Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2009, Vol. 45, No. 1, 58–69
DOI: 10.1214/07-AIHP149
© Association des Publications de l'Institut Henri Poincaré, 2009

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Spectral gap and convex concentration inequalities for birth–death processes

Wei Liu[a] and Yutao Ma[b,1]

[a]*Department of Mathematics and Statistics, Wuhan University, 430072-Hubei, China and Laboratoire de Mathématiques Appliquées, Université Blaise Pascal, France*
[b]*School of Mathematical Sciences, Beijing Normal University, China. E-mail: mayt@bnu.edu.cn*

Received 21 December 2006; revised 13 June 2007; accepted 27 September 2007

Abstract. In this paper, we consider a birth–death process with generator $\mathcal{L}$ and reversible invariant probability π. Given an increasing function ρ and the associated Lipschitz norm $\|\cdot\|_{\mathrm{Lip}(\rho)}$, we find an explicit formula for $\|(-\mathcal{L})^{-1}\|_{\mathrm{Lip}(\rho)}$. As a typical application, with spectral theory, we revisit one variational formula of M. F. Chen for the spectral gap of $\mathcal{L}$ in $L^2(\pi)$. Moreover, by Lyons–Zheng's forward-backward martingale decomposition theorem, we get convex concentration inequalities for additive functionals of birth–death processes.

Résumé. Dans ce travail, nous considérons un processus de naissance et de mort de générateur $\mathcal{L}$ et de probabilité invariante réversible π. Étant données une fonction strictement croissante ρ, et la norme lipschitzienne $\|\cdot\|_{\mathrm{Lip}(\rho)}$ par rapport à ρ, nous trouvons une représentation explicite de $\|(-\mathcal{L})^{-1}\|_{\mathrm{Lip}(\rho)}$. En guise d'une application typique, nous retrouvons une formule variationnelle de M. F. Chen pour le trou spectral de $\mathcal{L}$ dans $L^2(\pi)$. De plus, par la décomposition des martingales progressive-rétrogrades de Lyons–Zheng, nous obtenons des inégalités de concentration convexe pour des fonctionnelles additives de processus de naissance et de mort.

MSC: 60E15; 60G27

Keywords: Birth–death process; Spectral gap; Lipschitz function; Poisson equation; Convex concentration inequality

1. Introduction

Consider a birth–death process $(X_t)_{t\geq 0}$ on $\mathbb{N}=\{0,1,2,\ldots\}$ with birth rates $(b_i)_{i\in\mathbb{N}}$ and death rates $(a_i)_{i\in\mathbb{N}}$, i.e., its generator $\mathcal{L}$ is given for any real function f on $\mathbb{N}$ by,

$$\mathcal{L}f(i)=b_i\bigl(f(i+1)-f(i)\bigr)+a_i\bigl(f(i-1)-f(i)\bigr), \tag{1.1}$$

where b_i and a_i are positive for any $i\geq 1$, with furthermore $b_0>0$ and $a_0=0$. For any real function f, $f(-1)$ is supposed to be zero to simplify the notations. Throughout this paper, we assume that the process is positive recurrent, i.e.,

$$\sum_{n\geq 0}\mu_n\sum_{i\geq n}(\mu_i b_i)^{-1}=\infty \quad \text{and} \quad C:=\sum_{n=0}^{+\infty}\mu_n<+\infty,$$

[1]Supported in part by NSFC 10121101. This work was completed when the corresponding author did the post-doc research at MODAL'X, University Paris X from February to August 2007.

where μ given by

$$\mu_0 = 1, \qquad \mu_n = \frac{b_0 b_1 \cdots b_{n-1}}{a_1 a_2 \cdots a_n}, \quad n \ge 1,$$

is an invariant measure of the process. Define the normalized probability π of μ by $\pi_n = \frac{\mu_n}{C}$ for any $n \ge 0$, which is actually the reversible invariant probability of the process.

Our first problem of this paper is related with the spectral gap λ_1 of $\mathcal{L}$ in $L^2(\pi)$, i.e., the infimum of the spectrum of $-\mathcal{L}$ in $L^2(\pi)$. Let $(P_t)_{t\ge 0}$ be the corresponding semigroup of the process. Then λ_1 is the optimal constant ε in the inequality

$$\big\| P_t f - \pi(f) \big\|_{L^2(\pi)} \le \mathrm{e}^{-\varepsilon t} \big\| f - \pi(f) \big\|_{L^2(\pi)},$$

which characterizes the exponential decay of $(P_t)_{t\ge 0}$ to π. On the other hand, λ_1 is the best constant c in the following Poincaré inequality,

$$c \operatorname{Var}_\pi(f) \le \mathcal{E}_\pi(f),$$

where $\operatorname{Var}_\pi(f) := \sum_{i\ge 0} \pi_i (f(i) - \pi(f))^2$ and $\mathcal{E}_\pi(f) := \sum_{i\ge 0} \pi_i b_i (f(i+1) - f(i))^2$ are respectively the variance and Dirichlet form of f with respect to π.

Since Chen and Wang in their paper [8] (1993) used the coupling method to obtain the first eigenvalue on manifold, they and their working group obtained fruitful results (the reader is referred to [6,7] for an account of art). The work of Chen ([2], 1996) was the original one to prove two exact variational formulas of the spectral gap for birth–death processes by coupling method (the diffusion case is due to Chen and Wang ([9], 1997)). For both birth–death processes and diffusion processes, a simple analytic proof of those variational formulas was given by Chen ([3], 1999). Miclo ([17], 1999) extended the Muckenhoupt's generalized Hardy inequality from $\mathbb{R}$ to $\mathbb{N}$ and so derived upper and lower bounds on λ_1 which are different only by a factor 4, generalizing the previous work of Bobkov–Götze ([1], 1999) for one dimensional diffusion processes. Chen ([4], 2000) showed that the results of Bobkov–Götze and Miclo could be derived from his variational formulas. See [5–7] for further results and references on this well developed subject.

