Elect. Comm. in Probab. **16** (2011), 600–613

ELECTRONIC COMMUNICATIONS in PROBABILITY

TRANSPORTATION-INFORMATION INEQUALITIES FOR CONTINUUM GIBBS MEASURES

YUTAO MA [1]
School of Mathematical Sciences & *Lab. Math. Com. Sys., Beijing Normal University, 100875, Beijing China.*
email: mayt@bnu.edu.cn

RAN WANG[2]
School of Mathematics and Statistics, Wuhan University, 430072 Hubei, China.
email: whwangran@yahoo.com.cn

LIMING WU[3]
Institute of Applied Math., Chinese Academy of Sciences, 100190 Beijing, China and Laboratoire de Math. CNRS-UMR 6620, Université Blaise Pascal, 63177 Aubière, France.
email: Li-Ming.Wu@math.univ-bpclermont.fr

Submitted March 14, 2011, *accepted in final form* July 25, 2011

AMS 2000 Subject classification: 60E15. 60K35.
Keywords: transportation-information inequality, concentration inequality, Gibbs measure, Lyapunov function method.

Abstract

The objective of this paper is to establish explicit concentration inequalities for the Glauber dynamics related with continuum or discrete Gibbs measures. At first we establish the optimal transportation-information W_1I-inequality for the $M/M/\infty$-queue associated with the Poisson measure, which improves several previous known results. Under the Dobrushin's uniqueness condition, we obtain some explicit W_1I-inequalities for Gibbs measures both in the continuum and in the discrete lattice. Our method is a combination of Lipschitzian spectral gap, the Lyapunov test function approach and the tensorization technique.

1 Introduction

1.1 Transportation-information inequalities W_1I

Let $\mathscr{X}$ be a Polish space equipped with the Borel σ-field $\mathscr{B}$, and let d be a lower semi-continuous metric on the product space $\mathscr{X}\times\mathscr{X}$ (which does not necessarily generate the topology of $\mathscr{X}$). Let

[1]RESEARCH SUPPORTED BY THE FUNDAMENTAL RESEARCH FUNDS FOR THE CENTRAL UNIVERSITIES

[2]PARTIALLY SUPPORTED BY NSFC 11001208 AND THE FUNDAMENTAL RESEARCH FUNDS FOR THE CENTRAL UNIVERSITIES

[3]PARTIALLY SUPPORTED BY THOUSAND TALENTS PROGRAM OF THE CHINESE ACADEMY OF SCIENCES AND LE PROJET ANR EVOL

$\mathscr{M}_1(\mathscr{X})$ be the space of all probability measures on $\mathscr{X}$. Given $p \geq 1$ and two probability measures μ and ν on $\mathscr{X}$, we define the quantity

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

where the infimum is taken over all probability measures π on the product space $\mathscr{X} \times \mathscr{X}$ with marginal distributions μ and ν (say coupling of (μ,ν)). This infimum is finite once μ and ν belong to $\mathscr{M}_1^p(\mathscr{X},d) := \{\nu \in \mathscr{M}_1(\mathscr{X});\ \int d^p(x,x_0)d\nu < +\infty\}$, where x_0 is some fixed point of $\mathscr{X}$. This quantity is commonly referred to be as the L^p**-Wasserstein distance** between μ and ν. When $d(x,y) = 1_{x\neq y}$ (the trivial metric), it is known that $2W_{1,d}(\mu,\nu) = \|\mu-\nu\|_{TV}$, the total variation of the measure $\mu - \nu$.
Given a Dirichlet form $\mathscr{E}$ on $L^2(\mu) := L^2(\mathscr{X},\mu)$ with domain $\mathbb{D}(\mathscr{E})$, let $I(\nu|\mu)$ be the Fisher-Donsker-Varadhan information of ν with respect to μ

$$I(\nu|\mu) = \begin{cases} \mathscr{E}(\sqrt{f},\sqrt{f}) & \text{if } \nu = f\mu, \sqrt{f} \in \mathbb{D}(\mathscr{E}); \\ +\infty & \text{otherwise.} \end{cases} \tag{1}$$

Suppose that $((X_t)_{t\geq 0}, \mathbb{P}_\mu)$ is an $\mathscr{X}-$valued reversible Markov process associated with the Dirichlet form $(\mathscr{E},\mathbb{D}(\mathscr{E}))$. We always assume that it is ergodic, i.e., if $h \in \mathbb{D}(\mathscr{E})$ satisfies $\mathscr{E}(h,h) = 0$, then $h = 0$, $\mu - a.s.$.
Motivated by the concentration inequality for the empirical mean $\frac{1}{t}\int_0^t g(X_s)ds$ for a family $\mathscr{A}$ of bounded observables g, Guillin *et al.* [8] introduced the following transportation-information inequality

$$\alpha\left(\sup_{g\in\mathscr{A}}\left[\nu(g) - \mu(g)\right]\right) \leq I(\nu|\mu),\ \forall \nu \in \mathscr{M}_1^1(\mathscr{X}), \tag{2}$$

where $\alpha : \mathbb{R} \to [0,+\infty)$ is some non-decreasing and left-continuous function with $\alpha(0) = 0$. When $\mathscr{A}$ is the family of all bounded measurable and d-Lipschitzian functions g with $\|g\|_{\mathrm{Lip}(d)} := \sup_{x,y\in\mathscr{X}} \frac{|g(x)-g(y)|}{d(x,y)} \leq 1$, the previous inequality becomes by the Kantorovitch-Rubinstein duality,

$$\alpha(W_{1,d}(\nu,\mu)) \leq I(\nu|\mu),\ \forall \nu \in \mathscr{M}_1^1(\mathscr{X}). \tag{3}$$

More precisely Guillin *et al.* [8] obtained

Theorem 1.1. ([8, Theorem 2.4] or [5, Theorem 2.2]) *Let $\alpha : \mathbb{R} \to [0,+\infty)$ be some non-decreasing and left-continuous function with $\alpha(0) = 0$. Given a family $\mathscr{A}$ of bounded measurable functions g (say $g \in b\mathscr{B}$), the following properties are equivalent:*

(a) *The transportation-information inequality (2) holds.*

(b) *The following concentration inequality holds for each $g \in \mathscr{A}$ and any initial distribution $\nu \ll \mu$,*

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t g(X_s)ds > \mu(g) + r\right) \leq \|\frac{d\nu}{d\mu}\|_2 e^{-t\alpha(r)},\ \forall\, t, r > 0. \tag{4}$$

Here $\|\cdot\|_2$ is the norm of $L^2(\mu)$.

In particular, the $W_1 I$-inequality (3) is equivalent to

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t g(X_s)ds > \mu(g)+r\right) \le \|\frac{d\nu}{d\mu}\|_2 e^{-t\alpha(r)},\ \forall\, t,\, r>0 \tag{5}$$

for all $g \in b\mathscr{B}$ with $\|g\|_{\mathrm{Lip}(d)} \le 1$.

