Bernoulli **17**(1), 2011, 155–169
DOI: 10.3150/00-BEJ268

Transportation inequalities: From Poisson to Gibbs measures

YUTAO MA[1], SHI SHEN[2], XINYU WANG[3] and LIMING WU[4]

[1]*School of Mathematical Sciences and Lab. Math. Com. Sys., Beijing Normal University, 100875 Beijing, China. E-mail: mayt@bnu.edu.cn*
[2]*College of Science, Minzu University of China, 100081 Beijing, China. E-mail: mayt@bnu.edu.cn*
[3]*School of Mathematics, Wuhan University, 430072 Hubei, China. E-mail: mayt@bnu.edu.cn*
[4]*Laboratoire de Mathématiques, CNRS UMR 6620, Université Blaise Pascal, avenue des Landais 63177 Aubière, France and Institute of Applied Mathematics, Chinese Academy of Sciences, 100190 Beijing, China. E-mail: Li-Ming.Wu@math.univ-bpclermont.fr*

We establish an optimal transportation inequality for the Poisson measure on the configuration space. Furthermore, under the Dobrushin uniqueness condition, we obtain a sharp transportation inequality for the Gibbs measure on $\mathbb{N}^\Lambda$ or the continuum Gibbs measure on the configuration space.

Keywords: Gibbs measures; Poisson point processes; transportation inequalities

1. Introduction

Transportation inequality W_1H. Let $\mathcal{X}$ be a Polish space equipped with the Borel σ-field $\mathcal{B}$ and d be a lower semi-continuous metric on the product space $\mathcal{X}\times\mathcal{X}$ (which does not necessarily generate the topology of $\mathcal{X}$). Let $\mathcal{M}_1(\mathcal{X})$ be the space of all probability measures on $\mathcal{X}$. Given $p\geq 1$ and two probability measures μ and ν on $\mathcal{X}$, we define the quantity

$$W_{p,d}(\mu,\nu)=\inf\left(\int\int d(x,y)^p\,\mathrm{d}\pi(x,y)\right)^{1/p},$$

where the infimum is taken over all probability measures π on the product space $\mathcal{X}\times\mathcal{X}$ with marginal distributions μ and ν (say, coupling of (μ,ν)). This infimum is finite provided that μ and ν belong to $\mathcal{M}_1^p(\mathcal{X},d):=\{\nu\in\mathcal{M}_1(\mathcal{X});\int d^p(x,x_0)\,\mathrm{d}\nu<+\infty\}$, where x_0 is some fixed point of $\mathcal{X}$. This quantity is commonly referred to as the L^p-*Wasserstein distance* between μ and ν. When d is the trivial metric $d(x,y)=1_{x\neq y}$, $2W_{1,d}(\mu,\nu)=\|\mu-\nu\|_{\mathrm{TV}}$, the total variation of $\mu-\nu$.

The Kullback information (or relative entropy) of ν with respect to μ is defined as

$$H(\nu/\mu)=\begin{cases}\displaystyle\int\log\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\,\mathrm{d}\nu & \text{if } \nu\ll\mu,\\ +\infty & \text{otherwise.}\end{cases}\tag{1.1}$$

Let α be a non-decreasing left-continuous function on $\mathbb{R}^+=[0,+\infty)$ which vanishes at 0. If, moreover, α is convex, we write $\alpha\in\mathcal{C}$. We say that the probability measure μ satisfies the *transportation inequality* α-W_1H *with deviation function* α on $(\mathcal{X}, d)$ if

$$\alpha(W_{1,d}(\mu, \nu)) \le H(\nu/\mu) \qquad \forall \nu \in \mathcal{M}_1(\mathcal{X}). \tag{1.2}$$

This transportation inequality W_1H was introduced and studied by Marton [11] in relation with measure concentration, for quadratic deviation function α. It was further characterized by Bobkov and Götze [1], Djellout, Guillin and Wu [4], Bolley and Villani [2] and others. The latest development is due to Gozlan and Léonard [7], in which the general α-W_1H inequality above was introduced in relation to large deviations and characterized by concentration inequalities, as follows.

Theorem 1.1 (Gozlan and Léonard [7]). *Let* $\alpha \in \mathcal{C}$ *and* $\mu \in \mathcal{M}_1^1(\mathcal{X}, d)$. *The following statements are then equivalent*:

(a) *the transportation inequality* α-W_1H (1.2) *holds*;

(b) *for all* $\lambda \ge 0$ *and all* $F \in b\mathcal{B}$, $\|F\|_{\mathrm{Lip}(d)} := \sup_{x \neq y} \frac{|F(x)-F(y)|}{d(x,y)} \le 1$,

$$\log \int_{\mathcal{X}} \exp\bigl(\lambda[F - \mu(F)]\bigr)\mu(\mathrm{d}x) \le \alpha^*(\lambda),$$

where $\mu(F) := \int_{\mathcal{X}} F\,\mathrm{d}\mu$ *and* $\alpha^*(\lambda) := \sup_{r \ge 0}(\lambda r - \alpha(r))$ *is the semi-Legendre transformation of* α;

(b′) *for all* $\lambda \ge 0$ *and all* $F, G \in C_b(\mathcal{X})$ (*the space of all bounded and continuous functions on* $\mathcal{X}$) *such that* $F(x) - G(y) \le d(x, y)$ *for all* $x, y \in \mathcal{X}$,

$$\log \int_{\mathcal{X}} \mathrm{e}^{\lambda F} \mu(\mathrm{d}x) \le \lambda\mu(G) + \alpha^*(\lambda);$$

(c) *for any measurable function* F *such that* $\|F\|_{\mathrm{Lip}(d)} \le 1$, *the following concentration inequality holds true*: *for all* $n \ge 1, r \ge 0$,

$$\mathbb{P}\left(\frac{1}{n}\sum_1^n F(\xi_k) \ge \mu(F) + r\right) \le \mathrm{e}^{-n\alpha(r)}, \tag{1.3}$$

where $(\xi_n)_{n\ge1}$ *is a sequence of i.i.d.* $\mathcal{X}$*-valued random variables with common law* μ.

The estimate on the Laplace transform in (b) and the concentration inequality in (1.3) are the main motivations for the transportation inequality (α-W_1H).

