ON THE MONGE–AMPÈRE EQUATION IN INFINITE DIMENSIONS

VLADIMIR I. BOGACHEV, ALEXANDER V. KOLESNIKOV

ABSTRACT. We prove that the optimal transportation mapping that takes a Gaussian measure γ on an infinite dimensional space to an equivalent probability measure $g \cdot \gamma$ satisfies the Monge–Ampère equation provided that $\log g \in L^1(\gamma)$ and $g \log g \in L^1(\gamma)$.

1. INTRODUCTION AND MAIN RESULT

The Monge–Kantorovich problem and the Monge–Ampère equation have become a very popular object of research in the last decade (see [1], [15], [20], where one can find additional references). In the finite dimensional case, considerable progress has been achieved by Brenier [5] and McCann [13], whose works stimulated a growing flow of publications. Among important earlier contributions one should mention Sudakov's research [16]. In this paper we are interested in the infinite dimensional situation and extend several recent results from [9], [10], [11]. Our principal contribution is a derivation of the Monge–Ampère equation for transformations of Gaussian measures on infinite dimensional spaces. We shall use the following important existence result from [9]. Let X be a locally convex space and let γ be a centered Radon Gaussian measure on X with the Cameron–Martin space H. The natural inner product in H is denoted by $\langle \cdot, \cdot \rangle_H$; the corresponding norm is $|\cdot|_H$. One may assume without loss of generality that $X = \mathbb{R}^\infty$, the countable power of the real line, and that γ is the countable power of the standard Gaussian measure; then $H = l^2$. Suppose that we are given a probability measure $g \cdot \gamma$ such that

$$W_H(\gamma, g\cdot\gamma)^2 = \inf_{m \in \mathcal{P}(\gamma, g\cdot\gamma)} \int_{X\times X} |x_1 - x_2|_H^2 \, dm(x_1, x_2) < \infty,$$

where $\mathcal{P}(\gamma, g\cdot\gamma)$ is the set of all Radon probability measures on $X \times X$ whose projections on the first and second factors are γ and $g\cdot\gamma$. Then there exists a unique Borel mapping $T\colon\ X \to X$ sending γ to $g\cdot\gamma$ such that $W_H(\gamma, g\cdot\gamma)^2 = \displaystyle\int_X |T(x) - x|_H^2 \, d\gamma$. This mapping is called the optimal transportation plan or the optimal transportation mapping. An effective sufficient condition for $W_H(\gamma, g\cdot\gamma)$ to be finite is the finiteness of entropy

$$\mathrm{Ent}_\gamma g := \int_X g \log g \, d\gamma < \infty.$$

This is a consequence of the Talagrand inequality ([18], see also [12] for other inequalities of this type). This transportation plan has the form $T = I + \nabla\Phi$, where Φ belongs to the Sobolev class $W^{2,1}(\gamma)$ and is 1-convex (see the definition below). If $g > 0$ γ-a.e. and $\log g \in L^1(\gamma)$, then there exists a mapping S such that T and S are reciprocal, i.e., one has

$$T \circ S(x) = S \circ T(x) = x \quad \text{for } \gamma\text{-a.e. } x.$$

1991 *Mathematics Subject Classification.* 28C20, 35A30, 46G12, 49Q20, 60H05.

Key words and phrases. Gaussian measure, optimal transportation, Monge-Kantorovich problem, Monge-Ampère equation.

Moreover, S realizes the optimal transportation plan that takes $g\cdot\gamma$ to γ and $S = I + \nabla\Psi$, where $\Psi \in W^{2,1}(\gamma)$ is 1-convex.

The Monge–Kantorovich problem can be considered also from the point of view of partial differential equations. Suppose we are given two probability measures $f\,dx$ and $g\,dx$ on $\mathbb{R}^n$ and the corresponding optimal transportation plan T, which is known to be the gradient of a convex function V. Performing formally the change of variables we obtain $f = g(\nabla V)\det D^2V$. This formula is a partial case of the Monge–Ampère equation. The following rigorous result was obtained by McCann [14] (see also [20]).

Theorem 1.1. (McCann) *Let $\mu = f\,dx$ and $\nu = g\,dx$ be two absolutely continuous probability measures on $\mathbb{R}^n$ such that μ is equivalent to Lebesgue measure and let V be a convex function such that ∇V takes μ to ν. Let $\det(D^2_{\rm ac}V)$ be the determinant of the density $D^2_{\rm ac}V$ of the absolutely continuous part of D^2V (i.e., the determinant in Alexandroff's sense). Let M be the set of points where $D^2_{\rm ac}V$ is defined and invertible. Then M is of full μ-measure and for almost all $x\in M$ one has*

$$f(x) = g(\nabla V(x))\det D^2_{\rm ac}V(x).$$

Before we discuss the situation in the infinite dimensional case, let us introduce some notation. Let $L^2(\gamma, H)$ denote the Hilbert space of H-valued γ-measurable mappings v with finite norm

$$\|v\|_{L^2(\gamma,H)} = \Bigl(\int_X |v(x)|_H^2\,\gamma(dx)\Bigr)^{1/2}.$$

The Hilbert–Schmidt norm of a symmetric operator A on H is defined by

$$\|A\|_{\mathcal{H}} = \Bigl(\sum_{i=1}^{\infty}(Ae_i, Ae_i)\Bigr)^{1/2},$$

where $\{e_i\}$ is any orthonormal basis in H. Every vector $h\in H$ corresponds to a γ-measurable linear functional $\widehat{h}$ on X such that $\langle h,k\rangle_H = (\widehat{h},\widehat{k})_{L^2(\gamma)}$ for all $k\in H$ and

$$l(h) = \int_X l(x)\widehat{h}(x)\,\gamma(dx)$$

for all $l\in X^*$. The functional $\widehat{h}$ belongs to the closure of X^* in $L^2(\gamma)$; see [2] for details. Set $x_i := \widehat{e}_i(x)$. As noted above, one may assume that we deal with the standard Gaussian product-measure on $\mathbb{R}^\infty$ and then $\widehat{e}_i$ is the usual ith coordinate function. The σ-algebra generated by $\widehat{e}_1,\ldots,\widehat{e}_n$ is denoted by $\mathcal{F}_n$. The space of smooth cylindrical functions, denoted by $\mathcal{F}C_b^\infty$, consists of all functions of the form $\zeta(x_1,\ldots,x_n)$, where $\zeta\in C_b^\infty(\mathbb{R}^n)$ for some n.

The Sobolev class $W^{2,1}(\gamma)$ consists of all functions $f\in L^2(\gamma)$ that have a generalized gradient $\nabla f\in L^2(\gamma,H)$ along H such that

$$\int_X \partial_h\varphi\, f\,d\gamma = -\int_X \varphi\langle\nabla f, h\rangle_H\,d\gamma + \int_X \varphi f\widehat{h}\,d\gamma$$

for all $h\in H$ and $\varphi\in\mathcal{F}C_b^\infty$, where ∂_h is the partial derivative along h. The Sobolev class $W^{2,1}(\gamma,H)$ of H-valued mappings is defined in a similar way (see [2] for details). The Sobolev class $W^{2,2}(\gamma)$ consists of all twice weakly H-differentiable functions with finite norm

$$\|f\|_{W^{2,2}(\gamma)} = \Bigl(\int_X f^2\,d\gamma + \int_X |\nabla f|_H^2\,d\gamma + \int_X \|D^2 f\|_{\mathcal{H}}^2\,d\gamma\Bigr)^{\frac12}.$$

The Ornstein–Uhlenbeck semigroup $\{P_t\}$ on $L^p(\gamma)$, $1 \le p < \infty$, is defined by

$$P_t f(x) := \int_X f\Bigl(e^{-t}x + \sqrt{1-e^{-2t}}\, y\Bigr)\, \gamma(dy).$$

Let $\mathcal{L}$ be the generator of $\{P_t\}$ on $L^2(\gamma)$. We recall that $\mathcal{L}$ is an extension of the operator $\Delta_H f - \bigl\langle x, \nabla_H f\bigr\rangle_H := \sum_{n=1}^{\infty}(\partial^2_{e_i} f - x_i \partial_{e_i} f)$ acting on smooth cylindrical functions. The divergence of an H-valued vector field F is defined by

$$\delta F := \sum_{i=1}^{\infty}\bigl(\partial_{e_i} F^i - x_i F^i\bigr), \quad F^i = \langle F, e_i\rangle_H,$$

if F is smooth cylindrical; then divergence extends to vector fields from the Sobolev space $W^{2,1}(\gamma, H)$.

Given a function $f \in L^1(\gamma)$ such that $f \log f \in L^1(\gamma)$, one defines $\mathcal{L}f$ in the sense of distributions as the linear functional $\varphi \mapsto \int_X f \mathcal{L}\varphi\, d\gamma$ on $\mathcal{F}C_b^\infty$ (note that $f\mathcal{L}\varphi \in L^1(\gamma)$ by Young's inequality). If $f \in L^2(\gamma)$, then $\mathcal{L}f$ is a continuous linear functional on $W^{2,2}(\gamma)$. Convergence in the sense of distributions over (X, γ) is understood as pointwise convergence of linear functionals on $\mathcal{F}C_b^\infty$.

We recall the definition of a θ-convex function introduced in [8]. Let $F\colon X \to \mathbb{R} \cup \{\infty\}$ be a measurable mapping such that $\gamma(\{F < \infty\}) > 0$ and let $\theta \in \mathbb{R}^1$. Let

$$F_\theta\colon H \times X \to \mathbb{R} \cup \{\infty\}, \quad F_\theta(h, w+h) = \frac{\theta}{2}|h|_H^2 + F(w+h).$$

Then F is called θ-convex if for all $h, k \in H$ and $\alpha, \beta \in [0,1]$ with $\alpha + \beta = 1$, one has

$$F_\theta(\alpha h + \beta k, w + \alpha h + \beta k) \le \alpha F_\theta(h, w+h) + \beta F_\theta(k, w+k) \quad \gamma\text{-a.e.},$$

where the measure zero set on which this inequality fails may depend on h, k and α. See [8] for some equivalent definitions.

We recall that a Radon measure μ on X is called Skorohod differentiable along a vector $h \in X$ if there exists a Radon measure $d_h\mu$ such that for every smooth cylindrical function ζ one has

$$\int_X \partial_h \zeta(x)\, \mu(dx) = -\int_X \zeta(x)\, d_h\mu(dx).$$

Note that γ is differentiable along any $h \in H$ and $d_h\gamma = -\widehat{h} \cdot \gamma$. The second order derivative is defined inductively as $d^2_{hh}\mu := d_h(d_h\mu)$. In this paper we are especially interested in the second derivatives of the 1-convex potentials Φ and Ψ. In our case Φ and Ψ admit the first Sobolev derivatives $\nabla\Phi$ and $\nabla\Psi$ along H. We define Φ_{kh}, where $k, h \in H$, as a *Radon measure* satisfying the relation

$$\int_X \zeta(x)\, \Phi_{kh}(dx) = -\int_X \partial_h \zeta(x) \partial_k \Phi(x)\, \gamma(dx) + \int_X \zeta(x) \partial_k \Phi(x) \widehat{h}(x)\, \gamma(dx).$$

If $\partial_k \Phi$ is differentiable in the Sobolev sense, then, according to this definition, $\Phi_{kh} = \partial_h \partial_k \Phi \cdot \gamma$.

The density of the absolutely continuous part of Φ_{kh} (with respect to γ) is denoted by Φ^{ac}_{kh} and the singular part is denoted by $\Phi^{\mathrm{sing}}_{kh}$. Note that by 1-convexity $\Phi^{\mathrm{sing}}_{hh}$ is a nonnegative measure (Corollary 2.4). In the case when there exists an $\mathcal{H}$-valued measure with matrix elements $\Phi_{e_i e_j}$, we denote this measure by the symbol $D^2\Phi$. If $\sum_{i=1}^{\infty} |\Phi^{\mathrm{ac}}_{e_i e_i}(x)|^2 < \infty$ γ-a.e., then the $\mathcal{H}$-valued mapping with matrix elements $\Phi^{\mathrm{ac}}_{e_i e_j}$ is denoted by the symbol

$D^2_{\mathrm{ac}}\Phi$ (even if $D^2\Phi$ does not exist). If the measure $D^2\Phi$ exists and has bounded variation as an $\mathcal{H}$-valued measure, then $D^2_{\mathrm{ac}}\Phi$ is the density of its absolutely continuous part with respect to γ. Below we give some sufficient conditions for the existence of $D^2\Phi$. We recall that if a measure m on X with values in the Hilbert space $\mathcal{H}$ is of bounded variation, then it has the form $m = F\cdot m_0$, where m_0 is a bounded nonnegative measure on X (e.g., the total variation of m) and F is an m_0-integrable mapping with values in $\mathcal{H}$. Let m_0^{ac} be the absolutely continuous component of m_0 with respect to γ; the measure $F\cdot m_0^{\mathrm{ac}}$ is called the absolutely continuous component of m with respect to γ.

