^{-\beta\sum_{k=2}^{\ell+1} s_k} e^{\beta(j+\ell+1-1)s_1}\, ds_1 \cdots ds_{\ell+1} \\
&\quad= I_j^{\ell+1}(t),
\end{aligned}$$

as desired. ∎

It is prefactors of $I_j^\ell(t)$ that will give sufficient decay when iterating up the hierarchy to prove the bounds we require. As such, we need to understand how the I_j^ℓ decay as ℓ gets large. The following proposition is the first such estimate and follows from a simple induction.

Proposition 3.10. *For any $j, \ell \in \mathbb{N}$ and any $b \in \mathbb{N}$,*

$$I_j^\ell(t) \leqslant \Big(\frac{j+b}{j+\ell}\Big)^b e^{\beta b t}.$$

Proof. We note that for any $b \geqslant 0$,

$$\begin{aligned}
&\int_0^t \int_0^{s_\ell} \cdots \int_0^{s_2} e^{-\beta \sum_{k=2}^\ell s_k} e^{\beta(j+\ell-1)s_1}\, ds_1 \cdots ds_\ell \\
&\quad \leqslant \int_0^t \int_0^{s_\ell} \cdots \int_0^{s_3} e^{-\beta \sum_{k=2}^\ell s_k} \int_0^{s_2} e^{\beta(j+\ell-1+b)s_1}\, ds_1 \cdots ds_\ell \\
&\quad \leqslant \frac{1}{\beta(j+\ell-1+b)} \int_0^t \int_0^{s_\ell} \cdots \int_0^{s_3} e^{-\beta \sum_{k=2}^\ell s_k} e^{\beta(j+\ell-1+b)s_2} ds_2 \cdots ds_\ell \\
&\quad = \frac{1}{\beta(j+\ell-1+b)} \int_0^t \int_0^{s_\ell} \cdots \int_0^{s_4} e^{-\beta \sum_{k=3}^\ell s_k} \int_0^{s_3} e^{\beta(j+\ell-2+b)s_2} ds_2 \cdots ds_\ell \\
&\quad \leqslant \cdots \leqslant e^{\beta(j+b)t} \prod_{i=1}^{\ell} \frac{1}{\beta(j+\ell-i+b)} = e^{\beta(j+b)t} \beta^{-\ell} \prod_{i=0}^{\ell-1} \frac{1}{j+i+b}.
\end{aligned}$$

Thus, for $b \in \mathbb{N}$, exploiting cancellation in the product,

$$I_j^\ell(t) \leqslant e^{\beta b t} \prod_{i=0}^{\ell-1} \frac{j+i}{j+i+b} \leqslant \Big(\frac{j+b}{j+\ell}\Big)^b e^{\beta b t},$$

allowing us to conclude. □

For some of the estimates, the above polynomial decay will be sufficient; for others, we will need an exponential rate of decay. This exponential rate can be found by simply choosing the polynomial power b optimally in a time dependent way, as the below proposition shows.

Proposition 3.11. *For any $j, \ell \in \mathbb{N}$ and for any $t \geqslant 0$, if*

$$j \leqslant \tfrac{1}{3} e^{-2\beta t - 1} \ell,$$

then

$$I_j^\ell(t) \leqslant \exp(-\tfrac{1}{3} e^{-2\beta t - 1} \ell).$$

Remark 3.12. The above proposition is analogous to [Lac23, Proposition 5.1], although with a different proof using elementary techniques.

Proof. Let

$$\delta := \tfrac{1}{3} e^{-2\beta t - 1}.$$

We note that

$$1 \leqslant j \leqslant \delta \ell,$$

thus $\lceil \delta \ell \rceil \leqslant 2\delta\ell$. Then, letting $b = \lceil \delta \ell \rceil$, by Proposition 3.10 we have that

$$I_j^\ell(t) \leqslant \Big(\frac{j+b}{j+\ell}\Big)^b e^{\beta b t} \leqslant (3\delta)^{\lceil \delta \ell \rceil} e^{2\beta\delta\ell t} \leqslant \exp(\delta\ell(2\beta t + \log(3\delta))) = e^{-\delta\ell},$$

where we use that by definition

$$2\beta t + \log(3\delta) = -1.$$

Plugging the definition of δ into the bound, we conclude. □

Now that we have some control on the I_j^ℓ, we are ready to bound the hierarchies of differential inequalities. The following is the first step to getting the correct bound, given by inductively applying Grönwall's inequality and using Proposition 3.9. The first bound (3.5) is sufficient to give Theorem 1.4 for short times, but the second bound (3.6) is necessary to get the result for all times.

Proposition 3.13. *Suppose $x_k \geqslant 0$ satisfy the hierarchy of differential inequalities*

$$\begin{cases} \dot{x}_k \leqslant \beta k(\alpha_k x_{k+1} - x_k) + r_k \\ x_k(0) = 0, \end{cases}$$

for some $r_k, \alpha_k \geqslant 0$ constants. Then we have the bounds for any $j, \ell \in \mathbb{N}$, for any $t \geqslant 0$

$$x_j(t) \leqslant A_j^\ell I_j^\ell(t) \sup_{s\in[0,t]} x_{j+\ell}(s) + \frac{1}{\beta}\sum_{k=0}^{\ell-1} A_j^{k+1} I_j^{k+1}(t) \frac{r_{j+k}}{\alpha_{j+k}(j+k)}, \tag{3.5}$$

and for any $t_0 \geqslant 0, t \geqslant t_0$,

$$\begin{aligned} x_j(t) \leqslant A_j^\ell I_j^\ell(t-t_0) \sup_{s\in[t_0,t]} x_{j+\ell}(s) + \sum_{k=1}^{\ell} A_j^\ell I_{j+\ell-k}^k(t_0) I_j^{\ell-k}(t-t_0) \sup_{s\in[0,t_0]} x_{j+\ell}(s) \\ + \frac{1}{\beta}\sum_{k=0}^{\ell-1} A_j^{k+1} I_j^{k+1}(t) \frac{r_{j+k}}{\alpha_{j+k}(j+k)}, \end{aligned} \tag{3.6}$$

where

$$A_j^k := \prod_{i=j}^{j+k-1} \alpha_i,$$

and we take $A_j^0 = 1$ by convention.

Proof. We note that by Grönwall's inequality,

$$x_j \leqslant \beta\alpha_j j e^{-\beta j t}\int_0^t e^{\beta j s}\Big(x_{j+1}(s) + \frac{r_j}{\beta\alpha_j j}\Big)\, ds.$$

Note that

$$\alpha_j A_{j+1}^k = A_j^{k+1} \tag{3.7}$$

We first prove (3.5). We prove this bound inductively in ℓ, for all j. For $\ell = 0$, the bound is direct from $I_j^0(t) = A_j^0 = 1$. Then inductively, we use Grönwall's inequality together with the inductive hypothesis to give that x_j is bounded by

$$\begin{aligned} &\beta\alpha_j j e^{-\beta j t}\int_0^t e^{\beta j s}\Big(A_{j+1}^{\ell-1} I_{j+1}^{\ell-1}(s) \sup_{r\in[0,t]} x_{j+\ell}(r) + \frac{1}{\beta}\sum_{k=0}^{\ell-2} A_{j+1}^{k+1} I_{j+1}^{k+1}(s) \frac{r_{j+1+k}}{\alpha_{j+1+k}(j+1+k)} + \frac{r_j}{\beta\alpha_j j}\Big)\, ds \\ &= \beta j e^{-\beta j t}\int_0^t e^{\beta j s}\Big(A_j^\ell I_{j+1}^{\ell-1}(s) \sup_{r\in[0,t]} x_{j+\ell}(r) + \frac{1}{\beta}\sum_{k=0}^{\ell-1} A_j^{k+1} I_{j+1}^k(s) \frac{r_{j+k}}{\alpha_{j+k}(j+k)}\Big)\, ds, \end{aligned}$$

where we use (3.7) on the second line. The using Proposition 3.9, we get (3.5).

We now turn our attention to (3.6). Again we prove it inductively in ℓ, for all j. For $\ell = 0$, it is again direct from $I_j^0(t) = A_j^0 = 1$. Then inductively, we use Grönwall's inequality then (3.5) to control the integral on $[0, t_0]$ and the inductive hypothesis to control the integral on $[t_0, t]$. This gives

$$x_j \leqslant \beta\alpha_j j e^{-\beta j t}\int_0^{t_0} e^{\beta j s} A_{j+1}^{\ell-1} I_{j+1}^{\ell-1}(s) \sup_{r\in[0,t_0]} x_{j+\ell}(r)\, ds \tag{3.8}$$

$$+ \beta\alpha_j j e^{-\beta j t}\int_{t_0}^{t} e^{\beta j s}\Big(\sum_{k=1}^{\ell-1} A_{j+1}^{\ell-1} I_{j+\ell-k}^{k}(t_0) I_{j+1}^{\ell-1-k}(s-t_0) \sup_{r\in[0,t_0]} x_{j+\ell}(r)$$

$$+ A_{j+1}^{\ell-1} I_{j+1}^{\ell-1}(s-t_0) \sup_{r\in[t_0,t]} x_{j+\ell}(r)\Big)\, ds \tag{3.9}$$

$$+ \beta\alpha_j j e^{-\beta j t}\int_{0}^{t} e^{\beta j s}\Big(\frac{1}{\beta}\sum_{k=0}^{\ell-2} A_{j+1}^{k+1} I_{j+1}^{k+1}(s)\frac{r_{j+1+k}}{\alpha_{j+1+k}(j+1+k)} + \frac{r_j}{\beta\alpha_j j}\Big)\, ds. \tag{3.10}$$

We then note that top term (3.8) is equal to

$$e^{-\beta j(t-t_0)} A_j^{\ell} I_j^{\ell}(t_0) \sup_{s\in[0,t_0]} x_{j+\ell}(s) \leqslant A_j^{\ell} I_j^{\ell}(t_0) I_j^{0}(t-t_0) \sup_{s\in[0,t_0]} x_{j+\ell}(s),$$

where we use (3.7), Proposition 3.9, and the brutal bound $e^{-\beta j(t-t_0)} \leqslant 1$. Then the middle term (3.9) is equal to

$$\beta j e^{-\beta j(t-t_0)}\int_0^{t-t_0} e^{\beta j(s-t_0)}\Big(\sum_{k=1}^{\ell-1} A_j^{\ell} I_{j+\ell-k}^{k}(t_0) I_{j+1}^{\ell-1-k}(s) \sup_{r\in[0,t_0]} x_{j+\ell}(r) + A_j^{\ell} I_{j+1}^{\ell-1}(s) \sup_{r\in[t_0,t]} x_{j+\ell}(r)\Big)\, ds$$
$$= \sum_{k=1}^{\ell-1} A_j^{\ell} I_{j+\ell-k}^{k}(t_0) I_j^{\ell-k}(t-t_0) \sup_{s\in[0,t_0]} x_{j+\ell}(s) + A_j^{\ell} I_j^{\ell}(s) \sup_{s\in[t_0,t]} x_{j+\ell}(s),$$

we we again use (3.7) and Proposition 3.9. Lastly, we note that the third term (3.10) is equal to

$$\frac{1}{\beta}\sum_{k=0}^{\ell-1} A_j^{k+1} I_j^{k+1}(t)\frac{r_{j+k}}{\alpha_{j+k}(j+k)},$$

where the computation follows exactly as in the proof of (3.5). Combining these three equalities we get (3.6). ☐

Note. We remark that we take a very rough bound in the above argument, taking $e^{-\beta j(t-t_0)} \leqslant 1$. In other applications, one may wish to avoid taking this bound, but in this application, we will be interested in $t - t_0$ very small and $\sup_{s\in[0,t_0]} x_{j+\ell}$ already $O(1)$, as such we won't need the extra decay this exponential provides. Thus for simplicity, we discard it and get the above proposition.

We now can apply the exponential decay bound given by Proposition 3.11 to (3.6) to give the following.

Proposition 3.14. *Suppose $x_k \geqslant 0$ satisfy the hierarchy of differential inequalities*

$$\begin{cases} \dot{x}_k \leqslant \beta k(\alpha_k x_{k+1} - x_k) + r_k \\ x_k(0) = 0, \end{cases}$$

for some $r_k, \alpha_k \geqslant 0$ constants. Then for any $0 \leqslant t_0 \leqslant t$ and $j, \ell \in \mathbb{N}$ such that

$$j \leqslant e^{-2\beta t - 6}\ell, \tag{3.11}$$

we have the bound

$$x_j(t) \leqslant A_j^\ell \exp(-e^{-2\beta(t-t_0)-3}\ell) \sup_{s\in[t_0,t]} x_{j+\ell}(s) + A_j^\ell e^{2\beta t+7} \exp(-e^{-2\beta t-7}\ell) \sup_{s\in[0,t_0]} x_{j+\ell}(s)$$
$$+ \frac{1}{\beta}\sum_{k=0}^{\ell-1} A_j^{k+1} I_j^{k+1}(t) \frac{r_{j+k}}{\alpha_{j+k}(j+k)}.$$

Proof. By (3.6), it suffices to bound

$$I_j^\ell(t-t_0) \leqslant \exp(-e^{-2\beta(t-t_0)-3}\ell)$$

and

$$\sum_{k=1}^{\ell} I_{j+\ell-k}^k(t_0) I_j^{\ell-k}(t-t_0) \leqslant e^{2\beta t+7}\exp(-e^{-2\beta t-7}\ell).$$

The first bound is direct from Proposition 3.11 and the condition (3.11) on j. For the second, we let

$$\delta := \tfrac{1}{12}e^{-2\beta t_0-1}$$

and note that by (3.11),

$$j \leqslant \tfrac{1}{3}\delta e^{-2\beta(t-t_0)-1}\ell.$$

Then we have that

$$\sum_{k=1}^{\ell} I_{j+\ell-k}^k(t_0) I_j^{\ell-k}(t-t_0) = \sum_{k=1}^{\lfloor(1-\delta)\ell\rfloor} I_{j+\ell-k}^k(t_0) I_j^{\ell-k}(t-t_0) + \sum_{k=\lfloor(1-\delta)\ell\rfloor+1}^{\ell} I_{j+\ell-k}^k(t_0) I_j^{\ell-k}(t-t_0).$$

Then, since for $k \in \{1,\ldots,\lfloor(1-\delta)\ell\rfloor\}$, $\ell - k \geqslant \delta\ell$ and

$$j \leqslant \tfrac{1}{3}e^{-2\beta(t-t_0)-1}\delta\ell = \tfrac{1}{36}e^{-2\beta t-2}\ell,$$

we have from Proposition 3.11, using that $I_{j+\ell-k}^k(t_0) \leqslant 1$,

$$\begin{aligned}\sum_{k=1}^{\lfloor(1-\delta)\ell\rfloor} I_{j+\ell-k}^k(t_0) I_j^{\ell-k}(t-t_0) &\leqslant \sum_{k=1}^{\lfloor(1-\delta)\ell\rfloor} I_j^{\ell-k}(t-t_0)\\ &\leqslant \ell\exp(-\tfrac{1}{3}e^{-2\beta(t-t_0)-1}(\ell-k))\\ &\leqslant \ell\exp(-\tfrac{1}{3}\delta e^{-2\beta(t-t_0)-1}\ell) = \ell\exp(-\tfrac{1}{36}e^{-2\beta t-2}\ell). \end{aligned} \tag{3.12}$$

