Poisson-type deviation inequalities for curved continuous time Markov chains

Aldéric Joulin*
MODAL'X, Bât. G, Université Paris 10
200, Avenue de la République
92001 Nanterre Cedex
France

September 25, 2006

Abstract

In this paper, we present new Poisson-type deviation inequalities for continuous time Markov chains whose Wasserstein curvature or Γ-curvature is bounded below. Although these two curvatures are equivalent for Brownian motion on Riemannian manifolds, they are not comparable in discrete settings and yield different deviation bounds. In the case of birth-death processes, we provide some conditions on the transition rates of the associated generator for such discrete curvatures to be bounded below, and we extend the deviation inequalities established in [1] for continuous time random walks, seen as models in null curvature. Some applications of these tail estimates are given to Brownian driven Ornstein-Uhlenbeck processes and $M/M/1$ queues.

Key words: continuous time Markov chain, deviation inequality, semigroup, Wasserstein curvature, Γ-curvature, birth-death process.
Mathematics Subject Classification. 60E15, 60J27, 47D07, 41A25.

1 Introduction

Let μ be a probability measure on a metric space (E, d) and let $h : \mathbb{R}_+ \to \mathbb{R}_+$ be a function tending to 0 at infinity. The measure μ is said to satisfy a deviation

*ajoulin@u-paris10.fr

inequality of speed h if for any real Lipschitz function f on (E, d) with Lipschitz constant smaller than 1, the following inequality holds:

$$\mu\left(f - \mu(f) \geq y\right) \leq h(y), \quad y > 0.$$

Applying also the above inequality to $-f$ entails a concentration inequality stating that any Lipschitz map is concentrated around its mean under μ with high probability. In particular, the concentration is said to be Gaussian if h is of order $\exp(-y^2)$ for large y, whereas it is of Poisson type if h is of order $\exp(-y \log(y))$.

Actually, the concentration of measure phenomenon is useful to determine the rate of convergence of functionals involving a large number of random variables. In recent years, this area has been deeply investigated in the context of dependent random variables such as Markov chains. For instance, Gaussian concentration was put forward through transportation cost inequalities in the papers [11] then [6], whereas the authors in [1], [12] and [9] established some appropriate functional inequalities to derive Gaussian and Poisson-type deviation inequalities.

The purpose of the present paper is to give new deviation inequalities of Poisson type for continuous time Markov chains, which extend the tail estimates of [1]. Our approach is based on semigroup analysis and uses the notion of curvature for Markov processes on general metric measure spaces recently investigated in [14]. Although the various Brownian curvatures on a smooth Riemannian manifold are essentially equivalent and characterize the uniform lower bounds on the Ricci curvature of the manifold, such an equivalence does not hold for continuous time Markov chains since discrete gradients do not satisfy in general the chain rule formula. Thus, it is natural to study the role played by each type of discrete curvature in the concentration of measure phenomenon.

The paper is organized as follows. In Section 2, two different notions of curvatures of continuous time Markov chains are introduced: the Wasserstein curvature and the Γ-curvature. The next two sections are concerned with the main results of the paper. Namely, in Section 3, Theorem 3.1, a Poisson-type deviation inequality is established for continuous time Markov chains with Wasserstein curvature bounded below and bounded angle bracket, whereas a general estimate is derived in Section 4, Theorem 4.2, under the hypothesis of a lower bound on the Γ-curvature. With further assumptions on the chain, the latter upper bound is computed to yield Poisson tail probabilities involving the mixed Lipschitz seminorms $\|\cdot\|_{\mathrm{Lip}_d}$ and $f \mapsto \|\Gamma f\|_\infty^{1/2}$, with Γ the "carré du champ" operator, which allow us to relax the boundedness hypothesis on the angle bracket. The case of birth-death processes on $\mathbb{N}$ or on $\{0, 1, \ldots, n\}$ is investigated in Section 5. More precisely, we give some conditions on the transition rates of the associated generator for such discrete curvatures to be bounded below. Together with the tail estimates emphasized above, we extend to birth-death processes the deviation inequalities of [1] established for continuous time random walks on graphs, seen as models in null curvature since the transition rates of the generator do not depend on the space variable. Finally, some applications of these tail estimates are given to Brownian driven Ornstein-Uhlenbeck processes and $M/M/1$ queues.

2 Notation and preliminaries on curvatures

Throughout the paper, E is a countable set endowed with a metric d (different of the trivial one $\varrho(x,y) = 1_{x \neq y}$, $x, y \in E$), $\mathscr{F}(E)$ is the collection of all real-valued functions on E, $\mathscr{B}(E) \subset \mathscr{F}(E)$ is the subspace of bounded functions on E equipped with the supremum norm $\|f\|_\infty = \sup_{x \in E} |f(x)|$, and the space $\mathrm{Lip}_d(E)$ consists of Lipschitz functions on E, i.e.

$$\|f\|_{\mathrm{Lip}_d} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x,y)} < +\infty.$$

Consider on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \geq 0}, \mathbb{P})$ an E-valued regular non-explosive continuous time Markov chain $(X_t)_{t \geq 0}$, where regularity is understood in the sense of [5]. The generator $\mathcal{L}$ of the chain is given by

$$\mathcal{L}f(x) = \sum_{y \in E} (f(y) - f(x))\, Q(x,y), \quad x \in E,$$

where the transition rates $(Q(x,y))_{x \neq y}$ are non-negative and the function f lies in an algebra, say $\mathcal{A}$, containing the constant functions and which is stable by the action of

$\mathcal{L}$ and by the associated semigroup $(P_t)_{t\geq 0}$, which acts on elements of $\mathcal{A}$ as follows:

$$P_t f(x) := \mathbb{E}_x[f(X_t)] = \sum_{y\in E} f(y) P_t(x,y), \quad x \in E.$$

See for instance [2] for a discussion on the existence of this algebra.
If there exists $V > 0$ such that $\left\| \sum_{y\in E} d(\cdot,y)^2 Q(\cdot,y) \right\|_\infty \leq V^2$, then $\langle X, X\rangle_t \leq V^2 t$ and we say that $(X_t)_{t\geq 0}$ has angle bracket bounded by V^2. Moreover, we say that it has jumps bounded by some positive constant b if $\sup_{t>0} d(X_{t-}, X_t) \leq b$.

2.1 The Wasserstein curvature

Let us introduce the notion of curved Markov chains in the Wasserstein sense.

Definition 2.1 *The d-Wasserstein curvature at time $t > 0$ of the continuous time Markov chain $(X_t)_{t\geq 0}$ is defined by*

$$K_t := -\frac{1}{t} \sup \left\{ \log \left(\frac{\|P_t f\|_{\mathrm{Lip}_d}}{\|f\|_{\mathrm{Lip}_d}} \right) : f \in \mathcal{A} \cap \mathrm{Lip}_d(E),\ f \neq \mathrm{const} \right\} \quad \in [-\infty, +\infty).$$

It is said to be bounded below by $K \in \mathbb{R}$ if $\inf_{t>0} K_t \geq K$.

