On the Law of Large Numbers for the Empirical Measure Process of Generalized Dyson Brownian Motion

Songzi Li · Xiang-Dong Li · Yong-Xiao Xie

Received: 24 April 2019 / Accepted: 14 August 2020 / Published online: 21 September 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract
We study the generalized Dyson Brownian motion (GDBM) of an interacting \(N\)-particle system with logarithmic Coulomb interaction and general potential \(V\). Under reasonable condition on \(V\), we prove the existence and uniqueness of strong solution to SDE for GDBM. We then prove that the family of the empirical measures of GDBM is tight on \(C([0, T], \mathcal{P}(\mathbb{R}))\) and all the large \(N\) limits satisfy a nonlinear McKean–Vlasov equation. Inspired by previous works due to Biane and Speicher, Carrillo, McCann and Villani, and Blower, we prove that the McKean–Vlasov equation is indeed the gradient flow of the Voiculescu free entropy on the Wasserstein space of probability measures over \(\mathbb{R}\). Using the optimal transportation theory, we prove that if \(V'' \geq K\) for some constant \(K \in \mathbb{R}\), the McKean–Vlasov equation has a unique weak solution in the space of probability measures \(\mathcal{P}(\mathbb{R})\). This establishes the Law of Large Numbers and the propagation of chaos for the empirical measures of GDBM with non-quadratic external potentials which are not necessarily convex. Finally, we prove the longtime convergence of the McKean–Vlasov equation for \(C^2\)-convex potentials \(V\).

Keywords
Generalized Dyson Brownian motion · McKean–Vlasov equation · Gradient flow · Optimal transportation · Voiculescu free entropy · Law of Large Numbers · Propagation of chaos

1 Introduction

1.1 Background

In 1962, Dyson [17,18] observed that the eigenvalues of the \(N \times N\) Hermitian matrix valued Brownian motion is an interacting \(N\)-particle system with the logarithmic Coulomb interaction and derived their statistical properties. Since then, the Dyson Brownian motion has been used in various areas in mathematics and physics, including statistical physics and...
the quantum chaotic systems. See e.g. [29] and reference therein. In [34], Rogers and Shi proved that the empirical measure of the eigenvalues of the $N \times N$ Hermitian matrix valued Ornstein–Uhlenbeck process weakly converges to the nonlinear McKean–Vlasov equation with quadratic external potential as $N$ tends to infinity. This also gave a dynamic proof of Wigner’s famous semi-circle law for Gaussian Unitary Ensemble. See also [2,23].

The purpose of this paper is to study the generalized Dyson Brownian motion and the associated McKean–Vlasov equation with the logarithmic Coulomb interaction and with general external potential. More precisely, let $\beta \geq 1$ be a parameter, $V : \mathbb{R} \to \mathbb{R}^+$ be a continuous function, let $(W^1, \ldots, W^N)$ be an $N$-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathbb{F}_t)_{t \geq 0}, \mathbb{P})$ satisfying the usual conditions. Let $\lambda_N(0) = (\lambda^1_N(0), \ldots, \lambda^N_N(0)) \in \Delta_N = \{(x_i)_{1 \leq i \leq N} \in \mathbb{R}^N : x_1 < x_2 < \ldots < x_N\}$. The generalized Dyson Brownian motion (GDBM)$_V$ is an interacting $N$-particle system $\lambda_N(t) = (\lambda^1_N(t), \ldots, \lambda^N_N(t))$ with the logarithmic Coulomb interaction and with external potential $V$, and is defined as the solution to the following SDEs

$$
d\lambda^i_N(t) = \sqrt{\frac{2}{\beta N}} dW^i_t + \frac{1}{N} \sum_{j \neq i} \frac{1}{\lambda^j_N(t) - \lambda^i_N(t)} dt - \frac{1}{2} V'(\lambda^i_N(t)) dt, \quad i = 1, \ldots, N,
$$

with initial data $\lambda_N(0)$. It is a SDE for $N$-particles with a singular drift of the form $\frac{1}{x-y}$ due to the logarithmic Coulomb interaction, and an additional nonlinear drift due to non quadratic external potential. When $V = 0$ and $\beta = 1, 2, 4$, it is the standard Dyson Brownian motion [17,18]. When $V(x) = \frac{x^2}{2}$ and $\beta > 1$, it has been studied by Chan [13], Rogers and Shi [34], Cépa and Lépingle [12], Fontbona [20,21], Guionnet [23], Anderson, Guionnet and Zeitouni [2] and references therein.

By Itô’s calculus, $(\text{GDBM})_V$ is an interacting $N$-particle system with the Hamiltonian

$$H_N(x_1, \ldots, x_N) := -\frac{1}{N} \sum_{1 \leq i \neq j \leq N} \log |x_i - x_j| + \frac{1}{2} \sum_{i=1}^N V(x_i),$$

and the infinitesimal generator of $(\text{GDBM})_V$ is given by

$$L^\beta_N f = \frac{1}{\beta N} \sum_{k=1}^N \frac{\partial^2 f}{\partial x_k^2} + \sum_{k=1}^N \left( P \mathcal{F}_t, \int_{\mathbb{R}} \frac{L_N(dy)}{x_k - y} - \frac{1}{2} V'(y) \right) \frac{\partial f}{\partial x_k},$$

where $f \in C^2(\mathbb{R}^N)$ and $L_N = \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \in \mathcal{P}(\mathbb{R})$, the space of probability measures on $\mathbb{R}$.

Under suitable condition on $V$, we prove that the SDEs (1) for $(\text{GDBM})_V$ admit a unique strong solution $\lambda_N(t) \in \Delta_N$ with infinite lifetime. See Theorem 1.1. Let

$$L_N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda^i_N(t)} \in \mathcal{P}(\mathbb{R}), \quad t \in [0, \infty).$$

Standard argument shows that the family $\{L_N(t), t \in [0, T]\}$ is tight on $C([0, T], \mathcal{P}(\mathbb{R}))$, and the limit of any weakly convergent subsequence of $L_N(t)$, denoted by $\mu_t$, is a weak
solution in the space of probability measures $\mathcal{P}(\mathbb{R})$, to the following nonlinear McKean–Vlasov equation: for all $f \in C^2_b(\mathbb{R})$,

$$
\frac{d}{dt} \int_{\mathbb{R}} f(x) \mu_t(dx) = \frac{1}{2} \int_{\mathbb{R}^2} \frac{\partial_x f(x) - \partial_y f(y)}{x - y} \mu_t(dx) \mu_t(dy) - \frac{1}{2} \int_{\mathbb{R}} V'(x) f'(x) \mu_t(dx).
$$

(3)

In the case $\mu_t$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, integrating by parts, one can verify that the probability density $\rho_t = \frac{d\mu_t}{dx}$ satisfies the following nonlinear McKean–Vlasov equation (also called nonlinear Fokker–Planck equation in [5])

$$
\frac{\partial \rho_t}{\partial t} = -\frac{\partial}{\partial x} \left( \rho_t \left( \frac{1}{2} V' - H \rho_t \right) \right),
$$

(4)

where

$$
H \rho(x) = \text{P.V.} \int_{\mathbb{R}} \frac{\rho(y)}{x - y} dy
$$

is the Hilbert transform of $\rho$ (up to a multiplicative constant $\frac{1}{\pi}$).

The McKean–Vlasov equation (4) is a nonlinear singular integro-differential equation. In the case where $V$ is a quadratic function, for example, $V(x) = \theta x^2$, one can use the Stieltjes transform to study the McKean–Vlasov equation. Indeed, to characterize the McKean–Vlasov limit $\mu_t$, we need only to use the test function $f(x) = (z - x)^{-1}$, where $z \in \mathbb{C} \setminus \mathbb{R}$, instead of using all test functions $f \in C^2_b(\mathbb{R})$ in the McKean–Vlasov equation (3). Let

$$
G_t(z) = \int_{\mathbb{R}} \frac{\mu_t(dx)}{z - x}
$$

be the Stieltjes transform of $\mu_t$. Then $G_t(z)$ satisfies the following equation

$$
\frac{\partial}{\partial t} G_t(z) = -G_t(z) \frac{\partial}{\partial z} G_t(z) - \frac{1}{2} \int_{\mathbb{R}} \frac{V'(x)}{(z - x)^2} \mu_t(dx).
$$

(5)

In particular, in the case $V(x) = \frac{Kx^2}{2}$, since

$$
- \int_{\mathbb{R}} \frac{x}{(z - x)^2} \mu_t(dx) = z \frac{\partial}{\partial z} G_t(z) + G_t(z),
$$

the Stieltjes transform of $\mu_t$ satisfies the complex Burgers equation

$$
\frac{\partial}{\partial t} G_t(z) = (-G_t(z) + Kz) \frac{\partial}{\partial z} G_t(z) + KG_t(z).
$$

(6)

In [13,34], Chan and Rogers-Shi proved that the complex Burgers equation (6) has a unique solution, and $\lim_{t \to \infty} G_t(z)$ exists and coincides with the Stieltjes transform of the Wigner semicircle law $\mu_{SC}$. This gives a dynamic proof of the Wigner’s theorem, i.e., $L_N(\infty)$ weakly converges to $\mu_{SC}$. See also [2,23].

In general case of non quadratic potential $V$, Biane and Speicher [5] and Biane [4] proved the existence of weak solution to the McKean–Vlasov equation (called the free Fokker–Planck equation in their papers) by the method of free probability theory. However, as far as we know, it seems that one can not find well-established result in the literature on the uniqueness of weak solution to the McKean–Vlasov equation with general external potential $V$. By lack of this, one can not find established result in the literature on the Law of Large Numbers for the GDBM with non quadratic potentials. The difficulty for non quadratic potential $V$, at least to
us, is due to the fact that \( \int R V'(x)(z-x)2 \mu_t(dx) \) in (5) cannot be expressed in terms of \( G_t(z) \) and its derivatives with respect to \( z \). Thus, one cannot derive an analogue of the complex Burgers equation (6) for non quadratic potential \( V \), and we need to find a new approach to prove the uniqueness result of weak solution to the McKean–Vlasov equation for general potential \( V \).

One of the main contributions of this paper is to observe and to prove the fact that the McKean–Vlasov equation (4) is indeed the gradient flow of the Voiculescu free entropy \( \Sigma_V \) on the Wasserstein space \( \mathcal{P}_2(\mathbb{R}) \) equipped with Otto’s infinite dimensional Riemannian structure, and to use the optimal transportation theory to prove the uniqueness of weak solution to the McKean–Vlasov equation (4) for general potentials \( V \in C^2(\mathbb{R}) \) with \( V'' \geq K \) for \( K \in \mathbb{R} \). This establishes the Law of Large Numbers and the propagation of chaos for the empirical measures of GDBM with non-quadratic external potentials which are not necessarily convex. We also study the longtime asymptotic behavior of the weak solution to the McKean–Vlasov equation (4) for \( C^2 \)-convex potentials \( V \). The method of this paper might be further extended to study similar problems for other interacting particle systems in physics with singular interaction potential \( W \) and general external potential \( V \).

### 1.2 Statement of Results

To state our results, let us introduce some important quantities. Following Voiculescu [41], Biane-Speicher [5] and Biane [4], for every \( \mu \in \mathcal{P}(\mathbb{R}) \), we introduce the Voiculescu free entropy (called also the free energy in some literature) as follows

\[
\Sigma_V(\mu) = -\frac{1}{2} \int_{\mathbb{R}^2} \log |x-y|d\mu(x)d\mu(y) + \frac{1}{2} \int_{\mathbb{R}} V(x)d\mu(x).
\]

By [25], it is well-known that if \( V \) satisfies the growth condition

\[
V(x) \geq (1+\delta) \log(x^2+1), \quad x \in \mathbb{R},
\]

then there exists a unique minimizer (called the equilibrium measure) of \( \Sigma_V \), denoted by

\[
\mu_V = \arg\min_{\mu \in \mathcal{P}(\mathbb{R})} \Sigma_V(\mu).
\]

Moreover, \( \mu_V \) satisfies the Euler-Lagrange equation

\[
\mathcal{H}_{\mu_V}(x) = \frac{1}{2} V'(x), \quad \forall x \in \text{supp}(\mu_V).
\]

The relative free entropy is defined as follows

\[
\Sigma_V(\mu|\mu_V) = \Sigma_V(\mu) - \Sigma_V(\mu_V).
\]

Following [4,5], the relative free Fisher information is defined as follows

\[
I_V(\mu) = \int_{\mathbb{R}} \left( \mathcal{H}_{\mu}(x) - \frac{1}{2} V'(x) \right)^2 d\mu(x).
\]

Note that

\[
I_V(\mu_V) = 0.
\]

We now state the main results of this paper. Our first result establishes the existence and uniqueness of the strong solution to SDEs (1) and the tightness of the associated empirical measure for a class of \( \bar{V} \) with reasonable condition.
Theorem 1.1 Let $V$ be a $C^1$ function satisfying the growth condition (7) and the following conditions

(i) For all $R > 0$, there is a constant $K_R > 0$, such that for all $x, y \in \mathbb{R}$ with $|x|, |y| \leq R$,

$$(x - y)(V'(x) - V'(y)) \geq -K_R|x - y|^2,$$

(ii) There exists a constant $\gamma > 0$ such that

$$-x V'(x) \leq \gamma (1 + |x|^2), \; \forall x \in \mathbb{R}. \quad (8)$$

Then, for all $\beta \geq 1$, and for any given $\lambda_N(0) \in \Delta_N$, there exists a unique strong solution $(\lambda_N(t))_{t \geq 0}$ taking values in $\Delta_N$ with infinite lifetime to SDEs (1) with initial value $\lambda_N(0)$.

