



**HAL**  
open science

## Transport inequalities for random point measures

Nathaël Gozlan, Ronan Herry, Giovanni Peccati

► **To cite this version:**

Nathaël Gozlan, Ronan Herry, Giovanni Peccati. Transport inequalities for random point measures. Journal of Functional Analysis, 2021, 10.1016/j.jfa.2021.109141 . hal-02475798

**HAL Id: hal-02475798**

**<https://hal.science/hal-02475798v1>**

Submitted on 12 Feb 2020

**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.

# Transport inequalities for random point measures

Nathaël Gozlan, Ronan Herry, Giovanni Peccati

February 12, 2020

## Abstract

We derive transport-entropy inequalities for mixed binomial point processes, and for Poisson point processes. We show that when the finite intensity measure satisfies a Talagrand transport inequality, the law of the point process also satisfies a Talagrand type transport inequality. We also show that a Poisson point process (with arbitrary  $\sigma$ -finite intensity measure) always satisfies a universal transport-entropy inequality à la Marton. We explore the consequences of these inequalities in terms of concentration of measure and modified logarithmic Sobolev inequalities. In particular, our results allow one to extend a deviation inequality by Reitzner [31], originally proved for Poisson random measures with finite mass.

**Keywords:** Point Processes, Transport-Entropy Inequalities, Concentration of Measure.

**MSC2010:** 60G55; 60E15; 46E27.

## 1 Introduction

In this paper, we establish new transport-entropy inequalities for various random point measures including the important case of Poisson random measures.

The investigation of transport-entropy inequalities starts in the nineties with works by Marton [28, 27] and by Talagrand [39], in connection with the concentration of measure phenomenon for product measures. We refer to [24, Chapter 6], [41, Chapter 22], [5, Chapter 9], and [13, 16] for general introductions and surveys on these intimately related topics. In this field, the so-called *Talagrand's transport inequality* and *Marton's universal transport inequality*, that we now briefly present, are of prime importance.

In order to state these inequalities, we introduce some notations. Suppose that  $(Z, d)$  is some complete separable metric space, and denote by  $\mathcal{P}(Z)$  the set of Borel probability measures on  $Z$ . For  $k \geq 1$ , the *Monge-Kantorovich distance*  $\mathcal{W}_k$  (also called Wasserstein distance) between  $\nu_1, \nu_2 \in \mathcal{P}(Z)$  is defined by

$$(1.1) \quad \mathcal{W}_k^k(\nu_1, \nu_2) = \inf \mathbb{E} \left[ d^k(X_1, X_2) \right],$$

where the infimum runs over the set of all pairs  $(X_1, X_2)$  of random variables such that  $X_1 \sim \nu_1$  and  $X_2 \sim \nu_2$ . The *relative entropy*  $\mathcal{H}$  of  $\nu_1$  with respect to  $\nu_2$  is defined by

$$(1.2) \quad \mathcal{H}(\nu_1|\nu_2) = \int \log \frac{d\nu_1}{d\nu_2} d\nu_1,$$

if  $\nu_1 \ll \nu_2$ , and  $+\infty$  otherwise.

We say that a probability measure  $\mu \in \mathcal{P}(Z)$  satisfies *Talagrand's transport inequality*<sup>1</sup> with constant  $a > 0$  if, for all  $\nu_1, \nu_2 \in \mathcal{P}(Z)$ , it holds

$$(1.3) \quad \mathcal{W}_2^2(\nu_1, \nu_2) \leq a [\mathcal{H}(\nu_1|\mu) + \mathcal{H}(\nu_2|\mu)].$$

---

<sup>1</sup>Here, to be precise, we refer to a symmetric version of Talagrand's inequality, involving two probability measures  $\nu_1, \nu_2$ .

This inequality, first proved for the standard Gaussian measures in [39], now plays a central role in the literature devoted to concentration of measure, coercive inequalities for Markov semi-groups (logarithmic Sobolev and Poincaré inequalities, and their offshoots) [30, 7], and curvature-dimension conditions for weighted Riemannian manifolds or, more generally, metric measured spaces [37, 38, 25]. Various sufficient conditions are available to ensure that a given probability measure satisfies (1.3) (see [13] for an overview and [17] for a necessary and sufficient condition when  $Z = \mathbb{R}$ ). Dimension free concentration of measure is the main application of Talagrand’s inequality (and of its variants): if a probability measure  $\mu$  satisfies (1.3), then, for any  $n \geq 1$ , and for any vector  $(X_1, \dots, X_n)$  of i.i.d random variables with common law  $\mu$ , it holds

$$(1.4) \quad \mathbb{P}(f(X_1, \dots, X_n) \geq t) \leq e^{-\frac{t^2}{2a}}, \quad \forall t \geq 0,$$

for every function  $f: Z^n \rightarrow \mathbb{R}$  which is of mean 0 with respect to  $\mu^{\otimes n}$  and 1-Lipschitz with respect to the  $\ell^2$  product distance on  $Z^n$ ,

$$(1.5) \quad d_2(x, y) = \left( \sum_{i=1}^n d^2(x_i, y_i) \right)^{1/2}, \quad x, y \in Z^n.$$

We refer to e.g [13] for a presentation of the nice general argument due to Marton that enables to deduce (1.4) from (1.3). Remarkably, all the deviation inequalities in (1.4) hold simultaneously with the same constant  $a$  for all dimensions  $n \geq 1$  (a property which actually characterizes Talagrand’s inequality [14]). This type of dimension free bounds plays an important role in analysis, probability, or statistics in high dimensions [40, 24].

Marton’s inequality involves a variant of the Monge-Kantorovich costs, which we denote by  $\mathcal{M}$  in the sequel and we define as follows:

$$(1.6) \quad \mathcal{M}^2(\nu_1|\nu_2) = \inf \mathbb{E} [\mathbb{P}(X_1 \neq X_2|X_2)^2],$$

where the infimum runs over the set of all pairs  $(X_1, X_2)$  of random variables such that  $X_1 \sim \nu_1$  and  $X_2 \sim \nu_2$ . We refer to Section 2.2 for the presentation of the unifying framework of generalized transport costs introduced in [21] which contains in particular Monge-Kantorovich as well as Marton transport costs. Marton’s transport cost also admits the following explicit expression: if  $\nu_1$  and  $\nu_2$  are absolutely continuous with respect to some measure  $\mu$  on  $Z$ , with  $\nu_1 = f_1\mu$  and  $\nu_2 = f_2\mu$ , then the results of [28] imply that

$$(1.7) \quad \mathcal{M}^2(\nu_1|\nu_2) = \int \left[ 1 - \frac{f_1}{f_2} \right]_+^2 f_2 \, d\mu.$$

Contrary to Talagrand’s inequality, which holds only for some specific probability measures, according to [28], all probability measures satisfy Marton’s inequality. A classical version of Marton’s universal transport inequality reads as follows: for any probability measure  $\mu$  on  $Z$ , it holds that

$$(1.8) \quad \mathcal{M}^2(\nu_1|\nu_2) \leq 4\mathcal{H}(\nu_1|\mu) + 4\mathcal{H}(\nu_2|\mu),$$

for all  $\nu_1, \nu_2 \in \mathcal{P}(Z)$ . One can understand this inequality as a reinforcement of the classical Csiszar-Kullback-Pinsker inequality (see e.g [13] and the references therein) comparing the squared total variation distance to relative entropy. We refer to [11, 32, 33] for subsequent refinements of Marton’s inequality. Similarly to (1.3), Marton’s inequality has interesting consequences in terms of concentration of measure. As Marton [28] shows, (1.8) gives back the universal concentration of measure inequalities for product measures involving the so-called “convex distance” discovered by Talagrand in [40]. To avoid entering into too technical details in this introduction, let us recall a more concrete consequence of (1.8) in terms of deviation inequalities for convex functions (see [28, 11] for details). Namely, if we equip  $Z = \mathbb{R}^p$  with the standard Euclidean norm

and  $\mu \in \mathcal{P}(\mathbb{R}^p)$  has a bounded support with diameter  $D$ , then for any  $n \geq 1$ , and for any vector  $(X_1, \dots, X_n)$  of i.i.d random variables with common law  $\mu$ , we have that

$$(1.9) \quad \mathbb{P}(f(X_1, \dots, X_n) \geq t) \leq e^{-t^2/4D^2}, \quad \forall t \geq 0,$$

for all *convex* or *concave* function  $f: (\mathbb{R}^p)^n \rightarrow \mathbb{R}$  which is of mean 0 with respect to  $\mu^{\otimes n}$  and 1-Lipschitz with respect to the Euclidean norm on  $(\mathbb{R}^p)^n$ .

Transport-entropy inequalities share a pivotal tensorization mechanism; this mechanism enables one to recover the dimension-free deviation bounds (1.4) or (1.9) from (1.3) or (1.8). For instance, if  $\mu$  satisfies Talagrand's inequality on a space  $(Z, d)$  with a constant  $a$ , then its tensor product  $\mu^{\otimes n}$  also satisfies Talagrand's inequality on the space  $(Z^n, d_2)$  with  $d_2$  given by (1.5) and with the same constant  $a$ . This tensorization property, which comes from classical disintegration formulas for the functionals  $\mathcal{W}_2^2$  and  $\mathcal{H}$ , is explained in full generality, for instance in [13]. In this work, we use this stability by tensorization as a pervasive tool to obtain transport type inequalities for random point measures.

Let us now give a flavor of the content of this paper. In brief, understanding what happens to Talagrand's and Marton's transport inequalities in the framework of random point processes serves as the basic motivation behind this work. More precisely, in this article, we consider a particular class of point processes called *mixed binomial point processes*; they are of the following form:

$$(1.10) \quad \eta = \sum_{i=1}^N \delta_{X_i},$$

where  $N$  is a random variable taking values in  $\mathbb{N} := \{0, 1, 2, \dots\}$ ,  $(X_i)_{i \geq 1}$  is a sequence of i.i.d random variables with values in  $Z$  independent of  $N$ , and  $\delta_u$  is the Dirac mass at  $u \in Z$ . In the definition above, and everywhere in the rest of this work, by convention, an empty sum is equal to zero. This random point process  $\eta$  takes values in the set of all Borel finite measures on  $Z$ , written  $\mathcal{M}_b(Z)$  throughout the paper. We usually denote by  $\kappa \in \mathcal{P}(\mathbb{N})$  the law of  $N$ , and by  $\mu \in \mathcal{P}(Z)$  the common law of the  $X_i$ 's, which we often refer to as the "sampling measure" of the process  $\eta$ . We denote the law of  $\eta$  by  $B_{\mu, \kappa} \in \mathcal{P}(\mathcal{M}_b(Z))$ . Notably, when  $N$  is a Poisson random variable with mean  $\lambda \geq 0$ , the point process  $\eta$  is a *Poisson point process*, with (finite) intensity measure  $\mathbb{E}\eta = \nu = \lambda\mu \in \mathcal{M}_b(Z)$ . We also denote by  $\Pi_\nu \in \mathcal{P}(\mathcal{M}_b(Z))$  the law of such a Poisson point process. We can also consider Poisson point processes with a  $\sigma$ -finite intensity measure; for the sake of simplicity, in this introduction, we only consider the finite intensity case. We refer to [Section 3.1](#) for further information and definitions on these elementary random point processes.

In this work, we highlight a general principle leading to transport-entropy inequalities for point processes:  $B_{\mu, \kappa}$  inherits the transport inequalities satisfied by its sampling probability measure  $\mu$ . To state a first representative result illustrating this general rule, let us introduce some notations. Given  $\nu_1, \nu_2 \in \mathcal{M}_b(Z)$ , we extend the definition given in (1.1):

$$(1.11) \quad \mathcal{W}_2^2(\nu_1, \nu_2) = \begin{cases} m\mathcal{W}_2^2\left(\frac{\nu_1}{m}, \frac{\nu_2}{m}\right), & \text{if } \nu_1(Z) = \nu_2(Z) = m > 0; \\ 0, & \text{if } \nu_1(Z) = \nu_2(Z) = 0; \\ +\infty, & \text{if } \nu_1(Z) \neq \nu_2(Z). \end{cases}$$

Then, we define a process level Monge-Kantorovich cost  $\mathbb{W}_2^2$  on  $\mathcal{P}(\mathcal{M}_b(Z))$  as follows: for any  $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_b(Z))$ ,

$$\mathbb{W}_2^2(\Pi_1, \Pi_2) = \inf \mathbb{E} [\mathcal{W}_2^2(\eta_1, \eta_2)],$$

where the infimum runs over the set of couples  $(\eta_1, \eta_2)$  of random measures such that  $\eta_1 \sim \Pi_1$  and  $\eta_2 \sim \Pi_2$ . The condition  $\mathbb{W}_2^2(\Pi_1, \Pi_2) < +\infty$  is a strong assumption. Indeed, if it holds then there exists a couple  $(\eta_1, \eta_2)$ , such that  $\eta_1 \sim \Pi_1$  and  $\eta_2 \sim \Pi_2$ , and such that  $\eta_1(Z) = \eta_2(Z)$  almost surely.

With these notions at hand, we can now state an analogue of Talagrand's inequality for mixed binomial processes.

**Theorem 1.1.** *Suppose that  $\mu \in \mathcal{P}(Z)$  satisfies Talagrand's inequality (1.3) with a constant  $a > 0$ , then for any  $\kappa \in \mathcal{P}(\mathbb{N})$ , the probability measure  $B_{\mu, \kappa} \in \mathcal{P}(\mathcal{M}_b(Z))$  satisfies the following inequality: for all  $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_b(Z))$  such that  $\mathbb{W}_2^2(\Pi_1, \Pi_2) < +\infty$ ,*

$$(1.12) \quad \mathbb{W}_2^2(\Pi_1, \Pi_2) + 2a\mathcal{H}(\lambda|\kappa) \leq a\mathcal{H}(\Pi_1|B_{\mu, \kappa}) + a\mathcal{H}(\Pi_2|B_{\mu, \kappa}),$$

where  $\lambda \in \mathcal{P}(\mathbb{R}_+)$  is such that for all  $i \in \{1, 2\}$ ,  $\Pi_i(\{\eta \in \mathcal{M}(Z) : \eta(Z) \in A\}) = \lambda(A)$ , for all Borel  $A \subset \mathbb{R}_+$ .

This result is another instance of a phenomenon highlighted in a paper by Erbar & Huesmann [12]: geometric functional inequalities lift from the base space  $Z$  to the space of configurations. More precisely, if  $Z$  is a Riemannian manifold with Ricci curvature bounded by  $K \in \mathbb{R}$ , with infinite volume measure  $\text{vol}$ , and Riemannian distance  $d$ , the results of [12] show that

$$(1.13) \quad \mathcal{H}(\Pi_t|\Pi_{\text{vol}}) \leq (1-t)\mathcal{H}(\Pi_0|\Pi_{\text{vol}}) + t\mathcal{H}(\Pi_1|\Pi_{\text{vol}}) - \frac{K}{2}t(1-t)\mathbb{W}_2^2(\Pi_0, \Pi_1),$$

where  $\Pi_{\text{vol}}$  denotes the Poisson point process with intensity measure  $\text{vol}$ , and where  $\Pi_t$  is any  $\mathbb{W}_2$ -geodesic from  $\Pi_0$  to  $\Pi_1$ , for  $\Pi_0$  and  $\Pi_1 \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  with  $\mathcal{M}_{\mathbb{N}}(Z)$  denoting the space of measures  $\eta$  such that  $\eta(K) \in \mathbb{N}$  for all compact  $K \subset Z$ . This means that the curvature property of the metric measured space  $(Z, d, \text{vol})$  transfers to the (extended) metric measured space  $(\mathcal{M}_{\mathbb{N}}(Z), \mathbb{W}_2, \Pi_{\text{vol}})$ . Indeed, (1.13) states that  $(\mathcal{M}_{\mathbb{N}}(Z), \mathbb{W}_2, \Pi_{\text{vol}})$  has a synthetic Ricci curvature in the sense of Lott-Sturm-Villani bounded below by  $K$  (see, e.g. one of the seminal papers [37, 25] for definitions of synthetic Ricci lower bound in terms of convexity of the relative entropy, as well as [1, Definition 9.1] for a definition in the extended metric measured spaces setting). This lifting of Ricci lower bound is very natural, if we think of  $\Pi_{\text{vol}}$  as the invariant measure of a system of non-interacting  $N$  Brownian motions on  $Z$ , where  $N$  has a Poisson law with mean  $\text{vol}(M)$  (possibly infinite). In the case where  $K > 0$ , the manifold is compact so that  $\Pi_{\text{vol}}$  has a representation as a mixed binomial process, and, provided the results of [12] carry to the compact case, taking  $t = \frac{1}{2}$ , and using that the entropy is non-negative in (1.13) immediately yields, for all  $\Pi_0, \Pi_1 \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(M))$  such that  $\mathbb{W}_2^2(\Pi_0, \Pi_1) < \infty$ :

$$\frac{K}{4}\mathbb{W}_2^2(\Pi_0, \Pi_1) \leq \mathcal{H}(\Pi_0|\Pi_{\text{vol}}) + \mathcal{H}(\Pi_1|\Pi_{\text{vol}}).$$

Let us stress that our argument works under the sole assumption that the space  $Z$  supports a Talagrand inequality and is quite elementary.

To state an analogue of Marton's inequality, we introduce the following cost: for any  $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_b(Z))$ ,

$$\mathbb{M}^2(\Pi_1|\Pi_2) = \inf \mathbb{E} \left[ \int \mathbb{E} \left[ \left[ 1 - \frac{\eta_1(x)}{\eta_2(x)} \right]_+^2 \eta_2(dx) \right], \right.$$

where  $\eta_i(x)$  is a slight abuse of notation for  $\eta_i(\{x\})$ , and where the infimum runs over the set of couples  $(\eta_1, \eta_2)$  of random measures such that  $\eta_1 \sim \Pi_1$  and  $\eta_2 \sim \Pi_2$ .

**Theorem 1.2.** *For any  $\nu \in \mathcal{M}_b(Z)$ , the Poisson point process  $\Pi_\nu$  satisfies the following inequality: for all  $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_b(Z))$ ,*

$$\mathbb{M}^2(\Pi_1|\Pi_2) \leq 4\mathcal{H}(\Pi_1|\Pi_\nu) + 4\mathcal{H}(\Pi_2|\Pi_\nu).$$

A similar statement actually holds for Poisson point processes associated with a  $\sigma$ -finite sampling measure  $\nu$ , see Theorem 3.5 for a precise result.

Let us mention that other authors derived universal transport-entropy inequalities for Poisson point processes. For instance, Ma et al. [26] combine arguments of exponential integrability of [6] with the modified logarithmic Sobolev inequality for Poisson point processes of [42] in order to derive some  $L_1$ -transportation inequalities for Poisson point processes. To state their results, let us define

$$\mathbb{T}_{TV}(\Pi_1, \Pi_2) = \inf \mathbb{E} \|\eta_1 - \eta_2\|_{TV},$$

where the infimum is over all  $\eta_1 \sim \Pi_1$  and all  $\eta_2 \sim \Pi_2$ , and  $\|\cdot\|_{TV}$  is the total variation norm. Assuming, for simplicity,  $\nu(Z) = 1$  and choosing  $\phi = 1$  in [26, Theorem 2.6] yields

$$(1.14) \quad \alpha(\mathbb{T}_{TV}(\Pi, \Pi_\nu)) \leq \mathcal{H}(\Pi|\Pi_\nu),$$

where  $\alpha(r) = (1+r)\log(1+r) - r$ . In order to compare our inequalities in greater details, we would need to compare  $\alpha(\mathbb{T}_{TV})$  and  $\mathbb{M}^2$ . However, this does not seem to be possible in general. One reason for that is that since the  $\eta_i$ 's are not probability measures, one cannot apply Jensen's inequality.

