Bernoulli **26**(2), 2020, 1431–1452
https://doi.org/10.3150/19-BEJ1163

Around the entropic Talagrand inequality

GIOVANNI CONFORTI[1] and LUIGIA RIPANI[2]

[1]*Département de Mathématiques Appliquées, École Polytechnique, Route de Saclay, 91128, Palaiseau Cedex, France. E-mail: giovanni.conforti@polytechnique.edu*
[2]*Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France. E-mail: ripani@math.univ-lyon1.fr*

In this article, we study generalization of the classical Talagrand transport-entropy inequality in which the Wasserstein distance is replaced by the entropic transportation cost. This class of inequalities has been introduced in the recent work (*Probab. Theory Related Fields* **174** (2019) 1–47), in connection with the study of Schrödinger bridges. We provide several equivalent characterizations in terms of reverse hypercontractivity for the heat semigroup, contractivity of the Hamilton–Jacobi–Bellman semigroup and dimension-free concentration of measure. Properties such as tensorization and relations to other functional inequalities are also investigated. In particular, we show that the inequalities studied in this article are implied by a Logarithmic Sobolev inequality and imply Talagrand inequality.

Keywords: functional inequalities; large deviations; optimal transport; semigroups

1. Introduction and statements of the main results

A first probabilistic approach to transportation problems goes back to the early works of Schrödinger [29,30], who was interested in finding the most likely evolution of a cloud of independent brownian particles towards a given "unexpected" configuration. A rigorous formulation of Schrödinger's question is achieved through a constrained entropy minimization, known as the *Schrödinger problem* (SP). The optimal value in (SP) measures intuitively the asymptotic probability that the particles attain the desired configuration, and is called *entropic transportation cost*. Mikami discovered in [24] (see also [22]) a fundamental connection with deterministic optimal transport, by showing that the Monge–Kantorovich problem (MK) may be seen as a "small noise limit" of the Schrödinger problem. The study of the relations between these two transportation problems is nowadays an active field of research for at least two reasons: on the one hand the fact that (SP) provides with a regular convex approximation of (MK) has led to computational advantages [5,11]; on the other hand the goal is understanding what is the "stochastic" counterpart of the large body of results concerning the interplay between optimal transport, functional inequalities and curvature-like conditions [10,13,14]. The present article contributes to this second line of research by studying a family of functional inequalities introduced in [10] which naturally generalizes Talagrand's transportation inequality [31] to the entropic cost: for this reason we call them *entropic Talagrand inequalities*.

We recall that a probability measure m on $\mathbb{R}^d$ satisfies Talgrand's transportation inequality with constant C, if for any probability measure μ we have

$$W_2^2(\mu, m) \leq C\mathcal{H}(\mu|m), \tag{1.1}$$

1350-7265 © 2020 ISI/BS

where $W_2^2(\cdot,\cdot)$ is the squared Wasserstein distance of order two and $\mathcal{H}(\cdot|m)$ is the relative entropy w.r.t. m. This inequality was first introduced in [31] for the Gaussian measure in the Euclidean space by Talagrand, and then generalized in [27] by Otto and Villani. Later on we will adopt the notation TI(λ) for the classical Talagrand inequality (1.1) with constant $C = 1/\lambda$.

To introduce the entropic version of (1.1), we fix a probability measure $m(\mathrm{d}x) = \exp(-2U(x))\,\mathrm{d}x$ and a noise parameter $\varepsilon > 0$ and consider the Langevin dynamics for U

$$\mathrm{d}X_t = -\varepsilon\nabla U(X_t)\,\mathrm{d}t + \sqrt{\varepsilon}\,\mathrm{d}B_t, \quad X_0 \sim m. \tag{1.2}$$

Next, we call R^ε_{0t} the joint law at times $0, t$ of the Langevin dynamics: R^ε_{0t} acts as *reference measure* to define the entropic transportation cost $\mathcal{T}_{R^\varepsilon_{0t}}(\mu,\nu)$ via the associated Schrödinger problem. The latter consists in minimizing the relative entropy w.r.t. the reference measure R^ε_{0t} over the set of couplings of μ and ν. Leaving precise statements for later, let us just say that a probability measure m on $\mathbb{R}^d$ satisfies an entropic Talagrand inequality if

$$\forall\mu, \quad \mathcal{T}_{R^\varepsilon_{0t}}(\mu, m) \le C\mathcal{H}(\mu|m)$$

or, more generally,

$$\forall\mu,\nu, \quad \mathcal{T}_{R^\varepsilon_{0t}}(\mu,\nu) \le C\mathcal{H}(\mu|m) + C'\mathcal{H}(\nu|m).$$

These inequalities are stronger than the classical Talagrand inequality since the entropic transport cost dominates the Wasserstein, see Remark 1.2 below. Moreover, the classical Talagrand inequality is recovered in the limit when $\varepsilon \to 0$. The main results of this article include equivalent characterizations of the entropic Talagrand inequalities in terms of a weak form of reverse hypercontractivity for the semigroup associated with (1.2), contractivity properties for the Hamilton–Jacobi–Bellmann semigroup and a dimension-free concentration property, in the spirit of [17]; all these characterizations allow to recover well known results about Talagrand's inequality in the small noise limit. Furthermore, we show that the entropic Talagrand inequalities tensorize, and investigate relations with classical inequalities. In particular, we extend Otto-Villani's Theorem [27], by showing that the entropic transportation inequality is implied by a Logarithmic Sobolev inequality, and that it implies the classical Talagrand's inequality. As a byproduct, we obtain that the entropic Talagrand inequalities hold under the celebrated Bakry–Émery Γ_2 condition [2]. This fact has already been proven for measures on a compact Riemannian manifold in [10].

Transport-entropy inequalities for general costs have been studied in [20] (and also [8]). An observation we make here is that (a slight modification of) the entropic cost is indeed one of those general costs. This allows us to profit from the results contained in [20], thus simplifying some of our proofs. Conversely, we provide a novel concrete example of functional inequality which can be treated with the methods of [20]; moreover we can provide explicit conditions for this inequality to hold, something which cannot be achieved for the general costs considered there. Finally, let us remark that, to streamline exposition, we limit ourselves to take $\mathbb{R}^d$ as ambient space; however, it is very likely that the results we present here remain valid in a much wider setting.

Organization of the article. We recall at Section 1 some basic facts about (SP) and its connections to optimal transport. In Section 2, we first introduce the class of entropic Talagrand

inequalities at Definition 2.1, and prove two characterization results, Theorem 2.2 and Theorem 2.3. Next, we investigate different forms of tensorization at Propositions 2.1, 2.2 and 2.3. Then, we use these results to derive concentration of measure at Theorem 2.4. We establish at Corollary 2.1 connections with the classical Talagrand inequality and the Logarithmic Sobolev inequality. Finally, at Corollary 2.2 we show that an entropic Talagrand inequality implies an infimum convolution Logarithmic Sobolev inequality. The Appendix collects some useful results which are behind most of the proof presented here.

1.1. Schrödinger problem and entropic transportation cost

In order to define (SP), we shall first introduce a few notation. We fix a probability measure m on $\mathbb{R}^d$ whose density w.r.t. the Lebesgue measure is $\exp(-2U(x))$, where U is assumed to satisfy the minimal hypothesis which guarantee existence of a weak solution for the SDE (1.2). This is the case for instance, when there exists some constant $c > 0$ such that one of the following assumptions holds true:

(i) $\lim_{|x|\to\infty} U(x) = +\infty$ and $\inf\{|\nabla U|^2 - \Delta U/2\} > -\infty$, or
(ii) $-x \cdot \nabla U(x) \le c(1 + |x|^2)$, for all $x \in \mathbb{R}^d$.

See [28], Theorem 2.2.19, for the existence result under the assumptions (i) or (ii). For any $\varepsilon > 0$, we call R^ε the law of (1.2) on the space of continuous paths over $[0, +\infty]$ and for $t > 0$ we denote R^ε_{0t} the law of R^ε at times 0 and t:

$$R^\varepsilon_{0t}(\cdot) = R^\varepsilon\big((X_0, X_t) \in \cdot\big).$$

For any measurable space E, we denote by $\mathcal{P}(E)$ the space of probability measures over E and for any $p, q \in \mathcal{P}(E)$ $\Pi(p, q)$ is the set of couplings of p and q; finally $\mathcal{H}(q|p)$ is the relative entropy of q w.r.t. p defined as,

$$\mathcal{H}(q|p) = \begin{cases} \displaystyle\int \log \frac{\mathrm{d}q}{\mathrm{d}p}\,\mathrm{d}q & \text{if } q \ll p, \\ +\infty & \text{otherwise.} \end{cases}$$

We are now in position to define (SP). Given two marginal laws $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ and $\varepsilon, t > 0$, the (static) Schrödinger problem is the problem of finding the coupling of μ and ν which minimizes the relative entropy against R^ε_{0t},

$$\inf\big\{\mathcal{H}\big(\pi|R^\varepsilon_{0t}\big) : \pi \in \Pi(\mu, \nu)\big\}, \tag{SP}$$

we call the optimal value in (SP) the *entropic transportation cost* between μ and ν, and denote it $\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu)$.

