

The strong solution of the Monge-Ampère equation on Wiener space for general log-concave measures

La solution forte de l'équation de Monge-Ampère sur l'espace de Wiener pour les mesures log-concaves : cas général

D. Feyel
S. Üstünel

A large gray square containing the text "2007C001".

2007C001

2007

Département Informatique et Réseaux
Groupe Mathématiques de l'Information
des Communications et du Calcul

Solution of the Monge-Ampère Equation on Wiener Space for general log-concave measures

D. Feyel and A. S. Üstünel

Abstract

In this work we prove that the unique 1-convex solution of the Monge-Kantorovitch measure transportation problem between the Wiener measure and a target measure which has an H -log-concave density, in the sense of [9], w.r.to the Wiener measure is also the strong solution of the Monge-Ampère equation in the frame of infinite dimensional Fréchet spaces. We further enhance the polar factorization results of the mappings which transform a spread measure to another one in terms of the measure transportation of Monge-Kantorovitch and clarify the relation between this concept and the Itô-solutions of the Monge-Ampère equation.

1 Introduction

In 1781, G. Monge has launched his famous problem [18], which can be expressed in terms of the modern mathematics as follows: given two probability measures ρ and ν on \mathbb{R}^n , find the map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $T\rho = \nu$ ¹ and T is also the solution of the minimization problem

$$\inf_U \left\{ \int_{\mathbb{R}^n} c(x, U(x)) \rho(dx) \right\}, \quad (1.1)$$

where the infimum is taken between all the maps $U : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $U\rho = \nu$ and where $c : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_+$ is a positive, measurable function, called usually the cost function. In the original problem of Monge, the cost function $c(x, y)$ was $|x - y|$ and the dimension n was three. Later other costs have been considered, between them, the most popular one which is also abundantly studied, is the case where $c(x, y) = |x - y|^2$. After several tentatives (cf., [1, 2]), in the 1940's this highly nonlinear problem of Monge has been reduced to a linear problem by Kantorovitch, cf.[15], in the following way: let $\Sigma(\rho, \nu)$ be the set of probability measures on $\mathbb{R}^n \times \mathbb{R}^n$, whose first marginals are ρ and the second marginals are

¹ $T\rho$ means the image of the measure ρ under the map T

ν . Find the element(s) of $\Sigma(\rho, \nu)$ which are the solutions of the minimization problem:

$$\inf_{\beta \in \Sigma(\rho, \nu)} \left\{ \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) d\beta(x, y) \right\}. \quad (1.2)$$

It is obvious that $\Sigma(\rho, \nu)$ is a convex, compact set under the weak*-topology of measures, hence, in case, the cost function c has some regularity properties, like being lower semi-continuous, this problem would have solutions. If any one of them is supported by the graph of a map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, then obviously, T will also be a solution of the original problem of Monge 1.1. Since that time, the problem (1.2) is called the Monge-Kantorovitch problem (MKP). The program of Kantorovitch has been followed by several people and a major contribution has been done by Sudakov [24]. In the early 90's there has been another impetus to this problem, cf., [4, 19], where it has been discovered the important role played by the convex functions in the construction of the solutions of the MKP and of the problem of Monge (cf., [16, 17, 19]). We refer the reader to [8], [20] and to [30] for recent surveys.

In [11], we have solved the MKP and the problem of Monge in the infinite dimensional case, where the measures are concentrated in a Fréchet space W into which a Hilbert space H is injected densely and continuously. We call H the Cameron-Martin space in reference to the Gaussian case. The cost function is defined on $W \times W$ as

$$\begin{aligned} c(x, y) &= |x - y|_H^2 \text{ if } x - y \in H \\ &= \infty \text{ if } x - y \notin H, \end{aligned}$$

where $|\cdot|_H$ denotes the Euclidean norm of H . Because of this choice, in comparison to the finite dimensional space, the situation becomes quite singular, since, in general, the Cameron-Martin space H is a negligible set (i.e., of null measure) with respect to almost all reasonable measures for which one may wish to search the solutions of Monge's problem and of MKP. On the other hand, due to the potential applications to several problems of analysis and physics, this cost function is particularly important. For example, it is particularly well-adapted to the study of the absolute continuity of the image of the Wiener measure under the perturbations of identity, which is a subject under investigation since the early works of N. Wiener, R.H. Cameron and W.T. Martin and of several other mathematicians and engineers who have made worthy contributions (cf. the list of references of [28]).

This paper is devoted to the further developments of the subject. At first we give a generalization of the polar factorization of vector fields which map a probability measure on W to another one such that one of them is spread (cf. the preliminaries) and the two measures are at finite Wasserstein distance from each other. This is done without any absolute continuity hypothesis. As an example

we treat in detail the case of the infinite dimensional Gaussian measures: we construct two Gaussian measures which are at finite Wasserstein distance from each other, hence one can be transported on to the other one with an explicitly constructed transport map although they are mutually singular. Afterwards we prove that the forward transport map satisfies the functional analytic (or strong) Monge-Ampère equation, when the target measure has an H -log-concave density. We remark that this kind of measures are widely used in Physics. In [11], we have studied the Monge-Ampère equation for the upper and lower bounded densities with respect to the (infinite dimensional) Wiener measure. The main difficulty in this infinite dimensional case stems from the lack of regularity of the transport potentials, in fact we only know that these functions are in the Gaussian Sobolev space $ID_{2,1}$, i.e., they have only first order Sobolev derivatives. However, to write the Gaussian Jacobian, we need them to have second order Sobolev derivatives taking values in the space of Hilbert-Schmidt operators on the Cameron-Martin space H . This difficulty is worse than those we encounter in the finite dimensional case, since in the latter the Hilbert-Schmidt property holds automatically. Moreover, in the finite dimensional situation the lack of second order derivatives is solved with the help of the Alexandroff derivatives of the convex functions. In the infinite dimensional case the situation is more complicated: the transport potentials are not in general convex, nor H -convex (which is a more reasonable requirement than being convex, cf. [9]), but only 1-convex² in the Cameron-Martin space direction. Hence their second order derivatives in the sense of distributions are not in general Hilbert-Schmidt operators'-valued vector measures; even if the contrary occurs in some exceptional situations, their absolutely continuous parts do not take values necessarily in the space of Hilbert-Schmidt operators; a condition which is indispensable to write down the Jacobian of the transport map. Hence it is impossible in general to construct the strong solutions of the Monge-Ampère equation. In [13], this equation has been solved in a particular case where the target measure has a smooth density in the sense of the Malliavin calculus.

In Section 6, combining the finite dimensional results of Caffarelli [7] with Wiener space analysis, we solve completely this problem when the target measure is H -log-concave. More precisely, we show that the transport potential has a second order derivative as an Hilbert-Schmidt operator valued map, hence we can write the corresponding Jacobian which includes the modified Carleman-Fredholm determinant, cf. [28] and finally we prove that the transport potential is the unique 1-convex strong solution of the Monge-Ampère equation. In Section 7 we show that all these difficulties disappear if we use the natural Ito Calculus and we can calculate the Itô Jacobian (cf. Theorem 7.1) using the natural Brownian motion which is associated to the solution of the Monge problem. In fact, with Itô parametrization, the complications are absorbed by the filtrations of

²In finite dimensions this corresponds to the notion of semi-convexity.

forward and backward transport processes (i.e., maps). We give also the delicate relations between the polar factorization of the absolutely continuous transformations of the Wiener measure and the Brownian motions which appear in the semimartingale decomposition of the transport process with respect to its natural filtration.

2 Preliminaries and notations

Let W be a separable Fréchet space equipped with a Gaussian measure μ of zero mean whose support is the whole space³. 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, μ, 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 V_n be the σ -algebra generated by π_{F_n} , then for any $f \in L^p(\mu)$, the martingale sequence $(E[f|V_n], n \geq 1)$ converges to f (strongly if $p < \infty$) in $L^p(\mu)$. Observe that the function $f_n = E[f|V_n]$ can be identified with a function on the finite dimensional abstract Wiener space (F_n, μ_n, F_n) , where $\mu_n = \pi_n \mu$.

