arXiv:1011.2331v1 [math.PR] 10 Nov 2010

INTERTWINING AND COMMUTATION RELATIONS FOR BIRTH-DEATH PROCESSES

DJALIL CHAFAÏ AND ALDÉRIC JOULIN

ABSTRACT. Given a birth-death process on $\mathbb{N}$ with semigroup $(P_t)_{t\geq 0}$ and a discrete gradient ∂_u depending on a positive weight u, we establish intertwining relations of the form $\partial_u P_t = Q_t \partial_u$, where $(Q_t)_{t\geq 0}$ is the Feynman-Kac semigroup with potential V_u of another birth-death process. We provide applications when V_u is positive and uniformly bounded from below, including Lipschitz contraction and Wasserstein curvature, various functional inequalities, and stochastic orderings. The proofs are remarkably simple and rely on interpolation, commutation, and convexity.

1. Introduction

Commutation relations and convexity are useful tools for the fine analysis of Markov diffusion semigroups [BE, B, L]. The situation is more delicate on discrete spaces, due to the lack of a chain rule formula [Cha, CDPP]. In this work, we obtain intertwining and sub-commutation relations for a class of birth-death processes involving a discrete gradient and an auxiliary Feynman-Kac semigroup. We also provide various applications of these relations. More precisely, let us consider a birth-death process $(X_t)_{t\geq 0}$ on the state space $\mathbb{N} := \{0, 1, 2, \ldots\}$, i.e. a Markov process with transition probabilities given by

$$P_t^x(y) = \mathbb{P}_x(X_t = y) = \begin{cases} \lambda_x t + o(t) & \text{if } y = x+1, \\ \nu_x t + o(t) & \text{if } y = x-1, \\ 1 - (\lambda_x + \nu_x)t + o(t) & \text{if } y = x, \end{cases}$$

where $\lim_{t\to 0} t^{-1} o(t) = 0$. The transition rates λ and ν are respectively called the birth and death rates of the process $(X_t)_{t\geq 0}$. We assume that the process is irreducible, positive recurrent, and non-explosive. This holds when the rates satisfy to $\lambda > 0$ on $\mathbb{N}$ and $\nu > 0$ on $\mathbb{N}^*$ and $\nu_0 = 0$ and

$$\sum_{x=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{x-1}}{\nu_1 \nu_2 \cdots \nu_x} < \infty \quad \text{and} \quad \sum_{x=1}^{\infty} \left(\frac{1}{\lambda_x} + \frac{\nu_x}{\lambda_x \lambda_{x-1}} + \cdots + \frac{\nu_x \cdots \nu_1}{\lambda_x \cdots \lambda_1 \lambda_0} \right) = \infty.$$

Date: Preprint. June 2010. Compiled November 11, 2010.
2000 *Mathematics Subject Classification.* 60J80 ; 39C05.
Key words and phrases. Birth-death process; Feynman-Kac semigroup; discrete gradients; intertwining relation; functional inequalities.

The unique stationary distribution μ of the process is reversible and is given by

$$\mu(x) = \mu(0) \prod_{y=1}^{x} \frac{\lambda_{y-1}}{\nu_y}, \; x \in \mathbb{N} \quad \text{with} \quad \mu(0) := \left(1 + \sum_{x=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{x-1}}{\nu_1 \nu_2 \cdots \nu_x}\right)^{-1}. \tag{1.1}$$

Let us denote by $\mathcal{F}$ (respectively $\mathcal{F}_+$) the space of real-valued (respectively positive) functions f on $\mathbb{N}$, and let $b\mathcal{F}$ be the subspace of bounded functions. The associated semigroup $(P_t)_{t\geq 0}$ is defined for any function $f \in b\mathcal{F} \cup \mathcal{F}_+$ and $x \in \mathbb{N}$ as

$$P_t f(x) = \mathbb{E}_x[f(X_t)] = \sum_{y=0}^{\infty} f(y) P_t^x(y).$$

This family of operators is positivity preserving and contractive on $L^p(\mu)$, $p \in [1, \infty]$. Moreover, the semigroup is also symmetric in $L^2(\mu)$ since $\lambda_x \mu(x) = \nu_{1+x} \mu(1+x)$ for any $x \in \mathbb{N}$ (detailed balance equation). The generator $\mathcal{L}$ of the process is given for any $f \in \mathcal{F}$ and $x \in \mathbb{N}$ by

$$\begin{aligned} \mathcal{L} f(x) &= \lambda_x \, (f(x+1) - f(x)) + \nu_x \, (f(x-1) - f(x)) \\ &= \lambda_x \, \partial f(x) + \nu_x \, \partial^* f(x), \end{aligned}$$

where ∂ and ∂^* are respectively the forward and backward discrete gradients on $\mathbb{N}$:

$$\partial f(x) := f(x+1) - f(x) \quad \text{and} \quad \partial^* f(x) := f(x-1) - f(x).$$

Our approach is inspired from the remarkable properties of two special birth-death processes: the $M/M/1$ and the $M/M/\infty$ queues. The $M/M/\infty$ queue has rates $\lambda_x = \lambda$ and $\nu_x = \nu x$ for positive constants λ and ν. It is positive recurrent and its stationary distribution is the Poisson measure μ_ρ with mean $\rho = \lambda/\mu$. If $\mathscr{B}_{x,p}$ stands for the binomial distribution of size $x \in \mathbb{N}$ and parameter $p \in [0,1]$, the $M/M/\infty$ process satisfies for every $x \in \mathbb{N}$ and $t \geq 0$ to the Mehler type formula

$$\mathscr{L}(X_t | X_0 = x) = \mathscr{B}_{x, e^{-\nu t}} * \mu_{\rho(1-e^{-\nu t})}. \tag{1.2}$$

The $M/M/1$ queueing process has rates $\lambda_x = \lambda$ and $\nu_x = \nu \mathbf{1}_{\mathbb{N}\setminus\{0\}}$ where $0 < \lambda < \nu$ are constants. It is a positive recurrent random walk on $\mathbb{N}$ reflected at 0. Its stationary distribution μ is the geometric measure with parameter $\rho := \lambda/\nu$ given by $\mu(x) = (1-\rho)\rho^x$ for all $x \in \mathbb{N}$. A remarkable common property shared by the $M/M/1$ and $M/M/\infty$ processes is the intertwining relation

$$\partial \mathcal{L} = \mathcal{L}^V \partial \tag{1.3}$$

where $\mathcal{L}^V = \mathcal{L} - V$ is the discrete Schrödinger operator with potential V given by

- $V(x) := \nu$ in the case of the $M/M/\infty$ queue
- $V(x) := \nu \mathbf{1}_{\{0\}}(x)$ for the $M/M/1$ queue.

The operator $\mathcal{L}^V$ is the generator of a Feynman-Kac semigroup $(P_t^V)_{t\geq 0}$ given by

$$P_t^V f(x) = \mathbb{E}_x \left[f(X_t) \exp\left(- \int_0^t V(X_s) ds \right) \right].$$

The intertwining relation (1.3) is the infinitesimal version at time $t = 0$ of the semigroup intertwining

$$\partial P_t f(x) = P_t^V \partial f(x) = \mathbb{E}_x \left[\partial f(X_t) \exp \left(- \int_0^t V(X_s)\, ds \right) \right]. \tag{1.4}$$

Conversely, one may deduce (1.4) from (1.3) by using a semigroup interpolation. Namely, if we consider $s \in [0,t] \mapsto J(s) := P_s^V \partial P_{t-s} f$ with V as above, then (1.4) rewrites as $J(0) = J(t)$ and (1.4) follows from (1.3) since

$$J'(s) = P_s^V \left(\mathcal{L}^V \partial P_{t-s} f - \partial \mathcal{L} P_{t-s} f \right) = 0.$$

In section 2, we obtain by using semigroup interpolation an intertwining relation similar to (1.4) for more general birth-death processes. By using convexity as an additional ingredient, we also obtain sub-commutation relations. These results have several applications explored in section 3, including Lipschitz contraction and Wasserstein curvature (section 3.1), functional inequalities including Poincaré, entropic, isoperimetric and transportation-information inequalities (section 3.2), hitting time of the origin for the $M/M/1$ queue (section 3.3), convex domination and stochastic orderings (section 3.4).

