Transportation inequalities: From Poisson to Gibbs measures

YUTAO MA, SHI SHEN, XINYU WANG and LIMING WU

1School of Mathematical Sciences and Lab. Math. Com. Sys., Beijing Normal University, 100875 Beijing, China. E-mail: mayt@bnu.edu.cn
2College of Science, Minzu University of China, 100081 Beijing, China. E-mail: mayt@bnu.edu.cn
3School of Mathematics, Wuhan University, 430072 Hubei, China. E-mail: mayt@bnu.edu.cn
4Laboratoire 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 \( W_1 H \). Let \( \mathcal{X} \) be a Polish space equipped with the Borel \( \sigma \)-field \( B \) 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 \( \mu \) and \( \nu \) on \( \mathcal{X} \), we define the quantity

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

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

The Kullback information (or relative entropy) of \( \nu \) with respect to \( \mu \) is defined as

\[
H(\nu/\mu) = \begin{cases} 
\int \log \frac{d\nu}{d\mu} \ d\nu & \text{if } \nu \ll \mu, \\
+\infty & \text{otherwise.}
\end{cases}
\]

Let \( \alpha \) be a non-decreasing left-continuous function on \( \mathbb{R}^+ = [0, +\infty) \) which vanishes at 0. If, moreover, \( \alpha \) is convex, we write \( \alpha \in \mathcal{C} \). We say that the probability measure \( \mu \) satisfies the
transportation inequality \( \alpha-W_1 H \) with deviation function \( \alpha \) on \((\mathcal{X},d)\) if

\[
\alpha(W_1,d(\mu,\nu)) \leq H(\nu/\mu) \quad \forall \nu \in \mathcal{M}_1(\mathcal{X}).
\]

This transportation inequality \( W_1 H \) was introduced and studied by Marton [11] in relation with measure concentration, for quadratic deviation function \( \alpha \). 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 \( \alpha-W_1 H \) inequality above was introduced in relation to large deviations and characterized by concentration inequalities, as follows.

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

(a) the transportation inequality \( \alpha-W_1 H \) (1.2) holds;

(b) for all \( \lambda \geq 0 \) and all \( F \in bB, \|F\|_{\text{Lip}(d)} := \sup_{x \neq y} \frac{|F(x)-F(y)|}{d(x,y)} \leq 1 \),

\[
\log \int_{\mathcal{X}} \exp(\lambda[F - \mu(F)]) \mu(dx) \leq \alpha^*(\lambda),
\]

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

(b') for all \( \lambda \geq 0 \) and all \( F, G \in C_b(\mathcal{X}) \) (the space of all bounded and continuous functions on \( \mathcal{X} \)) such that \( F(x) - G(y) \leq d(x,y) \) for all \( x, y \in \mathcal{X} \),

\[
\log \int_{\mathcal{X}} e^{\lambda F} \mu(dx) \leq \lambda \mu(G) + \alpha^*(\lambda);
\]

(c) for any measurable function \( F \) such that \( \|F\|_{\text{Lip}(d)} \leq 1 \), the following concentration inequality holds true: for all \( n \geq 1, r \geq 0 \),

\[
\mathbb{P}\left( \frac{1}{n} \sum_{k=1}^{n} F(\xi_k) \geq \mu(F) + r \right) \leq e^{-n\alpha(r)},
\]

(1.3)

where \( (\xi_n)_{n \geq 1} \) is a sequence of i.i.d. \( \mathcal{X} \)-valued random variables with common law \( \mu \).

The estimate on the Laplace transform in (b) and the concentration inequality in (1.3) are the main motivations for the transportation inequality \( (\alpha-W_1 H) \).

