ON NONLINEAR TRANSFORMATIONS OF CONVEX MEASURES

V.I. BOGACHEV, A.V. KOLESNIKOV

ABSTRACT. Given a uniformly convex measure μ on $\mathbb{R}^\infty$ and a probability measure $\nu \ll \mu$, we show that there is a Borel transformation $T = (T_k)_{k=1}^\infty$ of $\mathbb{R}^\infty$ such that $\nu = \mu \circ T^{-1}$ and $F(x) := T(x) - x \in l^2$. Moreover, if ν is equivalent to its translation along $(1, 0, 0, \ldots)$, e.g., if ν is a product-measure, then T can be chosen triangular in the sense that each component T_k is a function of $x_1, \ldots, x_k$. In addition, if ν has finite entropy $\mathrm{Ent}_\mu(\nu)$ with respect to μ, then $\|F\|^2_{L^2(\mu,l^2)} \le C(\mu)\mathrm{Ent}_\mu(\nu)$. Several inverse results are proved. In particular, our results apply to the standard Gaussian product-measure. As an application we obtain a new sufficient condition for the absolute continuity of a nonlinear image of a convex measure and the membership of its Radon–Nikodym derivative in the class $L \log L$.

Let $X = \mathbb{R}^\infty$ be the space of all real sequences $x = (x_n)$ equipped with its natural product topology and the corresponding Borel σ-algebra $\mathcal{B}(X)$ and let $H = l^2$ be equipped with its usual Hilbert norm $|h|_H := \Bigl(\sum_{n=1}^\infty h_n^2\Bigr)^{1/2}$. Given a Borel measure μ on X and a μ-measurable mapping T on X, we denote by $\mu \circ T^{-1}$ the image of μ under T, i.e., $\mu \circ T^{-1}(B) = \mu(T^{-1}(B))$ for all $B \in \mathcal{B}(X)$. Suppose that μ is a Borel probability measure on X that is the product of countably many copies of a probability measure σ on the real line. When σ is the standard Gaussian measure, we call μ the standard Gaussian product-measure. It is well known that under broad assumptions on σ, the shifted measure $\mu_h(B) = \mu(B - h)$ is equivalent to μ if $h \in l^2$ and is mutually singular with μ otherwise (see Shepp [20]). Moreover, in many cases μ is equivalent to its images under nonlinear mappings of the form

$$T(x) = x + F(x), \quad F\colon\ X \to H, \tag{1}$$

under broad assumptions on F. The case where σ is Gaussian has been studied especially well. Although transformations of the above form are not the only ones under which the image of μ is absolutely continuous (e.g., there many automorphisms of μ not of that form), it is natural to ask which measures ν can be obtained as $\mu \circ T^{-1}$ with T of form (1). Measures $\nu \ll \mu$ with such a property are called representable in Definition 2.7.1 in Üstünel, Zakai [24] (where the Gaussian case is considered). Note that even in the Gaussian case, the condition $F(X) \subset H$ does not

2000 *Mathematics Subject Classification.* Primary 28C20, 46G12, 60B11; Secondary 47H99, 60E15, 60G30.

Key words and phrases. Convex measure, Gaussian measure, product-measure, Cameron–Martin space, absolute continuity, triangular mapping.

Partially supported by the Russian Foundation of Basic Research project 01-01-00858, Grant RFBR MAS 03-01-06270, Scientific Schools Grant 1758.2003.1 and the DFG Grant 436 RUS 113/343/0(R) .

guarantee that $\mu \ll \nu$. We prove below that if μ belongs to the class of uniformly convex measures (which includes the standard Gaussian product-measure), every probability measure $\nu \ll \mu$ is the image of μ under a triangular mapping T of type (1), i.e., is representable in the sense of Üstünel and Zakai. Moreover, under some additional, not very restrictive assumptions, the mapping T can be found with the property that $|F|_H \in L^2(\mu)$ and each component T_k of T is increasing in x_k. Finally, we give a sufficient condition on a mapping T of type (1) which guarantees that $\nu = \mu \circ T^{-1}$ is absolutely continuous with respect to μ. This condition is also necessary for increasing triangular mappings and measures ν with finite entropy $\mathrm{Ent}_\mu(\nu)$. We recall that in the case where $\nu \ll \mu$ and $f := d\nu/d\mu$ entropy $\mathrm{Ent}_\mu(\nu)$ is defined by

$$\mathrm{Ent}_\mu(\nu) := \mathrm{Ent}_\mu(f) := \int f \log f \, d\mu.$$

The main results are Theorems 2, 4 and 6. It should be noted that our results are new also for Gaussian measures; they extend previously known results with much shorter and elementary proofs.

A mapping $T = (T_k)_{k=1}^\infty \colon X \to X$ is said to be triangular if each T_k is a function of $x_1, \dots, x_k$: $T_k(x) = T_k(x_1, \dots, x_k)$. Throughout we consider all $\mathbb{R}^k$ as subspaces in X. Such a mapping is called increasing if the function $x_k \mapsto T_k(x_1, \dots, x_{k-1}, x_k)$ is increasing for all fixed $(x_1, \dots, x_{k-1}) \in \mathbb{R}^{k-1}$. Triangular mappings provide a simple and constructive way to transform one given measure into another. Such mappings were studied, e.g., in Knote [13], Talagrand [21], Bobkov [2]. In particular, Talagrand [21] proved that if μ is the standard Gaussian measure on $\mathbb{R}^n$ and $\nu \ll \mu$ is such that $\mathrm{Ent}_\mu(\nu)$ is finite, there exists an increasing triangular mapping T such that $\nu = \mu \circ T^{-1}$ and

$$\int |T(x) - x|^2 \, \mu(dx) \le 2\mathrm{Ent}_\mu(\nu). \tag{2}$$

Talagrand's inequality (2) has been generalized by several authors, see Otto, Villani [18], Cordero-Erausquin [8]. Very interesting links of this inequality to the mass transportation problems (see, e.g., Rachev, Rüschendorf [19]) are discussed in Fernique [9], Feyel, Üstünel [11], Feyel, Üstünel [12], Ledoux [16]. In particular, by using estimate (2), Feyel and Üstünel [12] show that if μ is the standard Gaussian product measure and $\nu \ll \mu$ has finite entropy, then one can find T of type (1) such that $\nu = \mu \circ T^{-1}$, estimate (2) holds, and $F = \nabla_H \psi$ for some function ψ in the Sobolev class $W^{2,1}(\mu)$. Fernique [9] shows that for any probability measure $\nu \ll \mu$, there is an automorphism U of μ and a mapping $F \colon X \to H$ such that $\nu = \mu \circ (U + F)^{-1}$. However, U may not be of type (1). In Kolesnikov [14] and Bogachev, Kolesnikov, Medvedev [5], a generalization of Talagrand's result is obtained in the form of equality (generalizing an identity established by Talagrand in the one dimensional case) for a couple of mappings. We recall this result (in the case of a single mapping), because it will be used in some of our proofs.