As it is the case for the Wasserstein distance, the entropic transportation cost admits a dual formulation. It is known that if $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ have finite relative entropy w.r.t. m, then

$$\varepsilon\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu) = \varepsilon\mathcal{H}(\mu|m) + \sup_{\varphi\in C_b(\mathbb{R}^d)} \left\{\int Q^\varepsilon_t \varphi\,\mathrm{d}\mu - \int \varphi\,\mathrm{d}\nu\right\}, \tag{1.3}$$

where for all $t \geq 0$, $x \in \mathbb{R}^d$

$$Q_t^\varepsilon \varphi(x) = \inf_{p \in \mathcal{P}(\mathbb{R}^d)} \left\{ \int \varphi(y) p(\mathrm{d}y) + \varepsilon \mathcal{H}\big(p | r_t^\varepsilon(x, \cdot)\big) \right\}, \tag{1.4}$$

where $x \mapsto r_t^\varepsilon(x, \cdot) \in \mathcal{P}(\mathbb{R}^d)$ is the m-a.s. defined Markov kernel such that

$$R_{0t}^\varepsilon(\mathrm{d}x\, \mathrm{d}y) = m(\mathrm{d}x) r_t^\varepsilon(x, \mathrm{d}y). \tag{1.5}$$

The semigroup $(Q_t^\varepsilon)_{t \geq 0}$ is the Hamilton–Jacobi–Bellman (HJB) semigroup characterizing the vanishing viscosity solutions for the Hamilton–Jacobi equation. Different proofs of (1.3) in more general contexts are by now available, see for instance [9,13,14,16,25]. Introducing the linear semigroup $(P_t^\varepsilon)_{t \geq 0}$ associated with (1.2) allows to give an alternative formulation of the HJB semigroup. We have

$$Q_t^\varepsilon \varphi(x) = -\varepsilon \log P_t^\varepsilon \exp(-\varphi/\varepsilon)(x), \quad x \in \mathbb{R}^d. \tag{1.6}$$

Note that (1.6) follows from the dual representation of the entropy (A.2).

1.2. The connection with optimal transport

A fundamental fact is that one recovers (MK) from (SP) as a small noise (or, equivalently, short time) limit. This was first proven in [24] when the reference measure is a Brownian motion and in [22] in a more general case using Γ-convergence. In particular, those results imply that for all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with second moment and relative entropy w.r.t. m finite,

$$\lim_{\varepsilon \to 0^+} \varepsilon \mathcal{T}_{R_{0t}^\varepsilon}(\mu, \nu) = \frac{W_2^2(\mu, \nu)}{2t}, \tag{1.7}$$

where $W_2(\mu, \nu)$ is Wasserstein distance of order two is defined for all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ with second moment as

$$W_2^2(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int |x - y|^2 \pi(\mathrm{d}x\, \mathrm{d}y).$$

Furthermore application of the Laplace principle [12], Theorem 4.3.1, yields

$$\forall x \in \mathbb{R}^d, \quad \lim_{\varepsilon \to 0} Q_t^\varepsilon \varphi(x) = \inf_{y \in \mathbb{R}^d} \left\{ \varphi(y) + \frac{1}{2t} |x - y|^2 \right\} := Q_t^0 \varphi(x). \tag{1.8}$$

Here $Q_t^0 \varphi$ is nothing but the Hopf–Lax semigroup that appears in the classical Kantorovich duality formula of optimal transport,

$$\frac{1}{2} W_2^2(\mu, \nu) = \sup_{\varphi \in C_b(\mathbb{R}^d)} \left\{ \int Q_t^0 \varphi \, \mathrm{d}\mu - \int \varphi \, \mathrm{d}\nu \right\}. \tag{1.9}$$

In [20], the authors study a general family of transportation costs. In particular, they look at costs which can be defined considering a measurable function $c : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to [0, +\infty]$ and setting

$$\mathcal{T}_c(\nu|\mu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^d} c\big(x, p(x, \cdot)\big)\mu(\mathrm{d}x), \tag{1.10}$$

where for $\pi \in \Pi(\mu, \nu)$ the map $x \mapsto p(x, \cdot)$ is the (μ-almost everywhere uniquely determined) probability kernel such that

$$\pi(\mathrm{d}x\,\mathrm{d}y) = \mu(\mathrm{d}x)p(x, \mathrm{d}y).$$

We observe that if we subtract the marginal entropy of μ to the entropic transportation cost, then we fall in the set of costs (1.10). This simple fact allows us to take advantage of the results in [20]. Inspired from their framework, we define

$$\mathcal{T}_{R^\varepsilon_{0t}}(\nu|\mu) = \inf\left\{\int_{\mathbb{R}^d} \mathcal{H}\big(p(x, \cdot)|r^\varepsilon_t(x, \cdot)\big)\mu(\mathrm{d}x) : \pi \in \Pi(\mu, \nu)\right\}, \tag{1.11}$$

which is nothing but the cost (1.10) with the choice

$$c(x, p) = \mathcal{H}\big(p|r^\varepsilon_t(x, \cdot)\big). \tag{1.12}$$

Lemma 1.1. *For all μ, ν such that $\mathcal{H}(\mu|m) < +\infty$ we have that*[1]

$$\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu) - \mathcal{H}(\mu|m) = \mathcal{T}_{R^\varepsilon_{0t}}(\nu|\mu). \tag{1.13}$$

Proof. Assume that $\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu) < +\infty$. In this case, the conclusion follows from the decomposition of the entropy formula (see [23], Theorem 2.4, or Lemma A.2 from the Appendix), valid for all $\pi \in \Pi(\mu, \nu)$

$$\mathcal{H}\big(\pi|R^\varepsilon_{0t}\big) = \mathcal{H}(\mu|m) + \int_{\mathbb{R}^d} \mathcal{H}\big(p(x, \cdot)|r^\varepsilon_t(x, \cdot)\big)\mu(\mathrm{d}x) \tag{1.14}$$

and by taking the infimum on both sides. On the other hand, if $\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu) = +\infty$, we find from (1.14) that

$$\forall \pi \in \Pi(\mu, \nu), \quad \mathcal{H}(\mu|m) + \int_{\mathbb{R}^d} \mathcal{H}\big(p(x, \cdot)|r^\varepsilon_t(x, \cdot)\big)\mu(\mathrm{d}x) = +\infty.$$

Using the fact that $\mathcal{H}(\mu|m) < +\infty$, we get that $\mathcal{T}_{R^\varepsilon_{0t}}(\nu|\mu) = +\infty$ as well, which is the desired conclusion. □

Remark 1.1. Note that the entropic transportation cost is symmetric, and together with (1.13) it implies

$$\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu) = \mathcal{T}_{R^\varepsilon_{0t}}(\nu|\mu) + \mathcal{H}(\mu|m) = \mathcal{T}_{R^\varepsilon_{0t}}(\mu|\nu) + \mathcal{H}(\nu|m), \tag{1.15}$$

[1] We adopt the standard convention that $+\infty - c = +\infty$, if $c < +\infty$

and taking $\mu = m$ (or equivalently $\nu = m$),

$$\mathcal{T}_{R^\varepsilon_{0t}}(m,\nu) = \mathcal{T}_{R^\varepsilon_{0t}}(\nu,m) = \mathcal{T}_{R^\varepsilon_{0t}}(\nu|m) = \mathcal{T}_{R^\varepsilon_{0t}}(m|\nu) + \mathcal{H}(\nu|m). \tag{1.16}$$

Remark 1.2. The entropic transportation cost is larger than the quadratic Wasserstein distance. Indeed, it follows from [13], Corollary 5.13, and the Benamou–Brenier formula [4] that for all $\varepsilon > 0$, $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$:

$$\varepsilon \mathcal{T}_{R^\varepsilon_{01}}(\mu,\nu) \geq \frac{\varepsilon}{2}\mathcal{H}(\mu|m) + \frac{\varepsilon}{2}\mathcal{H}(\nu|m) + \frac{1}{2}W_2^2(\mu,\nu).$$

2. Entropic Talagrand inequality and properties

The family of inequalities we consider in this article has been introduced in the recent article [10] where it was shown that, on a smooth compact manifold M satisfying the Bakry–Émery condition

$$\forall x \in M, \quad \mathfrak{Ric}_x + 2\,\mathrm{Hess}_x U \geq \lambda \mathrm{id} \tag{2.1}$$

we have that for all $\mu, \nu \in \mathcal{P}(M)$, $s \in (0,1)$ and $\varepsilon > 0$:

$$\mathcal{T}_{R^\varepsilon_{01}}(\mu,\nu) \leq \frac{1}{1-\exp(-\lambda\varepsilon s)}\mathcal{H}(\mu|m) + \frac{1}{1-\exp(-\lambda\varepsilon(1-s))}\mathcal{H}(\nu|m). \tag{2.2}$$

In view of Lemma 1.1 and (1.15), the latter is equivalent to

$$\mathcal{T}_{R^\varepsilon_{01}}(\nu|\mu) \leq \frac{1}{\exp(\lambda\varepsilon s)-1}\mathcal{H}(\mu|m) + \frac{1}{1-\exp(-\lambda\varepsilon(1-s))}\mathcal{H}(\nu|m). \tag{2.3}$$

Also, observe that setting $\nu = m$ and optimizing over s in (2.2) yields

$$\mathcal{T}_{R^\varepsilon_{01}}(\mu,m) \leq \frac{1}{1-\exp(-\lambda\varepsilon)}\mathcal{H}(\mu|m). \tag{2.4}$$

This motivates the following definition.

Assumption 2.1. Let $m = \exp(-2U(x))\,\mathrm{d}x \in \mathcal{P}(\mathbb{R}^d)$ with U such that (1.2) admits a weak solution and let R^ε_{0t} the joint law at time 0 and t of the path measure associated to (1.2).

Definition 2.1 (Entropic Talagrand inequalities). Let $m \in \mathcal{P}(\mathbb{R}^d)$ be such that Assumption 2.1 is satisfied and fix $\lambda > 0$, $0 \leq s < t$.

(i) We say that m satisfies the entropic Talagrand inequality $\mathrm{ETI}(\lambda,\varepsilon,s,t)$ if for all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$,

$$\mathcal{T}_{R^\varepsilon_{0t}}(\mu,\nu) \leq \frac{1}{1-\exp(-\lambda\varepsilon s)}\mathcal{H}(\mu|m) + \frac{1}{1-\exp(-\lambda\varepsilon(t-s))}\mathcal{H}(\nu|m).$$

(ii) We say that m satisfies the entropic Talagrand inequality $\mathrm{ETI}(\lambda, \varepsilon, t)$ if for all $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\mathcal{T}_{R^\varepsilon_{0t}}(\mu, m) \le \frac{1}{1-\exp(-\lambda\varepsilon t)}\mathcal{H}(\mu|m).$$

Let us recall that once the measure m is fixed, the law R^ε_{0t} is uniquely determined as the two-times marginal of the Langevin dynamics (1.2).

