VARIATIONAL SOLUTIONS OF THE MONGE TRANSPORT PROBLEM AND THE MONGE-AMPÈRE EQUATION IN ABSTRACT WIENER SPACE

A. S. ÜSTÜNEL

Abstract: Let \((W, H, \mu)\) be an abstract Wiener space, assume that \(T = I_W + \nabla \varphi\) is the solution of the Monge problem associated to the measures \(d\mu\) and \(d\nu = Ld\mu = e^{-f}d\mu\). Under the finite information hypothesis, using a variational method, we prove that the forward potential of the Monge-Kantorovitch problem satisfies the “structure equation"

\[
\delta((I_H + \nabla^2 \varphi)^{-1} - I_H) = \nabla \varphi + \nabla f \circ T
\]

and with it the Sobolev regularity of the backward Monge potential is proven map. A similar structure equations also holds for the forward Monge potential and it implies the regularity of it for \(1 - \varepsilon\) log-concave densities. We show that \(L = e^{-f}\) can be represented as

\[
L = \det_2(I + \nabla^2 \psi) \exp \left[ -L \psi - \frac{1}{2} |\nabla \psi|_H^2 \right],
\]

where \(\psi\) is the backward Monge potential. Moreover the forward potential gives the solution of the Monge-Ampère equation:

\[
L \circ T \Lambda = 1
\]

\(\mu\text{-a.s.},\) where

\[
\Lambda = \det_2(I + \nabla^2_a \varphi) \left[ -L^a \varphi - \frac{1}{2} |\varphi|_H^2 \right],
\]

and \(\nabla^2_a \varphi\) is the Radon-Nikodym derivative of the Hilbert-Schmidt-valued absolutely continuous part of the vector measure \(\nabla^2 \varphi\). In particular, for \(L \neq 0\) almost surely, the Girsanov identity holds:

\[
\int_W g \circ T \Lambda(\varphi) d\mu = \int_W g d\mu,
\]

for any \(g \in C_b(W)\).

Contents

1. Introductions 2
2. Preliminaries 4
3. Variational calculations 6
4. Approximations of the Monge Potentials 8
5. Sobolev regularity 19
6. Disintegration Results 23
7. Calculation of the Jacobians and Monge-Ampère Equation 26
8. Regularity of the forward potential 31
Keywords: Entropy, adapted perturbation of identity, Wiener measure, Monge and Monge Kantorovich problems, Monge potential, Monge-Brenier map, Monge-Ampère equation.

1. Introductions

Let $\nu$ be the probability measure defined by

$$d\nu = \frac{1}{c} e^{-f} d\mu$$

such that the relative entropy of $\nu$ w.r.t. the Wiener measure $\mu$, denoted as $H(\nu|\mu)$ is finite. Let $\Sigma(\mu, \nu)$ be the set of the probability measures on $(W \times W, \mathcal{B}(W \times W))$ whose first marginals are $\mu$ and the second ones are $\nu$. Consider the problem of minimization which defines also a strong Wasserstein distance between $\mu$ and $\nu$:

$$\inf \left( \int_{W \times W} |x - y|^2_H d\beta(x, y) : \beta \in \Sigma(\mu, \nu) \right) = d^2(\mu, \nu),$$

where $| \cdot |_H$ denotes the Cameron-Martin norm. In the finite dimensional case this problem has been extensively studied since almost three centuries and we refer to the texts [16] and [23] for history and references and also to [2] and [15].

In the infinite dimensional case, where the cost function is very singular, in the sense that the set on which the cost function is finite has zero measure w.r.t. the product measure $\mu \times \nu$ has been solved in a series of papers ([8, 9, 10]) and the answer can be summarized as follows: There exists a 1-convex function $\varphi : W \rightarrow \mathbb{R}$, in the Gaussian Sobolev space $\mathcal{D}_{2,1}$, called Monge potential or Monge-Brenier map such that the above infimum is attained at $\gamma = (I_W \times T)\mu$, i.e., the image of the measure $\mu$ under the map $I_W \times T$, where $T = I_W + \nabla \varphi$, where $\nabla \varphi$ is the $L^2(\mu)$-extended derivative of $\varphi$ in the direction of Cameron-Martin space. Moreover, there exists also a dual Monge potential $\psi : W \rightarrow \mathbb{R}$, which has an $L^2(\nu)$-extended derivative in the direction of Cameron-Martin space, such that, the map $S = I_W + \nabla \psi$ satisfies $(S \times I_W)\nu = (I_W \times T)\mu = \gamma$, hence $T \circ S = I_W \nu$-a.s. and $S \circ T = I_W$ $\mu$-a.s. The next important issue in this subject is to show the Sobolev regularity of the Monge-Brenier maps in such a way that one can write the Jacobian functions associated to the corresponding transformations $T$ and $S$. In finite dimensional case this problem has been treated by several authors (cf. [3] and the references given in [23]). In the infinite dimensional case there are also some results (cf. [14, 6, 11]) which are generalizations of the results given in [9, 10]. These results are generally some extensions of the methods developed especially by L. Caffarelli, though we have also given another method to calculate the Jacobian functions in infinite dimensions using the Itô calculus.

In this work we shall present a totally different method, namely, we shall prove the Sobolev regularity of the Monge-Brenier functions using the Calculus of variations. Let us begin by recalling a celebrated variational formula, which holds on any measurable space but we formulate it on a
Wiener space for the notational simplicity:

\[ J_f = \inf \{ (I_W + \nabla \varphi) \mu : \varphi \in \mathbb{D}_{2,1} \} \]

where \( M_1(W) \) denotes the set of probability measures on \((W, \mathcal{F})\), \( \mathcal{F} \) being the Borel sigma field of \( W \), \( \gamma, \mu \) are as described above and suppose that the measure \( e^{-f} d\mu \) is of finite relative entropy w.r.t. \( \mu \). Then the infimum is attained at \( d\nu = e^{-f} d\mu \) provided that \( H(\nu/\mu) \) is finite, cf. [20, 21].

On the other hand, we know from [8] that there exists some \( \varphi \in \mathbb{D}_{2,1} \), 1-convex function such that \((I_W + \nabla \varphi)\mu = \nu\), where we use the same notation for the image of a point and of a measure under a measurable map (here the map under question is \( T = I_W + \nabla \varphi \)). Consequently the following identity holds true:

\[ -\log \int e^{-f} d\mu = \inf \left( \int f \circ \nabla \varphi d\mu + H(\varphi \mu) : \varphi \in \mathbb{D}_{2,1} \right). \]

Therefore

\[ -\log \int e^{-f} d\mu \geq \inf \left( \int f \circ (I_W + \nabla \varphi) d\mu + H((I_W + \nabla \varphi)\mu) : \varphi \in \mathbb{D}_{2,1} \right). \]

For this infimum to be finite we need that \( H((I_W + \nabla \varphi)\mu) < \infty \), which implies \((I_W + \nabla \varphi)\mu \ll \mu\).

Besides the right hand side of the inequality (1.3) is always greater than

\[ \inf \left( \int f \circ (I_W + \nabla \varphi) d\mu + H((I_W + \nabla \varphi)\mu) : \varphi \in \mathbb{D}_{2,1} \right), \]

therefore we have equality between all these expressions:

**Theorem 1.** Assume that \( H(\nu/\mu) < \infty \), where \( d\nu = (E[e^{-f}])^{-1} e^{-f} d\mu \) and \( f \) is a measurable function. Then the infimum

\[ J_f = \inf \{ J_f(\xi) : \xi \in \mathbb{D}_{2,1} \} \]

is attained at the vector field \( \xi = \nabla \varphi \), where \( \varphi \) is the unique (up to an additive constant) Monge potential such that \((I_W + \nabla \varphi)\mu = \nu\) and that the \( L^2(\mu, H)\)-norm of \( \nabla \varphi \) is equal to the Wasserstein distance between \( \nu \) and \( \mu \):

\[ d_{H(\mu, \nu)}^2 = \inf \left( \int_{W \times W} |x-y|^2 d\gamma(x, y) : \gamma \in \Sigma_1(\mu, \nu) \right) \]

where \( \Sigma_1(\mu, \nu) \) denotes the set of probability measures on \( W \times W \), whose first marginals are \( \mu \) and the second ones are \( \nu \).

Note that if we could apply the variational principle above, namely, by taking the derivative of the functional \( J_f \) at the minimizing vector field \( \nabla \varphi \) in any admissible direction, we would obtain the following relation:

\[ \delta((I_H + \nabla^2 \varphi)^{-1} - I_H) = \nabla \varphi + \nabla f \circ (I_W + \nabla \varphi), \]

where \( \delta \) denotes the Gaussian divergence, i.e., the adjoint of the derivative \( \nabla \) w.r.t. the Gaussian measure \( \mu \) and this equation implies Sobolev regularity of \( \varphi \). A similar method can be used for
the dual Monge potential $\psi$ also. We shall realize this program in the sequel beginning from the finite dimensions and passing to the infinite dimensional case by a limiting argument. This limit procedure requires more general approximation-stability results about the convergence of Monge potentials corresponding to convergent sequences of target measures than those studied in [9], they are delicate (cf., Lemmas 2 and 4) and they are of independent interest.

Let us note that this method is applicable in other situations than the Gaussian case as one can see already in the case of dual potential.

We make a last remark: this work is devoted to the creation of a variational calculus by parametrizing the formula with the vector fields which are derivatives of scalar functionals. In another work, which has already appeared, [21], we have parametrized the same formula with adapted vector fields to obtain totally different results, like the existence, uniqueness and non-existence results of stochastic differential equations with past depending drift coefficients.

2. Preliminaries

Let $W$ be a separable Fréchet space equipped with a Gaussian measure $\mu$ of zero mean whose support is the whole space. The corresponding Cameron-Martin space is denoted by $H$. Recall that the injection $H \hookrightarrow W$ is compact and its adjoint is the natural injection $W^* \hookrightarrow H^* \subset L^2(\mu)$. The triple $(W, \mu, H)$ is called an abstract Wiener space. Recall that $W = H$ if and only if $W$ is finite dimensional. A subspace $F$ of $H$ is called regular if the corresponding orthogonal projection has a continuous extension to $W$, denoted again by the same letter. It is well-known that there exists an increasing sequence of regular subspaces $(F_n, n \geq 1)$, called total, such that $\cup_n F_n$ is dense in $H$ and in $W$. Let $\mathcal{V}_n$ be the $\sigma$-algebra generated by $\pi_{F_n}$, then for any $f \in L^p(\mu)$, the martingale sequence $(E[f|\mathcal{V}_n], n \geq 1)$ converges to $f$ (strongly if $p < \infty$) in $L^p(\mu)$. Observe that the function $f_n = E[f|\mathcal{V}_n]$ can be identified with a function on the finite dimensional abstract Wiener space $(F_n, \mu_n, F_n)$, where $\mu_n = \pi_n \mu$.

Since the translations of $\mu$ with the elements of $H$ induce measures equivalent to $\mu$, the Gâteaux derivative in $H$ direction of the random variables is a closable operator on $L^p(\mu)$-spaces and this closure will be denoted by $\nabla$ cf., for example [18]. The corresponding Sobolev spaces (the equivalence classes) of the real random variables will be denoted as $\mathbb{D}_{p,k}$, where $k \in \mathbb{N}$ is the order of differentiability and $p > 1$ is the order of integrability. If the random variables are with values in some separable Hilbert space, say $\Phi$, then we shall define similarly the corresponding Sobolev spaces and they are denoted as $\mathbb{D}_{p,k}(\Phi)$, $p > 1$, $k \in \mathbb{N}$. Since $\nabla : \mathbb{D}_{p,k} \rightarrow \mathbb{D}_{p,k-1}(H)$ is a continuous and linear operator its adjoint is a well-defined operator which we represent by $\delta$. In the case of classical Wiener space, i.e., when $W = C(\mathbb{R}_+, \mathbb{R}^d)$, then $\delta$ coincides with the Itô integral of the Lebesgue density of the adapted elements of $\mathbb{D}_{p,k}(H)$ (cf. [18]).

For any $t \geq 0$ and measurable $f : W \rightarrow \mathbb{R}_+$, we note by

$$P_t f(x) = \int_W f(e^{-t}x + \sqrt{1-e^{-2t}}y) \mu(dy),$$

it is well-known that $(P_t, t \in \mathbb{R}_+)$ is a hypercontractive semigroup on $L^p(\mu), p > 1$, which is called the Ornstein-Uhlenbeck semigroup (cf. [18]). Its infinitesimal generator is denoted by $-\mathcal{L}$ and we

---

1The reader may assume that $W = C(\mathbb{R}_+, \mathbb{R}^d), d \geq 1$ or $W = \mathbb{R}^\mathbb{N}$. 

4
call $\mathcal{L}$ the Ornstein-Uhlenbeck operator (sometimes called the number operator by the physicists).

Due to the Meyer inequalities (cf., for instance [18]), the norms defined by

$$
\|\varphi\|_{p,k} = \|(I + \mathcal{L})^{k/2}\varphi\|_{L^p(\mu)}
$$

are equivalent to the norms defined by the iterates of the Sobolev derivative $\nabla$. This observation permits us to identify the duals of the space $\mathcal{D}_{p,k}(\Phi); p > 1, k \in \mathbb{N}$ by $\mathcal{D}_{q,-k}(\Phi')$, with $q^{-1} = 1 - p^{-1}$, where the latter space is defined by replacing $k$ in (2.4) by $-k$, this gives us the distribution spaces on the Wiener space $W$ (in fact we can take as $k$ any real number). An easy calculation shows that, formally, $\delta \circ \nabla = \mathcal{L}$, and this permits us to extend the divergence and the derivative operators to the distributions as linear, continuous operators. In fact $\delta : \mathcal{D}_{q,k}(H \otimes \Phi) \rightarrow \mathcal{D}_{q,k-1}(\Phi)$ and $\nabla : \mathcal{D}_{q,k}(\Phi) \rightarrow \mathcal{D}_{q,k-1}(H \otimes \Phi)$ continuously, for any $q > 1$ and $k \in \mathbb{R}$, where $H \otimes \Phi$ denotes the completed Hilbert-Schmidt tensor product (cf., for instance [18]).

We shall use the following results about the divergence operator whose proof in the regular case is straightforward, for extensions with less regularity hypothesis, we refer the reader to Theorem B.6.4 in [22].

**Lemma 1.**

1. Assume that $\xi : W \rightarrow H$ is a smooth function and let $T = I_W + u$, with $u \in \mathcal{D}_{p,1}(H)$ for some $p > 1$. Suppose that $T \mu \ll \mu$, then the following identity holds $\mu$-a.s.:

$$
(\delta \xi) \circ T = \delta(\xi \circ T) + (\xi \circ T, u)_H + \text{ trace } ((\nabla \xi) \circ T, \nabla u).
$$

2. The second moment of a divergence w.r.t. the Gauss measure is given by

$$
E[(\delta(\xi))^2] = E[|\xi|^2_H] + E[\text{ trace } (\nabla \xi \cdot \nabla \xi)],
$$

where “$\cdot$” is the composition operation between two linear (Hilbert-Schmidt) operators.

The following assertion which has been proved by H. Sugita (cf. [17] or [18]) is useful: assume that $(Z_n, n \geq 1) \subset \mathcal{D}'$ converges to $Z$ in $\mathcal{D}'$, assume further that each each $Z_n$ is a probability measure on $W$, then $Z$ is also a probability and $(Z_n, n \geq 1)$ converges to $Z$ in the weak topology of measures. In particular, a lower bounded distribution (in the sense that there exists a constant $c \in \mathbb{R}$ such that $Z + c$ is a positive distribution) is a (Radon) measure on $W$.

