Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2009, Vol. 45, No. 3, 734–753
DOI: 10.1214/08-AIHP183

ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES

www.imstat.org/aihp

Convex entropy decay via the Bochner–Bakry–Emery approach

Pietro Caputo[a], Paolo Dai Pra[b] and Gustavo Posta[c]

[a]*Dipartimento di Matematica, Università di Roma Tre, L. go S. Murialdo 1, 00146 Roma, Italy. E-mail: caputo@mat.uniroma3.it*
[b]*Dipartimento di Matematica Pura e Applicata, Università di Padova, Via Belzoni 7, 35131 Padova, Italy. E-mail: daipra@math.unipd.it*
[c]*Dipartimento di Matematica, Politecnico di Milano, P. za Leonardo da Vinci 32, I-20133 Milano, Italy. E-mail: gustavo.posta@polimi.it*

Received 12 December 2007; accepted 9 June 2008

Abstract. We develop a method, based on a Bochner-type identity, to obtain estimates on the exponential rate of decay of the relative entropy from equilibrium of Markov processes in discrete settings. When this method applies the relative entropy decays in a convex way. The method is shown to be rather powerful when applied to a class of birth and death processes. We then consider other examples, including inhomogeneous zero-range processes and Bernoulli–Laplace models. For these two models, known results were limited to the homogeneous case, and obtained via the martingale approach, whose applicability to inhomogeneous models is still unclear.

Résumé. Nous développons une méthode, insipirée par une identité de Bochner, pour obtenir des estimées sur la decroissance exponentielle de l'entropie relative de processus de Markov avec sauts. Lorsque nous pouvons appliquer cette méthode, l'entropie relative est une fonction convexe du temps. On montre que la méthode s'applique de facon efficace à une large classe de processus de naissance et mort. On considère aussi d'autres exemples, comme les processus de zero-range et de Bernoulli–Laplace dans des cas non-homogènes. Pour ces derniers modèles les résultats connus, obtenus par la méthode de martingale, étaient limités au cas homogène.

MSC: 39B62; 60J80; 60K35

Keywords: Entropy decay; Modified logarithmic Sobolev inequality; Stochastic particle systems

1. Introduction

In the family of functional inequalities that are related to the convergence to equilibrium of Markov processes (Poincaré, logarithmic Sobolev and Nash inequalities) the *modified logarithmic Sobolev inequality* (MLSI) has been the last to attract interest among mathematicians, and certainly the less studied. Consider a time-homogeneous Markov process $(X_t)_{t\geq 0}$, with values on a measurable space $(S, \mathcal{S})$, having an invariant measure π. We assume the semigroup $(T_t)_{t\geq 0}$ defined on $L^2(\pi)$ by

$$T_t f(x) := E\big[f(X_t)|X_0 = x\big]$$

is strongly right-continuous, so that the infinitesimal generator $\mathcal{L}$ exists, i.e. $T_t = \mathrm{e}^{t\mathcal{L}}$. We also define the non-negative quadratic form on $\mathcal{D}(\mathcal{L}) \times \mathcal{D}(\mathcal{L})$, called *Dirichlet form* of $\mathcal{L}$,

$$\mathcal{E}(f, g) := -\pi[f\mathcal{L}g],$$

where $\mathcal{D}(\mathcal{L})$ is the domain of $\mathcal{L}$, and we use the notation $\pi[f]$ for $\int f\,\mathrm{d}\pi$. Given a probability measure μ on $(S,\mathcal{S})$, we denote by μT_t the distribution of X_t assuming X_0 is distributed according to μ, i.e.

$$\int f\,\mathrm{d}(\mu T_t) := \int (T_t f)\,\mathrm{d}\mu.$$

An ergodic Markov process, in particular a countable-state, irreducible and recurrent one, has a unique invariant measure π, and the rate of convergence of μT_t to π is a major topic of research. Quantitative estimates on this rate of convergence can be obtained by analyzing functional inequalities. To set up the necessary notations, define the *relative entropy* $h(\mu|\pi)$ of the probability μ with respect to π by

$$h(\mu|\pi) := \pi\left[\frac{\mathrm{d}\mu}{\mathrm{d}\pi}\log\frac{\mathrm{d}\mu}{\mathrm{d}\pi}\right],$$

where $h(\mu|\pi)$ is meant to be infinite whenever $\mu \not\ll \pi$ or $\frac{\mathrm{d}\mu}{\mathrm{d}\pi}\log\frac{\mathrm{d}\mu}{\mathrm{d}\pi} \notin L^1(\pi)$. Although $h(\cdot|\cdot)$ is not a metric in the usual sense, its use as “pseudo-distance” is motivated by a number of relevant properties, the most basic ones being:

$$h(\mu|\pi) = 0 \quad \Longleftrightarrow \quad \mu = \pi$$

and

$$\|\mu - \pi\|_{\mathrm{TV}}^2 \le h(\mu|\pi), \tag{1.1}$$

where $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation norm. For a generic measurable function $f \ge 0$ it is common to write

$$\mathrm{Ent}_\pi(f) := \begin{cases} \pi[f\log f] - \pi[f]\log\pi[f] & \text{if } f\log f \in L^1(\pi), \\ +\infty & \text{otherwise,} \end{cases}$$

so that $h(\mu|\pi) = \mathrm{Ent}_\pi(\frac{\mathrm{d}\mu}{\mathrm{d}\pi})$. Ignoring technical problems concerning the domains of Dirichlet forms, a simple formal computation shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}h(\mu T_t|\pi) = -\mathcal{E}\big(T_t^* f, \log T_t^* f\big), \tag{1.2}$$

where $f := \frac{\mathrm{d}\mu}{\mathrm{d}\pi}$, $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$ in $L^2(\pi)$, and $T_t^* := \mathrm{e}^{t\mathcal{L}^*}$. Therefore, assuming that, for each $f \ge 0$

$$\mathrm{Ent}_\pi(f) \le \frac{1}{\alpha}\mathcal{E}(f, \log f) \tag{1.3}$$

with $\alpha > 0$ (independent of f), then (1.2) can be closed to get a differential inequality, obtaining

$$h(\mu T_t|\pi) \le \mathrm{e}^{-\alpha t} h(\mu|\pi).$$

In other words, estimates on the best constant α for which the functional inequality (1.3) holds provide estimates for the rate of exponential convergence to equilibrium of the process, in the relative entropy sense.

We shall be interested in *reversible* dynamics, i.e. when $\mathcal{L} = \mathcal{L}^*$. When (1.3) holds we say that the pair $(\mathcal{L}, \pi)$ satisfies the *modified logarithmic Sobolev inequality* (MLSI) with constant α. This inequality turns out to be intermediate, in a sense that we will make precise in a moment, between two more “traditional” functional inequalities, namely the *logarithmic Sobolev inequality* (LSI)

$$\mathrm{Ent}_\pi\big(f^2\big) \le \frac{1}{\beta}\mathcal{E}(f, f), \tag{1.4}$$

and the *Poincaré inequality* (PI)

$$\mathrm{Var}_\pi(f) \le \frac{1}{\gamma}\mathcal{E}(f, f), \tag{1.5}$$

where $\mathrm{Var}_\pi(f) := \pi[(f - \pi[f])^2]$. It is well known that (LSI) is equivalent to hypercontractivity of the semigroup T_t, i.e. T_t is contractive as linear operator from $L^2(\pi)$ and $L^p(\pi)$ for some $p > 2$, and (PI) is equivalent to exponential convergence to equilibrium in L^2, i.e. $\|T_t f - \pi[f]\|_{L^2(\pi)} \le \mathrm{e}^{-\gamma t}\|f - \pi[f]\|_{L^2(\pi)}$. Moreover, if we let α, β, γ denote the best constant in the respective inequality (with the convention that the "best" constant is zero when the inequality fails for every positive constant), then for reversible systems

$$2\gamma \ge \alpha \ge 4\beta. \tag{1.6}$$

We refer to [12] and [3] for tutorial references on these inequalities (even though (MLSI) is never explicitly mentioned in [12]). It should also be noticed that in the case $\mathcal{L}$ is the generator of a reversible diffusion process, e.g. $\mathcal{L} = \frac{1}{2}\Delta + \nabla V \cdot \nabla$, (LSI) and (MLSI) coincide. This equivalence, that simply follows from the fact that $\nabla \log f = \nabla f / f$, does not extend to Markov processes with jumps; even for a two-state Markov chain, the best constants in (LSI) and (MLSI) behave quite differently in terms of the parameter of the invariant measure (see [2]). The case of processes with jumps leaves some freedom in deciding which inequality is the best analogue of the (LSI) and there are several inequalities which are commonly referred to as "modified logarithmic Sobolev" in the literature (all of them coincide with the (LSI) in the diffusion case). Here we only consider the (MLSI) defined in (1.3). Besides the exponential decay of entropy it is known that this estimate implies useful concentration bounds, see [3], and thanks to (1.1) it is a natural tool to estimate mixing times, see [12]. Furthermore, we mention that there is a further family of inequalities interpolating between the exponential decay in the L^2-sense of (PI) and the exponential decay in the $L\log L$-sense of (MLSI) that has received growing attention in the literature. These so-called Beckner inequalities deal with the exponential decay in the L^p-sense, for $p \in (1, 2)$, see [3] and references therein for more details.

While the study of (PI) and (LSI) for large-scale systems dates back to [19,21] and [23], a similar analysis for (MLSI) has been first proposed in [10], which deals with Glauber type dynamics with unbounded particle number per site; for a class of such systems (LSI) fails, while (MLSI) holds with a strictly positive constant. Dynamics with exchange of particles are less understood, with some relevant exceptions (see [3,13,14]).

The purpose of this paper is to partially extend to (MLSI) the general tools developed in [5] for obtaining estimates on the *spectral gap*, i.e. the best constant in (PI). The results in [5] apply to interacting particle systems ideas of S. Bochner (see [4]), who studied the spectral gap of the Laplacian in Riemannian manifolds. D. Bakry and M. Emery [1] have developed Bochner ideas, obtaining conditions for (LSI) to hold for diffusion processes. Our work can be interpreted as an attempt to complete the Bakry and Emery's program for processes with jumps. Although the basic tools are developed in full generality, useful estimates on the best constant in (MLSI) have been obtained, unfortunately, in a limited number of examples, if compared both with the diffusion case and the case of (PI) studied in [5]. We believe, however, that these examples are of interest, and that our approach is promising and not yet fully exploited.

