Time-uniform log-Sobolev inequalities and applications to propagation of chaos

Pierre Monmarché, Zhenjie Ren, Songbo Wang

Electron. J. Probab. **29** (2024), article no. 154, 1–38.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1217

Time-uniform log-Sobolev inequalities and applications to propagation of chaos*

Pierre Monmarché† Zhenjie Ren‡ Songbo Wang§

Abstract

Time-uniform log-Sobolev inequalities (LSI) satisfied by solutions of semi-linear mean-field equations have recently appeared to be a key tool to obtain time-uniform propagation of chaos estimates. This work addresses the more general settings of time-inhomogeneous Fokker–Planck equations. Time-uniform LSI are obtained in two cases, either with the bounded-Lipschitz perturbation argument with respect to a reference measure, or with a coupling approach at high temperature. These arguments are then applied to mean-field equations, where, on the one hand, sharp marginal propagation of chaos estimates are obtained in smooth cases and, on the other hand, time-uniform global propagation of chaos is shown in the case of vortex interactions with quadratic confinement potential on the whole space. In this second case, an important point is to establish global gradient and Hessian estimates, which is of independent interest. We prove these bounds in the more general situation of non-attractive logarithmic and Riesz singular interactions.

Keywords: log-Sobolev inequality; mean field limit; propagation of chaos.
MSC2020 subject classifications: 82C22; 60K35.
Submitted to EJP on January 30, 2024, final version accepted on September 26, 2024.
Supersedes arXiv:2401.07966v3.

1 Introduction

We are interested in families $(m_t)_{t\geqslant 0}$ of probability distributions solving time-inhomogeneous Fokker–Planck equations on $\mathbb{R}^d$ of the form

$$\partial_t m_t = \nabla \cdot (\sigma^2 \nabla m_t - b_t m_t)\,, \tag{1.1}$$

where $\sigma^2 > 0$ and $b_t : \mathbb{R}^d \to \mathbb{R}^d$ for $t \geqslant 0$. This describes the evolution of the law of the diffusion process

$$\mathrm{d}X_t = b_t(X_t)\,\mathrm{d}t + \sqrt{2}\sigma\,\mathrm{d}B_t\,, \tag{1.2}$$

*The work of P. M. is supported by the Project CONVIVIALITY ANR-23-CE40-0003 of the French National Research Agency. The work of Z. R. is supported by Finance For Energy Market Research Centre.
†LJLL & LCT, Sorbonne Université, Paris, France. E-mail: pierre.monmarche@sorbonne-universite.fr
‡LaMME, Université d'Évry, Université Paris-Saclay, Évry, France. E-mail: zhenjie.ren@univ-evry.fr
§CMAP, École polytechnique, IP Paris, Palaiseau, France. E-mail: songbo.wang@polytechnique.edu

where B is a standard d-dimensional Brownian motion. We have particularly in mind McKean–Vlasov equations, where b_t is in fact a function of m_t itself, namely

$$b_t(x) = F(x, m_t)\,, \tag{1.3}$$

for some suitable function F. Other examples are time-integrated McKean–Vlasov equations where $b_t(x) = F\big(x, \int_0^t m_s k_t(\mathrm{d}s)\big)$ for some kernel k_t (as in [7]).

Denoting by $\mathcal{C}^1_{\mathrm{c}}(\mathbb{R}^d)$ the set of compactly supported $\mathcal{C}^1$ functions from $\mathbb{R}^d$ to $\mathbb{R}$, a probability measure μ on $\mathbb{R}^d$ is said to satisfy a log-Sobolev inequality (LSI) with constant $C > 0$ if

$$\forall h \in \mathcal{C}^1_{\mathrm{c}}(\mathbb{R}^d) \text{ with } \int_{\mathbb{R}^d} h^2 \,\mathrm{d}\mu = 1\,, \qquad \int_{\mathbb{R}^d} h^2 \ln(h^2)\,\mathrm{d}\mu \leqslant C \int_{\mathbb{R}^d} |\nabla h|^2 \,\mathrm{d}\mu\,. \tag{1.4}$$

Equivalently, for all probability measure $\nu \in \mathcal{P}(\mathbb{R}^d)$ such that ν is absolutely continuous with respect to μ and $\sqrt{\mathrm{d}\nu/\mathrm{d}\mu} \in \mathcal{C}^1_{\mathrm{c}}$, we have

$$\mathcal{H}(\nu|\mu) \leqslant \frac{C}{4}\mathcal{I}(\nu|\mu)\,,$$

where $\mathcal{H}$, $\mathcal{I}$ are the relative entropy and Fisher information defined respectively as follows:

$$\mathcal{H}(\nu|\mu) \coloneqq \int_{\mathbb{R}^d} \ln \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\,\mathrm{d}\nu\,,$$
$$\mathcal{I}(\nu|\mu) \coloneqq \int_{\mathbb{R}^d} \left|\nabla \ln \frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right|^2 \mathrm{d}\nu\,.$$

We want to determine suitable conditions under which the family $(m_t)_{t\geqslant 0}$ solving (1.1) satisfies a uniform LSI, in the sense that (1.4) holds with $\mu = m_t$ and a constant C independent from t. As will be discussed below in details (in Sections 3 and 4), for McKean–Vlasov equations, this is an important tool to get uniform-in-time Propagation of Chaos (PoC) estimates [16, 23].

The paper is organized as follows. In the rest of this introduction we state our main results concerning time-uniform LSI (Theorem 1.3 and 1.4), which are proven in Section 2. In Section 3 we use them to extend the range of the work [23] of Lacker and Le Flem, obtaining sharp uniform in time PoC for McKean–Vlasov equations in cases of smooth interaction. Section 4 addresses the question of uniform-in-time LSI and PoC for singular (log or Riesz) interactions in $\mathbb{R}^d$.

Before stating our main results, we recall first the following result of Malrieu [26], based on the classical Bakry–Émery approach.

Proposition 1.1. *Assume that there exist $T > 0$, $L \in \mathbb{R}$ such that for all $t \in [0, T]$ and x, $y \in \mathbb{R}^d$,*

$$\big(b_t(x) - b_t(y)\big) \cdot (x - y) \leqslant L|x - y|^2\,, \tag{1.5}$$

and that m_0 satisfies an LSI with constant $C_0 > 0$. Then, for all $t \in [0, T]$, the measure m_t satisfies an LSI with constant

$$C_t = e^{2Lt}C_0 + \sigma^2 \int_0^t e^{2Ls}\,\mathrm{d}s\,.$$

For completeness, the proof is recalled in Section 2.1.

Remark 1.2. When the curvature lower-bound L in (1.5) is negative, this already gives an LSI uniform in t, but we are mostly interested in cases where (1.5) only holds with $L > 0$. Nevertheless, this first proposition means that, in the next results (Theorems 1.3 and 1.4), in fact, if the assumptions are only satisfied for $t \geqslant t_0$ for some $t_0 > 0$ large enough (for instance the condition (1.8)), we can apply Proposition 1.1 for times $t \in [0, t_0]$ and then apply the other results to $(m_{t+t_0})_{t \geqslant 0}$.

The next result addresses the high-diffusivity regime, namely when σ^2 is high enough (see (1.8)). It is proven in Section 2.2.

Theorem 1.3. *Assume that there exist ρ, L, R, $K > 0$ such that, for all $t \geqslant 0$,*

$$\big(b_t(x) - b_t(y)\big) \cdot (x - y) \leqslant \begin{cases} -\rho|x-y|^2 & \forall x, y \in \mathbb{R}^d \text{ with } |x| \geqslant R\,, \\ L|x-y|^2 & \forall x, y \in \mathbb{R}^d\,, \end{cases} \tag{1.6}$$

and, setting $R_* = R(2 + 2L/\rho)^{1/d}$,

$$\sup_{|x| \leqslant R_*} \{-x \cdot b_t(x)\} \leqslant K\,. \tag{1.7}$$

Then, provided m_0 satisfies an LSI and

$$\sigma^2 \geqslant \sigma_0^2 := 2(2L + \rho)\frac{(L + \rho/4)R_*^2 + K}{\rho d}\,, \tag{1.8}$$

the family $(m_t)_{t \geqslant 0}$ satisfies a uniform LSI.

Moreover, there exists $C_* > 0$ which depends on L, R, d and ρ but not on m_0, K nor σ such that, provided (1.8)*, there exists $t_* > 0$ (depending on all previous parameters, in particular m_0) such that the measure m_t satisfies an LSI with constant $\sigma^2 C_*$ for all $t \geqslant t_*$.*

More precisely, for any $\varepsilon > 0$, there exists $\sigma_0' > 0$ which depends only on L, R, d, ρ and ε such that for all $\sigma \geqslant \sigma_0'$, there exists $t_* > 0$ (depending on all previous parameters, in particular m_0) such that the measure m_t satisfies an LSI with constant $\sigma^2(\rho^{-1} + \varepsilon)$ for all $t \geqslant t_*$.

The restrictive high-diffusivity condition (1.8) is inherited from the method of [28] that we extend here to the time-inhomogeneous setting. It is unclear whether this condition is necessary for this method, based on parallel coupling, to work (even in the time-homogeneous case). A stronger result than in [28] is proven in [38], based on reflection coupling and without any constraint on σ^2, but the method crucially relies on a general spectral result from [27, 37], making the adaptation to the time-inhomogeneous case (where the law at time t is not given as the invariant measure of some explicit process) much less clear. Notice that, as detailed in Section 3, one of our main motivation is to apply the results of Lacker and Le Flem in [23] which anyway require a sufficiently large σ (which is then necessary for the result to hold: at small temperature, counter-examples are known where propagation of chaos cannot be uniform in time).

The next result is the adaptation in the time-inhomogeneous settings of the bounded-Lipschitz perturbation argument of [29]. Its proof is given in Section 2.3.

Theorem 1.4. *Assume that, for all $t \geqslant 0$, the drift writes $b_t = a_0 + g_t$ for some a_0, $g_t \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R}^d)$ with bounded derivatives such that the generator $\mathcal{L}_0 = a_0 \cdot \nabla + \sigma^2 \Delta$ admits a unique $\mathcal{C}^2$ invariant probability density μ_0 satisfying an LSI. Write $\tilde{b}_t := 2\nabla \ln \mu_0 - b_t$ and $\varphi_t := -\nabla \cdot g_t + g_t \cdot \nabla \ln \mu_0$. Assume that there exist L, R, M^φ, $L^\varphi \geqslant 0$ and $\rho > 0$ such that,*

- *for all $t \geqslant 0$, we have $\varphi_t = \varphi_{1,t} + \varphi_{2,t}$ with M^φ-bounded $\varphi_{1,t}$ and L^φ-Lipschitz $\varphi_{2,t}$;*

- *for all* $x, y \in \mathbb{R}^d$,

$$\big(\tilde{b}_t(x) - \tilde{b}_t(y)\big) \cdot (x-y) \leqslant \begin{cases} -\rho|x-y|^2 & \text{if } |x-y| \geqslant R, \\ L|x-y|^2 & \text{otherwise.} \end{cases} \tag{1.9}$$

Finally, assume that the logarithmic relative density $u_0 = \ln \mathrm{d}m_0/\mathrm{d}\mu_0$ *exists and is the sum of a bounded and a Lipschitz continuous functions. Then* $(m_t)_{t\geqslant 0}$ *satisfies a uniform LSI.*

Moreover there exists $C_* > 0$ *which depends on* L, R, M^φ, L^φ, σ^2, ρ *and the LSI constant of* μ_0 *but not on* m_0 *such that, for some* $t_* > 0$, m_t *satisfies an LSI with constant* C_* *for all* $t \geqslant t_*$.

Finally, denoting by C_0 *the LSI constant of* μ_0, *the following holds. For any* $\varepsilon > 0$, *there exists* $\eta > 0$ *(which depends only on* ρ, L, R *and* ε*) such that, if* $M^\varphi + L^\varphi \leqslant \eta$, *there exists* $t_* > 0$ *(depending on all previous parameters, in particular* m_0*) such that* m_t *satisfies an LSI with constant* $C_0 + \varepsilon$ *for all* $t \geqslant t_*$.

It may appear that the conditions of Theorem 1.4 are unnatural, in particular for its consideration of the quantities $\tilde{b}_t$, φ_t. Here we only mention that the proof of the theorem is based on the Hamilton–Jacobi–Bellman (HJB) equation (2.12) satisfied by the log-density $\ln \mathrm{d}m_t/\mathrm{d}\mu_0$ of the measure flow, where $\tilde{b}_t$, φ_t appear by computations. This analytical approach, despite being not very compatible with the theory of Markov diffusion processes, turns out to be robust enough for our treatment of singular interactions in Section 4.

2 Proofs of the general results

In this section we write $(P_{s,t})_{t\geqslant s\geqslant 0}$ the inhomogeneous Markov semi-group associated to (1.2), given by

$$P_{s,t}f(x) = \mathbb{E}[f(X_t)|X_s = x]\,.$$

In particular, the solution of (1.1) is then given by $m_t = m_0 P_{0,t}$. In the proofs of Proposition 1.1 and Theorem 1.3, we can additionally assume that b_t is smooth with all derivatives being bounded, and consider functions f which are for instance the sum of a positive constant and a compactly supported smooth non-negative function, which enable to justify the computations based on $\partial_t P_{s,t} f = P_{s,t}\mathcal{L}_t f$ and $\partial_s P_{s,t} f = -\mathcal{L}_s P_{s,t} f$ (using e.g. Proposition B.1). The conclusion is then obtained by approximation (as in e.g. [8]).

2.1 Proof of Proposition 1.1

Proof of Proposition 1.1. Considering X and X' two solutions of (1.2) driven by the same Brownian motion, the condition (1.5) gives

$$\mathrm{d}|X_t - X'_t|^2 \leqslant 2L|X_t - X'_t|^2\,\mathrm{d}t\,, \tag{2.1}$$

so that $|X_t - X'_t|^2 \leqslant e^{2L(t-s)}|X_s - X'_s|^2$ for all $t \geqslant s \geqslant 0$, which by [21] implies

$$|\nabla P_{s,t} f| \leqslant e^{L(t-s)} P_{s,t}|\nabla f|\,. \tag{2.2}$$

(In the initial work [26], the sub-commutation (2.2) is obtained by invoking [2], where it is proven directly through Bakry–Émery interpolations, requiring regularity or integrability assumptions, or approximation arguments to justify the time-differentation along the semi-group. By using a coupling and relying on the more recent [21], this issue is bypassed. Notice that the fact that a coupling bound implies a sub-commutation is the easiest implication in [21]; the converse implication is more involved.) Fix a function

$f \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}_+)$, globally Lipschitz continuous and lower bounded by a positive constant (it is sufficient to prove the LSI with these functions and conclude by approximation). For $t > s \geqslant 0$, we consider the interpolation $\Psi(u) = P_{s,u}(P_{u,t} f \ln P_{u,t} f)$ for $u \in [s,t]$, so that

$$\begin{aligned} P_{s,t}(f \ln f) - P_{s,t} f \ln P_{s,t} f &= \Psi(t) - \Psi(s) \\ &= \int_s^t \Psi'(u) \, \mathrm{d}u \\ &= \sigma^2 \int_s^t P_{s,u} \frac{|\nabla P_{u,t} f|^2}{P_{u,t} f} \, \mathrm{d}u \\ &\leqslant \sigma^2 \int_s^t e^{2L(t-u)} \, \mathrm{d}u P_{s,t} \left(\frac{|\nabla f|^2}{f} \right), \end{aligned}$$

where we used that $(P_{u,t}|\nabla f|)^2 \leqslant P_{u,t}(|\nabla f|^2/f) P_{u,t}(f)$ by Cauchy–Schwarz. Integrating with respect to m_s gives

$$m_t(f \ln f) \leqslant m_s(P_{s,t} f \ln P_{s,t} f) + m_t \left(\frac{|\nabla f|^2}{f} \right) \sigma^2 \int_s^t e^{2L(t-u)} \, \mathrm{d}u \,. \tag{2.3}$$

The proof is concluded by applying this with $s = 0$ and using the LSI for m_0, (2.2) and Cauchy–Schwarz to bound

$$\begin{aligned} m_0(P_{0,t} f \ln P_{0,t} f) &\leqslant m_t f \ln(m_t f) + C_0 m_0 \left(\frac{|\nabla P_{0,t} f|^2}{P_{0,t} f} \right) \\ &\leqslant m_t f m_t(\ln f) + C_0 e^{2Lt} m_t \left(\frac{|\nabla f|^2}{f} \right). \end{aligned}$$

□

2.2 Proof of Theorem 1.3

Proof of Theorem 1.3. The different steps of the proof are the following. First, using the coupling argument of [28] (at high diffusivity), we get a long-time L^2 contraction along the synchronous coupling of two solutions of (1.2). By contrast to the almost sure contraction (2.1), this L^2 contraction is not enough to get an LSI, but it gives a uniform Poincaré inequality following arguments similar to the proof of Proposition 1.1. It remains then to prove a so-called defective LSI, which together with the Poincaré inequality yields the desired LSI. The proof of the defective LSI follows the arguments of [29], except that in the present case the measure for which the LSI is proven is not an invariant measure of a time-homogeneous semi-group (which would solve $\mu = \mu P_t$, which in our case is replaced by $m_t = m_0 P_{0,t}$). These arguments combine a Wang–Harnak inequality for the operator $P_{0,t}$ with a Gaussian moment bound.

Step 1: Poincaré inequality. Let X, X' be two solutions of (1.2) driven by the same Brownian motion. Following the proof of [28, Theorem 1] (which is concerned with time-homogeneous processes, but the proof works readily in the non-homogeneous case under the time-uniform assumptions made in Theorem 1.3), we get for all $t \geqslant s \geqslant 0$,

$$\mathbb{E}\big[|X_t - X'_t|^2\big] \leqslant M e^{-\lambda(t-s)} \mathbb{E}\big[|X_s - X'_s|^2\big] \,,$$

where

$$\lambda = \frac{\rho}{2} \,, \qquad M = 1 + \frac{2(2L+\rho) R_*^2}{4 d \sigma^2} \,. \tag{2.4}$$

This implies, by [21],

$$|\nabla P_{s,t} f|^2 \leqslant M e^{-\lambda(t-s)} P_{s,t} |\nabla f|^2 \,.$$

Since m_0 satisfies a LSI, it satisfies a Poincaré inequality, and thus,

$$m_0(P_{0,t}f)^2 - (m_0P_{0,t}f)^2 \leqslant C_0 m_0|\nabla P_{0,t}f|^2 \leqslant C_0 M e^{-\lambda t} m_t|\nabla f|^2 \,.$$

Besides, for fixed $t \geqslant 0$ and $f \in \mathcal{C}^1(\mathbb{R}^d, \mathbb{R})$ globally Lipschitz continuous, considering the interpolation $\Psi(u) = P_{0,u}(P_{u,t}f)^2$ for $u \in [0,t]$, we get

$$\begin{aligned} P_{0,t}(f^2) - (P_{0,t}f)^2 &= \Psi(t) - \Psi(0) \\ &= \int_0^t \Psi'(u)\,\mathrm{d}u \\ &= \sigma^2 \int_0^t P_{0,u}|\nabla P_{u,t}f|^2\,\mathrm{d}u \\ &\leqslant \sigma^2 \int_0^t M e^{-\lambda(t-u)}\,\mathrm{d}u P_{0,t}|\nabla f|^2 \,. \end{aligned}$$

Combining these last two inequalities, we get

$$\begin{aligned} m_t(f^2) - \big(m_t(f)\big)^2 &= m_0\big(P_{0,t}(f^2) - (P_{0,t}f)^2\big) + m_0(P_{0,t}f)^2 - (m_0P_{0,t}f)^2 \\ &\leqslant M\bigg(\frac{\sigma^2}{\lambda} + e^{-\lambda t}C_0\bigg) m_t|\nabla f|^2 \,, \end{aligned} \tag{2.5}$$

which is a uniform Poincaré inequality for $(m_t)_{t\geqslant 0}$.

