

# MONTE CARLO METHODS FOR SENSITIVITY ANALYSIS OF POISSON-DRIVEN STOCHASTIC SYSTEMS, AND APPLICATIONS

CHARLES BORDENAVE,\* *University of California, Berkeley*

GIOVANNI LUCA TORRISI,\*\* *Istituto per le Applicazioni del Calcolo ‘Mauro Picone’*

## Abstract

We extend a result due to Zazanis (1992) on the analyticity of the expectation of suitable functionals of homogeneous Poisson processes with respect to the intensity of the process. As our main result, we provide Monte Carlo estimators for the derivatives. We apply our results to stochastic models which are of interest in stochastic geometry and insurance.

*Keywords:* Importance sampling; marked Poisson process; Monte Carlo estimator; sensitivity analysis; stabilizing functional; stopping set

2000 Mathematics Subject Classification: Primary 60D05; 65C05

Secondary 60G48; 60G55

## 1. Introduction

Let  $N$  be an independently marked homogeneous Poisson process (IMHPP) with points in  $\mathbb{R}^d$  and marks with distribution  $Q$  taking values on some complete separable metric space  $\mathbb{M}$ . Under the probability measure  $P_\lambda$ , the intensity of the Poisson point process is  $\lambda > 0$ . Moreover, let  $\varphi(N)$  be a real-valued functional of the process, and let  $E_\lambda$  be the expectation under  $P_\lambda$ . The function  $\lambda \mapsto E_\lambda[\varphi(N)]$  is known to be smooth in  $\lambda$  under several and different assumptions.

Zazanis [34] focused on functionals depending only on the configuration, up to a finite stopping time, of a homogeneous Poisson process on the half-line. For this class of functionals, he proved that the function  $\lambda \mapsto E_\lambda[\varphi(N)]$  is analytic under a specific moment condition on the functional, and a light-tailed assumption on the stopping time. However, he did not provide an explicit expression for the derivatives.

For one-dimensional IMHPPs, Baccelli *et al.* [5] provided sufficient conditions for the  $m$ -differentiability of  $E_\lambda[\varphi(N)]$  with respect to  $\lambda$ , in a neighborhood of the origin, and closed-form expressions for the derivatives in terms of multiple integrals. However, their method did not address the question of analyticity and their set of conditions was different from ours.

A more general framework was considered in [24]. Let  $N$  be a (not necessarily homogeneous) Poisson process on a locally compact separable metric space with intensity measure  $\Lambda$ . Moreover, let  $\varphi(N)$  be a suitable functional of the process. Molchanov and Zuyev [24] studied the analyticity of the expectation  $E_\Lambda[\varphi(N)]$  with respect to  $\Lambda$ . In particular, they proved that, under some assumptions on  $\varphi$ , the function  $\Lambda \mapsto E_\Lambda[\varphi(N)]$  is analytic on the cone of positive measures.

---

Received 2 November 2006; revision received 25 February 2007.

\* Postal address: Department of Electrical Engineering and Computer Science and Department of Statistics, University of California, 257 Cory Hall, Berkeley, CA 94720-1770, USA.

\*\* Postal address: Istituto per le Applicazioni del Calcolo ‘Mauro Picone’, Consiglio Nazionale delle Ricerche, Viale del Policlinico 137, 00161 Roma, Italy. Email address: torrisi@iac.rm.cnr.it

For Poisson processes with a finite intensity measure  $\Lambda$ , a relevant work is also that of [1], where it was proved that the expectation  $E_\Lambda[\varphi(N)]$  is analytic with respect to a perturbation of  $\Lambda$  by a semigroup.

In this paper we basically rely on Zazanis' paper [34] for the analyticity of  $\lambda \mapsto E_\lambda[\varphi(N)]$ , where  $N$  is an IMHPP on  $\mathbb{R}^d \times \mathbb{M}$ . As our main result, we derive explicit formulae for all the derivatives of  $\lambda \mapsto E_\lambda[\varphi(N)]$ . These formulae provide Monte Carlo methods for the sensitivity analysis of suitable Poisson-driven stochastic systems with respect to the intensity of the process.

There are several motivations for being interested in sensitivity analysis: the main reasons are the applications to optimization and control of complex systems occurring, for instance, in stochastic geometry and insurance. Sensitivity analysis was introduced in [16], and has been addressed by many authors (see, for instance, [14] and the references cited therein). There are mainly three ways to handle this problem: the infinitesimal perturbation analysis (IPA), the likelihood ratio method (LRM), and the rare perturbation analysis (RPA). We refer the reader to [19] and [32] for more insight into the IPA method, and to [30] for more details on the LRM. It is worthwhile to mention the work of Decreusefond [11], where, using Malliavin calculus, it was shown that the IPA, RPA, and LRM could be seen as a part of the stochastic calculus of variations. The main achievement of Decreusefond's paper was that he was potentially able to consider discrete-event systems to be more general than Poisson processes.

As already mentioned, we derive explicit formulae for all the derivatives of  $\lambda \mapsto E_\lambda[\varphi(N)]$ . For this we use the RPA method. Suppose that we wish to compute the derivative  $dE_\lambda[\varphi(N)]/d\lambda$ . We distinguish two different RPA methods: the virtual and the phantom. The virtual RPA method may be attributed to [29], and has been revisited in [4]. Following the ideas of these papers, we evaluate the limit

$$\lim_{\Delta\lambda \rightarrow 0} \frac{E_{\lambda+\Delta\lambda}[\varphi(N)] - E_\lambda[\varphi(N)]}{\Delta\lambda}.$$

The key idea is to use the superposition property of IMHPPs to generate an IMHPP of intensity  $\lambda + \Delta\lambda$  from a small perturbation of an IMHPP of intensity  $\lambda$ . By a coupling argument, an IMHPP of intensity  $\lambda + \Delta\lambda$  is generated from the superposition of two independent IMHPPs of respective intensity  $\lambda$  and  $\Delta\lambda$ . The phantom RPA method was introduced in [8]. Following the approach in this paper, we compute the limit

$$\lim_{\Delta\lambda \rightarrow 0} \frac{E_\lambda[\varphi(N)] - E_{\lambda-\Delta\lambda}[\varphi(N)]}{\Delta\lambda}.$$

The idea is to use the thinning property of IMHPPs to generate an IMHPP of intensity  $\lambda - \Delta\lambda$ : similarly to the previous case, this process is generated from a small perturbation of an IMHPP of intensity  $\lambda$  by a coupling argument. We generalize this approach to compute the  $n$ th-order derivatives  $d^n E_\lambda[\varphi(N)]/d\lambda^n$ .

Our results can be applied to suitable functionals of random sets arising in stochastic geometry. Furthermore, by using importance sampling and large deviations techniques, we show that our results can be applied to ruin probabilities of risk processes with Poisson arrivals and delayed or undelayed claims. In the case of classical risk processes (undelayed claims) we provide an asymptotically optimal Monte Carlo estimator for the first-order derivative of the ruin probability.

The paper is organized as follows. In Section 2 we fix the notation and extend Zazanis' result about analyticity of functionals of homogeneous Poisson processes. In Section 3 we

state our results about  $n$ th-order derivatives of functionals of homogeneous Poisson processes. In Section 4 we prove the results given in Section 3. Finally, in Section 5 we apply our results to stochastic models which are of interest in stochastic geometry and insurance.

## 2. Preliminaries

### 2.1. Notation

Let  $d \geq 1$  be an integer, let  $\mathbb{M}$  be a complete separable metric space, and let  $\mathcal{N}$  be the space of all counting measures on  $\mathbb{R}^d \times \mathbb{M}$ , defined on the Borel  $\sigma$ -field  $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{M})$ , such that each measure  $\mu \in \mathcal{N}$  is simple and locally finite, that is,  $\mu(\{(x, z)\})$  is equal to 0 or 1 for each  $(x, z) \in \mathbb{R}^d \times \mathbb{M}$  and  $\mu$  is finite on each set of the form  $B \times \mathbb{M}$ , where  $B$  is a bounded Borel set. We endow the space  $\mathcal{N}$  with its usual topology (see, for instance, [10] for the details). Any measure in  $\mathcal{N}$  can be represented as

$$\mu = \sum_{n=1}^{\mu(\mathbb{R}^d \times \mathbb{M})} \delta_{(x_n, z_n)} = \sum_{n=1}^{\mu(\mathbb{R}^d \times \mathbb{M})} \delta_{\mathbf{x}_n}, \tag{1}$$

where  $(x_n, z_n) = \mathbf{x}_n \in \mathbb{R}^d \times \mathbb{M}$  and  $\delta_{(x, z)}$ ,  $(x, z) \in \mathbb{R}^d \times \mathbb{M}$ , is the Dirac measure on  $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{M})$ : for any  $B \in \mathcal{B}(\mathbb{R}^d)$  and  $M \in \mathcal{B}(\mathbb{M})$ ,  $\delta_{(x, z)}(B \times M)$  is equal to 1 if  $(x, z) \in B \times M$  and equal to 0 otherwise. The elements of the sequence  $\{z_n\}_{n \geq 1} \subseteq \mathbb{M}$  are called marks. The symbol  $\text{supp}(\mu)$  will denote the support of the counting measure  $\mu \in \mathcal{N}$ , that is, if  $\mu$  is given by (1),  $\text{supp}(\mu) = \{(x_n, z_n)_{n \geq 1}\}$ .

Let  $B_r$  be the closed ball centered in 0 with radius  $r$ , and let  $B(x, r) = x + B_r$  be the closed ball centered at  $x$  with radius  $r$ . If  $K$  is a compact set then, throughout this paper, we denote by  $B_K$  the smallest closed ball centered in 0 which contains  $K$ .

Let  $\mu = \sum_{n \geq 1} \delta_{\mathbf{x}_n} = \sum_{n \geq 1} \delta_{(x_n, z_n)}$  be in  $\mathcal{N}$ . For  $B \in \mathcal{B}(\mathbb{R}^d)$  and  $M \in \mathcal{B}(\mathbb{M})$ , we set

$$\int_{B \times M} \psi(\mathbf{x}) \mu(d\mathbf{x}) = \int_{B \times M} \psi(x, z) \mu(dx \times dz) = \sum_{n \geq 1} \psi(x_n, z_n) \delta_{(x_n, z_n)}(B \times M)$$

for any measurable functional  $\psi: \mathbb{R}^d \times \mathbb{M} \rightarrow \mathbb{R}$  such that the sum is well defined.

Let  $B \subseteq \mathbb{R}^d$  be a Borel set. Throughout this paper, we denote the set of points of  $\text{supp}(\mu)$  in  $B \times \mathbb{M}$  by  $\mu|_B$  and the number of points of  $\mu$  in  $B \times \mathbb{M}$  by  $\mu_B$ . With an abuse of notation,  $|B|$  denotes the Lebesgue measure of  $B$  and, for real numbers  $x \in \mathbb{R}$ , the symbol  $|x|$  denotes the usual absolute value.

For any measurable functional  $\varphi: \mathcal{N} \rightarrow \mathbb{R}$  and  $\mathbf{x} \in \mathbb{R}^d \times \mathbb{M}$ , we define the increments

$$D_{\mathbf{x}}^+ \varphi(\mu) = \varphi(\mu + \delta_{\mathbf{x}}) - \varphi(\mu) \quad \text{and} \quad D_{\mathbf{x}}^- \varphi(\mu) = \varphi(\mu) - \varphi(\mu - \delta_{\mathbf{x}}),$$

where  $D_{\mathbf{x}}^- \varphi(\mu)$  is properly defined only if  $\mathbf{x} \in \text{supp}(\mu)$ . Similarly, if  $\mu' \in \mathcal{N}$ ,

$$D_{\mu'}^+ \varphi(\mu) = \varphi(\mu + \mu') - \varphi(\mu) \quad \text{and} \quad D_{\mu'}^- \varphi(\mu) = \varphi(\mu) - \varphi(\mu - \mu'),$$

where  $D_{\mu'}^- \varphi(\mu)$  is properly defined only if  $\text{supp}(\mu') \subseteq \text{supp}(\mu)$ .

Let  $(\Omega, \mathcal{F}, P)$  be a probability space. A (simple and locally finite) marked point process on  $\mathbb{R}^d$  with marks in  $\mathbb{M}$  is a measurable mapping from  $\Omega$  to  $\mathcal{N}$ . Throughout the paper, we fix a marked point process  $N$  on  $\mathbb{R}^d$  with marks in  $\mathbb{M}$ . For a Borel set  $B \subseteq \mathbb{R}^d$ , define the following  $\sigma$ -field on  $\Omega$ :

$$\mathcal{F}_B = \sigma\{N(C \times M) : C \in \mathcal{B}(\mathbb{R}^d), C \subseteq B, M \in \mathcal{B}(\mathbb{M})\}.$$

Let  $\mathbb{F}$  and  $\mathbb{K}$  respectively denote the family of closed and compact sets of  $\mathbb{R}^d$ . We endow these families with their standard topology (see [22] and [31]). Let  $S: \mathcal{N} \rightarrow \mathbb{F}$  be a measurable mapping. We say that  $S$  is a stopping set if  $S(N)$  is a measurable mapping from  $\Omega$  to  $\mathbb{K}$  such that  $\{S(N) \subseteq K\} \in \mathcal{F}_K$  for each  $K \in \mathbb{K}$ . The stopping  $\sigma$ -field is the following collection on  $\Omega$ :

$$\mathcal{F}_S = \sigma \left\{ F \in \bigvee_{K \in \mathbb{K}} \mathcal{F}_K : F \cap \{S(N) \subseteq K\} \in \mathcal{F}_K \text{ for all } K \in \mathbb{K} \right\}.$$

For details and properties of stopping sets and stopping  $\sigma$ -fields, we refer the reader to [36].

All the random elements considered in this paper are defined on the measurable space  $(\Omega, \mathcal{F})$ . We endow such a space with the family of probability measures  $\{P_\lambda\}_{\lambda>0}$  such that, under  $P_\lambda$ , the marked point process

$$N = \sum_{n \geq 1} \delta_{(X_n, Z_n)} = \sum_{n \geq 1} \delta_{X_n}$$

is an IMHPP of intensity  $\lambda > 0$ , that is, the ground point process  $\{X_n\}_{n \geq 1}$  is a homogeneous Poisson process with intensity  $\lambda$ , the random marks  $\{Z_n\}_{n \geq 1}$  are independent and identically distributed (i.i.d.) with law  $Q$ , and the sequences  $\{X_n\}_{n \geq 1}$  and  $\{Z_n\}_{n \geq 1}$  are independent. We denote by  $E_\lambda$  the expectation associated to  $P_\lambda$ . Note that  $N$  is actually a Poisson point process on  $\mathbb{R}^d \times \mathbb{M}$  with intensity measure  $\lambda\Lambda$ , where

$$\Lambda(dx \times dz) = dx Q(dz)$$

is the product measure on  $\mathbb{R}^d \times \mathbb{M}$  of the Lebesgue measure and  $Q$ .

