

# Regularity of Schrödinger’s functional equation in the weak topology and moment measures

By Toshio MIKAMI

(Received Jan. 3, 2019)

(Revised Aug. 29, 2019)

**Abstract.** We study the continuity and the measurability of the solution to Schrödinger’s functional equation, with respect to space, kernel and marginals, provided the space of all Borel probability measures is endowed with the weak topology. This is a continuation of our previous result where the space of all Borel probability measures was endowed with the strong topology. As an application, we construct a convex function of which the moment measure is a given probability measure, by the zero noise limit of a class of stochastic optimal transportation problems.

## 1. Introduction.

Schrödinger considered the following problem to find the statistical property of a particle on a finite time interval. Suppose that there exist  $N \geq 2$  particles in a set  $A := \{a_1, \dots, a_{n_0}\} \subset \mathbf{R}^3$  and each particle moves independently, with a given transition probability, to a set  $B := \{b_1, \dots, b_{n_1}\} \subset \mathbf{R}^3$ , where  $1 \leq n_0, n_1 \leq N$ . Find the maximal probability of such events, provided the numbers of particles in each point  $a_i, b_j$  are fixed (see Section 7 in [42] and also [41]). Though he did not succeed in finding the maximal probability, he obtained Euler’s equation for the variational problem above. The continuum limit is called Schrödinger’s functional equation (see [5], [9], [22], [24] for the solution of this problem). Bernstein [4] generalized Schrödinger’s idea and introduced the so-called Bernstein processes which are also called reciprocal processes. The theory of stochastic differential equation for Schrödinger’s functional equation was given by Jamison [25]. The solution is Doob’s h-path process (see [15]) with given two end point marginals. Schrödinger’s problem is also related to the theory of large deviations, the optimal mass transportation problem, entropic estimates and functional inequalities (see e.g. [1], [2], [11], [12], [16], [18], [23], [26], [27], [28], [30], [31], [35], [36], [37], [38], [44] and the references therein).

We describe Schrödinger’s functional equation (see e.g. [24]) in the setting considered in this paper. Let  $S$  be a  $\sigma$ -compact metric space and  $q \in C(S \times S; (0, \infty))$ . For Borel probability measures  $\mu_1, \mu_2$  on  $S$ , find nonnegative  $\sigma$ -finite Borel measures  $\nu_1, \nu_2$  on  $S$  for which the following holds:

---

2010 *Mathematics Subject Classification.* Primary 60G30; Secondary 93E20.

*Key Words and Phrases.* Schrödinger’s functional equation, regularity, moment measure, stochastic optimal transportation.

This work was supported by JSPS KAKENHI Grant Numbers JP26400136 and JP16H03948.

$$\begin{cases} \mu_1(dx) = \nu_1(dx) \int_S q(x, y) \nu_2(dy), \\ \mu_2(dy) = \nu_2(dy) \int_S q(x, y) \nu_1(dx). \end{cases} \quad (1.1)$$

It is known that there exists a solution  $(\nu_1, \nu_2)$  of (1.1) (see [9], [24]).  $(\nu_1, \nu_2)$  is unique up to a constant though the product measure  $\nu_1 \times \nu_2$  is unique. Indeed, for any  $C > 0$ ,  $(C\nu_1, C^{-1}\nu_2)$  is also a solution of (1.1). By the uniqueness of the solution to (1.1), we mean that of the product measure  $\nu_1 \times \nu_2$ . Let  $\{K_m\}_{m \geq 1}$  be a nondecreasing sequence of compact subsets of  $S$  such that  $S = \bigcup_{m \geq 1} K_m$ , where  $K_1 \equiv S$  when  $S$  is compact. When we consider  $\nu_1$  and  $\nu_2$  separately, considering  $(C(\mu_1, \mu_2)\nu_1, C(\mu_1, \mu_2)^{-1}\nu_2)$  if necessary, we assume that the following holds:

$$\nu_1(K_{m(\mu_1, \mu_2)}) = \nu_2(K_{m(\mu_1, \mu_2)}), \quad (1.2)$$

where

$$m(\mu_1, \mu_2) := \min\{m \geq 1 \mid \mu_1(K_m)\mu_2(K_m) > 0\}, \quad C(\mu_1, \mu_2) := \left( \frac{\nu_2(K_{m(\mu_1, \mu_2)})}{\nu_1(K_{m(\mu_1, \mu_2)})} \right)^{1/2}.$$

$$\mu(dx dy) := \nu_1(dx)q(x, y)\nu_2(dy), \quad (1.3)$$

$$u_i(x_i) := \log \left( \int_S q(x_1, x_2) \nu_j(dx_j) \right), \quad i, j = 1, 2, \quad i \neq j. \quad (1.4)$$

Then  $\exp(u_1(x))$  and  $\exp(u_2(x))$  are positive and

$$\mu(dx dy) = q(x, y) \exp(-u_1(x) - u_2(y)) \mu_1(dx) \mu_2(dy). \quad (1.5)$$

(1.1) can be rewritten as follows: for  $i, j = 1, 2, i \neq j$ ,

$$\mu_i(dx_i) = \exp(-u_i(x_i)) \mu_i(dx_i) \int_S q(x_1, x_2) \exp(-u_j(x_j)) \mu_j(dx_j). \quad (1.6)$$

In particular, Schrödinger's problem (1.1) is equivalent to finding functions  $u_1$  and  $u_2$  for which (1.6) holds.  $(u_1, u_2)$  is unique up to a constant though  $u_1(x_1) + u_2(x_2)$  is unique.

Let  $\mathcal{M}(S)$  and  $\mathcal{P}(S)$  denote the space of all Radon measures and that of all Borel probability measures on  $S$ , respectively, where a Radon measure means a locally finite and inner regular Borel measure. It is easy to see that  $\nu_1$  and  $\nu_2$  are functionals of  $\mu_1, \mu_2$  and  $q$ :

$$\nu_i(dx) = \nu_i(dx; q, \mu_1, \mu_2), \quad u_i(x) = u_i(x; q, \mu_1, \mu_2), \quad i = 1, 2. \quad (1.7)$$

In [33], we considered the case where  $\mathcal{P}(S)$  is endowed with the strong topology and showed that if  $S$  is compact, then the following is continuous:

$$\begin{aligned} \nu_i(dx; \cdot, \cdot, \cdot) &: C(S \times S) \times \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow \mathcal{M}(S), \\ \{u_i(x; \cdot, \cdot, \cdot)\}_{x \in S} &: C(S \times S) \times \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow C(S) \end{aligned}$$

and  $u_i \in C(S \times C(S \times S) \times \mathcal{P}(S) \times \mathcal{P}(S))$ . Here  $\mathcal{M}(S)$  is endowed with the strong topology and  $C(S \times S)$  and  $C(S)$  are endowed with the topology induced by the uniform convergence on  $S \times S$  and  $S$ , respectively. We also showed that if  $S$  is  $\sigma$ -compact, then the following is Borel measurable:

$$\int_S f(x) \nu_i(dx; \cdot, \cdot, \cdot) : C(S \times S) \times \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow \mathbb{R}, \quad f \in C_0(S),$$

$$u_i : S \times C(S \times S) \times \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow \mathbb{R} \cup \{\infty\}.$$

As an application of this measurability result, we showed that the coefficients of the mean field PDE system for the h-path process with given two end point marginals are measurable functions of space, time and marginal.

REMARK 1.1. (1.2) was assumed in [33] and implies that for  $i, j = 1, 2$ ,

$$\nu_i(dx_i; q, \mu_1, \mu_2) = \frac{\nu_i(dx_i; q, \mu_1, \mu_2) \nu_j(K_{m(\mu_1, \mu_2)}; q, \mu_1, \mu_2)}{(\nu_1(K_{m(\mu_1, \mu_2)}; q, \mu_1, \mu_2) \nu_2(K_{m(\mu_1, \mu_2)}; q, \mu_1, \mu_2))^{1/2}}.$$

In particular, the measurability of  $(q, \mu_1, \mu_2) \mapsto \nu_1(dx_1; q, \mu_1, \mu_2) \nu_2(dx_2; q, \mu_1, \mu_2)$  implies that of  $(q, \mu_1, \mu_2) \mapsto \nu_i(dx_i; q, \mu_1, \mu_2)$ .

In this paper we consider the case where  $\mathcal{P}(S)$  is endowed with the weak topology and show the continuity and measurability results on  $\nu_i$  and  $u_i$  (see Theorem 2.1 and Corollaries 2.1–2.3 in Section 2). Our continuity result in the weak topology is useful when one considers the existence of a minimizer of a variational problem (see [10] for the continuity result on optimal transport). Indeed, it is not easy to show that a minimizing sequence is compact in the strong topology. As an application (see Theorem 2.2 in Section 2), we give a stochastic optimal transportation approach to moment measures (see [13], [40]). The definition of a moment measure of a convex function is the following.

DEFINITION 1.1. Given a convex function  $u : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ , the following is called the moment measure of  $u$ :

$$\mu(dx) := (Du)_{\#}(\exp(-u(x))dx). \quad (1.8)$$

REMARK 1.2. If  $\mu$  is a moment measure of a convex function  $u : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ , then  $\exp(-u(x))dx \delta_{Du(x)}(dy)$  is the unique minimizer of the 2-Wasserstein distance  $W_2(\exp(-u(x))dx, \mu(dx))$ , provided  $\int_{\mathbb{R}^d} \exp(-u(x))dx = 1$  and  $W_2(\exp(-u(x))dx, \mu(dx))$  is finite (see [7], [8], [44]). Here  $\delta_x(dy)$  denotes the delta measure on  $\{x\}$  and for  $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$ ,

$$W_2(\mu_1, \mu_2) := \left( \inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 m(dx dy) \mid m \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d), \right. \right. \\ \left. \left. m(dx \times \mathbb{R}^d) = \mu_1(dx), m(\mathbb{R}^d \times dy) = \mu_2(dy) \right\} \right)^{1/2}. \quad (1.9)$$

We describe an application of our continuity result more precisely. Let  $\varepsilon > 0$  and

let  $W(t)$  and  $\gamma(t) = \gamma(t; \omega)$  denote a  $d$ -dimensional Brownian motion and a progressively measurable  $\mathbb{R}^d$ -valued stochastic process on a filtered probability space, respectively. Consider the following SDE in a weak sense (see e.g. [20]):

$$dX^{\varepsilon, \gamma}(t) = \gamma(t)dt + \sqrt{\varepsilon}dW(t). \quad (1.10)$$

For  $P_0, P_1 \in \mathcal{P}(\mathbb{R}^d)$ ,

$$V_\varepsilon(P_0, P_1) := \inf \left\{ E \left[ \int_0^1 \frac{1}{2\varepsilon} |\gamma(t)|^2 dt \right] \middle| PX^{\varepsilon, \gamma}(t)^{-1} = P_t, t = 0, 1 \right\}, \quad (1.11)$$

where  $V_\varepsilon(P_0, P_1) := \infty$  if the set over which the infimum is taken is empty (see [1], [2], [16], [18], [19], [27] for related problems on large deviations). For  $P \in \mathcal{P}(\mathbb{R}^d)$ ,

$$\mathcal{S}(P) := \begin{cases} \int_{\mathbb{R}^d} p(x) \log p(x) dx, & \text{if } p(x) := \frac{P(dx)}{dx} \text{ exists,} \\ \infty, & \text{otherwise.} \end{cases} \quad (1.12)$$

For  $\varepsilon, r > 0$ ,  $P_1 \in \mathcal{P}(\mathbb{R}^d)$ ,

$$\begin{aligned} & \Psi_{\varepsilon, r}(P_1) \\ & := \inf \left\{ \mathcal{S}(P) - \varepsilon V_\varepsilon(P, P_1) + \frac{1}{2} \int_{\mathbb{R}^d} |x|^2 P(dx) \middle| P(dx) = p(x)dx \in \mathcal{P}(B_r) \right\}, \end{aligned} \quad (1.13)$$

where

$$B_r := \{x \in \mathbb{R}^d : |x| \leq r\}.$$

By our weak continuity result of  $(q, \mu_1, \mu_2) \mapsto \mu(dxdy; q, \mu_1, \mu_2)$ , we can easily prove the existence of a minimizer  $P_{0, r, \varepsilon}$  of  $\Psi_{\varepsilon, r}(P_1)$  from the lower semicontinuity of a relative entropy and of  $\mathcal{S}$  with respect to the weak topology (see (1.20) and also Lemmas 3.4 and 3.5 in Section 3). We show that a subsequence of  $\{p_{0, r, \varepsilon}(x)dx\}_{\varepsilon > 0}$  weakly converges, as  $\varepsilon \rightarrow 0$ , to a Borel probability measure  $p_0(x)dx$  such that  $-\log p_0(x)$  is convex and  $P_1$  is a moment measure of  $-\log p_0(x)$ . This is formally implied by the representation of  $P_{0, r, \varepsilon}$  and the SDE for the minimizer of  $V_\varepsilon(P_{0, r, \varepsilon}, P_1)$  (see (2.11) and (1.17)). We also show that  $p_{0, r, \varepsilon}(x)$  has a subsequence which uniformly converges, as  $\varepsilon \rightarrow 0$ , to  $p_0(x)$ , provided  $P_1$  is compactly supported.

