



**HAL**  
open science

# ENTROPIC CURVATURE ON GRAPHS ALONG SCHRÖDINGER BRIDGES AT ZERO TEMPERATURE

Paul-Marie Samson

► **To cite this version:**

Paul-Marie Samson. ENTROPIC CURVATURE ON GRAPHS ALONG SCHRÖDINGER BRIDGES AT ZERO TEMPERATURE. *Probability Theory and Related Fields*, 2022, 10.1007/s00440-022-01167-4 . hal-02504530v6

**HAL Id: hal-02504530**

**<https://hal.science/hal-02504530v6>**

Submitted on 3 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 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 [25, 27, 26]. 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 [29, 11, 12], 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 [16]) associated to the so-called convex-hull method by Talagrand [41].

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 (see Definition 1.1), derived from C. Léonard approach [25, 27, 26], and dealing with Schrödinger bridges at zero temperature.

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 [37] and [16] 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 theorem of the present paper, Theorem 3.5 (with Lemma 3.1), is derived from this procedure, which consists of decreasing a temperature parameter  $\gamma$  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 Theorem 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. In a forthcoming paper, one will give sufficient geometric conditions to reach entropic curvature property on non-specific graphs from Theorem 3.5.

---

*Date:* October 3, 2022.

*2020 Mathematics Subject Classification.* 60E15, 32F32 and 39A12.

*Key words and phrases.* Displacement convexity property, Ricci curvature, graphs, Bernoulli Laplace model, discrete hypercube, 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	2
2. Main results : examples of entropic curvature bounds along Schrödinger bridges on graphs	11
2.1. The lattice $\mathbb{Z}^n$ endowed with the counting measure	14
2.2. The complete graph	15
2.3. Product measures on the discrete hypercube	16
2.4. The circle $\mathbb{Z}/N\mathbb{Z}$ endowed with a uniform measure	21
2.5. The Bernoulli-Laplace model	21
3. Proof of the main results	24
3.1. Strategy of proof, general statements to get entropic curvature results	24
3.2. Application to specific examples of graphs	31
4. Appendix A : Basic lemmas	42
5. Appendix B : Proofs of Lemmas 3.1, 3.2, 3.3, and 3.4	50
References	56

## 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  $\sigma$ -finite measures on  $\mathcal{Y}$  and  $\mathcal{P}(\mathcal{Y})$  the set of all probability measures on  $\mathcal{Y}$ .

Let  $(X, d)$  be a geodesic space equipped with a reference measure  $m \in \mathcal{M}_+(X)$ . According to Lott-Sturm-Villani theory of curvature on geodesic spaces [28, 39, 40, 43], a lower bound  $K \in \mathbb{R}$  on the entropic curvature of the space  $(X, d, m)$  is characterized by a *K-convexity property* of the *relative entropy* along *constant speed geodesics* of the *Wasserstein space*  $(\mathcal{P}_2(X), W_2)$ . 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 probability measure  $r \in \mathcal{P}(\mathcal{Y})$ , also called Kullback-Leibler distance between  $q$  and  $r$ , is given by

$$H(q|r) := \int_{\mathcal{Y}} \log(dq/dr) dq \quad \in [0, \infty],$$

if  $q$  is absolutely continuous with respect to  $r$  and  $H(q|r) := +\infty$  otherwise. As explained in [24], this definition extends to unbounded measures  $r \in \mathcal{M}_+(\mathcal{Y})$  as follows. Since  $r$  is a  $\sigma$ -finite measure, there exists some measurable function  $w : \mathcal{Y} \rightarrow [0, +\infty)$  such that

$$z_w := \int e^{-w} dr < \infty.$$

Define the probability measure  $r_w = \frac{e^{-w}}{z_w} r$ . Then the definition of  $H(q|r)$  is given for all  $q \in \mathcal{P}(\mathcal{Y})$  such that  $\int w dq < +\infty$  by

$$H(q|r) = H(q|r_w) - \int w dq - \log z_w \quad \in (-\infty, +\infty].$$

According to [24], this definition makes sense since the right-hand side does not depends on the function  $w$  satisfying  $z_w < \infty$  and  $\int w dq < +\infty$ . We refer to [24] for more details and properties about this definition of relative entropy with unbounded  $\sigma$ -finite measures. Let  $\mathcal{P}_2(X)$  denote the space of

probability measures with second moment, and let  $W_2$  be the *Wasserstein distance of order 2* on  $\mathcal{P}_2(\mathcal{X})$ : namely, for any  $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$ ,

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

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  $\nu_0$  and second marginal  $\nu_1$  (also called *transference plans* from  $\nu_0$  to  $\nu_1$ ). A path  $(\nu_t)_{t \in [0,1]}$  in  $\mathcal{P}_2(\mathcal{X})$  is a *constant speed  $W_2$ -geodesic* from  $\nu_0$  to  $\nu_1$  if for all  $0 \leq s < t \leq 1$ ,  $W_2(\nu_s, \nu_t) = (t - s)W_2(\nu_0, \nu_1)$ . 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  $\nu_0$  to  $\nu_1$  such that for all  $t \in [0, 1]$ ,

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

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 [31]. 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  $CD(K, \infty)$ :  $\text{Ricc} + \text{Hess}(V) \geq K$  (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 [7], Ollivier-Villani on the discrete cube [36], Erbar-Maas [29, 11, 12], Mielke [32], Léonard [25, 27, 26], Hillion [19, 20] and Gozlan-Roberto-Samson-Tetali [15].

This paper concerns Léonard entropic approach of curvature in discrete setting, from which we also recover results from [15] and [19]. In discrete spaces, several other notions of curvature have already been studied which are not considered in this paper : the coarse Ricci curvature [34, 35], the Bochner-Bakry-Emery approach with the (Bochner) curvature [8, 21] and the curvature dimension or exponential curvature dimension inequality [4].

For  $m$  as unique invariant 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 [29, 11, 12]. 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  $\pi$  as in the definition (1) of  $W_2^2$ . Erbar-Maas entropic Ricci curvature definition satisfies a tensorization property for product of graphs that allows to consider high dimensional spaces [11]. 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 [29, 11], the Bernoulli-Laplace model, the random transposition model [13, 14], birth and death processes, zero-range processes [14], Cayley graphs of non-abelian groups, weakly interacting Markov chains such as the Ising model [10]. The main strategy of all this papers is to prove an equivalent criterion of Erbar-Maas entropic curvature given in [11], 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 [33] and then C. Léonard [23, 25,

26, 27] 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

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

As explained below, each of these constant speed  $W_1$ -geodesics 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  $\gamma$  goes to zero. Given two probability measures  $\nu_0$  and  $\nu_1$ , this constant speed  $W_1$ -geodesic selected from this cooling down process (or slowing down procedure) is unique. According to its construction, we call it *Schrödinger brige at zero temperature* and we denote it  $(\widehat{Q}_t)_{t \in [0,1]}$  throughout this paper ( $\widehat{Q}_0 = \nu_0$  and  $\widehat{Q}_1 = \nu_1$ ). For  $x, y \in \mathcal{X}$ , one denotes  $(Q_t^{x,y})_{t \in [0,1]}$  the Schrödinger brige at zero temperature from the Dirac measure  $\delta_x = Q_0^{x,y}$  to the Dirac measure  $\delta_y = Q_1^{x,y}$ . Actually the bridge  $(\widehat{Q}_t)_{t \in [0,1]}$  is a mixture of Schrödinger briges from Dirac measures on the support of  $\nu_0$  to Dirac measures on the support of  $\nu_1$ , according to a selected transference plan denoted  $\widehat{\pi} \in \Pi(\nu_0, \nu_1)$ , that achieves  $W_1(\nu_0, \nu_1)$ . Namely, one has for any  $z \in \mathcal{X}$

$$(3) \quad \widehat{Q}_t(z) = \iint Q_t^{x,y}(z) d\widehat{\pi}(x, y), \quad \text{with} \quad \iint d(x, y) d\widehat{\pi}(x, y) = W_1(\nu_0, \nu_1).$$

Observe that the set of minimizers of  $W_1(\nu_0, \nu_1)$ , also called  $W_1$ -optimal couplings of  $\nu_0$  and  $\nu_1$ , is a convex set that is not necessarily reduced to a singleton. However, according to Leonard's paper [26], we know that given  $\nu_0, \nu_1, \widehat{\pi}$  is uniquely determined, as a minimizer of a strictly convex optimization problem (see [26, Result 0.2]). In our setting of property (2) on graphs, the curvature term  $W_2^2(\nu_0, \nu_1)$  is also replaced by some transport cost  $C_t(\widehat{\pi})$  that depends on the selected  $W_1$ -minimizer  $\widehat{\pi} \in \Pi(\nu_0, \nu_1)$ , and 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  $(\widehat{Q}_t)_{t \in [0,1]}$  from  $\nu_0$  to  $\nu_1$ , satisfies for any  $t \in (0, 1)$ ,*

$$(4) \quad H(\widehat{Q}_t|m) \leq (1-t)H(\nu_0|m) + tH(\nu_1|m) - \frac{t(1-t)}{2}C_t(\widehat{\pi}).$$

For some of the graphs studied in this paper, the cost  $C_t(\widehat{\pi})$  is bigger than  $K \left( \iint d(x, y) d\widehat{\pi}(x, y) \right)^2 = K W_1(\nu_0, \nu_1)^2$  for any  $t \in (0, 1)$  with  $K > 0$ . Such a property is also a consequence of Erbar-Maas positive entropic curvature since  $\mathcal{W}_2^2 \geq 2W_1^2$  (see [11, Proposition 2.12]). However, their convexity property of entropy deals with  $W_2$ -geodesics on  $\mathcal{P}(\mathcal{X})$ , whereas property (4) deals with  $W_1$ -geodesics. As a definition in this paper, the largest constant  $K \in \mathbb{R}$  so that (4) holds with  $C_t(\widehat{\pi}) = K W_1(\nu_0, \nu_1)^2$  for any  $\nu_0, \nu_1 \in \mathcal{P}_b(\mathcal{X})$  and any  $t \in (0, 1)$  is called, if it exists, *the  $W_1$ -entropic curvature* of the space  $(\mathcal{X}, d, m, L)$ .

Given a non negative cost function  $c : \mathbb{N} \rightarrow \mathbb{R}^+$ , let us denote

$$T_c(\widehat{\pi}) := \iint c(d(x, y)) d\widehat{\pi}(x, y)$$

and  $T_2 := T_c$  for the square function  $c(d) = d^2, d \geq 0$ . For some graphs in this paper, in order to compare our results with the  $W_2^2$  cost that appears in (2) on geodesics spaces, we prove that  $C_t(\widehat{\pi}) \geq K T_{c_2}(\widehat{\pi})$  with  $K \geq 0$ , where one denotes by  $c_2$  any universal cost function (independent of any characteristic of the graph) satisfying

$$\frac{d(d-1)}{2} \leq c_2(d) \leq d^2$$

and which is equivalent to the square function at infinity

$$\lim_{d \rightarrow \infty} \frac{c_2(d)}{d^2} = 1.$$

For such a cost function, one has for any  $\varepsilon \in (0, 1)$  and any  $d \in \mathbb{N}$ ,

$$c_2(d) \geq (1 - \varepsilon)d(d-1) - \alpha(\varepsilon)d,$$

where  $\alpha$  is the non-negative function given by  $\alpha(\varepsilon) := \sup_{k \in \mathbb{N}^*} \left\{ (1 - \varepsilon)(k-1) - \frac{c_2(k)}{k} \right\}$  ( $\alpha(\varepsilon) = 0$  for  $\varepsilon \in (1/2, 1)$ ). It follows that  $T_{c_2}(\widehat{\pi})$  is controlled by the Wasserstein distances as follows, for any  $\varepsilon \in (0, 1)$

$$T_{c_2}(\widehat{\pi}) \geq \sup_{\varepsilon \in (0,1)} \left\{ (1 - \varepsilon)W_2^2(\nu_0, \nu_1) - [(1 - \varepsilon) + \alpha(\varepsilon)]W_1(\nu_0, \nu_1) \right\} \geq \frac{1}{2} \left( W_2^2(\nu_0, \nu_1) - W_1(\nu_0, \nu_1) \right) \geq 0.$$

Therefore, the cost  $T_{c_2}(\widehat{\pi})$  can be interpreted as a discrete analogue of the cost  $W_2^2(\nu_0, \nu_1)$  in the usual  $K$ -convexity property (2) on geodesic spaces. As a definition in this paper, *the  $T_2$ -entropic curvature* of the space  $(X, d, m, L)$  is the largest constant  $K \in \mathbb{R}$  so that there exists a cost  $c_2$  satisfying the above conditions and such that (4) holds with  $C_t(\widehat{\pi}) = K T_{c_2}(\widehat{\pi})$  for any  $\nu_0, \nu_1 \in \mathcal{P}_b(X)$  and any  $t \in (0, 1)$ .

Due to the abstract definition of the cost  $\mathcal{W}_2^2$  with a discrete analogue of Benamou-Brenier formula, we don't know how to compare  $\mathcal{W}_2^2$  with costs involving transference plans and the discrete structure of the graph such as  $T_{c_2}$  or any other proposed costs  $C_t$  of this paper, excepted with  $W_1^2$  for which  $\mathcal{W}_2^2 \geq 2W_1^2$ . As a consequence, it is still a challenging problem to reach most of the results of the present paper from Erbar-Maas approach of entropic curvature on discrete spaces.

According to the property of the function  $c_2$ , the cost  $c_2(d(x, y)) = 0$  if  $x$  and  $y$  are neighbours. Therefore the transport-cost  $T_{c_2}$  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 [7] in their definition of rough (approximate) lower curvature.

For the graph with positive  $W_1$ -entropic curvature studied in this paper (the complete graph, the discrete hypercube and the Bernoulli-Laplace model), one may bound from below the cost  $C_t(\widehat{\pi})$  by different symmetric versions of weak transport cost denoted by  $\widetilde{T}_t(\widehat{\pi})$  and bounded from below by the so-called weak optimal transport costs introduced in the paper [16]. Weak transport-entropy inequalities were introduced by K. Marton in the seminal work [30] in order to get refined concentration properties for product measure, related to concentration's results derived from the so-called "Convex hull method" by M. Talagrand [41]. It was pushed forward in [15] that these costs are related to displacement convexity property of entropy along  $W_1$ -geodesics in the case of the complete graph and of the discrete hypercube. From the present paper, we learn that same observation extends to models without product structure with different kind of weak transport costs, like for the Bernoulli-Laplace model. Actually, our approach seems very efficient to prove (weak) transport-entropy inequalities since we discover new ones and get improved versions of the known ones.

As a guideline for other graphs, we present in this paper for the discrete hypercube and the Bernoulli Laplace model how to easily reach modified logarithmic Sobolev inequalities from the  $C$ -displacement convexity property (4). The strategy is to analyse the  $C$ -displacement convexity property (4) as  $t$  goes to zero when the cost  $C_t(\widehat{\pi})$  is lower bounded by some weak transport costs  $\widetilde{T}_t(\widehat{\pi})$ . It may provide different kinds of modified logarithmic Sobolev inequalities, depending on the model and the structure

of weak transport cost  $\widetilde{T}_t(\widehat{\pi})$ . Contrarily to the Erbar-Maas entropic curvature approach, connections and comparisons with other known modified logarithmic Sobolev inequalities with optimal constants are not always easy to handle. It still remains a challenge to improve our strategy or find other ways to reach modified logarithmic Sobolev inequalities from the use of Schrödinger bridges at zero temperature in discrete spaces.

Applying usual duality arguments, the  $C$ -displacement convexity property (4) also implies new kinds of curved Prékopa-Leindler inequalities, as opposed to Erbar-Maas entropic approach of curvature due to the abstract definition of  $\mathcal{W}_2$  (see Theorem 2.1).

Following the line of the paper [15], a tensorization property of the  $C$ -displacement convexity property holds involving Knothe-Rosenblatt coupling (see Theorem 2.2).

In the present 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 2.3), the discrete hypercube endowed with product probability measures (see Theorem 2.5), the discrete circle endowed with uniform measure (see Theorem 2.6), the complete graph (see Theorem 2.4), the Bernoulli-Laplace model (see Theorem 2.7). For all these graphs, one gets a non-negative lower bound for their  $W_1$  or  $T_2$ -entropic curvature.

In a forthcoming paper, starting from the key Theorem 3.5, one will give sufficient geometric conditions on balls of radius 2, available on any graph space  $(\mathcal{X}, d, m, L)$ , that give lower bounds on  $W_1$  or  $T_2$ -entropic curvature. Other examples of graphs will be studied, like the random transposition model on the symmetric group  $S_n$  (for which the  $W_1$ -entropic curvature is lower bounded by  $4/n^2$ ) or the multinomial distribution on the set  $\mathcal{X} := \{(x_1, \dots, x_d) \in \mathbb{N}^d \mid x_1 + \dots + x_d = N\}$ . Examples of graphs with negative entropic curvature like trees and also measures with interaction potential will be also considered.

For more comprehension, let us briefly explain the *slowing down procedure* in its original continuous setting before considering discrete spaces. Let  $R^\gamma$  be the law of a reversible Brownian motion with diffusion coefficient  $\gamma > 0$  on the set  $\Omega$  of continuous paths from  $[0, 1]$  to  $\mathcal{X} = \mathbb{R}^d$ . The coefficient  $\gamma$  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  $\Delta$  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  $F : \Omega \rightarrow \mathbb{R}$  with respect to  $Q$ , one notes

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

The result by T. Mikami [33] or C. Léonard [23] is the following: for any absolutely continuous measures  $\nu_0, \nu_1 \in \mathcal{P}_2(\mathcal{X})$ , for any sequences  $(\gamma_\ell)_{\ell \in \mathbb{N}}$  of temperature parameters going to zero,

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

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

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

As interpretation, the measure  $\widehat{Q}^\gamma$  is the law of the process with configuration  $\widehat{Q}_0^\gamma = \nu_0$  at time  $t = 0$  and  $\widehat{Q}_1^\gamma = \nu_1$  at time  $t = 1$ , which is the closest in some entropic meaning, to a reversible Brownian motion with diffusion coefficient  $\gamma$ . As a result (see [33, 23]), the sequence of minimizers  $(\widehat{Q}^{\gamma_\ell})_{\ell \in \mathbb{N}}$  converges to a single measure  $\widehat{Q} \in \mathcal{P}(\Omega)$ . For any  $t \in [0, 1]$ , let  $Q_t^\gamma := \widehat{Q}_t^\gamma$  and  $\nu_t := \widehat{Q}_t$ . By definition,  $(Q_t^\gamma)_{t \in [0, 1]}$  is a Schrödinger bridge from  $\nu_0$  to  $\nu_1$  at fixed temperature  $\gamma$ , and as a main result, as  $\gamma_\ell$  goes to zero, the limit path  $(\nu_t)_{t \in [0, 1]}$ , is a  $W_2$ -geodesic from  $\nu_0$  to  $\nu_1$  (see [25]). 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  $(Q_t^\gamma)_{t \in [0, 1]}$  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 [9].

Let us present the discrete analogue of this approach due to C. Léonard [25, 27, 26]. From now on, the space  $\mathcal{X}$  is a countable set endowed with the  $\sigma$ -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  $\Omega$  is endowed with the  $\sigma$ -algebra  $\mathcal{F}$  generated by the cylindrical sets. In all the paper, by convention, a sum indexed by an empty set is equal to zero.

According to C. Léonard's paper [26], 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 write  $x \sim y$  if  $(x, y) \in E$ . We assume that any vertex  $x \in \mathcal{X}$  has a finite number of neighbours  $d_x$  and that  $\sup_{x \in \mathcal{X}} d_x = d_{\max} < \infty$ . We note  $V(x)$  the set of neighbours of  $x$ . 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 [26, Hypothesis 2.1]), this means that for any  $x, y \in \mathcal{X}$ ,

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

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*  $\alpha$  of length  $\ell \in \mathbb{N}$  joining two vertices  $x$  and  $y$  is a sequence of  $\ell + 1$  neighbours  $\alpha = (z_0, \dots, z_\ell)$  so that  $z_0 = x$  and  $z_\ell = y$ . In the sequel, we note  $z \in \alpha$  if there exists  $i \in \{0, \dots, \ell\}$  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  $y$ ,

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

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

$$\begin{cases} L^\gamma(x, y) := \gamma^{d(x, y)} L(x, y) & \text{for } x \neq y, \\ L^\gamma(x, x) := - \sum_{y \in \mathcal{X}, y \neq x} L^\gamma(x, y), \end{cases}$$



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  $L^\gamma$  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  $\delta_x$  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  $f : \mathcal{X} \rightarrow \mathbb{R}$  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 :

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

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

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

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 [26], Theorem 2.1 ensures that under some assumptions (see [26, 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  $R^\gamma$ , where  $f^\gamma$  and  $g^\gamma$  are non-negative functions on  $\mathcal{X}$  satisfying the following so-called *Schrödinger system*

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

Since  $f^\gamma$  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  $\nu_0$  and  $\nu_1$  have finite support, then the Schrödinger system (7) implies that  $f^\gamma$  and  $g^\gamma$  have also finite support.

According to [27, Theorem 6.1.4.], from the Markov property, the law at time  $t$  of the Schrödinger bridge at fixed temperature  $\gamma$ ,  $\widehat{Q}_t^\gamma$ , is given by: for any  $z \in \mathcal{X}$ ,

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

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

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

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

$$(10) \quad \widehat{Q}_t^\gamma(z) = \iint Q_t^{\gamma,x,y}(z) d\widehat{\pi}^\gamma(x, y).$$

Actually, for any  $x, y \in \mathcal{X}$ ,  $(Q_t^{\gamma,x,y})_{t \in [0,1]}$  is the Schrödinger bridge joining the Dirac measures  $\delta_x$  and  $\delta_y$ . 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  $\text{supp}(\nu)$  denote the support of the measure  $\nu$ ,  $\text{supp}(\nu) := \{x \in \mathcal{X} \mid \nu(x) > 0\}$ . 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  $x \in \text{supp}(\nu_0)$ ,

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

and for any  $y \in \text{supp}(\nu_1)$ ,

$$(11) \quad \widehat{\pi}_\leftarrow^\gamma(x|y) := \frac{f^\gamma(x) P_1^\gamma(y, x)}{P_1^\gamma f^\gamma(y)}.$$

In order to fulfill 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 [26] 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

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

The main goal of this paper is to prove a convexity property for the function  $t \in [0, 1] \mapsto H(\widehat{Q}_t|m)$  by applying the *slowing down procedure*. Our strategy is first to differentiate twice at positive temperature  $\gamma > 0$  the function  $t \in [0, 1] \mapsto H(\widehat{Q}_t^\gamma|m)$  using backward equations for the Markov process. Then as a main contribution of this paper, we analyse the behavior of the second derivative of this functions as the temperature  $\gamma$  goes to zero (see Theorem 3.5). Considering different examples of graphs, any lower bound of this limit second derivative gives a convexity property of type (4).

We want this strategy to hold for a large class of graphs  $(\mathcal{X}, d, m, L)$ , with possibly infinite set of vertices  $\mathcal{X}$ . Mainly in order to justify the lower bounds on the second derivative as  $\gamma$  goes to zero, we make the following assumptions.

- The measure  $m$  is bounded,

$$(12) \quad \sup_{x \in \mathcal{X}} m(x) < \infty, \quad \text{and} \quad \inf_{x \in \mathcal{X}} m(x) > 0.$$

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

$$(13) \quad \sup_{x \in \mathcal{X}} |L(x, x)| \leq S,$$

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

$$(14) \quad \inf_{x, y \in \mathcal{X}, x \sim y} L(x, y) \geq I.$$

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

$$(15) \quad \sum_{y \in \mathcal{X}} \gamma_o^{d(x,y)} < \infty.$$

All these assumptions are obviously satisfy if  $\mathcal{X}$  is finite. One may also consider any infinite graph  $\mathcal{X}$  with bounded degree  $d_{\max}$  endowed with the counting measure  $m_0$ , which is reversible with respect to the generator  $L_0$  given by  $L_0(x, y) = 1$  for  $x \sim y$ ,  $L_0(x, x) = -d_x$ . On such graphs  $(\mathcal{X}, d, m_0, L_0)$ , a condition dealing with the geometry of balls of radius 2 will be given in a forthcoming paper to get lower bounds on the  $T_2$ -entropic curvature.

Unfortunately, the above assumptions are not fulfilled by example for the  $M/M/\infty$  process on  $\mathbb{N}$  with Poisson stationary measure. For such processes, the same strategy is expected to provide lower bounds on entropic curvature adapting proofs by the known specific expression of the Markov semi-group. A next challenge is to weak the assumptions of this paper for other specific classes of processes.

One of the main assets of Hypothesis (13) is to provide a simple expression for the semi-group  $(P_t^\gamma)_{t \geq 0}$ , namely

$$(16) \quad P_t^\gamma := e^{t\gamma L} = \sum_{k \in \mathbb{N}} \frac{(t\gamma)^k}{k!} L^k.$$

From this expression, on may simply derive a rather expression of Schrödinger bridges at zero temperature between Dirac measures. Namely, given  $x, y \in \mathcal{X}$ , as condition (13) holds, Lemma 4.4 (iv) gives the limit of the path  $(Q_t^{\gamma, x, y})_{t \in [0, 1]}$  defined by (9), namely for any  $z \in \mathcal{X}$ ,

$$(17) \quad \lim_{\gamma \rightarrow 0} Q_t^{\gamma, x, y}(z) = Q_t^{x, y}(z) := \mathbb{1}_{[x, y]}(z) r(x, z, z, y) \rho_t^{d(x, y)}(d(x, z)),$$

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

$$(18) \quad r(x, z, v, y) = \frac{L^{d(x, z)}(x, z) L^{d(v, y)}(v, y)}{L^{d(x, y)}(x, y)},$$

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

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

with the binomial coefficient  $\binom{d}{k} := \frac{d!}{k!(d-k)!}$ . Obviously one has  $Q_0^{x, y} = \delta_x$  and  $Q_1^{x, y} = \delta_y$ . Moreover, observe that for any  $t \in (0, 1)$ , the support of  $Q_t^{x, y}$  is  $[x, y]$ , the set of points on discrete geodesics from  $x$  to  $y$ . Observe that this limit Schrödinger bridge  $(Q_t^{x, y})_{t \in [0, 1]}$  is consistent with the metric graph structure. This is not surprising. As the temperature  $\gamma$  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. Roughly speaking,  $Q_t^{x, y}$  can be interpreted as the law of a process which is forced to go from  $x$  at time 0 to  $y$  at time 1 and that does not want to move or to jump too much between time 0 and 1. Therefore this process follows the geodesics of the graph from  $x$  to  $y$ .