2. Intertwining relations and sub-commutations

Let us fix some $u \in \mathcal{F}_+$. The u-modification of the original process $(X_t)_{t\geq 0}$ is a birth-death process $(X_{u,t})_{t\geq 0}$ with semigroup $(P_{u,t})_{t\geq 0}$ and generator $\mathcal{L}_u$ given by

$$\mathcal{L}_u f(x) = \lambda_x^u \, \partial f(x) + \nu_x^u \, \partial^* f(x),$$

where the birth and death rates are respectively given by

$$\lambda_x^u := \frac{u_{x+1}}{u_x} \lambda_{x+1} \quad \text{and} \quad \nu_x^u := \frac{u_{x-1}}{u_x} \nu_x.$$

One can check that the measure $\lambda u^2 \mu$ is reversible for $(X_{u,t})_{t\geq 0}$. As consequence, the process $(X_{u,t})_{t\geq 0}$ is positive recurrent if and only if λu^2 is μ-integrable. We define the discrete gradient ∂_u and the potential V_u by

$$\partial_u := (1/u)\partial \quad \text{and} \quad V_u(x) := \nu_{x+1} - \nu_x^u + \lambda_x - \lambda_x^u.$$

Let $\varphi : \mathbb{R} \to \mathbb{R}_+$ be a smooth convex function such that for some constant $c > 0$, and for all $r \in \mathbb{R}$,

$$\varphi'(r) r \geq c \varphi(r). \tag{2.1}$$

In particular, φ vanishes at 0, is non-increasing on $(-\infty, 0)$ and non-decreasing on $(0, \infty)$. Moreover, the behavior at infinity is at least polynomial of degree c. Note that one can easily find a sequence of such functions converging pointwise to the absolute value $|\cdot|$.

Theorem 2.1 (Intertwining and sub-commutation)**.** *Assume that for every* $x \in \mathbb{N}$ *and* $t \geq 0$*, we have*

$$\mathbb{E}_x\left[\exp\Big(-\int_0^t V_u(X_{u,s})\,ds\Big)\right] < \infty.$$

Then for every $f \in b\mathcal{F}$*,* $x \in \mathbb{N}$*,* $t \geq 0$*,*

$$\partial_u P_t f(x) \;=\; P_{u,t}^{V_u}\partial_u f(x) \;=\; \mathbb{E}_x\left[\partial_u f(X_{u,t})\,\exp\Big(-\int_0^t V_u(X_{u,s})\,ds\Big)\right]. \tag{2.2}$$

Moreover, if $V_u \geq 0$ *then for every* $f \in b\mathcal{F}$*,* $x \in \mathbb{N}$*,* $t \geq 0$*,*

$$\varphi\left(\partial_u P_t f\right)(x) \leq \mathbb{E}_x\left[\varphi(\partial_u f)(X_{u,t})\,\exp\Big(-\int_0^t cV_u(X_{u,s})\,ds\Big)\right]. \tag{2.3}$$

Proof. Let us prove (2.3). If we define

$$s \in [0,t] \mapsto J(s) := P_{u,s}^{cV_u}\varphi(\partial_u P_{t-s} f)$$

then (2.3) rewrites as $J(0) \leq J(t)$. Hence it suffices to show that J is non-decreasing. We have the intertwining relation

$$\partial_u \mathcal{L} = \mathcal{L}_u^{V_u}\partial_u, \tag{2.4}$$

where $\mathcal{L}_u$ is the generator of the u-modification process $(X_{u,t})_{t\geq 0}$ and where

$$\mathcal{L}_u^{V_u} := \mathcal{L}_u - V_u.$$

Now

$$J'(s) = P_{u,s}^{cV_u}(T) \quad \text{where} \quad T = \mathcal{L}_u^{cV_u}\varphi(\partial_u P_{t-s} f) - \varphi'(\partial_u P_{t-s} f)\,\partial_u \mathcal{L} P_{t-s} f.$$

Letting $g_u = \partial_u P_{t-s} f$, we obtain, by using (2.4),

$$\begin{aligned} T &= \mathcal{L}_u^{cV_u}\varphi(g_u) - \varphi'(g_u)\mathcal{L}_u^{V_u} g_u \\ &= \lambda^u\left(\partial\varphi(g_u) - \varphi'(g_u)\partial g_u\right) + \nu^u\left(\partial^*\varphi(g_u) - \varphi'(g_u)\partial^* g_u\right) + V_u\left(\varphi'(g_u)g_u - c\varphi(g_u)\right). \end{aligned}$$

Now (2.1) and $V_u \geq 0$ give $T \geq 0$. Since the Feynman-Kac semigroup $(P_{u,t}^{cV_u})_{t\geq 0}$ is positivity preserving, we get (2.3). The proof of (2.2) is similar but simpler (T is identically zero). □

Remark 2.2 (Propagation of monotonicity)**.** *The identity* (2.2) *provides a new proof of the propagation of monotonicity* [S, Proposition 4.2.10]*: if* f *is non-increasing then* $P_t f$ *is also non-increasing.*

Remark 2.3 (Other gradients)**.** *Theorem 2.1 possesses a natural analogue for the discrete backward gradient* ∂^**. We ignore if there exists a nice intertwining involving a combination of both.*

Our second result below complements the previous one for the case $u = 1$. Let $\mathcal{I}$ be an open interval of $\mathbb{R}$ and let $\varphi : \mathcal{I} \to \mathbb{R}$ be a smooth convex function such that $\varphi'' > 0$ and $-1/\varphi''$ is convex on $\mathcal{I}$. Following the notations of [Cha], we define on the convex subset $\mathcal{A}_\mathcal{I} := \{(r,s) \in \mathbb{R}^2 : (r, r+s) \in \mathcal{I} \times \mathcal{I}\}$ the non-negative function B^φ on $\mathcal{A}_\mathcal{I}$ by

$$B^\varphi(r,s) := (\varphi'(r+s) - \varphi'(r))\, s, \quad (r,s) \in \mathcal{A}_\mathcal{I}.$$

By Theorem 4.4 in [Cha], B^φ is convex on $\mathcal{A}_\mathcal{I}$. Some interesting examples of such functionals will be given in section 3.2 below.

Theorem 2.4 (Sub-commutation for 1-modification)**.** *If the transition rate λ is non-increasing and ν is non-decreasing then for any $f \in b\mathcal{F}$ and $t \geq 0$,*

$$B^\varphi\left(P_t f, \partial P_t f\right) \leq P_{1,t}^{V_1} B^\varphi(f, \partial f) \tag{2.5}$$

where the non-negative potential is $V_1 := \partial(\nu - \lambda)$.

Proof. If we define $s \in [0,t] \mapsto J(s) := P_{1,s}^{V_1} B^\varphi(P_{t-s} f, \partial P_{t-s} f)$ we see that (2.5) rewrites as $J(0) \leq J(t)$. Denote $F = P_{t-s} f$ and $G = \partial P_{t-s} f = \partial F$. Using (2.4) with the constant function $u = 1$, we have $J'(s) = P_{1,s}^{V_1}(T)$ with

$$\begin{aligned}
T &= \mathcal{L}_1^{V_1} B^\varphi(F,G) - \frac{\partial}{\partial x} B^\varphi(F,G)\, \mathcal{L} F - \frac{\partial}{\partial y} B^\varphi(F,G)\, \mathcal{L}_1^{V_1} G \\
&= \lambda^1 \partial B^\varphi(F,G) - \lambda \frac{\partial}{\partial x} B^\varphi(F,G)\, \partial F - \lambda^1 \frac{\partial}{\partial y} B^\varphi(F,G)\, \partial G \\
&\quad + \nu^1 \partial^* B^\varphi(F,G) - \nu \frac{\partial}{\partial x} B^\varphi(F,G)\, \partial^* F - \nu^1 \frac{\partial}{\partial y} B^\varphi(F,G)\, \partial^* G \\
&\quad + \partial(\nu - \lambda) \left(\frac{\partial}{\partial y} B^\varphi(F,G)\, G - B^\varphi(F,G) \right) \\
&\geq \partial \nu \left(\frac{\partial}{\partial y} B^\varphi(F,G)\, G - B^\varphi(F,G) \right) \\
&\quad - \partial \lambda \left(\frac{\partial}{\partial y} B^\varphi(F,G)\, G - \frac{\partial}{\partial x} B^\varphi(F,G)\, G - B^\varphi(F,G) \right),
\end{aligned}$$

and where in the last line we used the convexity of the bivariate function B^φ. Moreover, since the birth and death rates λ and ν are respectively non-increasing and non-decreasing on the one hand, and using once again convexity on the other hand, we get

$$\frac{\partial}{\partial y} B^\varphi(F,G)\, G \geq \begin{cases} \frac{\partial}{\partial x} B^\varphi(F,G)\, G + B^\varphi(F,G) \\ B^\varphi(F,G) \end{cases}$$

from which we deduce that T is non-negative and thus J is non-decreasing. □

Remark 2.5 (Diffusion case)**.** *Actually, the intertwining relations above have their counterpart in continuous state space. Let $\mathcal{A}$ be the generator of a one-dimensional real-valued diffusion $(X_t)_{t\geq 0}$ of the type*

$$\mathcal{A} f = \sigma^2 f'' + b f',$$

where f and the two functions σ, b are sufficiently smooth. Given a smooth positive function a on $\mathbb{R}$, the gradient of interest is $\nabla_a f = a\, f'$. Denote $(P_t)_{t\geq 0}$ the associated diffusion semigroup. Then it is not hard to adapt to the continuous case the argument of theorem 2.1 to show that the following intertwining relation holds:

$$\nabla_a P_t f(x) = \mathbb{E}_x \left[\nabla_a f(X_{a,t}) \, \exp\left(- \int_0^t V_a(X_{a,s})\, ds \right) \right].$$

