ENTROPIC CURVATURE ON GRAPHS ALONG SCHRÖDINGER BRIDGES AT ZERO TEMPERATURE

Paul-Marie Samson

ENTROPIC CURVATURE ON GRAPHS ALONG SCHRÖDINGER BRIDGES AT ZERO TEMPERATURE.

PAUL-MARIE SAMSON

ABSTRACT. Lott-Sturm-Villani theory of curvature on geodesic spaces has been extended to discrete graph spaces by C. Léonard by replacing W_2-Wasserstein geodesics by Schrödinger bridges in the definition of entropic curvature [23, 25, 24]. As a remarkable fact, as a temperature parameter goes to zero, these Schrödinger bridges are supported by geodesics of the space. We analyse this property on discrete graphs to reach entropic curvature on discrete spaces. Our approach provides lower bounds for the entropic curvature for several examples of graph spaces: the lattice $\mathbb{Z}^n$ endowed with the counting measure, the discrete cube endowed with product probability measures, the circle, the complete graph, the Bernoulli-Laplace model. Our general results also apply to a large class of graphs which are not specifically studied in this paper.

As opposed to Erbar-Maas results on graphs [27, 10, 11], entropic curvature results of this paper imply new Prékopa-Leindler type of inequalities on discrete spaces, and new transport-entropy inequalities related to refined concentration properties for the graphs mentioned above. For example on the discrete hypercube $\{0,1\}^n$ and for the Bernoulli Laplace model, a new $W_2 - W_1$ transport-entropy inequality is reached, that can not be derived by usual induction arguments over the dimension n. As a surprising fact, our method also gives improvements of weak transport-entropy inequalities (see [28, 15]) associated to the so-called convex-hull method by Talagrand [38].

The paper starts with a brief overview about known results concerning entropic curvature on discrete graphs. Then we introduce a specific entropic curvature property on graphs derived from C. Léonard approach [23, 25, 24], and dealing with Schrödinger bridges at zero temperature (see Definition 1.1).

The main curvature results are given in section 2, with their connections to new transport-entropy inequalities. The concentration properties following from such transport-entropy inequalities are not developed in the present paper. For that purpose, we refer to [15] by Gozlan & al, where the link between transport-entropy inequalities and concentration properties are widely investigated.

The strategy of proof, presented in section 3, uses the so called *slowing-down procedure* for Schrödinger bridges associated to jump processes on discrete spaces pushed forward by C. Léonard. The key proposition of the present paper, Proposition 3.5 (with Lemma 3.1), is derived from this procedure, which consists of decreasing a temperature parameter γ to 0 in order to construct W_1-Wasserstein geodesics on the set of probability measures on the graph. All the curvature results of this paper are derived from Proposition 3.5. Our strategy also applies for many other graph spaces which are not considered in this paper. The main goal of this work is to push forward Leonard's slowing-down procedure to reach entropic curvature on graphs through few significant new results.

Date: July 24, 2020.

1991 *Mathematics Subject Classification.* 60E15, 32F32 and 39A12.

Key words and phrases. Displacement convexity property, Ricci curvature, graphs, Bernoulli Laplace model, discrete hypercube, transposition model, Schrödinger bridges, transport-entropy inequalities, concentration of measure, Prékopa-Leindler inequalities.

This research is partly funded by the Bézout Labex, funded by ANR, reference ANR-10-LABX-58. The author is supported by a grant of the Simone and Cino Del Duca Foundation.

Contents

1. Introduction : Schrödinger bridges for entropic curvature

For any measurable space $\mathcal{Y}$, we note $\mathcal{M}(\mathcal{Y})$ the set of all non-negative σ-measures on $\mathcal{Y}$ and $\mathcal{P}(\mathcal{Y})$ the set of all probability measures on $\mathcal{Y}$.

Let $(\mathcal{X}, d)$ be a geodesic space equipped with a reference measure $m \in \mathcal{M}(\mathcal{X})$. According to Lott-Strum-Villani theory of curvature on geodesic spaces [26, 36, 37, 40], a lower bound $K \in \mathbb{R}$ on the entropic curvature of the space $(\mathcal{X}, d, m)$ is characterized by a *K-convexity property* of the *relative entropy* along *constant speed geodesics* of the *Wasserstein space* $(\mathcal{P}_2(\mathcal{X}), W_2)$. Let us precise this property for the non specialist reader. By definition, the *relative entropy* of a probability measure q on a measurable space $\mathcal{Y}$ with respect to a measure $r \in \mathcal{M}(\mathcal{Y})$ is given by

$$H(q|r) := \int_{\mathcal{Y}} \log(dq/dr)\, dq \qquad \in (-\infty, \infty],$$

if q is absolutely continuous with respect to r and $H(q|r) := +\infty$ otherwise. We refer to [22] for more details about this definition. The space $\mathcal{P}_2(\mathcal{X})$ is the set of probability measures with second moment and W_2 is the *Wasserstein distance of order 2* on $\mathcal{X}$: namely, for any $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$,

$$W_2(\nu_0, \nu_1) := \left(\inf_{\pi \in \Pi(\nu_0, \nu_1)} \iint d(x, y)^2 d\pi(x, y) \right)^{1/2}, \tag{1}$$

where $\Pi(\nu_0, \nu_1)$ is the set of all probability measures on the product space $\mathcal{X} \times \mathcal{X}$ with first marginal ν_0 and second marginal ν_1 (also called *transference plans* from ν_0 to ν_1). A path $(\nu_t)_{t\in[0,1]}$ in $\mathcal{P}_2(\mathcal{X})$ is a *constant speed W_2-geodesic* from ν_0 to ν_1 if for all $0 \le s < t \le 1$, $W_2(\nu_s, \nu_t) = (t - s)W_2(\nu_0, \nu_1)$. The *$K$-convexity property* of the relative entropy $H(\cdot|m)$ is expressed as follows: for any $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$ whose supports are included in the support of m, there exists a constant speed W_2-geodesic $(\nu_t)_{t\in[0,1]}$ from ν_0 to ν_1 such that for all $t \in [0, 1]$,

$$H(\nu_t|m) \le (1 - t)\, H(\nu_0|m) + t\, H(\nu_1|m) - \frac{K}{2}\, t(1 - t)\, W_2^2(\nu_0, \nu_1). \tag{2}$$

If such a property holds, one says that the Lott-Sturm-Villani entropic curvature of the space $(\mathcal{X}, d, m)$ is bounded from below by K.

Property (2) with $K = 0$ has been discovered by McCann on the Euclidean space $(\mathcal{X}, d) = (\mathbb{R}^d, |\cdot|_2)$ endowed with the Lebesgue measure [29]. More generally, as a remarkable fact, when $\mathcal{X}$ is a Riemannian manifold equipped with its geodesic distance d and a measure m with density e^{-V} with respect to the volume measure, property (2) is equivalent to the so-called Bakry-Emery curvature condition $CD(\infty, K)$: Ricc + Hess(V) $\geq$ K (see e.g. [3]). As a consequence, due to the wide range of implications of this notion of curvature, property (2) has been used as a guideline by Lott-Sturm-Villani to define curvature on geodesic spaces (see also [1, 2]) and then by different authors to propose entropic definitions of curvature on discrete spaces : Bonciocat-Sturm [6], Ollivier-Villani on the discrete cube [34], Erbar-Maas [27, 10, 11], Mielke [30], Léonard [23, 25, 24], Hillion [17, 18] and Gozlan-Roberto-Samson-Tetali [14].

This paper concerns Léonard entropic approach of curvature in discrete setting, from which we also recover results from [14] and [17]. In discrete spaces, several other notions of curvature have already been studied which are not considered in this paper : the caorse Ricci curvature [32, 33], the Bochner-Bakry-Emery approach with the (Bochner) curvature [7, 19] and the curvature dimension or exponential curvature dimension inequality [4].

As m is the unique invariante probability measure of a Markov kernel on a discrete space $\mathcal{X}$, a first global entropic approach has been proposed by M. Erbar and J. Maas [27, 10, 11]. The core of their approach is the construction of an abstract Wasserstein distance $\mathcal{W}_2$ on $\mathcal{P}(\mathcal{X})$, that replaces the Wasserstein distance W_2 in (2). This distance $\mathcal{W}_2$ is defined using a discrete analogue of the Benamou-Brenier formula for W_2, in order to provide a Riemannian structure for the probability space $\mathcal{P}(\mathcal{X})$. Unfortunately, there is no static definition of $\mathcal{W}_2^2$ as a minimum of a cost among transference plans π as in the definition (1) of W_2^2. Erbar-Maas entropic Ricci curvature definition satisfies a tensorisation property for product of graphs that allows to consider high dimensional spaces [10]. This definition has been used to get lower bounds on curvature for several models of graphs : the discrete circle, the complete graph, the discrete hypercube [27, 10], the Bernoulli-Laplace model, the random transposition model [12, 13], birth and death processes, zero-range processes [13], Cayley graphs of non-abelian groups, weakly interacting Markov chains such as the Ising model [9]. The main strategy of all this papers is to prove an equivalent criterion of Erbar-Maas entropic curvature given in [10], by identifying some discrete analogue of the Bochner identity in continuous setting.

Finding a minimizer in the definition of $W_2(\nu_0, \nu_1)$ is known as the quadratic Monge-Kantorovich problem. By the so-called *slowing down procedure*, T. Mikami [31] and then C. Léonard [21, 23, 24, 25] show that the quadratic Monge-Kantorovich problem in continuous, but also the W_1-Monge-Kantorovich problem in discrete, can be understood as the limit of a sequence of entropy minimization problems, the so-called *Schrödinger problems*.

In this paper, the *slowing down procedure*, described further, is used to prove entropic curvature properties of type (2) as $\mathcal{X}$ is a graph, endowed with its natural graph distance $d = d_\sim$, and with a measure m, reversible with respect to some generator L. More precisely, in property (2), constant speed W_2-geodesics $(\nu_t)_{t\in[0,1]}$ are replaced by constant speed W_1-geodesics where W_1 is *the Wasserstein distance of order 1* given by

$$W_1(\nu_0, \nu_1) := \inf_{\pi \in \Pi(\nu_0,\nu_1)} \iint d(x, y)\, d\pi(x, y), \qquad \nu_0, \nu_1 \in \mathcal{P}(\mathcal{X}).$$

As explained below, each of these constant speed W_1-geodesics, denoted by $(\widehat{Q}_t^0)_{t\in[0,1]}$ throughout this paper, is the limit path of a sequence of Schrödinger briges $(\widehat{Q}_t^\gamma)_{t\in[0,1]}$ indexed by a temperature parameter $\gamma > 0$, as γ goes to zero. We call it *Schrödinger brige at zero temperature*. In property (2), the curvature term $W_2^2(\nu_0, \nu_1)$ is also replaced by some transport cost $C_t(\nu_0, \nu_1)$ that may also depend on the parameter $t \in (0,1)$. Let $\mathcal{P}_b(\mathcal{X})$ denotes the set of probability measures on $\mathcal{X}$ with finite support. The analogue of property (2) on discrete graphs at the focus of this work is the following.

Definition 1.1. *On the discrete space $(\mathcal{X}, d, m, L)$, one says that the relative entropy is C-displacement convex where $C = (C_t)_{t\in[0,1]}$, if for any probability measure $\nu_0, \nu_1 \in \mathcal{P}_b(\mathcal{X})$, the Schrödinger bridge at zero temperature $(\widehat{Q}^0_t)_{t\in[0,1]}$ from ν_0 to ν_1, satisfies for any $t \in [0,1]$,*

$$H(\widehat{Q}^0_t|m) \leq (1-t)H(\nu_0|m) + t\,H(\nu_1|m) - \frac{t(1-t)}{2}C_t(\nu_0,\nu_1). \tag{3}$$

For some of the graphs studied in this paper, the cost $C_t(\nu_0,\nu_1)$ is bigger than $K\,W_1(\nu_0,\nu_1)^2$ for any $t \in [0,1]$ with $K \geq 0$. In that case one may say that *the W_1-entropic curvature* of the space $(\mathcal{X}, d, m, L)$ is bounded from below by K. Such a property is also a consequence of Erbar-Maas entropic curvature since $\mathcal{W}_2 \geq W_1$ but their property deals with different constant speed geodesics on $\mathcal{P}(\mathcal{X})$. Let us introduce another discrete analogue of the W_2-distance:

$$W_2^d(\nu_0,\nu_1) := \left(\inf_{\pi\in\Pi(\nu_0,\nu_1)} \iint d(x,y)(d(x,y)-1)\,d\pi(x,y)\right)^{1/2}, \quad \nu_0,\nu_1 \in \mathcal{P}_2(\mathcal{X}). \tag{4}$$

For some graphs in this paper, we also get

$$C_t(\nu_0,\nu_1) \geq K'\,\left(W_2(\nu_0,\nu_1)^2 - W_1(\nu_0,\nu_1)\right) \geq K' W_2^d(\nu_0,\nu_1)^2,$$

with $K' \geq 0$. In that case, one may say that *the W_2^d-entropic curvature*, of the space $(\mathcal{X}, d, m, L)$ is bounded from below by K'.

In the definition (4) of W_2^d, the cost $d(x,y)(d(x,y)-1)$ is zero if x and y are neighbours. Therefore the optimal transport-cost W_2^d does not well measure the distance between probabilities with close supports. Observe that such type of costs also appear in the paper by Bonciocat-Sturm [6] in their definition of rough (approximate) lower curvature.

In this paper, a C-displacement convexity property is proved for the following discrete spaces : the lattice $\mathbb{Z}^n$ endowed with the counting measure (see Theorem 2.2), the discrete hypercube endowed with product probability measures (see Theorem 2.4), the discrete circle endowed with uniform measure (see Theorem 2.5), the complete graph (see Theorem 2.3), the Bernoulli-Laplace model (see Theorem 2.6). For all these graphs, one gets a non-negative lower bound for their W_1 or W_2^d-entropic curvature. In a forthcoming paper one will also consider some cases of spaces with "negative" entropic curvature.

For more comprehension, let us briefly explain the *slowing down procedure* in its original continuous setting before considering discrete spaces. Let R^γ be the law of a reversible Brownian motion with diffusion coefficient $\gamma > 0$ on the set Ω of continuous paths from $[0,1]$ to $\mathcal{X} = \mathbb{R}^d$. The coefficient γ can be also interpreted as a temperature parameter. The measure $R^\gamma \in \mathcal{M}(\Omega)$ is a Markov measure with infinitesimal operator $L^\gamma = \gamma\Delta$ (where Δ denotes the Laplacian), and initial reversible measure $dm = dx$, the Lebesgue measure on $\mathbb{R}^d$.

In all the paper, we use the following notations. For any $t \in [0,1]$, X_t is *the projection map*

$$X_t : \omega \in \Omega \mapsto \omega_t \in \mathcal{X}.$$

Given $Q \in \mathcal{M}(\Omega)$, the measure $Q_t := X_t\#Q$ on $\mathcal{X}$ denotes the push-forward of the measure Q by X_t, and for any $0 \leq t < s \leq 1$, the measure $Q_{s,t} := (X_s, X_t)\#Q$ on $\mathcal{X}\times\mathcal{X}$ denotes the push forward of the measure Q by the projection map (X_s, X_t). For any integrable function $F : \Omega \to \mathbb{R}$ with respect to Q, one notes

$$\mathbb{E}_Q[F] := \int_\Omega F\,dQ.$$

The informal result by T. Mikami [31] or C. Léonard [21] is the following: for any absolutely continuous measures $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$, for any sequences $(\gamma_k)_{k\in\mathbb{N}}$ of temperature parameters going to zero,

$$\begin{aligned} W_2^2(\nu_0,\nu_1) &= \inf_{Q\in\mathcal{P}(\Omega)} \left\{ \mathbb{E}_Q[c] \,\middle|\, Q_0 = \nu_0, Q_1 = \nu_1 \right\} \\ &= \lim_{\gamma_k\to 0} \left[\gamma_k \min_{Q\in\mathcal{P}(\Omega)} \left\{ H(Q|R^{\gamma_k}) \,\middle|\, Q_0 = \nu_0, Q_1 = \nu_1 \right\} \right], \end{aligned}$$

where $c(\omega) := \int_0^1 |\, \dot{\omega}_t \,|^2 dt$, if the path $\omega = (\omega_t)_{t\in[0,1]}$ is absolutely continuous ($\dot{\omega}$ denotes its time derivative), and $c(\omega) := +\infty$ otherwise. The first equality is known as the Benamou-Brenier formula (see [5]). The second equality therefore relates W_2 to the so-called *dynamic Schrödinger minimization problems*. As a convex minimization problem, for any fixed $\gamma > 0$, it admits a single minimizer $\widehat{Q}^\gamma$, namely

$$\min_{Q\in\mathcal{P}(\Omega)} \left\{ H(Q|R^\gamma) \,\middle|\, Q_0 = \nu_0, Q_1 = \nu_1 \right\} = H(\widehat{Q}^\gamma|R^\gamma). \tag{5}$$

As interpretation, the measure $\widehat{Q}^\gamma$ is the law of the process with configuration $\widehat{Q}^\gamma_0 = \nu_0$ at time $t = 0$ and $\widehat{Q}^\gamma_1 = \nu_1$ at time $t = 1$, which is the closest one for the entropic distance, to a reversible Brownian motion with diffusion coefficient γ. As a result (see [31, 21]), the sequence of minimizers $(\widehat{Q}^{\gamma_k})_{k\in\mathbb{N}}$ converges to a single measure $\widehat{Q}^0 \in \mathcal{P}(\Omega)$. For any $t \in [0,1]$, let $\nu_t^\gamma := \widehat{Q}^\gamma_t$ and $\nu_t := \widehat{Q}^0_t$. By definition, $(\nu_t^\gamma)_{t\in[0,1]}$ is *a Schrödinger bridge from ν_0 to ν_1 at fixed temperature* γ, and as a main result, as γ_k goes to zero, the limit path $(\nu_t)_{t\in[0,1]}$, is a W_2-geodesic from ν_0 to ν_1 (see [23]). Therefore, it is natural to consider a relaxation of the curvature definition (2) by replacing the geodesic $(\nu_t)_{t\in[0,1]}$ by the bridge $(\nu_t^\gamma)_{t\in[0,1]}$ and by replacing $W_2^2(\nu_0,\nu_1)$ by $\gamma H(\widehat{Q}^\gamma|R^\gamma)$. This idea has been explored in continuous setting by G. Conforti in [8].

Let us present the discrete analogue of this approach due to C. Léonard [23, 25, 24]. From now on, the space $\mathcal{X}$ is a countable set endowed with the σ-algebra generated by singletons. The set $\Omega \subset \mathcal{X}^{[0,1]}$ denotes the space of all left-limited, right-continuous, piecewise constant paths $\omega = (\omega_t)_{t\in[0,1]}$ on $\mathcal{X}$, with finitely many jumps. The space Ω is endowed with the σ-algebra $\mathcal{F}$ generated by the cylindrical sets.

According to C. Léonard's paper [24], the discrete space $\mathcal{X}$ is equipped with a metric distance d. This distance is assumed to be *positively lower bounded*: for all $x \neq y$ in $\mathcal{X}$, $d(x,y) \geq 1$. The space $\mathcal{X}$ is also the set of vertices of a connected graph $G = (\mathcal{X}, E)$ where $E \subset \mathcal{X}\times\mathcal{X}$ denotes the set of directed edges of the graph. G is supposed to be an undirected graph so that for all $(x,y) \in E$, one has $(y,x) \in E$. Two vertices x and y are *neighbours* and we note $x \sim y$ if $(x,y) \in E$. We assume that any vertex $x \in \mathcal{X}$ has a finite number of neighbours d_x and that $\sup_{x\in\mathcal{X}} d_x = d_{\max} < \infty$. The length $\ell(\omega)$ of a piecewise constant path $\omega = (\omega_t)_{t\in[0,1]} \in \Omega$ is given by

$$\ell(\omega) := \sum_{0<t<1} d(\omega_{t^-}, \omega_t).$$

In C. Léonard's paper, the distance is assumed to be *intrinsic in the discrete sense* (see [24, Hypothesis 2.1]), this means that for any $x, y \in \mathcal{X}$,

$$d(x,y) := \inf \left\{ \ell(\omega) \,\middle|\, \omega \in \Omega, \omega_0 = x, \omega_1 = y \right\}.$$

In this paper, we only consider the simple case where $d = d_\sim$ is the graph distance for which the above assumptions are fulfilled: $d_\sim(x,y) = 1$ if and only if $x \sim y$.

A *discrete path α of length $\ell \in \mathbb{N}$* joining two vertices x and y is a sequence of $\ell + 1$ neighbours $\alpha = (z_0, \ldots, z_\ell)$ so that $z_0 = x$ and $z_\ell = y$. In the sequel, we note $z \in \alpha$ if there exists $i \in \{0, \ldots, \ell\}$ such that $z = z_i$, and we note $(z, z') \in \alpha$ if there exists $0 \leq i < j \leq \ell$ such that $z = z_i$ and $z' = z_j$. The distance $d(x,y)$ is also the minimal length of a path joining x and y. A *discrete geodesic path* joining x to y is a path of length $d(x,y)$ from x to y. We note $G(x,y)$ the set of all geodesic paths joining x to y, and we note $[x,y]$ the set of all points that belongs to a geodesic from x to y,

$$[x,y] := \left\{ z \in \mathcal{X} \,\middle|\, z \in \alpha, \alpha \in G(x,y) \right\}.$$

At fixed temperature $\gamma > 0$, as reference measure on Ω, we consider a Markov path measure R^γ with generator L^γ defined by

$$L^\gamma(x,y) := \gamma^{d(x,y)} L(x,y), \qquad x, y \in \mathcal{X},$$

and initial *reversible invariante* measure $R_0^\gamma = m$. More precisely, we assume that m is reversible with respect to L, which means that for any $x, y \in \mathcal{X}$

$$m(x)L(x,y) = m(y)L(y,x).$$

It implies that m is reversible with respect to L^γ for any $\gamma > 0$, and therefore $R_t^\gamma = m$ for all $t \in [0,1]$. We also assume that the Markov process is irreducible so that $m(x) > 0$ for all $x \in \mathcal{X}$. Recall that from the definition of a generator, for any $t \geq 0$ and any $x, y \in \mathcal{X}$, one has

$$R_{t,t+h}^\gamma(x,y) = R_t^\gamma(x)(\delta_x(y) + L^\gamma(x,y)h + o(h)),$$

where δ_x is the Dirac measure at point x. We note $P_t, t \geq 0$, the Markov semi-group associated to L, and $P_t^\gamma, t \geq 0$, the Markov semi-group associated to $L^\gamma, \gamma > 0$. By reversibility, one has for any $x, y \in \mathcal{X}$

$$R_{0,t}^\gamma(x,y) = m(x)P_t^\gamma(x,y) = m(y)P_t^\gamma(y,x),$$

and since the process is irreducible, $P_t^\gamma(x,y) > 0$ for all $t > 0$ and all $x, y \in \mathcal{X}$. For any integrable function $f : \mathcal{X} \to \mathbb{R}$ with respect to $P_t^\gamma(x,\cdot)$, we set

$$P_t^\gamma f(x) := \sum_{y \in \mathcal{X}} f(y)\, P_t^\gamma(x,y).$$

In this paper we only consider generator L satisfying :

$$L(x,y) > 0 \quad \text{if and only if} \quad x \sim y, \tag{6}$$

so that $P_t^\gamma = P_{\gamma t}$ for all $\gamma, t > 0$, but also for any $x \neq y$,

$$d(x,y) = \min\left\{ k \in \mathbb{N} \,\middle|\, L^k(x,y) > 0 \right\}.$$

Let $\nu_0, \nu_1 \in \mathcal{P}(\mathcal{X})$ with respective densities h_0 and h_1 according to m. In Léonard's paper [24], Theorem 2.1 ensures that under some assumptions (see [24, Hypothesis 2.1]), at fixed temperature $\gamma > 0$, the minimum value of the dynamic Schrödinger problem (5) is reached for a single probability measure $\widehat{Q}^\gamma$ which is Markov. This Markov property implies that the measure $\widehat{Q}^\gamma$ has density $f^\gamma(X_0)g^\gamma(X_1)$ with respect to R^γ, where f^γ and g^γ are measurable positive functions on $\mathcal{X}$ satisfying the following so-called *Schrödinger system*

$$\begin{cases} f^\gamma(x)\, P_1^\gamma g^\gamma(x) &= h_0(x), \\ g^\gamma(y)\, P_1^\gamma f^\gamma(y) &= h_1(y), \end{cases} \qquad \forall x, y \in \mathcal{X}. \tag{7}$$

Since f^γ is non-negative and $f^\gamma \neq 0$, by irreducibility one has $P_t^\gamma f^\gamma > 0$ for all $t > 0$, and for the same reason, $P_t^\gamma g^\gamma > 0$ for all $t > 0$. As a consequence, if ν_0 and ν_1 have finite support, then the Schrödinger system (7) implies that f^γ and g^γ have also finite support.

According to [25, Theorem 6.1.4.], from the Markov property, the law at time t of the Schrödinger bridge at fixed temperature γ, $\widehat{Q}_t^\gamma$, is given by: for any $z \in \mathcal{X}$,

$$\widehat{Q}_t^\gamma(z) = P_t^\gamma f^\gamma(z) P_{1-t}^\gamma g^\gamma(z) m(z) = \sum_{x,y \in \mathcal{X}} m(z) P_t^\gamma(z,x) P_{1-t}^\gamma(z,y) f^\gamma(x) g^\gamma(y). \tag{8}$$

Let us present another expression for $\widehat{Q}_t^\gamma$. First, by reversibility, one has

$$\sum_{z\in\mathcal{X}} m(z)P_t^\gamma(z,x)P_{1-t}^\gamma(z,y) = m(x)P_1^\gamma(x,y) = R_{0,1}^\gamma(x,y).$$

Therefore, setting

$$\nu_t^{\gamma x,y}(z) := \frac{m(z)P_t^\gamma(z,x)P_{1-t}^\gamma(z,y)}{m(x)P_1^\gamma(x,y)} = \frac{P_t^\gamma(x,z)P_{1-t}^\gamma(z,y)}{P_1^\gamma(x,y)} = \frac{P_{1-t}^\gamma(y,z)P_t^\gamma(z,x)}{P_1^\gamma(y,x)}, \tag{9}$$

and

$$\widehat{\pi}^\gamma(x,y) := \widehat{Q}_{0,1}^\gamma(x.y) = R_{0,1}^\gamma(x,y)f^\gamma(x)g^\gamma(y),$$

we get

$$\widehat{Q}_t^\gamma(z) = \sum_{x,y\in\mathcal{X}} \nu_t^{\gamma x,y}(z)\,\widehat{\pi}^\gamma(x,y), \qquad z\in\mathcal{X}.$$

Actually, for any $x,y\in\mathcal{X}$, $(\nu_t^{\gamma x,y})_{t\in[0,1]}$ is the Schrödinger bridge joining the Dirac measures δ_x and δ_y. The path $(\widehat{Q}_t^\gamma)_{[0,1]}$ is therefore a mixing of these Schrödinger bridges, according to the coupling measure $\widehat{\pi}^\gamma\in\Pi(\nu_0,\nu_1)$.