The coupling strategy of [2,8,9] etc. consists in finding some appropriate increasing function ρ so that $\|P_t f\|_{\mathrm{Lip}(\rho)} \le \mathrm{e}^{-\varepsilon t} \|f\|_{\mathrm{Lip}(\rho)}$ with $\varepsilon > 0$ as large as possible, and it is shown therein that the supremum of such $\varepsilon = \varepsilon(\rho)$ is exactly λ_1. Our approach here is different: for convex concentration inequalities (our other aim), we must control the norm in the space of Lipschitz functions w.r.t. ρ of the Poisson operator $(-\mathcal{L})^{-1}$. We are inspired by the work of Djellout and Wu ([10], 2007) who calculated explicitly $\|(-\mathcal{L})^{-1}\|_{\mathrm{Lip}(\rho)}$ for one dimensional diffusion processes (so yielding another proof of Chen–Wang's variational formula of λ_1), and applied that formula to give a bound of log-Sobolev constant for Gibbs measure. As in [10] we obtain an exact expression of $\|(-\mathcal{L})^{-1}\|_{\mathrm{Lip}(\rho)}$, but now for birth–death processes.

Our second aim is about convex concentration inequality. Two random variables F and G satisfy a convex concentration inequality if

$$\mathbb{E}\big[\phi(F)\big] \le \mathbb{E}\big[\phi(G)\big] \tag{1.2}$$

for all convex functions $\phi : \mathbb{R} \mapsto \mathbb{R}$ such that the inequality takes sense. This concept was firstly introduced by Hoeffding ([12], 1963) and was realized for general martingales by Klein et al. ([14], 2006). By a classical argument, the application of (1.2) to $\phi(x) = \exp(\lambda x)$, $\lambda > 0$, entails the deviation bound: for any $x > 0$,

$$\mathbb{P}(F \ge x) \le \inf_{\lambda > 0} \mathbb{E}\big[\mathrm{e}^{\lambda(F-x)} 1_{\{F \ge x\}}\big] \le \inf_{\lambda > 0} \mathbb{E}\big[\mathrm{e}^{\lambda(F-x)}\big] \le \inf_{\lambda > 0} \mathbb{E}\big[\mathrm{e}^{\lambda(G-x)}\big]. \tag{1.3}$$

Hence the deviation probabilities for F can be estimated via the Laplace transform of G and that is why lots of jobs have been done on this subject.

Given g a function on $\mathbb{N}$, we consider functionals $S_t = \int_0^t g(X_s)\,\mathrm{d}s$. We will prove concentration convex inequalities for $(S_t)_{t\ge 0}$ and then some deviation inequalities for $(t^{-1} S_t)_{t>0}$. This kind of deviation inequality will characterize the speed of the decay of the empirical measure $(L_t := t^{-1} \int_0^t \delta_{X_s}\,\mathrm{d}s)_{t>0}$ to π.

The remainder of the paper will be organized as follows. In Section 2 we concentrate on the representation of $\|(-\mathcal{L})^{-1}\|_{\mathrm{Lip}(\rho)}$ and in Section 3 we give another proof of Chen's variational formula of the spectral gap λ_1 for $\mathcal{L}$ in $L^2(\pi)$. The last section is devoted to convex concentration inequalities for additive functionals of birth–death processes.

2. Representation of $\|(-\mathcal{L})^{-1}\|_{\mathrm{Lip}(\rho)}$

Given an increasing function $\rho:\mathbb{N}\mapsto\mathbb{R}$, define $d_\rho(i,j)=|\rho(i)-\rho(j)|$ a metric on $\mathbb{N}$ with respect to ρ. We call a function f on $\mathbb{N}$ Lipschitz with respect to ρ (or ρ-Lipschitz) if

$$\|f\|_{\mathrm{Lip}(\rho)}:=\sup_{i\neq j}\frac{|f(j)-f(i)|}{|\rho(j)-\rho(i)|}<+\infty, \tag{2.1}$$

which is equivalent to

$$\|f\|_{\mathrm{Lip}(\rho)}=\sup_{i\geq 0}\frac{|f(i+1)-f(i)|}{\rho(i+1)-\rho(i)}<+\infty.$$

The space of all ρ-Lipschitz functions is denoted by $C_{\mathrm{Lip}(\rho)}$. Throughout this paper, we assume that $\rho\in L^1(\pi)$ and denote by $(C^0_{\mathrm{Lip}(\rho)},\|\cdot\|_{\mathrm{Lip}(\rho)})$ the space of all ρ-Lipschitz functions with $\pi(f):=\int f\,\mathrm{d}\pi=0$. In addition, $\|\cdot\|_{\mathrm{Lip}(\rho)}$ is a norm restricted to $C^0_{\mathrm{Lip}(\rho)}$.

By the equality (1.1), the equation $\mathcal{L}f=0$ admits constant solutions and so identically zero solution when $\pi(f)=0$ is required. Then for any function $g\in C^0_{\mathrm{Lip}(\rho)}$, there exists an unique solution f on $\mathbb{N}$ with $\pi(f)=0$ to the Poisson equation

$$-\mathcal{L}f=g.$$

Thereby $(-\mathcal{L})^{-1}$ is well defined on $C^0_{\mathrm{Lip}(\rho)}$. By definition, $\mathcal{L}$ has a spectral gap in $C^0_{\mathrm{Lip}(\rho)}$ if 0 is an isolated eigenvalue of $-\mathcal{L}$ in $C^0_{\mathrm{Lip}(\rho)}$, or equivalently $(-\mathcal{L})^{-1}:C^0_{\mathrm{Lip}(\rho)}\mapsto C^0_{\mathrm{Lip}(\rho)}$ is bounded.

Recall the usual Lipschitz norm of $(-\mathcal{L})^{-1}$ on $C^0_{\mathrm{Lip}(\rho)}$:

$$\left\|(-\mathcal{L})^{-1}\right\|_{\mathrm{Lip}(\rho)}\triangleq\sup_{\|g\|_{\mathrm{Lip}(\rho)}\leq 1}\left\|(-\mathcal{L})^{-1}g\right\|_{\mathrm{Lip}(\rho)}=\sup_{\|g\|_{\mathrm{Lip}(\rho)}=1}\left\|(-\mathcal{L})^{-1}g\right\|_{\mathrm{Lip}(\rho)}.$$

The main result of this section is

Theorem 2.1. *Let $\mathcal{L},\rho,\|\cdot\|_{\mathrm{Lip}(\rho)}$ be fixed as before and assume that $\rho\in L^1(\pi)$. We have*

$$\left\|(-\mathcal{L})^{-1}\right\|_{\mathrm{Lip}(\rho)}=\sup_{i\geq 1}\frac{\sum_{k=i}^{\infty}\pi_k(\rho(k)-\pi(\rho))}{\pi_i a_i(\rho(i)-\rho(i-1))}=:I(\rho). \tag{2.2}$$

Remarks 2.2. *The parameter $I(\rho)$ defined in (2.2) is not surely finite. Indeed, if $I(\rho)$ is finite, the operator $(-\mathcal{L})^{-1}$ maps the space $C^0_{\mathrm{Lip}(\rho)}$ to $C^0_{\mathrm{Lip}(\rho)}$ itself. Otherwise, there exists at least one function $g\in C^0_{\mathrm{Lip}(\rho)}$ such that $(-\mathcal{L})^{-1}g$ is not ρ-Lipschitz.*

Our approach to this theorem is similar to that of Djellout and Wu [10] while working for one dimensional diffusion processes. We begin the proof with two lemmas.

