C. R. Acad. Sci. Paris, Ser. I 357 (2019) 708–720

Contents lists available at ScienceDirect

C. R. Acad. Sci. Paris, Ser. I

www.sciencedirect.com

Partial differential equations/Mathematical physics

On mean-field limits and quantitative estimates with a large class of singular kernels: Application to the Patlak–Keller–Segel model

Limites de champ moyen pour des noyaux singuliers et applications au modèle de Patlak–Keller–Segel

Didier Bresch [a,1], Pierre-Emmanuel Jabin [b,2], Zhenfu Wang [c]

[a] *LAMA – UMR5127 CNRS, Bât. Le Chablais, Campus scientifique, 73376 Le Bourget-du-Lac, France*
[b] *CSCAMM and Dept. of Mathematics, University of Maryland, College Park, MD 20742, USA*
[c] *Dept. of Mathematics, University of Pennsylvania, Philadelphia, PA 19104, USA*

ARTICLE INFO

Article history:
Received 11 June 2019
Accepted 19 September 2019
Available online 30 September 2019

Presented by Laure Saint-Raymond

ABSTRACT

In this note, we propose a modulated free energy combination of the methods developed by P.-E. Jabin and Z. Wang [Inventiones (2018)] and by S. Serfaty [Proc. Int. Cong. Math. (2018) and references therein] to treat more general kernels in mean-field limit theory. This modulated free energy may be understood as introducing appropriate weights in the relative entropy developed by P.-E. Jabin and Z. Wang (in the spirit of what has been recently developed by D. Bresch and P.-E. Jabin [Ann. of Math. (2) (2018)]) to cancel the most singular terms involving the divergence of the flow. Our modulated free energy allows us to treat singular potentials that combine large smooth part, small attractive singular part, and large repulsive singular part. As an example, a full rigorous derivation (with quantitative estimates) of some chemotaxis models, such as the Patlak–Keller–Segel system in subcritical regimes, is obtained.

RÉSUMÉ

Dans cette note, on propose une énergie libre modulée combinant les méthodes développées par P.-E. Jabin et Z. Wang [Inventiones (2018)] et par S. Serfaty [voir l'article de revue Proc. Int. Cong. Math. (2018) et ses références] pour traiter des noyaux plus généraux en théorie de la limite champ moyen. Cette énergie libre modulée consiste en l'introduction d'une famille de poids appropriés dans l'entropie relative développée par P.-E. Jabin et Z. Wang (dans le même esprit que dans le travail récent de D. Bresch et P.-E. Jabin [Ann. of Math. (2) (2018)]) pour compenser les termes les plus singuliers qui font intervenir la divergence du champ de vitesse. Comme exemple, une preuve avec estimation quantitative de la limite champ moyen vers le modèle de Patlak–Keller–Segel en régime sous-critique est obtenue. Notre méthode permet également de couvrir des potentiels singuliers qui

E-mail addresses: Didier.Bresch@univ-smb.fr (D. Bresch), pjabin@cscamm.umd.edu (P.-E. Jabin), zwang423@math.upenn.edu (Z. Wang).

1 D. Bresch is partially supported by SingFlows project, grant ANR-18-CE40-0027.
2 P.-E. Jabin is partially supported by NSF DMS Grant 161453, 1908739, and NSF Grant RNMS (Ki-Net) 1107444.

https://doi.org/10.1016/j.crma.2019.09.007

peuvent combiner une partie réguliere, une petite partie singulière attractive et une grande partie singulière répulsive.

Version française abrégée

Dans cette note, on présente une entropie relative permettant d'encoder quantitativement la limite champ moyen $N \to +\infty$ pour un système de particules avec potentiels d'interaction singuliers pouvant contenir une partie régulière qui peut être grande, une partie singulière attractive de faible amplitude et une grande partie singulière répulsive. Une preuve avec estimation quantitative de la limite champ moyen vers le modèle de Patlak–Keller–Segel en régime sous-critique est, par exemple, obtenue en considérant le potentiel attractif de Poisson en dimension 2. Par souci de simplicité, on se restreint dans cette note au domaine périodique Π^d avec $d = 1, 2, 3$. On considère la limite champ moyen $N \to +\infty$ pour le système d'ordre 1 suivant à N particules

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}B_i, \quad i = 1, 2, \cdots, N, \tag{1}$$

où les B_i sont des mouvements browniens ou des processus de Wiener indépendants ; on considère un champ de vitesse type flot–gradient avec un noyau singulier donné par

$$K = -\nabla V. \tag{2}$$

Un exemple est donné par le modèle de Patlak–Keller–Segel, que l'on rencontre en biologie pour modéliser le chimiotactisme, où la position X_i de chaque micro-organisme suit le gradient d'une concentration d'un chimio-attractant représenté par $\frac{1}{N}\sum_i V(x - X_i)$. En dimension 2, V est typiquement donné par le noyau de Poisson

$$V = \lambda \log|x| + V_{\mathrm{e}}(x) \tag{3}$$

avec $\lambda > 0$ et où V_{e} est une correction régulière permettant à V d'être périodique. Le but est alors de donner une estimation quantitative de la convergence vers la solution du système de McKean–Vlasov

$$\partial_t \bar\rho + \mathrm{div}_x(\bar\rho\, K \star_x \bar\rho) = \sigma\,\Delta_x \bar\rho, \quad \text{avec} \quad K = -\nabla V, \quad \bar\rho(t=0, x) = \bar\rho^0 \in \mathcal{P}(\Pi^d), \tag{4}$$

autrement dit, de quantifier à quel point la distribution des positions aléatoires X_i obtenues par le système (1) peut être approchée par la densité déterministe $\bar\rho$, solution de (4). Ce type de résultat implique donc une preuve de la propagation du chaos.

Il existe différents moyens de comparer (1) avec la limite $\bar\rho$ donnée par (4). On suit ici [17,18] en introduisant la loi $\rho_N(t, x_1, \ldots, x_N)$ satisfaisant l'équation de Liouville

$$\partial_t \rho_N + \sum_{i=1}^N \mathrm{div}_{x_i}\Big(\rho_N \frac{1}{N}\sum_{j\neq i}^N K(x_i - x_j)\Big) = \sigma \sum_{i=1}^N \Delta_{x_i}\rho_N, \qquad \rho_N|_{t=0} = \rho_N^0 \text{ avec } \int_{\Pi^{dN}} \rho_N^0 = 1. \tag{5}$$

À titre d'exemple, dans le cas du potentiel (2) en dimension d, on obtient le résultat suivant.