Since the translations of μ with the elements of H induce measures equivalent to μ , 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 ∇ cf., for example [27]. 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 Φ , 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 δ . In the case of classical Wiener space, i.e., when $W = C(\mathbb{R}_+, \mathbb{R}^d)$, then δ coincides with the Itô integral of the Lebesgue density of the adapted elements of $\mathbb{D}_{p,k}(H)$ (cf.[27]).

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

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

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.[27]). Its infinitesimal

³The reader may assume that $W = C(\mathbb{R}_+, \mathbb{R}^d)$, $d \geq 1$ or $W = \mathbb{R}^{\mathbb{N}}$.

generator is denoted by $-\mathcal{L}$ and we call \mathcal{L} the Ornstein-Uhlenbeck operator (sometimes called the number operator by the physicists). Due to the Meyer inequalities (cf., for instance [27]), the norms defined by

$$\|\varphi\|_{p,k} = \|(I + \mathcal{L})^{k/2}\varphi\|_{L^p(\mu)} \quad (2.3)$$

are equivalent to the norms defined by the iterates of the Sobolev derivative ∇ . This observation permits us to identify the duals of the space $\mathbb{D}_{p,k}(\Phi)$; $p > 1$, $k \in \mathbb{N}$ by $\mathbb{D}_{q,-k}(\Phi')$, with $q^{-1} = 1 - p^{-1}$, where the latter space is defined by replacing k in (2.3) 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 $\delta : \mathbb{D}_{q,k}(H \otimes \Phi) \rightarrow \mathbb{D}_{q,k-1}(\Phi)$ and $\nabla : \mathbb{D}_{q,k}(\Phi) \rightarrow \mathbb{D}_{q,k-1}(H \otimes \Phi)$ continuously, for any $q > 1$ and $k \in \mathbb{R}$, where $H \otimes \Phi$ denotes the completed Hilbert-Schmidt tensor product (cf., for instance [27]). The following assertion is useful: assume that $(Z_n, n \geq 1) \subset \mathbb{D}'$ converges to Z in \mathbb{D}' , assume further that 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 , c.f. [27].

A measurable function $f : W \rightarrow \mathbb{R} \cup \{\infty\}$ is called H -convex (cf.[9]) if

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

is convex μ -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

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

is convex on the Cameron-Martin space H with values in $L^0(\mu)$. Note that all these notions are compatible with the μ -equivalence classes of random variables thanks to the Cameron-Martin theorem. It is proven in [9] 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 π_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

$$x_n \rightarrow \frac{1}{2}|x_n|_H^2 + f(x_n + x_n^\perp)$$

is $\pi_n^\perp \mu$ -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 [9].

3 Monge-Kantorovitch problem

Let us recall the definition of the Monge-Kantorovitch problem in our case:

Definition 3.1 *Let ρ and ν be two probability measures on W , let also $\Sigma(\rho, \nu)$ be the convex subset of the probability measures on the product space $W \times W$ whose first marginal is ρ and the second one is ν . The Monge-Kantorovitch problem for the couple (ρ, ν) consists of finding a measure $\gamma \in \Sigma(\rho, \nu)$ which realizes the following infimum:*

$$d_H^2(\rho, \nu) = \inf_{\beta \in \Sigma(\rho, \nu)} \int_{W \times W} |x - y|_H^2 d\beta(x, y).$$

The function $c(x, y) = |x - y|_H^2$ is called the cost function.

Remark 3.1 *Note that the cost function is only lower semi-continuous with respect to the product topology of $W \times W$ due to the dense and continuous injection of H into W . Hence it takes the value ∞ very often for the most notable measures, e.g., when ρ and ν are absolutely continuous with respect to the Wiener measure μ .*

The proof of the next theorem, for which we refer the reader to [11], can be done by choosing a proper disintegration of any optimal measure in such a way that the elements of this disintegration are the solutions of finite dimensional Monge-Kantorovitch problems. The latter is proven with the help of the measurable section-selection theorem.

Theorem 3.1 (General case) *Suppose that ρ and ν are two probability measures on W such that*

$$d_H(\rho, \nu) < \infty.$$

Let $(\pi_n, n \geq 1)$ be a total increasing sequence of regular projections (of H , converging to the identity map of H). Suppose that, for any $n \geq 1$, the regular conditional probabilities $\rho(\cdot | \pi_n^\perp = x^\perp)$ vanish $\pi_n^\perp \rho$ -almost surely on the subsets of $(\pi_n^\perp)^{-1}(W)$ with Hausdorff dimension $n-1$. Then there exists a unique solution of the Monge-Kantorovitch problem, denoted by $\gamma \in \Sigma(\rho, \nu)$ and γ is supported by the graph of a Borel map T which is the solution of the Monge problem. $T : W \rightarrow W$

is of the form $T = I_W + \xi$, where $\xi \in L^2(\rho, H)$ (it is an H -valued map with a ρ -square integrable H -norm). Besides we have

$$\begin{aligned} d_H^2(\rho, \nu) &= \int_{W \times W} |T(x) - x|_H^2 d\gamma(x, y) \\ &= \int_W |T(x) - x|_H^2 d\rho(x) = \int_W |\xi|_H^2 d\rho, \end{aligned}$$

and for $\pi_n^\perp \rho$ -almost almost all x_n^\perp , the map $u \rightarrow u + \xi(u + x_n^\perp)$ is cyclically monotone on $(\pi_n^\perp)^{-1}\{x_n^\perp\}$, in the sense that

$$\sum_{i=1}^N (u_i + \xi(x_n^\perp + u_i), u_{i+1} - u_i)_H \leq 0$$

$\pi_n^\perp \rho$ -almost surely, for any cyclic sequence $\{u_1, \dots, u_N, u_{N+1} = u_1\}$ from $\pi_n(W)$. Finally, if, for any $n \geq 1$, $\pi_n^\perp \nu$ -almost surely, $\nu(\cdot | \pi_n^\perp = y^\perp)$ also vanishes on the $n-1$ -Hausdorff dimensional subsets of $(\pi_n^\perp)^{-1}(W)$, then T is invertible, i.e., there exists $S : W \rightarrow W$ of the form $S = I_W + \eta$ such that $\eta \in L^2(\nu, H)$ satisfies a similar cyclic monotonicity property as ξ and that

$$\begin{aligned} 1 &= \gamma \{(x, y) \in W \times W : T \circ S(y) = y\} \\ &= \gamma \{(x, y) \in W \times W : S \circ T(x) = x\}. \end{aligned}$$

In particular we have

$$\begin{aligned} d_H^2(\rho, \nu) &= \int_{W \times W} |S(y) - y|_H^2 d\gamma(x, y) \\ &= \int_W |S(y) - y|_H^2 d\nu(y) = \int_W |\eta|_H^2 d\nu. \end{aligned}$$

Remark 3.2 In particular, for all the measures ρ which are absolutely continuous with respect to the Wiener measure μ , the second hypothesis is satisfied, i.e., the measure $\rho(\cdot | \pi_n^\perp = x_n^\perp)$ vanishes on the sets of Hausdorff dimension $n-1$.

Any probability measure satisfying the hypothesis of Theorem 3.1 is called a spread measure. Namely,

Definition 3.2 A probability measure m on $(W, \mathcal{B}(W))$ is called a spread measure if there exists a sequence of finite dimensional regular projections $(\pi_n, n \geq 1)$ converging to I_H such that the regular conditional probabilities $m(\cdot | \pi_n^\perp = x_n^\perp)$ which are concentrated in the n -dimensional spaces $\pi_n(W) + x_n^\perp$ vanish on the sets of Hausdorff dimension $n-1$ for $\pi_n^\perp(m)$ -almost all x_n^\perp and for any $n \geq 1$.