Recently, Gao and the third named author [6] proved a tensorization result for the Wasserstein distance (see Lemma 4.2 below) and established the “dimension-free" transportation-information inequalities $W_p I(p \ge 1)$ for the discrete Gibbs measure, under the Dobrushin's uniqueness condition ([3, 4]).

1.2 Continuum Gibbs measure and generator of the Glauber dynamic

Let $\mathscr{B}(\mathbb{R}^d)$ be the Borel $\sigma-$algebra on $\mathbb{R}^d$ $(d \ge 1)$. We denote by $\mathscr{B}_b(\mathbb{R}^d) \subset \mathscr{B}(\mathbb{R}^d)$ the collection of all bounded Borel sets. For each $A \in \mathscr{B}_b(\mathbb{R}^d)$, $|A|$ denotes the Lebesgue measure of A. We consider, as configuration space, the set Ω of all locally finite point measures on $\mathbb{R}^d$, i.e.,

$$\Omega := \left\{\omega = \sum_i \delta_{x_i} : \omega(A) < \infty \text{ for all } A \in \mathscr{B}_b(\mathbb{R}^d)\right\}.$$

with the $\sigma-$algebra $\mathscr{F}$ generated by the counting variables $N_A : \omega \to \omega(A)$, where $A \in \mathscr{B}_b(\mathbb{R}^d)$. Given the *activity* $z > 0$ (the name “*activity*" comes from Ruelle [18]), let P be the law of Poisson point process on $\mathbb{R}^d$ with intensity measure zdx.
Letting Λ be a bounded open subset of $\mathbb{R}^d$, we consider also the finite volume configuration space

$$\Omega_\Lambda := \{\omega \in \Omega : \mathrm{supp}(\omega) \subset \Lambda\} \tag{6}$$

with $\sigma-$algebra $\mathscr{F}_\Lambda$ generated by the function N_A, where A runs over the Borel $\sigma-$field of Λ and $\omega_\Lambda = \sum_{x\in\ \mathrm{supp}\omega\cap\Lambda}\delta_x$. The image measure P_Λ of P by $\omega \to \omega_\Lambda$ is the law of Poisson point process on Λ with intensity measure zdx. The configuration space Ω_Λ under the Prohorov metric, with the weak convergence topology, is a Polish space.
We say that an element η of Ω is a boundary condition on Λ^c, if

$$\eta = \sum_{k=1}^{+\infty}\delta_{y_k},\ y_k \in \Lambda^c,\ k \in \mathbb{N}.$$

Let $\varphi : \mathbb{R}^d \to \mathbb{R}^+ \cup \{+\infty\}$ be a nonnegative measurable even function, representing a repulsive pair interaction. The finite volume Gibbs measure in Λ for a given boundary condition η, at inverse temperature $\beta > 0$, is given by

$$\mu_\Lambda^\eta(d\omega_\Lambda) := (Z_\Lambda^\eta)^{-1}\exp\left\{-\beta H_\Lambda^\eta(\omega_\Lambda)\right\}P_\Lambda(d\omega_\Lambda) \tag{7}$$

where Z_Λ^η is the normalization constant and

$$H_\Lambda^\eta(\omega_\Lambda) := \frac{1}{2}\int_{\Lambda^2}\varphi(x-y)\omega_\Lambda(dx)\omega_\Lambda(dy) + \int_\Lambda \omega_\Lambda(dx)\int_{\Lambda^c}\varphi(x-y)\eta(dy)$$

is the Hamiltonian in Λ. This is the mathematical model for continuous gas in statistical physics, see the book of Ruelle [18].

Let $r\mathscr{F}$ be the space of real $\mathscr{F}-$measurable functions, and $b\mathscr{F}$ be the space of those $F \in r\mathscr{F}$ which are moreover bounded. For any $f \in r\mathscr{F}$, following Picard [16], consider the difference operators

$$\begin{aligned} D_x^+ f(\omega) &:= f(\omega + \delta_x) - f(\omega), \\ D_x^- f(\omega) &:= 1_{x\in \mathrm{supp}(\omega)}[f(\omega - \delta_x) - f(\omega)]. \end{aligned} \tag{8}$$

Recall that D_x^+ plays the same role in the Malliavin calculus over the Poisson space as the Malliavin derivative on the Wiener space ([16, 19] and references therein).

We shall work on the Glauber dynamic, which is formally generated by the pre-generator (see [1, 12, 20])

$$\mathscr{L}_\Lambda^\eta f(\omega_\Lambda) = \int_\Lambda D_x^- f(\omega_\Lambda)\omega_\Lambda(dx) + z\int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} D_x^+ f(\omega_\Lambda) dx, \quad f \in b\mathscr{F}_\Lambda. \tag{9}$$

It is easily checked that for all $f, g \in b\mathscr{F}_\Lambda$

$$\begin{aligned} \langle f, -\mathscr{L}_\Lambda^\eta g\rangle_{\mu_\Lambda^\eta} &= \int_{\Omega_\Lambda} d\mu_\Lambda^\eta(\omega_\Lambda) \int_\Lambda D_x^- f(\omega_\Lambda) D_x^- g(\omega_\Lambda)\omega_\Lambda(dx) \\ &= \int_{\Omega_\Lambda} d\mu_\Lambda^\eta(\omega_\Lambda) \int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} D_x^+ f(\omega_\Lambda) D_x^+ g(\omega_\Lambda) z dx \\ &=: \mathscr{E}_\Lambda^\eta(f, g). \end{aligned} \tag{10}$$