Using the Schrödinger system (7), the measure $\widehat{\pi}^\gamma$ can be rewritten as follows,

$$\widehat{\pi}^\gamma(x,y) = \nu_0(x)\,\frac{g^\gamma(y)P_1^\gamma(x,y)}{P_1^\gamma g^\gamma(x)} = \nu_1(y)\,\frac{f^\gamma(x)P_1^\gamma(y,x)}{P_1^\gamma f^\gamma(y)}.$$

For any $\nu\in\mathcal{P}(\mathcal{X})$, let supp$(\nu)$ denote the support of the measure ν, supp$(\nu) := \{x\in\mathcal{X}\,|\,\nu(x)>0\}$. The measure $\widehat{\pi}^\gamma$ admits the following decomposition,

$$\widehat{\pi}^\gamma(x,y) = \nu_0(x)\widehat{\pi}_{\rightarrow}^\gamma(y|x) = \nu_1(y)\widehat{\pi}_{\leftarrow}^\gamma(x|y),$$

where $\widehat{\pi}_{\rightarrow}^\gamma$ and $\widehat{\pi}_{\leftarrow}^\gamma$ are the Markov kernel defined by, for any $x\in$ supp(ν_0),

$$\widehat{\pi}_{\rightarrow}^\gamma(y|x) := \frac{g^\gamma(y)P_1^\gamma(x,y)}{P_1^\gamma g(x)},$$

and for any $y\in$ supp(ν_1),

$$\widehat{\pi}_{\leftarrow}^\gamma(x|y) := \frac{f^\gamma(x)P_1^\gamma(y,x)}{P_1^\gamma f^\gamma(y)}. \tag{10}$$

In order to fulfil this presentation, recall that the *static Schrödinger minimization problem* associated to $R_{0,1}^\gamma$ is to find the minimum value of $H(\pi|R_{0,1}^\gamma)$ over all $\pi\in\Pi(\nu_0,\nu_1)$. Theorem 2.1. by C. Léonard [24] ensures that under Hypothesis 2.1 of its paper, this minimum value is the same as the one of the dynamic Schrödinger minimization problem. Moreover it is reached for $\widehat{\pi}^\gamma = \widehat{Q}_{0,1}^\gamma \in \mathcal{P}(\mathcal{X}\times\mathcal{X})$ and therefore

$$\inf_{\pi\in\Pi(\nu_0,\nu_1)} H(\pi|R_{0,1}^\gamma) = H(\widehat{\pi}^\gamma|R_{0,1}^\gamma) = H(\widehat{Q}^\gamma|R^\gamma).$$

As in the continuous case, let us now apply the *slowing down procedure*. As the temperature γ decreases to zero, the jumps of the Markov process are less frequent, and the reference process is therefore a *lazy random walk* according to C. Léonard's terminology. In order to justify the behaviour of the Scrödinger bridge as the temperature goes to zero, for computational reasons, we need the following not so restrictive additional assumptions.

- The measure m is bounded,

$$\sup_{x\in\mathcal{X}} m(x) < \infty, \qquad \text{and} \qquad \inf_{x\in\mathcal{X}} m(x) > 0. \tag{11}$$

- The generator L is uniformly bounded : there exists $S \geq 1$ such that

$$\sup_{x\in\mathcal{X}} |L(x,x)| \leq S, \tag{12}$$

and there exists $I \in (0,1]$ such that

$$\inf_{x,y\in\mathcal{X}, x\sim y} L(x,y) \geq I. \tag{13}$$

- For any $x \in \mathcal{X}$, there exists $\gamma_0 \in (0,1]$ such that

$$\sum_{y\in\mathcal{X}} \gamma_0^{d(x,y)} < \infty. \tag{14}$$

Hypothesis (12) implies that the semi-group $(P_t^\gamma)_{t\geq 0}$ is given by

$$P_t^\gamma := e^{t\gamma L} = \sum_{k\in\mathbb{N}} \frac{(t\gamma)^k}{k!} L^k. \tag{15}$$

Let us now consider the behaviour of the Schrödinger bridges $(\widehat{Q}_t^\gamma)_{t\in[0,1]}$ as γ goes to zero. Assume ν_0 and ν_1 have finite support. As condition (12) holds, Lemma 4.3 (iv) gives the limit of the path $(\nu_t^{\gamma x,y})_{t\in[0,1]}$ defined by (9): namely, for any $z \in \mathcal{X}$,

$$\lim_{\gamma\to 0} \nu_t^{\gamma x,y}(z) = \nu_t^{0 x,y}(z) := \mathbb{1}_{[x,y]}(z)\, r(x,z,z,y)\, \rho_t^{d(x,y)}(d(x,z)), \tag{16}$$

where for any $x, z, v, y \in \mathcal{X}$,

$$r(x,z,v,y) = \frac{L^{d(x,z)}(x,z) L^{d(v,y)}(v,y)}{L^{d(x,y)}(x,y)}, \tag{17}$$

and ρ_t^d denotes the binomial law with parameter $t \in [0,1]$, $d \in \mathbb{N}$:

$$\rho_t^d(k) := \binom{d}{k} t^k (1-t)^{d-k}, \quad k \in \{0,\ldots,k\},$$

with the binomial coefficient $\binom{d}{k} := \frac{d!}{k!(d-k)!}$. This limit Schrödinger bridge $(\nu_t^{0 x,y})_{t\in[0,1]}$ is supported by $[x,y]$, the set of points on discrete geodesics from x to y. Therefore Schrödinger bridges at zero temperature are consistent with the metric graph structure. This is not surprising. Indeed, roughly speaking, $\nu_t^{0 x,y}$ can be interpreted as the law of a process going from x to y which is closest to a lazy random walk (since γ goes to 0). Therefore this process is forced to follow the geodesics of the graph from x to y.

For fixed $x \neq y$, the law $\nu_t^{0 x,y}$ on $[x,y]$ can be described as follows. Let S denote a binomial random variable with parameters $t \in [0,1]$ and $d = d(x,y) \in \mathbb{N}$, and let Γ be a random discrete geodesic in $G(x,y)$ whose law is given by

$$\mathbb{P}(\Gamma = \alpha) = \frac{L(\alpha_0,\alpha_1)\cdots L(\alpha_{d-1},\alpha_d)}{L^{d(x,y)}(x,y)}, \qquad \text{for all } \alpha = (\alpha_0,\alpha_1,\ldots,\alpha_d) \in G(x,y).$$

If S and $\Gamma = (\Gamma_0,\ldots,\Gamma_d)$ are independent then $\nu_t^{0 x,y}$ is the law of Γ_S.

Let us come back to the behaviour of the Schrödinger bridges at low temperature. C. Léonard [24, Theorem 2.1] proves that given a positive sequence $(\gamma_k)_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} \gamma_k = 0$, the sequence of optimal Schrödinger minimizers $(\widehat{Q}^{\gamma_k})_{k\in\mathbb{N}}$ converges to a single probability measure $\widehat{Q}^0 \in \mathcal{P}(\Omega)$ for the narrow convergence, provided Hypothesis 2.1 holds. In this paper, the measure $\widehat{Q}^0$ is named as *the limit Schrödinger problem optimizer at zero temperature, between* ν_0 *and* ν_1. In the framework of this work, choosing two probability measures ν_0 and ν_1 with finite supports, Hypothesis 2.1 in [24] is reduced to the following assumption (see condition (μ) in Hypothesis 2.1): for any $x, y \in \mathcal{X}$ and for any $\gamma > 0$

$$\mathbb{E}_{R^\gamma}[\ell \,|\, X_0 = x, X_1 = y] < \infty.$$

According to Lemma 4.3 (6), this assumption is fulfilled thanks to (12) since $P_1^\gamma(x,y) > 0$ for any $x, y \in \mathcal{X}$ and $\gamma > 0$.

As a main result of [24, Theorem 2.1], the measure $\widehat{Q}^0$ is also a solution of the following *dynamic Monge-Kantorovich problem* :

$$\inf\left\{\mathbb{E}_Q[\ell]\,\middle|\,Q \in \mathcal{P}(\Omega), Q_0 = \mu_0, Q_1 = \mu_1\right\} = \mathbb{E}_{\widehat{Q}^0}[\ell].$$

The sequence of coupling measures $(\widehat{\pi}^{\gamma_k})_{k\in\mathbb{N}}$ also weakly converges to

$$\widehat{\pi}^0 := \widehat{Q}^0_{0,1}.$$

Moreover, similarly to the continuous case, $\widehat{\pi}^0$ is a W_1-*optimal coupling* of ν_0 and ν_1, it means a minimizer of $W_1(\nu_0, \nu_1)$,

$$W_1(\nu_0,\nu_1) = \iint d(x,y)\,d\widehat{\pi}^0(x,y) = \mathbb{E}_{\widehat{Q}^0}(\ell).$$

The weak convergence of $(\widehat{Q}^{\gamma_k})_{k\in\mathbb{N}}$ to $\widehat{Q}^0$ also provides the convergence of $(\widehat{Q}^{\gamma_k}_t)_{k\in\mathbb{N}}$ to $\widehat{Q}^0_t$, and one has

$$\widehat{Q}^0_t(z) = \iint \nu_t^{0,x,y}(z)\,d\widehat{\pi}^0(x,y). \tag{18}$$

The path $(\widehat{Q}^0_t)_{t\in[0,1]}$ is joining ν_0 to ν_1. According to its construction, this bridge is called *Schrödinger bridge at zero temperature* from ν_0 to ν_1. As a main result, C. Leonard proves that with hypothesis (6), the path $(\widehat{Q}^0_t)_{t\in[0,1]}$ is a constant speed W_1-geodesic (see [24, Theorem 3.15]): for any $0 \le s \le t \le 1$,

$$W_1(\widehat{Q}^0_t, \widehat{Q}^0_s) = (t-s)W_1(\nu_0,\nu_1).$$

2. Main results : examples of entropic curvature bounds along Schrödinger bridges on graphs

The main purpose of this section is to present W_1 or W_2^d-entropic curvature bounds for several discrete graph spaces $(\mathcal{X}, d, m, L)$, in the framework of the first section. As explained before, these bounds follows from C-displacement convexity properties (3) of the relative entropy along *Schrödinger bridges at zero temperature* $(\widehat{Q}^0_t)_{t\in[0,1]}$, derived from the slowing down procedure.

As in the paper [14], C-displacement convexity properties imply a wide range of functional inequalities for the measure m on $\mathcal{X}$, such as Prékopa-Leindler type of inequalities, transport-entropy inequalities, and also discrete Poincaré or modified log-Sobolev inequalities.

To avoid lengths, discrete Poincaré and modified log-Sobolev inequalities are not considered in the present paper. However, we push forward new transport-entropy inequalities to emphasize the efficiency of the Schrödinger approach. Indeed, optimal transport costs derived from this method are well suited to get new concentration properties, using known connections between transport-entropy inequalities and concentration properties pushed forward in [15]. Observe that Erbar-Mass approach [11] does not allow to recover such concentration properties on discrete graphs.

New Prékopa-Leindler type of inequalities are also a straighforward dual consequence of the C-displacement convexity properties (3). Theorem 2.1 below is a general statement that applies for each of the discrete spaces $(\mathcal{X}, d, m, L)$ studied in this paper and presented next.

Theorem 2.1. *On a discrete space $(\mathcal{X}, d, m, L)$, assume that the relative entropy satisfies the C-displacement convexity property* (3) *with $C = (C_t)_{t\in[0,1]}$ given by : for any $\nu_0, \nu_1 \in \mathcal{P}_b(\mathcal{X})$*

$$C_t(\nu_0,\nu_1) = \iint c_t(x,y)\,d\widehat{\pi}^0(x,y),$$

where $\widehat{\pi}^0 = \widehat{Q}^0_{01}$, and $\widehat{Q}^0$ is the limit Schrödinger problem optimizer between ν_0 and ν_1. Then, the next property holds. If f, g, h are measurable functions on $\mathcal{X}$ satisfying

$$(1-t)f(x) + tg(y) \le \int h \, d\nu_t^{0x,y} + \frac{t(1-t)}{2}\, c_t(x,y), \qquad \forall x, y \in \mathcal{X},$$

then

$$\left(\int e^f \, dm\right)^{1-t} \left(\int e^g \, dm\right)^t \le \int e^h \, dm.$$

The proof of this result is an easy adaptation of the one of Theorem 6.3 in [15]. It is left to the reader.

For each of the discrete spaces $(\mathcal{X}, d, m, L)$ presented below, we describe the Schrödinger path at zero temperature and, as a main result, we give a C-displacement convexity property (3) satisfied by the reversible measure m by specifying the family of costs $C = (C_t)_{t\in[0,1]}$. The strategy of proof of these results is explained in section 3.

2.1. **The lattice $\mathbb{Z}^n$ endowed with the counting measure.** Let m denote the counting measure on $\mathcal{X} = \mathbb{Z}^n$. The graph structure on $\mathbb{Z}^n$ is given by the set of edges

$$E := \Big\{ (z, z+e_i), (z, z-e_i) \,\Big|\, z \in \mathbb{Z}^n, i \in [n] \Big\},$$

where $[n] := \{1, \dots, n\}$ and $(e_1, \dots, e_n)$ is the canonical base of $\mathbb{R}^n$. The graph distance is given by

$$d(x,y) = \sum_{i=1}^n |y_i - x_i|, \qquad x, y \in \mathbb{Z}^n.$$

The measure m is reversible with respect to the generator L defined by, for any $z \in \mathbb{Z}^n$, for any $i \in [n]$,

$$L(z, z+e_i) = L(z, z-e_i) = 1, \qquad L(z,z) = -2n.$$

For any integers $d, k_1, \dots, k_n$ such that $d = k_1 + \cdot + k_n$, $\binom{d}{k_1,\dots,k_n} = \frac{n!}{k_1!\cdots k_n!}$ denotes the multinomial coefficient. Since

$$L^{d(x,y)}(x,y) = \#G(x,y) = \binom{d(x,y)}{|y_1 - x_1|, \dots, |y_n - x_n|},$$

the Schrödinger bridge at zero temperature $(\widehat{Q}^0_t)_{t\in[0,1]}$ joining two measures $\nu_0, \nu_1 \in \mathcal{P}_b(\mathcal{X})$ is given by (18) with, according to (16),

$$\begin{aligned}
\nu_t^{0x,y}(z) &= \mathbb{1}_{[x,y]}(z) \frac{\binom{d(x,z)}{|z_1-x_1|,\dots,|z_n-x_n|}\binom{d(z,y)}{|y_1-z_1|,\dots,|y_n-z_n|}}{\binom{d(x,y)}{|y_1-x_1|,\dots,|y_n-x_n|}} \rho_t^{d(x,y)}(d(x,z)) \\
&= \mathbb{1}_{[x,y]}(z) \binom{|y_1 - x_1|}{|z_1 - x_1|} \cdots \binom{|y_n - x_n|}{|z_n - x_n|} t^{d(x,z)}(1-t)^{d(z,y)}, \qquad z \in \mathbb{Z}^n.
\end{aligned}$$

Observe that $(\nu_t^{0x,y})_{t\in[0,1]}$ is a binomial interpolation path as in the paper by E. Hillion [17].

Theorem 2.2. *On the space $(\mathbb{Z}^n, m, d, L)$, the relative entropy $H(\cdot|m)$ satisfies the 0-displacement convexity property* (3)*. In other words, for any Schrödinger bridge at zero temperature $(\widehat{Q}^0_t)_{t\in[0,1]}$ joining any two measures $\nu_0, \nu_1 \in \mathcal{P}_b(\mathbb{Z}^n)$, the map $t \mapsto H(\widehat{Q}^0_t|m)$ is convex.*

Therefore the space $(\mathbb{Z}^n, d, m, L)$ has positive W_1 or W_2^d-entropic curvature. It is a flat space. This convexity property along binomial interpolation paths has been first obtained by E. Hillion [17]. To compare with Hillion's method, the main interest of our approach is its simplicity. As explained in the next section, we first work at positive temperature $\gamma > 0$ so that the second derivative of the function $t \mapsto H(\widehat{Q}^\gamma_t|m)$ can be easily computed using Γ_2 calculus. Then we analyse the behaviour of the second derivative of this function as temperature goes to 0, and get a positive lower bound at zero temperature on $\mathbb{Z}^n$. This provides the convexity property of $t \mapsto H(\widehat{Q}^0_t|m)$. In Hillion's paper, one may say that computations are done directly at zero temperature. It leads to harder computations and the construction of the optimal coupling, related to a cyclic monotonicity property, is rather difficult to handle.

In the paper [16] by Gozlan & al., another kind of convexity property of entropy has been proposed that generalizes a new Prekopa-Leindler inequality on $\mathbb{Z}$ by Klartag-Lehec [20]. There convexity property is of different nature, it is only valid for $t = 1/2$. More precisely, given $\nu_0, \nu_1 \in \mathcal{P}_b(\mathbb{Z})$ they define two midpoint measures

$$\nu_- = m_-\#\pi \qquad \text{and} \qquad \nu_+ = m_+\#\pi,$$

where π is the monotone coupling between ν_0 and ν_1 (which is a W_1-optimizer), and for all $x, y \in \mathbb{Z}$,

$$m_-(x,y) := \left\lfloor \frac{x+y}{2} \right\rfloor, \qquad m_+(x,y) := \left\lceil \frac{x+y}{2} \right\rceil.$$

Gozlan & al. result [16, Theorem 8] states that

$$\frac{1}{2}H(\nu_-|m) + \frac{1}{2}H(\nu_+|m) \le \frac{1}{2}H(\nu_0|m) + \frac{1}{2}H(\nu_1|m).$$

As a main difference, the measures ν_+ and ν_- are only concentrated on the midpoints $m_-(x,y), m_+(x,y)$, for $x \in \operatorname{supp}(\nu_0)$ and $y \in \operatorname{supp}(\nu_1)$. Since ν_+ and ν_- are much more concentrated than $\widehat{Q}^0_{1/2}$, their result directly implies a Brunn-Minkovsky type of inequality. Unfortunately it seems that their approach do not extend to any values of $t \in (0,1)$.

2.2. **The complete graph.** Let $\mathcal{X}$ be a finite set and μ be any probability measure on $\mathcal{X}$. The set of edges of the complete graph $G = (\mathcal{X}, E)$ is $E := \mathcal{X} \times \mathcal{X}$ and the graph distance is the Hamming distance $d(x,y) := \mathbb{1}_{x\neq y}$ for any $x, y \in \mathcal{X}$. The measure μ is reversible with respect to the generator L given by : for any $z, z' \in \mathcal{X}$ with $z \neq z'$,

$$L(z,z') := \mu(z'), \qquad L(z,z) := -(1-\mu(z)).$$

The Schrödinger bridge at zero temperature $(\widehat{Q}^0_t)_{t\in[0,1]}$ given by (18), is the same as the bridge used in [14] for the complete graph (see section 2.1.1): for any $x, y \in \mathcal{X}$ one has

$$\nu_t^{0\,x,y}(z) = (1-t)\,\delta_x(z) + t\,\delta_y(z), \qquad z \in \mathcal{X}. \tag{19}$$

Theorem 2.3. *On the finite space $(\mathcal{X}, \mu, d, L)$, the relative entropy $H(\cdot|\mu)$ satisfies the C-displacement convexity property* (3)*, with $C = (C_t)_{t\in[0,1]}$ given by: for any $\nu_0, \nu_1 \in \mathcal{P}(\mathcal{X})$ with associated limit Schrödinger problem optimizer $\widehat{Q}^0 \in \mathcal{P}(\Omega)$,*

$$C_t(\nu_0,\nu_1) := \int h_t\left(\int \mathbb{1}_{w\neq x}\, d\widehat{\pi}^0_{\rightarrow}(w|x)\right) d\nu_0(x) + \int h_{1-t}\left(\int \mathbb{1}_{w\neq y}\, d\widehat{\pi}^0_{\leftarrow}(w|y)\right) d\nu_1(y),$$

where $\widehat{\pi}^0 = \widehat{Q}^0_{0,1}$ and $h_t(u) := 2\,\dfrac{t\ell(u) - \ell(tu)}{t(1-t)}$, $\ell(u) = (1-u)\log(1-u)$, $u \in [0,1)$.

The inequality $h_t(u) \ge u^2$, for all $u \in [0,1]$, $t \in (0,1)$, provides

$$C_t(\nu_0,\nu_1) \ge \widetilde{T}_2(\nu_0,\nu_1),$$

with

$$\widetilde{T}_2(\nu_0,\nu_1) := \int \left(\int \mathbb{1}_{w\neq x}\, d\widehat{\pi}_{\rightarrow}(w|x)\right)^2 d\nu_0(x) + \int \left(\int \mathbb{1}_{w\neq y}\, d\widehat{\pi}_{\leftarrow}(w|y)\right)^2 d\nu_1(y).$$

From this estimate, since $\widehat{Q}^0_t = (1-t)\nu_0 + t\nu_1$, one recovers a first convexity property of the relative entropy obtained by Gozlan & al. [14, Proposition 4.1].

Let us now compare $C_t(\nu_0,\nu_1)$ with a function of the total variation distance $\|\nu_0 - \nu_1\|_{TV}$. Recall that by Kantorovich duality

$$\|\nu_0 - \nu_1\|_{TV} := 2\sup_{A\subset X} |\nu_0(A) - \nu_1(A)| = 2\inf_{\pi\in\Pi(\nu_0,\nu_1)} \int \mathbb{1}_{x\neq y}\, d\pi(x,y) = 2W_1(\nu_0,\nu_1). \tag{20}$$

For any $x, y \in \mathcal{X}$, let $\Delta_{\rightarrow}(x) = \int \mathbb{1}_{w\neq x} d\widehat{\pi}^0_{\rightarrow}(w|x)$ and $\Delta_{\leftarrow}(y) = \int \mathbb{1}_{w\neq y} d\widehat{\pi}^0_{\leftarrow}(w|y)$. Define also

$$D_{\rightarrow} = \left\{x \in \mathrm{supp}(\nu_0) \,\middle|\, \Delta_{\rightarrow}(x) \neq 0\right\}, \qquad \text{and} \qquad D_{\leftarrow} = \left\{y \in \mathrm{supp}(\nu_1) \,\middle|\, \Delta_{\leftarrow}(y) \neq 0\right\}.$$

Observe that, using Lemma 4.2 (ii), since $\widehat{\pi}^0$ is a W_1-optimal coupling of ν_0 and ν_1, the sets $D_{\leftarrow}$ and $D_{\rightarrow}$ are disjoints. Since

$$\int \Delta_{\rightarrow} d\nu_0 = \int \Delta_{\leftarrow} d\nu_1 = W_1(\nu_0, \nu_1),$$

and h_t is convex, Jensen's inequality provides

$$\begin{aligned} C_t(\nu_0, \nu_1) &= \int_{D_{\rightarrow}} h_t\left(\Delta_{\rightarrow}\right) d\nu_0 + \int_{D_{\leftarrow}} h_{1-t}\left(\Delta_{\leftarrow}\right) d\nu_1 \\ &\geq \nu_0(D_{\rightarrow})\, h_t\left(\frac{W_1(\nu_0,\nu_1)}{\nu_0(D_{\rightarrow})}\right) + \nu_1(D_{\leftarrow})\, h_{1-t}\left(\frac{W_1(\nu_0,\nu_1)}{\nu_1(D_{\leftarrow})}\right). \end{aligned}$$

From (20), $W_1(\nu_0, \nu_1) \geq \nu_0(D_{\rightarrow}) - \nu_1(D_{\rightarrow})$, and therefore

$$\nu_0(D_{\rightarrow}) + \nu_1(D_{\leftarrow}) \leq W_1(\nu_0, \nu_1) + \nu_1(D_{\rightarrow}) + \nu_1(D_{\leftarrow}) + \leq W_1(\nu_0, \nu_1) + 1.$$

As a consequence, one gets the following result

$$C_t(\nu_0, \nu_1) \geq (1 + W_1(\nu_0, \nu_1))\, k_t\left(\frac{W_1(\nu_0,\nu_1)}{1 + W_1(\nu_0,\nu_1)}\right), \tag{21}$$

where for all $v \in [0, 1/2]$

$$k_t(v) := \inf_{\alpha,\beta,0<\alpha+\beta\leq 1} \left\{\alpha h_t\left(\frac{v}{\alpha}\right) + \beta h_{1-t}\left(\frac{v}{\beta}\right)\right\}.$$

In this definition, one needs to set $h_t(u) = +\infty$ for all $u > 1$ and $t \in (0, 1)$.

The function k_t can not be computed explicitly, however it can be estimated as follows. According to the proof of Theorem 2.3, for all $t \in (0, 1)$ and $v \in [0, 1]$,

$$h_t(v) = \int_0^1 \frac{v^2}{1 - uv} K_t(u)\, du,$$

with

$$K_t(u) = \frac{2u}{t}\mathbb{1}_{u\leq t} + \frac{2(1-u)}{1-t}\mathbb{1}_{u\geq t}, \qquad u \in [0, 1]. \tag{22}$$

As a consequence, since $K_{1-t}(u) = K_t(1 - u)$, one gets for any $v \in [0, 1/2]$,

$$k_t(v) \geq \int_0^1 \inf_{\alpha,\beta,\alpha>v,\beta>v,\alpha+\beta\leq 1} \left\{\frac{v^2}{\alpha - uv} + \frac{v^2}{\beta - (1-u)v}\right\} K_t(u) du.$$

Easy computations provide

$$\begin{aligned} \inf_{\alpha,\beta,\alpha>v,\beta>v,\alpha+\beta\leq 1} \left\{\frac{1}{\alpha - uv} + \frac{1}{\beta - (1-u)v}\right\} &= \inf_{\alpha',\beta',\alpha'>(1-u)v,\beta'>uv,\alpha'+\beta'\leq 1-v} \left\{\frac{1}{\alpha'} + \frac{1}{\beta'}\right\} \\ &\geq \inf_{\alpha',\beta',\alpha'>0,\beta'>0,\alpha'+\beta'\leq 1-v} \left\{\frac{1}{\alpha'} + \frac{1}{\beta'}\right\} = \frac{4}{1-v}. \end{aligned}$$

Since $\int_0^1 K_t(u)du = 1$, it implies

$$k_t(v) \geq \frac{4v^2}{1-v}, \tag{23}$$

and therefore

$$C_t(\nu_0, \nu_1) \geq 4W_1(\nu_0, \nu_1)^2 = \|\nu_0 - \nu_1\|_{TV}^2.$$

This lower estimate together with Theorem 2.3, also provides the second convexity property of the relative entropy given in [14, Proposition 4.1] with a different W_1-constant speed geodesic.