2. The Bochner–Bakry–Emery approach to (MLSI)

We recall here a simple argument relating (MLSI) to exponential decay of entropy and of its time derivative along the semigroup. To avoid problems that are not relevant for the applications we have in mind, we assume the state space S of the Markov chain to be finite or countable. By simple calculus one checks that, for $f > 0$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Ent}_\pi(T_t f) = -\mathcal{E}(T_t f, \log T_t f). \tag{2.1}$$

Therefore (MLSI) is equivalent to exponential decay of relative entropy in the sense that for every $\alpha \ge 0$ one has:

$$\alpha\,\mathrm{Ent}_\pi(f) \le \mathcal{E}(f, \log f) \quad \forall f > 0 \quad \Longleftrightarrow \quad \mathrm{Ent}_\pi(T_t f) \le \mathrm{e}^{-\alpha t}\,\mathrm{Ent}_\pi(f) \quad \forall f > 0, \forall t \ge 0. \tag{2.2}$$

Indeed, the implication $\Rightarrow$ is obtained by integrating (2.1) and the implication $\Leftarrow$ follows by subtracting $\mathrm{Ent}_\pi(f)$ from the right-hand side of (2.2), dividing by t, and taking $t \to 0$.

In other words, (MLSI) is equivalent to a control on the first time derivative of entropy. The following simple lemma (which is well known, see e.g. [18]) is based on a similar control of second derivatives.

Lemma 2.1. *Suppose the generator $\mathcal{L}$ is self-adjoint in $L^2(\pi)$ and the resulting Markov chain is irreducible. Then, for every $\kappa \geq 0$ we have the equivalence*:

$$\begin{aligned} &\kappa \mathcal{E}(f, \log f) \leq \pi[\mathcal{L} f \mathcal{L} \log f] + \pi\left[\frac{(\mathcal{L} f)^2}{f}\right] \quad \forall f > 0 \\ \iff \quad &\mathcal{E}(T_t f, \log T_t f) \leq \mathrm{e}^{-\kappa t} \mathcal{E}(f, \log f) \quad \forall f > 0, \forall t \geq 0. \end{aligned} \tag{2.3}$$

Moreover, if (2.3) *holds for some κ, then* (MLSI) *holds with $\alpha = \kappa$.*

Proof. Computing second derivatives we obtain,

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} \operatorname{Ent}_\pi(T_t f) = -\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(T_t f, \log T_t f) = \pi[\mathcal{L} T_t f \mathcal{L} \log T_t f] + \pi\left[\frac{(\mathcal{L} T_t f)^2}{T_t f}\right]. \tag{2.4}$$

The equivalence (2.3) therefore follows as in the case of (2.2) discussed above. To prove the last assertion, note that the inequality

$$\kappa \operatorname{Ent}_\pi(f) \leq \mathcal{E}(f, \log f),$$

follows by integrating from 0 to ∞ the inequality

$$-\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}(T_t f, \log T_t f) \geq -\kappa \frac{\mathrm{d}}{\mathrm{d}t} \operatorname{Ent}_\pi(T_t f).$$

□

Remark 2.2. *Inequality* (2.3) *is in general strictly stronger than* (MLSI): *it implies uniform exponential decay of entropy, but also that the decay is convex in time. While it is easily seen that $\mathcal{E}(f, \log f) \geq 0$ for all $f > 0$, nothing forces the second derivative of* $\operatorname{Ent}_\pi(T_t f)$ *to be non-negative. There are examples showing that* (MLSI) *may hold without convexity in time of entropy. An example in the continuous setting, due to B. Helffer, can be found in* [18]. *An example in the discrete setting will be given later in this paper, see Section* 4.

In order to investigate the validity of (2.3) in the discrete setting, we write the generators of our Markov chains in the form

$$\mathcal{L} f(\eta) = \sum_{\gamma \in G} c(\eta, \gamma)\big[f\big(\gamma(\eta)\big) - f(\eta)\big] =: \sum_{\gamma \in G} c(\eta, \gamma) \nabla_\gamma f(\eta), \tag{2.5}$$

where G is some (finite or countable) set of functions from S to S (the *allowed moves*) and $c: S \times G \to [0, \infty)$ represents the jump rates. It is easily seen that the generator of every finite or countable Markov chains can be written in this form; a form that, as we will see, becomes rather natural in many specific examples.

With these notations, reversibility is expressed as follows.

(Rev) For every $\gamma \in G$ there exists $\gamma^{-1} \in G$ such that $\gamma^{-1}(\gamma(\eta)) = \eta$ for every $\eta \in S$ such that $c(\eta, \gamma) > 0$. Moreover for every $f: S \to \mathbb{R}$ bounded

$$\pi\big[c(\eta, \gamma) f(\eta)\big] = \pi\big[c\big(\eta, \gamma^{-1}\big) f\big(\gamma^{-1}(\eta)\big)\big].$$

Under (Rev) it is easy to see that

$$\mathcal{E}(f, g) = \frac{1}{2} \pi\left[\sum_{\gamma \in G} c(\eta, \gamma) \nabla_\gamma f(\eta) \nabla_\gamma g(\eta)\right]. \tag{2.6}$$

In particular the Dirichlet form is symmetric, so $\mathcal{L}$ is self-adjoint in $L^2(\pi)$.

One of the key point of the Bochner–Bakry–Emery approach is the so called *Bochner identity*; a version of this identity in the discrete setting is given in the next lemma.

Lemma 2.3. *Let* $R : S \times G \times G \to [0, +\infty)$ *be such that*

$$\text{(P1):} \quad R(\eta, \gamma, \delta) = R(\eta, \delta, \gamma) \quad \forall \eta, \gamma, \delta \text{ with } R(\eta, \gamma, \delta) > 0,$$

$$\text{(P2):} \quad \pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\psi(\eta, \gamma, \delta)\Bigg] = \pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\psi\big(\gamma(\eta), \gamma^{-1}, \delta\big)\Bigg] \quad \forall \psi : S \times G \times G \to \mathbb{R} \text{ bounded},$$

$$\text{(P3):} \quad \gamma\big(\delta(\eta)\big) = \delta\big(\gamma(\eta)\big) \quad \forall \eta, \gamma, \delta \text{ with } R(\eta, \gamma, \delta) > 0.$$

Then, for every bounded f, g *the following* Bochner-*type identity holds*

$$\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\gamma f(\eta)\nabla_\delta g(\eta)\Bigg] = \frac{1}{4}\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\gamma\nabla_\delta f(\eta)\nabla_\gamma\nabla_\delta g(\eta)\Bigg].$$

Proof. First, by (P3), $\nabla_\gamma\nabla_\delta f(\eta)\nabla_\gamma\nabla_\delta g(\eta) = \nabla_\gamma\nabla_\delta f(\eta)\nabla_\delta\nabla_\gamma g(\eta)$ whenever $R(\eta, \gamma, \delta) > 0$. Then write

$$\nabla_\gamma\nabla_\delta f(\eta)\nabla_\delta\nabla_\gamma g(\eta) = \nabla_\delta f\big(\gamma(\eta)\big)\nabla_\gamma g\big(\delta(\eta)\big) - \nabla_\delta f\big(\gamma(\eta)\big)\nabla_\gamma g(\eta) - \nabla_\delta f(\eta)\nabla_\gamma g\big(\delta(\eta)\big) + \nabla_\delta f(\eta)\nabla_\gamma g(\eta).$$

We show that each one of the four summands in the r.h.s. of this last formula, when multiplied by $R(\eta, \gamma, \delta)$, summed over γ, δ and averaged over π gives

$$\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\delta f(\eta)\nabla_\gamma g(\eta)\Bigg].$$

For the fourth summand there is nothing to prove. Moreover, by (P2),

$$\begin{aligned}\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\delta f(\eta)\nabla_\gamma g(\eta)\Bigg] &= \pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\delta f\big(\gamma(\eta)\big)\nabla_{\gamma^{-1}} g\big(\gamma(\eta)\big)\Bigg] \\ &= -\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\delta f\big(\gamma(\eta)\big)\nabla_\gamma g(\eta)\Bigg]\end{aligned}$$

which takes care of the second and, by symmetry, of the third summand. For the first summand we use first (P2), then (P1), (P2) again and (P3):

$$\begin{aligned}\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\delta f(\eta)\nabla_\gamma g(\eta)\Bigg] &= \pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\delta f\big(\gamma(\eta)\big)\nabla_{\gamma^{-1}} g\big(\gamma(\eta)\big)\Bigg] \\ &= -\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\gamma f\big(\delta(\eta)\big)\nabla_\delta g(\eta)\Bigg] \\ &= -\pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_{\gamma^{-1}} f\big(\delta\gamma(\eta)\big)\nabla_\delta g\big(\gamma(\eta)\big)\Bigg] \\ &= \pi\Bigg[\sum_{\gamma,\delta} R(\eta, \gamma, \delta)\nabla_\gamma f\big(\delta(\eta)\big)\nabla_\delta g\big(\gamma(\eta)\big)\Bigg].\end{aligned}$$

□

Corollary 2.4. *Let* $R : S \times G \times G \to [0, +\infty)$ *be such that* (P1), (P2) *and* (P3) *from Lemma* 2.3 *hold. Define*

$$\Gamma(\eta, \gamma, \delta) := c(\eta, \gamma)c(\eta, \delta) - R(\eta, \gamma, \delta).$$

Then, for every bounded $f > 0$:

$$\pi[\mathcal{L}f\mathcal{L}\log f] + \pi\Bigg[\frac{(\mathcal{L}f)^2}{f}\Bigg] \geq \pi\Bigg[\sum_{\gamma,\delta} \Gamma(\eta, \gamma, \delta)\bigg(\nabla_\gamma f(\eta)\nabla_\delta \log f(\eta) + \frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\bigg)\Bigg].$$