Here $(X_{a,t})_{t\geq 0}$ is a new diffusion process with generator

$$\mathcal{A}_a f = \sigma^2 f'' + b_a f'$$

and drift b_a and potential V_a given by

$$b_a := 2\sigma\sigma' + b - 2\sigma^2 \frac{a'}{a} \quad \text{and} \quad V_a := \sigma^2 \frac{a''}{a} - b' + \frac{a'}{a}\, b_a.$$

In particular, if the weight $a = \sigma$, where σ is assumed to be positive, then the two processes above have the same distribution and by Jensen's inequality, we obtain

$$|\nabla_\sigma P_t f(x)| \leq \mathbb{E}_x \left[|\nabla_\sigma f(X_t)| \, \exp\left(- \int_0^t \left(\sigma\sigma'' - b' + b\, \frac{\sigma'}{\sigma} \right) (X_s)\, ds \right) \right].$$

Hence under the assumption that there exists a constant ρ such that

$$\inf \, \sigma\sigma'' - b' + b\, \frac{\sigma'}{\sigma} \geq \rho,$$

then we get $|\nabla_\sigma P_t f| \leq e^{-\rho t}\, P_t |\nabla_\sigma f|$. This type of sub-commutation relation is at the heart of the Bakry-Émery calculus [BE, B, L]. *See also* [MT] *for a nice study of functional inequalities for the invariant measure under the condition $\rho = 0$.*

3. Applications

This section is devoted to applications of theorems 2.1 and 2.4.

3.1. **Lipschitz contraction and Wasserstein curvature.** Theorem 2.1 allows to recover a result of Chen [Che1] on the contraction property of the semigroup on the space of Lipschitz functions. Indeed, the intertwining (2.2) can be used to derive bounds on the Wasserstein curvature of the birth-death process, without using the coupling technique emphasized by Chen. For a distance d on $\mathbb{N}$, we denote by $\mathcal{P}_d(\mathbb{N})$ the set of probability measures ξ on $\mathbb{N}$ such that $\sum_{x\in\mathbb{N}} d(x, x_0)\xi(x) < \infty$ for some (or equivalently for all) $x_0 \in \mathbb{N}$. We recall that the Wasserstein distance between two probability measures $\mu_1, \mu_2 \in \mathcal{P}_d(\mathbb{N})$ is defined by

$$\mathcal{W}_d(\mu_1, \mu_2) = \inf_{\gamma \in \mathrm{Marg}(\mu_1,\mu_2)} \int_{\mathbb{N}} \int_{\mathbb{N}} d(x,y) \gamma(dx, dy), \tag{3.1}$$

where $\mathrm{Marg}(\mu_1, \mu_2)$ is the set of probability measures on $\mathbb{N}^2$ such that the marginal distributions are μ_1 and μ_2, respectively. The Kantorovich-Rubinstein duality [V, Theorem 5.10] gives

$$\mathcal{W}_d(\mu_1,\mu_2) = \sup_{g\in \mathrm{Lip}_1(d)} \int_{\mathbb{N}} g\, d(\mu_1-\mu_2), \tag{3.2}$$

where Lip(d) is the set of Lipschitz function g with respect to the distance d, i.e.

$$\|g\|_{\mathrm{Lip}(d)} := \sup_{\substack{x,y\in\mathbb{N}\\ x\neq y}} \frac{|g(x)-g(y)|}{d(x,y)} < \infty,$$

and $\mathrm{Lip}_1(d)$ consists of 1-Lipschitz functions. We assume that the kernel $P_t^x \in \mathcal{P}_d(\mathbb{N})$ for every $x\in\mathbb{N}$ and $t\geq 0$ so that the semigroup is well-defined on Lip(d). In particular the intertwining relation of theorem 2.1 is available for any function $f\in \mathrm{Lip}(d)$. The Wasserstein curvature of $(X_t)_{t\geq 0}$ with respect to a given distance d is the optimal (largest) constant σ in the following contraction inequality:

$$\|P_t\|_{\mathrm{Lip}(d)\to \mathrm{Lip}(d)} \leq e^{-\sigma t}, \quad t\geq 0. \tag{3.3}$$

Here $\|P_t\|_{\mathrm{Lip}(d)\to\mathrm{Lip}(d)}$ denotes the supremum of $\|P_t f\|_{\mathrm{Lip}(d)}$ when f runs over $\mathrm{Lip}_1(d)$. It is actually equivalent to the property that

$$\mathcal{W}_d(P_t^x, P_t^y) \leq e^{-\sigma t}\, d(x,y), \quad x,y\in\mathbb{N}, \quad t\geq 0.$$

If the optimal constant is positive, then the process is positive recurrent and the semigroup converges exponentially fast in Wasserstein distance $\mathcal{W}_d$ to the stationary distribution μ [Che1].

Let $\rho\in\mathcal{F}_+$ be an increasing function and define $u\in\mathcal{F}_+$ as $u_x := \rho(x+1)-\rho(x)$. The metric under consideration in the forthcoming analysis is

$$d_u(x,y) = |\rho(x)-\rho(y)|.$$

Hence u remains for the distance between two consecutive points. Then it is shown in [Che1, J] by coupling arguments that the Wasserstein curvature σ_u with respect to the distance d_u is given by

$$\sigma_u = \inf_{x\in\mathbb{N}} \, \nu_{x+1} - \nu_x \frac{u_{x-1}}{u_x} + \lambda_x - \lambda_{x+1}\frac{u_{x+1}}{u_x}.$$

The following corollary of theorem 2.1 allows to recover this result via an intertwining relation.

Corollary 3.1 (Contraction and curvature)**.** *Assume that the Wasserstein curvature is finite, i.e.* $\sigma_u > -\infty$*. Then with the notations of theorem 2.1, for any* $t\geq 0$*,*

$$\|P_t\|_{\mathrm{Lip}(d_u)\to\mathrm{Lip}(d_u)} = \|P_t\rho\|_{\mathrm{Lip}(d_u)} = \sup_{x\in\mathbb{N}} \mathbb{E}_x\left[\exp\left(-\int_0^t V_u(X_{u,s})\,ds\right)\right]. \tag{3.4}$$

In particular, the contraction inequality (3.3) *is satisfied with the optimal constant*

$$\sigma_u = \inf_{y\in\mathbb{N}} V_u(y).$$

Proof. Let $f \in \mathrm{Lip}_1(d_u)$ be a 1-Lipschitz function with respect to the distance d_u. For any $y, z \in \mathbb{N}$ such that $y < z$ (without loss of generality), we have by the intertwining identity (2.2) of theorem 2.1 and Jensen's inequality,

$$\begin{aligned} |P_t f(z) - P_t f(y)| &\leq \sum_{x=y}^{z-1} u_x \, |\partial_u P_t f(x)| \\ &\leq \sum_{x=y}^{z-1} u_x \, \mathbb{E}_x \left[|\partial_u f(X_{u,t})| \exp\left(- \int_0^t V_u(X_{u,s}) \, ds \right) \right] \\ &\leq d_u(z,y) \sup_{x \in \mathbb{N}} \mathbb{E}_x \left[\exp\left(- \int_0^t V_u(X_{u,s}) \, ds \right) \right], \end{aligned}$$

so that dividing by $d_u(z,y)$ and taking suprema entail the inequality:

$$\|P_t\|_{\mathrm{Lip}(d_u) \to \mathrm{Lip}(d_u)} \leq \sup_{x \in \mathbb{N}} \mathbb{E}_x \left[\exp\left(- \int_0^t V_u(X_{u,s}) \, ds \right) \right].$$