**Objective and organization.** The objective of this paper is to prove the transportation inequality \( (\alpha-W_1 H) \) 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 \( \sigma \)-finite intensity measure \( m \) on \( E \), where \( E \) is some fixed locally compact space;
Transportation inequalities

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

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

where $\phi: \mathbb{R}^d \rightarrow [0, +\infty]$ is some pair-interaction non-negative even function (see Section 4 for notation) and $P^0$ is the Poisson measure with intensity $z \, dx$ 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 $P^\phi$, let $E := [-N, N]^d$ $(1 \leq N \in \mathbb{N})$ and $f: [-N, N]^d \rightarrow \mathbb{R}$ be measurable and periodic with period 1 at each variable so that $|f| \leq M$. Consider the empirical mean per volume $F(\omega) := \omega(f)/(2N)^d$ of $f$. Under Dobrushin’s uniqueness condition $D := z \int_{\mathbb{R}^d} (1 - e^{-\phi(y)}) \, dy < 1$, we have (see Remark 4.3 for proof)

$$P^\phi(F > P^\phi(F) + r) \leq \exp\left(-\frac{(2N)^d(1-D)r}{2M} \log\left(1 + \frac{(1-D)r}{zM}\right)\right), \quad r > 0, \quad (1.4)$$

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

The paper is organized as follows. In the next section, we prove $(\alpha-W_1 H)$ 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 $P^\phi$ 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 $W_1 H$ for mixed measure, Dobrushin’s uniqueness condition and the McDiarmid–Rio martingale method for dependent tensorization of the $W_1 H$-inequality. Finally, in the last section, by approximation, we obtain a sharp $(\alpha-W_1 H)$ inequality for the continuum Gibbs measure $P^\phi$ under Dobrushin’s uniqueness condition $D = z \int_{\mathbb{R}^d} (1 - e^{-\phi(y)}) \, dy < 1$. 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 $\sigma$-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. $\mathcal{F} = \sigma(\omega \rightarrow \omega(B) | B \in \mathcal{B}_E)$;
3. $\forall B \in \mathcal{B}_E, \forall k \in \mathbb{N}: P^0(\omega: \omega(B) = k) = e^{-m(B)} m(B)^k/k!$;
4. $\forall B_1, \ldots, B_n \in \mathcal{B}_E$ disjoint, $\omega(B_1), \ldots, \omega(B_n)$ are $P^0$-independent,

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

**Definition 2.1.** Letting \( \varphi \) 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

\[
d_\varphi(\omega, \omega') = \int_E \varphi \, d|\omega - \omega'|,
\]

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

**Lemma 2.2.** If \( \varphi \) is continuous, then the metric \( d_\varphi \) is lower semi-continuous on \( \Omega \).

**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 \( B_E \)-measurable functions \( f \) with compact support such that \( |f| \leq \varphi \). Now, as \( \varphi \) is continuous, we can approximate such \( f \) by \( f_n \in C_0(E) \) in \( L^1(E, \omega + \omega') \) and \( |f_n| \leq \varphi \). Then

\[
d_\varphi(\omega, \omega') = \sup_{f \in C_0(E), |f| \leq \varphi} |\omega(f) - \omega'(f)|.
\]

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

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

\[
W_{1,d_\varphi}(\mu, \nu) = \sup\left\{ \int F \, d\nu - \int G \, d\mu : F, G \in C_b(\Omega), F(\omega) - G(\omega') \leq d_\varphi(\omega, \omega') \right\}
\]

\[
= \sup\left\{ \int G \, d(\nu - \mu) : G \in b\mathcal{F}, \|G\|_{\text{Lip}(d_\varphi)} \leq 1 \right\}.
\]

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 \( D : L^0(\Omega, P^0) \rightarrow L^0(E \times \Omega, m \otimes P^0) \) given by

\[
F \rightarrow D_x F(\omega) := F(\omega + \delta_x) - F(\omega)
\]

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 $F : \Omega \to \mathbb{R}$, $\|F\|_{\text{Lip}(d_\varphi)} \leq 1$ if and only if $|D_x F(\omega)| \leq \varphi(x)$ for all $\omega \in \Omega$ and $x \in E$.

