

# Singular Limits for Plasmas with Thermalised Electrons

Megan Griffin-Pickering <sup>\*</sup>      Mikaela Iacobelli <sup>†‡</sup>

July 18, 2019

## Abstract

This work is concerned with the study of singular limits for the Vlasov-Poisson system in the case of massless electrons (VPME), which is a kinetic system modelling the ions in a plasma. Our objective is threefold: first, we provide a mean field derivation of the VPME system in dimensions  $d = 2, 3$  from a system of  $N$  extended charges. Secondly, we prove a rigorous quasineutral limit for initial data that are perturbations of analytic data, deriving the Kinetic Isothermal Euler (KIE) system from the VPME system in dimensions  $d = 2, 3$ . Lastly, we combine these two singular limits in order to show how to obtain the KIE system from an underlying particle system.

## Résumé

Ces travaux portent sur l'étude des limites singulières du système de Vlasov-Poisson dans le cas des électrons sans masse (VPME), qui est un système cinétique de modélisation des ions dans un plasma. Notre objectif est triple : premièrement, nous fournissons une dérivation du champ moyen du système VPME dans les dimensions  $d = 2, 3$  à partir d'un système de  $N$  charges étendues. Deuxièmement, nous prouvons une limite quasi neutre rigoureuse pour les données initiales qui sont des perturbations des données analytiques, dérivant le système Kinetic Isothermal Euler (KIE) du système VPME dans les dimensions  $d = 2, 3$ . Enfin, nous combinons ces deux limites singulières afin de montrer comment obtenir le système KIE à partir d'un système de particules sous-jacent.

*Keywords:* Plasma; Vlasov-Poisson for massless electrons; Mean-field derivation; Quasineutral limit; Kinetic isothermal Euler system

*MSC Codes:* 35Q83, 82B40, 82D10, 35Q35, 35Q70, 35Q82

## 1 Introduction

In this article, we will study a hierarchy of models for plasma. A plasma forms when a neutral gas undergoes a process of dissociation, so that some of the gas particles split into electrons and positively charged ions. These charged particles can interact with each other through the electromagnetic fields they generate. This long-range effect is the dominant form of interaction within the plasma.

As a mathematical idealisation, we can model a plasma as a system of point particles. We have two distinguished species of particle - ions and electrons. The ions have a much higher

---

<sup>\*</sup>University of Cambridge, DPMMS Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK. Email: [m.griffin-pickering@maths.cam.ac.uk](mailto:m.griffin-pickering@maths.cam.ac.uk)

<sup>†</sup>Durham University, Department of Mathematical Sciences, Lower Mountjoy, Stockton Road, Durham DH1 3LE, UK.

<sup>‡</sup>ETH Zürich, Department of Mathematics, Rämistrasse 101, 8092 Zürich, Switzerland. Email: [mikaela.iacobelli@math.ethz.ch](mailto:mikaela.iacobelli@math.ethz.ch)

mass than electrons, which leads to a separation of timescales between the two species. The electrons typically evolve on a much faster timescale than the ions. This makes it reasonable to consider the two species separately.

Suppose first that we would like to model the electrons. In this case, it is usual to take the ions to be fixed in a spatially uniform distribution. This is a reasonable assumption since the ions evolve on a much slower timescale than the electrons. A well-known model for this situation is the *classical Vlasov-Poisson system*:

$$(VP) := \begin{cases} \partial_t f + v \cdot \nabla_x f + E \cdot \nabla_v f = 0, \\ E = -\nabla U, \\ \Delta U = 1 - \int_{\mathbb{R}^d} f \, dv = 1 - \rho, \\ f|_{t=0} = f_0 \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f_0 \, dx \, dv = 1. \end{cases} \quad (1.1)$$

Here we are using a statistical description of the electrons, through a time dependent (probabilistic) density function  $f_t(x, v)$  on the phase space  $\mathbb{T}^d \times \mathbb{R}^d$ . We note that VP is an appropriate model in the electrostatic approximation when magnetic effects are negligible.

Here, we are interested in the reverse case: we wish instead to model the ions in the plasma. To do this, we need to account for the behaviour of the electrons. As we have explained, the electrons are much lighter than the ions and so evolve much more quickly. This means that they undergo collisions at a much faster rate, and hence rapidly approach thermodynamic equilibrium. We will therefore assume that the electrons are thermalised. They follow a Maxwell-Boltzmann law so that their spatial density is given by  $e^U$ , where  $U$  is the electrostatic potential in the plasma. The corresponding kinetic equation is the *Vlasov-Poisson system with massless electrons* (VPME):

$$(VPME) := \begin{cases} \partial_t f + v \cdot \nabla_x f + E \cdot \nabla_v f = 0, \\ E = -\nabla U, \\ \Delta U = e^U - \int_{\mathbb{R}^d} f \, dv = e^U - \rho, \\ f|_{t=0} = f_0 \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f_0 \, dx \, dv = 1. \end{cases} \quad (1.2)$$

The assumption of thermalised electrons is widely used in plasma physics (see for example [33]). VPME has appeared in the physics literature in, for instance, numerical studies of the formation of ion-acoustic shocks [37, 43] and the development of phase-space vortices behind such shocks [7], as well as in studies of the expansion of plasma into vacuum [38]. For a more physically oriented introduction to the model (1.2), see [24].

The nonlinearity in the Poisson equation is the key difference between the classical and VPME systems, and a source of additional mathematical richness. The VPME system has been studied less widely than the classical system, due to the additional difficulties created by this nonlinear coupling. In particular, while the global well-posedness in two and three dimensions is well understood in the classical case (see for example [46], [41],[44], [5], [34]), this problem remained completely open for massless electrons until very recently. Indeed, at least to our knowledge, the only general result on this model in dimension  $d = 2, 3$  was the global existence of weak solutions shown by Bouchut in [9]. In a very recent paper [20] we filled this gap by proving uniqueness for VPME in the class of solutions with bounded density, and global existence of solutions with bounded density for a general class of initial data, generalising to this setting all the previous results known for VP. See also the recent paper by Bardos-Golse-Nguyen-Sentis [3] for some related results on a variant of VPME.

Kinetic equations such as VP and VPME provide a medium scale or ‘mesoscopic’ description of physical systems. They lie between a microscopic description, in which one tracks the dynamics of individual particles, and a coarser macroscopic description. Our main goal in this

work is to study two types of limit that connect the VPME system (1.2) to related macroscopic and microscopic models: the *quasineutral limit* and the *mean field limit*. The limits we study were previously proved rigorously in particular regimes for the classical system (1.1). Here, we will prove analogous results for the VPME system (1.2). We will first derive the VPME system in dimensions  $d = 2, 3$  from a system of  $N$  extended charges, with assumptions on the choice of parameters that are the same as the ones found by Lazarovici for the classical VP system in [31]. Then we will show the validity of the quasineutral limit for initial data that are perturbations of analytic data, deriving the Kinetic Isothermal Euler (KIE) system from the VPME system in dimensions  $d = 2, 3$ . Finally, we will combine these two singular limits in order to recover the KIE system from the underlying particle system, in analogy with the results available for the VP system. To conclude our analysis we will show that the regime of parameters that allows the mean field and the quasineutral limit to “commute” is compatible with a “large” set of admissible configurations.

**The quasineutral limit** A plasma has a characteristic scale, known as the Debye length, that describes the scale of electrostatic interaction within the plasma. For instance, it characterises the scale of charge separation between electrons and ions. We introduce a small parameter  $\varepsilon := \frac{\lambda_D}{\Lambda}$ , where  $\lambda_D$  is the Debye length and  $\Lambda$  is the scale of observation. Under an appropriate scaling, the VPME system becomes

$$(VPME)_\varepsilon := \begin{cases} \partial_t f_\varepsilon + v \cdot \nabla_x f_\varepsilon + E_\varepsilon \cdot \nabla_v f_\varepsilon = 0, \\ E_\varepsilon = -\nabla U_\varepsilon, \\ \varepsilon^2 \Delta U_\varepsilon = e^{U_\varepsilon} - \int_{\mathbb{R}^d} f_\varepsilon dv = e^{U_\varepsilon} - \rho_\varepsilon, \\ f_\varepsilon|_{t=0} = f_{0,\varepsilon} \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f_{0,\varepsilon} dx dv = 1. \end{cases} \quad (1.3)$$

In real plasmas, the Debye length is typically very small. In this case, the plasma is called *quasineutral*, in reference to the fact that the plasma appears to be neutral overall at the observation scale. Quasineutrality is so common a property that it is included in some references in the very definition of plasma (e.g. [12]). In the plasma physics literature it is quite common to assume as an approximation that  $\varepsilon$  is effectively equal to zero. This has the advantage of simplifying the models. For instance, the presence of a small parameter introduces a multi-scale structure which can make numerical simulation computationally expensive, and it may therefore be preferable to eliminate  $\varepsilon$  if possible. Setting  $\varepsilon = 0$  in (1.3), we obtain the *kinetic isothermal Euler* system:

$$(KIE) := \begin{cases} \partial_t f + v \cdot \nabla_x f + E \cdot \nabla_v f = 0, \\ E = -\nabla U, \\ U = \log \rho, \\ f_0 \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f_0 dx dv = 1. \end{cases} \quad (1.4)$$

The system (1.4) was considered in a plasma physics context in [22, 23, 24]. A related equation, derived in the quasineutral limit from a version of (1.3) in which the electron density  $e^U$  is approximated by the linearisation  $1+U$ , was named *Vlasov-Dirac-Benney* by Bardos and studied in [2] and [4]. In the classical case, where  $E$  is replaced by a pressure term which is a Lagrange multiplier corresponding to the constraint  $\rho = 1$ , Bossy-Fontbona-Jabin-Jabir [8] showed local-in-time existence of analytic solutions in the one-dimensional case. Global existence of weak solutions is not known for (1.4).

Formal identification of the limiting system does not answer the question of whether (1.4) is indeed a good approximation for (1.3) when  $\varepsilon$  is small but non-zero. For this reason it is interesting to study the *quasineutral limit*, in which  $\varepsilon \rightarrow 0$  in (1.3). Indeed, Medvedev [38] describes a physical situation in which the quasineutral approximation  $U = \log \rho$  is not valid

everywhere, which provides some motivation for a study of the transition between (1.3) and (1.4).

The rigorous justification of the quasineutral limit is a non-trivial and subtle problem. It has a direct correspondence to a long-time limit for the Vlasov-Poisson system, and is therefore vulnerable to known instability mechanisms inherent to the physical system under consideration. In fact, for the classical system (1.1), it was shown by Hauray and Han-Kwan in [25] that the quasineutral limit is false in general if the initial data are assumed to have only Sobolev regularity.

Rigorous results on the quasineutral limit go back to the works of Brenier-Grenier [11] and Grenier [18] for the classical system (1.1). A result of particular relevance for our purposes is the work of Grenier [19], proving the limit for the classical system assuming uniformly analytic data. The works of Han-Kwan-Iacobelli [26, 27] extended this result to data that are very small, but possibly rough, perturbations of the uniformly analytic case, in dimension 1, 2 and 3.

In the massless electrons case, Han-Kwan-Iacobelli [27] showed a rigorous limit in dimension one, again for rough perturbations of analytic data, while Han-Kwan-Rousset [28] consider Penrose-stable data with sufficiently high Sobolev regularity. In this work, we extend the results of [27] to higher dimensions by showing a rigorous quasineutral limit for the VPME system (1.3) in dimension 2 and 3, for data that are very small, but possibly rough, perturbations of some uniformly analytic functions.

**The mean field limit** The second type of limit we will consider is the *mean field limit*. This refers to the general problem of deriving a Vlasov equation, such as the Vlasov-Poisson system, from an underlying microscopic particle system. In a typical formulation of this problem, one considers a system of  $N$  point particles evolving under the influence of binary interactions between the particles, described by an interaction force  $\nabla\Phi$  derived from a potential  $\Phi$ , and possibly an external force  $\nabla\Psi$  arising from a potential  $\Psi$ . The dynamics are modelled by a system of ODEs describing the phase space position  $(X_i, V_i)_{i=1}^N$  of each of the  $N$  particles:

$$\begin{cases} \dot{X}_i = V_i \\ \dot{V}_i = \frac{1}{N} \sum_{j \neq i} \nabla_x \Phi(X_i - X_j) + \nabla \Psi(X_i). \end{cases} \quad (1.5)$$

The factor  $1/N$  appears due to a choice of scaling, designed to be the appropriate one to obtain a Vlasov equation in the limit.

The connection to a Vlasov equation is formulated via the *empirical measure*  $\mu^N$  associated to (1.5), defined by:

$$\mu^N(t) := \frac{1}{N} \sum_{i=1}^N \delta_{(X_i(t), V_i(t))}. \quad (1.6)$$

If  $\mu^N$  converges to some measure  $f$  as  $N$  tends to infinity, then we expect  $f$  to be a solution of the associated *Vlasov equation*

$$\partial_t f + v \cdot \nabla_x f + (\nabla \Phi *_x \rho_f + \nabla \Psi) \cdot \nabla_v f = 0. \quad (1.7)$$

By ‘deriving (1.7) from the particle system (1.5)’, we mean showing rigorously that  $\mu^N$  converges to  $f$  in the sense of measures, where  $f$  is a solution of (1.7), assuming convergence of the initial data.

Both the VP and the VPME systems fit into this general framework. We can see this by introducing the Green function for the Laplacian on the torus. This is a function  $G$  satisfying

$$\Delta G = \delta_0 - 1. \quad (1.8)$$

We also define the Coulomb kernel  $K = \nabla G$ . More explicitly,  $G$  may be written in the form

$$G(x) = \begin{cases} \frac{1}{2\pi} \log |x| + G_0(x) & d = 2 \\ -\frac{1}{4\pi|x|} + G_0(x) & d = 3, \end{cases}$$

for some smooth function  $G_0$ . For a proof of this representation, see [45] or [21, Lemma 2.1].

In the case of the classical VP system (1.1), we have the representation

$$E = K * \rho.$$

Thus (1.1) is of the form (1.7), where we choose the kernel  $\Phi = G$ . Of course, in this case,  $\nabla\Phi = K$  has a very strong singularity at the origin, of order  $|x|^{-(d-1)}$ . This singularity is a key source of difficulty for the mathematical study of the Vlasov-Poisson system, particularly for the mean field limit.

Similarly, for the VPME system, we can represent the force in the form

$$E = K * \rho - K * e^U.$$

We can think of the VPME system as being of the form (1.7) by taking  $\Phi = G$  and an ‘external’ potential  $\Psi = G * e^U$ . Of course  $\Psi$  is not truly an external potential because  $U$  depends nonlinearly on  $f$ .

Early works on the mathematical justification of the mean field limit include, among others, Braun-Hepp [10] and Neunzert-Wick [40]. Dobrushin [13] proved a rigorous limit for Lipschitz force fields ( $\nabla\Phi, \nabla\Psi \in W^{1,\infty}$ ). The approach relies on the observation that the empirical measure  $\mu^N$  is a weak solution of the Vlasov equation (1.7), if the forces are sufficiently regular. Dobrushin proved the mean field limit by showing a stability result for (1.7) in the class of measure solutions, using an estimate in a Wasserstein distance. The paper [13] is a direct ancestor of much of the modern work on the subject. See the reviews of Golse [17] and Neunzert [39] for further background and literature.

In many physical systems, the force is described by an inverse power law  $|\nabla\Phi| \sim |x|^{-\alpha}$ . In this case, the singularity at the origin prevents the application of Dobrushin’s results. Notice that the VP and VPME systems correspond to the strongly singular case  $\alpha = d - 1$ , and no stability results are known in these cases. Hence, to deal with the mean field limit for VMPE, we will consider a suitable regularisation of the microscopic particle system.

There have been several works aimed at deriving Vlasov equations with singular forces from regularised particle systems. For instance, Hauray-Jabin [29] considered a truncation method in which the force is cut off below a certain distance from the origin  $r_N$ , dependent on the number of particles  $N$ . They showed that the mean field limit holds for a large set of initial configurations, for inverse power law forces with  $\alpha < d - 1$  (in particular not the Vlasov-Poisson case), from a particle system with force truncated at  $r_N$ , provided that  $r_N$  converges to zero sufficiently slowly as  $N$  tends to infinity. They also proved a true mean field limit, without truncation, for the case of ‘weakly singular’ forces in which  $\alpha < 1$ . More recently, Lazarovici-Pickl [32] achieved a similar result for the classical Vlasov-Poisson case  $\alpha = d - 1$  with this type of truncation, with  $r_N \sim N^{-\frac{1}{d}+\eta}$  for any  $\eta > 0$ . They use a law of large numbers approach to compare the mean field force from the particle system to the limiting force. Their results show that there exists a large set of initial configurations for which the mean field limit holds, but it is not possible to identify them from the initial configurations alone, since the argument relies on a law of large numbers throughout the evolution. In a different direction, Lazarovici [31] considered the alternative method of regularisation by convolution. In this approach, the point particles are replaced by delocalised packets of charge, with some smooth, compactly supported

shape  $\chi$ , fixed throughout the evolution. For the classical Vlasov-Poisson case, this results in the particle system

$$\begin{cases} \dot{X}_i = V_i \\ \dot{V}_i = \frac{1}{N} \sum_{j \neq i} [\chi *_x \nabla_x G *_x \chi](X_i - X_j). \end{cases}$$

The shape is then allowed to depend on a regularisation parameter  $r$  by taking  $\chi_r(x) := r^{-d} \chi\left(\frac{x}{r}\right)$ . Lazarovici showed that a mean field limit holds with high probability, provided that  $r_N \geq CN^{-\frac{1}{d(d+2)} + \eta}$  for some  $\eta > 0$ . The admissible configurations are identified by a condition on the initial configuration alone. The appearance of the double regularisation  $\chi_r *_x \nabla_x G *_x \chi_r$  is very important in this analysis. This type of regularisation was previously considered by Horst [30] in the Vlasov-Maxwell case and later used by Rein [42]; a version also appears in Bouchut [9]. It has the advantage that the microscopic dynamics correspond to a Hamiltonian system, for which the corresponding energy converges as  $r$  tends to zero to the energy of the true Vlasov-Poisson system. In this paper, we will prove a regularised mean field limit of this type for VPME that is, to the best of our knowledge, the first derivation of this system from the particle dynamics. As noted in Remark 2.7 we recover the same assumptions on  $r$  as the ones obtained by Lazarovici for the VP system in [31]. Furthermore, as already mentioned, we will prove a series of quantitative estimates that allow us to relate the mean field limit to the quasineutral one.

## 2 Structure of the paper and main results

In this section, we will state our main results. In section 2.1 we start by giving the basic definitions needed for the statements. In sections 2.2, 2.3, 2.4 we state, respectively, Theorem 2.2 about the convergence in the quasineutral limit, Theorem 2.4 concerning the mean field limit, and Theorem 2.5 where we combine these two regimes. Moreover in section 2.5 we state Theorems 2.6 and 2.8 where we prove typicality results that complete our analysis.

In Section 3, we summarise some basic information on the VPME system which is fundamental to the rest of the paper. We review results on the existence and uniqueness of solutions. We also describe the natural energy functional associated with (1.3) and explain its importance.

Section 4 is a key toolbox of the paper. In this section, we obtain regularity and stability estimates on the electric field. The remainder of the paper rests on these estimates. The methods in this section were established in [20]; in this paper, the focus is on quantifying the dependence of the estimates with respect to the Debye length  $\varepsilon$ .

In Sections 5-7, we establish the quasineutral limit for rough data (Theorem 2.2). In Section 5 we prove a strong-strong stability estimate à la Loeper between solutions of the VPME system, quantified in the  $W_2$  distance. This will be a crucial step towards the proof of the quasineutral limit in our perturbative setting, because it enables us to show that solutions with data close in  $W_2$  remain close at later times.

The stability estimate in Section 5 requires an  $L^\infty$  bound on the mass density associated to the solutions of the VPME system. To obtain such a bound, in Section 6 we study how the support of solutions of the VPME system grow over time. This is achieved through estimates on the characteristic flow, based on the regularity estimates on the electric field obtained in Section 4. We will need to develop two different proofs for the two and three dimensional case, respectively in Sections 6.1 and 6.2.

In Section 7, we put together the results of Sections 5 and 6 to complete the proof of the quasineutral limit with rough data.

In Section 8, we prove the regularised mean field limit for VPME (Theorem 2.4). We make use of the electric field estimates from Section 4 to adapt an approach used by Lazarovici for

the classical Vlasov–Poisson system [31] to the case of massless electrons.

In Section 9, we turn to the proof of the derivation of the KIE system (1.4) via a joint mean field and quasineutral limit. To do this, we combine the mean field limit theory from Section 8 with the quasineutral limit from Section 7. The mass density estimates proved in Section 6 are also needed.

Finally, in Section 10 we address the issue of whether the assumptions in our particle derivation results are reasonable, by proving Theorems 2.6 and 2.8.

## 2.1 Preliminary definitions

We begin by introducing some important quantities and technical tools needed to state our results.

**Energy:** We introduce the energy associated to the Vlasov-Poisson system for massless electrons (1.2). In quasineutral scaling, it is given by the functional

$$\mathcal{E}_\varepsilon[f_\varepsilon] := \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f_\varepsilon \, dx \, dv + \frac{\varepsilon^2}{2} \int_{\mathbb{T}^d} |\nabla U_\varepsilon|^2 \, dx + \int_{\mathbb{T}^d} U_\varepsilon e^{U_\varepsilon} \, dx. \quad (2.1)$$

Here, as before,  $U_\varepsilon$  denotes the electrostatic potential induced by the ion distribution  $f_\varepsilon$  in the case of massless electrons. That is,  $U_\varepsilon$  satisfies

$$\varepsilon^2 \Delta U_\varepsilon = e^{U_\varepsilon} - \rho[f_\varepsilon],$$

where  $\rho[f_\varepsilon]$  denotes the mass density associated to  $f_\varepsilon$  - see (2.2).

**Wasserstein distances:** Many of our results involve controlling the distance between solutions to some PDEs, for instance the VPME or KIE systems. For the equations considered in this article, it is often useful to look at solutions in the class of measures. A very natural way to quantify the distance between two such solutions is given by the Wasserstein distances, a particular family of metrics on measures. For our purposes it suffices to consider probability measures. To define the Wasserstein distances, we first need to introduce the notion of a *coupling* of two measures. If  $(\Omega, \mathcal{F})$  is a measurable space, and  $\mu$  and  $\nu$  are probability measures on this space, then a coupling of  $\mu$  and  $\nu$  is a measure  $\pi$  on the product space  $\Omega \times \Omega$  from which  $\mu$  and  $\nu$  may be recovered as the marginals of  $\pi$ . More precisely, this means that for any  $A \in \mathcal{F}$ , we have the two identities

$$\pi(A \times \Omega) = \mu(A), \quad \pi(\Omega \times A) = \nu(A).$$

We will use the notation  $\mathcal{P}(\Omega)$  to denote the space of probability measures on  $\Omega$ , and  $\Pi(\mu, \nu)$  to denote the set of possible couplings of  $\mu$  and  $\nu$ . A Wasserstein distance between  $\mu$  and  $\nu$  is constructed by optimising a particular cost functional over the set  $\Pi(\mu, \nu)$ .

**Definition 2.1** (Wasserstein distances). *Let  $(\Omega, d)$  be a Polish space with metric  $d$  and let  $\mathcal{F}$  denote its Borel  $\sigma$ -algebra. Let  $p \in [1, \infty)$ . The Wasserstein distance of order  $p$ , denoted  $W_p$ , is defined by*

$$W_p^p(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{\Omega \times \Omega} d(x, y)^p \, d\pi(x, y),$$

for all  $\mu, \nu \in \mathcal{P}(\Omega)$  such that the right hand side is finite. In particular this is well-defined for  $\mu, \nu \in \mathcal{P}_p$ , where  $\mathcal{P}_p$  denotes the set of probability measures  $\gamma$  for which

$$\int_{\Omega} d(x, x_0)^p \, d\gamma(x) < \infty,$$

for some  $x_0 \in \Omega$ .

In this work we will be using the cases  $\Omega = \mathbb{T}^d \times \mathbb{R}^d$  and  $p = 1, 2$ .

**Analytic norms:** For our results on the quasineutral limit, we will work with solutions that are close to being analytic, up to a small perturbation that may have much lower regularity. To measure analyticity we will use the norms  $\|\cdot\|_{B_\delta}$ , defined for  $\delta > 1$  by

$$\|g\|_{B_\delta} := \sum_{k \in \mathbb{Z}^d} |\hat{g}(k)| \delta^{|k|},$$

where  $\hat{g}(k)$  denotes the Fourier coefficient of  $g$  of index  $k$ . This type of analytic norm was notably used in the context of quasineutral limits by Grenier [19]. If  $\|g\|_{B_\delta}$  is finite, then the Fourier coefficients of  $g$  decay at an exponential rate with respect to  $k$ . This rate is quantified by the (large) parameter  $\delta > 1$ .

