Electron. Commun. Probab. **21** (2016), no. 19, 1–13.
DOI: 10.1214/16-ECP4762
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Bridges of Markov counting processes: quantitative estimates

Giovanni Conforti*

Abstract

In this paper we investigate the behavior of the bridges of a Markov counting process in several directions. We first characterize convexity(concavity) in time of the mean value in terms of lower (upper) bounds on the so called *reciprocal characteristics*. This result gives a natural criterion to determine whether bridges are "lazy" or "hurried". Under the hypothesis of global bounds on the reciprocal characteristics we prove sharp estimates for the marginal distributions and a comparison theorem for the jump times. When the height of the bridge tends to infinity we show the convergence to a deterministic curve, after a proper rescaling.

Keywords: bridges; counting processes; duality formula; tail estimates.
AMS MSC 2010: 60J27; 60J75.
Submitted to ECP on December 14, 2015, final version accepted on February 16, 2016.

Introduction

Counting processes are among the simplest stochastic processes but still feature a very broad set of applications in social sciences and engineering, see [1],[3],[7]. Moreover, some of the prototypes of jump processes, such as the Poisson processes, fall into this class, and the development of a stochastic calculus adjusted to these (and more general) pure jump processes led to important contributions, see [2],[8],[11]. In this paper we give a new criterion which allows to quantify the behavior of the *bridges* of a Markov counting process. One of the most popular interpretations is that counting processes model the number of people arriving in a line. Using this interpretation, our aim is to find robust and easy-to-check conditions on the intensity of the counting process under which we can pull out quantitative information on the conditional distribution of the i-th arrival time, or on how many customers have arrived by time t, given that n customers arrived in a day.

Despite the simplicity of the non pinned dynamics, bridges tend to be more complicated to study because their jump intensity is time-dependent and not explicit, but depends on the solution to a Kolmogorov-type equation, which is in general not available. Moreover, jump intensities and bridges are not in one-to-one correspondence, in the sense that different counting processes give raise to the same family of bridges. A well known example of this is the fact that all Poisson processes, independently from their jump intensity, have identical bridges. Good estimates have to take into account only those features of the jump intensity which play an effective role in the construction of bridges.

*University of Leipzig, Germany. E-mail: giovanniconfort@gmail.com

Our answers are inspired by the theory of reciprocal processes, which provides the main computational tool: the so called *reciprocal characteristics*. They should be thought at some invariants associated with bridges. Reciprocal characteristics have been first discovered for diffusions in [9] and are by now available for a much larger class of processes, including several classes of jump processes, see e.g.[5]. However, to the best of our knowledge, it is for the first time in this note that such invariants are used to make quantitative estimates on the behavior of bridges.

As a by-product of our analysis we also obtain a natural way of defining when a counting process has *lazy* bridges, meaning that "most customers tend to arrive at the end of the day".

It is interesting to remark that, since counting processes are among the building blocks for general queuing models, our results might open the way for a detailed quantitative study of queuing model under partial information, such as the final state of the queue.

The paper is organized as follows: in Section 1 we recall some basics about Markov counting processes and recall some results which we are going to use later on. In Section 2 we present our main results,Theorem 2.1, Theorem 2.2, and Theorem 2.5. In Section 3 we make their proofs.

1 Markov counting processes and their bridges

The sample path space Ω of the counting processes consists of all càdlàg step functions over the time interval $[0, 1]$ with finitely many jumps with amplitude $+1$ and an initial value in $\mathbb{N}$, from which we exclude paths which jumps either in $t = 0$ or $t = 1$. The space of probability measures over Ω is denoted $\mathcal{P}(\Omega)$. Any path $\omega \in \Omega$ is described by the collection $(x; t_1, \dots, t_n)$ of its initial position $x \in \mathbb{N}$ and its $n = \omega_1 - \omega_0$ instants of jumps $0 < t_1 < \cdots < t_n < 1$. We denote $T_i(\omega) := t_i$ the i-th instant of jump of ω. The canonical process is denoted by $X = (X_t)_{0 \le t \le 1}$, and we call counting process any probability law on Ω Any counting process P admits an almost surely unique non-decreasing process $A : [0, 1] \times \Omega \to \mathbb{R}_+$ such that $P(A(0) = 0) = 1$ and $t \mapsto X_t - X_0 - A(t)$ is a local P-martingale, (see Jacod [8, Thm. 2.1], for instance). A characterizes the dynamics of P, and it is called the compensator. When the compensator is absolutely continuous, we call its derivative the *intensity* of P. We consider *Markov* counting processes, meaning that the intensity depends only on time and the current state. We assume that the jump intensity $\ell : [0, 1] \times \mathbb{N} \to \mathbb{R}_+$ satisfies the following:

i) For all $z \in \mathbb{N}$, $t \mapsto \ell(t, z)$ is continuously differentiable in $[0, 1]$.
ii) The function $(t, z) \mapsto \ell(t, z + 1) - \ell(t, z)$ is globally upper and lower bounded on $[0, 1] \times \mathbb{N}$.
iii) For all $z \in \mathbb{N}$, $t \in [0, 1]$, $\ell(t, z) > 0$.

From now on, a reference counting process P with intensity satisfying i),ii),iii) and whose initial measure has full support is fixed. In the rest of the paper whenever we refer to the intensity of a Markov counting process, we always assume that it satisfies i),ii), and iii) even if we do not specify it. i),ii) ensure existence and non-explosion of the process, (for the non explosion part, see [10, Theorem 2.5.2]), while iii) ensures that bridges are well defined between any pair of states x, y with $x \le y$. We denote by P^{xy}_{su} the xy bridge between s and u.

$$P^{xy}_{su}(\cdot) := P(\cdot\,|X_s = x, X_u = y)$$

When $s = 0, u = 1$ we simply write P^{xy}. We will call $u - s$ the time-length of the bridge whereas $y - x$ is the height of the bridge. Moreover, for any $s \in [0, 1]$, $x \in \mathbb{N}$ we define

$$P^{x}_{s}(\cdot) = P(\cdot\,|X_s = x)$$

Expectation under P is denoted E_P. $E_{P^{xy}_{su}}$ and $E_{P^{xy}}$,$E_{P^x_s}$ are defined analogously.