$\Psi_{\varepsilon, r}(P_1)$  formally converges, as  $\varepsilon \rightarrow 0$ , to the functional considered in [40] where they take the infimum over  $\mathcal{P}(\mathbb{R}^d)$  instead of  $\mathcal{P}(B_r)$ . Our approach makes the proof easier than [40] since  $\mathcal{P}(B_r)$  is compact in the weak topology but can not be applied if we replace  $\mathcal{P}(B_r)$  by  $\mathcal{P}(\mathbb{R}^d)$ , which we regret.

In the proof of the representation of  $P_{0, r, \varepsilon}$  in (2.11), we also make use of properties of the solution to Schrödinger's functional equation and the duality theorem for  $V_\varepsilon(P_0, P_1)$ :

$$V_\varepsilon(P_0, P_1) = \sup \left\{ \int_{\mathbb{R}^d} f(x) P_1(dx) - \int_{\mathbb{R}^d} \varphi(0, x; f) P_0(dx) \middle| f \in C_b^\infty(\mathbb{R}^d) \right\}. \quad (1.14)$$

Here the supremum is taken over all classical solutions  $\varphi(t, x; f)$  to the following Hamilton–Jacobi–Bellman PDE:

$$\begin{aligned} \frac{\partial \varphi(t, x)}{\partial t} + \frac{\varepsilon}{2} \Delta_x \varphi(t, x) + \frac{\varepsilon}{2} |D_x \varphi(t, x)|^2 &= 0, \quad (t, x) \in [0, 1) \times \mathbb{R}^d, \\ \varphi(1, x) &= f(x) \end{aligned} \quad (1.15)$$

(see [30], [31], [34], [43] and the references therein).

$$\begin{aligned} g_\varepsilon(t, z) &:= \frac{1}{(2\pi\varepsilon t)^{d/2}} \exp\left(-\frac{|z|^2}{2\varepsilon t}\right), \quad t > 0, \quad z \in \mathbb{R}^d, \\ g_\varepsilon(t)(x, y) &:= g_\varepsilon(t, y - x), \quad t > 0, \quad x, y \in \mathbb{R}^d. \end{aligned} \quad (1.16)$$

It is known that for any  $P_0, P_1 \in \mathcal{P}(\mathbf{R}^d)$  for which  $P_1(dy) \ll dy$ , there exists the unique weak solution to the following two end points problem of SDE (see [25] and also [33], [34]):

$$\begin{aligned} dX^\varepsilon(t) &= \varepsilon D_x u_1(X^\varepsilon(t); g_\varepsilon(1-t), PX^\varepsilon(t)^{-1}, P_1) dt + \sqrt{\varepsilon} dW(t), \quad 0 < t < 1, \\ PX^\varepsilon(t)^{-1} &= P_t, \quad t = 0, 1. \end{aligned} \quad (1.17)$$

$X^\varepsilon(t)$  is called the h-path process for  $\sqrt{\varepsilon} W(t)$  on  $[0, 1]$  with initial and terminal distribution  $P_0$  and  $P_1$ , respectively. The following is also known:

$$P(X^\varepsilon(0), X^\varepsilon(1))^{-1}(dxdy) = \nu_1(dx; g_\varepsilon(1), P_0, P_1) g_\varepsilon(1, y-x) \nu_2(dy; g_\varepsilon(1), P_0, P_1). \quad (1.18)$$

Suppose that  $V_\varepsilon(P_0, P_1)$  is finite (see Remark 2.2 in Section 2 for a sufficient condition). Then  $X^\varepsilon$  in (1.17) is the unique minimizer of  $V_\varepsilon(P_0, P_1)$  (see [14], [21], [26]–[39], [43], [45] and the references therein). Besides, there exists  $f_o \in L^1(P_1)$  which is unique up to a constant such that the following holds (see [30], [31], [33], [34], [43] and the references therein and also (1.5)):

$$f_o(y) - \varphi(0, x; f_o) = \log p_1(y) - u_2(y; g_\varepsilon(1), P_0, P_1) - u_1(x; g_\varepsilon(1), P_0, P_1). \quad (1.19)$$

In particular, the following holds:

$$\begin{aligned} V_\varepsilon(P_0, P_1) &= \int_{\mathbb{R}^d} f_o(x) P_1(dx) - \int_{\mathbb{R}^d} \varphi(0, x; f_o) P_0(dx) \\ &= \mathcal{S}(P_1) - \int_{\mathbb{R}^d} u_2(x; g_\varepsilon(1), P_0, P_1) P_1(dx) \\ &\quad - \int_{\mathbb{R}^d} u_1(x; g_\varepsilon(1), P_0, P_1) P_0(dx) \\ &= H(P(X^\varepsilon(0), X^\varepsilon(1))^{-1}(dxdy) | P_0(dx) g_\varepsilon(1, y-x) dy) \\ &= \mathcal{S}(P_1) - H(P_0(dx) P_1(dy) | P(X^\varepsilon(0), X^\varepsilon(1))^{-1}(dxdy)) \\ &\quad - \int_{\mathbb{R}^d \times \mathbb{R}^d} \log g_\varepsilon(1, y-x) P_0(dx) P_1(dy). \end{aligned} \quad (1.20)$$

Here  $H$  denotes the relative entropy of two measures: for  $m, n \in \mathcal{P}(S \times S)$ ,

$$H(m|n) = \begin{cases} \int_{S \times S} \log \frac{m(dxdy)}{n(dxdy)} m(dxdy), & \text{if } m \ll n, \\ \infty, & \text{otherwise.} \end{cases} \quad (1.21)$$

REMARK 1.3. If  $V_\varepsilon(P_0, P_1)$  is finite, then  $P_1(dy) \ll dy$ . Indeed,  $V_\varepsilon(P_0, P_1)$  is the relative entropy of  $P(X^\varepsilon)^{-1}$  with respect to  $P_0 * P(\sqrt{\varepsilon}W)^{-1}$  on  $C([0, 1]; \mathbb{R}^d)$  and

$$P_0 * P(\sqrt{\varepsilon}W(1))^{-1}(dy) = \left( \int_{\mathbb{R}^d} g_\varepsilon(1, y-x) P_0(dx) \right) dy.$$

Here  $*$  denotes the convolution of two measures.

In Section 2 we state our main results and prove them in Section 4 by lemmas which are given in Section 3.

## 2. Main result.

In this section we state our main results. We first describe assumptions precisely.

(A1)  $S$  is a complete  $\sigma$ -compact metric space.

(A1)'  $S$  is a compact metric space.

(A2)  $q \in C(S \times S; (0, \infty))$ .

We remark that  $\mathcal{P}(S)$  is endowed with the weak topology and  $C(S \times S)$  is endowed with the topology induced by the uniform convergence on every compact subset of  $S$ .

Under (A1), let  $\{\varphi_m\}_{m \geq 1}$  be a nondecreasing sequence of functions in  $C_0(S; [0, 1])$  such that the following holds:

$$\varphi_m(x) = 1, \quad x \in K_m, \quad m \geq 1$$

(see (1.2)). If  $S = \mathbb{R}^d$ , then  $K_m := B_m$  and we assume that  $\varphi_m \in C_0(B_{m+1}; [0, 1])$ . For  $i \neq j$ ,  $i, j = 1, 2$ ,

$$u_{i|m}(x_i; q, \mu_1, \mu_2) := \log \left( \int_S q(x_1, x_2) \varphi_m(x_j) \nu_j(dx_j; q, \mu_1, \mu_2) \right), \quad (2.1)$$

provided the right hand side is well defined (see (1.7) and also (1.4)).

$$\mu(dxdy; q, \mu_1, \mu_2) := \nu_1(dx; q, \mu_1, \mu_2) q(x, y) \nu_2(dy; q, \mu_1, \mu_2). \quad (2.2)$$

The following is the continuity result on  $\nu_1 \times \nu_2$ ,  $\mu$  and  $u_{i|m}$ .

THEOREM 2.1. *Suppose that (A1) and (A2) hold and that  $q_n \in C(S \times S; (0, \infty))$ ,  $\mu_i, \mu_{i,n} \in \mathcal{P}(S)$ ,  $n \geq 1$ ,  $i = 1, 2$  and*

$$\lim_{n \rightarrow \infty} q_n = q, \quad \text{locally uniformly,} \quad (2.3)$$

$$\lim_{n \rightarrow \infty} \mu_{1,n} \times \mu_{2,n} = \mu_1 \times \mu_2, \quad \text{weakly.} \quad (2.4)$$

Then for any  $f \in C_0(S \times S)$ ,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{S \times S} f(x, y) \nu_1(dx; q_n, \mu_{1,n}, \mu_{2,n}) \nu_2(dy; q_n, \mu_{1,n}, \mu_{2,n}) \\ &= \int_{S \times S} f(x, y) \nu_1(dx; q, \mu_1, \mu_2) \nu_2(dy; q, \mu_1, \mu_2). \end{aligned} \quad (2.5)$$

In particular,

$$\lim_{n \rightarrow \infty} \mu(dx dy; q_n, \mu_{1,n}, \mu_{2,n}) = \mu(dx dy; q, \mu_1, \mu_2), \quad \text{weakly.} \quad (2.6)$$

For any  $\{x_{i,n}\}_{n \geq 1} \subset S$  which converges, as  $n \rightarrow \infty$ , to  $x_i \in S$ ,  $i = 1, 2$  and for sufficiently large  $m \geq 1$ ,

$$\lim_{n \rightarrow \infty} \sum_{i=1}^2 u_{i|m}(x_{i,n}; q_n, \mu_{1,n}, \mu_{2,n}) = \sum_{i=1}^2 u_{i|m}(x_i; q, \mu_1, \mu_2). \quad (2.7)$$

Since  $(\mu_1, \mu_2) \mapsto m(\mu_1, \mu_2)$  is measurable, Theorem 2.1 implies the following.

