Measurable Metrics and Gaussian Concentration

Francis HIRSCH
Equipe d'Analyse et Probabilités
Université d'Évry - Val d'Essonne
Boulevard F. Mitterrand
F-91025 EVRY CEDEX
`hirsch@maths.univ-evry.fr`

Abstract - *We introduce, in the general setting of a measure space endowed with a measurable metric in the sense of N. Weaver, the concentration function. We then extend some results concerning the Gaussian concentration property. In particular, we study relations between this property and transportation type inequalities or logarithmic Sobolev inequalities.*

1 Introduction

In the paper [18], N. Weaver introduced the notion of *measurable metric* and that of *Lipschitz function* with respect to such a metric. The study of these notions was pursued by the same author in subsequent papers ([19, 21, $\cdots$]) and in the book [20]. In our paper [11], we treated various important examples and, in particular, we studied the *intrinsic* measurable metric associated with a local Dirichlet form and notably the case of Wiener spaces.

This paper is a sequel of [11]. In the general setting of a measure space endowed with a measurable pseudometric, we shall first introduce (Section 2) the *distance function* to a set. This notion appears under a somewhat hidden form in [20] and plays an important rôle, in the case of Dirichlet forms, in the very nice paper [10]. From the distance function, we define in Section 3 the *concentration fonction* and we state an equivalent definition in terms of Lipschitz functions. Then, we study the Gaussian concentration property and its links, in the general framework, with Kantorovich metrics and inequalities involving the entropy (Section 4) or, in the framework of Dirichlet forms, with logarithmic Sobolev inequalities (Section 5). Finally, we give in Section 6 representation results for Wasserstein metrics in the case of Wiener spaces.

2 Distance function

We first recall some fundamental definitions and notation (see [11] and references therein).

In the following, we consider a σ-finite measure space (X, μ). The space L^∞ is related to this measure space. All functions that we consider in this paper are real functions. We denote by Ω the collection of all positive measure subsets of X. If A and B belong to Ω, we denote by $A \sim B$ the fact that A and B only differ from a null set ($\mu(A \setminus B) = \mu(B \setminus A) = 0$).

Definition 2.1 A *measurable pseudometric* is a map $\rho : \Omega^2 \longrightarrow [0, \infty]$ such that

1. $A' \sim A \Longrightarrow \rho(A', B) = \rho(A, B)$
2. $\rho(A, A) = 0$
3. $\rho(A, B) = \rho(B, A)$
4. $\rho(\cup_{n=1}^{\infty} A_n, B) = \inf_{n \geq 1} \rho(A_n, B)$
5. for any A, B, C, $\quad \rho(A, C) \leq \sup_{B' \subset B}(\rho(A, B') + \rho(B', C))$

where all subsets which appear above belong to Ω.

Let u be a μ-class of measurable real functions. If $A \in \Omega$, we denote by $F_A(u)$ the support of the image measure $(u_A)_*(\mu_A)$, where the index A indicates the restriction to A. This support $F_A(u)$ also is the essential image of the restriction u_A of u to A. We set, for $A, B \in \Omega$,

$$\rho_u(A, B) = \inf\{|x - y| : x \in F_A(u), y \in F_B(u)\}.$$

We also have

$$\rho_u(A, B) = \text{ess inf}\{|u(x) - u(y)| : x \in A, x \in B\},$$

this notation meaning

$$\max\{C \geq 0 : C \leq |\tilde{u}(x) - \tilde{u}(y)| \text{ for } \mu\text{-almost every } x \in A \text{ and for } \mu\text{-almost every } y \in B\},$$

where $\tilde{u}$ denotes any representative of the class u.

We henceforth consider a measurable pseudometric ρ, on the σ-finite measure space (X, μ).

Definition 2.2 Let u be a μ-class of measurable real functions. Then u is said to be a *ρ-Lipschitz function* if there exists a constant $C \geq 0$ such that

$$\forall A, B \in \Omega \quad \rho_u(A, B) \leq C\, \rho(A, B).$$

In this case,

$$L_\rho(u) = \sup\{\rho_u(A, B)/\rho(A, B) : A, B \in \Omega \text{ and } \rho(A, B) > 0\}$$

is finite and called the *Lipschitz constant* of u.

We shall denote by $Lip(\rho)$ the set of all ρ-Lipschitz functions, and by $\mathrm{Lip}^\infty(\rho)$ the set $Lip(\rho) \cap L^\infty$.

If S is a set of classes, we denote by $\bigvee S$ (resp. $\bigwedge S$) the essential supremum (resp. essential infimum) of S. Then $\bigvee S$ and $\bigwedge S$ are classes.

The following two Propositions come from [20]. Some proofs in [20] are partially erroneous but the main results hold true ([22]).

Proposition 2.3 ([20, Theorem 6.2.7]) *Let* $F = \mathrm{Lip}^\infty(\rho)$ *and, for* $f \in F$, *set* $\|f\|_0 = L_\rho(f)$. *Then* F *is a subspace of* L^∞ *containing the constant function* $\mathbf{1}$, *and* $\| \quad \|_0$ *is a seminorm on* F *such that* $\|\mathbf{1}\|_0 = 0$. *Moreover, for any* $M, N \geq 0$, *for any subset* S *of the set*

$$F_{M,N} = \{f \in F : \|f\|_\infty \leq M \text{ and } \|f\|_0 \leq N\},$$

the essential supremum $\bigvee S$ *belongs to* $F_{M,N}$.

In other words, the space $F = \mathrm{Lip}^\infty(\rho)$ equipped with the seminorm $\| \quad \|_0 = L_\rho$ satisfies the hypotheses of Proposition 2.6. in [11].

In what follows, if $B \subset X$, we denote by $\mathbf{1}_B$ the indicator function of B.

Proposition 2.4 ([20, Lemma 6.2.8]) *Let* $A \in \Omega$ *and* $a > 0$. *Set*

$$\rho_A^a = \bigvee \{\min(\rho(A,B), a)\, \mathbf{1}_B : B \in \Omega\}.$$

Then $\rho_A^a \in \mathrm{Lip}^\infty(\rho)$ *and* $L_\rho(\rho_A^a) \leq 1$. *Moreover,* $\rho_A^a(x) = 0$ *almost everywhere on* A.