If a number n is less than the dimension of a matrix B, we denote by $B_{n\times n}$ the $n\times n$-matrix defined by $B_{n\times n}(i,j) = B(i,j)$, $1\le i\le n, 1\le j\le n$.

The conditional expectation of $f\in L^1(\gamma)$ with respect to $\mathcal{F}_n$ is denoted by $\mathbb{E}(f|\mathcal{F}_n)$. Set $P_n x = \sum_{i=1}^n \widehat{e}_i(x)e_i$. The measure γ can be represented as a direct product $\gamma = \gamma_n\otimes\widetilde{\gamma}_n$, where $\gamma_n = \gamma\circ P_n^{-1}$ and $\widetilde{\gamma}_n$ is the image of γ under the projection $x\mapsto x-P_n x$ on the space $X^{(n)} = \{z\colon\ z = x - P_n x\}$. If one deals with the standard Gaussian product-measure, then γ_n and $\widetilde{\gamma}_n$ are product-measures on the corresponding spaces. It is known (see [2]) that

$$\mathbb{E}(f|\mathcal{F}_n)(x) = \int_{X^{(n)}} f(x+z)\widetilde{\gamma}_n(dz).$$

The operator $\mathbb{E}(\cdot|\mathcal{F}_n)$ extends to bounded Radon measures as follows: $\mathbb{E}(m|\mathcal{F}_n)$ is the restriction of a measure m to the σ-algebra $\mathcal{F}_n$. It is verified directly that $P_t\mathbb{E}(f|\mathcal{F}_n) = \mathbb{E}(P_t f|\mathcal{F}_n)$.

In this paper we consider the following problem: when do the potentials Φ and Ψ satisfy an infinite dimensional analog of the Monge–Ampère equation? The heuristic formulas for the Monge–Ampère equation are

$$g = \det{}_2(I + D^2\Psi)\exp\Bigl(\mathcal{L}\Psi - \frac{1}{2}|\nabla\Psi|_H^2\Bigr), \tag{1.1}$$

$$\frac{1}{g(T)} = \det{}_2(I + D^2\Phi)\exp\Bigl(\mathcal{L}\Phi - \frac{1}{2}|\nabla\Phi|_H^2\Bigr). \tag{1.2}$$

Here $\det_2$ denotes the Carleman–Fredholm determinant which is defined for any symmetric Hilbert–Schmidt operator Λ by the formula

$$\det{}_2(I+\Lambda) = \prod_{i=1}^{\infty}(1+\lambda_i)e^{-\lambda_i},$$

where λ_i are the eigenvalues of Λ counted with their multiplicities. Note that if $I+\Lambda\ge 0$, then $\det_2(I+\Lambda)\le 1$, because $(1+\lambda)e^{-\lambda}\le 1$ for all $\lambda\ge -1$.

Diverse results on the change of variables formula for general nonlinear shifts along the Cameron–Martin space can be found in [2], [19]. However, these results do not seem to be directly applicable to our case.

As the first step one has to show that all the objects involved in equalities (1.1) and (1.2) exist indeed. It has been shown by Feyel and Üstünel in [9] that $\mathcal{L}\Phi$ (considered as a distribution on the space (X,γ)) is a Radon measure if $g < C$. The density of its absolutely continuous part with respect to γ is denoted by $\mathcal{L}_{\mathrm{ac}}\Phi$. Similarly, if $g > c > 0$, then $\mathcal{L}\Psi$ is a Radon measure, and $\mathcal{L}_{\mathrm{ac}}\Psi$ is the density of its absolutely continuous part with respect to γ. Another result from [9] states that if $0 < c < g < C$, then

$$g(T)\underline{\lim}_n \det{}_2[I + D^2_{\mathrm{ac}}\mathbb{E}(\Phi|\mathcal{F}_n)]\exp\Bigl(\mathcal{L}_{\mathrm{ac}}\Phi - \frac{1}{2}|\nabla\Phi|_H^2\Bigr)\le 1,$$

and

$$g \geq \underline{\lim}_n \det{}_2[I + D^2_{\rm ac}\mathbb{E}(\Psi|\mathcal{F}_n)] \exp\Bigl(\mathcal{L}_{\rm ac}\Psi - \frac{1}{2}|\nabla\Psi|^2_H\Bigr).$$

We see that these results give only inequalities instead of the expected equalities. However, by another result of Feyel and Üstünel from [10], if $-\log g$ is an H-convex function, which is certainly a very strong restriction, then the infinite dimensional Monge–Ampère equation holds. A uniform estimate of the second derivative of the potential Φ established by Caffarelli [7] plays an important role in the proof.

The main result of this paper is the following theorem.

Theorem 1.2. *Suppose that* $\log g \in L^1(\gamma)$, $g \log g \in L^1(\gamma)$. *Then there exist* $\mathcal{H}$*-valued mappings* $D^2_{\rm ac}\Psi$ *and* $D^2_{\rm ac}\Phi$ *with matrix elements* $\Phi^{\rm ac}_{e_ie_j}$ *and* $\Psi^{\rm ac}_{e_ie_j}$ *and a subsequence* $\{n_k\}$ *such that* γ*-a.e. there exist finite limits*

$$\mathcal{L}_0\Psi = \lim_{m\to\infty} \frac{1}{m}\sum_{k=1}^{m}\sum_{i=1}^{n_k}\bigl(\Psi^{\rm ac}_{e_ie_i} - x_i\partial_{e_i}\Psi\bigr), \quad \mathcal{L}_0\Phi = \lim_{m\to\infty} \frac{1}{m}\sum_{k=1}^{m}\sum_{i=1}^{n_k}\bigl(\Phi^{\rm ac}_{e_ie_i} - x_i\partial_{e_i}\Phi\bigr). \quad (1.3)$$

In addition,

$$g = \det{}_2(I + D^2_{\rm ac}\Psi)\exp\Bigl(\mathcal{L}_0\Psi - \frac{1}{2}|\nabla\Psi|^2_H\Bigr), \quad (1.4)$$

$$\frac{1}{g(T)} = \det{}_2(I + D^2_{\rm ac}\Phi)\exp\Bigl(\mathcal{L}_0\Phi - \frac{1}{2}|\nabla\Phi|^2_H\Bigr). \quad (1.5)$$

Furthermore, $(I + D^2_{\rm ac}\Psi)(I + D^2_{\rm ac}\Phi(S)) = (I + D^2_{\rm ac}\Phi(S))(I + D^2_{\rm ac}\Psi) = I$.

(ii) *Suppose that* $g > c > 0$ *and* $g\log g \in L^1(\gamma)$. *Then* $\mathcal{L}_0\Psi = \mathcal{L}_{\rm ac}\Psi$ *and*

$$g = \det{}_2(I + D^2_{\rm ac}\Psi)\exp\Bigl(\mathcal{L}_{\rm ac}\Psi - \frac{1}{2}|\nabla\Psi|^2_H\Bigr). \quad (1.6)$$

(iii) *Suppose that* $0 < g < C$ *and* $\log g \in L^1(\gamma)$. *Then* $\mathcal{L}_0\Phi = \mathcal{L}_{\rm ac}\Phi$ *and*

$$\frac{1}{g(T)} = \det{}_2(I + D^2_{\rm ac}\Phi)\exp\Bigl(\mathcal{L}_{\rm ac}\Phi - \frac{1}{2}|\nabla\Phi|^2_H\Bigr). \quad (1.7)$$

Here and throughout the equality $f_1 = f_2$ for measurable functions means that $f_1(x) = f_2(x)$ a.e.

2. Auxiliary results and proofs

Before proving our main theorem we make several remarks and prove some auxiliary results. Let us consider a probability measure $g\cdot\gamma$ and an approximation of g by functions $g_n \to g$ such that every g_n is measurable with respect to $\mathcal{F}_n$. We shall consider the approximations $P_{\frac{1}{n}}\mathbb{E}(g|\mathcal{F}_n)$ and $\mathbb{E}(g|\mathcal{F}_n)$. Let $\{T_n\}$ and $\{S_n\}$ be two sequences of optimal transportation plans sending γ to $g_n\cdot\gamma$ and $g_n\cdot\gamma$ to γ, accordingly. By the finite dimensional case

$$T_n = I + \nabla\Phi_n, \quad S_n = I + \nabla\Psi_n,$$

where Φ_n and Ψ_n are 1-convex functions. It is clear that T_n and S_n are reciprocal, i.e.,

$$T_n \circ S_n(x) = S_n \circ T_n(x) = x \quad \gamma\text{-a.e.}$$

It has been shown in [9] that $T_n \to T$ and $S_n \to S$ in γ-measure, hence γ-a.e. for some subsequence (this is explained in more detail in Remark 2.1 below). By the regularity theory developed by Caffarelli (see, e.g., [6] or [20]) we obtain that Φ_n and Ψ_n are twice continuously differentiable in the case of $g_n = P_{\frac{1}{n}}\mathbb{E}(g|\mathcal{F}_n)$ (see Remark 2.1(iii)).

The following important identity was proved in [11]:

$$\int_X \log \frac{g_m}{g_n} g_m \, d\gamma = \frac{1}{2}\int_X |S_n - S_m|_H^2 g_m \, d\gamma + \int_X \Bigl[\mathrm{Tr}\bigl(DS_n(DS_m)^{-1} - I\bigr) - \log\det\bigl(DS_n(DS_m)^{-1}\bigr)\Bigr] g_m \, d\gamma, \quad (2.8)$$

where DF denotes the derivative of a mapping F. Note that both integrands are nonnegative. Letting $g_n = 1$ or $g_m = 1$ we obtain the following relations:

$$\int_X \log \frac{1}{g_n} \, d\gamma = \frac{1}{2}\int_X |S_n(x) - x|_H^2 \, \gamma(dx) - \int_X \log\det{}_2 DS_n \, d\gamma, \quad (2.9)$$

$$\int_X g_m \log g_m \, d\gamma = \frac{1}{2}\int_X |S_m(x) - x|_H^2 g_m \, \gamma(dx) - \int_X \log\det{}_2\bigl[(DS_m)^{-1}\bigr] g_m \, d\gamma. \quad (2.10)$$

These formulas give the following estimates of the transport cost:

$$\frac{1}{2}\int_X |\nabla\Psi_n|_H^2 \, d\gamma \le \int_X \log\frac{1}{g_n} \, d\gamma, \quad \frac{1}{2}\int_X |\nabla\Phi_n|_H^2 \, d\gamma \le \int_X g_n \log g_n \, d\gamma. \quad (2.11)$$

The second inequality is the well-known Talagrand inequality. An immediate consequence of (2.8) is the existence of an optimal transport S sending $g\cdot\gamma$ to γ. Another useful consequence of this identity is a result on convergence of DS_n and $(DS_n)^{-1}$ (see [11]). In Theorem 2.2 below we obtain an important extension of this result.

Let us write

$$(I + K_n)^2 := DS_n = I + D^2\Psi_n, \quad (I + L_n)^2 := (DS_n)^{-1},$$

where K_n and L_n are mappings with values in the space of symmetric operators.

Remark 2.1. (i) Suppose that $\log g \in L^1(\gamma)$ and $g\log g \in L^1(\gamma)$. Throughout the paper we consider the following two types of approximations of g by cylindrical functions:

$$\widetilde{g}_n := \mathbb{E}(g|\mathcal{F}_n) \quad \text{and} \quad g_n := \mathbb{E}(P_{\frac{1}{n}} g|\mathcal{F}_n).$$

By the martingale property $\widetilde{g}_n \to g$ in $L^1(\gamma)$ and γ-a.e. Moreover, it follows from Jensen's inequality that the sequences of entropies $\int_X \widetilde{g}_n \log \widetilde{g}_n \, d\gamma$ and $\int_X \log \frac{1}{\widetilde{g}_n} \, d\gamma$ are monotone and converge to $\int_X g\log g \, d\gamma$ and $\int_X \log g \, d\gamma$, respectively.