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

$$|D_x F(\omega)| = |F(\omega + \delta_x) - F(\omega)| \leq d_\varphi(\omega + \delta_x, \omega) = \int_E \varphi \, |\omega + \delta_x - \omega| = \varphi(x),$$

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

$$|F(\omega) - F(\omega')| \leq \sum_{k=1}^i |F(\omega \wedge \omega' + \sum_{l=1}^k \delta_{x_l}) - F(\omega \wedge \omega' + \sum_{l=1}^{k-1} \delta_{x_l})| + \sum_{k=1}^j |F(\omega \wedge \omega' + \sum_{l=1}^k \delta_{y_l}) - F(\omega \wedge \omega' + \sum_{l=1}^{k-1} \delta_{y_l})|$$

$$\leq \sum_{k=1}^i \varphi(x_k) + \sum_{k=1}^j \varphi(y_k) = \int_E \varphi \, d|\omega - \omega'| = d_\varphi(\omega, \omega'),$$

which implies that $\|F\|_{\text{Lip}(d_\varphi)} \leq 1$. \hfill \Box

Remark 2.4. When $\varphi = 1$, we denote $d_\varphi$ by $d$. Obviously, $d(\omega, \omega') = |\omega - \omega'|(E) = \|\omega - \omega'||_{TV}$, 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 $0 \leq \varphi \in L^2(E, m)$ such that $|D_x F(\omega)| \leq \varphi(x)$, $m \otimes P^0$-a.e., then for any $\lambda \geq 0$,

$$\mathbb{E}^P_0 e^{\lambda(F - P^0(F))} \leq \exp\left\{ \int_E (e^{\lambda \varphi} - \lambda \varphi - 1) \, dm \right\}.$$

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

$$\mathbb{E}^P_0 e^{\lambda(F - P^0(F))} \leq \exp\{(e^{\lambda} - \lambda - 1)m(E)\}.$$

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 $\varphi$ a bounded continuous function on $E$ such that $0 < \varphi \leq M$ and $\sigma^2 = \int_E \varphi^2 \, dm < +\infty$. Then
\[
\frac{1}{M} h_c(W_1, d_\varphi(Q, P^0)) \leq H(Q|P^0) \quad \forall Q \in \mathcal{M}_1(\Omega),
\] (2.1)

where $c = \sigma^2/M$ and
\[
h_c(r) = c \cdot h\left(\frac{r}{c}\right), \quad h(r) = (1 + r) \log(1 + r) - r.
\] (2.2)

Proof of Theorem 2.6. Since the function $(e^{\lambda \varphi} - \lambda \varphi - 1)/\varphi^2$ is increasing in $\varphi$, it is easy to see that
\[
\int_E (e^{\lambda \varphi} - \lambda \varphi - 1) \, dm \leq \frac{e^{\lambda M} - \lambda M - 1}{M^2} \int \varphi^2 \, dm.
\] (2.3)

Further, the Legendre transformation of the right-hand side of (2.3) is, for $r \geq 0$,
\[
\sup_{\lambda \geq 0} \left\{ \lambda r - \frac{e^{\lambda M} - \lambda M - 1}{M^2} \int \varphi^2 \, dm \right\} = \left( \frac{r}{M} + \frac{\int \varphi^2 \, dm}{M^2} \right) \log \left( \frac{Mr}{\int \varphi^2 \, dm} + 1 \right) - \frac{r}{M}
\]
\[
= \frac{1}{M} h_c(r).
\]

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

Remark 2.7. Let $\beta(\lambda) := \int_E (e^{\lambda \varphi} - \lambda \varphi - 1) \, dm$ and $\alpha(r) := \sup_{\lambda \geq 0} (\lambda r - \beta(\lambda))$. The proof above gives us
\[
\alpha(W_1, d_\varphi(Q, P^0)) \leq H(Q|P^0) \quad \forall Q \in \mathcal{M}_1(\Omega).
\]

This less explicit inequality is sharp. Indeed, assume that $E$ is compact and let $F(\omega) := \int_E \varphi(x)(\omega - m) \, dx$. We have $\|F\|_{\text{Lip}(d_\varphi)} = 1$ and
\[
\beta(\lambda) = \log \mathbb{E} P^0 e^{\lambda F}.
\]