Proof. First observe that

$$\pi[\mathcal{L}f\mathcal{L}\log f]+\pi\left[\frac{(\mathcal{L}f)^2}{f}\right]=\pi\left[\sum_{\gamma,\delta}c(\eta,\gamma)c(\eta,\delta)\nabla_\gamma f(\eta)\nabla_\delta\log f(\eta)\right]$$
$$+\pi\left[\sum_{\gamma,\delta}c(\eta,\gamma)c(\eta,\delta)\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right]. \tag{2.7}$$

Now, if we apply Bochner's identity to the first summand of the right-hand side of (2.7) we obtain

$$\pi[\mathcal{L}f\mathcal{L}\log f]+\pi\left[\frac{(\mathcal{L}f)^2}{f}\right]=\pi\left[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\nabla_\gamma f(\eta)\nabla_\delta\log f(\eta)\right]$$
$$+\frac{1}{4}\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\nabla_\gamma\nabla_\delta f(\eta)\nabla_\gamma\nabla_\delta\log f(\eta)\right]$$
$$+\pi\left[\sum_{\gamma,\delta}c(\eta,\gamma)c(\eta,\delta)\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right]. \tag{2.8}$$

Thus the conclusion follows immediately if we show that

$$\frac{1}{4}\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\nabla_\gamma\nabla_\delta f(\eta)\nabla_\gamma\nabla_\delta\log f(\eta)\right]+\pi\left[\sum_{\gamma,\delta}c(\eta,\gamma)c(\eta,\delta)\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right]$$
$$\geq\pi\left[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right], \tag{2.9}$$

or, equivalently

$$\frac{1}{4}\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\nabla_\gamma\nabla_\delta f(\eta)\nabla_\gamma\nabla_\delta\log f(\eta)\right]+\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right]\geq 0.$$

Now we apply to the second summand in (2.10) the same argument used in the proof of Lemma 2.3. The result is more cumbersome, due to the presence of the denominator $f(\eta)$:

$$\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right]$$
$$=\frac{1}{4}\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\left\{\nabla_\gamma\left(\frac{\nabla_\delta f(\eta)}{f(\delta(\eta))}\right)\nabla_\gamma\nabla_\delta f(\eta)-\nabla_\gamma\left(\frac{(\nabla_\delta f(\eta))^2}{f(\eta)f(\delta(\eta))}\right)\nabla_\gamma f(\eta)\right\}\right]. \tag{2.10}$$

Substituting in (2.10) we get

$$\frac{1}{4}\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\nabla_\gamma\nabla_\delta f(\eta)\nabla_\gamma\nabla_\delta\log f(\eta)\right]+\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right]$$
$$=\frac{1}{4}\pi\left[\sum_{\gamma,\delta}R(\eta,\gamma,\delta)\left\{\nabla_\gamma\nabla_\delta f(\eta)\nabla_\gamma\nabla_\delta\log f(\eta)\right.\right.$$
$$\left.\left.+\nabla_\gamma\left(\frac{\nabla_\delta f(\eta)}{f(\delta(\eta))}\right)\nabla_\gamma\nabla_\delta f(\eta)-\nabla_\gamma\left(\frac{(\nabla_\delta f(\eta))^2}{f(\eta)f(\delta(\eta))}\right)\nabla_\gamma f(\eta)\right\}\right]. \tag{2.11}$$

Setting $a := f(\eta), b := f(\delta(\eta)), c := f(\gamma(\eta)), d := f(\delta\gamma(\eta))$, one checks that the term in braces in the right-hand side of (2.11) equals the sum of the following 4 expressions:

$$\begin{aligned}
&d\log d - d\log(bc/a) + (bc/a) - d,\\
&c\log c - c\log(da/b) + (da/b) - c,\\
&b\log b - b\log(da/c) + (da/c) - b,\\
&a\log a - a\log(bc/d) + (bc/d) - a,
\end{aligned}$$

which are all non-negative, since $\alpha\log\alpha - \alpha\log\beta + \beta - \alpha \geq 0$ for every $\alpha, \beta > 0$. This shows that (2.11) is non-negative, which completes the proof. □

Summing up, from Lemma 2.1 and Corollary 2.4 we are led to the following result.

Proposition 2.5. *Consider the Markov chain defined by* (2.5). *Suppose there exists a constant* $\kappa > 0$ *such that for every* $f > 0$

$$\begin{aligned}
&\pi\Biggl[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\biggl(\nabla_\gamma f(\eta)\nabla_\delta\log f(\eta) + \frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\biggr)\Biggr]\\
&\quad\geq \frac{\kappa}{2}\pi\Biggl[\sum_{\gamma\in G}c(\eta,\gamma)\nabla_\gamma f(\eta)\nabla_\gamma\log f(\eta)\Biggr] = \frac{\kappa}{2}\mathcal{E}(f,\log f).
\end{aligned}$$

Then (2.3) *holds. In particular,* (MLSI) *holds with* $\alpha = \kappa$.

We do not have a general choice for the function R of Lemma 2.3 to use for applying the criterion in Proposition 2.5. One option is given in [5], Proposition 2.4. That choice, however, does not work for the examples in Section 4 below. We only mention that in all examples we obtain $R(\eta,\gamma,\delta)$ by "modifying" $c(\eta,\gamma)c(\gamma(\eta),\delta)$ in order to fulfill properties (P1)–(P3).

3. A warming-up example: birth and death processes

Consider a birth and death process on $\mathbb{N}$ with generator

$$\mathcal{L}f(n) = a(n)\nabla_+ f(n) + b(n)\nabla_- f(n). \tag{3.1}$$

In the language introduced in the previous section, $G = \{+,-\}$, where $+(n) = n+1$ and $-(n) = (n-1)\mathbf{1}_{n>0}$, and one is the inverse of the other. In particular, $\nabla_\pm f(n) = f(n\pm 1) - f(n)$. The rates a, b are non-negative functions on $\mathbb{N}$ such that $b(0) = 0$ and we assume that there exists a probability π on $\mathbb{N}$ such that the detailed balance equation

$$a(n)\pi(n) = b(n+1)\pi(n+1) \tag{3.2}$$

holds true. Moreover, we assume that the resulting Markov chain is irreducible.

Setting $c(n,+) = a(n)$, $c(n,-) = b(n)$ we see that condition (Rev) in the previous section is satisfied and

$$\mathcal{E}(f,g) = \pi\bigl[a(n)\nabla_+ f(n)\nabla_+ g(n)\bigr] = \pi\bigl[b(n)\nabla_- f(n)\nabla_- g(n)\bigr].$$

We define R as follows:

$$\begin{aligned}
&R(n,+,+) := a(n)a(n+1),\\
&R(n,-,-) := b(n)b(n-1),\\
&R(n,+,-) = R(n,-,+) := a(n)b(n).
\end{aligned} \tag{3.3}$$

It is a simple exercise to show that conditions (P1)–(P3) of Lemma 2.3 are satisfied. In particular, (P2) follows by application of reversibility. Then, letting as before $\Gamma(n,\delta,\gamma) = c(n,\gamma)c(n,\delta) - R(n,\gamma,\delta)$, Corollary 2.4 yields

$$\begin{aligned}
&\pi[\mathcal{L}f\mathcal{L}\log f] + \pi\left[\frac{(\mathcal{L}f)^2}{f}\right] \\
&\quad\geq \pi\left[\sum_{\gamma,\delta\in G}\Gamma(n,\delta,\gamma)\left(\nabla_\gamma f(n)\nabla_\delta \log f(n) + \frac{\nabla_\gamma f(n)\nabla_\delta f(n)}{f(n)}\right)\right] \\
&\quad= \pi\big[a(n)\big[a(n)-a(n+1)\big]\nabla_+ f(n)\nabla_+\log f(n) + b(n)\big[b(n)-b(n-1)\big]\nabla_- f(n)\nabla_-\log f(n)\big] \\
&\qquad+ \pi\left[a(n)\big[a(n)-a(n+1)\big]\frac{(\nabla_+ f(n))^2}{f(n)} + b(n)\big[b(n)-b(n-1)\big]\frac{(\nabla_- f(n))^2}{f(n)}\right].
\end{aligned} \tag{3.4}$$

We consider the following assumption:

(A) $a(n) \geq a(n+1)$ and $b(n+1) \geq b(n)$, and there exists $c > 0$ such that for every $n \geq 0$,

$$a(n) - a(n+1) + b(n+1) - b(n) \geq c. \tag{3.5}$$

Assuming (A), (3.4) can be further estimated as follows. Thanks to monotonicity of the rates we can drop the terms in the last line of (3.4). Moreover from the reversibility (3.2) we see that

$$\pi\big[b(n)\big[b(n)-b(n-1)\big]\nabla_- f(n)\nabla_-\log f(n)\big] = \pi\big[a(n)\big[b(n+1)-b(n)\big]\nabla_+ f(n)\nabla_+\log f(n)\big].$$

Therefore, using (3.5) we arrive at

$$\pi[\mathcal{L}f\mathcal{L}\log f] + \pi\left[\frac{(\mathcal{L}f)^2}{f}\right] \geq c\pi\big[a(n)\nabla_+ f(n)\nabla_+\log f(n)\big] = c\mathcal{E}(f,\log f).$$

Recalling Proposition 2.5 we have therefore proved the following result.

Theorem 3.1. *Under assumption* (A), *both* (2.3) *and* (MLSI) *hold with constant c.*

There are well known criteria for the validity of (PI) or (LSI) for one-dimensional processes as the ones considered above, see [20] for explicit estimates on the constants involved. On the other hand, we are not aware of any such result concerning (MLSI). As far as we know Theorem 3.1 is the first general sufficient condition for the validity of (MLSI). Moreover, despite its simplicity, this result is sharper than it may appear, as the following examples illustrate. It is worthy of note that condition (A) implies a contraction property for the Wasserstein metric, see the recent paper [16] where this observation is used to prove Poisson-like deviation inequalities.

