ASC Report No. 36/2015

Discrete Beckner inequalities via the Bochner-Bakry-Emery approach for Markov chains

A. Jüngel and W. Yue

Institute for Analysis and Scientific Computing — Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-05-6

Most recent ASC Reports

35/2015 *J. Burkotová, I. Rachunková, M. Hubner, E. Weinmüller*
Numerical evidence of Kneser solutions to a class of singular BVPs in ODEs.

34/2015 *O. Koch, S. Schirnhofer, E. Weinmüller*
Numerical simulation of the Korteweg-de Vries Equation for shallow water waves.

33/2015 *S. Börm and J.M. Melenk*
Approximation of the high-frequency Helmholtz kernel by nested directional interpolation

32/2015 *F. Achleitner, A. Arnold, E.A. Carlen*
On hypocoercive BGK models

31/2015 *M. Feischl, T. Führer, M. Niederer, S. Strommer, A. Steinboeck, and D. Praetorius*
Efficient numerical computation of direct exchange areas in thermal radiation analysis

30/2015 *B. Düring, P. Fuchs, and A. Jüngel*
A higher-order gradient flow scheme for a singular one-dimensional diffusion equation

29/2015 *C. Erath and D. Praetorius*
Adaptive finite volume methods with convergence rates

28/2015 *W. Auzinger, O. Koch, M. Schöbinger, E. Weinmüller*
A new version of the code bvpsuite for singular BVPs in ODEs: Nonlinear solver and its application to m-Laplacians.

27/2015 *C. Lehrenfeld, J. Schöberl*
High order exactly divergence-free Hybrid Discontinuous Galerkin Methods for unsteady incompressible flows

26/2015 *A. Jüngel, and S. Schuchnigg*
Entropy-dissipating semi-discrete Runge-Kutta schemes for nonlinear diffusion equations

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-05-6

DISCRETE BECKNER INEQUALITIES VIA THE BOCHNER-BAKRY-EMERY APPROACH FOR MARKOV CHAINS

ANSGAR JÜNGEL AND WEN YUE

Abstract. Discrete Beckner inequalities, which interpolate between the modified logarithmic Sobolev inequality and the Poincaré inequality, are derived for time-continuous Markov chains on countable state spaces. The proof is based on the Bakry-Emery approach and on discrete Bochner-type inequalities established by Caputo, Dai Pra, and Posta and recently extended by Fathi and Maas. The abstract result is applied to several Markov chains, including birth-death processes, zero-range processes, Bernoulli-Laplace models, and random transportation models, and to a finite-volume discretization of a one-dimensional Fokker-Planck equation, applying results by Mielke.

1. Introduction

Convex Sobolev inequalities such as Poincaré and logarithmic Sobolev inequalities play an important role in the analysis of the convergence to stationarity for Markov processes. Besides implying exponential decay of the entropy, it is known that these functional inequalities give useful concentration bounds [6] and hypercontractivity of the corresponding semigroup [14], and they are a natural tool to estimate mixing times [25]. There exists an extensive literature on the derivation of Poincaré inequalities (or spectral gap estimates) and logarithmic Sobolev (or shorter: log-Sobolev) inequalities in the discrete and continuous setting; see, e.g., the reviews [14, 19, 25] and the books [1, 3, 27]. An algorithm for the computation of the spectral gap is presented in [13], while corresponding estimates can be found in [8, 11, 9]. For log-Sobolev inequalities, we refer to [5, 10, 20].

There are much less results on Beckner inequalities for Markov chains, which interpolate between the Poincaré inequality and log-Sobolev inequality [4]. We are only aware of the paper by Bobkov and Tetali [6], where estimates on the constant of this inequality were derived for Bernoulli-Laplace and random transposition models. In this paper, we establish new bounds for discrete Beckner inequalities for stochastic processes not studied in [6].

The technique of proof is the Bochner-Bakry-Emery method of Caputo et al. [10], which was recently extended by Fathi and Maas in [15] in the context of Ricci curvature bounds. The idea of the Bakry-Emery approach is to relate the second time derivative of the entropy

Date: November 19, 2015.

2010 *Mathematics Subject Classification.* 60J27, 39B62, 60J80.

Key words and phrases. Time-continuous Markov chain, functional inequality, entropy decay, discrete Beckner inequality, stochastic particle systems.

The authors acknowledge partial support from the Austrian Science Fund (FWF), grants P24304, P27352, and W1245.

to its entropy production. This relation is achieved by employing a discrete Bochner-type equation which replaces the Bochner identity in the continuous case.

In order to make these ideas precise, consider a time-homogeneous Markov process $(X_t)_{t\geq 0}$ with values in a countable state space S, having an invariant measure π. We assume that the semigroup $(T_t)_{t\geq 0}$, defined on $L^2(\pi)$ by $T_t f(x) = \mathrm{E}[f(X_t) : X_0 = x]$, is strongly right continuous, so that the infinitesimal generator $\mathcal{L}$ exists, $T_t = e^{t\mathcal{L}}$. Given a probability measure μ on S, we denote by μT_t the distribution of X_t assuming that X_0 is distributed according to μ. The rate of convergence of μT_t to the invariant measure π is a major topic in probability theory. It can be achieved by estimating the time derivative of the relative entropy.

Before explaining the entropy decay, we introduce some notation. The relative entropy $h(\mu|\pi)$ of μ with respect to π is defined by

$$h(\mu|\pi) = \pi\left[\phi\left(\frac{d\mu}{d\pi}\right)\right] = \sum_{\eta\in S}\pi(\eta)\phi\left(\frac{d\mu}{d\pi}\right)(\eta),$$

where $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is a convex function with $\phi(1) = 0$, $\mathbb{R}_+ = [0,\infty)$, and $h(\mu|\pi)$ is meant to be infinite whenever $\mu \not\ll \pi$ or $\phi(d\mu/d\pi) \notin L^1(\pi)$. The entropy can be defined on the set of probability densities f such that $\phi(f) \in L^1(\pi)$ by

$$\mathrm{Ent}_\pi(f) = \pi[\phi(f)],$$

so that $h(\mu|\pi) = \mathrm{Ent}_\pi(d\mu/d\pi)$. If $\phi_1(s) = s(\log s - 1) + 1$, we obtain the logarithmic entropy and if $\phi_2(s) = s^2 - 1$, $\mathrm{Ent}_\pi(f)$ equals the variance of f, $\mathrm{Var}_\pi(f) = \pi[f^2] - \pi[f]^2$. In this paper, we choose $\phi_\alpha(s) = s^\alpha - 1$ for $1 < \alpha \leq 2$, which interpolates between ϕ_1 and ϕ_2.

Let $\rho_t = d(\mu T_t)/d\pi$ be the probability density of the Markov chain at time $t \geq 0$. We assume that the Markov chain is reversible, i.e., the generator is self-adjoint in $L^2(\pi)$. Then ρ_t solves the Kolmogorov equation $\partial_t\rho_t = \mathcal{L}\rho_t$, $t > 0$. The idea of Bakry and Emery [2] is to differentiate the entropy twice with respect to time. A formal computation gives

$$\begin{aligned} \frac{d}{dt}\mathrm{Ent}_\pi(\rho_t) &= -\mathcal{E}(\phi'(\rho_t), \rho_t), \\ \frac{d^2}{dt^2}\mathrm{Ent}_\pi(\rho_t) &= \pi\big[\mathcal{L}\phi'(\rho_t)\mathcal{L}\rho_t + \phi''(\rho_t)(\mathcal{L}\rho_t)^2\big], \end{aligned} \tag{1}$$

where $\mathcal{E}(f,g) = -\pi[f\mathcal{L}g]$ is the Dirichlet form of $\mathcal{L}$. Now suppose that the following inequality holds for some $\lambda > 0$:

$$\pi\big[\mathcal{L}\phi'(\rho)\mathcal{L}\rho + \phi''(\rho)(\mathcal{L}\rho)^2\big] \geq \lambda\mathcal{E}(\phi'(\rho), \rho), \quad t > 0. \tag{2}$$

This is equivalent to $\partial_t^2\mathrm{Ent}_\pi(\rho) + \lambda\partial_t\mathrm{Ent}_\pi(\rho) \geq 0$, and by Gronwall's lemma, we conclude that $\partial_t\mathrm{Ent}_\pi(\rho_t)$ converges to zero with exponential rate. Furthermore, integration over (t,∞) leads to

$$\frac{d}{dt}\mathrm{Ent}_\pi(\rho) + \lambda\mathrm{Ent}_\pi(\rho) \leq 0, \quad t > 0, \tag{3}$$

if we know that $\mathrm{Ent}_\pi(\rho_t) \to 0$ as $t \to \infty$. On the one hand, this implies exponential convergence of the relative entropy to zero and on the other hand, (3) is equivalent to the convex Sobolev inequality $\lambda \mathrm{Ent}_\pi(f) \le \mathcal{E}(\phi'(f), f)$, valid for all probability densities f.

For the special cases $\phi_1(s) = s(\log s - 1)$ and $\phi_2(s) = s^2 - 1$, we obtain the *modified log-Sobolev inequality* and *Poincaré inequality*, respectively,

$$\lambda_M \mathrm{Ent}_\pi(f) \le \mathcal{E}(\log f, f), \quad \lambda_P \mathrm{Var}_\pi(f) \le \mathcal{E}(f, f). \tag{4}$$

Note that if $\mathcal{L}$ is the generator of a reversible diffusion process, we may write $\mathcal{E}(\log f, f) = 4\mathcal{E}(f^{1/2}, f^{1/2})$, so the log-Sobolev inequality $\lambda_L \mathrm{Ent}_\pi(f) \le \mathcal{E}(f^{1/2}, f^{1/2})$ and the first inequality in (4) coincide with $\lambda_M = 4\lambda_L$. This is generally not true for Markov processes with jumps [5], but for reversible processes, the relations $4\lambda_L \le \lambda_M \le 2\lambda_P$ hold [6, 14].

The aim of this paper is to derive explicit constants $\lambda > 0$ such that the *Beckner inequality*

$$\lambda \mathrm{Ent}_\pi(\rho) \le \alpha \mathcal{E}(\rho^{\alpha-1}, \rho), \quad 1 < \alpha \le 2, \tag{5}$$

and the exponential entropy decay

$$\mathrm{Ent}_\pi(\rho_t) \le e^{-\lambda t} \mathrm{Ent}_\pi(f), \quad t > 0, \tag{6}$$

hold. According to the above discussion, this is achieved by proving (2), and the proof of this inequality is based on a discrete Bochner-type identity. The idea to employ such an identity was first presented in [8], elaborated later in [10, 15] and going back to [7]. The identity is obtained by identifying the Radon-Nikodym derivative of a measure involving the jump rates of the Markov chain [8, Section 2]. This allows one to relate terms with different orders of "discrete derivatives" occuring in $\mathcal{L}$. For details, we refer to Section 2. Our technique of proving (5) is completely different from the work [6], where an iteration method was used to derive discrete Beckner inequalities.

Fathi and Maas [15] extended the results of Caputo et al. [10]. The key idea of [15] (and, by the way, of [23]) is the use of the logarithmic mean

$$\rho^*(\eta, \xi) = \frac{\rho(\eta) - \rho(\xi)}{\log \rho(\eta) - \log \rho(\xi)}$$

in the analysis. The logarithmic mean allows for the discrete chain rule $\rho^* \nabla \log \rho = \nabla \rho$, where $\nabla \rho(\eta, \xi) = \rho(\eta) - \rho(\xi)$, which naturally holds in the continuous case. This chain rule is needed to treat the logarithmic entropy. In the case of power-type entropies, it is natural to replace the logarithmic mean by the power mean

$$\widehat{\rho}(\eta, \xi) = \frac{\rho(\eta) - \rho(\xi)}{\rho(\eta)^{\alpha-1} - \rho(\xi)^{\alpha-1}}, \quad 1 < \alpha \le 2,$$

which satisfies the discrete chain rule $\widehat{\rho} \nabla \rho^{\alpha-1} = \nabla \rho$ since $\widehat{\rho}$ "approximates" $\rho^{2-\alpha}/(\alpha-1)$. We remark that the idea to enforce a discrete chain rule is well known in the design of structure-preserving numerical schemes and was used, e.g., in the construction of entropy-conservative finite-volume fluxes [16] and in the discrete variational derivative method [17].