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

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

If  $N_t$  and  $\Gamma = (\Gamma_0, \dots, \Gamma_d)$  are independent then  $Q_t^{x, y}$  is the law of  $\Gamma_{N_t}$ .

Let us come back to the behavior of the Schrödinger bridges  $(\widehat{Q}_t^\gamma)_{t \in [0, 1]}$  as  $\gamma$  goes to zero. Assume  $\nu_0$  and  $\nu_1$  have finite support. C. Léonard [26, Theorem 2.1] proves that given a positive sequence  $(\gamma_\ell)_{\ell \in \mathbb{N}}$  with  $\lim_{\ell \rightarrow \infty} \gamma_\ell = 0$ , the sequence of optimal Schrödinger minimizers  $(\widehat{Q}^{\gamma_\ell})_{\ell \in \mathbb{N}}$  converges to

a single probability measure  $\widehat{Q} \in \mathcal{P}(\Omega)$  for the narrow convergence, provided Hypothesis 2.1 holds. In this paper, the measure  $\widehat{Q}$  is named as *the limit Schrödinger problem optimizer at zero temperature, between  $\nu_0$  and  $\nu_1$* . In the framework of this work, choosing two probability measures  $\nu_0$  and  $\nu_1$  with finite supports, Hypothesis 2.1 in [26] is reduced to the following assumption (see condition  $(\mu)$  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 4.4 (vi), this assumption is fulfilled thanks to (13) since  $P_1^\gamma(x, y) > 0$  for any  $x, y \in \mathcal{X}$  and  $\gamma > 0$ .

As a main result of [26, Theorem 2.1], the measure  $\widehat{Q}$  is also a solution of the following *dynamic Monge-Kantorovich problem* :

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

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

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

and similarly to the continuous case,  $\widehat{\pi}$  is a  $W_1$ -optimal coupling of  $\nu_0$  and  $\nu_1$ .

The weak convergence of  $(\widehat{Q}^{\ell})_{\ell \in \mathbb{N}}$  to  $\widehat{Q}$  also provides the convergence of  $(\widehat{Q}_t^{\ell})_{\ell \in \mathbb{N}}$  to  $\widehat{Q}_t$ , and (10) implies (3). According to its construction, this bridge is called *Schrödinger bridge at zero temperature* from  $\nu_0$  to  $\nu_1$ . Observe that for any  $t \in (0, 1)$ , the support of  $\widehat{Q}_t$  only depends on the support of the optimal coupling  $\widehat{\pi}$  of  $\nu_0$  and  $\nu_1$ ,

$$(19) \quad \text{supp}(\widehat{Q}_t) = \bigcup_{(x,y) \in \text{supp}(\widehat{\pi})} [x, y].$$

As a main result, C. Leonard proves that with hypothesis (6), the path  $(\widehat{Q}_t)_{t \in [0,1]}$  is a constant speed  $W_1$ -geodesic (see [26, Theorem 3.15]): for any  $0 \leq s \leq t \leq 1$ ,

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

Actually, from the above interpretation of the measure  $Q_t^{x,y}$  as the law of  $\Gamma_{N_t}$  where  $\Gamma$  is a random geodesic from  $x$  to  $y$ , independent of a binomial random variable  $N_t$  with parameters  $t \in [0, 1]$  and  $d(x, y)$ , one proves that any bridge  $(\widehat{Q}_t)_{t \in [0,1]}$  defined by (3) is a  $W_1$ -geodesic, as soon as  $\widehat{\pi}$  is a  $W_1$ -optimal coupling of  $\nu_0$  and  $\nu_1$ . The proof of this result is the same as the one of [15, Proposition 2.2].

## 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  $T_2$ -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 (4) of the relative entropy along *Schrödinger bridges at zero temperature*  $(\widehat{Q}_t)_{t \in [0,1]}$ , derived from the slowing down procedure.

As in the paper [15],  $C$ -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.

As mentioned before, our approach is efficient to reach new transport-entropy inequalities, transport cost well suited to get new concentration properties, using known connections between transport-entropy inequalities and concentration properties pushed forward in [16]. Although Erbar-Maas method does not allow to recover such concentration results on graphs, both approaches imply bounds on the so-called subgaussian constant  $\sigma^2(\mathcal{X})$  of the graph (see [6]), namely  $\sigma^2(\mathcal{X}) \leq 1/K$  if the  $W_1$ -entropic curvature is bounded from below by  $K > 0$ .

As a guideline for other graphs, connexions between  $C$ -displacement convexity properties along Schrödinger bridges at zero temperature and modified log-Sobolev inequalities are explained only in the case of the discrete hypercube or the Bernoulli-Laplace Model (see comments (d) after Theorem 2.5 and after Theorem 2.7). Even if this global strategy does not allow to recover exactly some known modified log-Sobolev inequality for the Bernoulli-Laplace model, preliminary computations look promising to apply it for measures on graphs with interaction potentials. A challenge is to improve it for that purpose.

New Prékopa-Leindler type of inequalities are also a straightforward dual consequence of the  $C$ -displacement convexity properties (4). Here is a general statement that applies for each of the discrete spaces  $(X, d, m, L)$  studied in this paper and presented next.

**Theorem 2.1.** *On a discrete space  $(X, d, m, L)$ , assume that the relative entropy satisfies a  $C$ -displacement convexity property (see Definition 1.1) with  $C = (C_t)_{t \in (0,1)}$  given by : for any  $\nu_0, \nu_1 \in \mathcal{P}_b(X)$*

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

where  $\widehat{\pi} = \widehat{Q}_{0,1}$ , and  $\widehat{Q}$  is the limit Schrödinger problem optimizer between  $\nu_0$  and  $\nu_1$ . Then, the next property holds for all  $t \in (0, 1)$ . If  $f, g, h$  are functions on  $X$  satisfying

$$(1-t)f(x) + tg(y) \leq \int h dQ_t^{x,y} + \frac{t(1-t)}{2} c_t(x, y), \quad \forall x, y \in X,$$

then

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

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

Following the paper [16, section 3.2], a tensorization property holds for the  $C$ -displacement property by using Knothe-Rosenblatt couplings. Let  $(X_i, d_i, m_i, L_i)$ ,  $i \in [n] := \{1, \dots, n\}$ , be  $n$  graphs satisfying the assumptions of the paper (12)-(15). Let  $(X, d, m, L)$  be the product graph space defined by  $X := X_1 \times \dots \times X_n$ ,  $m := m_1 \otimes \dots \otimes m_n$ , and for all  $x = (x_1, \dots, x_n) \in X$ ,  $y = (y_1, \dots, y_n) \in X$ ,

$$d(x, y) := \sum_{i=1}^n d_i(x_i, y_i).$$

If each measure  $m_i$  is reversible with respect to  $L_i$ , then the product measure  $m$  is reversible with respect to the generator

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

Namely  $L$  is defined by  $L(x, y) = 0$  if  $d(x, y) \geq 2$ ,  $L(x, x) = -\sum_{y \in X, y \neq x} L(x, y)$ , and for  $d(x, y) = 1$ , if  $i \in [n]$  is the index for which  $d_i(x_i, y_i) = 1$  (and  $x_j = y_j$  for all  $j \neq i$ ), then

$$L(x, y) = L_i(x_i, y_i).$$

The Markov semi-group  $(P_t)_{t \geq 0}$  associated to  $L$  has a product structure, for any  $x, y \in X$ , for any  $t \geq 0$ ,

$$P_t(x, y) = P_{1,t}(x_1, y_1) \cdots P_{n,t}(x_n, y_n),$$

where  $(P_{i,t})_{t \geq 0}$  denotes the semi-group associated to the generator  $L_i$  on  $X_i$ ,  $i \in \{1, \dots, n\}$ . By construction, it follows that the Schrödinger bridge at zero temperature between the Dirac measures  $\delta_x$  and  $\delta_y$  is a product of Schrödinger bridges at zero temperature between the Dirac measures  $\delta_{x_i}$  and  $\delta_{y_i}$  on  $X_i$ , namely for any  $z = (z_1, \dots, z_n) \in X$

$$(20) \quad Q_t^{x,y}(z) = Q_t^{x_1, y_1}(z_1) \cdots Q_t^{x_n, y_n}(z_n).$$

This can be also derived from the geometric structure of the graph. Since any discrete geodesic from  $x$  to  $y$  is made of  $d_i(x_i, y_i)$  jumps for the  $i$ 's coordinates picked from a discrete geodesic from  $x_i$  to  $y_i$  on  $\mathcal{X}_i$ , one has for  $x \neq y$ ,

$$L^{d(x,y)}(x, y) = \binom{d(x, y)}{d_1(x_1, y_1), \dots, d_n(x_n, y_n)} L_1^{d(x_1, y_1)}(x_1, y_1) \cdots L_n^{d(x_n, y_n)}(x_n, y_n),$$

where for any integers  $d, k_1, \dots, k_n$  such that  $d = k_1 + \dots + k_n$ ,  $\binom{d}{k_1, \dots, k_n} := \frac{d!}{k_1! \cdots k_n!}$  is the multinomial coefficient. The identity (20) then easily follows.

Using the notations of the paper [16], any measures  $\nu_0, \nu_1 \in \mathcal{P}(\mathcal{X})$  admit the following disintegration formulas: for all  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathcal{X}$ ,

$$\begin{aligned} \nu_0(x) &= \nu_0^1(x_1) \nu_0^2(x_2|x_1) \nu_0^3(x_3|x_1, x_2) \cdots \nu_0^n(x_n|x_1, \dots, x_{n-1}), \\ \nu_1(y) &= \nu_1^1(y_1) \nu_1^2(y_2|y_1) \nu_1^3(y_3|y_1, y_2) \cdots \nu_1^n(y_n|y_1, \dots, y_{n-1}), \end{aligned}$$

with  $\nu_0^1, \nu_1^1 \in \mathcal{P}(\mathcal{X}_1)$  and for any  $i \in \{2, \dots, n\}$ ,  $\nu_0^i(\cdot|x_1, \dots, x_{i-1}), \nu_1^i(\cdot|y_1, \dots, y_{i-1}) \in \mathcal{P}(\mathcal{X}_i)$ . For  $i \in [n]$ , let  $\pi^i(\cdot|x_1, \dots, x_{i-1}, y_1, \dots, y_{i-1}) \in \mathcal{P}(\mathcal{X}_i^2)$  be a coupling of  $\nu_0^i(\cdot|x_1, \dots, x_{i-1})$  and  $\nu_1^i(\cdot|y_1, \dots, y_{i-1})$ . Then, the Knothe-Rosenblatt coupling  $\pi^{(n)}$  of  $\nu_0$  and  $\nu_1$  associated to the collection of couplings  $\pi_i$ 's is defined by

$$\pi^{(n)}(x, y) := \pi_1(x_1, y_1) \pi_2(x_2, y_2|x_1, y_1) \cdots \pi_n(x_n, y_n|x_1, \dots, x_{n-1}, y_1, \dots, y_{n-1}).$$

One notices  $(Q_t^{(n)})_{t \in [0,1]}$  the bridge in  $\mathcal{P}(\mathcal{X})$  from  $Q_0^{(n)} = \nu_0$  to  $Q_1^{(n)} = \nu_1$ , associated to the coupling  $\pi^{(n)}$ , defined by

$$Q_t^{(n)}(z) = \iint Q_t^{x,y}(z) d\pi^{(n)}(x, y), \quad t \in [0, 1].$$

**Theorem 2.2.** *Let  $(\mathcal{X}_i, d_i, m_i, L_i)$ ,  $i \in [n]$ , be a collection of graph spaces. Assume that each space  $(\mathcal{X}_i, d_i, m_i, L_i)$  satisfies a  $C_i$ -displacement convexity property with  $C_i = (C_{i,t})_{t \in (0,1)}$ . Let  $(\mathcal{X}, d, m, L)$  be the product space defined as above. Given  $\nu_0, \nu_1 \in \mathcal{P}_b(\mathcal{X})$  with their disintegration formulas mentioned above, let  $\pi^{(n)}$  be the Knothe-Rosenblatt coupling of  $\nu_0$  and  $\nu_1$ , associated the collection of couplings  $\widehat{\pi}_i$ 's constructed as follows:  $\widehat{\pi}_1 := \widehat{Q}_{0,1}^1$  is the projection at time 0 and 1 of  $\widehat{Q}^1$ , the limit Schrödinger problem optimizer at zero temperature between  $\nu_0^1$  and  $\nu_1^1$ , and for  $i \in \{2, \dots, n\}$  and  $x_1, \dots, x_{i-1}, y_1, \dots, y_{i-1} \in \mathcal{X}$ ,*

$$\widehat{\pi}_i(\cdot|x_1, \dots, x_{i-1}, y_1, \dots, y_{i-1}) = \widehat{Q}_{0,1}^i(\cdot|x_1, \dots, x_{i-1}, y_1, \dots, y_{i-1})$$

*is the projection at time 0 and 1 of  $\widehat{Q}^i(\cdot|x_1, \dots, x_{i-1}, y_1, \dots, y_{i-1})$ , the limit Schrödinger problem optimizer at zero temperature between  $\nu_0^i(\cdot|x_1, \dots, x_{i-1})$  and  $\nu_1^i(\cdot|y_1, \dots, y_{i-1})$ . Then, the product space  $(\mathcal{X}, d, m, L)$  satisfies the following convexity property, for any  $\nu_0, \nu_1 \in \mathcal{P}_b(\mathcal{X}^2)$  and any  $t \in (0, 1)$ ,*

$$(21) \quad H(Q_t^{(n)}|m) \leq (1-t)H(\nu_0|m) + tH(\nu_1|m) - \frac{t(1-t)}{2} C_t(\pi^{(n)}),$$

where  $(Q_t^{(n)})_{t \in [0,1]}$  is the bridge from  $\nu_0$  to  $\nu_1$  associated to the coupling  $\pi^{(n)}$  and

$$C_t(\pi^{(n)}) := \sum_{i=1}^n \iint C_{i,t}(\widehat{\pi}_i(\cdot|x_1, \dots, x_{i-1}, y_1, \dots, y_{i-1})) d\pi^{(n)}(x, y).$$

The proof of this result is a simple adjustment of the proof of Theorem 1.1 in [15], which is left to the reader.

**Remarks.** • *Even if the  $\widehat{\pi}_i$ 's are  $W_1$ -optimal couplings in  $\mathcal{P}(\mathcal{X}_i^2)$ , there is no reason for  $\pi^{(n)}$  to be a  $W_1$ -optimal coupling of  $\nu_0$  and  $\nu_1$  in  $\mathcal{P}(\mathcal{X}^2)$ , and therefore for  $(Q_t^{(n)})_{t \in [0,1]}$  to be a  $W_1$ -geodesic. Therefore, the convexity property (21) on the product space  $(\mathcal{X}, d, m, L)$  slightly differs from the convexity property given by Definition 1.1.*

- *One will see on the discrete hypercube  $\mathcal{X} = \{0, 1\}^n$ , that working directly on the product space provides convexity properties that can not be derived from the tensorization property of Theorem 2.2.*

Let us now present results for specific discrete spaces  $(\mathcal{X}, d, m, L)$ . For each of these spaces, we describe the Schrödinger path at zero temperature and, as a main result, we give a  $C$ -displacement convexity property (4) 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

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

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

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

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

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

For any integers  $d, k_1, \dots, k_n$  such that  $d = k_1 + \dots + k_n$ ,  $\binom{d}{k_1, \dots, k_n} = \frac{d!}{k_1! \dots k_n!}$  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  $(\widehat{Q}_t)_{t \in [0,1]}$  joining two measures  $\nu_0, \nu_1 \in \mathcal{P}_b(\mathcal{X})$  is given by (3) with, according to (17),

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

Observe that  $(Q_t^{x,y})_{t \in [0,1]}$  is a binomial interpolation path as in the paper by E. Hillion [19].

**Theorem 2.3.** *On the space  $(\mathbb{Z}^n, m, d, L)$ , the relative entropy  $H(\cdot|m)$  satisfies the 0-displacement convexity property (4). In other words, for any Schrödinger bridge at zero temperature  $(\widehat{Q}_t)_{t \in [0,1]}$  joining any two measures  $\nu_0, \nu_1 \in \mathcal{P}_b(\mathbb{Z}^n)$ , the map  $t \mapsto H(\widehat{Q}_t|m)$  is convex.*

Therefore the space  $(\mathbb{Z}^n, d, m, L)$  has non-negative  $W_1$  or  $T_2$ -entropic curvature. Actually, it can not be positive and one may say that  $(\mathbb{Z}^n, d, m, L)$  is a flat space. Indeed, if property (4) holds with  $C_t(\widehat{\pi}^0) = KW_1^2(\nu_0, \nu_1)$ ,  $K > 0$ , then choosing  $\nu_0 = \delta_x$  and  $\nu_1 = \delta_y$  for  $x, y \in \mathcal{X}$ , one gets for  $t = 1/2$

$$-\log |[x, y]| = -\log |\text{supp}(Q_{1/2}^{x,y})| = H(Q_{1/2}^{x,y}|m) \leq -\frac{K}{8} d^2(x, y),$$

where for a finite set  $A$ ,  $|A|$  denotes its cardinality. Since  $|[x, y]| = \prod_{i=1}^n (|y_i - x_i| + 1)$ , the last inequality implies for any  $x, y \in \mathbb{Z}^n$ ,

$$\left( \sum_{i=1}^n |y_i - x_i| \right)^2 \leq \frac{8}{K} \sum_{i=1}^n \log(|y_i - x_i| + 1),$$

which is impossible for large values of  $|y_i - x_i|$ . A similar proof holds replacing  $W_1^2(\nu_0, \nu_1)$  by  $T_{c_2}(\widehat{\pi})$ .

The convexity property along binomial interpolation paths given by Theorem 2.3 has been first obtained by E. Hillion [19]. 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  $t \mapsto H(\widehat{Q}_t^{\gamma}|m)$  can be easily computed using  $\Gamma_2$  calculus. Then we analyse the behavior of the second derivative of this function as temperature goes to 0, and get a nonnegative lower bound at zero temperature on  $\mathbb{Z}^n$ . This provides the convexity property of  $t \mapsto H(\widehat{Q}_t|m)$ . 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 [17] 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 [22] (see also the more recent paper [18] by Halikias-Klartag-Slomka). Their 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

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

where  $\pi$  is the monotone coupling between  $\nu_0$  and  $\nu_1$  (which is a  $W_1$ -optimizer), and for all  $x, y \in \mathbb{Z}$ ,

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

Gozlan & al. result [17, Theorem 8] states that

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

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

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

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

The Schrödinger bridge at zero temperature  $(\widehat{Q}_t)_{t \in [0,1]}$  given by (3), is the same as the bridge used in [15] for the complete graph (see section 2.1.1): for any  $x, y \in \mathcal{X}$  one has

$$(22) \quad Q_t^{x,y}(z) = (1-t)\delta_x(z) + t\delta_y(z), \quad z \in \mathcal{X},$$

and therefore  $\widehat{Q}_t = (1-t)\nu_0 + t\nu_1$ .

**Theorem 2.4.** *On the finite space  $(\mathcal{X}, \mu, d, L)$ , the relative entropy  $H(\cdot|\mu)$  satisfies the C-displacement convexity property (4), 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  $\widehat{Q} \in \mathcal{P}(\Omega)$ ,*

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

where  $\widehat{\pi} \in \widehat{Q}_{0,1}$ , and for any  $t \in (0, 1)$ ,  $u \geq 0$ ,

$$h_t(u) := \frac{th(u) - h(tu)}{t(1-t)}, \quad \text{with} \quad h(u) = \begin{cases} 2[(1-u)\log(1-u) + u] & \text{for } 0 \leq u \leq 1, \\ +\infty & \text{for } u > 1. \end{cases}$$



The cost  $C_t(\widehat{\pi})$  can be compared with a function of the total variation distance

$$(23) \quad \|v_0 - v_1\|_{TV} := 2 \sup_{A \subset X} |v_0(A) - v_1(A)| = 2 \inf_{\pi \in \Pi(v_0, v_1)} \int \mathbb{1}_{x \neq y} d\pi(x, y) = 2W_1(v_0, v_1).$$

Namely, one has

$$(24) \quad C_t(\widehat{\pi}) \geq (1 + W_1(v_0, v_1)) k_t \left( \frac{W_1(v_0, v_1)}{1 + W_1(v_0, v_1)} \right),$$

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

$$(25) \quad k_t(v) := \inf_{\alpha, \beta, 0 < \alpha + \beta \leq 1} \left\{ \alpha h_t \left( \frac{v}{\alpha} \right) + \beta h_{1-t} \left( \frac{v}{\beta} \right) \right\} \geq \frac{4v^2}{1-v}.$$

**Comments.** (a) This result is an improved version of the convexity properties of the relative entropy obtained by Gozlan & al. [15, Proposition 4.1]. Indeed, from the estimate (24) and the inequality (25) (whose proofs are given at the end of the proof of Theorem 2.4), one gets

$$(26) \quad C_t(\widehat{\pi}) \geq 4W_1(v_0, v_1)^2 = \|v_0 - v_1\|_{TV}^2,$$

and from the inequality  $h_t(u) \geq u^2$ , for all  $u \in [0, 1]$ ,  $t \in (0, 1)$ , it follows that

$$C_t(\widehat{\pi}) \geq \widetilde{T}_2(v_0, v_1),$$

with

$$\widetilde{T}_2(v_0, v_1) := \inf_{\pi \in \Pi(v_0, v_1)} \left[ \int \left( \int \mathbb{1}_{w \neq x} d\pi_{\leftarrow}(w|x) \right)^2 dv_0(x) + \int \left( \int \mathbb{1}_{w \neq y} d\pi_{\leftarrow}(w|y) \right)^2 dv_1(y) \right].$$

These lower bounds on  $C_t(\widehat{\pi})$  exactly provide the convexity properties of Proposition 4.1 [15].

(b) Since  $\mu$  is a probability measure, by Jensen's inequality  $H(\widehat{Q}_t|\mu) \geq 0$ . Therefore, the displacement convexity property (4) together with the bound (26) imply the well-known Csiszar-Kullback-Pinsker inequality by optimizing over all  $t \in (0, 1)$  (see [15, Remark 4.2]), namely

$$\frac{1}{2} \|v_0 - v_1\|_{TV}^2 \leq \left( \sqrt{H(v_0|\mu)} + \sqrt{H(v_1|\mu)} \right)^2, \quad \forall v_0, v_1 \in \mathcal{P}(X).$$

The optimality of the constant 1/2 on the left-hand side of this inequality gives the optimality of the constant 4 in (26). Therefore the  $W_1$ -entropic curvature of the complete graph is 4.

Observe that (4) actually provides an improved version of the Csiszar-Kullback-Pinsker inequality, namely for any  $t \in (0, 1)$ ,

$$\frac{1}{2} (1 + W_1(v_0, v_1)) k_t \left( \frac{W_1(v_0, v_1)}{1 + W_1(v_0, v_1)} \right) \leq \frac{1}{t} H(v_0|\mu) + \frac{1}{1-t} H(v_1|\mu), \quad \forall v_0, v_1 \in \mathcal{P}(X).$$

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

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

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

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

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

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

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

and the graph distance is the *Hamming distance* :

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

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

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

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

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

and  $L_i(z_i, z_i) = -L_i(z_i, \bar{z}_i)$ , the Bernoulli measure  $\mu_i$  is reversible with respect to  $L_i$  and one has

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

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

$$(27) \quad L^{d(x,y)}(x, y) = d(x, y)! \prod_{i=1}^n (1 - \alpha_i)^{|x_i - y_i|_+} \alpha_i^{|y_i - x_i|_+},$$

and it follows that the Schrödinger bridge at zero temperature  $(\widehat{Q}_t^0)_{t \in [0,1]}$  joining two probability measures  $\nu_0$  and  $\nu_1$  is given by (3), with according to (17)

$$(28) \quad Q_t^{x,y}(z) = \mathbb{1}_{[x,y]}(z) t^{d(x,z)} (1-t)^{d(z,y)}, \quad z \in \{0, 1\}^n.$$

This path has exactly the same structure as the one used in [15] to establish entropic curvature bounds on the product space  $(\{0, 1\}^n, \mu)$  (see section 2.1.2).

**Theorem 2.5.** *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 (4), 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  $\widehat{Q} \in \mathcal{P}(\Omega)$ ,*

$$C_t(\widehat{\pi}) := \max \left[ \frac{4}{n} W_1^2(\nu_0, \nu_1), \frac{4}{n} T_{c_2}(\widehat{\pi}), \widetilde{T}_t(\widehat{\pi}) \right],$$

where  $\widehat{\pi} = \widehat{Q}_{0,1}$ , the cost function  $c_2$  of  $T_{c_2}$  is defined by

$$c_2(h) := \max \left\{ \frac{h(h-1)}{2}, h^2 - 2h(1 + \log h) \mathbb{1}_{h \neq 0} \right\}, \quad h \in \mathbb{N},$$

the cost  $\widetilde{T}_t$  is defined by

$$\widetilde{T}_t(\widehat{\pi}) := \int \sum_{i=1}^n h_t(\Pi_{\rightarrow}^i(x)) d\nu_0(x) + \int \sum_{i=1}^n h_{1-t}(\Pi_{\leftarrow}^i(y)) d\nu_1(y),$$

with the definition of the functions  $h_t$ ,  $t \in (0, 1)$  given in Theorem 2.4 and setting

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

**Comments.** (a) *The first lower bound  $C_t(\widehat{\pi}) \geq \frac{4}{n} W_1^2(\nu_0, \nu_1)^2$  gives the  $W_1$ -entropic curvature of the discrete hypercube  $\{0, 1\}^n$  bigger and asymptotically equal to  $4/n$  as  $n$  goes to infinity. Indeed, as in the previous part to recover the Csiszar-Kullback-Pinsker inequality, the well-known  $W_1$ -optimal transport-entropy inequality on the discrete hypercube for product probability measures is a consequence the displacement convexity property (4), using  $H(\widehat{Q}_t|\mu) \geq 0$  and optimizing over all  $t \in (0, 1)$ . Namely, one has*

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

From the central limit Theorem, the constant  $2/n$  (related to the subgaussian constant of the space as mentioned before) is known to be asymptotically optimal as  $n$  goes to infinity.