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
\[
h_m(E)(W_1, d_\varphi(Q, P^0)) \leq H(Q|P^0) \quad \forall Q \in \mathcal{M}_1(\Omega).
\] (2.4)

In particular, for the Poisson measure $\mathcal{P}(\lambda)$ with parameter $\lambda > 0$ on $\mathbb{N}$ equipped with the Euclidean distance $\rho$,
\[
h_\lambda(W_1, \rho(v, \mathcal{P}(\lambda))) \leq H(v|\mathcal{P}(\lambda)) \quad \forall v \in \mathcal{M}_1(\mathbb{N}).
\] (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 \( \Omega^0 \) 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 \( |\Psi(\omega) - \Psi(\omega')| = |\omega(E) - \omega'(E)| \leq d(\omega, \omega') \), \( \Psi \) is Lipschitzian with the Lipschitzian coefficient less than 1. Thus, (2.5) follows from (2.4) by [4], Lemma 2.1 and its proof. \( \square \)

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

\[
H(P(\lambda')|P(\lambda)) = \int_{\mathbb{N}} \log \frac{dP(\lambda')}{dP(\lambda)} dP(\lambda') = \sum_{n=0}^{\infty} P(\lambda')(n) \log \left( \frac{e^{-\lambda' \lambda^m}}{n!} / \frac{e^{-\lambda \lambda^m}}{n!} \right) = \lambda - \lambda' + \sum_{n=0}^{\infty} P(\lambda')(n) n \log \frac{\lambda'}{\lambda} = \lambda - \lambda' + \lambda' \log \frac{\lambda'}{\lambda}.
\]

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

\[
W_{1,\rho}(P(\lambda'), P(\lambda)) \leq \mathbb{E}|X - (X + Y)| = \mathbb{E}Y = r.
\]

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

\[
\mathbb{E}|X - X'| \geq \|\mathbb{E}X - \mathbb{E}X'\| = r,
\]

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

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

Namely, \( h_\lambda \) is the optimal deviation function for the Poisson distribution \( P(\lambda) \).

### 3. A discrete spin system

**The model and the Dobrushin interdependence coefficient.** Let \( \Lambda = \{1, \ldots, N\} \) \((2 \leq N \in \mathbb{N})\) 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

\[
P(x_1, \ldots, x_N) = e^{-\sum_{i<j} \gamma_{ij} x_i x_j} \prod_{i=1}^{N} P(\delta_i)(x_i) / C, \tag{3.1}
\]
where \( \mathcal{P}(\delta_i)(x_i) = e^{-\frac{\delta_i x_i}{x_i^!}}, x_i \in \mathbb{N} \), 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 \( P_i(dx_i|x_A) \) be the given regular conditional distribution of \( x_i \) given \( x_A \setminus \{i\} \), which is, in the present case, the Poisson distribution \( \mathcal{P}(\delta_i e^{-\sum_{j \neq i} \gamma_{ij} x_j}) \) with parameter \( \delta_i e^{-\sum_{j \neq i} \gamma_{ij} x_j} \), with the convention that the Poisson measure \( \mathcal{P}(0) \) with parameter \( \lambda = 0 \) is the Dirac measure \( \delta_0 \) at 0. Define the Dobrushin interdependence matrix \( C := (c_{ij})_{i,j \in \Lambda} \) w.r.t. the Euclidean metric \( \rho \) by

\[
  c_{ij} = \sup_{x_A = x_\Lambda^! \setminus j} \frac{W_1,\rho(P_i(dx_i|x_A), P_i(dx_i'|x_\Lambda^!))}{|x_j - x_j'|} \quad \forall i, j \in \Lambda \tag{3.2}
\]

(possibly, \( c_{ii} = 0 \)). 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

\[
  c_{ij} = \delta_i (1 - e^{-\gamma_{ij}}).
\]