Although in Section 2.2 and Section 3 we assume that  $\{X_n\}_{n \geq 1}$  is a homogeneous Poisson process on  $\Sigma = \mathbb{R}^d$ , the results therein still hold if  $\Sigma$  is a Borel subset of  $\mathbb{R}^d$  and  $\{X_n\}_{n \geq 1}$  is the restriction on  $\Sigma$  of a homogeneous Poisson process on  $\mathbb{R}^d$  (for instance, note that in Section 5.2 we apply the results of Section 2.2 and Section 3 to stochastic models where  $\{X_n\}_{n \geq 1}$  are the points of a homogeneous Poisson process on  $[0, \infty)$ ).

**2.2. Analyticity of functionals of IMHPPs**

Our analysis is based on a result, due to [34], which can be extended to the context of stopping sets as follows. Let  $\varphi$  be a measurable functional from  $\mathcal{N}$  to  $\mathbb{R}$ , let  $f(\lambda) = E_\lambda[\varphi(N)]$ , let  $f^{(n)}(\lambda) = d^n f(\lambda)/d\lambda^n$ , and let  $[a, b)$  be an interval of the positive half-line. We consider the following assumptions.

There exists a stopping set  $S$  such that  $\varphi(N)$  is  $\mathcal{F}_S$ -measurable. (2)

For any  $\lambda \in [a, b)$ , there exists  $\gamma = \gamma(\lambda) > 1$  such that  $E_\lambda[|\varphi(N)|^\gamma] < \infty$ . (3)

For any  $\lambda \in [a, b)$  there exists  $s = s(\lambda) > 0$  such that  $E_\lambda[\exp(s|B_{S(N)}|)] < \infty$ . (4)

The following theorem holds.

**Theorem 1.** *Assume that (2), (3), and (4) hold. Then  $f(\cdot)$  is analytic on  $[a, b)$ , that is, for a fixed  $x_0 \in [a, b)$ , we have*

$$f(x) = \sum_{n \geq 0} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n, \quad x \in [a, b).$$

Zazanis [34] considered a homogeneous Poisson process  $N$  on the half-line and stopping sets of the form  $S(N) = [0, T(N)]$ , where  $T(N)$  is a stopping time with respect to the natural filtration of the Poisson process. Moreover, he assumed the following stronger condition in place of (3):

$$E_\lambda[\varphi^4(N)] < \infty \quad \text{for any } \lambda \in [a, b].$$

To prove Theorem 1, we need the following lemmas.

**Lemma 1.** *Under assumptions (2) and (4), for all  $\lambda \in [a, b]$ , there exists  $s' = s'(\lambda) > 0$  such that  $E_\lambda[\exp(s'N_{S(N)})] < \infty$ .*

**Lemma 2.** *Under the assumptions of Theorem 1, we have*

$$E_\lambda \left[ |\varphi(N)| \left( \frac{\lambda + \varepsilon}{\lambda} \right)^{N_{S(N)}} \exp(\varepsilon |B_{S(N)}|) \right] < \infty, \quad \varepsilon \in \left( 0, \min \left\{ \frac{s(\gamma - 1)}{2\gamma}, \frac{s'\lambda(\gamma - 1)}{2\gamma} \right\} \right).$$

Here  $\gamma$  is given by assumption (3),  $s$  is given by assumption (4), and  $s'$  is determined by Lemma 1.

*Proof of Lemma 1.* For ease of notation, throughout this proof, we write  $S = S(N)$ . Let  $s$  be given by assumption (4), and set  $C = E_\lambda[e^{s|S|}]$  and  $\delta > e^2\lambda$ . For  $k \geq 0$ , let  $r_k$  be such that  $|B_{r_k}| = k/\delta$  (that is,  $r_k = (k/\delta\pi_d)^{1/d}$ , where  $\pi_d$  is the volume of the ball  $B_1$ ). We note that

$$P_\lambda(N_S > k) \leq P_\lambda \left( |B_S| > \frac{k}{\delta} \right) + P_\lambda(N_{B_{r_k}} > k) \quad \text{for all } k \geq 0. \tag{5}$$

By a standard large deviation estimate for the Poisson distribution (see, for instance, [26, Lemma 1.2]) we have, for all  $k \geq e^2\lambda$ ,

$$P_\lambda(N_{B_{r_k}} > k) \leq \exp \left( -\frac{k}{2} \log \left( \frac{k}{\lambda |B_{r_k}|} \right) \right) = \exp \left( -\frac{k}{2} \log \left( \frac{\delta}{\lambda} \right) \right). \tag{6}$$

Therefore, by (5), (6), and Markov’s inequality, it follows that, for all  $k \geq e^2\lambda$ ,

$$P_\lambda(N_S > k) \leq C \exp \left( -\frac{s}{\delta} k \right) + \exp \left( -\frac{k}{2} \log \left( \frac{\delta}{\lambda} \right) \right).$$

Finally, we easily deduce that, for  $0 < s' < \min\{s/\delta, \frac{1}{2} \log(\delta/\lambda)\}$ ,  $E_\lambda[\exp(s'N_S)] = 1 + (\exp(s') - 1) \sum_{k \geq 0} \exp(s'k) P_\lambda(N_S > k) < \infty$ .

*Proof of Lemma 2.* As above, we set  $S = S(N)$ . The proof is similar to the proof of Lemma 2 in [34]. Following the proof of Lemma 2 in [34] and using Hölder’s inequality (in place of the first application of the Cauchy–Schwarz inequality) and then the Cauchy–Schwarz inequality, we have

$$\begin{aligned} & E_\lambda \left[ |\varphi(N)| \left( \frac{\lambda + \varepsilon}{\lambda} \right)^{N_S} \exp(\varepsilon |B_S|) \right] \\ & \leq (E_\lambda[|\varphi(N)|^\gamma])^{1/\gamma} \left( E_\lambda \left[ \left( \frac{\lambda + \varepsilon}{\lambda} \right)^{(\gamma/(\gamma-1))N_S} \exp \left( \varepsilon \frac{\gamma}{\gamma-1} |B_S| \right) \right] \right)^{(\gamma-1)/\gamma} \\ & \leq (E_\lambda[|\varphi(N)|^\gamma])^{1/\gamma} \left( E_\lambda \left[ \left( \frac{\lambda + \varepsilon}{\lambda} \right)^{(2\gamma/(\gamma-1))N_S} \right] E_\lambda \left[ \exp \left( \varepsilon \frac{2\gamma}{\gamma-1} |B_S| \right) \right] \right)^{(\gamma-1)/2\gamma}. \end{aligned}$$

The claim follows by Lemma 1 and assumptions (3) and (4).

*Proof of Theorem 1.* Zazanis’ [34] result can be extended as stated by Theorem 1. We briefly outline the main changes in the proof: follow the proofs of Theorem 4 and Corollary 2 of [34] replacing Lemma 2 in [34] with Lemma 2 and the so-called Cameron–Martin–Girsanov change of measure by the change of measure

$$\frac{dP_{\lambda,S}}{dP_{a,S}} = \left(\frac{\lambda}{a}\right)^{N_{S(N)}} \exp(-|S(N)|(\lambda - a)), \tag{7}$$

where  $P_{\lambda,S}$  denotes the restriction of  $P_\lambda$  to  $\mathcal{F}_S$  (note that  $P_\lambda \ll P_a$  on the stopping  $\sigma$ -field  $\mathcal{F}_S$  with density (7) due to the results in [36]).

**Remark 1.** A function  $g: \mathbb{R} \rightarrow [0, \infty)$  is said to be absolutely monotonic in  $[a, b)$  if it has derivatives of all orders that satisfy  $g^{(k)}(x) \geq 0$  for all  $x \in (a, b)$  and  $k \geq 0$ . Consider a nonnegative functional  $\varphi(\mu, \lambda)$  on  $\mathcal{N} \times \mathbb{R}^+$ , which depends explicitly on  $\lambda$  in such a way that, for each  $\mu \in \mathcal{N}$ , the function  $\lambda \mapsto \varphi(\mu, \lambda)$  is absolutely monotonic in  $[a, b)$ . If, moreover, assumptions (2), (3), and (4) are satisfied with  $\varphi(\mu, \lambda)$  in place of  $\varphi(\mu)$ , then  $f(\lambda) = E_\lambda[\varphi(N, \lambda)]$  is analytic on  $[a, b)$ . The proof is similar to that of Theorem 1. In particular, note that the absolute monotonicity of  $\lambda \mapsto \varphi(N, \lambda)$  implies the absolute monotonicity of

$$\lambda \mapsto \varphi(N, \lambda) \left(\frac{\lambda}{a}\right)^{N_{S(N)}} \exp(-|S(N)|(\lambda - a))$$

in  $[a, b)$ . Indeed, similarly to [34], we can prove that the function

$$\lambda \mapsto \left(\frac{\lambda}{a}\right)^{N_{S(N)}} \exp(-|S(N)|(\lambda - a))$$

is absolutely monotonic. The claim follows using the fact that the product of two absolutely monotonic functions is an absolutely monotonic function.

### 3. Rare perturbation analysis

Sensitivity analysis is concerned with evaluating derivatives of cost functions with respect to parameters of interest. It plays a central role in identifying the most significant system parameters. In this section we give Monte Carlo methods to estimate the derivatives of the cost function  $f(\lambda) = E_\lambda[\varphi(N)]$ . An application of importance could be the use of such gradient estimates in stochastic gradient algorithms to find the optimal value  $\lambda_0$  that minimizes the cost function.

#### 3.1. Monotone mappings

The following notion of monotonicity is crucial in this work.

**Definition 1.** Let  $S$  be a measurable mapping from  $\mathcal{N}$  to  $\mathbb{F}$ . We say that  $S$  is nonincreasing or nondecreasing if, for any  $\mu_1, \mu_2$  in  $\mathcal{N}$ , the inclusion  $\text{supp}(\mu_1) \subseteq \text{supp}(\mu_2)$  implies that  $S(\mu_1) \supseteq S(\mu_2)$  or  $S(\mu_1) \subseteq S(\mu_2)$ , respectively. The mapping  $S$  is said to be monotone if it is nonincreasing or nondecreasing.

We give a couple of examples as a guide to intuition. Let  $\mu = \{x_n\}_{n \geq 1}$  be a locally finite counting measure on  $[0, \infty)$ . Define the functional  $\varphi(\mu) = 1 \wedge x_1$ , where  $x_1$  is the first point of  $\mu$  on  $[0, \infty)$ , and define the measurable mapping  $S(\mu) = [0, x_1]$ . Then  $S$  is nonincreasing,

but it is not nondecreasing. Instead, if we define  $S(\mu) = [0, 1]$  then  $S$  is nonincreasing and nondecreasing.

### 3.2. First-order derivative

In this subsection we state the result concerning the first-order derivative of  $f(\cdot)$ . Its proof is given in Section 4.

**Theorem 2.** *Under the assumptions of Theorem 1, with  $\gamma$  given in (3) such that  $\gamma > 2$ , and if, moreover, the mapping  $S$  is monotone then, for all  $\lambda \in [a, b]$ ,*

$$f'(\lambda) = E_\lambda[|S(N)|D_{\mathbf{X}}^+\varphi(N)] \tag{8}$$

$$= E_\lambda\left[\frac{N_{S(N)}}{\lambda}D_{\mathbf{X}'}^-\varphi(N)\right], \tag{9}$$

where  $\mathbf{X} = (\xi, \zeta)$  and  $\mathbf{X}' = (\xi', \zeta')$  are random variables on  $\mathbb{R}^d \times \mathbb{M}$ . Given  $S(N)$ ,  $\xi$  is uniformly distributed on  $S(N)$ ;  $\zeta$  is independent of  $N$  and  $\xi$  and has law  $\mathcal{Q}$ . Given the collection of points  $N_{|S(N)}$ ,  $\mathbf{X}' = (\xi', \zeta')$  is uniformly distributed on the collection.

The closed-form formulae provided by (8) and (9) both give a Monte Carlo method to simulate the derivative of  $f(\lambda)$ . We also note that if we consider an IMHPP on  $(-\infty, 0]$  with marks in  $(0, \infty)$ , and the assumptions of Theorem 2 hold, our formulae for the first-order derivative coincide with the corresponding formula in [5] (see [5, Equation 10] with  $k = 1$ ). This easily follows by equality (9) and the forthcoming equalities (37) and (38). A similar remark holds for the  $n$ th-order derivatives (see Theorem 3, below).

### 3.3. Higher-order derivatives

We now generalize Theorem 2, stating the result for the  $n$ th-order derivatives  $f^{(n)}(\lambda)$ . The details of the proof are given in Section 4.

Let  $\varphi$  be a measurable functional from  $\mathcal{N}$  to  $\mathbb{R}$ . As in [6] and [29] for  $\mu \in \mathcal{N}$ ,  $n \geq 1$ , and  $\mathbf{x}_i = (x_i, z_i) \in \mathbb{R}^d \times \mathbb{M}$ ,  $i = 1, \dots, n$ , define

$$\varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu) = \varepsilon(\mathbf{x}_1, \dots, \mathbf{x}_n) \sum_{k=0}^n (-1)^{n-k} \sum_{\pi \in \binom{[n]}{k}} \varphi\left(\mu + \sum_{i \in \pi} \delta_{\mathbf{x}_i}\right), \tag{10}$$

where  $\varepsilon(\mathbf{x}_1, \dots, \mathbf{x}_n) = \mathbf{1}(\{\mathbf{x}_1, \dots, \mathbf{x}_n \text{ are distinct}\})$  and  $\{\binom{[n]}{k}\}$  denotes the collection of all subsets with cardinality  $k$  of  $\{1, \dots, n\}$ . We will also consider the functionals

$$\varphi^{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu) = \varepsilon(\mathbf{x}_1, \dots, \mathbf{x}_n) \varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}\left(\mu - \sum_{i=1}^n \delta_{\mathbf{x}_i}\right), \tag{11}$$

which are properly defined only if  $\mathbf{x}_1, \dots, \mathbf{x}_n \in \text{supp}(\mu)$ . Note that  $\varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}$  (and therefore  $\varphi^{\mathbf{x}_1, \dots, \mathbf{x}_n}$ ) is invariant by permutations in the sense that, for any permutation  $\sigma$  of  $\{1, \dots, n\}$ ,  $\varphi_{\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)}}(\mu) = \varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu)$ . Furthermore, as can be easily seen, reasoning by induction on  $n \geq 1$ , we find that if  $\varepsilon(\mathbf{x}_1, \dots, \mathbf{x}_{n+1}) = 1$  then

$$\varphi_{\mathbf{x}_1, \dots, \mathbf{x}_{n+1}}(\mu) = \varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu + \delta_{\mathbf{x}_{n+1}}) - \varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu) = D_{\mathbf{x}_{n+1}}^+ \varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu) \tag{12}$$

and

$$\varphi^{x_1, \dots, x_{n+1}}(\mu) = \varphi^{x_1, \dots, x_n}(\mu) - \varphi^{x_1, \dots, x_n}(\mu - \delta_{x_{n+1}}) = D_{x_{n+1}}^- \varphi^{x_1, \dots, x_n}(\mu). \tag{13}$$

In particular,  $\varphi_{x_1}(\mu) = D_{x_1}^+ \varphi(\mu)$  and  $\varphi^{x_1}(\mu) = D_{x_1}^- \varphi(\mu)$ . In the following theorem we use the standard convention that the sum over an empty set is 0 and  $k!/(k - n)! = 0$  for  $n > k$ .