1.1 Duality formula for the bridge of a Markov counting process

The reciprocal characteristic is a map $\Xi_\ell : \mathbb{N} \times [0,1] \to \mathbb{R}$ derived from the intensity of P, which encodes all the necessary and sufficient information on ℓ to construct the family of bridges $\{P^{xy}_{su}\}_{x \leq y \in \mathbb{N}, s \leq u}$. It is defined as[1]:

$$\Xi_\ell(t,z) := \begin{cases} \partial_t \log \ell(t,z) + \ell(t,z+1) - \ell(t,z), & \text{if } \ell(t,z) \neq 0 \\ 0, & \text{if } \ell(t,z) = 0 \end{cases}$$

A short-time probabilistic interpretation of Ξ_ℓ can be found in a larger generality in [5, Theorem 2.6]: it describes the asymptotic distribution of jump times in a bridge whose time length is very short. It is part of this paper to extract quantitative bounds in a *non* asymptotic time scale. At this point, it is worth recalling a duality formula proved in [6], building on earlier work in [4], characterizing the bridges of P, as we are going to use it in the proofs. To streamline the presentation, we omit some minor details. For precise statements we refer to [6, Theorem 2.12]. The duality formula is an integration by parts formula on the path space Ω which puts in relation a derivative operator with a stochastic integral operator. The derivative operator acts on test functions $\Phi : \Omega \to \mathbb{R}$ of the form:

$$\Phi = \varphi\big(X_0; T_1, \ldots, T_m\big), \quad m \geq 1, \varphi : \mathbb{N} \times [0,1]^m \to \mathbb{R}$$

where φ is such that for all $x \in \mathbb{Z}$, the partial functions $\varphi(x; \cdot)$ are $\mathcal{C}^\infty$-differentiable. The directions of differentiation are given by the set of periodic functions

$$\mathcal{U} = \big\{u \in \mathcal{C}^1([0,1];\mathbb{R}) : u(0) = u(1) = 0\big\}$$

The derivative operator acts as follows on a test function Φ:

$$\mathcal{D}_u \Phi = -\sum_{j=1}^{m} \partial_{t_j} \varphi(X_0; T_1, \ldots, T_m) u(T_j) \tag{1.1}$$

After recalling that the stochastic integral $\int_0^1 f(t, X_{t^-}) dX_t$ is, as usual, $\sum_{T_i < 1} f(T_i, X_{T_i^-})$ we can state the formula.

Theorem 1.1. *Let P be a counting process of intensity ℓ. Then P^{xy} is the only element of $\mathcal{P}(\Omega)$ such that $P^{xy}(X_0 = x, X_1 = y) = 1$ and*

$$E_{P^{xy}}(\mathcal{D}_u \Phi) = E_{P^{xy}}\Big(\Phi \int_0^1 \big[\dot{u}(t) + \Xi_\ell(t, X_{t^-}) u(t)\big]\, dX_t\Big) \tag{1.2}$$

holds for any test function Φ and any $u \in \mathcal{U}$.

2 Quantitative estimates

2.1 Convexity and concavity in time of the mean value

Our first result is the equivalence between global bounds on Ξ_ℓ and the convexity/concavity of the mean value, as a function of time.

---

[1] In the definition of the reciprocal characteristic we allowed for $\ell(t,z) = 0$, which was excluded by our hypothesis iii). This is because in the proof of Theorem 2.1 we will consider the reciprocal characteristics of the intensities of the bridges of P, which fail to satisfy iii). However, the reference intensity ℓ always satisfies this condition

Theorem 2.1. *Let P be a counting process. If*

$$\inf_{t\in[0,1],z\in\mathbb{N}} \Xi_\ell(t,z) \geq 0,$$

then for all $0 \leq s \leq u \leq 1$ and for all $x \leq y \in \mathbb{N}$ the function $[s,u] \ni t \mapsto E_{P^{xy}_{su}}(X_t)$ is convex. Conversely, if

$$\sup_{t\in[0,1],z\in\mathbb{N}} \Xi_\ell(t,z) \leq 0,$$

then for all $0 \leq s \leq u \leq 1$, and for all $x \leq y \in \mathbb{N}$ the function $[s,u] \ni t \mapsto E_{P^{xy}_{su}}(X_t)$ is concave.

Thanks to this result, it makes sense to say that the bridges of a counting process are "lazy" when Ξ_ℓ is a non negative function. Indeed, using the interpretation that X_t counts the number of arrivals of customers, Theorem 2.1 states that, on average, customers arrive at a slow rate at the beginning of the day, and this rate always increases over time in order to match the information that $y-x$ customers arrived in the whole day: hence arrivals are concentrated towards the end of the day.

In particular, at time t, one expects to have observed less than $n\,t$ arrivals, as the curve $[0,1] \ni t \mapsto E_{P^{xy}}(X_t)$ lies below the straight line between x and y.

On the contrary, if Ξ_ℓ is non-positive we can say that bridges are "hurried". Customers arrive very quickly at the beginning, and the graph of $[0,1] \ni t \mapsto E_{P^{xy}}(X_t)$ lies above the straight line between x and y. The arrival of customers has to slow down as time grows (concavity), due to the constraint that $y-x$ customers, and not more, have to be arrived at the end of the day. We refer to Figure 1 for an illustration of these concepts.

2.2 Estimates for the marginals

Next result turns the qualitative statement of Theorem 2.1 into a quantitative one. We start from the observation that there are three main families of Markov counting processes such that Ξ_ℓ is a time-space constant function:

- The Poisson processes, i.e. $\ell(t,z) \equiv \alpha > 0$. Regardless of the value of α, their bridges are neither lazy nor hurried. The average rate which customers arrive is constantly equal to $y-x$.
- Markov counting processes with time homogeneous and space linear rates, i.e. $\ell(t,z) = \lambda z + \alpha$ for some $\lambda, \alpha > 0$.
- The space-homogeneous counting processes with exponential rate, i.e. $\ell(t,z) = \alpha \exp(\lambda t)$, for some $\lambda \in \mathbb{R}, \alpha > 0$. Their bridges may be lazy or hurried, depending on whether λ is positive or not.

For this kind of models all sort of explicit computations are possible: we will use them as a benchmark in the next theorem. As we will see in Lemma 3.3, the marginal at time t of the xy bridge of a process with $\Xi_\ell \equiv \lambda$ is a binomial distribution whose parameters are determined by λ, x, y and t. This is a generalization of the well known fact that marginals of Poisson bridges are binomial distributions.