A measurable function $f : W \rightarrow \mathbb{R} \cup \{\infty\}$ is called $H$-convex (cf. [7]) if

$$
h \rightarrow f(x + h)
$$

is convex $\mu$-almost surely, i.e., if for any $h, k \in H, s, t \in [0, 1], s + t = 1$, we have

$$
f(x + sh + tk) \leq sf(x + h) + tf(x + k),
$$

almost surely, where the negligible set on which this inequality fails may depend on the choice of $s, h$ and of $k$. We can rephrase this property by saying that $h \rightarrow (x \rightarrow f(x + h))$ is an $L^0(\mu)$-valued convex function on $H$. $f$ is called 1-convex if the map

$$
h \rightarrow \left(x \rightarrow f(x + h) + \frac{1}{2}|h|^2_H\right)
$$

is convex on the Cameron-Martin space $H$ with values in $L^0(\mu)$. Note that all these notions are compatible with the $\mu$-equivalence classes of random variables thanks to the Cameron-Martin theorem. It is proven in [7] that this definition is equivalent the following condition: Let $(\pi_n, n \geq 1)$
be a sequence of regular, finite dimensional, orthogonal projections of \( H \), increasing to the identity map \( I_H \). Denote also by \( \pi_n \) its continuous extension to \( W \) and define \( \pi_n^\perp = I_W - \pi_n \). For \( x \in W \), let \( x_n = \pi_n x \) and \( x_n^\perp = \pi_n^\perp x \). Then \( f \) is \( 1 \)-convex if and only if
\[
x_n \to \frac{1}{2}|x_n|_H^2 + f(x_n + x_n^\perp)
\]
is \( \pi_n \mu \)-almost surely convex. We define similarly the notion of \( H \)-concave and \( H \)-log-concave functions. In particular, one can prove that, for any \( H \)-log-concave function \( f \) on \( W \), \( P_t f \) and \( E[f|V_n] \) are again \( H \)-log-concave [7].

3. Variational calculations

Assume for a while that \( \varphi \in D_{2,1} \) is smooth; this can be achieved by replacing \( f \) by its regularization defined as
\[
e^{-f_n} = E[P_{1/n}e^{-f}|V_n],
\]
where \( (P_t, t \geq 0) \) is the Ornstein-Uhlenbeck semi-group, \( V_n \) is the sigma-algebra generated by \( \{\delta e_1, \ldots, \delta e_n\} \) and \( (e_n, n \geq 1) \) is a complete, orthonormal basis of \( H \). Since \( J_I = J(\nabla \varphi) \), if we take the Gateau derivative of \( J \) at \( \nabla \varphi \), it should give zero: Let \( L = (E[-f])^{-1} e^{-f} \) and denote by \( \Lambda \) the Gaussian Jacobian of \( T = I_W + \nabla \varphi \):
\[
\Lambda = \det_2(I_H + \nabla^2 \varphi) \exp \left( -L \varphi - \frac{1}{2} |\nabla \varphi|_H^2 \right)
\]
where \( L \) is the Ornstein-Uhlenbeck operator \( L = \delta \circ \nabla \), \( \det_2 \) denotes the modified Carleman-Fredholm determinant, \( \delta = \nabla^* \) where the adjoint is taken w.r.t. the Wiener measure \( \mu \), c.f. [22]. It follows from the change of variables formula, c.f. [22], that \( L \circ (I_W + \nabla \varphi) \Lambda = 1 \), hence
\[
H((I_W + \nabla \varphi)\mu|\mu) = E \left[ \frac{1}{2} |\nabla \varphi|_H^2 - \log \det_2(I_H + \nabla^2 \varphi) \right].
\]
In particular, thanks to the \( 1 \)-convexity of \( \varphi \), if we replace \( \varphi \) by \( t \varphi \), for small \( t \in [0,1] \), the shift \( T_t = I_W + t \nabla \varphi \) becomes strongly monotone and it is the solution of the Monge transportation problem for the measure \( \nu_t = T_t \mu \) (i.e., the image of \( \mu \) under \( T_t \)). Let \( f_t \) be defined as
\[
L_t = \frac{d\nu_t}{d\mu} = ce^{-f_t}.
\]
If \( \xi \in D_{2,1}(H) \) such that \( \nabla \xi \) has small \( L^\infty \)-norm as a Hilbert-Schmidt operator, then \( T_{t,\varepsilon} = I_W + t \nabla \varphi + \varepsilon \xi \) is a strongly monotone shift for small \( t, \varepsilon > 0 \), hence it is almost-surely invertible (cf. [22], Corollary 6.4.2). Note moreover that the shift \( I_W + t \nabla \varphi \) is the unique solution of another Monge problem, namely the one which corresponds to the measure \( ce^{-f_t} d\mu \). Here the multiplication with a small \( t \) permits us to have a sufficiently large set on which we calculate the Gateau derivative while preserving the \( 1 \)-convexity of the corresponding Monge potential, namely \( t \varphi \). Using again the change of variables formula for \( T_{t,\varepsilon} \), we get
\[
H(T_{t,\varepsilon}\mu|\mu) = E \left[ \frac{1}{2} |t \nabla \varphi + \varepsilon \xi|_H^2 - \log \det_2(I_H + t \nabla^2 \varphi + \varepsilon \nabla \xi) \right].
\]
Therefore
\[
J_t(t \nabla \varphi + \varepsilon \xi) = E \left[ f_t \circ T_{t,\varepsilon} + \frac{1}{2} |t \nabla \varphi + \varepsilon \xi|_H^2 - \log \det_2 \left( I_H + t \nabla^2 \varphi + \varepsilon \nabla \xi \right) \right].
\]
Since \( t \nabla \varphi \) minimizes the function \( J_1 \) between all the absolutely continuous shifts, we should have

\[
(3.6) \quad \frac{d}{d \varepsilon} J_1 ((t \nabla \varphi + \varepsilon \xi))_{\varepsilon=0} = E \left[ (t \nabla \varphi, \xi)_H - \operatorname{trace} \left( ((I + t \nabla^2 \varphi)^{-1} - I) \cdot (\nabla \xi) \right) + (\nabla f_t \circ (I_W + t \nabla \varphi), \xi)_H \right] = 0
\]

for any \( \xi \in D_2(1)(H) \) with \( \| \nabla \xi \|_2 \in L^\infty(\mu) \). Since the set of vector fields

\[
\Theta = \{ \xi \in D_2(1)(H) : \| \nabla \xi \|_2 \in L^\infty(\mu) \}
\]

is dense in any \( L^p(\mu, H) \), we have proved the following

**Theorem 2.** In the finite dimensional smooth case, the Monge potential \( \varphi \) satisfies the following relation

\[
\nabla \varphi + \nabla f \circ (I_W + t \nabla \varphi) - \delta [(I_H + t \nabla^2 \varphi)^{-1} - I_H] = 0
\]

almost surely, where \( \delta \) denotes the Gaussian divergence w.r.t. \( \mu \), i.e., the adjoint of \( \nabla \) w.r.t. \( \mu \).

**Proof:** In the equation (3.6) we have a term with trace, we just interpret it as a scalar product on the Hilbert-Schmidt operators on the Cameron-Martin space and the claim follows, for the case \( t \varphi \), from the definition of \( \delta \) as a mapping from Hilbert-Schmidt-valued operators to the vector fields under this scalar product. Hence we have the identity

\[
(3.7) \quad t \nabla \varphi + \nabla f_t \circ (I_W + t \nabla \varphi) - \delta [(I_H + t \nabla^2 \varphi)^{-1} - I_H] = 0.
\]

Since we have \( \Lambda_t e^{-f \circ T_t} = 1 \) a.s., where \( \Lambda_t = \det_2 (I_H + t \nabla^2 \varphi) \exp \left( -t \L \varphi - \frac{1}{2} t \| \nabla \varphi \|_H^2 \right) \) and \( T_t = I_W + t \nabla \varphi \), \( \lim_{t \to 1} \nabla f_t \circ T_t = \nabla f \circ T \) in probability, where \( T = T_1 = I_W + \nabla \varphi \). The justification of the other terms being trivial, the proof is completed. \( \square \)

**Theorem 3.** Suppose that \( f \) is a smooth, bounded function on \( \mathbb{R}^d \), denote the probability measure \( e^{-f} \, d\beta \) by \( \nu \), where \( \beta \) is the standard Gauss measure on \( \mathbb{R}^d \). Assume that the forward and backward potentials associated to the transport couple \( (\beta, \nu) \) denoted respectively \( (\varphi, \psi) \) are smooth functions. Let \( T \) be defined as \( T = I_{\mathbb{R}^d} + \nabla \varphi \). We have then the following control:

\[
(3.8) \quad E \left[ \| \nabla^2 \psi \circ T \|_2^2 \right] + \sum_{i} E \left[ \left( \| (I + \nabla^2 \varphi)^{1/2} \nabla^3 \psi \circ T (I + \nabla^2 \varphi)^{1/2} e_i \|_2 \right)^2 \right] = E \left[ \| \nabla \varphi + \nabla f \circ T \|_2^2 \right],
\]

where \( (e_i, i \leq d) \) is an orthonormal basis of \( \mathbb{R}^d \) and \( \| \cdot \|_2 \) denotes the Hilbert-Schmidt norm.

**Proof:** Let \( M = (I + \nabla^2 \varphi)^{-1} - I \), from Theorem 2 we have

\[
\delta(M) = \nabla \varphi + \nabla f \circ T
\]

and from Lemma 1

\[
E[\| \delta(M) \|_{\mathbb{R}^d}^2] = \sum_{i=1}^{d} E[\| \delta(M e_i) \|_2^2] = \sum_{i=1}^{d} E \left[ |M e_i|_{\mathbb{R}^d}^2 + \operatorname{trace} ((\nabla M e_i) \cdot (\nabla M e_i)) \right].
\]

7
Monge-Ampere Equation Variational Solutions

From the relation \((I + \nabla \psi) \circ T = I_{\mathbb{R}^d}\), we obtain

\[(I + \nabla^2 \psi) \circ T = (I + \nabla^2 \varphi)^{-1},\]

hence, differentiating both sides again we get

\[
\nabla((I + \nabla^2 \psi) \circ T) = \nabla^3 \psi \circ T(I + \nabla^2 \varphi) = \nabla(I + \nabla^2 \varphi)^{-1} = -(I + \nabla^2 \varphi)^{-1}\nabla^3 \varphi(I + \nabla^2 \varphi)^{-1},
\]

this gives the identity

\[(3.9) \quad -(I + \nabla^2 \varphi)^{-1}\nabla^3 \varphi(I + \nabla^2 \varphi)^{-1} = \nabla^3 \psi \circ T(I + \nabla^2 \varphi) \]

almost surely. Using the equality \[(3.9)\] in the calculation of \(\nabla M e_i\), we get

\[
\nabla(M e_i) = -(I + \nabla^2 \varphi)^{-1} \nabla^3 \varphi(I + \nabla^2 \varphi)^{-1} e_i = (\nabla e_i \nabla^2 \psi) \circ T(I + \nabla^2 \varphi).
\]

We finally get

\[
\text{trace } ((\nabla M e_i).(\nabla M e_i)) = \text{trace } ((\nabla e_i \nabla^2 \psi) \circ T(I + \nabla^2 \varphi) \cdot (\nabla e_i \nabla^2 \psi) \circ T(I + \nabla^2 \varphi))
\]

\[
= \text{trace } \left( \left(I + \nabla^2 \varphi \right)^{1/2}(\nabla e_i \nabla^2 \psi) \circ T(I + \nabla^2 \varphi)^{1/2} \right) \left[ (I + \nabla^2 \varphi)^{1/2}(\nabla e_i \nabla^2 \psi) \circ T(I + \nabla^2 \varphi)^{1/2} \right]
\]

\[
= \| (I + \nabla^2 \varphi)^{1/2}(\nabla e_i \nabla^2 \psi) \circ T(I + \nabla^2 \varphi)^{1/2} \|^2_2,
\]

and the proof follows. \(\square\)

4. Approximations of the Monge Potentials

The results given in this section are indispensable to study the stability and the approximation results of the forward and backward potentials in the finite dimensional situations whenever the target measures are approximated with more regular measures. There are some results in the literature (cf. \([5, 23]\)), but they do not cover the situations that we shall encounter.

**Lemma 2.** Let \(\beta\) be the standard Gaussian measure on \(\mathbb{R}^d\), \(f \in D_{2,1}\) s.t.

\[
\int_{\mathbb{R}^d} |\nabla f|^2 e^{-f} d\beta < \infty.
\]

Let \((\varphi, \psi)\) be the Monge potentials associated to the Monge-Kantorovich problem \(\Sigma(\beta, \nu)\), where \(d\nu = ce^{-f} d\beta\). Define \(f_n\) as to be \(Q_{1/n} e^{-f} = e^{-f_n}\), where \((Q_t, t \geq 0)\) denotes the Ornstein-Uhlenbeck semigroup on \(\mathbb{R}^d\). Let \((\varphi_n, \psi_n)\) be the Monge potentials corresponding to Monge-Kantorovich problem \(\Sigma(\beta, \nu_n)\), where \(d\nu_n = ce^{-f_n} d\beta\). Then \((\varphi_n, n \geq 1)\) converges to \(\varphi\) in \(D_{2,1}\), \((Q_{1/n} \psi_n, n \geq 1)\) converges to \(\psi\) in \(L^1(\nu)\) and \((Q_{1/n} \nabla \psi_n, n \geq 1)\) converges to \(\nabla \psi\) in \(L^2(\nu, \mathbb{R}^d)\).

**Proof:** In the sequel we replace \(\varphi_n\) by \(\varphi_n - E_\beta[\varphi_n]\) and \(Q_{1/n} \psi_n\) by \(Q_{1/n} \psi_n + E_\beta[\varphi_n]\) to avoid the ambiguities about the constants. Let \(\gamma_n, \gamma\) be the unique solutions of Monge-Kantorovich problems for \((\beta, \nu_n)\) and \((\beta, \nu)\) respectively. From Brenier’s theorem (cf. \([2]\))

\[(4.10) \quad F_n(x, y) = \varphi_n(x) + \psi_n(y) + \frac{1}{2} |x - y|^2 = 0 \quad \gamma_n - a.s.
\]
and $F_n(x, y) \geq 0$ for any $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. Similarly

\begin{equation}
F(x, y) = \varphi(x) + \psi(y) + \frac{1}{2} |x - y|^2 = 0 \quad \text{a.s.,}
\end{equation}

and $F(x, y) \geq 0$ for any $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. Let us denote by $p(x) = |x|^2$ (the Euclidean norm), then for any $\lambda < 1/2$, we have $\int \exp[\lambda Q_{1/n}p(x)]d\beta(x) \leq \int \exp[\lambda p(x)]d\beta(x) < \infty$ due to invariance of $\beta$ w.r.t. the Ornstein-Uhlenbeck semigroup $(Q_t, t \geq 0)$ and due to the Jensen inequality. From the definition of $\nu_n$

\begin{equation}
\int |y|^2d\nu_n(y) = \int |y|^2Q_{1/n}(e^{-f})d\beta.
\end{equation}

From the Young and Jensen inequalities, we obtain

\begin{equation}
|y|^2Q_{1/n}(e^{-f}) \leq e^{c|y|^2} + \frac{1}{\epsilon}Q_{1/n}(e^{-f}) \log Q_{1/n}(e^{-f}),
\end{equation}

finally, again from the Jensen inequalities, it follows that the sequence of integrands at the right of the equality (4.12) is uniformly integrable. Consequently it holds true that

\begin{equation}
\lim_{n \to \infty} \int_{\mathbb{R}^d} |y|^2d\nu_n(y) = \int |y|^2d\nu(y)
\end{equation}

and this implies that (cf. [1], Lemma 8.3)

\begin{align*}
\lim_n E_{\beta}[[\nabla \varphi_n]^2] &= \lim_n d_2(\beta, \nu_n)^2 \\
&= d_2(\beta, \nu)^2 \\
&= E_{\beta}[[\nabla \varphi]^2],
\end{align*}

where $d_2$ denotes the second order Wasserstein distance on the probability measures on $\mathbb{R}^d$. These relations imply that $(\varphi_n, n \geq 1)$ is bounded in $L^2(\gamma)$. Moreover, taking into account the relation $\nabla Q_{1/n}g = e^{-t}Q_{1/n}\nabla g$ for any smooth $g$, we have

\begin{align*}
E_{\nu_n}[[\nabla \psi_n]^2] &= E_{\beta}[[\nabla \varphi_n]Q_{1/n}e^{-f}] \geq E_{\beta}[[\nabla Q_{1/n}\psi_n]^2e^{-f}].
\end{align*}