**Theorem 3.** *Under the assumptions of Theorem 2, for all  $\lambda \in [a, b)$  and  $n \geq 1$ ,*

$$f^{(n)}(\lambda) = E_\lambda[|S(N)|^n \varphi_{X_1, \dots, X_n}(N)] \tag{14}$$

$$= E_\lambda \left[ \left( \frac{N_{S(N)}}{\lambda} \right)^n \varphi^{X'_1, \dots, X'_n}(N) \right] \tag{15}$$

$$= E_\lambda \left[ \frac{N_{S(N)}!}{(N_{S(N)} - n)! \lambda^n} \varphi^{X''_1, \dots, X''_n}(N) \right], \tag{16}$$

where, for  $1 \leq i \leq n$ ,  $X_i = (\xi_i, \zeta_i)$ ,  $X'_i = (\xi'_i, \zeta'_i)$ , and  $X''_i = (\xi''_i, \zeta''_i)$  are random variables on  $\mathbb{R}^d \times \mathbb{M}$ . Given  $S(N)$ ,  $(\xi_i)_{1 \leq i \leq n}$  are independent and uniformly distributed on  $S(N)$ , and independent of  $N$ ;  $(\zeta_i)_{1 \leq i \leq n}$  are independent, independent of  $N$  and  $(\xi_i)_{1 \leq i \leq n}$ , and have law  $Q$ . Given the collection of points  $N_{|S(N)}$ ,  $(X'_i)_{1 \leq i \leq n}$  are independent and uniformly distributed on the collection;  $\{X''_1, \dots, X''_n\}$  is uniformly distributed on the set of subsets of  $n$  distinct points of  $N_{|S(N)}$ .

Note that (16) implies that  $f^{(n)}(\lambda) = 0$  if  $N_{S(N)} < n$  with probability 1.

Putting together Theorems 1 and 3, we obtain the following corollary.

**Corollary 1.** *Under the assumptions of Theorem 2 and the notation of Theorem 3, for all  $\lambda \in [a, b)$ ,*

$$f(\lambda) = E_a \left[ \sum_{n=0}^{N_{S(N)}} \left( \frac{\lambda - a}{a} \right)^n \frac{N_{S(N)}^n}{n!} \varphi^{X'_1, \dots, X'_n}(N) \right].$$

### 4. Proofs of Theorems 2 and 3

#### 4.1. Integrability lemmas

In the core of the proof of Theorems 2 and 3 we use the integrability of some functionals. In this subsection we prove such integrability results.

We start with a simple continuity result.

**Lemma 3.** *Under the assumptions of Theorem 1, for all  $\alpha \in [0, \gamma(\lambda))$ , the function  $\lambda' \mapsto E_{\lambda'}[|\varphi(N)|^\alpha]$  is defined in an open neighborhood of  $\lambda$  and is continuous at  $\lambda$ .*

*Proof.* Throughout this proof, we set  $S = S(N)$ . The conclusion is trivial for  $\alpha = 0$ . Assume that  $\alpha > 0$ . We prove that

$$\lim_{\varepsilon \rightarrow 0^-} E_{\lambda+\varepsilon}[|\varphi(N)|^\alpha] = E_\lambda[|\varphi(N)|^\alpha]. \tag{17}$$

A similar argument can be used to prove the same limit as  $\varepsilon \rightarrow 0^+$ . Let  $s = s(\lambda) > 0$  be given by assumption (4),  $\beta = \gamma/\alpha > 1$ , and  $\varepsilon \in (\min\{-\lambda, -s(\beta - 1)/\beta\}, 0)$ . By the



Cameron–Martin–Girsanov change of measure (7), it follows that

$$E_{\lambda+\varepsilon}[|\varphi(N)|^\alpha] = E_\lambda \left[ |\varphi(N)|^\alpha \left( \frac{\lambda + \varepsilon}{\lambda} \right)^{N_S} e^{-\varepsilon|S|} \right].$$

By the choice of  $\varepsilon$  we have

$$|\varphi(N)|^\alpha \left( \frac{\lambda + \varepsilon}{\lambda} \right)^{N_S} e^{-\varepsilon|S|} \leq |\varphi(N)|^\alpha e^{-\varepsilon|S|} \leq |\varphi(N)|^\alpha \exp\left(\frac{s(\beta - 1)|S|}{\beta}\right).$$

Now, Hölder’s inequality and assumptions (3) and (4) give

$$E_\lambda \left[ |\varphi(N)|^\alpha \exp\left(\frac{s(\beta - 1)|S|}{\beta}\right) \right] \leq E_\lambda[|\varphi(N)|^\gamma]^{1/\gamma} E_\lambda[e^{s|S|}]^{(\beta-1)/\beta} < \infty.$$

The limit (17) is then a consequence of Lebesgue’s dominated convergence theorem.

For any  $\mu \in \mathcal{N}$ ,  $n \geq 1$ , and  $\mathbf{x}_i = (x_i, z_i) \in \mathbb{R}^d \times \mathbb{M}$ ,  $i = 1, \dots, n$ , define the functionals

$$\psi(\mu) = \int_{(\mathbb{R}^d \times \mathbb{M})^n} |\varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu)| \Lambda(d\mathbf{x}_1) \cdots \Lambda(d\mathbf{x}_n)$$

and (with the convention that the sum over an empty set is 0)

$$\chi(\mu) = \sum_{\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \text{supp}(\mu)} |\varphi^{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu)|,$$

where the sum is taken on sets of  $n$  distinct points of  $\mu$ . The following lemmas hold.

**Lemma 4.** *Under the assumptions of Theorem 1, and if, moreover, the mapping  $S$  is nonincreasing then, for all  $\lambda \in [a, b)$  and  $\alpha \in [1, \gamma)$ ,  $E_\lambda[\psi(N)^\alpha] < \infty$ .*

**Lemma 5.** *Under the assumptions of Theorem 1, and if, moreover the mapping  $S$  is nondecreasing then, for all  $\lambda \in [a, b)$  and  $\alpha \in [1, \gamma)$ ,  $E_\lambda[\chi(N)^\alpha] < \infty$ .*

*Proof of Lemma 4.* For ease of notation, set  $P = P_\lambda$  and  $E = E_\lambda$ . Let  $q > 1$  be such that  $q\alpha \leq \gamma$  and let  $p > 1$  such that  $1/p + 1/q = 1$ . Moreover, let  $\hat{N} = \sum_{n \geq 1} \delta_{(\hat{x}_n, \hat{z}_n)}$  be an IMHPP with intensity  $\Delta\lambda$ , such that  $\hat{Z}_1$  has law  $Q$  and  $\hat{N}$  is independent of  $N$ . Here  $\Delta\lambda$  is chosen so that  $\lambda + \Delta\lambda < b$  and  $E[\exp(2p\alpha\Delta\lambda|S(N)|)] < \infty$ . Reasoning by induction on  $n \geq 1$ , we find that assumption (2) and the monotonicity of  $S$  imply that

$$\varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(N) = 0 \quad \text{for any } (x_1, \dots, x_n) \notin S(N)^n. \tag{18}$$

Indeed, for  $n = 1$ , the  $\mathcal{F}_S$ -measurability of  $\varphi(N)$  and the inclusion  $S(N + \delta_{\mathbf{x}}) \subseteq S(N)$  for all  $\mathbf{x} = (x, z) \in \mathbb{R}^d \times \mathbb{M}$  imply that  $\varphi(N + \delta_{\mathbf{x}}) = \varphi(N)$  for each  $\mathbf{x} = (x, z) \in (\mathbb{R}^d \setminus S(N)) \times \mathbb{M}$ . The general case is proved similarly. Therefore,

$$\psi(N) = \int_{(S(N) \times \mathbb{M})^n} |\varphi_{\mathbf{x}_1, \dots, \mathbf{x}_n}(N)| \Lambda(d\mathbf{x}_1) \cdots \Lambda(d\mathbf{x}_n).$$

By the superposition property of Poisson processes,  $N + \hat{N}$  is an IMHPP with intensity  $\lambda + \Delta\lambda$ . It follows that

$$\begin{aligned} \psi(N) &\leq |S(N)|^n \sum_{k=0}^n \binom{n}{k} |S(N)|^{-k} \int_{(S(N) \times \mathbb{M})^k} \left| \varphi\left(N + \sum_{i=1}^k \delta_{\mathbf{x}_i}\right) \right| \Lambda(d\mathbf{x}_1) \cdots \Lambda(d\mathbf{x}_k) \\ &\leq |S(N)|^n \sum_{k=0}^n \binom{n}{k} \mathbb{E}[|\varphi(N + \hat{N})| \mid \hat{N}_{S(N)} = k, N] \\ &\leq |S(N)|^n \mathbb{E}[|\varphi(N + \hat{N})| \mid N] \sum_{k=0}^n \binom{n}{k} \mathbb{P}(\hat{N}_{S(N)} = k \mid N)^{-1} \\ &\leq |S(N)|^n \mathbb{E}[|\varphi(N + \hat{N})| \mid N] \sum_{k=0}^n \binom{n}{k} \frac{k!}{(\Delta\lambda |S(N)|)^k} e^{\Delta\lambda |S(N)|} \\ &= \frac{n!}{(\Delta\lambda)^n} e^{\Delta\lambda |S(N)|} \mathbb{E}[|\varphi(N + \hat{N})| \mid N] \sum_{k=0}^n \frac{(\Delta\lambda |S(N)|)^{n-k}}{(n-k)!} \\ &\leq \frac{n!}{(\Delta\lambda)^n} e^{2\Delta\lambda |S(N)|} \mathbb{E}[|\varphi(N + \hat{N})| \mid N]. \end{aligned}$$

Using Jensen’s inequality and Hölder’s inequality, we deduce that

$$\begin{aligned} \mathbb{E}[\psi(N)^\alpha] &\leq \left(\frac{n!}{(\Delta\lambda)^n}\right)^\alpha \mathbb{E}[e^{2\alpha\Delta\lambda |S(N)|} (\mathbb{E}[|\varphi(N + \hat{N})| \mid N])^\alpha] \\ &\leq \left(\frac{n!}{(\Delta\lambda)^n}\right)^\alpha \mathbb{E}[e^{2\alpha\Delta\lambda |S(N)|} \mathbb{E}[|\varphi(N + \hat{N})|^\alpha \mid N]] \\ &\leq \left(\frac{n!}{(\Delta\lambda)^n}\right)^\alpha \mathbb{E}[e^{2p\alpha\Delta\lambda |S(N)|}]^{1/p} \mathbb{E}[|\varphi(N + \hat{N})|^{q\alpha}]^{1/q} \\ &< \infty. \end{aligned}$$

*Proof of Lemma 5.* Set  $\mathbb{P} = \mathbb{P}_\lambda$  and  $\mathbb{E} = \mathbb{E}_\lambda$ , and let  $N^{\otimes n}$  and  $N_{|S(N)}^{\otimes n}$  respectively be the set of the  $n$ -tuples of  $n$  distinct points of  $N$  and  $N_{|S(N)}$ . Let  $p, q > 1$  be such that  $\alpha q \leq \gamma$  and  $1/p + 1/q = 1$ . Let  $\{\beta_n\}_{n \geq 1}$  be an i.i.d. sequence of Bernoulli random variables, independent of  $N$  and defined by

$$\mathbb{P}(\beta_n = 0) = 1 - \mathbb{P}(\beta_n = 1) = \frac{\Delta\lambda}{\lambda}.$$

Consider the thinned IMHPP of intensity  $\lambda - \Delta\lambda$  given by  $\tilde{N} = \sum_{n \geq 1} \beta_n \delta_{(X_n, Z_n)}$ . Let  $s > 0$  be such that  $\mathbb{E}[\exp(sN_{S(N)})] < \infty$  (see Lemma 1). Here we choose  $\Delta\lambda$  in such a way that  $2p\alpha \log(\lambda/(\lambda - \Delta\lambda)) < s$ . Reasoning by induction on  $n \geq 1$ , it can be proved that assumption (2) and the monotonicity of  $S$  imply that

$$\varphi^{X_1, \dots, X_n}(N) = 0 \quad \text{for } X_1, \dots, X_n \in \text{supp}(N): X_1, \dots, X_n \notin S(N). \tag{19}$$

Indeed, for  $n = 1$ , the  $\mathcal{F}_S$ -measurability of  $\varphi(N)$  and the inclusion  $S(N - \delta_X) \subseteq S(N)$  for all  $X = (X, Z) \in \text{supp}(N)$  imply that  $\varphi(N - \delta_X) = \varphi(N)$  for each  $X \in \text{supp}(N)$  such that