**Notation:** Throughout the paper, we shall use both  $\rho[f]$  and  $\rho_f$  to denote the density associated to  $f$ , namely

$$\rho[f](t, x) = \rho_f(t, x) := \int_{\mathbb{R}^d} f(t, x, v) dv. \quad (2.2)$$

Later on, we will use the notation  $\overline{\exp}_n$  to denote the  $n$ -fold iteration of the exponential function. For example

$$\overline{\exp}_3(x) := \exp \exp \exp(x).$$

In our proofs,  $C$  will denote a positive constant independent of the relevant parameters (i.e.,  $N, \varepsilon, r$ , depending on the context) and that may change from line to line. Subscripts may be used to indicate a constant depending on a given parameter - for instance,  $C_N$  denotes a constant depending on  $N$  but independent of other parameters.

## 2.2 Quasineutral limit

Our first result is a quasineutral limit for initial data that are small but rough perturbations of uniformly analytic functions. This extends the one dimensional result in [27] to higher dimensions, and generalises the results in [26] from the classical VP system to the massless electron case. With respect to the arguments in [26, 27], here we need to face several new challenges. Indeed, at the same time we need to deal with the nonlinear coupling in the Poisson equation (in [27] a solution to this issue was found only for  $d = 1$ ) and with the fact that weaker regularity estimates are available for the Poisson equation (in [26] this was solved in the case of the classical linear Poisson equation). As already discussed in [26, 27], we note that the exponential smallness (in  $\varepsilon$ ) of the perturbation is necessary. More precisely, the exponential smallness of the perturbation is already necessary at the ‘‘easier’’ level of the classical Vlasov-Poisson system [25]. In our case, the presence of an exponential nonlinearity naturally give rise to an extra exponential term, as already noted in the one-dimensional case [27]. With respect to the 1d case, the dependence on  $\varepsilon$  that we obtain here is worse, but this is what one usually expects in higher dimensions due to the much more singular character of the Poisson equation. Although we do not believe that our dependence on  $\varepsilon$  is sharp, it seems likely to us that one cannot go lower than a double exponential (see also Remark 2.9).

Let us observe that we work with compactly supported data in order to get control of the mass density of the solution. However we can allow the size of the support to grow at a controlled rate, exponential in  $\varepsilon^{-1}$ . These assumptions are chosen to match the results obtained in [26]. Although we believe that the hypothesis on the support may be slightly weakened to include densities that decay exponentially fast in velocity, achieving such extension here would



go completely beyond the scope of this paper. The interested reader is referred to the papers [26, 27] for a discussion about possible initial data that satisfy our assumptions.

**Theorem 2.2** (Quasineutral limit). *Let  $d = 2, 3$ . Consider initial data  $f_\varepsilon(0)$  satisfying the following conditions:*

- (Uniform bounds)  $f_\varepsilon(0)$  is bounded and has bounded energy, uniformly with respect to  $\varepsilon$ :

$$\|f_\varepsilon(0)\|_{L^\infty(\mathbb{T}^d \times \mathbb{R}^d)} \leq C_0, \quad \mathcal{E}_\varepsilon[f_\varepsilon(0)] \leq C_0, \quad (2.3)$$

for some constant  $C_0 > 0$ .

- (Control of support) There exists  $C_1 > 0$  such that

$$f_\varepsilon(0, x, v) = 0 \quad \text{for } |v| > \exp(C_1 \varepsilon^{-2}). \quad (2.4)$$

- (Perturbation of an analytic function) There exist  $g_\varepsilon(0)$  satisfying, for some  $\delta > 1$ ,  $\eta > 0$ , and  $C > 0$ ,

$$\begin{aligned} \sup_\varepsilon \sup_{v \in \mathbb{R}^d} (1 + |v|^2) \|g_\varepsilon(0, \cdot, v)\|_{B_\delta} &\leq C, \\ \sup_\varepsilon \left\| \int_{\mathbb{R}^d} g_\varepsilon(0, \cdot, v) \, dv - 1 \right\|_{B_\delta} &\leq \eta, \end{aligned}$$

as well as the support condition (2.4), such that

$$W_2(f_\varepsilon(0), g_\varepsilon(0)) \leq [\overline{\exp}_4(C\varepsilon^{-2})]^{-1} \quad (2.5)$$

for  $C$  sufficiently large with respect to  $C_0, C_1$ .

- (Convergence of data)  $g_\varepsilon(0)$  has a limit  $g(0)$  in the sense of distributions as  $\varepsilon \rightarrow 0$ .

Let  $f_\varepsilon$  denote the unique solution of (1.3) with bounded density and initial datum  $f_\varepsilon(0)$  (see Section 3.1). Then there exists a time horizon  $T_*$ , independent of  $\varepsilon$  but depending on the collection  $\{g_{0,\varepsilon}\}_\varepsilon$ , and a solution  $g$  of (1.4) on the time interval  $[0, T_*]$  with initial datum  $g(0)$ , such that

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T_*]} W_1(f_\varepsilon(t), g(t)) = 0.$$

**Remark 2.3.** The condition (2.4) should be understood as giving the fastest growth rate on the support for which the inverse quadruple exponential is an admissible rate in (2.5). In particular, this would still be the rate achievable by our methods even if the support of the data was uniform in  $\varepsilon$ .

### 2.3 Mean field limit

As mentioned in the introduction, the validity of a particle system approximation to kinetic equations is a fundamental problem. In Section 8 we will derive the VPME system as the mean field limit of a regularised microscopic particle system. In this paper we will use a regularisation of the kind proposed by Lazarovici [31], however we expect that it would be possible to adapt other similar regularisation methods to this setting.

More precisely, we fix a smooth, compactly supported, radially symmetric function  $\chi$  with unit mass and define

$$\chi_r(x) = r^{-d} \chi\left(\frac{x}{r}\right). \quad (2.6)$$

We then consider a microscopic system describing the dynamics of a system of ‘delocalised ions’ of shape  $\chi_r$ . For  $1 \leq i \leq N$ , let  $(X_i, V_i)$  denote the position and velocity of the centre of the  $i$ th delocalised ion. The system is described by the following system of ODEs:

$$\begin{cases} \dot{X}_i = V_i \\ \dot{V}_i = \frac{1}{N} \sum_{j \neq i} \chi_r * K * \chi_r(X_i - X_j) - \chi_r * K * e^U, \end{cases} \quad (2.7)$$

where  $U$  satisfies

$$\Delta U(x) = e^{U(x)} - \frac{1}{N} \sum_{i=1}^N \chi_r(x - X_i) \quad (2.8)$$

and  $K$  denotes the Coulomb kernel on  $\mathbb{T}^d$ . That is,  $K = \nabla G$  where  $G$  is defined by (1.8). Note that we can rewrite the velocity equation as

$$\dot{V}_i = -\chi_r * \nabla U(t, X_i).$$

and the equation for  $U$  as

$$\Delta U = e^U - \chi_r * \rho_{\mu^N},$$

where  $\mu^N$  denotes the empirical measure as defined in (1.6). This is valid because  $\chi_r * K * \chi_r(0) = 0$ , and so

$$\frac{1}{N} \sum_{j \neq i} \chi_r * K * \chi_r(X_i - X_j) = \frac{1}{N} \sum_{j=1}^N \chi_r * K * \chi_r(X_i - X_j) = \chi_r * K * \chi_r * \rho_{\mu^N}(X_i).$$

Indeed, since  $K$  is odd and  $\chi_r$  is even, we have

$$\begin{aligned} \chi_r * K * \chi_r(0) &= \int_{\mathbb{T}^d \times \mathbb{T}^d} \chi_r(-x) K(x - y) \chi_r(y) \, dx \, dy \\ &= \int_{\mathbb{T}^d \times \mathbb{T}^d} \chi_r(-y) K(y - x) \chi_r(x) \, dx \, dy \\ &= - \int_{\mathbb{T}^d \times \mathbb{T}^d} \chi_r(y) K(x - y) \chi_r(-x) \, dx \, dy \\ &= -\chi_r * K * \chi_r(0). \end{aligned}$$

The system (2.7)-(2.8) is of the form (1.5), but where the ‘external’ potential  $\Psi = \chi_r * G * e^U$  in fact depends nonlinearly on the solution of the system, via the Poisson equation (2.8).

The next result shows a general statement about the validity of the mean field limit for the regularised particle system. We shall discuss later, in the context of Theorem 2.6, assumptions on the regularisation parameter  $r$  for which one can find initial data to which this result applies.

**Theorem 2.4** (Regularised mean field limit). *Let  $d = 2, 3$ , and let  $f(0)$  be a choice of initial datum for (1.2) such that there exist a solution  $f$  of (1.2) and, for each  $r$ , a solution  $f_r$  of (8.1) such that on some time interval  $[0, T_*]$ , all these solutions have bounded density: for some  $M > 0$ ,*

$$\|\rho_f\|_{L^\infty([0, T_*] \times \mathbb{T}^d)}, \sup_r \|\rho_{f_r}\|_{L^\infty([0, T_*] \times \mathbb{T}^d)} \leq M.$$

*Assume that  $r = r_N$  and the initial configurations for (2.7) are chosen such that the corresponding empirical measures satisfy, for some constant  $C_{M, T_*}$ , sufficiently large depending on  $M$  and  $T_*$ ,*

$$\limsup_{r \rightarrow 0} \frac{W_2^2(f(0), \mu_r^N(0))}{r^{d+2+C_{M, T_*} |\log r|^{-1/2}}} < 1.$$

Then the empirical measure associated to the particle system dynamics starting from this configuration converges to  $f$ :

$$\lim_{r \rightarrow 0} \sup_{t \in [0, T_*]} W_2(f(t), \mu_r^N(t)) = 0.$$

## 2.4 Combined mean field and quasineutral limit

In the previous section, we considered the problem of deriving VPME (1.2) from an underlying microscopic particle system. It is natural to consider similarly whether KIE (1.4) can be derived from a particle system. It is clear that to obtain such a limit we will need to use a different scaling of the particle system from the mean field scaling  $\frac{1}{N}$ . A possible approach to this problem is to consider combining the mean field and quasineutral limits for VPME. This amounts to choosing a scaling of the form  $\frac{\varepsilon^{-2}}{N}$  and looking for a relationship  $\varepsilon = \varepsilon(N)$  which it makes it possible to take both limits simultaneously. Since  $\varepsilon$  can be related to physical constants of the plasma such as the temperature, a successful result could be interpreted as describing a physical regime in which the KIE model can be validated on the basis of the particle dynamics. In [21], the authors considered a strategy of this kind in order to derive a variant of (1.4) which corresponds to the classical VP system. Here, we extend this result to the case of massless electrons. The considerations regarding the singularity of the Coulomb force also apply to this problem. We therefore consider the following regularised system for  $(Z_i)_{i=1}^N = (X_i, V_i)_{i=1}^N \in (\mathbb{R}^{2d})^N$ :

$$\begin{cases} \dot{X}_i = V_i \\ \dot{V}_i = \frac{\varepsilon^{-2}}{N} \sum_{j \neq i} \chi_r * K * \chi_r(X_i - X_j) - \varepsilon^{-2} \chi_r * K * e^{U_\varepsilon}, \end{cases} \quad (2.9)$$

where  $U_\varepsilon$  satisfies

$$\varepsilon^2 \Delta U_\varepsilon(x) = e^{U_\varepsilon(x)} - \frac{1}{N} \sum_{i=1}^N \chi_r(x - X_i).$$

The goal is to find a relationship  $\varepsilon = \varepsilon(N), r = r(N)$  between the parameters such that the KIE system (1.4) is obtained from (2.9) in the limit as  $N$  tends to infinity. For this system, we can prove the following limit.

**Theorem 2.5** (From particles to KIE). *Let  $d = 2$  or  $3$ , and let  $f_\varepsilon(0), g_\varepsilon(0)$  and  $g(0)$  satisfy the assumptions of Theorem 2.2. Given  $\varepsilon, r, N$ , let  $(Z_{0,i}^{(\varepsilon,r)})_{i=1}^N \in (\mathbb{R}^{2d})^N$  be a choice of initial data for the regularised and scaled  $N$ -particle ODE system (2.9). Let  $(Z_{t,i}^{(\varepsilon,r)})_{i=1}^N$  be the solution of (2.9) with this initial data and let  $\mu_{\varepsilon,r}^N$  denote the associated empirical measure as defined in (1.6).*

*Let  $T_*$  be the maximal time of convergence from Theorem 2.2. There exists a constant  $C > 0$  depending on  $\{f_\varepsilon(0)\}_\varepsilon$  such that if the parameters  $(r, \varepsilon) = (r(N), \varepsilon(N))$  and the initial data  $\mu_{\varepsilon,r}^N(0)$  satisfy*

$$r \leq [\overline{\exp}_3(C\varepsilon^{-2})]^{-1}, \quad \lim_{N \rightarrow \infty} \frac{W_2(\mu_{\varepsilon,r}^N(0), f_\varepsilon(0))}{r^{(d+2+\eta)/2}} = 0,$$

*for some  $\eta > 0$ , then*

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T_*]} W_1(\mu_{\varepsilon,r}^N(t), g(t)) = 0,$$

*where  $g$  is a solution of the KIE system (1.4) with initial data  $g(0)$  on the time interval  $[0, T_*]$ .*

## 2.5 Existence of admissible configurations

Note that Theorems 2.4 and 2.5 hold only for special initial configurations for the particle system, i.e. those that converge sufficiently quickly to a given measure as  $N$  tends to infinity. This raises the natural question of whether configurations satisfying this rate of convergence exist. The goal of the results in this section is to identify ranges of the parameters  $r, N, \varepsilon$  for which suitable configurations exist. In fact we can show that there is, in some sense, a ‘large’ set of admissible configurations.

One way to approximate a fixed measure  $\nu$  by an empirical measure is to choose the points  $(Z_i)_{i=1}^N$  defining the empirical measure by drawing independent samples from  $\nu$ . That is,  $(Z_i)_{i=1}^N$  should have joint law  $\nu^{\otimes N}$ . This produces a random empirical measure  $\nu^N$ . Then, a law of large numbers result shows that  $\nu^N$  converges to  $\nu$  almost surely as  $N$  tends to infinity. In Theorems 2.6 and 2.8 we show that, for certain regimes of the parameters, this method of constructing the initial configurations will provide admissible configurations for Theorems 2.4 and 2.5 (respectively) with probability 1.

**Theorem 2.6** (Typicality for mean field limit with  $\varepsilon = 1$ ). *Let  $d = 2$  or  $3$ , and let  $f_0$  be a choice of initial datum for (1.2) satisfying the assumptions of Theorem 2.4 and having a finite  $k$ th moment for some  $k > 4$ :*

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} (|x|^k + |v|^k) f_0(dx dv) < +\infty.$$

Let  $r = cN^{-\gamma}$  for some  $\gamma$  satisfying

$$\gamma < \frac{1}{d+2} \min \left\{ \frac{1}{d}, 1 - \frac{4}{k} \right\}. \quad (2.10)$$

For each  $N$ , select initial configurations for the regularised  $N$ -particle system (2.7) by taking  $N$  independent samples from  $f_0$ . Then with probability 1, this gives an admissible set of configurations for Theorem 2.4, i.e. the regularised mean field limit holds:

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T_*]} W_2(\mu_r^N(t), f(t)) = 0.$$

**Remark 2.7.** *It is worth noticing that our assumptions on  $r$  are the same as the ones found by Lazarovici for the classical VP system in [31].*

**Theorem 2.8** (Typicality with quasineutral scaling). *Let  $d = 2$  or  $3$ , and  $f_\varepsilon(0), g_\varepsilon(0)$  and  $g(0)$  satisfy the assumptions of Theorem 2.2. Let  $(r, \varepsilon) = (r(N), \varepsilon(N))$  be chosen to satisfy*

$$r \leq [\overline{\exp}_3(C\varepsilon^{-2})]^{-1}$$

where  $C > 0$  is the constant from Theorem 2.5. Assume that  $r = r(N)$  are chosen such that

$$r = cN^{-\gamma},$$

where  $c > 0$  is an arbitrary constant and  $\gamma$  satisfies

$$0 \leq \gamma < \frac{1}{d(d+2)}. \quad (2.11)$$

For each  $N$ , let the initial configurations for the particle system (2.9) be chosen by taking  $N$  independent samples from  $f_\varepsilon(0)$ . Then, with probability 1, this procedure selects a set of configurations for which the combined mean field and quasineutral limit holds, that is

$$\lim_{N \rightarrow \infty} \sup_{t \in [0, T_*]} W_1(\mu_{\varepsilon, r}^N(t), g(t)) = 0,$$

where  $g$  is the solution of (1.4) with initial datum  $g(0)$  on the time interval  $[0, T_*]$  provided by Theorem 2.2.

**Remark 2.9.** Observe that these assumptions ultimately result in a relationship between  $N$  and  $\varepsilon$  of the form

$$\varepsilon > \frac{C}{\sqrt{\log \log \log N}}, \quad (2.12)$$

so that the Debye length must converge to zero very slowly in relation to  $N$ . This rate is much slower than the one found for the classical case in [21]. This is due to the singular nature of the interaction in the massless electrons case, which leads to the appearance of iterated exponential factors in our estimates for the electrostatic potential. On the other hand, the relation (2.11) is the same as the one found by Lazarovici in [31], where he only considers the mean field limit for the Vlasov-Poisson system. Hence, even if we are dealing with several singular limits at the same time, we do not lose any information (with respect to what is currently known) at the level of the dependence between  $r$  and  $N$ .

**Remark 2.10.** The rate (2.12) reflects more generally the difficulties in deriving macroscopic models from microscopic dynamics under physically realistic conditions. One of the best known examples arises in the study of fluid equations such as the Navier-Stokes system. A well-known strategy for the derivation of Navier-Stokes is to use the Boltzmann equation as an intervening step between the particle dynamics and the hydrodynamic system - in the same way that, in the present paper, VPME appears as an intervening model between the particle dynamics and the KIE system. In the fully nonlinear case, the program is not yet complete, because the derivation of the Boltzmann equation from a particle system has been proven only on a short timescale. This time of validity converges to zero under the hydrodynamic scaling from Boltzmann to Navier-Stokes, which obstructs the full limit. Results are available for linear versions of this limit, such as the microscopic derivation of the Stokes-Fourier system via the linearised Boltzmann equation [6]. In this latter case, the mean free path (which is the parameter defining the hydrodynamic scaling) is required to converge to zero with a slow (triply logarithmic) rate in the number of particles  $N$ .

## 3 Basic Properties of Solutions of VPME

### 3.1 Existence of Solutions

The existence of solutions for the VPME system (1.2) is provided by the results of [20]. We will work in the class of solutions with bounded density. Such solutions have a uniqueness property, and exist for compactly supported initial data. Moreover, the energy (2.1) is conserved for these solutions. See [20, Theorems 2.3-2.4], which we recall below:

**Theorem 3.1** (Uniqueness). *Let  $d = 2, 3$ . Consider an initial datum  $f_0 \in \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d)$ . Then there exists at most one solution  $f_t \in C([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d))$  of (1.2) with bounded density, i.e. such that the mass density  $\rho_f$  lies in the space  $L_{loc}^\infty([0, \infty); L^\infty(\mathbb{T}^d))$ .*

**Theorem 3.2** (Existence of bounded density solutions). *Let  $d = 2, 3$ . Consider an initial datum  $f_0 \in \mathcal{P} \cap L^1 \cap L^\infty(\mathbb{T}^d \times \mathbb{R}^d)$  with finite energy, namely  $\mathcal{E}[f_0] \leq C_0$  for some constant  $C_0 > 0$  and compact support in  $\mathbb{T}^d \times \mathbb{R}^d$ . Then there exists a global in time solution  $f_t \in C([0, \infty); \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d))$  of (1.2) with initial data  $f_0$ , which satisfies*

$$\|f_t\|_{L^p(\mathbb{T}^d \times \mathbb{R}^d)} \leq \|f_0\|_{L^p(\mathbb{T}^d \times \mathbb{R}^d)}$$

for all  $p \in [1, \infty]$ , and has conserved energy and locally bounded mass:

$$\mathcal{E}[f_t] = \mathcal{E}[f_0], \quad \rho_f \in L_{loc}^\infty([0, \infty); L^\infty(\mathbb{T}^d)).$$

### 3.2 The Role of the Energy

The energy (2.1) is conserved for all time by the bounded density solutions furnished by Theorem 3.2. The significance of this is that control of the energy implies a bound on the  $L^{\frac{d+2}{d}}(\mathbb{T}^d)$  norm of the mass density  $\rho$ . This means that estimates in terms of  $\|\rho\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)}$  are useful because this quantity is controlled uniformly in time. This observation will form a crucial part of our later arguments.

**Lemma 3.3.** (i) Let  $f(x, v)$  satisfy, for some constant  $C_0$ ,

$$\mathcal{E}_\varepsilon[f] \leq C_0.$$

Then there exists a constant  $C$ , depending on  $C_0$  but independent of  $\varepsilon$ , such that

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f(x, v) \, dx \, dv \leq C.$$

(ii) Assume that

$$\|f\|_{L^\infty(\mathbb{T}^d \times \mathbb{R}^d)} \leq C_0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f(x, v) \, dx \, dv \leq C_0.$$

Then the mass density

$$\rho(x) := \int_{\mathbb{R}^d} f(x, v) \, dv$$

lies in  $L^{(d+2)/d}(\mathbb{T}^d)$  with

$$\|\rho\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C,$$

for some constant  $C$  depending on  $C_0$  only.

(iii) In particular, if

$$\|f\|_{L^\infty(\mathbb{T}^d \times \mathbb{R}^d)} \leq C_0, \quad \mathcal{E}_\varepsilon[f] \leq C_0,$$

then there exists a constant  $C$  independent of  $\varepsilon$  such that

$$\|\rho\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C.$$

*Proof.* Observe that  $xe^x \geq -1$ . This implies that

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f \, dx \, dv \leq 2(C_0 + 1).$$

Part (ii) then follows from a standard interpolation argument. See for example [20, Lemma 5.1].  $\square$

## 4 Estimates on the electric field

In this section we prove a series of quantitative regularity estimates on the electric field for the VPME system. A crucial idea is to split the electrostatic potential  $U_\varepsilon$  into two parts:

$$U_\varepsilon = \bar{U}_\varepsilon + \widehat{U}_\varepsilon,$$

where  $\bar{U}_\varepsilon$  and  $\widehat{U}_\varepsilon$  are solutions of the equations

$$\varepsilon^2 \Delta \bar{U}_\varepsilon = 1 - \rho_\varepsilon, \quad \varepsilon^2 \Delta \widehat{U}_\varepsilon = e^{\bar{U}_\varepsilon + \widehat{U}_\varepsilon} - 1.$$

We will use the notation  $\bar{E}_\varepsilon = -\nabla \bar{U}_\varepsilon$  and  $\widehat{E}_\varepsilon = -\nabla \widehat{U}_\varepsilon$ . This approach was used previously in [27] and [20].  $\bar{U}_\varepsilon$  is the singular part of the potential, which behaves like the potential in the classical Vlasov-Poisson system, while  $\widehat{U}_\varepsilon$  is a smoother correction.

### 4.1 Regularity

The following is a version of Proposition 3.1 of [20] with quasineutral scaling. We will give the proof of these estimates below.

**Proposition 4.1** (Regularity estimates on  $\bar{U}_\varepsilon$  and  $\widehat{U}_\varepsilon$ ). *Let  $d = 2, 3$ . Let  $h \in L^\infty \cap L^{(d+2)/d}(\mathbb{T}^d)$ . Then there exist unique  $\bar{U}_\varepsilon \in W^{1,2}(\mathbb{T}^d)$  with zero mean and  $\widehat{U}_\varepsilon \in W^{1,2} \cap L^\infty(\mathbb{T}^d)$  satisfying*

$$\varepsilon^2 \Delta \bar{U}_\varepsilon = 1 - h, \quad \varepsilon^2 \Delta \widehat{U}_\varepsilon = e^{\bar{U}_\varepsilon + \widehat{U}_\varepsilon} - 1.$$

Moreover we have the following estimates for some constant  $C_{\alpha,d} > 0$ :

$$\begin{aligned} \|\bar{U}_\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^d)} &\leq C_{\alpha,d} \varepsilon^{-2} \left( 1 + \|h\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \right), & \alpha \in \begin{cases} (0, 1) & \text{if } d = 2 \\ (0, \frac{1}{5}] & \text{if } d = 3 \end{cases} \\ \|\bar{U}_\varepsilon\|_{C^{1,\alpha}(\mathbb{T}^d)} &\leq C_{\alpha,d} \varepsilon^{-2} \left( 1 + \|h\|_{L^\infty(\mathbb{T}^d)} \right), & \alpha \in (0, 1) \\ \|\widehat{U}_\varepsilon\|_{C^{1,\alpha}(\mathbb{T}^d)} &\leq C_{\alpha,d} \exp \left( C_{\alpha,d} \varepsilon^{-2} \left( 1 + \|h\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \right) \right) & \alpha \in (0, 1) \\ \|\widehat{U}_\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^d)} &\leq C_{\alpha,d} \overline{\exp}_2 \left( C_{\alpha,d} \varepsilon^{-2} \left( 1 + \|h\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \right) \right) & \alpha \in \begin{cases} (0, 1) & \text{if } d = 2 \\ (0, \frac{1}{5}] & \text{if } d = 3. \end{cases} \end{aligned}$$

#### 4.1.1 Regularity of $\bar{U}_\varepsilon$

The singular part of the potential,  $\bar{U}_\varepsilon$ , satisfies a classical Poisson equation on  $\mathbb{T}^d$ :

$$\varepsilon^2 \Delta \bar{U}_\varepsilon = 1 - h. \tag{4.1}$$

The existence of  $\bar{U}_\varepsilon$  for  $h \in L^\infty \cap L^{(d+2)/d}(\mathbb{T}^d)$  is well understood, and there is no loss of generality to assume that  $\bar{U}_\varepsilon$  has zero mean.