Corollary 2.5 *For any* $A, B \in \Omega$,

$$\rho(A,B) = \sup\{\rho_f(A,B) : f \in \mathrm{Lip}^\infty(\rho) \text{ and } L_\rho(f) \leq 1\}.$$

Proof: We clearly have

$$\rho_{\rho_A^a}(A,B) \geq \min(\rho(A,B), a).$$

Therefore, by Proposition 2.4,

$$\sup\{\rho_f(A,B) : f \in \mathrm{Lip}^\infty(\rho) \text{ and } L_\rho(f) \leq 1\} \geq \min(\rho(A,B), a)$$

for any $a > 0$. This yields one inequality and the other one is obvious. □

By previous Proposition 2.3 and Corollary 2.5, we therefore see that any measurable pseudometric may be defined by the method of Proposition 2.6. in [11].

Definition 2.6 Set, for $A \in \Omega$,

$$\rho_A = \bigvee \{\rho(A,B)\, \mathbf{1}_B : B \in \Omega\}.$$

The class ρ_A is called the *distance function to* A.

It is easy to verify that, for any $a > 0$, $\rho_A^a = \rho_A \wedge a$.

The characterization below shows in particular that this distance function coincides with that introduced in [10] in the framework of local Dirichlet forms.

Theorem 2.7 *Let* $A \in \Omega$. *Then*

$$\rho_A = \bigvee \{f : f \in \mathrm{Lip}^\infty(\rho), L_\rho(f) \leq 1, f \geq 0, f(x) = 0 \text{ a.e. on } A\}.$$

Moreover, ρ_A *is the unique class* φ *satisfying the following properties:*

1. $\varphi \geq 0$ *and* $\varphi(x) = 0$ *a.e. on* A

2. $\forall B\in\Omega$*,* $\text{ess inf}\{\varphi(x):x\in B\}=\rho(A,B)$

3. $\forall a>0$*,* $(\varphi\wedge a)\in Lip(\rho)$ *and* $L_\rho(\varphi\wedge a)\le 1$*.*

Proof: Consider first $\varphi=\rho_A$. By Proposition 2.4, φ satisfies Properties 1 and 3 of the above statement. By the definition, we also have, for $B\in\Omega$, $\text{ess inf}\{\varphi(x):x\in B\}\ge\rho(A,B)$. Now, Property 3 entails

$$\min(\text{ess inf}\{\varphi(x):x\in B\},a)\le\rho(A,B).$$

Finally, φ satisfies Property 2 too.

Consider next

$$\varphi=\bigvee\{f:f\in\text{Lip}^\infty(\rho),L_\rho(f)\le 1,f\ge 0,f(x)=0\text{ a.e. on }A\}.$$

Obviously φ satisfies Property 1 and, as

$$\varphi\wedge a=\bigvee\{f\wedge a:f\in\text{Lip}^\infty(\rho),L_\rho(f)\le 1,f\ge 0,f(x)=0\text{ a.e. on }A\},$$

by Proposition 2.3 φ satisfies Property 3. Then we have as before, for any $B\in\Omega$, $\text{ess inf}\{\varphi(x):x\in B\}\le\rho(A,B)$. Let now α be a real number with $\alpha<\rho(A,B)$. By Corollary 2.5, there exists $f\in\text{Lip}^\infty(\rho)$ with $L_\rho(f)\le 1$ such that $\rho_f(A,B)\ge\alpha$. Now, by Proposition 2.3 and Lemma 2.7. in [11], there exists $g\in\text{Lip}^\infty(\rho)$ with $L_\rho(g)\le 1$ such that $g\ge 0$, $g(x)=0$ on A and $g(x)=\rho_f(A,B)$ on B. Then by definition $\varphi\ge g$. Hence,

$$\text{ess inf}\{\varphi(x):x\in B\}\ge\rho_f(A,B)\ge\alpha.$$

Finally, φ satisfies Property 2 too.

Consider now a function φ satisfying Properties 1, 2 and 3. By Property 2 we have clearly $\varphi\ge\rho_A$. By Property 3, for any $B\in\Omega$, $\text{ess inf}\{\varphi(x):x\in B\}\le\rho(A,B)=\text{ess inf}\{\rho_A(x):x\in B\}$. Therefore, $\varphi\le\rho_A$, and finally, $\varphi=\rho_A$. □

Proposition 2.8 *Let* $A\in\Omega$*. Then* $\rho_A(x)<+\infty$ *a.e. if and only if, for any* $B\in\Omega$*,* $\rho(A,B)<+\infty$*. If this is satisfied,* $\rho_A\in Lip(\rho)$ *and* $L_\rho(\rho_A)\le 1$*.*

Proof: For $B\in\Omega$, $\text{ess inf}\{\rho_A(x):x\in B\}=\rho(A,B)$. Therefore, if $\rho_A(x)<+\infty$ a.e., $\rho(A,B)<+\infty$ for any $B\in\Omega$. Likewise, if $\rho_A(x)=+\infty$ for $x\in B$ with $B\in\Omega$, $\rho(A,B)=+\infty$.