Finally, since by remark 2.2 the semigroup $(P_t)_{t \geq 0}$ propagates monotonicity, the right-hand-side of the latter inequality is nothing but $\|P_t \rho\|_{\mathrm{Lip}(d_u)}$, showing that the supremum over $\mathrm{Lip}_1(d_u)$ is attained for the function ρ. The proof of corollary 3.1 is achieved. □

Remark 3.2 (Pointwise gradient estimates for the Poisson equation)**.** *The argument used in the proof of corollary 3.1 allows also to obtain pointwise gradient estimates for the solution of the Poisson equation at the heart of Chen-Stein methods* [BHJ]*. More precisely, let us assume that d_u is such that $\rho \in L^1(\mu)$. For any centered function $f \in \mathrm{Lip}_1(d_u)$, let us consider the Poisson equation $-\mathcal{L} g = f$, where the unknown is g. Then under the assumption $\sigma_u > 0$, there exists a unique centered solution $g_f \in \mathrm{Lip}(d_u)$ to this equation given by the formula $g_f = \int_0^\infty P_t f \, dt$. We have for any $x \in \mathbb{N}$ the following estimate (compare with* [LM, Theorem 2.1]*):*

$$\begin{aligned} \sup_{f \in \mathrm{Lip}_1(d_u)} |\partial g_f(x)| &= \sup_{f \in \mathrm{Lip}_1(d_u)} u_x \int_0^\infty |\partial_u P_t f(x)| \, dt \\ &= u_x \int_0^\infty \partial_u P_t \rho(x) \, dt \\ &= u_x \int_0^\infty \mathbb{E}_x \left[\exp\left(- \int_0^t V_u(X_{u,s}) \, ds \right) \right] dt \\ &\leq \frac{u_x}{\sigma_u}. \end{aligned}$$

3.2. **Functional inequalities.** Theorems 2.1 and 2.4 allow to establish a whole family of discrete functional inequalities. We define the bilinear symmetric form Γ on $\mathcal{F}$ by

$$\Gamma(f,g) := \frac{1}{2} \left(\mathcal{L}(fg) - f \mathcal{L} g - g \mathcal{L} f \right) = \frac{1}{2} \left(\lambda \, \partial f \, \partial g + \nu \, \partial^* f \, \partial^* g \right).$$

The associated Dirichlet form acting on its domain $\mathcal{D}(\mathcal{E}_\mu) \times \mathcal{D}(\mathcal{E}_\mu)$ is given by

$$\mathcal{E}_\mu(f,g) := \frac{1}{2} \int_{\mathbb{N}} \Gamma(f,g)\, d\mu = \int_{\mathbb{N}} \lambda\, \partial f\, \partial g\, d\mu$$

where the second equality comes from the reversibility of the process. The stationary distribution μ is said to satisfy the Poincaré inequality with constant c if for any function $f \in \mathcal{D}(\mathcal{E}_\mu)$,

$$c\, \mathrm{Var}_\mu(f) \leq \mathcal{E}_\mu(f,f), \tag{3.5}$$

where $\mathrm{Var}_\mu(f) := \mu(f^2) - \mu(f)^2$ and $\mu(f) := \int_{\mathbb{N}} f\, d\mu$. The optimal (largest) constant c_{P} is the spectral gap of $\mathcal{L}$, i.e. the first non-trivial eigenvalue of the operator $-\mathcal{L}$. The constant c_{P} governs the $L^2(\mu)$ exponential decay to the equilibrium of the semigroup: for all $f \in L^2(\mu)$ and $t \geq 0$,

$$\|P_t f - \mu(f)\|_{L^2(\mu)} \leq e^{-c_{\mathrm{P}} t}\, \|f - \mu(f)\|_{L^2(\mu)}.$$

Chen uses in [Che1] a coupling method which provides the following formula for the spectral gap:

$$c_{\mathrm{P}} = \sup_{u \in \mathcal{F}_+} \sigma_u$$

where σ_u is the Wasserstein curvature of section 3.1. The following corollary of theorem 2.1 allows to recover the $\geq$ part of Chen's formula.

Corollary 3.3 (Spectral gap and Wasserstein curvatures)**.** *Assume that there exists some function $u \in \mathcal{F}_+$ such that the associated Wasserstein curvature σ_u is positive. Then the Poincaré inequality* (3.5) *holds with constant* $\sup_{u \in \mathcal{F}_+} \sigma_u$*, or in other words*

$$c_{\mathrm{P}} \geq \sup_{u \in \mathcal{F}_+} \sigma_u.$$

Proof. Since there exists some function $u \in \mathcal{F}_+$ such that the Wasserstein curvature σ_u is positive, the process is positive recurrent. Assuming without loss of generality that $f \in b\mathcal{F} \cap \mathcal{D}(\mathcal{E}_\mu)$, we have

$$\begin{aligned}
\mathrm{Var}_\mu(f) &= - \int_{\mathbb{N}} \int_0^\infty \frac{d}{dt} (P_t f)^2\, dt\, d\mu \\
&= -2 \int_{\mathbb{N}} \int_0^\infty P_t f\, \mathcal{L} P_t f\, dt\, d\mu \\
&= 2 \int_0^\infty \int_{\mathbb{N}} \lambda\, u^2\, (\partial_u P_t f)^2\, d\mu\, dt \\
&\leq 2 \int_0^\infty e^{-2\sigma_u t} \int_{\mathbb{N}} \lambda\, u^2\, P_{u,t} (\partial_u f)^2\, d\mu\, dt,
\end{aligned}$$

where in the last line we used theorem 2.1 with the convex function $\varphi(x) = x^2$. Now the measure $\lambda u^2 \mu$ is invariant for the semigroup $(P_{u,t})_{t\geq 0}$, so that we have

$$\begin{aligned}\mathrm{Var}_\mu(f) &\leq 2\int_0^\infty e^{-2\sigma_u t}\int_{\mathbb{N}} \lambda\, u^2\, (\partial_u f)^2\, d\mu\, dt\\ &= \frac{1}{\sigma_u}\int_{\mathbb{N}} \lambda\, (\partial f)^2\, d\mu\\ &= \frac{1}{\sigma_u}\,\mathcal{E}_\mu(f,f),\end{aligned}$$

where in the second line we used $\sigma_u > 0$. The proof of the Poincaré inequality is complete. □

Remark 3.4 ($M/M/\infty$ and $M/M/1$)**.** *The spectral gap of the $M/M/\infty$ and $M/M/1$ processes is well-known* [Che1]*. Corollary 3.3 allows to recover it easily. Indeed, in the $M/M/\infty$ case, the value $c_{\mathrm{P}} = \nu$ can be obtained as follows: choose the constant weight $u = 1$ to get $c_{\mathrm{P}} \geq \nu$, and notice that the equality holds for affine functions. For a positive recurrent $M/M/1$ process, i.e. $\lambda < \nu$, we obtain $c_{\mathrm{P}} \geq (\sqrt{\lambda} - \sqrt{\nu})^2$ by choosing the weight $u_x := (\nu/\lambda)^{x/2}$, whereas the equality asymptotically holds in* (3.5) *as $\kappa \to \sqrt{\nu/\lambda}$ for the functions κ^x, $x \in \mathbb{N}$. We conclude that $c_{\mathrm{P}} = (\sqrt{\lambda} - \sqrt{\nu})^2$.*

Theorem 2.4 allows to derive functional inequalities more general than the Poincaré inequality. Let $\mathcal{I}$ be an open interval of $\mathbb{R}$ and for a convex function $\varphi : \mathcal{I} \to \mathbb{R}$ we define the φ-entropy of $f : \mathbb{N} \to \mathcal{I}$ as

$$\mathrm{Ent}_\mu^\varphi(f) = \mu\left(\varphi(f)\right) - \varphi\left(\mu(f)\right).$$

Following [Cha], we say that the stationary distribution μ satisfies a φ-entropy inequality with constant $c > 0$ if for any $\mathcal{I}$-valued function $f \in \mathcal{D}(\mathcal{E}_\mu)$ such that $\varphi'(f) \in \mathcal{D}(\mathcal{E}_\mu)$,

$$c\,\mathrm{Ent}_\mu^\varphi(f) \leq \mathcal{E}_\mu\left(f, \varphi'(f)\right). \tag{3.6}$$

See for instance [Cha] for an investigation of the properties of φ-entropies. The φ-entropy inequality (3.6) is satisfied if and only if the following entropy dissipation of the semigroup holds: for any sufficiently integrable $\mathcal{I}$-valued function f and every $t \geq 0$,

$$\mathrm{Ent}_\mu^\varphi(P_t f) \leq e^{-ct}\,\mathrm{Ent}_\mu^\varphi(f).$$

We have the following corollary of theorem 2.4.