Then $(-\mathscr{L}_\Lambda^\eta, b\mathscr{F}_\Lambda)$ is a nonnegative definite, symmetric operator on $L^2(\mu_\Lambda^\eta)$ (indeed it is essentially self-adjoint by Kondratiev and Lytvynov [12]). Hence $\mathscr{E}_\Lambda^\eta$ is a closable form and its closure $(\mathscr{E}_\Lambda^\eta, \mathbb{D}(\mathscr{E}_\Lambda^\eta))$ is a Dirichlet form on $L^2(\mu_\Lambda^\eta)$, generating a symmetric Markov semigroup $(P_t^{\Lambda,\eta})_{t\geq 0}$ on $L^2(\mu_\Lambda^\eta)$ such that $P_t^{\Lambda,\eta}1 = 1, \mu_\Lambda^\eta - a.s.$, associated with a reversible Markov process $((X_t^{\Lambda,\eta})_{t\geq 0}, \mathbb{P}_{\mu_\Lambda^\eta})$ such that its sample paths are $\mathbb{P}_{\mu_\Lambda^\eta}-$càdlàg. $(P_t^{\Lambda,\eta})_{t\geq 0}$ is a strongly continuous semigroup of contractions on $L^2(\mu_\Lambda^\eta)$, whose generator will be denoted by $(\mathscr{L}_\Lambda^\eta, \mathbb{D}(\mathscr{L}_\Lambda^\eta))$ ($\mathbb{D}(\mathscr{L}_\Lambda^\eta)$ being its domain in $L^2(\mu_\Lambda^\eta)$).
This dynamic, as a classical probabilistic model in statistical mechanics, was first introduced and studied by Preston in [17]. Bertini *et al.* [1] established the existence of a spectral gap, which is uniformly positive in the volume and boundary conditions, for the Glauber dynamic in the high temperature-low activity regime. The third named author [20] improved their work and extended to the hard core case by Poissonian approximation and Liggett's $M - \epsilon$ theorem for lattice gas. Kondratiev and Lytvynov [12] also obtained independently the spectral gap estimate in [20], by a different and simpler method.
In this paper we will always work on finite volume case for two reasons: 1) our $W_1 I$-inequality explodes in the infinite volume case even in the free case; 2) all interesting physical quantities (such as mean number of particles per unit volume) in the infinite volume case are calculated by approximation via finite volume ([18]).

Objective and organization. The objective of this paper is to establish some explicit transportation-information inequality $W_1 I$ for the Glauber dynamic above related with the continuum Gibbs measure μ_Λ^η, under the Dobrushin's uniqueness condition (cf. [20])

$$D := z\int_{\mathbb{R}^d} \left(1 - e^{-\beta\varphi(y)}\right) dy < 1. \tag{11}$$

As an interesting prelude to this end, we begin with the $M/M/\infty$ queue system in §2 (the jumps counterpart of the Ornstein-Uhlenbeck process), for which the optimal transportation-information inequality is obtained by means of the Lipschitzian spectral gap and Lyapunov test function method, improving some previous known results. In section 3, by generalizing the arguments of section 2, we obtain explicit W_1I inequality for the continuum Gibbs measure μ_Λ^η, under the Dobrushin's uniqueness condition. Section 4 is devoted to the discrete spin system. For this model we establish W_1I-inequality by the tensorization technique in Gao and Wu [6].

2 $M/M/\infty$ queue system

For the simplicity and the clarity of our presentation we begin with a simple model: $M/M/\infty$ queue system. Let μ be the Poisson measure with mean $\lambda > 0$ on $\mathbb{N}$ equipped with the Euclidean distance ρ. For each bounded measurable function f on $\mathbb{N}$, consider the Dirichlet form

$$\mathcal{E}(f,f) = \lambda \sum_{n\in\mathbb{N}} (f(n+1) - f(n))^2 \mu(n) \tag{12}$$

and the corresponding generator (with the convention $f(-1) := f(0)$)

$$\mathcal{L} f(n) = \lambda(f(n+1) - f(n)) + n(f(n-1) - f(n)), \ \forall n \in \mathbb{N}.$$

It is an ideal model for a queue system with a number of servers much larger than the number of clients (such as in an automatic computer service center). It is well known that the Poincaré constant c_P equals 1, but the log-Sobolev inequality does not hold (see [19]).

Theorem 2.1. *With respect to the Euclidean metric $\rho(x,y) = |x-y|$ on $\mathbb{N}$, for the Poisson measure μ with mean $\lambda > 0$, the following W_1I-inequality holds true:*

$$W_{1,\rho}(\nu,\mu) \le 2\sqrt{\lambda I} + I, \quad \forall \nu \in \mathcal{M}_1^1(\mathbb{N}), \tag{13}$$

where $I = I(\nu|\mu)$. This inequality is of the form (3) *with $\alpha(r) = (\sqrt{\lambda + r} - \sqrt{\lambda})^2$, which is optimal.*

Remark 2.2. By Theorem 1.1 the W_1I inequality (13) is equivalent to the following concentration inequality of Bernstein type: for any $g : \mathbb{N} \to \mathbb{R}$ with $\|g\|_{\mathrm{Lip}(\rho)} = 1$ and $\mu(g) = 0$,

$$\mathbb{P}_\nu\left(\frac{1}{t}\int_0^t g(X_s)ds > 2\sqrt{\lambda x} + x\right) \le \|\frac{d\nu}{d\mu}\|_2 e^{-tx}, \ \forall\, t, x > 0$$

for any initial measure $\nu \ll \mu$. For the function $g_0(n) := n - \lambda$, Gao *et al.* [5] showed that

$$\nu(g_0) - \mu(g_0) \le 2\sqrt{\lambda I} + I, \ I := I(\nu|\mu), \ \forall \nu \in \mathcal{M}_1^1(\mathbb{N})$$

is optimal (our result is motivated by this fact, of course). A different but direct way to see the optimality of (13) is to take ν as the Poisson measure with parameter $a\lambda$ where $a > 1$: $W_{1,\rho}(\nu,\mu) = \lambda(a-1)$ and $I := I(\nu|\mu) = \lambda[\sqrt{a} - 1]^2$. Then (13) becomes equality for such ν.

Remark 2.3. The optimal transportation-information inequality (13) is a definite improvement on the existing results on this model obtained by Gao *et al.* [5], Gao and Wu [6]. However our proof is largely inspired by those general works. For other known concentration inequalities on this model, see Joulin [10], Liu and Ma [13], Joulin and Ollivier [11] (for numerous other interesting models too). Chafaï [2] obtained the Φ-Sobolev inequalities (including the L^1-log-Sobolev inequalities) for the $M/M/\infty$ queue.

Proof of Theorem 2.1. Step 1. Lipschitzian spectral gap. First of all, we claim that

$$\|(-\mathscr{L})^{-1}\|_{\mathrm{Lip}(\rho)} := \sup_{\|g\|_{\mathrm{Lip}(\rho)}=1} \|(-\mathscr{L})^{-1}g\|_{\mathrm{Lip}(\rho)} = 1 \tag{14}$$

for this model. The simplest way to see this known fact is to remark the following commutation relation between the generator $\mathscr{L}$ and the difference operator $DG(n) := G(n+1) - G(n)$ (for a function G on $\mathbb{N}$):

$$D\mathscr{L}G = \mathscr{L}DG - DG.$$

Given any $g : \mathbb{N} \to \mathbb{R}$ with $\|g\|_{\mathrm{Lip}(\rho)} = 1$ and $\mu(g) = 0$, if $-\mathscr{L}G = g$, then $(1-\mathscr{L})DG = Dg$. By the resolvent of the infinitesimal generator $\mathscr{L}$, for any f with $\|f\|_\infty = 1$

$$\|(1-\mathscr{L})^{-1}f\|_\infty = \|\int_0^\infty e^{-s}P_s f\, ds\|_\infty \le \int_0^\infty e^{-s} ds = 1,$$

where it follows by taking $f \equiv 1$

$$\|(1-\mathscr{L})^{-1}\|_\infty := \sup_{\|f\|_\infty=1} \|(1-\mathscr{L})^{-1}f\|_\infty = 1.$$

Hence

$$\|G\|_{\mathrm{Lip}(\rho)} = \|DG\|_\infty \le \|-(1-\mathscr{L})^{-1}\|_\infty \cdot \|Dg\|_\infty = 1$$

and this inequality becomes equality if $Dg = 1$ (i.e. $g(n) = g_0(n) = n - \lambda$). That shows the fact.