Moreover, suppose that $L_N(0) \to \mu \in \mathcal{P}(\mathbb{R})$ as $N \to \infty$, and

$$\sup_{N \geq 0} \int_{\mathbb{R}} \log(x^2 + 1) dL_N(0) < \infty.$$ 

Then, the family $\{L_N(t), t \in [0, T]\}$ is tight in $C([0, T], \mathcal{P}(\mathbb{R}))$, and the limit of any weakly convergent subsequence of $\{L_N(t), t \in [0, T]\}$ is a weak solution of the McKean–Vlasov equation (3).

Remark 1.2 We refer the reader to [2, 24] for the standard definition of strong solution to stochastic differential equation (SDE). Theorem 1.1 can be viewed as an analogue of Krylov’s existence and uniqueness theorem of general SDEs (see Theorem 3.1.1 in [33]) for generalized Dyson Brownian motion of $N$-particles with logarithmic interacting interaction and external potential $V$. Let us mention that, under the condition $-x V'(x) \leq C$ for all $x \in \mathbb{R}$, Rogers and Shi [34] proved the non-collision of the strong solution to (1), but they did not precisely state the condition (i) which is needed for the existence of solution. In [22], Graczyk and Malecki proved the existence and uniqueness of strong solution to SDE (1) under the assumption that $V'$ is global Lipschitz. The conditions in Theorem 1.1 require that $V'$ satisfies the local monotonicity condition, i.e., (i), and one-side growth condition at infinity, i.e., (ii). We would like to point out that the local monotonicity condition (i) in Theorem 1.1 is weaker than the condition $V'$ is local Lipschitz, and the one-side growth condition (ii) is also weaker than the condition $V'$ is global Lipschitz.

In Theorem 1.1, we have proved the existence of weak solution to the McKean–Vlasov equation (3). By Theorem 3.1 and Theorem 5.2 in Biane and Speicher [5], for any initial distribution $\mu_0 = \rho_0 dx \in \mathcal{P}(\mathbb{R})$ with compact support, any weak solution $\mu_t$ of the McKean–Vlasov equation (3) is absolutely continuous with respect to the Lebesgue measure $dx$, and the density $\rho_t = \frac{d\mu_t}{dt}$ is a weak solution of the nonlinear McKean–Vlasov equation (4) (called the free Fokker–Planck equation in [5]). Moreover, they proved that there exists a constant $M > 0$ which does not depend on $t$ such that all the weak solutions to the McKean–Vlasov equation (4) satisfy

$$\text{supp}(\mu_t) \subseteq [-M, M],$$

and

$$\|\rho_t\|_\infty \leq \frac{K_1}{\sqrt{t}} + K_2,$$

$$\|D^{1/2} \rho_t\|_2 \leq \frac{K_1}{t} + K_2,$$
where $K_1$ and $K_2$ are two constants which are independent of $t$, $D^{1/2}$ is the fractional derivative of order $1/2$, and \( \|D^{1/2}\rho\|_2^2 = \int_{\mathbb{R}} |\hat{\rho}|^2 (\xi) d\xi \), $\hat{\rho}$ is the Fourier transform of $\rho$. Equivalently to say, for initial datum $\rho_0$ with compact support, all the weak solutions $\rho_t$ to the McKean–Vlasov equation (4) are compactly supported and for any $t > 0$, $\rho_t \in L^\infty(\mathbb{R}) \cap H^{1/2}(\mathbb{R})$. See Sect. 4 for the definition of $H^{1/2}(\mathbb{R})$. Note that, $u \in H^{1/2}(\mathbb{R})$ if and only if $u \in L^2(\mathbb{R})$ and
\[
\int_{\mathbb{R}^2} \frac{|u(x) - u(y)|^2}{|x - y|^2} dx dy < \infty.
\]

To prove the Law of Large Numbers for $L_N(t)$, we need only to show the uniqueness of weak solution to the nonlinear McKean–Vlasov equation (4). To do so, we use the optimal transportation theory, in particular, the theory of the gradient flows and Riemannian geometry on the $L^2$-Wasserstein space.

Now we briefly recall some basic notions of the theory of optimal transportation. Recall that, for any $p \geq 1$, the $W_p$-Wasserstein distance between two probability measures $\mu_0$ and $\mu_1$ on $\mathbb{R}^n$ is defined to be
\[
W_p(\mu_0, \mu_1) = \inf \left\{ \int_{\mathbb{R}^{2n}} |x - y|^p d\pi(x, y) : \pi \in \Pi(\mu_0, \mu_1) \right\}^{1/p},
\]
where $\Pi(\mu_0, \mu_1) = \{ \pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n) : \pi(\cdot, \mathbb{R}^n) = \mu_0, \pi(\mathbb{R}^n, \cdot) = \mu_1 \}$.

Let $\mathcal{P}_2(\mathbb{R}^n) = \{ \mu \in \mathcal{P}(\mathbb{R}^n) : \int_{\mathbb{R}^n} |x|^2 d\mu(x) < \infty \}$ be the $L^2$-Wasserstein space on $\mathbb{R}^n$. Let $\mathcal{P}_2^\infty(\mathbb{R}^n) = \{ \mathcal{P}_2(\mathbb{R}^n) : d\mu(x) = f(x) dx, f \text{ is positive, smooth and rapidly decay} \}$. Following Otto [30] and Otto-Villani [31], given $d\mu = f dx \in \mathcal{P}_2^\infty(\mathbb{R}^n)$, the tangent space of $\mathcal{P}_2^\infty(\mathbb{R}^n)$ at $\mu$ is the space of signed measures $s(x) dx$, where $s$ is a smooth and rapidly decay function on $\mathbb{R}^n$ such that $\int_{\mathbb{R}^n} s(x) dx = 0$. Let $\nabla$ be the divergence operator on $\mathbb{R}^n$. By the theory of elliptic PDEs, there exists a smooth and rapidly decreasing function $p$ such that
\[
s = -\nabla \cdot (f \nabla p).
\]
The tangent space $T_{f(x)dx} \mathcal{P}_2^\infty(\mathbb{R}^n)$ (in Otto’s Calculus) is defined by (see [30,31,38,39])
\[
T_{f(x)dx} \mathcal{P}_2^\infty(\mathbb{R}^n) = \{ s = -\nabla \cdot (f \nabla p) : p \text{ is smooth and rapidly decay} \}.
\]
By Ambrosio–Gigli–Savaré [1], the tangent space of $\mathcal{P}_2(\mathbb{R}^n)$ at $\mu = f(x) dx$, denoted by $T_{\mu} \mathcal{P}_2(\mathbb{R}^n)$, is the $L^2(\mu)$-closure of the set $\{ \nabla p : p \text{ is smooth and rapidly decay} \}$, and is identified to $\{ v \in L^2(\mu) : \int_{\mathbb{R}^n} v \mu = 0 \}$ for all $w \in L^2(\mu)$ such that $\nabla \cdot (w\mu) = 0$. Indeed, for smooth and rapidly decay functions $f$ and $p$, the above definitions are equivalent under the correspondence between $v = \nabla p$ and $s = -\nabla \cdot (f \nabla p)$.

Following Otto [30], the infinite dimensional Riemannian inner product on $T_{\mu} \mathcal{P}_2(\mathbb{R}^n)$ is defined as follows:
\[
\langle \langle v_1, v_2 \rangle \rangle = \int_{\mathbb{R}^n} \langle v_1(x), v_2(x) \rangle d\mu(x), \quad \forall v_1, v_2 \in T_{\mu} \mathcal{P}_2(\mathbb{R}^n).
\]
For any $v \in T_{\mu} \mathcal{P}_2(\mathbb{R}^n)$, the norm of $v$ is given by
\[
\|v\|^2 = \int_{\mathbb{R}^n} |v(x)|^2 d\mu(x).
\]
In view of this infinite dimensional inner product on the tangent space $T_{\mu} \mathcal{P}_2(\mathbb{R}^n)$, the Wasserstein space $\mathcal{P}_2(\mathbb{R}^n)$ is an infinite dimensional Riemannian manifold.
Let $F : \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}$ be a differentiable functional. The differential of $F$, denoted by $dF$, determines a unique tangent vector field on $\mathcal{P}_2(\mathbb{R}^n)$, called the gradient of $F$ and denoted by $\nabla \mathcal{P}_2(\mathbb{R}^n) F$, such that, for all $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ and for all tangent vector $v \in T_\mu \mathcal{P}_2(\mathbb{R}^n)$, it holds

$$
\langle \langle \nabla \mathcal{P}_2(\mathbb{R}^n) F(\mu), v \rangle \rangle = dF(\mu)s := \lim_{\varepsilon \to 0} \frac{F(\mu + \varepsilon s) - F(\mu)}{\varepsilon}.
$$

(9)

where $s = -\nabla \cdot (\mu v)$.

By [10,30,38,39], for any differentiable functional $F : \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}$, its gradient on $\mathcal{P}_2(\mathbb{R}^n)$ is given by

$$
\nabla \mathcal{P}_2(\mathbb{R}^n) F(\mu) = -\nabla \cdot \left( \rho \nabla \frac{\delta F}{\delta \rho} \right),
$$

(10)

where $\frac{\delta F}{\delta \rho}$ denotes the $L^2(dx)$-derivative of $F$ with respect to $\rho$, such that

$$
dF(\mu)s = \int_{\mathbb{R}^n} \frac{\delta F}{\delta \rho}(x)s(x)dx.
$$

**Definition 1.3** [10,30,38,39] Let $F : \mathcal{P}_2(\mathbb{R}^n) \to \mathbb{R}$ be a differentiable functional. The gradient flow of $F$ is a $C^1$-smooth curve $(\mu_t = \rho_t dx, t \in [-T, T])$ in $\mathcal{P}_2(\mathbb{R}^n)$ for some constant $T > 0$ such that

$$
\frac{d}{dt} \mu_t = -\nabla \mathcal{P}_2(\mathbb{R}^n) F(\mu_t).
$$

(11)

In view of (10), the gradient flow of $F$ is defined by

$$
\frac{\partial \rho_t}{\partial t} = \nabla \cdot \left( \rho_t \nabla \frac{\delta F}{\delta \rho} \right).
$$

(12)

By [9,11,28], given $\mu_0 = \rho_0 dx, \mu_1 = \rho_1 dx \in \mathcal{P}_2(\mathbb{R})$ with densities $\rho_0, \rho_1$ of compact supports, there exists a unique increasing function $\phi : \mathbb{R} \to \mathbb{R}$ such that $\phi_* \mu_0 = \mu_1$, i.e.,

$$
\int_{-\infty}^{x} \rho_0(y)dy = \int_{-\infty}^{\phi(x)} \rho_1(y)dy,
$$

and solves the Monge–Kantorovich minimizing problem

$$
W_2^2(\mu_0, \mu_1) = \int_{\mathbb{R}} |x - \phi(x)|^2 \rho_0(x)dx.
$$

In the case where $\rho_0$ and $\rho_1$ are $C^{0, \alpha}$-continuous, where $\alpha \in (0, 1)$ then $\phi$ is $C^{1, \alpha}$-smooth and it holds

$$
\rho_0(x) = \phi'(x) \rho_1(\phi(x)), \quad \forall x \in \mathbb{R}.
$$

Let $\phi_s = (1-s)\text{Id} + s\phi$, $s \in [0, 1]$. By McCann [28], the family of probability measures

$$
\mu_s = (\phi_s)_* \mu_0
$$

is the unique geodesic linking $\mu_0$ and $\mu_1$ on $\mathcal{P}_2(\mathbb{R})$ such that

$$
W_2^2(\mu_0, \mu_s) = \int_{\mathbb{R}} |x - \phi_s(x)|^2 \rho_0(x)dx = s^2 W_2^2(\mu_0, \mu_1).
$$
In the case where \( \rho_0 \) and \( \rho_1 \) are \( C^{0,\alpha} \)-continuous, where \( \alpha \in (0,1) \), then \( \rho_s(x) = \frac{d\mu_s}{dx} \) exists and satisfies
\[
\rho_0(x) = \phi'_s(x) \rho_s(\phi_s(x)), \quad \forall x \in \mathbb{R}. \tag{13}
\]

According to McCann [28], see also Otto [30], Carlen and Gangbo [10] and Villani [38, 39], if the curve \( s \to F(\mu_s) \) is twice differentiable along the geodesic \( (\mu_s, s \in [0,1]) \) on \( \mathcal{P}_2(\mathbb{R}^n) \) with the initial tangent vector \( v \in T_{\mu_0} \mathcal{P}_2(\mathbb{R}) \), we can formally define the Hessian of \( F \) along \( v \) as follows
\[
\text{Hess}_{\mathcal{P}_2(\mathbb{R}^n)} F(\mu_0)(v, v) := \frac{d^2}{ds^2} \bigg|_{s=0} F(\mu_s). \tag{14}
\]

According to McCann [28], we say that \( F \) is a geodesically \( K \)-convex (or \( K \)-displacement convex) functional on \( \mathcal{P}_2(\mathbb{R}^n) \) for some \( K \in \mathbb{R} \), if for any \( \mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^n) \) and for the unique geodesic \( \mu_s \) between \( \mu_0 \) and \( \mu_1 \), i.e., \( \mu_s = ((1-s)id + s\phi)\ast\mu_0, s \in [0,1] \), it holds
\[
F(\mu_s) \leq (1-s)F(\mu_0) + s F(\mu_1) - \frac{K}{2} s(1-s) W_2^2(\mu_0, \mu_1).
\]

In the case where \( F \) is twice differentiable, one can check that \( F \) is \( K \)-displacement convex on \( \mathcal{P}_2(\mathbb{R}^n) \) if the formally defined Hessian of \( F \) satisfies
\[
\text{Hess}_{\mathcal{P}_2(\mathbb{R}^n)} F(\mu)(v, v) \geq K \|v\|^2 = K \int_{\mathbb{R}^n} |v|^2 d\mu,
\]
for all \( \mu \in \mathcal{P}_2^\infty(\mathbb{R}^n) \) and \( v \in T_{\mu} \mathcal{P}_2^\infty(\mathbb{R}^n) \). See [10, 11, 30, 31, 38, 39].