The regularity theory of (4.1) is likewise well understood. If  $h \in L^p(\mathbb{T}^d)$ , then by Calderón-Zygmund estimates for the Laplacian [16] we have the estimate

$$\|\bar{U}_\varepsilon\|_{W^{2,p}(\mathbb{T}^d)} \leq C_d \varepsilon^{-2} \left( 1 + \|h\|_{L^p(\mathbb{T}^d)} \right).$$

We can then use Sobolev embedding [14] to deduce higher-order integrability or Hölder regularity for  $\bar{U}_\varepsilon$ . In the case  $p = \frac{d+2}{d}$ , we obtain

$$\|\bar{U}_\varepsilon\|_{C^{0,\alpha}} \leq C_{\alpha,d} \varepsilon^{-2} \left( 1 + \|h\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \right)$$

for  $\alpha \in (0, 1)$  if  $d = 2$ , or  $\alpha \in (0, \frac{1}{5}]$  if  $d = 3$ . In the case  $p = \infty$  we obtain

$$\|\bar{U}_\varepsilon\|_{C^{1,\alpha}(\mathbb{T}^d)} \leq C_{\alpha,d} \varepsilon^{-2} \left(1 + \|h\|_{L^\infty(\mathbb{T}^d)}\right)$$

for all  $\alpha \in (0, 1)$ .

Later we will perform estimates on the trajectories of the flow induced by  $E_\varepsilon$ . In order to do this, we would like to have Lipschitz regularity for  $E_\varepsilon$ . Unfortunately, for  $\bar{E}_\varepsilon$  we are not able to prove Lipschitz regularity under our assumptions. However, we do have a log-Lipschitz regularity estimate. See [36, Lemma 8.1] for the proof in the case where the spatial domain is  $\mathbb{R}^2$ , or [20, Lemma 3.3] for the case  $\mathbb{T}^d$  for general  $d$ .

**Lemma 4.2** (Log-Lipschitz regularity of  $\bar{E}_\varepsilon$ ). *Let  $\bar{U}_\varepsilon$  be a solution of*

$$\varepsilon^2 \Delta \bar{U}_\varepsilon = h$$

for  $h \in L^\infty(\mathbb{T}^d)$ . Then

$$|\nabla \bar{U}_\varepsilon(x) - \nabla \bar{U}_\varepsilon(y)| \leq \varepsilon^{-2} C \|h\|_{L^\infty} |x - y| \left(1 + \log \left(\frac{\sqrt{d}}{|x - y|}\right) \mathbb{1}_{|x - y| \leq \sqrt{d}}\right).$$

#### 4.1.2 Regularity of $\hat{U}_\varepsilon$

In [20, Proposition 3.4], we proved the existence of a unique  $W^{1,2}(\mathbb{T}^d)$  solution  $\hat{U}$  of the equation

$$\Delta \hat{U} = e^{\bar{U} + \hat{U}} - 1,$$

under the assumption that  $\bar{U} \in L^\infty(\mathbb{T}^d) \cap W^{1,2}(\mathbb{T}^d)$ . We showed that this solution in fact belongs to the Hölder space  $C^{1,\alpha}(\mathbb{T}^d)$  for  $\alpha \in (0, 1)$ . Furthermore, if  $\bar{U} \in C^{0,\alpha}(\mathbb{T}^d)$ , then  $\hat{U} \in C^{2,\alpha}(\mathbb{T}^d)$ . The argument in [20, Proposition 3.4] works for general  $\varepsilon$ . We will revisit part of the proof here in order to quantify how the constants depend on  $\varepsilon$ . This dependence will be crucial in our later proofs.

**Lemma 4.3.** *Consider the nonlinear Poisson equation*

$$\varepsilon^2 \Delta \hat{U}_\varepsilon = e^{\bar{U}_\varepsilon + \hat{U}_\varepsilon} - 1. \quad (4.2)$$

Assume that  $\bar{U}_\varepsilon \in L^\infty(\mathbb{T}^d) \cap W^{1,2}(\mathbb{T}^d)$  with

$$\|\bar{U}_\varepsilon\|_{L^\infty(\mathbb{T}^d)} \leq M_1.$$

Then there exists a unique solution  $\hat{U}_\varepsilon \in W^{1,2}(\mathbb{T}^d)$  of (4.2). Moreover,  $\hat{U}_\varepsilon \in C^{1,\alpha}(\mathbb{T}^d)$  for all  $\alpha \in (0, 1)$ , with the estimate

$$\|\hat{U}_\varepsilon\|_{C^{1,\alpha}(\mathbb{T}^d)} \leq C \varepsilon^{-2} (e^{2M_1} + 1). \quad (4.3)$$

If, in addition,  $\bar{U}_\varepsilon$  is Hölder regular with the estimate

$$\|\bar{U}_\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^d)} \leq M_2$$

for some  $\alpha \in (0, 1)$ , then  $\hat{U}_\varepsilon \in C^{2,\alpha}(\mathbb{T}^d)$  with the estimate

$$\|\hat{U}_\varepsilon\|_{C^{1,\alpha}(\mathbb{T}^d)} \leq [M_2 + C \varepsilon^{-2} (e^{2M_1} + 1)] \exp [C \varepsilon^{-2} (e^{2M_1} + 1)]. \quad (4.4)$$



*Proof.* The existence and uniqueness of  $\widehat{U}_\varepsilon \in W^{1,2}(\mathbb{T}^d)$  satisfying (4.2) can be proved by using a Calculus of Variations method, as in [20, Proposition 3.4]. Here we will focus on the regularity estimates (4.3) and (4.4).

First, we want to prove  $L^p(\mathbb{T}^d)$  estimates on  $e^{U_\varepsilon}$ . We will use this to deduce estimates on  $\widehat{U}_\varepsilon$  using regularity theory for the Poisson equation. Since  $\widehat{U}_\varepsilon$  is a weak solution of (4.2), we know that for all test functions  $\phi \in W^{1,2}(\mathbb{T}^d) \cap L^\infty(\mathbb{T}^d)$ ,

$$\int_{\mathbb{T}^d} \varepsilon^2 \nabla \widehat{U}_\varepsilon \cdot \nabla \phi + \left( e^{\bar{U}_\varepsilon + \widehat{U}_\varepsilon} - 1 \right) \phi \, dx = 0.$$

If we formally take  $e^{\widehat{U}_\varepsilon}$  as a test function, we obtain

$$\varepsilon^2 \int_{\mathbb{T}^d} |\nabla \widehat{U}_\varepsilon|^2 e^{\widehat{U}_\varepsilon} \, dx + \int_{\mathbb{T}^d} e^{\bar{U}_\varepsilon} \cdot e^{2\widehat{U}_\varepsilon} \, dx = \int_{\mathbb{T}^d} e^{\widehat{U}_\varepsilon} \, dx.$$

Then

$$\int_{\mathbb{T}^d} e^{\bar{U}_\varepsilon} \cdot e^{2\widehat{U}_\varepsilon} \, dx \leq \int_{\mathbb{T}^d} e^{\widehat{U}_\varepsilon} \, dx.$$

We can make this estimate rigorous by using a truncation argument, as described in [20]. Since

$$\|\bar{U}_\varepsilon\|_{L^\infty(\mathbb{T}^d)} \leq M_1,$$

we have

$$\int_{\mathbb{T}^d} e^{\bar{U}_\varepsilon} \cdot e^{2\widehat{U}_\varepsilon} \, dx \geq e^{-M_1} \|e^{\widehat{U}_\varepsilon}\|_{L^2(\mathbb{T}^d)}^2.$$

Similarly,

$$\|e^{U_\varepsilon}\|_{L^1(\mathbb{T}^d)} = \int_{\mathbb{T}^d} e^{\bar{U}_\varepsilon + \widehat{U}_\varepsilon} \, dx \geq e^{-M_1} \int_{\mathbb{T}^d} e^{\widehat{U}_\varepsilon} \, dx.$$

By (4.2),

$$0 = \varepsilon^2 \int_{\mathbb{T}^d} \Delta \widehat{U}_\varepsilon \, dx = \int_{\mathbb{T}^d} (e^{U_\varepsilon} - 1) \, dx,$$

which implies that

$$\|e^{U_\varepsilon}\|_{L^1(\mathbb{T}^d)} = 1.$$

Therefore,

$$\|e^{\widehat{U}_\varepsilon}\|_{L^2(\mathbb{T}^d)}^2 \leq e^{2M_1}.$$

Similarly, taking  $e^{(n-1)\widehat{U}_\varepsilon}$  as a test function, we have

$$\int_{\mathbb{T}^d} e^{(n-1)\widehat{U}_\varepsilon} \, dx = \varepsilon^2 (n-1) \int_{\mathbb{T}^d} |\nabla \widehat{U}_\varepsilon|^2 e^{\widehat{U}_\varepsilon} \, dx + \int_{\mathbb{T}^d} e^{\bar{U}_\varepsilon} \cdot e^{n\widehat{U}_\varepsilon} \, dx$$

and so

$$\|e^{\widehat{U}_\varepsilon}\|_{L^n(\mathbb{T}^d)}^n \leq e^{M_1} \|e^{\widehat{U}_\varepsilon}\|_{L^{n-1}(\mathbb{T}^d)}^{n-1};$$

again we can make this rigorous by a truncation argument. By induction, we conclude that for all integer  $n$ ,

$$\|e^{\widehat{U}_\varepsilon}\|_{L^n(\mathbb{T}^d)} \leq e^{M_1}. \quad (4.5)$$

Next, we use these estimates to deduce regularity for  $\widehat{U}_\varepsilon$ . By Calderón-Zygmund estimates for the Poisson equation,

$$\|\widehat{U}_\varepsilon\|_{W^{2,n}(\mathbb{T}^d)} \leq C\varepsilon^{-2} \|e^{\bar{U}_\varepsilon + \widehat{U}_\varepsilon}\|_{L^n(\mathbb{T}^d)} \leq C\varepsilon^{-2} (e^{2M_1} + 1).$$

By Sobolev embedding with  $n$  sufficiently large, for any  $\alpha \in (0, 1)$ ,

$$\|\widehat{U}_\varepsilon\|_{C^{1,\alpha}(\mathbb{T}^d)} \leq C\varepsilon^{-2} (e^{2M_1} + 1).$$

Now assume that for some  $\alpha$ ,

$$\|\bar{U}_\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^d)} \leq M_2.$$

Then  $U_\varepsilon \in C^{0,\alpha}(\mathbb{T}^d)$ , with

$$\|U_\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^d)} \leq \|\bar{U}_\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^d)} + \|\widehat{U}_\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^d)} \leq M_2 + C\varepsilon^{-2} (e^{2M_1} + 1).$$

Since

$$\left| e^{U_\varepsilon(x)} - e^{U_\varepsilon(y)} \right| \leq e^{\max\{U_\varepsilon(x), U_\varepsilon(y)\}} |U_\varepsilon(x) - U_\varepsilon(y)|,$$

it follows that

$$\|e^{U_\varepsilon}\|_{C^{0,\alpha}(\mathbb{T}^d)} \leq \|U_\varepsilon\|_{C^{0,\alpha}(\mathbb{T}^d)} \exp [C\varepsilon^{-2} (1 + e^{2M_1})].$$

By Schauder estimates [16, Chapter 4],

$$\begin{aligned} \|\widehat{U}_\varepsilon\|_{C^{2,\alpha}(\mathbb{T}^d)} &\leq C \left( \|\widehat{U}_\varepsilon\|_{L^\infty(\mathbb{T}^d)} + \varepsilon^{-2} \|e^{U_\varepsilon} - 1\|_{C^{0,\alpha}(\mathbb{T}^d)} \right) \\ &\leq [M_2 + C\varepsilon^{-2} (1 + e^{2M_1})] \exp [C\varepsilon^{-2} (1 + e^{2M_1})]. \end{aligned}$$

□

## 4.2 Stability

We are also interested in the stability of  $U_\varepsilon$  with respect to the charge density. We want to prove a quantitative version of [20, Proposition 3.5].

**Proposition 4.4.** *For each  $i = 1, 2$ , let  $\bar{U}_\varepsilon^{(i)}$  be a zero-mean solution of*

$$\varepsilon^2 \Delta \bar{U}_\varepsilon^{(i)} = h_i - 1,$$

where  $h_i \in L^\infty \cap L^{(d+2)/d}(\mathbb{T}^d)$ . Then

$$\|\nabla \bar{U}_\varepsilon^{(1)} - \nabla \bar{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq \varepsilon^{-4} \max_i \|h_i\|_{L^\infty(\mathbb{T}^d)} W_2^2(h_1, h_2). \quad (4.6)$$

Now, in addition, let  $\widehat{U}_\varepsilon^{(i)}$  be a solution of

$$\varepsilon^2 \Delta \widehat{U}_\varepsilon^{(i)} = e^{\bar{U}_\varepsilon^{(i)} + \widehat{U}_\varepsilon^{(i)}} - 1.$$

Then

$$\begin{aligned} \|\nabla \widehat{U}_\varepsilon^{(1)} - \nabla \widehat{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 &\leq \overline{\exp}_2 \left[ C_d \varepsilon^{-2} \left( 1 + \max_i \|h_i\|_{L^{(d+2)/d}(\mathbb{T}^d)} \right) \right] \\ &\quad \times \max_i \|h_i\|_{L^\infty(\mathbb{T}^d)} W_2^2(h_1, h_2). \end{aligned} \quad (4.7)$$

For the stability of  $\nabla \bar{U}_\varepsilon$ , we use a result proved originally by Loeper [35] in the whole space  $\mathbb{R}^d$ , and adapted to the torus  $\mathbb{T}^d$  in [26].

**Lemma 4.5** (Loeper-type estimate for Poisson's equation). *For each  $i = 1, 2$ , let  $\bar{U}_\varepsilon^{(i)}$  be a solution of*

$$\varepsilon^2 \Delta \bar{U}_\varepsilon^{(i)} = h_i - 1,$$

where  $h_i \in L^\infty(\mathbb{T}^d)$ . Then

$$\|\nabla \bar{U}_\varepsilon^{(1)} - \nabla \bar{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq \varepsilon^{-4} \max_i \|h_i\|_{L^\infty(\mathbb{T}^d)} W_2^2(h_1, h_2).$$

For  $\widehat{U}_\varepsilon$  we use a version of [20, Lemma 3.7] in which we quantify the dependence of the constants on  $\varepsilon$ .

**Lemma 4.6.** *For each  $i = 1, 2$ , let  $\widehat{U}_\varepsilon^{(i)} \in W^{1,2} \cap L^\infty(\mathbb{T}^d)$  be a solution of*

$$\varepsilon^2 \Delta \widehat{U}_\varepsilon^{(i)} = e^{\bar{U}_\varepsilon^{(i)} + \widehat{U}_\varepsilon^{(i)}} - 1,$$

for some given potentials  $\bar{U}_\varepsilon^{(i)} \in L^\infty(\mathbb{T}^d)$ . Then

$$\varepsilon^2 \|\nabla \widehat{U}_\varepsilon^{(1)} - \nabla \widehat{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq C_\varepsilon \|\bar{U}_\varepsilon^{(1)} - \bar{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2,$$

where  $C$  depends on the  $L^\infty$  norms of  $\widehat{U}_\varepsilon^{(i)}$  and  $\bar{U}_\varepsilon^{(i)}$ . More precisely,  $C_\varepsilon$  can be chosen such that

$$C_\varepsilon \leq \exp \left[ C \left( \max_i \|\bar{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} + \max_i \|\widehat{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} \right) \right],$$

for some sufficiently large constant  $C$ , independent of  $\varepsilon$ .

*Proof.* We look at the equation solved by the difference  $\widehat{U}_\varepsilon^{(1)} - \widehat{U}_\varepsilon^{(2)}$ . By subtracting the two equations in (4.6), we find that

$$\varepsilon^2 \Delta (\widehat{U}_\varepsilon^{(1)} - \widehat{U}_\varepsilon^{(2)}) = e^{\bar{U}_\varepsilon^{(1)} + \widehat{U}_\varepsilon^{(1)}} - e^{\bar{U}_\varepsilon^{(2)} + \widehat{U}_\varepsilon^{(2)}}. \quad (4.8)$$

By assumption,  $\widehat{U}_\varepsilon^{(i)} \in W^{1,2} \cap L^\infty(\mathbb{T}^d)$ . Hence  $\widehat{U}_\varepsilon^{(1)} - \widehat{U}_\varepsilon^{(2)}$  may be used as a test function in the weak form of (4.8). We multiply both sides by  $\widehat{U}_\varepsilon^{(1)} - \widehat{U}_\varepsilon^{(2)}$ , integrate over  $\mathbb{T}^d$ , and integrate by parts to find

$$-\varepsilon^2 \int_{\mathbb{T}^d} |\nabla \widehat{U}_\varepsilon^{(1)} - \nabla \widehat{U}_\varepsilon^{(2)}|^2 dx = \int_{\mathbb{T}^d} \left( e^{\bar{U}_\varepsilon^{(1)} + \widehat{U}_\varepsilon^{(1)}} - e^{\bar{U}_\varepsilon^{(2)} + \widehat{U}_\varepsilon^{(2)}} \right) (\widehat{U}_\varepsilon^{(1)} - \widehat{U}_\varepsilon^{(2)}) dx.$$

In the proof of [20, Lemma 3.7], we showed that for any  $\bar{U}, \widehat{U}, \bar{V}, \widehat{V} \in L^\infty(\mathbb{T}^d)$ ,

$$-\int_{\mathbb{T}^d} \left( e^{\bar{U} + \widehat{U}} - e^{\bar{V} + \widehat{V}} \right) (\widehat{U} - \widehat{V}) dx \leq C_{U,V} \|\bar{U} - \bar{V}\|_{L^2(\mathbb{T}^d)}^2,$$

where

$$C_{U,V} \leq \exp \left\{ C \left[ \|\bar{U}\|_{L^\infty(\mathbb{T}^d)} + \|\widehat{U}\|_{L^\infty(\mathbb{T}^d)} + \|\bar{V}\|_{L^\infty(\mathbb{T}^d)} + \|\widehat{V}\|_{L^\infty(\mathbb{T}^d)} \right] \right\}.$$

Thus

$$\|\nabla \widehat{U}_\varepsilon^{(1)} - \nabla \widehat{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq C_\varepsilon \varepsilon^{-2} \|\bar{U}_\varepsilon^{(1)} - \bar{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2,$$

where

$$C_\varepsilon \leq \exp \left[ C \left( \max_i \|\bar{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} + \max_i \|\widehat{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} \right) \right].$$

□

*Proof of Proposition 4.4.* It suffices to prove (4.7), since (4.6) follows immediately from Lemma 4.5. By the Poincaré inequality for zero-mean functions,

$$\|\bar{U}_\varepsilon^{(1)} - \bar{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq C \|\nabla \bar{U}_\varepsilon^{(1)} - \nabla \bar{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2.$$

By Lemma 4.5,

$$\|\nabla \bar{U}_\varepsilon^{(1)} - \nabla \bar{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq \varepsilon^{-4} \max_i \|h_i\|_{L^\infty(\mathbb{T}^d)} W_2^2(h_1, h_2).$$

Then, by Lemma 4.6,

$$\|\nabla \widehat{U}_\varepsilon^{(1)} - \nabla \widehat{U}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq C_\varepsilon \varepsilon^{-6} \max_i \|h_i\|_{L^\infty(\mathbb{T}^d)} W_2^2(h_1, h_2),$$

for

$$C_\varepsilon \leq \exp \left[ C \left( \max_i \|\bar{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} + \max_i \|\widehat{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} \right) \right].$$

For the  $L^\infty(\mathbb{T}^d)$  estimates, we use Proposition 4.1:

$$\begin{aligned} \|\bar{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} &\leq C_d \varepsilon^{-2} \left( 1 + \|h_i\|_{L^{\frac{d+2}{d}}} \right) \\ \|\widehat{U}_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} &\leq C_d \exp \left[ C_d \varepsilon^{-2} \left( 1 + \|h_i\|_{L^{\frac{d+2}{d}}} \right) \right]. \end{aligned}$$

Hence

$$C_\varepsilon \leq \overline{\exp}_2 \left[ C_d \varepsilon^{-2} \left( 1 + \max_i \|h_i\|_{L^{\frac{d+2}{d}}} \right) \right],$$

which implies (4.7). □

## 5 Strong-strong stability

To prove the quasineutral limit, we will need a quantified stability estimate between solutions of the VPME system (1.2). The following proposition is an  $\varepsilon$ -dependent version of [20, Proposition 4.1]. We revisit the proof here, keeping track of how all the constants depend on  $\varepsilon$ .

**Proposition 5.1** (Stability for solutions with bounded density). *For  $i = 1, 2$ , let  $f_\varepsilon^{(i)}$  be solutions of (1.3) satisfying for some constant  $M$  and all  $t \in [0, T]$ ,*

$$\rho[f_\varepsilon^{(i)}(t)] \leq M.$$

*Then there exists a constant  $C_\varepsilon$  such that, for all  $t \in [0, T]$ ,*

$$W_2 \left( f_\varepsilon^{(1)}(t), f_\varepsilon^{(2)}(t) \right) \leq \begin{cases} C \exp \left[ C \left( 1 + \log \frac{W_2(f_\varepsilon^{(1)}(0), f_\varepsilon^{(2)}(0))}{4\sqrt{d}} \right) e^{-C_\varepsilon t} \right] & \text{if } W_2 \left( f_\varepsilon^{(1)}(0), f_\varepsilon^{(2)}(0) \right) \leq d \\ W_2 \left( f_\varepsilon^{(1)}(0), f_\varepsilon^{(2)}(0) \right) e^{C_\varepsilon t} & \text{if } W_2 \left( f_\varepsilon^{(1)}(0), f_\varepsilon^{(2)}(0) \right) > d. \end{cases}$$

*If in addition, for some constant  $C_0$ ,*

$$\sup_{t \in [0, T]} \|f_\varepsilon^{(i)}(t, \cdot, \cdot)\|_{L_{x,v}^\infty} \leq C_0, \quad \sup_{t \in [0, T]} \mathcal{E}_\varepsilon[f_\varepsilon^{(i)}](t) \leq C_0, \quad (5.1)$$

*then  $C_\varepsilon$  may be chosen to satisfy*

$$C_\varepsilon \leq \overline{\exp}_2(C\varepsilon^{-2})(M + 1).$$

*Proof.* The  $W_2$  distance is defined as an infimum over couplings (2.1). Thus it suffices to estimate the  $L^2$  distance corresponding to a particular coupling of  $f_\varepsilon^{(1)}$  and  $f_\varepsilon^{(2)}$ . To do this we will first represent  $f_\varepsilon^{(i)}$  as the pushforward of  $f_\varepsilon^{(i)}(0)$  along the characteristic flow induced by  $f_\varepsilon^{(i)}$ . That is, consider the following system of ODEs for  $Z_z^{(i)} = (X_z^{(i)}, V_z^{(i)}) \in \mathbb{T}^d \times \mathbb{R}^d$ :

$$\begin{cases} \dot{X}_z^{(i)} = V_z^{(i)}, \\ \dot{V}_z^{(i)} = E_\varepsilon^{(i)}(X_z^{(i)}), \\ (X_z^{(i)}(0), V_z^{(i)}(0)) = (x, v) = z, \end{cases} \quad (5.2)$$

where the electric field  $E_\varepsilon^{(i)}$  is given by

$$\begin{aligned} E_\varepsilon^{(i)} &= -\nabla U_\varepsilon^{(i)}, \\ \varepsilon^2 \Delta U_\varepsilon^{(i)} &= e^{U_\varepsilon^{(i)}} - \rho_\varepsilon^{(i)} := e^{U_\varepsilon^{(i)}} - \rho[f_\varepsilon^{(i)}]. \end{aligned}$$

Since  $f_\varepsilon^{(i)}$  has bounded mass density, Proposition 4.1 and Lemma 4.2 imply that  $E_\varepsilon^{(i)}$  is a log-Lipschitz vector field, with a constant uniform on  $[0, T]$ . This regularity is enough to imply that there exists a unique solution of (5.2) for every initial condition  $z$ , resulting in a well-defined characteristic flow. Since the characteristic flow is unique, by [1, Theorem 3.1] the linear Vlasov equation

$$\begin{cases} \partial_t g + v \cdot \nabla_x g + E_\varepsilon^{(i)}(x) \cdot \nabla_v g = 0, \\ g|_{t=0} = f_\varepsilon^{(i)}(0) \geq 0 \end{cases} \quad (5.3)$$

has a unique solution. Moreover, this solution can be represented, in weak form, by the relation

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \phi g_t(dz) = \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \left( Z_z^{(i)} \right) f_\varepsilon^{(i)}(0)(dz), \quad (5.4)$$

for all  $\phi \in C_b(\mathbb{T}^d \times \mathbb{R}^d)$ . Since  $f_\varepsilon^{(i)}$  is certainly a solution of (5.3), we deduce that  $g = f_\varepsilon^{(i)}$  and so  $f_\varepsilon^{(i)}$  has the representation (5.4). We will use this representation to define a coupling between  $f_\varepsilon^{(1)}$  and  $f_\varepsilon^{(2)}$ .

