ON CONTINUITY EQUATIONS IN INFINITE DIMENSIONS WITH NON-GAUSSIAN REFERENCE MEASURE

ALEXANDER V. KOLESNIKOV AND MICHAEL RÖCKNER

Abstract. Let γ be a Gaussian measure on a locally convex space and H be the corresponding Cameron-Martin space. It has been recently shown by L. Ambrosio and A. Figalli that the linear first-order PDE

$$\dot{\rho} + \mathrm{div}_{\gamma}(\rho \cdot b) = 0, \ \ \rho|_{t=0} = \rho_0,$$

where $\rho_0 \cdot \gamma$ is a probability measure, admits a weak solution, in particular, under the following assumptions:

$$\|b\|_H \in L^p(\gamma), \ p > 1, \quad \exp\bigl(\varepsilon(\mathrm{div}_{\gamma} b)_-\bigr) \in L^1(\gamma).$$

Applying transportation of measures via triangular maps we prove a similar result for a large class of non-Gaussian probability measures ν on $\mathbb{R}^\infty$, under the main assumption that $\beta_i \in \cap_{n\in\mathbb{N}} L^n(\nu)$ for every $i \in \mathbb{N}$, where β_i is the logarithmic derivative of ν along the coordinate x_i. We also show uniqueness of the solution for a wide class of measures. This class includes uniformly log-concave Gibbs measures and certain product measures.

1. Introduction

In this paper we study infinite-dimensional continuity equations

$$\dot{\mu} + \mathrm{div}(\mu) = 0, \ \mu_0 = \zeta, \tag{1}$$

where $\mu = \mu_t(dx)$, $t \geq 0$, is a curve of probability measures on $\mathbb{R}^\infty$ equipped with the product σ-algebra induced by the Borel σ-algebra on $\mathbb{R}$ and $b : \mathbb{R}^\infty \to \mathbb{R}^\infty$. Furthermore, $\dot{\mu} = \frac{\partial}{\partial t}\mu$, div is meant in the sense of distributions and ζ is a probability measure on $\mathbb{R}^\infty$ serving as the initial datum. One approach to solve equation (1) is to choose a reference measure ζ and search for solutions for (1) with $\zeta = \rho_0 \cdot \nu$ which are of the form $\mu_t(dx) = \rho(t,x) \cdot \nu(dx)$. When (1) can be written as

$$\dot{\rho} + \mathrm{div}_{\nu}(\rho \cdot b) = 0, \ \rho(0,x) = \rho_0, \tag{2}$$

where div_ν is the divergence with respect to ν, i.e. (-1) times the adjoint of the gradient operator on $L^2(\mathbb{R}^\infty, \nu)$. We stress that the choice of the reference measure (even in the finite-dimensional case, where $\mathbb{R}^\infty$ is replaced by $\mathbb{R}^d$) is at our disposal and should be made depending on b. For instance, in the finite-dimensional case b might be in a weighted Sobolev class with respect to some measure ν absolutely continuous with respect to Lebesgue measure, but not weakly differentiable with

Key words and phrases. continuity equation, renormalized solution, Gaussian measures, Gibbs measures, triangular mappings, Sobolev a-piori estimates.

This study was carried out within "The National Research University Higher School of Economics" Academic Fund Program in 2012-2013, research grant No. 11-01-0175. The first author was supported by RFBR projects 10-01-00518, 11-01-90421-Ukr-f-a. Both authors would like to thank the DFG for financial support through CRC 701.

respect to Lebesgue measure itself. Then one should take ν to be the measure for which the components of b are in $W^{1,2}(\nu)$. So, singularities of b will thus be compensated by the zeros of the Lebesgue density of ν.

Likewise in the infinite dimensional case of $\mathbb{R}^\infty$, where one usually takes a Gaussian measure as reference measure, since they are best studied. However, in many cases this is not the best choice, for similar reasons as we have just seen, in the case of $\mathbb{R}^d$. For instance, there are interesting examples (presented in Section 7.2 below), where the reference measure should be taken to be a Gibbs measure, whose energy functional can be "read off" the given map b which determines equation (1), respectively (2).

The key point, that for such reference measures we can identify conditions so that (2) has a solution and/or that this solution is unique, lies in the fact that many probability measures on $\mathbb{R}^d$ are images of Gaussian measures under so-called triangular mappings which turn out to have sufficient regularity in many concrete situations. Therefore, we can reduce existence and uniqueness questions (1), respectively (2) to the case of a Gaussian reference measure, studied in [7] and [20].

To explain this and also to review a bit the history of the problem, let us return to equation (2) and recall that the associated Lagrangian flow has the form

$$\dot{X}(t,x) = b(X),\ X(0,x) = x. \tag{3}$$

A finite-dimensional theory of equations (1) and (2) for weakly differentiable drifts b has been deeply developed in a recent series of papers by L. Ambrosio. G. Crippa, C. De Lellis, G. Savaré, A. Figalli and others (see [5] and the references therein). This theory works under quite general assumptions and includes, in particular, existence and uniqueness results for BV (bounded variation) vector fields.

Relatively little is known, however, in the infinite-dimensional setting. The first results in this direction have been obtained by A.B. Cruzeiro [19], V.I. Bogachev and E Mayer-Wolf [15]. The starting point for us was the paper [7], where some finite-dimensional techniques (including the Di Perna-Lyons theory of renormalized solutions) have been generalized to the infinite-dimensional Gaussian case. Other recent developments can be found in Ambrosio-Figalli [6], Le Bris-Lions [28], Fang-Luo [20], Bogachev-Da Prato-Shaposhnikov-Röckner [16].

We stress that the uniqueness of the solution is a more difficult problem compared to the existence. The latter can be established under quite broad assumptions (see, for instance, [16] for the apparently most general results about existence). The uniqueness proof obtained in [7] relies very strongly on the Gaussian framework. An important technical point was smoothing by the Ornstein-Uhlenbeck semigroup which behaves very nicely with respect to many natural operations on the Wiener space (divergence, projections, conditional expectations, differentiation etc.). The absence of such a nice smoothing operator seems to be the main difficulty when one tries to solve (2) for non-Gaussian reference measures.

In this paper we prove an existence result for the case of reference measures ν on $\mathbb{R}^\infty$ with logarithmic derivatives integrable in any power. We also show uniqueness for a wide class of product measures, including log-concave ones. Another uniqueness result is proved for a class of uniformly log-concave Gibbs measures.

Our approach relies on the mass transportation method. The general scheme works as follows. Instead of directly solving (2) we consider a mass transportation mapping $T : \mathbb{R}^\infty \to \mathbb{R}^\infty$ pushing forward the standard Gaussian measure γ onto

ν: $\nu = \gamma \circ T^{-1}$. If $\nu_t = \rho_t \cdot \nu$ is the solution to (2), then the family of measures $\gamma_t = \nu_t \circ S^{-1}$ with $S = T^{-1}$ solves the continuity equation for the new vector field

$$c = DT^{-1} \cdot b(T), \tag{4}$$

here D denotes total derivative. Applying (slightly generalized) existence and uniqueness results for the Gaussian case from [7], we get a solution γ_t of the equation associated to the vector field c and transfer it back, i.e. $\nu_t = \gamma_t \circ T^{-1}$.

The main advantage of this approach is that the divergence operator commutes with T:

$$\mathrm{div}_\gamma c = \bigl[\mathrm{div}_\nu b\bigr] \circ T.$$

Hence the crucial assumptions on $\mathrm{div}_\gamma c$ can be directly transferred to $\mathrm{div}_\nu b$. On the other hand, assumptions on integral norms of c and Dc impose some restrictions on Sobolev norms of T and $S = T^{-1}$. To prove the corresponding a-priori bounds is the main technical difficulty of our approach.

Note that we are free to choose any type of transportation mappings provided they have sufficient regularity. In this paper we deal with triangular mass transportation. A short discussion about the optimal transportation approach can be found in the very last section of this paper. The advantage of these mappings is their simple form. Even in the infinite-dimensional case they have essentially finite-dimensional structure. We obtain some Sobolev estimates on S and deduce from them the existence result for (2). The key estimate for triangular mappings applied in this paper looks as follows. Let $S = \sum_i S_i \cdot e_i$ be the triangular mapping pushing forward the measure ν onto the standard Gaussian measure γ. Then

$$\int \|\partial_{x_j} S\|_{l^2}^2 \ d\nu = \sum_{i \ge j} \int \bigl(\partial_{x_j} S_i\bigr)^2 \ d\nu \le \int \beta_j^2 \ d\nu.$$

Here β_j is the logarithmic derivative of ν along x_j. For more details on triangular mappings see [8].

The paper is organized as follows. In Section 2 we prove an extension of the results from [7]. In particular, we weaken some assumptions in [7] by introducing a slightly weaker notion of solution (see Remark 2.1). In Section 3 we establish Sobolev estimates for triangular mappings. In Section 4 we prove the key technical relations between transport equations and mass transfer. The existence result is proved in Sections 5. Sections 6-7 deal with the uniqueness in the product and Gibbsian case. In particular, we prove a uniqueness result for log-concave Gibbs measures with the following formal Hamiltonian

$$\sum_{i=1}^{\infty} V_i(x_i) + \sum_{i,j=1}^{\infty} W_{i,j}(x_i, x_k).$$

In Section 8 we briefly discuss the approach via optimal transportation mappings and the finite-dimensional case. In particular, we prove an existence and uniqueness theorem for a broad class of log-concave measures under "dimension-free" assumptions. Furthermore, in Example 8.5 we give an example in the finite-dimensional case, for which our result (see Theorem 8.4) implies existence and uniqueness for (2), where $b : \mathbb{R}^\infty \to \mathbb{R}^\infty$ is not BV (hence the results of [7], [4] are not applicable).

Notations: Throughout the paper p^* is the dual numbers to $p \in [1, \infty[$: $\frac{1}{p} + \frac{1}{p^*} = 1$. We denote by $\mathcal{F}_n$ the σ-algebra generated by the projection $P_n(x) = (x_1, \cdots, x_n)$ and by $\mathbb{E}_\nu^{\mathcal{F}_n}$ the corresponding conditional expectation. Everywhere below $\|\cdot\|$

means the standard l^2-norm (finite and infinite dimensional). We denote by ∇ and D^2 the derivatives of first and second order along $H = l_2$ respectively. For every linear operator $A : l_2 \to l_2$ the notation $\|A\|$ means the standard operator norm and $\|A\|_{HS} = \sqrt{\mathrm{Tr}(A^*A)}$ the Hilbert-Schmidt norm. The time derivative of a function f is denoted by $\dot{f}$. We fix the standard orthogonal basis in $\mathbb{R}^\infty$ consisting of vectors $e_i = (\delta_{ij})_{j\in\mathbb{N}}$. We use the word "positive" in the sense of "strictly positive" (i.e. "> 0"), otherwise we say "nonnegative"

2. The Gaussian case

In this paper we use the following core of smooth cylindrical functions: $\mathcal{C}$ is the linear span of all infinitely differentiable functions $\varphi(x_1, \cdots, x_n)$ depending on a finite number of coordinates and having a compact (considered as functions on $\mathbb{R}^n$) support.

Remark 2.1.
- (i) The use of functions of the form $\varphi(x_1, \cdots, x_n)$, $\varphi \in C_0^\infty(\mathbb{R}^n)$, is natural for $\mathbb{R}^\infty$, but differs from the standard core in the Gaussian case, where φ usually depends on a finite collection of measurable functionals X_{h_i}, $h_i \in H$, which are $\mathcal{N}(0, \|h_i\|^2)$-distributed.
- (ii) Clearly, C separates the points of $\mathbb{R}^\infty$. Furthermore, a simple monotone class argument shows that C is dense in any $L^p(\nu)$, $p \in [1, \infty)$ and any finite measure ν on $\mathbb{R}^\infty$.

Let ν be a probability measure on $\mathbb{R}^\infty$.
Throughout the paper it is assumed that

H1) all the projections
$$\nu_n = \nu \circ P_n^{-1},$$
where $P_n(x) = (x_1, \cdots, x_n)$, have Lebesgue densities.

H2) for every $i \in \mathbb{N}$ there exists a function $\beta_i \in L^1(\nu)$ such that
$$\int \partial_{e_i}\varphi \, d\nu = -\int \varphi\beta_i \, d\nu.$$
for every $\varphi \in \mathcal{C}$.

β_i is called logarithmic derivative of ν along e_i

Remark 2.2. Note that these assumptions are not independent: H2) implies H1) for $n = 1$.

Remark 2.3. It is important to keep in mind that the projections ν_n also have logarithmic derivatives given by the conditional expectations $\mathbb{E}_\nu^{\mathcal{F}_n}\beta_i$.

We say that a mapping $b : \mathbb{R}^\infty \to \mathbb{R}^\infty$ has divergence $\mathrm{div}_\nu b \in L^1(\nu)$ if the following relation holds for every $\varphi \in \mathcal{C}$:

$$\int \mathrm{div}_\nu b \, \varphi \, d\nu = -\int \langle b, \nabla\varphi\rangle \, d\nu. \tag{5}$$

For an account in infinite-dimensional analysis on spaces with differentiable measures the readers are referred to [8], [9].

We study (2), where $\rho = \rho(t, x)$ is a family of probability densities with respect to ν with initial condition $\rho(0, \cdot) = \rho_0$, i.e. we are looking for solutions $\rho(t, x)$ given as densities of a family of probability measures $\mu_t(dx) = \rho(t, x) \cdot \nu(dx)$.

Definition 2.4. We say that ρ is a solution of (2) for $t \in [0,T]$ with initial value ρ_0 if for every $\varphi \in \mathcal{C}$ and $t \in [0,T]$ one has

$$\int \varphi\rho(t,x)\, d\nu = \int \varphi\rho(0,x)\, d\nu + \int_0^t \int \langle b, \nabla\varphi\rangle \rho(s,x)\, d\nu\, ds. \tag{6}$$

Remark 2.5. The solution in the finite-dimensional case is defined in the same way.

Remark 2.6. We note that the existence of the right-hand side is not obvious because it is not clear a-priori that $\langle b, \nabla\varphi\rangle\rho(s,x) \in L^1(I_{[0,t]}ds \times \nu)$. Nevertheless, we will see in the following Lemma that this is indeed the case if c defined in (4) satisfies some natural assumptions.

The following result has been proved by Ambrosio and Figalli in [7] (Theorem 6.1) for $\rho_0 \in L^\infty(\gamma)$. The proof of this result is the same and so we omit it here.

Lemma 2.7. *Consider the standard Gaussian measure γ on $\mathbb{R}^d$. Let $\|c\| \in L^p(\gamma)$, $p > 1$ and $\|\exp\bigl(\varepsilon(\mathrm{div}_\gamma c)_-\bigr)\|_{L^1(\gamma)} < \infty$ for some $\varepsilon > 0$. Then for any $\rho_0 \in L^{q'}(\gamma)$ with $q' > q = \frac{p}{p-1} = p^*$ there exists $T = T(\varepsilon, p, q') > 0$ such that the equation*

$$\dot\rho + \mathrm{div}_\gamma\bigl(c \cdot \rho\bigr) = 0$$

admits a solution ρ on $[0,T]$ satisfying $\sup_{t\in[0,T]} \|\rho_t\|_{L^q(\gamma)} < \infty$.

Let us give the idea how to control the L^p-norms of ρ_t via $\mathrm{div}_\gamma c$ needed in the proof of Lemma 2.7. Below we set for brevity

$$\rho_t = \rho(t,\cdot), \quad \text{and} \quad X_t := X(t,\cdot)$$

(see (3)). The well known change of variables formula for the mapping $x \to X_t(x)$ is given by the Liouville formula:

$$\rho_t(X_t) = \rho_0 \cdot \exp\Bigl(-\int_0^t \mathrm{div}_\gamma c(X_r)\, dr\Bigr) = \rho_s(X_s)\exp\Bigl(-\int_s^t \mathrm{div}_\gamma c(X_r)\, dr\Bigr).$$

One has for any $q \ge 1$

$$\begin{aligned}\int \rho_t^q\, d\gamma &= \int \rho_t^{q-1}(X_t)\rho_0\, d\gamma = \int \rho_0^q \exp\Bigl(-\int_0^t (q-1)\mathrm{div}_\gamma c(X_r)\, dr\Bigr)\, d\gamma \\ &\le \frac{1}{t}\int_0^t \int \rho_0^q \exp\Bigl(-t(q-1)\mathrm{div}_\gamma c(X_r)\Bigr)\, d\gamma\, dr.\end{aligned}$$

Applying the Hölder inequality and change of variables one gets that $\Lambda(t) = \int_0^t \int \rho_r^q\, d\gamma\, dr$ satisfies $\Lambda' \le C(\Lambda/t)^\delta$ for some $\delta > 0$. By the standard arguments one gets that Λ' is uniformly bounded. This finally gives the following key estimate for ρ_t: for any $q' > q$ there exist positive constants $q_1(q,q'), q_2(q,q')$ such that

$$\int \rho_t^q\, d\gamma \le \Bigl(\int \rho_0^{q'}\, d\gamma\Bigr)^{q_1} \int_0^t \int \exp\Bigl(q_2 t[\mathrm{div}_\gamma c]_-\Bigr)\, d\gamma\, dr. \tag{7}$$

Clearly, under assumptions of Lemma 2.7 the right-hand side of the inequality is finite for sufficiently small t.