As mentioned earlier, Marton introduces her inequality so one may derive concentration of measure with respect to the so-called Talagrand convex distance. In the setting of Poisson point processes, Reitzner [31] uses the Talagrand convex distance to prove concentration of measure results for Poisson random measures. In Section 4, we recover, in the spirit of Marton's work, the results of [31] using Theorem 1.2. Building on the ideas of [8], Bachmann & Peccati [2] considers a different approach towards concentration of measure for Poisson point processes. They obtain various general conditions on a functional  $F: \mathcal{M}_b(Z) \rightarrow \mathbb{R}$  under which the random variable  $F(\eta)$  satisfies some deviation inequality. Since the space  $\mathcal{M}_b(Z)$  does not come with a natural distance, a rather involved technical condition replaces the condition of being Lipschitz that appeared in (1.4). However, [2, 3] shows that the so-called *geometric U-statistics* always satisfy this condition, and hence always satisfy some concentration of measure estimate. Based on Theorem 1.2, we recover a deviation inequality for  $U$ -statistics in the spirit of [2, 3] with a simple argument. It would be interesting to know whether we could recover the main result of [2] for generic functionals from our transport inequality. This question will be examined elsewhere.

Theorems 1.1 and 1.2 are consequences of more general results presented in Sections 3.2 and 3.3. As already mentioned above, the tensorization property of the transport inequality satisfied by the sampling measure on the base space is crucial in our analysis. Heuristically, this dimension free structure matters, since a random point process  $\eta$  as in (1.10) can be seen as a random vector of random dimension (up to the permutation of its coordinates). It should be noted that in the work [12] on Ricci curvature bounds on the configuration space, even if the arguments are different, tensorization properties also play a pivotal role (in that case the tensorization of the Bakry-Emery condition).

The paper is organized as follows. Section 2 contains generalities about spaces of measure, optimal transport and functional inequalities. Section 2.1 recalls some basic facts about the weak topology of the space of measures over a Polish space. Sections 2.2 and 2.3 present, in a self-contained way, the structural properties of the generalized optimal transport and the related transport-entropy inequalities introduced by [21], and the subsequent results of [4]. This framework encompasses, as a particular case, Talagrand and Marton inequalities. In Section 2.4, we recall two further stability properties of transport-entropy inequalities: stability by tensorization (as already explained this property is at the heart of our argument), stability by push-forward, and stability by approximation. Section 3 is the core of this article and contains our main results that are Theorems 3.1, 3.3 and 3.5. We start, in Section 3.1, with some reminders on binomial point processes and Poisson point processes. In Section 3.2, we prove Theorem 3.1, that states that a Talagrand inequality for  $\mu$  implies a Talagrand-type inequality for  $B_{\mu, \kappa}$ , and that implies Theorem 1.1. We first show our result for the simple case of binomial processes of deterministic size, and then extend it to random size using some properties of optimal transport and relative entropy under conditioning. Section 3.3 contains several results about Marton-type transport-entropy inequalities on the configuration space. Our first result in that direction is Theorem 3.3, that states that a Marton inequality for  $\mu$  implies a transport-entropy inequality for binomial process of deterministic size. Since, for a properly chosen distance, all probability measures satisfy a Marton inequality, this produces Corollary 3.4 stating a universal Marton inequality for binomial process of fixed size. Contrary to the case of Talagrand inequality, we cannot extend our results to mixed binomial processes of arbitrary random size. In Theorem 3.5, we extend Corollary 3.4 to Poisson point processes with  $\sigma$ -finite intensity measure using a strong approximation of the Poisson point



process by thinning a binomial process. This result implies [Theorem 1.2](#). In [Section 4](#), we study the consequences of our transport-entropy inequalities in terms of concentration of measure. In [Section 4.1](#), we explain how to recover the main findings of [\[31\]](#) from [Theorem 3.5](#). In [Section 4.2](#), we also discuss concentration of measure for a particular class of functionals, containing in particular geometric  $U$ -statistics, and extend some results of [\[2, 3\]](#) to this class of functionals. Finally, in [Section 5](#) we discuss the consequences of [Theorem 3.3](#) in terms of modified logarithmic Sobolev inequality. Following [\[21\]](#), [Theorem 5.1](#) is a general modified logarithmic Sobolev inequality on the Poisson space where the energy term is given in term of an infimum-convolution operator. [Corollary 5.2](#) shows how we can partially deduce from this general modified logarithmic Sobolev inequality the modified logarithmic Sobolev inequality of [\[42\]](#) for monotonic functionals.

## Acknowledgments

N.G. is supported by a grant of the Simone and Cino Del Duca Foundation; R.H. gratefully acknowledges support of the European Union through the European Research Council Advanced Grant for *Karl-Theodor Sturm* “Metric measure spaces and Ricci curvature – analytic, geometric and probabilistic challenges”; G.P. is supported by the FNR grant FoRGES (R-AGR-3376-10) at Luxembourg University.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Preliminaries on measures and optimal transport</b>	<b>7</b>
2.1	Topology of spaces of measures . . . . .	7
2.2	Optimal transport between measures of arbitrary masses . . . . .	8
2.2.1	Couplings, transport costs and generalized transport costs . . . . .	8
2.2.2	Some examples . . . . .	8
2.2.3	Existence of optimal couplings, lower semicontinuity of transport costs and stability of optimal couplings . . . . .	9
2.2.4	Optimal (partial) transport between configurations . . . . .	10
2.3	Transport-entropy inequalities . . . . .	11
2.3.1	Definitions and examples . . . . .	11
2.4	Stability by tensorization, by push-forward, and by approximation . . . . .	12
<b>3</b>	<b>Transport inequalities for random point processes</b>	<b>15</b>
3.1	Random point processes . . . . .	15
3.2	Transport-entropy inequalities when the sampling measure satisfies a Talagrand type inequality . . . . .	15
3.2.1	Binomial process with deterministic size . . . . .	16
3.2.2	Binomial processes with general size . . . . .	17
3.3	Universal Marton-type inequalities for Poisson point processes . . . . .	18
3.3.1	Deterministic size case . . . . .	18
3.3.2	Poisson case . . . . .	19
<b>4</b>	<b>Applications to concentration estimates</b>	<b>23</b>
4.1	Generic concentration of measure inequalities . . . . .	23
4.2	Concentration of measure for convex functionals . . . . .	26
<b>5</b>	<b>Modified logarithmic Sobolev inequality</b>	<b>28</b>
5.1	Introduction . . . . .	28
5.2	Transport inequalities and variants of the log-Sobolev inequality on the Poisson space	28

<b>6</b>	<b>Some open questions</b>	<b>30</b>
6.1	From modified logarithmic Sobolev to the transport-entropy . . . . .	30
6.2	Links with displacement convexity . . . . .	31
6.3	Interacting point processes . . . . .	31

## 2 Preliminaries on measures and optimal transport

In what follows,  $(E, d)$  is a complete and separable metric space. We recall below basic topological properties of some sets of measures on  $E$ , then we introduce the framework and the basic properties of (generalized) optimal transport costs between two measures on  $E$  and finally, we recall definitions and basic properties of transport-entropy inequalities. In the subsequent sections, we apply the material of the present section with  $E = Z$ ,  $Z$  being the state space of our point process, or with  $E = \mathcal{M}_b(Z)$ .

### 2.1 Topology of spaces of measures

We always regard the space  $E$  as a measurable space equipped with its Borel  $\sigma$ -algebra  $\mathfrak{B}(E)$ . We recall that we denote by  $\mathcal{M}_b(E)$  the space of finite measures and by  $\mathcal{P}(E)$  the set of probability measures on  $E$ . We also write  $\mathcal{M}_0(E)$  for the space of Radon measures that are finite on balls. We write  $\mathcal{C}_b(E)$  for the space of bounded continuous functions, and  $\mathcal{C}_0(E)$  for the space of continuous functions vanishing outside of a ball. We endow the sets  $\mathcal{M}_b(E)$  and  $\mathcal{P}(E)$  with the *weak* topology that is generated by the maps  $f^* : \nu \mapsto \int f d\nu$  with  $f \in \mathcal{C}_b(E)$ . On the other hand, we endow the space  $\mathcal{M}_0(E)$  with the *vague* topology that is generated by the  $f^*$  with  $f \in \mathcal{C}_0(E)$ . According to [22, Section 4.1],  $\mathcal{M}_b(E)$  and  $\mathcal{P}(E)$  (with the weak topology), and  $\mathcal{M}_0(E)$  (with the vague topology) are Polish spaces (complete, separable, and metrizable). We often work with the following subset of  $\mathcal{M}_b(E)$ :

$$(2.1) \quad \mathcal{M}_{\mathbb{N}}(E) = \left\{ \sum_{i=1}^k \delta_{a_i} : k \in \mathbb{N}, a_1, \dots, a_k \in E \right\}$$

corresponding to finite configurations over  $E$ . We also work with  $\mathcal{M}_{\bar{\mathbb{N}}}(E)$  consisting of all  $\xi \in \mathcal{M}_0(E)$  such that, for every closed ball  $B \subset E$ ,  $\xi_{\upharpoonright B} \in \mathcal{M}_{\mathbb{N}}(B)$ . The spaces  $\mathcal{M}_{\mathbb{N}}(E)$  or  $\mathcal{M}_{\bar{\mathbb{N}}}(E)$  are the natural state space of our point processes.

**Lemma 2.1.** *The set  $\mathcal{M}_{\mathbb{N}}(E)$ , resp.  $\mathcal{M}_{\bar{\mathbb{N}}}(E)$ , is a closed subset of  $\mathcal{M}_b(E)$ , resp.  $\mathcal{M}_0(E)$ . In particular,  $\mathcal{M}_{\mathbb{N}}(E)$  and  $\mathcal{M}_{\bar{\mathbb{N}}}(E)$  are a Polish spaces.*

*Proof.* Let  $\{\xi_n; n \in \mathbb{N}\} \subset \mathcal{M}_{\mathbb{N}}(E)$  converging to some  $\xi \in \mathcal{M}_b(E)$ . Let us show that  $\xi$  belongs to  $\mathcal{M}_{\mathbb{N}}(E)$ . Since  $\xi_n(E)$  is a sequence of integers converging to  $\xi(E)$ , it must be that  $k := \xi(E) \in \mathbb{N}$  and that  $\xi_n(E) = k$  for all  $n$  sufficiently large. Then, since  $(\xi_n)_{n \geq 0}$  is converging, it is according to Prohorov's theorem a tight sequence: for any  $\varepsilon > 0$ , there exists a compact set  $K_\varepsilon \subset E$  such that  $\sup_{n \geq 0} \xi_n(E \setminus K_\varepsilon) \leq \varepsilon$ . Taking in particular  $\varepsilon = 1/2$  and using the fact that the  $\xi_n$ 's are sums of Dirac masses, one sees that  $\xi_n(E \setminus K_{1/2}) = 0$ . In other words, for all  $n$  large enough  $\xi_n = \sum_{i=1}^k \delta_{x_i^n}$ , with  $x_1^n, \dots, x_k^n \in K_{1/2}$ . Using a diagonal extraction argument, one can assume without loss of generality that  $x_i^n \rightarrow x_i \in K_{1/2}$  for all  $i \in \{1, \dots, k\}$ . This easily implies that  $\xi = \sum_{i=1}^k \delta_{x_i}$  and so  $\xi \in \mathcal{M}_{\mathbb{N}}(E)$ . Now let  $\{\xi_n; n \in \mathbb{N}\} \subset \mathcal{M}_{\bar{\mathbb{N}}}(E)$  converging to some  $\xi \in \mathcal{M}_0(E)$ . By what precedes for every ball  $B \subset E$ , there exists  $\xi_B \in \mathcal{M}_{\mathbb{N}}(B)$  such that

$$\xi_{n \upharpoonright B} \xrightarrow[n \rightarrow \infty]{\text{weakly}} \xi_B.$$

By a compatibility argument, we have that  $\xi_{\upharpoonright B} = \xi_B$ ; hence  $\xi \in \mathcal{M}_{\bar{\mathbb{N}}}(E)$ . As closed subsets of a Polish spaces,  $\mathcal{M}_{\mathbb{N}}(E)$  and  $\mathcal{M}_{\bar{\mathbb{N}}}(E)$  are themselves Polish.  $\square$



## 2.2 Optimal transport between measures of arbitrary masses

### 2.2.1 Couplings, transport costs and generalized transport costs

A *coupling* of  $\nu_1$  and  $\nu_2 \in \mathcal{M}_0(E)$  is an element  $N \in \mathcal{M}_0(E \times E)$  such that, for all  $A \in \mathfrak{B}(E)$ ,  $N(A \times E) = \nu_1(A)$  and  $N(E \times A) = \nu_2(A)$ . Observe that if  $\nu_1, \nu_2 \in \mathcal{M}_b(E)$ , then  $N \in \mathcal{M}_b(E \times E)$ . Imposing  $\nu_1(E) = \nu_2(E)$  is a necessary and sufficient condition for the existence of a coupling between  $\nu_1$  and  $\nu_2$ . Indeed, if there exists a coupling  $N$  of  $\nu_1$  and  $\nu_2$  then  $\nu_1(E) = N(E \times E) = \nu_2(E)$ . On the other hand, if  $\nu_1$  and  $\nu_2 \in \mathcal{M}_0(E)$  with  $\nu_1(E) = \nu_2(E) = n \in \mathbb{N} \cup \{\infty\}$ , then we can write  $\nu_1 = \sum_{q=1}^n \nu_{1,q}$  and  $\nu_2 = \sum_{q=1}^n \nu_{2,q}$  with  $\nu_{i,q} \in \mathcal{P}(E)$ . Then  $N = \sum_{q=1}^n (\nu_{1,q} \otimes \nu_{2,q})$  is a coupling of  $\nu_1$  and  $\nu_2$ . According to [9, Proposition 13, Section 2.7], for every coupling  $N$ , there exists a measurable application  $E \ni x \mapsto p_x \in \mathcal{P}(E)$  such that for all  $A$  and  $B \in \mathfrak{B}(E)$ :

$$(2.2) \quad N(A \times B) = \int_A p_x(B) \nu_1(dx).$$

We call  $p = \{p_x; x \in E\}$  the *disintegration kernel* of  $N$  along  $\nu_1$ . For short, we often abbreviate

$$N(dx dy) = p_x(dy) \nu_1(dx).$$

Following [21], given a bi-measurable *cost function*  $c: E \times \mathcal{P}(E) \rightarrow [0, \infty]$ , the (*generalized*) *optimal transport cost* associated to  $c$  from  $\nu_1$  to  $\nu_2 \in \mathcal{M}_0(E)$ , denoted by  $\mathcal{T}_c(\nu_2|\nu_1)$ , is given by

$$(2.3) \quad \mathcal{T}_c(\nu_2|\nu_1) = \inf \int c(x, p_x) \nu_1(dx),$$

where the infimum runs over all couplings  $N(dx dy) = p_x(dy) \nu_1(dx)$  between  $\nu_1$  and  $\nu_2$ . We use here the convention that  $\inf \emptyset = \infty$ . In particular, we have  $\mathcal{T}_c(\nu_2|\nu_1) = \infty$  as soon as  $\nu_1(E) \neq \nu_2(E)$ . We always implicitly assume that the bi-measurability of  $c$  is satisfied in the rest of the document. Note that,  $\mathcal{T}_c$  is, in general, not symmetric with respect to  $\nu_1$  and  $\nu_2$ . When  $c$  is of the form

$$(2.4) \quad c(x, p) = \int \omega(x, y) p(dy),$$

for some  $\omega: E \times E \rightarrow [0, \infty]$ , we say, with a slight abuse of language, that  $c$  is *linear*, and we set  $\mathcal{T}_\omega = \mathcal{T}_c$ . The cost  $\mathcal{T}_\omega$  is the transportation cost associated to  $\omega$  in the usual sense of optimal transport:

$$(2.5) \quad \mathcal{T}_\omega(\nu_2|\nu_1) = \inf \int \omega(x, y) N(dx dy),$$

where the infimum is running over all couplings  $N$  of  $\nu_1$  and  $\nu_2$ . Such objects have been intensively studied and play an important role in many different areas of mathematics. The reader can look at the reference [41] and the references therein.

### 2.2.2 Some examples

Let us recall some classical choices for transport costs.

**Monge-Kantorovich distance** Choosing  $\omega = d^k$  with  $k \geq 1$  yields  $\mathcal{T}_{d^k} = \mathcal{W}_k^k$ , the Monge-Kantorovich transport cost already introduced in (1.1) and (1.11).

**Marton-type transport distances** Given a measurable function  $\rho: E \times E \rightarrow [0, \infty]$  and a convex function  $\alpha: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , we introduce the Marton type cost function

$$(2.6) \quad c(x, p) = \alpha \left( \int \rho(x, y) p(dy) \right), \quad x \in E, p \in \mathcal{P}(E),$$

and, following the notations of [21], we denote by  $\tilde{\mathcal{T}}_{\alpha,\rho}$  the associated optimal transport cost, namely

$$\mathcal{T}_{\alpha,\rho}(\nu_2|\nu_1) = \inf \int \alpha \left( \int \rho(x,y)p_x(dy) \right) \nu_1(dx),$$

where the infimum runs over all couplings  $N(dx dy) = p_x(dy)\nu_1(dx)$ . When  $\nu_1, \nu_2 \in \mathcal{P}(E)$ , we can also write

$$(2.7) \quad \tilde{\mathcal{T}}_{\alpha,\rho}(\nu_2|\nu_1) = \inf \mathbb{E} [\alpha (\mathbb{E} [\rho(X_1, X_2)|X_1])],$$

where the infimum runs over the couples  $(X_1, X_2)$  with  $X_1 \sim \nu_1$  and  $X_2 \sim \nu_2$ . In particular, taking the *Hamming distance*  $\rho(x, y) = \mathbf{1}_{x \neq y}$  and  $\alpha(x) = x^2, x \geq 0$ , gives back Marton's cost  $\mathcal{M}^2$  already introduced in (1.6) and (1.7).

The costs  $\mathcal{W}_p$  are symmetric since  $d$  is symmetric. More generally  $\mathcal{T}_\omega$  is symmetric provided  $\omega$  is symmetric. On the other hand  $\tilde{\mathcal{T}}_{\alpha,\rho}$  is, in general, not symmetric even when  $\rho$  is symmetric.

Applying Jensen's inequality in (2.6), we easily check that, for all  $\nu_1, \nu_2 \in \mathcal{M}_b(E)$ ,

$$(2.8) \quad \tilde{\mathcal{T}}_{\alpha,\rho}(\nu_2|\nu_1) \leq \mathcal{T}_{\alpha \circ \rho}(\nu_1, \nu_2);$$

moreover, if  $\nu_1, \nu_2 \in \mathcal{P}(E)$  using Jensen's inequality in the integral with respect to  $\nu_1$  in the definition of  $\tilde{\mathcal{T}}_{\alpha,\rho}$  shows that:

$$(2.9) \quad \alpha(\mathcal{T}_\rho(\nu_2|\nu_1)) \leq \tilde{\mathcal{T}}_{\alpha,\rho}(\nu_2|\nu_1).$$

Choosing the *Hamming distance*  $\rho(x, y) = d_H(x, y) = \mathbf{1}_{x \neq y}, x, y \in E$ , we have for any  $\nu_1, \nu_2 \in \mathcal{P}(E)$ :

$$\mathcal{T}_{d_H}(\nu_2|\nu_1) = \mathcal{T}_{d_H^2}(\nu_2|\nu_1) = \inf_{X_1 \sim \nu_1, X_2 \sim \nu_2} \mathbb{P}(X_1 \neq X_2) = \|\nu_1 - \nu_2\|_{TV},$$

where  $\|\cdot\|_{TV}$  is the classical *total variation* norm. In this case, (2.8) and (2.9) yield

$$\|\nu_2 - \nu_1\|_{TV}^2 \leq \mathcal{M}^2(\nu_2|\nu_1) \leq \|\nu_2 - \nu_1\|_{TV}.$$

### 2.2.3 Existence of optimal couplings, lower semicontinuity of transport costs and stability of optimal couplings

Backhoff-Veraguas, Beiglöck & Pammer [4] prove the following result (the authors work with  $\nu_1, \nu_2 \in \mathcal{P}(E)$  but the generalization to  $\mathcal{M}_b(E)$  presents no difficulty).