- (b) The second lower bound  $C_t(\widehat{\pi}) \geq \frac{4}{n} T_{c_2}(\widehat{\pi})$  can not be derived from a tensorisation property such as in Theorem 2.2. Indeed, for  $n = 1$ , on the two points space, one has  $T_{c_2}(\widehat{\pi}) = 0$ . Therefore, the Schrödinger approach allows to capture a property of the hypercube that can not be derived from a tensorisation property as it is often the case.

This second lower bound also gives a new kind of curved Prékopa-Lindler inequality on the discrete hypercube by applying Theorem 2.1. It also implies the following new transport-entropy inequality on the discrete hypercube, for any  $\nu_0, \nu_1 \in \mathcal{P}(\{0, 1\}^n)$ ,

$$(29) \quad \frac{2}{n} T_{c_2}(\nu_0, \nu_1) \leq \left( \sqrt{H(\nu_0|\mu)} + \sqrt{H(\nu_1|\mu)} \right)^2.$$

As opposed to Marton's transport inequality or to  $W_2$ -Talagrand's transport inequality on Euclidean space, inequality (29) 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 (29), applying the central limit theorem, one exactly recovers the well-known  $W_2$ -transport entropy inequality for the standard Gaussian probability measure  $\gamma$  on  $\mathbb{R}$ , due to Talagrand [42]. Namely, one has for any absolutely continuous probability measure  $\nu \in \mathcal{P}_2(\mathbb{R})$ ,

$$(30) \quad W_2^2(\nu, \gamma) \leq 2H(\nu|\gamma).$$

For a sake of completeness, the proof of this implication is given in Appendix A (see Lemma 4.1). As a byproduct of this observation, since the constant 2 is optimal in Talagrand's inequality (30), the constant  $2/n$  in (29) and the constant  $4/n$  in  $C_t(\widehat{\pi}) \geq \frac{4}{n} T_{c_2}(\widehat{\pi})$  are also asymptotically optimal in  $n$ . Therefore the  $T_2$ -entropic curvature of the discrete hypercube is asymptotically equivalent to  $4/n$  as  $n$  goes to infinity.

Actually, according to the proof of Theorem 2.5, for each fixed  $t \in (0, 1)$ , the cost function  $c_2$  can be improved, one has

$$C_t(\widehat{\pi}) \geq \frac{4}{n} \iint w_t(d(x, y)) d\widehat{\pi}(x, y),$$

where for any  $d \in \mathbb{N}$

$$w_t(d) := \max \left\{ \frac{d(d-1)}{2}, \int_0^1 v_s(d) K_t(s) ds \right\} \geq c_2(d),$$

with

$$v_t(d) := \frac{1}{4} \left( \sum_{k=0}^d \sqrt{k(k-1)} \left( \frac{\rho_t^d(k)}{t} + \frac{\rho_{1-t}^d(k)}{1-t} \right) \right)^2.$$

- (c) The inequality  $h_t(u) \geq u^2$ , for all  $u \in [0, 1]$ ,  $t \in (0, 1)$  gives  $\widetilde{T}_t(\widehat{\pi}) \geq \widetilde{T}_2(\nu_0, \nu_1)$  with

$$(31) \quad \widetilde{T}_2(\nu_0, \nu_1) := \inf_{\pi \in \Pi(\nu_0, \nu_1)} \left[ \int \sum_{i=1}^n \left( \int \mathbb{1}_{w_i \neq x_i} d\pi_{\rightarrow}(w|x) \right)^2 d\nu_0(x) + \int \sum_{i=1}^n \left( \int \mathbb{1}_{w_i \neq y_i} d\pi_{\leftarrow}(w|y) \right)^2 d\nu_1(y) \right],$$

So, from the third lower bound  $\widetilde{T}_t(\widehat{\pi})$  of  $C_t(\widehat{\pi})$ , one recovers a similar convexity property as the one obtained for the discrete cube in [15, Corollary 4.4]. The only difference is the expression (3) of the path  $(\widehat{Q}_t)_{t \in [0, 1]}$ , the coupling measure  $\widehat{\pi}$  is replaced by an optimal Knothe-Rosenblatt coupling.

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

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

(d) The lower bound  $\widetilde{T}_t(\widehat{\pi})$  is also well adapted to recover modified logarithmic Sobolev inequality on the discrete hypercube as  $t$  goes to 0. Assume  $\nu_0$  is a probability measure with positive density  $f$ . Observe first that

$$\lim_{t \rightarrow 0} \widetilde{T}_t(\widehat{\pi}) = \int \sum_{i=1}^n h(\Pi_{\rightarrow}^i(x)) d\nu_0(x) + \int \sum_{i=1}^n h_1(\Pi_{\leftarrow}^i(y)) d\nu_1(y),$$

where for  $u \in [0, 1)$ ,  $h_1(u) := \lim_{t \rightarrow 1} h_t(u) = uh'(u) - h(u) = 2(-u - \log(1 - u))$ . For any real function  $g$  on  $\{0, 1\}^n$ , let us note

$$D_i g(x) := g(\sigma_i(x)) - g(x), \quad x \in \{0, 1\}^n.$$

Applying Lemma 4.2, since  $\Pi_{\rightarrow}^i(x) = \Pi_{\rightarrow}^{\sigma_i(x)}(x)$ , the convexity property (4) with  $C_t = \widetilde{T}_t$  given by Theorem 2.5 implies as  $t$  goes 0

$$\begin{aligned} H(\nu_0|\mu) &\leq H(\nu_1|\mu) + \sum_{x \in \mathcal{X}} \sum_{i=1}^n -D_i(\log f)(x) \Pi_{\rightarrow}^i(x) \nu_0(x) \\ &\quad - \frac{1}{2} \int \sum_{i=1}^n h(\Pi_{\rightarrow}^i(x)) d\nu_0(x) - \frac{1}{2} \int \sum_{i=1}^n h_1(\Pi_{\leftarrow}^i(y)) d\nu_1(y). \end{aligned}$$

Choosing then  $\nu_1 = \mu$  it follows that

$$(32) \quad \begin{aligned} H(\nu_0|\mu) &\leq \sum_{x \in \mathcal{X}} \sum_{i=1}^n -D_i(\log f)(x) \Pi_{\rightarrow}^i(x) \nu_0(x) \\ &\quad - \frac{1}{2} \int \sum_{i=1}^n h(\Pi_{\rightarrow}^i(x)) d\nu_0(x) - \frac{1}{2} \int \sum_{i=1}^n h_1(\Pi_{\leftarrow}^i(y)) d\mu(y). \end{aligned}$$

One may check that this inequality is optimal since for the two points space ( $n = 1$ ) this is an equality. The proof of this equality is left to the reader. It lies on the fact that since  $\widehat{\pi}$  is a  $W_1$  optimizer, one has  $\pi(x, x) = \min(\nu_0(x), \mu(x))$  for  $x = 0$  and  $x = 1$ . From this remark, starting from the tensorisation form of the one dimensional convexity property with  $C_t = \widetilde{T}_t$  given by Theorem 2.2 with the  $\pi^{(n)}$  be the Knothe-Rosenblatt coupling of  $\nu_0 \in \mathbb{P}(\{0, 1\}^n)$  and  $\nu_1 = \mu$ , one easily check that the same strategy as  $t$  goes to 0 implies

$$H(\nu_0|\mu) \leq H(\nu_0^1|\mu_1) + \int \sum_{i=2}^n H(\nu_0^i(\cdot|x_1, \dots, x_{i-1})|\mu_i) \nu_0(x),$$

which is still an equality due to the tensorisation property of entropy. However, without using the tensorisation argument, we don't know if (32) is an equality for dimension  $n$  bigger than 2.

From (32) in dimension  $n$ , using the identity

$$(33) \quad \sup_{p \in [0, 1)} \left\{ -Dp - \frac{1}{2} h(p) \right\} = \frac{1}{2} h^*(2D_-) = e^{-D_-} + D_- - 1,$$

and since

$$\begin{aligned} &\frac{1}{2} h^*(2[D_i(\log f)(x)]_-) f(x) \\ &= ([D_i f(x)]_- + f(\sigma_i(x)))[D_i(\log f)(x)]_- - [D_i f(x)]_- \\ &\leq [D_i(\log f)(x)]_- [D_i f(x)]_-, \end{aligned}$$

one gets the following modified logarithmic Sobolev inequality,

$$\begin{aligned}
(34) \quad H(f\mu|\mu) &\leq \int \sum_{i=1}^n \frac{1}{2} h^* (2[D_i(\log f)]_-) \nu_0(x) - \frac{1}{2} \sum_{i=1}^n \int h_1(\Pi_{\leftarrow}^i(y)) d\mu(y) \\
&\leq \sum_{i=1}^n \int \frac{1}{2} h^* (2[D_i(\log f)]_-) d\nu_0 \\
&\leq \int \sum_{i=1}^n [D_i(\log f)]_- [D_i f]_- d\mu.
\end{aligned}$$

Since  $h^*(2[D_i(\log f)]_-) \leq [D_i(\log f)]_-^2$ , one recovers the reinforced modified logarithmic Sobolev inequality of Corollary 5.5 in [15]. By means of the Central Limit Theorem, this reinforced modified log-Sobolev inequality actually leads to the usual logarithmic Sobolev inequality of Gross [18] for the standard Gaussian, with the optimal constant (see [15, Corollary 5.5]).

A simple way to improve the modified inequality (34) is to take into account the extra term involving  $h_1(\Pi_{\leftarrow}^i(y))$  in (32). Given  $x_i \in \{0, 1\}$  and for  $j \in [n] \setminus \{i\}$  given  $z_j \in \{0, 1\}$ , let us introduce the notations

$$z_{\bar{i}} x_i := (z_1, \dots, z_{i-1}, x_i, z_{i+1}, \dots, z_n) \in \{0, 1\}^n, \quad \text{and} \quad z_{\bar{i}} := (z_1, \dots, z_{i-1}, z_{i+1}, \dots, z_n) \in \{0, 1\}^{n-1}.$$

Applying Jensen's inequality, the convexity property of the function  $h_1$  provides

$$\int h_1(\Pi_{\leftarrow}^i(y)) d\mu(y) \geq \int h_1\left(1 - \frac{\sum_{z_{\bar{i}} \in \{0,1\}^{n-1}} \sum_{w_{\bar{i}} \in \{0,1\}^{n-1}} \widehat{\pi}(z_{\bar{i}} y_i, w_{\bar{i}} y_i)}{\mu_i(y_i)}\right) d\mu_i(y_i),$$

By setting  $\mu_{\bar{i}} = \otimes_{j \in [n] \setminus \{i\}} \mu_j$  and since

$$\sum_{z_{\bar{i}} \in \{0,1\}^{n-1}} \sum_{w_{\bar{i}} \in \{0,1\}^{n-1}} \widehat{\pi}(z_{\bar{i}} y_i, w_{\bar{i}} y_i) \leq \sum_{w \in \{0,1\}^n} \widehat{\pi}(z_{\bar{i}} y_i, w) = \sum_{z_{\bar{i}} \in \{0,1\}^{n-1}} f(z_{\bar{i}} y_i) \mu_{\bar{i}}(z_{\bar{i}}) \mu_i(y_i),$$

and

$$\sum_{z_{\bar{i}} \in \{0,1\}^{n-1}} \sum_{w_{\bar{i}} \in \{0,1\}^{n-1}} \widehat{\pi}(z_{\bar{i}} y_i, w_{\bar{i}} y_i) \leq \sum_{z \in \{0,1\}^n} \sum_{w_{\bar{i}} \in \{0,1\}^{n-1}} \widehat{\pi}(z, w_{\bar{i}} y_i) = \mu_i(y_i),$$

it follows that

$$\int h_1(\Pi_{\leftarrow}^i(y)) d\mu(y) \geq \int h_1\left(1 - \min\left\{1, \sum_{z_{\bar{i}} \in \{0,1\}^{n-1}} f(z_{\bar{i}} y_i) \mu_{\bar{i}}(z_{\bar{i}})\right\}\right) d\mu_i(y_i).$$

For any fixed  $y_i \in \{0, 1\}$ , one has

$$1 - \min\left\{1, \sum_{z_{\bar{i}} \in \{0,1\}^{n-1}} f(z_{\bar{i}} y_i) \mu_{\bar{i}}(z_{\bar{i}})\right\} = \left[\int f(z) - f(z_{\bar{i}} y_i) d\mu(z)\right]_+ = \left[\int D_i f(y) d\mu_{\bar{i}}(y_{\bar{i}})\right]_+ \mu_i(\bar{y}_i).$$

As a consequence (34) provides the following new modified logarithmic Sobolev inequality on the discrete hypercube,

$$H(f\mu|\mu) \leq \sum_{i=1}^n \int \frac{1}{2} h^* (2[D_i(\log f)]_-) d\nu_0 - \sum_{i=1}^n \frac{1}{2} \int h_1\left(\left[\int D_i f(y) d\mu_{\bar{i}}(y_{\bar{i}})\right]_+ \mu_i(\bar{y}_i)\right) d\mu_i(y_i).$$

As we will show in a forthcoming paper, the last strategy also simply provides modified logarithmic Sobolev inequalities for probability measures  $\mu_V = e^{-V} \mu$  on  $\{0, 1\}^n$  with interaction potentials  $V : \{0, 1\}^n \rightarrow \mathbb{R}$ , that can not be easily derived from tensorization property arguments.

**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  $\mu$ ,  $\mu(x) = 1/N$ . The measure  $\mu$  is reversible with respect to the generator  $L$  given by ,

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

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  $(\widehat{Q}_t)_{t \in [0,1]}$  joining two probability measures  $\nu_0$  and  $\nu_1$  on  $\mathbb{Z}/N\mathbb{Z}$  is given by (3), with according to (17)

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

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  $(\widehat{Q}_t)_{t \in [0,1]}$  is given by (3), with according to (17) : if  $d(x, y) < N/2$  then

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

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

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

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

**Theorem 2.6.** *On the space  $(\mathbb{Z}/N\mathbb{Z}, \mu, d, L)$ , the relative entropy  $H(\cdot|\mu)$  satisfies the 0-displacement convexity (4).*

Therefore the space  $(\mathbb{Z}/N\mathbb{Z}, d, \mu, L)$  has non-negative  $W_1$  or  $T_2$  entropic curvature.

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

$$\mathcal{X}_\kappa := \left\{ x = (x_1, \dots, x_n) \in \{0, 1\}^n \mid x_1 + \dots + x_n = \kappa \right\}.$$

For  $z \in \mathcal{X}_\kappa$ , let  $J_0(z) := \{i \in [n] \mid z_i = 0\}$  and  $J_1(z) := \{i \in [n] \mid z_i = 1\}$ . For any  $i \in J_0(z)$  and  $j \in J_1(z)$ , one denotes  $\sigma_{ij}(z)$  the neighbour of  $z$  in  $\mathcal{X}_\kappa$  defined by

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

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

$$E := \left\{ (z, \sigma_{ij}(z)) \mid z \in \mathcal{X}_\kappa, \{i, j\} \subset [n], z_i = 0, z_j = 1 \right\},$$

and the graph distance is given by

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

The measure  $\mu$  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  $L(z, z) := -\kappa(n - \kappa)$ .

Since  $L^{d(x,y)}(x, y) = (d(x, y)!)^2$ , the Schrödinger bridge at zero temperature  $(\widehat{Q}_t)_{t \in [0,1]}$  is given by (3), with according to (17),

$$(35) \quad Q_t^{x,y}(z) = \mathbb{1}_{[x,y]}(z) \left( \frac{d(x, y)}{d(x, z)} \right)^{-1} t^{d(x,z)} (1-t)^{d(z,y)}, \quad z \in \mathcal{X}_\kappa.$$

**Theorem 2.7.** *On the space  $(X_\kappa, \mu, d, L)$ , the relative entropy  $H(\cdot|\mu)$  satisfies the  $C$ -displacement convexity property (4), with  $C = (C_t)_{t \in (0,1)}$  defined by: for any  $\nu_0, \nu_1 \in \mathcal{P}(X_\kappa)$  with associated limit Schrödinger problem optimizer  $\widehat{Q} \in \mathcal{P}(\Omega)$ ,*

$$C_t(\widehat{\pi}) := \max \left[ \frac{4}{\min[\kappa, n - \kappa]} W_1^2(\nu_0, \nu_1), \frac{4}{\min[\kappa, n - \kappa]} T_{c_2}(\widehat{\pi}), \widetilde{T}_t(\widehat{\pi}) \right],$$

where  $\widehat{\pi} = \widehat{Q}_{0,1}$ , the cost function  $c_2$  of  $T_{c_2}$  is the same as in Theorem 2.5, and the cost  $\widetilde{T}_t$  is defined by

$$\begin{aligned} \widetilde{T}_t(\widehat{\pi}) := & \int \max \left[ \sum_{i \in J_0(x)} h_t(\Pi_{\rightarrow}^i(x)), \sum_{j \in J_1(x)} h_t(\Pi_{\rightarrow}^j(x)) \right] d\nu_0(x) \\ & + \int \max \left[ \sum_{i \in J_0(y)} h_{1-t}(\Pi_{\leftarrow}^i(y)), \sum_{j \in J_1(y)} h_{1-t}(\Pi_{\leftarrow}^j(y)) \right] d\nu_1(y), \end{aligned}$$

with the same definitions for the functions  $h_t$ ,  $t \in (0, 1)$  and the quantities  $\Pi_{\rightarrow}^i(x)$  and  $\Pi_{\leftarrow}^i(y)$  as in Theorem 2.5.

**Comments.** (a) Since  $C_t(\widehat{\pi}) \geq \frac{4}{\min(\kappa, n-\kappa)} W_1^2(\nu_0, \nu_1)$ , the  $W_1$ -entropic curvature of the space  $(X_\kappa, d, L)$  is bounded from below by  $\frac{4}{\min(\kappa, n-\kappa)}$ . Observe that this constant is optimal for  $\kappa = 1$  or  $\kappa = n - 1$ , since  $X_\kappa$  is the complete graph and one recovers its optimal lower curvature bound 4 (see Comment (b) of Theorem 2.4).

In the paper [13, Theorem 1.1] the Erbar-Maas entropic curvature of the Bernoulli Laplace model along  $\mathcal{W}_2$ -geodesics is bounded from below by  $\frac{n+2}{2\kappa(n-\kappa)}$ , therefore their curvature term is of order

$$\frac{n+2}{2\kappa(n-\kappa)} \mathcal{W}_2^2(\nu_0, \nu_1) \geq \frac{n+2}{\kappa(n-\kappa)} W_1^2(\nu_0, \nu_1).$$

Theorem 2.7 a slightly better constant as regards to the  $W_1$ -curvature term since  $\frac{n+2}{\kappa(n-\kappa)} \leq \frac{4}{\min[\kappa, n-\kappa]}$  with equality for  $(\kappa, n) = (1, 2)$ .

- (b) Since  $C_t(\widehat{\pi}) \geq \frac{4}{\min(\kappa, n-\kappa)} T_{c_2}(\widehat{\pi})$ , the  $T_2$ -entropic curvature of the space  $(X_\kappa, d, L)$  is bounded from below by  $\frac{4}{\min(\kappa, n-\kappa)}$ . Moreover, applying Theorem 2.1, this lower bound provides a new type of curved Prékopa-Leindler inequality on the slices of the discrete hypercube.
- (c) According to the definition (31) of  $\widetilde{T}_2(\nu_0, \nu_1)$ , as in the case of the hypercube, one has  $C_t(\widehat{\pi}) \geq \widetilde{T}_t(\widehat{\pi}) \geq \frac{1}{2} \widetilde{T}_2(\nu_0, \nu_1)$ . As a consequence, since  $H(\widehat{Q}_t|\mu) \geq 0$ , optimizing over all  $t \in (0, 1)$ , Theorem 2.7 implies the following weak transport-entropy inequality, for any  $\nu_0, \nu_1 \in \mathcal{P}(X_\kappa)$ ,

$$\frac{1}{4} \widetilde{T}_2(\nu_0, \nu_1) \leq \left( \sqrt{H(\nu_0|\mu)} + \sqrt{H(\nu_1|\mu)} \right)^2.$$

This inequality is a reinforced symmetric version of a transport entropy inequality given in [38, Theorem 1.8 (b)] with the worse constant 1/8 instead of 1/4. It was surprisingly obtained by projection of a transport-entropy inequality for the uniform measure on the symmetric group. The approach of the present paper is much more natural to reach such a result.

- (d) From the lower bound  $C_t(\widehat{\pi}) \geq \widetilde{T}_t(\widehat{\pi})$ , Theorem 2.7 also yields a modified logarithmic Sobolev. For any real function  $g$  on  $X_\kappa$ , let us note

$$D_{ij}g(x) := g(\sigma_{ij}(x)) - g(x), \quad x \in X_\kappa, \quad (i, j) \in J_0(x) \times J_1(x).$$

Assume  $\nu_0$  has positive density  $f$  with respect to  $\mu$  and let us choose  $\nu_1 = \mu$ . According to Lemma 4.2, setting  $\Pi_{\rightarrow}^{ij}(x) = \Pi_{\rightarrow}^{\sigma_{ij}(x)}(x)$ , the convexity property (4) with  $C_t = \tilde{T}_t$  given by Theorem 2.7 implies as  $t$  goes to 0

$$H(\nu_0|\mu) \leq \sum_{x \in X} \sum_{(i,j) \in J_0(x) \times J_1(x)} -D_{ij}(\log f)(x) \Pi_{\rightarrow}^{ij}(x) \nu_0(x) - \frac{1}{2} \int \max \left[ \sum_{i \in J_0(x)} h(\Pi_{\rightarrow}^i(x)), \sum_{j \in J_1(x)} h(\Pi_{\rightarrow}^j(x)) \right] d\nu_0(x).$$

Now, let us observe that for  $i \in J_0(x)$ , one has

$$\begin{aligned} \sum_{j \in J_1(x)} \Pi_{\rightarrow}^{ij}(x) &= \int \sum_{j \in J_1(x)} \mathbb{1}_{\sigma_{ij}(x) \in [x,y]} d(x,y) r(x, \sigma_{ij}(x), \sigma_{ij}(x), y) d\widehat{\pi}_{\rightarrow}(y|x) \\ &= \int \sum_{j \in J_1(x) \cap J_1(y)} \mathbb{1}_{x_i \neq y_i} d(x,y) r(x, \sigma_{ij}(x), \sigma_{ij}(x), y) d\widehat{\pi}_{\rightarrow}(y|x) \\ &= \int \mathbb{1}_{x_i \neq y_i} \frac{d^2(x,y)((d(x,y)-1)!)^2}{(d(x,y)!)^2} d\widehat{\pi}_{\rightarrow}(y|x) = \Pi_{\rightarrow}^i(x), \end{aligned}$$

and similarly for  $j \in J_1(x)$ , one has  $\sum_{i \in J_0(x)} \Pi_{\rightarrow}^{ij}(x) = \Pi_{\rightarrow}^j(x)$ . It follows that

$$H(f\mu|\mu) \leq \int \min \left[ \sum_{i \in J_0(x)} \left( \max_{j \in J_1(x)} [D_{ij}(\log f)(x)]_{-} \Pi_{\rightarrow}^i(x) - \frac{1}{2} h(\Pi_{\rightarrow}^i(x)) \right), \sum_{j \in J_1(x)} \left( \max_{i \in J_0(x)} [D_{ij}(\log f)(x)]_{-} \Pi_{\rightarrow}^j(x) - \frac{1}{2} h(\Pi_{\rightarrow}^j(x)) \right) \right] d\nu_0(x)$$

Finally the identity (33) gives the following modified logarithmic inequality

$$\begin{aligned} H(f\mu|\mu) &\leq \int \min \left[ \sum_{i \in J_0} \frac{1}{2} h^* \left( 2 \max_{j \in J_1} [D_{ij}(\log f)]_{-} \right), \sum_{j \in J_1} \frac{1}{2} h^* \left( 2 \max_{i \in J_0} [D_{ij}(\log f)]_{-} \right) \right] d\nu_0 \\ (36) \quad &\leq \int \min \left[ \sum_{i \in J_0} \max_{j \in J_1} ([D_{ij}(\log f)]_{-} - [D_{ij}f]_{-}), \sum_{j \in J_1} \max_{i \in J_0} ([D_{ij}(\log f)]_{-} - [D_{ij}f]_{-}) \right] d\mu \end{aligned}$$

From the lower bound  $\frac{n+2}{2\kappa(n-\kappa)}$  of Erbar entropic curvature given in [13, Theorem 1.1], we know from [11, Theorem 7.4] that the following modified logarithmic Sobolev inequality holds

$$(37) \quad H(f\mu|\mu) \leq c_n \int \sum_{(i,j) \in J_0 \times J_1} D_{ij}(\log f) D_{ij}f \, d\mu = 2c_n \int \sum_{(i,j) \in J_0 \times J_1} [D_{ij}(\log f)]_{-} - [D_{ij}f]_{-} \, d\mu,$$

with  $c_n = 1/2(n+2)$ , and the best constant  $c_n$  in this inequality is known to be greater than  $1/4n$  (see comments after [13, Theorem 1.1]). This inequality is stronger than (36). Indeed, one has

$$\begin{aligned} &\frac{1}{(n+2)} \int \sum_{(i,j) \in J_0 \times J_1} [D_{ij}(\log f)]_{-} - [D_{ij}f]_{-} \, d\mu \\ &\leq \frac{1}{(n+2)} \int \min \left[ \kappa \sum_{i \in J_0} \max_{j \in J_1} ([D_{ij}(\log f)]_{-} - [D_{ij}f]_{-}), (n-\kappa) \sum_{j \in J_1} \max_{i \in J_0} ([D_{ij}(\log f)]_{-} - [D_{ij}f]_{-}) \right] d\mu \\ &\leq \int \min \left[ \sum_{i \in J_0} \max_{j \in J_1} ([D_{ij}(\log f)]_{-} - [D_{ij}f]_{-}), \sum_{j \in J_1} \max_{i \in J_0} ([D_{ij}(\log f)]_{-} - [D_{ij}f]_{-}) \right] d\mu. \end{aligned}$$

Choosing the function  $f$  defined by  $f(x) := \alpha(x_1 + \beta)$ ,  $x \in X_\kappa$ , where  $\beta > 0$  and  $\alpha$  is a renormalisation constant, one may check that the right-hand side and the left-hand side of this



inequality are asymptotically equivalent as  $n$  goes to infinity. However it remains a challenge to improve our strategy in order to recover (37).

