



On the Diffusive-Mean Field Limit for Weakly Interacting Diffusions Exhibiting Phase Transitions

MATIAS G. DELGADINO , RISHABH S. GVALANI & GRIGORIOS A. PAVLIOTIS

Communicated by P.-L. LIONS

Abstract

The objective of this article is to analyse the statistical behaviour of a large number of weakly interacting diffusion processes evolving under the influence of a periodic interaction potential. We focus our attention on the combined mean field and diffusive (homogenisation) limits. In particular, we show that these two limits do not commute if the mean field system constrained to the torus undergoes a phase transition, that is to say, if it admits more than one steady state. A typical example of such a system on the torus is given by the noisy Kuramoto model of mean field plane rotators. As a by-product of our main results, we also analyse the energetic consequences of the central limit theorem for fluctuations around the mean field limit and derive optimal rates of convergence in relative entropy of the Gibbs measure to the (unique) limit of the mean field energy below the critical temperature.

Contents

1. Introduction
1.1. Overview
1.2. Set Up and Preliminaries
1.3. The Space $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ as the Limit of $\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$
1.4. Gradient Flow Formulation and the Mean Field Limit
1.5. Scaling and the Quotiented Process
1.6. The Diffusive Limit
1.7. The Limit $N \rightarrow \infty$ Followed by $\varepsilon \rightarrow 0$
1.8. The Limit $\varepsilon \rightarrow 0$ Followed by $N \rightarrow \infty$
1.9. The Effect of Phase Transitions
1.10. Application of the Fluctuation Theorem
2. Proof of Theorem 1.10
3. Proof of Theorem 1.16
4. Proofs of Section 1.9
References

1. Introduction

1.1. Overview

The study of large systems of interacting particles in the presence of noise has attracted a large amount of interest in recent years. This is largely due to the fact that they pose challenging mathematical questions and that they appear in several applications, ranging from the theory of random matrices [43] and the construction of Kähler–Einstein metrics [5] to the design of algorithms for global optimisation [28,40], biological models of chemotaxis [20], and models of opinion formation [22].

We place ourselves in the setting of a system of weakly interacting diffusion processes as in [36]. It is well-known that, under appropriate assumptions on the interaction and confining potentials, one can pass to the mean field limit as $N \rightarrow \infty$ to obtain the so-called McKean–Vlasov equation (1.14) for the limit of the N -particle empirical measure. More precisely, given chaotic initial data, the empirical measure associated to the system of particles converges weakly to the weak solution of the McKean–Vlasov equation. Formally, one can say that the law of the N -particle system decouples and converges to N copies of the mean field McKean–Vlasov equation. This corresponds to a strong law of large numbers (LLN) for the empirical measure. A natural question to ask then is whether one can obtain a second order characterisation of this convergence, i.e. a central limit theorem (CLT).

Partial results in this direction do exist: Fernandez and Méléard [19] obtained a finite-time horizon version of the CLT. They showed that the fluctuations around the mean field limit are described in the large N -limit by a Gaussian random field which itself is the solution of a linear stochastic PDE. Additionally, Dawson [11] proved an equilibrium CLT for the empirical measure of a system of particles in a bistable confining potential and Curie–Weiss interaction. The interesting feature of Dawson’s system is that exhibits a phase transition, i.e. for a certain value of the interaction strength the system transitions from having one invariant measure to having multiple. Dawson showed that below the phase transition point equilibrium fluctuations are described by Gaussian random field, similar to the result in [19]. However, at the critical temperature the fluctuations become non-Gaussian and are given by the invariant measure of a nonlinear SDE. These non-Gaussian fluctuations are persistent and are characterised by a longer time scale, exhibiting the well known phenomenon of critical slowing down (cf. [44] for a less rigorous derivation of similar results). We are not aware of any results on the limiting behaviour of the fluctuations that have been obtained ahead of the phase transition.

Fluctuations around the McKean–Vlasov mean field limit for a system of weakly interacting diffusions with an internal degree of freedom were also studied recently in [3]. Under the assumption of scale separation between the macroscopic and microscopic dynamics, a large deviations principle (LDP) was established for the slow dynamics, valid in the combined limit of infinite scale separation ($\varepsilon \rightarrow 0$, where ε measures the ratio between the small and the large scale in the problem) and of the number of particles going to infinity ($N \rightarrow \infty$). This LDP was then used to deduce information about the fluctuations around the mean field limit and

to also offer partial justification for the so-called Dean equation, a stochastic partial differential equation used in dynamical density functional theory which combines, formally, the mean field limit and central limit theorem results for the system of weakly interacting diffusions. Furthermore, the connection between the LDP framework and the Chapman–Enskog approach to the study of the hydrodynamic limit was discussed in detail. The crucial assumption made by the authors was that the microscopic dynamics has a unique stationary state, i.e. that no phase transitions occur.

The prototype of the systems we consider is the following system of N interacting SDEs on \mathbb{R} :

$$dX_t^i = -\frac{1}{N} \sum_{j=1, j \neq i}^N \sin\left(2\pi\left(X_t^i - X_t^j\right)\right) dt + \sqrt{2\beta^{-1}} dB_t^i. \quad (1.1)$$

Here the B_t^i are independent \mathbb{R} -valued Wiener processes. The interesting feature about the above system is that the interaction potential is 1-periodic. As a consequence of this, the behaviour of (1.1) is influenced heavily by the corresponding quotiented process on \mathbb{T} (the one dimensional unit torus). The quotiented system on the torus is in fact the noisy Kuramoto model for mean field plane rotators [6, 9].¹ Indeed (cf. Proposition 1.22), one can show that the corresponding mean field limit on the torus exhibits a phase transition. A more complete picture of the local bifurcations and phase transitions for the McKean–Vlasov equation on the torus can be found in [9].²

In the spirit of Dawson, our main objective is to study fluctuations in the presence of phase transitions. However, instead of the phase transitions of the system on \mathbb{R} , we will be concerned with the phase transitions of the quotiented system on \mathbb{T} . Furthermore, we study the diffusive limit which can be thought of as the first step in understanding fluctuations of the N -particle system. Although we do discuss the implications of a full CLT (cf. Sect. 1.10), we concern ourselves in this paper mainly with the combined diffusive-mean field limits.

The problem that we study in this paper is closely related, and simpler, to the one studied in [3]: scale separation arises naturally in our case due to the disparity between the period of the interaction potential which is the characteristic length scale of the microscopic dynamics, and the long, diffusive length/time scale. The

¹ Additionally, its reversible Gibbs measure corresponds to the classical Heisenberg XY model for lattice systems with continuous spins and mean field interaction. This is immediately apparent when one considers the associated Hamiltonian which is given by

$$H^N(x_1, \dots, x_N) = -(2N)^{-1} \sum_{i,j} \cos(2\pi(x_i - x_j)) = -(2N)^{-1} \sum_{i,j} S_i \cdot S_j,$$

where $S_j = (\cos(2\pi x_j), \sin(2\pi x_j))$, cf. [21, Chapter 9] or [6].

² In later sections, as a technical requirement, we will consider the same system with an additional confining potential in order to break the translation symmetry of the noisy Kuramoto system which leads to degeneracy of minimisers ahead of the phase transition (cf. Proposition 1.22).

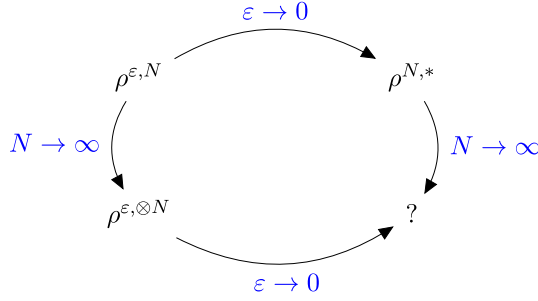


Fig. 1. The combined diffusive-mean field limit and the (possible) non-commutativity of the two limits

“hydrodynamics” in our problem is described by the (homogenised) heat equation, with the effective covariance matrix given by the standard homogenisation formula: compare Equation (1.19) below with formulas (3.14) and (3.15) in [3]. However, in contrast to [3] our main focus is on the effect of the presence of phase transitions at the microscopic scale on the effective/macroscopic dynamics. We are, in particular, interested on the effect of phase transitions on the (lack of) commutativity between the homogenisation and mean field limits.

Before we discuss what we mean by the combined limit, we remind the reader of what we mean by the diffusive limit. For a fixed number of particles $N > 0$ for the system in (1.1), a natural question to ask is how the law of the system behaves under the diffusive rescaling, i.e., if $\rho^{\varepsilon, N} = \text{Law}\left(\varepsilon X_{t/\varepsilon^2}^1, \dots, \varepsilon X_{t/\varepsilon^2}^N\right)$ then what is the limit as $\varepsilon \rightarrow 0$ of $\rho^{\varepsilon, N}$? The answer to this question can be obtained by using classical arguments from periodic homogenisation [38, Chapter 20] [4]. It turns out that $\rho^{\varepsilon, N}$ converges to $\rho^{N, *}$, the solution of the heat equation with a positive definite effective covariance matrix $A^{\text{eff}, N}$ (cf. Sect. 1.8 and Equation (1.18)), which can be obtained by solving a Poisson equation for the generator of the process on \mathbb{T}^N (cf. Equation (C)). Another way of reinterpreting this result is by saying that the system of particles (1.1) converges in law to an N -dimensional Brownian motion with covariance $A^{\text{eff}, N}$. A natural next question to ask is: how does the covariance matrix $A^{\text{eff}, N}$, and by extension the heat equation, behave in the limit as $N \rightarrow \infty$?

One could also ask the question the other way around. As discussed previously, for a fixed $\varepsilon > 0$, we can pass to the mean field limit as $N \rightarrow \infty$ in $\rho^{\varepsilon, N}$ to obtain N copies of the solution of the nonlinear McKean–Vlasov equation, $\rho^{\varepsilon, \otimes N}$. The natural question to ask now is whether we can understand the behaviour of $\rho^{\varepsilon, \otimes N}$ as $\varepsilon \rightarrow 0$. This dichotomy is illustrated in Figure 1. Starting from the rescaled law $\rho^{\varepsilon, N}$, we can take the limit $\varepsilon \rightarrow 0$ first followed by $N \rightarrow \infty$ if we move in the clockwise direction or the other way around in the anti-clockwise direction. Whether these two limits commute depends heavily on the ergodic properties of the quotiented process on \mathbb{T}^N and its behaviour in the mean field limit. Our main result asserts that the two limits commute at high temperatures (small β) and thus the combined limit is well-defined in this regime. However, at low temperatures

(large β) and in particular, in the presence of a phase transition (cf. Definition 1.18), we can construct initial data such that the two limits do not commute.

The problem of non-commutativity between the mean field and homogenisation limits was also studied in [24]. In this paper, a system of weakly interacting diffusions in a two-scale, locally periodic confining potential subject to a quadratic, Curie-Weiss, interaction potential was considered. It was shown that, although the combined homogenization-mean field limit leads to coarse-grained McKean-Vlasov dynamics that have the same functional form, the effective diffusion (mobility) tensor and the coarse-grained (Fixman) potential are different, depending on the order with which we consider these two limits (for non-separable two-scale potentials). In particular, the phase diagrams for the effective dynamics can be different, depending on the order with which we take the limits. A more striking manifestation of the non-commutativity between the two limits can be observed at small but finite values of ε , the parameter measuring scale separation: it is easy to construct examples where the mean field PDE, for small, finite ε can have arbitrarily many stationary states, the homogenised McKean–Vlasov equation (corresponding to the choice of sending first $\varepsilon \rightarrow 0$ and then $N \rightarrow \infty$) is characterised by a convex free energy functional and, thus, a unique steady state.

1.2. Set Up and Preliminaries

We denote by \mathbb{T}^d the d dimensional unit torus (which we identify with $[0, 1)^d$) and use the standard notation of $L^p(\mathbb{T}^d)$ and $H^s(\mathbb{T}^d)$ ($s \in \mathbb{R}$) for the Lebesgue and L^2 -Sobolev spaces, respectively. We will use $H_0^s(\mathbb{T}^d)$ ($s \in \mathbb{R}$) to denote the closed subspace of all $H^s(\mathbb{T}^d)$ functions with mean zero. These spaces are equipped with the following norms:

$$\|f\|_{H^s(\mathbb{T}^d)}^2 = \sum_{k \in \mathbb{Z}^d} (1 + 2\pi|k|)^{2s} |\hat{f}(k)|^2$$

$$\|f\|_{H_0^s(\mathbb{T}^d)}^2 = (2\pi)^{2s} \sum_{k \in \mathbb{Z}^d, k \neq 0} |k|^{2s} |\hat{f}(k)|^2.$$

We denote by the $C^k(\mathbb{T}^d)$, $C^\infty(\mathbb{T}^d)$ the space of k -times ($k \in \mathbb{N}$) continuously differentiable and smooth functions, respectively.

We denote by $\mathcal{P}(\Omega)$ the space of all Borel probability measures on Ω having finite second moment, with Ω some Polish metric space. We will use d_1 and d_2 to denote the 1 and 2-Wasserstein distances, respectively, on $\mathcal{P}(\mathbb{R}^d)$ and $\mathcal{P}(\mathbb{T}^d)$. Similarly we will use \mathfrak{D}_1 and \mathfrak{D}_2 for the 1 and 2-Wasserstein distances, respectively, on $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ and $\mathcal{P}(\mathcal{P}(\mathbb{T}^d))$. We will often identify a given $\rho \in \mathcal{P}(\mathbb{R}^d)$ (resp. $\mathcal{P}(\mathbb{T}^d)$) with its density with respect to the Lebesgue measure, $\rho \in L^1(\mathbb{R}^d)$ (resp. $L^1(\mathbb{T}^d)$). In the sequel, any limit of a sequence of measures $\{\rho_n\}_{n \in \mathbb{N}} \subset \mathcal{P}(\Omega)$ unless otherwise specified should be understood as a limit in the weak-* topology relative to $C_b(\Omega)$, i.e. tested against bounded, continuous functions. We will often use the same notation for a measure and its density if the density is well-defined.

We consider a large number $N \in \mathbb{N}$ of indistinguishable interacting particles $\{X_t^i\}_{i=1}^N$ in \mathbb{R}^d , where both the interaction and confining potentials are periodic and highly oscillatory. In particular, we consider the system

$$\begin{cases} dX_t^{\varepsilon,i} = -\varepsilon^{-1} \nabla V(\varepsilon^{-1} X_t^{\varepsilon,i}) - \frac{1}{N} \sum_{j \neq i}^N \varepsilon^{-1} \nabla W(\varepsilon^{-1} (X_t^{\varepsilon,i} - X_t^{\varepsilon,j})) dt \\ \quad + \sqrt{2\beta^{-1}} dB_t^i \\ \text{Law}((X_0^{\varepsilon,1}, \dots, X_0^{\varepsilon,N})) = \rho_0^{\varepsilon,N}, \end{cases} \quad (1.2)$$

where $W : \mathbb{R}^d \rightarrow \mathbb{R}$ and $V : \mathbb{R}^d \rightarrow \mathbb{R}$ are smooth 1-periodic interaction and confining potentials, respectively, $\varepsilon \ll 1$ is the period size, $\beta > 0$ is the inverse temperature, $\rho_0^{\varepsilon,N} \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ is the initial distribution of the particles which might depend on the period size, and $\{B_t^i\}_{i=1}^N$ are independent Wiener processes. Furthermore, we assume that W is even. We are interested in understanding the joint limit when the period of oscillations goes to 0 ($\varepsilon \rightarrow 0$) and the number of particles tends to infinity ($N \rightarrow \infty$).

We consider the joint law of the particle positions which is given by

$$\rho^{\varepsilon,N}(t) = \text{Law}((X_t^{\varepsilon,1}, \dots, X_t^{\varepsilon,N})) \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N),$$

where $\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ is as defined in (1.5). The law evolves through the linear forward Kolmogorov or Fokker–Planck equation

$$\begin{cases} \partial_t \rho^{\varepsilon,N} = \beta^{-1} \Delta \rho^{\varepsilon,N} + \nabla \cdot (\nabla H_\varepsilon^N \rho^{\varepsilon,N}) & \text{on } (0, \infty) \times (\mathbb{R}^d)^N \\ \rho^{\varepsilon,N}(0) = \rho_0^{\varepsilon,N} & \text{on } (\mathbb{R}^d)^N, \end{cases} \quad (1.3)$$

where $H_\varepsilon^N : (\mathbb{R}^d)^N \rightarrow \mathbb{R}$ is given by

$$H_\varepsilon^N(x_1, \dots, x_N) = \sum_{i=1}^N V(\varepsilon^{-1} x_i) + \frac{1}{2N} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N W(\varepsilon^{-1} (x_i - x_j)). \quad (1.4)$$

We note that under the conditions imposed of V , W , and $\rho_0^{\varepsilon,N}$, there exists a unique solution to (1.3). Indeed, we have the following result:

Theorem A. *Assume $\rho_0^{\varepsilon,N} \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ (with second moment bounded) and that $V, W : \mathbb{R}^d \rightarrow \mathbb{R}$ are smooth and 1-periodic. Then, there exists a unique weak solution to (1.3) $\rho^{\varepsilon,N} \in C([0, \infty); \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N))$ which is the gradient flow of (1.6) in the 2-Wasserstein metric. Furthermore, for all $t > 0$,*

$$\rho^{\varepsilon,N}(t) \in C^2(\mathbb{R}^d), \quad \partial_t \rho^{\varepsilon,N}(t) \in C^0(\mathbb{R}^d),$$

and $\rho^{\varepsilon,N}$ satisfies (1.3) in the classical sense.

Proof. The existence and uniqueness of gradient flow weak solutions follows from [1, Theorem 11.2.8] and the fact that H_ε^N is λ -convex. The required regularity then follows from classical parabolic regularity theory (cf., for example, [32, Theorem 5.6]) and the fact that V and W are both smooth. \square

We remind the reader that in the sequel we will always use the above notion of solution when we refer to (1.3).

The main objective of this paper is to study

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon, N} \quad \text{and} \quad \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho^{\varepsilon, N},$$

and understand under which regimes they coincide or differ. For the rest of this section we introduce the relevant notions that will play an important role in understanding these limits and present our main results. The result concerning the limit $N \rightarrow \infty$ followed by $\varepsilon \rightarrow 0$ can be found in Theorem 1.10, while the result concerning the limit $\varepsilon \rightarrow 0$ followed by $N \rightarrow \infty$ can be found in Theorem 1.16. We discuss the effect of the presence of a phase transition in Sect. 1.9. Finally, in Sect. 1.10 we discuss the implications of a CLT on the rate of convergence of the Gibbs measure before the phase transition. The proofs of the two main results, Theorems 1.10 and 1.16, can be found in Sects. 2 and 3, respectively. The proofs of other useful results related to the phenomenon of phase transitions are relegated to Sect. 4. Appendix A contains some coupling arguments which are useful for the proof of Theorem 1.10.

1.3. The Space $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ as the Limit of $\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$

The set up we consider is similar to that in [8]. We remark that due to the indistinguishability assumption on the particles their joint law is invariant under relabelling of the particles. In probability this is known as exchangeability, while in analysis this is referred to as symmetry and we denote the set of symmetric probability measures by $\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$, i.e.

$$\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N) := \left\{ \rho^N \in \mathcal{P}((\mathbb{R}^d)^N) : \rho^N(A) = \rho^N(\pi(A)), \forall \pi \in \Pi, A \text{ measurable} \right\}, \quad (1.5)$$

where A is any Borel set and Π is the set of permutations of the particle positions. Central to our work will be the classical result attributed to de Finetti [12] and Hewitt–Savage [27], that characterises the limit $N \rightarrow \infty$ of $\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$. Adapted to the set up of this paper, their result can be reformulated as follows:

Definition 1.1. Given a family $\{\rho^N\}_{N \in \mathbb{N}}$ such that $\rho^N \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ we say that

$$\rho^N \rightarrow X \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d)), \quad \text{as } N \rightarrow \infty,$$

if for every $n \in \mathbb{N}$ we have

$$\rho_n^N \xrightarrow{*} X^n \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^n), \quad \text{as } N \rightarrow \infty,$$

where $X^n \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^n)$ is defined by duality as follows

$$\langle X^n, \varphi \rangle = \int_{\mathcal{P}(\mathcal{P}(\mathbb{R}^d))} \int \varphi \, d\rho^{\otimes n} \, dX(\rho),$$

for all $\varphi \in C_b((\mathbb{R}^d)^n)$ and

$$\rho_n^N = \int_{(\mathbb{R}^d)^{N-n}} \rho^N dx_{N-n+1} \dots dx_N \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^n).$$

We will often suppress the $N \rightarrow \infty$ and just write $\rho^N \rightarrow X$.

In particular, we can relate this definition with the usual chaoticity assumption. We will say that $\{\rho^N\}_{N \in \mathbb{N}}$ is chaotic with limit $\rho \in \mathcal{P}(\mathbb{R}^d)$ if

$$\rho^N \rightarrow \delta_\rho \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d)),$$

in the sense of Definition 1.1. Additionally, the notion of convergence introduced in Definition 1.1 can also be interpreted in the following manner:

Definition 1.2. (*Empirical measure*) Given some $\rho^N \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ we define its empirical measure $\hat{\rho}^N \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ as follows:

$$\hat{\rho}^N := T_N \# \rho^N.$$

Here $T^N : (\mathbb{R}^d)^N \rightarrow \mathcal{P}(\mathbb{R}^d)$ is the measurable mapping $(x_1, \dots, x_N) \mapsto N^{-1} \sum_{i=1}^N \delta_{x_i}$. Furthermore, given a family $\{\rho^N\}_{N \in \mathbb{N}}$, we have that $\rho^N \rightarrow X \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ if and only if $\hat{\rho}^N \rightarrow^* X$.

We conclude this subsection with the following compactness result:

Lemma 1.3. (de Finetti–Hewitt–Savage) *Given a sequence $\{\rho^N\}_{N \in \mathbb{N}}$, with $\rho^N \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ for every N , assume that the sequence of the first marginals $\{\rho_1^N\}_{N \in \mathbb{N}} \in \mathcal{P}(\mathbb{R}^d)$ is tight. Then, up to a subsequence, not relabelled, there exists $X \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ such that $\rho^N \rightarrow X$ in the sense of Definition 1.1.*

For a proof and more details, see [8, 26, 41]. In the sequel, any limit of a sequence of symmetric measures $\{\rho^N\}_{N \in \mathbb{N}}$ with $\rho^N \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ should be understood in the sense of Definition 1.1.

Remark 1.4. The above notion of convergence, i.e. Definitions 1.1 and 1.2, can be naturally extended to $\mathcal{P}_{\text{sym}}((\mathbb{T}^d)^N)$.

1.4. Gradient Flow Formulation and the Mean Field Limit

In [8], the mean field limit (the limit $N \rightarrow \infty$) of the interacting particle system (1.3) is achieved by passing to the limit in the 2-Wasserstein gradient flow structure. The results of this article will build on this perspective which we briefly recall here:

The evolution of the joint law $\rho^{\varepsilon, N}$ given by (1.3) is the gradient flow (in the sense of [1, Definition 11.1.1]) of the energy $E^N : \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N) \rightarrow (-\infty, +\infty]$

$$E^N[\rho^N] := \frac{1}{N} \left(\beta^{-1} \int_{(\mathbb{R}^d)^N} \rho^N \log \rho^N dx + \int_{(\mathbb{R}^d)^N} H_\varepsilon^N(x) d\rho^N(x) \right), \quad (1.6)$$

under the rescaled 2-Wasserstein distance $\frac{1}{\sqrt{N}} d_2$ on $\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$. Moreover, we have the following classical result of Messer and Spohn [35]:

Lemma 1.5. *The N -particle free energy E^N Γ -converges to $E^\infty : \mathcal{P}(\mathcal{P}(\mathbb{R}^d)) \rightarrow (-\infty, +\infty]$, where*

$$E^\infty[X] = \int_{\mathcal{P}(\mathbb{R}^d)} E_{MF}[\rho] \, dX(\rho),$$

with $E_{MF} : \mathcal{P}(\mathbb{R}^d) \rightarrow (-\infty, +\infty]$ given by

$$\begin{aligned} E_{MF}[\rho] &= \beta^{-1} \int_{\mathbb{R}^d} \rho \log(\rho) \, dx + \int_{\mathbb{R}^d} V(\varepsilon^{-1}x) \, d\rho(x) \\ &\quad + \frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} W(\varepsilon^{-1}(x-y)) \, d\rho(y) \, d\rho(x). \end{aligned}$$

That is to say, for every $X \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ there exists a sequence $\{\rho^N\}_{N \in \mathbb{N}}, \rho^N \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ with $\rho^N \rightarrow X$ such that

$$\lim_{N \rightarrow \infty} E^N[\rho^N] = E^\infty[X].$$

Additionally, for every $X \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ and $\{\rho^N\}_{N \in \mathbb{N}}, \rho^N \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ with $\rho^N \rightarrow X$ it holds that

$$E^\infty[X] \leq \liminf_{N \rightarrow \infty} E^N[\rho^N].$$

On the other hand we have a similar convergence for the metrics: $\frac{1}{\sqrt{N}}d_2$ the rescaled 2-Wasserstein distance on $\mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ converges to \mathfrak{D}_2 the 2-Wasserstein distance on $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$. Specifically, given two sequences $\{\mu^N\}_{N \in \mathbb{N}}$ and $\{\nu^N\}_{N \in \mathbb{N}}$ of symmetric probability measures such that $\mu^N \rightarrow X_1$ and $\nu^N \rightarrow X_2$ and

$$\begin{aligned} \lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} |x|^2 \, d\mu_1^N &= \int_{\mathcal{P}(\mathcal{P}(\mathbb{R}^d))} \left(\int_{\mathbb{R}^d} |x|^2 \, d\rho(x) \right) \, dX_1(\rho) \\ \lim_{N \rightarrow \infty} \int_{\mathbb{R}^d} |x|^2 \, d\nu_1^N &= \int_{\mathcal{P}(\mathcal{P}(\mathbb{R}^d))} \left(\int_{\mathbb{R}^d} |x|^2 \, d\rho(x) \right) \, dX_2(\rho), \end{aligned}$$

then

$$\frac{1}{N} d_2^2(\mu^N, \nu^N) \rightarrow \mathfrak{D}_2^2(X_1, X_2).$$

We can now state our result concerning the mean field limit, i.e. the limit $N \rightarrow \infty$.