We make use of this explicit computation to provide estimates for the case when the marginal distributions are not known in closed form. We shall see that if Ξ_ℓ admits a non negative lower bound, say λ, then any bridge of P has marginals which have lighter tails than the marginals of the same bridge built from a process satisfying $\Xi_\ell \equiv \lambda$. Therefore we prove that the tails are lighter than the tails of a certain binomial law. In this sense, Theorem 2.2 strengthens the definition of lazy bridge, giving a parameter to measure quantitatively such laziness, and creating a partial order in the family of processes with lazy bridges. All statements can be reversed when lower bounds become upper bounds.

To state the next Theorem, some notation is needed: we denote by $\mathcal{B}_{n,p}$ the binomial distribution associated with n trials and probability of success p.

$$\forall\ 0 \le k \le n, \qquad \mathcal{B}_{n,p}(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

We define for all $\lambda \in \mathbb{R}$ the function $\pi_\lambda : [0,1] \to [0,1]$ as

$$\pi_\lambda(t) = \frac{\exp(\lambda t) - 1}{\exp(\lambda) - 1}. \tag{2.1}$$

Theorem 2.2. *Let P be a Markov counting process of intensity ℓ, $x \le y \in \mathbb{N}$. If for some $\lambda \in \mathbb{R}$,*

$$\inf_{t \in [0,1], x \le z \le y-1} \Xi_\ell(t,z) \ge \lambda \tag{2.2}$$

then for $1 \le i \le y-x$ and $t \in [0,1]$:

$$P^{xy}(X_t \ge x+i) \le \mathcal{B}_{y-x,\pi_\lambda(t)}(\{i, \cdots, y-x\}).$$

In other words, $P^{xy}(X_t - x \in \cdot)$ is stochastically dominated by the Bernoulli distribution $\mathcal{B}_{y-x,\pi_\lambda(t)}$. Conversely, if for some $\lambda \in \mathbb{R}$,

$$\sup_{t \in [0,1], x \le z \le y-1} \Xi_\ell(t,z) \le \lambda$$

then for any $1 \le i \le y-x$ and $t \in [0,1]$:

$$P^{xy}(X_t \ge x+i) \ge \mathcal{B}_{y-x,\pi_\lambda(t)}(\{i, \cdots, y-x\}).$$

Remark 2.3. The inequalities we obtain in this Theorem are equalities when $\Xi_\ell(t,z) \equiv \lambda$ (see Lemma 3.3) , and therefore are sharp.

A simple consequence of Theorem 2.2 is the following mean value estimate.

Corollary 2.1. Let P be a Markov counting process of intensity ℓ, $x \le y \in \mathbb{N}$. If

$$\inf_{t \in [0,1], z \in \mathbb{N}} \Xi_\ell(t,z) \ge \lambda$$

then for any $x \le y$ and $t \in [0,1]$:

$$E_{P^{xy}}(X_t) \le x + (y-x)\frac{\exp(\lambda t) - 1}{\exp(\lambda) - 1}.$$

Remark 2.4. About Theorem 2.1 and 2.2, we will provide a full proof only for the assertions concerning lower bounds on the characteristics. The assertions concerning the upper bounds can be proved with a time-reversal argument. Indeed, if P is a Markov counting process, then the time-reversed process P^* is pure-death Markov process (i.e. a process that can only make jumps of height -1) with values in $\mathbb{N}$. The reciprocal characteristics associated with P^* (in the precise sense of [5, Def 2.3]), which we denote by $\Xi^*_j(\cdot,\cdot)$ satisfy the following relation

$$\Xi^*_j(t,z) = -\Xi_j(1-t,z).$$

The proof of this fact is an easy consequence of the formula relating the intensities of a Markov process and its time-reversal. Therefore global upper bounds for Ξ_j are global lower bounds for Ξ^*_j, and one can follow the proofs we present in this article to get the statements about upper bounds on the reciprocal characteristics. One could also prove these statements by carefully repeating the proofs presented here, and changing the sign of the inequalities when needed. However, the former approach is more elegant.