Let
\[
F(\rho) = \int_{\mathbb{R}^n} \rho \log \rho dv + \int_{\mathbb{R}^n} \rho V dv + \frac{1}{2} \int_{\mathbb{R}^{2n}} W(x-y) \rho(x) \rho(y) dx dy.
\]

In [11], Carrillo, McCann and Villani proved that, if \( V \) and \( W \) are nice functions on \( \mathbb{R}^n \), and \( W(x) = W(-x) \) for all \( x \in \mathbb{R}^n \), then the following type McKean–Vlasov evolution equation of the granular media
\[
\partial_t \rho = \nabla \cdot (\rho \nabla (\log \rho + V + W * \rho))
\]
can be realized as a gradient flow of \( F \) on the Wasserstein space \( \mathcal{P}_2(\mathbb{R}^n) \) equipped with Otto’s infinite dimensional Riemannian metric. Moreover, they proved the following entropy dissipation formulae: Let \( \xi := \nabla (\log \rho + V + W * \rho) \). Then
\[
\frac{d}{dt} F(\rho_t) = -\int_{\mathbb{R}^n} \rho |\xi|^2 dv,
\]
\[
\frac{d^2}{dt^2} F(\rho_t) = 2 \int_{\mathbb{R}^n} \rho \text{Tr} (D\xi)^T (D\xi) dx + 2 \int_{\mathbb{R}^n} (D^2 V \cdot \xi, \xi) \rho dx
\]
\[
+ \int_{\mathbb{R}^{2n}} (D^2 W(x-y) \cdot [\xi(x) - \xi(y)], [\xi(x) - \xi(y)]) d\rho(x)d\rho(y).
\]

In our situation, the interaction function \( W(x) = -\log |x| \) is singular at the origin \( x = 0 \), and does not satisfy the condition required in Carrillo–McCann–Villani [11]. This causes the main difficulty for directly applying the result of Carrillo–McCann–Villani to study the singular nonlinear McKean–Vlasov integro-differential equation (4). In this paper, using some regularization argument involving the Hilbert transformation, and based on \( L^\infty \cap H^{1/2-} \) estimates for the McKean–Vlasov equation (4) due to Biane-Speicher [5], we prove the
following results which plays an important role in the proofs of other main results of this paper.

**Theorem 1.4** Let \( \mu_0 = \rho_0 dx \) with \( \rho_0 \in L^\infty(\mathbb{R}) \) and with compact support, and \( v \in H^{1/2}(\mathbb{R}) \). Then

\[
\text{Hess}_{\mathcal{P}_2(\mathbb{R})} \Sigma_V(\mu_0)(v, v) = \frac{1}{2} \int_{\mathbb{R}} V''(x)|v(x)|^2 \rho_0(x) dx + \frac{1}{2} \int_{\mathbb{R}^2} \frac{|v(x) - v(y)|^2}{(x-y)^2} \rho_0(x) \rho_0(y) dx dy.
\]

(15)

In particular, if \( V \in C^2(\mathbb{R}) \) and \( V'' \geq 2K \) for some constant \( K \in \mathbb{R} \), then for all \( v \in H^{1/2}(\mathbb{R}) \)

\[
\text{Hess}_{\mathcal{P}_2(\mathbb{R})} \Sigma_V(\mu_0)(v, v) \geq K \|v\|^2.
\]

(16)

As a byproduct of Theorem 1.4, we can recapture the following result which was proved earlier for \( K > 0 \) by an alternative approach in Blower [6].

**Corollary 1.5** [7] Assume that \( V \in C^2(\mathbb{R}) \) with \( V'' \geq 2K \). Then \( \Sigma_V \) is geodesically \( K \)-convex on \( \mathcal{P}_2(\mathbb{R}) \) in the sense that

\[
\Sigma_V(\mu_s) \leq (1-s)\Sigma_V(\mu_0) + s\Sigma_V(\mu_1) - \frac{K}{2} s(1-s)W^2_2(\mu_0, \mu_1), \quad \forall s \in [0, 1].
\]

(17)

**Theorem 1.6** Suppose that \( V \in C^1(\mathbb{R}, \mathbb{R}^+) \) and \( V \) satisfies the condition (8). Then the nonlinear McKean–Vlasov equation (4) is the gradient flow of \( \Sigma_V \) on the Wasserstein space \( \mathcal{P}_2(\mathbb{R}) \). Denote \( v_t = \frac{V}{2t} - H\rho_t \). We have

\[
\frac{d}{dt} \Sigma_V(\mu_t) = - \int_{\mathbb{R}} |v_t(x)|^2 \rho_t(x) dx,
\]

(18)

Moreover, when \( V \in C^2(\mathbb{R}) \), we have

\[
\frac{d^2}{dt^2} \Sigma_V(\mu_t) = \frac{1}{2} \int_{\mathbb{R}} V''(x)|v_t(x)|^2 \rho_t(x) dx + \frac{1}{2} \int_{\mathbb{R}^2} \frac{|v_t(x) - v_t(y)|^2}{(x-y)^2} \rho_t(x) \rho_t(y) dx dy.
\]

(19)

In particular, if \( V'' \geq 2K \) for some constant \( K \in \mathbb{R} \), then

\[
\frac{d^2}{dt^2} \Sigma_V(\mu_t) \geq K \int_{\mathbb{R}} |v_t(x)|^2 \rho_t(x) dx.
\]

(20)

**Remark 1.7** Biane and Speicher [5] proved the existence of weak solutions to the McKean–Vlasov equation (4) (called the free Fokker–Planck equation in [5]) for the case \( \beta = 2 \). See also Biane [4] and Blower [7]. To overcome the difficulty due to the singularity of the logarithmic function at origin, they first proved the existence of a diffusion process on \( N \times N \) Hermitian random matrices and then derived the gradient flow equation by showing that the subsequence of empirical measure of eigenvalues of the Hermitian diffusion matrices converges weakly to a weak solution of the McKean–Vlasov equation (4). For general \( \beta \), this approach does not work. The uniqueness of weak solution of the McKean–Vlasov equation (4) was not proved in [4–7]. Moreover, we would like to point out that the first order entropy dissipation formula (18) in Theorem 1.6 has been already proved by Biane and Speicher (see Proposition 6.2 in [5]) by an approximation approach based on the convergence of the Euler scheme to the free SDE related to the free Fokker–Planck equation (4). Our approach uses the optimal transportation theory.
Recall that, in the optimal transportation theory, it is well-known that if $F$ is a $K$-displacement convex functional on the Wasserstein space, then the $W_2$-Wasserstein distance between any two solutions of the gradient flow $\partial_t \mu = -\nabla \mathcal{P}_2(\mathbb{R}^n) F(\mu)$ with initial data $\mu_1(0)$ and $\mu_2(0)$ satisfies

$$W_2(\mu_1(t), \mu_2(t)) \leq e^{-Kt} W_2(\mu_1(0), \mu_2(0)).$$

In particular, the Cauchy problem of the gradient flow of $F$ has a unique weak solution. See [1,30,31,38,39]. In view of this and Corollary 1.5, we have the following result.

**Theorem 1.8** Suppose that $V \in C^2(\mathbb{R})$ satisfying the same condition as in Theorem 1.1, and there exists a constant $K \in \mathbb{R}$ such that

$$V''(x) \geq 2K, \quad \forall x \in \mathbb{R}.$$ 

Let $\mu_i(t)$ be any two weak solutions of the McKean–Vlasov equation (4) with initial data $\mu_i(0), i = 1, 2$. Then for all $t > 0$, we have

$$W_2(\mu_1(t), \mu_2(t)) \leq e^{-Kt} W_2(\mu_1(0), \mu_2(0)).$$

In particular, the Cauchy problem of the McKean–Vlasov equation (4) has a unique weak solution.

**Remark 1.9** We would like to mention that Cépa and Lépingle [12] proved the uniqueness of weak solution to the McKean–Vlasov equation (4) with quadratic potential function $V(x) = ax^2 + bx$ with two constants, $a \geq 0$ and $b \in \mathbb{R}$, and that Fontbona [21] proved the uniqueness of weak solution to the McKean–Vlasov equation (4) with external potential $V$ such that $V'(x) = \theta x + b_1(x)$, where $\theta \in \mathbb{R}$ is a constant and $b_1 \in C^1(\mathbb{R})$ is a bounded function with bounded derivative. See also [20]. Theorem 1.8 establishes the uniqueness of weak solution to the McKean–Vlasov equation (4) with more general external potentials $V$ satisfying the condition $V'' \geq 2K$ for some constant $K \in \mathbb{R}$. To do so, the authors of [12,20,21] used the Stieltjes transformation. Our method uses the optimal transportation theory and is different from the one in previous works [12,20,21].

We also extend Otto-Villani’s HWI inequality (see [11,31,38,39]) to the Voiculescu free entropy $\Sigma_V$, the $W_2$-Wasserstein distance and the free Fisher information $I_V$.

**Theorem 1.10** Suppose that there exists a constant $K \in \mathbb{R}$ such that

$$V''(x) \geq 2K, \quad \forall x \in \mathbb{R}.$$ 

Let $\mu_i \in \mathcal{P}_2(\mathbb{R}), i = 1, 2$. Then for all $t > 0$, the HWI inequality holds

$$\Sigma_V(\mu_1) - \Sigma_V(\mu_2) \leq W_2(\mu_1, \mu_2) \sqrt{I_V(\mu_1)} - \frac{K}{2} W_2^2(\mu_1, \mu_2).$$\hspace{1cm} (21)

In particular, for any weak solution to the McKean–Vlasov equation (4), we have

$$\Sigma_V(\mu_t) - \Sigma_V(\mu_V) \leq W_2(\mu_t, \mu_V) \sqrt{I_V(\mu_t)} - \frac{K}{2} W_2^2(\mu_t, \mu_V),$$\hspace{1cm} (22)

where

$$I_V(\mu) := \int_{\mathbb{R}} |V'(x)/2 - H\mu(x)|^2 d\mu(x).$$
As a consequence of Theorems 1.1 and 1.8, we can derive the Law of Large Numbers for the empirical measures of the generalized Dyson Brownian motion. We would like to mention that, as far as we know, one cannot find established result in the literature on the Law of the Large Numbers of the empirical measure processes for generalized Dyson Brownian motion with general external potential $V$ except the well-known case of Gaussian ensemble with quadratic potential.

**Theorem 1.11** Suppose that $L_N(0)$ converges to $\mu(0)$ in $\mathcal{P}_2(\mathbb{R})$. Let $V$ be a $C^2$ function satisfying the same condition as in Theorem 1.1 and $V'' \geq K$ for some constant $K \geq 0$. Then the empirical measure $L_N(t)$ weakly converges to the unique solution $\mu_t$ of the McKean–Vlasov equation (4). Moreover, for all $p \in [1, 2)$, we have

$$W_p(\mathbb{E}(L_N(t), \mu_t)) \to 0 \quad \text{as} \quad N \to \infty,$$

where the convergence is uniformly with respect to $t \in [0, T]$ for all fixed $T > 0$. In the case $V(x) = \frac{Kx^2}{2}$ with $K > 0$, the result also holds for $p = 2$.

The notion of propagation of chaos, which was introduced by M. Kac, plays a critical role in the study of the large $N$ limit of $N$-particle systems. According to Sznitman-Tanaka’s theorem [36], for exchangeable systems, propagation of chaos is equivalent to the Law of Large Numbers for the empirical measures of the system. In view of this and Theorem 1.11, we have the following result, which is a dynamic version of a result due to Johansson (Theorem 2 in [25]).

**Theorem 1.12** Assume the conditions in Theorem 1.11 holds. Let $M_N; k(t; dx_1, \ldots, dx_k)$ be the $k$th moment measure for the random probability measure $L_N(t, \cdot)$, that is, for any Borel sets $A_1, \ldots, A_k$,

$$M_N; k(t; A_1, \ldots, A_k) := \mathbb{E}(L_N(t, A_1) \cdots L_N(t, A_k)).$$

Then we have

$$\lim_{N \to \infty} \int_{\mathbb{R}^k} \varphi(x_1, \ldots, x_k) M_N; k(t; dx_1, \ldots, dx_k) = \int_{\mathbb{R}^k} \varphi(x_1, \ldots, x_k) \mu_t(dx_1) \cdots \mu_t(dx_k)$$

for any continuous, bounded $\varphi$ on $\mathbb{R}^k$.