Remark 2.2 Denote $\mathbb{P}_1(E)$ the space of probability measure μ on E such that $\sum_{y\in E} d(x,y)\mu(y) < +\infty$, $x \in E$. Given $\mu, \nu \in \mathbb{P}_1(E)$, define the d-Wasserstein distance between μ and ν by

$$W_d(\mu,\nu) := \inf_\pi \sum_{x,y\in E} d(x,y)\pi(x,y),$$

where the infimum runs over all $\pi \in \mathbb{P}_1(E \times E)$ with marginals μ and ν, cf. [5, Chapter 5]. The Kantorovich-Rubinstein duality theorem states that

$$W_d(\mu,\nu) = \sup \left\{ \left| \sum_{x\in E} f(x)(\mu(x) - \nu(x)) \right| : \|f\|_{\mathrm{Lip}_d} \leq 1 \right\}.$$

If $P_t(x,\cdot) \in \mathbb{P}_1(E)$, $x \in E$, $t > 0$, then the following assertions are equivalent:

(i) $\inf_{t>0} K_t \geq K$;

(ii) $\|P_t f\|_{\mathrm{Lip}_d} \leq e^{-Kt} \|f\|_{\mathrm{Lip}_d}$, for any $f \in \mathcal{A} \cap \mathrm{Lip}_d(E)$ and any $t > 0$;

(iii) $W_d(P_t(x,\cdot), P_t(y,\cdot)) \le e^{-Kt} d(x,y)$, for any $x, y \in E$ and any $t > 0$.

Hence, these assertions characterize the lower bounds on the d-Wasserstein curvature in terms of contraction properties of the semigroup in the metric W_d, which induces a coupling approach. Note that a version of the item (iii) above was introduced with the trivial metric ϱ in [11] and also through the condition (C1) in [6], in order to establish transportation and Gaussian concentration inequalities for weakly dependent sequences.

2.2 The Γ-curvature

The "carré du champ" operator Γ is defined on $\mathcal{A} \times \mathcal{A}$ by

$$\begin{aligned} \Gamma(f,g)(x) &:= \frac{1}{2}\left(\mathcal{L}(fg)(x) - f(x)\mathcal{L}g(x) - g(x)\mathcal{L}f(x)\right) \\ &= \frac{1}{2}\sum_{y\in E} \left(f(y) - f(x)\right)\left(g(y) - g(x)\right) Q(x,y). \end{aligned}$$

We set $\Gamma f = \Gamma(f,f)$ and introduce the notion of curved Markov chains in the Γ-sense:

Definition 2.3 *The Γ-curvature at time $t > 0$ of the continuous time Markov chain $(X_t)_{t\ge 0}$ is defined by*

$$\rho_t := -\frac{1}{t} \sup \left\{ \log\left(\frac{\left(\Gamma P_t f\right)^{1/2}(x)}{P_t\left(\Gamma f\right)^{1/2}(x)} \right) : f \in \mathcal{A},\ f \neq \text{const},\ x \in E \right\} \quad \in [-\infty, +\infty).$$

It is said to be bounded below by $\rho \in \mathbb{R}$ if $\inf_{t>0} \rho_t \ge \rho$.

Remark 2.4 By definition, the Γ-curvature is bounded below by $\rho \in \mathbb{R}$ if and only if for any $f \in \mathcal{A}$,

$$\left(\Gamma P_t f\right)^{1/2}(x) \le e^{-\rho t} P_t\left(\Gamma f\right)^{1/2}(x), \quad x \in E, \quad t > 0, \tag{2.1}$$

which is the discrete analogue of the commutation relation between local gradient and heat kernel on Riemannian manifolds with Ricci curvature bounded below, see [3].

As already mentioned in the introduction, both curvatures are equivalent in the continuous setting of Brownian motions on Riemannian manifolds, cf. [14]. This is no longer the case in discrete spaces since the discrete gradients do not satisfy in general the chain rule formula, and the curvatures defined above are not comparable.

To finish with the preliminaries, let us make some comments on the deviation inequalities we will establish in the remainder of this paper:

1) Our estimates are given for the distribution of X_t given $X_0 = x$, uniformly in $x \in E$ and for any $t > 0$. Hence, without risk of confusion, the range of validity of the parameters x and t will not be mentioned in our results.

2) In order to relieve the notation, our results are given with the function $u \mapsto u \log(1+u)/2$ in the upper bounds. However, sharper estimates are also available when replacing this function by $u \mapsto (1+u)\log(1+u) - u$, $u \geq 0$.

3 A deviation bound for curved Markov chains in the Wasserstein sense

In this part, we present a Poisson-type deviation estimate under the assumption of a lower bound on the d-Wasserstein curvature.

Theorem 3.1 *Assume that $(X_t)_{t\geq 0}$ has jumps and angle bracket bounded respectively by $b > 0$ and $V^2 > 0$. Suppose moreover that its d-Wasserstein curvature is bounded below by $K \in \mathbb{R}$. Let $f \in \mathrm{Lip}_d(E)$ and define $C_{t,K} := \sup_{0\leq s\leq t} e^{-K(t-s)}$ and $M_{t,K} := (1-e^{-2Kt})/(2K)$ ($M_{t,K} = t$ if $K = 0$). Then for any $y > 0$,*

$$\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) \;\leq\; \exp\left(-\frac{y}{2bC_{t,K}\|f\|_{\mathrm{Lip}_d}} \log\left(1 + \frac{bC_{t,K}y}{M_{t,K}V^2\|f\|_{\mathrm{Lip}_d}}\right)\right). \tag{3.1}$$