3.1. *Poisson case*

The Poisson case refers to the choice $a(n) = \lambda$, $b(n) = n$ and $\pi_\lambda(n) = \frac{\lambda^n}{n!}e^{-\lambda}$, with $\lambda > 0$. It is well known that (LSI) fails in this case. To see this, take $f_k(n) := \mathbf{1}_{(k,+\infty)}(n)$ in (1.4) and then let $k \to \infty$. On the other hand assumption (A) is satisfied with $c = 1$ so that Theorem 3.1 yields the following estimate for any $f > 0$:

$$\mathrm{Ent}_{\pi_\lambda}(f) \leq \lambda\pi_\lambda[\nabla_+ f\nabla_+\log f]. \tag{3.6}$$

Note that this estimate is sharp, in the sense that no better constant than λ can satisfy (3.6) for all $f > 0$. To see this, it suffices to take $f_k(n) = e^{kn}$, for fixed $k \in \mathbb{N}$; simple computations show that $\mathrm{Ent}_{\pi_\lambda}(f) = (k\lambda e^k - \lambda e^k + \lambda)e^{\lambda(e^k-1)}$ while $\pi_\lambda[\nabla_+ f\nabla_+\log f] = k(e^k-1)e^{\lambda(e^k-1)}$, and equality is approached in (3.6) as $k \to \infty$. In the Poisson case one can obtain inequality (3.6) also using the Poisson limit of the binomial distribution as in [2,10]. In this special case the analysis can be pushed beyond these statements, see e.g. [9,22] for further developments.

3.2. *Log-concave probabilities*

A non-negative function γ on $\mathbb{N}$ is called log-concave if

$$\gamma(n)^2 \geq \gamma(n+1)\gamma(n-1). \tag{3.7}$$

Suppose our measure π is such that $\gamma(n) := n!\pi(n)$ satisfies (3.7). Such a measure is sometimes called ultra log-concave. If we set $a(n) = 1$ for all $n \geq 0$, then it follows that

$$b(n+1) - b(n) = \frac{\gamma(n)}{\gamma(n+1)}(n+1) - \frac{\gamma(n-1)}{\gamma(n)}n \geq \frac{\gamma(n-1)}{\gamma(n)} \geq \cdots \geq \frac{\gamma(0)}{\gamma(1)} = b(1).$$

From Theorem 3.1 we obtain that (MLSI) holds in this model with $\alpha = b(1)$, i.e. that π satisfies

$$b(1)\,\mathrm{Ent}_\pi(f) \leq \pi[\nabla_+ f \nabla_+ \log f], \tag{3.8}$$

for any $f > 0$. Note that (3.6) is a special case of (3.8) – indeed, it suffices to choose $a(n) = 1$ and $b(n) = n/\lambda$ instead of $a(n) = \lambda$ and $b(n) = n$. It has been shown in [15] that Poisson measures maximize entropy in the class of ultra log-concave measures. It is interesting to note that the convexity results obtained in [15] can be derived in a simple way from the arguments in our proof of Theorem 3.1.

3.3. *Random walks*

Another example is the simple random walk on a segment $[0, N] \cap \mathbb{Z}$ with reflecting boundary conditions. Here the (MLSI) constant is known to be of order $1/N^2$ by (1.6), since both (LSI) and (PI) can be shown to hold with $\beta \sim 1/N^2$ and $\gamma \sim 1/N^2$ (see e.g. [23] for the proof of the (LSI)). Let us show that this can be deduced from the above bounds. Let μ denote the uniform probability over $[0, N] \cap \mathbb{Z}$. We want to prove that for some constant C, for all $f > 0$:

$$\mathrm{Ent}_\mu(f) \leq CN^2\mu[\nabla_+ f \nabla_+ \log f]. \tag{3.9}$$

Let π denote the probability on $[0, N] \cap \mathbb{Z}$ such that $\pi(n)$ is proportional to $\mathbf{1}_{\{0 \leq n \leq N\}} \mathrm{e}^{-n^2/N^2}$. It is easy to check that π is equivalent to μ, i.e. $\delta \leq \mu(n)/\pi(n) \leq \delta^{-1}$ for some $\delta > 0$ independent of N and n. Then, by a standard comparison argument (see e.g. Lemma 3.3 in [12]), it is sufficient to prove (3.9) for the measure π in place of μ. This in turn follows from Theorem 3.1. Indeed, setting $a(n) = 1$, for all $0 \leq n \leq N - 1$ and $a(N) = 0$, we have $b(n) = \pi(n-1)/\pi(n)$ for all $1 \leq n \leq N$, and $b(0) = 0$. Therefore (3.5) applies with $c^{-1} = \mathrm{O}(N^2)$, for all $0 \leq n \leq N - 1$ and Theorem 3.1 allows to prove the claim.

3.4. *Non-monotone rates*

By means of a perturbation argument we can relax the monotonicity requirement in assumption (A). More precisely, suppose that $a(n) = 1$ for all $n \geq 0$, so that the probability measure π satisfies $\pi(n+1)/\pi(n) = 1/b(n+1)$, or

$$\pi(n) = \frac{\pi(0)}{b(1)\cdots b(n)}, \quad n \geq 1. \tag{3.10}$$

In this case Theorem 3.1 shows that if $b(n+1) - b(n) \geq c$ for all $n \geq 0$ then (MLSI) holds with $\alpha = c$. The next result, which is a key ingredient in the proof of the main theorem in [7], shows that if we impose a Lipschitz condition on the rates $b(n)$, then it is sufficient to have monotonicity on a large scale.

Proposition 3.2. *Suppose that there exist $C_1 < \infty$, $\delta > 0$ and $n_0 \in \mathbb{N}$ such that*

$$\sup_{n\geq 0}\bigl|b(n+1)-b(n)\bigr| \leq C_1 \quad \text{and} \quad \inf_{n\geq 0}\bigl[b(n+n_0)-b(n)\bigr] \geq \delta.$$

Then, for some constant C which may depend on C_1, δ, and n_0 only, the probability measure (3.10) *satisfies*

$$\operatorname{Ent}_\pi(f) \leq C\pi[\nabla_+ f \nabla_+ \log f], \quad f > 0. \tag{3.11}$$

To prove the proposition we shall need a preliminary lemma. Define

$$\widetilde{b}(k) := b(k) + \frac{1}{n_0}\sum_{j=1}^{n_0-1}\frac{n_0-j}{n_0}\bigl[b(k+j)+b(k-j)-2b(k)\bigr], \quad k \geq n_0, \tag{3.12}$$

and, when $0 \leq k < n_0$, set $\widetilde{b}(k) = \widetilde{b}(n_0)k/n_0$. Let us call $\widetilde{\pi}$ the probability measure obtained from $\widetilde{b}$ by (3.10).

Lemma 3.3. *The rate function $\widetilde{b}$ is uniformly increasing: there exists $\delta_1 > 0$ such that $\nabla_+\widetilde{b} \geq \delta_1$. Moreover, π and $\widetilde{\pi}$ are equivalent: there exists $C > 0$ such that $C^{-1} \leq \widetilde{\pi}(n)/\pi(n) \leq C$, for all $n \in \mathbb{N}$.*

Proof. We rewrite $\widetilde{b}(k)$, $k \geq n_0$:

$$\widetilde{b}(k) = \frac{b(k)}{n_0} + \frac{1}{n_0}\sum_{j=1}^{n_0-1}\frac{n_0-j}{n_0}\bigl[b(k+j)+b(k-j)\bigr] = \frac{1}{n_0}\sum_{j=0}^{n_0-1}\left\{\frac{n_0-j}{n_0}b(k+j) + \frac{j}{n_0}b(k+j-n_0)\right\}.$$

To compute $\nabla_+\widetilde{b}$ we use summation by parts in the form

$$\sum_{j=\ell}^{m}\psi(j)\nabla_+\varphi(j) = \psi(m)\varphi(m+1) - \psi(\ell)\varphi(\ell) - \sum_{j=\ell+1}^{m}\varphi(j)\nabla_+\psi(j-1), \tag{3.13}$$

where $\ell < m$ and ψ, φ are arbitrary functions. We apply (3.13) with $\ell = 0, m = n_0 - 1$, first to the case $\psi(j) = (n_0 - j)/n_0$, $\varphi(j) = b(k+j)$ and then to the case $\psi(j) = j/n_0$, $\varphi(j) = b(k+j-n_0)$. The conclusion is that, for every $k \geq n_0$ we have

$$\nabla_+\widetilde{b}(k) = \frac{1}{n_0^2}\sum_{j=1}^{n_0}\bigl[b(k+j) - b(k+j-n_0)\bigr]. \tag{3.14}$$

Since $\nabla_+\widetilde{b}(k) \geq \widetilde{b}(n_0)/n_0$ for every $k < n_0$, the claim $\nabla_+\widetilde{b} \geq \delta_1$ follows from (3.14) and the hypothesis $b(n+n_0) - b(n) \geq \delta$.