**COROLLARY 2.1.** *Suppose that (A1) and (A2) hold. Then the following are Borel measurable: for  $i = 1, 2$ ,*

$$\begin{aligned} & \int_S f(x) \nu_i(dx; \cdot, \cdot, \cdot) : C(S \times S) \times \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow \mathbb{R}, \quad f \in C_0(S), \\ & u_i : S \times C(S \times S) \times \mathcal{P}(S) \times \mathcal{P}(S) \rightarrow \mathbb{R} \cup \{\infty\}. \end{aligned}$$

If  $S$  is compact, then  $\nu_1(S) = \nu_2(S)$  (see (1.2)). This implies, from Theorem 2.1, the following of which the proof is omitted.

**COROLLARY 2.2.** *Suppose that (A1)' and the assumptions of Theorem 2.1 except (A1) hold. Then the following holds: for  $i = 1, 2$ ,*

$$\lim_{n \rightarrow \infty} \int_S f(x) \nu_i(dx; q_n, \mu_{1,n}, \mu_{2,n}) = \int_S f(x) \nu_i(dx; q, \mu_1, \mu_2), \quad f \in C(S),$$

and for any  $\{x_n\}_{n \geq 1} \subset S$  which converges, as  $n \rightarrow \infty$ , to  $x \in S$ ,

$$\lim_{n \rightarrow \infty} u_i(x_n; q_n, \mu_{1,n}, \mu_{2,n}) = u_i(x; q, \mu_1, \mu_2).$$

A uniformly bounded sequence of convex functions on a convex neighborhood  $N_A$  of a convex subset  $A$  of  $\mathbb{R}^d$  is compact in  $C(A)$ , provided  $\text{dist}(A, N_A^c)$  is positive (see e.g. [3, Section 3.3]). We describe an additional assumption and state a stronger result than above, provided  $S \subset \mathbb{R}^d$ .

(A3.r) There exists  $C_r > 0$  for which  $x \mapsto C_r|x|^2 + \log q(x, y)$  and  $y \mapsto C_r|y|^2 + \log q(x, y)$  are convex on  $B_r$  for any  $y \in B_r$  and any  $x \in B_r$ , respectively.

REMARK 2.1. If  $\log q(x, y)$  has bounded second order partial derivatives on  $B_r$ , then (A3.r) holds.

$$\|f\|_{\infty, r} := \sup_{x \in B_r} |f(x)|, \quad f \in C(B_r). \quad (2.8)$$

The following is a stronger convergence result than Corollary 2.2.

COROLLARY 2.3. *Let  $r > 0$ . Suppose that (A3.r) and the assumptions of Corollary 2.2 with  $S = B_r$  hold. Then for any  $r' < r$ ,*

$$\lim_{n \rightarrow \infty} \sum_{i=1}^2 \|u_i(\cdot; q_n, \mu_{1,n}, \mu_{2,n}) - u_i(\cdot; q, \mu_1, \mu_2)\|_{\infty, r'} = 0. \quad (2.9)$$

$$\mathcal{P}_p(\mathbb{R}^d) := \left\{ P \in \mathcal{P}(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} |x|^p P(dx) < \infty \right\}, \quad p \geq 1. \quad (2.10)$$

As an application of our regularity result, we show that there exists a convex function of which the moment measure is a given probability measure.

THEOREM 2.2. *For any  $P_1(dx) = p_1(x)dx \in \mathcal{P}_2(\mathbb{R}^d)$  for which  $\mathcal{S}(P_1)$  is finite, there exists a minimizer of  $\Psi_{\varepsilon, r}(P_1)$ . For any minimizer  $P_{0, r, \varepsilon}(dx) = p_{0, r, \varepsilon}(x)dx$  of  $\Psi_{\varepsilon, r}(P_1)$ ,*

$$p_{0, r, \varepsilon}(x) = \frac{1}{C_\varepsilon} I_{B_r}(x) \exp\left(-\varepsilon u_1(x; g_\varepsilon(1), P_{0, r, \varepsilon}, P_1) - \frac{1}{2}|x|^2\right), \quad (2.11)$$

where  $C_\varepsilon$  is a normalizing constant. Besides, there exists a subsequence of  $p_{0, r, \varepsilon}(x)dx$  which weakly converges, as  $\varepsilon \rightarrow 0$ , to a probability measure  $p_0(x)dx$  such that  $p_1(x)dx$  is a moment measure of  $-\log p_0$ . Suppose, in addition, that  $P_1$  is compactly supported. Then there exists a subsequence of  $p_{0, r, \varepsilon}(x)$  which uniformly converges, as  $\varepsilon \rightarrow 0$ , to a probability density function  $p_0(x)$  such that  $p_1(x)dx$  is a moment measure of  $-\log p_0$ .

REMARK 2.2. If  $P_0, P_1(dx) = p_1(x)dx \in \mathcal{P}_2(\mathbb{R}^d)$  and  $\mathcal{S}(P_1)$  is finite, then  $V_\varepsilon(P_0, P_1)$  is finite. Indeed, from the last equality of (1.20),

$$V_\varepsilon(P_0, P_1) \leq \mathcal{S}(P_1) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \log g_\varepsilon(1, y - x) P_0(dx) p_1(y) dy$$

since, the relative entropy is nonnegative.

### 3. Lemmas.

In this section we state and prove lemmas. When it is not confusing, we omit the dependence of  $u_i, \nu_i$  on  $q, \nu_1, \nu_2$ .

#### 3.1. Lemmas for the proof of Theorem 2.1 and Corollary 2.3.

The following lemma will be used in the proof of Theorem 2.1.

LEMMA 3.1. *Suppose that (A1) and (A2) hold. Then for any  $\mu_1, \mu_2 \in \mathcal{P}(S)$ ,  $\mu$  defined by (1.3),*

$$\begin{aligned} \min_{x,y \in K_m} q(x,y)^{-1} \mu(K_m \times K_m) &\leq \int_S \varphi_m(x) \nu_1(dx) \int_S \varphi_m(y) \nu_2(dy) \\ &\leq \max_{x,y \in \text{supp}(\varphi_m)} q(x,y)^{-1}. \end{aligned} \quad (3.1)$$

PROOF. The proof is done by the following (see (1.3)):

$$\nu_1(dx) \nu_2(dy) = q(x,y)^{-1} \mu(dxdy). \quad \square$$

For  $r > 0$  and  $q \in C(B_r \times B_r; (0, \infty))$ ,

$$\begin{aligned} m_{q,r} &:= \min\{q(x,y) \mid |x|, |y| \leq r\}, \\ M_{q,r} &:= \max\{q(x,y) \mid |x|, |y| \leq r\}. \end{aligned} \quad (3.2)$$

Lemmas 3.2 and 3.3 will be used to prove Corollary 2.3.

LEMMA 3.2 ([5, p. 194]). *Let  $r > 0$ . Suppose that (A2) with  $S = B_r$  holds. Then, for any  $\mu_1, \mu_2 \in \mathcal{P}(B_r)$ , the following holds (see (1.4) for notation):*

$$\frac{m_{q,r}}{\sqrt{M_{q,r}}} \leq \exp(u_i(x)) \leq \frac{M_{q,r}}{\sqrt{m_{q,r}}}, \quad x \in B_r, \quad i = 1, 2. \quad (3.3)$$

By the method of proving the convexity of a log moment generating function, we obtain the following.

LEMMA 3.3. *Let  $C$  and  $\nu \in \mathcal{M}(\mathbb{R}^d)$  be a convex subset of  $\mathbb{R}^d$  and a nonnegative Radon measure, respectively. Suppose that  $C \ni x \mapsto f(x, y)$  is convex,  $\nu(dy)$ -a.e. Then  $C \ni x \mapsto \log \int_{\mathbb{R}^d} \exp(f(x, y)) \nu(dy)$  is convex.*

PROOF. For  $x, y \in C$  and  $\lambda \in (0, 1)$ , by Hölder's inequality,

$$\begin{aligned} &\int_{\mathbb{R}^d} \exp(f(\lambda x + (1-\lambda)y, x_2)) \nu(dx_2) \\ &\leq \int_{\mathbb{R}^d} \exp(\lambda f(x, x_2) + (1-\lambda)f(y, x_2)) \nu(dx_2) \\ &\leq \left( \int_{\mathbb{R}^d} \exp(f(x, x_2)) \nu(dx_2) \right)^\lambda \left( \int_{\mathbb{R}^d} \exp(f(y, x_2)) \nu(dx_2) \right)^{1-\lambda}. \end{aligned} \quad (3.4)$$

□

### 3.2. Lemmas for the proof of Theorem 2.2.

In this subsection, we prove lemmas for the proof of Theorem 2.2. Lemma 3.3 will be also used in the proof of Theorem 2.2.

$$B_{\mathcal{P}_p(\mathbb{R}^d), r} := \left\{ P \in \mathcal{P}_p(\mathbb{R}^d) \mid \int_{\mathbb{R}^d} |x|^p P(dx) \leq r \right\}.$$

The lower semicontinuity of a relative entropy and the continuity result in Theorem 2.1 imply the following.

LEMMA 3.4. *Suppose that Theorem 2.1 holds. Then for any  $r, \varepsilon > 0$ , the following is lower-semicontinuous on  $B_{\mathcal{P}_2(\mathbb{R}^d), r} \times B_{\mathcal{P}_2(\mathbb{R}^d), r}$  (see (1.4), (1.7) and (1.16) for notation):*

$$\mu_1 \times \mu_2 \mapsto -V_\varepsilon(\mu_1, \mu_2) + \mathcal{S}(\mu_2) + \sum_{i=1}^2 \int_{\mathbb{R}^d} \frac{|x|^2}{2\varepsilon} \mu_i(dx). \quad (3.5)$$

PROOF. From (1.20),

$$\begin{aligned} & -V_\varepsilon(\mu_1, \mu_2) + \mathcal{S}(\mu_2) + \sum_{i=1}^2 \int_{\mathbb{R}^d} \frac{|x|^2}{2\varepsilon} \mu_i(dx) \\ &= H(\mu_1(dx)\mu_2(dy)|\mu(dxdy); g_\varepsilon(1), \mu_1, \mu_2) \\ & \quad + \frac{1}{\varepsilon} \left\langle \int_{\mathbb{R}^d} x \mu_1(dx), \int_{\mathbb{R}^d} y \mu_2(dy) \right\rangle - \log(2\pi\varepsilon)^{d/2} \end{aligned} \quad (3.6)$$

(see (1.18) and (2.2) for notation). Since  $(m, n) \mapsto H(m(dxdy)|n(dxdy))$  is lower semicontinuous (see [17, Lemma 1.4.3]), the proof is over from Theorem 2.1.  $\square$

The following lemma can be proved by the lower semicontinuity of a relative entropy.