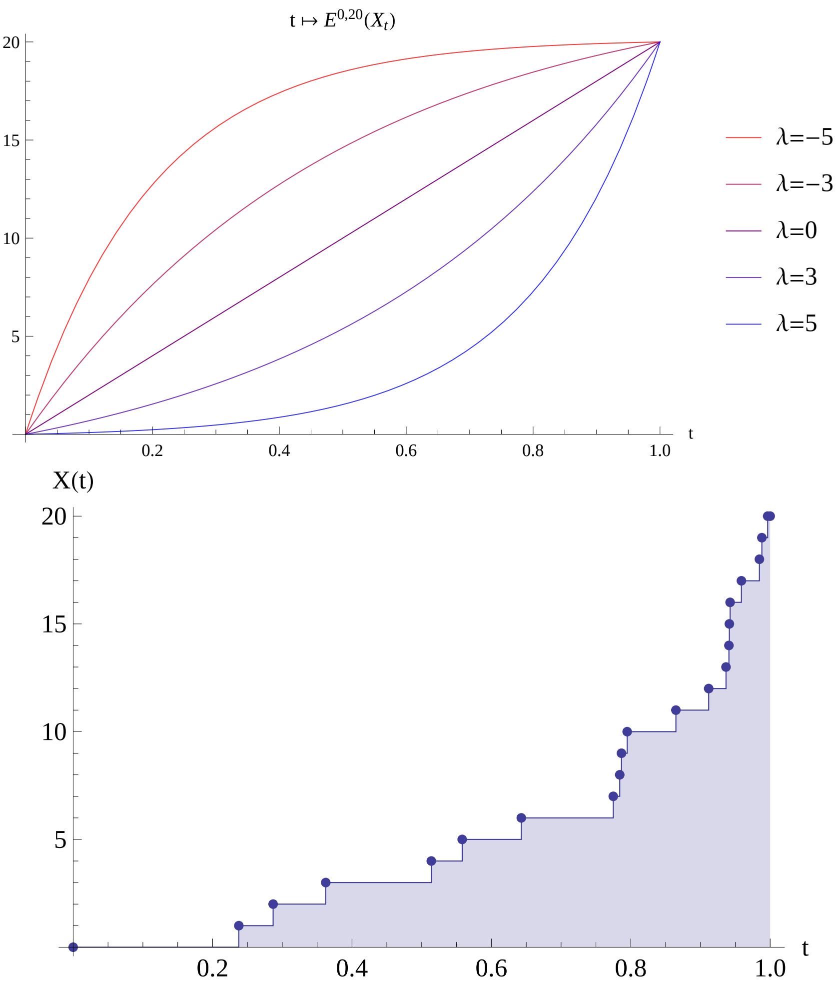

Figure 1: Top: We plot for $\lambda \in \{-5,-3,0,3,5\}$ the function $t \mapsto E_{P^{0,20}}(X_t) = 20\,\pi_\lambda(t)$, where the intensity ℓ of P satisfies $\Xi_\ell \equiv \lambda$. For negative values of λ we have a concave function, whereas for positive values a convex one. Moreover, we see that, as a function of λ, $E_{P^{0,20}}(X_t)$ is decreasing. Bottom: A sample path of a lazy bridge ($\lambda = 3$): it displays very little activity close to zero: almost all jump times accumulate around $t = 1$.

2.3 Law of large numbers

Here, we prove an asymptotic result under the hypothesis that the map $\Xi_\ell(t,z)$ converges as $z \to +\infty$ to $\lambda \in \mathbb{R}$. We will see that as $N \to +\infty$, the bridge of length one between 0 and N, after a suitable rescaling, converges to the deterministic curve $t \mapsto \pi_\lambda(t)$. In Figure 1 we plot the curve $20\,\pi_\lambda(t)$ for different values of λ.

Theorem 2.5. *Let P be a Markov counting process of intensity ℓ. Assume that for some $\lambda \in \mathbb{R}$:*

$$\lim_{z\to+\infty} \sup_{t\in[0,1]} |\Xi_\ell(t,z) - \lambda| = 0. \tag{2.3}$$

Then for every $\varepsilon > 0$:

$$\lim_{N\to+\infty} P^{0N}\left(\sup_{t\in[0,1]} \left| \frac{1}{N} X_t - \pi_\lambda(t) \right| \geq \varepsilon \right) = 0.$$

3 Proofs of the results

Proof of Theorem 2.1

The proof of Theorem 2.1 builds on the earlier results of [6] and on Lemma 3.1below. There, we prove the statement under P instead of P^{xy}. Unless in the rest of the paper, we do not assume ℓ to satisfy point iii) in the proof of this theorem. This is because we will eventually apply it to a bridge P^{xy}, whose intensity does not satisfy this condition.

Lemma 3.1. *Let P be a Markov counting process of intensity ℓ. Then*

$$\inf_{t\in[0,1],z\in\mathbb{N}} \Xi_\ell(t,z) \geq 0$$

if and only if for all $s,x \in [0,1]\times\mathbb{N}$ the function $[s,1]\ni t\mapsto E_{P^x_s}(X_t)$ is convex.

Conversely,

$$\sup_{t\in[0,1],z\in\mathbb{N}} \Xi_\ell(t,z) \leq 0$$

if and only if for all $s,x \in [0,1]\times\mathbb{N}$ the function $[s,1]\ni t\mapsto E_{P^x_s}(X_t)$ is concave.

Proof. We only prove the convexity statement, the other being analogous (see Remark 2.4).

($\Rightarrow$) Fix $s\in[0,1], x\in\mathbb{N}$ and define $\varphi_t := E_{P^x_s}(X_t)$ for $t\in[s,1]$. Differentiating twice we obtain:

$$\begin{aligned}
\dot\varphi_t &= E_{P^x_s}(\ell(t,X_{t^-}))\\
\ddot\varphi_t &= E_{P^x_s}(\partial_t\ell(t,X_{t^-}) + (\ell(t,X_{t^-}+1)-\ell(t,X_{t^-}))\ell(t,X_{t^-}))\\
&= E_{P^x_s}(\Xi_\ell(t,X_{t^-})\ell(t,X_{t^-})\mathbf{1}_{\{\ell(t,X_{t^-})>0\}} + \partial_t\ell(t,X_{t^-})\mathbf{1}_{\{\ell(t,X_{t^-})=0\}})
\end{aligned}$$

Since $\ell(s,z)$ is continuously differentiable and for all $t\in[0,1]$ $P(X_t - X_{t^-}\neq 0)=0$, we have that, for a fixed t, $\ell(t,X_{t^-})=0$ a.s. implies that $\partial_t\ell(t,X_{t^-})=0$. Therefore:

$$\ddot\varphi_t = E_{P^x_s}(\Xi_\ell(t,X_{t^-})\ell(t,X_{t^-}))$$

If $\inf_{t,z}\Xi_\ell(t,z)\geq 0$, we then have $\ddot\varphi_t\geq 0$, which gives the convexity.

($\Leftarrow$) Consider $s\in[0,1]$,$x\in\mathbb{N}$ such that $\ell(s,x)>0$ (note that if $\ell(s,x)=0$, $\Xi_\ell(s,x)=0$ and there is nothing to prove). Then, defining φ as above and repeating the same computations we obtain, by evaluating the second derivative at s, and exploiting the fact that $X_s = x$ almost surely:

$$\ddot\varphi_s = \Xi_\ell(s,x)\ell(s,x)$$

Since $\ddot\varphi_s\geq 0$ by assumption and $\ell(s,x)>0$, then $\Xi_\ell(s,x)\geq 0$. Because of the arbitrary choice of s and x the conclusion follows. □

The proof of Theorem 2.1 becomes now almost straightforward.