The case where one of the measures is the Wiener measure and the other is absolutely continuous with respect to μ is the most important one for the applications. Consequently we give the related results separately in the following theorem where the tools of the Malliavin calculus give more information about the maps ξ and η of Theorem 3.1:

Theorem 3.2 (Gaussian case) *Let ν be the measure $d\nu = Ld\mu$, where L is a positive random variable, with $E[L] = 1$. Assume that $d_H(\mu, \nu) < \infty$ (for instance $L \in \mathbb{L} \log \mathbb{L}$). Then there exists a 1-convex function $\varphi \in \mathbb{D}_{2,1}$ and a partially 1-convex function $\psi \in L^2(\nu)$, both are unique upto a constant and called, respectively, forward and backward Monge-Kantorovitch potentials, such that*

$$\varphi(x) + \psi(y) + \frac{1}{2}|x - y|_H^2 \geq 0$$

for all $(x, y) \in W \times W$ and that

$$\varphi(x) + \psi(y) + \frac{1}{2}|x - y|_H^2 = 0$$

γ -almost everywhere. The map $T = I_W + \nabla\varphi$ is the unique solution of the original problem of Monge. Moreover, its graph supports the unique solution of the Monge-Kantorovitch problem γ . Consequently

$$(I_W \times T)\mu = \gamma$$

In particular T maps μ to ν and T is almost surely invertible, i.e., there exists some $T^{-1} = I_W + \eta$ such that $T^{-1}\nu = \mu$, $\eta \in L^2(\nu)$ and that

$$\begin{aligned} 1 &= \mu \{x : T^{-1} \circ T(x) = x\} \\ &= \nu \{y \in W : T \circ T^{-1}(y) = y\} . \end{aligned}$$

Remark 3.3 Assume that the operator ∇ is closable with respect to ν , then we have $\eta = \nabla\psi$. In particular, if ν and μ are equivalent, then we have

$$T^{-1} = I_W + \nabla\psi ,$$

where ψ is a 1-convex function.

Remark 3.4 Let $(e_n, n \in \mathbb{N})$ be a complete, orthonormal in H , denote by V_n the sigma algebra generated by $\{\delta e_1, \dots, \delta e_n\}$ and let $L_n = E[L|V_n]$. If $\varphi_n \in \mathbb{D}_{2,1}$ is the function constructed in Theorem 3.2, corresponding to L_n , then, using the inequality (cf., [11])

$$d_H^2(\mu, \nu) \leq 2E[L \log L] ,$$

we can prove that the sequence $(\varphi_n, n \in \mathbb{N})$ converges to φ in $\mathbb{D}_{2,1}$.

Theorem 3.2 assumes the absolute continuity of ν with respect to μ . On the other hand Theorem 3.1 holds in the case where the Wasserstein distance between two measures is finite with any assumption about the absolute continuity. The following theorem fills this gap in the Gaussian case:

Theorem 3.3 (General Gaussian case) *Assume that the hypothesis of Theorem 3.1 are valid for the measures ν and μ , where as usual, μ denotes the Wiener measure. Then we can add to the conclusions of Theorem 3.1 the existence of a forward potential function $\varphi \in \mathbb{D}_{2,1}$ which is 1-convex and that $\xi = \nabla\varphi$ and $T = I_W + \nabla\varphi$.*

Proof: It is straightforward to see that $h \rightarrow h + P_t\xi(w + h)$ is cyclically monotone if $h \rightarrow h + \xi(w + h)$ is cyclically monotone, where P_t is the Ornstein-Uhlenbeck semigroup. Hence assume first that ξ is smooth. It follows from Poincaré-Helmholtz decomposition on the Wiener space (cf.[26], Theorem IV.1) that ξ has a unique decomposition

$$\xi = \nabla\varphi + \alpha,$$

where $\varphi \in \cap_k \mathbb{D}_{2,k} = \mathbb{D}_{2,\infty}$ and $\alpha \in \mathbb{D}_{2,\infty}(H) = \cap_k \mathbb{D}_{2,k}(H)$ such that $\delta\alpha = 0$. φ is defined by $\mathcal{L}\varphi = \delta\xi$, hence without loss of generality, we may assume that $E[\varphi] = 0$. By a theorem of Rockafellar [21]⁴, the map $h \rightarrow t_w(h) = h + \xi(w + h)$ is equal to the subdifferential of a convex function on H . Hence its derivative in H , which is μ -almost surely equal to $I_H + \nabla\xi(w + h)$ is a symmetric operator on H . By difference $\nabla\xi(w + h)$ is symmetric and by the quasi-invariance of μ , $\nabla\xi(w)$ is a symmetric operator on H μ -almost surely. Consequently $\nabla\alpha = \nabla\xi - \nabla^2\varphi$ is also symmetric. Since $\delta\alpha = 0$ and since we have

$$\begin{aligned} E[(\delta\alpha)^2] &= E[|\alpha|_H^2] + E[\text{trace}((\nabla\alpha) \cdot (\nabla\alpha))] \\ &= E[|\alpha|_H^2] + E[\text{trace}((\nabla\alpha)^* \cdot (\nabla\alpha))] \\ &= 0, \end{aligned}$$

where $*$ denotes the adjoint operator, we see that $\alpha = 0$ μ -almost surely, hence $\xi = \nabla\varphi$. Since $h \rightarrow h + \nabla\varphi(w + h)$ is cyclically monotone, again from [21], for μ -almost all $w \in W$, $h \rightarrow \frac{1}{2}|h|_H^2 + \varphi(w + h)$ is convex, i.e., φ is 1-convex and it has zero expectation. Let us denote this function with φ_t . For the general case, it suffices to take the limit when $t \rightarrow 0$, in this case we have, from the Poincaré inequality

$$E[|\varphi_t - \varphi_s|^2] \leq E[|P_t\xi - P_s\xi|^2] \rightarrow 0,$$

as $t - s \rightarrow 0$, hence $\lim_{t \rightarrow 0} \varphi_t$ exists in $\mathbb{D}_{2,1}$ and the limit is also 1-convex from [9]. \square

⁴[21] treats the finite dimensional case but the generalization to a Hilbert space is immediate.

4 Polar factorization of mappings between spread measures

In [11] we have proved the polar factorization of the mappings $U : W \rightarrow W$ such that the Wasserstein distance between $U\mu$ and the Wiener measure μ , denoted by $d_H(\mu, U\mu)$, is finite. We have also studied the particular case where U is a perturbations of identity, i.e., it is the form $I_W + u$, where u maps W to the Cameron-Martin space H . In this section we shall generalize this results in the frame of spread measures.

Theorem 4.1 *Assume that ρ and ν are spread measures with $d_H(\rho, \nu) < \infty$ and that $U\rho = \nu$, for some measurable map $U : W \rightarrow W$. Let T be the optimal transport map sending ρ to ν , whose existence and uniqueness is proven in Theorem 3.1. Then $R = T^{-1} \circ U$ is a ρ -rotation (i.e., $R\rho = \rho$) and $U = T \circ R$, moreover, if U is a perturbation of identity, then R is also a perturbation of identity. In both cases, R is the ρ -almost everywhere unique minimal ρ -rotation in the sense that*

$$\int_W |U(x) - R(x)|_H^2 d\rho(x) = \inf_{R' \in \mathcal{R}} \int_W |U(x) - R'(x)|_H^2 d\rho(x), \quad (4.4)$$

where \mathcal{R} denotes the set of ρ -rotations.

Proof: Let T be the optimal transportation of ρ to ν whose existence and uniqueness follows from Theorem 3.1. The unique solution γ of the Monge-Kantorovitch problem for $\Sigma(\rho, \nu)$ can be written as $\gamma = (I \times T)\rho$. Since ν is spread, T is invertible on the support of ν and we have also $\gamma = (T^{-1} \times I)\nu$. In particular $R\rho = T^{-1} \circ U\rho = T^{-1}\nu = \rho$, hence R is a rotation. Let R' be another rotation in \mathcal{R} , define $\gamma' = (R' \times U)\rho$, then $\gamma' \in \Sigma(\rho, \nu)$ and the optimality of γ implies that $J(\gamma) \leq J(\gamma')$, besides we have

$$\begin{aligned} \int_W |U(x) - R(x)|_H^2 d\rho(x) &= \int_W |U(x) - T^{-1} \circ U(x)|_H^2 d\rho(x) \\ &= \int_W |x - T^{-1}(x)|_H^2 d\nu(x) \\ &= \int_W |T(x) - x|_H^2 d\rho(x) \\ &= J(\gamma). \end{aligned}$$