Step 2: Gaussian moment. Since m_0 satisfies an LSI, there exists $\delta_0 > 0$ such that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} e^{\delta_0|x-y|^2} m_0(\mathrm{d}x)m_0(\mathrm{d}y) < \infty \,.$$

Write $V(x,y) = e^{\delta|x-y|^2}$ for some $0 < \delta < \min(\delta_0, \rho/5)$ and $\mathcal{L}_{2,t}$ the generator on $\mathbb{R}^d \times \mathbb{R}^d$ of two independent diffusion processes (1.2), namely

$$\mathcal{L}_{2,t}g(x,y) = b_t(x)\cdot\nabla_x + b_t(y)\cdot\nabla_y + \sigma^2\Delta_x + \sigma^2\Delta_y \,.$$

Using (1.6) (and that $|x-y| \geqslant 2R$ implies that either $|x| \geqslant R$ or $|y| \geqslant R$),

$$\begin{aligned} \frac{\mathcal{L}_{2,t}V(x,y)}{V(x,y)} &= 2\delta(x-y)\cdot\big(b_t(x) - b_t(y)\big) + 4\delta d + 8\delta^2|x-y|^2 \\ &\leqslant \begin{cases} 4\delta d + (8\delta^2 - 2\delta\rho)|x-y|^2 & \text{if } |x-y| \geqslant 2R \\ 4\delta d + 4(8\delta^2 + 2\delta L)R^2 & \text{otherwise} \end{cases} \\ &\leqslant -\delta\mathbb{1}_{|x-y|\geqslant R_*} + C_*\mathbb{1}_{|x-y|<R_*} \end{aligned}$$

with

$$R_*^2 = \max\bigg(\frac{1+4d}{2(\rho-4\delta)}, 4R^2\bigg)\,, \qquad C_* = \delta\big(4d + (2L+8\delta)R^2\big)\,.$$

Hence,

$$\mathcal{L}_{2,t}V(x,y) \leqslant -\delta V(x,y) + C_* e^{\delta R_*^2}\,,$$

and thus,

$$\partial_t(m_t\otimes m_t)(V) \leqslant -\delta(m_t\otimes m_t)(V) + C_* e^{\delta R_*^2}\,.$$

As a conclusion, for all $t \geqslant 0$,

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} e^{\delta|x-y|^2} m_t(\mathrm{d}x)m_t(\mathrm{d}y) \leqslant \delta^{-1}C_* e^{\delta R_*^2} + e^{-\delta t}\int_{\mathbb{R}^d\times\mathbb{R}^d} e^{\delta|x-y|^2} m_0(\mathrm{d}x)m_0(\mathrm{d}y)\,. \tag{2.6}$$

Step 3: Wang–Harnack inequality. In the following, fix $f > 0$ such that $m_t f = 1$. Using the Röckner–Wang argument [33] for the diffusion (1.2) we get, for all x, $y \in \mathbb{R}^d$, $\alpha > 1$ and $t > 0$,

$$\big(P_{0,t}f(y)\big)^\alpha \leqslant (P_{0,t}f^\alpha)(x) \exp\left(\frac{\alpha}{2\sigma^2(\alpha-1)}\left(L^2 t + \frac{|x-y|^2}{t}\right)\right), \tag{2.7}$$

so that

$$\int_{\mathbb{R}^d} (P_{0,t}f^\alpha)(x) m_0(\mathrm{d}x) \\ \geqslant \big(P_{0,t}f(y)\big)^\alpha \int_{\mathbb{R}^d} \exp\left(-\frac{\alpha}{2\sigma^2(\alpha-1)}\left(L^2 t + \frac{|x-y|^2}{t}\right)\right) m_0(\mathrm{d}x)\,,$$

and thus, for any $\beta > \alpha$,

$$\begin{aligned} &m_0(P_{0,t}f)^\beta \\ &\quad\leqslant \big(m_t(f^\alpha)\big)^{\beta/\alpha} \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} \exp\left(-\frac{\alpha\big(L^2 t + \frac{|x-y|^2}{t}\big)}{2\sigma^2(\alpha-1)}\right) m_0(\mathrm{d}x)\right]^{-\beta/\alpha} m_0(\mathrm{d}y) \\ &\quad\leqslant \big(m_t(f^\alpha)\big)^{\beta/\alpha} \int_{\mathbb{R}^{2d}} \exp\left(\frac{\beta\big(L^2 t + \frac{|x-y|^2}{t}\big)}{2\sigma^2(\alpha-1)}\right) m_0(\mathrm{d}x) m_0(\mathrm{d}y)\,. \end{aligned} \tag{2.8}$$

Using Jensen's inequality for the probability with density $P_{0,t}f$ with respect to m_0, taking $\alpha = 3/2$ so that $x \mapsto x^{\alpha-1}$ is concave, we get

$$m_t(f^\alpha) = m_0 P_{0,t} f^\alpha \leqslant \big(m_0(P_{0,t}f)^2\big)^{\alpha-1}\,.$$

Using (2.8) with $\beta = 2$ to bound the right hand side then gives

$$m_t(f^\alpha) \leqslant \big(m_t(f^\alpha)\big)^{2/3} \left[\int_{\mathbb{R}^{2d}} \exp\left(\frac{4}{2\sigma^2}\left(L^2 t + \frac{|x-y|^2}{t}\right)\right) m_0(\mathrm{d}x) m_0(\mathrm{d}y)\right]^{1/2}.$$

and we can divide by $(m_t f^\alpha)^{2/3}$ to end up with

$$m_t(f^\alpha) \leqslant \left[\int_{\mathbb{R}^{2d}} \exp\left(\frac{2}{\sigma^2}\left(L^2 t + \frac{|x-y|^2}{t}\right)\right) m_0(\mathrm{d}x) m_0(\mathrm{d}y)\right]^{3/2}.$$

Applying this result to $(m_{t+t_0})_{t\geqslant 0}$ for some $t_0 > 0$ we get that for all $t \geqslant 0$ and all $f > 0$ with $m_{t+t_0} f = 1$,

$$m_{t+t_0}\big(f^{3/2}\big) \leqslant \left[\int_{\mathbb{R}^{2d}} \exp\left(\frac{2}{\sigma^2}\left(L^2 t_0 + \frac{|x-y|^2}{t_0}\right)\right) m_t(\mathrm{d}x) m_t(\mathrm{d}y)\right]^{3/2}.$$

Taking $t_0 = 2/(\delta\sigma^2)$, the right hand side is bounded uniformly in $t \geqslant 0$ thanks to (2.6). As a conclusion, we have determined t_0, $C > 0$ such that

$$\forall t \geqslant t_0,\ \forall f > 0, \qquad m_t\big(f^{3/2}\big) \leqslant C (m_t f)^{3/2}\,. \tag{2.9}$$

Moreover, in view of (2.6), we can find $C > 0$ which depends on m_0 only through δ such that (2.9) holds with this C for all t large enough. To see that we can take δ independent from m_0, we can replace the function V above by the time-dependent $V_t(x,y) = e^{\delta_t |x-y|^2}$

where $t \mapsto \delta_t$ is slowly and smoothly increasing starting from some small $\delta_0 > 0$ (depending on m_0) and reaching $\rho/5$ after some time. Following similar computations as above we get that $(m_t \otimes m_t)(V_t)$ is non-increasing (taking $\mathrm{d}\delta_t/\,\mathrm{d}t$ sufficiently small), from which, replacing $(m_t)_{t\geqslant 0}$ by $(m_{t_0+t})_{t\geqslant 0}$ for some sufficiently large t_0, we can assume that (2.6) holds for $\delta = \rho/5$. As a conclusion, for times large enough, (2.9) holds with a constant C independent from m_0.

Step 4: Conclusion. For $t \geqslant t_0$, applying (2.3) with $s = t - t_0$ gives, for $f > 0$,

$$m_t(f \ln f) \leqslant 2m_s\big(P_{s,t} f \ln(P_{s,t} f)^{1/2}\big) + m_t\left(\frac{|\nabla f|^2}{f}\right)\sigma^2 \int_0^{t_0} e^{2Lu}\,\mathrm{d}u\,.$$

Assume that f is such that $m_t f = 1$. Applying Jensen's inequality twice (first with the probability measure with density $P_{s,t}f$ with respect to m_s) gives

$$m_s\big(P_{s,t} f \ln(P_{s,t} f)^{1/2}\big) \leqslant \ln\big(m_s(P_{s,t} f)^{3/2}\big) \leqslant \ln m_t\big(f^{3/2}\big)\,.$$

Thanks to (2.9), we have thus obtained that for all $t \geqslant t_0$ and all $f > 0$ with $m_t f = 1$,

$$m_t(f \ln f) \leqslant 2\ln C + m_t\left(\frac{|\nabla f|^2}{f}\right)\sigma^2 \int_0^{t_0} e^{2Lu}\,\mathrm{d}u\,, \tag{2.10}$$

which is called a defective LSI (and is uniform over $t \geqslant t_0$). According to [3, Proposition 5.1.3], combining this inequality with the (time uniform) Poincaré inequality (2.5) gives an LSI for m_t uniformly over $t \geqslant t_0$. For $t \in [0, t_0]$ we apply Proposition 1.1, which concludes the proof of the uniform LSI.

Finally, as mentioned above, the constant C may be taken independent from m_0, in which case the defective LSI (2.10) holds for sufficiently large times. Similarly, we see that the Poincaré inequality (2.5) holds with constant $M\sigma^2/\lambda + 1$ (which is independent from m_0) for t large enough. This shows that there exists $C'_* > 0$ independent from m_0 such that m_t satisfies an LSI with constant C'_* for t large enough. The fact that $C'_* \leqslant \sigma^2 C_*$ for some $C_* > 0$ independent from σ can be checked in the explicit expressions above. More precisely, taking $\delta = \rho/5$ and $t_0 = 2/(\delta\sigma^2)$, we get that, in (2.10) the constant C is uniformly bounded over $\sigma \geqslant \sigma_0$ by a constant that depends only on ρ, L, R, d, and similarly we can bound

$$\sigma^2 \int_0^{t_0} e^{2Lu}\,\mathrm{d}u \leqslant \sigma^2 t_0 e^{2Lt_0} \leqslant \frac{10}{\rho} e^{20L/\rho}$$

in (2.10) uniformly over $\sigma \geqslant 1$. As a consequence, for large values of σ^2, the leading term in the LSI constant for large times is $\sigma^2 M/\lambda$ from the Poincaré constant, with M and λ in (2.4). As $\sigma \to \infty$, M goes to 1, so we may take the LSI constant (for large times) to be $\sigma^2(\lambda^{-1} + \varepsilon)$ for any arbitrary $\varepsilon > 0$ for σ large enough. This estimate (with $\lambda = \rho/2$) is not sharp, as we expect an LSI of order σ^2/ρ (which is the Gaussian behavior). This is due to the $1/2$ factor in the definition of λ in [28], which is in fact arbitrary, in the sense that the computations of [28] work if we take $\lambda = \alpha\rho$ for an arbitrary $\alpha < 1$ (see the two first equations of [28, Section 2.1.2]), provided the lower bound on the temperature σ_0^2 is sufficiently large (depending on α). As a conclusion, we can get a Poincaré constant, and thus an LSI constant, equal to $\sigma^2(\rho^{-1} + \varepsilon)$ for an arbitrary ε for large times, provided σ is large enough. □

2.3 Proof of Theorem 1.4

Proof of Theorem 1.4. The proof closely follows the one of [29, Theorem 1] (in the time-homogeneous settings and with $m_0 = \mu_0$, i.e. $u_0 = 0$), the time dependencies appearing

along the proof being dealt with the uniform-in-time assumptions of Theorem 1.4. We recall the main arguments and refer to [29] for details. Starting from

$$\begin{aligned}\partial_t m_t &= -\nabla\cdot(b_t m_t) + \Delta m_t\,,\\ 0 = \partial_t \mu_0 &= -\nabla\cdot(a_0\mu_0) + \Delta\mu_0\,,\end{aligned}$$

we get that $h_t = m_t/\mu_0$ is a viscosity solution to

$$\partial_t h_t = \Delta h_t + \tilde{b}_t\cdot\nabla h_t + \varphi_t h_t\,. \tag{2.11}$$

This gives the Feynman–Kac representation

$$h_t(x) = \mathbb{E}\Big[h_0\big(X_t^{t,x}\big)\exp\Big(\int_0^t \varphi_s\big(X_s^{t,x}\big)\,\mathrm{d}s\Big)\Big]\,,$$

where $X^{t,x}$ solves

$$X_0^{t,x} = x\,,\qquad \mathrm{d}X_s^{t,x} = \tilde{b}_{t-s}\big(X_s^{t,x}\big)\,\mathrm{d}s + \sqrt{2}\,\mathrm{d}B_t \quad \text{for } s\in[0,t].$$

Suppose additionally that φ_t, h_0 and $1/h_0$ are bounded and Lipschitz continuous (the general case being obtained afterwards by an approximation argument, which we omit here, referring to [29]). Then, applying synchronous coupling to the Feynman–Kac formula above, for any $T>0$ we obtain a constant $M>0$ such that for every t, $s\in[0,T]$ and every x, $y\in\mathbb{R}^d$,

$$M^{-1}\leqslant h(t,x)\leqslant M \qquad\text{and}\qquad |h(t,x)-h(s,y)|\leqslant M\big(|t-s|^{1/2}+|x-y|\big)\,.$$

Taking the logarithm we obtain that $u_t \coloneqq \ln h_t$ is a bounded and uniformly continuous viscosity solution to the HJB equation,

$$\partial_t u_t = \Delta u_t + |\nabla u_t|^2 + \tilde{b}_t\cdot\nabla u_t + \varphi_t\,. \tag{2.12}$$

In order to use a stochastic control representation of the solutions of such equations, for $N\in\mathbb{N}$, consider the approximative HJB equation,

$$u_0^N = u_0\,,\qquad \partial_t u_t^N = \Delta u_t^N + \sup_{\alpha:|\alpha|\leqslant N}\{2\alpha\cdot\nabla u_t^N - |\alpha|^2\} + \tilde{b}\cdot\nabla u^N + \varphi_t\,, \tag{2.13}$$

and the associated control problem,

$$V^N(T,x) = \sup_{\nu}\sup_{\alpha:|\alpha_t|\leqslant N}\mathbb{E}\Big[u_0\big(X_T^{\alpha,x}\big) + \int_0^T\Big(\varphi_t\big(X_t^{\alpha,x}\big) - |\alpha_t|^2\Big)\,\mathrm{d}t\Big]\,, \tag{2.14}$$

where $\nu = \big(\Omega, F, (\mathcal{F}_\cdot), \mathbb{P}, (B_\cdot)\big)$ stands for a filter probability space with the usual conditions and an $(\mathcal{F}_\cdot)$-Brownian motion B, α is an $\mathbb{R}^d$-valued progressively measurable process such that $\int_0^T \mathbb{E}\big[|\alpha_t|^m\big]\,\mathrm{d}t$ is finite for every $m\in\mathbb{N}$, and $X^{\alpha,x}$ solves

$$X_0^{\alpha,x} = x\,,\qquad \mathrm{d}X_t^{\alpha,x} = \Big(\tilde{b}\big(X_t^{\alpha,x}\big) + 2\alpha_t\Big)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t\,. \tag{2.15}$$

By Theorem IV.7.1 and the results in Sections V.3 and V.9 of [15], the value function V^N defined by (2.14) is a bounded and uniformly continuous viscosity solution to (2.13).

Suppose $u_0 = \ln(m_0/\mu_0)$ is the sum of an M^{u_0}-bounded and an L^{u_0}-Lipschitz function. As shown in [29, Lemma 8], using a reflection coupling of two solutions of (2.15) with different initial conditions but using the same control α, we get that there exist C', $\kappa>0$,

depending only on ρ, L, R, such that for every x, $y \in \mathbb{R}^d$, $N \in \mathbb{N}$, $T > 0$ and $t > 0$, we have

$$|V^N(T,x) - V^N(T,y)| \leqslant 2M^{\varphi}t + 2M^{u_0}\mathbb{1}_{T<1} + C'\left(\mathbb{1}_{t<T}\frac{M^{\varphi}}{t} + L^{\varphi} + e^{-\kappa T}\left(L^{u_0} + \mathbb{1}_{T\geqslant 1}M^{u_0}\right)\right)|x-y|\,. \quad (2.16)$$

We simply take $t = 1$. Since both u and V^N are bounded and uniformly continuous on $[0,T]\times\mathbb{R}^d$, we can apply the parabolic comparison for viscosity solutions on the whole space [12, Theorem 1] to obtain $V^N(T,x) = u_T(x)$ for N sufficiently large. Therefore, we have obtained that there exists $C > 0$ such that for every $T > 0$ and every x, $y \in \mathbb{R}^d$, we have

$$|u_T(x) - u_T(y)| \leqslant C(1 + |x-y|)\,. \quad (2.17)$$

Besides, in view of (2.16), we can find $C > 0$ independent from m_0 such that (2.17) holds with this C for all T large enough. Moreover this C can be taken arbitrarily small provided $M^{\varphi} + L^{\varphi}$ is small enough.

We can then decompose u_T as the sum of a bounded and a Lipschitz continuous functions (with time uniform bounds for both functions). For instance we can consider a $2C(1+\sqrt{d})$-Lipschitz function v_T that coincides with u_T at all points $x \in \mathbb{Z}^d$ (thanks to (2.17)) and then $u_T - v_T$ is uniformly bounded (thanks to (2.17) again) uniformly in T. The proof is concluded by applying successively the Holley–Stroock and Aida–Shigekawa perturbation lemmas [17, 1]. □

3 Sharp PoC for McKean–Vlasov diffusions

3.1 Settings and notations

In this section, we consider the non-linear McKean–Vlasov equation on $\mathbb{R}^d$:

$$\partial_t m_t = \nabla\cdot\left(\sigma^2\nabla m_t - F(\cdot,m_t)m_t\right), \quad (3.1)$$

which corresponds to (1.1) in the case (1.3). In fact, since we want to apply the results of [23], we consider its settings, which reads

$$F(x,m) = b_0(x) + \int_{\mathbb{R}^d} b(x,y)m(\mathrm{d}y)$$

for some $b_0 : \mathbb{R}^d \to \mathbb{R}^d$ and $b : \mathbb{R}^d\times\mathbb{R}^d \to \mathbb{R}^d$ (which additionally may depend on time in [23], which we don't consider here for simplicity as it is not the case in the examples were are interested in, although it would work similarly). It is associated to the system of interacting particles $\boldsymbol{X} = (X^1,\dots,X^N)$ solving

$$\forall i \in [\![1,N]\!],\qquad \mathrm{d}X_t^i = b_0\big(X_t^i\big)\,\mathrm{d}t + \frac{1}{N-1}\sum_{j\in[\![1,N]\!]\setminus\{i\}} b\big(X_t^i,X_t^j\big)\,\mathrm{d}t + \sqrt{2}\sigma\,\mathrm{d}B_t^i\,, \quad (3.2)$$

where B^1, …, B^N are independent d-dimensional Brownian motions. Denote by m_t^N the law of $\big(X_t^1,\dots,X_t^N\big)$ and by $m_t^{k,N}$ the law of $\big(X_t^1,\dots,X_t^k\big)$ for $k \leqslant N$.

The PoC phenomenon describes the fact that, in the system of interacting particles, as $N\to\infty$, particles become more and more independent, so that $m_t^{k,N}$ converges to $m_t^{\otimes k}$ for a fixed k. Up to recently, known results were typically that, under suitable conditions, for a fixed $t > 0$, $\big\|m_t^{k,N} - m_t^{\otimes k}\big\|_{\mathrm{TV}} = \mathcal{O}\big(\sqrt{k/N}\big)$. This can be for instance obtained by showing the global estimate $\mathcal{H}\big(m_t^N\big|m_t^{\otimes N}\big) = \mathcal{O}(1)$ (which is optimal) using then that $\mathcal{H}\big(m_t^N\big|m_t^{\otimes N}\big) = (N/k)\mathcal{H}\big(m_t^{k,N}\big|m_t^{\otimes k}\big)$ (assuming for simplicity that $n/k \in \mathbb{N}$)

and concluding with Pinsker's inequality. This k/N rate for the marginal relative entropy (hence $\sqrt{k/N}$ in TV) was thought to be optimal until Lacker showed in [22] that it is possible to get a rate k^2/N^2 by working with a BBGKY hierarchy of entropic bounds instead of simply with the full entropy of the N particles system. We refer to such entropic estimates with a rate k^2/N^2 as *sharp* PoC, by comparison with other results (the k^2/N^2 rate being optimal, as it is reached, e.g., in Gaussian cases). The work [22] deals with finite-time intervals, and the technique is then refined by Lacker and Le Flem in [23] to get uniform-in-time sharp PoC in some cases (small interaction in the torus or convex potentials in $\mathbb{R}^d$). A crucial ingredient in their result is a uniform LSI for the solution of the non-linear equation (3.1). Our results can thus be applied to extend their results to more general cases, allowing for instance for non-convex potentials on $\mathbb{R}^d$.