Step 2. Lyapunov function method. For (13) we may assume that $\nu = f\mu$ with $\sqrt{f} \in \mathbb{D}(\mathscr{E})$ and $I := I(\nu|\mu) = \mathscr{E}(\sqrt{f}, \sqrt{f}) > 0$.

Given any function g on $\mathbb{N}$ with $\mu(g) = 0$ and $\|g\|_{\mathrm{Lip}(\rho)} = 1$, let G be the solution to the Poisson equation $-\mathscr{L}G = g$ with $\mu(G) = 0$. For any $\delta > 0$, we have (these few lines are the starting point of our approach)

$$\begin{aligned}
\nu(g) - \mu(g) &= \langle g, f\rangle_\mu = \mathscr{E}(G, f) \\
&= \sum_{n=0}^\infty \lambda\mu(n)(G(n+1) - G(n))(f(n+1) - f(n)) \\
&\le \sqrt{\sum_{n=0}^\infty \lambda\mu(n)(\sqrt{f(n+1)} - \sqrt{f(n)})^2} \\
&\quad \cdot \sqrt{\sum_{n=0}^\infty \lambda\mu(n)(G(n+1) - G(n))^2(\sqrt{f(n+1)} + \sqrt{f(n)})^2} \\
&\le \sqrt{I}\sqrt{\sum_{n=0}^\infty \lambda\mu(n)\Big((1+\delta)f(n+1) + (1+\frac{1}{\delta})f(n)\Big)}.
\end{aligned}$$

where the last inequality relies on the fact that $\|(-\mathscr{L})^{-1}\|_{\mathrm{Lip}(\rho)} = 1$ in Step 1, $\|g\|_{\mathrm{Lip}(\rho)} = 1$ and the elementary inequality $(x+y)^2 \le (1+\delta)x^2 + (1+\delta^{-1})y^2$ for any $x, y \in \mathbb{R}, \delta > 0$. The last term in the square root above, denoted by B, is (using $\lambda\mu(n) = (n+1)\mu(n+1)$)

$$\begin{aligned}
B &= (1+\frac{1}{\delta})\lambda\sum_{n=0}^\infty \mu(n)f(n) + (1+\delta)\sum_{n=0}^\infty (n+1)\mu(n+1)f(n+1) \\
&= \sum_{n=0}^\infty \mu(n)f(n)\Big((1+\delta)n + (1+\frac{1}{\delta})\lambda\Big).
\end{aligned}$$

We now employ the method of Lyapunov test function developed in Guillin *et al.* [8] for bounding the last term. The basic fact behind this approach is : for any function $V \geq 1$, if $-\frac{\mathscr{L}V}{V}$ is bounded from below, then

$$\int -\frac{\mathscr{L}V}{V} d\nu \leq I(\nu|\mu),\ \forall \nu \in \mathscr{M}_1^1(\mathbb{N}). \tag{15}$$

That was proved in [8, Lemma 5.6] for general reversible Markov processes. Our task now is to find a good function V such that

$$(1+\delta)n + \big(1+\frac{1}{\delta}\big)\lambda \leq -a\frac{\mathscr{L}V}{V}(n) + b \tag{16}$$

for two positive constants a, b, and (15) will imply

$$B \leq aI + b.$$

Taking $V(n) = \kappa^n$ for some constant $\kappa > 1$, the previous inequality holds with $a = (1+\delta)\kappa/(\kappa-1)$ and $b = \big((1+\delta)\kappa + (1+\frac{1}{\delta})\big)\lambda$ by simple algebra. As $\delta > 0, \kappa > 1$ are arbitrary, we get

$$\begin{aligned} \nu(g) - \mu(g) &\leq \sqrt{I} \inf_{\kappa>1,\delta>0} \sqrt{I(1+\delta)\kappa/(\kappa-1) + \big((1+\delta)\kappa + (1+\frac{1}{\delta})\big)\lambda} \\ &= I + 2\sqrt{\lambda I} \end{aligned}$$

where the equality is attained at $\kappa = 1 + \sqrt{I/\lambda}$ and $\delta = \kappa^{-1}$. Therefore the desired transportation-information inequality (22) follows by taking the supremum over all functions g such that $\mu(g) = 0$ and $\|g\|_{\mathrm{Lip}(\rho)} = 1$.

Remark 2.4. Given an increasing function w on $\mathbb{N}$ which induces a metric ρ_w as $\rho_w(x,y) = |w(x) - w(y)|$, the Lipschitzian norm of the Poisson operator $\|(-\mathscr{L})^{-1}\|_{\mathrm{Lip}(\rho_w)}$ is known for a general birth-death process (i.e. $\mathscr{L}f(n) = b_n(f(n+1)-f(n)) + a_n(f(n-1)-f(n))$ with the birth rate $b_n > 0$ for any $n \geq 0$ and the death rate $a_0 = 0$, $a_n > 0$ for any $n \geq 1$), due to Liu and the first named author [13]. In fact, consider the corresponding Poisson equation

$$-\mathscr{L}\varphi = w - \mu(w), \tag{17}$$

which admits a unique and explicit solution φ with zero mean ([13]). Theorem 2.1 in [13] says that

$$\|(-\mathscr{L})^{-1}\|_{\mathrm{Lip}(\rho_w)} = \|\varphi\|_{\mathrm{Lip}(\rho_w)}. \tag{18}$$

This fact together with the Lyapunov test function method above can produce the W_1I inequality for quite general birth-death processes. Notice also that for $w(n) = g_0(n) = n - \lambda$, the previous identification (18) of $\|(-\mathscr{L})^{-1}\|_{\mathrm{Lip}(\rho_w)}$ gives the result of Step 1 above, for $\varphi = g_0$.

3 W_1I-inequality for continuum Gibbs measure

In this section we generalize the arguments in §2 to study the W_1I-inequality for the continuum Gibbs measure μ_Λ^η.