Proof. We exploit the fact P^{xy}_{su} is also a counting process whose intensity $\ell^{x,y}_{s,u}(t,z)$ is continuously differentiable on $[0,1)\times\mathbb{N}$ and satisfies:

$$\Xi_{\ell^{x,y}_{s,u}}(t,z) = \begin{cases}\Xi_\ell(t,z), & \text{if } x\leq z\leq y-1,\ s<t<u\\ 0, & \text{otherwise}\end{cases}$$

This is proven in [6, Theorem 1.11]. An application of Lemma 3.1 gives then the conclusion. □

Proof of Theorem 2.2

Prior to the proof, we introduce some useful notation which simplifies many computations therein. We call $n := y - x$ the total number of jumps that the xy bridge makes. Moreover, we adopt the following conventions:

- For $(t_1, .., t_n) \in [0,1]^n$ we adopt the compact notation $\mathbf{t}$. For any $1 \le i \le k \le n$ the vector $(t_i, .., t_k)$ is denoted $\mathbf{t}_{i,k}$.
- For $1 \le i \le k \le n$, $0 \le s \le r \le 1$, we call $\Delta^{s,r}_{i,k}$ the set :

$$\Delta^{s,r}_{i,k} := \{(t_i, ..., t_k) \in [0,1]^{k-i+1} : s < t_i < t_{i+1} < .. < t_k < r\}$$

When $s = 0, r = 1, i = 1, k = n$, we simply write Δ_n .
- For any $1 \le j \le n$ we define $\xi_j : [0,1] \longrightarrow \mathbb{R}$ as the only primitive of $t \mapsto \Xi_\ell(t, x+j-1)$ such that $\xi_j(0) = 0$. We then define

$$\xi : \Delta_n \longrightarrow \mathbb{R}, \quad \xi(\mathbf{t}) = \sum_{j=1}^{n} \xi_j(t_j)$$

We also define for all $1 \le i \le k \le n$:

$$\xi_{i,k}(\mathbf{t}_{i,k}) = \sum_{j=i}^{k} \xi_j(t_j)$$

The proof is structured as follows: we first prove two Lemmas, Lemma 3.2 Lemma 3.3, which contain the two main ingredients of the proof. Then we do the proof of Theorem 2.2, relying on the more technical Lemma 3.4, which is proven later, together with the simple auxiliary Lemma 3.5.

Lemma 3.2. *The distribution of jump times $(T_1, .., T_n)$ under P^{xy} is given by:*

$$\forall A \subseteq \Delta_n, \quad P^{xy}\big((T_1, .., T_n) \in A\big) = \frac{1}{Z_\ell} \int_A \exp\big(\xi(\mathbf{t})\big) d\mathbf{t} \tag{3.1}$$

Proof. For convenience, we denote the vector $(T_1, .., T_n)$ by $\mathbf{T}$. Let us define the probability measure $Q \in \mathcal{P}(\Omega)$ by:

$$dQ = \frac{1}{Z_\ell} \exp(\xi(\mathbf{T})) dR^{xy}$$

where R^{xy} is the Poisson bridge from x to y. Thanks to the duality formula (1.2), the Poisson bridge is characterized by:

$$\forall \Phi, u \in \mathcal{U}, \quad E_{R^{xy}}(\mathcal{D}_u \Phi) = E_{R^{xy}}\Big(\Phi \int_0^1 \dot{u}(t)\, dX_t\Big) \tag{3.2}$$

Consider now a test function Φ, using the duality formula (1.2) and the fact that $\mathcal{D}_u$ is a true derivative operator satisfying Leibniz's product rule:

$$\begin{aligned}
E_Q(\mathcal{D}_u \Phi) &= E_{R^{xy}}\Big(\exp(\xi(\mathbf{T}))\, \mathcal{D}_u \Phi\Big) \\
&= E_{R^{xy}}\Big(\mathcal{D}_u\big[\exp(\xi(\mathbf{T}))\,\Phi\big]\Big) - E_{R^{xy}}\Big(\Phi\, \mathcal{D}_u\big[\exp(\xi(\mathbf{T}))\big]\Big) \\
&\underbrace{=}_{\text{eq. (3.2)+eq.(1.1)}} E_{R^{xy}}\Big(\exp(\xi(\mathbf{T}))\Phi \int_0^1 \dot{u}(t) dX_t\Big) \\
&\quad + E_{R^{xy}}\Big(\Phi \exp(\xi(\mathbf{T}) \sum_{j=1}^{n} \underbrace{\partial_{t_j}\xi(\mathbf{T})}_{=\Xi_\ell(T_j, X_{T_j^-})} u(T_j))\Big) \\
&= E_Q\Big(\Phi \int_0^1 \big[\dot{u}(t) + \Xi_\ell(t, X_{t^-}) u(t)\big] dX_t\Big)
\end{aligned}$$

An application of Theorem 1.1 allows to deduce that $Q = P^{xy}$. Therefore, the density of $P^{xy} \circ \mathbf{T}^{-1}$ with respect to $R^{xy} \circ \mathbf{T}^{-1}$ is proportional to $\exp(\xi(\mathbf{t}))$. It is well known that $R^{xy} \circ \mathbf{T}^{-1}$ is the normalized Lebesgue measure on Δ_n. The conclusion then follows. □

Let us compute explicitly the distribution of the jump times of a bridge of a counting process such that Ξ_ℓ is constantly equal to λ.

Lemma 3.3. *Let P be a Markov counting process of intensity ℓ, $x \leq y \in \mathbb{N}$. Let ℓ be such that for some $\lambda \in \mathbb{R}$:*

$$\forall t \in [0,1],\, x \leq z \leq y-1, \quad \Xi_\ell(t,z) = \lambda$$

Then, for all $t \in [0,1]$, the law of $X_t - x$ under P^{xy} is $\mathcal{B}_{y-x,\pi_\lambda(t)}$, where $\pi_\lambda(t)$ is defined by (2.1).

Proof. As a direct consequence of Lemma 3.2 combined with the current hypothesis we have:

$$P^{xy}(X_t \geq x+i) = P^{xy}(T_i \leq t) = \frac{1}{Z_\lambda} \int_{\Delta_n \cap \{t_i \leq t\}} \exp(\lambda \sum_{j=1}^n t_j) d\mathbf{t} \tag{3.3}$$

where Z_λ is a normalization constant. Thanks to the invariance of the integrand under permutations we have:

$$\frac{1}{Z_\lambda} \int_{\Delta_n \cap \{t_i \leq t\}} \exp\big(\lambda \sum_{j=1}^n t_j\big) d\mathbf{t} = \frac{1}{(\exp(\lambda)-1)^n} \int_A \exp(\lambda \sum_{j=1}^n t_j) d\mathbf{t} \tag{3.4}$$

with

$$A := \bigcup_{\sigma \in S_n} \{\mathbf{t} : t_{\sigma(1)} < .. < t_{\sigma(n)}, t_{\sigma(i)} \leq t\}$$

where S_n is the symmetric group with n elements. We can rewrite A in a convenient way, using some simple combinatorial arguments. We have:

$$A := \{\mathbf{t} \in [0,1]^n s.t. |\{j : t_j \leq t\}| \geq i\}$$

But then the right hand side of (3.4) is nothing but the probability that at least i among n i.i.d. random variables are less than t, each variable being supported on $[0,1]$ and having a density proportional to $\exp(\lambda s)$. The conclusion then follows after a simple calculation. □

We can now prove Theorem 2.2.