Lemma 2.8. *Let c satisfy the assumptions of Lemma 2.7 and f be a bounded Lipschitz function:*

$$|f|_{\mathrm{Lip}} = \sup_{x\neq y} \frac{|f(x)-f(y)|}{\|x-y\|} < \infty.$$

Then the solution ρ_t obtained in Lemma 2.7 satisfies the following property:

1) If, in addition,

$$\mathrm{div}_\gamma c \in L^N(\gamma), \qquad \|c\| \in L^{p'}(\gamma) \qquad \textit{for some } N > q^*,\ p' > p,$$

then there exist positive constants C, δ depending on

$$p, q, p', q', \quad \|c\|_{L^{p'}(\gamma)}, \quad \|\mathrm{div}_\gamma c\|_{L^N(\gamma)}, \quad \|\exp\bigl(\varepsilon(\mathrm{div}_\gamma c)_-\bigr)\|_{L^1(\gamma)}$$
$$\|\rho_0\|_{L^{q'}(\gamma)}, \ \sup|f|, \ |f|_{\mathrm{Lip}}$$

such that

$$\Bigl|\int f\rho_t^{1+\delta}\, d\gamma - \int f\rho_s^{1+\delta}\, d\gamma\Bigr| \le C|t-s|, \ \textit{for all } t, s \in [0,T]. \tag{8}$$

2) Without any extra assumption there exists a positive constant C depending on

$$p, q', \quad \|c\|_{L^p(\gamma)}, \quad \|\exp\bigl(\varepsilon(\mathrm{div}_\gamma c)_-\bigr)\|_{L^1(\gamma)}, \|\rho_0\|_{L^{q'}(\gamma)}, \ |f|_{\mathrm{Lip}}$$

such that

$$\Bigl|\int f\rho_t\, d\gamma - \int f\rho_s\, d\gamma\Bigr| \le C|t-s|, \ \textit{for all } t, s \in [0,T]. \tag{9}$$

Proof. We proof only 1) because the proof of 2) is easier and follows the same line. In the same way as in [7] we reduce the proof to the case when $X(t,x)$ is a globally defined smooth solution to $\dot{X} = c(X),\ X(0,x) = x$. We apply the change of variables formula for the mapping $x \to X_t(x)$. Let $s < t$, $\delta > 0$ and f be a bounded Lipschitz function.

$$\begin{aligned}
\int \rho_t^{1+\delta} f\, d\gamma &= \int \rho_t^\delta(X_t) f(X_t)\rho_0\, d\gamma \\
&= \int \rho_s^\delta(X_s) f(X_t)\exp\Bigl(-\delta\int_s^t \mathrm{div}_\gamma c(X_r)\, dr\Bigr)\rho_0\, d\gamma \\
&= \int \rho_s^\delta(X_s) f(X_t)\, d\gamma + \int \rho_s^\delta(X_s) f(X_t)\bigl[\exp\Bigl(-\delta\int_s^t \mathrm{div}_\gamma c(X_r)\, dr\Bigr) - 1\bigr]\rho_0\, d\gamma \\
&= \int \rho_s^{1+\delta} f\, d\gamma + \int \rho_s^\delta(X_s)\bigl(f(X_t) - f(X_s)\bigr)\ \rho_0\, d\gamma \\
&+ \int \rho_s^\delta(X_s) f(X_t)\bigl[\exp\Bigl(-\delta\int_s^t \mathrm{div}_\gamma c(X_r)\, dr\Bigr) - 1\bigr]\rho_0\, d\gamma.
\end{aligned}$$

Here we use that $\rho_s \cdot \gamma$ is the image of $\rho_0 \cdot \gamma$ under $x \to X(t,x)$. Note that $|e^{-t} - 1| \le u(t)$, where $u(t) = e^{\max\{-t,0\}}|t|$. Since f is convex one can apply the Jensen inequality. Then the last term in the right-hand side of the above inequality can be estimated by

$$\frac{\sup|f|}{t-s}\int \rho_s^\delta(X_s)\int_s^t u\bigl(\delta(t-s)\mathrm{div}_\gamma c(X_r)\bigr)\, dr\ \rho_0\, d\gamma.$$

The latter can be estimated by

$$\begin{aligned}
\sup|f|\Bigl[\int \rho_s^{q_1\delta+1}\, d\gamma\Bigr]^{\frac{1}{q_1}} &\Bigl[\frac{1}{t-s}\int\int_s^t u^{q_1^*}\bigl(\delta(t-s)\mathrm{div}_\gamma c\bigr)\ \rho_r\, dr\, d\gamma\Bigr]^{\frac{1}{q_1^*}} \\
&\le \sup|f|\Bigl[\int \rho_s^{q_1\delta+1}\, d\gamma\Bigr]^{\frac{1}{q_1}} \\
&\Bigl[\frac{1}{t-s}\int\int_s^t u^{q_1^* q^*}\bigl(\delta(t-s)\mathrm{div}_\gamma c\bigr)\ \, dr\, d\gamma\Bigr]^{\frac{1}{q_1^* q^*}}\Bigl[\frac{1}{t-s}\int\int_s^t \rho_r^q\, dr\, d\gamma\Bigr]^{\frac{1}{q_1^* q}}.
\end{aligned}$$

Applying again the Hölder inequality and Lemma 2.7 (see (7)) it is easy to show that the latter does not exceed $C|t-s|$ for some C, ε where $t-s, \delta$ are chosen sufficiently small and q_1^* close to 1.

Analogously, we estimate

$$\int \rho_s^\delta(X_s)\big(f(X_t)-f(X_s)\big)\rho_0 \, d\gamma \le \|f\|_{\mathrm{Lip}} \int \rho_s^\delta(X_s) \int_s^t \|c(X_r)\| \, dr\rho_0 \, d\gamma$$
$$\le (t-s)\|f\|_{\mathrm{Lip}} \Big[\int \rho_s^{p_2\delta+1} \, d\gamma\Big]^{\frac{1}{p_2}} \Big[\int \frac{1}{t-s}\int_s^t \|c\|^{p_2^*}\rho_r \;\; drd\gamma\Big]^{\frac{1}{p_2^*}}$$
$$\le (t-s)\|f\|_{\mathrm{Lip}} \Big[\int \rho_s^{p_2\delta+1} \, d\gamma\Big]^{\frac{1}{p_2}} \Big[\int \|c\|^{p_2^* p} \, d\gamma\Big]^{\frac{1}{p_2^* p}} \Big[\frac{1}{t-s}\int\int_s^t \rho_r^q \;\; drd\gamma\Big]^{\frac{1}{p_2^* q}}.$$

Choosing p_2^* close to 1 and a sufficiently small δ we get the desired result. $\square$

Remark 2.9. Below we generalize the existence result of [7] in infinite dimensions which has been established under the assumption that $\|c\| \in L^p(\gamma)$, $p>1$. We prove it under the weaker assumption $c_i \in L^p(\gamma)$. Furthermore, we work with our slightly weaker notion of solution from Definition 2.4 above.

Lemma 2.10. *Let $\nu = \gamma = \prod_{i=1}^\infty \gamma_i$ be a product of the standard Gaussian measures. Assume that $c=(c_i): \mathbb{R}^\infty \to \mathbb{R}^\infty$ is a mapping satisfying:*

1) *There exists $p>1$ such that $c_i \in L^p(\gamma)$ for every i.*
2) *The divergence $\mathrm{div}_\gamma c$ satisfies*

$$\exp\big(\varepsilon(\mathrm{div}_\gamma c)_-\big) \in L^1(\gamma). \tag{10}$$

Then there exists $T = T(\varepsilon, p, q')>0$ such that the equation $\dot\rho + \mathrm{div}_\gamma\big(c\cdot\rho\big)=0$ has a solution ρ_t on $[0,T]$ for every initial condition $\rho_0 \in L^{q'}(\gamma)$ with some $q' > \frac{p}{p-1} = p^*$. In addition, $\sup_{t\in[0,T]}\|\rho(t,\cdot)\|_{L^q(\gamma)} < \infty$.

Proof. Let us set:

$$c_{(n)} = \sum_{i=1}^n \mathbb{E}_\gamma^{\mathcal{F}_n} c_i \cdot e_i$$

and

$$\rho_{0,n}\cdot\big(\gamma\circ P_n^{-1}\big) = \big(\rho_0\cdot\gamma\big)\circ P_n^{-1}.$$

Equivalently, $\rho_{0,n} = \mathbb{E}_\gamma^{\mathcal{F}_n}\rho_0$. Note that assumption 1) ensures that $c_{(n)}$ is well-defined.

It is well-known (and easy to check) that

$$\mathrm{div}_\gamma c_{(n)} = \mathbb{E}_\gamma^{\mathcal{F}_n}[\mathrm{div}_\gamma c].$$

This relation easily implies that $[\mathrm{div}_\gamma c_{(n)}]_- \le \mathbb{E}_\gamma^{\mathcal{F}_n}[\mathrm{div}_\gamma c]_-$ and $|\mathrm{div}_\gamma c_{(n)}|^m \le \mathbb{E}_\gamma^{\mathcal{F}_n}|\mathrm{div}_\gamma c|^m$, $m\ge 1$.

By convexity and Jensen's inequality one has

$$c_{(n)} \in L^p\big(\gamma\circ P_n^{-1}\big).$$

Consider the equation

$$\dot{\rho_n} + \mathrm{div}_\gamma(\rho_n\cdot c_{(n)}) = 0$$

with $\rho_n|_{t=0} = \rho_{0,n}$. Since $\|\rho_{0,n}\|_{L^{q'}(\gamma)} \le \|\rho_0\|_{L^{q'}(\gamma)}$, we get by Lemma 2.8 that there exists $T = T(\varepsilon, p, q') > 0$ such that this equation admits a solution ρ_n on $[0, T]$ satisfying the following dimension-free bound

$$M = \sup_{t\in[0,T], n\in\mathbb{N}} \|\rho_n(t,\cdot)\|_{L^q(\gamma)} < \infty.$$

For any function $\varphi \in \mathcal{C}$, the following identity holds:

$$\int \varphi\rho_n(t,x)\, d\gamma = \int \varphi\rho_n(0,x)\, d\gamma + \int_0^t \int \langle c_n, \nabla\varphi\rangle \rho_n(s,x)\, d\gamma\, ds. \tag{11}$$

Applying a diagonal argument one can extract a subsequence (which is denoted in what follows again by ρ_n) such that $\{\rho_n(t_k, x)\}$ converges weakly in $L^q(\gamma)$ to some function $\rho(t_k, x)$ for any t_k from a dense countable subsequence $I = \{t_k\} \subset [0, T]$. Then $\{\rho_n(t, x)\}$ converges weakly in $L^q(\gamma)$ for any $t \in [0, T]$ to a function denoted in what follows by $\rho(t, x)$. Indeed, since

$$\sup_n \|\rho_n(t,x)\|_{L^q(\gamma)} < \infty,$$

by a standard subsequence argument it is enough to show that $\{\int f\rho_n(t,x)\, d\gamma\}$ is a convergent sequence for every $f \in L^p(\gamma)$. Clearly, it is sufficient to check the claim for functions from $\mathcal{C}$. Since for such a function $f \in \mathcal{C}$ the sequence $\{\int f\rho_n(t_k,x)\, d\gamma\}$ is convergent for every $t_k \in I$, it follows easily from the estimate (9) (we use here that f is cylindrical, hence the right-hand side of (9) depends on a finite collection of c_i) that $\{\int f\rho_n(t,x)\, d\gamma\}$ is a Cauchy sequence. Thus, we get that $\rho_n(t,x) \to \rho(t,x)$ weakly in $L^q(\gamma)$ for every $t \in [0, T]$.

One has for every smooth cylindrical function $\varphi = \varphi(x_1, \cdots, x_k)$

$$\lim_n \Bigl(\int \varphi\rho_n(t,x)\, d\gamma - \int \varphi\rho_n(0,x)\, d\gamma \Bigr) = \int \varphi\rho(t,x)\, d\gamma - \int \varphi\rho(0,x)\, d\gamma.$$

Set: $g_n(s) = \int \langle c_n, \nabla\varphi\rangle \rho_n(s,x)\, d\gamma$. Using the convergence $c_{(n)} \to c$ in $L^p(\gamma)$, one gets

$$\lim_n g_n(s) = \lim_n \int \langle c_n, \nabla\varphi\rangle \rho_n(s,x)\, d\gamma = \int \langle c, \nabla\varphi\rangle \rho(s,x)\, d\gamma = g(s)$$

for every $s \in [0, T]$. Clearly,

$$\sup_{s\in[0,T], n\in\mathbb{N}} |g_n(s)| \le \sup_{s\in[0,T], n\in\mathbb{N}} \|\rho_n(s,\cdot)\|_{L^q(\gamma)} \|P_k \circ c\|_{L^p(\gamma)} \|\nabla\varphi\|_{L^\infty(\gamma)}.$$

Then the Lebesgue dominated convergence theorem implies

$$\int_0^t \int \langle c_n, \nabla\varphi\rangle \rho_n(s,x)\, d\gamma\, ds \to \int_0^t \int \langle c, \nabla\varphi\rangle \rho(s,x)\, d\nu\, ds.$$

Passing to the limit in (11) we get that ρ is the desired solution. □

Before we proceed to the general case, let us explain the main idea of the proof. We construct a mapping T pushing forward another measure μ onto ν. If T is sufficiently smooth, one can define the following new drift:

$$c = DS(T) \cdot b(T),$$

where S is the inverse mapping to T. One has

$$\mu = \nu \circ S^{-1}$$

and

$$(DT)^{-1} = DS(T).$$

Let us give a heuristic proof of the key relation:

$$\mathrm{div}_\mu c \circ S = \mathrm{div}_\nu b. \tag{12}$$

Take a test function $\varphi \in \mathcal{C}$. One has

$$\int \langle \nabla\varphi, c\rangle \circ S \; d\nu = \int \langle \nabla\varphi, c\rangle \; d\mu = -\int \varphi \; \mathrm{div}_\mu c \; d\mu = -\int \varphi(S) \; \mathrm{div}_\mu c \circ S \; d\nu. \tag{13}$$

On the other hand, we note that by the chain rule

$$\nabla(\varphi(S)) = (DS)^* \nabla\varphi(S). \tag{14}$$

Hence $\int \langle \nabla\varphi, c\rangle \circ S \; d\nu$ is equal to

$$\int \langle (DS^*)^{-1} \nabla(\varphi(S)), c(S)\rangle \; d\nu = \int \langle \nabla(\varphi(S)), b\rangle \; d\nu = -\int \varphi(S) \mathrm{div}_\nu b \; d\nu. \tag{15}$$

Obviously, (13) and (15) imply (12).

Now let us try to solve the equation

$$\dot{\rho} + \mathrm{div}_\nu(\rho \cdot b) = 0$$

for a wide class of probability measures. Assume that ν is the image of the standard Gaussian measure γ under a mapping T. Setting $c = DT^{-1} \cdot b(T) = DS(T) \cdot b(T)$ we transform the equation into

$$\dot{g} + \mathrm{div}_\gamma(g \cdot c) = 0, \tag{16}$$

where every $\rho \cdot \mu$ is the image of $g \cdot \gamma$ under T. Applying Lemma 2.10 we obtain a solution to (16). Then the function

$$\rho(t, x) = g(t, T^{-1}(x))$$

presents the desired solution. This follows immediately from the definition of solution in Definition 2.4 and the change of variables formula.

3. Sobolev estimates for triangular mappings

Consider a Borel probability measure μ on $\mathbb{R}^\infty$. We denote by $\mu_i = \mu \circ P_i^{-1}$ the projection of μ onto the subspace generated by the first i basis vectors. Recall that throughout the paper μ_i is assumed to have a Lebesgue density, which will be denoted by ρ_{μ_i}. For every fixed $x = (x_1, \cdots, x_{i-1})$ we denote by $\mu^{\perp}_{x,i}$ the corresponding **one-dimensional** conditional measure obtained from the disintegration of μ_i with respect to μ_{i-1}. Note that $\mu_{i-1} = \mu_i \circ P_{i-1}^{-1}$. These measures are related by the following identity

$$\int \varphi \rho_{\mu_i} dx = \int \varphi(x, x_i) \; \mu_i(dx dx_i) = \int \Bigl(\int \varphi(x, x_i) \mu^{\perp}_{x,i}(dx_i) \Bigr) \; \mu_{i-1}(dx)$$

for all bounded Borel $\varphi : \mathbb{R}^i \to \mathbb{R}$.