Fix an arbitrary initial coupling  $\pi_0 \in \Pi[f_\varepsilon^{(1)}(0), f_\varepsilon^{(2)}(0)]$ . We define  $\pi_t$  to follow the corresponding characteristic flows: for  $\phi \in C_b((\mathbb{T}^d \times \mathbb{R}^d)^2)$ , let

$$\int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \phi(z_1, z_2) d\pi_t(z_1, z_2) = \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \left( Z_z^{(1)}, Z_z^{(2)} \right) d\pi_0(z_1, z_2).$$

We can verify that  $\pi_t$  is indeed a coupling of  $f_\varepsilon^{(1)}$  and  $f_\varepsilon^{(2)}$  by checking the marginals:

$$\begin{aligned} \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \phi(z_i) d\pi_t(z_1, z_2) &= \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \phi \left( Z_{z_i}^{(i)}(t) \right) d\pi_0(z_1, z_2) \\ &= \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \left( Z_z^{(i)}(t) \right) f_\varepsilon^{(i)}(0)(dz) \\ &= \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi(z) f_\varepsilon^{(i)}(t)(dz). \end{aligned} \quad (5.5)$$

Next, we define a functional which is greater than (or equal to) the squared Wasserstein distance between  $f_\varepsilon^{(1)}$  and  $f_\varepsilon^{(2)}$ . Let

$$D(t) = \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |x_1 - x_2|^2 + |v_1 - v_2|^2 d\pi_t(x_1, v_1, x_2, v_2).$$

Since the Wasserstein distance is an infimum over couplings, while  $\pi_t$  is a particular coupling,  $D$  must control the squared Wasserstein distance:

$$W_2^2(f_\varepsilon^{(1)}, f_\varepsilon^{(2)}) \leq D.$$

To prove our stability estimate, it therefore suffices to estimate  $D$ . We will do this using a Grönwall estimate. Taking a time derivative, we find

$$\begin{aligned} \dot{D} = 2 \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} & \left( X_{z_1}^{(1)}(t) - X_{z_2}^{(2)}(t) \right) \cdot \left( V_{z_1}^{(1)}(t) - V_{z_2}^{(2)}(t) \right) \\ & + \left( V_{z_1}^{(1)}(t) - V_{z_2}^{(2)}(t) \right) \cdot \left( E_\varepsilon^{(1)}(X_{z_1}^{(1)}(t)) - E_\varepsilon^{(2)}(X_{z_2}^{(2)}(t)) \right) d\pi_0(z_1, z_2). \end{aligned}$$

Using the Cauchy-Schwarz inequality, we obtain

$$\dot{D} \leq D + \sqrt{D} \left( \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \left| E_\varepsilon^{(1)}(X_{z_1}^{(1)}(t)) - E_\varepsilon^{(2)}(X_{z_2}^{(2)}(t)) \right|^2 d\pi_0(z_1, z_2) \right)^{1/2}.$$

In other words,

$$\dot{D} \leq D + \sqrt{D} \left( \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \left| E_\varepsilon^{(1)}(x_1) - E_\varepsilon^{(2)}(x_2) \right|^2 d\pi_t(z_1, z_2) \right)^{1/2}.$$

We split the electric field term into the form

$$\int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \left| E_\varepsilon^{(1)}(x_1) - E_\varepsilon^{(2)}(x_2) \right|^2 d\pi_t(z_1, z_2) \leq C \sum_{i=1}^4 I_i,$$

where

$$\begin{aligned} I_1 &:= \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |\bar{E}_\varepsilon^{(1)}(x_1) - \bar{E}_\varepsilon^{(1)}(x_2)|^2 d\pi_t, & I_2 &:= \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |\bar{E}_\varepsilon^{(1)}(x_2) - \bar{E}_\varepsilon^{(2)}(x_2)|^2 d\pi_t, \\ I_3 &:= \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |\hat{E}_\varepsilon^{(1)}(x_1) - \hat{E}_\varepsilon^{(1)}(x_2)|^2 d\pi_t, & I_4 &:= \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |\hat{E}_\varepsilon^{(1)}(x_2) - \hat{E}_\varepsilon^{(2)}(x_2)|^2 d\pi_t. \end{aligned}$$

**Control of  $I_1$ :** To estimate  $I_1$ , observe that by Lemma 4.2,

$$I_1 \leq C\varepsilon^{-4} \|\rho_\varepsilon^{(1)}\|_{L^\infty(\mathbb{T}^d)}^2 \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |x_1 - x_2|^2 \left( 1 + \log \left( \frac{\sqrt{d}}{|x_1 - x_2|} \right) \mathbb{1}_{|x_1 - x_2| \leq \sqrt{d}} \right)^2 d\pi_t.$$

We introduce the function

$$H(x) := \begin{cases} x \left( \log \frac{x}{16d} \right)^2 & \text{if } x \leq d \\ d \left( \log 16 \right)^2 & \text{if } x > d, \end{cases}$$

which is concave on  $\mathbb{R}_+$  (see [20, Lemma 4.4]). Then

$$I_1 \leq C\varepsilon^{-4} \|\rho_\varepsilon^{(1)}\|_{L^\infty(\mathbb{T}^d)}^2 \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} H(|x_1 - x_2|^2) d\pi_t.$$

By Jensen's inequality,

$$I_1 \leq C\varepsilon^{-4} \|\rho_\varepsilon^{(1)}\|_{L^\infty(\mathbb{T}^d)}^2 H \left( \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |x_1 - x_2|^2 d\pi_t \right) \leq C\varepsilon^{-4} M^2 H(D).$$

**Control of  $I_2$ :** For  $I_2$ , observe that

$$\begin{aligned} I_2 &\leq \int_{\mathbb{T}^d} |\bar{E}_\varepsilon^{(1)}(x) - \bar{E}_\varepsilon^{(2)}(x)|^2 \rho_\varepsilon^{(2)}(dx) \\ &\leq \|\rho_\varepsilon^{(2)}\|_{L^\infty(\mathbb{T}^d)} \|\bar{E}_\varepsilon^{(1)} - \bar{E}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2. \end{aligned}$$

We apply the Loeper stability estimate (Lemma 4.5) to obtain

$$I_2 \leq \varepsilon^{-4} \max_i \|\rho_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)}^2 W_2^2(\rho_\varepsilon^{(1)}, \rho_\varepsilon^{(2)}) \leq C\varepsilon^{-4} M^2 D.$$

**Control of  $I_3$ :** To estimate  $I_3$ , we recall a regularity estimate on  $\widehat{U}_\varepsilon^{(1)}$  from Proposition 4.1:

$$\|\widehat{E}_\varepsilon^{(1)}\|_{C^1(\mathbb{T}^d)} \leq \|\widehat{U}_\varepsilon^{(1)}\|_{C^2(\mathbb{T}^d)} \leq C_d \overline{\exp}_2 \left( C_d \varepsilon^{-2} \left( 1 + \|\rho_\varepsilon^{(1)}\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \right) \right).$$

Under conditon (5.1), by Lemma 3.3,

$$\|\rho_\varepsilon^{(1)}\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C$$

for some  $C$  depending on  $C_0$  only. Therefore

$$I_3 \leq \|\widehat{E}_\varepsilon^{(1)}\|_{C^1(\mathbb{T}^d)}^2 \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |x_1 - x_2|^2 d\pi_t \leq \overline{\exp}_2(C\varepsilon^{-2}) D,$$

for some  $C$  depending on  $C_0$  and  $d$  only. If (5.1) does not hold, we can use the fact that

$$\|\rho_\varepsilon^{(1)}\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq \|\rho_\varepsilon^{(1)}\|_{L^\infty(\mathbb{T}^d)} \leq M$$

and complete the proof in the same way, to find a constant depending on  $M$ .

**Control of  $I_4$ :** First, note that

$$I_4 = \int_{\mathbb{T}^d} |\widehat{E}_\varepsilon^{(1)}(x) - \widehat{E}_\varepsilon^{(2)}(x)|^2 \rho_\varepsilon^{(2)}(dx) \leq \|\rho_\varepsilon^{(2)}\|_{L^\infty(\mathbb{T}^d)} \|\widehat{E}_\varepsilon^{(1)} - \widehat{E}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2.$$

We apply the stability estimate on  $\widehat{U}_\varepsilon^{(1)}$  from Proposition 4.4 to find

$$\begin{aligned} \|\widehat{E}_\varepsilon^{(1)} - \widehat{E}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 &\leq \overline{\exp}_2 \left[ C_d \varepsilon^{-2} \left( 1 + \max_i \|\rho_\varepsilon^{(i)}\|_{L^{(d+2)/d}(\mathbb{T}^d)} \right) \right] \\ &\quad \times \max_i \|\rho_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} W_2^2(\rho_\varepsilon^{(1)}, \rho_\varepsilon^{(2)}). \end{aligned}$$

Once again, if conditon (5.1) holds then

$$\|\widehat{E}_\varepsilon^{(1)} - \widehat{E}_\varepsilon^{(2)}\|_{L^2(\mathbb{T}^d)}^2 \leq \overline{\exp}_2(C\varepsilon^{-2}) \max_i \|\rho_\varepsilon^{(i)}\|_{L^\infty(\mathbb{T}^d)} W_2^2(\rho_\varepsilon^{(1)}, \rho_\varepsilon^{(2)}),$$

for some  $C$  depending on  $C_0$  and  $d$  only. Thus

$$I_4 \leq \overline{\exp}_2(C\varepsilon^{-2}) M^2 D.$$

Altogether we find that

$$\dot{D} \leq \begin{cases} C_\varepsilon D (1 + |\log \frac{D}{16d}|) & \text{if } D < d \\ C_\varepsilon (1 + \log 16) D & \text{if } D \geq d. \end{cases}$$

If (5.1) holds, then  $C_\varepsilon$  may be chosen to satisfy

$$\begin{aligned} C_\varepsilon &\leq C\varepsilon^{-2}M + \overline{\exp}_2(C\varepsilon^{-2}) + \overline{\exp}_2(C\varepsilon^{-2})M \\ &\leq \overline{\exp}_2(C\varepsilon^{-2})(M+1). \end{aligned}$$

We conclude that

$$D(t) \leq C \exp \left[ \left( 1 + \log \frac{D(0)}{16d} \right) e^{-C_\varepsilon t} \right]$$

as long as  $D \leq d$ ; if  $D > d$  then

$$D(t) \leq (d \vee D(0)) e^{C_\varepsilon(1+\log 16)t}.$$

□

## 6 Growth estimates

In this section we will study how the support of a solution of the VPME system grows in time. This argument is very much dimension dependent and we shall present two different proofs for the two and three dimensional case, respectively in Sections 6.1 and 6.2. The result we obtain is the following:

**Proposition 6.1** (Mass bounds). *Let  $f_\varepsilon$  be a solution of (1.3) satisfying for some constant  $C_0$ ,*

$$\|f_\varepsilon\|_{L^\infty([0,T] \times \mathbb{T}^d \times \mathbb{R}^d)} \leq C_0, \quad \sup_{t \in [0,T]} \mathcal{E}_\varepsilon[f_\varepsilon(t)] \leq C_0.$$

*Assume that  $\rho_{f_\varepsilon} \in L^\infty([0,T]; L^\infty(\mathbb{T}^d))$ . Let  $R_0$  be any constant such that the support of  $f_\varepsilon(0, \cdot, \cdot)$  is contained in  $\mathbb{T}^d \times B_{\mathbb{R}^d}(0, R_0)$ .*

(i) *If  $d = 2$ , then:*

$$\sup_{t \in [0,T]} \|\rho_{f_\varepsilon}(t)\|_{L^\infty(\mathbb{T}^2)} \leq C_T e^{C_\varepsilon^{-2}} [R_0 + \varepsilon^{-2}]^2.$$

*The constant  $C$  depends on  $C_0$  only, while  $C_T$  depends on  $C_0$  and  $T$ .*

(ii) *If  $d = 3$ , then:*

$$\sup_{t \in [0,T]} \|\rho_\varepsilon(t)\|_{L^\infty(\mathbb{T}^3)} \leq \max\{T^{-81/8}, C(R_0^3 + e^{C_\varepsilon^{-2}} T^6)\}.$$

*The constant  $C$  depends on  $C_0$  only.*

### 6.1 Proof of Proposition 6.1 in the two dimensional case

In this section, we prove an estimate on the mass density  $\rho_\varepsilon$  in the case  $d = 2$ . Observe first that if the support of  $f_\varepsilon$  is contained in the set  $\mathbb{T}^2 \times B_{\mathbb{R}^2}(0; R_t)$ , then

$$\|\rho_\varepsilon(t)\|_{L^\infty(\mathbb{T}^d)} \leq CR_t^2, \tag{6.1}$$

where  $C$  is a constant depending on  $\|f_\varepsilon(0)\|_{L^\infty(\mathbb{T}^2 \times \mathbb{R}^2)}$ . Our argument will rely on controlling the growth of the support of  $f_\varepsilon$ . To do this we will find a bound on the growth rate of the velocity component of the characteristic trajectories. Since any characteristic trajectory  $(X_t, V_t)$  satisfies

$$\dot{V}_t = E_\varepsilon(X_t),$$



we look for a uniform estimate on the electric field  $E_\varepsilon$ .

By Proposition 4.1, for the smooth part  $\widehat{E}_\varepsilon$  we have the estimate

$$\begin{aligned}\|\widehat{E}_\varepsilon\|_{L^\infty(\mathbb{T}^2)} &\leq \exp\{C\varepsilon^{-2}(1 + \|\rho_\varepsilon\|_{L^2(\mathbb{T}^2)})\} \\ &\leq \exp(C\varepsilon^{-2}),\end{aligned}$$

where the constant  $C$  depends only on a bound on the initial energy. However, for the singular part  $\bar{E}_\varepsilon$ , Proposition 4.1 only gives us the estimate

$$\|\bar{E}_\varepsilon\|_{L^\infty(\mathbb{T}^2)} \leq C\varepsilon^{-2} \left(1 + \|\rho_\varepsilon\|_{L^\infty(\mathbb{T}^d)}\right),$$

which depends on an  $L^\infty$  bound on the mass density. If we use this estimate in combination with (6.1), this results in a bound on the size of the support of the form

$$R_t \leq R_0 + \exp(C\varepsilon^{-2})t + C\varepsilon^{-2} \int_0^t R_s^2 ds. \quad (6.2)$$

The solution of the ODE

$$\dot{y} = C(1 + y^2)$$

blows up in finite time and so the differential inequality (6.2) is not enough to imply a bound on  $R_t$ . We need to use a more careful estimate on  $\bar{E}_\varepsilon$ . In dimension two, we can make use of the fact that the conservation of energy gives us a uniform bound on  $\|\rho_\varepsilon\|_{L^2(\mathbb{T}^2)}$ , by Lemma 3.3, and use an interpolation argument.

**Lemma 6.2.** *Let  $\rho \in L^1 \cap L^\infty(\mathbb{T}^2)$  satisfy the bounds*

$$\|\rho\|_{L^1(\mathbb{T}^2)} = 1, \quad \|\rho\|_{L^2(\mathbb{T}^2)} \leq C_0, \quad \|\rho\|_{L^\infty(\mathbb{T}^2)} \leq M.$$

*Let  $\bar{E}_\varepsilon = -\nabla\bar{U}_\varepsilon$ , where  $\bar{U}_\varepsilon$  is the unique  $W^{1,2}(\mathbb{T}^2)$  solution of the Poisson equation*

$$\varepsilon^2 \Delta \bar{U}_\varepsilon = 1 - \rho.$$

*Then there exists a constant  $C$  depending only on  $C_0$  such that*

$$\|\bar{E}_\varepsilon\|_{L^\infty(\mathbb{T}^2)} \leq C\varepsilon^{-2} \left(1 + |\log M|^{1/2}\right).$$

*Proof.* We use the representation

$$\bar{E}_\varepsilon = K * (\rho - 1) = C\varepsilon^{-2} \int_{\mathbb{T}^d} \frac{x-y}{|x-y|^2} (\rho(y) - 1) dy + \varepsilon^{-2} K_0 * (\rho - 1),$$

where  $K_0$  is a  $C^1(\mathbb{T}^2)$  function. By Young's inequality,

$$\|K_0 * (\rho - 1)\|_{L^\infty(\mathbb{T}^2)} \leq \|K_0\|_{L^\infty(\mathbb{T}^2)} \|\rho - 1\|_{L^1(\mathbb{T}^2)} \leq C.$$

We split the integral term into a part where  $|x-y|$  is small and a part where  $|x-y|$  is large:

$$\int_{\mathbb{T}^d} \frac{x-y}{|x-y|^2} (\rho(y) - 1) dy = \int_{|x-y| \leq l} \frac{x-y}{|x-y|^2} (\rho(y) - 1) dy + \int_{|x-y| \geq l} \frac{x-y}{|x-y|^2} (\rho(y) - 1) dy.$$

For the part where  $|x - y|$  is large, we use Young's inequality with the  $L^2$  control on  $\rho$ :

$$\begin{aligned}
\left| \int_{|x-y| \geq l} \frac{x-y}{|x-y|^2} (\rho(y) - 1) dy \right| &\leq \int_{|x-y| \geq l} \frac{1}{|x-y|} |\rho(y) - 1| dy \\
&\leq \left\| \frac{1}{|x|} \mathbb{1}_{|x| \geq l} * |\rho - 1| \right\|_{L^\infty(\mathbb{T}^2)} \\
&\leq \left( \int_{|x| \geq l} \frac{1}{|x|^2} dx \right)^{1/2} \|\rho - 1\|_{L^2(\mathbb{T}^2)} \\
&\leq C(C_0 + 1) |\log l|^{1/2}.
\end{aligned}$$

Where  $|x - y|$  is small, we use Young's inequality with the  $L^\infty$  control on  $\rho$ :

$$\begin{aligned}
\left| \int_{|x-y| \leq l} \frac{x-y}{|x-y|^2} (\rho(y) - 1) dy \right| &\leq \int_{|x-y| \leq l} \frac{1}{|x-y|} |\rho(y) - 1| dy \\
&\leq \left\| \frac{1}{|x|} \mathbb{1}_{|x| \leq l} * |\rho - 1| \right\|_{L^\infty(\mathbb{T}^2)} \\
&\leq \|\rho - 1\|_{L^\infty(\mathbb{T}^2)} \int_{|x| \leq l} \frac{1}{|x|} dx \\
&\leq C M l.
\end{aligned}$$

Altogether we obtain

$$\|\bar{E}_\varepsilon\|_{L^\infty(\mathbb{T}^2)} \leq C\varepsilon^{-2} \left[ 1 + C_0 |\log l|^{1/2} + Ml \right].$$

We choose  $l = M^{-1}$  and conclude that

$$\|\bar{E}_\varepsilon\|_{L^\infty(\mathbb{T}^2)} \leq C\varepsilon^{-2} \left( 1 + |\log M|^{1/2} \right),$$

where  $C$  depends on  $C_0$  only. □

By using this estimate, we can deduce a differential inequality on  $R_t$  that can be closed.

**Lemma 6.3.** *Let  $f_\varepsilon$  be a solution of (1.3) with bounded energy and compact support contained in  $\mathbb{T}^2 \times B_{\mathbb{R}^2}(0; R_t)$  at time  $t$ . Then  $R$  satisfies the estimate*

$$R_t \leq e^{C\varepsilon^{-2}t} (1 + R_0 + (\log t \vee 0)).$$

*Proof.* We consider the velocity coordinate  $V_t(x, v)$  of an arbitrary characteristic trajectory starting from  $(x, v)$  at time  $t = 0$ . We have

$$\begin{aligned}
|V_t(x, v)| &\leq |v| + \int_0^t \|E_\varepsilon\|_{L^\infty(\mathbb{T}^2)} ds \\
&\leq |v| + \int_0^t \|\widehat{E}_\varepsilon\|_{L^\infty(\mathbb{T}^2)} + \|\bar{E}_\varepsilon\|_{L^\infty(\mathbb{T}^2)} ds \\
&\leq |v| + \int_0^t \exp(C\varepsilon^{-2}s) + C\varepsilon^{-2} \left( 1 + |\log R_s|^{1/2} \right) ds
\end{aligned}$$

The size of the support is controlled by the modulus of the furthest-reaching characteristic trajectory that starts within the support of  $f_\varepsilon(0)$ :

$$\begin{aligned} R_t &\leq \sup_{(x,v) \in \mathbb{T}^2 \times B_{\mathbb{R}^2}(0; R_0)} |V_t(x, v)| \\ &\leq \sup_{(x,v) \in \mathbb{T}^2 \times B_{\mathbb{R}^2}(0; R_0)} \left\{ |v| + \int_0^t \exp(C\varepsilon^{-2}) + C\varepsilon^{-2} \left(1 + |\log R_s|^{1/2}\right) ds \right\} \\ &\leq R_0 + \int_0^t \exp(C\varepsilon^{-2}) + C\varepsilon^{-2} |\log R_s|^{1/2} ds. \end{aligned}$$

We compare this with the function

$$z(t) = (1 + 2Ct) [R_0 + \log(1 + 2Ct)].$$

By [20, Lemma A.1], this satisfies

$$\dot{z} \geq C(1 + \log(1 + z)).$$

We deduce that

$$R_t \leq e^{C\varepsilon^{-2}t} (\varepsilon^{-2} + R_0 + (\log t \vee 0)).$$

□

*Proof of Proposition 6.1, case  $d = 2$ .* We combine Lemma 6.3 with the elementary estimate (6.1): for all  $t \in [0, T]$ ,

$$\begin{aligned} \|\rho_\varepsilon(t)\|_{L^\infty(\mathbb{T}^d)} &\leq CR_t^2 \\ &\leq e^{C\varepsilon^{-2}t} (\varepsilon^{-2} + R_0 + (\log t \vee 0))^2 \\ &\leq C_T e^{C\varepsilon^{-2}t} (R_0 + \varepsilon^{-2})^2. \end{aligned}$$

□

## 6.2 Proof of Proposition 6.1 in the three dimensional case

In this section, we prove a mass bound in the case  $d = 3$ . In this case, the conservation of energy gives us a uniform bound on  $\|\rho_\varepsilon\|_{L^{5/3}(\mathbb{T}^3)}$ . This integrability is not enough to allow us to use the elementary interpolation approach that we used in the two dimensional case. Instead, we will adapt estimates devised by Batt and Rein [5] for the classical Vlasov-Poisson equation on  $\mathbb{T}^3 \times \mathbb{R}^3$ . We used this approach in [20] to prove the existence of solutions of the VPME system with bounded density. Here we focus on identifying how the bounds on  $\|\rho_\varepsilon\|_{L^\infty(\mathbb{T}^3)}$  depend on  $\varepsilon$ .

As in the two dimensional case, Batt and Rein's argument relies on controlling the mass density using the characteristic trajectory with velocity component of greatest Euclidean norm starting within the support of  $f_\varepsilon(0)$  at time zero. They prove a bootstrap estimate on the convolution integral defining the singular part of the electric field. We recall this estimate in the following technical lemma.

**Lemma 6.4.** *Let  $(X(t; s, x, v), V(t; s, x, v))$  denote the solution at time  $t$  of an ODE*

$$\begin{pmatrix} \dot{X}(t) \\ \dot{V}(t) \end{pmatrix} = a(X(t), V(t)), \quad \begin{pmatrix} X(s) \\ V(s) \end{pmatrix} = \begin{pmatrix} x \\ v \end{pmatrix},$$

where  $a$  is of the form

$$a(X, V) = \begin{pmatrix} V \\ a_2(X, V) \end{pmatrix}.$$

Assume that  $f = f(t, x, v)$  is the pushforward of  $f_0$  along the associated characteristic flow; that is, for all  $\phi \in C_b(\mathbb{T}^3 \times \mathbb{R}^3)$ ,

$$\int_{\mathbb{T}^3 \times \mathbb{R}^3} f(t, x, v) \phi(x, v) \, dx \, dv = \int_{\mathbb{T}^3 \times \mathbb{R}^3} f(s, x, v) \phi(X(t; s, x, v), V(t; s, x, v)) \, dx \, dv,$$

and that  $f$  is bounded with a uniformly bounded second moment in velocity:

$$\|f\|_{L^\infty([0, T] \times \mathbb{T}^3 \times \mathbb{R}^3)} \leq C, \quad \sup_{t \in [0, T]} \|f_t |v|^2\|_{L^1(\mathbb{T}^3 \times \mathbb{R}^3)} \leq C.$$

Define the quantities  $h_\rho, h_\eta$  by

$$h_\rho(t) := \sup\{\|\rho_f(s)\|_{L^\infty(\mathbb{T}^3)}; 0 \leq s \leq t\}$$

$$h_\eta(t, \Delta) := \sup\{|V(t_1, \tau, x, v) - V(t_2, \tau, x, v)|; 0 \leq t_1, t_2, \tau \leq t, |t_1 - t_2| \leq \Delta, (x, v) \in \mathbb{T}^3 \times \mathbb{R}^3\}.$$

Assume that there exists  $C_* > 1$  such that

$$h_\eta(t, \Delta) \leq C_* h_\rho(t)^\beta \Delta. \tag{6.3}$$

Then for all  $t_1 < t_2 \leq t$ , if

$$h_\rho(t)^{-\beta/2} \leq \Delta \leq t$$

then

$$\int_{t_1}^{t_2} \int_{\mathbb{T}^3} |X(s) - y|^{-2} \rho_f(s, y) \, dy \, ds \leq C C_*^{4/3} \left( h_\rho(t)^{2\beta/3} + h_\rho(t)^{1/6} \right) \Delta.$$

We complete the proof of Proposition 6.1 by combining Lemma 6.4 with the estimates on  $\widehat{E}_\varepsilon$  from Proposition 4.1.