For any fixed $N$ and for a wide class of potentials $V$ (for example, if $\nabla^2 H_N \geq K$ for some constant $K > 0$, where $H_N$ is the Hamiltonian defined by (2)), one can use the Bakry-Emery calculus and functional inequalities such as the Poincaré inequality and the logarithmic Sobolev inequality, to show that $L_N(t)$ converge to $L_N$ as $t \to \infty$. On the other hand, the large $N$-limit of $L_N(t)$, i.e., $\mu_t(dx) = \rho_t(x) dx$, satisfies the nonlinear McKean–Vlasov equation (4). It is natural to ask the question whether $\mu_t$ converges to $\mu_V$ in the weak convergence topology or with respect to the $W_2$-Wasserstein distance for general potentials $V$. If this is true, then, with respect to the weak convergence on $\mathcal{P}(\mathbb{R})$ or the $W_2$-Wasserstein topology on $\mathcal{P}_2(\mathbb{R})$, the following diagram is commutative

$$L_N(t) \implies \mu_t \quad \text{and} \quad L_N \implies \mu_V.$$
In the literature, Chan [13] and Rogers-Shi [34] proved that this is true for \( V(x) = \frac{x^2}{2} \). See also [2,23]. In particular, this gives a dynamic proof of Wigner’s semi-circle law for the Gaussian Unitary Ensemble. The following result provides the positive answer to this problem for \( C^2 \)-convex potentials.

**Theorem 1.13** (i) Suppose that \( V \) is \( C^2 \)-convex, i.e., \( V'' \geq 0 \). Then \( \mu_t \) converges to \( \mu_V \) with respect to the Wasserstein distance in \( \mathcal{P}_p(\mathbb{R}) \) for all \( p \in [1, \infty) \), i.e.,

\[
W_p(\mu_t, \mu_V) \to 0 \quad \text{as} \quad t \to \infty.
\]

Moreover, for all \( t > s \geq 0 \), we have

\[
\begin{align*}
W_2(\mu_t, \mu_V) &\leq W_2(\mu_s, \mu_V), \\
\Sigma_V(\mu_t | \mu_V) &\leq \frac{W_2^2(\mu_s, \mu_V)}{t - s}, \\
I_V(\mu_t) &\leq I_V(\mu_0),
\end{align*}
\]

(ii) Suppose that \( V \) is \( C^2 \) and there exists a constant \( K \in \mathbb{R} \) such that

\[
V''(x) \geq 2K, \quad \forall x \in \mathbb{R}.
\]

Then for all \( t > 0 \), we have

\[
\begin{align*}
\Sigma_V(\mu_t | \mu_V) &\leq e^{-2Kt} \Sigma_V(\mu_0 | \mu_V), \\
W_2(\mu_t, \mu_V) &\leq e^{-Kt} W_2(\mu_0, \mu_V), \\
I_V(\mu_t) &\leq e^{-Kt/2} I_V(\mu_0).
\end{align*}
\]

In particular, if \( V \) is \( C^2 \)-uniform convex with \( V'' \geq 2K > 0 \), then \( \mu_t \) converges to \( \mu_V \) with the exponential rate \( K \) in the \( W_2 \)-Wasserstein topology on \( \mathcal{P}_2(\mathbb{R}) \).

(iii) Suppose that \( V \) is \( C^2 \)-convex and there exist two constants \( K > 0 \) and \( r > 0 \) such that

\[
V''(x) \geq K, \quad \forall x \in \mathbb{R} \setminus [-r, r].
\]

Then \( \mu_t \) converges to \( \mu_V \) with an exponential rate in the \( W_2 \)-Wasserstein topology on \( \mathcal{P}_2(\mathbb{R}) \). More precisely, there exist two constants \( C_1 > 0 \) and \( C_2 > 0 \) such that

\[
W_2^2(\mu_t, \mu_V) \leq e^{C_1t} \Sigma_V(\mu_0 | \mu_V), \quad t > 0.
\]

As a corollary of Theorem 1.13, for \( C^2 \)-convex potentials, we can give a dynamic proof of the well-known result due to Boutet de Monvel–Pastur–Shcherbina [8] and Johansson [25]. Their result says that, for \( V \) satisfying the growth condition (7), the empirical measure

\[
L_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}
\]

weakly converges to the equilibrium measure \( \mu_V \), where \((x_1, \ldots, x_N)\) satisfies the following probability distribution

\[
P_\beta^N(dx_1, \ldots, dx_N) = \frac{1}{Z_N^\beta} \prod_{i \neq j} |x_i - x_j|^\frac{\beta}{2} \exp \left( -\frac{\beta N}{2} \sum_{i=1}^{N} V(x_i) \right) \prod_{i=1}^{N} dx_i,
\]

where \( \beta > 0 \) is a parameter. We would like to mention that, for non-convex potentials \( V \), we do not know how to give a dynamic proof of the above result.
For $\beta = 2$ and for real analytic function $V$, we can prove that the generalized Dyson Brownian motion can be realized as the eigenvalues process of the $N \times N$ real Hermitian matrix valued diffusion process defined by
\[
dX^N_t = \frac{1}{\sqrt{N}} dB^N_t - \frac{1}{2} V'(X^N_t) dt,
\]
where $B^N_t$ is the $N \times N$ Hermitian matrix valued Brownian motion. Moreover, we can prove that $X^N_t$ converges in distribution to the free diffusion process $X_t$, which was defined by Biane and Speicher [5]. This extends a famous result, due to Voiculescu [40, 42] and Biane [3], which states that the renormalized Hermitian Brownian motion $\sqrt{N} B^N_t$ converges in distribution to the free Brownian motion $S_t$. See a forthcoming paper [27].

The rest of this paper is organized as follows. In Sect. 2, we prove Theorem 1.1. In Sect. 3, we prove Theorems 1.4 and 1.6. In Sect. 4, we prove Theorems 1.8 and 1.10. In Sect. 5, we prove Theorems 1.11 and 1.12. In Sect. 6, we prove Theorem 1.13. In Sect. 7, we discuss the case of double-well potentials.

Finally, let us mention that this paper is an improved version of our paper with the same title posted on arxiv(1407.7234v2), which was a revised version of our previous paper entitled “Generalized Dyson Brownian motion, McKean–Vlasov equation and eigenvalues of random matrices” (arXiv:1303.1240).

2 Proof of Theorem 1.1

The proof of Theorem 1.1 is adapted from classical argument coming back to McKean and exposed in [2, 12, 34].

**Proof of existence and uniqueness of GDBM** First, for fixed $R > 0$, let $\phi_R(x) = x^{-1}$ if $|x| \geq R^{-1}$, and $\phi_R(x) = R^2 x$ if $|x| < R^{-1}$. Since $\phi_R$ is uniformly Lipschitz and $V$ satisfies (i) and (ii), by Theorem 3.1.1 in [33], the following SDE for the truncated Dyson Brownian motion
\[
d\lambda^i_{N,R}(t) = \sqrt{\frac{2}{\beta N}} dW^i_t + \frac{1}{N} \sum_{j: j \neq i} \phi_R(\lambda^j_{N,R}(t) - \lambda^i_{N,R}(t)) dt - \frac{1}{2} V'(\lambda^i_{N,R}(t)) dt, \tag{24}
\]
with $\lambda^i_{N,R}(0) = \lambda^i_N(0)$ for $1 \leq i \leq N$, has a unique strong solution. Let
\[
\tau_R := \inf \left\{ t : \min_{i \neq j} | \lambda^i_{N,R}(t) - \lambda^j_{N,R}(t) | < R^{-1} \right\}.
\]
Then $\tau_R$ is monotone increasing in $R$ and $\lambda_{N,R}(t) = \lambda_{N,R'}(t)$ for all $t \leq \tau_R$ and $R < R'$.

Second, let $\lambda_N(t) = \lambda_{N,R}(t)$ on $t \in [0, \tau_R)$. To prove that $\lambda_N(t)$ is a global solution to SDE (1), we need only to prove $\lambda_N(t)$ does not explode, and $\lambda^i_N(t)$ and $\lambda^j_N(t)$ never collide for all $t > 0, i \neq j$.

To prove that $\lambda_N(t)$ does not explode, let $R_t = \frac{1}{N} \sum_{j=1}^N \lambda^j_N(t)^2$. By Itô’s formula, and by Levy’s characterization, we can introduce a new Brownian motion $B$, such that
\[
dR_t = \frac{2}{\beta N} \sqrt{\frac{R_t}{\beta}} dB_t + \left( \frac{1}{\beta N} + \frac{N - 1}{2N} - \frac{1}{2} L_N(t), x V'(x) \right) dt.
\]
Let $R'$ be the solution of
\[
dR'_t = \frac{2}{N} \sqrt{R'_t / \beta} dB_t + \left( \frac{1}{\beta N} + \frac{N - 1}{2N} + \frac{1}{2} \gamma + \gamma R'_t \right) dt,
\]
with $R'_0 = R_0$. Under the assumption (8), and using the comparison theorem of one-dimensional SDEs, cf. [24], we can derive that
\[
R_t \leq R'_t, \quad \forall t \geq 0, \text{ a.s.}
\]
Moreover, by Ikeda and Watanabe [24] (pp. 235–237), the process $R'$ never explodes. So the process $R$ (and hence $\lambda_N(t)$) does not explode in finite time.\(^1\)

To prove that $\lambda^i_N(t)$ and $\lambda^j_N(t)$ never collide for all $t > 0, i \neq j$, let us introduce the Lyapunov function $f(x_1, \ldots, x_N) = \frac{1}{N} \sum_{i=1}^{N} V(x_i) - \frac{1}{N^2} \sum_{i \neq j} \log |x_i - x_j|$. Similarly to [2,23], we can prove
\[
df(\lambda_N(t)) = dM_N(t) + \frac{1}{N^2} \left( \frac{1}{\beta} - 1 \right) \sum_{k \neq i} (\lambda^i_N(t) - \lambda^k_N(t))^2 dt - \frac{1}{2N} \sum_{i=1}^{N} |\nu'(\lambda^i_N(t))|^2 dt
\]
\[+ \frac{1}{N^2} \left( \frac{1}{\beta} \sum_{i=1}^{N} V''(\lambda^i_N(t)) + \frac{3}{2} \sum_{j \neq i} \left( \frac{V'(\lambda^i_N(t)) - V'(\lambda^j_N(t))}{\lambda^i_N(t) - \lambda^j_N(t)} \right) \right) dt,
\]
where $M_N$ is the following local martingale
\[
dM_N(t) = \frac{2}{\beta + 2N^2} \sum_{i=1}^{N} \left( V'(\lambda^i_N(t)) - \frac{1}{N} \sum_{k: k \neq i} \lambda^i_N(t) - \lambda^k_N(t) \right) dW_i^t.
\]
Fix $K > 0$ and $R > 0$ such that $\lambda^i_N(0) \in [-K, K]$ and $|\lambda^i_N(0) - \lambda^j_N(0)| \geq R$ for all $i \neq j$, $i, j = 1, \ldots, N$. Let $C_1(K) \geq 0$ be such that $\sup_{x \in [-K, K]} V''(x) \leq C_1(K)$. Let $A(t)dt = d(\lambda^i_N(t)) - dM_N(t)$, and $\xi_K = \inf \{ t \geq 0 : \lambda^i_N(t) \notin [-K, K] \text{ for some } i = 1, \ldots, N \}$, then for any fixed $T > 0$, $\sup_{t \in [0, T]} A(t) \leq C_1(K)$ and $\{ (\lambda^i_N(t \wedge \xi_K) - C_1(K) (t \wedge \xi_K), t \in [0, \tau_K \wedge T] \}$ is a supermartingale. Let $C_2(K) := \inf \{ V(x) : |x| \leq K \}$, we can prove
\[
\mathbb{P}(\tau_K \leq \xi_K \wedge T) \leq \frac{N^2(f(\lambda_N(0)) + T C_1(K) + N(N - 1) \log(2K) - C_2(K))}{\log(2K) + \log R}.
\]
Letting $R, T$ and $K$ tend to infinity, we can prove $\mathbb{P}(\tau_{\infty} < \xi) = 0$, where $\xi := \inf \{ t : \lambda^i_N(t) = \lambda^j_N(t) \ \exists \ 1 \leq i \neq j \leq N \}$. This proves that $\lambda^1_N(t), \ldots, \lambda^N_N(t)$ does not collide.

Finally, by the continuity of the trajectory of $\lambda_N(t)$, we have $\lambda_N(t) \in \Delta_N$ for all $t \geq 0$. The same argument as used in the proof of Theorem 12.1 in [23] proves the uniqueness of the weak solution to SDEs (1). The proof of Theorem 1.1 is completed.