(ii) Let $\widetilde{S}_n(x) := x + \nabla\widetilde{\Psi}_n(x)$ and $\widetilde{T}_n(x) := x + \nabla\widetilde{\Phi}_n(x)$ be the optimal transports taking $\widetilde{g}_n\cdot\gamma$ to γ and γ to $\widetilde{g}_n\cdot\gamma$, respectively. One has $\widetilde{\Psi}_n, \widetilde{\Phi}_n \in W^{2,1}(\gamma)$. By a result from [9] one has $\nabla\widetilde{\Psi}_n \to \nabla\Psi$ in $L^2(\gamma, H)$. Equality (2.8) yields that $\nabla\widetilde{\Phi}_n \to \nabla\Phi$ in $L^2(g\cdot\gamma, H)$. It follows from (2.11) that the sequences $\{\nabla\widetilde{\Psi}_n\}$, $\{\nabla\widetilde{\Phi}_n\}$ are bounded in the Hilbert space $L^2(\gamma, H)$. Hence $\nabla\widetilde{\Psi}_n \to \nabla\Psi$ and $\nabla\widetilde{\Phi}_n \to \nabla\Phi$ weakly in $L^2(\gamma, H)$. In particular, this implies that $\mathcal{L}\widetilde{\Psi}_n \to \mathcal{L}\Psi$ and $\mathcal{L}\widetilde{\Phi}_n \to \mathcal{L}\Phi$ in the sense of distributions on (X, γ).

(iii) The approximations g_n enjoy even better properties. It is well-known that if $g \in L^p(\gamma)$ for some $p > 1$, then $g_n = \mathbb{E}(P_{\frac{1}{n}} g|\mathcal{F}_n) \in \mathcal{F}C_b^\infty$. If we only have $\int_X g\log g \, d\gamma < \infty$, then $g_n = P_{\frac{1}{n}}\mathbb{E}(g|\mathcal{F}_n)$ is twice continuously differentiable. This is verified by Young's inequality using that fact that $\mathbb{E}(g|\mathcal{F}_n)\log\mathbb{E}(g|\mathcal{F}_n) \in L^1(\gamma)$. By the contracting property of the Ornstein–Uhlenbeck semigroup and convergence $P_{\frac{1}{n}} g \to g$ in $L^1(\gamma)$ we obtain that $g_n \to g$ in $L^1(\gamma)$. Passing to a subsequence one can assume without loss of generality that $g_n \to g$ γ-a.e. It can be proved by the same arguments as in (i) that convergence of the corresponding entropies holds also in this case. Apart from the well-known smoothing properties, the Ornstein–Uhlenbeck semigroup possesses other nice properties related to the optimal transport. It has been noted in [9] (however, without proof) that $\nabla\Phi_n \to \nabla\Phi$ in $L^2(\gamma, H)$, hence a subsequence converges γ-a.e. Convergence $\nabla\Psi_n \to \nabla\Psi$ in $L^2(g\cdot\gamma, H)$ follows easily from identity (2.8). Let us briefly discuss the case of Φ, which is needed for our purposes. In order to avoid a repetition of lengthy arguments in the proof of Theorem 4.1 in [9] we only comment on the steps where some difference between the two cases appears. Set $G_t(x,y) := g\bigl(e^{-t}x + \sqrt{1-e^{-2t}}\,y\bigr)$ and consider the optimal transport $\mathcal{T}\colon X \times X \to X \times X$ sending $\gamma\otimes\gamma$ to $G_t\cdot(\gamma\otimes\gamma)$. Since

$$W_2^2\bigl(\gamma\otimes\gamma, (g\cdot\gamma)\otimes\gamma\bigr) = W_2^2(\gamma, g\cdot\gamma) = \int_X |\nabla\Phi|_H^2\,d\gamma$$

and $G_t\cdot(\gamma\otimes\gamma)$ is the image of $(g\cdot\gamma)\otimes\gamma$ under a measure-preserving H-orthogonal linear operator, we have

$$W_2^2\bigl(\gamma\otimes\gamma, G_t\cdot(\gamma\otimes\gamma)\bigr) = W_2^2(\gamma, g\cdot\gamma).$$

The projection of $G_t\cdot(\gamma\otimes\gamma)$ onto the first factor is $P_tg\cdot\gamma$. Let $\mu\in\mathcal{P}(X^2\times X^2)$ be the solution of the Monge–Kantorovich problem for the couple of measures $\bigl(\gamma\otimes\gamma, G_t\cdot(\gamma\otimes\gamma)\bigr)$. The projection of μ on the first factor X^2 is a probability measure with the marginals γ and $P_tg\cdot\gamma$. By virtue of optimality one has

$$W_2^2(\gamma, P_tg\cdot\gamma) \le W_2^2\bigl(\gamma\otimes\gamma, G_t\cdot(\gamma\otimes\gamma)\bigr) = W_2^2(\gamma, g\cdot\gamma).$$

Since $P_{\frac{1}{n}}g \to g$ in $L^1(\gamma)$, by the semicontinuity of the function $x \to |x|_H$ we obtain that

$$\underline{\lim}_n W_2^2(\gamma, P_{\frac{1}{n}}g\cdot\gamma) \ge W_2^2(\gamma, g\cdot\gamma),$$

hence

$$\lim_{n\to\infty} W_2^2(\gamma, P_{\frac{1}{n}}g\cdot\gamma) = W_2^2(\gamma, g\cdot\gamma).$$

Then, following the proof of Theorem 4.1 in [9], one can show that $\nabla\Phi_n \to \nabla\Phi$ in $L^2(\gamma, H)$. In what follows we may assume without loss of generality that $\nabla\Phi_n \to \nabla\Phi$ and $\nabla\Psi_n \to \nabla\Psi$ γ-a.e. Obviously, $\mathcal{L}\Psi_n \to \mathcal{L}\Psi$ and $\mathcal{L}\Phi_n \to \mathcal{L}\Phi$ in the sense of distributions. These remarks will be employed below.

Theorem 2.2. *Assume that* $\log g \in L^1(\gamma)$ *and* $\mathrm{Ent}_\gamma g < \infty$. *Let* $g_n = P_{1/n}\mathbb{E}(g|\mathcal{F}_n)$. *Then there exists measurable mappings* K *and* L *with values in the space of symmetric Hilbert–Schmidt operators such that, for some subsequence* $\{n_k\}$, *the mappings* $DS_{n_k} - I = D^2\Psi_{n_k}$ *and* $(DS_{n_k})^{-1} - I$ *converge* γ*-a.e. to* $(I+K)^2 - I$ *and* $(I+L)^2 - I$ *in the Hilbert–Schmidt norm. Moreover,* $(I+K)(I+L) = (I+L)(I+K) = I$ *and the following inequalities hold:*

$$\int_X \log\frac{1}{g}\,d\gamma \ge \frac{1}{2}\int_X |\nabla\Psi|_H^2\,d\gamma - \int_X \log\det{}_2\bigl[(I+K)^2\bigr]\,d\gamma, \tag{2.12}$$

$$\int_X g\log g\,d\gamma \ge \frac{1}{2}\int_X |\nabla\Phi|_H^2\,d\gamma - \int_X \log\det{}_2\bigl[(I+L)^2\bigr]g\,d\gamma. \tag{2.13}$$

In particular,

$$\int_X \Bigl(\|K\|_{\mathcal{H}}^2 + \|L\|_{\mathcal{H}}^2\Bigr)\min(1,g)\,d\gamma < \infty. \tag{2.14}$$

Proof. Passing to a subsequence we may assume that $g_n \to g$ γ-a.e. The hypothesis $\log g \in L^1(\gamma)$ implies that $g > 0$ γ-a.e. Let $\underline{g} := \inf_n g_n$. As $g_n = P_{\frac{1}{n}} \mathbb{E}(g|\mathcal{F}_n) > 0$ γ-a.e., we obtain that $\underline{g} > 0$ γ-a.e. It has been proved in [11, Theorem 6.1] under the stronger assumption $g > c > 0$ that, passing to a subsequence which will be denoted by the same indices, one has $K_n \to K$ and $L_n \to L$ γ-a.e. in the uniform operator norm. In addition, $(I+K)(I+L) = (I+L)(I+K) = I$ γ-a.e. This result can be easily generalized to the present setting. It suffices to follow the proof in [11] and replace the measure $c \cdot \gamma$ by $\min(1, \underline{g}) \cdot \gamma$ in all the estimates. In particular, all the "almost surely" statements remain valid.

Let us show almost sure convergence in the Hilbert–Schmidt norm. It has been proved in [11] that

$$\mathrm{Tr}\Big((I+L)(I+K_n)^2(I+L) + (I+K)(I+L_n)^2(I+K) - 2I\Big) \to 0 \quad \gamma\text{-a.e.}$$

Let us write $(I+L)(I+K_n)^2(I+L) = I + Z_n$. Since

$$(I+K)(I+L_n)^2(I+K) = \Big[(I+L)(I+K_n)^2(I+L)\Big]^{-1},$$

we obtain

$$\mathrm{Tr}(I + Z_n + (I+Z_n)^{-1} - 2I) = \mathrm{Tr}\big[Z_n^2 (I+Z_n)^{-1}\big] \to 0 \quad \gamma\text{-a.e.}$$

Since Z_n tends to 0 in the operator norm, we obtain $\mathrm{Tr} Z_n^2 \to 0$, hence

$$(I+L)(I+K_n)^2(I+L) - I \to 0 \quad \gamma\text{-a.e.}$$

in the Hilbert–Schmidt norm. Taking into account that $I+K$ is bounded, we obtain

$$\big\|(I+K)\big((I+L)(I+K_n)^2(I+L) - I\big)(I+K)\big\|_{\mathcal{H}} = \|(I+K_n)^2 - (I+K)^2\|_{\mathcal{H}} \to 0 \quad \gamma\text{-a.e.}$$

Therefore,

$$\mathrm{Tr}\big(2(K_n - K) + K_n^2 - K^2\big)^2 = \mathrm{Tr}\big((K_n - K)(2I + K_n + K)\big)^2 \to 0 \quad \gamma\text{-a.e.}$$

Since $I+K$ is invertible and the operators K_n converge to K in the operator norm, we have $\mathrm{Tr}(K_n - K)^2 \to 0$. Hence $K_n \to K$. Clearly, $K_n^2 \to K^2$, hence

$$(I+K_n)^2 - I \to (I+K)^2 - I.$$

The case of $(I+L_n)^2 - I$ is handled in the same way. In particular, we obtain that

$$\det{}_2\big[(I+K_n)^2\big] \to \det{}_2\big[(I+K)^2\big], \quad \det{}_2\big[(I+L_n)^2\big] \to \det{}_2\big[(I+L)^2\big] \quad \gamma\text{-a.e.}$$

According to Remark 2.1 one has

$$\int_X g_n \log g_n \, d\gamma \to \int_X g \log g \, d\gamma, \quad \int_X \log g_n \, d\gamma \to \int_X \log g \, d\gamma$$

and $\nabla\Psi_n \to \nabla\Psi$ γ-a.e. Hence by the relations

$$(I + \nabla\Phi) \circ (I + \nabla\Psi) = I, \quad (g \cdot \gamma) \circ (I + \nabla\Psi)^{-1} = \gamma$$

and Fatou's theorem we obtain inequalities (2.12) and (2.13) from inequalities (2.9) and (2.10). Inequality (2.14) follows by (2.12), (2.13), and the estimate

$$-\log \det{}_2\big[(I+K)^2\big] - \log \det{}_2\big[(I+L)^2\big] \geq \|K\|_{\mathcal{H}}^2 + \|L\|_{\mathcal{H}}^2.$$

In order to prove this estimate we observe that its left-hand side equals

$$-\mathrm{Tr}[(K^2 + 2K)(L^2 + 2L)]$$

by the general formula

$$\det_2(I+A)\det_2(I+B) = \det_2[(I+A)(I+B)]\exp \mathrm{Tr}(AB)$$

and the identity $(I+K)(I+L)=I$. This identity yields $KL+K+L=0$. By using that K and L commute we find

$$\begin{aligned}-(K^2+2K)(L^2+2L) &= -KL(4I+2K+2L+KL) = (K+L)(4I+K+L)\\ &= K^2+L^2+2KL+4(K+L) = K^2+L^2-2KL\\ &= K^2+L^2+2K^2(I+K)^{-1} \geq K^2+L^2.\end{aligned}$$

The proof is complete. □

Remark 2.3. Note that $DT_n = (DS_n)^{-1}(T_n)$. Since by Jensen's inequality

$$\int_X g_n \log g_n \, d\gamma \leq \int_X g \log g \, d\gamma,$$

the Radon–Nikodym densities of $\gamma \circ T_n^{-1}$ with respect to γ form a γ-uniformly integrable sequence $\{g_n\}$. Hence one has $\|DT_n - (I+L)^2(T)\|_{\mathcal{H}} \to 0$ in measure. Passing to a subsequence we may assume that $\|DT_n - (I+L)^2(T)\|_{\mathcal{H}} \to 0$ γ-a.e.