Let us introduce necessary notation. Given a positive C^2 function ψ on $\mathbb{R}^n$ and two vectors v_1 and v_2, we define the operator

$$\Lambda\bigl[\psi, v_1, v_2\bigr] = \int_0^1 s D^2\bigl[-\log \psi\bigr]\bigl((1-s)v_1 + s v_2\bigr)\, ds,$$

where D^2 is the second derivative. If $\psi(x) = \exp\bigl(-|x|^2/2\bigr)$, then $\Lambda = I/2$. The Fredholm–Carleman determinant $\det_2 A$ of an operator A on $\mathbb{R}^n$ is defined by

$$\det_2 A = \exp\bigl(\mathrm{trace}(I - A)\bigr)\det A.$$

Suppose that μ is a probability measure on $\mathbb{R}^n$ with density $\exp(-\Phi)$, where $\Phi \in C^2$ and $T = I + F\colon \mathbb{R}^n \to \mathbb{R}^n$ is an invertible C^1-mapping such that $\det T(x) > 0$. Then $\nu := \mu\circ T^{-1}$ has a density f with respect to μ. Assume, in addition, that

$$\mathrm{trace} DT, \quad \log\det DT, \quad \partial_{x_i}\Phi\cdot F_i, \quad \bigl(\Lambda[\exp(-\Phi), T, I]F, F\bigr) \in L^1(\mu),$$

and $\mathrm{Ent}_\mu(\nu)$ is finite. Then, according to [5], one has

$$\mathrm{Ent}_\mu(\nu) = \int_{\mathbb{R}^n} \Bigl(\Lambda\bigl[\exp(-\Phi), T(x), x\bigr] F(x), F(x)\Bigr)\,\mu(dx) - \int_{\mathbb{R}^n} \log\det_2 DT(x)\,\mu(dx). \quad (3)$$

We shall need below the following modification of this result that is proved by the same reasoning: the assumption that T is C^1 can be replaced by the inclusion $T \in W^{2,1}_{loc}(\mathbb{R}^n,\mathbb{R}^n)$ provided that $f\circ T\exp(-\Phi\circ T)\det DT = \exp(-\Phi)$ a.e. For example, it is enough (along with the above integrability conditions and $\det DT > 0$ a.e.) that T be invertible and locally Lipschitzian (or $\|DT\| \in L^p_{loc}$ with some $p > n$).

If μ is the standard Gaussian measure on $\mathbb{R}^n$, then (3) becomes

$$\mathrm{Ent}_\mu(\nu) = \frac{1}{2}\int_{\mathbb{R}^n} |F(x)|^2\,\mu(dx) - \int_{\mathbb{R}^n} \log\det_2\bigl(I + DF(x)\bigr)\,\mu(dx). \quad (4)$$

This can be also seen from the formula for the Radon–Nikodym density:

$$f(x) = \frac{1}{\Lambda_F\bigl(T^{-1}(x)\bigr)}, \quad \Lambda_F = \det_2(I + DF)\exp\Bigl(\delta F - \frac{1}{2}|F|^2\Bigr), \quad (5)$$

where δF is the divergence of F with respect to μ, i.e., $\delta F(x) = \mathrm{trace} DF(x) - \bigl(F(x), x\bigr)$. Indeed, by the change of variable formula

$$\mathrm{Ent}_\mu(f) = \int \log f\bigl(T(x)\bigr)\,\mu(dx) = \int \Bigl[-\log\det_2(I + DF) - \delta F + \frac{1}{2}|F|^2\Bigr]\,d\mu,$$

but the integral of δF vanishes. Let us observe that formula (5) holds true if T is a continuous invertible mapping in the Sobolev class $W^{p,1}_{loc}(\mathbb{R}^n,\mathbb{R}^n)$ with some $p > n$ and $\det DT > 0$ a.e.

Let $P_n\colon \mathbb{R}^\infty \to \mathbb{R}^n$ be the natural projection. Given a Borel measure μ on $\mathbb{R}^\infty$, let $\mu_n := \mu\circ P_n^{-1}$. For concepts related to Gaussian measures (the Cameron–Martin space, etc.), see Bogachev [3]; however, without loss of generality one can assume that we deal with the standard Gaussian product-measure whose Cameron–Martin space is $H = l^2$.

A probability measure μ on $\mathbb{R}^n$ with density $\varrho_\mu = \exp(-\Phi)$, where Φ is a twice continuously differentiable (or of the class $W^{1,2}_{loc}$) convex function is said to be uniformly convex with constant $C > 0$ if $D^2\Phi(x) \geq C\cdot I$, where I is the identity operator.

We shall say that a Borel probability measure μ on X is uniformly convex with constant $C > 0$ if every projection μ_n is uniformly convex with constant C, i.e., one has $\mu_n = \exp(-\Phi_n)\,dx$ and $D^2\Phi_n \geq C\cdot I$ on $\mathbb{R}^n$.

By using (3), it is proved in [5] that given a probability measure μ that is uniformly convex with constant C and a probability measure $\nu \ll \mu$ with finite entropy $\mathrm{Ent}_\mu(\nu)$, there is a Borel increasing triangular mapping T such that $\nu = \mu\circ T^{-1}$ and

$$\int |F(x)|^2\,\mu(dx) \le \frac{2}{C}\mathrm{Ent}_\mu(\nu). \tag{6}$$

It should be emphasized that in all the results mentioned above the main point is to obtain certain additional properties of T such as estimate (2). If one admits arbitrary triangular mappings, then it is easy to show that any Borel probability measure ν on $\mathbb{R}^n$, not necessarily absolutely continuous, is the image of μ (or any other probability measure absolutely continuous with respect to Lebesgue measure) under an increasing triangular Borel mapping. The same is true for the space $\mathbb{R}^\infty$ if we take for μ any countable product of nonatomic measures. In infinite dimensions, also the requirement that $T(x) - x \in H$ becomes very restrictive. Let us describe a simple construction of such a mapping that we call the canonical triangular transformation of μ into ν. Let μ and ν be absolutely continuous probability measures on $\mathbb{R}^n$. The canonical increasing triangular Borel mapping $T_{\mu,\nu}$ transforming μ into ν is defined inductively as follows. If $n=1$, then we set $F_\mu(t) := \mu\bigl((-\infty,t)\bigr)$ for $t\in\mathbb{R}^1$, $G_\mu(u) := \inf\bigl\{s\colon\ F_\mu(s)\ge u\bigr\}$ for $u\in(0,1)$ and $T_{\mu,\nu} := G_\nu\circ F_\mu$. The function F_μ takes μ to Lebesgue measure λ on $(0,1)$, and G_ν takes λ to ν. This is true for any Borel probability measures provided that μ has no atoms. In the case $n=2$ we take the canonical mapping T_1 that takes μ_1 to ν_1. Let f and g be Borel measurable densities of μ and ν respectively. For μ_1-a.e. x_1 we have the conditional probability density

$$f^{x_1}(x_2) := f(x_1,x_2)\Bigl(\int_{\mathbb{R}} f(x_1,u)\,du\Bigr)^{-1}.$$

The measure with this density can be canonically transformed to the probability measure with density $g^{T_1(x_1)}(\,\cdot\,)$, where

$$g^{x_1}(x_2) := g(x_1,x_2)\Bigl(\int_{\mathbb{R}} g(x_1,u)\,du\Bigr)^{-1}.$$

Note that $\int g\bigl(T_1(x_1),u\bigr)\,du > 0$ for μ_1-a.e. x_1 due to the equality $\nu_1 = \mu_1\circ T_1^{-1}$. The constructed canonical mapping is denoted by $x_2\mapsto T_2(x_1,x_2)$. We set $T_2(x_1,x_2) = x_2$ for all x_1 such that $\int g(x_1,u)\,du = 0$. It is readily verified that $T := (T_1,T_2)$ is an increasing triangular Borel mapping and $\mu\circ T^{-1} = \nu$. We continue by induction on n by using the one dimensional conditional densities on the last coordinate line. Uniqueness in the class of μ-equivalent increasing triangular mappings is shown also by induction. Estimate (6) is proved in [5] for the canonical mapping. The mapping $T_{\mu,\nu}$ may be discontinuous even if both μ and ν have smooth densities. For example, if ν vanishes on some interval, then it cannot be a continuous image of the standard Gaussian measure. It is seen from the construction that if μ is equivalent to Lebesgue measure and ν is absolutely continuous, then there is a version of $T_{\mu,\nu}$ such that its k-th component is strictly increasing in x_k for all fixed $(x_1,\dots,x_{k-1})$. Hence this version is injective. Canonical mappings for general measures have been investigated recently by D. Alexandrova. It is proved in Bogachev, Kolesnikov, Medvedev [5] that if μ is equivalent to Lebesgue measure on $\mathbb{R}^n$ and probability measures ν_j converge to ν in variation, then $T_{\mu,\nu_j}\to T_{\mu,\nu}$ in measure μ.