Our analysis is similar to that of [10, 15], but the properties of the power mean and logarithmic mean partially differ and we need to take into account the dependence on the parameter α. Moreover, we derive new discrete convex Sobolev inequalities for a number of stochastic processes.

We detail the Bochner-Bakry-Emery method in Section 2. The validity of the discrete Beckner inequality (5) is reduced to the validity of an inequality which is a "discrete" version of (2). In Section 3, we apply the general technique to four stochastic processes (as in [15]): birth-death processes, zero-range processes, Bernoulli-Laplace models, and random transportation models. Furthermore, the results for birth-death processes are applied to a finite-volume discretization of a one-dimensional Fokker-Planck equation, yielding exponential decay of the discrete entropy. The proof consists of a combination of the Beckner inequality for birth-death processes and the results of Mielke [23], who proved exponential decay for the logarithmic entropy.

Our main conclusion is that the Bochner-Bakry-Emery approach is sufficiently flexible to be applicable to non-logarithmic functions.

2. The Bochner method

Let an irreducible and reversible Markov chain on a countable state space S be given and let π be the invariant measure. We write the generator $\mathcal{L}$ in the form

$$\mathcal{L}f(\eta) = \sum_{\gamma\in G} c(\eta,\gamma)\nabla_\gamma f(\eta),$$

where G is the set of allowed moves (represented by functions $\gamma : S \to S$), the mapping $c : S\times G \to [0,\infty)$ represents the jump rates, and $\nabla_\gamma f(\eta) = f(\gamma\eta)-f(\eta)$. We observe that the generator of every countable Markov chain can be written in this form. We assume the following two properties: For any $\gamma \in G$, there exists $\gamma^{-1} \in G$ satisfying $\gamma^{-1}\gamma\eta = \eta$ for all $\eta \in S$ with $c(\eta,\gamma) > 0$. Furthermore, the reversibility condition

$$\pi\bigg[\sum_{\gamma\in G} c(\eta,\gamma)F(\eta,\gamma)\bigg] = \pi\bigg[\sum_{\gamma\in G} c(\eta,\gamma)F(\gamma\eta,\gamma^{-1})\bigg]$$

holds for all $F : S \times G \to \mathbb{R}$. Under reversibility, the Dirichlet form can be written as

$$\mathcal{E}(f,g) = \frac{1}{2}\pi\bigg[\sum_{\gamma\in G} c(\eta,\gamma)\nabla_\gamma f(\eta)\nabla_\gamma g(\eta)\bigg]. \tag{7}$$

For the discrete Bochner-type identity, we suppose as in [10]:

Assumption 1. *There exists a function $R : S \times G \times G \to \mathbb{R}$ such that*
(i) $R(\eta,\gamma,\delta) = R(\eta,\delta,\gamma)$ *for all $\eta \in S$, γ, $\delta \in G$;*
(ii) *for all bounded functions $\psi : S \times G \times G \to \mathbb{R}$,*

$$\pi\bigg[\sum_{\gamma,\delta\in G} R(\eta,\gamma,\delta)\psi(\eta,\gamma,\delta)\bigg] = \pi\bigg[\sum_{\gamma,\delta\in G} R(\eta,\gamma,\delta)\psi(\gamma\eta,\gamma^{-1},\delta)\bigg].$$

(iii) $\gamma\delta\eta = \delta\gamma\eta$ *for all $\eta \in S$, γ, $\delta \in G$ with $R(\eta,\gamma,\delta) > 0$.*

The following lemma, which extends Lemma 2.3 in [10], was proven in [15, Lemma 3.3]. It expresses a discrete Bochner-type identity.

Lemma 1. *Let ϕ, $\psi : S \to \mathbb{R}$ and let $\beta : S \times S \to \mathbb{R}$ be symmetric. Then*

$$\pi\Big[\sum_{\gamma,\delta\in G} R(\eta,\gamma,\delta)\beta(\eta,\delta\eta)\nabla_\delta\phi(\eta)\nabla_\gamma\psi(\eta)\Big] = \frac{1}{4}\pi\Big[\sum_{\gamma,\delta\in G} R(\eta,\gamma,\delta)\nabla_\gamma\big(\beta(\eta,\delta\eta)\nabla_\delta\phi(\eta)\big)\nabla_\delta\nabla_\gamma\psi(\eta)\Big].$$

The key estimate is contained in the following proposition that is an extension of Theorem 3.5 in [15] from the logarithmic to the power-function case.

Proposition 2. *Assume that there exists a function R satisfying Assumption 1 and define $\Gamma(\eta,\gamma,\delta) = c(\eta,\gamma)c(\eta,\delta) - R(\eta,\gamma,\delta)$ for $\eta \in S$ and γ, $\delta \in G$. Then, for any positive probability density ρ,*

$$\pi\big[\mathcal{L}\rho^{\alpha-1}\mathcal{L}\rho\big] + (\alpha-1)\pi\big[(\mathcal{L}\rho)^2\rho^{\alpha-2}\big] \geq \pi\Big[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\Big(\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho(\eta) + (\alpha-1)\nabla_\gamma\rho(\eta)\nabla_\delta\rho(\eta)\rho^{\alpha-2}(\eta)\Big)\Big]. \tag{8}$$

Remark 3. In the proof of Proposition 2 and later, we employ some properties of the function $\theta(s,t) = (s-t)/(s^{\alpha-1} - t^{\alpha-1})$ for $s \neq t$, which are collected in the Appendix. We introduce the following notation:

$$\widehat{\rho}(\eta,\delta\eta) = \theta(\rho(\eta),\rho(\delta\eta)) = \frac{\rho(\delta\eta)-\rho(\eta)}{\rho^{\alpha-1}(\delta\eta)-\rho^{\alpha-1}(\eta)} = \frac{\nabla_\delta\rho(\eta)}{\nabla_\delta\rho^{\alpha-1}(\eta)}, \tag{9}$$

$$\widehat{\rho}_1(\eta,\delta\eta) = \partial_1\theta(\rho(\eta),\rho(\delta\rho)) = -\frac{1}{\nabla_\delta\rho^{\alpha-1}(\eta)} + (\alpha-1)\frac{\nabla_\delta\rho(\eta)\rho^{\alpha-2}(\eta)}{(\nabla_\delta\rho^{\alpha-1}(\eta))^2}, \tag{10}$$

$$\widehat{\rho}_2(\eta,\delta\eta) = \partial_2\theta(\rho(\eta),\rho(\delta\rho)) = \widehat{\rho}_1(\delta\eta,\eta), \tag{11}$$

where $\partial_1\theta$ and $\partial_2\theta$ are the partial derivatives of θ with respect to the first and second variable, respectively. □

Proof of Proposition 2. The first term on the left-hand side of (8) can be written as follows, using the definitions of $\mathcal{L}$, $\widehat{\rho}$, and Γ:

$$\begin{aligned}\pi\big[\mathcal{L}\rho^{\alpha-1}\mathcal{L}\rho\big] &= \pi\Big[\sum_{\gamma,\delta\in G} c(\eta,\gamma)c(\eta,\delta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho(\eta)\Big] \\ &= \pi\Big[\sum_{\gamma,\delta\in G} c(\eta,\gamma)c(\eta,\delta)\widehat{\rho}(\eta,\delta\eta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\Big] \\ &= \pi\Big[\sum_{\gamma,\delta\in G} R(\eta,\gamma,\delta)\widehat{\rho}(\eta,\delta\eta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\Big]\end{aligned}$$

$$+\pi\bigg[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\widehat{\rho}(\eta,\delta\eta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\bigg].$$

By Lemma 1 with $\beta(\eta,\delta\eta)=\widehat{\rho}(\eta,\delta\rho)$, the first term on the right-hand side of the previous equation can be rewritten, leading to $\pi[\mathcal{L}\rho^{\alpha-1}\mathcal{L}\rho]=A_1+A_2$, where

$$A_1=\frac{1}{4}\pi\bigg[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\nabla_\gamma\big(\widehat{\rho}(\eta,\delta\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\big)\nabla_\delta\nabla_\gamma\rho^{\alpha-1}(\eta)\bigg],$$

$$A_2=\pi\bigg[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\widehat{\rho}(\eta,\delta\eta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\bigg].$$

Next, we reformulate the second term on the left-hand side of (8), using the definitions of $\mathcal{L}$, $\widehat{\rho}_1$, and Γ:

$$\begin{aligned}(\alpha-1)\pi\big[(\mathcal{L}\rho)^2\rho^{\alpha-2}\big]&=(\alpha-1)\pi\bigg[\sum_{\gamma,\delta\in G}c(\eta,\gamma)c(\eta,\delta)\nabla_\gamma\rho(\eta)\nabla_\delta\rho(\eta)\rho^{\alpha-2}(\eta)\bigg]\\
&=\pi\bigg[\sum_{\gamma,\delta\in G}c(\eta,\gamma)c(\eta,\delta)\widehat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)(\nabla_\delta\rho^{\alpha-1}(\eta))^2\bigg]\\
&\quad+\pi\bigg[\sum_{\gamma,\delta\in G}c(\eta,\gamma)c(\eta,\delta)\nabla_\delta\rho^{\alpha-1}(\eta)\nabla_\gamma\rho(\eta)\bigg]\\
&=\pi\bigg[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\widehat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)(\nabla_\delta\rho^{\alpha-1}(\eta))^2\bigg]\\
&\quad+\pi\bigg[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\widehat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)(\nabla_\delta\rho^{\alpha-1}(\eta))^2\bigg]\\
&\quad+\pi\bigg[\sum_{\gamma,\delta\in G}c(\eta,\gamma)c(\eta,\delta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho(\eta)\bigg]\\
&=:B_1+B_2+(A_1+A_2).\end{aligned}$$

Then the left-hand side of (8) is given by

$$\pi\big[\mathcal{L}\rho^{\alpha-1}\mathcal{L}\rho\big]+(\alpha-1)\pi\big[(\mathcal{L}\rho)^2\rho^{\alpha-2}\big]=(B_1+2A_1)+(B_2+2A_2),$$

and we will estimate B_1+2A_1 and B_2+2A_2 separately.

First, we treat B_2+2A_2. Inserting the definition of $\widehat{\rho}(\eta,\delta\eta)$ and rearranging the terms, we find that

$$\begin{aligned}B_2+2A_2&=\pi\bigg[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\widehat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)(\nabla_\delta\rho^{\alpha-1}(\eta))^2\bigg]\\
&\quad+2\pi\bigg[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\widehat{\rho}(\eta,\delta\eta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\bigg]\end{aligned}$$

$$= \pi\Big[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho(\eta)\Big]$$
$$+ (\alpha-1)\pi\Big[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\nabla_\gamma\rho(\eta)\nabla_\delta\rho(\eta)\rho^{\alpha-2}(\eta)\Big],$$

which is exactly the right-hand side of (8). Thus, it remains to prove that $B_1 + 2A_1 \geq 0$.