If for μ_i-almost points x the corresponding conditional measures $\mu^{\perp}_{x,i}$ have Lebesgue densities, they will be denoted by $\rho_{\mu^{\perp}_{x,i}}$. In this case the latter formula reads as

$$\int \varphi \rho_{\mu_i} dx = \int \Bigl(\int \varphi(x, x_i) \rho_{\mu^{\perp}_{x,i}} dx_i \Bigr) \; \rho_{\mu_{i-1}}(x) \; dx.$$

In this section we study a-priori estimates for so-called triangular mappings. We call a mapping $T : \mathbb{R}^\infty \to \mathbb{R}^\infty$ triangular if it has the form

$$T = \sum_{i=1}^{\infty} T_i(x_1, \cdots, x_i) e_i$$

and, in addition, $x_i \to T_i(x_1, \cdots, x_i)$ is an increasing function.

Given two probability measures μ and ν on $\mathbb{R}^\infty$ we are looking for a triangular mapping $T : \mathbb{R}^\infty \to \mathbb{R}^\infty$ pushing forward μ onto ν. The proof of existence of mappings of such type on $\mathbb{R}^\infty$ for a broad class of measures can be found in ([12], [8]). It relies on the fact that T can be precisely described in terms of conditional probabilities of μ and ν. In the one-dimensional case $T = T_{\mu,\nu}$ is defined by the relation

$$\int_{-\infty}^{x} \rho_\mu(t)\ dt = \int_{-\infty}^{T(x)} \rho_\nu(t)\ dt.$$

In the finite- and infinite-dimensional case T is obtained by induction

1) T_1 is the increasing transport of the projections on the first coordinate

$$T_1(x_1) = T_{\mu_1,\nu_1}(x_1) \tag{17}$$

2) T_i, $i > 1$, is the increasing transport of the one-dimensional conditional measures $\tilde{\mu}$, $\tilde{\nu}$:

$$T_i(x_1, \cdots, x_i) = T_{\tilde{\mu},\tilde{\nu}}(x_i), \tag{18}$$

where $\tilde{\mu} = \mu^{\perp}_{x,i}$, $\tilde{\nu} = \nu^{\perp}_{T_{i-1}(x),i}$, $x = (x_1, \cdots, x_{i-1})$.

The existence result and the basic properties formulated in the following theorem have been proved in the ([11], [12]).

Theorem 3.1. *Let μ be a probability measure on $\mathbb{R}^\infty$ satisfying the following assumptions:*

1) *Any projection μ_i, $i \in \mathbb{N}$, is absolutely continuous measure with respect to Lebesgue measure on $\mathbb{R}^i$.*
2) *For μ_i-almost all x the corresponding conditional measures $\mu^{\perp}_{x,i}$ are absolutely continuous, with respect to Lebesgue measure on $\mathbb{R}$.*

Then there exists a triangular mapping T pushing forward μ onto ν. The mapping T is unique up to a set of μ-measure zero.

Let ν be another such measure. Then there exists a triangular mapping S pushing forward ν onto μ. In addition, they are reciprocal:

$$T \circ S = \mathrm{Id} \qquad \nu - \text{a.e.}$$

$$S \circ T = \mathrm{Id} \qquad \mu - \text{a.e.}$$

Remark 3.2. The absolute continuity of $\mu^{\perp}_{x,i}$ follows, in particular, from the absolute continuity of the conditional measures of μ in the direction e_i. The latter follows in turn from the existence of the logarithmic derivative β_i. In particular, the measures satisfying our general assumptions from Section 2 do satisfy the assumptions of Theorem 3.1.

Note that in the one-dimensional case T and S are just non-decreasing mappings which can be written exactly in terms of the distribution functions of μ and ν. Hence T (S) admits classical pointwise derivative T' μ (ν)-almost everywhere. One can easily check that T' is μ-a.e. positive, because otherwise ν has a non-trivial singular part. In particular

$$T'(S)\cdot S' = 1, \quad \mu - \text{a.e.}$$

Remark 3.3. Since every T_i is constructed as a one-dimensional increasing transportation of conditional measures, the following generalization of the above relation

$$\partial_{x_i} T_i(S)\cdot \partial_{x_i} S_i = 1 \quad \mu - \text{a.e..}$$

is valid in the finite- and infinite-dimensional case.

If T and S are smooth (meaning that every function T_i, S_i is smooth) then their Jacobian matrices are triangular:

$$DT = \begin{pmatrix} \partial_{x_1}T_1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \partial_{x_1}T_2 & \partial_{x_2}T_2 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \partial_{x_1}T_i & \partial_{x_2}T_i & \partial_{x_3}T_i & \cdots & \partial_{x_{i-1}}T_i & \partial_{x_i}T_i & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

$$DS = \begin{pmatrix} \partial_{x_1}S_1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \partial_{x_1}S_2 & \partial_{x_2}S_2 & 0 & \cdots & 0 & 0 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \partial_{x_1}S_i & \partial_{x_2}S_i & \partial_{x_3}S_i & \cdots & \partial_{x_{i-1}}S_i & \partial_{x_i}S_i & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

In addition, if all $\partial_{x_i}T_i(x) \ne 0$ at x (which happens μ-a.e.), then

$$DS(T(x)) = (DT)^{-1}.$$

In this section we establish global Sobolev estimates for triangular mappings. Note that some (dimension-dependent) Sobolev estimates for the triangular mappings have been obtained in [29]. See also [25] for similar results on optimal transportation.

Definition 3.4. Let ν be a probability measure on $\mathbb{R}^\infty$ and $f \in L^1(\nu)$. We say that $\partial_{x_i} f$ is a Sobolev partial derivative of f if $f\beta_i \in L^1(\nu)$ and

$$\int \partial_{x_i} f \cdot \varphi \, d\nu = -\int f\cdot \partial_{x_i}\varphi \, d\nu - \int f\varphi\beta_i \, d\nu$$

for every $\varphi \in \mathcal{C}$. Obviously, this defines $\partial_{x_i} f$ uniquely.

Definition 3.5. Let $p \in (1,\infty)$ and ν be a probability measure such that $\beta_i \in L^{p^*}(\nu)$ for all i. We say that a function f belongs to the Sobolev class $W^{1,p}(\nu)$ if

$$\|f\|_{L^p(\nu)} + \Big\| \Big(\sum_{i=1}^{\infty} f^2_{x_i}\Big)^{1/2} \Big\|_{L^p(\nu)} < \infty.$$

If, in addition, every f_{x_i} has all Sobolev partial derivatives in $L^p(\nu)$ and

$$\|D^2 f\|_{\mathcal{HS}} = \Big(\sum_{i,j} f^2_{x_i x_j}\Big)^{\frac{1}{2}} \in L^p(\nu),$$

we say that $f \in W^{2,p}(\nu)$.

Remark 3.6. Though we shall not use this below, it follows by [3] that due to the assumption that $\beta_i \in L^{p^*}(\nu)$ for all $i \in \mathbb{N}$ both $W^{1,p}(\nu)$ and $W^{2,p}(\nu)$ are complete.

All derivatives of S and T below will be understood in the Sobolev sense, with respect to ν and μ respectively.

We start with a one-dimensional estimate.

Proposition 3.7. *Let $\mu = e^{-V}dx$, $\nu = e^{-W}dx$ be two probability measures on $\mathbb{R}$ with continuously differentiable functions V and W. Consider the increasing mapping T pushing forward μ onto ν. Assume that*

$$xW'(x) \ge \frac{-1+\varepsilon}{p-1}$$

for some $p > 1$ and $\varepsilon > 0$. Assume, in addition, that $|x|^{\frac{p(p+\delta)}{\delta}} \in L^1(\nu)$, $|V'|^{p+\delta} \in L^1(\mu)$ for some $\delta > 0$. Then $T' := DT \in L^p(\mu)$.

Proof. By the change of variables formula $\int_{-\infty}^{x} e^{-V(t)}\, dt = \int_{-\infty}^{T(x)} e^{-W(t)}\, dt$. Clearly, this implies that T is differentiable and

$$e^{-V} = T' e^{-W(T)}.$$

Moreover, $T' = e^{W(T)-V}$ is continuously differentiable and satisfies

$$T'' = T'\bigl(W'(T)T' - V'\bigr).$$

Take a positive test function ξ. Integrating by parts one obtains

$$\begin{aligned}
&\int_{\mathbb{R}} (T')^p e^{-V} \xi\, dx = \\
&= -(p-1)\int_{\mathbb{R}} T(T')^{p-2}T'' e^{-V}\xi\, dx + \int_{\mathbb{R}} T(T')^{p-1}V' e^{-V}\xi\, dx - \int_{\mathbb{R}} (T')^{p-1}e^{-V}\xi'\, dx \\
&= -(p-1)\int_{\mathbb{R}} TW'(T)(T')^p e^{-V}\xi\, dx + p\int_{\mathbb{R}} T(T')^{p-1}V' e^{-V}\xi\, dx - \int_{\mathbb{R}} (T')^{p-1}e^{-V}\xi'\, dx.
\end{aligned}$$

Note that

$$\begin{aligned}
\varepsilon \int_{\mathbb{R}} (T')^p e^{-V}\xi\, dx &\le -p\int_{\mathbb{R}} T(T')^{p-1}V' e^{-V}\xi\, dx - \int_{\mathbb{R}} (T')^{p-1}e^{-V}\xi'\, dx \\
&\le \frac{\varepsilon}{2}\int_{\mathbb{R}} (T')^p e^{-V}\xi\, dx + N(\varepsilon,p)\int_{\mathbb{R}} T^p (V')^p e^{-V}\xi\, dx - \int_{\mathbb{R}} (T')^{p-1}e^{-V}\xi'\, dx.
\end{aligned}$$

By the Hölder inequality

$$\begin{aligned}
\int_{\mathbb{R}} T^p (V')^p e^{-V}\xi\, dx &\le \int_{\mathbb{R}} (V')^{p+\delta} e^{-V}\xi\, dx + C(\delta,p)\int_{\mathbb{R}} T^{\frac{p(p+\delta)}{\delta}} e^{-V}\xi\, dx \\
&= \int_{\mathbb{R}} (V')^{p+\delta} e^{-V}\xi\, dx + C(\delta,p)\int_{R} |x|^{\frac{p(p+\delta)}{\delta}} e^{-W}\xi\, dx, \\
-\int_{\mathbb{R}} (T')^{p-1}e^{-V}\xi'\, dx &\le \frac{\varepsilon}{4}\int_{\mathbb{R}} (T')^p e^{-V}\xi\, dx + c(p,\varepsilon)\int_{\mathbb{R}} \Bigl|\frac{\xi'}{\xi}\Bigr|^p \xi e^{-V}\, dx.
\end{aligned}$$

Thus, we obtain a bound for $\int_{\mathbb{R}} (T')^p (V')^p e^{-V}\xi\, dx$. Taking a suitable sequence $\{\xi_n\}$ with $\xi_n \to 1$ and $\lim_n \int_{\mathbb{R}} \Bigl|\frac{\xi_n'}{\xi_n}\Bigr|^p \xi_n e^{-V}\, dx = 0$ we complete the proof. $\square$

Now let us come back to the infinite-dimensional case. Below in the proofs we apply the following scheme.

1) Prove the statement for smooth positive densities.
2) Approximate the Sobolev densities by smooth positive densities and deduce the desired estimates.

First, we need an approximation lemma.

Lemma 3.8. *For every probability measure $\nu = \rho\ dx$ on $\mathbb{R}^d$ satisfying $\beta_i = \frac{\rho_{x_i}}{\rho} \in L^p(\nu)$ for some $p \geq 1$ (with $\beta_i := 0$ on $\{\rho = 0\}$) and all $1 \leq i \leq d$ there exists a sequence of probability measures $\nu_n = \rho_n\ dx$ such that*

1) *$\nu_n \to \nu$ in variation norm,*
2) $\lim_n \int \Bigl\| \frac{\nabla \rho_n}{\rho_n} \Bigr\|^p \rho_n\ dx = \int \Bigl\| \frac{\nabla \rho}{\rho} \Bigr\|^p \rho\ dx.$
3) *every ρ_n is smooth and nonnegative,*
4) *the partial derivatives $\partial_{x_i} \rho_n$ are uniformly bounded and integrable for every i, n.*

Moreover, if every logarithmic derivative β_i has Sobolev derivative along any coordinate x_j and, in addition, there exists $p \geq 1$ such that

$$\partial_{x_j} \beta_i = \frac{\partial_{x_i x_j} \rho}{\rho} - \frac{\partial_{x_i} \rho \cdot \partial_{x_j} \rho}{\rho^2} \in L^p(\nu), \quad \beta_i \in L^{2p}(\nu)$$

then item 2) can be strengthened as follows:

$$\lim_n \int \Bigl\| \frac{\nabla \rho_n}{\rho_n} \Bigr\|^{2p} \rho_n\ dx = \int \Bigl\| \frac{\nabla \rho}{\rho} \Bigr\|^{2p} \rho\ dx,\ \lim_n \int \Bigl| \frac{\partial_{x_i x_j} \rho_n}{\rho_n} \Bigr|^p \rho_n\ dx = \int \Bigl| \frac{\partial_{x_i x_j} \rho}{\rho} \Bigr|^p \rho\ dx.$$

Sketch of the proof: The proof is quite standard and we only give a sketch here. It consists of two steps: 1) approximate ρ by $\varphi_n \cdot \rho$, where $\{\varphi_n\}$ is a sequence of smooth nonnegative compactly supported functions satisfying $\sup_x \Bigl\| \frac{\nabla \varphi_n}{\varphi_n} \Bigr\|^p \varphi_n < \infty$ for every n such that $\varphi_n \to 1$ pointwise; 2) approximate compactly supported ρ by the ordinary convolutions with smooth kernels $(2\pi t)^{\frac{d}{2}} e^{-\frac{x^2}{2t^2}}$. For 2) we apply Jensen's inequality and Fatou's lemma.

Remark 3.9. It is straightforward to check using (17), (18) that T and S are continuously differentiable, $\mu = e^{-V}\ dx$, $\nu = e^{-W}\ dx$, and V, W have uniformly bounded derivatives. Note that in this case all conditional measures have positive densities and all the derivatives $\partial_{x_i} S_i, \partial_{x_i} T_i$ are positive. More precise statements about the regularity of triangular mappings can be found in [12] (Lemma 2.6) and [29].

Proposition 3.10. *Consider the triangular mapping S pushing forward ν onto γ. Assume that $\beta_i \in L^2(\nu)$ for all i. Then for every i the mapping S_i belongs to $W^{1,2}(\nu)$. In particular, the following estimates hold:*

$$\int \bigl(\partial_{x_i} S_i\bigr)^2 d\nu \leq \int \bigl(\mathbb{E}_\nu^{\mathcal{F}_i} \beta_i\bigr)^2 d\nu,$$

$$\int \bigl(\partial_{x_j} S_i\bigr)^2 d\nu \leq \int \Bigl(\mathbb{E}_\nu^{\mathcal{F}_i} \beta_j\Bigr)^2 d\nu - \int \Bigl(\mathbb{E}_\nu^{\mathcal{F}_{i-1}} \beta_j\Bigr)^2\ d\nu,\ i > j.$$

In particular,

$$\|\partial_{x_j} S\|^2 = \sum_{i \geq j} \int \bigl(\partial_{x_j} S_i\bigr)^2 d\nu \leq \int \beta_j^2 d\nu.$$

Proof. First, we note that due to the finite-dimensional structure of triangular mappings it is sufficient to establish the statement for finite-dimensional measures. We start with the case when $\rho_\nu = e^{-V}$, where V is a smooth function on $\mathbb{R}^d$ with uniformly bounded derivatives.