LEMMA 3.5. *For any  $r > 0$ ,  $\mathcal{S}$  is lower-semicontinuous on  $B_{\mathcal{P}_2(\mathbb{R}^d), r}$  in the weak topology.*

PROOF.

$$q(x) := \frac{(1 + |x|)^{-d-1}}{\int_{\mathbb{R}^d} (1 + |y|)^{-d-1} dy}, \quad x \in \mathbb{R}^d.$$

The proof is done by the following:

$$\mathcal{S}(P) = H(P(dx)|q(x)dx) + \int_{\mathbb{R}^d} \log q(x) P(dx) \quad (3.7)$$

(see e.g. [17, Lemma 1.4.3]).  $\square$

LEMMA 3.6. *Suppose that (A1) and (A2) hold. Then for any  $\mu_1, \mu_2 \in \mathcal{P}(S)$ ,  $\mu$  defined by (1.3) and sufficiently large  $m \geq 1$ ,  $m \mapsto u_{i|m}$  is nondecreasing,  $i = 1, 2$  and the following holds:*

$$\begin{aligned} & \min_{x, y \in \text{supp}(\varphi_m)} \frac{q(x_1, y)q(x, x_2)}{q(x, y)} \times \int_{S \times S} \varphi_m(x)\varphi_m(y)\mu(dxdy) \\ & \leq \exp(u_{1|m}(x_1) + u_{2|m}(x_2)) \\ & \leq \max_{x, y \in \text{supp}(\varphi_m)} \frac{q(x_1, y)q(x, x_2)}{q(x, y)}, \quad x_1, x_2 \in S. \end{aligned} \quad (3.8)$$

PROOF. The proof is done by the following (see (1.3) and (2.1)):

$$\exp(u_{1|m}(x_1) + u_{2|m}(x_2)) = \int_{S \times S} \frac{q(x_1, y)q(x, x_2)}{q(x, y)} \varphi_m(x) \varphi_m(y) \mu(dx dy), \quad (3.9)$$

provided the right hand side is positive.  $\square$

For  $i = 1, 2$ ,  $m \geq 1$ ,  $\varepsilon > 0$ ,  $x \in \mathbb{R}^d$ ,

$$\begin{aligned} \bar{u}_{i,\varepsilon}(x) &:= \varepsilon u_i(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) + \frac{1}{2}|x|^2, \\ \bar{u}_{i|m,\varepsilon}(x) &:= \varepsilon u_{i|m}(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) + \frac{1}{2}|x|^2. \end{aligned} \quad (3.10)$$

In the following lemma, the boundedness of the set  $B_r$  plays a crucial role.

LEMMA 3.7. For any  $\varepsilon, r > 0$  and  $P_1(dx) = p_1(x)dx \in \mathcal{P}(\mathbb{R}^d)$ ,

$$\Psi_{\varepsilon,r}(P_1) \leq -\log\{\text{Vol}(B_r)\} + \frac{1}{2} \int_{B_r} |x|^2 \frac{dx}{\text{Vol}(B_r)}. \quad (3.11)$$

Suppose that  $P_{0,r,\varepsilon}$  in (2.11) is a minimizer of  $\Psi_{\varepsilon,r}(P_1)$ . Then for  $y_0 := \int_{\mathbb{R}^d} x P_1(dx)$ ,

$$\exp\left(-\varepsilon \mathcal{S}(P_1) - \int_{\mathbb{R}^d} \frac{1}{2}|x|^2 P_1(dx) - \Psi_{\varepsilon,r}(P_1)\right) \leq C_\varepsilon \exp(-\bar{u}_{2,\varepsilon}(y_0)). \quad (3.12)$$

In particular, for any sequence  $\{\varepsilon_n\}_{n \geq 1}$  which converges to 0 as  $n \rightarrow \infty$ , the set  $\{x \in B_r \mid \liminf_{n \rightarrow \infty} (\bar{u}_{1,\varepsilon_n}(x) + \bar{u}_{2,\varepsilon_n}(y_0)) < \infty\}$  has a positive Lebesgue measure, provided  $P_1 \in \mathcal{P}_2(\mathbb{R}^d)$  and  $\mathcal{S}(P_1)$  is finite.

PROOF. Let  $p_{\text{uni},r}$  denote the probability density function of the uniform distribution on  $B_r$ . Then the following implies (3.11):

$$\Psi_{\varepsilon,r}(P_1) \leq \mathcal{S}(p_{\text{uni},r}(x)dx) + \frac{1}{2} \int_{B_r} |x|^2 \frac{dx}{\text{Vol}(B_r)}. \quad (3.13)$$

We prove (3.12). We only have to consider the case where  $\mathcal{S}(P_1)$  is finite and  $P_1 \in \mathcal{P}_2(\mathbb{R}^d)$ . From (1.20) and (2.11), by Jensen's inequality,

$$\begin{aligned} \Psi_{\varepsilon,r}(P_1) &= \mathcal{S}(P_{0,r,\varepsilon}) - \varepsilon \left( \mathcal{S}(P_1) - \int_{\mathbb{R}^d} u_2(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) P_1(dx) \right. \\ &\quad \left. - \int_{\mathbb{R}^d} u_1(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) P_{0,r,\varepsilon}(dx) \right) + \int_{\mathbb{R}^d} \frac{1}{2}|x|^2 P_{0,r,\varepsilon}(dx) \\ &= -\log C_\varepsilon - \varepsilon \mathcal{S}(P_1) + \int_{\mathbb{R}^d} \bar{u}_{2,\varepsilon}(x) P_1(dx) - \int_{\mathbb{R}^d} \frac{1}{2}|x|^2 P_1(dx) \\ &\geq -\log C_\varepsilon - \varepsilon \mathcal{S}(P_1) + \bar{u}_{2,\varepsilon}(y_0) - \int_{\mathbb{R}^d} \frac{1}{2}|x|^2 P_1(dx). \end{aligned} \quad (3.14)$$

Indeed, one can show that  $\bar{u}_{2,\varepsilon}$  is convex from Lemma 3.3 and that  $\bar{u}_{2,\varepsilon}$  is finite and continuous on  $\mathbb{R}^d$  since  $\nu_1(dx; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1)$  is a finite measure on  $B_r$ . The last part of this lemma can be shown by Fatou's lemma from (3.8) in Lemma 3.6 and from the following: for  $m > r$ ,

$$\bar{u}_{1,\varepsilon}(x) + \bar{u}_{2,\varepsilon}(y_0) \geq \bar{u}_{1|m,\varepsilon}(x) + \bar{u}_{2,\varepsilon}(y_0), \quad \bar{u}_{2,\varepsilon} = \bar{u}_{2|m,\varepsilon}$$

since  $\nu_1(dx; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1)$  is supported on  $B_r$ .  $\square$

For a convex function  $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ , the 0-sublevel set  $f^{-1}((-\infty, 0])$  is convex. Roughly speaking, the following lemma can be proved from the fact that a uniformly bounded sequence of convex functions defined on the same open set is compact in the sup norm on any compact subset of the open set (see Section 3.3 in [3]).

LEMMA 3.8. (i) For a convex set  $C \subset \mathbb{R}^d$ ,  $\text{dist}(x, C)$  is a convex function.

(ii) For a bounded sequence of convex sets  $\{C_n \subset \mathbb{R}^d\}_{n \geq 1}$ , there exists a closed convex set  $C_\infty$  and a subsequence  $\{C_{n_k}\}_{k \geq 1}$  of  $\{C_n\}_{n \geq 1}$  such that  $\{\text{dist}(x, C_{n_k})\}_{k \geq 1}$  converges, as  $k \rightarrow \infty$ , to  $\text{dist}(x, C_\infty)$  uniformly on every compact subset of  $\mathbb{R}^d$ .

(iii) For any  $\gamma > 0$ , the following holds: for sufficiently large  $k \geq 1$ ,

$$U_{-\gamma}(C_\infty) := \{y \in C_\infty \mid U_\gamma(y) \subset C_\infty\} \subset C_{n_k},$$

where  $U_\gamma(y) := \{x \in \mathbb{R}^d : |x - y| < \gamma\}$ .

PROOF. (i) For  $x_1, x_2 \in \mathbb{R}^d$ ,  $\lambda \in (0, 1)$ ,  $y_1, y_2 \in C$ , since  $\lambda y_1 + (1 - \lambda)y_2 \in C$ ,

$$\begin{aligned} \text{dist}(\lambda x_1 + (1 - \lambda)x_2, C) &\leq |\lambda x_1 + (1 - \lambda)x_2 - (\lambda y_1 + (1 - \lambda)y_2)| \\ &\leq \lambda|x_1 - y_1| + (1 - \lambda)|x_2 - y_2|. \end{aligned} \quad (3.15)$$

Taking the infimum over all  $y_1, y_2 \in C$ , the proof is done.

(ii) Since  $\{C_n\}_{n \geq 1}$  is bounded,  $\{\text{dist}(x, C_n)\}_{n \geq 1}$  is also locally bounded, which implies that there exists a convex function  $h(x)$  and a subsequence  $\{\text{dist}(x, C_{n_k})\}_{k \geq 1}$  such that

$$\text{dist}(x, C_{n_k}) \rightarrow h(x), \quad k \rightarrow \infty,$$

uniformly on every compact subset of  $\mathbb{R}^d$  (see e.g. [3, Section 3.3]).

$$C_\infty := h^{-1}(0).$$

Then it is easy to see that the set  $C_\infty$  is a closed convex set and  $h(x) = \text{dist}(x, C_\infty)$ .

(iii) We only have to consider the case where  $U_{-\gamma}(C_\infty) \neq \emptyset$ . From (ii), for sufficiently large  $k \geq 1$ ,

$$C_\infty \subset U_{\gamma_k}(C_{n_k}), \quad (3.16)$$

where

$$\gamma_k := \sup \left\{ \left| \text{dist}(x, C_{n_k}) \right| + \frac{\gamma}{2} \mid x \in C_\infty \right\} \rightarrow \frac{\gamma}{2}, \quad k \rightarrow \infty.$$

For  $x \in U_{-\gamma}(C_\infty)$ , if  $x \notin C_{n_k}$ , then the following which contradicts (3.16) holds: for  $\tilde{\gamma} < \gamma$ ,

$$\emptyset \neq U_\gamma(x) \cap U_{\tilde{\gamma}}(C_{n_k})^c \subset C_\infty \cap U_{\tilde{\gamma}}(C_{n_k})^c.$$

Indeed, since  $C_{n_k}$  is convex, for  $x \notin C_{n_k}$ , there exists  $p \in \mathbb{R}^d$  such that

$$\{y \mid \langle p, y - x \rangle \geq 0\} \subset C_{n_k}^c. \quad \square$$

#### 4. Proof of main results.

In this section we prove our main results.

PROOF OF THEOREM 2.1. We first prove (2.5). For the sake of simplicity,

$$\begin{aligned} \nu_{i,n}(dx) &:= \nu_i(dx; q_n, \mu_{1,n}, \mu_{2,n}), \\ \mu_n(dx dy) &:= \nu_{1,n}(dx) q_n(x, y) \nu_{2,n}(dy). \end{aligned} \quad (4.1)$$

Since  $\{\mu_{1,n}(dx) = \mu_n(dx \times S), \mu_{2,n}(dy) = \mu_n(S \times dy)\}_{n \geq 1}$  is convergent,  $\{\mu_n\}_{n \geq 1}$  is tight. Indeed, for any Borel sets  $A, B \in \mathbb{R}^d$ ,

$$\mu_n((A \times B)^c) \leq \mu_n(A^c \times S) + \mu_n(S \times B^c) = \mu_{1,n}(A^c) + \mu_{2,n}(B^c),$$

and a convergent sequence of probability measures on a complete separable metric space is tight by Prohorov's theorem (see e.g. [6]). Here notice that a  $\sigma$ -compact metric space is separable. By Prohorov's theorem, take a weakly convergent subsequence  $\{\mu_{n_k}\}_{k \geq 1}$  and denote the limit by  $\mu$ . Then it is easy to see that the following holds:

$$\mu_1(dx) = \mu(dx \times S), \quad \mu_2(dy) = \mu(S \times dy).$$

From (A2) and (2.3)–(2.4), the following holds: for any  $f \in C_0(S \times S)$ ,

$$\lim_{k \rightarrow \infty} \int_{S \times S} f(x, y) \nu_{1, n_k}(dx) \nu_{2, n_k}(dy) = \int_{S \times S} f(x, y) q(x, y)^{-1} \mu(dx dy). \quad (4.2)$$

Indeed,

$$\nu_{1,n}(dx) \nu_{2,n}(dy) = \left( \frac{1}{q_n(x, y)} - \frac{1}{q(x, y)} \right) \mu_n(dx dy) + \frac{1}{q(x, y)} \mu_n(dx dy).$$

The rest of the proof of (2.5) is divided into the following (4.3)–(4.4) which will be proved later.