Let us consider a simple example how (3) works.

Example 1. *Let μ be a Borel probability measure on X such that $\mu_n = \exp(-\Phi_n)\,dx$ for each n, where Φ_n is a C^2-function on $\mathbb{R}^n$.*

(i) *Suppose that $D^2\Phi_n(x) \le M \cdot I$ on $\mathbb{R}^n$ for all n, where $M > 0$ is a constant. Then for every $h \in H$, the measure μ_h that is the image of μ under the translation $T(x) = x - h$, is equivalent to μ and has finite entropy.*

(ii) *If for all n we have $D^2\Phi_n(x) \le M \exp q(x) \cdot I$ on $\mathbb{R}^n$, where*

$$q(x) = \Bigl(\sum_{n=1}^{\infty} \alpha_n x_n^2\Bigr)^{1/2}, \quad \alpha_n \ge 0, \quad \lim_{n\to\infty} \alpha_n = 0, \quad \text{and } \exp q \in L^1(\mu),$$

then $\mu_h \sim \mu$ for every $h \in H$ and μ_h has finite entropy.

(iii) *Under the assumptions in* (ii) *let $T_n(x) = x_n + F_n(x_1, \dots, x_{n-1})$. Assume that $|F|_H^2 \exp q \in L^1(\mu)$, where $F = (F_n)$. Then $\mu \circ T^{-1} \ll \mu$ and*

$$\mathrm{Ent}_\mu(\mu \circ T^{-1}) \le \frac{M}{2} \int_X |F(x)|_H^2\, \mu(dx).$$

Proof. Clearly, in all these cases T is an increasing triangular transformation. Set $\nu := \mu \circ T^{-1}$. Letting $S_n := P_n T$ on $\mathbb{R}^n$, we have $\log \det_2 DS_n = 0$. Note that $\nu_n = \mu_n \circ S_n^{-1}$, $\nu_n = f_n \cdot \mu_n$, and the sequence $\{f_n\}$ is a martingale on (X, μ) with respect to the filtration $\{\mathcal{F}_n\}$ generated by the projections P_n. In case (i) according to (3) one has $\mathrm{Ent}_{\mu_n}(\nu_n) \le M|h|_H^2/2$. Hence the martingale $\{f_n\}$ is uniformly integrable and converges to a function $f \in L^1(\mu)$, whence we obtain $\nu = f \cdot \mu$. By Fatou's theorem $\mathrm{Ent}_\mu(\nu) \le M|h|_H^2/2$. In case (ii) we have $q(h) < \infty$, hence

$$D^2\Phi_n\bigl(x - (1-s)P_n h\bigr) \le M \exp[q(P_n x) + q(P_n h)] \cdot I \le M \exp q(h) \exp q(P_n x) \cdot I$$

on $\mathbb{R}^n$ whenever $0 \le s \le 1$. This yields

$$\begin{aligned}\int \bigl(\Lambda[\exp(-\Phi_n), x - P_n h, x] P_n h, P_n h\bigr)\, \mu_n(dx) &\le \frac{M|P_n h|_H^2 \exp q(h)}{2} \int \exp q \circ P_n \, d\mu \\ &\le \frac{M|h|_H^2 \exp q(h)}{2} \int \exp q \, d\mu.\end{aligned}$$

Therefore, we have a uniform bound on $\mathrm{Ent}_{\mu_n}(\nu_n)$. A similar reasoning proves assertion (iii). □

It is worth noting that if in (ii) we have also a lower bound $C \cdot I \le D^2\Phi_n(x)$, where $C > 0$, and $q(x) < \infty$ a.e., then $\exp(\kappa q) \in L^1(\mu)$ for all $\kappa > 0$.

Theorem 2. *Suppose that a Borel probability measure μ on X is uniformly convex with constant $C > 0$. Let $\nu \ll \mu$ be a probability measure.*

(i) *If ν has finite entropy $\mathrm{Ent}_\mu(\nu)$, then the canonical triangular mapping $T_{\mu,\nu}$ has the property that*

$$\int_X |T_{\mu,\nu}(x) - x|_H^2\, \mu(dx) \le \frac{2}{C} \mathrm{Ent}_\mu(\nu). \tag{7}$$

(ii) *If μ is of the form $\mu_1 \otimes \mu'$, where μ' is a measure on the product of the remaining lines, then there exists a Borel triangular mapping T such that $T(x) - x \in H$ and $\nu = \mu \circ T^{-1}$.*

(iii) *If μ is equivalent to μ_{e_1}, where $e_1 = (1, 0, 0, \dots)$, then there exists a Borel mapping T of type* (1) *such that $\nu = \mu \circ T^{-1}$.*

Proof. (i) We observe that $\nu_n \ll \mu_n$ and $f_n := d\nu_n/d\mu_n$ is the conditional expectation of $f := d\nu/d\mu$ with respect to the σ-field $\mathcal{F}_n$ generated by P_n. By Jensen's inequality, $\mathrm{Ent}_{\mu_n}(\nu_n) \le \mathrm{Ent}_\mu(\nu)$. As explained above, letting $T^{(n)} := T_{\mu_n,\nu_n}$, we have

$$\int_{\mathbb{R}^n} |T^{(n)}(x) - x|^2\, \mu_n(dx) \le \frac{2}{C}\mathrm{Ent}_{\mu_n}(\nu_n) \le \frac{2}{C}\mathrm{Ent}_\mu(\nu).$$

It follows by the uniqueness property for the natural triangular mapping that one has $P_n \circ T^{(n+1)} = T^{(n)}$ on $\mathbb{R}^n$. Therefore, letting $T_k^{(n)}$ be the k-th component of $T^{(n)}$, we see that $T_k^{(n)} = T_k^{(m)} = T_k^{(k)}$ whenever $n, m \ge k$. So we get at once the desired infinite dimensional mapping T: its k-th component is just $T_k^{(k)}$. In order to see that $\mu \circ T^{-1} = \nu$, it suffices to verify the equality $\nu(B) = \mu(T^{-1}(B))$ on all Borel cylindrical sets of the form $B = B_0 \times \mathbb{R} \times \mathbb{R} \times \cdots$ with $B_0 \in \mathcal{B}(\mathbb{R}^n)$. This equality is obvious from the equality $\mu_n \circ (T^{(n)})^{-1} = \nu_n$ and our definition of T. Finally, (7) follows by Fatou's theorem.

(ii) Let us consider the case where $\mu = \mu_1 \otimes \mu'$ with some Borel probability measure μ' on the space Y that corresponds to the product with respect to the coordinates x_k with $k \ge 2$, i.e., we write $X = \mathbb{R}^1 \times Y$. Let us partition $\mathbb{R}^\infty$ into disjoint sets E_n of positive measure such that f is bounded on each E_n. For example, it suffices to pick all positive measure sets among $\{k \le f < k+1\}$. Next we partition the real line into intervals D_n with $\mu_1(D_n) = \nu(E_n)$. The measure $I_{D_n} \cdot \mu_1$ can be transformed into $\nu(E_n)\mu_1$ by an increasing mapping Ψ_n on the interior of D_n. Then the mapping $x \mapsto (\Psi_n(x_1), x_2, x_3, \ldots)$ takes the measure $(I_{D_n} \cdot \mu_1) \otimes \mu'$ to $\nu(E_n)\mu$. By the above, there exists a triangular mapping Λ_n with $\Lambda_n(x) - x \in H$ such that $\nu(E_n)\mu \circ \Lambda_n^{-1} = I_{E_n} \cdot \nu$. Hence we obtain the triangular mapping T that takes the measure $\mu = \sum_{n=1}^\infty (I_{D_n} \cdot \mu_1) \otimes \mu'$ to the measure $\sum_{n=1}^\infty I_{E_n} \cdot \nu = \nu$. On $D_n \times \mathbb{R}^1 \times \mathbb{R}^1 \times \cdots$ the mapping T is the composition of Ψ_n and Λ_n. Since Ψ_n has the indicated form, we obtain $x - T(x) \in H$.