Then, for $k \in \{\lfloor(1-\delta)\ell\rfloor+1,\ldots,\ell\}$,

$$j+\ell-k \leqslant j+\delta\ell \leqslant \tfrac{1}{6}e^{-2\beta t_0-1}\ell \leqslant \tfrac{1}{3}e^{-2\beta t_0-1}k,$$

using the definition of δ and that (3.11) implies

$$j \leqslant \tfrac{1}{12}e^{-2\beta t_0-1}\ell.$$

Thus Proposition 3.11 gives that

$$\begin{aligned}\sum_{k=\lceil(1-\delta)\ell\rceil}^{\ell} I_{j+\ell-k}^k(t_0) I_j^{\ell-k}(t-t_0) &\leqslant \sum_{k=\lceil(1-\delta)\ell\rceil}^{\ell} I_{j+\ell-k}^k(t_0)\\ &\leqslant \ell\exp(-\tfrac{1}{3}e^{-2\beta t_0-1}(\lfloor(1-\delta)\ell\rfloor+1)) \leqslant \ell\exp(-\tfrac{1}{6}e^{-2\beta t_0-1}\ell). \end{aligned}\tag{3.13}$$

Combining (3.12) and (3.13), we get

$$\sum_{k=1}^{\ell} I_{j+\ell-k}^{k}(t_0) I_j^{\ell-k}(t-t_0) \leqslant 2\ell \exp(-\tfrac{1}{36}e^{-2\beta t-2}\ell)$$
$$\leqslant 128 e^{2\beta t+2}\exp(-\tfrac{1}{72}e^{-2\beta t-2}\ell) \leqslant e^{2\beta t+7}\exp(-e^{-2\beta t-7}\ell),$$

allowing us to conclude. □

3.2 Proof of Theorem 1.4

With the bounds given by Proposition 3.14 in hand, we are now ready to prove Theorem 1.4. The heart of the proof is captured in the following lemma, which we will iterate to get the full result.

Lemma 3.15. *Suppose $f \in L^\infty(\mathbb{T}^d)$, $K \in L^\infty(\mathbb{T}^{2d})$, and there exists $m > 0$ such that $f \geqslant m$. There exists $C(\|K\|_{L^\infty}, i) < \infty$ such that for any $t_0 \geqslant 0$, $L > 0$ with $LN^{2/3} \leqslant N$ and for any $M \geqslant 1$ with*

$$\sup_{s\in[0,t_0]} \int \Big|\frac{\varphi^i_{LN^{2/3}} - f_{LN^{2/3}}}{\rho^{\otimes LN^{2/3}}}\Big|^2 \rho^{\otimes LN^{2/3}}\, dx \leqslant M,$$

then for $\delta > 0$ defined to be

$$\delta e^{2\beta\delta} := \frac{1}{48e^3\|K\|^2_{L^\infty}} \wedge 1,$$

we have for all $j \in \mathbb{N}$ with

$$j \leqslant Le^{-2\beta(t_0+1)-7}N^{2/3},$$

and for all $t_0 \leqslant t \leqslant t+\delta$, the bound

$$\int \Big|\frac{\varphi^i_j - f_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \leqslant Ce^{Ct_0+L^3}\Big(\Big(\frac{j}{N}\Big)^{2(i+1)} + \frac{M}{N^{2(i+1)}}(L^2N)^{-8i-1}\Big).$$

Proof. We let

$$x_k := \int \Big|\frac{\varphi^i_k - f_k}{\rho^{\otimes k}}\Big|^2 \rho^{\otimes k}\, dx,$$

so by (3.3), we have that

$$\dot{x}_k \leqslant 4\|K\|^2_{L^\infty} k(x_{k+1} - x_k) + 4\|K\|^2_{L^\infty}\frac{k^3}{N^2}x_k + r_k \leqslant \beta k(\alpha_k x_{k+1} - x_k) + r_k,$$

where

$$\alpha_k := 1 + \frac{k^2}{N^2}, \quad A_k^\ell := \prod_{i=k}^{k+\ell-1}\alpha_i, \quad r_k := 2\int\Big|\frac{r^i_k}{\rho^{\otimes k}}\Big|^2\rho^{\otimes k}\, dx.$$

Then we note that

$$\log A_k^\ell = \sum_{i=k}^{k+\ell-1}\log\Big(1 + \frac{i^2}{N^2}\Big) \leqslant \frac{1}{N^2}\sum_{i=k}^{k+\ell-1} i^2 \leqslant \frac{1}{N^2}\int_{k+1}^{k+\ell} x^2\, dx \leqslant \frac{(k+\ell)^3}{N^2}.$$

Thus, for $k+\ell \leqslant LN^{2/3}$,

$$A_k^\ell \leqslant \exp\Big(\frac{L^3N^2}{N^2}\Big) = e^{L^3}. \tag{3.14}$$

By Proposition 3.6, for all $t \in [0, t_0+\delta]$, we can bound

$$r_k(t) \leqslant Ce^{Ct}\left(\frac{k}{N}\right)^{2(i+1)} \leqslant Ce^{Ct_0}\left(\frac{k}{N}\right)^{2(i+1)}. \tag{3.15}$$

Then, for any

$$j \leqslant Le^{-2\beta t_0 - 7}N^{2/3},$$

letting

$$\ell := LN^{2/3} - j \geqslant \tfrac{1}{2}LN^{2/3},$$

we have that

$$j \leqslant e^{-2\beta(t_0+1)-6}\ell,$$

so for any $t \in [0, \delta]$, we have by Proposition 3.14, (3.14), and (3.15) that $x_j(t)$ is bounded by

$$\begin{aligned} e^{L^3}\exp(-e^{-2\beta(t-t_0)-3}\ell)\sup_{s\in[t_0,t]} x_{LN^{2/3}}(s) + e^{L^3}e^{2\beta(t_0+1)+7}\exp(-e^{-2\beta(t_0+1)-7}\ell)\sup_{s\in[0,t_0]} x_{LN^{2/3}}(s) \\ + Ce^{Ct_0}\frac{e^{L^3}}{\beta N^{2(i+1)}}\sum_{k=0}^{\ell-1} I_j^{k+1}(t)(j+k)^{2i+1}. \end{aligned} \tag{3.16}$$

We note that by Proposition 3.10,

$$\sum_{k=0}^{\ell-1} I_j^{k+1}(t)(j+k)^{2i+1} \leqslant e^{\beta(2i+3)t}\sum_{k=0}^{\ell-1}(j+2i+3)^{2i+3}(j+k)^{-2} \leqslant Ce^{Ct_0}j^{2i+3}\int_{x=j}^{\infty} x^{-2} \leqslant Ce^{Ct_0}j^{2(i+1)}.$$

Thus

$$Ce^{Ct_0}\frac{e^{L^3}}{\beta N^{2(i+1)}}\sum_{k=0}^{\ell-1} I_j^{k+1}(t)(j+k)^{2i+1} \leqslant Ce^{Ct_0+L^3}\left(\frac{j}{N}\right)^{2(i+1)}. \tag{3.17}$$

Then we note that

$$\begin{aligned} \exp(-e^{-2\beta(t_0+1)-7}\ell)\sup_{s\in[0,t_0]} x_{LN^{2/3}}(s) &\leqslant M\exp(-e^{-2\beta(t_0+1)-8}LN^{2/3}) \\ &\leqslant \frac{M((3(i+1)+12i+3/2)e^{2\beta(t_0+1)+8}L^{-1})^{3(i+1)+12i+3/2}}{(N^{2/3})^{3(i+1)+12i+3/2}} \\ &\leqslant \frac{Ce^{Ct_0}M}{N^{2(i+1)}}(L^2N)^{-8i-1}, \end{aligned}$$

where we use that

$$x^m e^{-ax} \leqslant \left(\frac{m}{a}\right)^m e^{-m}.$$

Thus

$$e^{L^3}e^{2\beta(t_0+1)+7}\exp(-e^{-2\beta(t_0+1)-7}\ell)\sup_{s\in[0,t_0]} x_{LN^{2/3}}(s) \leqslant Ce^{Ct_0+L^3}\frac{M}{N^{2(i+1)}}(L^2N)^{-8i-1}. \tag{3.18}$$

For the last term in (3.16), we need to control $x_{LN^{2/3}}(t)$ for $t \in [t_0, t_0+\delta]$. Let

$$y_{LN^{2/3}}(t) := \int \left|\frac{f_{LN^{2/3}}}{\rho^{\otimes LN^{2/3}}}\right|^2 \rho^{\otimes LN^{2/3}}\,dx.$$

We note that by the triangle inequality

$$y_{LN^{2/3}}(t_0) \leqslant 2x_{LN^{2/3}}(t_0) + 2\int \Big|\frac{\varphi^i_{LN^{2/3}}}{\rho^{\otimes LN^{2/3}}}\Big|^2 (t_0)\rho^{\otimes LN^{2/3}}(t_0)\,dx \leqslant Ce^{Ct_0}M,$$

where we use that $M \geqslant 1$ and Proposition 3.5 to bound the term involving φ^i_j. Then we note that (3.4) gives that

$$\dot{y}_{LN^{2/3}} \leqslant 12\|K\|^2_{L^\infty} LN^{2/3} y_{LN^{2/3}},$$

thus, for $t_0 \leqslant t \leqslant t_0 + \delta$

$$y_{LN^{2/3}}(t) \leqslant e^{12\|K\|^2_{L^\infty} LN^{2/3}(t-t_0)} y_{LN^{2/3}}(t_0) \leqslant Ce^{Ct_0} M e^{12\|K\|^2_{L^\infty} LN^{2/3}\delta}.$$

So

$$x_{LN^{2/3}}(t) \leqslant 2y_{LN^{2/3}}(t) + 2\int \Big|\frac{\varphi^i_{LN^{2/3}}}{\rho^{\otimes LN^{2/3}}}\Big|^2 \rho^{\otimes LN^{2/3}}\,dx \leqslant Ce^{Ct_0} M e^{12\|K\|^2_{L^\infty} LN^{2/3}\delta}.$$

Note that $\ell \geqslant j$ and $j + \ell = LN^{2/3}$, so

$$\ell \geqslant \tfrac{1}{2} LN^{2/3}.$$

Thus for $t_0 \leqslant t \leqslant t_0 + \delta$,

$$\begin{aligned} \exp(-e^{-2\beta(t-t_0)-3}\ell) \sup_{s\in[t_0,t]} x_{LN^{2/3}}(s) &\leqslant Ce^{Ct_0} M \exp((12\|K\|^2_{L^\infty}\delta - \tfrac{1}{2}e^{-2\beta\delta-3})LN^{2/3}) \\ &\leqslant Ce^{Ct_0} M \exp(-\tfrac{1}{4}e^{-2\beta\delta-3}LN^{2/3}), \end{aligned}$$

where we use that by the definition of δ,

$$12\|K\|^2_{L^\infty}\delta \leqslant \tfrac{1}{4}e^{-2\beta\delta-3}.$$

Then

$$\exp(-\tfrac{1}{4}e^{-2\beta\delta-3}LN^{2/3}) \leqslant \exp(-e^{-2\beta-5}LN^{2/3}) \leqslant \frac{C}{N^{2(i+1)}}(L^2N)^{-8i-1}.$$

Thus,

$$e^{L^3} \exp(-e^{-2\beta(t-t_0)-3}\ell) \sup_{s\in[t_0,t]} x_{LN^{2/3}}(s) \leqslant Ce^{Ct_0+L^3}\frac{M}{N^{2(i+1)}}(L^2N)^{-8i-1}. \tag{3.19}$$

Then combining (3.16), (3.17), (3.18), and (3.19), we see that for any $j \leqslant Le^{-2\beta t_0-7}N^{2/3}$ and any $t_0 \leqslant t \leqslant t_0 + \delta$,

$$x_j \leqslant Ce^{Ct_0+L^3}\Big(\Big(\frac{j}{N}\Big)^{2(i+1)} + \frac{M}{N^{2(i+1)}}(L^2N)^{-8i}\Big),$$

as desired. □

We now prove Theorem 1.4 by iterating Lemma 3.15. The main difficulty is controlling the constants that appear in the iteration.

Proof of Theorem 1.4. Fix δ as in Lemma 3.15, so that

$$\delta e^{2\beta\delta} := \frac{1}{48e^3\|K\|^2_{L^\infty}} \wedge 1.$$

Let $L_0 = 1$ and let

$$L_{k+1} := \lfloor L_k e^{-2\beta\delta k-7-2\beta} N^{2/3} \rfloor N^{-2/3}.$$

We then claim inductively that for

$$j \leqslant L_k N^{2/3}; \quad 0 \vee (k-1)\delta \leqslant t \leqslant k\delta,$$

we have the bound

$$\int \Big|\frac{\varphi_j^i - f_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \leqslant B_k \Big(\frac{j}{N}\Big)^{2(i+1)},$$

where $B_0 = 1$ and

$$B_{k+1} = Ce^{C\delta k + L_k^3}(1 + B_k(L_k^2 N)^{-8i-1}).$$

We note the bound is trivially true for $k = 0$. Then inductively, using Lemma 3.15 with $t_0 = \delta k$, $L = L_k$, and

$$M = B_k,$$

we get that for

$$j \leqslant L_{k+1} N^{2/3} \leqslant L_k e^{-2\beta(k\delta+1)-7} N^{2/3},$$

for $\delta k \leqslant t \leqslant \delta(k+1)$,

$$\int \Big|\frac{\varphi_j^i - f_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \leqslant Ce^{C\delta k + L_k^3}(1 + B_k(L_k^2 N)^{-8i-1})\Big(\frac{j}{N}\Big)^{2(i+1)} \leqslant B_{k+1}\Big(\frac{j}{N}\Big)^{2(i+1)},$$

Thus the induction closes.