On the other hand

$$J(\gamma') = \int_W |U(x) - R'(x)|_H^2 d\rho(x),$$

hence the relation (4.4) follows. Assume now that for the second rotation $R' \in \mathcal{R}$ we have the equality

$$\int_W |U(x) - R(x)|_H^2 d\rho(x) = \int_W |U(x) - R'(x)|_H^2 d\rho(x).$$

Then we have $J(\gamma) = J(\gamma')$, where γ' is defined above. By the uniqueness of the solution of Monge-Kantorovitch problem due to Theorem 3.1, we should have $\gamma = \gamma'$. Hence $(R \times U)\rho = (R' \times U)\rho = \gamma$, consequently, we have

$$\int_W f(R(x), U(x)) d\rho(x) = \int_W f(R'(x), U(x)) d\rho(x),$$

for any bounded, measurable map f on $W \times W$. This implies in particular

$$\int_W (a \circ T \circ R) (b \circ U) d\rho = \int_W (a \circ T \circ R') (b \circ U) d\rho$$

for any bounded measurable functions a and b . Let $U' = T \circ R'$, then the above expression reads as

$$\int_W a \circ U b \circ U d\rho = \int_W a \circ U' b \circ U d\rho.$$

Taking $a = b$, we obtain

$$\int_W (a \circ U) (a \circ U') d\rho = \|a \circ U\|_{L^2(\rho)} \|a \circ U'\|_{L^2(\rho)},$$

for any bounded, measurable a . This implies that $a \circ U = a \circ U'$ ρ -almost surely for any a , hence $U = U'$ i.e, $T \circ R = T \circ R'$ ρ -almost surely. Let us denote by S the left inverse of T whose existence follows from Theorem 3.1 and let $D = \{x \in W : S \circ T(x) = x\}$. Since $\rho(D) = 1$ and since R and R' are ρ -rotations, we have also

$$\rho(D \cap R^{-1}(D) \cap R'^{-1}(D)) = 1.$$

Let $x \in W$ be any element of $D \cap R^{-1}(D) \cap R'^{-1}(D)$, then

$$\begin{aligned} R(x) &= S \circ T \circ R(x) \\ &= S \circ T \circ R'(x) \\ &= R'(x), \end{aligned}$$

consequently $R = R'$ on a set of full ρ -measure.. □

5 Application to Gaussian measures

The next two paragraphs give two applications of the results illustrated in the preceding sections. The first one treats the transport of a Gaussian measure to a second Gaussian measure which is absolutely continuous with respect to the former one and in the second paragraph we suppress the hypothesis of absolute continuity.

5.1 Absolutely continuous case

Assume that $\rho = \mu$, i.e., the Wiener measure and let K be a Hilbert-Schmidt operator on H . Assume that the Carleman-Fredholm determinant $\det_2(I_H + K)$ is different than zero, hence the operator $I_H + K : H \rightarrow H$ is invertible. Moreover, it follows from the general theory that $I_H + K$ has a unique polar decomposition as $I_H + K = (I_H + \bar{K})(I_H + A)$, where $I_H + A$ is an isometry⁵ and $I_H + \bar{K}$ is a symmetric, positive operator. Note that \bar{K} is compulsorily Hilbert-Schmidt. Let us now define $U : W \rightarrow W$ as $U(x) = x + \delta K(x)$, where $\delta K(x)$ is the H -valued divergence, defined by $(\delta K(x), h)_H = \delta(K^*h)(x)$. Then it is known that the measure $U\mu$ is absolutely continuous with respect to μ , in fact $U\mu$ is even equivalent to μ since $|\Lambda_K| \neq 0$ μ -almost surely, where

$$\Lambda_K = \det_2(I_H + K) \exp \left\{ \delta^2(K) - \frac{1}{2} |\delta K|_H^2 \right\}.$$

Besides we have

$$L = \frac{dU\mu}{d\mu} = \frac{1}{|\Lambda_K| \circ V},$$

where V is the inverse of U , whose existence follows from the invertibility of $h \rightarrow h + \delta(K)(x) + Kh$ on H , cf. [28]. Consequently,

$$E[L \log L] = -E[\log |\Lambda_K|] < \infty,$$

hence $d_H(\mu, U\mu) < \infty$. We shall prove that the polar factorization of U is given by

$$U = (I_W + \delta \bar{K}) \circ (I_W + \delta A).$$

In fact, it follows from Theorem B.6.4 of [28], that

$$\begin{aligned} (I_W + \delta \bar{K}) \circ (I_W + \delta A) &= I_W + \delta \bar{K} + \delta A + \delta(\bar{K}A) \\ &= I_W + \delta(\bar{K} + A + \bar{K}A) \\ &= I_W + \delta K. \end{aligned}$$

Besides $\nabla^2 \delta^2 \bar{K} = 2\bar{K}$, and since $I_H + \bar{K}$ is a positive operator, the Wiener map $\frac{1}{2} \delta^2 \bar{K}$ is 1-convex, consequently, $T = I_W + \delta \bar{K}$ is the transport map and $I_W + \delta A$ is the unique rotation whose existence is proven in Theorem 4.1. The Kantorovitch potentials φ and ψ of Theorem 3.2 can be chosen as

$$\varphi(x) = \frac{1}{2} \delta^2 \bar{K}(x)$$

for T and

$$\psi(x) = -\frac{1}{2} \delta((I_H + \bar{K})^{-1} \bar{K})(x)$$

for $T^{-1} = I_W + \nabla \psi$.

⁵ A satisfies the relation $A + A^* + A^*A = 0$.

Remark 5.1 Let us denote by $P_{\ker \delta}$ the projection operator from $\mathbb{D}'(H)$ to the kernel of the divergence operator δ . Then, we have the following identity:

$$P_{\ker \delta} (\delta((I_H + \bar{K})A)) = \delta \hat{K} - \delta \bar{K} ,$$

where \hat{K} denotes the symmetrization of K . This shows that the polar decomposition and the Helmholtz decomposition are different in general.

5.2 The general case of two Gaussian measures at finite Wasserstein distance

We can also calculate the forward Monge-Kantorovitch potential function for the singular case as follows: assume that ν is a zero mean Gaussian measure on W such that $d_H(\mu, \nu) < \infty$. Then there exists a bilinear form q on W^* such that

$$\int_W e^{i\langle \alpha, x \rangle} d\nu(x) = \exp -\frac{1}{2}q(\alpha, \alpha) ,$$

for any $\alpha \in W^*$. On the other hand, from Theorem 3.3, there exists a $\varphi \in \mathbb{D}_{2,1}$, which is 1-convex, such that $T\mu = (I_W + \nabla\varphi)\mu = \nu$. Hence, rewriting the above relation with T , we obtain:

$$\int_W e^{i\langle t\alpha, T(x) \rangle} d\mu(x) = \exp -\frac{t^2}{2}q(\alpha, \alpha) , \quad (5.5)$$

for any $t \in \mathbb{R}$ and $\alpha \in W^*$. Taking the derivative of both sides twice at $t = 0$, we obtain

$$\begin{aligned} q(\alpha, \alpha) &= |\tilde{\alpha}|_H^2 + E [(\nabla\varphi, \tilde{\alpha})_H^2] + 2E [(\nabla\varphi, \tilde{\alpha})_H \delta\tilde{\alpha}] \\ &= |\tilde{\alpha}|_H^2 + E [(\nabla\varphi \otimes \nabla\varphi, \tilde{\alpha} \otimes \tilde{\alpha})_2] + 2E [(\nabla^2\varphi, \tilde{\alpha} \otimes \tilde{\alpha})_2] , \end{aligned}$$

where $\tilde{\alpha}$ denotes the image of α under the injection $W^* \hookrightarrow H$. Note that, here, $\nabla^2\varphi$ is to be interpreted as a distribution. Denote by M the Hilbert-Schmidt operator defined by

$$M = E [\nabla\varphi \otimes \nabla\varphi] + 2E [\nabla^2\varphi] .$$

We have

$$q(\alpha, \alpha) = ((I_H + M)\tilde{\alpha}, \tilde{\alpha})_H .$$

Let $I_H + N$ be the positive square root of the (positive) operator $I_H + M$, then N is a symmetric Hilbert-Schmidt operator. By the uniqueness of $\nabla\varphi$, it is clear that we can choose φ as

$$\varphi = \frac{1}{2}\delta^2 N$$

upto a constant: φ is a 1-convex element of $\mathbb{D}_{2,1}$, moreover the map T defined by $T = I_W + \nabla\varphi = I_W + \delta N$ satisfies the identity (5.5), hence T is the unique solution of the Monge problem and $(I_W \times T)\mu$ is the unique solution of MKP for $\Sigma(\mu, \nu)$.