Théorème 0.1. Supposons que $\rho_N \in L^\infty(0, T;\ L^1(\Pi^{Nd}))$ est solution faible de (5) avec la donnée initiale $\rho_N(t=0) = \bar\rho^{\otimes N}(t=0)$, que ρ_N vérifie une inégalité d'entropie et que le potentiel d'interaction V est donné par (3). Supposons que $\bar\rho \in L^\infty(0, T; W^{2,\infty}(\Pi^d))$ est solution de (4) avec $\inf \bar\rho > 0$ et que $\lambda < 2d\sigma$. Alors il existe une constante C et un exposant $\theta > 0$ indépendants de N tels que, pour tout k fixé :

$$\|\rho_{N,k} - \bar\rho^{\otimes k}\|_{L^\infty(0,T;\ L^1(\Pi^{kd}))} \le C\, k^{1/2}\, N^{-\theta}$$

où $\rho_{N,k}$ est la marginale du système au rang k, c'est-à-dire

$$\rho_{N,k}(t, x_1, \cdots, x_k) = \int_{\Pi^{(N-k)d}} \rho_N(t, x_1, \cdots, x_N)\,\mathrm{d}x_{k+1}\cdots\mathrm{d}x_N.$$

Remarque. Nous obtenons, en dimension 2, la constante optimale 4σ qui correspond à la masse critique $8\pi\sigma$ à partir de laquelle on a explosion en temps fini sur Patlak–Keller–Segel. Rappelons que nous avons normalisé $\overline{\rho}_0$ pour avoir une masse totale 1. Ce résultat précise le résultat obtenu dans [12] en donnant la première justification rigoureuse de la propagation du chaos avec un taux de convergence explicite.

Remarque. Un résultat plus général peut, en fait, être obtenu, permettant de considérer des potentiels V du type

$$V(-x) = V(x), \qquad V = V_a + V_r + V_s, \tag{6}$$

avec V_a un potentiel singulier attractif contraint, V_r un potentiel singulier répulsif assez général et V_s un potentiel suffisamment régulier. Par souci de clarté dans une note aux *Comptes rendus*, la preuve sur Patlak–Keller–Segel a été privilégiée et seule une remarque sera donnée sur V_s, V_r et V_a. Le résultat ainsi que sa preuve pour ce type de noyaux plus généraux seront donnés dans la version longue [5].

La preuve générale est basée sur une énergie libre modulée (14), qui est une combinaison des méthodes développées par P.-E. Jabin et Z. Wang [Inventiones (2018)] et par S. Serfaty [voir l'article de revue Proc. Int. Cong. Math. (2018) et ses références]. Elle consiste en l'introduction d'une famille de poids appropriés (équilibre de Gibbs G_N et distribution correspondante $G_{\overline{\rho}_N}$ définis dans (15)) dans l'entropie relative développée par P.-E. Jabin et Z. Wang pour compenser les termes les plus singuliers qui font intervenir la divergence du champ de vitesse. À noter qu'une énergie libre combinaison d'entropie et d'énergie modulée a été utilisée dans d'autres contextes, et notamment dans des problèmes de limites singulières en théorie cinétique : le lecteur est renvoyé notamment à [1] pour les limites vers des systèmes de MHD depuis Vlasov–Maxwell–Boltzmann et à [22], [14] pour les limites quasi neutres. La nouveauté ici réside dans la façon de définir la bonne combinaison entropie–énergie modulée dans le cadre des systèmes de particules pour obtenir la proposition 2.1 et dans les estimations spécifiques obtenues sur cette combinaison.

1. Introduction

In this note, we consider the limit when $N \to +\infty$ of the following system of N particles,

$$\mathrm{d}X_i = \frac{1}{N}\sum_{j\neq i} K(X_i - X_j)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}B_i, \quad i = 1, 2, \cdots, N, \tag{7}$$

where the B_i are independent Brownian Motions or Wiener processes. For simplicity, we restrict ourselves in this note to the periodic domain Π^d (with $d = 1, 2, 3, \cdots$). Bounded domains with the right boundary conditions can be treated in a similar manner, but the case of the whole space $\mathbb{R}^d$ presents technical difficulties.

We specifically consider gradient flows with the interaction kernel given by

$$K = -\nabla V. \tag{8}$$

A guiding example in this note is the *attractive* Poisson potential in dimension 2

$$V = \lambda \log|x| + V_e(x), \tag{9}$$

with $\lambda > 0$ and where V_e is a smooth correction such that V is periodic. Logarithmic potentials still play a critical role if the dimension $d > 2$; for this reason, we will still consider potentials like (9) in any dimension, even if there is no connection with the Poisson equation anymore.

Our main goal is to provide precise quantitative estimates for the convergence of (7) towards the limit McKean–Vlasov PDE

$$\partial_t\bar\rho + \mathrm{div}_x(\bar\rho\, K \star_x \bar\rho) = \sigma\Delta_x\bar\rho, \quad \text{with} \quad K = -\nabla V, \quad \bar\rho(t=0,x) = \bar\rho^0 \in \mathcal{P}(\Pi^d) \tag{10}$$

in the sense of Theorem 1.1. The density $\bar\rho$ is then in particular the limit of the 1-particle distribution of System (7).

In the case when V is given by (9) and $d = 2$, then (10) is the famous Patlak–Keller–Segel model, which is one of the first models of chemotaxis for micro-organisms. The potential $-V \star \bar\rho$ can then be seen as the concentration of some chemical species (one has typically $V \leq 0$ here). From (9), one has that $\Delta V - V = 2\pi\lambda\delta_0$, meaning that the chemical species are produced by the population. Moreover, (10) implies that the population follows the direction of higher chemical concentrations (more negative values of V).

Although it offers only a rough modeling of the biological processes involved in chemotaxis, the Patlak–Keller–Segel model is a good example of a singular attractive dynamics (all micro-organisms try to concentrate on a point) competing with the spreading effect due to diffusion. The potential $V \sim \log|x|$ is actually critical in dimension 2 in the sense that Eq. (10) may blow-up and form a Dirac mass in finite time. Still in dimension 2, one may exactly characterize that such a blow-up occurs if and only if $\lambda > 4\sigma$ since we normalize $\bar\rho^0$ to have total mass 1. Accordingly, Eq. (10) is subcritical when $\lambda < 4\sigma$ and global in time solutions then exist. We refer for instance to [3,8,9] and the references therein.

Because of this singular behavior of the potential, a full rigorous derivation of the Patlak–Keller–Segel model from the stochastic equation (7) has remained elusive, in spite of recent progress in [7,12] or [13,15]. The results in [12], for example, prove that any accumulation point as $N \to \infty$ of the random empirical measure associated with the system (7) is a weak solution to (10) in the so-called very subcritical regime with $\lambda < \sigma$. While this provides the mean-field limit, at least in some weak sense, it does not imply propagation of chaos. Of course, potentials like (9) are only one example of singular interactions between particles for which the mean-field limit remains poorly understood, especially in the stochastic cases.

There are several ways to quantitatively compare (7) with the limit $\bar\rho$ given by (10). We follow here [17,18] by using the joint law $\rho_N(t, x_1, \dots, x_N)$ of the process $(X_1, \dots, X_N)$, which solves the Liouville or Kolmogorov forward equation

$$\partial_t \rho_N + \sum_{i=1}^N \mathrm{div}_{x_i}\Big(\rho_N \frac{1}{N}\sum_{j\neq i}^N K(x_i - x_j)\Big) = \sigma \sum_{i=1}^N \Delta_{x_i}\rho_N, \qquad \rho_N|_{t=0} = \rho_N^0. \tag{11}$$

Eq. (11) offers a straightforward manner to understand solutions to system (7), which is actually non-trivial when dealing with singular potentials V. We mostly defer to the coming paper [5] for a complete discussion of the notion of solutions that we consider and which we simply call *entropy solutions* throughout this note. Those roughly correspond to solutions in the sense of distribution to (11) that also satisfy appropriate entropy and energy bounds.