In the proof we apply the following relation between the logarithmic derivatives and conditional densities of the corresponding projections

$$\mathbb{E}_{\nu}^{\mathcal{F}_i} \beta_i = \frac{\partial_{x_i} \rho_{\nu^{\perp}_{x,i}}}{\rho_{\nu^{\perp}_{x,i}}}.$$

We keep the notation ρ_{γ} for the Lebesgue density of the 1-dimensional standard Gaussian measure γ:

$$\rho_{\gamma} = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

According to Remark 3.9 all the functions S_i are continuously differentiable and $\partial_{x_i} S_i > 0$. It follows by the change of variables formula that $\int S_i^2 \ d\nu = \int x_i^2 \ d\gamma$, hence $S_i \in L^2(\nu)$. This implies that $\partial_{x_i} S_i \in L^1(\nu)$. Indeed

$$\int \partial_{x_i} S_i \ d\nu = -\int S_i \beta_i \ d\nu \le \|S_i\|_{L^2(\nu)} \|\beta_i\|_{L^2(\nu)}.$$

Let us estimate $\int \bigl(\partial_{x_i} S_i\bigr)^2 \ d\nu$. One has the following explicit formula for S_i (we stress that the expression below makes sense because ρ_ν is positive as well as the densities of its projections and conditional measures):

$$\int_{-\infty}^{S_i(x,x_i)} \rho_{\gamma}(t) \ dt = \int_{-\infty}^{x_i} \rho_{\nu^{\perp}_{x,i}}(t) \ dt, \tag{19}$$

where

$$\rho_{\nu^{\perp}_{x,i}}(t) = \frac{\rho_{\nu_i}(x,t)}{\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t) \ dt}.$$

Differentiating (19) along x_i one obtains

$$\frac{\rho_{\nu^{\perp}_{x,i}}}{\rho_{\gamma}(S_i)} = \partial_{x_i} S_i. \tag{20}$$

Formally applying integration by parts we get

$$\begin{aligned}\int \bigl(\partial_{x_i} S_i\bigr)^2 \ d\nu &= \int \partial_{x_i} S_i \frac{\rho_{\nu^{\perp}_{x,i}}}{\rho_{\gamma}(S_i)} \ d\nu = -\int \partial_{x_i} S_i \cdot S_i^2 \frac{\rho_{\nu^{\perp}_{x,i}}}{\rho_{\gamma}(S_i)} \ d\nu \\ &- 2\int S_i \frac{\partial_{x_i} \rho_{\nu^{\perp}_{x,i}}}{\rho_{\gamma}(S_i)} \ d\nu \le \int \frac{1}{\partial_{x_i} S_i} \Bigl(\frac{\partial_{x_i} \rho_{\nu^{\perp}_{x,i}}}{\rho_{\nu^{\perp}_{x,i}}}\Bigr)^2 \frac{\rho_{\nu^{\perp}_{x,i}}}{\rho_{\gamma}(S_i)} \ d\nu = \\ &= \int \Bigl(\frac{\partial_{x_i} \rho_{\nu^{\perp}_{x,i}}}{\rho_{\nu^{\perp}_{x,i}}}\Bigr)^2 \ d\nu = \int \bigl(\mathbb{E}_{\nu}^{\mathcal{F}_i} \beta_i\bigr)^2 \ d\nu.\end{aligned}$$

To justify the above computation we integrate not over ν but over $\xi \cdot \nu$, where ξ is a compactly supported smooth function on $\mathbb{R}^d$. By the same arguments one gets

$$\int \bigl(\partial_{x_i} S_i\bigr)^2 \xi \ d\nu \le (1+\varepsilon) \int \bigl(\mathbb{E}_{\nu}^{\mathcal{F}_i} \beta_i\bigr)^2 \xi \ d\nu + c(\varepsilon) \int \bigl(\frac{\partial_{x_i} \xi}{\xi}\bigr)^2 \xi \ d\nu.$$

Choosing "an appropriate" convergent sequence $\xi_k \to 1$ with $\lim_k \int \bigl(\frac{\partial_{x_i} \xi_k}{\xi_k}\bigr)^2 \xi_k \ d\nu = 0$ one easily gets the desired result.

Analogously, one has for $\partial_{x_j} S_i$, $i \ne j$:

$$\rho_{\gamma}(S_i) \partial_{x_j} S_i = \frac{\int_{-\infty}^{x_i} \partial_{x_j} \rho_{\nu_i}(x,t) \ dt}{\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t) \ dt} - \frac{\int_{-\infty}^{x_i} \rho_{\nu_i}(x,t) dt \ \int_{-\infty}^{\infty} \partial_{x_j} \rho_{\nu_i}(x,t) \ dt \cdot}{\Bigl(\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t) \ dt\Bigr)^2}.$$

Denoting the right-hand side by f one gets

$$\frac{\partial_{x_j} S_i}{\partial_{x_i} S_i} = \frac{f}{\rho_{\nu^{\perp}_{x,i}}}. \tag{21}$$

Consider the following formal computations

$$\begin{aligned}
\int \bigl(\partial_{x_j} S_i\bigr)^2 \, d\nu &= \int \bigl(\partial_{x_i} S_i\bigr)^2 \frac{f^2}{\rho^2_{\nu^{\perp}_{x,i}}} \, d\nu = \int \frac{\partial_{x_i} S_i}{\rho_\gamma(S_i)} \frac{f^2}{\rho_{\nu^{\perp}_{x,i}}} \, d\nu \\
&= \int \int \frac{\partial_{x_i} S_i}{\rho_\gamma(S_i)} f^2 \, d\nu_{i-1} \, dx_i = - \int \int \frac{\partial_{x_i} S_i}{\rho_\gamma(S_i)} S_i^2 f^2 \, d\nu_{i-1} \, dx_i \\
&- 2 \int \int \frac{S_i}{\rho_\gamma(S_i)} f f_{x_i} \, d\nu_{i-1} \, dx_i \le \int \int \frac{f^2_{x_i}}{\rho_\gamma(S_i) \partial_{x_i} S_i} \, d\nu_{i-1} \, dx_i \\
&= \int \Bigl(\frac{f_{x_i}}{\rho_{\nu^{\perp}_{x,i}}} \Bigr)^2 \, d\nu_i = \int \Bigl(\frac{\partial_{x_j} \rho_{\nu_i}}{\rho_{\nu_i}} - \frac{\int_{-\infty}^{\infty} \partial_{x_j} \rho_{\nu_i}(x,t) \, dt}{\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t) \, dt} \Bigr)^2 \, d\nu \\
&= \int \Bigl(\mathbb{E}_\nu^{\mathcal{F}_i} \beta_j - \mathbb{E}_\nu^{\mathcal{F}_{i-1}} \beta_j \Bigr)^2 \, d\nu = \int \Bigl(\mathbb{E}_\nu^{\mathcal{F}_i} \beta_j - \mathbb{E}_\nu^{\mathcal{F}_{i-1}} \beta_j \Bigr)^2 \, d\nu \\
&= \int \Bigl(\mathbb{E}_\nu^{\mathcal{F}_i} \beta_j \Bigr)^2 \, d\nu - \int \Bigl(\mathbb{E}_\nu^{\mathcal{F}_{i-1}} \beta_j \Bigr)^2 \, d\nu.
\end{aligned}$$

To justify the global integration above we integrate again with respect to $\xi \cdot \nu$, where ξ is a compactly supported smooth positive function on $\mathbb{R}^d$. Repeating the above arguments one gets

$$\begin{aligned}
\int \bigl(\partial_{x_j} S_i\bigr)^2 \xi \, d\nu &= - \int S_i \cdot \partial_{x_j} S_i \cdot \partial_{x_i} \xi \, d\nu - \int \int \frac{\partial_{x_i} S_i}{\rho_\gamma(S_i)} S_i^2 f^2 \xi \, d\nu_{i-1} \, dx_i \\
&- 2 \int \int \frac{S_i}{\rho_\gamma(S_i)} f f_{x_i} \xi \, d\nu_{i-1} \, dx_i
\end{aligned}$$

The term $- \int S_i \cdot \partial_{x_j} S_i \cdot \partial_{x_i} \xi \, d\nu$ can be estimated by

$$\varepsilon \int (\partial_{x_j} S_i)^2 \xi \, d\nu + \frac{4}{\varepsilon} \int S_i^2 \bigl(\frac{\partial_{x_i} \xi}{\xi} \bigr)^2 \xi \, d\nu.$$

Finally

$$(1-\varepsilon) \int \bigl(\partial_{x_j} S_i\bigr)^2 \xi \, d\nu \le \int \Bigl(\mathbb{E}_\nu^{\mathcal{F}_i} \beta_j - \mathbb{E}_\nu^{\mathcal{F}_{i-1}} \beta_j \Bigr)^2 \xi \, d\nu + \frac{4}{\varepsilon} \int S_i^2 \bigl(\frac{\partial_{x_i} \xi}{\xi} \bigr)^2 \xi \, d\nu.$$

Estimating the term $\int S_i^2 \bigl(\frac{\partial_{x_i} \xi}{\xi} \bigr)^2 \xi \, d\nu$ by the Hölder inequality and choosing an appropriate sequence $\xi_n \to 1$ we complete the justification of the above formal computation.

It remains to approximate an arbitrary density ρ on $\mathbb{R}^d$ with $\int \beta_i^2 \rho \, dx < \infty$ by smooth densities and prove that the desired a-priori estimate is preserved under taking the limit. Indeed, let $\rho_{(k)} = e^{-V_k}$ be approximating densities constructed in Lemma 3.8. Let $S^{(k)}$ be the triangular mappings pushing forward $\rho_{(k)} \, dx$ onto γ. Note that the functions $S_i^{(k)} \cdot \rho_{(k)}$ are in $W^{1,1}(\mathbb{R}^n)$. Indeed,

$$\int \| DS^{(k)} \|_{HS} \rho_{(k)} \, dx \le \Bigl[\int \| DS^{(k)} \|^2_{HS} \rho_{(k)} \, dx \Bigr]^{1/2} \le \sum_i \int (\beta_i^{(k)})^2 \rho_{(k)} \, dx \le \| \sqrt{\rho_{(k)}} \|^2_{W^{1,2}(\mathbb{R}^n)}$$

and

$$\int |S_i^{(k)}| \|\nabla \rho_{(k)}\| \, dx \le \Bigl[\int |S_i^{(k)}|^2 \rho_{(k)} \, dx\Bigr]^{1/2} \|\sqrt{\rho_{(k)}}\|_{W^{1,2}(\mathbb{R}^n)} = \Bigl[\int x_i^2 \, d\gamma\Bigr]^{1/2} \|\sqrt{\rho_{(k)}}\|_{W^{1,2}(\mathbb{R}^n)}.$$

Using Sobolev embeddings and extracting an almost everywhere convergent subsequence one can assume from the very beginning that $S^{(k)} \cdot \rho_{(k)}$ converges almost everywhere. Using that $\rho_{(k)}$ converges almost everywhere to ρ, one can easily see that $S^{(k)}$ converges to a triangular mapping S at ν-almost all points. Using almost everywhere convergence it is easy to check that S pushes forward ρ onto γ. From the almost everywhere convergence and the following change of variables formula

$$\int \|S^{(k)}\|^2 \rho_{(k)} \, dx = \int \|x\|^2 \, d\gamma = \int \|S\|^2 \rho \, dx,$$

one gets that $S^{(k)} \cdot \sqrt{\rho_{(k)}}$ converges to $S \cdot \sqrt{\rho}$ in $L^2(\mathbb{R}^d)$. Applying the estimates above one proves that $\nabla S_i^{(k)} \cdot \sqrt{\rho_{(k)}}$ converges (up to a subsequence) weakly in $L^2(\mathbb{R}^n)$ to a vector field v . Standard integration by parts arguments show that v can be identified with $\nabla S_i \cdot \sqrt{\rho}$. Indeed,

$$\int \varphi \cdot \partial_{x_i} S^{(k)} \rho^{(k)} \, dx = -\int \partial_{x_i} \varphi \cdot S^{(k)} \rho^{(k)} \, dx - \int \varphi \; S^{(k)} \cdot \partial_{x_i} \rho^{(k)} \, dx.$$

The left-hand side converges to $\int \varphi \cdot v \sqrt{\rho} \, dx$ and the right-hand side to

$$-\int \partial_{x_i} \varphi \cdot S\rho \, dx - \int \varphi \; S \cdot \partial_{x_i} \rho \, dx$$

(this follows from the strong convergence of $S^{(k)} \cdot \sqrt{\rho_{(k)}}$ and $\nabla \rho_{(k)} / \sqrt{\rho_{(k)}}$). Hence

$$\int |\partial_{x_i} S|^2 \rho \, dx \le \underline{lim}_k \int |\partial_{x_i} S^{(k)}|^2 \rho_{(k)} \, dx \le \lim_k \sum_i \int (\beta_i^{(k)})^2 \rho_{(k)} \, dx = \sum_i \int (\beta_i)^2 \rho \, dx.$$

The other estimates can be justified in the same way. Hence the proof is complete. □

Remark 3.11. It is clear, that formula (20) remains true in the non-smooth setting, for instance under the assumptions of Proposition 3.10. We understand $\partial_{x_i} S_i$ as the Sobolev derivative or just as the classical derivative of the one-dimensional increasing mapping $x_i \to S_i$. Taking product from $i = 1$ to d in (20) we obtain the change of variables formula

$$\rho_\nu = (2\pi)^{-d/2} e^{-\frac{1}{2}|S|^2} \det DS = \prod_{i=1}^{d} \rho_\gamma(S_i) \cdot \partial_{x_i} S_i. \tag{22}$$

Remark 3.12. In what follows we will give a proof for a-priori estimates only in the case of smooth and positive densities. The complete justification for Sobolev densities can be spelt out as in the proof of Proposition 3.10.

In particular, note that since all the densities are positive and smooth, all the expressions in the intermediate computations are well-defined.

We also note that in the general (i.e. Sobolev) case $\partial_{x_i} S_i$ remains positive ν-almost everywhere, because $\partial_{x_i} S_i = 0$ implies that the corresponding conditional density of ν vanishes, which can happen only on a set of ν-measure zero.

Remark 3.13. Another estimate of this type has been mentioned (without rigorous proof) in [25]

$$\int \beta_i^2 \ d\nu = \int \|\partial_{x_i} S\|^2 \ d\nu + \sum_k \int \Big(\frac{\partial_{x_i x_k} S_k}{\partial_{x_k} S_k} \Big)^2 \ d\nu.$$

Moreover, if the image measure μ is not Gaussian, but uniformly log-concave, i.e. has the form $\mu = e^{-W} \ dx$ with $D^2 W \ge K \cdot \mathrm{Id}$, $K > 0$, then

$$\int \beta_i^2 \ d\nu \ge K \int \|\partial_{x_i} S\|^2 \ d\nu + \sum_k \int \Big(\frac{\partial_{x_i x_k} S_k}{\partial_{x_k} S_k} \Big)^2 \ d\nu.$$

Remark 3.14. One can easily generalize Proposition 3.10 to the L^p-case. Under the same assumptions for every $p > 1$ there exists $C = C(p)$ such that

$$\int \big(\partial_{x_i} S_i\big)^p d\nu \le C(p) \int \big|\mathbb{E}_\nu^{\mathcal{F}_i} \beta_i\big|^p d\nu$$

and

$$\int \big|\partial_{x_j} S_i\big|^p d\nu \le C(p) \int \Big|\mathbb{E}_\nu^{\mathcal{F}_i} \beta_j - \mathbb{E}_\nu^{\mathcal{F}_{i-1}} \beta_j\Big|^p d\nu.$$

The proof follows along the same line of arguments as above.

We prove some L^p-estimates for higher order derivatives. Taking logarithm of both sides of the identity $\frac{\rho_{\nu^{\perp}_{x,i}}}{\rho_\gamma(S_i)} = \partial_{x_i} S_i$ and differentiating the result along x_j one gets

$$\frac{\partial_{x_j} \rho_{\nu^{\perp}_{x,i}}}{\rho_{\nu^{\perp}_{x,i}}} + S_i \cdot \partial_{x_j} S_i = \frac{\partial_{x_i x_j} S_i}{\partial_{x_i} S_i}.$$

Hence

$$\partial_{x_i x_j} S_i = \partial_{x_i} S_i \frac{\partial_{x_j} \rho_{\nu^{\perp}_{x,i}}}{\rho_{\nu^{\perp}_{x,i}}} + S_i \cdot \partial_{x_j} S_i \cdot \partial_{x_i} S_i.$$

Then applying the standard Hölder and Jensen inequalities and using that $S_i \in L^N(\eta)$ for every $N > 0$, we get trivially the following bound.

Proposition 3.15. *For every $p > 1$ and $\varepsilon > 0$ there exists $C(p,\varepsilon)$ such that under assumptions that $\beta_k \in L^p(\nu)$ for all k, one has*

$$\Big\| \frac{\partial_{x_i x_j} S_i}{\partial_{x_i} S_i} \Big\|_{L^p(\nu)} \le C(p,\varepsilon) \|\beta_j\|_{L^{p+\varepsilon}(\nu)}$$

and

$$\|\partial_{x_i x_j} S_i\|_{L^p(\nu)} \le C(p,\varepsilon) \ \||\beta_i|^p\|_{L^{2+\varepsilon}(\nu)} \cdot \||\beta_j|^p\|_{L^{2+\varepsilon}(\nu)}.$$

See also Remark 3.13.

It remains to estimate $\|\partial_{x_j x_m} S_i\|_{L^p(\nu)}$ for $j \ne i, m \ne i$.

Proposition 3.16. *Let $j < m < i$ and $p > 1$. Assume that $\beta_j, \beta_m \in L^{2p}(\nu)$ and β_j admits partial Sobolev derivative $\partial_{x_m} \beta_j \in L^p(\nu)$.*

Then there exists $C(p)$ such that

$$\int |\partial_{x_j x_m} S_i|^p \ d\nu \le C(p) \int \Big(\beta_j^{2p} + \beta_m^{2p} + |\partial_{x_m} \beta_j|^p \Big) \ d\nu.$$

Proof. In the same way as above the proof is reduced to the case where the densities are smooth and positive and admits integrable derivatives (see Proposition 3.10).