*Proof of Proposition 6.1, case  $d = 3$ .* For any characteristic trajectory  $(X_t, V_t)$ , we have for any  $0 \leq t_1 < t_2 \leq T$ ,

$$|V_{t_1} - V_{t_2}| \leq \int_{t_2}^{t_1} |E_\varepsilon(X_s)| \, ds.$$

We can write the total force  $E_\varepsilon$  in the form

$$E_\varepsilon(x) = \varepsilon^{-2} [K_0 * (\rho_\varepsilon - 1)](x) + C \varepsilon^{-2} \int_{\mathbb{T}^3} \frac{x - y}{|x - y|^3} \rho_\varepsilon(y) \, dy + \widehat{E}_\varepsilon.$$

Since  $K_0$  is a  $C^1(\mathbb{T}^3)$  function and  $\rho_\varepsilon$  has unit mass, the first term satisfies the bound

$$\begin{aligned} |\varepsilon^{-2} [K_0 * (\rho_\varepsilon - 1)](x)| &\leq \varepsilon^{-2} \|K_0\|_{L^\infty(\mathbb{T}^3)} \|\rho_\varepsilon - 1\|_{L^1(\mathbb{T}^3)} \\ &\leq C \varepsilon^{-2}. \end{aligned}$$

For the last term, we use Proposition 4.1:

$$|\widehat{E}_\varepsilon| \leq \exp \left[ C \left( 1 + \|\rho_\varepsilon\|_{L^{5/3}(\mathbb{T}^3)} \right) \right] \leq \exp(C \varepsilon^{-2}).$$

Therefore,

$$|V_{t_1} - V_{t_2}| \leq \int_{t_2}^{t_1} \left[ \exp(C\varepsilon^{-2}) + C\varepsilon^{-2} \int_{\mathbb{T}^3} |X(s) - y|^{-2} \rho_\varepsilon(y) \, dy \right] \, ds. \quad (6.4)$$

From [5] we have the estimate

$$\int_{\mathbb{T}^3} |x - y|^{-2} \rho_\varepsilon(s, y) \, dy \leq C \|\rho_\varepsilon(s, \cdot)\|_{L^{5/3}}^{5/9} \|\rho_\varepsilon(s, \cdot)\|_{L^\infty(\mathbb{T}^3)}^{4/9} \leq C \|\rho_\varepsilon(s, \cdot)\|_{L^\infty(\mathbb{T}^3)}^{4/9},$$

where  $C$  depends only on  $\|f_\varepsilon(0)\|_{L^\infty(\mathbb{T}^3 \times \mathbb{R}^3)}$  and the initial energy. Alternatively, by an elementary argument we can prove the estimate

$$\int_{\mathbb{T}^3} |x - y|^{-2} \rho_\varepsilon(s, y) \, dy \leq C \|\rho_\varepsilon(s, \cdot)\|_{L^\infty(\mathbb{T}^3)}^{4/9+\eta},$$

for any  $\eta > 0$ ; see [20, Remark 5.7]. The choice of exponent does not affect the number of iterations used in the next part of the proof.

By (6.4), we have

$$h_\eta(t, \Delta) \leq \left( C\varepsilon^{-2} h_\rho(t)^{\frac{4}{9}} + e^{C\varepsilon^{-2}} \right) \Delta.$$

Since  $\rho_\varepsilon$  has total mass 1,  $h_\rho \geq 1$ . Thus

$$h_\eta(t, \Delta) \leq e^{C\varepsilon^{-2}} \Delta h_\rho(t)^{\frac{4}{9}}.$$

This means that condition (6.3) is satisfied with  $C_* = e^{C\varepsilon^{-2}}$ . We apply Lemma 6.4 to improve our bound on

$$\int_{\mathbb{T}^3} |X(s) - y|^{-2} \rho_\varepsilon(y) \, dy \leq C C_*^{4/3} \left( h_\rho(t)^{\frac{2}{3} \cdot \frac{4}{9}} + h_\rho(t)^{1/6} \right) \Delta.$$

Feeding this new estimate into (6.4), we obtain

$$h_\eta(t, \Delta) \leq C\varepsilon^{-2} C_*^{4/3} \left( h_\rho(t)^{\frac{2}{3} \cdot \frac{4}{9}} + h_\rho(t)^{1/6} \right) \Delta + (C\varepsilon^{-2} + C e^{C\varepsilon^{-2}}) \Delta \leq e^{C\varepsilon^{-2}} h_\rho(t)^{\frac{8}{27}} \Delta,$$

as long as

$$h_\rho(t)^{-2/9} \leq \Delta \leq t.$$

We will iterate this process until we achieve the lowest possible exponent for  $h_\rho$ , i.e.  $\frac{1}{6}$ . Applying Lemma 6.4 a second time, we obtain

$$\int_{\mathbb{T}^3} |X(s) - y|^{-2} \rho_\varepsilon(y) \, dy \leq C C_*^{4/3} \left( h_\rho(t)^{\frac{2}{3} \cdot \frac{8}{27}} + h_\rho(t)^{1/6} \right) \Delta,$$

with  $C_* = e^{C\varepsilon^{-2}}$ , provided that

$$h_\rho(t)^{-4/27} \leq \Delta \leq t,$$

and therefore

$$h_\eta(t, \Delta) \leq C\varepsilon^{-2} C_*^{4/3} \left( h_\rho(t)^{\frac{2}{3} \cdot \frac{8}{27}} + h_\rho(t)^{1/6} \right) \Delta + (C\varepsilon^{-2} + C e^{C\varepsilon^{-2}}) \Delta \leq e^{C\varepsilon^{-2}} h_\rho(t)^{\frac{16}{81}} \Delta.$$

Applying Lemma 6.4 once more, we obtain

$$\int_{\mathbb{T}^3} |X(s) - y|^{-2} \rho_\varepsilon(y) \, dy \leq C C_*^{4/3} \left( h_\rho(t)^{\frac{2}{3} \cdot \frac{16}{81}} + h_\rho(t)^{1/6} \right) \Delta,$$

with  $C_* = e^{C\varepsilon^{-2}}$ , provided that

$$h_\rho(t)^{-8/81} \leq \Delta \leq t,$$

and therefore

$$h_\eta(t, \Delta) \leq C\varepsilon^{-2} C_*^{4/3} \left( h_\rho(t)^{\frac{2}{3} \cdot \frac{16}{81}} + h_\rho(t)^{1/6} \right) \Delta + (C\varepsilon^{-2} + Ce^{C\varepsilon^{-2}}) \Delta \leq e^{C\varepsilon^{-2}} h_\rho(t)^{\frac{1}{6}} \Delta, \quad (6.5)$$

since  $\frac{32}{243} < \frac{1}{6}$  and  $h_\rho \geq 1$ .

Finally, we use this new growth estimate on characteristic trajectories to control the mass density. Assuming that  $f_\varepsilon(0)$  is supported in  $\mathbb{T}^3 \times B_{\mathbb{R}^3}(0; R_t)$ , we have

$$h_\rho \leq C \|f_\varepsilon\|_{L^\infty(\mathbb{T}^d \times \mathbb{R}^d)} (R_0 + h_\eta(t, t))^3$$

Since we work with  $L^\infty(\mathbb{T}^d \times \mathbb{R}^d)$  solutions, we have a uniform bound on  $\|f_\varepsilon\|_{L^\infty(\mathbb{T}^d \times \mathbb{R}^d)}$  depending only on the initial data. Therefore, using (6.5), we find that if  $h_\rho(t)^{-8/81} \leq t$ ,

$$\begin{aligned} h_\rho &\leq C (R_0 + h_\eta(t, t))^3 \leq C \left( R_0 + e^{C\varepsilon^{-2}} h_\rho(t)^{\frac{1}{6}} t \right)^3 \\ &\leq CR_0^3 + e^{C\varepsilon^{-2}} h_\rho(t)^{1/2} t^3 \leq CR_0^3 + \frac{(e^{C\varepsilon^{-2}} t^3)^2 + h_\rho(t)}{2}. \end{aligned}$$

Hence

$$h_\rho \leq C \left( R_0^3 + e^{C\varepsilon^{-2}} t^6 \right).$$

If instead  $h_\rho(t)^{-8/81} \geq t$ , then

$$h_\rho(t) \leq t^{-\frac{81}{8}}.$$

Therefore, we may conclude that

$$h_\rho(t) \leq \max\{t^{-81/8}, C(R_0^3 + e^{C\varepsilon^{-2}} t^6)\}.$$

□

## 7 Quasineutral limit: proof of Theorem 2.2

In this section we will prove Theorem 2.2. The main idea is to consider the unique bounded density solution  $g_\varepsilon$  of the VPME system (1.3) with initial datum  $g_\varepsilon(0)$ . Since  $g_\varepsilon(0)$  is compactly supported, a bounded density solution exists by [20, Theorem 2.4] and is unique by [20, Theorem 2.3], so  $g_\varepsilon$  is well defined. We will use  $g_\varepsilon$  as a stepping stone between  $f_\varepsilon$  and a solution  $g$  of the KIE system (1.4).

We will need a monotonicity property of the Wasserstein distances with respect to the order  $p$ . We refer to [47] for proofs and further background.

**Lemma 7.1.** *Let  $\mu, \nu$  be probability measures on  $(\Omega, \mathcal{F})$ , each having a  $q$ th moment. Let  $p \leq q$ . Then*

$$W_p(\mu, \nu) \leq W_q(\mu, \nu).$$

*Proof of Theorem 2.2.* Let  $g_\varepsilon$  denote the solution of (1.3) with data  $g_\varepsilon(0)$ . We will use  $g_\varepsilon$  to interpolate between the solution  $f_\varepsilon$  of (1.3) starting from  $f_\varepsilon(0)$  and the solution  $g$  of (1.4) starting from  $g(0)$ . By the triangle inequality,

$$W_1(f_\varepsilon(t), g(t)) \leq W_1(f_\varepsilon(t), g_\varepsilon(t)) + W_1(g_\varepsilon(t), g(t)). \quad (7.1)$$

The quasineutral limit for the VPME system (1.3) with uniformly analytic data can be proved using the methods of Grenier [19], with the modifications for the massless electrons case described in [27, Proposition 4.1]. Grenier's result gives an  $H^s$  convergence of a representation of the VP system as a multi-fluid pressureless Euler system. In [27, Corollary 4.2], it is shown how this implies convergence in  $W_1$ . Since the initial data  $g_\varepsilon(0)$  satisfy Grenier's assumptions, there exists a solution  $g$  of (1.4) on a time interval  $[0, T_*]$  with initial data  $g(0)$  such that

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T_*]} W_1(g_\varepsilon(t), g(t)) = 0. \quad (7.2)$$

To deal with the first term of (7.1), we use a stability estimate around  $g_\varepsilon$  for the VPME system. By Proposition 5.1,

$$\begin{aligned} W_1(f_\varepsilon(t), g_\varepsilon(t)) &\leq W_2(f_\varepsilon(t), g_\varepsilon(t)) \\ &\leq \begin{cases} C \exp \left[ C \left( 1 + \log \frac{W_2(f_\varepsilon(0), g_\varepsilon(0))}{4\sqrt{d}} \right) e^{-C_\varepsilon t} \right] & \text{if } W_2(f_\varepsilon(0), g_\varepsilon(0)) \leq d \\ W_2(f_\varepsilon(0), g_\varepsilon(0)) e^{C_\varepsilon t} & \text{if } W_2(f_\varepsilon(0), g_\varepsilon(0)) > d. \end{cases} \end{aligned} \quad (7.3)$$

where  $C_\varepsilon$  may be chosen to satisfy

$$C_\varepsilon \leq \overline{\exp}_2(C\varepsilon^{-2})(M+1).$$

for any  $M$  satisfying

$$\sup_{[0, T_*]} \|\rho_{f_\varepsilon}(t)\|_{L_x^\infty}, \quad \sup_{[0, T_*]} \|\rho_{g_\varepsilon}(t)\|_{L_x^\infty} \leq M.$$

By Proposition 6.1, we may take  $M$  such that

$$M \leq C e^{C\varepsilon^{-2}}.$$

The constant  $C$  depends on  $T_*$ ,  $C_0$ , the dimension  $d$ , and  $C_1$ , the rate of growth of the initial support. We emphasise again that the appearance of an exponential rate here is a consequence of the form of the equation rather than because condition (2.4) allows fast growth of the initial support. It follows that we may estimate

$$C_\varepsilon t \leq C \overline{\exp}_2(C\varepsilon^{-2})$$

for all  $t \in [0, T_*]$ . Hence we have convergence if

$$\frac{\left| \log \frac{W_2(f_\varepsilon(0), g_\varepsilon(0))}{4\sqrt{d}} \right|}{\overline{\exp}_3(C\varepsilon^{-2})} \rightarrow \infty$$

as  $\varepsilon$  tends to zero. This holds if

$$W_2(f_\varepsilon(0), g_\varepsilon(0)) \leq (\overline{\exp}_4(C\varepsilon^{-2}))^{-1}$$

for sufficiently large  $C$ . In this case, it follows by (7.3) that

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0, T_*]} W_1(f_\varepsilon(t), g_\varepsilon(t)) = 0.$$

Combined with (7.2) and (7.1), this completes the proof. □

## 8 Mean field limit: proof of Theorem 2.4

In this section we will prove Theorem 2.4. Recall that we want to show that the empirical measure  $\mu_r^N$  corresponding to a solution of the regularised particle system (2.7) converges to a solution  $f$  of (1.2), in the Wasserstein sense:

$$\sup_{t \in [0, T]} W_2(f(t), \mu_r^N(t)) \rightarrow 0,$$

provided that the initial data  $\mu_r^N(0)$  converge sufficiently fast to  $f(0)$ .

The key observation is that the particle system (2.7) is constructed such that for each  $N$ ,  $\mu_r^N$  is a weak solution of a regularised version of the VPME system:

$$\begin{cases} \partial_t f_r + v \cdot \nabla_x f_r + E_r \cdot \nabla_v f_r = 0, \\ E_r = -\chi_r * \nabla U_r, \\ \Delta U_r = e^{U_r} - \chi_r * \rho[f_r], \\ f_r|_{t=0} = f(0) \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f(0) dx dv = 1. \end{cases} \quad (8.1)$$

The convergence of  $\mu_r^N$  to  $f$  is therefore a kind of stability result, in two stages.

First, as  $N$  tends to infinity with  $r$  fixed,  $\mu_r^N$  converges to  $f_r$ , the solution of (8.1) with initial datum  $f(0)$ . This holds because the force in (8.1) is regular enough that the equation has a stability property even in the class of measures (as investigated for example by Dobrushin [13]). Of course, the rate of this convergence will degenerate in the limit as  $r$  tends to zero. Our goal is therefore to quantify this convergence in  $W_2$ , optimising the constants so as to minimise the rate of blow-up as  $r$  tends to zero. This is the key difference between our approach, which is based on [31], and the approach of [13]. We use a weak-strong stability estimate, which holds for initial data converging sufficiently quickly, to improve the dependence of the constants on  $r$ .

Secondly, in [20] we proved that, if  $f(0)$  is bounded and compactly supported, then as  $r$  tends to zero,  $f_r$  converges to  $f$ , where  $f$  is the unique bounded density solution of the VPME system (1.2) with initial datum  $f(0)$ . This observation was part of our construction of solutions to (1.2). Our strategy will be to show that this convergence can be quantified in  $W_2$ , again aiming to optimise the rate.

By combining these two limits, we will identify a regime for the initial data in which  $\mu_r^N$  converges to  $f$ .

### 8.1 Behaviour of the $W_2$ distance under regularisation

We recall some useful results on the behaviour of Wasserstein distances under regularisation by convolution. See [47, Proposition 7.16] for proofs. Our first observation is that regularising two measures cannot increase the Wasserstein distance between them.

**Lemma 8.1.** *Let  $\mu, \nu$  be probability measures,  $r > 0$  any positive constant and  $\chi_r$  a mollifier as defined in (2.6). Then*

$$W_p(\chi_r * \mu, \chi_r * \nu) \leq W_p(\mu, \nu).$$

We also have an explicit control on the Wasserstein distance between a measure and its regularisation:

**Lemma 8.2.** *Let  $\mu$  be a probability measure and  $r > 0$ . Let  $\chi_r$  be a mollifier as defined in (2.6). Then*

$$W_p(\chi_r * \mu, \mu) \leq r.$$



If a measure  $\nu$  is close to an  $L^p$ -function in Wasserstein sense, then it is possible to estimate the  $L^p$ -norm of the regularised measure  $\chi_r * \nu$  in a way that exploits this fact. We will use this in our estimates to control the regularised mass density  $\chi_r * \rho_\mu$ . The following estimate is shown for  $p = \infty$  in [31, Lemma 4.3], but it is straightforward to adapt it to the case of general  $p$ .

**Lemma 8.3.** *Let  $\nu$  be a probability measure on  $\mathbb{T}^d$  and  $h \in L^\infty(\mathbb{T}^d)$  a probability density function. Then, for all  $r > 0$ ,  $p \in [1, \infty]$ ,  $q \in [1, \infty]$ ,*

$$\|\chi_r * \nu\|_{L^p(\mathbb{T}^d)} \leq C_d \left( \|h\|_{L^p(\mathbb{T}^d)} + r^{-(q+d)} W_q^q(h, \nu) \right).$$

## 8.2 The regularised VPME system

First, we must justify the existence of a unique solution for (8.1), under our assumptions on the initial datum. We recall a well-posedness result from [20, Lemma 6.1].

**Lemma 8.4** (Existence of regularised solutions). *For every  $f(0) \in \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d)$ , there exists a unique solution  $f_r \in C([0, \infty); \mathcal{P}(\mathbb{T}^d \times \mathbb{R}^d))$  of (8.1). If  $f(0) \in L^p(\mathbb{T}^d \times \mathbb{R}^d)$  for some  $p \in [1, \infty]$ , then for all  $t$*

$$\|f_r(t)\|_{L^p(\mathbb{T}^d \times \mathbb{R}^d)} \leq \|f(0)\|_{L^p(\mathbb{T}^d \times \mathbb{R}^d)}.$$

An important point about the regularisation approach we have chosen is that (8.1) has an associated conserved quantity. This is the regularised energy, defined by

$$\mathcal{E}_r[f] := \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f \, dx \, dv + \frac{1}{2} \int_{\mathbb{T}^d} |\nabla U_r|^2 \, dx + \int_{\mathbb{T}^d} U_r e^{U_r} \, dx. \quad (8.2)$$

For solutions with bounded mass and finite initial energy, the energy is conserved: if  $f_r$  is a solution of (8.1) with finite initial energy and  $\rho[f_r] \in L^\infty(0, T; L^\infty(\mathbb{T}^d))$ , then for all  $t \in [0, T]$ ,

$$\mathcal{E}_r[f_r(t)] = \mathcal{E}_r[f(0)].$$

Note that  $x e^x \geq -e^{-1}$  for all  $x \in \mathbb{R}$ . Hence if  $f_r$  is a solution of (8.1) with finite initial energy and  $\rho[f_r] \in L^\infty([0, T]; L^\infty(\mathbb{T}^d))$ , then for all  $t \in [0, T]$ ,

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f_r(t) \, dx \, dv \leq C (\mathcal{E}_r[f(0)] + 1).$$

Thus by Lemma 3.3,

$$\sup_{t \in [0, \infty)} \|\rho[f_r(t)]\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C (\mathcal{E}_r[f(0)] + 1). \quad (8.3)$$

Moreover, the growth estimates in Proposition 6.1 also apply to the regularised system. Hence, if the initial datum is compactly supported then the mass density is bounded on any compact time interval:  $\rho \in L_{\text{loc}}^\infty([0, \infty); L^\infty(\mathbb{T}^d))$ . In particular solutions of (8.1) beginning from compactly supported data have conserved energy.

## 8.3 $W_2$ stability for the regularised VPME system

In this section we follow the methods of [31] to prove a weak-strong stability estimate for the regularised equation (8.1) in the Wasserstein distance  $W_2$ , optimised to degenerate slowly as  $r$  tends to zero. This will allow us to use solutions of (8.1) as a bridge between the particle system (2.7) and the VPME system (1.2).

**Lemma 8.5** (Weak-strong stability for the regularised equation). *For each  $r > 0$ , let  $f_r, \mu_r$  be solutions of (8.1), where the  $f_r$  have uniformly bounded density and initial energy:*

$$\sup_r \sup_{t \in [0, T]} \|\rho_{f_r}\|_{L^\infty(\mathbb{T}^d)} \leq M, \quad (8.4)$$

$$\sup_r \mathcal{E}_r[f_r(0)] \leq C_0, \quad (8.5)$$

for some  $C_0, M > 0$ . Assume that the initial data satisfy, for some sufficiently large constant  $C > 0$  depending on  $T, C_0, M$  and  $d$ ,

$$\limsup_{r \rightarrow 0} \frac{W_2^2(f_r(0), \mu_r(0))}{r^{(d+2+C|\log r|^{-1/2})}} < 1.$$

Then

$$\lim_{r \rightarrow 0} \sup_{t \in [0, T]} W_2^2(f_r(t), \mu_r(t)) = 0.$$

*Proof.* To lighten the notation, we drop the subscript  $r$  from  $f_r$  and  $\mu_r$ . Fix an arbitrary coupling of the initial data  $\pi_0 \in \Pi(\mu_0, f_0)$ . As in the proof of Proposition 5.1, we define a special coupling that follows the characteristic flow of (8.1).

Consider the following systems of ordinary differential equations:

$$\begin{cases} \dot{X}_{x,v}^{(1)} = V_{x,v}^{(1)} \\ \dot{V}_{x,v}^{(1)} = E_r^{(\mu)}(X_{x,v}^{(1)}) \\ (X_{x,v}^{(1)}(0), V_{x,v}^{(1)}(0)) = (x, v) \\ E_r^{(\mu)} = -\chi_r * \nabla U_r^{(\mu)} \\ \Delta U_r^{(\mu)} = e^{U_r^{(\mu)}} - \chi_r * \rho_\mu \end{cases} \quad \begin{cases} \dot{X}_{x,v}^{(2)} = V_{x,v}^{(2)} \\ \dot{V}_{x,v}^{(2)} = E_r^{(f)}(X_{x,v}^{(2)}) \\ (X_{x,v}^{(2)}(0), V_{x,v}^{(2)}(0)) = (x, v) \\ E_r^{(f)} = -\chi_r * \nabla U_r^{(f)} \\ \Delta U_r^{(f)} = e^{U_r^{(f)}} - \chi_r * \rho_f. \end{cases} \quad (8.6)$$

First, we check that unique global solutions exist for both systems. The same argument applies to both  $\mu$  and  $f$ , so we will write it for  $\mu$  only. Observe that since  $\mu$  is a probability measure and  $\chi_r$  is smooth,  $\chi_r * \rho_\mu$  is a function with

$$\|\chi_r * \rho_\mu\|_{L^\infty(\mathbb{T}^d)} \leq \|\chi_r\|_{L^\infty(\mathbb{T}^d)}.$$

Hence by the regularity estimates in Proposition 4.1,  $U_r^{(\mu)}$  is a  $C^1$  function with

$$\|U_r^{(\mu)}\|_{C^1(\mathbb{T}^d)} \leq \exp \left[ C \left( 1 + \|\chi_r\|_{L^\infty(\mathbb{T}^d)} \right) \right]. \quad (8.7)$$

Then  $E_r^{(\mu)} = -\chi_r * \nabla U_r^{(\mu)}$  is a smooth function with bounded derivative

$$\|E_r^{(\mu)}\|_{C^1(\mathbb{T}^d)} \leq \|\chi_r\|_{C^1(\mathbb{T}^d)} \exp \left[ C \left( 1 + \|\chi_r\|_{L^\infty(\mathbb{T}^d)} \right) \right].$$

Therefore there is a unique  $C^1$  flow corresponding to (8.6).

As we argued in the proof of Proposition 5.1,  $\mu$  and  $f$  can be represented as the pushforwards of  $\mu(0)$  and  $f(0)$  respectively along these characteristic flows. That is, for  $\phi \in C_b(\mathbb{T}^d \times \mathbb{R}^d)$  we have

$$\begin{aligned} \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \, d\mu(t)(x, v) &= \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \left( X_{x,v}^{(1)}(t), V_{x,v}^{(1)}(t) \right) d\mu(0)(x, v), \\ \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \, f(t) \, dx \, dv &= \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \left( X_{x,v}^{(2)}(t), V_{x,v}^{(2)}(t) \right) f(0) \, dx \, dv. \end{aligned}$$

We define  $\pi_t$  to be the measure on  $(\mathbb{T}^d \times \mathbb{R}^d)^2$  such that for all  $\phi \in C_b [(\mathbb{T}^d \times \mathbb{R}^d)^2]$

$$\int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \phi(x_1, v_1, x_2, v_2) d\pi_t = \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi \left( X_{x_1, v_1}^{(1)}, V_{x_1, v_1}^{(1)}, X_{x_2, v_2}^{(2)}, V_{x_2, v_2}^{(2)} \right) d\pi_0. \quad (8.8)$$

Then  $\pi_t$  is a coupling of  $\mu(t)$  and  $f(t)$ . We can see this by using a calculation similar to (5.5) to check the marginals.