Suppose $\rho_A(x)<+\infty$ a.e. Set, for $n\in\mathbb{N}$, $H_n=\{\rho_A\le n\}$. If $B,C\in\Omega$ and $p\ge\max(m,n)$,

$$\rho_{\rho_A}(B,C)\le\rho_{\rho_A}(B\cap H_m,C\cap H_n)=\rho_{\rho_A\wedge p}(B\cap H_m,C\cap H_n)\le\rho(B\cap H_m,C\cap H_n).$$

Taking the infimum with respect to m and to n, we see that $\rho_A\in Lip(\rho)$ and $L_\rho(\rho_A)\le 1$. □

Particular case: Let ρ_0 be a map from X^2 into $[0,\infty]$ satisfying, for any $x,y,z\in X$,

$$\rho_0(x,x)=0,\quad \rho_0(x,y)=\rho_0(y,x)\quad\text{and}\quad\rho_0(x,z)\le\rho_0(x,y)+\rho_0(y,z).$$

For $A,B\in\Omega$, we set

$$\rho_0(A,B)=\inf\{\rho_0(x,y):x\in A,y\in B\},$$

and

$$\rho(A, B) = \sup\{\rho_0(A', B') : A' \sim A, B' \sim B\}.$$

We also have

$$\rho(A, B) = \text{ess } \inf\{\rho_0(x, y) : x \in A, y \in B\}.$$

For $A \subset X$, we denote by ρ_0^A the function

$$\rho_0^A(x) = \inf\{\rho_0(x, y) : y \in A\}.$$

Suppose that moreover ρ_0 satisfies the following assumption:
For all $A \in \Omega$, there exists $A' \in \Omega$ with $A' \subset A$, $A' \sim A$ and $\rho_0^{A'}$ measurable.
Then, by Proposition 2.3. of [11], ρ is a measurable pseudometric which we call the measurable pseudometric *associated with* ρ_0.

Proposition 2.9 *Under the above conditions, we have, for* $A \in \Omega$,

$$\rho_A = \bigvee\{\rho_0^{A'} : A' \subset A, A' \sim A, \rho_0^{A'} \textit{ measurable}\}.$$

Proof: Set

$$\varphi = \bigvee\{\rho_0^{A'} : A' \subset A, A' \sim A, \rho_0^{A'} \text{ measurable}\}.$$

Obviously, φ satisfies Property 1 of the statement of Theorem 2.7.

If $B \in \Omega$,

$$\text{ess } \inf\{\rho_0^{A'}(x) : x \in B\} = \sup\{\rho_0(A', B') : B' \sim B\}.$$

Hence,

$$\text{ess } \inf\{\varphi(x) : x \in B\} \geq \sup\{\rho_0(A', B') : A' \sim A, B' \sim B\} = \rho(A, B).$$

On the other hand, we have clearly that, for $a > 0$, $(\rho_0^{A'} \wedge a) \in \text{Lip}^\infty(\rho)$ and $L_\rho(\rho_0^{A'} \wedge a) \leq 1$. Therefore, by Proposition 2.3, φ satisfies Property 3 of the statement of Theorem 2.7. We conclude as previously that φ satisfies Property 2 too, and we may apply Theorem 2.7. □

3 Concentration function

We assume, in this section, that μ is a probability measure on X. As before, ρ denotes a measurable pseudometric on (X, μ).

Definition 3.1 For $A \in \Omega$ and $r > 0$, we denote by A_r the class (with respect to the relation $\sim$) $\{\rho_A < r\}$. We then define the *concentration fonction* α (related to the triplet (X, μ, ρ)) by

$$\forall r > 0 \quad \alpha(r) = 1 - \inf\{\mu(A_r) : A \in \Omega,\ \mu(A) \geq 1/2\}.$$

Consider first the particular case described in the previous section. For $A \subset X$ and $r > 0$, we set $A_r^0 = \{\rho_0^A < r\}$. We then define the "classical" concentration function α_0 (related to (X, μ, ρ_0)) by

$$\forall r > 0 \quad \alpha_0(r) = 1 - \inf\{\mu_*(A_r^0) : A \in \Omega,\ \mu(A) \geq 1/2\},$$

where μ_* denotes, for example, the interior measure. The next proposition shows that this classical concentration function coincides with the concentration function in the sense of Definition 3.1.

Proposition 3.2 *If α is the concentration function related to (X, μ, ρ), where ρ is the measurable pseudometric associated with ρ_0, then $\alpha_0 = \alpha$.*

Proof: Let $A \in \Omega$ with $\mu(A) \geq 1/2$. There exists $A' \in \Omega$, $A' \subset A$, $A' \sim A$, such that $\rho_0^{A'}$ is measurable. We have $\rho_0^{A'} \leq \rho_A$. Therefore, $A_r \subset (A')_r^0$ a.e. We also have $(A')_r^0 \subset A_r^0$. Hence, $\mu((A')_r^0) \leq \mu_*(A_r^0)$. Finally, for any $A \in \Omega$ with $\mu(A) \geq 1/2$, $\mu(A_r) \leq \mu_*(A_r^0)$, which implies $\alpha \geq \alpha_0$.

Conversely, suppose $A \in \Omega$ with $\mu(A) \geq 1/2$. If $(A_r)^c$ (the complement of A_r) belongs to Ω, then by Theorem 2.7

$$\text{ess } \inf\{\rho_A(x) : x \in (A_r)^c\} = \rho(A, (A_r)^c).$$

Therefore $\rho(A, (A_r)^c) \geq r$ and there exist $A', B \in \Omega$ with $A' \subset A$, $B \supset A_r$, $A' \sim A$, $B \sim A_r$, such that

$$\rho(A, (A_r)^c) = \rho_0(A', B^c).$$

Consequently, $B \supset (A')_r^0$ and $\mu_*((A')_r^0) \leq \mu(A_r)$. This implies $\alpha_0 \geq \alpha$. □

We now come back to the general situation. The following proposition is the analogue, in the general setting, of a classical result (see for example [1, Proposition 7.2.5]).

Proposition 3.3 *We have*

$$\alpha(r) = \sup\{\mu(F \geq m_F + r) : F \in \mathit{Lip}(\rho), L_\rho(F) \leq 1, m_F \textit{ median of } F\}$$

and we may replace above Lip(ρ) by $\mathrm{Lip}^\infty(\rho)$.