### 3. PROOF OF THE MAIN RESULTS

This section is divided into two parts. We first present general statements to prove displacement convexity property (4) 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.** In order to prove property (4), we fix two probability measures  $\nu_0$  and  $\nu_1$  in  $\mathcal{P}_b(\mathcal{X})$  in this part. As in the paper by G. Conforti [9] in continuous setting, the first step is to decompose the relative-entropy using the product structure given by (8): for any  $t \in [0, 1]$ ,

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

where

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

As recalled below, it is known that the function  $\varphi_\gamma$  is non-increasing and the function  $\psi_\gamma$  is non-decreasing (see [27, Theorem 6.4.2]).

Then, the strategy is to analyse the behaviour of the second order derivative  $\varphi_\gamma''$  and  $\psi_\gamma''$  as  $\gamma$  goes to 0, in order to apply the next Lemma. For any  $t \in (0, 1)$  let  $K_t : [0, 1] \rightarrow \mathbb{R}_+$ , be defined by

$$(38) \quad K_t(u) = \frac{2u}{t} \mathbb{1}_{u \leq t} + \frac{2(1-u)}{1-t} \mathbb{1}_{u \geq t}, \quad u \in [0, 1].$$

$K_t$  is a kernel function since  $\int_0^1 K_t(u) du = 1$ .

**Lemma 3.1.** *Assume that hypothesis (12), (13), (14) and (15) hold. Let  $(\gamma_\ell)_{\ell \in \mathbb{N}}$  be a sequence of positive numbers that converges to 0. If for any  $t \in (0, 1)$*

$$(39) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi_{\gamma_\ell}''(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi_{\gamma_\ell}''(t) \geq \xi''(t),$$

where  $\xi$  is a continuous functions on  $[0, 1]$ , twice differentiable on  $(0, 1)$ , depending on the coupling  $\widehat{\pi}$ , then the displacement convexity property (4) holds with

$$C_t(\widehat{\pi}) := \int_0^1 \xi''(u) K_t(u) du = \frac{2}{t(1-t)} [(1-t)\xi(0) + t\xi(1) - \xi(t)].$$

Observe that if  $\xi'' = K$  is a constant function, then  $C_t(\widehat{\pi}) = K$ . 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 (39). 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 [27, 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

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

and

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

where  $\zeta(s) := s \log s - s + 1$ ,  $s > 0$ , and  $G_t^\gamma$  and  $F_t^\gamma$  are the so-called Schrödinger potentials according to Léonard's paper terminology [27],

$$G_t^\gamma := \log P_{1-t}^\gamma g^\gamma, \quad \text{and} \quad F_t^\gamma := \log P_t^\gamma f^\gamma.$$

Since  $\zeta \geq 0$ , the function  $\varphi_\gamma$  is non-increasing and the function  $\psi_\gamma$  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

$$\begin{aligned} \varphi''_\gamma(t) &= \int \left[ \left( \sum_{z', z'' \sim z} e^{\nabla F_t^\gamma(z, z')} L^\gamma(z, z') \right)^2 + \sum_{z', z'' \sim z} \left( 1 + \nabla F_t^\gamma(z, z') \right) e^{\nabla F_t^\gamma(z, z')} \left( L^\gamma(z, z) - L^\gamma(z', z') \right) L^\gamma(z, z') \right. \\ &\quad \left. + \sum_{z', z'', z \sim z' \sim z''} \rho \left( e^{\nabla F_t^\gamma(z, z')}, e^{\nabla F_t^\gamma(z, z'')} \right) L^\gamma(z, z') L^\gamma(z', z'') \right] d\widehat{Q}_t^\gamma(z), \\ \psi''_\gamma(t) &= \int \left[ \left( \sum_{z', z'' \sim z} e^{\nabla G_t^\gamma(z, z')} L^\gamma(z, z') \right)^2 + \sum_{z', z'' \sim z} \left( 1 + \nabla G_t^\gamma(z, z') \right) e^{\nabla G_t^\gamma(z, z')} \left( L^\gamma(z, z) - L^\gamma(z', z') \right) L^\gamma(z, z') \right. \\ &\quad \left. + \sum_{z', z'', z \sim z' \sim z''} \rho \left( e^{\nabla G_t^\gamma(z, z')}, e^{\nabla G_t^\gamma(z, z'')} \right) L^\gamma(z, z') L^\gamma(z', z'') \right] d\widehat{Q}_t^\gamma(z). \end{aligned}$$

Let us now analyse the behavior of  $\varphi''_\gamma(t)$ ,  $\psi''_\gamma(t)$  as temperature  $\gamma$  goes to zero. Recall first that for  $t \in (0, 1)$ , the support of the Schrödinger bridge at zero temperature  $\widehat{Q}_t$  given by (19) is independent of  $t$ . For sake of simplicity, one denotes

$$\widehat{Z} := \text{supp}(\widehat{Q}_t), \quad t \in (0, 1).$$

As a consequence, one expects that the limit behavior of  $\varphi''_\gamma(t)$ ,  $\psi''_\gamma(t)$  is expressed in term of sums restricted to points of  $\widehat{Z}$ . Let us define, for any  $z \in \widehat{Z}$ ,

$$V_\rightarrow(z) := \{z' \in V(z) \mid (z, z') \in C_\rightarrow\} \quad \text{and} \quad V_\leftarrow(z) := \{z' \in V(z) \mid (z, z') \in C_\leftarrow\},$$

where

$$C_\rightarrow := \{(z, w) \in X \times X \mid z \neq w, \exists (x, y) \in \text{supp}(\widehat{\pi}), (z, w) \in [x, y]\},$$

and

$$C_\leftarrow := \{(z, w) \in X \times X \mid (w, z) \in C_\rightarrow\}.$$

Similarly, one also defines

$$\mathbb{V}_\rightarrow(z) := \{z'', \in \mathbb{V}(z) \mid (z, z'') \in C_\rightarrow\} \quad \text{and} \quad \mathbb{V}_\leftarrow(z) := \{z'', \in \mathbb{V}(z) \mid (z, z'') \in C_\leftarrow\},$$

where for any  $z \in X$

$$\mathbb{V}(z) := \{z'' \in X \mid d(z, z'') = 2\}.$$

As a remarkable fact, according to Lemma 4.3 postponed in Appendix A, from the  $d$ -cyclically monotone property of the  $W_1$ -optimal coupling  $\widehat{\pi}$ ,  $C_\rightarrow$  and  $C_\leftarrow$  are disjoint sets. This implies that  $V_\rightarrow(z)$  and  $V_\leftarrow(z)$  are disjoint, and also  $\mathbb{V}_\rightarrow(z)$  and  $\mathbb{V}_\leftarrow(z)$ , for any  $z \in \widehat{Z}$ .

According to the expression of  $\varphi''_\gamma(t)$ ,  $\psi''_\gamma(t)$  given in Lemma 3.3, a first step is to give the behavior as  $\gamma$  goes to zero of the quantities

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

for  $u = z'$  or  $u = z''$  with  $z \sim z' \sim z''$ . This is a key result of this paper. Let us briefly give the intuition behind it. From the Markov property, the quantity  $A_t^\gamma(z, u)$  can be interpreted as the mean ratio of transition probabilities under conditional law of the Schrödinger bridge, namely

$$(40) \quad A_t^\gamma(z, u) = \sum_{w \in \mathcal{X}} \frac{P_t^\gamma(u, w)}{P_t^\gamma(z, w)} \frac{f^\gamma(w) P_t^\gamma(z, w)}{P_t^\gamma f^\gamma(z)} = \sum_{w \in \mathcal{X}} \frac{P_t^\gamma(u, w)}{P_t^\gamma(z, w)} \widehat{Q}^\gamma(X_0 = w | X_t = z),$$

where  $\widehat{Q}^\gamma(X_0 = w | X_t = z)$  is the law of  $X_0$  given  $X_t = z$  under the law  $\widehat{Q}^\gamma$ . As  $\gamma$  goes to 0, the law  $\widehat{Q}^\gamma$  tends to  $\widehat{Q}$ , and the behavior of the ratio is given by the Taylor expansion of  $P_t^\gamma$  as  $\gamma$  goes to 0, namely according to Lemma 4.4 (iii),

$$\frac{P_t^\gamma(u, w)}{P_t^\gamma(z, w)} = (t\gamma)^{d(u, w) - d(z, w)} \left( \frac{L^{d(u, w)}(u, w) d(z, w)!}{L^{d(z, w)}(z, w) d(u, w)!} + o(1) \right).$$

Therefore, if  $\gamma$  goes to 0 then the main contribution in the sum given by (40) is for points  $w \in \mathcal{X}$  such that  $d(u, w) - d(z, w)$  has minimum value. This means that  $u \in [z, w]$ , so that  $d(u, w) - d(z, w) = -d(u, z)$ . It follows that for  $u = z'$  with  $z' \sim z$ ,

$$A_t^\gamma(z, z') \underset{\gamma \rightarrow 0}{\sim} \frac{1}{\gamma t} \sum_{w \in \mathcal{X}, z' \in [z, w]} \frac{L^{d(z', w)}(z', w) d(z, w)!}{L^{d(z, w)}(z, w) d(z', w)!} \widehat{Q}(X_0 = w | X_t = z),$$

and for  $u = z''$  with  $d(z, z'') = 2$ ,

$$A_t^\gamma(z, z'') \underset{\gamma \rightarrow 0}{\sim} \frac{1}{\gamma^2 t^2} \sum_{w \in \mathcal{X}, z'' \in [z, w]} \frac{L^{d(z'', w)}(z'', w) d(z, w)!}{L^{d(z, w)}(z, w) d(z'', w)!} \widehat{Q}(X_0 = w | X_t = z).$$

The quantity  $B_t^\gamma(z, u)$  can be similarly analysed as  $\gamma$  goes to 0.

Let us now formulate precise statements. One needs to define several quantities. For any  $z \in \mathcal{X}$ ,  $x \in \text{supp}(\nu_0)$ ,  $y \in \text{supp}(\nu_1)$  and any  $t \in (0, 1)$ , let

$$(41) \quad a_t(z, y) := \widehat{Q}(X_t = z | X_1 = y) = \int Q_t^{w, y}(z) d\widehat{\pi}_-(w | y),$$

and

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

Observe that for  $t \in (0, 1)$ ,  $a_t(z, y) > 0$  if and only if  $z \in \widehat{Z}$  and  $y \in \widehat{Y}_z$  with

$$\widehat{Y}_z := \{y \in \text{supp}(\nu_1) \mid \exists x \in \mathcal{X}, (x, y) \in \widehat{\pi}, z \in [x, y]\}.$$

Identically  $b_t(z, x) > 0$  if and only if  $z \in \widehat{Z}$  and  $x \in \widehat{X}_z$  with

$$\widehat{X}_z := \{x \in \text{supp}(\nu_0) \mid \exists x \in \mathcal{X}, (x, y) \in \widehat{\pi}, z \in [x, y]\}.$$

For further use, for any  $y \in \text{supp}(\nu_1)$  and  $x \in \text{supp}(\nu_0)$ , we also introduce the sets

$$\widehat{Z}^y := \{z \in \widehat{Z} \mid y \in \widehat{Y}_z\} \quad \text{and} \quad \widehat{Z}_x := \{z \in \widehat{Z} \mid x \in \widehat{X}_z\},$$

so that

$$(z \in \widehat{Z}, y \in \widehat{Y}_z) \Leftrightarrow (y \in \text{supp}(\nu_1), z \in \widehat{Z}^y),$$

and

$$(z \in \widehat{Z}, x \in \widehat{X}_z) \Leftrightarrow (x \in \text{supp}(\nu_0), z \in \widehat{Z}_x).$$

For any  $z \in \widehat{Z}$ ,  $z' \in V(z)$ , define

$$(42) \quad a_t(z, z', y) := \sum_{w \in \mathcal{X}, (z, z') \in [y, w]} r(y, z, z', w) d(y, w) \rho_t^{d(y, w)-1} (d(z, w) - 1) \widehat{\pi}_-(w|y),$$

and

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

where the function  $r$  is given by (18). One easily check that  $a_t(z, z', y) > 0$  if and only if  $z' \in V_-(z)$  and  $y \in \widehat{Y}_{(z, z')}$  with

$$\widehat{Y}_{(z, z')} = \left\{ y \in \text{supp}(v_1) \mid \exists x \in \mathcal{X}, (x, y) \in \widehat{\pi}, (z, z') \in [y, x] \right\} \subset \widehat{Y}_z \cap \widehat{Y}_{z'},$$

and identically  $b_t(z, z', x) > 0$  if and only if  $z' \in V_-(z)$  and  $x \in \widehat{X}_{(z, z')}$  with

$$\widehat{X}_{(z, z')} = \left\{ x \in \text{supp}(v_0) \mid \exists y \in \mathcal{X}, (x, y) \in \widehat{\pi}, (z, z') \in [x, y] \right\} \subset \widehat{X}_z \cap \widehat{X}_{z'}.$$

For any  $z \in \widehat{Z}$  and  $z'' \in \mathbb{V}(z)$ , define also

$$(43) \quad a_t(z, z'', y) := \sum_{w \in \mathcal{X}, (z, z'') \in [y, w]} r(y, z, z'', w) d(y, w) (d(y, w) - 1) \rho_t^{d(y, w)-2} (d(z, w) - 2) \widehat{\pi}_-(w|y),$$

and

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

We also have  $a_t(z, z'', y) > 0$  if and only if  $z'' \in \mathbb{V}_-(z)$  and  $y \in \widehat{Y}_{(z, z'')}$ , and  $b_t(z, z'', x) > 0$  if and only if  $z'' \in \mathbb{V}_-(z)$  and  $x \in \widehat{X}_{(z, z'')}$ .

**Lemma 3.4.** *Assume that conditions (13) and (14) are fulfilled. Let  $(\gamma_\ell)_{\ell \in \mathbb{N}}$  be a sequence of positive numbers converging to 0, and let  $\widehat{Q}_t$  denote the weak limit of the sequence of probability measures  $(\widehat{Q}_t^{\gamma_\ell})_{\ell \in \mathbb{N}}$ . Let  $z \in \widehat{Z}$ .*

- For any  $z' \in V(z)$ , it holds

$$(44) \quad \lim_{\gamma_\ell \rightarrow 0} \left( \gamma_\ell A_t^{\gamma_\ell}(z, z') \right) = A_t(z, z') \geq 0 \quad \text{and} \quad \lim_{\gamma_\ell \rightarrow 0} \left( \gamma_\ell B_t^{\gamma_\ell}(z, z') \right) = B_t(z, z') \geq 0,$$

with  $A_t(z, z') > 0$  if and only if  $z' \in V_-(z)$ , and  $B_t(z, z') > 0$  if and only if  $z' \in V_-(z)$ . Moreover, given  $z' \in V_-(z)$ , for any  $y \in \widehat{Y}_z$

$$A_t(z, z') := \frac{a_t(z, z', y)}{a_t(z, y)},$$

and given  $z' \in V_-(z)$ , for any  $x \in \widehat{X}_z$

$$B_t(z, z') := \frac{b_t(z, z', x)}{b_t(z, x)}.$$

- For any  $z'' \in \mathbb{V}(z)$ , it holds

$$(45) \quad \lim_{\gamma_\ell \rightarrow 0} \left( \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'') \right) = A_t(z, z'') \geq 0 \quad \text{and} \quad \lim_{\gamma_\ell \rightarrow 0} \left( \gamma_\ell^2 B_t^{\gamma_\ell}(z, z'') \right) = B_t(z, z'') \geq 0,$$

with  $A_t(z, z'') > 0$  if and only if  $z'' \in \mathbb{V}_-(z)$  and  $B_t(z, z'') > 0$  if and only if  $z'' \in \mathbb{V}_-(z)$ . Moreover, given  $z'' \in \mathbb{V}_-(z)$ , for any  $y \in \widehat{Y}_z$

$$A_t(z, z'') := \frac{a_t(z, z'', y)}{a_t(z, y)},$$

and given  $z'' \in \mathbb{V}_\rightarrow(z)$ , for any  $x \in \widehat{X}_z$

$$B_t(z, z'') := \frac{b_t(z, z'', x)}{b_t(z, x)}.$$

Lemma 3.4 provides the following Taylor estimates for the functions  $\varphi''_{\gamma_\ell}$  and  $\psi''_{\gamma_\ell}$  as  $\gamma_\ell$  goes to 0, which are a key result of this paper.

**Theorem 3.5.** *Assume that conditions (13), (14) and (15) are fulfilled. Let  $(\gamma_\ell)_{\ell \in \mathbb{N}}$  be a sequence of positive numbers converging to 0 and  $\widehat{Q}_t$  denotes the weak limit of the sequence of probability measures  $(\widehat{Q}_t^{\gamma_\ell})_{\ell \in \mathbb{N}}$ . With the notations of Lemma 3.4, one has for any  $t \in (0, 1)$*

$$\begin{aligned} & \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) \\ & \geq \int \left[ \left( \sum_{z' \in \mathbb{V}_\leftarrow(z)} A_t(z, z') L(z, z') \right)^2 + \sum_{z' \in \mathbb{V}_\leftarrow(z), z'' \in \mathbb{V}_\leftarrow(z), z' \sim z''} \rho(A_t(z, z'), A_t(z, z'')) L(z', z'') L(z, z') \right] d\widehat{Q}_t(z) \\ & = \int \left[ \left( \sum_{z' \in \mathbb{V}(z)} A_t(z, z') L(z, z') \right)^2 + \sum_{z' \in \mathbb{V}(z), z'' \in \mathbb{V}(z), z' \sim z''} \rho(A_t(z, z'), A_t(z, z'')) L(z', z'') L(z, z') \right] d\widehat{Q}_t(z), \end{aligned}$$

and

$$\begin{aligned} & \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \\ & \geq \int \left[ \left( \sum_{z' \in \mathbb{V}_\rightarrow(z)} B_t(z, z') L(z, z') \right)^2 + \sum_{z' \in \mathbb{V}_\rightarrow(z), z'' \in \mathbb{V}_\rightarrow(z), z' \sim z''} \rho(B_t(z, z'), B_t(z, z'')) L(z', z'') L(z, z') \right] d\widehat{Q}_t(z) \\ & = \int \left[ \left( \sum_{z' \in \mathbb{V}(z)} B_t(z, z') L(z, z') \right)^2 + \sum_{z' \in \mathbb{V}(z), z'' \in \mathbb{V}(z), z' \sim z''} \rho(B_t(z, z'), B_t(z, z'')) L(z', z'') L(z, z') \right] d\widehat{Q}_t(z). \end{aligned}$$

**Comments.** *Let us briefly explain how to use this result. First, adding the two above inequalities of this Theorem provides a lower bound on the second derivative of the relative entropy along the Schrödinger path at zero temperature. Then, it remains to find good estimates of this lower bound to apply Lemma 3.1 in order to get entropic curvature lower-bounds for the graph. The following equalities are a main guideline for this estimation, one has*

$$\begin{aligned} \int \sum_{z' \in \mathbb{V}_\leftarrow(z)} A_t(z, z') L(z, z') d\widehat{Q}_t(z) &= \int \sum_{z \in \widehat{Z}} \sum_{z' \in \mathbb{V}_\leftarrow(z)} a_t(z, z', y) L(z, z') d\nu_1(y) \\ &= \int \sum_{w \in \mathcal{X}} d(y, w) \widehat{\pi}_\leftarrow(w|y) d\nu_1(y) \\ &= W_1(\nu_0, \nu_1), \end{aligned} \tag{46}$$

and similarly

$$\int \sum_{z' \in \mathbb{V}_\rightarrow(z)} B_t(z, z') L(z, z') d\widehat{Q}_t(z) = W_1(\nu_0, \nu_1), \tag{47}$$

but also

$$\begin{aligned} \int \sum_{z'' \in \mathbb{V}_\leftarrow(z)} A_t(z, z'') L^2(z, z'') d\widehat{Q}_t(z) &= \int \sum_{w \in \mathcal{X}} d(y, w)(d(y, w) - 1) \widehat{\pi}_\leftarrow(w|y) d\nu_1(y) \\ &= \iint d(x, y)(d(x, y) - 1) d\widehat{\pi}(x, y), \end{aligned}$$

and

$$\int \sum_{z'' \in \mathbb{V}_-(z)} \mathbb{B}_t(z, z'') L^2(z, z'') d\widehat{\mathcal{Q}}_t(z) = \iint d(x, y)(d(x, y) - 1) d\widehat{\pi}(x, y).$$

The easy proof of these equalities is left to the reader.

*Proof of Theorem 3.5.* We only present the proof of the lower bound of  $\liminf_{\gamma \ell \rightarrow 0} \varphi''_{\gamma \ell}(t)$  since by symmetry, identical arguments provide the lower bound of  $\liminf_{\gamma \ell \rightarrow 0} \psi''_{\gamma \ell}(t)$ . We start with the expression of  $\varphi''_{\gamma}(t)$  given by Lemma, 3.3, for  $t \in (0, 1)$

$$(48) \quad \varphi''_{\gamma}(t) = \int (M_t^{\gamma} + R_t^{\gamma}) d\widehat{\mathcal{Q}}_t^{\gamma},$$

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

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

and

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

We will get the behaviour of  $\varphi''_{\gamma}(t)$  as  $\gamma$  goes to zero by applying Fatou's Lemma. For that purpose, we need first to bound from below the function  $(M_t^{\gamma} + R_t^{\gamma}) \widehat{\mathcal{Q}}_t^{\gamma}$  uniformly in  $\gamma$  by some integrable function with respect to the counting measure on  $\mathcal{X}$ . Let us first lower bound  $M_t^{\gamma}(z)$  and bound  $|R_t^{\gamma}(z)|$  uniformly in  $\gamma$ , for  $\gamma$  sufficiently small for any  $z \in \mathcal{X}$ .

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  $a \geq 0$ ,

$$(49) \quad \inf_{b \geq 0} \rho(a, b) = -a^2,$$

As a consequence, according to the definition of  $A_t^{\gamma}$ , one has

$$M_t^{\gamma}(z) \geq - \sum_{z', z'', z \sim z' \sim z''} A_t^{\gamma}(z, z')^2 L^{\gamma}(z, z') L^{\gamma}(z', z'').$$

From hypothesis (13) and then applying inequality (83), it follows that for any  $z \in \mathcal{X}$

$$(50) \quad M_t^{\gamma}(z) \geq -\gamma^2 S^2 d_{\max}^2 \max_{z', z' \sim z} A_t^{\gamma}(z, z')^2 \geq -\frac{(d^2(x_0, z) + 1) K^{2d(x_0, z)} O(1)}{t^2}.$$

where  $x_0$  is a fixed point of  $\mathcal{X}$ ,  $K = 2S/I$  and  $O(1)$  denotes a positive constant that does not depend on  $z, \gamma, t$ . Similarly, from (13) and (83), one may show that

$$(51) \quad |R_t^{\gamma}(z)| \leq \frac{\gamma}{t} \left[ \log \left( \frac{1}{\gamma} \right) + d(x_0, z) \right] d(x_0, z) K^{d(x_0, z)} O(1) \leq \frac{|\gamma \log \gamma|}{t} d^2(x_0, z) K^{d(x_0, z)} O(1).$$

Lemma 4.4 (vii) therefore implies for any  $z \in \mathcal{X}$  and any  $0 \leq \gamma < \bar{\gamma} < 1$ ,

$$(M_t^{\gamma}(z) + R_t^{\gamma}(z)) \widehat{\mathcal{Q}}_t^{\gamma}(z) \geq -O(1) \left( \mathbb{1}_B(z) + \mathbb{1}_{\mathcal{X} \setminus B}(z) \bar{\gamma} (\bar{\gamma} K^2)^{[2d(x_0, z) - 4D - 1]_+} \right) (d^2(x_0, z) + 1) K^{2d(x_0, z)},$$

where

$$B := \bigcup_{x \in \text{supp}(v_0), y \in \text{supp}(v_1)} [x, y] \supset \widehat{Z}.$$

It remains to choose  $\bar{\gamma}$  such that  $(\bar{\gamma} K^3)^2 < \gamma_0$  so that hypothesis (15) implies

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

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

$$(52) \quad \lim_{\gamma \rightarrow 0} \varphi_\gamma''(t) \geq \sum_{z \in \mathcal{X}} \liminf_{\gamma \rightarrow 0} \left[ \left( M_t^{\gamma_\ell}(z) + R_t^{\gamma_\ell}(z) \right) \widehat{Q}_t^{\gamma_\ell}(z) \right] > -\infty.$$

The weak convergence of  $(\widehat{Q}_t^{\gamma_\ell})_\ell$  to  $\widehat{Q}$  implies  $\lim_{\gamma_\ell \rightarrow 0} \widehat{Q}_t^{\gamma_\ell}(z) = \widehat{Q}_t(z)$ , and the inequality (51) gives  $\lim_{\gamma_\ell \rightarrow 0} R_t^{\gamma_\ell}(z) = 0$  for any  $z \in \mathcal{X}$ . As a consequence,

$$\liminf_{\gamma_\ell \rightarrow 0} \left[ \left( M_t^{\gamma_\ell}(z) + R_t^{\gamma_\ell}(z) \right) \widehat{Q}_t^{\gamma_\ell}(z) \right] = \liminf_{\gamma_\ell \rightarrow 0} \left[ M_t^{\gamma_\ell}(z) \right] \widehat{Q}_t(z).$$

In order to complete the proof Proposition 3.5, it remains to bound from below  $\liminf_{\gamma_\ell \rightarrow 0} \left[ M_t^{\gamma_\ell}(z) \right]$  for any  $z \in \widehat{Z}$  since otherwise  $\widehat{Q}_t(z) = 0$ . One has  $M_t^{\gamma_\ell} = E_t^{\gamma_\ell} + F_t^{\gamma_\ell}$ , where for any  $z \in \widehat{Z}$ ,

$$E_t^{\gamma_\ell}(z) := \left( \sum_{z', z'' \sim z} \gamma_\ell A_t^{\gamma_\ell}(z, z') L(z, z') \right)^2 - \sum_{z', z'', z \sim z' \sim z''} \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'')^2 L(z, z') L(z', z''),$$

and

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

Lemma 3.4 implies

$$(53) \quad \lim_{\gamma_\ell \rightarrow 0} E_t^{\gamma_\ell}(z) = \left( \sum_{z' \in V_-(z)} A_t(z, z') L(z, z') \right)^2 - \sum_{z' \in V_-(z), z'' \in \mathcal{X}, z'' \sim z'} A_t(z, z')^2 L(z, z') L(z', z'').$$

