Stability for a GNS inequality and the Log-HLS inequality, with application to the critical mass Keller-Segel equation

Eric A. Carlen[1] and Alessio Figalli[2]

1. Department of Mathematics, Hill Center,
Rutgers University, 110 Frelinghuysen Road Piscataway NJ 08854-8019 USA

2. Department of Mathematics, The University of Texas at Austin,
1 University Station C1200, Austin TX 78712 USA

July 18, 2012

Abstract

Starting from the quantitative stability result of Bianchi and Egnell for the 2-Sobolev inequality, we deduce several different stability results for a Gagliardo-Nirenberg-Sobolev inequality in the plane. Then, exploiting the connection between this inequality and a fast diffusion equation, we get stability for the Log-HLS inequality. Finally, using all these estimates, we prove a quantitative convergence result for the critical mass Keller-Segel system.

1 Introduction

Let $W^{1,2}(\mathbb{R}^n)$ denote the space of measurable functions on $\mathbb{R}^n$ that have a square integrable distributional gradient. The Gagliardo-Nirenberg-Sobolev (GNS) inequality states that, for $n \geq 2$ and all $1 \leq p \leq q < r(n)$ (with $r(2) := \infty$, and $r(n) := 2n/(n-2)$ if $n \geq 3$), there is a finite constant C such that for all $u \in W^{1,2}(\mathbb{R}^n)$,

$$\|u\|_q \leq C\|u\|_p^{1-\theta}\|\nabla u\|_2^{\theta} \tag{1.1}$$

where

$$\frac{1}{q} = \frac{\theta}{r(n)} + \frac{1-\theta}{p} \ . \tag{1.2}$$

For $n \geq 3$ (so that $r(n) < \infty$), (1.1) is valid also for $q = r(n)$, in which case (1.2) gives $\theta = 1$ and (1.1) reduces to the Sobolev inequality

$$\|u\|_{2n/(n-2)}^2 \leq S_n\|\nabla u\|_2^2 \ , \tag{1.3}$$

for which the sharp constant S_n is known.

[1]Work partially supported by U.S. National Science Foundation grant DMS 0901632.
[2]Work partially supported by U.S. National Science Foundation grant DMS 0969962.

There are a few other choices of the p and q for which sharp constants are known. For $p = 1$ and $q = 2$, (1.2) gives $\theta = n/(n+2)$, and (1.1) reduces to the sharp Nash inequality [11]

$$\|u\|_2^2 \leq C_n \|\nabla u\|_2^{n/(n+2)} \|u\|_1^{2/(n+2)} . \tag{1.4}$$

(This inequality is valid also for $n = 1$, even though $r(1)$ is negative.)

More recently, the sharp constant has been found [15] for a one-parameter family of GNS inequalities for each $n \geq 2$: For $t > 0$, let $p = t + 1$, and let $q = 2t$. Then

$$\|u\|_{2t} \leq A_{n,t} \|\nabla u\|_2^{\theta} \|u\|_{t+1}^{1-\theta} , \qquad \theta = \frac{n(t-1)}{t[2n - (1+t)(n-2)]} . \tag{1.5}$$

(This inequality is a trivial identity for $t = 1$, and is valid even for $t < 1/2$, in which case $p < 1$ so that strictly speaking, for $0 < t < 1/2$, the sharp inequality is not included in (1.1).)

It turns out that there is a close relation between the sharp Sobolev inequality (1.3) and the family of GNS inequalities (1.5). One aspect of this is that the functions u that saturate these inequalities are simply powers of one another: The optimal constant S_n in (1.3) is given by [1, 26, 27]

$$S_n = \frac{\|v\|_{2n/(n-2)}^2}{\|\nabla v\|_2^2} \qquad \text{where} \qquad v(x) = (1 + |x|^2)^{-(n-2)/2} , \tag{1.6}$$

and moreover, with this value of S_n, there is equality in (1.3) if and only if u is a multiple of $v(\mu(x - x_0))$ for some $\mu > 0$ and some $x_0 \in \mathbb{R}^n$.

Likewise, for $t > 1$ the optimal constant $A_{n,t}$ in (1.5) is given by [15]

$$A_{n,t} = \frac{\|v\|_{2t}}{\|v\|_{t+1}^{1-\theta} \|\nabla v\|_2^{\theta}} \qquad \text{where} \qquad v(x) = (1 + |x|^2)^{-1/(t-1)} , \tag{1.7}$$

and moreover, with this value of $A_{n,t}$, there is equality in (1.3) if and only if u is a multiple of $v(\mu(x - x_0))$ for some $\mu > 0$ and some $x_0 \in \mathbb{R}^n$. However, this is a very particular feature of this family: the sharp Nash inequality has optimizers of an entirely different form; see [11].

Another aspect of this close relation between (1.3) and (1.5) is that both inequalities can be proved using ideas coming from the theory of optimal mass transportation [14]: more precisely, one should consider (1.3) and (1.5) as inequalities for a mass density $\rho(x) := |u(x)|^q$. (Throughout this paper, by a *density* we mean a non-negative integrable function.) Then it turns out for $r \geq 1 - 1/n$, that the functional

$$\rho \mapsto \frac{1}{r-1} \int_{\mathbb{R}^n} \rho^r(x) \mathrm{d}x$$

is convex along the *displacement interpolation* ρ_t, $0 \leq t \leq 1$, between two densities ρ_0 and ρ_1 of the same mass on $\mathbb{R}^n$ [24]. Taking $\rho_0(x) = v^q(x)$ with v as above, and taking $\rho_1(x) = u^q(x)$ where u is a non-negative function with $\|u\|_q = \|v\|_q$, the "above the tangent line inequality" for convex functions translates into (1.3) and (1.5), as shown in [14].

In this paper we are concerned with the *stability* properties of the GNS inequalities (1.5), and the applications of this stability to certain partial differential equations. In fact, although one could generalize many of our arguments to the whole family in (1.5), because of its connection with the Keller-Segel equation that we consider here, we shall focus only on the $n = 2$, $t = 3$ case.

This case may be written explicitly as

$$\pi \int_{\mathbb{R}^2} u^6(x)\mathrm{d}x \le \left(\int_{\mathbb{R}^2} |\nabla u(x)|^2 \mathrm{d}x\right)\left(\int_{\mathbb{R}^2} u^4(x)\mathrm{d}x\right) , \tag{1.8}$$

where u, here and throughout the rest of the paper, is a non-negative function on $\mathbb{R}^2$.

1.1 DEFINITION (GNS deficit functional)**.** Given a non-negative function u in $W^{1,2}(\mathbb{R}^2)$, define $\delta_{\mathrm{GNS}}[u]$ by

$$\delta_{\mathrm{GNS}}[u] := \left(\int_{\mathbb{R}^2} |\nabla u|^2 \mathrm{d}y\right)^{1/2}\left(\int_{\mathbb{R}^2} u^4 \mathrm{d}y\right)^{1/2} - \left(\pi \int_{\mathbb{R}^2} u^6 \mathrm{d}y\right)^{1/2} . \tag{1.9}$$

Also, for $\lambda > 0$ and $x_0 \in \mathbb{R}^2$, define

$$v_{\lambda,x_0} := (1 + \lambda^2|x - x_0|^2)^{-1/2} . \tag{1.10}$$

Also, throughout the paper, we use $v(x)$ to denote the function $v_{1,0}$; i.e.,

$$v(x) := (1 + |x|^2)^{-1/2} . \tag{1.11}$$

By [15, Theorem 1] of Del Pino and Dolbeault, $\delta_{\mathrm{GNS}}[u] > 0$ unless u is a multiple of v_{λ,x_0} for some $\lambda > 0$ and some $x_0 \in \mathbb{R}^2$. The question addressed in this paper is:

- *When $\delta_{\mathrm{GNS}}[u] = 0$ is small, in what sense must u be close to some multiple of v_{λ,x_0}?*

As indicated above, it is natural to think of the GNS inequality as an inequality concerning densities ρ, and hence it is natural to think of δ_{GNS} in this way too. However, associated to each u there are *two* natural densities to consider: $\rho(x) = u^6(x)$ and $\sigma(x) = u^4(x)$. Indeed, u^6 is the density with appears in the optimal transportation proof (and, for scaling reasons, it is the "natural" quantity to control using the deficit), while u^4 is the density which appears in the application to the Keller-Segel equations. Therefore, our notation refers to δ_{GNS} as a function of u.

Our first main result is:

1.2 THEOREM. *Let $u \in W^{1,2}(\mathbb{R}^2)$ be a non-negative function such that $\|u\|_6 = \|v\|_6$. Then there exist universal constants $K_1, \delta_1 > 0$ such that, whenever $\delta_{\mathrm{GNS}}[u] \le \delta_1$,*

$$\inf_{\lambda>0, x_0\in\mathbb{R}^2} \|u^6 - \lambda^2 v^6_{\lambda,x_0}\|_1 \le K_1 \delta_{\mathrm{GNS}}[u]^{1/2}. \tag{1.12}$$

1.3 Remark. Actually, since $\|\lambda^2 v^6_{\lambda,x_0}\|_1 = \|v^6\|_1$ and $\|u^6 - \lambda^2 v^6_{\lambda,x_0}\|_1 \le \|u^6\|_1 + \|v^6\|_1 = \pi$ for all $\lambda > 0$, (2.17) holds with $K_1 = \pi/\delta_1^{1/2}$ whenever $\delta_{\mathrm{GNS}}[u] \ge \delta_1$. So, up to enlarging K_1, (2.17) always holds without any restriction on $\delta_{\mathrm{GNS}}[u]$. Moreover, as can be be seen for instance from the argument in [23, 13, 18], the sign restriction on u is superfluous. However, for the applications we have in mind u will always be nonnegative and we are only interested in the regime when $\delta_{\mathrm{GNS}}[u]$ is small. Moreover, it is interesting to note that in Theorems 1.4 and 1.9, while the multiplicative constant depends on several parameters, the smallness of the deficit is universal. For these reasons, we have chosen to state our theorems in this simple form.

To obtain a similar result for the density $u^4(x)$, we need to require additional *a-priori* bounds ensuring some uniform integrability of the class of densities satisfying the bounds. For the PDE applications we have in mind, it is natural to use *moment bounds* and *entropy bounds*.

Define

$$N_p(u) = \int_{\mathbb{R}^2} |y|^p u^4(y)\mathrm{d}y \qquad \text{and} \qquad S(u) = \int_{\mathbb{R}^2} u^4 \log(u^4)\mathrm{d}y \ . \tag{1.13}$$

1.4 THEOREM. *Let $u \in W^{1,2}(\mathbb{R}^2)$ be a non-negative function such that $\|u\|_4 = \|v\|_4$. Suppose also that for some $A, B < \infty$ and some $1 < p < 2$,*

$$S[u] = \int_{\mathbb{R}^2} u^4 \log(u^4)\mathrm{d}x \le A < \infty \qquad \text{and} \qquad N_p[u] := \int_{\mathbb{R}^2} |y|^p u^4(y)\mathrm{d}y \le B < \infty\,, \tag{1.14}$$

and assume also that

$$\int_{\mathbb{R}^2} y u^4 \mathrm{d}y = 0 \ . \tag{1.15}$$

Then there are constants $K_2, \delta_2 > 0$, with δ_2 universal and K_2 depending only on p, A, and B, so that whenever $\delta_{\mathrm{GNS}}[u] \le \delta_2$,

$$\inf_{\lambda>0} \|u^4 - \lambda^2 v_\lambda^4\|_1 \le K_2 \delta_{\mathrm{GNS}}[u]^{(p-1)/(4p)} \ . \tag{1.16}$$

Moreover, there is a constant $a > 0$, depending only on A and B, such that the infimum in (1.16) is achieved at some $\lambda \in [a, 1/a]$.

To explain how to prove these results, let us first recall that a stability result for the sharp Sobolev inequality (1.3) has been proved some time ago by Bianchi and Egnell [4]. It states that there is a constant C_n, $n \ge 3$, so that for all $f \in W^{1,2}(\mathbb{R}^n)$,

$$C_n \left(\|\nabla f\|_2^2 - S_n \|f\|_{2n/(n-2)}^2 \right) \ge \inf_{c,\mu>0,\ x_0\in\mathbb{R}^n} \|\nabla f - c\nabla h_{\mu,x_0}\|_2^2 \tag{1.17}$$

where

$$h_{\mu,x_0}(x) := (1 + \mu^2|x - x_0|^2)^{-(n-2)/2} \ .$$

The proof uses a compactness argument so there is no information on the value of C_n. On the other hand, the metric used on the right hand side in (1.17) is as strong as one could hope for, and in this sense the result of Bianchi and Egnell is remarkably strong.

Unfortunately, the fact that typical GNS inequalities involve three norms and not two prevents any direct adaptation of the proof of Bianchi and Egnell to any of the other cases of the GNS inequality for which the optimizers are known. Moreover, other recent proofs for stability based on optimal transportation [17] or symmetrization techniques [19, 20, 13, 18] did not produce (at least up to now) any results in this situation.

However, it has recently been shown [3] that one may deduce the *sharp forms* of the GNS inequalities in (1.5) from the sharp Sobolev inequality (1.3). Of course, it is quite easy to deduce the GNS inequalities with a non-optimal constant from the Sobolev inequality and Hölder's inequality. The argument in [3], which we learned from Dominique Bakry, is more subtle: In particular, as we explain in the next section, one deduces the particular *two-dimensional* GNS inequality (1.8) from the *four-dimensional* Sobolev inequality.

This derivation of (1.8) provides the beginnings of a bridge between the Bianchi-Egnell stability result for the Sobolev inequality and our theorems on stability for (1.8). Building and crossing the bridge still requires further work, and this is carried out in Section 2 of the paper where we prove Theorems 1.2 and 1.4.

The third section of the paper concerns two evolution equations and three functionals, all with close connection to the GNS inequality (1.8). The two equations, both describing the evolution of mass densities on $\mathbb{R}^2$, are:

(1) A two dimensional fast diffusion equation:

$$\frac{\partial \sigma}{\partial t}(t,x) = \Delta\sqrt{\sigma(t,x)} + 2\sqrt{\frac{\pi}{\kappa M}}\,\mathrm{div}(x\,\sigma(t,x))\ . \tag{1.18}$$

Here κ and M are positive parameters that set the scale and mass of stationary solutions, as we shall explain. (It will be convenient to keep them separate).

(2) The *Keller-Segel equation*:

$$\frac{\partial \rho}{\partial t}(t,x) = \mathrm{div}\big[\nabla\rho(t,x) - \rho(t,x)\nabla c(t,x)\big]\ , \tag{1.19}$$

where

$$c(t,x) = -\frac{1}{2\,\pi}\int_{\mathbb{R}^2}\log|x-y|\rho(t,y)\mathrm{d}y\,.$$

The fast diffusion equation (1.18) has the steady state solutions

$$\sigma_{\kappa,M}(x) := \frac{M}{\pi}\frac{\kappa}{(\kappa+|x|^2)^2}\ . \tag{1.20}$$

Note that $\int_{\mathbb{R}^2}\sigma_{\kappa,M}(x)\mathrm{d}x = M$ for all κ.

The densities in (1.20) with $M = 8\pi$ are also the steady states of the Keller-Segel system (1.19), and $M = 8\pi$ is the *critical mass* for (1.19): If the initial data has a mass less than 8π, diffusion dominates and the solution diffuses away to infinity; if the initial data has a mass greater than 8π, the restoring drift dominates and the solution collapses in finite time [16].

We now remark that each $\sigma_{\kappa,M}$ is the fourth power of a GNS optimizer; equivalently, they are multiples of the densities v_λ^4 that figure in Theorem 1.4. This is the first indication of a close connection of these two equations to one another and to the GNS inequality (1.8).

To go further, we note that both of these equations are gradient flow for the 2-Wasserstein metric W_2 in the sense of Otto [25]. (For this fact, and further background on the Wasserstein metric, gradient flow, and these equations, see [6].)

The fast diffusion equation is gradient flow for the functional $\mathcal{H}_{\kappa,M}$, where:

1.5 DEFINITION (Fast diffusion entropy)**.**

$$\mathcal{H}_{\kappa,M}[\sigma] := \int_{\mathbb{R}^2}\frac{|\sqrt{\sigma(y)}-\sqrt{\sigma_{\kappa,M}(y)}|^2}{\sqrt{\sigma_{\kappa,M}(y)}}\mathrm{d}y \tag{1.21}$$

It is evident that $\mathcal{H}_{\kappa,M}[\sigma]$ is uniquely minimized at $\sigma = \sigma_{\kappa,M}$, and it is very easy to deduce an L^1 stability result for this functional; see [6].

On the other hand, the Keller-Segel system is gradient flow for the following "free energy" functional:

$$\mathcal{F}_{\mathrm{KS}}[\rho] = \int_{\mathbb{R}^2} \rho \log \rho(x) \mathrm{d}x + \frac{1}{4\pi} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} \rho(x) \log |x-y| \rho(y) \mathrm{d}x \mathrm{d}y \ . \tag{1.22}$$

We are concerned with the critical mass case $M = 8\pi$, in which case this coincides with the logarithmic Hardy-Littlewood-Sobolev (Log-HLS) functional:

1.6 DEFINITION (Log-HLS Functional)**.** The Log-HLS functional $\mathcal{F}$ is defined by

$$\mathcal{F}[\rho] := \int_{\mathbb{R}^2} \rho \log \rho(x) \mathrm{d}x + 2 \left(\int_{\mathbb{R}^2} \rho(x) \mathrm{d}x \right)^{-1} \iint_{\mathbb{R}^2 \times \mathbb{R}^2} \rho(x) \log |x-y| \rho(y) \mathrm{d}x \mathrm{d}y$$

on the domain consisting of densities ρ on $\mathbb{R}^2$ such that both $\rho \ln \rho$ and $\rho \ln(e + |x|^2)$ belong to $L^1(\mathbb{R}^2)$ (we define $\mathcal{F}[\rho] := +\infty$ otherwise).

The logarithmic HLS functional $\mathcal{F}$ is invariant under scale changes: for $a > 0$ and ρ in the domain of $\mathcal{F}$, $\mathcal{F}[\rho] = \mathcal{F}[\rho_{(a)}]$ for all $a > 0$, where $\rho_{(a)} := a^2 \rho(ax)$. In particular, $\mathcal{F}[\sigma_{\kappa,M}]$ is independent of κ. One computes [7, 10]

$$\mathcal{F}[\sigma_{\kappa,M}] := C(M) = M(1 + \log \pi - \log(M)) \ . \tag{1.23}$$

The sharp Log-HLS inequality [7, 10] states that $\mathcal{F}[\rho] \geq C(M)$ for all densities of mass $M > 0$. Moreover, there is equality if and only if $\rho(x) = \sigma_{\kappa,M}(x - x_0)$ for some $\kappa > 0$ and some $x_0 \in \mathbb{R}^2$. Thus, among densities of fixed mass M, the $\sigma_{\kappa,M}$ are the unique minimizers of $\mathcal{F}$. However, in contrast with the fast diffusion entropy $\mathcal{H}_{\kappa,M}$, is not so simple to deduce an L^1 stability result for the Log-HLS inequality (i.e., for the minimization problem associated to $\mathcal{F}$ at fixed mass). One of the main results proved in Section 3 is a stability result for this inequality; see Theorem 1.9 below.

The fact that the fast diffusion equation (1.18) is a gradient flow for $\mathcal{H}_{\kappa,M}$ implies that $\mathcal{H}_{\kappa,M}[\sigma(t,\cdot)]$ is monotone decreasing along solutions of (1.18) with initial data for which $\mathcal{H}_{\kappa,M}[\sigma(0,\cdot)]$ is finite. Likewise, the fact that the Keller-Segel equation is a gradient flow for the functional $\mathcal{F}_{\mathrm{KS}}$ implies that the Log-HLS functional $\mathcal{F}[\rho(t,\cdot)]$ is decreasing along solutions of (1.19) for initial data with the critical mass $M = 8\pi$ such that $\mathcal{F}[\rho(0,\cdot)]$ is finite.

There is, nonetheless, a fundamental difference: *The functional $\mathcal{H}_{\kappa,M}$ is uniformly displacement convex [6]*, and as shown by Otto [25], evolution equations that are W_2-gradient flows of uniformly displacement convex functionals have an exponential rate of convergence to equilibrium; i.e., the minimizers of the functional. This yields an exponential rate of convergence to equilibrium for the fast diffusion equation.

However, the Log-HLS functional *is not* displacement convex (nor is it even convex in the usual sense), and hence the gradient flow structure by itself does not provide any sort of rate of convergence for this equation. We shall show that our quantitative stability estimates for the GNS inequality lead to a stability result for the Log-HLS inequality, and combining these results we get a quantitative rate of convergence estimate for the Keller-Segel equation; see Theorem 3.5.

A key to this is a surprising interplay between $\mathcal{H}_{\kappa,M}$ and $\mathcal{F}$ along our two evolutions. As noted above, by their nature as gradient flow evolutions, it is naturally true that $\mathcal{F}[\rho(t,\cdot)]$ decreases along critical mass solutions of the Keller-Segel equation, and it is naturally true that $\mathcal{H}_{\kappa,M}[\rho(t,\cdot)]$ decreases along solutions of the fast diffusion equation (1.18).

More surprisingly, it has recently been shown [9, 6] that, in fact, $\mathcal{F}[\sigma(t,\cdot)]$ is also decreasing along solutions of (1.18), and that $\mathcal{H}_{\kappa,8\pi}[\rho(t,\cdot)]$ is also decreasing along critical mass solutions of the Keller-Segel equation (1.19).

In fact, as shown in [9], for any solution $\sigma(t,x)$ of (1.18),

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}[\sigma(t,\cdot)] = -\frac{8\pi}{M}\mathcal{D}[\sigma(t,\cdot)]\ , \tag{1.24}$$

where $\mathcal{D}$ denotes the dissipation functional defined as follows:

1.7 DEFINITION (Dissipation functional)**.** For any density σ on $\mathbb{R}^2$, let $u := \sigma^{1/4}$. If u has a square integrable distributional gradient, define

$$\mathcal{D}[\sigma] := \frac{1}{\pi}\left(\|\nabla u\|_2^2\|u\|_4^4 - \pi\|u\|_6^6\right)\ . \tag{1.25}$$

Otherwise, define $\mathcal{D}[\sigma]$ to be infinite.

Note that $\mathcal{D}[\sigma] \geq 0$ as a consequence (actually, a restatement) of the sharp GNS inequality (1.8).

Since uniqueness of solutions to the critical mass Keller-Segel equation is not known, what is actually proved in this case is somewhat less: in [6], a natural class of solutions called "properly dissipative solutions" is constructed, along which

$$\mathcal{H}_{\kappa,8\pi}[\rho(T,\cdot)] + \int_0^T \mathcal{D}[\rho(t,\cdot)]\mathrm{d}t \leq \mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)] \tag{1.26}$$

for all $T > 0$. (This is evidently an analog of (1.24) in integrated form.)

Our goal here is to understand the asymptotic behavior of a properly dissipative solution $\rho(t)$ of (1.19) starting from some ρ such that $\mathcal{H}_{\kappa,8\pi}[\rho] < \infty$. Without loss of generality we can assume that $\int_{\mathbb{R}^2} x\rho(x)\mathrm{d}x = 0$, a condition which is preserved along the flow.

The first observation is that, as an immediate consequence of (1.26), for any $T > 1$

$$\inf_{t\in[1,T]}\mathcal{D}[\rho(t,\cdot)] \leq \frac{1}{T-1}\int_1^T \mathcal{D}[\rho(t,\cdot)]\mathrm{d}t \leq \frac{1}{T-1}\mathcal{H}_{\kappa,8\pi}[\rho]\ . \tag{1.27}$$

(As we will see in Section 3, the reason for considering $t \geq 1$ is to ensure that some time passes so that the solution enjoys some further regularity properties needed to apply our estimates.)

Now, observe that for any density σ on $\mathbb{R}^2$ such that $\|\nabla\sigma^{1/4}\|_2 < \infty$,

$$\begin{aligned}\mathcal{D}[\sigma] &= \left(\|\nabla\sigma^{1/4}\|_2\|\sigma^{1/4}\|_4^2 + \sqrt{\pi}\|\sigma^{1/4}\|_6^3\right)\delta_{\mathrm{GNS}}(\sigma^{1/4})\\ &= \left(\|\nabla\sigma^{1/4}\|_2\|\sigma\|_1^{1/2} + \sqrt{\pi}\|\sigma\|_{3/2}^2\right)\delta_{\mathrm{GNS}}(\sigma^{1/4})\ .\end{aligned} \tag{1.28}$$

Hence, granted (for now) an *a-priori* bound on $\int_{\mathbb{R}^2}|x|^p\rho(t,x)\mathrm{d}x$ for some $1 < p < 2$ for $t \geq 1$, we have a lower bound on $\|\rho(t,\cdot)\|_{3/2}$ depending only on the pth moment bound. From this and (1.27) we deduce that, for any $T \geq 2$, there exists *some* $\bar{t} \in [1,T]$ such that

$$\delta_{\mathrm{GNS}}[\rho^{1/4}(\bar{t},\cdot)] \leq \frac{C}{T}\mathcal{H}_{\kappa,8\pi}[\rho]\ ,$$

where C is universal (as it depends only on the pth moment bound).

Then, granted also an a-priori upper bound on the entropy $\int_{\mathbb{R}^2} \rho \log \rho(t,x)\mathrm{d}x$ for $t \geq 1$, applying Theorem 1.4 we conclude that for *some* $\mu > 0$,

$$\|\rho(\bar{t},\cdot) - \sigma_{\mu,8\pi}\|_1 \leq C\left(\frac{1}{T}\mathcal{H}_{\kappa,8\pi}[\rho]\right)^{(p-1)/4p} , \tag{1.29}$$

(recall that the density v_λ^4 is a multiple of some $\sigma_{\mu,8\pi}$).

The inequality (1.29) bounds the time it takes a solution of the critical mass Keller-Segel equation to approach $\sigma_{\mu,8\pi}$ for *some* μ. However, to get a quantitative convergence result, we must do two more things: First, show that $\rho(t,\cdot)$ approaches $\sigma_{\mu,8\pi}$ for $\mu = \kappa$, and then show that eventually it *remains* close.

The first point is relatively easy, since $\mathcal{H}_{\kappa,8\pi}[\sigma_{\mu,8\pi}] = \infty$ for $\mu \neq \kappa$ (because of the sensitivity of $\mathcal{H}_{\kappa,M}[\rho]$ to the tail of ρ; see [6]).

The second requires more work: The strategy used in [6] was to show that eventually $\mathcal{F}[\rho(t,\cdot)]$ becomes small. Since this quantity is monotone, once small, it stays small. Then one uses a stability inequality for the Log-HLS inequality to conclude that $\|\rho(t,\cdot) - \sigma_{\kappa,8\pi}\|_1$ stays small. The stability inequality for the Log-HLS inequality used in [6] relied on a compactness argument, and thus gave L^1 convergence to the steady state, but without any rate estimate. Moreover, the argument in [6] also used compactness arguments to deduce that $\|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1$ eventually becomes small for some μ, so that there was no quantitative estimate on the time to first approach the set of densities $\{\sigma_{\kappa,8\pi} \ \ \kappa > 0\}$.

To provide a convergence result with quantitative bounds we do the following: First we show almost Lipschitz regularity of $\mathcal{F}$ in L^1 (Theorem 3.7), and we combine it with (1.29) and the fact that p can be chosen close to 2, to deduce that

$$\mathcal{F}[\rho(\bar{t},\cdot)] - C(8\pi) \leq CT^{-(1-\epsilon)/8},$$

where $C(M)$ is defined in (1.23). Since $\bar{t} \leq T$ and $\mathcal{F}[\rho(t,\cdot)]$ is decreasing, we deduce that

$$\mathcal{F}[\rho(T,\cdot)] - C(8\pi) \leq CT^{-(1-\epsilon)/8} \tag{1.30}$$

for all $T \geq 2$.

This brings us to our final stability result:

1.8 DEFINITION (Log-HLS deficit)**.** For any density ρ on $\mathbb{R}^2$ with $\int_{\mathbb{R}^2} \rho(x)\mathrm{d}x = M$, we define $\delta_{\mathrm{HLS}}[\rho]$, the deficit in the Log-HLS inequality, as

$$\delta_{\mathrm{HLS}}[\rho] := \mathcal{F}[\rho] - M(1 + \log\pi - \log(M)) . \tag{1.31}$$

1.9 THEOREM (Stability for Log-HLS)**.** *Let* ρ *be a density of mass* M *on* $\mathbb{R}^2$ *such that* $\int_{\mathbb{R}^2} x\rho(x)\mathrm{d}x = 0$ *and, for some* $\kappa > 0$ *and all* $q \in [1,\infty)$*,*

$$\mathcal{H}_{\kappa,M}[\rho] =: B_{\mathcal{H}} < \infty , \qquad \mathcal{F}[\rho] =: B_{\mathcal{F}} < \infty , \qquad \|\rho\|_q := B_q < \infty \quad \text{and} \qquad 1 + \mathcal{D}[\rho] =: B_{\mathcal{D}} < \infty .$$

Then, there exists a universal constant $\delta_3 > 0$ *such that the following holds: for all* $\epsilon > 0$ *there exist* $q(\epsilon) \in [1,\infty)$*, and constant* C *depending only on* ϵ*,* M*,* κ*,* $B_{\mathcal{H}}$*,* $B_{\mathcal{F}}$ *and* $B_{q(\epsilon)}$*, but not on* $B_{\mathcal{D}}$*, such that*

$$\|\rho - \sigma_{\mu,M}\|_1 \leq C\left(1 + B_{\mathcal{D}}^{1/6}\delta_{\mathrm{HLS}}[\rho]^{(1-\epsilon)/20}\right)\delta_{\mathrm{HLS}}[\rho]^{(1-\epsilon)/20}$$

for some $\mu > 0$*, provided* $\delta_{\mathrm{HLS}}[\rho] \leq \delta_3$*.*

By (1.30), $\delta_{\rm HLS}[\rho(t,\cdot)]$ is decreasing to zero at a rate of (essentially) $t^{-1/8}$. Then, by Theorem 1.9, there exists some $\mu(t)$ such that

$$\|\rho(t,\cdot) - \sigma_{\mu(t),8\pi}\|_1$$

converges to zero at a rate of essentially $t^{-1/160}$ at all times for which $B_{\mathcal{D}}$ is bounded. Combining this with the one-sided Lipschitz estimate provided by Lemma 3.8, we obtain a convergence *for all* t at a rate of essentially $t^{-1/320}$. Finally, a simple argument using the sensitive dependence of $\mathcal{H}_{\kappa,8\pi}$ on tails allows us to show that $\mu(t)$ converges at a logarithmic rate to κ.

It is interesting that the approach to equilibrium described by these quantitative bounds takes place on two separate time scales: The solution approaches the one-parameter family of (centered) stationary states with at least a polynomial rate. Then, perhaps much more gradually, at only a logarithmic rate, the solution adjusts its spatial scale to finally converge to the unique stationary solution within its basis of attraction. It is reasonable to expect such behavior: The initial data may, for example, be exactly equal to $\sigma_{\kappa,8\pi}$ on the complement of a ball of very large radius R, and yet may “look much more like” $\sigma_{\mu,8\pi}$ on a ball of smaller radius for some $\mu \neq \kappa$. One can then expect the solution to first approach $\sigma_{\mu,8\pi}$, and then only slowly begin to feel its distant tails and make the necessary adjustments to the spatial scale.

The precise statement of our results on the rates of convergence for the critical mass Keller-Segel equation is given in Theorem 3.5 below.

We close this introduction by remarking that the key to the proof of Theorem 1.9 is (1.24), which, upon integration, yields an expression for the Log-HLS deficit that can be related to the GNS deficit studied in Section 2.

2 Stability results for GNS inequalities

In the forth-coming book [3] the authors present a very elegant argument to deduce the family of sharp Gagliardo-Nirenberg inequalities (1.5) as a simple corollary of the sharp Sobolev inequality (1.3). The argument has been known for some time in certain circles, and is referred to as a result of D. Bakry in the third part of the remark following [14, Theorem 4]. We are grateful to D. Bakry for communicating this proof to us, and for providing us with a draft of the relevant chapter of [3].

Here, starting from this proof and combining it with the quantitative stability result (1.17) of Bianchi and Egnell, we deduce several stability results for the Gagliardo-Nirenberg-Sobolev inequality (1.8) of that family.

Although much of the argument below could be carried out for this whole family (modulo being able to extend the argument of Bianchi-Egnell to a slightly more general situation), we prefer to consider only the one particular GNS inequality which is important for the applications we consider here. In this way we also avoid the risk of making the paper excessively involved and hiding the main ideas.

2.1 From Sobolev to GNS

We begin by explaining the argument of [3] specialized to our particular case of interest.

The four-dimensional version of the sharp Sobolev inequality (1.3) has the explicit form

$$\|f\|_4^2 \leq \frac{1}{4\pi}\sqrt{\frac{3}{2}}\|\nabla f\|_2^2 \ , \tag{2.1}$$

and equality holds if $f = g$, where

$$g(x,y) := \frac{1}{1+|y|^2+|x|^2} \qquad x, y \in \mathbb{R}^2 \ . \tag{2.2}$$

The key observation, which is at the core of the proof of the next result, is that g can be written as

$$g(x,y) = \frac{1}{G(y)+|x|^2} \qquad \text{with} \quad G(y) := v^{-2}(y) = 1+|y|^2.$$

The following result, which is a particular case of the results in [3, Chapter 7], relates (1.8) and (2.1).

2.1 PROPOSITION. *Let $u \in W^{1,2}(\mathbb{R}^2)$ be a non-negative function satisfying*

$$\|u\|_6 = \|v\|_6 = \frac{\pi}{2} \ , \qquad \sqrt{2}\|\nabla u\|_2 = \|u\|_4^2 \ , \tag{2.3}$$

and define $f : \mathbb{R}^4 \to \mathbb{R}$ as

$$f(x,y) := \frac{1}{F(y)+|x|^2}, \qquad F(y) := u^{-2}(y), \qquad x, y \in \mathbb{R}^2.$$

Then

$$\delta_{\mathrm{GNS}}[u] = \left(\int_{\mathbb{R}^2} |\nabla u|^2 \mathrm{d}y\right)^{1/2} \left(\int_{\mathbb{R}^2} u^4 \mathrm{d}y\right)^{1/2} - \left(\pi\int_{\mathbb{R}^2} u^6 \mathrm{d}y\right)^{1/2} = \sqrt{3}\left(\frac{1}{4\pi}\sqrt{\frac{3}{2}}\|\nabla f\|_2^2 - \|f\|_4^2\right) . \tag{2.4}$$

2.2 Remark. Observe that, given $u \in W^{1,2}(\mathbb{R}^2)$ with $u \not\equiv 0$, we can always multiply it by a constant so that $\|u\|_6 = \|v\|_6$, and then scale it as $\mu^{1/3}u(\mu y)$ choosing μ to ensure that $\sqrt{2}\|\nabla u\|_2 = \|u\|_4^2$. Since (1.8) is invariant under this scaling, this proves (1.8). This is the use of the identity (2.4) made in [3]. Our interest in this proposition is that it relates the GNS deficit to the Sobolev deficit.

Proof. We compute

$$\begin{aligned}\|\nabla f\|_2^2 &= \int_{\mathbb{R}^2}\left(\int_{\mathbb{R}^2}\frac{|\nabla F(y)|^2}{(F(y)+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y + \int_{\mathbb{R}^2}\left(\int_{\mathbb{R}^2}\frac{4|x|^2}{(F(y)+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y \\ &= \frac{\pi}{3}\int_{\mathbb{R}^2}|\nabla F(y)|^2F^{-3}(y)\mathrm{d}y + \frac{2\pi}{3}\int_{\mathbb{R}^2}F^{-2}(y)\mathrm{d}y\end{aligned}$$

and

$$\|f\|_4^2 = \left(\frac{\pi}{3}\int_{\mathbb{R}^2}F^{-3}(y)\mathrm{d}y\right)^{1/2} \ . \tag{2.5}$$

Thus

$$0 \leq \frac{1}{4\pi}\sqrt{\frac{3}{2}}\|\nabla f\|_2^2 - \|f\|_4^2 = \frac{1}{2\sqrt{6}}\left(2\int_{\mathbb{R}^2}|\nabla u|^2\mathrm{d}y + \int_{\mathbb{R}^2}u^4\mathrm{d}y\right) - \left(\frac{\pi}{3}\int_{\mathbb{R}^2}u^6\mathrm{d}y\right)^{1/2} \ ,$$

or equivalently (using the identity $2\sqrt{AB} = A + B - (\sqrt{A} - \sqrt{B})^2$)

$$\left(\int_{\mathbb{R}^2} |\nabla u|^2 \mathrm{d}y\right)^{1/2} \left(\int_{\mathbb{R}^2} u^4 \mathrm{d}y\right)^{1/2} - \left(\pi \int_{\mathbb{R}^2} u^6 \mathrm{d}y\right)^{1/2}$$
$$= \sqrt{3}\left(\frac{1}{4\pi}\sqrt{\frac{3}{2}}\|\nabla f\|_2^2 - \|f\|_4^2\right) - \frac{1}{2\sqrt{2}}\left(\sqrt{2}\|\nabla u\|_2 - \|u\|_4^2\right)^2 .$$

Recalling that $\sqrt{2}\|\nabla u\|_2 = \|u\|_4^2$ by assumption, and recalling the definition (1.9) of the GNS deficit, the proof is complete. □

2.2 Controlling the infimum in the Bianchi-Egnell Theorem.

The family of functions

$$g_{c,\mu,x_0,y_0}(x,y) := \frac{c\mu}{1 + \mu^2|x + x_0|^2 + \mu^2|y + y_0|^2}, \qquad c \in \mathbb{R},\ \mu > 0,\ x_0, y_0 \in \mathbb{R}^2.$$

consists of all of the optimizers of the Sobolev inequality (2.1). Observe that, with this definition, $g = g_{1,1,0,0}$, where g is the function defined in (2.2).

The Bianchi-Egnell stability result [4] combined with the Sobolev inequality (2.1) asserts the existence of a universal constant C_0 such that

$$C_0\sqrt{3}\left(\frac{1}{4\pi}\sqrt{\frac{3}{2}}\|\nabla f\|_2^2 - \|f\|_4^2\right) \geq \inf_{c,\mu,x_0,y_0} \|f - g_{c,\mu,x_0,y_0}\|_4^2 . \tag{2.6}$$

Hence, whenever u satisfies the conditions (2.3) of Proposition 2.1,

$$C_0\delta_{\mathrm{GNS}}[u] \geq \inf_{c,\mu,x_0,y_0} \|f - g_{c,\mu,x_0,y_0}\|_4^2 . \tag{2.7}$$

Let us observe that the renormalization $\|u\|_6 = \|v\|_6$ is equivalent to $\|f\|_4 = \|g\|_4$.

Our main goal in this subsection is to show first that, up to enlarging the constant C_0, we can assume that $c = \mu = 1$ and $x_0 = 0$ (see Lemma 2.3 below). This paves the way for the estimation of the infimum on the right hand side of (2.7) in terms of u and v.

2.3 LEMMA. *Let f be given by $f(x,y) = 1/(F(y) + |x|^2)$, with $F : \mathbb{R}^2 \to \mathbb{R}$ non-negative, and g be given by (2.2). Suppose that $\|f\|_4 = \|g\|_4$. Then there is a universal constant C_1 so that, for all real numbers $\delta > 0$ with*

$$\delta^{1/2} \leq \frac{1}{2400} , \tag{2.8}$$

whenever

$$\|f - g_{c,\mu,x_0,y_0}\|_4 \leq \delta^{1/2} \qquad \text{for some} \qquad c, \mu, x_0, y_0 ,$$

then

$$\|f - g_{1,1,0,y_0}\|_4 \leq C_1\delta^{1/2} .$$

As can be seen from the proof, a possible choice for C_1 is 4800.

Proof. Suppose that $\|f - g_{c,\mu,x_0,y_0}\|_4 < \delta^{1/2}$ for some $\delta^{1/2} > 0$ satisfying (2.8).

- *Step 1: we can assume* $c = 1$. First of all notice that $c \geq 0$, as otherwise

$$\delta^2 \geq \int_{\mathbb{R}^4} |f - g_{c,\mu,x_0,y_0}|^4 \mathrm{d}x\mathrm{d}y \geq \int_{\mathbb{R}^4} \left(|f|^4 + |g_{c,\mu,x_0,y_0}|^4\right) \mathrm{d}x\mathrm{d}y \geq \|f\|_4^4 = \|g\|_4^4 = \frac{\pi^2}{6} \,,$$

which is in contradiction with (2.8).

Now, for any $c, \mu > 0$ and $x_0, y_0 \in \mathbb{R}^2$, $\|g_{c,\mu,x_0,y_0}\|_4 = c\|g\|_4 = c\|f\|_4$. Hence,

$$|c - 1|\|g\|_4 = |\|g_{c,\mu,x_0,y_0}\|_4 - \|f\|_4| \leq \|f - g_{c,\mu,x_0,y_0}\|_4 < \delta^{1/2} \,,$$

and by the triangle inequality we get

$$\begin{aligned} \|f - g_{1,\mu,x_0,y_0}\|_4 &\leq \|f - g_{c,\mu,x_0,y_0}\|_4 + \|g_{1,\mu,x_0,y_0} - g_{c,\mu,x_0,y_0}\|_4 \\ &= \|f - g_{c,\mu,x_0,y_0}\|_4 + |c-1|\|g\|_4 \\ &\leq 2\,\delta^{1/2} \,. \end{aligned} \tag{2.9}$$

Thus, up to enlarging the constant, we may replace c by 1.

- *Step 2: we can assume* $x_0 = 0$. Observe that, by construction, f is even in x. Therefore (2.9) implies

$$\|f - g_{1,\mu,x_0,y_0}\|_4 = \|f - g_{1,\mu,-x_0,y_0}\|_4 \leq 2\,\delta^{1/2} \,,$$

and by the triangle inequality,

$$\|g_{1,\mu,2x_0,y_0} - g_{1,\mu,0,y_0}\|_4 = \|g_{1,\mu,x_0,y_0} - g_{1,\mu,-x_0,y_0}\|_4 \leq 4\,\delta^{1/2} \,.$$

However, a simple argument using the unimodality and symmetry properties of $g = g_{1,1,0,0}$ shows that

$$a \mapsto \|g_{1,\mu,ax_0,y_0} - g_{1,\mu,0,y_0}\|_4$$

is increasing in $a > 0$, thus

$$\|g_{1,\mu,x_0,y_0} - g_{1,\mu,0,y_0}\|_4 \leq \|g_{1,\mu,2x_0,y_0} - g_{1,\mu,0,y_0}\|_4 \leq 4\,\delta^{1/2}.$$

One more use of the triangle inequality gives

$$\|f - g_{1,\mu,0,y_0}\|_4 \leq 6\,\delta^{1/2} \,. \tag{2.10}$$

Hence, up to further enlarging the constant, we may replace x_0 by 0.

- *Step 3: we can assume* $\mu = 1$. Making a change of scale, we can rewrite (2.10) as

$$\left\| \frac{1}{\mu} \frac{1}{F(y/\mu) + |x|^2/\mu^2} - \frac{1}{1 + |y - \mu y_0|^2 + |x|^2} \right\|_4 \leq 6\,\delta^{1/2} \tag{2.11}$$

Let $A := \{(x,y) \in \mathbb{R}^4 \ : \ |x| \leq 1, \ |y - \mu y_0| \leq 1\}$. Note that the Lebesgue measure of A is π^2. Moreover, by a simple Fubini argument, for any set $B \subset A$ with measure greater than $(15/16)\pi^2$ there exists $\bar{y} \in \{y \ : \ |y - \mu y_0| \leq 1\}$ such that the set $B \cap \left(\mathbb{R}^2 \times \{\bar{y}\}\right)$ must intersect both

$$A \cap \{(x,y) \ : \ |x| < 1/4\} \quad \text{and} \quad A \cap \{(x,y) \ : \ |x| > 3/4\} \,.$$

(Indeed, if this was not the case, by Fubini Theorem the measure of B would be smaller than $(15/16)\pi^2$.)

Now, applying Chebyshev's inequality, by (2.11) we get the existence of a set $B \subset A$ of measure at least $(31/32)\pi^2$ such that

$$\left|\frac{1}{\mu}\frac{1}{F(y/\mu)+|x|^2/\mu^2} - \frac{1}{1+|y-\mu y_0|^2+|x|^2}\right| \leq 12\,\delta^{1/2} \qquad \forall\,(x,y)\in B \tag{2.12}$$

(as the complement of the above set has measure less or equal than $1/16$, which is less than $\pi^2/32$).

Set $\alpha := 1+|y-\mu y_0|^2+|x|^2$ and $\beta := \mu\big(F(y/\mu)+|x|^2/\mu^2\big)$, so that (2.12) becomes

$$\left|\frac{1}{\beta}-\frac{1}{\alpha}\right| \leq 12\,\delta^{1/2} \qquad \forall\,(x,y)\in B. \tag{2.13}$$

We observe that $\alpha \leq 3$ on B. Moreover, thanks to (2.8),

$$\frac{1}{\beta} \geq \frac{1}{\alpha} - \left|\frac{1}{\beta}-\frac{1}{\alpha}\right| \geq \frac{1}{3} - 12\,\delta^{1/2} \geq \frac{1}{4} \qquad \text{inside } B\ ,$$

that is $\beta \leq 4$ on B. Hence (2.13) gives

$$|\alpha-\beta| \leq 12\alpha\beta\,\delta^{1/2} \leq 144\,\delta^{1/2} \qquad \text{inside } B\ ,$$

or equivalently

$$\big|1+|y-\mu y_0|^2+|x|^2-\mu F(y/\mu)-|x|^2/\mu\big| \leq 144\,\delta^{1/2} \qquad \forall\,(x,y)\in B.$$

By the observation above, we have chosen B large enough that there exists $\bar y \in \{y\ : |y-\mu y_0| \leq 1\}$ such that $(x_1,\bar y),(x_2,\bar y)\in B$, with $|x_1|\in[0,1/4]$ and $|x_2|\in[3/4,1]$. Then the above estimate gives

$$\begin{aligned}
\frac{1}{2}\left|1-\frac{1}{\mu}\right| &\leq \big(|x_2|^2-|x_1|^2\big)\left|1-\frac{1}{\mu}\right| \\
&\leq \big|1+|\bar y-\mu y_0|^2+|x_1|^2-\mu F(\bar y/\mu)-|x_1|^2/\mu\big| \\
&\quad + \big|1+|\bar y-\mu y_0|^2+|x_2|^2-\mu F(\bar y/\mu)-|x_2|^2/\mu\big| \\
&\leq 288\,\delta^{1/2} \leq 300\,\delta^{1/2}\,.
\end{aligned}$$

Using (2.8) and the identity $(\mu-1)(1-(1-1/\mu)) = (1-1/\mu)$, we easily deduce

$$|\mu-1| \leq 1200\delta^{1/2}\ . \tag{2.14}$$

Since (as it is easy to check by a direct computation)

$$|\partial_\mu g_{1,\mu,0,y_0}(x,y)| \leq \frac{|g_{1,\mu,0,y_0}(x,y)|}{\mu} \leq 2|g_{1,\mu,0,y_0}(x,y)| \qquad \forall\,\mu\in[1/2,2]$$

we get

$$\|g_{1,\mu,0,y_0}-g_{1,1,0,y_0}\|_4 \leq 2|\mu-1|\|g\|_4 \qquad \forall\,\mu\in[1/2,2].$$

Combining this with (2.10) and (2.14), we finally obtain

$$\|f-g_{1,1,0,y_0}\|_4 \leq (6+2400\|g\|_4)\,\delta^{1/2} \leq 4800\,\delta^{1/2}\ , \tag{2.15}$$

concluding the proof. □

2.3 Bounding $\|u^6 - v^6\|_1$ in terms of $\|f - g\|_4$.

Our goal in this subsection is to bound $\|u^6 - v^6(\cdot - y_0)\|_1$ from above in terms of $\|f - g_{1,1,0,y_0}\|_4$.

2.4 LEMMA. *Let $u \in W^{1,2}(\mathbb{R}^2)$ be a non-negative function satisfying* (2.3)*, and let f be defined as in Proposition 2.1. Suppose that $\|f - g_{1,1,0,y_0}\|_4 \le 1$. Then*

$$\|u^6 - v^6(\cdot - y_0)\|_1 \le C_2 \|f - g_{1,1,0,y_0}\|_4$$

for some universal constant C_2.

As can be seen for the proof, a possible choice for C_2 is 1000. We also remark that, by considering u of the form $v + \varepsilon\phi$ with $\varepsilon > 0$ small, one sees that the unit in the above estimate is optimal.

Proof. Up to replace u and f by $u(\cdot - y_0)$ and $f(\cdot - y_0)$ respectively, we can assume that $y_0 = 0$.

We write

$$\begin{aligned} \|f-g\|_4^4 &= \int_{\mathbb{R}^2}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y \\ &\ge \int_{\{F<G\}}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y \\ &\quad + \int_{\{F>G\}}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y. \end{aligned}$$

By symmetry, it suffices to estimate the first integral in the last expression. We split $\{F < G\} = \{G/2 \le F < G\} \cup \{F < G/2\} =: A_1 \cup A_2$.

On A_1 we compute

$$\begin{aligned} \int_{A_1}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y &\ge \int_{A_1}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(G+|x|^2)^8}\mathrm{d}x\right)\mathrm{d}y \\ &= \frac{\pi}{7}\int_{A_1}\frac{|F-G|^4}{G^7}\mathrm{d}y. \end{aligned}$$

Now, since $1/G^3 = u^6 \in L^1$, and since $\|u\|_6 = \|v\|_6 = \pi/2 \le 2$, Hölder's inequality yields

$$\int_{A_1}\frac{|F-G|}{G^4}\mathrm{d}y \le 2^{3/4}\left(\int_{A_1}\frac{|F-G|^4}{G^7}\mathrm{d}y\right)^{1/4} .$$

Also, pointwise on A_1,

$$|u^6 - v^6| = \left|\frac{1}{F^3} - \frac{1}{G^3}\right| = \left|\frac{(G-F)(G^2+GF+F^2)}{F^3G^3}\right| \le 14\,\frac{|G-F|}{G^4} .$$

Combining the last three estimates, we have

$$\int_{A_1}|u^6 - v^6|\mathrm{d}y \le 28\left(\frac{7}{\pi}\int_{A_1}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y\right)^{1/4} \tag{2.16}$$

For A_2, we observe that

$$\begin{aligned}
\int_{A_2}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y &\geq \int_{A_2}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y\\
&\geq \int_{A_2}\frac{|F-G|^4}{G^7}\left(\int_{\mathbb{R}^2}\frac{1}{(1+|x|^2)^4(F/G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y\\
&\geq \frac{1}{16}\int_{A_2}\frac{|F-G|^4}{G^7}\left(\int_{B_1}\frac{1}{(F/G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y\\
&= \frac{\pi}{48}\int_{A_2}\frac{|F-G|^4}{G^7}\left[\frac{1}{(F/G)^3}-\frac{1}{1+(F/G)^3}\right]\mathrm{d}y\\
&\geq \frac{\pi}{92}\int_{A_2}\frac{|G-F|^4}{F^3G^4}\mathrm{d}y,
\end{aligned}$$

where we used that $(1+|x|^2)^4 \leq 16$ on B_1, and that $\frac{1}{2(F/G)^3} \geq \frac{1}{1+(F/G)^3}$. Since $G/2 > F$ on A_2, $\dfrac{|G-F|^4}{F^3G^4} \geq \dfrac{1}{16}\dfrac{1}{F^3}$ on A_2. Therefore (using that $16\cdot 96 \leq 500\pi$)

$$\begin{aligned}
\int_{A_2}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y &\geq \frac{1}{500}\int_{A_2}\frac{1}{F^3}\mathrm{d}y\\
&\geq \frac{1}{500}\int_{A_2}\left(\frac{1}{F^3}-\frac{1}{G^3}\right)\mathrm{d}y\\
&= \frac{1}{500}\int_{A_2}|u^6-v^6|\mathrm{d}y.
\end{aligned}$$

When $\|f-g\|_4 \leq 1$, the left hand side is not greater than 1, and hence, taking the fourth root on the left, we obtain

$$\left(\int_{A_2}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\mathrm{d}y\right)^{1/4} \geq \frac{1}{500}\int_{A_2}|u^6-v^6|\mathrm{d}y$$

Combining this with (2.16), we have

$$\|f-g\|_4^4 \geq \left(\int_{\{F<G\}}\left(\int_{\mathbb{R}^2}\frac{|F-G|^4}{(F+|x|^2)^4(G+|x|^2)^4}\mathrm{d}x\right)\right)^{1/4}\mathrm{d}y \geq \frac{1}{500}\int_{\{u>v\}}(u^6-v^6)\mathrm{d}y\ .$$

By symmetry we also get

$$\|f-g\|_4^4 \geq \frac{1}{500}\int_{\{u<v\}}(u^6-v^6)\mathrm{d}y\ ,$$

which concludes the proof. □

2.4 Proof of Theorem 1.2

First, suppose that $u \in W^{1,2}(\mathbb{R}^2)$ is a non-negative function satisfying (2.3). Collecting together (2.7) and Lemmas 2.3 and 2.4, we deduce that there exist universal constants $K_1, \delta_1 > 0$ such that, whenever $\delta_{\mathrm{GNS}}[u] \leq \delta_1$,

$$\|u^6 - v^6(\cdot - y_0)\|_1 \leq K_1\delta_{\mathrm{GNS}}[u]^{1/2}. \tag{2.17}$$

Next, $\delta_{\rm GNS}[u]$ and $\|u\|_6$ are both unchanged if $u(y)$ is replaced by $u_\mu := \mu^{1/3}u(\mu y)$. Thus, assuming only that $\|u\|_6 = \|v\|_6$, we may choose a scale parameter μ so that $\sqrt{2}\|\nabla u_\mu\|_2 = \|u_\mu\|_4^2$. We then learn that

$$\int_{\mathbb{R}^2} \left|\mu^2 u^6(\mu y) - v(y-y_0)\right| {\rm d}y \le K_1 \delta_{\rm GNS}[u]^{1/2} \ .$$

Changing variables once more, and taking $\lambda := 1/\mu$, we obtain

$$\int_{\mathbb{R}^2} \left|u^6(y) - \lambda^2 v(\lambda y-y_0)\right| {\rm d}y \le K_1 \delta_{\rm GNS}[u]^{1/2} \ ,$$

which proves (1.12) and concludes the proof.

2.5 Controlling the translation

So far we know that if u satisfies (2.3) there is *some translate* $\widetilde{u}(y) = u(y-y_0)$ of u such that

$$\|\widetilde{u}^6 - v^6\|_1 \le K_1 \delta_{\rm GNS}[u]^{1/2} \tag{2.18}$$

for some universal constant K_1 (see Theorem 1.2 and Remark 1.3).

Our goal in this section is to show that under the additional hypotheses that

$$M_p[u] := \int_{\mathbb{R}^2} |y|^p u^6(y){\rm d}y < \infty \qquad \text{for some } p > 1 \tag{2.19}$$

and

$$\int_{\mathbb{R}^2} y u^6(y){\rm d}y = 0 \ , \tag{2.20}$$

then $\|u^6 - v^6\|_1$ will be bounded by a multiple of some fractional power of $\delta_{\rm GNS}[u]^{1/2}$, with the fractional power depending on how large p can be taken in (2.19). The power would still be $\delta_{\rm GNS}[u]^{1/2}$ if we could take $p = \infty$. However, since $M_4(v) = +\infty$, the useful values of p are those in the range $1 < p < 4$.

We warn the reader that the following result is not needed for our application to the Keller-Segel equation. However, we present it for two reasons: First, we believe that the result and the argument for its proof have an interest in their own; as we have explained, u^6 is in some ways the most natural density associated to our GNS inequality.. Second, an analogous argument we will also be needed in Step 7 in the proof of Theorem 1.4. Since the proof of Theorem 1.4 is somewhat involved, we believe it should be easier for the reader to have already seen the argument employed here.

2.5 PROPOSITION. *Let $u \in W^{1,2}(\mathbb{R}^2)$ be a non-negative function satisfying* (2.3)*, and suppose that $M_p(u) < \infty$ for some $1 < p < 4$, and that (2.20) is satisfied. Then there are constants $\tilde{K}, \tilde{\delta} > 0$, with $\tilde{\delta}$ universal and $\tilde{K}$ depending only on p and $M_p(u)$, so that whenever $\delta_{\rm GNS}[u] \le \tilde{\delta}$,*

$$\|u^6 - v^6\|_1 \le \tilde{K}\delta_{\rm GNS}[u]^{(p-1)/2p} \ .$$

Proof. First note that

$$y_0\|v\|_6^6 = \int_{\mathbb{R}^2} y\widetilde{u}^6(y){\rm d}y = \int_{\mathbb{R}^2} y v^6(y){\rm d}y + \int_{\mathbb{R}^2} y[\widetilde{u}^6(y) - v^6(y)]{\rm d}y \ .$$

By the symmetry of v, the first term on the right is zero. By Hölder's inequality,

$$\begin{aligned}\int_{\mathbb{R}^2} |y||\widetilde{u}^6(y) - v^6(y)|\mathrm{d}y &= \int_{\mathbb{R}^2} |y||\widetilde{u}^6(y) - v^6(y)|^{1/p}|\widetilde{u}^6(y) - v^6(y)|^{1/q}\mathrm{d}y \\ &\leq \left(\int_{\mathbb{R}^2} |y|^p|\widetilde{u}^6(y) - v^6(y)|\mathrm{d}y\right)^{1/p} \|\widetilde{u}^6 - v^6\|_1^{1/q} \\ &\leq \left(M_p(\widetilde{u})^{1/p} + M_p(v)^{1/p}\right) \|\widetilde{u}^6 - v^6\|_1^{1/q} \ ,\end{aligned}$$

where $q = p/(p-1)$. Next we note that

$$\begin{aligned}M_p(\widetilde{u})^{1/p} &= \left(\int_{\mathbb{R}^2} |y|^p\widetilde{u}^6(y)\mathrm{d}y\right)^{1/p} \\ &= \left(\int_{\mathbb{R}^2} |y+y_0|^p u^6(y)\mathrm{d}y\right)^{1/p} \leq M_p(u)^{1/p} + |y_0|\|v\|_6^{6/p}\end{aligned}$$

Combining the last three estimates, we obtain

$$|y_0|\|v^6\|_1 \leq \left[M_p(u)^{1/p} + M_p(v)^{1/p} + |y_0|\|v\|_6^{6/p}\right] \|\widetilde{u}^6 - v^6\|_1^{1/q} \ ,$$

and therefore by (2.18)

$$|y_0| \leq \frac{\left(M_p(u)^{1/p} + M_p(v)^{1/p}\right) K_1\delta_{\mathrm{GNS}}^{1/2q}[u]}{\|v^6\|_1 - \|v^6\|_1^{1/p}K_1\delta_{\mathrm{GNS}}^{1/2q}[u]} \ .$$

Hence, there exist a constant K' depending only on p and $M_p(u)$, and a universal constant δ' (recall that K_1 is universal and $p \in (1,4)$), such that whenever $\delta_{\mathrm{GNS}}[u] < \delta'$,

$$|y_0| \leq K'\delta^{1/2q}(u) \ . \tag{2.21}$$

Now note that, by the Fundamental Theorem of Calculus and Hölder's inequality,

$$\|\widetilde{u}^6 - u^6\|_1 \leq 6|y_0|\|\nabla u\|_2\|u\|_{10}^5 \ .$$

Then by the GNS inequality $\|u\|_{10} \leq C\|\nabla u\|_2^{2/5}\|u\|_6^{3/5}$ we obtain

$$\|\widetilde{u}^6 - u^6\|_1 \leq 6C|y_0|\|\nabla u\|_2^3\|u\|_6^3 \ . \tag{2.22}$$

We now want to control the right hand side. By hypothesis, $\sqrt{2}\|\nabla u\|_2 = \|u\|_4^2$ and $\|\nabla u\|_2\|u\|_4^2 = \sqrt{\pi}\|u\|_6^3 + \delta_{\mathrm{GNS}}[u]$. Therefore

$$\|\nabla u\|_2^3 = \sqrt{\frac{\pi}{2}}\|u\|_6^3 + \frac{\delta_{\mathrm{GNS}}[u]}{\sqrt{2}} \ .$$

Again using the fact that $\|u\|_6 = \|v\|_6 = \pi/2$ and that $\delta_{\mathrm{GNS}}[u]$ is small (so in particular we can assume $\delta_{\mathrm{GNS}}[u] \leq 1$) from (2.22) we obtain that

$$\|\widetilde{u}^6 - u^6\|_1 \leq K''|y_0| \ ,$$

for some universal constant K''. Combining this with (2.21) yields the result. □

2.6 Bounding $\inf_{\lambda>0}\|u^4 - v_\lambda^4\|_1$

As noted in the introduction, in certain PDE applications of stability estimate for the GNS inequality (1.8), u^4 will play the role of a mass density, and it will be of interest to control $\inf_{\lambda>0}\|u^4 - v_\lambda^4\|_1$ assuming that some moments of u^4 exist, and that the "normalization" assumptions

$$\int_{\mathbb{R}^2} u^4(y)\mathrm{d}y = \int_{\mathbb{R}^2} v^4(y)\mathrm{d}y \qquad \text{and} \qquad \int_{\mathbb{R}^2} yu^4(y)\mathrm{d}y = 0 \tag{2.23}$$

hold. Since the GNS deficit functional is scale invariant, we cannot hope to get information on the minimizing value of λ out of a bound on $\delta_{\mathrm{GNS}}[u]$ alone. However, knowing that the density u^4 is close to v_λ^4 for *some* λ is a strong information that can be combined with other ones, specific to a particular application, that then fix the scale λ. We shall see an example of this in the next section. Here we concentrate on proving Theorem 1.4 which bounds $\inf_{\lambda>0}\|u^4 - v_\lambda^4\|_1$ in terms of $\delta_{\mathrm{GNS}}[u]$.

We recall that Theorem 1.4 refers to non-negative functions $u \in W^{1,2}(\mathbb{R}^2)$ that satisfy (2.23) and also certain moment and entropy conditions: Recall we have defined

$$N_p(u) = \int_{\mathbb{R}^2} |y|^p u^4(y)\mathrm{d}y \qquad \text{and} \qquad S(u) = \int_{\mathbb{R}^2} u^4 \log(u^4)\mathrm{d}y\ .$$

and Theorem 1.4 also requires that for some $A, B < \infty$ and some $1 < p < 2$,

$$S(u) \le A \qquad \text{and} \qquad N_p(u) \le B\ . \tag{2.24}$$

2.6 Remark. Conditions (2.24) provide some "uniform integrability control" on the class of densities u^4 that satisfy them. The proof that we give would yield similar results for essentially any other pair of conditions that quantify uniform integrability. The one we have chosen, moments and entropy, are natural in PDE applications. It is natural that some such condition is required: a bound on the deficit does not supply any compactness, as is clear from the scale-invariance.

Proof of Theorem 1.4: The proof is divided in several steps.

• *Step 1: We show that $\|u\|_6$ cannot be too small provided $N_p(u)$ is not too large.* Indeed,

$$\int_{B_R} |u|^4 \mathrm{d}y \ge \|v\|_4^4 - R^{-p}N_p(u)\ .$$

Choosing $R > 0$ such that $R^{-p}N_p(u) = \|v\|_4^4/2$ and using Hölder's inequality, we get

$$\frac{1}{2}\|v\|_4^4 \le \int_{B_R} |u|^4 \le \|u\|_6^2(\pi R^2)^{2/3}\ ,$$

that is

$$\|u\|_6^6 \ge c_1 N_p(u)^{4/p} \tag{2.25}$$

for some universal constant $c_1 > 0$.

• *Step 2: To apply our previous results, we must multiply u by a constant and rescale. In this step we show that these modifications do not seriously affect the size of the deficit $\delta_{\mathrm{GNS}}[u]$.*

Define

$$\widetilde{u}(y) := \frac{\|v\|_6}{\|u\|_6}\lambda^{1/3}u(\lambda y)\ .$$

where λ is chosen so that $\sqrt{2}\|\nabla\widetilde{u}\|_2 = \|\widetilde{u}\|_4^2$. Note that $\|\widetilde{u}\|_6 = \|v\|_6$. Since the rescaling does not affect the L^6 norm, it does not affect the deficit, but the constant multiple does: we have

$$\delta_{\rm GNS}[\widetilde{u}] = \frac{\|v\|_6^3}{\|u\|_6^3}\delta_{\rm GNS}[u] \ , \tag{2.26}$$

By what we have noted in Step 1, we have an *a-priori* upper bound on the factor $\|v\|_6^3/\|u\|_6^3$ (see (2.25)), which gives the bound

$$\delta_{\rm GNS}[\widetilde{u}] \le C\delta_{\rm GNS}[u] \ . \tag{2.27}$$

• *Step 3: We now relate the constant multiple and the scale factor when the deficit is small.* First, we claim that

$$\left|\|\widetilde{u}^4\|_4^4 - \|v\|_4^4\right| \le C\delta_{\rm GNS}[u] \ . \tag{2.28}$$

To see this note that, since $\sqrt{2}\|\nabla\widetilde{u}\|_2 = \|\widetilde{u}\|_4^2$ (see Step 2),

$$\left|2\|\widetilde{u}\|_4^6 - \pi\|\widetilde{u}\|_6^6\right| = \delta(\widetilde{u}) \ ,$$

The claim then follows by (2.27) together with

$$\pi\|\widetilde{u}\|_6^6 = \pi\|v\|_6^6 = 2\|v\|_4^6 \ .$$

Let us also observe that, since

$$\|v\|_4^4 = \|u\|_4^4 = \lambda^{2/3}\frac{\|u\|_6^4}{\|v\|_6^4}\|\widetilde{u}\|_4 \ ,$$

by (2.28) we also get

$$\left|\lambda^{2/3}\|u\|_6^4 - \|v\|_6^4\right| \le C\delta_{\rm GNS}[u] \ . \tag{2.29}$$

• *Step 4: We now show that some translate $\hat{u}^4$ of $\widetilde{u}^4$ is close to v^4 when the deficit is small.* Theorem 1.2 shows that there is a translate $\hat{u}(y) = \widetilde{u}(y-y_0)$ of $\widetilde{u}$ such that

$$\|\hat{u}^6 - v^6\|_1 \le K_1\delta_{\rm GNS}[\widetilde{u}]^{1/2} \ . \tag{2.30}$$

Note that for positive numbers a and b,

$$|a^4-b^4| = |a-b|(a^3+a^2b+a^2b+b^3) \quad \text{and} \quad |a^6-b^6| = |a-b|(a^5+a^4b+a^3b^2+a^2b^3+ab^4+b^5) \ .$$

Hence, since $(a^2+b^2)(a^3+a^2b+a^2b+b^3) \le 2(a^5+a^4b+a^3b^2+a^2b^3+ab^4+b^5)$, it follows that

$$|a^4-b^4| \le \frac{2}{a^2+b^2}|a^6-b^6| \ . \tag{2.31}$$

So, observing that

$$\frac{1}{u^2+v^2} \le \frac{1}{v^2} \le 1+R^2 \qquad \text{on } B_R,$$

by (2.31), (2.30), and (2.27), we obtain

$$\int_{B_R}|\hat{u}^4 - v^4|{\rm d}y \le 2(1+R^2)\|\hat{u}^6 - v^6\|_1 \le C(1+R^2)\delta_{\rm GNS}[u]^{1/2} \ .$$

Next, using (2.28),

$$\begin{aligned}\int_{|x|\geq R} \hat{u}^4 \mathrm{d}y &= \|\hat{u}\|_4^4 - \int_{|x|\leq R} \hat{u}^4 \mathrm{d}y \\ &\leq \|v\|_4^4 - \int_{|x|\leq R} \hat{u}^4 \mathrm{d}y + C\delta_{\mathrm{GNS}}[u] \\ &= \int_{|x|>R} v^4 \mathrm{d}y + \int_{|x|\leq R} (v^4 - \hat{u}^4)\mathrm{d}y + C\delta_{\mathrm{GNS}}[u] \\ &\leq \frac{\pi}{1+R^2} + \int_{|x|\leq R} |v^4 - \hat{u}^4|\mathrm{d}y + C\delta_{\mathrm{GNS}}[u]\end{aligned}$$

Combining results, we then get

$$\|\hat{u}^4 - v^4\|_1 \leq C(1+R^2)\delta_{\mathrm{GNS}}[u]^{1/2} + C(1+R^2)^{-1} \ ,$$

which (optimizing with respect to R) leads to the estimate

$$\|\widetilde{u}^4(\cdot - y_0) - v^4\|_1 = \|\hat{u}^4 - v^4\|_1 \leq C\delta_{\mathrm{GNS}}[u]^{1/4} \ . \tag{2.32}$$

• *Step 5:* Set $u_{1/\lambda} := \lambda^{1/2} u(\lambda y)$. Note that $\|u_{1/\lambda}\|_4 = \|u\|_4$ and

$$\int_{\mathbb{R}^2} |\widetilde{u}^4 - u_{1/\lambda}^4|\mathrm{d}y = \left| \frac{\|v\|_6^4}{\|u\|_6^4} \lambda^{-2/3} - 1 \right| \|u\|_4^4 .$$

Now, by (2.29), $\lambda^{2/3}\|u\|_6^4$ is uniformly bounded away from zero (for $\delta_{\mathrm{GNS}}[u]$ smaller than some universal constant). Therefore

$$\int_{\mathbb{R}^2} |\widetilde{u}^4 - u_{1/\lambda}^4|\mathrm{d}y \leq \frac{C\delta_{\mathrm{GNS}}[u]}{\lambda^{2/3}\|u\|_6^4} \|u\|_4^4 \leq C\delta_{\mathrm{GNS}}[u]\|u\|_4^4 \ , \tag{2.33}$$

which combined with (2.32) gives

$$\|u_{1/\lambda}^4(y - y_0) - v^4\|_1 \leq C\delta_{\mathrm{GNS}}[u]^{1/4}. \tag{2.34}$$

• *Step 6: We obtain upper and lower bounds on the scaling parameter λ.* We already have an upper bound since (2.29) says that $\lambda^{-1} \sim \|u\|_6^6$, and (2.25) gives a lower bound for $\|u\|_6^6$ in terms of $N_p(u)$. Our assumption that $S(u)$, the entropy of u^4 (see (1.13)), is finite enters at this point.

Since $\|v\|_4 = \pi/2$, if follows from (2.34) that, if $\delta_{\mathrm{GNS}}[u]$ is smaller than some universal constant,

$$\int_{B_1} \lambda^2 u^4(\lambda(y - y_0))\mathrm{d}y \geq \frac{\pi}{4},$$

or equivalently

$$\int_{B_\lambda(\lambda y_0)} u^4(y)\mathrm{d}y \geq \frac{\pi}{4}.$$

Thus, the average value of u^4 on $B_\lambda(\lambda y_0)$ is at least $\lambda^{-2}/2$. Hence by Jensen's inequality,

$$\begin{aligned}\frac{1}{\pi\lambda^2}\int_{B_\lambda(\lambda y_0)} u^4(y)\log(u^4(y))\mathrm{d}y &\geq \left(\frac{1}{\pi\lambda^2}\int_{B_\lambda(\lambda y_0)} u^4(y)\mathrm{d}y\right)\log\left(\frac{1}{\pi\lambda^2}\int_{B_\lambda(\lambda y_0)} u^4(y)\mathrm{d}y\right) \\ &\geq \frac{1}{4\lambda^2}\log\left(\frac{1}{4\lambda^2}\right),\end{aligned}$$

that is

$$\int_{B_\lambda(\lambda y_0)} u^4(y)\log(u^4(y))\mathrm{d}y \geq \pi\left(-\frac{\log 2}{2}-\frac{\log\lambda}{2}\right) .$$

Next we recall a standard estimate, valid for any non-negative integrable function ρ on $\mathbb{R}^2$ with finite first moment (see for instance [6, Lemma 2.4]):

$$\int_{\mathbb{R}^2}\rho(x)\,\log_-(\rho(x))\,\mathrm{d}x \leq \int_{\mathbb{R}^2}|x|\rho(x)\,\mathrm{d}x + \frac{1}{e}\int_{\mathbb{R}^2}e^{-|x|}\mathrm{d}x = N_1(u)+\frac{2\pi}{e} ,$$

where $\log_-(s)=\max\{-\log(s),0\}$. Combining all the estimates together, we arrive at

$$-\log\lambda \leq \frac{2}{\pi}(S(u)+N_1(u))+\frac{4}{e}+\log 2 . \tag{2.35}$$

Since $N_1(u)\leq \|u\|_4^4+N_p(u)$ for $p\geq 1$, the above inequality provides the desired lower bound on λ.

• *Step 7: We now reabsorb y_0.* Arguing as in the proof of Proposition 2.5 and using (1.15) and (2.32), we get

$$\begin{aligned}
|y_0|\|v^4\|_4 &\leq \left[N_p(\hat{u})^{1/p}+N_p(v)^{1/p}\right]\|\hat{u}^4-v^4\|_1^{1/q}\\
&\leq C\left[N_p(\widetilde{u})^{1/p}+|y_0|\|v^4\|_1^{1/p}+N_p(v)^{1/p}\right]\delta_{\mathrm{GNS}}[u]^{1/(4q)}\\
&\leq C\left[\lambda^{-p-2/3}N_p(u)^{1/p}+|y_0|\|v^4\|_1^{1/p}+N_p(v)^{1/p}\right]\delta_{\mathrm{GNS}}[u]^{1/(4q)} .
\end{aligned}$$

By the bound (2.35) this implies $|y_0|\leq C\delta_{\mathrm{GNS}}[u]^{1/(4q)}$. So, since

$$\|\widetilde{u}^4-\hat{u}^4\|_1 = \|\widetilde{u}^4-\widetilde{u}^4(\cdot-y_0)\|_1 \leq 4|y_0|\|\nabla\widetilde{u}\|_2\|\widetilde{u}\|_6^3$$

and $\sqrt{2}\|\nabla\widetilde{u}\|_2=\|\widetilde{u}\|_4^2=\|v\|_4^2$, as in the proof of Proposition 2.5 we get

$$\|\widetilde{u}^4-v^4\|_1\leq C\delta_{\mathrm{GNS}}[u]^{(p-1)/(4p)} .$$

In particular, by (2.33) we obtain

$$\|u_{1/\lambda}^4-v^4\|_1\leq C\delta_{\mathrm{GNS}}[u]^{(p-1)/(4p)},$$

which is equivalent to (1.16). □

3 Application to stability for the Log-HLS inequality and to Keller-Segel equation

3.1 *A-priori* estimates

In this section we apply the results proved in the previous section, carrying out the strategy for quantitatively bounding the rate of approach to equilibrium for critcal mass solutions of the Keller-Segel equation, and, along the way, proving a stability result for the Log-HLS inequality. This and several other results obtained here may be of interest apart from their particular application to the Keller-Segel equation.

First of all, we recall some *a-priori* regularity results concerning functions in level sets of the various functional $\mathcal{F}$, $\mathcal{D}$ and $\mathcal{H}_{\kappa,M}$ that have been defined in the introduction.

As we have seen $\mathcal{F}[\sigma_{\kappa,M}] = \mathcal{D}[\sigma_{\kappa,M}] = 0$ for all κ and M. But as κ tends to 0, the measures $\sigma_{\kappa,M}(x)\mathrm{d}x$ tend to a point mass (of mass M). Hence the level sets of neither $\mathcal{F}$ nor $\mathcal{D}[\sigma_{\kappa,M}]$ are compact in $L^1(\mathbb{R}^2)$ or even uniformly integrable. It is also easy to see that the level sets of $\mathcal{H}_{\kappa,M}$ are not compact in $L^1(\mathbb{R}^2)$, or even uniformly integrable. However, as shown in [6], taken together bounds on various combinations of $\mathcal{F}[\rho]$, $\mathcal{H}_{\kappa,M}[\rho]$ and $\mathcal{D}[\rho]$ do yield strong estimates on ρ.

First, we recall that $\mathcal{F}$ and $\mathcal{H}_\kappa$ provide control of the entropy [6, Theorem 1.9]. The two theorems that follow were proved for $M = 8\pi$, in which case the log-HLS functional coincides with the Keller-Segel. However, it is easy to see that the proofs are valid for the log-HLS functional*for any* mass $M > 0$. Of course, our main application here is to the critical mass Keller-Segel equation, and the reader interested only in this equation can assume $M = 8\pi$. Here and in the sequel, $\log_+$ denotes the positive part of the natural logarithm function.

3.1 THEOREM (Entropy bound via $\mathcal{F}$ and $\mathcal{H}_\kappa$)**.** *Let ρ be any density on $\mathbb{R}^2$ with mass M satisfying $\mathcal{H}_{\kappa,M}[\rho] < \infty$ for some $\kappa > 0$. Then there exist positive computable constants γ_1 and $C_{\mathcal{FH}}$, depending only on M, κ and $\mathcal{H}_{\kappa,M}[\rho]$, such that*

$$\gamma_1 \int_{\mathbb{R}^2} \rho \log_+ \rho(x)\mathrm{d}x \le \mathcal{F}[\rho] + C_{\mathcal{FH}}\,. \tag{3.1}$$

Likewise, [6, Theorem 1.10] shows that a bound on $\mathcal{F}$, $\mathcal{H}_{\kappa,M}$ and $\mathcal{D}$ together controls the energy integral $\|\nabla u\|_2^2$.

3.2 THEOREM (Energy bound via $\mathcal{F}$, $\mathcal{H}_{\kappa,M}$ and $\mathcal{D}$)**.** *Let ρ be any density on $\mathbb{R}^2$ with mass M, with $\mathcal{F}[\rho]$ finite, and $\mathcal{H}_{\kappa,M}[\rho]$ finite for some $\kappa > 0$. Then there exist positive computable constants γ_2 and $C_{\mathcal{FHD}}$, depending only on M, κ, $\mathcal{H}_{\kappa,M}[\rho]$ and $\mathcal{F}[\rho]$, such that*

$$\gamma_2 \int_{\mathbb{R}^2} |\nabla \rho^{1/4}|^2 \mathrm{d}x \le \pi\mathcal{D}[\rho] + C_{\mathcal{FHD}}\ . \tag{3.2}$$

Recall the classical Gagliardo-Nirenberg inequality

$$\int_{\mathbb{R}^2} |v|^p \mathrm{d}x \le D_p \left[\int_{\mathbb{R}^2} |\nabla v|^2 \mathrm{d}x\right]^{p/2-2} \int_{\mathbb{R}^2} |v|^4 \mathrm{d}x \qquad \forall\, p \in [4,\infty)\,. \tag{3.3}$$

Combining this with (3.2), we see that together $\mathcal{F}[\rho]$, $\mathcal{H}_{\kappa,M}[\rho]$ and $\mathcal{D}[\rho]$ give us a quantitative bound on $\|\rho\|_q$ for all $q < \infty$:

3.3 COROLLARY (L^q bound $\mathcal{F}$, $\mathcal{H}_{\kappa,M}$ and $\mathcal{D}$)**.** *Let ρ be any density on $\mathbb{R}^2$ with mass M, with $\mathcal{H}_{\kappa,M}[\rho]$ finite for some $\kappa > 0$, such that also $\mathcal{F}[\rho]$ and $\mathcal{D}[\rho]$ are finite. Then, for all $q \ge 1$, there is a constant C depending only on M, q, κ, $\mathcal{F}[\rho]$, $\mathcal{H}_{\kappa,M}[\rho]$ and $\mathcal{D}[\rho]$, such that*

$$\|\rho\|_q \le C\ . \tag{3.4}$$

Next, with an argument analogous to the one used in [6], we can use the functional $\mathcal{H}_{\kappa,M}$ to control pth moments for all $p < 2$:

3.4 THEOREM (Moments and lower bounds on the $L^{3/2}$-norm via $\mathcal{H}_{\kappa,M}$). *Let ρ be a density on $\mathbb{R}^2$ with mass M. For all $0 \le p < 2$, there is a constant C, depending only on M, p and κ, such that*

$$\int_{\mathbb{R}^2} |x|^p \rho(x) \mathrm{d}x \le C\big(1 + \mathcal{H}_{\kappa,M}[\rho]\big)\,, \tag{3.5}$$

$$\|\rho\|_{3/2} \ge \frac{C}{\big(1 + \mathcal{H}_{\kappa,M}[\rho]\big)^{1/p}}\,, \tag{3.6}$$

Proof. Since $\sigma_{\kappa,M}$ has finite pth moments for all $p < 2$, to prove (3.5) it suffices to estimate

$$\int_{\mathbb{R}^2} |x|^p \big|\rho(x) - \sigma_{\kappa,M}(x)\big| \mathrm{d}x.$$

Observing that $|x|^p \le C/\sqrt{\sigma_{\kappa,M}(x)}$ and that

$$\frac{\big|\rho - \sigma_{\kappa,M}\big|}{\sqrt{\sigma_{\kappa,M}}} \le \frac{\big|\sqrt{\rho} - \sqrt{\sigma_{\kappa,M}}\big|^2}{\sqrt{\sigma_{\kappa,M}}} + 2\sigma_{\kappa,M}^{1/4} \frac{\big|\sqrt{\rho} - \sqrt{\sigma_{\kappa,M}}\big|}{\sigma_{\kappa,M}^{1/4}},$$

(3.5) follows easily using Hölder inequality.

Finally, (3.6) is a consequence of (3.5): indeed, for any $\theta \in (0,1)$ and $q > 1$,

$$\begin{aligned} \|\rho\|_1 &\le \left(\int_{\mathbb{R}^2} (1 + |x|^p)\rho(x)\mathrm{d}x\right)^{\theta} \left(\int_{\mathbb{R}^2} \frac{\rho(x)}{(1+|x|^p)^{\theta/(1-\theta)}} \mathrm{d}x\right)^{1-\theta} \\ &\le \left(\int_{\mathbb{R}^2} (1 + |x|^p)\rho(x)\mathrm{d}x\right)^{\theta} \|(1+|x|^p)^{-\theta/(1-\theta)}\|_{q/(q-1)}^{1-\theta} \|\rho\|_q^{1-\theta} \end{aligned}$$

Choosing $q = 3/2$ and $\theta = 1/(p+1)$ (so that $(1 + |x|^p)^{-\theta/(1-\theta)} \in L^3(\mathbb{R}^2)$), we get

$$M \le C\|\rho\|_{3/2}^{p/(p+1)},$$

which proves (3.6). □

We close this subsection with the following observation that will be used later: unlike $\mathcal{F}$ and $\mathcal{D}$, the functional $\mathcal{H}_{\kappa,M}$ is not scale invariant. Indeed, for any $M > 0$, $\mathcal{H}_{\kappa,M}[\sigma_{\mu,M}] < \infty$ if and only if $\mu = \kappa$. In fact, later we shall need a somewhat more precise version of this estimate, which can be easily proved by a direct computation: there exists a constant $c_0 > 0$ depending only on κ and M such that

$$\int_{|y|<R} \frac{|\sqrt{\sigma_{\mu,M}(y)} - \sqrt{\sigma_{\kappa,M}(y)}|^2}{\sqrt{\sigma_{\kappa,M}(y)}} \mathrm{d}y \ge c_0(\sqrt{\mu} - \sqrt{\kappa})^2 \log(1 + R) \tag{3.7}$$

3.2 A quantitative convergence result for the critical mass Keller-Segel equation

We now state and prove our a quantitative bound on the rate of relaxation to equilibrium for the critical mass Keller-Segel equation. Recall that $C(M)$ is the constant appearing in (1.23).

3.5 THEOREM. *Let $\rho(t, x)$ be any properly dissipative solution of the Keller-Segel equation of critical mass $M = 8\pi$ in the sense of [6], so that in particular $\mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)] < \infty$ for some $\kappa > 0$,* *and* $\mathcal{F}[\rho(0,\cdot)] < \infty$. *Let us suppose also that* $\int_{\mathbb{R}^2} x\rho(0,x)\mathrm{d}x = 0$. *Then, for all* $\epsilon > 0$, *there are constants* C_1 *and* C_2, *depending only on* ϵ, κ, $\mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)]$ *and* $\mathcal{F}[\rho(0,\cdot)]$, *such that, for all* $t > 0$,

$$\mathcal{F}[\rho(t,\cdot)] - C(8\pi) \le C_1(1+t)^{-(1-\epsilon)/8} \tag{3.8}$$

$$\inf_{\mu>0} \|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1 \le C_2(1+t)^{-(1-\epsilon)/320}. \tag{3.9}$$

Moreover, there is a positive number $a > 0$, *depending only on* $\mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)]$ *and* $\mathcal{F}[\rho(0,\cdot)]$, *so that for each* $t > 0$,

$$\inf_{\mu>0} \|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1 = \min_{a<\mu<1/a} \|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1.$$

Finally, for each $t > 0$, *the above minimum is achieved at some value* $\mu(t)$ *such that*

$$(\mu(t) - \kappa)^2 \le \frac{C}{\log(e+t)} \ . \tag{3.10}$$

In particular

$$\|\rho(t,\cdot) - \sigma_{\kappa,8\pi}\|_1 \le \frac{C}{\sqrt{\log(e+t)}} \,. \tag{3.11}$$

As indicated in the introduction, to carry out the proof of Theorem 3.5, we need an "almost Lipschitz" property of the functional $\mathcal{F}$. We introduce this next, before turning to the proof of Theorem 3.5.

To obtain continuity properties of the Log-HLS functional, we will need to impose some restrictions on the set of densities. In view of the wider interest of the almost Lipschitz continuity of the entropic part of $\mathcal{F}$ (Theorem 3.10 below), our next definition refers to densities on $\mathbb{R}^n$.

3.6 DEFINITION. For $p > 0$, $q > 1$ and $A, B < \infty$, let $\mathcal{M}_{n,p,q,A,B}$ denote the set of mass densities ρ on $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n} |x|^p \rho(x)\mathrm{d}x \le A \qquad \text{and} \qquad \int_{\mathbb{R}^n} |\rho(x)|^q \mathrm{d}x \le B \ . \tag{3.12}$$

Note that we *do not* specify the mass of the densities in $\mathcal{M}_{p,q,A,B}$ though of course they can be bounded above in terms of A and B. A key result (which will be proved later in Subsection 3.4) is that, for any p, q, A and B, the Log-HLS functional is almost Lipschitz continuous on $\mathcal{M}_{2,p,q,A,B}$:

3.7 THEOREM. *For all* $0 < \epsilon < 1$, *and all* $M > 0$, *there is a constant* C *depending only on* ϵ, M, p, q, A *and* B *such that for any* $\rho, \sigma \in \mathcal{M}_{2,p,q,A,B}$ *both of mass* M,

$$|\mathcal{F}[\rho] - \mathcal{F}[\sigma]| \le C\|\rho - \sigma\|_1^{1-\epsilon} \ .$$

Proof of Theorem 3.5: Of course it suffices to prove the estimates in Theorem 3.5 for t large.

As shown in [6], for all $\tau > 0$, and all $p < 2$ and $q < \infty$, there exist finite constants A and B such that

$$\text{for all} \quad t \ge \tau \ , \qquad \rho(t,\cdot) \in \mathcal{M}_{2,p,q,A,B} \ ,$$

see also Subsection 3.1.

Choose $\tau = 1$. As noted earlier in this section (see (1.27)), by the definition of properly dissipative solution

$$\mathcal{H}_{\kappa,8\pi}[\rho(T,\cdot)] + \int_0^T \mathcal{D}[\rho(t,\cdot)]\mathrm{d}t \le \mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)] \ , \tag{3.13}$$

we immediately deduce that, for all $T > 1$,

$$\frac{1}{T-1}\int_1^T \mathcal{D}[\rho(t,\cdot)]\mathrm{d}t \le \frac{1}{T-1}\mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)] \ . \tag{3.14}$$

Then, Theorem 3.4 together with (3.13) ensures a uniform bound on $\int_{\mathbb{R}^2}|x|^p\rho(t,x)\mathrm{d}x$ for any $p < 2$, which also ensure a lower bound on $\|\rho(t,\cdot)\|_{3/2}$.

Hence, by (1.28) and (3.14), we deduce that for any $T \ge 2$ there exists some $1 \le t \le T$ such that

$$\delta_{\mathrm{GNS}}[\rho^{1/4}(t,\cdot)] \le \frac{C}{T}\mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)] \ .$$

Next, Theorem 3.1 gives us an *a-priori* upper bound on the entropy $S[\rho(t,\cdot)]$, and thus permits us to apply Theorem 1.4 for T sufficiently large (observe that the baricenter condition on $\rho(0,\cdot)$ is preserved along the flow). Hence we conclude that, for any $p < 2$, there exist $a > 0$, *some* $\mu \in [a, 1/a]$, and *some* $1 \le t \le T$, such that

$$\|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1 \le C\left(\frac{1}{T}\mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)]\right)^{(p-1)/4p} \ , \tag{3.15}$$

(recall that the density v_λ^4 is a multiple of some $\sigma_{\mu,8\pi}$).

Next, since we can choose p arbitrarily close to 2, by Theorem 3.7

$$\mathcal{F}[\rho(t,\cdot)] - C(8\pi) \le CT^{-(1-\epsilon)/8}$$

for *some* $1 \le t \le T$. However we can now use that $\mathcal{F}[\rho(t,\cdot)]$ is monotone decreasing to deduce that

$$\mathcal{F}[\rho(T,\cdot)] - C(8\pi) = \delta_{\mathrm{HLS}}[\rho(T,\cdot)] \le CT^{-(1-\epsilon)/8} \tag{3.16}$$

for *all* T sufficiently large. Hence, up to adjusting the constant C we obtain (3.8).

We next wish to apply Theorem 1.9, the stability theorem for the Log-HLS inequality to show that there is *some* $\mu = \mu(t) \in [a, 1/a]$ such that $\|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1$ is controlled by a power of $\mathcal{F}[\rho(T,\cdot)] - C(8\pi)$ for all large t. By what we have just proved, this would give us algebraic decay, in t, of $\|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1$. However, we cannot immediately apply Theorem 1.9 since under the Keller-Segel evolution we do not have uniform-in-time control on $\mathcal{D}[\rho(\cdot,t)]$, though we do have uniform-in-time control on all of the other quantities whose bounds enter into Theorem 1.9. (In particular, we have uniform bounds on all L^q norms for all times $t > 1$, and note in addition that $\mathcal{F}$ and $\mathcal{H}_{\kappa,8\pi}$ are non-increasing.)

The only bound on $\mathcal{D}$ that we have is that for all $0 < t - t_0 < t$,

$$\min\{D[\rho(s,\cdot)] \ : \ t - t_0 \le s \le t\} \le \frac{1}{t_0}\int_{t-t_0}^t \mathcal{D}[\rho(s,\cdot)]\mathrm{d}s \le \frac{1}{t_0}\mathcal{H}_{\kappa,8\pi}[\rho] \ .$$

This will suffice since the next lemma (which we believe to have an interest on its own) provides a one-sided Lipschitz estimate on the function $t \mapsto \|\sqrt{\rho(t,\cdot)} - \sqrt{\sigma_{\mu,8\pi}}\|_2^2$:

3.8 LEMMA. *Let $\rho(t,\cdot)$ be a properly dissipative solution on the Keller-Segel equation satisfying the hypotheses of Theorem 3.5. Then for any $\tau > 0$ and $a > 0$ there is a constant C depending only on τ, a, κ, $\mathcal{H}_{\kappa,8\pi}[\rho]$ and $\mathcal{F}[\rho]$ such that for any $\mu \in [a, 1/a]$,*

$$\|\sqrt{\rho(t,\cdot)} - \sqrt{\sigma_{\mu,8\pi}}\|_2^2 \leq \|\sqrt{\rho(s,\cdot)} - \sqrt{\sigma_{\mu,8\pi}}\|_2^2 + C(t-s) \qquad \forall\, t > s \geq \tau \ .$$

We postpone the proof of this lemma to Subsection 3.5.

Since $|\sqrt{\alpha} - \sqrt{\beta}|^2 \leq |\alpha - \beta| = |\sqrt{\alpha} - \sqrt{\beta}|(\sqrt{\alpha} + \sqrt{\beta})$ for all $\alpha, \beta > 0$, we can combine the result from Lemma 3.8 with Hölder inequality to get

$$\begin{aligned} \|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1 &\leq \sqrt{2\|\rho(t,\cdot)\|_1 + 2\|\sigma_{\mu,8\pi}\|_1}\, \|\sqrt{\rho(t,\cdot)} - \sqrt{\sigma_{\mu,8\pi}}\|_2 \\ &\leq 4\sqrt{2\pi}\left(\|\sqrt{\rho(s,\cdot)} - \sqrt{\sigma_{\mu,8\pi}}\|_2^2 + C(t-s)\right)^{1/2} \\ &\leq 4\sqrt{2\pi}\Big(\|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1 + C(t-s)\Big)^{1/2} \end{aligned} \tag{3.17}$$

(recall that both $\rho(t,\cdot)$ and $\sigma_{\mu,8\pi}$ have mass equal to 8π). We may now apply Theorem 1.9 and (3.16) to conclude that for *all* $t > 2$, and all $t_0 < 1$, there is some $\mu = \mu(t) \in [a, 1/a]$ and some $s \in [t - t_0, t]$ such that

$$\begin{aligned} \|\rho(s,\cdot) - \sigma_{\mu,8\pi}\|_1 &\leq C\left(\left(\frac{1}{t_0}\mathcal{H}_{\kappa,8\pi}[\rho]\right)^{1/6} \delta_{\rm HLS}[\rho(s,\cdot)]^{(1-\epsilon)/20} + 1\right)\delta_{\rm HLS}[\rho(s,\cdot)]^{(1-\epsilon)/20} \\ &\leq C\left(\left(\frac{1}{t_0}\mathcal{H}_{\kappa,8\pi}[\rho]\right)^{1/6} t^{-(1-\epsilon)/160} + 1\right) t^{-(1-\epsilon)/160} \end{aligned}$$

Combining this with (3.17), we obtain

$$\|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1 \leq C\left[\left(\left(\frac{1}{t_0}\mathcal{H}_{\kappa,8\pi}[\rho]\right)^{1/6} t^{-(1-\epsilon)/160} + 1\right) t^{-(1-\epsilon)/160} + t_0\right]^{1/2} \ .$$

Now choose $t_0 = t^{-(1-\epsilon)3/80}$ to get

$$\|\rho(t,\cdot) - \sigma_{\mu,8\pi}\|_1 \leq C(1+t)^{-(1-\epsilon)/320} \ . \tag{3.18}$$

Finally, we use the bound on $\mathcal{H}_{\kappa,8\pi}[\rho(t,\cdot)]$ to fix the scale: using again that $|\sqrt{\alpha} - \sqrt{\beta}|^2 \leq |\alpha - \beta|$ for all $\alpha, \beta > 0$, estimate (3.18) gives

$$\|\sqrt{\rho(t,\cdot)} - \sqrt{\sigma_{\mu,8\pi}}\|_2^2 \leq C(1+t)^{-(1-\epsilon)/320} \tag{3.19}$$

By the triangle inequality, (3.7), (3.19), and using that $\sqrt{\sigma_{\mu,8\pi}} \geq \kappa^{1/2}(\kappa + R^2)^{-1}$ inside B_R, we get

$$\begin{aligned} \sqrt{\mathcal{H}_{\kappa,8\pi}[\rho(t,\cdot)]} &\geq \left(\int_{|y|<R} \frac{\sqrt{\rho(t,y)} - \sqrt{\sigma_{\kappa,8\pi}(y)}}{\sqrt{\sigma_{\kappa,8\pi}(y)}}{\rm d}y\right)^{1/2} \\ &\geq \left(\int_{|y|<R} \frac{\sqrt{\sigma_{\mu,8\pi}(y)} - \sqrt{\sigma_{\kappa,8\pi}(y)}}{\sqrt{\sigma_{\kappa,8\pi}(y)}}{\rm d}y\right)^{1/2} - C\sqrt{\frac{\kappa + R^2}{\kappa^{1/2}}}(1+t)^{-(1-\epsilon)/320} \\ &\geq c(\kappa - \mu)^2 \log(1+R) - C\sqrt{\frac{\kappa + R^2}{\kappa^{1/2}}}(1+t)^{-(1-\epsilon)/320} \ , \end{aligned}$$

where in the last line we have used (3.7) and the fact that we have an *a-priori* lower bound on μ.

Thus,

$$(\kappa-\mu)^2 \le \frac{C}{\log(1+R)}\left[\sqrt{\mathcal{H}_{\kappa,8\pi}[\rho(0,\cdot)]}+\sqrt{\frac{\kappa+R^2}{\kappa^{1/2}}(1+t)^{-(1-\epsilon)/320}}\right].$$

Choosing $1+R=(e+t)^{(1-\epsilon)/320}$ we get

$$(\kappa-\mu)^2 \le \frac{C}{\log(e+t)},$$

as desired. Finally (3.11) follows from (3.9) and (3.10), observing that

$$\|\sigma_{\mu,8\pi}-\sigma_{\kappa,8\pi}\|_1 \le C_\kappa|\mu-\kappa| \qquad \forall\,\mu>0\,,$$

with C_κ depending on κ only. □

3.3 Stability for the Logarithmic HLS inequality: proof of Theorem 1.9

We now prove a stability result for the Log-HLS inequality. The proof of Theorem 1.9 is based on the recently discovered fact [9] that $\mathcal{F}$ is decreasing along the fast diffusion flow. Moreover, since the fast diffusion flow is gradient flow for $\mathcal{H}_{\kappa,M}$, also $\mathcal{H}_{\kappa,M}$ is decreasing along the fast diffusion flow. While $\mathcal{D}$ is not decreasing along the flow, the dissipation relation gives us

$$\int_0^T \mathcal{D}[\sigma(t,\cdot)]\mathrm{d}t \le \mathcal{F}[\rho]-C(M)\;, \tag{3.20}$$

where $\sigma(t,x)$ is the solution to (1.18) with initial data $\sigma(0,\cdot)=\rho$. The estimate (3.20) is proved in [9] for initial data ρ such that, for some $C,R>0$, $\rho(x)\le C|x|^{-4}$ for all $|x|>R$. Then, regularity estimates from [8] permit one to integrate by parts and prove that $\lim_{t\to\infty}\mathcal{F}[\rho(t,\cdot)]=C(M)$, which leads to (3.20). However, the regularity provided by [8] is only used in a qualitative way, and the values of R and C do not matter. Hence, a simple truncation and replacement argument can be used to achieve these bounds while making an arbitrarily small effect on $\mathcal{F}[\rho]$ and $\mathcal{H}_\kappa[\rho]$, and moving ρ an arbitrarily small distance in the L^1 norm. So, we may freely assume the bound $\rho(x)\le C|x|^{-4}$ for all $|x|>R$ for some finite constants C and R.

Proof of Theorem 1.9: Let $\sigma(t,\cdot)$ be the solution of (1.18) with initial data ρ. As it is immediately checked, the condition $\int_{\mathbb{R}^2}x\rho(x)\mathrm{d}x=0$ is preserved in time, i.e.,

$$\int_{\mathbb{R}^2} x\sigma(t,x)\mathrm{d}x = 0. \tag{3.21}$$

In addition, as explained above the dissipation relation (3.20) holds, therefore

$$\delta_{\mathrm{HLS}}[\rho] \ge \int_0^T \mathcal{D}[\sigma(t,\cdot)]\mathrm{d}t \qquad \forall\,T>0\;. \tag{3.22}$$

In the proof that follows, C denotes a constant depending at most on only on the quantities ϵ, M, κ, $B_{\mathcal{H}}$, $B_{\mathcal{F}}$ and B_q specified in Theorem 1.9, but not on $B_{\mathcal{D}}$, and changing from step to step, as the proof proceeds.

• *Step 1: The HLS deficit of ρ controls the GNS deficit of $\sigma(t,\cdot)$ for some t close to* 0. Pick some $T \in (0,1]$, with $\delta_{\rm HLS}[\rho] \ll T \ll 1$, to be chosen later. Then by (3.22), there exists some $t \in (0,T)$ such that

$$\mathcal{D}[\sigma(t,\cdot)] \le \frac{\delta_{\rm HLS}[\rho]}{T} \tag{3.23}$$

Since $\mathcal{H}_{\kappa,M}[\sigma(t,\cdot)]$ is decreasing along the flow, Theorem 3.4 gives us a lower bound on $\|\sigma(t,\cdot)\|_{3/2}$. Then, by (1.28), there is a constant $C > 0$ such that

$$\mathcal{D}[\sigma(t,\cdot)] \ge C\delta_{\rm GNS}[\sigma^{1/4}(t,\cdot)] \ .$$

Hence, by (3.23) we get

$$\delta_{\rm GNS}[\sigma^{1/4}(t,\cdot)] \le C\frac{\delta_{\rm HLS}[\rho]}{T} \ .$$

• *Step 2: Application of stability for the GNS inequality.* Recalling (3.21) and that v_λ^4 is a multiple of $\sigma_{\mu,M}$ for some μ, by Theorem 1.4 and Step 1, there exists some $\mu > 0$ (on which we have *a-priori* bounds above and below) such that

$$\|\sigma(t,\cdot) - \sigma_{\mu,M}\|_1 \le C\left(\frac{\delta_{\rm HLS}[\rho]}{T}\right)^{(p-1)/4p} ,$$

and by the triangle inequality,

$$\|\rho - \sigma_{\mu,M}\|_1 \le C\|\rho - \sigma(t,\cdot)\|_1 + C\left(\frac{\delta_{\rm HLS}[\rho]}{T}\right)^{(p-1)/4p} . \tag{3.24}$$

• *Step 3: Controlling $\|\rho - \sigma(t,\cdot)\|_1$.* We claim that for all $\epsilon > 0$ and all $1 < p < 2$, there is a constant C such that

$$\|\rho - \sigma(t,\cdot)\|_1 \le C\left(1 + B_{\mathcal{D}}^{p/4(p+1)} + \left(\frac{\delta_{\rm HLS}[\rho]}{T}\right)^{p/4(p+1)}\right) T^{p(1-\epsilon)/4(p+1)} + CT^{p(1-\epsilon)/8(p+1)} . \tag{3.25}$$

This will be proven below. Assuming this for now, we complete the proof in the next step.
• *Step 4: Optimizing in T.* Combining (3.24) and (3.25), we get that for all $T > \delta_{\rm HLS}[\rho]$,

$$\|\rho - \sigma_{\mu,M}\|_1 \le C\left(\frac{\delta_{\rm HLS}[\rho]}{T}\right)^{(p-1)/4p} + CB_{\mathcal{D}}^{p/4(p+1)}T^{p(1-\epsilon)/4(p+1)} + CT^{p(1-\epsilon)/8(p+1)}.$$

Since $p < 2$ and $B_{\mathcal{D}} \ge 1$, we can bound $B_{\mathcal{D}}^{p/4(p+1)}$ with $B_{\mathcal{D}}^{1/6}$. Then, setting $r = \dfrac{p(1-\epsilon)}{8(p+1)}$ and $s = \dfrac{p-1}{4p}$, we choose $T := \delta_{\rm HLS}[\rho]^{s/(r+s)}$ to obtain

$$\|\rho - \sigma_{\mu,M}\|_1 \le C\left[1 + B_{\mathcal{D}}^{1/6}\delta_{\rm HLS}[\rho]^{rs/(r+s)}\right]\delta_{\rm HLS}[\rho]^{rs/(r+s)}.$$

Since p can be chosen arbitrarily close to 2, and $\epsilon > 0$ is arbitrarily small, we obtain the result. □

We close this subsection by proving (3.25). To this aim, we make use of the fact that, for each $\kappa > 0$, the equation (1.18) is a gradient flow of the functional $\mathcal{H}_{\kappa,M}$, with respect to the

2-Wasserstein metric W_2, on the space of densities of mass M. This has the standard consequence that

$$\mathrm{W}_2^2(\sigma(s,\cdot),\sigma(t,\cdot)) \le \mathcal{H}_{\kappa,M}[\rho](t-s) \tag{3.26}$$

for all $t > s \ge 0$, see for instance [2] or also [6, Lemma 5.3]. That is, the fact that the equation is gradient flow for W_2 automatically yields a Hölder$-1/2$ modulus of continuity bound in this metric. What we need now is to improve this bound into a L^1 continuity.

In the proof, we use (3.26) together with the following interpolation result, see [6, Theorem 5.11]:

3.9 THEOREM (Interpolation bound)**.** *Let σ_0 and σ_1 be two densities of mass M on $\mathbb{R}^2$ such that for some $q > 2$, $\|\sigma_0\|_{q+1}^{q+1}$, $\|\sigma_1\|_{q+1}^{q+1} \le K$. Suppose also that $\sigma_0^{1/4}$ and $\sigma_1^{1/4}$ have square integrable distributional gradients. Then*

$$\begin{aligned}\|\sigma_0-\sigma_1\|_2^2 \;&\le\; \left(\|\nabla(\sigma_0^{1/4})\|_2 + \|\nabla(\sigma_1^{1/4})\|_2\right)(2^{5/2}+2^{9/2}K)(\mathrm{W}_2(\sigma_0,\sigma_1))^{(4q-3)/(4q+2)} \\ &+\; 16M^{(q-1)/q}K^{(q+2)/2q}(\mathrm{W}_2(\sigma_0,\sigma_1))^{(q-1)/(2q+1)}\ .\end{aligned}$$

Proof of (3.25)*:* We apply the interpolation bound quoted above with $\sigma_0 = \rho$ and $\sigma_1 = \sigma(t,\cdot)$. By Theorem 3.2 and (3.23), we have

$$\|\nabla\sigma_0^{1/4}\|_2^2 + \|\nabla\sigma_1^{1/4}\|_2^2 \le C\left(1 + B_{\mathcal{D}} + \frac{\delta_{\mathrm{HLS}}[\rho]}{T}\right)\ ,$$

where C depends only on κ, $B_{\mathcal{F}}$, and $B_{\mathcal{H}}$. Moreover, since all L^p norms are propagated along the evolution locally in time, any bound on some L^q norm of $\sigma_0 = \rho$ is valid also for σ_0 (up to universal multiplicative constants). Furthermore, by (3.26), $\mathrm{W}_2(\sigma_0,\sigma_1) \le C\sqrt{T}$. Pick $\epsilon > 0$, and choose $q = q(\epsilon)$ so large that

$$\frac{4q-3}{4q+2} > 1-\epsilon \qquad \text{and} \qquad \frac{q-1}{2q+1} > \frac{1-\epsilon}{2}\ .$$

We then have

$$\|\sigma_0-\sigma_1\|_2 \le C\left(1 + B_{\mathcal{D}}^{1/4} + \left(\frac{\delta_{\mathrm{HLS}}[\rho]}{T}\right)^{1/4}\right)T^{(1-\epsilon)/4} + CT^{(1-\epsilon)/8}\ .$$

Next, for non-negative functions f on $\mathbb{R}^2$ and all $R > 0$, we estimate

$$\begin{aligned}\|f\|_1 \;&=\; \int_{|x|\le R} f(x)\mathrm{d}x + \int_{|x|\ge R} f(x)\mathrm{d}x \\ &\le\; (\pi R^2)^{1/2}\|f\|_2 + \frac{1}{R^p}\int_{\mathbb{R}^2}|x|^p f(x)\mathrm{d}x\ .\end{aligned}$$

Optimizing in R yields

$$\|f\|_1 \le C\|f\|_2^{p/(p+1)}\left(\int_{\mathbb{R}^2}|x|^p|f(x)|\mathrm{d}x\right)^{1/(1+p)}\ ,$$

Applying this, we finally obtain

$$\|\sigma_0-\sigma_1\|_1 \le C\left(1 + B_{\mathcal{D}}^{p/4(p+1)} + \left(\frac{\delta_{\mathrm{HLS}}[\rho]}{T}\right)^{p/4(p+1)}\right)T^{p(1-\epsilon)/4(p+1)} + CT^{p(1-\epsilon)/8(p+1)}\ .$$

□

3.4 Continuity properties of the Log-HLS functional: proof of Theorem 3.7

In order to prove Theorem 3.7, we begin with a continuity result for the entropy that is of interest in its own right. In this section, p, q, A and B are as in Definition 3.6.

The next result states the almost Lipschitz continuity of the entropic part of $\mathcal{F}$, and is not restricted to dimension two.

3.10 THEOREM. *There is a constant C, depending only on n, p, q, A and B such that, for any $\rho, \sigma \in \mathcal{M}_{n,p,q,A,B}$ with $\|\rho - \sigma\|_1 \leq 1/e$,*

$$\left| \int_{\mathbb{R}^n} \rho \log \rho(x) \mathrm{d}x - \int_{\mathbb{R}^n} \sigma \log \sigma(x) \mathrm{d}x \right| \leq C \|\rho - \sigma\|_1 \big| \log \|\rho - \sigma\|_1 \big| \ .$$

We first prove two lemmas.

3.11 LEMMA. *For all $0 < s < p$ there is a constant C, depending only on s, n, p, q, A and B, such that*

$$\int_{|x|>R} \rho |\log \rho|(x) \mathrm{d}x \leq C R^{-s} \ ,$$

for all $\rho \in \mathcal{M}_{n,p,q,A,B}$ and all $R > 0$.

Proof. We begin by recalling the following elementary inequality: for all $r > 0$,

$$|\log s| \leq \frac{1}{r} \max\{s^r, s^{-r}\} \qquad \forall\, s > 0\,. \tag{3.27}$$

Now, pick $0 < \gamma < 1$ and set

$$r := (q-1)\gamma > 0\ , \qquad q = \frac{p}{p-1}\ .$$

We claim that

$$\int_{\mathbb{R}^n} |x|^{p(1-\gamma)} \rho^{1+r}(x) \mathrm{d}x \leq \left(\int_{\mathbb{R}^n} |x|^p \rho(x) \mathrm{d}x \right)^{1-\gamma} \left(\int_{\mathbb{R}^n} |\rho(x)|^q \mathrm{d}x \right)^{\gamma} \ . \tag{3.28}$$

To see this, define $\beta = q\gamma$ and note that $\beta + (1-\gamma) = 1 + r$. Thus by Hölder's inequality,

$$\begin{aligned} \int_{\mathbb{R}^n} |x|^{p(1-\gamma)} \rho^{1+r}(x) \mathrm{d}x &= \int_{\mathbb{R}^n} |x|^{p(1-\gamma)} \rho^{1-\gamma}(x) \rho^{\beta}(x) \mathrm{d}x \\ &\leq \left(\int_{\mathbb{R}^n} (|x|^{p(1-\gamma)} \rho^{1-\gamma}(x))^{1/(1-\gamma)} \mathrm{d}x \right)^{1-\gamma} \left(\int_{\mathbb{R}^n} (\rho^{\beta}(x))^{1/\gamma} \mathrm{d}x \right)^{\gamma}\ , \end{aligned}$$

from which the claim follows.

Thus, under the conditions (3.12), with γ and r chosen as above we have a uniform bound on $\int_{\mathbb{R}^n} |x|^{p(1-\gamma)} \rho^{1+r}(x) \mathrm{d}x$. Hence, using (3.27) with $s = \rho(x)$ we get

$$\begin{aligned} \int_{\{|x| \geq R\} \cap \{\rho \geq 1\}} \rho |\log \rho|(x) \mathrm{d}x &\leq \frac{1}{r} \int_{|x|>R} \rho^{1+r}(x) \mathrm{d}x \\ &= \left(\frac{1}{r} \int_{|x|>R} |x|^{p(1-\gamma)} \rho^{1+r}(x) \mathrm{d}x \right) R^{-p(1-\gamma)}. \end{aligned} \tag{3.29}$$

Next, we want to consider the set $\{\rho \leq 1\}$. Pick $0 < \delta < 1$ and set

$$\alpha := p(1-\delta) \ .$$

We shall require $\alpha > n\delta$ (or equivalently $p/n > \delta/(1-\delta)$), which is always satisfied for δ sufficiently small.

Then, by (3.27) (with $r = \delta$ and $s = \rho(x)$) and Hölder's inequality, for all $\alpha > n\delta$ we have

$$\begin{aligned}
\int_{\{|x|\geq R\}\cap\{\rho\leq 1\}} \rho|\log\rho|(x)\mathrm{d}x &\leq \frac{1}{\delta}\int_{|x|>R} \rho^{1-\delta}(x)\mathrm{d}x \\
&= \frac{1}{\delta}\int_{|x|>R} |x|^{\alpha}\rho^{1-\delta}(x)|x|^{-\alpha}\mathrm{d}x \\
&\leq \frac{1}{\delta}\left(\int_{\mathbb{R}^n} |x|^{\alpha/(1-\delta)}\rho(x)\mathrm{d}x\right)^{1-\delta}\left(\int_{|x|>R}|x|^{-\alpha/\delta}\mathrm{d}x\right)^{\delta} \\
&= \frac{1}{\delta}\left(\int_{\mathbb{R}^n} |x|^{p}\rho(x)\mathrm{d}x\right)^{1-\delta}\left(\int_{|x|>R}|x|^{-\alpha/\delta}\mathrm{d}x\right)^{\delta} .
\end{aligned}$$

Computing

$$\int_{|x|>R}|x|^{-\alpha/\delta}\mathrm{d}x = \frac{n|B_1^n|\delta}{\alpha-n\delta}R^{-(\alpha-n\delta)/\delta}$$

and recalling the definition of α, we obtain

$$\int_{\{|x|\geq R\}\cap\{\rho\leq 1\}} \rho|\log\rho|(x)\mathrm{d}x \leq \frac{1}{\delta}\left(\int_{\mathbb{R}^n} |x|^{p}\rho(x)\mathrm{d}x\right)^{1-\delta}\left(\frac{n|B_1^n|r}{\alpha-n\delta}\right)^{\delta} R^{-(p(1-\delta)-n\delta)} .$$

(Here and in the sequel, $|B_1^n|$ denotes the Lebesgue measure of the unit ball in $\mathbb{R}^n$.) Combining this bound with (3.29) and choosing both γ and δ sufficiently small, we have the result. □

3.12 LEMMA. *For all* $0 < t < q$, $\rho \in \mathcal{M}_{n,p,q,A,B}$, $R > 0$, *and* $0 < \epsilon < 1/e$, *it holds*

$$\left|\int_{|x|\leq R}\rho\log\rho(x)\mathrm{d}x - \int_{|x|\leq R}\sigma\log\sigma(x)\mathrm{d}x\right| \leq$$
$$2|\log\epsilon|\left[2\epsilon|B_1^n|R^n + \frac{2}{t-1}\epsilon^{q-t}(|\rho\|_q^q + \|\sigma\|_q^q) + \|\rho-\sigma\|_1\right] \ . \quad (3.30)$$

Proof. Pick ϵ with $0 < \epsilon < 1/e$ and define

$$\rho_\epsilon(x) := \begin{cases} \epsilon & \text{if } \rho(x) \leq \epsilon \\ \rho(x) & \text{if } \epsilon < \rho(x) \leq 1/\epsilon \\ 1/\epsilon & \text{if } \rho(x) \geq 1/\epsilon \end{cases} .$$

Note that $s \mapsto s\log s$ is decreasing on $(0, 1/e)$, so

$$\big|s\log s - \epsilon\log\epsilon\big| \leq \epsilon\big|\log\epsilon\big| \qquad \forall\, s \in (0,\epsilon]\,.$$

Therefore,

$$\left| \int_{\{|x|\le R\}\cap\{\rho\le\epsilon\}} \rho \log \rho(x) \mathrm{d}x - \int_{\{|x|\le R\}\cap\{\rho\le\epsilon\}} \rho_\epsilon \log \rho_\epsilon(x) \mathrm{d}x \right| \le \epsilon |\log \epsilon| |B_1^n| R^n \ .$$

In an analogous way,

$$\left| \int_{\{|x|\le R\}\cap\{\rho\le\epsilon\}} \rho(x) \mathrm{d}x - \int_{\{|x|\le R\}\cap\{\rho\le\epsilon\}} \rho_\epsilon(x) \mathrm{d}x \right| \le \epsilon |B_1^n| R^n \ .$$

Next, by Chebychev's inequality, $|\{\rho \ge 1/\epsilon\}| \le \epsilon^q \|\rho\|_q^q$. Thus, for any $1 < t < q$, applying (3.27) with $r = t-1$ we get

$$\begin{aligned} \int_{\{|x|\le R\}\cap\{\rho\ge 1/\epsilon\}} \rho \log \rho(x) \mathrm{d}x &\le \frac{1}{t-1} \int_{\{|x|\le R\}\cap\{\rho\ge 1/\epsilon\}} \rho^t(x) \mathrm{d}x \\ &\le \frac{1}{t-1} \|\rho\|_q^t \left(|\{\rho \ge 1/\epsilon\}| \| \right)^{(q-t)/q} \\ &\le \frac{1}{t-1} \epsilon^{q-t} \|\rho\|_q^q \ . \end{aligned}$$

Hence,

$$\left| \int_{\{|x|\le R\}\cap\{\rho\ge 1/\epsilon\}} \rho \log \rho(x) \mathrm{d}x - \int_{\{|x|\le R\}\cap\{\rho\ge 1/\epsilon\}} \rho_\epsilon \log \rho_\epsilon(x) \mathrm{d}x \right| \le \frac{1}{t-1} \epsilon^{q-t} \|\rho\|_q^q \ .$$

In a similar way,

$$\left| \int_{\{|x|\le R\}\cap\{\rho\ge 1/\epsilon\}} \rho(x) \mathrm{d}x - \int_{\{|x|\le R\}\cap\{\rho\ge 1/\epsilon\}} \rho_\epsilon(x) \mathrm{d}x \right| \le \epsilon^{q-1} \|\rho\|_q^q \ .$$

Thus, combining all these estimates together, we have

$$\left| \int_{\{|x|\le R\}} \rho \log \rho(x) \mathrm{d}x - \int_{\{|x|\le R\}} \rho_\epsilon \log \rho_\epsilon(x) \mathrm{d}x \right| \le \epsilon |\log \epsilon| |B_1^n| R^n + \frac{1}{t-1} \epsilon^{q-t} \|\rho\|_q^q \ , \tag{3.31}$$

and

$$\int_{|x|\le R} |\rho - \rho_\epsilon| \mathrm{d}x \le \epsilon |B_1^n| R^n + \epsilon^{q-1} \|\rho\|_q^q \ . \tag{3.32}$$

Of course, we have the analogous estimates for σ.

Next, we observe that the derivative of $s \mapsto s \log s$ on $[\epsilon, 1/\epsilon]$ is bounded by $2|\log \epsilon|$ (recall that $\epsilon \le 1/e$). Hence, since ρ_ϵ and σ_ϵ are bounded below by ϵ and above by $1/\epsilon$,

$$|\rho_\epsilon \log \rho_\epsilon(x) - \sigma_\epsilon \log \sigma_\epsilon(x)| \le 2 |\log \epsilon| |\rho_\epsilon(x) - \sigma_\epsilon(x)| \qquad \forall\, x \in \mathbb{R}^n \ .$$

Integrating over $\{|x| \le R\}$ we find

$$\int_{\{|x|\le R\}} |\rho_\epsilon \log \rho_\epsilon(x) - \sigma_\epsilon \log \sigma_\epsilon(x)| \le 2 |\log \epsilon| \int_{\{|x|\le R\}} |\rho_\epsilon(x) - \sigma_\epsilon(x)| \mathrm{d}x \ .$$

Combining this with (3.31), (3.32), and the corresponding estimates for σ, we obtain (3.30). □

Proof of Theorem 3.10: Combining the last two lemmas, for any $s \in (0,p)$, $t \in (1,q)$ and $R > 0$, we have

$$\int_{\mathbb{R}^n} |\rho \log \rho(x) - \sigma \log \sigma(x)| \leq CR^{-s} + 2|\log \epsilon| \left[2\epsilon |B_1^n| R^n + \frac{2}{t-1} \epsilon^{q-t} (|\rho\|_q^q + \|\sigma\|_q^q) + \|\rho - \sigma\|_1 \right] .$$

Choosing $R = (\epsilon|\log \epsilon|)^{-1/(n+s)}$, and recalling that we can take s close to p and t close to 1, we obtain

$$\int_{\mathbb{R}^n} |\rho \log \rho(x) - \sigma \log \sigma(x)| \leq C\epsilon^m + |\log \epsilon| \|\rho - \sigma\|_1 \ ,$$

for any $0 < m < \min \left\{ \frac{p}{n+p} \ , \ q - 1 \right\}$. Assuming without loos of generality $m < 1$, we can choose $\epsilon := \|\rho - \sigma\|_1^{1/m}$ (recall that by assumption $\|\rho - \sigma\|_1 \leq 1/e$) to conclude the proof. □

In the rest of this section we are concerned only with $n = 2$. Given a mass density ρ on $\mathbb{R}^2$ such that

$$\int_{\mathbb{R}^2} \log(e + |x|^2)\rho(x)\mathrm{d}x < \infty\,, \tag{3.33}$$

the *Newtonian potential energy* of ρ is given by

$$\mathcal{U}[\rho] := \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \rho(x) \log |x - y| \rho(y) \mathrm{d}x\mathrm{d}y \ . \tag{3.34}$$

By the elementary inequality

$$\log_+ |x - y|_+ \leq \log 2 + \log_+ |x| + \log_+ |y| \ ,$$

(recall that $\log_+$ denotes the positive part of log), the condition (3.33) ensures that the integral in (3.34) is well-defined, though possibly with the value $-\infty$.

3.13 LEMMA. *With p, q, A and B as in Definition 3.6, for all $0 < \epsilon < 1$, there is an explicitly computable constant C depending only on ϵ, p, q, A and B such that, for any $\rho, \sigma \in \mathcal{M}_{2,p,q,A,B}$ with $\|\rho - \sigma\|_1 \leq 1$,*

$$|\mathcal{U}[\rho] - \mathcal{U}[\sigma]| \leq C\|\rho - \sigma\|_1^{1-\epsilon} \ . \ .$$

Proof. We define

$$\mathcal{U}_+[\rho] := \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \rho(x) \log_+ |x - y| \rho(y) \mathrm{d}x\mathrm{d}y \quad \text{and} \quad \mathcal{U}_-[\rho] := \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \rho(x) \log_- |x - y| \rho(y) \mathrm{d}x\mathrm{d}y \ .$$

Then using $\log_+ |x| \leq (1/r)|x|^r$ (see (3.27)) and likewise for y, we obtain

$$\begin{aligned} |\mathcal{U}_+[\rho] - \mathcal{U}_+[\sigma]| &\leq \int_{\mathbb{R}^2} |\rho(x) - \sigma(x)| \left(\int_{\mathbb{R}^2} (2 + \log_+ |x| + \log_+ |y|)[\rho(y) + \sigma(y)]\mathrm{d}y \right) \mathrm{d}x \\ &\leq C \int_{\mathbb{R}^2} |\rho(x) - \sigma(x)| \left(1 + \frac{1}{r}|x|^r \right) \mathrm{d}x \ . \end{aligned}$$

Hence, if $r \in (0,p)$, using Hölder's inequality we can estimate

$$\int_{\mathbb{R}^2} |\rho(x) - \sigma(x)|^{(p-r)/p} |\rho(x) - \sigma(x)|^{r/p} |x|^r \mathrm{d}x \leq \|\rho - \sigma\|_1^{(p-r)/p} \left(\int_{\mathbb{R}^2} [\rho(x)|x|^p + \sigma(x)|x|^p \mathrm{d}x \right)^{r/p} \ .$$

Choosing $r = \epsilon p$, we get

$$|\mathcal{U}_+[\rho] - \mathcal{U}_+[\sigma]| \leq C\|\rho - \sigma\|_1^{1-\epsilon} \ , \tag{3.35}$$

for some constant C depending only on ϵ, p, A and B.