Using  $\pi_t$ , we define an anisotropic functional  $D$  which controls the squared Wasserstein distance  $W_2^2(f_t, \mu_t)$ . For  $\lambda > 0$ , let

$$D(t) = \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \lambda^2 |x_1 - x_2|^2 + |v_1 - v_2|^2 d\pi_t(x_1, v_1, x_2, v_2). \quad (8.9)$$

We will choose  $\lambda$  later in order to optimise our eventual rate. We note some relationships between  $D$  and the Wasserstein distance. By Definition 2.1, since  $\pi_t$  is a particular coupling of  $\mu_t$  and  $f_t$ , as long as we choose  $\lambda^2 > 1$ , we have

$$W_2^2(\mu_t, f_t) \leq D(t). \quad (8.10)$$

If we only look at the spatial variables, we can get a sharper estimate:

$$W_2^2(\rho_\mu(t), \rho_f(t)) \leq \lambda^{-2} D(t). \quad (8.11)$$

Since  $\pi_0 \in \Pi(\mu_0, f_0)$  was arbitrary we may take the infimum to obtain

$$\inf_{\pi_0} D(0) \leq \lambda^2 W_2^2(\mu_0, f_0). \quad (8.12)$$

We now perform a Grönwall estimate on  $D$ .  $D$  is differentiable with respect to time, since  $E_r$  is uniformly bounded (8.7) and  $\mu(0)$  and  $f(0)$  have finite second moments. Taking a time derivative, we obtain

$$\dot{D} = 2 \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \lambda^2 (x_1 - x_2) \cdot (v_1 - v_2) + (v_1 - v_2) \cdot \left( E_r^{(\mu)}(x_1) - E_r^{(f)}(x_2) \right) d\pi_t.$$

Using a weighted Cauchy inequality, we find that for any  $\alpha > 0$ ,

$$\begin{aligned} \dot{D} &= \lambda \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \lambda^2 |x_1 - x_2|^2 + |v_1 - v_2|^2 d\pi_t \\ &\quad + \alpha \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |v_1 - v_2|^2 d\pi_t + \frac{1}{\alpha} \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \left| E_r^{(\mu)}(x_1) - E_r^{(f)}(x_2) \right|^2 d\pi_t. \end{aligned}$$

Therefore

$$\dot{D} \leq (\alpha + \lambda) D + \frac{C}{\alpha} \sum_{i=1}^4 I_i,$$

where

$$\begin{aligned} I_1 &:= \int |\bar{E}_r^{(\mu)}(X_t^{(1)}) - \bar{E}_r^{(\mu)}(X_t^{(2)})|^2 d\pi_0, & I_2 &:= \int |\bar{E}_r^{(f)}(X_t^{(2)}) - \bar{E}_r^{(\mu)}(X_t^{(2)})|^2 d\pi_0, \\ I_3 &:= \int |\widehat{E}_r^{(\mu)}(X_t^{(1)}) - \widehat{E}_r^{(\mu)}(X_t^{(2)})|^2 d\pi_0, & I_4 &:= \int |\widehat{E}_r^{(f)}(X_t^{(2)}) - \widehat{E}_r^{(\mu)}(X_t^{(2)})|^2 d\pi_0. \end{aligned}$$

We have again used the decomposition  $E_r^{(f)} = \bar{E}_r^{(f)} + \widehat{E}_r^{(f)}$ , and the analogous form for  $E_r^{(\mu)}$ .

To estimate these quantities, we first note some basic regularity properties, which follow from Proposition 4.1. First, we wish to control the regularised mass density  $\chi_r * \rho_\mu$  with an estimate that behaves well as  $r$  tends to zero. For this we use Lemma 8.3 with  $q = 2$  and (8.11):

$$\begin{aligned} \|\chi_r * \rho_\mu\|_{L^p(\mathbb{T}^d)} &\leq C_d \left( \|\rho_f\|_{L^p(\mathbb{T}^d)} + r^{-(d+2)} W_2^2(\rho_\mu, \rho_f) \right) \\ &\leq C_d \left( \|\rho_f\|_{L^p(\mathbb{T}^d)} + r^{-(d+2)} \lambda^{-2} D \right). \end{aligned} \quad (8.13)$$

We will use this estimate in the cases  $p = \frac{d+2}{d}$  and  $p = \infty$ . For  $p = \infty$ , by assumption (8.4) we obtain

$$\|\chi_r * \rho_\mu\|_{L^\infty(\mathbb{T}^d)} \leq C_d \left( M + r^{-(d+2)} \lambda^{-2} D \right). \quad (8.14)$$

For  $p = \frac{d+2}{d}$ , by the initial energy assumption (8.5) and (8.3) we have

$$\|\rho_f\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C_d,$$

where  $C_d$  depends on  $C_0$  and  $d$ . Thus

$$\|\chi_r * \rho_\mu\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C_d \left( 1 + r^{-(d+2)} \lambda^{-2} D \right) \quad (8.15)$$

for  $C_d$  depending on  $C_0$  and  $d$ .

We also wish to control the regularity of  $\widehat{U}_r^{(\mu)}$ . Using (8.15) and Proposition 4.1, we obtain

$$\|\widehat{U}_r^{(\mu)}\|_{C^2(\mathbb{T}^d)} \leq C_d \overline{\exp}_2 \left( C_d \left( 1 + r^{-(d+2)} \lambda^{-2} D \right) \right). \quad (8.16)$$

We estimate  $I_1$  and  $I_2$  in the same way as in [31]. For  $I_1$  we use the regularity estimate

$$\|\chi_r * K * h\|_{\text{Lip}} \leq C |\log r| (1 + \|h\|_{L^\infty}); \quad (8.17)$$

see [31, Lemma 4.2(ii)]. We combine this with the mass density estimate (8.14) to obtain

$$I_1 \leq C (\log r)^2 \left( M + r^{-(d+2)} \lambda^{-2} D \right)^2 \lambda^{-2} D. \quad (8.18)$$

For  $I_2$ , we use Proposition 4.4 and (8.14) to obtain

$$I_2 \leq CM (M + r^{-(d+2)} \lambda^{-2} D) \lambda^{-2} D. \quad (8.19)$$

For  $I_3$  we compute:

$$\begin{aligned} I_3 &= \int |\widehat{E}_r^{(\mu)}(X_t^{(1)}) - \widehat{E}_r^{(\mu)}(X_t^{(2)})|^2 d\pi_0 = \int |\chi_r * (\nabla \widehat{U}_r^{(\mu)}(x) - \nabla \widehat{U}_r^{(\mu)}(y))|^2 d\pi_t \\ &\leq \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \|\chi_r * \nabla \widehat{U}_r^{(\mu)}\|_{\text{Lip}}^2 |x - y|^2 d\pi_t \leq \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \|\widehat{U}_r^{(\mu)}\|_{C^2(\mathbb{T}^d)}^2 |x - y|^2 d\pi_t. \end{aligned} \quad (8.20)$$

We apply the regularity estimate (8.16) to obtain

$$\begin{aligned} I_3 &\leq C \overline{\exp}_2 \left[ C_d (1 + r^{-(d+2)} \lambda^{-2} D) \right] \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |x - y|^2 d\pi_t \\ &\leq C \overline{\exp}_2 \left[ C_d (1 + r^{-(d+2)} \lambda^{-2} D) \right] \lambda^{-2} D. \end{aligned}$$

For  $I_4$  we compute

$$\begin{aligned} I_4 &= \int |\widehat{E}_r^{(f)}(X_t^{(2)}) - \widehat{E}_r^{(\mu)}(X_t^{(2)})|^2 d\pi_0 = \int |\chi_r * (\nabla \widehat{U}_r^{(f)}(X_t^{(2)}) - \nabla \widehat{U}_r^{(\mu)}(X_t^{(2)}))|^2 d\pi_0 \\ &= \int_{\mathbb{T}^d} |\chi_r * (\nabla \widehat{U}_r^{(f)}(x) - \nabla \widehat{U}_r^{(\mu)}(x))|^2 \rho_f(x) dx \leq \|\rho_f\|_{L^\infty(\mathbb{T}^d)} \|\nabla \widehat{U}_r^{(f)} - \nabla \widehat{U}_r^{(\mu)}\|_{L^2(\mathbb{T}^d)}^2 \end{aligned} \quad (8.21)$$

By Proposition 4.4,

$$\begin{aligned} \|\nabla \widehat{U}_r^{(f)} - \nabla \widehat{U}_r^{(\mu)}\|_{L^2(\mathbb{T}^d)}^2 &\leq C \overline{\text{exp}}_2 \left[ C_d (1 + r^{-(d+2)} \lambda^{-2} D) \right] \|\bar{U}_r^{(f)} - \bar{U}_r^{(\mu)}\|_{L^2(\mathbb{T}^d)}^2 \\ &\leq C \overline{\text{exp}}_2 \left[ C_d (1 + r^{-(d+2)} \lambda^{-2} D) \right] (M + r^{-(d+2)} \lambda^{-2} D) \lambda^{-2} D, \end{aligned}$$

thus

$$I_4 \leq CM \overline{\text{exp}}_2 \left[ C_d (1 + r^{-(d+2)} \lambda^{-2} D) \right] (M + r^{-(d+2)} \lambda^{-2} D) \lambda^{-2} D.$$

We summarise all these bounds as

$$\begin{aligned} \dot{D} &\leq (\lambda + \alpha) D + \frac{1}{4\alpha} C (1 + |\log r|)^2 \left( M + r^{-(d+2)} \lambda^{-2} D \right)^2 \lambda^{-2} D \\ &\quad + C \overline{\text{exp}}_2 \left[ C_d (1 + r^{-(d+2)} \lambda^{-2} D) \right] \left( 1 + M (M + r^{-(d+2)} \lambda^{-2} D) \right) \lambda^{-2} D. \end{aligned} \quad (8.22)$$

Unfortunately, this is a nonlinear estimate and so cannot be closed in its current form. To deal with this, we introduce a truncated functional, rescaled to be of order 1:

$$\hat{D} = 1 \wedge \left( r^{-(d+2)} \lambda^{-2} D \right)$$

Since  $D$  is differentiable,  $\hat{D}$  is at least Lipschitz. In particular, at almost all times  $t$ ,  $\hat{D}$  is differentiable, with either

$$\frac{d}{dt} \hat{D}(t) = 0 \quad \text{or} \quad \frac{d}{dt} \hat{D}(t) = r^{-(d+2)} \lambda^{-2} \frac{d}{dt} D(t).$$

Thus (8.22) implies that, for almost all  $t$ ,

$$\frac{d}{dt} \hat{D} \leq (\lambda + \alpha) \hat{D} + \frac{1}{\alpha} C_d [1 + |\log r|^2] M^2 \lambda^{-2} \hat{D}.$$

We then optimise the exponent by choosing

$$\alpha_* = C_d (1 + |\log r|) M \lambda_*^{-1}, \quad \lambda_* = C_d (1 + |\log r|)^{1/2} \sqrt{M}. \quad (8.23)$$

Then

$$\frac{d}{dt} \hat{D} \leq C_d \lambda_* \hat{D}.$$

Since  $\hat{D}$  is absolutely continuous,

$$\hat{D}(t) = \hat{D}(0) + \int_0^t g(s) ds,$$

where the function  $g(t) = \frac{d}{dt} \hat{D}(t)$  wherever this derivative exists, and  $g(t) = 0$  otherwise. Thus

$$\hat{D}(t) \leq \hat{D}(0) + C_d \lambda_* \int_0^t \hat{D}(s) ds.$$

Then by Grönwall's inequality we deduce

$$\begin{aligned} \sup_{[0,T]} \hat{D}(t) &\leq \exp(C_d \lambda_* T) \hat{D}(0) \leq r^{-(d+2)} \lambda_*^{-2} \exp(C_d \lambda_* T) D(0) \\ &\leq C_d \exp \left[ |\log r| \left( (d+2) + C_d T \sqrt{M} |\log r|^{-1/2} \right) \right] \lambda_*^{-2} D(0). \end{aligned}$$

In order to use this estimate to control the Wasserstein distance, we need to ensure that for  $r$  sufficiently small,

$$\inf_{\pi_0} \sup_{t \in [0,T]} \hat{D}(t) < 1. \quad (8.24)$$

If (8.24) holds, then by (8.10), for all  $t \in [0, T]$ ,

$$W_2^2(\mu_t, f_t) \leq r^{d+2} \lambda_*^2 \inf_{\pi_0} \sup_{t \in [0,T]} \hat{D}(t) \leq r^{d+2} \lambda_*^2 \rightarrow 0$$

as  $r$  tends to zero, since  $\lambda_*^2$  only grows like  $|\log r|$  by definition (8.23). Using (8.12), we obtain for any  $\pi_0$ ,

$$\begin{aligned} \sup_{[0,T]} \hat{D}(t) &\leq C_d \exp \left[ |\log r| \left( (d+2) + C_d T \sqrt{M} |\log r|^{-1/2} \right) \right] \lambda_*^{-2} \inf_{\pi_0} D(0) \\ &\leq C_d \exp \left[ |\log r| \left( (d+2) + C_d T \sqrt{M} |\log r|^{-1/2} \right) \right] W_2^2(\mu_0, f_0). \end{aligned}$$

Since we assumed that the initial data satisfy

$$\limsup_{r \rightarrow 0} \frac{W_2^2(\mu_0, f_0)}{r^{d+2+C_{d,M,T} |\log r|^{-1/2}}} < 1,$$

for large  $C_{d,M,T}$ , (8.24) holds for sufficiently small  $r$ . This completes the proof.  $\square$

## 8.4 Convergence to the original equation

**Lemma 8.6** (Approximation of (VPME)). *Fix  $f(0) \in L^1 \cap L^\infty(\mathbb{T}^d \times \mathbb{R}^d)$  satisfying a uniform bound on the energy as defined in (8.2):*

$$\sup_r \mathcal{E}_r[f(0)] \leq C_0, \quad (8.25)$$

for some  $C_0 > 0$ .

For each  $r > 0$ , let  $f_r$  be a solution of (8.1) with initial datum  $f(0)$ . Let  $f$  be a solution of (1.2) with the same initial datum  $f(0)$ . Assume that  $(f_r)_{r>0}$ ,  $f$  have uniformly bounded density:

$$\sup_r \sup_{t \in [0,T]} \|\rho_{f_r}\|_{L^\infty(\mathbb{T}^d)} \leq M, \quad \sup_{t \in [0,T]} \|\rho_f\|_{L^\infty(\mathbb{T}^d)} \leq M$$

for some  $M > 0$ . Then

$$\lim_{r \rightarrow 0} \sup_{t \in [0,T]} W_2^2(f(t), f_r(t)) = 0.$$

*Proof.* We fix an initial coupling  $\pi_0 \in \Pi(f(0), f(0))$  and construct  $\pi_t \in \Pi(f(t), f_r(t))$  as in (8.8), using the characteristic flows

$$\begin{cases} \dot{X}_{x,v}^{(1)} = V_{x,v}^{(1)} \\ \dot{V}_{x,v}^{(1)} = E(X_{x,v}^{(1)}) \\ (X_{x,v}^{(1)}(0), V_{x,v}^{(1)}(0)) = (x, v) \\ E = -\nabla U \\ \Delta U = e^U - \rho_f \end{cases} \quad \begin{cases} \dot{X}_{x,v}^{(2)} = V_{x,v}^{(2)} \\ \dot{V}_{x,v}^{(2)} = E_r(X_{x,v}^{(2)}) \\ (X_{x,v}^{(2)}(0), V_{x,v}^{(2)}(0)) = (x, v) \\ E_r = -\chi_r * \nabla U_r \\ \Delta U_r = e^{U_r} - \chi_r * \rho_{f_r}. \end{cases} \quad (8.26)$$

We define  $D$  as in (8.9). As in Lemma 8.5, we obtain for any  $\alpha > 0$ ,

$$\dot{D} \leq (\alpha + \lambda)D + \frac{C}{\alpha} \sum_{i=1}^4 I_i,$$

where

$$\begin{aligned} I_1 &:= \int |\bar{E}_r(X_t^{(1)}) - \bar{E}_r(X_t^{(2)})|^2 d\pi_0, & I_2 &:= \int |\bar{E}_r(X_t^{(1)}) - \bar{E}(X_t^{(1)})|^2 d\pi_0, \\ I_3 &:= \int |\hat{E}_r(X_t^{(1)}) - \hat{E}_r(X_t^{(2)})|^2 d\pi_0, & I_4 &:= \int |\hat{E}(X_t^{(1)}) - \hat{E}_r(X_t^{(1)})|^2 d\pi_0. \end{aligned}$$

For  $I_1$  we use the regularity estimate (8.17) to deduce

$$I_1 \leq |\log r|^2 M^2 \int |X_t^{(1)} - X_t^{(2)}|^2 d\pi_0 \leq |\log r|^2 M^2 \lambda^{-2} D. \quad (8.27)$$

For  $I_2$  we use the stability estimate from Lemma 4.5:

$$\|\nabla \bar{U} - \nabla \bar{U}_r\|_{L^2(\mathbb{T}^d)} \leq \sqrt{M} W_2(\chi_r * \rho_{f_r}, \rho_f).$$

By Lemma 8.2,

$$W_2(\chi_r * \rho_{f_r}, \rho_f) \leq r + W_2(\rho_{f_r}, \rho_f).$$

Hence

$$\|\nabla \bar{U} - \nabla \bar{U}_r\|_{L^2(\mathbb{T}^d)} \leq \sqrt{M}(r + \lambda^{-1} \sqrt{D}). \quad (8.28)$$

We account for the extra regularisation by elementary methods: first, for  $g \in C^1(\mathbb{T}^d)$ ,

$$\begin{aligned} \|g - \chi_r * g\|_{L^2(\mathbb{T}^d)}^2 &= \int_{\mathbb{T}^d} \left| \int_{\mathbb{T}^d} \chi_r(y) [g(x) - g(x-y)] dy \right|^2 dx \\ &\leq \int_{\mathbb{T}^d} \left| \int_{\mathbb{T}^d} \int_0^1 -y \chi_r(y) \nabla g(x-hy) dh dy \right|^2 dx. \end{aligned}$$

Hence, by Jensen's inequality applied to the probability measure  $\chi_r(y) dy dh$  on  $\mathbb{T}^d \times [0, 1]$ ,

$$\begin{aligned} \|g - \chi_r * g\|_{L^2(\mathbb{T}^d)}^2 &\leq \int_0^1 \int_{\mathbb{T}^d} |y|^2 \chi_r(y) \int_{\mathbb{T}^d} |\nabla g(x-hy)|^2 dx dy dh \\ &\leq r^2 \|\nabla g\|_{L^2(\mathbb{T}^d)}^2 \int_{\mathbb{T}^d} |y|^2 \chi_1(y) dy \leq Cr^2 \|\nabla g\|_{L^2(\mathbb{T}^d)}^2. \end{aligned} \quad (8.29)$$

This estimate extends by density to  $g \in W^{1,2}(\mathbb{T}^d)$ .

Next, standard estimates for the Poisson equation and  $L^p$  interpolation inequalities imply that

$$\|\nabla^2 \bar{U}_r\|_{L^2(\mathbb{T}^d)} \leq C \|\rho_{f_r}\|_{L^2(\mathbb{T}^d)} \leq C \|\rho_{f_r}\|_{L^\infty(\mathbb{T}^d)}^{\frac{d-2}{2d}} \|\rho_{f_r}\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)}^{\frac{d+2}{2d}} \leq CM^{\frac{d-2}{2d}}, \quad (8.30)$$

since by (8.25) and (8.3) we have

$$\|\rho_{f_r}\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C \quad (8.31)$$

for some  $C$  depending on  $d$  and  $C_0$  only.

Therefore, using (8.29),

$$\|\nabla \bar{U}_r - \chi_r * \nabla \bar{U}_r\|_{L^2(\mathbb{T}^d)} \leq C \|\nabla^2 \bar{U}_r\|_{L^2(\mathbb{T}^d)} r \leq CM^{\frac{d-2}{2d}} r,$$

and we conclude that

$$I_2 \leq CM^2(r + \lambda^{-1}\sqrt{D})^2.$$

The term  $I_3$  is estimated like  $I_1$ , using the regularity estimate from Proposition 4.1 and (8.31):

$$\|\widehat{U}_r\|_{C^{2,\alpha}(\mathbb{T}^d)} \leq C,$$

where  $C$  depends only on the constant  $C_0$  controlling the initial energy in (8.25). We obtain

$$I_3 \leq C \int |X_t^{(1)} - X_t^{(2)}|^2 d\pi_0 \leq C \lambda^{-2} D.$$

Finally,  $I_4$  is estimated in the same way as  $I_2$ , using the stability estimate Lemma 4.6, Proposition 4.1 and (8.31):

$$\|\nabla\widehat{U} - \nabla\widehat{U}_r\|_{L^2(\mathbb{T}^d)} \leq C\|\nabla\bar{U} - \nabla\bar{U}_r\|_{L^2(\mathbb{T}^d)} \leq C\sqrt{M}(r + \lambda^{-1}\sqrt{D}).$$

By Proposition 4.1 and (8.31), we have

$$\|U_r\|_{L^\infty} \leq \|\bar{U}_r\|_{C^{0,\alpha}(\mathbb{T}^d)} + \|\widehat{U}_r\|_{C^{1,\alpha}(\mathbb{T}^d)} \leq \exp\left(C\left(1 + \|\chi_r * \rho_{f_r}\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)}\right)\right) \leq C,$$

where  $C$  depends on  $C_0$  and  $d$  only. Hence

$$\|e^{U_r}\|_{L^2(\mathbb{T}^d)} \leq C.$$

Since  $\Delta\widehat{U}_r = e^{U_r} - 1$ , by standard regularity results for the Poisson equation this implies that

$$\|D^2\widehat{U}_r\|_{L^2(\mathbb{T}^d)} \leq C.$$

Therefore, using (8.29) again,

$$\|\nabla\widehat{U}_r - \chi_r * \nabla\widehat{U}_r\|_{L^2(\mathbb{T}^d)} \leq C\|D^2\widehat{U}_r\|_{L^2(\mathbb{T}^d)} r \leq Cr,$$

and we conclude that

$$I_4 \leq CM^2(r + \lambda^{-1}\sqrt{D})^2.$$

Altogether we have

$$\dot{D} \leq (\alpha + \lambda)D + \frac{C}{\alpha\lambda^2} (|\log r|^2 + 1) M^2 D + CM^2 r^2.$$

Optimising the exponent, we deduce that

$$\dot{D} \leq C\lambda_* D + CM^2 r^2,$$

where

$$\lambda_* = \sqrt{M}(1 + |\log r|)^{1/4}.$$

Hence

$$\begin{aligned} D(t) &\leq (D(0) + CM^2 r^2) \exp\left[C\sqrt{M}(1 + |\log r|)^{1/4} t\right] \\ &\leq (D(0) + CM^2 r^2) e^{C\sqrt{M|\log r|} t}. \end{aligned}$$

Since  $D$  controls the squared Wasserstein distance (8.10),

$$W_2^2(f_r(t), f(t)) \leq (D(0) + CM^2 r^2) e^{C\sqrt{M|\log r|} t}.$$



Then, by (8.12), and since  $f$  and  $f_r$  share the same initial datum,

$$\begin{aligned} W_2^2(f_r(t), f(t)) &\leq (\inf_{\pi_0} D(0) + CM^2r^2)e^{C\sqrt{M|\log r|}t} \\ &\leq (\lambda_*^2 W_2^2(f(0), f(0)) + CM^2r^2)e^{C\sqrt{M|\log r|}t} \\ &\leq CM^2r^2e^{C\sqrt{M|\log r|}t}. \end{aligned}$$

Since

$$r^2e^{C\sqrt{M|\log r|}t} = r^{2-C\sqrt{\frac{Mt^2}{|\log r|}}},$$

we conclude that for any compact time interval  $[0, T]$ ,

$$\lim_{r \rightarrow 0} \sup_{t \in [0, T]} W_2^2(f_r(t), f(t)) = 0.$$

□

*Proof of Theorem 2.4.* Let  $f_r$  be the unique solution of (8.1) with initial datum  $f_0$ . By the triangle inequality for  $W_2$ , we have

$$\sup_{t \in [0, T]} W_2^2(\mu_r^N(t), f(t)) \leq \sup_{t \in [0, T]} W_2^2(\mu_r^N(t), f_r(t)) + \sup_{t \in [0, T]} W_2^2(f_r(t), f(t)).$$

We apply Lemma 8.5 to the first term using the assumption on the initial configurations and deduce that it converges to zero as  $r$  tends to zero. For the second term we apply Lemma 8.6, since  $f$  and  $f_r$  have the same initial datum. This completes the proof.