$X \in \mathbb{R}^d \setminus S(N)$ . The general case is proved similarly. Therefore,

$$\chi(N) = \sum_{\{X_1^*, \dots, X_n^*\} \subset \text{supp}(N)} |\varphi^{X_1^*, \dots, X_n^*}(N)| \tag{20}$$

$$= \frac{1}{n!} \sum_{(X_1^*, \dots, X_n^*) \in N^{\otimes n}} |\varphi^{X_1^*, \dots, X_n^*}(N)| \tag{21}$$

$$= \frac{1}{n!} \sum_{(X_1^*, \dots, X_n^*) \in N_{|S(N)}^{\otimes n}} |\varphi^{X_1^*, \dots, X_n^*}(N)| \tag{22}$$

$$= \frac{1}{n!} \mathbb{E}[(N_{S(N)})^n |\varphi^{X'_1, \dots, X'_n}(N)| \mid N], \tag{23}$$

where the equality in (21) follows from the invariance by permutations of  $\varphi^{x_1, \dots, x_n}(\mu)$ , the equality in (22) follows by (19), and the equality in (23) follows by the definition of  $(X'_i)_{1 \leq i \leq n}$  (see the statement of Theorem 3). If  $N_{S(N)} < n$  then  $\chi(N) = 0$ . On the other hand, if  $N_{S(N)} \geq n$ , we deduce that

$$\begin{aligned} \chi(N) &\leq \mathbb{E}[(N_{S(N)})^n |\varphi^{X'_1, \dots, X'_n}(N)| \mid N] \\ &\leq (N_{S(N)})^n \mathbb{E} \left[ \sum_{k=0}^n \sum_{\{i_1, \dots, i_k\} \in \binom{[n]}{k}} \left| \varphi \left( N - \sum_{j=1}^k \delta_{X'_{i_j}} \right) \right| \mid N \right] \\ &= (N_{S(N)})^n \sum_{k=0}^n \binom{n}{k} \mathbb{E}[|\varphi(\tilde{N})| \mid N, N_{S(N)} - \tilde{N}_{S(N)} = k] \\ &\leq (N_{S(N)})^n \mathbb{E}[|\varphi(\tilde{N})| \mid N] \sum_{k=0}^n \binom{n}{k} \mathbb{P}(N_{S(N)} - \tilde{N}_{S(N)} = k \mid N)^{-1} \\ &= (N_{S(N)})^n \mathbb{E}[|\varphi(\tilde{N})| \mid N] \sum_{k=0}^n \binom{n}{k} \frac{k! (N_{S(N)} - k)!}{N_{S(N)}!} \left( \frac{\Delta\lambda}{\lambda} \right)^{-k} \left( 1 - \frac{\Delta\lambda}{\lambda} \right)^{k - N_{S(N)}} \\ &\leq K (N_{S(N)})^n \left( \frac{\lambda}{\lambda - \Delta\lambda} \right)^{N_{S(N)}} \mathbb{E}[|\varphi(\tilde{N})| \mid N], \end{aligned}$$

where  $K = (\lambda/\Delta\lambda)^n \sum_{k=0}^n k! \binom{n}{k}$ . Finally, using Jensen’s inequality, Hölder’s inequality, and Cauchy–Schwartz’s inequality, we obtain

$$\begin{aligned} \mathbb{E}[\chi(N)^\alpha] &\leq K^\alpha \mathbb{E} \left[ (N_{S(N)})^{\alpha n} \left( \frac{\lambda}{\lambda - \Delta\lambda} \right)^{\alpha N_{S(N)}} |\varphi(\tilde{N})|^\alpha \right] \\ &\leq K^\alpha \mathbb{E} \left[ (N_{S(N)})^{p\alpha n} \left( \frac{\lambda}{\lambda - \Delta\lambda} \right)^{p\alpha N_{S(N)}} \right]^{1/p} \mathbb{E}[|\varphi(\tilde{N})|^{q\alpha}]^{1/q} \\ &\leq K^\alpha \mathbb{E}[(N_{S(N)})^{2p\alpha n}]^{1/2p} \mathbb{E} \left[ \left( \frac{\lambda}{\lambda - \Delta\lambda} \right)^{2p\alpha N_{S(N)}} \right]^{1/2p} \mathbb{E}[|\varphi(\tilde{N})|^{q\alpha}]^{1/q} \\ &< \infty. \end{aligned}$$

**4.2. Case of nonincreasing mappings**

In this subsection we prove the closed-form formulae given by (8) and (14) in the case of nonincreasing mappings  $S$ . More precisely, the following propositions hold.

**Proposition 1.** *Under the assumptions of Theorem 2, and if, moreover, the mapping  $S$  is non-increasing then, for all  $\lambda \in [a, b)$ , (8) holds.*

**Proposition 2.** *Under the assumptions of Theorem 2, and if, moreover, the mapping  $S$  is non-increasing then, for all  $\lambda \in [a, b)$  and  $n \geq 1$ , (14) holds.*

We start by proving Proposition 1. The proof is based on the virtual rare perturbation method considered in [4].

*Proof of Proposition 1.* For ease of notation, we set  $P = P_\lambda$  and  $E = E_\lambda$ . A straightforward computation gives

$$\begin{aligned} E[|D_{\mathbf{X}}^+ \varphi(N)| \mid N] &= \frac{1}{|S(N)|} \int_{S(N) \times \mathbb{M}} |D_{\mathbf{x}}^+ \varphi(N)| \Lambda(d\mathbf{x}) \\ &= \frac{1}{|S(N)|} \int_{\mathbb{R}^d \times \mathbb{M}} |D_{\mathbf{x}}^+ \varphi(N)| \Lambda(d\mathbf{x}) \end{aligned} \tag{24}$$

almost surely (a.s.), where the latter equality follows by assumption (2) and by the assumption that  $S$  is nonincreasing. Indeed, as in the proof of Lemma 4, the  $\mathcal{F}_S$ -measurability of  $\varphi(N)$  and the inclusion  $S(N + \delta_{\mathbf{x}}) \subseteq S(N)$  for each  $\mathbf{x} \in \mathbb{R}^d \times \mathbb{M}$  imply that  $\varphi(N + \delta_{\mathbf{x}}) = \varphi(N)$  for each  $\mathbf{x} = (x, z) \in (\mathbb{R}^d \setminus S(N)) \times \mathbb{M}$ . Thus, the integrability of the random variable  $|S(N)|D_{\mathbf{X}}^+ \varphi(N)$  follows by Lemma 4. Now, as in Lemma 4, let  $\hat{N} = \sum_{n \geq 1} \delta_{(\hat{x}_n, \hat{z}_n)}$  be an IMHPP with intensity  $\Delta\lambda$  such that  $\hat{Z}_1$  has law  $Q$  and  $\hat{N}$  is independent of  $N$ . By the superposition property of Poisson processes,  $N + \hat{N}$  is an IMHPP with intensity  $\lambda + \Delta\lambda$ . Here we choose small enough  $\Delta\lambda$  so that  $\lambda + \Delta\lambda < b$ . Owing to the monotonicity of  $S$ , we have  $S(N + \hat{N}) \subseteq S(N)$ , and so the  $\mathcal{F}_S$ -measurability of  $\varphi(N)$  yields

$$\varphi(N) = \varphi(N|_{S(N)}) \quad \text{and} \quad \varphi(N + \hat{N}) = \varphi((N + \hat{N})|_{S(N)}). \tag{25}$$

We then note that

$$\begin{aligned} \frac{f(\lambda + \Delta\lambda) - f(\lambda)}{\Delta\lambda} &= \frac{E[D_{\hat{N}}^+ \varphi(N)]}{\Delta\lambda} \\ &= \frac{1}{\Delta\lambda} E \left[ \sum_{k \geq 1} \mathbf{1}(\hat{N}_{S(N)} = k) D_{\hat{N}_{S(N)}}^+ \varphi(N) \right] \\ &= \frac{1}{\Delta\lambda} E[\mathbf{1}(\hat{N}_{S(N)} = 1) D_{\hat{N}_{S(N)}}^+ \varphi(N)] \end{aligned} \tag{26}$$

$$+ \frac{1}{\Delta\lambda} E[\mathbf{1}(\hat{N}_{S(N)} \geq 2) D_{\hat{N}_{S(N)}}^+ \varphi(N)], \tag{27}$$

where the second equality follows noticing that by (25) on  $\{\hat{N}_{S(N)} = 0\}$  we have  $\varphi(N) = \varphi(N + \hat{N})$ . Fix  $\alpha \in (2, \gamma(\lambda))$ , by Lemma 3, the function  $\lambda' \mapsto E[|\varphi(N)|^\alpha]$  is continuous at  $\lambda$ . Therefore, there exists a positive constant  $C > 0$  such that  $E[|\varphi(N)|^\alpha] < C^\alpha$  and  $E[|\varphi(N + \hat{N})|^\alpha] < C^\alpha$  for small enough  $\Delta\lambda$ . Using Hölder’s inequality and Minkowski’s inequality, we have

$$\begin{aligned} |E[\mathbf{1}(\hat{N}_{S(N)} \geq 2) D_{\hat{N}_{S(N)}}^+ \varphi(N)]| &\leq (P(\hat{N}_{S(N)} \geq 2))^{1-1/\alpha} (E[|\varphi(N + \hat{N}_{S(N)}) - \varphi(N)|^\alpha])^{1/\alpha} \\ &\leq 2C \left( E \left[ \sum_{k \geq 2} \frac{(\Delta\lambda)^k |S(N)|^k}{k!} e^{-\Delta\lambda |S(N)|} \right] \right)^{1-1/\alpha} \\ &\leq 2C (\Delta\lambda)^{2(1-1/\alpha)} (E[|S(N)|^2])^{1-1/\alpha}. \end{aligned} \tag{28}$$

By assumption (4) we have  $E[|S(N)|^2] < \infty$ . Therefore, by inequality (28), it follows that the term in (27) goes to 0 as  $\Delta\lambda \rightarrow 0$ . Since  $\hat{N}$  is independent of  $N$ , it follows that

$$\begin{aligned} E[\mathbf{1}(\hat{N}_{S(N)} = 1)D_{\hat{N}_{S(N)}}^+ \varphi(N)] &= E[E[\mathbf{1}(\hat{N}_{S(N)} = 1)D_{\hat{N}_{S(N)}}^+ \varphi(N) \mid N]] \\ &= E[\Delta\lambda|S(N)|e^{-\Delta\lambda|S(N)|} E[D_{\hat{N}_{S(N)}}^+ \varphi(N) \mid N, \hat{N}_{S(N)} = 1]] \\ &= E[\Delta\lambda|S(N)|e^{-\Delta\lambda|S(N)|} E[D_{\mathbf{X}}^+ \varphi(N) \mid N]] \\ &= \Delta\lambda E[|S(N)|e^{-\Delta\lambda|S(N)|} D_{\mathbf{X}}^+ \varphi(N)]. \end{aligned}$$

Thus, by the dominated convergence theorem, the term in (26) converges to  $E[|S(N)|D_{\mathbf{X}}^+ \varphi(N)]$  as  $\Delta\lambda \rightarrow 0$ .

*Proof of Proposition 2.* Set  $E = E_\lambda$ , and note that by (18) we have

$$\psi(N) = E[|S(N)|^n |\varphi_{X_1, \dots, X_n}(N)| \mid N].$$

Thus, the integrability of  $|S(N)|^n \varphi_{X_1, \dots, X_n}(N)$  for any  $n \geq 1$  follows by Lemma 4. We prove (14) by induction on  $n \geq 1$ . As already shown, it holds for  $n = 1$ . Let  $\tilde{\psi}$  be the functional defined as  $\psi$  without the absolute value. By (2) and (18), it follows that  $\tilde{\psi}(N)$  is  $\mathcal{F}_S$ -measurable. Assume the inductive hypothesis  $f^{(n)}(\lambda) = E[\tilde{\psi}(N)]$  for  $n > 1$ . Fix  $\alpha \in (2, \gamma)$ , by Lemma 4 we have  $E[|\tilde{\psi}(N)|^\alpha] < \infty$ . Define the random variable  $\mathbf{X}_{n+1} = (\xi_{n+1}, \zeta_{n+1})$  with values on  $\mathbb{R}^d \times \mathbb{M}$  as follows: given  $S(N)$ ,  $\xi_{n+1}$  is uniformly distributed on  $S(N)$  and is independent of  $N$  and  $\mathbf{X}_1, \dots, \mathbf{X}_n$ ;  $\zeta_{n+1}$  has law  $Q$  and is independent of  $N, \mathbf{X}_1, \dots, \mathbf{X}_n$ , and  $\xi_{n+1}$ . By Proposition 1 we obtain

$$f^{(n+1)}(\lambda) = E[|S(N)|D_{\mathbf{X}_{n+1}}^+ \tilde{\psi}(N)].$$

The conclusion follows noting that, by (12) and (18), we have

$$\begin{aligned} E[|S(N)|D_{\mathbf{X}_{n+1}}^+ \tilde{\psi}(N)] &= \int_{\mathbb{R}^d \times \mathbb{M}} E[D_{\mathbf{x}}^+ \tilde{\psi}(N)] \Lambda(d\mathbf{x}) \\ &= \int_{(\mathbb{R}^d \times \mathbb{M})^{n+1}} E[\varphi_{\mathbf{x}_1, \dots, \mathbf{x}_{n+1}}(N)] \Lambda(d\mathbf{x}_1) \cdots \Lambda(d\mathbf{x}_{n+1}) \\ &= E[|S(N)|^{n+1} \varphi_{\mathbf{X}_1, \dots, \mathbf{X}_{n+1}}(N)]. \end{aligned}$$

### 4.3. Case of nondecreasing mappings

In this subsection we prove the closed-form formulae given by (9) and (15) in the case of nondecreasing mappings  $S$ . More precisely, the following propositions hold.

**Proposition 3.** *Under the assumptions of Theorem 2, and if, moreover, the mapping  $S$  is non-decreasing then, for all  $\lambda \in [a, b)$ ,  $f'(\lambda)$  equals the term in (9).*

**Proposition 4.** *Under the assumptions of Theorem 2, and if, moreover, the mapping  $S$  is non-decreasing then, for all  $\lambda \in [a, b)$  and  $n \geq 1$ ,  $f^{(n)}(\lambda)$  equals the term in (15).*

We first prove Proposition 3. For this we use the so-called phantom rare perturbation method introduced in [8].

*Proof of Proposition 3.* Set  $P = P_\lambda$  and  $E = E_\lambda$ . As in the proof of Lemma 5, the  $\mathcal{F}_S$ -measurability of  $\varphi(N)$  and the inclusion  $S(N - \delta_X) \subseteq S(N)$  for all  $X = (X, Z) \in \text{supp}(N)$  imply that  $\varphi(N - \delta_X) = \varphi(N)$  for each  $X \in \text{supp}(N)$  such that  $X \in \mathbb{R}^d \setminus S(N)$ . Therefore,

$$\int_{S(N) \times \mathbb{M}} |D_x^- \varphi(N)| N(d\mathbf{x}) = \int_{\mathbb{R}^d \times \mathbb{M}} |D_x^- \varphi(N)| N(d\mathbf{x}). \tag{29}$$

Thus, the integrability of the random variable  $\int_{S(N) \times \mathbb{M}} D_x^- \varphi(N) N(d\mathbf{x})$  follows from Lemma 5. Now note that

$$\begin{aligned} \mathbb{E} \left[ \frac{N_{S(N)}}{\lambda} D_{X'}^- \varphi(N) \right] &= \mathbb{E} \left[ \mathbb{E} \left[ \frac{N_{S(N)}}{\lambda} D_{X'}^- \varphi(N) \mid N_{|S(N)} \right] \right] \\ &= \mathbb{E} \left[ \frac{1}{\lambda} \int_{S(N) \times \mathbb{M}} D_x^- \varphi(N) N(d\mathbf{x}) \right]. \end{aligned} \tag{30}$$

Finally, we show that

$$f'(\lambda) = \mathbb{E} \left[ \frac{1}{\lambda} \int_{S(N) \times \mathbb{M}} D_x^- \varphi(N) N(d\mathbf{x}) \right].$$

Let  $\{\beta_n\}_{n \geq 1}$  be the sequence of Bernoulli random variables defined in the proof of Lemma 5. Consider the thinned IMHPP of intensity  $\lambda - \Delta\lambda$  given by  $\tilde{N} = \sum_{n \geq 1} \beta_n \delta_{(X_n, Z_n)}$ . By assumption (2) and the monotonicity of  $S$ , it follows that  $\varphi(N) = \varphi(N_{|S(N)})$  and  $\varphi(\tilde{N}) = \varphi(\tilde{N}_{|S(N)})$ . By the independence of  $\{\beta_n\}_{n \geq 1}$  and  $N$ , we have, for  $0 \leq k \leq N_{S(N)}$ ,

$$P(N_{S(N)} - \tilde{N}_{S(N)} = k \mid N) = \binom{N_{S(N)}}{k} \left( \frac{\Delta\lambda}{\lambda} \right)^k \left( 1 - \frac{\Delta\lambda}{\lambda} \right)^{N_{S(N)} - k}.$$