Remark 2.1 (Hierarchy of the entropic Talagrand inequalities). It can be deduced from the Benamou–Brenier formulation of the entropic transportation cost (see, e.g., [16]) that the function $t \mapsto t\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu)$ is increasing. Therefore, if $\mathrm{ETI}(\lambda, \varepsilon, s, \tau)$ holds for some $\lambda, \varepsilon, s, \tau > 0$ with $s < \tau$, then for any fixed $s < t < \tau$ we have

$$\mathcal{T}_{R^\varepsilon_{0t}}(\mu, \nu) \le \frac{\tau}{t}\frac{1}{1-\exp(-\lambda\varepsilon s)}\mathcal{H}(\mu|m) + \frac{\tau}{t}\frac{1}{1-\exp(-\lambda\varepsilon(t-s))}\mathcal{H}(\nu|m) \tag{2.5}$$

uniformly in $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. Consider now any λ' with the property that

$$\lambda' \le \min\left\{-\frac{\log(1-\frac{t}{\tau}(1-\exp(-\lambda\varepsilon s)))}{\varepsilon s}, -\frac{\log(1-\frac{t}{\tau}(1-\exp(-\lambda\varepsilon(t-s))))}{\varepsilon(t-s)}\right\},$$

where we remark that both quantities on the right hand side are strictly positive under the current assumptions. It is easily verified that this choice of λ' yields

$$\begin{aligned}\frac{\tau}{t}\frac{1}{1-\exp(-\lambda\varepsilon s)}\mathcal{H}(\mu|m) &+ \frac{\tau}{t}\frac{1}{1-\exp(-\lambda\varepsilon(t-s))}\mathcal{H}(\nu|m)\\ &\le \frac{1}{1-\exp(-\lambda'\varepsilon s)}\mathcal{H}(\mu|m) + \frac{1}{1-\exp(-\lambda'\varepsilon(t-s))}\mathcal{H}(\nu|m).\end{aligned}$$

uniformly in $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$. In conclusion, we have obtained that $\mathrm{ETI}(\lambda, \varepsilon, s, \tau)$ implies $\mathrm{ETI}(\lambda', \varepsilon, s, t)$.

2.1. Equivalent forms of the entropic Talagrand inequalities

In this section, we state and prove several equivalent characterizations of $\mathrm{ETI}(\lambda, \varepsilon, s, t)$ and $\mathrm{ETI}(\lambda, \varepsilon, t)$ in terms of reverse hypercontractivity for the heat semigroup (Theorem 2.2) contractivity of the HJB semigroup (Theorem 2.3) and dimension-free concentration of measure (Theorem 2.4).

A weak form of reverse hypercontractivity

To recall the notions of hypercontractivity ([26]) and reverse hypercontractivity we first recall the definition of the heat semigroup $(P^\varepsilon_t)_{t\ge 0}$ associated with (1.2),

$$\forall f > 0, \quad P^\varepsilon_t f(x) := \int_{\mathbb{R}^d} f(y) r^\varepsilon_t(x, \mathrm{d}y),$$

where $r_t^\varepsilon(x, \mathrm{d}y)$ is the transition kernel for R_{0t}^ε. Note that we have the scaling relation

$$\forall \varepsilon, t > 0, x \in \mathbb{R}^d, f > 0, \quad P_t^\varepsilon f(x) = P_{\varepsilon t}^1 f(x). \tag{2.6}$$

For $f \geq 0$, $p \in \mathbb{R}$, $p \neq 0$, we set

$$\|f\|_p := \left(\int_{\mathbb{R}^d} f^p \,\mathrm{d}m\right)^{1/p}. \tag{2.7}$$

Note that we do not ask $p > 1$. For an $f > 0$ such that $\log f$ is integrable, the norm $\|f\|_0$ is defined by

$$\|f\|_0 = \exp\left(\int_{\mathbb{R}^d} \log f \,\mathrm{d}m\right).$$

Definition 2.2 (Hypercontractivity and reverse hypercontractivity). Let $\lambda, \varepsilon > 0$. The semigroup $(P_t^\varepsilon)_{t\geq 0}$ is λ-hypercontractive if for all $t > 0$, $p > 1$ and $f > 0$ it holds that

$$\left\|P_t^\varepsilon f\right\|_q \leq \|f\|_p, \quad \text{where } \frac{q-1}{p-1} = e^{2\lambda\varepsilon t}.$$

On the other hand, λ-reverse hypercontractivity is defined asking that for all $t > 0$, $p < 1$ and $f > 0$,

$$\left\|P_t^\varepsilon f\right\|_q \geq \|f\|_p, \quad \text{where } \frac{q-1}{p-1} = e^{2\lambda\varepsilon t}.$$

Next, the theorem shows that the dual form of $\mathrm{ETI}(\lambda, \varepsilon, s, t)$ encodes a weaker form of $\frac{\lambda}{2}$-reverse hypercontractivity; we recall that Gross established in [21] equivalence between the logarithmic Sobolev inequality and hypercontractivity. The equivalence between (full) reverse hypercontractivity and Log Sobolev is also known [1], Theorem 3.3. In the proof, and in the rest of the article, we take advantage of the notation

$$\theta_{\lambda\varepsilon}(s) := \frac{1}{1 - \exp(-\lambda\varepsilon s)}. \tag{2.8}$$

Theorem 2.2 ($\mathrm{ETI}(\lambda, \varepsilon, s, t)$ **and reverse hypercontractivity).** *Let $m \in \mathcal{P}(\mathbb{R}^d)$ be such that Assumption 2.1 is satisfied. For $t \geq 0$, the following are equivalent*

(i) *m satisfies $\mathrm{ETI}(\lambda, \varepsilon, s, t)$ for all $s \in [0, t]$.*
(ii) *For all $f > 0$ and $p, q \in [0, 1) \times (-\infty, 0]$ such that $\frac{q-1}{p-1} = \exp(\lambda\varepsilon t)$ we have*

$$\left\|P_t^\varepsilon f\right\|_q \geq \|f\|_p. \tag{2.9}$$

Remark 2.2. Because of the restrictions on (p, q), we were not able to conclude that the weak form of reverse hypercontractivity of Theorem 2.2 is equivalent to the log-Sobolev inequality. This would be true if the constraint $(p, q) \in (0, 1) \times (-\infty, 0)$ could be dropped, see the classical

reference [1], Theorem 3.3. However, we will see at Corollary 2.1 below that the weak reverse hypercontractivity implies a Poincaré inequality.

Proof. In the proof, we set for simplicity $t = 1$. Inspired by [20], Proposition 4.5, which generalizes some of the results in [7], we look for the dual formulation of ETI$(\lambda, \varepsilon, s, 1)$. First, we rewrite it multiplying by ε as,

$$\forall \mu, \nu \in \mathcal{P}\big(\mathbb{R}^d\big), \quad \varepsilon \mathcal{T}_{R^\varepsilon_{01}}(\mu, \nu) \le \varepsilon\theta_{\lambda\varepsilon}(s)\mathcal{H}(\mu|m) + \varepsilon\theta_{\lambda\varepsilon}(1-s)\mathcal{H}(\nu|m). \tag{2.10}$$

The dual formulation (1.3) tells that (i) is equivalent to say that for all $s \in (0,1)$, $\varphi \in C_b(\mathbb{R}^d)$, $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ we have

$$\varepsilon\mathcal{H}(\mu|m) + \int Q^\varepsilon_t \varphi \,\mathrm{d}\mu - \int \varphi \,\mathrm{d}\nu \le \varepsilon\theta_{\lambda\varepsilon}(s)\mathcal{H}(\mu|m) + \varepsilon\theta_{\lambda\varepsilon}(1-s)\mathcal{H}(\nu|m).$$

Rearranging the terms, we can rewrite the latter as

$$\begin{aligned}\varepsilon\big(\theta_{\lambda\varepsilon}(s)-1\big)&\bigg(\int \frac{Q^\varepsilon_1\varphi}{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)}\,\mathrm{d}\mu - \mathcal{H}(\mu|m)\bigg)\\ &+ \varepsilon\theta_{\lambda\varepsilon}(1-s)\bigg(\int -\frac{\varphi}{\varepsilon\theta_{\lambda\varepsilon}(1-s)}\,\mathrm{d}\nu - \mathcal{H}(\nu|m)\bigg) \le 0.\end{aligned}$$

We now take the suprema over μ and ν and use the variational formula (A.1), to obtain that (i) is equivalent to the fact that for all $s \in (0,1)$ and $\varphi \in C_b(\mathbb{R}^d)$

$$\begin{aligned}\varepsilon\big(\theta_{\lambda\varepsilon}(s)-1\big)\log&\int \exp\bigg(\frac{Q^\varepsilon_1\varphi}{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)}\bigg)\,\mathrm{d}m\\ &+ \varepsilon\theta_{\lambda\varepsilon}(1-s)\log\int \exp\bigg(-\frac{\varphi}{\varepsilon\theta_{\lambda\varepsilon}(1-s)}\bigg)\,\mathrm{d}m \le 0.\end{aligned}$$

Taking exponentials we get

$$\bigg(\int \exp\bigg(\frac{Q^\varepsilon_1\varphi}{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)}\bigg)\,\mathrm{d}m\bigg)^{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)}\bigg(\int \exp\bigg(-\frac{\varphi}{\varepsilon}\theta_{\lambda\varepsilon}(1-s)\bigg)\,\mathrm{d}m\bigg)^{\varepsilon\theta_{\lambda\varepsilon}(1-s)} \le 1. \tag{2.11}$$