There exists a subsequence  $\{\bar{n}_k\} \subset \{n_k\}$  and finite measures  $\bar{\nu}_{1,m}, \bar{\nu}_{2,m} \in \mathcal{M}(\text{supp}(\varphi_m))$  such that for sufficiently large  $m \geq 1$  and any  $f \in C_0(S \times S)$ ,

$$\begin{aligned} & \lim_{k \rightarrow \infty} \int_{S \times S} f(x, y) \varphi_m(x) \varphi_m(y) \nu_{1, \bar{n}_k}(dx) \nu_{2, \bar{n}_k}(dy) \\ &= \int_{S \times S} f(x, y) \bar{\nu}_{1,m}(dx) \bar{\nu}_{2,m}(dy). \end{aligned} \quad (4.3)$$

From (4.3), for sufficiently large  $m \geq 1$  and any Borel sets  $A_1, A_2 \subset S$ ,

$$\int_{A_1 \times A_2} q(x, y)^{-1} \mu(dxdy) = \frac{\int_{A_1 \times K_m} q(x, y)^{-1} \mu(dxdy) \int_{K_m \times A_2} q(x, y)^{-1} \mu(dxdy)}{\bar{\nu}_{1,m}(K_m) \bar{\nu}_{2,m}(K_m)}. \quad (4.4)$$

(4.4) implies that  $q(x, y)^{-1} \mu(dxdy)$  is a product measure which satisfies (1.1). (4.2) and the uniqueness of the solution to (1.1) implies that (2.5) is true.

We prove (4.3)–(4.4) to complete the proof of (2.5). (4.3) can be proved by the diagonal method, since  $\{\mu_n\}_{n \geq 1}$  is tight and since for sufficiently large  $m \geq 1$ ,

$$\begin{aligned} & \varphi_m(x_1) \varphi_m(x_2) \nu_{1,n_k}(dx_1) \nu_{2,n_k}(dx_2) \\ &= \int_S \varphi_m(x) \nu_{1,n_k}(dx) \int_S \varphi_m(y) \nu_{2,n_k}(dy) \frac{\varphi_m(x_1) \nu_{1,n_k}(dx_1)}{\int_S \varphi_m(x) \nu_{1,n_k}(dx)} \frac{\varphi_m(x_2) \nu_{2,n_k}(dx_2)}{\int_S \varphi_m(x) \nu_{2,n_k}(dx)} \end{aligned} \quad (4.5)$$

has a convergent subsequence from (3.1) in Lemma 3.1 by Prohorov's theorem and since any weak limit is a product measure. We prove (4.4). From (4.2) and (4.3), for sufficiently large  $\tilde{m} \geq 1$ ,

$$\begin{aligned} & \int_{(A_1 \times A_2) \cap (K_{\tilde{m}} \times K_{\tilde{m}})} q(x, y)^{-1} \mu(dxdy) \\ &= \int_{(A_1 \times A_2) \cap (K_{\tilde{m}} \times K_{\tilde{m}})} \varphi_{\tilde{m}}(x) \varphi_{\tilde{m}}(y) q(x, y)^{-1} \mu(dxdy) \\ &= \int_{(A_1 \times A_2) \cap (K_{\tilde{m}} \times K_{\tilde{m}})} \bar{\nu}_{1,\tilde{m}}(dx) \bar{\nu}_{2,\tilde{m}}(dy) \\ &= \bar{\nu}_{1,\tilde{m}}(A_1 \cap K_{\tilde{m}}) \bar{\nu}_{2,\tilde{m}}(A_2 \cap K_{\tilde{m}}). \end{aligned} \quad (4.6)$$

From (4.6), for  $\tilde{m} \geq m$ , setting  $A_i = K_m$ ,

$$\begin{aligned} \bar{\nu}_{1,\tilde{m}}(A_1 \cap K_{\tilde{m}}) &= \frac{\int_{(A_1 \times K_m) \cap (K_{\tilde{m}} \times K_{\tilde{m}})} q(x, y)^{-1} \mu(dxdy)}{\bar{\nu}_{2,\tilde{m}}(K_m)}, \\ \bar{\nu}_{2,\tilde{m}}(A_2 \cap K_{\tilde{m}}) &= \frac{\int_{(K_m \times A_2) \cap (K_{\tilde{m}} \times K_{\tilde{m}})} q(x, y)^{-1} \mu(dxdy)}{\bar{\nu}_{1,\tilde{m}}(K_m)}, \end{aligned} \quad (4.7)$$

$$\bar{\nu}_{1,\tilde{m}}(K_m) \bar{\nu}_{2,\tilde{m}}(K_m) = \bar{\nu}_{1,m}(K_m) \bar{\nu}_{2,m}(K_m) = \int_{K_m \times K_m} q(x, y)^{-1} \mu(dxdy).$$

Substitute (4.7) to (4.6) and let  $\tilde{m} \rightarrow \infty$ . Then we obtain (4.4). (2.7) can be shown from (2.5) by the following: from (2.1),

$$\begin{aligned} & \exp \left( \sum_{i=1}^2 u_{i|m}(x_{i,n}; q_n, \mu_{1,n}, \mu_{2,n}) \right) \\ &= \int_{S \times S} q_n(x_{1,n}, y) q_n(x, x_{2,n}) \varphi_m(x) \varphi_m(y) \nu_{1,n}(dx) \nu_{2,n}(dy), \end{aligned} \quad (4.8)$$

provided the right hand side is positive.  $\square$

For a compact set  $K \subset S$ ,  $\mathcal{P}(S) \ni \nu \mapsto \nu(K)$  is upper semicontinuous in the weak topology and is hence measurable. Corollary 2.1 can be proved in the same way as in [33] and we omit the proof.

As we mentioned in Section 2, we omit the proof of Corollary 2.2. Corollary 2.2 and Lemmas 3.2 and 3.3 immediately imply Corollary 2.3 (see [3, Section 3.3]) and we omit the proof. Indeed, if any subsequence of a sequence of pointwise convergent continuous functions has a uniformly convergent subsequence, then it is uniformly convergent.

Before we prove Theorem 2.2, we briefly describe the idea of the proof. Theorem 2.1 and Lemmas 3.4–3.5 imply the lower semicontinuity of the functional that we minimize in  $\Psi_{\varepsilon,r}(P_1)$ . (3.11) in Lemma 3.7 implies the finiteness of  $\Psi_{\varepsilon,r}(P_1)$ . In particular, the existence of a minimizer  $p_{0,r,\varepsilon}$  of  $\Psi_{\varepsilon,r}(P_1)$  is obtained. (2.11) can be proved by the Duality theorem (1.14) for  $V_\varepsilon(P_{0,r,\varepsilon}, P_1)$  and by the fact that the relative entropy of two probability measures is nonnegative and is equal to zero if and only if two probability measures are the same. The characterization of the limit  $p_0$  of  $p_{0,r,\varepsilon}$ , as  $\varepsilon \rightarrow 0$ , can be inferred from the following. Roughly speaking, from [28],

$$(2\varepsilon V_\varepsilon(P_{0,r,\varepsilon}, P_1))^{1/2} \sim W_2(p_0(x)dx, P_1), \quad \varepsilon \rightarrow 0$$

(see (1.9) and (1.11) for notation). Besides, there exists a convex function  $u : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$  such that for the minimizer  $X^\varepsilon$  of  $V_\varepsilon(P_{0,r,\varepsilon}, P_1)$ , as  $\varepsilon \rightarrow 0$ ,

$$\begin{aligned} \varepsilon D_x u_1(x; g_\varepsilon(1-t), PX^\varepsilon(t)^{-1}, P_1) + x &\sim Du(x), \quad 0 \leq t \leq 1, \\ \mu(dxdy; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) &\sim P_0(dx)\delta_{Du(x)}(dy). \end{aligned}$$

In particular,

$$\varepsilon u_1(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) + \frac{|x|^2}{2} \sim u(x) + \text{Constant}.$$

PROOF OF THEOREM 2.2. Since  $\mathcal{P}(B_r)$  is tight, by Prohorov's theorem (see e.g. [6]), Lemmas 3.4–3.5 and (3.11) in Lemma 3.7 imply the existence of a minimizer  $P_{0,r,\varepsilon}(dx) = p_{0,r,\varepsilon}(x)dx$  of  $\Psi_{\varepsilon,r}(P_1)$ . By (1.14),

$$\begin{aligned} \Psi_{\varepsilon,r}(P_1) = \inf \left\{ \mathcal{S}(p(x)dx) - \varepsilon \left( \int_{\mathbb{R}^d} f(x)p_1(x)dx - \int_{\mathbb{R}^d} \varphi(0, x; f)p(x)dx \right) \right. \\ \left. + \frac{1}{2} \int_{\mathbb{R}^d} |x|^2 p(x)dx \mid p(x)dx \in \mathcal{P}(B_r), f \in C_b^\infty(\mathbb{R}^d) \right\}. \end{aligned} \quad (4.9)$$

Let  $f_{0,r,\varepsilon}$  denote  $f_0$  in (1.19) with  $P_0 = P_{0,r,\varepsilon}$ . Then

$$\begin{aligned} \Psi_{\varepsilon,r}(P_1) = \inf \left\{ \mathcal{S}(p(x)dx) + \int_{\mathbb{R}^d} \left( \varepsilon \varphi(0, x; f_{0,r,\varepsilon}) + \frac{|x|^2}{2} \right) p(x)dx \right. \\ \left. \mid p(x)dx \in \mathcal{P}(B_r) \right\} - \varepsilon \int_{\mathbb{R}^d} f_{0,r,\varepsilon}(x)p_1(x)dx \end{aligned} \quad (4.10)$$

(see (1.19), (1.5) and Remark 2.2). Indeed, for  $p(x)dx \in \mathcal{P}(B_r)$ ,

$$\int_{\mathbb{R}^d} \varphi(0, x; f_{0,r,\varepsilon}) p(x) dx - \int_{\mathbb{R}^d} f_{0,r,\varepsilon}(x) p_1(x) dx \geq -V_\varepsilon(p(x) dx, P_1),$$

since

$$\begin{aligned} & \int_{\mathbb{R}^d} \varphi(0, x; f_{0,r,\varepsilon}) p(x) dx - \int_{\mathbb{R}^d} f_{0,r,\varepsilon}(x) p_1(x) dx \\ &= - \int_{\mathbb{R}^d \times \mathbb{R}^d} \log \left( \frac{\mu(dxdy; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1)}{P_{0,r,\varepsilon}(dx) g_\varepsilon(1, y-x)} \right) \mu(dxdy; g_\varepsilon(1), p(x) dx, P_1), \end{aligned}$$

$$\begin{aligned} & V_\varepsilon(p(x) dx, P_1) \\ &= - \int_{\mathbb{R}^d \times \mathbb{R}^d} \log \left( \frac{p(x) dx g_\varepsilon(1, y-x)}{\mu(dxdy; g_\varepsilon(1), p(x) dx, P_1)} \right) \mu(dxdy; g_\varepsilon(1), p(x) dx, P_1), \end{aligned}$$

and by Jensen's inequality,

$$\begin{aligned} & \int_{\mathbb{R}^d \times \mathbb{R}^d} \log \left( \frac{\mu(dxdy; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1)}{P_{0,r,\varepsilon}(dx) g_\varepsilon(1, y-x)} \frac{p(x) dx g_\varepsilon(1, y-x)}{\mu(dxdy; g_\varepsilon(1), p(x) dx, P_1)} \right) \\ & \quad \times \mu(dxdy; g_\varepsilon(1), p(x) dx, P_1) \leq 0. \end{aligned}$$