By the boundedness of $(\varphi_n, n \geq 1)$ in $L^2(\beta)$ there exists $\varphi' \in L^2(\beta)$ such that $(\varphi_n, n \geq 1)$ converges weakly to $\varphi'$ (upto a subsequence) in $L^2(\beta)$, hence also in $L^2(\gamma)$. Since $F_n \geq 0$ everywhere, by applying $Q_{1/n}$ in $y$-variable, we get

\begin{equation}
Q_{1/n}F_n(x, \cdot)(y) = \varphi_n(x) + Q_{1/n}\psi_n(y) + \frac{1}{2} Q_{1/n}(|x| - |y|)^2(y) \geq 0
\end{equation}

for any $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. Integrating this inequality w.r.t. the measure $\gamma$ and taking the limit, we get

\begin{align*}
\lim_n \int \left( \varphi_n(x) + Q_{1/n}\psi_n(y) + \frac{1}{2} Q_{1/n}(|x| - |y|)^2(y) \right) d\gamma \\
= \lim_n \left( \int \varphi_n d\beta + \int \psi_n d\nu_n + \frac{1}{2} Q_{1/n}(|x| - |y|)^2(y) d\gamma \right).
\end{align*}

Let us calculate the third term:

\begin{align*}
\int Q_{1/n}(|x| - |y|)^2(y) d\gamma &= \int \int |x - e^{-1/n}y - \sqrt{1 - e^{-2/n}}z|^2 d\beta(z) d\gamma(x, y) \\
&= \int \int \left( |x - e^{-1/n}y|^2 + (1 - e^{-2/n})|z|^2 \right) d\beta(z) d\gamma(x, y)
\end{align*}
As
\[ \lim_{n \to \infty} \int (1 - e^{-1/n}) |z|^2 \beta(dz) = 0 \]
and as
\[ \int |y|^2 d\gamma = \int |y|^2 d\nu = \int |y|^2 e^{-f} d\beta(y) \leq \int e^{\epsilon |y|^2} d\beta(y) + \frac{1}{\epsilon} H(\nu|\beta), \]
we obtain
\[ \lim_n \int Q_{1/n}(|x - \cdot|^2)(y) d\gamma = \int |x - y|^2 d\gamma(x, y). \]
Let us note for later use that the inequality (4.13) combined with the triangle inequality implies the \( \gamma \)-uniform integrability of the sequence \((Q_{1/n}(|x - \cdot|^2)(y), n \geq 1)\) and from the dominated convergence theorem, we see that \((Q_{1/n}(|x - \cdot|^2)(y), n \geq 1)\) converges to \(|x - y|^2\) in \( \gamma \)-probability (i.e., in \( L^0(\gamma) \)) since
\[ |x - e^{-1/n} y + \sqrt{1 - e^{-2/n}} z|^2 \leq 4(|x|^2 + |y|^2) + 2|z|^2. \]

By the Poincaré inequality, \((\varphi_n, n \geq 1)\) is bounded in \( D_{2,1} \) w.r.t. the Gaussian measure \( \beta \), hence also it is bounded in \( L^2(\gamma) \). Therefore, up to a subsequence, it converges weakly to some \( a' \in L^2(\gamma) \) (note that \( a' \in D_{2,1}(\beta) \) also). Moreover, from the relation (4.10)
\[ \lim_n \int Q_{1/n} \psi_n(y) d\gamma(x, y) = \lim \int \psi_n(y) d\nu_n(y) \]
\[ \quad = - \lim \int \varphi_n d\beta - \frac{1}{2} \lim \int |x - y|^2 d\gamma_n(x, y) \]
\[ \quad = - \lim \int |a d\beta - \frac{1}{2} \int |x - y|^2 d\gamma(x, y) \]
where the last equality above follows from
\[ \lim \frac{1}{2} d_2^2(\beta, \nu_n) = \lim \frac{1}{2} \int |x - y|^2 d\gamma_n(x, y) = \int |x - y|^2 d\gamma(x, y) \]
which is a consequence of Lemma 8.3 of [1]. Consequently
\[ \lim_n \int Q_{1/n} F_n(x, \cdot)(y) d\gamma(x, y) = 0. \]

Therefore the sequence \((\varphi_n(x) + Q_{1/n} \psi_n(y) + \frac{1}{2} |x - y|^2, n \geq 1)\) converges to 0 in the norm topology of \( L^1(\gamma) \), consequently \((Q_{1/n} \psi_n, n \geq 1)\) is also uniformly integrable in \( L^1(\gamma) \), therefore there exists some \( b' \in L^1(\nu) \) which is a weak adherent point of \((Q_{1/n} \psi_n, n \geq 1)\). Therefore
\[ a'(x) + b'(y) + \frac{1}{2} |x - y|^2 = 0 \]
\( \gamma \)-a.s. Let \((\varphi'_n, n \geq 1)\) and \((Q_{1/n} \psi'_n, n \geq 1)\) be the convex combinations of the sequences \((\varphi_n)\) and \((Q_{1/n} \psi_n)\) respectively, which converge strongly in \( L^2(\gamma) \) and \( L^1(\gamma) \) respectively. Let \( a(x) = \limsup_n \varphi'_n(x)\) and \( b(y) = \limsup Q_{1/n} \psi'_n(y)\). We have then
\[ a(x) + b(y) + \frac{1}{2} |x - y|^2 \geq 0 \]
Monge-Ampere Equation

As $\phi$ of the derivative operator, and the weak limit is equal to

$$\gamma \xi_{\text{converges weakly in }} L^2$$

In the next section we shall need the following result which is a corollary of Lemma 2 whose notations are used without further explanation:

**Corollary 1.** Assume the hypothesis of Lemma 2 are valid and assume that $(\nabla^2 \psi_n \circ T_n, n \geq 1)$ converges weakly in $L^2(\beta, \mathbb{R}^d \otimes \mathbb{R}^d)$, then $\nabla^2 \psi \in L^2(\nu, \mathbb{R}^d \otimes \mathbb{R}^d)$, where $\nabla$ is the closure in $L^2(\nu)$ of the derivative operator, and the weak limit is equal to $\nabla^2 \psi \circ T$, i.e.,

$$w - \lim_n \nabla^2 \psi_n \circ T_n = \nabla^2 \psi \circ T.$$

**Proof:** As $S \circ T = I_{\mathbb{R}^d}$ $\beta$-a.s., where $S = I + \nabla \psi$, $T = I + \nabla \varphi$, we can represent $\lim_n \nabla^2 \psi_n \circ T_n$ as $\xi \circ T$, where $\xi \in L^2(\nu, \mathbb{R}^d \otimes \mathbb{R}^d)$. As $\lim_n Q_{1/n} \psi_n = \psi$ in $L^1(\nu)$, $\lim_n Q_{1/n} \nabla \psi_n = \nabla \psi$ and as $\nabla$ is closable in $L^2(\nu)$, $(\nabla^2 Q_{1/n} \psi_n, n \geq 1)$ converges as a distribution, i.e., for any smooth function of compact support $\eta$, we have

$$\lim_n E_\nu[\nabla^2 Q_{1/n} \psi_n \cdot \eta] = E_\nu[\nabla^2 \psi \cdot \eta].$$

On the other hand, letting $L_n = Q_{1/n} L,$

$$E_\nu[\nabla^2 Q_{1/n} \psi_n \cdot \eta] = E[\nabla^2 Q_{1/n} \psi_n \circ T \cdot \eta \circ T]$$

$$= e^{-2/n} E[Q_{1/n} \nabla^2 \psi_n \cdot \eta L]$$

$$= e^{-2/n} E[\nabla^2 \psi_n \cdot Q_{1/n} (\eta L)]$$

$$= e^{-2/n} \left[ \nabla^2 \psi_n \cdot \frac{Q_{1/n} (\eta L)}{L_n} L_n \right]$$

$$= e^{-2/n} \left[ \nabla^2 \psi_n \circ T_n \cdot \left( \frac{Q_{1/n} (\eta L)}{L_n} \right) \circ T_n \right].$$

As

$$\left| \frac{Q_{1/n} (\eta L)}{L_n} \right| \leq \| \eta \|_{\infty}$$

we get

$$\lim_n E_\nu[\nabla^2 Q_{1/n} \psi_n \cdot \eta] = E_\nu[\nabla^2 \psi \cdot \eta]$$

$$= E[\xi \circ T \cdot \eta \circ T]$$

$$= E[\xi \cdot \eta L].$$

---

2This follows from $\int |\nabla f|^2 \nu < \infty$. 

---
consequently \( \nabla \psi \) belongs to the \( L^2(\nu) \)-extended domain of the Gateaux derivative operator and \( \xi = \nabla^2 \psi \) \( \nu \)-almost surely.

\[ \]

\textbf{Lemma 3.} Let \( \beta \) be the standard Gaussian measure on \( \mathbb{R}^d \), \( L \in L^1(\beta) \) be a probability density such that
\[ \int_{\mathbb{R}^d} L \log L d\beta < \infty. \]

Let \((\varphi, \psi)\) be the Monge potentials associated to the Monge-Kantorovich problem \( \Sigma(\beta, \nu) \), where \( d\nu = Ld\beta \). Define \( L_n(y) = c_n L(y) \tilde{\theta}_n(y) = \theta_n(y)L(y) \) as another density, where \( \tilde{\theta}_n \in C_{\infty}(\mathbb{R}^d) \) is approximating the constant 1, \( c_n \) is the normalization constant and \( \theta_n = c_n \tilde{\theta}_n \). Let \((\varphi_n, \psi_n)\) be the Monge potentials corresponding to Monge-Kantorovich problem with quadratic cost over \( \Sigma(\beta, \nu_n) \), where \( d\nu_n = L_n d\nu \). Then \((\varphi_n, n \geq 1)\) converges to \( \varphi \) in \( D_{2,1} \), \((\theta_n \psi_n, n \geq 1)\) converges to \( \psi \) in \( L^1(\nu) \) and also \((\sqrt{\theta_n} \nabla \psi_n, n \geq 1)\) converges to \( \nabla \psi \) in \( L^2(\gamma) \), in particular
\[ \lim_n E_{\nu}[\theta_n |\nabla \psi_n|^2] = E_{\nu}[|\nabla \psi|^2]. \]

\textbf{Proof:} The proof is similar to the proof of Lemma [2]. Let \( \gamma \) and \( \gamma_n \) be the transport plans corresponding to the Monge-Kantorovich problems for \( (\beta, \nu) \) and \( (\beta, \nu_n) \) respectively. We replace \( \varphi_n \) by \( \varphi_n - E_\beta[\varphi_n] \) and \( \psi_n \) with \( \psi_n + E_\beta[\varphi_n] \) to fix the ideas. We denote also by \( \theta_n(y) \) the function \( c_n \tilde{\theta}_n(y) \). As in the Lemma [2], we have
\[ F_n(x, y) = \varphi_n(x) + \psi_n(y) + \frac{1}{2} |x - y|^2 \geq 0 \]
for any \( x, y \in \mathbb{R}^d \) and
\[ F_n(x, y) = \varphi_n(x) + \psi_n(y) + \frac{1}{2} |x - y|^2 = 0 \quad \gamma_n \text{--} a.s. \]

It follows from (4.14) that
\[ 0 \leq \int \theta_n(y) F_n(x, y) d\gamma(x, y) \]
\[ = \int \theta_n(y) \varphi_n(x) d\gamma(x, y) + \int \theta_n(y) \psi_n(y) d\gamma(x, y) \]
\[ + \frac{1}{2} \int \theta_n(y) |x - y|^2 d\gamma(x, y) \]
\[ = I_n + II_n + III_n \]

Let us observe the second term in more detail:
\[ II_n = \int \psi_n(y) \theta_n(y) d\gamma(x, y) = \int \psi_n(y) \theta_n(y) d\nu(y) \]
\[ = \int \psi_n(y) d\nu_n(y) = \int \psi_n(y) d\gamma_n(x, y) \]
\[ = -\frac{1}{2} \int |x - y|^2 d\gamma_n(x, y) - \int \varphi_n d\beta. \]
Consequently
\[
\int \theta_n(y) F_n(x, y) d\gamma(x, y) = \int \theta_n(y) \varphi_n(x) d\gamma(x, y) - \frac{1}{2} \int |x-y|^2 d\gamma_n(x, y) \\
- \int \varphi_n d\beta + \frac{1}{2} \int \theta_n(y)|x-y|^2 d\gamma(x, y) \\
= \int \theta_n(y) \varphi_n(x) d\gamma(x, y) - \int \varphi_n(x) d\gamma(x, y) \\
- \frac{1}{2} \int |x-y|^2 d\gamma_n(x, y) \\
+ \frac{1}{2} \int \theta_n(y)|x-y|^2 d\gamma(x, y) \\
= \int (\theta_n(y) - 1) \varphi_n(x) d\gamma(x, y) \\
- \frac{1}{2} \int |x-y|^2 d\gamma_n(x, y) + \frac{1}{2} \int \theta_n(y)|x-y|^2 d\gamma(x, y)
\]

Recall that \( (\tilde{\theta}_n, n \geq 1) \) is a sequence of positive smooth functions of compact support increasing to one, hence the sequence \( (c_n, n \geq 1) \) defined as \( c_n^{-1} = \int \tilde{\theta}_n(y)L(y)d\beta(y) \), increases to one. Therefore, for any \( \varepsilon > 0 \) there exists some \( n_\varepsilon \in \mathbb{N} \) such that \( 1 \leq c_n \leq 1 + \varepsilon \) for any \( n \geq n_\varepsilon \). Remark also that, from Poincaré and Talagrand inequalities,

\[
\int \varphi_n^2 d\gamma \leq \int |\nabla \varphi_n|^2 d\beta \leq 2 \int L_n \log L_n d\beta = 2 \int (\theta_n L \log L + \theta_n L \log \theta_n) d\beta \leq \frac{2}{1-\varepsilon} \int L \log L d\beta + \frac{2}{1-\varepsilon},
\]

where we have used the fact that \( \theta_n \leq \frac{1}{1-\varepsilon} \) for \( n \geq n_\varepsilon \). As this estimation is uniform w.r.t. \( n \geq n_\varepsilon \), we conclude that

\[
\lim_{n \to \infty} \int \varphi_n(x)(\theta_n(y) - 1) d\gamma(x, y) = 0.
\]

It follows from the monotone convergence theorem that

\[
\lim_{n \to \infty} \int \theta_n(y)|x-y|^2 d\gamma(x, y) = \int |x-y|^2 d\gamma(x, y).
\]

The Young inequality can be applied as

\[
|y|^2 \theta_n(y) L \leq \frac{1}{1-\varepsilon}(e^{\varepsilon|y|^2} + \frac{1}{\varepsilon} L \log L),
\]

for \( n \geq n_\varepsilon \) as explained above. Hence the sequence \( (|y|^2 \theta_n(y)L(y), n \geq 1) \) is uniformly integrable w.r.t. the measure \( \beta \), therefore

\[
\lim_{n \to \infty} \int |y|^2 L_n d\beta = \lim_{n \to \infty} \int |y|^2 d\nu_n = \int |y|^2 d\nu.
\]

The relation (4.17), combined with \[\text{II}\] implies

\[
\lim_n \frac{1}{2} \int |x-y|^2 d\gamma_n(x, y) = \lim_n \frac{1}{2} d^2_2(\nu_n, \beta) \\
= \frac{1}{2} d^2_2(\nu, \beta) = \frac{1}{2} \int |x-y|^2 d\gamma(x, y).
\]
The relations (4.15), (4.16) and (4.17) imply that
\[ \lim_{n} \int \theta_n(y) F_n(x, y) d\gamma(x, y) = 0. \]
As \((\theta_n, n \geq n_0)\) converges to one and non-negative, there exists a subsequence \((F_{n_k}, k \geq 1)\) which converges to zero \(\gamma\)-a.s. Moreover \((\theta_n(y) \varphi_n(x), n \geq 1)\) is \(\gamma\)-uniformly integrable, hence \((\theta_n \psi_n, n \geq 1)\) is also \(\gamma\)-uniformly integrable. Consequently, up to a subsequence, the sequences \((\theta_n(y) \varphi_n(x), n \geq 1)\) and \((\theta_n \psi_n, n \geq 1)\) converge weakly in \(L^1(\gamma)\) respectively to \(a'\) and \(\psi'\). Moreover \((\varphi_n, n \geq 1)\) is bounded in the Sobolev space \(\mathcal{D}^2_{2,1}\), hence it has a subsequence which converges to some \(\varphi'\) weakly in \(\mathcal{D}^2_{2,1}\), hence also weakly in \(L^2(\gamma)\). For any \(h \in L^\infty(\gamma)\), we have
\[ \int (\varphi_n \theta_n - \varphi') h d\gamma = \int \varphi_n(\theta_n - 1) h d\gamma + \int (\varphi_n - \varphi') h d\gamma, \]
As \((\theta_n, n \geq 1)\) converges to one in all \(L^p\)-spaces, the first terms at the right converges to zero, the second one also converges to zero as \(n\) tends to infinity, therefore \(a' = \varphi'\) \(\gamma\)-a.s. Let us take convex combinations of these weakly convergent sequences to obtain the strong convergence these combinations, which assured by Mazur’s Lemma:

- The sequence of convex combinations \((co(\theta_n \varphi_n), n \geq 1)\) converges to \(\varphi'\) in \(L^2(\gamma)\),
- the sequence of convex combinations \((co(\theta_n \psi_n), n \geq 1)\) converges to \(\psi'\) in \(L^1(\gamma)\),
- \((\theta_n(y)|x-y|^2, n \geq 1)\) converges to \(|x-y|^2\) in \(L^1(\gamma)\).