The rest of this section is organized as follows. In Section 3.2 for the reader's convenience we give a brief overview of the general result of Lacker and Le Flem. In Sections 3.3 and 3.4 we apply respectively Theorems 1.4 and 1.3 to get, under suitable conditions, uniform-in-time LSI for solutions of the McKean–Vlasov equation, and thus uniform-in-time sharp PoC as a corollary, in cases which are not covered by [23].

3.2 Lacker and Le Flem's result

First, for the reader's convenience, we recall [23, Theorem 2.1]. There are two sets of assumptions to apply this result: Assumption **E** of [23] is technical conditions related to well-posedness of m and m^N and we omit them as they are not important in our discussion (see Proposition 3.5 below). The second set of assumptions of [23] is the following.

Assumption 3.1 (Assumption **A** of [23])**.** *The following holds.*

1. $(m_t)_{t\geqslant 0}$ *satisfies a uniform LSI with constant* $\eta > 0$*.*
2. $(m_t)_{t\geqslant 0}$ *satisfies a uniform transport inequality: there exists* $\gamma > 0$ *such that, for all* $t \geqslant 0$*,* $x \in \mathbb{R}^d$ *and* $\nu \in \mathcal{P}(\mathbb{R}^d)$*,*
$$\big|\nu\big(b(x,\cdot)\big) - m_t\big(b(x,\cdot)\big)\big|^2 \leqslant \gamma \mathcal{H}(\nu|m_t)\,. \tag{3.3}$$
3. $(m_t)_{t\geqslant 0}$ *and* $\big(m_t^N\big)_{t\geqslant 0}$ *satisfy this uniform* L^2 *boundedness:*
$$\sup_{N\in\mathbb{N}}\sup_{t\geqslant 0}\int_{\mathbb{R}^{dN}} \big|b(x_1,x_2) - m_t\big(b(x_1,\cdot)\big)\big|^2 m_t^N(\mathrm{d}x) < \infty\,. \tag{3.4}$$

When b is bounded, (3.4) is trivial and (3.3) follows from Pinsker's inequality. When $y \mapsto b(x,y)$ is Lipschitz continuous uniformly in x, (3.4) follows from time-uniform second moment bounds, which are classically obtained by Lyapunov arguments, and (3.3) is implied by the uniform LSI.

Theorem 3.2 (From Theorem 2.1 of [23])**.** *Under Assumptions* ***A*** *and* ***E*** *of [23], assume moreover that* $\sigma^4 > 8\gamma\eta$ *and that*
$$\exists C_0 > 0,\ \forall N \geqslant 2,\ \forall k \in [\![1,N]\!], \qquad \mathcal{H}\big(m_0^{k,N}\big|m_0^{\otimes k}\big) \leqslant C_0\frac{k^2}{N^2}\,. \tag{3.5}$$
Then,
$$\exists C > 0,\ \forall N \geqslant 2,\ \forall k \in [\![1,N]\!],\ \forall t \geqslant 0, \qquad \mathcal{H}\big(m_t^{k,N}\big|m_t^{\otimes k}\big) \leqslant C\frac{k^2}{N^2}\,. \tag{3.6}$$

Remark 3.3. As in Remark 1.2, it is in fact sufficient to enforce Assumption **A** with the condition $\sigma^4 > 8\gamma\eta$ for times $t \geqslant t_0$ for some t_0 and apply Theorem 3.2 to $(m_{t+t_0})_{t\geqslant 0}$. More precisely, for some t_0, assume that (3.3) and (3.4) holds uniformly over $t \in [0,t_0]$. Then, assuming the initial chaos (3.5), [22, Theorem 2.2] gives (3.6) for some constant $C > 0$ uniformly over $t \in [0,t_0]$. In particular, the initial chaos (3.5) holds for $(m_{t+t_0})_{t\geqslant 0}$.

In [23], the assumptions of Theorem 3.2 are shown to hold in two cases: either convex potentials on $\mathbb{R}^d$, or models on the torus. In any cases, the condition $\sigma^4 > 8\gamma\eta$ (corresponding to $r_c > 1$ with the notation of [23, Theorem 2.1]) means that the PoC estimates require that either the temperature σ^2 is high enough or the strength of the interaction is small enough. In Sections 3.3 and 3.4 we extend the range of application of [23, Theorem 2.1] to some cases with non-convex potentials on $\mathbb{R}^d$.

Before that, in order to focus on the uniform LSI afterwards, let us state a result concerning the other technical conditions, which is sufficient for the cases considered in the two next sections.

Assumption 3.4. *The initial conditions m_0 and m_0^N have finite moments of all orders, m_0^N is exchangeable and there exists C independent from N such that $\int_{\mathbb{R}^d} |x_1|^2 m_0^{1,N}(\mathrm{d}x_1) \leqslant C$.*

We omit the proof of the next result, the arguments are the same as in [23, Corollary 2.7].

Proposition 3.5. *Assume that b_0 and b are $\mathcal{C}^1$, that $|b_0|$ grows at most polynomially, that b is the sum of a bounded and a Lipchitz continuous function, and that there exist c, $C > 0$ such that for all x, $y \in \mathbb{R}^d$,*

$$\big(b_0(x) + b(x,y)\big) \cdot x \leqslant -c|x|^2 + C(1 + |y|)\,.$$

*Then, under Assumption 3.4, $(m_t)_{t\geqslant 0}$ and $\left(m_t^N\right)_{t\geqslant 0}$ are well defined and Assumption **E** of [23] and the uniform L^2 boundedness* (3.4) *holds.*

3.3 Convergent trajectories

In this section we focus on the cases where m_t converges as $t \to \infty$ towards a stationary solution m_* of the non-linear equation (3.1). This is known to hold in various cases of interest, like the granular media equation with convex potentials, or repulsive interaction, or high temperature, or small interaction, or other models like the adaptive biasing force method [24] or the mean-field gradient descent ascent [25]. So, assume that

$$\|m_t - m_*\|_{\mathrm{TV}} \underset{t\to\infty}{\longrightarrow} 0\,. \tag{3.7}$$

Remark 3.6. Under suitable conditions, [32, Theorem 4.1] allows to obtain (3.7) from a $\mathcal{W}_2$ convergence.

We now discuss suitable conditions to apply Theorem 1.4 with M^φ, L^φ arbitrarily small for large times, where we decompose the drift $F(x, m_t) = a_0(x) + g_t(x)$ with $a_0(x) = F(x, m_*)$ and $g_t(x) = F(x, m_t) - F(x, m_*)$. For simplicity we focus on the case where

$$F(x,m) = -\nabla V(x) - \int_{\mathbb{R}^d} \nabla_x W(x,y) m(\mathrm{d}y)\,, \tag{3.8}$$

for some $V \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$ and $W \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R})$. The next result would be easily adapted to other cases where the density of the stationary solutions of (3.1) are explicit or solve an explicit fixed-point equation (namely when the invariant measure of $\sigma^2\Delta + F(\cdot, m)\cdot\nabla$ is explicit for each m), which is for instance the case in [24, 25].

Proposition 3.7. *Let $(m_t)_{t\geqslant 0}$ be a solution to* (3.1) *(in the case* (3.8)*) which converges in TV in long time towards a stationary solution m_*. Assume that m_0 admits a density e^{u_0} with respect to m_*, with u_0 being the sum of a bounded and a Lipschitz continuous function. Assume furthermore that there exists L, $\alpha > 0$ such that, for all x, $y \in \mathbb{R}^d$,*

$$|\Delta_x W(x,y)| \leqslant L\,, \qquad |\nabla_x W(x,y)| \leqslant \frac{L}{1 + |x-y|^\alpha}\,, \qquad |\nabla V(x)| \leqslant L(1 + |x|^\alpha)\,. \tag{3.9}$$

Finally, assume that V is strongly convex outside of a compact set. Then, $(m_t)_{t\geqslant 0}$ satisfies a uniform LSI. Moreover, as $t\to\infty$, the optimal LSI constant of m_t converges to the optimal LSI constant of m_*.

Notice that, V being strongly convex outside a compact set, the last condition of (3.9) can only hold with some $\alpha\geqslant 1$. Hence, the second condition of (3.9) on ∇W means that we only consider local interactions.

Proof. Considering the decomposition $F(x,m_t)=a_0(x)+g_t(x)$ with $a_0(x)=F(x,m_*)$ and $g_t(x)=F(x,m_t)-F(x,m_*)$, we have to show that Theorem 1.4 applies to $(m_{t+t_0})_{t\geqslant 0}$ with M^φ, L^φ arbitrarily small provided t_0 is large enough. Indeed, the last part of Theorem 1.4 will then give that, for any $\varepsilon>0$, the optimal LSI constant of m_t is less than $C_0+\varepsilon$ for t large enough, where C_0 is the optimal LSI constant of m_*. On the other hand, for any $\varepsilon>0$, there exists a non-constant $\mathcal{C}^\infty$ function f with compact support such that

$$m_*(f^2\ln f^2)-m_*(f^2)\ln m_*(f^2)\geqslant (C_0-\varepsilon)m_*|\nabla f|^2\,.$$

The weak convergence implied by (3.7) leads to

$$m_t(f^2\ln f^2)-m_t(f^2)\ln m_t(f^2)\geqslant (C_0-2\varepsilon)m_t|\nabla f|^2$$

for t large enough, which implies that the optimal LSI constant of m_t is larger than $C_0-2\varepsilon$.

Hence, we turn to the application of Theorem 1.4 using its notations. We write $m\star W(x)=\int_{\mathbb{R}^d}W(x,y)m(\mathrm{d}y)$. The invariant measure of $a_0\cdot\nabla+\sigma^2\Delta$ is $\mu_0=m_*$, with $\nabla\ln m_*=-\nabla(V+m_*\star W)=F(\cdot,m_*)$, so that

$$\tilde{b}_t(x)=-\nabla(V+2m_*\star W-m_t\star W)\,.$$

Since $\nabla_x W$ is bounded by (3.9), the contribution of W in $\tilde{b}_t$ is bounded (uniformly in t) and thus (1.9) holds thanks to the convexity of V outside a compact set. From (3.9),

$$|\nabla\cdot g_t(x)|=|(m_t-m_*)\star\Delta_x W(x)|\leqslant L\|m_t-m_*\|_{\mathrm{TV}}\,,$$

and, given (Y,Y') an optimal TV coupling of m_t and m_* and using the Cauchy–Schwarz inequality,

$$\begin{aligned}
&|g_t(x)\cdot\nabla\ln m_*(x)|\\
&\quad\leqslant\big|\mathbb{E}[\nabla_x W(x,Y)-\nabla_x W(x,Y')]\big|L(2+|x|^\alpha)\\
&\quad\leqslant\mathbb{E}\left[\mathbb{1}_{Y\neq Y'}\left(\frac{1}{1+|x-Y|^\alpha}+\frac{1}{1+|x-Y'|^\alpha}\right)\right]L^2(2+|x|^\alpha)\\
&\quad\leqslant\|m_t-m_*\|_{\mathrm{TV}}^{1/2}\,\mathbb{E}\left[\left(\frac{1}{1+|x-Y|^\alpha}+\frac{1}{1+|x-Y'|^\alpha}\right)^2\right]^{1/2}L^2(2+|x|^\alpha)\,.
\end{aligned}$$

Then we bound, for the first term in the expection,

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{(1+|x-Y|^\alpha)^2}\right]&\leqslant\frac{1}{(1+|x/2|^\alpha)^2}+\mathbb{P}[|Y|\geqslant|x|/2]\\
&\leqslant\frac{1}{(1+|x/2|^\alpha)^2}+\frac{1+\mathbb{E}[|Y|^{2\alpha}]}{1+|x/2|^{2\alpha}}\,,
\end{aligned}$$

and similarly for the second term involving m_*. Using that V is convex outside a compact set and that $\nabla_x W$ is bounded we easily get by Lyapunov arguments that the

moments of m_t are bounded uniformly in time. As a consequence, we have obtained, for $\varphi_t := -\nabla \cdot g_t + g_t \cdot \nabla \ln \mu_0$, a bound

$$\|\varphi_t\|_\infty \leqslant L' \|m_t - m_*\|_{\mathrm{TV}}^{1/2}$$

for some L' independent from t. The TV convergence (3.7) concludes the proof. □

Corollary 3.8. *Under Assumption 3.4 and the settings of Proposition 3.7, assume furthermore that W is bounded and $V = V_1 + V_2$ where V_1 is ρ-strongly convex and V_2 is bounded. Assume that*

$$\sigma^2 > \frac{4}{\rho} \|\nabla_x W\|_\infty^2 \exp\left(\frac{\|V_2\|_\infty + \|W\|_\infty}{\sigma^2} \right). \tag{3.10}$$

Then, provided the initial PoC (3.5) *holds, so does the uniform in time sharp PoC* (3.6)*.*

This applies to cases on $\mathbb{R}^d$ where V is not convex, which are not covered by [23]. In general cases where V may have several local minima, a condition in the spirit of (3.10), that states that either temperature is large enough or interaction is small enough, is necessary to have a uniform-in-time propagation of chaos estimate.

Proof. The assumptions of Proposition 3.7 imply those of Proposition 3.5. Since $\nabla_x W$ is bounded, Pinsker's inequality gives the transport inequality (3.3) with $\gamma = \|\nabla_x W\|_\infty^2/2$. Proposition 3.7 provides the uniform LSI for $(m_t)_{t \geqslant 0}$. Moreover, for large times, the LSI constant of m_t converges to the LSI constant C_* of m_*, which by the Bakry–Émery and Holley–Stroock results is less than $\sigma^2 \rho^{-1} \exp\big((\|V_2\|_\infty + \|W\|_\infty)/\sigma^2\big)$. Corollary 3.8 thus follows from Theorem 3.2 (since, as noticed in Remark 3.3, the condition $\sigma^4 > 8\gamma\eta$ only has to be verified for sufficiently long times). □

3.4 High temperature regime

Instead of Corollary 3.8, using Theorem 1.3, we can get an alternative result, which doesn't require the a priori knowledge that m_t converges in large time and with weaker assumptions on W, but which only works at high temperature and is less explicit (an explicit condition on σ^2 can be obtained in principle by checking the proofs, but it wouldn't be as nice as (3.10)). In the next statement we consider a solution $(m_t)_{t \geqslant 0}$ of (3.1) in the case (3.8).

Proposition 3.9. *Under Assumption 3.4, assume furthemore that $|\nabla U|$ grows at most polynomially, that there exist ρ, L, $R > 0$ such that, for all $z \in \mathbb{R}^d$, $\psi_z := -\nabla U - \nabla_x W(\cdot, z)$ satisfies*

$$\big(\psi_z(x) - \psi_z(y)\big) \cdot (x - y) \leqslant \begin{cases} -\rho |x-y|^2 & \forall x, y \in \mathbb{R}^d \text{ with } |x| \geqslant R\,, \\ L|x-y|^2 & \forall x, y \in \mathbb{R}^d\,, \end{cases} \tag{3.11}$$

and that $\nabla_x W = F_1 + F_2$ where F_1 is bounded and $y \mapsto F_2(x, y)$ is L_W-Lipschitz with $8L_W^2 < \rho$, uniformly in x. Then, there exists $\sigma_*^2 > 0$ (which depends on U, W and d) such that, assuming $\sigma^2 \geqslant \sigma_*^2$ and the initial sharp PoC (3.5)*, we have that the uniform in time sharp PoC* (3.6) *holds.*

In particular, if U is strongly convex outside a compact set and $x \mapsto W(x, z)$ is convex for all z with $\nabla_x W$ being bounded, then Proposition 3.9 applies, without any further smallness condition on the interaction. For instance, with $W(x, z) = a\sqrt{1 + |x - z|^2}$, it applies for any $a > 0$. However, in that case, the temperature threshold σ_*^2 in Proposition 3.9 will depend on a and will become large when a is large (i.e. when the interaction is strong).

Proof. We verify the conditions of Theorem 3.2. Using (3.11) with $y = 0$ we see that Proposition 3.5 holds. The uniform LSI in the high temperature regime $\sigma^2 \geqslant \sigma_0^2$ is ensured by Theorem 1.3, and for times large enough it holds with a constant $\eta = \sigma^2\eta'$ for some $\eta' > 0$ independent from σ, and which can be taken arbitrarily close to $1/\rho$ for σ^2 large enough. Here we have used that $\sup\{-x \cdot b_t(x) : |x| \leqslant R_*\}$ can be bounded by a constant K independent from t and such that (1.8) holds for σ large enough (for t large enough). Indeed, we can bound

$$|b_t(x)| \leqslant |\nabla U(x)| + \|F_1\|_\infty + |F_2(x,0)| + L_W \int_{\mathbb{R}^d} |y| m_t(\mathrm{d}y)\,.$$

Then, the condition (3.11) implies that $s_t \coloneqq \int_{\mathbb{R}^d} |y|^2 m_t(\mathrm{d}y)$ satisfies $\mathrm{d}s_t/\,\mathrm{d}t \leqslant -\rho s_t/2 + q + 2d\sigma^2$ for some $q > 0$ independent from t and σ^2. From this, for t large enough, we get $\int_{\mathbb{R}^d} |y| m_t(\mathrm{d}y) \leqslant C(1+\sigma)$ where C depends only on d, ρ, L, R. As a consequence, in (1.7) we can take $K = C'(1+\sigma)$ for some C' (independent from t and σ), so that (1.8) holds for σ large enough, as claimed.

It remains to check the transport inequality (3.3). For any $\theta > 0$ we can bound, for all $t \geqslant 0$, $x \in \mathbb{R}^d$ and $\nu \in \mathcal{P}(\mathbb{R}^d)$,

$$\begin{aligned}
&\big|\nu\big(b(x,\cdot)\big) - m_t\big(b(x,\cdot)\big)\big|^2 \\
&\quad\leqslant (1+\theta)\big|\nu\big(F_1(x,\cdot)\big) - m_t\big(F_1(x,\cdot)\big)\big|^2 + (1+\theta^{-1})\big|\nu\big(F_2(x,\cdot)\big) - m_t\big(F_2(x,\cdot)\big)\big|^2 \\
&\quad\leqslant (1+\theta)\|F_1\|_\infty^2 \|\nu - m_t\|_{\mathrm{TV}}^2 + (1+\theta^{-1}) L_W^2 \mathcal{W}_2^2(\nu, m_t) \\
&\quad\leqslant \gamma \mathcal{H}(\nu|m_t)\,,
\end{aligned}$$

where we used Pinsker's and Talagrand's inequalities, and γ on the last line is defined by

$$\gamma = \frac{1+\theta}{2}\|F_1\|_\infty^2 + \sigma^2\eta'(1+\theta^{-1})L_W^2\,.$$

Fixing θ (independent from σ) large enough so that $8(1+\theta^{-1})L_W^2 < \rho$, the condition $\sigma^4 > 8\gamma\eta$ holds for σ large enough, which concludes. □

4 Application to log and Riesz interactions

In this section, we still consider McKean–Vlasov equations (3.1), but now we impose the following condition on the non-linear drift.

Assumption 4.1. *We have $d \geqslant 2$, $s \in [0, d-1)$ and the McKean–Vlasov drift in* (3.1) *reads*

$$F(x,m) = -\nabla U(x) + M\nabla g_s \star m(x)\,,$$

where U, M, g_s satisfy the following conditions:

- *the function $U : \mathbb{R}^d \to \mathbb{R}$ has bounded Hessian $\nabla^2 U \in L^\infty$ and satisfies the weak convexity condition: there exist $\kappa_U > 0$ and $R \geqslant 0$ such that for all x, $y \in \mathbb{R}^d$ with $|x-y| \geqslant R$, we have*

$$\langle \nabla U(x) - \nabla U(y), x - y\rangle \geqslant \kappa_U |x-y|^2\,;$$

- *$g_s : \mathbb{R}^d \to \mathbb{R}$ is the logarithmic or Riesz potential:*

$$g_s(x) = \begin{cases} -\ln|x| & \text{when } s = 0, \\ |x|^{-s} & \text{when } s > 0; \end{cases}$$

- *in the sub-Coulombic case where $s < d-2$, M is a $d \times d$ real matrix such that $M : \nabla^2 g(x) \geqslant 0$ for $x \neq 0$; in the Coulombic and the super-Coulombic cases where $s \in [d-2, d-1)$, M is anti-symmetric.*

These models have raised a high interest over the recent years, in particular with a series of work by Rosenzweig, Serfaty and coauthors on the one hand (see e.g. [34, 9, 35] and references within) and Bresch, Jabin, Wang and coauthors on the other hand (see e.g. [20, 5, 4] and references within). The main result of the section, to be stated in Theorem 4.12 in Section 4.4, addresses the McKean–Vlasov drift force above with $d \geqslant 2$, $s = 0$, M being anti-symmetric and U being isotropically quadratic. We show that in this case the dynamics exhibits the time-uniform propagation of chaos. This result is a continuation of a recent work of Guillin, Le Bris and one of the author [16], where the uniform PoC is shown for the dynamics on the torus (thus in a periodic setting). We also note that a non-time-uniform result on the whole space have also been obtained very recently by Feng and Wang [14]. In terms of methodology, the main addition of our work is that we employ the reflection coupling technique of Conforti [10] to get regularity bounds for the mean field flow on the whole space (Theorems 4.6 and 4.11), which enable to apply the Jabin–Wang method.