(2.11) holds since  $\varphi(0, x; f_{0,r,\varepsilon}) - u_1(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1)$  is a constant  $C$  (see (1.19)) and since, for  $p(x) dx \in \mathcal{P}(B_r)$ ,

$$\begin{aligned} & S(p(x) dx) + \int_{\mathbb{R}^d} \left( \varepsilon \varphi(0, x; f_{0,r,\varepsilon}) + \frac{|x|^2}{2} \right) p(x) dx \\ &= \int_{B_r} p(x) dx \log \frac{p(x)}{C_\varepsilon^{-1} I_{B_r}(x) \exp(-\varepsilon u_1(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) - (1/2)|x|^2)} - \log C_\varepsilon + C \\ &\geq -\log C_\varepsilon + C. \end{aligned}$$

Here

$$C_\varepsilon := \int_{B_r} \exp \left( -\varepsilon u_1(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) - \frac{1}{2}|x|^2 \right) dx,$$

and the equality holds if and only if

$$p(x) = \frac{1}{C_\varepsilon} I_{B_r}(x) \exp \left( -\varepsilon u_1(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) - \frac{1}{2}|x|^2 \right).$$

We prove the second part of Theorem 2.2. For  $i = 1, 2$ ,  $m \geq 1$  and  $x \in \mathbb{R}^d$ ,

$$\begin{aligned} \nu_{i,\varepsilon}(dx) &:= \nu_i(dx; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1), \\ \mu_\varepsilon(dxdy) &:= \nu_{1,\varepsilon}(dx) g_\varepsilon(1, y-x) \nu_{2,\varepsilon}(dy), \\ u_{i|m,\varepsilon}(x) &:= u_{i|m}(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1) \end{aligned} \tag{4.11}$$

(see (2.1) for notation). Since  $\mathcal{P}(B_r)$  is compact,  $\{P_{0,r,\varepsilon}\}_{\varepsilon>0}$  and  $\{\mu_\varepsilon\}_{\varepsilon>0}$  has a weakly convergent subsequence by Prohorov's theorem in the same way as in the proof of Theorem 2.1 (see [6]). Let  $P_0$  and  $\mu$  denote the weak limit along the same subsequence, as

$\varepsilon \rightarrow 0$ , of  $P_{0,r,\varepsilon}$  and  $\mu_\varepsilon$ , respectively. For sufficiently large  $m \geq 1$ , by the diagonal method,  $\bar{u}_{1|m,\varepsilon}(x) + \bar{u}_{2|m,\varepsilon}(y)$  has a subsequence which is uniformly convergent, as  $\varepsilon \rightarrow 0$ , on every compact subset of  $\mathbb{R}^d \times \mathbb{R}^d$  (see (3.10) for notation). Indeed, for sufficiently large  $m \geq 1$  and small  $\varepsilon > 0$ ,  $\bar{u}_{i|m,\varepsilon}$ ,  $i = 1, 2$  are convex from Lemma 3.3, and  $\bar{u}_{1|m,\varepsilon}(x) + \bar{u}_{2|m,\varepsilon}(y)$  is uniformly bounded on every compact subset of  $\mathbb{R}^d \times \mathbb{R}^d$ , from (3.8) in Lemma 3.6:

$$\begin{aligned} & \left| \bar{u}_{1|m,\varepsilon}(x) + \bar{u}_{2|m,\varepsilon}(y) + \varepsilon \log(2\pi\varepsilon)^{d/2} \right| \\ & \leq (m+1)^2 + (m+1)(|x| + |y|) - \varepsilon \log \mu_\varepsilon(B_m \times B_m), \quad x, y \in \mathbb{R}^d. \end{aligned} \quad (4.12)$$

Let  $\{\varepsilon_n\}_{n \geq 1}$  denote a sequence which converges to 0, as  $n \rightarrow \infty$  and along which the above sequences are all convergent.

$$u_m(x, y) := \lim_{n \rightarrow \infty} (\bar{u}_{1|m,\varepsilon_n}(x) + \bar{u}_{2|m,\varepsilon_n}(y)), \quad m \geq 1, \quad x, y \in \mathbb{R}^d. \quad (4.13)$$

There exists the limit

$$u(x, y) := \lim_{m \rightarrow \infty} u_m(x, y), \quad x, y \in \mathbb{R}^d. \quad (4.14)$$

Indeed,  $m \mapsto u_m$  is nondecreasing since

$$\begin{aligned} & \exp(u_{1|m,\varepsilon}(x_1) + u_{2|m,\varepsilon}(x_2)) \\ & = \int_{S \times S} \frac{g_\varepsilon(1, x_1 - y) g_\varepsilon(1, x - x_2)}{g_\varepsilon(1, x - y)} \varphi_m(x) \varphi_m(y) \mu_\varepsilon(dx dy). \end{aligned} \quad (4.15)$$

From the last statement of Lemma 3.7, there exists  $x_0 \in B_r$  such that  $u(x_0, y_0) < \infty$ , since

$$\bar{u}_{1,\varepsilon_n}(x) + \bar{u}_{2,\varepsilon_n}(y_0) \geq \bar{u}_{1|m,\varepsilon_n}(x) + \bar{u}_{2|m,\varepsilon_n}(y_0).$$

To complete the proof of Theorem 2.2, we show that the following holds:

$$\langle x, y \rangle = u(x, y), \quad \mu\text{-a.s.}, \quad (4.16)$$

$$x = D_y u(x_0, y), \quad y = D_x u(x, y_0), \quad \mu\text{-a.s.}, \quad (4.17)$$

$$P_0(dx) = \frac{I_D(x)}{C} \exp(-u(x, y_0)) dx, \quad (4.18)$$

where  $D$  is a convex subset of  $B_r$  and  $C$  is a normalizing constant. Notice that  $u(x, y)$  is convex and is differentiable almost everywhere on its domain.

PROOF OF (4.16). The following implies that (4.16) holds: for sufficiently large  $m \geq r$ ,

$$\langle x, y \rangle = u_m(x, y), \quad (x, y) \in \mathbb{R}^d \times \text{Int}(\text{supp}(\varphi_m)), \quad \mu\text{-a.s.} \quad (4.19)$$

Indeed, from (4.14) and (4.19), for sufficiently large  $m > r$ ,

$$\langle x, y \rangle = u_m(x, y) = u_{m'}(x, y) = u(x, y), \quad m' > m,$$

$(x, y) \in \mathbb{R}^d \times \text{Int}(\text{supp}(\varphi_m)) \subset \mathbb{R}^d \times \text{Int}(\text{supp}(\varphi_{m'}))$ ,  $\mu$ -a.s. To prove (4.19), we first prove that the following holds: for sufficiently large  $m \geq r$ ,

$$\langle x, y \rangle \leq u_m(x, y), \quad (x, y) \in \text{supp}(P_0) \times (\text{supp}(P_1) \cap \text{Int}(\text{supp}(\varphi_m))). \quad (4.20)$$

For  $i \neq j$ ,  $i, j = 1, 2$ ,

$$\mu_{i|m,\varepsilon}(dx_i) := \nu_{i,\varepsilon}(dx_i) \exp(u_{i|m,\varepsilon}(x_i)) = \int_{\{x_j \in \mathbb{R}^d\}} \varphi_m(x_j) \mu_\varepsilon(dx_1 dx_2). \quad (4.21)$$

Then for  $\delta > 0$  and  $(x, y) \in \text{supp}(P_0) \times (\text{supp}(P_1) \cap \text{Int}(\text{supp}(\varphi_m)))$ ,

$$\begin{aligned} & \exp\left(\frac{\bar{u}_{1|m,\varepsilon}(x) + \bar{u}_{2|m,\varepsilon}(y)}{\varepsilon}\right) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{(2\pi\varepsilon)^d} \exp\left(\frac{\langle x_1, y \rangle + \langle x, x_2 \rangle - \bar{u}_{1,\varepsilon|m}(x_1) - \bar{u}_{2,\varepsilon|m}(x_2)}{\varepsilon}\right) \\ & \quad \times \varphi_m(x_1) \varphi_m(x_2) \mu_{1|m,\varepsilon}(dx_1) \mu_{2|m,\varepsilon}(dx_2) \\ & \geq \int_{U_\delta(x) \times U_\delta(y)} \frac{1}{(2\pi\varepsilon)^d} \exp\left(\frac{\langle x_1, y \rangle + \langle x, x_2 \rangle - \bar{u}_{1,\varepsilon|m}(x_1) - \bar{u}_{2,\varepsilon|m}(x_2)}{\varepsilon}\right) \\ & \quad \times \varphi_m(x_2) \mu_{1|m,\varepsilon}(dx_1) p_1(x_2) dx_2 \end{aligned} \quad (4.22)$$

(see (4.15)). Indeed, for  $m > r$ ,  $\mu_{1|m,\varepsilon}(dx)$  is supported on  $B_r$  since  $P_{0,r,\varepsilon} \in \mathcal{P}(B_r)$  and

$$\mu_{2|m,\varepsilon}(dy) = \left( \int_{B_r} g_\varepsilon(1, y - x) \varphi_m(x) \nu_{1,\varepsilon}(dx) \right) \nu_{2,\varepsilon}(dy) = p_1(y) dy.$$

(4.22) implies (4.20) since

$$\begin{aligned} & \int_{U_\delta(y)} \varphi_m(x_2) p_1(x_2) dx_2 > 0, \\ & \liminf_{n \rightarrow \infty} \mu_{1|m,\varepsilon_n}(U_\delta(x)) \geq \int_{U_\delta(x) \times \mathbb{R}^d} \varphi_m(x_2) \mu(dx_1 dx_2) \\ & \geq P_0(U_\delta(x)) - P_1(B_m^c) > 0, \quad \text{for sufficiently large } m. \end{aligned}$$

Next we prove that the following holds: for sufficiently large  $m \geq 1$ ,

$$\langle x, y \rangle \geq u_m(x, y), \quad \mu\text{-a.s.} \quad (4.23)$$

$$A_{m,\delta,k} := \{(x, y) \in B_r \times U_k(0) \mid \langle x, y \rangle - u_m(x, y) < -\delta\}, \quad \delta > 0, \quad k \geq 1.$$

Then  $A_{m,\delta,k}$  is open since  $u_m$  is convex and finite (see (4.12)–(4.13)) and is continuous. The following implies that (4.23) is true: from (4.13), for sufficiently large  $m \geq 1$ ,

$$\mu(A_{m,\delta,k}) \leq \liminf_{n \rightarrow \infty} \mu_{\varepsilon_n}(A_{m,\delta,k}), \quad (4.24)$$

$$\begin{aligned} \mu_{\varepsilon_n}(A_{m,\delta,k}) &= \int_{A_{m,\delta,k}} \frac{1}{(2\pi\varepsilon_n)^{d/2}} \exp\left(\frac{\langle x, y \rangle - \bar{u}_{1|m,\varepsilon_n}(x) - \bar{u}_{2|m,\varepsilon_n}(y)}{\varepsilon_n}\right) \\ &\quad \times \mu_{1|m,\varepsilon_n}(dx) \mu_{2|m,\varepsilon_n}(dy) \\ &\rightarrow 0, \quad n \rightarrow \infty. \end{aligned} \quad \square$$

PROOF OF (4.17). For  $(x, y) \in \text{supp}(P_0) \times \text{supp}(P_1)$ ,

$$\langle x, y \rangle \leq u(x, y) = u(x, y_0) + u(x_0, y) - u(x_0, y_0). \quad (4.25)$$