**Theorem 2.2.** *Let  $c: E \times \mathcal{P}(E) \rightarrow [0, \infty]$  be convex in its second argument, and jointly lower semi-continuous, then, for all  $\nu_1$  and  $\nu_2 \in \mathcal{M}_b(E)$  with same total mass, there exists a coupling  $N(dx dy) = p_x(dy)\nu_1(dx)$  of  $\nu_1$  and  $\nu_2$  with disintegration kernel  $p$  such that*

$$(2.10) \quad \mathcal{T}_c(\nu_2|\nu_1) = \int c(x, p_x)\nu_1(dx).$$

*Proof.* Let  $m = \nu_1(E) = \nu_2(E)$  be the common mass of the measures. If  $m = 0$ , then  $\nu_1 = \nu_2 = 0$  and  $\mathcal{T}_c(\nu_2|\nu_1) = 0$  and the coupling  $N = 0$  is optimal. Now, assume that  $m > 0$ . Since  $\mathcal{T}_c(\nu_2|\nu_1) = m\mathcal{T}_c(\nu_2/m|\nu_1/m)$ , we use [4, Theorem 1.1] to conclude.  $\square$

Rather classically, this yields lower semi-continuity of  $\mathcal{T}_c$  as shown in [4]. Below we simply adapt the argument to finite measures of arbitrary total mass.

**Theorem 2.3.** *Let  $c: E \times \mathcal{P}(E) \rightarrow [0, \infty]$  be jointly lower semi-continuous, and convex in its second argument, then  $\mathcal{T}_c: \mathcal{M}_b(E) \times \mathcal{M}_b(E) \rightarrow [0, \infty]$  is lower semi-continuous.*

*Proof.* Let  $(\mu_k)_{k \geq 0}$  and  $(\nu_k)_{k \geq 0}$  respectively weakly converging to  $\mu$  and  $\nu$  in  $\mathcal{M}_b(E)$ . Let us show that

$$(2.11) \quad \liminf_{k \rightarrow +\infty} \mathcal{T}_c(\mu_k | \nu_k) \geq \mathcal{T}_c(\mu | \nu).$$

By definition of the weak convergence, we have  $\mu_k(E) \rightarrow \mu(E)$  and  $\nu_k(E) \rightarrow \nu(E)$ . If  $\mu(E) \neq \nu(E)$ , then for all  $k$  sufficiently large, it holds  $\mu_k(E) \neq \nu_k(E)$  and so  $\mathcal{T}_c(\mu_k | \nu_k) = +\infty$ . Thus (2.11) holds in this case. If  $\mu(E) = \nu(E) = 0$ , then  $\mathcal{T}_c(\mu | \nu) = 0$  and (2.11) trivially holds. Let us now assume that  $\mu(E) = \nu(E) = m > 0$ . Extracting a subsequence if necessary, one can assume that  $\mu_k(E) = \nu_k(E) = m_k > 0$  for all  $k \geq 0$ . Since  $\mathcal{T}_c(\mu_k | \nu_k) = m_k \mathcal{T}_c(\mu_k/m_k | \nu_k/m_k)$  and using [4, Theorem 2.9], we see that

$$\liminf_{k \rightarrow +\infty} \mathcal{T}_c(\mu_k | \nu_k) = \liminf_{k \rightarrow +\infty} m_k \mathcal{T}_c(\mu_k/m_k | \nu_k/m_k) \geq m \mathcal{T}_c(\mu/m | \nu/m) = \mathcal{T}_c(\mu | \nu),$$

which completes the proof.  $\square$

When  $c$  is lower semi-continuous, we do not know if  $\mathcal{T}_c: \mathcal{M}_0(E) \times \mathcal{M}_0(E) \rightarrow [0, \infty]$  is lower semi-continuous.

Finally, let us recall the following stability result taken from [41, Theorem 4.6].

**Theorem 2.4.** *Let  $\omega: E \times E \rightarrow [0, \infty]$  be a lower semi-continuous cost function. Let  $\mu, \nu \in \mathcal{P}(E)$  be such that  $\mathcal{T}_\omega(\mu, \nu) < \infty$  and let  $N$  be an optimal transport plan. Let  $\tilde{N} \in \mathcal{M}_b(E \times E)$  such that  $\tilde{N} \leq N$  and  $\tilde{N} \neq 0$ , then the probability measure*

$$(2.12) \quad N' = \frac{\tilde{N}}{\tilde{N}(E \times E)},$$

is an optimal transport plan between its marginals.

## 2.2.4 Optimal (partial) transport between configurations

In this section, we study in more details  $\mathcal{T}_c(\chi | \xi)$ , in the case where  $\xi, \chi \in \mathcal{M}_{\mathbb{N}}(E)$  are configurations.

**Proposition 2.5.** *Let  $\omega: E \times E \rightarrow [0, \infty]$  be lower semi-continuous. Let  $\xi, \chi \in \mathcal{M}_{\mathbb{N}}(E)$  such that  $\xi = \sum_{i=1}^k \delta_{a_i}$  and  $\chi = \sum_{i=1}^k \delta_{b_i}$  for some  $k \geq 1$  and  $a_1, \dots, a_k, b_1, \dots, b_k \in E$ . Then,*

$$(2.13) \quad \mathcal{T}_\omega(\chi | \xi) = \inf \left\{ \sum_{i=1}^k \omega(a_i, b_{\sigma(i)}) \right\}.$$

where the infimum runs over the set of all permutations  $\sigma$  of  $\{1, \dots, k\}$ .

This result is very classical (but usually stated for pairwise distinct  $a_i$ 's or  $b_j$ 's); we include a proof for completeness.

*Proof.* The inequality  $\leq$  is clear, since for any permutation  $\sigma$ , the measure  $N = \sum_{i=1}^k \delta_{(a_i, b_{\sigma(i)})}$  is a coupling between  $\xi$  and  $\chi$ . To prove the converse, let us consider a coupling  $N$  between  $\xi$  and  $\chi$  and define the matrix  $M = [M_{i,j}]_{1 \leq i, j \leq k}$  as follows:

$$M_{i,j} = \frac{N(\{a_i\} \times \{b_j\})}{\xi(a_i) \nu(b_j)}, \quad \forall 1 \leq i, j \leq k.$$

The matrix  $M$  is bi-stochastic, since denoting by  $S = \{b_1, \dots, b_k\}$  the support of  $\chi$ ,

$$\begin{aligned} \sum_{j=1}^k M_{i,j} &= \frac{1}{\xi(a_i)} \sum_{j=1}^k \frac{N(\{a_i\} \times \{b_j\})}{\chi(b_j)} = \frac{1}{\xi(a_i)} \int_S \frac{N(\{a_i\} \times \{y\})}{\chi(y)} \chi(dy) \\ &= \frac{1}{\xi(a_i)} \sum_{y \in S} N(\{a_i\} \times \{y\}) = 1 \end{aligned}$$

and similarly,  $\sum_{i=1}^k M_{i,j} = 1$ . Since  $\int \omega(x, y) N(dx dy) = \sum_{i,j} \omega(x_i, y_j) M_{i,j}$ , the other inequality follows from Birkhoff Theorem on extremal points of doubly stochastic matrices.  $\square$

Even when  $c$  is bounded, the cost  $\mathcal{T}_c$  is infinite as soon as the measures are of different total masses. One way to avoid this, is to work with *partial* optimal transport that we now define. For  $\xi, \chi \in \mathcal{M}_{\mathbb{N}}(E)$ , we define

$$(2.14) \quad \mathcal{T}_{c,0}(\chi|\xi) = \begin{cases} \min \{ \mathcal{T}_c(\chi|\xi') : \xi' \in \mathcal{M}_{\mathbb{N}}(E), \xi' \leq \xi, \xi'(Z) = \chi(Z) \}, & \text{if } \xi(Z) \geq \chi(Z); \\ \min \{ \mathcal{T}_c(\chi'| \xi) : \chi' \in \mathcal{M}_{\mathbb{N}}(E), \chi' \leq \chi, \chi'(Z) = \xi(Z) \}, & \text{if } \xi(Z) < \chi(Z). \end{cases}$$

Here and in the rest of this paper, the notation  $\nu' \leq \nu$ , for  $\nu$  and  $\nu' \in \mathcal{M}_0(E)$ , means that  $\nu'(A) \leq \nu(A)$  for every Borel set  $A$ . The following result follows immediately from [Proposition 2.5](#).

**Proposition 2.6.** *Let  $\xi, \chi \in \mathcal{M}_{\mathbb{N}}(E)$  be such that  $\xi = \sum_{i=1}^n \delta_{a_i}$  and  $\chi = \sum_{i=1}^k \delta_{b_i}$ , for some  $k, n \geq 1$  and  $a_1, \dots, a_n, b_1, \dots, b_k \in E$ . Assume  $n \geq k$  then*

$$(2.15) \quad \mathcal{T}_{\omega,0}(\xi, \chi) = \inf \left\{ \sum_{i=1}^k \omega(x_i, y_i) : \sum_{i=1}^k \delta_{x_i} \leq \xi, \chi = \sum_{i=1}^k \delta_{y_i} \right\}.$$

## 2.3 Transport-entropy inequalities

### 2.3.1 Definitions and examples

Recall the definition of the relative entropy functional  $\mathcal{H}$  given in [\(1.2\)](#). Given a cost function  $c: E \times \mathcal{P}(E) \rightarrow [0, \infty]$ , we say that a probability measure  $\gamma \in \mathcal{P}(E)$  satisfies the *transport-entropy inequality*  $(\mathbf{T}_c)$  with constants  $a_1, a_2 > 0$  if

$$(\mathbf{T}_c) \quad \mathcal{T}_c(\nu_2|\nu_1) \leq a_1 \mathcal{H}(\nu_1|\gamma) + a_2 \mathcal{H}(\nu_2|\gamma), \quad \forall \nu_1, \nu_2 \in \mathcal{P}(E).$$

When  $c(x, p) = \int \omega(x, y) p(dy)$ , for  $\omega: E \times E \rightarrow [0, \infty]$  we use  $(\mathbf{T}_\omega)$  to designate the corresponding transport-entropy inequality, which reads as follows

$$(\mathbf{T}_\omega) \quad \mathcal{T}_\omega(\nu_1, \nu_2) \leq a_1 \mathcal{H}(\nu_1|\gamma) + a_2 \mathcal{H}(\nu_2|\gamma), \quad \forall \nu_1, \nu_2 \in \mathcal{P}(E).$$

When  $c(x, p) = \alpha \left( \int \rho(x, y) p(dy) \right)$  for some cost function  $\rho: E \times E \rightarrow [0, \infty]$  and some convex function  $\alpha$  as in [\(2.6\)](#) on  $\mathbb{R}^+$ , we say that  $\gamma$  satisfies the inequality  $(\tilde{\mathbf{T}}_{\alpha,\rho})$  with constants  $a_1, a_2 > 0$  if

$$(\tilde{\mathbf{T}}_{\alpha,\rho}) \quad \tilde{\mathcal{T}}_{\alpha,\rho}(\nu_2|\nu_1) \leq a_1 \mathcal{H}(\nu_1|\gamma) + a_2 \mathcal{H}(\nu_2|\gamma), \quad \forall \nu_1, \nu_2 \in \mathcal{P}(E).$$

In these definitions,  $a_1$  or  $a_2$  can assume the value  $\infty$ , and we use the convention that  $0 \cdot \infty = 0$ . For instance, if, for instance,  $a_2 = \infty$  the previous inequalities are non-trivial if and only if we take  $\nu_2 = \gamma$ . In that case,  $(\tilde{\mathbf{T}}_{\alpha,\rho})$  is interpreted as

$$(\tilde{\mathbf{T}}_{\alpha,\rho}) \quad \tilde{\mathcal{T}}_{\alpha,\rho}(\gamma|\nu_1) \leq a_1 \mathcal{H}(\nu_1|\gamma), \quad \forall \nu_1 \in \mathcal{P}(E).$$

When  $a_1 = a_2 = a < \infty$ , we simply say that  $\gamma$  satisfies a transport-entropy inequality with constant  $a$ .

We refer to [\[13, 21\]](#) for a panorama of cost functions and transport inequalities. We refer to inequalities of the form  $(\mathbf{T}_\omega)$  as *Talagrand type* transport-entropy inequalities; whereas we use *Marton type* transport-entropy inequalities for inequalities of the form  $(\tilde{\mathbf{T}}_{\alpha,\rho})$ .

**Talagrand's inequality** Choosing  $\omega = d^2$  in  $(\mathbf{T}_\omega)$  yields Talagrand's inequality [\(1.3\)](#).

**Marton's universal transport inequalities** Choosing  $\rho = d_H$ ,  $\alpha(x) = x^2$ ,  $x \geq 0$  and  $a = 4$  in  $(\tilde{\mathbf{T}}_{\alpha,\rho})$  gives Marton's inequality (1.8), which holds true for any probability measure  $\gamma$ . Actually, Marton's inequality (1.8) can be slightly improved. Dembo [11] shows that any probability measure  $\gamma$  satisfies a family of inequalities  $(\tilde{\mathbf{T}}_{\alpha,\rho})$  with respect to the Hamming distance  $d_H$ , namely:

$$\tilde{\mathcal{T}}_{\alpha_t, d_H}(\nu_2 | \nu_1) \leq \frac{1}{t} \mathcal{H}(\nu_1 | \gamma) + \frac{1}{1-t} \mathcal{H}(\nu_2 | \gamma), \quad \forall t \in [0, 1];$$

where

$$(2.16) \quad \alpha_t(u) = \frac{t(1-u) \log(1-u) - (1-tu) \log(1-tu)}{t(1-t)}, \quad \forall u \in [0, 1].$$

For  $t = 0$  or  $t = 1$ ,  $\alpha_t$  is defined by taking the corresponding limit in the definition above, namely:

$$\begin{aligned} \alpha_0(u) &= (1-u) \log(1-u) + u; \\ \alpha_1(u) &= -u - \log(1-u). \end{aligned}$$

We easily check that  $\alpha_t(u) \geq \alpha_0(u) \geq \frac{u^2}{2}$ ,  $u \in [0, 1]$ . In particular, with  $t = \frac{1}{2}$ , Dembo's result gives back Marton's inequality (1.8).

## 2.4 Stability by tensorization, by push-forward, and by approximation

We now state the three main operations that enable us to obtain transport inequalities for mixed binomial processes: the tensorization, push-forward, and strong approximations. Importantly for us, transport-entropy inequalities are closed under these operations.

Transport-entropy inequalities enjoy the following well known tensorization property [21, Theorem 4.11].

**Proposition 2.7.** *Consider, for  $i = 1, \dots, n$ , costs functions  $c_i: E_i \times \mathcal{P}(E_i) \rightarrow [0, \infty]$  for some Polish spaces  $E_i$ . Suppose that the functions  $c_i$  are convex with respect to their second variables. If, for all  $1 \leq i \leq n$ ,  $\gamma_i \in \mathcal{P}(E_i)$  satisfies  $(\mathbf{T}_{c_i})$  with constant  $a_1, a_2 > 0$ , then  $\gamma_1 \otimes \dots \otimes \gamma_n$  satisfies  $(\mathbf{T}_{\bar{c}})$  (with the same constant  $a_1, a_2$ ) where*

$$(2.17) \quad \bar{c}(x, p) = \sum_{i=1}^n c_i(x_i, p_i),$$

where  $x = (x_1, \dots, x_n) \in E_1 \times \dots \times E_n$  and  $p \in \mathcal{P}(\times E_i)$  has  $i$ -th marginal  $p_i$ .

*Remark 1.* The reference [21] proves the theorem only for the case  $a_1, a_2 < \infty$  but if  $a_1 = 1$  (resp.  $a_2 = \infty$ ) their proof still works, replacing  $\nu'$  (resp.  $\nu$ ) by  $\mu^2$ .

Now let us turn to the stability by push-forward. Let us fix two Polish spaces  $X$  and  $Y$ , and a measurable map  $S: X \rightarrow Y$ . Recall that for  $\gamma \in \mathcal{P}(X)$  the *push forward* of  $\gamma$  by  $S$  is the element of  $\mathcal{P}(Y)$  defined by  $S_{\#}\gamma(A) = \gamma(S^{-1}A)$ , for all  $A \in \mathfrak{B}(Y)$ . Given a cost  $c: X \times \mathcal{P}(X) \rightarrow [0, \infty]$  we define its *push forward* by

$$(2.18) \quad S_{\#}c(y, q) = \inf \{c(x, p) : S(x) = y, S_{\#}p = q\},$$

for all  $y \in Y$  and  $q \in \mathcal{P}(Y)$ . Likewise, for  $\omega: X \times X \rightarrow [0, \infty]$ , we write

$$(2.19) \quad S_{\#}\omega(y_1, y_2) = \inf \{\omega(x_1, x_2) : S(x_1) = y_1, S(x_2) = y_2\}, \quad y_1, y_2 \in Y.$$

We now show that the push-forward preserves transport-entropy inequalities.

**Proposition 2.8.** *Suppose that  $\gamma \in \mathcal{P}(X)$  satisfies  $(\mathbf{T}_c)$  (with constants  $a_1, a_2 > 0$ ) for some convex cost function  $c: X \times \mathcal{P}(X) \rightarrow \mathbb{R}_+$ . Then, the probability measure  $S_{\#}\gamma$  satisfies  $(\mathbf{T}_{S_{\#}c})$  with the same constants  $a_1, a_2$ . In particular, if  $\gamma$  satisfies  $(\mathbf{T}_{\omega})$  for some  $\omega: X \times X \rightarrow [0, \infty]$ , then  $S_{\#}\gamma$  satisfies  $(\mathbf{T}_{S_{\#}\omega})$ .*

This property is rather classical for transport inequalities of the form  $(\mathbf{T}_\omega)$ ; it goes back at least to [29] (in the dual language of infimum convolution inequalities). See [15, 17, 21] for applications of this transfer principle. However, at this level of generality, **Proposition 2.8** is new.