The joint law ρ_N is compared to the chaotic law $\bar\rho_N := \bar\rho^{\otimes^N} = \Pi_{i=1}^N \bar\rho(t, x_i)$, built from the limit $\bar\rho$. Of course $\bar\rho_N$ cannot be an exact solution to (11) but instead solves

$$\partial_t \bar\rho_N + \sum_{i=1}^N \mathrm{div}_{x_i}\Big(\bar\rho_N\, K \star_x \bar\rho(x_i)\Big) = \sigma \sum_{i=1}^N \Delta_{x_i}\bar\rho_N. \tag{12}$$

As probability densities, both ρ_N and $\bar\rho_N$ are initially normalized by

$$\int_{\Pi^{dN}} \rho_N|_{t=0} = 1 = \int_{\Pi^d} \bar\rho|_{t=0}, \tag{13}$$

which is formally preserved by either (11) or (10)/(12). The method leads in particular to direct estimates between $\bar\rho^{\otimes k}$ and any observable or marginal of the system at a fixed rank k,

$$\rho_{N,k}(t, x_1, \dots, x_k) = \int_{\Pi^{(N-k)\,d}} \rho_N(t, x_1, \dots, x_N)\, \mathrm{d}x_{k+1} \dots x_N.$$

A corollary of the analysis sketched in this note and fully developed in the coming paper [5] is a rigorous derivation of the Patlak–Keller–Segel system in some subcritical regimes, typically the following result.

Theorem 1.1. *Assume that $\rho_N \in L^\infty(0,T;\ L^1(\Pi^{Nd}))$ is an entropy solution to Eq. (11) normalized by (13), with initial condition $\rho_N(t=0) = \bar\rho^{\otimes^N}(t=0)$, and for the potential V given by (9). Assume that $\bar\rho \in L^\infty(0,T; W^{2,\infty}(\Pi^d))$ solves Eq. (10) with $\inf\bar\rho > 0$. Assume finally that $\lambda < 2d\sigma$. Then there exists a constant $C > 0$ and an exponent $\theta > 0$ independent of N such that, for any fixed k,*

$$\|\rho_{N,k} - \bar\rho^{\otimes^k}\|_{L^\infty(0,T;\ L^1(\Pi^{kd}))} \le C\, k^{1/2}\, N^{-\theta}.$$

Remark. We get, in dimension 2, the optimal constant 4σ that corresponds to the critical mass $8\pi\sigma$ for which we have blow-up in finite time for Patlak–Keller–Segel. We do not obtain the optimal exponent $\theta = 1/2$ however, at least for potentials with logarithmic singularity such as given by (9). We cannot prove the limit if $\lambda > 2d\sigma$, though we conjecture that such a limit should still hold before the blow-up.

Theorem 1.1 follows directly from Theorem 2.2 stated later in this note and the classical Csiszár–Kullback–Pinsker inequality. The exponent θ could be made fully explicit and actually depends only on $2d\sigma - \lambda$.

The note is organized as follows: in Section 2, we introduce the new relative entropy with weights related to the Gibbs equilibrium G_N and the corresponding distribution $G_{\bar\rho_N}$ given by (15) in Section 2. We present the explicit expression for the time evolution of this modulated free energy in Proposition 2.1 and we explain the mathematical strategy to get, from this inequality, a quantitative estimate on the free energy (see Theorem 2.2) that provides the explicit propagation of chaos. In Section 3, we briefly sketch the main arguments in the proof of Theorem 2.2; the details and the full proof will be given in the paper [5]. We conclude with an appendix that contains some mathematical tools related to the large-deviation estimates that are used.

2. A new relative entropy and the mathematical strategy

2.1. The modulated free energy

The main idea of the method is to obtain quantitative estimates on the free energy, which is the natural physical notion for stochastic systems such as (7). Such a modulated free energy can actually be written as a relative entropy between the joint law ρ_N, the chaotic law $\bar{\rho}_N$ and the corresponding Gibbs equilibria, leading to

$$E_N(\frac{\rho_N}{G_N} \mid \frac{\bar{\rho}_N}{G_{\bar{\rho}_N}}) = \frac{1}{N} \int_{\Pi^{dN}} \rho_N(t, X^N) \log\Big(\frac{\rho_N(t, X^N)}{G_N(X^N)} \frac{G_{\bar{\rho}_N}(t, X^N)}{\bar{\rho}_N(t, X^N)} \Big) \, \mathrm{d}X^N, \tag{14}$$

where we denote by G_N the Gibbs equilibrium of the system (7), and by $G_{\bar{\rho}_N}$ the corresponding distribution where the exact field is replaced by the mean-field limit according to the law $\bar{\rho}$, *i.e.*

$$\begin{aligned} G_N(t, X^N) &= \exp\Big(-\frac{1}{2N\sigma} \sum_{i \neq j} V(x_i - x_j) \Big), \\ G_{\bar{\rho}}(t, x) &= \exp\Big(-\frac{1}{\sigma} V \star \bar{\rho}(x) + \frac{1}{2\,\sigma} \int_{\Pi^d} V \star \bar{\rho}\, \bar{\rho} \Big), \\ G_{\bar{\rho}_N}(t, X^N) &= \exp\Big(-\frac{1}{\sigma} \sum_{i=1}^{N} V \star \bar{\rho}(x_i) + \frac{N}{2\sigma} \int_{\Pi^d} V \star \bar{\rho}\, \bar{\rho} \Big). \end{aligned} \tag{15}$$

One may decompose E_N as follows:

$$E_N(\frac{\rho_N}{G_N} \mid \frac{\bar{\rho}_N}{G_{\bar{\rho}_N}}) = \mathcal{H}_N(\rho_N|\bar{\rho}_N) + \mathcal{K}_N(G_N|G_{\bar{\rho}_N}),$$

where

$$\mathcal{H}_N(\rho_N|\bar{\rho}_N) = \frac{1}{N} \int_{\Pi^{dN}} \rho_N(t, X^N) \log\Big(\frac{\rho_N(t, X^N)}{\bar{\rho}_N(t, X^N)} \Big) \, \mathrm{d}X^N$$

is exactly the relative entropy introduced in [17,18] and

$$\mathcal{K}_N(G_N|G_{\bar{\rho}_N}) = -\frac{1}{N} \int_{\Pi^{dN}} \rho_N(t, X^N) \log(\frac{G_N(t, X^N)}{G_{\bar{\rho}_N}(t, X^N)}) \, \mathrm{d}X^N$$

is the expectation of the modulated energy developed, for instance, in [10,25,24].

Combining the relative entropy with a modulated energy has already been very successfully used for various singular limits in kinetic theory. A first example concerns the so-called quasineutral limit for plasmas for which we refer to [14] and [22]. Another example is the seminal derivation of the incompressible viscous electro-magneto-hydrodynamics from the Vlasov–Maxwell–Boltzmann system in [1]; one issue in that monograph is in particular to prove the asymptotic positivity of the combined free energy, which is a problem that we are facing as well, as explained below.

In the more specific context of the mean-field limits, our approach may hence be seen *as a combination of the two methods respectively introduced in* [18] *and in* [25].