Lemma 2.3. *Given a function g on $\mathbb{N}$ with $\pi(g)=0$, consider the Poisson equation*

$$-\mathcal{L}f=g. \tag{2.3}$$

For any $i \geq 0$, the solution of the above Eq. (2.3) *satisfies the following relation*:

$$f(i+1) - f(i) = -\frac{\sum_{j=0}^{i} \pi_j g(j)}{\pi_{i+1} a_{i+1}}. \tag{2.4}$$

Proof. *This lemma, to some extent, is known in the sense that one only needs an obvious change in the proof of Lemma* 4.1 *in* [2]. *But we still state it for the convenience of the reader. Indeed, the formula* (2.4) *follows from*

$$\begin{aligned} -\sum_{j=0}^{i} \pi_j g(j) &= \sum_{j=0}^{i} \pi_j \mathcal{L} f(j) = \sum_{j=0}^{i} \big[\pi_j a_j \big(f(j-1) - f(j)\big) + \pi_j b_j \big(f(j+1) - f(j)\big)\big] \\ &= \sum_{j=0}^{i} \big[-\pi_j a_j \big(f(j) - f(j-1)\big) + \pi_{j+1} a_{j+1} \big(f(j+1) - f(j)\big)\big] \\ &= \pi_{i+1} a_{i+1} \big(f(i+1) - f(i)\big). \end{aligned}$$

□

Now we prove the crucial point of this section:

Lemma 2.4. *Provided that $\|g\|_{\mathrm{Lip}(\rho)} = 1$ and $\pi(g) = 0$, we have for any $k \geq 0$,*

$$\sum_{i \geq k} \pi_i g(i) \leq \sum_{i \geq k} \pi_i \big(\rho(i) - \pi(\rho)\big). \tag{2.5}$$

Proof. Set

$$F(k) = \sum_{i=k}^{+\infty} \pi_i g(i) - \sum_{i=k}^{+\infty} \pi_i \big(\rho(i) - \pi(\rho)\big),$$

which satisfies that $F(0) = 0$ and $\lim_{k\to\infty} F(k) = 0$. It suffices to show either $F \equiv 0$ or there exists some $K \in \mathbb{N}$, such that $F(k)$ is nonincreasing for $k \leq K$ and $F(k)$ is nondecreasing for $k > K$. Below we suppose that F is not identically zero.

Simple calculus gives us

$$F(k+1) - F(k) = -\pi_k g(k) + \pi_k \big(\rho(k) - \pi(\rho)\big) = \pi_k \big(\rho(k) - g(k) - \pi(\rho)\big).$$

Define

$$G(k) := \frac{F(k+1) - F(k)}{\pi_k} = \rho(k) - g(k) - \pi(\rho).$$

Since $\|g\|_{\mathrm{Lip}(\rho)} = 1$, we have

$$G(k+1) - G(k) = \rho(k+1) - \rho(k) - \big(g(k+1) - g(k)\big) \geq 0.$$

If $G(0) > 0$, then $G(k) > 0$ for any $k \geq 0$, which implies that F is increasing. We have

$$0 = \lim_{k\to\infty} F(k) \geq F(1) = \pi_0 G(0) > 0,$$

a contradiction. Thus $G(0) \leq 0$.

Since G is nondecreasing, there is at most one time to change its sign. If G does not change its sign, it means $G(k) \leq 0$ for any $k \geq 0$, then F is nonincreasing. For any $N \in \mathbb{N}$,

$$0 = \lim_{k\to\infty} F(k) \leq F(N) \leq \cdots \leq F(0) = 0,$$

which shows $F \equiv 0$, it is not our case. Hence G changes its sign, i.e., there exists some $K \geq 0$, such that

$$G(k) > 0, \quad k > K \quad \text{and} \quad G(k) \leq 0, \quad k \leq K.$$

That is to say, for any $1 \leq k \leq K$, $F(k) \leq F(k-1)$ and $F(k-1) \leq F(k)$ when $k > K$, which completes the proof. $\square$

Remarks 2.5. *Applying (2.5) to $-g$, we get*

$$\left|\sum_{i \geq k} \pi_j g(j)\right| \leq \sum_{i \geq k} \pi_i \big(\rho(i) - \pi(\rho)\big). \tag{2.6}$$

Proof of Theorem 2.1. By Eq. (2.3), for any function $g \in C^0_{\mathrm{Lip}(\rho)}$ with $\|g\|_{\mathrm{Lip}(\rho)} = 1$, we have

$$(-\mathcal{L})^{-1} g = f. \tag{2.7}$$

Therefore with the definition of $\|\cdot\|_{\mathrm{Lip}(\rho)}$ and Lemma 2.3, we have

$$\begin{aligned}
\|f\|_{\mathrm{Lip}(\rho)} &= \sup_{i \geq 0} \frac{|f(i+1) - f(i)|}{\rho(i+1) - \rho(i)} \\
&= \sup_{i \geq 0} \frac{|\sum_{j=i+1}^{+\infty} \pi_j g(j)|}{(\rho(i+1) - \rho(i))\pi_{i+1} a_{i+1}}.
\end{aligned}$$