Theorem B. (Mean field limit [8]) *Under the regularity assumptions on W and V , fix any $\varepsilon > 0$ and assume that $\lim_{N \rightarrow \infty} \rho_0^{\varepsilon, N} = X_0^\varepsilon \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ and*

$$\sup_{N \in \mathbb{N}} \frac{1}{N^2} \int_{(\mathbb{R}^d)^N} |x|^4 \, d\rho_0^{\varepsilon, N}(x) < \infty.$$

Then, for every $t > 0$,

$$\lim_{N \rightarrow \infty} \rho^{\varepsilon, N}(t) = X^\varepsilon(t) \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d)),$$

where $X^\varepsilon : [0, \infty) \rightarrow \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ is the unique gradient flow of E^∞ under the 2-Wasserstein metric \mathfrak{D}_2 with initial condition X_0^ε .

Moreover,

$$X^\varepsilon(t) = S_t^\varepsilon \# X_0^\varepsilon, \quad (1.7)$$

where $X_0^\varepsilon = \lim_{N \rightarrow \infty} \rho_0^{\varepsilon, N}$ and $S_t^\varepsilon : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d)$ is the solution semigroup associated to the nonlinear McKean–Vlasov evolution equation

$$\partial_t \rho^\varepsilon = \beta^{-1} \Delta \rho^\varepsilon + \nabla \cdot (\rho^\varepsilon (\nabla W_\varepsilon * \rho^\varepsilon + \nabla V_\varepsilon)), \quad (1.8)$$

with $W_\varepsilon(x) = W(\varepsilon^{-1}x)$ and $V_\varepsilon(x) = V(\varepsilon^{-1}x)$.

Remark 1.6. As W and V are assumed to be smooth and periodic, using similar arguments as in the proof of Theorem A, one can show that for any $\mathcal{P}(\mathbb{R}^d)$ -valued initial condition there exists a unique gradient flow solution to (1.8) which becomes smooth for any positive time. Similar arguments can be used for the notions of solution to the PDEs discussed in the next subsection.

1.5. Scaling and the Quotiented Process

We notice that the Fokker–Planck equation (1.3) behaves well under the parabolic scaling, i.e. given a solution $\rho^{\varepsilon, N}$ of (1.3) we have that

$$v^N(s, y) = \varepsilon^{Nd} \rho^{\varepsilon, N}(\varepsilon^2 s, \varepsilon y), \quad (1.9)$$

is the solution to the Fokker–Planck equation at scale $\varepsilon = 1$, i.e.

$$\partial_s v^N = \beta^{-1} \Delta v^N + \nabla \cdot (\nabla H_1^N v^N), \quad (s, y) \in (0, \infty) \times (\mathbb{R}^d)^N. \quad (1.10)$$

The above equation naturally describes the evolution of the law of N -particle system (1.2) at scale $\varepsilon = 1$:

$$\begin{cases} dX_t^i = -\nabla V(X_t^i) - \frac{1}{N} \sum_{j \neq i}^N \nabla W(X_t^i - X_t^j) dt + \sqrt{2\beta^{-1}} dB_t^i \\ \text{Law}(X_0^1, \dots, X_0^N) = \varepsilon^{Nd} \rho_0^{\varepsilon, N}(\varepsilon x) := v_0^N. \end{cases} \quad (1.11)$$

Since W and V are periodic, in order to that to understand the behaviour of $\rho^{\varepsilon, N}$ in the limit as $\varepsilon \rightarrow 0$, we must first understand the behaviour of the quotiented process $\{\tilde{X}_t^i\}_{i=1}^N$ of (1.11) which lives on $(\mathbb{T}^d)^N$ [31, Section 9.1] [4, Section 3.3.2]. Before we introduce the quotiented process, we define the following notion which will play an important role in the rest of the paper:

Definition 1.7. Given a measure $\rho \in \mathcal{P}(\mathbb{R}^d)$ we define its periodic rearrangement at scale $\varepsilon > 0$ to be the measure $\tilde{\rho} \in \mathcal{P}(\mathbb{T}^d)$, such that for any measurable $A \subset \mathbb{T}^d$ it holds that

$$\tilde{\rho}(A) := \varepsilon^d \sum_{k \in \mathbb{Z}^d} \rho(\varepsilon(A + k)).$$

We will often just use the words periodic rearrangement when $\varepsilon = 1$.

Analogously, we have that $\{\dot{X}_t^i\}_{i=1}^N$ the quotient of the process $\{X_t^i\}_{i=1}^N$ is the unique Markov process associated to the generator $L : H^2((\mathbb{T}^d)^N) \rightarrow L^2((\mathbb{T}^d)^N)$

$$Lf := \beta^{-1} \Delta f - \nabla H_1^N \cdot \nabla f, \quad (1.12)$$

with initial law given by \tilde{v}_0^N , the periodic rearrangement of v_0^N in the sense of Definition 1.7. One can check that the process $\{\dot{X}_t^i\}_{i=1}^N$ is a reversible ergodic diffusion process with its unique invariant or Gibbs measure $M_N \in \mathcal{P}_{\text{sym}}((\mathbb{T}^d)^N)$ given by

$$M_N(x) = \frac{e^{-\beta H_1^N(x)}}{\int_{(\mathbb{T}^d)^N} e^{-\beta H_1^N(y)} dy}.$$

As expected, the law $\tilde{v}^N(t)$ of the quotiented system (1.12) can be obtained by considering the periodic rearrangement of $v^N(t)$, the solution of (1.10) and it evolves according to the following PDE:

$$\partial_s \tilde{v}^N = \beta^{-1} \Delta \tilde{v}^N + \nabla \cdot \left(\nabla H_1^N \tilde{v}^N \right), \quad (s, y) \in (0, \infty) \times (\mathbb{T}^d)^N. \quad (1.13)$$

In analogy to the discussion in Sect. 1.4, the above PDE is the gradient flow of the N -particle periodic free energy

$$\tilde{E}^N[\tilde{v}^N] := \frac{1}{N} \left(\beta^{-1} \int_{(\mathbb{T}^d)^N} \tilde{v}^N \log \tilde{v}^N dx + \int_{(\mathbb{T}^d)^N} H_1^N(x) d\tilde{v}^N(x) \right)$$

under the rescaled 2-Wasserstein distance $\frac{1}{\sqrt{N}} d_2$ on $\mathcal{P}_{\text{sym}}((\mathbb{T}^d)^N)$. Furthermore, the Gibbs measure M_N of the process (1.12) is the unique minimiser of \tilde{E}^N .

Similarly, we also notice that the nonlinear McKean–Vlasov equation (1.8) behaves well under the parabolic scaling. Specifically, given ρ^ε a solution to (1.8), then

$$v(s, y) = \varepsilon^d \rho^\varepsilon(\varepsilon^2 s, \varepsilon y)$$

is a solution to the McKean–Vlasov equation at scale $\varepsilon = 1$,

$$\partial_s v = \beta^{-1} \Delta v + \nabla \cdot (v \nabla W * v + \nabla V), \quad (s, y) \in (0, \infty) \times \mathbb{R}^d. \quad (1.14)$$

It is well known that this describes the law of the corresponding mean field McKean SDE which is given by

$$\begin{cases} dY_t^\varepsilon = -\nabla V(Y_t^\varepsilon) dt - \nabla(W * v(t))(Y_t^\varepsilon) dt + \sqrt{2\beta^{-1}} dB_t \\ \text{Law}(Y_0^\varepsilon) = v_0^\varepsilon = \varepsilon^d \rho_0^\varepsilon(\varepsilon x) \in \mathcal{P}(\mathbb{R}^d). \end{cases}$$

Again, we notice that all the coefficients in (1.14) are 1-periodic. Therefore, the nonlinearity $\nabla W * v$ only depends on the law of the quotiented process. We can thus understand the behaviour of the nonlinearity $\nabla W * v$ by considering the evolution

of the periodic rearrangement $\tilde{\nu}(t)$ of $\nu(t)$, which solves the periodic nonlinear McKean–Vlasov equation:

$$\partial_s \tilde{\nu} = \beta^{-1} \Delta \tilde{\nu} + \nabla \cdot (\tilde{\nu}(\nabla W * \tilde{\nu} + \nabla V)), \quad (s, y) \in (0, \infty) \times \mathbb{T}^d. \quad (1.15)$$

An important role is thus played by the limiting behaviour of solutions $\tilde{\nu}(t)$ of the above equation and its steady states. As in Sect. 1.4, the equation (1.15) is the gradient flow of the periodic mean field free energy

$$\begin{aligned} \tilde{E}_{MF}[\tilde{\nu}] &= \beta^{-1} \int_{\mathbb{T}^d} \tilde{\nu}(x) \log(\tilde{\nu}(x)) \, dx + \int_{\mathbb{T}^d} V(x) \, d\tilde{\nu}(x) \\ &\quad + \frac{1}{2} \int_{\mathbb{T}^d} W * \tilde{\nu}(x) \, d\tilde{\nu}(x) \end{aligned} \quad (1.16)$$

with respect to the the 2-Wasserstein metric on \mathbb{T}^d and the energies \tilde{E}^N and \tilde{E}^{MF} are related in the same way as the energies E^N and E_{MF} , i.e. through the result of Messer and Spohn [35].

Lemma 1.8. *The N -particle periodic free energy \tilde{E}^N Γ -converges (in the sense of Lemma 1.5) to $\tilde{E}^\infty : \mathcal{P}(\mathcal{P}(\mathbb{T}^d)) \rightarrow (-\infty, +\infty]$, where*

$$\tilde{E}^\infty[X] = \int_{\mathcal{P}(\mathbb{T}^d)} \tilde{E}_{MF}[\tilde{\nu}] \, dX(\tilde{\nu}).$$

As a consequence, if $\{M_N\}_{N \in \mathbb{N}}$ is the sequence of minimisers of \tilde{E}^N , then any accumulation point $X \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ of this sequence is a minimiser of \tilde{E}^∞ .

We can use the gradient flow structure to provide a useful characterisation of the steady states of the periodic McKean–Vlasov system (1.15).

Proposition 1.9. *Let $\tilde{\nu} \in \mathcal{P}(\mathbb{T}^d)$. Then, the following statements are equivalent:*

- (1) $\tilde{\nu}$ is a steady state of (1.15).
- (2) $\tilde{\nu}$ is a critical point of the mean field free energy, \tilde{E}_{MF} , i.e. the metric slope (cf. [1, Definition 1.2.4]) $|\partial \tilde{E}_{MF}|(\tilde{\nu}) = 0$.
- (3) $\tilde{\nu}$ is a zero of the dissipation functional $D : \mathcal{P}(\mathbb{T}^d) \rightarrow (-\infty, +\infty]$, i.e.

$$D(\tilde{\nu}) = \int_{\mathbb{T}^d} \left| \nabla \log \frac{\tilde{\nu}}{e^{-\beta(W * \nu + V)}} \right|^2 \tilde{\nu} \, dx = 0.$$

- (4) $\tilde{\nu}$ satisfies the self-consistency equation

$$\tilde{\nu} = \frac{e^{-\beta(V + W * \tilde{\nu})}}{Z}, \quad (1.17)$$

with the partition function given by

$$Z = \int_{\mathbb{T}^d} e^{-\beta(V + W * \tilde{\nu}(y))} \, dy.$$

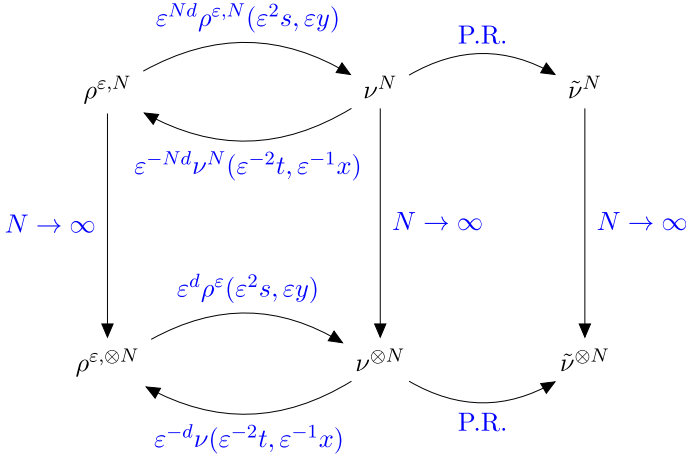


Fig. 2. A schematic of the notation. The P.R. denotes periodic rearrangement in the sense of Definition 1.7

A proof of this result can be found, for example, in [9, Proposition 2.4] or in [45]. It is evident from this characterisation that the behaviour of the system (1.15) on the torus will affect the distinguished limits (either $N \rightarrow \infty$ or $\varepsilon \rightarrow 0$) of the system (1.8) on \mathbb{R}^d . In particular, if (1.15) has multiple steady states then the distinguished limits will be influenced by steady states attained in the long-time dynamics. We refer to the phenomenon of nonuniqueness of steady states as a phase transition and discuss its effect on the limits in Sect. 1.9.

To conclude this subsection, for the reader's convenience, we include Figure 2 which provides a useful schematic of the notation that will be used for the rest of this paper. Starting with $\rho^{\varepsilon, N}$ the solution of (1.3), one can obtain $\tilde{\nu}^N$, the solution of (1.10), by using the scaling in (1.9). One can then pass to the limit $N \rightarrow \infty$ in $\rho^{\varepsilon, N}$ and ν^N , to obtain the McKean–Vlasov equation at scale ε (1.8) or scale 1 (1.14), respectively. Alternatively one can consider the periodic rearrangement $\tilde{\nu}^N$ of ν^N which solves (1.13) and pass to the limit $N \rightarrow \infty$ to obtain a solution of the periodic McKean–Vlasov equation (1.15). The rest of the figure follows in a similar fashion.

1.6. The Diffusive Limit

We have already discussed the limit $N \rightarrow \infty$ in Sect. 1.4. Here, we discuss the diffusive limit, i.e. $\varepsilon \rightarrow 0$. For a fixed number of particles N , we can use techniques from the theory of periodic homogenisation to pass to the limit $\varepsilon \rightarrow 0$ in (1.3), see for instance [38, Chapter 20] [4, 13, 30]. In particular, we have the following result:

Theorem C. (The diffusive limit) *Consider $\rho^{\varepsilon, N}$ the solution to (1.3) with initial data $\rho_0^{\varepsilon, N} \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$. Then, for all $t > 0$ the limit*

$$\rho^{N,*}(t) = \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon, N}(t)$$

exists. Furthermore, the curve of measures $\rho^{N,*} : [0, \infty) \rightarrow \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N)$ satisfies the heat equation

$$\partial_t \rho^{N,*} = \nabla \cdot (A^{\text{eff},N} \nabla \rho^{N,*}), \quad (1.18)$$

with initial data $\rho^{N,*}(0) = \lim_{\varepsilon \rightarrow 0} \rho_0^{\varepsilon,N}$ and where the covariance matrix is given by the formula

$$A^{\text{eff},N} = \beta^{-1} \int_{(\mathbb{T}^d)^N} (I + \nabla \Psi^N(y)) M_N(y) \, dy, \quad (1.19)$$

with

$$M_N(x) = \frac{e^{-\beta H_1^N(x)}}{\int_{(\mathbb{T}^d)^N} e^{-\beta H_1^N(y)} \, dy},$$

the Gibbs measure of the quotiented N -particle system (1.12) and $\Psi^N : (\mathbb{T}^d)^N \rightarrow (\mathbb{R}^d)^N$ the unique mean zero solution to the associated corrector problem

$$\nabla \cdot (M_N \nabla \Psi^N) = -\nabla M_N.$$

Here, H_1^N is the Hamiltonian of the associated particle system and is as defined in (1.4).

1.7. The Limit $N \rightarrow \infty$ Followed by $\varepsilon \rightarrow 0$

We have discussed the mean field limit $N \rightarrow \infty$ in Sect. 1.4. Now, we are ready to state our first result that characterises the limit $\lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho^{\varepsilon,N}$:

Theorem 1.10. *Consider the set of initial data given by $\{\rho_0^\varepsilon\}_{\varepsilon>0} \subset \mathcal{P}(\mathbb{R}^d)$, and consider the periodic rearrangement at scale $\varepsilon > 0$, i.e.*

$$\tilde{\nu}_0^\varepsilon(A) = \varepsilon^d \sum_{k \in \mathbb{Z}^d} \rho_0^\varepsilon(\varepsilon(A+k)) \quad \text{for } \varepsilon > 0.$$

Assume that there exists $C > 0$, $p > 1$ and a steady state $\tilde{\nu}^* \in \mathcal{P}(\mathbb{T}^d)$ such that $\tilde{\nu}^\varepsilon(t)$, the solution to the $\varepsilon = 1$ periodic nonlinear evolution (1.15) with initial data $\tilde{\nu}_0^\varepsilon(x)$, satisfies

$$\sup_{\varepsilon>0} d_2^2(\tilde{\nu}^\varepsilon(t), \tilde{\nu}^*) \leq Ct^{-p}. \quad (A1)$$

Then,

$$\lim_{\varepsilon \rightarrow 0} d_2^2(S_t^\varepsilon \rho_0^\varepsilon, S_t^* \rho_0^*) = 0, \quad (1.20)$$

where S_t^ε is the solution semigroup associated to (1.8), $\rho_0^* \in \mathcal{P}(\mathbb{R}^d)$ is the weak-* limit of ρ_0^ε , and S_t^* is the solution semigroup of the heat equation

$$\partial_t \rho = \nabla \cdot (A_*^{\text{eff}} \nabla \rho),$$

where the covariance matrix

$$A_*^{\text{eff}} = \beta^{-1} \int_{\mathbb{T}^d} (I + \nabla \Psi^*(y)) \, d\tilde{v}^*(y),$$

with $\Psi^* : \mathbb{T}^d \rightarrow \mathbb{R}^d$, $\Psi_i^* \in H^1(\mathbb{T}^d)$ for $i = 1, \dots, d$, is the unique mean zero solution to the associated corrector problem

$$\nabla \cdot (\tilde{v}^* \nabla \Psi^*) = -\nabla \tilde{v}^*. \quad (1.21)$$

Furthermore, assume that $X(t)^\varepsilon$ is as defined in (1.7) and that $\lim_{N \rightarrow \infty} \rho_0^{\varepsilon, N} = X_0^\varepsilon = \delta_{\rho_0^\varepsilon}$. Then it holds that

$$\lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho^{\varepsilon, N} = \lim_{\varepsilon \rightarrow 0} X(t)^\varepsilon = S_t^* \# X_0,$$

where $X_0 = \delta_{\rho_0^*}$.

In particular, we can apply this theorem to obtain the following result:

Corollary 1.11. *Assume that the periodic mean field energy (1.16) admits a unique minimiser (and hence critical point) \tilde{v}^{\min} and that it is an exponential attractor for arbitrary initial data of the evolution of (1.15), i.e. $d_2(\tilde{v}(t), \tilde{v}^{\min}) \leq d_2(\tilde{v}_0^\varepsilon, \tilde{v}_{\min}^\varepsilon) e^{-Ct}$ for some fixed constant $C > 0$. Then, the conclusions of Theorem 1.10 are valid for arbitrary initial data.*

Proof. The proof of this follows from the fact that $d_2(\tilde{v}(t), \tilde{v}^{\min}) \leq d_2(\tilde{v}_0^\varepsilon, \tilde{v}_{\min}^\varepsilon) e^{-Ct}$ implies that assumption (A1) holds. \square

Remark 1.12. (Non-chaotic initial data) Although Theorem 1.10 requires that the initial data be chaotic, we can deal with non-chaotic initial data by tweaking assumption (A1) to read as

$$\sup_{\varepsilon > 0} \sup_{\rho \in \text{supp } X_0^\varepsilon} d_2(\tilde{v}_\rho(t), \tilde{v}^*) \leq Ct^{-p}$$

for $p > 1$ and $C > 0$, with $\tilde{v}_\rho(t)$ being the solution of (1.15) starting with initial data \tilde{v}_0 which is the periodic rearrangement of ρ .

Remark 1.13. We cannot expect convergence $S_t^\varepsilon \rho_0^\varepsilon$ to $S_t^* \rho_0^*$ in a strong sense. By performing a formal multiscale expansion, we expect that

$$S_t^\varepsilon \rho_0^\varepsilon(x) = S_t^* \rho_0^*(x) \tilde{v}^*(x/\varepsilon) + \varepsilon S_t^*(\nabla \rho_0^*)(x) \cdot \Phi(x/\varepsilon) + \mathcal{O}(\varepsilon^2).$$

In particular, whenever \tilde{v}^* is not trivial, the leading term $S_t^* \rho_0^*(x) \tilde{v}^*(x/\varepsilon) \rightharpoonup S_t^* \rho_0^*$ converges only weakly to its limit.

Remark 1.14. The effective covariance matrix A_*^{eff} is strictly positive definite and we have the following bound on the ellipticity of the effective covariance matrix

$$\frac{\beta^{-1}}{Z_* Z_*^-} I \leq A_*^{\text{eff}} \leq \beta^{-1} I,$$

where

$$Z_* = \int_{\mathbb{T}^d} e^{-\beta(V+W*\tilde{v}^*(y))} \, dy \quad \text{and} \quad Z_*^- = \int_{\mathbb{T}^d} e^{\beta(V+W*\tilde{v}^*(y))} \, dy;$$

see [38, Theorem 13.12].

Remark 1.15. If we consider rapidly varying initial data, that is to say, if there exists $\rho_{in} \in \mathcal{P}(\mathbb{R}^d)$ such that

$$\rho_0^\varepsilon(x) = \varepsilon^{-d} \rho_{in}(\varepsilon^{-1}x).$$

Then, the hypothesis of Theorem 1.10 reduces to checking the speed of convergence to v^* of the solution to (1.15) with the periodic rearrangement of ρ_{in} as initial data, and $\rho_0^* = \delta_0$.

Here we can easily see how the phase transition matters for the limiting behaviour. If the evolution (1.15) admits more than one steady state \tilde{v}_1^* and \tilde{v}_2^* , then the diffusive limit will be different if we consider $\rho_{in} = \tilde{v}_1^*$ or $\rho_{in} = \tilde{v}_2^*$, see Corollary 1.27 for an explicit example.

1.8. The Limit $\varepsilon \rightarrow 0$ Followed by $N \rightarrow \infty$

Now that we have discussed the diffusive limit in Sect. 1.6, we characterise the limit $N \rightarrow \infty$ of $\rho^{N,*}(t)$.

Theorem 1.16. *Assume that the periodic mean field energy \tilde{E}_{MF} (1.16) admits a unique minimiser \tilde{v}^{\min} , then we have that $\rho^{N,*}$ the solution of (1.18) satisfies, for any fixed $t > 0$,*

$$\lim_{N \rightarrow \infty} \rho^{N,*}(t) = X(t) = S_t^{\min} \# X_0,$$

where $X_0 \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ is the limit of $\rho^{N,*}(0)$ in the sense of Definition (1.1), and $S_t^{\min} : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d)$ is the solution semigroup of the heat equation

$$\partial_t \rho = \nabla \cdot (A_{\min}^{\text{eff}} \nabla \rho),$$

where the covariance matrix

$$A_{\min}^{\text{eff}} = \beta^{-1} \int_{\mathbb{T}^d} (I + \nabla \Psi^{\min}(y)) \, d\tilde{v}^{\min}(y),$$

with $\Psi^{\min} : \mathbb{T}^d \rightarrow \mathbb{R}^d$, $\Psi_i^{\min} \in H^1(\mathbb{T}^d)$ for $i = 1, \dots, d$, the unique mean zero solution to the associated corrector problem

$$\nabla \cdot (\tilde{v}^{\min} \nabla \Psi^{\min}) = -\nabla \tilde{v}^{\min}. \quad (1.22)$$

It follows then, that for any fixed $t > 0$, the solution $\rho^{\varepsilon,N}(t)$ of (1.3) satisfies

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon,N}(t) = \lim_{N \rightarrow \infty} \rho^{N,*}(t) = S_t^{\min} \# X_0.$$

Remark 1.17. By Γ -convergence, the assumption that the periodic mean field energy \tilde{E}_{MF} defined in (1.16) admits a unique minimiser implies chaoticity of the Gibbs measure, that is to say $M_N \rightarrow \delta_{\tilde{v}^{\min}} \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$, see Lemma 1.8. We note that the assumption that \tilde{E}_{MF} admits a unique minimiser can be replaced by the weaker chaoticity assumption on M_N , i.e. $M_N \rightarrow \delta_{\tilde{v}_0^{\min}}$ for some specific minimiser \tilde{v}_0^{\min} .

