A new Poisson-type deviation inequality for the empirical distribution of ergodic birth-death processes

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

October 13, 2006

Abstract

In this paper, we present a new Poisson-type deviation inequality from equilibrium of the empirical distribution of an ergodic birth-death process, which generalizes the results of Lezaud [9, 10]. Our approach relies on the notion recently developed in [7] of Wasserstein curvatures of the process with respect to a suitable distance on $\mathbb{N}$.

Key words: birth-death process, deviation inequality, Wasserstein curvature, empirical distribution, Markov semigroup, $M/M/\infty$ queueing process.
Mathematics Subject Classification. 60E15, 60J27, 47D07, 41A25.

1 Introduction

Let $(X_t)_{t\geq 0}$ be an ergodic Markov process on a Polish state space E with stationary distribution π. The weak law of large numbers asserts that for any function $\phi \in L^1(\pi)$ and any $y > 0$, the probability

$$\Lambda(t, \phi, x, y) := \mathbb{P}_x \left(\left| \frac{1}{t} \int_0^t \phi(X_s) ds - \int \phi d\pi \right| \geq y \right), \tag{1.1}$$

*ajoulin@u-paris10.fr

tends to 0 as t goes to infinity. Actually, large deviations theory gives asymptotic bounds on the quantity $t^{-1} \log \Lambda(t, \phi, x, y)$, but it is unsatisfactory when one wants to control the probability (1.1) for fixed parameters. In recent years, such a problem has been investigated by several authors. For instance, using the Lumer-Philips theorem, Wu derived in [12] a non-asymptotic estimate on the probability (1.1), which is sharp for symmetric semigroups, however not really tractable. More recently, various authors obtained for diffusion processes explicit upper bounds on (1.1) under regularity assumptions on the function ϕ together with some functional inequalities satisfied by the stationary distribution π, see for instance [1, 5, 6]. On the other hand, in the case of continuous time Markov chains having a spectral gap, Lezaud established Poisson-type deviation bounds for bounded functions ϕ on finite or infinite countable state spaces, whose proofs rely on Kato's perturbation theory for linear operators, cf. [9, 10].

The purpose of the present paper is to extend in two ways the estimates provided in the articles [9, 10] in the case of ergodic birth-death processes on $\mathbb{N}$. The first improvement is to relax the boundedness assumption on the transition rates of the generator of the process, whereas the second one allows us to consider not only bounded but Lipschitz functions ϕ with respect to a suitable distance. Our approach relies on the notion of Wasserstein curvatures recently investigated by the author in [7], which characterize exponential contraction properties of the associated semigroup on the space of Lipschitz functions with respect to this metric. As a result, we establish a tail estimate which is convenient for large time, in contrast to the bounds obtained in [7] involving the classical distance on $\mathbb{N}$.

The paper is organized as follows. In Section 2, basic material on birth-death processes is recalled, as the discrete curvatures defined in [7]. Namely, given an ergodic birth-death process $(X_t)_{t \geq 0}$ on $\mathbb{N}$, we introduce its Wasserstein curvature associated to a suitable metric δ on $\mathbb{N}$ and we provide in Proposition 2.6 some conditions on its generator for this discrete curvature to be bounded below by a positive constant. Under these criteria, we state in the second part of this section our main contribution of the paper which is contained in Theorem 2.7, where a Poisson-type upper bound on a deviation probability similar to (1.1) is established for a (not necessarily bounded) Lipschitz function ϕ with respect to the distance δ. In particular, no boundedness assumption on the transition rates of the generator is required. The whole Section 3 is devoted to the proof of Theorem 2.7, which is rather technical and is divided into several lemma. The main part is given in Lemma 3.3, where an upper bound on the Laplace transform of a Lipschitz cylindrical function of the process $(X_t)_{t\geq 0}$ with respect to the ℓ^1-metric generated by δ, is provided through a tensorization procedure of the one-dimensional case. Finally, the example of the $M/M/\infty$ queueing process is investigated in Section 4.

2 Preliminaries and main result

On a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$, we consider throughout the paper an irreducible ergodic birth-death process $(X_t)_{t\geq 0}$ on the infinite state space $\mathbb{N} := \{0, 1, \ldots\}$, with stationary distribution π. The process $(X_t)_{t\geq 0}$ is a stable conservative continuous time Markov chain with generator defined on the set $\mathscr{F}(\mathbb{N})$ of all real-valued functions on $\mathbb{N}$ by

$$\mathcal{L}f(x) = \lambda_x \left(f(x+1) - f(x)\right) + \nu_x \left(f(x-1) - f(x)\right), \quad x \in \mathbb{N}, \tag{2.1}$$

where the transition rates λ and ν are positive with 0 as the unique reflecting state, i.e. $\nu_0 = 0$, conditions ensuring irreducibility. Denote $(P_t)_{t\geq 0}$ the homogeneous semigroup operator whose transition probabilities are given for any $x \in \mathbb{N}$ by