We will write $g = g_s$ if that does not lead to ambiguities. For simplicity, we also set $\sigma = 1$ in this section. Under the assumptions above, we denote $K = M\nabla g$, and the McKean–Vlasov dynamics writes

$$\partial_t m_t = \Delta m_t - \nabla \cdot \big(m_t (K \star m_t - \nabla U)\big) \,. \tag{4.1}$$

Note that the interaction kernel K is divergence-free when the matrix M is anti-symmetric.

Consider now the system of N particles in interaction:

$$\mathrm{d}X_t^i = -\nabla U\big(X_t^i\big)\,\mathrm{d}t + \frac{1}{N-1} \sum_{j \in [\![1,N]\!] \setminus \{i\}} K\big(X_t^i - X_t^j\big)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t^i \,, \quad i = 1, \dots, N, \tag{4.2}$$

where W_t^i are N independent Brownian motions. The flow of probabilities $m_t^N = \mathrm{Law}(\boldsymbol{X}_t) = \mathrm{Law}(X_t^1, \dots, X_t^N)$ in $\mathbb{R}^{dN}$ satisfies the Fokker–Planck equation at least formally:

$$\partial_t m_t^N = \sum_{i=1}^N \left(\Delta_i m_t^N - \nabla_i \cdot \left(\Big(\frac{1}{N-1} \sum_{j \in [\![1,N]\!] \setminus \{i\}} K(x_i - x_j) - \nabla U(x^i) \Big) m_t^N \right) \right). \tag{4.3}$$

In this section, η^ε denotes a $\mathcal{C}^\infty$ mollifier with support in $B(0,\varepsilon)$ that is also invariant by rotation. We set $g^\varepsilon \coloneqq g \star \eta^\varepsilon$ and $K^\varepsilon \coloneqq M\nabla g^\varepsilon = M\nabla g \star \eta^\varepsilon$. Since under Assumption 4.1, we are restricted to the case where $s < d-1$, the interaction potential $g \propto |x|^{-s}$ is integrable around zero, so g^ε is infinitely differentiable with bounded derivatives. Notice that the rotational invariance of η^ε implies that the value $g^\varepsilon(x)$ depends only on $|x|$ and thus, $\nabla g^\varepsilon(x)$ is parallel to x. We also work with the approximation of the confinement $U^\varepsilon \coloneqq U \star \eta^\varepsilon$.

Sometimes, in the rest of this section, for conciseness, we write $A \lesssim B$ when there exists a constant C such that $A \leqslant CB$.

4.1 Well-posedness of the mean field and particle systems

For a function $f : \mathbb{R}^d \to \mathbb{R}$ and $\theta \in (0,1]$, we denote the homogeneous θ-Hölder (semi-)norm of f by

$$[f]_{\mathcal{C}^\theta} = \sup_{x,y \in \mathbb{R}^d \,:\, x \neq y} \frac{|f(x) - f(y)|}{|x-y|^\theta} \,.$$

In order to study the singular interaction kernel K, we use the following crucial estimate. This generalizes the estimate in (2.9) of [34] (which corresponds to the case $p = \infty$). We refer readers to Lemma 4.5.4 and Theorem 4.5.10 of [18] for the proof, where the statement of the latter should be accompanied with an interpolation.

Proposition 4.2. *Let $s > 0$. For all $m \in L^1 \cap L^p(\mathbb{R}^d)$ with $\left(1 - \frac{s}{d}\right)^{-1} < p \leqslant \infty$, we have*

$$\left\| |\cdot|^{-s} \star m \right\|_{L^\infty} \lesssim \|m\|_{L^1}^{1-qs/d} \|m\|_{L^p}^{qs/d},$$

where $p^{-1} + q^{-1} = 1$. If additionally, for some $\theta \in (0,1)$, we have $\left(1 - \frac{s+\theta}{d}\right)^{-1} < p \leqslant \infty$, then

$$\left[|\cdot|^{-s} \star m \right]_{\mathcal{C}^\theta} \lesssim \|m\|_{L^1}^{1-q(s+\theta)/d} \|m\|_{L^p}^{q(s+\theta)/d}.$$

Then, we present the well-posedness results for the mean field and the particle system.

Proposition 4.3 (Well-posedness of the mean field system)**.** *Let Assumption 4.1 hold. Then we have the following results:*

- *For each initial value $m_0 \in L^1 \cap L^\infty \cap \mathcal{P}(\mathbb{R}^d)$, there exists a unique solution to the mean field flow (4.1) in $C\big([0,\infty); L^1(\mathbb{R}^d) \cap \mathcal{P}\big) \cap L^\infty\big([0,\infty); L^\infty(\mathbb{R}^d)\big)$ depending continuously on the initial value. In particular, we have the time-uniform bound:*

$$\sup_{t \in [0,\infty)} \|m_t\|_{L^\infty} \leqslant C_1(U, \|m_0\|_{L^\infty}) < \infty. \tag{4.4}$$

- *If additionally the initial value m_0 has finite k-th moment for some $k > 0$, then the mean field flow m_t has finite k-th moment, uniformly in time:*

$$\sup_{t \in [0,\infty)} \int_{\mathbb{R}^d} |x|^k m_t(\mathrm{d}x) \leqslant C_2\left(U, K, k, \|m_0\|_{L^\infty}, \int_{\mathbb{R}^d} |x|^k m_0(\mathrm{d}x)\right)$$

- *Finally, let $K^\varepsilon = K \star \eta^\varepsilon$, $U^\varepsilon = U \star \eta^\varepsilon$ be the mollified kernel and confinement. If m_0^ε converges to m_0 in L^1 and if $\sup_\varepsilon \|m_0^\varepsilon\|_{L^\infty} < \infty$, then the solution m_t^ε of the approximate mean field flow*

$$\partial_t m_t^\varepsilon = \Delta m_t^\varepsilon - \nabla \cdot \left(m_t^\varepsilon (K^\varepsilon \star m_t^\varepsilon - \nabla U^\varepsilon) \right) \tag{4.5}$$

  converges to m_t in L^1 for all $t \geqslant 0$. Moreover, the L^∞ norm and the k-th moment bounds above hold when we replace m by m^ε.

Proposition 4.4 (Well-posedness of the particle system)**.** *Let Assumption 4.1 hold with $s \leqslant d-2$ and suppose that for all $x \in \mathbb{R}^d$, we have $x^\top M x \leqslant 0$. Then, for any initial value $\boldsymbol{X}_0 = \left(X_0^1, \ldots, X_0^N\right)$ such that $X_0^i \neq X_0^j$ almost surely for $i \neq j$, the SDE system (4.2) has a global unique strong solution. Moreover, setting $K^\varepsilon = K \star \eta^\varepsilon$, $U^\varepsilon = U \star \eta^\varepsilon$, and considering the approximate SDE system*

$$\mathrm{d}X_t^{\varepsilon,i} = -\nabla U^\varepsilon\left(X_t^{\varepsilon,i}\right) \mathrm{d}t + \frac{1}{N-1} \sum_{j \in [\![1,N]\!] \setminus \{i\}} K^\varepsilon\left(X_t^{\varepsilon,i} - X_t^{\varepsilon,j}\right) \mathrm{d}t + \sqrt{2}\, \mathrm{d}W_t^i, \tag{4.6}$$

for $i \in [\![1,N]\!]$, with the initial condition $X_0^{\varepsilon,i} = X_0^i$, we have, for all $t \geqslant 0$ and $i = 1, \ldots, N$,

$$X_t^{\varepsilon,i} \to X_t^i \text{ a.s.}, \qquad \text{when } \varepsilon \to 0.$$

These results may be considered mathematical folklore and we do not claim originality from them. Their proofs are postponed to Appendix A.

4.2 Uniform Lipschitz and Hessian bounds, and LSI

We introduce the invariant measure μ_0 of the reversible diffusion generated by $\Delta - \nabla U \cdot \nabla$, whose density is explicit:

$$\mu_0(x) = Z(\mu_0)^{-1} \exp\bigl(-U(x)\bigr)\,, \quad Z(\mu_0) = \int_{\mathbb{R}^d} \exp\bigl(-U(x)\bigr)\,\mathrm{d}x\,.$$

Note that, under Assumption 4.1, using the HJB flow method of Conforti (see Theorem 1.3 and Remark 1.7 of [11]), we can show that the measure μ_0 is the image of a Gaussian measure under a transport mapping with an explicit Lipschitz constant, and thus satisfies an LSI with an explicit constant.

We use the following result on the Lipschitz and Hessian bounds on the solution to a class of HJB equations.

Theorem 4.5. *Let $T > 0$. Let $u \in \mathcal{C}^{1,2}_{\mathrm{p}}([0,T] \times \mathbb{R}^d; \mathbb{R})$ be a classical solution to the HJB equation*

$$\partial_t u_t = \Delta u_t - |\nabla u_t|^2 + \tilde{b}_t \cdot \nabla u_t + \varphi_t\,,$$

for some $\tilde{b} \in \mathcal{C}^{0,2}_{\mathrm{p}}([0,T] \times \mathbb{R}^d; \mathbb{R}^d)$, $\varphi \in \mathcal{C}^{0,2}_{\mathrm{p}}([0,T] \times \mathbb{R}^d; \mathbb{R})$. Suppose the initial condition $u_0 \in \mathcal{C}^3_{\mathrm{Lip}}(\mathbb{R}^d; \mathbb{R})$. Suppose the drift $\tilde{b}$ satisfies the weak convexity condition

$$\bigl(\tilde{b}_t(x) - \tilde{b}_t(y), x - y\bigr) \leqslant \kappa_{\tilde{b}}(|x-y|)|x-y|$$

for some $\mathcal{C}^1$-continuous function $\kappa_{\tilde{b}} : (0,\infty) \to \mathbb{R}$ such that $\int_0^1 r\bigl(\kappa_{\tilde{b}}(r) \vee 0\bigr)\,\mathrm{d}r < \infty$ and $\liminf_{r\to\infty} \kappa_{\tilde{b}}(r) < 0$. Suppose $\sup_{t\in[0,T]} \|\nabla \tilde{b}_t\|_{L^\infty} < \infty$. Then, we have the following results:

- *If $\varphi_t \in L^\infty$ for all $t \in [0,T]$, then, we have, for all $t \in [0,T]$,*

$$\|\nabla u_t\|_{L^\infty} \leqslant Ce^{-ct}\|\nabla u_0\|_{L^\infty} + \int_0^t \frac{Ce^{-cv}}{\sqrt{v \wedge 1}} \|\varphi_{t-v}\|_{L^\infty}\,\mathrm{d}v\,, \tag{4.7}$$

  where C, $c > 0$ and depend only on $\kappa_{\tilde{b}}$.

- *If additionally, $\nabla\varphi_t \in L^\infty$ for all $t \in [0,T]$, then we have, for all $t \in [0,T]$,*

$$\begin{aligned}\|\nabla^2 u_t\|_{L^\infty} \leqslant{}& \frac{C'e^{-c't}}{\sqrt{t \wedge 1}}\|\nabla u_0\|_{L^\infty} \\ &+ \int_0^t \frac{C'e^{-c'v}}{\sqrt{v \wedge 1}} \bigl(\|\nabla\varphi_{t-v}\|_{L^\infty} + \|\nabla\tilde{b}_{t-v} \cdot \nabla u_{t-v}\|_{L^\infty}\bigr)\,\mathrm{d}v\,,\end{aligned} \tag{4.8}$$

  where C', $c' > 0$ and depend only on $\kappa_{\tilde{b}}$, $\|\nabla u_0\|_{L^\infty}$ and $\sup_{t\in[0,T]} \|\varphi_t\|_{L^\infty}$.

The theorem is only an enhancement to the result of Conforti [10] by using the short-time gradient estimates obtained by Priola and Wang [31], and by Porretta and Priola [30]. Thus we only provide a sketch of proof here.

Sketch of proof of Theorem 4.5. The stated results differ from the main result of [10], i.e. Theorem 1.3 of it, only in two aspects: first, we work in a time-non-homogeneous setting; and second, the uniform gradient estimate that we utilize in the proof has explosion $t^{-1/2}$ instead of t^{-1} when $t \to 0$.

Following the method in the proof of Theorem 1.4 (and ignoring technical issues about the correspondance to stochastic control problems), for every x, $y \in \mathbb{R}^d$ and

$t \in [0,T]$, we can find stochastic processes $X^{\alpha,x}_\cdot$, $X^{\alpha,y}_\cdot$, α, all defined on $[0,t]$ and taking values in $\mathbb{R}^d$, such that

$$\begin{aligned} X_0^{\alpha,z} = z\,, \qquad \mathrm{d}X_v^{\alpha,z} &= \Big(\tilde{b}\big(X_v^{\alpha,z}\big) + 2\alpha_v\Big)\,\mathrm{d}v + \sqrt{2}\,\mathrm{d}B_v^z\,, \qquad \text{for } z = x, y\,, \\ u(t,x) &= \mathbb{E}\bigg[\int_0^t \Big(|\alpha_v|^2 + \varphi_{t-v}\big(X_v^{\alpha,x}\big)\Big)\,\mathrm{d}v + u\big(0, X_t^{\alpha,x}\big)\bigg]\,, \\ u(t,y) &\leqslant \mathbb{E}\bigg[\int_0^t \Big(|\alpha_v|^2 + \varphi_{t-v}\big(X_v^{\alpha,y}\big)\Big)\,\mathrm{d}v + u\big(0, X_t^{\alpha,y}\big)\bigg]\,, \end{aligned}$$

where B^x, B^y are Brownian motions coupled by reflection until $X^{\alpha,x}$, $X^{\alpha,y}$ collide:

$$\mathrm{d}B_v^{\alpha,y} = \left(1 - \frac{2\big(X_v^{\alpha,y} - X_v^{\alpha,x}\big)\big(X_v^{\alpha,y} - X_v^{\alpha,x}\big)^{\mathsf{T}}}{\big|X_v^{\alpha,y} - X_v^{\alpha,x}\big|^2}\right)\mathrm{d}B_v^{\alpha,x}\,,$$

for $v \leqslant \tau \coloneqq \inf\big\{w \geqslant 0 : X_w^{\alpha,x} = X_w^{\alpha,y}\big\}$, and $\mathrm{d}B_v^{\alpha,x} = \mathrm{d}B_v^{\alpha,y}$ for $v > \tau$. Then, by subtracting the dynamics of $X^{\alpha,x}$ and $X^{\alpha,y}$, we find that their difference process $\big|X^{\alpha,x}_\cdot - X^{\alpha,y}_\cdot\big|$ is stochastically dominated by a one-dimensional Itō process $(r_t)_{t\geqslant 0}$ solving

$$\mathrm{d}r_v = -r_v \kappa_{\tilde{b}}(r_v)\,\mathrm{d}v + 2\sqrt{2}\,\mathrm{d}W_v$$

with an absorbing boundary at 0 with initial value $r_0 = |x-y|$. It is shown in [31] that

$$\mathbb{P}[r_v > 0] \leqslant \frac{C r_0}{\sqrt{v \wedge 1}}$$

for some C depending only on $\kappa_{\tilde{b}}$. Then, by combining the result above with the long-time Wasserstein contraction studied in [13], we get, for all $v \in [0,t]$,

$$\mathbb{P}[r_v > 0] \leqslant \frac{C' e^{-c'v} r_0}{\sqrt{v \wedge 1}}$$

for some C', $c' > 0$ depending only on $\kappa_{\tilde{b}}$. Therefore, by subtracting the stochastic representation for $u(t,x)$, $u(t,y)$ and applying the bound above on r_v, we get the first claim.

For the second-order estimate, we take spatial derivatives in the HJB and find that ∇u_t solves the $\mathbb{R}^d$-valued equation

$$\partial_t \nabla u_t = \Delta \nabla u_t + (\tilde{b}_t - 2\nabla u_t)\cdot\nabla^2 u_t + \nabla \tilde{b}_t \cdot \nabla u_t + \nabla \varphi_t\,.$$

Thus, ∇u_t solves a second-order equation with the weakly semi-monotone drift term $\tilde{b}_t - 2\nabla u_t$ (as $\tilde{b}_t$ is weakly semi-monotone and ∇u_t is bounded by the first claim), and a bounded source term $\nabla \tilde{b}_t \cdot \nabla u_t + \nabla \varphi_t$. Writing the Feynman–Kac formula for ∇u_t and using the coupling by reflection as above, we get the second claim. We refer readers to [10] for a rigorous justification of the Hessian bound. □

Theorem 4.6. *Let Assumption 4.1 hold. Let $m_0 \in \mathcal{P}(\mathbb{R}^d)$ be such that*

$$u_0 \coloneqq -\ln \frac{\mathrm{d}m_0}{\mathrm{d}\mu_0} = -\ln m_0 - U - \ln Z(\mu_0)$$

is Lipschitz continuous and let $(m_t)_{t\geqslant 0}$ be the solution to (4.1)*. Denote $u_t \coloneqq -\ln \mathrm{d}m_t/\mathrm{d}\mu_0$. Then we have, for all $t > 0$,*

$$\sup_{x\in\mathbb{R}^d} \big|K \star m_t(x)(1+|x|)\big| \leqslant C\,, \quad \|\nabla u_t\|_{L^\infty} \leqslant C\,, \quad \|\nabla^2 u_t\|_{L^\infty} \leqslant \frac{C}{\sqrt{t \wedge 1}}$$

for some C depending only on d, s, U, $|M|$, $\|m_0\|_{L^\infty}$ and $\|\nabla u_0\|_{L^\infty}$. Moreover, when $|M|$ increases and all other dependencies are kept constant, C increases. Consequently, the flow $(m_t)_{t\geqslant 0}$ satisfies a uniform LSI whose constant has the same dependency as above and is increasing in $|M|$.

The proof of Theorem 4.6 is postponed to Section 4.5.1.

Remark 4.7 (Modulated free energy and LSI, and kinetic case)**.** We remark that since we have obtained the L^∞ bound of $\nabla^2 u_t$ in the theorem above (and also in Theorem 4.11 below), we can control the Lipschitz norm of the time-dependent vector field

$$x \mapsto \sigma^2 \nabla \ln \frac{m_t(x)}{e^{-U(x)}} - K \star m_t(x). \tag{4.9}$$

The control of this quantity, as remarked in [9, Section 1.2], is crucial for the modulated free energy method since it appears in the "commutator estimates". See e.g. [36, Proposition 1.1] or [9, Proposition 2.13]. We note that unfortunately our method to obtain this control exploits the long-time contractivity of Brownian motions coupled by reflection, and relies fundamentally on the diffusivity of the dynamics, so it is not useful for deterministic dynamics (i.e. $\sigma = 0$) considered originally in [36]. Nevertheless, since similar results for the kinetic case in the time-homogeneous setting have been established by two of the authors in [29, Theorems 2 and 13], our method provides bounds on $\nabla^2 \ln m_t$, which is of interest in the perspective of applying the arguments [20, Theorem 2] in such hypoelliptic cases.

Besides, together with the control of the Lipschitz norm of (4.9), a key ingredient to get uniform-in-time estimates when using modulated free energy instead of relative entropy is the modulated log-Sobolev inequalities discussed in [35]. These modulated LSI are in fact classical LSI satisfied uniformly over a specific family of measures (called the modulated Gibbs measures, and distinct from the law m_t that we consider in Theorem 4.6; but a similar time-uniformity is required). The arguments of the time-uniform LSI of Theorem 4.6 may thus be useful to establish time-uniform modulated LSI (although additional difficulties appear in the latter case, in particular a uniformity in the number of particles is required). On the topic of modulated free energy and modulated LSI, we mention that an upcoming work [19] is announced in [9].