1.9. The Effect of Phase Transitions

As mentioned in Sect. 1.5, we expect the presence of phase transition to affect the commutativity of the limits, especially since the results of Theorems 1.10 and 1.16 depend on the steady states of (1.15) and the minimisers of the periodic mean field energy \tilde{E}_{MF} . Before proceeding any further, we define what we mean by a phase transition.

Definition 1.18. (*Phase transition*) The periodic mean field system (1.15) is said to undergo a phase transition at some $0 < \beta_c < \infty$, if

- (1) For all $\beta < \beta_c$, there exists a unique steady state of (1.15).
- (2) For $\beta > \beta_c$, there exist at least two steady states of (1.15).

The temperature β_c is referred to as the point of phase transition or the critical temperature.

The above definition would not make sense without the following result:

Proposition 1.19. (Uniqueness at high temperature) *For all $0 < \beta < \infty$, the periodic mean field system (1.15) has at least one steady state, which is a minimiser of the periodic mean field energy \tilde{E}_{MF} . Furthermore, for β small enough, there exists a unique steady state \tilde{v}^{\min} of (1.15), which corresponds to the unique minimiser of \tilde{E}_{MF} .*

The proof of this result follows from standard fixed point and compactness arguments and can be found in [9, Theorem 2.3 and Proposition 2.8] or [35, Theorem 3].

Remark 1.20. The reader may have noticed that in Definition 1.18 we do not discuss what happens at $\beta = \beta_c$. This is due to the fact that this depends on the nature of the phase transition, i.e. whether it is continuous or discontinuous. A detailed discussion of these phenomena and the conditions under which they arise can be found in [9, 10].

In the absence of a confining potential, i.e. for $V = 0$, the existence and properties of phase transitions were studied in detail in [9, 10]. It turns out that a key role in understanding this phenomenon is played by the notion of H -stability. We refer to an interaction potential W as H -stable, denoted by $W \in \mathbf{H}_s$, if its Fourier coefficients are nonnegative, i.e.

$$\hat{W}(k) := \int_{\mathbb{T}^d} W(x) e^{i2\pi kx} dx \geq 0 \quad \forall k \in \mathbb{Z}^d, k \neq 0.$$

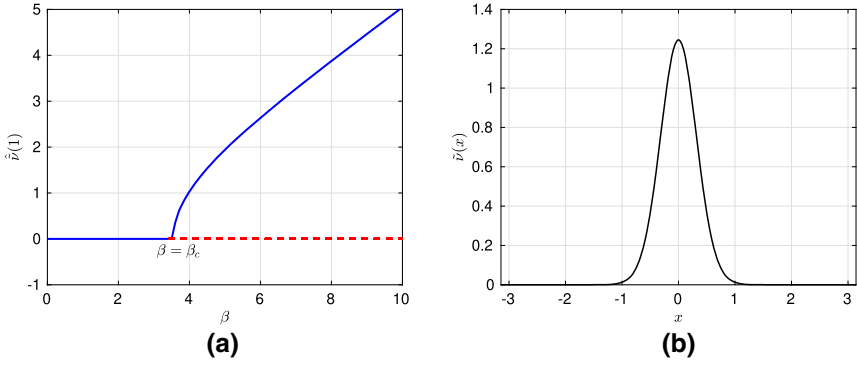


Fig. 3. **a** The bifurcation diagram for the noisy Kuramoto system: the solid blue line denotes the stable branch of solutions while the dotted red line denotes the unstable branch of solutions **b** An example of a clustered steady state \tilde{v}_β^{\min} representing phase synchronisation of the oscillators

This notion of H -stability is closely related to a similar concept used in the statistical mechanics of lattice spin systems (cf. [42]). Indeed, it provides us with a sharp criterion for the existence of a phase transition in the absence of the term V .

Proposition 1.21. (Existence of phase transitions, [9, 10]) *Assume $V = 0$. Then the periodic mean field system (1.15) undergoes a phase transition in the sense of Definition 1.18 if and only if $W \notin \mathbf{H}_s$.*

As discussed in the introduction, a prototypical example of a system that exhibits a phase transition is given by the potentials $V = 0$, $W = -\cos(2\pi x)$. The corresponding particle system is referred to the noisy Kuramoto model. The structure of phase transitions for this system is remarkably simple and is discussed in the following proposition:

Proposition 1.22. *Consider the quotiented periodic mean field system (1.15) with $d = 1$, $W = -\cos(2\pi x)$, and $V = 0$. Then for $\beta \leq 2$, $\tilde{v}_\infty \equiv 1$ is the unique minimiser and steady state of (1.15). For $\beta > 2$, the steady states of (1.15) are given by $\tilde{v}_\infty \equiv 1$ and the family of translates of the measure \tilde{v}_β^{\min} which is given by the expression*

$$\tilde{v}_\beta^{\min} = Z^{-1} e^{a \cos(2\pi x)}, \quad Z = \int_{\mathbb{T}} e^{a \cos(2\pi x)} dx,$$

with $a = a(\beta)$ the solution of the following nonlinear equation $a = \beta I_1(a)/I_0(a)$, where $I_1(a)$, $I_0(a)$ are the modified Bessel functions of the first and zeroth kind respectively. Moreover for $\beta > 2$, \tilde{v}_β^{\min} (and its translates) are the only minimisers of the periodic mean field energy \tilde{E}_{MF} . Thus, $\beta_c = 2$ is the critical temperature of (1.15).

A proof of the above result can be found in [9, Proposition 6.1]. A depiction of the bifurcation diagram of the noisy Kuramoto system can be found in Figure 3.

We can now start stating our results concerning the effect of the presence of a phase transition on the combined diffusive-mean field limit. In general, we have that for the large temperature regime the limits commute; that is, we have

Corollary 1.23. *Assume that $\rho_0^{\varepsilon, N} = (\rho_0^\varepsilon)^{\otimes N}$ for some $\rho_0^\varepsilon \in \mathcal{P}(\mathbb{R}^d)$ and that*

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho_0^{\varepsilon, N} = \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho_0^{\varepsilon, N} = X_0 = \delta_{\rho_0^*} \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d)),$$

where $\rho_0^* \in \mathcal{P}(\mathbb{R}^d)$ is the weak-* limit of ρ_0^ε . Then, there exists an explicit $\beta_0 \in (0, \beta_c]$ depending on $\|V\|_{C^2(\mathbb{T}^d)}$ and $\|W\|_{C^2(\mathbb{T}^d)}$ such that, for $\beta < \beta_0$, the limits commute:

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon, N}(t) = \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho^{\varepsilon, N}(t) = S_t^{\min} \# X_0.$$

Moreover, for rapidly varying initial data and $V = 0$, we can show that the limits commute all the way up to the phase transition. We have the following result:

Corollary 1.24. *Assume $V = 0$ and $\beta < \beta_c$, the critical temperature. Assume further that*

$$\rho_0^{\varepsilon, N} = \varepsilon^{-d} \rho_0(\varepsilon^{-1} x_1) \otimes \cdots \otimes \varepsilon^{-d} \rho_0(\varepsilon^{-1} x_N) \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^{\times N})$$

for some fixed $\rho_0 \in \mathcal{P}(\mathbb{R}^d)$. Then the limits commute, i.e.

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon, N}(t) = \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho^{\varepsilon, N}(t) = S_t^{\min} \# X_0,$$

where $X_0 = \delta_{\rho_0} \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$. If W is H -stable, this result holds for all $0 < \beta < \infty$ and arbitrary chaotic initial data.

The proof of Corollaries 1.23 and 1.24 can be found in Sect. 4.

Remark 1.25. The results of the preceding corollaries apply to the noisy Kuramoto model.

We are now ready to present our results above the critical temperature. As we are interested in illustrating our results in a clear way, we consider a simple system that undergoes a phase transition and show that the limits do not commute ahead of the phase transition. We do not consider the noisy Kuramoto model because, as demonstrated in Proposition 1.22, the minimisers of the \tilde{E}_{MF} are not unique ahead of the phase transition; the entire family of translates of \tilde{v}_β^{\min} are minimisers. Thus we cannot apply the results of Theorem 1.16 directly. Indeed, applying Lemma 1.8, one can show that the N -particle Gibbs measure M_N converges, in the sense of Definitions 1.1 and 1.2, to $X \in \mathcal{P}(\mathcal{P}(\mathbb{T}))$, where X is supported uniformly on the set of translates of \tilde{v}_β^{\min} , see [35, Proposition 3].

The alternative is to work in a quotient space as in [34, 39] or to add a small confinement to break the translation invariance of the problem. We choose to do the latter. However, we do expect our results to hold true even in the translation-invariant setting but we do not deal with what we feel is essentially a technical issue in this paper.

In particular, we consider in 1 space dimension the dynamics generated by the potentials $V = -\eta \cos(2\pi x)$ and $W = -\cos(2\pi x)$ with $0 < \eta < 1$. In this case, we have the following characterisation of phase transitions:

Lemma 1.26. *Consider the periodic mean field system (1.15) with $d = 1$, $W = -\cos(2\pi x)$, and $V = -\eta \cos(2\pi x)$ for a fixed $\eta \in (0, 1)$. Then there exists a value of the parameter $\beta = \beta_c$ such that*

- For $\beta < \beta_c$, there exists a unique steady state of the quotiented periodic system (1.15) given by

$$\tilde{v}^{\min}(x) = Z_{\min}^{-1} e^{a^{\min} \cos(2\pi x)}, \quad Z_{\min} = \int_{\mathbb{T}} e^{a^{\min} \cos(2\pi x)} dx, \quad (1.23)$$

for some $a^{\min} = a^{\min}(\beta)$, $a^{\min} > 0$, which is the unique minimiser of the periodic mean field energy \tilde{E}_{MF} (1.16).

- For $\beta > \beta_c$, there exist at least 2 steady states of the quotiented periodic system (1.15) given by

$$\begin{aligned} \tilde{v}^{\min}(x) &= Z_{\min}^{-1} e^{a^{\min} \cos(2\pi x)}, & Z_{\min} &= \int_{\mathbb{T}} e^{a^{\min} \cos(2\pi x)} dx, \\ \tilde{v}^*(x) &= Z_*^{-1} e^{a^* \cos(2\pi x)}, & Z_*^{-1} &= \int_{\mathbb{T}} e^{a^* \cos(2\pi x)} dx, \end{aligned}$$

where $a^* < 0 < a^{\min}$ and both constants depend on β . Here \tilde{v}^{\min} is the unique minimiser and \tilde{v}^* is a non-minimising critical points of the periodic mean field energy \tilde{E}_{MF} (1.16). Moreover, $a^* \neq -a^{\min}$.

The proof of Lemma 1.26 can be found in Sect. 4.

Now, we are ready to state our results in this specific case, i.e. above the phase transition we can choose specific initial data for which the limits do not commute.

Corollary 1.27. *Assume that $V = -\eta \cos(2\pi x)$, $W = \cos(2\pi x)$ for a fixed $\eta \in (0, 1)$, and that we are above the phase transition $\beta > \beta_c$. As in Proposition 1.26, we denote by \tilde{v}^{\min} and \tilde{v}^* the minimiser and the nonminimising critical point of \tilde{E}_{MF} . We choose the initial data*

$$\rho_0^{\varepsilon, N}(x) = (\varepsilon^{-1} \rho_0^*(\varepsilon^{-1} x_1)) \otimes \dots \otimes (\varepsilon^{-1} \rho_0^*(\varepsilon^{-1} x_N)) \in \mathcal{P}_{\text{sym}}((\mathbb{R}^d)^N),$$

where $\rho_0^* \in \mathcal{P}(\mathbb{R})$ satisfies

$$\tilde{v}^*(A) = \sum_{k \in \mathbb{Z}^d} \rho_0^*(A + k) \quad (1.24)$$

for any measurable A , i.e. its periodic rearrangement is \tilde{v}^* . Then, for every $t > 0$, $\rho^{\varepsilon, N}(t)$, the solution to (1.3) satisfies

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon, N}(t) = \delta_{\rho^{\min}(t)} \in \mathcal{P}(\mathcal{P}(\mathbb{R})),$$

where

$$\rho^{\min}(t) = \frac{\sqrt{\beta} I_0(a^{\min})}{\sqrt{2\pi t}} \exp\left(-\frac{\beta I_0(a^{\min})^2 |x|^2}{2t}\right).$$

On the other hand, we have that

$$\lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho^{\varepsilon, N} = \delta_{\rho^*(t)} \in \mathcal{P}(\mathcal{P}(\mathbb{R})),$$

where

$$\rho^*(t) = \frac{\sqrt{\beta} I_0(-a^*)}{\sqrt{2\pi t}} \exp\left(-\frac{\beta I_0(-a^*)^2 |x|^2}{2t}\right). \quad (1.25)$$

Finally, by Lemma 1.26 $a^* \neq -a^{\min}$, and therefore by the strict monotonicity of the modified zeroth Bessel function I_0 we obtain that

$$\rho^{\min}(t) \neq \rho^*(t) \quad \text{for any } t > 0.$$

Proof. We first note that

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho_0^{\varepsilon, N} = \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho_0^{\varepsilon, N} = \delta_{\delta_0} =: X_0 \in \mathcal{P}(\mathcal{P}(\mathbb{R})).$$

For the limit $\varepsilon \rightarrow 0$ followed by $N \rightarrow \infty$, we use that by Proposition 1.26 $\tilde{\nu}^{\min}$ is the unique minimiser of \tilde{E}_{MF} , hence we can apply Theorem 1.16 to obtain that

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon, N}(t) = \delta_{\rho^{\min}(t)}$$

with ρ^{\min} satisfying

$$\begin{cases} \partial_t \rho^{\min} = \partial_x (A_{\min}^{\text{eff}} \partial_x \rho^{\min}) \\ \rho^{\min}(0) = \delta_0, \end{cases}$$

where

$$A_{\min}^{\text{eff}} = \frac{\beta^{-1}}{Z\hat{Z}} = \frac{\beta^{-1}}{\int_{\mathbb{T}} e^{a^{\min} \cos(2\pi y)} dy \int_{\mathbb{T}} e^{-a^{\min} \cos(2\pi y)} dy} = \frac{\beta^{-1}}{I_0(a^{\min})^2}.$$

To obtain this formula, we have used that in 1-D we can solve the corrector problem (1.22) explicitly, see for instance [38, Equation (13.6.13)]. The explicit expression for $\rho^{\min}(t)$ now follows.

Now we turn to the other limit. As discussed in Sect. 1.4, passing to the limit $N \rightarrow \infty$, we obtain that for a fixed $t > 0$

$$\lim_{N \rightarrow \infty} \rho^{\varepsilon, N}(t) = X^\varepsilon(t) = S_t^\varepsilon \# X_0^\varepsilon,$$

with S_t^ε the solution semigroup of (1.8) and $X_0^\varepsilon = \delta_{\varepsilon^{-1} \rho_0^*(\varepsilon^{-1} x)}$. Using (1.24), we have that the initial data for the $\varepsilon = 1$ periodic mean field equation (1.15) is given by

$$\tilde{\nu}_0^\varepsilon = \sum_{k \in \mathbb{Z}^d} \rho_0^*(x+k) = \tilde{\nu}^*.$$

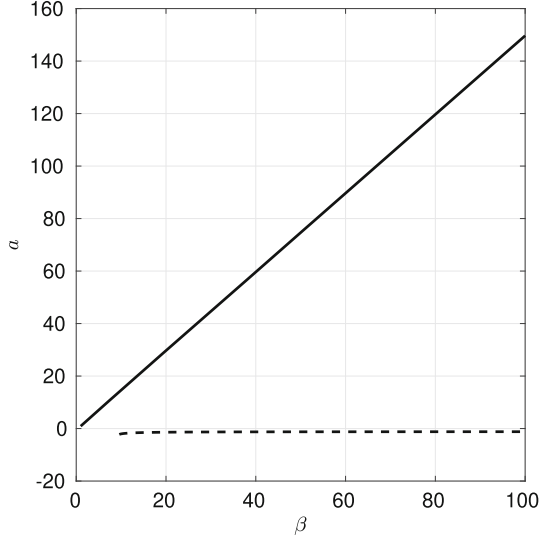


Fig. 4. The two solutions of Equation (4.2), i.e. a^{\min} (the solid line) and a^* (the dotted line) for $\eta = 0.5$. The figure demonstrates how stark the difference between the two solutions and hence between the two effective solutions, ρ^{\min} and ρ^* of Corollary 1.27, is

We know from Proposition 1.26 that \tilde{v}^* is steady state of (1.15), thus the hypothesis (A1) is trivially satisfied. Therefore, we can pass to the limit as $\varepsilon \rightarrow 0$ using Theorem 1.10 and obtain for a fixed $t > 0$

$$\lim_{\varepsilon \rightarrow 0} X^\varepsilon(t) = \delta_{\rho^*(t)},$$

with ρ^* satisfying

$$\begin{cases} \partial_t \rho^* = \partial_x (A_*^{\text{eff}} \partial_x \rho^*) \\ \rho^*(0) = \delta_0, \end{cases}$$

where

$$A_*^{\text{eff}} = \frac{\beta^{-1}}{Z\hat{Z}} = \frac{\beta^{-1}}{\int_{\mathbb{T}} e^{a^* \cos(2\pi y)} dy \int_{\mathbb{T}} e^{-a^* \cos(2\pi y)} dy} = \frac{\beta^{-1}}{I_0(-a^*)^2},$$

thus proving (1.25) and completing the proof of the result. □

Remark 1.28. The result of Corollary 1.27 can be generalized to other rapidly varying initial data that is exponentially attracted to \tilde{v}^* .

Remark 1.29. A simple choice of initial data which satisfies (1.24) is $\rho_0^* = \chi_{[0,1]} \tilde{v}^*$, with χ_A the indicator function of the set A .

1.10. Application of the Fluctuation Theorem

In this subsection we will assume without proof that we have a characterisation of the fluctuations around the mean field limit, in the spirit of [11, 19], as the solution to a linear SPDE and use this together the energy minimisation property of the Gibbs measure to obtain a rate of convergence in relative entropy of the Gibbs measure to the minimiser of the periodic mean field energy (1.15). We also characterise the asymptotic behaviour of the partition function. At the end of the subsection we present a provisional result in which we show that this rate of convergence does hold at high temperatures without using the central limit theorem (the characterisation of fluctuations) but instead conditional on a certain rate of convergence in a weaker topology (cf. (1.35)).

We start by restating the classical result by Messer–Spohn [35] (cf. Lemma 1.8). We consider the unique minimiser of $\tilde{E}^N : \mathcal{P}_{\text{sym}}((\mathbb{T}^d)^N) \rightarrow (-\infty, +\infty]$

$$\begin{aligned} \tilde{E}^N[\tilde{\nu}^N] &= \frac{1}{N} \left(\beta^{-1} \int_{(\mathbb{T}^d)^N} \tilde{\nu}^N(x) \log \tilde{\nu}^N(x) \, dx \right. \\ &\quad \left. + \int_{(\mathbb{T}^d)^N} \frac{1}{2N} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N W(x_i - x_j) + \sum_{i=1}^N V(x_i) \, d\tilde{\nu}^N(x) \right), \end{aligned}$$

which is given by the Gibbs measure

$$M_N(x) = \frac{e^{-\beta \left(\frac{1}{2N} \sum_{i=1}^N \sum_{j \neq i}^N W(x_i - x_j) + \sum_{i=1}^N V(x_i) \right)}}{Z_N},$$

with the partition function

$$Z_N = \int_{(\mathbb{T}^d)^N} e^{-\beta \left(\frac{1}{2N} \sum_{i=1}^N \sum_{j \neq i}^N W(y_i - y_j) + \sum_{i=1}^N V(y_i) \right)} \, dy.$$

Then any accumulation point $X^\infty \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ of the sequence of minimisers M_N is a minimiser of

$$\tilde{E}^\infty[X] = \int_{\mathcal{P}(\mathcal{P}(\mathbb{T}^d))} \tilde{E}_{MF}[\tilde{\nu}] \, dX(\tilde{\nu})$$

with $\tilde{E}_{MF} : \mathcal{P}(\mathbb{T}^d) \rightarrow (-\infty, +\infty]$ given by

$$\begin{aligned} \tilde{E}_{MF}[\tilde{\nu}] &= \beta^{-1} \int_{\mathbb{T}^d} \tilde{\nu} \log(\tilde{\nu}) \, dx + \frac{1}{2} \int_{\mathbb{T}^d} \int_{\mathbb{T}^d} W(x - y) \tilde{\nu}(x) \tilde{\nu}(y) \, dx \, dy \\ &\quad + \int_{\mathbb{T}^d} V(x) \tilde{\nu}(x) \, dx, \end{aligned} \tag{1.26}$$

which implies that

$$\text{supp } X^\infty \subset \{ \tilde{\nu} \in \mathcal{P}(\mathbb{T}^d) : \tilde{E}_{MF}[\tilde{\nu}] = \inf \tilde{E}_{MF} \}.$$

In particular, if we are below the phase transition $\beta < \beta_c$, we have, by Definition 1.18 and Proposition 1.9, that \tilde{E}_{MF} admits a unique minimiser, which we denote by $\tilde{v}^{\min} \in \mathcal{P}(\mathbb{T}^d)$ and thus $X^\infty = \delta_{\tilde{v}^{\min}} \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$. In the subsequent calculations, we will use $(M_N)_n$ to refer to the n^{th} marginal of the N -particle Gibbs measure M_N .