To this end, we reformulate B_1, employing Assumption 1 (i)-(ii) and identity (11):

$$B_1 = \pi\Big[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\widehat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta)(\nabla_\delta\rho^{\alpha-1}(\eta))^2\Big] \tag{12}$$
$$= \pi\Big[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\widehat{\rho}_1(\delta\eta,\eta)\nabla_\gamma\rho(\delta\eta)(\nabla_{\delta^{-1}}\rho^{\alpha-1}(\delta\eta))^2\Big]$$
$$= \pi\Big[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\widehat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta)(\nabla_\delta\rho^{\alpha-1}(\eta))^2\Big], \tag{13}$$

since $\nabla_{\delta^{-1}}\rho^{\alpha-1}(\delta\eta) = -\nabla_\delta\rho^{\alpha-1}(\eta)$. Averaging (12) and (13) gives

$$B_1 = \frac{1}{2}\pi\Big[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\Big(\widehat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta) + \widehat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta)\Big)(\nabla_\delta\rho^{\alpha-1}(\eta))^2\Big].$$

By Lemma 16 (i) (see Appendix) with $u = \rho(\gamma\eta)$, $v = \rho(\gamma\delta\eta)$, $s = \rho(\eta)$, and $t = \rho(\delta\eta)$, it follows that

$$\widehat{\rho}_1(\eta,\delta\eta)\nabla_\gamma\rho(\eta) + \widehat{\rho}_2(\eta,\delta\eta)\nabla_\gamma\rho(\delta\eta) \geq \nabla_\gamma\widehat{\rho}(\eta,\delta\eta),$$

and we infer from the definition of A_1 that

$$B_1 + 2A_1 \geq \frac{1}{2}\pi\Big[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\Big\{\nabla_\gamma\widehat{\rho}(\eta,\delta\eta)(\nabla_\delta\rho^{\alpha-1}(\eta))^2 + \nabla_\gamma\big(\widehat{\rho}(\eta,\delta\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\big)\nabla_\delta\nabla_\gamma\rho^{\alpha-1}(\eta)\Big\}\Big]. \tag{14}$$

The following identity has been used in the proof of Theorem 3.5 in [15]:

$$\nabla_\gamma\widehat{\rho}(\eta,\delta\eta)(\nabla_\delta\psi(\eta))^2 + \nabla_\gamma\big(\widehat{\rho}(\eta,\delta\eta)\nabla_\delta\psi(\eta)\big)\nabla_\delta\nabla_\gamma\psi(\eta)$$
$$= \widehat{\rho}(\gamma\eta,\gamma\delta\eta)(\nabla_\gamma\nabla_\delta\psi(\eta))^2 - \widehat{\rho}(\eta,\delta\eta)\nabla_\delta\psi(\gamma\eta)\nabla_\delta\psi(\eta) + \widehat{\rho}(\gamma\eta,\delta\gamma\eta)\nabla_\delta\psi(\gamma\eta)\nabla_\delta\psi(\eta). \tag{15}$$

It can be verified by elementary computations. Taking $\psi(\eta) = \rho^{\alpha-1}(\eta)$, the left-hand side of (15) equals the expression in the curly brackets of (14), and we conclude that

$$B_1 + 2A_1 \geq \frac{1}{2}\pi\Big[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\widehat{\rho}(\gamma\eta,\gamma\delta\eta)(\nabla_\gamma\nabla_\delta\rho^{\alpha-1}(\eta))^2\Big]$$
$$- \frac{1}{2}\pi\Big[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\widehat{\rho}(\eta,\delta\eta)\nabla_\delta\rho^{\alpha-1}(\gamma\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\Big]$$

$$+\frac{1}{2}\pi\bigg[\sum_{\gamma,\delta\in G}R(\eta,\gamma,\delta)\widehat{\rho}(\gamma\eta,\delta\gamma\eta)\nabla_\delta\rho^{\alpha-1}(\gamma\eta)\nabla_\delta\rho^{\alpha-1}(\eta)\bigg].$$

It follows from Assumption 1 (ii)-(iii) that the second and third term on the right-hand side cancel. The first term being nonnegative, we infer that $B_1+2A_1\geq 0$, which concludes the proof. □

The following corollary is a consequence of Proposition 2.

Corollary 4. *Suppose that there exists a constant $\lambda>0$ such that for all positive probability densities ρ,*

$$\begin{aligned}\pi\bigg[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\Big(\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho(\eta)+(\alpha-1)\nabla_\gamma\rho(\eta)\nabla_\delta\rho(\eta)\rho^{\alpha-2}(\eta)\Big)\bigg]\\ \geq\frac{\lambda}{2}\pi\bigg[\sum_{\gamma\in G}c(\eta,\gamma)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\gamma\rho(\eta)\bigg].\end{aligned}\tag{16}$$

Then the Beckner inequality (5)*, the decay of the Dirichlet form*

$$\mathcal{E}((e^{t\mathcal{L}}\rho)^{\alpha-1},e^{t\mathcal{L}}\rho)\leq e^{-\lambda t}\mathcal{E}(\rho^{\alpha-1},\rho),\quad t>0,\tag{17}$$

and the decay of the entropy (6) *hold for all positive probability densities ρ.*

Proof. By Proposition 2 and representation (7) of the Dirichlet form, it follows from (16) that

$$\pi[\mathcal{L}\rho^{\alpha-1}\mathcal{L}\rho]+(\alpha-1)\pi[(\mathcal{L}\rho)^2\rho^{\alpha-2}]\geq\lambda\mathcal{E}(\rho^{\alpha-1},\rho).$$

Taking into account (1), this inequality is equivalent to

$$\frac{d^2}{dt^2}\mathrm{Ent}_\pi(\rho_t)\geq-\lambda\frac{d}{dt}\mathrm{Ent}_\pi(\rho_t).\tag{18}$$

Using Gronwall's lemma, we infer that $0=\lim_{t\to\infty}(-\partial_t\mathrm{Ent}_\pi(\rho_t))$. Furthermore, as π is an invariant measure, $\rho_t\to 1$ and $\mathrm{Ent}_\pi(\rho_t)\to 0$ as $t\to\infty$. Therefore, integrating (18) over $(0,\infty)$, we conclude that

$$-\alpha\mathcal{E}(\rho_0^{\alpha-1},\rho_0)=\frac{d}{dt}\mathrm{Ent}_\pi(\rho_0)\leq-\lambda\mathrm{Ent}_\pi(\rho_0),$$

and this is exactly the Beckner inequality (5). □

3. Examples

In this section, we consider some stochastic processes analyzed in [10, 15] but for logarithmic entropies only. Our notation follows that of [10].

3.1. **Birth-death processes.** We investigate birth-death processes on $\mathbb{N} = \{0, 1, 2, \ldots\}$ with generator

$$\mathcal{L}f(n) = a(n)\nabla_+ f(n) + b(n)\nabla_- f(n), \quad n \in \mathbb{N},$$

where a and b are nonnegative functions on $\mathbb{N}$ satisfying $b(0) = 0$. The function a represents the rate of birth, the function b the rate of death. The set of allowed moves is given by $G = \{+, -\}$, where $+(n) = n + 1$ for $n \in \mathbb{N}$ and $-(n) = n - 1$ for $n \geq 1$, $-(0) = 0$. In particular, $\nabla_\pm f(n) = f(n \pm 1) - f(n)$. According to the notation of Section 2, $c(n, +) = a(n)$ and $c(n, -) = b(n)$.

We suppose that this Markov chain is irreducible and reversible, i.e., there exists a probability measure π on $\mathbb{N}$ satisfying the detailed-balance condition

$$a(n)\pi(n) = b(n + 1)\pi(n + 1), \quad n \in \mathbb{N}. \tag{19}$$

The following theorem is a consequence of Corollary 4, applied to birth-death processes.

Theorem 5. *Let $\lambda > 0$. Assume that a is nonincreasing, b is nondecreasing, and*

$$a(n) - a(n + 1) + b(n + 1) - b(n) + \Theta\big(a(n) - a(n + 1), b(n + 1) - b(n)\big) \geq \lambda \tag{20}$$

for all $n \in \mathbb{N}$, where

$$\Theta(A, B) = (\alpha - 1) \inf_{s,t>0} \theta(s, t)(As^{\alpha-2} + Bt^{\alpha-2}), \quad A, B \geq 0,$$

and $\theta(s, t) = (s - t)/(s^{\alpha-1} - t^{\alpha-1})$ for $s \neq t$. Then the Beckner inequality (5) *and the decay estimates* (6) *and* (17) *hold with constant λ.*

The mapping Θ generalizes the function in [15, Section 4.1] to $1 < \alpha \leq 2$. Lemma 17 in the Appendix shows that $\Theta(A, B) \geq (\alpha - 1)(A + B)$. Moreover, $\Theta(A, B) = A + B$ if $\alpha = 2$. Figure 1 illustrates the "sharpness" of the inequality $\Theta(A, B) \geq (\alpha - 1)(A + B)$ for α close to one.

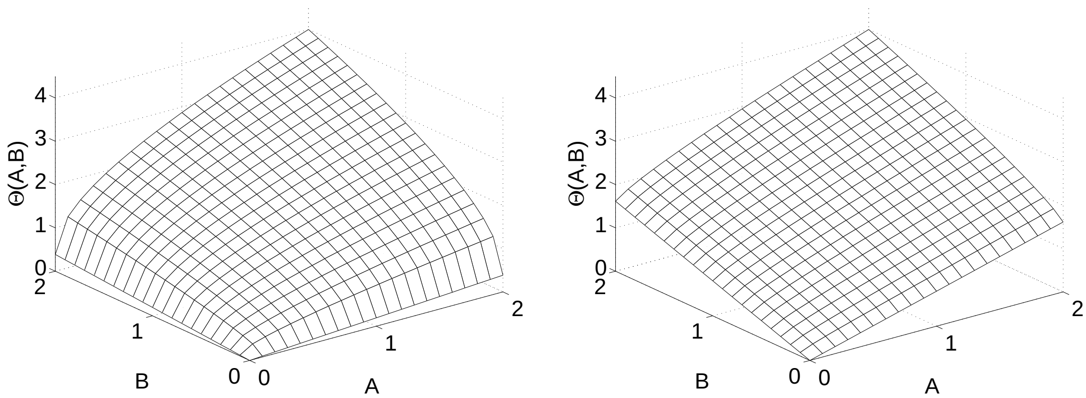

FIGURE 1. Illustration of $\Theta(A, B)$ for $\alpha = 1.01$ (left) and $\alpha = 1.8$ (right).

Remark 6. Estimates for Poincaré inequalities for Markov chains are given in, e.g., [11, 12, 22]. The same criterion as in (20) was obtained in [23, Theorem 5.1] and [15, Theorem 4.1] for the logarithmic entropy ($\alpha \to 1$). From Lemma 17 we conclude that the Beckner constant can be estimated by $\lambda \geq \alpha(a(n)-a(n+1)+b(n-1)-b(n))$. There exist sufficient and necessary conditions on π and $a(n)$ such that an interpolation between the Poincaré and log-Sobolev inequality holds, but without estimates on the constant [27, Theorem 6.2.4]. □

Proof. We define as in [10, Section 3]