Proof. Again, because of Remark 2.4, we only prove the Theorem when $\Xi_j(\cdot,\cdot)$ is bounded below. Because of Lemma 3.2 and Lemma 3.3, and the elementary fact that $\{X_t \geq x+i\} = \{T_i \leq t\}$, all what we need to do is to prove that for all $1 \leq i \leq n, 0 \leq t \leq 1$:

$$\frac{1}{Z_\ell} \int_{\mathbf{t} \in \Delta_n \cap \{t_i \leq t\}} \exp(\xi(\mathbf{t})) d\mathbf{t} \leq \frac{1}{Z_\lambda} \int_{\mathbf{t} \in \Delta_n \cap \{t_i \leq t\}} \exp\big(\lambda \sum_{j=1}^n t_j\big) d\mathbf{t} \tag{3.5}$$

where Z_ℓ, Z_λ are normalization constants.

To this aim, we define the function $\rho : [0,1] \to \mathbb{R}_+$ by:

$$\rho(r) := \frac{\int_{\mathbf{t} \in \Delta_n \cap \{t_i \leq r\}} \exp(\xi(\mathbf{t})) d\mathbf{t}}{\int_{\mathbf{t} \in \Delta_n \cap \{t_i \leq r\}} \exp\big(\lambda \sum_{j=1}^n t_j\big) d\mathbf{t}}$$

Observing that $\rho(1) = \frac{Z_\ell}{Z_\lambda}$, (3.5) is equivalent to $\rho(t) \leq \rho(1)$, and therefore it is sufficient to prove that ρ is non-decreasing. For this, we rewrite $\rho(t)$ in a convenient way, by conditioning on the value of t_i:

$$\rho(r) = \frac{\int_0^r p(s)ds}{\int_0^r q(s)ds} \tag{3.6}$$

with:

$$\begin{aligned}
p(s) &:= \int_{\mathbf{t}\in\Delta_n\cap\{t_i=s\}} \exp\big(\xi(\mathbf{t})\big) d\mathbf{t}_{1,i-1} d\mathbf{t}_{i,n} \\
&= \exp(\xi_i(s)) \int_{\Delta^{0,s}_{1,i-1}} \exp\big(\xi_{1,i-1}(\mathbf{t}_{1,i-1})\big) d\mathbf{t}_{1,i-1} \int_{\Delta^{s,1}_{i+1,n}} \exp\big(\xi_{i+1,n}(\mathbf{t}_{i+1,n})\big) d\mathbf{t}_{i+1,n} \\
q(s) &:= \int_{\mathbf{t}\in\Delta_n\cap\{t_i=s\}} \exp\big(\lambda \sum_{j=1}^n t_j\big) d\mathbf{t}_{1,i-1} d\mathbf{t}_{i+1,n} \\
&= \exp(\lambda s) \int_{\Delta^{0,s}_{1,i-1}} \exp\big(\lambda \sum_{j=1}^{i-1} t_j\big) d\mathbf{t}_{1,i-1} \int_{\Delta^{s,1}_{i+1,n}} \exp\big(\lambda \sum_{j=i+1}^{n} t_j\big) d\mathbf{t}_{i+1,n}
\end{aligned}$$

Let us first show that $p(s)/q(s)$ is non-decreasing. Lemma 3.4 ensures that both

$$s \mapsto \frac{\int_{\Delta^{0,s}_{1,i-1}} \exp(\xi_{1,i-1}(\mathbf{t}_{1,i-1})) d\mathbf{t}_{1,i-1}}{\int_{\Delta^{0,s}_{1,i-1}} \exp(\lambda \sum_{j=1}^{i-1}) d\mathbf{t}_{1,i-1}} \quad \text{and} \quad s \mapsto \frac{\int_{\Delta^{s,1}_{i+1,n}} \exp(\xi_{i+1,n}(\mathbf{t}_{i+1,n})) d\mathbf{t}_{i+1,n}}{\int_{\Delta^{s,1}_{i+1,n}} \exp(\lambda \sum_{j=i+1}^{n} t_j) d\mathbf{t}_{i+1,n}}$$

are non-decreasing in s. Moreover, $\exp(\xi_i(s) - \lambda s)$ is also non-decreasing in s since $\dot{\xi}_i(\cdot) = \Xi_\ell(\cdot, x+i-1) \geq \lambda$ by hypothesis. We have thus shown that $p(s)/q(s)$ is non-decreasing. By Lemma 3.5 and (3.6) we get that ρ is non-decreasing too, and the conclusion follows. □

Lemma 3.4. *The functions*

$$s \mapsto \frac{\int_{\Delta^{0,s}_{1,i-1}} \exp(\xi_{1,i-1}(\mathbf{t}_{1,i-1})) d\mathbf{t}_{1,i-1}}{\int_{\Delta^{0,s}_{1,i-1}} \exp(\lambda \sum_{j=1}^{i-1}) d\mathbf{t}_{1,i-1}}$$

and

$$s \mapsto \frac{\int_{\Delta^{s,1}_{i+1,n}} \exp(\xi_{i+1,n}(\mathbf{t}_{i+1,n})) d\mathbf{t}_{i+1,n}}{\int_{\Delta^{s,1}_{i+1,n}} \exp(\lambda \sum_{j=i+1}^{n} t_j) d\mathbf{t}_{i+1,n}}$$

are non-decreasing in s.