Define \(\tilde{\varphi} = \limsup_n (co(\theta_n \varphi_n))\) and \(\tilde{\psi} = \limsup_n (co(\theta_n \psi_n))\). We have then
\[ \tilde{\varphi}(x) + \tilde{\psi}(y) + \frac{1}{2} |x - y|^2 \geq 0 \]
for any \(x, y \in \mathbb{R}^d\) and we have also that
\[ \tilde{\varphi}(x) + \tilde{\psi}(y) + \frac{1}{2} |x - y|^2 = 0 \]
\(\gamma\)-almost surely. By the uniqueness of the solutions of the dual Monge-Kantorovitch, we should have \(\varphi = \tilde{\varphi}\) and \(\psi = \tilde{\psi}\) \(\gamma\)-almost surely. As \((\varphi_n, n \geq 1)\) converges to \(\varphi\) weakly in \(L^2(\beta)\) and as \(\lim_n E_\beta(\|\nabla \varphi_n\|^2) = E_\beta(\|\nabla \varphi\|^2), (\varphi_n, n \geq 1)\) converges to \(\varphi\) strongly in \(\mathcal{D}^2_{2,1}\). As \((\theta_n F_n, n \geq 1)\) converges to zero strongly in \(L^1(\gamma)\), \((\theta_n \psi_n, n \geq 1)\) converges strongly to \(\psi\) in \(L^1(\gamma)\). Moreover, we have
\[ \int_{\mathbb{R}^d \times \mathbb{R}^d} \theta_n |\nabla \psi_n|^2 d\gamma = \int_{\mathbb{R}^d} \theta_n |\nabla \psi_n|^2 d\nu = \int |\nabla \psi_n|^2 d\nu_n = \int |\nabla \psi_n \circ T_n|^2 d\beta = \int |\nabla \varphi_n|^2 d\beta, \]
where \(T_n = I_{\mathbb{R}^d} + \nabla \varphi_n\) is the forward transport map, and we obtain at once
\[ (4.18) \quad \lim_{n} \int_{\mathbb{R}^d \times \mathbb{R}^d} \theta_n |\nabla \psi_n|^2 d\gamma = \int |\nabla \psi|^2 d\gamma. \]
In particular, the relation (4.18) implies the \(\gamma\)-uniform integrability of the sequence \((\sqrt{\theta_n} \nabla \psi_n, n \geq 1)\). We shall obtain the strong convergence of the latter as soon as we prove that its weak limit is equal
to $\nabla \psi$. To see this, let $\xi$ be a smooth vector field, then

\[
(4.19) \quad \int (\nabla \psi_n, \xi) \sqrt{\theta_n} d\gamma = \int_{\{\theta_n < c\}} (\nabla \psi_n, \xi) \sqrt{\theta_n} d\gamma + \int_{\{\theta_n \geq c\}} (\nabla \psi_n, \xi) \sqrt{\theta_n} d\gamma.
\]

As $\theta_n \to 1$, by the uniform integrability of $(\sqrt{\theta_n} \nabla \psi_n, n \geq 1)$ and the boundedness of $\xi$, we see that the first integral at the right of the equality (4.19) can be made arbitrarily small for $c < 1$ for any $n \geq n_c$, for some $n_c \in \mathbb{N}$. The second integral at the right of the equality (4.19) is equal to

\[
\int_{\{\theta_n \circ T_n \geq c\}} \frac{(\nabla \varphi_n, \xi \circ T_n)}{\sqrt{\theta_n \circ T_n}} d\beta.
\]

By the uniform integrability of $(L_n, n \geq 1)$, the sequence $(T_n, n \geq 1)$ is equi-concentrated on compacta and consequently $\lim \theta_n \circ T_n = 1$ in $\beta$-probability, also $\lim \xi \circ T_n = \xi \circ T$ in $\beta$-probability.

The dominated convergence theorem and the $L^2$-boundedness of $(\nabla \varphi_n, n \geq 1)$ imply at one that

\[
\lim_{n \to \infty} \int_{\{\theta_n \circ T_n \geq c\}} \frac{(\nabla \varphi_n, \xi \circ T_n)}{\sqrt{\theta_n \circ T_n}} d\beta = -\int (\nabla \varphi, \xi \circ T) d\beta
\]

\[
= \int (\nabla \psi \circ T, \xi \circ T) d\beta
\]

\[
= \int (\nabla \psi, \xi) d\gamma.
\]

We also have the following, where we use the same notations as in the preceding lemmas:

**Lemma 4.** Assume that $d\nu = ce^{-f} d\beta$, where $f : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ is a measurable function with the property that $\nu(\mathbb{R}^d) = 1$ and that

\[
(4.20) \quad \int (|f| + |\nabla f|^2) d\nu < \infty.
\]

Define $f_n = f \wedge n$, $n \in \mathbb{N}$ and define $d\nu_n = c_n e^{-f_n} d\beta$. Let $(\varphi_n, \psi_n)$ and $(\varphi, \psi)$ be the Monge potentials corresponding to the Monge-Kantorovitch problem over $\Sigma(\beta, \nu)$ and over $\Sigma(\beta, \nu)$ respectively (with quadratic cost). Then $(\varphi_n, n \geq 1)$ converges to $\varphi$ in $D_{x,1}$ and $(\psi_n, n \geq 1)$ converges to $\psi$ in $L^1(\nu)$ as well as $(\nabla \psi_n, n \geq 1)$ converges $\nabla \psi$ in $L^2(\nu)$.

**Proof:** We apply the same conventions about the expectations of $\varphi_n$ as in the preceding lemmas. Let $F_n(x, y) = \varphi_n(x) + \psi_n(y) + \frac{1}{2} |x - y|^2$, also define $F(x, y) = \varphi(x) + \psi(y) + \frac{1}{2} |x - y|^2$. Denote by $\gamma$ and $\gamma_n$ the Monge-Kantorovitch plans for $\Sigma(\beta, \nu)$ and $\Sigma(\beta, \nu)$ respectively. From the properties of the solutions of the dual problem, we know that $F_n(x, y) \geq 0$ for all $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ and $F_n = 0$ $\gamma_n$-a.e., similar properties hold true for $\gamma$ and $F$, i.e., $F \geq 0$ everywhere and $F = 0$ $\gamma$-a.e. In the sequel, to simplify the notations, we shall equate the normalizing constants $c$ and $c_n$ to unity. Let $D_f = \{x \in \mathbb{R}^d : f(x) < \infty\}$, we have

\[
\int \psi_n d\gamma = \int \psi_n e^{-f} d\beta = \int \psi_n e^{-(f-f_n)} d\nu_n
\]

\[
= \int \psi_n e^{-(f-f_n)} d\gamma_n = \int \left[ -\varphi_n(x) - \frac{1}{2} |x - y|^2 \right] e^{-(f-f_n)} d\gamma_n
\]

\[
= -\int \varphi_n(x) e^{-(f(y)-f_n(y))} d\gamma_n - \int \frac{1}{2} |x - y|^2 e^{-(f-f_n)} d\gamma_n.
\]
Therefore, using the relation \((I \times T_n)\beta = \gamma_n\) and inserting the expression for \(\int \psi_n d\gamma\) that we have calculated just above, we get
\[
\int F_n d\gamma = \int \psi_n d\gamma + \int \varphi_n d\beta + \frac{1}{2} \int |x - y|^2 d\gamma(x, y)
\]
\[
= \int \varphi_n(x) \left(1 - e^{-(f(y) - f_n(y))}\right) d\gamma_n(x, y) - \frac{1}{2} \int |\nabla \varphi_n|^2 (1 - e^{-(f_n)\circ T_n}) d\beta
\]
\[
= I_n + II_n,
\]
written in the respective order. We have
\[
I_n = \int_{D_f} \varphi_n(x) \left(1 - e^{-(f(y) - f_n(y))}\right) d\gamma_n(x, y)
\]
\[
+ \int_{D_f^c} \varphi_n(x) d\gamma_n(x, y).
\]
Since \((\varphi_n, n \geq 1)\) is bounded in \(L^2(\beta)\) and as
\[
(4.21) \quad \lim_{n \to \infty} 1_{T_n^{-1}(D_f)}(f \circ T_n - f_n \circ T_n) = 0
\]
in \(L^0(\beta)\) and \(\lim_n \beta(T_n^{-1}(D_f^c)) = 0\) as well, we conclude that \(\lim_n I_n = 0\). To calculate \(\lim_n II_n\), we proceed similarly: we have
\[
\int |x - y|^2 e^{-(f - f_n)(y)} d\gamma_n = \int_{T_n^{-1}(D_f)} |\nabla \varphi_n|^2 e^{-(f - f_n)\circ T_n} d\beta
\]
\[
= \int |\nabla \varphi_n|^2 e^{-(f - f_n)\circ T_n} d\beta.
\]
From the relation (4.21), the exponential term converges to 1 in probability (i.e., in \(L^0(\beta)\)) boundedly. It remains to show that the sequence \(|\nabla \varphi_n|^2, n \geq 1\) is uniformly integrable (w.r.t. \(\beta\)). By the triangle inequality, it suffices to prove that the sequence of functions \(|T_n(x)|^2, n \geq 1\) is \(\beta\)-uniformly integrable and to achieve this, using the notation \(L_n = e^{-f_n}\), we write:
\[
\int_{\{|T_n| > c\}} |T_n|^2 d\beta = \int_{\{|x| > c\}} |x|^2 L_n d\beta
\]
\[
\leq \int_{\{|x| > c\}} e^{c|x|^2} d\beta + \frac{1}{\varepsilon} \int_{\{|x| > c\}} L_n \log L_n d\beta,
\]
the first term at the last line converges to zero as \(c \to \infty\) from Fernique’s Lemma (cf. [18, 19]). It is also easy to see that \((L_n \log L_n, n \geq 1)\) is \(\beta\)-uniformly integrable, hence
\[
\lim_{c \to \infty} \sup_n \int_{\{|T_n| > c\}} |T_n|^2 d\beta = 0.
\]
Finally we see that
\[
\lim_{n \to \infty} \int F_n(x, y) d\gamma(x, y) = \lim_n (I_n + II_n) = 0.
\]
Since \(F_n \geq 0\) for any \(n \geq 1\), \(\lim \int F_n d\gamma = 0\) implies the convergence of \((F_n, n \geq 1)\) in \(L^1(\gamma)\). As \((\varphi_n, n \geq 1)\) is bounded in \(L^2(\beta)\), we can choose a weak cluster point of it, say \(a'\), as \((\psi_n, n \geq 1)\) is \(\gamma\)-uniformly integrable, it has also a weak cluster point, say \(b'\). This implies that
\[
a'(x) + b'(y) + \frac{1}{2} |x - y|^2 = 0
\]
Lemma 4, although it turns the density of the derivative operator, and we have
\[ \xi \] have
\[ L \]
converges weakly in \( L^2(\gamma) \) and in \( L^1(\gamma) \). Define \( a \) and \( b \) respectively as \( a = \lim sup_n \phi_n', \ b = \lim sup_n \phi_n' \), then, \( a = a', \ b = b' \) \( \gamma \)-a.s. and
\[ a(x) + b(y) + \frac{1}{2} |x - y|^2 \geq 0 \]
for all \((x, y) \in \mathbb{R}^d \times \mathbb{R}^d \) and
\[ a(x) + b(y) + \frac{1}{2} |x - y|^2 = 0 \]
\( \gamma \)-a.s. From the uniqueness (up to constants) of the Monge potential functions, it follows that
\( a = \varphi \) and \( b = \psi \) \( \gamma \)-a.s. Moreover, the same construction works for any cluster points of the sequences \( (\varphi_n) \) and \( (\psi_n) \), consequently these two sequences have each a unique cluster point and this proves the convergence in \( L^2(\gamma) \) of \( (\varphi_n, n \geq 1) \) and the convergence in \( L^1(\gamma) \) of \( (\psi_n, n \geq 1) \). Since \( \lim_n E_\beta[[\nabla \varphi_n]^2] = E_\beta[[\nabla \varphi]^2] \), \( (\varphi_n, n \geq 1) \) converges to \( \varphi \) in \( D^2_{2,1} \). To show the convergence of \( \nabla \psi_n \), we have, as \( \gamma(D_f) = 1 \),
\[
E_\gamma[[\nabla \psi_n]^2] = E_\gamma[1_{D_f}\nabla \psi_n]^2 \]
\[
= E_\nu[1_{D_f}\nabla \psi_n]^2 \]
\[
= E_\nu [1_{D_f} \nabla \psi_n]^2 e^{-(f-f_n)} \]
\[
= E_\beta [1_{D_f} \circ T_n \nabla \varphi_n]^2 e^{-(f-f_n) \circ T_n} \].
As \( \nabla \varphi_n \to \nabla \varphi \) in \( L^2(\beta) \) and as \( 1_{D_f} \circ T_n \exp[-(f-f_n) \circ T_n] \to 1 \) in probability as \( n \to \infty \), we conclude that
\[
\lim_n E_\gamma[[\nabla \psi_n]^2] = \lim_n E_\gamma[1_{D_f}\nabla \psi_n]^2 \]
\[
= \lim_n E_\beta [1_{D_f} \circ T_n \nabla \varphi_n]^2 e^{-(f-f_n) \circ T_n} \]
\[
= \lim_n E_\beta[[\nabla \varphi_n]^2] = E_\beta[[\nabla \varphi]^2] \]
\[
= E_\nu[[\nabla \psi]^2] = E_\gamma[[\nabla \psi]^2].
\]
Assume that \( \xi \in L^2(\nu) \) be any weak cluster point of the sequence \( (\nabla \psi_n, n \geq 1) \), i.e., let \( \xi = \lim_n \nabla \psi_{n_k} \). As the derivative operator is closed in \( L^2(\nu) \) due to the hypothesis \( 1.20 \), we should have \( \xi = \nabla \psi \), where the latter is defined in the Sobolev sense. Consequently \( (\nabla \psi_n, n \geq 1) \) converges to \( \nabla \psi \) in \( L^2(\nu) \).

The proof of the following is exactly as the proof of Corollary 1 hence it is omitted:

**Corollary 2.** Assume the hypothesis of Lemma 4 are valid and assume that \( (\nabla^2 \psi_n \circ T_n, n \geq 1) \) converges weakly in \( L^2(\beta, \mathbb{R}^d \otimes \mathbb{R}^d) \), then \( \nabla^2 \psi \in L^2(\nu, \mathbb{R}^d \otimes \mathbb{R}^d) \), where \( \nabla \) is the closure in \( L^2(\nu) \) of the derivative operator, and we have
\[
w - \lim_n \nabla^2 \psi_n \circ T_n = \nabla^2 \psi \circ T.
\]

Lemma 4 although it turns the density \( L = e^{-f} \) into a non-degenerate one, it spoils the convexity of \( f \), in other words the log-concavity of \( L \). The following results heals this default:
Lemma 5. Assume that \( L = e^{-f} \) and \( f \) satisfy the properties common to the lemmata of this section. Define \( L_\varepsilon \) as
\[
L_\varepsilon = \frac{\varepsilon + L}{1 + \varepsilon},
\]
where \( \varepsilon > 0 \). Let \( d\nu_\varepsilon = L_\varepsilon d\beta \) and \( d\nu = L d\beta \) then
\[
\sup_{\varepsilon \in [0,1]} H(\nu_\varepsilon | \beta) < \infty.
\]
Moreover the forward and backward Monge potentials \((\varphi_\varepsilon, \psi_\varepsilon)\) associated to the quadratic transport problem on \( \Sigma(\beta, \nu_\varepsilon) \) satisfy the following properties:
- \( \lim_{\varepsilon \to 0} \varphi_\varepsilon = \varphi \) in the Gaussian Sobolev space \( \mathbb{D}_{2,1} \),
- \( \lim_{\varepsilon \to 0} \psi_\varepsilon = \psi \) in \( L^1(\nu) \) and \( \lim_{\varepsilon \to 0} \nabla \psi_\varepsilon = \nabla \psi \) in \( L^2(\nu) \).