(iii) In the general case, where μ may not be represented as a product measure, it suffices to find an auxiliary measure μ_0 of the form $\mu_0 = \mu_1 \otimes \mu'$, where μ' is a measure on Y, such that μ can be transformed into μ_0 by a mapping T_0 and μ_0 can be transformed into ν by a mapping T' such that T_0 and T' are Borel mappings of type (1). Let μ' be the projection of μ on Y. Then it is readily verified that μ' is uniformly convex with constant C. Indeed, it suffices to show that if $\mu = \exp(-\Phi)\, dx$ is uniformly convex on $\mathbb{R}^n$ with constant C, then its projection μ_{n-1} to $\mathbb{R}^{n-1}$ is of the same type. Let $\varepsilon > 0$ and $\Phi_\varepsilon(y, x_n) = \Phi(y, x_n) - C(1-\varepsilon)|y|^2/2$, $y \in \mathbb{R}^{n-1}$. We observe that Φ_ε is convex and integrable. Then

$$\int \Phi(y, x_n)\, dx_n = \exp\Bigl(-\frac{1}{2}C(1-\varepsilon)|y|^2\Bigr) \int \Phi_\varepsilon(y, x_n)\, dx_n.$$

By Prékopa's theorem (see, e.g., Corollary 1.8.3 in Bogachev [3]), the integral on the right has the form $\exp[-W(y)]$ with some convex function W. Hence $\mu_{n-1} = \exp(-\Phi_{n-1})$, where $D^2\Phi_{n-1} \ge (C - \varepsilon) \cdot I$. Letting $\varepsilon \to 0$, we prove our claim. If we show that the conditional measures μ^y on the straight lines $L_y := y + \mathbb{R}^1 e_1$, where $e_1 = (1, 0, 0, \ldots)$, have densities ϱ_y with respect to the natural Lebesgue measure on L_y, then we can transform μ into μ_0 by a Borel mapping $T_0 = (T_{0,k})$ such that $T_{0,k}(x) = x_k$ for all $k \ge 2$. To this end, it suffices to transform the measure with density ϱ_y into μ_1 by an increasing function θ_y and set $T_{0,1}(x) = \theta_y(x_1)$, where $x = (x_1, y)$. Clearly, T_0 is not triangular, but has form (1). Finally, densities ϱ_y exist indeed, because according to Borell [6] the conditional measures μ^y on L_y are convex, hence μ^y is either absolutely continuous or Dirac's measure at some point a_y. It follows from our assumption that μ^y cannot be Dirac's measure for all y from a set of positive μ'-measure. □

We observe that according to Example 1 we have $\mu_{e_1} \sim \mu$ under the additional assumption that $D^2\Phi_n(x) \leq M \cdot I$ on $\mathbb{R}^n$ for some common constant M.

Part (ii) applies to the case considered in Bogachev, Kolesnikov, Medvedev [5] where μ is the product of countably many copies of a probability measure σ on the real line that is uniformly convex with constant $C > 0$. In this case, for any Borel probability measure ν absolutely continuous with respect to μ, there exists a Borel triangular mapping T such that $\nu = \mu \circ T^{-1}$. For example, μ may be the standard Gaussian product-measure.

A mapping that is increasing in the above sense may not be monotone in the sense of Feyel, Üstünel [10] and Feyel, Üstünel [12]. Also, a triangular mapping is typically not a gradient-type mapping. So mappings considered here are quite different from those employed in Brenier [7], Cordero-Erausquin [8], Feyel, Üstünel [12], McCann [17] and many other works on optimal transport. Note that in the Gaussian case, the last part of the above proof can be repeated for the gradient-type mappings constructed in Feyel, Üstünel [12], which will produce a resulting mapping of type (1), although not necessarily of gradient-type. However, it remains unclear whether one can find a monotone transformation without extra assumptions on ν. We also do not know whether T in (iii) can be always chosen increasing triangular or at least triangular if we insist on (1). Also in (ii) it is not clear whether one can find an increasing mapping (again with (1)). As noted above, there are no problems if we drop (1). We shall now see, however, that (1) is necessary if T is increasing triangular and $\mu \circ T^{-1}$ has finite entropy. Finally, we shall show that if in the infinite dimensional case the analog of the right hand side of (3) is finite, then the measure $\mu \circ T^{-1}$ is absolutely continuous with respect to μ and has finite entropy such that (3) holds true. The next assertion is seen directly from the proof of Theorem 2.

Proposition 3. *Let a Borel probability measure μ on X be uniformly convex with constant $C > 0$ and let T be a μ-measurable increasing triangular mapping such that $\nu := \mu \circ T^{-1}$ is absolutely continuous with respect to μ and has finite entropy. Then $T(x) - x \in H$ for μ-a.e. x and one has inequality* (2).

If T is an increasing triangular mapping on X such that its components T_n are absolutely continuous in x_n (i.e., the functions $x_n \mapsto T_n(x_1, \ldots, x_n)$ are absolutely continuous on every compact interval if $(x_1, \ldots, x_{n-1})$ is fixed), then $\det_2 DT(x)$ is defined by

$$\det_2 DT(x) := \prod_{n=1}^{\infty} \partial_{x_n} T_n(x) \exp\bigl(1 - \partial_{x_n} T_n(x)\bigr).$$

Observe that it is not assumed that DT makes sense separately. In particular, with such a definition, one has $\det_2 DT(x) = 1$ if $T_n(x) = x_n + F_n(x_1, \ldots, x_{n-1})$ whatever is F_n.

Theorem 4. *Let a Borel probability measure μ on X be uniformly convex with constant $C>0$ and let T be a μ-measurable increasing triangular mapping such that every T_n is absolutely continuous in x_n and the quantities on the right-hand side of (3) for μ_n and $P_n\circ T$ are uniformly bounded in n. Then $\nu:=\mu\circ T^{-1}$ is absolutely continuous with respect to μ and has finite entropy and $\mathrm{Ent}_\mu(\nu)=\lim\limits_{n\to\infty}\mathrm{Ent}_{\mu_n}(\nu_n)$.*