Proof. Assume first that f is bounded. The process $\left(Z_s^f\right)_{0\leq s\leq t}$ given by $Z_s^f := P_{t-s}f(X_s) - P_tf(X_0)$ is a real $\mathbb{P}_x$-martingale with respect to the truncated filtration $(\mathscr{F}_s)_{0\leq s\leq t}$ and we have by Itô's formula:

$$Z_s^f \;=\; \sum_{y,z\in E} \int_0^s \left(P_{t-\tau}f(y) - P_{t-\tau}f(z)\right) 1_{\{X_{\tau-}=z\}}(N_{z,y} - \sigma_{z,y})(d\tau),$$

where $(N_{z,y})_{z,y\in E}$ is a family of independent Poisson processes on $\mathbb{R}_+$ with respective intensity $\sigma_{z,y}(dt) = Q(z,y)dt$. Since the d-Wasserstein curvature is bounded below, the jumps of $\left(Z_s^f\right)_{0\le s\le t}$ are bounded for any $s\in[0,t]$:

$$\begin{aligned}\left|Z_s^f - Z_{s-}^f\right| &= |P_{t-s}f(X_s) - P_{t-s}f(X_{s-})| \\ &\le d(X_s, X_{s-})\|f\|_{\mathrm{Lip}_d} C_{t,K} \\ &\le b\|f\|_{\mathrm{Lip}_d} C_{t,K},\end{aligned}$$

as its angle bracket:

$$\begin{aligned}\langle Z^f, Z^f\rangle_s &= \sum_{y,z\in E}\int_0^s \left(P_{t-\tau}f(y) - P_{t-\tau}f(z)\right)^2 1_{\{X_{\tau-}=z\}}\,\sigma_{z,y}(d\tau) \\ &\le \|f\|^2_{\mathrm{Lip}_d}\sum_{y,z\in E}\int_0^s e^{-2K(t-\tau)} d(z,y)^2 1_{\{X_{\tau-}=z\}}\,Q(z,y)d\tau \\ &\le \|f\|^2_{\mathrm{Lip}_d} M_{t,K} V^2.\end{aligned}$$

By [10, Lemma 23.19], for any positive λ, the process $(Y_s^{(\lambda)})_{0\le s\le t}$ given by

$$Y_s^{(\lambda)} := \exp\left\{\lambda Z_s^f - \lambda^2\psi(\lambda b\|f\|_{\mathrm{Lip}_d} C_{t,K})\langle Z^f, Z^f\rangle_s\right\}$$

is a $\mathbb{P}_x$-supermartingale with respect to $(\mathscr{F}_s)_{0\le s\le t}$, where $\psi(z) = z^{-2}\left(e^z - z - 1\right)$, $z > 0$. Thus, we get for any $\lambda > 0$:

$$\begin{aligned}\mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] &= \mathbb{E}_x\left[e^{\lambda Z_t^f}\right] \\ &\le \exp\left\{\lambda^2\|f\|^2_{\mathrm{Lip}_d} M_{t,K} V^2\,\psi(\lambda b\|f\|_{\mathrm{Lip}_d} C_{t,K})\right\}\mathbb{E}_x\left[Y_t^{(\lambda)}\right] \\ &\le \exp\left\{\lambda^2\|f\|^2_{\mathrm{Lip}_d} M_{t,K} V^2\,\psi(\lambda b\|f\|_{\mathrm{Lip}_d} C_{t,K})\right\} \\ &= \exp\left\{\frac{M_{t,K}V^2}{b^2 C_{t,K}^2}\left(e^{\lambda b\|f\|_{\mathrm{Lip}_d} C_{t,K}} - \lambda b\|f\|_{\mathrm{Lip}_d} C_{t,K} - 1\right)\right\}.\end{aligned}$$

Using then Chebychev's inequality and optimizing in $\lambda > 0$ in the exponential estimate above, the deviation inequality (3.1) is established in the bounded case. Finally, the boundedness assumption on f is removed by a classical argument. ■

Remark 3.2 If $K = 0$, then the estimate in Theorem 3.1 is similar to the deviation inequalities of [8, 13] established for infinitely divisible distributions with compactly supported Lévy measure. If $K < 0$, the decay in (3.1) is slower, due to some exponential factors, whereas if $K > 0$, the chain is ergodic, cf. [5, Theorem 5.23], and such an estimate can be extended as $t \to +\infty$ to the stationary distribution, as illustrated in Section 5.2.2. On the other hand, the sign of K has no influence in small time in (3.1).

To conclude this section, note that Theorem 3.1 allows us to consider neither continuous time Markov chains with unbounded angle bracket nor another Lipschitz seminorm than $\|\cdot\|_{\mathrm{Lip}_d}$. To overcome this difficulty, one has to require some assumptions on a different curvature of the chain, namely the Γ-curvature.

4 Estimates for curved Markov chains in the Γ-sense

In this section, we adapt to the Markovian case the covariance method of [8] to derive deviation inequalities for curved continuous time Markov chains in the Γ-sense. Although the Wasserstein and Γ-curvatures are not comparable in discrete spaces, the results we give in this part are more general than Theorem 3.1.

4.1 A general bound

Before turning to Theorem 4.2 below, let us establish the following

Lemma 4.1 *Assume that $(X_t)_{t\geq 0}$ has Γ-curvature bounded below by $\rho \in \mathbb{R}$. Let $g_1, g_2 \in \mathscr{B}(E)$ with $\|\Gamma g_1\|_\infty < +\infty$ and define $L_{t,\rho} = (1-e^{-2\rho t})/(2\rho)$ if $\rho \neq 0$, and $L_{t,\rho} = t$ otherwise. Then we have the covariance inequality*

$$\begin{aligned}\mathrm{Cov}_x\left[g_1(X_t), g_2(X_t)\right] &:= \mathbb{E}_x\left[\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right]\right)\left(g(X_t) - \mathbb{E}_x\left[g(X_t)\right]\right)\right] \\ &\leq 2L_{t,\rho}\|\Gamma g_1\|_\infty^{1/2}\,\mathbb{E}_x\left[(\Gamma g_2)^{1/2}(X_t)\right].\end{aligned}$$

Proof. As in the proof of Theorem 3.1, we have for $i = 1, 2$:

$$g_i(X_t) - \mathbb{E}_x\left[g_i(X_t)\right] = \sum_{y,z\in E}\int_0^t \left(P_{t-s}g_i(y) - P_{t-s}g_i(z)\right) 1_{\{X_{s-}=z\}}(N_{z,y} - \sigma_{z,y})(ds).$$

By the Cauchy-Schwarz inequality,

$$\mathrm{Cov}_x\left[g_1(X_t), g_2(X_t)\right] = 2\int_0^t P_s\left(\Gamma(P_{t-s}g_1, P_{t-s}g_2)\right)(x)\,ds$$

$$
\begin{aligned}
&\leq 2 \int_0^t P_s \left((\Gamma P_{t-s} g_1)^{1/2} (\Gamma P_{t-s} g_2)^{1/2} \right) (x) \, ds \\
&\leq 2 \int_0^t e^{-2\rho(t-s)} P_s \left(P_{t-s} (\Gamma g_1)^{1/2} \, P_{t-s} (\Gamma g_2)^{1/2} \right) (x) \, ds,
\end{aligned}
$$

where in the latter inequality we used the assumption of a lower bound ρ on the Γ-curvature. Since $(P_t)_{t\geq 0}$ is a contraction operator on $\mathscr{B}(E)$, we have

$$
\begin{aligned}
\mathrm{Cov}_x \left[g_1(X_t), g_2(X_t) \right] &\leq 2 \, \|\Gamma g_1\|_\infty^{1/2} \int_0^t e^{-2\rho(t-s)} P_s \left(P_{t-s} (\Gamma g_2)^{1/2} \right) (x) \, ds \\
&= 2 L_{t,\rho} \|\Gamma g_1\|_\infty^{1/2} \, \mathbb{E}_x \left[(\Gamma g_2)^{1/2} (X_t) \right].
\end{aligned}
$$