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

This function satisfies Assumption 1. In particular, (ii) follows from the detailed-balance condition (19). As before, we set $\Gamma(n,\gamma,\delta) = c(n,\gamma)c(n,\delta) - R(n,\gamma,\delta)$ for γ, $\delta \in G$. According to Corollary 4, we only need to verify (16). The left-hand side equals

$$\begin{aligned} \pi\Big[&\sum_{\gamma,\delta\in G}\Gamma(n,\gamma,\delta)\Big(\nabla_\gamma\rho^{\alpha-1}(n)\nabla_\delta\rho(n) + (\alpha-1)\nabla_\gamma\rho(n)\nabla_\delta\rho(n)\rho^{\alpha-2}(n)\Big)\Big] \\ &= \pi\Big[a(n)(a(n)-a(n+1))\Big(\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n) + (\alpha-1)(\nabla_+\rho(n))^2\rho^{\alpha-2}(n)\Big)\Big] \\ &\quad + \pi\Big[b(n)(b(n)-b(n-1))\Big(\nabla_-\rho^{\alpha-1}(n)\nabla_-\rho(n) + (\alpha-1)(\nabla_-\rho(n))^2\rho^{\alpha-2}(n)\Big)\Big], \end{aligned}$$

since the sum over all γ, $\delta \in G$ consists of four terms $(+,+)$, $(-,-)$, $(+,-)$, and $(-,+)$, and because of $\Gamma(n,+,-) = \Gamma(n,-,+) = 0$, only two terms do not vanish. Now, we perform the change $n \mapsto n+1$ in the second term and replace $\pi(n+1)b(n+1)$ by $\pi(n)a(n)$, according to the detailed-balance condition (19). Observing that $b(0) = 0$ and $\nabla_-\rho(n+1) = -\nabla_+\rho(n)$, this leads to

$$\begin{aligned} \pi\Big[&\sum_{\gamma,\delta\in G}\Gamma(n,\gamma,\delta)\Big(\nabla_\gamma\rho^{\alpha-1}(n)\nabla_\delta\rho(n) + (\alpha-1)\nabla_\gamma\rho(n)\nabla_\delta\rho(n)\rho^{\alpha-2}(n)\Big)\Big] \\ &= \pi\Big[a(n)(a(n)-a(n+1))\Big(\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n) + (\alpha-1)(\nabla_+\rho(n))^2\rho^{\alpha-2}(n)\Big)\Big] \\ &\quad + \pi\Big[a(n)(b(n+1)-b(n))\Big(\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n) + (\alpha-1)(\nabla_+\rho(n))^2\rho^{\alpha-2}(n+1)\Big)\Big] \\ &= \pi\Big[a(n)\Big(a(n)-a(n+1)+b(n+1)-b(n)\Big)\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n)\Big] \\ &\quad + (\alpha-1)\pi\Big[a(n)\Big((a(n)-a(n+1))\rho^{\alpha-2}(n) + (b(n+1)-b(n))\rho^{\alpha-2}(n+1)\Big) \\ &\quad \times \widehat{\rho}(n,n+1)\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n)\Big] \\ &\geq \lambda\pi\big[a(n)\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n)\big], \end{aligned}$$

where in the last step we employed (20). Using again the detailed-balance condition (19) and the identity $\nabla_-\rho(n) = -\nabla_+\rho(n-1)$, the right-hand side of (16) becomes

$$\begin{aligned}
&\frac{\lambda}{2}\pi\Big[\sum_{\gamma\in G} c(n,\gamma)\nabla_\gamma\rho^{\alpha-1}(n)\nabla_\gamma\rho(n)\Big] \\
&\quad = \frac{\lambda}{2}\pi\Big[a(n)\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n)\Big] + \frac{\lambda}{2}\pi\Big[b(n)\nabla_-\rho^{\alpha-1}(n)\nabla_-\rho(n)\Big] \\
&\quad = \frac{\lambda}{2}\pi\Big[a(n)\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n)\Big] + \frac{\lambda}{2}\pi\Big[a(n)\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n)\Big] \\
&\quad = \lambda\pi\big[a(n)\nabla_+\rho^{\alpha-1}(n)\nabla_+\rho(n)\big].
\end{aligned}$$

Combining the above computations, inequality (16) follows. □

3.2. **Zero-range processes.** A zero-range process describes a stochastically interacting particle system that may exhibit phase separation; see, e.g., [24]. The system consists of finitely many particles moving in a finite number of sites $\{1,2,\ldots,L\}$. We adopt the notation of [10]. Let $\eta_x \in \mathbb{N}$ denote the number of particles at $x \in \{1,2,\ldots,L\}$. Then the state space is $S = \mathbb{N}^L$. The configuration is changed by moving a particle from an (occupied) site x to another site y. Correspondingly, the set G of allowed moves is given by self-mappings of S which are of the form $\eta \mapsto \eta^{xy}$, where x, $y \in \{1,2,\ldots,L\}$, $x \neq y$, and

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

We denote by xy the mapping $\eta \mapsto \eta^{xy}$ (such that $xy(\eta) = \eta^{xy}$) and set $\nabla_{xy}f(\eta) = f(\eta^{xy}) - f(\eta)$ for $\eta \in S$.

The jump rates are functions $c_x : \mathbb{N} \to \mathbb{R}_+$ for $x \in \{1,2,\ldots,L\}$ satisyfing $c_x(0) = 0$ and $c_x(n) > 0$ for $n > 0$. They describe the rate at which a particle is moved from site x to site y, with randomly chosen y, with uniform probability on $\{1,2,\ldots,L\}$. Then the rate $c(\eta, xy)$ for moving a particle from x to y is $c_x(\eta_x)/L$, and the generator of the Markov chain becomes

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

where the sum extends to all x, $y \in \{1,2,\ldots,L\}$. The number of particles $N = \sum_{1\le x\le L}\eta_x$ is conserved. We define the probability measure π_N on configurations with N particles by

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

where $Z_N > 0$ the (finite) normalization constant. Since

$$\pi[c_x(\eta_x)g(\eta)] = \pi[c_y(\eta_y)g(\eta^{yx})] \tag{21}$$

holds for all functions $g : S \to \mathbb{R}$, the Markov chain is reversible with respect to π_N. In the following, we fix the number of particles N and omit the subscript N.

Theorem 7. *Assume that there exist constants* $0 \le \delta < 2^{2-\alpha}c$ *such that*

$$c \le c_x(n+1) - c_x(n) \le c + \delta \quad \text{for } x \in \{1,2,\ldots,G\},\ n \ge 0. \tag{22}$$

Then the Beckner inequality (5) *and the decay estimates* (6) *and* (17) *hold with* $\lambda = \alpha c - (3 + 2^{\alpha-2} - \alpha)\delta$.

Remark 8. In the case of constant rates, the spectral gap is of the order of $L^2/(L^2+N^2)$ [26]. Note that our bound $\lambda = 2(c-\delta)$ for $\alpha = 2$ does not depend on either L or N. It was shown in [8] that the lower bound in (22) is sufficient to derive the spectral-gap estimate $\lambda \ge c$. In the homogeneous case $\delta = 0$, we have even $\lambda = 2c$. As pointed out in [10], a condition on the growth of the rates is necessary for the modified logarithmic Sobolev inequality. Our bound $\lambda = c - 5\delta/2$ for $\alpha \to 1$ is the same as in [15, Theorem 4.3]. □

Proof. We define as in [10, Section 4] the function

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

which satisfies Assumption 1. It follows that $\Gamma(\eta, xy, uv) = 0$ if $x \ne u$ and

$$\Gamma(\eta, xy, uv) = L^{-2}c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x - 1)\big) \quad \text{if } x = u,$$

and the left-hand side of (16) can be written as

$$\begin{aligned} &\pi\Bigg[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\Big(\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho(\eta) + (\alpha-1)\nabla_\gamma\rho(\eta)\nabla_\delta\rho(\eta)\rho^{\alpha-2}(\eta)\Big)\Bigg] \\ &\quad = \frac{1}{L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x-1)\big)\nabla_{xv}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta)\Bigg] \\ &\qquad + \frac{\alpha-1}{L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x-1)\big)\nabla_{xy}\rho(\eta)\nabla_{xv}\rho(\eta)\rho^{\alpha-2}(\eta)\Bigg] \\ &\quad = C_1 + C_2. \end{aligned}$$

For future reference, we denote the right-hand side of (16) (without the constant λ) by

$$A = \frac{1}{2L}\pi\Bigg[\sum_{x,y} c_x(\eta_x)\nabla_{xy}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta)\Bigg].$$

The estimate of the term C_1 is similar to $\widetilde{\mathcal{B}}_1(\rho,\psi)$ in the proof of Theorem 4.3 in [15] (take $\psi(\eta) = \rho^{\alpha-1}(\eta)$). First, we interchange y and v and then use $\nabla_{xv}\rho^{\alpha-1}(\eta) = \nabla_{xy}\rho^{\alpha-1}(\eta) + \nabla_{yv}\rho^{\alpha-1}(\eta^{xy})$ as well as the lower bound $c_x(\eta_x) - c_x(\eta_x - 1) \ge c$:

$$\begin{aligned} C_1 &= \frac{1}{L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x-1)\big)\big(\nabla_{xy}\rho^{\alpha-1}(\eta) + \nabla_{yv}\rho^{\alpha-1}(\eta^{xy})\big)\nabla_{xy}\rho(\eta)\Bigg] \\ &\ge 2cA + \frac{1}{L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x-1)\big)\nabla_{yv}\rho^{\alpha-1}(\eta^{xy})\nabla_{xy}\rho(\eta)\Bigg]. \end{aligned} \tag{23}$$

Note that the term involving $\nabla_{xy}\rho^{\alpha-1}(\eta)$ does not depend on v, so the sum over x, y, v equals L times the sum over x, y. Employing the reversibility condition (21) and exchanging x and y in the second term yields

$$\begin{aligned} C_1 &\geq 2cA + \frac{1}{L^2}\pi\Bigg[\sum_{x,y,v} c_y(\eta_y)\big(c_x(\eta_x^{yx}) - c_x(\eta_x^{yx}-1)\big)\nabla_{yv}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta^{yx})\Bigg] \\ &= 2cA - \frac{1}{L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)\big(c_y(\eta_y+1) - c_y(\eta_y)\big)\nabla_{xv}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta)\Bigg]. \end{aligned} \tag{24}$$

We average (23) and (24) and employ again the identity $\nabla_{xy}\rho^{\alpha-1}(\eta) + \nabla_{yv}\rho^{\alpha-1}(\eta^{xy}) = \nabla_{xv}\rho^{\alpha-1}(\eta)$:

$$\begin{aligned} C_1 \geq cA + \frac{1}{2L^2}\pi\Bigg[&\sum_{x,y,v} c_x(\eta_x)\Big((c_x(\eta_x) - c_x(\eta_x-1)) - (c_y(\eta_y+1) - c_y(\eta_y))\Big) \\ &\times \nabla_{xv}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta)\Bigg]. \end{aligned}$$

Setting $C_0 := (c_x(\eta_x) - c_x(\eta_x-1)) - (c_y(\eta_y+1) - c_y(\eta_y))$, the bounds (22) imply that $|C_0| \leq \delta$. Hence, by Young's inequality,

$$\begin{aligned} C_0\nabla_{xv}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta) &= C_0\widehat{\rho}(\eta,\eta^{xy})\nabla_{xv}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho^{\alpha-1}(\eta) \\ &\geq -\frac{1}{2}|C_0|\widehat{\rho}(\eta,\eta^{xy})\Big((\nabla_{xy}\rho^{\alpha-1}(\eta))^2 + (\nabla_{xv}\rho^{\alpha-1}(\eta))^2\Big) \\ &\geq -\frac{\delta}{2}\Big(\nabla_{xy}\rho(\eta)\nabla_{xy}\rho^{\alpha-1}(\eta) + (\nabla_{xv}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta,\eta^{xy})\Big). \end{aligned}$$

This yields

$$C_1 \geq \left(c - \frac{\delta}{2}\right)A - \frac{\delta}{4L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta,\eta^{xy})\Bigg]. \tag{25}$$

Next, we rewrite $B = (C_2 - C_1)/2$. By definition (10) of $\widehat{\rho}_1$ and the reversibility condition (21),

$$\begin{aligned} B &= \frac{1}{2L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)\big(c_x(\eta_x) - c_x(\eta_x-1)\big)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\widehat{\rho}_1(\eta,\eta^{xy})\nabla_{xv}\rho(\eta)\Bigg] \\ &= \frac{1}{2L^2}\pi\Bigg[\sum_{x,y,v} c_y(\eta_y)\big(c_x(\eta_x+1) - c_x(\eta_x)\big)(\nabla_{xy}\rho^{\alpha-1}(\eta^{yx}))^2\widehat{\rho}_1(\eta^{yx},\eta)\nabla_{xv}\rho(\eta^{yx})\Bigg] \\ &= \frac{1}{2L^2}\pi\Bigg[\sum_{x,y,v} c_x(\eta_x)\big(c_y(\eta_y+1) - c_y(\eta_y)\big)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\widehat{\rho}_2(\eta,\eta^{xy})\big(\rho(\eta^{xv}) - \rho(\eta^{xy})\big)\Bigg]. \end{aligned}$$