Then

$$\begin{aligned}
\left\|(-\mathcal{L})^{-1}\right\|_{\mathrm{Lip}(\rho)} &= \sup_{\|g\|_{\mathrm{Lip}(\rho)}=1} \sup_{i \geq 0} \frac{|\sum_{j=i+1}^{+\infty} \pi_j (g(j) - \pi(g))|}{(\rho(i+1) - \rho(i))\pi_{i+1} a_{i+1}} \\
&= \sup_{i \geq 0} \frac{1}{\pi_{i+1} a_{i+1} (\rho(i+1) - \rho(i))} \sup_{\|g\|_{\mathrm{Lip}(\rho)}=1} \left| \sum_{j=i+1}^{+\infty} \pi_j \big(g(j) - \pi(g)\big) \right| \\
&\leq \sup_{i \geq 1} \frac{\sum_{j=i}^{\infty} \pi_j (\rho(j) - \pi(\rho))}{\pi_i a_i (\rho(i) - \rho(i-1))},
\end{aligned} \tag{2.8}$$

where the last inequality follows from (2.6). On the other hand, $\|\rho - \pi(\rho)\|_{\mathrm{Lip}(\rho)} = 1$, then

$$\left\|(-\mathcal{L})^{-1}\right\|_{\mathrm{Lip}(\rho)} \geq \sup_{i \geq 1} \frac{\sum_{j=i}^{\infty} \pi_j (\rho(j) - \pi(\rho))}{\pi_i a_i (\rho(i) - \rho(i-1))}. \tag{2.9}$$

Combining (2.8) and (2.9), we have

$$\left\|(-\mathcal{L})^{-1}\right\|_{\mathrm{Lip}(\rho)} = \sup_{i \geq 1} \frac{\sum_{j=i}^{\infty} \pi_j (\rho_j - \pi(\rho))}{\pi_i a_i (\rho(i) - \rho(i-1))},$$

the desired result. $\square$

3. Application to spectral gap on $L^2(\pi)$

Let $\mathcal{A}$ be the set of all real increasing functions ρ on $\mathbb{N}$ such that $\rho \in L^1(\pi)$. As an application of Theorem 2.1, we revisit one of the variational formulas of the spectral gap λ_1 for birth–death processes, due to M. F. Chen [2]:

Theorem 3.1. *Let λ_1 be the spectral gap of $\mathcal{L}$ in $L^2(\pi)$, then we have*

$$\lambda_1 = \sup_{\rho\in\mathcal{A}} I(\rho)^{-1}, \tag{3.1}$$

where $I(\rho)$ is the same as in Theorem 2.1.

Proof of Theorem 3.1 (Following [19,21]). We prove at first the "$\geq$" of the formula (3.1). Taking any $\rho \in \mathcal{A}$, let

$$\mathcal{D} = L^\infty(\pi) \cap C^0_{\mathrm{Lip}(\rho)},$$

a dense subset of $L^2_0(\pi)$, where 0 represents zero mean under π. We may assume that $I(\rho)$ is finite (trivial otherwise).

The operator $\mathcal{L}$ is self-adjoint and negative definite, so it admits a spectral decomposition on $L^2_0(\pi)$ (see [22]),

$$-\mathcal{L} = \int_{(0,+\infty)} \lambda \, \mathrm{d}E_\lambda, \tag{3.2}$$

then

$$(-\mathcal{L})^{-1} = \int_{(0,+\infty)} \lambda^{-1} \, \mathrm{d}E_\lambda. \tag{3.3}$$

We prove now the following assertion:

$$\text{given } \lambda_0 \in \bigl(0, I(\rho)^{-1}\bigr), \qquad E_{\lambda_0} f = 0 \quad \text{for any } f \in \mathcal{D}. \tag{3.4}$$

For any function g with $\|g\|_{\mathrm{Lip}(\rho)} = 1$ and $\pi(g) = 0$, we have

$$\begin{aligned}
\|g\|_1 &:= \sum_{k=0}^{\infty} \bigl|g(k)\bigr|\pi_k \leq \sum_{k=0}^{\infty} \bigl(\rho(k) - \rho(0)\bigr)\pi_k + \bigl|g(0)\bigr| \\
&= \pi(\rho) - \rho(0) + \frac{|\sum_{k=1}^{\infty} \pi_k g(k)|}{\pi_0} \\
&\leq \pi(\rho) - \pi(0) + \frac{\sum_{k=1}^{\infty} \pi_k (\rho(k) - \pi(\rho))}{\pi_0} \\
&= 2\bigl(\pi(\rho) - \rho(0)\bigr),
\end{aligned} \tag{3.5}$$

where the last but one inequality is ensured by the inequality (2.6) and $\pi(\rho) - \rho(0)$ is positive because ρ is increasing. Theorem 2.1 and the finiteness of $I(\rho)$ guarantee that for any $n \geq 1$, $(-\mathcal{L})^{-n} f$ is in $C^0_{\mathrm{Lip}(\rho)}$ once f belongs to $C^0_{\mathrm{Lip}(\rho)}$. Precisely for any $f \in \mathcal{D}$,

$$\bigl\|(-\mathcal{L})^{-n} f\bigr\|_{\mathrm{Lip}(\rho)} \leq I(\rho)^n \|f\|_{\mathrm{Lip}(\rho)}. \tag{3.6}$$

Thus with (3.5), we have

$$\begin{aligned}
\bigl\langle f, (-\mathcal{L})^{-n} f\bigr\rangle_\pi &\leq \|f\|_\infty \bigl\|(-\mathcal{L})^{-n} f\bigr\|_1 \leq 2\bigl(\pi(\rho) - \rho(0)\bigr)\|f\|_\infty \bigl\|(-\mathcal{L})^{-n} f\bigr\|_{\mathrm{Lip}(\rho)} \\
&\leq 2\bigl(\pi(\rho) - \rho(0)\bigr)\|f\|_\infty \|f\|_{\mathrm{Lip}(\rho)} I(\rho)^n.
\end{aligned}$$

On the other hand, in $[0, +\infty]$ we always have

$$\begin{aligned}
\bigl\langle f, (-\mathcal{L})^{-n} f\bigr\rangle_\pi &= \int_{(0,+\infty)} \lambda^{-n} \, \mathrm{d}\langle E_\lambda f, f\rangle_\pi \\
&\geq \lambda_0^{-n} \langle E_{\lambda_0} f, f\rangle_\pi.
\end{aligned}$$