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

where the function o is such that the ratio $o(t)/t$ converges to 0 as t tends to 0. Denoting $\mathbb{P}_x$ the distribution of the process starting from $x \in \mathbb{N}$ and $\mathbb{E}_x$ the expectation with respect to $\mathbb{P}_x$, the family of operators

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

are positivity preserving contractions on $L^1(\pi)$ and also on $L^\infty(\pi)$, hence by interpolation on every space $L^p(\pi)$, $p \in [1, +\infty]$. In particular, if the stationary distribution satisfies for some metric ρ the moment condition

$$\sum_{x \in \mathbb{N}} \rho(x,y)\pi(x) < +\infty, \quad y \in \mathbb{N}, \tag{2.2}$$

then the semigroup is well-defined on the space Lip_ρ of Lipschitz function $f : \mathbb{N} \to \mathbb{R}$ endowed with the Lipschitz seminorm

$$\|f\|_{\mathrm{Lip}_\rho} := \sup_{x,y \in \mathbb{N}} \frac{|f(x) - f(y)|}{\rho(x,y)} < +\infty.$$

Let us recall the definition given in [7] of the Wasserstein curvature of the birth-death process $(X_t)_{t \geq 0}$ with respect to some distance ρ.

Definition 2.1. *We assume that the stationary distribution π satisfies the moment condition (2.2) with a metric ρ. The ρ-Wasserstein curvature at time $t > 0$ of the process $(X_t)_{t\geq 0}$ is defined by*

$$\alpha_t := -\frac{1}{t} \sup \left\{ \log \left(\frac{\|P_t f\|_{\mathrm{Lip}_\rho}}{\|f\|_{\mathrm{Lip}_\rho}} \right) : f \in \mathrm{Lip}_\rho,\ f \neq \mathrm{const} \right\} \quad \in [-\infty, +\infty).$$

It is said to be bounded below by $\alpha \in \mathbb{R}$ if $\inf_{t>0} \alpha_t \geq \alpha$. In other words, the semigroup $(P_t)_{t\geq 0}$ is exponentially contractive in the following sense:

$$\|P_t f\|_{\mathrm{Lip}_\rho} \leq e^{-\alpha t} \|f\|_{\mathrm{Lip}_\rho}, \quad t > 0.$$

If the stationary distribution π satisfies the moment condition (2.2) with the distance ρ, then by the Kantorovich-Rubinstein duality theorem, see e.g. Theorem 5.10 in [4], the ρ-Wasserstein curvature of the process is bounded below by α if and only if for any $x \in \mathbb{N}$ and any $t > 0$, the kernel $P_t(x, \cdot)$ verifies the moment condition (2.2) with the metric ρ and

$$W_\rho(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-\alpha t} \rho(x,y), \quad x, y \in \mathbb{N}, \quad t > 0.$$

Here, $W_\rho(\mu, \nu)$ stands for the Wasserstein distance between two probability measures μ and ν on $\mathbb{N}$, endowed with the cost function ρ:

$$W_\rho(\mu,\nu) := \inf_\eta \sum_{x,y \in \mathbb{N}} \rho(x,y)\eta(x,y),$$

where the infimum runs over any probability measure η on $\mathbb{N}^2$ having marginals μ and ν. Therefore, if α is positive, then the process has good ergodicity properties since the semigroup $(P_t)_{t\geq 0}$ converges exponentially fast to the stationary distribution π with respect to the metric W_ρ, cf. Theorem 5.23 in [4].

In the paper [7], some Poisson-type deviation inequalities are established for birth-death processes through the discrete curvatures approach and with the classical distance on $\mathbb{N}$ given by

$$d(x,y) = |x-y|, \quad x,y \in \mathbb{N}.$$

However, this metric does not allow us to obtain positive lower bounds on the associated Wasserstein curvature, and such tail probabilities do not involve any information on the chain in large time. In order to provide a better estimate as the time parameter is large, the idea is to consider the Wasserstein curvature of the process with respect to another metric than the classical one d on $\mathbb{N}$.
We denote in the sequel $a \wedge b := \min\{a,b\}$ and $a \vee b := \max\{a,b\}$, $a,b \in \mathbb{R}$.

Definition 2.2. *Given a positive function $u \in \mathscr{F}(\mathbb{N})$, define the distance $\delta : \mathbb{N}\times\mathbb{N} \to [0,+\infty)$ as*

$$\delta(x,y) := \left| \sum_{k=0}^{x-1} u(k) - \sum_{k=0}^{y-1} u(k) \right|, \quad u(-1) = 1.$$

Remark 2.3. Note that this distance has been used by Chen in [3] in order to obtain variational formulae for spectral gaps of birth-death processes.