Corollary 3.5 (Entropic inequalities and Wasserstein curvature)**.** *If the birth rate λ is non-increasing and the Wasserstein curvature σ_1 (with the constant weight $u = 1$) is positive, then the φ-entropy inequality* (3.6) *holds with constant σ_1.*

Proof. The assertion $\sigma_1 > 0$ entails the positive recurrence of the process. By reversibility, we have for any $\mathcal{I}$-valued function $f \in b\mathcal{F} \cap \mathcal{D}(\mathcal{E}_\mu)$ such that $\varphi'(f) \in$

$\mathcal{D}(\mathcal{E}_\mu)$,

$$\begin{aligned}
\mathrm{Ent}_\mu^\varphi(f) &= \int_{\mathbb{N}} \left(\varphi(P_0 f) - \varphi(\mu(f))\right) d\mu \\
&= -\int_{\mathbb{N}} \int_0^\infty \frac{d}{dt} \varphi(P_t f)\, dt\, d\mu \\
&= -\int_0^\infty \int_{\mathbb{N}} \varphi'(P_t f)\, \mathcal{L} P_t f\, d\mu\, dt \\
&= \int_0^\infty \int_{\mathbb{N}} \lambda\, \partial P_t f\, \partial \varphi'(P_t f)\, d\mu\, dt \\
&= \int_0^\infty \int_{\mathbb{N}} \lambda\, B^\varphi\left(P_t f, \partial P_t f\right) d\mu\, dt,
\end{aligned}$$

where B^φ is as in theorem 2.4. Using now theorem 2.4 together with the invariance of the measure $\lambda\mu$ for the 1-modification semigroup $(P_{1,t})_{t\geq 0}$, we obtain

$$\begin{aligned}
\mathrm{Ent}_\mu^\varphi(f) &\leq \int_0^\infty \int_{\mathbb{N}} e^{-\sigma_1 t}\, \lambda\, P_{1,t} B^\varphi\left(f, \partial f\right) d\mu\, dt \\
&= \int_0^\infty \int_{\mathbb{N}} e^{-\sigma_1 t}\, \lambda\, B^\varphi\left(f, \partial f\right) d\mu\, dt \\
&= \frac{1}{\sigma_1} \int_{\mathbb{N}} \lambda\, B^\varphi\left(f, \partial f\right) d\mu \\
&= \frac{1}{\sigma_1}\, \mathcal{E}_\mu\left(f, \varphi'(f)\right).
\end{aligned}$$

□

Remark 3.6 (Examples of entropic inequalities)**.** *The constant in the φ-entropy inequality provided by corollary 3.5 is not optimal in general (compare for instance with the Poincaré inequality of corollary 3.3 when $\varphi(r) = r^2$ with $\mathcal{I} = \mathbb{R}$). The choice $\varphi(r) = r \log r$ with $\mathcal{I} = (0,\infty)$ allows us to recover the modified* log*-Sobolev inequality of* [CDPP, Theorem 3.1]*: for any positive function $f \in \mathcal{D}(\mathcal{E}_\mu)$ such that* $\log f \in \mathcal{D}(\mathcal{E}_\mu)$,

$$\sigma_1\, \mathrm{Ent}_\mu^\varphi(f) \leq \mathcal{E}_\mu(f, \log f). \tag{3.7}$$

For the $M/M/\infty$ process, the estimate of corollary 3.5 is sharp since $\sigma_1 = \nu$ and the equality in (3.7) *holds as $\alpha \to \infty$ for the function $x \in \mathbb{N} \mapsto e^{\alpha x}$. Note that the $M/M/1$ process and its invariant distribution, which is geometric, do not satisfy a modified* log*-Sobolev inequality. Another φ-entropy inequality of interest is that obtained when considering the convex function $\phi(r) := r^p$, $p \in (1,2]$, with $\mathcal{I} = (0,\infty)$: for any positive function $f \in \mathcal{D}(\mathcal{E}_\mu)$ such that $f^{p-1} \in \mathcal{D}(\mathcal{E}_\mu)$,*

$$\mu(f^p) - \mu(f)^p \leq \frac{p}{\sigma_1}\, \mathcal{E}_\mu(f, f^{p-1}). \tag{3.8}$$

Such an inequality has been studied in [BT] *in the case of Markov processes on a finite state space and also in* [Cha] *for the $M/M/\infty$ queueing process. In particular, it can be seen as an interpolation between Poincaré and modified* log*-Sobolev inequalities.*

Theorem 2.1 implies also other type of functional inequalities such as discrete isoperimetry and transportation-information inequalities. Given a positive function u, we focus on the distance d_u constructed in section 3.1, where we assume moreover that $\rho \in \mathcal{D}(\mathcal{E}_\mu)$. The invariant measure μ is said to satisfy a weighted isoperimetric inequality with weight u and constant $h_u > 0$ if for any absolutely continuous probability measure π with density $f \in \mathcal{D}(\mathcal{E}_\mu)$ with respect to μ,

$$h_u \, \mathcal{W}_{d_u}(\pi, \mu) \leq \int_{\mathbb{N}} \lambda \, u \, |\partial f| \, d\mu, \tag{3.9}$$

where the Wasserstein distance $\mathcal{W}_{d_u}$ is defined in (3.1) with respect to the distance d_u. The terminology of isoperimetry is employed here because it is a generalization of the classical isoperimetry, which states that the centered L^1-norm is dominated by an energy of L^1-type. Indeed, if the weight u is identically 1, then the distance d_1 between two different points is at least 1, so that (3.9) entails

$$2h_1 \int_{\mathbb{N}} |f - 1| \, d\mu = h_1 \, \mathcal{W}_d(\pi, \mu) \leq h_1 \, \mathcal{W}_{d_1}(\pi, \mu) \leq \int_{\mathbb{N}} \lambda \, |\partial f| \, d\mu,$$

where d is the trivial distance 0 or 1. Note that the L^1-energy emphasized above differs from the discrete version of the diffusion case, since our discrete gradient does not derive from Γ.

On the other hand, let us introduce the transportation-information inequalities emphasized in [GLWY]. Let α be a continuous positive and increasing function on $[0, \infty)$ vanishing at 0. The invariant measure μ satisfies a transportation-information inequality with deviation function α if for any absolutely continuous probability measure π with density f with respect to μ, we have

$$\alpha \left(\mathcal{W}_{d_u}(\pi, \mu) \right) \leq \mathcal{I}(\pi, \mu), \tag{3.10}$$

where the so-called Fisher-Donsker-Varadhan information of π with respect to μ is defined as

$$\mathcal{I}(\pi, \mu) := \begin{cases} \mathcal{E}_\mu(\sqrt{f}, \sqrt{f}) & \text{if } \sqrt{f} \in \mathcal{D}(\mathcal{E}_\mu); \\ \infty & \text{otherwise.} \end{cases}$$

Note that $\mathcal{I}(\cdot, \mu)$ is nothing but the rate function governing the large deviation principle in large time of the empirical measure $L_t := t^{-1} \int_0^t \delta_{X_s} \, ds$, where δ_x is the Dirac mass at point x. In other words, the Fisher-Donsker-Varadhan information rewrites as the variational identity [Che2, Theorem 8.8]:

$$\mathcal{I}(\pi, \mu) = \sup_{V \in \mathcal{F}_+} \int_{\mathbb{N}} -\frac{\mathcal{L} V}{V} \, d\pi.$$

The interest of the transportation-information inequality resides in the equivalence with the following tail estimate of the empirical measure [GLWY, Theorem 2.4]: for any absolutely continuous probability measure π with density $f \in L^2(\mu)$ with respect to μ, and any $g \in \mathrm{Lip}_1(d_u)$,

$$\mathbb{P}_\pi\left(L_t(g) - \mu(g) > r\right) \leq \|f\|_{L^2(\mu)}\, e^{-\alpha(r)}, \quad r > 0, \quad t > 0.$$

We have the following corollary of theorem 2.1.