■

Now, we are able to state Theorem 4.2 which presents a general deviation bound for curved continuous time Markov chains in the Γ-sense:

Theorem 4.2 *Assume that $(X_t)_{t\geq 0}$ has Γ-curvature bounded below by $\rho \in \mathbb{R}$. Let $f \in \mathrm{Lip}_d(E)$ with $\|\Gamma f\|_\infty < +\infty$, and define the function $\psi_{f,t} : \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$ by*

$$
\psi_{f,t}(\lambda) := \sqrt{2} L_{t,\rho} \|\Gamma f\|_\infty^{1/2} \left\| \sum_{y \in E} (f(y) - f(\cdot))^2 \left(\frac{e^{\lambda \, \|f\|_{\mathrm{Lip}_d} d(\cdot, y)} - 1}{\|f\|_{\mathrm{Lip}_d} d(\cdot, y)} \right)^2 Q(\cdot, y) \right\|_\infty^{1/2},
$$

where $L_{t,\rho}$ is defined in Lemma 4.1. If $M_{f,t} := \sup\{\lambda > 0 : \psi_{f,t}(\lambda) < +\infty\}$, then

$$
\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \leq \exp \inf_{\lambda \in (0, M_{f,t})} \int_0^\lambda \left(\psi_{f,t}(\tau) - y \right) d\tau, \quad y > 0. \tag{4.1}
$$

Remark 4.3 Note that $\psi_{f,t}$ is bijective from $(0, M_{f,t})$ to $(0, +\infty)$, so that the term in the exponential is negative and the inequality (4.1) makes sense.

Proof. Using a standard argument, we are reduced to establish the result for f bounded Lipschitz. Applying the covariance inequality of Lemma 4.1 with the functions $g_1(z) = f(z) - \mathbb{E}_x[f(X_t)]$ and $g_2(z) = \exp\{\lambda(f(z) - \mathbb{E}_x[f(X_t)])\}$, $z \in E$, $\lambda \in (0, M_{f,t})$, we have

$$
\begin{aligned}
&\mathbb{E}_x \left[(f(X_t) - \mathbb{E}_x[f(X_t)]) \, e^{\lambda (f(X_t) - \mathbb{E}_x[f(X_t)])} \right] \\
\leq\ & 2 L_{t,\rho} \|\Gamma f\|_\infty^{1/2} \, e^{-\lambda \mathbb{E}_x[f(X_t)]} \mathbb{E}_x \left[\left(\Gamma e^{\lambda f} \right)^{1/2} (X_t) \right]
\end{aligned}
$$

$$\leq \sqrt{2}\, L_{t,\rho}\|\Gamma f\|_\infty^{1/2}\, \mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])} \left(\sum_{y,z\in E} \left(e^{\lambda|f(y)-f(z)|}-1\right)^2 1_{\{X_t=z\}} Q(z,y) \right)^{1/2} \right]$$

$$\leq \psi_{f,t}(\lambda)\mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right].$$

Letting $H_{f,t,x}(\lambda) := \mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right]$, then the latter inequality rewrites as $H'_{f,t,x}(\lambda) \leq \psi_{f,t}(\lambda)H_{f,t,x}(\lambda)$, from which follows the bound

$$\mathbb{E}_x\left[e^{\lambda(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] = H_{f,t,x}(\lambda) \leq e^{\int_0^\lambda \psi_{f,t}(\tau)d\tau}, \quad \lambda \in (0, M_{f,t}).$$

Finally, using Chebychev's inequality, Theorem 4.2 is established. ■

4.2 Some explicit tail estimates

Since the estimate (4.1) is very general, let us make further assumptions on $(X_t)_{t\geq 0}$ to get Poisson-type deviation inequalities. Denote in the sequel $L_{t,\rho} = (1-e^{-2\rho t})/(2\rho)$ if $\rho \neq 0$, and $L_{t,\rho} = t$ otherwise. Using the notation of Theorem 4.2, we have the

Corollary 4.4 *Under the hypothesis of Theorem 4.2, suppose moreover that $(X_t)_{t\geq 0}$ has jumps bounded by $b > 0$. Then for any $y > 0$,*

$$\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) \leq \exp\left(-\frac{y}{2b\|f\|_{\mathrm{Lip}_d}} \log\left(1 + \frac{yb\|f\|_{\mathrm{Lip}_d}}{2L_{t,\rho}\|\Gamma f\|_\infty}\right)\right).$$

Proof. Under the notation of Theorem 4.2, the boundedness of the jumps implies $M_{f,t} = +\infty$, and $\psi_{f,t}$ is bounded by

$$\psi_{f,t}(\lambda) \leq 2L_{t,\rho}\|\Gamma f\|_\infty \frac{e^{\lambda b\|f\|_{\mathrm{Lip}_d}}-1}{b\|f\|_{\mathrm{Lip}_d}}, \quad \lambda > 0.$$

Using then Theorem 4.2 and optimizing in $\lambda > 0$, the proof is achieved. ■

Note that the latter deviation inequality is more general than (3.1), since the finiteness assumption on $\|\Gamma f\|_\infty$ allows us to relax the boundedness assumption on the angle bracket. Thus, when the angle bracket of $(X_t)_{t\geq 0}$ is bounded, the next corollary exhibits an estimate comparable to that of Theorem 3.1:

Corollary 4.5 *Assume that* $(X_t)_{t\geq 0}$ *has jumps and angle bracket bounded respectively by* $b > 0$ *and* $V^2 > 0$*. Suppose moreover that its* Γ*-curvature is bounded below by* $\rho \in \mathbb{R}$*, and let* $f \in \mathrm{Lip}_d(E)$*. Then for any* $y > 0$*,*

$$\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \leq \exp \left(-\frac{y}{2b \|f\|_{\mathrm{Lip}_d}} \log \left(1 + \frac{by}{L_{t,\rho} V^2 \|f\|_{\mathrm{Lip}_d}} \right) \right).$$

Proof. By the boundedness of the jumps and of the angle bracket, the function $\psi_{f,t}$ in Theorem 4.2 is bounded by

$$\psi_{f,t}(\lambda) \leq L_{t,\rho} V^2 \|f\|_{\mathrm{Lip}_d} \frac{e^{\lambda b \|f\|_{\mathrm{Lip}_d}} - 1}{b}, \quad \lambda > 0.$$

Finally, applying Theorem 4.2 yields the result. ∎

5 The case of birth-death processes

In the paper [1], some deviation inequalities are established for continuous time random walks on graphs. Such processes may be seen as models in null curvature since the transition rates of the associated generator do not depend on the space variable. Using the results of Sections 3 and 4, the purpose of this part is to extend these tail estimates to birth-death processes whose discrete curvatures are bounded below.