Assume that  $z' \in V_-(z)$ , or equivalently  $\lim_{\gamma_\ell \rightarrow 0} \gamma_\ell A_t^{\gamma_\ell}(z, z') \neq 0$ . According to Lemma 3.4, for any  $z'' \sim z'$ , one has  $\lim_{\gamma_\ell \rightarrow 0} \left( \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'') \right) = 0$  if  $d(z, z'') \leq 1$  and  $\lim_{\gamma_\ell \rightarrow 0} \left( \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'') \right) = A_t(z, z'')$  if  $z'' \in \mathbb{V}(z)$ . As a consequence the continuity of the function  $\rho$  on the set  $(0, \infty) \times [0, \infty)$ , implies

$$\lim_{\gamma_\ell \rightarrow 0} \left[ \rho \left( \gamma_\ell A_t^{\gamma_\ell}(z, z'), \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'') \right) + \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'')^2 \right] = \rho \left( A_t(z, z'), A_t(z, z'') \right) \mathbb{1}_{z'' \in \mathbb{V}(z)} + A_t(z, z')^2.$$

If  $z' \in V(z) \setminus V_-(z)$ , or equivalently  $\lim_{\gamma_\ell \rightarrow 0} \gamma_\ell A_t^{\gamma_\ell}(z, z') = A_t(z, z') = 0$ , then identity (49) provides, according to the definition of the function  $\rho$ ,

$$\begin{aligned} \liminf_{\gamma_\ell \rightarrow 0} \left[ \rho \left( \gamma_\ell A_t^{\gamma_\ell}(z, z'), \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'') \right) + \gamma_\ell^2 A_t^{\gamma_\ell}(z, z'')^2 \right] \\ \geq 0 = \rho(0, A_t(z, z'')) = \rho(A_t(z, z'), A_t(z, z'')) \mathbb{1}_{z'' \in \mathbb{V}(z)} + A_t(z, z')^2. \end{aligned}$$

As a consequence, one gets

$$\begin{aligned} \liminf_{\gamma_\ell \rightarrow 0} F_t^{\gamma_\ell}(z) &\geq \sum_{z', z'', z \sim z' \sim z''} \left[ \rho(A_t(z, z'), A_t(z, z'')) \mathbb{1}_{z'' \in \mathbb{V}(z)} + A_t(z, z')^2 \right] \\ &= \sum_{z' \in V_-(z), z'' \in \mathbb{V}_-(z), z' \sim z''} \rho(A_t(z, z'), A_t(z, z'')) + \sum_{z' \in V_-(z), z'' \in \mathcal{X}, z'' \sim z'} A_t(z, z')^2 L(z, z') L(z', z''). \end{aligned}$$

This inequality together with (52) and (53) ends the proof of Theorem 3.5.  $\square$

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

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

*Proof of Theorem 2.3.* 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

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

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

$$\begin{aligned} \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) &\geq \int \left( \sum_{i=1}^n (A_{i+} + A_{i-}) \right)^2 d\widehat{Q}_t + \int \sum_{i=1}^n \left( \rho(A_{i+}, A_{i+i+}) + \rho(A_{i-}, A_{i-i-}) \right) d\widehat{Q}_t \\ &\quad + \frac{1}{2} \int \sum_{i,j,i \neq j} \left( \rho(A_{i+}, A_{j+i+}) + \rho(A_{j+}, A_{j+i+}) \right) + \left( \rho(A_{i-}, A_{j-i-}) + \rho(A_{j-}, A_{j-i-}) \right) \\ &\quad + \left( \rho(A_{i+}, A_{j-i+}) + \rho(A_{j-}, A_{j-i+}) \right) + \left( \rho(A_{i-}, A_{j+i-}) + \rho(A_{j+}, A_{j+i-}) \right) d\widehat{Q}_t. \end{aligned}$$

Identity (49) implies for any  $a, a', b \in \mathbb{R}_+$ ,

$$(54) \quad \rho(a, b) + \rho(a', b) = 2\rho(\sqrt{aa'}, b) \geq -2aa'.$$

It follows that

$$\begin{aligned} \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) &\geq \int \left( \sum_{i=1}^n (A_{i+} + A_{i-}) \right)^2 d\widehat{Q}_t - \int \sum_{i=1}^n (A_{i+}^2 + A_{i-}^2) d\widehat{Q}_t \\ &\quad - \int \sum_{i,j,i \neq j} (A_{i+}A_{j+} + A_{i-}A_{j-} + A_{i+}A_{j-} + A_{i-}A_{j+}) d\widehat{Q}_t \\ &= 2 \int \sum_{i=1}^n A_{i+}A_{i-} d\widehat{Q}_t \geq 0. \end{aligned}$$

Identically one proves that  $\liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq 0$ . Applying then Lemma 3.1 ends the proof of Theorem 2.3.  $\square$

#### 3.2.2. The complete graph.

*Proof of Theorem 2.4.* Since for any  $x, y \in X$ ,  $d(x, y) = 1$ , Theorem 3.5 and Lemma 3.4 provide for any  $t \in (0, 1)$

$$\begin{aligned} \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) &\geq \int \left( \sum_{z' \in V_-(z)} A_t(z, z') L(z, z') \right)^2 d\widehat{Q}_t(z) = \iint \left( \sum_{z' \in V_-(z)} A_t(z, z') L(z, z') \right)^2 d\widehat{Q}_{t,1}(z, y) \\ &= \int \sum_{z \in \widehat{Z}^y} \left( \sum_{z' \in V_-(z)} \frac{a_t(z, z', y)}{a_t(z, y)} L(z, z') \right)^2 a_t(z, y) d\nu_1(y) \end{aligned}$$

With the expression (22) of  $Q_t^{x,y}$ , one easily check that for any  $z \in \widehat{Z}, y \in \widehat{Y}_z$ , or equivalently for any  $y \in \text{supp}(\nu_1), z \in \widehat{Z}^y$ ,

$$a_t(z, y) = (1-t)\widehat{\pi}_-(z|y) + t\delta_y(z),$$

and with (42), for any  $z' \in V_-(z)$ ,

$$a_t(z, z', y) = \mathbf{1}_{z=y} \frac{\widehat{\pi}_-(z'|y)}{\mu(z')}.$$



As a consequence, one gets

$$\begin{aligned} \int \sum_{z \in \mathcal{Z}'} \left( \sum_{z' \in V_{\leftarrow}(z)} \frac{a_t(z, z', y)}{a_t(z, y)} L(z, z') \right)^2 a_t(z, y) dv_1(y) &= \int \left( \sum_{z' \in V_{\leftarrow}(y)} \frac{a_t(y, z', y)}{a_t(y, y)} \mu(z') \right)^2 a_t(y, y) dv_1(y) \\ &= \int \frac{(1 - \widehat{\pi}_{\leftarrow}(y|y))^2}{1 - (1-t)(1 - \widehat{\pi}_{\leftarrow}(y|y))} dv_1(y) = \int \frac{1}{2} (1 - \widehat{\pi}_{\leftarrow}(y|y))^2 h''((1-t)(1 - \widehat{\pi}_{\leftarrow}(y|y))) dv_1(y) = \xi_{\leftarrow}''(t), \end{aligned}$$

where for any  $t \in [0, 1]$ ,

$$\xi_{\leftarrow}(t) := \frac{1}{2} \int h((1-t)(1 - \widehat{\pi}_{\leftarrow}(y|y))) dv_1(y).$$

One similarly shows that for any  $t \in (0, 1)$ ,

$$\liminf_{\gamma_t \rightarrow 0} \psi_{\gamma_t}''(t) \geq \xi_{\rightarrow}''(t),$$

with  $\xi_{\rightarrow}(t) := \frac{1}{2} \int h(t(1 - \widehat{\pi}_{\rightarrow}(x|x))) dv_0(x)$ . The proof of Theorem 2.4 ends applying Lemma 3.1 and the two following identities

$$(1-t)\xi_{\leftarrow}(0) + t\xi_{\leftarrow}(1) - \xi_{\leftarrow}(t) = \frac{t(1-t)}{2} \int h_{1-t} \left( \int \mathbb{1}_{w \neq y} d\widehat{\pi}_{\leftarrow}(w|y) \right) dv_1(y),$$

and

$$(1-t)\xi_{\rightarrow}(0) + t\xi_{\rightarrow}(1) - \xi_{\rightarrow}(t) = \frac{t(1-t)}{2} \int h_t \left( \int \mathbb{1}_{w \neq x} d\widehat{\pi}_{\rightarrow}(w|x) \right) dv_0(x).$$

Let us now compare  $C_t(\widehat{\pi})$  with a function of  $W_1(\nu_0, \nu_1)$ . Observe that for any  $y \in \text{supp}(\nu_1)$ ,  $\int \mathbb{1}_{w \neq y} d\widehat{\pi}_{\leftarrow}(w|y) \neq 0$ , if and only if  $y$  belongs to the set

$$D_{\leftarrow} := \left\{ w \in \text{supp}(\nu_1) \mid \exists x \in \mathcal{X}, w \neq x, (x, w) \in \text{supp}(\widehat{\pi}) \right\}.$$

Since  $h_{1-t}(0) = 0$  and  $h_{1-t}$  is convex, Jensen's inequality provides

$$\int h_{1-t} \left( \int \mathbb{1}_{w \neq y} d\widehat{\pi}_{\leftarrow}(w|y) \right) dv_1(y) \geq \nu_1(D_{\leftarrow}) h_{1-t} \left( \frac{\iint \mathbb{1}_{w \neq y} d\widehat{\pi}_{\leftarrow}(w|y) dv_1(y)}{\nu_1(D_{\leftarrow})} \right) = \nu_1(D_{\leftarrow}) h_{1-t} \left( \frac{W_1(\nu_0, \nu_1)}{\nu_1(D_{\leftarrow})} \right).$$

Similarly one has

$$\int h_t \left( \int \mathbb{1}_{w \neq x} d\widehat{\pi}_{\rightarrow}(w|x) \right) dv_0(x) \geq \nu_0(D_{\rightarrow}) h_t \left( \frac{\iint \mathbb{1}_{w \neq x} d\widehat{\pi}_{\rightarrow}(w|x) dv_0(x)}{\nu_0(D_{\rightarrow})} \right) = \nu_0(D_{\rightarrow}) h_t \left( \frac{W_1(\nu_0, \nu_1)}{\nu_0(D_{\rightarrow})} \right),$$

with

$$D_{\rightarrow} := \left\{ w \in \text{supp}(\nu_0) \mid \exists y \in \mathcal{X}, w \neq y, (w, y) \in \text{supp}(\widehat{\pi}) \right\}.$$

According to (23),  $W_1(\nu_0, \nu_1) \geq \nu_0(D_{\rightarrow}) - \nu_1(D_{\leftarrow})$ , and we know from Lemma 4.3 (iii) that the sets  $D_{\leftarrow}$  and  $D_{\rightarrow}$  are disjoint. As a consequence,

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

This leads to the expected result (24) :

$$\begin{aligned} C_t(\widehat{\pi}) &\geq (1 + W_1(\nu_0, \nu_1)) \inf_{\alpha, \beta, 0 < \alpha + \beta \leq 1} \left\{ \alpha h_t \left( \frac{W_1(\nu_0, \nu_1)}{\alpha(1 + W_1(\nu_0, \nu_1))} \right) + \beta h_{1-t} \left( \frac{W_1(\nu_0, \nu_1)}{\beta(1 + W_1(\nu_0, \nu_1))} \right) \right\}, \\ &= (1 + W_1(\nu_0, \nu_1)) k_t \left( \frac{W_1(\nu_0, \nu_1)}{1 + W_1(\nu_0, \nu_1)} \right). \end{aligned}$$

In order to prove the estimate (25) of the function  $k_t$ , one first observes that by construction, for any  $t \in (0, 1)$  and  $v \in [0, 1]$ ,

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

and since  $K_t(u) = K_{1-t}(1-u)$ ,

$$h_{1-t}(v) = \int_0^1 \frac{v^2}{1-(1-u)v} K_t(u) du.$$

Since  $h_t(u) = +\infty$ , for  $u > 1$ , it follows that for any  $v \in [0, 1/2]$ ,

$$\begin{aligned} k_t(v) &= \inf_{\alpha, \beta, 0 < \alpha + \beta \leq 1} \left\{ \alpha h_t\left(\frac{v}{\alpha}\right) + \beta h_{1-t}\left(\frac{v}{\beta}\right) \right\} \geq \inf_{\alpha, \beta, \alpha > v, \beta > v, \alpha + \beta \leq 1} \left\{ \alpha h_t\left(\frac{v}{\alpha}\right) + \beta h_{1-t}\left(\frac{v}{\beta}\right) \right\} \\ &\geq \int_0^1 v^2 \inf_{\alpha, \beta, \alpha > v, \beta > v, \alpha + \beta \leq 1} \left\{ \frac{1}{\alpha - uv} + \frac{1}{\beta - (1-u)v} \right\} K_t(u) du. \end{aligned}$$

Easy computations give

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

It provides the expected estimate (25), namely  $k_t(v) \geq \frac{4v^2}{1-v}$ .  $\square$

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

*Proof of Theorem 2.5.* The first step of the proof is to express the lower bounds on  $\liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t)$  and  $\liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t)$  given by Theorem 3.5 using the symmetries of the graph structure of the hypercube, and keeping in mind the comments given next to Theorem 3.5. This leads to the estimates (57) and (58). The second step is to prove that each of the lower bound on  $C_t(\widehat{\pi})$  in Theorem 2.5 is a consequence of these estimates.

**Step 1 :** Given  $z \in \widehat{Z}$ , let us define the sets

$$I^{\leftarrow}(z) := \{i \in [n] \mid \sigma_i(z) \in V_{\leftarrow}(z)\} = \{i \in [n] \mid (z, \sigma_i(z)) \in C_{\leftarrow}\},$$

$$I^{\rightarrow}(z) := \{i \in [n] \mid \sigma_i(z) \in V_{\rightarrow}(z)\} = \{i \in [n] \mid (z, \sigma_i(z)) \in C_{\rightarrow}\},$$

$$\mathbb{I}^{\leftarrow}(z) := \{(i, j) \in [n] \times [n] \mid i \neq j, \sigma_i \sigma_j(z) \in \mathbb{V}_{\leftarrow}(z)\} = \{(i, j) \in [n] \times [n] \mid (z, \sigma_i \sigma_j(z)) \in C_{\leftarrow}\},$$

$$\mathbb{I}^{\rightarrow}(z) := \{(i, j) \in [n] \times [n] \mid \sigma_i \sigma_j(z) \in \mathbb{V}_{\rightarrow}(z)\} = \{(i, j) \in [n] \times [n] \mid (z, \sigma_i \sigma_j(z)) \in C_{\rightarrow}\}.$$

and  $\mathbb{I}_1^{\leftarrow}(z) := \{i \in [n] \mid \exists j \in [n], (i, j) \in \mathbb{I}^{\leftarrow}(z)\}$ ,  $\mathbb{I}_1^{\rightarrow}(z) := \{i \in [n] \mid \exists j \in [n], (i, j) \in \mathbb{I}^{\rightarrow}(z)\}$ . Observe that if  $\mathbb{I}^{\leftarrow}(z) \neq \emptyset$  then  $|\mathbb{I}_1^{\leftarrow}(z)| \geq 2$ . Obviously one has  $\mathbb{I}_1^{\leftarrow}(z) \subset I^{\leftarrow}(z)$  and since  $\sigma_i \sigma_j = \sigma_j \sigma_i$ , one has  $\mathbb{I}^{\leftarrow}(z) = \{(i, j) \mid i, j \in \mathbb{I}_1^{\leftarrow}(z), i \neq j\}$ . Same remarks hold with the sets  $I^{\rightarrow}(z), \mathbb{I}_1^{\rightarrow}(z), \mathbb{I}^{\rightarrow}(z)$ . The sets  $C_{\leftarrow}$  and  $C_{\rightarrow}$  are disjoint and therefore  $I^{\leftarrow}(z)$  and  $I^{\rightarrow}(z)$  are also disjoint. To simplify, for  $z \in \widehat{Z}$  and  $i, j \in [n], i \neq j$  let us denote

$$A_i(z) := A_t(z, \sigma_i(z)), \quad A_{ij}(z) := A_t(z, \sigma_i \sigma_j(z)), \quad L_i(z) := L(z, \sigma_i(z)).$$

Since for any  $i \neq j$ ,  $\sigma_i \sigma_j = \sigma_j \sigma_i$ , one has  $A_{ij} = A_{ji}$ , and observing that  $L(\sigma_i(z), \sigma_j \sigma_i(z)) = L_j(z)$  Theorem 3.5 provides after symmetrization

$$(55) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) \geq \int \left( \sum_{i \in I^{\leftarrow}} A_i L_i \right)^2 d\widehat{Q}_t + \int \sum_{\{i, j\} \subset \mathbb{I}_1^{\leftarrow}} [\rho(A_i, A_{ij}) + \rho(A_j, A_{ij})] L_i L_j d\widehat{Q}_t.$$

Let  $\mathbb{A} := \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} 2\mathbb{A}_{ij}L_iL_j$  and  $\beta_{ij} := \frac{2\mathbb{A}_{ij}L_iL_j}{\mathbb{A}}$ . According to the definition of the function  $\rho$  given in Lemma 3.3, computations provide

$$\begin{aligned}
& \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} \left[ \rho(A_i, A_{ij}) + \rho(A_j, A_{ij}) \right] L_i L_j \\
&= \mathbb{A} \log \mathbb{A} - \mathbb{A} + \mathbb{A} \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} \beta_{ij} \log(\beta_{ij}) - \mathbb{A} \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} \log(2A_i A_j) \beta_{ij} \\
(56) \quad & \geq \mathbb{A} \log \mathbb{A} - \mathbb{A} + \mathbb{A} \log \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} 2A_i A_j,
\end{aligned}$$

where the last inequality follows from the duality formula between the log-Laplace transform and the entropy, namely in this case

$$\sup_{\beta} \left\{ \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} \log(2A_i A_j) \beta_{ij} - \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} \beta_{ij} \log(\beta_{ij}) \right\} = \log \sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} 2A_i A_j,$$

where the supremum runs over all probabilities  $\beta$  on  $\mathbb{I}_1^{\leftarrow}$ . Note that  $\mathbb{A} \neq 0$  if and only if  $|\mathbb{I}_1^{\leftarrow}| \geq 2$  and therefore  $\sum_{\{i,j\} \subset \mathbb{I}_1^{\leftarrow}} 2A_i A_j > 0$ . It follows that all quantities above are well defined. Setting  $A := \sum_{i \in I^{\leftarrow}} A_i L_i$  and  $\widetilde{A}^2 := \sum_{i \in I^{\leftarrow}} (A_i L_i)^2$ , since  $\mathbb{I}_1^{\leftarrow} \subset I^{\leftarrow}$ , (55) and (56) finally give the following lower-bound

$$(57) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) \geq \int \left[ A^2 - \mathbb{A} + \mathbb{A} \left( \log \mathbb{A} - \log(A^2 - \widetilde{A}^2) \right) \mathbb{1}_{|I^{\leftarrow}| \geq 2} \right] d\widehat{Q}_t.$$

From the lower-bound of  $\liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t)$  given by Theorem 2.5, following the same lines of proof one gets

$$(58) \quad \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \left[ B^2 - \mathbb{B} + \mathbb{B} \left( \log \mathbb{B} - \log(B^2 - \widetilde{B}^2) \right) \mathbb{1}_{|I^{\leftarrow}| \geq 2} \right] d\widehat{Q}_t,$$

where we set for any  $z \in \widetilde{Z}$

$$B(z) := \sum_{i \in I^{\rightarrow}} B_i(z, \sigma_i(z)) L_i(z), \quad \widetilde{B}^2(z) := \sum_{i \in I^{\rightarrow}} B_i^2(z, \sigma_i(z)) L_i^2(z), \quad \mathbb{B}(z) := \sum_{\{i,j\} \subset \mathbb{I}_1^{\rightarrow}} \mathbb{B}_t(z, \sigma_i \sigma_j(z)) L_i(z) L_j(z).$$

**Step 2 :** By the Cauchy-Schwarz inequality  $\widetilde{A}^2 \geq A^2/|I^{\leftarrow}|$  and therefore (57) gives

$$\begin{aligned}
(59) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) & \geq \int \left[ A^2 - \mathbb{A} + \mathbb{A} \left( \log \mathbb{A} - \log(A^2) - \log \left( 1 - \frac{1}{|I^{\leftarrow}|} \right) \right) \mathbb{1}_{|I^{\leftarrow}| \geq 2} \right] d\widehat{Q}_t \\
& \geq \int \max \left\{ \frac{A^2}{|I^{\leftarrow}|}, -\log \left( 1 - \frac{1}{|I^{\leftarrow}|} \right) \mathbb{A} \right\} d\widehat{Q}_t \\
& \geq \int \frac{1}{|I^{\leftarrow}|} \max \{ A^2, \mathbb{A} \} d\widehat{Q}_t
\end{aligned}$$

where the last inequalities follows from the concavity property of the logarithmic function and since  $A^2/|I^{\leftarrow}| = A^2 \mathbb{1}_{|I^{\leftarrow}|=1} + A^2/|I^{\leftarrow}| \mathbb{1}_{|I^{\leftarrow}| \geq 2}$ . Identically (58) implies

$$(60) \quad \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \frac{1}{|I^{\rightarrow}|} \max \{ B^2, \mathbb{B} \} d\widehat{Q}_t$$

Keeping only the quantities involving  $A$  and  $B$  in (59) and (60), and applying Cauchy-Schwarz inequality, the identities (46) and (47) yield

$$\liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq W_1^2(\nu_0, \nu_1) \left[ \frac{1}{\int |I^{\leftarrow}| d\widehat{Q}_t} + \frac{1}{\int |I^{\rightarrow}| d\widehat{Q}_t} \right].$$

Since, the sets  $I^\leftarrow$  and  $I^\rightarrow$  are disjoint  $\int |I^\leftarrow| d\widehat{Q}_t + \int |I^\rightarrow| d\widehat{Q}_t \leq n$ , and therefore the identity  $\min_{\alpha, \beta > 0, \alpha + \beta \leq 1} \left\{ \frac{1}{\alpha} + \frac{1}{\beta} \right\} = 4$  implies

$$\liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \frac{4}{n} W_1^2(\nu_0, \nu_1).$$

Then applying Lemma 3.1, this estimate give the first lower bound of  $C_t(\widehat{\pi})$  in Theorem 2.5.