6 Strong solutions of the Monge-Ampère equation for H -log-concave densities

Assume that $L \in \mathbb{L}_{+,1}^1(\mu)$ is of the form

$$L = \frac{1}{E[e^{-f}]} e^{-f},$$

where f is an H -convex function. We suppose that L is in $\mathbb{L} \log \mathbb{L}(\mu)$, $p > 1$. Denote by $\varphi \in \mathbb{D}_{2,1}$ the potential of the transport problem between μ and $\nu = L \cdot \mu$ which is a 1-convex function. This means that the mapping defined by $T = I_W + \nabla \varphi$ satisfies $T\mu = L \cdot \mu$ and $(I_W \times T)\mu$ is the unique solution of the Monge-Kantorovitch problem in $\Sigma(\mu, \nu)$ with the singular quadratic cost function $c(x, y) = |x - y|_H^2$. Let $\Lambda = 1/L \circ T$, it is immediate that Λ is well-defined since $L \circ T \neq 0$ μ -almost surely. There exists an inverse map T^{-1} , of the form $I_W + \eta$ defined ν -almost everywhere such that $T^{-1}\nu = \mu$ and that $\eta \in L^2(\nu, H)$. If ∇ is closable with respect to ν , then η is of the form $\eta = \nabla \psi$ with $\psi \in \mathbb{L}^2(\nu)$ (cf. Remark 3.3). Let $L_n = E[P_{1/n} L | V_n]$, where V_n is the sigma algebra generated by the first n elements of an orthonormal basis $(e_n, n \geq 1)$ of H and $P_{1/n}$ is the Ornstein-Uhlenbeck semigroup at $t = 1/n$. It follows from [9] and from the positivity improving property of the Ornstein-Uhlenbeck semigroup, that $L_n > 0$ μ -a.s. and that it is of the form $\frac{1}{c} e^{-f_n}$, where f_n is a smooth H -convex function on W and $c = E[e^{-f}]$. We denote by $\varphi_n, \Lambda_n, \psi_n$ the maps associated to L_n , i.e., $T_n = I_W + \nabla \varphi_n$ maps μ to the measure $L_n \cdot \mu$ and $S_n = I_W + \nabla \psi_n$ maps $L_n \cdot \mu$ to μ . Besides, from [7], $x \rightarrow x + \nabla \varphi_n(x)$ is a 1-Lipschitz map, i.e.,

$$|x + \nabla \varphi_n(x) - y - \nabla \varphi_n(y)| \leq |x - y|,$$

for any $x, y \in \mathbb{R}^n$, here it is remarkable that the Lipschitz constant is one and it is independent of the dimension of the underlying space. Moreover, we know already that the operator valued map $I_H + \nabla^2 \varphi_n \geq 0$, hence φ_n is also a concave function in $\mathbb{D}_{2,2}$. Therefore $\mathcal{L}\varphi_n$ is a well-defined element of $L^2(\mu)$, $|\nabla \varphi_n|_H^2$ is uniformly exponentially integrable, i.e., there exists some $t > 0$ such that

$$\sup_n E [\exp t |\nabla \varphi_n|_H^2] < \infty, \quad (6.6)$$

then the Fatou Lemma implies that

$$E [\exp t |\nabla \varphi|_H^2] < \infty.$$

It follows in particular that $(\varphi_n, n \geq 1) \subset \mathbb{D}_{p,2}$ and it converges to φ in $\mathbb{D}_{p,1}$ for any $p \geq 1$, cf., [11]. Moreover, from the Jacobi formula, we have

$$\int_W g \circ T_n \Lambda(\varphi_n) d\mu = \int_W g d\mu,$$

for any $g \in C_b(W)$, where

$$\Lambda(\varphi_n) = \det_2(I_H + \nabla^2 \varphi_n) \exp \left\{ -\mathcal{L}\varphi_n - \frac{1}{2} |\nabla \varphi_n|_H^2 \right\}.$$

In this case we have $\Lambda(\varphi_n) = 1/L_n \circ T_n$, and since $(L_n, n \geq 1)$ is uniformly integrable, using the Lusin theorem and the convergence in $\mathbb{L}^0(\mu)$ of $(T_n, n \geq 1)$ to T , we can show as in Section 5.3 of [28], that $(\Lambda(\varphi_n), n \geq 1)$ converges in $\mathbb{L}^0(\mu)$ to $1/L \circ T$. To make the things self-contained, let us give a quick proof of this assertion:

Lemma 6.1 *The sequence $(L_n \circ T_n, n \geq 1)$ converges to $L \circ T$ in $L^0(\mu)$.*

Proof: Let

$$l_n = \frac{L_n}{1 + L_n},$$

then it suffices to show that the sequence $(l_n \circ T_n, n \geq 1)$ converges to $l \circ T$ in $L^2(\mu)$, where $l = L/(1 + L)$. We have

$$E[(l_n \circ T_n)^2] = E[l_n^2 L_n] \rightarrow E[l^2 L] = E[(l \circ T)^2], \quad (6.7)$$

as $n \rightarrow \infty$, since the sequence $(l_n, n \geq 1)$ is bounded by one and since the sequence $(L_n, n \geq 1)$ converges to L in $L^1(\mu)$ by the martingale convergence theorem. Assume now that g is a smooth, cylindrical function, we have

$$E[l_n \circ T_n g] = E[l_n L_n g \circ S_n],$$

where S_n is the inverse of T_n , whose existence is given by Theorem 3.2. Since $(l_n, n \geq 1)$ is bounded by one and since $(L_n, n \geq 1)$ converges to L in $L^1(\mu)$, the sequence $(l_n L_n, n \geq 1)$ converges to $l L$ in $L^1(\mu)$. Moreover $(g \circ S_n, n \geq 1)$ converges in probability to $g \circ S$ and it is bounded by the bound of g , hence, it follows from the Lebesgue dominated convergence theorem that

$$\begin{aligned} \lim_n E[l_n \circ T_n g] &= \lim_n E[l_n L_n g \circ S_n] \\ &= E[l L g \circ S] \\ &= E[l \circ T g]. \end{aligned}$$

Therefore, the relatively weakly compact sequence $(l_n \circ T_n, n \geq 1)$ has a unique accumulation point in $L^2(\mu)$. It then follows from the relation (6.7) that it converges in the norm topology of $L^2(\mu)$. \square

The next result about the regularity of φ is of fundamental importance:

Theorem 6.1 *Suppose that $L \in \mathbb{L} \log \mathbb{L}(\mu)$ is H -log-concave. Then the forward transport potential φ associated to Monge-Kantorovitch problem on $\Sigma(\mu, L \cdot \mu)$, belongs to the Sobolev space $\mathbb{D}_{2,2}$. In particular $\mathcal{L}\varphi \in L^2(\mu)$ and $\det_2(I_H + \nabla^2 \varphi)$ is a well-defined function.*

We shall first prove a lemma of general interest who will imply directly Theorem 6.1:

Lemma 6.2 *With the hypothesis of Theorem 6.1, it holds that*

$$E \left[|\nabla\varphi|_H^2 + \|\nabla^2\varphi\|_2^2 \right] \leq 2E[L \log L], \quad (6.8)$$

where $\|\cdot\|_2$ denotes the Hilbert-Schmidt norm.