An improved version of the Csiszar-Kullback-Pinsker inequality follows from (21). Indeed, since by Jensen's inequality $H(\widehat{Q}_t^0|\mu) \geq 0$, the displacement convexity property (3) and (21) imply, for any $t \in (0,1)$,

$$\frac{1}{2}(1 + W_1(\nu_0,\nu_1))\, k_t\left(\frac{W_1(\nu_0,\nu_1)}{1 + W_1(\nu_0,\nu_1)}\right) \leq \frac{1}{t}\, H(\nu_0|\mu) + \frac{1}{1-t}\, H(\nu_1|\mu), \qquad \forall \nu_0,\nu_1 \in \mathcal{P}(\mathcal{X}).$$

The well-known Csiszar-Kullback-Pinsker inequality is obtained using (23) and then optimizing over all $t \in (0,1)$ (see [14, Remark 4.2]):

$$\frac{1}{2}\|\nu_0 - \nu_1\|_{TV}^2 \leq \left(\sqrt{H(\nu_0|\mu)} + \sqrt{H(\nu_1|\mu)}\right)^2, \qquad \forall \nu_0,\nu_1 \in \mathcal{P}(\mathcal{X}).$$

2.3. **Product measures on the discrete hypercube.** In this section, the reference space is the discrete hypercube $\mathcal{X} = \{0,1\}^n$ equipped with a product of Bernoulli measure

$$\mu = \mu_1 \otimes \cdots \otimes \mu_n,$$

with for any $i \in [n]$, $\mu_i(1) = 1 - \mu_i(0) := \alpha_i$, $\alpha_i \in (0,1)$.

For any $z = (z_1,\ldots,z_n) \in \{0,1\}^n$ and any $i \in [n]$ let $\sigma_i(z)$ denotes the neighbour of z according to the i's coordinate defined by

$$\sigma_i(z) := (z_1,\ldots,z_{i-1},\overline{z}_i,z_{i+1},\ldots,z_n),$$

where $\overline{z_i} := 1 - z_i$. The set of edges on $\{0,1\}^n$ is

$$E := \left\{(z,\sigma_i(z)) \,\middle|\, z \in \{0,1\}^n, i \in [n]\right\},$$

and the graph distance is the *Hamming distance* :

$$d(x,y) := \sum_{i=1}^n \mathbb{1}_{x_i \neq y_i}, \qquad x,y \in \{0,1\}^n.$$

The measure μ is reversible with respect to the generator L given by: for all $z \in \{0,1\}^n$,

$$L(z,\sigma_i(z)) := (1-\alpha_i)\, z_i + \alpha_i \overline{z_i}, \quad \forall i \in [n],$$

and $L(z,z) := -\sum_{i=1}^n L(z,\sigma_i(z))$. Observe that setting

$$L_i(z_i,\overline{z_i}) := (1-\alpha_i)\, z_i + \alpha_i \overline{z_i}, \quad z_i \in \{0,1\},$$

and $L_i(z_i,z_i) = -L_i(z_i,\overline{z_i})$, the Bernoulli measure μ_i is reversible with respect to L_i and one has

$$L := L_1 \oplus \cdots \oplus L_n.$$

Easy computations give that for any $x,y \in \{0,1\}^n$,

$$L^{d(x,y)}(x,y) = d(x,y)! \prod_{i=1}^n (1-\alpha_i)^{[x_i - y_i]_+} \alpha_i^{[y_i - x_i]_+},$$

and one gets that the Schrödinger bridge at zero temperature $(\widehat{Q}_t^0)_{t\in[0,1]}$ joining two probability measures ν_0 and ν_1 is given by (18), with according to (16)

$$\nu_t^{0\,x,y}(z) = \mathbb{1}_{[x,y]}(z)\, t^{d(x,z)}(1-t)^{d(z,y)}, \qquad z \in \{0,1\}^n.$$

This path has exactly the same structure as the one used in [14] to establish entropic curvature bounds on the product space $(\{0,1\}^n,\mu)$ (see section 2.1.2).

Theorem 2.4. *Let $\mu = \mu_1 \otimes \cdots \otimes \mu_n$ be a product probability measure on the discrete hypercube $\mathcal{X} = \{0,1\}^n$. On the space $(\{0,1\}^n, \mu, d, L)$, the relative entropy $H(\cdot|\mu)$ satisfies the C-displacement convexity property* (3)*, with $C = (C_t)_{t\in[0,1]}$ defined by: for any $\nu_0, \nu_1 \in \mathcal{P}(\{0,1\}^n)$ with associated limit Schrödinger problem optimizer $\widehat{Q}^0 \in \mathcal{P}(\Omega)$,*

$$C_t(\nu_0,\nu_1) := \max\Big[\int \sum_{i=1}^n h_t\left(\Pi^i_{\rightarrow}(x)\right) d\nu_0(x) + \int \sum_{i=1}^n h_{1-t}\left(\Pi^i_{\leftarrow}(y)\right) d\nu_1(y)\ ,\ 4\sum_{i=1}^n (\Pi^i)^2,$$
$$2C_n \iint d(x,w)(d(x,w)-1)\, d\widehat{\pi}^0(x,w)\Big],$$

where

$$\Pi^i_{\rightarrow}(x) := \int \mathbb{1}_{w_i\neq x_i} d\widehat{\pi}^0_{\rightarrow}(w|x),\ \Pi^i_{\leftarrow}(y) := \int \mathbb{1}_{w_i\neq y_i} d\widehat{\pi}^0_{\leftarrow}(w|y),\ \Pi^i := \iint \mathbb{1}_{x_i\neq y_i} d\widehat{\pi}^0(x,y),$$

and $\widehat{\pi}^0 = \widehat{Q}^0_{0,1}$, $C_n = -\log(1-1/n) \geq 1/n$, $h_t(u) := 2\,\dfrac{t\ell(u)-\ell(tu)}{t(1-t)}$, *and* $\ell(u) := (1-u)\log(1-u)$, $u \in [0,1)$.

Comments. • *By the Cauchy-Schwarz inequality, one has*

$$\sum_{i=1}^n (\Pi^i)^2 \geq \frac{1}{n} W_1^2(\nu_0,\nu_1).$$

As a consequence $C_t(\nu_0,\nu_1)$ is bounded from below by $4W_1^2(\nu_0,\nu_1)^2/n$, and the W_1-entropic curvature of the discrete hypercube $\{0,1\}^n$ is bounded from below by $4/n$.

As in the previous part to recover the Csiszar-Kullback-Pinsker inequality, the well-know W_1-optimal transport inequality on the discrete cube for product probability measures follows from the displacement convexity property (3)*, using $H(\widehat{Q}^0_t|\mu) \geq 0$ and optimizing over all $t \in (0,1)$:*

$$\frac{2}{n} W_1^2(\nu_0,\nu_1) \leq \left(\sqrt{H(\nu_0|\mu)} + \sqrt{H(\nu_1|\mu)}\right)^2, \qquad \forall \nu_0,\nu_1 \in \mathcal{P}(\{0,1\}^n).$$

Actually, Theorem 2.4 provides the following improvement of the W_1-optimal transport inequality

$$2 \inf_{\pi\in\Pi(\nu_0,\nu_1)} \sum_{i=1}^n \left(\iint \mathbb{1}_{x_i\neq y_i} d\pi(x,y)\right)^2 \leq \left(\sqrt{H(\nu_0|\mu)} + \sqrt{H(\nu_1|\mu)}\right)^2.$$

• *By bounding from below the cost $C_t(\nu_0,\nu_1)$ by*

$$\widetilde{T}_2(\nu_0,\nu_1) := \inf_{\pi\in\Pi(\nu_0,\nu_1)} \Big[\int \sum_{i=1}^n \left(\int \mathbb{1}_{w_i\neq x_i} d\pi_{\rightarrow}(w|x)\right)^2 d\nu_0(x)$$
$$+ \int \sum_{i=1}^n \left(\int \mathbb{1}_{w_i\neq y_i} d\pi_{\leftarrow}(w|y)\right)^2 d\nu_1(y)\Big],$$

one recovers a similar convexity property as the one obtained for the discrete cube in [14, Corollary 4.4]. *The only difference is in the expression* (18) *of the path $(\widehat{Q}^0_t)_{t\in[0,1]}$, the coupling measure $\widehat{\pi}^0$ is replaced by an optimal Knothe-Rosenblatt coupling.*

Marton's transport entropy inequality on the discrete hypercube is a consequence of the last lower bound on $C_t(\nu_0,\nu_1)$: for any $\nu_0,\nu_1 \in \mathcal{P}(\{0,1\}^n)$,

$$\frac{1}{2}\widetilde{T}_2(\nu_0,\nu_1) \leq \left(\sqrt{H(\nu_0|\mu)} + \sqrt{H(\nu_1|\mu)}\right)^2.$$

- *The cost $C_t(\nu_0,\nu_1)$ can be also bounded from below by*

$$2C_n\left(W_2^2(\nu_0,\nu_1)-W_1(\nu_0,\nu_1)\right)\geq 2C_n\,W_2^{d^2}(\nu_0,\nu_1).$$

with W_2^d defined by (4). *Therefore the discrete hypercube has also W_2^d-entropic curvature bounded from below by $2C_n\geq 2/n$. From this estimate, Theorem 2.4 provides the following new transport-entropy inequality, for any $\nu_0,\nu_1\in\mathcal{P}(\{0,1\}^n)$,*

$$C_nW_2^d(\nu_0,\nu_1)^2\leq C_n\left(W_2^2(\nu_0,\nu_1)-W_1(\nu_0,\nu_1)\right)\leq\left(\sqrt{H(\nu_0|\mu)}+\sqrt{H(\nu_1|\mu)}\right)^2. \tag{24}$$

As opposed to Marton's transport inequality or to W_2-Talagrand's transport inequality on Euclidean space, inequality (24) *on the hypercube does not tensorize. Nevertheless, it can be interpreted as a discrete analogue on the hypercube of the W_2-Talagrand's transport inequality. Indeed, from* (24)*, applying the central limit theorem, one recover, up to constant, the well-known W_2-transport entropy inequality for the standard Gaussian probability measure γ on $\mathbb{R}$, due to Talagrand* [39]. *Namely, for any absolutely continuous probability measure $\nu\in\mathcal{P}_2(\mathbb{R})$,*

$$W_2^2(\nu,\gamma)\leq 2H(\nu|\gamma). \tag{25}$$

For a sake of completeness, the proof of this implication is given in Appendix (see Lemma 4.1). Unfortunately, to recover (25)*, the constant $2/n$ is expected, instead C_n in the left-hand side of* (24)*, like in the W_1-transport entropy inequality. Improving the transport-inequality* (24) *in order to recover* (25) *is a remaining problem.*

2.4. **The circle $\mathbb{Z}/N\mathbb{Z}$ endowed with a uniform measure.** Let $N\in\mathbb{N}$ and $\mathcal{X}$ be the space $\mathbb{Z}/N\mathbb{Z}$, endowed with the uniform probability measure μ, $\mu(x)=1/N$. The measure μ is reversible with respect to the generator L given by ,

$$L(z,z+1)=L(z,z-1)=1,\qquad L(z,z)=-2,$$

for any $z\in\mathbb{Z}/N\mathbb{Z}$. One always have $d(x,y)\leq\lfloor N/2\rfloor=n$ where $\lfloor\cdot\rfloor$ denotes the floor function.

If N is odd then for any $x,y\in\mathbb{Z}/N\mathbb{Z}$, $L^{d(x,y)}(x,y)=1$ and therefore the Schrödinger bridge at zero temperature $(\widehat{Q_t^0})_{t\in[0,1]}$ joining two probability measures ν_0 and ν_1 on Z/NZ is given by (18), with according to (16)

$$\nu_t^{0^{x,y}}(z)=\mathbb{1}_{z\in[x,y]}\,\rho_t^{d(x,y)}(d(x,z)).$$

If N is even then for any $x,y\in\mathbb{Z}/N\mathbb{Z}$ such that $d(x,y)<N/2$, $L^{d(x,y)}(x,y)=1$ and $L^{d(x,x+n)}(x,x+n)=2$. The Schrödinger bridge at zero temperature $(\widehat{Q_t^0})_{t\in[0,1]}$ is given by (18), with according to (16) : if $d(x,y)<N/2$ then

$$\nu_t^{0^{x,y}}(z)=\mathbb{1}_{z\in[x,y]}\,\rho_t^{d(x,y)}(d(x,z)),$$

and if $d(x,y)=N/2$ ($y=x+n$), for any $z\in\mathbb{Z}/N\mathbb{Z}\setminus\{x,x+n\}$,

$$\nu_t^{0^{x,x+n}}(z)=\frac{1}{2}\mathbb{1}_{z\in[x,x+n]}\,\rho_t^{d(x,x+n)}(d(x,z)),$$

and $\nu_t^{0^{x,x+n}}(x)=(1-t)^{d(x,x+n)}$, $\nu_t^{0^{x,x+n}}(x+n)=t^{d(x,x+n)}$.

Theorem 2.5. *On the space $(\mathbb{Z}/N\mathbb{Z},\mu,d,L)$, the relative entropy $H(\cdot|\mu)$ satisfies the 0-displacement convexity* (3).

Therefore the space $(\mathbb{Z}/N\mathbb{Z},d,\mu,L)$ has positive entropic curvature, it is a flat space.

2.5. **The Bernoulli-Laplace model.** Let $\mathcal{X} = \mathcal{X}_k$ denotes the slice of the discrete hypercube $\{0,1\}^n$ of order $k \in [n-1]$, endowed with the uniform probability measure μ, namely

$$\mathcal{X}_k := \Big\{x = (x_1, \ldots, x_n) \in \{0,1\} \,\Big|\, x_1 + \ldots + x_n = k\Big\}.$$

For $z \in \mathcal{X}_k$, we note $J_0(z) := \{i \in [n] \,|\, z_i = 0\}$ and $J_1(z) := \{i \in [n] \,|\, z_i = 1\}$. For any $i \in J_0(z)$ and $j \in J_1(z)$, one notes $\sigma_{ij}(z)$ the neighbour of z in $\mathcal{X}_k$ defined by

$$\Big(\sigma_{ij}(z)\Big)_i = 1, \quad \Big(\sigma_{ij}(z)\Big)_j = 0,$$

and for any $\ell \in [n] \setminus \{i,j\}$, $\Big(\sigma_{ij}(z)\Big)_\ell = z_\ell$. The set of edges of the graph is

$$E := \Big\{(z, \sigma_{ij}(z)) \,\Big|\, z \in \mathcal{X}_k, \{i,j\} \subset [n], z_i = 0, z_j = 1\Big\},$$

and the graph distance is given by

$$d(x,y) := \frac{1}{2}\sum_{i=1}^n \mathbb{1}_{x_i \neq y_i}, \qquad x, y \in \mathcal{X}_k.$$

The measure μ is reversible with respect to the generator L given by $L(z, \sigma_{ij}(z)) := 1$ for any i, j such that $z_i = 0$ and $z_j = 1$, and $L(z,z) := -k(n-k)$.

Since $L^{d(x,y)}(x,y) = (d(x,y)!)^2$, the Schrödinger bridge at zero temperature $(\widehat{Q}^0_t)_{t\in[0,1]}$ is given by (18), with according to (16),

$$\nu_t^{0^{x,y}}(z) = \mathbb{1}_{[x,y]}(z)\binom{d(x,y)}{d(x,z)}^{-1} t^{d(x,z)}(1-t)^{d(z,y)}, \qquad z \in \mathcal{X}_k.$$

Theorem 2.6. *On the space $(\mathcal{X}_k, \mu, d, L)$, the relative entropy $H(\cdot|\mu)$ satisfies the C-displacement convexity property (3), with $C = (C_t)_{t\in[0,1]}$ defined by: for any $\nu_0, \nu_1 \in \mathcal{P}(\mathcal{X}_k)$ with associated limit Schrödinger problem optimizer $\widehat{Q}^0 \in \mathcal{P}(\Omega)$,*

$$C_t(\nu_0,\nu_1) := \max\Big[\frac{4}{\min(k,n-k)} W_1^2(\nu_0,\nu_1)\,,\ \tilde{c}_t(\widehat{\pi}^0)\,,\ 2C_{n,k}\,\hat{c}(\widehat{\pi}^0)\Big]$$

where $\widehat{\pi}^0 = \widehat{Q}^0_{0,1}$, $C_{n,k} := -\log\Big(1 - \frac{1}{\min(k,n-k)}\Big) \geq \frac{1}{\min(k,n-k)}$, and

$$\hat{c}(\widehat{\pi}^0) := \iint d(x,w)(d(x,w)-1)\, d\widehat{\pi}^0(x,w),$$

$$\tilde{c}_t(\widehat{\pi}^0) := \max\Big[\int \sum_{i\in J_0(x)} h_t\Big(\Pi^i_{\rightarrow}(x)\Big)\, d\nu_0(x), \int \sum_{i\in J_1(x)} h_t\Big(\Pi^i_{\rightarrow}(x)\Big)\, d\nu_0(x)\Big]$$
$$+ \max\Big[\int \sum_{i\in J_0(y)} h_{1-t}\Big(\Pi^i_{\leftarrow}(y)\Big)\, d\nu_1(y), \int \sum_{i\in J_1(y)} h_{1-t}\Big(\Pi^i_{\leftarrow}(y)\Big)\, d\nu_1(y)\Big],$$

with

$$\Pi^i_{\rightarrow}(x) := \int \mathbb{1}_{w_i\neq x_i} d\widehat{\pi}^0_{\rightarrow}(w|x), \qquad \Pi^i_{\leftarrow}(y) := \int \mathbb{1}_{w_i\neq y_i} d\widehat{\pi}^0_{\leftarrow}(w|y),$$

and $h_t(u) := 2\,\dfrac{t\ell(u) - \ell(tu)}{t(1-t)}$, $\ell(u) := (1-u)\log(1-u)$, $u \in [0,1)$.

Comments. • *Let*

$$\widetilde{T}_2(\nu_0,\nu_1) := \inf_{\pi\in\Pi(\nu_0,\nu_1)} \Big[\int \sum_{i=1}^n \left(\int \mathbb{1}_{w_i\neq x_i}\, d\pi_{\rightarrow}(w|x) \right)^2 d\nu_0(x) + \int \sum_{i=1}^n \left(\int \mathbb{1}_{w_i\neq y_i}\, d\pi_{\leftarrow}(w|y) \right)^2 d\nu_1(y) \Big].$$

One has

$$C_t(\nu_0,\nu_1) \geq \tilde{c}_t(\widehat{\pi}^0) \geq \frac{1}{2}\int \sum_{i\in[n]} h_t\left(\Pi^i_{\rightarrow}(x)\right) d\nu_0(x) + \frac{1}{2}\int \sum_{i\in[n]} h_{1-t}\left(\Pi^i_{\leftarrow}(y)\right) d\nu_1(y)$$
$$\geq \widetilde{T}_2(\nu_0,\nu_1).$$

As a consequence, since $H(\widehat{Q}^0_t|\mu) \geq 0$, optimizing over all $t \in (0,1)$, Theorem 2.6 implies the following weak transport-entropy inequality, for any $\nu_0,\nu_1 \in \mathcal{P}(\mathcal{X}_k)$,

$$\frac{1}{2}\,\widetilde{T}_2(\nu_0,\nu_1) \leq \left(\sqrt{H(\nu_0|\mu)} + \sqrt{H(\nu_1|\mu)} \right)^2.$$

This inequality has been first surprisingly obtained in [35, Theorem 1.8 (b)] *by projection of a transport-entropy inequality for the uniform measure on the symmetric group, but with the worse constant* $1/8$ *instead of* $1/2$*. Our approach is much more natural to reach such a result.*

- *Since $C_t(\nu_0,\nu_1) \geq \frac{4}{\min(k,n-k)}\, W_1^2(\nu_0,\nu_1)$, the W_1-entropic curvature of the space $(\mathcal{X}_k,d,L)$ is bounded from below by $\frac{4}{\min(k,n-k)}$. Observe that this constant is optimal since for $k=1$ or $k=n-1$, $\mathcal{X}_k$ is the complete graph and one recovers its optimal lower curvature bound* 4*. Similarly since*

$$\hat{c}(\widehat{\pi}^0) \geq W_2(\nu_0,\nu_1)^2 - W_1(\nu_0,\nu_1) \geq W_2^d(\nu_0,\nu_1)^2,$$

the W_2^d-entropic curvature of the space $(\mathcal{X}_k,d,L)$ is bounded from below by $\frac{2}{\min(k,n-k)}$.

3. Proof of the main results

This section is divided into two parts. We first present general statements to prove displacement convexity property (3) along Schrödinger bridges at zero temperature. Then we show how it applies for each involved discrete space of the last part.

3.1. Strategy of proof, general statements to get entropic curvature results.

As in the paper by G. Conforti [8] in continuous setting, the first step is to decompose the relative-entropy using the product structure given by (8): for any $t\in[0,1]$,

$$H(\widehat{Q}^\gamma_t|m) = \varphi_\gamma(t) + \psi_\gamma(t),$$

where

$$\varphi_\gamma(t) := \int \log(P^\gamma_t f^\gamma) P^\gamma_t f^\gamma\, P^\gamma_{1-t} g^\gamma\, dm \quad \text{and} \quad \psi_\gamma(t) := \int \log(P^\gamma_{1-t} g^\gamma) P^\gamma_{1-t} g^\gamma\, P^\gamma_t f^\gamma\, dm.$$

As recalled below, it is known that the function φ_γ is non-increasing and the function ψ_γ is non-decreasing (see [25, Theorem 6.4.2]).

Then, the strategy is to analyse the behaviour of the second order derivative φ''_γ and ψ''_γ as γ goes to 0, in order to apply the next Lemma.

Lemma 3.1. *Let ν_0 and ν_1 in $\mathcal{P}_b(\mathcal{X})$ and assume that hypothesis* (11), (12), (13) *and* (14) *hold. Let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of positive numbers that converges to 0. If for any $t \in (0,1)$*

$$\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \geq \varphi''_0(t), \quad \textit{and} \quad \liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t) \geq \psi''_0(t), \tag{26}$$

where φ_0 and ψ_0 are continuous functions on [0,1], twice differentiable on $(0,1)$, then the displacement convexity property (3) *holds with*

$$C_t(\nu_0,\nu_1) := \frac{2}{t(1-t)}\Big[(1-t)\varphi_0(0) + t\varphi_0(1) - \varphi_0(t) + (1-t)\psi_0(0) + t\psi_0(1) - \psi_0(t)\Big].$$

Observe that if $\varphi''_0 = K_\varphi$ and $\psi''_0 = K_\psi$ are constant functions, then

$$C_t(\nu_0,\nu_1) = \Big(K_\varphi + K_\psi\Big).$$

The proof of this lemma is postponed in Appendix B.

In order to apply Lemma 3.1, we need first to compute $\varphi'_\gamma, \psi'_\gamma$ and $\varphi''_\gamma, \psi''_\gamma$ in a suitable form so as to get (26). For any real function u on $\mathcal{X}$, we note

$$\nabla u(z,w) = u(w) - u(z), \quad z,w \in \mathcal{X},$$

and

$$Lu(z) := \sum_{w\in\mathcal{X}} u(w)\, L(z,w) = \sum_{w, w\sim z} \nabla u(z,w)\, L(z,w).$$

The expressions of $\varphi'_\gamma, \psi'_\gamma$ and $\varphi''_\gamma, \psi''_\gamma$ are given by the next lemmas. These expressions can be found in Léonard's paper [25, section 6.4] in a more general framework (for stationary non-reversible Markov processes). For completeness, the proof of the next result is recalled in Appendix B.

Lemma 3.2. *For any $t \in (0,1)$, one has*

$$\varphi'_\gamma(t) = -\int \sum_{z', z'\sim z} \zeta(e^{\nabla F^\gamma_t(z,z')})\, L^\gamma(z,z')\, d\widehat{Q}^\gamma_t(z),$$

and

$$\psi'_\gamma(t) = \int \sum_{z', z'\sim z} \zeta(e^{\nabla G^\gamma_t(z,z')})\, L^\gamma(z,z')\, d\widehat{Q}^\gamma_t(z),$$

where $\zeta(s) := s\log s - s + 1$, $s > 0$, and G^γ_t and F^γ_t are the so-called Schrödinger potentials according to Léonard's paper terminology [25],

$$G^\gamma_t := \log P^\gamma_{1-t} g^\gamma, \qquad \textit{and} \qquad F^\gamma_t := \log P^\gamma_t f^\gamma.$$

Since $\zeta \geq 0$, the function φ_γ is non-increasing and the function ψ_γ is non-decreasing.