Proof: Since
\[
L_\varepsilon = \frac{\varepsilon}{1 + \varepsilon} + \frac{1}{1 + \varepsilon} L,
\]
from the convexity of the function \( x \to x \log x \), we get
\[
H(\nu_\varepsilon | \beta) = \int_{\mathbb{R}^d} L_\varepsilon \log L_\varepsilon d\beta \\
\leq \frac{1}{1 + \varepsilon} \int_{\mathbb{R}^d} L \log L d\beta,
\]
hence
\[
(4.22) \quad \sup_{\varepsilon \in [0,1]} H(\nu_\varepsilon | \beta) \leq H(\nu | \beta).
\]
Let now \( \gamma_\varepsilon \) and \( \gamma \) be optimal transport plans associated to \( \Sigma(\beta, \nu_\varepsilon) \) and \( \Sigma(\beta, \nu) \) respectively, and let \( F_\varepsilon(x, y) = \varphi_\varepsilon(x) + \psi_\varepsilon(y) + \frac{1}{2} |x - y|^2 \), where, as usual we replace \( \varphi_\varepsilon \) and \( \psi_\varepsilon \) by \( \varphi_\varepsilon - E_\beta [\varphi_\varepsilon] \) and \( \psi_\varepsilon + E_\beta [\varphi_\varepsilon] \) respectively. We have \( F_\varepsilon(x, y) \geq 0 \) for all \( x, y \in \mathbb{R}^d \) and \( F_\varepsilon = 0 \) \( \gamma_\varepsilon \)-a.s. We claim that
\[
(4.23) \quad \lim_{\varepsilon \to 0} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} F_\varepsilon(x, y) d\gamma(x, y) = 0.
\]
At first as in the proof of Lemma 2, \( \lim_{\varepsilon \to 0} \int |y|^2 d\nu_\varepsilon = \int |y|^2 d\nu \), hence we get
\[
\lim_{\varepsilon \to 0} E[|\nabla \varphi_\varepsilon|^2] = \lim_{\varepsilon \to 0} d^2_W(\beta, \nu_\varepsilon) = d^2_W(\beta) = E[|\nabla \varphi|^2].
\]
To prove (4.23) we write first
\[
\int F_\varepsilon d\gamma + \varepsilon \int \varphi_\varepsilon d\beta = \int_\mathbb{R}^d \psi_\varepsilon(\varepsilon + L) d\beta + \frac{1}{2} \int_\mathbb{R}^d |x - y|^2 d\gamma(x, y)
\]
\[
= (1 + \varepsilon) \int \psi_\varepsilon d\nu_\varepsilon + \frac{1}{2} \int |x - y|^2 d\gamma(x, y).
\]
As
\[
\lim_{\varepsilon \to 0} (1 + \varepsilon) \int \psi_\varepsilon d\nu_\varepsilon \quad = \quad - \lim_{\varepsilon \to 0} \frac{1}{2} \int |x - y|^2 d\gamma_\varepsilon(x, y)
\]
\[
= \quad - \lim_{\varepsilon \to 0} d^2_W(\beta, \nu_\varepsilon) = -d^2_W(\beta, \nu)
\]
\[
= \quad -1 \int |x - y|^2 d\gamma(x, y),
\]
18
we conclude that
\begin{equation}
\lim_{\varepsilon \to 0} \left( \int F_\varepsilon d\gamma + \varepsilon \int \psi_\varepsilon d\beta \right) = 0.
\end{equation}
Recall that, $F_\varepsilon \geq 0$ everywhere in $\mathbb{R}^d \times \mathbb{R}^d$, hence it is also positive on the diagonal, i.e., $\varphi_\varepsilon(x) + \psi_\varepsilon(x) \geq 0$ for any $x \in \mathbb{R}^d$, if $\int \varphi_\varepsilon d\beta = 0$, we have $\int \psi_\varepsilon d\beta \geq 0$, consequently the relation (4.23) is proved. Note that, due to the inequality (4.22) and the Talagrand inequality, the set $(\varphi_\varepsilon, \varepsilon \in [0, 1])$ is bounded in $D_{2,1}$. The rest of the proof goes exactly as the proofs of other lemmata of this section, i.e., use of Mazur’s Lemma, etc., hence it is omitted.

We have also the obvious corollary of Lemma 5 whose proof is very similar to that of Corollary 1.

**Corollary 3.** Assume the hypothesis of Lemma 4 are valid and assume, with $\varepsilon = 1/n$, that $(\nabla^2 \psi_n \circ T_n, n \geq 1)$ converges in $L^2(\beta, \mathbb{R}^d \otimes \mathbb{R}^d)$, then $\nabla^2 \psi \in L^2(\nu, \mathbb{R}^d \otimes \mathbb{R}^d)$, where $\nabla$ is the closure in $L^2(\nu)$ of the derivative operator, and we have
\[
\lim_n \nabla^2 \psi_n \circ T_n = \nabla^2 \psi \circ T.
\]

5. **Sobolev regularity**

To approximate the Monge potentials constructed on the Wiener space $(W, H, \mu)$ (cf. [8, 9]), as explained in preliminaries, we choose any complete orthonormal basis $(e_n, n \geq 1)$ in the Cameron-Martin space $H$ and construct the increasing sequence of $\sigma$-algebras $(V_n, n \geq 1)$, where $V_n$ is the $\sigma$-algebra generated by the Gaussian random variables $\{\delta e_1, \ldots, \delta e_n\}$.

The following lemma is placed in this section although it is an approximation result, since it is more interesting in the infinite dimensional case and it is proved in [9], we recall it here for the sake of completeness:

**Lemma 6.**

1. Let $L \in L^1_+, \nu = Ld\mu = e^{-f}d\mu$. Assume that $E[L|\nabla f|^2] < \infty$ and denote by $(\varphi, \psi)$ forward and backward potentials corresponding to the transport problem from $\mu$ to $\nu$. Let $\psi_n$ be defined as $\psi_n = L_n d\mu$, $L_n = E[L|V_n]$ and denote by $(\varphi_n, \psi_n)$ the Monge potentials corresponding to $(\mu, \nu_n)$. Then $(\varphi_n, n \geq 1)$ converges to $\varphi$ in $D_{2,1}$, $(\psi_n, n \geq 1)$ converges to $\psi$ in $L^1(\nu)$ and $(\nabla \psi_n, n \geq 1)$ converges to $\nabla \psi$ in $L^2(\nu, H)$.
2. In particular, letting $T_n$ and $T$ be the forward transport maps defined as $T_n = I_W + \nabla \varphi_n$, $T = I_W + \nabla \varphi$, if $(\nabla^2 \psi_n \circ T_n, n \geq 1)$ converges weakly in $L^2(\mu, H \otimes H)$, then, $\nabla^2 \psi \in L^2(\nu, H \otimes H)$ and we have
\[
w - \lim_n \nabla^2 \psi_n \circ T_n = \nabla^2 \psi \circ T,
\]
$\mu$-almost surely.

**Proof:** Only the second part requires a proof but it is immediate due to the closability of the $H$-Gateaux derivative w.r.t. the target measure $\nu$.

We then define the smooth target measures approximating the original one in three steps:

1. $d\nu_n = L_n d\mu$ as $L_n = E[e^{-f}|V_n] = e^{-f_n}$,
2. $d\nu_{n,k} = L_{n,k} d\mu = e^{-f_{n,k}} d\mu = c_{n,k} e^{-f_{n,k}} d\mu$,
3. $d\nu_{n,k,l} = L_{n,k,l} d\mu = e^{-f_{n,k,l}} d\mu = P_{j/l}(e^{-f_{n,k}}) d\mu$, 

19
where \((P_t, t \geq 0)\) denotes the Ornstein-Uhlenbeck semigroup on the Wiener space, whose version on \(\mathbb{R}^d\) was denoted by \((Q_t, t \geq 0)\) in the preceding pages. Let us note again the relation which will be used without any warning in the sequel: if \(G : W \to Z\) is a Wiener function in \(L^p(\mu, Z)\), where \(Z\) is a separable Hilbert space, then \(E[G|V_n]\) can be represented as \(g_n(\delta e_1, \ldots, \delta e_n)\), where \(g_n : \mathbb{R}^n \to Z\) is in \(L^p(\mu_n, Z)\), and \(\mu_n\) is the Gauss measure on \(\mathbb{R}^n\). In this situation we have

\[
E[P_tG|V_n] = (Q_tg_n)(\delta e_1, \ldots, \delta e_n)
\]

\(\mu\)-a.s., where \((Q_t, t \geq 0)\) is the Ornstein-Uhlenbeck defined on \(\mathbb{R}^n\). Consequently \(L_{n,k,l}\) can be written as \(e^{-\tilde{f}_n(\delta e_1, \ldots, \delta e_n)}\) where \(\tilde{f}_n : \mathbb{R}^n \to \mathbb{R}\) is a smooth, lower and upper bounded function. Consequently the classical results of transport like \([3, 23]\) affirm the existence of smooth Monge potentials corresponding to the Monge-Kantorovitch problem on the set \(\Sigma(\mu, \nu)\). We shall denote by \((\varphi, \psi, \varphi_n, \psi_n)\) respectively. Recall that, from the series of lemmas of Section \(4\), we know that \(\lim_{n \to \infty} \varphi_n, \varphi_n = \varphi\) in the Sobolev spaces \(\mathbb{D}_{2,1}\). For the dual or backward potentials the situation is more involved. We begin with

\begin{lemma}
Let \(T_{nkl}\) be the transport map defined as \(T_{nkl} = I_W + \nabla \varphi_{nkl}\), then we have

\[
\lim_{n \to \infty} \lim_{k \to \infty} \lim_{l \to \infty} \nabla f_{nkl} \circ T_{nkl} = \nabla f \circ T
\]

in \(L^2(\mu, H)\).
\end{lemma}

\begin{proof}
We have

\[
E \left[ \nabla f_{nkl}^2 \right] = 4E \left[ \frac{1}{2} \frac{\nabla P_{1/t}(e^{-f_{nkl}})}{P_{1/t}(e^{-f_{nkl}})} \right]_H^2
\]

\[
= 4E \left[ \frac{1}{2} \frac{\nabla P_{1/t}(e^{-f_{nkl}})}{P_{1/t}(e^{-f_{nkl}})}}^{1/2} \right]_H^2
\]

\[
= E \left[ \frac{P_{1/t}(e^{-f_{nkl}})}{P_{1/t}(e^{-f_{nkl}})}}^{1/2} \right]_H^2
\]

\[
= E \left[ \frac{P_{1/t}(\nabla e^{-f_{nkl}})^2}{P_{1/t}(e^{-f_{nkl}})}} \right]_H^2
\]

\[
\leq e^{-2/t} E \left[ \frac{P_{1/t}(\nabla e^{-f_{nkl}})^2}{P_{1/t}(e^{-f_{nkl}})}} \right]_H^2
\]

\[\text{In the sequel we shall often omit the commas between the lower indices for typographical reasons.}\]

20
as, due to the local character of the Sobolev derivative, \( \nabla f_{nk} = 0 \) \( \mu \)-a.s. on the set \( \{ f_{nk} = k \} = \{ f_n \geq k \} \), the last line of the above inequality can be upper bounded as follows

\[
\begin{align*}
& e^{-2l/4}E \left[ P_{1/l}(1(f_{nk} \leq k) e^{-f_{nk}^2}) \left( P_{1/l}(e^{-f_{nk}^2}) \right)^{-1} \right] \\
\leq & \ e^{-2l/4}E \left[ P_{1/l}(1(f_{nk} \leq k) e^{-f_{nk}^2}) (P_{1/l}(e^{-f_{nk}^2}))^{-1} \right] \\
\leq & \ E \left[ P_{1/l}(1(f_{nk}^2 e^{-f_{nk}^2} 1(f_{nk} \leq k))) \frac{P_{1/l}(e^{-f_{nk}^2} 1(f_{nk} \leq k))}{P_{1/l}(e^{-f_{nk}^2})} \right] \\
\leq & \ e^{-2l/4}E \left[ |\nabla \psi_{nk}^2| e^{-f_{nk}^2} \right] \\
= & \ e^{-2l/4}E\left[ \frac{|\nabla \psi_{nk}^2| e^{-f_{nk}^2}}{E[e^{-f}]} \right].
\end{align*}
\]

Therefore \( \nabla \psi_{nk}^2, n \geq 1; k \geq 1; l \geq 1 \) is bounded in \( L^2(\mu, H) \). As \( \lim_{n,k,l} f_{nk} = f \) in \( L^0(\mu, \mathbb{N}) \) and as \( E[|e^{-f_{nk}^2/2}|^2] \rightarrow E[|e^{-f/2}|^2] \), \( \nabla e^{-f_{nk}^2/2}, n \geq 1; k \geq 1; l \geq 1 \) converges weakly to \( \nabla e^{-f/2} \) in \( L^2(\mu, H) \). Hence

\[
E[|\nabla e^{-f/2}|^2] \leq \lim \inf_{n,k,l} E[|\nabla e^{-f_{nk}^2/2}|^2] \\
\leq \lim \sup_{n,k,l} E[|\nabla e^{-f_{nk}^2/2}|^2] \\
\leq E[|\nabla e^{-f/2}|^2],
\]

from the above calculations, consequently the strong convergence in \( \mathbb{D}_{2,1} \) holds true. Finally we have

\[
\begin{align*}
\lim_{n,k,l} E[|\nabla f_{nk} \circ T_{nk}|^2] & = \lim_{n,k,l} E[|\nabla f_{nk}^2 e^{-f_{nk}}|^2] \\
& = E[|\nabla f|^2 e^{-f}] = E[|\nabla f \circ T|^2],
\end{align*}
\]

moreover \( \lim_{n,k,l} \nabla f_{nk} \circ T_{nk} = \nabla f \circ T \) in \( L^0(\mu) \) (i.e., in probability), hence the \( L^2 \)-convergence follows.

**Theorem 4.** The sequence \( (\nabla^2 \psi_{nk} \circ T_{nk} \colon n \geq 1; k \geq 1; l \geq 1) \) converges to \( \nabla^2 \psi \circ T \) weakly in \( L^2(\mu, H \otimes H) \) as \( l \to \infty \), then as \( k \to \infty \) and then as \( n \to \infty \).

**Proof:** The proof follows from applications of Corollary [1], Corollary [2] and Lemma [3].

**Lemma 8.** The set of functions \( \log L_{nk} \circ T_{nk} : n, k, l \in \mathbb{N} \) is uniformly integrable.

---

\footnote{Here the order of the limits is important}
Proof: We claim first that \((L_{nkl} \log L_{nkl} : n, k, l \in \mathbb{N})\) is uniformly integrable: in fact, it follows from the Jensen inequality

\[
L_{nkl} \log L_{nkl} \leq P_{1/l} E[L_k \log L_k | V_n] \\
\leq P_{1/l} E[L \log L 1_{\{f \leq k\}} + k e^{-k} 1_{\{f > k\}} | V_n] \\
\leq P_{1/l} E[L \log L | V_n] + 1,
\]

as \(L \log L \in L^1(\mu)\), the set \((P_{1/l} E[L \log L | V_n]; n, l \in \mathbb{N})\) is uniformly integrable, therefore \((L_{nkl} \log L_{nkl} : n, k, l \in \mathbb{N})\) is also uniformly integrable. To complete the proof it suffices to see that

\[
E[\log L_{nkl} \circ T_{nkl} 1_{\{\log L_{nkl} \circ T_{nkl} > c\}}] = E[L_{nkl} \log L_{nkl} 1_{\{\log L_{nkl} > c\}}] \\
= E[L_{nkl} \log L_{nkl} 1_{\{\log L_{nkl} > c\}}] \\
= E[L_{nkl} \log L_{nkl} 1_{\{\log L_{nkl} > c\}}] \to 0
\]
as \(c \to \infty\) uniformly w.r.t. \(n, k, l \in \mathbb{N}\) by the uniform integrability of \((L_{nkl} \log L_{nkl} : n, k, l \in \mathbb{N})\).

\[\square\]

Theorem 5. The sequence \((\mathcal{L} \psi_{nkl} \circ T_{nkl}, n, k, l)\) converges to \(\mathcal{L} \circ T\) in the sense of distributions. Moreover, \(\mathcal{L} \psi \circ T\) is in fact an element of \(L^1(\mu)\).