In the last step, we interchanged x and y and used the identity $\widehat{\rho}_1(\eta^{xy},\eta) = \widehat{\rho}_2(\eta,\eta^{xy})$. Averaging the expressions for B involving $\widehat{\rho}_1$ and $\widehat{\rho}_2$ gives

$$\begin{aligned} B &= \frac{1}{4L^2}\pi\bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\rho(\eta^{xv}) \\ &\qquad \times \Big(\big(c_x(\eta_x)-c_x(\eta_x-1)\big)\widehat{\rho}_1(\eta,\eta^{xy}) + \big(c_y(\eta_y+1)-c_y(\eta_y)\big)\widehat{\rho}_2(\eta,\eta^{xy})\Big)\bigg] \\ &\quad - \frac{1}{4L^2}\pi\bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2 \\ &\qquad \times \Big(\big(c_x(\eta_x)-c_x(\eta_x-1)\big)\widehat{\rho}_1(\eta,\eta^{xy})\rho(\eta) + \big(c_y(\eta_y+1)-c_y(\eta_y)\big)\widehat{\rho}_2(\eta,\eta^{xy})\rho(\eta^{xy})\Big)\bigg] \\ &= B_1 + B_2. \end{aligned}$$

The term B_1 is estimated by using condition (22) (note that $\widehat{\rho}_1$, $\widehat{\rho}_2 \geq 0$ since θ is nondecreasing in both variables) and then employing the assumption $c \geq 2^{\alpha-2}\delta$ and interchanging y and v:

$$\begin{aligned} B_1 &\geq \frac{c}{4L^2}\pi\bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\rho(\eta^{xv})\big(\widehat{\rho}_1(\eta,\eta^{xy}) + \widehat{\rho}_2(\eta,\eta^{xy})\big)\bigg] \\ &\geq \frac{2^{\alpha-2}\delta}{4L^2}\pi\bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\rho(\eta^{xy})\Big(\widehat{\rho}_1(\eta,\eta^{xv}) + \widehat{\rho}_2(\eta,\eta^{xv})\Big)\bigg] \\ &= B_3. \end{aligned}$$

We employ condition (22) once more and Lemma 16 (ii) (see the Appendix) to estimate B_2:

$$\begin{aligned} B_2 &\geq -\frac{c+\delta}{4L^2}\pi\bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\Big(\widehat{\rho}_1(\eta,\eta^{xy})\rho(\eta) + \widehat{\rho}_2(\eta,\eta^{xy})\rho(\eta^{xy})\Big)\bigg] \\ &= -\frac{c+\delta}{4L^2}(2-\alpha)\pi\bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta,\eta^{xy})\bigg] = -\frac{1}{2}(2-\alpha)(c+\delta)A. \end{aligned}$$

Consequently,

$$B \geq -\frac{1}{2}(2-\alpha)(c+\delta)A + B_3. \tag{26}$$

We add (25) and (26):

$$\begin{aligned} C_1 + B &\geq \left(c - \frac{\delta}{2} - \frac{1}{2}(2-\alpha)(c+\delta)\right)A + B_4, \quad \text{where} \\ B_4 &= B_3 - \frac{\delta}{4L^2}\pi\bigg[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta,\eta^{xy})\bigg]. \end{aligned} \tag{27}$$

We wish to estimate B_4 from below by a multiple of A. To this end, we employ the reversibility and interchange x and v in the second term in B_4:

$$\pi\Big[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta,\eta^{xy})\Big] = \pi\Big[\sum_{x,y,v} c_v(\eta_v)(\nabla_{xv}\rho^{\alpha-1}(\eta^{vx}))^2\widehat{\rho}(\eta^{vx},\eta^{vy})\Big]$$
$$= \pi\Big[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta^{xv},\eta^{xy})\Big].$$

Then, averaging those two expressions for B_4 that involve $\widehat{\rho}(\eta,\eta^{xy})$ and $\widehat{\rho}(\eta^{xv},\eta^{xy})$,

$$B_4 = \frac{\delta}{8L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2(2^{\alpha-1}\rho(\eta^{xy}))\Big(\widehat{\rho}_1(\eta,\eta^{xv})+\widehat{\rho}_2(\eta,\eta^{xv})\Big)\Big]$$
$$-\frac{\delta}{8L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\Big(\widehat{\rho}(\eta,\eta^{xy})+\widehat{\rho}(\eta^{xv},\eta^{xy})\Big)\Big].$$

We employ Lemma 16 (iii) in the form

$$2^{\alpha-1}\rho(\eta^{xy}))\big(\widehat{\rho}_1(\eta,\eta^{xv})+\widehat{\rho}_2(\eta,\eta^{xv})\big)-\big(\widehat{\rho}(\eta,\eta^{xy})+\widehat{\rho}(\eta^{xv},\eta^{xy})\big)\geq -2^{\alpha-1}\widehat{\rho}(\eta,\eta^{xv}),$$

which leads to

$$B_4 \geq -\frac{2^{\alpha-1}\delta}{8L^2}\pi\Big[\sum_{x,y,v} c_x(\eta_x)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta,\eta^{xv})\Big] = -\frac{2^{\alpha-1}\delta}{4}A.$$

Hence, we infer from (27) that

$$C_1 + B \geq \left(c - \frac{\delta}{2} - \frac{1}{2}(2-\alpha)(c+\delta) - \frac{\delta}{4}2^{\alpha-1}\right)A.$$

Finally, by definition of B,

$$C_1 + C_2 = 2(C_1 + B) \geq \big(2c - \delta - (2-\alpha)(c+\delta) - 2^{\alpha-2}\delta\big)A = \lambda A.$$

This shows (16), and an application of Corollary 4 finishes the proof. □

3.3. **Bernoulli-Laplace models.** We consider again a system of particles moving in a finite set of sizes $\{1,2,\ldots,L\}$ but in contrast to the previous subsection, we assume that at most one particle per site is allowed, i.e. $S=\{0,1\}^L$. The set of allowed moves is $G=\{xy : x,y\in\{1,2,\ldots,L\},\ x\neq y\}$, and the moves are of the form $xy:\eta\mapsto\eta^{xy}$ for $\eta\in S$, where $\eta^{xy}=\eta$ if $\eta_x(1-\eta_y)=0$ and otherwise,

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

We associate to each site x a Poisson clock of constant intensity $\lambda_x>0$. When the clock of site x rings, we choose randomly a site y. If $\eta_x=1$ and $\eta_y=0$ (i.e. if $\eta_x(1-\eta_y)=1$), the particle at x moves to y; otherwise (i.e. if $\eta_x(1-\eta_y)=0$), nothing happens. Therefore, the transition rates are given by $c(\eta,xy)=(\lambda_x/L)\eta_x(1-\eta_y)$, and the generator reads as

$$\mathcal{L}f(\eta)=\frac{1}{L}\sum_{xy\in G}\lambda_x\eta_x(1-\eta_y)\nabla_{xy}f(\eta),$$

where, as in the previous subsection, $\nabla_{xy}f(\eta)=f(\eta^{xy})-f(\eta)$.

Let $N\le L$ be the number of particles in the system. There exists a unique stationary distribution π_N, which is given by [10, Section 5]

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

where $Z_{L,N}>0$ is a normalization constant. In the following, we write π instead of π_N, as the number of particles is fixed. Reversibility holds for π, and it reads as

$$\pi\bigg[\sum_{xy\in G}c(\eta,xy)F(\eta,xy)\bigg]=\pi\bigg[\sum_{xy\in G}c(\eta,xy)F(\eta^{xy},yx)\bigg] \tag{28}$$

for arbitrary functions $F:S\times G\to\mathbb{R}$.

Theorem 9. *Assume that there exist constants $0\le\delta\le 2^{2-\alpha}c$ such that*

$$c\le\lambda_x\le c+\delta\quad\textit{for } x\in\{1,2,\ldots,L\}. \tag{29}$$

Then the Beckner inequality (5) *and the decay estimates* (6) *and* (17) *hold with* $\lambda=\alpha c-(\frac{5}{2}+2^{\alpha-3}-\alpha)\delta$.

Remark 10. For the modified log-Sobolev inequality, the bound in [10] reads as $\lambda=c-\delta$, and the bound in [15] equals $\lambda=c-7\delta/4$ (for $\delta<4c/7$). Our result coincides with that in [15] for $\alpha\to1$. In [18], the bound $1\le\lambda\le2$ was proved in case $c=1$, $\delta=0$. Further bounds, depending on L and N, were collected in [6, Examples 3.11].

Concerning the Beckner inequality, Bobkov and Tetali [6, Section 4] derived for the homogeneous case $c=L/(N(L-N))$ and $\delta=0$ the constant $\lambda\ge\alpha(L+2)/(2N(L-N)$. Our constant $\lambda=(\alpha L-2\alpha+4)/(N(L-N))$ (see the proof below) is larger for $L>2$ and all $1<\alpha\le2$. □

Proof. We need to verify the condition in Corollary 4. As in [10], we choose

$$R(\eta,xy,uv)=L^{-2}\lambda_x\lambda_u\eta_x(1-\eta_y)\eta_u(1-\eta_v)\quad\text{for } |\{x,y,u,v\}|=4$$

and $R(\eta,xy,uv)=0$ otherwise. The notation $|\{x,y,u,v\}|=4$ means that the four variables are pairwise different. Then $\Gamma(\eta,xy,uv)=0$ if $|\{x,y,u,v\}|=4$ and

$$\Gamma(\eta,xy,uv)=L^{-2}\lambda_x\lambda_u\eta_x(1-\eta_y)\eta_u(1-\eta_v)$$

otherwise. The sum of $\Gamma(\eta,\gamma,\delta)$ over γ, $\delta\in G$ in the left-hand side of (16) vanishes if (x,y,u,v) are pairwise different. Therefore, the sum consists of three terms: $(\gamma,\delta)=$

(xy, xy), $(\gamma,\delta) = (xy, uy)$, and $(\gamma,\delta) = (xy, xv)$, and it follows that

$$\begin{aligned}
&\pi\Big[\sum_{\gamma,\delta\in G}\Gamma(\eta,\gamma,\delta)\Big(\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\delta\rho(\eta) + (\alpha-1)\nabla_\gamma\rho(\eta)\nabla_\delta\rho(\eta)\rho^{\alpha-2}(\eta)\Big)\Big] \\
&\quad = \frac{1}{L^2}\pi\Big[\sum_{x,y}\lambda_x^2\nabla_{xy}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta) + \sum_{|\{x,y,u\}|=3}\lambda_x\lambda_u\nabla_{xy}\rho^{\alpha-1}(\eta)\nabla_{uy}\rho(\eta) \\
&\qquad + \sum_{|\{x,y,v\}|=3}\lambda_x^2\nabla_{xy}\rho^{\alpha-1}(\eta)\nabla_{xv}\rho(\eta)\Big] + \frac{\alpha-1}{L^2}\pi\Big[\sum_{x,y}\lambda_x^2\nabla_{xy}\rho(\eta)\nabla_{xy}\rho(\eta)\rho^{\alpha-2}(\eta) \\
&\qquad + \sum_{|\{x,y,u\}|=3}\lambda_x\lambda_u\nabla_{xy}\rho(\eta)\nabla_{uy}\rho(\eta)\rho^{\alpha-2}(\eta) + \sum_{|\{x,y,v\}|=3}\lambda_x^2\nabla_{xy}\rho(\eta)\nabla_{xv}\rho(\eta)\rho^{\alpha-2}(\eta)\Big] \\
&\quad = C_1 + C_2.
\end{aligned}$$

Observe that the right-hand side of (16) (without the constant λ) reads as

$$A = \frac{1}{2}\pi\Big[\sum_{\gamma\in G}c(\eta,\gamma)\nabla_\gamma\rho^{\alpha-1}(\eta)\nabla_\gamma\rho(\eta)\Big] = \frac{1}{2L}\pi\Big[\sum_{xy\in G}\lambda_x\nabla_{xy}\rho^{\alpha-1}(\eta)\nabla_{xy}\rho(\eta)\Big], \tag{30}$$

since $\nabla_{xy}\rho(\eta) = 0$ whenever $\eta_x(1-\eta_y) = 0$, so the factor $\eta_x(1-\eta_y)$ can be omitted.