Lemma 3.3. *For any $a > 0, b > 0$, let*

$$\rho(a,b) := (\log b - 2\log a - 1)\, b,$$

and let $\rho(a,b)=0$ if either $a=0$ or $b=0$. For any $t\in(0,1)$, one has

$$\begin{aligned}\varphi_\gamma''(t) = \int \Big[\Big(\sum_{z',z'\sim z} e^{\nabla F_t^\gamma(z,z')} L^\gamma(z,z')\Big)^2 \\ + \sum_{z',z'\sim z} \Big(1+\nabla F_t^\gamma(z,z')\Big)\, e^{\nabla F_t^\gamma(z,z')}\Big(L^\gamma(z,z)-L^\gamma(z',z')\Big) L^\gamma(z,z') \\ + \sum_{z',z'',z\sim z'\sim z''} \rho\Big(e^{\nabla F_t^\gamma(z,z')}, e^{\nabla F_t^\gamma(z,z'')}\Big) L^\gamma(z,z')L^\gamma(z',z'')\Big]\, d\widehat{Q}_t^\gamma(z),\end{aligned}$$

$$\begin{aligned}\psi_\gamma''(t) = \int \Big[\Big(\sum_{z',z'\sim z} e^{\nabla G_t^\gamma(z,z')} L^\gamma(z,z')\Big)^2 \\ + \sum_{z',z'\sim z} \Big(1+\nabla G_t^\gamma(z,z')\Big)\, e^{\nabla G_t^\gamma(z,z')}\Big(L^\gamma(z,z)-L^\gamma(z',z')\Big) L^\gamma(z,z') \\ + \sum_{z',z'',z\sim z'\sim z''} \rho\Big(e^{\nabla G_t^\gamma(z,z')}, e^{\nabla G_t^\gamma(z,z'')}\Big) L^\gamma(z,z')L^\gamma(z',z'')\Big]\, d\widehat{Q}_t^\gamma(z).\end{aligned}$$

Let us now analyse the behaviour of $\varphi_\gamma''(t)$, $\psi_\gamma''(t)$ as temperature γ goes to zero. For any $z,w\in\mathcal{X}$, we set

$$A_t^\gamma(z,w) := e^{\nabla F_t^\gamma(z,w)} = \frac{P_t^\gamma f^\gamma(w)}{P_t^\gamma f^\gamma(z)}, \quad \text{and} \quad B_t^\gamma(z,w) := e^{\nabla G_t^\gamma(z,w)} = \frac{P_{1-t}^\gamma g^\gamma(w)}{P_{1-t}^\gamma g^\gamma(z)}.$$

The next key lemma gives Taylor expansions as γ goes to zero of the quantities $A_t^\gamma(z,w)$, $B_t^\gamma(z,w)$ if $z\sim w$ or if $d(z,w)=2$. Its proof is postponed in Appendix B. For any $z,y\in\mathcal{X}$, and any $t\in(0,1)$, one notes

$$a_t(z,y) := \widehat{Q}^0(X_t=z|X_1=y) = \int \nu_t^{0^{w,y}}(z)\, d\widehat{\pi}^0_{\leftarrow}(w|y), \tag{27}$$

and

$$b_t(z,x) := \widehat{Q}^0(X_t=z|X_0=x) = \int \nu_t^{0^{x,w}}(z)\, d\widehat{\pi}^0_{\rightarrow}(w|x).$$

Lemma 3.4. *Assume that conditions* (12) *and* (13) *are fulfilled. Let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of positive numbers converging to 0, and let $\widehat{Q}_t^0$ denote the weak limit of the sequence of probability measures $(\widehat{Q}_t^{\gamma_k})_{k\in\mathbb{N}}$. Let $z\in\mathcal{X}$, $y\in\operatorname{supp}(\nu_1)$, $x\in\operatorname{supp}(\nu_0)$ such that $a_t(z,y)\neq 0$ and $b_t(z,x)\neq 0$.*

- *For any $z'\in\mathcal{X}$ such that $z'\sim z$, define*

  $$\mathrm{a}_t(z,z',y) \;:=\; \textstyle\sum_{w\in\mathcal{X},(z,z')\in[y,w]} r(y,z,z',w)\, d(y,w)\, \rho_t^{d(y,w)-1}(d(z,w)-1)\, \widehat{\pi}^0_{\leftarrow}(w|y),$$

  and

  $$\mathrm{b}_t(z,z',x) \;:=\; \textstyle\sum_{w\in\mathcal{X},(z,z')\in[x,w]} r(x,z,z',w)\, d(x,w)\, \rho_t^{d(x,w)-1}(d(x,z))\, \widehat{\pi}^0_{\rightarrow}(w|x),$$

  where the function r is given by (17). *It holds*

  $$\lim_{k\to\infty}\Big(\gamma_k A_t^{\gamma_k}(z,z')\Big) = A_t(z,z',y) \quad \text{and} \quad \lim_{k\to\infty}\Big(\gamma_k B_t^{\gamma_k}(z,z')\Big) = B_t(z,z',x), \tag{28}$$

  with

  $$A_t(z,z',y) := \frac{\mathrm{a}_t(z,z',y)}{a_t(z,y)} \quad \text{and} \quad B_t(z,z',x) := \frac{\mathrm{b}_t(z,z',x)}{b_t(z,x)}.$$

- *For any $z,z''\in\mathcal{X}$ with $d(z,z'')=2$, define*

$$\mathbb{a}_t(z,z'',y)$$

$$:= \sum_{w\in\mathcal{X},(z,z'')\in[y,w]} r(y,z,z'',w)\, d(y,w)(d(y,w)-1)\, \rho_t^{d(y,w)-2}(d(z,w)-2)\, \widehat{\pi}^0_{\leftarrow}(w|y),$$

and

$$\mathbb{b}_t(z,z'',x) := \sum_{w\in\mathcal{X},(z,z'')\in[x,w]} r(x,z,z'',w)\, d(x,w)(d(x,w)-1)\,\rho_t^{d(x,w)-2}(d(x,z))\,\widehat{\pi}^0_{\rightarrow}(w|x).$$

It holds

$$\lim_{k\to\infty}\left({\gamma_k}^2 A_t^{\gamma_k}(z,z'')\right) = \mathbb{A}_t(z,z'',y) \text{ and } \lim_{k\to\infty}\left({\gamma_k}^2 B_t^{\gamma_k}(z,z'')\right) = \mathbb{B}_t(z,z'',y), \tag{29}$$

with

$$\mathbb{A}_t(z,z'',y) := \frac{\mathbb{a}_t(z,z'',y)}{a_t(z,y)}, \quad \text{and} \quad \mathbb{B}_t(z,z'',x) := \frac{\mathbb{b}_t(z,z'',x)}{b_t(z,x)}.$$

Remark. *(1) For any $t\in(0,1)$, $z\in\mathcal{X}$ and $y\in\operatorname{supp}(\nu_1)$, according to the definition* (27)*, $a_t(z,y)\neq 0$ if and only if there exists $w\in\operatorname{supp}(\nu_0)$ such that $(w,y)\in\operatorname{supp}(\widehat{\pi}^0)$.*
Identically, for any $t\in(0,1)$, $z\in\mathcal{X}$ and $x\in\operatorname{supp}(\nu_0)$, $b_t(z,x)\neq 0$ if and only if there exists $w\in\operatorname{supp}(\nu_1)$ such that $(x,w)\in\operatorname{supp}(\widehat{\pi}^0)$.

(2) For any $t\in(0,1)$, $z\in\mathcal{X}$ and $y\in\operatorname{supp}(\nu_1)$, if $\mathrm{a}_t(z,z',y)\neq 0$ for some $z'\sim z$ or if $\mathbb{a}_t(z,z'',y)\neq 0$ for some z'' with $d(z,z'')=2$ then $a_t(z,y)\neq 0$. The same property holds with $b_t, \mathrm{b}_t, \mathbb{b}_t$.

Lemma 3.4 provides the following Taylor estimates for the functions φ''_{γ_k} and ψ''_{γ_k} as γ_k goes to 0.

Proposition 3.5. *Assume that conditions* (12), (13) *and* (14) *are fulfilled. Let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of positive numbers converging to 0 and $\widehat{Q}_t^0$ denotes the weak limit of the sequence of probability measures $(\widehat{Q}_t^{\gamma_k})_{k\in\mathbb{N}}$. According to the notations of Lemma 3.4, for any $t\in(0,1)$, one has*

$$\liminf_{\gamma_k\to0}\varphi''_{\gamma_k}(t) \geq \iint \Big[\Big(\sum_{z',z'\sim z} A_t(z,z',y)\,L(z,z')\Big)^2 + \sum_{z',z'',z\sim z'\sim z'',d(z,z'')=2} \rho\Big(A_t(z,z',y),\mathbb{A}_t(z,z'',y)\Big)L(z',z'')L(z,z')\Big]d\widehat{Q}^0_{t1}(z,y),$$

$$\liminf_{\gamma_k\to0}\psi''_{\gamma_k}(t) \geq \iint \Big[\Big(\sum_{z',z'\sim z} B_t(z,z',x)\,L(z,z')\Big)^2 + \sum_{z',z'',z\sim z'\sim z'',d(z,z'')=2} \rho\Big(B_t(z,z',x),\mathbb{B}_t(z,z'',x)\Big)L(z',z'')L(z,z')\Big]d\widehat{Q}^0_{0t}(x,z).$$

Proof of Proposition 3.5. We only prove the lower bound of $\liminf_{\gamma_k\to0}\varphi''_{\gamma_k}(t)$ since by symmetry, identical arguments provides the lower bound of $\liminf_{\gamma_k\to0}\psi''_{\gamma_k}(t)$. We start with the expression of $\varphi''_\gamma(t)$, $t\in(0,1)$, given by Lemma 3.3,

$$\varphi''_\gamma(t) = \int\left(M_t^\gamma + R_t^\gamma\right)d\widehat{Q}_t^\gamma, \tag{30}$$

with for any $z\in\mathcal{X}$,

$$M_t^\gamma(z) := \Big(\sum_{z',z'\sim z} e^{\nabla F_t^\gamma(z,z')}L^\gamma(z,z')\Big)^2 + \sum_{z',z'',z\sim z'\sim z''} \rho\left(e^{\nabla F_t^\gamma(z,z')}, e^{\nabla F_t^\gamma(z,z'')}\right)L^\gamma(z,z')L^\gamma(z',z''),$$

and

$$R_t^\gamma(z) := \sum_{z',z'\sim z}\left(1+\nabla F_t^\gamma(z,z')\right)e^{\nabla F_t^\gamma(z,z')}\left(L^\gamma(z,z)-L^\gamma(z',z')\right)L^\gamma(z,z').$$

We will get the behaviour of $\varphi_\gamma''(t)$ as γ goes to zero by applying Fatou's Lemma. Let us first bound $|M_t^\gamma(z)|$ and $|R_t^\gamma(z)|$ uniformly in γ, for γ sufficiently small. According to the definition of A_t^γ and from hypothesis (12), for any $z \in \mathcal{X}$

$$|M_t^\gamma(z)| \leq \gamma^2 S^2 d_{\max}^2 \max_{z',z'\sim z} |A_t^\gamma(z,z')|^2 + \gamma^2 S^2 d_{\max}^2 \max_{z',z'',z''\sim z'\sim z} |\rho(A_t^\gamma(z,z'), A_t^\gamma(z,z''))|.$$

Using the convexity inequality $\log b - \log a \leq (b-a)/a$, one easily check that

$$|\rho(A_t^\gamma(z,z'), A_t^\gamma(z,z''))| \leq \max\left(A_t^\gamma(z,z')^2, 2A_t^\gamma(z,z'')\right).$$

Applying inequality (51), it follows that

$$|M_t^\gamma(z)| \leq \frac{\left(d^2(x_0,z)+1\right) K^{2d(x_0,z)}\, O(1)}{t^2}. \tag{31}$$

Similarly, from (12) and (51), one may show that

$$|R_t^\gamma(z)| \leq \frac{\gamma}{t}\Big[\log\left(\frac{1}{\gamma}\right) + d(x_0,z)\Big]\, d(x_0,z)\, K^{d(x_0,z)}\, O(1) \leq \frac{|\gamma\log\gamma|}{t}\, d^2(x_0,z)\, K^{d(x_0,z)}\, O(1). \tag{32}$$

Lemma 4.3 (vii) therefore implies for any $z \in \mathcal{X}$ and any $0 \leq \gamma < \gamma_1$,

$$\begin{aligned}&|M_t^\gamma(z) + R_t^\gamma(z)|\, \widehat{Q}_t^\gamma(z)\\ &\qquad\leq O(1)\left(\mathbb{1}_B(z) + \mathbb{1}_{\mathcal{X}\setminus B}(z)\,\gamma_1\, \left(\gamma_1 K^2\right)^{[2d(x_0,z)-4D-1]_+}\right)\left(d^2(x_0,z)+1\right) K^{2d(x_0,z)}.\end{aligned}$$

It remains to choose γ_1 such that $(\gamma_1 K^3)^2 < \gamma_0$ so that hypothesis (14) implies

$$\sum_{z\in\mathcal{X}}\left(\mathbb{1}_B(z) + \mathbb{1}_{\mathcal{X}\setminus B}(z)\,\gamma_1\, \left(\gamma_1 K^2\right)^{[2d(x_0,z)-4D-1]_+}\right)\left(d^2(x_0,z)+1\right) K^{2d(x_0,z)} < \infty.$$

Now, conditions for Fatou's Lemma are fulfilled and one has

$$\lim_{\gamma_k\to 0} \varphi_\gamma''(t) \geq \sum_{z\in\mathcal{X}} \liminf_{\gamma_k\to 0}\Big[\left(M_t^{\gamma_k}(z) + R_t^{\gamma_k}(z)\right)\widehat{Q}_t^{\gamma_k}(z)\Big].$$

The weak convergence of $(\widehat{Q}^{\gamma_k})$ to $\widehat{Q}^0$ implies $\lim_{\gamma_k\to 0}\widehat{Q}_t^{\gamma_k}(z) = \widehat{Q}_t^0(z)$, and the inequality (32) gives $\lim_{\gamma_k\to 0} R_t^{\gamma_k}(z) = 0$. As a consequence,

$$\liminf_{\gamma_k\to 0}\Big[\left(M_t^{\gamma_k}(z) + R_t^{\gamma_k}(z)\right)\widehat{Q}_t^{\gamma_k}(z)\Big] = \liminf_{\gamma_k\to 0} M_t^{\gamma_k}(z)\,\widehat{Q}_t^0(z).$$

Since

$$\sum_{z\in\mathcal{X}} M_t^{\gamma_k}(z)\, Q_t^0(z) = \sum_{y\in\mathrm{supp}(\nu_1)} M_t^{\gamma_k}(z)\, a_t(z,y)\,\nu_1(y),$$

in order to complete the proof Proposition 3.5, it remains to bound from below $\liminf_{\gamma_k\to 0} M_t^{\gamma_k}(z)$ for any z, y such that $a_t(z,y) \neq 0$. One has

$$M_t^{\gamma_k} = E_t^{\gamma_k} - F_t^{\gamma_k} + G_t^{\gamma_k},$$

where for any $z \in \mathcal{X}$,

$$E_t^{\gamma_k}(z) := \Big(\sum_{z',z'\sim z} \gamma_k A_t^{\gamma_k}(z,z')\, L(z,z')\Big)^2,\ F_t^{\gamma_k}(z) = 2\sum_{z',z'',z\sim z'\sim z''} \gamma_k^2 A_t^{\gamma_k}(z,z'')\, L(z,z')L(z',z''),$$

and

$$G_t^{\gamma_k}(z) = \sum_{z',z'',z\sim z'\sim z''} \gamma_k^2\left[\rho\left(A_t^{\gamma_k}(z,z'), A_t^{\gamma_k}(z,z'')\right) + 2A_t^{\gamma_k}(z,z'')\right] L(z,z')L(z',z'').$$

Since $a_t(z,y)\neq 0$, Lemma 3.4 implies

$$\lim_{\gamma_k\to 0} E_t^{\gamma_k}(z)=\Big(\sum_{z',\,z'\sim z} A_t(z,z',y)\,L(z,z')\Big)^2,$$

and since $\lim_{\gamma_k\to 0}\left(\gamma_k^2 A_t^{\gamma_k}(z,z'')\right)=0$ if $d(z,z'')\leq 1$ and $\lim_{\gamma_k\to 0}\left(\gamma_k^2 A_t^{\gamma_k}(z,z'')\right)=A_t^{\gamma_k}(z,z'')$ if $d(z,z'')=2$, it also implies

$$\lim_{\gamma_k\to 0} F_t^{\gamma_k}(z)=2\sum_{z',\,z'',\,z\sim z'\sim z'',\,d(z,z'')=2} \mathbb{A}_t(z,z'',y)\,L(z,z')L(z',z'').$$

Observing that $\rho(a,b)+2b\geq 0$ and $\gamma^2\rho(a,b)=\rho(\gamma a,\gamma^2 b)$ for any $a>0, b>0, \gamma>0$, one gets

$$\liminf_{\gamma_k\to 0} G_t^{\gamma_k}(z)$$
$$\geq \sum_{z',\,z'',\,z\sim z'\sim z'',\,d(z,z'')=2} \liminf_{\gamma_k}\left[\rho\left(\gamma_k A_t^{\gamma_k}(z,z'),\gamma_k^2 A_t^{\gamma_k}(z,z'')\right)+2\gamma_k^2 A_t^{\gamma_k}(z,z'')\right]L(z,z')L(z',z'').$$

Let $z',z'',z\in\mathcal{X}$ such that $z\sim z'\sim z''$ and $d(z,z'')=2$. If $\lim_{\gamma_k\to 0}\gamma_k A_t^{\gamma_k}(z,z')\neq 0$, then by continuity of the function ρ on the set $(0,\infty)\times[0,\infty)$, Lemma 3.4 provides

$$\lim_{\gamma_k\to 0}\left[\rho\left(\gamma_k A_t^{\gamma_k}(z,z'),\gamma_k^2 A_t^{\gamma_k}(z,z'')\right)+2\gamma_k^2 A_t^{\gamma_k}(z,z'')\right]$$
$$=\rho\Big(A_t(z,z',y),\mathbb{A}_t(z,z'',y)\Big)+2\mathbb{A}_t(z,z'',y). \tag{33}$$

If $\lim_{\gamma_k\to 0}\gamma_k A_t^{\gamma_k}(z,z')=A_t(z,z',y)=0$, then according the definition of $A_t(z,z',y)$ in Lemma 3.4, one has, for any $w\in\mathcal{X}$ such that $(z,z')\in[y,w]$, $\widehat{\pi}^0_{\leftarrow}(w|y)=0$. Therefore, observing that if $(z,z'')\in[y,w]$ and $z\sim z'\sim z''$ with $d(z,z'')=2$ then $(z,z')\in[y,w]$, one also gets $A_t(z,z'',y)=0$ or equivalently $\lim_{\gamma_k\to 0}\left(\gamma_k^2 A_t^{\gamma_k}(z,z'')\right)=0$. By convexity arguments, for any $a,b,\gamma>0$

$$0\leq\rho\left(\gamma a,\gamma^2 b\right)+2\gamma^2 b\leq(\gamma a)^2+2\gamma^2 b.$$

It follows that

$$\lim_{\gamma_k\to 0}\left[\rho\Big(\gamma_k A_t^{\gamma_k}(z,z'),\gamma_k^2 A_t^{\gamma_k}(z,z'')\Big)+2\gamma_k^2 A_t^{\gamma_k}(z,z'')\right]=0=\rho(0,0).$$

Therefore, (33) holds in any cases and

$$\liminf_{\gamma_k\to 0} G_t^{\gamma_k}(z)$$
$$\geq \sum_{z',\,z'',\,z\sim z'\sim z'',\,d(z,z'')=2}\left[\rho\Big(A_t(z,z',y),\mathbb{A}_t(z,z'',y)\Big)+2\mathbb{A}_t(z,z'',y)\right]L(z,z')L(z',z'').$$

The proof of Proposition 3.5 is completed. □

3.2. **Application to specific examples of graphs.**

3.2.1. *The lattice $\mathbb{Z}^n$* .

Proof of Theorem 2.2. For any $z\in\mathbb{Z}^n$ and any $i\in[n]$, we note $\sigma_{i+}(z)=z+e_i$ and $\sigma_{i-}(z)=z-e_i$. One has $\sigma_{i+}\sigma_{i-}=id$ and for $j\neq i$, $\sigma_{i+}\sigma_{j+}=\sigma_{j+}\sigma_{i+}$, $\sigma_{i+}\sigma_{j-}=\sigma_{j-}\sigma_{i+}$, $\sigma_{i-}\sigma_{j-}=\sigma_{j-}\sigma_{i-}$. We note

$$A_{i+}(z,y):=A_t(z,\sigma_{i+}(z),y),\quad A_{i+j+}(z,y):=\mathbb{A}_t(z,\sigma_{i+}\sigma_{j+}(z),y),\qquad z\in\mathbb{Z}^n.$$

We define similarly $A_{i-}, A_{i-j-}, A_{i-j+}$. Applying Proposition 3.5, by symmetrisation one gets

$$
\begin{aligned}
&\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t)\\
&\geq \iint \Big(\sum_{i=1}^n (A_{i+}+A_{i-})\Big)^2 d\widehat{Q}^0_{t,1} + \iint \sum_{i=1}^n \rho\,(A_{i+}, A_{i+i+}) + \rho\,(A_{i-}, A_{i-i-})\; d\widehat{Q}^0_{t,1}\\
&+\frac{1}{2}\iint \sum_{i,j,i\neq j} \Big(\rho(A_{i+}, A_{j+i+}) + \rho(A_{j+}, A_{j+i+})\Big) + \Big(\rho(A_{i-}, A_{j-i-}) + \rho(A_{j-}, A_{j-i-})\Big)\\
&\qquad + \Big(\rho(A_{i+}, A_{j-i+}) + \rho(A_{j-}, A_{j-i+})\Big) + \Big(\rho(A_{i-}, A_{j+i-}) + \rho(A_{j+}, A_{j+i-})\Big)\; d\widehat{Q}^0_{t,1}
\end{aligned}
$$

Recall that $\rho(a,b)=0$ as soon as $a=0$ or $b=0$, and $\rho(a,b)=(\log b - 2\log a - 1)b$. Therefore, easy computations give for any $a,a'\geq 0$,

$$\inf_{b\geq 0} \rho(a,b) = -a^2, \tag{34}$$

and

$$\inf_{b\geq 0} \Big(\rho(a,b)+\rho(a',b)\Big) = -2aa'. \tag{35}$$

It follows that

$$
\begin{aligned}
\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) &\geq \iint \Big(\sum_{i=1}^n (A_{i+}+A_{i-})\Big)^2 d\widehat{Q}^0_{t,1} - \iint \sum_{i=1}^n \Big(A_{i+}^2 + A_{i-}^2\Big) d\widehat{Q}^0_{t,1}\\
&\quad - \iint \sum_{i,j,i\neq j} \Big(A_{i+}A_{j+} + A_{i-}A_{j-} + A_{i+}A_{j-} + A_{i-}A_{j+}\Big) d\widehat{Q}^0_{t,1}\\
&= 0
\end{aligned}
$$

Identically one may prove that $\liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t)\geq 0$. Applying then Lemma 3.1 ends the proof of Theorem 2.2.

□

3.2.2. *The complete graph.*

Proof of Theorem 2.3. Since for any $x,y\in X$, $d(x,y)=1$, Proposition 3.5 provides for any $t\in(0,1)$

$$\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \geq \iint \Big(\sum_{z',\,z'\sim z} A_t(z,z',y)\,L(z,z')\Big)^2 d\widehat{Q}^0_{t,1}(z,y).$$

From the expression (19) of $\nu_t^{0x,y}$, one easily check that for any $z,y\in\mathcal{X}$,

$$a_t(z,y) = (1-t)\,\overset{\frown}{\pi}{}^0_{\leftarrow}(z|y) + t\,\delta_y(z),$$

and for any $z'\in\mathcal{X}$ with $z'\sim z$,

$$\mathrm{a}_t(z,z',y) = \mathbb{1}_{z=y}\frac{\overset{\frown}{\pi}{}^0_{\leftarrow}(z'|y)}{\mu(z')}.$$

As a consequence one gets

$$
\begin{aligned}
&\iint \Big(\sum_{z',\,z'\sim z} A_t(z,z',y)L(z,z')\Big)^2 d\widehat{Q}^0_{t,1}(z,y) = \sum_{y\in\mathcal{X}}\Big(\sum_{z',z'\neq y}\frac{\mathrm{a}_t(y,z',y)}{a_t(y,y)}\mu(z')\Big)^2 a_t(y,y)\,\nu_1(y)\\
&= \sum_{y\in\mathcal{X}} \frac{\Big(1-\overset{\frown}{\pi}{}^0_{\leftarrow}(y|y)\Big)^2}{\overset{\frown}{\pi}{}^0_{\leftarrow}(y|y) + t\Big(1-\overset{\frown}{\pi}{}^0_{\leftarrow}(y|y)\Big)}\,\nu_1(y) = \varphi''_0(t),
\end{aligned}
$$

where we set

$$\varphi_0(t) := \int f\left(\widehat{\pi}^0_{\leftarrow}(y|y) + t\,(1-\widehat{\pi}^0_{\leftarrow}(y|y)\right)\,d\nu_1(y),$$

for any $t \in [0,1]$, with $f(s) := s\log s - s$, $s > 0$. One may similarly show that for any $t \in (0,1)$,

$$\liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t) \geq \psi''_0(t),$$

with $\psi_0(t) := \int f\left(\widehat{\pi}^0_{\rightarrow}(x|x) + (1-t)(1-\widehat{\pi}^0_{\rightarrow}(x|x)\right)\,d\nu_0(x)$. The proof of Theorem 2.3 ends by applying Lemma 3.1 and since

$$(1-t)\varphi_0(0) + t\varphi_0(1) - \varphi_0(t) = \frac{t(1-t)}{2}\int h_{1-t}\left(\int \mathbb{1}_{w\neq y} d\widehat{\pi}^0_{\leftarrow}(w|y)\right)d\nu_1(y),$$

and

$$(1-t)\psi_0(0) + t\psi_0(1) - \psi_0(t) = \frac{t(1-t)}{2}\int h_t\left(\int \mathbb{1}_{w\neq x} d\widehat{\pi}^0_{\rightarrow}(w|x)\right)d\nu_0(x).$$

□

3.2.3. *Product probability measures on the discrete hypercube.*

Proof of Theorem 2.4. According to Lemma 3.4, one has for any $i, j \in [n]$ with $i \neq j$, for any $y, z \in \{0,1\}^n$ such that $a_t(z,y) \neq 0$,

$$A_t(z,\sigma_i(z),y) := \frac{\mathrm{a}_t(z,\sigma_i(z),y)}{a_t(z,y)}, \quad \text{and} \quad \mathbb{A}_t(z,\sigma_j\sigma_i(z),y) := \frac{\mathbb{a}_t(z,\sigma_j\sigma_i(z),y)}{a_t(z,y)}$$

with

$$\mathrm{a}_t(z,\sigma_i(z),y) := \sum_{w,(z,\sigma_i(z))\in[y,w]} \frac{\mathbb{1}_{y_i\neq w_i}\mathbb{1}_{z_i=y_i}}{L_i(z_i,\overline{z_i})}\,(1-t)^{d(y,z)}t^{d(z,w)-1}\,\widehat{\pi}^0_{\leftarrow}(w|y),$$

and

$$\mathbb{a}_t(z,\sigma_j\sigma_i(z),y) := \sum_{w,(z,\sigma_i\sigma_j(z))\in[y,w]} \frac{\mathbb{1}_{y_i\neq w_i}\mathbb{1}_{z_i=y_i}}{L_i(z_i,\overline{z_i})}\,\frac{\mathbb{1}_{y_j\neq w_j}\mathbb{1}_{z_j=y_j}}{L_j(z_j,\overline{z_j})}\,(1-t)^{d(y,z)}t^{d(z,w)-2}\,\widehat{\pi}^0_{\leftarrow}(w|y).$$

For any $y \in \mathrm{supp}(\nu_1)$ and $z \in \{0,1\}^n$, let $I^{\leftarrow}(z,y)$ be the possibly empty set of indices such that $a_t(z,\sigma_i(z),y) \neq 0$,

$$\begin{aligned} I^{\leftarrow}(z,y) &:= \left\{ i\in[n] \,\middle|\, \exists\, v\in\{0,1\}^n, (z,\sigma_i(z))\in[y,w], \widehat{\pi}^0(v,y)>0 \right\} \\ &= \left\{ i\in[n] \,\middle|\, z_i = y_i, \text{ and } \exists\, v\in\{0,1\}^n, v_i\neq y_i, z\in[y,v], \widehat{\pi}^0(v,y)>0 \right\}. \end{aligned}$$