3.1 Lipschitzian norm of $(-\mathscr{L}_\Lambda^\eta)^{-1}$

We consider the total variation metric d on Ω_Λ : for any $\omega, \omega' \in \Omega_\Lambda$,

$$d(\omega, \omega') = \|\omega - \omega'\|_{\mathrm{TV}}. \tag{19}$$

Given any functional $F \in r\mathscr{F}_\Lambda$, we call F is Lipschitzian with respect to d if

$$\|F\|_{\mathrm{Lip}(d)} := \sup_{\omega \neq \omega'} \frac{|F(\omega) - F(\omega')|}{d(\omega, \omega')} < \infty.$$

By Lemma 2.2 in [14],

$$\|F\|_{\mathrm{Lip}(d)} = \sup_{x \in \Lambda, \omega_\Lambda \in \Omega_\Lambda} |D_x^+ F(\omega_\Lambda)|. \tag{20}$$

Denote by $\mathbf{C}_{\mathrm{Lip}}^{\mathbf{0}}$ the set of functionals $F \in r\mathscr{F}_\Lambda$ with $\|F\|_{\mathrm{Lip}(d)} < \infty$ and $\mu_\Lambda^\eta(F) = 0$.
Recall the usual Lipschitzian norm of $(-\mathscr{L}_\Lambda^\eta)^{-1}$ on C_{Lip}^0:

$$\|(-\mathscr{L}_\Lambda^\eta)^{-1}\|_{\mathrm{Lip}(d)} = \sup_{\|g\|_{\mathrm{Lip}(d)} \leq 1} \|(-\mathscr{L}_\Lambda^\eta)^{-1} g\|_{\mathrm{Lip}(d)}. \tag{21}$$

First we give a key lemma which provides a sharp estimate of the Lipschitzian norm of $(-\mathscr{L}_\Lambda^\eta)^{-1}$ and which is essentially due to the third named author [20].

Lemma 3.1. *Suppose that the Dobrushin's uniqueness condition holds, i.e.,*

$$D = z \int_{\mathbb{R}^d} (1 - e^{-\beta\varphi(x)}) dx < 1.$$

We have

$$\|(-\mathscr{L}_\Lambda^\eta)^{-1}\|_{\mathrm{Lip}(d)} \leq \frac{1}{1-D}.$$

Proof. By Theorem 5.1 in [20], for any functional $F \in b\mathscr{F}_\Lambda \cap C_{\mathrm{Lip}}^0$,

$$\|P_t^{\Lambda,\eta} F\|_{\mathrm{Lip}(d)} \leq e^{-(1-D)t} \|F\|_{\mathrm{Lip}(d)}.$$

Hence

$$\begin{aligned}
\|(-\mathscr{L}_\Lambda^\eta)^{-1} F\|_{\mathrm{Lip}(d)} &= \sup_{x \in \Lambda, \omega_\Lambda \in \Omega_\Lambda} |D_x^+ (-\mathscr{L}_\Lambda^\eta)^{-1} F(\omega_\Lambda)| \\
&= \sup_{x \in \Lambda, \omega_\Lambda \in \Omega_\Lambda} |D_x^+ \int_0^\infty P_t^{\Lambda,\eta} F(\omega_\Lambda) dt| \\
&\leq \int_0^\infty \sup_{x \in \Lambda, \omega_\Lambda \in \Omega_\Lambda} |D_x^+ P_t^{\Lambda,\eta} F(\omega_\Lambda)| dt \\
&\leq \int_0^\infty e^{-(1-D)t} dt \|F\|_{\mathrm{Lip}(d)} = \frac{1}{1-D} \|F\|_{\mathrm{Lip}(d)}.
\end{aligned}$$

For general $F \in C_{\mathrm{Lip}}^0$, let $F_n = (F \wedge n) \vee (-n)$, we can approximate F by $F_n - \mu_\Lambda^\eta(F_n)$, then the desired result follows.

3.2 $W_1 I$-inequality

The main result of this paper is the following theorem

Theorem 3.2. *For the continuum Gibbs measure μ_Λ^η given in (7) with the nonnegative even pair interaction φ, suppose that the Dobrushin's uniqueness condition holds, i.e.,*

$$D = z \int_{\mathbb{R}^d} (1 - e^{-\beta\varphi(x)}) dx < 1.$$

Then the transportation-information inequality below holds

$$W_{1,d}(\nu, \mu_\Lambda^\eta) \le \frac{1}{1-D}\left(I + 2\sqrt{z|\Lambda| I}\right), \quad \forall \nu \in \mathscr{M}_1^1(\Omega_\Lambda) \tag{22}$$

where $I = I(\nu|\mu_\Lambda^\eta)$ is the Fisher-Donsker-Varadhan's information related with $\mathscr{E}_\Lambda^\eta$ given in (10) and the metric d is the total variation metric defined in (19).

Remark 3.3. When $\varphi = 0$ (no interaction case), the inequality (22) is optimal. Since in this case $D = 0$ and $N_\Lambda(X_t)$ is just the $M/M/\infty$ queue with $\lambda = z|\Lambda|$ and then Theorem 2.1 guarantees its optimality.

Remark 3.4. Since the Lipschitzian norm w.r.t. d of $F(\omega) = \frac{1}{|\Lambda|} N_\Lambda(\omega)$ (the mean number of particles per unit volume of ω) is $1/|\Lambda|$, hence by (22) and Theorem 1.1 we have for all $t, r > 0$ and initial distribution $\nu \ll \mu_\Lambda^\eta$,

$$\mathbb{P}_\nu\left(\frac{1}{t|\Lambda|}\int_0^t N_\Lambda(X_s)ds - \frac{\mu_\Lambda^\eta(N_\Lambda)}{|\Lambda|} > r\right) \le \|\frac{d\nu}{d\mu}\|_2 \exp\left(-t|\Lambda|\left[\sqrt{z + (1-D)r} - \sqrt{z}\right]^2\right).$$

This concentration inequality shows that the Glauber dynamics here is a very efficient tool for estimating $\mu_\Lambda^\eta(N_\Lambda)/|\Lambda|$.

The same argument as Theorem 2.1, namely estimating $\|(-\mathscr{L}_\Lambda^\eta)^{-1}\|_{\mathrm{Lip}(d)}$ plus Lyapunov condition (16), works for proving Theorem 3.2. Then with Lemma 3.1, it remains to find some good function V such that Lyapunov condition is verified. For this aim, we begin by introducing the generalized domain $\mathbb{D}_e(\mathscr{L}_\Lambda^\eta)$.