Corollary 3.7 (Weighted isoperimetry and transportation-information inequality). *With the notations of theorem 2.1, assume that the following quantity is well-defined:*

$$\kappa_u := \int_0^\infty \sup_{x\in\mathbb{N}} \mathbb{E}_x\left[\exp\left(-\int_0^t V_u(X_{u,s})\,ds\right)\right] dt < \infty.$$

Then the weighted isoperimetric inequality (3.9) is satisfied with constant $h_u = 1/\kappa_u$.
If moreover there exists two constants $\varepsilon > 0$ and $\theta > 1$ such that

$$(1+\varepsilon)\lambda_x u_x^2 + (1+1/\varepsilon)\nu_x u_{x-1}^2 \leq -a\left(\lambda_x(\theta-1) + \nu_x(1/\theta - 1)\right) + b, \quad x \in \mathbb{N}, \tag{3.11}$$

where $a := a_{\varepsilon,\theta} \geq 0$ and $b := b_{\varepsilon,\theta} > 0$ are two other constants depending on both ε and θ, then the transportation-information inequality (3.10) is satisfied with deviation function

$$\alpha(r) := \sup_{\varepsilon>0,\theta>1} \frac{\sqrt{b^2 + 2a(r/\kappa_u)^2} - b}{2a}.$$

Remark 3.8 (The case of positive Wasserstein curvature). *In particular if the Wasserstein curvature σ_u with respect to the distance d_u is positive, then under the same assumptions as above,*

$$\sigma_u\, \mathcal{W}_{d_u}(\pi,\mu) \leq \int_{\mathbb{N}} \lambda\, u\, |\partial f|\, d\mu \quad \text{and} \quad \alpha\left(\mathcal{W}_{d_u}(\pi,\mu)\right) \leq \mathcal{I}(\pi,\mu),$$

with the deviation function

$$\alpha(r) := \sup_{\varepsilon>0,\theta>1} \frac{\sqrt{b^2 + 2a(r\sigma_u)^2} - b}{2a}.$$

Proof. For every $f, g \in \mathcal{D}(\mathcal{E}_\mu)$ we have, by reversibility,

$$\begin{aligned}
\mathrm{Cov}_\mu(f,g) &:= \int_{\mathbb{N}} \left(g - \int_{\mathbb{N}} g\,d\mu\right) f\,d\mu \\
&= \int_{\mathbb{N}} \left(-\int_0^\infty \mathcal{L}P_t g\,dt\right) f\,d\mu \\
&= \int_0^\infty \left(-\int_{\mathbb{N}} P_t g\,\mathcal{L}f\,d\mu\right) dt \\
&= \int_0^\infty \mathcal{E}_\mu(P_t g, f)\,dt.
\end{aligned} \tag{3.12}$$

Now, for every probability measure $\pi \ll \mu$ with $d\pi = f d\mu$, $f \in b\mathcal{F} \cap \mathcal{D}(\mathcal{E}_\mu)$, we get, using (3.12),

$$\begin{aligned}
\mathcal{W}_{d_u}(\pi,\mu) &= \sup_{g\in \mathrm{Lip}_1(d_u)} \mathrm{Cov}_\mu(f,g) \\
&= \sup_{g\in \mathrm{Lip}_1(d_u)} \int_0^\infty \mathcal{E}_\mu(P_t g, f)\, dt \\
&= \sup_{g\in \mathrm{Lip}_1(d_u)} \int_0^\infty \int_{\mathbb{N}} \lambda\, u\, \partial f\ \partial_u P_t g\, d\mu\, dt \\
&= \int_0^\infty \int_{\mathbb{N}} \lambda\, u\, |\partial f|\ \partial_u P_t \rho\, d\mu\, dt \\
&\leq \int_0^\infty \sup_{x\in\mathbb{N}} \mathbb{E}_x \left[\exp\left(- \int_0^t V_u(X_{u,s})\, ds \right)\right] dt \int_{\mathbb{N}} \lambda\, u\, |\partial f|\, d\mu,
\end{aligned}$$

where in the last inequality we used theorem 2.1. This concludes the proof of the weighted isoperimetric inequality.

Using now Cauchy-Schwarz inequality, reversibility and then (3.11) with $V_\theta(x) := \theta^x$, $x \in \mathbb{N}$,

$$\begin{aligned}
\mathcal{W}_{d_u}(\pi,\mu) &\leq \kappa_u \sqrt{\mathcal{I}(\pi,\mu)} \sqrt{\int_{\mathbb{N}} \lambda\, u^2 \left(\sqrt{f(\cdot+1)} + \sqrt{f} \right)^2 d\mu} \\
&\leq \kappa_u \sqrt{\mathcal{I}(\pi,\mu)} \sqrt{\int_{\mathbb{N}} ((1+\varepsilon)\lambda u^2 + (1+1/\varepsilon)\nu u_{\cdot-1}^2)\, f\, d\mu} \\
&\leq \kappa_u \sqrt{\mathcal{I}(\pi,\mu)} \sqrt{\int_{\mathbb{N}} \left(-a \frac{\mathcal{L}V_\theta}{V_\theta} + b \right) f\, d\mu} \\
&\leq \kappa_u \sqrt{\mathcal{I}(\pi,\mu)} \sqrt{a\mathcal{I}(\pi,\mu) + b},
\end{aligned}$$

from which the desired transportation-information inequality holds. □

Remark 3.9 ($M/M/\infty$ and $M/M/1$ revisited)**.** *Corollary 3.7 exhibits optimal functional inequalities, at least in the $M/M/\infty$ case and its stationary distribution, the Poisson measure of mean λ/ν. Choosing the weight $u = 1$, we obtain the optimal constant $\hbar_1 = \nu$ in the isoperimetric inequality. Indeed, corollary 3.7 entails $\hbar_1 \geq \nu$, whereas the other inequality is obtained by choosing π a Poisson measure of different parameter. For the transportation-information inequality, we recover theorem 2.1 in* [MWW] *since the choice of $a := \theta(1+1/\varepsilon)/(\theta-1)$ and $b := \lambda(1+\varepsilon+(1+1/\varepsilon)\theta)$ allows us to obtain the deviation function $\alpha(r) := \lambda(\sqrt{1+\nu r/\lambda}-1)^2$, $r > 0$. Note that it is optimal in view of Example 4.5 in* [GGW]*: for any absolutely continuous probability measure π with square-integrable density with respect to μ,*

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\pi \left(\frac{1}{t} \int_0^t X_s\, ds - \frac{\lambda}{\nu} > r \right) = -\lambda \left(\sqrt{1 + \frac{\nu r}{\lambda}} - 1 \right)^2, \quad r > 0.$$

For the $M/M/1$ process, we have the following inequalities for the optimal isoperimetric constant $\hbar_u$, with $u_x = (\nu/\lambda)^{x/2}$ (a quantity that will appear again in section 3.3):

$$(\sqrt{\lambda} - \sqrt{\nu})^2 \le \hbar_u \le (\sqrt{\nu} - \sqrt{\lambda})\sqrt{\nu}.$$

To get the second inequality, we choose the density $f = (\nu/\lambda)(1 - 1_{\{0\}})$ and the 1-Lipschitz test function $g = \rho$. In particular as the ratio λ/ν is small, we obtain $\hbar_u \approx \nu$. However, we ignore if such a process satisfies a transportation-information inequality.

3.3. **Hitting time of the origin by the $M/M/1$ process.** Recall that we consider the ergodic $M/M/1$ process ($\lambda < \nu$) for which the stationary distribution is geometric of parameter λ/ν. Since the process behaves as a random walk outside 0, the ergodic property relies essentially on its behavior at point 0. Using the notation of theorem 2.1, the intertwining relation (2.2) applied with a positive function u entails the identity

$$\partial_u P_t f(x) = \mathbb{E}_x \left[\partial_u f(X_t) \, \exp\left(- \int_0^t V_u(X_{u,s}) \, ds \right) \right]$$

where the potential is given for every $x \in \mathbb{N}$ by

$$V_u(x) := \nu - \frac{u_{x-1}}{u_x} \nu \mathbf{1}_{\{x \neq 0\}} + \lambda - \frac{u_{x+1}}{u_x} \lambda.$$

Following Robert [R], the process $(X_t^y)_{t \ge 0}$ is the solution of the stochastic differential equation

$$X_0^y = y \quad \text{and} \quad dX_t^y = dN_t^{(\lambda)} - \mathbf{1}_{\{X_{t-}^y > 0\}} \, dN_t^{(\nu)}, \quad t > 0, \tag{3.13}$$

where $(N_t^{(\lambda)})_{t \ge 0}$ and $(N_t^{(\nu)})_{t \ge 0}$ are two independent Poisson processes with parameter λ and ν, respectively. Since the process is assumed to be positive recurrent, the hitting time of 0,

$$T_0^y := \inf\{t > 0 : X_t^y = 0\}$$

is finite almost surely. We have the following corollary of theorem 2.1.