Let $(X_t)_{t\geq 0}$ be a birth-death process on the state space $E = \mathbb{N}$ or $E = \{0, 1, \ldots, n\}$, with stationary distribution π. It is a regular continuous time Markov chain with generator defined on $\mathscr{F}(E)$ by

$$\mathcal{L}f(x) = \lambda_x \left(f(x+1) - f(x) \right) + \nu_x \left(f(x-1) - f(x) \right), \quad x \in E, \tag{5.1}$$

where the transition rates λ and ν are positive with 0 as reflecting state, i.e. $\nu_0 = 0$ (if $E = \{0, 1, \ldots, n\}$, the state n is also reflecting: $\lambda_n = 0$), ensuring irreducibility. Denote $(P_t)_{t\geq 0}$ the homogeneous semigroup whose transition probabilities are given for any $x \in E$ by

$$P_t(x,y) = \begin{cases} \lambda_x t + o(t) & \text{if } y = x+1, \\ \nu_x t + o(t) & \text{if } y = x-1, \\ 1 - (\lambda_x + \nu_x)t + o(t) & \text{if } y = x, \\ 0 & \text{if } y \in E \setminus \{x-1, x, x+1\}, \end{cases}$$

where the function o is such that $o(t)/t$ converges to 0 as t tends to 0. Since P_t is a contraction on $L^1(\pi)$ for any $t \geq 0$, this operator is well-defined on the space $\mathrm{Lip}_d(E)$ of Lipschitz functions on E with respect to the classical distance $d(x,y) := |x-y|$, $x, y \in E$, provided the stationary distribution π satisfies the moment condition $\sum_{y\in E} y\pi(y) < +\infty$, that we suppose in the sequel.

5.1 Criteria for lower bounded curvatures

Let us give some criteria on the generator of the process $(X_t)_{t\geq 0}$ which ensure that the different discrete curvatures are bounded below.

Proposition 5.1 *Assume that there exists a real number K such that*

$$\inf_{x\in E\setminus\{0\}} \lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} \geq K. \tag{5.2}$$

Then the d-Wasserstein curvature of the process $(X_t)_{t\geq 0}$ is bounded below by K.

Remark 5.2 If $E = \mathbb{N}$ and the transition rates of the generator are bounded and satisfy the assumptions of Proposition 5.1, then necessarily $K \leq 0$.

Proof. Let us establish the result via coupling methods. Consider $(X_t^x)_{t\geq 0}$ and $(X_t^y)_{t\geq 0}$ two independent copies of $(X_t)_{t\geq 0}$, starting respectively from x and y. Then the generator $\tilde{\mathcal{L}}$ of the process $(X_t^x, X_t^y)_{t\geq 0}$ is given for any $f \in \mathscr{F}(E\times E)$ by

$$\tilde{\mathcal{L}} f(z,w) = (\mathcal{L} f(\cdot, w))(z) + (\mathcal{L} f(z,\cdot))(w), \quad z, w \in E.$$

Since the transition rates of the generator satisfy (5.2), we have immediately the bound $\tilde{\mathcal{L}} d(z, z+1) \leq -K$, $z \in E$, which is equivalent to the inequality $\tilde{\mathcal{L}} d(z,w) \leq -K d(z,w)$ for any $z, w \in E$. Therefore, we obtain the estimate $\mathbb{E}\left[d(X_t^x, X_t^y)\right] \leq e^{-Kt} d(x,y)$ which in turn implies

$$W_d(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-Kt} d(x,y).$$

Finally, by the equivalent statements of Remark 2.2, the d-Wasserstein curvature of $(X_t)_{t\geq 0}$ is bounded below by K. ∎

In order to establish modified logarithmic Sobolev inequalities for continuous time random walks on $\mathbb{Z}$, the authors in [1] used a suitable Γ_2-calculus to give a criterion under which the Γ-curvature is bounded below by 0. Actually, this criterion can be generalized to any real lower bound on the Γ-curvature via Lemma 5.3 below.
Define the Γ_2-operator on $\mathscr{F}(E)$ by

$$\Gamma_2 f(x) := \frac{1}{2}\left(\mathcal{L}\Gamma f(x) - 2\Gamma(f, \mathcal{L}f)(x)\right), \quad x \in E.$$

By adapting the proof in [1] mentioned above, we get the

Lemma 5.3 *Assume that there exists $\rho \in \mathbb{R}$ such that the inequality*

$$\Gamma_2 f(x) - \Gamma\left(\Gamma f\right)^{1/2}(x) \geq \rho \Gamma f(x), \quad x \in E, \tag{5.3}$$

is satisfied for any $f \in \mathscr{F}(E)$. Then $(X_t)_{t\geq 0}$ has Γ-curvature bounded below by ρ.

Remark 5.4 Differentiating in a neighborhood of 0_+ the function $t \mapsto e^{-\rho t} P_t (\Gamma f)^{1/2} - (\Gamma P_t f)^{1/2}$ shows that the equivalence holds in Lemma 5.3.

Proposition 5.5 *Assume that the transition rates λ and ν are respectively non-increasing and non-decreasing and that there exists some $\rho \geq 0$ such that*

$$\inf_{x \in E \setminus \{0, \sup E\}} \min\{\lambda_{x-1} - \lambda_x, \nu_{x+1} - \nu_x\} \geq \rho. \tag{5.4}$$

Then the Γ-curvature of $(X_t)_{t\geq 0}$ is bounded below by ρ.

Remark 5.6 If $E = \mathbb{N}$ and the transition rates of the generator satisfy the assumptions of Proposition 5.5, then necessarily $\rho = 0$.