This equation implies that

$$\begin{aligned} &\mathbb{E}[\mathbf{1}(N_{S(N)} - \tilde{N}_{S(N)} = k)(\varphi(N) - \varphi(\tilde{N}))] \\ &= \mathbb{E} \left[ \binom{N_{S(N)}}{k} \left( \frac{\Delta\lambda}{\lambda} \right)^k \left( 1 - \frac{\Delta\lambda}{\lambda} \right)^{N_{S(N)} - k} \right. \\ &\quad \left. \times \mathbb{E}[\varphi(N) - \varphi(\tilde{N}) \mid N, N_{S(N)} - \tilde{N}_{S(N)} = k] \right]. \end{aligned}$$

Since  $\varphi(N) = \varphi(N_{|S(N)})$  and  $\varphi(\tilde{N}) = \varphi(\tilde{N}_{|S(N)})$ , we have  $\mathbb{E}[\mathbf{1}(N_{S(N)} - \tilde{N}_{S(N)} = 0)(\varphi(N) - \varphi(\tilde{N}))] = 0$ . Therefore,

$$\begin{aligned} &\frac{f(\lambda) - f(\lambda - \Delta\lambda)}{\Delta\lambda} \\ &= \frac{\mathbb{E}[\varphi(N) - \varphi(\tilde{N})]}{\Delta\lambda} \\ &= \frac{1}{\Delta\lambda} \mathbb{E} \left[ \sum_{k \geq 1} \mathbf{1}(N_{S(N)} - \tilde{N}_{S(N)} = k)(\varphi(N) - \varphi(\tilde{N})) \right] \\ &= \frac{1}{\Delta\lambda} \mathbb{E} \left[ N_{S(N)} \left( \frac{\Delta\lambda}{\lambda} \right) \left( 1 - \frac{\Delta\lambda}{\lambda} \right)^{N_{S(N)} - 1} \mathbb{E}[\varphi(N) - \varphi(\tilde{N}) \mid N, N_{S(N)} - \tilde{N}_{S(N)} = 1] \right] \end{aligned} \tag{31}$$

$$+ \frac{1}{\Delta\lambda} \mathbb{E}[\mathbf{1}(N_{S(N)} - \tilde{N}_{S(N)} \geq 2)(\varphi(N) - \varphi(\tilde{N}))]. \tag{32}$$

We note that, given  $N$  and the event  $\{N_{S(N)} - \tilde{N}_{S(N)} = 1\}$ , the law of the random variable  $\varphi(\tilde{N})$  is equal to the law of  $\varphi(N - \delta_{X'})$ . Thus,

$$E[\varphi(N) - \varphi(\tilde{N}) \mid N, N_{S(N)} - \tilde{N}_{S(N)} = 1] = \frac{1}{N_{S(N)}} \int_{S(N) \times \mathbb{M}} (\varphi(N) - \varphi(N - \delta_x)) N(\mathbf{d}x). \tag{33}$$

By the dominated convergence theorem and (33), it follows that, as  $\Delta\lambda \rightarrow 0$ , the term in (31) goes to

$$E\left[\frac{1}{\lambda} \int_{S(N) \times \mathbb{M}} D_x^- \varphi(N) N(\mathbf{d}x)\right].$$

The proof of the proposition is complete if we prove that the term in (32) goes to 0 as  $\Delta\lambda \rightarrow 0$ . Fix  $\alpha \in (2, \gamma(\lambda))$ , by Lemma 3, the function  $\lambda' \rightarrow E[|\varphi(N)|^\alpha]$  is continuous at  $\lambda$ . Therefore, there exists a positive constant  $C > 0$  such that  $E[|\varphi(N)|^\alpha] < C^\alpha$  and  $E[|\varphi(\tilde{N})|^\alpha] < C^\alpha$  for small enough  $\Delta\lambda$ . Using Hölder’s inequality and Minkowski’s inequality, we have

$$\begin{aligned} & |E[\mathbf{1}(N_{S(N)} - \tilde{N}_{S(N)} \geq 2)(\varphi(N) - \varphi(\tilde{N}))]| \\ & \leq (P(N_{S(N)} - \tilde{N}_{S(N)} \geq 2))^{1-1/\alpha} (E[|\varphi(N) - \varphi(\tilde{N})|^\alpha])^{1/\alpha} \\ & \leq 2C \left( E\left[ \sum_{k=2}^{N_{S(N)}} \binom{N_{S(N)}}{k} \left(\frac{\Delta\lambda}{\lambda}\right)^k \left(1 - \frac{\Delta\lambda}{\lambda}\right)^{N_{S(N)}-k} \right] \right)^{1-1/\alpha}. \end{aligned} \tag{34}$$

As can be easily checked, for any  $n \geq 2$  and  $p \in (0, 1)$ ,

$$\sum_{m=2}^n \binom{n}{m} p^m (1-p)^{n-m} \leq \frac{1}{2} n^2 p^2. \tag{35}$$

Thus, by (34) and (35), the absolute value of the term in (32) can be bounded from above by

$$\frac{2C}{\Delta\lambda} \left( \frac{1}{2} \left(\frac{\Delta\lambda}{\lambda}\right)^2 E[N_{S(N)}^2] \right)^{1-1/\alpha},$$

and this quantity goes to 0 as  $\Delta\lambda \rightarrow 0$ , since  $E[N_{S(N)}^2] < \infty$  by Lemma 1.

*Proof of Proposition 4.* Again, set  $E = E_\lambda$ . By (20)–(23), we obtain

$$E\left[\left(\frac{N_{S(N)}}{\lambda}\right)^n |\varphi^{X'_1, \dots, X'_n}(N)| \mid N\right] = \frac{n! \chi(N)}{\lambda^n}.$$

Thus, the integrability of  $(N_{S(N)}/\lambda)^n \varphi^{X'_1, \dots, X'_n}(N)$  for any  $n \geq 1$  follows by Lemma 5. We prove (15) by induction on  $n \geq 1$ . By Proposition 3, (15) holds for  $n = 1$ . Let  $\tilde{\chi}$  be the functional defined by

$$\tilde{\chi}(\mu) = \frac{n!}{\lambda^n} \sum_{\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \text{supp}(\mu)} \varphi^{\mathbf{x}_1, \dots, \mathbf{x}_n}(\mu).$$

Let  $N_{|S(N)}^{\otimes n}$  denote the set of the  $n$ -tuples of  $n$  distinct points of  $N_{|S(N)}$ . Since

$$\tilde{\chi}(N) = \frac{1}{\lambda^n} \sum_{(X_1^*, \dots, X_n^*) \in N_{|S(N)}^{\otimes n}} \varphi^{X_1^*, \dots, X_n^*}(N)$$

(see (20)–(22)), we know that  $\tilde{\chi}(N)$  is  $\mathcal{F}_S$ -measurable. Moreover, for each  $n \geq 1$ ,

$$\tilde{\chi}(N) = \mathbb{E} \left[ \left( \frac{N_{S(N)}}{\lambda} \right)^n \varphi^{X'_1, \dots, X'_n}(N) \mid N \right] \tag{36}$$

(see (22)–(23)). Assume the inductive hypothesis  $f^{(n)}(\lambda) = \mathbb{E}[\tilde{\chi}(N)]$  for  $n > 1$ . Fix  $\alpha \in (2, \gamma)$ . By Lemma 5 we have  $\mathbb{E}[|\tilde{\chi}(N)|^\alpha] < \infty$ . Let  $X'_{n+1}$  be a random variable on  $\mathbb{R}^d \times \mathbb{M}$  such that, given  $N_{|S(N)}$ ,  $X'_{n+1}$  is independent of  $(X'_i)_{1 \leq i \leq n}$  and uniformly distributed on the collection  $N_{|S(N)}$ . By Proposition 3 we obtain

$$f^{(n+1)}(\lambda) = \mathbb{E} \left[ \frac{N_{S(N)}}{\lambda} D_{X'_{n+1}}^- \tilde{\chi}(N) \right].$$

The conclusion follows noticing that, by (13) and (36), we have

$$\begin{aligned} & \mathbb{E}[D_{X'_{n+1}}^- \tilde{\chi}(N) \mid N] \\ &= \mathbb{E} \left[ \left( \frac{N_{S(N)}}{\lambda} \right)^n (\varphi^{X'_1, \dots, X'_n}(N) - \varphi^{X'_1, \dots, X'_n}(N - \delta_{X'_{n+1}})) \mid N \right] \\ &= \mathbb{E} \left[ \left( \frac{N_{S(N)}}{\lambda} \right)^n \varphi^{X'_1, \dots, X'_{n+1}}(N) \mid N \right]. \end{aligned}$$

**4.4. Proof of Theorem 2**

For ease of notation, we again set  $\mathbb{E} = \mathbb{E}_\lambda$ . In view of Propositions 1 and 3, it is sufficient to show that

$$\mathbb{E}[|S(N)| D_X^+ \varphi(N)] = \mathbb{E} \left[ \frac{N_{S(N)}}{\lambda} D_X^- \varphi(N) \right].$$

Arguing as for (24), we have

$$\mathbb{E}[D_X^+ \varphi(N) \mid N] = \frac{1}{|S(N)|} \int_{\mathbb{R}^d \times \mathbb{M}} D_x^+ \varphi(N) \Lambda(\mathbf{d}\mathbf{x}) \quad \text{a.s.}$$

Therefore,

$$\mathbb{E}[|S(N)| D_X^+ \varphi(N)] = \int_{\mathbb{R}^d \times \mathbb{M}} \mathbb{E}[D_x^+ \varphi(N)] \Lambda(\mathbf{d}\mathbf{x}). \tag{37}$$

On the other hand, using the same argument as for (29) and the Slivnyak–Mecke theorem (see, for instance, [10]), we obtain

$$\begin{aligned} \mathbb{E} \left[ \int_{S(N) \times \mathbb{M}} D_x^- \varphi(N) N(\mathbf{d}\mathbf{x}) \right] &= \mathbb{E} \left[ \int_{\mathbb{R}^d \times \mathbb{M}} D_x^- \varphi(N) N(\mathbf{d}\mathbf{x}) \right] \\ &= \lambda \int_{\mathbb{R}^d \times \mathbb{M}} \mathbb{E}[D_x^+ \varphi(N)] \Lambda(\mathbf{d}\mathbf{x}). \end{aligned} \tag{38}$$

The conclusion follows by (30), (37), and (38), which does not depend on the monotonicity of  $S$ .



### 4.5. Proof of Theorem 3

As usual, set  $E = E_\lambda$ . Let  $\tilde{\psi}$  and  $\tilde{\chi}$  be the functionals defined in the proofs of Proposition 2 and Proposition 4, respectively. Equations (14)–(15) will follow if we prove that

$$E[\tilde{\psi}(N)] = E[\tilde{\chi}(N)]. \tag{39}$$

Indeed, by the proof of Proposition 2, if  $S$  is nonincreasing, we have  $f^{(n)}(\lambda) = E[\tilde{\psi}(N)]$ , and, by the proof of Proposition 4, if  $S$  is nondecreasing, we have  $f^{(n)}(\lambda) = E[\tilde{\chi}(N)]$ . Equality (39) follows since, by the extended Slivnyak–Campbell theorem (see [25]) and the invariance by permutation of  $\varphi^{x_1, \dots, x_n}(\mu)$ , we have

$$\begin{aligned} E[\tilde{\chi}(N)] &= \frac{n!}{\lambda^n} E \left[ \sum_{\{X_1^*, \dots, X_n^*\} \subset \text{supp}(N)} \varphi^{X_1^*, \dots, X_n^*}(N) \right] \\ &= \int_{(\mathbb{R}^d \times \mathbb{M})^n} E \left[ \varphi^{x_1, \dots, x_n} \left( N + \sum_{i=1}^n \delta_{x_i} \right) \right] \Lambda(\mathbf{d}x_1) \cdots \Lambda(\mathbf{d}x_n) \\ &= \int_{(\mathbb{R}^d \times \mathbb{M})^n} E[\varphi_{x_1, \dots, x_n}(N)] \Lambda(\mathbf{d}x_1) \cdots \Lambda(\mathbf{d}x_n) \\ &= E[\tilde{\psi}(N)], \end{aligned}$$

where we have used (11).

It remains to show (16). To this end, we write

$$\begin{aligned} \tilde{\chi}(N) &= \frac{n!}{\lambda^n} \sum_{\{X_1^*, \dots, X_n^*\} \subset N|_{S(N)}} \varphi^{X_1^*, \dots, X_n^*}(N) \\ &= \frac{N_{S(N)}!}{\lambda^n (N_{S(N)} - n)!} \binom{N_{S(N)}}{n}^{-1} \sum_{\{X_1^*, \dots, X_n^*\} \subset N|_{S(N)}} \varphi^{X_1^*, \dots, X_n^*}(N) \\ &= \frac{N_{S(N)}!}{\lambda^n (N_{S(N)} - n)!} E[\varphi^{X_1'', \dots, X_n''}(N) \mid N], \end{aligned}$$

where the latter equality follows from the invariance by permutations of  $\varphi^{x_1, \dots, x_n}(\mu)$  and the fact that  $\varphi^{x_1, \dots, x_n}(\mu) = 0$  if  $\varepsilon(x_1, \dots, x_n) = 0$ .

## 5. Applications

### 5.1. Stochastic geometry

Stabilizing functionals are widely used in stochastic geometry. This class of functionals was first introduced in [20] and further developed by Penrose and Yukich (see, for instance, [27] and [28]). Assumption (2) is closely related to assuming that  $\varphi$  is stabilizing. The main difference is that in (2) we require that  $S$  is a stopping set. Thus, stochastic geometry is a natural field of application of Theorems 1, 2, and 3. In the next two subsections we develop two examples of application in this field.

5.1.1. *Cluster in subcritical continuum percolation.* Let  $N = \sum_{n \geq 1} \delta_{(X_n, Z_n)}$  be an IMHPP on  $\mathbb{R}^d$  of intensity  $\lambda$  with marks in  $[0, r]$ ,  $r > 0$ . Consider the Boolean model

$$\Xi = \bigcup_{n \geq 1} B(X_n, Z_n).$$

Continuum percolation deals with the existence of an infinite connected component in  $\Xi$ . It is well known that there exists a critical value of  $\lambda$ , say  $\lambda_c > 0$ , such that if  $\lambda < \lambda_c$  a.s., there are not infinite connected components in  $\Xi$ , and if  $\lambda > \lambda_c$  a.s., there is a unique infinite connected component in  $\Xi$  (see [23] as a general reference on continuum percolation).