Corollary 3.10 (Hitting time of the origin for the ergodic $M/M/1$ process)**.** *Given $x \in \mathbb{N}$, consider a positive recurrent $M/M/1$ process $(X_t^{x+1})_{t \ge 0}$ starting at point $x+1$, and denote $(X_{u,t}^x)_{t \ge 0}$ its u-modification process starting at point x, where*

$$u_x := \left(\frac{\nu}{\lambda} \right)^{\frac{x}{2}} \ge 1.$$

Then we have the following tail estimate: for any $t \ge 0$,

$$\begin{aligned} \mathbb{P}(T_0^{x+1} > t) &= u_x \, e^{-t \left(\sqrt{\lambda} - \sqrt{\nu}\right)^2} \, \mathbb{E} \left[\frac{1}{u(X_{u,t}^x)} \, \exp\left(-\sqrt{\lambda\nu} \int_0^t \mathbf{1}_{\{0\}}(X_{u,s}^x) \, ds \right) \right] \\ &\le u_x \, e^{-t \left(\sqrt{\lambda} - \sqrt{\nu}\right)^2}. \end{aligned}$$

Proof. Let us use a coupling argument. Let $(X_t^x)_{t\geq 0}$ be a copy of $(X_t^{x+1})_{t\geq 0}$, starting at point x. We assume that it constructed with respect to the same driving Poisson processes $(N_t^{(\lambda)})_{t\geq 0}$ and $(N_t^{(\nu)})_{t\geq 0}$ as the process $(X_t^{x+1})_{t\geq 0}$. Hence the stochastic differential equation (3.13) satisfied by the two coupling processes entails that the difference between $(X_t^{x+1})_{t\geq 0}$ and $(X_t^x)_{t\geq 0}$ remains constant, equal to 1, until time T_0^{x+1}, the first hitting time of the origin by $(X_t^{x+1})_{t\geq 0}$. After time T_0^{x+1}, the processes are identically the same, so that the following identity holds:

$$X_t^{x+1} = X_t^x + \mathbf{1}_{\{T_0^{x+1} > t\}}, \quad t \geq 0.$$

Since the original process is assumed to be positive recurrent, the coupling is successful, i.e. the coupling time is finite almost surely. Therefore we have for any function $f \in b\mathcal{F} \cup \mathrm{Lip}(d_1)$, where d_1 is the distance $d_1(x,y) = |x-y|$,

$$\partial P_t f(x) = P_t f(x+1) - P_t f(x) = \mathbb{E}\left[f(X_t^{x+1}) - f(X_t^x)\right] = \mathbb{E}\left[\partial f(X_t^x)\,\mathbf{1}_{\{T_0^{x+1}>t\}}\right]$$

so that if we denote the function $\rho(x) = x$, we obtain

$$\mathbb{P}(T_0^{x+1} > t) = \partial P_t \rho(x) = u_x\, \partial_u P_t \rho(x).$$

Using now (2.2) with the function u, we get

$$\mathbb{P}(T_0^{x+1} > t) = u_x\, \mathbb{E}\left[\frac{1}{u(X_{u,t}^x)} \exp\left(-\int_0^t V_u(X_{u,s}^x)\, ds\right)\right],$$

where $V_u := (\sqrt{\lambda} - \sqrt{\nu})^2 + \sqrt{\lambda\nu}\,\mathbf{1}_{\{0\}}$. □

Remark 3.11 (Sharpness)**.** *Using a completely different approach, Van Doorn established in* [VD]*, through his theorem 4.2 together with his Example 5, the following asymptotics*

$$\lim_{t\to\infty} \frac{1}{t} \log \mathbb{P}(T_0^{x+1} > t) = -(\sqrt{\lambda} - \sqrt{\nu})^2, \quad x \in \mathbb{N}.$$

Hence one deduces that the exponential decay in the result of corollary 3.10 is sharp. On the other hand, Proposition 5.4 in [R] *states that* T_0^{x+1} *has exponential moment bounded as follows:*

$$\mathbb{E}\left[e^{(\sqrt{\lambda}-\sqrt{\nu})^2 T_0^{x+1}}\right] \leq \left(\frac{\nu}{\lambda}\right)^{(x+1)/2},$$

so that Chebyshev's inequality yields a tail estimate somewhat similar to ours - although with a worst constant depending on the initial point $x+1$*.*

Remark 3.12 (Other approach)**.** *The proof of corollary 3.10 suggests also a martingale approach. First, note that we have the identity*

$$-\nu\,\mathbf{1}_{\{0\}} = -\frac{\mathcal{L}u}{u} - V_u$$

which entails as in the previous proof and since $u \geq 1$, the following computations:

$$\begin{aligned}\mathbb{P}(T_0^{x+1} > t) &= \partial P_t \rho(x) \\ &= \mathbb{E}\left[\exp\left(-\int_0^t \nu\, \mathbf{1}_{\{0\}}(X_s^x)\, ds\right)\right] \\ &\leq \mathbb{E}\left[u(X_t^x)\, \exp\left(-\int_0^t \left(\frac{\mathcal{L}u}{u} + V_u\right)(X_s^x)\, ds\right)\right] \\ &\leq u_x\, e^{-t\,(\sqrt{\lambda}-\sqrt{\nu})^2},\end{aligned}$$

since the process $(M_t^u)_{t\geq 0}$ given by

$$M_t^u := u(X_t^x)\, \exp\left(-\int_0^t \frac{\mathcal{L}u}{u}\,(X_s^x)\, ds\right), \quad t \geq 0,$$

is a supermartingale. Indeed, denoting

$$Z_t^u := \exp\left(-\int_0^t \frac{\mathcal{L}u}{u}\,(X_s^x)\, ds\right),$$

we have by Ito's formula:

$$\begin{aligned}dM_t^u &= Z_t^u du(X_t^x) + u(X_t^x)\, dZ_t^u \\ &= Z_t^u(dM_t + \mathcal{L}u(X_t^x)dt) - u(X_t^x)\frac{\mathcal{L}u}{u}\,(X_t^x)\, Z_t^u dt \\ &= Z_t^u dM_t,\end{aligned}$$

where $(M_t)_{t\geq 0}$ is a local martingale. Therefore, the process $(M_t^u)_{t\geq 0}$ is a positive local martingale and thus a supermartingale.

3.4. **Convex domination of birth-death processes.** Let $(X_t^x)_{t\geq 0}$ be the $M/M/\infty$ process starting from $x \in \mathbb{N}$. The Mehler-type formula (1.2) states that the random variable X_t^x has the same distribution as the independent sum of the variable X_t^0, which follows the Poisson distribution of parameter $\rho(1-e^{-\nu t})$, and a binomial random variable $B_t^{(x)}$ of parameters $(x, e^{-\nu t})$. By convention, $B_t^{(0)}$ is assumed to be 0 in the sequel. Hence we have for any function $f \in b\mathcal{F} \cup \mathcal{F}_+$ and any $x \in \mathbb{N}$,

$$\mathbb{E}\left[f(X_t^x)\right] = \mathbb{E}[f(X_t^0 + B_t^{(x)})], \quad t \geq 0. \tag{3.14}$$

Such an identity can be provided by using the commutation relation (1.4). Indeed we have

$$\mathbb{E}\left[f(X_t^{x+1})\right] = (1 - e^{-\nu t})\, \mathbb{E}\left[f(X_t^x)\right] + e^{-\nu t}\, \mathbb{E}[f(X_t^x + 1)],$$

so that a recursive argument on the initial state provides the required result. Therefore, one may ask if for a more general birth-death process, the intertwining relation of type (2.2) may imply a relation similar to (3.14). This leads the notion of stochastic ordering. Given two random variables X, Y, we say that X is convex dominated by Y, and we note $X \leq_{\mathrm{c}} Y$, if for any non-decreasing convex function $f \in \mathcal{F}_+$, we have

$$\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)].$$

Here the convexity of f is understood as $\partial^2 f \geq 0$. We may deduce typically from the convex domination a comparison of moments or Laplace transforms of X and Y. Coming back to our birth-death framework, we observe that if we want to use the intertwining relation (2.2) of theorem 2.1 in order to obtain stochastic domination, then a first difficulty arises. Indeed, another birth-death process appears in the right-hand-side of (2.2), namely the u-modification of the original process. Therefore, let us provide first a lemma which allows us to compare stochastically two birth-death processes.