It will be shown in Lemma 2.9 that $(I+K(x))^2 - I$ and $\big[I+L(T(x))\big]^2 - I$ coincide a.e. with $D^2_{\mathrm{ac}}\Psi(x)$ and $D^2_{\mathrm{ac}}\Phi(x)$, respectively.

Corollary 2.4. *Let $\log g \in L^1(\gamma)$ and $g\log g \in L^1(\gamma)$. Then, for any $h,k \in H$, there exist bounded Radon measures Ψ_{hk} and Φ_{hk} and one has*

$$\Psi_{hk} = \frac{1}{2}\big[\Psi_{(h+k)(h+k)} - \Psi_{hh} - \Psi_{kk}\big], \quad \Phi_{hk} = \frac{1}{2}\big[\Phi_{(h+k)(h+k)} - \Phi_{hh} - \Phi_{kk}\big].$$

In addition, the measures $\Phi^{\mathrm{sing}}_{hh}$ and $\Psi^{\mathrm{sing}}_{hh}$ are nonnegative and one has

$$\Phi^{\mathrm{ac}}_{hh} \geq l_h(T) \quad \gamma\text{-a.e.}, \quad \Psi^{\mathrm{ac}}_{hh} \geq k_h \quad \gamma\text{-a.e.},$$

where

$$k_h := \big\langle ((I+K)^2 - I)h, h\big\rangle_H, \quad l_h := \big\langle ((I+L)^2 - I)h, h\big\rangle_H.$$

Proof. We recall that $\Phi, \Psi \in W^{2,1}(\gamma)$. By a result from [4] the measure $\big[\Phi+\widehat{h}^2\big]\cdot\gamma := F^h\cdot\gamma$ is twice Skorohod differentiable and the following inequality for its variation norm holds:

$$\|d^2_{hh}(F^h\cdot\gamma)\| \leq 2\|F^h\|_{L^1(d^2_{hh}\gamma)} + 2\|\partial_h F^h\|_{L^1(d_h\gamma)}.$$

It follows easily by the Cauchy inequality that $\|d^2_{hh}(F^h\cdot\gamma)\| < \infty$, hence $\|d^2_{hh}(\Phi\cdot\gamma)\| < \infty$. Similarly, $\|d^2_{hh}(\Psi\cdot\gamma)\| < \infty$. For any smooth cylindrical function η one has

$$\int_X \eta(x)\,\Psi_{(h+k)(h+k)}(dx) = -\int_X (\partial_h\eta + \partial_k\eta)\partial_{h+k}\Psi\,d\gamma + \int_X (\widehat{h}+\widehat{k})\eta\partial_{h+k}\Psi\,d\gamma. \qquad (2.15)$$

One has $\partial_{h+k}\Psi = \partial_h\Psi + \partial_k\Psi$. By using (2.15) and the identities

$$\int_X \eta(x)\,\Psi_{hh}(dx) = -\int_X \partial_h\eta\partial_h\Psi\,d\gamma + \int_X \widehat{h}\eta\partial_h\Psi\,d\gamma,$$

$$\int_X \eta(x)\,\Psi_{kk}(dx) = -\int_X \partial_k\eta\partial_k\Psi\,d\gamma + \int_X \widehat{k}\eta\partial_k\Psi\,d\gamma,$$

we find

$$\int_X \eta(x)\,\Big[\Psi_{(h+k)(h+k)} - \Psi_{hh} - \Psi_{kk}\Big](dx)$$
$$= -\int_X \big(\partial_h\eta\partial_k\Psi + \partial_k\eta\partial_h\Psi\big)\,d\gamma + \int_X \eta\big(\widehat{k}\partial_h\Psi + \widehat{h}\partial_k\Psi\big)\,d\gamma.$$

Note that

$$-\int_X \partial_h\eta\partial_k\Psi\,d\gamma + \int_X \widehat{h}\eta\partial_k\Psi\,d\gamma = -\int_X \partial_k\eta\partial_h\Psi\,d\gamma + \int_X \widehat{k}\eta\partial_h\Psi\,d\gamma$$
$$= \int_X \Psi\big(\partial_h\partial_k\eta - \partial_h\eta\widehat{k} - \partial_k\eta\widehat{h} + \eta\widehat{h}\widehat{k} - \eta\langle h,k\rangle_H\big)\,d\gamma.$$

Hence

$$\frac{1}{2}\int_X \eta(x)\,\Big[\Psi_{(h+k)(h+k)} - \Psi_{hh} - \Psi_{kk}\Big](dx) = -\int_X \partial_h\eta\partial_k\Psi\,d\gamma + \int_X \widehat{h}\eta\partial_k\Psi\,d\gamma.$$

Therefore, $\frac{1}{2}\Big[\Psi_{(h+k)(h+k)} - \Psi_{hh} - \Psi_{kk}\Big] = \Psi_{hk}$. The case of Φ is analogous.

Let us show that the measures $\Psi_{hh}^{\rm sing}$ and $\Phi_{hh}^{\rm sing}$ are nonnegative. We may assume that $\widehat{h}(h) = 1$. The function

$$t \mapsto \Psi(x+th) + \widehat{h}(x+th)^2/2$$

is convex, hence its derivative $t \mapsto \partial_h\Psi(x+th) + \widehat{h}(x+th)$ is increasing. Suppose that B is a Borel set such that $\gamma(B) = 0$ and $\Psi_{hh}^{\rm sing}(B) < 0$. One can find a sequence of smooth cylindrical functions f_j such that $0 \le f_j \le 1$, $f_j \to I_B$ a.e. with respect to the measure $\gamma + |\Psi_{hh}|$. Hence $f_j \to 0$ γ-a.e. By the Lebesgue dominated convergence theorem we obtain

$$\Psi_{hh}^{\rm sing}(B) = -\lim_{j\to\infty}\int_X \partial_h f_j(x)\partial_h\Psi(x)\,\gamma(dx).$$

We show that the right-hand side is nonnegative. To this end, we note that

$$\lim_{j\to\infty}\int_X \partial_h f_j(x)\partial_h\Psi(x)\,\gamma(dx)$$
$$= \lim_{j\to\infty}\int_X \Big(\partial_h f_j(x)[\partial_h\Psi(x) + \widehat{h}] + f_j(x)\widehat{h}[\partial_h\Psi(x) + \widehat{h}]\Big)\,\gamma(dx),$$

because

$$\int_X \partial_h f_j(x)\widehat{h}\,\gamma(dx) = -\int_X f_j(x)(1 - \widehat{h}^2)\,\gamma(dx)$$

tends to zero as $j \to \infty$ and the same is true for the integral of $f_j(x)\widehat{h}[\partial_h\Psi(x) + \widehat{h}]$, which is clear from the integrability of $\widehat{h}^2 + |\widehat{h}\partial_h\Psi(x)|$ and the Lebesgue dominated convergence theorem. Finally, we observe that

$$\int_X \Big(\partial_h f_j(x)[\partial_h\Psi(x) + \widehat{h}] + f_j(x)\widehat{h}[\partial_h\Psi(x) + \widehat{h}]\Big)\,\gamma(dx) \le 0.$$

Indeed, one can approximate Ψ in $W^{2,1}(\gamma)$ by functions $\Psi_k \in W^{2,2}(\gamma)$ with the property that the functions $t \mapsto \partial_h\Phi_k(x+th) + \widehat{h}(x+th)$ are increasing. Then, by the integration by parts formula, the integral on the left for Ψ_k in place of Ψ equals the integral of $-f_j\partial_h[\partial_h\Psi_k(x) + \widehat{h}]$, which is nonpositive, because $f_j \ge 0$, $\partial_h[\partial_h\Psi_k(x) + \widehat{h}] \ge 0$.

Let us fix a nonnegative function $\zeta \in \mathcal{F}C_b^\infty$. By Remark 2.1, Fatou's theorem and Theorem 2.2 we obtain

$$\int_X k_h \zeta \, d\gamma \le \lim_{n\to\infty} \int_X (\Psi_n)_{hh} \zeta \, d\gamma = -\lim_{n\to\infty} \int_X \bigl(\partial_h \Psi_n \partial_h \zeta - \widehat{h} \partial_h \Psi_n \zeta\bigr) \, d\gamma$$
$$= -\int_X \bigl(\partial_h \Psi \partial_h \zeta - \widehat{h} \partial_h \Psi \zeta\bigr) \, d\gamma = \int_X \Psi^{\rm ac}_{hh} \zeta \, d\gamma + \int_X \zeta(x) \, \Psi^{\rm sing}_{hh}(dx).$$

By the singularity of $\Psi^{\rm sing}_{hh}$ we obtain $k_h \le \Psi^{\rm ac}_{hh}$ a.e. The case of Φ is similar. □

Corollary 2.5. (i) *Suppose that $g \ge c > 0$ for some constant c and $g \log g \in L^1(\gamma)$. Then there exists an $\mathcal{H}$-valued measure $D^2\Psi$ of bounded variation.*

(ii) *Suppose that $0 < g \le C$ for some constant C and $\log g \in L^1(\gamma)$. Then there exists an $\mathcal{H}$-valued measure $D^2\Phi$ of bounded variation.*

Proof. (i) Let us show that the finite dimensional measures $D^2\Psi_n$ have uniformly bounded variations regarded as $\mathcal{H}$-valued measures. Since these measures are given by $\mathcal{H}$-valued densities $(I + K_n)^2 - I$ with respect to γ, it suffices to have a uniform bound of the integrals of $\|2K_n + K_n^2\|_{\mathcal{H}}$ with respect to γ. It is clear from the proof of Theorem 2.2 that the integrals of $\|K_n\|^2_{\mathcal{H}}$ against the measure $\min(1, g) \cdot \gamma$, hence against γ, are uniformly bounded. It remains to observe that $\|K_n^2\|_{\mathcal{H}} \le \|K_n\|^2_{\mathcal{H}}$.

Now let us show that there exists an $\mathcal{H}$-valued measure $D^2\Psi$ of bounded variation whose matrix elements are $\Psi_{e_ie_j}$. For every $h \in H$, the derivative of $\partial_h\Psi$ along h in the sense of distributions over Wiener space is nonnegative, hence is represented by a nonnegative Radon measure ν_h (see [17]). This measure is the limit of the sequence of functions $\partial_h^2\Psi_n = \langle D^2\Psi_n h, h\rangle_H$ in the sense of distributions. Let us define the operator-valued measure $D^2\Psi$ by the equality

$$\langle D^2\Psi h, k\rangle_H := \frac{1}{2}(\nu_{h+k} - \nu_h - \nu_k).$$

The value of the right-hand side on every fixed Borel set is a symmetric bilinear form. This is clear from the fact that the integral of any test function θ against $(\nu_{h+k} - \nu_h - \nu_k)/2$ coincides with the limit of the integrals of the functions $\theta[\partial^2_{h+k}\Psi_n - \partial^2_h\Psi_n - \partial^2_k\Psi_n]$ against γ. The uniform estimate of variations with respect to the $\mathcal{H}$-norm yields that this bilinear form is generated by a symmetric Hilbert–Schmidt operator and that the obtained $\mathcal{H}$-valued measure is of bounded variation. Assertion (ii) is analogous. We only note that the integral of $\|K_n(T_n(x))\|^2_{\mathcal{H}} + \|L_n(T_n(x))\|^2_{\mathcal{H}}$ against γ (which appears when we consider the second derivative of Φ) equals the integral of $\|K_n(x)\|^2_{\mathcal{H}} + \|L_n(x)\|^2_{\mathcal{H}}$ against $g_n \cdot \gamma$, hence is estimated by a constant. □

The following lemma is a generalization of [9, Lemma 7.2].