- The article [18] compares ρ_N with $\bar{\rho}_N$ through the rescaled relative entropy $\mathcal{H}_N(\rho_N|\bar{\rho}_N)$. This leads to a quantitative version of chaos propagation for systems (7) under the assumption that the kernel K belongs to $\dot{W}^{-1,\infty}$ with $\operatorname{div} K \in \dot{W}^{-1,\infty}$. This, in particular, implied the first quantitative derivation of the incompressible Navier–Stokes system in dimension 2 from point vortices with diffusion (in that case $\operatorname{div} K = 0$), where only qualitative limits were previously available in [11,20,21].
  More precisely, the Navier–Stokes case corresponds to taking $K = \operatorname{curl} V$, with V still given by (9) and, at first glance, it may hence appear that the derivation of the Patlak–Keller–Segel with the kernel K given by $K = -\nabla V$ should follow similarly. However, the requirement $\operatorname{div} K \in \dot{W}^{-1,\infty}$ breaks the comparison: $\operatorname{div} K = 0$ if $K = \operatorname{curl} V$, while $\operatorname{div} K \in \dot{W}^{-1,\infty}$ forces V to be log-Lipschitz if $K = -\nabla V$. In general and because of this condition on the divergence, the result in [18] *underperforms in the gradient flow setting that we consider here*. The interested reader is further referred to [23] for an introduction on this work in the Bourbaki seminar.

– The strategy followed in [25,24] is implemented for deterministic system ($\sigma = 0$ in (7)) and consists in using the modulated potential energy. Written in terms of the empirical measure $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta(x - X_i(t))$, the modulated energy reads

$$\frac{1}{2}\int_{\Pi^{2d}\cap\{x\neq y\}} V(x-y)\,(\mu_N(\mathrm{d}x) - \bar\rho(x)\,\mathrm{d}x)\,(\mu_N(\mathrm{d}y) - \bar\rho(y)\,\mathrm{d}y),$$

so that $\sigma\,\mathcal{K}_N(G_N|G_{\bar\rho_N})$ is the expectation w.r.t. ρ_N

$$\mathcal{K}_N(G_N|G_{\bar\rho_N}) = \frac{1}{2\sigma}\int_{\Pi^{N\,d}}\int_{\Pi^{2\,d}\cap\{x\neq y\}} V(x-y)\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}\mathrm{d}\rho_N. \tag{16}$$

In the deterministic setting, the corresponding method provides the quantitative rate of convergence for any Riesz potential $V(x) = c\,|x|^{-\alpha}$, with $c > 0$ and $\alpha < d$. This not only applies to gradient flows, but also to some Hamiltonian flows such as $K = c\,\mathrm{curl}\,|x|^{-\alpha}$ in dimension 2. Because one may use $\alpha > d-2$, the result goes beyond Coulombian interactions and the type of singularity that could be handled in [16], for example.
The approach in [25,24] is however limited to repulsive potentials (the modulated energy is otherwise not positive) and to exact Riesz potentials with the possible addition of a smooth perturbation.

The hope is hence that adding $\mathcal{K}_N(G_N|G_{\bar\rho_N})$ and $\mathcal{H}_N(\rho_N|\bar\rho_N)$ will allow us to combine the control on potentials with very singular but specific structure in [25] with estimates in [18] that require less structure but cover less singular potentials. In some sense, the proposed modulated free energy may also be understood as introducing appropriate weights in the relative entropy used in [18] (in the spirit of what has been recently developed by some of the authors in [4] in an other framework) to cancel the most singular terms involving the divergence of the flow, namely $\mathrm{div}\,K = -\Delta V$.

2.2. Our bound on the modulated free energy

The first step is obviously to calculate the time evolution of the modulated free energy E_N.

Proposition 2.1. *Assume that V is an even function, and that ρ_N is an entropy solution to (11) and $\bar\rho$ solves (10). Then the modulated free energy defined by (14) satisfies*

$$E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)(t) - E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)(0) \le -\frac{\sigma}{N}\int_0^t\int_{\Pi^{d\,N}} \mathrm{d}\rho_N\left|\nabla\log\frac{\rho_N}{\bar\rho_N} - \nabla\log\frac{G_N}{G_{\bar\rho_N}}\right|^2$$
$$-\frac{1}{2}\int_0^t\int_{\Pi^{dN}}\int_{\Pi^{2\,d}\cap\{x\neq y\}} \nabla V(x-y)\cdot\left(\nabla\log\frac{\bar\rho}{G_{\bar\rho}}(x) - \nabla\log\frac{\bar\rho}{G_{\bar\rho}}(y)\right)(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}\mathrm{d}\rho_N, \tag{17}$$

where $\mu_N = \frac{1}{N}\sum_{i=1}^N \delta(x - x_i)$ is the empirical measure.

The strategy for obtaining the convergence to the chaotic solution $\bar\rho^{\otimes N}$ from Proposition 2.1 then relies on proving the two following points.

– Show that $E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)$ effectively controls the distance between ρ_N and $\bar\rho_N$. Note that E_N is not a priori a positive quantity, since ρ_N/G_N and $\bar\rho_N/G_{\bar\rho_N}$ do not have the same mass. In this note, our goal will be to obtain estimates like

$$E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right) \ge \frac{1}{C}\,\mathcal{H}_N(\rho_N\,|\,\bar\rho_N) - \frac{C}{N^\theta}, \tag{18}$$

for some constant C and exponent θ independent of N. As a matter of fact, even for *smooth, attractive* potentials, it may occur that E_N is negative with $\mathcal{K}_N$ negative and dominating $\mathcal{H}_N$. This issue forced the limitation to repulsive potentials in [25], but fortunately for our present method, a straightforward solution in the stochastic case is simply to remove from V any large smooth part that could create issues.
– Control the right-hand side in Proposition 2.1 by $E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)$ or individually by $\mathcal{H}_N(\rho_N\,|\,\bar\rho_N)$ or $\mathcal{K}_N\left(G_N\,|\,G_{\bar\rho_N}\right)$ plus some vanishing in N correction of the form C/N^θ.

While it may first appear that the second point is more complex, it is not necessarily so. In fact, in the Patlak–Keller–Segel case with V given by (9), the right-hand side in Proposition 2.1 may be directly controlled by the large-deviation inequality provided in [18] and it is Estimate (18) of the modulated free energy E_N that is the delicate point. Note that the renormalized modulated entropy may be negative for fixed parameters, but it recovers asymptotically a non-negative quantity. This feature can be observed in other settings and we mention in particular the recent seminal work [1] that was cited above (see also references therein).

However, specific tools are needed for the present (and different) context of the mean-field limit for many-particle systems, and those tools constitute the main contribution of this note and the upcoming paper [5]. Details regarding the proof in the Patlak–Keller–Segel case will be given in Section 3 with the explanation on the conclusion of the proof in Subsection 3.4.