Proof. Let us keep the notation of the proof of Theorem 2. If we show that $\sup\limits_n \mathrm{Ent}_{\mu_n}(\nu_n)<\infty$, then the sequence $f_n=d\nu_n/d\mu_n$, which is a martingale, is uniformly integrable. Hence $\nu\ll\mu$. The equality $\mathrm{Ent}_\mu(\nu)=\lim\limits_{n\to\infty}\mathrm{Ent}_{\mu_n}(\nu_n)$ follows from the fact that $\mathrm{Ent}_\mu(\nu)\le\liminf\mathrm{Ent}_{\mu_n}(\nu_n)$ by Fatou's theorem and lower boundedness of $f_n\log f_n$ combined with the fact that $\mathrm{Ent}_{\mu_n}(\nu_n)=\mathrm{Ent}_\mu(f_n)\le\mathrm{Ent}_\mu(f)$ by the properties of conditional expectations. Now it suffices to consider the case of $\mathbb{R}^n$ and show that if the right-hand side of (3) is finite, then $\mathrm{Ent}_\mu(\nu)$ is finite as well (then (3) holds). This is not completely obvious. So we deal with $T=(T_1,\dots,T_n)$ and omit indices at μ and ν. One can find a sequence of smooth increasing triangular mappings $S_j=(S_{j,1},\dots,S_{j,n})$ on $\mathbb{R}^n$ with the following properties: the function $S_{j,k}-x_k$ are compactly supported on $\mathbb{R}^k$, $S_j(x)\to T(x)$ a.e., and

$$\sup_n\int\Bigl[\Bigl(\Lambda\bigl(e^{-\Phi},S_j(x),x\bigr)F_j(x),F_j(x)\Bigr)-\log\det{}_2 DS_j(x)\Bigr]\mu(dx)<\infty,$$

where $F_j:=S_j-I$. Observe that one cannot expect to find S_j such that S_j-I would be of compact support on $\mathbb{R}^n$, because the k-th component depends only on $x_1,\dots,x_k$, hence can be of compact support only if it is identically zero. It follows that $\mu\circ S_j^{-1}$ has a positive density f_j with respect to μ such that $f_j=1$ outside of some cube K_j. It is clear that $\mu\circ S_j^{-1}$ has finite entropy η_j satisfying equality (3). Therefore, $\eta:=\sup\limits_j\eta_j<\infty$. Since the measures $\mu\circ S_j^{-1}$ converge weakly to $\mu\circ T^{-1}$, we obtain that $\mu\circ T^{-1}\ll\mu$ and $\mathrm{Ent}_\mu(\nu)\le\eta$. Indeed, by the Kómlos theorem (see, e.g., §4.7 in Bogachev [3]), there is a subsequence $\{f_{n_j}\}$ such that the functions $\varrho_k:=(f_{n_1}+\cdots+f_{n_k})/k$ converge a.e. to some function $f\in L^1(\mu)$. By convexity one has $\mathrm{Ent}_\mu(\varrho_k)\le\eta$. Hence $\varrho_k\to f$ in $L^1(\mu)$. By Fatou's theorem $\mathrm{Ent}_\mu(f)\le\eta$. Since $\varrho_k\cdot\mu\to\mu\circ T^{-1}$ weakly, one has $\mu\circ T^{-1}=f\cdot\mu$. □

Example 5. *Let a Borel probability measure μ on X be uniformly convex with constant C and let T be a μ-measurable increasing triangular mapping such that every T_n is absolutely continuous in x_n. Assume, in addition, that for the densities $\exp(-\Phi_n)$ of the projections μ_n we have $D^2\Phi_n(x)\le M\cdot I$ on $\mathbb{R}^n$. Then the integrability of $|T(x)-x|_H^2$ and $\log\det_2 DT(x)$ with respect to μ implies that $\mu\circ T^{-1}\ll\mu$ and $\mathrm{Ent}_\mu(\mu\circ T^{-1})<\infty$. In particular, this applies to the standard Gaussian product-measure.*

Let us give an interesting application to the problem of absolute continuity of induced measures in the case of not necessarily triangular mappings. Let γ be a centered Radon Gaussian measure on a locally convex space E and let $H:=H(\gamma)$ be its Cameron–Martin space. Suppose that $\{e_n\}$ is an orthonormal basis in H, $\widehat{e_n}$ the measurable linear functional associated with e_n, and let $\mathrm{E}^{\mathcal{F}_n}$ denote the conditional expectation with respect to the σ-field $\mathcal{F}_n$ generated by $\widehat{e_i}$, $i\le n$. One can assume without loss of generality that we deal with the standard Gaussian product-measure and the standard basis in l^2 (then $\widehat{e_n}$ is the n-th coordinate function). The Hilbert–Schmidt norm of an operator A on H is denoted by $\|A\|_{\mathcal{H}}$ and $\|A\|_{L(H)}$ is its operator norm. The space of all Hilbert–Schmidt operators on H is denoted by $\mathcal{H}$. We use the standard notation for the Sobolev class $W^{2,1}(\gamma, H)$ of mappings $F\colon E \to H$ from $L^2(\gamma, H)$ such that the derivative $D_{H}F$ of F along H belongs to $L^2(\gamma, \mathcal{H})$.

Theorem 6. *Let $T = I + F$ with $F \in W^{2,1}(\gamma, H)$ be such that*

$$\Big(\big(I + D_{H}F(x)\big)h, h\Big)_{H} \geq \eta(x)|h|_{H}^2 \quad \text{for all } h \in H,$$

where η is an a.e. positive bounded measurable function. Suppose that there is a measurable function θ such that $\big(\mathrm{E}^{\mathcal{F}_n}\eta\big)^{-1} \leq \theta$ a.e. for all n and $\theta^4\|D_{H}F\|_{\mathcal{H}}^2 \in L^1(\gamma)$. Then $\nu := \gamma\circ T^{-1} \ll \gamma$, $\mathrm{Ent}_\gamma(\nu) < \infty$, and

$$\mathrm{Ent}_\gamma(\nu) \leq \frac{1}{2}\int \big[|F|_{H}^2 + 2\|D_{H}F\|_{\mathcal{H}}^2 + \theta^4\|D_{H}F\|_{\mathcal{H}}^2\big]\, d\gamma. \tag{8}$$

If $\eta(x) = \alpha > 0$ is constant and $\|D_{H}F\|_{L(H)} \in \bigcap_{p\geq 1} L^p(\gamma)$, e.g., if $\|D_{H}F\|_{L(H)} \leq q < 1$, where q is a constant, then there exist a Borel mapping S on X and a set Ω of full ν-measure such that $S(T(x)) = x$ for γ-a.e. x, $T(S(y)) = y$ for ν-a.e. y and on the set Ω one has

$$\frac{d\nu}{d\gamma} = \frac{1}{\Lambda_F\circ S}, \quad \Lambda_F := \det{}_2(I + D_{H}F)\exp\Big(\delta F - \frac{1}{2}|F|_{H}^2\Big), \tag{9}$$

and $d\nu/d\gamma = 0$ on the complement of Ω. In addition, one has

$$\mathrm{Ent}_\gamma(\nu) = \frac{1}{2}\int |F|_{H}^2\, d\gamma - \int \log\det{}_2(I + D_{H}F)\, d\gamma. \tag{10}$$

Proof. For any operators A and B on $\mathbb{R}^n$ one has (see, e.g., [3] or [24])

$$\det{}_2(I + A)\det{}_2(I + B) = \det{}_2\big((I + A)(I + B)\big)\exp\mathrm{trace}(AB).$$

If $I + A$ is invertible, then letting $B = -A(I + A)^{-1}$ we have $I + B = (I + A)^{-1}$. By Carleman's inequality $\det_2(I + B) \leq \exp\big(\|B\|_{\mathcal{H}}^2/2\big)$, we obtain

$$|\det{}_2(I + A)|^{-1} \leq \exp\Big(\frac{1}{2}\|A(I + A)^{-1}\|_{\mathcal{H}}^2 + \mathrm{trace}\big(A^2(I + A)^{-1}\big)\Big).$$

Since $\|AC\|_{\mathcal{H}} \leq \|C\|_{L(H)}\|A\|_{\mathcal{H}}$ and $\mathrm{trace}(A^2C) \leq \|C\|_{L(H)}\|A\|_{\mathcal{H}}$, one has

$$-\log|\det{}_2(I + A)| \leq \Big(\|(I + A)^{-1}\|_{L(H)}^2 + \frac{1}{2}\Big)\|A\|_{\mathcal{H}}^2.$$