A continuous function h is said to be in the $\mu_\Lambda^\eta-$extended domain $\mathbb{D}_e(\mathscr{L}_\Lambda^\eta)$ of the generator of the Markov process $((X_t^{\Lambda,\eta}), \mu_\Lambda^\eta)$ if there is some measurable function g such that $\int_0^t |g|(X_s^{\Lambda,\eta})ds < +\infty, \mu_\Lambda^\eta-$a.s., and

$$M_t := h(X_t^{\Lambda,\eta}) - h(X_0^{\Lambda,\eta}) - \int_0^t g(X_s^{\Lambda,\eta})ds$$

is a local μ_Λ^η-martingale. It is obvious that g is uniquely determined up to $\mu_\Lambda^\eta-$equivalence. In such case one writes $h \in \mathbb{D}_e(\mathscr{L}_\Lambda^\eta)$ and $\mathscr{L}_\Lambda^\eta h = g$.

Lemma 3.5. *There exists a function $V : \Omega_\Lambda \to [1, \infty)$ in $\mathbb{D}_e(\mathscr{L}_\Lambda^\eta)$ such that for any $\delta > 0$,*

$$\begin{aligned} &(1+\delta)N_\Lambda(\omega_\Lambda) + (1+\frac{1}{\delta})z|\Lambda| \le -a\frac{\mathscr{L}_\Lambda^\eta V(\omega_\Lambda)}{V(\omega_\Lambda)} + b, \quad \omega_\Lambda \in \Omega_\Lambda \\ &a = (1+\delta)\frac{\kappa}{\kappa-1}, \quad b = \left((1+\delta)\kappa + (1+\frac{1}{\delta})\right)z|\Lambda|. \end{aligned} \tag{23}$$

Proof. For a constant $\kappa > 1$, take $V(\omega_\Lambda) = \kappa^{N_\Lambda(\omega_\Lambda)}$. Then

$$-\frac{\mathscr{L}_\Lambda^\eta V(\omega_\Lambda)}{V(\omega_\Lambda)} = (1-\kappa^{-1})N_\Lambda(\omega_\Lambda) - (\kappa-1)z\int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)}dx.$$

As $\varphi \geq 0$, we see that (23) holds.
Proof of Theorem 3.2 In order to establish (22), we may assume that $\nu = f\mu_\Lambda^\eta$ with $\sqrt{f} \in \mathbb{D}(\mathscr{E})$ and $I = I(\nu|\mu_\Lambda^\eta) > 0$.
Given any $g \in C_{\text{Lip}}^0$ with $\|g\|_{\text{Lip}(d)} = 1$, let $G = (-\mathscr{L}_\Lambda^\eta)^{-1}g$. By Cauchy-Schwarz inequality and (10), we have

$$\begin{aligned}
\nu(g) - \mu_\Lambda^\eta(g) &= \langle g, f\rangle_{\mu_\Lambda^\eta} = \langle -\mathscr{L}_\Lambda^\eta G, f\rangle_{\mu_\Lambda^\eta} = \mathscr{E}_\Lambda^\eta(G,f)\\
&= \int_{\Omega_\Lambda} d\mu_\Lambda^\eta \int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} D_x^+ G(\omega_\Lambda) D_x^+ f(\omega_\Lambda) z dx\\
&\leq \sqrt{I}\sqrt{\int_{\Omega_\Lambda} d\mu_\Lambda^\eta \int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} (D_x^+ G(\omega_\Lambda))^2 \left(\sqrt{f(\omega_\Lambda+\delta_x)} + \sqrt{f(\omega_\Lambda)}\right)^2 z dx}.
\end{aligned}$$

We treat the term in the last square root as in the proof of Theorem 2.1,

$$\begin{aligned}
&\int_{\Omega_\Lambda} d\mu_\Lambda^\eta \int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} (D_x^+ G(\omega_\Lambda))^2 \left(\sqrt{f(\omega_\Lambda+\delta_x)} + \sqrt{f(\omega_\Lambda)}\right)^2 z dx\\
&\leq \int_{\Omega_\Lambda} d\mu_\Lambda^\eta \int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} (D_x^+ G(\omega_\Lambda))^2 \left((1+\delta)f(\omega_\Lambda+\delta_x) + (1+\frac{1}{\delta})f(\omega_\Lambda)\right) z dx\\
&\leq \|G\|_{\text{Lip}(d)}^2 \int_{\Omega_\Lambda} d\mu_\Lambda^\eta \int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} \left((1+\delta)f(\omega_\Lambda+\delta_x) + (1+\frac{1}{\delta})f(\omega_\Lambda)\right) z dx\\
&= \|G\|_{\text{Lip}(d)}^2 \int_{\Omega_\Lambda} f(\omega_\Lambda) d\mu_\Lambda^\eta \left((1+\delta)\int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda - \delta_x)}\omega_\Lambda(dx) + (1+\frac{1}{\delta})\int_\Lambda e^{-\beta D_x^+ H_\Lambda^\eta(\omega_\Lambda)} z dx\right)\\
&\leq \frac{1}{(1-D)^2}\int_{\Omega_\Lambda}\left((1+\delta)N_\Lambda(\omega_\Lambda) + (1+\frac{1}{\delta})z|\Lambda|\right)\nu(d\omega_\Lambda)\\
&\leq \frac{1}{(1-D)^2}\int_{\Omega_\Lambda}\left(-a\frac{\mathscr{L}_\Lambda^\eta V(\omega_\Lambda)}{V(\omega_\Lambda)} + b\right)\nu(d\omega_\Lambda)\\
&\leq \frac{1}{(1-D)^2}\left(aI + b\right),
\end{aligned}$$

where $\delta > 0$ is arbitrary, the third crucial equality is due to the duality formula in the Malliavin calculus on the Poisson space ([16]) saying for any measurable functional $F : \Omega_\Lambda \times \Lambda \mapsto [0, +\infty]$,

$$\int_{\Omega_\Lambda} d\mu_\eta^\Lambda \int_\Lambda \omega_\Lambda(dx)F(\omega_\Lambda, x) = \int_{\Omega_\Lambda} d\mu_\Lambda^\eta \int_\Lambda \exp\{-\beta E(x,\omega_\Lambda)\}F(\omega_\Lambda+\delta_x, x) z dx$$

with

$$E(x,\omega_\Lambda) := \begin{cases} \int_\Lambda \varphi(x-y)\omega_\Lambda(dy), & \text{if } \int_\Lambda |\varphi(x-y)|\omega_\Lambda(dy) < \infty;\\ +\infty, & \text{otherwise ,}\end{cases}$$

the fourth inequality is true by the Lipschitzian spectral gap estimate in Lemma 3.1, the last but second inequality is an application of (23) with constants a, b given there and the last one follows by [8, Lemma 5.6] as recalled in (15).
Now by the same optimization procedure over $\kappa > 1, \delta > 0$ as in the proof of Theorem 2.1, we obtain

$$\nu(g) - \mu_\Lambda^\eta(g) \le \frac{1}{1-D}\left(I + 2\sqrt{z|\Lambda|I}\right)$$

where the desired result (22) follows, since g in C^0_{Lip} with $\|g\|_{\mathrm{Lip}(d)} = 1$ is arbitrary.