*Proof.* For short, we write  $\bar{\gamma} = S_{\#}\gamma$  and  $\bar{c} = S_{\#}c$ . Let  $\bar{\nu}_1, \bar{\nu}_2 \in \mathcal{P}(Y)$  be such that  $\mathcal{H}(\bar{\nu}_1|\bar{\gamma}) < \infty$  and  $\mathcal{H}(\bar{\nu}_2|\bar{\gamma}) < \infty$ . Let  $\bar{h}_1$  be the density of  $\bar{\nu}_1$  with respect to  $\bar{\gamma}$ ; then for  $f \in \mathcal{C}_b(Y)$

$$(2.20) \quad \int f d\bar{\nu}_1 = \int f \bar{h}_1 d\bar{\gamma} = \int f(S) \bar{h}_1(S) d\gamma = \int f(S) d\nu_1,$$

where  $d\nu_1 = \bar{h}_1(S) d\gamma$ . Therefore, there exists at least one probability measure  $\nu_1$  on  $X$  such that  $\bar{\nu}_1 = S_{\#}\nu_1$ . On the other hand,

$$(2.21) \quad \mathcal{H}(\bar{\nu}_1|\bar{\gamma}) = \int \bar{h}_1 \log \bar{h}_1 d\bar{\gamma} = \int \bar{h}_1(S) \log \bar{h}_1(S) d\gamma = \mathcal{H}(\nu_1|\gamma).$$

Let us consider the function  $\bar{\mathcal{T}}_c(\cdot|\cdot): \mathcal{P}(X) \times \mathcal{P}(X) \rightarrow [0, \infty]$  defined by

$$(2.22) \quad \bar{\mathcal{T}}_c(\bar{\nu}_1|\bar{\nu}_2) = \inf \{ \mathcal{T}_c(\nu_1|\nu_2) : \bar{\nu}_1 = S_{\#}\nu_1 \text{ and } \bar{\nu}_2 = S_{\#}\nu_2 \}.$$

According to what precedes, for all  $\bar{\nu}_i, i = 1, 2$ , such that  $\mathcal{H}(\bar{\nu}_i|\bar{\gamma}) < \infty$ , there exist  $\nu_i, i = 1, 2$ , on  $\mathcal{P}(X)$  such that  $\bar{\nu}_i = S_{\#}\nu_i$  and so

$$(2.23) \quad \bar{\mathcal{T}}_c(\bar{\nu}_1|\bar{\nu}_2) \leq \mathcal{T}_c(\nu_1|\nu_2) \leq a_1 \mathcal{H}(\nu_1|\gamma) + a_2 \mathcal{H}(\nu_2|\gamma) = a_1 \mathcal{H}(\bar{\nu}_1|\bar{\gamma}) + a_2 \mathcal{H}(\bar{\nu}_2|\bar{\gamma}).$$

Now let us prove that

$$(2.24) \quad \bar{\mathcal{T}}_c(\bar{\nu}_1|\bar{\nu}_2) \geq \mathcal{T}_{\bar{c}}(\bar{\nu}_1|\bar{\nu}_2).$$

Let  $\bar{\nu}_1, \bar{\nu}_2$  such that  $\mathcal{H}(\bar{\nu}_i|\bar{\gamma}) < +\infty, i = 1, 2$ ; there exist  $\nu_1, \nu_2$  such that  $\bar{\nu}_i = S_{\#}\nu_i$ . Let  $p$  be a kernel such that  $\nu_1 = \nu_2 p$ . Equivalently, there exists a pair of random variables  $(X_1, X_2)$  with  $X_2 \sim \nu_2$  and  $\text{Law}(X_1|X_2 = x) = p_x$ . Consider  $Y_1 = S(X_1)$  and  $Y_2 = S(X_2)$ ; for all bounded continuous functions  $f_1, f_2$  on  $Y$  it holds

$$\begin{aligned} \mathbb{E}[f_1(Y_1)f_2(Y_2)] &= \mathbb{E}[f_1(S(X_1))f_2(S(X_2))] = \mathbb{E} \left[ \int f_1(S(x_1)) dp_{X_2}(x_1) f_2(S(X_2)) \right] \\ &= \mathbb{E} \left[ \int f_1(y_1) d(S_{\#}p_{X_2})(y_1) f_2(S(X_2)) \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ \int f_1(y_1) d(S_{\#}p_{X_2})(y_1) | S(X_2) \right] f_2(S(X_2)) \right] \end{aligned}$$

Consider a regular conditional probability  $k$  for  $\text{Law}(X_2|Y_2)$ ; then

$$(2.25) \quad \begin{aligned} \mathbb{E}[f_1(Y_1)f_2(Y_2)] &= \mathbb{E} \left[ \int \left( \int f_1(y_1) d(S_{\#}p_{x_2})(y_1) \right) k_{Y_2}(dx_2) f_2(Y_2) \right] \\ &= \iint f_1(y_1) f_2(y_2) \bar{p}_{y_2}(dy_1) \bar{\nu}_2(dy_2), \end{aligned}$$

with

$$(2.26) \quad \bar{p}_{y_2} = \int (S_{\#}p_{x_2}) k_{y_2}(dx_2).$$



This proves that  $\bar{\nu}_2 \bar{p} = \nu_1$ .

$$\begin{aligned}
\int c(x_2, p_{x_2}) \nu_2(dx_2) &\geq \int \bar{c}(S(x_2), S_{\#}p_{x_2}) \nu_2(dx_2) \\
&= \iint \bar{c}(S(x_2), S_{\#}p_{x_2}) k_{y_2}(dx_2) \bar{\nu}_2(dy_2) \\
&= \iint \bar{c}(y_2, S_{\#}p_{x_2}) k_{y_2}(dx_2) \bar{\nu}_2(dy_2) \\
&\geq \int \bar{c}\left(y_2, \int S_{\#}p_{x_2} k_{y_2}\right) \bar{\nu}_2(dy_2) \\
&\geq \mathcal{T}_{\bar{c}}(\bar{\nu}_1 | \bar{\nu}_2),
\end{aligned}$$

where the first inequality comes from the definition of  $\bar{c}$ , the third is a consequence of the fact that  $S(x_2) = y_2$  for  $k_{y_2}$  almost all  $x_2$ , and the fourth follows from the convexity of  $\bar{c}$  (which is itself a simple consequence of the convexity of  $c$ ). Therefore, taking the infimum over  $p$  yields to  $\mathcal{T}_c(\nu_1 | \nu_2) \geq \mathcal{T}_{\bar{c}}(\bar{\nu}_1 | \bar{\nu}_2)$ . Taking the infimum over all  $\nu_1, \nu_2$  such that  $S_{\#}\nu_i = \bar{\nu}_i$  finally gives (2.24) and completes the proof.  $\square$

In applications, one often approximates a Poisson point process of interest by a simpler point process for which one can establish a transport-entropy inequality. The following general lemma allows us to transfer a transport-entropy inequality from the simpler process to the other process.

**Lemma 2.9.** *Let  $E$  be a Polish space and let  $c: E \times \mathcal{P}(E) \rightarrow \mathbb{R}_+$  be a lower semi-continuous cost. Let  $\{\gamma_n; n \in \mathbb{N}\} \cup \{\gamma\}$  such that for all  $A \in \mathfrak{B}(E)$ :*

$$\gamma_n(A) \xrightarrow{n \rightarrow \infty} \gamma(A).$$

*If for all  $n \in \mathbb{N}$ ,  $\gamma_n$  satisfies  $(\mathbf{T}_c)$  with constants  $a_{1,n}, a_{2,n} > 0$ , then  $\gamma$  satisfies  $(\mathbf{T}_c)$  with the constants  $a_i = \liminf_n a_{i,n}$ ,  $i = 1, 2$ .*

*Proof.* Thanks to the strong convergence, for any  $f: E \rightarrow \mathbb{R}$  bounded and measurable, it holds

$$(2.27) \quad \int_E f(x) \gamma_n(dx) \xrightarrow{n \rightarrow \infty} \int_E f(x) \gamma(dx).$$

Take  $\nu^1, \nu^2 \in \mathcal{P}(E)$ , absolutely continuous with respect to  $\gamma$  with *bounded* respective densities  $f_1$  and  $f_2$ . Let us define for all  $i = 1, 2$  and for all Borel set  $A \subset E$ ,

$$\nu_n^i(A) = \frac{\int_A f_i(x) \gamma_n(dx)}{\int f_i(x) \gamma_n(dx)}.$$

By assumption on  $\gamma_n$ , we have that

$$(2.28) \quad \mathcal{T}_c(\nu_n^1 | \nu_n^2) \leq a_{1,n} \mathcal{H}(\nu_n^1 | \gamma_n) + a_{2,n} \mathcal{H}(\nu_n^2 | \gamma_n).$$

Using (2.27), we see that  $\nu_n^i \xrightarrow{n \rightarrow \infty} \nu^i$  weakly (and even setwise). Since, by assumption, the cost function  $c$  is lower semi-continuous, by [Theorem 2.3](#),  $\mathcal{T}_c: \mathcal{P}(E) \times \mathcal{P}(E) \rightarrow [0, \infty]$  is also lower semi-continuous, and so

$$\liminf_{n \rightarrow \infty} \mathcal{T}_c(\nu_n^1 | \nu_n^2) \geq \mathcal{T}_c(\nu^1 | \nu^2).$$

On the other hand, using (2.27), we see that for  $i = 1, 2$

$$\mathcal{H}(\nu_n^i | \gamma_n) = \frac{\gamma_n(f_i \log f_i)}{\gamma_n(f_i)} - \log \gamma_n(f_i) \xrightarrow{n \rightarrow \infty} \gamma(f_i \log f_i) = \mathcal{H}(\nu^i | \gamma).$$

Letting  $n \rightarrow \infty$  in (2.28) thus completes the proof in the case of bounded densities. The general case follows by standard approximations arguments.  $\square$

### 3 Transport inequalities for random point processes

In this section, we establish transport-entropy inequalities for a mixed binomial point process of the form (1.10) provided its sampling measure itself satisfies some transport-entropy inequalities. We essentially base our proofs on the material of Section 2.4. The section is organized as follows: first we present some generalities about point processes, then we establish transport-entropy inequalities for mixed binomial point processes under the assumption that the sampling measure satisfies an inequality of the form  $(\mathbf{T}_\omega)$  (a case which covers in particular our Theorem 1.1), and finally we consider the case where the sampling measure satisfies an inequality of the form  $(\tilde{\mathbf{T}}_{\alpha,\rho})$  (which includes our Theorem 1.2).

#### 3.1 Random point processes

From now on, we fix a complete separable metric space  $(Z, d)$  that we regard as our reference state space. Recall the notation  $\mathcal{M}_{\mathbb{N}}(Z)$  (resp.  $\mathcal{M}_{\overline{\mathbb{N}}}(Z)$ ) introduced in (2.1) for the space of finite configurations (resp. configurations); equipped with the weak (resp. vague) topology (introduced in Section 2.1),  $\mathcal{M}_{\mathbb{N}}(Z)$  (resp.  $\mathcal{M}_{\overline{\mathbb{N}}}(Z)$ ) is a Polish space according to Lemma 2.1. We always equip the space  $\mathcal{M}_{\mathbb{N}}(Z)$  and  $\mathcal{M}_{\overline{\mathbb{N}}}(Z)$  with their Borel  $\sigma$ -field (without further mention). For  $n \in \mathbb{N}$ , we also conveniently write  $\mathcal{M}_n(Z) = \{\xi \in \mathcal{M}_{\mathbb{N}}(Z) : \xi(Z) = n\}$ . This space is a closed subset of  $\mathcal{M}_{\mathbb{N}}(Z)$  and is therefore also Polish when equipped with the weak topology.

A *point process* (resp. *finite point process*) on  $Z$  is a  $\mathcal{M}_{\overline{\mathbb{N}}}(Z)$ -valued (resp.  $\mathcal{M}_{\mathbb{N}}(Z)$ -valued) random variable. In this paper, we focus, on the one hand, on mixed binomial process defined in (1.10). For reader's convenience, we recall that this is the class of finite point processes of the form  $\eta = \sum_{i=1}^N \delta_{X_i}$  where  $(X_i)_{i \geq 1}$  is an i.i.d sequence of law  $\mu \in \mathcal{P}(Z)$  independent of the  $\mathbb{N}$ -valued random variable  $N$  whose law is denoted by  $\kappa$ . We also recall that the law of  $\eta$  is denoted  $B_{\mu,\kappa}$ . When  $\kappa = \delta_n$ , where  $n \in \mathbb{N}$ , we write  $B_{\mu,n} \in \mathcal{P}(\mathcal{M}_n(Z))$  instead of  $B_{\mu,\delta_n}$ . A process  $\eta \sim B_{\mu,n}$  is called a *binomial process of size  $n$* .

On the other hand, we consider the important class of *Poisson point processes*. Given a measure  $\nu$  on  $Z$ , we say that  $\eta$  is a *Poisson point process with intensity measure  $\nu$*  whenever: for all pairwise disjoint measurable sets  $A_1, \dots, A_l \in \mathfrak{Z}$ , with  $\nu(A_i) < \infty$  for all  $i = 1, \dots, l$ , the random vector  $(\eta(A_1), \dots, \eta(A_l))$  is a vector of independent Poisson random variables with mean  $(\nu(A_1), \dots, \nu(A_l))$ . Provided such process exists, the intensity measure  $\nu$  characterizes the law of such a process; we write  $\Pi_\nu$  to denote this law. Existence of Poisson point process on arbitrary state space and arbitrary reference measure is in general a non-trivial fact. However, if  $\nu \in \mathcal{M}_b(Z)$ , one can easily check that a mixed binomial process  $\eta$  constructed using  $(X_i)_{i \geq 1}$  of common law  $\mu = \frac{\nu}{\nu(Z)}$  and a random variable  $N$  with a Poisson distribution of mean  $\nu(Z)$  is a Poisson process with intensity measure  $\nu$ . In other words: for all  $\nu \in \mathcal{M}_b(Z)$ ,

$$(3.1) \quad \Pi_\nu = B_{\mu,\kappa},$$

where  $\mu = \frac{\nu}{\nu(Z)} \in \mathcal{P}(Z)$  and  $\kappa \in \mathcal{P}(\mathbb{N})$  is the Poisson distribution of mean  $\nu(Z)$ . By a simple gluing procedure [23, Theorem 3.6], one can also construct Poisson point processes with  $\sigma$ -finite intensity measure. More precisely, writing  $\nu = \sum_{i=1}^{\infty} \nu_n$  with  $\nu_n \in \mathcal{M}_b(Z)$ , we have that

$$\Pi_\nu = \Pi_{\nu_1} * \Pi_{\nu_2} * \dots,$$

where  $*$  is the convolution product. In general,  $\Pi_\nu \in \mathcal{P}(\mathcal{M}_{\overline{\mathbb{N}}}(Z))$ ;  $\Pi_\nu \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  if and only if  $\nu \in \mathcal{M}_b(Z)$ ; and  $\Pi_\nu \in \mathcal{P}(\mathcal{M}_{\overline{\mathbb{N}}}(Z) \setminus \mathcal{M}_{\mathbb{N}}(Z))$  if and only if  $\nu(Z) = \infty$ .

#### 3.2 Transport-entropy inequalities when the sampling measure satisfies a Talagrand type inequality

In this section, we derive some transport-entropy inequalities of the type  $(\mathbf{T}_\omega)$  on  $\mathcal{M}_{\mathbb{N}}(Z)$ , where  $\omega$  is of the form  $\mathcal{T}_\rho: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{M}_{\mathbb{N}}(Z) \rightarrow [0, \infty]$  for some  $\rho: Z \times Z \rightarrow [0, \infty]$ . We denote by  $\mathcal{T}_\rho$  the

transport cost  $\mathcal{T}_{\rho}$ , given explicitly by

$$\mathbb{T}_{\rho}(\Pi_1, \Pi_2) = \inf \mathbb{E} [\mathcal{T}_{\rho}(\xi_1, \xi_2)], \quad \Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)),$$

where the infimum runs over the couples  $(\xi_1, \xi_2)$  of point processes such that  $\xi_1 \sim \Pi_1$  and  $\xi_2 \sim \Pi_2$ . As already observed in [Section 1](#), the condition  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2) < +\infty$  is very strong and it implies in particular that there exists a coupling  $(\xi_1, \xi_2)$  of  $\Pi_1, \Pi_2$  such that  $\xi_1(Z) = \xi_2(Z)$  almost surely. In particular, the finiteness of  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2)$  implies that the law of the total mass is the same under  $\Pi_1$  and  $\Pi_2$ .

**Theorem 3.1.** *Let  $\kappa \in \mathcal{P}(\mathbb{N})$ ,  $\mu \in \mathcal{P}(Z)$ , and  $\rho: Z \times Z \rightarrow [0, \infty]$  be lower semi-continuous. Assume  $\mu \in \mathcal{P}(Z)$  satisfies [\(T \$\_{\rho}\$ \)](#) with constants  $a_1, a_2 > 0$ . Then, for all  $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  such that  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2) < \infty$ , it holds*

$$(3.2) \quad \mathbb{T}_{\rho}(\Pi_1, \Pi_2) + (a_1 + a_2)\mathcal{H}(\lambda|\kappa) \leq a_1\mathcal{H}(\Pi_1|B_{\mu, \kappa}) + a_2\mathcal{H}(\Pi_2|B_{\mu, \kappa}),$$

where  $\lambda \in \mathcal{P}(\mathbb{N})$  is such that for all  $i \in \{1, 2\}$ ,  $\Pi_i(\{\eta \in \mathcal{M}_{\mathbb{N}}(Z) : \eta(Z) \in A\}) = \lambda(A)$ , for all Borel  $A \subset \mathbb{N}$ .

Note that the preceding result could be alternatively stated on the bigger state space  $\mathcal{M}_b(Z)$  instead of  $\mathcal{M}_{\mathbb{N}}(Z)$ . This actually does not make any difference since  $\mathcal{H}(\Pi|B_{\mu, \kappa}) < +\infty$  implies that  $\Pi \in \mathcal{M}_{\mathbb{N}}(Z)$ . Specifying  $\rho = d^2$  in [Theorem 3.1](#) gives [Theorem 1.1](#).

*Remark 2.* Let us comment on [\(3.2\)](#).

This inequality is a transport-entropy inequality ([T \$\_{\omega}\$](#) ), where  $\omega = \mathcal{T}_{\rho}$  (as introduced in [Section 2.3.1](#)) with the *additional constraint* of finiteness of the transport cost:  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2) < +\infty$ . This condition is crucial since otherwise it would be possible to have a finite right hand side and an infinite left hand side. For instance, take  $\mu \in \mathcal{P}(Z)$  satisfying [\(T \$\_{\rho}\$ \)](#) with some constant  $a > 0$ , take  $n_1 \neq n_2$  and consider  $B_{\mu, n_i}$  the binomial point process of size  $n_i$ . Then  $\mathbb{T}_{\rho}(B_{\mu, n_1}, B_{\mu, n_2}) = \infty$ . On the other hand, denoting by  $\Pi_{\mu}$  the law of the Poisson point process with intensity measure  $\mu$  and taking  $\eta \sim \Pi_{\mu}$ , we easily see that  $B_{\mu, n_i} = \text{Law}(\eta|\eta(Z) = n_i)$ . Therefore,  $B_{\mu, n_i}$  is absolutely continuous with respect to  $\Pi_{\mu}$  with Radon-Nikodym derivative given by  $\frac{1}{\mathbb{P}(\eta(Z) = n_i)} \mathbf{1}_{\{\eta(Z) = n_i\}}$  and, thus,  $\mathcal{H}(B_{\mu, n_i}|\Pi_{\mu}) = -\log \mathbb{P}(\eta(Z) = n_i) = -\log \frac{e^{-1}}{n_i!} < +\infty$ .

### 3.2.1 Binomial process with deterministic size

As a first step, we prove [Theorem 3.1](#) in the particular case of binomial point process of fixed size  $n \in \mathbb{N}$ .