A natural next step is to consider the next order of convergence:

$$\begin{aligned}
& N(\tilde{E}^N(M_N) - \tilde{E}_{MF}(\tilde{v}^{\min})) \\
&= \beta^{-1} \int_{\mathbb{T}^{Nd}} M_N \log M_N \, dx - \beta^{-1} \int_{\mathbb{T}^d} (\tilde{v}^{\min})^{\otimes N} \log(\tilde{v}^{\min})^{\otimes N} \, dx \\
&\quad + \frac{N-1}{2} \int_{\mathbb{T}^{2d}} W(x-y)(M_N)_2 \, dx \, dy + N \int_{\mathbb{T}^d} V(x)(M_N)_1 \, dx \\
&\quad - \frac{N}{2} \int_{\mathbb{T}^{2d}} W(x-y)\tilde{v}^{\min}(x)\tilde{v}^{\min}(y) \, dx \, dy - N \int_{\mathbb{T}^d} V(x)\tilde{v}^{\min} \, dx. \quad (1.27)
\end{aligned}$$

The idea is to massage the previous expression to obtain something we can control with the fluctuations. To do this we first need to use the empirical measure $\hat{M}_N \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ associated to $M_N \in \mathcal{P}_{\text{sym}}((\mathbb{T}^d)^N)$, as defined in Definition 1.2. We can compare the second marginal $(M_N)_2$ of M_N with the products of the empirical measure. We notice that for any test function $\varphi \in C^\infty(\mathbb{T}^{2d})$, we have

$$\begin{aligned}
& \left(1 - \frac{1}{N}\right) \int_{\mathbb{T}^{2d}} \varphi(x, y)(M_N)_2 \, dx \, dy + \frac{1}{N} \int_{\mathbb{T}^d} \varphi(x, x)(M_N)_1 \, dx \\
&= \mathbb{E} \left\langle \varphi, \left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} \right)^{\otimes 2} \right\rangle, \quad (1.28)
\end{aligned}$$

where the expectation is taken with respect to the measure M_N on the probability space $((\mathbb{T}^d)^N, M_N)$ (for more details on these type identities for higher order marginals see [14]). We know from Proposition 1.9 that the minimiser of the mean field energy must satisfy the following condition:

$$\beta^{-1} \log \tilde{v}^{\min} = -W * \tilde{v}^{\min} - V + C. \quad (1.29)$$

Putting (1.27), (1.28), and (1.29) together, adding and subtracting

$$\beta^{-1} \int_{(\mathbb{T}^d)^N} M_N \log(\tilde{v}^{\min})^{\otimes N} \, dx,$$

and completing the square, we obtain

$$\begin{aligned}
N(\tilde{E}^N(M_N) - \tilde{E}_{MF}(\tilde{v}^{\min})) &= \beta^{-1} \mathcal{H}(M_N | (\tilde{v}^{\min})^{\otimes N}) \\
&\quad + \frac{1}{2} \mathbb{E} \left\langle W(x-y), (\mathcal{G}^N)^{\otimes 2} \right\rangle - \frac{W(0)}{2},
\end{aligned}$$

where $\mathcal{H}(\cdot | \cdot)$ denotes the relative entropy or Kullback–Leibler divergence and

$$\mathcal{G}^N := \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \tilde{v}^{\min} \right)$$

is a Radon measure-valued random variable defined on the probability space $((\mathbb{T}^d)^N, M_N)$. We refer to \mathcal{G}^N as the fluctuations around the mean field limit. Using the fact that

$$\begin{aligned} N\tilde{E}_{MF}(\tilde{\nu}^{\min}) &= N\tilde{E}^N((\tilde{\nu}^{\min})^{\otimes N}) + \frac{1}{2} \int_{\mathbb{T}^{2d}} W(x-y) \tilde{\nu}^{\min}(x) \tilde{\nu}^{\min}(y) \, dx \, dy \\ &\geq N\tilde{E}^N(M_N) + \frac{1}{2} \int_{\mathbb{T}^{2d}} W(x-y) \tilde{\nu}^{\min}(x) \tilde{\nu}^{\min}(y) \, dx \, dy, \end{aligned}$$

we obtain the bound

$$\begin{aligned} 0 \leq \mathcal{H}(M_N | (\tilde{\nu}^{\min})^{\otimes N}) &\leq \frac{W(0)}{2} - \frac{1}{2} \mathbb{E} \left\langle W(x-y), (\mathcal{G}^N)^{\otimes 2} \right\rangle \\ &\quad - \frac{1}{2} \int_{\mathbb{T}^{2d}} W(x-y) \tilde{\nu}^{\min}(x) \tilde{\nu}^{\min}(y) \, dx \, dy. \end{aligned} \quad (1.30)$$

In a similar way, we can also obtain the bound on the quotient between the N particle partition function and the limiting partition function Z_{\min} for ν_{\min} :

$$\begin{aligned} & - \frac{1}{2} \int_{\mathbb{T}^{2d}} W(x-y) \tilde{\nu}^{\min}(x) \tilde{\nu}^{\min}(y) \, dx \, dy \\ & \geq N \log \left(\frac{Z_{\min}}{Z_N} \right) = N \left(\tilde{E}^N[M_N] - \tilde{E}_{MF}[\tilde{\nu}^{\min}] \right) \\ & \geq - \frac{W(0)}{2} + \frac{1}{2} \mathbb{E} \left\langle W(x-y), (\mathcal{G}^N)^{\otimes 2} \right\rangle, \end{aligned} \quad (1.31)$$

where we have used the positivity of the relative entropy. Therefore, to obtain useful information from (1.30) and (1.31), we need to show that

$$\liminf_{N \rightarrow \infty} \mathbb{E} \left\langle W(x-y), (\mathcal{G}^N)^{\otimes 2} \right\rangle > -\infty.$$

To simplify the discussion and obtain sharp bounds all the way up to the phase transition, we consider the specific example of $d = 1$, $V = 0$ and $W = -\cos(2\pi x)$, which undergoes a phase transition at $\beta_c = 2$ (cf. Proposition 1.22). We now make our main assumption that we have an equilibrium version of the central limit theorem before the phase transition, i.e. \mathcal{G}^N converges in law to \mathcal{G}^∞ whose law is the unique invariant measure of the following linear stochastic PDE

$$d\mathcal{G}^\infty = \left(\beta^{-1} \partial_{xx} \mathcal{G}^\infty + (2\pi)^2 \cos(2\pi x) * \mathcal{G}^\infty \right) dt + \sqrt{2\beta^{-1}} d\xi, \quad (1.32)$$

where we have used that $\tilde{\nu}^{\min} = d\mathcal{L}$ and that $W = -\cos(2\pi x)$ has zero average to simplify the linearisation of the nonlinear PDE (1.8) and ξ is the space and time derivative of the cylindrical Wiener process. More specifically, if we consider $\{e_k\}_{k \in \mathbb{Z}}$ the standard orthonormal Fourier basis of $L^2(\mathbb{T})$ given by

$$e_k(x) = \begin{cases} \sqrt{2} \sin(2\pi kx) & k > 0 \\ 1 & k = 0 \\ \sqrt{2} \cos(2\pi kx) & k < 0, \end{cases}$$

then we can express

$$\xi(t, x) = \sum_{k \in \mathbb{Z}} 2\pi k e_k(x) \dot{B}_k(t),$$

where $\{\dot{B}_k\}_{k \in \mathbb{Z}}$ is a countable family of independent \mathbb{T} -valued Wiener processes. In particular, we can decompose (1.32) by projecting it onto each mode to obtain a family of uncoupled SDEs given by

$$d\langle \mathcal{G}^\infty, e_k \rangle = \begin{cases} (-\beta^{-1}(2\pi)^2 \langle \mathcal{G}^\infty, e_k \rangle + \frac{1}{2}(2\pi)^2 \langle \mathcal{G}^\infty, e_k \rangle) dt + (\sqrt{2\beta^{-1}} 2\pi k) d\dot{B}_k & |k| = 1 \\ (-\beta^{-1}(2\pi k)^2 \langle \mathcal{G}^\infty, e_k \rangle) dt + (\sqrt{2\beta^{-1}} 2\pi k) d\dot{B}_k & |k| \neq 1, \end{cases}$$

where we have used the trigonometric identity

$$\cos(2\pi x) * \mathcal{G}^\infty = \frac{1}{2} (e_1 \langle \mathcal{G}^\infty, e_1 \rangle + e_{-1} \langle \mathcal{G}^\infty, e_{-1} \rangle). \quad (1.33)$$

In particular, we can find the invariant measure explicitly for each mode

$$\text{Law}(\langle \mathcal{G}^\infty, e_k \rangle) = \begin{cases} \mathcal{N}\left(0, \frac{2}{2-\beta}\right) & |k| = 1 \\ \mathcal{N}(0, 1) & |k| \neq 1, \end{cases} \quad (1.34)$$

where \mathcal{N} is the normal distribution. From (1.34) we can clearly identify the phase transition $\beta_c = 2$ when the SPDE (1.32) does no longer support an invariant measure.

Taking limits in (1.30) and using that $W(0) = -1$ we obtain that for this specific system we have the bound

$$\limsup_{N \rightarrow \infty} \mathcal{H}(M_N | (\tilde{v}^{\min})^{\otimes N}) \leq \frac{W(0)}{2} + \mathbb{E} \frac{|e_1, \mathcal{G}^\infty|^2 + |e_{-1}, \mathcal{G}^\infty|^2}{4} = \frac{\beta}{2(2-\beta)},$$

where we have used the trigonometric identity (1.33) and the law (1.34) of the projections of \mathcal{G}^∞ . Decomposing the M_N into its marginals, we can use the subadditivity of the relative entropy to conclude that

$$\mathcal{H}((M_N)_n, (\tilde{v}^{\min})^n) \leq \left\lfloor \frac{n}{N} \right\rfloor \frac{\beta}{2(2-\beta)},$$

where $\lfloor n/N \rfloor$ is the largest integer less than n/N . We note that this estimate holds all the way up to the phase transition for this system $\beta_c = 2$. Similarly, using (1.31) we obtain that for every $\delta > 0$, we have the estimate

$$1 \leq \frac{Z_N}{Z_{\min}} \leq e^{\frac{1}{N} \left(\frac{\beta}{2(2-\beta)} + \delta \right)} \xrightarrow{N \rightarrow \infty} 1$$

for N large enough. To conclude this subsection, we rewrite these bounds into a general provisional theorem (cf. Remark 1.31).

Theorem 1.30. Consider \tilde{v}^{\min} , the unique minimiser of the periodic mean field energy \tilde{E}_{MF} (1.26) and Z_{\min} , its associated partition function. Assume that W is smooth and there exists $l \in \mathbb{R}$ such that

$$\sup_N \mathbb{E} \|\mathcal{G}^N\|_{H^{-l}(\mathbb{T}^d)}^2 < \infty. \quad (1.35)$$

Then, there exists $C > 0$ such that the following estimates hold:

$$\mathcal{H}\left((M_N)_n, (\tilde{v}^{\min})^{\otimes n}\right) \leq C \left\lfloor \frac{n}{N} \right\rfloor$$

and

$$1 \leq \frac{Z_N}{Z_{\min}} \leq e^{\frac{C}{N}}.$$

Remark 1.31. Note that if the formal central limit theorem discussed at the start of the subsection (in the spirit of [19]) could be proved rigorously then (1.35) would hold for some exponent l which depends on dimension. For the case $\beta = \beta_c$, we cannot expect (1.35) to hold (cf. [11]).

Proof. By (1.30), we need to show that there exists a C depending on β , V , and W such that

$$\liminf_{N \rightarrow \infty} \mathbb{E} \langle W(x - y), (\mathcal{G}^N)^{\otimes 2} \rangle \geq -C$$

for some $C > 0$. By using the smoothness of W , we obtain the following estimate:

$$|\mathbb{E} \langle W(x - y), (\mathcal{G}^N)^{\otimes 2} \rangle| \leq \mathbb{E} \|W * \mathcal{G}^N\|_{H^l(\mathbb{T}^d)} \|\mathcal{G}^N\|_{H^{-l}(\mathbb{T}^d)} \leq C_W \mathbb{E} \|\mathcal{G}^N\|_{H^{-l}(\mathbb{T}^d)}^2.$$

The result now follows by applying hypothesis (1.35). \square

2. Proof of Theorem 1.10

We start the proof of Theorem 1.10 with some basic elliptic estimates on a time-dependent corrector problem.

Lemma 2.1. Consider the elliptic equations

$$\nabla \cdot (\tilde{\mu}^\varepsilon \nabla \chi) = -\nabla \tilde{\mu}^\varepsilon, \quad (t, x) \in [0, \infty) \times \mathbb{T}^d, \quad (2.1)$$

where

$$\tilde{\mu}^\varepsilon(t, x) = Z^{-1} \exp(-\beta(W * \tilde{v}^\varepsilon + V)), \quad Z = \int_{\mathbb{T}^d} \exp(-\beta(W * \tilde{v}^\varepsilon + V)) dx,$$

where $\tilde{v}^\varepsilon(t, x)$ is a solution to the evolution (1.15) with initial data \tilde{v}_0^ε . Then, there exists a unique (up to an additive constant) smooth solution $\chi : [0, \infty) \times \mathbb{T}^d \rightarrow \mathbb{R}^d$, $\chi_i \in H^1(\mathbb{T}^d)$ to (2.1). Additionally, this satisfies the estimates

$$\|\chi_i\|_{H^m(\mathbb{T}^d)} \leq C_1 \quad (2.2)$$

$$\|\partial_t \chi_i(t)\|_{H^m(\mathbb{T}^d)} \leq \sum_{k=1}^m c_k \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}^k(t) \quad (2.3)$$

for all $i = 1, \dots, d$ and $t > 0$, where $C^{-3}(\mathbb{T}^d)$ is the dual of $C^3(\mathbb{T}^d)$, and the constants $C_1, c_k > 0$ depend only on $m, d, \|V\|_{C^m(\mathbb{T}^d)}$, and $\|W\|_{C^m(\mathbb{T}^d)}$.

Proof. Existence and uniqueness We consider the equation component-wise for any $i = 1, \dots, d$:

$$\nabla \cdot (\tilde{\mu}^\varepsilon \nabla \chi_i) = -\partial_{x_i} \tilde{\mu}^\varepsilon. \quad (2.4)$$

Note that $\tilde{\mu}^\varepsilon$ is smooth and is bounded above and below uniformly in time:

$$e^{-2\beta(\|W\|_{L^\infty(\mathbb{T}^d)} + \|V\|_{L^\infty(\mathbb{T}^d)})} \leq \mu^\varepsilon(t, x) \leq e^{2\beta(\|W\|_{L^\infty(\mathbb{T}^d)} + \|V\|_{L^\infty(\mathbb{T}^d))}. \quad (2.5)$$

Thus, by standard elliptic theory, for each $t \geq 0$ and $i = 1, \dots, d$, there exists a unique smooth solution $\chi_i \in H_0^1(\mathbb{T}^d)$ to (2.4). We can check that χ_i is continuously differentiable in time, as $\xi_i = \partial_t \chi_i$ satisfies

$$\nabla \cdot (\tilde{\mu}^\varepsilon \nabla \xi_i) = -\partial_{x_i} \partial_t \tilde{\mu}^\varepsilon - \nabla \cdot (\partial_t \tilde{\mu}^\varepsilon \nabla \chi_i). \quad (2.6)$$

Similar arguments imply that there exists a unique smooth solution of the above equation $\xi_i \in H_0^1(\mathbb{T}^d)$.

Regularity We note that it is sufficient to prove the bounds (2.2) and (2.3) in the weighted space $H^m(\mathbb{T}^d, \tilde{\mu}^\varepsilon)$ since by (2.5) these norms are equivalent to the flat space up to a time-independent multiplicative constant. We deal first with the regularity of (2.4). Testing against χ_i , we obtain

$$\begin{aligned} \int_{\mathbb{T}^d} |\nabla \chi_i|^2 \tilde{\mu}^\varepsilon dx &= - \int_{\mathbb{T}^d} \partial_{x_i} \chi_i \tilde{\mu}^\varepsilon dx \\ &\leq \|\partial_{x_i} \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \leq \|\nabla \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)}. \end{aligned}$$

It follows then that

$$\|\chi_i\|_{H_0^1(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \leq 1. \quad (2.7)$$

Now let $\alpha \in \mathbb{N}^d$ be any multi-index of order $m - 1$ for some $m \geq 1$. Testing (2.4) against $\partial_{2\alpha} \chi_i$, we obtain, for the left hand side,

$$\begin{aligned} \int_{\mathbb{T}^d} (\partial_{2\alpha} \chi_i) \nabla \cdot (\tilde{\mu}^\varepsilon \nabla \chi_i) \, dx &= (-1)^m \int_{\mathbb{T}^d} (\nabla \partial_\alpha \chi_i) \cdot \partial_\alpha (\tilde{\mu}^\varepsilon \nabla \chi_i) \, dx \\ &= (-1)^m \int_{\mathbb{T}^d} (\nabla \partial_\alpha \chi_i) \cdot \sum_{\gamma \leq \alpha} C_{\alpha, \gamma} (\partial_\gamma \tilde{\mu}^\varepsilon) (\partial_{\alpha-\gamma} \nabla \chi_i) \, dx \\ &= (-1)^m \int_{\mathbb{T}^d} |\nabla \partial_\alpha \chi_i|^2 \tilde{\mu}^\varepsilon \, dx \\ &\quad + (-1)^m \int_{\mathbb{T}^d} (\nabla \partial_\alpha \chi_i) \cdot \sum_{\substack{\gamma \leq \alpha \\ \gamma \neq 0}} C_{\alpha, \gamma} (\partial_\gamma \tilde{\mu}^\varepsilon) (\partial_{\alpha-\gamma} \nabla \chi_i) \, dx, \end{aligned}$$

where the coefficients are given by

$$C_{\alpha, \gamma} = \prod_{l=1}^d \binom{\alpha_l}{\gamma_l}.$$

Similarly, for the right hand side of (2.4), we obtain

$$- \int_{\mathbb{T}^d} (\partial_{2\alpha} \chi_i) (\partial_{x_i} \tilde{\mu}^\varepsilon) \, dx = (-1)^{m-1} \int_{\mathbb{T}^d} (\partial_{x_i} \partial_\alpha \chi_i) \partial_\alpha \tilde{\mu}^\varepsilon \, dx.$$

Putting the previous two equations together and multiplying by $(-1)^{-m}$, we have

$$\begin{aligned} \int_{\mathbb{T}^d} |\nabla \partial_\alpha \chi_i|^2 \tilde{\mu}^\varepsilon \, dx &= - \int_{\mathbb{T}^d} (\partial_{x_i} \partial_\alpha \chi_i) \partial_\alpha \tilde{\mu}^\varepsilon \, dx \\ &\quad - \int_{\mathbb{T}^d} (\nabla \partial_\alpha \chi_i) \cdot \sum_{\substack{\gamma \leq \alpha \\ \gamma \neq 0}} C_{\alpha, \gamma} (\partial_\gamma \tilde{\mu}^\varepsilon) (\partial_{\alpha-\gamma} \nabla \chi_i) \, dx. \end{aligned} \quad (2.8)$$

Using the exponential form of μ^ε , we note that for any multi-index $\alpha \in \mathbb{N}^d$ we have $\partial_\alpha \tilde{\mu}^\varepsilon = f^\alpha \tilde{\mu}^\varepsilon$, where f^α is a smooth function which is a linear combination of $\partial_\gamma W * \tilde{v}^\varepsilon + \partial_\gamma V$ for $\gamma \leq \alpha$. This implies that we can obtain the bound

$$\|f^\alpha\|_{L^\infty(\mathbb{T}^d)} \leq C_\alpha,$$

where C_α depends only on $\|W\|_{C^{m-1}(\mathbb{T}^d)}$, $\|V\|_{C^{m-1}(\mathbb{T}^d)}$. Applying Hölder's inequality and bounding in (2.8), we obtain

$$\begin{aligned} \int_{\mathbb{T}^d} |\nabla \partial_\alpha \chi_i|^2 \tilde{\mu}^\varepsilon \, dx &\leq \|\partial_{x_i} \partial_\alpha \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \|f^\alpha\|_{L^\infty(\mathbb{T}^d)} \\ &\quad + \|\nabla \partial_\alpha \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \sum_{\substack{\gamma \leq \alpha \\ \gamma \neq 0}} C_{\alpha, \gamma} \|f^\gamma\|_{L^\infty(\mathbb{T}^d)} \|\partial_{\alpha-\gamma} \nabla \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \end{aligned}$$

Simplifying, we obtain

$$\begin{aligned} \|\nabla \partial_\alpha \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} &\leq \|f^\alpha\|_{L^\infty(\mathbb{T}^d)} + \sum_{\substack{\gamma \leq \alpha \\ \gamma \neq 0}} C_{\alpha, \gamma} \|f^\gamma\|_{L^\infty(\mathbb{T}^d)} \|\partial_{\alpha-\gamma} \nabla \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \\ &\leq C_\alpha + \sum_{\substack{\gamma \leq \alpha \\ \gamma \neq 0}} C_{\alpha, \gamma} C_\gamma \|\chi_i\|_{H^{m-1}(\tilde{\mu}^\varepsilon)}. \end{aligned}$$

We can sum over all such α and recursively apply this bound along with (2.7) to obtain (2.2). Note that the $\|\chi_i\|_{L^2(\mathbb{T}^d)}$ norm can be controlled by the Poincaré inequality since χ_i is mean zero.

Before we turn to the regularity of (2.6), we derive the estimates

$$|\partial_t \tilde{\mu}^\varepsilon| = \beta(W * \partial_t \tilde{v}^\varepsilon) \tilde{\mu}^\varepsilon \leq \beta \|W\|_{C^3(\mathbb{T}^d)} \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)} \tilde{\mu}^\varepsilon, \quad (2.9)$$

where we denote by $C^{-3}(\mathbb{T}^d)$ the dual of $C^3(\mathbb{T}^d)$ and equip it with the norm $\|g\|_{C^{-3}(\mathbb{T}^d)} = \sup_{\|f\|_{C^3(\mathbb{T}^d)} \leq 1} \langle f, g \rangle$. Similarly, for $\alpha \in \mathbb{N}^d$, the following estimate holds:

$$|\partial_\alpha \partial_t \tilde{\mu}^\varepsilon| \leq \beta \|W\|_{C^{3+|\alpha|}} \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)} \tilde{\mu}^\varepsilon.$$

Next, we test (2.6) against ξ_i to obtain

$$\begin{aligned} \int_{\mathbb{T}^d} |\nabla \xi_i|^2 d\tilde{\mu}^\varepsilon dx &= - \int_{\mathbb{T}^d} \partial_{x_i} \xi_i \partial_t \tilde{\mu}^\varepsilon dx - \int_{\mathbb{T}^d} \nabla \xi_i \cdot \nabla \chi_i \partial_t \tilde{\mu}^\varepsilon dx \\ &\leq \beta \|W\|_{C^3(\mathbb{T}^d)} \|\nabla \xi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}(t) \\ &\quad + \beta \|W\|_{C^3(\mathbb{T}^d)} \|\nabla \xi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \|\nabla \chi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}(t), \end{aligned}$$

where we have simply used (2.9) and applied the Cauchy–Schwartz inequality. It follows that

$$\|\xi_i\|_{H^1(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \leq C \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}(t),$$

where the constant C is independent of t and depends on $\|\chi_i\|_{H^1(\mathbb{T}^d, \tilde{\mu}^\varepsilon)}$, V , and W . We omit the details but an essentially similar argument to the one used for (2.4) will give us an estimate of the form

$$\|\nabla \partial_\alpha \xi_i\|_{L^2(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \leq C' \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}(t) + \sum_{l=1}^m C_l \|\xi_i\|_{H^{m-l}(\mathbb{T}^d, \tilde{\mu}^\varepsilon)} \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}(t),$$

where $|\alpha| = m - 1$, and the constants C' , C_l are independent of t and depend on the norms of χ_i , W , V , and their derivatives. Recursively applying these bounds one obtains (2.3). \square

Next, we bound $\|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}$ by $d_2(\tilde{v}^\varepsilon, \tilde{v}^*)$.

Lemma 2.2. *Assuming that \tilde{v}^ε and \tilde{v}^* are a solution and a steady state of (1.15), respectively, then*

$$\|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)} \leq C d_2(\tilde{v}^\varepsilon, \tilde{v}^*),$$

where the constant C depends on dimension, β , $\|W\|_{C^2(\mathbb{T}^d)}$, and $\|V\|_{C^2(\mathbb{T}^d)}$.

Proof. Using (1.15) and that \tilde{v}^* is a steady state, we obtain that, for any test function φ ,

$$\int_{\mathbb{T}^d} \partial_t \tilde{v}^\varepsilon \varphi \, dx = \int_{\mathbb{T}^d} \beta^{-1} \Delta \varphi \tilde{v}^\varepsilon - \nabla \varphi \cdot (\nabla W * \tilde{v}^\varepsilon + \nabla V) \tilde{v}^\varepsilon \, dx$$

and

$$0 = \int_{\mathbb{T}^d} \partial_t \tilde{v}^* \varphi \, dx = \int_{\mathbb{T}^d} \beta^{-1} \Delta \varphi \tilde{v}^* - \nabla \varphi \cdot (\nabla W * \tilde{v}^* + \nabla V) \tilde{v}^* \, dx.$$

Therefore,

$$\begin{aligned} \int_{\mathbb{T}^d} \partial_t \tilde{v}^\varepsilon \varphi \, dx &= \int_{\mathbb{T}^d} \beta^{-1} \Delta \varphi (\tilde{v}^\varepsilon - \tilde{v}^*) - \nabla \varphi \cdot (\nabla W * \tilde{v} + \nabla V) (\tilde{v}^\varepsilon - \tilde{v}^*) \\ &\quad + \nabla \varphi \cdot \nabla W * (\tilde{v}^\varepsilon - \tilde{v}^*) \tilde{v}^* \, dx. \end{aligned}$$

Using the dual formulation of the 1-Wasserstein distance we can obtain the following bound:

$$\begin{aligned} \left| \int_{\mathbb{T}^d} \partial_t \tilde{v}^\varepsilon \varphi \, dx \right| &= \left(\beta^{-1} \|\varphi\|_{C^3(\mathbb{T}^d)} + \|\varphi\|_{C^2(\mathbb{T}^d)} (\|W\|_{C^2(\mathbb{T}^d)} + \|V\|_{C^2(\mathbb{T}^d)}) \right. \\ &\quad \left. + \|\varphi\|_{C^1(\mathbb{T}^d)} \|W\|_{C^2(\mathbb{T}^d)} \right) d_1(\tilde{v}^\varepsilon, \tilde{v}^*). \end{aligned}$$

Finally, bounding the 1-Wasserstein distance, by the 2-Wasserstein distance we obtain

$$d_1(\tilde{v}^\varepsilon, \tilde{v}^*) \leq d_2(\tilde{v}^\varepsilon, \tilde{v}^*).$$

Thus we have the desired estimate. \square

We now study the behaviour of the underlying SDE associated to (1.14).

Lemma 2.3. *Consider the mean field SDE*

$$\begin{cases} dY_t^\varepsilon = -\nabla V(Y_t^\varepsilon) dt - \nabla(W * v(t))(Y_t^\varepsilon) dt + \sqrt{2\beta^{-1}} dB_t \\ \text{Law}(Y_0^\varepsilon) = \nu_0^\varepsilon = \varepsilon^d \rho_0^\varepsilon(\varepsilon x) \in \mathcal{P}(\mathbb{R}^d), \end{cases} \quad (2.10)$$

where $v(t)$ is a solution of (1.14) with initial data ν_0^ε and B_t is a standard d -dimensional Wiener process. Then for fixed $\varepsilon > 0$, the random variables $\{t^{-1/2} Y_t^\varepsilon\}_{t>0}$ converge in law (specifically in d_2) as $t \rightarrow \infty$ to a mean zero Gaussian random variable Y with covariance matrix $2A_*^{\text{eff}} \in \mathbb{R}^{d \times d}$.