Since for any $i \neq j$, $\sigma_i\sigma_j = \sigma_j\sigma_i$, and observing that

$$L_j(z_j,\overline{z_j}) = L\left(\sigma_i(z),\sigma_j\sigma_i(z)\right) = L\left(z,\sigma_j(z)\right),$$

Proposition 3.5 provides after symmetrization,

$$\begin{aligned} \liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \geq &\iint \Big(\sum_{i=1}^n A_t(z,\sigma_i(z),y)\,L_i(z_i,\overline{z_i}))\Big)^2\,d\widehat{Q}^0_{t,1}(z,y) \\ &+ \frac{1}{2}\iint\sum_{(i,j),\,i\neq j} \Big[\rho\Big(A_t(z,\sigma_i(z),y),\mathbb{A}_t(z,\sigma_j\sigma_i(z),y)\Big) + \rho\Big(A_t(z,\sigma_i(z),y),\mathbb{A}_t(z,\sigma_j\sigma_i(z),y)\Big)\Big] \\ &\qquad L_j(z_j,\overline{z_j})\,L_i(z_i,\overline{z_i})\,d\widehat{Q}^0_{t,1}(z,y). \end{aligned} \tag{36}$$

By applying identity (35), one gets

$$\begin{aligned}\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t) &\geq \iint \sum_{i=1}^n \Big(A_t(z,\sigma_i(z),y)\,L_i(z_i,\overline{z_i})\Big)^2\,d\widehat{Q}^0_{t,1}(z,y)\\ &= \iint \sum_{i\in I^{\leftarrow}(z,y)} \Big(A_t(z,\sigma_i(z),y)\,L_i(z_i,\overline{z_i})\Big)^2\,d\widehat{Q}^0_{t,1}(z,y)\\ &= \sum_{i=1}^n \int \sum_{z\in\{0,1\}^n} \mathbb{1}_{E_i^{\leftarrow}(y)}(z)\Big(A_t(z,\sigma_i(z),y)\,L_i(z_i,\overline{z_i})\Big)^2 a_t(z,y)\,d\nu_1(y),\end{aligned}$$

where in the last inequality, for any $y\in\mathrm{supp}(\nu_1)$, the set $E_i^{\leftarrow}(y)$ is defined by

$$\begin{aligned}E_i^{\leftarrow}(y) &:= \Big\{z\in\{0,1\}^n \Big|\, i\in I^{\leftarrow}(z,y)\Big\}\\ &= \Big\{z\in\{0,1\}^n \Big|\, z_i=y_i,\ \text{and } \exists\, v\in\{0,1\}^n, v_i\neq y_i, z\in[y,v], \widehat{\pi}^0(v,y)>0\Big\}.\end{aligned}$$

From the definition of $a_t(z,y)$, one has

$$\sum_{z\in\{0,1\}^n} \mathbb{1}_{E_i^{\leftarrow}(y)}(z)\,a_t(z,y) = \frac{\widehat{Q}^0_{t,1}\Big(E_i^{\leftarrow}(y)\times\{y\}\Big)}{\nu_1(y)},$$

and simple computations give

$$\sum_{z\in\{0,1\}^n} A_t(z,\sigma_i(z),y)\,L_i(z_i,\overline{z_i})\,a_t(z,y) = \sum_{z\in\{0,1\}^n} \mathrm{a}_t(z,\sigma_i(z),y)\,L_i(z_i,\overline{z_i}) = \Pi^i_{\leftarrow}(y).$$

Therefore Cauchy-Schwarz inequality provides

$$\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t) \geq \sum_{i=1}^n\sum_{y\in\{0,1\}^n} \frac{\Pi^i_{\leftarrow}(y)^2\,\nu_1(y)^2}{\widehat{Q}^0_{t,1}\Big(E_i^{\leftarrow}(y)\times\{y\}\Big)}. \tag{37}$$

At this level, a first lower bound is obtained using the fact that

$$\frac{\widehat{Q}^0_{t,1}\Big(E_i^{\leftarrow}(y)\times\{y\}\Big)}{\nu_1(y)} \leq \sum_{z\in\{0,1\}^n} \mathbb{1}_{z_i=y_i} a_t(z,y) = 1-\Pi^i_{\leftarrow}(y)+t\,\Pi^i_{\leftarrow}(y).$$

This inequality implies (as in the last section) for any $t\in(0,1)$

$$\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t)\geq \varphi''_0(t), \tag{38}$$

where

$$\varphi_0(t) := \int\sum_{i=1}^n f\Big(1-\Pi^i_{\leftarrow}(y)+t\,\Pi^i_{\leftarrow}(y)\Big)\,d\nu_1(y),$$

with $f(s):=s\log s - s$. One may identically show that that

$$\liminf_{\gamma_k\to 0}\psi''_{\gamma_k}(t)\geq \psi''_0(t), \tag{39}$$

where

$$\psi_0(t) := \int\sum_{i=1}^n f\Big(1-\Pi^i_{\rightarrow}(x)+(1-t)\,\Pi^i_{\rightarrow}(x)\Big)d\nu_0(x).$$

A second lower bound can be reached from (37) applying again Cauchy-Schwarz inequality. Setting

$$\alpha_i(t) := \widehat{Q}^0_{t,1}\Big(\bigcup_{y\in\mathrm{supp}(\nu_1)}\Big(E_i^{\leftarrow}(y)\times\{y\}\Big)\Big),$$

one gets, for any $t \in (0,1)$

$$\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \geq \sum_{i=1}^n \frac{1}{\alpha_i(t)} \left(\int \Pi^i_{\leftarrow}(y)\, d\nu_1(y) \right)^2 = \sum_{i=1}^n \frac{(\Pi^i)^2}{\alpha_i(t)},$$

where $\Pi^i := \widehat{\pi}_0\left(\{(x,y) \in \{0,1\}^n \,|\, x_i \neq y_i\}\right)$. By symmetry, one may identically show that for any $t \in (0,1)$

$$\liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t) \geq \sum_{i=1}^n \frac{(\Pi^i)^2}{\beta_i(t)},$$

with

$$\beta_i(t) := \widehat{Q}^0_{0t}\Big(\bigcup_{x\in \mathrm{supp}(\nu_0)} \Big(\{x\} \times E_i^{\rightarrow}(x) \Big) \Big),$$

and

$$E_i^{\rightarrow}(x) := \Big\{ z \in \{0,1\}^n \Big| z_i = x_i, \text{ and } \exists\, w \in \{0,1\}^n, w_i \neq x_i, z \in [x,w], \widehat{\pi}^0(x,w) > 0 \Big\}.$$

Observe that for any $x \in \mathrm{supp}(\nu_1), y \in \mathrm{supp}(\nu_0)$, one has

$$E_i^{\rightarrow}(x) \cap E_i^{\leftarrow}(y) = \emptyset.$$

Indeed if $z \in E_i^{\rightarrow}(x) \cap E_i^{\leftarrow}(y)$ then there exist v and w in $\{0,1\}^n$ such that $z \in [x,w]$, $z \in [v,y]$ and $\widehat{\pi}^0(x,w) > 0, \widehat{\pi}^0(v,y) > 0$. According to Lemma 4.2 (i), it follows that $z \in [v,w]$. But since $v_i = w_i$ and $z_i \neq v_i$, this leads to a contradiction. As a consequence, one gets

$$\alpha_i(t) + \beta_i(t) = \sum_{x,y\in\{0,1\}^n} \sum_{z \in E_i^{\rightarrow}(x)\cup E_i^{\leftarrow}(y)} \nu_t^{0\,x,y}(z)\widehat{\pi}^0(x,y) \leq 1.$$

Since $\min_{\alpha,\beta>0,\alpha+\beta\leq 1} \left\{ \frac{1}{\alpha} + \frac{1}{\beta} \right\} = 4$, this property implies

$$\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) + \liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t) \geq 4 \sum_{i=1}^n (\Pi^i)^2. \tag{40}$$

Let us now suggest another type of lower bound for $\liminf_{\gamma\to 0} \varphi''_\gamma(t)$ starting again from (36). For that purpose, we define

$$\begin{aligned}
\mathbb{I}^{\leftarrow}(z,y) &:= \Big\{ (i,j) \in [n]\times[n] \Big| \exists\, v \in \{0,1\}^n, (z, \sigma_j\sigma_i(z)) \in [y,v], \widehat{\pi}^0(v,y) > 0 \Big\} \\
&= \Big\{ (i,j) \in [n]\times[n] \Big| z_i = y_i, z_j = y_j, \exists\, v \in \{0,1\}^n, v_i \neq y_i, v_j \neq y_j z \in [y,v], \widehat{\pi}^0(v,y) > 0 \Big\},
\end{aligned}$$

By symmetry, since $\sigma_j\sigma_i = \sigma_i\sigma_j$, if $(i,j) \in \mathbb{I}^{\leftarrow}(z,y)$ then $(j,i) \in \mathbb{I}^{\leftarrow}(z,y)$. We also note

$$\mathbb{I}_1^{\leftarrow}(z,y) := \Big\{ i \in [n] \Big| \exists\, j \in [n], (i,j) \in \mathbb{I}^{\leftarrow}(z,y) \Big\} = \Big\{ i \in [n] \Big| \exists\, j \in [n], (j,i) \in \mathbb{I}^{\leftarrow}(z,y) \Big\},$$

and given $i \in \mathbb{I}_1^{\leftarrow}(z,y)$, we note

$$\mathbb{I}_{2,i}^{\leftarrow}(z,y) := \Big\{ j \in [n] \Big| (i,j) \in \mathbb{I}^{\leftarrow}(z,y) \Big\}.$$

One may observe that for any y,z, $\mathbb{I}_1^{\leftarrow}(z,y) \subset I^{\leftarrow}(z,y)$ and for any $i \in I^{\leftarrow}(z,y)$, $\mathbb{I}_{2,i}^{\leftarrow}(z,y) \subset I^{\leftarrow}(z,y) \setminus \{i\}$.

To simplify the notations, let $L_i(z) := L_i(z_i, \overline{z_i})$, $A_i(z,y) := A_t(z, \sigma_i(z), y)$ and $A_{ij}(z,y) = \mathbb{A}_t(z, \sigma_j\sigma_i(z), y)$. Observing that $\Big(\rho(A_i, A_{ij}) + \rho(A_j, A_{ij})\Big) L_iL_j = \rho(A_iL_i, A_{ij}L_iL_j) + \rho(A_jL_j, A_{ji}L_iL_j)$, one gets that (36) is equivalent to

$$\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \geq \iint \Big[\Big(\sum_{i\in I^{\leftarrow}} A_i\, L_i \Big)^2 + \sum_{(i,j)\in \mathbb{I}^{\leftarrow}} \rho\Big(A_iL_i, A_{ij}L_iL_j \Big) \Big] d\widehat{Q}^0_{t,1}. \tag{41}$$

The idea is now to minimize the expression inside the integral in the right-hand side over all A_iL_i, $i \in I^{\leftarrow}$. For any fixed $\beta_{ij} := A_{ij}L_iL_j$, $(i,j) \in \mathbb{I}^{\leftarrow}$, and let

$$F((\beta_i)_{i\in I^{\leftarrow}}) := \Big(\sum_{i\in I^{\leftarrow}} \beta_i\Big)^2 + \sum_{(i,j)\in \mathbb{I}^{\leftarrow}} \rho\Big(\beta_i, \beta_{ij}\Big), \qquad \beta_i > 0, i \in I^{\leftarrow}.$$

Since $\mathbb{I}_1^{\leftarrow} \subset I^{\leftarrow}$, one has

$$\inf_{\beta_i>0, i\in I^{\leftarrow}\setminus \mathbb{I}_1^{\leftarrow}} F((\beta_i)_{i\in I}) = \Big(\sum_{i\in \mathbb{I}_1^{\leftarrow}} \beta_i\Big)^2 + \sum_{i\in \mathbb{I}_1^{\leftarrow}} \sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} \rho\Big(\beta_i, \beta_{ij}\Big).$$

Observe that if $\mathbb{I}_1^{\leftarrow} = \emptyset$ then $\inf_{\beta_i>0, i\in I^{\leftarrow}} F((\beta_i)_{i\in I^{\leftarrow}}) = 0$. We assume now that $\mathbb{I}_1^{\leftarrow} \neq \emptyset$. The function of $(\beta_i)_{i\in \mathbb{I}_1{}^{\leftarrow}} \in (\mathbb{R}_+^*)^{|\mathbb{I}_1^{\leftarrow}|}$ on the right-hand side is convex. By differentiating, its minimum value is reached at the point $(\beta_i)_{i\in \mathbb{I}_1^{\leftarrow}}$ satisfying for all $i \in \mathbb{I}_1^{\leftarrow}$,

$$2 \sum_{i'\in \mathbb{I}_1^{\leftarrow}} \beta_{i'} - 2 \sum_{j\in \in \mathbb{I}_{2i}^{\leftarrow}} \frac{\beta_{ij}}{\beta_i} = 0.$$

Therefore, one has $\beta_i = \frac{\sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} \beta_{ij}}{\sum_{i'\in \mathbb{I}_1^{\leftarrow}} \beta_{i'}}$. Summing the last equality over all $i \in \mathbb{I}_1^{\leftarrow}$, one gets

$$\Big(\sum_{i'\in \mathbb{I}_1^{\leftarrow}} \beta_{i'}\Big)^2 = \sum_{(i,j)\in \mathbb{I}^{\leftarrow}} \beta_{ij} := S,$$

and it follows that $\beta_i = \frac{\sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} \beta_{ij}}{\sqrt{S}}$. Finally, setting $S_i = \sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} \beta_{j,i}$, one gets

$$\begin{aligned}
\inf_{\beta_i>0, i\in I^{\leftarrow}} F((\beta_i)_{i\in I^{\leftarrow}}) &= \frac{1}{S}\Big(\sum_{i\in \mathbb{I}_1^{\leftarrow}} S_i\Big)^2 + \sum_{(i,j)\in \mathbb{I}^{\leftarrow}} \beta_{ij} \left(\log \beta_{ij} - 2\log \frac{S_i}{\sqrt{S}}\right) - \beta_{ij} \\
&= \sum_{(i,j)\in \mathbb{I}^{\leftarrow}} \beta_{ij} \Big(\log(\beta_{ij} S) - \log(S_i S_j)\Big) \\
&= \sum_{i\in \mathbb{I}_1^{\leftarrow}} \sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} \beta_{ij} \log \frac{\beta_{ij} S}{S_i S_j}
\end{aligned}$$

By convexity of the function $H : t \mapsto t\log t$, applying Jensen inequality, one gets

$$\begin{aligned}
&\inf_{\beta_i>0, i\in I^{\leftarrow}} F((\beta_i)_{i\in I^{\leftarrow}}) = \frac{1}{S} \sum_{i\in \mathbb{I}_1^{\leftarrow}} S_i \sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} H\left(\frac{\beta_{ij} S}{S_i S_j}\right) S_j \\
&\geq \frac{1}{S} \sum_{i\in \mathbb{I}_1^{\leftarrow}} S_i \Big(\sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} S_j\Big) H\Big(\sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} \frac{\beta_{ij} S}{S_i \sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} S_j}\Big) = \sum_{i\in \mathbb{I}_1^{\leftarrow}} S_i \log\Big(\frac{S}{\sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} S_j}\Big) \\
&\geq -\sum_{i\in \mathbb{I}_1^{\leftarrow}} S_i \log\left(1 - \frac{S_i}{S}\right)
\end{aligned}$$

where the last inequality holds since $\sum_{j\in \mathbb{I}_{2i}^{\leftarrow}} S_j \leq S - S_i$. Applying Jensen's inequality with the convex increasing function $s \in (0,1) \mapsto -\log(1-s)$ and using Cauchy-Schwarz inequality, one gets

$$-\sum_{i\in \mathbb{I}_1^{\leftarrow}} S_i \log\left(1 - \frac{S_i}{S}\right) \geq -S \log\left(1 - \frac{\sum_{i\in \mathbb{I}_1^{\leftarrow}} S_i^2}{S^2}\right) \geq -S \log\left(1 - \frac{1}{|\mathbb{I}_1^{\leftarrow}|}\right),$$

and therefore

$$\inf_{\beta_i>0, i\in I^{\leftarrow}} F((\beta_i)_{i\in I^{\leftarrow}}) \geq C_n S,$$

with $C_n = -\log\left(1-\frac{1}{n}\right)$. Observe that this inequality also holds for $I^{\leftarrow} = \emptyset$ since $S = 0$ in that case. Finally, (41) provides,

$$\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \geq C_n \iint S \, d\widehat{Q}^0_{t,1}.$$

Simple computations give for any $(i,j), i \neq j$,

$$\sum_{z\in\{0,1\}^n} \mathbb{A}_t(z, \sigma_j\sigma_i(z), y)\, L_i(z_i, \overline{z_i})\, L_j(z_j, \overline{z_j})\, a_t(z,y) = \int \mathbb{1}_{y_i\neq w_i}\mathbb{1}_{y_j\neq w_j}\, d\widehat{\pi}^0_{\leftarrow}(w|y),$$

and therefore

$$\iint S \, d\widehat{Q}^0_{t1} = \iint \sum_{(i,j),i\neq j} \mathbb{1}_{y_i\neq w_i}\mathbb{1}_{y_j\neq w_j}\, d\widehat{\pi}^0(w,y) = \iint d(y,w)(d(y,w)-1)\, d\widehat{\pi}^0(w,y).$$

Identically, one may prove that

$$\liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t) \geq C_n \iint d(x,w)(d(x,w)-1)\, d\widehat{\pi}^0(x,w).$$

The proof of Theorem 2.4 ends as the proof of Theorem 2.3, by applying Lemma 3.1 using the last estimates together with (38), (39) and (40). □

3.2.4. *The circle $\mathbb{Z}/N\mathbb{Z}$.*

Proof of Theorem 2.5. Let us note $n' = \lceil N/2 \rceil$ where $\lceil\cdot\rceil$ denotes the ceiling function. Let $y, z \in \mathbb{Z}/N\mathbb{Z}$. We observe that if $\{w \in \mathbb{Z}/N\mathbb{Z} \,|\, (z, z-1) \in [y, n]\} \neq \emptyset$ then necessarily $(z-1, z) \in [y+n', y]$ and if $\{w \in \mathbb{Z}/N\mathbb{Z} \,|\, (z, z+1) \in [y, n]\} \neq \emptyset$ then necessarily $(z, z+1) \in [y, y+n]$. As a consequence, since the sets $\{z \in \mathbb{Z}/N\mathbb{Z} \,|\, (z, z+1) \in [y, y+n]\}$ and $\{z \in \mathbb{Z}/N\mathbb{Z} \,|\, (z-1, z) \in [y+n', y]\}$ are disjoints, according to the definition of $\mathrm{a}_t(z, z+1, y)$ and $\mathrm{a}_t(z, z+1, y)$ in Lemma 3.4, the two following sets $\{z \in \mathbb{Z}/N\mathbb{Z} \,|\, \mathrm{a}_t(z, z+1, y)\}$ and $\{z \in \mathbb{Z}/N\mathbb{Z} \,|\, \mathrm{a}_t(z, z-1, y)\}$ are also disjoints. It follows that

$$\begin{aligned}\iint \Big(A_t(z, z+1, y) + A_t(z, z-1, y)\Big)^2 d\widehat{Q}^0_{t,1}(z,y) \\ = \iint \Big(A_t^2(z, z+1, y) + A_t^2(z, z-1, y)\Big)\, d\widehat{Q}^0_{t,1}(z,y).\end{aligned}$$

Therefore Proposition 3.5 together with (34) provide

$$\begin{aligned}&\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \\ &\geq \iint \Big(A_t^2(z, z+1, y) + A_t^2(z, z-1, y)\Big) + \rho\Big(A_t(z, z+1, y), \mathbb{A}_t(z, z+2, y)\Big) \\ &\qquad + \rho\Big(A_t(z, z-1, y), \mathbb{A}_t(z, z-2, y)\Big)\, d\widehat{Q}^0_{t,1}(z,y) \\ &\geq 0\end{aligned}$$

Identically one proves that $\liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t) \geq 0$. The proof of Theorem 2.5 ends by applying Lemma 3.1. □

3.2.5. *The Bernoulli-Laplace model.*

Proof of Theorem 2.6. According to Lemma 3.4, one has for any $y \in \mathrm{supp}(\nu_1), z \in \mathcal{X}_k$ such that $a_t(z,y) \neq 0$, for any $i, k \in J_0(z)$ and any $j, l \in J_1(z)$ with $i \neq k$ and $j \neq l$

$$A_t(z, \sigma_{ij}(z), y) := \frac{\mathrm{a}_t(z, \sigma_{ij}(z), y)}{a_t(z,y)}, \quad \text{and} \quad \mathbb{A}_t(z, \sigma_{kl}\sigma_{ij}(z), y) := \frac{\mathbb{a}_t(z, \sigma_{kl}\sigma_{ij}(z), y)}{a_t(z,y)},$$

with

$$\mathrm{a}_t(z,\sigma_{ij}(z),y) \quad := \sum_{w\in\mathcal{X}_k,\,(z,\sigma_{ij}(z))\in[y,w]} r(y,z,\sigma_{ij}(z),w)\,d(y,w)\,\rho_t^{d(y,w)-1}(d(z,w) \quad - \quad 1)\,\widehat{\pi}^0_{\leftarrow}(w|y),$$

and

$$\mathbb{a}_t(z,\sigma_{kl}\sigma_{ij}(z),y)$$
$$:= \sum_{w\in\mathcal{X}_k,\,(z,\sigma_{ij}\sigma_{kl}(z))\in[y,w]} r(y,z,\sigma_{kl}\sigma_{ij}(z),w)\,d(y,w)(d(y,w)-1)\,\rho_t^{d(y,w)-2}(d(z,w)-2)\,\widehat{\pi}^0_{\leftarrow}(w|y).$$

To simplify the notations, let us note $A_t(z,\sigma_{ij}(z),y) = A_{ij}(z,y)$ and $\mathbb{A}_t(z,\sigma_{kl}\sigma_{ij}(z),y) = A_{kl,ij}(z,y)$. Observe that $\sigma_{ki}\sigma_{ij}(z) = \sigma_{kj}(z)$ so that $d(z,\sigma_{ki}\sigma_{ij}(z)) = 1$ and similarly $d(z,\sigma_{jl}\sigma_{ij}(z)) = 1$. As a consequence, Proposition 3.5 provides

$$\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t) \geq \iint \Big(\sum_{(i,j)\in J_0(z)\times J_1(z)} A_{ij}(z,y)\Big)^2\, d\widehat{Q}^0_{t,1}(z,y)$$
$$+ \iint \sum_{(i,j),(k,l)\in J_0(z)\times J_1(z),i\neq k,j\neq l} \rho(A_{ij}(z,y),A_{kl,ij}(z,y))\,d\widehat{Q}^0_{t,1}(z,y).$$

For $y\in\mathrm{supp}(\nu_1)$ and $z\in\mathcal{X}_k$, let us define

$$\begin{aligned}
I^{\leftarrow}(z,y) &:= \Big\{(i,j)\in J_0\times J_1 \,\Big|\, A_{ij}(z,y)>0\Big\}\\
&= \Big\{(i,j)\in J_0\times J_1 \,\Big|\, z_i=y_i=0, z_j=y_j=1, \exists v\in\mathcal{X}_k, v_i=1, v_j=0, z\in[y,v], \widehat{\pi}^0(v,y)>0\Big\},\\
\mathbb{I}^{\leftarrow}(z,y) &:= \Big\{((i,j),(k,l))\in(J_0(z)\times J_1(z))^2 \,\Big|\, i\neq k, j\neq l, A_{kl,ij}(z,y)>0\Big\},\\
\mathbb{I}^{\leftarrow}_1(z,y) &:= \Big\{(i,j)\in J_0(z)\times J_1(z) \,\Big|\, \exists (k,l)\in J_0(z)\times J_1(z), ((i,j),(k,l))\in\mathbb{I}^{\leftarrow}(z,y)\Big\},
\end{aligned}$$

and for $(i,j)\in\mathbb{I}^{\leftarrow}_1(z,y)$,

$$\mathbb{I}^{\leftarrow}_{2,ij}(z,y) := \Big\{(k,l)\in J_0(z)\times J_1(z) \,\Big|\, ((i,j),(k,l))\in\mathbb{I}^{\leftarrow}(z,y)\Big\}.$$

If the indices k,l,i,j all differ, then $\sigma_{kl}\sigma_{ij}(z) = \sigma_{ij}\sigma_{kl}(z)$, and therefore $A_{kl,ij}(z,y) = A_{ij,kl}(z,y)$ and $((i,j),(k,l))\in\mathbb{I}^{\leftarrow}(z,y)$ implies $((k,l),(i,j))\in\mathbb{I}^{\leftarrow}(z,y)$. Moreover, one may easily check that $\mathbb{I}^{\leftarrow}_1(z,y)\subset I^{\leftarrow}(z,y)$. As a consequence, by symmetrisation it follows

$$\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t) \geq \iint \Big(\sum_{(i,j)\in I^{\leftarrow}} A_{ij}\Big)^2\, d\widehat{Q}^0_{t,1}$$
$$+ \frac{1}{2}\iint \sum_{((i,j),(k,l))\in\mathbb{I}^{\leftarrow}} \Big(\rho(A_{ij},A_{kl,ij}) + \rho(A_{kl},A_{kl,ij}\Big)\, d\widehat{Q}^0_{t,1}. \tag{42}$$

Let us compute a first lower bound of the right hand side of this inequality. Applying identity (35) yields

$$
\begin{aligned}
\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t) &\geq \iint \Big[\Big(\sum_{(i,j)\in I^{\leftarrow}} A_{ij}\Big)^2 - \sum_{((i,j),(k,l))\in \mathbb{I}^{\leftarrow}} A_{ij}A_{kl}\Big]\, d\widehat{Q}^0_{t,1}\\
&\geq \iint \Big[\sum_{i\in J_0}\Big(\sum_{j\in J_1} A_{ij}\Big)^2 + \sum_{j\in J_1}\Big(\sum_{i\in J_0} A_{ij}\Big)^2 - \sum_{(i,j)\in J_0\times J_1} A_{ij}^2\Big]\, d\widehat{Q}^0_{t,1}\\
&\geq \max\Big[\iint \sum_{i\in J_0}\Big(\sum_{j\in J_1} A_{ij}\Big)^2 d\widehat{Q}^0_{t1}, \iint \sum_{j\in J_1}\Big(\sum_{i\in J_0} A_{ij}\Big)^2 d\widehat{Q}^0_{t,1}\Big]\\
&= \max\Big[\iint \sum_{z\in\mathcal{X}_k}\sum_{i\in[n]}\Big(\sum_{j\in[n]} A_{ij}(z,y)\mathbb{1}_{(i,j)\in I^{\leftarrow}(z,y)}\Big)^2 d\widehat{Q}^0_{t,1}(z,y),\\
&\qquad\qquad \iint \sum_{z\in\mathcal{X}_k}\sum_{j\in[n]}\Big(\sum_{i\in[n]} A_{ij}(z,y)\mathbb{1}_{(i,j)\in I^{\leftarrow}(z,y)}\Big)^2 d\widehat{Q}^0_{t,1}(z,y)\Big].
\end{aligned}
$$