Lemma 2.6. *Let $\log g \in L^2(\gamma)$. Then $\mathcal{L}\Psi$ is a bounded Radon measure and γ-a.e. one has $\mathcal{L}_{\rm ac}\mathbb{E}(\Psi|\mathcal{F}_n) \to \mathcal{L}_{\rm ac}\Psi$. If $|\log g|^2 g \in L^1(\gamma)$, then $\mathcal{L}\Phi$ is a bounded Radon measure and γ-a.e. one has $\mathcal{L}_{\rm ac}\mathbb{E}(\Phi|\mathcal{F}_n) \to \mathcal{L}_{\rm ac}\Phi$.*

Proof. Let us approximate g by the functions $\widetilde{g}_n := \mathbb{E}(g|\mathcal{F}_n)$ and denote by $\widetilde{\Psi}_n$ the corresponding potentials. By the finite dimensional change of variables formula one has

$$\log \frac{1}{\widetilde{g}_n} = -\mathcal{L}_{\rm ac}\widetilde{\Psi}_n + \frac{1}{2}|\nabla\widetilde{\Psi}_n|^2_H - \log \det{}_2(I + D^2_{\rm ac}\widetilde{\Psi}_n).$$

It is known (see Remark 2.1) that $\mathcal{L}\widetilde{\Psi}_n \to \mathcal{L}\Psi$ in the sense of distributions. Let us show that $\log \widetilde{g}_n \to \log g$ in the sense of distributions. Indeed, according to Jensen's inequality for every fixed bounded nonnegative $\mathcal{F}_N$-measurable function η and $n > N$ one has

$$-\int_X \eta\log\widetilde{g}_n\,d\gamma \le -\int_X \eta\mathbb{E}(\log g|\mathcal{F}_n)\,d\gamma = -\int_X \eta\log g\,d\gamma.$$

As $\widetilde{g}_n \to g$ a.e. and the function $x\log x$ is bounded from below, we obtain by Fatou's theorem

$$\begin{aligned}\underline{\lim}_n \int_X \eta\log\frac{1}{\widetilde{g}_n}\,d\gamma &= \underline{\lim}_n \int_X \eta\Bigl(\log\frac{1}{\widetilde{g}_n}\Bigr)\frac{1}{\widetilde{g}_n}\widetilde{g}_n\,d\gamma = \underline{\lim}_n \int_X \eta\Bigl(\log\frac{1}{\widetilde{g}_n}\Bigr)\frac{1}{\widetilde{g}_n}g\,d\gamma\\ &\ge \int_X \eta\Bigl(\log\frac{1}{g}\Bigr)\frac{1}{g}g\,d\gamma = \int_X \eta\log\frac{1}{g}\,d\gamma.\end{aligned}$$

Hence $\lim\limits_{n\to\infty}\displaystyle\int_X \eta\log\widetilde{g}_n\,d\gamma = \int_X \eta\log g\,d\gamma$. Taking into account that

$$\log\frac{1}{\widetilde{g}_n} + \mathcal{L}_{\rm ac}\widetilde{\Psi}_n \ge 0,$$

we obtain that $\log\frac{1}{g} + \mathcal{L}\Psi \ge 0$ in the sense of distributions. We observe that $\log g$ defines an element of the dual to the Sobolev space $W^{2,1}(\gamma)$ and $\mathcal{L}\Psi$ belongs to the dual to $W^{2,2}(\gamma)$. Hence $\log\frac{1}{g} + \mathcal{L}\Psi$ is a bounded Radon measure (see [17]). It was shown in [9] that $\{\mathcal{L}_{\rm ac}\mathbb{E}(\Phi|\mathcal{F}_n)\}$ is a submartingale convergent γ-a.e. to $\mathcal{L}_{\rm ac}\Phi$. The analogous assertion for Ψ is proved along the same lines. The proof of the remaining assertions in the case of Φ is similar. We only note that it follows from our hypotheses that the integrals of $|\log g_n(T_n)|^2$ against γ are uniformly bounded. As $g_n(T_n) \to g(T)$ in measure, we obtain that $\log g_n(T_n) \to \log g(T)$ in $L^1(\gamma)$. The rest of the proof is the same as in the case of Ψ. □

Lemma 2.7. *Let $A = (a_{i,j})$ be a symmetric $(n+1)\times(n+1)$ matrix such that $I + A > 0$ and let $B = A_{n\times n}$. Then $-\log\det_2(I + A) \ge -\log\det_2(I + B)$.*

Proof. Let us take a new orthonormal basis $v_1,\dots,v_n$ in $\mathbb{R}^n$ such that B is diagonal in this basis and has the eigenvalues $b_1,\dots,b_n$ and consider the matrix A in the basis $v_1,\dots,v_n,e_{n+1}$. Then ${\rm Tr}A = {\rm Tr}B + a_{n+1,n+1}$. One can easily show that

$$\det(I + A) = \det(I + B)\Bigl[1 + a_{n+1,n+1} - \sum_{i=1}^n \frac{a_{i,n+1}^2}{1 + b_i}\Bigr].$$

Hence

$$\begin{aligned}&{\rm Tr}A - \log\det(I + A)\\ &\quad= {\rm Tr}B + a_{n+1,n+1} - \log\det(I + B) - \log\Bigl(1 + a_{n+1,n+1} - \sum_{i=1}^n \frac{a_{i,n+1}^2}{1 + b_i}\Bigr)\\ &\quad\ge {\rm Tr}B - \log\det(I + B).\end{aligned}$$

The proof is complete. □

We need also the following technical lemma. Let λ denote Lebesgue measure.

Lemma 2.8. *Let $F\colon \mathbb{R}^d \to \mathbb{R}^d$ be a locally integrable mapping such that its derivative DF in the sense of generalized functions is a locally bounded measure with values in the space of nonnegative symmetric matrices. Let $D_{\rm ac}F$ be the operator-valued density of the absolutely continuous component of DF and let $\Omega := \{x\colon\ \det D_{\rm ac}F(x) > 0\}$. Then the measure $\lambda|_\Omega \circ F^{-1}$ is absolutely continuous.*

Proof. It suffices to show that there is a sequence of measurable sets $\Omega_k \subset \Omega$ such that $\Omega\backslash\bigcup_{k=1}^{\infty}\Omega_k$ has measure zero and each measure $\lambda|_{\Omega_k}\circ F^{-1}$ has a density. Therefore, denoting by $m(A)$ the minimal eigenvalue of a matrix A, it suffices to prove our claim for the restrictions of F to the sets $\Omega_\alpha := \bigl\{x\in\Omega\colon\ m\bigl(D_{\rm ac}F(x)\bigr)\ge\alpha\bigr\}$, $\alpha>0$. Moreover, it suffices to consider bounded subsets of Ω_α. We fix numbers $\alpha>0$ and $\delta>0$, a ball B, and a probability density $\theta\in C_0^\infty(\mathbb{R}^n)$. Let $\theta(x) := k^d\theta(x/k)$ and $F_k = F*\theta_k$. Then $F_k(x)\to F(x)$ and $D_{\rm ac}F*\theta_k(x)\to D_{\rm ac}F(x)$ a.e., since $D_{\rm ac}F$ is a locally integrable operator-valued mapping (as a density of the absolutely continuous part of a locally bounded operator-valued measure). By Egoroff's theorem, there exists a measurable set $E_\delta\subset\Omega_\alpha\cap B$ such that $\lambda((\Omega_\alpha\cap B)\backslash E_\delta)<\delta$ and the sequence $D_{\rm ac}F*\theta_k(x)$ converges uniformly on E_δ. Hence we may assume that $m\bigl((D_{\rm ac}F*\theta_k)(x)\bigr)\ge\alpha/2$ for all k and all $x\in E_\delta$. We observe that

$$
\begin{aligned}
D(F*\theta_k)(x) = DF*\theta_k(x) &= \int_{\mathbb{R}^n}\theta_k(x-y)\,DF(dy)\\
&\ge \int_{\mathbb{R}^n}\theta_k(x-y)D_{\rm ac}F(y)\,dy = D_{\rm ac}F*\theta_k(x)
\end{aligned}
$$

in the sense of quadratic forms, since the singular component of DF also takes values in the space of nonnegative symmetric operators. Therefore,

$$
m\bigl(D(F*\theta_k)(x)\bigr) = m\bigl(DF*\theta_k(x)\bigr)\ge m\bigl(D_{\rm ac}F*\theta_k(x)\bigr)\ge\alpha/2,\quad x\in E_\delta.
$$

It follows that $\det[D(F*\theta_k)(x)]\ge(\alpha/2)^d$ for all $x\in E_\delta$, which yields that the measure $\mu_k := \lambda|_{E_\delta}\circ(F*\theta_k)^{-1}$ admits a density $\varrho_k\le(2/\alpha)^d$. Since the measures μ_k converge weakly to the measure $\lambda|_{E_\delta}\circ F^{-1}$, we conclude that the latter has a density too. Letting $\delta\to 0$, we arrive at the desired conclusion. □

In the proof of the next lemma we employ two important results from measure theory (see [3]). Let μ be a finite nonnegative measure on a measurable space $(X,\mathcal{A})$ and let $\{f_n\}\subset L^1(\mu)$ be a norm bounded sequence. Then, according to the Komlós theorem, there exist a subsequence $\{h_n\}\subset\{f_n\}$ and a function $f\in L^1(\mu)$ such that the sequence of averages $n^{-1}\sum_{i=1}^n h_n$ converges to f μ-a.e. In addition, by the Gaposhkin theorem, such a subsequence can be found with the property that, for every $\varepsilon>0$, there exists a subset $X_\varepsilon\subset X$ such that $\mu(X\setminus X_\varepsilon)<\varepsilon$ and $h_n\to f$ weakly in $L^1(\mu|_{X_\varepsilon})$.

Lemma 2.9. *Suppose that* $\log g\in L^1(\gamma)$, $g\log g\in L^1(\gamma)$. *Then there exist* $\mathcal{H}$*-valued mappings* $D^2_{\rm ac}\Psi$ *and* $D^2_{\rm ac}\Phi$ *with matrix elements* $\Phi^{\rm ac}_{e_ie_j}$ *and* $\Psi^{\rm ac}_{e_ie_j}$ *such that*

$$
I + D^2_{\rm ac}\Psi = (I+K)^2,\ I + D^2_{\rm ac}\Phi = (I+L(T))^2\quad \gamma\text{-a.e.}
$$

In addition, there exist finite limits in (1.3) *and formulas* (1.4) *and* (1.5) *hold.*

Proof. Let us consider the approximations $g_n = P_{1/n}\mathbb{E}(g|\mathcal{F}_n)\to g$ and let Ψ_n be the corresponding potentials such that $\nabla\Psi_n\to\nabla\Psi$ weakly in $L^2(\gamma;H)$ according to Remark 2.1. By the finite dimensional change of variables formula one has

$$
\log\frac{1}{g_n} = -\mathcal{L}\Psi_n + \frac{1}{2}|\nabla\Psi_n|_H^2 - \log\det{}_2(I+D^2\Psi_n).
$$

By Theorem 2.2, passing to a subsequence, we have $\|2K+K^2-D^2\Psi_n\|_{\mathcal{H}}\to 0$ γ-a.e., hence we have $\log\det_2(I+D^2\Psi_n)\to\log\det_2\bigl[(I+K)^2\bigr]$ γ-a.e. Moreover, by Remark 2.1, we have $g_n\to g$ and $|\nabla\Psi_n|_H^2\to|\nabla\Psi|_H^2$ γ-a.e. Hence for γ-almost all x, there exists a finite limit

$$
\widetilde{\mathcal{L}}\Psi(x) := \lim_{n\to\infty}\mathcal{L}\Psi_n(x)
$$

and the following formula holds:

$$g = \det_2\bigl[(I+K)^2\bigr]\exp\Bigl(\widetilde{\mathcal{L}}\Psi - \frac{1}{2}|\nabla\Psi|_H^2\Bigr). \tag{2.16}$$

Analogously, taking into account that $\{\frac{\gamma\circ T_n^{-1}}{\gamma}\} = \{g_n\}$ is a γ-uniformly integrable sequence and extracting a suitable subsequence we obtain

$$\frac{1}{g(T)} = \det_2\bigl[(I+L(T))^2\bigr]\exp\Bigl(\widetilde{\mathcal{L}}\Phi - \frac{1}{2}|\nabla\Phi|_H^2\Bigr), \tag{2.17}$$

where $\widetilde{\mathcal{L}}\Phi(x) = \lim\limits_{n\to\infty}\mathcal{L}\Phi_n(x)$ for γ-a.e. x.

We divide the subsequent proof into several steps.