Proof. By Lemma 5.3, the result holds true if the Γ_2-inequality (5.3) above is satisfied, that we prove now. Fix $x \in E$ and let $a = f(x) - f(x+1)$, $b = f(x) - f(x-1)$, $c = f(x+2) - f(x+1)$ and $d = f(x-2) - f(x-1)$. We have

$$2\Gamma_2 f(x) - 2\Gamma\left(\Gamma f\right)^{1/2}(x) = \lambda_x(\nu_{x+1} - \nu_x)a^2 + \nu_x(\lambda_{x-1} - \lambda_x)b^2 + I(x) + J(x),$$

where

$$I(x) := \lambda_x \lambda_{x+1} ac + \lambda_x \nu_x ab + \lambda_x \left(\lambda_{x+1}c^2 + \nu_{x+1}a^2\right)^{1/2} \left(\lambda_x a^2 + \nu_x b^2\right)^{1/2},$$

$$J(x) := \nu_x\nu_{x-1}bd + \lambda_x\nu_x ab + \nu_x\left(\lambda_{x-1}b^2 + \nu_{x-1}d^2\right)^{1/2}\left(\lambda_x a^2 + \nu_x b^2\right)^{1/2}.$$

Since the transition rates λ and ν are respectively non-increasing and non-decreasing and satisfy furthermore (5.4), we get

$$2\Gamma_2 f(x) - 2\Gamma\left(\Gamma f\right)^{1/2}(x) \geq 2\rho\Gamma f(x) + I(x) + J(x).$$

By symmetry with the function J, it is sufficient to establish $I \geq 0$. We have

$$\begin{aligned} I(x) &\geq \lambda_x\left(\lambda_{x+1}c^2 + \nu_{x+1}a^2\right)^{1/2}\left(\lambda_x a^2 + \nu_x b^2\right)^{1/2} - \lambda_x\lambda_{x+1}|ac| - \lambda_x\nu_x|ab| \\ &= \lambda_x\left(I_1(x) - I_2(x)\right), \end{aligned}$$

where

$$I_1(x) := \left(\lambda_{x+1}c^2 + \nu_{x+1}a^2\right)^{1/2}\left(\lambda_x a^2 + \nu_x b^2\right)^{1/2} \quad \text{and} \quad I_2(x) := \lambda_{x+1}|ac| + \nu_x|ab|.$$

Using again the monotonic assumptions on the transition rates of the generator,

$$\begin{aligned} &\left(I_1(x)\right)^2 - \left(I_2(x)\right)^2 \\ &\quad = \lambda_{x+1}(\lambda_x - \lambda_{x+1})a^2c^2 + \nu_x(\nu_{x+1} - \nu_x)a^2b^2 + \lambda_x\nu_{x+1}a^4 + \lambda_{x+1}\nu_x b^2c^2 - 2\nu_x\lambda_{x+1}a^2bc \\ &\quad \geq \nu_x\lambda_{x+1}(a^2 - bc)^2 \geq 0. \end{aligned}$$

■

5.2 Applications

The proofs of the following results are omitted since they are immediate applications of Theorems 3.1 and 4.2, once the assumptions of Propositions 5.1 and 5.5 are satisfied, respectively.

5.2.1 The case $E = \mathbb{N}$

Corollary 5.7 *Assume that $\lambda, \nu \in \mathscr{B}(\mathbb{N})$ and that there exists $K \leq 0$ such that $\inf_{x\in\mathbb{N}\setminus\{0\}} \lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} \geq K$. Let $f \in \mathrm{Lip}_d(\mathbb{N})$. Then for any $y > 0$,*

$$\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) \leq \exp\left(-\frac{ye^{tK}}{2\|f\|_{\mathrm{Lip}_d}}\log\left(1 + \frac{yK}{\sinh(tK)\|\lambda + \nu\|_\infty\|f\|_{\mathrm{Lip}_d}}\right)\right).$$

If $K = 0$, then replace the latter inequality by its limit as $K \to 0$.

In Corollary 5.8 below, no particular boundedness assumption is made on the transition rates of the generator of the birth-death process $(X_t)_{t\geq 0}$.

Corollary 5.8 *Assume that λ and ν are respectively non-increasing and non-decreasing. Let $f \in \mathrm{Lip}_d(\mathbb{N})$ with furthermore $\|\Gamma f\|_\infty < +\infty$. Then for any $y > 0$,*

$$\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \quad \leq \quad \exp \left(- \frac{y}{2\|f\|_{\mathrm{Lip}_d}} \log \left(1 + \frac{y\|f\|_{\mathrm{Lip}_d}}{2t\|\Gamma f\|_\infty} \right) \right).$$

Remark 5.9 In order to obtain deviation bounds for stationary distributions, the positivity of lower bounds of discrete curvatures is crucial and thus does not allow us to extend such estimates to birth-death processes on the infinite state space $E = \mathbb{N}$, see the Remarks 5.2 and 5.6. In particular, it excludes the $M/M/\infty$ queueing process recently investigated in [4] and whose stationary distribution is the Poisson measure on $\mathbb{N}$. Therefore, we expect to recover the classical deviation inequality satisfied by the Poisson distribution by taking the limit as $t \to +\infty$ in an appropriate deviation estimate, and such an interesting problem will be addressed in a subsequent paper. Note also that Corollary 5.8 applies for such a process, but it does not reflect its positive curvature emphasized in [4].

5.2.2 The case $E = \{0, 1, \dots, n\}$

The purpose of this part is to refine Corollaries 5.7 and 5.8 when the state space is finite, in order to establish by a limiting argument Poisson-type deviation estimates for the stationary distribution π. To do so, the crucial point is to obtain positive lower bounds on the discrete curvatures.
Our estimates below may be compared to that of [9, Proposition 4] established under reversibility assumptions and without notion of discrete curvatures.

Corollary 5.10 *Assume that there exists $K > 0$ such that $\min_{x\in\{1,\dots,n\}} \lambda_{x-1} - \lambda_x + \nu_x - \nu_{x-1} \geq K$, and let $f \in \mathrm{Lip}_d(\{0, 1, \dots, n\})$. Then for any $y > 0$,*

$$\mathbb{P}_x \left(f(X_t) - \mathbb{E}_x \left[f(X_t) \right] \geq y \right) \quad \leq \quad \exp \left(- \frac{y}{2\|f\|_{\mathrm{Lip}_d}} \log \left(1 + \frac{2Ky}{(1 - e^{-2Kt})\|\lambda + \nu\|_\infty \|f\|_{\mathrm{Lip}_d}} \right) \right).$$

In particular, letting $t \to +\infty$ above yields under the stationary distribution π:

$$\pi \left(f - \mathbb{E}_\pi[f] \geq y \right) \leq \exp \left(- \frac{y}{2\|f\|_{\mathrm{Lip}_d}} \log \left(1 + \frac{2Ky}{\|\lambda + \nu\|_\infty \|f\|_{\mathrm{Lip}_d}} \right) \right).$$

Under different assumptions on the generator, we get a somewhat similar estimate:

Corollary 5.11 *Assume that the transition rates* λ *and* ν *satisfy* $\min_{x\in\{1,\ldots,n-1\}}\min\{\lambda_{x-1}-\lambda_x, \nu_{x+1}-\nu_x\} \geq \rho > 0$*, and let* $f \in \mathrm{Lip}_d(\{0,1,\ldots,n\})$*. Then for any* $y > 0$*,*

$$\mathbb{P}_x\left(f(X_t) - \mathbb{E}_x\left[f(X_t)\right] \geq y\right) \;\leq\; \exp\left(-\frac{y}{2\|f\|_{\mathrm{Lip}_d}}\log\left(1+\frac{2\rho y}{(1-e^{-2\rho t})(\lambda_0+\nu_n)\|f\|_{\mathrm{Lip}_d}}\right)\right).$$

As $t \to +\infty$*, we obtain the deviation inequality*

$$\pi\left(f - \mathbb{E}_\pi[f] \geq y\right) \leq \exp\left(-\frac{y}{2\|f\|_{\mathrm{Lip}_d}}\log\left(1+\frac{2\rho y}{(\lambda_0+\nu_n)\|f\|_{\mathrm{Lip}_d}}\right)\right).$$

As an application of Corollary 5.10, let us recover the classical Gaussian deviation inequality for a Brownian driven Ornstein-Uhlenbeck process constructed as a fluid limit of rescaled continuous time Ehrenfest chains.