Proof. Consider $v^\varepsilon(t)$ the solution to the mean field PDE (1.14) with initial data given by v_0^ε . (we add the ε superscript to $v(t)$ to emphasise the dependence of the initial data on ε). As V and W are smooth 1-periodic functions, it follows that $(W * v(t))(x)$ is also 1-periodic and is equal to $(W * \tilde{v}^\varepsilon(t))(x)$, where \tilde{v}^ε is the periodic rearrangement of $v^\varepsilon(t)$. Thus the SDE in (2.10) can be rewritten as

$$\begin{cases} dY_t^\varepsilon = -\nabla V(\dot{Y}_t^\varepsilon) dt - \nabla(W * \tilde{v}^\varepsilon(t))(\dot{Y}_t^\varepsilon) dt + \sqrt{2\beta^{-1}} dB_t \\ \text{Law}(Y_0^\varepsilon) = v_0^\varepsilon = \varepsilon^d \rho_0^\varepsilon(\varepsilon x) \in \mathcal{P}(\mathbb{R}^d), \end{cases}$$

where \dot{Y}_t^ε is the quotient process, i.e. $(\dot{Y}_t^\varepsilon)_j = (Y_t^\varepsilon)_j \pmod{1}$ for all $j = 1, \dots, d$. Furthermore, \dot{Y}_t^ε satisfies the SDE

$$\begin{cases} d\dot{Y}_t^\varepsilon = -\nabla V(\dot{Y}_t^\varepsilon) dt - \nabla(W * \tilde{v}^\varepsilon(t))(\dot{Y}_t^\varepsilon) dt + \sqrt{2\beta^{-1}} d\dot{B}_t \\ \text{Law}(\dot{Y}_0^\varepsilon) = \tilde{v}_0^\varepsilon \in \mathcal{P}(\mathbb{T}^d), \end{cases} \quad (2.11)$$

where \dot{B}_t is a \mathbb{T}^d -valued Wiener process. Now consider the unique solution $\chi(t, \cdot) \in H^1(\mathbb{T}^d, \tilde{\mu}^\varepsilon)$ of the time-dependent corrector problem in (2.1) given by Lemma 2.1. Applying Itô's lemma to $\chi(t, Y_t^\varepsilon)$, we obtain

$$\begin{aligned} \chi(t, Y_t^\varepsilon) &= \chi(0, Y_0^\varepsilon) + \int_0^t \partial_s \chi(s, \dot{Y}_s^\varepsilon) ds \\ &\quad + \int_0^t \left(-(\nabla V(\dot{Y}_s^\varepsilon) + \nabla(W * \tilde{v}^\varepsilon(s))(\dot{Y}_s^\varepsilon)) \cdot \nabla + \beta^{-1} \Delta \right) \chi(s, \dot{Y}_s^\varepsilon) ds \\ &\quad + \sqrt{2\beta^{-1}} \int_0^t \nabla \chi(s, \dot{Y}_s^\varepsilon) dB_s \\ &= \chi(0, Y_0^\varepsilon) + \int_0^t \partial_s \chi(s, \dot{Y}_s^\varepsilon) ds + \int_0^t \beta^{-1} (\tilde{\mu}^\varepsilon)^{-1} (\nabla \cdot (\tilde{\mu}^\varepsilon \nabla \chi))(s, \dot{Y}_s^\varepsilon) ds \\ &\quad + \sqrt{2\beta^{-1}} \int_0^t \nabla \chi(s, \dot{Y}_s^\varepsilon) dB_s, \end{aligned}$$

where we have used the fact that $f(Y_t^\varepsilon) = f(\dot{Y}_t^\varepsilon)$ for any 1-periodic function f and the equation for $\tilde{\mu}^\varepsilon$. Using the fact that χ satisfies (2.1), the above expression simplifies to

$$\begin{aligned} \chi(t, Y_t^\varepsilon) &= \chi(0, Y_0^\varepsilon) + \int_0^t \partial_s \chi(s, \dot{Y}_s^\varepsilon) ds - \int_0^t \beta^{-1} (\tilde{\mu}^\varepsilon)^{-1} \nabla \tilde{\mu}^\varepsilon(s, \dot{Y}_s^\varepsilon) ds \\ &\quad + \sqrt{2\beta^{-1}} \int_0^t \nabla \chi(s, \dot{Y}_s^\varepsilon) dB_s \\ &= \chi(0, \dot{Y}_0^\varepsilon) + \int_0^t \partial_s \chi(s, \dot{Y}_s^\varepsilon) ds + \int_0^t \nabla V(\dot{Y}_s^\varepsilon) + \nabla(W * \tilde{v}^\varepsilon(s))(\dot{Y}_s^\varepsilon) ds \\ &\quad + \sqrt{2\beta^{-1}} \int_0^t \nabla \chi(s, \dot{Y}_s^\varepsilon) dB_s. \end{aligned}$$

Integrating (2.11) from 0 to t and adding the above expression, we obtain

$$Y_t = Y_0 + \chi(0, Y_0^\varepsilon) - \chi(t, Y_t^\varepsilon) + \int_0^t \partial_s \chi(s, \dot{Y}_s^\varepsilon) ds \\ + \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \chi(s, \dot{Y}_s^\varepsilon)) dB_s.$$

Multiplying by $t^{-1/2}$, we obtain

$$t^{-1/2} Y_t^\varepsilon = t^{-1/2} \left(\chi(0, Y_0^\varepsilon) - \chi(t, Y_t^\varepsilon) + \int_0^t \partial_s \chi(s, \dot{Y}_s^\varepsilon) ds \right) + t^{-1/2} Y_0^\varepsilon \\ + t^{-1/2} \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \chi(s, \dot{Y}_s^\varepsilon)) dB_s.$$

To analyse the limit of $t^{-1/2} Y_t^\varepsilon$ we start by showing that the first three terms on the RHS of the above expression go to zero in $L^\infty(\mathbb{P})$ as $t \rightarrow \infty$. Picking $m > d/2$ and applying the results of Lemma 2.1 along with Morrey's inequality, we have

$$t^{-1/2} \left(\chi(0, Y_0^\varepsilon) - \chi(t, Y_t^\varepsilon) + \int_0^t \partial_s \chi(s, \dot{Y}_s^\varepsilon) ds \right) \\ \leq t^{-1/2} C_d \left(2 \|\chi(t, \cdot)\|_{H^m(\mathbb{T}^d)} + \int_0^t \|\partial_s \chi(t, \cdot)\|_{H^m(\mathbb{T}^d)} ds \right) \\ \leq t^{-1/2} C_d \left(2C_1 + \int_0^t \sum_{k=1}^m c_k \|\partial_t \tilde{v}^\varepsilon\|_{C^{-3}(\mathbb{T}^d)}^k(s) ds \right) \xrightarrow{t \rightarrow \infty} 0, \quad (2.12)$$

where in the last step we have used Lemma 2.2 and applied assumption (A1). For the fourth term we simply use the fact that Y_0 has finite second moment to argue that it goes to zero in $L^2(\mathbb{P})$. Thus studying the behaviour of $t^{-1/2} Y_t$, in law, as $t \rightarrow \infty$ is equivalent to studying the asymptotic behaviour of the martingale term Z_t , where

$$Z_t := t^{-1/2} \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \chi(s, \dot{Y}_s^\varepsilon)) dB_s.$$

We will proceed in steps: In *Step 1*, we will argue that the χ in the above expression can be replaced by Ψ^* , where Ψ^* solves (1.21). In *Step 2*, we will compute the limiting covariance matrix of Z_t as $t \rightarrow \infty$ and show that it is precisely $2A_*^{\text{eff}}$. Finally, in *Step 3*, we will argue that the limiting random variable is a mean zero Gaussian.

Step 1. First note that $\mu^\varepsilon(t) \rightarrow \tilde{v}^*$ in L^∞ as $t \rightarrow \infty$. Indeed, we have that

$$\begin{aligned}
& \left| Z^{-1}(\tilde{\mu}^\varepsilon) \exp(-\beta(W * \tilde{v}^\varepsilon + V)) - Z^{-1}(\tilde{v}^*) \exp(-\beta(W * \tilde{v}^* + V)) \right| \\
& \leq Z^{-1}(\tilde{\mu}^\varepsilon) \left| \exp(-\beta(W * \tilde{v}^\varepsilon + V)) - \exp(-\beta(W * \tilde{v}^* + V)) \right| \\
& \quad + \left| Z^{-1}(\tilde{\mu}^\varepsilon) - Z^{-1}(\tilde{v}^*) \right| e^{\beta(\|W\|_{L^\infty(\mathbb{T}^d)} + \|V\|_{L^\infty(\mathbb{T}^d)})} \\
& \leq e^{2\beta(\|W\|_{L^\infty(\mathbb{T}^d)} + \|V\|_{L^\infty(\mathbb{T}^d)})} \beta |W * \tilde{v}^\varepsilon - W * \tilde{v}^*| \\
& \quad + e^{3\beta(\|W\|_{L^\infty(\mathbb{T}^d)} + \|V\|_{L^\infty(\mathbb{T}^d)})} \beta |W * \tilde{v}^\varepsilon - W * \tilde{v}^*| \\
& \leq Cd_2(\tilde{v}^\varepsilon, \tilde{v}^*) \xrightarrow{t \rightarrow \infty} 0, \tag{2.13}
\end{aligned}$$

where we have used (A1). We now argue that $\nabla\chi$ converges to $\nabla\Psi^*$ in $L^2(\mathbb{T}^d; \mathbb{R}^d)$. We perform the proof component-wise using the weak formulations of (2.1) and (1.21)

$$\begin{aligned}
& \int_{\mathbb{T}^d} \nabla\chi_i \cdot \nabla\phi \tilde{\mu}^\varepsilon(t) \, dx - \int_{\mathbb{T}^d} \nabla\Psi_i^* \cdot \nabla\phi \tilde{\mu}^\varepsilon(t) \, dx \\
& = \int_{\mathbb{T}^d} \partial_{x_i}(\tilde{\mu}^\varepsilon - \tilde{v}^*)\phi \, dx + \int_{\mathbb{T}^d} \nabla\Psi_i^* \cdot \nabla\phi \tilde{v}^* \, dx - \int_{\mathbb{T}^d} \nabla\Psi_i^* \cdot \nabla\phi \tilde{\mu}^\varepsilon(t) \, dx \\
& = - \int_{\mathbb{T}^d} (\tilde{\mu}^\varepsilon - \tilde{v}^*)\partial_{x_i}\phi \, dx + \int_{\mathbb{T}^d} \nabla\Psi_i^* \cdot \nabla\phi(\tilde{v}^* - \tilde{\mu}^\varepsilon(t)) \, dx \\
& \leq \|\tilde{\mu}^\varepsilon - \tilde{v}^*\|_{L^\infty(\mathbb{T}^d)}(1 + \|\nabla\Psi_i^*\|_{L^2(\mathbb{T}^d)})\|\nabla\phi\|_{L^2(\mathbb{T}^d)}.
\end{aligned}$$

Choosing $\phi = \chi_i - \Psi_i^*$ and using the uniform lower bound from (2.5), we obtain that

$$\|\nabla(\chi_i - \Psi_i^*)\|_{L^2(\mathbb{T}^d)} \leq C\|\tilde{\mu}^\varepsilon - \tilde{v}^*\|_{L^\infty(\mathbb{T}^d)}(1 + \|\nabla\Psi_i^*\|_{L^2(\mathbb{T}^d)}) \xrightarrow{t \rightarrow \infty} 0, \tag{2.14}$$

using (2.13). Thus we can now simply apply Ito's isometry as follows:

$$\begin{aligned}
& \mathbb{E} \left[t^{-1} \left| \sqrt{2\beta^{-1}} \int_0^t (I + \nabla\chi(s, \dot{Y}_s^\varepsilon)) dB_s - \sqrt{2\beta^{-1}} \int_0^t (I + \nabla\Psi^*(\dot{Y}_s^\varepsilon)) dB_s \right|^2 \right] \\
& = 2\beta^{-1} t^{-1} \mathbb{E} \left[\left| \int_0^t (\nabla(\chi - \Psi^*)) (s, \dot{Y}_s^\varepsilon) dB_s \right|^2 \right] \\
& = 2\beta^{-1} t^{-1} \mathbb{E} \left[\int_0^t \sum_{i=1}^d |\nabla(\chi_i - \Psi_i^*)|^2 (s, \dot{Y}_s^\varepsilon) \, ds \right] \\
& \leq 2\beta^{-1} t^{-1} \int_0^t \sum_{i=1}^d \|\nabla(\chi_i - \Psi_i^*)\|_{L^\infty(\mathbb{T}^d)}^2 (s) \, ds.
\end{aligned}$$

Picking some $m > d/2$ and applying Morrey's inequality, we obtain

$$\begin{aligned} & \mathbb{E} \left[t^{-1} \left| \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \chi(s, \dot{Y}_s^\varepsilon)) dB_s - \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \Psi^*(\dot{Y}_s^\varepsilon)) dB_s \right|^2 \right] \\ & \leq 2\beta^{-1} t^{-1} C_d^2 \int_0^t \sum_{i=1}^d \|\nabla(\chi_i - \Psi_i^*)\|_{H_0^m(\mathbb{T}^d)}^2(s) ds \\ & \leq 2\beta^{-1} t^{-1} C_d^2 C^2 \int_0^t \sum_{i=1}^d \|\nabla(\chi_i - \Psi_i^*)\|_{H_0^{m+1}(\mathbb{T}^d)}^{2\alpha}(s) \|\nabla(\chi_i - \Psi_i^*)\|_{L^2(\mathbb{T}^d)}^{2-2\alpha}(s) ds, \end{aligned}$$

where we have applied the Gagliardo–Nirenberg–Sobolev inequality and $\alpha = m/(m+1)$. We bound the H^{m+1} -norm in the above expression by a uniform constant using Lemma 2.1 and the fact that Ψ^* is the solution of a uniformly elliptic PDE with smooth coefficients. Hence, using (2.14), we obtain

$$\mathbb{E} \left[t^{-1} \left| \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \chi(s, \dot{Y}_s^\varepsilon)) dB_s - \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \Psi^*(\dot{Y}_s^\varepsilon)) dB_s \right|^2 \right] \xrightarrow{t \rightarrow \infty} 0. \quad (2.15)$$

Step 2. In this step, we compute the limiting covariance as $t \rightarrow \infty$ of the following term:

$$G_t^\varepsilon := t^{-1/2} \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \Psi^*(\dot{Y}_s^\varepsilon)) dB_s. \quad (2.16)$$

Applying Itô's isometry again, we have

$$\begin{aligned} & \mathbb{E}[(G_t^\varepsilon)_i (G_t^\varepsilon)_j] \\ & = 2\beta^{-1} t^{-1} \int_0^t \mathbb{E} \left[\left(\sum_{k=1}^d (\delta_{ik} + \partial_{x_k} \Psi_i^*) (\delta_{jk} + \partial_{x_k} \Psi_j^*) \right) (\dot{Y}_s^\varepsilon) \right] ds \\ & = 2\beta^{-1} t^{-1} \int_0^t \int_{\mathbb{T}^d} \left(\sum_{k=1}^d (\delta_{ik} + \partial_{x_i} \Psi_k^*) (\delta_{jk} + \partial_{x_j} \Psi_k^*) \right) (x) \tilde{v}^\varepsilon(s, x) dx ds \\ & = 2\beta^{-1} t^{-1} \int_0^t \int_{\mathbb{T}^d} \left(\sum_{k=1}^d (\delta_{ik} + \partial_{x_i} \Psi_k^*) (\delta_{jk} + \partial_{x_j} \Psi_k^*) \right) (x) (\tilde{v}^\varepsilon - \tilde{v}^*)(s, x) dx ds \\ & \quad + 2\beta^{-1} t^{-1} \int_0^t ds \left(\int_{\mathbb{T}^d} \left(\sum_{k=1}^d (\delta_{ik} + \partial_{x_i} \Psi_k^*) (\delta_{jk} + \partial_{x_j} \Psi_k^*) \right) (x) \tilde{v}^*(x) dx \right). \end{aligned} \quad (2.17)$$

We can bound the first term as

$$\begin{aligned} & 2\beta^{-1} t^{-1} \int_0^t \int_{\mathbb{T}^d} \left(\sum_{k=1}^d (\delta_{ik} + \partial_{x_i} \Psi_k^*) (\delta_{jk} + \partial_{x_j} \Psi_k^*) \right) (x) (\tilde{v}^\varepsilon - \tilde{v}^*)(s, x) dx ds \\ & \lesssim t^{-1} \int_0^t d_2(\tilde{v}^\varepsilon(s), \tilde{v}^*) ds \xrightarrow{t \rightarrow \infty} 0, \end{aligned} \quad (2.18)$$

where we have used (A1) and the fact $\nabla\Psi^*$ is Lipschitz. Then, from (2.17), it follows that

$$\begin{aligned}
\lim_{t \rightarrow \infty} \mathbb{E}[(G_t^\varepsilon)_i (G_t^\varepsilon)_j] &= 2\beta^{-1} \left(\int_{\mathbb{T}^d} \left(\sum_{k=1}^d (\delta_{ik} + \partial_{x_i} \Psi_k^*) (\delta_{jk} + \partial_{x_j} \Psi_k^*) \right) (x) \tilde{\nu}^*(x) dx \right) \\
&= 2\beta^{-1} \left(\int_{\mathbb{T}^d} \left(\delta_{ij} + \partial_{x_i} \Psi_j^* + \partial_{x_j} \Psi_i^* + \sum_{k=1}^d \partial_k \Psi_i^* \partial_k \Psi_j^* \right) (x) \tilde{\nu}^*(x) dx \right) \\
&= 2\beta^{-1} \left(\int_{\mathbb{T}^d} \left(\delta_{ij} + \partial_{x_i} \Psi_j^* + \partial_{x_j} \Psi_i^* + \nabla \Psi_i^* \cdot \nabla \Psi_j^* \right) (x) \tilde{\nu}^*(x) dx \right) \\
&= 2\beta^{-1} \left(\int_{\mathbb{T}^d} \left(\delta_{ij} + \partial_{x_i} \Psi_j^* \right) (x) \tilde{\nu}^*(x) dx \right) = 2 \left(A_*^{\text{eff}} \right)_{ij},
\end{aligned}$$

where in the penultimate step we have used the fact that Ψ^* satisfies (1.21).

Step 3. In the final step, we will show that the limit in law of G_t^ε as $t \rightarrow \infty$ is a Gaussian random variable. The key step involves replacing \dot{Y}_t^ε in the expression for G_t^ε in (2.16) by \dot{X}_t , where \dot{X}_t solves

$$\begin{cases} d\dot{X}_t = -\nabla V(\dot{X}_t) dt - \nabla(W * \tilde{\nu}^*)(\dot{X}_t) dt + \sqrt{2\beta^{-1}} d\dot{B}_t \\ \text{Law}(\dot{X}_0) = \tilde{\nu}^* \in \mathcal{P}(\mathbb{T}^d). \end{cases} \quad (2.19)$$

Here \dot{X}_t is a stationary, ergodic process with invariant measure $\tilde{\nu}^*$ and is precisely the process \dot{Y}_t started from the invariant measure $\tilde{\nu}^*$. We assert now (cf. Lemma A.1) that (A1) implies that there exists a coupling of (\dot{X}_t, \dot{Y}_t) (indeed a reflection coupling) such that $\sup_{\varepsilon > 0} \mathbb{E}[d_{\mathbb{T}^d}(\dot{X}_t, \dot{Y}_t)^2] \rightarrow 0$ as $t \rightarrow \infty$. Using this, we obtain,

$$\begin{aligned}
&\mathbb{E} \left[\left(G_t^\varepsilon - t^{-1/2} \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \Psi^*(\dot{X}_s)) dB_s \right)^2 \right] \\
&\leq 2C\beta^{-1} t^{-1} \int_0^t \mathbb{E} [d_{\mathbb{T}^d}(\dot{X}_s, \dot{Y}_s)^2] ds \xrightarrow{t \rightarrow \infty} 0, \quad (2.20)
\end{aligned}$$

where we simply use the Itô's isometry and the fact that Ψ^* has a Lipschitz regular gradient. Thus we are left to analyse the following term:

$$t^{-1/2} \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \Psi^*(\dot{X}_s)) dB_s.$$

Here \dot{X}_s is stationary ergodic process. Additionally, we know the limiting covariance of the above term, i.e. $2A_*^{\text{eff}}$. We now apply the Birkhoff ergodic theorem followed by the martingale central limit theorem (cf. [31, Theorem 2.1] or [46, Theorem 2.1]) to complete the proof. The fact that the convergence is also in d_2 follows from the fact that the covariance matrices also converge. \square

The above result holds for a fixed $\varepsilon > 0$, however we can improve it by using the fact that the convergence in (A1) is uniform in $\varepsilon > 0$. A consequence of the analysis in the previous result is the following corollary:

Corollary 2.4. *Consider the process*

$$M_t^\varepsilon := t^{-\frac{1}{2}} \left(\chi(0, Y_0^\varepsilon) - \chi(t, Y_t^\varepsilon) + \int_0^t \partial_s \chi(s, Y_s^\varepsilon) ds + \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \chi(s, \dot{Y}_s^\varepsilon)) dB_s \right),$$

where χ is the solution of the time-dependent corrector problem (2.1) and Y_t^ε solves (2.10). Then

$$\lim_{t \rightarrow \infty} \sup_{\varepsilon > 0} \left\| M_t^\varepsilon - t^{-1/2} \sqrt{2\beta^{-1}} \int_0^t (I + \nabla \Psi^*(\dot{X}_s)) dB_s \right\|_{L^2(\mathbb{P})} \rightarrow 0,$$

where \dot{X}_t is the solves and is coupled to \dot{Y}_t^ε as in the proof of Lemma 2.3.

Proof. The proof of this result follows from the fact that the convergence in (2.12), (2.13), (2.15), (2.18), and (2.20), are all controlled by (A1) which is uniform in $\varepsilon > 0$. \square

We can finally put all the pieces together and complete the proof of Theorem 1.10

Proof of Theorem 1.10. We would like to understand the behaviour of the trajectory $S_t^\varepsilon \rho_0^\varepsilon$ where S_t^ε is the solution semigroup associated to (1.8). However, we know that $S_t^\varepsilon \rho_0^\varepsilon = \text{Law}(\varepsilon Y_{t/\varepsilon^2}^\varepsilon)$, where Y_t^ε is the solution (2.10) of with initial law $\varepsilon^d \rho_0^\varepsilon(\varepsilon x)$. Fix $t = 1$ and set $\varepsilon^{-2} = s$. Thus we have that

$$\begin{aligned} \varepsilon Y_{1/\varepsilon^2}^\varepsilon &= s^{-\frac{1}{2}} Y_s^{s^{-\frac{1}{2}}} \\ &= s^{-\frac{1}{2}} Y_0^{s^{-\frac{1}{2}}} + s^{-\frac{1}{2}} \left(\chi(0, Y_0) - \chi(s, Y_s^{s^{-1/2}}) + \int_0^s \partial_u \chi(s, \dot{Y}_u^{s^{-1/2}}) du \right. \\ &\quad \left. + \sqrt{2\beta^{-1}} \int_0^s (I + \nabla \chi(u, \dot{Y}_u^{s^{-1/2}})) dB_u \right). \end{aligned}$$

Applying Corollary 2.4, we can pass to the limit $s \rightarrow \infty$ for the second term on the right hand side of the above expression. Since the convergence is in $L^2(\mathbb{P})$, we can replace the second term in the limit as $s \rightarrow \infty$ as follows:

$$\begin{aligned} &\lim_{s \rightarrow \infty} \text{Law}(s^{-\frac{1}{2}} Y_s^{s^{-\frac{1}{2}}}) \\ &= \lim_{s \rightarrow \infty} \text{Law} \left(s^{-\frac{1}{2}} Y_0^{s^{-\frac{1}{2}}} + s^{-1/2} \sqrt{2\beta^{-1}} \int_0^s (I + \nabla \Psi^*(\dot{X}_u)) dB_u \right). \end{aligned}$$

Here \dot{X}_u solves (2.19). The two random variables on the right hand side of the above expression are independent. Thus we can rewrite the above expression as

$$\lim_{s \rightarrow \infty} \text{Law}(s^{-\frac{1}{2}} Y_s^{s^{-\frac{1}{2}}}) = \lim_{s \rightarrow \infty} F_s * \rho_0^{s^{-1/2}},$$

where F_s is the law of $s^{-1/2} \sqrt{2\beta^{-1}} \int_0^s (I + \nabla \Psi^*(\dot{X}_u)) dB_u$. Since both laws converge in d_2 , their convolution converges to the convolution of the individual limits in d_2 as $s \rightarrow \infty$. The limit of F_s can be obtained by the martingale central limit

theorem as in the proof of Lemma 2.3 while the limit of $\rho_0^{s^{-1/2}}$ is ρ_0^* . Thus we have that

$$\begin{aligned} & \lim_{s \rightarrow \infty} d_2(\text{Law}(s^{-\frac{1}{2}} Y_s^{s^{-\frac{1}{2}}}), \mathcal{N}(0, A_*^{\text{eff}}) * \rho_0^*) \\ &= \lim_{s \rightarrow \infty} d_2(S_1^{s^{-\frac{1}{2}}} \rho_0^{s^{-\frac{1}{2}}}, \mathcal{N}(0, A_*^{\text{eff}}) * \rho_0^*) = 0. \end{aligned}$$

Rewriting the same in terms of the laws, we have

$$\lim_{\varepsilon \rightarrow 0} d_2(S_1^\varepsilon \rho_0^\varepsilon, \mathcal{N}(0, A_*^{\text{eff}}) * \rho_0^*) = 0.$$

The choice of time $t = 1$ was arbitrary. The same arguments can be repeated for arbitrary $t \geq 0$ to complete the proof of (1.20). Assume now that the initial data of (1.3), $\rho_0^{\varepsilon, N}$ is such that $\lim_{N \rightarrow \infty} \rho_0^{\varepsilon, N} = X_0^\varepsilon = \delta_{\rho_0^\varepsilon}$. We can then apply Theorem B to first assert that, for a fixed $t > 0$,

$$\lim_{N \rightarrow \infty} \rho^{\varepsilon, N}(t) = X^\varepsilon(t) = S_t^\varepsilon \# X_0^\varepsilon.$$

Let $X_0 = \delta_{\rho_0^*}$ and consider some $\Phi \in \text{Lip}(\mathcal{P}(\mathbb{R}^d))$. Then we have

$$\begin{aligned} & \int_{\mathcal{P}(\mathbb{R}^d)} \Phi(\rho) d(X(t)^\varepsilon - S_t^* \# X_0)(\rho) \\ &= \int_{\mathcal{P}(\mathbb{R}^d)} \Phi(\rho) d(S_t^\varepsilon \# X_0^\varepsilon - S_t^* \# X_0)(\rho) \\ &= \Phi(S_t^\varepsilon \rho_0^\varepsilon) - \Phi(S_t^* \rho_0^*) \leq d_2(S_t^\varepsilon \rho_0^\varepsilon, S_t^* \rho_0^*) \xrightarrow{\varepsilon \rightarrow 0} 0, \end{aligned}$$

where in the last step we have simply applied (1.20), thus completing the proof of the theorem. \square

3. Proof of Theorem 1.16

Strategy of proof. We first need to pass to the limit in the covariance matrix

$$A^{\text{eff}, N} = \beta^{-1} \int_{(\mathbb{T}^d)^N} (I + \nabla \Psi^N) M_N \, dx. \quad (3.1)$$

To do this, we first pass to the limit in the Poisson equation for $\Psi^N : (\mathbb{T}^d)^N \rightarrow (\mathbb{R}^d)^N$

$$\nabla \cdot (M_N \nabla \Psi^N) = -\nabla M_N, \quad (3.2)$$

with

$$\int_{(\mathbb{T}^d)^N} \Psi^N M_N \, dx = 0. \quad (3.3)$$

Once we have obtained the limit of the covariance matrix, we then need to pass to the limit in the equation

$$\partial_t \rho^{N,*} = \nabla \cdot (A^{\text{eff},N} \nabla \rho^{N,*}).$$

We do this by testing against cylindrical test functions, that is to say functions that depend on a finite number of variables, which is enough to determine the limit in $\mathcal{P}(\mathcal{P}(\mathbb{R}^d))$.