Proof: We set

$$\alpha_1(r) = \sup\{\mu(F \geq m_F + r) : F \in \mathit{Lip}(\rho), L_\rho(F) \leq 1, m_F \text{ median of } F\},$$

$$\alpha_2(r) = \sup\{\mu(F \geq m_F + r) : F \in \mathrm{Lip}^\infty(\rho), L_\rho(F) \leq 1, m_F \text{ median of } F\}.$$

Suppose $F \in \mathit{Lip}(\rho)$ with $L_\rho(F) \leq 1$ and let m_F be a median of F. We have, if $\{F \geq m_F + r\} \in \Omega$,

$$r \leq \rho_F(\{F \leq m_F\}, \{F \geq m_F + r\}) \leq \rho(\{F \leq m_F\}, \{F \geq m_F + r\}).$$

This entails

$$\{F \geq m_F + r\} \subset (\{F \leq m_F\}_r)^c.$$

Therefore,

$$\mu(\{F \geq m_F + r\}) \leq 1 - \mu(\{F \leq m_F\}_r).$$

As $\mu(\{F \leq m_F\}) \geq 1/2$ (m_F is a median of F), we have by the definition of α

$$\mu(\{F \geq m_F + r\}) \leq \alpha(r).$$

Consequently, $\alpha_1 \leq \alpha$.

Let now $A \in \Omega$ with $\mu(A) \geq 1/2$. By Proposition 2.4, for any $r > 0$, $\rho_A^r \in \mathrm{Lip}^\infty(\rho)$, $L_\rho(\rho_A^r) \leq 1$ and 0 is a median of ρ_A^r. Therefore

$$\mu(\rho_A^r \geq r) \leq \alpha_2(r).$$

Now,

$$\{\rho_A^r \geq r\} = \{\rho_A \geq r\} = (A_r)^c.$$

Then $1 - \mu(A_r) \leq \alpha_2(r)$, which yields $\alpha \leq \alpha_2$. □

We finish this section with the definition of the Gaussian concentration property.

Definition 3.4 Let c be a positive constant. We say that the triplet (X, μ, ρ) satisfies the *Gaussian concentration property* denoted by $G(c)$ if the associated concentration function α satisfies

$$\exists C \geq 0 \quad \forall r > 0 \qquad \alpha(r) \leq C \exp\left(-\frac{r^2}{c}\right).$$

We see by Proposition 3.2 that, in the particular case where the measurable pseudometric ρ is associated with a pointwise semimetric ρ_0 (see Section 2), the above Gaussian concentration property coincides with the classical one related to (X, μ, ρ_0) (see for example [1, Définition 7.2.3]).

4 Kantorovich metric and Gaussian concentration

We keep, in this section, the general framework and the notation of Section 3.

We denote by $\Pi(\mu)$ the set of all probability measures on X which are absolutely continuous with respect to μ. We first define, in this setting, the Kantorovich metric (see, for example, [16, p.88]).

Definition 4.1 We define the *Kantorovich metric* κ on $\Pi(\mu)$ by

$$\forall \xi, \eta \in \Pi(\mu) \quad \kappa(\xi, \eta) = \sup\{\int f \, \mathrm{d}\xi - \int f \, \mathrm{d}\eta : f \in \mathrm{Lip}^\infty(\rho), L_\rho(f) \leq 1\} \leq +\infty.$$

Generally, κ is not a true metric on $\Pi(\mu)$. We have

$$\forall \xi, \eta \in \Pi(\mu) \qquad \kappa(\xi, \eta) = 0 \quad \Longrightarrow \quad \xi = \eta$$

if and only if ρ is what is called a measurable metric (see [11, Definition 2.9.]).

We now define transportation type inequalities (this terminology being taken from [1]). We recall that, if $\xi = \varphi \, \mathrm{d}\mu$ belongs to $\Pi(\mu)$, the *entropy of* ξ is defined by

$$\mathrm{Ent}_\mu(\xi) = \mathrm{E}_\mu(\varphi \log \varphi)$$

where E_μ denotes the expectation with respect to μ.

Definition 4.2 Let c be a positive constant. We say that the triplet (X, μ, ρ) satisfies the *transportation inequality* denoted by $T(c)$ if

$$\forall \xi \in \Pi(\mu) \qquad \kappa(\xi, \mu) \leq \sqrt{c \operatorname{Ent}_\mu(\xi)}.$$

We can now state in this framework a result of S.G. Bobkov and F. Götze ([2]).

Proposition 4.3 *The transportation inequality $T(c)$ is equivalent to the following property: For all $\psi \in \operatorname{Lip}^\infty(\rho)$ such that $L_\rho(\psi) \leq 1$, one has*

$$\forall t \in \mathbb{R} \qquad \mathrm{E}_\mu(\exp(t\,\psi)) \leq \exp\left(\frac{c}{4}\,t^2 + t\,\mathrm{E}_\mu(\psi)\right).$$

Proof: According to our definition of the transportation inequality from the Kantorovich metric, the proof of [2, Theorem 3.1.]) (see also [1, Théorème 8.3.2]) works without modification. □

The following theorem shows the equivalence between a transportation inequality and the Gaussian concentration property.

Theorem 4.4 *Let c be a positive real number. If $T(c)$ holds, then $G(c')$ holds for any $c' > c$. Conversely, if $G(c)$ holds, then there exists $c' > 0$ such that $T(c')$ holds.*

Proof: Assume first that $T(c)$ holds. We can use Marton's argument ([14, 2, 13, 1, ...]). Suppose $A, B \in \Omega$ and consider the conditional probabilities $\mu_A = \mu(\cdot|A)$ and $\mu_B = \mu(\cdot|B)$. By $T(c)$ we have

$$\kappa(\mu_A, \mu_B) \leq \kappa(\mu_A, \mu) + \kappa(\mu, \mu_B) \leq \sqrt{-c\,\log(\mu(A))} + \sqrt{-c\,\log(\mu(B))}.$$

Let $r > 0$ and set $B = (A_r)^c$. If $B \in \Omega$, we have by the definition of κ and Proposition 2.4, for $a > 0$,

$$\kappa(\mu_A, \mu_B) \geq \int \rho_A^a \,\mathrm{d}\mu_B \geq \rho(A, B) \wedge a.$$

Now, as we have seen in the proof of Proposition 3.2, $\rho(A, (A_r)^c) \geq r$. Therefore

$$r \leq \sqrt{-c\,\log(\mu(A))} + \sqrt{-c\,\log[1 - \mu(A_r)]}.$$

This implies

$$r \geq \sqrt{-c\,\log(\mu(A))} \Longrightarrow [1 - \mu(A_r)] \leq \exp\left[-\frac{1}{c}\left(r - \sqrt{-c\,\log(\mu(A))}\right)^2\right].$$

Consequently

$$r \geq \sqrt{c\,\log 2} \Longrightarrow \alpha(r) \leq \exp\left[-\frac{1}{c}\left(r - \sqrt{c\,\log 2}\right)^2\right],$$

which implies $G(c')$ for any $c' > c$.