**Proposition 3.2.** *Let  $n \in \mathbb{N}$ ,  $\mu \in \mathcal{P}(Z)$  and  $\rho: Z \times Z \rightarrow [0, \infty]$  be lower semi-continuous. Assume  $\mu \in \mathcal{P}(Z)$  satisfies [\(T \$\_{\rho}\$ \)](#) with constants  $a_1, a_2 > 0$ . Then, for all  $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_n(Z))$  such that  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2) < \infty$ , it holds*

$$\mathbb{T}_{\rho}(\Pi_1, \Pi_2) \leq a_1\mathcal{H}(\Pi_1|B_{\mu, n}) + a_2\mathcal{H}(\Pi_2|B_{\mu, n}).$$

*Proof.* For  $n = 0$  there is nothing to prove. Fix  $n \geq 1$ ; by the tensorization property [Proposition 2.7](#), the probability measure  $\mu^{\otimes n} \in \mathcal{P}(Z^n)$  satisfies the transport-entropy inequality ([T \$\_{\rho^n}\$](#) ) with constant  $a$  and with the cost function  $\rho^n$  defined by

$$\rho^n(x, y) = \sum_{i=1}^n \rho(x_i, y_i), \quad x, y \in Z^n.$$

Now, consider the map  $S^n: Z^n \rightarrow \mathcal{M}_n(Z)$  defined by

$$(3.3) \quad S^n(z) = \sum_{i=1}^n \delta_{z_i}, \quad \forall z = (z_1, \dots, z_n) \in Z^n.$$

By construction,  $S_{\#}^n \mu^{\otimes n} = B_{\mu,n}$ . Furthermore, for any  $\xi_1, \xi_2 \in \mathcal{M}_n(Z)$ , it follows from [Proposition 2.5](#) that

$$\begin{aligned} \mathcal{T}_\rho(\xi_2|\xi_1) &= \inf \left\{ \sum_{i=1}^n \rho(x_i, y_i) : \xi_1 = S^n(x), \xi_2 = S^n(y) \right\} \\ &= S_{\#}^n \rho^n, \end{aligned}$$

using the notation introduced in [\(2.19\)](#). Finally, we obtain from [Proposition 2.8](#) that  $B_{\mu,n}$  satisfies a transport-entropy inequality on  $\mathcal{M}_n(Z)$  with respect to the cost function  $\mathcal{T}_\rho$  (and with the same constants  $a_1, a_2$ ).  $\square$

### 3.2.2 Binomial processes with general size

Now we are ready to prove [Theorem 3.1](#) in the general case.

*Proof of [Theorem 3.1](#).* Let  $\Pi_1, \Pi_2 \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  such that  $\mathbb{T}_\rho(\Pi_1, \Pi_2) < \infty$  and  $\mathcal{H}(\Pi_i|B_{\mu,\kappa}) < +\infty$ ,  $i = 1, 2$ . Recall that, according to [Lemma 2.1](#),  $\mathcal{M}_{\mathbb{N}}(Z)$  endowed with the weak topology is a Polish space. According to [Theorem 2.3](#), the transport cost  $\mathcal{T}_\rho: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{M}_{\mathbb{N}}(Z) \rightarrow [0, \infty]$  is lower semi-continuous. Therefore, according to [\[41, Theorem 4.1\]](#) (or [Theorem 2.2](#)), there exists an optimal coupling for  $\mathbb{T}_\rho(\Pi_1, \Pi_2)$ ; we write  $(\xi_1, \xi_2)$  for such an optimal coupling. From the finiteness assumption of the transport distance, we have that  $\xi_1(Z) = \xi_2(Z) := K$  almost surely. We denote by  $\lambda$  the law of  $K$ . For  $n \in \mathbb{N}$  such that  $\lambda(n) > 0$ , we write

$$\Pi^n = \text{Law}((\xi_1, \xi_2)|K = n),$$

we consider  $(\xi_1^n, \xi_2^n) \sim \Pi^n$ , and we write  $\Pi_1^n$  (resp.  $\Pi_2^n$ ) for the law of  $\xi_1^n$  (resp.  $\xi_2^n$ ), that is  $\Pi_1^n$  and  $\Pi_2^n$  are the marginals of  $\Pi^n$ . Applying [Theorem 2.4](#), we see that  $(\xi_1^n, \xi_2^n)$  is an optimal coupling for  $\mathbb{T}_\rho(\Pi_1^n, \Pi_2^n)$ . By construction, we have that

$$\begin{aligned} \mathbb{T}_\rho(\Pi_1, \Pi_2) &= \mathbb{E}[\mathcal{T}_\rho(\xi_1, \xi_2)] \\ &= \sum_{n \in \mathbb{N}} \mathbb{E}[\mathcal{T}_\rho(\xi_1, \xi_2)|K = n] \mathbb{P}(K = n) \\ &= \sum_{n \in \mathbb{N}} \mathbb{E}[\mathcal{T}_\rho(\xi_1^n, \xi_2^n)] \mathbb{P}(K = n) \\ &= \sum_{n \in \mathbb{N}} \mathbb{T}_\rho(\Pi_1^n, \Pi_2^n) \mathbb{P}(K = n). \end{aligned}$$

According to [Proposition 3.2](#), we know that  $B_{\mu,n}$  satisfies a transport-entropy inequality with the cost function  $\mathcal{T}_\rho$  on  $\mathcal{M}_n(Z)$  and constants  $a_1, a_2$ . Thus

$$(3.4) \quad \mathbb{T}_\rho(\Pi_1, \Pi_2) \leq \sum_{n \in \mathbb{N}} (a_1 \mathcal{H}(\Pi_1^n|B_{\mu,n}) + a_2 \mathcal{H}(\Pi_2^n|B_{\mu,n})) \mathbb{P}(K = n).$$

It follows from the chain rule for relative entropy (recalled below) that

$$(3.5) \quad \mathcal{H}(\Pi_i|B_{\mu,\kappa}) = \mathcal{H}(\lambda|\kappa) + \sum_{n \in \mathbb{N}} \mathcal{H}(\Pi_i^n|B_{\mu,n}) \mathbb{P}(K = n).$$

Plugging [\(3.5\)](#) into [\(3.4\)](#) gives [\(3.2\)](#) and completes the proof.

For the sake of completeness, we recall the proof of [\(3.5\)](#). Let  $\eta \sim B_{\mu,\kappa}$  with  $N \sim \kappa$  (see [\(1.10\)](#)). Recall that we assume that  $\mathcal{H}(\Pi_i|B_{\mu,\kappa}) < +\infty$ . Denoting by  $f_i: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}_+$  the Radon-Nikodym derivative of  $\Pi_i$  with respect to  $B_{\mu,\kappa}$ , for all measurable set  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$ , we have that

$$\begin{aligned} \mathbb{P}(\xi_i^n \in A) \mathbb{P}(K = n) &= \mathbb{P}(\xi_i \in A, K = n) \\ (3.6) \quad &= \mathbb{E}[\mathbf{1}_{\{\eta \in A \cap \mathcal{M}_n(Z)\}} f_i(\eta)] \\ &= \mathbb{E}[\mathbf{1}_{\{N=n\}} \mathbf{1}_{\{\eta \in A\}} f_i(\eta)] \\ &= \mathbb{P}(N = n) \mathbb{E}[\mathbf{1}_{\{\eta \in A\}} f_i(\eta)|N = n]. \end{aligned}$$

This shows that if  $\mathbb{P}(K = n) > 0$  then  $\mathbb{P}(N = n) > 0$  and for such  $n$  we have that

$$(3.7) \quad \frac{d\Pi_i^n}{dB_{\mu,n}} = \frac{d\Pi_i}{dB_{\mu,\kappa}} \frac{\mathbb{P}(N = n)}{\mathbb{P}(K = n)} \mathbf{1}_{\{N=n\}}.$$

Hence,

$$(3.8) \quad \mathcal{H}(\Pi_i^n | B_{\mu,n}) \mathbb{P}(K = n) = \mathbb{E}[\log f_i(\xi_i) | N = n] \mathbb{P}(N = n) - \mathbb{P}(K = n) \log \frac{\mathbb{P}(K = n)}{\mathbb{P}(N = n)}.$$

Summing the previous expression for  $n \in \mathbb{N}$  yields (3.5).  $\square$

### 3.3 Universal Marton-type inequalities for Poisson point processes

#### 3.3.1 Deterministic size case

Recall the definition of the partial transport cost  $\mathcal{T}_{\rho,0}$  introduced at the end of Section 2.2.4.

**Theorem 3.3.** *Let  $n \in \mathbb{N}$ ,  $\mu \in \mathcal{P}(Z)$ . Assume that  $\mu$  satisfies  $(\tilde{\mathbf{T}}_{\alpha,\rho})$  on  $Z$  for  $a_1, a_2 > 0$  and for some lower semi-continuous cost function  $\rho: Z \times Z \rightarrow [0, \infty]$  and  $\alpha: [0, \infty] \rightarrow [0, \infty]$  convex. Then,  $B_{\mu,n}$  satisfies the transport-entropy inequality  $(\mathbf{T}_c)$  (with the same constants) on  $\mathcal{M}_n(Z)$  with the cost function  $c: \mathcal{M}_n(Z) \times \mathcal{P}(\mathcal{M}_n(Z)) \rightarrow \mathbb{R}_+$  defined, for all  $\xi \in \mathcal{M}_n(Z)$  and  $\Pi \in \mathcal{P}(\mathcal{M}_n(Z))$ , by*

$$(3.9) \quad c(\xi, \Pi) = \int \alpha \left( \frac{1}{\xi(x)} \int \mathcal{T}_{\rho,0}(\chi, \xi(x)\delta_x) \Pi(d\chi) \right) \xi(dx).$$

Before proving Theorem 3.3, let us highlight a particular case. Thanks to the universal transport inequalities by Marton and Dembo recalled in Section 2.3.1, binomial processes always satisfy some transport-entropy inequality, as shown in the following result.

**Corollary 3.4.** *Let  $t \in [0, 1]$  and let  $\alpha_t$  be the function defined by (2.16). Let  $\mu \in \mathcal{P}(Z)$  and  $n \in \mathbb{N}$ . Then,  $B_{\mu,n}$  satisfies a transport-entropy inequality  $(\mathbf{T}_c)$ , namely*

$$\mathcal{T}_{c_t}(\Pi_2 | \Pi_1) \leq \frac{1}{t} \mathcal{H}(\Pi_1 | B_{\mu,n}) + \frac{1}{1-t} \mathcal{H}(\Pi_2 | B_{\mu,n}),$$

where, for all  $\xi \in \mathcal{M}_n(Z)$  and  $\Pi \in \mathcal{P}(\mathcal{M}_n(Z))$ :

$$(3.10) \quad c_t(\xi, \Pi) = \int \alpha_t \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx).$$

Since  $\alpha_{\frac{1}{2}}(u) \geq \frac{u^2}{2}$ ,  $u \in [0, 1]$ , when  $t = \frac{1}{2}$ , the same conclusion holds with  $\frac{u^2}{2}$  replacing  $\alpha_t$  in (3.10).

*Proof of Corollary 3.4.* According to Section 2.3.1, the probability measure  $\mu$  satisfies the inequality  $(\tilde{\mathbf{T}}_{\alpha_t, d_H})$  with  $\alpha_t$  given by (2.16) and the constants  $a_1 = \frac{1}{t}$  and  $a_2 = \frac{1}{1-t}$ . Moreover, if  $\chi, \xi \in \mathcal{M}_n(Z)$ , then using that  $\xi(x) \leq \xi(Z) = \chi(Z)$ , we find that

$$(3.11) \quad \mathcal{T}_{d_H,0}(\chi, \xi(x)\delta_x) = (\xi(x) - \chi(x))_+.$$

Applying Theorem 3.3 concludes the proof.  $\square$

*Proof of Theorem 3.3.* We follow the proof of Proposition 3.2. According to Proposition 2.7, the probability measure  $\mu^{\otimes n}$  satisfies the transport inequality  $(\mathbf{T}_c)$  on  $Z^n$  with constants  $a_1, a_2$  with respect to the cost function  $c$  defined by

$$(3.12) \quad c(x, p) = \sum_{i=1}^n \alpha \left( \int \rho(x_i, y) p_i(dy) \right), \quad x \in Z^n, p \in \mathcal{P}(Z^n).$$

The probability measure  $B_{\mu,n}$  is the push-forward of  $\mu^{\otimes n}$  by the map  $S^n$  defined by (3.3). Therefore, according to Proposition 2.8, we see that  $B_{\mu,n}$  satisfies the transport-entropy inequality with the cost function  $\bar{c}$  defined for all  $(\xi, \Pi) \in \mathcal{M}_n(Z) \times \mathcal{P}(\mathcal{M}_n(Z))$  by

$$(3.13) \quad \bar{c}(\xi, \Pi) = \inf \{c(x, p) : (x, p) \text{ such that } S^n(x) = \xi, S^n_\# p = \Pi\}.$$

In other words, if  $\xi = \sum_{i=1}^n \delta_{a_i}$ , then

$$(3.14) \quad \bar{c}(\xi, \Pi) = \inf \left\{ \sum_{i=1}^n \alpha(\mathbb{E}[\rho(Y_i, x_i)]) \right\},$$

where the infimum runs over all random vectors  $Y = (Y_1, \dots, Y_n)$  such that  $\sum_{i=1}^n \delta_{Y_i} \sim \Pi$  (whose existence is given by [23, Proposition 6.3]) and all  $x = (x_1, \dots, x_n) \in Z^n$  such that  $\xi = \sum_{i=1}^n \delta_{x_i}$ . The constraint  $\xi = \sum_{i=1}^n \delta_{x_i}$  determines the  $x_i$ 's up to permutation. Since the other constraint  $\sum_{i=1}^n \delta_{Y_i} \sim \Pi$  is permutation invariant, one can assume without loss of generality that  $a_i = x_i$  for all  $i \in \{1, \dots, n\}$ . In this way, we obtain that

$$(3.15) \quad \bar{c}(\xi, \Pi) = \inf \left\{ \sum_{i=1}^n \alpha(\mathbb{E}[\rho(Y_i, a_i)]) : \sum_{i=1}^n \delta_{Y_i} \sim \Pi \right\}.$$

Let us prove that  $\bar{c}$  is bounded from below by the cost given in (3.9). For all  $a \in Z$ , such that  $\xi(a) > 0$ , we define  $I(a) = \{i : a_i = a\}$ . Then, letting  $\chi = \sum_{i=1}^n \delta_{Y_i}$ , we obtain

$$\begin{aligned} \sum_{i=1}^n \alpha(\mathbb{E}[\rho(Y_i, a_i)]) &= \sum_{a:\xi(a)>0} \xi(a) \frac{\sum_{i \in I(a)} \alpha(\mathbb{E}[\rho(Y_i, a_i)])}{\xi(a)} \\ &\geq \sum_{a:\xi(a)>0} \xi(a) \alpha \left( \mathbb{E} \left[ \frac{\sum_{i \in I(a)} \rho(Y_i, a)}{\xi(a)} \right] \right) \\ &\geq \sum_{a:\xi(a)>0} \xi(a) \alpha \left( \mathbb{E} \left[ \frac{\mathcal{T}_{\rho,0}(\chi, \xi(a)\delta_a)}{\xi(a)} \right] \right), \end{aligned}$$

where the first inequality comes from Jensen inequality and the second by definition of (2.14). This completes the proof.  $\square$

### 3.3.2 Poisson case

First, a process level ‘‘binomial to Poisson argument’’, adapted from [31], allows us to extend the conclusion of Theorem 3.3 to Poisson processes with finite intensity measure. Then we will use another approximation argument to pass from the case of finite intensity measure to the case of  $\sigma$ -finite intensity measure.

**Theorem 3.5.** *Let  $\alpha_t$  be the function defined by (2.16). Let  $\nu$  be a  $\sigma$ -finite Radon measure on  $Z$ . Then  $\Pi_\nu$  satisfies a transport-entropy inequality ( $\mathbf{T}_c$ ), namely*

$$\mathcal{T}_{c_t}(\Pi_2 | \Pi_1) \leq \frac{1}{t} \mathcal{H}(\Pi_1 | \Pi_\nu) + \frac{1}{1-t} \mathcal{H}(\Pi_2 | \Pi_\nu),$$

where the cost function  $c_t: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \rightarrow [0, \infty]$  is defined by

$$(3.16) \quad c_t(\xi, \Pi) = \int \alpha_t \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx), \quad \xi \in \mathcal{M}_{\mathbb{N}}(Z), \Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$$

In particular, this result gives Theorem 1.2.



*Proof.* We first assume that there exists  $\mu \in \mathcal{P}(Z)$  and  $\lambda > 0$  such that  $\nu = \lambda\mu$ . For  $\mu \in \mathcal{P}(Z)$ ,  $n \in \mathbb{N}$ , and  $t \in [0, 1]$ , the  $t$ -thinning of  $B_{\mu,n}$  is the law denoted by  $B_{\mu,n}^{(t)}$  of the point process obtained by removing each point of  $\eta \sim B_{\mu,n}$  with probability  $(1-t)$  independently of the others. Formally, this family of point processes can be realized for instance as follows: for  $t \in [0, 1]$ , set

$$\tilde{\eta}^{(t)} = \sum_{i=1}^n \varepsilon_i \delta_{X_i}$$

with  $X_1, \dots, X_n$  i.i.d of law  $\mu$  and  $\varepsilon_1, \dots, \varepsilon_n$  i.i.d with a Bernoulli distribution of parameter  $t \in [0, 1]$ . Then  $B_{\mu,n}^{(t)}$  is by definition the law of  $\tilde{\eta}^{(t)}$ . We conveniently use a slightly different construction taken from [31]. Let us add a cemetery point  $\infty$  to  $Z$  by considering  $\hat{Z} = Z \cup \{\infty\}$  where  $\infty \notin Z$  and  $\{\infty\}$  is isolated. Define

$$\hat{\eta}^{(t)} = \sum_{i=1}^n \delta_{\hat{X}_i^{(t)}},$$

where  $\hat{X}_1^{(t)}, \dots, \hat{X}_n^{(t)}$  are i.i.d random variables with law  $t\mu + (1-t)\delta_\infty$ . Defining

$$\eta^{(t)} = \hat{\eta}^{(t)}|_Z = \sum_{i=1}^n \mathbf{1}_{\{\hat{X}_i^{(t)} \neq \infty\}} \delta_{\hat{X}_i^{(t)}},$$

we then easily see that  $\eta^{(t)}$  and  $\tilde{\eta}^{(t)}$  have the same law.

Let  $\lambda > 0$ , and  $n$  sufficiently large so that  $\lambda_n := \lambda/n \leq 1$ ; we write  $\Pi_n = B_{\mu,n}^{(\lambda_n)}$ . This sequence  $\Pi_n$  converges to  $\Pi_\nu$ , the law of the Poisson process of intensity measure  $\nu := \lambda\mu$ . Indeed, according to [22, Theorem 4.34], one gets that for all Borel set  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$ :

$$\Pi_n(A) \xrightarrow[n \rightarrow \infty]{} \Pi_\nu(A).$$

Since according to Lemma 3.6 below, the cost function  $c$  is lower semi-continuous, in view of the previous convergence and Lemma 2.9, to conclude the proof in the case of finite intensity measure, we simply establish a transport-entropy inequality for  $\Pi_n$  ( $n \in \mathbb{N}$ ). Let  $\hat{\mu}_n = \lambda_n\mu + (1 - \lambda_n)\delta_\infty$ ; according to Corollary 3.4, we see that  $B_{\hat{\mu}_n,n} \in \mathcal{P}(\mathcal{M}_n(\hat{Z}))$  satisfies the transport-entropy inequality with constants  $a_1 = \frac{1}{t}$  and  $a_2 = \frac{1}{1-t}$ , and with cost function  $\hat{c}: \mathcal{M}_n(\hat{Z}) \times \mathcal{P}(\mathcal{M}_n(\hat{Z})) \rightarrow \mathbb{R}_+$  defined by

$$\hat{c}(\hat{\xi}, \hat{\Pi}) = \int \alpha_t \left( \int \left[ 1 - \frac{\hat{\chi}(x)}{\hat{\xi}(x)} \right]_+ \hat{\Pi}(d\hat{\chi}) \right) \hat{\xi}(dx), \quad \hat{\xi} \in \mathcal{M}_n(\hat{Z}), \hat{\Pi} \in \mathcal{P}(\mathcal{M}_n(\hat{Z})),$$

where  $\alpha_t$  is the function defined in (2.16). By construction,  $\Pi_n$  is the push-forward of  $B_{\hat{\mu}_n,n}$  under the map  $S: \mathcal{M}_n(\hat{Z}) \rightarrow \mathcal{M}_{\mathbb{N}}(Z)$  defined, for all  $\hat{\xi} = \sum_{i=1}^n \delta_{\hat{x}_i}$ ,  $\hat{x}_1, \dots, \hat{x}_n \in \hat{Z}$ , by

$$S(\hat{\xi}) = \hat{\xi}|_Z = \sum_{i=1}^n \mathbf{1}_{\{\hat{x}_i \in Z\}} \delta_{\hat{x}_i}.$$

According to Proposition 2.8, it follows that  $\Pi_n$  satisfies the transport-entropy inequality with the same constants and with the cost function  $\bar{c}: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \rightarrow \mathbb{R}_+$  defined by

$$\bar{c}(\xi, \Pi) = \inf \left\{ c(\hat{\xi}, \hat{\Pi}) : S(\hat{\xi}) = \xi, S_\# \hat{\Pi} = \Pi \right\}, \quad \xi \in \mathcal{M}_{\mathbb{N}}(Z), \Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)),$$

Now, if  $S(\hat{\xi}) = \xi$  and  $S_{\#}\hat{\Pi} = \Pi$ , then

$$\begin{aligned} \hat{c}(\hat{\xi}, \hat{\Pi}) &= \int_{\hat{Z}} \alpha_t \left( \int \left[ 1 - \frac{\hat{\chi}(x)}{\hat{\xi}(x)} \right]_+ \hat{\Pi}(d\hat{\chi}) \right) \hat{\xi}(dx) \\ &\geq \int_Z \alpha_t \left( \int \left[ 1 - \frac{S(\hat{\chi})(x)}{\xi(x)} \right]_+ \hat{\Pi}(d\hat{\chi}) \right) \xi(dx) \\ &= \int_Z \alpha_t \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx) \\ &= c(\xi, \Pi), \end{aligned}$$

where, for short, we write  $c = c_t$  for the cost function defined in (3.16). This concludes the proof for Poisson point process with finite intensity measure.