Combining those two inequalities, we get

$$\langle E_{\lambda_0} f, f\rangle_\pi \le C(f)\big(\lambda_0 I(\rho)\big)^n, \tag{3.7}$$

where $C(f) = 2(\pi(\rho) - \rho(0))\|f\|_\infty \|f\|_{\mathrm{Lip}(\rho)}$ is a finite constant independent of n. Then the assertion (3.4) follows from (3.7) by letting $n \to +\infty$. Since $\mathcal{D}$ is dense in $L_0^2(\pi)$ and E_{λ_0} is bounded, then $E_{\lambda_0} f = 0$ for any $f \in L^2(\pi)$, which means $E_{\lambda_0} = 0$. Equivalently

$$\lambda_1 \ge I(\rho)^{-1}$$

and so immediately

$$\lambda_1 \ge \sup_{\rho \in \mathcal{A}} I(\rho)^{-1}.$$

Now take $\bar\rho$ the eigenfunction of $\mathcal{L}$ corresponding to λ_1 in weak sense, i.e.,

$$\mathcal{L}\bar\rho = -\lambda_1 \bar\rho \quad \text{pointwise}.$$

It is known that $\bar\rho$ is an increasing function on $\mathbb{N}$ in $L^1(\pi)$ (see Lemma 4.2 in [2] for details) and then belongs to $\mathcal{A}$. As showed in the proof of Theorem 2.1, the parameter

$$I(\bar\rho) := \big\|(-\mathcal{L})^{-1}\big\|_{\mathrm{Lip}(\bar\rho)} = \sup_{\|g\|_{\mathrm{Lip}(\bar\rho)}=1} \big\|(-\mathcal{L})^{-1}\big(g - \pi(g)\big)\big\|$$

attains the supremum at $\bar\rho$, then equals to $1/\lambda_1$. The proof is complete now. $\square$

Remarks 3.2. *In his paper* [2], *Chen established the variational formula* (3.1). *Then in* [3–5], *he reproved it by different methods. And furthermore with this variational formula, Chen in* [4] *derived explicit bounds on the spectral gap, which recovered the lower and upper bounds, differing up to a factor* 4, *obtained originally by Miclo in* [17] *via generalized Hardy's inequality.*

Now we are in position to state convex concentration inequalities.

4. Convex concentration inequalities for additional functionals: The method of Lyons–Zheng forward–backward martingale decomposition

4.1. *A general result*

In this section, $(N_t)_{t\ge 0}$ is always supposed to be a standard Poisson process independent of $(X_t)_{t\ge 0}$ and ρ is an increasing function in $L^1(\pi)$. Let

$$\mathcal{C}_c := \big\{\phi : \mathbb{R} \mapsto \mathbb{R}, \phi \text{ is convex and } \phi'' \text{ is nondecreasing}\big\}.$$

In the sequel, we will add the notation of the probability to be precise with respect to which the expectation is considered. Firstly, we recall one result for pure jump martingales (see [14,16] for details):

Theorem 4.1. *Let $(M_t)_{t\ge 0}$ be a pure jump martingale on some probability space $(E, \mathcal{E}, \mathbb{P})$ satisfying for all $t \ge 0$,*

$$|\Delta M_t| \le K \quad \textit{and} \quad \big\|\langle M\rangle_t\big\|_\infty < +\infty,$$

where $\|\langle M\rangle_t\|_\infty := \operatorname{ess\,sup}_{\omega\in E} |\langle M\rangle_t(\omega)|$. Then for any function $\phi \in \mathcal{C}_c$ and any $t \ge 0$, we have

$$\mathbb{E}_{\mathbb{P}}\big[\phi\big(M_t - \mathbb{E}_{\mathbb{P}}[M_t]\big)\big] \le \mathbb{E}\bigg[\phi\bigg(K N_{\|\langle M\rangle_t\|_\infty/K^2} - \frac{\|\langle M\rangle_t\|_\infty}{K}\bigg)\bigg]. \tag{4.1}$$

Moreover, (4.1) *still holds if $\|\langle M\rangle_t\|_\infty$ is replaced by some deterministic function bounding above such a quantity.*

Now we return to the birth–death process $(X_t)_{t\geq 0}$. Recall the operator Γ of birth–death processes with respect to $\mathcal{L}$, which is defined for any functions f and g on $\mathbb{N}$ as,

$$\Gamma(f,g) = \frac{1}{2}\big[\mathcal{L}(fg) - \mathcal{L}fg - f\mathcal{L}g\big].$$

Noting $\Gamma(f,f) := \Gamma(f)$, we have for any $k \geq 0$,

$$\Gamma(f)(k) = a_k\big(f(k-1) - f(k)\big)^2 + b_k\big(f(k+1) - f(k)\big)^2.$$

For any $t \geq 0$, define

$$M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,\mathrm{d}s.$$

Theorem 4.1 entails the following

Proposition 4.2. *Let f be a function satisfying $K = \sup_{k\geq 1}|f(k) - f(k-1)| < \infty$ and $\|\Gamma(f)\|_\infty := \sup_{k\geq 0}\Gamma(f)(k) < \infty$. Then for any function $\phi \in \mathcal{C}_c$ and $t \geq 0$,*

$$\mathbb{E}\big[\phi(M_t)\big] \leq \mathbb{E}\left[\phi\left(KN_{\|\Gamma(f)\|_\infty t/K^2} - \frac{\|\Gamma(f)\|_\infty t}{K}\right)\right]. \tag{4.2}$$

Proof. Given any $T > 0$, the process $(M_t)_{0\leq t\leq T}$ is a pure jump martingale. By definition, for any $0 \leq t \leq T$, we have

$$|\Delta M_t| \leq K, \quad \text{since } |\Delta M_t| \leq \sup_{k\geq 0}\big|f(k+1) - f(k)\big|.$$

On the other hand, for any $0 \leq t \leq T$, by Ito's formula, $\mathbb{E}[M_t^2] = \mathbb{E}\int_0^t \Gamma(f)(X_s)\,\mathrm{d}s$, hence

$$\|\langle M\rangle_t\|_\infty = \left\|\int_0^t \Gamma f(X_s)\,\mathrm{d}s\right\|_\infty \leq \big\|\Gamma(f)\big\|_\infty t.$$

Then the inequality (4.2) is satisfied for any $0 \leq t \leq T$ and in particular holds for T. The arbitrariness of T completes the proof. □