Remark 4.8 (Non-conservative flow and more singularity)**.** Two natural extensions to the setting considered in Assumption 4.1 are to consider a not necessarily anti-symmetric M (notably $M = -I_{d\times d}$ which corresponds to the gradient flow) and a more singular interaction with $s \in [d-1, d)$. We note that in the first case, a not anti-symmetric M poses challenges in mathematical analysis since the divergence term

$$\nabla \cdot (K \star m_t) = M : \nabla^2 g \star m_t$$

appears and is more singular than the flow m_t itself when $s \geqslant d-2$. By re-examining the proof, we find that the method of Theorem 4.6 will continue to work if $(m_t)_{t\geqslant 0}$ satisfies a uniform θ-Hölder bound for some $\theta > s - d + 3$ without the anti-symmetry of M, or for some $\theta > s - d + 2$ with an anti-symmetric M. The authors are unfortunately unaware of such results for Riesz flows with confinement in the whole space, which are possibly worthy of independent studies in the future.

4.3 Global PoC for log interaction with general confinement potential

As a consequence of Theorem 4.6, we get the strong uniform-in-time propagation of chaos result.

Theorem 4.9. *Let Assumption 4.1 hold and suppose additionally that $s = 0$ and M is anti-symmetric. Let $(m_t)_{t\geqslant 0}$ be a solution to* (4.1) *whose initial value m_0 satisfies the*

conditions of Theorem 4.6 and let $(m_t^N)_{t\geqslant 0}$ *be a solution to* (4.3)*. Then, there exist* C, C', $\rho > 0$, *depending only on* d, U, $|M|$ *and* m_0, *such that*

$$\mathcal{H}\big(m_t^N \big| m_t^{\otimes N}\big) \leqslant C\exp\big(-(\rho - C')t\big)\mathcal{H}\big(m_0^N \big| m_0^{\otimes N}\big) + C\Big(1 + \exp\big(-(\rho - C')t\big)\Big) \tag{4.10}$$

for all $t \geqslant 0$, *once* $\mathcal{H}\big(m_0^N \big| m_0^{\otimes N}\big) < \infty$. *Moreover, when* $|M|$ *decreases and all other dependencies are kept constant,* C *decreases,* ρ *increases, and* $\lim_{|M|\to 0} C' = 0$.

By the dependency of C, C' and ρ in $|M|$, we find $C' < \rho$ when $|M|$ is sufficiently small, and in this case the bound (4.10) becomes uniform in time. Even when $|M|$ is not small, we get a global PoC estimate for the dissipative log-interaction on $\mathbb{R}^d$ with a confinement potential, which is new to our knowledge (the case $U = 0$ is addressed in [14]).

The proof of Theorem 4.9 is postponed to Section 4.5.2.

4.4 Uniform PoC for log interaction with quadratic confinement potential

In this subsection, we impose the additional assumption.

Assumption 4.10. *The confinement potential reads* $U(x) = \kappa_U|x|^2/2$ *for some* $\kappa_U > 0$.

Under Assumptions 4.1 and 4.10, we easily verify that the Gaussian measure m_* with density

$$m_*(x) = \exp\big(-U(x)\big) = \exp\bigg(-\frac{\kappa_U|x|^2}{2}\bigg)$$

is invariant to the mean field flow (4.1). The first result that we obtain is the exponential convergence of the mean field flow towards m_*.

Theorem 4.11. *Let Assumptions 4.1 and 4.10 hold and suppose additionally that* M *is anti-symmetric. Let* $m_0 \in \mathcal{P}(\mathbb{R}^d)$ *satisfy the conditions of Theorem 4.6 and let* $(m_t)_{t\geqslant 0}$ *be the solution to* (4.1)*. Then, we have, for all* $t \geqslant 0$,

$$\mathcal{H}(m_t|m_*) \leqslant \exp(-2\kappa_U t)\mathcal{H}(m_0|m_*)\,.$$

Moreover, setting $u_t = -\ln \mathrm{d}m_t/\mathrm{d}m_*$, *we have, for all* $t > 0$,

$$\sup_{x\in\mathbb{R}^d} |x\cdot K \star (m_t - m_*)| \leqslant Ce^{-ct}\,, \quad \|\nabla u\|_{L^\infty} \leqslant Ce^{-ct}\,, \quad \|\nabla^2 u\|_{L^\infty} \leqslant \frac{Ce^{-ct}}{\sqrt{t\wedge 1}}$$

for some C, $c > 0$ *that depend only on* d, s, κ_U, M *and* $\|\nabla u_0\|_{L^\infty}$.

The proof of Theorem 4.11 is postponed to Section 4.5.1.

Building upon the exponential convergence above, we obtain the uniform-in-time propagation of chaos without restriction on the strength of the interaction.

Theorem 4.12. *Let Assumptions 4.1 and 4.10 hold and suppose additionally that* $s = 0$ *and* M *is anti-symmetric. Let* $(m_t)_{t\geqslant 0}$ *be a solution to* (4.1) *whose initial value* m_0 *satisfies the conditions of Theorem 4.6 and let* $(m_t^N)_{t\geqslant 0}$ *be a solution to* (4.3)*. Then, there exists* $C > 0$, *depending only on* d, κ_U, M *and* m_0, *such that*

$$\mathcal{H}\big(m_t^N \big| m_t^{\otimes N}\big) \leqslant C\exp(-2\kappa_U t)\Big(\mathcal{H}\big(m_0^N \big| m_0^{\otimes N}\big) + 1\Big) \tag{4.11}$$

for all $t \geqslant 0$, *once* $\mathcal{H}\big(m_0^N \big| m_0^{\otimes N}\big) < \infty$.

The proof of Theorem 4.12 is postponed to Section 4.5.2. Notice that, as discussed in e.g. [35], this result describes a *generation* of chaos property (not only *propagation*) since it implies that $\mathcal{H}\big(m_t^N \big| m_t^{\otimes N}\big)$ is of order 1 (in terms of N) for large times even if it is not the case at time $t = 0$. Here, moreover, and more surprisingly, the right hand

side of (4.11) vanishes at $t \to \infty$, which is due to the fact that in the specific case of an isotropic Gaussian confining potential, the invariant measure of the system of interacting particles is a tensorized Gaussian distribution, which is thus also the long-time limit of the product of solutions of the non-linear equation. Finally, in contrast with the results stated in Section 3, here (as in Theorem 4.9) we only state a result on the relative entropy of the full system, and thus by sub-additivity of the relative entropy this yields PoC estimates on the k-particles marginals which are not sharp in the sense of [22, 23].

4.5 Proofs

4.5.1 Proofs of uniform bounds and LSI

Proof of Theorem 4.6. Set $\mu_0 = Z^{-1}\exp(-U)$ with $Z = \int \exp(-U)$.

Step 1: Construction of a regular approximation. Recall that the initial condition m_0 is such that

$$u_0 = -\ln \frac{\mathrm{d}m_0}{\mathrm{d}\mu_0} = -\ln m_0 - \ln Z - U$$

is Lipschitz continuous. We construct, for $\varepsilon > 0$, the approximative initial value

$$m_0^\varepsilon = \frac{\exp(-u_0 \star \eta^\varepsilon)\mu_0}{\int \exp(-u_0 \star \eta^\varepsilon)\mu_0^\varepsilon}\,,$$

where $\mu_0^\varepsilon \propto \exp(-U^\varepsilon)$. Construct as well the approximative dynamics (4.5) with the mollified kernel $K^\varepsilon = K \star \eta^\varepsilon$ and mollified confinement $U^\varepsilon = U \star \eta^\varepsilon$. By construction, the initial value m_0^ε converges to m_0 in L^1 and is uniformly bounded in L^∞, thus the last claim of Proposition 4.3 indicates that $m_t^\varepsilon \to m_t$ in L^1 for all $t \geqslant 0$. Using the uniqueness of the solution of the Fokker–Planck equation satisfied by the relative density $\mathrm{d}m_t^\varepsilon/\mathrm{d}\mu_0^\varepsilon$ and a Feynman–Kac argument similar to that of Proposition B.1, we obtain that $u_t^\varepsilon \coloneqq -\ln \mathrm{d}m_t^\varepsilon/\mathrm{d}\mu_0^\varepsilon$ is $\mathcal{C}^2$ in space and its derivatives ∇u_t^ε, $\nabla^2 u_t^\varepsilon$ grow at most linearly in space (locally in time). As a consequence, u_t^ε is a classical solution to the HJB equation

$$\partial_t u_t^\varepsilon = \Delta u_t^\varepsilon - |\nabla u_t^\varepsilon|^2 + \tilde{b}_t^\varepsilon \cdot \nabla u_t^\varepsilon + \varphi_t^\varepsilon\,,$$

where $\tilde{b}_t^\varepsilon$, φ_t^ε are given by

$$\begin{aligned}\tilde{b}_t^\varepsilon &= -\nabla U^\varepsilon - K^\varepsilon \star m_t^\varepsilon\,,\\ \varphi_t^\varepsilon &= -\nabla \cdot (K^\varepsilon \star m_t^\varepsilon) - (K^\varepsilon \star m_t^\varepsilon)\cdot \nabla U^\varepsilon\,.\end{aligned}$$

Step 2: Uniform bound on $K \star m_t$ and ∇u_t, and uniform LSI. We verify that the drift $\tilde{b}_t^\varepsilon$ satisfies the semi-monotonicity condition of Theorem 4.5, as the contribution from the interaction $K^\varepsilon \star m_t^\varepsilon$ is controlled by Proposition 4.2:

$$\|K^\varepsilon \star m_t^\varepsilon\|_{L^\infty} \leqslant \|K \star m_t^\varepsilon\|_{L^\infty} \lesssim \|m_t^\varepsilon\|_{L^1}^{1-(s+1)/d}\|m_t^\varepsilon\|_{L^\infty}^{(s+1)/d}\,,$$

and U (along with its approximation U^ε) is already weakly convex. Now we focus on proving the uniform L^∞ bound on φ_t^ε. For the first term in φ_t^ε, we find that in the Coulombic and the super-Coulombic cases, due to the anti-symmetry of M, we have

$$\nabla \cdot (K^\varepsilon \star m_t^\varepsilon) = \nabla \cdot (M\nabla g \star m_t^\varepsilon \star \eta^\varepsilon) = M : g \star \nabla^2(m_t^\varepsilon \star \eta^\varepsilon) = 0\,;$$

for the sub-Coulombic case where $s < d-2$, applying Proposition 4.2 with $p = \infty$, we get

$$\|\nabla \cdot (K^\varepsilon \star m_t^\varepsilon)\|_{L^\infty} \lesssim \|m_t^\varepsilon\|_{L^1}^{1-(s+2)/d}\|m_t^\varepsilon\|_{L^\infty}^{(s+2)/d}\,,$$

so the first term is uniformly bounded in L^∞ in both cases. To treat the second term, we note that

$$|K^\varepsilon \star m^\varepsilon(x)| \leqslant \sup_{x' \in B(x,\varepsilon)} |K \star m^\varepsilon(x')|\,,$$

so it suffices to prove the bound uniformly:

$$|K \star m^\varepsilon(x')| \lesssim (1+|x|)^{-1}\,.$$

Decompose the kernel in the following way:

$$K(x) = K(x)\mathbb{1}_{|x|<R} + K(x)\mathbb{1}_{|x|\geqslant R} =: K_1(x) + K_2(x)\,.$$

For the exploding part K_1, we have

$$\begin{aligned}
|K_1 \star m_t^\varepsilon(x)| &= \left| \int_{B(x,R)} K(x-y) m_t^\varepsilon(y)\,\mathrm{d}y \right| \\
&\lesssim \int_{B(x,R)} |x-y|^{-s-1} m_t^\varepsilon(y)\,\mathrm{d}y \\
&\leqslant \left(\int_{B(x,R)} |x-y|^{-p(s+1)}\,\mathrm{d}y \right)^{1/p} \|m_t^\varepsilon \mathbb{1}_{B(x,R)}\|_{L^q} \\
&\lesssim R^{d/p-s-1} \|m_t^\varepsilon \mathbb{1}_{B(x,R)}\|_{L^q}\,,
\end{aligned}$$

where $p \in \left(1, \frac{d}{s+1}\right)$ and $p^{-1}+q^{-1}=1$. For $|x|>R$, we observe

$$\int_{B(x,R)} (m_t^\varepsilon)^q \leqslant \|m_t^\varepsilon\|_{L^\infty}^{q-1} \int_{B(x,R)} m_t^\varepsilon \leqslant \|m_t^\varepsilon\|_{L^\infty}^{q-1} (|x|-R)^{-q} \int_{\mathbb{R}^d} |x|^q m_t^\varepsilon(x)\,\mathrm{d}x\,.$$

For the non-exploding part K_2, we have

$$\begin{aligned}
|K_2 \star m_t^\varepsilon(x)| &= \left| \int_{\mathbb{R}^d \setminus B(x,R)} K(x-y) m_t^\varepsilon(y)\,\mathrm{d}y \right| \\
&\lesssim \int_{\mathbb{R}^d \setminus B(x,R)} |x-y|^{-s-1} m_t^\varepsilon(y)\,\mathrm{d}y \\
&= |x|^{-s-1} \int_{\mathbb{R}^d \setminus B(x,R)} \frac{|x|^{s+1}}{|x-y|^{s+1}} m_t^\varepsilon(y)\,\mathrm{d}y \\
&\lesssim |x|^{-s-1} \int_{\mathbb{R}^d \setminus B(x,R)} \frac{|x-y|^{s+1}+|y|^{s+1}}{|x-y|^{s+1}} m_t^\varepsilon(y)\,\mathrm{d}y \\
&\leqslant |x|^{-s-1} \int_{\mathbb{R}^d} (1+R^{-s-1}|y|^{s+1}) m_t^\varepsilon(y)\,\mathrm{d}y\,.
\end{aligned}$$

Thanks to Proposition 4.3, the mean field flow $(m_t^\varepsilon)_{t\geqslant 0}$ enjoys uniform bounds on its L^∞ norm and its moments, as all moments of its initial value m_0^ε are finite. Thus, we have, uniformly in t,

$$\sup_{t\geqslant 0} |K \star m_t^\varepsilon(x)| \leqslant \sup_{t\geqslant 0} |K_1 \star m_t^\varepsilon(x) + K_2 \star m_t^\varepsilon(x)| \lesssim (1+|x|)^{-1}\,.$$

So we have obtained $\sup_{t\geqslant 0} \|\varphi_t^\varepsilon\|_{L^\infty} < \infty$, and the first claim of Theorem 4.5 implies that $\|\nabla u_t^\varepsilon\|_{L^\infty}$ is uniformly bounded. Taking the limit $\varepsilon \to 0$, we recover the uniform spatial Lipschitz bound on u_t and thus by the perturbation lemma of Aida–Shigekawa, the flow $(m_t)_{t\geqslant 0}$ satisfies a uniform LSI.

Step 3: Uniform bound on $\nabla^2 u_t$. We want to apply the second claim of Theorem 4.5 to the HJB solution u_t^ε, and it suffices to control uniformly in time the following quantities:

$$\nabla \tilde{b}_t^\varepsilon \cdot \nabla u_t^\varepsilon = (-\nabla^2 U^\varepsilon - K^\varepsilon \star \nabla m_t^\varepsilon) \cdot \nabla u_t^\varepsilon \,,$$
$$\nabla \varphi_t^\varepsilon = -\nabla^2 \cdot (K^\varepsilon \star m_t^\varepsilon) - \nabla (K^\varepsilon \star m_t^\varepsilon) \cdot \nabla U^\varepsilon - (K^\varepsilon \star m_t^\varepsilon) \cdot \nabla^2 U^\varepsilon \,.$$

The first quantity can be bounded by

$$\|\nabla \tilde{b}_t^\varepsilon \cdot \nabla u_t^\varepsilon\|_{L^\infty} \leqslant \|\nabla \tilde{b}_t^\varepsilon\|_{L^\infty} \|\nabla u_t^\varepsilon\|_{L^\infty} \leqslant \big(\|\nabla^2 U\|_{L^\infty} + \|K^\varepsilon \star \nabla m_t^\varepsilon\|_{L^\infty}\big) \|\nabla u_t^\varepsilon\|_{L^\infty} \,,$$

where $\|K^\varepsilon \star \nabla m_t\|_{L^\infty}$ is uniformly bounded as

$$\nabla m_t^\varepsilon = m_t^\varepsilon (-\nabla U^\varepsilon + \nabla u_t^\varepsilon) = \frac{\exp(-U^\varepsilon - u_t^\varepsilon)}{\int \exp(-U^\varepsilon - u_t^\varepsilon)} (-\nabla U^\varepsilon + \nabla u_t^\varepsilon) \in L^1 \cap L^\infty$$

uniformly in time, thanks to the uniform bound on ∇u_t^ε. Now consider the second quantity $\nabla \varphi_t^\varepsilon$. In the case $s \in [d-2, d-1)$, we have $K = M \nabla g$ with an anti-symmetric M, so the first term $\nabla^2 \cdot (K^\varepsilon \star m_t^\varepsilon)$ vanishes. In the case $s < d-2$, we have

$$\|\nabla^2 \cdot (K^\varepsilon \star m_t^\varepsilon)\|_{L^\infty} \leqslant \|\nabla K \star \nabla m_t^\varepsilon\|_{L^\infty} \lesssim \|\nabla m_t^\varepsilon\|_{L^1}^{1-(s+2)/d} \|\nabla m_t^\varepsilon\|_{L^\infty}^{(s+2)/d} \,,$$

and by the uniform L^1 and L^∞ bound on ∇m_t^ε, this term is uniformly bounded. That is to say, in both cases, the first term $\nabla^2 \cdot (K^\varepsilon \star m_t^\varepsilon)$ is uniformly bounded in L^∞. As we have $\|\nabla^2 U\|_{L^\infty} < \infty$, the third term $(K^\varepsilon \star m_t^\varepsilon) \cdot \nabla^2 U$ is equally uniformly bounded. So it remains to obtain a uniform bound on the second term $\nabla (K^\varepsilon \star m_t^\varepsilon) \cdot \nabla U$. Since ∇U^ε is of linear growth, it suffices to prove

$$\nabla (K \star m_t^\varepsilon)(x) = (K \star \nabla m_t^\varepsilon)(x) \lesssim (1 + |x|)^{-1}$$

uniformly in time. For this, we use again the decomposition $K = K_1 + K_2$ in the end of the previous step, and redoing all the computations, we find that it is sufficient to uniformly control

$$\begin{aligned} \int_{\mathbb{R}^d} |x|^q |\nabla m_t^\varepsilon(x)| \,\mathrm{d}x &= \int_{\mathbb{R}^d} |x|^q |-\nabla U(x) + \nabla u_t^\varepsilon(x)| m_t^\varepsilon(x) \,\mathrm{d}x \\ &\lesssim \int_{\mathbb{R}^d} |x|^q (1 + |x|) m_t^\varepsilon(x) \,\mathrm{d}x \end{aligned}$$

for some $q > \big(1 - \frac{s+1}{d}\big)^{-1}$. But from Proposition 4.3 we know that the q and $(q+1)$-th moments of m_t^ε are uniformly bounded. Hence, $\nabla \varphi_t^\varepsilon$ is uniformly bounded in L^∞ and by the second claim of Theorem 4.5, we get that $\|\nabla^2 u_t^\varepsilon\|_{L^\infty}$ is uniformly bounded. Thus $\nabla^2 \ln m_t^\varepsilon = -\nabla^2 U^\varepsilon - \nabla^2 u_t^\varepsilon$ is uniformly bounded as well, and taking the limit $\varepsilon \to 0$, we get the desired result for u_t. □

Proof of Theorem 4.11. The proof is similar to that of Theorem 4.6, except that now the Lipschitz and Hessian bounds converge to zero. Thus, we first show that the mean field flow m_t converges to the invariant measure m_* and then redo the estimates on the log-density.