4 $W_1 I$-inequality for the discrete spin system

The discrete spin system and the Dobrushin's interdependence coefficient. Let T be a finite subset of $\mathbb{Z}^d$ and $\gamma : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^+$ be a nonnegative interaction function satisfying $\gamma_{ij} = \gamma_{ji}$ and $\gamma_{ii} = 0$ for all $i, j \in \mathbb{Z}^d$. The Gibbs measure on $\mathbb{N}^T$ with boundary condition $(x_k)_{k\in T^c}$ is defined by

$$\mu_T(dx_T|x) = \frac{e^{-\frac{1}{2}\sum_{\{i,j\}\cap T\neq\emptyset}\gamma_{ij}x_i x_j}}{Z(x_{T^c})}\Pi_{i\in T}\sigma_{\lambda_i}(dx_i) \tag{24}$$

where $\left\{\sigma_{\lambda_i}(\cdot)\right\}_{i\in\mathbb{Z}^d}$ are the given Poisson measures on $\mathbb{N}$ with means $\{\lambda_i > 0\}_{i\in\mathbb{Z}^d}$, and $Z(x_{T^c})$ is the normalization factor. When $T = \{i\}$, $\mu_T(dx_T|x)$ is simply denoted by $\mu_i := \mu_i(dx_i|x)$, which is the conditional distribution of x_i knowing $(x_j)_{j\neq i}$. In the present case, $\mu_i(dx_i|x)$ is the Poisson distribution $\mathscr{P}(\lambda_i e^{-\sum_{j\neq i}\gamma_{ij}x_j})$ with parameter $\lambda_i e^{-\sum_{j\neq i}\gamma_{ij}x_j}$.
The purpose of this section is to propose another approach : tensorization technique, to establish the $W_1 I$-inequality for the discrete Gibbs measure $\mu_T(dx_T|x)$ from (13) for Poisson measure. For this dependent tensorization, the key tool is the Dobrushin's interdependence matrix $C := (c_{ij})_{i,j\in T}$ w.r.t. the Euclidean metric ρ on $\mathbb{N}$, defined by

$$c_{ij} = \sup_{x=x'\mathrm{off} j}\frac{W_{1,\rho}\left(\mu_i(dx_i|x), \mu_i(dx_i'|x')\right)}{|x_j - x_j'|}, \quad \forall\, i, j \in \mathbb{Z}^d \tag{25}$$

(obviously $c_{ii} = 0$). Then the Dobrushin's uniqueness condition [3, 4] is

$$D := \sup_{j\in T}\sum_{i\in T} c_{ij} < 1. \tag{26}$$

The Dobrushin's interdependence coefficient c_{ij} can be easily identified for this model.

Lemma 4.1. *([14, Lemma 3.1]) For $i \neq j$ in $\mathbb{Z}^d$,*

$$c_{ij} = \lambda_i(1 - e^{-\gamma_{ij}}). \tag{27}$$

The transportation-information inequality $W_1 I$ for the discrete spin system. Consider the metric

$$d_{l^1}(x, y) := \sum_{i\in T}|x_i - y_i|, \quad \forall x, y \in \mathbb{N}^T \tag{28}$$

on $\mathbb{N}^T$. The following disintegration of W_1-metric is our starting point.

Lemma 4.2. (*Gao-Wu [6, Theorem 3.1]*) *Let μ_T be the discrete Gibbs measure given in (24). Assume the Dobrushin's uniqueness condition*

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

Then for all $\nu_T \in \mathscr{M}_1^1(\mathbb{N}^T)$,

$$W_{1,d_{l^1}}(\nu_T, \mu_T) \le \frac{1}{1-D} \mathbb{E}^{\nu_T} \sum_{i\in T} W_{1,\rho}(\nu_i, \mu_i) \tag{29}$$

where ν_i is the conditional distribution of x_i knowing $(x_j)_{j\neq i}$.

We now introduce the Glauber dynamic. For each $i \in T$ and $\hat{x}_i := x_{T\setminus\{i\}}$ fixed, consider the site's Dirichlet form associated with the Poisson measure $\mu_i(dx_i|x)$:

$$\mathscr{E}_i(f,f) := \lambda_i e^{-\sum_{j\neq i}\gamma_{ij}x_j} \sum_{x_i\in\mathbb{N}} (f(x_i+1) - f(x_i))^2 \mu_i(x_i|x),$$
$$\mathbb{D}(\mathscr{E}_i) := \{f \in L^2(\mu_i);\ \mathscr{E}_i(f,f) < +\infty\}.$$

which corresponds to the $M/M/\infty$ queue with parameter $\lambda = \lambda_i e^{-\sum_{j\neq i}\gamma_{ij}x_j}$. Define the global Dirichlet form $\mathscr{E}_T$ on T by

$$\mathbb{D}(\mathscr{E}_T) := \left\{ g \in L^2(\mu_T):\ g_i \in \mathbb{D}(\mathscr{E}_i), \text{ for } \mu_T - \text{a.e. } \hat{x}_i \text{ and } \int_{\mathbb{N}^T} \sum_{i\in T} \mathscr{E}_i(g_i, g_i) d\mu_T < +\infty \right\},$$
$$\mathscr{E}_T(g,g) := \int_{\mathbb{N}^T} \sum_{i\in T} \mathscr{E}_i(g_i, g_i) d\mu_T, \quad g \in \mathbb{D}(\mathscr{E}_T) \tag{30}$$

where $g_i(x_i) := g(x_i, \hat{x}_i)$ with $\hat{x}_i := x_{T\setminus\{i\}}$ fixed.
The following additivity property of the Fisher information will be needed.

Lemma 4.3. (*Guillin et al. [8, Lemma 2.12]*) *Let ν_T, μ_T be probability measures on $\mathbb{N}^T$ such that $I_T(\nu_T|\mu_T) < +\infty$, and let μ_i, ν_i be the conditional distributions of x_i knowing $\hat{x}_i$ under μ, ν respectively. Then*

$$I_T(\nu_T|\mu_T) = \mathbb{E}^{\nu_T} \sum_{i\in T} I_i(\nu_i|\mu_i) \tag{31}$$

where $I_i(\nu_i|\mu_i)$ is the Fisher-Donsker-Varadhan information related to the Dirichlet form $(\mathscr{E}_i, \mathbb{D}(\mathscr{E}_i))$.