Now we just need to control B_k, L_k. First note that

$$L_k \leqslant L_{k-1} e^{-2\beta\delta(k-1)-7-2\beta} \leqslant L_{k-1} \leqslant \cdots \leqslant L_0 = 1,$$

and also that

$$\begin{aligned}
L_k &\geqslant L_{k-1} e^{-2\beta\delta(k-1)-7-2\beta} - N^{-2/3} \\
&\geqslant \cdots \geqslant \exp\Big(-(7+2\beta)k - 2\beta\delta \sum_{i=0}^{k-1} i\Big) L_0 - N^{-2/3} \sum_{\ell=0}^{k-1} \exp(-7\ell) \\
&\geqslant \exp\Big(-(7+2\beta)k - 2\beta\delta \sum_{i=0}^{k-1} i\Big) - 2N^{-2/3}.
\end{aligned}$$

Recalling $\sum_{i=0}^{k-1} i = \frac{1}{2}(k^2 - k)$, we have

$$L_k \geqslant \exp(-(7+2\beta)k - \beta\delta(k^2-k)) - 2N^{-2/3} \geqslant \exp(-10(1+\beta)k^2) - 2N^{-2/3}.$$

Thus, for $k \leqslant C^{-1}\sqrt{\log(N)} - C$, we have,

$$2N^{-2/3} \leqslant \tfrac{1}{2}\exp(-10(1+\beta)k^2)$$

so that

$$L_k \geqslant \frac{1}{2}\exp(-10(1+\beta)k^2) \geqslant \exp(-11(1+\beta)k^2).$$

Thus for $k \leqslant C^{-1}\sqrt{\log(N)} - C$,

$$\frac{1}{L_k^2 N} \leqslant \frac{1}{N}\exp(22(1+\beta)k^2) \leqslant N^{-1/2},$$

which implies that

$$B_{k+1} \leqslant Ce^{Ck}(1 + B_k(L_k^2 N)^{-8i-1}) \leqslant Ce^{Ck} + \frac{Ce^{Ck}}{N^{1/2}} B_k \leqslant C(e^{Ck} + B_k).$$

Iterating this bound then gives

$$B_k \leqslant Ce^{Ck}.$$

Therefore, for all $k \leqslant C^{-1}\sqrt{\log(N)} - C$, if

$$j \leqslant \exp(-11(1+\beta)k^2)L \leqslant L_k,$$

and $0 \leqslant t \leqslant k\delta$,

$$\int \Big|\frac{\varphi_j^i - f_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \leqslant Ce^{Ck}\Big(\frac{j}{N}\Big)^{2(i+1)}.$$

Choosing the optimal k, we get that for any $t \leqslant C^{-1}\sqrt{\log(N)} - C$, if $j \leqslant C^{-1}\exp(-Ct^2)N^{2/3}$, we get the bound

$$\int \Big|\frac{\varphi_j^i - f_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \leqslant Ce^{Ct}\Big(\frac{j}{N}\Big)^{2(i+1)}.$$

This is almost precisely the result, except with an additional restriction on t. We note though that the t bound is superfluous as by choosing C large enough, if $t \geqslant C_0\sqrt{\log(N)} - C_0$, then

$$C^{-1}\exp(-Ct^2)N^{2/3} < 1,$$

so the result holds vacuously in this case. Thus we can remove the t restriction and conclude. ☐

Note. We note by choosing the starting point of the induction L_0 to be larger, one can slightly expand the range of j for which one can prove the bound. This adds complication without being of any particular interest, so we omit this argument.

3.3 Proofs of bounds on the g_j^i, f_j^i, and R_j^i

To bound the R_j^i and f_j^i we must first bound the g_j^i. The following proof follows similarly to Proposition 2.10, where we inductively iterate up the hierarchy of equations satisfied by g_j^i to find estimates.

Proposition 3.16. *Suppose $f \in L^\infty(\mathbb{T}^d), K \in L^\infty(\mathbb{T}^{2d})$, and there exists $m > 0$ such that $f \geqslant m$. For all $i \geqslant 0$, letting*

$$\widetilde{g}_j^i := \frac{g_j^i}{\rho^{\otimes j}}, \tag{3.20}$$

there exists a constant $C(\|K\|_{L^\infty}, i)$ such that

$$\int |\widetilde{g}_j^i|^2 \rho^{\otimes j}\, dx \leqslant Ce^{Ct}. \tag{3.21}$$

Proof. We will inductively show this bound holds for $(i,j) \in T$ under the order given in Definition 1.1.

The bound trivially holds in the base case $(i,j) = (0,1)$ since $\widetilde{g}_1^0 = 1$, thus

$$\int |\widetilde{g}_1^0|^2 \rho^{\otimes j}\, dx = 1.$$

Assuming that for all $(k,\ell) < (i,j)$ the bound (3.21) holds, we group the terms on the right hand side of the equation for g_j^i to write

$$\partial_t g_j^i - \Delta g_j^i + \sum_{k=1}^{j} H_k g_{\{k\}}^0 g_{[j]\cup\{*\}-\{k\}}^i + \sum_{k=1}^{j} H_k g_{[j]}^i g_{\{*\}}^0 = \nabla \cdot F_j^i$$

where

$$F_j^i = \sum_{k=1}^{8} F_{j,k}^i$$

with

$$F_{j,1}^i = -\sum_{k=1}^{j} e_k \otimes \int K(x_k, x_*) g_{[j]\cup\{*\}}^i \, dx_*$$

$$F_{j,2}^i = -\sum_{k=1}^{j} \sum_{W\subseteq[j]-\{k\}} \sum_{m=1}^{i-1} e_k \otimes \int K(x_k, x_*) g_{W\cup\{k\}}^m g_{[j]\cup\{*\}-W-\{k\}}^{i-m} \, dx_*$$

and the other $F_{j,k}^i$ are defined similarly in the order of the equation (2.4). Taking a derivative we find

$$\begin{aligned}
\frac{d}{dt}\int |\widetilde{g}_j^i|^2 \rho^{\otimes j} &= \int 2\partial_t g_j^i \widetilde{g}_j^i - \partial_t \rho^{\otimes j} (\widetilde{g}_j^i)^2 \, dx \\
&= \int 2\Delta g_j^i \widetilde{g}_j^i - \Delta \rho^{\otimes j} (\widetilde{g}_j^i)^2 \, dx + 2\int \nabla \cdot F_j^i \widetilde{g}_j^i \, dx \\
&\quad - 2\int \sum_{k=1}^{j} H_k g_{\{k\}}^0 g_{[j]\cup\{*\}-\{k\}}^i \widetilde{g}_j^i \, dx - 2\int \sum_{k=1}^{j} H_k g_{[j]}^i g_{\{*\}}^0 \widetilde{g}_j^i \, dx \\
&\quad + \int \nabla \cdot \left(\sum_{k=1}^{j} e_k \otimes H_k \rho^{\otimes(j+1)} \right) (\widetilde{g}_j^i)^2 \, dx \\
&= -2\int |\nabla \widetilde{g}_j^i|^2 \rho^{\otimes j} \, dx - 2\int \frac{F_j^i}{\rho^{\otimes j}} \cdot \nabla \widetilde{g}_j^i \rho^{\otimes j} \, dx \\
&\quad + 2\int \left(\sum_{k=1}^{j} e_k \otimes \int K(x_k, x_*) \widetilde{g}_{[j]\cup\{*\}-\{k\}}^i \rho(x_*) \, dx_* \right) \cdot \nabla \widetilde{g}_j^i \rho^{\otimes j} \, dx \\
&\quad + 2\int \widetilde{g}_j^i \left(\sum_{k=1}^{j} e_k \otimes \int K(x_k, x_*) \rho(x_*) \, dx_* \right) \cdot \nabla \widetilde{g}_j^i \rho^{\otimes j} \, dx \\
&\quad - 2\int \widetilde{g}_j^i \left(\sum_{k=1}^{j} e_k \otimes \int K(x_k, x_*) \rho(x_*) \, dx_* \right) \cdot \nabla \widetilde{g}_j^i \rho^{\otimes j} \, dx \\
&\leqslant 2\int \left| \frac{F_j^i}{\rho^{\otimes j}} \right|^2 \rho^{\otimes j} \, dx + 2j \int \left| \int K(x_1, x_*) \rho(x_*) \widetilde{g}_{[j]\cup\{*\}-\{1\}}^i \, dx_* \right|^2 \rho^{\otimes j} \, dx,
\end{aligned}$$

where the last line follows via Young's inequality. Using Jensen's inequality

$$\begin{aligned}
\int \left| \int K(x_1, x_*) \rho(x_*) \widetilde{g}_{[j]\cup\{*\}-\{1\}}^i \, dx_* \right|^2 \rho^{\otimes j} \, dx &\leqslant \int |K(x_1, x_*)|^2 \left| \widetilde{g}_{[j]\cup\{*\}-\{k\}}^i \right|^2 \rho^{\otimes(j+1)} \, dx dx_* \\
&\leqslant \|K\|_{L^\infty} \int |\widetilde{g}_j^i|^2 \rho^{\otimes j} \, dx.
\end{aligned}$$

This has the form of a Grönwall term, thus all that remains is to bound the term involving F^i_j. First we use the triangle inequality to bound

$$\int \Big|\frac{F^i_j}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\,dx \leqslant C\sum_{m=1}^{8}\int \Big|\frac{F^i_{m,j}}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\,dx.$$

The most intimidating term is $F^i_{6,j}$ which equals

$$\sum_{k=1}^{j} e_k \sum_{W\subseteq[j]-\{k\}}\ \sum_{R\subseteq[j]-\{k\}-W}(j-1-|W|-|R|)\sum_{m=0}^{i-1}\sum_{n=0}^{i-1-m}\int K(x_k,x_*)g^m_{W\cup\{k\}}g^n_{R\cup\{*\}}g^{i-1-m-n}_{[j]-R-W-\{k\}}\,dx_*.$$

Using the triangle inequality over the sums and using exchangeability we find there exists a j dependent constant such that

$$\begin{aligned}&\int \Big|\frac{F^i_{k,j}}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\,dx\\ &\qquad\leqslant C\sum_{\substack{W\subseteq[j]-\{1\}\\ R\subseteq[j]-\{1\}-W}}\ \sum_{m=0}^{i-1}\sum_{n=0}^{i-1-m}\int\Big|\int K(x_1,x_*)\widetilde{g}^m_{W\cup\{1\}}\widetilde{g}^n_{R\cup\{*\}}\widetilde{g}^{i-1-m-n}_{[j]-R-W-\{1\}}\rho(x_*)\,dx_*\Big|^2\rho^{\otimes j}\,dx.\end{aligned}$$

We note that $(m,|W|+1),(n,|R|+1)$, and $(i-1-m-n, j-|R|-|W|-1)$ are all less than (i,j). We can thus bound

$$\begin{aligned}\int\Big|\int K(x_1,x*)&\widetilde{g}^m_{W\cup\{1\}}\widetilde{g}^n_{R\cup\{*\}}\widetilde{g}^{i-1-m-n}_{[j]-R-W-\{1\}}d\rho(x_*)\Big|^2\rho^{\otimes j}\,dx\\
&\leqslant \int\Big|\int K(x_1,x*)\widetilde{g}^n_{R\cup\{*\}}d\rho(x_*)\Big|^2|\widetilde{g}^m_{W\cup\{1\}}|^2|\widetilde{g}^{i-1-m-n}_{[j]-R-W-\{1\}}|^2\rho^{\otimes j}\,dx\\
&\leqslant \|K\|^2_{L^\infty}\int|\widetilde{g}^n_{R\cup\{*\}}|^2|\widetilde{g}^m_{W\cup\{1\}}|^2|\widetilde{g}^{i-1-m-n}_{[j]-R-W-\{1\}}|^2\rho^{\otimes(j+1)}\,dxdx_*\\
&= \|K\|^2_{L^\infty}\int|\widetilde{g}^n_{R\cup\{*\}}|^2\rho^{\otimes|R|+1}\,dx^{R\cup\{*\}}\int|\widetilde{g}^m_{W\cup\{1\}}|^2\rho^{\otimes|W|+1}\,dx^{W\cup\{1\}}\\
&\qquad\times\int|\widetilde{g}^{i-1-m-n}_{[j]-R-W-\{1\}}|^2\rho^{\otimes j-|R|-|W|-1}\,dx^{[j]-R-W-\{1\}}\\
&\leqslant Ce^{Ct}\Big(\sup_{(k,\ell)<(i,j)}\int|\widetilde{g}^k_\ell|^2\rho^{\otimes\ell}\,dx\Big)^3,\end{aligned}$$

where the second inequality follows by Jensen's inequality. Terms $F^i_{1,j}$ to $F^i_{5,j}$ are bounded similarly. The bounds on $F^i_{6,j}$ and $F^i_{8,j}$ are also straightforward and rely on bounding for $W\subset[j]-\{k,\ell\}$ integrals of the form

$$\begin{aligned}\int\Big|K(x_k,x_\ell)&\widetilde{g}^m_{W\cup\{k\}}\widetilde{g}^{i-1-m}_{[j]-\{1\}-W}\Big|^2\rho^{\otimes j}\,dx\\
&\leqslant\|K\|^2_{L^\infty}\int|\widetilde{g}^m_{W\cup\{k\}}|^2\rho^{\otimes|W|+1}\,dx^{W\cup\{k\}}\int|\widetilde{g}^{i-1-m}_{[j]-\{k\}-W}|^2\rho^{\otimes j-1-|W|}\,dx^{[j]-\{k\}-W}\\
&\leqslant Ce^{Ct}\Big(\sup_{(k,\ell)<(i,j)}\int|\widetilde{g}^k_\ell|^2\rho^{\otimes\ell}\,dx\Big)^2.\end{aligned}$$

All together these bounds imply that

$$\int \Big|\frac{F_j^i}{\rho^{\otimes j}}\Big|^2 \rho^{\otimes j}\, dx \leqslant Ce^{Ct}\Big(\sup_{(k,\ell)<(i,j)} \int |\widetilde{g}_\ell^k|^2 \rho^{\otimes \ell}\, dx\Big)^3.$$

Thus, in total we've found that

$$\frac{d}{dt}\frac{1}{2}\int |\widetilde{g}_j^i|^2\rho^{\otimes j}\, dx \leqslant 3j\int |\widetilde{g}_j^i|^2\rho^{\otimes j}\, dx + Ce^{Ct}\Big(\sup_{(k,\ell)<(i,j)} \int |\widetilde{g}_\ell^k|^2 \rho^{\otimes \ell}\, dx\Big)^3.$$

Applying Grönwall's inequality and inducting allows us to conclude, noting that $j \leqslant i+1$. □

With the bounds on g_j^i given by Proposition 3.16 in hand, we can now show the bounds on f_j^i, φ_j^i, and R_j^i given in Proposition 3.5 and Proposition 3.6. Before continuing, we prove a useful representation of the f_j^i.