Using (1.6) and setting $\exp(-\varphi/\varepsilon) = f$, we obtain

$$\bigg(\int \big(P^\varepsilon_1 f\big)^{-1/(\theta_{\lambda\varepsilon}(s)-1)}\,\mathrm{d}m\bigg)^{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)}\bigg(\int f^{1/\theta_{\lambda\varepsilon}(1-s)}\,\mathrm{d}m\bigg)^{\varepsilon\theta_{\lambda\varepsilon}(1-s)} \le 1.$$

Raising to the power of $1/\varepsilon$, using (2.7) and setting $q(\lambda\varepsilon, s) = -1/(\theta_{\lambda\varepsilon}(s)-1)$, $p(\lambda\varepsilon, s) = 1/\theta_{\lambda\varepsilon}(1-s)$ we obtain a new equivalent formulation of (i) after a simple approximation argument:

$$\forall s \in (0,1),\ f > 0, \quad \big\|P^\varepsilon_1 f\big\|_{q(\lambda\varepsilon,s)} \ge \|f\|_{p(\lambda\varepsilon,s)}.$$

To conclude the proof, we first observe that

$$\big\{\big(p(\lambda\varepsilon,s),q(\lambda\varepsilon,s)\big):s\in(0,1)\big\}=\left\{(p,q)\in(0,1)\times(-\infty,0):\frac{q-1}{p-1}=\exp(\lambda\varepsilon)\right\}.$$

The case $p=0$ is obtained with a standard approximation argument. □

The dual formulation of TI(λ) is equivalent to some contraction properties for the Hopf-Lax semigroup, see [3], Proposition 9.2.3. Here we show that ETI(λ,ε,t) admits a dual formulation in terms of contraction properties for the HJB semigroup.

Theorem 2.3 (ETI(λ,ε,t) **and the HJB semigroup).** *Let $m\in\mathcal{P}(\mathbb{R}^d)$ be such that Assumption 2.1 is satisfied. The following are equivalent*

(i) ETI(λ,ε,t) *holds.*
(ii) *For all $\varphi\in C_b(\mathbb{R}^d)$,*

$$\int\exp\left(-\frac{1}{\varepsilon\theta_{\lambda\varepsilon}(t)}\varphi\right)\mathrm{d}m\le\exp\left(-\frac{1}{\varepsilon\theta_{\lambda\varepsilon}(t)}\int Q_t^\varepsilon\varphi\,\mathrm{d}m\right).$$

(iii) *For all $\psi\in C_b(\mathbb{R}^d)$,*

$$\int\exp\big(Q_{Ct}^{\varepsilon/C}\psi\big)\,\mathrm{d}m\le\exp\left(\int\psi\,\mathrm{d}m\right),\tag{2.12}$$

where

$$C=\varepsilon\big(\theta_{\lambda\varepsilon}(t)-1\big).$$

Remark that letting $\varepsilon\to0$ in (2.12) gives back, at least formally, the above mentioned characterization of TI(λ).

Proof of Theorem 2.3. We follow the same arguments as in the proof of Theorem 2.2. Again, w.l.o.g. we fix $t=1$. To prove (ii), we multiply ETI($\lambda,\varepsilon,1$) by ε and recall that according to the Kantorovich dual formulation (1.3) and the symmetric property for the entropic cost we have,

$$\varepsilon\mathcal{T}_{R_{01}^\varepsilon}(\mu,m)=\sup_{\varphi\in C_b(\mathbb{R}^d)}\left\{\int Q_1^\varepsilon\varphi\,\mathrm{d}m-\int\varphi\,\mathrm{d}\mu\right\}.$$

Plugging this into ETI($\lambda,\varepsilon,1$) yields the equivalent formulation

$$\forall\varphi\in C_b\big(\mathbb{R}^d\big),\quad\int Q_1^\varepsilon\varphi\,\mathrm{d}m-\int\varphi\,\mathrm{d}\mu-\varepsilon\theta_{\lambda\varepsilon}(1)\mathcal{H}(\mu|m)\le0.$$

This can be rewritten as,

$$\forall\varphi\in C_b\big(\mathbb{R}^d\big),\quad\frac{1}{\varepsilon\theta_{\lambda\varepsilon}(1)}\int Q_1^\varepsilon\varphi\,\mathrm{d}m+\left(\int-\frac{\varphi}{\varepsilon\theta_{\lambda\varepsilon}(1)}\,\mathrm{d}\mu-\mathcal{H}(\mu|m)\right)\le0.$$

Taking the supremum over μ and exponentiating, we obtain the desired result thanks to (A.1). The proof of (iii) is analogue. We start from the Kantorovich formulation of the entropic cost (1.3) to obtain that ETI($\lambda, \varepsilon, 1$) is equivalent to the property that for all $\varphi \in C_b(\mathbb{R}^d)$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$,

$$\varepsilon \mathcal{H}(\mu|m) + \sup_{\varphi \in C_b(\mathbb{R}^d)} \left\{ \int Q_1^\varepsilon \varphi \, \mathrm{d}\mu - \int \varphi \, \mathrm{d}m \right\} \leq \varepsilon \theta_{\lambda\varepsilon}(1) \mathcal{H}(\mu|m).$$

Rearranging terms, taking sumpremum over μ, using (A.1) we arrive at the following equivalent form of ETI($\lambda, \varepsilon, 1$)

$$\forall \varphi \in C_b\big(\mathbb{R}^d\big), \quad \varepsilon\big(\theta_{\lambda\varepsilon}(1) - 1\big) \log \int \exp\left(\frac{Q_1^\varepsilon \varphi}{\varepsilon(\theta_{\lambda\varepsilon}(1) - 1)} \right) \mathrm{d}m - \int \varphi \, \mathrm{d}m \leq 0.$$

The conclusion follows by exponentiating, setting $\psi = \varphi / C$ and an application of the scaling relation (see (2.6))

$$\frac{1}{C} Q_t^\varepsilon (C\psi) = Q_{tC}^{\varepsilon/C}(\psi). \qquad \square$$

2.2. Properties of entropic Talagrand inequalities

In the next lines, we investigate tensorization of ETI(λ, ε, t) and ETI($\lambda, \varepsilon, s, t$). In what follows we adopt the following convention: if $p(x, \cdot)$ is a probability kernel on $\mathbb{R}^{d_1} \times \cdots \times \mathbb{R}^{d_n}$ we write $p_i(x, \cdot) \in \mathcal{P}(\mathbb{R}^{d_i})$ for the i-th marginal distribution of $p(x, \cdot)$.

Proposition 2.1 (Tensorization: first form). *Let* $n \in \mathbb{N}$, $1 \leq i \leq n$ *and* $m_i(\mathrm{d}x) = \exp(-2U_i(x))\,\mathrm{d}x \in \mathcal{P}(\mathbb{R}^{d_i})$ *such that Assumption 2.1 is satisfied and satisfy* ETI($\lambda, \varepsilon, s, t$). *Then* $m = m_1 \otimes \cdots \otimes m_n$ *satisfies* ETI($\lambda, \varepsilon, s, t$).

We recall that in the above proposition the entropic cost for $m_1 \otimes \cdots \otimes m_n$ is the one corresponding to the law of n independent diffusions of the form (1.2) associated with the potentials U_i, $i = 1, \ldots, n$.

Proof of Proposition 2.1. We assume again w.l.o.g. that $t = 1$. Recall that ETI($\lambda, \varepsilon, s, 1$) for m_i has the equivalent form

$$\mathcal{T}_{R_{01}^{\varepsilon,i}}(\nu|\mu) \leq \big(\theta_{\lambda\varepsilon}(s) - 1\big)\mathcal{H}(\mu|m_i) + \theta_{\lambda\varepsilon}(1-s)\mathcal{H}(\nu|m_i),$$

where $R_{01}^{\varepsilon,i}$ is the two times law of the Langevin dynamics for m_i. By induction, it is also enough to consider only the case $n = 2$. Consider now $\mu, \nu \in \mathcal{P}(\mathbb{R}^{d_1+d_2})$, and assume that $\mathcal{T}_{R_{01}^\varepsilon}(\mu, \nu) < +\infty$. Then, there exist an optimal kernel $\tilde{p}_1(x_1, \mathrm{d}y_1)$ such that

$$\int_{\mathbb{R}^{d_1}} \mathcal{H}\big(\tilde{p}_1(x_1, \cdot)|r_1^{\varepsilon,1}(x_1, \cdot)\big)\mu_1(\mathrm{d}x_1) = \mathcal{T}_{R_{01}^{\varepsilon,1}}(\nu_1|\mu_1),$$

where we denoted μ_1, ν_1 the image laws of μ and ν through the projection on the first d_1 coordinates.

Moreover, for any fixed $x_1, y_1 \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_1}$ there exist an optimal kernel $q^{x_1,y_1}(x_2, \mathrm{d}y_2)$ on $\mathbb{R}^{d_2}$ such that

$$\int_{\mathbb{R}^{d_2}} \mathcal{H}\big(q^{x_1,y_1}(x_2,\cdot)|r_1^{\varepsilon,2}(x_2,\cdot)\big)\mu(x_1,\mathrm{d}x_2) = \mathcal{T}_{R_{01}^{\varepsilon,2}}\big(\nu(\cdot|y_1)|\mu(x_1,\cdot)\big),$$

where $\mu(x_1,\cdot)$ (resp. $\nu(y_1,\cdot)$) is the kernel defined via $\mu(\mathrm{d}x_1\,\mathrm{d}x_2) = \mu_1(\mathrm{d}x_1)\mu(x_1,\mathrm{d}x_2)$ (resp. $\nu(\mathrm{d}y_1\,\mathrm{d}y_2) = \nu_1(\mathrm{d}y_1)\nu(y_1,\mathrm{d}y_2)$).