Proof. We only prove the first statement, the other case being completely analogous. We work by induction on i. For the case $i = 2$, we have to show that

$$s \mapsto \frac{\int_0^s \exp(\xi_1(t_1)) dt_1}{\int_0^s \exp(\lambda t_1) dt_1}$$

is non-decreasing. The assumption of Theorem 2.2 ensures that $\xi_1(t_1) - \lambda t_1$ is non-decreasing, since $\lambda \leq \Xi_\ell(\cdot, x) = \dot{\xi}_1(\cdot)$ by hypothesis. But then we conclude by applying Lemma 3.5. For the inductive step we rewrite

$$\begin{aligned}
&\int_{\Delta^{0,s}_{1,i-1}} \exp(\xi_{1,i-1}(\mathbf{t}_{1,i-1})) d\mathbf{t}_{1,i-1} \\
&\quad = \int_0^s \exp(\xi_{i-1}(t_{i-1})) \left\{ \int_{\Delta^{0,t_{i-1}}_{1,i-2}} \exp(\xi_{1,i-2}(\mathbf{t}_{1,i-2})) d\mathbf{t}_{1,i-2} \right\} dt_{i-1}
\end{aligned}$$

and similarly

$$\int_{\Delta^{0,s}_{1,i-1}} \exp(\lambda \sum_{j=1}^{i-1} t_j) d\mathbf{t}_{1,i-1} = \int_0^s \exp(\lambda t_{i-1}) \left\{ \int_{\Delta^{0,t_{i-1}}_{1,i-2}} \exp(\sum_{j=1}^{i-2} t_j) d\mathbf{t}_{1,i-2} \right\} dt_{i-1}$$

Using the inductive hypothesis, we have that

$$t_{i-1} \mapsto \frac{\int_{\Delta^{0,t_{i-1}}_{1,i-2}} \exp(\xi_{1,i-2}(\mathbf{t}_{1,i-2})) d\mathbf{t}_{1,i-2}}{\int_{\Delta^{0,t_{i-2}}_{1,i-1}} \exp(\sum_{j=1}^{i-2} t_j) d\mathbf{t}_{1,i-2}}$$

is non-decreasing. Moreover, $t_{i-1} \mapsto \frac{\exp(\xi_{i-1}(t_{i-1}))}{\exp(\lambda t_{i-1})}$ is also non-decreasing, since by assumption $\lambda \leq \Xi_\ell(\cdot, x+i-2) = \dot{\xi}_{i-1}(\cdot)$. The conclusion follows by Lemma 3.5. □

Lemma 3.5. *Let p, q be continuously differentiable strictly positive functions such that $\frac{p}{q}(s)$ is non-decreasing. Then*

$$t \mapsto \frac{\int_0^t p(s)ds}{\int_0^t q(s)ds} \tag{3.7}$$

is also non-decreasing.

Proof. We differentiate the map defined by (3.7) in t. We obtain:

$$\frac{1}{\left(\int_0^t q(s)ds\right)^2} \left[p(t) \int_0^t q(s)ds - q(t) \int_0^t p(s)ds \right]$$

We observe that the conclusion follows if we can prove that

$$\forall s \leq t \quad p(t)q(s) \geq q(t)p(s)$$

Since $s \leq t$, our hypothesis ensures this. □

Let us prove Corollary 2.1

Proof. We simply write:

$$E_{P^{xy}}(X_t) = x + \sum_{i=1}^n P^{xy}(T_i \leq t) \underbrace{\leq}_{\text{Th. 2.2}} \sum_{i=1}^n \mathcal{B}_{n,\pi_\lambda(t)}(\{i, \cdots, n\}) = x + n \frac{\exp(\lambda t) - 1}{\exp(\lambda) - 1}.$$

Recalling that $n = y - x$ the conclusion follows. □

Proof of Theorem 2.5

Theorem 2.2 can be adapted to the case when we consider bridges of non-unitary length. The proof is basically identical, and we do not include it here. However, we will need this fact.

Proposition 3.1. Let P be a Markov counting process of intensity ℓ, $x \leq y \in \mathbb{N}$ and $0 \leq s < u \leq 1$. If for some $\lambda \in \mathbb{R}$,

$$\inf_{t \in [s,u], x \leq z \leq y-1} \Xi_\ell(t, z) \geq \lambda$$

then for $1 \leq i \leq y - x$ and $t \in [s, u]$:

$$P^{xy}_{su}(X_t \geq x + i) \leq \mathcal{B}_{y-x, \pi^{s,u}_\lambda(t)}(\{i, ., y - x\}) \tag{3.8}$$

where

$$\pi^{s,u}_\lambda(t) = \frac{\exp(\lambda(t-s)) - 1}{\exp(\lambda(u-s)) - 1}$$

Given the proposition, we can prove Theorem 2.5

Proof. Fix $\varepsilon > 0$ and $k \geq \frac{2}{\varepsilon} \sup_{t\in[0,1]} \partial_t \pi_\lambda(t)$. We show that:

$$A := \left\{ \sup_{0\leq j\leq k} |\frac{1}{N} X_{\frac{j}{k}} - \pi_\lambda(\frac{j}{n}))| \leq \frac{\varepsilon}{2} \right\} \subseteq \left\{ \sup_{t\in[0,1]} |\frac{1}{N} X_t - \pi_\lambda(t)| \leq \varepsilon \right\} =: B$$

Indeed, assume that $\omega \in A$ and take $t \in [0,1]$. Then $t \in [\frac{j}{k}, \frac{j+1}{k}]$ for some $0 \leq j \leq k-1$. Because of the fact that both $\frac{1}{N} X_\cdot$ and $\pi_\lambda(\cdot)$ are non-decreasing functions we have:

$$|\frac{1}{N} X_t - \pi_\lambda(t)| \leq \max \left\{ \frac{1}{N} X_{\frac{j+1}{k}} - \pi_\lambda(\frac{j}{k}), \pi_\lambda(\frac{j+1}{k}) - \frac{1}{N} X_{\frac{j}{k}} \right\}$$

Since $X_\cdot \in A$ we have:

$$\frac{1}{N} X_{\frac{j+1}{k}} - \pi_\lambda(\frac{j}{k}) = \frac{1}{N} X_{\frac{j+1}{k}} - \pi_\lambda(\frac{j+1}{k}) + \pi_\lambda(\frac{j+1}{k}) - \pi_\lambda(\frac{j}{k}) \leq \frac{\varepsilon}{2} + \frac{1}{k} \sup_t \partial_t \pi_\lambda(t) \leq \varepsilon$$

where the last inequality is justified by the choice of k. In the same way one shows that $\pi_\lambda(\frac{j+1}{k}) - \frac{1}{N} X_{\frac{j}{k}} \leq \varepsilon$. Since t can be chosen arbitrarily in $[0,1]$, we have shown that $A \subseteq B$. Because of this, to prove the theorem, it is sufficient to prove that for any finite collection $(t_1, .., t_k)$ and any $\varepsilon > 0$ we have:

$$\lim_{N\to+\infty} P^{0N} (\sup_{1\leq l\leq k} |\frac{1}{N} X_{t_l} - \pi_\lambda(t_l)| \geq \varepsilon) = 0$$