This estimate remains true for any Hilbert–Schmidt operators A and B whenever $I + A$ is invertible. If T is an invertible C^1-mapping on $\mathbb{R}^n$ such that

$$\big\|\big(I + DF(x)\big)^{-1}\big\|_{L(H)} \leq \theta(x),$$

we obtain (8). In order to justify (8) in the finite dimensional case under our more general assumptions, we consider the mapping $T_\varepsilon = I + (1 - \varepsilon)F$, $\varepsilon > 0$. Note that $\big(DT_\varepsilon(x)h, h\big)_{H} \geq \big(\varepsilon + \eta(x)\big)|h|_{H}^2$, hence $\|DT_\varepsilon(x)^{-1}\|_{L(H)} \leq \big(\varepsilon + \eta(x)\big)^{-1}$. It suffices to show (8) for T_ε (with the objects corresponding to T_ε, of course), because letting $\varepsilon \to 0$, we arrive at (8) for T. Let us consider the smooth mapping $T_{\varepsilon,k} := P_{1/k}T_\varepsilon$, where $\{P_t\}$ is the Ornstein–Uhlenbeck semigroup. One has the estimate $\big(DT_{\varepsilon,k}(x)h, h\big)_{H} \geq \varepsilon|h|_{H}^2$. Hence $T_{\varepsilon,k}$ is invertible and $\|DT_{\varepsilon,k}(x)^{-1}\|_{L(H)} \leq$

$\big(\varepsilon+P_{1/k}\eta(x)\big)^{-1}$. Therefore, we obtain (8) in the case of $T_{\varepsilon,k}$. Letting $k\to\infty$ we arrive at (8) for T_ε. The general case reduces to the case of the standard Gaussian product measure, because if $H(\gamma)$ is infinite dimensional, then there is a measure preserving linear isomorphism J between (E,γ) and (X,μ), where μ is the standard Gaussian product-measure, such that J is an isometry between $H(\gamma)$ and H (see Chapter 6 in Bogachev [3]). Let $F_n:=\mathrm{E}^{\mathcal{F}_n}P_nF$. Then $D_HF_n=\mathrm{E}^{\mathcal{F}_n}P_nD_HF$. For any $h\in P_n(H)$ with $|h|_H=1$ we have

$$\big((I+D_HF_n)h,h\big)_H=\mathrm{E}^{\mathcal{F}_n}\big((I+P_nD_HF)h,h\big)_H=\mathrm{E}^{\mathcal{F}_n}\big((I+D_HF)h,h\big)_H\ge \mathrm{E}^{\mathcal{F}_n}\eta.$$

It follows that $\|(I+D_HF_n)^{-1}\|_{L(H)}\le\big(\mathrm{E}^{\mathcal{F}_n}\eta\big)^{-1}\le\theta$ a.e. Therefore,

$$\begin{aligned}\mathrm{Ent}_{\gamma_n}(\nu_n)&\le\frac12\int|F_n|_H^2\,d\gamma_n+\int\|D_HF_n\|^2_{\mathcal{H}}\big[\frac12+\big(\mathrm{E}^{\mathcal{F}_n}\eta\big)^{-2}\big]\,d\gamma_n\\&=\frac12\int|F_n|_H^2\,d\gamma+\frac12\int\|D_HF_n\|^2_{\mathcal{H}}\,d\gamma+\int(D_HF,D_HF_n)_{\mathcal{H}}\big(\mathrm{E}^{\mathcal{F}_n}\eta\big)^{-2}\,d\gamma_n\\&\le\frac12\int|F|_H^2\,d\gamma+\frac12\int\|D_HF\|^2_{\mathcal{H}}+\Big(\int\|D_HF\|^2_{\mathcal{H}}\theta^4\,d\gamma\Big)^{1/2}\Big(\int\|D_HF\|^2_{\mathcal{H}}\,d\gamma\Big)^{1/2}\\&\le\frac12\int|F|_H^2\,d\gamma+\int\|D_HF\|^2_{\mathcal{H}}\,d\gamma+\frac12\int\|D_HF\|^2_{\mathcal{H}}\theta^4\,d\gamma.\end{aligned}$$

This yields that $\nu\ll\gamma$ and (8) holds.

Let us verify (9) in the case where $\eta(x)=\alpha>0$ and $\|D_HF\|_{L(H)}$ belongs to all $L^p(\gamma)$. Suppose first that F takes values in $\mathbb{R}^n$ and that

$$\big((I+DF_H)h,h\big)_H\ge\alpha|h|_H^2\quad\text{for all } h\in\mathbb{R}^n. \tag{11}$$

The claim reduces to the mappings $T_y:=z+F_y(z)$, $F_y(z):=F(z,y)$, on $\mathbb{R}^n$ if we write $x=(z,y)$, $z\in\mathbb{R}^n$. It is known (see, e.g., Chapter 5 in Bogachev [3]) that F_y belongs to the class $W^{2,1}(\gamma_n,\mathbb{R}^n)$ for γ_0-a.e. y, where we write γ as the product of the standard Gaussian measure γ_n on $\mathbb{R}^n$ and the standard Gaussian-product measure γ_0 on the product of the remaining lines. In addition, one has (11) and $\|D_HF_y\|_{L(H)}\in\bigcap_{p>1}L^p(\gamma_n)$ for γ_0-a.e. y. We can deal with a version of F such that F_y possesses these properties for all y. Now it suffices to consider the case of a single mapping F on $\mathbb{R}^n$ that belongs to $W^{2,1}(\gamma_n,\mathbb{R}^n)$ with the operator norm of DF in all $L^p(\gamma_n)$ and satisfies (11) and apply this to the mappings F_y. By the Sobolev embedding theorem F has a continuous modification denoted also by F. Since one has $\big(T(x+h)-T(x),h\big)\ge\alpha|h|^2$, we conclude by Minty's theorem that T is a homeomorphism of $\mathbb{R}^n$. Moreover, $S:=T^{-1}$ is Lipschitzian with constant α^{-1}. This ensures formula (5) for the Radon–Nikodym density. Thus we obtain a mapping G with values in $\mathbb{R}^n$ such that for any y, the mapping $S_y\colon z\mapsto z+G(z,y)$ is Lipschitzian with constant α^{-1} and $S_y(T_y(z))=T_y(S_y(z))=z$ on $\mathbb{R}^n$. Note that in this case $\gamma\circ T^{-1}\sim\gamma$, since $\gamma_n\circ T_y^{-1}\sim\gamma_n$.

The next step is to take finite dimensional approximations $F_n:=P_nF$. The mapping F_n takes values in $\mathbb{R}^n$ and belongs to the class $W^{2,1}(\gamma,\mathbb{R}^n)$. In addition, (11) is fulfilled. Hence we obtain (9) for each F_n. Set $T_n:=I+F_n$, $S_n:=I+G_n:=(I+F_n)^{-1}$. We have $\nu=f\cdot\gamma$, $\nu_n:=\gamma\circ T_n^{-1}=f_n\cdot\gamma$. The sequence $\{f_n\}$ is a uniformly integrable martingale with respect to γ. Hence it converges to f in $L^1(\gamma)$. We observe that $|G_k\circ T_k-G_k\circ T_n|_H\le(1+\alpha^{-1})|F_k-F_n|_H$, since the mapping $I+G_k$ is Lipschitzian along $\mathbb{R}^k$ with constant α^{-1}. Let us show that $\int \min(1, |G_n - G_k|_H)\, d\nu \to 0$ as $n,k\to\infty$. It suffices to show that $\int \min(1, |G_n - G_k|_H)\, d\nu_n \to 0$ as $n,k\to\infty$, since $\|\nu-\nu_n\|\to 0$. Let $k>n$. Then

$$\int \min(1, |G_n - G_k|_H)\, d\nu_n = \int \min(1, |G_n\circ T_n - G_k\circ T_n|_H)\, d\gamma.$$