□

## 9 Combined quasineutral and mean field limit: proof of Theorem 2.5

In this section, we prove Theorem 2.5. The idea is to use the scaled Vlasov-Poisson equation (1.3) as a bridge between the particle system (2.9) and the kinetic isothermal Euler system (1.4). This approach was previously used for the classical Vlasov-Poisson system in [21]. In order to pass from the particle system to the Vlasov-Poisson system, we need to revisit the estimates from Section 8 in order to quantify the dependence of the constants on  $\varepsilon$ .

We will be using a scaled version of the regularised VPME system (8.1):

$$\left\{ \begin{array}{l} \partial_t f_\varepsilon^{(r)} + v \cdot \nabla_x f_\varepsilon^{(r)} + E_{\varepsilon, r} \cdot \nabla_v f_\varepsilon^{(r)} = 0, \\ E_{\varepsilon, r} = -\chi_r * \nabla U_{\varepsilon, r}, \\ \varepsilon^2 \Delta U_{\varepsilon, r} = e^{U_{\varepsilon, r}} - \chi_r * \rho[f_\varepsilon^{(r)}], \\ f_\varepsilon^{(r)}|_{t=0} = f_\varepsilon(0) \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f_\varepsilon(0) dx dv = 1. \end{array} \right. \quad (9.1)$$

We begin with a few remarks about this system. The well-posedness theory is clearly the same as for the case  $\varepsilon = 1$ . The associated energy needs to be scaled with  $\varepsilon$  in the following way:

$$\mathcal{E}_{\varepsilon, r}[f] := \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |v|^2 f dx dv + \frac{\varepsilon^2}{2} \int_{\mathbb{T}^d} |\nabla U_r|^2 dx + \int_{\mathbb{T}^d} U_r e^{U_r} dx.$$

The conservation of this energy implies that solutions with bounded mass density and finite initial energy satisfy

$$\sup_{t \in [0, \infty)} \|\rho[f_\varepsilon^{(r)}(t)]\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C \left( \mathcal{E}_{\varepsilon, r}[f_\varepsilon^{(r)}(0)] + 1 \right), \quad (9.2)$$

where  $C$  is independent of  $\varepsilon$ . This follows from the same argument as we gave before (8.3).

The following lemma is a quantified version of Lemma 8.5.

**Lemma 9.1** (Weak-strong stability for the regularised equation, with quasineutral scaling). *Let  $f_\varepsilon(0)$  satisfy the assumptions of Theorem 2.2. Let  $f_\varepsilon^{(r)}$  be the solution of (9.1) with data  $f_\varepsilon(0)$ . Fix any  $T > 0$ . Then there exists a constant  $K$  depending on  $T$  and  $\{f_\varepsilon(0)\}$  such that the following holds. For each  $(\varepsilon, r)$ , let  $\mu_\varepsilon^{(r)}$  be a measure solution of (9.1). Assume that for some  $\eta > 0$ ,  $r = r(\varepsilon)$  satisfies*

$$r \leq [\overline{\text{exp}}_3(K\varepsilon^{-2})]^{-1} \quad \text{and} \quad \lim_{r \rightarrow 0} \frac{W_2\left(\mu_\varepsilon^{(r)}(0), f_\varepsilon(0)\right)}{r^{(d+2+\eta)/2}} = 0. \quad (9.3)$$

Then

$$\lim_{r \rightarrow 0} \sup_{t \in [0, T]} W_2\left(\mu_\varepsilon^{(r)}(t), f_\varepsilon^{(r)}(t)\right) = 0.$$

*Proof.* For ease of notation we drop the sub and superscripts on  $f_\varepsilon^{(r)}$  and  $\mu_\varepsilon^{(r)}$ . We use the same method as for Lemma 8.5, tracking the dependence on  $\varepsilon$ . Let  $\pi_0 \in \Pi(f_0, \mu_0)$ . We define a time dependent coupling  $\pi_t$  as in (8.8) with  $(X_{x,v}^{(i)}, V_{x,v}^{(i)})$  being the characteristic flows corresponding to  $f$  and  $\mu$ :

$$\begin{cases} \dot{X}_{x,v}^{(1)} = V_{x,v}^{(1)} \\ \dot{V}_{x,v}^{(1)} = E_{\varepsilon,r}^{(\mu)}(X_{x,v}^{(1)}) \\ (X_{x,v}^{(1)}(0), V_{x,v}^{(1)}(0)) = (x, v) \\ E_{\varepsilon,r}^{(\mu)} = -\chi_r * \nabla U_{\varepsilon,r}^{(\mu)} \\ \varepsilon^2 \Delta U_{\varepsilon,r}^{(\mu)} = e^{U_{\varepsilon,r}^{(\mu)}} - \chi_r * \rho_\mu \end{cases} \quad \begin{cases} \dot{X}_{x,v}^{(2)} = V_{x,v}^{(2)} \\ \dot{V}_{x,v}^{(2)} = E_{\varepsilon,r}^{(f)}(X_{x,v}^{(2)}) \\ (X_{x,v}^{(2)}(0), V_{x,v}^{(2)}(0)) = (x, v) \\ E_{\varepsilon,r}^{(f)} = -\chi_r * \nabla U_{\varepsilon,r}^{(f)} \\ \varepsilon^2 \Delta U_{\varepsilon,r}^{(f)} = e^{U_{\varepsilon,r}^{(f)}} - \chi_r * \rho_f \end{cases}$$

Then, since  $f$  and  $\mu$  are the pushforwards of  $f_0$  and  $\mu_0$  along their respective characteristic flows,  $\pi_t$  is a coupling of  $f_t$  and  $\mu_t$ .

Again we define an anisotropic functional  $D$ :

$$D(t) = \frac{1}{2} \int \lambda^2 |x - y|^2 + |v - w|^2 \, d\pi_t(x, v, y, w).$$

As before,

$$\dot{D} \leq (\alpha + \lambda)D + \frac{C}{\alpha} \sum_{i=1}^4 I_i,$$

where

$$\begin{aligned} I_1 &:= \int |\bar{E}_{\varepsilon,r}^{(\mu)}(X_t^{(1)}) - \bar{E}_{\varepsilon,r}^{(\mu)}(X_t^{(2)})|^2 \, d\pi_0, & I_2 &:= \int |\bar{E}_{\varepsilon,r}^{(f)}(X_t^{(2)}) - \bar{E}_{\varepsilon,r}^{(\mu)}(X_t^{(2)})|^2 \, d\pi_0, \\ I_3 &:= \int |\hat{E}_{\varepsilon,r}^{(\mu)}(X_t^{(1)}) - \hat{E}_{\varepsilon,r}^{(\mu)}(X_t^{(2)})|^2 \, d\pi_0, & I_4 &:= \int |\hat{E}_{\varepsilon,r}^{(f)}(X_t^{(2)}) - \hat{E}_{\varepsilon,r}^{(\mu)}(X_t^{(2)})|^2 \, d\pi_0. \end{aligned}$$

To estimate these quantities, we first note some basic  $L^p(\mathbb{T}^d)$  estimates on the regularised mass density  $\chi_r * \rho_\mu$ , using (8.13). For  $p = \frac{d+2}{d}$ , since  $f_\varepsilon(0)$  satisfies (2.3), by (9.2) we obtain

$$\|\rho_f\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C,$$

where  $C$  depends on  $C_0$  and  $d$  only. Thus

$$\|\chi_r * \rho_\mu\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)} \leq C_d \left(1 + r^{-(d+2)} \lambda^{-2} D\right), \quad (9.4)$$

for  $C_d$  depending on  $C_0$  and  $d$ .

For  $p = \infty$  we obtain

$$\|\chi_r * \rho_\mu\|_{L^\infty(\mathbb{T}^d)} \leq C_d \left( M_\varepsilon + r^{-(d+2)} \lambda^{-2} D \right). \quad (9.5)$$

where  $M_\varepsilon$  is a constant such that

$$\sup_{t \in [0, T]} \|\rho_{f_t}\|_{L^\infty(\mathbb{T}^d)} \leq M_\varepsilon.$$

By Proposition 6.1 (which also applies to the regularised system), there exists a constant  $C$  depending on  $T$  such that

$$\sup_{t \in [0, T]} \|\rho_{f_t}\|_{L^\infty(\mathbb{T}^d)} \leq \exp(C\varepsilon^{-2}).$$

Therefore  $M_\varepsilon$  may be chosen to satisfy

$$M_\varepsilon \leq \exp(C\varepsilon^{-2}). \quad (9.6)$$

We estimate  $I_1$  as in (8.18). An extra factor of  $\varepsilon^{-4}$  appears due to the quasineutral scaling on the force:

$$I_1 \leq C\varepsilon^{-4} (\log r)^2 \left( M_\varepsilon + r^{-(d+2)} \lambda^{-2} D \right)^2 \lambda^{-2} D.$$

Similarly,  $I_2$  is estimated as in (8.19) using Lemma 4.5 and (9.5) to obtain

$$I_2 \leq C\varepsilon^{-4} M_\varepsilon (M_\varepsilon + r^{-(d+2)} \lambda^{-2} D) \lambda^{-2} D.$$

For  $I_3$  the same computation as in (8.20) implies that

$$I_3 \leq \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} \|\widehat{U}_{\varepsilon, r}(\mu)\|_{C^2(\mathbb{T}^d)}^2 |x - y|^2 d\pi_t.$$

Using (9.4) we apply Proposition 4.1 to obtain

$$\|\widehat{U}_{\varepsilon, r}(\mu)\|_{C^2(\mathbb{T}^d)} \leq C_d \overline{\exp}_2 \left( C_d \varepsilon^{-2} \left( 1 + r^{-(d+2)} \lambda^{-2} D \right) \right).$$

Thus

$$\begin{aligned} I_3 &\leq C \overline{\exp}_2 \left[ C_d \varepsilon^{-2} (1 + r^{-(d+2)} \lambda^{-2} D) \right] \int_{(\mathbb{T}^d \times \mathbb{R}^d)^2} |x - y|^2 d\pi_t \\ &\leq C \overline{\exp}_2 \left[ C_d \varepsilon^{-2} (1 + r^{-(d+2)} \lambda^{-2} D) \right] \lambda^{-2} D. \end{aligned}$$

For  $I_4$ , as in (8.21) we obtain

$$I_4 \leq \|\rho_f\|_{L^\infty(\mathbb{T}^d)} \|\nabla \widehat{U}_{\varepsilon, r}^{(f)} - \nabla \widehat{U}_{\varepsilon, r}^{(\mu)}\|_{L^2(\mathbb{T}^d)}^2.$$

By Lemma 4.6 and (9.5),

$$\begin{aligned} \|\nabla \widehat{U}_{\varepsilon, r}^{(f)} - \nabla \widehat{U}_{\varepsilon, r}^{(\mu)}\|_{L^2(\mathbb{T}^d)}^2 &\leq C\varepsilon^{-2} \overline{\exp}_2 \left[ C_d \varepsilon^{-2} (1 + r^{-(d+2)} \lambda^{-2} D) \right] \|\bar{U}_{\varepsilon, r}^{(f)} - \bar{U}_{\varepsilon, r}^{(\mu)}\|_{L^2(\mathbb{T}^d)}^2 \\ &\leq C\varepsilon^{-6} \overline{\exp}_2 \left[ C_d \varepsilon^{-2} (1 + r^{-(d+2)} \lambda^{-2} D) \right] (M_\varepsilon + r^{-(d+2)} \lambda^{-2} D) \lambda^{-2} D. \end{aligned}$$

Thus

$$I_4 \leq C\varepsilon^{-6} M_\varepsilon \overline{\exp}_2 \left[ C_d \varepsilon^{-2} (1 + r^{-(d+2)} \lambda^{-2} D) \right] (M_\varepsilon + r^{-(d+2)} \lambda^{-2} D) \lambda^{-2} D.$$

We summarise this as

$$\begin{aligned} \dot{D} &\leq (\lambda + \alpha)D + \frac{1}{\alpha}C\varepsilon^{-4}(1 + |\log r|)^2 \left( M_\varepsilon + r^{-(d+2)}\lambda^{-2}D \right)^2 \lambda^{-2}D \\ &\quad + C\overline{\exp}_2 \left[ C_d\varepsilon^{-2}(1 + r^{-(d+2)}\lambda^{-2}D) \right] \left( 1 + M_\varepsilon(M_\varepsilon + r^{-(d+2)}\lambda^{-2}D) \right) \lambda^{-2}D. \end{aligned} \quad (9.7)$$

Again we introduce the truncated functional

$$\hat{D} = 1 \wedge \left( r^{-(d+2)}\lambda^{-2}D \right),$$

which is Lipschitz and so differentiable almost everywhere. Where the derivative exists, (9.7) implies that

$$\frac{d}{dt}\hat{D} \leq (\lambda + \alpha)\hat{D} + \frac{1}{4\alpha}C_d \left[ \varepsilon^{-4}(1 + |\log r|)^2 + \overline{\exp}_2(C\varepsilon^{-2}) \right] M_\varepsilon^2 \lambda^{-2} \hat{D}.$$

We optimise the exponent by choosing

$$\begin{aligned} \alpha_* &= C_d \left[ \varepsilon^{-4}(1 + |\log r|)^2 + \overline{\exp}_2(C\varepsilon^{-2}) \right]^{1/2} M_\varepsilon \lambda_*^{-1} \\ \lambda_* &= C_d \left[ \varepsilon^{-4}(1 + |\log r|)^2 + \overline{\exp}_2(C\varepsilon^{-2}) \right]^{1/4} \sqrt{M_\varepsilon}. \end{aligned} \quad (9.8)$$

Then

$$\frac{d}{dt}\hat{D} \leq C_d \lambda_* \hat{D}.$$

By absolute continuity of  $\hat{D}$ , it follows that

$$\hat{D}(t) \leq \hat{D}(0) + C_d \lambda_* \int_0^t \hat{D}(s) ds.$$

From this we deduce that

$$\begin{aligned} \sup_{[0,T]} \hat{D}(t) &\leq \exp(C_d \lambda_* T) \hat{D}(0) \\ &\leq r^{-(d+2)} \lambda_*^{-2} \exp(C_d \lambda_* T) D(0) \\ &\leq r^{-(d+2)} \exp(C_d \lambda_* T) D(0) \\ &\leq C_d \exp \left[ |\log r| \left( (d+2) + C_d T \sqrt{M_\varepsilon} \varepsilon^{-1} |\log r|^{-1/2} \right) + \sqrt{M_\varepsilon} \overline{\exp}_2(C\varepsilon^{-2}) \right] D(0). \end{aligned}$$

By the estimate (9.6) on  $M_\varepsilon$ , we obtain

$$\sup_{[0,T]} \hat{D}(t) \leq C_d \exp \left[ |\log r| \left( (d+2) + C_d T \exp(C\varepsilon^{-2}) |\log r|^{-1/2} \right) \right] \overline{\exp}_3(C\varepsilon^{-2}) D(0).$$

Since by assumption (9.3)

$$|\log r|^{-1/2} \leq \exp \left[ -\frac{1}{2} \exp(K\varepsilon^{-2}) \right],$$

we have

$$\sup_{[0,T]} \hat{D}(t) \leq C_d \exp \left[ |\log r| \left( (d+2) + \exp(C\varepsilon^{-2}) - \frac{1}{2} \exp(K\varepsilon^{-2}) \right) \right] \overline{\exp}_3(C\varepsilon^{-2}) D(0).$$

By assumption on  $W_2(\mu_{\varepsilon,r}^N(0), f_\varepsilon^{(r)}(0))$ , for all sufficiently small  $r$  there exists a choice of initial coupling  $\pi_0^{(r)}$  such that

$$D(0) < r^{d+2+\eta}.$$

Then

$$\sup_{[0,T]} \hat{D}(t) \leq C_d \exp \left[ |\log r| \left( \exp(C\varepsilon^{-2}) - \frac{1}{2} \exp(K\varepsilon^{-2}) \right) - \eta \right] + \overline{\exp}_2(C\varepsilon^{-2}).$$

For sufficiently small  $\varepsilon$ ,

$$\begin{aligned} \sup_{[0,T]} \hat{D}(t) &\leq C_d \exp \left[ -\frac{1}{2} \eta |\log r| + \overline{\exp}_2(C\varepsilon^{-2}) \right] \\ &\leq C_d \exp \left[ -\frac{1}{2} \eta \overline{\exp}_2(K\varepsilon^{-2}) + \overline{\exp}_2(C\varepsilon^{-2}) \right]. \end{aligned}$$

Thus if  $K > C$ , then

$$\sup_{[0,T]} \hat{D}(t) \rightarrow 0$$

as  $\varepsilon$  tends to zero. In particular, for  $\varepsilon$  sufficiently small,

$$\inf_{\pi_0} \sup_{t \in [0,T]} \hat{D}(t) < 1.$$

Hence, for  $\varepsilon$  sufficiently small,

$$\inf_{\pi_0} \sup_{t \in [0,T]} \hat{D}(t) = \inf_{\pi_0} \sup_{t \in [0,T]} r^{-(d+2)} \lambda_*^{-2} D(t).$$

Thus

$$\sup_{t \in [0,T]} W_2^2(\mu_{\varepsilon,r}^N(t), f_\varepsilon^{(r)}(t)) \leq \inf_{\pi_0} \sup_{t \in [0,T]} D(t) = r^{d+2} \lambda_*^2 \inf_{\pi_0} \sup_{t \in [0,T]} \hat{D}(t) \leq r^{d+2} \lambda_*^2.$$

By (9.8),

$$\lambda_*^2 = C_d \left[ \varepsilon^{-4} (1 + |\log r|)^2 + \overline{\exp}_2(C\varepsilon^{-2}) \right]^{1/2} \exp(C\varepsilon^{-2}).$$

Hence, for any  $\alpha > 0$ ,

$$r^{d+2} \lambda_*^2 \leq C r^{d+2-\alpha} \overline{\exp}_2(C\varepsilon^{-2}) \leq C \exp \{ \exp(C\varepsilon^{-2}) - (d+2-\alpha) \overline{\exp}_2(K\varepsilon^{-2}) \}.$$

The right hand side converges to zero as  $\varepsilon$  tends to zero. Therefore

$$\lim_{\varepsilon \rightarrow 0} \sup_{t \in [0,T]} W_2^2(f_t, \mu_t) = 0.$$

□

The next lemma is a quantified version of Lemma 8.6.

**Lemma 9.2** (Approximation of (VPME) in quasineutral scaling). *Let  $f_\varepsilon(0)$  satisfy the assumptions of Theorem 2.2. Let  $f_\varepsilon^{(r)}$  be the solution of the scaled and regularised Vlasov equation (9.1) with initial datum  $f_\varepsilon(0)$ . Let  $f_\varepsilon$  be the unique bounded density solution of (1.3).*

*Fix  $T > 0$ . Then there exists a constant  $C$  depending on  $T$  and on  $\{f_\varepsilon(0)\}_\varepsilon$  such that the following holds. If  $r$  and  $\varepsilon$  satisfy*

$$r \leq [\overline{\exp}_3(C\varepsilon^{-2})]^{-1},$$

*then*

$$\lim_{r \rightarrow 0} \sup_{[0,T]} W_2 \left( f_\varepsilon^{(r)}(t), f_\varepsilon(t) \right) = 0.$$

*Proof.* By (2.3), Lemma 3.3 and (9.2) there exists a constant  $C$  depending on  $C_0$  and  $d$  only such that

$$\|\rho_{f_\varepsilon^{(r)}}\|_{L^\infty([0,T];L^{\frac{d+2}{d}}(\mathbb{T}^d))}, \|\rho_{f_\varepsilon}\|_{L^\infty([0,T];L^{\frac{d+2}{d}}(\mathbb{T}^d))} \leq C. \quad (9.9)$$

By Proposition 6.1, there exists a constant  $C$  depending on the initial data and  $T$  such that

$$\|\rho_{f_\varepsilon^{(r)}}\|_{L^\infty([0,T];L^\infty(\mathbb{T}^d))}, \|\rho_{f_\varepsilon}\|_{L^\infty([0,T];L^\infty(\mathbb{T}^d))} \leq \exp(C\varepsilon^{-2}). \quad (9.10)$$

We will control the Wasserstein distance between  $f_\varepsilon^{(r)}$  and  $f_\varepsilon$  using a particular coupling  $\pi_t$ . Since both solutions share the same initial datum, we take  $\pi_0$  to be the trivial coupling  $f_\varepsilon(0)(x, v)\delta((x, v) - (y, w)) dx dv dy dw$ . We construct  $\pi_t \in \Pi(f_\varepsilon(t), f_\varepsilon^{(r)}(t))$  as in (8.8), using the scaled version of the characteristic systems in (8.26):

$$\begin{cases} \dot{X}_{x,v}^{(1)} = V_{x,v}^{(1)} \\ \dot{V}_{x,v}^{(1)} = E(X_{x,v}^{(1)}) \\ (X_{x,v}^{(1)}(0), V_{x,v}^{(1)}(0)) = (x, v) \\ E = -\nabla U \\ \varepsilon^2 \Delta U = e^U - \rho_{f_\varepsilon} \end{cases} \quad \begin{cases} \dot{X}_{x,v}^{(2)} = V_{x,v}^{(2)} \\ \dot{V}_{x,v}^{(2)} = E_r(X_{x,v}^{(2)}) \\ (X_{x,v}^{(2)}(0), V_{x,v}^{(2)}(0)) = (x, v) \\ E_r = -\chi_r * \nabla U_r \\ \varepsilon^2 \Delta U_r = e^{U_r} - \chi_r * \rho_{f_\varepsilon^{(r)}}. \end{cases}$$

We define  $D$  as in (8.9). As in Lemma 8.5, we obtain for any  $\alpha > 0$ ,

$$\dot{D} \leq (\alpha + \lambda)D + \frac{C}{\alpha} \sum_{i=1}^5 I_i,$$

where

$$\begin{aligned} I_1 &:= \int |\nabla \bar{U}_r(X_t^{(1)}) - \nabla \bar{U}_r(X_t^{(2)})|^2 d\pi_0, & I_2 &:= \int |\nabla \bar{U}_r(X_t^{(1)}) - \nabla \bar{U}(X_t^{(1)})|^2 d\pi_0, \\ I_3 &:= \int |\nabla \hat{U}_r(X_t^{(1)}) - \nabla \hat{U}_r(X_t^{(2)})|^2 d\pi_0, & I_4 &:= \int |\nabla \hat{U}(X_t^{(1)}) - \nabla \hat{U}_r(X_t^{(1)})|^2 d\pi_0 \\ I_5 &:= \int |\chi_r * \nabla U_r(X_t^{(2)}) - \nabla U_r(X_t^{(2)})|^2 d\pi_0. \end{aligned}$$

$I_1$  is estimated as in (8.27). There is an extra factor of  $\varepsilon^{-4}$  due to the quasineutral scaling and the mass bound  $M = M_\varepsilon$  depends on  $\varepsilon$ :

$$I_1 \leq C\varepsilon^{-4} |\log r|^2 M_\varepsilon^2 \lambda^{-2} D.$$

We estimate  $I_2$  as in (8.28), keeping track of the dependence on  $\varepsilon$  in Lemma 4.5:

$$\|\nabla \bar{U} - \nabla \bar{U}_r\|_{L^2(\mathbb{T}^d)} \leq C\varepsilon^{-2} \sqrt{M_\varepsilon} (r + \lambda^{-1} \sqrt{D}).$$

We conclude that

$$I_2 \leq C\varepsilon^{-4} M_\varepsilon^2 (r + \lambda^{-1} \sqrt{D})^2.$$

For  $I_3$ , by the regularity estimate from Proposition 4.1 and the uniform  $L^{\frac{d+2}{d}}(\mathbb{T}^d)$  estimate on the density (9.9) we have

$$\|\hat{U}_r\|_{C^{2,\alpha}(\mathbb{T}^d)} \leq \overline{\exp}_2(C\varepsilon^{-2}),$$

where  $C$  depends only on  $C_0$ . We obtain

$$\begin{aligned} I_3 &\leq \overline{\exp}_2(C\varepsilon^{-2}) \int_{\mathbb{T}^d \times \mathbb{R}^d} |X_t^{(1)} - X_t^{(2)}|^2 d\pi_0 \\ &\leq \overline{\exp}_2(C\varepsilon^{-2}) \lambda^{-2} D. \end{aligned}$$

For  $I_4$  we use the stability estimate from Lemma 4.6 and the  $L^{\frac{d+2}{d}}(\mathbb{T}^d)$  estimate (9.9):

$$\|\nabla\widehat{U} - \nabla\widehat{U}_r\|_{L^2(\mathbb{T}^d)} \leq \overline{\text{exp}}_2(C\varepsilon^{-2})\|\nabla\bar{U} - \nabla\bar{U}_r\|_{L^2(\mathbb{T}^d)} \leq \overline{\text{exp}}_2(C\varepsilon^{-2})\sqrt{M_\varepsilon}(r + \lambda^{-1}\sqrt{D}).$$

Hence

$$I_4 \leq \overline{\text{exp}}_2(C\varepsilon^{-2})M_\varepsilon^2(r + \lambda^{-1}\sqrt{D})^2.$$

For  $I_5$ , by (8.29) we have

$$\|\chi_r * \nabla U_r - \nabla U_r\|_{L^2(\mathbb{T}^d)}^2 \leq Cr^2\|U_r\|_{W^{2,2}(\mathbb{T}^d)}^2.$$

Thus

$$\begin{aligned} I_5 &:= \int_{\mathbb{T}^d} |\chi_r * \nabla U_r(x) - \nabla U_r(x)|^2 \rho_{f_\varepsilon^{(r)}}(dx) \\ &\leq \|\rho_{f_\varepsilon^{(r)}}\|_{L^\infty(\mathbb{T}^d)} \|\chi_r * \nabla U_r - \nabla U_r\|_{L^2(\mathbb{T}^d)}^2 \\ &\leq Cr^2 M_\varepsilon \|U_r\|_{W^{2,2}(\mathbb{T}^d)}^2 \\ &\leq Cr^2 \varepsilon^{-4} M_\varepsilon \|e^{U_r} - \rho_{f_\varepsilon^{(r)}}\|_{L^2(\mathbb{T}^d)}^2. \end{aligned}$$

As in (8.30), we have

$$\|\rho_{f_\varepsilon^{(r)}}\|_{L^2(\mathbb{T}^d)} \leq \|\rho_{f_\varepsilon^{(r)}}\|_{L^\infty(\mathbb{T}^d)}^{\frac{d-2}{2d}} \|\rho_{f_\varepsilon^{(r)}}\|_{L^{\frac{d+2}{d}}(\mathbb{T}^d)}^{\frac{d+2}{2d}} \leq CM_\varepsilon^{\frac{d-2}{2d}}.$$

To estimate  $e^{U_r}$ , first note that by Proposition 4.1 and (9.9),

$$\|\bar{U}_r\|_{L^\infty(\mathbb{T}^d)} \leq C\varepsilon^{-2}.$$

By (4.5),

$$\|e^{\widehat{U}_r}\|_{L^2(\mathbb{T}^d)} \leq \exp(C\varepsilon^{-2}).$$

Thus

$$\|e^{U_r}\|_{L^2(\mathbb{T}^d)} \leq C \exp\left(\|\bar{U}_r\|_{L^\infty(\mathbb{T}^d)}\right) \|e^{\widehat{U}_r}\|_{L^2(\mathbb{T}^d)} \leq \exp(C\varepsilon^{-2}).$$

Therefore

$$I_5 \leq Cr^2 \varepsilon^{-4} M_\varepsilon \left( M_\varepsilon^{\frac{d-2}{d}} + e^{C\varepsilon^{-2}} \right).$$

Putting these five estimates together, we obtain

$$\begin{aligned} \dot{D} &\leq (\alpha + \lambda)D + \frac{C}{\alpha\lambda^2} r^2 \overline{\text{exp}}_2(C\varepsilon^{-2}) M_\varepsilon^2 \\ &\quad + \frac{C}{\alpha\lambda^2} \left\{ \overline{\text{exp}}_2(C\varepsilon^{-2}) M_\varepsilon^2 + \varepsilon^{-4} |\log r|^2 M_\varepsilon^2 + \overline{\text{exp}}_2(C\varepsilon^{-2}) \right\} D. \end{aligned}$$

By (9.10), we may estimate that

$$M_\varepsilon \leq \exp(C\varepsilon^{-2}).$$

From this we deduce

$$\dot{D} \leq (\alpha + \lambda)D + \frac{C}{\alpha\lambda^2} \left\{ \exp(C\varepsilon^{-2}) |\log r|^2 + \overline{\text{exp}}_2(C\varepsilon^{-2}) \right\} D + \frac{C}{\alpha\lambda^2} \overline{\text{exp}}_2(C\varepsilon^{-2}) r^2.$$

After choosing  $\alpha$  and  $\lambda$  so as to minimise the constant in front of  $D$  we obtain

$$\dot{D} \leq C\lambda_* D + \overline{\text{exp}}_2(C\varepsilon^{-2})r^2,$$

where

$$\lambda_* = [\exp(C\varepsilon^{-2})|\log r|^2 + \overline{\text{exp}}_2(C\varepsilon^{-2})]^{1/4} \geq 1,$$

for  $r, \varepsilon$  sufficiently small.