We can construct a coupling π of μ and ν setting,

$$\pi(\mathrm{d}x\,\mathrm{d}y) = \mu(\mathrm{d}x_1\,\mathrm{d}x_2)p(x,\mathrm{d}y), \qquad p(x,\mathrm{d}y) = \tilde{p}_1(x_1,\mathrm{d}y_1)q^{x_1,y_1}(x_2,\mathrm{d}y_2). \tag{2.13}$$

Note that for any x we have $p_1(x,\cdot) = \tilde{p}_1(x_1,\cdot)$ and

$$p_2(x,\cdot) = \int_{\mathbb{R}^{d_1}} \tilde{p}_1(x_1,\mathrm{d}y_1)q^{x_1,y_1}(x_2,\cdot).$$

Since the Langevin dynamics for $m_1 \times m_2$ is the product of the Langevin dynamics for m_1 and m_2 we have

$$\mathcal{T}_{R_{01}^{\varepsilon}}(\mu,\nu) \le \int_{\mathbb{R}^{d_1+d_2}} \mathcal{H}\big(\tilde{p}_1(x_1,\cdot)q^{x_1,y_1}(x_2,\cdot)|r_1^{\varepsilon,1}(x_1,\cdot)\otimes r_1^{\varepsilon,2}(x_2,\cdot)\big)\mu(\mathrm{d}x_1\,\mathrm{d}x_2).$$

Thanks to the decomposition of the entropy formula (A.2) we have for all μ almost all x_1, x_2

$$\begin{aligned}
&\mathcal{H}\big(\tilde{p}_1(x_1,\cdot)q^{x_1,y_1}(x_2,\cdot)|r_1^{\varepsilon,1}(x_1,\cdot)\otimes r_1^{\varepsilon,2}(x_2,\cdot)\big)\\
&\quad= \mathcal{H}\big(\tilde{p}_1(x_1,\cdot)|r_1^{\varepsilon,1}(x_1,\cdot)\big) + \int_{\mathbb{R}^{d_1}} \mathcal{H}\big(q^{x_1,y_1}(x_2,\cdot)|r_1^{\varepsilon,2}(x_2,\cdot)\big)\tilde{p}_1(x_1,\mathrm{d}y_1).
\end{aligned}$$

Plugging this into the above formula and using the optimality of the couplings yields

$$\begin{aligned}
\mathcal{T}_{R_{01}^{\varepsilon}}(\mu,\nu) &\le \mathcal{T}_{R_{01}^{\varepsilon,1}}(\nu_1|\mu_1)\\
&\quad+ \int_{\mathbb{R}^{d_1+d_1}}\left[\int_{\mathbb{R}^{d_2}} \mathcal{H}\big(q^{x_1,y_1}(x_2,\cdot)|r_1^{\varepsilon,2}(x_2,)\big)\mu(x_1,\mathrm{d}x_2)\right]\tilde{p}_1(x_1,\mathrm{d}y_1)\mu_1(\mathrm{d}x_1)\\
&= \mathcal{T}_{R_{01}^{\varepsilon,1}}(\nu_1|\mu_1) + \int_{\mathbb{R}^{d_1+d_1}} \mathcal{T}_{R_{01}^{\varepsilon,2}}\big(\nu(y_1,\cdot)|\mu(x_1,\cdot)\big)\tilde{p}_1(x_1,\mathrm{d}y_1)\mu_1(\mathrm{d}x_1).
\end{aligned}$$

Applying ETI$(\lambda,\varepsilon,s,1)$ and the fact that $\mu_1(x_1)\tilde{p}_1(x_1,\mathrm{d}y_1) = \nu_1(\mathrm{d}y_1)$ we get

$$\begin{aligned}
\mathcal{T}_{R_{01}^{\varepsilon}}(\mu,\nu) &\le \big(\theta_{\lambda\varepsilon}(s)-1\big)\left[\mathcal{H}(\mu_1|m_1) + \int_{\mathbb{R}^{d_1+d_1}} \mathcal{H}\big(\mu(x_1,\cdot)|m_2\big)\tilde{p}_1(x_1,\mathrm{d}y_1)\mu_1(\mathrm{d}x_1)\right]\\
&\quad+ \theta_{\lambda\varepsilon}(1-s)\left[\mathcal{H}(\nu_1|m_1) + \int_{\mathbb{R}^{d_1+d_1}} \mathcal{H}\big(\nu(y_1,\cdot)|m_2\big)\tilde{p}_1(x_1,\mathrm{d}y_1)\mu_1(\mathrm{d}x_1)\right]
\end{aligned}$$

$$
\begin{aligned}
&= \big(\theta_{\lambda\varepsilon}(s) - 1\big)\Big[\mathcal{H}(\mu_1|m_1) + \int_{\mathbb{R}^{d_1}} \mathcal{H}\big(\nu(\cdot|y_1)|m_2\big)\mu_1(\mathrm{d}x_1)\Big] \\
&\quad + \theta_{\lambda\varepsilon}(1-s)\Big[\mathcal{H}(\nu_1|m_1) + \int_{\mathbb{R}^{d_1}} \mathcal{H}\big(\nu(\cdot|y_1)|m_2\big)\nu_1(\mathrm{d}y_1)\Big] \\
&= \big(\theta_{\lambda\varepsilon}(s) - 1\big)\mathcal{H}(\mu|m_1 \otimes m_2) + \theta_{\lambda\varepsilon}(1-s)\mathcal{H}(\nu|m_1 \otimes m_2),
\end{aligned}
$$

where the last equality follows from the decomposition of the entropy formula (A.2). □

A second form of tensorization holds, following [20].

Proposition 2.2 (Tensorization: second form). *Let $n \in \mathbb{N}$, $1 \le i \le n$ and $m_i \in \mathcal{P}(\mathbb{R}^{d_i})$ such that Assumption 2.1 is satisfied and satisfy* $\mathrm{ETI}(\lambda, \varepsilon, s, t)$ *. Then $m = m_1 \otimes \cdots \otimes m_n$ satisfies the following inequality*

$$
\forall \mu, \nu \in \mathcal{P}\big(\mathbb{R}^{d_1+\cdots+d_n}\big), \quad \bar{\mathcal{T}}(\nu|\mu) \le \big(\theta_{\lambda\varepsilon}(s) - 1\big)\mathcal{H}(\mu|m) + \theta_{\lambda\varepsilon}(1-s)\mathcal{H}(\nu|m), \tag{2.14}
$$

where

$$
\bar{\mathcal{T}}(\nu|\mu) = \inf_{\pi \in \Pi(\mu,\nu)} \int \sum_{i=1}^{n} \mathcal{H}\big(p_i(x,\cdot)|r_1^{\varepsilon,i}(x_i,\cdot)\big)\mu(\mathrm{d}x).
$$

Proof. The proof follows the same lines as the former one. As before, we can restrict to $n = 2$, and construct the coupling π via (2.13). Note that

$$
p_2(x,\cdot) = \int_{\mathbb{R}^{d_1}} \tilde{p}_1(x_1, \mathrm{d}y_1) q^{x_1,y_1}(x_2,\cdot).
$$

We have

$$
\begin{aligned}
\bar{\mathcal{T}}(\nu|\mu) &\le \int_{\mathbb{R}^{d_1+d_2}} \mathcal{H}\big(p_1(x,\cdot)|r_1^{\varepsilon,1}(x_1,\cdot)\big)\mu(\mathrm{d}x) + \int_{\mathbb{R}^{d_1+d_2}} \mathcal{H}\big(p_2(x,\cdot)|r_1^{\varepsilon,2}(x_2,\cdot)\big)\mu(\mathrm{d}x) \\
&= \int_{\mathbb{R}^{d_1+d_2}} \mathcal{H}\big(\tilde{p}_1(x_1,\cdot)|r_1^{\varepsilon,1}(x_1,\cdot)\big)\mu_1(\mathrm{d}x_1) \\
&\quad + \int_{\mathbb{R}^{d_1+d_2}} \mathcal{H}\Big(\int_{\mathbb{R}^{d_1}} \tilde{p}_1(x_1, dy_1) q^{x_1,y_1}(x_2,\cdot)|r_1^{\varepsilon,2}(x_2,\cdot)\Big)\mu(x_1,\mathrm{d}x_2)\mu_1(\mathrm{d}x_1) \\
&\le \mathcal{T}_{R_{01}^{\varepsilon,1}}(\nu_1|\mu_1) \\
&\quad + \int_{\mathbb{R}^{d_1+d_1+d_2}} \mathcal{H}\big(q^{x_1,y_1}(x_2,\cdot)|r_1^{\varepsilon,2}(x_2,\cdot)\big)\mu(x_1,\mathrm{d}x_2)\tilde{p}_1(x_1,dy_1)\mu_1(\mathrm{d}x_1) \\
&= \mathcal{T}_{R_{01}^{\varepsilon,1}}(\nu_1|\mu_1) \\
&\quad + \int_{\mathbb{R}^{d_1+d_1}} \mathcal{T}_{R_{01}^{\varepsilon,1}}\big(\nu(y_1,\cdot)|\mu(x_1,\cdot)\big)\tilde{p}_1(x_1,\mathrm{d}y_1)\mu_1(\mathrm{d}x_1),
\end{aligned}
$$

where the last inequality is a consequence of the convexity of the relative entropy. From now on, the proof goes as in the former proposition. □

It can be easily seen that Propositions 2.1 and 2.2 are valid also for $\mathrm{ETI}(\lambda, \varepsilon, t)$. However, we propose here an alternative proof of the tensorization property for $\mathrm{ETI}(\lambda, \varepsilon, t)$, in the same spirit of [3], Proposition 9.2.4.

Proposition 2.3 (Tensorization: third form). *Let $n \in \mathbb{N}$, $1 \le i \le n$ and $m_i \in \mathcal{P}(\mathbb{R}^{d_i})$ such that Assumption 2.1 is satisfied and satisfy* $\mathrm{ETI}(\lambda, \varepsilon, t)$. *Then $m_1 \otimes \cdots \otimes m_n$ satisfies* $\mathrm{ETI}(\lambda, \varepsilon, t)$.