Step 1: Convergence in entropy. For the initial value m_0 such that

$$\nabla u_0 = -\nabla \ln m_0 - \nabla U \in L^\infty,$$

we find, as in the beginning of the proof of Theorem 4.6, an approximation m_0^ε defined by the following:

$$m_0^\varepsilon = \frac{\exp(-u_0 \star \eta^\varepsilon - U)}{\int \exp(-u_0 \star \eta^\varepsilon - U)} \,.$$

Set $u_t^\varepsilon := -\ln \mathrm{d}m_0^\varepsilon/\mathrm{d}m_*$. We also consider the approximative flow $(m_t^\varepsilon)_{t\geqslant 0}$ solving the mean field Fokker–Planck equation (4.5). Notice that, in the case of quadratic potential, the mollified potential $U^\varepsilon = U \star \eta^\varepsilon$ is nothing but U translated by a constant, due to the symmetry of η^ε. By Feynman–Kac arguments, we get that m_t^ε is a classical solution to the Fokker–Planck and $\nabla^i u_t^\varepsilon$ grows at most linearly for $i = 0, 1, 2$. Thus, we can derive $t \mapsto \mathcal{H}(m_t^\varepsilon|m_*)$ and get

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{H}(m_t^\varepsilon|m_*)}{\mathrm{d}t} &= -\mathcal{I}(m_t^\varepsilon|m_*) + \int_{\mathbb{R}^d} \nabla \ln \frac{m_t^\varepsilon(x)}{m_*(x)} \cdot K^\varepsilon \star (m_t^\varepsilon - m_*)(x) m_t^\varepsilon(\mathrm{d}x) \\
&= -\mathcal{I}(m_t^\varepsilon|m_*) - \int_{\mathbb{R}^d} \nabla \ln m_*(x) \cdot K^\varepsilon \star (m_t^\varepsilon - m_*)(x) m_t^\varepsilon(\mathrm{d}x) \\
&= -\mathcal{I}(m_t^\varepsilon|m_*) - \int_{\mathbb{R}^d} \nabla \ln m_*(x) \cdot K^\varepsilon \star m_t^\varepsilon(x) m_t^\varepsilon(\mathrm{d}x) \\
&= -\mathcal{I}(m_t^\varepsilon|m_*) + \kappa_U \int x \cdot K^\varepsilon \star m_t^\varepsilon(x) m_t(\mathrm{d}x) \\
&= -\mathcal{I}(m_t^\varepsilon|m_*) + \kappa_U \iint_{\mathbb{R}^d\times\mathbb{R}^d} x \cdot K^\varepsilon(x-y) m_t^\varepsilon(\mathrm{d}x) m_t^\varepsilon(\mathrm{d}y) \\
&= -\mathcal{I}(m_t^\varepsilon|m_*) + \frac{\kappa_U}{2} \iint_{\mathbb{R}^d\times\mathbb{R}^d} (x-y) \cdot K^\varepsilon(x-y) m_t^\varepsilon(\mathrm{d}x) m_t^\varepsilon(\mathrm{d}y) \\
&= -\mathcal{I}(m_t^\varepsilon|m_*) \leqslant -2\kappa_U \mathcal{H}(m_t^\varepsilon|m_*)\,.
\end{aligned}$$

Here the second inequality is due to the integration by parts and the fact that $\nabla \cdot K^\varepsilon = 0$; the third to the fact that $\nabla \ln m_*(x)$ is parallel to x and $K^\varepsilon \star m_*(x) = K \star (m_* \star \eta^\varepsilon)(x)$ is always orthogonal to x, as $m_* \star \eta^\varepsilon$ is invariant by rotation; the sixth due to the oddness of K^ε; and the last due to

$$x \cdot K^\varepsilon(x) = x^\top M \nabla g^\varepsilon(x)$$

and $\nabla g^\varepsilon(x)$ is parallel to x. Then applying Grönwall's lemma and the log-Sobolev inequality for m_*, we get

$$\mathcal{H}(m_t^\varepsilon|m_*) \leqslant \mathcal{H}(m_0^\varepsilon|m_*) \exp(-2\kappa_U t)\,,$$

and taking the limit $\varepsilon \to 0$ and using the lower semi-continuity of relative entropy, we recover the first claim.

Step 2: Decaying bound on $x\cdot K\star m_t(x)$ and ∇u_t. In the following, C, c will denote positive reals that has the same dependency as stated in the theorem and may change from line to line. Working again with the approximation m_t^ε, we get by Pinsker's inequality,

$$\begin{aligned}
\|m_t^\varepsilon - m_*\|_{L^1} &\leqslant \exp(-\kappa_U t)\sqrt{2\mathcal{H}(m_0^\varepsilon|m_*)} \\
&\leqslant \exp(-\kappa_U t)\sqrt{2\kappa_U^{-1}\mathcal{I}(m_0^\varepsilon|m_*)} \\
&\leqslant \exp(-\kappa_U t)\sqrt{2\kappa_U^{-1}\|\nabla u_0^\varepsilon\|_{L^\infty}^2} = C\exp(-\kappa_U t)\,.
\end{aligned}$$

Then, applying Proposition 4.2, we get

$$\|K^\varepsilon \star (m_t^\varepsilon - m_*)\|_{L^\infty} \leqslant C\|m_t^\varepsilon - m_*\|_{L^1}^{1-(s+1)/d}\|m_t^\varepsilon - m_*\|_{L^\infty}^{(s+1)/d} \leqslant Ce^{-ct}\,.$$

We know that $u_t^\varepsilon = -\ln \mathrm{d}m_t^\varepsilon/\mathrm{d}m_*$ solves the HJB equation

$$\partial_t u_t^\varepsilon = \Delta u_t^\varepsilon - |\nabla u_t^\varepsilon|^2 + \tilde{b}_t^\varepsilon \cdot \nabla u_t + \varphi_t^\varepsilon$$

for $\tilde{b}_t^\varepsilon(x) = -\kappa_U x - K^\varepsilon \star m_t^\varepsilon(x)$ and $\varphi_t^\varepsilon(x) = -\kappa_U x \cdot K^\varepsilon \star m_t^\varepsilon(x)$. Note that φ_t satisfies

$$\varphi_t^\varepsilon(x) = -\kappa_U x \cdot K^\varepsilon \star (m_t^\varepsilon - m_*)(x)\,,$$

since $x \cdot K^\varepsilon \star m_*(x) = 0$ according to the argument in Step 1. Thus, we have

$$\begin{aligned}\varphi_t^\varepsilon(x) &= -\kappa_U \int_{\mathbb{R}^d} x^\top M \nabla g^\varepsilon(x-y)(m_t^\varepsilon - m_*)(\mathrm{d}y) \\ &= -\kappa_U \int_{\mathbb{R}^d} y^\top M \nabla g^\varepsilon(x-y)(m_t^\varepsilon - m_*)(\mathrm{d}y)\,,\end{aligned}$$

as $x^\top M \nabla g^\varepsilon(x) = 0$ for all $x \in \mathbb{R}^d$. So φ_t^ε satisfies the bound

$$\begin{aligned}\|\varphi_t^\varepsilon\|_{L^\infty} &= \left| \int_{\mathbb{R}^d} y^\top K^\varepsilon(x-y)(m_t^\varepsilon - m_*)(\mathrm{d}y) \right| \\ &\lesssim \int_{B(0,1)} \frac{|y|}{|x-y|^{s+1}} |m_t^\varepsilon - m_*|(\mathrm{d}y) + \sup_{y:|y-x|\geqslant 1} |y^\top K^\varepsilon(x-y)|\, \|m_t^\varepsilon - m_*\|_{L^1} \\ &\lesssim \|(m_t^\varepsilon - m_*)\mathbb{1}_{B(x,1)}\|_{L^q} + \|m_t^\varepsilon - m_*\|_{L^1}\end{aligned}$$

for $q > \left(1 - \frac{s+1}{d}\right)^{-1}$, according to the argument in the proof of Theorem 4.6. For the L^q norm we have, by interpolation,

$$\|(m_t^\varepsilon - m_*)\mathbb{1}_{B(x,1)}\|_{L^q} \leqslant \|m_t^\varepsilon - m_*\|_{L^q} \leqslant \|m_t^\varepsilon - m_*\|_{L^1}^{1/q} \|m_t^\varepsilon - m_*\|_{L^\infty}^{1/p} \lesssim \|m_t^\varepsilon - m_*\|_{L^1}^{1/q}$$

for $p^{-1} + q^{-1} = 1$. Thus, $\|\varphi_t^\varepsilon\|_{L^\infty} \leqslant Ce^{-ct}$, and applying the first claim of Theorem 4.5, we get $\|\nabla u_t^\varepsilon\|_{L^\infty} \leqslant Ce^{-ct}$. The first claim is then proved by taking the limit $\varepsilon \to 0$.

Step 3: Decaying bound on $\nabla^2 u_t$. First, we have

$$\|\nabla \tilde{b}_t^\varepsilon \cdot \nabla u_t^\varepsilon\|_{L^\infty} \leqslant \|\nabla \tilde{b}_t^\varepsilon\|_{L^\infty} \cdot Ce^{-ct} \leqslant \left(\|\nabla^2 U\|_{L^\infty} + \|K^\varepsilon \star \nabla m_t^\varepsilon\|_{L^\infty}\right) \cdot Ce^{-ct}\,,$$

where

$$\|K^\varepsilon \star \nabla m_t^\varepsilon\|_{L^\infty} \lesssim \|\nabla m_t^\varepsilon\|_{L^1}^{1-(s+1)/d} \|\nabla m_t^\varepsilon\|_{L^\infty}^{(s+1)/d}\,.$$

As we have

$$\nabla m_t^\varepsilon = -(\nabla U + \nabla u_t^\varepsilon) m_t^\varepsilon = -\frac{\nabla U \exp(-U - u_t^\varepsilon)}{\int \exp(-U - u_t^\varepsilon)} - \nabla u_t^\varepsilon\, m_t^\varepsilon$$

with ∇U of linear growth and ∇u_t^ε being uniformly bounded, we find that $\nabla m_t^\varepsilon \in L^1 \cap L^\infty$ uniformly. Thus,

$$\|\nabla \tilde{b}_t^\varepsilon \cdot \nabla u_t^\varepsilon\|_{L^\infty} \leqslant Ce^{-ct}\,.$$

The gradient of φ_t^ε reads

$$\begin{aligned}\nabla \varphi_t^\varepsilon(x) &= -\nabla\big(\kappa_U x \cdot K^\varepsilon \star (m_t^\varepsilon - m_*)(x)\big) \\ &= -\kappa_U K^\varepsilon \star (m_t^\varepsilon - m_*)(x) - \kappa_U x \cdot K^\varepsilon \star \nabla(m_t^\varepsilon - m_*)(x)\,.\end{aligned}$$

The first term on the right hand side is already controlled:

$$|K^\varepsilon \star (m_t - m_*)(x)| \lesssim \|m_t - m_*\|_{L^1}^{1-(s+1)/d} \|m_t - m_*\|_{L^\infty}^{(s+1)/d} \leqslant Ce^{-ct}\,,$$

and in the following we show that the same is true for the second term. Again, using the fact that $x \cdot K^\varepsilon(x) = 0$, we get

$$\begin{aligned}x \cdot K^\varepsilon \star \nabla(m_t^\varepsilon - m_*)(x) &= \int_{\mathbb{R}^d} x^\top M \nabla g^\varepsilon(x-y) \nabla(m_t^\varepsilon - m_*)(\mathrm{d}y) \\ &= \int_{\mathbb{R}^d} y^\top M \nabla g^\varepsilon(x-y) \nabla(m_t^\varepsilon - m_*)(\mathrm{d}y)\,.\end{aligned}$$

Following the argument in Step 2, we separate the two cases $|y-x| < 1$ and $\geqslant 1$, and get

$$\sup_{x \in \mathbb{R}^d} |x \cdot K^\varepsilon \star \nabla(m_t^\varepsilon - m_*)(x)| \lesssim \|\nabla(m_t^\varepsilon - m_*)\|_{L^q} + \|\nabla(m_t^\varepsilon - m_*)\|_{L^1}$$

for $q > \big(1 - \frac{s+1}{d}\big)^{-1}$. Using the explicit density of m_t^ε, we get

$$\nabla m_t^\varepsilon - \nabla m_* = -\nabla u_t^\varepsilon \, m_t^\varepsilon - \nabla U(m_t^\varepsilon - m_*)\,.$$

The first term satisfies

$$\|\nabla u_t^\varepsilon \, m_t^\varepsilon\|_{L^1} \leqslant \|\nabla u_t^\varepsilon\|_{L^\infty} \|m_t^\varepsilon\|_{L^1} \leqslant Ce^{-ct}\,,$$

and the second satisfies

$$\|\nabla U(m_t^\varepsilon - m_*)\|_{L^1} \leqslant \|\nabla^2 U\|_{L^\infty} W_1(m_t^\varepsilon, m_*) \lesssim \sqrt{\mathcal{H}(m_t^\varepsilon | m_*)} \leqslant Ce^{-ct}\,.$$

Finally, their densities have the L^∞ bounds:

$$\begin{aligned}
\|\nabla u_t^\varepsilon \, m_t^\varepsilon\|_{L^\infty} &\leqslant \|\nabla u_t^\varepsilon\|_{L^\infty} \|m_t^\varepsilon\|_{L^\infty} \leqslant C\,,\\
\|\nabla U(m_t^\varepsilon - m_*)\|_{L^\infty} &\lesssim \sup_{x \in \mathbb{R}^d} (1+|x|) \bigg(\frac{\exp\big(-U(x)\big)}{\int \exp(-U)} + \frac{\exp\big(-u_t^\varepsilon(x) - U(x)\big)}{\int \exp(-u_t^\varepsilon - U)} \bigg) \leqslant C\,.
\end{aligned}$$

Then, by the same interpolation as in Step 2, we get $\|\nabla \varphi_t^\varepsilon\|_{L^\infty} \leqslant Ce^{-ct}$. Applying the second claim of Theorem 4.5, we get

$$\|\nabla^2 u_t^\varepsilon\|_{L^\infty} \leqslant \frac{Ce^{-ct}}{\sqrt{t \wedge 1}} + \int_0^t \frac{Ce^{-cv}}{\sqrt{v \wedge 1}} \cdot Ce^{-c(t-v)} \, \mathrm{d}v \leqslant \frac{Ce^{-ct}}{\sqrt{t \wedge 1}}\,.$$

Taking the limit $\varepsilon \to 0$, we recover the second claim and this concludes the proof. □

4.5.2 Proofs of propagation of chaos

Proof of Theorem 4.9. According to Propositions 4.3 and 4.4, given the initial values m_0, m_0^N, we find respectively approximating sequences m_0^ε, $m_0^{\varepsilon,N}$ such that $\ln m_0^\varepsilon + U^\varepsilon \in \mathcal{C}_\mathrm{b}^\infty$ and $\boldsymbol{x} \mapsto \ln m_0^{\varepsilon,N}(\boldsymbol{x}) + \sum_{i=1}^N U^\varepsilon(x^i) \in \mathcal{C}_\mathrm{b}^\infty$. The solutions of (4.5) and of the forward Kolmogorov equation associated to (4.6) being unique, we can use the Feynman–Kac representation of Proposition B.1 to find that the densities and their classically derivatives

$$\nabla^i \big(\ln m_t^\varepsilon + U^\varepsilon\big)\,, \quad \nabla^i \bigg(\ln m_t^{\varepsilon,N}(\boldsymbol{x}) + \sum_{i=1}^N U^\varepsilon(x^i)\bigg)\,, \qquad i \geqslant 1$$

exist and grow at most linearly in space (locally in time). Then in the following we can justify all the exchanges between limit and integration, and all the integrations by parts. Taking the derivative of the relative entropy $\mathcal{H}_t^\varepsilon = \mathcal{H}\big(m_t^{\varepsilon,N} \big| (m_t^\varepsilon)^{\otimes N}\big)$, and denoting the relative Fisher information by

$$\mathcal{I}_t^\varepsilon = \mathcal{I}\big(m_t^{\varepsilon,N} \big| (m_t^\varepsilon)^{\otimes N}\big) = \int_{\mathbb{R}^{dN}} \bigg| \nabla \ln \frac{m_t^{\varepsilon,N}(\boldsymbol{x})}{(m_t^\varepsilon)^{\otimes N}(\boldsymbol{x})} \bigg|^2 m_t^{\varepsilon,N}(\mathrm{d}\boldsymbol{x})\,,$$

we get

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{H}_t^\varepsilon}{\mathrm{d}t} &= -\mathcal{I}_t^\varepsilon + \frac{1}{N-1} \sum_{i \neq j} \int_{\mathbb{R}^{dN}} \nabla_i \ln \frac{m_t^{\varepsilon,N}(\boldsymbol{x})}{m_t^\varepsilon(x^i)} \\
&\qquad\qquad \cdot \big(K^\varepsilon(x^i - x^j) - K^\varepsilon \star m_t^\varepsilon(x^i)\big) m_t^{\varepsilon,N}(\mathrm{d}\boldsymbol{x}) \\
&= -\mathcal{I}_t^\varepsilon - \frac{1}{N-1} \sum_{i \neq j} \int_{\mathbb{R}^{dN}} \nabla \ln m_t^\varepsilon(x^i) \\
&\qquad\qquad \cdot \big(K^\varepsilon(x^i - x^j) - K^\varepsilon \star m_t^\varepsilon(x^i)\big) m_t^{\varepsilon,N}(\mathrm{d}\boldsymbol{x})\,,
\end{aligned}$$

where i, j are summed over $[\![1, N]\!]$ and the second equality is due to integration by parts and the fact that $\nabla \cdot K^\varepsilon = 0$. Noting that the regularized N-particle measure $m_t^{\varepsilon,N}$ has density and has no mass on the sets $\{\boldsymbol{x} : x^i = x^j\}$ for $i \neq j$, we find that the second term is equal to, by symmetrization,

$$-\frac{1}{N-1}\sum_{i,j=1}^{N}\int_{\mathbb{R}^{dN}}\phi_t(x^i,x^j)m_t^{\varepsilon,N}(\mathrm{d}\boldsymbol{x})\,,$$

where the function $\phi_t(\cdot,\cdot)$ is given by

$$\begin{aligned}\phi_t(x,y) = \frac{1}{2}K^\varepsilon(x-y)\cdot\big(\nabla\ln m_t^\varepsilon(x) - \nabla\ln m_t^\varepsilon(y)\big)\mathbb{1}_{x\neq y}\\ -\frac{1}{2}K^\varepsilon\star m_t^\varepsilon(x)\cdot\nabla\ln m_t^\varepsilon(x) - \frac{1}{2}K^\varepsilon\star m_t^\varepsilon(y)\cdot\nabla\ln m_t^\varepsilon(y)\,. \quad (4.12)\end{aligned}$$

The function ϕ_t satisfies

$$\begin{aligned}\int_{\mathbb{R}^d}\phi_t(x,y)m_t^\varepsilon(\mathrm{d}y) = 0\,,\\ \int_{\mathbb{R}^d}\phi_t(y,x)m_t^\varepsilon(\mathrm{d}y) = 0\end{aligned}$$

for all $x \in \mathbb{R}^d$. From now on, the symbols C_i, $i \in \mathbb{N}$ denote a positive number that has the same dependency as C, ρ have in the statement of the theorem. For the first term in (4.12), we have by Theorem 4.6,

$$\sup_{x,y\,:\,x\neq y}\big|K^\varepsilon(x-y)\cdot\big(\nabla\ln m_t^\varepsilon(x) - \nabla\ln m_t^\varepsilon(y)\big)\big| \leqslant C_1|M|\|\nabla^2\ln m_t\|_{L^\infty} \leqslant \frac{C_2|M|}{\sqrt{t\wedge 1}}\,.$$

For the last two terms in the definition (4.12) of ϕ_t, we have by the same theorem,

$$\begin{aligned}|K\star m_t^\varepsilon(x)| \leqslant C_3|M|(1+|x|)^{-1}\,,\\ |\nabla\ln m_t^\varepsilon(x)| = |\nabla u_t^\varepsilon(x)| + |\nabla U(x)| \leqslant C_4(1+|x|)\,.\end{aligned}$$

Thus,

$$|K\star m_t^\varepsilon(x)\cdot\nabla\ln m_t^\varepsilon(x)| \leqslant C_6|M|\,.$$

So the functions ϕ_t satisfies

$$\|\phi_t\|_{L^\infty} \leqslant \frac{C_7|M|}{\sqrt{t\wedge 1}}\,.$$

Therefore, using the convex duality for relative entropy, we get

$$\frac{\mathrm{d}\mathcal{H}_t^\varepsilon}{\mathrm{d}t} = -\mathcal{I}_t^\varepsilon + \delta_t\mathcal{H}_t^\varepsilon + \delta_t\ln\int_{\mathbb{R}^{dN}}\exp\bigg(\frac{1}{\delta_t(N-1)}\sum_{i,j=1}^{N}\phi_t(x^i,x^j)\bigg)(m_t^\varepsilon)^{\otimes N}(\mathrm{d}\boldsymbol{x})\,,$$

where we set

$$\delta_t = \frac{3(1600^2+36e^4)C_7|M|}{\sqrt{t\wedge 1}}\,.$$

Then, applying the "concentration" estimate [20, Theorem 4] (whose constant is given explicitly in [16, Theorem 5]), we obtain

$$\frac{\mathrm{d}\mathcal{H}_t^\varepsilon}{\mathrm{d}t} \leqslant -\mathcal{I}_t^\varepsilon + \frac{C_8|M|}{\sqrt{t\wedge 1}}\mathcal{H}_t^\varepsilon + \frac{C_8|M|}{\sqrt{t\wedge 1}} \leqslant -C_9\mathcal{H}_t^\varepsilon + \frac{C_8|M|}{\sqrt{t\wedge 1}}\mathcal{H}_t^\varepsilon + \frac{C_8|M|}{\sqrt{t\wedge 1}}\,.$$