The heart of the proof, both to control non-negativity of E_N and to control the right-hand sides of (17) in Proposition 2.1, consists in estimating terms of the form

$$\int_{\Pi^{dN}} \mathrm{d}\rho_N \int_{\Pi^{2d}} f(x,y)\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2} \tag{19}$$

in terms of $\mathcal{H}_N$ or $\mathcal{K}_N$, and for suitably smooth functions f. The bounds are sketched in Section 3 and rely on large-deviation tools, some of which are briefly presented in Appendix A. Since we focus in this note on attractive singular potentials, the main issue is actually to control E_N through lower bounds on $\mathcal{K}_N$ given by (16). This leads us to take $f(x,y) = V(x-y)\,1_{\{x\neq y\}}$ above and to define the specific functional

$$F(\mu) = -\int_{\Pi^{2d}\cap\{x\neq y\}} V(x-y)(\mathrm{d}\mu - \mathrm{d}\overline{\rho})^{\otimes 2}. \tag{20}$$

We emphasize in particular that

$$\mathcal{K}_N = -\frac{1}{2\sigma}\mathbb{E}_{\rho_N}(F(\mu_N)) = -\frac{1}{2\sigma}\int_{\Pi^{dN}} \rho_N\, F(\mu_N)\,\mathrm{d}x_1\ldots\mathrm{d}x_N.$$

The main bounds that can be obtained on $F(\mu)$ are sketched in the next section.

We now state our precise main result (Theorem 2.2).

Theorem 2.2. *Let $\sigma > 0$ and let V be an even potential $V(-x) = V(x)$ satisfying*

$$V \in L^p(\Pi^d) \cap C^2(\Pi^d \setminus \{0\}) \quad \textit{for some } p > 1, \tag{21}$$

$$V(x) \geq \gamma\,\log|x| + C \quad \textit{for some } 0 \leq \gamma < 2\,d\,\sigma, \tag{22}$$

$$|\nabla V(x)| \leq \frac{C}{|x|}, \tag{23}$$

with $C > 0$ constant. Let $\bar\rho$ be a solution to Eq. (10) with $\inf\bar\rho > 0$, $\bar\rho \in L^\infty\big(0,T;W^{2,\infty}(\Pi^d)\cap W^{d+3,1}(\Pi^d)\cap\mathcal{P}(\Pi^d)\big)$. There exists a constant $C > 0$ and an exponent $\theta > 0$ independent of N such that, for any entropy solution ρ_N to Equation (11), we have

$$E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)(t) \leq \left(E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)(t=0) + C\,\mathcal{H}_N\,(\rho_N\,|\,\bar\rho_N)\,(t=0) + \frac{C}{N^\theta}\right)\exp(Ct),$$

and the explicit bound on the relative entropy

$$\mathcal{H}_N\,(\rho_N\,|\,\bar\rho_N)\,(t) \leq C\left(E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)(t=0) + C\,\mathcal{H}_N\,(\rho_N\,|\,\bar\rho_N)\,(t=0) + \frac{C}{N^\theta}\right)\exp(Ct).$$

Remark 1. We want to emphasize that Theorem 2.2 holds in the whole sub-critical regime for the Patlak–Keller–Segel system since (22) only requires sub-critical masses $\gamma < 2\,d\,\sigma$.

Remark 2. Assumptions (21)–(22) are specifically required to obtain the lower bound on E_N namely (18) while Assumption (23) is the only one needed to control the right-hand side in (17).

This last assumption can be considerably relaxed if the corresponding part is *repulsive*. More precisely, in [5], we can have $V = V_a + V_r$, with V_a a possibly attractive part satisfying (21)–(22)–(23) and $V_r \geq 0$ a purely repulsive part that still satisfies (21), but with only $|\nabla V_r| \leq \frac{C}{|x|^k}$ for any exponent k provided that a specific control on the Fourier transform is provided, namely

$$|\nabla_\xi \hat V_r(\xi)| \leq C\,\frac{\hat V_r(\xi)}{1+|\xi|} + \frac{C}{1+|\xi|^{d+1}}, \tag{24}$$

which is still required to bound the right-hand side in Proposition 2.1. Note that such class of singular repulsive potentials include all Riesz potentials $V_r = |x|^{-k}$ with $k < d$ considered for instance in [25], [24] and references cited therein, but also allow mildly singular perturbation $V_r = |x|^{-k} + W(x)$, where $|\nabla_\xi \hat{W}| \le C(1 + |\xi|)^{d-1}$. Condition (24) can replace the extension theorem from [6] used in [25,24].

Remark 3. All constants and exponents in Theorem 2.2 can be computed with explicit dependence on the various norms of $\bar\rho$ and V, which are assumed to be finite.

We finally stress that the scaled relative entropy $\mathcal{H}_{\mathcal{N}}$ controls the relative entropy of all marginals with

$$\frac{1}{k}\int_{\Pi^{dk}} \mathrm{d}\rho_{N,k} \log\frac{\rho_{N,k}}{\bar\rho^{\otimes k}} \le \mathcal{H}_N(\rho_N \mid \bar\rho_N).$$

By the Csiszár–Kullback–Pinsker inequality, we hence have that

$$\|\rho_{N,k} - \bar\rho^{\otimes k}\|_{L^1} \le \sqrt{2k}\,(\mathcal{H}_N(\rho_N \mid \bar\rho_N))^{1/2}$$

which leads to Theorem 1.1.

3. The Patlak–Keller–Segel potential: proof of Theorem 2.2

3.1. Upper bound of the right-hand side in Proposition 2.1

Under assumption (23), already existing results allow us to directly bound the right-hand side of (17) in Proposition 2.1. We refer more specifically to the large-deviation inequality first proved in [18] and recalled in Theorem A.2. Defining

$$f(x,y) = -\nabla V(x-y)\cdot\left(\nabla\log\frac{\bar\rho}{G_{\bar\rho}}(x) - \nabla\log\frac{\bar\rho}{G_{\bar\rho}}(y)\right),$$

we observe that Assumption (23), namely $|\nabla V(x)| \le C/|x|$, ensures that $f \in L^\infty$, simply by using that $\nabla\log\bar\rho$, $\nabla V \star \bar\rho \in W^{1,\infty}$ (see Definition (15) of $G_{\bar\rho}$). We may then combine Theorem A.2 with Lemma A.1 to get

$$-\int_{\Pi^{dN}}\int_{\Pi^{2d}\cap\{x\neq y\}} \nabla V(x-y)\cdot\left(\nabla\log\frac{\bar\rho}{G_{\bar\rho}}(x) - \nabla\log\frac{\bar\rho}{G_{\bar\rho}}(y)\right)(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}\,\mathrm{d}\rho_N \le C\,\mathcal{H}_N(\rho_N \mid \bar\rho_N) + \frac{C}{N}. \tag{25}$$

3.2. Lower-bound control of the free energy E_N

Unfortunately, we cannot use again Theorem A.2 to provide a lower bound on the modulated free energy E_N when V has attractive singularities like in (9). Applied to $f(x,y) = V(x-y)$, Theorem A.2 would require V to be bounded. Instead, we directly study the behavior of $\mathcal{K}_N$ to prove the following proposition.