Indeed, from (4.14) and (4.20), for sufficiently large  $m > r$  such that  $y \in \text{Int}(\text{supp}(\varphi_m))$ ,

$$\langle x, y \rangle \leq u_m(x, y) \leq u(x, y).$$

(4.14) and the following imply (4.25): from (4.13),

$$u_m(x, y) = u_m(x, y_0) + u_m(x_0, y) - u_m(x_0, y_0).$$

$$A := \{(x, y) \in \mathbb{R}^d \times \mathbb{R}^d \mid \langle x, y \rangle = u(x, y)\}.$$

$u(x, y_0)$  and  $u(x_0, y)$  are finite for  $(x, y) \in A$  from (4.12), since from (4.14) and the equality in (4.25),

$$\begin{aligned} \langle x, y \rangle &= u(x, y) \\ &\geq \max(u(x, y_0) + u_m(x_0, y) - u(x_0, y_0), u_m(x, y_0) + u(x_0, y) - u(x_0, y_0)). \end{aligned}$$

For a set  $B \subset \mathbb{R}^d$  and a function  $f : B \rightarrow \mathbb{R}$ ,

$$\begin{aligned} \text{co } B &:= \left\{ \sum_{i=1}^{d+1} \lambda_i x_i \mid \sum_{i=1}^{d+1} \lambda_i = 1, \lambda_i \geq 0, x_i \in B, 1 \leq i \leq d+1 \right\}, \\ \text{con } f(x) &:= \begin{cases} \inf \left\{ \sum_{i=1}^{d+1} \lambda_i f(x_i) \mid x = \sum_{i=1}^{d+1} \lambda_i x_i, \sum_{i=1}^{d+1} \lambda_i = 1, \right. \\ \quad \left. \lambda_i \geq 0, x_i \in B, 1 \leq i \leq d+1 \right\}, & x \in \text{co } B, \\ \infty, & x \notin \text{co } B. \end{cases} \end{aligned}$$

Then, from (4.25), for  $x \in \text{supp}(P_0)$ ,

$$\begin{aligned} u(x, y_0) - u(x_0, y_0) &\geq \sup\{\langle x, y \rangle - u(x_0, y) \mid y \in \text{supp}(P_1)\} \\ &= \sup\{\langle x, y \rangle - \text{con}(u|_{\text{supp}(P_1)})(x_0, y) \mid y \in \mathbb{R}^d\}. \end{aligned} \quad (4.26)$$

Here  $(u|_{\text{supp}(P_1)})(x_0, y)$  denotes the restriction of  $u(x_0, y)$  on  $\text{supp}(P_1)$  and the equality holds if  $(x, y_x) \in A$  for some  $y_x \in \text{supp}(P_1)$ , in which case  $x \in \partial_y \text{con}(u|_{\text{supp}(P_1)})(x_0, y_x)$ , where for a function  $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{\infty\}$ ,

$$\partial_y f(y) := \{x \in \mathbb{R}^d \mid f(z) \geq f(y) + \langle x, z - y \rangle, \text{ for any } z \in \mathbb{R}^d\}.$$

In particular,  $x \in \partial_y \operatorname{con}(u|_{\operatorname{supp}(P_1)})(x_0, y)$ ,  $\mu$ -a.s. from (4.16).  $x = D_y u(x_0, y)$ ,  $\mu$ -a.s. since

$$\begin{aligned} \operatorname{con}(u|_{\operatorname{supp}(P_1)})(x_0, y) &= u(x_0, y), \quad y \in \operatorname{supp}(P_1), \\ \partial_y \operatorname{con}(u|_{\operatorname{supp}(P_1)})(x_0, y) &= \{D_y u(x_0, y)\}, \quad dy\text{-a.e. on } \operatorname{supp}(P_1) \end{aligned}$$

and since  $P_1(dx)$  has a probability density function. In the same way, one can show that  $y = D_x u(x, y_0)$ ,  $\mu$ -a.s.  $\square$

PROOF OF (4.18).

$$D_{R,\varepsilon} := \{x \in B_r \mid \bar{u}_{1,\varepsilon}(x) + \bar{u}_{2,\varepsilon}(y_0) \leq R\}, \quad R, \varepsilon > 0$$

(see (3.10) for notation). Then, from Lemma 3.7,

$$\lim_{R \rightarrow \infty} \limsup_{\varepsilon \rightarrow 0} \int_{D_{R,\varepsilon}^c} p_{0,r,\varepsilon}(x) dx = 0. \quad (4.27)$$

Indeed,

$$\begin{aligned} \int_{D_{R,\varepsilon}^c} p_{0,r,\varepsilon}(x) dx &= \frac{1}{C_\varepsilon \exp(-\bar{u}_{2,\varepsilon}(y_0))} \int_{B_r \cap D_{R,\varepsilon}^c} \exp(-\bar{u}_{1,\varepsilon}(x) - \bar{u}_{2,\varepsilon}(y_0)) dx \\ &\leq \exp(-R) \frac{\operatorname{Vol}(B_r)}{C_\varepsilon \exp(-\bar{u}_{2,\varepsilon}(y_0))}. \end{aligned}$$

For  $\delta > 0$ ,

$$\psi_{\delta,R,\varepsilon}(x) := \max\left(0, 1 - \frac{\operatorname{dist}(x, D_{R,\varepsilon})}{\delta}\right).$$

Then, from Lemma 3.8, there exists a convergent subsequence  $\{\psi_{\delta,R,\varepsilon_{n_k}}(x)\}_{k \geq 1}$  in  $C(B_r)$  and a closed convex set  $D_{R,0} \subset B_r$  such that

$$\lim_{k \rightarrow \infty} \|\psi_{\delta,R,\varepsilon_{n_k}} - \psi_{\delta,R,0}\|_{\infty,r} = 0. \quad (4.28)$$

$$D := \bigcup_{R>0} D_{R,0}.$$

Then we prove that the following holds: for a closed set  $B \subset B_r$ ,

$$\begin{aligned} \lim_{R \rightarrow \infty} \limsup_{k \rightarrow \infty} \int_{B \cap D_{R,\varepsilon_{n_k}}} p_{0,r,\varepsilon_{n_k}}(x) dx \\ \leq \frac{1}{\int_D \exp(-u(x, y_0)) dx} \int_{B \cap D} \exp(-u(x, y_0)) dx. \end{aligned} \quad (4.29)$$

The proof of (4.29) is done by the following (4.30)–(4.31) which will be proved later.

$$\begin{aligned} & \lim_{R \rightarrow \infty} \limsup_{k \rightarrow \infty} \int_{B \cap D_{R, \varepsilon_{n_k}}} \exp(-\bar{u}_{1, \varepsilon_{n_k}}(x) - \bar{u}_{2, \varepsilon_{n_k}}(y_0)) dx \\ & \leq \int_{B \cap D} \exp(-u(x, y_0)) dx, \end{aligned} \quad (4.30)$$

$$\liminf_{k \rightarrow \infty} C_{\varepsilon_{n_k}} \exp(-\bar{u}_{2, \varepsilon_{n_k}}(y_0)) \geq \int_D \exp(-u(x, y_0)) dx. \quad (4.31)$$

Notice that, from (4.13)–(4.14) and Lemma 3.7, the following holds:

$$\begin{aligned} & \int_D \exp(-u(x, y_0)) dx \\ & > \exp\left(-\int_{\mathbb{R}^d} \frac{1}{2} |x|^2 P_1(dx) + \log\{\text{Vol}(B_r)\} - \frac{1}{2} \int_{B_r} |x|^2 \frac{dx}{\text{Vol}(B_r)}\right) > 0. \end{aligned}$$

We prove (4.30). For sufficiently large  $m \geq 1$ ,

$$\int_{B \cap D_{R, \varepsilon}} \exp(-\bar{u}_{1, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx \leq \int_B \psi_{\delta, R, \varepsilon}(x) \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx$$

(see (3.10) for notation). Let  $\psi_\delta$  denote the function  $\psi_{\delta, R, 0}$  with  $D_{R, 0}$  replaced by  $D$ . Then for  $m > r$ , from (4.13) and (4.28),

$$\begin{aligned} & \lim_{k \rightarrow \infty} \int_B \psi_{\delta, R, \varepsilon_{n_k}}(x) \exp(-\bar{u}_{1|m, \varepsilon_{n_k}}(x) - \bar{u}_{2, \varepsilon_{n_k}}(y_0)) dx \\ & = \int_B \psi_{\delta, R, 0}(x) \exp(-u_m(x, y_0)) dx \rightarrow \int_B \psi_\delta(x) \exp(-u(x, y_0)) dx, \quad m, R \rightarrow \infty, \\ & \rightarrow \int_B I_D(x) \exp(-u(x, y_0)) dx, \quad \delta \rightarrow 0, \end{aligned} \quad (4.32)$$

since  $R \mapsto D_{R, \varepsilon}$  is nondecreasing.

We prove (4.31).

$$\tilde{D}_{\delta, m, \varepsilon} := \{x \in B_r \mid \bar{u}_{1, \varepsilon}(x) - \bar{u}_{1|m, \varepsilon}(x) < \delta\}, \quad \delta > 0.$$

Then

$$\begin{aligned} & C_\varepsilon \exp(-\bar{u}_{2, \varepsilon}(y_0)) \\ & \geq \int_{B_r} \psi_{\delta, R, \varepsilon}(x) \exp(-\bar{u}_{1, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx \\ & \geq \exp(-\delta) \int_{\tilde{D}_{\delta, m, \varepsilon}} \psi_{\delta, R, \varepsilon}(x) \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx \\ & = \exp(-\delta) \int_{B_r} \psi_{\delta, R, \varepsilon}(x) \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx \\ & \quad - \exp(-\delta) \int_{B_r \cap \tilde{D}_{\delta, m, \varepsilon}^c} \psi_{\delta, R, \varepsilon}(x) \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx. \end{aligned} \quad (4.33)$$

From (4.32), we only have to prove that the following holds:

$$\begin{aligned} \lim_{\delta \rightarrow 0} \lim_{R \rightarrow \infty} \limsup_{m \rightarrow \infty} \limsup_{k \rightarrow \infty} \int_{B_r \cap \tilde{D}_{\delta, m, \varepsilon_{n_k}}^c} \psi_{\delta, R, \varepsilon_{n_k}}(x) \\ \times \exp(-\bar{u}_{1|m, \varepsilon_{n_k}}(x) - \bar{u}_{2, \varepsilon_{n_k}}(y_0)) dx = 0. \end{aligned} \quad (4.34)$$

$$\begin{aligned} & \int_{B_r \cap \tilde{D}_{\delta, m, \varepsilon}^c} \psi_{\delta, R, \varepsilon}(x) \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx \\ & \leq \int_{B_r \cap U_\delta(D_{R, \varepsilon}) \cap D_{R, \varepsilon}^c} \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx \\ & \quad + \int_{B_r \cap \tilde{D}_{\delta, m, \varepsilon}^c \cap D_{R, \varepsilon}} \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx, \end{aligned}$$

since  $\psi_{\delta, R, \varepsilon}^{-1}((0, 1]) = U_\delta(D_{R, \varepsilon})$ . For any  $\gamma > 0$ , sufficiently large  $m \geq m_0 \geq 1$  and  $k$ , from Lemma 3.8 and (4.13),

$$\begin{aligned} & \int_{B_r \cap U_\delta(D_{R, \varepsilon_{n_k}}) \cap D_{R, \varepsilon_{n_k}}^c} \exp(-\bar{u}_{1|m, \varepsilon_{n_k}}(x) - \bar{u}_{2, \varepsilon_{n_k}}(y_0)) dx \\ & \leq \int_{B_r \cap U_{\delta+\gamma}(D_{R, 0}) \cap U_{-\gamma}(D_{R, 0})^c} \exp(-\bar{u}_{1|m_0, \varepsilon_{n_k}}(x) - \bar{u}_{2, \varepsilon_{n_k}}(y_0)) dx \\ & \rightarrow \int_{B_r \cap U_{\delta+\gamma}(D_{R, 0}) \cap U_{-\gamma}(D_{R, 0})^c} \exp(-u_{m_0}(x, y_0)) dx, \quad k \rightarrow \infty, \\ & \rightarrow \int_{B_r \cap U_{\delta+\gamma}(D) \cap U_{-\gamma}(D)^c} \exp(-u_{m_0}(x, y_0)) dx, \quad R \rightarrow \infty, \\ & \rightarrow 0, \quad \delta, \gamma \rightarrow 0. \end{aligned}$$