Moreover, since we since the $\sup$ which appears in the above equation is taken over finitely many terms, it suffices to consider one-time marginals. All we need to show is that for all $t \in [0,1]$, $\varepsilon \geq 0$

$$\lim_{N\to+\infty} P^{0N} \left(|\frac{1}{N} X_t - \pi_\lambda(t)| \geq \varepsilon \right) = 0$$

Fix now $t \in [0,1], \varepsilon > 0$, and choose $\delta > 0$ small enough such that

$$\pi_\lambda(t) + \varepsilon \geq \pi_{\lambda-\delta}(t) \tag{3.9}$$

Because of the assumption (2.3), there exist k such that for all

$$\forall z \geq k \quad \sup_{s\in[0,1]} |\Xi_\ell(s,z) - \lambda| \leq \delta. \tag{3.10}$$

We choose one such k. Let us observe that, if N is large enough we have

$$\left\{ \frac{1}{N} X_t - \pi_\lambda(t) \geq \varepsilon \right\} \subseteq \{T_k \leq t\} \tag{3.11}$$

Using the Markov property we have, by conditioning on T_k and using (3.11):

$$\begin{aligned} P^{0N} \left(\frac{1}{N} X_t - \pi_\lambda(t) \geq \varepsilon \right) &= E_{P^{0N}} \left(E_P^{0,N} \left(\mathbf{1}_{\{\frac{1}{N} X_t - \pi_\lambda(t) \geq \varepsilon\}} \middle| T_k \right) \mathbf{1}_{\{T_k \leq t\}} \right) \\ &= E_{P^{0N}} \left(P_{T_k,1}^{k,N} \left(\frac{1}{N} X_t - \pi_\lambda(t) \geq \varepsilon \right) \mathbf{1}_{\{T_k \leq t\}} \right) \\ &\leq \sup_{s\in[0,t]} P_{s,1}^{k,N} \left(\frac{1}{N} X_t - \pi_\lambda(t) \geq \varepsilon \right) \end{aligned}$$

Thanks to our choice of k, see (3.10), we are entitled to apply Proposition 3.1. We get:

$$\sup_{s\in[0,t]} P^{k,N}_{s,1}\left(\frac{1}{N}X_t - \pi_\lambda(t) \geq \varepsilon\right) \leq \sup_{s\in[0,t]} \mathcal{B}_{N-k,\pi^{s,1}_{\lambda-\delta}(t)}(\{N(\underbrace{\pi_\lambda(t)+\varepsilon-k/N}_{:=\xi_N}),..,N-k\})$$

Since $\pi^{s,1}_{\lambda-\delta}(t)$ is decreasing in s, we have:

$$\sup_{s\in[0,1]} \mathcal{B}_{N-k,\pi^{s,1}_{\lambda-\delta}(t)}(\{N\xi_N,..,N-k\}) = \mathcal{B}_{N-k,\pi_{\lambda-\delta}(t)}(\{N\xi_N,..,N-k\})$$

With an application of the Law of large numbers, we have that, as N goes to infinity:

$$\mathcal{B}_{N-k,\pi_{\lambda-\delta}(t)}(NA) \to \delta_{\pi_{\lambda-\delta}(t)}(A), \quad \forall A \subseteq [0,1]$$

Therefore $\lim_{N\to+\infty} \mathcal{B}_{N-k,\pi_{\lambda-\delta}(t)}(\{N\xi_N,..,N-k\}) = 0$ if $\liminf_{N\to+\infty} \xi_N > \pi_{\lambda-\delta}(t)$. But this is ensured by our choice of δ, see (3.9). Therefore we have proven that $\lim_{N\to+\infty} P^{0N}(\frac{1}{N}X_t - \pi_\lambda(t) \geq \varepsilon) = 0$. With a similar argument one shows that $\lim_{N\to+\infty} P^{0N}(\frac{1}{N}X_t - \pi_\lambda(t) \leq -\varepsilon) = 0$, from which the conclusion follows. $\square$

Acknowledgments. The author wishes to thank Sylvie Roelly for reading a preliminary version of the article and two anonymous referees for several useful comments and remarks

References

[1] Per Kragh Andersen, Ørnulf Borgan, Richard D. Gill, and Niels Keiding, *Statistical models based on counting processes*, Springer Series in Statistics, Springer-Verlag, New York, 1993. MR-1198884

[2] P. Brémaud, *The martingale theory of point processes over the real half line admitting an intensity*, Control theory, numerical methods and computer systems modelling (Internat. Sympos., IRIA LABORIA, Rocquencourt, 1974), Springer, Berlin, 1975, pp. 519–542. Lecture Notes in Econom. and Math. Systems, Vol. 107. MR-0372984

[3] Pierre Brémaud, *Markov chains*, Texts in Applied Mathematics, vol. 31, Springer-Verlag, New York, 1999, Gibbs fields, Monte Carlo simulation, and queues. MR-1689633

[4] Eric A. Carlen and Étienne Pardoux, *Differential calculus and integration by parts on Poisson space*, Stochastics, algebra and analysis in classical and quantum dynamics (Marseille, 1988), Math. Appl., vol. 59, Kluwer Acad. Publ., Dordrecht, 1990, pp. 63–73. MR-1052702

[5] G. Conforti and C. Léonard, *Reciprocal classes of a random walk on graphs*, arXiv:1505.01323

[6] Giovanni Conforti, Christian Léonard, Rüdiger Murr, and Sylvie Rœlly, *Bridges of Markov counting processes. Reciprocal classes and duality formulas*, Electron. Commun. Probab. **20** (2015), no. 18, 12. MR-3320406

[7] Thomas R. Fleming and David P. Harrington, *Counting processes and survival analysis*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, John Wiley & Sons, Inc., New York, 1991. MR-1100924

[8] Jean Jacod, *Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **31** (1974/75), 235–253. MR-0380978

[9] Arthur J. Krener, *Reciprocal diffusions and stochastic differential equations of second order*, Stochastics **24** (1988), no. 4, 393–422. MR-972972

[10] J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original. MR-1600720

[11] Shinzo Watanabe, *On discontinuous additive functionals and Lévy measures of a Markov process*, Japan. J. Math. **34** (1964), 53–70. MR-0185675

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)
- Secure publication (LOCKSS[1])
- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]
- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm