AFST

Annales de la Faculté des Sciences de Toulouse

MATHÉMATIQUES

Matthew Rosenzweig and Sylvia Serfaty

Modulated logarithmic Sobolev inequalities and generation of chaos

Tome XXXIV, nº 1 (2025), p. 107–134.

https://doi.org/10.5802/afst.1807

Publication membre du centre Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 2258-7519

Annales de la faculté des sciences de Toulouse Volume XXXIV, n° 1, 2025
pp. 107-134

Modulated logarithmic Sobolev inequalities and generation of chaos [(*)]

Matthew Rosenzweig [(1)] and Sylvia Serfaty [(2)]

Abstract. — We consider mean-field limits for overdamped Langevin dynamics of N particles with possibly singular interactions. It has been shown that a modulated free energy method can be used to prove the mean-field convergence or propagation of chaos for a certain class of interactions, including Riesz kernels. We show here that generation of chaos, i.e. exponential in time convergence to a tensorized (or iid) state starting from a nontensorized one, can be deduced from the modulated free energy method provided a uniform-in-N "modulated logarithmic Sobolev inequality" holds. Proving such an inequality is a question of independent interest, which is generally difficult. As an illustration, we show that uniform modulated logarithmic Sobolev inequalities can be proven for a class of situations in one dimension.

Résumé. — On considère la limite de champ moyen pour la dynamique de Langevin suramortie de N particules en interaction (possiblement) singulière. Il a été montré qu'on peut utiliser une méthode d'énergie libre modulée pour traiter une certaine classe d'interactions qui inclut les potentiels de Riesz. Nous montrons ici que l'on peut déduire de la méthode d'énergie libre modulée la génération du chaos, c.à.d. la convergence exponentielle en temps vers un état tensorisé partant d'une situation initiale non tensorisée, à condition qu'une « inégalité de Sobolev logarithmique modulée » uniforme en N soit vraie. Prouver une telle inégalité est une question indépendante qui est en général difficile. Comme illustration, nous montrons que des inégalités de Sobolev logarithmiques modulées peuvent être prouvées pour une classe de problèmes unidimensionnels.

(*) Reçu le 21 juillet 2023, accepté le 6 novembre 2023.

(1) Carnegie Mellon University, Department of Mathematical Sciences, Pittsburgh, PA, USA — mrosenz2@andrew.cmu.edu

M. R. is supported in part by the Simons Foundation through the Simons Collaboration on Wave Turbulence and by NSF grants DMS-2052651, DMS-2206085.

(2) Courant Institute of Mathematical Sciences, New York University, New York City, NY, USA — serfaty@cims.nyu.edu

S. S. is supported by the Simons Foundation through a Simons Investigator award, by NSF grant DMS-2247846, and by an excellence chair of the Fondation Sciences Mathématiques de Paris and PSL.

Article proposé par Radu Ignat.

1. Introduction

Consider a canonical Gibbs measure for N particles with energy $\mathcal{H}_N(X_N)$, with $X_N \coloneqq (x_1, \ldots, x_N)$, $x_i \in \mathbb{R}^{\mathsf{d}}$, of the form

$$\mathrm{d}\mathbb{P}_{N,\beta}(X_N) = \frac{1}{Z_{N,\beta}} e^{-\beta\mathcal{H}_N(X_N)} \mathrm{d}x_1 \ldots x_N. \tag{1.1}$$

It is well know that if $\mathbb{P}_{N,\beta}$ satisfies a Poincaré (or spectral gap) or logarithmic Sobolev inequality (LSI), with a constant independent of N, then the joint law of the particles under the overdamped Langevin (Glauber) dynamics

$$\mathrm{d}x_i^t = -\nabla_i \mathcal{H}_N(X_N) + \sqrt{\frac{2}{\beta}} \mathrm{d}W_i^t \qquad i \in \{1, \ldots, N\}, \tag{1.2}$$

where the W_i^t are independent standard Brownian motions, converges exponentially fast in time to the steady state $\mathbb{P}_{N,\beta}$. See, for instance, [5, 53]. A Poincaré inequality or LSI is satisfied as soon as $\mathcal{H}_N$ satisfies a uniform strict convexity condition of the form $\operatorname{Hess} \mathcal{H}_N \geqslant c I_{\mathrm{d}N \times \mathrm{d}N}$ with $c > 0$, with the Poincaré/LSI constant only depending on c [4]. Proving uniform LSIs meaningfully beyond this uniformly convex case is in general hard and the object of current efforts. We refer to [6, 7] for some instances of progress. For a taste of the extensive literature on LSIs, we refer to [1].

In this note, we are interested in the particular case of pair interaction energies of the form

$$\mathcal{H}_N(X_N) = \frac{1}{2N} \sum_{1 \leqslant i \neq j \leqslant N} \mathsf{g}(x_i, x_j) + \sum_{i=1}^{N} V(x_i), \tag{1.3}$$

where again $X_N = (x_1, \ldots, x_N) \in (\mathbb{R}^{\mathsf{d}})^N$; $\mathsf{g} : (\mathbb{R}^{\mathsf{d}})^2 \to [-\infty, \infty]$ is some symmetric interaction potential belonging to a class to be specified later and is similar to that considered in [11, 12, 19, 31], which includes repulsive Coulomb and Riesz interactions of the form

$$\mathsf{g}(x,y) = \begin{cases} -\log|x-y|, & s = 0 \\ \dfrac{1}{s}|x-y|^{-s}, & s < \mathsf{d}, \end{cases} \tag{1.4}$$

as well as moderately attractive ones; and V is some confinement potential.

In that case, the overdamped Langevin dynamics is of the form

$$\begin{cases} \mathrm{d}x_i^t = \left(-\frac{1}{N} \sum_{1\leqslant j\leqslant N: j\neq i} \nabla_1 \mathsf{g}(x_i^t, x_j^t) - \nabla V(x_i^t) \right) \mathrm{d}t + \sqrt{\frac{2}{\beta}} \mathrm{d}W_i^t \\ x_i^t|_{t=0} = x_i^0 \end{cases} \qquad i \in \{1, \dots, N\}, \quad (1.5)$$

with $x_i^0 \in \mathbb{R}^{\mathsf{d}}$ the pairwise distinct initial positions. Here, ∇_1 denotes the gradient with respect to the first argument of g. The mean-field limit, or equivalently *propagation of chaos*, for such evolutions has been proved for g sufficiently regular by many classical methods [30], for g possibly singular (with $V = 0$) by the relative entropy method in [31], and for g even more singular and Coulomb/Riesz-like by the modulated free energy method [11, 12, 19] using the modulated energy of [21, 51]. We will recall these methods below but suffice it to say that propagation of chaos means that if the initial data is distributed according to the probability distribution $f_N^0(X_N) = \mu^0(x_1)\dots\mu^0(x_N)$ on $(\mathbb{R}^{\mathsf{d}})^N$ (i.e., the particles are iid with common law μ^0), then the solution f_N^t of the N-particle Liouville/forward Kolmogorov equation associated to the dynamics (1.5) is such that

$$f_{N,k}^t \rightharpoonup (\mu^t)^{\otimes k} \quad \text{as } N \to \infty, \qquad (1.6)$$

where $f_{N,k}^t$ denotes the k-point marginal of f_N^t and μ^t is a solution to the mean-field evolution(1)

$$\begin{cases} \partial_t \mu^t - \operatorname{div}((\nabla \mathsf{g} * \mu^t + \nabla V)\mu^t) = \frac{1}{\beta}\Delta \mu^t \\ \mu^t|_{t=0} = \mu^0. \end{cases} \qquad (1.7)$$

The convergence (1.6) is for fixed k, as $N \to \infty$. There has been recent progress on understanding the optimal rate of this convergence in the context of the relative entropy method [33]. The modulated free energy method yields convergence in relative entropy, which in turn implies convergence of all the fixed marginals. Here, the (normalized) relative entropy is defined by

$$H_N(f_N|g_N) := \frac{1}{N} \int_{(\mathbb{R}^{\mathsf{d}})^N} \log\left(\frac{f_N}{g_N}\right) \mathrm{d}f_N. \qquad (1.8)$$

There has also been progress on showing bounds for the relative entropy which vanish as $N \to \infty$ and hold uniformly in time, hence proving uniform-in-time propagation of chaos in [19, 25, 26, 34, 46]. Informally, the distance between the laws $f_N^t, (\mu^t)^{\otimes N}$ does not grow arbitrarily large as time becomes large.

(1) In (1.7) and the remainder of the paper, we abuse the convolution notation by defining $\mathsf{g} * \mu(x) := \int_{\mathbb{R}^{\mathsf{d}}} \mathsf{g}(x,y)\mathrm{d}\mu(y) = \int_{\mathbb{R}^{\mathsf{d}}} \mathsf{g}(y,x)\mathrm{d}\mu(y)$, since g is assumed to be symmetric.

The notion of *generation of chaos*, a term coined recently by Lukkarinen [37], consists in a similar convergence as time gets large, even when the initial data f_N^0 is not tensorized, i.e. does not exhibit chaos or independence. Interpreted in an entropic sense (see [28] for a discussion of various notions of chaos), we have generation of chaos if $H_N(f_N^t|(\mu^t)^{\otimes N}) = o_t(1)+o_N(1)$ as $N \to \infty$ and $t \to \infty$, where $o_t(1)$ depends on N only through $H_N(f_N^0|(\mu^0)^{\otimes N})$ and is uniform in N assuming $H_N(f_N^0|(\mu^0)^{\otimes N}) = O(1)$, and $o_N(1)$ is uniform in t. This is what we wish to demonstrate here holds, under a *uniform-in-N modulated LSI* condition, that we will define below.

1.1. Modulated energies and modulated Gibbs measures

Before going further, let us review the notion of modulated energy. This object was first introduced as a next-order electric energy in [41, 49, 50] and used in the dynamics context as a modulated energy in [21, 51] and following works (in the spirit of [9]). Given a probability density μ on $\mathbb{R}^{\mathsf{d}}$, we define the modulated energy of the configuration X_N as

$$F_N(X_N,\mu) \\ := \frac{1}{2}\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle} \mathsf{g}(x,y)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i}-\mu\right)(x)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i}-\mu\right)(y), \quad (1.9)$$

where $\triangle$ denotes the diagonal in $(\mathbb{R}^{\mathsf{d}})^2$. This is the total interaction of the system of N discrete charges at x_i against a negative (neutralizing) background charge μ, with the self-interaction of the points, which is infinite if $\mathsf{g}(x,x)=\infty$,[(2)] removed. As shown in the aforementioned prior works, F_N is not necessarily positive; however, under appropriate assumptions on g, it acts in effect as a squared distance between the empirical measure $\frac{1}{N}\sum_{i=1}^N \delta_{x_i}$ and μ. Given the density μ, we can also define the *modulated Gibbs measure*

$$\mathbb{Q}_{N,\beta}(\mu) := \frac{1}{K_{N,\beta}(\mu)} e^{-\beta N F_N(X_N,\mu)}\mathrm{d}\mu(x_1)\ldots\mathrm{d}\mu(x_N), \quad (1.10)$$

where

$$K_{N,\beta}(\mu) := \int_{(\mathbb{R}^{\mathsf{d}})^N} e^{-\beta N F_N(X_N,\mu)}\mathrm{d}\mu(x_1)\ldots\mathrm{d}\mu(x_N) \quad (1.11)$$

is the associated partition function. An example of use of such a modulated Gibbs measure is provided in [2] in the study of (1.1) for the energy (1.3) in the case where g is the Coulomb interaction.

[(2)] If $\mathsf{g}(x,x)$ is finite, then the renormalization is unnecessary. See Section 1.5 for elaboration.

Following for instance [2, 3], we may introduce in the context of (1.3) the *thermal equilibrium measure* μ_β, which is defined as the minimizer among probability densities of the mean-field free energy

$$\mathcal{E}_\beta(\mu) := \frac{1}{2}\int_{(\mathbb{R}^\mathsf{d})^2} \mathsf{g}(x,y)\mathrm{d}\mu(x)\mathrm{d}\mu(y) + \int_{\mathbb{R}^\mathsf{d}} V(x)\mathrm{d}\mu(x) + \frac{1}{\beta}\int_{\mathbb{R}^\mathsf{d}} \log\mu(x)\mathrm{d}\mu(x). \quad (1.12)$$

If V grows sufficiently fast at infinity, then $\mathcal{E}_\beta$ has a unique minimizer, which is characterized by the existence of a constant $c_\beta \in \mathbb{R}$ such that

$$\mathsf{g} * \mu_\beta + V + \frac{1}{\beta}\log\mu_\beta = c_\beta \quad \text{in } \mathbb{R}^\mathsf{d}. \quad (1.13)$$

In the Coulomb case, [3] studied how μ_β converges to the usual equilibrium measure as $\beta \to \infty$.

The thermal equilibrium measure allows, as seen in [2], for a nice *splitting of the energy* and thus of the Gibbs measure, as follows. For any X_N, using (1.9), (1.13), and direct computations, we have

$$\mathcal{H}_N(X_N) = N\mathcal{E}_\beta(\mu_\beta) + NF_N(X_N,\mu_\beta) - \frac{1}{\beta}\sum_{i=1}^N \log\mu_\beta(x_i). \quad (1.14)$$

Inserting this identity into (1.1), we find that

$$\mathrm{d}\mathbb{P}^V_{N,\beta}(X_N) = \frac{e^{-\beta N\mathcal{E}_\beta(\mu_\beta)}}{Z^V_{N,\beta}} e^{-\beta NF_N(X_N,\mu_\beta)}\mathrm{d}\mu_\beta(x_1)\ldots\mathrm{d}\mu_\beta(x_N). \quad (1.15)$$

In other words, comparing with (1.10), we have found that

$$\mathbb{P}^V_{N,\beta} = \mathbb{Q}_{N,\beta}(\mu_\beta) \quad (1.16)$$

and

$$Z^V_{N,\beta} = K_{N,\beta}(\mu_\beta)e^{-\beta N\mathcal{E}_\beta(\mu_\beta)}. \quad (1.17)$$

Thus, the Gibbs measure is itself a modulated Gibbs measure, relative to the thermal equilibrium measure.

Conversely, given a probability measure μ, it is easy to see the modulated Gibbs measure $\mathbb{Q}_{N,\beta}(\mu)$ as a Gibbs measure through a change of the confining potential. Following (1.13), let

$$V_{\mu,\beta} := -\mathsf{g} * \mu - \frac{1}{\beta}\log\mu. \quad (1.18)$$

Then retracing the steps of the splitting formula above, one has

$$\mathbb{Q}_{N,\beta}(\mu) = \mathbb{P}^{V_{\mu,\beta}}_{N,\beta}. \quad (1.19)$$

With the rewriting (1.16), a crucial condition, appearing in all that follows, is

$$|\log K_{N,\beta}(\mu)| = o(N) \tag{1.20}$$

with a $o(N)$ uniform in $\beta \in [\frac{1}{2}\beta_0, 2\beta_0]$, for some fixed β_0, which corresponds for instance to the "large deviations estimates" in [31]. We will call it a *smallness of the free energy*. This condition (and even a stronger quantitative one) can be proven in the Riesz cases (1.4) and for bounded continuous interactions. We give a short proof of this fact in the appendix. In the attractive log case, it is proven in [12] and later streamlined in [18].

In several cases of interest, including in particular (1.4) (cf. [39, 51]), F_N is positive up to a small additive constant and controls a form of distance (e.g., a squared Sobolev norm). In such cases one may easily obtain a concentration estimate around μ as follows.[(3)] By definition (1.10) of $\mathbb{Q}_{N,\beta}(\mu)$, we may rewrite the exponential moments of the modulated energy F_N as

$$\log \mathbb{E}_{\mathbb{Q}_{N,\beta}(\mu)}\left[e^{\frac{\beta}{2}NF_N(X_N,\mu)}\right] = \log \frac{K_{N,\beta/2}(\mu)}{K_{N,\beta}(\mu)}. \tag{1.21}$$

If (1.20) holds, we obtain the exponential moment control

$$\left|\log \mathbb{E}_{\mathbb{Q}_{N,\beta}(\mu)}\left[e^{\frac{\beta}{2}NF_N(X_N,\mu)}\right]\right| \leqslant o(N). \tag{1.22}$$

Thus, using the almost positivity of F_N and the fact that it controls a squared distance between the empirical measure and reference density, this provides a concentration estimate around μ and implies a law of large numbers in the form

$$\mathbb{E}_{\mathbb{Q}_{N,\beta}(\mu)}\left[\left\|\frac{1}{N}\sum_{i=1}^{N}\delta_{x_i} - \mu\right\|^2\right] \to 0, \tag{1.23}$$

where $\|\cdot\|$ is a suitable norm. By standard arguments, this convergence also implies propagation of chaos for the statistical equilibrium $\mathbb{Q}_{N,\beta}(\mu)$ (see for instance [16, 49]):

$$\mathbb{Q}_{N,\beta}^{(k)}(\mu) \rightharpoonup \mu^{\otimes k} \quad \text{as } N \to \infty, \tag{1.24}$$

where $\mathbb{Q}_{N,\beta}^{(k)}(\mu)$ denotes the k-point marginal of $\mathbb{Q}_{N,\beta}(\mu)$ and k is fixed.

---

(3) If this is true for $-F_N$ instead of F_N, the same reasoning below applies, using 2β and β instead of $\beta/2$ and β in (1.21).

1.2. Modulated free energy

We may now define the modulated free energy, as introduced in [10, 11, 12]. Given a reference probability density μ on $\mathbb{R}^\mathsf{d}$ as above and a probability density f_N on $(\mathbb{R}^\mathsf{d})^N$, the modulated free energy is defined by

$$E_N(f_N,\mu) := \frac{1}{\beta} H_N(f_N|\mu^{\otimes N}) + \mathbb{E}_{f_N}\left[F_N(X_N,\mu)\right], \tag{1.25}$$

where H_N is the relative entropy as in (1.8) and $\mathbb{E}_{f_N}$ denotes the expectation with respect to the measure f_N, viewing the X_N as a random variable. Let us remark here that using the explicit form of (1.10), the modulated free energy can be rewritten as

$$E_N(f_N,\mu) = \frac{1}{\beta}\left(H_N(f_N|\mathbb{Q}_{N,\beta}(\mu)) - \frac{\log K_{N,\beta}(\mu)}{N}\right). \tag{1.26}$$

In other words, up to a constant related to the smallness of free energy condition (1.20), the modulated free energy is another relative entropy. Note that this provides an easy proof of the fact that $E_N(f_N,\mu)$ is essentially positive if the smallness of free energy condition (1.20) holds. Moreover, controlling the relative entropy from f_N to $\mathbb{Q}_{N,\beta}$ proves closeness of the particle density to $\mathbb{Q}_{N,\beta}(\mu)$ and is, in reality, more precise than the mean-field limit and propagation of chaos provided by the control of $H_N(f_N|\mu^{\otimes N})$. As $t\to\infty$, the solution μ^t to (1.7) converges to the thermal equilibrium measure μ_β, and $\mathbb{Q}_{N,\beta}(\mu_\beta)$ is, as already noticed in (1.19), equal to $\mathbb{P}_{N,\beta}$, so we retrieve the fact, provided by usual LSI, that there is convergence in large time to $\mathbb{P}_{N,\beta}$, the invariant measure for the dynamics (1.5). See Section 1.6 below for a further discussion on the advantages of $\mathbb{Q}_{N,\beta}(\mu)$ over $\mathbb{P}_{N,\beta}$. Finally, if one wishes to retrieve closeness of f_N to $\mu^{\otimes N}$, one may either use a control of the negative part of the modulated energy by the relative entropy, as ensured by condition (ii) below, or use the concentration inequality via its consequence (1.24).

1.3. Evolution of modulated energy, Fisher information, and uniform LSI

The crucial computation of [10, 11, 12] (performed on the torus, but the whole-space with confining potential case is similar) is that when differentiating in time $E_N(f_N^t,\mu^t)$, for f_N^t solving the forward Kolmogorov equation and μ^t solving the mean-field evolution equation (1.7), a cancellation occurs, leading to

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}E_N(f_N^t,\mu^t)\\
&\leqslant -\frac{1}{2}\int_{(\mathbb{R}^{\mathsf{d}})^N}\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle}(u^t(x)-u^t(y))\cdot\nabla_1\mathsf{g}(x,y)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu^t\right)^{\otimes 2}(x,y)\,\mathrm{d}f_N^t\\
&\quad-\frac{1}{\beta^2 N}\int_{(\mathbb{R}^{\mathsf{d}})^N}\sum_{i=1}^N\left|\nabla_i\log\left(\frac{f_N^t}{(\mu^t)^{\otimes N}}\right)+\frac{\beta}{N}\sum_{j\neq i}\nabla_1\mathsf{g}(x_i,x_j)-\beta\nabla\mathsf{g}*\mu^t(x_i)\right|^2\mathrm{d}f_N^t,
\end{aligned}\tag{1.27}$$

where

$$u^t \coloneqq \frac{1}{\beta}\nabla\log\mu^t+\nabla V+\nabla\mathsf{g}*\mu^t \tag{1.28}$$

is the velocity field associated to the mean-field dynamics (1.7).

At first pass, the second term on the right-hand side of (1.27), which is nonpositive, may be discarded, and, assuming g is translation-invariant, the first term in the right-hand side can be controlled, for instance in Riesz cases (1.4) via the second author's inequality from [51] and its refinements and generalizations [39, 47], by the modulated energy itself, allowing to close a Grönwall loop. When $V=0$, this is what is done in [11] and revisited in [18, 19]. More precisely, the following type of inequality is used: for any sufficiently regular vector field v and any pairwise distinct $X_N\in(\mathbb{R}^{\mathsf{d}})^N$,

$$\begin{aligned}
&\left|\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle}(v(x)-v(y))\cdot\nabla_1\mathsf{g}(x,y)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu\right)^{\otimes 2}(x,y)\right|\\
&\qquad\qquad\leqslant C\|v\|_*\left(F_N(X_N,\mu)+o_N(1)\right),
\end{aligned}\tag{1.29}$$

where $\|v\|_*$ is some homogeneous Sobolev norm of v and $o_N(1)$ depends only on (and is increasing with respect to) the L^∞ norm of μ and vanishes as $N\to\infty$. This inequality was first proven in full generality in [51] for all Coulomb/super-Coulombic Riesz potentials, following a previous work for the $\mathsf{d}=2$ Coulomb case [35]. A sharp additive error $o_N(1)=O(\|\mu\|_{L^\infty}^{\frac{s}{\mathsf{d}}}N^{\frac{s}{\mathsf{d}}-1})$ with $\|v\|_*=\|\nabla v\|_{L^\infty}$ was proven in [47], following earlier Coulomb results [35, 44, 52]. The estimate (1.29) was generalized to Riesz-like kernels in [39]. For g satisfying $|(x-y)\cdot\nabla_1\mathsf{g}(x,y)|\leqslant C$, one may extract from [31], as was done in [12], the averaged inequality

$$\left|\int_{(\mathbb{R}^{\mathsf{d}})^N}\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle}(v(x)-v(y))\cdot\nabla_1\mathsf{g}(x,y)\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu\right)^{\otimes 2}(x,y)\,\mathrm{d}f_N\right| \\ \leqslant \|\nabla v\|_{L^\infty}\left(C_1H_N(f_N|\mu^{\otimes N})+\frac{C_2}{N}\right). \quad (1.30)$$

Let us now examine the nonpositive term in the right-hand side of (1.27). We rewrite it as

$$-\frac{1}{\beta^2N}\int_{(\mathbb{R}^{\mathsf{d}})^N}\sum_{i=1}^N\left|\nabla\log\left(\frac{f_N^t}{(\mu^t)^{\otimes N}}\right)+\frac{\beta}{N}\sum_{j\neq i}\nabla_1\mathsf{g}(x_i,x_j)-\nabla\mathsf{g}*\mu^t(x_i)\right|^2\mathrm{d}f_N^t$$
$$=-\frac{1}{\beta^2N}\int_{(\mathbb{R}^{\mathsf{d}})^N}\left|\nabla\log\frac{f_N^t}{\mathbb{Q}_{N,\beta}(\mu^t)}\right|^2\mathrm{d}f_N^t$$
$$=-\frac{1}{\beta^2N}\int_{(\mathbb{R}^{\mathsf{d}})^N}\left|\nabla\sqrt{\frac{f_N^t}{\mathbb{Q}_{N,\beta}(\mu^t)}}\right|^2\mathrm{d}\mathbb{Q}_{N,\beta}(\mu^t). \quad (1.31)$$

Indeed, one may check that by definition (1.10) of $\mathbb{Q}_{N,\beta}(\mu)$,

$$\nabla_i\log\mathbb{Q}_{N,\beta}(\mu)=-\beta N\nabla_iF_N(X_N,\mu)+\nabla\log\mu(x_i), \quad (1.32)$$

and in view of the definition (1.9) of $F_N(X_N,\mu)$,

$$\nabla_iF_N(X_N,\mu)=\frac{1}{N^2}\sum_{1\leqslant j\leqslant N:j\neq i}\nabla_1\mathsf{g}(x_i,x_j)-\frac{1}{N}\nabla(\mathsf{g}*\mu)(x_i). \quad (1.33)$$

For any f_N and any reference probability density μ, we call the quantity

$$\frac{1}{N}\int_{(\mathbb{R}^{\mathsf{d}})^N}\left|\nabla\sqrt{\frac{f_N}{\mathbb{Q}_{N,\beta}(\mu)}}\right|^2\mathrm{d}\mathbb{Q}_{N,\beta}(\mu) \quad (1.34)$$

the *modulated Fisher information*, which is nothing but the normalized relative Fisher information $I_N(f_N|\mathbb{Q}_{N,\beta}(\mu))$, and the relation (1.27) transforms into

$$\frac{\mathrm{d}}{\mathrm{d}t}E_N(f_N^t,\mu^t)\leqslant-\frac{1}{\beta^2N}\int_{(\mathbb{R}^{\mathsf{d}})^N}\left|\nabla\sqrt{\frac{f_N^t}{\mathbb{Q}_{N,\beta}(\mu^t)}}\right|^2\mathrm{d}\mathbb{Q}_{N,\beta}(\mu^t) \\ -\frac{1}{2}\int_{(\mathbb{R}^{\mathsf{d}})^N}\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle}(u^t(x)-u^t(y))\cdot\nabla_1\mathsf{g}(x,y)\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu^t\right)^{\otimes 2}(x,y)\,\mathrm{d}f_N^t. \quad (1.35)$$

The goal is then to exploit a functional inequality relating the modulated Fisher information to the modulated free energy to take advantage of the negative term in (1.35).

Definition 1.1. — *We say that a family of probability measures* $\{P_N\}_{N\geqslant 1}$ *satisfies a* uniform logarithmic Sobolev inequality *(LSI) if there exists a constant* $C_{LS} > 0$*, such that for any* $N \geqslant 1$ *and* $f \in C^1((\mathbb{R}^{\mathsf{d}})^N)$*, we have*

$$\int_{(\mathbb{R}^{\mathsf{d}})^N} f^2 \log \frac{f^2}{\int f^2 \mathrm{d}P_N} \mathrm{d}P_N \leqslant C_{LS} \int_{(\mathbb{R}^{\mathsf{d}})^N} |\nabla f|^2 \mathrm{d}P_N. \tag{1.36}$$

Given data (g, V, β)*, we say that a* uniform μ-modulated LSI (μ-LSI) *holds if the family of probability measures* $\{\mathbb{Q}_{N,\beta}(\mu)\}_{N\geqslant 1}$ *of the form* (1.10) *satisfies a uniform LSI.*

Our main observation is that if $\mathbb{Q}_{N,\beta}(\mu)$ satisfies a uniform LSI, then applying (1.36) to $f = \sqrt{\frac{f_N}{\mathbb{Q}_{N,\beta}(\mu)}}$, with f_N a probability density on $(\mathbb{R}^{\mathsf{d}})^N$, we find

$$\int_{(\mathbb{R}^{\mathsf{d}})^N} \left|\nabla \sqrt{\frac{f_N}{\mathbb{Q}_{N,\beta}(\mu)}}\right|^2 \mathrm{d}\mathbb{Q}_{N,\beta}(\mu) \geqslant \frac{1}{C_{LS}} \int_{(\mathbb{R}^{\mathsf{d}})^N} \log \left(\frac{\frac{f_N}{\mathbb{Q}_{N,\beta}(\mu)}}{\int_{(\mathbb{R}^{\mathsf{d}})^N} \frac{f_N}{\mathbb{Q}_{N,\beta}(\mu)} \mathrm{d}\mathbb{Q}_{N,\beta}(\mu)} \right) \mathrm{d}f_N. \tag{1.37}$$

Using that f_N is a probability density, we recognize on the right-hand side $NH_N(f_N|\mathbb{Q}_{N,\beta}(\mu))$. In light of (1.26), we then have

$$\frac{1}{N} \int_{(\mathbb{R}^{\mathsf{d}})^N} \left|\nabla \sqrt{\frac{f_N}{\mathbb{Q}_{N,\beta}(\mu)}}\right|^2 \mathrm{d}\mathbb{Q}_{N,\beta}(\mu) \geqslant \frac{1}{C_{LS}} \left(\beta E_N(f_N, \mu) + \frac{1}{N} \log K_{N,\beta}(\mu) \right). \tag{1.38}$$

In other words, a uniform LSI for $\mathbb{Q}_{N,\beta}(\mu)$ implies that the modulated Fisher information is bounded below by the modulated free energy and an additive error that is $o_N(1)$ assuming smallness of free energy. If (1.38) holds for all μ^t along the flow, then it can be inserted into (1.35) to obtain an exponential decay of the modulated free energy, provided (1.29) or (1.30) holds.

In [26], in the context of conservative dynamics on the torus $\mathbb{T}^d$ (see remarks at the end of Section 1.4), a uniform LSI is used in the context of the relative entropy method [31]. In that method, one differentiates in time $H_N(f_N^t|(\mu^t)^{\otimes N})$ instead of (1.26), leading to a Fisher information relative to the reference measure $(\mu^t)^{\otimes N}$ instead of $\mathbb{Q}_{N,\beta}(\mu^t)$. Proving the needed uniform LSI holds is straightforward, as it follows from upper and lower bounds on μ^t (a consequence of maximum principle and only possible on compact domains) and the Holley–Stroock perturbation lemma. See also [34] for a similar idea applied to the hierarchal relative entropy method of [33].

1.4. Main result

To present the main result of this note, we list some assumptions that we make on the potential $\mathsf{g} : (\mathbb{R}^{\mathsf{d}})^2 \to [-\infty, \infty]$. We will explain below specific cases in which these assumptions hold.

(i) $\mathsf{g} \in C^2((\mathbb{R}^{\mathsf{d}})^2 \setminus \triangle)$ is symmetric and for some $s < \mathsf{d}$, satisfies

$$|\mathsf{g}(x,y)| \leqslant C \begin{cases} 1 + |\log|x-y||\,, & s = 0 \\ 1 + |x-y|^{-s}, & s > 0 \end{cases} \tag{1.39}$$

for some constant $C > 0$.

(ii) There exists a constant $C_\beta \in [0, \frac{1}{\beta})$ such that for any $f_N \in \mathcal{P}_{ac}((\mathbb{R}^{\mathsf{d}})^N)$ and $\mu \in \mathcal{P}(\mathbb{R}^{\mathsf{d}}) \cap L^\infty(\mathbb{R}^{\mathsf{d}})$, with $\int_{\mathbb{R}^{\mathsf{d}}} \log(1+|x|)\mathrm{d}\mu(x) < \infty$ if $s = 0$,

$$\mathbb{E}_{f_N}\left[F_N(X_N, \mu)\right] \geqslant -C_\beta H_N(f_N|\mu^{\otimes N}) - o_N(1), \tag{1.40}$$

where $o_N(1)$ only depends (in an increasing fashion) on μ through $\|\mu\|_{L^\infty}$.

(iii) There exist constants $C_{RE}, C_{ME} \geqslant 0$, such that

$$\left| \int_{(\mathbb{R}^{\mathsf{d}})^N} \int_{(\mathbb{R}^{\mathsf{d}})^2 \setminus \triangle} (v(x) - v(y)) \cdot \nabla_1 \mathsf{g}(x,y)\, \mathrm{d}\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu \right)^{\otimes 2} (x,y) \mathrm{d}f_N \right| \\ \leqslant \|v\|_* \left(C_{RE} H_N(f_N|\mu^{\otimes N}) + C_{ME} \mathbb{E}_{f_N}\left[F_N(X_N,\mu)\right] + o_N(1) \right) \tag{1.41}$$

for all pairwise distinct configurations $X_N \in (\mathbb{R}^{\mathsf{d}})^N$, densities $f_N \in \mathcal{P}_{ac}((\mathbb{R}^{\mathsf{d}})^N)$ and $\mu \in \mathcal{P}(\mathbb{R}^{\mathsf{d}}) \cap L^\infty(\mathbb{R}^{\mathsf{d}})$, and continuous vector fields v with finite homogeneous Sobolev norm $\|\cdot\|_*$ of some order.

Remark 1.2. — Assumption (i) is to ensure that all energy expressions are well-defined and that all differential identities can be justified. Assumption (ii) ensures that the modulated energy does overwhelm the relative entropy, which is not *a priori* forbidden, since we make no sign assumptions on g. Since $C_\beta < \frac{1}{\beta}$, it ensures that the modulated free energy is nonnegative up to $o_N(1)$ error. In fact, it shows that the modulated free energy controls the relative entropy.

Let us introduce the quantity

$$\mathcal{E}_N^t \coloneqq E_N(f_N^t, \mu^t) + o_N^t(1) \tag{1.42}$$

as a substitute for the modulated free energy. The additive error $o_N^t(1)$ is a constant multiple of the maximum of the additive errors in assumptions (ii), (iii) and ensures that $\mathcal{E}_N^t \geqslant 0$, which allows to perform a Grönwall argument on this quantity. It depends only on μ^t through the L^∞ norm, hence the t superscript, and is increasing in this dependence. Also, it is easier to write the statements with $\mathcal{E}_N^t$, as these additive constants appear as the errors $o_N(1)$ in (1.29).

Theorem 1.3. — *Let $\beta > 0$. Assume that equation* (1.7) *admits a solution $\mu \in C([0,\infty), \mathcal{P}(\mathbb{R}^{\mathsf{d}}) \cap L^\infty(\mathbb{R}^{\mathsf{d}}))$, such that $\|\mu^t\|_{L^\infty}$ is bounded uniformly in t and $\nabla u^t \in L^\infty$ locally uniformly in t. If $s = 0$, further assume that $\int_{\mathbb{R}^{\mathsf{d}}} \log(1+|x|) \mathrm{d}\mu^t < \infty$ for every $t \geqslant 0$. If $\mathbb{Q}_{N,\beta}(\mu^t)$ satisfies a uniform LSI with constant $C_{LS} > 0$ for every $t \geqslant 0$, then, for every $t \geqslant 0$,*

$$\begin{aligned}\mathcal{E}_N^t \leqslant\ & e^{-\frac{4t}{\beta C_{LS}} + \int_0^t \frac{\mathsf{C}\|u^\tau\|_*}{2}\mathrm{d}\tau} \mathcal{E}_N^0 \\ & + e^{-\frac{4t}{\beta C_{LS}} + \int_0^t \frac{\mathsf{C}\|u^\tau\|_*}{2}\mathrm{d}\tau} \\ & \times \int_0^t e^{\frac{4\tau}{\beta C_{LS}} - \int_0^\tau \frac{\mathsf{C}\|u^{\tau'}\|_*}{2}\mathrm{d}\tau'} \left[\dot{o}_N^\tau + \frac{4}{\beta C_{LS}} \left(o_N^\tau(1) - \frac{\log K_{N,\beta}(\mu^\tau)}{\beta N} \right) \right] \mathrm{d}\tau,\end{aligned} \tag{1.43}$$

where $K_{N,\beta}(\mu^\tau)$ is as in (1.11)*, $o_N^\tau(1)$ is as above, and $\dot{o}_N^\tau(1)$ denotes the derivative of $o_N^\tau(1)$ with respect to time.*

We see here that provided $\int_0^\infty \|u^\tau\|_* \mathrm{d}\tau < \infty$, the first term on the right-hand side converges exponentially fast to 0 as $t \to \infty$, while the second term is $o_N(1)$ uniformly bounded in t, assuming $\log K_{N,\beta}(\mu^\tau) = o(N)$ uniformly in τ and that $\int_0^\infty |\dot{o}_N^\tau(1)| \mathrm{d}\tau < \infty$, by the fundamental theorem of calculus and our assumption that $\|\mu^t\|_{L^\infty}$ is uniformly bounded. Since $\mathcal{E}_N$ differs from E_N only by additive constants which are $o_N(1)$, and the modulated free energy E_N controls the relative entropy H_N, as explained in Remark 1.2, it follows that the estimate (1.43) implies entropic generation of chaos and also gives a uniform-in-time propagation of chaos if the initial data is such that $\mathcal{E}_N^0 = o_N(1)$. In the next subsection, we give cases of interest to which Theorem 1.3 applies.

Generation of chaos for potentials g with $\nabla \mathsf{g}$ in L^∞, which does not allow for singular potentials, was shown in [34], with a rate of convergence in N that is sharp for relative entropy, under smallness assumptions on β. In [26], a generation of chaos result was shown for *conservative* dynamics (replace ∇ with $\mathbb{M}\nabla$ for an antisymmetric matrix $\mathbb{M}$) with g having a log-type singularity. Both [26, 34] are restricted to the torus $\mathbb{T}^{\mathsf{d}}$. A weaker generation of chaos result in 2-Wasserstein distance was shown in [25] for the Riesz case on $\mathbb{R}$ with uniformly convex confinement via coupling methods. We mention that convergence in relative entropy implies convergence in W_2 by a theorem of Otto–Villani [40].

Remark 1.4. — The long-time analysis of equation (1.7) that allows to show in the Riesz case that $\int_0^\infty \|\nabla u^\tau\|_{L^\infty} d\tau < \infty$ and $K_{N,\beta}(\mu^\tau) = o(N)$ uniformly in τ is the subject of forthcoming work with J. Huang [29]. In fact, this work shows that solutions converge as $t \to \infty$ to the thermal equilibrium μ_β in a strong sense at a quantifiable rate and even covers the case of $\mathbb{R}^{\mathsf{d}}$ without confinement, which has been an outstanding problem.

Remark 1.5. — One could also consider the periodic setting $\mathbb{T}^{\mathsf{d}}$, as in [11, 12, 19, 26]. But the case of $\mathbb{R}^{\mathsf{d}}$ is mathematically more interesting.

In the case where $\mathsf{d} = 1$ and $\mathsf{g}(x) = -\log|x|$ or $|x|^{-s}$ for $s \in (0,1)$, V is a C^2 uniformly convex potential (e.g., $V(x) = |x|^2$), and μ is a probability density which is not too far from the thermal equilibrium μ_β, we are able to verify a uniform μ-modulated LSI. The general d-dimensional Riesz case is challenging: it is at least as difficult as the uniform LSI for $\mathbb{P}^V_{N,\beta}$, which is a well-known open problem.

1.5. Applications

We can give a more precise form of the estimate (1.43) in the repulsive singular Riesz case (1.4) so that $(-\Delta)^{\frac{\mathsf{d}-s}{2}}\mathsf{g} = \mathsf{c}_{\mathsf{d},s}\delta_0$. One has that

$$F_N(\underline{x}_N,\mu) \geqslant -\begin{cases} \dfrac{\log(N\|\mu\|_{L^\infty})}{2N\mathsf{d}}\mathbf{1}_{s=0} + \mathsf{C}\|\mu\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{\frac{s}{\mathsf{d}}-1}, & s \geqslant \mathsf{d}-2 \\ \dfrac{\mathsf{C}\log(N\|\mu\|_{L^\infty})}{N}\mathbf{1}_{s=0} + \mathsf{C}\|\mu\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{-\frac{2(\mathsf{d}-s)}{2(\mathsf{d}-s)+s(\mathsf{d}+2)}}, & s < \mathsf{d}-2. \end{cases} \tag{1.44}$$

Here, $\mathsf{C} > 0$ is an absolute constant. The additive errors for the sub-Coulomb case $s < \mathsf{d}-2$ are expected to be suboptimal, while they are sharp in the Coulomb/super-Coulomb case $s \geqslant \mathsf{d}-2$.[4] For details, we refer to [46] in the case $s < \mathsf{d}-2$ and [19, 47] in the case $s \geqslant \mathsf{d}-2$. In particular, (1.44)

(4) The L^∞ condition here (and by implication, the L^∞ condition in Theorem 1.3) can be relaxed quite a bit (e.g., see [42, 43]) at the cost of increasing the additive errors; but we will not concern ourselves with such generality.

shows that

$$E_N(f_N,\mu) \geqslant \frac{1}{\beta} H_N(f_N|\mu^{\otimes N}) - \begin{cases} \dfrac{\log(N\|\mu\|_{L^\infty})}{2N\mathsf{d}}\mathbf{1}_{s=0} + \mathsf{C}\|\mu\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{\frac{s}{\mathsf{d}}-1}, & s \geqslant \mathsf{d}-2 \\ \dfrac{\mathsf{C}\log(N\|\mu\|_{L^\infty})}{N}\mathbf{1}_{s=0} + \mathsf{C}\|\mu\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{-\frac{2(\mathsf{d}-s)}{2(\mathsf{d}-s)+s(\mathsf{d}+2)}}, & s < \mathsf{d}-2. \end{cases} \tag{1.45}$$

We take

$$\mathcal{E}_N^t \coloneqq E_N(f_N^t,\mu^t) + \begin{cases} \dfrac{\log(N\|\mu^t\|_{L^\infty})}{2N\mathsf{d}}\mathbf{1}_{s=0} + \mathsf{C}\|\mu^t\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{\frac{s}{\mathsf{d}}-1}, & s \geqslant \mathsf{d}-2 \\ \dfrac{\mathsf{C}\log(N\|\mu^t\|_{L^\infty})}{N}\mathbf{1}_{s=0} + \mathsf{C}\|\mu^t\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{-\frac{2(\mathsf{d}-s)}{2(\mathsf{d}-s)+s(\mathsf{d}+2)}}, & s < \mathsf{d}-2 \end{cases} \tag{1.46}$$

The estimate (1.41) holds with $C_{RE}=0$,

$$\|v\|_* = \begin{cases} \|\nabla v\|_{L^\infty}, & s \geqslant \mathsf{d}-2 \\ \|\nabla v\|_{L^\infty} + \|(-\Delta)^{\frac{\mathsf{d}-s}{4}} v\|_{L^{\frac{2\mathsf{d}}{\mathsf{d}-2-s}}}, & s < \mathsf{d}-2, \end{cases} \tag{1.47}$$

and

$$o_N^t(1) = \begin{cases} \dfrac{\log(N\|\mu^t\|_{L^\infty})}{2N\mathsf{d}}\mathbf{1}_{s=0} + \mathsf{C}\|\mu^t\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{\frac{s}{\mathsf{d}}-1}, & s \geqslant \mathsf{d}-2 \\ \|(-\Delta)^{\frac{s+1-\mathsf{d}}{2}}\mu^t\|_{L^\infty} N^{-\frac{s+1+\frac{(2(\mathsf{d}-s)}{\mathsf{d}+2}}{\left(s+\frac{(2(\mathsf{d}-s)}{\mathsf{d}+2}\right)(1+s)}} + \|\mu^t\|_{L^\infty}^{\frac{2+s}{\mathsf{d}+2}} N^{-\frac{\frac{(2(\mathsf{d}-s)}{\mathsf{d}+2}}{\left(s+\frac{(2(\mathsf{d}-s)}{\mathsf{d}+2}\right)(1+s)}}, & s < \mathsf{d}-2. \end{cases} \tag{1.48}$$

In the attractive log case on the torus $\mathbb{T}^{\mathsf{d}}$, it is shown in [18] (building on [12]) that there exists $\beta_{\mathsf{d}} > 0$ such that for any $0 \leqslant \beta < \beta_{\mathsf{d}}$, there are constants $\varepsilon_\beta \in (0,1)$ and $C_\beta > 0$ such that

$$\beta \mathbb{E}_{f_N}\left[F_N(X_N,\mu)\right] \leqslant \varepsilon_\beta H_N(f_N|\mu^{\otimes N}) + \frac{C_\beta}{N}. \tag{1.49}$$

Therefore, assumption (ii) is satisfied. It is shown in [18] that the optimal value $\beta_{\mathsf{d}} \leqslant \min(2\mathsf{d}, \frac{2\pi^{\mathsf{d}/2}}{\Gamma(\mathsf{d}/2)})$, where Γ denotes the usual Gamma function, as the inequality (1.49) is false for $\beta > \frac{2\pi^{\mathsf{d}/2}}{\Gamma(\mathsf{d}/2)}$.

For potentials $\mathsf{g} : (\mathbb{R}^{\mathsf{d}})^2 \to \mathbb{R}$ that are continuous along the diagonal, one can skip the renormalization and simply define

$$F_N(X_N,\mu) = \int_{(\mathbb{R}^{\mathsf{d}})^2} \mathsf{g}(x,y)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)(x)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N \delta_{x_i} - \mu\right)(y). \tag{1.50}$$

If g is repulsive in the sense that $\mathsf{g}(x,y)$ is the integral kernel of a positive semidefinite operator on the space of finite Borel measures, as in the case of the equations used for neural networks parameters evolution [17, 38, 48], then $F_N(X_N,\mu) \geqslant 0$.

Continuing to assume that g is continuous at the origin, but dropping the repulsive assumption, we may use the Donsker–Varadhan lemma to estimate

$$\mathbb{E}_{f_N}\left[F_N(X_N,\mu)\right] \leqslant \frac{1}{\eta}\left(H_N(f_N|\mu^{\otimes N}) + \frac{1}{N}\log \mathbb{E}_{\mu^{\otimes N}}\left[e^{N\eta F_N(X_N,\mu)}\right]\right) \tag{1.51}$$

for any $\eta > 0$. If $\mathsf{g} \in L^\infty$, then one may use [31, Theorem 4] (see also [36, Section 5] for a simpler proof) with

$$\begin{aligned}\phi(x,z) \coloneqq \Bigg(\mathsf{g}(x,z) - \int_{\mathbb{R}^{\mathsf{d}}} \mathsf{g}(x,y)\mathrm{d}\mu(y) - \int_{\mathbb{R}^{\mathsf{d}}} \mathsf{g}(y,z)\mathrm{d}\mu(y)& \\ + \int_{(\mathbb{R}^{\mathsf{d}})^2} \mathsf{g}(y,y')\mathrm{d}\mu(y)\mathrm{d}\mu(y')\Bigg).&\end{aligned} \tag{1.52}$$

The conclusion is that if $\sqrt{C_0}\eta\|\phi\|_{L^\infty} < 1$, where C_0 is a universal constant, then

$$\log \mathbb{E}_{\mu^{\otimes N}}\left[e^{N\eta F_N(X_N,\mu)}\right] \leqslant \log\left(\frac{2}{1 - C_0\eta^2\|\phi\|^2_{L^\infty}}\right). \tag{1.53}$$

Replacing F_N by $-F_N$ and repeating the preceding reasoning, we then find that

$$\left|\mathbb{E}_{f_N}\left[F_N(X_N,\mu)\right]\right| \leqslant \frac{1}{\eta}H_N(f_N|\mu^{\otimes N}) + \frac{1}{\eta N}\log\left(\frac{2}{1 - C_0\eta^2\|\phi\|^2_{L^\infty}}\right). \tag{1.54}$$

If $\frac{1}{\beta} > \frac{1}{\sqrt{C_0}\|\phi\|_{L^\infty}}$, then we may choose $\frac{1}{\eta} \in \left(\frac{1}{\sqrt{C_0}\|\phi\|_{L^\infty}}, \frac{1}{\beta}\right)$, implying that assumption (ii) holds.

1.6. Advantages of modulated LSI over LSI

Let us explain the advantage of a uniform modulated LSI over merely a uniform LSI for $\mathbb{P}_{N,\beta}$. For simplicity, let us assume that g is translation-invariant. Ignoring regularity questions,

$$\frac{\mathrm{d}}{\mathrm{d}t} H_N\left(f_N^t|\mathbb{P}_{N,\beta}\right) = -\frac{1}{\beta} I_N(f_N^t|\mathbb{P}_{N,\beta}), \tag{1.55}$$

where we recall that I_N is the normalized relative Fisher information. If there is a uniform LSI constant C_{LS} for $\mathbb{P}_{N,\beta}$, then by Grönwall's lemma,

$$H_N\left(f_N^t|\mathbb{P}_{N,\beta}\right) \leqslant e^{-\frac{t}{C_{LS}\beta}} H_N\left(f_N^0|\mathbb{P}_{N,\beta}\right). \tag{1.56}$$

By subadditivity of relative entropy and Pinsker's inequality, for any fixed $1 \leqslant k \leqslant N$,

$$\left\| f_{N,k}^t - \mathbb{P}_{N,\beta}^{(k)} \right\|_{TV}^2 \leqslant 2k e^{-\frac{t}{C_{LS}\beta}} H_N\left(f_N^0|\mathbb{P}_{N,\beta}\right). \tag{1.57}$$

If μ^t is a solution of the mean-field evolution (1.7), then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathcal{E}_\beta(\mu^t) - \mathcal{E}_\beta(\mu^\beta)\right] = -\int_{\mathbb{R}^\mathsf{d}} \left| \frac{1}{\beta}\nabla \log \mu^t + \nabla \mathsf{g} * \mu^t + \nabla V \right|^2 \mathrm{d}\mu^t, \tag{1.58}$$

where $\mathcal{E}_\beta$ is the mean-field free energy as defined in (1.12). One may check by direct computation that

$$\lim_{N\to\infty} \frac{1}{\beta} H_N(\mu^{\otimes N}|\mathbb{P}_{N,\beta}) = \mathcal{E}_\beta(\mu) - \mathcal{E}_\beta(\mu_\beta) \tag{1.59}$$

and

$$\lim_{N\to\infty} \frac{1}{\beta} I_N(\mu^{\otimes N}|\mathbb{P}_{N,\beta}) = \beta \int_{\mathbb{R}^\mathsf{d}} \left| \frac{1}{\beta} \log \mu + \nabla V + \nabla \mathsf{g} * \mu \right|^2 \mathrm{d}\mu, \tag{1.60}$$

which together with the uniform LSI for $\mathbb{P}_{N,\beta}$ imply the infinite-volume LSI

$$\begin{aligned} \beta\left[\mathcal{E}_\beta(\mu) - \mathcal{E}_\beta(\mu_\beta)\right] &= \lim_{N\to\infty} H_N(\mu^{\otimes N}|\mathbb{P}_{N,\beta}) \\ &\leqslant \lim_{N\to\infty} C_{LS} I_N(\mu^{\otimes N}|\mathbb{P}_{N,\beta}) \\ &= C_{LS}\beta^2 \int_{\mathbb{R}^\mathsf{d}} \left| \frac{1}{\beta} \log \mu + \nabla V + \nabla \mathsf{g} * \mu \right|^2 \mathrm{d}\mu. \end{aligned} \tag{1.61}$$

Inserting this inequality into the right-hand side of (1.58) and applying Grönwall again,

$$\left[\mathcal{E}_\beta(\mu^t) - \mathcal{E}_\beta(\mu^\beta)\right] \leqslant e^{-\frac{t}{C_{LS}\beta}} \left[\mathcal{E}_\beta(\mu^0) - \mathcal{E}_\beta(\mu^\beta)\right]. \tag{1.62}$$

Using (1.13) and direct computation, one may also check that

$$\mathcal{E}_\beta(\mu) - \mathcal{E}_\beta(\mu_\beta) = \frac{1}{\beta}\int_{\mathbb{R}^\mathsf{d}} \log\left(\frac{\mu}{\mu_\beta}\right) \mathrm{d}\mu + \frac{1}{2}\int_{(\mathbb{R}^\mathsf{d})^2} \mathsf{g}(x-y)\,\mathrm{d}(\mu-\mu_\beta)^{\otimes 2}(x,y). \tag{1.63}$$

Assuming, say, that the Fourier transform $\widehat{\mathsf{g}} \geqslant 0$, we may discard the potential energy term and then apply Pinsker's inequality again to obtain

$$\left\|\mu^t - \mu_\beta\right\|_{TV}^2 \leqslant 2\beta e^{-\frac{t}{C_{LS}\beta}} \left[\mathcal{E}_\beta(\mu^0) - \mathcal{E}_\beta(\mu^\beta)\right]. \tag{1.64}$$

Considering just the case $k=1$ to simplify the analysis, we have by triangle inequality that

$$\begin{aligned} &\left\|f_{N,1}^t - \mu^t\right\|_{TV} \\ &\quad\leqslant \left\|f_{N,1}^t - \mathbb{P}_{N,\beta}^{(1)}\right\|_{TV} + \left\|\mathbb{P}_{N,\beta}^{(1)} - \mu_\beta\right\|_{TV} + \left\|\mu^t - \mu_\beta\right\|_{TV} \\ &\quad\leqslant \sqrt{2e^{-\frac{t}{C_{LS}\beta}} H_N\left(f_N^0|\mathbb{P}_{N,\beta}\right)} \\ &\qquad + \sqrt{2\beta e^{-\frac{t}{C_{LS}\beta}}\left(\mathcal{E}_\beta(\mu^0) - \mathcal{E}_\beta(\mu^\beta)\right)} + \left\|\mathbb{P}_{N,\beta}^{(1)} - \mu_\beta\right\|_{TV}. \end{aligned} \tag{1.65}$$

Supposing that[5]

$$\left\|\mathbb{P}_{N,\beta}^{(1)} - \mu_\beta\right\|_{TV} = o_N(1), \tag{1.66}$$

the right-hand side of (1.65) tends to zero as $t\to\infty$ and $N\to\infty$. But this estimate does not imply propagation of chaos, even locally in time, as the second term does not vanish as $N\to\infty$. To address this unsatisfactory feature, one would also need a local-in-time estimate with which to interpolate, say, of the form

$$\left\|f_{N,1}^t - \mu^t\right\|_{TV} \leqslant e^{Ct} o_N(1), \tag{1.67}$$

where $o_N(1)$ vanishes as $N\to\infty$ assuming some form of chaos for the initial data.

The above described argument is rather inefficient. We had to pass from relative entropy to a genuine metric, total variation distance, to implement this triangle inequality argument. In doing so, one loses the optimality of the rate in N [33]. Moreover, by trying to balance t and N, the rate of convergence further deteriorates. In contrast, a uniform modulated LSI addresses propagation/generation of chaos in one swoop, because it is dynamic: it not only depends on N but also allows for dependence on t through the flowing of μ according to (1.7).

[5] Such a bound is known (with a sharp estimate for $o_N(1)$), for instance, in the high-temperature case where g has bounded gradient [32].

We mention that this classical triangle inequality/interpolation idea was used in [20] for energies with regular interactions, except with total variation distance replaced by 2-Wasserstein distance, which works just as well since LSI implies a Talagrand inequality [40]. Though, only a statement of uniform-in-time propagation of chaos (with suboptimal rate), as opposed to generation of chaos, is presented in [20].

1.7. Organization of the paper

Let us conclude the introduction with some remarks on the organization of the body of the paper. In Section 2, we give the details of the proof of the main result, Theorem 1.3. Then in Section 3, we turn to proving that a uniform modulated LSI holds in the log/Riesz case for a certain class of densities μ in dimension $\mathsf{d} = 1$.

1.8. Acknowledgments

The authors thank Djalil Chafaï for helpful discussion and references. The first author also thanks Daniel Lacker and Yair Shenfeld for discussion. The second author acknowledges the Fondation Sciences Mathématiques de Paris and PSL Research University who supported her visit to ENS-PSL, where this work was completed.

2. Proof of the main theorem

Applying the uniform LSI for $\mathbb{Q}_{N,\beta}(\mu^t)$ to the first term in the right-hand side of (1.35) via (1.38), we find, abbreviating $K_N^t \coloneqq K_{N,\beta}(\mu^t)$,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t} E_N(f_N^t, \mu^t) \\
&\leqslant -\frac{1}{2} \int_{(\mathbb{R}^{\mathsf{d}})^N} \int_{(\mathbb{R}^{\mathsf{d}})^2 \setminus \triangle} (u^\tau(x) - u^\tau(y)) \cdot \nabla_1 \mathsf{g}(x,y) \, \mathrm{d}\left(\frac{1}{N} \sum_{i=1}^N \delta_{x_i} - \mu^t \right)^{\otimes 2} (x,y) \, \mathrm{d}f_N^t \\
&\qquad\qquad - \frac{4}{\beta C_{LS}} \left(E_N(f_N^t, \mu^t) + \frac{\log K_N^t}{\beta N} \right). \quad (2.1)
\end{aligned}$$

Under the assumption (iii), we have

$$\int_{(\mathbb{R}^{\mathsf{d}})^N}\left|\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle}(u^\tau(x)-u^\tau(y))\cdot\nabla_1\mathsf{g}(x,y)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu^t\right)^{\otimes 2}(x,y)\right|\mathrm{d}f_N^t$$
$$\leqslant \|u^t\|_*\Big(C_{RE}H_N(f_N^t|(\mu^t)^{\otimes N})+C_{ME}\mathbb{E}_{f_N^t}\left[F_N(X_N,\mu^t)\right]+o_N^t(1)\Big). \quad (2.2)$$

If $\frac{C_{RE}}{C_{ME}}\leqslant\frac{1}{\beta}$, then since $H_N(f_N^t|(\mu^t)^{\otimes N})\geqslant 0$, we may assume without loss of generality that $\frac{C_{RE}}{C_{ME}}=\frac{1}{\beta}$. If $\frac{C_{RE}}{C_{ME}}>\frac{1}{\beta}$, then using assumption (ii),

$$\begin{aligned}C_{ME}&\mathbb{E}_{f_N^t}\left[F_N(X_N^t,\mu^t)\right]\\&\leqslant C_{ME}\Big(\mathbb{E}_{f_N^t}\left[F_N(X_N,\mu^t)\right]+C_\beta H_N(f_N^t|(\mu^t)^{\otimes N})+o_N^t(1)\Big)\\&\leqslant C'_{ME}\Big(\mathbb{E}_{f_N^t}\left[F_N(X_N,\mu^t)\right]+C_\beta H_N(f_N^t|(\mu^t)^{\otimes N})+o_N^t(1)\Big)\end{aligned} \quad (2.3)$$

for any $C'_{ME}\geqslant C_{ME}$. So, choosing C'_{ME} sufficiently large so that $\frac{C_{RE}}{C'_{ME}}+C_\beta\leqslant\frac{1}{\beta}$ (remember that $C_\beta<\frac{1}{\beta}$ by assumption), we see that in all cases,

$$\int_{(\mathbb{R}^{\mathsf{d}})^N}\left|\int_{(\mathbb{R}^{\mathsf{d}})^2\setminus\triangle}(u^\tau(x)-u^\tau(y))\cdot\nabla_1\mathsf{g}(x,y)\,\mathrm{d}\left(\frac{1}{N}\sum_{i=1}^N\delta_{x_i}-\mu^t\right)^{\otimes 2}(x,y)\right|\mathrm{d}f_N^t$$
$$\leqslant \mathsf{C}\|u^t\|_*\left(\frac{1}{\beta}H_N(f_N^t|(\mu^t)^{\otimes N})+\mathbb{E}_{f_N^t}\left[F_N(X_N,\mu^t)\right]+o_N^t(1)\right), \quad (2.4)$$

for some constant $\mathsf{C}>0$. To establish a Grönwall relation, we use the quantity (1.42). We see from combining (2.1) and (2.2) that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_N^t&\leqslant-\frac{4}{\beta C_{LS}}\left(E_N(f_N^t,\mu^t)+\frac{\log K_N^t}{\beta N}\right)+\frac{\mathsf{C}}{2}\|u^t\|_*\mathcal{E}_N^t+\dot{o}_N^t(1)\\&=\left(-\frac{4}{\beta C_{LS}}+\frac{\mathsf{C}\|u^t\|_*}{2}\right)\mathcal{E}_N^t+\frac{4}{\beta C_{LS}}\left(o_N^t(1)-\frac{\log K_N^t}{\beta N}\right).\end{aligned} \quad (2.5)$$

Recall that $\dot{o}_N^t(1)$ denotes the time derivative. Multiplying both sides by $e^{\int_0^t(\frac{4}{\beta C_{LS}}-\frac{\mathsf{C}\|u^{\tau'}\|_*}{2})\mathrm{d}\tau'}$, we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}&\left[e^{\int_0^t(\frac{4}{\beta C_{LS}}-\frac{\mathsf{C}\|u^{\tau'}\|_*}{2})\mathrm{d}\tau'}\mathcal{E}_N^t\right]\\&\leqslant e^{\int_0^t(\frac{4}{\beta C_{LS}}-\frac{\mathsf{C}\|u^{\tau'}\|_*}{2})\mathrm{d}\tau'}\left[\dot{o}_N^t(1)+\frac{4}{\beta C_{LS}}\left(o_N^t(1)-\frac{\log K_N^t}{\beta N}\right)\right].\end{aligned} \quad (2.6)$$

Now using the fundamental theorem of calculus followed by a little rearrangement,

$$\begin{aligned}\mathcal{E}_N^t \leqslant e^{-\frac{4t}{\beta C_{LS}}+\int_0^t \frac{\mathsf{C}\|u^\tau\|_*}{2}\mathrm{d}\tau}\mathcal{E}_N^0 \\ + e^{-\frac{4t}{\beta C_{LS}}+\int_0^t \frac{\mathsf{C}\|u^\tau\|_*}{2}\mathrm{d}\tau} \\ \times \int_0^t e^{\frac{4\tau}{\beta C_{LS}}-\int_0^\tau \frac{\mathsf{C}\|u^{\tau'}\|_*}{2}\mathrm{d}\tau'}\left[\dot{o}_N^\tau(1)+\frac{4}{\beta C_{LS}}\left(o_N^\tau(1)-\frac{\log K_N^\tau}{\beta N}\right)\right]\mathrm{d}\tau. \quad (2.7)\end{aligned}$$

This gives the estimate (1.43) and therefore completes the proof of Theorem 1.3.

3. Uniform LSI for $\mathsf{d} = 1$ Riesz case

We show in this section that a uniform modulated LSI holds in the $\mathsf{d} = 1$ repulsive Riesz case (1.4) for uniformly convex confinement V. Using the notation from the introduction,

$$\mathcal{H}_N(X_N) := \sum_{i=1}^N V(x_i) + \frac{1}{N}\sum_{1\leqslant i<j\leqslant N} \mathsf{g}(x_j - x_i). \quad (3.1)$$

Remark 3.1. — In fact, the proof will show that a modulated LSI holds for any interaction potential g which is convex or for any C^2 interaction potential with $\|\mathsf{g}\|_{\dot{C}^2}$ sufficiently small depending on the convexity of V. We leave the details as an exercise for the reader. We expect that one could generalize further by following the proof of Zegarlinski's theorem [54], as used to show uniform LSIs in [27], or the two-scale approach of [24], but will not pursue this.

Remark 3.2. — As explained in [45], the approach of [6] implies the LSI up to the critical inverse temperature for the Gibbs measure of the mean-field classical XY/$O(2)$/planar rotator/Kuramoto model, whose energy is far from convex. It is straightforward to adapt the reasoning of Section 3.2 to obtain a modulated LSI for μ close enough to $\mu_\beta = 1$.

3.1. Uniform LSI for $\mathbb{P}^V_{N,\beta}$

Following Chafaï–Lehec [15],[6] we present the LSI for the Gibbs measure $\mathbb{P}^V_{N,\beta}$ in the $\mathsf{d}=1$ Riesz case with uniformly convex confinement V. This is a warm-up for proving the modulated LSI in the next subsection.

Proposition 3.3. — *Let $V:\mathbb{R}\to\mathbb{R}$ be κ-convex for some $\kappa>0$. For $\beta>0$, the probability measure $\mathbb{P}^V_{N,\beta}$ has LSI constant $\frac{2}{\beta\kappa}$.*

Proof. — As V is fixed, we omit the superscript in $\mathbb{P}^V_{N,\beta}$ in what follows. By exchangeability, it suffices to restrict to the *Weyl chamber*[7] $\Delta_N := \{X_N\in\mathbb{R}^N : x_1\leqslant\cdots\leqslant x_N\}$. More precisely, define

$$\widetilde{\mathsf{g}}(x) := \begin{cases}\mathsf{g}(x), & x>0\\ \infty, & x\leqslant 0\end{cases}$$

$$\text{and}\quad \widetilde{\mathcal{H}}_N(X_N) := \sum_{i=1}^N V(x_i) + \frac{1}{N}\sum_{1\leqslant i<j\leqslant N}\widetilde{\mathsf{g}}(x_j-x_i), \quad (3.2)$$

and $d\widetilde{\mathbb{P}}_{N,\beta} = \frac{e^{-\beta\widetilde{\mathcal{H}}_N}}{\widetilde{Z}_{N,\beta}}\mathrm{d}X_N$. Since

$$\int_{\mathbb{R}^N} e^{-\beta\mathcal{H}_N}\mathrm{d}X_N = N!\int_{\Delta_N} e^{-\beta\mathcal{H}_N}\mathrm{d}X_N, \quad (3.3)$$

it follows that if φ is invariant under permutation of coordinates, then

$$\int_{\mathbb{R}^N}\varphi^2\log\left(\frac{\varphi^2}{\int\varphi^2\mathrm{d}\widetilde{\mathbb{P}}_{N,\beta}}\right)\mathrm{d}\widetilde{\mathbb{P}}_{N,\beta} = \int_{\mathbb{R}^N}\varphi^2\log\left(\frac{\varphi^2}{\int\varphi^2\mathrm{d}\widetilde{\mathbb{P}}_{N,\beta}}\right)\mathrm{d}\widetilde{\mathbb{P}}_{N,\beta}. \quad (3.4)$$

So, $\widetilde{\mathbb{P}}_{N,\beta}$ has LSI constant C_{LS} if and only if $\mathbb{P}_{N,\beta}$ has LSI constant C_{LS}. Going forward, we drop the $\widetilde{}$ superscript in $\widetilde{\mathsf{g}},\widetilde{\mathcal{H}}_N,\widetilde{\mathbb{P}}_{N,\beta}$.

Assuming that V is κ-convex, for some $\kappa>0$, we claim that $\mathcal{H}_N$ is κ-convex. Indeed, let $X_N,Y_N\in\Delta_N$ and $\rho\in(0,1)$. We want to show that

$$\mathcal{H}_N\left(\rho X_N+(1-\rho)Y_N\right) \leqslant \rho\mathcal{H}_N(X_N)+(1-\rho)\mathcal{H}_N(Y_N) - \frac{\kappa\rho(1-\rho)}{2}|Y_N-X_N|^2. \quad (3.5)$$

---

[6] Strictly speaking, [15] considers the $\mathsf{d}=1$ log case; but the argument works with trivial modification in the general Riesz case. Furthermore, Chafaï–Lehec present more than one proof; but we choose to highlight the one based on Caffarelli's contraction theorem.

[7] This ability to order is, of course, a special feature of the one-dimensional setting.

If $x_i = x_j$ or $y_i = y_j$ for some $1 \leqslant i < j \leqslant N$, then the right-hand side is infinite and the inequality holds trivially; so, suppose otherwise. Since V is κ-convex, we have for each i,

$$V(\rho x_i + (1-\rho)y_i) \leqslant \rho V(x_i) + (1-\rho)V(y_i) - \frac{\kappa\rho(1-\rho)}{2}|y_i - x_i|^2. \quad (3.6)$$

So, it only remains to show that for each pair $i < j$,

$$\begin{aligned}\mathsf{g}\left([\rho x_j + (1-\rho)y_j] - [\rho x_i + (1-\rho)y_i]\right)\\ &= \mathsf{g}\left(\rho(x_j - x_i) + (1-\rho)(y_j - y_i)\right)\\ &\leqslant \rho\mathsf{g}(x_j - x_i) + (1-\rho)\mathsf{g}(y_j - y_i). \quad (3.7)\end{aligned}$$

Fix a pair $i < j$. If $x_j - x_i = y_j - y_i$, then there is nothing further to show; so, suppose otherwise. Without loss of generality, suppose $y_j - y_i > x_j - x_i > 0$. Then by the fact that

$$\forall\, x > 0, \qquad \mathsf{g}''(x) = \begin{cases} \dfrac{1}{x^2}, & s = 0 \\ \dfrac{s(s+1)}{|x|^{s+2}}, & s \neq 0, \end{cases} \quad (3.8)$$

and therefore g is convex on $\mathbb{R}_+$, we see that (3.7) holds.

We perform a qualitative regularization argument that reduces us to the case when $\mathbb{P}_{N,\beta}$ has full support $\mathbb{R}^N$ and $\mathcal{H}_N \in C^2(\mathbb{R}^N)$. Let $\mathbb{G}_N$ be the Gaussian measure with covariance $(\beta\kappa)^{-1/2} I_{N\times N}$,

$$\mathrm{d}\mathbb{G}_N = (2\pi/\beta\kappa)^{-\frac{N}{2}} e^{-\frac{\beta\kappa|X_N|^2}{2}} \mathrm{d}X_N. \quad (3.9)$$

Since

$$\log\left(\frac{\mathrm{d}\mathbb{P}_{N,\beta}}{\mathrm{d}\mathbb{G}_N}\right) = -\beta\mathcal{H}_N - \log(Z_{N,\beta}) + \frac{N}{2}\log\left(\frac{2\pi}{\beta\kappa}\right) + \frac{\beta\kappa|X_N|^2}{2}, \quad (3.10)$$

we see that $\mathcal{H}_N$ is κ-convex if and only if $\log\big(\frac{\mathrm{d}\mathbb{P}_{N,\beta}}{\mathrm{d}\mathbb{G}_N}\big)$ is concave. Let $\{Q_t\}_{t\geqslant 0}$ be the Ornstein–Uhlenbeck semigroup with stationary measure $\mathbb{G}_N$: for any test function f,

$$\forall\, X_N \in \mathbb{R}^N, \quad (Q_t f)(X_N) := \int_{\mathbb{R}^N} f\Big(e^{-t}X_N + \sqrt{1-e^{-2t}}Y_N\Big)\mathrm{d}\mathbb{G}_N(Y_N). \quad (3.11)$$

The measure $\mathbb{G}_N$ is reversible for $\{Q_t\}_{t\geqslant 0}$. Therefore, $Q_t\#\mathbb{P}_{N,\beta}$ is absolutely continuous with respect to $\mathbb{G}_N$, and its Radon–Nikodym derivative $\frac{\mathrm{d}Q_t\#\mathbb{P}_{N,\beta}}{\mathrm{d}\mathbb{G}_N} = Q_t\big(\frac{\mathrm{d}\mathbb{P}_{N,\beta}}{\mathrm{d}\mathbb{G}_N}\big)$. Moreover, as consequence of the Prékopa–Leindler inequality, Q_t preserves log concavity. Hence, $\mathcal{H}_N^t := -\frac{1}{\beta}\log\left(Q_t\#\mathbb{P}_{N,\beta}\right)$ is κ-convex and belongs to $C^\infty(\mathbb{R}^N)$. Finally, since $\lim_{t\to 0}(Q_t f)(x) = f(x)$ for any continuous f, it follows that $Q_t\#\mathbb{P}_{N,\beta} \rightharpoonup \mathbb{P}_{N,\beta}$ as $t \to 0$. Thus, if $Q_t\#\mathbb{P}_{N,\beta}$ has LSI constant C_{LS} for every $t > 0$, then so does $\mathbb{P}_{N,\beta}$.

We proceed under the C^2 and full support assumptions. According to Caffarelli's contraction theorem [13, 14] (see also [22] for an alternative proof), if $\mathcal{H}_N$ is $\beta\kappa$-convex, then the Brenier map [8] T from $\mathbb{G}_N$ to $\mathbb{P}_{N,\beta}$ (i.e., $T\#\mathbb{G}_N = \mathbb{P}_{N,\beta}$) is 1-Lipschitz. So, for any test function $\varphi \geqslant 0$,

$$\begin{aligned}
\int_{\mathbb{R}^N} \varphi \log(\varphi) \mathrm{d}\mathbb{P}_{N,\beta} &= \int_{\mathbb{R}^N} \varphi \log(\varphi) \mathrm{d}(T\#\mathbb{G}_N) \\
&= \int_{\mathbb{R}^N} (\varphi \circ T) \log(\varphi \circ T) \mathrm{d}\mathbb{G}_N \\
&\leqslant \frac{2}{\beta\kappa} \int_{\mathbb{R}^N} |\nabla(\varphi \circ T)|^2 \mathrm{d}\mathbb{G}_N \\
&\leqslant \frac{2}{\beta\kappa} \int_{\mathbb{R}^N} |(\nabla\varphi) \circ T|^2 |\nabla T|^2 \mathrm{d}\mathbb{G}_N \\
&\leqslant \frac{2}{\beta\kappa} \int_{\mathbb{R}^N} |\nabla\varphi|^2 \mathrm{d}\mathbb{P}_{N,\beta}.
\end{aligned} \tag{3.12}$$

In the third line, we have used the well-known LSI for $\mathbb{G}_N$ [23]; and in the final line we have used that $\|\nabla T\|_{L^\infty} \leqslant 1$ together with another application of $T\#\mathbb{G}_N = \mathbb{P}_{N,\beta}$. This completes the proof. ☐

3.2. Uniform LSI for $\mathbb{Q}_{N,\beta}(\mu)$

Given a density μ, recall from (1.19) and (1.18) that

$$\mathbb{Q}_{N,\beta}(\mu) = \mathbb{P}_{N,\beta}^{V_{\mu,\beta}}, \qquad \text{where} \quad V_{\mu,\beta} := -\mathsf{g} * \mu - \frac{1}{\beta} \log \mu. \tag{3.13}$$

We recycle the notation $\mathcal{H}_N$, so that

$$\mathcal{H}_N(X_N) = \sum_{i=1}^N V_{\mu,\beta}(x_i) + \frac{1}{2N} \sum_{1 \leqslant i \neq j \leqslant N} \mathsf{g}(x_i - x_j). \tag{3.14}$$

The advantage of this notation is that assuming $V_{\mu,\beta}$ is κ-convex, for some $\kappa > 0$, we may apply Proposition 3.3 with V replaced by $V_{\mu,\beta}$ to obtain a uniform LSI for $\mathbb{Q}_{N,\beta}(\mu)$.

Proposition 3.4. — *Suppose that $\mu \in \mathcal{P}(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and if $s = 0$, also suppose that $\int \log(1 + |x|) \mathrm{d}\mu < \infty$.*[8] *For $\beta > 0$, suppose that $V_{\mu,\beta}$ is κ-convex, for some $\kappa > 0$. Then the probability measure $\mathbb{Q}_{N,\beta}(\mu)$ has LSI constant $\frac{2}{\beta\kappa}$.*

---

[8] The L^∞ and log moment assumptions are just to ensure that the convolution $\mathsf{g} * \mu$ is well-defined.

To give meaning to Proposition 3.4, we now specify conditions under which $V_{\mu,\beta}$ is uniformly convex.

Lemma 3.5. — *Let $\mu \in \mathcal{P}(\mathbb{R})$ be such that $\log \frac{\mu}{\mu_\beta} \in C^2(\mathbb{R})$.*(9) *Suppose $V \in C^2$ and $\beta > 0$. Then $V_{\mu,\beta}$ is κ-convex with*

$$\kappa := \inf V'' - \left(\frac{1}{\beta}\|\log \frac{\mu}{\mu_\beta}\|_{\dot{C}^2} + \|\mathsf{g} * (\mu - \mu_\beta)\|_{\dot{C}^2}\right). \tag{3.15}$$

Proof. — Recalling the definition of $V_{\mu,\beta}$,

$$\begin{aligned} V_{\mu,\beta} &= -\mathsf{g} * (\mu - \mu_\beta + \mu_\beta) - \frac{1}{\beta}\log\left(\frac{\mu}{\mu_\beta}\mu_\beta\right) \\ &= -\mathsf{g} * (\mu - \mu_\beta) - \frac{1}{\beta}\log\frac{\mu}{\mu_\beta} - \left(\mathsf{g} * \mu_\beta + \frac{1}{\beta}\log \mu_\beta\right) \\ &= -\mathsf{g} * (\mu - \mu_\beta) - \frac{1}{\beta}\log\frac{\mu}{\mu_\beta} + V - c_\beta, \end{aligned} \tag{3.16}$$

where to obtain the third line, we have applied (1.13) to the last term of the second line. By triangle inequality,

$$V''_{\mu,\beta} \geqslant V'' - \|\mathsf{g} * (\mu - \mu_\beta)\|_{\dot{C}^2} - \frac{1}{\beta}\left\|\log\frac{\mu}{\mu_\beta}\right\|_{\dot{C}^2}, \tag{3.17}$$

from which the desired conclusion is immediate. □

Remark 3.6. — One can produce probability measures μ such that $\log \frac{\mu}{\mu_\beta} \in C^2$ by choosing $h \in C^2$ and then setting $\mu := \frac{e^h \mu_\beta}{\int e^h \mathrm{d}\mu_\beta}$, which is tautologically a probability density. One can make the quantities $\|\log \frac{\mu}{\mu_\beta}\|_{\dot{C}^2}$, $\|\mathsf{g} * (\mu - \mu_\beta)\|_{\dot{C}^2}$ arbitrarily small by taking $\|e^h - 1\|_{C^2}$ arbitrarily small. In particular, we see that there exist non-equilibrium densities μ such that $V_{\mu,\beta}$ is uniformly convex.

As a corollary in this one-dimensional Riesz case with uniformly convex confinement, suppose we start the dynamics (1.5) from an initial data μ^0 that is close enough to μ_β in the sense that (3.15) with $\mu = \mu^0$ is strictly positive. If this closeness persists throughout the dynamics (1.7) in the sense that (3.15) with $\mu = \mu^t$ is bounded from below by some $\kappa_0 > 0$ uniformly in t (this is a consequence of the aforementioned forthcoming work [29]), then Theorem 1.3 applies, showing entropic generation of chaos.

(9) Since $\mu_\beta \in C^2$, this assumption implies by the chain rule that $\mu \in C^2$.

Appendix A. Proof of the smallness of free energy in Riesz and regular cases

In this appendix, we prove the smallness of the free energy (1.20) in the cases (1.4) and in the case of bounded continuous nonnegative interactions, by showing

$$|\log K_{N,\beta}(\mu)| \leqslant \beta\, o(N), \tag{A.1}$$

for $o(N)$ independent of β.

The upper bound is obtained straightforwardly in the Riesz cases from inserting (1.44) into (1.11), then using that μ is a probability measure:

$$\log K_{N,\beta}(\mu) \leqslant \beta \begin{cases} \dfrac{\log(N\|\mu\|_{L^\infty})}{2\mathsf{d}}\mathbf{1}_{s=0} + \mathsf{C}\|\mu\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{\frac{s}{\mathsf{d}}}, & \mathsf{d}-2 \leqslant s < \mathsf{d} \\ \mathsf{C}\log(N\|\mu\|_{L^\infty})\mathbf{1}_{s=0} + \mathsf{C}\|\mu\|_{L^\infty}^{\frac{s}{\mathsf{d}}} N^{1-\frac{2(\mathsf{d}-s)}{2(\mathsf{d}-s)+s(\mathsf{d}+2)}}, & s < \mathsf{d}-2. \end{cases} \tag{A.2}$$

In all cases, the preceding right-hand side is $\beta\, o(N)$. When g is nonnegative and continuous, one can insert the diagonal back into the definition of F_N, which implies that

$$F_N(X_N, \mu) \geqslant -\frac{1}{2N}\mathsf{g}(0,0), \tag{A.3}$$

and the proof of the upper bound is concluded in the same way.