Keeping only the quantities involving  $\mathbb{A}$  and  $\mathbb{B}$  in (59) and (60), one gets

$$(61) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \frac{\mathbb{A}}{|I^\leftarrow|} + \frac{\mathbb{B}}{|I^\rightarrow|} d\widehat{Q}_t.$$

According to Lemma 3.4 and (43), for any  $i, j \in [n]$  with  $i \neq j$ , for any  $z \in \widehat{Z}$  and  $y \in \widehat{Y}_z$ ,

$$2A_t(z, \sigma_j \sigma_i(z)) L_i(z) L_j(z) := \frac{a_t(z, \sigma_j \sigma_i(z), y) 2L_i(z) L_j(z)}{a_t(z, y)}$$

with

$$a_t(z, \sigma_j \sigma_i(z), y) := \sum_{w, (z, \sigma_i \sigma_j(z)) \in [y, w]} r(y, z, \sigma_i \sigma_j(z), w) d(y, w) (d(y, w) - 1) \rho_t^{d(y, w) - 2} (d(z, w) - 2) \widehat{\pi}_-(w|y).$$

From the identity

$$\sum_{\{i, j\} \subset \mathbb{I}_t^\leftarrow} r(y, z, \sigma_i \sigma_j(z), w) L^2(z, \sigma_i \sigma_j(z)) = r(y, z, z, w),$$

and since  $L^2(z, \sigma_i \sigma_j(z)) = 2L_i(z) L_j(z)$  one has for any  $z \in \widehat{Z}$  and  $y \in \widehat{Y}_z$ ,

$$\mathbb{A}(z) = \frac{1}{a_t(z, y)} \sum_{w \in X} \sum_{z, z \in [y, w]} r(y, z, z, w) d(y, w) (d(y, w) - 1) \rho_t^{d(y, w) - 2} (d(z, w) - 2) \widehat{\pi}_-(w|y).$$

Working identically with  $\mathbb{B}(z)$  one finally gets

$$(62) \quad \int \frac{\mathbb{A}}{|I^\leftarrow|} + \frac{\mathbb{B}}{|I^\rightarrow|} d\widehat{Q}_t \geq \iint c_t(x, y) d\widehat{\pi}(x, y),$$

where

$$c_t(x, y) = \sum_{z \in [x, y]} \frac{1}{|I^\leftarrow(z)|} r(x, z, z, y) d(x, y) (d(x, y) - 1) \rho_t^{d(x, y) - 2} (d(x, z) - 2) \\ + \sum_{z \in [x, y]} \frac{1}{|I^\rightarrow(z)|} r(x, z, z, y) d(x, y) (d(x, y) - 1) \rho_t^{d(x, y) - 2} (d(x, z)).$$

Since  $|I^\leftarrow| \leq n$  and  $|I^\rightarrow| \leq n$  and for any  $k \in \{0, \dots, d(x, y)\}$ ,  $\sum_{z \in [x, y], d(x, z) = k} r(x, z, z, y) = 1$ , it follows that

$$(63) \quad c_t(x, y) \geq \frac{2}{n} d(x, y) (d(x, y) - 1).$$

For large values of  $d(x, y)$ , this lower bound can be improved using the fact that  $I^\leftarrow$  and  $I^\rightarrow$  are disjoint and therefore  $|I^\leftarrow| + |I^\rightarrow| \leq n$ . By first rewriting  $c_t(x, y)$ , applying Cauchy-Schwarz inequality, and then

using the identity  $\inf_{\alpha>0,\beta>0,\alpha+\beta\leq 1} \left\{ \frac{u^2}{\alpha} + \frac{v^2}{\beta} \right\} = (u+v)^2$ ,  $u, v \geq 0$ , one gets

$$\begin{aligned} c_t(x, y) &= \sum_{z, z \in [x, y]} \left( \frac{d(x, z)(d(x, z) - 1)}{|I^{\leftarrow}(z)|t^2} + \frac{d(z, y)(d(z, y) - 1)}{|I^{\rightarrow}(z)|(1-t)^2} \right) Q_t^{x, y}(z) \\ &\geq \frac{\left( \int \sqrt{d(x, z)(d(x, z) - 1)} dQ_t^{x, y}(z) \right)^2}{t^2 \int I^{\leftarrow} dQ_t^{x, y}} + \frac{\left( \int \sqrt{d(z, y)(d(z, y) - 1)} dQ_t(z)^{x, y}(z) \right)^2}{(1-t)^2 \int I^{\rightarrow} dQ_t^{x, y}} \\ &\geq \frac{1}{n} \left( \int \frac{\sqrt{d(x, z)(d(x, z) - 1)}}{t} dQ_t^{x, y}(z) + \int \frac{\sqrt{d(z, y)(d(z, y) - 1)}}{1-t} dQ_t^{x, y}(z) \right)^2 \\ &= \frac{4}{n} v_t(d(x, y)), \end{aligned}$$

with for any  $d \in \mathbb{N}$ ,

$$v_t(d) = \frac{1}{4} \left( \sum_{k=0}^d \sqrt{k(k-1)} \left( \frac{\rho_t^d(k)}{t} + \frac{\rho_{1-t}^d(k)}{1-t} \right) \right)^2.$$

Then applying Lemma 3.1 together with (61), (62), (63) provides the following lower bound on the cost  $C_t(\widehat{\pi})$ ,

$$C_t(\widehat{\pi}) \geq \frac{4}{n} \iint w_t(d(x, y)) d\widehat{\pi}(x, y)$$

with

$$w_t(d) := \max \left\{ \frac{d(d-1)}{2}, \int_0^1 v_s(d) K_t(s) ds \right\}, \quad d \in \mathbb{N}.$$

The proof of the second lower bound on  $C_t(\widehat{\pi})$  ends from the next estimate of the quantity  $\int_0^1 v_s(d) K_t(s) ds$ . Since for any  $s \in (0, 1)$  and  $d \in \mathbb{N}$ , one has

$$\begin{aligned} v_s(d) &= \frac{1}{4} \left( 2d - \sum_{k=1}^d \frac{k}{k + \sqrt{k(k-1)}} \left( \frac{\rho_t^d(k)}{t} + \frac{\rho_{1-t}^d(k)}{1-t} \right) \right)^2 \\ &\geq d^2 - d \sum_{k=1}^d \left( \frac{\rho_t^d(k)}{t} + \frac{\rho_{1-t}^d(k)}{1-t} \right) \geq d^2 - d \left( \frac{1 - \rho_t^d(0)}{t} + \frac{1 - \rho_{1-t}^d(0)}{1-t} \right), \end{aligned}$$

it follows that for any  $t \in (0, 1)$

$$\int_0^1 v_s(d) K_t(s) ds \geq d^2 - d \int_0^1 \left( \frac{1 - \rho_s^d(0)}{s} + \frac{1 - \rho_{1-s}^d(0)}{1-s} \right) K_t(s) ds,$$

with for  $d \geq 1$ ,

$$\begin{aligned} \int_0^1 \left( \frac{1 - \rho_s^d(0)}{s} + \frac{1 - \rho_{1-s}^d(0)}{1-s} \right) K_t(s) ds &= 2 \sum_{k=1}^d \left( (1-t)^{k-1} + t^{k-1} \right) \left( \frac{1}{k} - \frac{1}{d+1} \right) \\ &\leq 2 \left( 1 + \sum_{k=2}^d \frac{1}{k} \right) \leq 2 + 2 \log d. \end{aligned}$$

For the proof of third lower bound on  $C_t(\widehat{\pi})$ , one uses again (57) and (58) with the concavity of the logarithmic function to obtain

$$(64) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \widetilde{A}^2 d\widehat{Q}_t + \int \widetilde{B}^2 d\widehat{Q}_t.$$

According to the definition of  $\widetilde{A}^2$ ,

$$\int \widetilde{A}^2 d\widehat{Q}_t = \sum_{i=1}^n \int \sum_{z \in E_i^{\leftarrow}(y)} (A_i(z) L_i(z))^2 a_t(z, y) dv_1(y),$$

where  $E_i^{\leftarrow}(y) := \{z \in \widehat{Z}^y \mid y \in \widehat{Y}_{(z, \sigma_i(z))}\}$  for any  $y \in \text{supp}(v_1)$ . Easy computations give

$$\sum_{z \in E_i^{\leftarrow}(y)} A_i(z) L_i(z) a_t(z, y) = \sum_{z \in E_i^{\leftarrow}(y)} a_t(z, \sigma_i(z), y) L_i(z) = \Pi_{\leftarrow}^i(y).$$

and therefore by the Cauchy-Schwarz inequality

$$\int \widetilde{A}^2 d\widehat{Q}_t \geq \sum_{i=1}^n \frac{\Pi_{\leftarrow}^i(y)^2}{\sum_{z \in E_i^{\leftarrow}(y)} a_t(z, y)}.$$

If  $z \in E_i^{\leftarrow}(y)$  then  $z_i = y_i$  and therefore

$$\sum_{z \in E_i^{\leftarrow}(y)} a_t(z, y) \leq 1 - \sum_{z \in \widehat{Z}^y} \mathbb{1}_{z_i \neq y_i} a_t(z, y).$$

From the definition (41) of  $a_t(z, y)$ , and observing that if  $z \in [y, w]$  and  $z_i \neq y_i$  then necessarily  $z_i = w_i$ , one gets

$$\sum_{z \in \widehat{Z}^y} \mathbb{1}_{z_i \neq y_i} a_t(z, y) = \sum_{w \in \{0,1\}^n} \mathbb{1}_{w_i \neq y_i} \left( \sum_{z \in [y, w]} \mathbb{1}_{z_i = w_i} (1-t)^{d(y,z)} t^{d(z,w)} \right) \widehat{\pi}_{\leftarrow}(w|y) = (1-t) \sum_{w \in \{0,1\}^n} \mathbb{1}_{w_i \neq y_i} \widehat{\pi}_{\leftarrow}(w|y),$$

and therefore

$$\int \widetilde{A}^2 d\widehat{Q}_t \geq \sum_{i=1}^n \sum_{y \in \{0,1\}^n} \frac{\Pi_{\leftarrow}^i(y)^2 v_1(y)}{1 - (1-t) \Pi_{\leftarrow}^i(y)}.$$

This inequality implies (as in the proof of Theorem 2.4) for any  $t \in (0, 1)$

$$\int \widetilde{A}^2 d\widehat{Q}_t \geq \xi_{\leftarrow}''(t), \quad \text{with} \quad \xi_{\leftarrow}(t) := \frac{1}{2} \int \sum_{i=1}^n h((1-t) \Pi_{\leftarrow}^i(y)) dv_1(y).$$

Similar computations with the quantity  $\int \widetilde{B}^2 d\widehat{Q}_t$  and (64) finally provide

$$\liminf_{\gamma_\ell \rightarrow 0} \varphi_{\gamma_\ell}''(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi_{\gamma_\ell}''(t) \geq \xi_{\leftarrow}''(t) + \xi_{\rightarrow}''(t),$$

with  $\xi_{\rightarrow}(t) := \frac{1}{2} \int \sum_{i=1}^n h(t \Pi_{\rightarrow}^i(x)) dv_0(x)$ . Following the proof of Theorem 2.4, the two above estimates yield the third lower bound of  $C_t(\widehat{\pi})$ .  $\square$

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

*Proof of Theorem 2.6.* Let us note  $n' = \lceil N/2 \rceil$  where  $\lceil \cdot \rceil$  denotes the ceiling function. Let  $y \in \text{supp}(v_1) \subset \mathbb{Z}/N\mathbb{Z}$ , and  $z \in \widehat{Z}^y$ . We observe that if  $\{w \in \mathbb{Z}/N\mathbb{Z} \mid (z, z-1) \in [y, w]\} \neq \emptyset$  then necessarily  $(z-1, z) \in [y+n', y]$  and if  $\{w \in \mathbb{Z}/N\mathbb{Z} \mid (z, z+1) \in [y, w]\} \neq \emptyset$  then necessarily  $(z, z+1) \in [y, y+n]$ . As a consequence, since the sets  $\{z \in \mathbb{Z}/N\mathbb{Z} \mid (z, z+1) \in [y, y+n]\}$  and  $\{z \in \mathbb{Z}/N\mathbb{Z} \mid (z-1, z) \in [y+n', y]\}$  are disjoint, the sets  $\{z \in \widehat{Z}^y \mid y \in \widehat{Y}_{(z, z+1)}\}$  and  $\{z \in \widehat{Z}^y \mid y \in \widehat{Y}_{(z, z-1)}\}$  are also disjoint. It follows that

$$\int \sum_{z \in \widehat{Z}^y} (A_t(z, z+1) + A_t(z, z-1))^2 a_t(z, y) dv_1(y) = \int \sum_{z \in \widehat{Z}^y} (A_t^2(z, z+1) + A_t^2(z, z-1)) a_t(z, y) dv_1(y).$$

Therefore Theorem 3.5 together with (49) provide

$$\begin{aligned} \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) &\geq \int \sum_{z \in \widehat{\mathcal{Z}}} \left( A_t^2(z, z+1) + A_t^2(z, z-1) \right) + \rho(A_t(z, z+1), A_t(z, z+2)) \\ &\quad + \rho(A_t(z, z-1), A_t(z, z-2)) a_t(z, y) d\nu_1(y) \\ &\geq 0 \end{aligned}$$

Identically one proves that  $\liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq 0$ . The proof of Theorem 2.6 ends applying Lemma 3.1.  $\square$

### 3.2.5. The Bernoulli-Laplace model.

*Proof of Theorem 2.7.* One follows the same strategy as for the proof of Theorem 2.5. As a first step, the geometric structure of the slices of the cube provides estimates of the lower lower bounds on  $\liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t)$  and  $\liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t)$  given by Theorem 3.5. In the second step, one explains how these estimates (namely (67) and (68)) imply each of the lower bound on  $C_t(\widehat{\pi})$  given by Theorem 2.7.

**Step 1 :** For  $z \in \widehat{\mathcal{Z}}$ , one defines the sets

$$\begin{aligned} I^{\leftarrow}(z) &:= \left\{ (i, j) \in J_0(z) \times J_1(z) \mid (z, \sigma_{ij}(z)) \in C_{\leftarrow} \right\}, \\ I^{\rightarrow}(z) &:= \left\{ (i, j) \in J_0(z) \times J_1(z) \mid (z, \sigma_{ij}(z)) \in C_{\rightarrow} \right\}, \\ \mathbb{I}^{\leftarrow}(z) &:= \left\{ ((i, j), (k, l)) \in (J_0(z) \times J_1(z))^2 \mid (i, j) \neq (k, l), \sigma_{kl}\sigma_{ij}(z) \in \mathbb{V}_{\leftarrow}(z) \right\}, \\ \mathbb{I}^{\rightarrow}(z) &:= \left\{ ((i, j), (k, l)) \in (J_0(z) \times J_1(z))^2 \mid (i, j) \neq (k, l), \sigma_{kl}\sigma_{ij}(z) \in \mathbb{V}_{\rightarrow}(z) \right\}, \\ \mathbb{I}_1^{\leftarrow}(z) &:= \left\{ (i, j) \in J_0(z) \times J_1(z) \mid \exists (k, l) \in J_0(z) \times J_1(z), ((i, j), (k, l)) \in \mathbb{I}^{\leftarrow}(z) \right\}, \\ \mathbb{I}_1^{\rightarrow}(z) &:= \left\{ (i, j) \in J_0(z) \times J_1(z) \mid \exists (k, l) \in J_0(z) \times J_1(z), ((i, j), (k, l)) \in \mathbb{I}^{\rightarrow}(z) \right\}, \end{aligned}$$

The sets  $I^{\leftarrow}(z)$  and  $I^{\rightarrow}(z)$  are disjoint since  $C_{\leftarrow} \cap C_{\rightarrow} = \emptyset$ . Obviously one has  $\mathbb{I}_1^{\leftarrow}(z) \subset I^{\leftarrow}(z)$ . Observe that  $\sigma_{ki}\sigma_{ij}(z) = \sigma_{kj}(z)$  so that  $d(z, \sigma_{ki}\sigma_{ij}(z)) = 1$  and similarly  $d(z, \sigma_{jl}\sigma_{ij}(z)) = 1$ . It follows that if  $((i, j), (k, l)) \in \mathbb{I}^{\leftarrow}(z)$  or  $((i, j), (k, l)) \in \mathbb{I}^{\rightarrow}(z)$ , then the indices  $i, j, k, l$  all differ and  $\sigma_{kl}\sigma_{ij}(z) = \sigma_{ij}\sigma_{kl}(z)$ . As a consequence one has  $\mathbb{I}^{\rightarrow}(z) = \{((i, j), (k, l)) \mid ((i, j), (k, l)) \in \mathbb{I}_1^{\leftarrow}(z), i \neq k, j \neq l\}$ . Same remarks hold with the sets  $I^{\rightarrow}(z), \mathbb{I}_1^{\rightarrow}(z), \mathbb{I}^{\leftarrow}(z)$ . To simplify, one denotes  $A_{ij}(z) := A_t(z, \sigma_{ij}(z))$  and  $A_{kl,ij}(z) := A_t(z, \sigma_{kl}\sigma_{ij}(z))$ . After symmetrization, Theorem 3.5 provides

$$(65) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) \geq \int \left( \sum_{(i,j) \in I^{\leftarrow}} A_{ij} \right)^2 d\widehat{\mathcal{Q}}_t + \int \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} \left[ \rho(A_{ij}, A_{kl,ij}) + \rho(A_{ij}, A_{kl,ij}) \right] d\widehat{\mathcal{Q}}_t.$$

Setting  $\mathbb{A} := \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} 2A_{kl,ij}$ ,  $\beta_{kl,ij} := \frac{2A_{kl,ij}}{\mathbb{A}}$ , according to the definition of the function  $\rho$  given in Lemma 3.3, easy computations provides

$$\begin{aligned} &\sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} \left[ \rho(A_{ij}, A_{kl,ij}) + \rho(A_{ij}, A_{kl,ij}) \right] \\ &= \mathbb{A} \log \mathbb{A} - \mathbb{A} + \mathbb{A} \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} \beta_{kl,ij} \log(\beta_{kl,ij}) - \mathbb{A} \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} \log(2A_{ij}A_{kl}) \beta_{ij,kl} \\ (66) \quad &\geq \mathbb{A} \log \mathbb{A} - \mathbb{A} - \mathbb{A} \log \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} 2A_{ij}A_{jk}, \end{aligned}$$

where the last inequality follows from the duality formula between the log-Laplace transform and the entropy. For  $z \in \widehat{\mathcal{Z}}$ , let

$$J_0^{\leftarrow}(z) := \{i \in J_0(z) \mid \exists j \in J_1(z), (i, j) \in I^{\leftarrow}\}, \quad J_1^{\leftarrow}(z) := \{j \in J_1(z) \mid \exists i \in J_0(z), (i, j) \in I^{\leftarrow}\},$$

and let us define identically  $J_0^\rightarrow(z)$  and  $J_1^\rightarrow(z)$  by replacing the set  $I^\leftarrow$  by the  $I^\rightarrow$ . If  $i \in J_0^\leftarrow(z) \cap J_0^\rightarrow(z)$  then there exists  $j$  and  $l$  in  $J_1(z)$  such that  $(z, \sigma_{ij}(z))$  and  $(\sigma_{il}(z), z)$  are points of  $C_{\rightarrow}$ . According to Lemma 4.3  $i$ ), this is impossible since  $d(\sigma_{ij}(z), \sigma_{il}(z)) \leq 1$ . It follows that  $J_0^\leftarrow(z) \cap J_0^\rightarrow(z) = \emptyset$  and identically one proves that  $J_1^\leftarrow(z) \cap J_1^\rightarrow(z) = \emptyset$ . Let  $A := \sum_{(i,j) \in I^\leftarrow} A_{ij}$ . Since  $\mathbb{I}_1^\leftarrow \subset I^\leftarrow$ , one checks that

$$\begin{aligned} \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^\leftarrow, i \neq k, j \neq l} 2A_{ij}A_{jk} &\leq \sum_{((i,j),(k,l)) \in I^\leftarrow \times I^\leftarrow, i \neq k, j \neq l} A_{ij}A_{kl} \\ &= A^2 + \sum_{(i,j) \in I^\leftarrow} A_{ij}^2 - \sum_{i \in J_0^\leftarrow} \left( \sum_{j \in J_1^\leftarrow} A_{ij} \right)^2 - \sum_{j \in J_1^\leftarrow} \left( \sum_{i \in J_0^\leftarrow} A_{ij} \right)^2. \end{aligned}$$

Therefore, setting

$$\widetilde{A}^2 := \sum_{i \in J_0^\leftarrow} \left( \sum_{j \in J_1^\leftarrow} A_{ij} \right)^2 + \sum_{j \in J_1^\leftarrow} \left( \sum_{i \in J_0^\leftarrow} A_{ij} \right)^2 - \sum_{(i,j) \in I^\leftarrow} A_{ij}^2,$$

(65) and (66) imply

$$(67) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) \geq \int \left[ A^2 - \mathbb{A} + \mathbb{A} \left( \log \mathbb{A} - \log (A^2 - \widetilde{A}^2) \right) \right] d\widehat{Q}_t.$$

Identically, the lower-bound of  $\liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t)$  given by Theorem 3.5 provides

$$(68) \quad \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \left[ B^2 - \mathbb{B} + \mathbb{B} \left( \log \mathbb{B} - \log (B^2 - \widetilde{B}^2) \right) \right] d\widehat{Q}_t,$$

where we set for any  $z \in \widehat{Z}$ ,  $B_t(z) := \sum_{(i,j) \in I^\rightarrow} B_t(z, \sigma_{ij}(z))$ ,  $\mathbb{B}(z) := \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^\rightarrow, i \neq k, j \neq l} 2\mathbb{B}_t(z, \sigma_{kl}\sigma_{ij}(z))$  and

$$\widetilde{B}^2(z) := \sum_{i \in J_0^\rightarrow} \left( \sum_{j \in J_1^\rightarrow} B_t(z, \sigma_{ij}(z)) \right)^2 + \sum_{j \in J_1^\rightarrow} \left( \sum_{i \in J_0^\rightarrow} B_t(z, \sigma_{ij}(z)) \right)^2 - \sum_{(i,j) \in I^\rightarrow} B_t(z, \sigma_{ij}(z))^2.$$

**Step 2 :** By the Cauchy-Schwarz inequality, one has

$$\widetilde{A}^2 \geq \max \left[ \sum_{i \in J_0^\leftarrow} \left( \sum_{j \in J_1^\leftarrow} A_{ij} \right)^2, \sum_{j \in J_1^\leftarrow} \left( \sum_{i \in J_0^\leftarrow} A_{ij} \right)^2 \right] \geq \max \left[ \frac{1}{|J_0^\leftarrow|}, \frac{1}{|J_1^\leftarrow|} \right] A^2,$$

and therefore, (67) together with the concavity property of the logarithmic function yield

$$(69) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) \geq \int \max \left[ \frac{1}{|J_0^\leftarrow|}, \frac{1}{|J_1^\leftarrow|} \right] \max \{ A^2, \mathbb{A} \} d\widehat{Q}_t.$$

Identically (68) gives

$$(70) \quad \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \max \left[ \frac{1}{|J_0^\rightarrow|}, \frac{1}{|J_1^\rightarrow|} \right] \max \{ B^2, \mathbb{B} \} d\widehat{Q}_t.$$

Keeping the quantities involving  $A$  and  $B$  in (69) and (70), and applying Cauchy-Schwarz inequality, the identities (46) and (47) yield

$$\liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq W_1^2(\nu_0, \nu_1) \left[ \frac{1}{\int \min[|J_0^\leftarrow|, |J_1^\leftarrow|] d\widehat{Q}_t} + \frac{1}{\int \min[|J_0^\rightarrow|, |J_1^\rightarrow|] d\widehat{Q}_t} \right].$$

Since  $J_0^\leftarrow \cap J_0^\rightarrow = \emptyset$  and  $J_1^\leftarrow \cap J_1^\rightarrow = \emptyset$ , one has

$$(71) \quad \min[|J_0^\leftarrow|, |J_1^\leftarrow|] + \min[|J_0^\rightarrow|, |J_1^\rightarrow|] \leq \min[|J_0^\leftarrow| + |J_0^\rightarrow|, |J_1^\leftarrow| + |J_1^\rightarrow|] \leq \min[n - \kappa, \kappa].$$

and therefore the identity  $\min_{\alpha, \beta > 0, \alpha + \beta \leq 1} \left\{ \frac{1}{\alpha} + \frac{1}{\beta} \right\} = 4$  implies

$$\liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \frac{4}{\min[n - \kappa, \kappa]} W_1^2(\nu_0, \nu_1).$$



The first lower bound of  $C_t(\widehat{\pi})$  in Theorem 2.7 then follows applying Lemma 3.1.

Keeping only the quantities involving  $\mathbb{A}$  and  $\mathbb{B}$  in (69) and (70) gives

$$(72) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \frac{\mathbb{A}}{\min[|J_0^{\leftarrow}|, |J_1^{\leftarrow}|]} + \frac{\mathbb{B}}{\min[|J_0^{\rightarrow}|, |J_1^{\rightarrow}|]} d\widehat{Q}_t.$$

According to Lemma 3.4 one has for any  $z \in \widehat{Z}$ ,  $y \in \widehat{Y}_z$ , and for any  $((i, j), (k, l)) \in \mathbb{I}^{\leftarrow}(z)$ ,

$$2\mathbb{A}_{kl,ij}(z) := \frac{a_t(z, \sigma_{kl}\sigma_{ij}(z), y)L(z, \sigma_{kl}\sigma_{ij}(z))}{a_t(z, y)}.$$

Therefore the expression (43) of  $a_t(z, \sigma_{kl}\sigma_{ij}(z), y)$  with the identity

$$\sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} r(y, z, \sigma_{kl}\sigma_{ij}(z), w)L^2(z, \sigma_{kl}\sigma_{ij}(z)) = r(y, z, z, w),$$

give

$$\begin{aligned} \mathbb{A}(z) &= \sum_{\{(i,j),(k,l)\} \subset \mathbb{I}_1^{\leftarrow}, i \neq k, j \neq l} 2\mathbb{A}_{kl,ij}(z) \\ &= \frac{1}{a_t(z, y)} \sum_{w \in \mathcal{X}} \sum_{z, z \in [y, w]} r(y, z, z, w)d(y, w)(d(y, w) - 1)\rho_t^{d(y,w)-2}(d(z, w) - 2)\widehat{\pi}_-(w|y). \end{aligned}$$

Working identically with  $\mathbb{B}(z)$  we finally get

$$\int \frac{\mathbb{A}}{\min[|J_0^{\leftarrow}|, |J_1^{\leftarrow}|]} + \frac{\mathbb{B}}{\min[|J_0^{\rightarrow}|, |J_1^{\rightarrow}|]} d\widehat{Q}_t \geq \iint c_t(x, y) d\widehat{\pi}(x, y),$$

where

$$\begin{aligned} c_t(x, y) &= \sum_{z \in [x, y]} \frac{1}{\min[|J_0^{\leftarrow}(z)|, |J_1^{\leftarrow}(z)|]} r(x, z, z, y)d(x, y)(d(x, y) - 1)\rho_t^{d(x,y)-2}(d(x, z) - 2) \\ &\quad + \sum_{z \in [x, y]} \frac{1}{\min[|J_0^{\rightarrow}(z)|, |J_1^{\rightarrow}(z)|]} r(x, z, z, y)d(x, y)(d(x, y) - 1)\rho_t^{d(x,y)-2}(d(x, z)). \end{aligned}$$

Using the inequality (71), the end of the proof of the second lower bound of  $C_t(\widehat{\pi})$  involving  $T_{c_2}(\widehat{\pi})$  is exactly the same as in the proof Theorem 2.5. It is left to the reader.