Proof: We shall prove only the convergence of a subsequence, which is chosen by a double diagonalization method. The proof of convergence in the sense of distributions is straightforward by duality. To show that \(\mathcal{L} \psi \circ T \in L^1(\mu)\) or equivalently that \(\mathcal{L} \psi \in L^1(\mu)\) is more delicate: as \((\nabla^2 \psi_n \circ T_n, n \geq 1)\) converges weakly in \(L^2(\mu, H \otimes H)\), from Mazur’s Theorem, we can form a sequence of its convex combinations, denoted as \((\nabla^2 \psi_n' \circ T_n', n \geq 1)\), which converges strongly in \(L^2(\mu, H \otimes H)\), in particular, the sequence

\[
(\|\nabla^2 \psi_n' \circ T_n'\|_2^2, n \geq 1)
\]
is uniformly integrable in \(L^1(\mu)\). Let us denote by \((\mathcal{L} \psi_n' \circ T_n', n \geq 1)\) the corresponding convex combinations of \((\mathcal{L} \psi_n \circ T_n, n \geq 1)\). We have

\[
\mathcal{L} \psi_n' \circ T_n' \geq \sum_i \lambda_i \left[ -\log L_{n_i} \circ T_{n_i} - \frac{1}{2} \|\nabla \psi_{n_i} \circ T_{n_i}\|_H^2 - \frac{1}{2} \|\nabla^2 \psi_{n_i} \circ T_{n_i}\|_2^2 \right] \\
\geq \sum_i \lambda_i \left[ -\log L_{n_i} \circ T_{n_i} - \frac{1}{2} \|\nabla \psi_{n_i} \circ T_{n_i}\|_H^2 - \frac{1}{2} \|\nabla^2 \psi_{n_i} \circ T_{n_i}\|_2^2 \right],
\]

where we have used the inequality 
|\text{det} \nabla^2 (I + A)| \leq \frac{1}{2} \|A\|_2^2, for any Hilbert-Schmidt operator \(A\). From Lemma\[\text{Lemma}\[\text{Lemma}\] the sequence \((\mathcal{L} \psi_n' \circ T_n', n \geq 1)\) is lower bounded by a uniformly integrable sequence. We also have

\[
\mathcal{L} \psi_n' \circ T_n' = \sum_i \lambda_i \mathcal{L} \psi_{n_i} \circ T_{n_i},
\]

\[
\leq \sum_i \lambda_i \left[ -\log L_{n_i} \circ T_{n_i} + \log \text{det} \nabla^2 (I + \nabla^2 \psi_{n_i} \circ T_{n_i}) - \frac{1}{2} \|\nabla \psi_{n_i} \circ T_{n_i}\|_H^2 \right] \\
\leq \sum_i \lambda_i \left[ -\log L_{n_i} \circ T_{n_i} - \frac{1}{2} \|\nabla \psi_{n_i} \circ T_{n_i}\|_H^2 \right].
\]
Hence, from again Lemma 8 the sequence \( (\mathcal{L}\psi'_n \circ T'_n, n \geq 1) \) is also upperbounded by a uniformly integrable sequence, therefore it is itself uniformly integrable. Consequently, \( (\mathcal{L}\psi'_n \circ T'_n, n \geq 1) \) converges weakly in \( L^1(\mu) \) and the limit is equal to \( \mathcal{L}\psi \circ T \).

Remark:

6. Disintegration Results

Let \((\varphi, \psi)\) be the forward and backward potentials corresponding to the solution of the Monge-Kantorovitch problem corresponding to \((\mu, \nu)\), let \(\pi_n : H \to H_n\) be an orthogonal projection from \(H\) to a finite dimensional subspace of \(H\), denoted by \(H_n\), we may assume the existence of a complete orthonormal basis of \(H\), denoted \(\{e_k, k \geq 1\} \subset W^*\) with \(H_n = \text{span}\{e_1, \ldots, e_n\}\). As each \(e_i \in W^*\), \(\pi_n\) has a continuous extension to the whole space \(W\) that we shall denote again with the same notation. The following result has been proved in [9] (Lemma 6.2), in much more general case:

**Lemma 9.** \(\gamma \in \Sigma(\mu, \nu)\) be the optimal measure (i.e., the transport plan). Let \(\pi_n\) be defined as above and let \(\pi_n^\perp = I_W - \pi_n\). Define \(p_n\) as the projection from \(W \times W\) onto \(H_n\) with \(p_n(x, y) = \pi_n x\) and let \(p_n^\perp(x, y) = \pi_n^\perp x\). Consider the Borel disintegration

\[
\gamma(\cdot) = \int_{H_n^* \times W} \gamma(\cdot | x_n^\perp) \gamma_n^\perp(dz_n^\perp)
\]

along the projection of \(W \times W\) on \(H_n^*\), where \(\mu_n^\perp\) is the measure \(\pi_n^\perp \mu\), \(\gamma(\cdot | x_n^\perp)\) denotes the regular conditional probability \(\gamma(\cdot | p_n^\perp = x_n^\perp)\) and \(\gamma_n^\perp\) is the measure \(p_n^\perp \gamma\). Then, \(\mu_n^\perp\) and \(\gamma_n^\perp\) almost surely \(\gamma(\cdot | x_n^\perp)\) is optimal on \((x_n^\perp + H_n) \times W\), in the sense that it realizes the quadratic Wasserstein distance between the measures \(P_1 \gamma(\cdot | x_n^\perp)\) and \(P_2 \gamma(\cdot | x_n^\perp)\), where \(P_i\) are the projection maps defined on \((x_n^\perp + H_n) \times W\) by \(P_i(\xi_1, \xi_2) = \xi_i, i = 1, 2\).

The following result is just the Bayes’ formula of classical probability:

**Lemma 10.** For any \(g \in C_b(W)\), we have

\[
E_\mu \left[ g \circ T | \pi_n^\perp = x_n^\perp \right] = \frac{E_\mu \left[ gL | \pi_n^\perp = x_n^\perp \right]}{E_\mu \left[ L | \pi_n^\perp = x_n^\perp \right]} = E_\nu \left[ g | \pi_n^\perp = x_n^\perp \right],
\]

in other words the image of the regular conditional probability measure \(\mu(\cdot | \pi_n^\perp = x_n^\perp)\) under the transport map \(x_n \to x_n + \pi_n \varphi(x_n^\perp + x_n)\) is equal to the regular conditional probability measure \(\nu(\cdot | \pi_n^\perp = x_n^\perp)\) and this latter measure is absolutely continuous w.r.t. \(\mu(\cdot | \pi_n^\perp = x_n^\perp)\) with the corresponding Radon-Nikodym density given by

\[
x_n \to \frac{L(x_n + x_n^\perp)}{E_\mu \left[ L | \pi_n^\perp = x_n^\perp \right]}.\]
Proof: For typographical simplicity, we shall omit the lower index “n” in the proof. Let now \( g, h \in C_b(W) \), then from the very definition of the conditional probability, we have
\[
\int g \circ T(x^+ + x) \mu(dx|x^+) h(x^+) \mu^+(dx^+) = \int E_\mu[g \circ T|\pi^+ = x^+] h(\pi^+ x) \mu(dx)
\]
\[
= \int g \circ T(x) h(\pi^+ x) \mu(dx)
\]
\[
= \int (g \circ T)(h \circ \pi^+ \circ S \circ T) d\mu
\]
\[
= \int g(x) h(\pi^+ S(x)) L(x) \mu(dx)
\]
\[
= \int E_\mu[g L|\pi^+ S] h(\pi^+ S) d\mu
\]
\[
= \int E_\mu[g L|\pi^+=x^+] \frac{L}{E_\mu[L|\pi^+=x^+]} h(\pi^+ S) d\mu
\]
\[
= \int E_\mu[g L|\pi^+=x^+] \frac{L}{E_\mu[L|\pi^+=x^+]} h(x^+) d\mu^+
\]
and the proof follows.

\[\square\]

**Lemma 11.** Let \( t(w^+_n, \cdot) : H_n \rightarrow H_n \) be defined as \( t(w^+_n, x) = x + \pi_n \nabla \varphi(w^+_n + x) \) and let \( s(w^+_n, \cdot) : H_n \rightarrow H_n \) be defined as \( s(w^+_n, x) = x + \pi_n \nabla \psi(w^+_n + x) \). Then it holds that, \( \mu^+_n \)-almost everywhere \( \mu_n^+ \)-almost everywhere
\[
s(w^+_n, t(w^+_n, x)) = x \quad \mu(\cdot | \pi_n^+ = w_n^+) \text{ - a.s.}
\]
\[
t(w^+_n, s(w^+_n, x)) = x \quad \nu(\cdot | \pi_n^+ = w_n^+) \text{ - a.s.}
\]

In particular, we have, for \( \mu_n^+ \)-almost all \( w_n^+ \),
\[
\frac{L(w^+_n + x)}{E_\mu[L|\pi_n^+ = w_n^+]} = \det_2(I_{H_n} + D^2 \psi(w^+_n + x)) \exp[-\delta \pi_n \nabla \psi(w^+_n + x) - \frac{1}{2} |\pi_n \nabla \psi(w^+_n + x)|^2],
\]
\( \mu(\cdot | \pi_n^+ = w_n^+) \) almost surely.

**Proof:** Note that we have
\[
\pi_n \nabla \varphi \circ T(w) = -\pi_n \nabla \varphi(w) = -D\varphi(w^+_n + x)
\]
\( \mu \text{ a.s.}, \) where \( D \) is the derivative on \( H_n \) which is regarded as the Euclidean space \( \mathbb{R}^n \), \( w^+_n = w - \pi_n(w) = w - x \). The left hand side of the relation (6.26) can also be written as\( D \psi(w^+_n + t(w^+_n, x)) \),

hence we get, for \( \mu_n^+ \)-almost all \( w_n^+ \),
\[
D\psi(w^+_n + t(w^+_n, x)) = -D\varphi(w^+_n + x)
\]
\( \mu(\cdot | \pi_n^+ = w_n^+) \)-almost surely. Consequently the partial maps \( t(w^+_n, \cdot) \) and \( s(w^+_n, \cdot) \) are inverse to each other on \( H_n \), consequently the representation of the density follows from finite dimensional results about the derivatives of the convex functions and their Legendre transformations corresponding to
the Monge-Kantorovitch problem corresponding to the measures which are absolutely continuous w.r.t. Lebesgue measure (cf. [15, 23]).

\[ \int L(\pi_n^+ + \pi_n^-(w)) \frac{\mu}{E_\mu[L|\pi_n^+]} \]

\[ = \det_2(I + \pi_n \nabla^2 \psi \pi_n(\pi_n^+(w) + x) \exp \left[ -\delta \pi_n \psi(\pi_n^+(w) + x) - \frac{1}{2} |\pi_n \nabla \psi(\pi_n^+(w) + x)|^2 \right] \]

\[ \mu_n^+ \times \mu_n = \mu - \text{almost surely}. \]

**Lemma 12.** The sequence \((\delta(\pi_n \nabla \psi), n \geq 1)\) is \(\nu\)-uniformly integrable and it converges to \(L \psi\) in \(L^1(\nu)\). In particular

\[ \lim_{n \to \infty} \frac{L(\pi_n^+(w) + \pi_n^-(w))}{E_\mu[L|\pi_n^+]} = \det_2(I + \pi_n \nabla^2 \psi(w)) \exp \left[ -\mathcal{L}_\psi(w) - \frac{1}{2} |\nabla \psi(w)|^2 \right] \]

\(\nu\)-a.s.

**Proof:** Let

\[ \lambda_n(w_n, w_n^+) = \frac{L(w_n^+ + w_n)}{E_\mu[L|\pi_n^+]}, \]

from Martingale Convergence Theorem of Doob, \((\lambda_n(\pi_n^+(w) + \pi_n(w)), n \geq 1)\) converges to \(L\) on \(D = \{L > 0\}\) \(\mu\)-a.s., or \(\nu\)-a.s. Therefore \((\delta(\pi_n \nabla \psi), n \geq 1)\) converges \(\mu\)-a.s. on \(D\) (or \(\nu\)-a.s.). We have, by writing \(w = \pi_n(w) + \pi_n^-(w)\), from Lemma 11

\[ -\log \lambda_n(w) = -\log \det_2(I + \pi_n \nabla^2 \psi \pi_n) + \delta \pi_n \nabla \psi + \frac{1}{2} |\pi_n \nabla \psi|_H^2 \]

\(\mu\)-a.s. on the set \(D\). Consequently we can write

\[ \delta \pi_n \nabla \psi = -\log L + \log E[L|\pi_n^+] + \log \det_2(I + \pi_n \nabla^2 \psi \pi_n) - \frac{1}{2} |\pi_n \nabla \psi|_H^2 \]

\(\nu\)-a.s. and \(\mu\)-a.s. on the set \(D\). To show the uniform integrability, we write, for any \(A \in B(W)\)

\[ E_\nu[1_A |\delta \pi_n \nabla \psi|] \leq E_\nu[-1_A \log L] + E_\nu[1_A \log E[L|\pi_n^+]] + \frac{1}{2} E_\nu[1_A \|\pi_n \nabla^2 \psi \pi_n\|^2] + \frac{1}{2} E_\nu[\pi_n \nabla \psi|_H^2] \]

\[ = I + II_n + III_n + IV_n \]

As \(E_\mu[\log L] < \infty\), we have \(I < \varepsilon/4\), provided \(\mu(A) < \delta\). For \(II_n\) we have, from Jensen inequality

\[ E_\nu \left[ 1_A \log E_\mu[L|\pi_n^+] \right] = E_n \left[ E_\nu[1_A |\pi_n^+|] \log E_\mu[L|\pi_n^+] \right] \]

\[ = E_\mu \left[ E_\nu[1_A |\pi_n^+|] E_\mu[L|\pi_n^+] \log E_\mu[L|\pi_n^+] \right] \]

\[ \leq E_\mu \left[ E_\nu[1_A |\pi_n^+|] L \log L \right] \]

as \(E_\nu[\|\log L\|] < \infty\), the sequence \((E_\nu[\log L|\pi_n^+], n \geq 1)\) is \(\nu\)-uniformly integrable, hence \(\sup_n II_n \leq \varepsilon/4\) for \(\mu(A) < \delta\). As \((\|\pi_n \nabla^2 \psi \pi_n\|_2, n \geq 1)\) is a monotone, increasing sequence, it follows from the Monotone Convergence Theorem that \(\sup_n III_n < \varepsilon/4\) for \(\mu(A) < \delta\). The fourth term \(IV_n\) is trivial
to control since \( E_\nu[1_A|\pi_n \nabla \psi|_{H}^2] \leq E_\nu[1_A|\nabla \psi|_{H}^2] \) for any \( n \geq 1 \). Consequently \( (\delta(\pi_n \nabla \psi), n \geq 1) \) is \( \nu \)-uniformly integrable, hence \( \lim_{n \to \infty} \delta(\pi_n \nabla \psi) = L^1(\nu) \).

\[ \nabla E(7.29) \]

7. Calculation of the Jacobians and Monge-Ampère Equation

We are now at a position to express the density \( L \) in terms of the backward potential for the original problem:

**Theorem 6.** The target density has the following representation:

\[
L = \det_2(I + \nabla^2 \psi) \exp \left[ -\mathcal{L} \psi - \frac{1}{2} |\nabla \psi|_{H}^2 \right]
\]

\( \nu \)-almost surely.