Remarks 4.3. *Taking ρ as $\rho(i) = i$ for all $i \in \mathbb{N}$ deriving the classical metric, suppose that $f \in C_{\mathrm{Lip}(\rho)}$ and $\|\Gamma(f)\|_\infty < \infty$. Thereby we have for any function $\phi \in \mathcal{C}_c$ and $t \geq 0$,*

$$\mathbb{E}\big[\phi(M_t)\big] \leq \mathbb{E}\left[\phi\left(\|f\|_{\mathrm{Lip}(\rho)}N_{\|\Gamma(f)\|_\infty t/\|f\|^2_{\mathrm{Lip}(\rho)}} - \frac{\|\Gamma(f)\|_\infty t}{\|f\|_{\mathrm{Lip}(\rho)}}\right)\right].$$

Now given g a function on $\mathbb{N}$ with $\pi(g) = 0$ for simplicity. Consider functionals $S_t = \int_0^t g(X_s)\,\mathrm{d}s$, we want to give convex concentration inequalities for $(S_t)_{t\geq 0}$. Lyons–Zheng's forward backward martingale decomposition theorem (see [15,20]) inspires us for any $t \geq 0$ to define

$$\overrightarrow{M}_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)\,\mathrm{d}s$$

and

$$\overleftarrow{M}_t = f(X_0) - f(X_t) - \int_0^t \mathcal{L}f(X_s)\,\mathrm{d}s,$$

where f satisfies $-\mathcal{L}f = g$. Obviously $S_t = \frac{1}{2}(\overleftarrow{M_t} + \overrightarrow{M_t})$ and moreover by the reversibility of $((X_t)_{t\geq 0}, \mathbb{P}_\pi)$, $\overrightarrow{M_t}$ and $\overleftarrow{M_t}$ have the same distribution under $\mathbb{P}_\pi$. Therefore we have for any convex function ϕ,

$$\mathbb{E}_\pi\big[\phi(S_t)\big] = \mathbb{E}_\pi\bigg[\phi\bigg(\frac{\overleftarrow{M_t} + \overrightarrow{M_t}}{2}\bigg)\bigg] \leq \frac{\mathbb{E}_\pi[\phi(\overleftarrow{M_t}) + \phi(\overrightarrow{M_t})]}{2} = \mathbb{E}_\pi\big[\phi(\overrightarrow{M_t})\big].$$

If we impose some hypotheses on g such that the solution f to the Poisson equation $-\mathcal{L}f = g$ verifies the conditions of Proposition 4.2, we could derive convex concentration inequalities for $(S_t)_{t\geq 0}$. The following essential hypotheses appear here just for this aim.

- Hypothesis A:

$$K := \sup_{k\geq 1} \frac{|\sum_{i=k}^\infty \pi_i g(i)|}{\pi_k a_k} < \infty,$$

- Hypothesis B:

$$\sup_{k\geq 0}\bigg\{a_k\bigg(\frac{\sum_{i=k}^\infty \pi_i g(i)}{\pi_k a_k}\bigg)^2 + b_k\bigg(\frac{\sum_{i=k+1}^\infty \pi_i g(i)}{\pi_{k+1} a_{k+1}}\bigg)^2\bigg\} < \infty.$$

Remarks 4.4. *By Lemma 2.3, we have $K = \sup_{k\geq 1}|f(k) - f(k-1)|$ and the Hypothesis* B *is equivalent to $\|\Gamma(f)\|_\infty < \infty$. Moreover if $\rho - \pi(\rho)$ verifies the Hypotheses* A *and* B, *so does any $g \in C^0_{\mathrm{Lip}(\rho)}$.*

The functions f, g used together thereafter are supposed to satisfy the Poisson equation $-\mathcal{L}f = g$. We have

Theorem 4.5. *For any function g verifying the Hypotheses* A *and* B, *we have for any $\phi \in \mathcal{C}_c$ and $t \geq 0$,*

$$\mathbb{E}_\pi\bigg[\phi\bigg(\int_0^t g(X_s)\,\mathrm{d}s\bigg)\bigg] \leq \mathbb{E}\bigg[\phi\bigg(K N_{\|\Gamma(f)\|_\infty t/K^2} - \frac{\|\Gamma(f)\|_\infty t}{K}\bigg)\bigg]. \tag{4.3}$$

Furthermore, the inequality (4.3) still holds for any $g \in C^0_{\mathrm{Lip}(\rho)}$ if $\rho - \pi(\rho)$ satisfies the Hypotheses A *and* B.

Remarks 4.6. *Suppose that $g - \pi(g)$ satisfies the Hypotheses* A *and* B. *As introduced before, the Laplace transform of the right-hand side of (4.3) offers a deviation inequality for $t^{-1}\int_0^t g(X_s)\,\mathrm{d}s$, precisely for any $x > 0$ and $t \geq 0$,*

$$\begin{aligned}\mathbb{P}_\pi\bigg(t^{-1}\int_0^t g(X_s)\,\mathrm{d}s - \pi(g) \geq x\bigg) &\leq \inf_{\lambda>0}\exp\bigg\{-\lambda t x + \big\|\Gamma(f)\big\|_\infty \frac{\mathrm{e}^{\lambda K} - \lambda K - 1}{K^2}\bigg\} \\ &= \exp\bigg\{-\frac{\|\Gamma(f)\|_\infty t}{K^2} h\bigg(\frac{Kx}{\|\Gamma(f)\|_\infty}\bigg)\bigg\},\end{aligned} \tag{4.4}$$

where $h(u) = (1+u)\log(1+u) - u$. Similarly, the inequality (4.4) is true for any $g \in C_{\mathrm{Lip}(\rho)}$ when $\rho - \pi(\rho)$ satisfies the Hypotheses A *and* B.