Let us introduce the following set of assumptions on the transition rates of the generator:

(A) There exists two constants $K>0$ and $C>0$ such that

$$\left(\inf_{x\geq 0} \lambda_x\right) \wedge \left(\inf_{x\geq 1} \nu_x\right) \geq K \quad \text{and} \quad u(x) \leq C\left(\frac{1}{\sqrt{\nu_{x+1}}} \wedge \frac{1}{\sqrt{\lambda_x}}\right), \quad x \in \mathbb{N}.$$

(B) The stationary distribution π satisfies the moment condition (2.2) with the metric δ, and there exists a positive constant α such that

$$\inf_{x\in\mathbb{N}} \left\{ \nu_{x+1} + \lambda_x - \nu_x \frac{u(x-1)}{u(x)} - \lambda_{x+1}\frac{u(x+1)}{u(x)} \right\} \geq \alpha. \tag{2.3}$$

Under the assumption (A), we have a control on the distance δ as follows:

Lemma 2.4. *Under the assumption (A), the two inequalities below hold:*

(1) $\delta(x,y) \le \frac{C}{\sqrt{K}} d(x,y)$, $x,y \in \mathbb{N}$;

(2) $\sup_{x\in\mathbb{N}} \lambda_x \delta(x,x+1)^2 + \nu_x \delta(x,x-1)^2 \le 2C^2$.

Proof. On the one hand, it is sufficient by symmetry to prove the inequality (1) for $x<y$, $x,y \in \mathbb{N}$. Under the assumption (A), we have

$$\delta(x,y) = \left| \sum_{k=0}^{x-1} u(k) - \sum_{k=0}^{y-1} u(k) \right| = \sum_{k=x}^{y-1} u(k) \le \sum_{k=x}^{y-1} \frac{C}{\sqrt{K}} = \frac{C}{\sqrt{K}}(y-x).$$

Hence (1) is established. On the other hand, using the second inequality of the assumption (A), the proof of (2) is immediate. ∎

Remark 2.5. If at least one of the transition rates of the generator is unbounded, then the distances δ and d are not equivalent, and the proper inclusion $\mathrm{Lip}_\delta \subsetneq \mathrm{Lip}_d$ holds. In particular, the identity function $f(x)=x$ is not Lipschitz on $\mathbb{N}$ with respect to the metric δ.

The assumption (B) allows us to establish positive lower bounds on the δ-Wasserstein curvature of the process.

Proposition 2.6. *Assume that the assumption (B) is fulfilled. Then the δ-Wasserstein curvature of the process is bounded below by α.*

Proof. Consider $(X_t^x)_{t\ge0}$ and $(X_t^y)_{t\ge0}$ two independent copies of $(X_t)_{t\ge0}$, starting respectively from x and y. Then the generator $\tilde{\mathcal{L}}$ of the two-dimensional process $(X_t^x, X_t^y)_{t\ge0}$ is given for any real function f on $\mathbb{N}^2$ by

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

We have

$$\tilde{\mathcal{L}}\delta(x,y) \;=\; \sum_{k=x\wedge y}^{x\vee y-1} \tilde{\mathcal{L}}\delta(k,k+1)$$

$$
\begin{aligned}
&= \sum_{k=x\wedge y}^{x\vee y-1} (\lambda_{k+1}u(k+1) - \nu_{k+1}u(k) + \nu_k u(k-1) - \lambda_k u(k)) \\
&\leq -\alpha \sum_{k=x\wedge y}^{x\vee y-1} u(k) = -\alpha \sum_{k=x\wedge y}^{x\vee y-1} \delta(k,k+1) \leq -\alpha\delta(x,y),
\end{aligned}
$$

where in the first inequality we used the assumption (B). Since the stationary distribution π satisfies the moment condition (2.2) with the metric δ, the process $(\delta(X_t^x, X_t^y))_{t\geq 0}$ has finite expectation and we obtain from the latter inequality

$$
\mathbb{E}[\delta(X_t^x, X_t^y)] \leq e^{-\alpha t}\delta(x,y),
$$

which yields immediately the bound

$$
W_\delta(P_t(x,\cdot), P_t(y,\cdot)) \leq e^{-\alpha t}\delta(x,y).
$$

Finally, by the Kantorovich-Rubinstein duality theorem, the δ-Wasserstein curvature of the process is bounded below by the positive constant α. ■

Now we are able to state the main result of this paper, whose proof is given in the next section. Denote in the sequel the function $g(u) := (1+u)\log(1+u) - u$, $u > 0$.

Theorem 2.7. *Under the assumptions (A) and (B), then for any Lipschitz function $\phi \in \mathrm{Lip}_\delta$, any $t > 0$, any initial state $x \in \mathbb{N}$ and any deviation level $y > 0$, we have the following Poisson-type deviation inequality:*

$$
\begin{aligned}
\mathbb{P}_x\left(\left| \frac{1}{t}\int_0^t (\phi(X_s) - \mathbb{E}_x[\phi(X_s)])\, ds \right| \geq y \right) &\leq 2e^{-2Ktg\left(\frac{y\alpha}{2\sqrt{K}C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}\right)} \qquad (2.4)\\
&\leq 2e^{-\frac{ty\alpha\sqrt{K}}{2C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}\log\left(1+\frac{y\alpha}{2\sqrt{K}C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}\right)}.
\end{aligned}
$$