**Proof:** Let us extract a sequence \((\Psi_n, n \geq 1)\) from \((\psi_{nkl})\) of the form \( \Psi_n = \psi_{n,k_n,l_n} \) applying twice the diagonal sequence selection. By the uniform integrability results we can assume that \( w - \lim_n \nabla^2 \Psi_n \circ T_n = \nabla^2 \psi \circ T, w - \lim_n \mathcal{L} \Psi_n \circ T_n = \mathcal{L} \psi \circ T \) weakly in \( L^2(\mu, H \otimes H) \) and weakly in \( L^1(\mu) \) respectively, where \( T_n = I_W + \nabla \Phi_n \) and \( \Phi_n = \varphi_{n,k_n,l_n} \). Note that \( \nabla \Psi_n \circ T_n \) converges strongly in \( L^2(\mu, H) \). From Mazur’s theorem, we can construct sequence of convex combinations\(^5\)

\[
\left( \sum_{m_i \geq n} \lambda_i \nabla_2 \Psi_{m_i} \circ T_{m_i}, n \geq 1 \right)
\]

which converges strongly to \( \nabla^2 \psi \circ T \) in \( L^2(\mu, H \otimes H) \), hence

\[
\lim_n \left[ - \log \det_2(I + \sum_{m_i \geq n} \lambda_i \nabla^2 \Psi_{m_i} \circ T_{m_i}) + \frac{1}{2} \sum_{m_i \geq n} \lambda_i \mathcal{L} \Psi_{m_i} \circ T_{m_i} \right] \\
= - \log \det_2(I + \nabla^2 \psi \circ T) + \mathcal{L} \psi \circ T + \frac{1}{2} |\nabla \psi \circ T|_{H}^2 \\
= - \log \Lambda(\psi) \circ T
\]

in the weak topology. From the convexity of the \( A \to - \log \det_2(I_H + A) \), it follows that

\[
- \log \Lambda(\psi) \circ T \leq - \log L \circ T,
\]

hence

\[
\Lambda(\psi) \circ T \geq L \circ T
\]

\( \mu \)-almost surely or \( \Lambda(\psi) \geq L \) \( \nu \)-almost surely. To show that they are equal \( \nu \)-a.e., it suffices to prove that

\[
(7.29) \quad E_\mu[1_D \Lambda(\psi)] \leq 1,
\]

where \( D \) is defined as \( D = \{ w \in W : L(w) > 0 \} \). If the last claim were true we would have

\[
1 = E[1_D] \leq E[\Lambda(\psi)1_D] \leq 1,
\]

\(^5\)eventually taking twice convex combinations, the first combinations for assuring the weak convergence of \( (\mathcal{L} \Psi_n \circ T_n, n \geq 1) \) in \( L^1(\mu) \) and the second ones to assure its strong convergence in \( L^1(\mu) \) as well as the strong convergence of \( \nabla \Psi_n \circ T_n \) in \( L^2(\mu, H \otimes H) \).
which would imply $L_{1,D} = \Lambda(\psi)1_D$ $\mu$-a.s., as $\mu$ and $\nu$ are equivalent on $D$, the claim would have been proved. Let us now prove (7.29): From Lemma [12] the Fatou Lemma and the Fubini Theorem, we have

$$E_\mu[\Lambda(\psi)] \leq \liminf_n E_\mu[\Lambda(\pi_n^+ + \pi_n)] = \liminf_n \int \frac{\Lambda(\psi(w_n^+ + x))}{E_\mu[\pi_n^+ = w_n^+]} d\mu_n(x)d\mu_n(w_n^+) = 1$$

and the proof follows.

Recall that we denote the forward potential of the original problem by $\varphi$ which is a 1-convex function, and it is an element of the Sobolev space $D_{2,1}$. Its 1-convexity implies that the operator valued distribution $I_H + \nabla^2 \varphi$ is positive, hence it is a vector measure in the sense that, for any finite rank operator $\kappa$ on $H$, trace $[((I + \nabla^2 \varphi)\kappa]$ is a signed measure and if $\kappa$ is also positive, then trace $((I + \nabla^2 \varphi)\kappa)$ is a positive measure. Hence, due to the Lebesgue decomposition theorem, we can write it as the sum of an absolutely continuous measure and a singular measure w.r.t. the Wiener measure $\mu$. The absolutely continuous part is well-defined, denoted by $\nabla_a^2 \varphi$ and it satisfies

$$[\text{trace } ((I + \nabla^2 \varphi)\kappa)]_a = [\text{trace } ((I + \nabla_a^2 \varphi)\kappa)]$$

Similarly, we can look at $\pi_n\nabla^2 \varphi \pi_n$ as the second order derivative in the direction of the $n$-dimensional space $H_n$, hence, if we denote $w = \pi_n^+(w) + \pi_n(w) = w_n^+ + w_n$, then the partial map $w_n^+ \to \pi_n\nabla^2 \varphi \pi_n(w_n^+ + \cdot)$ can be interpreted as a measurable map with values in the space of measures on $H_n$, whose absolutely continuous part w.r.t. Lebesgue measure is denoted as $\pi_n\nabla^2 \varphi \pi_n(w_n^+ + \cdot)_a$.

With these notations we can announce

**Lemma 13.** The following relations hold true $\mu$-a.s.:

1. $$\lim_n (\pi_n \nabla^2 \varphi \pi_n)_a = \nabla_a^2 \varphi,$$
2. $$\lim_n ((\pi_n \nabla^2 \varphi \pi_n)(\pi_n^+(w) + \cdot))_a (\pi_n(w)) = \nabla_a^2 \varphi$$

**Proof:** Let us recall that the inequalities that we use below are to be understood in the sense of operators, i.e., we say that $A \geq B$, where $A$ and $B$ are two symmetric, positive operators on $H$ if $A - B$ is a positive operator. As $(\pi_n, n \geq 1)$ increases to the identity operator of $H$, the sequence $(\pi_n \nabla^2 \varphi \pi_n, n \geq 1)$ increases to $\nabla^2 \varphi$ as a sequence of lower bounded operator-valued measures, by the maximality of the absolutely continuous part of a measure in Lebesgue decomposition theorem, it follows that $\lim_n (\pi_n \nabla^2 \varphi \pi_n)_a \geq \nabla_a^2 \varphi$ almost surely. We have also naturally $\nabla^2 \varphi \geq \nabla_a^2 \varphi$, in particular, if $0 \leq g \in D$ and if $\gamma$ is a positive, symmetric, Hilbert-Schmidt or nuclear operator on $H$, then

$$E[ \text{trace } (\nabla^2 \varphi(g\gamma))] \geq E[ \text{trace } (\nabla_a^2 \varphi \gamma)g],$$

and $\pi_n \nabla^2 \varphi \pi_n \geq \pi_n \nabla_a^2 \varphi \pi_n$, by maximality we have also

$$(\pi_n \nabla^2 \varphi \pi_n)_a \geq \pi_n \nabla_a^2 \varphi \pi_n$$

---

Note that the lower index “$a$” depends also on the dimension $n$ and we omit this dependence for typographical simplicity.
almost surely. Taking the limit of both sides we get
\[ \lim_n (\pi_n \nabla^2 \varphi \pi_n)_a \geq \lim_n \pi_n \nabla^2 \varphi \pi_n = \nabla^2 \varphi , \]
which proves the first claim. To prove the second relation, let \( w_n = \pi_n w, w_n^+ = \pi_n^+ w \). We have, as explained above,
\[ (\pi_n \nabla^2 \varphi \pi_n)_a (w_n^+ + \cdot) \leq (\pi_n \nabla^2 \varphi \pi_n)(w_n^+ + \cdot)_a \]
\[ \leq (\pi_n \nabla^2 \varphi \pi_n)(w_n^+ + \cdot) \leq \nabla^2 \varphi (w_n^+ + \cdot) \]
\( \mu_n \)-a.s. for \( \mu_n^+ \)-almost all \( w_n^+ \). Writing \( w_n = \pi_n w, w_n^+ = \pi_n^+ w \), via Fubini, we get
\[ \limsup_n (\pi_n \nabla^2 \varphi \pi_n)(w_n^+ + \cdot)_a \leq \nabla^2 \varphi (w) \]
\( \mu \)-a.s., and from the maximality of the absolutely continuous part, we have also
\[ \limsup_n (\pi_n \nabla^2 \varphi \pi_n)(w_n^+ + \cdot)_a \leq \nabla^2 \varphi (w) , \]
almost surely. Writing all the inequalities together we get
\[ \nabla^2 \varphi \geq \limsup_n (\pi_n \nabla^2 \varphi \pi_n)(w_n^+ + \cdot)_a \]
\[ \geq \limsup_n (\pi_n \nabla^2 \varphi \pi_n)(w_n^+ + \cdot) \]
\[ = \limsup_n (\pi_n \nabla^2 \varphi \pi_n)(w) \]
\[ = \nabla^2 \varphi \]
\( \mu \)-a.s. \( \square \)

As a consequence of the above result we have

**Theorem 7.** The second Alexandroff derivative of the forward transport potential \( \varphi \), which is denoted as \( \nabla^2_a \varphi \), is a map with values in the space of the Hilbert-Schmidt operators on the Cameron-Martin space \( H \). It satisfies the following identity:
\[ (I_H + \nabla^2 \psi \circ T)(I_H + \nabla^2 a \varphi) = I_H \]
\( \mu \)-almost surely.

**Proof:** From the finite dimensional results and with the notations explained in Section 6 and in this section, cf., [15] [23], we have, for \( \mu_n^+ \)-a.a. \( w_n^+ \) the relation
\[ (I_{\mathbb{R}^n} + D^2_n \psi(w_n^+ + \pi_n T(w_n^+ + x)) \cdot (I_{\mathbb{R}^n} + D^2_n \varphi(w_n^+ + \cdot)_a(x)) = I_{\mathbb{R}^n} , \]
where, as we have mentioned before, \( D_n \) is the derivation operator on \( n \)-dimensional Euclidean space and the notation \( (D^2_n \varphi(w_n^+ + \cdot)_a(x) \) means the second Alexandroff derivative of the partial map \( x \to \varphi(w_n^+ + x) \) while \( w_n^+ \) is kept fixed. From the equation (7.31), we get
\[ D^2_n \psi(w_n^+ + \pi_n T(w_n^+ + x))(I_{\mathbb{R}^n} + \pi_n \nabla^2 \varphi \pi_n(w_n^+ + \cdot)_a(x)) = -\pi_n \nabla^2 \varphi \pi_n(w_n^+ + \cdot)_a . \]
As
\[ \lim_n D^2_n \psi(\pi_n^+ (w) + \pi_n T(\pi_n^+ (w) + \pi_n (w))) = \nabla^2 \psi \circ T(w) \]
\( \mu \)-a.s., and
\[
\lim_n (D_n^2 \varphi(\pi_n^+ + \cdot))_\omega(\pi_n(w)) = \nabla^2 \varphi(w)
\]
\( \mu \)-a.s., as proven in Lemma \( \text{I} \), the proof follows once we recall that \( D_n \) is the restriction of \( \nabla \) to the \( n \)-dimensional subspaces of \( H \).

**Remark 2.** It follows from the last lines of the above proof, namely from the relation \((7.33)\), that
\[
\lim_n \pi_n \nabla^2 \varphi(\pi_n^+ (w + \cdot)_n)(\pi_n(w)) = \nabla^2 \varphi(w)
\]
\( \mu \)-a.s. in the strong topology of Hilbert-Schmidt operators.

Let us define, with \( \mu \), the functional
\[
\delta_n(\pi_n \nabla \varphi)(w) = \sum_i (\nabla \varphi, e_i)_H \delta e_i - \text{trace } D_n^2(\pi_n^+ + \cdot)_n(w_n)
\]
where the lower index represents the absolutely continuous component of the measure on the Euclidean space \( H_n \) spanned by the orthonormal set \( \{e_1, \ldots, e_n\} \). It follows then from Lemmas \( \text{II} \) and \( \text{III} \) the following relation:

\[
\begin{align*}
(7.33) \\
&= -\delta_n(\pi_n \nabla \varphi)(w_n^1, w_n) + \log \det_2(J_{\mathbb{R}^n} + \pi_n \nabla^2 \varphi(\pi_n^+ (w_n^1 + \cdot)_n)) - \frac{1}{2} |\pi_n \nabla \varphi|^2 \\
&= -\log l_n(w_n^1 + \cdot) \circ (w_n + \pi_n \nabla \varphi(w_n^1 + w_n))
\end{align*}
\]

where
\[
l_n(w_n^1 + w_n) = \frac{L(w_n^1 + w_n)}{E[L|\pi_n^+ = w_n^1]}
\]
The following result gives the key information for the sequel:

**Theorem 8.** Let \( T_n(w) = w + \pi_n \nabla \varphi(w) \), then \( T_n \mu \) is absolutely continuous w.r.t. \( \mu \) with the corresponding Radon-Nikodym derivative
\[
M_n = \frac{dT_n \mu}{d\mu} = \frac{L}{E[L|\pi_n^+ = w_n^1]}
\]
\( \mu \)-almost surely. Moreover \( (M_n, n \geq 1) \) is uniformly integrable (w.r.t. \( \mu \)).

**Proof:** Let \( g \in C_b(W) \), it follows from Lemma \( \text{IIIA} \) and Lemma \( \text{I} \) that
\[
E[g \circ T_n] = E[E[g \circ T_n|\pi_n^+]]
\]
\[
= \int E[g \circ T_n|\pi_n^+] \mu_n^+(dw_n^1)
\]
\[
= \int E \left[ \frac{g L}{E[L|\pi_n^+]} \right] \mu_n^+(dw_n^1)
\]
\[
= \int E[g L|\pi_n^+] \frac{1}{E[L|\pi_n^+]} \mu_n^+(dw_n^1)
\]
\[
= E \left[ \frac{g L}{E[L|\pi_n^+]} \right].
\]
Let us remark that, from the elementary properties of the conditional expectation, \( E[L|\pi_n^+] \neq 0 \) on the set \( \{L \neq 0\} \) \( \mu \)-almost surely. Finally, it follows from the martingale convergence theorem.
\[(M_n = \frac{L_n}{E[L_n]}, n \geq 1)\] converges to \(L\) almost surely and \(E[M_n] = 1 = E[L]\) for any \(n \geq 1\), hence the uniform integrability follows.