**Proof of tightness and identification of McKean–Vlasov limit**

We follow the argument used in [34] to prove the tightness of $\{L_N(t), t \in [0, T]\}$. Let us pick functions $f_j \in C^\infty_b(\mathbb{R}, \mathbb{C}), j = 1, 2, \ldots$, which is dense in $C_b(\mathbb{R})$. Thus
\[
\langle \mu, f_j \rangle = \langle \mu', f_j \rangle, \quad \forall j \Rightarrow \mu = \mu'.
\]
\(^1\) In [34], Rogers and Shi proved the non-explosion of GDBM for $V$ satisfying $-x V'(x) \leq \gamma, \forall x \in \mathbb{R}.
We also pick a $C^\infty$ function $f_0 : \mathbb{R} \to [1, \infty)$ with the properties

$$f_0(x) = f_0(-x), \quad f_0(x) \to \infty \text{ as } x \to \infty, \ x \in \mathbb{R}^+.$$ 

Taking test functions in the Schwartz class of smooth functions whose derivatives (up to second order) are rapidly decreasing, we may assume that

$$f_j, \ f_j'', \ V'f_j' \text{ are uniformly bounded for all } j \geq 1.$$ 

By Ethier and Kurtz [19] (p.107), to prove the tightness of $\{L_N(t), \ t \in [0, T], \ N \geq 1\}$, it is sufficient to prove that for each $j$ the sequence of continuous real-valued functions

$$\{(L_N(t), f_j), \ t \in [0, T], \ N \geq 1\}$$

is relatively compact. To this end, note that, by the first part of Theorem 1.1, there is non-collision and non-explosion for the particles $\lambda_N^i(t)$ for all $t \in [0, \infty)$. By Itô’s formula, we have

$$d\langle L_N(t), f \rangle = \frac{1}{\sqrt{2}} \frac{1}{\beta N} \sum_{i=1}^{N} f'(\lambda_N^i(t))dW_i + \left(L_N(t), \left(2 - \frac{1}{\beta} - 1\right) \frac{1}{2N} f'' - \frac{1}{2} V'f'ight)dt + \frac{1}{2} \int_{\mathbb{R}^2} \frac{f''(x) - f''(y)}{x-y} L_N(t, dx)L_N(t, dy)dt. \quad (25)$$

This yields

$$\langle L_N(t), f_j \rangle = \langle L_N(0), f_j \rangle + \frac{1}{2} \int_0^t \int_{\mathbb{R}^2} \frac{f_j(x) - f_j(y)}{x-y} L_N(s, dx)L_N(s, dy)ds - \frac{1}{2} \int_0^t \langle L_N(s), V'f_j \rangle ds + \int_0^t \left(L_N(s), \left(2 - \frac{1}{\beta} - 1\right) \frac{1}{2N} f'' \right) ds + M_N^{f_j}(t) \quad (26)$$

where

$$M_N^{f_j}(t) = \frac{1}{\sqrt{2}} \frac{1}{\beta N} \sum_{i=1}^{N} f_j'\lambda_N^i(s)ds.$$ 

Note that, as $L_N(0)$ is weakly convergent, $I_1(N)$ is convergent. By the assumption that $f_j$ and $f_j''$ are uniformly bounded (hence $f_j'$ are uniformly bounded) , we can easily show that $\{M_N^{f_j}(t), t \in [0, T]\}$ and $I_4(N)$ converge to zero. Moreover, by the assumption that $V'f_j'$ and $f_j''$ are uniformly bounded, the Arzela-Ascoli theorem implies that $I_2(N)$ and $I_3(N)$ are relatively compact in $C([0, T], \mathbb{R})$. Thus the sequence $\{(L_N(t))_{t \geq 0}, \ N \geq 1\}$ is tight in $C([0, T], \mathbb{R})$. Tightness also follows for $j = 0$ if we have

$$\langle L_N(0), f_0 \rangle \to \text{ finite limit as } N \to \infty.$$ 

So let us suppose that the initial distribution $L_N(0)$ have the property $\langle L_N(0), f_0 \rangle \leq K$ for some $K$, for all $N$. For given $\mu_0$, we could always find $L_N(0)$ and $f_0$ to satisfy this and the other conditions, and this gives the tightness for $j = 0$ also.

Finally, we identify the limit process of any weakly convergent subsequence of $\{L_N(t)\}$. Assuming that $\{L_N(t), t \in [0, T]\}$ is a weakly convergent subsequence in $C([0, T], \mathcal{P}(\mathbb{R}))$. 

\[ \text{Springer} \]
Then, for all \( f \in C^2_b(\mathbb{R}) \), the Itô’s formula (26) and the above argument show that \( \langle \mu_t, f \rangle = \lim_{j \to \infty} \langle L_{N_j}(t), f \rangle \) satisfies the following equation

\[
\int_{\mathbb{R}} f(x) \mu_t(dx) = \int_{\mathbb{R}} f(x) \mu_0(dx) + \frac{1}{2} \int_0^t \int_{\mathbb{R}^2} \frac{\partial_x f(x) - \partial_y f(y)}{x - y} \mu_s(dx) \mu_s(dy) ds
\]

\[
- \frac{1}{2} \int_0^t \int_{\mathbb{R}} V'(x) f'(x) \mu_s(dx) ds.
\]

This proves that \( \mu_t \) is a weak solution in the space of probability measures \( \mathcal{P}(\mathbb{R}) \) to the McKean–Vlasov equation (3). The proof of Theorem 1.1 is completed. \( \Box \)

3 Proofs of Theorems 1.4 and 1.6

3.1 Some Preliminary Results

Recall the definition of the fractional Sobolev space: for any \( s > 0 \),

\[ H^s(\mathbb{R}) = \left\{ u \in L^2(\mathbb{R}) : (1 + |\xi|^2)^{s/2} |\hat{u}| \in L^2(d\xi) \right\}, \]

with the Sobolev norm

\[ \|u\|_{H^s(\mathbb{R})}^2 = \int_{\mathbb{R}} (1 + |\xi|^2)^s |\hat{u}(\xi)|^2 d\xi. \]

Moreover, see e.g. [15], \( u \in H^s(\mathbb{R}) \) if and only if \( u \in L^2(\mathbb{R}) \) and

\[ \int_{\mathbb{R}^2} \frac{|u(x) - u(y)|^2}{|x - y|^{1+2s}} dxdy < \infty. \]

Let \( u \in C_c^\infty(\mathbb{R}) \). The Hilbert transform of \( u \) is defined by

\[ H u(x) = \frac{1}{\pi} \lim_{\varepsilon \to 0} \int_{|x - y| \geq \varepsilon} \frac{u(y)}{x - y} dy, \quad a.s. \ x \in \mathbb{R}. \]

It is well-known that the Hilbert transform \( H \) can be extended to be a bounded operator in \( L^p(\mathbb{R}) \) for all \( p \in (1, \infty) \).

**Lemma 3.1** For any \( s > 0 \), the Hilbert transform is an isometry on \( H^s(\mathbb{R}) \). That is to say, for any \( u \in H^s(\mathbb{R}) \), we have \( Hu \in H^s(\mathbb{R}) \), and

\[ \|Hu\|_{H^s(\mathbb{R})} = \|u\|_{H^s(\mathbb{R})}. \]

**Proof** Note that, for any \( u \in L^2(\mathbb{R}) \), \( Hu \in L^2(\mathbb{R}) \) and \( \widehat{Hu}(\xi) = -i\text{sgn}(\xi)\hat{u}(\xi) \), for all \( \xi \in \mathbb{R} \). This yields, for any \( u \in H^s(\mathbb{R}) \) with \( s > 0 \),

\[
\|Hu\|_{H^s(\mathbb{R})}^2 = \int_{\mathbb{R}} (1 + |\xi|^2)^s |\widehat{Hu}(\xi)|^2 d\xi
\]

\[
= \int_{\mathbb{R}} (1 + |\xi|^2)^s |i\text{sgn}(\xi)\hat{u}(\xi)|^2 d\xi
\]

\[
= \int_{\mathbb{R}} (1 + |\xi|^2)^s |\hat{u}(\xi)|^2 d\xi
\]

\[
= \|u\|_{H^s(\mathbb{R})}^2.
\]

\( \Box \)
To simplify the notation, we will omit the multiplicative constant $\frac{1}{\pi}$ in the definition of the Hilbert transform.

**Lemma 3.2** Let $f \in L^\infty(\mathbb{R})$ with compact support $\text{supp}(f) = [-M, M]$, and $v \in L^2(\mathbb{R})$. Then

$$
\int_{-M}^{M} \int_{-M}^{M} \frac{v(x) - v(y)}{x - y} f(x)f(y)dydx = \int_{-M}^{M} Hf(x)f(x)v(x)dx.
$$

**Proof** When $v \in C^1(\mathbb{R})$, we can prove the result by dividing $[-M, M]^2$ into $\{(x, y) \in [-M, M]^2 : |x - y| \geq \varepsilon\}$ and $\{(x, y) \in [-M, M]^2 : |x - y| < \varepsilon\}$ and using the Lipschitz continuity of $v$ on $[-M, M]$. Due to limit of the paper, we omit the detail of the proof. By the Cauchy–Schwarz inequality and the $L^2$-isometry of the Hilbert transform, we have

$$
\left| \int_{\mathbb{R}} Hf(x)f(x)v(x)dx \right| \leq \left( \int_{\mathbb{R}} |Hf(x)|^2 f(x)dx \right)^{1/2} \left( \int_{\mathbb{R}} |v(x)|^2 f(x)dx \right)^{1/2} \\
\leq \|f\|_\infty \|Hf\|_2 \|v\|_2 \leq \|f\|_\infty \|f\|_2 \|v\|_2 \leq \sqrt{2} M \|f\|_\infty^2 \|v\|_2.
$$

Thus, the density argument proves the lemma for all $v \in L^2(\mathbb{R})$. \hfill \Box

### 3.2 Proof of Theorem 1.4

Note that there exists a constant $C > 0$ such that for $v \in H^{1/2}(\mathbb{R})$ it holds

$$
\int_{\mathbb{R}^2} \frac{|v(x) - v(y)|^2}{(x-y)^2} dxdy \leq C \|v\|_{H^{1/2}(\mathbb{R})}^2.
$$

See e.g. [15]. Hence, the second term in the right hand side of (15) is bounded above by $C \|\rho_0\|_\infty^2 \|v\|_{H^{1/2}(\mathbb{R})}^2$. By density argument, we need only to prove (15) for smooth and rapidly decay velocity $v$. In this case, let $\phi_s(x) = x + sv(x), s \in [0, 1], x \in \mathbb{R}$, and let $\mu_s = \rho_s dx$ be a geodesic passing through $\mu_0$ at $s = 0$ and with the smooth initial velocity $v$. By (13) and changing of variables, we have

$$
\Sigma_V(\mu_s) = -\frac{1}{2} \int_{\mathbb{R}^2} \log|x-y| \rho_s(x) \rho_s(y) dxdy + \frac{1}{2} \int_{\mathbb{R}} V(x) \rho_s(x) dx = -\frac{1}{2} \int_{\mathbb{R}^2} \log|\phi_s(x) - \phi_s(y)| \rho_0(x) \rho_0(y) dxdy + \frac{1}{2} \int_{\mathbb{R}} V(\phi_s(x)) \rho_0(x) dx.
$$

Taking derivatives with respect to $s$, and noting that $\rho_0$ has compact support (so that we can exchange the order of differential and integration), a standard argument yields

$$
\frac{d}{ds} \Sigma_V(\mu_s) = -\frac{1}{2} \int_{\mathbb{R}^2} \frac{\partial_s \phi_s(x) - \partial_s \phi_s(y)}{\phi_s(x) - \phi_s(y)} \rho_0(x) \rho_0(y) dxdy + \frac{1}{2} \int_{\mathbb{R}} V'(\phi_s(x)) \partial_s \phi_s(x) \rho_0(x) dx,
$$

and

$$
\frac{d^2}{ds^2} \Sigma_V(\mu_s) = \frac{1}{2} \int_{\mathbb{R}^2} \frac{|\partial_s \phi_s(x) - \partial_s \phi_s(y)|^2}{|\phi_s(x) - \phi_s(y)|^2} \rho_0(x) \rho_0(y) dxdy + \frac{1}{2} \int_{\mathbb{R}} V''(\phi_s(x)) (\partial_s \phi_s(x))^2 \rho_0(x) dx.
$$
Thus, at the point \( s = 0 \), using Lemma 3.2, we have

\[
\frac{d}{ds}\bigg|_{s=0} \Sigma_V(\mu_s) = -\frac{1}{2} \int_{\mathbb{R}^2} \frac{v(x) - v(y)}{x - y} \rho_0(x)\rho_0(y) \, dx \, dy + \frac{1}{2} \int_{\mathbb{R}} V'(x)v(x)\rho_0(x) \, dx
\]

\[
= \int_{\mathbb{R}} \left( \frac{1}{2} V'(x) - H\rho_0(x) \right) v(x)\rho_0(x) \, dx,
\]

(27)

and

\[
\frac{d^2}{ds^2}\bigg|_{s=0} \Sigma_V(\mu_s) = \frac{1}{2} \int_{\mathbb{R}^2} \frac{|v(x) - v(y)|^2}{|x - y|^2} \rho_0(x)\rho_0(y) \, dx \, dy
\]

\[
+ \frac{1}{2} \int_{\mathbb{R}} V''(x)|v(x)|^2 \rho_0(x) \, dx.
\]

(28)

Combining (34) with the definition formula (14) of the Hessian, we obtain (15).

\( \square \)

**Proof of Corollary 1.5** By the fact that the tangent space of \( \mathcal{P}_2(\mathbb{R}) \) is the \( L^2 \)-closure of the tangent space of the smooth Wasse-derstein space \( \mathcal{P}_\infty(\mathbb{R}) \), and using the density argument, we need only to prove (17) when \( \mu_0, \mu_1 \in \mathcal{P}_\infty(\mathbb{R}) \) and when \( \mu_1 \) and \( \mu_2 \) is connected by a unique smooth geodesic with smooth and rapidly decay initial velocity \( v \). In this case, we can use a standard argument based on the Taylor formula and the Hessian formula in Theorem 1.4 to give a proof of (17).