For simplicity let us consider only the case $p = 2$. We use relation (21): $\partial_{x_j} S_i = \partial_{x_i} S_i \frac{f}{\rho_{\nu^{\perp}_{x,i}}}$ with

$$f = \frac{\int_{-\infty}^{x_i} \partial_{x_j} \rho_{\nu_i}(x,t)\ dt}{\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t)\ dt} - \frac{\int_{-\infty}^{x_i} \rho_{\nu_i}(x,t) dt\ \int_{-\infty}^{\infty} \partial_{x_j} \rho_{\nu_i}(x,t)\ dt \cdot}{\Bigl(\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t)\ dt\Bigr)^2}. \tag{23}$$

Differentiating $\partial_{x_j} S_i = \partial_{x_i} S_i \frac{f}{\rho_{\nu^{\perp}_{x,i}}}$ along x_m we get

$$\begin{aligned}
\partial_{x_j x_m} S_i =& \partial_{x_i x_m} S_i \frac{f}{\rho_{\nu^{\perp}_{x,i}}} + \partial_{x_i} S_i \cdot \frac{\partial}{\partial x_m}\Bigl[\frac{f}{\rho_{\nu^{\perp}_{x,i}}}\Bigr] = \partial_{x_j} S_i \frac{\partial_{x_i x_m} S_i}{\partial_{x_i} S_i} + \partial_{x_i} S_i \cdot \frac{\partial}{\partial x_m}\Bigl[\frac{f}{\rho_{\nu^{\perp}_{x,i}}}\Bigr] \\
&= \partial_{x_j} S_i \frac{\partial_{x_i x_m} S_i}{\partial_{x_i} S_i} - \partial_{x_j} S_i \frac{\partial_{x_m} \rho_{\nu^{\perp}_{x,i}}}{\rho_{\nu^{\perp}_{x,i}}} + \partial_{x_i} S_i \frac{\partial_{x_m} f}{\rho_{\nu^{\perp}_{x,i}}}.
\end{aligned}$$

The bounds for the first two terms follow immediately from the previous estimates. Let us estimate $\int \Bigl(\partial_{x_i} S_i \frac{\partial_{x_m} f}{\rho_{\nu^{\perp}_{x,i}}}\Bigr)^2$. One has

$$\begin{aligned}
&\int (\partial_{x_i} S_i)^2 \frac{f^2_{x_m}}{\rho^2_{\nu^{\perp}_{x,i}}}\ d\nu = \int \frac{\partial_{x_i} S_i}{\rho_\gamma(S_i)} \frac{f^2_{x_m}}{\rho_{\nu^{\perp}_{x,i}}}\ d\nu = \int\int \frac{\partial_{x_i} S_i}{\rho_\gamma(S_i)} f^2_{x_m}\ d\nu_{i-1}\ dx_i \\
&= -\int\int \frac{\partial_{x_i} S_i}{\rho_\gamma(S_i)} S_i^2 f^2_{x_m}\ d\nu_{i-1}\ dx_i - 2\int\int \frac{S_i}{\rho_\gamma(S_i)} f_{x_m} f_{x_i x_m}\ d\nu_{i-1}\ dx_i \\
&\le \int\int \frac{f^2_{x_i x_m}}{\rho_\gamma(S_i)\partial_{x_i} S_i}\ d\nu_{i-1}\ dx_i = \int \Bigl(\frac{f_{x_i x_m}}{\rho_{\nu^{\perp}_{x,i}}}\Bigr)^2\ d\nu_i.
\end{aligned}$$

Differentiating (23) one gets

$$\begin{aligned}
\frac{f_{x_i x_m}}{\rho_{\nu^{\perp}_{x,i}}} =& \frac{\partial_{x_m x_j} \rho_{\nu_i}}{\rho_{\nu_i}} - \frac{\partial_{x_j} \rho_{\nu_i}}{\rho_{\nu_i}} \frac{\int_{-\infty}^{\infty} \partial_{x_m} \rho_{\nu_i}(x,t) dt}{\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t)\ dt} - \frac{\partial_{x_m} \rho_{\nu_i}}{\rho_{\nu_i}} \frac{\int_{-\infty}^{\infty} \partial_{x_j} \rho_{\nu_i}(x,t) dt}{\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t)\ dt} \\
&- \frac{\int_{-\infty}^{\infty} \partial_{x_j x_m} \rho_{\nu_i}(x,t) dt}{\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t)\ dt} + 2\frac{\int_{-\infty}^{\infty} \partial_{x_m} \rho_{\nu_i}(x,t) dt\ \int_{-\infty}^{\infty} \partial_{x_j} \rho_{\nu_i}(x,t)\ dt}{\Bigl(\int_{-\infty}^{\infty} \rho_{\nu_i}(x,t)\ dt\Bigr)^2}.
\end{aligned}$$

Arguing as above, we easily get

$$\mathbb{E}\Bigl(\frac{f_{x_i x_m}}{\rho_{\nu^{\perp}_{x,i}}}\Bigr)^2 \le C \int \Bigl(\beta_j^4 + \beta_m^4 + (\partial_{x_m}\beta_j)^2\Bigr)\ d\nu.$$

Hence the proof is complete. □

4. Transfer of solutions

We consider in this section a probability measure ν on the space X, where $X = \mathbb{R}^d$ or $X = \mathbb{R}^\infty$. We denote by γ the standard Gaussian measure if $X = \mathbb{R}^d$ and the product of the standard Gaussian measures on $\mathbb{R}^1$

$$\gamma = \prod_{i=1}^{\infty} \gamma_i(dx_i),$$

if $X = \mathbb{R}^\infty$.

Everywhere in this section S is the triangular mapping pushing forward ν onto γ. As usual, we set: $T = S^{-1}$ and $c = DS(T) \cdot b(T)$.

It will be assumed throughout that ν admits logarithmic derivatives $\beta_i \in L^p(\nu)$, $i \in \mathbb{N}$, at least for some $p > 1$ (independent on i). Thus by Remark 3.14 the functions S_i are all Sobolev, more precisely $S_i \in L^n(\nu)$ for all n and $|\nabla S_i| \in L^p(\nu)$.

We also apply systematically the

Chain rule: for every $f \in W^{1,p}(\nu)$ and every smooth compactly supported function φ on $\mathbb{R}$ one has $\varphi(f) \in W^{1,p}(\nu)$ (see Lemma 2.6.9 [9]).

We also need the following important fact (see Theorem 2.6.11 [9]).

Theorem 4.1. *Assume that $d < \infty$, $p \geq 1$. The set of smooth compactly supported functions is dense in the weighted Sobolev space $W^{1,p}(\nu)$, $\nu = \rho\, dx$ provided* $\log \rho \in W^{1,p}(\nu)$.

Everywhere below β_i is the logarithmic derivative of ν along e_i.

Lemma 4.2. *Assume that $X = \mathbb{R}^\infty$ and $\beta_i \in L^p(\nu)$, $b_i \in L^{p^*}(\nu)$ for some $p > 1$ and all i. Assume that $\mathrm{div}_\nu b \in L^1(\nu)$. Then every $c_i \in L^1(\gamma)$ and, in addition, c admits a divergence and the following relation holds: $\mathrm{div}_\gamma c \circ S = \mathrm{div}_\nu b$.*

Proof. It is easy to see that for every $\varphi \in C_0^\infty(\mathbb{R}^n)$ the function $\varphi(S)$ is cylindrical and belongs to $W^{1,p}(\nu)$ by the chain rule. Clearly $\nabla\varphi(S) = (DS)^*\nabla\varphi(S)$. By Theorem 4.1 there exists a sequence of $C_0^\infty(\mathbb{R}^n)$-functions $\{\psi_k\}$ such that $\psi_k \to \varphi(S)$ and $\nabla\psi_k \to \nabla\bigl[\varphi(S)\bigr]$ in $L^p(\nu)$. This, in particular, implies that the relation $-\int f \cdot \mathrm{div}_\nu b\, d\nu = \int \langle \nabla f, b\rangle\, d\nu$ holds for $f = \varphi(S)$. The assumptions of this lemma now imply that $c_i \in L^1(\gamma)$ for every i. Hence

$$\begin{aligned}\int \langle \nabla\varphi, c\rangle\, d\gamma &= \int \langle \nabla\varphi, c\rangle \circ S\, d\nu = \int \langle (DS^*)^{-1}\nabla(\varphi(S)), c(S)\rangle\, d\nu \\ &= \int \langle \nabla(\varphi(S)), b\rangle\, d\nu = -\int \varphi(S)\mathrm{div}_\nu b\, d\nu = -\int \varphi\ \mathrm{div}_\nu b(T)\, d\gamma.\end{aligned}$$

This implies the last assertion. $\square$

Lemma 4.3. *Assume that $X = \mathbb{R}^d$ and $\beta_i \in L^p(\nu)$, for any $p \geq 1$ and all $i \in \{1, \cdots, d\}$. Assume, in addition, that $\frac{1}{\rho_\nu}$ is locally integrable in any power. Then for every $\varphi \in C_0^\infty(\mathbb{R}^n)$ there exists a sequence of $C_0^\infty(\mathbb{R}^n)$-functions $\{\psi_k\}$ such that $\psi_k \to \varphi(T)$ and $\nabla\psi_k \to \nabla\bigl[\varphi(T)\bigr]$ γ-almost everywhere and weakly in $L^n(\gamma)$ for any $n \geq 1$.*

Proof. In the same way as in Lemma 4.2 we show first that $\varphi(T)$ belongs to $W^{1,n}(\gamma)$ for any $n \geq 1$. Then one can find approximating sequences for any n and choose the desired one by a diagonal argument. To show that $\varphi(T)$ belongs to $W^{1,n}(\gamma)$ we apply Remark 3.14. One gets that the functions $\langle DS^* e_i, e_j\rangle$ are integrable in any power. One has

$$\int \|\nabla\bigl[\varphi(T)\bigr]\|^n\, d\gamma \leq \int \|DT^* \cdot \nabla\varphi(T)\|^n\, d\gamma = \int \|(DS^*)^{-1} \cdot \nabla\varphi\|^n\, d\nu.$$

It remains to show that all the functions $\langle (DS^*)^{-1} e_i, e_j\rangle$ are locally ν-integrable in any power. Since $\langle DS^* e_i, e_j\rangle$ admits the same property, we only need to show the local integrability of $\frac{1}{(\det S)^n}$. Taking into account that $\frac{1}{\det S} = \frac{\rho_\gamma(S)}{\rho_\nu}$ (see (22)) and the assumptions of this lemma we only need to show that S is locally bounded. Applying again the local integrability of $\frac{1}{\rho_\nu}$ we get that S belongs to the standard local Sobolev class $W^{1,p}_{loc}$ for every p. Then the boundedness of S follows from the classical embedding theorem. $\square$

Proposition 4.4. *Assume that $X = \mathbb{R}^\infty$. Let $\rho(t,x)$ be a solution to the equation $\dot\rho + \operatorname{div}_\nu(\rho \cdot b) = 0$. Assume that there exists $p > 1$ such that $\sup_{t\in[0,T]} \|\rho(t,\cdot) \cdot b_i\|_{L^{p^*}(\nu)} < \infty$ and, in addition, $\beta_i \in L^p(\nu)$ for all i.*

Then the function $g(t,x)$ defined by the relation $g\cdot\gamma = (\rho\cdot\nu)\circ S^{-1}$ is the solution to the equation

$$\dot g + \operatorname{div}_\gamma(g \cdot c) = 0$$

Proof. We know that

$$\int \varphi\rho(t,x)\ d\nu = \int \varphi\rho(0,x)\ d\nu + \int_0^t \int \langle b, \nabla\varphi\rangle \rho(s,x)\ d\nu\ ds, \quad \text{for all } t \in [0,T].$$

Let us apply this identity to the function $\varphi = \psi(S)$, $\psi \in C_0^\infty(\mathbb{R}^n)$. This is possible, because $\nabla\varphi \in L^p(\nu)$ by Proposition 3.10 and one can approximate φ and its gradient in $L^p(\nu)$ by smooth cylindrical functions (see the proof of Lemma 4.2).

By the change of variables formula

$$\int \varphi\rho(t,x)\ d\nu = \int \psi g(t,x)\ d\gamma, \quad \int \varphi\rho(0,x)\ d\nu = \int \psi g(0,x)\ d\gamma.$$

Taking into account the chain rule $\nabla\varphi = (DS)^*\nabla\psi(S)$ we immediately get for all $t \in [0,T]$

$$\int_0^t \int \langle b, \nabla\varphi\rangle\rho(s,x)\ d\nu\ ds = \int_0^t \int \langle DS{\cdot}b, \nabla\psi(S)\rangle\rho(s,x)\ d\nu\ ds = \int_0^t \int \langle c, \nabla\psi\rangle\ g(s,x)\ d\gamma\ ds.$$

Hence g satisfies the desired integral relation and the proof is complete. $\square$

Proposition 4.5. *Assume that $X = \mathbb{R}^\infty$ and the following assumptions hold*

1) *$\beta_i \in L^m(\nu)$, for all $m, i \in \mathbb{N}$*
2) *$\frac{1}{\rho_{\nu_n}}$ is locally integrable in any power for every $n \in \mathbb{N}$, where ρ_{ν_n} is the Lebesgue density of the projection $\nu \circ P_n^{-1} = \rho_{\nu_n}\ dx$*

Assume, in addition, that that g solves the equation $\dot g + \operatorname{div}_\gamma(g\cdot c) = 0$ for some c satisfying $\sup_{t\in[0,T]} \|g(t,\cdot)\cdot c_i\|_{L^{1+\varepsilon}(\gamma)} < \infty$ for some $\varepsilon > 0$ and all i.

Then the function ρ defined by the relation $\rho\cdot\nu = (g\cdot\gamma)\circ T^{-1}$ is the solution to the equation

$$\dot\rho + \operatorname{div}_\nu(\rho\cdot b) = 0.$$

Proof. We apply the same arguments as in the proof of the previous proposition. We note that $\int_0^t \int \langle b, \nabla\varphi\rangle\rho(s,x)\ d\nu\ ds$ is well-defined for every $\varphi \in \mathcal{C}$, $t \in [0,T]$, because $b_i = \langle (DS)^{-1}c(S), e_i\rangle$, any function $\langle (DS)^{-1}e_i, e_j\rangle$ (depending on a finite number of variables) is locally integrable in any power by the previous proposition and

$$\sup_{0\le s\le T} \|c_i(S)\rho(s,\cdot)\|_{L^{1+\varepsilon}(\nu)} < \infty$$

by the change of variables formula.

The relation

$$\int_0^t \int \langle c, \nabla\psi\rangle g(s,x)\ d\gamma\ ds = \int_0^t \int \langle b, \nabla\varphi\rangle\rho(s,x)\ d\nu\ ds,\ t\in[0,T],$$

with $\psi = \varphi(T)$, $\varphi \in \mathcal{C}$, can be easily justified with the help of Lemma 4.3. Indeed, since ψ can be approximated in the corresponding Sobolev norm by smooth functions, the integral in the left hand side is well-defined and the chain rule is applicable. $\square$

5. Existence

In this section we prove the existence result by transferring a solution in the Gaussian case (whose existence was established in [7]) with the help of a triangular mapping.

Theorem 5.1. *Assume that ν is a probability measure on $\mathbb{R}^\infty$ such that:*

1)
$$\beta_i \in L^m(\nu) \quad \textit{for all } m, i \in \mathbb{N};$$
2) *there exists $p > 1$ such that*
$$b_i \in L^p(\mu) \quad \textit{for all } i \in \mathbb{N};$$
3) *there exists $\varepsilon > 0$ such that*
$$\exp\bigl(\varepsilon(\mathrm{div}_\nu b)_-\bigr) \in L^1(\nu);$$
4) *$\frac{1}{\rho_{\nu_n}}$ is locally integrable in any power for every $n \in \mathbb{N}$, where ρ_{ν_n} is the Lebesgue density of the projection $\nu \circ P_n^{-1} = \rho_{\nu_n}\, dx$.*

Then for every $\rho_0 \in L^{q'}(\nu)$, and $\tilde{q}$ with $q' > \tilde{q} > p^*$ there exists $t_0 > 0$ depending on the above parameters such that the equation

$$\dot{\rho} + \mathrm{div}_\nu\bigl(b \cdot \rho\bigr) = 0$$

has a solution on $[0, t_0]$ satisfying $\rho|_{t=0} = \rho_0$ and

$$\sup_{t \in [0, t_0]} \|\rho(t, \cdot)\|_{L^{\tilde{q}}} < \infty. \tag{24}$$

Remark 5.2. One can easily see that assumptions 1), 4) together with Sobolev embedding imply that $1/\rho_{\nu_n}$ is Hölder continuous. This may be sometimes restrictive for applications. We stress that we need 1) and 4) mainly for a-priori estimates on DT (see Lemma 4.3). There are some possibilities to weaken these assumption. Some (weaker) sufficient conditions for T to be locally Sobolev one can find in [29]. This result is applicable if one has high integrability of ρ_0 and b_i. Some bounds on DT are available under the assumption that ν is log-concave. They work even better if instead of triangular mapping one applies optimal transportation. See Theorem 8.4 and Example 8.5 below.

Proof. Consider the triangular mapping T sending γ to ν. Let us show that $c = DS(T) \cdot b(T)$ satisfy all the assumptions of Lemma 2.10. One has

$$c_i = \sum_{j=1}^{i} \frac{\partial S_i}{\partial x_j}(T) b_j(T).$$

It follows immediately from the assumption of this theorem and Remark 3.14 that $c_i \in L^{p'}(\nu)$ for every i and $p' < p$.