Corollary 5.12 *Let* $(U_t)_{t\geq 0}$ *be the Brownian driven Ornstein-Uhlenbeck process given by*

$$U_t = z_0 e^{-t} + \sqrt{2\lambda\nu}\int_0^t e^{-(t-s)}dB_s, \quad t > 0,$$

where $z_0 \in \mathbb{R}$ *and the positive parameters* λ *and* ν *are such that* $\lambda+\nu = 1$*. Then for any Lipschitz function* f *on* $\mathbb{R}$ *with Lipschitz constant* $\|f\|_{\mathrm{Lip}}$*, we recover the classical Gaussian deviation inequality*

$$\mathbb{P}_{z_0}\left(f(U_t) - \mathbb{E}_{z_0}\left[f(U_t)\right] \geq y\right) \leq \exp\left(-\frac{y^2}{(1-e^{-2t})\nu\|f\|^2_{\mathrm{Lip}}}\right), \quad y > 0.$$

Proof. Let $(X^n_t)_{t\geq 0}$ be the continuous time Ehrenfest chain on $\{0,1,\ldots,n\}$ starting from some $x_n \in \{0,1,\ldots,n\}$ and with generator given by

$$\mathcal{L}_n f(x) = \lambda(n-x)\left(f(x+1)-f(x)\right) + \nu x\left(f(x-1)-f(x)\right), \quad x \in \{0,1,\ldots,n\}.$$

Suppose that $\lim_{n\to+\infty} x_n/n = \lambda$ and define the process $(Z^n_t)_{t\geq 0}$ by $Z^n_t = (X^n_t - \lambda n)/\sqrt{n}$, $t > 0$. Assume furthermore that the sequence of initial states $(Z^n_0)_{n\in\mathbb{N}}$ converges to z_0. By the central limit theorem in [7, Chapter 11], the sequence $(Z^n_t)_{t\geq 0}$ converges as n goes to infinity to the process $(U_t)_{t\geq 0}$.

Now, fix $n \in \mathbb{N}\backslash\{0\}$, $t > 0$, and consider the function $h_n = f\circ\phi_n$, where ϕ_n is defined on $\{0,1,\ldots,n\}$ by $\phi_n(x) = (x-n\lambda)/\sqrt{n}$. Then $h_n \in \mathrm{Lip}_d(\{0,1,\ldots,n\})$ with constant at most $n^{-1/2}\|f\|_{\mathrm{Lip}}$. Therefore we can apply Corollary 5.10 to $(X_t^n)_{t\geq 0}$ and h_n, with $K=\lambda+\nu=1$, to get for any $y>0$:

$$\mathbb{P}_{x_n}\left(h_n(X_t^n)-\mathbb{E}_{x_n}\left[h_n(X_t^n)\right]\geq y\right) \leq \exp\left(-\frac{y\sqrt{n}}{2\|f\|_{\mathrm{Lip}}}\log\left(1+\frac{2y}{(1-e^{-2t})\sqrt{n}\nu\|f\|_{\mathrm{Lip}}}\right)\right).$$

Finally, letting n going to infinity in the above inequality yields the result. ■

5.2.3 A multidimensional deviation inequality for the $M/M/1$ queue

In this final part, we give a Poisson-type deviation estimate for sample vectors of the $M/M/1$ queueing process. It is an irreducible birth-death process $(X_t)_{t\geq 0}$ whose generator is given by

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

where λ and ν are positive. The existence of an integration by parts formula for the associated semigroup, together with a tensorization procedure of the Laplace transform of a Lipschitz function of the process, allow us to provide with Corollary 5.13 below a multidimensional deviation inequality for the $M/M/1$ queue.
We say in the sequel that a function $f:\mathbb{N}^n\to\mathbb{R}$ is ℓ^1-Lipschitz if

$$\|f\|_{\mathrm{Lip}(n)} := \sup_{x\neq y}\frac{|f(x)-f(y)|}{\|x-y\|_1} < +\infty,$$

where $\|\cdot\|_1$ denotes the ℓ^1-norm $\|z\|_1=\sum_{i=1}^n |z_i|$, $z\in\mathbb{N}^n$.

Corollary 5.13 *Consider the sample vector $X^n=(X_{t_1},\ldots,X_{t_n})$, $0=t_0<t_1<\cdots<t_n=T$, and let f be ℓ^1-Lipschitz on $\mathbb{N}^n$. Then for any $y>0$,*

$$\mathbb{P}_x\left(f(X^n)-\mathbb{E}_x[f(X^n)]\geq y\right) \leq \exp\left(-\frac{y}{2n\|f\|_{\mathrm{Lip}(n)}}\log\left(1+\frac{y}{Tn(\lambda+\nu)\|f\|_{\mathrm{Lip}(n)}}\right)\right). \tag{5.5}$$

Proof. Let u be a Lipschitz function on $\mathbb{N}$ with Lipschitz constant $\|u\|_{\mathrm{Lip}(1)}$ and let $t>0$. Rewriting the proof of Theorem 4.2 for the $M/M/1$ queue yields for any $\tau>0$:

$$\mathbb{E}_x\left[e^{\tau u(X_t)}\right] \leq \exp\left\{\tau\mathbb{E}_x[u(X_t)]+h(\tau,t,\|u\|_{\mathrm{Lip}(1)})\right\}, \tag{5.6}$$

where h is the function defined on $(\mathbb{R}_+)^3$ by $h(\tau,t,z) = t(\lambda+\nu)\left(e^{\tau z} - \tau z - 1\right)$.
To obtain a multidimensional version of (5.6), the idea is to tensorize with respect to the ℓ^1-metric the Laplace transform via an integration by parts formula satisfied by the semigroup $(P_t)_{t\geq 0}$ of the $M/M/1$ queueing process. We sketch now the argument for the dimension $n=2$. Let $0<s<t$ and denote f_y the function $f_y(z) = f(y,z)$ and $f_1(y) = \sum_{z\in\mathbb{N}} f(y,z)P_{t-s}(y,z)$. By the Markov property together with (5.6), we have