Remark 2.8. Under the assumption (B) and by invariance of the stationary distribution π, we have

$$
\begin{aligned}
\left| \frac{1}{t}\int_0^t P_s\phi(x)ds - \int \phi d\pi \right| &= \left| \frac{1}{t}\int_0^t \sum_{z\in\mathbb{N}} (P_s\phi(x) - P_s\phi(z))\,\pi(z)ds \right| \\
&\leq \|\phi\|_{\mathrm{Lip}_\delta} \sum_{z\in\mathbb{N}} \delta(x,z)\pi(z)\,\frac{1}{t}\int_0^t e^{-\alpha s}ds
\end{aligned}
$$

$$= \|\phi\|_{\mathrm{Lip}_\delta} \sum_{z\in\mathbb{N}} \delta(x,z)\pi(z)\, \frac{1-e^{-\alpha t}}{t\alpha},$$

so that for large t, we have the following order of magnitude

$$\left| \frac{1}{t}\int_0^t P_s\phi(x)ds - \int \phi d\pi \right| = O\left(\frac{1}{t}\right).$$

Hence, the deviation probability in the left-hand-side of (2.4) is comparable to the probability (1.1), at the price of strengthening the range of the deviation level y, and the the inequality (2.4) is understood as an estimate on the speed of convergence to equilibrium.

Remark 2.9. Since the function g in Theorem 2.7 above is equivalent to $u^2/2$ as u is close to 0 and to $u\log(u)$ as u tends to infinity, the Bennett-type inequality (2.4) exhibits a Gaussian tail for small values of the deviation level y, in accordance with the central limit theorem, and a Poisson tail for its large values. Therefore, the estimate (2.4) generalizes in two ways the inequalities given by Lezaud in [10], in the case of birth-death processes: on the one hand, the boundedness assumption on the transition rates of the generator is not required, and on the other hand this estimate is available for (not necessarily bounded) Lipschitz functions with respect to the metric δ. However, the price to pay here is to suppose that the assumption (B) is fulfilled, which is stronger than the existence of a spectral gap assumed by Lezaud in [10], see for instance Theorem 9.18 in [4].

3 Proof of Theorem 2.7

This section is devoted to the proof of Theorem 2.7, which is divided into several lemma. First, we establish a convenient upper bound in large time on the Laplace transform of a Lipschitz function of the process with respect to the distance δ, cf. Lemma 3.1. The key point of the proof of Theorem 2.7 is contained in Lemma 3.3 with the extension of such Laplace transform estimate to the multi-dimensional case by considering Lipschitz cylindrical functions with respect to the ℓ^1-metric. Finally, the last part is devoted to the approximation of the empirical distribution of the birth-death process by a suitable Lipschitz cylindrical function.

3.1 A Laplace transform estimate

Let us start with an upper bound on the Laplace transform of a Lipschitz function of the birth-death process $(X_t)_{t\geq 0}$ with respect to the distance δ.

Lemma 3.1. *Suppose that the assumptions (A) and (B) are satisfied. Then for any Lipschitz function $f \in \mathrm{Lip}_\delta$, any $t > 0$, any initial state $x \in \mathbb{N}$ and any $\tau > 0$, we have the following estimate on the Laplace transform:*

$$\mathbb{E}_x\left[e^{\tau(f(X_t)-\mathbb{E}_x[f(X_t)])}\right] \leq \exp\left\{\frac{K(1-e^{-2\alpha t})}{\alpha}\left(e^{\frac{\tau C\|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}} - \frac{\tau C\|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}} - 1\right)\right\}. \quad (3.1)$$

Proof. We adapt to the case of birth-death processes and with the metric δ the proof of Theorem 3.1 in [7].

Assume first that f is bounded. The process $\left(Z_s^f\right)_{0\leq s\leq t}$ given by $Z_s^f := P_{t-s}f(X_s) - P_t f(X_0)$ 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:

$$\begin{aligned} Z_s^f &= \int_0^s \left(P_{t-\tau}f(z+1) - P_{t-\tau}f(z)\right) 1_{\{X_{\tau-}=z\}}(N_z^\uparrow - \sigma_z^\uparrow)(d\tau) \\ &\quad + \int_0^s \left(P_{t-\tau}f(z-1) - P_{t-\tau}f(z)\right) 1_{\{X_{\tau-}=z\}}(N_z^\downarrow - \sigma_z^\downarrow)(d\tau), \end{aligned}$$

where $(N_z^\uparrow)_{z\in\mathbb{N}}$ and $(N_z^\downarrow)_{z\in\mathbb{N}}$ are two independent families of independent Poisson processes on $\mathbb{R}_+$ with respective intensities $\sigma_z^\uparrow(dt) = \lambda_z dt$ and $\sigma_z^\downarrow(dt) = \nu_z dt$. Since the δ-Wasserstein curvature is bounded below by $\alpha > 0$, i.e. for any function $f \in \mathrm{Lip}_\delta$,

$$\|P_t f\|_{\mathrm{Lip}_\delta} \leq e^{-\alpha t}\|f\|_{\mathrm{Lip}_\delta}, \quad t > 0,$$

the jumps of $\left(Z_s^f\right)_{0\leq s\leq t}$ are bounded as follows:

$$\begin{aligned} \sup_{0<s\leq t}\left|Z_s^f - Z_{s-}^f\right| &= \sup_{0<s\leq t}|P_{t-s}f(X_s) - P_{t-s}f(X_{s-})| \\ &\leq \|f\|_{\mathrm{Lip}_\delta}\sup_{0<s\leq t} e^{-\alpha(t-s)}\delta(X_s, X_{s-}) \\ &\leq \|f\|_{\mathrm{Lip}_\delta}\sup_{z\in\mathbb{N}}\delta(z, z+1) \\ &\leq \frac{C\|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}, \end{aligned}$$

where in the last inequality we used the inequality (1) of Lemma 2.4. Moreover, the angle bracket process satisfies the bound for any $s \in [0,t]$:

$$\begin{aligned}
&\langle Z^f, Z^f \rangle_s \\
&= \int_0^s \left\{ \lambda_{X_{\tau-}} \left(P_{t-\tau} f(X_{\tau-}+1) - P_{t-\tau} f(X_{\tau-}) \right)^2 + \nu_{X_{\tau-}} \left(P_{t-\tau} f(X_{\tau-}-1) - P_{t-\tau} f(X_{\tau-}) \right)^2 \right\} d\tau \\
&\leq \|f\|^2_{\mathrm{Lip}_\delta} \int_0^s e^{-2\alpha(t-\tau)} \left\{ \lambda_{X_{\tau-}} \delta(X_{\tau-}, X_{\tau-}+1)^2 + \nu_{X_{\tau-}} \delta(X_{\tau-}, X_{\tau-}-1)^2 \right\} d\tau \\
&\leq \frac{C^2(1-e^{-2\alpha t}) \|f\|^2_{\mathrm{Lip}_\delta}}{\alpha},
\end{aligned}$$

where in the last inequality we used the estimate (2) of Lemma 2.4.
Now, by Lemma 23.19 in [8], the process $(Y_s^{(\tau)})_{0 \leq s \leq t}$ given for any $\tau > 0$ by

$$Y_s^{(\tau)} := \exp\left\{ \tau Z_s^f - \tau^2 \psi\left(\frac{\tau C \|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}} \right) \langle Z^f, Z^f \rangle_s \right\}$$

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

$$\begin{aligned}
\mathbb{E}_x\left[e^{\tau(f(X_t) - \mathbb{E}_x[f(X_t)])} \right] &= \mathbb{E}_x\left[e^{\tau Z_t^f} \right] \\
&\leq \exp\left\{ \frac{\tau^2 C^2 (1-e^{-2\alpha t}) \|f\|^2_{\mathrm{Lip}_\delta}}{\alpha} \psi\left(\frac{\tau C \|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}} \right) \right\} \mathbb{E}_x\left[Y_t^{(\tau)} \right] \\
&\leq \exp\left\{ \frac{\tau^2 C^2 (1-e^{-2\alpha t}) \|f\|^2_{\mathrm{Lip}_\delta}}{\alpha} \psi\left(\frac{\tau C \|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}} \right) \right\} \\
&= \exp\left\{ \frac{K(1-e^{-2\alpha t})}{\alpha} \left(e^{\frac{\tau C \|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}} - \frac{\tau C \|f\|_{\mathrm{Lip}_\delta}}{\sqrt{K}} - 1 \right) \right\}.
\end{aligned}$$

Finally, the boundedness assumption on f is removed by a classical argument. ■

Remark 3.2. The upper bound in (3.1) allows us to sharpen in large time the deviation inequalities given in [7] for birth-death processes on $\mathbb{N}$. Indeed, under the notation of Lemma 3.1, we get easily from (3.1), the exponential Chebychev inequality and an optimization in $\tau > 0$, the following Poisson-type deviation inequality:

$$\mathbb{P}_x\left(|f(X_t) - \mathbb{E}_x[f(X_t)]| \geq y \right) \leq 2 e^{-\frac{K(1-e^{-2\alpha t})}{\alpha} g\left(\frac{\alpha y}{C\sqrt{K}(1-e^{-2\alpha t}) \|f\|_{\mathrm{Lip}_\delta}} \right)}, \quad y > 0,$$

where $g(u) := (1+u)\log(1+u) - u$, $u > 0$. In particular, letting t tend to infinity in the latter inequality entails the estimate under the stationary distribution π:

$$\pi\left(|f - \mathbb{E}_\pi[f]| \geq y\right) \quad \leq \quad 2e^{\frac{y\sqrt{K}}{C\|f\|_{\mathrm{Lip}_\delta}} - \left(\frac{K}{\alpha} + \frac{y\sqrt{K}}{C\|f\|_{\mathrm{Lip}_\delta}}\right)\log\left(1 + \frac{\alpha y}{C\sqrt{K}\|f\|_{\mathrm{Lip}_\delta}}\right)},$$

where $\mathbb{E}_\pi[f] = \sum_{z\in\mathbb{N}} f(z)\pi(z)$. However, in contrast to the deviation inequalities given in [7], the price to pay here is to require stronger regularity assumptions on the function f, namely $f \in \mathrm{Lip}_\delta$.