Proof: Assume first that $W = H = \mathbb{R}^d$ and define

$$\begin{aligned} m(t) &= -\log \Lambda(t\varphi) \\ &= t\mathcal{L}\varphi - \log \det_2(I + t\nabla^2\varphi) + \frac{t^2}{2}|\nabla\varphi|_H^2. \end{aligned}$$

It is immediate to see that

$$m'(t) = \mathcal{L}\varphi + t \operatorname{trace} \left[(I + t\nabla^2\varphi)^{-1} \nabla^2\varphi \right] + t|\nabla\varphi|_H^2$$

and that

$$m''(t) = \|(I + t\nabla^2\varphi)^{-1} \nabla^2\varphi\|_2^2 + |\nabla\varphi|_H^2.$$

Note that, since the eigenvalues of $\nabla^2\varphi$ are in the interval $[-1, 0]$, we have $m''(t) \geq m''(0)$. Hence $m(1) = m'(0) + \frac{1}{2}m''(\tau) \geq m'(0) + \frac{1}{2}m''(0)$, which implies that

$$|\nabla\varphi|_H^2 + \|\nabla^2\varphi\|_2^2 \leq 2m(1) = -2\log \Lambda(\varphi) - 2\mathcal{L}\varphi.$$

Taking the expectation of both sides gives

$$\begin{aligned} E \left[|\nabla\varphi|_H^2 + \|\nabla^2\varphi\|_2^2 \right] &\leq -2E[\log \Lambda(\varphi)] \\ &= 2E[L \log L]. \end{aligned} \quad (6.9)$$

□

Proof of Theorem 6.1: Recall now that φ_n denotes the transport potential associated to the measure $d\nu_n = L_n d\mu$, where $L_n = E[P_{1/n}L|V_n]$, hence it follows from the proof of Theorem 4.1 of [11], that $(\varphi_n, n \geq 1)$ converges to φ in $\mathbb{D}_{2,1}$. By the Jensen inequality, we get from (6.9)

$$\begin{aligned} \sup_n E \left[|\nabla\varphi_n|^2 + \|\nabla^2\varphi_n\|_2^2 \right] &\leq \sup_n -2E[\log \Lambda(\varphi_n)] \\ &= \sup_n 2E[L_n \log L_n] \\ &\leq 2[L \log L], \end{aligned}$$

consequently φ belongs to $\mathbb{D}_{2,2}$.

□

Corollary 6.1 *Under the hypothesis of Theorem 6.1, we have*

$$E[g \circ T \Lambda(\varphi)] \leq E[g],$$

for any positive, measurable function g . In particular $\Lambda(\varphi)$ is a sub-solution of the Monge-Ampère equation in the sense that

$$\Lambda(\varphi) L \circ T \leq 1$$

μ -almost surely.

Proof: Since $T = I_W + \nabla \varphi$ is a monotone shift, the first inequality follows from Theorem 6.3.1 of [28]. For the second one we use the first one:

$$\begin{aligned} E[g \circ T L \circ T \Lambda(\varphi)] &\leq E[g L] \\ &= E[g \circ T], \end{aligned}$$

for any positive, measurable function g . Since T has a left inverse, the sigma algebra generated by it is equal to the Borel sigma algebra of W upto μ -negligeable sets. Therefore we get the second claim. \square

We can prove now the main theorem of this section:

Theorem 6.2 *Let $L \in \mathbb{L} \log \mathbb{L}(\mu)$ be given as $c^{-1} e^{-f}$, where f is an H -convex Wiener function and define the probability measure ν as $d\nu = L d\mu$, where $c = E[e^{-f}]$ is the normalization constant. Let $T = I_W + \nabla \varphi$ be the optimal transportation of μ to ν in the sense of Wasserstein distance, where $\varphi \in \mathbb{D}_{2,1}$ is the forward 1-convex potential function associated to the Monge-Kantorovitch problem on $\Sigma(\mu, \nu)$. Then $\varphi \in \mathbb{D}_{2,2}$ and the Gaussian Jacobian of T exists and it satisfies the Monge-Ampère equation:*

$$\Lambda(\varphi) L \circ T = 1$$

μ -almost surely, where

$$\Lambda(\varphi) = \det_2(I_H + \nabla^2 \varphi) \exp \left\{ -\mathcal{L}\varphi - \frac{1}{2} |\nabla \varphi|_H^2 \right\}. \quad (6.10)$$

Proof: We have prepared everything necessary for the proof. First, we can form a sequence, denoted by $(\varphi'_n, n \geq 1)$ such that each φ'_n is obtained as a convex combination from the elements of the tail sequence $(\varphi_k, k \geq n)$ and that the sequence $(\varphi'_n, n \geq 1)$ converges to φ in $\mathbb{D}_{2,2}$. Let us denote the Jacobian written with φ'_n by $\Lambda_n(\varphi'_n)$ whose explicit expression is given as

$$\Lambda(\varphi'_n) = \det_2(I + \nabla^2 \varphi'_n) \exp \left\{ -\mathcal{L}\varphi'_n - \frac{1}{2} |\nabla \varphi'_n|_H^2 \right\}$$

Let $T'_n = I_W + \nabla \varphi'_n$ and $S'_n = I_W + \nabla \psi'_n$. Since $A \rightarrow -\log \det_2(I_H + A)$ is a convex function on the space of symmetric Hilbert-Schmidt operators which are lower bounded by $-I_H$ (cf. [3], p.63), we have

$$\begin{aligned} -\log \Lambda(\varphi'_n) &= -\log \det_2 \left(I_H + \sum_i t_i \nabla^2 \varphi_{n_i} \right) \\ &\quad + \sum_i t_i \mathcal{L} \varphi_{n_i} + \frac{1}{2} \left\| \sum_i t_i \nabla \varphi_{n_i} \right\|_H^2 \\ &\leq \sum_i -t_i \log \Lambda(\varphi_{n_i}). \end{aligned}$$

It follows from Lemma 6.1 that the sequence $(L_n \circ T_n, n \geq 1)$ converges to $L \circ T$ in probability, moreover, for any $n \geq 1$,

$$\Lambda(\varphi_n) L_n \circ T_n = 1$$

almost surely. Therefore the sequence $(-\log \Lambda(\varphi_n), n \geq 1)$ converges to $\log L \circ T$ in probability. Besides, from the construction, $(-\log \Lambda_n(\varphi'_n), n \geq 1)$ converges to $-\log \Lambda(\varphi)$ in probability. Consequently, it follows from the convexity inequality above that

$$-\log \Lambda(\varphi) \leq \log L \circ T$$

almost surely. Hence $(L \circ T)^{-1} \leq \Lambda(\varphi)$ almost surely. Finally Corollary 6.1 implies that $(L \circ T)^{-1} = \Lambda(\varphi)$ almost surely and this completes the proof. \square

Remark 6.1 *In our proof we have used the fact that $\nabla^2 \varphi$ has its eigenvalues in the interval $[-1, 0]$. However, proceeding as in [13], one can show the Hilbert-Schmidt property of $\nabla^2 \varphi$ from the essential boundedness of the operator norm $\|I_H + \nabla^2 \varphi\|$ under some regularity hypothesis on L but without the log-concavity assumption.*

The following corollary gives the exact value of the Wasserstein distance:

Corollary 6.2 *With the hypothesis of Theorem 6.2, we have*

$$\frac{1}{2} d_H^2(\mu, L \cdot \mu) = E[L \log L] + E[\log \det_2(I_H + \nabla^2 \varphi)].$$

In particular, if

$$L = \frac{1_A}{\mu(A)},$$

where A is an H -convex set, then we have

$$\Lambda(\varphi) = \mu(A) \quad \mu - a.s..$$

Consequently

$$\mu(A) = \exp \left\{ -\frac{1}{2} d_H^2(\mu, L \cdot \mu) + E[\log \det_2(I + \nabla^2 \varphi)] \right\}. \quad (6.11)$$

In particular we have

$$\mu(A) \leq \exp \left(-\frac{1}{2} E[q_A^2] + E[\log \det_2(I_H + \nabla^2 \varphi)] \right),$$

where q_A is the H -gauge function of A defined by

$$q_A(w) = \inf(|h|_H : h \in A - w).$$

Proof: Since $\Lambda = c e^{f \circ T}$, it follows from the theorem that

$$\begin{aligned} \frac{1}{2} d_H^2(\mu, L \cdot \mu) &= \frac{1}{2} E[|\nabla \varphi|_H^2] \\ &= -E[f \circ T] - \log c + E[\log \det_2(I_H + \nabla^2 \varphi)] \\ &= E[L \log L] + E[\log \det_2(I_H + \nabla^2 \varphi)]. \end{aligned}$$

In particular, the fact that $E[\log \det_2(I_H + \nabla^2 \varphi)]$ is always negative explains the defect in the Talagrand inequality [25]. To prove the last part, note that we have $\mu(T^{-1}(A)) = 1$, i.e., $1_A \circ T = 1$ μ -almost surely. Consequently

$$\begin{aligned} 1 &= \Lambda(\varphi) \frac{1_A \circ T}{\mu(A)} \\ &= \Lambda(\varphi) \frac{1}{\mu(A)}. \end{aligned}$$

Taking the logarithm of this equality and taking its expectation afterwards gives immediately the formula (6.11). The last claim follows from the relation (6.11) and from the fact that

$$E[q_A^2] \leq d_H^2(\mu, L \cdot \mu) = E[|\nabla \varphi|_H^2].$$

□

Let us give an interesting result about the upper bound of the interpolated density whose proof makes use also the convexity results as in the proof of Theorem 6.2 :

Proposition 6.1 *Assume the validity of the hypothesis of Theorem 6.2, suppose furthermore that the density L is almost surely bounded, i.e., its exponent is almost surely lower bounded by some $-\alpha$, $\alpha > 0$. Let T_t be defined as $T_t = I_W + t \nabla \varphi$, $t \in [0, 1]$, then the Radon-Nikodym density $L_t = d(T_t \mu)/d\mu$ is also bounded:*

$$L_t \leq \frac{1}{c} \exp \alpha t$$

almost surely, where $c = E[\exp -f]$.