Proof. For any $\varepsilon > 0$, let P_t^ε, Q_t^ε be the heat and HJB semigroups for $m_1 \times m_2$. Also, we note $P_t^{\varepsilon,1}$(resp. $P_t^{\varepsilon,2}$) and $Q_t^{\varepsilon,1}$ (resp. $Q_t^{\varepsilon,2}$) the same semigroups for m_1 (resp. m_2). To obtain the result, we show that the equivalent form (iii) in Theorem 2.3 of $\mathrm{ETI}(\lambda, \varepsilon, t)$ holds. To this aim, we observe that, thanks to the fact that the Langevin dynamics for $m_1 \times m_2$ is the product of the Langevin dynamics for m_1 and m_2, we have for all $\varepsilon, t > 0$ and $x_1, x_2 \in \mathbb{R}^d$:

$$Q_{ct}^{\varepsilon/c}\varphi(x_1, x_2) = Q_{ct}^{\varepsilon/c,1}\big(Q_{ct}^{\varepsilon/c,2,\cdot}\psi(x_2)\big)(x_1), \tag{2.15}$$

where, for any $(y_1, x_2) \in \mathbb{R}^d \times \mathbb{R}^d$,

$$Q_{ct}^{\varepsilon/c,2,y_1}\psi(x_2) = Q_{ct}^{\varepsilon/c,2}\big(\psi(y_1, \cdot)\big)(x_2).$$

Using (2.15) and $\mathrm{ETI}(\lambda, \varepsilon, t)$ for m_1 we obtain,

$$\begin{aligned}
&\int \exp\big(Q_{ct}^{\varepsilon/c}\big)\varphi(x_1, x_2) m_1 \otimes m_2(\mathrm{d}x_1\,\mathrm{d}x_2) \\
&\quad \le \int \exp\bigg(\int Q_{ct}^{\varepsilon/c,2,x_1}\psi(x_2)\mu(\mathrm{d}x_1)\bigg)\mu(\mathrm{d}x_2).
\end{aligned}$$

Using the definition of $Q_{ct}^{\varepsilon/c,2}$ as an infimum (1.4), we obtain

$$\int Q_{ct}^{\varepsilon/c,2,x_1}\mu(\mathrm{d}x_1) \le Q_{ct}^{\varepsilon/c,2}\bigg(\int \psi(x_1, \cdot)\mu(\mathrm{d}x_1)\bigg)(x_2).$$

Using this and $\mathrm{ETI}(\lambda, \varepsilon, t)$ for m_2 we get

$$\begin{aligned}
&\int \exp\bigg(\int Q_{ct}^{\varepsilon/c,2,x_1}\psi(x_2) m_1(\mathrm{d}x_1)\bigg) m_2(\mathrm{d}x_2) \\
&\quad \le \int \exp\bigg(Q_{ct}^{\varepsilon/c,2}\bigg(\int \psi(x_1, \cdot) m_1(\mathrm{d}x_1)\bigg)(x_2)\bigg) m_2(\mathrm{d}x_2) \\
&\quad \le \exp\bigg(\int \psi(x_1, x_2) m_1 \otimes m_2(\mathrm{d}x_1\,\mathrm{d}x_2)\bigg)
\end{aligned}$$

which is the desired conclusion. □

The tensorization property allows us to give a further characterization of $\mathrm{ETI}(\lambda, \varepsilon, s, t)$ via a dimension free concentration property. Let us first introduce some notation. For $m \in \mathcal{P}(\mathbb{R}^d)$ we denote $m^n = \underbrace{m \otimes \cdots \otimes m}_{n \text{ times}} \in \mathcal{P}(\mathbb{R}^{d\times n}) = \mathcal{P}(\underbrace{\mathbb{R}^d \times \cdots \times \mathbb{R}^d}_{n \text{ times}})$; for any $t > 0$ $R_{0t}^{\varepsilon,n}$ is the joint law of the reference measure with reversing measure m^n and generator $\mathcal{L}^{\varepsilon,n} = \mathcal{L}^\varepsilon \oplus \cdots \oplus \mathcal{L}^\varepsilon$, $r_t^{\varepsilon,n}(x,\cdot)$ its Markov kernel and $(P_t^{\varepsilon,n})_{t\geq 0}$ the associated product Markov semigroup. Finally, in accordance with what we did above we define the corresponding HJB semigroup:

$$Q_t^{\varepsilon,n}\varphi(x) = -\varepsilon \log P_t^{\varepsilon,n} \exp(-\varphi/\varepsilon)(x), \quad \text{for } x \in \mathbb{R}^{d\times n}.$$

For any Borel set $A \subset \mathbb{R}^{d\times n}$ following [20], we consider

$$c_A^n(x) := \inf\big\{\mathcal{H}\big(p|r_1^{\varepsilon,n}(x,\cdot)\big), p \in \mathcal{P}\big(\mathbb{R}^{d\times n}\big), p(A) = 1\big\}, \quad x \in \mathbb{R}^{d\times n}. \tag{2.16}$$

A standard calculation shows that

$$c_A^n(x) = -\log r_1^{\varepsilon,n}(x, A).$$

Moreover, we define for all $u \geq 0$,

$$A_u := \big\{x \in \mathbb{R}^{d\times n} : c_A^n(x) \leq u\big\} = \big\{x \in \mathbb{R}^{d\times n} : r_1^{\varepsilon,n}(x, A) \geq e^{-u}\big\}. \tag{2.17}$$

Remark 2.3. Note that A_u is not in general an enlargement of A, i.e. $A \nsubseteq A_u$.

In the next theorem, we provide an equivalent characterization of $\mathrm{ETI}(\lambda, \varepsilon, s, 1)$ in terms of dimension-free concentration. Note that the tensorization result we use here is Proposition 2.1 and not Proposition 2.2, as it is more natural in this context. Thus, our Theorem 2.4 is close in spirit, but different from Theorem 5.1 in [20].

Theorem 2.4 (Dimension free concentration). *Let R^ε be the stationary Markov process for the generator $\mathcal{L}^\varepsilon$ and $m \in \mathcal{P}(\mathbb{R}^d)$ such that Assumption 2.1 is satisfied. The following are equivalent for $\lambda > 0$ and $s \in [0, 1]$,*

(i) *m satisfies* $\mathrm{ETI}(\lambda, \varepsilon, s, 1)$.
(ii) *For any integer $n \geq 1$, for all Borel set $A \subset \mathbb{R}^{d\times n}$ and any $u \geq 0$ it holds,*

$$m^n\big(\mathbb{R}^{d\times n} \setminus A_u^n\big)^{\theta_{\lambda\varepsilon}(s)-1} m^n(A)^{\theta_{\lambda\varepsilon}(1-s)} \leq e^{-u},$$

with $\theta_{\lambda\varepsilon}(s)$ defined at (2.8).
(iii) *For all integers $n \geq 1$, for all non-negative $\varphi \in C_b(\mathbb{R}^{d\times n})$, it holds,*

$$m^n\big(Q_1^{\varepsilon,n}\varphi > u\big)^{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)} m^n(\varphi \leq v)^{\varepsilon\theta_{\lambda\varepsilon}(1-s)} \leq e^{v-u}$$

for all $v \in \mathbb{R}$ and u s.t. $u - v > 0$.

Proof. The proof follows the one of [20], Theorem 5.1. For completeness we recall here some key points. The implication (i) $\Rightarrow$ (ii) is a generalization to the entropic transportation inequality of Marton's argument. Since m satisfies ETI$(\lambda, \varepsilon, s, 1)$ then thanks to Proposition 2.1 the same holds for m^n. As observed at Remark 1.1, ETI$(\lambda, \varepsilon, s, 1)$ can be equivalently written as

$$\mathcal{T}_{R_{01}^{\varepsilon,n}}(\nu|\mu) \leq \big(\theta_{\lambda\varepsilon}(s) - 1\big)\mathcal{H}\big(\mu|m^n\big) + \theta_{\lambda\varepsilon}(1-s)\mathcal{H}\big(\nu|m^n\big)$$

for all $\mu, \nu \in \mathcal{P}(\mathbb{R}^{d\times n})$. For $A \subset \mathbb{R}^{d\times n}$ we choose the couple of probability measures $\mu(\mathrm{d}x) = \mathbf{1}_B/m^n(B)m^n(\mathrm{d}x)$ and $\nu(\mathrm{d}x) = \mathbf{1}_A/m^n(A)m^n(\mathrm{d}x)$ where $B = \mathbb{R}^{d\times n} \setminus A_u$ and A_u is defined at (2.17). Hence direct computations show that $\mathcal{H}(\mu|m^n) = -\log m^n(B)$ and $\mathcal{H}(\nu|m^n) = -\log m^n(A)$. Also, observe that the infimum value in (2.16) can be easily computed, providing $c_A^n(x) = -\log r_1^{\varepsilon,n}(x, A)$. Moreover the set A_u can be rewritten as,

$$A_u = \big\{x \in \mathbb{R}^{d\times n} : r_1^{\varepsilon,n}(x, A) \geq e^{-u}\big\}.$$

To conclude, take any $\pi \in \Pi(\mu, \nu)$ with disintegration kernel $(p_x)_{x\in\mathbb{R}^{d\times n}}$ then

$$\int \mathcal{H}\big(p_x|r_1^{\varepsilon,n}(x,\cdot)\big)\mu(\mathrm{d}x) \geq \int c_A^n(x)\mu(\mathrm{d}x) > u.$$

The conclusion follows by taking the infimum on the set of couplings of μ and ν.