We turn to the proof of the equivalence of $\pi, \widetilde{\pi}$. We have to prove that there exists $C \in [1, \infty)$ such that for every $n \in \mathbb{N}$

$$C^{-1} \leq \prod_{k=1}^{n}\frac{\widetilde{b}(k)}{b(k)} \leq C.$$

We shall prove the left inequality above. The right inequality is obtained with the same proof by interchanging the role of b and $\widetilde{b}$. Passing to logarithms it suffices to prove

$$\sup_n \sum_{k=1}^{n}\frac{b(k)-\widetilde{b}(k)}{\widetilde{b}(k)} < \infty. \tag{3.15}$$

From (3.12), writing

$$b(k+j)+b(k-j)-2b(k)=\sum_{i=0}^{j-1}\big[\nabla_+b(k+i)-\nabla_+b(k+i-j)\big],$$

we have

$$\sum_{k=n_0}^{n}\frac{b(k)-\widetilde{b}(k)}{\widetilde{b}(k)}=\frac{1}{n_0}\sum_{j=1}^{n_0-1}\frac{n_0-j}{n_0}\sum_{i=0}^{j-1}\sum_{k=n_0}^{n}\frac{\nabla_+b(k+i-j)-\nabla_+b(k+i)}{\widetilde{b}(k)}. \tag{3.16}$$

Now, for every fixed $i<j$ we can use summation by parts as in (3.13), with $\ell=n_0, m=n$ and $\psi(k)=1/\widetilde{b}(k)$, $\varphi(k)=b(k+i-j)$ to obtain

$$\sum_{k=n_0}^{n}\frac{\nabla_+b(k+i-j)}{\widetilde{b}(k)}=\frac{b(n+1+i-j)}{\widetilde{b}(n)}-\frac{b(n_0+i-j)}{\widetilde{b}(n_0)}+\sum_{k=n_0+1}^{n}\frac{b(k+i-j)\nabla_+\widetilde{b}(k-1)}{\widetilde{b}(k)\widetilde{b}(k-1)}.$$

Another application of (3.13) with $\varphi(k)=b(k+i)$ yields therefore the identity

$$\begin{aligned}\sum_{k=n_0}^{n}\frac{\nabla_+b(k+i-j)-\nabla_+b(k+i)}{\widetilde{b}(k)}&=\frac{b(n+1+i-j)-b(n+1+i)}{\widetilde{b}(n)}-\frac{b(n_0+i-j)-b(n_0+i)}{\widetilde{b}(n_0)}\\&\quad+\sum_{k=n_0+1}^{n}\frac{[b(k+i-j)-b(k+i)]\nabla_+\widetilde{b}(k-1)}{\widetilde{b}(k)\widetilde{b}(k-1)}.\end{aligned} \tag{3.17}$$

Since $\widetilde{b}(k)\geq\delta_1 k$, the sequence $(\widetilde{b}(k)\widetilde{b}(k-1))^{-1}$ is summable. By hypothesis the increments of b (and $\widetilde{b}$) are uniformly bounded and therefore the sum in (3.17) is uniformly bounded in n, for every $i<j<n_0$. Now (3.15) follows from (3.16). □

The proof of Proposition 3.2 now follows by an application of the perturbation argument recalled in Section 3.3. Namely, due to Lemma 3.3 and Theorem 3.1 we know that $\widetilde{\pi}$ satisfies the inequality (3.11). Therefore (3.11) follows from the equivalence between π and $\widetilde{\pi}$.

3.5. *Extension to* $\mathbb{Z}$

It is not difficult to extend the result of Theorem 3.1 to processes on $\mathbb{Z}$ rather than $\mathbb{N}$. Namely, consider the process with generator (3.1) for all $n\in\mathbb{Z}$. Here, of course, we do not require $b(0)=0$. Again, we assume reversibility in the form (3.2), which holds now for every $n\in\mathbb{Z}$. Similarly, we choose the function R as in (3.3) for all $n\in\mathbb{Z}$. It is easily checked that all the arguments given in the proof of Theorem 3.1 apply to this case without modification, provided the requirements of assumption (A) are extended to all $n\in\mathbb{Z}$. For instance, this can be used to show that the double sided Poisson measures $\widetilde{\pi}_\lambda(n)=(2e^\lambda-1)^{-1}\lambda^{|n|}/|n|!$, $n\in\mathbb{Z}$, $\lambda\in(0,1)$ satisfy the inequality

$$\mathrm{Ent}_{\widetilde{\pi}_\lambda}(f)\leq\frac{1}{1-\lambda}\mathcal{E}(f,\log f).$$

Indeed, here we may choose $a(n)=\lambda$ for $n\geq 0$ and $b(n)=\lambda$ for $n\leq 0$. This gives $b(n)=n$ for all $n\geq 1$ and $a(n)=-n$ for all $n\leq -1$. In particular, $-\nabla_+a(n)+\nabla_+b(n)=1$ for all $n\neq 0,-1$, in which cases it is equal to $c=1-\lambda$ so that Theorem 3.1 implies the above estimate. Several improvements of this type of estimates can be obtained along the lines discussed in the previous subsections. On the other hand an extension to processes on $\mathbb{Z}^d$, $d\geq 2$, does not appear to be straightforward.

4. Zero range processes

In this section we consider a class of interacting particle systems consisting of finitely many particles moving in a finite set of sites. Particles are neither created nor destroyed. The elements of the set $\{1, 2, \dots, L\}$ label the sites; for $x \in \{1, 2, \dots, L\}$, $\eta_x \in \mathbb{N}$ denotes the number of particles at x. The whole configuration will be denoted by $\eta \in S := \mathbb{N}^L$.

The set G of allowed moves is given by the set of maps from S to S of the form $\eta \mapsto \eta^{xy}$, with $x \neq y \in \{1, 2, \dots, L\}$, and

$$\eta_z^{xy} = \begin{cases} \eta_z & \text{if } z \notin \{x, y\} \text{ or } \eta_x = 0, \\ \eta_x - 1 & \text{for } z = x, \eta_x > 0, \\ \eta_y + 1 & \text{for } z = y, \eta_x > 0. \end{cases}$$

In other words η^{xy} is obtained from η by moving a particle (if any) from the site x to the site y. We simply denote by xy the map $\eta \mapsto \eta^{xy}$, and by ∇_{xy} the corresponding discrete gradient.

For $x \in \{1, 2, \dots, L\}$ consider functions $c_x : \mathbb{N} \to [0, +\infty)$ such that $c_x(0) = 0$, $c_x(n) > 0$ for $n > 0$. $c_x(\eta_x)$ is the rate at which a particle is moved from the site x to a site y chosen with uniform probability. Thus we consider dynamics on S for which the rate $c(\eta, xy)$ for moving a particle from x to y is $L^{-1} c_x(\eta_x)$. Therefore (2.5) becomes

$$\mathcal{L} f(\eta) = \frac{1}{L} \sum_{x,y} c_x(\eta_x) \nabla_{xy} f(\eta), \tag{4.1}$$

where the sum extends to all $x, y \in \{1, \dots, L\}$. The continuous time Markov chain with generator (4.1) is the zero-range process on the complete graph with L vertices.

Note that the total number of particles $N := \sum_x \eta_x$ is conserved. Set $p_x(n) := \prod_{k=1}^n \frac{1}{c_x(k)}$ for $n \geq 1$, $p_x(0) = 1$, and consider the probability π_N, defined on configurations with N particles, with $N = \sum_x \eta_x$, given by

$$\pi_N(\eta) := \frac{1}{Z_N} \prod_{x=1}^{L} p_x(\eta_x),$$

where $Z_N := \sum_{\eta \in S : \sum_x \eta_x = N} \prod_{x=1}^L p_x(\eta_x)$ is the normalization. In what follows the subscripts N will be omitted. In the context of the class of models in Section 2, we see easily that $(xy)^{-1} = yx$ and that the reversibility condition (Rev) holds, because of the identity

$$\pi\big[c_x(\eta_x) g(\eta)\big] = \pi\big[c_y(\eta_y) g\big(\eta^{yx}\big)\big], \tag{4.2}$$

valid for arbitrary functions $g : S \to \mathbb{R}$. We now define the function $R(\eta, \gamma, \delta)$ to be

$$R(\eta, xy, uv) := \frac{1}{L^2} \begin{cases} c_x(\eta_x) c_u(\eta_u) & \text{for } x \neq u, \\ c_x(\eta_x) c_x(\eta_x - 1) & \text{for } x = u, \end{cases} \tag{4.3}$$

where $c(-1)$ is meant to be zero. The symmetry condition (P1) is checked immediately. Also condition (P3) is simple to check. Indeed, xy and uv commute when applied to η unless $\eta_x \eta_u = 0$, but in this latter case $R(\eta, xy, uv) = 0$. Condition (P2) can be checked by direct inspection using (4.2).

In order to use Proposition 2.5 we shall assume:

(A1) All functions $c_x(\cdot)$ are nondecreasing.

Lemma 4.1. *Assume* (A1). *Set* $\Gamma(\eta, \gamma, \delta) = c(\eta, \gamma) c(\eta, \delta) - R(\eta, \gamma, \delta)$, *where* $c(\eta, \gamma) = L^{-1} c_x(\eta_x)$, *whenever* $\gamma = xy$. *Then, for all* $f > 0$

$$\pi\Bigg[\sum_{\gamma,\delta} \Gamma(x, \gamma, \delta) \bigg(\nabla_\gamma f(\eta) \nabla_\delta \log f(\eta) + \frac{\nabla_\gamma f(\eta) \nabla_\delta f(\eta)}{f(\eta)}\bigg)\Bigg]$$

$$\geq \frac{1}{L} \sum_{x,y} \pi\big[c_x(\eta_x) A_x(\eta) \nabla_{xy} f(\eta) \nabla_{xy} \log f(\eta)\big], \tag{4.4}$$

where

$$A_x(\eta) := \big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\left(1 - \frac{1}{2L}\right) - \frac{1}{2L}\sum_{v:v\neq x}\big(c_v(\eta_v + 1) - c_v(\eta_v)\big).$$

Proof. We write

$$\begin{aligned}
&\pi\Bigg[\sum_{\gamma,\delta}\Gamma(\eta,\gamma,\delta)\left(\nabla_\gamma f(\eta)\nabla_\delta \log f(\eta) + \frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right)\Bigg] \\
&\quad = \frac{1}{L^2}\sum_{x,y,v}\pi\left[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\left(\nabla_{xy} f(\eta)\nabla_{xv}\log f(\eta) + \frac{\nabla_{xy} f(\eta)\nabla_{xv} f(\eta)}{f(\eta)}\right)\right] \\
&\quad \geq \frac{1}{L^2}\sum_{x,y,v}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\nabla_{xy} f(\eta)\nabla_{xv}\log f(\eta)\big],
\end{aligned} \tag{4.5}$$

where in (4.5) we used (A1) and the fact that, for every x and η:

$$\sum_{y,v}\nabla_{xy} f(\eta)\nabla_{xv} f(\eta) = \Big[\sum_y \nabla_{xy} f(\eta)\Big]^2 \geq 0.$$