As in the previous subsection, we estimate $B = (C_2 - C_1)/2$, recalling definition (10) of $\widehat{\rho}_1$:

$$\begin{aligned}
B &= \frac{1}{2L^2}\pi\Big[\sum_{x,y}\lambda_x^2(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\widehat{\rho}_1(\eta,\eta^{xy})\nabla_{xy}\rho(\eta)\Big] \\
&\quad + \frac{1}{2L^2}\pi\Big[\sum_{|\{x,y,u\}|=3}\lambda_x\lambda_u(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\widehat{\rho}_1(\eta,\eta^{xy})\nabla_{uy}\rho(\eta)\Big] \\
&\quad + \frac{1}{2L^2}\pi\Big[\sum_{|\{x,y,v\}|=3}\lambda_x^2(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\widehat{\rho}_1(\eta,\eta^{xy})\nabla_{xv}\rho(\eta)\Big] \\
&= B_1 + B_2 + B_3.
\end{aligned} \tag{31}$$

The estimations of B_1, B_2, and B_3 are the same as in the proof of Theorem 4.6 in [15] after taking $\psi(\eta) = \rho^{\alpha-1}(\eta)$ in $\widetilde{\mathcal{B}}_2(\rho,\psi)$. The key point is the use of Lemma 16 (iv). In constrast to [15], the factor $2-\alpha$ appears. Therefore, following [15] and taking into account (30), we conclude that

$$\begin{aligned}
B_1 &\geq -\frac{\delta}{2L}(2-\alpha)A, \\
B_2 &\geq -\frac{1}{2L}(N-1)(c+\delta)(2-\alpha)A, \\
B_3 &\geq \frac{c}{4L^2}\pi\Big[\sum_{|\{x,y,v\}|=3}\lambda_x\eta_x(1-\eta_y)(1-\eta_v)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\rho(\eta^{xv})
\end{aligned}$$

$$\times \Big(\widehat{\rho}_1(\eta,\eta^{xy}) + \widehat{\rho}_2(\eta,\eta^{xy})\Big)\Big] - \frac{1}{2L}(L-N-1)(c+\delta)(2-\alpha)A. \tag{32}$$

Since we assumed that $\delta \le 2^{2-\alpha}c$, we can estimate the factor in the first term of B_3 by $c/(4L^2) \ge 2^{\alpha-4}\delta/L^2$.

Next, we estimate C_1. This expression consists of three terms. We interchange x and u in the second term and y and v in the third term. Then $C_1 = B_4 + B_5 + B_6$, where

$$\begin{aligned}
B_4 &= \frac{1}{L^2}\pi\bigg[\sum_{x,y}\lambda_x^2\nabla_{xy}\rho(\eta)\nabla_{xy}\rho^{\alpha-1}(\eta)\bigg],\\
B_5 &= \frac{1}{L^2}\pi\bigg[\sum_{|\{x,y,u\}|=3}\lambda_x\lambda_u\nabla_{xy}\rho(\eta)\nabla_{uy}\rho^{\alpha-1}(\eta)\bigg],\\
B_6 &= \frac{1}{L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x^2\nabla_{xy}\rho(\eta)\nabla_{xv}\rho^{\alpha-1}(\eta)\bigg].
\end{aligned}$$

By condition (29), $B_4 \ge (2c/L)A$. The term B_6 is estimated by employing the reversibility (28), averaging, and using (29), similar to the estimate of J_6 in the proof of Theorem 4.6 in [15]. The result is

$$\begin{aligned}
B_6 &\ge \frac{1}{2L}(L-N-1)(2c-\delta)A - B_7, \quad \text{where}\\
B_7 &= \frac{\delta}{4L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x\eta_x(1-\eta_y)(1-\eta_v)(\nabla_{xv}\rho^{\alpha-1}(\eta))^2\widehat{\rho}(\eta,\eta^{xy})\bigg].
\end{aligned} \tag{33}$$

Similarly, replacing $\psi(\eta)$ by $\rho^{\alpha-1}(\eta)$ in J_5 in the proof of Theorem 4.6 in [15], we have $B_5 \ge (c/L)(N-1)A$.

It remains to rewrite B_7. For this, we employ the reversibility, average the original and the resulting expressions, and interchange y and v. This yields (see the computation of J_7 in [15])

$$B_7 = \frac{\delta}{8L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x\eta_x(1-\eta_y)(1-\eta_v)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\Big(\widehat{\rho}(\eta^{xv},\eta^{xy}) + \widehat{\rho}(\eta,\eta^{xv})\Big)\bigg].$$

Combining estimate (32) for B_3 and (33), together with the above estimate for B_7 and applying Lemma 16 (iii), we infer that

$$\begin{aligned}
B_3 + B_6 &\ge \frac{1}{2L}(L-N-1)\big(\alpha c - (3-\alpha)\delta\big)A\\
&\quad + \frac{\delta}{8L^2}\pi\bigg[\sum_{|\{x,y,v\}|=3}\lambda_x\eta_x(1-\eta_y)(1-\eta_v)(\nabla_{xy}\rho^{\alpha-1}(\eta))^2\\
&\quad \times\Big(2^{\alpha-1}\rho(\eta^{xv})\big(\widehat{\rho}_1(\eta,\eta^{xy}) + \widehat{\rho}_2(\eta,\eta^{xy})\big) - \big(\widehat{\rho}(\eta^{xv},\eta^{xy}) + \widehat{\rho}(\eta^{xv},\eta)\big)\Big)\bigg]
\end{aligned}$$

$$\geq \frac{1}{4L}(L-N-1)\big(2\alpha c - 2(3-\alpha)\delta - 2^{\alpha-1}\delta\big)A.$$

It remains to summarize the estimates:

$$\begin{aligned} C_1 + C_2 &= 2B + 2C_1 = 2(B_1+B_2) + 2(B_4+B_5) + 2(B_3+B_6) \\ &\geq -\frac{(2-\alpha)}{L}\big(\delta + (N-1)(c+\delta)\big)A + \frac{2}{L}\big(2c + (N-1)c\big)A \\ &\quad + \frac{1}{2L}(L-N-1)\big(2\alpha c - 2(3-\alpha)\delta - 2^{\alpha-1}\delta\big)A \\ &= \frac{1}{L}\Big((\alpha L + 4 - 2\alpha)c + \big((\alpha - 2^{\alpha-2} - 3)L + (1+2^{\alpha-2})N + (3 + 2^{\alpha-2} - \alpha)\big)\delta\Big)A. \end{aligned}$$

Arguing as in [15], we may suppose that $N \geq L/2$. Because of $4 - 2\alpha \geq 0$, $(1+2^{\alpha-2})N/L \geq (1+2^{\alpha-2})/2$, and $3 + 2^{\alpha-2} - \alpha \geq 0$, we infer that

$$\begin{aligned} C_1 + C_2 &\geq \left(\frac{1}{L}(\alpha L + 4 - 2\alpha)c + \left(\alpha - \frac{5}{2} - 2^{\alpha-3}\right)\delta\right)A \\ &\geq \left(\alpha c - \left(\frac{5}{2} + 2^{\alpha-3} - \alpha\right)\delta\right)A \end{aligned}$$

which concludes the proof. □

3.4. **Random transposition model.** The random transposition model is a random walk on the group of permutations. Let S_n be the set of permutations on $\{1, 2, \ldots, n\}$ and T_n the set of all transpositions in S_n. Given $1 \leq i, j \leq n$, we denote by $\tau_{ij} \in T_n$ the transposition that interchanges i and j, i.e. $\tau_{ij}(i) = j$, $\tau_{ij}(j) = i$, and $\tau_{ij}(k) = k$ for $k \neq i, j$. The composition of two permutations σ_1, $\sigma_2 \in S_n$ is denoted by $\sigma_1\sigma_2$.

We define a graph structure on the group S_n by saying that two permutations are neighbors if they differ by precisely one transposition. Thus every vertex $\sigma \in S_n$ has $\binom{n}{2} = n(n-1)/2$ neighbors given by $\{\tau_{ij}\sigma\}_{1\leq i,j\leq n}$, and the set of edges is $E_n = \{\{\sigma, \tau_{ij}\sigma\} : 1 \leq i, j \leq n,\ \sigma \in S_n\}$. We write $\sigma \leftrightarrow \tau\sigma$ if $\{\sigma, \tau\sigma\} \in E_n$. The random walk on (S_n, E_n) is then defined by the transition rates $c(\sigma, \tau) = 2/(n(n-1))$ if $\sigma \leftrightarrow \tau\sigma$ and $c(\sigma, \tau) = 0$ otherwise. The generator of the Markov chain reads as

$$\mathcal{L}f(\sigma) = \frac{2}{n(n-1)} \sum_{\tau \in T_n} \nabla_\tau f(\sigma),$$

where $\nabla_\tau f(\sigma) = f(\tau \circ \sigma) - f(\sigma)$. The uniform measure $\pi(\sigma) = 1/n!$ for all $\sigma \in S_n$ is reversible for the above transition rates $c(\sigma, \tau)$. To simplify the notation, we write $\nabla_{ij} = \nabla_\tau$ if $\tau = \tau_{ij}$, $\sigma_{ij} = \tau_{ij} \circ \sigma$, and $\sigma_{ijk} = \tau_{ij} \circ \tau_{jk} \circ \sigma$.

Theorem 11. *For $n \geq 2$, the Beckner inequality* (5) *and the decay estimates* (6) *and* (17) *hold with constant $\lambda = 8/(n(n-1))$.*

Remark 12. Diaconis and Saloff-Coste [14, Section 4.3] report that the logarithmic Sobolev constant satisfies the bounds $1/(3n\log n) \leq \lambda \leq 1/(n-1)$; also see [18, Theorem 1]. Our bound is worse by a factor of $1/n$. The bound $\lambda \geq \alpha(n+2)/(n(n-1))$ was derived in

[6, Section 4]. It is usually better than our bound $\lambda = 8/(n(n-1))$; for very small numbers of n (namely $n < (8/\alpha) - 2$), our result is superior. ☐

Proof. The right-hand side of (16) (except the factor λ) can be written as

$$A = \frac{1}{n(n-1)}\pi\bigg[\sum_{\tau\in T_n}\nabla_\tau\rho^{\alpha-1}(\sigma)\nabla_\tau\rho(\sigma)\bigg] = \frac{1}{2n(n-1)}\pi\bigg[\sum_{i\neq j}\nabla_{ij}\rho^{\alpha-1}(\sigma)\nabla_{ij}\rho(\sigma)\bigg], \tag{34}$$

where the factor 1/2 takes into account that every transposition (i,j) is counted twice. As in [15, Section 4.4], we define $R(\sigma,(i,j),(k,\ell)) = 4/(n^2(n-1)^2)$ if $|\{i,j,k,\ell\}| = 4$ and $R(\sigma,(i,j),(k,\ell)) = 0$ otherwise. Then $\Gamma(\sigma,(i,j),(k,\ell)) = 0$ if $|\{i,j,k,\ell\}| = 4$ and

$$\Gamma(\sigma,(i,j),(k,\ell)) = \frac{4}{n^2(n-1)^2}$$

otherwise. The left-hand side of (16) then becomes

$$\begin{aligned}
&\pi\bigg[\sum_{\gamma,\delta}\Gamma(\sigma,\gamma,\delta)\Big(\nabla_\gamma\rho^{\alpha-1}(\sigma)\nabla_\delta\rho(\sigma) + (\alpha-1)\nabla_\gamma\rho(\sigma)\nabla_\delta\rho(\sigma)\rho^{\alpha-2}(\sigma)\bigg] \\
&\quad = \frac{2}{n^2(n-1)^2}\pi\bigg[\sum_{i\neq j}\nabla_{ij}\rho^{\alpha-1}(\sigma)\nabla_{ij}\rho(\sigma) + 2\sum_{|\{i,j,k\}|=3}\nabla_{ij}\rho(\sigma)\nabla_{ik}\rho^{\alpha-1}(\sigma)\bigg] \\
&\qquad + \frac{2(\alpha-1)}{n^2(n-1)^2}\pi\bigg[\sum_{i\neq j}\nabla_{ij}\rho(\sigma)\nabla_{ij}\rho(\sigma)\rho^{\alpha-2}(\sigma) + 2\sum_{|\{i,j,k\}|=3}\nabla_{ij}\rho(\sigma)\nabla_{ik}\rho(\sigma)\rho^{\alpha-2}(\sigma)\bigg] \\
&\quad = C_1 + C_2.
\end{aligned}$$