Since

$$\begin{aligned}|G_n\circ T_n - G_k\circ T_n|_H &\le |G_n\circ T_n - G_k\circ T_k|_H + |G_k\circ T_k - G_k\circ T_n|_H \\ &\le |F_n - F_k|_H + (1+\alpha^{-1})|F_n - F_k|_H,\end{aligned}$$

our claim follows by pointwise convergence of $|F_n - F_k|_H$ to 0. Hence there exists an H-valued mapping G such that $|G_n - G|_H \to 0$ in ν-measure. Passing to a subsequence we may assume that $|G_n - G|_H \to 0$ ν-a.e. Clearly, we can take a Borel version of G. Set $S := I + G$. We have $S_n \to S$ ν a.e. and $F_n(S_n(y)) = -G_n(y)$. This yields that $F(S(y)) = -G(y)$ ν-a.e. Indeed, since $F_n\to F$ pointwise, it suffices to show that the sequence of measures $\nu\circ S_n^{-1}$ is uniformly countably additive. If this is done, then, given $\varepsilon>0$ there is a compact set K and a number N_ε such that $F|_K$ is continuous, $\sup\limits_{x\in K}|F_n(x)-F(x)|_H \le \varepsilon$ for all $n\ge N_\varepsilon$ and $\nu\circ S^{-1}(K) > 1-\varepsilon$, $\nu\circ S_n^{-1}(K)>1-\varepsilon$, which is possible due to the uniform countable additivity. We recall that taking any measure ν' with respect to which all $\nu\circ S_n^{-1}$ are absolutely continuous, one can find $\delta>0$ such that $\nu\circ S_n^{-1}(B)<\varepsilon$ whenever $\nu'(B)<\delta$. In fact, one can take $\nu'=\gamma$, because $\nu\ll\gamma$ and $\gamma\circ S_n^{-1}\ll\gamma$. One can extend $F|_K$ to a continuous H-valued mapping F_ε on all of X. Then, by convergence $F_\varepsilon(S_n)\to F_\varepsilon(S)$ ν-a.e., we obtain $N>N_\varepsilon$ such that $\nu\bigl(|F_\varepsilon(S)-F_\varepsilon(S_n)|_H>\varepsilon\bigr)<\varepsilon$ for all $n\ge N$. This yields $\nu\bigl(|F(S)-F(S_n)|_H>\varepsilon\bigr)<3\varepsilon$ for all $n\ge N$. Combining this estimate with $\nu\bigl(|F(S_n)-F_n(S_n)|>\varepsilon\bigr)\le \nu\circ S_n^{-1}(X\backslash K)\le\varepsilon$, we obtain convergence in measure. Let us establish the uniform countable additivity of the measures $\nu\circ S_n^{-1}$. It suffices to show that given a sequence of Borel sets $B_k\downarrow\emptyset$ and $\varepsilon>0$, there is k_1 such that $\nu\bigl(S_n^{-1}(B_{k_1})\bigr)\le\varepsilon$ for all n. We pick N such that $\|\nu_n-\nu\|<\varepsilon/2$ for all $n\ge N$. Next we pick k_1 such that $\gamma(B_{k_1})<\varepsilon/2$ and $\nu\bigl(S_n^{-1}(B_{k_1})\bigr)\le\varepsilon$ for each $n\le N$. If $n\ge N$, we obtain

$$\nu\bigl(S_n^{-1}(B_{k_1})\bigr)\le \nu_n\bigl(S_n^{-1}(B_{k_1})\bigr)+\|\nu-\nu_n\| = \gamma(B_{k_1})+\|\nu-\nu_n\|\le\varepsilon.$$

Thus we have $T(S(y)) = y$ for ν-a.e. y. In a similar manner one verifies that $S(T(x)) = x$ for γ-a.e. x. Namely, $|G_n\circ T - G\circ T|_H\to 0$ γ-a.e. by the above construction. Since $\|\gamma\circ T_n^{-1}-\gamma\circ T^{-1}\|\to 0$, this yields that $|G_n\circ T_n - G\circ T_n|_H\to 0$ in measure with respect γ. As above we have also $|G\circ T_n - G\circ T|_H\to 0$ in measure with respect to γ. Hence $|G_n\circ T_n - G\circ T|_H\to 0$ in γ-measure, whence by the identity $G_n\circ T_n = -F_n$ we obtain $-F = G\circ T$ γ-a.e. By norm convergence of F_n to F in $W^{2,1}(\gamma,H)$ we have convergence in measure of the functions $\det_2(I+D_H F_n)$, δF_n and $|F_n|_H$ to the corresponding expressions for F, i.e., $\Lambda_{F_n}\to\Lambda_F$ in measure with respect to γ. Passing to a subsequence we may assume that all these sequences converge γ-a.e. Since $f_n = 1/\Lambda_{F_n}\circ S_n\to f$ a.e. and the functions $\Lambda_{F_n}\circ S_n$ converge to $\Lambda_F\circ S$ in ν-measure by the above argument, we arrive at the desired expression for the Radon–Nikodym density of ν that holds on some set of full ν-measure (it is not clear whether ν is equivalent to γ). Finally, (10) follows from (9) as explained above in the derivation of (4), but can be also deduced from the finite dimensional case along the same lines. □

Note that the condition $\|D_H F\|_{L(H)} \le q$ is equivalent to the existence of a version of F such that $|F(x+h)-F(x)|_H \le q|h|_H$ for all $x\in E$ and $h\in H$ (see Section 5.8 in Bogachev [3]). It also suffices to have the estimate $\|D_H F(x)\|_{L(H)} \le q(x) < 1$ provided that $\|D_H F\|^2_{\mathcal{H}}(1-q)^{-4}$ is integrable.

This theorem extends the results of Üstünel, Zakai [23], Üstünel, Zakai [24] on the membership of the Radon–Nikodym density of the induced measure in the class $L\log L$ proved under the additional assumption that $\exp\Bigl(c|D_H F|^2_{\mathcal{H}}\Bigr) \in L^1(\gamma)$ for some $c > (2+\alpha)/(2\alpha)$. It is worth noting that our estimates enable one to give a shorter justification of the equivalence of ν and γ (i.e., to show that Ω is a set of full γ-measure) under this additional assumption, which was established in Üstünel, Zakai [23] (see also Üstünel, Zakai [24]).

A closer look at the proof of the previous theorem shows that it applies also to some other measures. For example, we have the following result, where the Sobolev class $W^{2,1}(\mu,H)$ is defined exactly as in the Gaussian case.

Theorem 7. *Let μ be the product of countably many copies of a probability measure σ on the real line with a density $\exp(-\Phi)$ such that $0 < C \le \Phi''(t) \le M$. Let $F\in W^{2,1}(\mu,H)$ satisfy the same conditions as in the previous theorem. Then one has $\mu\circ T^{-1} \ll \mu$ and $\mathrm{Ent}_\mu(\mu\circ T^{-1})$ is finite.*

Remark 8. It is worth noting that in the case where μ is the product of countably many copies of a strictly convex measure σ on $\mathbb{R}^1$ (e.g., of a nondegenerate Gaussian measure) and ν is a probability measure absolutely continuous with respect to μ, the existence of a triangular mapping T of type (1) with $\nu = \mu\circ T^{-1}$ can be established by using the inductive argument of Talagrand [21]. For the reader's convenience, we provide the details. The first step is to obtain (2) in $\mathbb{R}^n$ by induction on n as explained in Talagrand [21] in the Gaussian case. In fact, the only nontrivial thing is to justify the case $n=1$. It is enough to consider measures with smooth densities, which along with the condition $n=1$ makes the proof quite easy (see Bogachev, Kolesnikov, Medvedev [5]). Next, assume that (2) is true for some $n\ge 1$ and consider the product μ of $n+1$ copies of σ. Points in $\mathbb{R}^{n+1}$ will be written as $x=(P_n x, x_{n+1})$. Let $\nu = f\cdot\mu$. For $y\in\mathbb{R}^n$ set

$$g(y) := \int_{\mathbb{R}^1} f(y,t)\,\sigma(dt), \quad f_y(t) := f(y,t)/g(y),$$

where $f_y(t)=0$ for all t whenever $g(y)=0$. Then

$$\int_{\mathbb{R}^{n+1}} f\log f\,d\mu = \int_{\mathbb{R}^n} g(y)\log g(y)\,\mu_n(dy) + \int_{\mathbb{R}^n}\Bigl(\int_{\mathbb{R}} f_y(t)\log f_y(t)\,\sigma(dt)\Bigr) g(y)\,\mu_n(dy).$$