Proof: Let g be any positive, measurable function on W , by the convexity of $t \rightarrow -\log \Lambda_t$, we have $-\log \Lambda_t \leq -t \log \Lambda$. Therefore

$$\begin{aligned} E[L_t \log L_t g] &= E[-\log \Lambda_t g \circ T_t] \\ &\leq E[-t \log \Lambda g \circ T_t] \\ &= E[-t(f \circ T + \log c)g \circ T_t] \\ &\leq E[(t\alpha - \log c)g \circ T_t] \\ &= E[(t\alpha - \log c)L_t g]. \end{aligned}$$

Consequently

$$L_t \log L_t \leq (t\alpha - \log c) L_t$$

almost surely. □

7 Itô-solutions of the Monge-Ampère equation

In the following calculations we shall take W as the classical Wiener space $W = C_0([0, 1], \mathbb{R})$, $H = H^1$, i.e., the Sobolev space $W_{2,1}([0, 1])$. We note that this choice does not entail any restriction of generality as indicated in [28], Chapter 2.6. Suppose we are given a positive random variables $L = \frac{1}{c}e^{-f}$ whose expectation is equal to one, c being the normalization constant. Define the measure ν as $d\nu = Ld\mu$. We shall suppose that the Wasserstein distance $d_H(\mu, \nu)$ is finite, hence the conclusions of Theorem 3.1 are valid. In order to simplify the discussion we shall assume that L is strictly positive. The transport map T can be represented as $T = I_W + \nabla\varphi$ again with $\varphi \in \mathbb{D}_{2,1}$. Define now

$$\Lambda = \frac{1}{L \circ T}.$$

We have

$$\int g \circ T \Lambda d\mu = \int g d\mu,$$

for any $g \in C_b(W)$. This implies that the process $(T_t, t \in [0, 1])$ defined on $[0, 1] \times W$ by

$$(t, x) \rightarrow T_t(x) = x(t) + \int_0^t D_\tau \varphi(x) d\tau,$$

is a Wiener process under the measure $\Lambda d\mu$ with respect to its natural filtration $(\mathcal{F}_t^T, t \in [0, 1])$, where $D_t \varphi$ represents the Lebesgue density of the map $t \rightarrow \nabla \varphi(x)(t) \in H$ on $[0, 1]$. Since T is invertible, we have also

$$\bigvee_{t \in [0, 1]} \mathcal{F}_t^T = \mathcal{B}(W),$$

upto μ -negligeable sets. Since $\Lambda d\mu$ is equivalent to the Wiener measure, the process $(T_t, t \in [0, 1])$ is a μ -semimartingale with respect to its natural filtration. It is clear that it has a decomposition of the form

$$T_t = B_t^T + A_t,$$

with respect to μ , where B^T is a μ -Brownian motion and A is a process of finite variation. Since we are dealing with the Brownian filtrations, $(A_t, t \in [0, 1])$ should be absolutely continuous with respect to the Lebesgue measure dt of $[0, 1]$. In order to calculate its Lebesgue density it suffices to calculate the limit

$$\lim_{h \rightarrow 0} \frac{1}{h} E [T_{t+h} - T_t | \mathcal{F}_t^T].$$

To calculate this limit, it is enough to test it on the functions of the type $g \circ T_t$ with a smooth function g :

$$\begin{aligned} E [(T_{t+h} - T_t) g \circ T_t] &= E [(W_{t+h} - W_t) g \circ W_t L] \\ &= E [(\delta U_{[t, t+h]}) g \circ W_t L] \\ &= E [(U_{[t, t+h]}, \nabla(L g \circ W_t))_H] \end{aligned} \quad (7.12)$$

$$= E \left[g \circ W_t \int_t^{t+h} D_\tau L d\tau \right], \quad (7.13)$$

where $U_{[t, t+h]}$ is the element of H whose Lebesgue density is equal to the indicator function of the interval $[t, t+h]$. Note that for the equality (7.12), we have used the fact that $\delta = \nabla^*$ under the Wiener measure μ and the equality (7.13) follows from the fact that the support of $\nabla(g(W_t))$ lies in the interval $[0, t]$, hence its scalar product in H with $U_{[t, t+h]}$ is zero (cf.[27]). Hence we have

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{1}{h} E [T_{t+h} - T_t | \mathcal{F}_t^T] &= -E[D_t f \circ T | \mathcal{F}_t^T] \\ &= -E_\nu[D_t f | \mathcal{F}_t] \circ T, \end{aligned}$$

$dt \times d\mu$ -almost surely, where the last inequality follows from the fact that $T^{-1}(\mathcal{F}_t) = \mathcal{F}_t^T$. Hence we have proven

Proposition 7.1 *The transport process $(T_t, t \in [0, 1])$ is a $(\mu, (\mathcal{F}_t^T))$ -semimartingale with its canonical decomposition*

$$\begin{aligned} T_t &= B_t^T - \int_0^t E_\nu [D_\tau f | \mathcal{F}_\tau] \circ T d\tau \\ &= B_t^T - \int_0^t E [D_\tau f \circ T | \mathcal{F}_\tau^T] d\tau. \end{aligned}$$

We can give now the Itô solution of the Monge-Ampère equation:

Theorem 7.1 Assume that $f \in \mathbb{D}_{2,1}$ be such that $c = E[\exp(-f)] < \infty$, denote by L the probability density defined by $\frac{1}{c}e^{-f}$ and by ν the probability $d\nu = Ld\mu$. Assume that $d_H(\mu, \nu) < \infty$ and let $T = I_W + \nabla\varphi$ be the transport map whose properties are announced in Theorem 3.2. We have then

$$\Lambda = \exp \left\{ \int_0^1 E_\nu[D_t f | \mathcal{F}_t] \circ T dB_t^T - \frac{1}{2} \int_0^1 E_\nu[D_t f | \mathcal{F}_t]^2 \circ T dt \right\}. \quad (7.14)$$

Proof: From the Itô representation formula [26], we have

$$L = \exp \left\{ - \int_0^1 E_\nu[D_t f | \mathcal{F}_t] dW_t - \frac{1}{2} \int_0^1 E_\nu[D_t f | \mathcal{F}_t]^2 dt \right\}.$$

Since the Girsanov measure for T has the density Λ given by

$$\Lambda = \frac{1}{L \circ T},$$

we have, using the identity $T^{-1}(\mathcal{F}_t) = \mathcal{F}_t^T$ and Proposition 7.1,

$$\begin{aligned} L \circ T &= \exp \left\{ - \int_0^1 E_\nu[D_t f | \mathcal{F}_t] \circ T dT_t - \frac{1}{2} \int_0^1 E_\nu[D_t f | \mathcal{F}_t]^2 \circ T dt \right\} \\ &= \exp \left\{ - \int_0^1 E_\nu[D_t f | \mathcal{F}_t] \circ T (dB_t^T - E_\nu[D_t f | \mathcal{F}_t] \circ T dt) \right. \\ &\quad \left. - \frac{1}{2} \int_0^1 E_\nu[D_t f | \mathcal{F}_t]^2 \circ T dt \right\}, \end{aligned}$$

which is exactly the inverse of the expression given by the relation (7.14). \square

The following proposition explains the relation between the semimartingale representation of T and the polar factorization studied in Section 4:

Proposition 7.2 Let X be the process defined by

$$X_t = W_t + \int_0^t E_\nu[D_\tau f | \mathcal{F}_\tau] d\tau,$$

then $T \circ X$ is a ν -rotation, i.e., $T \circ X(\nu) = \nu$, in fact it is the minimal ν -rotation in the sense that

$$\inf_{O \in \mathcal{R}_\nu} E_\nu[|O - X|_H^2] = E_\nu[|T \circ X - X|_H^2],$$

where \mathcal{R}_ν denotes the set of transformations preserving the measure ν . Finally the Brownian motion B^T is the rotation corresponding to $X \circ T$.