Lemma 3.17.

$$f_j^i = \sum_{\substack{P\subseteq[j]\\ |P|\leqslant 2i}} \sum_{\pi\vdash P} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q = i\\ i_Q\geqslant 1}} \rho^{\otimes(j-|P|)}(x^{[j]-P}) \prod_{Q\in\pi} g_Q^{i_Q}.$$

Proof. By the definition (1.5),

$$f_j^i = \sum_{\sigma\vdash[j]} \sum_{\substack{(i_R)_{R\in\sigma}\\ \sum i_R = i}} \prod_{R\in\pi} g_R^{i_R}.$$

Since $g_\ell^k = 0$ if $\ell > k+1$, the product

$$\prod_{R\in\sigma} g_R^{i_R} = 0$$

unless $|R| \leqslant i_R + 1$ for all $R \in \sigma$. Suppose that σ corresponds to a nonzero product. Since $\sum_{R\in\pi} i_R = i$, we have that $i_R \neq 0$ for at most i sets $R \in \sigma$. Thus it must be the case that

$$\sum_{\substack{R\in\sigma\\ i_R\neq 0}} |R| \leqslant \sum_{\substack{R\in\pi\\ i_R\neq 0}} i_R + 1 \leqslant 2i.$$

Letting $P = \bigcup_{i_R\neq 0} R$, then $|P| \leqslant 2i$, $\sigma = \pi \cup \{\{k\} : k \in [j] - Q\}$ where $\pi \vdash P$, $\sum_{Q\in\pi} i_Q = i$, $i_Q \geqslant 1$ for $Q \in \pi$ and $i_{\{k\}} = 0$ for $k \notin P$.

Re-indexing the sum which defines f_j^i and using that $g_1^0 = \rho$ we thus get the above claimed representation of f_j^i. □

We now show the bounds on f_j^i. This will be a warm up for the more involved bounds on R_j^i.

Proof of Proposition 3.5. Using Lemma 3.17, and the definition of $\widetilde{g}_i^j$ given in Proposition 3.16

$$\frac{f_j^i}{\rho^{\otimes j}} = \sum_{\substack{P\subseteq[j]\\ |P|\leqslant 2i}} \sum_{\pi\vdash P} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q = i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}.$$

Thus expanding out the sums

$$\int \left|\frac{f_i^j}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\, dx = \sum_{\substack{P,R\subseteq[j]\\ |P|,|R|\leqslant 2i}} \sum_{\substack{\pi\vdash P\\ \sigma\vdash R}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \sum_{\substack{(i_W)_{W\in\pi}\\ \sum i_W=i\\ i_W\geqslant 1}} \int \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \times \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \rho^{\otimes j}\, dx.$$

Suppose that $P \neq R$, $\pi \vdash P$, $\sigma \vdash R$, and $i_Q, i_W \geqslant 1$ where $Q \in \pi$ and $W \in \sigma$. Then it must be the case that

$$\int \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \times \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \rho^{\otimes j}\, dx = 0.$$

Indeed, if $x_k \in Q \in \pi$, but x_k is not in R, then the marginalization given by Proposition 2.10 of $g_Q^{i_Q}$ implies this. We thus find that in fact

$$\int \left|\frac{f_i^j}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\, dx = \sum_{\substack{P\subseteq[j]\\ |P|\leqslant 2i}} \sum_{\pi,\sigma\vdash P} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \sum_{\substack{(i_W)_{W\in\pi}\\ \sum i_W=i\\ i_W\geqslant 1}} \int \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \times \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \rho^{\otimes j}\, dx.$$

Hölder's inequality with Proposition 3.16 imply that

$$\begin{aligned} \left|\int \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \rho^{\otimes j}\, dx\right|^2 &\leqslant \prod_{Q\in\pi} \int |\widetilde{g}_Q^{i_Q}|^2 \rho^{\otimes|Q|}\, dx^Q \times \prod_{W\in\pi} \int |\widetilde{g}_W^{i_W}|^2 \rho^{\otimes|W|}\, dx^W \\ &\leqslant Ce^{Ct}, \end{aligned}$$

where this constant only depends on i since there are at most i terms in the products. On the other hand

$$\sum_{\substack{P\subseteq[j]\\ |P|\leqslant 2i}} \sum_{\pi,\sigma\vdash P} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \sum_{\substack{(i_W)_{W\in\pi}\\ \sum i_W=i\\ i_W\geqslant 1}} 1 \leqslant \sum_{\pi,\sigma\vdash P} C \leqslant Cj^{2i}$$

where the constant C just depends on i. This completes the bound on the f_j^i.

The bound on φ_j^i is a direct consequence of this bound and the triangle inequality. □

Proof of Proposition 3.6. Throughout this proof, we somewhat abuse notation and denote

$$K * \rho(x) := \int K(x,y)\rho(y)\, dy.$$

First we note that

$$\frac{R_j^i}{\rho^{\otimes j}} = \frac{1}{N^{i+1}} \sum_{k=1}^{j} e_k \otimes \sum_{\ell=1}^{j} \int K(x_k, x_*) \frac{f^i_{[j]\cup\{*\}}}{\rho^{\otimes j}} dx_* - K(x_k, x_\ell) \frac{f_j^i}{\rho^{\otimes j}}.$$

Lemma 3.17 implies that

$$\frac{f_j^i}{\rho^{\otimes j}} = \sum_{m=1}^{2i} \sum_{\substack{P\subseteq[j]\\ |P|=m}} \sum_{\pi\vdash P} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}.$$

Thus

$$\frac{f^i_{[j]\cup\{*\}}}{\rho^{\otimes j}} = \frac{f^i_j}{\rho^{\otimes j}}\rho(x_*) + \rho(x_*) \sum_{m=1}^{2i} \sum_{\substack{P\subseteq[j]\\|P|=m-1}} \sum_{\pi\vdash P\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}.$$

Using exchangeability we find

$$\begin{aligned}
\int\left|\frac{R^i_j}{\rho^{\otimes j}}\right|^2 \rho^{\otimes j}\,dx &\leqslant \frac{j}{N^{2(i+1)}}\left|\sum_{\ell=1}^{j}\int K(x_1,x_*)\frac{f^i_{[j]\cup\{*\}}}{\rho^{\otimes j}}dx_* - K(x_1,x_\ell)\frac{f^i_j}{\rho^{\otimes j}}\right|^2\rho^{\otimes j}\,dx\\
&\leqslant \frac{2j}{N^{2(i+1)}}\int\left|\sum_{\ell=1}^{j}(K*\rho(x_1)-K(x_1,x_\ell))\frac{f^i_j}{\rho^{\otimes j}}\right|^2\rho^{\otimes j}\,dx\\
&\quad+\frac{2j^3}{N^{2(i+1)}}\int\left|\int K(x_1,x_*)\rho(x_*)\sum_{m=1}^{2i} \sum_{\substack{P\subseteq[j]\\|P|=m-1}} \sum_{\pi\vdash P\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}\,dx_*\right|^2\rho^{\otimes j}\,dx.
\end{aligned}$$

We first consider the second term. Applying Jensen's inequality, we have

$$\begin{aligned}
&\int\left|\int K(x_1,x_*)\rho(x_*)\sum_{m=1}^{2i} \sum_{\substack{P\subseteq[j]\\|P|=m-1}} \sum_{\pi\vdash P\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}\,dx_*\right|^2\rho^{\otimes j}\,dx\\
&\leqslant \int\left|K(x_1,x_*)\sum_{m=1}^{2i} \sum_{\substack{P\subseteq[j]\\|P|=m-1}} \sum_{\pi\vdash P\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}\right|^2\rho^{\otimes(j+1)}\,dxdx_*\\
&\leqslant 2i\sum_{m=1}^{2i}\int\left|K(x_1,x_*) \sum_{\substack{P\subseteq[j]\\|P|=m-1}} \sum_{\pi\vdash P\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}\right|^2\rho^{\otimes(j+1)}\,dxdx_*.
\end{aligned}$$

We now fix m and analyze the term under the integral, expanding the square

$$\begin{aligned}
&\int\left|K(x_1,x_*) \sum_{\substack{P\subseteq[j]\\|P|=m-1}} \sum_{\pi\vdash P\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}\right|^2\rho^{\otimes(j+1)}\,dxdx_*\\
&\leqslant \sum_{\substack{P\subseteq[j]\\|P|=m-1}}\sum_{\substack{R\subseteq[j]\\|R|=m-1}} \sum_{\pi\vdash P\cup\{*\}}\sum_{\sigma\vdash R\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}}\sum_{\substack{(i_W)_{W\in\sigma}\\ \sum i_W=i\\ i_W\geqslant 1}}\int|K(x_1,x_*)|^2 \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}\prod_{W\in\sigma} \widetilde{g}_W^{i_W}\,\rho^{\otimes(j+1)}\,dxdx_*.
\end{aligned}$$

Note then that unless $P=R$,

$$\int|K(x_1,x_*)|^2 \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q}\prod_{W\in\sigma} \widetilde{g}_W^{i_W}\,\rho^{\otimes(j+1)}\,dxdx_* = 0.$$

To see this, suppose $P \neq R$. Since $|P| = |R|$, then there exists $p \in P$ such that $p \notin R$ and there exists $r \in R$ such that $r \notin P$. We must have that $p \neq 1$ or $r \neq 1$. Let us suppose that $p \neq 1$, the other case follows symmetrically. Then let $S \in \pi$ such that $p \in S$. Then

$$\begin{aligned}
&\int |K(x_1, x_*)|^2 \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \rho^{\otimes j} \rho(x_*)\, dx dx_* \\
&\quad = \int |K(x_1, x_*)|^2 \prod_{Q\in\pi-\{S\}} \widetilde{g}_Q^{i_Q} \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \int \widetilde{g}_S^{i_S} \rho^{\otimes(j+1)}\, dx_p dx_1 \cdots dx_{p-1} dx_{p+1} \cdots dx_j dx_* = 0,
\end{aligned}$$

where we use that

$$\int \widetilde{g}_S^{i_S} \rho^{\otimes|S|}(x^S)\, dx_p = \int g_S^{i_S}\, dx_p = 0,$$

by Proposition 2.10, since $i_S \geqslant 1$.

Using Hölder's inequality

$$\begin{aligned}
&\int |K(x_1, x_*)|^2 \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \rho^{\otimes(j+1)}\, dx dx_* \\
&\quad \leqslant \|K\|_{L^\infty}^2 \left(\int \prod_{Q\in\pi} |\widetilde{g}_Q^{i_Q}|^2 \rho^{\otimes(j+1)}\, dx dx_* \right)^{\frac{1}{2}} \left(\int \prod_{W\in\sigma} |\widetilde{g}_W^{i_W}|^2 \rho^{\otimes(j+1)}\, dx dx_* \right)^{\frac{1}{2}} \\
&\quad = \|K\|_{L^\infty}^2 \prod_{Q\in\pi} \left(\int |\widetilde{g}_Q^{i_Q}|^2 \rho^{\otimes|Q|}\, dx^Q \right)^{\frac{1}{2}} \prod_{W\in\sigma} \left(\int |\widetilde{g}_W^{i_W}|^2 \rho^{\otimes|W|}\, dx^W \right)^{\frac{1}{2}}
\end{aligned}$$

Since $i_Q \leqslant i$ for all $Q \in \pi$, Proposition 3.21 implies that

$$\prod_{Q\in\pi} \left(\int |\widetilde{g}_Q^{i_Q}|^2 \rho^{\otimes|Q|}\, dx^Q \right)^{\frac{1}{2}} \prod_{W\in\sigma} \left(\int |\widetilde{g}_W^{i_W}|^2 \rho^{\otimes|W|}\, dx^W \right)^{\frac{1}{2}} \leqslant (Ce^{Ct})^{4i} \leqslant Ce^{Ct}. \tag{3.22}$$

Thus we always have that

$$\int |K(x_1, x_*)|^2 \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \prod_{W\in\sigma} \widetilde{g}_W^{i_W} \rho^{\otimes j} \rho(x_*)\, dx dx_* \leqslant Ce^{Ct}.$$

We also have that for any $P, R \subseteq [j]$ such that $|P| = |R| = m-1 \leqslant 2i$,

$$\sum_{\pi \vdash P\cup\{*\}} \sum_{\sigma \vdash R\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi} \\ \sum i_Q = i \\ i_Q \geqslant 1}} \sum_{\substack{(i_W)_{W\in\sigma} \\ \sum i_W = i \\ i_W \geqslant 1}} 1 \leqslant C.$$

Thus

$$\begin{aligned}
&\int \left| K(x_1, x_*) \sum_{\substack{P\subseteq[j] \\ |P|=m-1}} \sum_{\pi \vdash P\cup\{*\}} \sum_{\substack{(i_Q)_{Q\in\pi} \\ \sum i_Q = i \\ i_Q \geqslant 1}} \prod_{Q\in\pi} \widetilde{g}_Q^{i_Q} \right|^2 \rho^{\otimes j} \rho(x_*)\, dx dx_* \\
&\quad \leqslant \sum_{\substack{P\subseteq[j] \\ |P|=m-1}} \sum_{\substack{R\subseteq[j] \\ |R|=m-1}} Ce^{Ct} \delta_{P=R} = Ce^{Ct} \binom{j}{m-1} \leqslant Ce^{Ct} j^{2i-1}.
\end{aligned}$$

Putting it together, we so far have that

$$\int\left|\frac{R_j^i}{\rho^{\otimes j}}\right|^2\rho^{\otimes j}\,dx \leqslant Ce^{Ct}\Big(\frac{j}{N}\Big)^{2(i+1)} + \frac{2j}{N^{2(i+1)}}\int\left|\sum_{\ell=1}^{j}(K*\rho(x_1)-K(x_1,x_\ell))\frac{f_j^i}{\rho^{\otimes j}}\right|^2\rho^{\otimes j}\,dx. \tag{3.23}$$

All that remains therefore is to bound the second term above, which is somewhat more involved. Expanding $\frac{f_j^i}{\rho^{\otimes j}}$, pulling out one of the sums then expanding the square, we get

$$\begin{aligned}
&\int\left|\sum_{\ell=1}^{j}(K*\rho(x_1)-K(x_1,x_\ell))\frac{f_j^i}{\rho^{\otimes j}}\right|^2\rho^{\otimes j}\,dx \\
&\quad\leqslant 2i\sum_{m=1}^{2i}\int\left|\sum_{\ell=1}^{j}(K*\rho(x_1)-K(x_1,x_\ell))\sum_{\substack{P\subseteq[j]\\|P|=m}}\sum_{\pi\vdash P}\sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}}\prod_{Q\in\pi}\widetilde{g}_Q^{i_Q}\right|^2\rho^{\otimes j}\,dx \\
&\quad\leqslant 2i\sum_{m=1}^{2i}\sum_{\ell=1}^{j}\sum_{k=1}^{j}\sum_{\substack{P\subseteq[j]\\|P|=m}}\sum_{\substack{R\subseteq[j]\\|R|=m}}\sum_{\pi\vdash P}\sum_{\sigma\vdash R}\sum_{\substack{(i_Q)_{Q\in\pi}\\ \sum i_Q=i\\ i_Q\geqslant 1}}\sum_{\substack{(i_W)_{W\in\sigma}\\ \sum i_W=i\\ i_W\geqslant 1}} 1 \\
&\qquad\times\int(K*\rho(x_1)-K(x_1,x_\ell))\cdot(K*\rho(x_1)-K(x_1,x_k))\prod_{Q\in\pi}\widetilde{g}_Q^{i_Q}\prod_{W\in\sigma}\widetilde{g}_W^{i_W}\,\rho^{\otimes j}\,dx.
\end{aligned}\tag{3.24}$$