Now observe that

$$\begin{aligned}
&\frac{1}{L^2}\sum_{x,y,v}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\nabla_{xy} f(\eta)\nabla_{xv}\log f(\eta)\big] \\
&\quad = \frac{1}{L^2}\sum_{x,y,v}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\nabla_{xy} f(\eta)\nabla_{xy}\log f(\eta)\big] \\
&\qquad + \frac{1}{L^2}\sum_{x,y,v}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\nabla_{xy} f(\eta)\big(\log f(\eta^{xv}) - \log f(\eta^{xy})\big)\big] \\
&\quad = \frac{1}{L}\sum_{x,y}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big)\nabla_{xy} f(\eta)\nabla_{xy}\log f(\eta)\big] \\
&\qquad + \frac{1}{L^2}\sum_{x,y,v}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big) f(\eta^{xy})\big(\log f(\eta^{xv}) - \log f(\eta^{xy})\big)\big],
\end{aligned} \tag{4.6}$$

where in the last step we simply observed that, by symmetry,

$$\sum_{y,v}\big(\log f(\eta^{xv}) - \log f(\eta^{xy})\big) = 0.$$

We now use reversibility in the form (4.2) to rewrite the last term in (4.6):

$$\begin{aligned}
&\frac{1}{L^2}\sum_{x,y,v}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big) f(\eta^{xy})\big(\log f(\eta^{xv}) - \log f(\eta^{xy})\big)\big] \\
&\quad = \frac{1}{L^2}\sum_{x,y,v:y\neq x}\pi\big[c_y(\eta_y)\big(c_x(\eta_x + 1) - c_x(\eta_x)\big) f(\eta)\big(\log f(\eta^{yv}) - \log f(\eta)\big)\big] \\
&\qquad + \frac{1}{L^2}\sum_{x,v}\pi\big[c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big) f(\eta)\big(\log f(\eta^{xv}) - \log f(\eta)\big)\big]
\end{aligned} \tag{4.7}$$

$$= \frac{1}{L^2} \sum_{x,y,v:v\neq x} \pi\big[c_v(\eta_v)\big(c_x(\eta_x+1)-c_x(\eta_x)\big)f\big(\eta^{vy}\big)\big(\log f(\eta)-\log f\big(\eta^{vy}\big)\big)\big]$$

$$+ \frac{1}{L^2} \sum_{x,y} \pi\big[c_x(\eta_x)\big(c_x(\eta_x)-c_x(\eta_x-1)\big)f\big(\eta^{xy}\big)\big(\log f(\eta)-\log f\big(\eta^{xy}\big)\big)\big]. \tag{4.8}$$

Therefore, exchanging the labels y and v in (4.8) and summing this expression with (4.7) we obtain

$$\frac{1}{L^2} \sum_{x,y,v} \pi\big[c_x(\eta_x)\big(c_x(\eta_x)-c_x(\eta_x-1)\big)f\big(\eta^{xy}\big)\big(\log f\big(\eta^{xv}\big)-\log f\big(\eta^{xy}\big)\big)\big]$$

$$= -\frac{1}{2L^2} \sum_{x,y,v:y\neq x} \pi\big[c_y(\eta_y)\big(c_x(\eta_x+1)-c_x(\eta_x)\big)\nabla_{yv} f(\eta)\nabla_{yv}\log f(\eta)\big]$$

$$- \frac{1}{2L^2} \sum_{x,v} \pi\big[c_x(\eta_x)\big(c_x(\eta_x)-c_x(\eta_x-1)\big)\nabla_{xv} f(\eta)\nabla_{xv}\log f\big].$$

The desired conclusion now follows from (4.5) and (4.6). □

The previous lemma allows us to obtain (MLSI) under the following condition:

(B1) There exist $0 \le \delta < c$ such that for every $x \in \{1, 2, \dots, L\}$ and $n \ge 0$

$$c \le c_x(n+1) - c_x(n) \le c + \delta.$$

Indeed, it is immediately seen that, under (B1),

$$A_x(\eta) \ge \frac{c-\delta}{2}.$$

Therefore, using Proposition 2.5, we have proved the following result.

Theorem 4.2. *If the rates c_x satisfy assumption* (B1) *then the inequality* (2.3) *holds with $\kappa = c - \delta$, uniformly in the number of vertices and the number of particles. In particular,* (MLSI) *holds with the same constant.*

The following remarks give some elements to test the strength of the theorem we have just derived.

4.1. *The independent case*

Consider the case of linear rates, i.e. $c_x(\eta_x) = a_x\eta_x$, for some coefficients $a_x \in (0,\infty)$, $x \in \{1, \dots, L\}$. In this special case the process describes N independent random walks on the complete graph, where each particle jumps from a vertex x to a vertex y with rate a_x. The equilibrium measure π becomes a product of identical single-particle measures π_1, the π_1-probability that the particle is at vertex x being proportional to a_x^{-1}. By the tensorization property of entropy one can then reduce the problem to establishing (MLSI) for a single random walk. Already in this case, Theorem 4.2 gives a non-trivial result. Note that, in the homogeneous case $a_x \equiv 1$, our estimate reduces to the well known bound $\alpha = 1$ for the simple random walk on the unweighted complete graph, see e.g. Example 3.10 in [3].

4.2. *Non-convex decay of entropy*

It is natural to wonder about the necessity of the restriction $\delta < c$ in our assumption (B1). It was shown in [5] that as far as the spectral gap is concerned, inequality (PI) holds for this model with $\gamma \ge c$ as soon as $c_x(\eta_x) - c_x(\eta_x - 1) \ge c$ for all x and η, without further restriction. While we suspect that a similar condition should be sufficient for (MLSI) it is interesting to note that in order to have convexity of the relative entropy along the semigroup some restriction on the growth of the rates is necessary. To see this we consider the following simple example of zero-range process exhibiting non-convex decay of entropy, i.e. such that

$$\pi[\mathcal{L}f\mathcal{L}\log f]+\pi\left[\frac{(\mathcal{L}f)^2}{f}\right]<0, \tag{4.9}$$

for some $f>0$. Take $N=1$ particle only. Note that, since $N=1$ we must have $R(\eta,xy,uv)=0$ for all η and all xy,uv in (4.3). Moreover, set $c_x:=c_x(\eta_x)$, $\pi_x:=\pi(\eta)$ and $f_x:=f(\eta)$ when the particle is at x (i.e. when $\eta_x=1$). Since $\pi_x=Z^{-1}c_x^{-1}$, $Z:=\sum_x c_x^{-1}$, we see that the left-hand side of (4.9) equals

$$\begin{aligned}
&\frac{1}{L^2}\sum_{x,y,z}\pi\left[c_x(\eta_x)^2\left\{\nabla_{xy}f(\eta)\nabla_{xz}\log f(\eta)+\frac{\nabla_{xy}f(\eta)\nabla_{xz}f(\eta)}{f(\eta)}\right\}\right]\\
&\quad=\frac{1}{ZL^2}\sum_x c_x\sum_{y,z}\left\{(f_y-f_x)\log(f_z/f_x)+\frac{(f_y-f_x)(f_z-f_x)}{f_x}\right\}=:\frac{1}{ZL^2}\sum_x c_xQ_x.
\end{aligned}\tag{4.10}$$

This expression can be shown to be negative for suitable choices of $\{f_x\}$ and $\{c_x\}$. A simple example is obtained if e.g. $L=3$, $f_1=1$, $f_2=2$, $f_3=\varepsilon>0$ and $c_1>c_2=c_3=1$. If ε is sufficiently small, in this case we see that $Q_1=\varepsilon(\log 2+\varepsilon+\log\varepsilon)<0$, so that $\sum_x c_xQ_x=c_1Q_1+Q_2+Q_3$ must become negative when c_1 is large. Thus (4.9) holds and the entropy of T_tf is not convex in $t\ge 0$. Clearly, (4.10) can be used to construct many other examples of such a behavior.

4.3. *Non-monotone versus non-homogeneous rates*

The case of non-monotone rates refers to the situation where the rates c_x satisfy the assumptions appearing in Proposition 3.2. Unfortunately, Theorem 4.2 does not extend to this case by simple perturbation arguments. Zero range processes with non-monotone rates have been thoroughly studied in the literature, under the further assumption that the model is homogeneous, i.e. $c_x=c_y$ for all x,y. For the nearest neighbor version of this model, both Poincaré and logarithmic Sobolev inequalities have been established [11,17]. Moreover, the corresponding complete graph model has been shown to satisfy the (MLSI) inequality [7]. These results, all based on some version of the so-called martingale decomposition method, do not extend to non-homogeneous models in a standard way and, as far as we know Theorem 4.2 represents the only criterium available in non-homogeneous models.

5. Bernoulli–Laplace models

As in previous section we consider a system of particles moving in the finite set of sites $\{1,2,\dots,L\}$; here we assume that particles are subject to an exclusion rule, namely at most one particle per site is allowed. Thus $S:=\{0,1\}^L$. The set of allowed moves is $G:=\{xy\colon x,y\in\{1,2,\dots,L\}, x\neq y\}$, where, for $\eta\in S$, $\eta^{xy}=\eta$ unless $\eta_x(1-\eta_y)=1$, and in this case

$$\eta_z^{xy}=\begin{cases}\eta_z & \text{if } z\notin\{x,y\},\\ 0 & \text{for } z=x,\\ 1 & \text{for } z=y.\end{cases}$$

To each site x we associate a Poisson clock of constant intensity $\lambda_x>0$; when the clock of site x rings, a site y is chosen at random: if $\eta_x=1$ and $\eta_y=0$ then the particle at x moves to y, otherwise nothing happens. This dynamics corresponds to the infinitesimal generator

$$\mathcal{L}f(\eta):=\frac{1}{L}\sum_{x,y=1}^{L}\lambda_x\eta_x(1-\eta_y)\nabla_{xy}f(\eta).$$

In other words, we set $c(\eta, xy) = L^{-1}\lambda_x \eta_x(1-\eta_y)$ in (2.5). Denote by $N \le L$ the number of particles in the system; since it is conserved by the dynamics, we can consider the restriction of the dynamics to configurations with N particles. In this restricted state space there is a unique stationary distribution π_N, given by conditioning to configurations with N particles the product of Bernoulli measures with parameters $\frac{1}{1+\lambda_x}$. More precisely

$$\pi_N(\eta) = \frac{1}{Z_{L,N}} \prod_{x=1}^{L} \left(\frac{1}{1+\lambda_x}\right)^{\eta_x} \left(\frac{\lambda_x}{1+\lambda_x}\right)^{1-\eta_x}.$$

We will refer to this measure as the *canonical* measure on $\{1, 2, \dots, L\}$. The reversibility condition (Rev) holds true for π_N; the subscript N will be omitted from now on. Note that, for $N = 1$, zero range processes coincide with Bernoulli–Laplace models. In particular, the counterexample in Section 4.2 concerning the non-convex decay of entropy applies here too. Thus, some bound on the non homogeneity of the model is needed. The following condition, which is presumably not optimal, is analogous to condition (B1) for zero range processes:

(B2) There exists $0 \le \delta < c$ such that for every $x = 1, 2, \dots, L$

$$c \le \lambda_x \le c + \delta.$$

Theorem 5.1. *Assume* (B2). *Then the inequality* (2.3) *holds with* $\kappa = c - \delta$. *In particular,* (MLSI) *holds with* $\alpha = c - \delta$.