Proposition 3.1. *Assume that V satisfies (21)–(22) together with*

$$|\nabla V(x)| \le \frac{C}{|x|^k} \qquad \text{for some } k.$$

Then there exist $\delta < 1$ and $\eta > 0$ depending only on $\|V\|_{L^p}$ and γ such that, for any smooth function χ with $\chi(x) = 1$ if $x < 1/2$ and supp $\chi \in [0,1]$,

$$-\frac{1}{2\sigma}\int_{\Pi^{dN}}\int_{\Pi^{2}\cap\{x\neq y\}} V(x-y)\,\chi(|x-y|/\eta)\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}\,\rho_N\,\mathrm{d}X^N$$
$$\le \delta\,\mathcal{H}_N(\rho_N \mid \bar\rho_N) + \frac{C}{N^{1/(2d+1)}}(\|\log\bar\rho\|_{W^{1,\infty}} + \eta^{-1}) + \frac{C}{N^{1/(2(2d+1))}}.$$

Remark 4. By definition of the truncation χ, Proposition 3.1 controls the singular part of V, but it does not control the long-range part. It is actually unclear and even doubtful that an equivalent of Proposition 3.1 would hold without such a truncation. This requires a specific treatment of the long-range part of V, which will be explained in the next subsection.

Sketch of the proof. To give a brief sketch of the proof of Proposition 3.1, denote similarly to $F(\mu)$:

$$F_\eta(\mu) = -\int_{\Pi^{2d}\cap\{x\neq y\}} V(x-y)\,\chi(|x-y|/\eta)\,(\mu(\mathrm{d}x)-\bar\rho(x)\,\mathrm{d}x)\,(\mu(\mathrm{d}y)-\bar\rho(y)\,\mathrm{d}y).$$

Choosing some $\delta<1$ such that $\gamma/\delta<2\,d\,\sigma$, a straightforward use of Lemma A.1 shows that Proposition 3.1 is implied by the large-deviation inequality:

$$\frac{1}{N}\log\int_{\Pi^{dN}}\exp\left(\frac{N}{2\sigma\,\delta}\,F_\eta(\mu_N)\right)\bar\rho_N\,\mathrm{d}X^N \le \frac{C}{N^{1/(2d+1)}}\,(\|\log\bar\rho\|_{W^{1,\infty}}+\eta^{-1})+\frac{C}{N^{1/2\,(2d+1)}}. \tag{26}$$

The proof of (26) is intricate and the main difficulty lies in replacing $F_\eta(\mu)$ in (26) by $F_\eta(L_\varepsilon\star\mu)$ for some classical convolution kernel L_ε. More specifically, we can show that, for $\lambda>1$,

$$\begin{aligned}\frac{1}{N}\log\int_{\Pi^{dN}}\exp\left(\frac{N}{2\sigma\,\delta}\,F_\eta(\mu_N)\right)\bar\rho_N\,\mathrm{d}X^N &\le C\,\varepsilon^{1/2}+C\,\varepsilon\,(\|\bar\rho\|_{L^\infty}+\eta^{-1})\\ &+\frac{C_\lambda}{N}\log\int_{\Pi^{dN}}\exp\left(\frac{\lambda N}{2\sigma\,\delta}\,F_\eta(L_\varepsilon\star\mu_N)\right)\bar\rho_N\,\mathrm{d}X^N.\end{aligned} \tag{27}$$

The price to pay for such a regularization is the extra factor λ in the exponential with the requirement $\lambda>1$ (as $C_\lambda\to\infty$ as $\lambda\to1$). In our present subcritical setting, we choose some $\lambda>1$, so that we still have that $\frac{\lambda\,\gamma}{\delta}<2\,d\,\sigma$.

The functional $\mu\to F(L_\varepsilon\star\mu)$ is now continuous on measures, so that we are left with quantifying classical large-deviation results. An example of such a quantified inequality is Theorem A.3, stated in the appendix. Applying it, one obtains that

$$\begin{aligned}\frac{C_\lambda}{N}\log\int_{\Pi^{dN}}\exp\left(\frac{\lambda N}{2\sigma\,\delta}\,F_\eta(L_\varepsilon\star\mu_N)\right)\bar\rho_N\,\mathrm{d}X^N &\le I\left(\frac{\lambda}{2\sigma\,\delta}\,F_\eta\right)\\ &+\frac{C}{N^{1/(d+1)}\,\varepsilon^{d/(d+1)}}\,(\log N+|\log\varepsilon|+\|\log\bar\rho\|_{L^\infty})+C\,\varepsilon\,\|\log\bar\rho\|_{W^{1,\infty}},\end{aligned} \tag{28}$$

which connects the desired bound with the large-deviation functional $I\left(\frac{\lambda}{2\sigma\,\delta}\,F_\eta\right)$, which reads

$$I\left(\frac{\lambda}{2\sigma\,\delta}\,F_\eta\right)=\sup_{\mu\in\mathcal{P}(\Pi^d)}\frac{\lambda}{2\sigma\,\delta}\,F_\eta(\mu)-\int_{\Pi^d}\mathrm{d}\mu\,\log\frac{\mu}{\bar\rho}. \tag{29}$$

The final step in the proof is hence to show that $I\left(\frac{\lambda}{2\sigma\,\delta}\,F_\eta\right)=0$ (given that $\frac{\lambda\gamma}{2\sigma\delta}<d$), which requires the truncation η. By employing the logarithmic Hardy–Littlewood–Sobolev inequality, which has long been recognized as a key estimate for the Patlak–Keller–Segel system (see [8], for example), we first show that the supremum is attained in $I\left(\frac{\lambda}{2\sigma\,\delta}\,F_\eta\right)$. Then, by choosing η small enough, we can ensure that the maximum is only attained for $\mu=\bar\rho$. The proof of Proposition 3.1 is concluded by inserting (28) into (27) to deduce (26) by optimizing in ε.

3.3. *Large smooth part away from zero*

As mentioned above, Proposition 3.1 does not control the long-range part in V, which we need to remove and deal with on its own. Hence denote this long-range part

$$W(x)=V(x)\,(1-\chi(|x|/\eta)).$$

Then define

$$\mathcal{K}_N^W(G_N^W\mid G_{\bar\rho_N}^W)=\frac{1}{N}\int_{\Pi^{dN}}\rho_N\log\frac{G_{\bar\rho_N}^W}{G_N^W}\,\mathrm{d}X^N,$$

with, similarly to the potential V,

$$G_N^W(X^N) = \exp\Big(-\frac{1}{2N\sigma}\sum_{i\neq j} W(x_i - x_j)\Big),$$

$$G_{\bar\rho}^W(x) = \exp\Big(-\frac{1}{\sigma}\, W\star\bar\rho(x) + \frac{1}{2\,\sigma}\int_{\Pi^d} W\star\bar\rho\,\bar\rho\Big),$$

$$G_{\bar\rho_N}^W(t, X^N) = \exp\Big(-\frac{1}{\sigma}\sum_{i=1}^N W\star\bar\rho(x_i) + \frac{N}{2\sigma}\int_{\Pi^d} W\star\bar\rho\,\bar\rho\Big).$$

Now Proposition 3.1 actually provides a lower bound on $\mathcal{K}_N - \mathcal{K}_N^W$. We then need to calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(E_N\Big(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\Big) - \mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W)\Big),$$

using Proposition 2.1 together with the following lemma.