Objective and organization. The objective of this paper is to prove the transportation inequality (α-W_1H) for:

(1) (the free case) the Poisson measure P^0 on the configuration space consisting of Radon point measures $\omega = \sum_i \delta_{x_i}, x_i \in E$ with some σ-finite intensity measure m on E, where E is some fixed locally compact space;

(2) (the interaction case) the continuum Gibbs measure over a compact subset E of $\mathbb{R}^d$,

$$P^\phi(\mathrm{d}\omega) = \frac{\mathrm{e}^{-(1/2)\sum_{x_i,x_j\in\operatorname{supp}\omega, i\neq j}\phi(x_i-x_j)-\sum_{k,x_i\in\operatorname{supp}(\omega)}\phi(x_i-y_k)}}{Z}P^0(\mathrm{d}\omega),$$

where $\phi:\mathbb{R}^d \to [0,+\infty]$ is some pair-interaction non-negative even function (see Section 4 for notation) and P^0 is the Poisson measure with intensity $z\,\mathrm{d}x$ on E.

For Poisson measures on $\mathbb{N}$, Liu [10] obtained the optimal deviation function by means of Theorem 1.1. For transportation inequalities of Gibbs measures on discrete sites, see [12] and [17].

For an illustration of our main result (Theorem 4.1) on the continuum Gibbs measure P^ϕ, let $E := [-N, N]^d$ ($1 \le N \in \mathbb{N}$) and $f : [-N, N]^d \to \mathbb{R}$ be measurable and periodic with period 1 at each variable so that $|f| \le M$. Consider the empirical mean per volume $F(\omega) := \omega(f)/(2N)^d$ of f. Under Dobrushin's uniqueness condition $D := z\int_{\mathbb{R}^d}(1-\mathrm{e}^{-\phi(y)})\,\mathrm{d}y < 1$, we have (see Remark 4.3 for proof)

$$P^\phi\bigl(F > P^\phi(F) + r\bigr) \le \exp\biggl(-\frac{(2N)^d(1-D)r}{2M}\log\biggl(1+\frac{(1-D)r}{zM}\biggr)\biggr), \qquad r > 0, \tag{1.4}$$

an explicit Poissonian concentration inequality which is sharp when $\phi = 0$.

The paper is organized as follows. In the next section, we prove $(\alpha\text{–}W_1H)$ for the Poisson measure on the configuration space with respect to two metrics: in both cases, we obtain optimal deviation functions. Our main tool is Gozlan and Leonard's Theorem 1.1 and a known concentration inequality in [15]. Section 3, as a prelude to the study of the continuum Gibbs measure P^ϕ on the configuration space, is devoted to the study of a Gibbs measure on $\mathbb{N}^\Lambda$. Our method is a combination of a lemma on W_1H for mixed measure, Dobrushin's uniqueness condition and the McDiarmid–Rio martingale method for dependent tensorization of the W_1H-inequality. Finally, in the last section, by approximation, we obtain a sharp $(\alpha\text{–}W_1H)$ inequality for the continuum Gibbs measure P^ϕ under Dobrushin's uniqueness condition $D = z\int_{\mathbb{R}^d}(1-\mathrm{e}^{-\phi(y)})\,\mathrm{d}y < 1$. The latter is a sharp sufficient condition, both for the analyticity of the pressure functional and for the spectral gap; see [16].

2. Poisson point processes

Poisson space. Let E be a metric complete locally compact space with the Borel field $\mathcal{B}_E$ and m a σ-finite positive Radon measure on E. The Poisson space $(\Omega, \mathcal{F}, P^0)$ is given by:

(1) $\Omega := \{\omega = \sum_i \delta_{x_i}(\text{Radon measure}); x_i \in E\}$ (the so-called configuration space over E);
(2) $\mathcal{F} = \sigma(\omega \to \omega(B) | B \in \mathcal{B}_E)$;
(3) $\forall B \in \mathcal{B}_E, \forall k \in \mathbb{N}$: $P^0(\omega : \omega(B) = k) = \mathrm{e}^{-m(B)}\frac{m(B)^k}{k!}$;
(4) $\forall B_1, \dots, B_n \in \mathcal{B}_E$ disjoint, $\omega(B_1), \dots, \omega(B_n)$ are P^0-independent,

where δ_x denotes the Dirac measure at x. Under P^0, ω is exactly the Poisson point process on E with intensity measure $m(\mathrm{d}x)$. On Ω, we consider the vague convergence topology, that is, the coarsest topology such that $\omega \to \omega(f)$ is continuous, where f runs over the space $C_0(E)$ of all continuous functions with compact support on E. Equipped with this topology, Ω is a Polish space and this topology is the weak convergence topology (of measures) if E is compact.

Definition 2.1. *Letting φ be a positive measurable function on E, we define a metric $d_\varphi(\cdot,\cdot)$ (which may be infinite) on the Poisson space $(\Omega, \mathcal{F}, P^0)$ by*

$$d_\varphi(\omega, \omega') = \int_E \varphi \, \mathrm{d}|\omega - \omega'|,$$

where $|\nu| := \nu^+ + \nu^-$ for a signed measure ν ($\nu^\pm$ are, respectively, the positive and negative parts of ν in the Hahn–Jordan decomposition).

Lemma 2.2. *If φ is continuous, then the metric d_φ is lower semi-continuous on Ω.*

Proof. Indeed, for any $\omega, \omega' \in \Omega$,

$$d_\varphi(\omega, \omega') = \sup_f |\omega(f) - \omega'(f)|,$$

where the supremum is taken over all bounded $\mathcal{B}_E$-measurable functions f with compact support such that $|f| \le \varphi$. Now, as φ is continuous, we can approximate such f by $f_n \in C_0(E)$ in $L^1(E, \omega + \omega')$ and $|f_n| \le \varphi$. Then

$$d_\varphi(\omega, \omega') = \sup_{f \in C_0(E), |f| \le \varphi} |\omega(f) - \omega'(f)|.$$

As $(\omega, \omega') \to |\omega(f) - \omega'(f)|$ is continuous on $\Omega \times \Omega$, $d_\varphi(\omega, \omega')$ is lower semi-continuous on $\Omega \times \Omega$. $\square$

Assume from now on that φ is continuous. Then, for any $\nu, \mu \in \mathcal{M}_1(\Omega)$, we have the Kantorovitch–Rubinstein equality [8,9,14],

$$\begin{aligned} W_{1,d_\varphi}(\mu, \nu) &= \sup\left\{ \int F \, \mathrm{d}\nu - \int G \, \mathrm{d}\mu \,\middle|\, F, G \in C_b(\Omega), F(\omega) - G(\omega') \le d_\varphi(\omega, \omega') \right\} \\ &= \sup\left\{ \int G \, \mathrm{d}(\nu - \mu) : G \in b\mathcal{F}, \|G\|_{\mathrm{Lip}(d_\varphi)} \le 1 \right\}. \end{aligned}$$

Here, $b\mathcal{F}$ is the space of all real, bounded and $\mathcal{F}$-measurable functions.