Assume conversely that $G(c)$ holds. By Proposition 3.3, if $F \in \operatorname{Lip}^\infty(\rho)$ with $L_\rho(F) \leq 1$ and if m_F is a median of F, we have for $r > 0$

$$\mu(F \geq m_F + r) \leq C\,\exp\left(-\frac{r^2}{c}\right).$$

Therefore, applying this inequality to F and to $-F$, we get

$$\mathrm{E}_\mu(|F-m_F|) \le 2\,C \int_0^\infty \exp\left(-\frac{r^2}{c}\right)\,\mathrm{d}r = C\,\sqrt{\pi\,c}.$$

We then have $|\mathrm{E}_\mu(F)-m_F| \le C'$ (with $C' = C\,\sqrt{\pi\,c}$) and, for $r \ge C'$,

$$\mu(F \ge \mathrm{E}_\mu(F)+r) \le C\,\exp\left(-\frac{(r-C')^2}{c}\right).$$

We deduce therefrom that, if $c_1 > c$, there exists $C_1 > 0$ such that, for any $F \in \mathrm{Lip}^\infty(\rho)$ with $L_\rho(F) \le 1$,

$$\forall r > 0 \qquad \mu(F \ge \mathrm{E}_\mu(F)+r) \le C_1\,\exp\left(-\frac{r^2}{c_1}\right).$$

Let now $\psi \in \mathrm{Lip}^\infty(\rho)$ with $L_\rho(\psi) \le 1$ and $\mathrm{E}_\mu(\psi) = 0$. We have

$$\mathrm{E}_\mu(\exp(t\,\psi)) = 1 + t\int_0^\infty \mathrm{e}^{tr}\,\mu(\psi \ge r)\,\mathrm{d}r - t\int_0^\infty \mathrm{e}^{-tr}\,\mu(\psi \le -r)\,\mathrm{d}r.$$

As

$$\mathrm{E}_\mu(\psi) = \int_0^\infty \mu(\psi \ge r)\,\mathrm{d}r - \int_0^\infty \mu(\psi \le -r)\,\mathrm{d}r = 0,$$

we also have

$$\mathrm{E}_\mu(\exp(t\,\psi)) = 1 + t\int_0^\infty (\mathrm{e}^{tr}-1)\,\mu(\psi \ge r)\,\mathrm{d}r + t\int_0^\infty (1-\mathrm{e}^{-tr})\,\mu(\psi \le -r)\,\mathrm{d}r.$$

Using what precedes, we get

$$\begin{aligned}\mathrm{E}_\mu(\exp(t\,\psi)) &\le 1 + t\,C_1\,\exp\left(c_1\,\frac{t^2}{4}\right)\int_{-\frac{tc_1}{2}}^{\frac{tc_1}{2}} \exp\left(-\frac{r^2}{c_1}\right)\,\mathrm{d}r,\\ &\le 1 + t^2\,C_1\;c_1\,\exp\left(c_1\,\frac{t^2}{4}\right) \le \exp(c_2\,t^2)\end{aligned}$$

for some constant c_2 independent of t. This implies $T(c')$ for $c' = 4c_2$ by Proposition 4.3.

5 Logarithmic Sobolev inequality and Gaussian concentration

In this section, we particularize the previous framework. We consider a Dirichlet form $(\mathbb{D},\mathcal{E})$ on the probability space (X,μ), in the sense of [3] to which we refer for the main definitions and properties. We assume that the Dirichlet form $\mathcal{E}$ is local (which means e.g.

$$\forall f,g \in \mathbb{D}\ \forall a \in \mathbb{R} \quad (f+a)\,g = 0 \Longrightarrow \mathcal{E}(f,g) = 0)$$

and there exists a carré du champ operator Γ (which means that there exists a continuous map $\Gamma : \mathbb{D}\times\mathbb{D} \longrightarrow L^1(\mu)$ such that

$$\forall f,g,h \in \mathbb{D}\cap L^\infty \quad \mathcal{E}(fh,g)+\mathcal{E}(gh,f)-\mathcal{E}(fg,h) = \int h\,\Gamma(f,g)\,\mathrm{d}\mu).$$

In what follows, we write $\Gamma(f)$ instead of $\Gamma(f,f)$ and $\mathcal{E}(f)$ instead of $\mathcal{E}(f,f)$. We recall that, under the above conditions,

$$\mathcal{E}(f) = \frac{1}{2}\int \Gamma(f)\,\mathrm{d}\mu.$$

We assume moreover that the constant function $\mathbf{1}$ belongs to $\mathbb{D}$.
We set

$$\mathbb{D}^\infty = \{f \in \mathbb{D} \cap L^\infty : \Gamma(f) \in L^\infty\}.$$

Following [11], we now define the *intrinsic metric* ρ by

$$\forall A, B \in \Omega \quad \rho(A,B) = \sup\{\rho_f(A,B) : f \in \mathbb{D}^\infty \text{ and } \Gamma(f) \le 1\}.$$

Then, by [11, Corollary 3.2.], ρ is a measurable pseudometric on (X,μ), $\mathrm{Lip}^\infty(\rho) = \mathbb{D}^\infty$ and, for any $f \in \mathbb{D}^\infty$,

$$L_\rho(f) = \|\Gamma(f)^{1/2}\|_\infty.$$

We recall that, if $f \in L^2(\mu)$, the *entropy of* f^2 is defined by

$$\mathrm{Ent}_\mu(f^2) = \mathrm{E}_\mu(f^2 \log f^2) - \mathrm{E}_\mu(f^2)\log \mathrm{E}_\mu(f^2).$$

Definition 5.1 Let c be a positive constant. We say that the Dirichlet form $(X,\mu,\mathbb{D},\mathcal{E})$ satisfies the *logarithmic Sobolev inequality* denoted by $LS(c)$ if

$$\forall f \in \mathbb{D} \qquad \mathrm{Ent}_\mu(f^2) \le c\,\mathcal{E}(f).$$

We can then prove that, as in classical cases, a logarithmic Sobolev inequality implies a transportation inequality.
In the rest of this section, c denotes a positive real number.