It should be stressed that because in the present case only one particle per site is allowed, the proof of Theorem 5.1 requires some arguments that were not needed in the proof of Theorem 4.2.

Proof of Theorem 5.1. We use, of course, Proposition 2.5. The choice of R is essentially forced by the commutation condition (P3). We set

$$R(\eta, xy, zu) := \frac{1}{L^2} \begin{cases} \lambda_x \lambda_z \eta_x(1-\eta_y)\eta_z(1-\eta_u) & \text{for } \left|\{x, y, z, u\}\right| = 4, \\ 0 & \text{otherwise.} \end{cases}$$

We obtain, after having noticed that $\nabla_{xy} f \nabla_{yz} g \equiv 0 \equiv \nabla_{xy} f \nabla_{ux} g$ for any choice of x, y, z, u and f, g,

$$\begin{aligned} &\pi\left[\sum_{\gamma,\delta} \Gamma(\eta, \gamma, \delta)\left(\nabla_\gamma f(\eta) \nabla_\delta \log f(\eta) + \frac{\nabla_\gamma f(\eta) \nabla_\delta f(\eta)}{f(\eta)}\right)\right] \\ &\quad = \frac{1}{L^2} \sum_{\substack{x,y,z:\\ |\{x,y,z\}|=3}} \pi\left[\lambda_x^2 \nabla_{xy} f \nabla_{xz} \log f\right] + \frac{1}{L^2} \sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi[\lambda_x \lambda_u \nabla_{xy} f \nabla_{uy} \log f] \\ &\qquad + \frac{1}{L^2} \sum_{x,y} \pi\left[\lambda_x^2 \nabla_{xy} f \nabla_{xy} \log f\right] + \frac{1}{L^2} \sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\left[\lambda_x \lambda_u \frac{\nabla_{xy} f \nabla_{uy} f}{f}\right] \\ &\qquad + \frac{1}{L^2} \sum_{\substack{x,y,z:\\ |\{x,y,z\}|=3}} \pi\left[\lambda_x^2 \frac{\nabla_{xy} f \nabla_{xz} f}{f}\right] + \frac{1}{L^2} \sum_{x,y} \pi\left[\lambda_x^2 \frac{(\nabla_{xy} f)^2}{f}\right]. \end{aligned} \tag{5.1}$$

The sum of the last two terms in (5.1) equals

$$\frac{1}{L^2} \sum_x \lambda_x^2 \pi\left[\frac{(\sum_y \nabla_{xy} f)^2}{f}\right] \ge 0. \tag{5.2}$$

We now claim that

$$\sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\left[\lambda_x\lambda_u \frac{\nabla_{xy} f \nabla_{uy} f}{f}\right] \geq 0 \tag{5.3}$$

for every $f > 0$. To prove (5.3) we observe that

$$\begin{aligned}\sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\left[\lambda_x\lambda_u \frac{\nabla_{xy} f \nabla_{uy} f}{f}\right] &= \sum_T \sum_{\substack{x,y,z\in T:\\ |\{x,y,z\}|=3}} \pi\left[\lambda_x\lambda_u \frac{\nabla_{xy} f \nabla_{uy} f}{f}\right]\\ &= \sum_T \sum_{\substack{x,y,z\in T:\\ |\{x,y,z\}|=3}} \pi\left\{\pi\left[\lambda_x\lambda_u \frac{\nabla_{xy} f \nabla_{uy} f}{f}\Big|\eta_{T^c}\right]\right\},\end{aligned}$$

where T varies over $\{T \subseteq \{1, 2, \dots, L\}\colon |T| = 3\}$, and $\pi[\cdot|\eta_{T^c}]$ denotes the conditional expectation with respect to the configuration outside T, which we denote by η_{T^c}. Note that the corresponding conditional measure is the canonical measure on T with $N - \sum_{x\notin T} \eta_x$ particles. Inequality (5.3) is proved if we show that for every fixed T and η_{T^c}

$$\sum_{\substack{x,y,z\in T:\\ |\{x,y,z\}|=3}} \pi\left[\lambda_x\lambda_u \frac{\nabla_{xy} f \nabla_{uy} f}{f}\Big|\eta_{T^c}\right] \geq 0. \tag{5.4}$$

Set $T = \{a, b, c\}$ and note that if $\sum_{x\notin T} \eta_x \in \{N, N-3\}$ then T contains either no particle or no hole, so that $\nabla_{xy} f \equiv 0$ for every f and every $x, y \in T$. Similarly, for $\sum_{x\notin T} \eta_x = N - 1$ there is only one particle in T, so $\nabla_{xy} f \nabla_{uy} f \equiv 0$ for $x \neq u$. So we only need to consider the case $\sum_{x\notin T} \eta_x = N - 2$, which means there is exactly one hole in T. For a given configuration η_{T^c} outside of T, we denote by α the value of f on the configuration with the hole in a. Similarly, β is the value of f when the hole is in b and γ when the hole is in c. Then, by direct computation we get

$$\begin{aligned}&\frac{Z_{T,2}}{2} \sum_{\substack{x,y,z\in T:\\ |\{x,y,z\}|=3}} \pi\left[\lambda_x\lambda_u \frac{\nabla_{xy} f \nabla_{uy} f}{f}\Big|\eta_{T^c}\right]\\ &\quad= \frac{\lambda_a}{1+\lambda_a}\frac{1}{1+\lambda_b}\frac{1}{1+\lambda_c}\lambda_b\lambda_c\frac{(\beta-\alpha)(\gamma-\alpha)}{\alpha}\\ &\qquad+ \frac{\lambda_b}{1+\lambda_b}\frac{1}{1+\lambda_a}\frac{1}{1+\lambda_c}\lambda_a\lambda_c\frac{(\gamma-\beta)(\alpha-\beta)}{\beta} + \frac{\lambda_c}{1+\lambda_c}\frac{1}{1+\lambda_a}\frac{1}{1+\lambda_b}\lambda_a\lambda_b\frac{(\beta-\gamma)(\alpha-\gamma)}{\gamma}\\ &\quad= \frac{\lambda_a\lambda_b\lambda_c}{(1+\lambda_a)(1+\lambda_b)(1+\lambda_c)}\left[\frac{\beta\gamma}{\alpha} + \frac{\alpha\gamma}{\beta} + \frac{\alpha\beta}{\gamma} - \alpha - \beta - \gamma\right],\end{aligned} \tag{5.5}$$

where $Z_{T,2}$ is the normalization factor of the canonical measure on T with 2 particles, and the further factor $1/2$ in the l.h.s. of (5.5) is due to the fact that the sum over the "particles" x, u is a sum over ordered pairs. We need to show that the r.h.s. of (5.5) is non-negative, for every $\alpha, \beta, \gamma > 0$. Since the expression is homogeneous of degree one and invariant for permutations of the variables, we may restrict to $\gamma = 1$, $\alpha, \beta \geq 1$. In other words we need to show that

$$F(\alpha, \beta) := \frac{\beta}{\alpha} + \frac{\alpha}{\beta} + \alpha\beta - \alpha - \beta - 1 \geq 0, \tag{5.6}$$

for every $\alpha, \beta \geq 1$. Since $z + z^{-1} \geq 2$ for all $z > 0$, we have

$$F(\alpha, \beta) \geq 1 + \alpha\beta - \alpha - \beta = (\alpha - 1)(\beta - 1) \geq 0,$$

and (5.6), follows. This shows (5.4).

From (5.1)–(5.4) we get

$$\begin{aligned}
&\pi\Big[\sum_{\gamma,\delta}\Gamma(x,\gamma,\delta)\Big(\nabla_\gamma f(\eta)\nabla_\delta \log f(\eta)+\frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\Big)\Big]\\
&\quad\geq \frac{1}{L^2}\sum_{\substack{x,y,z:\\ |\{x,y,z\}|=3}}\pi\big[\lambda_x^2\nabla_{xy} f\nabla_{xz}\log f\big]+\frac{1}{L^2}\sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}}\pi[\lambda_x\lambda_u\nabla_{xy} f\nabla_{uy}\log f]\\
&\qquad+\frac{1}{L^2}\sum_{x,y}\pi\big[\lambda_x^2\nabla_{xy} f\nabla_{xy}\log f\big]\\
&\quad=\frac{1}{L^2}\sum_{x,y,z}\pi\big[\lambda_x^2\nabla_{xy} f\nabla_{xz}\log f\big]+\frac{1}{L^2}\sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}}\pi[\lambda_x\lambda_u\nabla_{xy} f\nabla_{uy}\log f]. \qquad (5.7)
\end{aligned}$$