Proof. For the completeness we reproduce the proof. Let $f = d\nu_T/d\mu_T$. Then $d\nu_i/d\mu_i = f/\mu_i(f) = f_i/\mu_i(f_i)$, ν_T-a.s. where $f_i(x_i) = f(x_i, \hat{x}_i)$. For μ_T- a.e. $\hat{x}_i$ fixed,

$$I_i(\nu_i|\mu_i) = \mathscr{E}_i\left(\sqrt{\frac{f_i}{\mu_i(f_i)}}, \sqrt{\frac{f_i}{\mu_i(f_i)}}\right) = \frac{1}{\mu_i(f_i)} \mathscr{E}_i(\sqrt{f_i}, \sqrt{f_i})$$

(for $\mu_i(f_i)$ is constant with $\hat{x}_i$ fixed). We obtain

$$\mathbb{E}^{\nu_T} \sum_{i\in T} I_i(\nu_i|\mu_i) = \mathbb{E}^{\mu_T} f \sum_{i\in T} \frac{1}{\mu_i(f_i)} \mathscr{E}_i(\sqrt{f_i}, \sqrt{f_i}) = \mathbb{E}^{\mu_T} \sum_{i\in T} \mathscr{E}_i(\sqrt{f_i}, \sqrt{f_i}),$$

which completes the proof.
We can now state the main result of this section.

Theorem 4.4. *Let μ_T be the Gibbs measure given in (24). Assume the Dobrushin's uniqueness condition*

$$D = \sup_{j\in T}\sum_{i\in T}\lambda_i(1-e^{-\gamma_{ij}}) < 1. \tag{32}$$

Then for any $\nu_T \in \mathscr{M}_1^1(\mathbb{N}^T, d_{l^1})$, it holds that

$$W_{1,d_{l^1}}(\nu_T,\mu_T) \le \frac{1}{1-D}\left(2\sqrt{\left(\sum_{i\in T}\lambda_i\right)I} + I\right) \tag{33}$$

where $I = I_T(\nu_T|\mu_T)$.

Proof of Theorem Gibbs By Theorem 2.1, we know that for each $\mu_i = \mu_i(\cdot|x)$, it holds that

$$W_{1,\rho}(\nu_i,\mu_i) \le 2\sqrt{\lambda_i I_i(\nu_i|\mu_i)} + I_i(\nu_i|\mu_i),\ \forall\, \nu_i \in \mathscr{M}_1^1(\mathbb{N}). \tag{34}$$

Under the Dobrushin's uniqueness condition (32), by Lemma 4.2 and (34), we have by Cauchy-Schwarz inequality,

$$\begin{aligned}
(1-D)W_{1,d_{l^1}}(\nu_T,\mu_T) &\le \mathbb{E}^{\nu_T}\sum_{i\in T}W_{1,\rho}(\nu_i,\mu_i)\\
&\le 2\mathbb{E}^{\nu_T}\sum_{i\in T}\sqrt{\lambda_i I_i(\nu_i|\mu_i)} + \mathbb{E}^{\nu_T}\sum_{i\in T}I_i(\nu_i|\mu_i)\\
&\le 2\sqrt{\sum_{i\in T}\lambda_i\cdot\mathbb{E}^{\nu_T}\sum_{i\in T}I_i(\nu_i|\mu_i)} + \mathbb{E}^{\nu_T}\sum_{i\in T}I_i(\nu_i|\mu_i)
\end{aligned}$$

where the desired inequality follows by Lemma 4.3.

Remark 4.5. The inequality (33) in the free case is again sharp. Indeed if $\gamma_{ij} = 0$ (no interaction case) it is optimal, as seen by applying Theorem 2.1 to the function $\sum_{i\in T} x_i$.

References

[1] L. Bertini, N. Cancrini and F. Cesi. The spectral gap for a Glauber-type dynamic in a continuous gas. *Ann. Inst. Henri Poincaré Probab. Stat.,* 38(1): 91-108, 2002. MR1899231

[2] D. Chafaï. Binomial-Poisson inequalities and $M/M/\infty$ queue, *ESAIMS Probab. Stat.,* 10: 317-339, 2006. MR2247924

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

[4] R. L. Dobrushin. Prescribing a system of random variables by conditional distributions. *Theory Probab. Appl.,* 15: 458-486, 1970.

[5] F. Q. Gao, A. Guillin and L. M. Wu. Berstein type's concentration inequalities for symmetric Markov processes. To appear in *SIAM: Theory of Probability and its Applications.* 2011.

[6] F. Q. Gao and L. M. Wu. Transportation information inequalities for Gibbs measures. Preprint, 2007.

[7] A. Guillin., A. Joulin, C. Leonard and L. M. Wu. Transportation-information inequalities for Markov processes (Ⅲ): processes with jumps. Preprint, 2008.

[8] A. Guillin, C. Léonard, L. M. Wu and N. Yao. Transportation-information inequalities for Markov processes. *Probab. Theory Relat. Fields*, 144: 669-695, 2009. MR2496446

[9] A. Guillin, C. Léonard, F.-Y. Wang and L. M. Wu. Transportation-information inequalities for Markov processes (Ⅱ): relations with other functional inequalities. Available online at http://arxiv.org/abs/0902.2101.

[10] A. Joulin. A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature. *Bernoulli,* 15(2): 532–549, 2009. MR2543873

[11] A. Joulin and Y. Ollivier. Curvature, concentration, and error estimates for Markov chain Monte Carlo. *Ann. Probab.* 38(6): 2418-2442, 2010. MR2683634

[12] Yu. Kondratiev and E. Lytvynov. Glauber dynamics of continuous particle systems. *Ann. Inst. Henri Poincaré Probab. Stat.,* 41: 685-702, 2005. MR2144229

[13] W. Liu and Y.T. Ma. Spectral gap and convex concentration inequalities for birth-death processes. *Ann. Inst. Henri Poincaré Probab. Stat.,* 45(1): 58–69 (2009). MR2500228

[14] Y. T. Ma, S. Shen, X. Y. Wang and L. M. Wu. Transportation inequalities: from Poisson to Gibbs measures. *Bernoulli*. 17(1): 155-169, 2011. MR2797986

[15] Y. Ollivier. Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.,* 256(3): 810–864, 2009. MR2484937

[16] J. Picard. Formules de dualité sur l'espace de Poisson. *Ann. Inst. Henri Poincaré Probab. Stat.*, 32(4): 509–548, 1996. MR1411270

[17] C. Preston. Spatial birth-and-death processes, in: Proceedings of the 40th Session of the International Statistical Institute (Warsaw), vol. 2, in *Bull. Inst. Internat. Statist*., 46: 371-391, 1975. MR0474532

[18] D. Ruelle, *Statistical Mechanics: Rigorous Results,* Benjamin, 1969 MR0289084

[19] L. M. Wu. A new modified logarithmic Sobolev inequality for Poisson processes and several applications. *Probab. Theory Relat. Fields,* 118: 427-438, 2000. MR1800540

[20] L. M. Wu. Estimate of the spectral gap for continuous gas. *Ann. Inst. Henri Poincaré Probab. Stat.,* 40: 387-409, 2004. MR2070332