The lower bound follows from Jensen's inequality. Indeed,

$$\log K_{N,\beta}(\mu) \geqslant -\beta N \mathbb{E}_{\mu^{\otimes N}}\left[F_N(X_N,\mu)\right]. \tag{A.4}$$

We then expand out the definition (1.9) of F_N and use the symmetry of g to find that

$$\begin{aligned} \mathbb{E}_{\mu^{\otimes N}}\left[F_N(X_N,\mu)\right] &= \int_{(\mathbb{R}^\mathsf{d})^N} \Bigg(\frac{1}{2N^2}\sum_{i\neq j}\mathsf{g}(x_i,x_j) - \frac{1}{N}\sum_{i=1}^N \int_{\mathbb{R}^\mathsf{d}} \mathsf{g}(x_i,y)\mathrm{d}\mu(y) \\ &\qquad + \frac{1}{2}\int_{(\mathbb{R}^\mathsf{d})^2}\mathsf{g}(x,y)\mathrm{d}\mu^{\otimes 2}(x,y)\Bigg)\mathrm{d}\mu^{\otimes N}(X_N) \\ &= -\frac{1}{2N}\int_{(\mathbb{R}^\mathsf{d})^2}\mathsf{g}(x,y)\mathrm{d}\mu^{\otimes 2}(x,y). \end{aligned} \tag{A.5}$$

Inserting the last line back into the right-hand side of (A.4) yields

$$\log K_{N,\beta}(\mu) \geqslant \frac{\beta}{2}\int_{(\mathbb{R}^\mathsf{d})^2}\mathsf{g}(x-y)\mathrm{d}\mu^{\otimes 2}(x,y), \tag{A.6}$$

which gives the desired lower bound in all cases (1.4) and all cases where g is bounded.

Bibliography

[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto & G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses, vol. 10, Société Mathématique de France, 2000, xiii+217 pages.

[2] S. Armstrong & S. Serfaty, "Local laws and rigidity for Coulomb gases at any temperature", *Ann. Probab.* **49** (2021), no. 1, p. 46-121.

[3] ———, "Thermal approximation of the equilibrium measure and obstacle problem", *Ann. Fac. Sci. Toulouse, Math.* **31** (2022), no. 4, p. 1085-1110.

[4] D. Bakry & M. Émery, "Diffusions hypercontractives", in *Séminaire de probabilités, XIX, 1983/84*, Lecture Notes in Mathematics, vol. 1123, Springer, 1985, p. 177-206.

[5] D. Bakry, I. Gentil & M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Mathematischen Wissenschaften, vol. 348, Springer, 2014, xx+552 pages.

[6] R. Bauerschmidt & T. Bodineau, "A very simple proof of the LSI for high temperature spin systems", *J. Funct. Anal.* **276** (2019), no. 8, p. 2582-2588.

[7] ———, "Log-Sobolev inequality for the continuum sine-Gordon model", *Pure Appl. Math.* **74** (2064), no. 10, p. 2064-2113.

[8] Y. Brenier, "Polar factorization and monotone rearrangement of vector-valued functions", *Commun. Pure Appl. Math.* **44** (1991), no. 4, p. 375-417.

[9] ———, "Convergence of the Vlasov–Poisson system to the incompressible Euler equations", *Commun. Partial Differ. Equations* **25** (2000), no. 3-4, p. 737-754.

[10] D. Bresch, P.-E. Jabin & Z. Wang, "On mean-field limits and quantitative estimates with a large class of singular kernels: application to the Patlak–Keller–Segel model", *C. R. Math.* **357** (2019), no. 9, p. 708-720.

[11] ———, "Modulated free energy and mean field limit", *Sémin. Laurent Schwartz, EDP Appl.* **2019-2020** (2020), article no. 2 (22 pages).

[12] ———, "Mean field limit and quantitative estimates with singular attractive kernels", *Duke Math. J.* **172** (2023), no. 13, p. 2591-2641.

[13] L. A. Caffarelli, "Monotonicity properties of optimal transportation and the FKG and related inequalities", *Commun. Math. Phys.* **214** (2000), no. 3, p. 547-563.

[14] ———, "Erratum: "Monotonicity of optimal transportation and the FKG and related inequalities" [Comm. Math. Phys. **214** (2000), no. 3, 547–563; MR1800860 (2002c:60029)]", *Commun. Math. Phys.* **225** (2002), no. 2, p. 449-450.

[15] D. Chafaï & J. Lehec, "On Poincaré and logarithmic Sobolev inequalities for a class of singular Gibbs measures", in *Geometric aspects of functional analysis. Israel seminar (GAFA) 2017–2019. Volume 1*, Lecture Notes in Mathematics, vol. 2256, Springer, 2020, p. 219-246.

[16] L.-P. Chaintron & A. Diez, "Propagation of chaos: a review of models, methods and applications", 2021, `https://arxiv.org/abs/2106.14812v1`.

[17] L. Chizat & F. Bach, "On the global convergence of gradient descent for over-parameterized models using optimal transport", in *Proceedings of the 32nd International Conference on Neural Information Processing Systems, NIPS18*, Curran Associates, Inc., 2018, p. 3040-3050.

[18] A. C. de Courcel, M. Rosenzweig & S. Serfaty, "The attractive log gas: stability, uniqueness, and propagation of chaos", 2023, `https://arxiv.org/abs/2311.14560`.

[19] ———, "Sharp uniform-in-time mean-field convergence for singular periodic Riesz flows", *Ann. Inst. Henri Poincaré C, Anal. Non Linéaire* **42** (2025), no. 2, p. 391-472.

[20] M. G. Delgadino, R. S. Gvalani, G. A. Pavliotis & S. A. Smith, "Phase transitions, logarithmic Sobolev inequalities, and uniform-in-time propagation of chaos for weakly interacting diffusions", *Commun. Math. Phys.* **401** (2023), no. 1, p. 275-323.

[21] M. Duerinckx, “Mean-field limits for some Riesz interaction gradient flows”, *SIAM J. Math. Anal.* **43** (2016), no. 2, p. 2269-2300.
[22] M. Fathi, N. Gozlan & M. Prod'homme, “A proof of the Caffarelli contraction theorem via entropic regularization”, *Calc. Var. Partial Differ. Equ.* **59** (2020), no. 3, article no. 96 (18 pages).
[23] L. Gross, “Logarithmic Sobolev inequalities”, *Am. J. Math.* **97** (1975), no. 4, p. 1061-1083.
[24] N. Grunewald, F. Otto, C. Villani & M. G. Westdickenberg, “A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit”, *Ann. Inst. Henri Poincaré, Probab. Stat.* **45** (2009), no. 2, p. 302-351.
[25] A. Guillin, P. Le Bris & P. Monmarché, “On systems of particles in singular repulsive interaction in dimension one: log and Riesz gas”, *J. Éc. Polytech., Math.* **10** (2023), p. 867-916.
[26] ———, “Uniform in time propagation of chaos for the 2d vortex model and other singular stochastic systems”, *J. Eur. Math. Soc.* (2024), Online First. doi:10.4171/JEMS/1413.
[27] A. Guillin, W. Liu, L. Wu & C. Zhang, “Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems”, *Ann. Appl. Probab.* **32** (2022), no. 3, p. 1590-1614.
[28] M. Hauray & S. Mischler, “On Kac's chaos and related problems”, *J. Funct. Anal.* **266** (2014), no. 10, p. 6055-6157.
[29] J. Huang, M. Rosenzweig & S. Serfaty, “The modulated free energy method on $\mathbb{R}^d$”, in preparation.
[30] P.-E. Jabin & Z. Wang, “Mean field limit for stochastic particle systems”, in *Active Particles, Volume 1. Advances in Theory, Models, and Applications*, Modeling and Simulation in Science, Engineering and Technology, Birkhäuser/Springer, 2017, p. 379-402.
[31] ———, “Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels”, *Invent. Math.* **214** (2018), no. 1, p. 523-591.
[32] D. Lacker, “Quantitative approximate independence for continuous mean field Gibbs measures”, *Electron. J. Probab.* **27** (2022), article no. 15 (21 pages).
[33] ———, “Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions”, *Probab. Math. Phys.* **4** (2023), no. 2, p. 377-432.
[34] D. Lacker & L. Le Flem, “Sharp uniform-in-time propagation of chaos”, *Probab. Theory Relat. Fields* **187** (2023), no. 1-2, p. 443-480.
[35] T. Leblé & S. Serfaty, “Fluctuations of two dimensional Coulomb gases”, *Geom. Funct. Anal.* **28** (2018), no. 2, p. 443-508.
[36] T. S. Lim, Y. Lu & J. H. Nolen, “Quantitative propagation of chaos in a bimolecular chemical reaction-diffusion model”, *SIAM J. Math. Anal.* **52** (2020), no. 2, p. 2098-2133.
[37] J. Lukkarinen, “Generation and propagation of chaos in the stochastic Kac model”, 2023, talk at Rutgers.
[38] S. Mei, T. Misiakiewicz & A. Montanari, “Mean-field theory of two-layers neural networks: dimension-free bounds and kernel limit”, in *Proceedings of the Thirty-Second Conference on Learning Theory* (A. Beygelzimer & D. Hsu, eds.), Proceedings of Machine Learning Research, vol. 99, 2019, p. 2388-2464.
[39] Q. H. Nguyen, M. Rosenzweig & S. Serfaty, “Mean-field limits of Riesz-type singular flows”, *Ars Inven. Anal.* (2022), article no. 4 (45 pages).
[40] F. Otto & C. Villani, “Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality”, *J. Funct. Anal.* **173** (2000), no. 3, p. 361-400.

[41] M. Petrache & S. Serfaty, "Next order asymptotics and renormalized energy for Riesz interactions", *J. Inst. Math. Jussieu* **16** (2017), no. 3, p. 501-569.

[42] M. Rosenzweig, "The mean-field approximation for higher-dimensional Coulomb flows in the scaling-critical L^∞ space", *Nonlinearity* **35** (2022), no. 6, p. 2722-2766.

[43] ———, "Mean-field convergence of point vortices to the incompressible Euler equation with vorticity in L^∞", *Arch. Ration. Mech. Anal.* **243** (2022), no. 3, p. 1361-1431.

[44] ———, "On the rigorous derivation of the incompressible Euler equation from Newton's second law", *Lett. Math. Phys.* **113** (2023), no. 1, article no. 13 (32 pages).

[45] ———, "Remarks on the logarithmic Sobolev inequality for the Kuramoto model", 2023, available at https://cmu.box.com/s/s2pmhykuolezwkjn65kt6loa70jt93ul.

[46] M. Rosenzweig & S. Serfaty, "Global-in-time mean-field convergence for singular Riesz-type diffusive flows", *Ann. Appl. Probab.* **33** (2023), no. 2, p. 754-798.

[47] ———, "Sharp commutator estimates of all order for Coulomb and Riesz modulated energies", 2024, https://arxiv.org/abs/2407.15650.

[48] G. Rotskoff & E. Vanden-Eijnden, "Trainability and accuracy of artificial neural networks: an interacting particle system approach", *Commun. Pure Appl. Math.* **75** (2022), no. 9, p. 1889-1935.

[49] N. Rougerie & S. Serfaty, "Higher-dimensional Coulomb gases and renormalized energy functionals", *Pure Appl. Math.* **69** (2016), no. 3, p. 519-605.

[50] É. Sandier & S. Serfaty, "2D Coulomb gases and the renormalized energy", *Ann. Probab.* **43** (2015), no. 4, p. 2026-2083.

[51] S. Serfaty, "Mean field limit for Coulomb-type flows", *Duke Math. J.* **169** (2020), no. 15, p. 2887-2935, appendix with Mitia Duerinckx.

[52] ———, "Gaussian fluctuations and free energy expansion for Coulomb gases at any temperature", *Ann. Inst. Henri Poincaré, Probab. Stat.* **59** (2023), no. 2, p. 1074-1142.

[53] C. Villani, "Trend to equilibrium for dissipative equations, functional inequalities and mass transportation", in *Recent advances in the theory and applications of mass transport*, Contemporary Mathematics, vol. 353, American Mathematical Society, 2004, p. 95-109.

[54] B. Zegarlinski, "Dobrushin uniqueness theorem and logarithmic Sobolev inequalities", *J. Funct. Anal.* **105** (1992), no. 1, p. 77-111.