Step 1. We start by showing a few a priori estimates. First, we show that for every $1 \leq i \leq Nd$ and $k \in \{1, \dots, N\}$ such that i does not belong to the particle k (i.e. $i \notin [(k-1)d+1, kd]$) the solution of the corrector problem (3.2) satisfies

$$\int_{(\mathbb{T}^d)^N} |\nabla \Psi_i^N|^2 M_N \, dx \leq 1 \quad \text{and} \quad \int_{(\mathbb{T}^d)^N} |\nabla_{x_k} \Psi_i^N|^2 M_N \, dx \leq \frac{1}{N-1}, \quad (3.4)$$

where

$$\nabla \phi := \sum_{j=1}^{Nd} (\partial_j \phi) e_j \in \mathbb{R}^{Nd} \quad \text{and} \quad \nabla_{x_k} \phi := \sum_{j=(k-1)d+1}^{kd} (\partial_j \phi) e_j \in \mathbb{R}^d$$

for any smooth function $\phi : (\mathbb{T}^d)^N \rightarrow \mathbb{R}$. Testing the i -th equation of (3.2) against Ψ_i^N , applying the Cauchy-Schwarz inequality, and using the fact that M_N has mass one, we obtain

$$\begin{aligned} \int_{(\mathbb{T}^d)^N} M_N |\partial_i \Psi_i^N|^2 \, dx &\leq \int_{(\mathbb{T}^d)^N} |\nabla \Psi_i^N|^2 M_N \, dx \\ &= \int_{(\mathbb{T}^d)^N} M_N \partial_i \Psi_i^N \, dx \\ &\leq \left(\int_{(\mathbb{T}^d)^N} M_N |\partial_i \Psi_i^N|^2 \, dx \right)^{1/2}, \end{aligned}$$

which implies the first inequality in (3.4). The second inequality follows due to the exchangeability of the particles. In fact, for any $k_1, k_2 \in \{1, \dots, N\}$ and $i \notin [(k_1-1)d+1, k_1d] \cup [(k_2-1)d+1, k_2d]$ we have that up to exchanging the k_1 and k_2 particles (i.e. changing variables)

$$\nabla_{x_{k_1}} \Psi_i^N = \nabla_{x_{k_2}} \Psi_i^N.$$

Combining this with the first inequality of (3.4) we obtain the second inequality of (3.4).

Next, we show that there exists $C(\beta, \|W\|_{C^1(\mathbb{T}^d)}, \|V\|_{C^1(\mathbb{T}^d)})$ such that

$$\|(M_N)_1\|_{C^1(\mathbb{T}^d)} \leq C \quad \text{and} \quad C^{-1} \leq M_N / M_{N-1} \leq C, \quad (3.5)$$

where M_{N-1} is the Gibbs measure associated to the quotiented $(N - 1)$ -particle system trivially extended to $(\mathbb{T}^d)^N$ and $(M_N)_1$ is the first marginal of M_N . We start by rewriting

$$\begin{aligned} M_N &= \frac{e^{-\beta\left(\frac{1}{2N}\sum_{i=1}^N\sum_{j\neq i}W(x_i-x_j)+\sum_{i=1}^NV(x_i)\right)}}{Z_N} \\ &= e^{-\beta\left(\frac{1}{N}\sum_{j=1}^NW(x_1-x_j)-\frac{1}{2N(N-1)}\sum_{i,j}^NW(x_i-x_j)+V(x_1)\right)}M_{N-1}\frac{Z_{N-1}}{Z_N}. \end{aligned} \quad (3.6)$$

Differentiating the previous expression with respect to x_1 , we obtain

$$\begin{aligned} \nabla_{x_1}M_N &= -\beta\left(\left(\frac{1}{N}-\frac{1}{N(N-1)}\right)\sum_{j=1}^N\nabla W(x_1-x_j)+\nabla V(x_1)\right) \\ &\quad e^{-\beta\left(\frac{1}{N}\sum_{j=1}^NW(x_1-x_j)-\frac{1}{2N(N-1)}\sum_{i,j}^NW(x_i-x_j)+V(x_1)\right)}M_{N-1}\frac{Z_{N-1}}{Z_N}. \end{aligned} \quad (3.7)$$

By (3.6) and (3.7) and the fact that V and W are sufficiently regular, the desired estimates follows if we can show that Z_{N-1}/Z_N is bounded above and below. This follows from the estimate

$$\begin{aligned} Z_N &= \int_{(\mathbb{T}^d)^N} e^{-\beta\left(\frac{1}{2N}\sum_{i=1}^N\sum_{j\neq i}W(y_i-y_j)+\sum_{i=1}^NV(y_i)\right)} dy \\ &\geq e^{-\beta\left(\left(\frac{N}{2(N-1)}+1\right)\|W\|_\infty+\|V\|_\infty\right)} \int_{(\mathbb{T}^d)^N} e^{-\beta\left(\frac{1}{2(N-1)}\sum_{i=2}^N\sum_{j\neq i}W(y_i-y_j)+\sum_{i=2}^NV(y_i)\right)} dy \\ &\geq C^{-1}Z_{N-1} \end{aligned}$$

and its analogue for the reverse bound.

Step 2. Next we show that we can in a suitable sense decompose M_N by the product $(M_N)_1M_{N-1}$. To be precise, we show that, for every $x_1 \in \mathbb{T}^d$,

$$\lim_{N \rightarrow \infty} \int_{(\mathbb{T}^d)^{N-1}} \left(\frac{M_N}{M_{N-1}} - (M_N)_1 \right)^2 M_{N-1} dx_2 \dots dx_N = 0. \quad (3.8)$$

We notice that

$$(M_N)_1 = \int_{(\mathbb{T}^d)^{N-1}} M_N dx_2 \dots dx_N = \int_{(\mathbb{T}^d)^{N-1}} \frac{M_N}{M_{N-1}} M_{N-1} dx_2 \dots dx_N,$$

and we rewrite

$$\begin{aligned} &\int_{(\mathbb{T}^d)^{N-1}} \left(\frac{M_N}{M_{N-1}} - (M_N)_1 \right)^2 M_{N-1} \\ &= \left| \frac{Z_{N-1}}{Z_N} \right|^2 \int_{(\mathbb{T}^d)^{N-1}} \left(u^N - \int_{(\mathbb{T}^d)^{N-1}} u^N M_{N-1} \right)^2 M_{N-1} \\ &= \left| \frac{Z_{N-1}}{Z_N} \right|^2 \left(\int_{(\mathbb{T}^d)^{N-1}} (u^N)^2 M_{N-1} - \left(\int_{(\mathbb{T}^d)^{N-1}} u^N M_{N-1} \right)^2 \right), \end{aligned}$$

where the function $u^N : \mathbb{T}^d \times (\mathbb{T}^d)^{N-1} \rightarrow \mathbb{R}$ is given by

$$u^N = \frac{Z_N}{Z_{N-1}} \frac{M_N}{M_{N-1}} = e^{-\beta \left(\frac{1}{N} \sum_{j=2}^N W(x_1 - x_j) - \frac{1}{2N(N-1)} \sum_{i,j=2}^N W(x_i - x_j) + V(x_1) \right)}.$$

Therefore, by (3.5), we can show (3.8) by showing that

$$\begin{aligned} & \lim_{N \rightarrow \infty} \int_{(\mathbb{T}^d)^{N-1}} (u^N)^2 M_{N-1} \, dx_2 \dots dx_N \\ & - \left(\int_{(\mathbb{T}^d)^{N-1}} u^N M_{N-1} \, dx_2 \dots dx_N \right)^2 = 0. \end{aligned} \quad (3.9)$$

This will follow from the chaoticity assumption on M_N and a version of the Arzela-Ascoli theorem for the limit of symmetric functions, where we employ an idea that was proposed by Lions [33] in the context of mean field games (cf. [7, 23]).

We show that the sequence of functions $\{u^N\}_{N \in \mathbb{N}}$ induces a compact sequence $\{U^N\}_{N \in \mathbb{N}} \subset C(\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d))$ and that (3.9) can be written in terms of the limit of U^N . We start by noticing that u_N is continuous and symmetric in the variables x_2 through x_N such that there exists $C(\beta, W, V)$ such that

$$\|u^N\|_{L^\infty((\mathbb{T}^d)^N)} \leq C, \quad |\nabla_{x_1} u^N| \leq C \quad \text{and} \quad |\nabla_{x_j} u^N| \leq \frac{C}{N} \quad x_j \neq x_1.$$

Using the symmetry of u^N and the previous bound, we can estimate

$$\begin{aligned} & |u^N(x_1, x_2, \dots, x_N) - u^N(y_1, \dots, y_N)| \\ & = |u^N(x_1, x_{\sigma(2)}, \dots, x_{\sigma(N)}) - u^N(y_1, \dots, y_N)| \\ & \leq C d_{\mathbb{T}^d}(x_1, y_1) + \frac{C}{N} \sum_{i=2}^N d_{\mathbb{T}^d}(x_{\sigma(i)}, y_i), \end{aligned}$$

with σ an arbitrary permutation of the indices $\{2, 3, \dots, N\}$. Taking the infimum over σ , we obtain that

$$\begin{aligned} & |u^N(x_1, x_2, \dots, x_N) - u^N(y_1, \dots, y_N)| \\ & \leq C \left(|x_1 - y_1| + d_1 \left(\frac{1}{N-1} \sum_{i=2}^N \delta_{x_i}, \frac{1}{N-1} \sum_{i=2}^N \delta_{y_i} \right) \right), \end{aligned} \quad (3.10)$$

where d_1 denotes the 1-Wasserstein distance on $\mathcal{P}(\mathbb{T}^d)$. For any $(x_1, \mu) \in \mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)$, we define

$$\begin{aligned} U^N(x_1, \mu) := & \inf_{(z_2, \dots, z_N) \in (\mathbb{T}^d)^{N-1}} 2C d_1 \left(\mu, \frac{1}{N-1} \sum_{i=2}^N \delta_{z_i} \right) \\ & + u^N(x_1, z_2, \dots, z_N) \in C^0(\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)). \end{aligned}$$

It follows directly from (3.10) that

$$U^N \left(x_1, \frac{1}{N-1} \sum_{i=2}^N \delta_{x_i} \right) = u^N(x_1, x_2, \dots, x_N). \quad (3.11)$$

Using (3.11), we can rewrite (3.9) as

$$\lim_{N \rightarrow \infty} \int_{\mathcal{P}(\mathbb{T}^d)} (U^N)^2(x_1, \rho) d\hat{M}_{N-1}(\rho) - \left(\int_{\mathcal{P}(\mathbb{T}^d)} U^N(x_1, \rho) d\hat{M}_{N-1}(\rho) \right)^2 = 0, \quad (3.12)$$

where $\hat{M}_{N-1} \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ is the empirical measure associated with M_{N-1} as defined in Definition 1.2.

Next, we show that U^N is Lipschitz with respect to the 1-Wasserstein distance, i.e.

$$|U^N(x_1, \mu) - U^N(y_1, \nu)| \leq 2C(d_{\mathbb{T}^d}(x_1, y_1) + d_1(\mu, \nu)). \quad (3.13)$$

By the definition of $U^N(y_1, \nu)$, for every $\delta > 0$ there exists (z_2, \dots, z_N) such that

$$U^N(y_1, \nu) + \delta \geq 2Cd_1 \left(\nu, \frac{1}{N-1} \sum_{i=2}^N \delta_{z_i} \right) + u^N(y_1, z_2, \dots, z_N).$$

By the definition of $U^N(x_1, \mu)$, the Lipschitz property of u^N (3.10), and the triangle inequality for d_1 , we have

$$\begin{aligned} U^N(x_1, \mu) &\leq 2Cd_1 \left(\mu, \frac{1}{N-1} \sum_{i=2}^N \delta_{z_i} \right) + u^N(x_1, z_2, \dots, z_N) \\ &\leq Cd_{\mathbb{T}^d}(x_1, y_1) + 2Cd_1 \left(\mu, \frac{1}{N-1} \sum_{i=2}^N \delta_{z_i} \right) + u^N(y_1, z_2, \dots, z_N) \\ &\leq Cd_{\mathbb{T}^d}(x_1, y_1) + 2C \left(d_1 \left(\mu, \frac{1}{N-1} \sum_{i=2}^N \delta_{z_i} \right) \right. \\ &\quad \left. - d_1 \left(\nu, \frac{1}{N-1} \sum_{i=2}^N \delta_{z_i} \right) \right) + U^N(y_1, \nu) + \delta \\ &\leq Cd_{\mathbb{T}^d}(x_1, y_1) + 2Cd_1(\mu, \nu) + U^N(y_1, \nu) + \delta. \end{aligned}$$

Using the fact that $\delta > 0$, (x_1, μ) and (y_1, ν) are arbitrary, (3.13) follows. Due to the compactness of \mathbb{T}^d , the space $\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d)$ equipped with the metric $d_{\mathbb{T}^d} + d_1$ is also compact. Therefore, by the Arzela–Ascoli theorem and the uniform Lipschitz bound in (3.13), we have that, up to subsequence, there exists $U \in C^0(\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d))$ such that

$$\lim_{N \rightarrow \infty} \|U^N - U\|_{C^0(\mathbb{T}^d \times \mathcal{P}(\mathbb{T}^d))} = 0.$$

Finally, we use the assumption that $\hat{M}_{N-1} \rightarrow \delta_{\tilde{v}^{\min}} \in \mathcal{P}(\mathcal{P}(\mathbb{T}^d))$ (see Remark 1.17), to obtain that, up to subsequence,

$$\begin{aligned} & \lim_{N \rightarrow \infty} \int_{\mathcal{P}(\mathbb{T}^d)} (U^N)^2(x_1, \rho) d\hat{M}_{N-1}(\rho) - \left(\int_{\mathcal{P}(\mathbb{T}^d)} U^N(x_1, \rho) d\hat{M}_{N-1}(\rho) \right)^2 \\ &= \int_{\mathcal{P}(\mathbb{T}^d)} U^2(x_1, \rho) d\delta_{\tilde{v}^{\min}}(\rho) - \left(\int_{\mathcal{P}(\mathbb{T}^d)} U(x_1, \rho) d\delta_{\tilde{v}^{\min}}(\rho) \right)^2 \\ &= U^2(x_1, \tilde{v}^{\min}) - U^2(x_1, \tilde{v}^{\min}) \\ &= 0. \end{aligned}$$

As the limit is independent of the subsequence we have chosen, we obtain (3.12), which implies (3.8).

Step 3. Now we are ready to pass to the limit in the Poisson equation (3.2). As the dimension where the problem is posed grows, we consider test functions that depend on a finite number of variables. We take some $\varphi \in [C^1(\mathbb{T}^d)]^d$ and consider its trivial extension by 1 to $[C^1((\mathbb{T}^d)^N)]^{Nd}$ to test the first d equations in (3.2):

$$\int_{(\mathbb{T}^d)^N} M_N(x) \nabla \Psi^N(x) : \nabla_{x_1} \varphi \, dx = \int_{(\mathbb{T}^d)^N} M_N(x) \nabla_{x_1} \cdot \varphi(x_1) \, dx. \quad (3.14)$$

Here $\nabla \Psi^N(x) : \nabla_{x_1} \varphi$ denotes the inner product between matrices and we notice that $\nabla_{x_1} \varphi$ has non-trivial entries only for $1 \leq i, j \leq d$. Integrating the variables x_2 to x_N in the right hand side of (3.14) we obtain

$$\int_{(\mathbb{T}^d)^N} M_N(x) \nabla_{x_1} \cdot \varphi(x_1) \, dx = \int_{\mathbb{T}^d} (M_N)_1(x_1) \nabla_{x_1} \cdot \varphi(x_1) \, dx_1, \quad (3.15)$$

where $(M_N)_1$ is the first marginal of M_N .

For the left hand side of (3.14), we notice that by *Step 1.* (3.4) and *Step 2.* (3.8) we can exchange M_N in the integrand with the product $M_{N-1}(M_N)_1$

$$\begin{aligned} & \left| \int_{(\mathbb{T}^d)^N} (M_N - (M_N)_1 M_{N-1}) \nabla_{x_1} \Psi^N : \nabla_{x_1} \varphi \, dx \right| \\ & \leq \|\varphi\|_{C^1(\mathbb{T}^d)} \sum_{i=1}^d \int_{(\mathbb{T}^d)^N} \left(\frac{M_N}{M_{N-1}} - (M_N)_1 \right) |\nabla_{x_1} \Psi_i^N| M_{N-1} \, dx \\ & \leq \|\varphi\|_{C^1(\mathbb{T}^d)} \sum_{i=1}^d \|\nabla_{x_1} \Psi_i^N\|_{L^2(M_{N-1})} \\ & \quad \times \int_{\mathbb{T}^d} \left(\int_{(\mathbb{T}^d)^{N-1}} \left(\frac{M_N}{M_{N-1}} - (M_N)_1 \right)^2 M_{N-1} \, dx_2 \dots dx_N \right) dx_1, \\ & = o(1) \xrightarrow{N \rightarrow \infty} 0, \end{aligned} \quad (3.16)$$

where in the last equality we have used (3.8) and that

$$\sup_{x_1 \in \mathbb{T}^d} \left| \int_{(\mathbb{T}^d)^{N-1}} \left(\frac{M_N}{M_{N-1}} - (M_N)_1 \right)^2 M_{N-1} \, dx_2 \dots dx_N \right| < C,$$

independently of N to be able to apply Lebesgue dominated convergence to pass to the limit in the outer integral.

Hence, putting together (3.14), (3.15) and (3.16) we obtain

$$\begin{aligned} & \int_{\mathbb{T}^d} (M_N)_1 \nabla_{x_1} \left(\int_{(\mathbb{T}^d)^{N-1}} \Psi^N M_{N-1} \, dx_2 \dots dx_N \right) : \nabla_{x_1} \varphi \, dx_1 \\ &= - \int_{\mathbb{T}^d} (M_N)_1 \nabla_{x_1} \cdot \varphi \, dx_1 + o(1). \end{aligned} \quad (3.17)$$

To pass to the limit in (3.17), we use of the a priori estimates which we proved in *Step 1.*, (3.4) and (3.5), which say that there exists $C > 0$ such that for every $N \in \mathbb{N}$ and $i \leq d$ we have

$$\left\| \int_{(\mathbb{T}^d)^{N-1}} \Psi_i^N M_{N-1} \, dx_2 \dots dx_N \right\|_{H^1(\mathbb{T}^d)} \leq C \quad \text{and} \quad \|(M_N)_1\|_{C^1(\mathbb{T}^d)} \leq C.$$

Note that we have used the Poincaré inequality on \mathbb{T}^d to extend the gradient bound from (3.4) to an $H^1(\mathbb{T}^d)$ bound uniform in N . Moreover, by the chaoticity assumption $M_N \rightarrow \delta_{\tilde{\nu}^{\min}}$ (see Remark 1.17) we can conclude that as $N \rightarrow \infty$ we have $(M_N)_1 \rightarrow \tilde{\nu}^{\min}$ in $C^0(\mathbb{T}^d)$. Then, passing to the limit in (3.17) we can see that any weak- H^1 accumulation point $\Psi^{\min} \in [H^1(\mathbb{T}^d)]^d$ of the function

$$\left(\int_{(\mathbb{T}^d)^{N-1}} \Psi_1^N M_{N-1} \, dx_2 \dots dx_N, \dots, \int_{(\mathbb{T}^d)^{N-1}} \Psi_d^N M_{N-1} \, dx_2 \dots dx_N \right),$$

satisfies the equation

$$\int_{\mathbb{T}^d} \tilde{\nu}^{\min} \nabla_{x_1} \Psi^{\min} : \nabla_{x_1} \varphi \, dx_1 = - \int_{\mathbb{T}^d} \tilde{\nu}^{\min} \nabla_{x_1} \cdot \varphi \, dx_1$$

with the condition

$$\int_{\mathbb{T}^d} \Psi^{\min} \tilde{\nu}^{\min} \, dx_1 = 0,$$

which follows from passing to the limit in (3.3) in the same fashion as above. This uniquely determines the limit Ψ^{\min} . Therefore, up to exchanging the coordinates, we can pass to the limit in the diagonal of (3.1). That is to say for every set of indices i, j satisfying $(k-1)d \leq i, j \leq kd$, we have

$$A_{i,j}^{\text{eff},N} \rightarrow \int_{\mathbb{T}^d} \beta^{-1} \left(\delta_{i,j} + \partial_j \Psi_i^{\min} \right) d\tilde{\nu}^{\min}, \quad (3.18)$$

where \tilde{i} and \tilde{j} are respectively i and j modulo d . We also notice that using the a priori estimate (3.4), we have that for every pair of indices i and j satisfying $(k_1 - 1)d \leq i \leq k_1d$ and $(k_2 - 1)d \leq j \leq k_2d$ with $k_1 \neq k_2$,

$$|A_{i,j}^{\text{eff},N}| \leq \frac{1}{N-1} \xrightarrow{N \rightarrow \infty} 0. \quad (3.19)$$

Step 4. Finally, we show that we can pass to the limit in the equation

$$\partial_t \rho^{N,*} = \nabla \cdot (A^{\text{eff},N} \nabla \rho^{N,*}) \quad \text{on } (0, \infty) \times (\mathbb{R}^d)^N. \quad (3.20)$$

We consider a test function $\varphi \in C^2((\mathbb{R}^d)^n)$ and extend it trivially to $C^2((\mathbb{R}^d)^N)$. Testing (3.20) against φ we obtain that, for every $t > 0$,

$$\begin{aligned} & \int_{(\mathbb{R}^d)^N} \varphi \rho^N(t) \, dx - \int_{(\mathbb{R}^d)^N} \varphi \rho^N(0) \, dx \\ &= \int_0^t \int_{(\mathbb{R}^d)^N} \nabla \cdot (A^{\text{eff},N} \nabla \varphi) \rho^{N,*}(s) \, dx \, ds. \end{aligned}$$

Next, we use that φ only depends on the first nd variables to obtain

$$\begin{aligned} & \int_{(\mathbb{R}^d)^n} \varphi \rho_n^N(t) \, dz - \int_{(\mathbb{R}^d)^n} \varphi \rho_n^N(0) \, dz \\ &= \int_0^t \int_{(\mathbb{R}^d)^n} \nabla \cdot ([A^{\text{eff},N}]_{1 \leq i, j \leq nd} \nabla \varphi) \rho_n^N(s) \, dz \, ds, \end{aligned} \quad (3.21)$$

where $\rho_n^N \in \mathcal{P}((\mathbb{R}^d)^n)$ is the n -th marginal of $\rho^{N,*}$ and $[A^{\text{eff},N}]_{1 \leq i, j \leq nd} \in \mathbb{R}^{nd \times nd}$ is the first $nd \times nd$ coordinates of $A^{\text{eff},N}$. Therefore, ρ_n^N is a weak solution of

$$\partial_t \rho_n^{N,*} = \nabla \cdot ([A^{\text{eff},N}]_{1 \leq i, j \leq nd} \nabla \rho_n^N) \quad \text{on } (0, \infty) \times (\mathbb{R}^d)^n.$$

By *Step 3.*, we note that

$$[A^{\text{eff},N}]_{1 \leq i, j \leq nd} \rightarrow [A_{\min}^{\infty, \text{eff}}]_{1 \leq i, j \leq nd}$$

where $A_{\min}^{\infty, \text{eff}} \in \mathbb{R}^{\infty \times \infty}$ denotes the matrix which if considered in $d \times d$ blocks is diagonal, which has the constant matrix

$$A_{\min}^{\text{eff}} = \beta^{-1} \int_{\mathbb{T}^d} (I + \nabla \Psi^{\min}) \, d\tilde{v}^{\min} \in \mathbb{R}^{d \times d},$$

which is non-degenerate elliptic by Remark 1.14. Therefore, for N large enough $[A^{\text{eff},N}]_{1 \leq i, j \leq nd}$ is uniformly elliptic and we can use standard parabolic techniques to obtain compactness of the curve $\rho_n^{N,*}$ in $C([0, T]; \mathcal{P}((\mathbb{R}^d)^n))$. By *Step 3.*, we can use (3.18) and (3.19) to pass to the limit (3.21) and obtain

$$\begin{aligned} & \int_{(\mathbb{R}^d)^n} \varphi \rho_n^\infty(t) \, dz - \int_{(\mathbb{R}^d)^n} \varphi \rho_n^\infty(0) \, dz \\ &= \int_0^t \int_{(\mathbb{R}^d)^n} \nabla \cdot ([A_{\min}^{\infty, \text{eff}}]_{1 \leq i, j \leq nd} \nabla \varphi) \rho_n^\infty(s) \, dz \, ds. \end{aligned} \quad (3.22)$$

Equation (3.22) completely characterises $\lim_{N \rightarrow \infty} \rho_n^{N,*}$. In particular, we notice that

$$\rho_n^\infty = (S_t^{\min} \# X_0)^n,$$

where $S_t^{\min} : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d)$ is the solution semigroup associated to

$$\partial_t \rho = \nabla \cdot (A_{\min}^{\text{eff}} \nabla \rho) \quad \text{on } (0, \infty) \times \mathbb{R}^d.$$

As the marginals characterise $\lim_{N \rightarrow \infty} \rho^{N,*}(t) = X(t) \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ (cf. [8, Lemma 3]), we obtain the desired result:

$$X(t) = S_t^{\min} \# X_0.$$

Combining this with Theorem C, we have that, for a fixed $t > 0$, the solution $\rho^{\varepsilon,N}(t)$ of (1.3) satisfies

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon,N}(t) = \lim_{N \rightarrow \infty} \rho^{N,*}(t) = X(t) = S_t^{\min} \# X_0.$$

□

4. Proofs of Section 1.9

In this section we include the proofs of Corollary 1.23, Corollary 1.24 and Lemma 1.26.