We then claim that

$$\begin{aligned}
&\int(K*\rho(x_1)-K(x_1,x_\ell))\cdot(K*\rho(x_1)-K(x_1,x_k))\prod_{Q\in\pi}\widetilde{g}_Q^{i_Q}\prod_{W\in\sigma}\widetilde{g}_W^{i_W}\,\rho^{\otimes j}\,dx \\
&\qquad\qquad\leqslant Ce^{Ct}\delta_{\substack{\ell\in P\cup R\cup\{1,k\}\\ k\in P\cup R\cup\{1,\ell\}}}\delta_{\substack{P\subseteq R\cup\{1,\ell,k\}\\ R\subseteq P\cup\{1,\ell,k\}}}.
\end{aligned}\tag{3.25}$$

The bound by Ce^{Ct} follows by (3.22) as

$$\begin{aligned}
&\left|\int(K*\rho(x_1)-K(x_1,x_\ell))\cdot(K*\rho(x_1)-K(x_1,x_k))\prod_{Q\in\pi}\widetilde{g}_Q^{i_Q}\prod_{W\in\sigma}\widetilde{g}_W^{i_W}\,\rho^{\otimes j}\,dx\right| \\
&\qquad\leqslant 4\|K\|_{L^\infty}^2\prod_{Q\in\pi}\Big(\int|\widetilde{g}_Q^{i_Q}|^2\rho^{\otimes|Q|}\,dx^Q\Big)^{\frac12}\prod_{W\in\sigma}\Big(\int|\widetilde{g}_W^{i_W}|^2\rho^{\otimes|W|}\,dx^W\Big)^{\frac12}\leqslant Ce^{Ct}.
\end{aligned}$$

Thus we just need to show that if if any of the above four conditions fails to hold, the integral is 0. The integral and conditions are symmetric in ℓ,k and also symmetric in P,R, so we just need to check the two conditions. If $\ell\notin P\cup R\cup\{1,k\}$, then

$$\begin{aligned}
&\int(K*\rho(x_1)-K(x_1,x_\ell))\cdot(K*\rho(x_1)-K(x_1,x_k))\prod_{Q\in\pi}\widetilde{g}_Q^{i_Q}\prod_{W\in\sigma}\widetilde{g}_W^{i_W}\,\rho^{\otimes j}\,dx \\
&\quad=\int(K*\rho(x_1)-K(x_1,x_k))\prod_{Q\in\pi}\widetilde{g}_Q^{i_Q}\prod_{W\in\sigma}\widetilde{g}_W^{i_W} \\
&\qquad\times\int(K*\rho(x_1)-K(x_1,x_\ell))\rho^{\otimes j}\,dx_\ell dx_1\cdots dx_{\ell-1}dx_{\ell+1}\cdots dx_j=0,
\end{aligned}$$

where we use that

$$\int (K * \rho(x_1) - K(x_1, x_\ell))\rho(x_\ell)\, dx_\ell = K * \rho(x_1) - K * \rho(x_1) = 0,$$

as $\ell \neq 1$.

Thus we see we get the term $\delta_{\ell \in P \cup R \cup \{1,k\}}$ in the bound and applying the argument with k and ℓ switched, we get the the term $\delta_{k \in P \cup R \cup \{1,\ell\}}$. Now suppose that $P \nsubseteq R \cup \{1, \ell, k\}$, i.e. there exists $p \in P$ s.t. $p \notin R \cup \{1, \ell, k\}$. Then let $S \in \pi$ such that $p \in S$. Then we have that

$$\begin{aligned}
&\int (K * \rho(x_1) - K(x_1, x_\ell)) \cdot (K * \rho(x_1) - K(x_1, x_k)) \prod_{Q \in \pi} \tilde{g}_Q^{i_Q} \prod_{W \in \sigma} \tilde{g}_W^{i_W} \rho^{\otimes j}\, dx \\
&= \int (K * \rho(x_1) - K(x_1, x_\ell)) \cdot (K * \rho(x_1) - K(x_1, x_k)) \\
&\quad \times \prod_{Q \in \pi - \{S\}} \tilde{g}_Q^{i_Q} \prod_{W \in \sigma} \tilde{g}_W^{i_W} \int \tilde{g}_S^{i_S} \rho^{\otimes j}\, dx_p dx_1 \cdots dx_{p-1} dx_{p+1} \cdots dx_j = 0,
\end{aligned}$$

where we use that

$$\int \tilde{g}_S^{i_S} \rho^{\otimes |S|}(x^S)\, dx_p = \int g_S^{i_S}\, dx_p = 0,$$

by Proposition 2.10, using $i_S \geqslant 1$. Thus we get the term $\delta_{P \subseteq R \cup \{1,\ell,k\}}$ and symmetrically get the term $\delta_{R \subseteq P \cup \{1,\ell,k\}}$, thus showing the claim (3.25).

Thus we are left with bounding

$$\begin{aligned}
&2i \sum_{m=1}^{2i} \sum_{\ell=1}^{j} \sum_{k=1}^{j} \sum_{\substack{P \subseteq [j] \\ |P|=m}} \sum_{\substack{R \subseteq [j] \\ |R|=m}} \sum_{\pi \vdash P} \sum_{\sigma \vdash R} \sum_{\substack{(i_Q)_{Q \in \pi} \\ \sum i_Q = i \\ i_Q \geqslant 1}} \sum_{\substack{(i_W)_{W \in \sigma} \\ \sum i_W = i \\ i_W \geqslant 1}} \delta_{\substack{\ell \in P \cup R \cup \{1,k\} \\ k \in P \cup R \cup \{1,\ell\}}} \delta_{\substack{P \subseteq R \cup \{1,\ell,k\} \\ R \subseteq P \cup \{1,\ell,k\}}} \\
&= 2i \sum_{m=1}^{2i} \sum_{\ell=1}^{j} \sum_{k=1}^{j} \sum_{\substack{P \subseteq [j] \\ |P|=m}} \sum_{\substack{R \subseteq [j] \\ |R|=m}} \delta_{\substack{\ell \in P \cup R \cup \{1,k\} \\ k \in P \cup R \cup \{1,\ell\}}} \delta_{\substack{P \subseteq R \cup \{1,\ell,k\} \\ R \subseteq P \cup \{1,\ell,k\}}} \sum_{\pi \vdash P} \sum_{\sigma \vdash R} \sum_{\substack{(i_Q)_{Q \in \pi} \\ \sum i_Q = i \\ i_Q \geqslant 1}} \sum_{\substack{(i_W)_{W \in \sigma} \\ \sum i_W = i \\ i_W \geqslant 1}} 1 \\
&\leqslant C \sum_{m=1}^{2i} \sum_{\ell=1}^{j} \sum_{k=1}^{j} \sum_{\substack{P \subseteq [j] \\ |P|=m}} \sum_{\substack{R \subseteq [j] \\ |R|=m}} \delta_{\substack{\ell \in P \cup R \cup \{1,k\} \\ k \in P \cup R \cup \{1,\ell\}}} \delta_{\substack{P \subseteq R \cup \{1,\ell,k\} \\ R \subseteq P \cup \{1,\ell,k\}}},
\end{aligned} \tag{3.26}$$

where we use that

$$\sum_{\pi \vdash P} \sum_{\sigma \vdash R} \sum_{\substack{(i_Q)_{Q \in \pi} \\ \sum i_Q = i \\ i_Q \geqslant 1}} \sum_{\substack{(i_W)_{W \in \sigma} \\ \sum i_W = i \\ i_W \geqslant 1}} 1 = \sum_{\pi \vdash [m]} \sum_{\sigma \vdash [m]} \sum_{\substack{(i_Q)_{Q \in \pi} \\ \sum i_Q = i \\ i_Q \geqslant 1}} \sum_{\substack{(i_W)_{W \in \sigma} \\ \sum i_W = i \\ i_W \geqslant 1}} 1 = C(m, i) \leqslant C.$$

We now claim that

$$\delta_{\substack{P \subseteq R \cup \{1,\ell,k\} \\ R \subseteq P \cup \{1,\ell,k\}}} = 0$$

unless $P = R$ or $P = R - \{a\} \cup \{b\}$ with $a \in R; a \neq b; a, b \in \{1, \ell, k\}$.

To see this, suppose $P \subseteq R \cup \{1, \ell, k\}$ and $R \subseteq P \cup \{1, \ell, k\}$. Then note that the symmetric difference $P \,\Delta\, R = (P - R) \cup (R - P) \subseteq \{1, \ell, k\}$. Then, since $|P| = |R|$, we have that

$$|P - R| = |P| - |R \cap P| = |R| - |R \cap P| = |R - P|.$$

Thus

$$|P \,\Delta\, R| = |P - R| + |R - P| = 2|P - R|.$$

Thus $P \,\Delta\, R$ is an even sized subset of $\{1, \ell, k\}$, hence either $P = R$ or $P = R - \{a\} \cup \{b\}$ with $a \in R; a \neq b; a, b \in \{1, \ell, k\}$, as claimed. Let us first deal with the case that $P = R$, in which case the sum becomes

$$\sum_{\substack{P\subseteq[j]\\|P|=m}} \sum_{\ell=1}^{j} \sum_{k=1}^{j} \delta_{\ell\in P\cup\{1,k\}} \delta_{k\in P\cup\{1,\ell\}} \leqslant \sum_{\substack{P\subseteq[j]\\|P|=m}} \sum_{\ell=1}^{j} m + 2 \leqslant Cj^{2i+1},$$

using that $m \leqslant 2i$.

For $P \neq R$ the remaining part of the sum to bound is

$$\begin{aligned}
&\sum_{\ell=1}^{j} \sum_{k=1}^{j} \sum_{\substack{R\subseteq[j]\\|R|=m}} \sum_{a\in\{1,k,\ell\}\cap R} \sum_{b\in\{1,k,\ell\}-\{a\}} \delta_{\substack{\ell\in(R-\{a\}\cup\{b\})\cup R\cup\{1,k\}\\ k\in(R-\{a\}\cup\{b\})\cup R\cup\{1,\ell\}}}\\
&= \sum_{\substack{R\subseteq[j]\\|R|=m}} \sum_{\ell=1}^{j} \sum_{k=1}^{j} \sum_{a\in\{1,k,\ell\}\cap R} \sum_{b\in\{1,k,\ell\}-\{a\}} \delta_{\substack{\ell\in R\cup\{1,k,b\}\\ k\in R\cup\{1,\ell,b\}}}\\
&\leqslant C_i j^{2i+1} + \sum_{\substack{R\subseteq[j]\\|R|=m}} \sum_{\ell=2}^{j} \sum_{k=2,k\neq\ell}^{j} \sum_{a\in\{1,k,\ell\}\cap R} \sum_{b\in\{1,k,\ell\}-\{a\}} \delta_{\substack{\ell\in R\cup\{b\}\\ k\in R\cup\{b\}}},
\end{aligned}$$

where on the last line we split off the three cases $\ell = 1, k = 1$, and $\ell = k$ and apply the straightforward bounds to them separately. We lastly split the remaining term along the cases the $a = 1, a = k$, and $a = \ell$. The first case $a = 1$ gives

$$\sum_{\substack{R\subseteq[j]\\|R|=m}} \sum_{\ell=2}^{j} \sum_{k=2,k\neq\ell}^{j} \sum_{b\in\{k,\ell\}} \delta_{\substack{\ell\in R\cup\{b\}\\ k\in R\cup\{b\}}} \delta_{1\in R} \leqslant \sum_{\substack{W\subseteq[j]-\{1\}\\|W|=m-1}} \sum_{\ell=2}^{j} \sum_{k=2,k\neq\ell}^{j} 2 \leqslant Cj^{2i+1}.$$

The second case $a = k$ gives

$$\sum_{\substack{R\subseteq[j]\\|R|=m}} \sum_{\ell=2}^{j} \sum_{k\in R,k\neq\ell} \sum_{b\in\{1,\ell\}} \delta_{\substack{\ell\in R\cup\{b\}\\ k\in R\cup\{b\}}} \leqslant 2m \sum_{\substack{R\subseteq[j]\\|R|=m}} \sum_{\ell=2}^{j} \leqslant Cj^{2i+1}.$$

The third case $a = \ell$ follows symmetrically. Thus

$$\sum_{m=1}^{2i} \sum_{\ell=1}^{j} \sum_{k=1}^{j} \sum_{\substack{P\subseteq[j]\\|P|=m}} \sum_{\substack{R\subseteq[j]\\|R|=m}} \delta_{\substack{\ell\in P\cup R\cup\{1,k\}\\ k\in P\cup R\cup\{1,\ell\}}} \delta_{\substack{P\subseteq R\cup\{1,\ell,k\}\\ R\subseteq P\cup\{1,\ell,k\}}} \leqslant \sum_{m=1}^{2i} Cj^{2i+1} \leqslant Cj^{2i+1}. \tag{3.27}$$

Combining (3.23), (3.24), (3.25), (3.26), and (3.27), we conclude. $\square$

4. Proofs of cluster expansions and perturbation theory

We give some additional notation for partitions.

Definition 4.1. We define the following partial order on partitions. If $\sigma, \pi \vdash A$, we say that $\sigma \leqslant \pi$ if for every $P \in \pi$, there exists $Q \in \sigma$ such that $P \subseteq Q$. If $\sigma \leqslant \pi$ we say σ is a *combining of* π. We note that if $\sigma \leqslant \pi \vdash A$, then $|\sigma| \leqslant |\pi| \leqslant |A|$.

We note the following combinatoric lemma which we will appeal to frequently in the below proofs.