3.2 Tensorization procedure

The second step in the proof of Theorem 2.7 is devoted to the extension to the multi-dimensional case of the inequality (3.1). Such a tensorization procedure is the continuous time analogous of the method used by Rio in [11], then by Djellout, Guillin and Wu in the article [5], in order to establish Gaussian concentration inequalities for weakly dependent sequences.

Define $\mathrm{Lip}_\delta(n)$ the space of real Lipschitz functions on the product space $\mathbb{N}^n$, $n \in \mathbb{N} \setminus \{0,1\}$, endowed with the Lipschitz seminorm

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

where δ_n is the ℓ^1-distance on the product space $\mathbb{N}^n$ with respect to the metric δ, i.e. $\delta_n(y,z) := \sum_{i=1}^n \delta(y_i, z_i)$, $y, z \in \mathbb{N}^n$. We have the

Lemma 3.3. *Suppose that the assumptions (A) and (B) are satisfied. Denote X^n the sample $X^n = (X_{t_1}, \dots, X_{t_n})$, $0 = t_0 < t_1 < \cdots < t_n$, and let $f \in \mathrm{Lip}_\delta(n)$. Then for any initial state $x \in \mathbb{N}$ and any $\tau > 0$, we have the multi-dimensional Laplace transform estimate:*

$$\mathbb{E}_x\left[e^{\tau(f(X^n) - \mathbb{E}_x[f(X^n)])}\right] \leq \exp\left\{\sum_{k=1}^n h\left(\tau, t_k - t_{k-1}, \frac{M_k C\|f\|_{\mathrm{Lip}_\delta(n)}}{\sqrt{K}}\right)\right\}, \tag{3.2}$$

where $M_k := \sum_{l=k}^n e^{-\alpha(t_l - t_k)}$ and h is the function defined on $(\mathbb{R}_+)^3$ by

$$h(\tau, t, z) := \frac{K\left(1 - e^{-2\alpha t}\right)}{\alpha}\left(e^{\tau z} - \tau z - 1\right).$$

Proof. Fix the initial state $x \in \mathbb{N}$. If g is a one-dimensional Lipschitz function with respect to the metric δ, then by Lemma 3.1, we have for any $t > 0$ and any $\tau > 0$ the inequality:

$$\mathbb{E}_x\left[e^{\tau g(X_t)}\right] \leq \exp\left\{\tau \mathbb{E}_x[g(X_t)] + h\left(\tau, t, \frac{C\|g\|_{\mathrm{Lip}_\delta}}{\sqrt{K}}\right)\right\}. \tag{3.3}$$

Let us now extend this estimate to the multi-dimensional case by tensorization of the Laplace transform. We sketch the argument for $n = 2$. Let $0 < s < t$ and denote the functions $f_y(z) := f(y, z)$ and $f_1(y) := \sum_{z\in\mathbb{N}} f(y, z)P_{t-s}(y, z)$. It is clear that $\|f_y\|_{\mathrm{Lip}_\delta} \leq \|f\|_{\mathrm{Lip}_\delta(2)}$. Let us verify that the function f_1 is also Lipschitz with respect to δ, with furthermore the bound

$$\|f_1\|_{\mathrm{Lip}_\delta} \leq \left(1 + e^{-\alpha(t-s)}\right)\|f\|_{\mathrm{Lip}_\delta(2)}. \tag{3.4}$$

We have for any $y \in \mathbb{N}$:

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

where in the second inequality we used Proposition 2.6, i.e. the δ-Wasserstein curvature is bounded below by α. Since the Lipschitz constant of a Lipschitz function $v \in \mathrm{Lip}_\delta$ rewrites as

$$\|v\|_{\mathrm{Lip}_\delta} = \sup_{x\in\mathbb{N}} \frac{|v(x+1) - v(x)|}{\delta(x, x+1)},$$

the function f_1 is Lipschitz with respect to the metric δ and the inequality (3.4) is satisfied.

Now, using the Markov property, the estimate (3.3) applied to the Lipschitz functions f_y then f_1, entails

$$\mathbb{E}_x\left[e^{\tau f(X_s, X_t)}\right]$$

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

since the function h is non-decreasing in its last variable. Hence, the latter inequality is exactly the estimate (3.2) for the dimension $n = 2$.

In the general case, we show similarly that for any $k = 1, \ldots, n$, the function f_k defined on $\mathbb{N}^k$ by

$$
f_k(x_1, \ldots, x_k) := \sum_{x_{k+1},\ldots,x_n\in\mathbb{N}} f(x_1, \ldots, x_k, \ldots, x_n) \prod_{l=k}^{n-1} P_{t_{l+1}-t_l}(x_l, x_{l+1}),
$$

has Lipschitz seminorm with respect to the k^{th} variable, in the metric δ, smaller than $M_k \|f\|_{\mathrm{Lip}_\delta(n)}$. Therefore, using again that h is non-decreasing in its last variable, we obtain (3.2) in full generality with a simple recursive argument on the dimension n. ■

3.3 Proof of Theorem 2.7

Now we are able to prove Theorem 2.7.