Now, we assume that  $\nu$  is  $\sigma$ -finite, and we let  $(Z_n)$  be an increasing measurable exhaustion of  $Z$  in measurable sets of finite measure. Let  $\Pi_n = \Pi_{\nu_n}$  where  $\nu_n = \nu(\cdot \cap Z_n)$ . We have that for all  $m \in \mathbb{N}$ ,

$$\nu_n(\cdot \cap Z_m) \xrightarrow[n \rightarrow \infty]{TV} \nu(\cdot \cap Z_m).$$

According to [22, Theorem 4.33], we find that  $\Pi_n \xrightarrow[n \rightarrow \infty]{TV} \Pi$ . By Lemma 3.6,  $c_t$  is lower semi-continuous, and by the previous analysis  $\Pi_n$  satisfies  $(\mathbf{T}_c)$  for all  $n \in \mathbb{N}$ . Thus, by the strong convergence of  $\Pi_n$  to  $\Pi$  and Lemma 2.9, we find that  $\Pi_{\nu}$  satisfies  $(\mathbf{T}_c)$ .  $\square$

Now we turn to the proof of the lower semi-continuity of  $c$ .

**Lemma 3.6.** *Let  $\alpha: \mathbb{R} \rightarrow [0, \infty]$  be a any lower semi-continuous function. The function  $c: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  defined by*

$$c(\xi, \Pi) = \int \alpha \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx)$$

*is lower semi-continuous. Moreover, the restriction of  $c$  to  $\mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  is also lower semi-continuous.*

*Proof.* We will first show that  $c: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \rightarrow [0, \infty]$  is lower semi-continuous and then show that  $c: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \rightarrow [0, \infty]$  is lower semi-continuous. Let  $\{\xi_n; n \in \mathbb{N}\} \cup \{\xi\} \subset \mathcal{M}_{\mathbb{N}}(Z)$  and  $\{\Pi_n; n \in \mathbb{N}\} \cup \{\Pi\} \subset \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  such that, as  $n \rightarrow \infty$ ,  $(\xi_n)$  converges weakly to  $\xi$ , and  $(\Pi_n)$  converges weakly (with respect to the weak topology on  $\mathcal{M}_{\mathbb{N}}(Z)$ ) to  $\Pi$ . Let us show that  $\liminf_{n \rightarrow +\infty} c(\xi_n, \Pi_n) \geq c(\xi, \Pi)$ . By definition there exists  $q \in \mathbb{N}$  and  $x_1, \dots, x_q \in Z$  such that  $\xi = \sum_{i=1}^q \delta_{x_i}$ . According to the proof of Lemma 2.1, we can without loss of generality assume that  $\xi_n = \sum_{i=1}^q \delta_{x_i^n}$  where  $x_i^n \rightarrow x_i$  when  $n \rightarrow \infty$ . Let  $p$  be the number of points in the support of  $\xi$  and write  $\text{Supp}(\xi) = \{a_1, \dots, a_p\}$  (with therefore pairwise distinct  $a_i$ 's) and denote by  $k_j = \xi(\{a_j\})$ , for all  $1 \leq j \leq p$ . For all  $j \in \{1, \dots, p\}$  denote by  $I_j = \{i : x_i = a_j\}$  and by  $S_j^n = \{x_i^n : i \in I_j\}$ . By Jensen's inequality, and the fact that  $\alpha$  and  $[0, 1] \ni u \mapsto [1 - u]_+$  are convex, we find that

$$\xi(a_j) \sum_{z \in S_j^n} \alpha \left( \int \left[ 1 - \frac{\chi(z)}{\xi_n(z)} \right]_+ \Pi_n(d\chi) \right) \frac{\xi_n(z)}{\xi(a_j)} \geq \xi(a_j) \alpha \left( \int \left[ 1 - \frac{\sum_{z \in S_j^n} \chi(z)}{\xi(a_j)} \right]_+ \Pi_n(d\chi) \right).$$

Fix some  $\varepsilon > 0$ ; for  $n$  large enough, one has  $S_j^n \subset \bar{B}(a_j, \varepsilon)$  for all  $1 \leq j \leq p$ , where  $\bar{B}(a, \varepsilon)$  denotes the closed ball of center  $a$  and radius  $\varepsilon$ . Since the points of  $S_j^n$  are all distinct, we find  $\sum_{z \in S_j^n} \chi(z) = \chi(S_j^n) \leq \chi(\bar{B}(a_j, \varepsilon))$ . Hence, by the monotonicity of  $\alpha$ :

$$\xi(a_j) \alpha \left( \int \left[ 1 - \frac{\sum_{z \in S_j^n} \chi(z)}{k_j} \right]_+ \Pi_n(d\chi) \right) \geq \xi(a_j) \alpha \left( \int \left[ 1 - \frac{\chi(\bar{B}(a_j, \varepsilon))}{k_j} \right]_+ \Pi_n(d\chi) \right).$$

Since  $\xi_n = \sum_{j=1}^p \sum_{z \in S_n^j} \xi_n(z) \delta_z$ , summing the two previous estimates over  $j \in [p]$  gives

$$c(\xi_n, \Pi_n) \geq \int \alpha \left( \int \left[ 1 - \frac{\chi(\bar{B}(a, \varepsilon))}{\xi(a)} \right]_+ \Pi_n(d\chi) \right) d\xi(a).$$

According to Portmanteau theorem, since  $\bar{B}(a, \varepsilon)$  is a closed set, the map  $\mathcal{M}_{\mathbb{N}}(Z) \ni \chi \mapsto \chi(\bar{B}(a, \varepsilon))$  is upper-semicontinuous (for the weak topology on  $\mathcal{M}_{\mathbb{N}}(Z)$ ); hence, the lower semi-continuity (for the weak topology on  $\mathcal{M}_{\mathbb{N}}(Z)$ ) of the map

$$\mathcal{M}_{\mathbb{N}}(Z) \ni \chi \mapsto \left[ 1 - \frac{\chi(\bar{B}(a, \varepsilon))}{\xi(a)} \right]_+.$$

Therefore, using again the Portmanteau theorem, we obtain, for any fixed  $a$ , the lower semi-continuity of the map

$$(3.17) \quad \Pi \mapsto \int \left[ 1 - \frac{\chi(\bar{B}(a, \varepsilon))}{\xi(a)} \right]_+ \Pi(d\chi).$$

Therefore, letting  $n \rightarrow \infty$ , and using that  $\liminf a_n + b_n \geq \liminf a_n + \liminf b_n$  yield

$$\liminf_{n \rightarrow +\infty} c(\xi_n, \Pi_n) \geq \int \alpha \left( \int \left[ 1 - \frac{\chi(\bar{B}(a, \varepsilon))}{\xi(a)} \right]_+ \Pi(d\chi) \right) d\xi(a).$$

Finally, letting  $\varepsilon \rightarrow 0$  and using the dominated convergence theorem, we have that

$$\chi(\bar{B}(a, \varepsilon)) \xrightarrow{\varepsilon \rightarrow 0} \chi(a).$$

Using again the dominated convergence theorem, we thus find that

$$\int \left[ 1 - \frac{\chi(\bar{B}(a, \varepsilon))}{\xi(a)} \right]_+ \Pi(d\chi) \xrightarrow{\varepsilon \rightarrow 0} \int \left[ 1 - \frac{\chi(a)}{\xi(a)} \right]_+ \Pi(d\chi).$$

Since  $\alpha$  is lower semi-continuous, we finally find

$$\liminf_{n \rightarrow +\infty} c(\xi_n, \Pi_n) \geq c(\xi, \Pi).$$

This proves that  $c: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{M}_{\mathbb{N}}(Z) \rightarrow [0, \infty]$  is weakly lower semi-continuous. Now let us consider

$$\xi_n \xrightarrow[n \rightarrow \infty]{\mathcal{M}_{\mathbb{N}}(Z)} \xi; \quad \text{and} \quad \Pi_n \xrightarrow[n \rightarrow \infty]{\mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))} \Pi.$$

Fix  $o \in Z$ , and, for  $L > 0$ , write  $B_L = B(o, L)$ . By definition of the vague convergence, for all  $L > 0$ ,

$$\xi_{n|B_L} \xrightarrow[n \rightarrow \infty]{\text{weakly}} \xi|_{B_L}.$$

By [22, Lemma 4.1],  $\mathcal{M}_{\mathbb{N}}(Z) \ni \chi \mapsto \chi(\bar{B}(a, \varepsilon))$  is upper semi-continuous (for the vague topology on  $\mathcal{M}_{\mathbb{N}}(Z)$ ); hence, using again Portmanteau theorem, one gets that for any fixed  $a$ , the map  $\mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \ni \Pi \mapsto \int \left[ 1 - \frac{\chi(\bar{B}(a, \varepsilon))}{\xi(a)} \right]_+ \Pi(d\chi)$  is lower-semicontinuous. Thus, arguing as before, we find that:

$$\liminf_{n \rightarrow \infty} c(\xi_{n|B_L}, \Pi_n) \geq \int \alpha \left( \int \left[ 1 - \frac{\chi(a)}{\xi(a)} \right]_+ \Pi(d\chi) \right) 1_{B_L}(a) \xi(da).$$

On the one hand, since  $\xi_n \geq \xi_{n|B_L}$ , we have that  $c(\xi_n, \Pi_n) \geq c(\xi_{n|B_L}, \Pi_n)$ . On the other hand, by the monotone convergence theorem, we find that

$$\int \alpha \left( \int \left[ 1 - \frac{\chi(a)}{\xi(a)} \right]_+ \Pi(d\chi) \right) 1_{B_L}(a) \xi(da) \xrightarrow{L \rightarrow \infty} c(\xi, \Pi).$$

Combining these observations, we obtain that

$$\liminf_{n \rightarrow \infty} c(\xi_n, \Pi_n) \geq c(\xi, \Pi).$$

This concludes the proof.  $\square$

## 4 Applications to concentration estimates

With the help of the transport-entropy inequalities proved in [Section 3](#), we derive concentration results for Poisson point processes and other mixed binomial processes. Below, we highlight some particular cases.

### 4.1 Generic concentration of measure inequalities

First let us draw some consequences of [Theorem 3.1](#).

**Corollary 4.1.** *Let  $\kappa \in \mathcal{P}(\mathbb{N})$ ,  $\mu \in \mathcal{P}(Z)$ , and  $\eta \sim B_{\mu, \kappa}$ . Assume  $\mu \in \mathcal{P}(Z)$  satisfies [\(T \$\_{\rho}\$ \)](#), for some lower semi-continuous cost function  $\rho: Z \times Z \rightarrow [0, \infty]$ , with constant  $a > 0$ . For all Borel sets  $A_1, A_2 \subset \mathcal{M}_{\mathbb{N}}(Z)$  such that  $\mathbb{P}(\eta \in A_i) > 0$  and  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2) < +\infty$ , where  $\Pi_i = \frac{1_{A_i}}{B_{\mu, \kappa}(A_i)} B_{\mu, \kappa}$ , it holds that*

$$\mathbb{P}(\eta \in A_1)\mathbb{P}(\eta \in A_2) \leq \exp\left(-\frac{1}{a} \inf_{\xi_1 \in A_1, \xi_2 \in A_2} \mathcal{T}_{\rho}(\xi_1, \xi_2)\right).$$

*Proof.* Simply apply the transport-entropy of [Theorem 3.1](#) to the probability measures  $\Pi_1$  and  $\Pi_2$  defined above and use that  $\mathcal{H}(\Pi_i | B_{\mu, \kappa}) = -\log \mathbb{P}(\eta \in A_i)$ ,  $i = 1, 2$ , and the lower bound  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2) \geq \inf_{\xi_1 \in A_1, \xi_2 \in A_2} \mathcal{T}_{\rho}(\xi_1, \xi_2)$ . Details of this argument can be found in e.g [\[13\]](#).  $\square$

Due to the requirement  $\mathbb{T}_{\rho}(\Pi_1, \Pi_2) < +\infty$ , this concentration bound seems delicate to manipulate.

Now let us examine the consequences of [Theorem 3.3](#) and [Theorem 3.5](#). To state them, let us recall some definitions from [\[21\]](#) adapted to our Poisson setting. Let  $c: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \rightarrow [0, \infty]$  be a cost. For a Borel set  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$ , we write

$$(4.1) \quad c_A(\xi) = \inf_{\Pi(A)=1} c(\xi, \Pi), \quad \forall \xi \in \mathcal{M}_{\mathbb{N}}(Z),$$

for the *Talagrand convex distance* associated with  $A$  and  $c$  (which, despite its name, is not really a distance), and

$$(4.2) \quad A_r = \{x \in \mathcal{M}_{\mathbb{N}}(Z) : c_A(x) \leq r\}, \quad r \geq 0,$$

for the *enlargement* of  $A$  with respect to the Talagrand convex distance.

With these notations, we state the following concentration result.

**Corollary 4.2.** *Let  $t \in (0, 1)$  and  $\alpha_t$  be the function defined in [\(2.16\)](#).*

(a) *Let  $\eta$  be a binomial process (of size  $n$ ). Then for every Borel set  $A \subset \mathcal{M}_n(Z)$*

$$\mathbb{P}(\eta \in A)^{\frac{1}{t}} \mathbb{P}(\eta \notin A_r)^{\frac{1}{1-t}} \leq e^{-r}, \quad \forall r \geq 0,$$

where  $A_r$  is the enlargement of  $A$  (in  $\mathcal{M}_n(Z)$ ) given in [\(4.2\)](#) for the choice of  $c = c_t$  given in [\(3.10\)](#), that is

$$c(\xi, \Pi) = \int \alpha_t \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx), \quad \xi \in \mathcal{M}_n(Z), \Pi \in \mathcal{P}(\mathcal{M}_n(Z)).$$

(b) *Let  $\eta$  be a Poisson point process with finite intensity measure  $\nu \in \mathcal{M}_b(Z)$ . Then for every Borel set  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$*

$$\mathbb{P}(\eta \in A)^{\frac{1}{t}} \mathbb{P}(\eta \notin A_r)^{\frac{1}{1-t}} \leq e^{-r}, \quad \forall r \geq 0,$$

where  $A_r$  is the enlargement of  $A$  (in  $\mathcal{M}_{\mathbb{N}}(Z)$ ) given in [\(4.2\)](#) for the choice of  $c = c_t$  given in [\(3.16\)](#), that is

$$c(\xi, \Pi) = \int \alpha_t \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx) \quad \xi \in \mathcal{M}_{\mathbb{N}}(Z), \Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)).$$

(c) Let  $\eta$  be a Poisson point process with  $\sigma$ -finite intensity measure  $\nu$ . Then for every Borel set  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$

$$\mathbb{P}(\eta \in A)^{\frac{1}{t}} \mathbb{P}(\eta \notin A_r)^{\frac{1}{1-t}} \leq e^{-r}, \quad \forall r \geq 0,$$

where  $A_r$  is the enlargement of  $A$  (in  $\mathcal{M}_{\mathbb{N}}(Z)$ ) given in (4.2) for the choice of  $c = c_t$  given in (3.16), that is

$$c(\xi, \Pi) = \int \alpha_t \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx) \quad \xi \in \mathcal{M}_{\mathbb{N}}(Z), \Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)).$$

*Proof.* Combine [21, Theorem 5.1] with Corollary 3.4 or Theorem 3.5.  $\square$

For further discussions, let us first recall some definitions and state a theorem of [31] to be compared with Corollary 4.2. Given a point measure  $\xi \in \mathcal{M}_{\mathbb{N}}(Z)$ , we write  $\mathcal{L}^2(\xi)$  for the space of functions  $g: Z \rightarrow \mathbb{R}$  such that  $g(x) = 0$  if  $\xi(x) = 0$  and such that:

$$|g|_{\mathcal{L}^2(\xi)}^2 := \sum_{x \in \xi} \xi(x) g(x)^2 < \infty,$$

where the notation  $x \in \xi$  is shorthand for  $x \in \text{supp } \xi$ . When equipped with the norm  $|\cdot|_{\mathcal{L}^2(\xi)}$ , the linear space  $\mathcal{L}^2(\xi)$  is a Hilbert space. Given another  $\chi \in \mathcal{M}_{\mathbb{N}}(Z)$ , we write  $\xi \setminus \chi$  for the point measure given by  $\sum_{x \in \xi} (\xi(x) - \chi(x))_+ \delta_x$ . For  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$  and  $\xi \in \mathcal{M}_{\mathbb{N}}(Z)$ , following [31] ([31] only works on  $\mathcal{M}_{\mathbb{N}}(Z)$ ), we define

$$(4.3) \quad d_A(\xi) = \sup_{|g|_{\mathcal{L}^2(\xi)} \leq 1} \inf_{\chi \in A} \int g d(\xi \setminus \chi),$$

where the supremum runs over non-negative  $g$  only, and we also define

$$A_r^d = \{\xi \in \mathcal{M}_{\mathbb{N}}(Z) : d_A(\xi) \leq r\}, \quad r \geq 0.$$

Then, we have the following extension of [31, Theorem 1.1].

**Theorem 4.3.** *Let  $\eta$  be a Poisson point process on  $Z$  with a  $\sigma$ -finite intensity. Then, for every Borel set  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$ ,*

$$\mathbb{P}(\eta \in A) \mathbb{P}(\eta \notin A_r^d) \leq e^{-\frac{r^2}{4}}.$$

The case of a Poisson point processes with finite intensity measure is proved established in [31], whereas the case of an intensity with infinite mass is new. The proof of Theorem 4.3 follows immediately from Corollary 4.2 (with  $t = 1/2$ ) and the following lemma (using the fact that  $\alpha_{1/2}(u) \geq u^2/2$ ).

**Lemma 4.4.** *With the notation introduced above, and the functions  $\alpha(u) = \frac{u^2}{2}$ ,  $u \geq 0$ , and*

$$c(\xi, \Pi) = \int \alpha \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx), \quad \xi \in \mathcal{M}_{\mathbb{N}}(Z), \Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)).$$

*For any  $A \subset \mathcal{M}_{\mathbb{N}}(Z)$ , it holds  $A_{2\sqrt{r}}^d = A_r$ ; equivalently  $c_A(\xi) = \frac{1}{2} d_A(\xi)^2$ , for all  $\xi \in \mathcal{M}_{\mathbb{N}}(Z)$ .*

*Proof.* The argument is quite classical and goes back to Talagrand. We give a proof for completeness. First, we recall this well-known duality formula on Hilbert spaces: if  $\mathcal{H}$  is an Hilbert space with inner product  $\langle \cdot, \cdot \rangle$  and induced norm  $|\cdot|$ , we have

$$|x| = \sup_{|y| \leq 1} \langle x, y \rangle.$$

Second, we recall this well-known fact about randomization of infimum

$$\inf_{x \in A} f(x) = \inf_{X \in A} \mathbb{E}[f(X)],$$

where the second infimum runs on random variables  $X$  taking values in  $A$ .