The following theorem is the Monge-Ampère equation satisfied by \((\varphi, \psi)\):

**Theorem 9.** The sequence of Wiener functionals \((\delta_n \pi_n \nabla \varphi, n \geq 1)\) converges \(\mu\)-almost surely to a Wiener functional \(\mathcal{L}^n \varphi\). Moreover the Wiener Jacobian defined by

\[
\Lambda(\varphi) = \det_2 (I_H + \nabla^2 \varphi) \exp \left[ -\mathcal{L}^n \varphi - \frac{1}{2} |\nabla \varphi|^2_H \right],
\]

satisfies the Monge-Ampère equation:

\[
(7.34) \quad L \circ T \Lambda(\varphi) = 1
\]

\(\mu\)-a.s., where \(T = I_W + \nabla \varphi\). In particular, if \(\mu\{L > 0\} = 1\), then, for any \(g \in C_b(W)\), we have the Jacobi-Girsanov relation:

\[
\int_W g \circ T \Lambda(\varphi) d\mu = \int_W g d\mu.
\]

**Proof:** We shall use the notations of this section without further recall. First note that, due to the uniform integrability of \((M_n, n \geq 1)\) proven in Theorem 8, an application of Lusin theorem implies that

\[
\lim_{n \to \infty} L \circ T_n = L \circ T
\]

\(\mu\)-a.s. It follows from Remark 2 that

\[
\lim_{n \to \infty} \det_2 (I_{H^n} + \pi_n \nabla^2 \varphi) \pi_n(\pi_n^\perp(w) + \cdot)_{\pi_n}(\pi_n(w)) = \det_2 (I + \nabla^2 \varphi(w))
\]

\(\mu\)-a.s. As \(\lim_n |\pi_n \nabla \varphi|^2_H = |\nabla \varphi|^2_H\) \(\mu\)-a.s., also, it follows from the equality \((7.33)\), the sequence \((\delta_n \pi_n \nabla \varphi, n \geq 1)\) converges \(\mu\)-almost surely to a Wiener functional that we denote as \(\mathcal{L}^n \varphi\) and the relation \((7.34)\) follows. To show the Jacobi-Girsanov relation, for any \(g \in C_b(W)\), we have, from \((7.34)\),

\[
\int g \circ T \Lambda(\varphi) d\mu = \int g \circ T \frac{1}{L \circ T} d\mu = \lim_{\varepsilon \to 0} \int g \circ T \frac{1}{L \circ T + \varepsilon} d\mu = \lim_{\varepsilon \to 0} \int g \frac{1}{L + \varepsilon} d\mu = \int_{\{L \neq 0\}} g d\mu = \int g d\mu,
\]

since \(\mu(\{L \neq 0\}) = 1\).
8. Regularity of the forward potential $\varphi$

Let us show now the regularity of the forward Monge potential $\varphi$: assume first that, we have reduced the problem to the case where everything is smooth using the approximation results that we have proven before. Let $\nu$ be the measure defined by $d\nu = e^{-f}d\mu$. The following relation holds then true:

$$-\log\nu(e^f) = \inf_{\alpha} \left( \int -f d\alpha + H(\alpha|\nu) \right)$$

$$= \inf_U \left( \int -f \circ U d\nu + H(U\nu|\nu) \right).$$

(8.35)  

(8.36)

It is important to remark that in the equation 8.35, the infimum is taken over the set of probability measures and in the equation 8.36, the infimum is taken over the perturbations of identity of the form $U = I_W + u$ when $u$ runs in the set of the gradients of 1-convex functions, cf. [9]. Moreover, denoting $\frac{dU\mu}{d\nu}$ by $l_U$, we have

$$(l_U e^{-f}) \circ U \Lambda_u = e^{-f},$$

where $\Lambda_u$ is the Gaussian Jacobian associated to $U = I_W + u$. Therefore $\log l_U \circ U = f \circ U - \log \Lambda_u - f$ and we get

$$H(U\nu|\nu) = \int (f \circ U - \log \Lambda_u - f)d\nu$$

$$= \int (f \circ U - \log \Lambda_u - f)e^{-f}d\mu.$$ 

Consequently

$$-\log\nu(e^f) = \inf_U \left( \int -f \circ U d\nu + \int f \circ U e^{-f}d\mu - \int (f + \log \Lambda_u)e^{-f}d\mu \right)$$

$$= \inf_U \left( \int f e^{-f}d\mu - \int \log \Lambda_u e^{-f}d\mu \right)$$

$$= \inf_U J_b(U).$$

We know that the above infimum is attained at $S = T^{-1} = I_W + \nabla \psi$, hence we should have

$$J_b'(S) : \xi = \frac{d}{d\lambda} J_b(S + \lambda \xi)|_{\lambda=0} = 0$$

for any smooth $\xi : W \rightarrow H$ such that $\|\nabla \xi\|_2 \in L^\infty(\mu)$. A similar calculation as performed before implies that

$$\frac{d}{d\lambda} J_b(S + \lambda \xi)|_{\lambda=0} = \frac{d}{d\lambda} \left( \int -\log \Lambda_{S+\lambda \xi} d\nu \right)|_{\lambda=0}$$

$$= \int \left[ -\text{trace} \left( ((I + \nabla^2 \psi)^{-1} - I) \cdot \nabla \xi \right) + \delta \xi + (\nabla \psi, \xi) \right] d\nu$$

$$= 0$$

for any $\xi$ as above. Consequently we have

Theorem 10. The backward Monge potential satisfies the relation

$$\delta_{\nu}(I + \nabla^2 \psi)^{-1} = \nabla \psi + \nabla f,$$

where $\delta_{\nu}$ denotes the adjoint of $\nabla$ w.r.t. the measure $\nu$.

We need a couple of techical results:
Lemma 14. Let $\xi : W \to H$ be a smooth vector field, then the following results hold true:

1. $\delta_v \xi = \delta \xi + (\nabla f, \xi)_H$.
2. For any $h \in H$,
   \[ E_\nu[(\delta_v h)^2] = E_\nu[[h]^2 + (\nabla^2 f, h \otimes h)]. \]
3. For any $h \in H$ and smooth $\alpha : W \to \mathbb{R}$,
   \[ E_\nu[(\alpha(\delta_v h))^2] = E_\nu[(\alpha I_H + \nabla^2 \alpha + \alpha \nabla^2 f, h \otimes h)]_{H^\otimes 2}. \]

Lemma 15. For any smooth $\xi : W \to H$, we have

\[ E_\nu[(\delta_v \xi)^2] = E_\nu[(I_H + \nabla^2 f, \xi \otimes \xi)]_{H^\otimes 2} + \text{trace } (\nabla \xi \cdot \nabla \xi). \]

Proof: By the definition of $\delta_v$, we have

\[ E_\nu[\alpha(\delta_v h)^2] = E_\nu[[|\alpha|^2 + (\xi, \delta \otimes \nabla \xi) + (\xi, \nabla \nabla f, \xi)]]. \]

Besides $(\xi, \delta \otimes \nabla \xi) = \delta \nabla \xi + \text{trace } (\nabla \xi \cdot \nabla \xi)$ (cf. [22]). Hence

\[ E_\nu[\alpha(\delta_v h)^2] = E_\nu[|\xi|^2 + \delta \nabla \xi + \text{trace } (\nabla \xi \cdot \nabla \xi) + \delta \nabla \xi (\nabla \xi, \xi)]. \]

We also have

\[ \delta \nabla \xi = \delta \nabla \xi - (\nabla f, \nabla \xi, \xi). \]

Substituting this expression in the above calculation gives

\[
E_\nu[(\delta_v \xi)^2] = E_\nu[|\xi|^2_H + \delta_v u(\nabla \xi) - (\nabla f, \nabla \xi)_H + \text{trace } (\nabla \xi \cdot \nabla \xi) + \delta_v (\nabla f, \xi)_H] \\
= E_\nu[|\xi|^2_H - (\nabla f, \nabla \xi)_H + \text{trace } (\nabla \xi \cdot \nabla \xi) + (\xi, \nabla (\nabla f, \xi)_H)_H] \\
= E_\nu[|\xi|^2_H + \text{trace } (\nabla \xi \cdot \nabla \xi) - (\nabla f, \nabla \xi)_H + (\nabla \xi, \nabla \xi)_H + (\nabla f, \nabla \xi)_H] \\
= E_\nu[|\xi|^2_H + \text{trace } (\nabla \xi \cdot \nabla \xi) + (\nabla \xi, \nabla \xi)_H] \\
= E_\nu[|\xi|^2_H + \text{trace } (\nabla \xi \cdot \nabla \xi) + (\nabla^2 f, \xi \otimes \xi)_{H^\otimes 2}] \\
= E_\nu[|\xi|^2_H + \text{trace } (\nabla \xi \cdot \nabla \xi) + (\nabla^2 f, \xi \otimes \xi)_{H^\otimes 2}].
\]

The following theorem extends a result of Caffarelli [4], from log-concave densities to $1 - \varepsilon$-log-concave densities with a different proof:

Theorem 11. Assume that $f \in L^p(\mu)$ for some $p > 1$, satisfying $E[|\nabla f|_H^2 e^{-f}] < \infty$. Assume moreover that it is $(1 - \varepsilon)$-convex for some $\varepsilon > 0$, in the sense that the mapping

\[ h \to \frac{1 - \varepsilon}{2} |h|^2_{H^\otimes 2} + f(w + h) \]

is a convex map from the Cameron-Martin space $H$ to $L^0(\mu)$ (i.e., the equivalence class of real-valued Wiener functionals under the topology of convergence in probability). Then the forward Monge potential $\varphi$ belongs to the Gaussian Sobolev space $\mathbb{D}_{2,2}$. 

32
Consequently we get the corresponding forward potential \( \varphi \). Let \( d\nu_n = e^{-f_n}d\mu \), then we have, from Theorem 10

\[
\delta_{\nu_n}((I_H + \nabla^2\psi_n)^{-1} - I_H) = \nabla\psi_n + \nabla f_n.
\]

From Lemma 15 and denoting \( (I_H + \nabla^2\psi_n)^{-1} \) by \( M_n \), we get

\[
E_{\nu_n} \left[ \|\delta_{\nu_n}((I_H + \nabla^2\psi_n)^{-1} - I_H)\|_H^2 \right] = \sum_{k=1}^{\infty} E_{\nu_n} \left[ (\delta_{\nu_n}(M_n - I)(e_k))^2 \right]
\]

\[
= \sum_{k=1}^{\infty} E_{\nu_n} \left[ (I_H + \nabla^2 f_n, (M_n - I_H)e_k \otimes (M_n - I_H)e_k) \right] + \sum_{k=1}^{\infty} E_{\nu_n} \left[ \text{trace} \left( \nabla(M_n e_k) \cdot (\nabla M_n e_k) \right) \right].
\]

Since, due to the \((1 - \varepsilon)\)-convexity of \( f \) and as the second terms at the right of the second line is positive, we obtain

\[
E_{\nu_n} \|\delta_{\nu_n}(M_n - I_H)\|_H^2 \geq \varepsilon \sum_{k=1}^{\infty} E_{\nu_n} \| (M_n - I_H) e_k \|_H^2.
\]

Hence

\[
\varepsilon E_{\nu_n} \left[ \| (I_H + \nabla^2\psi_n)^{-1} - I_H \|_H^2 \right] \leq 2 E_{\nu_n} \left[ \|\nabla\psi_n\|_H^2 + \|\nabla f_n\|_H^2 \right],
\]

but \( E_{\nu_n} \|\nabla\psi_n\|_H^2 = E[\|\nabla\varphi_n\|_H^2] \) and

\[
E_{\nu_n} \left[ \| (I_H + \nabla^2\psi_n)^{-1} - I_H \|_H^2 \right] = E[\| (I_H + \nabla^2\psi_n)^{-1} \circ T_n - I_H \|_H^2] = E[\|\nabla^2\varphi_n\|_H^2].
\]

We also have

\[
E_{\nu_n} \left[ \|\nabla f_n\|_H^2 \right] = 4 E[\|e^{-f_n/2}\|^2]
\]

\[
= 4 E \left[ \frac{1}{e^{-f_n}} \|\nabla f_n e^{-f_n}\|^2 \right]
\]

\[
= 4 E \left[ \frac{1}{e^{-f_n}} \|\nabla e^{-f_n}\|^2 \right]
\]

\[
= 4 E \left[ \frac{1}{e^{-f_n}} \|\nabla P_{1/n} E[e^{-f}\|V_n\|]\|^2 \right]
\]

\[
\leq 4 e^{-1/n} E \left[ \frac{1}{e^{-f_n}} \|P_{1/n} E[\nabla e^{-f}\|V_n\|]\|^2 \right]
\]

\[
\leq 4 e^{-1/n} E \left[ \frac{1}{e^{-f_n}} P_{1/n} (E[\|\nabla f\|^2 e^{-f}\|V_n\|]) P_{1/n} E[e^{-f}\|V_n\|] \right]
\]

\[
= 4 e^{-1/n} E[P_{1/n} E[\|\nabla f\|^2 e^{-f}\|V_n\|]]
\]

\[
= 4 e^{-1/n} E[\|\nabla f\|^2 e^{-f}] .
\]

Consequently we get

\[
\varepsilon E[\|\nabla^2\varphi_n\|_H^2] \leq 2 E[\|\nabla\varphi_n\|_H^2] + 8 E[\|\nabla f\|^2 e^{-f}]
\]
Monge-Ampere Equation

Variational Solutions

and the claim follows by taking the limit at the r.h.s. and he limit inferior at the l.h.s. even with an explicit bound:

$$
\varepsilon E[\|\nabla^2 \varphi\|_2^2] \leq 2E[\|\nabla \varphi\|_H^2] + 8E[|\nabla f|^2 e^{-f}].
$$

The next corollary which follows from Theorem 4 and from Lemma 15 is about the regularity of the dual potential $\psi$:

**Corollary 4.** Assume that $E[\|\nabla^2 f\|_H^2 e^{-f}] < \infty$, where $\| \cdot \|_\infty$ denotes the operator norm on $H$. Then $(\delta_\nu \circ \nabla)\psi = L_\nu \psi$ belongs to $L^2(\nu)$.

**References**

[1] P.J. Bickel and D.A. Freedman: “Some asymptotic theory for the bootstrap”. Ann. Statis. 9(6), 1196–1217 (1981)
[2] Y. Brenier: “Polar factorization and monotone rearrangement of vector valued functions”. Comm. Pure Appl. Math. 44, 375-417, 1991.
[3] L. A. Caffarelli: “The regularity of mappings with a convex potential”. J. Am. Math. Soc. 5, 99–104 (1992)
[4] L. A. Caffarelli: “Monotonicity properties of optimal transportation and the FKG and related inequalities”. Commun. Math. Phys. 214, 547-563, 2000.
[5] V. Chernozhukov: “Uniform convergence of transport maps”. arXiv:1412.8434v1.
[6] S. Fang and V. Nolot: “Sobolev estimates for optimal transport maps on Gaussian spaces”. J. Funct. Anal. 266, no. 8, 5045–5084 (2014).
[7] D. Feyel and A.S. Üstünel. The notion of convexity and concavity on Wiener space. *Journal of Functional Analysis*, vol. 176, pp. 400-428, 2000.
[8] D. Feyel and A. S. Üstünel: “Transport of measures on Wiener space and the Girsanov theorem”. Comptes Rendus Mathématiques, Vol. 334, Issue 1, 1025-1028, 2002.
[9] D. Feyel, A.S. Üstünel: Monge-Kantorovich measure transportation and Monge-Ampère equation on Wiener space. *Probab. Theor. Relat. Fields*, 128, no. 3, pp. 347–385, 2004.
[10] D. Feyel and A. S. Üstünel: “Monge-Kantorovich measure transportation, Monge-Ampère equation and the Itô calculus”. Advanced Studies in Pure Mathematics, Math. Soc. of Japan, Vol. 41, p. 32-49, 2004. Mathematical Society of Japan.
[11] D. Feyel and A. S. Üstünel: “Solution of the Monge-Ampère equation on Wiener space for general log-concave measures”. J. Funct. Anal. 232 (2006), no. 1, 29–55.
[12] D. Feyel, A.S. Üstünel and M. Zakai: “Realization of Positive Random Variables via Absolutely Continuous Transformations of Measure on Wiener Space”. Probability Surveys,Vol. 3, (electronic) p.170-205, 2006.
[13] L. Gross: “Logarithmic Sobolev inequalities”. Amer. Jour. Math.97, p.1061-1083 (1975).
[14] A. V. Kolesnikov: “Convexity inequalities and optimal transport of infinite-dimensional measures”. J. Math. Pures Appl. (9), 83, n 11, 1373-1404, 2004.
[15] R.J McCann: “Existence and uniqueness of monotone measure-preserving maps”. Duke Math. J. 80, 309–323, 1995.
[16] S.T. Rachev and L. Rüschendorf: *Mass transportation problems. Vol. I, II. Theory. Probability and its Applications* (New York). Springer-Verlag, New York, 1998.
[17] H. Sugita: “Positive generalized Wiener functions and potential theory over abstract Wiener spaces”. Osaka J. Math. 25 (1988), no. 3, 665–696.
[18] A. S. Üstünel: *Introduction to Analysis on Wiener Space*. Lecture Notes in Math. Vol. 1610. Springer, 1995.
[19] A. S. Üstünel: *Analysis on Wiener Space and Applications*. Electronic text at the site arXiv.org. math. arXiv:1003.1649
[20] A. S. Üstünel: “Entropy, invertibility and variational calculus of adapted shifts on Wiener space”. J. Funct. Anal. 257 (2009), no. 11, 3655–3689.
[21] A. S. Üstünel: “Variational calculation of Laplace transforms via entropy on Wiener space and applications”. J. Funct. Anal. 267 (2014), no. 8, 3058–3083.
Monge-Ampere Equation

Variational Solutions

[22] A. S. Üstünel and M. Zakai: Transformation of Measure on Wiener Space. Springer Verlag, 1999.
[23] C. Villani: Topics in Optimal Transportation. Graduate Series in Math., 58. Amer. Math. Soc., 2003.

A. S. Üstünel, Bilkent University, Math. Dept., Ankara, Turkey
ustunel@fen.bilkent.edu.tr