**Proof.** By Remark 2.9, if \( x_A = x_\Lambda^! \setminus j \), then

\[
  W_1,\rho(P_i(dx_i|x_A), P_i(dx_i'|x_\Lambda^!)) = \delta_i |e^{-\sum_k \gamma_{ik} x_k} - e^{-\sum_k \gamma_{ik} x_k'}|.
\]

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

\[
  c_{ij} = \delta_i \sup_{x_A = x_\Lambda^! \setminus j} \frac{|e^{-\sum_k \gamma_{ik} x_k} - e^{-\sum_k \gamma_{ik} x_k'}|}{|x_j - x_j'|} = \delta_i \sup_{x \geq 1} \frac{1 - e^{-\gamma_{ij} x}}{x} (\text{taking } x_k = x_k' = 0 \text{ for } k \neq j, x_j' = 0)
\]

\[
  = \delta_i (1 - e^{-\gamma_{ij}}).
\]

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

The transportation inequality \( W_1 H \) 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 \( \mu := \int_I \mu_\lambda \, d\sigma(\lambda) \) on \( \mathcal{X} \), where, for each \( \lambda \in I \), \( \mu_\lambda \) is a probability on \( \mathcal{X} \) and \( \sigma \) is a probability measure on another Polish space \( I \). Let \( \rho \) be a lower semi-continuous metric on \( I \).

**Proposition 3.2.** Suppose that:
Transportation inequalities

(i) for any $\lambda \in I$, $\mu_\lambda$ satisfies $\alpha - W_1 H$ with deviation function $\alpha \in C$, 

$$\alpha(W_{1,d}(v, \mu_\lambda)) \leq H(v|\mu_\lambda) \quad \forall v \in M_1(X);$$

(ii) $\sigma$ satisfies a $\beta - W_1 H$ inequality on $I$ with deviation function $\beta \in C,

$$\beta(W_{1,\rho}(\eta, \sigma)) \leq H(\eta|\sigma) \quad \forall \eta \in M_1(I);$$

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

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

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

$$\tilde{\alpha}(W_{1,d}(v, \mu)) \leq H(v|\mu) \quad \forall v \in M_1(X), \quad (3.3)$$

where 

$$\tilde{\alpha}(r) = \sup_{b \geq 0} \{ br - [\alpha^*(b) + \beta^*(bM)] \}; \quad r \geq 0.$$

**Proof.** By Gozlan and Leonard’s Theorem 1.1, it is enough to show that for any Lipschitzian function $f$ on $X$ with $\|f\|_{\text{Lip}(d)} \leq 1$ and $b \geq 0$, 

$$\int_X e^{b[f(x) - \mu(f)]} \, d\mu(x) \leq \exp(\alpha^*(b) + \beta^*(bM)).$$

Let $g(\lambda) := \int_X f(x) \, d\mu_\lambda(x) = \mu_\lambda(f)$. We have $\sigma(g) = \mu(f)$ and, by Kantorovitch’s duality equality and our condition (iii), $|g(\lambda) - g(\lambda')| \leq M\rho(\lambda, \lambda')$. Using Theorem 1.1 and our conditions (i) and (ii), we then get, for any $b \geq 0$,

$$\int_X e^{b[f(x) - \mu(f)]} \, d\mu = \int_I \left( \int_X e^{b[f(x) - \mu_\lambda(f)]} \, d\mu_\lambda(x) \right) e^{b[g(\lambda) - \sigma(g)]} \, d\sigma(\lambda),$$

$$\leq e^{\alpha^*(b) + \beta^*(bM)}$$

the desired result. \qed

We now turn to a mixed Poisson distribution,

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

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

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

and $W_{1,\rho}(\mathcal{P}(\lambda), \mathcal{P}(\lambda')) = |\lambda - \lambda'|$. Since $h_\lambda$ is decreasing in $\lambda$, the hypotheses in Proposition 3.2 with $E = \mathbb{N}$, $I = [0, a]$, both equipped with the Euclidean metric $\rho$, are satisfied with
\[\alpha(r) = h_a(r) = a h_a \left( \frac{r}{a} \right) \] and \[\beta(r) = 2r^2/a^2\] (the well-known CKP inequality). On the other hand, obviously,
\[h(r) = (1 + r) \log(1 + r) - r \leq \frac{r^2}{2}, \quad r \geq 0,\]
which implies that
\[h_{a^2/4} = \frac{a^2}{4} h \left( \frac{4r}{a^2} \right) \leq \frac{2r^2}{a^2} = \beta(r).\]