We now deal separately with the last two terms in (5.7), similarly to what done in Section 4. For the first term we have

$$\begin{aligned}
&\sum_{x,y,z}\pi\big[\lambda_x^2\nabla_{xy} f\nabla_{xz}\log f\big]=\sum_{x,y,z}\pi\big[\lambda_x^2\eta_x(1-\eta_y)(1-\eta_z)\nabla_{xy} f\nabla_{xz}\log f\big]\\
&\quad=\sum_{x,y,z}\pi\big[\lambda_x^2\eta_x(1-\eta_y)(1-\eta_z)\nabla_{xy} f\nabla_{xy}\log f\big]\\
&\qquad+\sum_{x,y,z}\pi\big[\lambda_x^2\eta_x(1-\eta_y)(1-\eta_z)\nabla_{xy} f(\eta)\big[\log f\big(\eta^{xz}\big)-\log f\big(\eta^{xy}\big)\big]\big]\\
&\quad=(L-N)\sum_{x,y}\pi\big[\lambda_x^2\nabla_{xy} f\nabla_{xy}\log f\big]\\
&\qquad+\sum_{x,y,z}\pi\big[\lambda_x^2\eta_x(1-\eta_y)(1-\eta_z)f\big(\eta^{xy}\big)\big[\log f\big(\eta^{xz}\big)-\log f\big(\eta^{xy}\big)\big]\big], \qquad (5.8)
\end{aligned}$$

where we use the fact that, for any z and η,

$$\sum_z(1-\eta_z)=L-N,$$

and the fact that, for any x and η:

$$\sum_{y,z}(1-\eta_y)(1-\eta_z)\big[\log f\big(\eta^{xz}\big)-\log f\big(\eta^{xy}\big)\big]=0.$$

Using reversibility we have

$$\begin{aligned}
&\sum_{x,y,z}\pi\big[\lambda_x^2\eta_x(1-\eta_y)(1-\eta_z)f\big(\eta^{xy}\big)\big[\log f\big(\eta^{xz}\big)-\log f\big(\eta^{xy}\big)\big]\big]\\
&\quad=\sum_{x,y,z:z\neq x}\pi\big[\lambda_x\lambda_y\eta_y(1-\eta_x)(1-\eta_z)f(\eta)\nabla_{yz}\log f(\eta)\big]\\
&\quad=-\sum_{x,y,z:y\neq x}\pi\big[\lambda_x\lambda_z\eta_z(1-\eta_y)(1-\eta_x)f\big(\eta^{zy}\big)\nabla_{zy}\log f(\eta)\big]\\
&\quad=-\frac{1}{2}\sum_{x,y,z:z\neq x}\pi\big[\lambda_x\lambda_y\eta_y(1-\eta_x)(1-\eta_z)\nabla_{yz} f\nabla_{yz}\log f\big], \qquad (5.9)
\end{aligned}$$

where we use permutation of x, y, z. By condition (B2) we easily have

$$\sum_{x:z\neq x} \lambda_x(1-\eta_x)(1-\eta_z) \le (c+\delta)(L-N-1)(1-\eta_z),$$

which, together with (5.8) and (5.9) yields

$$\sum_{x,y,z} \pi\big[\lambda_x^2 \nabla_{xy} f \nabla_{xz} \log f\big] \ge \left[c(L-N) - \frac{c+\delta}{2}(L-N-1)\right] \sum_{x,y} \pi[\lambda_x \nabla_{xy} f \nabla_{xy} \log f]. \tag{5.10}$$

The last summand in (5.7) is dealt with similarly:

$$\begin{aligned}
&\sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi[\lambda_x\lambda_u \nabla_{xy} f \nabla_{uy} \log f] \qquad (5.11)\\
&\quad= \sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\big[\lambda_x\lambda_u\eta_x\eta_u(1-\eta_y)\nabla_{xy} f \nabla_{uy}\log f\big]\\
&\quad= \sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\big[\lambda_x\lambda_u\eta_x\eta_u(1-\eta_y)\nabla_{xy} f \nabla_{xy}\log f\big]\\
&\qquad+ \sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\big[\lambda_x\lambda_u\eta_x\eta_u(1-\eta_y) f\big(\eta^{xy}\big)\big[\log f\big(\eta^{uy}\big) - \log f\big(\eta^{xy}\big)\big]\big]\\
&\quad\ge c(N-1)\sum_{x,y}\pi[\lambda_x \nabla_{xy} f \nabla_{xy}\log f] + \sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\big[\lambda_y\lambda_u\eta_y(1-\eta_x)\eta_u f(\eta)\nabla_{ux}\log f(\eta)\big]\\
&\quad= c(N-1)\sum_{x,y}\pi[\lambda_x \nabla_{xy} f \nabla_{xy}\log f] - \sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\big[\lambda_y\lambda_x\eta_y\eta_x(1-\eta_u) f\big(\eta^{xu}\big)\nabla_{xu}\log f(\eta)\big]\\
&\quad= c(N-1)\sum_{x,y}\pi[\lambda_x \nabla_{xy} f \nabla_{xy}\log f] - \frac{1}{2}\sum_{\substack{x,y,u:\\ |\{x,y,u\}|=3}} \pi\big[\lambda_x\lambda_u\eta_x\eta_u(1-\eta_y)\nabla_{xy} f \nabla_{xy}\log f\big]\\
&\quad\ge \left[c(N-1) - \frac{c+\delta}{2}(N-1)\right]\sum_{x,y}\pi[\lambda_x \nabla_{xy} f \nabla_{xy}\log f]. \qquad (5.12)
\end{aligned}$$

By (5.7), (5.10) and (5.11) we end up with

$$\pi\left[\sum_{\gamma,\delta}\Gamma(x,\gamma,\delta)\left(\nabla_\gamma f(\eta)\nabla_\delta \log f(\eta) + \frac{\nabla_\gamma f(\eta)\nabla_\delta f(\eta)}{f(\eta)}\right)\right] \ge \frac{c-\delta}{2L}\sum_{x,y}\pi[\lambda_x \nabla_{xy} f \nabla_{xy}\log f]$$

which, together with Proposition 2.5, completes the proof. □

We conclude with a comparison of our bound with previously known results, that are limited to the homogeneous case $\lambda_x \equiv 1$. In this case we can take $c = 1$ and $\delta = 0$ in Theorem 5.1, therefore obtaining the (MLSI) with $\alpha = 1$. By using Yau's martingale method, Gao and Quastel [13] and Goel [14] have proved (MLSI) for this model with $\alpha = 1/2$ (asymptotically as $L \to \infty$). The same estimate has been obtained with a similar approach by Bobkov and Tetali [3]. We mention that a proof of our bound $\alpha = 1$ in the homogeneous case can also be obtained by a slightly different method [8].

It is not hard to show that the homogeneous Bernoulli–Laplace model has a spectral gap equal to 1, independent of the number of particles (see e.g. [5]). Therefore, from (1.6) we have that the best constant α in (MLSI) satisfies $\alpha \in [1, 2]$. The optimal value is not known, even in the case of one particle (random walk on the complete graph), see e.g. the discussion of Example 3.10 in [3]. Moreover, it is not known whether the optimal constant depends on the number of particles.

For the non-homogeneous model considered here only the spectral gap has been obtained before, see [6], where a uniform Poincaré inequality is established under the assumption that $C^{-1} \le \lambda_x \le C$ for some $C > 0$, for all x.

References

[1] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX* 177–206. *Lecture Notes in Math.* **1123**. Springer, Berlin, 1985. MR0889476

[2] S. G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.* **156** (1998) 347–365. MR1636948

[3] S. G. Bobkov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.* **19** (2006) 289–336. MR2283379

[4] S. Bochner. Vector fields and Ricci curvature. *Bull. Amer. Math. Soc.* **52** (1946) 776–797. MR0018022

[5] A. S. Boudou, P. Caputo, P. Dai Pra and G. Posta. Spectral gap estimates for interacting particle systems via a Bochner-type identity. *J. Funct. Anal.* **232** (2006) 222–258. MR2200172

[6] P. Caputo. Spectral gap inequalities in product spaces with conservation laws. In *Advanced Studies in Pure Mathematics* **39**. H. Osada and T. Funaki (Eds). Math. Soc. Japan, Tokyo, 2004. MR2073330

[7] P. Caputo and G. Posta. Entropy dissipation estimates in a zero-range dynamics. *Probab. Theory Related Fields* **139** (2007) 65–87. MR2322692

[8] P. Caputo and P. Tetali. Unpublished notes, 2005.

[9] D. Chafai. Binomial–Poisson entropic inequalities and the $M/M/\infty$ queue. *ESAIM Probab. Stat.* **10** (2006) 317–339. MR2247924

[10] P. Dai Pra, A. M. Paganoni and G. Posta. Entropy inequalities for unbounded spin systems. *Ann. Probab.* **30** (2002) 1959–1976. MR1944012

[11] P. Dai Pra and G. Posta. Logarithmic Sobolev inequality for zero-range dynamics. *Ann. Probab.* **33** (2005) 2355–2401. MR2184099

[12] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** (1996) 695–750. MR1410112

[13] F. Gao and J. Quastel. Exponential decay of entropy in the random transposition and Bernoulli–Laplace models. *Ann. Appl. Probab.* **13** (2003) 1591–1600. MR2023890

[14] S. Goel. Modified logarithmic Sobolev inequalities for some models of random walk. *Stochastic Process. Appl.* **114** (2004) 51–79. MR2094147

[15] O. Johnson. Log-concavity and the maximum entropy property of the Poisson distribution. *Stochastic Process. Appl.* **117** (2007) 791–802. MR2327839

[16] A. Joulin. Poisson-type deviation inequalities for curved continuous-time Markov chains. *Bernoulli* **13** (2007) 782–798. MR2348750

[17] C. Landim, S. Sethuraman and S. R. S. Varadhan. Spectral gap for zero-range dynamics. *Ann. Probab.* **24** (1996) 1871–1902. MR1415232

[18] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités XXXV* 167–194. *Lecture Notes in Math.* **1755**. Springer, Berlin, 2001. MR1837286

[19] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case. *Comm. Math. Phys.* **161** (1994) 487–514. MR1269388

[20] L. Miclo. An example of application of discrete Hardy's inequalities. *Markov Process. Related Fields* **5** (1999) 319–330. MR1710983

[21] D. W. Stroock and B. Zegarliński. The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Comm. Math. Phys.* **149** (1992) 175–193. MR1182416

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

[23] H.-T. Yau. Logarithmic Sobolev inequality for lattice gases with mixing conditions. *Comm. Math. Phys.* **181** (1996) 367–408. MR1414837