$$\begin{aligned}
\mathbb{E}_x\left[\exp\left(\tau f(X_s,X_t)\right)\right] &= \sum_{y,z\in\mathbb{N}} \exp\left(\tau f_y(z)\right) P_{t-s}(y,z)P_s(x,y) \\
&\leq \sum_{y\in\mathbb{N}} \exp\left(\tau \sum_{z\in\mathbb{N}} f_y(z)P_{t-s}(y,z) + h(\tau, t-s, \|f_y\|_{\mathrm{Lip}(1)})\right) P_s(x,y) \\
&\leq \exp\left\{h(\tau,t-s,\|f\|_{\mathrm{Lip}(2)})\right\} \sum_{y\in\mathbb{N}} \exp\left(\tau f_1(y)\right) P_s(x,y) \\
&\leq \exp\left\{h(\tau,t-s,2\|f\|_{\mathrm{Lip}(2)}) + h(\tau,s,\|f_1\|_{\mathrm{Lip}(1)}) + \tau\mathbb{E}_x[f(X_s,X_t)]\right\},
\end{aligned} \tag{5.7}$$

since the function $z\mapsto h(\cdot,\cdot,z)$ is non-decreasing on $\mathbb{R}_+$. Now, let us bound $\|f_1\|_{\mathrm{Lip}(1)}$ by $2\|f\|_{\mathrm{Lip}(2)}$. To do so, observe that the commutation relation $\mathcal{L}\mathrm{d}^+ = \mathrm{d}^+\mathcal{L}$ holds, where d^+ is the forward gradient $\mathrm{d}^+f(x) = f(x+1) - f(x)$, $x\in\mathbb{N}$. It implies $P_t\mathrm{d}^+ = \mathrm{d}^+P_t$, $t>0$, which in turn entails the integration by parts formula

$$\sum_{y\in\mathbb{N}} u(y)P_t(x+1,y) = \sum_{y\in\mathbb{N}} u(y+1)P_t(x,y), \quad x\in\mathbb{N}.$$

Thus, we have

$$\begin{aligned}
\|f_1\|_{\mathrm{Lip}(1)} &= \sup_{y\in\mathbb{N}} |f_1(y+1) - f_1(y)| \\
&= \sup_{y\in\mathbb{N}} \left| \sum_{z\in\mathbb{N}} f(y+1,z)P_{t-s}(y+1,z) - \sum_{z\in\mathbb{N}} f(y,z)P_{t-s}(y,z) \right| \\
&= \sup_{y\in\mathbb{N}} \left| \sum_{z\in\mathbb{N}} \left(f(y+1,z+1) - f(y,z)\right) P_{t-s}(y,z) \right| \\
&\leq 2\|f\|_{\mathrm{Lip}(2)}.
\end{aligned}$$

Therefore, plugging this into (5.7) entails

$$\mathbb{E}_x\left[\exp\left(\tau f(X_s,X_t)\right)\right] \leq \exp\left\{t(\lambda+\nu)\left(e^{2\tau\|f\|_{\mathrm{Lip}(2)}} - 2\tau\|f\|_{\mathrm{Lip}(2)} - 1\right) + \tau\mathbb{E}_x[f(X_s,X_t)]\right\}.$$

In the general case, we show similarly that the function f_i defined on $\mathbb{N}^i$ by

$$f_i(x_1,\ldots,x_i) := \sum_{x_{i+1},\ldots,x_n\in\mathbb{N}} f(x_1,\ldots,x_i,\ldots,x_n) P_{t_{i+1}-t_i}(x_i,x_{i+1})\cdots P_{t_n-t_{n-1}}(x_{n-1},x_n),$$

has Lipschitz seminorm (with respect to the i^{th} variable) smaller than $(n-i+1)\|f\|_{\mathrm{Lip}(n)}$, and thus than $n\|f\|_{\mathrm{Lip}(n)}$. Therefore, since h is non-decreasing in its third variable, we obtain by using recursively (5.7):

$$\begin{aligned}\mathbb{E}_x\left[e^{\tau f(X^n)}\right] &\leq \exp\left(\tau\mathbb{E}_x[f(X^n)] + \sum_{i=1}^n h(\tau, t_i - t_{i-1}, n\|f\|_{\mathrm{Lip}(n)})\right) \\ &= \exp\left\{\tau\mathbb{E}_x[f(X^n)] + T(\lambda+\nu)\left(e^{\tau n\|f\|_{\mathrm{Lip}(n)}} - \tau n\|f\|_{\mathrm{Lip}(n)} - 1\right)\right\}.\end{aligned}$$

Dividing by $e^{\tau\mathbb{E}_x[f(X^n)]}$ and using Chebychev's inequality achieves the proof. ■

Remark 5.14 To conclude this work, note that Corollary 5.13 does not allow us to extend the deviation inequality (5.5) to functionals on path spaces. Thus, it would be an interesting project to refine suitably such an estimate in terms of the increments $\Delta_i = t_i - t_{i-1}$, as $\Delta_i \to 0$.

References

[1] C. Ané and M. Ledoux. On logarithmic Sobolev inequalities for continuous time random walks on graphs. *Probab. Theory Related Fields*, 116(4):573–602, 2000.

[2] D. Bakry. On Sobolev and logarithmic Sobolev inequalities for Markov semi-groups. In *New trends in stochastic analysis (Charingworth, 1994)*, pages 43–75. World Sci. Publishing, River Edge, NJ, 1997.

[3] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

[4] D. Chafaï. Binomial-Poisson entropic inequalities and the $M/M/\infty$ queue. *ESAIM Probab. Stat.*, 10:317–339, 2006.

[5] M.F. Chen. *From Markov chains to non-equilibrium particle systems.* World Scientific Publishing Co. Inc., River Edge, NJ, second edition, 2004.

[6] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, 32(3B):2702–2732, 2004.

[7] S.N. Ethier and T.G. Kurtz. *Markov processes, Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[8] C. Houdré. Remarks on deviation inequalities for functions of infinitely divisible random vectors. *Ann. Probab.*, 30(3):1223–1237, 2002.

[9] C. Houdré and P. Tetali. Concentration of measure for products of Markov kernels and graph products via functional inequalities. *Combin. Probab. Comput.*, 10(1):1–28, 2001.

[10] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, 1997.

[11] K. Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.*, 6(3):556–571, 1996.

[12] P.M. Samson. Concentration of measure inequalities for Markov chains and Φ-mixing processes. *Ann. Probab.*, 28(1):416–461, 2000.

[13] M. Schmuckenschläger. Martingales, Poincaré type inequalities, and deviation inequalities. *J. Funct. Anal.*, 155(2):303–323, 1998.

[14] K.T. Sturm and M.K. Von Renesse. Transport inequalities, gradient estimates, entropy, and Ricci curvature. *Comm. Pure Appl. Math.*, 58(7):923–940, 2005.