Proof: Since $E[L] = 1$, ν is the Girsanov measure for the transformation X , consequently, we have

$$\begin{aligned} E_\nu[g(T \circ X)] &= E[g(T \circ X) L] \\ &= E[g(T)] \\ &= E_\nu[g], \end{aligned}$$

for any $g \in C_b(W)$ and this implies $(T \circ X)\nu = \nu$. Let now $\mathcal{O} \in \mathcal{R}_\nu$, then the measure $(\mathcal{O} \times X)\nu$ belongs to $\Sigma(\nu, \mu)$. Since $(T \times I_W)\mu$ is the solution of MKP in $\Sigma(\nu, \mu)$, we have

$$E_\nu[|\mathcal{O} - X|_H^2] \geq E_\nu[|T \circ X - X|_H^2] = d_H(\mu, \nu)^2.$$

The uniqueness follows from the same argument as used in the proof of Theorem 4.1. The last claim is obvious since $X \circ T$ is a μ -rotation, hence as a process it is a (μ -) Brownian motion, then by comparing it with the result of Proposition 7.1, we see that $B^T = X \circ T$. \square

Remark 7.1 It is to be noted that

$$\begin{aligned} E[|X \circ T - T|_H^2] &= E_\nu[|X - I_W|_H^2] \\ &= E[L|X - I_W|_H^2] \\ &= E\left[L \int_0^1 E[D_s f | \mathcal{F}_s]^2 ds\right] \\ &= 2E[L \log L] \\ &\geq d_H^2(\mu, \nu) \\ &= E_\nu[|T \circ X - X|_H^2]. \end{aligned}$$

Let us give some immediate consequences of these results whose proofs follow from the results of this section and from Theorem 6.2 :

Corollary 7.1 *We have the following identity*

$$\begin{aligned} -\log E[e^{-f}] &= E\left[f \circ T + \frac{1}{2} \int_0^1 E_\nu[D_t f | \mathcal{F}_t]^2 \circ T dt\right] \\ &= E\left[f \circ T + \frac{1}{2} \int_0^1 E[D_t f \circ T | \mathcal{F}_t^T]^2 dt\right]. \end{aligned}$$

If, furthermore, f is H -convex, then we also have

$$-\log E[e^{-f}] = E\left[f \circ T - \log \det_2(I_H + \nabla^2 \varphi) + \frac{1}{2} |\nabla \varphi|_H^2\right].$$

In particular we have the exact characterization of the Wasserstein distance between μ and ν :

$$\begin{aligned}\frac{1}{2}d_H^2(\mu, \nu) &= E [\log \det_2(I_H + \nabla^2 \varphi)] + \frac{1}{2}E \left[\int_0^1 E_\nu[D_t f | \mathcal{F}_t]^2 \circ T dt \right] \\ &= E [\log \det_2(I_H + \nabla^2 \varphi)] + E[L \log L].\end{aligned}$$

References

- [1] P. Appell: “Mémoire sur déblais et les remblais des systèmes continus ou discontinus”. Mémoires présentées par divers savants à l’Académie des Sciences de l’Institut de France. Paris, I. N. **29**, 1-208, 1887.
- [2] P. Appell: “Le problème géométrique des déblais et des remblais”. Mémorial des Sciences Mathématiques, fasc. **XXVII**, Paris, 1928.
- [3] E.F. Beckenbach and R. Bellman: *Inequalities*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 30. Springer-Verlag, 1983.
- [4] Y. Brenier: “Polar factorization and monotone rearrangement of vector valued functions”. Comm. pure Appl. Math, **44**, 375-417, 1991.
- [5] N. Dunford and J. T. Schwartz: *Linear Operators* **2**, Interscience 1963.
- [6] L. A. Caffarelli: “The regularity of mappings with a convex potential”. Jour. Amer. Math. Soc., **5**, 99-104, 1992.
- [7] L. A. Caffarelli: “Monotonicity properties of optimal transportation and the FKG and related inequalities”. Commun. Math. Phys. **214**, 547-563, 2000.
- [8] D. Feyel: “A survey on the Monge transport problem”. Preprint, 2004.
- [9] D. Feyel and A. S. Üstünel: “The notion of convexity and concavity on Wiener space”. Journal of Functional Analysis, **176**, 400-428, 2000.
- [10] 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.
- [11] D. Feyel and A. S. Üstünel: “Monge-Kantorovitch measure transportation and Monge-Ampère equation on Wiener space”. Prob. Theory and Related Fields, **128**, 347-385, 2004.
- [12] D. Feyel and A. S. Üstünel: “Monge-Kantorovitch measure transportation, Monge-Ampère equation and the Itô calculus”. Advanced Studies in Pure Mathematics, pp. 49-74 Vol. **41**. Mathematical Society of Japan, 2004.

- [13] D. Feyel and A. S. Üstünel: “The strong solution of the Monge-Ampère equation on the Wiener space for log-concave densities”. *Comptes Rendus Mathématiques*, Vol. **339**, Issue 1, 49-53, 2004.
- [14] W. Gangbo and R. J. McCann: “The geometry of optimal transportation”. *Acta Mathematica*, **177**, 113-161, 1996.
- [15] L. V. Kantorovitch: “On the transfer of masses”. *Dokl. Acad. Nauk. SSSR* **37**, 227-229, 1942.
- [16] R. J. McCann: “Existence and uniqueness of monotone measure-preserving maps”. *Duke Math. Jour.*, **80**, 309-323, 1995.
- [17] R. J. McCann: “A convexity principle for interacting gases”. *Advances in Mathematics*, **128**, 153-179, 1997.
- [18] G. Monge: “Mémoire sur la théorie des déblais et des remblais”. *Histoire de l’Académie Royale des Sciences*, Paris, 1781.
- [19] S.T. Rachev and L. Rüchendorf: “A characterization of random variables with minimum L^2 -distance”. *Journal of Multivariate Analysis*, **32**, 48-54, 1990.
- [20] S.T. Rachev and L. Rüchendorf: *Mass Transportation Problems*, Vol. I and II. Probability and its Applications, Springer, 1998.
- [21] T. Rockafellar: *Convex Analysis*. Princeton University Press, Princeton, 1972.
- [22] J. Schwartz: *Nonlinear Functional Analysis*. Gordon and Breach, 1969.
- [23] B. Simon: *Functional Integration and Quantum Physics*. Academic Press, 1979.
- [24] V. N. Sudakov: “Geometric problems in the theory of infinite dimensional probability distributions”. *Proc. Steklov Inst. Math.*, **141**, 1-178, 1979.
- [25] M. Talagrand: “Transportation cost for Gaussian and other product measures”. *Geom. Funct. Anal.*, **6**, 587-600, 1996.
- [26] A. S. Üstünel: “Representation of distributions on Wiener space and Stochastic Calculus of Variations”. *Journal of Functional Analysis*, **70**, 126-139, 1987.
- [27] A. S. Üstünel: *Introduction to Analysis on Wiener Space*. Lecture Notes in Math. Vol. **1610**. Springer, 1995.

- [28] A. S. Üstünel and M. Zakai: *Transformation of Measure on Wiener Space*. Springer Monographs in Mathematics. Springer Verlag, 1999.
 - [29] A. S. Üstünel: “Damped logarithmic Sobolev inequality on the Wiener space”. Stochastic Analysis and Related Topics VII, The Silivri Workshop, p. 245-249. Progress in Probability, Vol. **48**. Birkhäuser, 2001.
 - [30] C. Villani: *Topics in Optimal Transportation*. Graduate Series in Math., **58**. Amer. Math. Soc., 2003.
- D. Feyel, Université d’Evry-Val-d’Essone, 91025 Evry Cedex, France. E-mail: feyel@maths.univ-evry.fr
 - A. S. Üstünel, ENST, Dépt. Infres, 46, rue Barrault, 75634 Paris Cedex 13, France. E-mail: ustunel@enst.fr