Define by  $W(N)$  the connected component (or cluster) of  $\Xi$  containing the origin (note that  $W(N)$  is possibly empty) and, with a little abuse of notation, define by  $N_{|W(N)}$  the restriction of the random measure  $N$  to the cluster  $W(N)$  (this is indeed an abuse of notation since usually throughout this paper  $\mu_{|B}$  denotes the set of points of  $\mu$  on  $B \times \mathbb{M}$ ). If  $\lambda < \lambda_c$  then  $W(N)$  is a.s. a compact set. We define  $\varphi(N) = \mathbf{1}(N_{|W(N)} \in A)$  for some measurable set  $A \subseteq \mathcal{N}$ . It is of general interest to analyze the function

$$f(\lambda) = E_\lambda[\varphi(N)], \quad \lambda \in (0, \lambda_c).$$

In this paragraph we provide a continuous analog of Russo’s formula for Poisson point fields; see [35]. More precisely, for  $\mu \in \mathcal{N}$ , define the sets of pivotal points of  $A$  by

$$\begin{aligned} \mathcal{P}^+(\mu) &= \{x \in (\mathbb{R}^d \times [0, r]) \setminus \text{supp}(\mu) : \varphi(\mu + \delta_x) = 1, \varphi(\mu) = 0\} \\ &\cup \{x \in \text{supp}(\mu) : \varphi(\mu) = 1, \varphi(\mu - \delta_x) = 0\} \end{aligned}$$

and

$$\begin{aligned} \mathcal{P}^-(\mu) &= \{x \in (\mathbb{R}^d \times [0, r]) \setminus \text{supp}(\mu) : \varphi(\mu + \delta_x) = 0, \varphi(\mu) = 1\} \\ &\cup \{x \in \text{supp}(\mu) : \varphi(\mu) = 0, \varphi(\mu - \delta_x) = 1\}. \end{aligned}$$

The following theorem holds.

**Theorem 4.** *The function  $f(\lambda) = P_\lambda(N_{|W(N)} \in A)$  is analytic on the interval  $(0, \lambda_c)$  and*

$$f'(\lambda) = E_\lambda[\Lambda(\mathcal{P}^+(N)) - \Lambda(\mathcal{P}^-(N))] = \frac{1}{\lambda} E_\lambda[N_{\mathcal{P}^+(N)} - N_{\mathcal{P}^-(N)}] \quad \text{for } 0 < \lambda < \lambda_c.$$

The proof of Theorem 4 is based on Theorems 1 and 2. The main difficulty in applying these theorems to  $f(\lambda)$  is that  $W$  is not a stopping set. This difficulty can be circumvented as follows. The Minkowski addition is defined by

$$A \oplus B = \{a + b : a \in A, b \in B\}, \quad A, B \in \mathbb{K}$$

(see [22] for a complete treatment of the Minkowski operations). Next, Lemma 6, below, provides a stopping set  $S$  which satisfies assumptions (2) and (4).

**Lemma 6.** *Define the random compact set  $S(N) = W(N) \oplus B_r$ . Then*

- (i)  *$S$  is a stopping set such that  $W(N)$  (and therefore  $\varphi(N)$ ) is  $\mathcal{F}_S$ -measurable;*
- (ii) *for each  $\lambda > 0$ , there exists  $s = s(\lambda) > 0$  such that  $E_\lambda[\exp(s|B_{S(N)}|)] < \infty$ .*

*Proof.* We first prove that  $S$  is a stopping set. First note that if  $X_k = (X_k, Z_k) \in \text{supp}(N_{|W(N)})$  then either  $X_k$  is a distance at most  $Z_k + Z_n$  from at least one other point  $X_n$  with  $X_n = (X_n, Z_n) \in \text{supp}(N_{|W(N)})$  or  $X_k$  is a distance at most  $Z_k$  from the origin. Note that

$$W(N) = \bigoplus_{X_n \in \text{supp}(N_{|W(N)})} (X_n + B_{Z_n})$$

and, therefore, by the definition of  $S(N)$  we have

$$S(N) = \left( \bigoplus_{X_n \in \text{supp}(N|_{W(N)})} (X_n + B_{Z_n}) \right) \oplus B_r.$$

Letting  $\|\cdot\|$  denote the Euclidean norm in  $\mathbb{R}^d$  and  $A^c$  denote the complement of the set  $A$ , we deduce that, for any  $K \in \mathbb{K}$ ,

$$\begin{aligned} \{S(N) \subseteq K\}^c &= \{\text{there exists } y \in K^c, X_1, \dots, X_n \in \text{supp}(N): \\ &\|X_1\| \leq Z_1, \|X_{i+1} - X_i\| \leq Z_i + Z_{i+1}, 1 \leq i \leq n - 1, \\ &\|X_n - y\| \leq Z_n + r\}. \end{aligned}$$

Now, set in the above expression  $m = \min\{k \in \{1, \dots, n + 1\}: X_k \in K^c\}$  with the convention  $X_{n+1} = y$ . Since, by assumption,  $Z_{i+1} \leq r$ ,  $\|X_{i+1} - X_i\| \leq Z_i + Z_{i+1}$  implies that  $\|X_{i+1} - X_i\| \leq Z_i + r$ ; hence, the event  $\{S(N) \subseteq K\}^c$  can be rewritten as

$$\begin{aligned} \{S(N) \subseteq K\}^c &= \{\text{there exists } X_1, \dots, X_m \in \text{supp}(N): X_1, \dots, X_m \in K \text{ and} \\ &\|X_1\| \leq Z_1, \|X_i - X_{i+1}\| \leq Z_i + Z_{i+1}, 1 \leq i \leq m - 1, \\ &(X_m + B_{Z_m+r}) \cap K^c \neq \emptyset\}. \end{aligned}$$

It follows that  $\{S(N) \subseteq K\}^c \in \mathcal{F}_K$  and, thus,  $S$  is a stopping set. Now, by construction,  $N|_{W(N)}$  is  $\mathcal{F}_S$ -measurable. We deduce that  $W(N)$  is also  $\mathcal{F}_S$ -measurable and part (i) is proved. It remains to prove part (ii). Define  $r_{W(N)} = \inf\{r \geq 0: W(N) \subseteq B_r\}$ . Then  $B_{S(N)} = B_{r_{W(N)}+r}$ . Thus, part (ii) is a consequence of the exponential decrease of the subcritical cluster; see Section 3.7 and Lemma 3.3 of [23].

*Proof of Theorem 4.* Clearly, the functional  $\varphi$  satisfies assumption (3). Moreover, by Lemma 6, assumptions (2) and (4) are satisfied with  $S(N) = W(N) \oplus B_r$ . Thus, by Theorem 1, the function  $f(\cdot)$  is analytic on  $(0, \lambda_c)$ . Note that  $S$  is nondecreasing; thus, using (8), it follows that

$$\begin{aligned} f'(\lambda) &= E_\lambda[|S(N)|(\mathbf{1}(X \in \mathcal{P}^+(N)) - \mathbf{1}(X \in \mathcal{P}^-(N)))] \\ &= E_\lambda[|S(N)| E_\lambda[\mathbf{1}(X \in \mathcal{P}^+(N)) - \mathbf{1}(X \in \mathcal{P}^-(N)) \mid N]] \\ &= E_\lambda \left[ |S(N)| \frac{1}{|S(N)|} \int_{S(N) \times [0,r]} (\mathbf{1}(x \in \mathcal{P}^+(N)) - \mathbf{1}(x \in \mathcal{P}^-(N))) \Lambda(dx) \right] \\ &= E_\lambda[\Lambda(\mathcal{P}^+(N)) - \Lambda(\mathcal{P}^-(N))], \end{aligned}$$

where  $X = (\xi, \zeta)$  is defined in Theorem 2. In particular, the first equality above follows since, given  $S(N)$ ,  $\xi$  is uniformly distributed on  $S(N)$  and, therefore,  $X \notin \text{supp}(N)$  a.s. This proves the first equality of the claim. The second equality of the claim can be proved similarly, using (9).

5.1.2. *Typical cell of Poisson–Voronoi tessellation.* Let  $A \subset \mathbb{R}^d$  be a locally finite point set and let  $\mu_A = \sum_{a \in A} \delta_a$ . The Voronoi cell with respect to  $A$  with nucleus  $y \in A$  is by definition

$$C(y, \mu_A) = \{x \in \mathbb{R}^d: \|x - y\| \leq \|x - a\| \text{ for all } a \in A\}.$$

Let  $N = \sum_{n \geq 1} \delta_{X_n}$  be a Poisson point process on  $\mathbb{R}^d$  of intensity  $\lambda$ . The Poisson–Voronoi cell with nucleus  $X_k$  is by definition the random convex set  $C(X_k, N)$  (see, for instance, [31]). Let

$\mathcal{K}_0^d$  be the set of convex bodies of  $\mathbb{R}^d$  containing the origin, equipped with the Hausdorff metric and the related Borel  $\sigma$ -field. Moreover, let  $N_0$  be the point process obtained by  $N$  adding a point at the origin. The typical Poisson–Voronoi cell is defined by  $C(0, N_0)$ . This cell is called typical since, by Slivnyak’s theorem (see, for instance, [31]),  $P_\lambda^0(C(0, N) \in A) = P_\lambda(C(0, N_0) \in A)$ , where  $P_\lambda^0$  is the Palm version of  $P_\lambda$  and  $A$  is a Borel set of  $\mathcal{K}_0^d$ .

Let  $\phi$  be a measurable functional from  $\mathcal{K}_0^d$  to  $\mathbb{R}$ . Define  $\varphi(N) = \phi(C(0, N_0))$  and

$$f(\lambda) = E_\lambda[\phi(C(0, N_0))] = E_\lambda[\varphi(N)].$$

The Voronoi flower  $V(N)$  is the union of the closed balls that have the origin and  $d$  points of  $N$  on their boundary, and no points of  $N$  inside. It is known that the centers of the balls which form  $V(N)$  are the vertices of the typical Poisson–Voronoi cell. Then  $V$  is a stopping set and  $\varphi(N) = \phi(C(0, N_0))$  is  $\mathcal{F}_V$ -measurable (see, for instance, [36]). It is also known that  $B_{V(N)}$  satisfies (4) for all  $\lambda > 0$ ; indeed, by Lemma 1 and Remark 5 of [13], it follows that, for all  $\lambda > 0$ ,

$$P_\lambda(|B_{V(N)}| > 2^{-d}t) < \exp(-c_d t) \quad \text{for each } t > 0,$$

for some positive constant  $c_d$  depending only on the dimension  $d$ . Furthermore, it can be easily realized that the mapping  $V$  is monotone nonincreasing. Hence, our results can be applied to  $f(\lambda)$  provided that  $E_\lambda[|\phi(C(0, N_0))|^\gamma] < \infty$  for some  $\gamma = \gamma(\lambda) > 2$ .

Note that this latter condition holds if  $\varphi(N) = \phi(C(0, N_0)) = \mathbf{1}(C(0, N_0) \in A)$  for some measurable set  $A \subset \mathcal{K}_0^d$ . In particular, in this case the following analog of Theorem 4 holds. For each  $\mu \in \mathcal{N}$  such that  $0 \in \text{supp}(\mu)$ , consider the following sets of pivotal points of the measurable set  $A \subset \mathcal{K}_0^d$ :

$$\begin{aligned} \mathcal{P}^+(\mu) &= \{x \in \mathbb{R}^d \setminus \text{supp}(\mu) : C(0, \mu + \delta_x) \in A, C(0, \mu) \notin A\} \\ &\cup \{x \in \text{supp}(\mu) : C(0, \mu) \in A, C(0, \mu - \delta_x) \notin A\} \end{aligned}$$

and

$$\begin{aligned} \mathcal{P}^-(\mu) &= \{x \in \mathbb{R}^d \setminus \text{supp}(\mu) : C(0, \mu + \delta_x) \notin A, C(0, \mu) \in A\} \\ &\cup \{x \in \text{supp}(\mu) : C(0, \mu) \notin A, C(0, \mu - \delta_x) \in A\}. \end{aligned}$$

The following theorem holds.

**Theorem 5.** *The function  $f(\lambda) = P_\lambda(C(0, N_0) \in A)$  is analytic on  $(0, \infty)$  and*

$$f'(\lambda) = E_\lambda[|\mathcal{P}^+(N)| - |\mathcal{P}^-(N)|] = \frac{1}{\lambda} E_\lambda[N_{\mathcal{P}^+(N)} - N_{\mathcal{P}^-(N)}], \quad \lambda > 0.$$

The proof is similar to that of Theorem 4 and therefore omitted.

### 5.2. Insurance

In this subsection we apply our results to risk processes described in terms of Poisson shot noise and compound Poisson processes. The former have been introduced in [17] and [18] to model delayed claims, the latter correspond to the classical Cramér–Lundberg model (see, for instance, [2]). The main results of this subsection are Theorems 6 and 7. Under suitable light-tailed conditions on the claims, Theorems 6 and 7 respectively provide closed-form formulae for the  $n$ th-order derivative of the ruin probability of risk processes with delayed (and undelayed) claims, and an efficient Monte Carlo estimator for the first-order derivative of the ruin probability

of the classical Cramér–Lundberg model. The estimator proposed in Section 5.2.2 is alternative to that of [3] (see Remark 2).

Now we briefly recall the notion of an asymptotically optimal estimator, which will be considered in this subsection. Let  $z(u)$  be a positive function such that  $z(u) \rightarrow 0$  as  $u \rightarrow \infty$ . To obtain an asymptotically efficient estimator of  $z(u)$ , we look for an unbiased estimator  $\hat{r}_u$  of  $z(u)$  whose relative error is asymptotically bounded. In the following we focus on a weaker concept of efficiency. We say that  $\hat{r}_u$  is asymptotically optimal (as  $u \rightarrow \infty$ ) if

$$\liminf_{u \rightarrow \infty} \frac{\log \sqrt{E[\hat{r}_u^2]}}{\log z(u)} \geq 1$$

(see [2] and [3]).

All the random variables considered in this subsection are defined on a measurable space  $(\Omega, \mathcal{F})$ . Here we consider marked point processes on  $[0, \infty)$  with marks in  $[0, \infty)$ . We endow  $(\Omega, \mathcal{F})$  with the family of probability measures  $\{P_\lambda\}_{\lambda > 0}$  such that, under  $P_\lambda$ ,  $X_1 < X_2 < \dots$  are the points of a homogeneous Poisson process on  $[0, \infty)$  with intensity  $\lambda > 0$ , and  $\{Z_n\}_{n \geq 1}$  are i.i.d. nonnegative random variables with distribution  $Q$ , and independent of the Poisson process. We denote by  $N$  the IMHPP  $\sum_{n \geq 1} \delta_{(X_n, Z_n)}$ , by  $N_t$  the number of points of  $N$  on  $[0, t] \times [0, \infty)$ , by  $N|_t$  the set of points of  $\text{supp}(N)$  on  $[0, t] \times [0, \infty)$ , and by  $E_\lambda$  the expectation with respect to  $P_\lambda$ .