Let $T := T_{\mu,\nu} = (T_1, \dots, T_{n+1})$. Then $T_{\mu_n,\nu_n} = P_n T$ on $\mathbb{R}^n$. For every $y \in \mathbb{R}^n$ the function $t \mapsto T_{n+1}(y,t)$ is increasing and takes σ to the measure $f_{P_nT(y)} \cdot \sigma$. We have

$$\int_{\mathbb{R}^n} \Bigl(\int_{\mathbb{R}} f_y(t) \log f_y(t)\, \sigma(dt) \Bigr) g(y)\, \mu_n(dy) = \int_{\mathbb{R}^n} \Bigl(\int_{\mathbb{R}} f_{P_nT(y)}(t) \log f_{P_nT(y)}(t)\, \sigma(dt) \Bigr)\, \mu_n(dy).$$

For every fixed y, by the case $n = 1$ we have

$$\int_{\mathbb{R}} f_{P_nT(y)}(t) \log f_{P_nT(y)}(t)\, \sigma(dt) = \int_{\mathbb{R}} \log f_{P_nT(y)}(t)\, \sigma \circ T_{n+1}(y,\cdot)^{-1}(dt) \geq \frac{C}{2} \int_{\mathbb{R}} \bigl| T_{n+1}(y,t) - t \bigr|^2 \sigma(dt).$$

By the inductive assumption we obtain

$$\int_{\mathbb{R}^{n+1}} f \log f\, d\mu \geq \frac{C}{2} \int_{\mathbb{R}^n} \bigl| PT_n(y) - y \bigr|^2 \mu_n(dy) + \frac{C}{2} \int_{\mathbb{R}^n} \int_{\mathbb{R}} \bigl| T_{n+1}(y,t) - t \bigr|^2 \sigma(dt)\, \mu_n(dy) = \frac{C}{2} \int_{\mathbb{R}^{n+1}} \bigl| T(z) - z \bigr|^2 \mu(dz).$$

For the standard Gaussian measure one has $C = 1$. As above, this immediately yields the claim in the case of $\mathbb{R}^\infty$ provided that ν has finite entropy. The last step is the same as in Theorem 2(ii).

Remark 9. We do not know which measures ν on X can be obtained as the images of the standard Gaussian product-measure μ under mappings T of type (1). It is easy to see that any such measure must vanish on all finite dimensional subspaces. Moreover, $\nu(L) = 0$ for any Borel linear space L such that $\mu(L + H) = 0$, in particular, $\nu(H) = 0$.

Finally, we observe that most of the above results can be formulated and proved along the same lines in a more abstract setting where X is a general locally convex space equipped with a measure μ, $H \subset X$ is a continuously embedded separable Hilbert space (an analog of the Cameron–Martin space for Gaussian measures), and the coordinate functions are replaced by suitable μ-measurable linear functionals. For example, a natural concept of a uniformly convex measure in this setting was studied in Albeverio, Kondratiev, Röckner [1] and Kulik [15]. To be more specific, assume that H is dense in X. Then one has the embedding $j_H \colon X^* \to H$, $\bigl(j_H(l), h\bigr)_H = l(h)$, $l \in X^*$, $h \in H$. Let $\{l_n\} \subset X^*$ be such that $\{e_n := j_H(l_n)\}$ is an orthonormal basis in H. Suppose that μ is a Radon probability measure on X such that the functionals l_n separate the points in a linear subspace E of full μ-measure. Then, taking E, $\{l_n\}$ and H in place of $\mathbb{R}^\infty$ with the coordinate functions and l^2, one has the analogs of the above results in this more general setting.

This work has been completed during a visit to the university of Bielefeld. We thank M. Röckner for useful discussions.

References

[1] Albeverio S., Kondratiev Yu.G., Röckner M. (1995) Dirichlet operators via stochastic analysis. J. Funct. Anal. **128** 102–138.

[2] Bobkov S.G. (2003) Large deviations via transfer plans. Preprint.
[3] Bogachev V.I. (1998) Gaussian measures. Amer. Math. Soc. Providence, Rhode Island.
[4] Bogachev V.I. (2003) Foundations of measure theory. V. 1,2. Regular and Chaotic Dynamics, Moscow (in Russian).
[5] Bogachev V.I., Kolesnikov A.V., Medvedev K.V. (2003) On triangular transformations of measures. Russian Math. Dokl. (to appear).
[6] Borell C. (1978) A note on conditional probabilities of a convex measure. Vector space measures and applications (Proc. Conf., Univ. Dublin, Dublin, 1977), II, pp. 68–72, Lecture Notes in Phys. **77**, Springer, Berlin – New York.
[7] Brenier Y. (1991) Polar factorization and monotone rearrangement of vector valued functions. Comm. Pure Appl. Math. **44** 375–417.
[8] Cordero-Erausquin D. (2002) Some applications of mass transport to Gaussian-type inequalities. Arch. Rat. Mech. Anal. **161** 257–269.
[9] Fernique X. (2003) Extension du théorème de Cameron–Martin aux translations aléatoires. Ann. Probab. **31** 1296–1304.
[10] Feyel D., Üstünel A.S. (2000) The notion of convexity and concavity on Wiener space. J. Funct. Anal. **176**, no. 2, 400–428.
[11] Feyel D., Üstünel A.S. (2002) Transport of measures on Wiener space and the Girsanov theorem. C. R. Acad. Sci. Paris. **334** 1025–1028.
[12] Feyel D., Üstünel A.S. (2003) Monge–Kantorovitch measure transportation and Monge–Ampère equation on Wiener space. Probab. Theor. Relat. Fields.
[13] Knote H. (1957) Contributions to the theory of convex bodies. Michigan Math. J. **4** 39–52.
[14] Kolesnikov A.V. (2003) Convexity inequalities and nonlinear transformations of measures. Russian Math. Dokl. (to appear).
[15] Kulik A.M. (2002) Log-Sobolev inequality, exponential integrability and large deviation estimates for $C(\alpha,\beta)$ log-concave measures. Random Oper. Stochastic Equations **10**, no. 2, 105–122.
[16] Ledoux M. (2001) The concentration of measure phenomenon. Amer. Math. Soc., Providence, Rhode Island.
[17] McCann R.J. (1995) Existence and uniqueness of monotone measure-preserving maps. Duke Math. J. **80** 309–323.
[18] Otto F., Villani C. (2000) Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173** 361–400.
[19] Rachev S.T., Rüschendorf L. (1998) Mass transportation problems. Vol. I. Springer-Verlag, New York.
[20] Shepp L. (1965) Distingunishing a sequence of random variables from a translate of itself. Ann. Math. Statist. **36** 1107–1112.
[21] Talagrand M. (1996) Transportation cost for Gaussian and other product measures. Geom. Funct. Anal. **6** 587–600.
[22] Üstünel A.S., Zakai M. (1996) Measures induced on Wiener space by monotone shifts. Probab. Theory Relat. Fields **105** 545–563.
[23] Üstünel A.S., Zakai M. (1998) On the uniform integrability of the Radon–Nikodym densities for Wiener measure. J. Funct. Anal. **159** 642–663.
[24] Üstünel A.S., Zakai M. (2000) Transformation of measure on Wiener space. Springer, Berlin.

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