Since \[\gamma_{ij} \geq 0\] and Dobrushin’s uniqueness condition \[\sup_{j} \sum_{i} \delta_i (1 - e^{-\gamma_{ij}}) < 1.\]

For any probability measure \(Q\) on \(\mathbb{N}^\Lambda\) equipped with the metric \(\rho_H(x_i, y_i) := \sum_{i \in \Lambda} |x_i - y_i|\) (the index \(H\) refers to Hamming), we then have, for \(c := \sum_{i \in \Lambda} (\delta_i + \delta^2_i/4),\)
\[h_c \left( (1 - D) W_{1, \rho_H(Q, P)} \right) \leq H(Q|P) \quad \forall Q \in \mathcal{M}_1(\mathbb{N}^\Lambda).\]

This result, without the extra constants \(\delta^2_i/4\), 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 \(\rho_H,\)
\[\log \mathbb{E}^P e^{\lambda(F - \mathbb{E}^PF)} \leq h^*_c \left( \frac{\lambda}{1 - D} \right) = c h^*_c \left( \frac{\lambda}{1 - D} \right) \quad \forall \lambda > 0.\]  \hspace{1cm} (3.6)

We prove the inequality (3.6) by the McDiarmid–Rio martingale method (as in [4, 17]). Consider the martingale
\[M_0 = \mathbb{E}^P (F), \quad M_k(x^k_1) = \int F(x^k_1, x^N_{k+1}) P(dx^N_{k+1}|x^k_1), \quad 1 \leq k \leq N,\]
where \( x_j^i = (x_k)_{i \leq k \leq j} \), \( P(dx_{k+1}^N|x_1^k) \) is the conditional distribution of \( x_{k+1}^N \) given \( x_k^k \). 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 \( k = 1, \ldots, N \), \( P \)-a.s.,

\[
\log \int \exp \left( \lambda (M_k(x_1^k, x_k) - M_{k-1}(x_1^{k-1})) \right) P(dx_1^k|x_1^{k-1}) \leq (\delta_k + \delta_k^2/4) h^* \left( \frac{\lambda}{1 - D} \right). \tag{3.7}
\]

By (3.5), \( P(dx_1^k|x_1^{k-1}) \), being a convex combination of Poisson measures \( P_k(dx_k|x_k) = \mathcal{P}(\delta_k e^{-\sum_j k \gamma_j x_j}) \) (over \( x_{k+1}^N \)), satisfies the \( W_{1,H} \)-inequality with the deviation function \( h_{\delta_k + \delta_k^2/4} \). Hence, by Theorem 1.1, (3.7) holds if

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

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. \( \square \)

**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) \( W_{1,H} \) 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)

\[
\text{Var}_{\mathcal{P}(\lambda)}(f) \leq \lambda \int_N (Df(x))^2 d\mathcal{P}(\lambda)(x) \quad \forall f \in L^2(N, \mathcal{P}(\lambda)),
\]

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

\[
\text{Var}_P(F) \leq \max_{1 \leq i \leq N} \delta_i \int_{\mathbb{N}^\Lambda} \sum_{i \in \Lambda} (D_i F)^2(x) dP(x) \quad \forall F \in L^2(\mathbb{N}^\Lambda, P),
\]

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

\[
H(FP|P) \leq C \int_{\mathbb{N}^\Lambda} \sum_{i \in \Lambda} D_i F \cdot D_i \log F dP \quad \text{for all } P\text{-probability densities } F
\]
(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 $m(dx) = z dx$, where the Lebesgue measure $|E|$ of $E$ is positive and finite, and $z > 0$ represents the activity.