The expression C_1 can be estimated exactly as in the proof of Theorem 4.8 in [15] using the reversibility and averaging (see the estimate for $\widetilde{\mathcal{B}}_1(\rho,\psi)$ for $\psi = \rho^{\alpha-1}$):

$$C_1 \geq \frac{2}{n-1}A - \frac{1}{n^2(n-1)^2}\pi\bigg[\sum_{|\{i,j,k\}|=3}\big(\rho^{\alpha-1}(\sigma_{ij}) - \rho^{\alpha-1}(\sigma)\big)^2\widehat{\rho}(\sigma_{ik},\sigma_{ijk})\bigg].$$

We estimate now $B = (C_2 - C_1)/2$:

$$\begin{aligned}
B = \frac{1}{n^2(n-1)^2}\pi\bigg[&\sum_{i\neq j}(\nabla_{ij}\rho^{\alpha-1}(\sigma))^2\nabla_{ij}\rho(\sigma)\widehat{\rho}_1(\sigma,\sigma_{ij}) \\
&+ 2\sum_{|\{i,j,k\}|=3}(\nabla_{ik}\rho^{\alpha-1}(\sigma))^2\nabla_{ij}\rho(\sigma)\widehat{\rho}_1(\sigma,\sigma_{ik})\bigg].
\end{aligned}$$

Arguing as for $\widetilde{\mathcal{B}}_2(\rho,\psi)$ with $\psi = \rho^{\alpha-1}$ in the proof of Theorem 4.8 in [15], it follows that

$$\begin{aligned}
B = &\frac{1}{n^2(n-1)^2}\pi\bigg[\sum_{|\{i,j,k\}|=3}(\nabla_{ij}\rho^{\alpha-1}(\sigma))^2\Big(\rho(\sigma_{ik})\widehat{\rho}_1(\sigma,\sigma_{ij}) + \rho(\sigma_{ijk})\widehat{\rho}_2(\sigma,\sigma_{ij})\Big)\bigg] \\
&- \frac{1}{n^2(n-1)^2}\pi\bigg[\sum_{|\{i,j,k\}|=3}(\nabla_{ij}\rho^{\alpha-1}(\sigma))^2\Big(\rho(\sigma)\widehat{\rho}_1(\sigma,\sigma_{ij}) + \rho(\sigma_{ij})\widehat{\rho}_2(\sigma,\sigma_{ij})\Big)\bigg]
\end{aligned}$$

$$+\frac{1}{2n^2(n-1)^2}\pi\left[\sum_{i\neq j}(\nabla_{ij}\rho^{\alpha-1}(\sigma))^2\nabla_{ij}\rho(\sigma)\Big(\widehat{\rho}_1(\sigma,\sigma_{ij})-\widehat{\rho}_2(\sigma,\sigma_{ij})\Big)\right]$$
$$= B_1 + B_2 + B_3.$$

Property (iv) of Lemma 16 (applied with $\lambda_1 = \lambda_2 = 1$) implies that $B_3 \geq 0$. Combining B_1 and B_2, we can apply Lemma 16 (i) with $s = \rho(\sigma)$, $t = \rho(\sigma_{ij})$, $u = \rho(\sigma_{ik})$, and $v = \rho(\sigma_{ijk})$, leading to

$$B \geq B_1 + B_2 \geq \frac{1}{n^2(n-1)^2}\pi\left[\sum_{|\{i,j,k\}|=3}(\nabla_{ij}\rho^{\alpha-1}(\sigma))^2\big(\widehat{\rho}(\sigma_{ik},\sigma_{ijk})-\widehat{\rho}(\sigma,\sigma_{ij})\big)\right]$$
$$= \frac{1}{n^2(n-1)^2}\pi\left[\sum_{|\{i,j,k\}|=3}(\nabla_{ij}\rho^{\alpha-1}(\sigma))^2\widehat{\rho}(\sigma_{ik},\sigma_{ijk})\right] - \frac{2(n-2)}{n(n-1)}A.$$

Adding the estimations for C_1 and B, one term cancels and we end up with

$$C_1 + C_2 = 2(C_1 + B) \geq 2\left(\frac{2}{n-1} - \frac{2(n-2)}{n(n-1)}\right)A = \frac{8}{n(n-1)}A.$$

This concludes the proof. □

4. Application: Finite-volume discretization of a Fokker-Planck equation

The Bakry-Emery method has been originally applied to Markov diffusion operators or associated Fokker-Planck equations, and the exponential decay for the probability densities with an explicit decay rate was shown. In numerical analysis, the aim is to prove this equilibration property also for numerical discretizations of Fokker-Planck equations. As these discretizations can, at least in some cases, be interpreted as a Markov chain, one may apply Markov chain theory to achieve this goal. This was done by Mielke [23, Section 5.3] to prove exponential decay of the logarithmic entropy for a finite-volume approximation of a Fokker-Planck equation. The proof is based on diagonal dominance properties of the matrices appearing in (2). Our aim is to extend the exponential decay to power-type entropies by combining Mielkes results and the estimate for birth-death processes from Theorem 5. As a by-product, this provides an alternative proof for the case $\alpha \to 1$ without using matrix algebra.

More specifically, we consider a finite-volume approximation of the one-dimensional Fokker-Planck equation

$$\partial_t u = \partial_x(\partial_x u + u\partial_x V), \quad t > 0, \quad u(\cdot, 0) = u_0 \quad \text{in } \mathbb{R}, \tag{35}$$

where $u(x,t)$ describes some probability density and $V(x)$ is a given potential satisfying $e^{-V} \in L^1(\mathbb{R})$. We introduce the uniform grid $x_n = n/N$, $n \in \mathbb{Z}$, where $N \in \mathbb{N}$. The quantity $h = 1/N$ is the grid size. The Fokker-Planck equation has the unique steady state $\pi(x) = Ze^{-V(x)}$, where $Z > 0$ is a normalization constant. The symmetric form of (35),

$$\partial_t \rho = \frac{1}{\pi}\partial_x(\pi\partial_x\rho), \quad \rho = \frac{u}{\pi},$$

motivates the following numerical scheme. We integrate this equation over $[x_{n-1}, x_n]$:

$$\frac{d}{dt}\frac{1}{h}\int_{x_{n-1}}^{x_n}\rho(x,t)dx = \frac{1}{h}\frac{1}{\pi}\big[\pi\partial_x\rho\big]_{x_{n-1}}^{x_n}.$$

We choose ρ_n to approximate $\int_{x_{n-1}}^{x_n}\rho(x,t)dx/h$, $\pi_n = \int_{x_{n-1}}^{x_n}\pi(x)dx/h$, and the numerical flux q_n to approximate $[\pi\partial_x\rho](x_n)/h$. We choose as in [23]

$$q_n = \frac{\kappa_n}{h^2}(\rho_{n+1}-\rho_n), \quad \kappa_n = (\pi_n\pi_{n+1})^{1/2}.$$

The numerical scheme reads as

$$\begin{aligned}\partial_t\rho_n &= \frac{1}{\pi_n}(q_n - q_{n-1}) = \frac{\kappa_n}{h^2\pi_n}(\rho_{n+1}-\rho_n) + \frac{\kappa_{n-1}}{h^2\pi_n}(\rho_{n-1}-\rho_n)\\ &= a(n)\nabla_+\rho_n + b(n)\nabla_-\rho_n,\end{aligned}$$

where we employed the notation of Section 3.1 and $a(n) = \kappa_n/(h^2\pi_n)$, $b(n) = \kappa_{n-1}/(h^2\pi_n)$. The right-hand side can be interpreted as the generator of a birth-death process on $\mathbb{Z}$. According to [10, Section 3.5], the results of Theorem 5 still hold in that case, and the assumption $b(0) = 0$ is clearly not needed. We note that the entropy reads explicitly as

$$\mathrm{Ent}_\pi(\rho) = \sum_{n\in\mathbb{Z}}\pi_n(\rho_n^\alpha - 1), \quad 1 < \alpha \le 2.$$

Theorem 13. *Let $V \in C^2([0,1])$ and $V''(x) \ge \lambda > 0$ for $x \in [0,1]$. Then*

$$\mathrm{Ent}_\pi(\rho_n) \le \mathrm{Ent}_\pi(\rho_0)e^{-2\alpha\lambda_h t}, \quad n \in \mathbb{N},$$

where $\rho_0 = u_0/\pi_0$, $\lambda_h = 2h^{-2}\Phi(h^2\lambda/8)$, and

$$\Phi(s^2) = \frac{3\mathrm{erf}(s) - \mathrm{erf}(3s)}{2\mathrm{erf}(s)} \quad \textit{with} \quad \mathrm{erf}(s) = \frac{2}{\sqrt{p}}\int_0^s e^{-t^2}dt$$

and $p = 3.14159\ldots$ is the number pi (to avoid confusion with the invariant measure π). Moreover, the following discrete Beckner inequality holds:

$$\sum_{n\in\mathbb{Z}}\pi_n(\rho_n^\alpha - 1) \le \alpha\lambda_h\sum_{n\in\mathbb{Z}}\frac{\sqrt{\pi_{n+1}\pi_n}}{h^2}\big(\rho_{n+1}^{\alpha-1} - \rho_n^{\alpha-1}\big)\big(\rho_{n+1}-\rho_n\big).$$

Remark 14. We remark that $\lambda_h \nearrow \lambda$ as $h \to 0$ [23, Corollary 5.5]. Thus, the decay rate is asymptotically sharp. A modified log-Sobolev inequality with constant λ for a finite-difference approximation was proved in [21] for λ-log-concave potentials by translating the Bakry-Emery condition to the discrete case. □

Proof. Note that $a(n)$ and $b(n)$ satisfy the detailed-balance condition (19). The proof is a consequence of Theorem 5 and the results of Mielke [23, Section 5]. In particular, he has shown that $(1-\lambda_h)\pi_n \ge \sqrt{\pi_{n-1}\pi_{n+1}}$. Consequently,

$$a(n) - a(n+1) = \sqrt{\frac{\pi_{n+1}}{\pi_n}} - \sqrt{\frac{\pi_{n+2}}{\pi_{n+1}}} \ge \lambda_h\sqrt{\frac{\pi_{n+1}}{\pi_n}},$$

$$b(n+1)-b(n)=\sqrt{\frac{\pi_n}{\pi_{n+1}}}-\sqrt{\frac{\pi_{n-1}}{\pi_n}}\geq\lambda_h\sqrt{\frac{\pi_n}{\pi_{n+1}}}.$$

Using Lemma 17 and the relation between the arithmetic and geometric mean, it follows that

$$\begin{aligned}
&a(n)-a(n+1)+b(n+1)-b(n)+\Theta\big(a(n)-a(n+1),b(n+1)-b(n)\big)\\
&\quad\geq\alpha\big(a(n)-a(n+1)+b(n+1)-b(n)\big)\\
&\quad\geq 2\alpha\sqrt{(a(n)-a(n+1))(b(n+1)-b(n))}\geq 2\alpha\lambda_h.
\end{aligned}$$

Applying Theorem 5 concludes the proof. □

Appendix A. Properties of the power mean

We show some properties of the power mean

$$\theta(s,t)=\frac{s-t}{s^{\alpha-1}-t^{\alpha-1}},\quad 0<s,t<\infty,\ s\neq t,\quad 1<\alpha\leq 2,$$

which generalizes the logarithmic mean $(s,t)\mapsto(s-t)/(\log s-\log t)$. We define $\theta(s,s)=s^{2-\alpha}/(\alpha-1)$ for $s>0$.