Proof of Corollary 1.23. The proof follows by combining Corollary 1.11 and Lemma A.2. Indeed, we can first apply Theorem 1.10 which gives us:

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon,N} = S_t^* \# X_0.$$

However, from Lemma A.2, we know that (1.15) can have only one steady state. But Propositions 1.9 and 1.19 tell us that steady states must be minimisers and minimisers always exist. Thus, for $\beta \in (0, \beta_0]$, we have that $\tilde{v}^* = \tilde{v}^{\min}$, the unique minimiser. It follows that

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho^{\varepsilon,N} = S_t^* \# X_0 = S_t^{\min} \# X_0.$$

The limit the other way around follows by applying Theorem 1.16 and using the fact that \tilde{E}_{MF} has a unique minimiser. □

Proof of Corollary 1.24. Since $\lim_{\varepsilon \rightarrow 0} \varepsilon^{-d} \rho_0(\varepsilon^{-1}x) = \delta_0 \in \mathcal{P}(\mathbb{R}^d)$, one can check that

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \rho_0^{\varepsilon,N} = \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \rho_0^{\varepsilon,N} = \delta_{\delta_0} = X_0.$$

The proof of the limit $\varepsilon \rightarrow 0$ followed by $N \rightarrow \infty$ follows by simply applying Theorem 1.16 and using the fact that for $\beta < \beta_c$, \tilde{E}_{MF} has a unique minimiser \tilde{v}^{\min} (cf. Definition 1.18 and Proposition 1.9).

For the other limit, since the initial data $\rho_0^{\varepsilon, N}$ is rapidly varying, the corresponding initial data for (1.14) is precisely $\rho_0^{\otimes N}$ and is independent of $\varepsilon > 0$. Thus we need to show that (A1) holds for some fixed initial data $\tilde{v}_0 \in \mathcal{P}(\mathbb{T}^d)$ independent of $\varepsilon > 0$. Here \tilde{v}_0 is the periodic rearrangement of ρ_0 . We will prove this by using the fact that if $\beta < \beta_c$, then (1.15) has a unique steady state, namely $\tilde{v}_\infty \equiv 1$. This follows simply from our definition of a phase transition and by plugging \tilde{v}_∞ into the right hand side of (1.15). We divide the proof into two steps. In *Step 1*, we show that solutions of (1.15) enjoy certain compactness properties and converge to \tilde{v}_∞ along some time-divergent subsequences. In *Step 2*, we will show that if \tilde{v}_0 is close to \tilde{v}_∞ in an appropriate topology, then this convergence happens exponentially fast and along the whole trajectory, if $\beta < \beta_c$. Combining these together will then establish (A1).

Step 1. By parabolic regularity theory, for any positive time $t > 0$ the solution $\tilde{v}(t)$ of (1.15) is smooth for any positive time. Thus we can assume without loss of generality that $\tilde{v}_0 \in C^\infty(\mathbb{T}^d)$. Furthermore, as discussed in Sect. 1.4, (1.15) is a gradient flow of \tilde{E}_{MF} with respect to the 2-Wasserstein metric d_2 on $\mathcal{P}(\mathbb{T}^d)$. It follows from [1, Theorem 11.1.3], that we have the following energy-dissipation identity along solutions of (1.15):

$$\frac{d}{dt} \tilde{E}_{MF}[\tilde{v}(t)] = -D(\tilde{v}(t)) = - \int_{\mathbb{T}^d} \left| \nabla \log \frac{\tilde{v}(t)}{e^{-\beta W * \tilde{v}(t)}} \right|^2 \tilde{v} \, dx .$$

Integrating from 0 to ∞ and using the fact that the periodic mean field free energy \tilde{E}_{MF} is bounded below, we obtain

$$\int_0^\infty \int_{\mathbb{T}^d} \left| \nabla \log \frac{\tilde{v}(t)}{e^{-\beta W * \tilde{v}(t)}} \right|^2 \tilde{v}(t) \, dx \, dt \leq C$$

for some constant $C > 0$. Thus, there must exist a sequence of times $t_n \rightarrow \infty$ such that

$$\lim_{n \rightarrow \infty} \int_{\mathbb{T}^d} \left| \nabla \log \frac{\tilde{v}(t_n)}{e^{-\beta W * \tilde{v}(t_n)}} \right|^2 \tilde{v}(t_n) \, dx = 0 .$$

Since $|\nabla W * \tilde{v}(t)| \leq \|\nabla W\|_{L^\infty(\mathbb{T}^d)}$, the above limit implies the following bound along the sequence t_n :

$$\int_{\mathbb{T}^d} \left| \nabla \sqrt{\tilde{v}(t_n)} \right|^2 \, dx = \int_{\mathbb{T}^d} |\nabla \log \tilde{v}(t)|^2 \tilde{v}(t_n) \, dx \leq C .$$

Using the fact that $\tilde{v}(t) \in \mathcal{P}(\mathbb{T}^d)$, we have that $\left\| \sqrt{\tilde{v}(t_n)} \right\|_{H^1(\mathbb{T}^d)} \leq C$. Thus, there exists a subsequence of times t_{n_k} and a function $f \in H^1(\mathbb{T}^d)$ such that

$$\sqrt{\tilde{v}(t_{n_k})} \xrightarrow{k \rightarrow \infty} f \quad \text{strongly in } L^2(\mathbb{T}^d), \text{ weakly in } H^1(\mathbb{T}^d) .$$

Furthermore, we have that

$$\begin{aligned}
\left\| \tilde{v}(t_{n_k}) - f^2 \right\|_{L^1(\mathbb{T}^d)} &= \left\| \left(\sqrt{\tilde{v}(t_{n_k})} + f \right) \left(\sqrt{\tilde{v}(t_{n_k})} - f \right) \right\|_{L^1(\mathbb{T}^d)} \\
&\leq \left\| \sqrt{\tilde{v}(t_{n_k})} + f \right\|_{L^2(\mathbb{T}^d)} \left\| \sqrt{\tilde{v}(t_{n_k})} - f \right\|_{L^2(\mathbb{T}^d)} \\
&\leq (1 + \|f\|_{L^2(\mathbb{T}^d)}) \left\| \sqrt{\tilde{v}(t_{n_k})} - f \right\|_{L^2(\mathbb{T}^d)} \xrightarrow{k \rightarrow \infty} 0.
\end{aligned}$$

Thus, $\int_{\mathbb{T}^d} f^2 dx = 1$, $f^2 \geq 0$, and thus $f^2 \in \mathcal{P}(\mathbb{T}^d)$. Next, we check that the dissipation is lower semicontinuous with respect to L^1 convergence. We follow the idea in [37, Section 5] and notice that the dissipation can be written as the supremum of weakly continuous functionals as follows:

$$\begin{aligned}
&\int_{\mathbb{T}^d} \left| \nabla \log \frac{\tilde{v}(t_{n_k})}{e^{-\beta W * \tilde{v}(t_{n_k})}} \right|^2 \tilde{v}(t_{n_k}) dx \\
&= \sup_{\psi \in C^\infty(\mathbb{T}^d)} \int_{\mathbb{T}^d} \left(2 \nabla \log \frac{\tilde{v}(t_{n_k})}{e^{-\beta W * \tilde{v}(t_{n_k})}} \nabla \psi - |\nabla \psi|^2 \right) \tilde{v}(t_{n_k}) dx \\
&= \sup_{\psi \in C^\infty(\mathbb{T}^d)} \int_{\mathbb{T}^d} \left(2(-\Delta)\psi + 2\beta(\nabla W * \tilde{v}(t_{n_k}))\nabla \psi - |\nabla \psi|^2 \right) \tilde{v}(t_{n_k}) dx.
\end{aligned}$$

This implies that it is lower semicontinuous with respect to weak-* convergence

$$\int_{\mathbb{T}^d} \left| \nabla \log \frac{f^2}{e^{-\beta W * f^2}} \right|^2 f^2 dx \leq \liminf_{k \rightarrow \infty} \int_{\mathbb{T}^d} \left| \nabla \log \frac{\tilde{v}(t_{n_k})}{e^{-\beta W * \tilde{v}(t_{n_k})}} \right|^2 \tilde{v}(t_{n_k}) dx = 0.$$

It follows then that $D(f^2) = 0$ and from Proposition 1.9, that $f^2 \in \mathcal{P}(\mathbb{T}^d)$ is a steady state of (1.15). Since \tilde{v}_∞ is the only stationary solution for $\beta < \beta_c$, it must hold that $f^2 = \tilde{v}_\infty$ and that

$$\lim_{k \rightarrow \infty} \left\| \tilde{v}_{t_{n_k}} - \tilde{v}_\infty \right\|_{L^1(\mathbb{T}^d)} = 0.$$

Step 2. We now use [10, Theorem 2.11] which tells us that if $\beta < \beta_* := -(\min_k \hat{W}(k))^{-1}$ and $\|\tilde{v}_0 - \tilde{v}_\infty\|_{L^1(\mathbb{T}^d)} < \varepsilon_0$, then

$$\|\tilde{v}(t) - \tilde{v}_\infty\|_{L^1(\mathbb{T}^d)} \leq \|\tilde{v}_0 - \tilde{v}_\infty\|_{L^1(\mathbb{T}^d)} e^{-Ct}$$

for some $C > 0$, $\varepsilon_0 > 0$, and all $t \geq 0$. Since we know from the previous step that $\lim_{k \rightarrow \infty} \|\tilde{v}(t_{n_k}) - \tilde{v}_\infty\|_{L^1(\mathbb{T}^d)} = 0$, there must exist some time $T > 0$ such that $\|\tilde{v}(T) - \tilde{v}_\infty\|_{L^1(\mathbb{T}^d)} < \varepsilon_0$. We also know from [9, Proposition 5.3] that $\beta_c \leq \beta_*$. Thus for all $\beta < \beta_c$, we have that

$$\|\tilde{v}(t) - \tilde{v}_\infty\|_{L^1(\mathbb{T}^d)} \leq C_T e^{-C(t-T)},$$

where $C_T := \max_{s \in [0, T]} \|\tilde{v}(s) - \tilde{v}_\infty\|_{L^1(\mathbb{T}^d)} \leq 2$. Thus, we have shown that (A1) holds, completing the proof of the first part of the result.

We remind the reader that $W \in \mathbf{H}_s$ means that $\hat{W}(k) \geq 0$ for all $k \in \mathbb{Z}^d$. For the second half of the result, we use the fact $W \in \mathbf{H}_s$ implies, by Proposition 1.21, that $\beta_c = +\infty$ and thus the result of the corollary necessarily holds for all $\beta < +\infty$ and rapidly varying initial data. We now sketch how to extend the result to all chaotic initial data. We note that by applying Duhamel's formula for the solution of (1.15) one can show that there exists a time, say $t' = 1 > 0$, such that for all initial data $\mathcal{H}(\tilde{v}(1)|\tilde{v}_\infty) < C$ for some fixed constant $C \geq 0$. Additionally, we can apply [9, Proposition 3.1], to assert that for $W \in \mathbf{H}_s$ and all $\beta < \infty$, we have that

$$\mathcal{H}(\tilde{v}(t)|\tilde{v}_\infty) \leq \mathcal{H}(\tilde{v}(1)|\tilde{v}_\infty)e^{-C_1(t-1)}$$

for all $t \geq 1$. Since the relative entropy controls the 2-Wasserstein distance, we can apply Corollary 1.11 to complete the proof of the result. \square

Proof of Lemma 1.26. We know from Proposition 1.9 that steady states of the quotiented periodic system (1.15) are equivalent to solutions of the self-consistency equation (1.17), which we rewrite as

$$\tilde{v} = \frac{e^{-\beta(V+W*\tilde{v})}}{Z}, \quad Z = \int_{\mathbb{T}} e^{-\beta(V+W*\tilde{v})} dx. \quad (4.1)$$

We also know from Proposition 1.19 that for β sufficiently small the map in the above expression has a unique fixed point. Thus (1.15) has a unique steady state for β sufficiently small. Since minimisers of \tilde{E}_{MF} exist and are always steady states (cf. Propositions 1.19 and 1.9), it must also be the unique minimiser of \tilde{E}_{MF} . We argue further that any minimiser of \tilde{E}_{MF} must be symmetric about $x = 1/2$ and decreasing from 0 to $1/2$. This follows directly from the Baernstein–Taylor inequality for spherical rearrangements of functions [2].

To investigate the problem ahead of the phase transition, we consider (4.1). Plugging our choice of V and W and testing against $\cos(2\pi x)$ we can simplify this to

$$\tilde{v}_1 = Z^{-1} \int_{\mathbb{T}^d} \cos(2\pi x) \exp(\beta(\cos(2\pi x)(\eta + \tilde{v}_1) + \sin(2\pi x)\tilde{v}_{-1})) dx,$$

where $\tilde{v}_1 = \langle \tilde{v}, \cos(2\pi x) \rangle$ and $\tilde{v}_{-1} = \langle \tilde{v}, \sin(2\pi x) \rangle$. Let us consider the problem when $\tilde{v}_{-1} = 0$, as this corresponds to the setting when \tilde{v} is symmetric about $x = 1/2$. Simplifying further, we obtain

$$\tilde{v}_1 = Z^{-1} \int_{\mathbb{T}^d} \cos(2\pi x) \exp(\beta(\cos(2\pi x)(\eta + \tilde{v}_1))) dx.$$

Using the fact the modified Bessel functions of the first kind can be expressed as $I_n(y) = \int_{\mathbb{T}} \cos(2\pi nx) e^{y \cos(2\pi x)} dx$, we obtain,

$$\tilde{v}_1 = r_0(\beta(\eta + \tilde{v}_1))$$

where $r_0(x) := I_1(x)/I_0(x)$, and I_1, I_0 are first and zeroth modified Bessel functions of the first kind. Setting $\beta(\eta + \tilde{v}_1) = a$ we simplify the above expression to

$$a = \beta(\eta + r_0(a)). \quad (4.2)$$

The function $r_0(a)$ has the following properties [9, Proposition 6.1]:

$$\begin{aligned} r_0(0) &= 0 \\ \lim_{a \rightarrow \infty} r_0(a) &= 1 \quad \lim_{a \rightarrow -\infty} r_0(a) = -1 \end{aligned} \tag{4.3}$$

$$r_0''(a) < 0, \quad a > 0 \tag{4.4}$$

We notice that assuming the symmetry around $x = 1/2$ (or $\tilde{v}_{-1} = 0$), the solutions of (4.1) are completely determined by the value of its first Fourier mode \tilde{v}_1 :

$$\tilde{v} = \frac{\exp(\beta(\cos(2\pi x)(\eta + \tilde{v}_1)))}{\int_{\mathbb{T}^d} \exp(\beta(\cos(2\pi y)(\eta + \tilde{v}_1))) dy}.$$

Therefore, obtaining a solution to (4.1) is equivalent to finding solutions $\tilde{v}_1 = a$ to (4.2). By the Baernstein-Taylor re-arrangement inequality, a necessary condition for \tilde{v} to be a minimizer of the free energy, is that $\tilde{v}_1 \geq -\eta$, (if $\tilde{v}_1 < -\eta$ then \tilde{v} is decreasing from 0 to $1/2$). We argue now that (4.2) has exactly one solution for $a > 0$, for all $\beta > 0$. Consider the function $F : \mathbb{R} \rightarrow \mathbb{R}$ defined as follows:

$$F(a) := \beta(\eta + r_0(a)) - a.$$

We know that $F(0) = \beta\eta > 0$. Furthermore for a large enough and positive we have that $F(a) < 0$, using (4.3). Thus, by the intermediate value theorem, for every fixed $\beta > 0$, we can find at least one $a^{\min} > 0$ such that $F(a^{\min}) = 0$. Now if $F(a^{\min}) = 0$ for some $a^{\min} > 0$, we must have that $\beta r_0'(a^{\min}) < 1$. If not, we would have that

$$\begin{aligned} F(a^{\min}) &= \beta\eta + \int_0^{a^{\min}} (\beta r_0'(a) - 1) da \\ &\geq \beta\eta + a^{\min}(\beta r_0'(a^{\min}) - 1) > 0, \end{aligned}$$

which is a contradiction. In the last inequality we have used (4.4). This implies that

$$F'(a^{\min}) = \beta r_0'(a^{\min}) - 1 < 0.$$

Also,

$$F''(a) = \beta r_0''(a) < 0.$$

Thus, once $F'(a) < 0$, it remains negative for all $a > 0$. It follows that $F(a^{\min}) = 0$ for only one $a^{\min} > 0$. Since this is the only symmetric decreasing solution of (1.17) it corresponds to the unique minimiser of \tilde{E}_{MF} through the expression in (1.23). It is also must be the unique steady state obtained using the contraction argument earlier in the proof.

We will now show that for β large enough we can find another solution of (4.2) for $a < 0$. Let $\eta = 1 - \delta$ for some $\delta \in (0, 1)$. From (4.3) we know that there exists some $a' < 0$ such that for all $a \leq a'$, $r_0(a) < -1 + \delta/2$. We then have that

$$\begin{aligned} F(a') &= \beta - \beta\delta + \beta r_0(a') - a' \\ &< -\beta \frac{\delta}{2} - a'. \end{aligned}$$

Furthermore, since $r_0(a)$ is an odd function and $\eta > 0$, if a^{\min} is a solution of (4.2), then $-a^{\min}$ cannot be a solution. It follows that $a^* \neq -a^{\min}$. \square

Acknowledgements. MGD was partially supported by EPSRC grant number EP/P031587/1, CNPq-Brazil(#308800/2019-2) and Instituto Serrapilheira. RSG is funded by an Imperial College President’s PhD Scholarship, partially through EPSRC Award Ref. 1676118. Part of this work was carried out at the “Junior Trimester Programme in Kinetic Theory” held at the Hausdorff Research Institute for Mathematics, Bonn. RSG is grateful to the institute for its hospitality. GAP was partially supported by the EPSRC through grant numbers EP/P031587/1, EP/L024926/1, and EP/L020564/1. This research was funded in part by JP-Morgan Chase & Co. Any views or opinions expressed herein are solely those of the authors listed, and may differ from the views and opinions expressed by JPMorgan Chase & Co. or its affiliates. This material is not a product of the Research Department of J.P. Morgan Securities LLC. This material does not constitute a solicitation or offer in any jurisdiction. The authors would like to thank Martin Hairer and Felix Otto for useful discussions during the course of this work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Appendix A. Coupling Arguments

In this section we will use coupling techniques introduced by Eberle and co-authors [15–18] to show some necessary results for our proofs. Following the previous strategy we construct a new metric which is equivalent to the Wasserstein metric. We define the constant

$$\kappa := \inf_{x \in \mathbb{T}} V''(x) + \inf_{x \in \mathbb{T}} W''(x) \leq 0,$$

which gives a lower bound of the semi-convexity of the function of V and $V + W * \tilde{v}^*$ on \mathbb{T} . Next, we define the following functions on $[0, 1/2]$:

$$\begin{aligned} \psi(r) &:= \exp\left(\frac{\beta \kappa r^2}{8}\right), \quad \Phi(r) := \int_0^r \psi(s) \, ds, \\ g(r) &:= 1 - \frac{c}{2} \int_0^r \Phi(s)(\psi(s))^{-1} \, ds, \\ c &:= \left(\int_0^{1/2} \Phi(s)(\psi(s))^{-1} \, ds\right)^{-1} \geq \frac{\beta |\kappa|}{4 \left(e^{\frac{\beta |\kappa|}{32}} - 1\right)}. \end{aligned} \tag{A.1}$$

We note that $g(r) \in [1/2, 1]$ for all $r \in [0, 1/2]$ and $\lim_{\beta \rightarrow 0^+} c = 1/8$. Additionally, both ψ and g are decreasing functions of r . Thus the function $f : [0, 1/2] \rightarrow [0, 1/2]$, defined as

$$f(r) := \int_0^r g(s)\psi(s) \, ds,$$

is increasing and subadditive. Furthermore, we have the bounds

$$\frac{\psi(1/2)}{2}r \leq f(r) \leq \Phi(r) \leq r. \quad (\text{A.2})$$

Thus $d_f(x, y) := f(d_{\mathbb{T}}(x, y))$ defines a metric on \mathbb{T} which is equivalent to $d_{\mathbb{T}}$. The main point of this construction is to obtain the following inequality:

$$f''(r) - \beta r \kappa \frac{f'(r)}{4} \leq -\frac{c}{2}f(r) \quad \text{for all } r \in [0, 1/2]. \quad (\text{A.3})$$

This easily follows from the following computation:

$$f''(r) - \beta r \kappa \frac{f'(r)}{4} = \beta r \frac{\kappa}{4} f'(r) - \frac{c}{2}\Phi(r) - \beta r \kappa \frac{f'(r)}{4} = -\frac{c}{2}\Phi(r) \leq -\frac{c}{2}f(r).$$

Moreover, we define the Lipschitz functions $\varphi_r^\delta, \varphi_s^\delta : \mathbb{R} \rightarrow \mathbb{R}$ for some $\delta > 0$, such that

$$(\varphi_r^\delta)^2(x) + (\varphi_s^\delta)^2(x) = 1 \quad \varphi_r^\delta(x) = \begin{cases} 0 & \gamma(|x|) \leq \delta/2 \\ 1 & \gamma(|x|) > \delta, \end{cases} \quad (\text{A.4})$$

where the function $\gamma : \mathbb{R}_+ \rightarrow [0, 1/2]$ maps Euclidean distances to distances on the torus

$$\gamma(|x|) := \begin{cases} (|x| \bmod 1) & (|x| \bmod 1) \leq 1/2 \\ 1 - (|x| \bmod 1) & \text{otherwise} \end{cases}.$$

The introduction of the function γ to account for the periodic setting is the main difference with the results in the literature [15–18].