Proof of Theorem 2.7. We use the notation of Lemma 3.3. Fix the initial state $x \in \mathbb{N}$ of the birth-death process $(X_t)_{t\geq 0}$ and a finite time horizon $t > 0$. Define $t_k = kt/n$, $k = 0, \ldots, n$, a regular subdivision of the time interval $[0, t]$ and let X^n be the sample $X^n = (X_{t_1}, \ldots, X_{t_n})$. Letting the cylindrical function

$$
f(z) := \frac{1}{n} \sum_{k=1}^{n} \phi(z_k), \quad z = (z_1, \ldots, z_n),
$$

then f is Lipschitz on the product space $\mathbb{N}^n$ with respect to the ℓ^1-metric δ_n and we have the bound

$$
\|f\|_{\mathrm{Lip}_\delta(n)} \leq \frac{1}{n} \|\phi\|_{\mathrm{Lip}_\delta}.
$$

Hence, since

$$\sup_{k=1,\ldots,n} M_k = \sup_{k=1,\ldots,n} \sum_{l=k}^{n} e^{-\alpha t(l-k)/n} = \frac{1-e^{-\alpha t}}{1-e^{-\alpha t/n}},$$

and that the function h is non-decreasing in its last variable, Lemma 3.3 entails for any $\tau > 0$:

$$\mathbb{E}_x\left[e^{\tau(f(X^n)-\mathbb{E}_x[f(X^n)])}\right] \leq \exp\left\{nh\left(\tau, \frac{t}{n}, \frac{C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}{n\sqrt{K}(1-e^{-\alpha t/n})}\right)\right\}. \tag{3.5}$$

As we have the following approximation of the empirical distribution

$$\frac{1}{t}\int_0^t \phi(X_s)ds = \lim_{n\to+\infty} \frac{1}{n}\sum_{k=1}^{n} \phi(X_{kt/n}), \quad \mathbb{P}_x - a.s.,$$

we obtain by using Fatou's lemma in (3.5):

$$\begin{aligned} \mathbb{E}_x\left[e^{\frac{\tau}{t}\int_0^t(\phi(X_s)-\mathbb{E}_x[\phi(X_s)])ds}\right] &\leq \liminf_{n\to+\infty} \mathbb{E}_x\left[e^{\tau(f(X^n)-\mathbb{E}_x[f(X^n)])}\right] \\ &\leq \exp\left\{2Kt\left(e^{\frac{\tau C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}{\alpha\sqrt{K}t}} - \frac{\tau C(1-e^{-\alpha t})\|\phi\|_{\mathrm{Lip}_\delta}}{\alpha\sqrt{K}t} - 1\right)\right\}. \end{aligned}$$

The exponential Chebychev inequality together with an optimization in $\tau > 0$ achieves the proof of Theorem 2.7 in the one-sided case. Finally, applying the same reasoning to the function $-\phi$ yields the general result. ■

Remark 3.4. Since the method emphasized in the proof of Theorem 2.7 is available for birth-death processes with values in a finite set, we also recover the results given by Lezaud in the paper [9]. Moreover, we point out that this approach works for diffusion processes satisfying the Bakry-Emery curvature condition, and would imply a Hoeffding-type deviation inequality. However, it is well-known that this curvature criterion implies a logarithmic Sobolev inequality which in turn entails immediately the result by the Corollary 4 of [12].

Remark 3.5. We mention that the problem to find a similar rate of convergence as (2.4) for the identity function $\phi(x) = x$ on $\mathbb{N}$ is unsolved when the generator of the birth-death process is unbounded. Indeed, our estimate is only available for Lipschitz function ϕ lying in the space Lip_δ, which excludes in that case the identity function.

4 Application to the $M/M/\infty$ queueing process

Consider a ticket booth system where each customer arriving in front of some stands is immediately served. Denoting X_t the number of customers in the system at time $t > 0$, we assume that the arrival process is a Poisson process of intensity $\lambda > 0$ and that conditionally on the event $\{X_s = x\}$, the service time $T := \inf\{t > s : X_t \neq X_s\}$ follows an exponential distribution with parameter $\lambda + \nu x$, $\nu > 0$. Then the stochastic process $(X_t)_{t\geq 0}$ is a $M/M/\infty$ queueing process. It is an ergodic birth-death process whose generator is given by

$$\mathcal{L}f(x) = \lambda(f(x+1) - f(x)) + \nu x(f(x-1) - f(x)), \quad x \in \mathbb{N}.$$

The stationary distribution is the Poisson measure $\mathscr{P}(\sigma)$ on $\mathbb{N}$ with parameter $\sigma := \lambda/\nu$, i.e.

$$\mathscr{P}(\sigma)(x) = e^{-\sigma}\frac{\sigma^x}{x!}, \quad x \in \mathbb{N}.$$