**5.2.1. Derivatives of the ruin probability of risk processes with delayed claims.** Consider the following risk model. Let  $u - Y(t)$  be the surplus of the insurance portfolio described by the shot noise process with drift

$$Y(t) = \sum_{n \geq 1} H(t - X_n, Z_n) \mathbf{1}_{(0,t]}(X_n) - ct, \quad t \geq 0.$$

Here  $u > 0$  is the initial capital,  $c > 0$  is the premium density (which is assumed to be constant), and  $H: \mathbb{R} \times [0, \infty) \rightarrow [0, \infty)$  is a nondecreasing, continuous function such that  $H(t, z) = 0$  for  $t \leq 0$ . Throughout this subsection, we assume that  $H(\infty, z) = z$  and  $P_\lambda(Z_1 > 0) > 0$ . Since the law of  $Z_1$  under  $P_\lambda$  does not depend on  $\lambda$ , from now on, for a measurable function  $g$ , we set  $E_\lambda[g(Z_1)] = E[g(Z_1)]$ .

Note that the function  $H$  models the delay in claim settlement in the sense that the insurance company honors a claim at time  $X_n$  paying the quantity  $H(t - X_n, Z_n)$  at time  $t$ . The associated ruin probability is defined by the quantity

$$f_u(\lambda) = P_\lambda(T_u(N) < \infty), \quad u \geq 0,$$

where

$$T_u(N) = \inf\{t \geq 0: Y(t) > u\}, \quad T_u(N) = \infty \text{ if } \{\dots\} = \emptyset,$$

is the ruin time. Brémaud [7] proved that under the assumptions

$$\kappa(\theta) = E[\exp(\theta Z_1)] < \infty \text{ for all } \theta \text{ in a neighborhood of } 0, \text{ say } (0, \eta) \text{ with } \eta \leq \infty \quad (40)$$

and

$$c > \lambda E[Z_1], \quad (41)$$

it holds that

$$f_u(\lambda) \leq e^{-wu} \text{ for all } u \geq 0$$

and

$$\lim_{u \rightarrow \infty} \frac{1}{u} \log f_u(\lambda) = -w,$$

where  $w$  (called the Lundberg parameter) is the unique positive 0 of the function

$$\Lambda(\theta) = \lambda(\kappa(\theta) - 1) - c\theta.$$

(Note that the function  $\Lambda$  should not be confused with the intensity measure  $\Lambda$  considered in the previous sections. In the remaining part of the paper, the symbol  $\Lambda$  will not be used anymore to denote the intensity measure). Thus, under (40) and (41), the event  $\{T_u(N) < \infty\}$  is rare as  $u \rightarrow \infty$  and this yields problems if we want to estimate  $f_u(\lambda)$  by an efficient Monte Carlo simulation (we refer the reader to [9] for an introduction to rare event simulations). Such difficulties can be overcome using importance sampling. Define the stochastic process

$$C(t) = \sum_{n \geq 1} Z_n \mathbf{1}_{(0,t]}(X_n),$$

and consider the family of laws  $\{P_\lambda^\theta\}_{\theta: \kappa(\theta) < \infty}$  defined as follows: the probability measure  $P_\lambda^\theta$  is absolutely continuous with respect to the original law  $P_\lambda$  on the  $\sigma$ -field  $\mathcal{F}_{[0,t]}$  for each  $t \geq 0$  and the corresponding density is

$$\ell_t^{P_\lambda^\theta, P_\lambda} = \frac{e^{\theta C(t)}}{E_\lambda[e^{\theta C(t)}]} = \exp(\theta C(t) - \lambda t(\kappa(\theta) - 1)).$$

We point out (see, for instance, [2]) that, under  $P_\lambda^\theta$ , the process  $\{X_n\}_{n \geq 1}$  is a homogeneous Poisson process with intensity  $\lambda\kappa(\theta)$ , independent of the sequence  $\{Z_n\}_{n \geq 1}$  of i.i.d. random variables, whose common law  $Q^\theta$  is absolutely continuous with respect to their common law  $Q$  under  $P_\lambda$ , with density

$$\frac{dQ^\theta}{dQ}(z) = \frac{e^{\theta z}}{\kappa(\theta)}.$$

Throughout this subsection, we denote by  $E_\lambda^\theta$  the expectation under  $P_\lambda^\theta$ . Furthermore, since the law of  $Z_1$  under  $P_\lambda^\theta$  does not depend on  $\lambda$ , for a measurable function  $g$ , we set  $E_\lambda^\theta[g(Z_1)] = E^\theta[g(Z_1)]$ .

The following result can be found in [33].

**Proposition 5.** ([33].) *Assume that (40) and (41) hold. Then  $P_\lambda^w(T_u(N) < \infty) = 1$  for all  $u > 0$  and, under  $P_\lambda^w$ ,*

$$\hat{r}_u(N) = \ell_{T_u(N)}^{P_\lambda, P_\lambda^w}$$

*is an asymptotically optimal estimator of  $f_u(\lambda)$ .*

Let  $\mathbf{x}_n = (x_n, z_n) \in (0, \infty) \times (0, \infty)$ ,  $n \geq 1$ . For each locally finite counting measure  $\mu = \sum_{n \geq 1} \delta_{\mathbf{x}_n}$ , define the functionals

$$\varphi_\theta(\mu, \lambda) = \exp\left(-\theta \sum_{n \geq 1} z_n \mathbf{1}_{(0, T_u(\mu)]}(x_n) + \lambda(\kappa(\theta) - 1)T_u(\mu)\right), \quad 0 \leq \theta < w,$$

and

$$\varphi_w(\mu) = \exp\left(-w \left(\sum_{n \geq 1} z_n \mathbf{1}_{(0, T_u(\mu)]}(x_n) - cT_u(\mu)\right)\right).$$

Moreover, we consider the functionals  $\varphi_{w, x_1, \dots, x_n}(\mu)$  and  $\varphi_w^{x_1, \dots, x_n}(\mu)$ , which are respectively defined by (10) and (11) with  $\varphi_w$  in place of  $\varphi$ .

The following theorem provides closed-form expressions for the  $n$ th-order derivatives of the ruin probability. As usual, we use the standard convention that the sum over an empty set is 0 and that  $k!/(k - n)! = 0$  for  $n > k$ .

**Theorem 6.** *Under the assumptions of Proposition 5, we find that, for a fixed  $u > 0$ , the function  $f_u(\cdot)$  is analytic in a neighborhood of  $\lambda$  and that, for all  $n \geq 1$ ,*

$$\begin{aligned} f_u^{(n)}(\lambda) - (\kappa(w) - 1)^n E_\lambda^w [(T_u(N))^n \varphi_w(N)] &= E_\lambda^w [(T_u(N))^n \varphi_{w, X_1, \dots, X_n}(N)] \\ &= E_\lambda^w \left[ \left( \frac{N_{T_u(N)}}{\lambda} \right)^n \varphi_w^{X'_1, \dots, X'_n}(N) \right] \\ &= E_\lambda^w \left[ \frac{N_{T_u(N)}!}{(N_{T_u(N)} - n)! \lambda^n} \varphi_w^{X''_1, \dots, X''_n}(N) \right], \end{aligned}$$

where, for  $1 \leq i \leq n$ ,  $X_i = (\xi_i, \zeta_i)$ ,  $X'_i = (\xi'_i, \zeta'_i)$ , and  $X''_i = (\xi''_i, \zeta''_i)$  are random variables on  $(0, \infty) \times (0, \infty)$ . Given  $T_u(N)$ ,  $(\xi_i)_{1 \leq i \leq n}$  are independent and uniformly distributed on  $[0, T_u(N)]$ , and independent of  $N$ ;  $(\zeta_i)_{1 \leq i \leq n}$  are independent, independent of  $N$  and  $(\xi_i)_{1 \leq i \leq n}$ , and have law  $Q^w$ . Given the collection of points  $N_{|T_u(N)}$ ,  $(X'_i)_{1 \leq i \leq n}$  are independent and uniformly distributed on the collection;  $\{X''_1, \dots, X''_n\}$  is uniformly distributed on the set of subsets of  $n$  distinct points of  $N_{|T_u(N)}$ .

To prove Theorem 6, we need Lemmas 7 and 8, below. Here we consider the notion of a large deviation principle for which we refer the reader to [12].

**Lemma 7.** *Assume that (40) holds. If, moreover, the function  $\theta \mapsto \kappa(\theta)$ ,  $\theta \in (0, \eta)$ , is steep, namely  $\lim_{n \rightarrow \infty} \kappa'(\theta_n) = \infty$  whenever  $\{\theta_n\}$  is a sequence converging to  $\eta$ , we find that the stochastic process  $\{Y(t)/t\}_{t>0}$  satisfies a large deviation principle with rate function  $\Lambda^*(x) = \sup_{\theta \in \mathbb{R}} (\theta x - \Lambda(\theta))$ .*

The proof of Lemma 7 can be found in [21] (see Proposition 3.1 therein).

**Lemma 8.** *Under the assumptions of Proposition 5, we find that, for a fixed  $u > 0$  and  $\bar{\theta} \in (0, w]$  such that  $\lambda \kappa'(\bar{\theta}) - c > 0$ , there exists  $s = s(\lambda) > 0$  such that  $E_\lambda^{\bar{\theta}}[\exp(sT_u(N))] < \infty$ .*

*Proof.* In this proof we write  $T_u$  in place of  $T_u(N)$ . Since

$$\kappa_{\bar{\theta}}(\alpha) = E^{\bar{\theta}}[\exp(\alpha Z_1)] = \frac{\kappa(\alpha + \bar{\theta})}{\kappa(\bar{\theta})},$$

by the assumptions, it follows that  $\kappa_{\bar{\theta}}(\alpha) < \infty$  for  $\alpha \in (0, \eta - \bar{\theta})$  and that the function  $\alpha \mapsto \kappa_{\bar{\theta}}(\alpha)$  is steep. Therefore, by Lemma 7, the stochastic process  $\{Y(t)/t\}_{t>0}$  satisfies a large deviation principle with respect to  $P_\lambda^{\bar{\theta}}$  with rate function  $\Lambda_{\bar{\theta}}^*(x) = \sup_{\theta \in \mathbb{R}} (\theta x - \Lambda_{\bar{\theta}}(\theta))$ , where  $\Lambda_{\bar{\theta}}(\theta) = \lambda \kappa(\bar{\theta})(\kappa_{\bar{\theta}}(\theta) - 1) - c\theta$ . Since  $\lambda \kappa'(\bar{\theta}) - c > 0$ , by assumption we can choose  $\beta \in (0, \kappa'(\bar{\theta}))$  such that  $\gamma = \lambda\beta - c > 0$ . By the large deviation principle of  $\{Y(t)/t\}_{t>0}$  with respect to  $P_\lambda^{\bar{\theta}}$  and the regularity properties of the rate function  $\Lambda_{\bar{\theta}}^*(\cdot)$ , we have

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log P_\lambda^{\bar{\theta}} \left( \frac{Y(t)}{t} < \gamma \right) = -\Lambda_{\bar{\theta}}^*(\gamma). \tag{42}$$

Moreover, for any  $u > 0$ , we find that there exists  $\bar{t}_1 = \bar{t}_1(u, \gamma)$  such that

$$P_{\lambda}^{\bar{\theta}}(T_u > t) \leq P_{\lambda}^{\bar{\theta}}(Y(t) < u) \leq P_{\lambda}^{\bar{\theta}}\left(\frac{Y(t)}{t} < \gamma\right) \quad \text{for all } t \geq \bar{t}_1. \tag{43}$$

Therefore, by (42) and (43), it follows that, for any  $\varepsilon, u > 0$ , there exists  $\bar{t} = \bar{t}(\varepsilon, u, \gamma)$  such that

$$P_{\lambda}^{\bar{\theta}}(T_u > t) < \exp(-(\Lambda_{\bar{\theta}}^*(\gamma) - \varepsilon)t) \quad \text{for all } t \geq \bar{t}. \tag{44}$$

Now, take  $0 < s < \Lambda_{\bar{\theta}}^*(\gamma) - \varepsilon$ . The conclusion follows noticing that by (44) we have

$$\begin{aligned} E_{\lambda}^{\bar{\theta}}[\exp(sT_u)] &= 1 + s \int_0^{\infty} e^{st} P_{\lambda}^{\bar{\theta}}(T_u > t) dt \\ &\leq e^{s\bar{t}} + s \int_{\bar{t}}^{\infty} \exp(-((\Lambda_{\bar{\theta}}^*(\gamma) - \varepsilon) - s)t) dt \\ &< \infty. \end{aligned}$$

*Proof of Theorem 6.* We start by noting that, by the properties of the function  $\theta \mapsto \Lambda(\theta)$ ,  $\theta \in (0, \eta)$ , there exists a strictly increasing sequence  $\{\theta_k\}$  converging to  $w$  such that  $\theta_k \in (0, w)$  and  $\lambda\kappa'(\theta_k) - c > 0$ . By the implicit function theorem, the function  $\lambda \mapsto w(\lambda)$  is continuous. Therefore, for each  $k$ , there exists a neighborhood of  $\lambda$ , say  $I_k = (\lambda - \varepsilon_k, \lambda + \varepsilon_k)$ , such that, for all  $\lambda' \in I_k$ , we have  $\theta_k < w(\lambda')$  and  $\lambda'\kappa'(\theta_k) - c > 0$ . We note that, since  $\theta_k \in (0, w)$  is such that  $\lambda\kappa'(\theta_k) - c > 0$ , it holds that  $P_{\lambda}^{\theta_k}(T_u(N) < \infty) = 1$  (see Lemma 3.2 of [33] for details). Therefore,

$$f_u(\lambda) = E_{\lambda}[\mathbf{1}(T_u(N) < \infty)] = E_{\lambda}^{\theta_k}[\varphi_{\theta_k}(N, \lambda)].$$

Note that  $[0, T_u]$  is a stopping set. Furthermore, the functional  $\varphi_{\theta_k}(N, \lambda)$  is  $\mathcal{F}_{[0, T_u]}$ -measurable and absolutely monotonic in  $\lambda$  (see Remark 1 for the definition of an absolutely monotonic function). Also, note that  $\Lambda(\theta) \leq 0$  for each  $\theta \in [0, w]$ . Therefore, by the definition of  $T_u(N)$  and the assumption  $H(t, z) \nearrow z$  as  $t \nearrow \infty$ , we find that, for each  $u$ ,

$$\varphi_{\theta}(N, \lambda) \leq e^{-\theta u} \quad \text{and} \quad \varphi_w(N) \leq e^{-wu}.$$