We will now bound from below the right hand side of this inequality using Cauchy-Schwarz inequality. For any $y\in \operatorname{supp}(\nu_1)$, and $i\in J_0(y)$ we note

$$
E^{\leftarrow}_{i,0}(y) := \Big\{z\in\mathcal{X}_k \,\Big|\, \exists j\in J_1(y), \exists\, v\in\mathcal{X}_k, (z,\sigma_{ij}(z))\in[y,v], \widehat{\pi}^0(v,y)>0\Big\},
$$

and for $j\in J_1(y)$

$$
E^{\leftarrow}_{j,1}(y) := \Big\{z\in\mathcal{X}_k \,\Big|\, \exists i\in J_0(y), \exists\, v\in\mathcal{X}_k, (z,\sigma_{ij}(z))\in[y,v], \widehat{\pi}^0(v,y)>0\Big\}.
$$

Since $(i,j)\in I(z,y)$ implies $z\in E^{\leftarrow}_{i,0}(y)$ and $z\in E^{\leftarrow}_{j,1}(y)$, one has

$$
\begin{aligned}
&\iint \sum_{z\in\mathcal{X}_k}\sum_{i\in[n]}\Big(\sum_{j\in[n]} A_{ij}(z,y)\mathbb{1}_{(i,j)\in I(z,y)}\Big)^2 d\widehat{Q}^0_{t,1}(z,y)\\
&\qquad\qquad = \int \sum_{i\in J_0(y)}\sum_{z\in E^{\leftarrow}_{i,0}(y)}\Big(\sum_{j\in J_1(y)} A_{ij}(z,y)\Big)^2 a_t(z,y)\, d\nu_1(y),
\end{aligned}
$$

and therefore by Cauchy-Schwarz inequality,

$$
\begin{aligned}
&\iint \sum_{z\in\mathcal{X}_k}\sum_{i\in[n]}\Big(\sum_{j\in[n]} A_{ij}(z,y)\mathbb{1}_{(i,j)\in I^{\leftarrow}(z,y)}\Big)^2 d\widehat{Q}^0_{t,1}(z,y)\\
&\qquad\qquad \geq \int \sum_{i\in J_0(y)} \frac{\Big(\sum_{j\in J_1(y)}\sum_{z\in E^{\leftarrow}_{i,0}(y)} A_{ij}(z,y)a_t(z,y)\Big)^2}{\sum_{z\in E^{\leftarrow}_{i,0}(y)} a_t(z,y)}\, d\nu_1(y).
\end{aligned}
$$

For $(i,j)\in J_0(y)\times J_1(y)$, one may compute the quantity $\sum_{z\in E^{\leftarrow}_{i,0}(y)} A_{ij}(z,y)a_t(z,y)$ using the two following observations. First $(z,\sigma_{ij}(z))\in[y,w]$ holds if and only if one has $y_i=z_i=w_j=0$, $y_j=z_j=w_i=1$ and $z\in[y,\sigma_{ij}(w)]$. Secondly, the generator L is translation invariant which implies that $r(y,z,\sigma_{ij}(z),w)=$

$r(y,z,z,\sigma_{ij}(w))\,\frac{L^{d(y,\sigma_{ij}(w))}(y,\sigma_{ij}(w))}{L^{d(y,w)}(y,w)}$. Therefore, one gets for any $(i,j)\in J_0(y)\times J_1(y)$,

$$\begin{aligned}
\sum_{z\in E^{\leftarrow}_{i,0}(y)} A_{ij}(z,y)\,a_t(z,y) &= \sum_{z\in\mathcal{X}_k} A_{ij}(z,y)\,a_t(z,y)\\
&= \sum_{w\in\mathcal{X}_k} \mathbb{1}_{y_i=w_j=0}\mathbb{1}_{y_j=w_i=1}\sum_{s=0}^{d(y,\sigma_{ij}(w))}\ \sum_{z\in[y,\sigma_{ij}(w)],d(y,z)=s} r(y,z,z,\sigma_{ij}(w))\frac{L^{d(y,\sigma_{ij}(w))}(y,\sigma_{ij}(w))}{L^{d(y,w)}(y,w)}\\
&\qquad\qquad d(y,w)\rho_t^{d(y,w)-1}(d(y,w)-1-s)\,\widehat{\pi}^0_{\leftarrow}(w|y)\\
&= \sum_{w\in\mathcal{X}_k} \mathbb{1}_{y_i=w_j=0}\mathbb{1}_{y_j=w_i=1}\frac{L^{d(y,\sigma_{ij}(w))}(y,\sigma_{ij}(w))}{L^{d(y,w)}(y,w)}\,d(y,w)\,\widehat{\pi}^0_{\leftarrow}(w|y)\\
&= \sum_{w\in\mathcal{X}_k} \frac{\mathbb{1}_{y_i=w_j=0}\mathbb{1}_{y_j=w_i=1}}{d(y,w)}\,\widehat{\pi}^0_{\leftarrow}(w|y).
\end{aligned}$$

Since for $i\in J_0(y)$, $\displaystyle\sum_{j\in J_1(y)} \mathbb{1}_{y_i=w_j=0}\mathbb{1}_{y_j=w_i=1} = d(y,w)\,\mathbb{1}_{w_i\neq y_i}$, it follows that

$$\sum_{j\in J_1(y)}\sum_{z\in E^{\leftarrow}_{i,0}(y)} A_{ij}(z,y)\,a_t(z,y) = \sum_{w\in\mathcal{X}_k} \mathbb{1}_{w_i\neq y_i}\,\widehat{\pi}^0_{\leftarrow}(w|y),$$

and therefore

$$\begin{aligned}
&\iint \sum_{z\in\mathcal{X}_k}\sum_{i\in[n]}\Big(\sum_{j\in[n]} A_{ij}(z,y)\mathbb{1}_{(i,j)\in I^{\leftarrow}(z,y)}\Big)^2\, d\widehat{Q}^0_{t1}(z,y)\\
&\qquad\qquad \geq \int\sum_{i\in J_0(y)} \frac{\Big(\sum_{w\in\mathcal{X}_k}\mathbb{1}_{w_i\neq y_i}\,\widehat{\pi}^0_{\leftarrow}(w|y)\Big)^2}{\sum_{z\in E^{\leftarrow}_{i,0}(y)} a_t(z,y)}\,d\nu_1(y).
\end{aligned}$$

With same computations, by exchanging the role of i and j, (35) finally implies

$$\begin{aligned}
\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t) \geq \max\Big[&\int\sum_{i\in J_0(y)} \frac{\Big(\sum_{w\in\mathcal{X}_k}\mathbb{1}_{w_i\neq y_i}\,\widehat{\pi}^0_{\leftarrow}(w|y)\Big)^2}{\sum_{z\in E^{\leftarrow}_{i,0}(y)} a_t(z,y)}d\nu_1(y),\\
&\int\sum_{j\in J_1(y)} \frac{\Big(\sum_{w\in\mathcal{X}_k}\mathbb{1}_{w_j\neq y_j}\,\widehat{\pi}^0_{\leftarrow}(w|y)\Big)^2}{\sum_{z\in E^{\leftarrow}_{j,1}(y)} a_t(z,y)}d\nu_1(y)\Big]
\end{aligned}\tag{43}$$

Working on $\psi''_{\gamma_k}(t)$, on may identically show that

$$\begin{aligned}
\liminf_{\gamma_k\to 0}\psi''_{\gamma_k}(t) \geq \max\Big[&\int\sum_{i\in I_0(x)} \frac{\Big(\sum_{w\in\mathcal{X}_k}\mathbb{1}_{w_i\neq x_i}\,\widehat{\pi}^0_{\rightarrow}(w|x)\Big)^2}{\sum_{z\in E^{\rightarrow}_{i,0}(x)} b_t(z,x)}d\nu_0(x),\\
&\int\sum_{j\in I_1(x)} \frac{\Big(\sum_{w\in\mathcal{X}_k}\mathbb{1}_{w_j\neq x_j}\,\widehat{\pi}^0_{\rightarrow}(w|x)\Big)^2}{\sum_{z\in E^{\rightarrow}_{j,1}(x)} b_t(z,x)}d\nu_0(x)\Big]
\end{aligned}\tag{44}$$

where for any $i\in J_0(x)$ we note

$$E^{\rightarrow}_{i,0}(x) := \Big\{z\in\mathcal{X}_k\,\Big|\,\exists j\in I_1(x), \exists\, w\in\mathcal{X}_k, (z,\sigma_{ij}(z))\in[x,w], \widehat{\pi}^0(x,w)>0\Big\},$$

and for any $j\in J_1(x)$

$$E^{\rightarrow}_{j,1}(x) := \Big\{z\in\mathcal{X}_k\,\Big|\,\exists i\in I_0(x), \exists\, w\in\mathcal{X}_k, (z,\sigma_{ij}(z))\in[y,v], \widehat{\pi}^0(x,w)>0\Big\}.$$

From this two estimates, we will derive two different lower-bounds. A first strategy is to apply again Cauchy-Schwarz inequality, (43) and (44) implies

$$
\begin{aligned}
&\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t)+\liminf_{\gamma_k\to 0}\psi''_{\gamma_k}(t)\\
&\geq \max\Bigg[\sum_{i\in[n]}\Big[\frac{1}{\alpha_{i,0}(t)}\left(\iint \mathbb{1}_{x_i\neq y_i}\mathbb{1}_{y_i=0}\,d\widehat{\pi}^0(x,y)\right)^2+\frac{1}{\beta_{i,0}(t)}\left(\iint \mathbb{1}_{x_i\neq y_i}\mathbb{1}_{x_i=0}\,d\widehat{\pi}^0(x,y)\right)^2\Big],\\
&\qquad\sum_{j\in[n]}\Big[\frac{1}{\alpha_{i,1}(t)}\left(\iint \mathbb{1}_{x_i\neq y_i}\mathbb{1}_{y_i=1}\,d\widehat{\pi}^0(x,y)\right)^2+\frac{1}{\beta_{i,1}(t)}\left(\iint \mathbb{1}_{x_i\neq y_i}\mathbb{1}_{x_i=1}\,d\widehat{\pi}^0(x,y)\right)^2\Big]\Bigg]
\end{aligned}
$$

with

$$
\alpha_{i,0}(t):=Q_{t1}\left(\{(z,y)\,|\,y_i=0,z\in E^{\leftarrow}_{i,0}(y)\}\right),\qquad \beta_{i,0}(t):=Q_{0t}\left(\{(x,z)\,|\,x_i=0,z\in E^{\rightarrow}_{i,0}(x)\}\right),
$$

and

$$
\alpha_{i,1}(t):=Q_{t1}\left(\{(z,y)\,|\,y_i=1,z\in E^{\leftarrow}_{i,1}(y)\}\right),\qquad \beta_{i,1}(t):=Q_{0t}\left(\{(x,z)\,|\,x_i=1,z\in E^{\rightarrow}_{i,1}(x)\}\right).
$$

Observe that the sets $\{(x,z,y)\,|\,z\in[x,y],y_i=0,z\in E^{\leftarrow}_{i,0}(y)\}$ and $\{(x,z,y)\,|\,z\in[x,y]x_i=0,z\in E^{\rightarrow}_{i,0}(x)\}$ are disjoint. Indeed, if it is not the case, there exists $x,z,y,v,w\in\mathcal{X}_k$ such that $z\in[x,y]\cap[v,y]\cap[x,w]$, $x_i=z_i=y_i=0$, $v_i=w_i=1$, $\widehat{\pi}^0(v,y)>0$ and $\widehat{\pi}^0(x,w)>0$. Lemma 4.2 (i) implies that $z\in[v,w]$. This leads to a contradiction since $z_i=0$ and $v_i=w_i=1$. It follows that

$$
\alpha_{i,0}(t)+\beta_{i,0}(t)\leq\sum_{x,y\in\mathcal{X}_k}\sum_{z\in E^{\leftarrow}_{i,0}(y)\cup E^{\rightarrow}_{i,0}(x)}\nu_t^{0\,x,y}(z)\widehat{\pi}^0(x,y)\leq\widehat{Q}^0_t(\{z\in\mathcal{X}_k\,|\,z_i=0\}).
$$

Similarly one proves that

$$
\alpha_{i,1}(t)+\beta_{i,1}(t)\leq\widehat{Q}^0_t(\{z\in\mathcal{X}_k\,|\,z_i=1\}).
$$

As a consequence, from the identity $\inf_{\alpha>0,\beta>0,\alpha+\beta\leq 1}\left\{\frac{u^2}{\alpha}+\frac{v^2}{\beta}\right\}=u^2+v^2$, $u,v\geq 0$, one gets

$$
\begin{aligned}
&\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t)+\liminf_{\gamma_k\to 0}\psi''_{\gamma_k}(t)\\
&\geq\max\Big[\sum_{i\in[n]}\frac{1}{\widehat{Q}^0_t(\{z\in X_k\,|\,z_i=0\})}\left(\iint\mathbb{1}_{x_i\neq y_i}\,d\widehat{\pi}^0(x,y)\right)^2,\\
&\qquad\sum_{j\in[n]}\frac{1}{Q^0_t(\{z\in\mathcal{X}_k\,|\,z_j=1\})}\left(\iint\mathbb{1}_{x_j\neq y_j}\,d\widehat{\pi}^0(x,y)\right)^2\Big]
\end{aligned}
$$

Applying again Cauchy-Schwarz inequality and since $\sum_{j\in[n]}Q^0_t(\{z\in\mathcal{X}_k\,|\,z_j=1\})=k$, $\sum_{i\in[n]}Q^0_t(\{z\in X_k\,|\,z_i=0\})=n-k$ and $2d(x,y)=\sum_{i\in[n]}\mathbb{1}_{x_i\neq y_i}$, one obtains

$$
\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t)+\liminf_{\gamma_k\to 0}\psi''_{\gamma_k}(t)\geq\frac{4}{\min(k,n-k)}W_1^2(\nu_0,\nu_1). \tag{45}
$$

Let us start again from (43) and (44) to reach another lower-bound. For any $i\in J_0(y)$, one has

$$
\begin{aligned}
\sum_{z\in E^{\leftarrow}_{i,0}(y)}a_t(z,y)&=\sum_{w\in\mathcal{X}_k}\sum_{z\in[y,w]}\mathbb{1}_{z\in E^{\leftarrow}_{i,0}(y)}\nu_t^{0\,w,y}(z)\widehat{\pi}^0_{\leftarrow}(w|y)\\
&\leq\sum_{w\in\mathcal{X}_k}\sum_{z\in[y,w]}\mathbb{1}_{z_i=y_i=0}\nu_t^{0\,w,y}(z)\widehat{\pi}^0_{\leftarrow}(w|y)\\
&=\sum_{w\in\mathcal{X}_k}\mathbb{1}_{y_i=w_i=0}\widehat{\pi}^0_{\leftarrow}(w|y)+\sum_{w\in\mathcal{X}_k}\mathbb{1}_{y_i\neq w_i}\Big(\sum_{z\in[y,w]}\mathbb{1}_{z_i=y_i=0}\nu_t^{0\,w,y}(z)\Big)\widehat{\pi}^0_{\leftarrow}(w|y).
\end{aligned}
$$

Since for $y_i = 0$ and $w_i = 1$,

$$\begin{aligned}\sum_{z\in[y,w]} \mathbb{1}_{z_i=y_i=0}\nu_t^{0w,y}(z) &= \sum_{k=0}^{d(y,w)-1}\Big(\sum_{z,z\in[y,w],z_i=0} \mathbb{1}_{d(y,z)=k}\Big)\frac{(1-t)^k t^{d(y,w)-k}}{\binom{d(y,w)}{k}}\\ &= \sum_{k=0}^{d(y,w)-1}\binom{d(y,w)}{k}\binom{d(y,w)-1}{k}\frac{(1-t)^k t^{d(y,w)-k}}{\binom{d(y,w)}{k}}\\ &= t\end{aligned}$$

one gets for any $i \in J_0(y)$

$$\sum_{z\in E^{\leftarrow}_{i,0}(y)} a_t(z,y) \le 1-(1-t)\int \mathbb{1}_{y_i\neq w_i} d\widehat{\pi}^0_{\leftarrow}(w|y).$$

One identically shows that for any $i \in J_1(y)$,

$$\sum_{z\in E^{\leftarrow}_{i,1}(y)} a_t(z,y) \le 1-(1-t)\int \mathbb{1}_{y_i\neq w_i} d\widehat{\pi}^0_{\leftarrow}(w|y).$$

As a consequence, setting $\Pi^i_{\leftarrow}(y) := \int \mathbb{1}_{y_i\neq w_i} d\widehat{\pi}^0_{\leftarrow}(w|y)$, (43) provides

$$\liminf_{\gamma_k\to 0} \varphi''_{\gamma_k}(t) \ge \varphi''_0(t), \tag{46}$$

where

$$\varphi_0(t) := \max\Big[\int\sum_{i=1}^n \mathbb{1}_{y_i=0} f\Big(1-(1-t)\Pi^i_{\leftarrow}(y)\Big)\,d\nu_1(y),\qquad \int\sum_{j=1}^n \mathbb{1}_{y_j=1} f\Big(1-(1-t)\Pi^j_{\leftarrow}(y)\Big)\,d\nu_1(y)\Big]$$

with $f(s) := s\log s - s$. One may identically show from (44) that

$$\liminf_{\gamma_k\to 0} \psi''_{\gamma_k}(t) \ge \psi''_0(t), \tag{47}$$

where

$$\psi_0(t) := \max\Big[\int\sum_{i=1}^n \mathbb{1}_{x_i=0} f\Big(1-t\,\Pi^i_{\to}(x)\Big)\,d\nu_0(x), \int\sum_{j=1}^n \mathbb{1}_{x_j=1} f\Big(1-t\,\Pi^j_{\to}(x)\Big)\,d\nu_0(x)\Big].$$

As in the case of the hypercube, another lower bound for $\liminf_{\gamma\to 0}\varphi''_\gamma(t)$ can be reached by estimating differently the right-hand side of inequality (42).

For any fixed positive reals $A_{kl,ij}(z,y), ((i,j),(k,l)) \in \mathbb{I}(z,y)$, let us define the convex function $F : (\mathbb{R}^*_+)^{I(z,y)} \to \mathbb{R}$ defined by

$$F((\beta_{ij})_{(i,j)\in I}) := \Big(\sum_{(i,j)\in I}\beta_{ij}\Big)^2 + \frac{1}{2}\sum_{((i,j),(k,l))\in\mathbb{I}}\Big(\rho(\beta_{ij},A_{kl,ij})+\rho(\beta_{kl},A_{kl,ij})\Big),\ \beta_{ij}\in\mathbb{R}^*_+$$

(the dependence in z, y is omitted to simplify the notations). As in the proof of Theorem 2.4, after some computations, its minimum value is given by: if $\mathbb{I}^{\leftarrow} = \emptyset$, then

$$\inf_{(\beta_{ij})_{(i,j)\in I}\in(\mathbb{R}^*_+)^{I^{\leftarrow}}} F((\beta_{ij})_{(i,j)\in I^{\leftarrow}}) = 0,$$

and if $\mathbb{I} \neq \emptyset$

$$\inf_{(\beta_{ij})_{(i,j)\in I}\in(\mathbb{R}^*_+)^{I^{\leftarrow}}} F((\beta_{ij})_{(i,j)\in I}) = \sum_{((i,j),(k,l))\in\mathbb{I}^{\leftarrow}} A_{kl,ij}\log\frac{A_{kl,ij}S}{S_{ij}S_{kl}},$$

where $S_{ij} := \sum_{(k,l)\in\mathbb{I}^{\leftarrow}_{2,ij}} A_{kl,ij}$ and $S := \sum_{(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}$. Let $W := \sum_{((i,j),(k,l))\in\mathbb{I}^{\leftarrow}} S_{ij}S_{kl}$. Assume that $\mathbb{I}^{\leftarrow} \neq \emptyset$. By convexity of the function $H : t \mapsto t\log t$, applying Jensen inequality, it follows that

$$\inf_{(\beta_{ij})_{(i,j)\in I}\in(\mathbb{R}^*_+)^{I^{\leftarrow}}} F((\beta_{ij})_{(i,j)\in I}) = \frac{1}{S}\sum_{((i,j),(k,l))\in\mathbb{I}^{\leftarrow}} H\left(\frac{A_{ij,kl}S}{S_{ij}S_{kl}}\right)S_{ij}S_{kl}$$
$$\geq \frac{W}{S}H\Big(\sum_{((i,j),(k,l))\in\mathbb{I}^{\leftarrow}} \frac{A_{ij,kl}S}{W}\Big) = -S\,\log\left(\frac{W}{S^2}\right).$$

For any $(i,j)\in\mathbb{I}^{\leftarrow}_1$, one has

$$\mathbb{I}^{\leftarrow}_{2,ij} \subset \mathbb{I}^{\leftarrow}_1 \setminus \left\{\left\{(i,j)\right\}\cup\left\{(i,l')\,\middle|\, l'\in[n]\setminus\{j\}\right\}\cup\left\{(k',j)\,\middle|\,k'\in[n]\setminus\{i\}\right\}\right\},$$

and therefore

$$\begin{aligned}
W &= \sum_{(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\sum_{(k,l)\in\mathbb{I}^{\leftarrow}_{2,ij}} S_{kl}\\
&\leq \sum_{(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\Big[\Big(\sum_{(k,l)\in\mathbb{I}^{\leftarrow}_1} S_{kl}\Big) + S_{ij} - \Big(\sum_{l',(i,l')\in\mathbb{I}^{\leftarrow}_1} S_{il'}\Big) - \Big(\sum_{k',(k',j)\in\mathbb{I}^{\leftarrow}_1} S_{k'j}\Big)\Big]\\
&= S^2 + \widetilde{S}^2 - \sum_{i\in J_0}\Big(\sum_{j,(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\Big)^2 - \sum_{j\in J_1}\Big(\sum_{i,(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\Big)^2,
\end{aligned}$$

where we set $\widetilde{S}^2 := \sum_{(i,j)\in\mathbb{I}_1} S^2_{ij}$. By Cauchy Schwarz inequality, since $|J_0| = n-k$ and $|J_1| = k$, one has

$$\sum_{i\in J_0}\Big(\sum_{j,(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\Big)^2 \geq \frac{S^2}{n-k} \quad\text{and}\quad \sum_{j\in J_1}\Big(\sum_{i,(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\Big)^2 \geq \frac{S^2}{k}.$$

As a consequence, since $\sum_{i\in J_0}\left(\sum_{j,(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\right)^2 \geq \widetilde{S}^2$ and $\sum_{j\in J_1}\left(\sum_{i,(i,j)\in\mathbb{I}^{\leftarrow}_1} S_{ij}\right)^2 \geq \widetilde{S}^2$, we get

$$W \leq S^2\left(1 - \max\Big[\frac{1}{n-k},\frac{1}{k}\Big]\right),$$

and therefore

$$\inf_{(\beta_{ij})_{(i,j)\in I^{\leftarrow}}\in(\mathbb{R}^*_+)^{I^{\leftarrow}}} F((\beta_{ij})_{(i,j)\in I^{\leftarrow}}) \geq C_{n,k}S.$$

This lower estimate also holds if $\mathbb{I}^{\leftarrow} = \emptyset$ since $S = 0$ in that case. As a consequence (42) imply

$$\begin{aligned}
\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(t) &\geq C_{n,k}\int \sum_{((i,j),(k,l))\in\mathbb{I}^{\leftarrow}} A_{kl,ij}\, d\widehat{Q}^0_{t,1}\\
&= C_{n,k}\int \sum_{(i,j),(k,l)\in J_0(y)\times J_1(y), i\neq k, j\neq l}\ \sum_{z\in\mathcal{X}_k} A_{kl,ij}(z,y)\,a_t(z,y)\,d\nu_1(y).
\end{aligned}$$

Observing that for $(i,j),(k,l)\in J_0(y)\times J_1(y)$ with $i\neq k$ and $j\neq l$, $(z,\sigma_{kl}\sigma_{ij}(z))\in[y,w]$ if and only if one has $y_i = w_j = y_k = w_l = 0$, $y_j = w_i = y_l = w_k = 1$ and $z\in[y,\sigma_{kl}\sigma_{ij}(w)]$, and using the fact that $L^{d(\sigma_{kl}\sigma_{ij}(z),w)}(\sigma_{kl}\sigma_{ij}(z),w) = L^{d(z,\sigma_{kl}\sigma_{ij}(w))}(z,\sigma_{kl}\sigma_{ij}(w))$, one gets for any $y\in\mathcal{X}_k$, and

$(i,j),(k,l)\in J_0(y)\times J_1(y)$ with $i\neq k$ and $j\neq l$,

$$\sum_{z\in\mathcal{X}_k} A_{kl,ij}(z,y)\,a_t(z,y)$$

$$= \sum_{w\in\mathcal{X}_k, w_j=w_l=0, w_i=w_k=1}\ \sum_{s=0}^{d(y,\sigma_{kl}\sigma_{ij}(w))}\ \sum_{z\in[y,\sigma_{kl}\sigma_{ij}(w)], d(y,z)=s} r(y,z,z,\sigma_{kl}\sigma_{ij}(w))$$

$$d(y,w)(d(y,w)-1)\,\rho_t^{d(y,w)-2}(d(y,w)-2-s)\,\widehat{\pi}^0_{\leftarrow}(w|y)$$

$$= \sum_{w\in\mathcal{X}_k} \mathbb{1}_{y_i=w_j=y_k=w_l=0}\mathbb{1}_{y_j=w_i=y_l=w_k=1}\frac{L^{d(y,\sigma_{ij}(w))}(y,\sigma_{ij}(w))}{L^{d(y,w)}(y,w)}\,d(y,w)(d(y,w)-1)\,\widehat{\pi}^0_{\leftarrow}(w|y)$$

$$= \sum_{w\in\mathcal{X}_k} \frac{\mathbb{1}_{y_i=w_j=y_k=w_l=0}\mathbb{1}_{y_j=w_i=y_l=w_k=1}}{d(y,w)(d(y,w)-1)}\,\widehat{\pi}^0_{\leftarrow}(w|y).$$