We conclude by applying Grönwall's lemma and taking the limit $\varepsilon \to 0$. □

Proof of Theorem 4.12. The argument is largely the same as the proof above, i.e. the proof of Theorem 4.9. So here we only indicate the differences. Defining the same ϕ_t function as in (4.12), we find that in the quadratic case, we have the following bounds by Theorem 4.11:

$$\begin{aligned}\left\|\nabla \ln \frac{m_t^\varepsilon}{m_*}\right\|_{L^\infty} &\leqslant C_1 e^{-ct}\,,\\ \left\|\nabla^2 \ln \frac{m_t^\varepsilon}{m_*}\right\|_{L^\infty} &\leqslant \frac{C_1 e^{-ct}}{\sqrt{t\wedge 1}}\,,\\ \sup_x |x\cdot K^\varepsilon \star (m_t^\varepsilon - m_*)(x)| &\leqslant C_1 e^{-ct}\,.\end{aligned}$$

So for the first term in the definition (4.12) of ϕ, we have, for all $x \neq y$,

$$\begin{aligned}&\left|K^\varepsilon(x-y)\cdot\left(\nabla \ln m_t^\varepsilon(x) - \nabla \ln m_t^\varepsilon(y)\right)\right|\\ &\quad= \left|K^\varepsilon(x-y)\cdot\left(\nabla \ln \frac{m_t^\varepsilon(x)}{m_*(x)} - \nabla \ln \frac{m_t^\varepsilon(y)}{m_*(y)}\right)\right|\\ &\quad\lesssim \left\|\nabla^2 \ln \frac{m_t^\varepsilon}{m_*}\right\|_{L^\infty} \leqslant \frac{C_1 e^{-ct}}{\sqrt{t\wedge 1}}\,.\end{aligned}$$

For the second term, we have

$$K^\varepsilon \star m_t^\varepsilon(x)\cdot\nabla \ln m_t^\varepsilon(x) = K^\varepsilon \star (m^\varepsilon - m_*)(x)\cdot \nabla \ln m_* + K^\varepsilon \star m_t^\varepsilon(x)\cdot \nabla \ln \frac{m_t^\varepsilon(x)}{m_*(x)}\,,$$

therefore,

$$\begin{aligned}\|K^\varepsilon \star m_t^\varepsilon \cdot \nabla \ln m_t^\varepsilon\|_{L^\infty} \leqslant \kappa_U \sup_x |x\cdot K^\varepsilon \star (m_t^\varepsilon - m_*)(x)| + \|K^\varepsilon \star m_t^\varepsilon\|_{L^\infty}\left\|\nabla \ln \frac{m_t^\varepsilon}{m_*}\right\|_{L^\infty}\\ \leqslant C_2 e^{-ct}\,.\end{aligned}$$

Combining the two results above, we derive the decaying L^∞ bound for ϕ_t:

$$\|\phi_t\|_{L^\infty} \leqslant \frac{C_3 e^{-ct}}{\sqrt{t\wedge 1}}\,.$$

Thus, taking the alternative

$$\delta_t = \frac{3(1600^2 + 36e^4)C_3 e^{-ct}}{\sqrt{t\wedge 1}}\,,$$

we get

$$\frac{\mathrm{d}\mathcal{H}_t^\varepsilon}{\mathrm{d}t} \leqslant -\mathcal{I}_t^\varepsilon + \frac{C_4 e^{-ct}}{\sqrt{t\wedge 1}}\mathcal{H}_t^\varepsilon + \frac{C_4 e^{-ct}}{\sqrt{t\wedge 1}}\,.$$

Finally, we note that, as the Lipschitz constant of $\ln m_t^\varepsilon/m_*$ tends to zero exponentially, the perturbed measure m_t^ε satisfies a k_t-LSI with

$$k_t = 2\kappa_U \exp(-C_5 e^{-c't})\,.$$

Thus, for all $t \geqslant 0$, we have

$$\mathcal{I}_t^\varepsilon \geqslant 2\kappa_U \exp(-C_5 e^{-c't})\mathcal{H}_t^\varepsilon\,.$$

We conclude by applying Grönwall's lemma and taking the limit $\varepsilon \to 0$. □

A Well-posedness of singular dynamics

The mean field well-posedness proof will mainly be based on the estimates on the convolution with the kernel K in Proposition 4.2 and the following elementary result.

Proposition A.1 (Growth and stability estimates)**.** *Let $T > 0$ and $\beta : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a vector field that is the sum of a Lipschitz and a bounded part, that is, $\beta = \beta_{\mathrm{Lip}} + \beta_{\mathrm{b}}$ with $\nabla\beta_{\mathrm{Lip}}$, $\beta_{\mathrm{b}} \in L^\infty$. Suppose its divergence is lower bounded: $(\nabla \cdot \beta)_- \in L^\infty$. Let $m : [0,T] \to \mathcal{P}(\mathbb{R}^d)$ be a probability solution to the parabolic equation*

$$\partial_t m_t = \Delta m_t - \nabla \cdot (\beta_t m_t)\,.$$

Then, for all $p \in [2,\infty]$, we have

$$\|m_t\|_{L^p} \leqslant C_p\big(\|m_0\|_{L^p} + 1\big)$$

for some C_p depending only on p, d and $\|(\nabla \cdot \beta)_-\|_{L^\infty}$ (notably independent of t and T).

Moreover, let β' be another vector field satisfying the same conditions as β, and let m' be a probability solution to the equation corresponding to β'. Then, for all $p \in \{1\} \cup [2,\infty)$, we have

$$\|m_t - m'_t\|_{L^p} \leqslant e^{C'_p t}\|m_0 - m'_0\|_{L^p} + C'_p\Big((e^{C'_p t} - 1)\mathbb{1}_{p\geqslant 2} + \sqrt{t}\mathbb{1}_{p=1}\Big) \sup_{v\in[0,t]} \|\beta_v - \beta'_v\|_{L^\infty}$$

for some C'_p depending only on p, d, $\|(\nabla \cdot \beta)_-\|_{L^\infty}$, $\|(\nabla \cdot \beta')_-\|_{L^\infty}$, $\|m_0\|_{L^p}$ and $\|m'_0\|_{L^p}$.

Proof. First, consider the SDE

$$\mathrm{d}X_t = \beta_t(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t\,.$$

Since its drift is the sum of a bounded and a Lipschitz part, we have the existence of the strong solution and we find that if $\mathrm{Law}(X_0) = m_0$, then we have the correspondence $\mathrm{Law}(X_t) = m_t$, by the uniqueness of the PDE. Moreover, it is known (see e.g. [6]) that if we take a mollified sequence approaching towards β, the SDE solution will also tend to the original one, i.e. X, and we have the continuous dependency on the initial value as well. So without loss of generality, we can suppose that $\nabla\beta \in \mathcal{C}^\infty_{\mathrm{b}}$ and m_0 belongs to the Schwartz class. By a Feynman–Kac argument similar to that of Proposition B.1, we know that m_t belongs also to the Schwartz class. Thus, in the following we perform only formal calculations.

Step 1: Growth estimates. Let $p \geqslant 2$. The L^p norm of m_t satisfies

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} m_t^p &= p\int_{\mathbb{R}^d} m_t^{p-1}\partial_t m_t \\
&= p\int_{\mathbb{R}^d} m_t^{p-1}\Big(\Delta m_t - \nabla\cdot(\beta_t m_t)\Big) \\
&= \int_{\mathbb{R}^d}\Big(-p(p-1)m_t^{p-2}|\nabla m_t|^2 - (p-1)(\nabla\cdot\beta_t)m_t\Big) \\
&\leqslant -p(p-1)\int_{\mathbb{R}^d} m_t^{p-2}|\nabla m_t|^2 + (p-1)\|(\nabla\cdot\beta)_-\|_{L^\infty}\int_{\mathbb{R}^d} m_t^p \\
&\leqslant (p-1)\|(\nabla\cdot\beta)_-\|_{L^\infty}\int_{\mathbb{R}^d} m_t^p\,,
\end{aligned}$$

where here and in the following C_p denotes a constant having the same dependencies as in the statement, and may change from line to line. We would also denote by C

a constant that does not depend on p, but having the same other dependencies. By Grönwall's lemma, we get

$$\|m_t\|_{L^p} \leqslant \exp\left(\frac{p-1}{p}(\nabla\cdot\beta)_-\|_{L^\infty} t\right)\|m_0\|_{L^p}\,, \tag{A.1}$$

and taking $p\to\infty$, we get

$$\|m_t\|_{L^\infty} \leqslant \exp\big(\|(\nabla\cdot\beta)_-\|_{L^\infty} t\big)\|m_0\|_{L^\infty}\,. \tag{A.2}$$

Now we show that the two estimates above can be improved into time-uniform ones. To this end, define the operator $\mathcal{L}_t=\Delta+\beta_t\cdot\nabla$ and its dual $\mathcal{L}_t^*=\Delta-\nabla\cdot(\beta_t\cdot)$. Denote by $(P_{u,t})_{0\leqslant u\leqslant t\leqslant T}$ the time-dependent semi-group generated. Specializing to $p=2$ in the L^p computations above, we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} m_t^2 \leqslant -2\int_{\mathbb{R}^d}|\nabla m_t|^2 + \|(\nabla\cdot\beta)_-\|_{L^\infty}\int_{\mathbb{R}^d} m_t^2\,.$$

The Nash inequality indicates

$$\|m_t\|_{L^2}^{1+2/d} \leqslant C_d\|m_t\|_{L^1}^{2/d}\|\nabla m_t\|_{L^2}\,,$$

where C_d depends only on d. So by Grönwall's lemma, we get the uniform-in-time bound over $\|m_t\|_{L^2}$:

$$\|m_t\|_{L^2}^2 \leqslant \left(\frac{C_d^2\|(\nabla\cdot\beta)_-\|_{L^\infty}}{2}\right)^{d/2}(1-e^{-\kappa t})^{-d/2}\|m_t\|_{L^1}^2\,,$$

for $\kappa = 2\|(\nabla\cdot\beta)_-\|_{L^\infty}/d$. Note that this bound is independent of $\|m_0\|_{L^\infty}$. Now we take an arbitrary $h_0:\mathbb{R}^d\to[0,\infty)$ of the Schwartz class and consider the dual evolution $\partial_u h_u = \mathcal{L}_{t-u}h_u$, that is,

$$\partial_u h_u = \Delta h_u + \beta_{t-u}\cdot\nabla h_u\,,$$

for $u\in[0,t]$, where $t\in[0,T]$. Deriving the L^1 norm of h_s and integrating by parts, we get

$$\|h_u\|_{L^1} \leqslant \exp\big(\|(\nabla\cdot\beta)_-\|_{L^\infty} u\big)\|h_0\|_{L^1}\,.$$

Doing the same for the L^2 norm, we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}u}\int_{\mathbb{R}^d} h_u^2 &= 2\int_{\mathbb{R}^d} h_u\mathcal{L}_{T-u}h_u\\ &= 2\int_{\mathbb{R}^d} h_u\big(\Delta h_u+\beta_{t-u}\cdot\nabla h_u\big)\\ &= -2\int_{\mathbb{R}^d}|\nabla h_u|^2 - \int_{\mathbb{R}^d} h_u^2\nabla\cdot\beta_{t-u}\\ &\leqslant -2\int_{\mathbb{R}^d}|\nabla h_u|^2 + \|(\nabla\cdot\beta)_-\|_{L^\infty}\int_{\mathbb{R}^d} h_u^2\,.\end{aligned}$$

Again, using the Nash inequality:

$$\|h_u\|_{L^2}^{1+2/d} \leqslant C_d\|h_u\|_{L^1}^{2/d}\|\nabla h_u\|_{L^2} \leqslant C_d\exp\big(2\|(\nabla\cdot\beta)_-\|_{L^\infty}u/d\big)\|h_0\|_{L^1}^{2/d}\|\nabla h_u\|_{L^2}\,,$$

we derive the bound over $\|h_t\|_{L^2}$:

$$\|P_{t-u,t}h_0\|_{L^2}^2 = \|h_u\|_{L^2}^2 \leqslant \left(\frac{C_d^2\|(\nabla\cdot\beta)_-\|_{L^\infty}}{2}\right)^{d/2}(e^{-\kappa u}-e^{-2\kappa u})^{-d/2}\|h_0\|_{L^1}^2\,,$$

from which follows the bound on $\|P_{t-u,t}\|_{L^1\to L^2}$. So, taking $u = \max(t/2, t-\kappa^{-1})$, we get

$$\|m_t\|_{L^\infty} = \|P^*_{u,t} m_u\|_{L^\infty} \leqslant \|m_u\|_{L^2}\|P^*_{u,t}\|_{L^2\to L^\infty} \\ = \|m_u\|_{L^2}\|P_{u,t}\|_{L^1\to L^2} \leqslant C(t\wedge 1)^{-d/2}\|m_0\|_{L^1}\,. \quad \text{(A.3)}$$

So, combining (A.2) and (A.3), we get a uniform-in-time bound over $\|m_t\|_{L^\infty}$:

$$\sup_{t\in[0,T]} \|m_t\|_{L^\infty} \leqslant C\big(\|m_0\|_{L^1} + \|m_0\|_{L^\infty}\big)\,. \quad \text{(A.4)}$$

Finally, by differentiating $\int m_t$ and integrating by parts, we get

$$\|m_t\|_{L^1} = \|m_0\|_{L^1}\,. \quad \text{(A.5)}$$

Similarly, interpolating between (A.3) and (A.5), we get

$$\|m_t\|_{L^p} \leqslant C^{(p-1)/p}(t\wedge 1)^{-(p-1)d/2p}\|m_0\|_{L^1}\,,$$

and combing with (A.1), we get

$$\sup_{t\in[0,T]} \|m_t\|_{L^p} \leqslant C_p\big(\|m_0\|_{L^1} + \|m_0\|_{L^p}\big)\,. \quad \text{(A.6)}$$

Step 2: Stability estimates. Now let β', m' be the other vector field and the probability solution. Recall that m, m' correspond respectively to the SDE

$$\mathrm{d}X_t = \beta_t(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t\,, \qquad \mathrm{Law}(X_0) = m_0\,,$$
$$\mathrm{d}X'_t = \beta'_t(X'_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t\,, \qquad \mathrm{Law}(X'_0) = m'_0\,.$$

Now introduce the third SDE, whose drift term is identical to the first, but initial condition identical to the second:

$$\mathrm{d}X''_t = \beta_t(X''_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t\,, \qquad \mathrm{Law}(X''_0) = m'_0\,.$$

such that $\mathbb{P}[X_0 \neq X''_0] = \frac{1}{2}\|m_0 - m'_0\|_{L^1}$. Denote $m''_t = \mathrm{Law}(X''_t)$. Thus, conditioning on the initial condition, we get

$$\|m_t - m''_t\|_{L^1} \leqslant 2\,\mathbb{P}[X_t \neq X''_t] \leqslant 2\,\mathbb{P}[X_0 \neq X''_0] = \|m_0 - m'_0\|_{L^1}\,.$$

On the other hand, by Pinsker's inequality and Girsanov's theorem, we have

$$\|m'_t - m''_t\|^2_{L^1} \leqslant 2H(m'_t|m''_t) \leqslant \frac{1}{2}\int_0^t \|\beta_v - \beta'_v\|^2_{L^\infty}\,\mathrm{d}v\,.$$

Combining the two inequalities above yields the L^1-stability estimate. Now, let $p \geqslant 2$ and

let us calculate:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}|m_t-m_t'|^p \\
&\quad= p\int_{\mathbb{R}^d}|m_t-m_t'|^{p-2}(m_t-m_t')\Big(\Delta(m_t-m_t')-\nabla\cdot(m_t\beta_t)+\nabla\cdot(m_t'\beta_t')\Big) \\
&\quad= -p(p-1)\int_{\mathbb{R}^d}|m_t-m_t'|^{p-2}|\nabla(m_t-m_t')|^2 \\
&\qquad+ p\int_{\mathbb{R}^d}|m_t-m_t'|^{p-2}(m_t-m_t')\Big(-\nabla\cdot(m_t\beta_t)+\nabla\cdot(m_t'\beta_t')\Big) \\
&\quad= -p(p-1)\int_{\mathbb{R}^d}|m_t-m_t'|^{p-2}|\nabla(m_t-m_t')|^2 \\
&\qquad+ p\int_{\mathbb{R}^d}|m_t-m_t'|^{p-2}(m_t-m_t')\Big(-\nabla\cdot\big((m_t-m_t')\beta_t'\big)+\nabla\cdot\big(m_t(\beta_t'-\beta_t)\big)\Big) \\
&\quad\leqslant (p-1)\int_{\mathbb{R}^d}|m_t-m_t'|^p\cdot(-\nabla\cdot\beta_t')+\frac{p(p-1)}{4}\int_{\mathbb{R}^d}|m_t-m_t'|^{p-2}m_t^2|\beta_t-\beta_t'|^2 \\
&\quad\leqslant \frac{(p-1)}{2}\|(\nabla\cdot\beta')_-\|_{L^\infty}\int_{\mathbb{R}^d}|m_t-m_t'|^p+\frac{(p-1)(p-2)}{4}\int_{\mathbb{R}^d}|m_t-m_t'|^p \\
&\qquad+\frac{p-1}{2}\|m_t\|_{L^p}^p\|\beta_t-\beta_t'\|_{L^\infty}^p\,.
\end{aligned}$$

Then, using the uniform L^p estimate in the first step, applying Grönwall's lemma and taking the p-th root, we get the desired result. □

Now we are ready to prove the well-posedness of the mean field dynamics.

Proof of Proposition 4.3. Take a $p\in(1,\infty)$ such that $p^{-1}<1-\frac{s+1}{d}$ and let q be its conjugate: $p^{-1}+q^{-1}=1$. We also take a $\theta\in(0,d-s-1)$.

Step 1: Well-posedness. Let $T\in(0,\infty)$. We define the functional space:

$$\mathcal{X}:=\mathcal{C}([0,T];L^1\cap L^p\cap\mathcal{P})\,.$$

The space $\mathcal{X}$ is a complete metric space. Given $m\in X$, we let $\mathcal{T}[m]$ be the uniqueness probability solution to the Cauchy problem

$$\partial_t\mathcal{T}[m]_t=\Delta\mathcal{T}[m]_t-\nabla\cdot\big((K\star m_t-\nabla U)\mathcal{T}[m]_t\big)\,,\quad \mathcal{T}[m]_0=m_0\,.$$

According to Proposition A.1 we know that $\mathcal{T}[m]\in\mathcal{X}$, where the continuity in $L^1\cap L^p$ follows from a density argument. Moreover, by the stability estimate in the proposition, for all m, $m'\in X$, we have

$$\begin{aligned}
\|\mathcal{T}[m]_t-\mathcal{T}[m']_t\|_{L^r}&\leqslant e^{C_r't}\|m_0-m_0'\|_{L^r}\\
&\quad+C_r'\Big((e^{C_r't}-1)\mathbb{1}_{r=p}+\sqrt{t}\mathbb{1}_{r=1}\Big)\sup_{v\in[0,t]}\|K\star(m_v-m_v')\|_{L^\infty}
\end{aligned}$$

for $r=1$, p. But by Proposition 4.2, we have

$$\|K\star(m_v-m_v')\|_{L^\infty}\lesssim\|m_v-m_v'\|_{L^1}^{1-q(s+1)/d}\|m_v-m_v'\|_{L^p}^{q(s+1)/d}\,.$$

Thus, restricting to the subspace of $\mathcal{X}$ of common initial value and letting T be small enough, we get that the mapping $\mathcal{T}$ is a contraction in $\mathcal{X}$. So a time-local solution exists and is unique. Thanks to the uniform growth estimates, this short time interval can be extended infinitely by iteration. So a unique global solution is recovered and it satisfies

the uniform L^∞ bound thanks to Proposition A.1. For the continuous dependency on the initial value, we use the stability estimates on a small time interval without restricting the initial values to be the same and iterate infinitely as well.