Next, for all $0 < r < 2(q-1)/q$, by (3.27)

$$\begin{aligned}
|\mathcal{U}_-[\rho] - \mathcal{U}_-[\sigma]| &\leq \int_{\mathbb{R}^2} |\rho(x) - \sigma(x)| \left(\int_{\mathbb{R}^2} \log_- |x-y|[\rho(y) + \sigma(y)]\mathrm{d}y\right) \mathrm{d}x \\
&\leq \int_{\mathbb{R}^2} |\rho(x) - \sigma(x)| \left(\int_{\{|x-y|\leq 1\}} \frac{1}{r}|x-y|^{-r}[\rho(y) + \sigma(y)]\mathrm{d}y\right) \mathrm{d}x \\
&\leq \int_{\mathbb{R}^2} |\rho(x) - \sigma(x)|[\|\rho\|_q + \|\sigma\|_q] \left(\int_{\{|y|\leq 1\}} \frac{1}{r}|y|^{-rq/(q-1)}\mathrm{d}y\right)^{(q-1)/q} \mathrm{d}x \\
&= \|\rho - \sigma\|_1[\|\rho\|_q + \|\sigma\|_q] \left(\int_{\{|y|\leq 1\}} \frac{1}{r}|y|^{-rq/(q-1)}\mathrm{d}y\right)^{(q-1)/q} .
\end{aligned}$$

The integral on the right is clearly finite for our choice of r, and we conclude that

$$|\mathcal{U}_-[\rho] - \mathcal{U}_-[\sigma]| \leq C\|\rho - \sigma\|_1 \ , \tag{3.36}$$

for some C depending only on q and B. Combining (3.35) and (3.36) we obtain the result. □

Proof of Theorem 3.7: The theorem follows directly from the results proved in this subsection. □

3.5 Proof of Lemma 3.8

We begin with a formal calculation: To simplify the notation, let ρ denote a solution of the Keller-Segel equation as in the statement of the lemma, and let σ denote $\sigma_{\kappa,8\pi}$. Moreover, whenever we say that ρ (or quantities related to ρ) are bounded in some L^p space, we mean that the bound is uniform in time.

Recall that, since we are considering the solution ρ on the time interval $[\tau, +\infty)$ with $\tau > 0$, ρ is bounded in all L^q spaces for $q \in [1, \infty)$, and it has finite pth moments for all $p < 2$ (see the discussion at the beginning of the proof of Theorem 3.5).

We compute

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^2} |\sqrt{\rho} - \sqrt{\sigma}|^2 &= \int_{\mathbb{R}^2} (\sqrt{\rho} - \sqrt{\sigma}) \frac{1}{\sqrt{\rho}} \frac{\partial}{\partial t}\rho \\
&= -\int_{\mathbb{R}^2} \frac{\sqrt{\sigma}}{\sqrt{\rho}} \Delta\rho - \int_{\mathbb{R}^2} \frac{\sqrt{\sigma}}{\sqrt{\rho}} \mathrm{div}(\rho\nabla\Delta^{-1}\rho) \\
&= -\frac{1}{2}\int_{\mathbb{R}^2} \sqrt{\sigma}\frac{|\nabla\rho|^2}{\rho^{3/2}} + \int_{\mathbb{R}^2} \nabla\sqrt{\sigma} \cdot \frac{\nabla\rho}{\sqrt{\rho}} \\
&\quad + \int_{\mathbb{R}^2} \sqrt{\rho}\nabla\sqrt{\sigma} \cdot \nabla\Delta^{-1}\rho - \frac{1}{2}\int_{\mathbb{R}^2} \sqrt{\sigma}\frac{\nabla\rho}{\sqrt{\rho}} \cdot \nabla\Delta^{-1}\rho
\end{aligned}$$

Now, if we denote by $D := \int \sqrt{\sigma}\frac{|\nabla\rho|^2}{\rho^{3/2}}$, we have that the first term is $-D/2$, and using Hölder inequality the second is bounded by

$$\sqrt{D}\sqrt{\int_{\mathbb{R}^2} \frac{|\nabla\sqrt{\sigma}|^2}{\sqrt{\sigma}}\sqrt{\rho}} = 2\sqrt{D}\sqrt{\int_{\mathbb{R}^2} |\nabla\sigma^{1/4}|^2\sqrt{\rho}} \leq C\sqrt{D}$$

since both $|\nabla\sigma^{1/4}|^2$ and $\sqrt{\rho}$ are bounded in L^2. For the third and the fourth, we observe that since $\rho \in L^q$ for all $q > 1$, and the operator $\frac{\partial^2}{\partial x_i \partial x_j}\Delta^{-1}$ is bounded from $L^s(\mathbb{R}^2)$ into itself for all $s \in (1,\infty)$, we have that

$$\frac{\partial^2}{\partial x_i \partial x_j}\Delta^{-1}\rho \in L^q(\mathbb{R}^2) \qquad \forall\, q \in (1,\infty),$$

for each i, j, which by Sobolev embedding implies

$$\nabla\Delta^{-1}\rho \in L^q(\mathbb{R}^2) \qquad \forall\, q \in (2,\infty). \tag{3.37}$$

Hence a simple Hölder inequality argument allows us to bound the third term with

$$\|\rho\|_1^{1/2}\|\nabla\sqrt{\sigma}\|_4^{1/4}\|\nabla\Delta^{-1}\rho\|_4^{1/4} \leq C.$$

Finally, arguing similarly as we did for the second term, we obtain that also the fourth is bounded by $C\sqrt{D}$.

Hence, this implies that we can use the first term to reabsorb both the second and the fourth (up to an additive constant), and we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^2} |\sqrt{\rho} - \sqrt{\sigma}|^2 \leq C\,,$$

from which the bound claimed in Lemma 3.8 follows. Hence, it remains to justify the formal argument.

Proof of Lemma 3.8. To justify the formal computation, we use a two-parameter regularization: we first convolve ρ with a radial convolution kernel η_ϵ supported in B_ϵ, and we set $\rho_\epsilon = \rho * \eta_\epsilon$, and then for $\epsilon, \gamma > 0$ small we define

$$\rho_{\epsilon,\gamma} := (1-\gamma)\rho_\epsilon + \gamma\sigma.$$

Then, since σ is a stationary solution of the Keller-Segel equation, we get

$$\frac{\partial \rho_{\epsilon,\gamma}}{\partial t} = \Delta\rho_{\epsilon,\gamma} + \mathrm{div}\Big((1-\gamma)[\rho\nabla\Delta^{-1}\rho]_\epsilon + \gamma[\sigma\nabla\Delta^{-1}\sigma]\Big),$$

where $[\rho\nabla\Delta^{-1}\rho]_\epsilon := [\rho\nabla\Delta^{-1}\rho] * \eta_\epsilon$. As we did before, we now differentiate $\int |\sqrt{\rho_{\epsilon,\gamma}} - \sqrt{\sigma}|^2$ in time to get

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^2} |\sqrt{\rho_{\epsilon,\gamma}} - \sqrt{\sigma}|^2 &= -\int_{\mathbb{R}^2} \frac{\sqrt{\sigma}}{\sqrt{\rho}_{\epsilon,\gamma}}\Delta\rho_{\epsilon,\gamma} - \int_{\mathbb{R}^2} \frac{\sqrt{\sigma}}{\sqrt{\rho}_{\epsilon,\gamma}}\mathrm{div}\Big((1-\gamma)[\rho\nabla\Delta^{-1}\rho]_\epsilon + \gamma[\sigma\nabla\Delta^{-1}\sigma]\Big) \\
&= -\frac{1}{2}\int_{\mathbb{R}^2} \sqrt{\sigma}\frac{|\nabla\rho_{\epsilon,\gamma}|^2}{\rho_{\epsilon,\gamma}^{3/2}} + \int_{\mathbb{R}^2} \nabla\sqrt{\sigma}\cdot\frac{\nabla\rho_{\epsilon,\gamma}}{\sqrt{\rho}_{\epsilon,\gamma}} \\
&\quad + \int_{\mathbb{R}^2} \frac{1}{\sqrt{\rho_{\epsilon,\gamma}}}\nabla\sqrt{\sigma}\cdot\Big((1-\gamma)[\rho\nabla\Delta^{-1}\rho]_\epsilon + \gamma[\sigma\nabla\Delta^{-1}\sigma]\Big) \\
&\quad -\frac{1}{2}\int_{\mathbb{R}^2} \sqrt{\sigma}\frac{\nabla\rho_{\epsilon,\gamma}}{\rho_{\epsilon,\gamma}^{3/2}}\cdot\Big((1-\gamma)[\rho\nabla\Delta^{-1}\rho]_\epsilon + \gamma[\sigma\nabla\Delta^{-1}\sigma]\Big) \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

We now observe that

$$|I_2| \leq \left| \int_{\mathbb{R}^2} |\nabla(\sigma^{1/4})| \rho_{\epsilon,\gamma}^{1/4} \frac{|\nabla \rho_{\epsilon,\gamma}|}{\rho_{\epsilon,\gamma}^{3/4}} \right| \leq \sqrt{|I_1|} \sqrt{\int_{\mathbb{R}^2} |\nabla(\sigma^{1/4})|^2 \rho_{\epsilon,\gamma}^{1/2}} \leq C\sqrt{|I_1|},$$

so I_2 can be reabsorbed into the negative term I_1. Concerning I_3 we notice that

$$\frac{|\nabla \sqrt{\sigma}|}{\sqrt{\rho_{\epsilon,\gamma}}} \leq \frac{|\nabla \sqrt{\sigma}|}{\gamma \sigma} \leq \frac{C}{\gamma}(1+|x|).$$

Hence, since the oscillation of the function $1+|x|$ inside a ball of radius ϵ (the support of η_ϵ) is ϵ, the first integrand inside I_3 is bounded from above by

$$\frac{C}{\gamma}(1+|x|)[\rho \nabla \Delta^{-1} \rho]_\epsilon \leq \frac{C}{\gamma}[(1+|x|)\rho \nabla \Delta^{-1}\rho]_\epsilon + \frac{C}{\gamma}\epsilon[\rho|\nabla \Delta^{-1}\rho|]_\epsilon.$$

Now, since ρ has finite pth moments for any $p < 2$, and both ρ and $\nabla \Delta^{-1}\rho$ belong to all L^q spaces for $q \geq 2$ (see (3.37)), by Hölder inequality $(1+|x|)\rho \nabla \Delta^{-1}\rho$ belongs to $L^s(\mathbb{R}^2)$ for all $s \in [1, 3/2]$. Thus the quantity $[(1+|x|)\rho \nabla \Delta^{-1}\rho]_\epsilon$ is pointwise controlled by its maximal function, which also belongs to $L^s(\mathbb{R}^2)$ for all $s \in [1, 3/2]$. Since both terms $\frac{C}{\gamma}\epsilon[\rho|\nabla \Delta^{-1}\rho|]_\epsilon$ and $\gamma[\sigma \nabla \Delta^{-1}\sigma]$ are easy to control, we have proved that, for $\gamma > 0$ fixed, the integrand inside I_3 is dominated by an integrable function, uniformly with respect to $\epsilon > 0$. Hence we can let $\epsilon \to 0$ to obtain that

$$\begin{aligned} \lim_{\epsilon \to 0} I_3 &= \int_{\mathbb{R}^2} \nabla \sqrt{\sigma} \cdot \Big(\frac{(1-\gamma)\rho \nabla \Delta^{-1}\rho}{\sqrt{(1-\gamma)\rho + \gamma\sigma}} + \frac{\gamma[\sigma \nabla \Delta^{-1}\sigma]}{\sqrt{(1-\gamma)\rho + \gamma\sigma}} \Big) \\ &\leq \int_{\mathbb{R}^2} |\nabla \sqrt{\sigma}| \left(\sqrt{1-\gamma}\sqrt{\rho}|\nabla \Delta^{-1}\rho| + \sqrt{\gamma}\sqrt{\sigma}|\Delta^{-1}\sigma| \right) \leq C\,. \end{aligned}$$

Finally, a similar argument can be used to estimate I_4 with $C\sqrt{|I_1|}$.

Hence, by taking first the limit as $\epsilon \to 0$ and then as $\gamma \to 0$ we have rigorously justified the previous formal computation. □

References

[1] Aubin, Th (1976) Problémes isoperimétriques et espaces de Sobolev. J. Diff. Geom. 11: 573–598.

[2] Ambrosio L, Gigli N, Savaré G (2005) Gradient flows in metric spaces and in the space of probability measures. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel.

[3] Barky D, Gentil Y, Ledoux M, Analysis and geometry of diffusion semigroups. Monograph in preparation.

[4] Bianchi G, Egnell H (1991) A note on the Sobolev inequality, J. Funct. Anal., 100: 18–24.

[5] Blanchet A, Bonforte M, Dolbeault J, Grillo G, Vázquez JL (2009) Asymptotic of the fast diffusion equation via entropy estimates, Arch. Ration. Mech. Anal., 191: 347–385.

[6] Blanchet A, Carlen EA, Carrillo JA (2010) Functional inequalities, thick tails and asymptotics for the critical mass Patlak-Keller-Segel model, to appear in Jour. Funct. Anal.

[7] Beckner W (1993) Sharp Sobolev inequalities on the sphere and the Moser-Trudinger inequality, Ann. of Math. 2, 138: 213–242.

[8] Bonforte M, Vázquez JL (2006) Global positivity estimates and Harnack inequalities for the fast diffusion equation, Jour. Funct. Anal., 240: 399–428.

[9] Carlen EA, Carrilo JA, Loss M (2010) Hardy-Littlewood-Sobolev Inequalities via Fast Diffusion Flows, P.N.A.S. 107 (46): 19696–19701

[10] Carlen EA, Loss M (1992) Competing symmetries, the logarithmic HLS inequality and Onofri's inequality on S^n, Geom. Funct. Anal., 2: 90–104.

[11] Carlen EA, Loss M. (1993) Sharp constant in Nash's inequality. Internat. Math. Res. Notices, 7: 213–215

[12] Carrillo JA, Vázquez JL (2003) Fine asymptotics for fast diffusion equations, Comm. Partial Differential Equations, 28: 1023–1056.

[13] Cianchi A, Fusco N, Maggi F, Pratelli A (2009) The sharp Sobolev inequality in quantitative form. J. Eur. Math. Soc. (JEMS) 11: 1105–1139.

[14] Cordero -Erausqauin D, Nazaret B, Villani C (2004) A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities, Adv. Math. 182 (2): 307–332.

[15] Del Pino M, Dolbeault J (2002) Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions, J. Math. Pures Appl., 81: 301–342.

[16] Dolbeault J, Perthame B (2004) Optimal critical mass in the two-dimensional Keller- Segel model in $\mathbb{R}^2$, C. R. Math. Acad. Sci. Paris, 339: 611–616.

[17] Figalli A, Maggi F, Pratelli A (2010) A mass transportation approach to quantitative isoperimetric inequalities. Invent. Math., 182: 167–211.

[18] Figalli A, Maggi F, Pratelli A, Sharp stability theorems for the anisotropic Sobolev and log-Sobolev inequalities on functions of bounded variation. Preprint.

[19] Fusco N, Maggi F, Pratelli A (2008) The sharp quantitative isoperimetric inequality. Ann. of Math., 168: 941–980.

[20] Fusco N, Maggi F, Pratelli A (2007) The sharp quantitative Sobolev inequality for functions of bounded variation. J. Funct. Anal., 244: 315–341.

[21] Herrero MA, Pierre M (1985) The Cauchy problem for $u_t = \Delta u^m$ when $0 < m < 1$, Trans. Amer. Math. Soc., 291: 145–158.

[22] Lieb EH (1983) Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities, Ann. of Math., 118: 349–374.

[23] Lieb EH, Loss M (1997) Analysis. Graduate Studies in Mathematics, 14. American Mathematical Society, Providence, RI.

[24] McCann RJ (1997) A convexity principle for interacting gases, Adv. Math., 128: 153–179.

[25] Otto F (2001) The geometry of dissipative evolution equations: the porous medium equation, Comm. Partial Differential Equations, 26: 101–174.

[26] Rosen G (1971) Minimum value for c in the Sobolev inequality $\|\phi^3\| \le c\|\nabla\phi\|^3$. SIAM J. Appl. Math. 21 (1971), 30-32.

[27] Talenti G (1976) Best constants in Sobolev inequality. Ann. Mat. Pura Appl. 110 , 353-372.

[28] Vázquez JL (2007) The Porous Medium Equation. Mathematical theory, Oxford Mathematical Monographs, The Clarendon Press/Oxford University Press, Oxford/New York.

[29] Vázquez JL (2003) Asymptotic behaviour for the porous medium equation posed in the whole space, J. Evol. Equations, 3: 67–118.