By Lemma 4.2 $\mathrm{div}_\gamma c \circ S = \mathrm{div}_\nu b$. Consequently, the assumption 2) of Lemma 2.10 is satisfied. Hence, there exists a solution to the equation $\dot{g} + \mathrm{div}_\gamma(g \cdot c) = 0$

with $g(0,x)=\rho(0,T(x))$. Proposition 4.5 now implies that $\rho = g(S)$ is the desired solution.

Property (24) is a slight extension of the corresponding statement of Lemma 2.10 and can be easily checked. Hence the proof is complete. $\square$

Example 5.3. Let us give an example of a probability measure on $\mathbb{R}^\infty$ with integrable logarithmic derivatives which is typical for applications and satisfies the assumption of the above theorem. We consider Gibbs measures on a lattice $\mathbb{R}^{\mathbb{Z}^d}$, which can be formally written in the following way

$$\nu ='' \exp\Big(-\sum_{k\in\mathbb{Z}^d} V_k(x_k) - \sum_{k,j\in\mathbb{Z}^d} W_{k,j}(x_k,x_j)\Big)\ dx'', \tag{25}$$

where $''dx''$ denotes infinite-dimensional Lebesgue measure on $\mathbb{R}^{\mathbb{Z}^d}$ (which does not exist). The following existence result has been established in [1]. Assume that there exist a number $N\geq 2$ and a symmetric matrix $J=\{J_{k,j}\}_{k,j\in\mathbb{Z}^{\mathbb{R}^d}}$ such that

$$W_{k,j}(x_k,x_j) = W_{j,k}(x_j,x_k)$$

$$|W_{k,j}(x_k,x_j)| \leq J_{k,j}(1+|x_k|+|x_j|)^N$$

$$|\partial_{x_k} W_{k,j}(x_k,x_j)| \leq J_{k,j}(1+|x_k|+|x_j|)^{N-1}$$

$$|V_k(x_k)| \leq C(1+|x_k|)^L, \quad |\partial_{x_k} V_k(x_k)| \leq C(1+|x_k|)^{L-1}$$

$$\partial_{x_k} V_k(x_k)\cdot x_k \geq A|x_k|^{N+\sigma} - B$$

for some $A,B,C,\sigma>0, L\geq 1$.

The matrix J is also assumed to be fastly decreasing (see [1] for details), in particular the finite range case $J_{k,j}=0$ if $|j-k|>N_0$ for some N_0 is included. Then there exists a probability ("Gibbs") measure ν on $\mathbb{R}^{\mathbb{Z}^d}$ with exponentially integrable logarithmic derivatives

$$\beta_k = \partial_{x_k} V_k(x_k) + \sum_{j\in\mathbb{Z}^d} \partial_{x_k}(W_{k,j}+W_{j,k}) \quad k\in\mathcal{Z}^d.$$

It was shown in [1] that such ν is a rigorous definition of ν in (25) via the Dobrushin-Lanford-Ruelle equations. See [2] for uniqueness results.

6. Uniqueness in the Gaussian case

The following result was essentially established in [7]. We give below a slightly modified version with a sketch of the proof.

Theorem 6.1. *Assume that there exist $p>1$, $q>1$ such that $\|c\|\in L^p(\gamma)$, $\|Dc\|_{\mathcal{HS}}\in L^q(\gamma)$, $\mathrm{div}_\gamma c\in L^q(\gamma)$.*

Then for every $t_0>0$ there exists at most one solution to (2) satisfying

$$\sup_{0\leq t\leq t_0} \|\rho(t,\cdot)\|_{L^r(\gamma)} < \infty,$$

where $r\geq \max(p^*,q^*)$.

If, in addition, $d<\infty$ and $p\geq q$, the assumption $\|Dc\|_{\mathcal{HS}}\in L^q(\gamma)$ can be replaced by $\|Dc\|\in L^q_{loc}(\gamma)$.

Sketch of the proof. We discuss only the case $d < \infty$ (the proof for $d = \infty$ is almost the same). Fix a non-negative $C_0^\infty(\mathbb{R}^d)$-function φ. Let ρ be a solution to (2). Set: $\hat{\rho} = \varphi \cdot \rho$ (this is important only for $d < \infty$, since in this case we apply the local assumption). Take any $\psi \in C_0^\infty(\mathbb{R}^d)$ such that $\psi = 1$ on $\operatorname{supp}(\varphi)$ and set $\hat{c} = \psi \cdot c$. Note that $\hat{\rho}$ solves the following equation

$$\frac{d}{dt}\hat{\rho} + \operatorname{div}_\gamma(\hat{c}, \hat{\rho}) = \langle \nabla\varphi, \hat{c}\rangle\rho.$$

Let us smoothen $\hat{\rho}$ with the Ornstein-Uhlenbeck semigroup: $\rho^\varepsilon = e^{-\varepsilon}T_\varepsilon(\hat{\rho})$. One has

$$\frac{d}{dt}\rho^\varepsilon + \operatorname{div}_\gamma(\hat{c}\cdot\rho^\varepsilon) = e^{-\varepsilon}r^\varepsilon(\hat{\rho}, \hat{c}) + \rho^\varepsilon \operatorname{div}_\gamma(\hat{c}) + e^{-\varepsilon}T_\varepsilon(\langle\nabla\varphi, \hat{c}\rangle\rho),$$

where $r^\varepsilon(v, b) = e^\varepsilon\langle b, \nabla(T_\varepsilon(v)) - T_\varepsilon(\operatorname{div}_\gamma(v\cdot b)).$

The uniqueness proof relies on the concept of the so-called renormalizing solutions. Take a continuously differentiable globally Lipschitz function v and using smoothness of ρ^ε compute $\frac{d}{dt}v(\rho^\varepsilon)$:

$$\frac{d}{dt}v(\rho^\varepsilon) + \operatorname{div}_\gamma(\hat{c}\cdot v(\rho^\varepsilon)) = \bigl[v(\rho^\varepsilon) - \rho^\varepsilon v'(\rho^\varepsilon)\bigr]\cdot \operatorname{div}_\gamma(\hat{c})$$
$$+ v'(\rho^\varepsilon)\bigl(e^{-\varepsilon}r^\varepsilon(\hat{\rho}, \hat{c}) + \rho^\varepsilon\operatorname{div}_\gamma(\hat{c}) + e^{-\varepsilon}T_\varepsilon(\langle\nabla\varphi, \hat{c}\rangle\rho)\bigr).$$

According to estimate (68) from [7] there exists $C = C(p, q)$ such that for $r = \max(p^*, q^*)$ and small values of ε one has

$$\|r^\varepsilon\|_{L^1(\gamma)} \le C\|\hat{\rho}\|_{L^r(\gamma)}(\sqrt{\varepsilon}\|\hat{c}\|_{L^p(\gamma)} + \|\operatorname{div}_\gamma\hat{c}\|_{L^q(\gamma)} + \|(D\hat{c})^{sym}\|_{L^q(\gamma)}).$$

It follows from the assumptions of this theorem that the right-hand side is finite. In addition, $r^\varepsilon \to -\operatorname{div}_\gamma(\hat{c})\cdot\hat{\rho}$ in $L^1(\gamma)$ as $\varepsilon \to 0$ (Proposition 3.5 in [7]).

Passing to the limit one obtains that

$$\frac{d}{dt}v(\hat{\rho}) + \operatorname{div}_\gamma(\hat{c}\cdot v(\hat{\rho})) = \bigl[v(\hat{\rho}) - \hat{\rho}v'(\hat{\rho})\bigr]\cdot\operatorname{div}_\gamma(\hat{c}) + v'(\hat{\rho})(\langle\nabla\varphi, \hat{c}\rangle\rho)$$

in the distributional sense (i.e., ρ is a *renormalizing* solution). Assume that there exists two different solutions ρ_1, ρ_2 in $L^r(\gamma)$ with the same initial condition. Applying this relation to the difference $\rho = \rho_1 - \rho_2$ and $v(t) = \max(0, t)$, we get that for every $\varphi \in C_0^\infty(\mathbb{R}^d)$ one has $\frac{d}{dt}(\varphi\cdot\rho_+) + \operatorname{div}_\gamma(c\cdot(\rho)_+\varphi) = \langle\nabla\varphi, c\rangle\rho_+$. Finally,

$$\frac{d}{dt}\rho_+ + \operatorname{div}_\gamma(c\cdot(\rho)_+) = 0$$

in the distributional sense. Clearly, $\frac{d}{dt}\int\rho_+\,d\gamma = 0$, hence $\int\rho_+\,d\gamma = 0$ and $\rho = 0.\square$

7. Examples of uniqueness

In this section we study uniqueness problem for transport equations. As in Theorem 5.1 we reduce the proof to the Gaussian case (see Theorem 6.1).

Recall that

$$c = DS(T)\cdot b(T).$$

Since the assumption on the divergence can be directly transferred, we need only to find some sufficient conditions for

$$\|c\|,\ \|Dc\|_{HS} \in L^p(\gamma).$$

One can try to apply the trivial operator norm estimate

$$\|c\| \le \|DS(T)\|\|b(T)\|.$$

Let us stress, however, that operator norm estimates do not seem to be available in the case of triangular mappings (unlike optimal transportation ones). In spite of this let us give another estimate of c which does not use operator norms.

Lemma 7.1. *For every $1 \le p \le 2$ and $q \ge 1$ one has*

$$\int \|c\|^p \ d\gamma \le C(p,q)\Bigl(\sup_i \int \beta_i^{pq} \ d\nu\Bigr)^{\frac{1}{q}} \cdot \Bigl[\sum_i^{\infty}\Bigl(\int |b_i|^{pq^*} \ d\nu\Bigr)^{\frac{1}{2q^*}}\Bigr]^2.$$

Proof. Trivially we have

$$\|c\| \le \sum_{i=1}^{\infty} |b_i(T)| \cdot \|\partial_{e_i} S(T)\|.$$

Applying the inequality

$$\sum_i |a_i| \le \bigl(\sum_i |a_i|^{\frac{1}{q}}\bigr)^q$$

which holds for every $q \ge 1$, we get for every $1 \le p \le 2$

$$\begin{aligned}\|c\|^p &= \Bigl(\sum_{i=1}^{\infty} |b_i(T)| \cdot \|\partial_{e_i} S(T)\|\Bigr)^p \le \Bigl(\sum_{i=1}^{\infty} |b_i(T)|^{\frac{p}{2}} \cdot \|\partial_{e_i} S(T)\|^{\frac{p}{2}}\Bigr)^2 \\ &= \sum_{i,j=1}^{\infty} |b_i(T)|^{\frac{p}{2}} \cdot |b_j(T)|^{\frac{p}{2}} \cdot \|\partial_{e_i} S(T)\|^{\frac{p}{2}} \cdot \|\partial_{e_j} S(T)\|^{\frac{p}{2}}.\end{aligned}$$

By the Hölder inequality and Proposition 3.10

$$\begin{aligned}\int \|c\|^p \ d\gamma &\le \sum_{i,j=1}^{\infty} \Bigl(\int |b_i(T)|^{\frac{pq^*}{2}} \cdot |b_j(T)|^{\frac{pq^*}{2}} \ d\gamma\Bigr)^{\frac{1}{q^*}} \Bigl(\int \|\partial_{e_i} S(T)\|^{\frac{pq}{2}} \cdot \|\partial_{e_j} S(T)\|^{\frac{pq}{2}} \ d\gamma\Bigr)^{\frac{1}{q}} \\ &\le \sum_{i,j=1}^{\infty} \Bigl(\int |b_i|^{pq^*} \ d\nu\Bigr)^{\frac{1}{2q^*}} \cdot \Bigl(\int |b_j|^{pq^*} \ d\nu\Bigr)^{\frac{1}{2q^*}} \cdot \Bigl(\int \|\partial_{e_i} S\|^{pq} \ d\nu\Bigr)^{\frac{1}{2q}} \cdot \Bigl(\int \|\partial_{e_j} S\|^{pq} \ d\nu\Bigr)^{\frac{1}{2q}} \\ &\le C(p,q)\Bigl(\sup_i \int \beta_i^{pq} \ d\nu\Bigr)^{\frac{1}{q}} \cdot \Bigl[\sum_i^{\infty}\Bigl(\int |b_i|^{pq^*} \ d\nu\Bigr)^{\frac{1}{2q^*}}\Bigr]^2.\end{aligned}$$

□

7.1. **Product case.**

Theorem 7.2. *Assume that ν is a product measure on $\mathbb{R}^\infty$*

$$\nu = \prod_{i=1}^{\infty} e^{-w_i(x_i)} \ dx_i.$$

Assume that for some $\varepsilon > 0, \delta > 0, 1 < p \le 2, q > 1$

1)

$$t\omega_i'(t) \ge \frac{(-1+\delta)\varepsilon}{pq^*(pq^*+\varepsilon)-\varepsilon} \quad \text{for every } i$$

and

$$\sup_i \int \beta_i^{pq^*+\varepsilon} \ d\nu < \infty;$$

2) $\|Db\|_{HS} \in L^p(\nu)$ *and*

$$\sum_{i=1}^{\infty}\Bigl(\int |b_i|^{pq}\ d\nu\Bigr)^{\frac{1}{q}} + \sum_{i\neq j}\Bigl(\int |\partial_{x_j} b_i|^{pq}\ d\nu\Bigr)^{\frac{1}{q}} < \infty;$$

3) $\mathrm{div}_\nu(b) \in L^p(\nu)$.

Then $\|c\|,\ \|Dc\|_{HS} \in L^p(\gamma)$. *In particular, for every* $t_0 > 0$ *there exists at most one solution to the equation (2) satisfying*

$$\sup_{0\le t\le t_0} \|\rho(t,\cdot)\|_{L^{p^*}(\nu)} < \infty.$$

Proof. First we check that the assumptions of Lemma 4.2 and Proposition 4.4 are satisfied. This is clear except for the estimate $\sup_{0\le t\le t_0} \|\rho\cdot b_i\|_{L^{(pq^*)^*}} < \infty$. To prove this we apply the Hölder inequality $\|\rho\cdot b_i\|_{L^{(pq^*)^*}} = \|\rho\cdot b_i\|_{L^{pq/(pq-q+1)}} \le \|\rho\|_{L^{p^*}}\|b_i\|_{L^{pq}}$.

By Theorem 6.1 and Proposition 4.4 the problem is now reduced to the uniqueness problem in the Gaussian case.

Thus, it is sufficient to show that $\|c\|,\ \|Dc\|_{HS} \in L^p(\gamma)$ for some $p > 1$. Since we deal with a product measure, the transportation mapping has a simple structure

$$T = (T_1(x_1), T_2(x_2), \cdots, T_n(x_n), \cdots).$$

Hence

$$c_i = \partial_{x_i} S_i(T_i) b_i(T).$$

We apply Lemma 7.1. Note that that in this case $\partial_{e_j} S_i = 0$. Taking this into account and following the proof of Lemma 7.1 we can get a more precise estimate:

$$\int \|c\|^p d\gamma \le C(p,q)\sum_{i=1}^{\infty}\Bigl(\int |b_i|^{pq}\ d\nu\Bigr)^{\frac{1}{q}}\cdot\Bigl(\int \beta_i^{pq^*}\ d\nu\Bigr)^{\frac{1}{q^*}}$$
$$\le C(p,q)\sup_i\Bigl(\int \beta_i^{pq^*}\ d\nu\Bigr)^{\frac{1}{q^*}}\sum_{i=1}^{\infty}\Bigl(\int |b_i|^{pq}\ d\nu\Bigr)^{\frac{1}{q}},\quad 1<p<2.$$

Let us estimate DC. Taking into account that S_i and T_i are reciprocal, one easily gets

$$\partial_{x_i} c_i = \Bigl(\frac{\partial_{x_i x_i} S_i}{\partial_{x_i} S_i}\cdot b_i + \partial_{x_i} b^i\Bigr)\circ T.$$

and

$$\partial_{x_j} c_i = \Bigl(\frac{\partial_{x_i} S_i}{\partial_{x_j} S_j}\partial_{x_j} b^i\Bigr)\circ T,\quad i\neq j.$$

One can estimate $\|(\partial_{x_i} c_i)\|$ in the same way as $\|c\|$ and applying Proposition 3.15 one gets

$$\int \|(\partial_{x_i} c_i)\|^p\ d\gamma \le C(p)\Bigl[\int \|(\partial_{x_i} b^i)\|^p\ d\nu + \sup_i\Bigl(\int \bigl(\frac{\partial_{x_i x_i} S_i}{\partial_{x_i} S_i}\bigr)^{pq^*}\ d\nu\Bigr)^{\frac{1}{q^*}}\sum_{i=1}^{\infty}\Bigl(\int |b_i|^{pq}\ d\nu\Bigr)^{\frac{1}{q}}\Bigr]$$
$$\le C(p,q)\Bigl[\int \|(\partial_{x_i} b^i)\|^p\ d\nu + \sup_i\Bigl(\|\beta_i\|_{L^{pq^*+\varepsilon}(\nu)}\Bigr)^p\sum_{i=1}^{\infty}\Bigl(\int |b_i|^{pq}\ d\nu\Bigr)^{\frac{1}{q}}\Bigr].$$

Similarly

$$\int \|B\|^p_{\mathcal{HS}}\ d\gamma \le \sup_{i,j}\Bigl(\int \Bigl[\frac{\partial_{x_i} S_i}{\partial_{x_j} S_j}\Bigr]^{pq^*}\ d\nu\Bigr)^{\frac{1}{q^*}}\sum_{i\neq j}\Bigl(\int |\partial_{x_j} b_i|^{pq}\ d\nu\Bigr)^{\frac{1}{q}},$$

where $B_{i,j} = \partial_{x_i} c_i$, $i \neq j$, $B_{i,i} = 0$.