Lemma 4.2. *Let S be a finite set, $\pi \vdash S$. Then*

$$\sum_{\sigma \leqslant \pi} (-1)^{|\sigma|-1}(|\sigma|-1)! = \begin{cases} 1 & |\pi| = 1, \\ 0 & |\pi| \geqslant 2. \end{cases}$$

Proof. In order to evaluate these sums, we take advantage of the natural isomorphism from partitions of π to combinings of π. For $\Pi \vdash \pi$, we let

$$\sigma(\Pi) = \left\{ \bigcup_{P \in \alpha} P : \alpha \in \Pi \right\}.$$

Note that σ defines a bijection between partitions of π and combinings of π, and further that $|\sigma(\Pi)| = |\Pi|$. This immediately implies that

$$\sum_{\sigma \leqslant \pi} (-1)^{|\sigma|-1}(|\sigma|-1)! = \sum_{\Pi \vdash \pi} (-1)^{|\Pi|-1}(|\Pi|-1)!$$

The lemma then follows after applying the following fact

$$\sum_{\alpha \vdash [j]} (-1)^{|\alpha|-1}(|\alpha|-1)! = \begin{cases} 1 & j = 1, \\ 0 & j \geqslant 2, \end{cases}$$

which follows by the Faà di Bruno's formula applied to $\log e^x$. ∎

Proof of Proposition 2.4. We prove the equality inductively in j. The case $j = 1$ is clear. For $j \geqslant 2$, we have

$$\begin{aligned} g_j &= \sum_{\pi \vdash [j]} (-1)^{|\pi|-1}(|\pi|-1)! \prod_{P \in \pi} f_P \\ &= f_j + \sum_{\pi \vdash [j], |\pi| \geqslant 2} (-1)^{|\pi|-1}(|\pi|-1)! \prod_{P \in \pi} f_P \\ &= f_j + \sum_{\pi \vdash [j], |\pi| \geqslant 2} (-1)^{|\pi|-1}(|\pi|-1)! \prod_{P \in \pi} \sum_{\sigma \vdash P} \prod_{Q \in \sigma} g_Q \\ &= f_j + \sum_{\pi \vdash [j], |\pi| \geqslant 2} (-1)^{|\pi|-1}(|\pi|-1)! \sum_{\sigma \geqslant \pi} \prod_{Q \in \sigma} g_Q \\ &= f_j + \sum_{\sigma \vdash [j]} \sum_{\pi \leqslant \sigma, |\pi| \geqslant 2} (-1)^{|\pi|-1}(|\pi|-1)! \prod_{P \in \sigma} g_P. \end{aligned}$$

We have now collected all the terms $\prod_{P\in\sigma} g_P$ so all that remains is to compute the combinatoric constants. To that end, using Lemma 4.2, we have

$$\sum_{\pi\leqslant\sigma,|\pi|\geqslant 2}(-1)^{|\pi|-1}(|\pi|-1)! = -1+\sum_{\pi\leqslant\sigma}(-1)^{|\pi|-1}(|\pi|-1)! = \begin{cases}0 & |\sigma|=1,\\ -1 & |\sigma|\geqslant 2.\end{cases}$$

Plugging this in above allows us to conclude. $\square$

Proof of Proposition 2.6. We start by just directly computing $(\partial_t-\Delta)g_j$ using the definition of g_j in terms of the f_j,

$$\begin{aligned}\partial_t g_j-\Delta g_j &= \sum_{\pi\vdash[j]}(-1)^{|\pi|-1}(|\pi|-1)!\sum_{P\in\pi}(\partial_t f_P-\Delta f_P)\prod_{\substack{Q\in\pi\\ Q\neq P}} f_Q\\ &= -\sum_{\pi\vdash[j]}(-1)^{|\pi|-1}(|\pi|-1)!\sum_{P\in\pi}\left(\frac{N-|P|}{N}\sum_{k\in P}H_k f_{P\cup\{*\}}+\frac{1}{N}\sum_{k,\ell\in P}S_{k,\ell}f_P\right)\prod_{\substack{Q\in\pi\\ Q\neq P}} f_Q.\end{aligned} \tag{4.1}$$

Note that, consistent with the definition of H_k, the variable $*$ is always the coordinate being integrated over. We consider the H_k terms and the $S_{k,\ell}$ terms separately. We first consider the $S_{k,\ell}$ terms. We use Proposition 2.4 to expand each of the f_R in terms of g_Q

$$\begin{aligned}&\sum_{\pi\vdash[j]}(-1)^{|\pi|-1}(|\pi|-1)!\sum_{P\in\pi}\sum_{k,\ell\in P}S_{k,\ell}f_P\prod_{\substack{Q\in\pi\\ Q\neq P}} f_Q\\ &\quad=\sum_{\pi\vdash[j]}(-1)^{|\pi|-1}(|\pi|-1)!\sum_{P\in\pi}\sum_{k,\ell\in P}\sum_{\pi\leqslant\sigma}S_{k,\ell}\prod_{Q\in\sigma}g_Q\\ &\quad=\sum_{k,\ell=1}^{j}\sum_{\sigma\vdash[j]}\Bigg(\sum_{\substack{\pi\leqslant\sigma\\ \exists P\in\pi,\{k,\ell\}\subseteq P}}(-1)^{|\pi|-1}(|\pi|-1)!\Bigg)S_{k,\ell}\prod_{Q\in\sigma}g_Q\\ &\quad=:\sum_{k,\ell=1}^{j}\sum_{\sigma\vdash[j]}a^{\sigma}_{k,\ell}S_{k,\ell}\prod_{P\in\sigma}g_P.\end{aligned}$$

We now compute $a^{\sigma}_{k,\ell}$. Fix σ,k,ℓ. We split into two cases, the first being that there exists $Q\in\sigma$ such that $\{k,\ell\}\subseteq Q$ and second being that there exists $Q,R\in\sigma, Q\neq R$ such that $k\in Q,\ell\in R$. Note that in the second case, it must be that $k\neq\ell$.

In the first case, since for any $\pi\leqslant\sigma$, by definition of the order, there exists $P\in\pi$ such that $Q\subseteq P$, as such $\{k,\ell\}\subseteq P$. Thus

$$\begin{aligned}a^{\sigma}_{k,\ell} &= \sum_{\substack{\pi\leqslant\sigma\\ \exists P\in\pi,\{k,\ell\}\subseteq P}}(-1)^{|\pi|-1}(|\pi|-1)!\\ &= \sum_{\pi\leqslant\sigma}(-1)^{|\pi|-1}(|\pi|-1)!\\ &= \begin{cases}1 & |\sigma|=1\\ 0 & |\sigma|\geqslant 2,\end{cases}\end{aligned}$$

where we use Lemma 4.2 to conclude.

For the second case, we first write $\sigma = \{Q, R, W_1, ..., W_m\}$ such that $k \in Q, \ell \in R$. Then define

$$\widetilde{\sigma} := \{Q \cup R, W_1, ..., W_m\}.$$

Then we note that $\pi \leqslant \sigma$ for which there exists $P \in \pi$ such that $\{k, \ell\} \subseteq P$ if and only if $\pi \leqslant \widetilde{\sigma}$. Thus

$$a_{k,\ell}^{\sigma} = \sum_{\pi \leqslant \widetilde{\sigma}} (-1)^{|\pi|-1}(|\pi| - 1)! = \begin{cases} 1 & |\sigma| = 2 \\ 0 & |\sigma| \geqslant 3, \end{cases}$$

once again using Lemma 4.2 and noting $|\sigma| = |\widetilde{\sigma}| + 1$.

Combining these two cases, we note the complete formula for $a_{k,\ell}^{\sigma}$ is given by

$$a_{k,\ell}^{\sigma} = \begin{cases} 1 & |\sigma| = 1 \\ 1 & \sigma = \{Q, R\}, k \in Q, \ell \in R \\ 0 & \sigma = \{Q, R\}, k, \ell \in Q \\ 0 & |\sigma| \geqslant 3. \end{cases} \tag{4.2}$$

We have thus dealt with the $S_{k,\ell}$ terms completely. We now proceed to the H_k terms. Similarly, we compute

$$\begin{aligned}
&\sum_{\pi \vdash [j]} (-1)^{|\pi|-1}(|\pi| - 1)! \sum_{P \in \pi} \frac{N - |P|}{N} \sum_{k \in P} H_k f_{P \cup \{*\}} \prod_{\substack{Q \in \pi \\ Q \neq P}} f_Q \\
&\quad = \sum_{\pi \vdash [j]} (-1)^{|\pi|-1}(|\pi| - 1)! \sum_{P \in \pi} \frac{N - |P|}{N} \sum_{k \in P} \sum_{\widetilde{\pi} \leqslant \sigma} H_k \prod_{Q \in \sigma} g_Q,
\end{aligned}$$

where if $\pi = \{P, W_1, ..., W_m\}$, then, letting $\widetilde{P} := P \cup \{*\}$, we define

$$\widetilde{\pi} := \{\widetilde{P}, W_1, ..., W_m\} \vdash [j] \cup \{*\}.$$

Continuing the above computation and reindexing sums, we get

$$\begin{aligned}
&\sum_{\pi \vdash [j]} (-1)^{|\pi|-1}(|\pi| - 1)! \sum_{P \in \pi} \frac{N - |P|}{N} \sum_{k \in P} \sum_{\widetilde{\pi} \leqslant \sigma} H_k \prod_{Q \in \sigma} g_Q \\
&\quad = \sum_{k=1}^{j} \sum_{\sigma \vdash [j] \cup \{*\}} \left(\sum_{\substack{\widetilde{\pi} \leqslant \sigma \\ \exists \widetilde{P} \in \widetilde{\pi}, \{k, *\} \subseteq \widetilde{P}}} (-1)^{|\widetilde{\pi}|-1}(|\widetilde{\pi}| - 1)! \frac{N - |\widetilde{P}| + 1}{N} \right) H_k \prod_{Q \in \sigma} g_Q \\
&\quad =: \sum_{k=1}^{j} \sum_{\sigma \vdash [j] \cup \{*\}} b_k^{\sigma} H_k \prod_{Q \in \sigma} g_Q,
\end{aligned}$$

where we note $\widetilde{\pi}$ is a partition of the larger set $[j] \cup \{*\}$. We now compute b_k^{σ}. Similarly to above, we split according to whether the relevant variables $k, *$ are in the same block of σ. Writing $\sigma = \{Q, R, W_1, ..., W_m\}$, the first case is that $\{k, *\} \subseteq Q$ and the second case is that $k \in Q, * \in R$.

In the first case, we note that if $\widetilde{\pi} \leqslant \sigma$, then by definition there exists $\widetilde{P} \in \widetilde{\pi}$ such that $Q \subseteq \widetilde{P}$, and as such $\{k, *\} \subseteq \widetilde{P}$. Thus

$$\begin{aligned} b_k^\sigma &= \sum_{\widetilde{\pi} \leqslant \sigma} (-1)^{|\widetilde{\pi}|-1}(|\widetilde{\pi}|-1)! \frac{N - |\widetilde{P}| + 1}{N} \\ &= \sum_{\rho \leqslant \sigma - \{Q\}} (-1)^{|\rho|}|\rho|! \frac{N - |Q| + 1}{N} + \sum_{S \in \rho} (-1)^{|\rho|-1}(|\rho|-1)! \frac{N - |S \cup Q| + 1}{N}, \end{aligned} \tag{4.3}$$

where for the second equality, in order to deal with the term $|\widetilde{P}|$, we look at possible ways of constructing $\widetilde{P}$. We note that any combining $\widetilde{\pi} \leqslant \sigma$ is generated by first taking a combing $\rho \leqslant \sigma - \{Q\}$ and then either adding Q as its own block or unioning Q with a block of ρ. This corresponds to the first and second term in the sum respectively.

The above computation is valid in the case that $Q = [j] \cup \{*\}$, but for the sake the analysis to follow let us deal with this edge case now. A direct computation verifies that in that case, we have that $|\sigma| = 1$ and

$$b_k^\sigma = \frac{N - j}{N}.$$

Continuing the above computation with the additional assumption that $|\sigma| \geqslant 2$, we first note that

$$\begin{aligned} &\sum_{S \in \rho} (-1)^{|\rho|-1}(|\rho|-1)! \frac{N - |S \cup Q| + 1}{N} \\ &\quad = \sum_{S \in \rho} (-1)^{|\rho|-1}(|\rho|-1)! \frac{N - |Q| + 1}{N} - \sum_{S \in \rho} (-1)^{|\rho|-1}(|\rho|-1)! \frac{|S|}{N} \\ &\quad = (-1)^{|\rho|-1}|\rho|! \frac{N - |Q| + 1}{N} - (-1)^{|\rho|-1}(|\rho|-1)! \frac{j + 1 - |Q|}{N}. \end{aligned} \tag{4.4}$$

For the last equality, we use that the first term doesn't depend on S, and as such we just get a multiplicative factor of $|\rho|$, which then goes into the factorial. For the second term, we use that

$$\sum_{S \in \rho} |S| = |[j] \cup \{*\} - Q| = j + 1 - |Q|.$$

Then, plugging (4.4) into (4.3) and noting the cancellation of the first two terms, we have that

$$b_k^\sigma = -\frac{j + 1 - |Q|}{N} \sum_{\rho \leqslant \sigma - \{Q\}} (-1)^{|\rho|-1}(|\rho|-1)! = \begin{cases} -\frac{j+1-|Q|}{N} & |\sigma| = 2 \\ 0 & |\sigma| \geqslant 3, \end{cases}$$

where we have once again used Lemma 4.2. Then recalling the above remarks on the case that $|\sigma| = 1$, we have that

$$b_k^\sigma = \begin{cases} \frac{N-j}{N} & |\sigma| = 1 \\ -\frac{j+1-|Q|}{N} & |\sigma| = 2 \\ 0 & |\sigma| \geqslant 3. \end{cases}$$

We now consider the other case, that $k \in Q, * \in R$. We then, similarly to the analysis for the $S_{k,\ell}^\sigma$ terms, define

$$\widetilde{\sigma} := \{Q \cup R, W_1, \ldots, W_m\}.$$

Then we note that

$$b_k^\sigma = \sum_{\substack{\widetilde{\pi} \leqslant \sigma \\ \exists \widetilde{P} \in \widetilde{\pi}, \{k,*\} \subseteq \widetilde{P}}} (-1)^{|\widetilde{\pi}|-1}(|\widetilde{\pi}|-1)! \frac{N - |\widetilde{P}| + 1}{N} = \sum_{\widetilde{\pi} \leqslant \widetilde{\sigma}} (-1)^{|\widetilde{\pi}|-1}(|\widetilde{\pi}|-1)! \frac{N - |\widetilde{P}| + 1}{N}.$$

We note now that we are in the same setting as we were for the previous case, except with $\widetilde{\sigma}$ in place of σ and $Q \cup R$ in place of Q. As such, the same computations demonstrate that, in this case,

$$b_k^\sigma = \begin{cases} \frac{N-j}{N} & |\sigma| = 2 \\ -\frac{j+1-|Q|-|R|}{N} & |\sigma| = 3 \\ 0 & |\sigma| \geqslant 4, \end{cases}$$

where we note that $|\sigma| = |\widetilde{\sigma}| + 1$. Thus, in total, we have that

$$b_k^\sigma = \begin{cases} \frac{N-j}{N} & |\sigma| = 1 \\ -\frac{j+1-|Q|}{N} & \sigma = \{Q, R\}, \{k, *\} \subseteq Q \\ \frac{N-j}{N} & \sigma = \{Q, R\}, k \in Q, * \in R \\ 0 & \sigma = \{Q, R, W\}, \{k, *\} \subseteq Q \\ -\frac{j+1-|Q|-|R|}{N} & \sigma = \{Q, R, W\}, k \in Q, * \in R \\ 0 & |\sigma| \geqslant 4. \end{cases} \tag{4.5}$$

We have thus computed all the coefficients, so we can plug in (4.2) and (4.5) into (4.1) to give the PDE g_j solves.