\( \square \)

### 3.3 Proof of Theorem 1.6

By (33), the gradient of \( \Sigma_V \) is given by

\[
\nabla \mathcal{P}_2(\mathbb{R}) \Sigma_V(\rho) = -\frac{\partial}{\partial x} \left( \rho(x) \left( \frac{1}{2} V'(x) - H\rho(x) \right) \right).
\]

(29)

Indeed, by standard calculation of variation, we can prove

\[
\frac{\partial}{\partial x} \frac{\delta \Sigma_V}{\delta \rho}(\rho) = \frac{V'}{2} - H\rho.
\]

Combining this with (10), we derive again (29). Thus the gradient flow of \( \Sigma_V \) on \( \mathcal{P}_2(\mathbb{R}) \) is given by

\[
\frac{d\mu_t}{dt} = -\nabla \mathcal{P}_2(\mathbb{R}) \Sigma_V(\mu_t).
\]

That is to say, the density function \( \rho_t = \frac{d\mu_t}{dt} \) satisfies the McKean–Vlasov equation (4), i.e.,

\[
\partial_t \rho = \partial_x (\rho v_t).
\]

where \( v_t = \frac{V'}{2} - H\rho_t \). This proves the first statement of Theorem 1.6.

Taking derivative of \( \Sigma_V \) along the gradient flow and by definition of the norm on \( \mathcal{P}_2(\mathbb{R}) \), we obtain (18). Indeed,

\[
\frac{d}{dt} \Sigma_V(\mu_t) = \left\langle \nabla \mathcal{P}_2(\mathbb{R}) \Sigma_V(\mu_t), \frac{d\mu_t}{dt} \right\rangle
\]

\[
= -\left\| \nabla \mathcal{P}_2(\mathbb{R}) \Sigma_V(\mu_t) \right\|^2 = -\int_{\mathbb{R}} |v_t(x)|^2 \rho_t(x) \, dx.
\]

\( \square \) Springer
When \( V \in C^2 \), by [5], see also Sect. 4.1, \( \rho_t \) is compactly supported, and \( \rho_t \in L^\infty(\mathbb{R}) \cap H^{1/2}(\mathbb{R}). \) Using the fact that the Hilbert transform \( H \) is an isometry on \( H^{1/2}(\mathbb{R}) \) (see Lemma 3.1), we see that \( v_t = \frac{V'}{2} - H\rho_t \in H^{1/2}(\mathbb{R}) \) when restricting on the compact support of \( \rho_t \). Hence \( v_t \) satisfies the condition required in Theorem 1.4. Taking differentiation on the both sides of (30) with respect to \( t \), we have

\[
\frac{d^2}{dt^2} \Sigma_V(\mu_t | \mu_V) = 2Hess_{\mathcal{P}_2(\mathbb{R})} \Sigma_V(\mu_t)(v_t, v_t). \quad (31)
\]

By (31) and substituting \( v_t = \frac{V'}{2} - H\rho_t \) into the Hessian formula (15), we prove (20). \( \square \)

### 4 Proofs of Theorems 1.8 and 1.10

**Proof of Theorem 1.8** Under the assumption \( V'' \geq 2K \), \( \Sigma_V \) is a proper, lower-semicontinuous and geodesically \( K \)-convex functional on \( \mathcal{P}_2(\mathbb{R}) \). In view of this, we can directly apply Theorem 11.1.4 in Ambrosio–Gigli–Savaré [1] to conclude Theorem 1.8. To the reader who is not familiar with the theory of gradient flow on metric spaces as required in Ambrosio–Gigli–Savare’s book [1], we think that it would be nice to give a direct proof of Theorem 1.8 for weak solutions of the gradient flow of \( \Sigma_V \), which in some sense is similar to the way that one can do via the Otto Calculus for smooth solutions of the gradient flow for any geodesically \( K \)-convex differentiable functional on \( \mathcal{P}_2(\mathbb{R}^d) \) (see [10,11,30,31,38,39]). Indeed, a heuristic proof of Theorem 1.8 via Otto and Carrillo-McCann-Villani’s approach was given in the previous versions of this paper (see arxiv 1303.1240 and 1407.7234).

Recall that, see Sect. 11.1.2 in [1], for any geodesically \( K \)-convex functional \( \phi \) on \( \mathcal{P}_2(\mathbb{R}) \), a gradient flow of \( \phi \) is an absolutely continuous curve \( \mu(t) : (0, \infty) \rightarrow \mathcal{P}_2(\mathbb{R}) \) with a velocity vector field \( v_t \) such that \( \|v_t\|_{L^2(\mu(t))}^2 \in L^1_{loc}(0, \infty) \),

\[
\partial_t \mu(t) + \nabla \cdot (\mu(t)v_t) = 0
\]

in the weak sense, and

\[
\phi(\sigma) - \phi(\mu(t)) \geq \int_{\mathbb{R}^2} \langle v_t, \text{id} - T^\sigma_{\mu(t)} \rangle d\mu(t) + \frac{K}{2} W_2^2(\sigma, \mu(t)), \quad (32)
\]

where \( T^\sigma_{\mu(t)} \) denotes the unique optimal transport map from \( \mu(t) \) to \( \sigma \).

Applying (32) to \( \phi := \Sigma_V, \mu(t) := \mu_1(t), \) and \( \sigma := \mu_2(t) \), we have

\[
\Sigma_V(\mu_2(t)) - \Sigma_V(\mu_1(t)) \geq \int_{\mathbb{R}^2} \langle v_1(t), \text{id} - T^{\mu_2(t)}_{\mu_1(t)} \rangle d\mu_1(t) + \frac{K}{2} W_2^2(\mu_2(t), \mu_1(t)), \quad (33)
\]

where \( T^{\mu_2(t)}_{\mu_1(t)} \) denotes the unique optimal transport map from \( \mu_1(t) \) to \( \mu_2(t) \).

Exchanging the role of \( \mu_1(t) \) and \( \mu_2(t) \), we have

\[
\Sigma_V(\mu_1(t)) - \Sigma_V(\mu_2(t)) \geq \int_{\mathbb{R}^2} \langle v_2(t), \text{id} - T^{\mu_1(t)}_{\mu_2(t)} \rangle d\mu_2(t) + \frac{K}{2} W_2^2(\mu_1(t), \mu_2(t)). \quad (34)
\]

Summing (33) and (34), we obtain

\[
KW_2^2(\mu_1(t), \mu_2(t)) + \int_{\mathbb{R}^2} \langle v_1(t), \text{id} - T^{\mu_2(t)}_{\mu_1(t)} \rangle d\mu_1(t)
\]

\[
+ \int_{\mathbb{R}^2} \langle v_2(t), \text{id} - T^{\mu_1(t)}_{\mu_2(t)} \rangle d\mu_2(t) \leq 0. \quad (35)
\]
On the other hand, by the differential formula of $W^2$ in Lemma 8.4.7 in [1], we can derive

$$\frac{1}{2} \frac{d}{dt} W^2_2(\mu_1(t), \mu_2(t)) = \int_{\mathbb{R}^2} (v_1(t), x_1) - T^{\mu_2(t)}_{\mu_1(t)} d\mu_1(t)$$

$$+ \int_{\mathbb{R}^2} (v_2(t), x_2) - T^{\mu_1(t)}_{\mu_2(t)} d\mu_2(t).$$

(36)

Combining (35) with (36), we have

$$\frac{d}{dt} W^2_2(\mu_1(t), \mu_2(t)) \leq -2KW^2_2(\mu_1(t), \mu_2(t)).$$

The Gronwall inequality implies

$$W_2(\mu_1(t), \mu_2(t)) \leq e^{-Kt} W_2(\mu_1(0), \mu_2(0)).$$

In particular, the McKean–Vlasov equation (3) has a unique weak solution.

\(\square\)

**Proof of Theorem 1.10** By [1], the gradient flow $\mu_t$ of $\Sigma_V$ can be also characterized by the system of “Evolution Variational Inequality” in the sense that

$$\frac{1}{2} \frac{d}{dt} W^2_2(\mu_t, \sigma) + \frac{K}{2} W^2_2(\mu_t, \sigma) \leq \Sigma_V(\sigma) - \Sigma_V(\mu_t)$$

for almost every $t > 0$ and for all $\sigma \in \text{Dom}(\Sigma_V)$. Taking $\sigma = \mu_V$, we have

$$\frac{1}{2} \frac{d}{dt} W^2_2(\mu_t, \mu_V) + \frac{K}{2} W^2_2(\mu_t, \mu_V) \leq \Sigma_V(\mu_V) - \Sigma_V(\mu_t).$$

By Lemma 8.4.7 in [1] and the Cauchy–Schwarz inequality,

$$\frac{1}{2} \frac{d}{dt} W^2_2(\mu_t, \mu_V) = \int_{\mathbb{R}^2} (v_t(x_1), x_1 - x_2) d\gamma_t(x_1, x_2)$$

$$\geq -\left( \int_{\mathbb{R}^2} |v_t|^2(x_1) d\gamma_t(x_1, x_2) \right)^{1/2} \left( \int_{\mathbb{R}^2} |x_1 - x_2|^2 d\gamma_t(x_1, x_2) \right)^{1/2},$$

where $\gamma_t \in \mathcal{P}(\mathbb{R}^2)$ is the unique optimal transport plane between $\mu_t$ and $\mu_V$, which gives

$$W^2_2(\mu_t, \mu_V) = \int_{\mathbb{R}^2} |x_1 - x_2|^2 d\gamma_t.$$
5 Proofs of Theorems 1.11 and 1.12

Proof of Theorem 1.11 By Theorem 1.1, the family \( \{L_N(t), t \in [0, T]\} \) is tight with respect to the weak convergence topology on \( \mathcal{P}(\mathbb{R}) \), and the limit of any weakly convergent subsequence of \( \{L_N(t), t \in [0, T]\} \) is a weak solution of (4). By the uniqueness of weak solution to the McKean–Vlasov equation (4), we conclude that \( L_N(t) \) weakly converges to \( \mu_t \), and hence \( \mathbb{E}[L_N(t)] \) weakly converges to \( \mu_t \) as \( N \to \infty \).

Taking \( f(x) = x^2 \) in (3) and (25) respectively, we can derive that

\[
\frac{d}{dt} \int_{\mathbb{R}} x^2 \mu_t(dx) = 1 - \int_{\mathbb{R}} xV'(x)\mu_t(dx),
\]

and

\[
d(L_N(t), x^2) = \frac{2}{N} \sqrt{\frac{2}{\beta N} \sum_{i=1}^N \lambda_i^2(t)dW_i^i} + \left(L_N(t), \left(\frac{2}{\beta} - 1\right) \frac{1}{N} - xV'\right) dt + 1.
\]

Taking expectation, we have

\[
\frac{d}{dt} \int_{\mathbb{R}} x^2 \mathbb{E}[L_N(t, dx)] = 1 + \left(\frac{2}{\beta} - 1\right) \frac{1}{N} - \int_{\mathbb{R}} xV'(x)\mathbb{E}[L_N(t, dx)]
\]

\[
\leq 1 + \left(\frac{2}{\beta} - 1\right) \frac{1}{N} + \gamma \int_{\mathbb{R}} (1 + x^2)\mathbb{E}[L_N(t, dx)]
\]

\[
\leq 1 + \frac{1}{\beta} + \gamma + \gamma \int_{\mathbb{R}} x^2 \mathbb{E}[L_N(t, dx)].
\]

Note that \( m_0 := \sup_N \int_{\mathbb{R}} x^2 \mathbb{E}[L_N(0, dx)] < \infty \). The Gronwall inequality implies

\[
\sup_{t \in [0, T]} \sup_N \int_{\mathbb{R}} x^2 d\mathbb{E}[L_N](t, dx) \leq C(\gamma, \beta, m_0, T)e^{\gamma T} < \infty.
\]

By Hölder inequality, for all \( p \in [1, 2) \),

\[
\int_{|x| \geq A} x^p d\mathbb{E}[L_N(t)](x) \leq \left(\int_{\mathbb{R}} x^2 d\mathbb{E}[L_N(t)](x)\right)^{p/2} (\mathbb{E}[L_N(t)](|X| \geq A))^{(2-p)/2}.
\]

By the tightness of \( \mathbb{E}[L_N(t)] \), we have

\[
\lim_{A \to \infty} \sup_{t \in [0, T]} \sup_N \int_{|x| \geq A} x^p d\mathbb{E}[L_N(t)](x) = 0.
\]

By the characterization of the \( W_p \)-convergence on \( \mathcal{P}_p(\mathbb{R}) \), see [1,38,39], for all \( p \in [1, 2) \), we have

\[
\lim_{N \to \infty} \sup_{0 \leq t \leq T} W_p(\mathbb{E}[L_N(t)], \mu_t) = 0.
\]

When \( V(x) = \frac{Kx^2}{2} \), we have

\[
\frac{d}{dt} \int_{\mathbb{R}} x^2 \mu_t(dx) = 1 - K \int_{\mathbb{R}} x^2 \mu_t(dx),
\]
\[
d(L_N(t), x^2) = \frac{2}{N} \sqrt{\frac{2}{\beta N} \sum_{i=1}^{N} \lambda_i^N(t) dW_i^t} - K \langle L_N(t), x^2 \rangle dt + \left(\frac{2}{\beta} - 1\right) \frac{1}{N} + 1,
\]
and
\[
\frac{d}{dt} \int_{\mathbb{R}} x^2 \mathbb{E}[L_N(t, dx)] = 1 + \left(\frac{2}{\beta} - 1\right) \frac{1}{N} - K \int_{\mathbb{R}} x^2 \mathbb{E}[L_N(t, dx)].
\]
Hence
\[
\int_{\mathbb{R}} x^2 \mathbb{E}[L_N(t, dx)] - \int_{\mathbb{R}} x^2 \mu_t(dx) = e^{-Kt} \left[ \int_{\mathbb{R}} x^2 \mathbb{E}[L_N(0, dx)] - \int_{\mathbb{R}} x^2 \mu_0(dx) \right] + \frac{1}{N} \left(\frac{2}{\beta} - 1\right) \frac{1 - e^{-Kt}}{K}.
\]
The proof of Theorem 1.11 is completed.