We have the following result:

Lemma A.1. *Assume that (A1) holds and consider the two SDEs in (2.11) and (2.19). Then there exists a coupling of (\dot{X}_t, \dot{Y}_t) and a metric d_f on \mathbb{T}^d which is equivalent to $d_{\mathbb{T}^d}$ such that*

$$\sup_{\varepsilon > 0} \mathbb{E} \left[d_f(\dot{X}_t, \dot{Y}_t)^2 \right] \rightarrow 0$$

as $t \rightarrow \infty$.

Proof. For convenience we write the proofs in 1 space dimension. The generalization to higher dimensions follows along similar lines. We start the proof by using the metric d_f defined previously. We now proceed to construct the coupling between the two processes by considering the corresponding processes on \mathbb{R} , i.e. Y_t and X_t . We assume that $\text{Law}(X_0) = \nu^* \in \mathcal{P}(\mathbb{R})$ such that the periodic rearrangement of ν^* is precisely $\tilde{\nu}^*$.

Let B_t^1 and B_t^2 be two independent standard Wiener processes which are also independent of the initial conditions. We then couple the processes in a similar manner to [18] as follows:

$$\begin{cases} dY_t = -V'(Y_t) dt - W' * \tilde{\nu}^\varepsilon(t)(Y_t) dt + \sqrt{2\beta^{-1}} \left(\varphi_r^\delta(E_t) dB_t^1 + \varphi_s^\delta(E_t) dB_t^2 \right) \\ \text{Law}(Y_0) = \nu_0^\varepsilon \in \mathcal{P}(\mathbb{R}), \end{cases} \quad (\text{A.5})$$

$$\begin{cases} dX_t = -V'(X_t) dt - W' * \tilde{\nu}^*(X_t) dt + \sqrt{2\beta^{-1}} \left(-\varphi_r^\delta(E_t) dB_t^1 + \varphi_s^\delta(E_t) dB_t^2 \right) \\ \text{Law}(X_0) = \nu^* \in \mathcal{P}(\mathbb{R}). \end{cases} \quad (\text{A.6})$$

Here $E_t := Y_t - X_t$ and X_0, Y_0 are independent. This above coupling corresponds to a combination of reflection and synchronous coupling. Note that we have suppressed the dependence on δ for the sake of notational convenience. However, in the limit as $\delta \rightarrow 0_+$ the processes X_t and Y_t converge \mathbb{P} -a.s. to corresponding limits with only reflection coupling. We also define the following function:

$$e_t := \begin{cases} \frac{E_t}{|E_t|} & |E_t| > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Subtracting (A.6) from (A.5) and using the same arguments as in [18] we obtain

$$\begin{aligned} d|E_t| &= -(V'(Y_t) + W' * \tilde{v}^\varepsilon(t)(Y_t) - V'(X_t) - W' * \tilde{v}^*(X_t))(e_t) dt \\ &\quad + 2\sqrt{2\beta^{-1}}\varphi_r^\delta(E_t)e_t dB_t^1. \end{aligned}$$

Note now that the function $\mathbb{R}_+ \ni x \rightarrow \gamma(x)$ is a function whose derivatives are of locally bounded variation. Thus it can be expressed as the difference of two convex functions [25, Theorem (I)]. We can thus apply the Meyer–Tanaka formula [29, Theorem 6.22] to it, to obtain

$$\begin{aligned} \gamma(|E_t|) &= \gamma(|E_0|) - \int_0^t \gamma'_\ell(|E_s|)(V'(Y_s) - W' * \tilde{v}^\varepsilon(s)(Y_s) - V'(X_s) - W' * \tilde{v}^*(X_s))(e_s) ds \\ &\quad + \int_0^t \gamma'_\ell(|E_s|)2\sqrt{2\beta^{-1}}\varphi_r^\delta(E_s)e_s dB_s^1 + \int_{\mathbb{R}_+} \Lambda_t(a) d\gamma''_-(a), \end{aligned}$$

where γ'_ℓ is the left derivative of γ , Λ_t is the local time of the process $|E_t|$, and γ''_- is the negative part γ'' the distributional derivative of γ . We can throw away the positive part as $\varphi_r^\delta(0) = 0$. The reader should note that $\gamma(|E_t|) = d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)$, i.e. it is distance on the torus between the quotiented processes. Since the local time is an adapted non-decreasing continuous process it follows that $A_t := \int_{\mathbb{R}_+} \Lambda_t(a) d\gamma''_-(a)$ is an adapted nonincreasing continuous process.

Since $\gamma_t := \gamma(|E_t|)$ is a continuous semimartingale we can apply Itô's formula to $f(\gamma_t)$ to obtain

$$\begin{aligned} df(\gamma_t) &= -f'(\gamma_t)\gamma'_\ell(|E_t|)(V'(Y_t) + W' * \tilde{v}^\varepsilon(t)(Y_t) - V'(X_t) - W' * \tilde{v}^*(X_t))(e_t) dt \\ &\quad + f'(\gamma_t)dA_t + \gamma'_\ell(|E_t|)2\sqrt{2\beta^{-1}}\varphi_r^\delta(E_t)e_t dB_t^1 \\ &\quad + 4f''(\gamma_t)\beta^{-1}(\varphi_r^\delta(E_t))^2(\gamma'_\ell(|E_t|))^2 dt \end{aligned}$$

Next, we note that since $f'(x) \geq 0$ and A_t is nonincreasing we have the bound

$$\begin{aligned} df(\gamma_t) &\leq -f'(\gamma_t)\gamma'_\ell(|E_t|)(V'(Y_t) + W' * \tilde{v}^*(Y_t) - V'(X_t) - W' * \tilde{v}^*(X_t))(e_t) dt \\ &\quad + f'(\gamma_t)\gamma'_\ell(|E_t|)(-W' * \tilde{v}^\varepsilon(t)(Y_t) + W' * \tilde{v}^*(Y_t))(e_t) \\ &\quad + \gamma'_\ell(|E_t|)2\sqrt{2\beta^{-1}}\varphi_r^\delta(E_t)e_t dB_t^1 \\ &\quad + 4f''(\gamma_t)\beta^{-1}(\varphi_r^\delta(E_t))^2 dt \\ &\leq -f'(\gamma_t)\gamma'_\ell(|E_t|)(V'(Y_t) + W' * \tilde{v}^*(Y_t) - V'(X_t) - W' * \tilde{v}^*(X_t))(e_t) dt \\ &\quad + \|f'\|_{L^\infty(\mathbb{T})}\|\gamma'_\ell\|_{L^\infty(\mathbb{T})}\|W''\|_{L^\infty(\mathbb{T})}d_1(\tilde{v}^\varepsilon(t), \tilde{v}^*) dt \\ &\quad + \gamma'_\ell(|E_t|)2\sqrt{2\beta^{-1}}\varphi_r^\delta(E_t)e_t dB_t^1 \\ &\quad + 4f''(\gamma_t)\beta^{-1}(\varphi_r^\delta(E_t))^2 dt, \end{aligned} \tag{A.7}$$

where in the second inequality we have used the dual formulation of the 1-Wasserstein distance. Consider now the 1-periodic function $F := V + W * \tilde{v}^*$. Using the definition of γ and κ , we can obtain the following inequality:

$$\gamma'_\ell(|E_t|)(F'(Y_t) - F'(X_t))(e_t) \geq \kappa d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t) = \kappa \gamma_t.$$

Indeed, for $0 < (|E_t| \bmod 1) \leq 1/2$, we have that $\gamma'_\ell(|E_t|) = 1$, from which it follows that

$$\begin{aligned} \gamma'_\ell(|E_t|)(F'(Y_t) - F'(X_t))(e_t) &= (F'(Y_t) - F'(X_t))e_t \frac{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)}{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)} \\ &= (F'(\dot{Y}_t) - F'(\dot{X}_t))(\dot{Y}_t - \dot{X}_t) \frac{1}{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)} \\ &\geq \kappa \frac{|\dot{Y}_t - \dot{X}_t|^2}{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)} \\ &= \kappa d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t), \end{aligned}$$

where we have used the fact that

$$|\dot{Y}_t - \dot{X}_t| = d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t),$$

if $(|E_t| \bmod 1) \leq 1/2$. On the other hand if $1/2 < (|E_t| \bmod 1) < 1$, we have that $\gamma'_\ell(|E_t|) = -1$, and so

$$\begin{aligned} \gamma'_\ell(|E_t|)(F'(Y_t) - F'(X_t))(e_t) &= -(F'(Y_t) - F'(X_t))e_t \frac{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)}{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)} \\ &= (F'(\dot{Y}_t) - F'(\dot{X}_t + 1))(\dot{Y}_t - \dot{X}_t - 1) \frac{1}{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)} \\ &\geq \kappa \frac{|\dot{Y}_t - \dot{X}_t - 1|^2}{d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)} \\ &= \kappa d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t), \end{aligned}$$

where we have used the fact that

$$|\dot{Y}_t - \dot{X}_t - 1| = d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)$$

if $(|E_t| \bmod 1) > 1/2$. The proof for $(|E_t| \bmod 1) = 0$ follows trivially since both sides of the inequality are zero.

Applying this estimate to (A.7) and using the fact that $f' > 0$, we obtain

$$\begin{aligned} df(\gamma_t) &\leq -\kappa f'(\gamma_t)\gamma_t dt + 4f''(\gamma_t)\beta^{-1}(\varphi_r^\delta(E_t))^2 dt \\ &\quad + \|W''\|_{L^\infty(\mathbb{T})}d_1(\tilde{v}^\varepsilon(t), \tilde{v}^*) dt + \gamma'_\ell(|E_t|)2\sqrt{2\beta^{-1}}\varphi_r^\delta(E_t)e_t dB_t^1. \\ &= -\kappa f'(\gamma_t)\gamma_t(\varphi_r^\delta(E_t))^2 dt + 4f''(\gamma_t)\beta^{-1}(\varphi_r^\delta(E_t))^2 dt \\ &\quad + \kappa f'(\gamma_t)\gamma_t((\varphi_r^\delta(E_t))^2 - 1) dt \\ &\quad + \|W''\|_{L^\infty(\mathbb{T})}d_1(\tilde{v}^\varepsilon(t), \tilde{v}^*) dt + \gamma'_\ell(|E_t|)2\sqrt{2\beta^{-1}}\varphi_r^\delta(E_t)e_t dB_t^1. \end{aligned}$$

Applying the differential inequality for f (A.3), $f' \leq 1$, and the definition of φ_r^δ (A.4), we obtain

$$\begin{aligned} df(\gamma_t) &\leq -2c\beta^{-1} f(\gamma_t) (\varphi_r^\delta(E_t))^\delta dt + \frac{|\kappa|}{2} \delta dt + \|W''\|_{L^\infty(\mathbb{T})} d_1(\tilde{v}^\varepsilon(t), \tilde{v}^*) dt \\ &\quad + \gamma'_\ell(|E_t|) 2\sqrt{2\beta^{-1}} \varphi_r^\delta(E_t) e_t dB_t^1 \\ &\leq -2c\beta^{-1} f(\gamma_t) dt + 2c\beta^{-1} f(\delta) dt + \frac{|\kappa|}{2} \delta dt + \|W''\|_{L^\infty(\mathbb{T})} d_1(\tilde{v}^\varepsilon(t), \tilde{v}^*) dt \\ &\quad + \gamma'_\ell(|E_t|) 2\sqrt{2\beta^{-1}} \varphi_r^\delta(E_t) e_t dB_t^1. \end{aligned}$$

Taking the expectation of the above expression and passing to the limit as $\delta \rightarrow 0^+$, we obtain

$$\frac{d}{dt} \mathbb{E}[f(\gamma_t)] \leq -2c\beta^{-1} \mathbb{E}[f(\gamma_t)] + \|W''\|_{L^\infty(\mathbb{T})} d_1(\tilde{v}^\varepsilon(t), \tilde{v}^*).$$

It follows by Gronwall's Lemma that

$$\mathbb{E}[f(\gamma_t)] \leq e^{-2c\beta^{-1}t} \mathbb{E}[f(\gamma_0)] + \|W''\|_{L^\infty(\mathbb{T})} e^{-2c\beta^{-1}t} \int_0^t e^{2c\beta^{-1}s} d_1(\tilde{v}^\varepsilon(s), \tilde{v}^*) ds. \quad (\text{A.8})$$

Applying (A1), we have that $\mathbb{E}[d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)] = \mathbb{E}[f(\gamma_t)] \rightarrow 0$ as $t \rightarrow \infty$, \dot{X}_t and \dot{Y}_t are the quotiented processes obtained in the limit as $\delta \rightarrow 0_+$. This completes the proof. \square

Lemma A.2. *Given V and W , there exists an explicit β_0 depending V and W such that for*

$$\beta \leq \beta_0$$

there exists a unique minimiser and critical point, $\tilde{v}^{\min} \in \mathcal{P}(\mathbb{T}^d)$, of the periodic mean field energy (1.16) and C_2 depending on β , W and V such that

$$d_1(\tilde{v}(t), \tilde{v}^{\min}) \leq e^{-C_2 t},$$

where $v(t)$ is the solution to the periodic McKean–Vlasov equation (1.15) with arbitrary initial data $\tilde{v}_0 \in \mathcal{P}(\mathbb{T}^d)$.

Proof. As done previously, we state the proof in 1 dimension for the sake of simplicity. Clearly for β small enough, by Proposition 1.19, the periodic mean field energy has a unique minimiser, \tilde{v}^{\min} . Similar to the proof of (A.1), we consider the processes on \mathbb{R}

$$\begin{cases} dY_t = -V'(Y_t) dt - W' * \tilde{v}(t)(Y_t) dt + \sqrt{2\beta^{-1}} \left(\varphi_r^\delta(E_t) dB_t^1 + \varphi_s^\delta(E_t) dB_t^2 \right) \\ \text{Law}(Y_0) = \nu_0 \in \mathcal{P}(\mathbb{R}), \\ \\ dX_t = -V'(X_t) dt - W' * \tilde{v}^{\min}(X_t) dt + \sqrt{2\beta^{-1}} \left(-\varphi_r^\delta(E_t) dB_t^1 + \varphi_s^\delta(E_t) dB_t^2 \right) \\ \text{Law}(X_0) = \nu^{\min} \in \mathcal{P}(\mathbb{R}). \end{cases}$$

such that $\tilde{v}_0, \tilde{v}^{\min}$ are the periodic rearrangements of ν_0, ν^{\min} , respectively. We obtain the inequality (A.8)

$$\begin{aligned} \mathbb{E}[f(\gamma_t)] &\leq e^{-2c\beta^{-1}t} \mathbb{E}[f(\gamma_0)] \\ &\quad + \|W''\|_{L^\infty(\mathbb{T})} e^{-2c\beta^{-1}t} \int_0^t e^{2c\beta^{-1}s} d_1(\tilde{v}(s), \tilde{v}^{\min}) ds, \end{aligned}$$

where $\gamma_t = d_{\mathbb{T}}(\dot{X}_t, \dot{Y}_t)$. Using the fact that $\text{Law}(\dot{Y}_t) = \tilde{\nu}(t)$ and $\text{Law}(\dot{X}_t) = \tilde{\nu}^{\min}$ and applying the bounds from (A.2), we obtain

$$\frac{e^{\frac{\beta k}{32}}}{2} d_1(\tilde{\nu}(t), \tilde{\nu}^{\min}) \leq d_f(\tilde{\nu}(t), \tilde{\nu}^{\min}) \leq \mathbb{E}[f(\gamma_t)].$$

Combining this with the previous identity and applying the integral version of Gronwall's lemma, we obtain

$$d_1(\tilde{\nu}(t), \tilde{\nu}^{\min}) \leq e^{-2c\beta^{-1}t} + t e^{-t} \left(\beta^{-1} c - e^{-\frac{\beta k}{32} \|W''\|_{L^\infty(\mathbb{T})}} \right).$$

Using the lower bound for c (A.1), we have the following: if

$$\frac{|\kappa|}{4e^{\frac{\beta|\kappa|}{32}} \left(e^{\frac{\beta|\kappa|}{32}} - 1 \right)} \geq \|W''\|_{L^\infty(\mathbb{T})},$$

then there exists $C_2 > 0$ such that

$$d_1(\tilde{\nu}(t), \tilde{\nu}^{\min}) \leq e^{-C_2 t}.$$

By making β smaller than some β_0 , this can be achieved. \square

References

1. AMBROSIO, L., GIGLI, N., SAVARÉ, G.: *Gradient Flows in Metric Spaces and in the Space of Probability measures*. Lectures in Mathematics Birkhäuser Verlag, Basel 2005
2. BAERNSTEIN II, A., TAYLOR, B.A.: Spherical rearrangements, subharmonic functions, and $*$ -functions in n -space. *Duke Math. J.* **43**(2), 245–268, 1976
3. BARRÉ, J., BERNARDIN, C., CHÉTRITE, R., CHOPRA, Y., MARIANI, M.: From fluctuating kinetics to fluctuating hydrodynamics: a Γ -convergence of large deviations functionals approach. *J. Stat. Phys.* **180**(1–6), 1095–1127, 2020
4. BENSOUSSAN, A., LIONS, J.-L., PAPANICOLAOU, G.: *Asymptotic analysis for periodic structures*. AMS Chelsea Publishing, Providence, 2011. [Corrected reprint of the 1978 original [MR0503330]]
5. BERMAN, R.J., ÖNNHEIM, M.: Propagation of chaos, Wasserstein gradient flows and toric Kähler–Einstein metrics. *Anal. PDE* **11**(6), 1343–1380, 2018
6. BERTINI, L., GIACOMIN, G., PAKDAMAN, K.: Dynamical aspects of mean field plane rotators and the Kuramoto model. *J. Stat. Phys.* **138**(1–3), 270–290, 2010
7. CARDIALAGUET, P.: *Notes on mean-field games (from pl. lions' lectures at college de france)*, 2013
8. CARRILLO, J.A., DELGADINO, M.G., PAVLIOTIS, G.A.: A λ -convexity based proof for the propagation of chaos for weakly interacting stochastic particles. *J. Funct. Anal.* **279**(10), 108734, 2020
9. CARRILLO, J.A., GVALANI, R.S., PAVLIOTIS, G.A., SCHLICHTING, A.: Long-time behaviour and phase transitions for the McKean–Vlasov equation on the torus. *Arch. Ration. Mech. Anal.* 2019
10. CHAYES, L., PANFEROV, V.: The McKean–Vlasov equation in finite volume. *J. Stat. Phys.* **138**(1–3), 351–380, 2010
11. DAWSON, D.A.: Critical dynamics and fluctuations for a mean-field model of cooperative behavior. *J. Stat. Phys.* **31**(1), 29–85, 1983
12. de FINETTI, B.: La prévision?: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7**(1), 1–68, 1937

13. De MASI, A., FERRARI, P.A., GOLDSTEIN, S., WICK, W.D.: An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Stat. Phys.* **55**(3–4), 787–855, 1989
14. DIACONIS, P., FREEDMAN, D.: Finite exchangeable sequences. *Ann. Probab.* **8**(4), 745–764, 1980
15. DURMUS, A., EBERLE, A., GUILLIN, A., ZIMMER, R.: An elementary approach to uniform in time propagation of chaos. *Proc. Am. Math. Soc.* **148**(12), 5387–5398, 2020
16. EBERLE, A.: Reflection coupling and Wasserstein contractivity without convexity. *C. R. Math. Acad. Sci. Paris* **349**(19–20), 1101–1104, 2011
17. EBERLE, A.: Reflection couplings and contraction rates for diffusions. *Probab. Theory Relat. Fields* **166**(3–4), 851–886, 2016
18. EBERLE, A., GUILLIN, A., ZIMMER, R.: Couplings and quantitative contraction rates for Langevin dynamics. *Ann. Probab.* **47**(4), 1982–2010, 2019
19. FERNANDEZ, B., MÉLÉARD, S.: A Hilbertian approach for fluctuations on the McKean–Vlasov model. *Stochast. Process. Appl.* **71**(1), 33–53, 1997
20. FOURNIER, N., JOURDAIN, B.: Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes. *Ann. Appl. Probab.* **27**(5), 2807–2861, 2017
21. FRIEDLI, S., VELENIK, Y.: *Statistical Mechanics of Lattice Systems*. Cambridge University Press, Cambridge 2018
22. GARNIER, J., PAPANICOLAOU, G., YANG, T.-W.: Consensus convergence with stochastic effects. *Vietnam J. Math.* **45**(1–2), 51–75, 2017
23. GOLSE, F.: On the dynamics of large particle systems in the mean field limit, 2013. arXiv e-prints [arXiv:1301.5494](https://arxiv.org/abs/1301.5494)
24. GOMES, S.N., PAVLIOTIS, G.A.: Mean field limits for interacting diffusions in a two-scale potential. *J. Nonlinear Sci.* **28**(3), 905–941, 2018
25. HARTMAN, P.: On functions representable as a difference of convex functions. *Pacific J. Math.* **9**, 707–713, 1959
26. HAURAY, M., MISCHLER, S.: On Kac’s chaos and related problems. *J. Funct. Anal.* **266**(10), 6055–6157, 2014
27. HEWITT, E., SAVAGE, L.J.: Symmetric measures on Cartesian products. *Trans. Am. Math. Soc.* **80**, 470–501, 1955
28. KANTAS, N., PARPAS, P., PAVLIOTIS, G.A.: The sharp, the flat and the shallow: Can weakly interacting agents learn to escape bad minima? 2019. arXiv e-prints [arXiv:1905.04121](https://arxiv.org/abs/1905.04121)
29. KARATZAS, I., SHREVE, S.E.: *Brownian Motion and Stochastic Calculus, Volume 113 of Graduate Texts in Mathematics*, 2nd edn. Springer, New York, 1991
30. KIPNIS, C., VARADHAN, S.R.S.: Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun. Math. Phys.* **104**(1), 1–19, 1986
31. KOMOROWSKI, T., LANDIM, C., OLLA, S.: *Fluctuations in Markov Processes, Volume 345 of Grundlehren der Mathematischen Wissenschaften* [Fundamental Principles of Mathematical Sciences]. Springer, Heidelberg, 2012
32. LIEBERMAN, G.M.: *Second Order Parabolic Differential Equations*. World Scientific Publishing Co. Inc., River Edge 1996
33. LIONS, P.: *Mean-Field Games and Applications*. Lectures at the College de France, 2007
34. MALRIEU, F., et al.: Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* **13**(2), 540–560, 2003
35. MESSER, J., SPOHN, H.: Statistical mechanics of the isothermal Lane–Emden equation. *J. Stat. Phys.* **29**(3), 561–578, 1982
36. OELSCHLÄGER, K.: A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann. Probab.* **12**(2), 458–479, 1984
37. OTTO, F.: The geometry of dissipative evolution equations: the porous medium equation, 2001
38. PAVLIOTIS, G.A., STUART, A.M.: *Multiscale Methods*, vol. 53. Texts in Applied MathematicsSpringer, New York 2008

39. REYGNER, J.: Equilibrium large deviations for mean-field systems with translation invariance. *Ann. Appl. Probab.* **28**(5), 2922–2965, 2018
40. ROTSKOFF, G.M., VANDEN-EIJNDEN, E.: Trainability and accuracy of neural networks: an interacting particle system approach, 2018. arXiv e-prints [arXiv:1805.00915](https://arxiv.org/abs/1805.00915)
41. ROUGERIE, N.: De finetti theorems, mean-field limits and Bose-Einstein condensation, 2015. arXiv e-prints [arXiv:1506.05263](https://arxiv.org/abs/1506.05263)
42. RUELLE, D.: *Statistical Mechanics: Rigorous Results*. W. A. Benjamin Inc., New York 1969
43. SERFATY, S.: *Coulomb Gases and Ginzburg-Landau Vortices*. Zurich Lectures in Advanced Mathematics European Mathematical Society (EMS), Zürich 2015
44. SHIINO, M.: Dynamical behavior of stochastic systems of infinitely many coupled nonlinear oscillators exhibiting phase transitions of mean-field type: H theorem on asymptotic approach to equilibrium and critical slowing down of order-parameter fluctuations. *Phys. Rev. A* **36**, 2393–2412, 1987
45. TAMURA, Y.: On asymptotic behaviors of the solution of a nonlinear diffusion equation. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **31**(1), 195–221, 1984
46. WHITT, W.: Proofs of the martingale FCLT. *Probab. Surv.* **4**, 268–302, 2007

MATIAS G. DELGADINO
Mathematical Institute,
University of Oxford,
Oxford
UK.

e-mail: matias.delgadino@maths.ox.ac.uk

and

MATIAS G. DELGADINO
Department of Mathematics,
Pontifical Catholic University of Rio de Janeiro,
Rio de Janeiro
Brazil.

and

RISHABH S. GVALANI
Max-Planck-Institut für Mathematik in den Naturwissenschaften,
Leipzig
Germany.

e-mail: gvalani@mis.mpg.de

and

GRIGORIOS A. PAVLIOTIS
Department of Mathematics,
Imperial College London,
London
UK.

e-mail: g.pavliotis@imperial.ac.uk

(Received January 20, 2020 / Accepted March 23, 2021)

© The Author(s) (2021)