For the initial conditions, we remark that as $f_j(0, \cdot) = f^{\otimes j}$, the equation (2.2) gives that

$$g_j = f^{\otimes j} \sum_{\pi \vdash [j]} (-1)^{|\pi|-1}(|\pi|-1)!$$

thus Lemma 4.2 gives the stated initial conditions. □

Proof of Proposition 2.7. Computing $(\partial_t - \Delta) f_j^i$ using its definition we get

$$(\partial_t - \Delta) f_j^i = \sum_{\pi \vdash [j]} \sum_{\substack{(i_P)_{P \in \pi} \\ \sum i_P = i}} \sum_{P \in \pi} (\partial_t - \Delta) g_P^{i_P} \prod_{Q \in \pi - \{P\}} g_Q^{i_Q}.$$

Then using (2.4), this becomes

$$
\begin{aligned}
\sum_{\pi \vdash [j]} \sum_{\substack{(i_P)_{P\in\pi} \\ \sum i_P = i}} \sum_{P\in\pi} \Bigg(& -\sum_{k\in P} H_k g_{P\cup\{*\}}^{i_P} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \\
& - \sum_{k\in P} \sum_{W\subseteq P-\{k\}} \sum_{m=0}^{i_P} H_k g_{W\cup\{k\}}^m g_{P\cup\{*\}-W-\{k\}}^{i_P-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \\
& + \sum_{k\in P} |P| H_k g_{P\cup\{*\}}^{i_P-1} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \\
& + \sum_{k\in P} \sum_{W\subseteq P-\{k\}} (|P|-1-|W|) \sum_{m=0}^{i_P-1} H_k g_{W\cup\{k,*\}}^m g_{|P|-\{k\}-W}^{i_P-1-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \\
& + \sum_{k\in P} \sum_{W\subseteq P-\{k\}} \sum_{m=0}^{i_P-1} |P| H_k g_{W\cup\{k\}}^m g_{P\cup\{*\}-W-\{k\}}^{i_P-1-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \\
& + \sum_{k\in P} \sum_{W\subseteq P-\{k\}} \sum_{R\subseteq P-\{k\}-W} (|P|-1-|W|-|R|) \\
& \qquad \times \sum_{m=0}^{i_P-1} \sum_{n=0}^{i_P-1-m} H_k g_{W\cup\{k\}}^m g_{R\cup\{*\}}^n g_{P-R-W-\{k\}}^{i_P-1-m-n} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \\
& - \sum_{k,\ell\in P} S_{k,\ell} g_P^{i_P-1} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \\
& - \sum_{\substack{k,\ell\in P \\ k\neq \ell}} \sum_{W\subseteq P-\{k,\ell\}} \sum_{m=0}^{i_P-1} S_{k,\ell} g_{W\cup\{k\}}^m g_{P-\{k\}-W}^{i_P-1-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \Bigg).
\end{aligned}
$$

In order to conclude, we must show the above is equal to

$$
\begin{aligned}
-\sum_{k=1}^{j} H_k f_{[j]\cup\{*\}}^i & + j \sum_{k=1}^{j} H_k f_j^{i-1} - \sum_{k,\ell=1}^{j} S_{k,\ell} f_j^{i-1} \\
& = -\sum_k \sum_{\pi\vdash[j]\cup\{*\}} \sum_{\substack{(i_P)_{P\in\pi} \\ \sum i_P = i}} H_k \prod_{P\in\pi} g_P^{i_P} \\
& \quad + j \sum_k \sum_{\pi\vdash[j]\cup\{*\}} \sum_{\substack{(i_P)_{P\in\pi} \\ \sum i_P = i-1}} H_k \prod_{P\in\pi} g_P^{i_P} \\
& \quad - \sum_{k,\ell} \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi} \\ \sum i_P = i-1}} S_{k,\ell} \prod_{P\in\pi} g_P^{i_P}.
\end{aligned}
$$

In particular, we show the following three claims.

Claim 1:

$$\sum_k \sum_{\pi\vdash[j]\cup\{*\}} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} H_k \prod_{P\in\pi} g_P^{i_P} = \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{k\in P} H_k g_{P\cup\{*\}}^{i_P} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q}$$

$$+ \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{k\in P} \sum_{W\subseteq P-\{k\}} \sum_{m=0}^{i_P} H_k g_{W\cup\{k\}}^m g_{P\cup\{*\}-W-\{k\}}^{i_P-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q}.$$

Claim 2:

$$j\sum_k \sum_{\pi\vdash[j]\cup\{*\}} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i-1}} H_k \prod_{P\in\pi} g_P^{i_P}$$

$$= \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{k\in P} |P| H_k g_{P\cup\{*\}}^{i_P-1} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \tag{4.6}$$

$$\sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{k\in P} \sum_{W\subseteq P-\{k\}} (|P|-1-|W|) \sum_{m=0}^{i_P-1} H_k g_{W\cup\{k,*\}}^m g_{|P|-\{k\}-W}^{i_P-1-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \tag{4.7}$$

$$\sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{k\in P} \sum_{W\subseteq P-\{k\}} \sum_{m=0}^{i_P-1} |P| H_k g_{W\cup\{k\}}^m g_{P\cup\{*\}-W-\{k\}}^{i_P-1-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q} \tag{4.8}$$

$$\sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{k\in P} \sum_{W\subseteq P-\{k\}} \sum_{R\subseteq P-\{k\}-W} (|P|-1-|W|-|R|) \tag{4.9}$$

$$\times \sum_{m=0}^{i_P-1} \sum_{n=0}^{i_P-1-m} H_k g_{W\cup\{k\}}^m g_{R\cup\{*\}}^n g_{P-R-W-\{k\}}^{i_P-1-m-n} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q}.$$

Claim 3:

$$\sum_{k,\ell} \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i-1}} S_{k,\ell} \prod_{P\in\pi} g_P^{i_P} = \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{k,\ell\in P} S_{k,\ell} g_P^{i_P-1} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q}$$

$$+ \sum_{\pi\vdash[j]} \sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}} \sum_{P\in\pi}\sum_{\substack{k,\ell\in P\\ k\neq\ell}} \sum_{W\subseteq P-\{k,\ell\}} \sum_{m=0}^{i_P-1} S_{k,\ell} g_{W\cup\{k\}}^m g_{P-\{k\}-W}^{i_P-1-m} \prod_{Q\in\pi-\{P\}} g_Q^{i_Q}.$$

For Claim 1, we note that the first term of the right hand side is simply a sum over all partitions $\pi\vdash[j]\cup\{*\}$ such that k and $*$ are in the same block of π (together with all choices of orders i_P). Then the second term on the right hand side is a sum over all partitions π such that k and $*$ are in the different blocks of π. Thus together they give a sum over all partitions, which is equal then to the left hand side. Symbolically

$$\begin{aligned}
&\sum_{\pi\vdash[j]}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{P\in\pi}\sum_{k\in P}H_k g_{P\cup\{*\}}^{i_P}\prod_{Q\in\pi-\{P\}}g_Q^{i_Q}\\
&\qquad+\sum_{\pi\vdash[j]}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{P\in\pi}\sum_{k\in P}\sum_{W\subseteq P-\{k\}}\sum_{m=0}^{i_P}H_k g_{W\cup\{k\}}^{m}g_{P\cup\{*\}-W-\{k\}}^{i_P-m}\prod_{Q\in\pi-\{P\}}g_Q^{i_Q}\\
&=\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists P\in\pi,\{k,*\}\subseteq P}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}H_k\prod_{P\in\pi}g_P^{i_P}\\
&\qquad+\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists P\in\pi,k\in P,*\notin P}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}H_k\prod_{P\in\pi}g_P^{i_P}\\
&=\sum_k\sum_{\pi\vdash[j]\cup\{*\}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}H_k\prod_{P\in\pi}g_P^{i_P}.
\end{aligned}$$

For Claim 2, we note the first and second terms, (4.6) and (4.7), sum over the same partitions, namely those partitions $\pi\vdash[j]\cup\{*\}$ such that $k,*$ are in the same block of π. Thus there is "overcounting" and we have to compute the correct constant prefactor on each such partition. Reindexing (4.6), we get

$$\sum_{\pi\vdash[j]}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{P\in\pi}\sum_{k\in P}|P|H_k g_{P\cup\{*\}}^{i_P-1}\prod_{Q\in\pi-\{P\}}g_Q^{i_Q}=\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists P\in\pi,\{k,*\}\subseteq P}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i-1}}|P|H_k\prod_{P\in\pi}g_P^{i_P}.$$

Then reindexing (4.7), we get

$$\begin{aligned}
&\sum_{\pi\vdash[j]}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{P\in\pi}\sum_{k\in P}\sum_{W\subseteq P-\{k\}}(|P|-1-|W|)\sum_{m=0}^{i_P-1}H_k g_{W\cup\{k,*\}}^{m}g_{P-\{k\}-W}^{i_P-1-m}\prod_{Q\in\pi-\{P\}}g_Q^{i_Q}\\
&\quad=\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists A\in\pi,\{k,*\}\subseteq A}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i-1}}\sum_{B\in\pi-\{A\}}|B|H_k g_A^{i_A}g_B^{i_B}\prod_{Q\in\pi-\{A,B\}}g_Q^{i_Q}\\
&\quad=\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists P\in\pi,\{k,*\}\subseteq P}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i-1}}(j-|P|)H_k\prod_{Q\in\pi}g_Q^{i_Q}.
\end{aligned}$$

Thus adding them together, we get

$$j\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists P\in\pi,\{k,*\}\subseteq P}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i-1}}H_k\prod_{P\in\pi}g_P^{i_P}. \tag{4.10}$$

Similarly, the third and fourth terms, (4.8) and (4.9), sum over the same set of partitions, namely those partitions $\pi\vdash[j]\cup\{*\}$ such that $k,*$ are in different blocks. So we again reindex to compute the constant prefactors. Reindexing (4.8), we get

$$\sum_{\pi\vdash[j]}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{P\in\pi}\sum_{k\in P}\sum_{W\subseteq P-\{k\}}\sum_{m=0}^{i_P-1}|P|H_k g^m_{W\cup\{k\}}g^{i_P-1-m}_{P\cup\{*\}-W-\{k\}}\prod_{Q\in\pi-\{P\}}g_Q^{i_Q}$$
$$=\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists A,B\in\pi,k\in A,*\notin B,A\neq B}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}(|A|+|B|-1)H_k\prod_{P\in\pi}g_P^{i_P}.$$

Finally, reindexing (4.9), we get

$$\sum_{\pi\vdash[j]}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{P\in\pi}\sum_{k\in P}\sum_{W\subseteq P-\{k\}}\sum_{R\subseteq P-\{k\}-W}(|P|-1-|W|-|R|)$$
$$\times\sum_{m=0}^{i_P-1}\sum_{n=0}^{i_P-1-m}H_k g^m_{W\cup\{k\}}g^n_{R\cup\{*\}}g^{i_P-1-m-n}_{P-R-W-\{k\}}\prod_{Q\in\pi-\{P\}}g_Q^{i_Q}$$
$$=\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists A,B\in\pi,k\in A,*\notin B,A\neq B}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{C\in\pi-\{A,B\}}|C|H_k g_A^{i_A}g_B^{i_B}g_C^{i_C}\prod_{Q\in\pi-\{A,B,C\}}g_Q^{i_Q}$$
$$=\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists A,B\in\pi,k\in A,*\notin B,A\neq B}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i}}\sum_{C\in\pi-\{A,B\}}(j+1-|A|-|B|)H_k\prod_{P\in\pi}g_P^{i_P}.$$

So adding these together, we get

$$j\sum_k\sum_{\substack{\pi\vdash[j]\cup\{*\}\\ \exists P\in\pi,k\in P,*\notin P}}\sum_{\substack{(i_P)_{P\in\pi}\\ \sum i_P=i-1}}H_k\prod_{P\in\pi}g_P^{i_P}.\tag{4.11}$$

Thus, adding (4.10) and (4.11), we have shown Claim 2.

Lastly, Claim 3 follows exactly as Claim 1. ☐

A. Existence of the mean-field limit

Proof of Proposition 2.9. Let us use the notation $K*\rho(x):=\int K(x,x_*)\rho(x_*)\,dx_*$, so that the equation becomes

$$\partial_t\rho-\Delta\rho+\nabla\cdot(K*\rho\rho)=0.\tag{A.1}$$

Note that for $\rho\in C([0,\infty),L^2(\mathbb{T}^d))\cap L^2_{loc}([0,T],H^1(\mathbb{T}^d))$ such that ρ solves (A.1), treating $K*\rho\in L^\infty$ as a drift, we can view ρ as solving a drift-diffusion equation. Standard linear parabolic theory gives that, since $\rho(0,\cdot)=f\geqslant 0$, $\rho\geqslant 0$ for all times. Then, since the equation is mean preserving, we get that for all times

$$\|\rho\|_{L^1(\mathbb{T}^d)}=\int\rho\,dx=\int f\,dx=1.$$

Further, we have that

$$\begin{aligned}\frac{d}{dt}\|\rho\|_{L^2(\mathbb{T}^d)}^2 &= 2\int -|\nabla\rho|^2 + \nabla\rho\cdot(K*\rho\rho)\\ &\leqslant \int (K*\rho\rho)^2\\ &\leqslant \|K\|_{L^\infty}^2\|\rho\|_{L^1(\mathbb{T}^d)}^2\|\rho\|_{L^2(\mathbb{T}^d)}^2\\ &= \|K\|_{L^\infty}^2\|\rho\|_{L^2(\mathbb{T}^d)}^2,\end{aligned}$$

so by Grönwall's inequality,

$$\|\rho\|_{L^2(\mathbb{T}^d)}^2(t) \leqslant e^{\|K\|_{L^\infty}^2 t}\|f\|_{L^2(\mathbb{T}^d)}^2.$$

For uniqueness, supposing that $\rho, \rho' \in C([0,\infty), L^2(\mathbb{T}^d)) \cap L^2_{loc}([0,T], H^1(\mathbb{T}^d))$ are both solutions to (A.1). Then

$$\partial_t(\rho-\rho') = \Delta(\rho-\rho') - \nabla\cdot\Big(K*(\rho-\rho')\rho + K*\rho'(\rho-\rho')\Big),$$

so

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}\|\rho-\rho'\|_{L^2(\mathbb{T}^d)}^2 &\leqslant \int \Big(K*(\rho-\rho')\rho\Big)^2 + \Big(K*\rho'(\rho-\rho')\Big)^2\,dx\\ &\leqslant C\|K\|_{L^\infty}^2\|\rho-\rho'\|_{L^2(\mathbb{T}^d)}^2\|\rho\|_{L^2(\mathbb{T}^d)}^2 + \|K\|_{L^\infty}^2\|\rho'\|_{L^1(\mathbb{T}^d)}^2\|\rho-\rho'\|_{L^2(\mathbb{T}^d)}^2\\ &\leqslant C\|K\|_{L^\infty}^2 e^{\|K\|_{L^\infty}^2 t}\|f\|_{L^2(\mathbb{T}^d)}^2\|\rho-\rho'\|_{L^2(\mathbb{T}^d)}^2,\end{aligned}$$

thus we can conclude by a Grönwall argument.