For the implication (ii) $\Rightarrow$ (iii), let $\varphi \in C_b(\mathbb{R}^{d\times n})$ and consider $A = \{\varphi \leq v\}$ for some real v. We show that $\{Q_1^{\varepsilon,n}\varphi > u\} \subset \{c_A^n > u - v\}$. Take $x \in \{Q_1^{\varepsilon,n}\varphi > u\}$, then for all $p \in \mathcal{P}(\mathbb{R}^{d\times n})$ with $p(A) = 1$, and thanks to (1.4) it holds,

$$u < \int \varphi\,\mathrm{d}p + \varepsilon\mathcal{H}\big(p|r_1^{\varepsilon,n}(x,\cdot)\big) \leq v + \varepsilon\mathcal{H}\big(p|r_1^{\varepsilon,n}(x,\cdot)\big).$$

The conclusion follows by optimizing among all the probability $p \in \mathcal{P}(\mathbb{R}^{d\times n})$ such that $p(A) = 1$. To show the last implication (iii) $\Rightarrow$ (i) we fix for simplicity $n = 2$. Let $\delta \in (0, 1)$, f a non-negative function on $\mathbb{R}^d$. Define $\varphi(x) = f(x_1) + f(x_2)$, $x \in \mathbb{R}^{2\times d}$. Then according to [3] it can be verified that $Q_1^{\varepsilon,2}\varphi(x) = Q_1^{\varepsilon}f(x_1) + Q_1^{\varepsilon}f(x_2)$. Hence one has,

$$\begin{aligned}
&\left(\int \exp\left(\frac{Q_1^{\varepsilon}f}{(1+\delta)\varepsilon(\theta_{\lambda\varepsilon}(s)-1)}\right)\mathrm{d}m\right)^{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)} \\
&\qquad\times\left(\int \exp\left(-\frac{f}{(1-\delta)\varepsilon\theta_{\lambda\varepsilon}(1-s)}\right)\mathrm{d}m\right)^{\varepsilon\theta_{\lambda\varepsilon}(1-s)} \\
&= \left(\int \exp\left(\frac{Q_1^{\varepsilon,2}\varphi}{(1+\delta)\varepsilon(\theta_{\lambda\varepsilon}(s)-1)}\right)\mathrm{d}m^2\right)^{\varepsilon(\theta_{\lambda\varepsilon}(s)-1)/2} \\
&\qquad\times\left(\int \exp\left(-\frac{\varphi}{(1-\delta)\varepsilon\theta_{\lambda\varepsilon}(1-s)}\right)\mathrm{d}m^2\right)^{\varepsilon\theta_{\lambda\varepsilon}(1-s)/2}
\end{aligned}$$

the rest of the proof is the same as [20], Theorem 5.1. $\square$

2.3. Relation with other functional inequalities

In this section, we shall see how the entropic Talagrand inequality relates to other well known functional inequalities. First, we provide a new proof via the entropic Talagrand inequality of the fact that the Logarithmic Sobolev inequality implies Talagrand's inequality. This seminal result was first proven by Otto and Villani in [27]. In particular, we show that

$$\text{log-Sobolev ineq.} \Rightarrow \mathrm{ETI}(\lambda, \varepsilon, s, t) \Rightarrow \mathrm{TI}(\lambda).$$

Our argument may be seen as a generalization to the HJB semigroup of the alternative proof of Otto and Villani's result given in [6].

Corollary 2.1. *For any $\lambda, \varepsilon, t > 0$ we have the following relations*

(i) *If m satisfies the log-Sob. inequality with constant $1/\lambda$ then it satisfies* $\mathrm{ETI}(\lambda, \varepsilon, s, t)$ *for any $s \in [0, t]$.*
(ii) *If m satisfies* $\mathrm{ETI}(\lambda, \varepsilon, 1)$, *then it satisfies* $\mathrm{TI}(1/(2\varepsilon\theta_{\lambda\varepsilon}(1) - \varepsilon))$.
(iii) *If the potential U in $m = \exp(-2U)$ is two times continuously differentiable and λ-convex, then m satisfies* $\mathrm{ETI}(2\lambda, \varepsilon, t)$.
(iv) *If m satisfies* $\mathrm{ETI}(\lambda, \varepsilon, s, t)$ *for all $s \in [0, t]$, then it satisfies the Poincaré inequality,* $\mathrm{PI}(\lambda)$

$$\forall g \in \mathcal{D}, \quad \int_{\mathbb{R}^d} g^2(x) m(\mathrm{d}x) - \left(\int_{\mathbb{R}^d} g(x) m(\mathrm{d}x) \right)^2 \leq \frac{2}{\lambda} \int_{\mathbb{R}^d} |\nabla g|^2(x) m(\mathrm{d}x),$$

where $\mathcal{D}$ is the domain of the generator $\mathcal{L} = \frac{1}{2}\Delta - \nabla U \cdot \nabla$.

Proof. The statement (i) is a natural consequence the equivalence between reverse hypercontractivity and the log-Sob inequality [1], Theorem 3.3. Statement (ii) follows by Remark 1.2, while statement (iii) is a direct consequence of statement (i). Statement (iv) is obtained following the proof of Gross' Theorem; for this reason, we do not provide full detail. We consider a bounded positive function f, bounded away from 0 and with bounded derivatives of order two. The relation (2.9) with the choices $p = 0$ and $q(s) = 1 - \exp(\lambda\varepsilon s)$ implies that $\frac{\mathrm{d}}{\mathrm{d}s}\Lambda(s)|_{s=0} \geq 0$, where $\Lambda(s) = \|f\|_{q(s)}$. One obtains that

$$\left(\|f\|_0\right)^{-1} \frac{\mathrm{d}}{\mathrm{d}s}\Lambda(s)\bigg|_{s=0} = -\frac{\lambda\varepsilon}{2}\left(\int_{\mathbb{R}^d} (\log f)^2 \,\mathrm{d}m - \left(\int_{\mathbb{R}^d} \log f \,\mathrm{d}m \right)^2 \right) + \int_{\mathbb{R}^d} \frac{1}{f} \mathcal{L}^\varepsilon f \,\mathrm{d}m,$$

where $\mathcal{L}^\varepsilon$ is the generator $\frac{\varepsilon}{2}\Delta - \varepsilon\nabla U \cdot \nabla$. The desired conclusion is obtained by mean of some simple algebraic manipulations, the basic rules of Γ-calculus and upon setting $g = \log f$.

Note that Statement (iv) can be seen also as a direct byproduct of the relation $\mathrm{TI}(\lambda) \Rightarrow \mathrm{PI}(\lambda)$ proved in [27]. □

Remark 2.4. Adopting the notation of Bakry's notes [1] and using Theorem 3.3 therein, we get that the relation (2.9) with the choices $p = 0$ and $q(s) = 1 - \exp(\lambda\varepsilon s)$ implies the inequality LogS(0) with constant 0. However, as it can be seen from its Definition at the bottom of page 37,

such inequality is a trivial one and therefore we cannot conclude that it implies the classical Log Sobolev inequality LogS(2). The degeneracy is due to the fact that $q(0) = 0$ and is forced by the restrictions imposed in Theorem 2.2 on the parameters p, q.

Combining statements (i) and (ii) and taking the limit $\varepsilon \to 0$ we obtain the classical result of Otto and Villani [27]

$$\text{log-Sobolev ineq.} \Rightarrow \mathrm{TI}(\lambda)\ .$$

Remark 2.5. In [27], the authors also introduce a stronger inequality which implies both the log-Sobolev and the Talagrand inequality, leading the quadratic Wasserstein distance, the Entropy and the Fisher information together:

$$\mathcal{H}(\mu|m) \le W_2(\mu, m)\sqrt{I(\mu|m)} - \frac{\lambda}{2} W_2^2(\mu, m).$$

It is interesting to point out that we can derive the entropic counterpart of this result, by differentiating in $s = 0$ the convexity estimate for the entropy along Schrödinger bridges (see [10], Theorem 1.4.).

Let us mention that an alternative proof of the classical HWI inequality is given in [15] via the Schrödinger problem. In particular, it is based on the Otto-Villani heuristics applied to Schrödinger bridges.

The next result is a generalization to the entropic transportation inequality of [19], Theorem 2.1, in which it is introduced an inf-convolution log-Sobolev inequality that is implied by a transportation inequality with a general cost.

Corollary 2.2 (ETI$(\lambda, \varepsilon, t)$ **and inf-convolution log-Sobolev inequality).** *For* $\varepsilon, \lambda > 0$ *and* t *such that*,

$$1 + \frac{\varepsilon}{\exp(\lambda\varepsilon t) - (1+\varepsilon)} \ge 0,$$

ETI$(\lambda, \varepsilon, t)$ *implies the following inf-convolution log-Sobolev inequality. For any* $f : \mathbb{R}^d \to \mathbb{R}$,

$$\mathrm{Ent}_m\left(e^f\right) \le \left(1 + \frac{\varepsilon}{\exp(\lambda\varepsilon t) - (1+\varepsilon)}\right) \int \left(f - Q_t^\varepsilon f\right) e^f \,\mathrm{d}m,$$

where we used the standard notation $\mathrm{Ent}_m(f) = \int f \log f \,\mathrm{d}m - \int f \,\mathrm{d}m \log \int f \,\mathrm{d}m$.