Remarks 4.7. *Suppose that ν is a probability on $\mathbb{N}$ absolutely continuous with respect to π with $\frac{\mathrm{d}\nu}{\mathrm{d}\pi} \in L^2(\pi)$ and $g - \pi(g)$ verifies the Hypotheses* A *and* B. *With the inequality (4.4) and Cauchy–Schwarz inequality, we could have for any $x > 0$ and $t \geq 0$,*

$$\mathbb{P}_\nu\bigg(t^{-1}\int_0^t g(X_s)\,\mathrm{d}s - \pi(g) \geq x\bigg) \leq \bigg\|\frac{\mathrm{d}\nu}{\mathrm{d}\pi}\bigg\|_{L^2(\pi)} \exp\bigg\{-\frac{\|\Gamma(f)\|_\infty t}{2K^2} h\bigg(\frac{Kx}{\|\Gamma(f)\|_\infty}\bigg)\bigg\}.$$

Remarks 4.8. *In [13], Joulin obtained also deviation inequalities, which are somewhat similar to (4.4), with respect to the probability measure $\mathbb{P}_x$. His proof relied on Wasserstein's curvature whose positivity is guaranteed by the discrete*

Bakry–Émery criterion. In fact, this criterion ensures that $I(\rho)$ is finite and so when the space of Lipschitz functions is considered, the second condition in his Lemma 5.4 *is equivalent to our Hypothesis* B. *However our Hypothesis* A *and the first condition in his Lemma* 5.4 *are not comparable.*

Even though our hypotheses are quite natural with our method, they seem complicated to be verified. Next we give two classical birth–death processes to show how such hypotheses are satisfied.

4.2. *Two classical examples*

*The M/M/*1 *queueing process*
The M/M/1 queueing process is a simple birth–death process whose generator is given for any function f on $\mathbb{N}$ by,

$$\mathcal{L}f(i) = \lambda\big(f(i+1) - f(i)\big) + \nu 1_{i\neq 0}\big(f(i-1) - f(i)\big), \quad i \in \mathbb{N},$$

where the positive numbers λ and ν correspond respectively to the input rate and service rate of the queue: the independent and identically distributed interarrival times and independent and identically distributed service times of the customers follow an exponential law with respective parameters λ and ν. We assume here that $\sigma := \lambda/\nu < 1$, then the process is ergodic and its reversible invariant probability π is the geometric distribution with parameter σ, i.e.,

$$\pi_k = (1-\sigma)\sigma^k, \quad \forall k \geq 0.$$

For this simple example, we have

Proposition 4.9. *Let $(X_t)_{t\geq 0}$ be the M/M/*1 *queueing process defined above. Suppose that $\pi(g) = 0$ and $K := \frac{1}{\nu-\lambda}\sup_{k\geq 1}|\sum_{i=0}^{\infty}\pi_i g(i+k)| < \infty$. Then for any function $\phi \in \mathcal{C}_c$ and $t \geq 0$, we have*

$$\mathbb{E}_\pi\left[\phi\left(\int_0^t g(X_s)\,\mathrm{d}s\right)\right] \leq \mathbb{E}\big[\phi\big(KN_{(\lambda+\nu)t} - (\lambda+\nu)Kt\big)\big]. \tag{4.5}$$

Proof. By Theorem 4.5, it is sufficient to verify the Hypotheses A and B. We have

$$\begin{aligned}\sup_{k\geq 1}\left|\frac{\sum_{i=k}^{\infty}\pi_i g(i)}{\pi_k a_k}\right| &= \sup_{k\geq 1}\left|\frac{1}{\nu}\sum_{i=k}^{\infty}\frac{\sigma^i g(i)}{\sigma^k}\right| = \frac{1}{\nu}\sup_{k\geq 1}\left|\sum_{i=0}^{\infty}\sigma^i g(i+k)\right| \\ &= \frac{1}{\nu}(1-\sigma)^{-1}\sup_{k\geq 1}\left|\sum_{i=0}^{+\infty}\pi_i g(i+k)\right| = K < \infty.\end{aligned}$$

Furthermore, the Hypothesis B is verified as the Hypothesis A since we obtain

$$\big\|\Gamma(f)\big\|_\infty \leq (\lambda+\nu)K^2 < +\infty. \qquad \square$$

Remarks 4.10. *Suppose that ρ satisfies*

$$C(\rho) := \sup_{k\geq 1}\sum_{i=0}^{\infty}\pi_i\big(\rho(i+k) - \rho(i)\big) < \infty.$$

Then Proposition 4.9 is verified with $K = C(\rho)(\nu-\lambda)$ for any $g \in C^0_{\mathrm{Lip}(\rho)}$.

The M/M/∞ queueing process
The $M/M/\infty$ model is a particular birth–death process whose generator $\mathcal{L}$ is given for any functional f on $\mathbb{N}$ by,

$$\mathcal{L}f(i) = \lambda\big(f(i+1) - f(i)\big) + \nu i\big(f(i-1) - f(i)\big),$$

where λ, ν are two positive numbers. Then this process is ergodic with reversible invariant probability π, the Poisson measure on $\mathbb{N}$ with parameter $\sigma := \lambda/\nu$, i.e.

$$\pi_k = \mathrm{e}^{-\sigma}\frac{\sigma^k}{k!}, \quad k \in \mathbb{N}.$$

For this model, we have

Proposition 4.11. *Given a function g satisfying $\pi(g) = 0$ and*

$$\sup_{k\ge 1}\sqrt{k}\frac{|\sum_{i=k}^{+\infty}\sigma^i g(i)/i!|}{\nu k\sigma^k/k!} \le K, \tag{4.6}$$

where K is a positive constant, we have for any $\phi \in \mathcal{C}_c$ and $t \ge 0$,

$$\mathbb{E}_\pi\left[\phi\left(\int_0^t g(X_s)\,\mathrm{d}s\right)\right] \le \mathbb{E}\big[\phi\big(KN_{(\lambda+\nu)t} - (\lambda+\nu)Kt\big)\big]. \tag{4.7}$$

Proof. As for M/M/1 model, we verify the Hypotheses A and B. Indeed,

$$\sup_{k\ge 1}\left|\sum_{i=k}^{+\infty}\frac{\pi_i g(i)}{\pi_k a_k}\right| = \sup_{k\ge 1}\left|\sum_{i=k}^{+\infty}\frac{\sigma^i}{i!}g(i)\right|\frac{(k-1)!}{\nu\sigma^k} \le \sup_{k\ge 1}\frac{K}{\sqrt{k}} = K,$$

which implies the Hypothesis A and moreover $|f(k-1) - f(k)| \le \frac{K}{\sqrt{k}}$. As a consequence,

$$\|\Gamma(f)\|_\infty \le (\lambda+\nu)K^2,$$

the Hypothesis B. □

Here we translate again the above proposition to Lipschitz space.