**Proof of Theorem 1.12** By the conditions in Theorem 1.11 and the Theorem of Sznitman and Tanaka’s [36], we know that, \(M_N(0)\) is \(\mu_0\)-chaotic. Since \(L_N(t)\) weakly converges to the deterministic measure \(\mu_t\) for every \(t \in [0, T]\), and the systems (GDBM)\(\nu\) are exchangeable systems, then we have the propagation of chaos by Sznitman and Tanaka’s Theorem [36].

### 6 Proof of Theorem 1.13

**Proof of Theorem 1.13 (i)** By Corollary 3.2 in Biane [4], for any \(C^2\)-convex \(V\), \(\Sigma_V\) has a unique minimizer \(\mu_V\). As was pointed out by a referee, this classical problem was essentially discussed at length by Staff–Totik [36]. By [7] (see also Theorem 1.5), as \(V\) is \(C^2\)-convex, \(\Sigma_V\) is geodesically convex on \(\mathcal{P}_2(\mathbb{R})\). Moreover, by the fact that \(\Sigma_V\) is lower semi-continuous and with respect to the weak convergence topology, see e.g. [2,23], we see that it is also lower semi-continuous with respect to the Wasserstein metric \(W_p\) on \(\mathcal{P}_p(\mathbb{R})\) for any \(p \in [1, \infty)\).

Under the condition \(V'' \geq 0\), the Wasserstein distance \(W_2(\mu_t, \mu_V)\) is nonincreasing in \(t \geq 0\). Let \(\gamma_t\) be the unique optimal transport plan between \(\mu_t\) and \(\mu_V\). Then
\[
W_2^2(\mu_t, \mu_V) = \int_{\mathbb{R}^2} |x - y|^2 d\gamma_t(x, y).
\] (37)

By the inequality \(|x|^2 \leq 2[|x - y|^2 + |y|^2]\), we have
\[
\int_{\mathbb{R}^2} |x|^2 d\gamma_t(x, y) \leq 2 \int_{\mathbb{R}^2} [|x - y|^2 + |y|^2] d\gamma_t(x, y) = 2W_2^2(\mu_t, \mu_V) + 2 \int_{\mathbb{R}^2} |y|^2 d\gamma_t(x, y) \leq 2W_2^2(\mu_0, \mu_V) + 2 \int_{\mathbb{R}^2} |y|^2 d\gamma_t(x, y).
\]
Note that
\[
\int_{\mathbb{R}^2} |x|^2 d\gamma_t(x, y) = \int_{\mathbb{R}} |x|^2 d\mu_t(x), \quad \int_{\mathbb{R}^2} |y|^2 d\gamma_t(x, y) = \int_{\mathbb{R}} |y|^2 d\mu_V(y).
\]
Therefore
\[
A := \sup_{t > 0} \int_{\mathbb{R}} |x|^2 d\mu_t(x) \leq 2 W_2^2(\mu_0, \mu_V) + 2 \int_{\mathbb{R}} |y|^2 d\mu_V(y).
\]

By the Chebyshev inequality, for any \( R > 0 \), it holds
\[
\mu_t(|x| \geq R) \leq \frac{\mu_t(|x|^2)}{R^2} \leq \frac{A}{R^2}.
\]

Indeed, using the fact that \( \text{supp} (\mu_t) \subset [-M, M] \), where \( M > 0 \) is a constant independent of \( t \), we have \( A \leq M^2 \). Moreover, for all \( p \geq 1 \) and \( 0 < R \leq M \), we have
\[
\int_{|x| \geq R} |x|^p d\mu_t(x) = \int_{\mathbb{R} \setminus [|x| \leq M]} |x|^p d\mu_t(x) \leq \frac{AM^p}{R^2},
\]
and for all \( R > M \), \( \int_{|x| \geq R} |x|^p d\mu_t(x) = 0 \).

Thus \( \{\mu_t, t \geq 0\} \) is tight and has uniformly integrable \( p \)-th moments in the sense of [1]. By the compactness criterion in \( \mathcal{P}(\mathbb{R}) \) with the \( W_p \)-metric, see e.g. [1, 38, 39], we see that \( \{\mu_t, t \geq 0\} \) is relatively compact in \( (\mathcal{P}(\mathbb{R}), W_p) \) for all \( p \in [1, \infty) \).

By the inequality (23) which will be proved in the part of Proof of Theorem 1.13 (ii) below, we see that \( \Sigma_V(\mu_t) \to \Sigma_V(\mu_V) \) as \( t \to \infty \). Fix \( p \in [1, 2] \). Let \( \{\mu_{t_n}, n \in \mathbb{N}\} \) be any convergent subsequence of \( \{\mu_t, t \geq 0\} \) in \( (\mathcal{P}(\mathbb{R}), W_p) \), and denote \( \mu_* \) its limit in \( (\mathcal{P}(\mathbb{R}), W_p) \). By the lower semicontinuity of \( \Sigma_V \) on \( (\mathcal{P}(\mathbb{R}), W_p) \), it holds
\[
0 \leq \Sigma_V(\mu_*) - \Sigma_V(\mu_V) \leq \lim_{n \to \infty} \Sigma_V(\mu_{t_n}) - \Sigma_V(\mu_V) \leq \lim_{n \to \infty} \frac{W_2^2(\mu_0, \mu_V)}{t_n} = 0.
\]
Thus, \( \mu_* \) achieves the minimum of \( \Sigma_V \). As \( \Sigma_V \) has the unique minimizer \( \mu_V \), we conclude that every convergent subsequence \( \{\mu_{t_n}, n \in \mathbb{N}\} \) in \( (\mathcal{P}(\mathbb{R}), W_p) \) must converge to the unique limit \( \mu_* = \mu_V \) for all \( p \in [1, 2] \). This proves that
\[
W_p(\mu_t, \mu_V) \to 0.
\]

For the proof of the inequalities in (i), see the proof of Theorem 1.13(ii). When \( p > 2 \), as \( \{\mu_t\} \) is relatively compact in \( (\mathcal{P}(\mathbb{R}), W_p) \) and \( W_2(\mu_t, \mu_V) \to 0 \) as \( t \to \infty \), we can prove that all accumulation points of \( \{\mu_t\} \) in \( (\mathcal{P}(\mathbb{R}), W_p) \) must be \( \mu_V \). This shows that \( W_p(\mu_t, \mu_V) \to 0 \) as \( t \to \infty \) for all \( p \in [1, \infty) \) if \( \mu \) is a convex \( C^2 \)-potential.

**Proof of Theorem 1.13 (ii)** Taking \( \mu_1(t) = \mu_t \) and \( \mu_2(t) = \mu_V \) in Theorem 1.8, we have
\[
W_2(\mu_t, \mu_V) \leq e^{-Kt} W_2(\mu(0), \mu_V).
\]

This proves the \( W_2 \)-convergence with exponential rate of \( \mu_t \) to the equilibrium \( \mu_V \) if \( K > 0 \).

By the HWI inequality that we proved in Theorem 1.10, we have
\[
\Sigma_V(\mu_t|\mu_V) \leq \sqrt{-\frac{d}{dt} \Sigma_V(\mu_t|\mu_V) W_2(\mu_t, \mu_V) - \frac{K}{2} W_2^2(\mu_t, \mu_V).}
\]

(38)

When \( K > 0 \), we have max \( x \geq 0 \left\{ ax - \frac{x^2}{2K} \right\} = \frac{a^2}{2K} \). Hence
\[
\Sigma_V(\mu_t|\mu_V) \leq \frac{1}{2K} \left( -\frac{d}{dt} \Sigma_V(\mu_t|\mu_V) \right).
\]

\[\text{Springer}\]
which implies

$$
\Sigma_V(\mu_t|x) \leq e^{-2Kt} \Sigma_V(\mu_0|x).
$$

When $K = 0$, by Theorem 1.6, we have

$$
\frac{d}{dt} \Sigma_V(\mu_t|x) = -\int_{\mathbb{R}} |\xi|^2 d\mu_t \leq 0.
$$

On the other hand, taking $K = 0$ in (38), we have

$$
\Sigma_V(\mu_t|x) \leq \sqrt{-\frac{d}{dt} \Sigma_V(\mu_t|x) W_2^2(\mu_t, \mu_0)},
$$

which yields

$$
\frac{d}{dt} \Sigma_V(\mu_t|x) = -\frac{d}{dt} \Sigma_V(\mu_t|x) \geq \frac{1}{W_2^2(\mu_t, \mu_0)},
$$

and therefore for any $0 \leq s < t$,

$$
\Sigma_V(\mu_t|x) \leq \Sigma_V(\mu_s|x) \left[ 1 + \Sigma_V(\mu_s|x) \int_s^t dr \frac{d}{dt} \Sigma_V(\mu_r|x) W_2^2(\mu_s, \mu_0) \right]^{-1}.
$$

Using the fact that $\Sigma_V(\mu_t|x)$ is non-negative and $W_2^2(\mu_t, \mu_0)$ is nonincreasing, we have

$$
\Sigma_V(\mu_t|x) \leq \Sigma_V(\mu_s|x) \left[ 1 + \frac{(t-s) \Sigma_V(\mu_s|x)}{W_2^2(\mu_s, \mu_0)} \right]^{-1}.
$$

Using \( \frac{1}{1+x^2} \leq \frac{1}{x} \) for $x > 0$, we have

$$
\Sigma_V(\mu_t|x) \leq \frac{W_2^2(\mu_s, \mu_0)}{t-s}, \quad \forall t > s \geq 0.
$$

On the other hand, under the condition $V'' \geq K$, Theorem 1.6 yields

$$
\frac{d}{dt} \Sigma_V(\mu_t|x) = -\int_{\mathbb{R}} |v_t(x)|^2 \rho_t(x) dx = -\|v_t\|^2,
$$

$$
\frac{d^2}{dt^2} \Sigma_V(\mu_t|x) \geq K \int_{\mathbb{R}} |v_t(x)|^2 \rho_t(x) dx = K \|v_t\|^2.
$$

Therefore

$$
\frac{d}{dt} \|v_t\|^2 = -\frac{d^2}{dt^2} \Sigma_V(\mu_t|x) \leq -K \|v_t\|^2.
$$

The Gronwall inequality yields

$$
\|v_t\| \leq e^{-Kt/2} \|v_0\|.
$$

Equivalently, we have

$$
I_V(\mu_t) \leq e^{-Kt/2} I_V(\mu_0).
$$

The proof of Theorem 1.13 (i) and (ii) is completed. \( \square \)
To prove Theorem 1.13 (iii), we need the following free logarithmic Sobolev inequality and free Talagrand transportation cost inequality due to Ledoux and Popescu [26].

**Theorem 6.1** [26] Suppose that $V$ is $C^2$-convex and there exists a constant $r > 0$ such that

$$V''(x) \geq K > 0, \quad \text{for } |x| \geq r.$$ 

Then there exists a constant $c = C(K, r) > 0$ such that the free Log-Sobolev inequality holds: for all probability measure $\mu$ with $I_V(\mu) < \infty$,

$$\Sigma_V(\mu|\mu_V) \leq \frac{2}{c} I_V(\mu).$$

Moreover, the free Talagrand transportation inequality holds: there exists a constant $C = C(K, r) > 0$ such that

$$CW^2_2(\mu, \mu_V) \leq \Sigma_V(\mu|\mu_V).$$

**Proof of Theorem 1.13 (iii)** By Biane and Speicher [5] (see also Theorem 1.6), we have

$$\frac{d}{dt} \Sigma_V(\mu_t|\mu_V) = -\frac{1}{2} I_V(\mu_t).$$

By Theorem 6.1, there exists a constant $C_1 > 0$ such that the free LSI holds

$$\Sigma_V(\mu|\mu_V) \leq \frac{2}{C_1} I_V(\mu),$$

which yields

$$\frac{d}{dt} \Sigma_V(\mu_t|\mu_V) \leq -\frac{C_1}{4} \Sigma_V(\mu_t|\mu_V).$$

By the Gronwall inequality, we have

$$\Sigma_V(\mu_t|\mu_V) \leq e^{-\frac{C_1}{4} t} \Sigma_V(\mu_0|\mu_V).$$

By Theorem 6.1 again, there exists a constant $C_2 > 0$ such that the free Talagrand transportation inequality holds

$$W^2_2(\mu_t, \mu_V) \leq \frac{1}{C_2} \Sigma_V(\mu_t|\mu_V).$$

Therefore

$$W^2_2(\mu_t, \mu_V) \leq e^{-\frac{C_1}{4} t/2} \Sigma_V(\mu_0|\mu_V).$$

This finishes the proof of Theorem 1.13 (iii). $\square$

# 7 Double-Well Potentials and Some Conjectures

In this section we discuss again the problem of the longtime convergence of the McKean–Vlasov equation towards the equilibrium measure. More precisely, we want to study the question under which condition on the external potential $V$ the following double limits are exchangeable. That is, can we prove

$$\lim_{N \to \infty} \lim_{t \to \infty} L_N(t) = \lim_{t \to \infty} \lim_{N \to \infty} L_N(t)\ ?$$
Equivalently, can we prove that the following diagram is commutative?, i.e.,

\[
\begin{array}{ccc}
L_N(t) & \implies & \mu_t \\
\downarrow & & \downarrow \\
L_N & \implies & \mu_V?
\end{array}
\]

By [13,34], see also [2,23], this is the case when \( V(x) = \frac{x^2}{2} \).