Third, for two convex spaces  $C_1$  and  $C_2$ , consider a bilinear functional  $\Lambda: C_1 \times C_2 \rightarrow \mathbb{R}$  that is upper semi-continuous and concave in its first variable, and lower semi-continuous and convex in its second variable. If  $C_1$  is compact, then by the min-max theorem [36, Corollary 3.3]:

$$\sup_{C_1} \inf_{C_2} \Lambda = \inf_{C_2} \sup_{C_1} \Lambda.$$

We equip  $\mathcal{L}^2(\xi)$  with the weak topology  $\sigma(\mathcal{L}^2(\xi), \mathcal{L}^2(\xi))$  defined by duality, and we write  $\mathcal{L}_+^2(\xi)$  for the convex cone of non-negative  $g \in \mathcal{L}^2(\xi)$ . We write  $\mathcal{P}_\xi$  for the convex set of all  $\Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  such that

$$\left( Z \ni x \mapsto \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \in \mathcal{L}_+^2(\xi).$$

We have that

$$\mathcal{P}_\xi = \{ \Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) : c(\xi, \Pi) < \infty \}.$$

We now define the function  $\Lambda: \mathcal{L}_+^2(\xi) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \rightarrow [0, \infty]$  by

$$\Lambda(g, \Pi) = \int g d(\xi \setminus \chi) \Pi(d\chi) = \int g(x) \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \xi(dx), \quad \text{if } \Pi \in \mathcal{P}_\xi,$$

and  $\Lambda(g, \Pi) = \infty$  otherwise. We claim that  $\Lambda$  is upper semi-continuous and concave in its first variable. The concavity is immediate. Let us show the upper semi-continuity. Let  $\Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z))$  and  $\{g_n; n \in \mathbb{N}\} \cup \{g\} \subset \mathcal{L}_+^2(\xi)$  such that:

$$g_n \xrightarrow[n \rightarrow \infty]{\sigma(\mathcal{L}^2(\xi), \mathcal{L}^2(\xi))} g,$$

Without loss of generality, we can assume that  $\Pi \in \mathcal{P}_\xi$  (otherwise  $\Lambda(g, \Pi) = \infty$  and there is nothing to prove). From the definition of the  $\sigma(\mathcal{L}^2(\xi), \mathcal{L}^2(\xi))$  convergence, and the fact that  $\Pi \in \mathcal{P}_\xi$ , we find

$$\limsup \Lambda(g_n, \Pi) = \lim \Lambda(g_n, \Pi) = \Lambda(g, \Pi).$$

We now claim that  $\Lambda$  is lower semi-continuous and convex in its second variable. The convexity is also immediate. By [22, Theorem 4.1], for all  $x \in \xi$  and  $g \in \mathcal{L}_+^2(\xi)$ , the function

$$\mathcal{M}_{\mathbb{N}}(Z) \ni \chi \mapsto g(x)(\xi(x) - \chi(x))_+,$$

is lower semi-continuous; hence by Portmanteau's theorem:

$$\mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \ni \Pi \mapsto \int g(x)(\xi(x) - \chi(x))_+ \Pi(d\chi),$$

is lower semi-continuous. By Tonelli's theorem, we find that

$$\Lambda(g, \Pi) = \sum_{x \in \xi} \int g(x)(\xi(x) - \chi(x))_+ \Pi(d\chi).$$

We conclude since a sum of lower semi-continuous functions is again lower semi-continuous. By the Banach-Alaoglu theorem,  $\mathcal{B}$ , the unit ball of  $\mathcal{L}^2(\xi)$ , is weakly-\* compact; hence, since  $\mathcal{L}^2(\xi)$  is reflexive,  $\mathcal{B}$  is weakly compact. Let us also write  $\mathcal{B}_+$  for  $\mathcal{B} \cap \mathcal{L}_+^2(\xi)$ . We can therefore apply the randomization of infimum, the min-max theorem (remark that  $\{ \Pi \in \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) : \Pi(A) = 1 \}$  is



convex), and the duality formula for the norm in Hilbert spaces (we recall these three arguments at the beginning of the proof) to obtain that

$$\begin{aligned}
d_A(\xi) &= \sup_{g \in \mathcal{B}_+} \inf_{\chi \in A} \int g d(\xi \setminus \chi) \\
&= \sup_{g \in \mathcal{B}_+} \inf_{\Pi(A)=1} \int g(x) \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx) \\
&= \inf_{\Pi(A)=1} \sup_{g \in \mathcal{B}_+} \int g(x) \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx) \\
&= \inf_{\Pi(A)=1} \left( \int \left( \int \left[ 1 - \frac{\nu(x)}{\xi(x)} \right]_+ \Pi(d\nu) \right)^2 \xi(dx) \right)^{1/2} \\
&= (2c_A(\xi))^{1/2}.
\end{aligned}$$

□

## 4.2 Concentration of measure for convex functionals

On  $\mathbb{R}^p$ , Marton's inequality implies a deviation inequality for convex 1-Lipschitz functions (see (1.9)). The goal of this section is twofold:

- (i) Introduce a reasonable notion for convex and Lipschitz functions on  $\mathcal{M}_{\mathbb{N}}(Z)$  and show that those functions also satisfy a deviation estimate.
- (ii) Show how this new deviation estimate sheds light on concentration results obtained for  $U$ -statistics by [3].

Despite the lack of a geodesic structure and of a metric structure on  $\mathcal{M}_{\mathbb{N}}(Z)$ , we define convex and Lipschitz functionals in a rather straightforward way thanks to an analogy with the euclidean case. Recall that for a differentiable function  $f: \mathbb{R}^p \rightarrow \mathbb{R}$ ,  $f$  is convex if and only if its Hessian is non-negative, while  $f$  is Lipschitz if and only if the norm of the gradient of  $f$  is bounded by a constant.

On the Poisson space, the *difference operators* provide an ersatz of a differentiable structure. For a functional  $F$  defined on  $\mathcal{M}_{\mathbb{N}}(Z)$ , we define these operators as follows:

$$\begin{aligned}
D_x^- F(\xi) &= F(\xi) - F(\xi - \delta_x), & \forall x \in \xi; \\
D_x^+ F(\xi) &= F(\xi + \delta_x) - F(\xi), & \forall x \in Z.
\end{aligned}$$

In the sequel, we show concentration of measure for *convex* functionals  $F: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}$ . By definition, we say that a functional  $F$  is *non-decreasing* if

$$(4.4) \quad \forall \xi, \chi \in \mathcal{M}_{\mathbb{N}}(Z), (\xi \leq \chi \implies F(\xi) \leq F(\chi)).$$

We can alternatively state this monotonicity property with the  $D^-$  or  $D^+$  operators:  $F$  is non-decreasing if and only if  $D_x^- F(\xi) \geq 0$  for all  $\xi \in \mathcal{M}_{\mathbb{N}}(Z)$  and all  $x \in \xi$  if and only if  $D_x^+ F(\xi) \geq 0$  for all  $\xi \in \mathcal{M}_{\mathbb{N}}(Z)$  and all  $x \in Z$ . We say that the functional  $F$  is *convex* whenever  $D_x^+ F$  is non-decreasing for all fixed  $x \in Z$ . Alternatively,  $F$  is convex if and only if its "second derivative" satisfies:  $D_{xy}^+ F(\xi) \geq 0$  for all  $\xi \in \mathcal{M}_{\mathbb{N}}(Z)$ , and all  $x$  and  $y \in Z$ . The class of non-decreasing convex functionals contains in particular the class of  $U$ -statistics with non-negative kernel. The  $U$ -statistics of order  $q \in \mathbb{N}_{>0}$  associated with a symmetric kernel  $h: Z^q \rightarrow \mathbb{R}_+$  is a functional of the form

$$F(\xi) = \sum_{\substack{\neq \\ i_1, \dots, i_q \leq n}} h(x_{i_1}, \dots, x_{i_q}) := \xi^{(q)}(h), \quad \forall \xi = \sum_{i=1}^n \delta_{x_i} \in \mathcal{M}_{\mathbb{N}}(Z),$$

where the superscript  $\neq$  indicates summation over pairwise disjoint indices. By Mecke's formula [23, Theorem 4.4], when  $h \in \mathcal{L}^1(\nu^q)$ , we have that  $\mathbb{E}|F(\eta)| \leq \nu^q(|h|) < \infty$ , and the random variable  $F(\eta)$  is almost surely finite. It is immediate that, for all  $x \in Z$ ,  $D_x^+ F(\xi) = q\xi^{(q-1)}(h(x, \cdot))$ . Thus, we see that when  $h \geq 0$ , then the  $U$ -statistics  $F$  is non-decreasing in the sense of (4.4). We also have that  $D_y^+ D_x^+ F(\xi) = q(q-1)\xi^{(q-2)}(h(x, y, \cdot))$ . So that, when  $h \geq 0$ , the  $U$ -statistics  $F$  is also convex in our sense. Similarly, we define a notion of being Lipschitz by imposing a bound on the square of the "norm of the gradient" which in our setting is given by  $\sum_{\eta} (D_x^- F(\eta))^2$ .

We now give a slight extension of a result of [3] to convex functionals and to Poisson point processes with intensity measure possibly having atoms.

**Theorem 4.5** ([3, Proposition 5.6]). *Suppose that  $\eta$  is a Poisson process with a  $\sigma$  finite intensity measure. Let  $F: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}_+$  be a non-negative convex functional such that there exist  $\delta > 0$  and  $\beta \in [0, 2)$  such that*

$$(4.5) \quad \sum_{x \in \eta} (D_x^- F(\eta))^2 \leq \delta F(\eta)^\beta \quad a.s.$$

Let  $m$  be a median of  $F$ . Then, for all  $r \geq 0$ , we have that

$$\begin{aligned} \mathbb{P}(F(\eta) \geq m + r) &\leq 2 \exp\left(-\frac{r^2}{4\delta(r+m)^\alpha}\right); \\ \mathbb{P}(F(\eta) < m - r) &\leq 2 \exp\left(-\frac{r^2}{4\delta m^\alpha}\right). \end{aligned}$$

The proof of the theorem goes simply by observing that [3, Lemma 5.7] still holds for convex functionals. The rest of the proof is similar and is omitted.

**Lemma 4.6.** *Let  $F: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}_+$  be a convex functional. Then, for all  $\xi$  and  $\chi \in \mathcal{M}_{\mathbb{N}}(Z)$ , such that  $\chi \leq \xi$ :*

$$F(\xi) - F(\chi) \leq \int D_x^- F(\xi) (\xi \setminus \chi)(dx).$$

*Proof.* Without loss of generality, we assume that  $\xi \neq 0$ . Since  $\chi \leq \xi$ , we can write  $\xi = \chi + \sum_{i=1}^n \delta_{x_i}$ , where  $n \in \mathbb{N} \cup \{+\infty\}$  and the  $x_i$ 's are elements of  $Z$ . Then, denoting by  $\chi_0 = \chi$  and for all  $1 \leq i \leq n-1$ ,  $\chi_{i+1} = \chi_i + \delta_{x_{i+1}}$ , it holds

$$\begin{aligned} F(\xi) - F(\chi) &= \sum_{i=0}^{n-1} F(\chi_{i+1}) - F(\chi_i) \\ &= \sum_{i=0}^{n-1} D_{x_{i+1}}^+ F(\chi_i) \leq \sum_{i=0}^{n-1} D_{x_{i+1}}^+ F(\xi - \delta_{x_{i+1}}) = \sum_{i=0}^{n-1} D_{x_{i+1}}^- F(\xi), \end{aligned}$$

where the inequality comes from the fact that  $D_x^+ F$  is non-decreasing.  $\square$

In [3], provided that  $F$  satisfies an inequality as of the one of Lemma 4.6, they give a lower bound on  $d_A(\eta)$  by some quantity depending on  $F$  only. Then, they use the deviation inequality of [31] (see Theorem 4.3) involving the convex distance  $d_A$  defined in (4.3) to conclude in the case of a finite intensity measure and use an approximation by thinning to conclude. Following, our previous results, we could follow an alternative approach rather based on the distance  $c_A$  defined in (4.1). The distance  $c_A$  intimately relates to the infimum-convolution operator  $R_c$  (see [21, Section 5] for a discussion of this link in the general case). We define  $R_c$  in the next section to study logarithmic Sobolev inequalities linked to our transport-entropy inequality. As shown by Lemma 5.3, we are only able to bound  $R_c$  for convex functions so that with the approach based on  $c_A$ , we could not improve upon Theorem 4.5.

## 5 Modified logarithmic Sobolev inequality

In this section we investigate the links between the transport inequality obtained in [Theorem 3.5](#) and modified logarithmic Sobolev inequalities on the Poisson space.

### 5.1 Introduction

On the Euclidean space (and more generally on a Riemannian manifold), a celebrated result by Otto & Villani [\[30\]](#) (see also [\[7\]](#)) shows that a probability measure  $\gamma \in \mathcal{P}(\mathbb{R}^d)$  satisfies the quadratic Talagrand's inequality

$$(5.1) \quad \mathcal{W}_2^2(\nu, \gamma) \leq C\mathcal{H}(\nu|\gamma), \quad \forall \nu$$

whenever  $\gamma$  satisfies the logarithmic Sobolev inequality: for all function  $f$  sufficiently smooth,

$$\text{Ent} \left( e^{f(X)} \right) \leq \frac{C}{4} \mathbb{E} \left[ e^{f(X)} |\nabla f(X)|^2 \right], \quad X \sim \gamma,$$

denoting

$$\text{Ent} \left( e^Z \right) = \mathbb{E}[Z e^Z] - \mathbb{E}[Z] \log \mathbb{E}[e^Z],$$

for any random variable  $Z$  such that  $\mathbb{E}[|Z|e^Z] < +\infty$ .

The work by [\[18\]](#) supplements the Otto-Villani theorem: it shows that  $\gamma$  satisfies [\(5.1\)](#) if and only if  $\gamma$  satisfies the following *restricted* logarithmic Sobolev inequality

$$\text{Ent}(e^{f(X)}) \leq C' \mathbb{E} \left[ |\nabla f(X)|^2 e^{f(X)} \right], \quad X \sim \gamma,$$

for all  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $x \mapsto f(x) + C''|x|^2$  is convex, where  $C', C'' > 0$  depend quantitatively on  $C$ . The reference [\[19\]](#) further investigates such an equivalence between transport inequalities and modified versions of the logarithmic Sobolev inequality, and proves that on a complete separable metric space  $(E, d)$ , a probability measure  $\gamma \in \mathcal{P}(E)$  satisfies [\(5.1\)](#) (where  $\mathcal{W}_2$  is defined with respect to the metric  $d$ ) if and only if there exists  $C', \lambda > 0$  (that can be precisely related to  $C$ ) such that, for all bounded measurable function  $f$

$$\text{Ent} \left( e^{f(X)} \right) \leq C' \mathbb{E} \left[ e^{f(X)} (f(X) - Q^\lambda f(X)) \right], \quad X \sim \gamma,$$

where  $Q^\lambda$  is the infimum convolution operator defined by

$$Q^\lambda f(x) = \inf_{y \in E} \left\{ f(y) + \lambda d(x, y)^2 \right\}, \quad x \in E.$$

We refer to [\[21, 34, 35\]](#) for other results connecting transport inequalities of the form [\(T<sub>c</sub>\)](#) and variants of the logarithmic Sobolev inequality.

Our goal in what follows is to do a first step in extending these results to the Poisson framework.

### 5.2 Transport inequalities and variants of the log-Sobolev inequality on the Poisson space

In [\[21\]](#), it is shown that a transport-entropy inequality of the form [\(T<sub>c</sub>\)](#), involving some cost function  $c: E \times \mathcal{P}(E)$  with  $E$  being an arbitrary Polish space, always implies some sort of logarithmic Sobolev inequality, whose energy term contains a "gradient" defined using the following infimum-convolution operator  $R_c$

$$R_c F(\xi) = \inf_{\Pi \in \mathcal{P}(E)} \left\{ \Pi(F) + c(\xi, \Pi) \right\}, \quad \forall \xi \in E,$$

for all  $F: E \rightarrow \mathbb{R}$ .

In our case, [Theorem 3.5](#) with  $t = 1$  (and  $E = \mathcal{M}_{\mathbb{N}}(Z)$ ) combined with [\[21, Theorem 3.8\]](#) immediately gives the following result.

**Theorem 5.1** (Infimum-convolution logarithmic Sobolev inequality). *Let  $\nu$  be a  $\sigma$ -finite measure on  $Z$  and let  $\eta \sim \Pi_\nu$ . Then, for all  $F: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}$  such that  $\mathbb{E}[|F|(\eta) e^{F(\eta)}] < \infty$ , we have, for all  $\lambda \in (0, 1)$ :*

$$(5.2) \quad \mathbb{E}_{\text{nt}} \left( e^{F(\eta)} \right) := \mathbb{E} \left[ e^{F(\eta)} F(\eta) \right] - \mathbb{E} [F(\eta)] \mathbb{E} \left[ e^{F(\eta)} \right] \leq \frac{1}{1-\lambda} \mathbb{E} \left[ (F(\eta) - R_{\lambda c_1} F(\eta)) e^{F(\eta)} \right]$$

where  $c_1$  is the cost given in (3.16) associated with  $\alpha_1$  given in (2.16).

We do not know if Inequality (5.2) implies back the transport inequality of Theorem 3.5 (for  $t = 1$ ).

In the following result, we deduce from Theorem 5.1 a modified logarithmic Sobolev inequality reminiscent of Wu's inequality [42] in restriction to the class of non-decreasing convex functions. For all  $\lambda \in (0, 1)$ , we will denote by

$$\phi_\lambda(s) = \frac{s}{1-\lambda} - \frac{\lambda}{1-\lambda} \log \left( 1 + \frac{s}{\lambda} \right), \quad s \geq 0.$$

We also consider

$$\phi_0(s) = \lim_{\lambda \rightarrow 0} \phi_\lambda(s) = s, \quad s \geq 0.$$

**Corollary 5.2.** *Let  $\nu$  be a  $\sigma$ -finite measure on  $Z$  with no atoms and let  $\eta \sim \Pi_\nu$ . Let  $F: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}$  be convex non-decreasing. Then, for all  $0 \leq \lambda < 1$ :*

$$\mathbb{E}_{\text{nt}} \left( e^{F(\eta)} \right) \leq \mathbb{E} \left[ e^{F(\eta)} \int \phi_\lambda(D_x^- F(\eta)) \eta(dx) \right].$$

As mentioned above, this inequality is close to the modified logarithmic Sobolev inequality by Wu [42]: for all  $F: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}$ ,

$$\mathbb{E}_{\text{nt}} \left( e^{F(\eta)} \right) \leq \mathbb{E} \left[ e^{F(\eta)} \int \phi_w(D_x^- F(\eta)) \eta(dx) \right],$$

where

$$\phi_w(s) = e^{-s} + s - 1, \quad s \geq 0.$$

The functions  $\phi_w(s)$  and  $\phi_\lambda(s)$ ,  $\lambda \in (0, 1)$ , are of the same order : quadratic for small values of  $s$  and linear for large values  $s$ . Nevertheless, one can observe that for  $s$  sufficiently large  $\phi_w(s) < \phi_\lambda(s)$ , so that our inequality does not improve upon the one by [42].

To prove Corollary 5.2, we will need the following lemma.

**Lemma 5.3.** *Let  $\nu$  be a  $\sigma$ -finite measure on  $Z$  with no atoms and let  $\eta \sim \Pi_\nu$ . Let  $\alpha: \mathbb{R} \rightarrow [0, \infty]$  be any convex lower semi-continuous function. Consider the function  $c: \mathcal{M}_{\mathbb{N}}(Z) \times \mathcal{P}(\mathcal{M}_{\mathbb{N}}(Z)) \rightarrow [0, \infty]$  defined by*

$$c(\xi, \Pi) = \int \alpha \left( \int \left[ 1 - \frac{\chi(x)}{\xi(x)} \right]_+ \Pi(d\chi) \right) \xi(dx).$$

For any non-decreasing convex  $F: \mathcal{M}_{\mathbb{N}}(Z) \rightarrow \mathbb{R}_+$ , we have that

$$(5.3) \quad F(\eta) - R_c F(\eta) \leq \int \alpha^*(D_x^- F(\eta)) \eta(dx) \quad \text{a.s.},$$

where  $\alpha^*(s) = \sup_u \{us - \alpha(u)\}$  is the Fenchel-Legendre conjugate of  $\alpha$ .

*Proof.* Since the intensity measure of  $\eta$  has no atoms, with probability one  $\eta$  is simple, i.e for all  $x \in Z$ ,  $\eta(x) \in \{0, 1\}$ . Let us consider  $\xi = \sum_{i=1}^n \delta_{x_i} \in \mathcal{M}(Z)$ , with pairwise distinct  $x'_i$  and  $n = \xi(Z) \in \mathbb{N} \cup \{\infty\}$ ; we have the following simplification:

$$c(\xi, \Pi) = \sum_{x \in \xi} \alpha(\mathbb{P}(\chi(x) = 0)).$$

Now we observe that

$$\sum_{x \in \xi} \alpha(\mathbb{P}(\chi(x) = 0)) = \sum_{x \in \xi} \alpha(\mathbb{P}((\chi \cap \xi)(x) = 0)),$$

where  $\chi \cap \xi = \sum_{x \in \xi} \min(\xi(x), \chi(x))\delta_x$ . Since  $\chi \geq \chi \cap \xi$  and  $F$  is non-decreasing, we have that  $F(\chi) \geq F(\chi \cap \xi)$ . Therefore, denoting by  $\Xi(\xi)$  the set of random variables  $\chi$  taking values in  $\mathcal{M}_{\mathbb{N}}(Z)$  such that  $\mathbb{P}(\chi \leq \xi) = 1$ , we get

$$R_c F(\xi) = \inf_{\chi \in \Xi(\xi)} \left\{ \mathbb{E}[F(\chi)] + \int \alpha(\mathbb{P}(\chi(x) = 0))\xi(dx) \right\}.$$

Invoking [Lemma 4.6](#), we thus find that

$$R_c F(\xi) - F(\xi) \geq \inf_{\chi \in \Xi(\xi)} \left\{ \int \alpha(\mathbb{P}(\chi(x) = 0))\xi(dx) - \mathbb{E} \left[ \int D_x^- F(\xi)(\xi \setminus \chi)(dx) \right] \right\}.$$

Now observe that a random measure  $\chi \in \Xi(\xi)$  if and only if it is of the form

$$\chi = \sum_{i=1}^n (1 - \varepsilon_i)\delta_{x_i},$$

for some Bernoulli random variables  $(\varepsilon_i)_{1 \leq i \leq n}$ . Therefore

$$\begin{aligned} R_c F(\xi) - F(\xi) &\geq \inf_{\varepsilon = (\varepsilon_i)_{1 \leq i \leq n}} \sum_{i=1}^n (\alpha(\mathbb{E}\varepsilon_i) - D_{x_i}^- F(\xi)\mathbb{E}\varepsilon_i) \\ &= \inf_{p \in [0,1]^n} \sum_{i=1}^n (\alpha(p_i) - D_{x_i}^- F(\xi)p_i) \\ &\geq \sum_{i=1}^n \inf_{u \in \mathbb{R}} (\alpha(u) - D_{x_i}^- F(\xi)u) \\ &= - \sum_{i=1}^n \alpha^*(D_{x_i}^- F(\xi)), \end{aligned}$$

which completes the proof.  $\square$

*Proof of [Theorem 5.1](#).* Since, for  $\lambda > 0$ ,  $(\lambda\alpha_1)^*(s) = \lambda\alpha_1^*(t/\lambda)$ ,  $t \in \mathbb{R}$ , [Lemma 5.3](#) gives that

$$F(\eta) - R_{\lambda c_1} F(\eta) \leq \lambda \int \alpha_1^* \left( \frac{D_x^- F(\eta)}{\lambda} \right) \eta(dx),$$

for all non-decreasing convex function  $F$ . Therefore, by [Theorem 5.1](#), for such  $F$  it holds

$$\mathbb{E}nt \left( e^{F(\eta)} \right) \leq \frac{\lambda}{1-\lambda} \mathbb{E} \left[ e^{F(\eta)} \int \alpha_1^* \left( \frac{D_x^- F(\eta)}{\lambda} \right) \eta(dx) \right], \quad \forall \lambda \in (0, 1).$$

A simple calculation shows that  $\alpha_1^*(s) = s - \log(1+s)$  for all  $s > -1$ . This concludes the proof.  $\square$

## 6 Some open questions

### 6.1 From modified logarithmic Sobolev to the transport-entropy

We ask whether it is possible to recover our [Theorem 3.3](#) directly from the modified logarithmic Sobolev inequalities by Wu [[42](#)] or from the infimum-convolution logarithmic Sobolev inequality ([5.2](#)). Following the ideas of [[7](#)], doing so would require a better understanding of the infimum-convolution operator  $R_c$  on the Poisson; as well as new techniques regarding Hamilton-Jacobi equations in the setting of generalized optimal transport, a question that has its own independent interest.

## 6.2 Links with displacement convexity

[20] introduces a notion of discrete displacement convexity of the entropy for finite graphs with respect to  $\tilde{\mathcal{T}}_2$ . This displacement convexity, that is one of the many possible adaptations of the Lott-Sturm-Villani synthetic curvature bound to the discrete setting, entails many other functional inequalities, such as a modified logarithmic Sobolev inequality and a transport-entropy inequality. In view of [Theorem 1.2](#), defining, in the spirit of [20], a notion of discrete curvature on the Poisson space would probably involve our cost  $\mathbb{M}$ , and we hope that the present work could help clarifying the situation.

## 6.3 Interacting point processes

The mixed binomial processes and Poisson point processes studied in this work exhibit a very strong independence. By definition, if  $A$  and  $B$  are disjoint measurable sets with finite  $\nu$ -measure then  $\eta(A)$  is independent of  $\eta(B)$ . More generally, the  $q$ -points correlation measure of a mixed binomial process with sampling measure  $\mu$  is proportional to  $\mu^q$ . We use this independence by identifying a mixed binomial process with the image of independent and identically distributed random variables. Developing tools to study transport-entropy inequalities for point processes not relying on the independence seems an attractive path, as it would allow to consider other point processes such as determinantal point processes. Let us point out that, to the best of our knowledge, the understanding of transport inequalities for point processes with interaction remains for the moment very partial. There exists some  $W_1$  transport-entropy inequality for Gibbs interaction as shown by [26]. In the setting of empirical measures (normalized point process), we can also mention the recent work by [10] about Coulomb gases.

## References

- [1] L. Ambrosio, N. Gigli & G. Savaré. “Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below”. In: *Invent. Math.* 195.2 (2014), pp. 289–391. DOI: [10.1007/s00222-013-0456-1](https://doi.org/10.1007/s00222-013-0456-1).
- [2] S. Bachmann & G. Peccati. “Concentration bounds for geometric Poisson functionals: logarithmic Sobolev inequalities revisited”. In: *Electron. J. Probab.* 21 (2016), Paper No. 6, 44. DOI: [10.1214/16-EJP4235](https://doi.org/10.1214/16-EJP4235).
- [3] S. Bachmann & M. Reitzner. “Concentration for Poisson  $U$ -statistics: subgraph counts in random geometric graphs”. In: *Stochastic Process. Appl.* 128.10 (2018), pp. 3327–3352. DOI: [10.1016/j.spa.2017.11.001](https://doi.org/10.1016/j.spa.2017.11.001).
- [4] J. Backhoff-Veraguas, M. Beiglböck & G. Pammer. “Existence, duality, and cyclical monotonicity for weak transport costs”. In: *Calc. Var. Partial Differential Equations* 58.6 (2019), Art. 203, 28. DOI: [10.1007/s00526-019-1624-y](https://doi.org/10.1007/s00526-019-1624-y).
- [5] D. Bakry, I. Gentil & M. Ledoux. *Analysis and geometry of Markov diffusion operators*. Vol. 348. Grundlehren der Mathematischen Wissenschaften. Springer, Cham, 2014, pp. xx+552. DOI: [10.1007/978-3-319-00227-9](https://doi.org/10.1007/978-3-319-00227-9).
- [6] S. G. Bobkov & F. Götze. “Exponential integrability and transportation cost related to logarithmic Sobolev inequalities”. In: *J. Funct. Anal.* 163.1 (1999), pp. 1–28. DOI: [10.1006/jfan.1998.3326](https://doi.org/10.1006/jfan.1998.3326).
- [7] S. G. Bobkov, I. Gentil & M. Ledoux. “Hypercontractivity of Hamilton-Jacobi equations”. In: *J. Math. Pures Appl.* (9) 80.7 (2001), pp. 669–696. DOI: [10.1016/S0021-7824\(01\)01208-9](https://doi.org/10.1016/S0021-7824(01)01208-9).
- [8] S. Boucheron, G. Lugosi & P. Massart. “Concentration inequalities using the entropy method”. In: *Ann. Probab.* 31.3 (2003), pp. 1583–1614. DOI: [10.1214/aop/1055425791](https://doi.org/10.1214/aop/1055425791).



- [9] N. Bourbaki. *Éléments de mathématique. Fasc. XXXV. Livre VI: Intégration. Chapitre IX: Intégration sur les espaces topologiques séparés*. Actualités Scientifiques et Industrielles, No. 1343. Hermann, Paris, 1969, p. 133.
- [10] D. Chafaï, A. Hardy & M. Maïda. “Concentration for Coulomb gases and Coulomb transport inequalities”. In: *J. Funct. Anal.* 275.6 (2018), pp. 1447–1483. DOI: [10.1016/j.jfa.2018.06.004](https://doi.org/10.1016/j.jfa.2018.06.004).
- [11] A. Dembo. “Information inequalities and concentration of measure”. In: *Ann. Probab.* 25.2 (1997), pp. 927–939. DOI: [10.1214/aop/1024404424](https://doi.org/10.1214/aop/1024404424).
- [12] M. Erbar & M. Huesmann. “Curvature bounds for configuration spaces”. In: *Calc. Var. Partial Differential Equations* 54.1 (2015), pp. 397–430. DOI: [10.1007/s00526-014-0790-1](https://doi.org/10.1007/s00526-014-0790-1).
- [13] N. Gozlan & C. Léonard. “Transport inequalities. A survey”. In: *Markov Process. Related Fields* 16.4 (2010), pp. 635–736.
- [14] N. Gozlan. “A characterization of dimension free concentration in terms of transportation inequalities”. In: *Ann. Probab.* 37.6 (2009), pp. 2480–2498. DOI: [10.1214/09-AOP470](https://doi.org/10.1214/09-AOP470).
- [15] N. Gozlan. “Characterization of Talagrand’s like transportation-cost inequalities on the real line”. In: *J. Funct. Anal.* 250.2 (2007), pp. 400–425. DOI: [10.1016/j.jfa.2007.05.025](https://doi.org/10.1016/j.jfa.2007.05.025).
- [16] N. Gozlan. “Transport inequalities and concentration of measure”. In: *Modélisation Aléatoire et Statistique—Journées MAS 2014*. Vol. 51. ESAIM Proc. Surveys. EDP Sci., Les Ulis, 2015, pp. 1–23. DOI: [10.1051/proc/201551001](https://doi.org/10.1051/proc/201551001).
- [17] N. Gozlan. “Transport-entropy inequalities on the line”. In: *Electron. J. Probab.* 17 (2012), no. 49, 18. DOI: [10.1214/EJP.v17-1864](https://doi.org/10.1214/EJP.v17-1864).
- [18] N. Gozlan, C. Roberto & P.-M. Samson. “A new characterization of Talagrand’s transport-entropy inequalities and applications”. In: *Ann. Probab.* 39.3 (2011), pp. 857–880. DOI: [10.1214/10-AOP5](https://doi.org/10.1214/10-AOP5).
- [19] N. Gozlan, C. Roberto & P.-M. Samson. “Characterization of Talagrand’s transport-entropy inequalities in metric spaces”. In: *Ann. Probab.* 41.5 (2013), pp. 3112–3139. DOI: [10.1214/12-AOP757](https://doi.org/10.1214/12-AOP757).
- [20] N. Gozlan, C. Roberto, P.-M. Samson & P. Tetali. “Displacement convexity of entropy and related inequalities on graphs”. In: *Probab. Theory Related Fields* 160.1-2 (2014), pp. 47–94. DOI: [10.1007/s00440-013-0523-y](https://doi.org/10.1007/s00440-013-0523-y).
- [21] N. Gozlan, C. Roberto, P.-M. Samson & P. Tetali. “Kantorovich duality for general transport costs and applications”. In: *J. Funct. Anal.* 273.11 (2017), pp. 3327–3405. DOI: [10.1016/j.jfa.2017.08](https://doi.org/10.1016/j.jfa.2017.08).
- [22] O. Kallenberg. *Random measures, theory and applications*. Vol. 77. Probability Theory and Stochastic Modelling. Springer, Cham, 2017, pp. xiii+694. DOI: [10.1007/978-3-319-41598-7](https://doi.org/10.1007/978-3-319-41598-7).
- [23] G. Last & M. Penrose. *Lectures on the Poisson process*. Vol. 7. Institute of Mathematical Statistics Textbooks. Cambridge University Press, Cambridge, 2018, pp. xx+293.
- [24] M. Ledoux. *The concentration of measure phenomenon*. Vol. 89. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2001, pp. x+181.
- [25] J. Lott & C. Villani. “Ricci curvature for metric-measure spaces via optimal transport”. In: *Ann. of Math. (2)* 169.3 (2009), pp. 903–991. DOI: [10.4007/annals.2009.169.903](https://doi.org/10.4007/annals.2009.169.903).
- [26] Y. Ma, S. Shen, X. Wang & L. Wu. “Transportation inequalities: From Poisson to Gibbs measures”. In: *Bernoulli* 17.1 (Feb. 2011), pp. 155–169. DOI: [10.3150/00-BEJ268](https://doi.org/10.3150/00-BEJ268).
- [27] K. Marton. “Bounding  $\bar{d}$ -distance by informational divergence: a method to prove measure concentration”. In: *Ann. Probab.* 24.2 (1996), pp. 857–866. DOI: [10.1214/aop/1039639365](https://doi.org/10.1214/aop/1039639365).
- [28] K. Marton. “A measure concentration inequality for contracting Markov chains”. In: *Geom. Funct. Anal.* 6.3 (1996), pp. 556–571. DOI: [10.1007/BF02249263](https://doi.org/10.1007/BF02249263).
- [29] B. Maurey. “Some deviation inequalities”. In: *Geom. Funct. Anal.* 1.2 (1991), pp. 188–197. DOI: [10.1007/BF01896377](https://doi.org/10.1007/BF01896377).

- [30] F. Otto & C. Villani. “Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality”. In: *J. Funct. Anal.* 173.2 (2000), pp. 361–400. DOI: [10.1006/jfan.1999.355](https://doi.org/10.1006/jfan.1999.355)
- [31] M. Reitzner. “Poisson point processes: large deviation inequalities for the convex distance”. In: *Electron. Commun. Probab.* 18 (2013), no. 96, 7. DOI: [10.1214/ECP.v18-2851](https://doi.org/10.1214/ECP.v18-2851).
- [32] P.-M. Samson. “Concentration inequalities for convex functions on product spaces”. In: *Stochastic inequalities and applications*. Vol. 56. Progr. Probab. Birkhäuser, Basel, 2003, pp. 33–52.
- [33] P.-M. Samson. “Infimum-convolution description of concentration properties of product probability measures, with applications”. In: *Ann. Inst. H. Poincaré Probab. Statist.* 43.3 (2007), pp. 321–338. DOI: [10.1016/j.anihpb.2006.05.003](https://doi.org/10.1016/j.anihpb.2006.05.003).
- [34] Y. Shu. “Hamilton-Jacobi equations on graph and applications”. In: *Potential Anal.* 48.2 (2018), pp. 125–157. DOI: [10.1007/s11118-017-9628-8](https://doi.org/10.1007/s11118-017-9628-8).
- [35] Y. Shu & M. Strzelecki. “A characterization of a class of convex log-Sobolev inequalities on the real line”. In: *Ann. Inst. Henri Poincaré Probab. Stat.* 54.4 (2018), pp. 2075–2091. DOI: [10.1214/17-AIHP865](https://doi.org/10.1214/17-AIHP865).
- [36] M. Sion. “On general minimax theorems”. In: *Pacific J. Math.* 8 (1958), pp. 171–176. URL: <http://projecteuclid.org/euclid.pjm/1103040253>.
- [37] K.-T. Sturm. “On the geometry of metric measure spaces. I”. In: *Acta Math.* 196.1 (2006), pp. 65–131. DOI: [10.1007/s11511-006-0002-8](https://doi.org/10.1007/s11511-006-0002-8).
- [38] K.-T. Sturm. “On the geometry of metric measure spaces. II”. In: *Acta Math.* 196.1 (2006), pp. 133–177. DOI: [10.1007/s11511-006-0003-7](https://doi.org/10.1007/s11511-006-0003-7).
- [39] M. Talagrand. “Transportation cost for Gaussian and other product measures”. In: *Geom. Funct. Anal.* 6.3 (1996), pp. 587–600. DOI: [10.1007/BF02249265](https://doi.org/10.1007/BF02249265).
- [40] M. Talagrand. “Concentration of measure and isoperimetric inequalities in product spaces”. In: *Inst. Hautes Études Sci. Publ. Math.* 81 (1995), pp. 73–205. URL: <http://www.numdam.org/item/>
- [41] C. Villani. *Optimal transport*. Vol. 338. Grundlehren der Mathematischen Wissenschaften. Springer-Verlag, Berlin, 2009, pp. xxii+973. DOI: [10.1007/978-3-540-71050-9](https://doi.org/10.1007/978-3-540-71050-9).
- [42] L. Wu. “A new modified logarithmic Sobolev inequality for Poisson point processes and several applications”. In: *Probab. Theory Related Fields* 118.3 (2000), pp. 427–438. DOI: [10.1007/PL000087](https://doi.org/10.1007/PL000087)

N.G. is a member of MAP5 (UMR CNRS 8145), UNIVERSITÉ PARIS DESCARTES  
45 rue des Saints-Pères, 75270 Paris cedex 6, France.  
[natael.gozlan@parisdescartes.fr](mailto:natael.gozlan@parisdescartes.fr)

R.H. is a member of INSTITUT FÜR ANGEWANDTE MATHEMATIK, BONN UNIVERSITÄT  
Endenicher Allee 60, D-53115, Bonn, Germany.  
[herry@iam.uni-bonn.de](mailto:herry@iam.uni-bonn.de)

G.P. is a member of MRU, UNIVERSITÉ DU LUXEMBOURG  
6 avenue de la Fonte, L-4363 Esch-sur-Alzette, Luxembourg.  
[giovanni.peccati@uni.lu](mailto:giovanni.peccati@uni.lu)