The difference operator D. We denote by $L^0(\Omega, P^0)$ the space of all P^0-equivalent classes of real measurable functions w.r.t. the completion of $\mathcal{F}$ by P^0. Hence, the difference operator $D : L^0(\Omega, P^0) \to L^0(E \times \Omega, m \otimes P^0)$ given by

$$F \to D_x F(\omega) := F(\omega + \delta_x) - F(\omega)$$

is well defined (see [15]) and plays a crucial role in the Malliavin calculus on the Poisson space.

Lemma 2.3. *Given a measurable function* $F : \Omega \to \mathbb{R}$, $\|F\|_{\mathrm{Lip}(d_\varphi)} \le 1$ *if and only if* $|D_x F(\omega)| \le \varphi(x)$ *for all* $\omega \in \Omega$ *and* $x \in E$.

Proof. If $\|F\|_{\mathrm{Lip}(d_\varphi)} \le 1$, since

$$|D_x F(\omega)| = |F(\omega + \delta_x) - F(\omega)| \le d_\varphi(\omega + \delta_x, \omega) = \int_E \varphi \,\mathrm{d}|(\omega + \delta_x) - \omega| = \varphi(x),$$

the necessity is true. We now prove the sufficiency. For any $\omega, \omega' \in \Omega$, we write $\omega = \sum_{k=1}^{i} \delta_{x_k} + \omega \wedge \omega'$ and $\omega' = \sum_{k=1}^{j} \delta_{y_k} + \omega \wedge \omega'$, where $\omega \wedge \omega' := \frac{1}{2}(\omega + \omega' - |\omega - \omega'|)$. We then have

$$\begin{aligned}
|F(\omega) - F(\omega')| &\le |F(\omega) - F(\omega \wedge \omega')| + |F(\omega') - F(\omega \wedge \omega')| \\
&\le \sum_{k=1}^{i} \left| F\left(\omega \wedge \omega' + \sum_{l=1}^{k} \delta_{x_l}\right) - F\left(\omega \wedge \omega' + \sum_{l=1}^{k-1} \delta_{x_l}\right) \right| \\
&\quad + \sum_{k=1}^{j} \left| F\left(\omega \wedge \omega' + \sum_{l=1}^{k} \delta_{y_l}\right) - F\left(\omega \wedge \omega' + \sum_{l=1}^{k-1} \delta_{y_l}\right) \right| \\
&\le \sum_{k=1}^{i} \varphi(x_k) + \sum_{k=1}^{j} \varphi(y_k) = \int_E \varphi \,\mathrm{d}|\omega - \omega'| = d_\varphi(\omega, \omega'),
\end{aligned}$$

which implies that $\|F\|_{\mathrm{Lip}(d_\varphi)} \le 1$. □

Remark 2.4. When $\varphi = 1$, we denote d_φ by d. Obviously, $d(\omega, \omega') = |\omega - \omega'|(E) = \|\omega - \omega'\|_{\mathrm{TV}}$, that is, d is exactly the total variation distance.

The following result, due to the fourth-named author [15], was obtained by means of the L^1-log-Sobolev inequality and will play an important role.

Lemma 2.5 ([15], Proposition 3.2). *Let* $F \in L^1(\Omega, P^0)$. *If there is some* $0 \le \varphi \in L^2(E, m)$ *such that* $|D_x F(\omega)| \le \varphi(x)$, $m \otimes P^0$*-a.e., then for any* $\lambda \ge 0$,

$$\mathbb{E}^{P^0} \mathrm{e}^{\lambda(F - P^0(F))} \le \exp\left\{ \int_E (\mathrm{e}^{\lambda\varphi} - \lambda\varphi - 1) \,\mathrm{d}m \right\}.$$

In particular, if m *is finite and* $|D_x F(\omega)| \le 1$ *for* $m \times P^0$*-a.e.* (x, ω) *on* $E \times \Omega$ *(i.e.,* $\varphi(x) = 1$*), then*

$$\mathbb{E}^{P^0} \mathrm{e}^{\lambda(F - P^0(F))} \le \exp\{(\mathrm{e}^{\lambda} - \lambda - 1) m(E)\}.$$

We now state our main result on the Poisson space.

Theorem 2.6. *Let $(\Omega, \mathcal{F}, P^0)$ be the Poisson space with intensity measure $m(dx)$ and φ a bounded continuous function on E such that $0 < \varphi \le M$ and $\sigma^2 = \int_E \varphi^2 \, \mathrm{d}m < +\infty$. Then*

$$\frac{1}{M} h_c(W_{1,d_\varphi}(Q, P^0)) \le H(Q|P^0) \qquad \forall Q \in \mathcal{M}_1(\Omega), \tag{2.1}$$

where $c = \sigma^2/M$ and

$$h_c(r) = c \cdot h\left(\frac{r}{c}\right), \qquad h(r) = (1+r)\log(1+r) - r. \tag{2.2}$$

Note that $h^*(\lambda) := \sup_{r\ge 0}(\lambda r - h(r)) = \mathrm{e}^\lambda - \lambda - 1$ and $h_c^*(\lambda) = ch^*(\lambda)$.

Proof of Theorem 2.6. Since the function $(\mathrm{e}^{\lambda\varphi} - \lambda\varphi - 1)/\varphi^2$ is increasing in φ, it is easy to see that

$$\int_E (\mathrm{e}^{\lambda\varphi} - \lambda\varphi - 1)\, \mathrm{d}m \le \frac{\mathrm{e}^{\lambda M} - \lambda M - 1}{M^2} \int \varphi^2 \, \mathrm{d}m. \tag{2.3}$$

Further, the Legendre transformation of the right-hand side of (2.3) is, for $r \ge 0$,

$$\begin{aligned}
\sup_{\lambda\ge 0}\left\{\lambda r - \frac{\mathrm{e}^{\lambda M} - \lambda M - 1}{M^2}\int \varphi^2\,\mathrm{d}m\right\} &= \left(\frac{r}{M} + \frac{\int \varphi^2\,\mathrm{d}m}{M^2}\right)\log\left(\frac{Mr}{\int\varphi^2\,\mathrm{d}m} + 1\right) - \frac{r}{M} \\
&= \frac{1}{M} h_c(r).
\end{aligned}$$

The desired result then follows from Theorem 1.1, by Lemma 2.5. □

Remark 2.7. Let $\beta(\lambda) := \int_E (\mathrm{e}^{\lambda\varphi} - \lambda\varphi - 1)\,\mathrm{d}m$ and $\alpha(r) := \sup_{\lambda\ge 0}(\lambda r - \beta(\lambda))$. The proof above gives us

$$\alpha(W_{1,d_\varphi}(Q, P^0)) \le H(Q|P^0) \qquad \forall Q \in \mathcal{M}_1(\Omega).$$

This less explicit inequality is sharp. Indeed, assume that E is compact and let $F(\omega) := \int_E \varphi(x)(\omega - m)(\mathrm{d}x)$. We have $\|F\|_{\mathrm{Lip}(d_\varphi)} = 1$ and

$$\log \mathbb{E}^{P^0} \mathrm{e}^{\lambda F} = \beta(\lambda).$$

The sharpness is then ensured by Theorem 1.1.

Proposition 2.8. *If $\varphi = 1$ and m is finite, then the inequality (2.1) turns out to be*

$$h_{m(E)}(W_{1,d}(Q, P^0)) \le H(Q|P^0) \qquad \forall Q \in \mathcal{M}_1(\Omega). \tag{2.4}$$

In particular, for the Poisson measure $\mathcal{P}(\lambda)$ with parameter $\lambda > 0$ on $\mathbb{N}$ equipped with the Euclidean distance ρ,

$$h_\lambda(W_{1,\rho}(\nu, \mathcal{P}(\lambda))) \le H(\nu|\mathcal{P}(\lambda)) \qquad \forall \nu \in \mathcal{M}_1(\mathbb{N}). \tag{2.5}$$

Proof. The inequality (2.4) is a particular case of (2.1) with $\varphi = 1$ and it holds on $\Omega^0 := \{\omega \in \Omega; \omega(E) < +\infty\}$ (for P^0 is actually supported in Ω^0 as m is finite). For (2.5), let $m(E) = \lambda$ and consider the mapping $\Psi : \Omega^0 \to \mathbb{N}$, $\Psi(\omega) = \omega(E)$. Since $|\Psi(\omega) - \Psi(\omega')| = |\omega(E) - \omega'(E)| \le d(\omega, \omega')$, Ψ is Lipschitzian with the Lipschitzian coefficient less than 1. Thus, (2.5) follows from (2.4) by [4], Lemma 2.1 and its proof. □

Remark 2.9. The transportation inequality (2.5) was shown by Liu [10] by means of a tensorization technique and the approximation of $\mathcal{P}(\lambda)$ by binomial distributions. It is optimal (therefore, so is (2.4)). In fact, consider another Poisson distribution $\mathcal{P}(\lambda')$ with parameter $\lambda' > \lambda$. On the one hand,

$$\begin{aligned} H(\mathcal{P}(\lambda')|\mathcal{P}(\lambda)) &= \int_{\mathbb{N}} \log \frac{\mathrm{d}\mathcal{P}(\lambda')}{\mathrm{d}\mathcal{P}(\lambda)} \, \mathrm{d}\mathcal{P}(\lambda') = \sum_{n=0}^{\infty} \mathcal{P}(\lambda')(n) \log\left(\frac{\mathrm{e}^{-\lambda'}\lambda'^n}{n!} \Big/ \frac{\mathrm{e}^{-\lambda}\lambda^n}{n!} \right) \\ &= \lambda - \lambda' + \sum_{n=0}^{\infty} \mathcal{P}(\lambda')(n) n \log \frac{\lambda'}{\lambda} \\ &= \lambda - \lambda' + \lambda' \log \frac{\lambda'}{\lambda}. \end{aligned}$$

On the other hand, let $r := \lambda' - \lambda > 0$. Let X, Y be two independent random variables having distributions $\mathcal{P}(\lambda)$ and $\mathcal{P}(r)$, respectively. Obviously, the law of $X + Y$ is $\mathcal{P}(\lambda')$. Then

$$W_{1,\rho}(\mathcal{P}(\lambda'), \mathcal{P}(\lambda)) \le \mathbb{E}|X - (X + Y)| = \mathbb{E}Y = r.$$

Now, supposing that (X, X') is a coupling of $\mathcal{P}(\lambda')$ and $\mathcal{P}(\lambda)$, we have

$$\mathbb{E}|X - X'| \ge |\mathbb{E}X - \mathbb{E}X'| = r,$$

which implies that $W_{1,\rho}(\mathcal{P}(\lambda'), \mathcal{P}(\lambda)) \ge r$. Then $W_{1,\rho}(\mathcal{P}(\lambda'), \mathcal{P}(\lambda)) = r$ (and $(X, X + Y)$ is an optimal coupling for $\mathcal{P}(\lambda)$ and $\mathcal{P}(\lambda')$). Therefore,

$$h_\lambda(W_{1,\rho}(\mathcal{P}(\lambda'), \mathcal{P}(\lambda))) = h_\lambda(r) = H(\mathcal{P}(\lambda')|\mathcal{P}(\lambda)).$$

Namely, h_λ is the optimal deviation function for the Poisson distribution $\mathcal{P}(\lambda)$.

3. A discrete spin system

The model and the Dobrushin interdependence coefficient. Let $\Lambda = \{1, \ldots, N\}$ ($2 \le N \in \mathbb{N}$) and $\gamma : \Lambda \times \Lambda \mapsto [0, +\infty]$ be a *non-negative* interaction function satisfying $\gamma_{ij} = \gamma_{ji}$ and $\gamma_{ii} = 0$ for all $i, j \in \Lambda$. Consider the Gibbs measure P on $\mathbb{N}^\Lambda$ with

$$P(x_1, \ldots, x_N) = \mathrm{e}^{-\sum_{i<j} \gamma_{ij} x_i x_j} \prod_{i=1}^{N} \mathcal{P}(\delta_i)(x_i) \Big/ C, \tag{3.1}$$

where $\mathcal{P}(\delta_i)(x_i) = \mathrm{e}^{-\delta_i}\frac{\delta_i^{x_i}}{x_i!}$, $x_i \in \mathbb{N}$, is the Poisson distribution with parameter $\delta_i > 0$ and C is the normalization constant. Here and hereafter, the convention that $0 \cdot \infty = 0$ is used. Let $P_i(\mathrm{d}x_i|x_\Lambda)$ be the given regular conditional distribution of x_i given $x_{\Lambda\setminus\{i\}}$, which is, in the present case, the Poisson distribution $\mathcal{P}(\delta_i \mathrm{e}^{-\sum_{j\neq i}\gamma_{ij}x_j})$ with parameter $\delta_i \mathrm{e}^{-\sum_{j\neq i}\gamma_{ij}x_j}$, with the convention that the Poisson measure $\mathcal{P}(0)$ with parameter $\lambda = 0$ is the Dirac measure δ_0 at 0. Define the Dobrushin interdependence matrix $C := (c_{ij})_{i,j\in\Lambda}$ w.r.t. the Euclidean metric ρ by

$$c_{ij} = \sup_{x_\Lambda = x'_\Lambda \,\mathrm{off}\, j} \frac{W_{1,\rho}(P_i(\mathrm{d}x_i|x_\Lambda), P_i(\mathrm{d}x'_i|x'_\Lambda))}{|x_j - x'_j|} \qquad \forall i, j \in \Lambda \tag{3.2}$$

(obviously, $c_{ii} = 0$). The Dobrushin uniqueness condition [5,6] is then

$$D := \sup_j \sum_i c_{ij} < 1.$$

For this model, we can identify c_{ij}.

Lemma 3.1. *Recall that $\gamma_{ij} \geq 0$. We have*

$$c_{ij} = \delta_i(1 - \mathrm{e}^{-\gamma_{ij}}).$$

Proof. By Remark 2.9, if $x_\Lambda = x'_\Lambda$ off j, then

$$W_{1,\rho}(P_i(\mathrm{d}x_i|x_\Lambda), P_i(\mathrm{d}x'_i|x'_\Lambda)) = \delta_i\left|\mathrm{e}^{-\sum_k \gamma_{ik}x_k} - \mathrm{e}^{-\sum_k \gamma_{ik}x'_k}\right|.$$

Without loss of generality, suppose that $x_j = x'_j + x$ with $x \geq 1$. We have then

$$\begin{aligned} c_{ij} &= \delta_i \sup_{x_\Lambda = x'_\Lambda \,\mathrm{off}\, j} \frac{|\mathrm{e}^{-\sum_k \gamma_{ik}x_k} - \mathrm{e}^{-\sum_k \gamma_{ik}x'_k}|}{|x_j - x'_j|} \\ &= \delta_i \sup_{x\geq 1} \frac{1 - \mathrm{e}^{-\gamma_{ij}x}}{x} \qquad (\text{taking } x_k = x'_k = 0 \text{ for } k \neq j,\ x'_j = 0) \\ &= \delta_i(1 - \mathrm{e}^{-\gamma_{ij}}). \end{aligned}$$

Here, the first equality holds since γ_{ij} is non-negative and the last equality is due to the fact that $(1 - \mathrm{e}^{-\gamma_{ij}x})/x$ is decreasing in $x > 0$. $\square$

The transportation inequality W_1H for mixed measure. We return to the general framework of the Introduction. Let $\mathcal{X}$ be a general Polish space and d be a metric on $\mathcal{X}$ which is lower semi-continuous on $\mathcal{X} \times \mathcal{X}$. Consider a mixed probability measure $\mu := \int_I \mu_\lambda \,\mathrm{d}\sigma(\lambda)$ on $\mathcal{X}$, where, for each $\lambda \in I$, μ_λ is a probability on $\mathcal{X}$ and σ is a probability measure on another Polish space I. Let ρ be a lower semi-continuous metric on I.

Proposition 3.2. *Suppose that*:

(i) *for any* $\lambda \in I$, μ_λ *satisfies* α–W_1H *with deviation function* $\alpha \in \mathcal{C}$,

$$\alpha(W_{1,d}(\nu, \mu_\lambda)) \le H(\nu|\mu_\lambda) \qquad \forall \nu \in \mathcal{M}_1(\mathcal{X});$$

(ii) σ *satisfies a* β–W_1H *inequality on* I *with deviation function* $\beta \in \mathcal{C}$,

$$\beta(W_{1,\rho}(\eta, \sigma)) \le H(\eta|\sigma) \qquad \forall \eta \in \mathcal{M}_1(I);$$

(iii) $\lambda \to \mu_\lambda$ *is Lipschitzian, that is, for some constant* $M > 0$,

$$W_{1,d}(\mu_\lambda, \mu_{\lambda'}) \le M\rho(\lambda, \lambda') \qquad \forall \lambda, \lambda' \in I.$$

The mixed probability $\mu = \int_I \mu_\lambda \, \mathrm{d}\sigma(\lambda)$ *then satisfies*

$$\tilde{\alpha}(W_{1,d}(\nu, \mu)) \le H(\nu|\mu) \qquad \forall \nu \in \mathcal{M}_1(\mathcal{X}), \tag{3.3}$$

where

$$\tilde{\alpha}(r) = \sup_{b \ge 0}\{br - [\alpha^*(b) + \beta^*(bM)]\}, \qquad r \ge 0.$$

Proof. By Gozlan and Leonard's Theorem 1.1, it is enough to show that for any Lipschitzian function f on $\mathcal{X}$ with $\|f\|_{\mathrm{Lip}(d)} \le 1$ and $b \ge 0$,

$$\int_{\mathcal{X}} \mathrm{e}^{b[f(x) - \mu(f)]} \, \mathrm{d}\mu(x) \le \exp\bigl(\alpha^*(b) + \beta^*(bM)\bigr).$$

Let $g(\lambda) := \int_{\mathcal{X}} f(x) \, \mathrm{d}\mu_\lambda(x) = \mu_\lambda(f)$. We have $\sigma(g) = \mu(f)$ and, by Kantorovitch's duality equality and our condition (iii), $|g(\lambda) - g(\lambda')| \le M\rho(\lambda, \lambda')$. Using Theorem 1.1 and our conditions (i) and (ii), we then get, for any $b \ge 0$,

$$\begin{aligned} \int_{\mathcal{X}} \mathrm{e}^{b[f(x) - \mu(f)]} \, \mathrm{d}\mu &= \int_I \left(\int_{\mathcal{X}} \mathrm{e}^{b[f(x) - \mu_\lambda(f)]} \, \mathrm{d}\mu_\lambda(x) \right) \mathrm{e}^{b[g(\lambda) - \sigma(g)]} \, \mathrm{d}\sigma(\lambda), \\ &\le \mathrm{e}^{\alpha^*(b) + \beta^*(bM)} \end{aligned}$$

the desired result. □

We now turn to a mixed Poisson distribution,

$$\mu = \int_0^a \mathcal{P}(\lambda)\sigma(\mathrm{d}\lambda), \tag{3.4}$$

where $a > 0$. By Proposition 2.8, we know that w.r.t. the Euclidean metric ρ,

$$h_\lambda(W_{1,\rho}(\nu, \mathcal{P}(\lambda))) \le H(\nu|\mathcal{P}(\lambda))$$

and $W_{1,\rho}(\mathcal{P}(\lambda), \mathcal{P}(\lambda')) = |\lambda - \lambda'|$. Since h_λ is decreasing in λ, the hypotheses in Proposition 3.2 with $E = \mathbb{N}$, $I = [0, a]$, both equipped with the Euclidean metric ρ, are satisfied with

$\alpha(r) = h_a(r) = ah(\frac{r}{a})$ and $\beta(r) = 2r^2/a^2$ (the well-known CKP inequality). On the other hand, obviously,

$$h(r) = (1+r)\log(1+r) - r \le \frac{r^2}{2}, \qquad r \ge 0,$$

which implies that

$$h_{a^2/4}(r) = \frac{a^2}{4} h\left(\frac{4r}{a^2}\right) \le \frac{2r^2}{a^2} = \beta(r).$$

Since $h_c^*(\lambda) = c(e^\lambda - \lambda - 1)$,

$$\sup_{b\ge 0}\{br - [(h_a(b))^* + (h_{a^2/4}(b))^*]\} = \sup_{b\ge 0}\{br - (a + a^2/4)(e^b - b - 1)\} = h_{a+a^2/4}(r).$$

By Proposition 3.2, we have, for the mixed Poisson measure μ given in (3.4),

$$h_{a+a^2/4}(W_{1,d}(\nu, \mu)) \le H(\nu|\mu) \qquad \forall \nu \in \mathcal{M}_1(\mathbb{N}). \tag{3.5}$$

See Chafai and Malrieu [3] for fine analysis of transportation or functional inequalities for mixed measures. We can now state the main result of this section.

Theorem 3.3. *Let P be the Gibbs measure given in (3.1) with $\gamma_{ij} \ge 0$. Assume Dobrushin's uniqueness condition*

$$D := \sup_{j\in\Lambda} \sum_{i\in\Lambda} \delta_i (1 - e^{-\gamma_{ij}}) < 1.$$

For any probability measure Q on $\mathbb{N}^\Lambda$ equipped with the metric $\rho_H(x_\Lambda, y_\Lambda) := \sum_{i\in\Lambda} |x_i - y_i|$ (the index H refers to Hamming), we then have, for $c := \sum_{i\in\Lambda}(\delta_i + \delta_i^2/4)$,

$$h_c\big((1-D)W_{1,\rho_H}(Q, P)\big) \le H(Q|P) \qquad \forall Q \in \mathcal{M}_1(\mathbb{N}^\Lambda).$$

This result, without the extra constants $\delta_i^2/4$, would become sharp if $\gamma = 0$ (i.e., without interaction) or $P = \mathcal{P}(\delta)^{\otimes\Lambda}$.

Proof of Theorem 3.3. By Theorem 1.1, it is equivalent to prove that for any 1-Lipschitzian functional F w.r.t. the metric ρ_H,

$$\log \mathbb{E}^P e^{\lambda(F - \mathbb{E}^P F)} \le h_c^*\left(\frac{\lambda}{1-D}\right) = ch^*\left(\frac{\lambda}{1-D}\right) \qquad \forall \lambda > 0. \tag{3.6}$$

We prove the inequality (3.6) by the McDiarmid–Rio martingale method (as in [4,17]). Consider the martingale

$$M_0 = \mathbb{E}^P(F), \qquad M_k(x_1^k) = \int F(x_1^k, x_{k+1}^N) P(dx_{k+1}^N | x_1^k), \qquad 1 \le k \le N,$$

where $x_i^j = (x_k)_{i\le k\le j}$, $P(dx_{k+1}^N|x_1^k)$ is the conditional distribution of x_{k+1}^N given x_1^k. Since $M_N = F$, we have

$$\mathbb{E}^P e^{\lambda(F-\mathbb{E}^P F)} = \mathbb{E}^P \exp\left(\lambda \sum_{k=1}^N (M_k - M_{k-1})\right).$$

By induction, for (3.6), it suffices to establish that for each $k = 1,\dots,N$, P-a.s.,

$$\log \int \exp\big(\lambda\big(M_k(x_1^{k-1}, x_k) - M_{k-1}(x_1^{k-1})\big)\big) P(\mathrm{d}x_k|x_1^{k-1}) \le (\delta_k + \delta_k^2/4) h^*\left(\frac{\lambda}{1-D}\right). \tag{3.7}$$

By (3.5), $P(\mathrm{d}x_k|x_1^{k-1})$, being a convex combination of Poisson measures $P_k(\mathrm{d}x_k|x_\Lambda) = \mathcal{P}(\delta_k e^{-\sum_{j\ne k}\gamma_{kj}x_j})$ (over x_{k+1}^N), satisfies the W_1H-inequality with the deviation function $h_{\delta_k+\delta_k^2/4}$. Hence, by Theorem 1.1, (3.7) holds if

$$|M_k(x_1^{k-1}, x_k) - M_k(x_1^{k-1}, y_k)| \le \frac{1}{1-D}|x_k - y_k|. \tag{3.8}$$

In fact, the inequality (3.8) has been proven in [17], step 2 in the proof of Theorem 4.3. The proof is thus complete. □

Remark 3.4. For a previous study on transportation inequalities for Gibbs measures on discrete sites, see Marton [12] and Wu [17]. Our method here is quite close to that in [17], but with two new features: (1) W_1H for mixed probability measures; (2) Gozlan and Léonard's Theorem 1.1 as a new tool.

Remark 3.5. Every Poisson distribution $\mathcal{P}(\lambda)$ satisfies the Poincaré inequality ([15], Remark 1.4)

$$\mathrm{Var}_{\mathcal{P}(\lambda)}(f) \le \lambda \int_{\mathbb{N}} (Df(x))^2 \,\mathrm{d}\mathcal{P}(\lambda)(x) \qquad \forall f \in L^2(\mathbb{N}, \mathcal{P}(\lambda)),$$

where $Df(x) := f(x+1) - f(x)$ and $\mathrm{Var}_\mu(f) := \mu(f^2) - [\mu(f)]^2$ is the variance of f w.r.t. μ. By [17], Theorem 2.2 we have the following Poincaré inequality for the Gibbs measure P: if $D < 1$, then

$$\mathrm{Var}_P(F) \le \frac{\max_{1\le i\le N}\delta_i}{1-D} \int_{\mathbb{N}^\Lambda} \sum_{i\in\Lambda} (D_i F)^2(x)\,\mathrm{d}P(x) \qquad \forall F \in L^2(\mathbb{N}^\Lambda, P),$$

where $D_i F(x_1,\dots,x_N) := F(x_1,\dots,x_{i-1}, x_i + 1, x_{i+1},\dots,x_N) - F(x_1,\dots,x_N)$. We remind the reader that an important open question is to prove the L^1-log-Sobolev inequality (or entropy inequality)

$$H(FP|P) \le C \int_{\mathbb{N}^\Lambda} \sum_{i\in\Lambda} D_i F \cdot D_i \log F \,\mathrm{d}P \qquad \text{for all } P\text{-probability densities } F$$

(which is equivalent to the exponential convergence in entropy of the corresponding Glauber system) under Dobrushin's uniqueness condition, or at least for high temperature.

4. $W_1 H$-inequality for the continuum Gibbs measure

We now generalize the result for the discrete sites Gibbs measure in Section 3 to the continuum Gibbs measure (continuous gas model), by an approximation procedure.

Let $(\Omega, \mathcal{F}, P^0)$ be the Poisson space over a compact subset E of $\mathbb{R}^d$ with intensity $m(\mathrm{d}x) = z\,\mathrm{d}x$, where the Lebesgue measure $|E|$ of E is positive and finite, and $z > 0$ represents the *activity*. Given a *non-negative* pair-interaction function $\phi : \mathbb{R}^d \mapsto [0, +\infty]$, which is measurable and even over $\mathbb{R}^d$, the corresponding Poisson space is denoted by $(\Omega, \mathcal{F}, P^0)$ and the associated Gibbs measure is given by

$$P^\phi(\mathrm{d}\omega) = \frac{\mathrm{e}^{-(1/2)\sum_{x_i, x_j \in \mathrm{supp}(\omega), i \neq j} \phi(x_i - x_j) - \sum_{k, x_i \in \mathrm{supp}(\omega)} \phi(x_i - y_k)}}{Z} P^0(\mathrm{d}\omega),$$

where Z is the normalization constant and $\{y_k, k\}$ is an at most countable family of points in $\mathbb{R}^d \backslash E$ such that $\sum_k \phi(x - y_k) < +\infty$ for all $x \in E$ (boundary condition). The main result of this section is the following theorem.

Theorem 4.1. *Assume that the Dobrushin uniqueness condition holds, that is,*

$$D := z \int_{\mathbb{R}^d} \big(1 - \mathrm{e}^{-\phi(y)}\big)\,\mathrm{d}y < 1. \tag{4.1}$$

Then, w.r.t. the total variation distance $d = d_\varphi$ with $\varphi = 1$ on Ω,

$$h_{z|E|}\big((1 - D) W_{1,d}(Q, P^\phi)\big) \le H(Q|P^\phi) \qquad \forall Q \in \mathcal{M}_1(\Omega). \tag{4.2}$$

Remark 4.2. Without interaction (i.e., $\phi = 0$), $D = 0$ and the $W_1 H$-inequality (4.2) is exactly the optimal $W_1 H$-inequality for the Poisson measure P^0 in Proposition 2.8. In the presence of non-negative interaction ϕ, it is well known that $D < 1$ is a sharp condition for the analyticity of the pressure functional $p(z)$: indeed, the radius R of convergence of the entire series of $p(z)$ at $z = 0$ satisfies $R \int_{\mathbb{R}^d} (1 - \mathrm{e}^{-\phi(y)})\,\mathrm{d}y < 1$; see [13], Theorem 4.5.3. The corresponding sharp Poincaré inequality for P^ϕ was established in [16].

Proof of Theorem 4.1. We shall establish this sharp α–$W_1 H$ inequality for P^ϕ by approximation.

By part (b′) of Theorem 1.1, it is equivalent to show that for any $F, G \in C_b(\Omega)$ such that $F(\omega) - G(\omega') \le d(\omega, \omega')$, $\omega, \omega' \in \Omega$, and for any $\lambda > 0$,

$$\log \int_\Omega \mathrm{e}^{\lambda F}\,\mathrm{d}P^\phi \le \lambda P^\phi(G) + z|E| h^*\left(\frac{\lambda}{1 - D}\right), \tag{4.3}$$

where $h^*(\lambda) = \mathrm{e}^\lambda - \lambda - 1$.

Step 1. *ϕ is continuous and $\{y_k, k\}$ is finite.* We want to approximate P^ϕ by the discrete sites Gibbs measures given in the previous section. To this end, assume first that ϕ is continuous ($+\infty$ is regarded as the one-point compactification of $\mathbb{R}^+$) or, equivalently, that $\mathrm{e}^{-\phi} : \mathbb{R}^d \to [0, 1]$ is continuous with the convention that $\mathrm{e}^{-\infty} := 0$.

For each $N \geq 2$, let $\{E_1, \dots, E_N\}$ be a measurable decomposition of E such that, as N goes to infinity, $\max_{1\leq i\leq N} \mathrm{Diam}(E_i) \to 0$ and $\max_{1\leq i\leq N} |E_i| \to 0$, where $|E|$ is the Lebesgue measure of E and $\mathrm{Diam}(E_i) = \sup_{x,y\in E_i} |x - y|$ is the diameter of E_i. Fix $x_i^0 \in E_i$ for each i. Consider the probability measure P_N on $\mathbb{N}^\Lambda$ ($\Lambda := \{1, \dots, N\}$) given by, for all $(n_1, \dots, n_N) \in \mathbb{N}^\Lambda$,

$$P_N(n_1, \dots, n_N) = (1/Z)\mathrm{e}^{-(1/2)\sum_{i\neq j} \phi(x_i^0 - x_j^0)n_i n_j - \sum_{i,k} \phi(x_i^0 - y_k)n_i} \prod_{i=1}^N \mathcal{P}(z|E_i|)(n_i)$$

$$= (1/Z')\mathrm{e}^{-\sum_{i<j} \phi(x_i^0 - x_j^0)n_i n_j} \prod_{i=1}^N \mathcal{P}(\delta_{N,i})(n_i),$$

where Z, Z' are normalization constants and $\delta_{N,i} = z|E_i|\mathrm{e}^{-\sum_k \phi(x_i^0 - y_k)} \leq z|E_i|$. Consider the mapping $\Phi : \mathbb{N}^\Lambda \to \Omega$ given by

$$\Phi(n_1, \dots, n_N) = \sum_{i=1}^N n_i \delta_{x_i^0}.$$

Φ is isometric from $(\mathbb{N}^\Lambda, \rho_H)$ to (Ω, d), where $d = d_\varphi$ with $\varphi = 1$ (given in Section 2). Finally, let P^N be the push-forward of P_N by Φ. It is quite direct to see that $P^N \to P$ weakly.

The Dobrushin constant D_N associated with P_N is given by

$$D_N = \sup_j \sum_i \delta_{N,i}\bigl(1 - \mathrm{e}^{-\phi(x_i^0 - x_j^0)}\bigr) \leq \sup_j \sum_i z|E_i|\bigl(1 - \mathrm{e}^{-\phi(x_i^0 - x_j^0)}\bigr).$$

When N goes to infinity,

$$\limsup_{N\to\infty} D_N \leq \sup_{y\in\mathbb{R}^d} z \int_E \bigl(1 - \mathrm{e}^{-\phi(x-y)}\bigr)\,\mathrm{d}x = z \int_{\mathbb{R}^d} \bigl(1 - \mathrm{e}^{-\phi(x)}\bigr)\,\mathrm{d}x = D.$$

Therefore, if $D < 1$ and $D_N < 1$ for all N large enough, then the $W_1 H$-inequality in Theorem 3.3 holds for P_N. By the isometry of the mapping Φ, P^N satisfies the same $W_1 H$-inequality on Ω w.r.t. the metric d, which gives us, by Theorem 1.1(b′),

$$\log \mathbb{E}^{P^N} \mathrm{e}^{\lambda F} \leq \lambda P^N(G) + \biggl(\sum_{i\in\Lambda} [\delta_{N,i} + \delta_{N,i}^2/4]\biggr) h^*\biggl(\frac{\lambda}{1 - D_N}\biggr).$$

By letting N go to infinity, this yields (4.3), for $P^N \to P^\phi$ weakly and

$$\sum_{i\in\Lambda} [\delta_{N,i} + \delta_{N,i}^2/4] \leq \sum_{i\in\Lambda} z|E_i|(1 + z|E_i|/4) \to z|E|.$$

Step 2. *General ϕ and $\{y_k, k\}$ is finite.* For general measurable non-negative and even interaction function ϕ, we take a sequence of continuous, even and non-negative functions (ϕ_n) such that $1 - e^{-\phi_n} \to 1 - e^{-\phi}$ in $L^1(\mathbb{R}^d, dx)$. Now, note that $\frac{dP^{\phi_n}}{dP^0} \to \frac{dP^{\phi}}{dP^0}$ in $L^1(\Omega, P^0)$, that is, $P^{\phi_n} \to P^{\phi}$ in total variation. Hence, (4.3) for P^{ϕ_n} (proved in step 1) yields (4.3) for P^{ϕ}.

Step 3. *General case.* Finally, if the set of points $\{y_k, k\}$ is infinite, approximating $\sum_{k=1}^{\infty} \phi(x_i - y_k)$ by $\sum_{k=1}^{n} \phi(x_i - y_k)$ in the definition of P^{ϕ}, we get (4.3) for P^{ϕ}, as in step 2. □

Remark 4.3. The explicit Poissonian concentration inequality (1.4) follows from Theorem 4.1 by Theorem 1.1(c) (with $n = 1$) by noting that the observable $F(\omega) = \omega(f)/(2N)^d$ there is Lipschitzian w.r.t. d with $\|F\|_{\mathrm{Lip(d)}} \le M/(2N)^d$ and $h(r) \ge (r/2)\log(1+r)$.

Remark 4.4. A quite curious phenomena occurs in the continuous gas model: the *extra* constant $\delta_i^2/4$ coming from the mixture of measures now disappears.

Acknowledgments

We are grateful to the referee for his conscientious comments. Yutao Ma was supported by NSFC Grant No. 10721091.

References

[1] Bobkov, S.G. and Götze, F. (1999). Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** 1–28. MR1682772

[2] Bolley, F. and Villani, C. (2005). Weighted Csiszár–Kullback–Pinsker inequalities and applications to transportation inequalities. *Ann. Fac. Sci. Toulouse* **14** 331–352. MR2172583

[3] Chafai, D. and Malrieu, F. (2010). On fine properties of mixtures with respect to concentration and Sobolev type inequalities. *Ann. Inst. H. Poincaré.* Preprint. To appear.

[4] Djellout, H., Gullin, A. and Wu, L.M. (2004). Transportation cost-information inequalities for random dynamical systems and diffusions. *Ann. Probab.* **32** 2702–2732. MR2078555

[5] Dobrushin, R.L. (1968). The description of a random field by means of conditional probabilities and condition of its regularity. *Theory Probab. Appl.* **13** 197–224. MR0231434

[6] Dobrushin, R.L. (1970). Prescribing a system of random variables by conditional distributions. *Theory Probab. Appl.* **15** 458–486.

[7] Gozlan, N. and Léonard, C. (2007). A large deviation approach to some transportation cost inequalities. *Probab. Theory Related Fields* **139** 235–283. MR2322697

[8] Kellerer, H.G. (1984). Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete* **67** 399–432. MR0761565

[9] Léonard, C. (2007). Transport inequalities: A large deviation point of view. In *Course in Chinese Summer School for Ph.D. Students, Wuhan*.

[10] Liu, W. Optimal transportation-entropy inequalities for several usual distributions on $\mathbb{R}$. Preprint. Submitted.

[11] Marton, K. (1996). Bounding $\bar{d}$-distance by informational divergence: A way to prove measure concentration. *Ann. Probab.* **24** 857–866. MR1404531

[12] Marton, K. (2004). Measure concentration for Euclidean distance in the case of dependent random variables. *Ann. Probab.* **32** 2526–2544. MR2078549

[13] Ruelle, D. (1969). *Statistical Mechanics: Rigorous Results*. New York: Benjamin. MR0289084

[14] C. Villani. (2003). *Topics in Optimal Transportation*. Providence, RI: Amer. Math. Soc. MR1964483

[15] Wu, L.M. (2000). A new modified logarithmic Sobolev inequality for Poisson processes and several applications. *Probab. Theory Related Fields* **118** 427–438. MR1800540

[16] Wu, L.M. (2004). Estimate of the spectral gap for continuous gas. *Ann. Inst. H. Poincaré Probab. Statist.* **40** 387–409. MR2070332

[17] Wu, L.M. (2006). Poincaré and transportation inequalities for Gibbs measures under the Dobrushin uniqueness condition. *Ann. Probab.* **34** 1960–1989. MR2271488

Received March 2009 and revised February 2010