Step 2: Control of moments. Given the uniform L^∞ bound obtained above, we have, according to Proposition 4.2,

$$\|K \star m_t\|_{L^\infty} \lesssim \|m_t\|_{L^1}^{1-(s+1)/d} \|m_t\|_{L^\infty}^{(s+1)/d}.$$

So the contribution from the interaction kernel is bounded. Then we construct, for $k > 0$, the Lyapunov function

$$V_k(x) = \sqrt{1 + |x|^{2k}}\,,$$

and we can easily verify

$$\big(\Delta - \nabla U \cdot \nabla + (K \star m_t) \cdot \nabla\big) V_k \leqslant -c_k V_k + C_k\,,$$

for some $c_k > 0$, $C_k \geqslant 0$. This implies the uniform bound on the k-th moment.

Step 3: Approximation. Let $(m_t^\varepsilon)_{t \geqslant 0}$ be the flow corresponding to the mollified kernel K^ε and potential U^ε. Applying the stability estimates in Proposition A.1, we get

$$\begin{aligned}\|m_t - m_t^\varepsilon\|_{L^r} &\leqslant e^{C_r' t} \|m_0 - m_0^\varepsilon\|_{L^r} \\ &\quad + C_r' \Big((e^{C_r' t} - 1) \mathbb{1}_{r=p} + \sqrt{t} \mathbb{1}_{r=1} \Big) \sup_{v \in [0,t]} \Big(\|K \star m_v - K^\varepsilon \star m_v^\varepsilon\|_{L^\infty} + \|\nabla U - \nabla U^\varepsilon\|_{L^\infty} \Big)\,.\end{aligned}$$

Note that the initial L^p error $\|m_t - m_t^\varepsilon\|_{L^p} \to 0$ by interpolation between L^1 and L^∞. For the first term in the supremem, we have

$$\begin{aligned}\|K \star m_v - K^\varepsilon \star m_v^\varepsilon\|_{L^\infty} &\leqslant \|K \star (m_v - m_v^\varepsilon)\|_{L^\infty} + \|K \star m_v^\varepsilon - K \star m_v^\varepsilon \star \eta^\varepsilon\|_{L^\infty} \\ &\leqslant \|K \star (m_v - m_v^\varepsilon)\|_{L^\infty} + \varepsilon^\theta [K \star m_v^\varepsilon]_{\mathcal{C}^\theta}\,.\end{aligned}$$

By the L^∞ and Hölder estimates in Proposition 4.2, we have the following controls:

$$\begin{aligned}\|K \star (m_v - m_v^\varepsilon)\|_{L^\infty} &\lesssim \|m_v - m_v^\varepsilon\|_{L^1}^{1-q(s+1)/d} \|m_v - m_v^\varepsilon\|_{L^p}^{q(s+1)/d}\,, \\ [K \star m_v^\varepsilon]_{\mathcal{C}^\theta} &\lesssim \|m_v^\varepsilon\|_{L^1}^{1-(s+1+\theta)/d} \|m_v^\varepsilon\|_{L^\infty}^{(s+1+\theta)/d}\,.\end{aligned}$$

For the second term we simply bound $\|\nabla U - \nabla U^\varepsilon\|_{L^\infty} \leqslant \|\nabla^2 U\|_{L^\infty} \varepsilon$. Since m^ε is again uniformly bounded in $L^1 \cap L^\infty$, we get an error bound between m_t and m_t^ε for small t and we iterate infinitely. □

Finally, we prove the well-posedness of the particle system in the non-attractive sub-Coulombic and Coulombic cases.

Proof of Proposition 4.4. Define for $n \in \mathbb{N}$ the sequence of stopping times:

$$\tau_n := \inf\big\{ t \geqslant 0 : \big|X_t^i - X_t^j\big| \leqslant 1/n \text{ for some } i \neq j \big\}\,.$$

Then the original SDE system (4.2) stopped at τ_n is well defined according to Cauchy–Lipschitz theory. Consider the "energy" functional

$$E(\boldsymbol{x}) = E(x^1, \dots, x^N) = \frac{1}{2} \sum_{\substack{i,j \in [\![1,N]\!] \\ i \neq j}} g_s(x^i - x^j) + \frac{N \mathbb{1}_{s=0}}{2} \sum_{i=1}^N |x^i|^2\,.$$

The energy functional is always lower bounded, and by Itō calculus, we find that $\mathbb{E}\big[E\big(\boldsymbol{X}_{t\wedge\tau_n}\big)\big]$ is upper bounded uniformly in n. Then using the Markov inequality for the energy, we show that $\mathbb{P}[\tau_n \leqslant t] \to 0$ when $n \to \infty$. This implies that $\lim_{n\to\infty} \tau_n = \infty$ almost surely, thus the local well-posedness of the SDE extends to the half line $[0,\infty)$. That is to say the first claim is proved.

Now prove the second claim. For each $n \in \mathbb{N}$, we construct a Lipschitz kernel $\tilde{K}_n : \mathbb{R}^d \to \mathbb{R}$ such that $\tilde{K}_n(x) = K(x)$ for $x \in \mathbb{R}^d$ with $|x| \geqslant 1/n$. Define the convolution $\tilde{K}_n^\varepsilon = \tilde{K}_n \star \eta^\varepsilon$ and consider the SDE system

$$\mathrm{d}\tilde{X}_{n,t}^{\varepsilon,i} = -\nabla U^\varepsilon\big(\tilde{X}_{n,t}^{\varepsilon,i}\big)\,\mathrm{d}t + \frac{1}{N-1}\sum_{j\in[\![1,N]\!]\setminus\{i\}} \tilde{K}_n^\varepsilon\big(\tilde{X}_{n,t}^{\varepsilon,i} - \tilde{X}_{n,t}^{\varepsilon,j}\big)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t^i\,,$$

for $i \in [\![1,N]\!]$, with initial condition $\tilde{\boldsymbol{X}}_{n,0}^\varepsilon = \boldsymbol{X}_0$. Define the stopping time

$$\tau_n^\varepsilon \coloneqq \inf\big\{t \geqslant 0 : \big|\tilde{X}_{n,t}^{\varepsilon,i} - \tilde{X}_{n,t}^{\varepsilon,j}\big| \leqslant 1/n + \varepsilon \text{ for some } i \neq j\big\}\,.$$

By construction, we know

$$\tilde{\boldsymbol{X}}_{n,t\wedge\tau_n^\varepsilon}^\varepsilon = \boldsymbol{X}_{t\wedge\tau_n^\varepsilon}^\varepsilon \text{ a.s.}$$

On the other hand, by Cauchy–Lipschitz theory, we know

$$\sup_{t\in[0,T]}\big|\tilde{\boldsymbol{X}}_{n,t\wedge\tau_n}^\varepsilon - \boldsymbol{X}_{t\wedge\tau_n}\big| \leqslant C(n,N,K,U,T)\varepsilon \text{ a.s.}$$

Thus, for each $n \in \mathbb{N}$, there exists $\varepsilon_0(n,N,K,U,T) > 0$ such that for all $\varepsilon \leqslant \varepsilon_0$, we have

$$\sup_{t\in[0,T]}\big|\tilde{\boldsymbol{X}}_{n,t\wedge\tau_n}^\varepsilon - \boldsymbol{X}_{t\wedge\tau_n}\big| \leqslant \frac{1}{3n} \text{ a.s.}$$

In particular, we get for all $\varepsilon \leqslant \varepsilon_0$, $t \leqslant T\wedge\tau_n$ and $i \neq j$,

$$\big|\tilde{X}_{n,t}^{\varepsilon,i} - \tilde{X}_{n,t}^{\varepsilon,j}\big| \geqslant \frac{1}{3n} \text{ a.s.}$$

Consequently, for $\varepsilon \leqslant \varepsilon_1(n,N,K,U,T) \coloneqq \varepsilon_0 \wedge 1/(13n)$, we have $T\wedge\tau_n \leqslant \tau_{4n}^\varepsilon$, and therefore,

$$\begin{aligned}\sup_{t\leqslant T\wedge\tau_n}\big|\boldsymbol{X}_t^\varepsilon - \boldsymbol{X}_t\big| = \sup_{t\leqslant T\wedge\tau_n}\big|\boldsymbol{X}_{t\wedge\tau_{4n}^\varepsilon}^\varepsilon - \boldsymbol{X}_t\big| = \sup_{t\leqslant T\wedge\tau_n}\big|\tilde{\boldsymbol{X}}_{4n,t}^\varepsilon - \boldsymbol{X}_t\big| \\ \leqslant C(4n,N,K,U,T)\varepsilon \text{ a.s.}\end{aligned}$$

Thus, taking $\varepsilon \to 0$, we get $\boldsymbol{X}_{t\wedge\tau_n}^\varepsilon \to \boldsymbol{X}_{t\wedge\tau_n}$ a.s. for all $t \leqslant T$. We recover the second claim by using the arbitrariness of T and the fact that $\lim_{n\to\infty}\tau_n = \infty$ a.s. □

B Feynman–Kac formula

Proposition B.1. *Let $T > 0$. Suppose $\beta : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\varphi : [0,T]\times\mathbb{R}^d \to \mathbb{R}$ are measurable functions and suppose that there exists $C > 0$ such that for all $t \in [0,T]$ and $x \in \mathbb{R}^d$, we have*

$$\begin{aligned}|\beta(t,x)| &\leqslant C(1+|x|)\,,\\ |\varphi(t,x)| &\leqslant C(1+|x|)\,,\\ \big|\nabla_x^k\beta(t,x)\big| &\leqslant C\,, \quad \textit{for } k \in [\![1,3]\!]\,,\\ \big|\nabla_x^k\varphi(t,x)\big| &\leqslant C\,, \quad \textit{for } k \in [\![1,2]\!]\,.\end{aligned}$$

Suppose in addition that $f_0 : \mathbb{R}^d \to \mathbb{R}$ is measurable and satisfies, for the same constant C, and for all $x \in \mathbb{R}^d$,

$$|\nabla^k f_0(x)| \leqslant C \exp\big(C(1+|x|)\big)\,, \quad \textit{for } k \in [\![0,2]\!]\,.$$

Then, the function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ defined by

$$f(t,x) = \mathbb{E}\left[\exp\left(\int_0^t \varphi\big(t-u, X_u^{t,x}\big)\,\mathrm{d}u\right) f_0\big(X_t^{t,x}\big)\right],$$

where $X_\cdot^{t,x}$ solves

$$\mathrm{d}X_u^{t,x} = \beta\big(t-s, X_u^{t,x}\big)\,\mathrm{d}u + \sqrt{2}\,\mathrm{d}B_t\,, \quad u \in [0,t]\,, \qquad X_0^{t,x} = x\,,$$

is a strong solution to the Cauchy problem

$$\partial_t f = \Delta f + \beta \cdot \nabla f + \varphi f\,, \qquad f|_{t=0} = f_0$$

with the following bound: there exists $C' > 0$ such that for all $t \in [0,T]$ and $x \in \mathbb{R}^d$, we have

$$|\nabla^k f_0(t,x)| \leqslant C' \exp\big(C'(1+|x|)\big)\,, \quad \textit{for } k \in [\![0,2]\!]\,.$$

The result can be easily obtained by differentiating the defining SDE of the process $X_\cdot^{t,x}$. We refer readers to e.g. the appendix of [8] for details.

References

[1] Shigeki Aida and Ichiro Shigekawa. Logarithmic Sobolev inequalities and spectral gaps: Perturbation theory. *J. Funct. Anal.*, 126(2):448–475, 1994. ISSN: 0022-1236. DOI: 10.1006/jfan.1994.1154. MR1305076

[2] Dominique Bakry. On Sobolev and logarithmic Sobolev inequalities for Markov semigroups. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 43–75. World Sci. Publ., River Edge, NJ, 1997. ISBN: 981-02-2867-8. MR1654503

[3] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren Math. Wiss.* Springer, Cham, 2014. ISBN: 978-3-319-00226-2; 978-3-319-00227-9. DOI: 10.1007/978-3-319-00227-9. MR3155209

[4] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. Mean field limit and quantitative estimates with singular attractive kernels. *Duke Math. J.*, 172(13):2591–2641, 2023. ISSN: 0012-7094,1547-7398. DOI: 10.1215/00127094-2022-0088. MR4658923

[5] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. Modulated free energy and mean field limit. In *Sémin. Laurent Schwartz, EDP Appl., Année 2019-2020*, Exposé n° II. DOI: 10.5802/slsedp.135.

[6] Nicolas Champagnat and Pierre-Emmanuel Jabin. Strong solutions to stochastic differential equations with rough coefficients. *Ann. Probab.*, 46(3):1498–1541, 2018. ISSN: 0091-1798. DOI: 10.1214/17-AOP1208.

[7] Paul-Éric Chaudru de Raynal, Manh Hong Duong, Pierre Monmarché, Milica Tomašević, and Julian Tugaut. Reducing exit-times of diffusions with repulsive interactions. *ESAIM, Probab. Stat.*, 27:723–748, 2023. ISSN: 1292-8100. DOI: 10.1051/ps/2023012. MR4621420

[8] Fan Chen, Zhenjie Ren, and Songbo Wang. Uniform-in-time propagation of chaos for mean field Langevin dynamics. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 2024. arXiv: 2212.03050 [math.PR]. Forthcoming.

[9] Antonin Chodron de Courcel, Matthew Rosenzweig, and Sylvia Serfaty. Sharp uniform-in-time mean-field convergence for singular periodic Riesz flows. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire*, December 1, 2023. DOI: 10.4171/AIHPC/105. Pre-published.

[10] Giovanni Conforti. Coupling by reflection for controlled diffusion processes: turnpike property and large time behavior of Hamilton–Jacobi–Bellman equations. *Ann. Appl. Probab.*, 33(6A):4608–4644, 2023. ISSN: 1050-5164. DOI: 10.1214/22-AAP1927. MR4674060

[11] Giovanni Conforti. Weak semiconvexity estimates for Schrödinger potentials and logarithmic Sobolev inequality for Schrödinger bridges. *Probab. Theory Relat. Fields*, February 2024. ISSN: 0178-8051, 1432-2064. DOI: 10.1007/s00440-024-01264-6. MR4771110

[12] Joscha Diehl, Peter K. Friz, and Harald Oberhauser. Parabolic comparison revisited and applications. 2011. arXiv: 1102.5774 [math.AP].

[13] Andreas Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Relat. Fields*, 166(3-4):851–886, 2016. ISSN: 0178-8051. DOI: 10.1007/s00440-015-0673-1.

[14] Xuanrui Feng and Zhenfu Wang. Quantitative Propagation of Chaos for 2D Viscous Vortex Model on the Whole Space. 2023. arXiv: 2310.05156 [math.AP].

[15] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stoch. Model. Appl. Probab.* New York, NY: Springer, 2nd ed. Edition, 2006. ISBN: 0-387-26045-5; 978-1-4419-2078-2; 0-387-31071-1. DOI: 10.1007/0-387-31071-1.

[16] Arnaud Guillin, Pierre Le Bris, and Pierre Monmarché. Uniform in time propagation of chaos for the 2D vortex model and other singular stochastic systems. *J. Eur. Math. Soc.*, January 22, 2024. DOI: 10.4171/JEMS/1413. Pre-published.

[17] Richard Holley and Daniel Stroock. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.*, 46(5-6):1159–1194, 1987. ISSN: 0022-4715. DOI: 10.1007/BF01011161.

[18] Lars Hörmander. *The analysis of linear partial differential operators. I. Distribution theory and Fourier analysis.* Volume 256 of *Grundlehren Math. Wiss.* Springer-Verlag, Berlin etc., 2nd ed. Edition, 1990. ISBN: 3-540-52343-X; 3-540-52345-6. MR1065993

[19] Jiaoyang Huang, Matthew Rosenzweig, and Sylvia Serfaty. The modulated free energy method on $\mathbb{R}^d$. In preparation.

[20] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018. ISSN: 0020-9910. DOI: 10.1007/s00222-018-0808-y.

[21] Kazumasa Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.*, 258(11):3758–3774, 2010. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2010.01.010. MR2606871

[22] Daniel Lacker. Hierarchies, entropy, and quantitative propagation of chaos for mean field diffusions. *Probab. Math. Phys.*, 4(2):377–432, 2023. ISSN: 2690-0998. DOI: 10.2140/pmp.2023.4.377. MR4595391

[23] Daniel Lacker and Luc Le Flem. Sharp uniform-in-time propagation of chaos. *Probab. Theory Relat. Fields*, 187(1-2):443–480, 2023. ISSN: 0178-8051. DOI: 10.1007/s00440-023-01192-x.

[24] Tony Lelièvre, Mathias Rousset, and Gabriel Stoltz. Long-time convergence of an adaptive biasing force method. *Nonlinearity*, 21(6):1155–1181, 2008. ISSN: 0951-7715. DOI: 10.1088/0951-7715/21/6/001. MR2422373

[25] Yulong Lu. Two-Scale Gradient Descent Ascent Dynamics Finds Mixed Nash Equilibria of Continuous Games: A Mean-Field Perspective. In *Proc. 40th Int. Conf. Mach. Learn.* Volume 202 of *Proc. Mach. Learn. Res.* Pages 22790–22811. PMLR, 2023. URL: https://proceedings.mlr.press/v202/lu23b.html.

[26] Florent Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Processes Appl.*, 95(1):109–132, 2001. ISSN: 0304-4149. DOI: 10.1016/S0304-4149(01)00095-3. MR1847094

[27] Laurent Miclo. On hyperboundedness and spectrum of Markov operators. *Invent. Math.*, 200(1):311–343, 2015. ISSN: 0020-9910. DOI: 10.1007/s00222-014-0538-8.

[28] Pierre Monmarché. Wasserstein contraction and Poincaré inequalities for elliptic diffusions with high diffusivity. *Ann. Henri Lebesgue*, 6:941–973, 2023. DOI: 10.5802/ahl.182. MR4678939

[29] Pierre Monmarché and Songbo Wang. Logarithmic Sobolev inequalities for non-equilibrium steady states. 2023. arXiv: 2309.07858 [math.PR].

[30] Alessio Porretta and Enrico Priola. Global Lipschitz regularizing effects for linear and nonlinear parabolic equations. *J. Math. Pures Appl. (9)*, 100(5):633–686, 2013. ISSN: 0021-7824. DOI: 10.1016/j.matpur.2013.01.016. MR3115828

[31] Enrico Priola and Feng-Yu Wang. Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.*, 236(1):244–264, 2006. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2005.12.010.

[32] Panpan Ren and Feng-Yu Wang. Entropy Estimate Between Diffusion Processes with Application to Nonlinear Fokker–Planck Equations. 2023. arXiv: 2302.13500 [math.PR].

[33] Michael Röckner and Feng-Yu Wang. Log-Harnack inequality for stochastic differential equations in Hilbert spaces and its consequences. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 13(1):27–37, 2010. ISSN: 0219-0257. DOI: 10.1142/S0219025710003936.

[34] Matthew Rosenzweig and Sylvia Serfaty. Global-in-time mean-field convergence for singular Riesz-type diffusive flows. *Ann. Appl. Probab.*, 33(2):954–998, 2023. ISSN: 1050-5164. DOI: 10.1214/22-AAP1833. MR4564418

[35] Matthew Rosenzweig and Sylvia Serfaty. Modulated logarithmic Sobolev inequalities and generation of chaos. *Ann. Fac. Sci. Toulouse, Math. (6)*, 2023. arXiv: 2307.07587 [math.PR]. Forthcoming.

[36] Sylvia Serfaty. Mean field limit for Coulomb-type flows. *Duke Math. J.*, 169(15):2887–2935, 2020. ISSN: 0012-7094. DOI: 10.1215/00127094-2020-0019. Appendix by Mitia Duerinckx and Sylvia Serfaty. MR4158670

[37] Feng-Yu Wang. Criteria of spectral gap for Markov operators. *J. Funct. Anal.*, 266(4):2137–2152, 2014. ISSN: 0022-1236. DOI: 10.1016/j.jfa.2013.11.016. MR3150155

[38] Feng-Yu Wang. Exponential contraction in Wasserstein distances for diffusion semigroups with negative curvature. *Potential Anal.*, 53(3):1123–1144, 2020. ISSN: 0926-2601. DOI: 10.1007/s11118-019-09800-z.