Theorem 5.2 *If the Dirichlet form $(X,\mu,\mathbb{D},\mathcal{E})$ satisfies the logarithmic Sobolev inequality $LS(c)$, then the triplet (X,μ,ρ) satisfies the transportation inequality $T(c/2)$.*

Proof: We use Herbst's argument (see [13, 1]). Let $\psi \in \mathbb{D}^\infty$ with $\Gamma(\psi) \le 1$ and set, for $t > 0$,

$$H(t) = \mathrm{E}_\mu(\exp(t\,\psi)) \quad \text{and} \quad f_t = \exp(\frac{t}{2}\,\psi).$$

By [3, Corollary 6.1.3] (functional calculus of Γ)

$$\Gamma(f_t) = \frac{t^2}{4}\exp(t\,\psi)\,\Gamma(\psi).$$

Therefore,

$$\Gamma(f_t) \le \frac{t^2}{4}\exp(t\,\psi).$$

By the logarithmic Sobolev inequality $LS(c)$ applied to f_t, we then obtain

$$t\,H'(t) - H(t)\,\log H(t) \le \frac{c}{8}\,t^2\,H(t)$$

or

$$\frac{d}{dt}\left(\frac{\log H(t)}{t}\right) \leq \frac{c}{8}.$$

On the other hand, we have clearly

$$\lim_{t\to 0} \frac{\log H(t)}{t} = \mathrm{E}_\mu(\psi).$$

Consequently, for all $t > 0$,

$$H(t) \leq \exp\left(\frac{c}{8}\, t^2 + t\, \mathrm{E}_\mu(\psi)\right).$$

It is now enough to apply Proposition 4.3. □

Corollary 5.3 *If the Dirichlet form $(X, \mu, \mathbb{D}, \mathcal{E})$ satisfies the logarithmic Sobolev inequality $LS(c)$, then the triplet (X, μ, ρ) satisfies the Gaussian concentration inequality $G(c')$ for any $c' > (c/2)$. More precisely, one has*

$$r \geq \sqrt{\frac{c}{2}\log 2} \Longrightarrow \alpha(r) \leq \exp\left[-\frac{2}{c}\left(r - \sqrt{\frac{c}{2}\log 2}\right)^2\right],$$

where α denotes the concentration function related to (X, μ, ρ).

Proof: See the proof of the first part of Theorem 4.4. □

Corollary 5.4 *If the Dirichlet form $(X, \mu, \mathbb{D}, \mathcal{E})$ satisfies the logarithmic Sobolev inequality $LS(c)$, then, for any $F \in \mathbb{D}$ such that $\Gamma(F) \leq 1$, one has*

$$\forall r > 0 \qquad \mu(F \geq \mathrm{E}_\mu(F) + r) \leq \exp\left(-\frac{2}{c}\, r^2\right).$$

Proof: We follow again [13, 1]. Let $F \in \mathbb{D}$ such that $\Gamma(F) \leq 1$ and set, for $n \in \mathbb{N}$, $F_n = (F \wedge n) \vee (-n)$. By the proof of Theorem 5.2, for all $t > 0$ and for all $n \in \mathbb{N}$,

$$\mathrm{E}_\mu(\exp(t\, F_n)) \leq \exp\left(\frac{c}{8}\, t^2 + t\, \mathrm{E}_\mu(F_n)\right).$$

We may pass to the limit for n tending to ∞ in the above inequality. Thus, for all $t > 0$,

$$\mathrm{E}_\mu(\exp(t\, F)) \leq \exp\left(\frac{c}{8}\, t^2 + t\, \mathrm{E}_\mu(F)\right).$$

By Chebyshev's inequality, for every $t, r > 0$,

$$\mu(F \geq \mathrm{E}_\mu(F) + r) \leq \exp\left(\frac{c}{8}\, t^2 - t\, r\right).$$

Optimizing in t, we obtain the result. □

6 Wasserstein metrics on Wiener spaces

We consider in this section the particular setting of an abstract Wiener space (X, H, μ) in the sense of L. Gross [8] (see also [17, 3, ...]. We recall that, in particular, X is a separable Banach space, μ is a centered Gaussian probability measure on X with support X, and $(H, |\ \ |_H)$ is a Hilbert space compactly embedded in X, which is called the Cameron-Martin space. We refer to [3] for a precise definition and for the notation we adopt here. Actually, we could take the more general framework considered in [5].
As shown in [3], there is a canonical Dirichlet form $(\mathbb{D}, \mathcal{E})$ on (X, μ), which is local and admits a carré du champ Γ. Moreover, $\mathbf{1} \in \mathbb{D}$. The results of the previous section are available and we keep the same notation. In particular the intrinsic metric ρ is defined as before.
We now set, for $x, y \in X$,

$$\begin{array}{rcll} \rho_0(x,y) & = & |x-y|_H & \text{if } x-y \in H \\ & = & +\infty & \text{otherwise.} \end{array}$$

We are then in the situation of the particular case described in Section 2, whose we keep the notation, and, by [11, Theorem 4.6.], ρ is actually the measurable pseudometric associated with ρ_0. In fact, ρ is a measurable metric and, for all $A, B \in \Omega$, $\rho(A, B) < \infty$ (see [11]). Moreover, we know ([9]) that the logarithmic Sobolev inequality $LS(4)$ holds.