For existence, we use a fixed-point argument. Let

$$\rho_0(t,\cdot) := f$$

and for any $j \in \mathbb{N}$, let $\rho_{j+1} \in C([0,\infty), L^2(\mathbb{T}^d)) \cap L^2_{loc}([0,T], H^1(\mathbb{T}^d))$ solve

$$\begin{cases}\partial_t\rho_{j+1} - \Delta\rho_{j+1} + \nabla\cdot(K*\rho_j\rho_{j+1}) = 0\\ \rho_{j+1}(0,\cdot) = f.\end{cases}$$

Note that, inductively, for all t, $\|\rho_j\|_{L^1(\mathbb{T}^d)} = 1$, since ρ_{j+1} solves a drift-diffusion equation with L^∞ drift, the equation is L^1 non-increasing, and the initial data satisfies $\|f\|_{L^1(\mathbb{T}^d)} = 1$. Then we also have the estimates

$$\begin{aligned}\frac{d}{dt}\|\rho_{j+1}\|_{L^2(\mathbb{T}^d)}^2 &= \int -2|\nabla\rho_{j+1}|^2 + 2\nabla\rho_j\cdot(K*\rho_j\rho_{j+1})\,dx\\ &\leqslant \int -|\nabla\rho_{j+1}|^2 + (K*\rho_j\rho_{j+1})^2\,dx\\ &\leqslant -\|\nabla\rho_{j+1}\|_{L^2(\mathbb{T}^d)}^2 + \|K\|_{L^\infty}^2\|\rho_{j+1}\|_{L^2(\mathbb{T}^d)}^2.\end{aligned}$$

Thus, by Grönwall's inequality,

$$\|\rho_{j+1}\|_{L^2(\mathbb{T}^d)}^2(t) \leqslant e^{\|K\|_{L^\infty}^2 t}\|f\|_{L^2(\mathbb{T}^d)}^2.$$

Also, we have that

$$\begin{aligned}\|\nabla\rho_{j+1}\|^2_{L^2([0,t],L^2(\mathbb{T}^d))} &\leqslant \int_0^t \|K\|^2_{L^\infty}\|\rho_{j+1}\|^2_{L^2(\mathbb{T}^d)}(s) - \frac{d}{ds}\|\rho_{j+1}\|^2_{L^2(\mathbb{T}^d)}(s)\,ds \\ &\leqslant \|f\|^2_{L^2(\mathbb{T}^d)}\Big(\int_0^t \|K\|^2_{L^\infty}e^{\|K\|^2_{L^\infty}s}\,ds+1\Big) \\ &= e^{\|K\|^2_{L^\infty}t}\|f\|^2_{L^2(\mathbb{T}^d)}. \end{aligned}\tag{A.2}$$

Then we note that

$$\partial_t(\rho_{j+1}-\rho_j) = \Delta(\rho_{j+1}-\rho_j) - \nabla\cdot(K*\rho_j(\rho_{j+1}-\rho_j) + K*(\rho_j-\rho_{j-1})\rho_j).$$

Thus

$$\begin{aligned}\frac{d}{dt}\|\rho_{j+1}-\rho_j\|^2_{L^2(\mathbb{T}^d)} &\leqslant \|K*\rho_j(\rho_{j+1}-\rho_j)\|^2_{L^2(\mathbb{T}^d)} + \|K*(\rho_j-\rho_{j-1})\rho_j\|^2_{L^2(\mathbb{T}^d)} \\ &\leqslant \|K\|^2_{L^\infty}\|\rho_{j+1}-\rho_j\|^2_{L^2(\mathbb{T}^d)} + C\|K\|^2_{L^\infty}\|\rho_j\|^2_{L^2(\mathbb{T}^d)}\|\rho_j-\rho_{j-1}\|^2_{L^2(\mathbb{T}^d)} \\ &\leqslant \|K\|^2_{L^\infty}\|\rho_{j+1}-\rho_j\|^2_{L^2(\mathbb{T}^d)} + C\|K\|^2_{L^\infty}e^{\|K\|^2_{L^\infty}t}\|f\|^2_{L^2(\mathbb{T}^d)}\|\rho_j-\rho_{j-1}\|^2_{L^2(\mathbb{T}^d)}.\end{aligned}$$

Thus by Grönwall's inequality,

$$\begin{aligned}\|\rho_{j+1}-\rho_j\|^2_{L^2(\mathbb{T}^d)}(t) &\leqslant C\|K\|^2_{L^\infty}\|f\|^2_{L^2(\mathbb{T}^d)}e^{\|K\|^2_{L^\infty}t}\int_0^t\|\rho_j-\rho_{j-1}\|^2_{L^2(\mathbb{T}^d)}(s)\,ds \\ &\leqslant C\|f\|^2_{L^2(\mathbb{T}^d)}e^{2\|K\|^2_{L^\infty}t}t\sup_{s\in[0,t]}\|\rho_j-\rho_{j-1}\|^2_{L^2(\mathbb{T}^d)}(s),\end{aligned}$$

therefore

$$\|\rho_{j+1}-\rho_j\|_{C([0,t],L^2(\mathbb{T}^d))} \leqslant C\|f\|_{L^2(\mathbb{T}^d)}e^{\|K\|^2_{L^\infty}t}\sqrt{t}\|\rho_j-\rho_{j-1}\|_{C([0,t],L^2(\mathbb{T}^d))}.$$

Let

$$t_* := \frac{1}{4C\|f\|^2_{L^2(\mathbb{T}^d)}e^{2\|K\|^2_{L^\infty}}},$$

then $0<t_*\leqslant 1$ and so

$$C\|f\|_{L^2(\mathbb{T}^d)}e^{\|K\|^2_{L^\infty}t_*}\sqrt{t_*}\leqslant \tfrac{1}{2},$$

thus

$$\|\rho_{j+1}-\rho_j\|_{C([0,t_*],L^2(\mathbb{T}^d))}\leqslant \tfrac{1}{2}\|\rho_j-\rho_{j-1}\|_{C([0,t_*],L^2(\mathbb{T}^d))}.$$

This contraction then implies that there exists $\rho\in C([0,t_*],L^2(\mathbb{T}^d))$ such that, in this norm, $\rho_j\to\rho$. Note that, by (A.2), ρ_j is also uniformly bounded in $L^2([0,t_*],H^1(\mathbb{T}^d))$, thus by weak compactness and taking a subsequence, we get that $\rho\in L^2([0,t_*],H^1(\mathbb{T}^d))$. That ρ distributionally solves (A.1) is direct from testing the equation for ρ_j against a C_c^∞ function and using the strong convergence. Thus we have a solution ρ for a short time $t_* = \frac{1}{C\|f\|^2_{L^2}}$. We can iterate this result to get existence for all time, as long as the existence time t_* doesn't go to zero, which happens as long as $\|\rho\|_{L^2(\mathbb{T}^d)}$ stays bounded, uniformly in time. To that end, we note that for some $a\in(1/2,1)$, by the Gagliardo-Nirenberg embeddings,

$$\|\rho\|_{L^2(\mathbb{T}^d)} - C\leqslant\|\rho-1\|_{L^2(\mathbb{T}^d)}\leqslant C\|\nabla\rho\|^a_{L^2(\mathbb{T}^d)}\|\rho\|^{1-a}_{L^1(\mathbb{T}^d)} = C\|\nabla\rho\|^a_{L^2(\mathbb{T}^d)},$$

so that

$$\|\nabla\rho\|^2_{L^2(\mathbb{T}^d)}\geqslant C^{-1}(\|\rho\|_{L^2(\mathbb{T}^d)}-C)^{2a^{-1}}\geqslant C^{-1}\|\rho\|^{2a^{-1}}_{L^2(\mathbb{T}^d)} - C.$$

Then we have that

$$\begin{aligned}\frac{d}{dt}\|\rho\|_{L^2(\mathbb{T}^d)}^2 &= 2\int -|\nabla\rho|^2 + \nabla\rho\cdot(K*\rho\rho)\\ &\leqslant -\|\nabla\rho\|_{L^2(\mathbb{T}^d)}^2 + \|K\|_{L^\infty}^2\|\rho\|_{L^2(T^d)}^2\\ &\leqslant -C^{-1}\|\rho\|_{L^2(\mathbb{T}^d)}^{2a^{-1}} + C + \|K\|_{L^\infty}^2\|\rho\|_{L^2(T^d)}^2.\end{aligned}$$

Then we note that since $a<1$, the right hand side is negative for $\|\rho\|_{L^2(\mathbb{T}^d)}^2$ big enough. Thus there exists C such that

$$\|\rho\|_{L^2(\mathbb{T}^d)} \leqslant C \vee \|f\|_{L^2(\mathbb{T}^d)}.$$

Thus we have the global-in-time bound on $\|\rho\|_{L^2(\mathbb{T}^d)}$, so the existence times stays bounded below, and we can iterate the local existence argued above to get global-in-time existence, allowing use to conclude. □

References

[BAZ99] Gerard Ben Arous and Ofer Zeitouni. Increasing propagation of chaos for mean field models. *Annales de l'I.H.P. Probabilités et statistiques*, 35(1):85–102, 1999.

[BGSRS20] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. Statistical dynamics of a hard sphere gas: fluctuating Boltzmann equation and large deviations, August 2020. arXiv:2008.10403 [math.AP].

[BJS22] Didier Bresch, Pierre-Emmanuel Jabin, and Juan Soler. A new approach to the mean-field limit of Vlasov-Fokker-Planck equations, March 2022. arXiv:2203.15747 [math.AP].

[BJW19] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. On Mean Field Limit and Quantitative Estimates with a Large Class of Singular Kernels: Application to the Patlak-Keller-Segel Model, June 2019. arXiv:1906.04093 [math.AP].

[Bog60] Nikolai Nikolaevich Bogolyubov. *Problems of a Dynamical Theory in Statistical Physics*. Thermal Radiation Laboratory, Geophysics Research Directorate, AF Cambridge Research Laboratories, Air Force Research Division (ARDC) U.S. Air Force, 1960. Google-Books-ID: 56PvAAAAMAAJ.

[CB18] Lénaïc Chizat and Francis Bach. On the Global Convergence of Gradient Descent for Over-parameterized Models using Optimal Transport. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

[CD22] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. I. Models and methods. *Kinetic and Related Models*, 15(6):895–1015, November 2022.

[dCRS23] Antonin Chodron de Courcel, Matthew Rosenzweig, and Sylvia Serfaty. Sharp uniform-in-time mean-field convergence for singular periodic Riesz flows, April 2023. arXiv:2304.05315 [math.AP].

[DEGZ20] Alain Durmus, Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. An elementary approach to uniform in time propagation of chaos. *Proceedings of the American Mathematical Society*, 148(12):5387–5398, December 2020.

[DMP91] Anna De Masi and Errico Presutti. *Mathematical Methods for Hydrodynamic Limits*, volume 1501 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 1991.

[Due21] Mitia Duerinckx. On the Size of Chaos via Glauber Calculus in the Classical Mean-Field Dynamics. *Communications in Mathematical Physics*, 382(1):613–653, February 2021.

[GBM23] Arnaud Guillin, Pierre Le Bris, and Pierre Monmarché. On systems of particles in singular repulsive interaction in dimension one: log and Riesz gas. *Journal de l'École polytechnique — Mathématiques*, 10:867–916, 2023.

[Jab19] Jean-Francois Jabir. Rate of propagation of chaos for diffusive stochastic particle systems via Girsanov transformation, July 2019. arXiv:1907.09096 [math].

[JW16] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit and propagation of chaos for Vlasov systems with bounded forces. *Journal of Functional Analysis*, 271(12):3588–3627, December 2016.

[JW18] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Inventiones mathematicae*, 214(1):523–591, October 2018.

[Lac18] Daniel Lacker. On a strong form of propagation of chaos for McKean-Vlasov equations. *Electronic Communications in Probability*, 23:1–11, January 2018.

[Lac23] Daniel Lacker. Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probability and Mathematical Physics*, 4(2):377–432, May 2023.

[LLX20] Songzi Li, Xiang-Dong Li, and Yong-Xiao Xie. On the Law of Large Numbers for the Empirical Measure Process of Generalized Dyson Brownian Motion. *Journal of Statistical Physics*, 181(4):1277–1305, November 2020.

[LM72] Jacques-Louis Lions and Enrico Magenes. *Non-Homogeneous Boundary Value Problems and Applications*. Springer, Berlin, Heidelberg, 1972.

[MM01] Laurent Miclo and Pierre Del Moral. Genealogies and Increasing Propagation of Chaos For Feynman-Kac and Genetic Models. *The Annals of Applied Probability*, 11(4):1166–1198, November 2001.

[MP94] Carlo Marchioro and Mario Pulvirenti. *Mathematical Theory of Incompressible Nonviscous Fluids*, volume 96 of *Applied Mathematical Sciences*. Springer, New York, NY, 1994.

[MV21] Yuliya Mishura and Alexander Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*, 103:59–101, June 2021.

[Ons49] Lars Onsager. Statistical hydrodynamics. *Il Nuovo Cimento (1943-1954)*, 6(2):279–287, March 1949.

[Per07] Benoît Perthame. *Transport Equations in Biology*. Frontiers in Mathematics. Birkhäuser, Basel, 2007.

[PP19] Thierry Paul and Mario Pulvirenti. Asymptotic expansion of the mean-field approximation. *Discrete and Continuous Dynamical Systems*, 39(4):1891–1921, March 2019.

[PPS19] Thierry Paul, Mario Pulvirenti, and Sergio Simonella. On the Size of Chaos in the Mean Field Dynamics. *Archive for Rational Mechanics and Analysis*, 231(1):285–317, January 2019.

[PS17] Mario Pulvirenti and Sergio Simonella. The Boltzmann–Grad limit of a hard sphere system: analysis of the correlation error. *Inventiones mathematicae*, 207(3):1135–1237, March 2017.

[RS23] Matthew Rosenzweig and Sylvia Serfaty. Modulated logarithmic Sobolev inequalities and generation of chaos, July 2023. arXiv:2307.07587 [math-ph].

[RVE22] Grant Rotskoff and Eric Vanden-Eijnden. Trainability and Accuracy of Artificial Neural Networks: An Interacting Particle System Approach. *Communications on Pure and Applied Mathematics*, 75(9):1889–1935, 2022.

[SV16] Igal Sason and Sergio Verdú. f -Divergence Inequalities. *IEEE Transactions on Information Theory*, 62(11):5973–6006, November 2016.

[TBL06] Chad M. Topaz, Andrea L. Bertozzi, and Mark A. Lewis. A Nonlocal Continuum Model for Biological Aggregation. *Bulletin of Mathematical Biology*, 68(7):1601–1623, October 2006.