Corollary 4.12. *Take ρ as $\rho(i) = \sqrt{i}$ for any $i \ge 0$. Then the condition (4.6) is satisfied with $K = \mathrm{e}^\sigma\|g\|_{\mathrm{Lip}(\rho)}/\nu$ for any function $g \in C^0_{\mathrm{Lip}(\rho)}$.*

Proof. For any $i \ge 1, k \ge 1$, the following inequality holds

$$\binom{k+i}{k} := \frac{(k+i)!}{i!k!} \ge \binom{i+k}{1} = i + k, \tag{4.8}$$

where $\binom{k+i}{k}$ is the combination function. Therefore we have

$$\begin{aligned}
\sup_{k\ge 1}\sqrt{k}\frac{|\sum_{i=k}^{+\infty}\sigma^i g(i)/i!|}{\sigma^k\nu k/k!} &\le \sup_{k\ge 1}\frac{\sum_{i=k}^{+\infty}\sigma^i(\rho(i)-\pi(\rho))/i!}{\sigma^k\nu\sqrt{k}/k!}\|g\|_{\mathrm{Lip}(\rho)} \\
&\le \sup_{k\ge 1}\frac{1}{\nu}\sum_{i=0}^{\infty}\frac{\sigma^i k!}{(i+k)!}\frac{\sqrt{i+k}}{\sqrt{k}}\|g\|_{\mathrm{Lip}(\rho)} \\
&\le \frac{1}{\nu}\left(1 + \sup_{k\ge 1}\sum_{i=1}^{\infty}\frac{\sigma^i\sqrt{i+k}}{i!(i+k)\sqrt{k}}\right)\|g\|_{\mathrm{Lip}(\rho)} \\
&\le \frac{1}{\nu}\left(1 + \sum_{i=1}^{\infty}\frac{\sigma^i}{i!}\right)\|g\|_{\mathrm{Lip}(\rho)} = \frac{\mathrm{e}^\sigma}{\nu}\|g\|_{\mathrm{Lip}(\rho)},
\end{aligned}$$

where the first inequality follows from Lemma 2.5, the positivity of $\pi(\rho)$ guarantees the second one, and the third is due to the inequality (4.8). The proof is now complete. □

Remarks 4.13. *In fact, such a function ρ of order $1/2$ is optimal in the sense that Corollary 4.12 fails when $\rho(k) = k^a, k \geq 0$ for any $a > 1/2$. Joulin [13] worked on this model with the metric $d(i,j) = \sum_{k=i}^{j} \frac{1}{\sqrt{k}}$ for any $1 \leq i \leq j$. Since for any $i \geq 1$,*

$$\frac{1}{2\sqrt{i+1}} \leq \sqrt{i+1} - \sqrt{i} \leq \frac{1}{2\sqrt{i}},$$

his metric is equivalent to ours. In addition, Guillin et al. in [11] studied this model with the same metric while they obtained a Gaussian tail estimation via information inequality.

Acknowledgments

The authors are grateful to the anonymous referees for their remarks and questions which help to greatly improve the last version of this article. The authors would also like to thank Professor Liming Wu for many helps and constructive suggestions.

References

[1] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** (1999) 1–28. MR1682772
[2] M. F. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sin. New Ser.* **12** (1996) 337–360. MR1457859
[3] M. F. Chen. Analytic proof of dual variational formula for the first eigenvalue in dimension one. *Sci. Sin. (A)* **42** (1999) 805–815. MR1738551
[4] M. F. Chen. Explicit bounds of the first eigenvalue. *Sci. China (A)* **43** (2000) 1051–1059. MR1802148
[5] M. F. Chen. Variational formulas and approximation theorems for the first eigenvalue. *Sci. China (A)* **44** (2001) 409–418. MR1831443
[6] M. F. Chen. *From Markov Chains to Non-equilibrium Particle Systems*, 2nd edition. Springer, 2004. MR2091955
[7] M. F. Chen. *Eigenvalues, Inequalities and Ergodic Theory.* Springer, 2005. MR2105651
[8] M. F. Chen and F. Y. Wang. Application of coupling method to the first eigenvalue on manifold. *Sci. Sin. (A)* **23** (1993) 1130–1140 (Chinese Edition); **37** (1994) 1–14 (English Edition). MR1308707
[9] M. F. Chen and F. Y. Wang. Estimation of spectral gap for elliptic operators. *Trans. Amer. Math. Soc.* **349** (1997) 1239–1267. MR1401516
[10] H. Djellout and L. M. Wu. Spectral gap of one dimensional diffusions in Lipschitz norm and application to log-Sobolev inequalities for Gibbs measures. Preprint, 2007.
[11] A. Guillin, C. Léonard, L. M. Wu and N. Yao. Transportation-information inequalities for Markov processes. Preprint, 2007.
[12] W. Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Stat. Assoc.* **58** (1963) 13–30. MR0144363
[13] A. Joulin. A new Poisson-type deviation inequality for Markov jump process with positive Wasserstein curvature. Preprint, 2007.
[14] T. Klein, Y. T. Ma and N. Privault. Convex concentration inequalities and forward/backward stochastic calculus. *Electron. J. Probab.* **11** (2006) 486–512. MR2242653
[15] T. J. Lyons and W. A. Zheng. A crossing estimate for the canonical process on a Dirichlet space and a tightness result. *Astérique* **157–158** (1988) 249–271.
[16] Y. T. Ma. Grandes déviations et concentration convexe en temps continu et discret. PhD thesis, Université de La Rochelle (France) et Université de Wuhan (Chine), 2006. Available at http://perso.univ-lr.fr/yma/thesis.pdf.
[17] L. Miclo. An exemple of application of discrete Hardy's inequalities. *Markov Process. Related Fields* **5** (1999) 319–330. MR1710983
[18] Z. K. Wang and X. Q. Yang. *Birth–Death Processes and Markov Chains.* Academic Press of China, Beijing, 2005 (in Chinese).
[19] L. M. Wu. Moderate deviations of dependent random variables related to CLT. *Ann. Probab.* **23** (1995) 420–445. MR1330777
[20] L. M. Wu. Forward–backward martingale decomposition and compactness results for additive functionals of stationary ergodic Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.* **35** (1999) 121–141. MR1678517
[21] L. M. Wu. Essential spectral radius for Markov semigroups (I): discrete time case. *Probab. Theory Related Fields* **128** (2004) 255–321. MR2031227
[22] K. Yosida. *Functional Analysis*, 6th edition. Spring, 1999.