Step 1. Let us show that

$$\widetilde{\mathcal{L}}\Psi + \widetilde{\mathcal{L}}\Phi(S) = \sum_{i=1}^{\infty}\bigl[k_i^2 + 2k_i + l_i^2 + 2l_i - x_i\partial_{e_i}\Psi - S_i\partial_{e_i}\Phi(S)\bigr]. \tag{2.18}$$

Indeed, taking into account that $(I+K)(I+L) = I$ we find

$$\det_2(I+K)^2\det_2(I+L)^2 = \exp\bigl[-\mathrm{Tr}(K^2+2K+L^2+2L)\bigr].$$

Now (2.17) yields

$$\frac{1}{g} = \frac{1}{g(T\circ S)} = \frac{\exp\Bigl(-\mathrm{Tr}(K^2+2K+L^2+2L)\Bigr)}{\det_2\bigl[(I+K)^2\bigr]}\exp\Bigl(\widetilde{\mathcal{L}}\Phi(S) - \frac{1}{2}|\nabla\Phi(S)|_H^2\Bigr).$$

Hence by (2.16) we have

$$\mathrm{Tr}(K^2+2K+L^2+2L) = \widetilde{\mathcal{L}}\Psi + \widetilde{\mathcal{L}}\Phi(S) - \frac{1}{2}|\nabla\Psi|_H^2 - \frac{1}{2}|\nabla\Phi(S)|_H^2.$$

Finally, we obtain

$$\widetilde{\mathcal{L}}\Psi + \widetilde{\mathcal{L}}\Phi(S) = \sum_{i=1}^{\infty}\Bigl[k_i^2 + l_i^2 + 2k_i + 2l_i + \frac{1}{2}(\partial_{e_i}\Psi)^2 + \frac{1}{2}(\partial_{e_i}\Phi(S))^2\Bigr].$$

Taking into account that $S_i(x) = x_i + \partial_{e_i}\Psi(x)$ and $\partial_{e_i}\Phi\bigl(S(x)\bigr) = x_i - S_i(x)$ by the equality $S(x) + \nabla\Phi\bigl(S(x)\bigr) = x$ we find

$$\frac{1}{2}(\partial_{e_i}\Psi)^2 + \frac{1}{2}(\partial_{e_i}\Phi(S))^2 = -x_i\partial_{e_i}\Psi - S_i\partial_{e_i}\Phi(S).$$

The proof of (2.18) is complete.

Step 2. Equality (2.18) yields

$$\widetilde{\mathcal{L}}\Psi + \widetilde{\mathcal{L}}\Phi(S) = \lim_{m\to\infty}\Bigl[\mathcal{L}_{K,m}\Psi + \mathcal{L}_{L,m}\Phi(S)\Bigr]\ \gamma\text{-a.e.}, \tag{2.19}$$

where

$$\mathcal{L}_{K,m}\Psi := \sum_{i=1}^{m}\bigl[k_i^2 + 2k_i - x_i\partial_{e_i}\Psi\bigr], \quad \mathcal{L}_{L,m}\Phi := \sum_{i=1}^{m}\bigl[l_i^2(T) + 2l_i(T) - x_i\partial_{e_i}\Phi\bigr].$$

Let us show that for some subsequence $\{n_k\}$ one has γ-a.e.

$$\widetilde{\mathcal{L}}\Psi = \lim_{m\to\infty}\frac{1}{m}\sum_{k=1}^{m}\mathcal{L}_{K,n_k}\Psi, \quad \widetilde{\mathcal{L}}\Phi = \lim_{m\to\infty}\frac{1}{m}\sum_{k=1}^{m}\mathcal{L}_{L,n_k}\Phi. \tag{2.20}$$

To this end we consider new approximations of g and $S = I + \nabla\Psi$. Set $Q_n = I + F_n$, where

$$F_n = \sum_{i=1}^n F_n^i e_i, \quad F_n^i := \langle F_n, e_i\rangle_H = \begin{cases} \partial_{e_i}\Psi, & i \le n \\ 0, & i > n. \end{cases}$$

Let

$$u_n = \det{}_2\big[I + (\Psi^{\rm ac}_{e_ie_j})_{n\times n}\big] \exp\Big(\sum_{i=1}^n (\Psi^{\rm ac}_{e_ie_i} - x_i\partial_{e_i}\Psi) - \frac{1}{2}\sum_{i=1}^n (\partial_{e_i}\Psi)^2\Big).$$

Let $\widetilde{x}_n$ be the image of $x = \sum_{i=1}^\infty \widehat{e}_i(x)e_i$ under the projection $x \mapsto x - P_nx$, i.e., $\widetilde{x}_n = \sum_{i=n+1}^\infty \widehat{e}_i(x)e_i$. The measure γ can be represented as a product measure $\gamma = \gamma_n \otimes \widetilde{\gamma}_n$, where $\gamma_n = \gamma \circ P_n^{-1}$. For any fixed $\widetilde{x}_n$ the mapping F_n can be considered as a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$. Moreover, it is the gradient of a 1-convex function. As by Corollary 2.4) we have

$$(I + K_{n\times n})^2 \le I + (\Psi^{\rm ac}_{e_ie_j})_{n\times n} \tag{2.21}$$

and $I + K(x)$ is invertible γ-almost everywhere, we obtain that $I + (D^2_{\rm ac}\Psi)_{n\times n}$ on $\mathbb{R}^n$ is almost surely invertible. Hence by Lemma 2.8, for $\widetilde{\gamma}_n$-almost every fixed $\widetilde{x}_n$, the image of the measure $u_n(\,\cdot\,, \widetilde{x}_n)\cdot\gamma_n$, where

$$u_n(z, \widetilde{x}_n) = u_n(z_1e_1 + \cdots + z_ne_n + \widetilde{x}_n)$$

and $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$, under the mapping

$$R_{\widetilde{x}_n}\colon\ \mathbb{R}^n \to \mathbb{R}^n, \quad z \mapsto Q_n(z_1e_1 + \cdots + z_ne_n + \widetilde{x}_n),$$

admits a density with respect to the n-dimensional Lebesgue measure. By Theorem 1.1 one has

$$\big[u_n(\,\cdot\,, \widetilde{x}_n)\cdot\gamma_n\big] \circ R^{-1}_{\widetilde{x}_n} = \gamma_n|_{R_{\widetilde{x}_n}(\mathbb{R}^n)}$$

for $\widetilde{\gamma}_n$-almost every fixed $\widetilde{x}_n$ and by Fubini's theorem $(u_n\cdot\gamma)\circ Q_n^{-1} = \gamma|_{Q_n(X)}$.

In the same way we define $R_n = I + U_n$, where

$$U_n = \sum_{i=1}^n U_n^i e_i, \quad U_n^i := \langle U_n, e_i\rangle_H = \begin{cases} \partial_{e_i}\Phi, & i \le n \\ 0, & i > n. \end{cases}$$

Exactly as above we prove that the measure $\gamma\circ R_n^{-1}$ is absolutely continuous with respect to γ and its density v_n satisfies the relation

$$\frac{1}{v_n\circ R_n} = \det{}_2\big[I + (\Phi^{\rm ac}_{e_ie_j})_{n\times n}\big] \exp\Big(\sum_{i=1}^n (\Phi^{\rm ac}_{e_ie_i} - x_i\partial_{e_i}\Phi) - \frac{1}{2}\sum_{i=1}^n (\partial_{e_i}\Phi)^2\Big).$$

We set $v_n(y) := 0$ if $y \notin R_n(X)$.

Let us apply the above mentioned Komlós and Gaposhkin theorems to the sequence $\{u_n\}$ and the measure γ (note that $\|u_n\|_{L^1(\gamma)} \le 1$ for every n). For the sake of simplicity we denote the new subsequence obtained from those theorems again by $\{u_n\}$. Let u_∞ be the corresponding limit. Repeating this procedure for the sequence of functions $1/v_n(R_n)$ we may assume that it also admits a limit f_∞ in the sense of the cited theorems. We set $v_\infty := 1/f_\infty(S)$, where $v_\infty(x) := \infty$ if $f_\infty(S)(x) = 0$. Hence $1/v_n(R_n) \to 1/v_\infty(T)$ weakly in $L_1(\gamma|_{X_\varepsilon})$ for every ε. It suffices to show that

$$\widetilde{\mathcal{L}}\Psi \ge \overline{\lim}_n \frac{1}{n}\sum_{m=1}^n \mathcal{L}_{K,m}\Psi, \quad \widetilde{\mathcal{L}}\Phi \ge \overline{\lim}_n \frac{1}{n}\sum_{m=1}^n \mathcal{L}_{L,m}\Phi \quad \gamma\text{-a.e.} \tag{2.22}$$

Indeed, if (2.22) holds, then (2.19) yields that γ-a.e. one has

$$\widetilde{\mathcal{L}}\Psi + \widetilde{\mathcal{L}}\Phi(S) = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n}\bigl(\mathcal{L}_{K,m}\Psi + \mathcal{L}_{L,m}\Phi(S)\bigr)$$
$$\le \overline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m}\Psi + \overline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m}\Phi(S),$$

and the strict inequality in (2.22) is impossible. Furthermore,

$$\widetilde{\mathcal{L}}\Psi + \widetilde{\mathcal{L}}\Phi(S) = \overline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m}\Psi + \underline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m}\Phi(S) = \widetilde{\mathcal{L}}\Psi + \underline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m}\Phi(S).$$

Hence we obtain

$$\widetilde{\mathcal{L}}\Phi = \underline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m} = \overline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m} = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m}.$$

Analogous relations hold for $\widetilde{\mathcal{L}}\Psi$.

Now let us prove (2.22). First we show that $u_\infty \le g$ γ-a.e. Fix a bounded nonnegative continuous function φ. One has

$$\int_X \varphi(S) g\, d\gamma = \int_X \varphi\, d\gamma \ge \int_{Q_n(X)} \varphi\, d\gamma = \int_X \varphi(Q_n) u_n\, d\gamma \ge \int_{X_\varepsilon} \varphi(Q_n) u_n\, d\gamma.$$

As $Q_n \to S$ γ-a.e., we have $\varphi(Q_n) \to \varphi(S)$ γ-a.e. By the Egoroff theorem one can choose a compact set $K_\varepsilon \subset X_\varepsilon$ such that $\gamma(X_\varepsilon \setminus K_\varepsilon) \le \varepsilon$ and $\varphi(Q_n) \to \varphi(S)$ uniformly on K_ε. By using that $u_n \to u_\infty$ weakly in $L_1(\gamma|_{X_\varepsilon})$, we obtain

$$\int_{X_\varepsilon} \varphi(Q_n) u_n\, d\gamma \ge \int_{K_\varepsilon} \varphi(Q_n) u_n\, d\gamma \to \int_{K_\varepsilon} \varphi(S) u_\infty\, d\gamma.$$

Hence

$$\int_X \varphi(S) g\, d\gamma \ge \int_{K_\varepsilon} \varphi(S) u_\infty\, d\gamma.$$

Then

$$\int_X \varphi(S) g\, d\gamma \ge \int_X \varphi(S) u_\infty\, d\gamma,$$

since we have $\gamma(\bigcup_\varepsilon K_\varepsilon) = 1$. As S has an inverse mapping T, for every measurable set B one can find a uniformly bounded sequence of nonnegative smooth cylindrical functions η_j such that $\eta_j \to I_B \circ T$ a.e., which gives $\eta_j \circ S \to I_B$ a.e., whence

$$\int_B u_\infty\, d\gamma \le \int_B g\, d\gamma.$$

This implies the desired estimate $u_\infty \le g$ γ-a.e. In the same way the relations

$$\int_X \frac{\varphi(T)}{g(T)}\, d\gamma = \int_X \varphi\, d\gamma \ge \int_{R_n(X)} \varphi \frac{v_n}{v_n}\, d\gamma \ge \int_X \frac{\varphi(R_n)}{v_n(R_n)}\, d\gamma$$

yield that $1/v_\infty(T) \le 1/g(T)$ γ-a.e, hence $v_\infty \ge g$ γ-a.e.

Now suppose that (2.22) does not hold. Assume that on a positive measure set M one has $\widetilde{\mathcal{L}}\Psi + \delta \le \overline{\lim}_n \frac{1}{n}\sum_{m=1}^{n}\mathcal{L}_{K,m}\Psi$ for some $\delta > 0$. Then, taking into account that $I + D^2_{\mathrm{ac}}\Phi \ge (I+K)^2$, we obtain that $\gamma|_M$-a.e.

$$
\begin{aligned}
u_\infty &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} u_k \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \det{}_2\bigl[I + (\Psi^{\rm ac}_{e_ie_j})_{k\times k}\bigr] \exp\Bigl(\sum_{i=1}^{k}(\Psi^{\rm ac}_{e_ie_i} - x_i\partial_{e_i}\Psi) - \frac{1}{2}\sum_{i=1}^{k}(\partial_{e_i}\Psi)^2\Bigr) \\
&\ge \lim_{n\to\infty} \exp\Bigl(\frac{1}{n}\sum_{k=1}^{n}\Bigl[\log\det{}_2\bigl[I + (\Psi^{\rm ac}_{e_ie_j})_{k\times k}\bigr] + \sum_{i=1}^{k}(\Psi^{\rm ac}_{e_ie_i} - x_i\partial_{e_i}\Psi) - \frac{1}{2}\sum_{i=1}^{k}(\partial_{e_i}\Psi)^2\Bigr]\Bigr) \\
&= \exp\Bigl(\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\Bigl[\log\det\bigl[I + (\Psi^{\rm ac}_{e_ie_j})_{k\times k}\bigr] - \sum_{i=1}^{k} x_i\partial_{e_i}\Psi - \frac{1}{2}\sum_{i=1}^{k}(\partial_{e_i}\Psi)^2\Bigr]\Bigr) \\
&\ge \exp\Bigl(\overline{\lim}_n \frac{1}{n}\sum_{k=1}^{n}\Bigl[\log\det\bigl[(I+K)^2_{k\times k}\bigr] - \sum_{i=1}^{k} x_i\partial_{e_i}\Psi - \frac{1}{2}\sum_{i=1}^{k}(\partial_{e_i}\Psi)^2\Bigr]\Bigr) \\
&= \exp\Bigl(\overline{\lim}_n \frac{1}{n}\sum_{k=1}^{n}\Bigl[\log\det{}_2\bigl[(I+K)^2_{k\times k}\bigr] + \sum_{i=1}^{k} k_i^2 + 2k_i - x_i\partial_{e_i}\Psi - \frac{1}{2}\sum_{i=1}^{k}(\partial_{e_i}\Psi)^2\Bigr]\Bigr) \\
&= \det{}_2(I+K)^2 \exp\Bigl[\overline{\lim}_n \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{k}(k_i^2 + 2k_i - x_i\partial_{e_i}\Psi)\Bigr] \exp\Bigl(-\frac{1}{2}|\nabla\Psi|_H^2\Bigr) \\
&\ge \det{}_2(I+K)^2 \exp\Bigl(\widetilde{\mathcal{L}}\Psi + \delta - \frac{1}{2}|\nabla\Psi|_H^2\Bigr) = ge^{\delta}.
\end{aligned}
$$

This contradicts the estimate $u_\infty \le g$ γ-a.e. Hence $\widetilde{\mathcal{L}}\Psi \ge \overline{\lim}_n \frac{1}{n}\sum_{m=1}^{n} \mathcal{L}_{K,m}\Psi$. The case of $\widetilde{\mathcal{L}}\Phi$ is considered in the same way.