$$\begin{aligned} & \int_{B_r \cap \tilde{D}_{\delta, m, \varepsilon}^c \cap D_{R, \varepsilon}} \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0)) dx \\ & \leq \int_{\tilde{D}_{\delta, m, \varepsilon}^c} \exp(-\bar{u}_{1|m, \varepsilon}(x) - \bar{u}_{2, \varepsilon}(y_0) + R - \bar{u}_{2, \varepsilon}(y_0)) C_\varepsilon p_{0, r, \varepsilon}(x) dx \\ & \leq \exp(-2 \inf\{\bar{u}_{1|m_0, \varepsilon}(x) + \bar{u}_{2, \varepsilon}(y_0) | x \in B_r\} + R) \text{Vol}(B_r) \int_{\tilde{D}_{\delta, m, \varepsilon}^c} p_{0, r, \varepsilon}(x) dx, \end{aligned} \quad (4.35)$$

$$\begin{aligned} & \int_{\tilde{D}_{\delta, m, \varepsilon}^c} p_{0, r, \varepsilon}(x) dx = \int_{\tilde{D}_{\delta, m, \varepsilon}^c \times \mathbb{R}^d} \mu_\varepsilon(dx dy) \leq \mu_{1|m, \varepsilon}(\tilde{D}_{\delta, m, \varepsilon}^c) + P_1(B_m^c) \\ & \leq \int_{\tilde{D}_{\delta, m, \varepsilon}^c} \exp\left(\frac{\bar{u}_{1, \varepsilon}(x) - \bar{u}_{1|m, \varepsilon}(x) - \delta}{\varepsilon}\right) \mu_{1|m, \varepsilon}(dx) + P_1(B_m^c) \\ & \leq \exp\left(-\frac{\delta}{\varepsilon}\right) \int_{B_r} p_{0, r, \varepsilon}(x) dx + P_1(B_m^c) \rightarrow P_1(B_m^c), \quad \varepsilon \rightarrow 0, \\ & \rightarrow 0, \quad m \rightarrow \infty. \end{aligned} \quad (4.36)$$

Here, from (3.10) and (4.21) (see also (1.1)),

$$\exp\left(\frac{\bar{u}_{1,\varepsilon}(x) - \bar{u}_{1|m,\varepsilon}(x)}{\varepsilon}\right) \mu_{1|m,\varepsilon}(dx) = \exp(u_1(x; g_\varepsilon(1), P_{0,r,\varepsilon}, P_1)) \nu_{1,\varepsilon}(dx) = p_{0,r,\varepsilon}(x) dx.$$

(4.12) and (4.13) complete the proof of (4.34).

If  $P_1$  is compactly supported, then  $u_{i|m,\varepsilon} = u_{i|m',\varepsilon}$  and  $u(x, y) = u_{m'}(x, y)$  for  $m' \geq m$ , provided  $B_r \cup \text{supp}(P_1) \subset B_m$ . (4.11)–(4.13) imply that the last statement of Theorem 2.2 holds.  $\square$

ACKNOWLEDGEMENTS. The author would like to thank an anonymous referee for useful suggestions.

## References

- [ 1 ] S. Adams, N. Dirr, M. A. Peletier and J. Zimmer, From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage, *Comm. Math. Phys.*, **307** (2011), 791–815.
- [ 2 ] J. Backhoff, G. Conforti, I. Gentil and C. Léonard, The mean field Schrödinger problem: ergodic behavior, entropy estimates and functional inequalities, arXiv:1905.02393v1.
- [ 3 ] I. J. Bakelman, *Convex Analysis and Nonlinear Geometric Elliptic Equations*, Springer-Verlag, 1994.
- [ 4 ] S. Bernstein, Sur les liaisons entre les grandeurs aléatoires, *Verhandlungen des Internationalen Math.-Kongress, Zurich*, **1** (1932), 288–309.
- [ 5 ] A. Beurling, An automorphism of product measures, *Ann. of Math. (2)*, **72** (1960), 189–200.
- [ 6 ] P. Billingsley, *Convergence of Probability Measures*, Wiley Ser. Probab. Stat., A Wiley-Interscience Publ., John Wiley & Sons, 1999.
- [ 7 ] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs, *C. R. Acad. Sci. Paris Sér. I Math.*, **305** (1987), 805–808.
- [ 8 ] Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.*, **44** (1991), 375–417.
- [ 9 ] Y. Chen, T. Georgiou and M. Pavon, Entropic and displacement interpolation: a computational approach using the Hilbert metric, *SIAM J. Appl. Math.*, **76** (2016), 2375–2396.
- [ 10 ] M. Colombo, S. Di Marino and F. Stra, Continuity of multimarginal optimal transport with repulsive cost, *SIAM J. Math. Anal.*, **51** (2019), 2903–2926.
- [ 11 ] G. Conforti, A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost, *Probab. Theory Related Fields*, **174** (2019), 1–47.
- [ 12 ] G. Conforti and L. Ripani, Around the entropic Talagrand inequality, *Bernoulli*, **26** (2020), 1431–1452.
- [ 13 ] D. Cordero-Erausquin and B. Klartag, Moment measures, *J. Funct. Anal.*, **268** (2015), 3834–3866.
- [ 14 ] P. Dai Pra, A stochastic control approach to reciprocal diffusion processes, *Appl. Math. Optim.*, **23** (1991), 313–329.
- [ 15 ] J. L. Doob, Conditional Brownian motion and the boundary limits of harmonic functions, *Bull. Soc. Math. France*, **85** (1957), 431–458.
- [ 16 ] M. H. Duong, V. Laschos and M. Renger, Wasserstein gradient flows from large deviations of many-particle limits, *ESAIM Control Optim. Calc. Var.*, **19** (2013), 1166–1188.
- [ 17 ] P. Dupuis and R. S. Ellis, *A Weak Convergence Approach to the Theory of Large Deviations*, Wiley Ser. Probab. Stat., A Wiley-Interscience Publ., John Wiley & Sons, 1997.
- [ 18 ] M. Erbar, J. Maas and M. Renger, From large deviations to Wasserstein gradient flows in multiple dimensions, *Electron. Commun. Probab.*, **20** (2015), no. 89, pp. 12.
- [ 19 ] J. Feng and T. Nguyen, Hamilton–Jacobi equations in space of measures associated with a system of conservation laws, *J. Math. Pures Appl.*, **97** (2012), 318–390.
- [ 20 ] W. H. Fleming and H. M. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed., *Stoch. Model. Appl. Probab.*, **25**, Springer, 2006.

- [21] H. Föllmer, Random fields and diffusion processes, In: *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, (ed. P.-L. Hennequin), *Lecture Notes in Math.*, **1362**, Springer, 2006, 101–203.
- [22] R. Fortet, Résolution d'un Système d'équations de M. Schrödinger, *J. Math. Pures Appl.*, **IX** (1940), 83–105.
- [23] I. Gentil, C. Léonard, L. Ripani and L. Tamanini, An entropic interpolation proof of the HWI inequality, *Stochastic Process. Appl.*, **130** (2020), 907–923.
- [24] B. Jamison, Reciprocal processes, *Z. Wahrsch. Verw. Gebiete*, **30** (1974), 65–86.
- [25] B. Jamison, The Markov processes of Schrödinger, *Z. Wahrsch. Verw. Gebiete*, **32** (1975), 323–331.
- [26] C. Léonard, From the Schrödinger problem to the Monge–Kantorovich problem, *J. Funct. Anal.*, **262** (2012), 1879–1920.
- [27] C. Léonard, A survey of the Schrödinger problem and some of its connections with optimal transport, Special Issue on Optimal Transport and Applications, *Discrete Contin. Dyn. Syst.*, **34** (2014), 1533–1574.
- [28] T. Mikami, Monge's problem with a quadratic cost by the zero-noise limit of  $h$ -path processes, *Probab. Theory Related Fields*, **129** (2004), 245–260.
- [29] T. Mikami, Semimartingales from the Fokker–Planck equation, *Appl. Math. Optim.*, **53** (2006), 209–219.
- [30] T. Mikami, Marginal problem for semimartingales via duality, In: *International Conference for the 25th Anniversary of Viscosity Solutions*, (eds. Y. Giga, K. Ishii, S. Koike et al.), *GAKUTO Internat. Ser. Math. Sci. Appl.*, **30**, Gakkotosho, 2008, 133–152.
- [31] T. Mikami, Optimal transportation problem as stochastic mechanics, In: *Selected Papers on Probability and Statistics*, *Amer. Math. Soc. Transl. Ser. 2*, **227**, Amer. Math. Soc., 2009, 75–94.
- [32] T. Mikami, Two end points marginal problem by stochastic optimal transportation, *SIAM J. Control Optim.*, **53** (2015), 2449–2461.
- [33] T. Mikami, Regularity of Schrödinger's functional equation and mean field PDEs for  $h$ -path processes, *Osaka J. Math.*, **56** (2019), 831–842.
- [34] T. Mikami and M. Thieullen, Duality theorem for the stochastic optimal control problem, *Stochastic Process. Appl.*, **116** (2006), 1815–1835.
- [35] S. Pal, On the difference between entropic cost and the optimal transport cost, arXiv:1905.12206v1.
- [36] S. T. Rachev and L. Rüschendorf, *Mass Transportation Problems. Vol. I: Theory*, *Probab. Appl. (N.Y.)*, Springer-Verlag, 1998.
- [37] S. T. Rachev and L. Rüschendorf, *Mass Transportation Problems. Vol. II: Applications*, *Probab. Appl. (N.Y.)*, Springer-Verlag, 1998.
- [38] L. Ripani, Convexity and regularity properties for entropic interpolations, *J. Funct. Anal.*, **277** (2019), 368–391.
- [39] L. Rüschendorf and W. Thomsen, Note on the Schrödinger equation and  $I$ -projections, *Statist. Probab. Lett.*, **17** (1993), 369–375.
- [40] F. Santambrogio, Dealing with moment measures via entropy and optimal transport, *J. Funct. Anal.*, **271** (2016), 418–436.
- [41] E. Schrödinger, Über die Umkehrung der Naturgesetze, *Sitz.-Ber. der Preuss. Akad. Wiss., Berlin, Phys. Math.*, (1931), 144–153.
- [42] E. Schrödinger, Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique, *Ann. Inst. H. Poincaré*, **2** (1932), 269–310.
- [43] X. Tan and N. Touzi, Optimal transportation under controlled stochastic dynamics, *Ann. Probab.*, **41** (2013), 3201–3240.
- [44] C. Villani, *Optimal Transport: Old and New*, *Grundlehren Math. Wiss.*, **338**, Springer-Verlag, Berlin, 2009.
- [45] J.-C. Zambrini, Variational processes, In: *Stochastic Processes in Classical and Quantum Systems*, Ascona, 1985, (eds. S. Albeverio, G. Casati, D. Merlini), *Lecture Notes in Phys.*, **262**, Springer, 1986, 517–529.

Toshio MIKAMI  
Department of Mathematics  
Tsuda University  
2-1-1 Tsuda-machi, Kodaira  
Tokyo 187-8577, Japan  
E-mail: t.mikami@tsuda.ac.jp