For the sake of simplicity, we assume in the sequel that the process is normalized, i.e. $\lambda = \nu$. The knowledge of its distribution at time $t > 0$ allows us to make explicit computations. Indeed, by the Mehler-type convolution formula given by Chafaï in [2],

$$\mathcal{L}(X_t|X_0 = x) = \mathscr{B}\left(x, e^{-\nu t}\right) * \mathscr{P}\left(1 - e^{-\nu t}\right), \quad t > 0, \tag{4.1}$$

where $\mathscr{B}(n,p)$ denotes a binomial distribution with parameters $n \in \mathbb{N}$ and $p \in (0,1)$, we get for any $\tau > 0$,

$$\begin{aligned}
\mathbb{E}_x\left[e^{\tau(X_t - \mathbb{E}_x[X_t])}\right] &= \exp\left\{x\log\left(1 + e^{-\nu t}(e^\tau - 1)\right) - \tau x e^{-\nu t} + \left(1 - e^{-\nu t}\right)\left(e^\tau - \tau - 1\right)\right\} \\
&\leq \exp\left\{\left(x e^{-\nu t} + 1 - e^{-\nu t}\right)\left(e^\tau - \tau - 1\right)\right\} \\
&= \exp\left\{\mathbb{E}_x[X_t]\ \left(e^\tau - \tau - 1\right)\right\},
\end{aligned} \tag{4.2}$$

where we used the inequality $\log(1+x) \leq x$, $x > 0$. Thus, by Chebychev's inequality, we obtain for any $y > 0$ the Poisson-type deviation inequality

$$\begin{aligned}
\mathbb{P}_x\left(X_t - \mathbb{E}_x[X_t] \geq y\right) &\leq \inf_{\tau > 0} e^{-\tau y}\, \mathbb{E}_x\left[e^{\tau(X_t - \mathbb{E}_x[X_t])}\right] \\
&\leq \exp\left\{y - \left(\mathbb{E}_x[X_t] + y\right)\log\left(1 + \frac{y}{\mathbb{E}_x[X_t]}\right)\right\},
\end{aligned} \tag{4.3}$$

which entails by ergodicity as $t \to +\infty$ the estimate

$$\mathbb{P}\left(X - \mathbb{E}[X] \geq y\right) \leq \exp\left\{y - (1+y)\log\left(1+y\right)\right\},$$

where X is a Poisson random variable of intensity 1. Therefore, we expect that the Poisson-type deviation inequality (4.3) might be extended to the empirical distribution of the $M/M/\infty$ queueing process.

Choosing the function $u(x) := 1/\sqrt{x+1}$, $x \in \mathbb{N}$, in the definition of the metric δ, the transition rates of the generator satisfy the assumption (A) with the constants $C = \sqrt{K} = \sqrt{\nu}$. Moreover, a short computation shows that the assumption (B) is verified with $\alpha = \nu/2 > 0$, which is the half of the exact curvature of the $M/M/\infty$ queueing process, see [2]. Hence, Theorem 2.7 entails for any Lipschitz function $f \in \mathrm{Lip}_\delta$, any $t > 0$, any initial state $x \in \mathbb{N}$ and any $y > 0$, the Poisson-type deviation inequality

$$\begin{aligned}
\mathbb{P}_x\left(\left|\frac{1}{t}\int_0^t \left(\phi(X_s) - \mathbb{E}_x[\phi(X_s)]\right) ds\right| \geq y\right) &\leq 2e^{-2\nu t g\left(\frac{y}{4(1-e^{-\nu t/2})\|\phi\|_{\mathrm{Lip}_\delta}}\right)} \\
&\leq 2e^{-\frac{ty}{4(1-e^{-\nu t/2})\|\phi\|_{\mathrm{Lip}_\delta}}\log\left(1+\frac{y}{4(1-e^{-\nu t/2})\|\phi\|_{\mathrm{Lip}_\delta}}\right)},
\end{aligned}$$

where $g(u) = (1+u)\log(1+u) - u$, $u > 0$.

Remark 4.1. We mention that the Laplace transform estimate (4.2) cannot be tensorized by using the general method of Section 3, since its upper bound depends strongly on the initial condition $x \in \mathbb{N}$.

References

[1] P. Cattiaux and A. Guillin. Deviation bounds for additive functionals of Markov processes. Preprint, 2006.

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

[3] M.F. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sinica*, 12(4):337–360, 1996.

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

[5] 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.

[6] M. Gourcy and L. Wu. Logarithmic Sobolev inequalities of diffusions for the L2-metric. *Potential Anal.*, 25(1):77–102, 2006.

[7] A. Joulin. Poisson-type deviation inequalities for curved continuous time Markov chains. Preprint, 2006.

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

[9] P. Lezaud. Chernoff-type bound for finite Markov chains. *Ann. Appl. Probab.*, 8(3):849–867, 1998.

[10] P. Lezaud. Chernoff and Berry-Esséen inequalities for Markov processes. *ESAIM Probab. Statist.*, 5:183–201 (electronic), 2001.

[11] E. Rio. Inégalités de Hoeffding pour les fonctions lipschitziennes de suites dépendantes. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(10):905–908, 2000.

[12] L. Wu. A deviation inequality for non-reversible Markov processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 36(4):435–445, 2000.