Step 3. Let us show that $I + D^2_{\rm ac}\Psi = (I+K)^2$ and $I + D^2_{\rm ac}\Phi = (I + L(T))^2$. Suppose that at a point x the infinite matrix $\bigl(\Psi^{\rm ac}_{e_ie_j}(x)\bigr)$ does not coincide with $K(x)^2 + 2K(x)$. Then there exist a natural number N and a nontrivial nonnegative symmetric operator B of finite rank such that B has the same eigenbasis as $K(x)$ and at the point x one has

$$
(I + K_{n\times n})^2 + B_{n\times n} \le I + (D^2_{\rm ac}\Psi)_{n\times n}
$$

for all $n \ge N$ and there is no equality for $n = N$. Hence, at the point x, there is no equality for all $n \ge N$ and

$$
\det(I + K_{n\times n})^2 < \det\bigl((I + K_{n\times n})^2 + B_{n\times n}\bigr) \le \det\bigl(I + (D^2_{\rm ac}\Psi)_{n\times n}\bigr)
$$

for all $n > N$. Since Tr is additive, the relation $\det_2(I + A) = \det(I + A)\exp(-\mathrm{Tr}A)$ yields

$$
\det{}_2\bigl[(I + K_{n\times n})^2\bigr] < \det{}_2\Bigl[\bigl((I + K_{n\times n})^2 + B_{n\times n}\bigr)\Bigr] \exp \mathrm{Tr} B_{n\times n}.
$$

This gives the strict inequality

$$
\det{}_2\bigl[(I+K)^2\bigr] < \det{}_2\bigl[(I+K)^2 + B\bigr] \exp \mathrm{Tr} B
$$

at the point x. Therefore, exactly as above, we obtain at x the relations

$$
\begin{aligned}
u_\infty &\geq \lim_{n\to\infty} \exp\Bigl(\frac{1}{n}\sum_{k=1}^{n}\Bigl[\log\det\bigl[I + (\Psi^{\mathrm{ac}}_{e_ie_j})_{k\times k}\bigr] - \sum_{i=1}^{k} x_i\partial_{e_i}\Psi - \frac{1}{2}\sum_{i=1}^{k}(\partial_{e_i}\Psi)^2\Bigr]\Bigr) \\
&\geq \lim_{n\to\infty} \exp\Bigl(\frac{1}{n}\sum_{k=1}^{n}\Bigl[\log\det\bigl[(I + K_{k\times k})^2 + B_{k\times k}\bigr] - \sum_{i=1}^{k} x_i\partial_{e_i}\Psi - \frac{1}{2}\sum_{i=1}^{k}(\partial_{e_i}\Psi)^2\Bigr]\Bigr) \\
&\geq \lim_{n\to\infty} \exp\Bigl(\frac{1}{n}\sum_{k=1}^{n}\Bigl[\log\det{}_2\bigl[(I + K_{k\times k})^2 + B_{k\times k}\bigr] \\
&\quad + \sum_{i=1}^{k} k_i^2 + 2k_i + (Be_i, e_i) - x_i\partial_{e_i}\Psi\Bigr]\Bigr) \exp\Bigl(-\frac{1}{2}|\nabla\Psi|_H^2\Bigr) \\
&= \log\det{}_2\bigl[(I+K)^2 + B\bigr] \exp\Bigl(\mathrm{Tr}B + \widetilde{\mathcal{L}}\Psi - \frac{1}{2}|\nabla\Psi|_H^2\Bigr) \\
&> \log\det{}_2(I+K)^2 \exp\Bigl(\widetilde{\mathcal{L}}\Psi - \frac{1}{2}|\nabla\Psi|_H^2\Bigr) = g
\end{aligned}
$$

This contradiction implies our claim. The case of Φ is analogous. Note that the equality

$$(I + D^2_{\mathrm{ac}}\Psi)(I + D^2_{\mathrm{ac}}\Phi(S)) = (I + D^2_{\mathrm{ac}}\Phi(S))(I + D^2_{\mathrm{ac}}\Psi) = I$$

follows from the relation $(I+K)(I+L) = I$. ☐

Lemma 2.10. (i) *Let $g > c > 0$ and $g\log g \in L^1(\gamma)$. Then the series $\sum_{i=1}^\infty \Psi^{\mathrm{sing}}_{e_ie_i}$ converges in variation to a bounded nonnegative Borel measure on X.*

(ii) *If $0 < g \leq C$ and $\log g \in L^1(\gamma)$, then the series $\sum_{i=1}^\infty \Phi^{\mathrm{sing}}_{e_ie_i}$ converges in variation to a bounded nonnegative Borel measure on X.*

Proof. (i) Let us consider the sequence of functions

$$\mathrm{Tr}\bigl(D^2\Psi_n + D^2\Phi_n(S_n)\bigr) = \mathrm{Tr}\bigl(2(K_n + L_n) + K_n^2 + L_n^2\bigr).$$

Since $(I + K_n)(I + L_n) = I$, we have $K_nL_n = -K_n - L_n$ and

$$\mathrm{Tr}\bigl(D^2\Psi_n + D^2\Phi_n(S_n)\bigr) = \mathrm{Tr}(K_n - L_n)^2.$$

By Theorem 2.2 and the assumption $g > c$, we have

$$\sup_n \int_X \mathrm{Tr}\Bigl(K_n^2 + L_n^2\Bigr)\, d\gamma < \infty,$$

hence

$$\sup_n \int_X \mathrm{Tr}\bigl(D^2\Psi_n + D^2\Phi_n(S_n)\bigr)\, d\gamma := M < \infty.$$

By the integration by parts formula

$$\lim_{n\to\infty}\int_X (\Psi_n)_{e_ie_i}\, d\gamma = \lim_{n\to\infty}\int_X \partial_{e_i}\Psi_n\widehat{e}_i\, d\gamma = \int_X \partial_{e_i}\Psi\widehat{e}_i\, d\gamma = \int_X \Psi^{\mathrm{ac}}_{e_ie_i}\, d\gamma + \Psi^{\mathrm{sing}}_{e_ie_i}(X).$$

Note that $\gamma\circ S_n^{-1} = \frac{1}{g_n(T_n)}\, d\gamma$. Since $\sup_n \int_X \log\frac{1}{g_n}\, d\gamma < \infty$, the sequence of densities $d\gamma\circ S_n^{-1}/d\gamma = 1/g_n(T_n)$ is uniformly integrable. Since $(\Phi_n)_{e_ie_i} \to \Phi^{\mathrm{ac}}_{e_ie_i}$ γ-a.e. (see

Theorem 2.2 and Lemma 2.9), we obtain $(\Phi_n)_{e_ie_i}\circ S_n \to \Phi^{\rm ac}_{e_ie_i}\circ S$ in measure. By Fatou's theorem one has

$$M \ge \underline{\lim}_n \int_X \mathrm{Tr}\bigl(D^2\Psi_n + D^2\Phi_n(S_n)\bigr)\,d\gamma$$
$$\ge \int_X \mathrm{Tr}\bigl(D^2_{\rm ac}\Psi + D^2_{\rm ac}\Phi(S)\bigr)\,d\gamma + \sum_{i=1}^\infty \Psi^{\rm sing}_{e_ie_i}(X) = \int_X \mathrm{Tr}(K-L)^2\,d\gamma + \sum_{i=1}^\infty \Psi^{\rm sing}_{e_ie_i}(X).$$

Since $\sum_{i=1}^\infty \Psi^{\rm sing}_{e_ie_i} \ge 0$, we obtain $\|\sum_{i=1}^\infty \Psi^{\rm sing}_{e_ie_i}\| < \infty$. Case (ii) is considered in a similar way. □

Proof of Theorem 1.2. In view of Lemma 2.9 it remains to show (ii) and (iii). Consider case (ii). We have to prove that if $g > c > 0$ and $g\log g \in L^1(\gamma)$, then $\mathcal{L}_0\Psi$ coincides with $\mathcal{L}_{\rm ac}\Psi$, the density of the absolutely continuous part of the distributional divergence of $\nabla\Psi$. To this end we consider yet another approximation of Ψ by the conditional expectations $\Lambda_n := \mathbb{E}(\Psi|\mathcal{F}_n)$. It is readily verified that Λ_n is 1-convex. By Corollary 2.5 there exists the $\mathcal{H}$-valued measure $D^2\Psi$ of bounded variation. We have the decomposition $D^2\Psi = D^2_{\rm ac}\Psi\cdot\gamma + D^2_{\rm sing}\Psi$. Let us denote by $D^2_{{\rm ac},n}\Psi$ the matrix whose (i,j)-element equals $\langle D^2_{\rm ac}\Psi e_i, e_j\rangle_H$ if $i,j \le n$ and is zero otherwise. Let $D^2_{{\rm sing},n}\Psi$ be defined similarly for the singular part. Given a bounded Borel measure m on X, let $[m]_{\rm ac}$ denote the density of the absolutely continuous part of m with respect to γ. The same notation is used for $\mathcal{H}$-valued measures. It follows from the integration by parts formula that

$$D^2_{\rm ac}\Lambda_n = \mathbb{E}(D^2_{{\rm ac},n}\Psi + D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)$$
$$= \mathbb{E}(D^2_{{\rm ac},n}\Psi|\mathcal{F}_n) + \bigl[\mathbb{E}(D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)\bigr]_{\rm ac}.$$

By the Jessen theorem (see [19, Theorem 1.2.1]) and Lemma 2.10 we obtain

$$\lim_{n\to\infty}\Bigl[\mathbb{E}\Bigl(\sum_{i=1}^\infty \Psi^{\rm sing}_{e_ie_i}|\mathcal{F}_n\Bigr)\Bigr]_{\rm ac} = 0 \quad \text{a.e.}$$

Hence $\mathrm{Tr}\bigl[\mathbb{E}(D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)\bigr]_{\rm ac} \to 0$. Since $\bigl[\mathbb{E}(D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)\bigr]_{\rm ac}$ is nonnegative, the sequence of $\mathcal{H}$-valued mappings $\bigl[\mathbb{E}(D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)\bigr]_{\rm ac}$ converges to zero also in the Hilbert–Schmidt norm. We have

$$\|D^2_{\rm ac}\Lambda_n\|_{\mathcal{H}} \le \bigl\|\mathbb{E}(D^2_{{\rm ac},n}\Psi|\mathcal{F}_n)\bigr\|_{\mathcal{H}} + \bigl\|\bigl[\mathbb{E}(D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)\bigr]_{\rm ac}\bigr\|_{\mathcal{H}}$$
$$\le \mathbb{E}\bigl(\|D^2_{{\rm ac},n}\Psi\|_{\mathcal{H}}|\mathcal{F}_n\bigr) + \bigl\|\bigl[\mathbb{E}(D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)\bigr]_{\rm ac}\bigr\|_{\mathcal{H}}$$
$$\le \mathbb{E}\bigl(\|D^2_{\rm ac}\Psi\|_{\mathcal{H}}|\mathcal{F}_n\bigr) + \bigl\|\bigl[\mathbb{E}(D^2_{{\rm sing},n}\Psi|\mathcal{F}_n)\bigr]_{\rm ac}\bigr\|_{\mathcal{H}}.$$