It remains to estimate

$$\int \Bigl[\frac{\partial_{x_i} S_i}{\partial_{x_j} S_j}\Bigr]^{pq^*} d\nu \leq \Bigl[\int \bigl[\partial_{x_i} S_i\bigr]^{pq^*+\varepsilon} d\nu\Bigr]^{\frac{pq^*}{pq^*+\varepsilon}} \Bigl[\int \Bigl[\frac{1}{\partial_{x_j} S_j}\Bigr]^{\frac{pq^*(\varepsilon+pq^*)}{\varepsilon}} d\nu\Bigr]^{\frac{\varepsilon}{pq^*+\varepsilon}}$$
$$= \Bigl[\int \bigl[\partial_{x_i} S_i\bigr]^{pq^*+\varepsilon} d\nu\Bigr]^{\frac{pq^*}{pq^*+\varepsilon}} \Bigl[\int \bigl[\partial_{x_j} T_j\bigr]^{\frac{pq^*(\varepsilon+pq^*)}{\varepsilon}} d\gamma\Bigr]^{\frac{\varepsilon}{pq^*+\varepsilon}}.$$

According to Proposition 3.7 and Proposition 3.10 we get that the right-hand side is finite if $|\beta_i|_{L^{pq^*+\varepsilon}(\nu)} < \infty$ and $t\omega_i'(t) \geq \frac{(-1+\delta)\varepsilon}{pq^*(\varepsilon+pq^*)-\varepsilon}$. The proof is complete. $\square$

Corollary 7.3. *Let ν be as in Theorem 7.2 satisfying $tw_i'(t) \geq 0$. Assume that*

1)
$$\sup_i \int |w_i'(t)|^N e^{-w_i(t)} dt < \infty$$
for every $N > 1$;

2) *for some $1 < p \leq 2$ and $q > 1$ one has $(\partial_{x_i} b_i) \in L^p(\nu)$;*

3) $\operatorname{div}_\nu(b) \in L^p(\nu)$ *and*
$$\sum_{i=1}^{\infty} \Bigl(\int |b_i|^{pq}\, d\nu\Bigr)^{\frac{1}{q}} + \sum_{i\neq j} \Bigl(\int |\partial_{x_j} b_i|^{pq}\, d\nu\Bigr)^{\frac{1}{q}} < \infty.$$

Then $\|c\|$, $\|Dc\|_{HS} \in L^p(\gamma)$. In particular, for every $t_0 > 0$ there exists at most one solution to the equation (2) satisfying

$$\sup_{0\leq t\leq t_0} \|\rho(t,\cdot)\|_{L^{p^*}(\nu)} < \infty.$$

7.2. **Gibbs measures.** In this section we prove uniqueness for the measures described in Example 5.3. More generally, we will assume:

Assumption (A): There exist smooth functions $V_i(x_i), W_{i,j}(x_i, x_j)$ such that

$$\beta_i = V_i'(x_i) + \sum_{j=1}^{\infty} \partial_{x_i} W_{i,j} \tag{26}$$

and there exists $N_0 \geq 1$ such that $W_{i,j} = 0$ if $|i-j| > N_0$.

Clearly, in this case the corresponding mapping S has a special structure

$$S(x_1, \cdots, x_n, \cdots) = (S_1(x_1), S_2(x_1, x_2), \cdots, S_n(x_{n-N_0}, \cdots, x_{n-1}, x_n), \cdots).$$

Example 7.4. Let us consider a Gibbs measure $\nu ='' e^{-H} dx''$ with Hamiltonian

$$H = \sum_{i=1}^{\infty} V_i(x_i) + \sum_{i,j=1}^{\infty} W_{i,j}(x_i, x_j).$$

Under the assumptions of Remark 5.3 there exists a unique Gibbs measure ν satisfying (26), as explained above.

We will also need the following 1-dimensional version of the Caffarelli contraction theorem (which holds true for optimal transport mappings in any dimension, see [18], [23]).

Theorem 7.5. *Let $T : \mathbb{R} \to \mathbb{R}$ be the canonical increasing mapping pushing forward a probability measure $e^{-V(x)}\, dx$ onto a probability measure $e^{-W(x)}\, dx$. Assume that V and W are twice continuously differentiable and $V'' \le C$, $W'' \ge K$. Then T is Lipschitz satisfying $T' \le \sqrt{\frac{C}{K}}$*

We recall that a probability measure μ on $\mathbb{R}^d$ is called log-concave if it has the form $e^{-V} \cdot \mathcal{H}^k|_L$, where $\mathcal{H}^k$ is the k-dimensional Haussdorff measure, $k \in \{0, 1, \cdots, d\}$, L is an affine subspace, and V is a convex function. We call a measure μ uniformly log-concave if $\frac{1}{Z} e^{K|x|^2} \cdot \mu$ is a log-concave measure for some $K > 0$ and a suitable renormalization factor Z. It is well-known (C. Borell) that the projections and conditional measures of log-concave measures are log-concave. The same holds for uniformly log-concave measures. We can extend this notion to the infinite-dimensional case. Namely, we call a probability measure μ on a locally convex space X log-concave (uniformly log-concave with $K > 0$) if its images $\mu \circ l^{-1}$, $l \in X^*$, under all linear continuous functionals are all log-concave (uniformly log-concave with $K > 0$). We will also use the fact that the (one-dimensional) conditional measures of uniformly log-concave measures are uniformly log-concave with the same constant.

Theorem 7.6. *Assume that assumption* **(A)** *is satisfied and*

1) *for every $n \ge 1$*

$$\sup_i \int \Big(|\beta_i|^n + \|\nabla \beta_i\|^n \Big)\, d\mu < \infty;$$

2) *ν is uniformly log-concave;*

3) *for some $1 < p_0 \le 2, q_0 > 1$*

$$\sum_{i=1}^{\infty} \Big(\int |b_i|^{p_0 q_0}\, d\nu \Big)^{\frac{1}{2q_0}} < \infty;$$

4) *for some $1 < p < 2$*

$$\sum_{j,k=1}^{\infty} k^{\frac{p}{2}} \Big(\int |\partial_{x_k} b_j|^{\frac{4p}{2-p}}\, d\nu \Big)^{\frac{2-p}{8}} < \infty.$$

Then $\|c\|$, $\|Dc\|_{HS} \in L^q(\gamma)$ for some $q > 1$. If, in addition, $\mathrm{div}_\nu(b) \in L^q(\nu)$, then for every t_0 there exists at most one solution to (2) satisfying $\sup_{0 \le t \le t_0} \|\rho(t, \cdot)\|_{L^{q^*}(\nu)} < \infty$.

Remark 7.7. a) According to results of [1] there exist probability measures satisfying the assumptions of the theorem (see Remark 5.3). In particular, 1) is automatically satisfied for these measures.

b) We believe that the factor $k^{\frac{p}{2}}$ in 4) can be removed. This factor arise just because we deal with triangular transportations and the quantity ∇T_k is difficult to control (see Lemma 7.8). One can get a better control applying optimal transportation mappings. Unfortunately, the existence of such a mapping in the infinite dimensional case for mutually singular measures is an open problem to the best of our knowledge.

c) The assumption of uniformly log-concavity of the Gibbs measure ν can be expressed in terms of the potentials V_i, W_{ij}. Note that it is sufficient to require the uniformly log-concavity of the approximations.

Proof. It is easy to check that b, ν, and ρ satisfy the assumptions of Lemma 4.2, Lemma 4.3, and Proposition 4.4. Thus the problem is reduced to the uniqueness problem in the Gaussian case. We have to show that $\|c\|$, $\|Dc\|_{HS} \in L^q(\gamma)$ for some $q > 1$. The first part follows from Lemma 7.1. The second part follows from Lemmata 7.8, 7.9. Indeed, note that

$$Dc = \bigl(DS \cdot Db \cdot (DS)^{-1}\bigr) \circ T + \Bigl(\sum_{i=1}^{\infty} b_i \cdot (DS_{x_i})(DS)^{-1}\Bigr) \circ T. \tag{27}$$

In the same way as in Lemma 7.1 and applying Lemma 7.9 we get the desired estimate for $\sum_{i=1}^{\infty} b_i (DS_{x_i})(DS)^{-1}$. Note that $\frac{1}{\partial_{x_i} S_i} = \partial_{x_i} T_i(S)$ is bounded by $\frac{1}{\sqrt{K}}$ as a one-dimensional optimal mapping of a Gaussian measure onto a uniformly log-concave measure by the Caffarelli theorem. Here we use the fact that the corresponding conditional measures are uniformly log-concave.

We apply Lemma 7.8 to estimate $DS \cdot Db \cdot (DS)^{-1}$. To complete the proof we need to estimate $\int |\nabla T_k|^2 \, d\gamma$. Indeed, since ν is uniformly log-concave, we can apply Remark 3.13. We get $K \int |\partial_{x_i} T|^2 \, d\nu \le \int x_i^2 \, d\gamma$. Hence

$$\int \|\nabla T_k\|^2 \, d\gamma = \int \sum_{i=1}^{k} (\partial_{x_i} T_k)^2 \, d\gamma \le \sum_{i=1}^{k} \int \|\partial_{x_i} T\|^2 \, d\gamma \le \frac{1}{K} \sum_{i=1}^{k} \int x_i^2 \, d\gamma = \frac{k}{K}.$$

□

Lemma 7.8. *For every $1 < p < 2$ there exists C, depending on p, N_0, and $\sup_i \int |\beta_i|^{\frac{4p}{2-p}} \, d\nu$ such that*

$$\int \|DS \cdot Db \cdot (DS)^{-1}\|^p_{\mathcal{HS}} \le C \Bigl[\sum_{j,k=1}^{\infty} \Bigl(\int |\partial_{x_k} b_j|^{\frac{4p}{2-p}} \, d\nu \Bigr)^{\frac{2-p}{8}} \Bigl(\int \|\nabla T_k\|^2 \, d\gamma \Bigr)^{\frac{p}{4}} \Bigr]^2.$$

Proof. We estimate $\|DS \cdot Db \cdot (DS)^{-1}\|^2_{\mathcal{HS}} = \sum_{i=1}^{\infty} \|DS \cdot Db \cdot (DS)^{-1} \cdot e_i\|^2$.

For simplicity set $M = DS \cdot DB$, $L = (DS)^{-1}$. Then

$$\begin{aligned}
\int \Bigl(\sum_{i=1}^{\infty} \|MLe_i\|^2 \Bigr)^{\frac{p}{2}} \, d\nu &\le \int \Bigl(\sum_{i=1}^{\infty} \Bigl(\sum_{k=1}^{\infty} |L_{i,k}| \|M \cdot e_k\| \Bigr)^2 \Bigr)^{\frac{p}{2}} \, d\nu \\
&\le \int \Bigl[\sum_{i=1}^{\infty} \sum_{k,j=1}^{\infty} |L_{i,k}||L_{i,j}| \|M \cdot e_k\| \|M \cdot e_j\| \Bigr]^{\frac{p}{2}} \, d\nu \le \int \sum_{k,j=1}^{\infty} \sum_{i=1}^{\infty} |L_{i,k}|^{\frac{p}{2}} |L_{i,j}|^{\frac{p}{2}} \|M \cdot e_k\|^{\frac{p}{2}} \|M \cdot e_j\|^{\frac{p}{2}} \, d\nu \\
&\le \sum_{k,j=1}^{\infty} \Bigl(\int \|M \cdot e_k\|^{\frac{p}{2-p}} \|M \cdot e_j\|^{\frac{p}{2-p}} \, d\nu \Bigr)^{\frac{2-p}{2}} \Bigl(\int \sum_{i=1}^{\infty} |L_{i,k}||L_{i,j}| \, d\nu \Bigr)^{\frac{p}{2}} \\
&\le \sum_{k,j=1}^{\infty} \Bigl[\Bigl(\int \|M \cdot e_k\|^{\frac{2p}{2-p}} \, d\nu \Bigr)^{\frac{2-p}{4}} \Bigl(\int \|M \cdot e_j\|^{\frac{2p}{2-p}} \, d\nu \Bigr)^{\frac{2-p}{4}} \Bigl(\int \sum_{i=1}^{\infty} |L_{i,k}|^2 \, d\nu \Bigr)^{\frac{p}{4}} \Bigl(\int \sum_{i=1}^{\infty} |L_{i,j}|^2 \, d\nu \Bigr)^{\frac{p}{4}} \Bigr] \\
&= \Bigl(\sum_{k=1}^{\infty} \Bigl(\int \|M \cdot e_k\|^{\frac{2p}{2-p}} \, d\nu \Bigr)^{\frac{2-p}{4}} \Bigl(\int \sum_{i=1}^{\infty} |L_{i,k}|^2 \, d\nu \Bigr)^{\frac{p}{4}} \Bigr)^2.
\end{aligned}$$

Furthermore,

$$\sum_{i=1}^{\infty} \int |L_{i,k}|^2 \, d\nu = \sum_{i=1}^{\infty} \int |(DT)_{i,k}|^2 \, d\gamma = \sum_{i=1}^{k} \int |\partial_{x_i} T_k|^2 \, d\gamma = \int \|\nabla T_k\|^2 \, d\gamma$$

and

$$\int \|M\cdot e_k\|^{\frac{2p}{2-p}}\,d\nu = \int \Big(\sum_{i=1}^{\infty}\big(\sum_{r=0}^{N_0}\partial_{x_i}S_{i-r}\cdot\partial_{x_k}b_{i-r}\big)^2\Big)^{\frac{p}{2-p}}\,d\nu$$

$$\le C(p,N_0)\int\sum_{i=1}^{\infty}\Big(\sum_{r=0}^{N_0}|\partial_{x_i}S_{i-r}|^{\frac{2p}{2-p}}|\partial_{x_k}b_{i-r}|^{\frac{2p}{2-p}}\Big)\,d\nu$$

$$\le C(p,N_0)\sum_{i=1}^{\infty}\sum_{r=0}^{N_0}\Big(\int|\partial_{x_i}S_{i-r}|^{\frac{4p}{2-p}}\,d\nu\Big)^{1/2}\Big(\int|\partial_{x_k}b_{i-r}|^{\frac{4p}{2-p}}\,d\nu\Big)^{1/2}$$

$$\le C(p,N_0)\sup_i\Big(\sum_{j=i-N_0}^{i}\Big(\int|\partial_{x_i}S_j|^{\frac{4p}{2-p}}\,d\nu\Big)^{1/2}\Big)\sum_{j=1}^{\infty}\Big(\int|\partial_{x_k}b_j|^{\frac{4p}{2-p}}\,d\nu\Big)^{1/2}$$

$$\le C\sum_{j=1}^{\infty}\Big(\int|\partial_{x_k}b_j|^{\frac{4p}{2-p}}\,d\nu\Big)^{1/2}.$$

In the last estimate we apply Proposition 3.14 and the special structure of S .

Finally, we obtain

$$\int\|DS\cdot Db\cdot(DS)^{-1}\|^p_{\mathcal{HS}}\le C\Big[\sum_{j,k=1}^{\infty}\Big(\int|b_{j,k}|^{\frac{4p}{2-p}}\,d\nu\Big)^{\frac{2-p}{8}}\Big(\int\|\nabla T_k\|^2\,d\gamma\Big)^{\frac{p}{4}}\Big]^2.$$

$\square$

Lemma 7.9. *Assume that for every $p\ge 1$*

$$\sup_i\int\Big(|\beta_i|^p+\|\nabla\beta_i\|^p+\frac{1}{(\partial_{x_i}S_i)^p}\Big)\,d\mu\le C(p).$$

Then $\sup_i\int\|(DS_{x_i})(DS)^{-1}\|^p_{\mathcal{HS}}\,d\nu\le D$ with D depending on p and N_0.

Proof. Due to the special structure of S the matrix DS_{x_i} only has a finite number of non-zero entries. The result now follows immediately from Remark 3.14 and Propositions 3.15, 3.16. $\square$

8. Finite-dimensional case and optimal transportation

Instead of triangular mappings one can also apply the optimal transportation mappings. For a detailed account on optimal transportation see [13], [30]. In this case the available a-priori estimates are essentially better in many respects. For instance, there exist L^p-estimates on operator norms of DT which do not depend on dimension (see [25], [13]). Unfortunately, this approach has certain disadvantages: 1) unlike the triangular mappings, the optimal transportation mappings do not have an explicit form and the a-priori estimates for them are usually hard to prove, 2) the existence problem for optimal transportation mappings in infinite dimensions is solved in sufficient generality only for the case when the measures μ and ν have a finite Kantorovich distance $W_2(\mu,\nu)$ (see [21], [22])). If $\mu=\gamma$ is Gaussian, this limitation means that basically we should restrict ourselves to measures which are absolutely continuous with respect to γ, i.e. $\nu=g\cdot\gamma$ and, moreover, have finite entropy, that is $\int g\log g\,d\gamma<\infty$.