We now turn to the proof of third lower bound on  $C_t(\widehat{\pi})$ . Using again (67) and (68) and the concavity of the logarithmic function, one gets

$$(73) \quad \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(t) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(t) \geq \int \widetilde{A}^2 d\widehat{Q}_t + \int \widetilde{B}^2 d\widehat{Q}_t.$$

According to the definition of  $\widetilde{A}^2$ , one has

$$\begin{aligned} \int \widetilde{A}^2 d\widehat{Q}_t &\geq \int \max \left[ \sum_{i \in J_0^{\leftarrow}} \left( \sum_{j \in J_1^{\leftarrow}} A_{ij} \right)^2, \sum_{j \in J_1^{\leftarrow}} \left( \sum_{i \in J_0^{\leftarrow}} A_{ij} \right)^2 \right] d\widehat{Q}_t \\ &= \int \sum_{z \in \widehat{Z}^y} \max \left[ \sum_{i \in [n]} \left( \sum_{j \in [n]} A_{ij}(z) \mathbb{1}_{(i,j) \in I^{\leftarrow}(z)} \right)^2, \sum_{j \in [n]} \left( \sum_{i \in [n]} A_{ij}(z) \mathbb{1}_{(i,j) \in I^{\leftarrow}(z)} \right)^2 \right] a_t(z, y) dv_1(y) \\ (74) \quad &\geq \int \max \left[ \sum_{z \in \widehat{Z}^y} \sum_{i \in [n]} \left( \sum_{j \in [n]} A_{ij}(z) \mathbb{1}_{(i,j) \in I^{\leftarrow}(z)} \right)^2 a_t(z, y), \sum_{z \in \widehat{Z}^y} \sum_{j \in [n]} \left( \sum_{i \in [n]} A_{ij}(z) \mathbb{1}_{(i,j) \in I^{\leftarrow}(z)} \right)^2 a_t(z, y) \right] dv_1(y) \end{aligned}$$

For any  $y \in \text{supp}(v_1)$ , and any  $i \in J_0(y)$ ,  $j \in J_1(y)$  we note

$$E_{i,0}^{\leftarrow}(y) := \{z \in \mathcal{X}_k \mid \exists l \in J_1(y), y \in \widehat{Y}_{(z, \sigma_{il}(z))}\}, \quad E_{j,1}^{\leftarrow}(y) := \{z \in \mathcal{X}_k \mid \exists k \in J_0(y), y \in \widehat{Y}_{(z, \sigma_{kj}(z))}\}.$$

Since  $(i, j) \in I^-(z)$  and  $z \in \widehat{Z}^y$  imply  $z \in E_{i,0}^-(y)$  and  $z \in E_{j,1}^-(y)$ , one has

$$\sum_{z \in \widehat{Z}^y} \sum_{i \in [n]} \left( \sum_{j \in [n]} A_{ij}(z) \mathbb{1}_{(i,j) \in I^-(z)} \right)^2 a_t(z, y) = \sum_{i \in J_0(y)} \sum_{z \in E_{i,0}^-(y)} \left( \sum_{j \in J_1(y)} A_{ij}(z) \right)^2 a_t(z, y),$$

and therefore by Cauchy-Schwarz inequality,

$$(75) \quad \sum_{z \in \widehat{Z}^y} \sum_{i \in [n]} \left( \sum_{j \in [n]} A_{ij}(z) \mathbb{1}_{(i,j) \in I^-(z)} \right)^2 a_t(z, y) \geq \sum_{i \in J_0(y)} \frac{\left( \sum_{j \in J_1(y)} \sum_{z \in E_{i,0}^-(y)} A_{ij}(z) a_t(z, y) \right)^2}{\sum_{z \in E_{i,0}^-(y)} a_t(z, y)}.$$

For  $(i, j) \in J_0(y) \times J_1(y)$ , one may compute the quantity  $\sum_{z \in E_{i,0}^-(y)} A_{ij}(z) a_t(z, y)$  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 for any  $(z, \sigma_{ij}(z)) \in [y, w]$ ,

$$r(y, z, \sigma_{ij}(z), w) = r(y, z, z, \sigma_{ij}(w)) \frac{L^{d(y, \sigma_{ij}(w))}(y, \sigma_{ij}(w))}{L^{d(y, w)}(y, w)}.$$

Therefore, using (42), one gets for any  $(i, j) \in J_0(y) \times J_1(y)$ ,

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

where the last equality holds since  $L^{d(x, y)}(x, y) = (d(x, y)!)^2$  for any  $x, y \in \mathcal{X}_k$ . Since for  $i \in J_0(y)$ ,  $\sum_{j \in J_1(y)} \mathbb{1}_{y_i=w_j=0} \mathbb{1}_{y_j=w_i=1} = d(y, w) \mathbb{1}_{w_i \neq y_i}$ , it follows that

$$(76) \quad \sum_{j \in J_1(y)} \sum_{z \in E_{i,0}^-(y)} A_{ij}(z) a_t(z, y) = \sum_{w \in \mathcal{X}_k} \mathbb{1}_{w_i \neq y_i} \widehat{\pi}_-(w|y).$$

Similar computations also provide, for any  $i \in J_0(y)$ ,

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

Moreover from the expression of  $Q_t^{w,y}(z)$  given by (35), one has for  $y_i = 0$  and  $w_i = 1$ ,

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

It follows that for any  $i \in J_0(y)$

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

As a consequence, since  $\Pi_{\leftarrow}^i(y) := \int \mathbb{1}_{y_i \neq w_i} d\bar{\pi}_{\leftarrow}(w|y)$ , (75) and (76) implies

$$\sum_{z \in \bar{Z}^y} \sum_{i \in [n]} \left( \sum_{j \in [n]} A_{ij}(z) \mathbb{1}_{(i,j) \in I^{\leftarrow}(z)} \right)^2 a_t(z, y) \geq \sum_{i \in J_0(y)} \frac{\Pi_{\leftarrow}^i(y)^2}{1 - (1-t)\Pi_{\leftarrow}^i(y)}.$$

By symmetry, the same inequality holds exchanging the role of  $i$  and  $j$ , and therefore (74) gives

$$\int \bar{A}^2 d\bar{Q}_t \geq \int \max \left[ \sum_{i \in J_0(y)} \frac{\Pi_{\leftarrow}^i(y)^2}{1 - (1-t)\Pi_{\leftarrow}^i(y)}, \sum_{j \in J_1(y)} \frac{\Pi_{\leftarrow}^j(y)^2}{1 - (1-t)\Pi_{\leftarrow}^j(y)} \right] d\nu_1(y).$$

As in the proof of Theorem 2.4, this inequality implies for any  $t \in (0, 1)$

$$\int \bar{A}^2 d\bar{Q}_t \geq \xi_{\leftarrow}''(t),$$

with

$$\xi_{\leftarrow}(t) := \frac{1}{2} \int \max \left[ \sum_{i \in J_0(y)} h((1-t)\Pi_{\leftarrow}^i(y)), \sum_{j \in J_1(y)} h((1-t)\Pi_{\leftarrow}^j(y)) \right] d\nu_1(y).$$

Identically, one proves that

$$\int \bar{B}^2 d\bar{Q}_t \geq \xi_{\rightarrow}''(t),$$

where

$$\xi_{\rightarrow}(t) := \frac{1}{2} \int \max \left[ \sum_{i \in J_0(x)} h(t\Pi_{\rightarrow}^i(x)), \sum_{j \in J_1(x)} h(t\Pi_{\rightarrow}^j(x)) \right] d\nu_0(x).$$

From (73) and the two last estimates, applying Lemma 3.1 provides the third lower bound of  $C_t(\bar{\pi})$  in Theorem 2.7.  $\square$

#### 4. APPENDIX A : BASIC LEMMAS

**Lemma 4.1.** *The transport-entropy inequality (29) implies the  $W_2$  transport-entropy inequality (30) for the standard Gaussian measure  $\gamma$ .*

*Proof.* The result follows from the transport-entropy inequality (29) for the uniform probability measure  $\mu$  on the hypercube ( $\alpha_i = 1/2$  for all  $i \in [n]$ ), and by using the central limit Theorem with the projection map

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

By density, it is sufficient to prove (30) for any probability measure  $\nu$  on  $\mathbb{R}$  with continuous density  $f$  and compact support  $K$ . Let  $\nu^n$  denotes the probability measure on  $\{0, 1\}^n$  with density  $f_n$  with respect to  $\mu$  given by

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

Applying (29) with  $\nu_0 := \mu$  and  $\nu_1 := \nu^n$ , one gets

$$(77) \quad \frac{2}{n} T_{c_2}(\mu, \nu^n) \leq H(\nu^n | \mu).$$

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

$$(78) \quad \lim_{n \rightarrow \infty} H(\nu^n | \mu) = H(\nu | \gamma),$$

and for  $k = 1$  or  $k = 2$ ,

$$(79) \quad \lim_{n \rightarrow \infty} \int |w|^k d(T_n \# \nu^n)(w) = \lim_{n \rightarrow \infty} \int |T_n(x)|^k f_n(T_n(x)) d\mu(x) = \int |w|^k d\nu(w).$$

Since  $d(x, y) \geq \frac{\sqrt{n}}{2} |T_n(x) - T_n(y)|$  and the monotonicity property of the function  $c_2 : \mathbb{R} \rightarrow \mathbb{R}^+$  on  $[2, +\infty)$  implies

$$\frac{2}{n} c_2(d(x, y)) \geq \frac{2}{n} c_2\left(\frac{\sqrt{n}}{2} |T_n(x) - T_n(y)|\right) \mathbb{1}_{\frac{\sqrt{n}}{2} |T_n(x) - T_n(y)| \geq 2},$$

and therefore

$$\frac{2}{n} T_{c_2}(\mu, \nu^n)^2 \geq \frac{1}{2} \inf_{\pi_n \in \Pi(T_n \# \mu, T_n \# \nu^n)} \iint c_n(z, w) d\pi_n(z, w),$$

where for any  $z, w \in \mathbb{R}$

$$\begin{aligned} c_n(z, w) &:= 4nc_2\left(\frac{\sqrt{n}|z-w|}{2}\right) \mathbb{1}_{|z-w| \geq 4/\sqrt{n}} \\ &= \left[ |z-w|^2 - \frac{4(1 + \log(\sqrt{n}/2))}{\sqrt{n}} |z-w| - \frac{4}{\sqrt{n}} |z-w| \log |z-w| \right] \mathbb{1}_{|z-w| \geq 4/\sqrt{n}}. \end{aligned}$$

Let  $c(z, w) := |z-w|^2$ ,  $z, w \in \mathbb{R}$ . One has, for any  $z, w \in \mathbb{R}$ ,  $c(z, w) \geq c_n(z, w)$  and

$$\begin{aligned} c(z, w) - c_n(z, w) &= |z-w|^2 \mathbb{1}_{|z-w| < 4/\sqrt{n}} + \left[ \frac{4(1 + \log(\sqrt{n}/2))}{\sqrt{n}} |z-w| + \frac{4}{\sqrt{n}} |z-w| \log |z-w| \right] \mathbb{1}_{|z-w| \geq 4/\sqrt{n}} \\ &\leq \frac{16}{n} + \frac{4(1 + \log n)}{\sqrt{n}} \left[ |z| + |w| + 1 + 2|z|^2 + 2|w|^2 \right], \end{aligned}$$

where the last inequality follows from  $|u \log u| \leq 1 + u^2$ ,  $u > 0$ . Since

$$\int |z| d(T_n \# \mu)(z) \leq \left( \int |z|^2 d(T_n \# \mu)(z) \right)^{1/2} = \left( \int T_n^2 d\mu \right)^{1/2} = 1,$$

it follows that for any  $\pi_n \in \Pi(T_n \# \mu, T_n \# \nu^n)$ ,

$$\begin{aligned} \iint c_n d\pi_n &\geq \iint c d\pi_n - \frac{16}{n} - \frac{4(1 + \log n)}{\sqrt{n}} \iint \left[ |z| + |w| + 1 + 2|z|^2 + 2|w|^2 \right] d\pi_n(z, w) \\ &\geq \iint c d\pi_n - \frac{32(1 + \log n)}{\sqrt{n}} \left[ 1 + \int |w| d(T_n \# \nu^n)(w) + \int |w|^2 d(T_n \# \nu^n)(w) \right], \end{aligned}$$

and therefore

$$\frac{2}{n} T_{c_2}(\mu, \nu^n)^2 \geq \frac{1}{2} W_2^2(T_n \# \mu, T_n \# \nu^n) - \frac{16(1 + \log n)}{\sqrt{n}} \left[ 1 + \int |w| d(T_n \# \nu^n)(w) + \int |w|^2 d(T_n \# \nu^n)(w) \right].$$

From the weak convergence in  $\mathcal{P}_2(\mathbb{R})$  of the sequences  $(T_n\#\mu)$  and  $(T_n\#v^n)$  and using (79), the last inequality implies as  $n$  goes to infinity

$$\liminf_{n \rightarrow +\infty} \frac{2}{n} T_{c_2}(\mu, v^n) \geq \frac{1}{2} W_2^2(v, \gamma).$$

Finally, Talagrand's inequality  $W_2^2(v, \gamma) \leq 2H(v|\gamma)$ , follows from (77) and (78).  $\square$

**Lemma 4.2.** *If the convexity property (4) holds, then for any  $v_0, v_1 \in \mathbb{P}_b(\mathcal{X})$ ,*

$$H(v_0|\mu) \leq H(v_1|\mu) + \sum_{x \in \mathcal{X}} \sum_{x' \in \mathcal{X}, x' \sim x} (\log(f(x) - \log f(x'))) \Pi_{\rightarrow}^{x'}(x) v_0(x) - \frac{1}{2} \liminf_{t \rightarrow 0} C_t(\widehat{\pi}).$$

where  $\Pi_{\rightarrow}^{x'}(x) := \int \mathbb{1}_{x' \in [x, y]} d(x, y) r(x, x', x', y) d\widehat{\pi}_{\rightarrow}(y|x)$ .

*Proof.* The convexity property (4) implies, for any  $v_0, v_1 \in \mathbb{P}_b(\mathcal{X})$  and for any  $t \in (0, 1)$

$$(80) \quad H(v_0|\mu) \leq H(v_1|\mu) - \frac{H(\widehat{Q}_t|\mu) - H(v_0|\mu)}{t} - \frac{(1-t)}{2} C_t(\widehat{\pi}).$$

The first step is to compute the left-hand side of this inequality as  $t$  goes to zero. According to the expression (28) of  $Q_t^{x,y}$ , for any  $x, y, z \in \{0, 1\}^n$ ,

$$\partial_t Q_t^{x,y}(z) = r(x, z, z, y) \begin{pmatrix} d(x, y) \\ d(x, z) \end{pmatrix} \mathbb{1}_{[x, y]}(z) \left( d(x, z) t^{d(x, z)-1} (1-t)^{d(z, y)} - d(z, y) t^{d(x, z)} (1-t)^{d(z, y)-1} \right),$$

and therefore

$$\begin{aligned} \partial_t Q_t^{x,y}(z)|_{t=0} &= r(x, z, z, y) \begin{pmatrix} d(x, y) \\ d(x, z) \end{pmatrix} \left( \mathbb{1}_{[x, y]}(z) \mathbb{1}_{z \sim x} - d(x, y) \mathbb{1}_{x=z} \right) \\ &= \sum_{x' \in [x, y], x' \sim x} d(x, y) r(x, x', x', y) (\delta_{x'}(z) - \delta_x(z)). \end{aligned}$$

Since  $\partial_t \widehat{Q}_t^y(z)|_{t=0} = \sum_{x, y \in \mathcal{X}} \partial_t Q_t^{x,y}(z)|_{t=0} \widehat{\pi}(x, y)$ , it follows that

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{H(\widehat{Q}_t|\mu) - H(v_0|\mu)}{t} &= \partial_t H(\widehat{Q}_t|\mu)|_{t=0} = \sum_{z \in \mathcal{X}} \partial_t \widehat{Q}_t(z)|_{t=0} \log f(z) \mu(z) \\ &= \sum_{x, y \in \mathcal{X}} \sum_{x' \in [x, y], x' \sim x} d(x, y) (\log f(x') - \log f(x)) d(x, y) r(x, x', x', y) \widehat{\pi}(x, y) \\ &= \sum_{x \in \mathcal{X}} \sum_{x' \in \mathcal{X}, x' \sim x} (\log(f(x')) - \log f(x)) \left( \sum_{y \in \mathcal{X}, x' \in [x, y]} d(x, y) r(x, x', x', y) \widehat{\pi}_{\rightarrow}(y|x) \right) v_0(x) \end{aligned}$$

The proof of Lemma 4.2 ends from (80) as  $t$  goes to 0.  $\square$

**Lemma 4.3.** *Let  $\mathcal{X}$  be a graph with graph distance  $d$ . Let  $v_0, v_1 \in \mathcal{P}(\mathcal{X})$  and assume that  $\widehat{\pi} \in \mathcal{P}(\mathcal{X} \times \mathcal{X})$  is a  $W_1$ -optimal coupling of  $v_0$  and  $v_1$ , namely*

$$W_1(v_0, v_1) = \iint d(x, y) d\widehat{\pi}(x, y).$$

(i) Let

$$C_{\rightarrow} := \{(z, w) \in \mathcal{X} \times \mathcal{X} \mid z \neq w, \exists (x, y) \in \text{supp}(\widehat{\pi}), (z, w) \in [x, y]\}.$$

If  $(z_1, w) \in C_{\rightarrow}$  and  $(w, z_2) \in C_{\rightarrow}$  then  $d(z_1, z_2) \geq 2$  and  $w \in [z_1, z_2]$ .

(ii) Let

$$C_{\leftarrow} := \{(z, w) \in \mathcal{X} \times \mathcal{X} \mid (w, z) \in C_{\rightarrow}\}.$$

The sets  $C_{\rightarrow}$  and  $C_{\leftarrow}$  are disjoint.

(iii) If  $d$  is the Hamming distance then the following sets  $D_{\rightarrow}$  and  $D_{\leftarrow}$  are disjoint,

$$D_{\leftarrow} := \left\{ w \in \text{supp}(\nu_1) \mid \exists x \in \mathcal{X}, w \neq x, (x, w) \in \text{supp}(\widehat{\pi}) \right\},$$

and

$$D_{\rightarrow} := \left\{ w \in \text{supp}(\nu_0) \mid \exists y \in \mathcal{X}, w \neq y, (w, y) \in \text{supp}(\widehat{\pi}) \right\}.$$

*Proof.* (i) Let  $(z_1, w) \in C_{\rightarrow}$  and  $(w, z_2) \in C_{\leftarrow}$ . There exists  $(x, y) \in \text{supp}(\widehat{\pi})$  such that  $(z_1, w) \in [x, y]$  and there exists  $(x', y') \in \text{supp}(\widehat{\pi})$  such that  $(w, z_2) \in [x', y']$ . One has

$$d(z_1, w) + d(w, z_2) = ((d(x, y) - d(x, z_1) - d(w, y)) + (d(x', y') - d(x', w) - d(z_2, y'))).$$

It is well known that the support of any optimizer of  $W_1(\nu_0, \nu_1)$  is  $d$ -cyclically monotone (see [43, Theorem 5.10]). By definition, it means that for any family  $(x_1, y_1), \dots, (x_N, y_N)$  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  $y_{N+1} = y_1$ . It follows that

$$d(x, y) + d(x', y') \leq d(x, y') + d(x', y),$$

and therefore, from the above identity,

$$d(z_1, w) + d(w, z_2) \leq d(x, y') + d(x', y) - d(x, z_1) - d(w, y) - d(x', w) - d(z_2, y').$$

By the triangular inequality, it follows that

$$\begin{aligned} 2 \leq d(z_1, w) + d(w, z_2) &\leq (d(x, z_1) + d(z_1, z_2) + d(z_2, y')) \\ &\quad + (d(x', w) + d(w, y)) - d(x, z_1) - d(w, y) - d(x', w) - d(z_2, y') = d(z_1, z_2). \end{aligned}$$

This implies that  $d(z_1, z_2) \geq 2$  and  $w \in [z_1, z_2]$ .

- (ii) Assume there exists  $(z, w) \in C_{\rightarrow} \cap C_{\leftarrow}$ . Then  $(w, z) \in C_{\rightarrow}$  and therefore, according to (i),  $z \in [w, w] = \{w\}$ . This is impossible since  $z \neq w$ .
- (iii) We assume that  $d(x, y) = \mathbb{1}_{x \neq y}$  for any  $x, y \in \mathcal{X}$ . If the two sets  $D_{\rightarrow}$  and  $D_{\leftarrow}$  intersect, then there exists  $(x, w) \in C_{\rightarrow}$  and  $(w, y) \in C_{\leftarrow}$ . Point (i) implies  $w \in [x, y]$ , and since  $d(x, y) = 1$ , we get either  $w = x$  or  $w = y$ , which is impossible.

□

**Lemma 4.4.** Let  $\nu_0$  and  $\nu_1$  some probability measures in  $\mathcal{P}(\mathcal{X})$  with bounded support.

(i) If (13) holds ( $\exists S \geq 1, \sup_{x \in \mathcal{X}} |L(x, x)| \leq S$ ), then for any  $x, y \in \mathcal{X}$  and any integer  $k$ ,

$$L^k(x, y) \leq (2S)^k.$$

(ii) If (14) holds ( $\exists I \in (0, 1], \inf_{x, y \in \mathcal{X}, x \sim y} L(x, y) \geq I$ ), then for any  $x, y \in \mathcal{X}$ ,  $L^{d(x, y)}(x, y) \geq I^{d(x, y)}$ .

(iii) If (13) and (14) 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  $x, y, t$  and  $\gamma$ .

(iv) If (13) holds then for any  $x, y, z \in \mathcal{X}$  and for any  $t \in [0, 1]$

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

(v) If (13) holds then for any  $x, y \in \mathcal{X}$ ,

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

For a fixed  $x_0 \in \mathcal{X}$ , let  $D := \max_{x \in \text{supp}(v_0), y \in \text{supp}(v_1)} (d(x_0, x), d(x_0, y))$ . It follows that if (13) and (14) hold then for any  $\gamma \in (0, 1)$  and  $t \in (0, 1)$ ,

$$(81) \quad 0 < e^{-S} \left( \frac{t\gamma I}{d(x_0, z) + 1 + D} \right)^{d(x_0, z) + 1 + D} \min_{w \in \text{supp}(v_0)} f^\gamma(w) \leq P_t^\gamma f^\gamma(z) \leq \max_{w \in \text{supp}(v_0)} f^\gamma(w).$$

(vi) If (13) holds then  $\mathbb{E}_{P^\gamma}[\ell | X_0 = x, X_1 = y] \leq \frac{\gamma S}{P_1^\gamma(x, y)}$ .

(vii) Assume (13) and (14) hold. For a fixed  $x_0 \in \mathcal{X}$ , let  $D := \max_{x \in \text{supp}(v_0), y \in \text{supp}(v_1)} (d(x_0, x), d(x_0, y))$ . For any  $x \in \text{supp}(v_0)$  and  $y \in \text{supp}(v_1)$ , one has for any  $t \in (0, 1)$  and any  $\gamma \in (0, 1)$

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

where  $K := 2S/I$  and  $O(1)$  denotes a constant that only depends on  $S, I, D$  and  $K := 2S/I$ .

As a consequence, setting

$$B := \bigcup_{x \in \text{supp}(v_0), y \in \text{supp}(v_1)} [x, y],$$

one has

$$(82) \quad \widehat{Q}_t^\gamma(z) \leq O(1) \gamma (\gamma K^2)^{[2d(x_0, z) - 4D - 1]_+}, \quad \forall z \in \mathcal{X} \setminus B.$$

(viii) Assume (13) and (14) hold. Let  $x_0 \in \mathcal{X}$ ,  $t \in (0, 1)$  and  $\gamma \in (0, 1)$ . For any  $w, z, z' \in \mathcal{X}$  with  $d(z, z') \leq 2$  and  $w \in \text{supp}(v_0)$  one has

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

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

$$(83) \quad \frac{(\gamma t)^{d(z, z')}}{\max(1, d(x_0, z)^{d(z, z')}) K^{d(x_0, z)} O(1)} \leq \frac{P_t^\gamma f^\gamma(z')}{P_t^\gamma f^\gamma(z)} \leq \frac{\max(1, d(x_0, z)^{d(z, z')}) K^{d(x_0, z)} O(1)}{(\gamma t)^{d(z, z')}}.$$

(ix) Let  $(\gamma_\ell)_{\ell \in \mathbb{N}}$  be a sequence of positive numbers converging to zero. If (12), (13), (14) and (15) hold, then for any  $t \in [0, 1]$

$$\lim_{\gamma_\ell \rightarrow 0} H(\widehat{Q}_t^{\gamma_\ell} | m) = H(\widehat{Q}_t^0 | m).$$

*Proof.* (i) Given (13), 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  $\alpha$  from  $x$  to  $y$  of length  $d(x, y)$ ,  $\alpha = (z_0, \dots, z_{d(x,y)})$  with  $z_0 = x$  and  $z_{d(x,y)} = y$ , and

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

Such a path  $\alpha$  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 (13), for such a path  $\alpha$ ,  $L_\alpha \geq I^{d(x,y)}$ . As a consequence we get  $L^{d(x,y)}(x, y) \geq I^{d(x,y)}$ .

(iii) According to (16), for any  $x, y \in \mathcal{X}$ ,

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

Applying Lemma 4.4 (i) and (ii), we get

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

from which the expected result follows.

(iv) Let  $x, y, z \in \mathcal{X}$  and  $t \in [0, 1]$ . If (13) holds, according to (16), the Taylor expansion of  $P_t^\gamma(x, y)$  as  $\gamma$  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  $Q_t^{\gamma,x,y}(z)$ , defined by (9), is

$$\begin{aligned} Q_t^{\gamma,x,y}(z) &= \gamma^{d(x,z)+d(z,y)-d(x,y)} \frac{L^{d(x,z)}(x, z) L^{d(z,y)}(z, y)}{L^{d(x,y)}(x, y)} \frac{d(x, y)!}{d(x, z)! d(z, y)!} t^{d(x,z)} (1-t)^{d(z,y)} \\ & \quad + o(\gamma^{d(x,z)+d(z,y)-d(x,y)}). \end{aligned}$$

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

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

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

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  $R^\gamma$  given  $X_0 = x$  is the same as the law of the process  $(\tilde{X}_t)_{t \geq 0}$  under  $\mathbb{P}$  given  $\tilde{X}_0 = x$  defined by  $\tilde{X}_t := Y_{N_t}$ . As a consequence, one has for any  $y \in \mathcal{X}$ ,

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



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

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

This ends the proof of the first part of (v). Observe that from the Schrödinger system (7),  $f^\gamma(w) > 0$  if and only if  $w \in \text{supp}(v_0)$ . Since  $v_0$  has bounded support, it follows that for any  $w \in \text{supp}(v_0)$ ,

$$0 < \min_{u \in \text{supp}(v_0)} f^\gamma(u) \leq f^\gamma(w) \leq \max_{u \in \text{supp}(v_0)} f(u),$$

and therefore for any  $z \in \mathcal{X}$ ,

$$\min_{u \in \text{supp}(v_0)} f(u) \min_{w \in \text{supp}(v_0)} P_t^\gamma(z, w) \leq \sum_{w \in \text{supp}(v_0)} f^\gamma(w) P_t^\gamma(z, w) = P_t^\gamma f^\gamma(z) \leq \max_{u \in \text{supp}(v_0)} f(u).$$

From (14) and (ii) and since  $d(z, w) \leq d(z, x_0) + 1 + D$  for any  $w \in \text{supp}(v_0)$ , one gets

$$\min_{w \in \text{supp}(v_0)} P_t^\gamma(z, w) \geq e^{-S} \left( \frac{t\gamma I}{d(x_0, z) + 1 + D} \right)^{d(x_0, z) + 1 + D},$$

from which the second part of (v) follows.