From the identities

$$\sum_{i\in J_0(y)}\ \sum_{k\in J_0(y)\setminus\{i\}} \mathbb{1}_{y_i=0,w_i=1}\mathbb{1}_{y_k=0,w_k=1} = d(y,w)(d(y,w)-1),$$

and

$$\sum_{j\in J_1(y)}\ \sum_{l\in J_1(y)\setminus\{j\}} \mathbb{1}_{y_j=1,w_j=0}\mathbb{1}_{y_l=1,w_l=0} = d(y,w)(d(y,w)-1),$$

we finally obtain

$$\liminf_{\gamma_k\to0}\varphi''_{\gamma_k}(t)\geq C_{n,k}\int\sum_{w\in\mathcal{X}_k} d(y,w)(d(y,w)-1)\,\widehat{\pi}^0_{\leftarrow}(w|y)\,d\nu_1(y)$$

$$= C_{n,k}\iint d(y,w)(d(y,w)-1)\,d\widehat{\pi}^0(y,w).$$

By symmetry the same estimate holds for $\liminf_{\gamma\to0}\psi''_\gamma(t)$. The proof of Theorem 2.6 ends by applying Lemma 3.1 and using the last estimates together with the other ones (45),(46) and (47). □

4. Appendix A : Basic lemmas

Lemma 4.1. *The transport-entropy inequality* (24) *implies the W_2 transport-entropy inequality* (25) *for the standard Gaussian measure with the constant* 4 *instead of 2.*

Proof. The result follows from the transport-entropy inequality (24) for the uniform probability measure μ on the hypercucube ($\alpha_i=1/2$ for all $i\in[n]$), by using the central limit Theorem with the projection map

$$T_n(x) := \frac{2}{\sqrt{n}}\Big(\sum_{i=1}^n x_i-\frac{n}{2}\Big),\qquad x,y\in\{0,1\}^n.$$

Let $\nu\in\mathcal{P}_2(\mathbb{R})$ with continuous density f with respect to γ. Let ν^n denotes the probability measure on $\{0,1\}^n$ with density f_n with respect to μ given by

$$f_n(x) := \frac{f(T_n(x))}{\int f\circ T_n\,d\mu},\qquad x\in\{0,1\}^n.$$

Applying (24) with $\nu_0:=\mu$ and $\nu_1:=\nu^n$, one gets

$$\frac{1}{n}W_2^d(\mu,\nu^n)^2\leq H(\nu^n|\mu). \tag{48}$$

By the weak convergence of $T_n\#\mu$ to the standard Gaussian law γ, one has

$$\lim_{n\to\infty} H(\nu^n|\mu) = H(\nu|\gamma). \tag{49}$$

Easy computations give for any $x, y \in \{0,1\}^n$,

$$\frac{1}{n}\, d(x,y)(d(x,y)-1) \geq \frac{1}{4}\,|T_n(x)-T_n(y)|\left(|T_n(x)-T_n(y)| - \frac{2}{\sqrt{n}}\right),$$

and therefore

$$\frac{1}{n}\, W_2^d(\mu,\nu^n)^2 \geq \frac{1}{4} \inf_{\pi_n\in\Pi(T_n\#\mu,T_n\#\nu^n)} \iint c_n(z,w)\, d\pi_n(z,w),$$

where $c_n(z,w) = |z-w|\left(|z-w| - \frac{2}{\sqrt{n}}\right)$. Let $\varepsilon > 0$. Since $T_n\#\mu$ weakly converges to γ and $T_n\#\nu^n$ weakly converges to ν, one checks that any sequence $\pi_n \in \Pi(T_n\#\mu, T_n\#\nu^n)$ is relatively compact, there exists a compact set K_ε such that

$$\sup_{n\in\mathbb{N}} \int_{\mathbb{R}^2\setminus K_\varepsilon} (|z|+|w|)\, d\pi_n(z,w) \leq \varepsilon.$$

Let $c(z,w) := |z-w|^2$. The cost c_n uniformly converges to the quadratic cost c on K_ε. It follows that for n sufficiently large

$$\iint c_n\, d\pi_n \geq \iint c\, d\pi_n - \iint_{K_\varepsilon} |c-c_n|\, d\pi_n - \iint_{\mathbb{R}^2\setminus K_\varepsilon} (c-c_n)\, d\pi_n \geq \iint c\, d\pi_n - 2\varepsilon.$$

and therefore

$$\frac{1}{n}\, W_2^d(\mu,\nu^n)^2 \geq \frac{1}{4}\, W_2^2(T_n\#\mu, T_n\#\nu^n) - \frac{\varepsilon}{2}.$$

From the weak convergence in $\mathcal{P}_2(\mathbb{R})$ of the sequences $(T_n\#\mu)$ and $(T_n\#\nu^n)$ and then letting ε goes to 0, one gets

$$\liminf_{n\to+\infty} \frac{1}{n}\, W_2^d(\mu,\nu^n)^2 \geq \frac{1}{4}\, W_2^2(\nu,\gamma).$$

Finally, (48) and (49) imply $W_2^2(\nu,\gamma) \leq 4H(\nu|\gamma)$ as n goes to $+\infty$. □

Lemma 4.2. *Let $\mathcal{X}$ be a graph with graph distance d. Let $\nu_0, \nu_1 \in \mathcal{P}(\mathcal{X})$ and assume that $\widehat{\pi} \in \mathcal{P}(\mathcal{X}\times\mathcal{X})$ is a W_1-optimal coupling of ν_0 and ν_1,*

$$W_1(\nu_0,\nu_1) = \iint d(x,y)\, d\widehat{\pi}(x,y).$$

(i) Let $x,y,z,v,w \in \mathcal{X}$ such that $z\in[x,y]$, $z\in[x,w]$, $z\in[v,y]$ and $\widehat{\pi}(v,y)>0$, $\widehat{\pi}(x,w)>0$. Then one has $z\in[v,w]$.

(ii) On the complete graph $\mathcal{X}$, the graph distance is the Hamming distance $d(x,y) = \mathbb{1}_{x\neq y}$, $x,y\in\mathcal{X}$. Setting for any $x,y\in\mathcal{X}$, $\Delta_\rightarrow(x) = \int \mathbb{1}_{w\neq x} d\widehat{\pi}_\rightarrow(w|x)$ and $\Delta_\leftarrow(y) = \int \mathbb{1}_{w\neq y} d\widehat{\pi}_\leftarrow(w|y)$, the two sets

$$D_\rightarrow = \left\{x\in \mathrm{supp}(\nu_0)\,\middle|\,\Delta_\rightarrow(x)\neq 0\right\} \quad \textit{and} \quad D_\leftarrow = \left\{y\in \mathrm{supp}(\nu_1)\,\middle|\,\Delta_\leftarrow(y)\neq 0\right\}$$

do not intersect.

Proof. (i) In order to get the property, $z\in[v,w]$, or equivalently, $d(v,z)+d(z,w) = d(v,w)$, it suffices to show that

$$d(v,z)+d(z,w) \leq d(v,w).$$

and the equality follows from the triangle inequality. The assumption $z\in[x,y]\cap[v,y]\cap[x,w]$ implies

$$d(v,z)+d(z,w) = (d(v,y)-d(z,y)) + (d(x,w)-d(x,z)) = d(v,y)+d(x,w)-d(x,y). \tag{50}$$

It is well known that the support of any optimizer of $W_1(\nu_0, \nu_1)$ is d-cyclically monotone (see [40, Theorem 5.10]. By definition, it means that for any family $(x_1, y_1), \ldots, (x_N, y_N)$ of points in the support of $\widehat{\pi}$

$$\sum_{i=1}^{N} d(x_i, y_i) \leq \sum_{i=1}^{N} d(x_i, y_{i+1}),$$

with the convention $y_{N+1} = y_1$. Therefore assumptions $\widehat{\pi}(v, y) > 0$ and $\widehat{\pi}(x, w) > 0$ imply

$$d(v, y) + d(x, w) \leq d(v, w) + d(x, y),$$

and the expected inequality follows from (50).

(ii) If the two sets $D_{\rightarrow}$ and $D_{\leftarrow}$ intersect, then there exists $x, w', w \in \mathcal{X}$ such that $x \neq w$, $x \neq w'$ $\widehat{\pi}(x, w) > 0$ and $\widehat{\pi}(w', x) > 0$. As above, since the support of $\widehat{\pi}$ is d-cyclically monotone, one should have

$$2 = d(x, w) + d(w', x) \leq d(x, x) + d(w', w) \leq 1,$$

which is impossible.

□

Lemma 4.3. *Let ν_0 and ν_1 some probability measures in $\mathcal{P}(\mathcal{X})$ with bounded support.*

(i) If (12) *holds, then for any $x, y \in \mathcal{X}$ and any integer k, $L^k(x, y) \leq (2S)^k$.*

(ii) If (13) *holds, then for any $x, y \in \mathcal{X}$, $L^{d(x,y)}(x, y) \geq I^{d(x,y)}$.*

(iii) If (12) *and* (13) *hold, then for any $x, y \in \mathcal{X}$, any $t \in [0, 1]$, and any $\gamma \in (0, 1)$, one has*

$$P_t^{\gamma}(x, y) = \frac{L^{d(x,y)}(x, y)}{d(x, y)!} (\gamma t)^{d(x,y)} \left(1 + \gamma K^{d(x,y)} O(1)\right),$$

where $K := 2S/I$ and $O(1)$ denotes a quantity uniformly bounded in x, y and t.

(iv) If (12) *holds then for any $x, y, z \in \mathcal{X}$ and for any $t \in [0, 1]$*

$$\lim_{\gamma \to 0} \nu_t^{\gamma x,y}(z) = \nu_t^{0 x,y}(z) := \mathbb{1}_{[x,y]}(z)\, r(x, z, z, y)\, \rho_t^{d(x,y)}(d(x, z)).$$

(v) If (12) *holds then for any $x, y \in \mathcal{X}$,*

$$P_t^{\gamma}(x, y) \geq \frac{L^{d(x,y)}(x, y)}{d(x, y)!} (t\gamma)^{d(x,y)} e^{-\gamma t S}.$$

(vi) If (12) *holds then $\mathbb{E}_{R^{\gamma}}[\ell | X_0 = x, X_1 = y] \leq \frac{\gamma S}{P_1^{\gamma}(x,y)}$.*

(vii) Assume (12) *and* (13) *hold. Let $D := \max_{x \in \mathrm{supp}(\nu_0), y \in \mathrm{supp}(\nu_1)} (d(x_0, x), d(x_0, y))$. For any $x \in \mathrm{supp}(\nu_0)$ and $y \in \mathrm{supp}(\nu_1)$, one has*

$$\nu_t^{\gamma x,y}(z) \leq O(1) \left(\mathbb{1}_{[x,y]}(z) + \left(1 - \mathbb{1}_{[x,y]}(z)\right) \gamma \left(\gamma K^2\right)^{[2d(x_0,z)-4D-1]_+}\right),$$

where $K := 2S/I$ and $O(1)$ denotes a constant that does not depend on x, y, z, γ, t and $K := 2S/I$.

As a consequence, if B denotes the finite set

$$B := \left\{z \in \mathcal{X} \,\middle|\, z \in [x, y], x \in \mathrm{supp}(\nu_0), y \in \mathrm{supp}(\nu_1)\right\},$$

then

$$\widehat{Q}_t^{\gamma}(z) \leq O(1)\, \gamma \left(\gamma K^2\right)^{[2d(x_0,z)-4D-1]_+}, \qquad \forall z \in \mathcal{X} \setminus B.$$

(viii) Assume (12) *and* (13) *hold. For any $w, z, z' \in \mathcal{X}$ with $d(z, z') \leq 2$ and $w \in \mathrm{supp}(\nu_0)$ one has*

$$\frac{P_t^{\gamma}(z', w)}{P_t^{\gamma}(z, w)} \leq \frac{\max\left(1, d(x_0, z)^{d(z,z')}\right) K^{d(x_0,z)}\, O(1)}{(\gamma t)^{d(z,z')}},$$

where $K := 2S/I$ and $O(1)$ is a positive constant that does not depend on z, z', γ, t. It follows that

$$\frac{(\gamma t)^{d(z,z')}}{\max\left(1, d(x_0,z)^{d(z,z')}\right) K^{d(x_0,z)}\, O(1)} \leq \frac{P_t^\gamma f^\gamma(z')}{P_t^\gamma f^\gamma(z')} \leq \frac{\max\left(1, d(x_0,z)^{d(z,z')}\right) K^{d(x_0,z)}\, O(1)}{(\gamma t)^{d(z,z')}} \tag{51}$$

(ix) *Let $(\gamma_k)_{k\in\mathbb{N}}$ be a sequence of positive numbers converging to zero. If* (11)*,* (12)*,* (13) *and* (14) *hold, then for any $t \in [0,1]$*

$$\lim_{\gamma_k\to 0} H(\widehat{Q}_t^{\gamma_k}|m) = H(\widehat{Q}_t^0|m).$$

Proof. (i) Given (12), we want to show that for any $x \in \mathcal{X}$, $S_k(y) := \sup_{x\in\mathcal{X}} |L^k(x,y)| \leq (2S)^k$. It follows by induction on k from the inequality

$$S_{k+1}(y) = \sup_{x\in\mathcal{X}} \left| L(x,x)L^k(x,y) + \sum_{z, z\sim x} L(x,z)L^k(z,y)\right| \leq 2 \sup_{x\in\mathcal{X}} |L(x,x)|\ S_k(y).$$

(ii) For $x = y$, one has $L^{d(x,y)}(x,y) = 1$ and by definition for $x \neq y$,

$$L^{d(x,y)}(x,y) := \sum_\alpha L_\alpha,$$

where the sum is over all path α from x to y of length $d(x,y)$, $\alpha = (z_0, \ldots, z_{d(x,y)})$ with $z_0 = x$ and $z_{d(x,y)} = y$, and

$$L_\alpha := L(z_0,z_1)L(z_1,z_2)\ldots L(z_{d(x,y)-1}, z_{d(x,y)}).$$

Such a path α is a geodesic. Since we assume in this paper that $L(x,y) > 0$ if and only if x and y are neighbour, one has $L_\alpha > 0$. By irreducibility it always exists at most one geodesic path from x to y, and from assumption (12), for such a path α, $L_\alpha \geq I^{d(x,y)}$. As a consequence we get $L^{d(x,y)}(x,y) \geq I^{d(x,y)}$.

(iii) According to (15), for any $x, y \in \mathcal{X}$,

$$P_t^\gamma(x,y) = \frac{L^{d(x,y)}(x,y)}{d(x,y)!} (\gamma t)^{d(x,y)} \cdot \left(1 + \gamma \sum_{k, k\geq d(x,y)+1} \frac{L^k(x,y)}{L^{d(x,y)}(x,y)} \frac{d(x,y)!}{k!} t^{k-d(x,y)} \gamma^{k-d(x,y)-1}\right).$$

Applying Lemma 4.3 (i) and (ii), we get

$$\begin{aligned}
&\left| P_t^\gamma(x,y) - \frac{L^{d(x,y)}(x,y)}{d(x,y)!} (\gamma t)^{d(x,y)}\right| \\
&\quad\leq \gamma \frac{L^{d(x,y)}(x,y)}{d(x,y)!} (\gamma t)^{d(x,y)} \sum_{k, k\geq d(x,y)+1} K^{d(x,y)} \frac{(2S)^{k-d(x,y)}}{(k-d(x,y))!} \\
&\quad\leq \gamma \frac{L^{d(x,y)}(x,y)}{d(x,y)!} (\gamma t)^{d(x,y)} K^{d(x,y)} e^{2S},
\end{aligned}$$

from which the expected result follows.

(iv) Let $x, y, z \in \mathcal{X}$ and $t \in [0,1]$. If (12) holds, according to (15), the Taylor expansion of $P_t^\gamma(x,y)$ as γ goes to zero is given by

$$P_t^\gamma(x,y) = \frac{L^{d(x,y)}(x,y)}{d(x,y)!} (\gamma t)^{d(x,y)} + o(\gamma^{d(x,y)}),$$

As a consequence, the Taylor expansion of $\nu_t^{\gamma x,y}(z)$, defined by (9), is

$$\nu_t^{\gamma\, x,y}(z)$$
$$= \gamma^{d(x,z)+d(z,y)-d(x,y)} \frac{L^{d(x,z)}(x,z)L^{d(z,y)}(z,y)}{L^{d(x,y)}(x,y)} \frac{d(x,y)!}{d(x,z)!d(z,y)!} t^{d(x,z)}(1-t)^{d(z,y)} + o(\gamma^{d(x,z)+d(z,y)-d(x,y)}).$$

The expected result follows since one has $\gamma^{d(x,z)+d(z,y)-d(x,y)} = 1$ if $z \in [x,y]$, and $\lim_{\gamma\to 0}\gamma^{d(x,z)+d(z,y)-d(x,y)} = 0$ otherwise.

(v) On some probability space $(\Omega', \mathcal{A}, \mathbb{P})$, let $(N_s)_{s\geq 0}$ be a Poisson process with parameter γS and $(Y_n)_{n\in\mathbb{N}}$ be a Markov chain on $\mathcal{X}$ with transition matrix Q given by

$$Q(z,w) := \frac{L^\gamma(x,w)}{\gamma S}, \quad \text{for } w \neq z \in \mathcal{X}, \text{ and } \quad Q(z,z) := \frac{\gamma S + L^\gamma(z,z)}{\gamma S}.$$

We assume that $(Y_n)_{n\in\mathbb{N}}$ and $(N_s)_{s\geq 0}$ are independent. It is well known that the law of the process $(X_t)_{t\geq 0}$ under R^γ given $X_0 = x$ is the same as the law of the process $(\widetilde{X}_t)_{t\geq 0}$ under $\mathbb{P}$ given $\widetilde{X}_0 = x$ defined by $\widetilde{X}_t := Y_{N_t}$. As a consequence, one has for any $y \in \mathcal{X}$,

$$P_t^\gamma(x,y) = R^\gamma\left(X_t = y \,|\, X_0 = x\right) = \mathbb{P}\left(\widetilde{X}_t = y \,|\, \widetilde{X}_0 = x\right).$$

Let $n = d(x,y)$ and $\widetilde{N}_t$ denotes the number of jumps of the process $\widetilde{X}_t$, one has

$$\begin{aligned}
P_t^\gamma(x,y) &\geq \mathbb{P}\left(\widetilde{X}_t = y, \widetilde{N}_t = n \,|\, \widetilde{X}_0 = x\right)\\
&= \mathbb{P}\left(Y_1,\ldots,Y_n \text{ are all differents}, Y_n = y, N_t = n \,|\, \widetilde{X}_0 = x\right)\\
&= \mathbb{P}\left(N_t = n\right)\mathbb{P}(Y_1,\ldots,Y_n \text{ are all differents}, Y_n = y \,|\, \widetilde{X}_0 = x\right)\\
&= \frac{(\gamma t S)^n}{n!} e^{-\gamma t S} \sum_{\alpha=(x_0,\ldots,x_n),\, \alpha \text{ geodesic from } x \text{ to } y} Q(x_0,x_1)\cdots Q(x_{n-1},x_n)\\
&= \frac{(\gamma t)^n}{n!} e^{-\gamma t S} L^{d(x,y)}(x,y).
\end{aligned}$$

(vi) The length $\ell(\omega)$ of a path $\omega \in \Omega$ represents the number of jumps of the process X_t between times 0 and 1. Therefore according to the definition of the process $(\widetilde{X}_t)_{t\geq 0}$ above,

$$\begin{aligned}
\mathbb{E}_{R^\gamma}[\ell \,|\, X_0 = x, X_1 = y] &= \mathbb{E}_{\mathbb{P}}\left[\widetilde{N}_1 \,|\, \widetilde{X}_0 = x, \widetilde{X}_1 = y\right]\\
&\leq \mathbb{E}_{\mathbb{P}}\left[N_1 \,|\, \widetilde{X}_0 = x, \widetilde{X}_1 = y\right] = \frac{\mathbb{E}_{\mathbb{P}}\left[N_1 \mathbb{1}_{\widetilde{X}_1 = y} \,|\, \widetilde{X}_0 = x\right]}{\mathbb{P}\left(\widetilde{X}_1 = y \,|\, \widetilde{X}_0 = x\right)} \leq \frac{\mathbb{E}_{\mathbb{P}}\left[N_1\right]}{P_1^\gamma(x,y)},
\end{aligned}$$

which ends the proof since $\mathbb{E}_{\mathbb{P}}\left[N_1\right] = \gamma S$.

(vii) From (iii) and (v), one gets for any $x,z,y \in \mathcal{X}$,

$$\begin{aligned}
\nu_t^{\gamma\, x,y}(z) &= \frac{P_t^\gamma(x,z)P_t^\gamma(z,y)}{P_1^\gamma(x,y)}\\
&\leq \gamma^{d(x,z)+d(z,y)-d(x,y)} r(x,z,z,y)\, \rho_t^{d(x,y)}(d(x,z))\, e^{\gamma S}\\
&\qquad \left(1+\gamma K^{d(x,z)}O(1)\right)\left(1+\gamma K^{d(z,y)}O(1)\right).
\end{aligned} \tag{52}$$

If $z \in [x,y]$ then thanks to (i) and (ii), the right-hand side of this inequality is bounded from above by

$$\left(\frac{S}{I}\right)^{d(x,y)} e^{d(x,y)} e^{\gamma S} 4K^{2d(x,y)} O(1),$$

and the maximum of this quantity over all $x \in \mathrm{supp}(\nu_0)$ and $y \in \mathrm{supp}(\nu_1)$ is a constant $O(1)$, independent of x, z, y and γ.
If $z \notin [x, y]$, then $d(x, z) + d(z, y) - d(x, y) \geq 1$, and the right-hand side of (52) is bounded by

$$\gamma^{d(x,z)+d(z,y)-d(x,y)} \frac{S^{d(x,z)+d(z,y)}}{I^{d(x,y)}} d(x, y)!\, e^{\gamma S}\, 4K^{d(x,z)+d(z,y)} O(1)$$
$$\leq \gamma^{1+[2d(x_0,z)-4D-1]_+} \frac{S^{2d(x_0,z)+2D}}{I^{d(x,y)}} d(x, y)!\, e^{\gamma S}\, 4K^{2d(x_0,z)+2D} O(1).$$

The maximum over all $x \in \mathrm{supp}(\nu_0)$ and $y \in \mathrm{supp}(\nu_1)$ of the right-hand side quantity is bounded by $O(1)\,\gamma^{1+[2d(x_0,z)-4D-1]_+} K^{4d(x_0,z)}$. This ends the proof of the first inequality of (vii). The second inequality easily follows since

$$\widehat{Q}_t^{\gamma}(z) = \sum_{x\in\mathrm{supp}(\nu_0), y\in\mathrm{supp}(\nu_1)} \nu_t^{\gamma x,y}(z)\, \widehat{\pi}^{\gamma}(x, y).$$

(viii) Using (iii) and (v), one gets for any $z, z' \in \mathcal{X}$ and any $w \in \mathrm{supp}(\nu_0)$,

$$\frac{P_t^{\gamma}(z', w)}{P_t^{\gamma}(z, w)} \leq \frac{L^{d(z',w)}(z', w)}{L^{d(z,w)}(z, w)} \frac{d(z, w)!}{d(z', w)!} \left(\frac{1}{\gamma t}\right)^{d(z',w)-d(z,w)} e^{\gamma t S} \left(1 + \gamma K^{d(z',w)} O(1)\right)$$
$$\leq K^{d(z,z')+d(z,x_0)+d(x_0,w)} \max\left(1, d(z, w)^2\right) \left(\frac{1}{\gamma t}\right)^{d(z,z')} 2e^{S} K^{d(z,z')+d(z,x_0)+d(x_0,w)} O(1)$$
$$\leq \frac{K^{2d(z,x_0)} \max\left(1, d(z, x_0)^2\right) O(1)}{(\gamma t)^{d(z,z')}},$$

where one maximizes over all $w \in \mathrm{supp}(\nu_0)$ to get the last inequality. Inequality (51) follows since

$$\frac{P_t^{\gamma} f^{\gamma}(z')}{P_t^{\gamma} f^{\gamma}(z)} = \sum_{w\in\mathrm{supp}(\nu_0)} \frac{P_t^{\gamma}(z', w)}{P_t^{\gamma}(z, w)} \frac{f^{\gamma}(w) P_t^{\gamma}(z, w)}{P_t^{\gamma} f^{\gamma}(z)},$$

with $\displaystyle\sum_{w\in\mathrm{supp}(\nu_0)} \frac{f^{\gamma}(w) P_t^{\gamma}(z, w)}{P_t^{\gamma} f^{\gamma}(z)} = 1.$

(ix) Recall that

$$H(\widehat{Q}_t^{\gamma}|m) = \sum_{z\in\mathcal{X}\setminus B} \log \frac{\widehat{Q}_t^{\gamma_k}(z)}{m(z)} \widehat{Q}_t^{\gamma_k}(z).$$

Let us consider the finite set B defined in Lemma 4.3 (vii). From the weak convergence of the sequence $(\widehat{Q}_t^{\gamma_k})$ to $\widehat{Q}_t^{0}$ and since $\mathrm{supp}(\widehat{Q}_t^{0}) \subset B$, one has

$$\lim_{\gamma_k\to 0} \sum_{z\in B} \log \frac{\widehat{Q}_t^{\gamma_k}(z)}{m(z)} \widehat{Q}_t^{\gamma_k}(z) = H(\widehat{Q}_t^{0}|m)$$

Therefore it remains to prove that

$$\lim_{\gamma_k\to 0} \sum_{z\in\mathcal{X}\setminus B} \log \frac{\widehat{Q}_t^{\gamma_k}(z)}{m(z)} \widehat{Q}_t^{\gamma_k}(z) = 0.$$

From Lemma 4.3(vii) and hypothesis (11) one has, for any $z \in \mathcal{X} \setminus B$,

$$\frac{\widehat{Q}_t^{\gamma_k}(z)}{m(z)} \leq \frac{O(1)\,\gamma\, \left(\gamma K^2\right)^{[2d(x_0,z)-4D-1]_+}}{\inf_{z\in\mathcal{X}} m(z)}.$$

Using the inequality $|v\log v| \leq \sqrt{v}$ for $v \in (0,1]$, we get for $0 < \gamma \leq \min\left(\frac{\inf_{z\in\mathcal{X}} m(z)}{O(1)}, \frac{1}{K^2}\right)$,

$$\sum_{z\in\mathcal{X}\setminus B} \log \frac{\widehat{Q}_t^{\gamma_k}(z)}{m(z)}\, \widehat{Q}_t^{\gamma_k}(z) \leq O(1) \sup_{z\in\mathcal{X}} m(z)\, \sqrt{\gamma} \sum_{z\in\mathcal{X}} \left(\gamma K^2\right)^{[2d(x_0,z)-4D-1]_+/2}.$$

Hypothesis (14) then implies that there exists $\gamma_1 > 0$ such that for any $0 < \gamma < \gamma_1$

$$\sum_{z\in\mathcal{X}\setminus B} \log \frac{\widehat{Q}_t^{\gamma_k}(z)}{m(z)}\, \widehat{Q}_t^{\gamma_k}(z) \leq O(1)\, \sqrt{\gamma},$$

and the expected result follows.