Therefore, by a Grönwall estimate

$$D(t) \leq (D(0) + \overline{\text{exp}}_2(C\varepsilon^{-2})r^2) \exp[Ct\lambda_*].$$

Since  $\pi_0$  was trivial,  $D(0) = 0$ . Since  $D$  controls the squared Wasserstein distance,

$$W_2^2(f_\varepsilon^{(r)}(t), f_\varepsilon(t)) \leq \exp(C\varepsilon^{-2})r^2 \exp[Ct\lambda_*].$$

Then, by definition of  $\lambda_*$ ,

$$\sup_{t \in [0, T]} W_2^2(f_\varepsilon^{(r)}(t), f_\varepsilon(t)) \leq r^2 \exp\left[\exp(C_T\varepsilon^{-2})|\log r|^{1/2}\right] \cdot \overline{\text{exp}}_3(C_T\varepsilon^{-2}).$$

If  $r \leq [\overline{\text{exp}}_3(K\varepsilon^{-2})]^{-1}$ , then

$$e^{C\varepsilon^{-2}}|\log r|^{-1/2} \leq \exp\left[C_T\varepsilon^{-2} - \frac{1}{2}\exp(K\varepsilon^{-2})\right] \rightarrow 0$$

as  $\varepsilon$  tends to zero, for any  $K \geq 0$ . Hence, for any  $\eta > 0$ , for  $\varepsilon$  sufficiently small,

$$\begin{aligned} r^2 \exp\left[\exp(C_T\varepsilon^{-2})|\log r|^{1/2}\right] &\leq r^{2-\exp(C_T\varepsilon^{-2})|\log r|^{-1/2}} \\ &\leq r^{2-\eta}. \end{aligned}$$

Moreover,

$$r^{2-\eta} \overline{\text{exp}}_3(C_T\varepsilon^{-2}) \leq \exp\left[\overline{\text{exp}}_2(C_T\varepsilon^{-2}) - (2-\eta)\overline{\text{exp}}_2(K\varepsilon^{-2})\right],$$

which converges to zero for any  $\eta < 2$  as  $\varepsilon$  tends to zero, as long as  $K \geq C_T$ .

Therefore, if  $r \leq [\overline{\text{exp}}_3(K\varepsilon^{-2})]^{-1}$  for  $K \geq C_T$ , then as  $\varepsilon$  tends to zero (and so  $r$  also tends to zero),

$$\sup_{t \in [0, T]} W_2(f_\varepsilon^{(r)}(t), f_\varepsilon(t)) \rightarrow 0.$$

□

*Proof of Theorem 2.5.* In the following,  $\mu_{\varepsilon, r}^N$  denotes the empirical measure associated to the solution of the particle system (2.9), while  $f_\varepsilon^{(r)}$  denotes the solution of the scaled version of (9.1) with initial datum  $f_\varepsilon(0)$ . Let  $g$  be the solution of the KIE system (1.4), on some time interval  $[0, T_*]$ , obtained in the quasineutral limit from  $f_\varepsilon$  using Theorem 2.2. By the triangle inequality for  $W_1$ ,

$$\begin{aligned} \sup_{t \in [0, T_*]} W_1(\mu_{\varepsilon, r}^N(t), g(t)) &\leq \sup_{t \in [0, T_*]} W_1(\mu_{\varepsilon, r}^N(t), f_\varepsilon^{(r)}(t)) \\ &\quad + \sup_{t \in [0, T_*]} W_1(f_\varepsilon^{(r)}(t), f_\varepsilon(t)) + \sup_{t \in [0, T_*]} W_1(f_\varepsilon(t), g(t)). \end{aligned} \quad (9.11)$$

The last term converges to zero as  $\varepsilon$  tends to zero, by Theorem 2.2.

For the other two terms, we first observe that

$$W_1(\mu_{\varepsilon, r}^N(t), f_\varepsilon^{(r)}(t)) \leq W_2(\mu_{\varepsilon, r}^N(t), f_\varepsilon^{(r)}(t)), \quad W_1(f_\varepsilon^{(r)}(t), f_\varepsilon(t)) \leq W_2(f_\varepsilon^{(r)}(t), f_\varepsilon(t)).$$

Then the second term of (9.11) converges to zero by Lemma 9.1 and the third term of (9.11) converges to zero by Lemma 9.2, provided that (9.3) is satisfied for  $C$  depending on  $T_*$ .

□



## 10 Typicality: proof of Theorems 2.6 and 2.8

In this last section, we prove Theorems 2.6 and 2.8 concerning the relation between the choice of parameters in the mean field (or combined mean field-quasineutral) limit and the initial configurations. Our method follows the approach of Lazarovici [31] for the mean field case, with modifications for the quasineutral case similar to those used in [21].

The underlying observation is that if one constructs a collection of empirical measures  $(\nu^N)_N$  by drawing  $N$  independent samples from a reference measure  $\nu$ , then, by the Glivenko-Cantelli theorem, almost surely  $\nu^N$  will converge to  $\nu$  as  $N$  tends to infinity in the sense of weak convergence of measures. The idea is to use a quantitative version of this result to find configurations for which the associated empirical measures converge to our reference data  $f_0$  sufficiently quickly. A result of this type was proved by Fournier and Guillin in [15], with the distance between  $\nu^N$  and  $\nu$  measured in Wasserstein sense. We use two slightly different versions of their concentration estimates.

In the unscaled case  $\varepsilon = 1$  (Theorem 2.6), we give a result for more general data  $f_0$  satisfying a moment condition. For this we need the following result from [15, Theorem 2].

**Theorem 10.1.** *Let  $\nu$  be a probability measure on  $\mathbb{R}^m$  and let  $\nu^N$  denote the empirical measure of  $N$  independent samples from  $\nu$ . Assume that  $\nu$  has a finite  $k$ th moment for some  $k > 2p$ :*

$$M_k(\nu) := \int_{\mathbb{R}^m} |x|^k d\nu(x) < +\infty.$$

Then there exist constants  $c, C$  depending on  $p, m$  and  $M_k(\nu)$  such that for any  $x > 0$ ,

$$\mathbb{P}(W_p^p(\nu^N, \nu) \geq x) \leq a(N, x)\mathbb{1}_{\{x \leq 1\}} + b(N, x),$$

where

$$a(N, x) = C \begin{cases} \exp(-cNx^2) & p > \frac{m}{2} \\ \exp\left(-cN \left[\frac{x}{\log(2+\frac{1}{x})}\right]^2\right) & p = \frac{m}{2} \\ \exp(-cNx^{m/p}) & p < \frac{m}{2} \end{cases} \quad (10.1)$$

and

$$b(N, x) = CN(Nx)^{-(k-\alpha)/p}$$

for any  $\alpha \in (0, k)$ .

For the combined mean field and quasineutral limit, we work with a different initial datum  $f_\varepsilon(0)$  for each  $\varepsilon$ . To use the Wasserstein concentration estimates, we need to take care of the dependence of the constants  $c, C$  on (the moments of)  $f_\varepsilon(0)$ . In fact, since in Theorem 2.2 we work with compactly supported data, we will find it more convenient to use a slightly different version of the estimates, designed for compactly supported measures. The following result is from [15, Proposition 10].

**Theorem 10.2.** *Let  $\nu$  be a probability measure supported on  $(-1, 1]^m$ . Let  $\nu^N$  denote the empirical measure of  $N$  independent samples from  $\nu$ . Then there exist constants  $c, C$  depending on  $p$  and  $m$  only such that for any  $x > 0$ ,*

$$\mathbb{P}(W_p^p(\nu^N, \nu) \geq x) \leq a(N, x)\mathbb{1}_{\{x \leq 1\}},$$

where  $a(N, x)$  is defined by (10.1).

We can now prove Theorems 2.6 and 2.8. Since the proofs are rather similar to the ones used in [31, 21] for the classical VP system, we invite the reader to consult those papers for more details.

*Proof of Theorem 2.6.* We follow the approach of [31, Theorem 3.3]. The idea of the proof is to show that, for the choice  $r = cN^{-\gamma}$ ,

$$\mathbb{P} \left( \limsup_{N \rightarrow \infty} \frac{W_2^2(\mu_0^N, f_0)}{r^{d+2+C_{T^*,M}|\log r|^{-1/2}}} < 1 \right) = 1. \quad (10.2)$$

Then we may apply Theorem 2.4 to conclude that the mean field limit holds on this full probability event. To prove (10.2), observe that

$$\bigcup_n \bigcap_{N \geq n} A_N^c \subset \left\{ \limsup_{N \rightarrow \infty} \frac{W_2^2(\mu_0^N, f_0)}{r^{d+2+C_{T^*,M}|\log r|^{-1/2}}} < 1 \right\},$$

where  $A_N$  is the event

$$A_N := \left\{ W_2^2(\mu_0^N, f_0) > \frac{1}{2} r^{d+2+C_{T^*,M}|\log r|^{-1/2}} \right\}.$$

Since  $\left( \bigcup_n \bigcap_{N \geq n} A_N^c \right)^c = \bigcap_n \bigcup_{N \geq n} A_N$ , by the Borel-Cantelli lemma it suffices to show that

$$\sum_N \mathbb{P}(A_N) < \infty. \quad (10.3)$$

We estimate  $\mathbb{P}(A_N)$  using Theorem 10.1, with

$$x_N = \frac{1}{2} r^{d+2+C_{T^*,M}|\log r|^{-1/2}} = cN^{-\gamma(d+2)-C_{T^*,M,\gamma}|\log N|^{-1/2}}.$$

Note that  $p = 2$  and  $m = 2d$ . The assumptions on  $\gamma$  in (2.10) are chosen such that

$$\sum_N a(N, x_N) + b(N, x_N) < \infty.$$

In this way (10.3) holds and the result follows.  $\square$

*Proof of Theorem 2.8.* We use the same strategy as for Theorem 2.6. However, since the constants in Theorem 10.1 depend on the  $k$ th moment of  $f_\varepsilon(0)$  which may change with  $\varepsilon$ , we will instead use the compact version of the estimate from Theorem 10.2. Recall that we already assumed that  $f_\varepsilon(0)$  were compactly supported with the support in velocity growing no faster than  $e^{C\varepsilon^{-2}}$  for some  $C$ . We perform a scaling argument in order to work with measures that are supported in  $(-1, 1]^{2d}$ .

Given a probability measure  $\nu$ , let  $\mathcal{S}_R[\nu]$  be the measure such that for any  $A_X \in \mathcal{B}(\mathbb{T}^d)$  and  $B_V \in \mathcal{B}(\mathbb{R}^d)$ ,

$$\mathcal{S}_R[\nu](A_X \times B_V) = \nu(A_X \times RB_V).$$

This maps measures supported in  $[-1, 1]^d \times [-R, R]^d$  to measures supported in  $[-1, 1]^{2d}$ . By [21, Lemma 7.6], if  $\nu_1, \nu_2$  are measures on  $\mathbb{T}^d \times \mathbb{R}^d$ , then for any  $p \in [1, \infty)$ ,

$$W_p(\nu_1, \nu_2) \leq R W_p(\mathcal{S}_R[\nu_1], \mathcal{S}_R[\nu_2]).$$

Note also that if  $(Z_i^{(R)})_{i=1}^N$  are  $N$  independent samples from  $\mathcal{S}_R[\nu]$ , then  $(Z_i^{(R)})_{i=1}^N$  has the same law as  $(X_i, \frac{1}{R}V_i)_{i=1}^N$ , where  $(Z_i)_{i=1}^N = (X_i, V_i)_{i=1}^N$  are  $N$  independent samples from  $\nu$ . It is therefore enough to show that

$$\sum_N \mathbb{P}(A_N) < \infty, \quad (10.4)$$

where  $A_N$  denotes the event

$$A_N := \left\{ W_2^2(\mathcal{S}_{e^{-C\varepsilon^{-2}}}[\mu_\varepsilon^N(0)], \mathcal{S}_{e^{-C\varepsilon^{-2}}}[f_{0,\varepsilon}]) > \frac{1}{2} r^{d+2+\eta} \exp(-2C\varepsilon^{-2}) \right\}.$$

We observe that the assumption

$$r < [\overline{\exp}_3(K\varepsilon^{-2})]^{-1},$$

implies that

$$\exp(-C\varepsilon^{-2}) > c(\log \log N)^{-\zeta} > cN^{-\alpha}$$

for  $\zeta$  depending on  $C$  and  $K$ , any  $\alpha > 0$  and  $c$  depending on  $\alpha$ ,  $C$  and  $K$ . We then apply Theorem 10.2 with the choice

$$x_N = cN^{-\gamma(d+2+\eta)-\alpha}.$$

The assumption (2.11) on  $\gamma$  implies that it is possible to find  $\eta > 0$  such that

$$\sum_N a(N, x_N) < \infty.$$

This yields (10.4), which completes the proof. □

*Acknowledgements:* This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/L016516/1 for the University of Cambridge Centre for Doctoral Training, the Cambridge Centre for Analysis; and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 726386). The authors would like to thank the anonymous referees for their very useful and thoughtful comments.

## References

- [1] L. Ambrosio. Transport equation and Cauchy problem for Non-smooth Vector Fields. In B. Dacorogna and P. Marcellini, editors, *Calculus of Variations and Nonlinear Partial Differential Equations: Lectures given at the C.I.M.E. Summer School held in Cetraro, Italy, June 27 - July 2, 2005*. Springer-Verlag Berlin Heidelberg, 2008.
- [2] C. Bardos and N. Besse. The Cauchy problem for the Vlasov-Dirac-Benney equation and related issues in fluid mechanics and semi-classical limits. *Kinet. Relat. Models*, 6(4):893–917, 2013.
- [3] C. Bardos, F. Golse, T. T. Nguyen, and R. Sentis. The Maxwell-Boltzmann approximation for ion kinetic modeling. *Phys. D*, 376/377:94–107, 2018.
- [4] C. Bardos and A. Nouri. A Vlasov equation with Dirac potential used in fusion plasmas. *J. Math. Phys.*, 53(11):115621, 2012.

- [5] J. Batt and G. Rein. Global classical solutions of the periodic Vlasov-Poisson system in three dimensions. *C. R. Acad. Sci. Paris Sér. I Math.*, 313(6):411–416, 1991.
- [6] T. Bodineau, I. Gallagher, and L. Saint-Raymond. From hard sphere dynamics to the stokes–fourier equations: An analysis of the boltzmann–grad limit. *Annals of PDE*, 3(1):2, Feb 2017.
- [7] G. Bonhomme, T. Pierre, G. Leclert, and J. Trulsen. Ion phase space vortices in ion beam-plasma systems and their relation with the ion acoustic instability: numerical and experimental results. *Plasma Physics and Controlled Fusion*, 33(5):507–520, may 1991.
- [8] M. Bossy, J. Fontbona, P.-E. Jabin, and J.-F. Jabir. Local existence of analytical solutions to an incompressible Lagrangian stochastic model in a periodic domain. *Comm. Partial Differential Equations*, 38(7):1141–1182, 2013.
- [9] F. Bouchut. Global weak solution of the Vlasov-Poisson system for small electrons mass. *Comm. Partial Differential Equations*, 16(8-9):1337–1365, 1991.
- [10] W. Braun and K. Hepp. The Vlasov dynamics and its fluctuations in the  $1/N$  limit of interacting classical particles. *Comm. Math. Phys.*, 56(2):101–113, 1977.
- [11] Y. Brenier and E. Grenier. Limite singulière du système de Vlasov-Poisson dans le régime de quasi neutralité: le cas indépendant du temps. *C. R. Acad. Sci. Paris Sér. I Math.*, 318(2):121–124, 1994.
- [12] F. F. Chen. *Introduction to Plasma Physics and Controlled Fusion*. Springer International Publishing, 3 edition, 2016.
- [13] R. L. Dobrushin. Vlasov Equations. *Funktsional. Anal. i Prilozhen.*, 13(2):48–58, 1979.
- [14] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 2010.
- [15] N. Fournier and A. Guillin. On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields*, 162(3-4):707–738, 2015.
- [16] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag Berlin Heidelberg, 1977.
- [17] F. Golse. On the dynamics of large particle systems in the mean field limit. In A. Muntean, J. Rademacher, and A. Zagaris, editors, *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 1–144. Springer, 2016.
- [18] E. Grenier. Defect measures of the Vlasov-Poisson system in the quasineutral regime. *Comm. Partial Differential Equations*, 20(7-8):1189–1215, 1995.
- [19] E. Grenier. Oscillations in quasineutral plasmas. *Comm. Partial Differential Equations*, 21(3-4):363–394, 1996.
- [20] M. Griffin-Pickering and M. Iacobelli. Global well-posedness in 3-dimensions for the Vlasov–Poisson system with massless electrons. Preprint, 2018, 1810.06928.

- [21] M. Griffin-Pickering and M. Iacobelli. A mean field approach to the quasi-neutral limit for the Vlasov–Poisson equation. *SIAM J. Math. Anal.*, 50(5):5502–5536, 2018.
- [22] A. Gurevich, L. Pariiskaya, and L. Pitaevskii. Self-similar motion of rarefied plasma. *Soviet Phys. JETP*, 22(2):449–454, 1966.
- [23] A. Gurevich, L. Pariiskaya, and L. Pitaevskii. Self-similar motion of a low-density plasma. II. *Soviet Phys. JETP*, 27(3):476–482, 1968.
- [24] A. V. Gurevich and L. P. Pitaevsky. Non-linear dynamics of a rarefied ionized gas. *Progress in Aerospace Sciences*, 16(3):227 – 272, 1975.
- [25] D. Han-Kwan and M. Hauray. Stability issues in the quasineutral limit of the one-dimensional Vlasov-Poisson equation. *Comm. Math. Phys.*, 334(2):1101–1152, 2015.
- [26] D. Han-Kwan and M. Iacobelli. Quasineutral limit for Vlasov-Poisson via Wasserstein stability estimates in higher dimension. *J. Differential Equations*, 263(1):1–25, 2017.
- [27] D. Han-Kwan and M. Iacobelli. The quasineutral limit of the Vlasov-Poisson equation in Wasserstein metric. *Commun. Math. Sci.*, 15(2):481–509, 2017.
- [28] D. Han-Kwan and F. Rousset. Quasineutral limit for Vlasov-Poisson with Penrose stable data. *Ann. Sci. Éc. Norm. Supér. (4)*, 49(6):1445–1495, 2016.
- [29] M. Hauray and P.-E. Jabin. Particle approximation of Vlasov equations with singular forces: propagation of chaos. *Ann. Sci. Éc. Norm. Supér. (4)*, 48(4):891–940, 2015.
- [30] E. Horst. Global solutions of the relativistic Vlasov-Maxwell system of plasma physics. *Dissertationes Math. (Rozprawy Mat.)*, 292, 1990.
- [31] D. Lazarovici. The Vlasov-Poisson dynamics as the mean field limit of extended charges. *Comm. Math. Phys.*, 347(1):271–289, 2016.
- [32] D. Lazarovici and P. Pickl. A mean field limit for the Vlasov-Poisson system. *Arch. Ration. Mech. Anal.*, 225(3):1201–1231, 2017.
- [33] M. A. Lieberman and A. J. Lichtenberg. *Principles of Plasma Discharges and Materials Processing*. John Wiley & Sons, 2005.
- [34] P. L. Lions and B. Perthame. Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. *Invent. Math.*, 105(2):415–430, 1991.
- [35] G. Loeper. Uniqueness of the solution to the Vlasov-Poisson system with bounded density. *J. Math. Pures Appl. (9)*, 86(1):68–79, 2006.
- [36] A. Majda and A. Bertozzi. *Vorticity and Incompressible Flow*, volume 27 of *Cambridge Texts in Applied Mathematics*. Cambridge University Press, 2002.
- [37] R. J. Mason. Computer simulation of ion-acoustic shocks. the diaphragm problem. *The Physics of Fluids*, 14(9):1943–1958, 1971.
- [38] Y. V. Medvedev. Ion front in an expanding collisionless plasma. *Plasma Physics and Controlled Fusion*, 53(12):125007, nov 2011.

- [39] H. Neunzert. An introduction to the nonlinear Boltzmann-Vlasov equation. In *Kinetic Theories and the Boltzmann Equation*, volume 1048 of *Lecture Notes in Math.*, pages 60–110. Springer, Berlin, Heidelberg, 1984.
- [40] H. Neunzert and J. Wick. Die Approximation der Lösung von Integro-Differentialgleichungen durch endliche Punktmengen. In *Numerische Behandlung nichtlinearer Integrodifferential-und Differentialgleichungen.*, volume 395 of *Lecture Notes in Math.*, pages 275–290, Berlin, Heidelberg, 1974. Springer.
- [41] K. Pfaffelmoser. Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. *J. Differential Equations*, 95(2):281–303, 1992.
- [42] G. Rein. Collisionless kinetic equations from astrophysics - the Vlasov-Poisson system. In *Handbook of Differential Equations: Evolutionary Equations*, volume III, chapter 5, pages 383–476. Elsevier/North Holland, Amsterdam, 2007.
- [43] P. Sakanaka, C. Chu, and T. Marshall. Formation of ion-acoustic collisionless shocks. *The Physics of Fluids*, 14(611), 1971.
- [44] J. Schaeffer. Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions. *Comm. Partial Differential Equations*, 16(8-9):1313–1335, 1991.
- [45] E. Titchmarsh. *Eigenfunction Expansions Associated with Second-Order Differential Equations, Part II*. Oxford University Press, 1958.
- [46] S. Ukai and T. Okabe. On classical solutions in the large in time of two-dimensional Vlasov's equation. *Osaka J. Math.*, 15(2):245–261, 1978.
- [47] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.