Lemma 3.2. *For $W \in W^{2,\infty}$ even, one has that*

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W) &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{N}\int_{\Pi^{dN}} \rho_N \log\frac{G_{\bar\rho_N}^W}{G_N^W}\,\mathrm{d}X^N\\
&= \int_{\Pi^{Nd}} \rho_N \int_{\Pi^{2d}} \Delta W(x-y)\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}\,\mathrm{d}X^N\\
&\quad - \frac{1}{\sigma}\int_{\Pi^{Nd}} \rho_N \int_{\Pi^{2d}}\int_{\Pi^d} \nabla W(z-x)\cdot\nabla V(z-y)\,\bar\rho(z)\,\mathrm{d}z\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}\\
&\quad - \frac{1}{2\,\sigma}\int_{\Pi^{Nd}} \rho_N \int_{\Pi^{2d}} \nabla V(x-y)\,(\nabla W\star(\mu_N-\bar\rho)(x) - \nabla W\star(\mu_N-\bar\rho)(y))\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}\\
&\quad - \frac{1}{2\,\sigma}\int_{\Pi^{Nd}} \rho_N \int_{\Pi^{2d}} \nabla W(x-y)\,(\nabla V\star\bar\rho(x) - \nabla V\star\bar\rho(y))\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2}.
\end{aligned} \tag{30}$$

Since W is not directly connected to the dynamics, Lemma 3.2 introduces several more terms on W than Proposition 2.1 did on V. For example, the right-hand side now involves ΔW. However, thanks to Assumption (21), we know that $W \in W^{2,\infty}$ and therefore the right-hand side in Lemma 3.2 can be directly bounded by Theorem A.2 in a manner similar to that described in the first subsection. Consequently, we can deduce from Lemma 3.2 the following bound

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W) \le C\,\mathcal{H}_N(\rho_N\,|\,\bar\rho_N) + \frac{C}{N}. \tag{31}$$

3.4. Conclusion of the proof

We are now ready to conclude the proof of Theorem 2.2 and summarize the main steps:

- combine Proposition 3.1 and Lemma 3.2 to calculate

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(E_N\Big(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\Big) - \mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W)\Big);$$

- use Theorem A.2 to bound the right-hand side; Specifically, from (25) and (31), one obtains that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(E_N\Big(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\Big) - \mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W)\Big) \le C\,\mathcal{H}_N(\rho_N\,|\,\bar\rho_N) + \frac{C}{N}; \tag{32}$$

- use Proposition 3.1 to deduce that

$$E_N\Big(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\Big) - \mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W) \ge (1-\delta)\,\mathcal{H}_N(\rho_N\,|\,\bar\rho_N) - \frac{C}{N^\theta}; \tag{33}$$

  for some exponent $\theta > 0$ and a constant C that depends on various norms of $\bar\rho$;

– combining (32) and (33), we obtain that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)-\mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W)\right)\le C\left(E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)-\mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W)\right)+\frac{C}{N^\theta};$$

for some $\theta>0$. By Gronwall's lemma, this implies that

$$\left(E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)-\mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W)\right)\le \mathrm{e}^{C\,t}\left(E_N(t=0)-\mathcal{K}_N^W(t=0)+\frac{C}{N^\theta}\right); \tag{34}$$

– it only remains to re-arrange the various terms. We note first that by applying Theorem A.2, we have that

$$-\mathcal{K}_N^W(G_N^W\,|\,G_{\bar\rho_N}^W)\le C\,\mathcal{H}_N(\rho_N\,|\,\bar\rho_N)+\frac{C}{N}.$$

Combined with (33), this proves that

$$\mathcal{H}_N(\rho_N\,|\,\bar\rho_N)\le \mathrm{e}^{C\,t}\left(E_N(t=0)+C\,\mathcal{H}_N(t=0)+\frac{C}{N^\theta}\right),$$

and therefore, using the smoothness of W and $\overline{\rho}$, we get that

$$E_N\left(\frac{\rho_N}{G_N}\,|\,\frac{\bar\rho_N}{G_{\bar\rho_N}}\right)\le \mathrm{e}^{C\,t}\left(E_N(t=0)+C\,\mathcal{H}_N(t=0)+\frac{C}{N^\theta}\right),$$

finishing the proof.

Appendix A. Explicit basic large-deviation estimate

A.1. Some results from the paper [18]

We first recall the Jensen-type inequality, which was Lemma 1 in [18] for similar purposes.

Lemma A.1. *For any $\rho_N,\ \bar\rho_N\in\mathcal{P}(\Pi^{dN})$, any test function $\psi\in L^\infty(\Pi^{dN})$, one has that, for any $\alpha>0$,*

$$\int_{\Pi^{dN}}\psi(X^N)\,\mathrm{d}\rho_N\le\frac{1}{\alpha}\frac{1}{N}\int\mathrm{d}\rho_N\,\log\frac{\rho_N}{\bar\rho_N}+\frac{1}{\alpha}\frac{1}{N}\log\int_{\Pi^{dN}}\exp\big(\alpha N\,\psi(X^N)\big)\,\mathrm{d}\bar\rho_N.$$

Lemma A.1 directly connects bounds on quantities like (19) to the relative entropy $\mathcal{H}_N$ and estimates on quantities that can be seen as partition functions:

$$\int_{\Pi^{dN}}\exp\Big(N\int_{\Pi^{2d}}f(x,y)\,(\mathrm{d}\mu_N-\mathrm{d}\bar\rho)^{\otimes 2}\Big)\,\bar\rho^{\otimes N}\,\mathrm{d}X^N. \tag{A.1}$$

It is hence natural to try to use large-deviation tools to bound (A.1). Note, however, that our goals here are different from classical large-deviation approaches. We do not try to calculate the limit as $N\to\infty$ of (A.1) but instead to obtain bounds that are uniform in N. As a first example of such a result, we recall the estimate from [18], which reads as follows in the present context.

Theorem A.2. (Theorem 4 in [18]). *There exists a constant c_d depending only on d such that, for any $\bar\rho\in L^\infty(\Pi^d)\cap\mathcal{P}(\Pi^d)$ and for any $f\in L^\infty(\Pi^{2d})$ with $f(x,y)=f(y,x)$ and*

$$\eta:=c_d\left(\sup_{p\ge1}\frac{\|\sup_y|f(x,y)|\,\|_{L^p_x}}{p}\right)^2+c_d\,(\|\bar\rho\|_{L^\infty}\,\|f\|_{L^\infty})^2<1,$$

then

$$\sup_{N\ge2}\int_{\Pi^{dN}}\exp^{\left(N\int_{\Pi^{2d}}f(x,y)\,(\mathrm{d}\mu_N-\mathrm{d}\bar\rho)^{\otimes2}\right)}\,\bar\rho^{\otimes N}\,\mathrm{d}X^N\le\frac{2}{1-\eta}\le C<\infty.$$

Remark 5. Observe that $\int_{\Pi^{2d}} f(x,y)\,(\mathrm{d}\mu_N - \mathrm{d}\bar\rho)^{\otimes 2} = \frac{1}{N^2}\sum_{i,j}\psi(x_i,x_j)$ for

$$\psi(x,y) = f(x,y) - \int f(x,z)\,\bar\rho(z)\,\mathrm{d}z - \int f(z,y)\,\bar\rho(z)\,\mathrm{d}z + \int f(z,z')\,\bar\rho(z)\,\bar\rho(z')\,\mathrm{d}z\,\mathrm{d}z',$$

so that $\int \psi(x,y)\,\bar\rho(y)\,\mathrm{d}y = 0$ and of course by symmetry so does $\int \psi(x,y)\,\bar\rho(x)\,\mathrm{d}x = 0$, which were the two cancellations required in [18].