□

5. Appendix B : Proofs of Lemmas 3.1, 3.2, 3.3, 3.4

Proof of Lemma 3.1. Let $\varepsilon \in (0,1/2)$. We first prove that if (12), (13) and (14) hold then $|\varphi_\gamma''(t)|$ is uniformly bounded over all $t \in [\varepsilon, 1]$ and $\gamma \in (0,\gamma_1)$ for some $\gamma_1 \in (0,1)$. According to (30) and inequality (31) and (32), for any $t \in [\varepsilon,1]$ and $\gamma > 0$,

$$\begin{aligned}
&|\varphi_\gamma''(t)| \\
&\leq O(1)\left[\frac{|\gamma\log\gamma|}{\varepsilon}\int d^2(x_0,z)K^{d(x_0,z)}d\widehat{Q}_t^{\gamma}(z) + \frac{1}{\varepsilon^2}\int \left(d^2(x_0,z)+1\right)K^{2d(x_0,z)}d\widehat{Q}_t^{\gamma}(z)\right] \\
&\leq O(1)\int d^2(x_0,z)K^{2d(x_0,z)}d\widehat{Q}_t^{\gamma}(z).
\end{aligned}$$

Using Lemma 4.3 (vii) and the fact that ν_0 and ν_1 have bounded support, it follows that

$$\begin{aligned}
&|\varphi_\gamma''(t)| \\
&\leq O(1) \sum_{x\in \mathrm{supp}(\nu_0), y\in\mathrm{supp}(\nu_1)} \max_{z\in[x,y]} \left(d^2(x_0,z)K^{2d(x_0,z)}\right) + O(1)\sum_{z\in\mathcal{X}} d^2(x_0,z)\left(\gamma K^3\right)^{2d(x_0,z)} \\
&= O(1) + \sum_{z\in\mathcal{X}} d^2(x_0,z)\left(\gamma K^3\right)^{[2d(x_0,z)-4D-1]_+}
\end{aligned}$$

Using hypothesis (14) and choosing $\gamma_1 > 0$ so that $(\gamma_1 K^3)^2 < \gamma_0$, one gets

$$\sup_{\gamma\in(0,\gamma_1), t\in[\varepsilon,1]} |\varphi_\gamma''(t)| \leq O(1).$$

One may similarly proved by symmetry that if (12), (12) and (14) hold, then $|\psi_\gamma''(t)|$ is also uniformly bounded, namely

$$\sup_{\gamma\in(0,\gamma_1), t\in[0,1-\varepsilon]} |\varphi_\gamma''(t)| \leq O(1).$$

Let $\varepsilon \in (0,1/2)$, and for $\gamma \in [0,1)$, let

$$F_\gamma^\varepsilon(t) = H(\widehat{Q}^\gamma_{(1-\varepsilon)t+\varepsilon(1-t)}|m), \qquad t\in[0,1].$$

We will first prove a convexity property for the function F_0^ε from a convexity property of $F_\varepsilon^{\gamma_k}$ as the sequence (γ_k) goes to zero. We use the identity, for any $t \in (0,1)$

$$(1-t)F^\varepsilon_{\gamma_k}(0) + tF^\varepsilon_{\gamma_k}(1) - F^\varepsilon_{\gamma_k}(t) = \frac{t(1-t)}{2}\int_0^1 K_t(s)(F^\varepsilon_{\gamma_k})''(s)\,ds, \tag{53}$$

where the kernel K_t is defined by (22). Observe that

$$\int_0^1 K_t(s)(F^\varepsilon_{\gamma_k})''(s)\,ds = (1-2\varepsilon)\int_\varepsilon^{1-\varepsilon} K_t\left(\frac{u-\varepsilon}{1-2\varepsilon}\right)\left(\varphi''_{\gamma_k}(u) + \psi''_{\gamma_k}(u)\right)du.$$

The above uniform bounds on φ''_γ and φ''_{γ_k} allow to apply Fatou's Lemma. Together with Lemma 4.3 (ix) it implies, for any $\varepsilon \in (0, 1/2)$

$$(54)\quad (1-t)F_0^\varepsilon(0) + tF_0^\varepsilon(1) - F_0^\varepsilon(t) \geq \frac{t(1-t)}{2}(1-2\varepsilon)\int_\varepsilon^{1-\varepsilon} K_t\left(\frac{u-\varepsilon}{1-2\varepsilon}\right)\liminf_{\gamma_k\to 0}\left(\varphi''_{\gamma_k}(u) + \psi''_{\gamma_k}(u)\right)du.$$

For any $t \in [0,1]$ the support of the measure $\widehat{Q}_t^0$ is finite, included in the set B defined Lemma 4.3 (vii). As a consequence, the function $t \in [0,1] \to H(\widehat{Q}_t^0|m)$ is continuous as a finite sum of continuous function. It follows that for any $t \in [0,1]$,

$$\lim_{\varepsilon\to 0} F_0^\varepsilon(t) = H(\widehat{Q}_t^0|m).$$

Consequently, using hypothesis (26) and applying Fatou's Lemma as ε goes to zero, equality (54) provides

$$\begin{aligned}
&(1-t)H(\nu_0|m) + tH(\nu_1|m) - H(\widehat{Q}_t^0|m)\\
&\quad\geq \frac{t(1-t)}{2}\int_0^1 K_t(u)\left(\liminf_{\gamma_k\to 0}\varphi''_{\gamma_k}(u) + \liminf_{\gamma_k\to 0}\psi''_{\gamma_k}(u)\right)du\\
&\quad\geq \frac{t(1-t)}{2}\int_0^1 K_t(u)\left(\varphi''_0(u) + \psi''_0(u)\right)du\\
&\quad= [(1-t)\varphi_0(0) + t\varphi_0(1) - \varphi_0(t)] + [(1-t)\psi_0(0) + t\psi_0(1) - \psi_0(t)]
\end{aligned}$$

were the last equality is a consequence of identity (53) applied with φ_0 and ψ_0. □

Proof of Lemma 3.2 and Lemma 3.3. To simplify the notations, the dependence in the temperature parameter γ is omitted. Let us note $f_t := P_t f$ and $g_t := P_{1-t}g$ and recall that $F_t := \log f_t$, $G_t := \log g_t$ and

$$\varphi(t) = \int F_t f_t\, g_t\, dm, \qquad \psi(t) = \int G_t f_t\, g_t\, dm.$$

The proof is based on Γ_2-calculus by using backward equations, $\partial_t f_t = Lf_t$, $\partial_t g_t = -Lg_t$, and integration by parts formula : for any functions $h : \mathcal{X} \to \mathbb{R}$, $k : \mathcal{X} \to \mathbb{R}$,

$$\int h\, Lk\, dm = \int k\, Lh\, dm.$$

We only present the proof of the expression of $\varphi'(t)$ and $\varphi''(t)$. Same arguments provide the expression of $\psi'(t)$ and $\psi''(t)$. We start with a general statement that we will apply twice. Let $(t,z) \to V_t(z)$ be some differentiable function of t, then for any $t \in (0,1)$,

$$\begin{aligned}
\partial_t\left(\int V_t f_t\, g_t\, dm\right) &= \int (\partial_t V_t)\, f_t\, g_t + V_t\,(Lf_t)\, g_t - V_t\, f_t\,(Lg_t)\, dm\\
&= \int (\partial_t V_t)\, f_t\, g_t + V_t\,(Lf_t)\, g_t - L(V_t f_t) g_t\, dm\\
&= \int \Big[\partial_t V_t(z) - \sum_{z',\, z'\sim z} e^{\nabla F_t(z,z')}\nabla V_t(z,z')\, L(z,z')\Big] f_t(z) g_t(z)\, dm(z). \qquad (55)
\end{aligned}$$

The first equality is due to the backward equation and the last equality holds by integration by part formula.

Applying (55) with $V_t = F_t$, and since

$$\partial_t F_t(z) = \sum_{z'\in\mathcal{X}} e^{\nabla F_t(z,z')} L(z,z') = \sum_{z',\, z'\sim z}\left(e^{\nabla F_t(z,z')} - 1\right) L(z,z'), \qquad z \in \mathcal{X},$$

one gets the expected result

$$\varphi'(t) = \int \sum_{z',\, z'\sim z} \left(e^{\nabla F_t(z,z')} - 1 - \nabla F_t(z,z') e^{\nabla F_t(z,z')}\right) L(z,z') f_t(z) g_t(z)\, dm(z)$$
$$= -\int \sum_{z',\, z'\sim z} \zeta\left(e^{\nabla F_t(z,z')}\right) L(z,z')\, d\widehat{Q}_t^\gamma(z).$$

Applying again (55) with $V_t(z) = \sum_{z',z'\sim z} \zeta\left(e^{\nabla F_t(z,z')}\right) L(z,z')$, $z\in\mathcal{X}$, and using

$$\partial_t V_t(z) = \sum_{z',\, z'\sim z} \left(\frac{Lf_t(z')}{f_t(z)} - \frac{f_t(z')Lf_t(z)}{f_t^2(z)}\right) \zeta'\left(e^{\nabla F_t(z,z')}\right) L(z,z')$$
$$= \sum_{z',\, z'\sim z} e^{\nabla F_t(z,z')} \left(\frac{Lf_t(z')}{f_t(z')} - \frac{Lf_t(z)}{f_t(z)}\right) \nabla F_t(z,z')\, L(z,z')$$
$$= \sum_{z',\, z'',\, z\sim z'\sim z''} \nabla F_t(z,z')\, e^{\nabla F_t(z,z')} \left(e^{\nabla F_t(z',z'')} - 1\right) L(z,z')\, L(z',z'')$$
$$- \sum_{z',\, w',\, z'\sim z,\, w'\sim z} \nabla F_t(z,z')\, e^{\nabla F_t(z,z')} \left(e^{\nabla F_t(z,w')} - 1\right) L(z,z')\, L(z,w'),$$

and

$$\sum_{z',\, z'\sim z} e^{\nabla F_t(z,z')} \nabla V_t(z,z')\, L(z,z') = \sum_{z',\, z'',\, z\sim z'\sim z''} e^{\nabla F_t(z,z')} \zeta\left(e^{\nabla F_t(z',z'')}\right) L(z,z')\, L(z',z'')$$
$$- \sum_{z',\, w',\, z'\sim z,\, w'\sim z} e^{\nabla F_t(z,z')} \zeta\left(e^{\nabla F_t(z,w')}\right) L(z,z')\, L(z,w'),$$

one gets

$$\varphi''(t)$$
$$= -\int \Big[\sum_{z',\, w',\, z'\sim z,\, w'\sim z} \left[\zeta\left(e^{\nabla F_t(z,w')}\right) - \nabla F_t(z,z') \left(e^{\nabla F_t(z,w')} - 1\right)\right] e^{\nabla F_t(z,z')} L(z,z')\, L(z,w')$$
$$+ \sum_{z',\, z'',\, z\sim z'\sim z''} \left[\nabla F_t(z,z') \left(e^{\nabla F_t(z',z'')} - 1\right) - \zeta\left(e^{\nabla F_t(z',z'')}\right)\right] e^{\nabla F_t(z,z')} L(z,z')\, L(z',z'') \Big] d\widehat{Q}_t^\gamma(z)$$
$$= -\int \Big[\sum_{z',\, w',\, z'\sim z,\, w'\sim z} \left((\nabla F_t(z,w') - \nabla F_t(z,z')) - 1\right) e^{\nabla F_t(z,w') + \nabla F_t(z,z')} L(z,z')\, L(z,w')$$
$$+ \sum_{z',\, w',\, z'\sim z,\, w'\sim z} (\nabla F_t(z,z') + 1)\, e^{\nabla F_t(z,z')} L(z,z')\, L(z,w')$$
$$- \sum_{z',\, z'',\, z\sim z'\sim z''} (\nabla F_t(z,z') + 1)\, e^{\nabla F_t(z,z')} L(z,z')\, L(z',z'')$$
$$- \sum_{z',\, z'',\, z\sim z'\sim z''} \rho\left(e^{\nabla F_t(z,z')}, e^{\nabla F_t(z,z'')}\right) L(z,z')\, L(z',z'') \Big] d\widehat{Q}_t^\gamma(z)$$

The expected expression of $\varphi''(t)$ follows by symmetrization of the first sum in z' and w', and since $\sum_{w',\, w'\sim z} L(z,w') = -L(z,z)$. □

Proof of Lemma 3.4. Let $z, z' \in \mathcal{X}$ such that $z\sim z'$. One will only compute the expression of $\lim_{\gamma_k\to 0}\left(\gamma_k A_t^{\gamma_k}(z,z')\right)$ and similar calculations provide $\lim_{\gamma_k\to 0}\left(\gamma_k B_t^{\gamma_k}(z,z')\right)$. For any $\gamma > 0$, let

$$a_t^\gamma(z,y) := \widehat{Q}^\gamma(X_t = z | X_1 = y) = \int \nu_t^{\gamma w, y}(z)\, d\widehat{\pi}^\gamma(w|y),$$

and

$$\mathrm{a}_t^\gamma(z,z',y) := \int \gamma\alpha_t^\gamma(y,z,z',w)\, d\widehat{\pi}^\gamma_{\leftarrow}(w|y), \quad \text{with} \quad \alpha_t^\gamma(y,z,z',w) = \frac{P^\gamma_{1-t}(y,z)P_t^\gamma(z',w)}{P_1^\gamma(y,w)}.$$

Using equality (10) and since $P_1^\gamma f^\gamma(y) > 0$ for any $\gamma > 0$, one easily check that for any $\gamma > 0$,

$$\gamma A_t^\gamma(z,z') = \gamma\,\frac{P_t^\gamma f^\gamma(z')}{P_t^\gamma f^\gamma(z)} = \frac{\gamma\,\mathrm{a}_t^\gamma(z,z',y)}{a_t^\gamma(z,y)}.$$

From the expression (27) of $a_t(z,y)$ and since $\mathrm{supp}(\widehat{\pi}^{\gamma_k}(\cdot|y)) \subset \mathrm{supp}(\nu_0)$, one has

$$\left|a_t^{\gamma_k}(z,y) - a_t(z,y)\right| \leq \sup_{w\in\mathrm{supp}(\nu_0)} \left|\nu_t^{\gamma_k w,y}(z) - \nu_t^{0 w,y}(z)\right| + \sum_{w\in\mathrm{supp}(\nu_0)} \left|\widehat{\pi}^{\gamma_k}(w|y) - \widehat{\pi}^0(w|y)\right|.$$

Therefore, the weak convergence of $(\widehat{\pi}^{\gamma_k})_{k\in\mathbb{N}}$ to $\widehat{\pi}^0$ and Lemma 4.3 (4) imply

$$\lim_{\gamma_k\to 0} a_t^{\gamma_k}(z,y) = a_t(z,y). \tag{56}$$

Let us now consider the behaviour of $\gamma_k \mathrm{a}_t^{\gamma_k}(z,z',y)$ as γ_k goes to zero. Lemma 4.3 (3) provides the following Taylor expansion,

$$\gamma\alpha_t^\gamma(y,z,z',w) = \gamma^{d(y,z)+1+d(z',w)-d(y,w)} \cdot\left(r(y,z,z',w)\,\frac{d(y,w)!}{d(y,z)!d(z',w)!}\,(1-t)^{d(y,z)}t^{d(z',w)} + \gamma O(1)\right),$$

where $O(1)$ is a quantity uniformly bounded in t,γ,z,z',x,y. By the triangular inequality and since $z\sim z'$, one has $d(y,w) \leq d(y,z)+1+d(z',w)$, with equality if and only if $(z,z')\in[y,w]$. Therefore, one gets

$$\lim_{\gamma\to 0}\gamma\alpha_t^\gamma(y,z,z',w) = \alpha_t^0(y,z,z',w),$$

with

$$\alpha_t^0(y,z,z',w) := \mathbb{1}_{(z,z')\in[y,w]}\, r(y,z,z',w)\rho_t^{d(y,w)-1}(d(z,w)-1).$$

Moreover, Lemma 4.3 (1) and (2) ensures that for any $w\in\mathrm{supp}(\nu_0)$ and $y\in\mathrm{supp}(\nu_1)$,

$$\begin{aligned}\gamma\alpha_t^\gamma(y,z,z',w) &\leq \gamma^{d(y,z)+1+d(z',w)-d(y,w)}\\ &\quad\cdot\left((2S)^{d(y,z)+d(z',w)-d(y,w)}\max_{w\in\mathrm{supp}(\nu_0),y\in\mathrm{supp}(\nu_1)}\frac{(2S)^{d(y,w)}d(y,w)!}{I^{d(y,w)}} + O(1)\right)\\ &\leq O(1)\,(\gamma 2S)^{d(y,z)+d(z',w)+1-d(y,w)},\end{aligned}$$

where $O(1)$ is a constant independent of t,y,z,z',w. Therefore $\gamma\alpha_t^\gamma(y,z,z',w) \leq O(1)$ as soon as $\gamma < 1/(2S)$. As a consequence, for any $\gamma_k < 1/(2S)$, it holds

$$\begin{aligned}&\left|\gamma_k\mathrm{a}_t^{\gamma_k}(z,z',y) - \mathrm{a}_t(z,z',y)\right|\\ &\qquad\leq \sup_{w\in\mathrm{supp}(\nu_0)}\left|\gamma_k\alpha_t^{\gamma_k}(y,z,z',w) - \alpha_t^0(y,z,z',w)\right| + O(1)\sum_{w\in\mathrm{supp}(\nu_0)}\left|\widehat{\pi}^{\gamma_k}(w|y) - \widehat{\pi}^0(w|y)\right|,\end{aligned}$$

As γ_k goes to 0, this inequality with the weak convergence of $\widehat{\pi}^{\gamma_k}$ to $\widehat{\pi}^0$ implies

$$\lim_{\gamma_k\to 0}\gamma_k\,\mathrm{a}_t^{\gamma_k}(z,z',y) = \mathrm{a}_t(z,z',y),$$

Together with (56), it completes the proof of (28).

We now turn to the proof of (29). One will compute $\lim_{\gamma_k\to 0}\left(\gamma_k^2 A_t^{\gamma_k}(z,z'')\right)$ for $z,z''\in\mathcal{X}$ such that $d(z,z'')=2$ and the expression of $\lim_{\gamma_k\to 0}\left(\gamma_k^2 B_t^{\gamma_k}(z,z'')\right)$ follows from similar calculations. For any $y\in\mathcal{X}$ and any $t>0$, one has

$$\gamma^2 A_t^\gamma(z,z'')=\frac{\gamma^2 \mathrm{a}_t^\gamma(z,z'',y)}{a_t^\gamma(z,y)},$$

with

$$\gamma^2 \mathrm{a}_t^\gamma(z,z'',y):=\int \gamma^2\,\alpha_t^\gamma(y,z,z'',w)\,d\widehat{\pi}^\gamma_{\leftarrow}(w|y) \text{ and } \alpha_t^\gamma(y,z,z'',w)=\frac{P_{1-t}^\gamma(y,z)P_t^\gamma(z'',w)}{P_1^\gamma(y,w)}.$$

It remains to compute $\lim_{\gamma_k\to 0}\gamma_k^2\mathrm{a}_t^{\gamma_k}(z,z'',y)$ to prove (29). Lemma 4.3 (3) implies

$$\gamma^2\alpha_t^\gamma(y,z,z'',w)=\gamma^{d(y,z)+2+d(z',w)-d(y,w)}\cdot\left(r(y,z,z'',w)\,\frac{d(y,w)!}{d(y,z)!d(z'',w)!}\,(1-t)^{d(y,z)}t^{d(z'',w)}+\gamma O(1)\right),$$

where $O(1)$ is a quantity uniformly bounded in t,γ,z,z'',x,y. Since $d(y,w)\le d(y,z)+2+d(z'',w)$ with equality if and only if $(z,z'')\in[y,w]$, it follows that

$$\lim_{\gamma\to 0}\gamma^2\alpha_t^\gamma(y,z,z'',w)=\alpha_t^0(y,z,z'',w):=\mathbb{1}_{(z,z'')\in[y,w]}\,r(y,z,z'',w)\,\rho_t^{d(y,w)-2}(d(z,w)-2).$$

Moreover, Lemma 4.3 (1) and (2) gives that for any $w\in\mathrm{supp}(\nu_0)$ and $y\in\mathrm{supp}(\nu_1)$,

$$\gamma^2\alpha_t^\gamma(y,z,z'',w)\le O(1)\,(\gamma 2S)^{d(y,z)+d(z',w)+2-d(y,w)},$$

where $O(1)$ is a constant independent of t,y,z,z'',w. As above, the proof ends as γ_k goes to 0 from the inequality

$$\begin{aligned}&\left|\gamma_k^2\mathrm{a}_t^{\gamma_k}(z,z'',y)-\mathbb{a}_t(z,z'',y)\right|\\&\qquad\le\sup_{w\in\mathrm{supp}(\nu_0)}\left|\gamma_k^2\alpha_t^{\gamma_k}(y,z,z'',w)-\alpha_t^0(y,z,z'',w)\right|+O(1)\sum_{w\in\mathrm{supp}(\nu_0)}\left|\widehat{\pi}^{\gamma_k}(w|y)-\widehat{\pi}^0(w|y)\right|,\end{aligned}$$

for all $\gamma_k<1/(2S)$. □

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[2] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.*, 195(2):289–391, 2014.

[3] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.

[4] F. Bauer, P. Horn, Y. Lin, G. Lippner, D. Mangoubi, S.-T. Yau, et al. Li-yau inequality on graphs. *Journal of Differential Geometry*, 99(3):359–405, 2015.

[5] J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[6] A.I. Bonciocat and K.T. Sturm. Mass transportation and rough curvature bounds for discrete spaces. *J. Funct. Anal.*, 256(9):2944–2966, 2009.

[7] P. Caputo, P. Dai Pra, and G. Posta. Convex entropy decay via the bochner-bakry-emery approach. *Annales de l'I.H.P. Probabilités et statistiques*, 45(3):734–753, 2009.

[8] G. Conforti. A second order equation for schrödinger bridges with applications to the hot gas experiment and entropic transportation cost. *Probability Theory and Related Fields*, 174(1-2):1–47, 2019.

[9] M. Erbar, C. Henderson, G. Menz, and P. Tetali. Ricci curvature bounds for weakly interacting markov chains. *Electronic Journal of Probability*, 22, 2017.

[10] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.*, 206(3):997–1038, 2012.

[11] M. Erbar and J. Maas. Gradient flow structures for discrete porous medium equations. *Discrete & Continuous Dynamical Systems-A*, 34(4):1355–1374, 2014.

[12] M. Erbar, J. Maas, and P. Tetali. Discrete ricci curvature bounds for bernoulli-laplace and random transposition models. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 24, pages 781–800, 2015.

[13] M. Fathi and J. Maas. Entropic ricci curvature bounds for discrete interacting systems. *The Annals of Applied Probability*, 26(3):1774–1806, 2016.

[14] N. Gozlan, C. Roberto, P.-M. Samson, and P. Tetali. Displacement convexity of entropy and related inequalities on graphs. *Probability Theory and Related Fields*, 160(1-2):47–94, 2014.

[15] N. Gozlan, C. Roberto, P.-M. Samson, and P. Tetali. Kantorovich duality for general transport costs and applications. *J. Funct. Anal.*, 273(11):3327–3405, 2017.

[16] N. Gozlan, C. Roberto, P.-M. Samson, and P. Tetali. Transport proofs of some discrete variants of the prékopa-leindler inequality. *arXiv preprint arXiv:1905.04038*, 2019.

[17] E. Hillion. $W_{1,+}$-interpolation of probability measures on graphs. *Electron. J. Probab.*, 19:no. 92, 29, 2014.

[18] E. Hillion. Interpolation of probability measures on graphs. In *Convexity and concentration*, volume 161 of *IMA Vol. Math. Appl.*, pages 3–32. Springer, New York, 2017.

[19] B. Klartag, G. Kozma, P. Ralli, and P. Tetali. Discrete curvature and abelian groups. *Canadian Journal of Mathematics*, 68(3):655–674, 2016.

[20] B. Klartag and J. Lehec. Poisson processes and a log-concave Bernstein theorem. *Studia Math.*, 247(1):85–107, 2019.

[21] C. Léonard. From the schrödinger problem to the monge–kantorovich problem. *Journal of Functional Analysis*, 262(4):1879–1920, 2012.

[22] C. Léonard. Some properties of path measures. In *Séminaire de Probabilités XLVI*, pages 207–230. Springer, 2014.

[23] C. Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *Discrete & Continuous Dynamical Systems-A*, 34(4):1533–1574, 2014.

[24] C. Léonard. Lazy random walks and optimal transport on graphs. *Ann. Probab.*, 44(3):1864–1915, 2016.

[25] C. Léonard. On the convexity of the entropy along entropic interpolations. In *Measure Theory in Non-Smooth Spaces*, pages 194–242. Sciendo Migration, 2017.

[26] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[27] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

[28] K. Marton. Bounding $\overline{d}$-distance by informational divergence: a method to prove measure concentration. *Ann. Probab.*, 24(2):857–866, 1996.

[29] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[30] A. Mielke. Geodesic convexity of the relative entropy in reversible markov chains. *Calculus of Variations and Partial Differential Equations*, 48(1-2):1–31, 2013.

[31] T. Mikami. Monges problem with a quadratic cost by the zero-noise limit of h-path processes. *Probability theory and related fields*, 129(2):245–260, 2004.

[32] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256(3):810–864, 2009.

[33] Y. Ollivier. A visual introduction to Riemannian curvatures and some discrete generalizations. In *Analysis and geometry of metric measure spaces*, volume 56 of *CRM Proc. Lecture Notes*, pages 197–220. Amer. Math. Soc., Providence, RI, 2013.

[34] Y. Ollivier and C. Villani. A curved Brunn-Minkowski inequality on the discrete hypercube. *Siam Journal on Discrete Mathematics*, 26(3):983–996, 2012.

[35] P.-M. Samson. Transport-entropy inequalities on locally acting groups of permutations. *Electronic Journal of Probability*, 22, 2017.

[36] K.T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.

[37] K.T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.

[38] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Publications Mathématiques de l'Institut des Hautes Etudes Scientifiques*, 81(1):73–205, 1995.

[39] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.

[40] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.

P.-M. Samson, LAMA, Univ Gustave Eiffel, UPEM, Univ Paris Est Creteil, CNRS, F-77447 Marne-la-Vallée, France

E-mail address: paul-marie.samson@univ-eiffel.fr