Let $\{y_k, k\}$ be an at most countable family of points in $\mathbb{R}^d \setminus E$ 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,

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

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

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

**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 $\phi$, 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 $R \int_{\mathbb{R}^d} (1 - e^{-\phi(y)}) \, dy < 1$; see [13], Theorem 4.5.3. The corresponding sharp Poincaré inequality for $P^\phi$ was established in [16].

**Proof of Theorem 4.1.** We shall establish this sharp $\alpha-W_1 H$ inequality for $P^\phi$ 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 $F(\omega) - G(\omega') \leq d(\omega, \omega')$, $\omega, \omega' \in \Omega$, and for any $\lambda > 0$,

$$\log \int_{\Omega} e^{\lambda F} \, dP^\phi \leq \lambda P^\phi(G) + z|E|h^*(\frac{\lambda}{1-D}), \quad (4.3)$$
where $h^*(\lambda) = e^\lambda - \lambda - 1$.

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

For each $N \geq 2$, let $\{E_1, \ldots, E_N\}$ be a measurable decomposition of $E$ such that, as $N$ goes to infinity, $\text{max}_{1 \leq i \leq N} \text{Diam}(E_i) \to 0$ and $\text{max}_{1 \leq i \leq N} |E_i| \to 0$, where $|E|$ is the Lebesgue measure of $E$ and $\text{Diam}(E_i) = \sup_{x, y \in E_i} |x - y|$ is the diameter of $E_i$. Fix $x_0^i \in E_i$ for each $i$. Consider the probability measure $P_N$ on $\mathbb{N}^\Lambda$ ($\Lambda := \{1, \ldots, N\}$) given by, for all $(n_1, \ldots, n_N) \in \mathbb{N}^\Lambda$,

$$P_N(n_1, \ldots, n_N) = (1/Z)e^{-\sum_{i,j} \phi(x^0_i - x^0_j)n_in_j - \sum_k \phi(x^0_k - y_k)n_i} \prod_{i=1}^N P(z|E_i)(n_i)$$

$$= (1/Z')e^{-\sum_{i,j} \phi(x^0_i - x^0_j)n_in_j} \prod_{i=1}^N P(\delta_{N,i})(n_i),$$

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

$$\Phi(n_1, \ldots, n_N) = \sum_{i=1}^N n_i \delta_{y_i}.$$  

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

The Dobrushin constant $D_N$ associated with $P_N$ is given by

$$D_N = \sup \sum_i \delta_{N,i} \left(1 - e^{-\phi(x^0_i - x^0_j)}\right) \leq \sup \sum_i z|E_i| \left(1 - e^{-\phi(x^0_i - x^0_j)}\right).$$

When $N$ goes to infinity,

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

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 $\Phi$, $P_N$ satisfies the same $W_1 H$-inequality on $\Omega$ w.r.t. the metric $d$, which gives us, by Theorem 1.1(b'),

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

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 $\phi$ and $\{y_k, k\}$ is finite. For general measurable non-negative and even interaction function $\phi$, we take a sequence of continuous, even and non-negative functions $(\phi_n)$ such that $1 - e^{-\phi_n} \to 1 - e^{-\phi}$ in $L^1(\mathbb{R}^d, dx)$. Now, note that $dP_{\phi_n}/dP_0 \to dP_{\phi}/dP_0$ in $L^1(\Omega, P^0)$, that is, $P_{\phi_n} \to P_{\phi}$ in total variation. Hence, (4.3) for $P_{\phi_n}$ (proved in step 1) yields (4.3) for $P_{\phi}$.

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 $P_{\phi}$, we get (4.3) for $P_{\phi}$, as in step 2.

\[\square\]

**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 $\|F\|_{\text{Lip}(d)} \leq M/(2N)^d$ and $h(r) \geq (r/2) \log(1 + r)$.

**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 $\tilde{d}$-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] Villani, C. (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