As $\bigl\|D^2_{\rm ac}\Psi\bigr\|_{\mathcal{H}} \in L^1(\gamma)$, we obtain $\overline{\lim}_n\|D^2_{\rm ac}\Lambda_n\|_{\mathcal{H}} \le \bigl\|D^2_{\rm ac}\Psi\bigr\|_{\mathcal{H}}$ γ-a.e. On the other hand, for every $h \in H$ one has

$$(\Lambda_n)^{\rm ac}_{hh} = \mathbb{E}(\Psi^{\rm ac}_{hh}|\mathcal{F}_n) + \bigl[\mathbb{E}(\Psi^{\rm sing}_{hh}|\mathcal{F}_n)\bigr]_{\rm ac} \to \Psi^{\rm ac}_{hh} \quad \gamma\text{-a.e.}$$

Hence for γ-almost all x we have $D^2_{\rm ac}\Lambda_n(x) \to D^2_{\rm ac}\Psi(x)$ weakly in $\mathcal{H}$ and

$$\underline{\lim}_n\|D^2_{\rm ac}\Lambda_n(x)\|_{\mathcal{H}} \ge \|D^2_{\rm ac}\Psi(x)\|_{\mathcal{H}}.$$

Combining this inequality with the previous one we conclude that

$$\lim_{n\to\infty}\|D^2_{\rm ac}\Lambda_n\|_{\mathcal{H}} = \|D^2_{\rm ac}\Psi\|_{\mathcal{H}} \quad\text{and}\quad \lim_{n\to\infty}\|D^2_{\rm ac}\Lambda_n - D^2_{\rm ac}\Psi\|_{\mathcal{H}} = 0 \quad \gamma\text{-a.e.}$$

In particular,

$$\lim_{n\to\infty}\det{}_2(I + D^2_{\rm ac}\Lambda_n) = \det{}_2(I + D^2_{\rm ac}\Psi) \quad \gamma\text{-a.e.} \tag{2.23}$$

The justification of convergence $|\nabla\Lambda_n|_H^2 \to |\nabla\Psi|_H^2$ γ-a.e. follows the same ideas and is even simpler. According to Lemma 2.6 (the hypotheses of that lemma are fulfilled in the situation we consider) one has $\mathcal{L}_{\rm ac}\Lambda_n \to \mathcal{L}_{\rm ac}\Psi$ γ-a.e. Let us define w_n by

$$\log\frac{1}{w_n} = -\mathcal{L}_{\rm ac}\Lambda_n + \frac{1}{2}|\nabla\Lambda_n|_H^2 - \log\det{}_2(I + D^2_{\rm ac}\Lambda_n).$$

Clearly, w_n admits a limit for γ-almost all x. We denote this limit by w_∞ and write

$$\log\frac{1}{w_\infty} = -\mathcal{L}_{\rm ac}\Psi + \frac{1}{2}|\nabla\Psi|_H^2 - \log\det{}_2(I + D^2_{\rm ac}\Psi).$$

Exactly as above we prove with the help of Theorem 1.1 and Lemma 2.8 that the mapping $G_n := I + \nabla\Lambda_n$ sends $w_n\cdot\gamma$ to $I_{G_n(X)}\cdot\gamma$. Let us show that $w_\infty = g$. Let us fix a nonnegative function $\zeta \in \mathcal{F}C_b^\infty$. Let $\mathcal{L}_{\rm sing}\Psi := \mathcal{L}\Psi - \mathcal{L}_{\rm ac}\Psi\cdot\gamma$ be the singular part of the measure $\mathcal{L}\Psi$. As shown in Lemma 2.9 (see (2.20)) one has $\mathcal{L}\Psi_n \to \mathcal{L}_0\Psi$ γ-a.e. In addition, $\mathcal{L}\Psi_n \ge \log g_n > \log c$, because $\det_2(I + D^2\Psi_n) \le 1$. By Fatou's theorem we have

$$\begin{aligned}\int_X \zeta\mathcal{L}_0\Psi\,d\gamma &\le \lim_{n\to\infty}\int_X \zeta\mathcal{L}\Psi_n\,d\gamma \le -\lim_{n\to\infty}\int_X \langle\nabla\Psi_n,\nabla\zeta\rangle_H\,d\gamma\\ &= \int_X \zeta\mathcal{L}_{\rm ac}\Psi\,d\gamma + \int_X \zeta(x)\,\mathcal{L}_{\rm sing}\Psi(dx).\end{aligned}$$

Since $\mathcal{L}_{\rm sing}\Psi$ is singular with respect to γ, we have $\mathcal{L}_0\Psi \le \mathcal{L}_{\rm ac}\Psi$. Hence $w_\infty \ge g$ γ-a.e. On the other hand, since $w_n \to w_\infty$, one has

$$\begin{aligned}\int_X \zeta(S)w_\infty\,d\gamma &\le \lim_{n\to\infty}\int_X \zeta\big(x + \nabla\Lambda_n(x)\big)w_n(x)\,\gamma(dx)\\ &= \lim_{n\to\infty}\int_{G_n(X)}\zeta\,d\gamma \le \int_X \zeta\,d\gamma = \int_X \zeta(S)g\,d\gamma.\end{aligned}$$

Since S is invertible, this implies the opposite inequality by the same reasoning as in Lemma 2.9. Therefore, we obtain the equality $w_\infty = g$, which completes the proof of (1.6). Case (iii) is analogous. ☐

One can ask about conditions ensuring the absolute continuity of $D^2\Psi$ and $D^2\Phi$. A result of this type can be deduced from Caffarelli's estimate. It has been shown in [7] that if the optimal mapping $I + \nabla\Phi$ takes the standard Gaussian measure γ on $\mathbb{R}^d$ to the measure $e^{-V}\cdot\gamma$, where V is convex, then $I + D^2\Phi \le I$. Following Caffarelli's techniques it is not hard to prove that if $V_{hh} \ge -1 + \varepsilon$ for some $h \in \mathbb{R}^d$, $|h| = 1$ and $\varepsilon > 0$, then $1 + \Phi_{hh} \le \frac{1}{\varepsilon}$. Analogously one can prove that if a second partial derivative of V is bounded *from above*, then the corresponding second partial derivative of $\frac{x^2}{2} + \Psi$ is bounded (see [11] for details). These estimates can be generalized to the infinite dimensional case. In particular, under this type of restriction on g one obtains the absolute continuity of $D^2\Phi$ and $D^2\Psi$ (see [11] for details). For example, if $D^2(-\log g)$ is bounded either from above or from below, then the corresponding potential (Φ or Ψ) belongs to the Sobolev class $W^{2,2}(\gamma)$. The precise statements are given below. The following proposition generalizes Lemma 5.1 from [10]. The proof is similar.

Proposition 2.11. *Let $X = \mathbb{R}^d$ and let γ be the standard Gaussian measure.*

(i) *Let $g = e^\Phi$ and $D^2\Phi \le M$, where $M < 1$. Then $I + D^2\Phi \le \frac{1}{\sqrt{1-M}}I$. In addition,*

$$\int_{\mathbb{R}^d}|\nabla\Phi|^2\,d\gamma + (1 - M)\int_{\mathbb{R}^d}\|D^2\Phi\|_{\mathcal{H}}^2\,d\gamma \le 2\int_{\mathbb{R}^d} g\log g\,d\gamma.$$

(ii) *Let* $g = e^{\Psi}$ *and* $D^2\Psi \geq -M$, *where* $M > -1$. *Then* $I + D^2\Psi \leq \sqrt{1+M} \cdot I$. *In addition,*

$$\int_{\mathbb{R}^d} |\nabla \Psi|^2 \, d\gamma + \frac{1}{1+M} \int_X \|D^2\Psi\|_{\mathcal{H}}^2 \, d\gamma \leq 2 \int_{\mathbb{R}^d} \log \frac{1}{g} \, d\gamma.$$

These finite dimensional estimates can be easily generalized to the infinite dimensional situation. As a result we obtain the following statement.

Corollary 2.12. *Let* g *be a probability density with respect to* γ. *Suppose that* $g = e^{-V}$, *where the function* V *is* $(1-\varepsilon)$*-convex,* $\varepsilon > 0$. *Then* $\Phi \in W^{2,2}(\gamma)$. *If* $g = e^{W}$, *where* W *is an* M*-convex function for some* $M > -1$, *then* $\Psi \in W^{2,2}(\gamma)$.

Acknowledgements. This work was completed during our very pleasant visit to the University of Bielefeld. We thank M. Röckner for useful discussions. This research was supported by the RFBR project 04-01-00748, the DFG Grant 436 RUS 113/343/0(R), the INTAS project 03-51-5018, the Scientific Schools Grant 1758.2003.1, the DFG–Forschergruppe "Spectral Analysis, Asymptotic Distributions, and Stochastic Dynamics", the BiBoS–reseach centre. The second author holds the Junior Visiting Position of the Centro di Ricerca Matematica Ennio De Giorgi, Scuola Normale Superiore di Pisa.

References

[1] L. Ambrosio, A. Pratelli *Existence and stability results in the L^1 theory of optimal transportation.* Optimal transportation and applications (Martina Franca, 2001), 123–160, Lecture Notes in Math., 1813, Springer, Berlin, 2003.

[2] V.I. Bogachev, Gaussian measures. Amer. Math. Soc., Rhode Island, 1998.

[3] V.I. Bogachev, Foundations of measure theory. V. 1,2. Regular and Chaotic Dynamics, Moscow (in Russian), 2003.

[4] V.I. Bogachev, B. Goldys, *On second order derivatives of convex functions.* Preprints der Forschergruppe "Spektrale Analysis und stochastische Dynamik" N 04-046 (2004), Universität Bielefeld.

[5] Y. Brenier, *Polar factorization and monotone rearrangement of vector valued functions.* Comm. Pure Appl. Math. **44** (1991), 375–417.

[6] L.A. Caffarelli, *Some regularity properties of solutions of Monge–Ampère equation.* Comm. Pure Appl. Math. **44** (1991), no. 8–9, 965-969.

[7] L.A. Caffarelli, *Monotonicity properties of optimal transportation and the FKG and related inequalities.* Comm. Math. Phys. **214** (2000), no. 3, 547–563.

[8] D. Feyel, A.S. Üstünel, *The notions of convexity and concavity on Wiener space.* J. Funct. Anal. **176** (2000), 400–428.

[9] D. Feyel, A.S. Üstünel, *Monge–Kantorovitch measure transportation and Monge–Ampère equation on Wiener space.* Probab. Theor. Relat. Fields **128** (2004), no. 3, 347–385.

[10] D. Feyel, A.S. Üstünel, *Solution of the Monge–Ampère equation on Wiener space for log-concave measures: general case.* Preprint Arxiv.org (2004).

[11] A.V. Kolesnikov, *Convexity inequalities and optimal transport of infinite-dimensional measures.* J. Math. Pures Appl. **83** (2004), no. 11, 1373–1404.

[12] M. Ledoux, The concentration of measure phenomenon. Amer. Math. Soc., Providence, Rhode Island, 2001.

[13] R.J. McCann, *Existence and uniqueness of monotone measure-preserving maps.* Duke Math. J. **80** (1995), 309–323.

[14] R.J. McCann, *A convexity principle for interacting gases.* Adv. Math. **128** (1997), no. 1, 153–179.

[15] S.T. Rachev, L. Rüschendorf, Mass transportation problems. V. I, II, Springer, New York, 1998.

[16] V.N. Sudakov, Geometric problems of the theory of infinite–dimensional probability distributions. Trudy Mat. Inst. Steklov. 1976. V. 141. P. 1–190 (in Russian); English transl.: Proc. Steklov Inst. Math. 1979. N 2. P. 1–178.

[17] H. Sugita, *Positive generalized Wiener functions and potential theory over abstract Wiener spaces.* Osaka J. Math. **25** (1988), no. 3, 665–696.

[18] M. Talagrand, *Transportation cost for Gaussian and other product measures.* Geom. Funct. Anal. **6** (1996), 587–600.

[19] A.S. Üstünel, M. Zakai, Transformation of measure on Wiener space. Springer, Berlin, 2000.

[20] C. Villani, Topics in optimal transportation. Amer. Math. Soc., Providence, Rhode Island, 2003.

Department of Mechanics and Mathematics, Moscow State University, 119992 Moscow, Russia

Scuola Normale Superiore, Centro di Ricerca Matematica Ennio De Giorgi, I-56100 Pisa, Italy, E-mail: sascha77@mail.ru