In particular, this implies that the functional  $\varphi_{\theta_k}(N, \lambda)$  is bounded. Therefore, by Lemma 8 and Remark 1, it follows that  $f_u(\cdot)$  is analytic on  $I_k$ . Consider the functions

$$F_k(x, y) = E_x^{\theta_k}[\varphi_{\theta_k}(N, y)], \quad (x, y) \in I_k \times I_k.$$

Using obvious notation, we will show later that

$$\partial_x^n F_k(x, y) = E_x^{\theta_k}[(T_u(N))^n \varphi_{\theta_k, X_1, \dots, X_n}(N, y)], \quad n \geq 1, \tag{45}$$

and

$$\partial_y^n F_k(x, y) = (\kappa(\theta_k) - 1)^n E_x^{\theta_k}[(T_u(N))^n \varphi_{\theta_k}(N, y)], \quad n \geq 1. \tag{46}$$

Therefore,

$$\begin{aligned} f_u^{(n)}(\lambda) - (\kappa(\theta_k) - 1)^n E_{\lambda}^{\theta_k}[(T_u(N))^n \varphi_{\theta_k}(N, \lambda)] \\ = E_{\lambda}^{\theta_k}[(T_u(N))^n \varphi_{\theta_k, X_1, \dots, X_n}(N, \lambda)], \quad n \geq 1. \end{aligned}$$



We now show that

$$\lim_{k \rightarrow \infty} E_{\lambda}^{\theta_k} [(T_u(N))^n \varphi_{\theta_k, X_1, \dots, X_n}(N, \lambda)] = E_{\lambda}^w [(T_u(N))^n \varphi_{w, X_1, \dots, X_n}(N)]. \tag{47}$$

Using the exponential tilting, we have

$$\begin{aligned} & |E_{\lambda}^{\theta_k} [(T_u(N))^n \varphi_{\theta_k, X_1, \dots, X_n}(N, \lambda)] - E_{\lambda}^w [(T_u(N))^n \varphi_{w, X_1, \dots, X_n}(N)]| \\ & \leq E_{\lambda}^w [(T_u(N))^n |\exp(-(w - \theta_k)(C(T_u(N)) - cT_u(N)) - \Lambda(\theta_k)T_u(N)) \\ & \quad \times \varphi_{\theta_k, X_1, \dots, X_n}(N, \lambda) - \varphi_{w, X_1, \dots, X_n}(N)|]. \end{aligned} \tag{48}$$

Note that the argument of the mean in (48) is less than or equal to  $(2T_u(N))^n (\exp(-\Lambda(\theta_k) \times T_u(N)) + 1)$ . By Lemma 8, there exists  $s = s(\lambda) > 0$  such that  $E_{\lambda}^w [\exp(sT_u(N))] < \infty$ . Fix  $\varepsilon \in (0, s)$  and choose  $\bar{k} = \bar{k}(\varepsilon)$  such that, for all  $k > \bar{k}$ , it holds that  $0 < -\Lambda(\theta_k) < \varepsilon < s$  (such a  $\bar{k}$  exists since  $\lim_{k \rightarrow \infty} \Lambda(\theta_k) = \Lambda(w) = 0$ ). Then, for all  $k > \bar{k}$ , the argument of the mean in (48) is less than or equal to

$$(2T_u(N))^n (\exp(\varepsilon T_u(N)) + 1),$$

which is integrable under  $P_{\lambda}^w$  since  $E_{\lambda}^w [\exp(sT_u(N))] < \infty$ . Thus, (47) follows by the dominated convergence theorem. The limit

$$\lim_{k \rightarrow \infty} E_{\lambda}^{\theta_k} [(T_u(N))^n \varphi_{\theta_k}(N, \lambda)] = E_{\lambda}^w [(T_u(N))^n \varphi_w(N)]$$

can be proved similarly. This shows the first equality in the statement. The other equalities can be proved as in Proposition 4 and Theorem 3.

It remains to show that  $F_k(\cdot, \cdot)$  has partial derivatives (45) and (46). Equality (45) follows by Proposition 2. Indeed, by Lemma 8, for each  $x \in I_k$ , there exists  $s = s(x) > 0$  such that  $E_x^{\theta_k} [\exp(sT_u(N))] < \infty$ , and the mapping  $S_u := [0, T_u]$  is nonincreasing. We now show (46) with  $n = 1$ . The general case follows along similar lines, reasoning by induction. If we justify the interchange between the sign of limit and the sign of mean in the expression,

$$\lim_{h \rightarrow 0^+} \frac{E_x^{\theta_k} [\varphi_{\theta_k}(N, y + h) - \varphi_{\theta_k}(N, y)]}{h},$$

then the right-hand derivative equals the right-hand side of (46). In fact, we can pass the limit into the sign of the expectation in that a straightforward computation gives

$$\left| \frac{\varphi_{\theta_k}(N, y + h) - \varphi_{\theta_k}(N, y)}{h} \right| \leq (\kappa(\theta_k) - 1)T_u(N) \exp(h(\kappa(\theta_k) - 1)T_u(N)).$$

Here, again by Lemma 8, the right-hand side of the above inequality is integrable under  $P_x^{\theta_k}$  and, therefore, we can apply the dominated convergence theorem. Similarly, we can show that the left-hand derivative equals the right-hand side of (46). This completes the proof.

5.2.2. *Classical risk processes: an efficient Monte Carlo algorithm for the first-order derivative of the ruin probability.* The classical risk model is defined by the surplus  $u - Y(t)$  of the insurance portfolio described by the following compound process with drift:

$$Y(t) = \sum_{n \geq 1} Z_n \mathbf{1}_{(0,t]}(X_n) - ct.$$

The interpretation of the quantities in the above formula is exactly as in the previous subsection. Moreover, we consider the same statistical assumptions and the same notation (clearly, the ruin probability  $f_u(\lambda)$  and the ruin time  $T_u(N)$  are now defined with respect to the classical risk process). For the Cramér–Lundberg model, it is well known that, under the assumptions (40) and (41), we have

$$\lim_{u \rightarrow \infty} \frac{e^{wu}}{u} f'_u(\lambda) = \frac{c^2 w(c - \lambda E[Z_1])}{\lambda(\lambda\kappa'(w) - c)^2}, \tag{49}$$

where  $w$  is the unique positive 0 of the function  $\Lambda(\cdot)$  (see Proposition 9.4 of [2]). Moreover, note that in the case of classical risk processes the corresponding Theorem 6 can be proved along similar lines, and provides the following unbiased estimator of  $f'_u(\lambda)$  under  $P_\lambda^w$ :

$$\hat{s}_u(N) = (\kappa(w) - 1)T_u(N)\varphi_w(N) + T_u(N)(\varphi_w(N + \delta_X) - \varphi_w(N)).$$

Here,  $X = (\xi, \zeta)$  is a random variable on  $(0, \infty) \times (0, \infty)$ ; given  $T_u(N)$ ,  $\xi$  is uniformly distributed on  $[0, T_u(N)]$ , and independent of  $N$ ;  $\zeta$  is independent of  $N$  and  $\xi$ , and has law  $Q^w$ .

The following theorem holds.

**Theorem 7.** *Assume that (40) and (41) hold. Then  $\hat{s}_u(N)$  is an asymptotically optimal estimator of  $f'_u(\lambda)$  as  $u \rightarrow \infty$  under the law  $P_\lambda^w$ .*

*Proof.* We only need to prove that

$$\liminf_{u \rightarrow \infty} \frac{\log \sqrt{E_\lambda^w[(\hat{s}_u(N))^2]}}{\log f'_u(\lambda)} \geq 1. \tag{50}$$

For any (locally finite) counting measure  $\mu$  on  $(0, \infty) \times (0, \infty)$  and  $u > 0$ , we find that  $\varphi_w(\mu) \leq e^{-wu}$ ; thus

$$|\hat{s}_u(N)| \leq (\kappa(w) + 1)e^{-wu}T_u(N)$$

and, therefore,

$$E_\lambda^w[(\hat{s}_u(N))^2] \leq (\kappa(w) + 1)^2 e^{-2wu} E_\lambda^w[T_u(N)^2]. \tag{51}$$

Denote by  $\{\tilde{X}_i\}$  the interarrivals of the Poisson process  $\{X_i\}$ . Under  $P_\lambda^w$ ,  $\sum_{i=1}^n (Z_i - c\tilde{X}_i)$  is a random walk with positive drift; indeed

$$E_\lambda^w[Z_1 - c\tilde{X}_1] = \frac{\lambda\kappa'(w) - c}{\lambda\kappa(w)} > 0.$$

Furthermore,  $T_u(N)$  is the hitting time of this random walk. Therefore, by the results in [15], it follows that

$$E_\lambda^w[T_u(N)^2] = O(u^2) \quad \text{as } u \rightarrow \infty. \tag{52}$$

Finally, (50) follows by (49), (51), and (52).

While it is tempting to conjecture that a similar optimality result holds for risk processes with delay in claim settlement, we do not have a proof of this claim.

**Remark 2.** Note that, under the assumptions of Theorem 7, Asmussen and Rubinstein [3] (see also [2]) proved that

$$\hat{\sigma}_u(N) = \left( \frac{N_{T_u(N)}}{\lambda} - T_u(N) \right) \exp(-w(C(T_u(N)) - cT_u(N)))$$

is asymptotically optimal for  $f'_u(\lambda)$  under  $P_\lambda^w$ . The estimator  $\hat{s}_u(N)$  is alternative to  $\hat{\sigma}_u(N)$ .

## Acknowledgements

The authors thank the Editor, two anonymous referees, and Nicolas Privault for a careful reading of the paper and many valuable remarks.

## References

- [1] ALBEVERIO, S., KONDRATIEV, Y. G. AND RÖCKNER, M. (1996). Differential geometry of Poisson spaces. *C. R. Acad. Sci. Paris Sér. I* **323**, 1129–1134.
- [2] ASMUSSEN, S. (2000). *Ruin Probabilities*. World Scientific, Singapore.
- [3] ASMUSSEN, S. AND RUBINSTEIN, R. Y. (1999). Sensitivity analysis of insurance risk models. *Manag. Sci.* **45**, 1125–1141.
- [4] BACCELLI, F. AND BRÉMAUD, P. (1993). Virtual customers in sensitivity and light traffic analysis via Campbell's formula for point processes. *Adv. Appl. Prob.* **25**, 221–234.
- [5] BACCELLI, F., HASENFUSS, S. AND SCHMIDT, V. (1999). Differentiability of functionals of Poisson processes via coupling with applications to queueing theory. *Stoch. Process. Appl.* **81**, 299–321.
- [6] BLASZCZYSHYN, B. (1995). Factorial moment expansion for stochastic systems. *Stoch. Process. Appl.* **56**, 321–335.
- [7] BRÉMAUD, P. (2000). An insensitivity property of Lundberg's estimate for delayed claims. *J. Appl. Prob.* **37**, 914–917.
- [8] BRÉMAUD, P. AND VAZQUEZ-ABAD, F. J. (1992). On the pathwise computation of derivatives with respect to the rate of a point process: the phantom RPA method. *Queueing Systems Theory Appl.* **10**, 249–270.
- [9] BUCKLEW, J. A. (2004). *Introduction to Rare Event Simulation*. Springer, New York.
- [10] DALEY, D. J. AND VERE-JONES, D. (2003). *An Introduction to the Theory of Point Processes*. Springer, New York.
- [11] DECREUSEFOND, L. (1998). Perturbation analysis and Malliavin calculus. *Ann. Appl. Prob.* **8**, 496–523.
- [12] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*. Springer, New York.
- [13] FOSS, S. AND ZUYEV, S. (1996). On a Voronoi aggregative process related to a bivariate Poisson process. *Adv. Appl. Prob.* **28**, 965–981.
- [14] GLASSERMANN, P. (1990). *Gradient Estimation via Perturbation Analysis*. Kluwer, Dordrecht.
- [15] GUT, A. (1974). On the moments and limit distributions of some first passage times. *Ann. Prob.* **2**, 277–308.
- [16] HO, Y. C. AND CAO, X. R. (1983). Perturbation analysis and optimization of queueing networks. *J. Optimization Theory Appl.* **40**, 559–582.
- [17] KLÜPPELBERG, C. AND MIKOSCH, T. (1995). Delay in claim settlement and ruin probability approximations. *Scand. Actuarial J.* **2**, 154–168.
- [18] KLÜPPELBERG, C. AND MIKOSCH, T. (1995). Explosive Poisson shot noise processes with applications to risk reserves. *Bernoulli* **1**, 125–147.
- [19] L'ECUYER, P. (1990). A unified version of the IPA, SF, and LR gradient estimation techniques. *Manag. Sci.* **36**, 1364–1383.
- [20] LEE, S. (1997). The central limit theorem for Euclidean minimal spanning trees. *Ann. Appl. Prob.* **7**, 996–1020.
- [21] MACCI, C., STABILE, G. AND TORRISI, G. L. (2005). Lundberg parameters for non standard risk processes. *Scand. Actuarial J.* **6**, 417–432.
- [22] MATHERON, G. (1975). *Random Sets and Integral Geometry*. John Wiley, New York.
- [23] MEESTER, R. AND ROY, R. (1996). *Continuum Percolation*. Cambridge University Press.
- [24] MOLCHANOV, I. AND ZUYEV, S. (2000). Variational analysis of functionals of Poisson processes. *Math. Operat. Res.* **25**, 485–508.
- [25] MØLLER, J. AND WAAGEPETERSEN, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton, FL.
- [26] PENROSE, M. (2003). *Random Geometric Graphs*. Oxford University Press.
- [27] PENROSE, M. AND YUKICH, J. (2001). Limit theory for random sequential packing and deposition. *Ann. Appl. Prob.* **12**, 272–301.
- [28] PENROSE, M. AND YUKICH, J. (2003). Weak laws of large numbers in geometric probability. *Ann. Appl. Prob.* **13**, 277–303.
- [29] REIMAN, M. I. AND SIMON, B. (1989). Open queueing systems in light traffic. *Math. Operat. Res.* **14**, 26–59.
- [30] REIMAN, M. I. AND WEISS, A. (1989). Sensitivity analysis for simulations via likelihood ratios. *Operat. Res.* **37**, 830–844.
- [31] STOYAN, D., KENDALL, W. S. AND MECKE, J. (1995). *Stochastic Geometry and Its Applications*. John Wiley, Chichester.
- [32] SURI, R. AND ZAZANIS, F. M. (1988). Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queues. *Manag. Sci.* **34**, 39–64.

- [33] TORRISI, G. L. (2004). Simulating the ruin probability of risk processes with delay in claim settlement. *Stoch. Process. Appl.* **112**, 225–244.
- [34] ZAZANIS, F. M. (1992). Analyticity of Poisson-driven stochastic systems. *Adv. Appl. Prob.* **24**, 532–541.
- [35] ZUYEV, S. (1993). Russo's formula for Poisson point fields and its applications. *Discrete Math. Appl.* **3**, 63–73.
- [36] ZUYEV, S. (1999). Stopping-sets: gamma-type results and hitting properties. *Adv. Appl. Prob.* **31**, 355–366.