A measurable real function f is called *H-Lipschitz continuous* if

$$\exists C \geq 0 \quad \forall x, y \in X \qquad |f(x) - f(y)| \leq \rho_0(x, y).$$

A class f is called *μ-a.e. H-Lipschitz continuous* ([4]) if f admits a representative $\tilde{f}$ which is H-Lipschitz continuous. In this case, the *H-Lipschitz constant* of f is defined by

$$L_H(f) = \text{ess sup}\{ \sup_{h \in H \setminus \{0\}} \frac{1}{|h|_H} |\tilde{f}(x+h) - \tilde{f}(x)| : x \in X\}.$$

We then recall the following characterization ([4, 11]):
Let $f \in L^\infty$. Then f is μ-a.e. H-Lipschitz continuous if and only if $f \in \text{Lip}^\infty(\rho)$ and, in this case, $L_H(f) = L_\rho(f)$.

If ξ and η are probability measures on X, we denote by $\Sigma(\xi, \eta)$ the set of all probability measures on $X \times X$ with marginal distributions ξ and η.

Definition 6.1 For $p \in [1, \infty[$ we define the *L^p-Wasserstein metric* by

$$W_p(\xi, \eta) = \left[\inf\{ \int [\rho_0(x,y)]^p \, \mathrm{d}\pi(x,y) : \pi \in \Sigma(\xi, \eta)\} \right]^{1/p}.$$

We shall first extend to this situation the classical Kantorovich representation theorem (see for example [16, Theorem 2.5.6]).

Theorem 6.2 *For every $\xi, \eta \in \Pi(\mu)$,*

$$W_1(\xi, \eta) = \kappa(\xi, \eta),$$

where κ denotes the Kantorovich metric defined in Definition 4.1

Proof: Let $\xi, \eta \in \Pi(\mu)$. As ρ_0 is lower semicontinuous on $X \times X$, the duality theorem [12, Theorem 2.6] gives,

$$\begin{aligned} W_1(\xi,\eta) = \sup\{\textstyle\int f\,\mathrm{d}\xi - \int g\,\mathrm{d}\eta \quad & : f \in \mathcal{L}^1(\xi), g \in \mathcal{L}^1(\eta) \text{ and} \\ & \forall x,y \in X \ f(x) < +\infty, g(y) > -\infty, f(x) - g(y) \le \rho_0(x,y)\}. \end{aligned}$$

Replacing in the sup, f and g respectively by $(f \wedge M) \vee (-M)$ and $(g \wedge M) \vee (-M)$ and letting M tend to infinity, we may assume, in the sup, that f and g are bounded. By the regularity of η, we may assume, moreover, that g is lower semicontinuous. Suppose then that f and g satisfy:

$$\forall x,y \in X \quad |f(x)| \le M,\ |g(y)| \le M,\ f(x) - g(y) \le \rho_0(x,y)$$

and that g is lower semicontinuous. Let U be a $\mathcal{K}_\sigma$ of X with $\mu(U) = 1$ and set, for $x \in X$,

$$f_1(x) = \inf\{g(y) + \rho_0(x,y) : y \in U\} \wedge M.$$

As g is l.s.c., f_1 is a limit of a decreasing sequence of l.s.c. functions. Hence, f_1 is a Borel function. Moreover we have clearly $f_1 \ge f$, $|f_1| \le M$ and $f_1(y) \le g(y)$ for $y \in U$ and therefore μ-a.e. We denote by φ the class of f_1. Clearly, φ is μ-a.e. H-Lipschitz continuous and $L_H(\varphi) \le 1$. Therefore $\varphi \in \mathrm{Lip}^\infty(\rho)$ and $L_\rho(\varphi) \le 1$. Now

$$\int f\,\mathrm{d}\xi - \int g\,\mathrm{d}\eta \le \int f_1\,\mathrm{d}\xi - \int f_1\,\mathrm{d}\eta = \int \varphi\,\mathrm{d}\xi - \int \varphi\,\mathrm{d}\eta.$$

This implies $W_1(\xi,\eta) \le \kappa(\xi,\eta)$. The converse inequality is easy, using again the characterization of μ-a.e. H-Lipschitz continuous functions. □

For $p > 1$, we have the following extension of [16, Corollary 2.5.2] (see also [15, Theorem 5.2.1]). We also mention the deep study [6, 7] of the L^2- Wasserstein metric on abstract Wiener spaces related to transportation problems. In what follows, if $h \in \mathrm{Lip}^\infty(\rho)$, $\tilde{h}$ denotes an H-Lipschitz continuous representative of h.

Theorem 6.3 *Let $p > 1$. For every $\xi, \eta \in \Pi(\mu)$,*

$$\begin{aligned} [W_p(\xi,\eta)]^p = \sup\{\textstyle\int f\,\mathrm{d}\xi - \int g\,\mathrm{d}\eta \quad & : f \in \mathrm{Lip}^\infty(\rho),\ g \in \mathrm{Lip}^\infty(\rho) \text{ and} \\ & \forall x,y \in X \quad \tilde{f}(x) - \tilde{g}(y) \le [\rho_0(x,y)]^p\}. \end{aligned}$$

Proof: Let $\xi, \eta \in \Pi(\mu)$. As before, the duality theorem [12, Theorem 2.6] gives,

$$\begin{aligned} [W_p(\xi,\eta)]^p = \sup\{\textstyle\int f\,\mathrm{d}\xi - \int g\,\mathrm{d}\eta \quad & : f \in \mathcal{L}^1(\xi), g \in \mathcal{L}^1(\eta) \text{ and} \\ & \forall x,y \in X \ f(x) < +\infty, g(y) > -\infty, f(x) - g(y) \le [\rho_0(x,y)]^p\}. \end{aligned}$$