Remark 8.1. Some new existence results on optimal transportation of certain Gibbs measures are obtained in [26]. These results together with estimates from [25], [13] can be used to obtain infinite-dimensional uniqueness/existence statements for the case where ν is uniformly log-concave. But we don't consider this approach in this paper.

We assume in the rest of this section that $d < \infty$. We consider the optimal transportation mapping T pushing forward the standard Gaussian measure γ onto $\nu = e^{-W}dx$. In particular, T has the form $T = \nabla\Psi$, where Ψ is a convex function. The inverse mapping $S = T^{-1}$ is optimal too and has the form $S = \nabla\Phi$, where Φ is the convex conjugate to Ψ.

The drifts c and b are related in the same way as above

$$c = DS(T) \cdot b(T) = D^2\Phi(T) \cdot b(T).$$

Let us illustrate how our methods work in the finite-dimensional case.

Theorem 8.2. (Uniqueness) *Assume that W is locally Hölder, $\|Db\| \in L^2_{loc}(\nu)$, $|b| \in L^4(\nu)$, $|\nabla W| \in L^4(\nu)$. Then $\|c\| \in L^2(\gamma), \|D^2c\| \in L^2_{loc}(\gamma)$ and for every t_0 and every fixed initial condition $\rho_0 \in L^2(\nu)$ there exists at most one solution to (2) satisfying $\sup_{0\le t\le t_0} \|\rho(t,\cdot)\|_{L^2(\nu)} < \infty$.*

Remark 8.3. 1) Note that we do not need any bounds on the *second* derivatives of W.

2) The assumption of Hölder continuity is made only to assure high enough local integrability (even boundedness) of $\|D^2\Phi\| \cdot \|(D^2\Phi)^{-1}\|$. We believe that this can be achieved in some other (more efficient) way without Hölder continuity. We note, however, that Sobolev estimates for optimal transportation in $W^{1,p}_{loc}$ with big p are hard to prove. See in this respect the recent paper [27] and the references therein.

3) Some estimates applied in the proof are valid in the infinite-dimensional setting.

Proof. First we note that the second partial derivative $\partial_{ev}\Phi$ of Φ is locally Hölder for any vectors $e, v \in \mathbb{R}^d$, because both measures ν and γ have Lebesgue densities which are locally Hölder. This follows from the well-known results of Caffarelli [17] (see [24] for technical improvements for unbounded domains). Clearly, the same holds for $D^2\Phi$. Let us apply Proposition 4.4. We need to show that $\|c\| \in L^2(\gamma)$, $\|Dc\| \in L^2_{loc}(\gamma)$.

For $\|c\|$ one has the trivial estimate

$$\|c\| \le \|D\Phi(T)\| \cdot \|b(T)\|,$$

which implies

$$2\int \|c\|^2 \, d\nu \le \int \|D^2\Phi\|^4 d\nu + \int \|b\|^4 \, d\nu.$$

By Theorem 6.1 from [25] assumption $\|\nabla W\| \in L^4(\nu)$ implies that $\|D^2\Phi\| \in L^4(\nu)$. Hence $\|c\| \in L^2(\nu)$.

Let us estimate Dc:

$$Dc = \big[D^2\Phi \cdot Db \cdot (D^2\Phi)^{-1}\big](\nabla\Psi) + \big[(D^2\Phi)_b(D^2\Phi)^{-1}\big](\nabla\Psi),$$

where for brevity we set

$$(D^2\Phi)_b = \sum_{i=1}^{d} b_i \cdot \partial_{x_i}(D^2\Phi).$$

Using the standard HS-norm inequalities

$$\|AB\|_{HS} \le \|A\| \cdot \|B\|_{HS}, \quad \|BA\|_{HS} \le \|A\| \cdot \|B\|_{HS},$$

applied to an arbitrary A and symmetric B, we get

$$\|Dc\|_{HS} \le \Big(\mathcal{L}\|Db\|_{HS} + \mathcal{L}^{1/2} \cdot \|(D^2\Phi)^{-1/2}(D^2\Phi)_b(D^2\Phi)^{-1/2}\|_{HS}\Big)(\nabla\Psi),$$

where $\mathcal{L} = \|D^2\Phi\| \cdot \|(D^2\Phi)^{-1}\|$.

Since $\|D^2\Phi\|, \|(D^2\Phi)^{-1}\|$ are locally bounded functions, we only need to locally estimate

$$\|(D^2\Phi)^{-1/2}(D^2\Phi)_b(D^2\Phi)^{-1/2}\|_{HS};$$

To this end we apply the following inequality

(28)

$$\int \langle D^2W \cdot b, b\rangle\eta \, d\nu \ge \int \|D^2\Phi \cdot b\|^2\eta \, d\nu + \int \langle (D^2\Phi)_b \cdot b, (D^2\Phi)^{-1}\nabla\eta\rangle \, d\nu$$
$$+ 2\int \mathrm{Tr}\big((D^2\Phi)_b \cdot Db \cdot (D^2\Phi)^{-1}\big)\eta \, d\nu + \int \|(D^2\Phi)^{-1/2}(D^2\Phi)_b(D^2\Phi)^{-1/2}\|^2_{HS}\eta \, d\nu,$$

proved in Lemma 7.1 [25]. Here $\eta \in C_0^\infty(\mathbb{R}^d)$ with $\int \Big\|\frac{\nabla\eta}{\eta}\Big\|^p \eta \, d\nu < \infty$ for a sufficiently big p.

Since we do not assume existence of D^2W, we apply the integration by parts formula to get rid of this:

$$\int \langle D^2W \cdot b, b\rangle\eta \, d\nu = \int \langle\nabla W, b\rangle^2\eta \, d\nu - \int \langle\nabla W, b\rangle \mathrm{div}_\mu b \cdot \eta \ \ d\nu - \int \langle\nabla W, Db \cdot b\rangle\eta \, d\nu$$
$$+ \int \langle\nabla W, b\rangle\langle\nabla W, \nabla\eta\rangle \, d\nu.$$

Clearly, the right-hand side is finite by the Cauchy inequality.

The elementary estimates

$$\text{(29)} \quad \int \langle (D^2\Phi)_b \cdot b, (D^2\Phi)^{-1}\nabla\eta\rangle \, d\nu \le \varepsilon \int \|(D^2\Phi)^{-1/2}(D^2\Phi)_b(D^2\Phi)^{-1/2}\|^2_{\mathcal{HS}}\eta \, d\nu$$
$$+ C(d,\varepsilon)\int \|\mathcal{L}\||b|^2\frac{\|\nabla\eta\|^2}{\eta} \, d\nu,$$

and

(30)

$$2\int \mathrm{Tr}\big((D^2\Phi)_b \cdot Db \cdot (D^2\Phi)^{-1}\big)\eta \, d\nu \le \varepsilon \int \|(D^2\Phi)^{-1/2}(D^2\Phi)_b(D^2\Phi)^{-1/2}\|^2_{HS}\eta \, d\nu$$
$$+ C(\varepsilon)\int \|\mathcal{L}\|\|Db\|^2_{HS}\eta d\nu.$$

imply that $\|(D^2\Phi)^{-1/2}(D^2\Phi)_b(D^2\Phi)^{-1/2}\|^2_{HS} \in L^2_{loc}(\nu)$, hence $\|Dc\| \in L^2_{loc}(\gamma)$. $\square$

The following theorem is formulated in a "dimension-free" manner. This means in particular that this formulation makes sense in the infinite-dimensional setting. Actually, we believe that an appropriate generalization of the theorem always holds in the infinite-dimensional case provided the corresponding optimal transportation map of ν to γ does exist.

Theorem 8.4. **(Existence and uniqueness)** *Let ν be a probability measure with Lebesgue density on $\mathbb{R}^d$. Assume that ν is a uniformly log-concave measure, so, in particular, there exists $K > 0$ such that $D^2 W \geq K \cdot \mathrm{Id}$ in the interior of the support of ν. Assume that W belongs to $W^{2,p}(\nu)$ for some $p \geq 1$ and, moreover,*

$$\partial_{x_i} W \in L^{2p}(\nu), \quad 1 \leq i \leq d, \quad \|D^2 W\|^p \in L^1(\nu).$$

In addition, we assume that

$$\int \langle D^2 W b, b\rangle \ d\nu < \infty, \ b_i \in L^{\frac{4p}{2p-1}}(\nu), \quad 1 \leq i \leq d,$$

and

$$\mathrm{div}_\nu b \in L^2(\nu), \ \|Db\|_{HS}^{\frac{4p}{2p-1}} \in L^1(\nu).$$

Then

1) *under the additional assumption that $e^{\varepsilon(\mathrm{div}_\nu \mathrm{b})_-} \in L^1(\nu)$, there exists a solution to (2) for any initial $\rho_0 \in L^{2+\varepsilon}(\nu)$ satisfying $\sup_{0 \leq t \leq t_0} \|\rho(t,\cdot)\|_{L^2(\nu)} < \infty$;*
2) *any two solutions to (2) satisfying $\sup_{0 \leq t \leq t_0} \|\rho(t,\cdot)\|_{L^2(\nu)} < \infty$ with the same initial condition ρ_0 coincide.*

Proof. The proof follows the same line as the proof of the previous theorem. According to the Caffarelli's theorem $\|(D^2\Phi)^{-1}\| \leq \frac{1}{\sqrt{K}}$. According to a result of [25] $\int \|D^2\Phi\|^{2p} \ d\nu \leq \int \|D^2 W\|^p \ d\nu < \infty$. This implies that $\mathcal{L} \in L^{2p}(\nu)$.

We use the estimates from the proof of the previous theorem. By Hölder's inequality $\int \|\mathcal{L}\| \|Db\|_{HS}^2 \ d\nu < \infty$. Applying that $\int \langle D^2 W b, b\rangle \ d\nu < \infty$ we get from (28) and (30) that

$$\int \|(D^2\Phi)^{-1/2}(D^2\Phi)_b(D^2\Phi)^{-1/2}\|_{HS}^2 d\nu < \infty, \quad \int \|D^2\Phi \cdot b\|^2 \ d\nu < \infty,$$

and, finally $\|c\|^2, \|Dc\|_{HS}^2 \in L^1(\nu)$. The uniqueness statement follows easily from Proposition 4.4 with the help of Hölder's inequality.

For proving existence, we cannot use Theorem 5.1 directly, because assumptions 1) and 4) are not fulfilled in general. Note, however, that we need 1) and 4) only to apply Proposition 4.5. But the statement of Proposition 4.5 holds trivially in our case because of the uniform bound $\|(D^2\Phi)^{-1}\| \leq \frac{1}{\sqrt{K}}$ (see the proof of Lemma 4.3). $\square$

Example 8.5. Let

$$\nu = \rho(x) \ dx, \quad \rho(x) = C \prod_{i=1}^{d} e^{-(\frac{1}{x_i^2} + \frac{x_i^2}{2})} I_{\{x_i > 0\}}.$$

Note that $\nu = e^{-W}$ is uniformly log-concave and $\|\nabla W\|, \|D^2 W\|$ belongs to $L^p(\nu)$ for any $p > 0$. Fix an arbitrary $q > 0$ and set $b = \frac{x}{\|x\|^q}$. It is easy to check that $\mathrm{div}_\nu \mathrm{b}$ is bounded from below and the other assumptions on b are satisfied. Hence for every $\rho_0 \in L^{2+\varepsilon}(\nu)$ there exists a unique short-time (even long-time) solution to (2).

Note that if q is big, then b is not a BV function with respect to Lebesgue measure. This makes inapplicable the finite-dimensional theory from [5].

References

[1] Albeverio S., Kondratiev Yu.G., Röckner M., Tsikalenko T.V., A-priori estimates for symmetrizing measures and their applications to Gibbs states, JFA 171 (2000) 366–400.

[2] Albeverio S., Kondratiev Yu.G., Röckner M., Tsikalenko T.V., Uniqueness of Gibbs states for quantum lattice systems, Prob. Theory Related Fields 108 (1997) 193–218.

[3] Albeverio S., Röckner M., Classical Dirichlet forms on topological vector spaces. Closability and a Cameron–Martin formula, J. Funct. Anal. 88 (1990) 395–436.

[4] Ambrosio L., Transport equation and Cauchy problem for BV vector fields, Invent. Math. 158 (2004), 227–260.

[5] Ambrosio L., Transport equations and Cauchy vector problems for non-smooth vector fields, Lecture Notes of the CIME Summer school in Cetrary, June 27-July 2, 2005

[6] Ambrosio L., Figalli A., Almost everywhere well-posedness of continuity equations with measure initial data, Comp. Rend. Math. 348 (5-6) (2010) 249–252.

[7] Ambrosio L., Figalli A., On flows associated to Sobolev vector fields in Wiener spaces: An approach a la DiPerna-Lions, JFA, 256 (1) (2009) 179–214.

[8] Bogachev V.I., Measure theory. V. 1,2. Springer, Berlin – New York, 2007.

[9] Bogachev V.I., Differentiable measures and Malliavin calculus, AMS, Math. Surv. Monogr., vol. 164, 2010.

[10] Bogachev V.I., Kolesnikov A.V., On the Monge–Ampere equation in infinite dimensions, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 8 (4) (2005) 547–572.

[11] Bogachev V.I., Kolesnikov A.V., On nonlinear transformation of convex measures, Probab. Theory and Appl. 50(1) (2005) 34–52.

[12] Bogachev V.I., Kolesnikov A.V., Medvedev K.V., Triangular transformations of measures, Sbornik Math. 196(3) (2005) 3–30.

[13] Bogachev V.I., Kolesnikov A.V., Sobolev regularity for the Monge–Ampere equation in the Wiener space. arXiv: 1110.1822. (to appear in Kyoto Jour. Math.).

[14] Bogachev V.I., Kolesnikov A.V., The Monge–Kantorovich problem: achievements, connections, and perspectives, Russian Mathematical Surveys, 67 (5) (2012) 785–890.

[15] Bogachev V.I., E. Mayer-Wolf, Absolutely continuous flows generated by Sobolev class vector fields in finite in infinite dimensions, JFA 167 (1999) 1–68.

[16] Bogachev V.I., Da Prato G., Röckner M., Shaposhnikov S.V., Nonlinear evolution equations for measures on infinite dimensional spaces, In Stochastic Part. Diff. Equations and Appl., Quaderni di Matematica 25 (2010) 51–64.

[17] Caffarelli L.A., Interior $W^{2,p}$-estimates for solutions of the Monge–Ampère equation, Ann. of Math. (2) 131(1) (1990) 135–150.

[18] Caffarelli L.A., Monotonicity properties of optimal transportation and the FKG and related inequalities, Comm. Math. Phys. 214(3) (2000) 547–563.

[19] Cruzeiro A.B., Èquations différentielles sur l'espace de Wiener et formules de Cameron-Martin non linéaires, J. Func. Anal. 54 (1983) 206–227.

[20] Fang S., Luo D. Transport equations and quasi-invariant flows on the Wiener space, Bull. Sci. Math. 134 (2010) 295–328.

[21] Feyel D., Üstünel A.S., Monge–Kantorovitch measure transportation and Monge–Ampère equation on Wiener spaces. Prob. Theory and Related Fields, 128 (2004) 347–385.

[22] Kolesnikov A.V., Convexity inequalities and optimal transport of infinite-dimensional measures, J. Math. Pures Appl. (9) 83:11 (2004) 1373–1404.

[23] Kolesnikov A.V., Mass transportation and contractions, MIPT Proc. 2(4) (2010) 90-99.

[24] Kolesnikov A.V., On global Hölder estimates for optimal transportation, Mat. zametki, 88:5 (2010) 708–728.

[25] Kolesnikov A.V., On Sobolev regularity of mass transport and transportation inequalities, Theory of Probability and its Applications, 57:2 (2012) 296–321.

[26] Kolesnikov A.V., Zaev D., Optimal transportation of processes with infinite Kantorovich distance. Independence and symmetry. (in preparation).

[27] Kolesnikov A.V., Tikhonov S.Yu., Regularity of the Monge–Ampère equation in Besov's spaces. arXiv: 1203.3457. To appear in Calc. Variations PDE.

[28] Le Bris C., Lions P.L., Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients, Comm. in PDE's, 33(7) (2007) 1272–1317.

[29] Ovsienko Yu., Zhdanov R., Estimates for Sobolev norms of triangular mappings, Moscow University Math. Bull., 62(1) (2010) 1–4.
[30] Villani C., Topics in Optimal Transportation, Amer. Math. Soc. Providence, Rhode Island, 2003.

Higher School of Economics, Moscow, Russia
E-mail address: Sascha77@mail.ru

Universität Bielefeld, Bielefeld, Germany
E-mail address: roeckner@math.uni-bielefeld.de