(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,

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

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

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

$$\begin{aligned}
Q_t^{\gamma, x, y}(z) &= \frac{P_t^\gamma(x, z) P_{1-t}^\gamma(z, y)}{P_1^\gamma(x, y)} \\
&\leq \gamma^{d(x, z) + d(z, y) - d(x, y)} r(x, z, z, y) \frac{d(x, y)!}{d(x, z)! d(z, y)!} t^{d(x, z)} (1-t)^{d(z, y)} e^{\gamma S} \\
(84) \quad &\quad \quad \quad (1 + \gamma K^{d(x, z)} O(1)) (1 + \gamma K^{d(z, y)} O(1)).
\end{aligned}$$

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

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

and the maximum of this quantity over all  $x \in \text{supp}(v_0)$  and  $y \in \text{supp}(v_1)$  is a constant  $O(1)$ , independent of  $x, z, y$  and  $\gamma$ .

If  $z \notin [x, y]$ , then  $d(x, z) + d(z, y) - d(x, y) \geq \max\{1, 2d(x_0, z) - 4D\}$ , and the right-hand side of (84) is bounded by

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

The maximum over all  $x \in \text{supp}(v_0)$  and  $y \in \text{supp}(v_1)$  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

$$\widehat{Q}_t^\gamma(z) = \sum_{x \in \text{supp}(v_0), y \in \text{supp}(v_1)} Q_t^{\gamma x, y}(z) \widehat{\pi}^\gamma(x, y).$$

(viii) Using (iii) and (v), one gets for any  $z, z' \in \mathcal{X}$  and any  $w \in \text{supp}(v_0)$ ,

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

where one maximizes over all  $w \in \text{supp}(v_0)$  to get the last inequality. Inequality (83) follows since

$$\frac{P_t^\gamma f^\gamma(z')}{P_t^\gamma f^\gamma(z)} = \sum_{w \in \text{supp}(v_0)} \frac{P_t^\gamma(z', w)}{P_t^\gamma(z, w)} \frac{f^\gamma(w) P_t^\gamma(z, w)}{P_t^\gamma f^\gamma(z)},$$

$$\text{with } \sum_{w \in \text{supp}(v_0)} \frac{f^\gamma(w) P_t^\gamma(z, w)}{P_t^\gamma f^\gamma(z)} = 1.$$

(ix) Recall that

$$H(\widehat{Q}_t^{\gamma_\ell} | m) = \sum_{z \in \mathcal{X}} \log \frac{\widehat{Q}_t^{\gamma_\ell}(z)}{m(z)} \widehat{Q}_t^{\gamma_\ell}(z).$$

Let us consider the finite set  $B$  defined in Lemma 4.4 (vii). From the weak convergence of the sequence  $(\widehat{Q}_t^{\gamma_\ell})$  to  $\widehat{Q}_t^0$  and since  $\text{supp}(\widehat{Q}_t^0) \subset B$ , one has

$$\lim_{\gamma_\ell \rightarrow 0} \sum_{z \in B} \log \frac{\widehat{Q}_t^{\gamma_\ell}(z)}{m(z)} \widehat{Q}_t^{\gamma_\ell}(z) = H(\widehat{Q}_t^0 | m).$$

Therefore it remains to prove that

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

From Lemma 4.4 (vii) and hypothesis (12) one has, for any  $z \in \mathcal{X} \setminus B$ ,

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

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

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

Hypothesis (15) then implies that there exists  $\tilde{\gamma} > 0$  such that for any  $0 < \gamma_\ell < \tilde{\gamma}$

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

and the expected result follows. □

## 5. APPENDIX B : PROOFS OF LEMMAS 3.1, 3.2, 3.3, AND 3.4

*Proof of Lemma 3.2 and Lemma 3.3.* Let  $\gamma$  denotes a fixed parameter of temperature that can be chosen as small as we want. To simplify the notations, the dependence in the temperature parameter  $\gamma$  is sometimes omitted. For  $t \in (0, 1)$ , let us note  $f_t := P_t^\gamma f^\gamma$  and  $g_t := P_{1-t}^\gamma g^\gamma$  and recall that  $F_t := \log f_t$ ,  $G_t := \log g_t$  and

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

Observe that for  $\gamma$  sufficiently small, these two functions are well defined on  $(0, 1)$  since (81) and (82) implies

$$\begin{aligned} \int |F_t| f_t g_t dm &= \sum_{z \in \mathcal{X}} |\log(P_t^\gamma f^\gamma(z))| \widehat{Q}_t^\gamma(z) \\ &\leq O(1) + O(1) \sum_{z \in \mathcal{X} \setminus B} (d(x_0, z) + 1 + D) \left( \log \frac{1}{t\gamma I} + \log(d(x_0, z) + 1 + D) \right) \gamma (\gamma K^2)^{[2d(x_0, z) - 4D - 1]_+}. \end{aligned}$$

According to hypothesis (15), the right-hand side of this inequality is finite if  $(\gamma K^2)^2 < \gamma_o$ . Identically, one could check that  $\int |G_t| f_t g_t dm$  is finite for  $\gamma$  sufficiently small.

The proof is based on  $\Gamma_2$ -calculus by using backward equations,  $\partial_t f_t = L f_t$ ,  $\partial_t g_t = -L g_t$ . 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  $(t, z) \in (0, 1) \times \mathcal{X} \rightarrow V_t(z) \in \mathbb{R}$  denotes some differentiable function in  $t$  (that also depends of the parameter  $\gamma$ ) satisfying for any  $\varepsilon \in (0, 1/2)$ , and any  $x_0 \in \mathcal{X}$ ,

$$(85) \quad \sup_{t \in (\varepsilon, 1-\varepsilon)} |V_t(z)| \leq O(1) \frac{A^{d(x_0, z)}}{\gamma^{10}},$$

and

$$(86) \quad \sup_{t \in (\varepsilon, 1-\varepsilon)} |\partial_t V_t(z)| \leq O(1) \frac{B^{d(x_0, z)}}{\gamma^{10}},$$

for all  $z \in \mathcal{X}$  where  $O(1), A, B$  denote constants that do not depend on  $t, \gamma$  and  $z$ . Then the following identity holds: for any  $t \in (0, 1)$ ,

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

It suffices to justify this identity for any  $\varepsilon \in (0, 1/2)$  and any  $t \in (\varepsilon, 1 - \varepsilon)$ . The second equality of (87) is due to the backward equations. The first equality of (87) is justified by applying Lebesgue's theorem with hypothesis (15), provided that for  $\gamma$  sufficiently small, one has

$$\sup_{t \in (\varepsilon, 1 - \varepsilon)} |\partial_t (V_t f_t g_t)(z) m(z)| \leq O(1) \gamma_o^{d(x_0, z)}.$$

This is indeed the case, since for any  $z \in \mathcal{X}$ ,

$$\partial_t (V_t f_t g_t)(z) m(z) = \left[ (\partial_t V_t)(z) + V_t(z) \frac{LP_t^\gamma f^\gamma(z)}{P_t^\gamma f^\gamma(z)} - V_t(z) \frac{LP_{1-t}^\gamma g^\gamma(z)}{P_{1-t}^\gamma g^\gamma(z)} \right] \widehat{Q}_t^\gamma(z),$$

with according to (83), for any  $t \in (\varepsilon, 1)$ ,

$$\left| \frac{LP_t^\gamma f^\gamma(z)}{P_t^\gamma f^\gamma(z)} \right| \leq S d_{\max} \left( 1 + \max_{z', z' \sim z} \left| \frac{P_t^\gamma f^\gamma(z')}{P_t^\gamma f^\gamma(z)} \right| \right) \leq S d_{\max} \frac{\max(1, d(x_0, z)) K^{d(x_0, z)} O(1)}{\gamma \varepsilon} \leq O(1) \frac{K^{d(x_0, z)}}{\gamma}.$$

One identically shows that  $\left| \frac{LP_{1-t}^\gamma g^\gamma(z)}{P_{1-t}^\gamma g^\gamma(z)} \right| \leq O(1) \frac{K^{d(x_0, z)}}{\gamma}$ , for any  $t \in (0, 1 - \varepsilon)$  and  $z \in \mathcal{X}$ . Together with (82), we get the bound, for any  $z \in \mathcal{X}$  and  $t \in (\varepsilon, 1 - \varepsilon)$ ,

$$|\partial_t (V_t f_t g_t)(z) m(z)| \leq O(1) \left( B^{d(x_0, z)} + (AK)^{d(x_0, z)} \right) \frac{(\gamma K^2)^{2d(x_0, z)}}{\gamma^{11}} \leq O(1) \gamma_o^{d(x_0, z)},$$

for any  $\gamma > 0$  with  $\gamma^2 (B + AK) K^4 \leq \gamma_o$ . The third equality of (87) is due to Fubini's theorem together with the reversibility property of  $m$  with respect to  $L$ . The last equality of (87) is a simple rearrangement of the terms.

At first, one applies (87) with  $V_t = F_t$ , since according to (81), for any  $t \in (\varepsilon, 1 - \varepsilon)$ , for any  $z \in \mathcal{X}$ ,

$$|F_t(z)| \leq O(1) (d(x_0, z) + 1 + D) \left( \log \frac{1}{\varepsilon \gamma I} + \log (d(x_0, z) + 1 + D) \right) \leq O(1) \frac{2^{d(x_0, z)}}{\gamma},$$

and

$$|\partial_t F_t(z)| = \left| \frac{LP_t^\gamma f^\gamma(z)}{P_t^\gamma f^\gamma(z)} \right| \leq O(1) \frac{K^{d(x_0, z)}}{\gamma}.$$

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

one gets the expected result

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

We want now to apply again (87) with  $V_t(z) = \sum_{z', z' \sim z} \zeta \left( e^{\nabla F_t(z, z')} \right) L(z, z')$ ,  $z \in \mathcal{X}$ . From the inequality,  $|\zeta(a)| \leq 2 + a^2$ ,  $a > 0$  and using (83), one may check as above that (85) holds. The backward equations ensure that

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

Simple computations together with (83) show that (86) holds too.

Applying the identity (87), since

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

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

$$\begin{aligned}\varphi''(t) &= - \int \left[ \sum_{z', w', z' \sim z, w' \sim z} \left[ \zeta \left( e^{\nabla F_t(z, w')} \right) - \nabla F_t(z, z') \left( e^{\nabla F_t(z, w')} - 1 \right) \right] e^{\nabla F_t(z, z')} L(z, z') L(z, w') \right. \\ &\quad \left. + \sum_{z', z'', z \sim z' \sim z''} \left[ \nabla F_t(z, z') \left( e^{\nabla F_t(z', z'')} - 1 \right) - \zeta \left( e^{\nabla F_t(z', z'')} \right) \right] e^{\nabla F_t(z, z')} L(z, z') L(z', z'') \right] d\widehat{Q}_t^Y(z) \\ &= - \int \left[ \sum_{z', w', z' \sim z, w' \sim z} \left( \nabla F_t(z, w') - \nabla F_t(z, z') \right) e^{\nabla F_t(z, w') + \nabla F_t(z, z')} L(z, z') L(z, w') \right. \\ &\quad + \sum_{z', w', z' \sim z, w' \sim z} \left( \nabla F_t(z, z') + 1 \right) e^{\nabla F_t(z, z')} L(z, z') L(z, w') \\ &\quad - \sum_{z', z'', z \sim z' \sim z''} \left( \nabla F_t(z, z') + 1 \right) e^{\nabla F_t(z, z')} L(z, z') L(z', z'') \\ &\quad \left. - \sum_{z', z'', z \sim z' \sim z''} \rho \left( e^{\nabla F_t(z, z')}, e^{\nabla F_t(z, z'')} \right) L(z, z') L(z', z'') \right] d\widehat{Q}_t^Y(z),\end{aligned}$$

where the last equality holds since  $\nabla F_t(z, z') + \nabla F_t(z', z'') = \nabla F_t(z, z'')$ . The expected expression of  $\varphi''(t)$  follows by symmetrization of the first sum in  $z'$  and  $w'$ , and since  $\sum_{w', w' \sim z} L(z, w') = -L(z, z)$ .  $\square$

*Proof of Lemma 3.1.* Let  $\varepsilon \in (0, 1/2)$ . We first prove that if (13), (14) and (15) hold then  $\varphi''_t(t)$  is uniformly lower bounded over all  $t \in [\varepsilon, 1]$  and  $\gamma \in (0, \bar{\gamma}]$  for some  $\bar{\gamma} \in (0, 1)$ . According to (48) and

inequality (50) and (51), for any  $t \in [\varepsilon, 1]$  and  $\gamma > 0$ ,

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

where  $O(1)$  denotes a positive constant that only depends on  $\bar{\gamma}$  and  $\varepsilon$ . Using Lemma 4.4 (vii) and the fact that  $\nu_0$  and  $\nu_1$  have bounded support, it follows that

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

From hypothesis (15), choosing  $\bar{\gamma}$  so that  $(\bar{\gamma} K^3)^2 < \gamma_0$ , one gets

$$\inf_{\gamma \in (0, \bar{\gamma}), t \in [\varepsilon, 1]} \varphi''_\gamma(t) \geq -O(1).$$

One may similarly prove by symmetry that if (13), (14) and (15) hold, then  $-\psi''_\gamma(t)$  is also uniformly lower bounded, namely

$$\inf_{\gamma \in (0, \bar{\gamma}), t \in [0, 1 - \varepsilon]} \psi''_\gamma(t) \geq -O(1).$$

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

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

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

$$(88) \quad (1-t)F_{\gamma_\ell}^\varepsilon(0) + tF_{\gamma_\ell}^\varepsilon(1) - F_{\gamma_\ell}^\varepsilon(t) = \frac{t(1-t)}{2} \int_0^1 K_t(s) (F_{\gamma_\ell}^\varepsilon)''(s) ds,$$

where the kernel  $K_t$  is defined by (38). Observe that

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

The above uniform bounds on  $\varphi''_\gamma$  and  $\psi''_\gamma$  for  $\gamma \in (0, \bar{\gamma})$  allow to apply Fatou's Lemma. Together with Lemma 4.4 (ix) it implies, for any  $\varepsilon \in (0, 1/2)$

$$(89) \quad (1-t)F_0^\varepsilon(0) + tF_0^\varepsilon(1) - F_0^\varepsilon(t) \geq \frac{t(1-t)}{2} (1-2\varepsilon) \int_\varepsilon^{1-\varepsilon} K_t\left(\frac{u-\varepsilon}{1-2\varepsilon}\right) \liminf_{\gamma_\ell \rightarrow 0} (\varphi''_{\gamma_\ell}(u) + \psi''_{\gamma_\ell}(u)) du.$$

For any  $t \in [0, 1]$  the support of the measure  $\widehat{Q}_t$  is finite, included in the set  $B$  defined Lemma 4.4 (vii). As a consequence, the function  $t \in [0, 1] \rightarrow H(\widehat{Q}_t | m)$  is continuous as a finite sum of continuous functions. It follows that for any  $t \in [0, 1]$ ,

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

Consequently, using hypothesis (39) and applying Fatou's Lemma as  $\varepsilon$  goes to zero, equality (89) provides

$$\begin{aligned} (1-t)H(\nu_0|m) + tH(\nu_1|m) - H(\widehat{Q}_t|m) &\geq \frac{t(1-t)}{2} \int_0^1 K_t(u) \left( \liminf_{\gamma_\ell \rightarrow 0} \varphi''_{\gamma_\ell}(u) + \liminf_{\gamma_\ell \rightarrow 0} \psi''_{\gamma_\ell}(u) \right) du \\ &\geq \frac{t(1-t)}{2} \int_0^1 K_t(u) \xi''(u) du \\ &= (1-t)\xi(0) + t\xi(1) - \xi(t) \end{aligned}$$

were the last equality is a consequence of identity (88) applied with  $\xi$ .  $\square$

*Proof of Lemma 3.4.* Let  $z \in \widehat{Z}$  and  $z' \in V(z)$ . One will only compute the expression of  $\lim_{\gamma_\ell \rightarrow 0} (\gamma_\ell A_t^{\gamma_\ell}(z, z'))$  and similar calculations provide  $\lim_{\gamma_\ell \rightarrow 0} (\gamma_\ell B_t^{\gamma_\ell}(z, z'))$ . For any  $\gamma > 0$ , let

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

and

$$a_t^\gamma(z, z', y) := \int \alpha_t^\gamma(y, z, z', w) d\widehat{\pi}_-^\gamma(w|y), \quad \text{with} \quad \alpha_t^\gamma(y, z, z', w) = \frac{P_{1-t}^\gamma(y, z) P_t^\gamma(z', w)}{P_1^\gamma(y, w)}.$$

Using equality (11) and since  $P_1^\gamma f^\gamma(y) > 0$  for any  $\gamma > 0$ , one easily check that for any  $\gamma > 0$ ,

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

From the expression (41) of  $a_t(z, y)$  and since  $\text{supp}(\widehat{\pi}_-^{\gamma_\ell}(\cdot|y)) \subset \text{supp}(\nu_0)$ , one has

$$\left| a_t^{\gamma_\ell}(z, y) - a_t(z, y) \right| \leq \sup_{w \in \text{supp}(\nu_0)} \left| Q_t^{\gamma_\ell w, y}(z) - Q_t^{w, y}(z) \right| + \sum_{w \in \text{supp}(\nu_0)} \left| \widehat{\pi}_-^{\gamma_\ell}(w|y) - \widehat{\pi}_-(w|y) \right|.$$

Therefore, the weak convergence of  $(\widehat{\pi}_-^{\gamma_\ell})_{k \in \mathbb{N}}$  to  $\widehat{\pi}$  and Lemma 4.4 (iv) imply

$$(90) \quad \lim_{\gamma_\ell \rightarrow 0} a_t^{\gamma_\ell}(z, y) = a_t(z, y).$$

Let us now consider the behaviour of  $\gamma_\ell a_t^{\gamma_\ell}(z, z', y)$  as  $\gamma_\ell$  goes to zero. Lemma 4.4 (iii) provides the following Taylor expansion,

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

where  $O(1)$  is a quantity uniformly bounded in  $t, \gamma, z, z', x, y$ . 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

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

with

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

Moreover, Lemma 4.4 (i), (ii) and (iii) ensures that for any  $w \in \text{supp}(v_0)$  and  $y \in \text{supp}(v_1)$ ,

$$\begin{aligned} \gamma \alpha_t^\gamma(y, z, z', w) &\leq O(1) \gamma^{d(y,z)+1+d(z',w)-d(y,w)} (2S)^{d(y,z)+d(z',w)-d(y,w)} K^{d(y,z)+d(z',w)} \\ &\quad \cdot \max_{w \in \text{supp}(v_0), y \in \text{supp}(v_1)} \frac{(2S)^{d(y,w)} d(y, w)! K^{d(y,w)}}{I^{d(y,w)}} \\ &\leq O(1) (\gamma 2S K)^{d(y,z)+d(z',w)+1-d(y,w)}, \end{aligned}$$

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  $\gamma < 1/(2SK)$ . As a consequence, for any  $\gamma_\ell < 1/(2SK)$ , it holds

$$\begin{aligned} & \left| \gamma_\ell \mathfrak{a}_t^{\gamma_\ell}(z, z', y) - \mathfrak{a}_t(z, z', y) \right| \\ & \leq \sup_{w \in \text{supp}(v_0)} \left| \gamma_\ell \alpha_t^{\gamma_\ell}(y, z, z', w) - \alpha_t(y, z, z', w) \right| + O(1) \sum_{w \in \text{supp}(v_0)} \left| \widehat{\pi}_-^{\gamma_\ell}(w|y) - \widehat{\pi}_-(w|y) \right|, \end{aligned}$$

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

$$\lim_{\gamma_\ell \rightarrow 0} \gamma_\ell \mathfrak{a}_t^{\gamma_\ell}(z, z', y) = \mathfrak{a}_t(z, z', y),$$

The set  $\widehat{Y}_z$  is not empty since  $z \in \widehat{Z}$ . Since for any  $y \in \widehat{Y}_z$ ,  $\mathfrak{a}_t(z, y) \neq 0$ , it follows from (90) that  $\gamma_\ell \mathfrak{A}_t^{\gamma_\ell}(z, z')$  converges as  $\gamma_\ell$  goes to zero with for any  $y \in \widehat{Y}_z$ ,

$$\lim_{\gamma_\ell \rightarrow 0} \gamma_\ell \mathfrak{A}_t^{\gamma_\ell}(z, z') = \frac{\mathfrak{a}_t(z, z', y)}{\mathfrak{a}_t(z, y)}.$$

The proof of the first part of Lemma 3.4 is completed.

We now turn to the proof of the second part of Lemma 3.4. One will only compute  $\lim_{\gamma_\ell \rightarrow 0} (\gamma_\ell^2 \mathfrak{A}_t^{\gamma_\ell}(z, z''))$  for  $z \in \widehat{Z}, z'' \in \mathbb{V}(z)$  and the expression of  $\lim_{\gamma_\ell \rightarrow 0} (\gamma_\ell^2 \mathfrak{B}_t^{\gamma_\ell}(z, z''))$  follows from similar calculations. For any  $y \in \mathcal{X}$  and any  $t > 0$ , one has

$$\mathfrak{A}_t^\gamma(z, z'') = \frac{\mathfrak{a}_t^\gamma(z, z'', y)}{\mathfrak{a}_t^\gamma(z, y)},$$

with

$$\mathfrak{a}_t^\gamma(z, z'', y) := \int \alpha_t^\gamma(y, z, z'', w) d\widehat{\pi}_-^\gamma(w|y).$$

It remains to compute  $\lim_{\gamma_\ell \rightarrow 0} \gamma_\ell^2 \mathfrak{a}_t^{\gamma_\ell}(z, z'', y)$  to prove (45). As above, Lemma 4.4 (iii) provides

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

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

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

Moreover, Lemma 4.4 (i), (ii) and (iii) gives that for any  $w \in \text{supp}(v_0)$  and  $y \in \text{supp}(v_1)$ ,

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

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

$$\begin{aligned} & \left| \gamma_\ell^2 \mathfrak{a}_t^{\gamma_\ell}(z, z'', y) - \mathfrak{a}_t(z, z'', y) \right| \\ & \leq \sup_{w \in \text{supp}(v_0)} \left| \gamma_\ell^2 \alpha_t^{\gamma_\ell}(y, z, z'', w) - \alpha_t(y, z, z'', w) \right| + O(1) \sum_{w \in \text{supp}(v_0)} \left| \widehat{\pi}_-^{\gamma_\ell}(w|y) - \widehat{\pi}_-(w|y) \right|, \end{aligned}$$



for all  $\gamma_\ell < 1/(2SK)$ . The end of the proof of the second part of Lemma 3.4 is identical to the one the first part.  $\square$

## 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] S. Bobkov, C. Houdré, and P. Tetali. The subgaussian constant and concentration inequalities. *Israel Journal of Mathematics*, 156(1):255–283, 2006.
- [7] A.I. Bonciocat and K.T. Sturm. Mass transportation and rough curvature bounds for discrete spaces. *J. Funct. Anal.*, 256(9):2944–2966, 2009.
- [8] 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.
- [9] 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.
- [10] M. Erbar, C. Henderson, G. Menz, and P. Tetali. Ricci curvature bounds for weakly interacting markov chains. *Electronic Journal of Probability*, 22, 2017.
- [11] 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.
- [12] M. Erbar and J. Maas. Gradient flow structures for discrete porous medium equations. *Discrete & Continuous Dynamical Systems-A*, 34(4):1355–1374, 2014.
- [13] 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.
- [14] M. Fathi and J. Maas. Entropic ricci curvature bounds for discrete interacting systems. *The Annals of Applied Probability*, 26(3):1774–1806, 2016.
- [15] 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.
- [16] 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.
- [17] N. Gozlan, C. Roberto, P.-M. Samson, and P. Tetali. Transport proofs of some discrete variants of the prékopa-leindler inequality. *Annali della Scuola Normale Superiore di Pisa. Classe di scienze*, 22(3):1207–1232, 2021.
- [18] D. Halikias, B. Klartag, and B. A Slomka. Discrete variants of brunn–minkowski type inequalities. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 30, pages 267–279, 2021.
- [19] E. Hillion.  $W_{1,+}$ -interpolation of probability measures on graphs. *Electron. J. Probab.*, 19:no. 92, 29, 2014.
- [20] 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.
- [21] B. Klartag, G. Kozma, P. Ralli, and P. Tetali. Discrete curvature and abelian groups. *Canadian Journal of Mathematics*, 68(3):655–674, 2016.
- [22] B. Klartag and J. Lehec. Poisson processes and a log-concave Bernstein theorem. *Studia Math.*, 247(1):85–107, 2019.
- [23] C. Léonard. From the schrödinger problem to the monge–kantorovich problem. *Journal of Functional Analysis*, 262(4):1879–1920, 2012.
- [24] C. Léonard. Some properties of path measures. In *Séminaire de Probabilités XLVI*, pages 207–230. Springer, 2014.
- [25] 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.
- [26] C. Léonard. Lazy random walks and optimal transport on graphs. *Ann. Probab.*, 44(3):1864–1915, 2016.
- [27] C. Léonard. On the convexity of the entropy along entropic interpolations. In *Measure Theory in Non-Smooth Spaces*, pages 194–242. Sciendo Migration, 2017.
- [28] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.
- [29] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.

- [30] K. Marton. Bounding  $\bar{d}$ -distance by informational divergence: a method to prove measure concentration. *Ann. Probab.*, 24(2):857–866, 1996.
- [31] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
- [32] 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.
- [33] T. Mikami. Monge’s problem with a quadratic cost by the zero-noise limit of h-path processes. *Probability theory and related fields*, 129(2):245–260, 2004.
- [34] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.*, 256(3):810–864, 2009.
- [35] 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.
- [36] Y. Ollivier and C. Villani. A curved Brunn-Minkowski inequality on the discrete hypercube. *Siam Journal on Discrete Mathematics*, 26(3):983–996, 2012.
- [37] P.-M. Samson. Concentration of measure principle and entropy-inequalities. In *Convexity and concentration*, pages 55–105. Springer, 2017.
- [38] P.-M. Samson. Transport-entropy inequalities on locally acting groups of permutations. *Electronic Journal of Probability*, 22, 2017.
- [39] K.T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.
- [40] K.T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.
- [41] 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.
- [42] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.
- [43] 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, UNIV GUSTAVE EIFFEL, UNIV PARIS EST CRETEIL, CNRS, LAMA UMR8050 F-77447 MARNE-LA-VALLÉE, FRANCE

Email address: paul-marie.samson@univ-eiffel.fr