Lemma 15. *The function θ is C^∞, symmetric, positive, increasing and concave on $(0,\infty)^2$. Furthermore, θ and its first partial derivatives are positive homogenous, i.e., $\theta(\lambda s,\lambda t)=\lambda^{2-\alpha}\theta(s,t)$, $\partial_1\theta(\lambda s,\lambda t)=\lambda^{1-\alpha}\partial_1\theta(s,t)$, and $\partial_2\theta(\lambda s,\lambda t)=\lambda^{1-\alpha}\partial_2\theta(s,t)$ for all s, $t>0$ and $\lambda>0$.*

Proof. The regularity, symmetry, and positivity of θ follow by elementary computations. The monotonicity is equivalent to the positivity of the first derivatives. Since

$$\partial_1\theta(s,t)=\frac{s^{\alpha-1}-t^{\alpha-1}-(\alpha-1)(s-t)s^{\alpha-2}}{(s^{\alpha-1}-t^{\alpha-1})^2}\tag{36}$$

(and similarly for $\partial_t\theta(s,t)$), is is sufficient to prove the positivity of $G(s,t)=s^{\alpha-1}-t^{\alpha-1}-(\alpha-1)(s-t)s^{\alpha-2}$. A computation shows that

$$\partial_sG(s,t)=-(\alpha-1)(2-\alpha)(t-s)s^{\alpha-3},$$

and this derivative is decreasing if $s\in(0,t)$ and increasing otherwise. Then $G(s,t)\geq G(t,t)=0$, and the conclusion follows.

For the proof of the concavity, we observe that

$$\theta(s,t)=\frac{1}{\alpha-1}\int_0^1\big((1-m)s^{\alpha-1}+mt^{\alpha-1}\big)^{(2-\alpha)/(\alpha-1)}dm.\tag{37}$$

The concavity of θ is equivalent to the concavity of

$$F(s,t)=((1-m)s^{\alpha-1}+mt^{\alpha-1})^{(2-\alpha)/(\alpha-1)}$$

for fixed $m\in[0,1]$. The Hessian of F, denoted by $\mathrm{Hess}(F)$, is given by

$$\mathrm{Hess}(F)=\frac{2-\alpha}{\alpha-1}\big((1-m)s^{\alpha-1}+mt^{\alpha-1}\big)^{(2-\alpha)/(\alpha-1)-2}(st)^{\alpha-2}M,$$

where $M = (M_{ij})$ with

$$M_{11} = -(1-m)^2(\alpha-1)\left(\frac{t}{s}\right)^{2-\alpha} - m(1-m)(2-\alpha)\frac{t}{s} \le 0,$$
$$M_{12} = M_{21} = m(1-m)(3-2\alpha),$$
$$M_{22} = -m^2(\alpha-1)\left(\frac{s}{t}\right)^{2-\alpha} - m(1-m)(2-\alpha)\frac{s}{t} \le 0.$$

The determinant of M is nonnegative since

$$\det M = (\alpha-1)(2-\alpha)m(1-m)\left(2m(1-m) + m^2\left(\frac{s}{t}\right)^{1-\alpha} + (1-m)^2\left(\frac{t}{s}\right)^{1-\alpha}\right),$$

which, together with M_{11}, $M_{22} \le 0$, shows the concavity of F. ☐

Lemma 16. *The function θ satisfies for all r, s, $t > 0$ and λ_1, $\lambda_2 > 0$,*
(i) $\theta(u,v) - \theta(s,t) \le \partial_1\theta(s,t)(u-s) + \partial_2\theta(s,t)(v-t)$ *(i.e., θ is concave);*
(ii) $s\partial_1\theta(s,t) + t\partial_2\theta(s,t) = (2-\alpha)\theta(s,t)$*;*
(iii) $2^{\alpha-1}r(\partial_1\theta(s,t) + \partial_2\theta(s,t)) - (\theta(r,s) + \theta(r,t)) \ge -2^{\alpha-1}\theta(s,t)$*;*
(iv) $\lambda_1\partial_1\theta(s,t)(s-t) - \lambda_2\partial_2\theta(s,t)(s-t) \le (2-\alpha)|\lambda_1 - \lambda_2|\theta(s,t)$*.*

Proof. Inequality (i) is an immediate consequence of the concavity of θ and Taylor's theorem. Identity (ii) can be obtained by an elementray computation using (36). The proof of (iii) is similar to the proof of Lemma A.2 in [15]. Indeed, setting $u = s/r$ and $v = t/r$ and using the homogeneity properties of θ and its first partial derivatives, inequality (iii) is equivalent to

$$2^{\alpha-1}\big(\partial_1\theta(u,v) + \partial_2\theta(u,v)\big) - \big(\theta(1,u) + \theta(1,v)\big) \ge -2^{\alpha-1}\theta(u,v).$$

This inequality follows from the concavity and the $(2-\alpha)$-homogeneity property of θ and from (i):

$$\theta(1,u) + \theta(1,v) \le 2\theta\left(\frac{u+1}{2}, \frac{v+1}{2}\right) = 2^{\alpha-1}\theta(u+1, v+1)$$
$$\le 2^{\alpha-1}\big(\theta(u,v) + \partial_1\theta(u,v) + \partial_2\theta(u,v)\big).$$

Finally, identity (ii) yields inequality (iv):

$$\lambda_1\partial_1\theta(s,t)(s-t) - \lambda_2\partial_2\theta(s,t)(s-t)$$
$$\le \max\{\lambda_1,\lambda_2\}\big(s\partial_1\theta(s,t) + t\partial_2\theta(s,t)\big) - \min\{\lambda_1,\lambda_2\}\big(t\partial_1\theta(s,t) + s\partial_2\theta(s,t)\big)$$
$$= \big(\max\{\lambda_1,\lambda_2\} - \min\{\lambda_1,\lambda_2\}\big)(2-\alpha)\theta(s,t) = |\lambda_1 - \lambda_2|(2-\alpha)\theta(s,t).$$

This concludes the proof. ☐

Lemma 17. *It holds for all A, $B \ge 0$,*

$$\Theta(A,B) = (\alpha-1)\inf_{s,t>0}\theta(s,t)(As^{\alpha-2} + Bt^{\alpha-2}) \ge (\alpha-1)(A+B).$$

Proof. By (37), it follows that

$$\begin{aligned}(\alpha-1)\theta(s,t)(As^{\alpha-2}+Bt^{\alpha-2}) &= A\int_0^1\left((1-m)+m\left(\frac{t}{s}\right)^{\alpha-1}\right)^{(2-\alpha)/(\alpha-1)}dm \\ &\quad + B\int_0^1\left((1-m)\left(\frac{s}{t}\right)^{\alpha-1}+m\right)^{(2-\alpha)/(\alpha-1)}dm \\ &\geq A\int_0^1(1-m)^{(2-\alpha)/(\alpha-1)}dm + B\int_0^1 m^{(2-\alpha)/(\alpha-1)}dm \\ &= (\alpha-1)(A+B),\end{aligned}$$

which finishes the proof. □

References

[1] C. Ané, S. Blachère, D. Chafa, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques.* Soc. Math. France, Paris, 2000.

[2] D. Bakry and M. Emery. Diffusions hypercontractives. *Séminaire de probabilités* 19 (1983/84), 177-206. Lect. Notes Math. 1123. Springer, Berlin, 1985.

[3] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators.* Springer, Cham, 2014.

[4] W. Beckner. A generalized Poincaré inequality for Gaussian measures. *Proc. Amer. Math. Soc.* 105 (1989), 397-400.

[5] S. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.* 156 (1998), 347-365.

[6] S. Bobkov and P. Tetali. Modified logarithmic Sobolev inequalities in discrete settings. *J. Theor. Prob.* 19 (2006), 289-336.

[7] S. Bochner. Vector fields and Ricci Curvature. *Bull. Amer. Math. Soc.* 52 (1946), 776-797.

[8] 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.

[9] K. Burdzy and W. Kendall. Efficient Markovian couplings: Examples and counterexamples. *Ann. Appl. Prob.* 10 (2000), 362-409.

[10] P. Caputo, P. Dai Pra, and G. Posta. Convex entropy decay via the Bochner-Bakry-Emery approach. *Ann. Inst. H. Poincaré Prob. Stat.* 45 (2009), 734-753.

[11] M. F. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sinica, Engl. Ser.* 12 (1996), 337-360.

[12] M. F. Chen. Variational formulas of Poincaré-type inequalities for birth-death processes. *Acta Math. Sinica, Engl. Ser.* 19 (2003), 625-644.

[13] G.-Y. Chen and L. Saloff-Coste. Spectral computations for birth and death chains. *Stoch. Processes Appl.* 124 (2014), 848-882.

[14] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Prob.* 6 (1996), 695-750.

[15] M. Fathi and J. Maas. Entropic Ricci curvature bounds for discrete interacting systems. To appear in *Ann. Appl. Prob.*, 2015. `arXiv:1501.00562`.

[16] U. Fjordholm, S. Mishra, and E. Tadmor. Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws. *SIAM J. Numer. Anal.* 50 (2012), 544-573

[17] D. Furihata and T. Matsuo. *The Discrete Variational Method. A Structure-Preserving Numerical Method for Partial Differential Equations.* Chapman and Hall/CRC, Boca Raton, 2011.

[18] F. Gao and J. Quastel. Exponential decay of entropy in the randon transposition and Bernoulli-Laplace models. *Ann. Appl. Prob.* 13 (2003), 1591-1600.

[19] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In: J. Azéma et al. (eds.), *Séminaire de Probabilités* 36 (2002), 1-134. Lect. Notes Math. 1801, Springer, Berlin, 2003.

[20] M. Jerrum, J.-B. Son, P. Tetali, and E. Vigoda. Elementary bounds on Poincaré and log-Sobolev constants for decomposable Markov chains. *Ann. Appl. Prob.* 14 (2004), 1741-1765.

[21] O. Johnson. A discrete log-Sobolev inequality under a Bakry-Emery type condition. Preprint, 2015. arXiv:1507.06268..

[22] L. Miclo. An exemple of application of discrete Hardy's inequalities. *Markov Processes Related Fields* 5 (1999), 319-330.

[23] A. Mielke. Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Part. Diff. Eqs.* 48 (2013), 1-31.

[24] L. del Molino, P. Chleboun, and S. Grosskinsky. Condensation in randomly perturbed zero-range processes. J. Phys. A: Math. Theor. 45 (2012), 205001 (17 pp.).

[25] R. Montenegro and P. Tetali. Mathematical Aspects of Mixing Times in Markov Chains. *Found. Trends Theor. Comput. Sci.* 1 (2006), 121 pp.

[26] B. Morris. Spectral gap for the zero range process with constant rate. *Ann. Prob.* 34 (2006), 1645-1664.

[27] F.-Y. Wang. *Functional Inequalities, Markov Semigroups and Spectral Theory.* Science Press, Beijing, 2005.

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria

E-mail address: juengel@tuwien.ac.at

Institute for Analysis and Scientific Computing, Vienna University of Technology, Wiedner Hauptstrasse 8–10, 1040 Wien, Austria

E-mail address: wen.yue@tuwien.ac.at