As in the previous proof, we may assume that, in the sup, f and g satisfy:

$$\forall x,y \in X \quad |f(x)| \le M,\ |g(y)| \le M,\ f(x) - g(y) \le [\rho_0(x,y)]^p$$

and that f is upper semicontinuous. Let U be a $\mathcal{K}_\sigma$ of X with $U = U + H$ and $\mu(U) = 1$. We set, for $y \in X$ and $N \in \mathbb{N}$,

$$g_N(y) = \sup\{f(x) - (\rho_0(x,y) \wedge N)^p : x \in U\}.$$

Then g_N is a Borel function, $-M - N^p \leq g_N \leq M$ and g_N is H-Lipschitz continuous. We set

$$f_N = \mathbf{1}_U \, f - (M + N^p) \, \mathbf{1}_{U^c}.$$

We have

$$\forall x, y \in X \qquad f_N(x) - g_N(y) \leq [\rho_0(x,y)]^p.$$

We see easily that (g_N) is a decreasing sequence which converges to k defined by

$$k(y) = \sup\{f(x) - [\rho_0(x,y)]^p : x \in U\}.$$

We have $k \leq g$. Hence,

$$\lim_{N \to \infty} \int g_N \, \mathrm{d}\eta = \int k \, \mathrm{d}\eta \leq \int g \, \mathrm{d}\eta$$

and, for all N,

$$\int f_N \, \mathrm{d}\xi = \int f \, \mathrm{d}\xi.$$

Therefore, replacing in the sup the couple (f, g) by (f_N, g_N), we may suppose that f and g are bounded and that g is H-Lipschitz continuous. Let $C \geq 0$ such that

$$\forall x, y \in X \qquad |g(x) - g(y)| \leq C \, \rho_0(x,y)$$

and let h be an l.s.c. function such that $g \leq h$. We set

$$\hat{h}(x) = \inf\{h(y) + C \, \rho_0(x,y) : y \in U\}.$$

Then $g \leq \hat{h}$ and $\hat{h}(x) \leq h(x)$ for $x \in U$. This implies

$$\int g \, \mathrm{d}\eta \leq \int \hat{h} \, \mathrm{d}\eta \leq \int h \, \mathrm{d}\eta.$$

Therefore, by the regularity of η, we may replace g by $\hat{h}$. Now, if $U = \cup_{n \geq 0} K_n$ with each K_n being a compact set, then $\hat{h}$ is the decreasing limit of the sequence (h_n) with

$$h_n(x) = \inf\{h(y) + C \, \rho_0(x,y) : y \in K_n\}.$$

As h_n is an l.s.c. H-Lipschitz continuous function, this shows that we may assume that the H-Lipschitz continuous function g is moreover a lower semicontinuous function. Then, symmetrizing the construction of the first part of the proof, we may finally assume that f and g are bounded H-Lipschitz continuous functions. $\square$

References

[1] C. ANÉ, S. BLACHÈRE, D. CHAFAÏ, P. FOUGÈRES, I. GENTIL, F. MALRIEU, C. ROBERTO, G. SCHEFFER *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses 10, Société Mathématique de France, 2000.

[2] S.G. BOBKOV, F. GÖTZE Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, J. Funct. Anal., 163 (1999), 1-28.

[3] N. BOULEAU, F. HIRSCH *Dirichlet forms and analysis on Wiener space*, de Gruyter, Berlin, 1991.

[4] O. ENCHEV, D.W. STROOCK Rademacher's theorem for Wiener functionals, Ann. Probab. 21 (1993), 25-33.

[5] D. FEYEL, A. de LA PRADELLE Espaces de Sobolev gaussiens, Ann. Inst. Fourier, 39-4 (1989), 875-908.

[6] D. FEYEL, A. S. ÜSTÜNEL Measure Transport on Wiener Space and the Girsanov Theorem, C.R. Acad. Sci. Paris, t. 334, Série I, p. 1025-1028, 2002.

[7] D. FEYEL, A. S. ÜSTÜNEL Monge-Kantorovich Measure Transportation and Monge-Ampère Equation on Wiener Space, Preprint 2003.

[8] L. GROSS Abstract Wiener Spaces, in: *Proc. Fifth Berkeley Sympos. Math. Statist. Proba.*, p. 31-42, Univ. of California Press, 1965.

[9] L. GROSS Logarithmic Sobolev inequalities, Amer. J. Math., 97 (1976), 1061-1083.

[10] M. HINO, J.A. RAMIREZ Small time Gaussian behavior of symmetric diffusion semigroups, Ann. Probab., 31 (2003), 1254-1295.

[11] F. HIRSCH Intrinsic metrics and Lipschitz functions, J. evol. equ., 3 (2003), 11-25.

[12] H.G. KELLERER Duality Theorems for Marginal Problems, Z. Wahrscheinlichkeitstheorie verw. Gebiete, 67 (1984), 399-432.

[13] M. LEDOUX Concentration of measure and logarithmic Sobolev inequalities, in: *Séminaire de Probabilités XXXIII*, p. 120-216, Lecture Notes in Math. 1709, Springer, Berlin,1999.

[14] K. MARTON Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration, Ann. Probab., 24 (1996), 857-866.

[15] S.T. RACHEV *Probability Metrics and the Stability of Stochastic Models*, Wiley, Chichester-New York, 1991.

[16] S.T. RACHEV, L. RÜSCHENDORF *Mass transportation problems, Vol. 1*, Springer, Berlin, 1998.

[17] S. WATANABE *On stochastic differential equations and Malliavin calculus*, Tata Intitute of Fund. Research, Vol.73, Springer, Berlin, 1984.

[18] N. WEAVER Nonatomic Lipschitz Spaces, Studia Math., 115 (1995), 277-289.

[19] N. WEAVER Lipschitz Algebras and Derivations of von Neumann Algebras, J. Funct. Anal., 139 (1996), 261-300.

[20] N. WEAVER *Lipschitz Algebras*, World Scientific, Singapore, 1999.

[21] N. WEAVER Lipschitz Algebras and Derivations II. Exterior Differentiation, J. Funct. Anal., 178 (2000), 64-112.

[22] N. WEAVER Personal communication.