Lemma 3.13 (Stochastic comparison of birth-death processes). *Let $(X_t^x)_{t\geq 0}$ and $(\tilde{X}_t^x)_{t\geq 0}$ be two birth-death processes both starting from $x \in \mathbb{N}$. Denoting respectively λ, ν and $\tilde{\lambda}, \tilde{\nu}$ the transition rates of the associated generators $\mathcal{L}$ and $\tilde{\mathcal{L}}$, we assume that they satisfy the following assumption:*

$$\tilde{\lambda} \leq \lambda \quad \textit{and} \quad \tilde{\nu} \geq \nu$$

Then the result reads as follows: for any non-decreasing function $g \in \mathcal{F}_+$, we have, for every $t \geq 0$,

$$\mathbb{E}\Big[g(\tilde{X}_t^x)\Big] \leq \mathbb{E}[g(X_t^x)].$$

Proof. Let us define $s \in [0,t] \mapsto J(s) := \tilde{P}_s P_{t-s} g$ where $(P_t)_{t\geq 0}$ and $(\tilde{P}_t)_{t\geq 0}$ are the semigroups of $(X_t^x)_{t\geq 0}$ and $(\tilde{X}_t^x)_{t\geq 0}$ respectively. By differentiation, we have

$$J'(s) = \tilde{P}_s\left(\tilde{\mathcal{L}} P_{t-s} g - \mathcal{L} P_{t-s} g\right) = \tilde{P}_s\left((\tilde{\lambda} - \lambda)\, \partial P_{t-s} g + (\tilde{\nu} - \nu)\, \partial^* P_{t-s} g\right),$$

which is non-positive since the semigroup $(P_t)_{t\geq 0}$ satisfies the propagation of monotonicity, cf. remark 2.2. Hence the function J is non-increasing and the desired result holds. □

Now we are able to state the following corollary of theorem 2.1.

Corollary 3.14 (Convex domination). *Let $(X_t^x)_{t\geq 0}$ be a birth-death process starting from $x \in \mathbb{N}$. We assume that the birth rate λ is non-increasing and that there exists $\kappa > 0$ such that*

$$\partial(\nu - \lambda) \geq \kappa.$$

Then for any $t \geq 0$, the random variable X_t^x is convex dominated by the independent sum of X_t^0 and a binomial random variable $B_t^{(x)}$ of parameters $(x, e^{-\kappa t})$. In other words, we have

$$X_t^x \leq_{\mathrm{c}} X_t^0 + B_t^{(x)},$$

as in the case of the $M/M/\infty$ queueing process.

Proof. We have to show that for any non-decreasing convex function $f \in \mathcal{F}_+$ and any $t \geq 0$,

$$\mathbb{E}[f(X_t^x)] \leq \mathbb{E}\Big[f(X_t^0 + B_t^{(x)})\Big]. \tag{3.15}$$

Let us provide a simple recursive proof. First, note that at rank $x = 0$, the result holds trivially since $B_t^{(0)} = 0$ by convention. Assume now that (3.15) is verified at rank x. Using then the intertwining relation (2.2) of theorem 2.1 and then lemma 3.13 with the 1-modification $(X_{1,t})_{t\geq 0}$ of the process $(X_t)_{t\geq 0}$, we have:

$$\begin{aligned}
\mathbb{E}\Big[f(X_t^{x+1})\Big] &\leq \mathbb{E}[f(X_t^x)] + e^{-\kappa t}\,\mathbb{E}\Big[\partial f(X_{1,t}^x)\Big] \\
&\leq \mathbb{E}[f(X_t^x)] + e^{-\kappa t}\,\mathbb{E}[\partial f(X_t^x)] \\
&= (1-e^{-\kappa t})\,\mathbb{E}[f(X_t^x)] + e^{-\kappa t}\,\mathbb{E}[f(X_t^x+1)] \\
&\leq (1-e^{-\kappa t})\,\mathbb{E}\Big[f(X_t^0 + B_t^{(x)})\Big] + e^{-\kappa t}\,\mathbb{E}\Big[f(X_t^0 + B_t^{(x)} + 1)\Big] \\
&= \mathbb{E}\Big[f(X_t^0 + B_t^{(x)} + B_t^{(1)})\Big] \\
&= \mathbb{E}\Big[f(X_t^0 + B_t^{(x+1)})\Big],
\end{aligned}$$

where we used for the last inequality the fact that the functions f and $f(\cdot+1)$ are non-decreasing and convex on $\mathbb{N}$. Therefore, (3.15) is established at rank $x+1$, hence in full generality. □

Acknowledgements. The authors are grateful to Arnaud Guillin and Laurent Miclo for their remarks during the ANR EVOL meeting held in Hammamet (2010).

References

[B] D. Bakry, On Sobolev and logarithmic Sobolev inequalities for Markov semigroups, *New trends in stochastic analysis*, Charingworth 1994: 43-75, World Sci. Publ., River Edge, 1997.

[BE] D. Bakry and M. Émery, Diffusion hypercontractives, *Séminaire de Probabilités* XIX, 1983/84, *Lecture Notes in Math.* 1123:177-206, Springer, Berlin, 1985.

[BHJ] A. D. Barbour, L. Holst, and S. Janson, *Poisson approximation*, Oxford Studies in Probability (2), The Clarendon Press Oxford University Press, New York, 1992.

[BT] S. Bobkov and P. Tetali, Modified logarithmic Sobolev inequalities in discrete settings, *J. Theor. Probab.*, 19(2):289-336, 2006.

[BX] T. C. Brown and A. Xia, Stein's method and birth-death processes, *Ann. Probab.*, 29(3):1373-1403, 2001.

[CDPP] P. Caputo, P. Dai Pra and G. Posta, Convex entropy decay via the Bochner-Bakry-Emery approach, *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(3):734-753, 2009.

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

[Che1] M. F. Chen, Estimation of spectral gap for Markov chains, *Acta Math. Sin.*, 12(4):337-360, 1996.

[Che2] M. F. Chen, *From Markov chains to non-equilibrium particle systems*, World Scientific Publishing Co. Inc., River Edge, NJ, 2004.

[GGW] F. Gao, A. Guillin and L. Wu, Bernstein type's concentration inequalities for symmetric Markov processes, Preprint, 2010.

[GLWY] A. Guillin, C. Léonard, L. Wu and N. Yao, Transportation-information inequalities for Markov processes, *Probab. Theory Related Fields*, 144(3-4):669-695, 2009.

[J] A. Joulin, A new Poisson-type deviation inequality for Markov jump processes with positive Wasserstein curvature, *Bernoulli*, 15(2):532-549, 2009.

[L] M. Ledoux, The geometry of Markov diffusion generators, *Ann. Fac. Sci. Toulouse Math.* (6) 9(2): 305-366, (2000)

[LM] W. Liu and Y. Ma, Spectral gap and convex concentration inequalities for birth-death processes, *Ann. Inst. H. Poincaré Probab. Statist.*, 45(1):58-69, 2009.

[MWW] Y. Ma, R. Wang and L. Wu, Transportation-information inequalities for continuum Gibbs measures, Preprint, 2010.

[MT] F. Malrieu and D. Talay, Concentration inequalities for Euler schemes, *Monte Carlo and Quasi-Monte Carlo Methods* 2004: 355-371, Springer-Verlag, 2006.

[R] P. Robert, *Stochastic networks and queues*, Stochastic Modelling and Applied Probability Series, Springer-Verlag, New York, 2003.

[S] D. Stoyan, *Comparison methods for queues and other stochastic models*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley and Sons, Chichester, 1983.

[VD] E. A. Van Doorn, On associated polynomials and decay rates for birth-death processes, *J. Math. Anal. Appl.*, 278, 500-511, 2003.

[V] C. Villani, *Optimal transport: old and new*, Grundlehren der mathematischen Wissenschaften, Springer, Berlin, 2009.

(D. Chafaï) Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS UMR8050, Université Paris-Est Marne-la-Vallée, France

E-mail address: `mailto:djalil(at)chafai.net`

URL: `http://djalil.chafai.net/`

(A. Joulin, corresponding author) Université de Toulouse, Institut National des Sciences Appliquées de Toulouse, Institut de Mathématiques de Toulouse, CNRS UMR5219, France

E-mail address: `ajoulin(at)insa-toulouse.fr`

URL: `http://www-gmm.insa-toulouse.fr/~ajoulin/`