While the proof of Theorem A.2 in [18] was combinatorial, quite recently a probabilistic proof of this theorem was given in [19] and the method then applied to a chemical reaction–diffusion model.

A.2. Explicit large-deviation approach through regularization

The last part of the proof of Proposition 3.1 relies on a quantified version of the classical large-deviation approaches as developed for example in [2,26]. Given a possibly unbounded functional $F:\mathcal{P}(\Pi^d)\to\mathbb{R}\cup\{\pm\infty\}$, we recall that the large-deviation functional associated with F is:

$$I(F) = \max_{\mu\in\mathcal{P}(\Pi^d)} F(\mu) - \int_{\Pi^d} \mu\,\log\frac{\mu}{\bar\rho}\,\mathrm{d}x. \tag{A.2}$$

It is well known that, if F is continuous, then in the limit $N\to\infty$, at first order the partition function given by

$$\int_{\Pi^{dN}} \exp(N\,F(\mu_N))\,\bar\rho_N\,\mathrm{d}X^N$$

behaves like $\mathrm{e}^{N\,I(F)}$. We essentially have to quantify such estimates to make them uniform in N and applicable to possibly unbounded F.

The strategy followed in the proof of Proposition 3.1 consists in trying to replace $F(\mu)$ by $F(L_\varepsilon\star\mu_N)$ as it is then reasonably straightforward to obtain the following theorem.

Theorem A.3. *Assume that* $\log\bar\rho\in W^{1,\infty}$*, then there exists a constant C depending only on d, L, so that one has:*

$$\frac{1}{N}\log\int_{\Pi^{dN}} \exp(N\,F(L_\varepsilon\star\mu_N))\,\bar\rho_N\,\mathrm{d}X^N \le I(F)$$
$$+\frac{C}{N^{1/(d+1)}\,\varepsilon^{d/(d+1)}}\,(\log N + |\log\varepsilon| + \|\log\bar\rho\|_{L^\infty}) + C\,\varepsilon\,\|\log\bar\rho\|_{W^{1,\infty}}.$$

References

[1] D. Arsenio, L. Saint-Raymond, From the Vlasov Maxwell Boltzmann System to Incompressible Viscous Electro–Magneto–Hydrodynamics, EMS Monographs in Mathematics, 2019.
[2] G. Ben Arous, M. Brunaud, Méthode de Laplace : étude variationnelle des fluctuations de diffusions de type «champ moyen», Stoch. Stoch. Rep. 31 (1990) 79–144.
[3] A. Blanchet, J. Dolbeault, B. Perthame, Two-dimensional Keller–Segel model: optimal critical mass and qualitative properties of the solutions, Electron. J. Differ. Equ. 2006 (44) (2006).
[4] D. Bresch, P.-E. Jabin, Global existence of weak solutions for compressible Navier–Stokes equations: thermodynamically unstable pressure and anisotropic viscous stress tensor, Ann. of Math. (2) 188 (2018) 577–684.
[5] D. Bresch, P.-E. Jabin, Z. Wang, Mean field limit and quantitative estimates with a large class of singular kernels, in preparation.
[6] L. Caffarelli, L. Silvestre, An extension problem related to the fractional Laplacian, Commun. Partial Differ. Equ. 32 (8) (2007) 1245–1260.
[7] P. Cattiaux, L. Pédèches, The 2-D stochastic Keller–Segel particle model: existence and uniqueness, ALEA Lat. Am. J. Probab. Math. Stat. 13 (1) (2016) 447–463.
[8] J. Dolbeault, J. Campos Serrano, A functional framework for the Keller–Segel system: logarithmic Hardy–Littlewood–Sobolev and related spectral gap inequalities, C. R. Acad. Sci. Paris, Ser. I 350 (21–22) (2012) 949–954.
[9] J. Dolbeault, B. Perthame, Optimal critical mass in the two-dimensional Keller–Segel model in $\mathbb{R}^2$, C. R. Acad. Sci. Paris, Ser. I 339 (9) (2004) 611–616.
[10] M. Duerinckx, Mean-field limits for some Riesz interaction gradient flows, SIAM J. Math. Anal. 48 (3) (2016) 2269–2300.
[11] N. Fournier, M. Hauray, S. Mischler, Propagation of chaos for the 2d viscous vortex model, J. Eur. Math. Soc. 16 (2014) 1425–1466.
[12] N. Fournier, B. Jourdain, Stochastic particle approximation of the Keller–Segel equation and two-dimensional generalization of Bessel processes, Ann. Appl. Probab. 5 (2017) 2807–2861.
[13] D. Godinh, C. Quininao, Propagation of chaos for a sub-critical Keller–Segel model, Ann. Inst. Henri Poincaré Probab. Stat. 51 (2015) 965–992.
[14] D. Han-Kwan, Quasineutral limit of the Vlasov–Poisson system with massless electrons, Commun. Partial Differ. Equ. 1385 (1425) (2010).
[15] J. Haškovec, C. Schmeiser, Convergence of a stochastic particle approximation for measure solutions of the 2D Keller–Segel system, Commun. Partial Differ. Equ. 36 (2011) 940–960.
[16] M. Hauray, Wasserstein distances for vortices approximation of Euler-type equations, Math. Models Methods Appl. Sci. 19 (2009) 1357–1384.
[17] P.-E. Jabin, Z. Wang, Mean field limit and propagation of chaos for Vlasov systems with bounded forces, J. Funct. Anal. 271 (2016) 3588–3627.
[18] P.-E. Jabin, Z. Wang, Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels, Inventiones 214 (1) (2018) 523–591.

[19] T-S. Lim, Y. Lu, J. Nolen, Quantitative propagation of chaos in the bimolecular chemical reaction–diffusion model, arXiv:1906.01051, 2019.
[20] H. Osada, A stochastic differential equation arising from the vortex problem, Proc. Jpn. Acad., Ser. A, Math. Sci. 61 (1986) 333–336.
[21] H. Osada, Propagation of chaos for the two dimensional Navier–Stokes equation, in: Probabilistic Methods in Mathematical Physics, Katata Kyoto, 1985, Academic Press, Boston, MA, USA, 1987, pp. 303–334.
[22] M. Puel, L. Saint-Raymond, Quasineutral limit for the relativistic Vlasov–Maxwell system, Asymptot. Anal. 40 (3, 4) (2004) 303–352.
[23] L. Saint-Raymond, Des points vortex aux équations de Navier–Stokes [d'après P.-E. Jabin et Z. Wang], in: Séminaire Bourbaki, 70^{e} année, 2017–2018 (1143).
[24] S. Serfaty, Systems of points with Coulomb interactions, in: Proc. Int. Cong. of Math, Rio de Janeiro, vol. 1, 2018, pp. 935–978.
[25] S. Serfaty (Appendix with Mitia Duerinckx), Mean field limit for Coulomb-type flows, submitted for publication.
[26] S.R.S. Varadhan, Large deviations and applications, in: C.B.M.S., vol. 46, S.I.A.M., Philadelphia, PA, USA, 1984.