Theorem 1.13 ensures the longtime convergence of the weak solution of the McKean–Vlasov equation to the equilibrium measure \( \mu_V \) for \( C^2 \)-convex potentials \( V \). In particular, Theorem 1.13 applies to \( V(x) = a|x|^p \) with \( a > 0 \) and \( p \geq 2 \). When \( V(x) = \frac{x^2}{2} \) and \( \beta = 1, 2, 4 \), this corresponds to the cases of GUE, GOE and GSE. Moreover, Theorem 1.13 also applies to the Kontsevich-Penner model on the Hermitian random matrices ensemble with external potential (cf. [14])

\[
V(x) = \frac{ax^4}{12} - \frac{bx^2}{2} - c \log |x|.
\]

Note that, for all \( x \neq 0 \),

\[
V''(x) = ax^2 + \frac{c}{x^2} - b \geq 2\sqrt{ac} - b \geq 0
\]

provided that \( a > 0, c > 0 \) and \( 4ac \geq b^2 \).

Can we establish the longtime convergence of the McKean–Vlasov equation in the non-convex case of external potential? In [4,5], Biane and Speicher gave a non-convex potential \( \mu \) which is not necessary the global minimizer \( \mu \). Equivalently, can we prove that the following diagram is commutative?, i.e.,

\[
\begin{array}{ccc}
1 & \to & 2 \\
\downarrow & & \downarrow \\
\text{LNG} & \to & \text{Vlasov equation}
\end{array}
\]

\( \Rightarrow \) implies that

\[
\begin{array}{c}
\text{LNG} \to \text{Vlasov equation} \\
\text{LNG} \to \text{Vlasov equation}
\end{array}
\]

\( \text{LNG} \to \text{Vlasov equation} \)

When \( c \in [0, \infty) \), \( V \) is \( C^2 \) convex and \( V''(x) \geq 3 \) for \( |x| \geq 1 \). In this case, Theorem 1.13 (ii) implies that \( W_2(\mu_t, \mu_V) \to 0 \) with an exponential convergence rate.

When \( c \in (-\infty, -2) \), \( \mu_V \) has two supports \([-b, -a] \) and \([a, b]\) which are disjoint. By Sect. 7.1 in Biane-Speicher [5], it is known that \( \mu_t \) does not converge to \( \mu_V \). See also Biane [4]. This also indicates that one cannot simultaneously prove a free version of the Holley-Stroock logarithmic Sobolev inequality and a free version of the Talagrand \( T_2 \)-transportation cost inequality under bounded perturbations of potentials \( V \) in \( \mu_V \) and \( \Sigma_V \), or \( p_N(dx) = Z_N^{-1} \prod_{i<j} |x_i - x_j|^2 \prod_{i=1}^N e^{-NV(x_i)}dx \). Otherwise, if both the free LSI and the free Talagrand transport inequality hold when a convex potential \( V \) is replaced by a non-convex potential

\( \square \) Springer
On the Law of Large Numbers for the... 1303

V + U, where U is a bounded potential, then by analogue of the proof of Theorem 1.13 (iii), we may prove that the gradient flow of \( \Sigma_{V+U} \) converges to \( \mu_{V+U} \) with respect the \( W_2 \)-Wasserstein distance and hence in the weak convergence topology on \( \mathcal{P}(\mathbb{R}) \). Indeed, if a convex potential V is replaced by a non-convex potential \( V + U \), where U is a bounded potential, the support of the equilibrium of \( \Sigma_{V+U} \) may have several disjoint intervals, and the solution of the equilibrium problem is given by a different type of integral equation that has sides conditions involving the supported intervals.

In the case \( c \in [-2, 0) \), as the global minimizer \( \mu_V \) of \( \Sigma_V \) has a unique support, and all stationary points of \( \Sigma_V \) must satisfy the Euler-Lagrange equation \( H \mu = \frac{1}{4} V' \), one can see that the Voiculescu free entropy \( \Sigma_V \) has a unique stationary point which is \( \mu_V \). As \( \mu_t \) is the gradient flow of \( \Sigma_V \) on \( \mathcal{P}_2(\mathbb{R}) \), and since \( \frac{d}{dt} \Sigma_V(\mu_t) = -2 \int_{\mathbb{R}} [V'(x) - 2H \rho_t(x)]^2 \rho_t(x) dx \), we see that \( \Sigma_V(\mu_t) \) is strictly decreasing in time \( t \) unless \( \mu_t \) achieves the minimizer \( \mu_V \). This yields that the limit of \( \Sigma_V(\mu_t) \) exists as \( t \to \infty \). If one can prove that \( \{\mu_t\} \) is tight, and \( \lim_{t \to \infty} \Sigma_V(\mu_t) = \Sigma_V(\mu_V) \), then one can derive that \( \mu_t \) weakly converges to \( \mu_V \). In the first version of this paper, we raised the following conjectures.

**Conjecture 7.1** Consider the double-well potential \( V(x) = \frac{1}{4} x^4 + \frac{c}{2} x^2 \) with \( c \in [-2, 0) \). Then \( \mu_t \) converges to \( \mu_V \) with respect the \( W_2 \)-Wasserstein distance and hence in the weak convergence topology on \( \mathcal{P}(\mathbb{R}) \).

**Conjecture 7.2** Suppose that \( V \) is a \( C^2 \) potential function with \( V''(x) \geq K \) for all \( x \in \mathbb{R} \), where \( K \in \mathbb{R} \) is a constant. Suppose further that \( \Sigma_V \) has a unique stationary point which has a unique connected compact support. Then \( \mu_t \) converges to \( \mu_V \) with respect the \( W_2 \)-Wasserstein distance and in the weak convergence topology on \( \mathcal{P}(\mathbb{R}) \).

In [16], Donarti-Martin et al proved the following result which gives the affirmative answer to Conjecture 7.1.

**Theorem 7.3** [16] Let \( V(x) = \frac{1}{4} x^4 + \frac{c}{2} x^2 \) with \( c \in [-2, 0) \). Then \( \mu_t \) converges to \( \mu_V \) with respect the \( W_p \)-Wasserstein distance for all \( p \geq 1 \).

We would like to point out that, as we have seen in the proof of Theorem 1.13 (i), under the condition \( V \in C^2 \) with \( V'' \geq K \) for some \( K \in \mathbb{R} \), the family \( \{\mu_t\} \) is relatively compact in \( (\mathcal{P}_p(\mathbb{R}), W_p) \) for all \( p \in [1, \infty) \). Moreover, using the same argument as in the end of [43], we can derive the relative compactness of \( \{\rho_t = \frac{d\mu_t}{dx}, t \geq 1\} \) in \( L^p \) for all \( p \in [1, \infty) \) based on the fact that \( \{\rho_t, t \geq 1\} \) is uniformly bounded in \( H^{1/2}(\mathbb{R}) \). This fact plays an important role in the proof of Theorem 7.3 (i.e., Theorem 1.1 in [16]). Taking any convergent subsequence of \( \mu_t \) in \( (\mathcal{P}_p(\mathbb{R}), W_p) \), its limit must be a stationary point of the Voiculescu free entropy \( \Sigma_V \), which satisfies the Euler-Lagrange equation \( H \mu = \frac{1}{4} V' \) (see also [16]). In view of this, if \( \Sigma_V \) has a unique stationary point which has a unique connected compact support, we can prove that \( W_p(\mu_t, \mu_V) \to 0 \) as \( t \to \infty \) for all \( p \in [1, \infty) \) and hence \( \mu_t \) converges to \( \mu_V \) in the weak convergence topology on \( \mathcal{P}(\mathbb{R}) \). This proves Conjecture 7.2.

Finally, let us mention the following conjecture due to Biane and Speicher [5].

**Conjecture 7.4** [5] Consider the double-well potential given by \( V(x) = \frac{1}{2} x^2 + \frac{g}{4} x^4 \), where \( g \) is a negative constant but very close to zero. Then \( \mu_t \) weakly converges to \( \mu_V \).

**Acknowledgements** The authors would like to express their sincere gratitude to the editor for his interest on this work, and to anonymous referees, who pointed out some technical problems in the earlier versions of this paper, and whose questions, comments and suggestions lead us to clarify and improve many places of this paper.
References

1. Ambrosio, L., Gigli, N., Savré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. Birkhäuser-Verlag, Berlin (2005)
2. Anderson, G., Guionnet, A., Zeitouni, O.: An Introduction to Random Matrices. Cambridge University Press, Cambridge (2010)
3. Biane, P.: Free Brownian motions, free stochastic calculus and random matrices. In: Voiculescu, D.V. (ed.) Free Probability Theory, pp. 1–19. American Mathematical Society, Providence (1995)
4. Biane, P.: Logarithmic Sobolev inequalities, matrix models and free entropy. Acta. Math. Sin. 19(3), 497–506 (2003)
5. Biane, P., Speicher, R.: Free diffusions, free energy and free Fisher information. Ann. Inst. H. Poincaré Probab. Stat. 37, 581–606 (2001)
6. Blower, G.: Displacement convexity for the generalized orthogonal ensemble. J. Stat. Phys. 116(5/6), 1359–1387 (2004)
7. Blower, G.: Random Matrices: High Dimensional Phenomena, London Math. Soc., Lect. Notes Ser. 367, Cambridge University Press, Cambridge (2009)
8. Boutet de Monvel, A., Pastur, L., Shcherbina, M.: On the statistical mechanics approach in the random matrix theory: Integrated Density of States. J. Stat. Phys. 79, 585–611 (1995)
9. Brenier, Y.: Polar factorization and monotone rearrangement of vector-valued functions. Commun. Pure Appl. Math. 44(4), 375–417 (1991)
10. Carlen, E.A., Gangbo, W.: Constrained steepest descent in the 2-Wasserstein metric. Ann. Math 157, 807–846 (2003)
11. Carrillo, J., McCann, R., Villani, C.: Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. Rev. Mat. Iberoamericana 19, 971–1018 (2003)
12. Cépa, E., Lépingle, D.: Diffusing particles with electrostatic repulsion. Probab. Theory Relat. Fields 107(4), 429–449 (1997)
13. Chan, T.: The Wigner semi-circle law and eigenvalues of matrix-valued diffusions. Probab. Theory Relat. Fields 93, 249–272 (1992)
14. Chekhov, L., Makeenko, Y.: The multicritical Kontsevich–Penner model. Mod. Phys. Lett. A. 7(14), 1223–1236 (1992)
15. Demengel, F., Demengel, G.: Functional Spaces for the Theory of Elliptic Partial Differential Equations. Springer, Berlin (2007)
16. Donati-Martin, C., Groux, B.: Benjamin, Maïda, Convergence to equilibrium in the free Fokker-Planck equation with a double-well potential. Ann. Inst. Henri Poincaré Probab. Stat. 54(4), 1805–1818 (2018)
17. Dyson, F.J.: Statistical theory of the energy levels of complex systems. I, II, and III. J. Math. Phys. 3, 140–156, 157–165, 166–175 (1962)
18. Dyson, F.J.: A Brownian-motion model of the eigenvalues of a random matrix. J. Math. Phys. 3, 1191–1198 (1962)
19. Ethier, S.N., Kurtz, T.G.: Markov Processes: Characterization and Convergence. Wiley, New York (1986)
20. Fontbona, J.: Nonlinear martingale problems involving singular integrals. J. Funct. Anal. 200, 198–236 (2003)
21. Fontbona, J.: Uniqueness for a weak nonlinear evolution equation and large deviations for diffusing particles with electrostatic repulsion. Stoch. Process. Appl. 112, 119–144 (2004)
22. Graczyk, P., Maecki, J.: Multidimensional Yamada-Watanabe theorem and its applications to particle systems. J. Math. Phys. 54(021503), 1–15 (2013)
23. Guionnet, A.: Large Random Matrices: Lectures on Macroscopic Asymptotics. Springer, Berlin (2008)
24. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes. North-Holland Publishing Company, Amsterdam (1981)
25. Johansson, K.: On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. 91, 151–204 (1998)
26. Ledoux, M., Popescu, I.: Mass transportation proofs of free functional inequalities, and free Poincaré inequalities. J. Funct. Anal. 257, 1175–1221 (2009)
27. Li, S., Li, X., Xie, Y.: Matrix diffusion processes and the generalized Dyson Brownian motion, in preparation
28. McCann, R.: A convexity principle for interacting gases. Adv. Math. 128, 153–179 (1997)
29. Mehta, M.: Random Matrices, 3rd edn. Elsevier, Singapore (2006)
30. Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. Commun. Partial Differ. Equ. 26(1/2), 101–174 (2001)
31. Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. 173, 361–400 (2000)
On the Law of Large Numbers for the...