Proof. We start by following the proof of [19], Theorem 2.1. We fix $f \in C_b(\mathbb{R}^d)$ and define $\mathrm{d}\nu_f = \frac{e^f}{\int e^f \,\mathrm{d}m} \,\mathrm{d}m$, hence we have

$$\begin{aligned}\mathcal{H}(\nu_f|m) &= \int \log\left(\frac{e^f}{\int e^f \,\mathrm{d}m}\right) \frac{e^f}{\int e^f \,\mathrm{d}m} \,\mathrm{d}m \\ &= \int f \,\mathrm{d}\nu_f - \log \int e^f \,\mathrm{d}m\end{aligned}$$

$$\leq \int f \, \mathrm{d}\nu_f - \int f \, \mathrm{d}m = \int \big(f - Q_t^\varepsilon f\big) \, \mathrm{d}\nu_f + \int Q_t^\varepsilon f \, \mathrm{d}\nu_f - \int f \, \mathrm{d}m$$
$$\leq \int \big(f - Q_t^\varepsilon f\big) \, \mathrm{d}\nu_f + \varepsilon \mathcal{T}_{R_{0t}^\varepsilon}(\nu_f, m) - \varepsilon \mathcal{H}(\nu_f | m),$$

where the first inequality is given by Jensen's inequality, while the last inequality is due to the Kantorovich dual formulation for the entropic transportation cost (1.3). Now ETI$(\lambda, \varepsilon, t)$ implies,

$$\mathcal{H}(\nu_f | m) \leq \int \big(f - Q_t^\varepsilon f\big) \, \mathrm{d}\nu_f + \varepsilon\big(\theta_{\lambda\varepsilon}(t) - 1\big)\mathcal{H}(\nu_f | m).$$

Hence,

$$\mathcal{H}(\nu_f | m)\big(1 + \varepsilon - \varepsilon\theta_{\lambda\varepsilon}(t)\big) \leq \int \big(f - Q_t^\varepsilon f\big) \, \mathrm{d}\nu_f,$$

that is

$$\mathcal{H}(\nu_f | m)\big(1 - \varepsilon\big(\theta_{\lambda\varepsilon}(t) - 1\big)\big) \leq \int \big(f - Q_t^\varepsilon f\big) \, \mathrm{d}\nu_f.$$

To conclude, we remark that $\mathcal{H}(\nu_f | m) = \mathrm{Ent}_m(e^f) \int e^f \, \mathrm{d}m$, thus we obtain the announced inequality. □

Appendix

We briefly collect here some known and fundamental result that we made use of in the previous sections.

Lemma A.1 (Dual representation of the entropy). *Let $\mathcal{Z}$ be a measurable space and $p \in \mathcal{P}(\mathcal{Z})$. For any measurable function $\psi : \mathcal{Z} \to [-\infty, \infty)$ it holds*

$$\log \int \exp(\psi) \, \mathrm{d}p = \sup\left\{ \int \psi \, \mathrm{d}q - \mathcal{H}(q|p); q \in \mathcal{P}(\mathcal{Z}) : \int_{\mathcal{Z}} f_+ \, \mathrm{d}q < \infty \right\} \in [-\infty, \infty]. \quad \text{(A.1)}$$

The proof can be found for instance in [18,23].

Lemma A.2 (Additive property of the relative entropy). *Let Ω, Z two Polish spaces. For any $p, r \in \mathcal{P}(\Omega)$ and any measurable function $\phi : \Omega \to Z$,*

$$\mathcal{H}(p|r) = \mathcal{H}(p_\phi | r_\phi) + \int \mathcal{H}\big(p(\cdot|\phi = z) | r(\cdot|\phi = z)\big) p_\phi(dz), \quad \text{(A.2)}$$

where $p_\phi = \phi_\# p$.

For the proof see [23], Theorem 2.4.

Theorem A.5 (Dual formulation of the entropic transportation cost).

$$\varepsilon \mathcal{T}_{R^\varepsilon_{01}}(\mu,\nu) = \varepsilon \mathcal{H}(\mu|m) + \sup_{\varphi \in C_b(\mathbb{R}^d)} \left\{ \int Q^\varepsilon_1 \varphi \, \mathrm{d}\mu - \int \varphi \, \mathrm{d}\nu \right\},$$

where for all $t \geq 0$,

$$Q^\varepsilon_t \varphi(x) = -\varepsilon \log P^\varepsilon_t \exp(-\varphi/\varepsilon)(x), \quad x \in \mathbb{R}^d.$$

Several proofs are available [9,13,14,16,25].

Acknowledgements

This research was partially supported by the LABEX MILYON ANR-10-LABX-0070. Giovanni Conforti acknowledges financial support from the CMAP, École Polytechnique.

References

[1] Bakry, D. (1994). L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on Probability Theory* (*Saint-Flour*, 1992). *Lecture Notes in Math*. **1581** 1–114. Berlin: Springer. MR1307413 https://doi.org/10.1007/BFb0073872

[2] Bakry, D. and Émery, M. (1985). Diffusions hypercontractives. In *Séminaire de Probabilités*, *XIX*, 1983/84. *Lecture Notes in Math*. **1123** 177–206. Berlin: Springer. MR0889476 https://doi.org/10.1007/BFb0075847

[3] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **348**. Cham: Springer. MR3155209 https://doi.org/10.1007/978-3-319-00227-9

[4] Benamou, J.-D. and Brenier, Y. (2000). A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math*. **84** 375–393. MR1738163 https://doi.org/10.1007/s002110050002

[5] Benamou, J.-D., Carlier, G., Cuturi, M., Nenna, L. and Peyré, G. (2015). Iterative Bregman projections for regularized transportation problems. *SIAM J. Sci. Comput*. **37** A1111–A1138. MR3340204 https://doi.org/10.1137/141000439

[6] Bobkov, S.G., Gentil, I. and Ledoux, M. (2001). Hypercontractivity of Hamilton–Jacobi equations. *J. Math. Pures Appl*. (9) **80** 669–696. MR1846020 https://doi.org/10.1016/S0021-7824(01)01208-9

[7] Bobkov, S.G. and Götze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal*. **163** 1–28. MR1682772 https://doi.org/10.1006/jfan.1998.3326

[8] Bowles, M. and Ghoussoub, N. (2018). A theory of transfers: Duality and convolution. Preprint. Available at arXiv:1804.08563.

[9] Chen, Y., Georgiou, T.T. and Pavon, M. (2016). On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. *J. Optim. Theory Appl*. **169** 671–691. MR3489825 https://doi.org/10.1007/s10957-015-0803-z

[10] Conforti, G. (2019). A second order equation for Schrödinger bridges with applications to the hot gas experiment and entropic transportation cost. *Probab. Theory Related Fields* **174** 1–47. MR3947319 https://doi.org/10.1007/s00440-018-0856-7

[11] Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems* 2292–2300.

[12] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications*. *Stochastic Modelling and Applied Probability* **38**. Berlin: Springer. Corrected reprint of the second (1998) edition. MR2571413 https://doi.org/10.1007/978-3-642-03311-7

[13] Gentil, I., Léonard, C. and Ripani, L. (2017). About the analogy between optimal transport and minimal entropy. *Ann. Fac. Sci. Toulouse Math.* (6) **26** 569–601. MR3669966 https://doi.org/10.5802/afst.1546

[14] Gentil, I., Léonard, C. and Ripani, L. (2019). Dynamical aspects of generalized Schrödinger problem via Otto calculus – A heuristic point of view. *Rev. Mat. Iberoam.* To appear. Preprint. Avaialble at arXiv:1806.01553.

[15] Gentil, I., Léonard, C., Ripani, L. and Tamanini, L. (2019). An entropic interpolation proof of the HWI inequality. *Stochastic Process. Appl.* To appear. Preprint. Avaialble at arXiv:1807.06893. https://doi.org/10.1016/j.spa.2019.04.002

[16] Gigli, N. and Tamanini, L. (2019). Benamou–Brenier and duality formulas for the entropic cost on $RCD^*(K,N)$ spaces. *Probab. Theory Related Fields*. To appear. Available at arXiv:1805.06325. https://doi.org/10.1007/s00440-019-00909-1

[17] Gozlan, N. (2009). A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.* **37** 2480–2498. MR2573565 https://doi.org/10.1214/09-AOP470

[18] Gozlan, N. and Léonard, C. (2010). Transport inequalities. A survey. *Markov Process. Related Fields* **16** 635–736. MR2895086

[19] Gozlan, N., Roberto, C. and Samson, P.-M. (2011). A new characterization of Talagrand's transport-entropy inequalities and applications. *Ann. Probab.* **39** 857–880. MR2789577 https://doi.org/10.1214/10-AOP570

[20] Gozlan, N., Roberto, C., Samson, P.-M. and Tetali, P. (2017). Kantorovich duality for general transport costs and applications. *J. Funct. Anal.* **273** 3327–3405. MR3706606 https://doi.org/10.1016/j.jfa.2017.08.015

[21] Gross, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083. MR0420249 https://doi.org/10.2307/2373688

[22] Léonard, C. (2012). From the Schrödinger problem to the Monge–Kantorovich problem. *J. Funct. Anal.* **262** 1879–1920. MR2873864 https://doi.org/10.1016/j.jfa.2011.11.026

[23] Léonard, C. (2014). Some properties of path measures. In *Séminaire de Probabilités XLVI. Lecture Notes in Math.* **2123** 207–230. Cham: Springer. MR3330819 https://doi.org/10.1007/978-3-319-11970-0_8

[24] Mikami, T. (2004). Monge's problem with a quadratic cost by the zero-noise limit of h-path processes. *Probab. Theory Related Fields* **129** 245–260. MR2063377 https://doi.org/10.1007/s00440-004-0340-4

[25] Mikami, T. and Thieullen, M. (2006). Duality theorem for the stochastic optimal control problem. *Stochastic Process. Appl.* **116** 1815–1835. MR2307060 https://doi.org/10.1016/j.spa.2006.04.014

[26] Nelson, E. (1967). *Dynamical Theories of Brownian Motion*. Princeton, NJ: Princeton Univ. Press. MR0214150

[27] Otto, F. and Villani, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400. MR1760620 https://doi.org/10.1006/jfan.1999.3557

[28] Royer, G. (1999). *Une Initiation aux Inégalités de Sobolev Logarithmiques*. *Cours Spécialisés* [*Specialized Courses*] **5**. Paris: Société Mathématique de France. MR1704288

[29] Schrödinger, E. (1931). Über die Umkehrung der Naturgesetze. *Sitzungsberichte Preuss. Akad. Wiss. Berlin. Phys. Math.* **144** 144–153. https://doi.org/10.1063/1.527002

[30] Schrödinger, E. (1932). Sur la théorie relativiste de l'électron et l'interprétation de la mécanique quantique. *Ann. Inst. Henri Poincaré* **2** 269–310. MR1508000

[31] Talagrand, M. (1996). Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.* **6** 587–600. MR1392331 https://doi.org/10.1007/BF02249265

Received September 2018 and revised September 2019