The Annals of Applied Probability
1998, Vol. 8, No. 2, 496–523

PERTURBATION ANALYSIS AND MALLIAVIN CALCULUS

By L. Decreusefond

École Nationale Supérieure des Télécommunications

Using the Malliavin calculus, we give a unified treatment of the so-called perturbation analysis of dynamic systems. Several applications are also given.

1. Introduction. Given a marked point process (MPP for short) N whose law depends on a parameter $\theta \in \Theta \subset \mathbf{R}$ and a functional F of the sample paths of N, the so-called perturbation (or sensitivity) analysis is concerned with the evaluation of the derivative with respect to θ of $\mathbf{E}_\theta\,[F(N)]$, where $\mathbf{E}_\theta$ is the expectation under $\mathbf{P}_\theta$, the law of N for the value θ of the parameter. In other words, the objective is to compute (or at least to find some estimates of) the sensitivity of the mean value of F with respect to a slight change in the law of N. The simplest but generic example follows.

Example 1. Given a standard Poisson process N of intensity θ and a functional of it, say $F = N_t$ for t fixed, how can we compute $d/d\theta\ E_\theta[N_t]$? The result is straightforward here since we know that $\mathbf{E}_\theta[N_t] = \theta t$, but what happens for a more complex functional?

There are several motivations for being interested in such a question: the main reasons are the applications to optimization and control of systems; see, for instance, Devetsikiotis, Wael, Freebersyser and Townsend (1993). The concept of perturbation analysis was introduced in a paper by Ho and Cao (1983) and has been addressed by many authors [Glasserman (1990); Glynn (1987); Heidelberger (1987); Ho, Cao and Cassandros (1983); Reiman and Weiss (1989a, b); Suri and Zazanis (1988); Suri (1989)], mainly in the context of queuing networks. There are essentially three ways to handle this problem: the so-called infinitesimal perturbation analysis (IPA), rare perturbation analysis (RPA) and likelihood ratio method (LRM). Our motivation here is not to discuss these methods in detail, but to show how they can be seen as a part of the stochastic calculus of variations. This theory, initiated by Malliavin (1978) in the context of the Brownian motion, aims to define a differential calculus for stochastic processes mimicking the differential calculus of usual numerical functions; see, for example, Üstünel (1995) and references therein. Besides the aesthetics of this new point of view, the known results of the

Received October 1995; revised January 1997.
AMS 1991 *subject classifications*. Primary 60H07, 60H30; secondary 60G55.
Key words and phrases. Chaos decomposition, IPA, light traffic, LRM, Malliavin calculus, RPA, sensitivity analysis, simulation.

Malliavin calculus also allow us to obtain somewhat deeper results in perturbation analysis.

Section 2 contains a brief description of IPA and RPA and a rather detailed description of the LR method. Actually, this latter approach plays a key role to exhibit relationships between the stochastic calculus of variations and the sensitivity analysis. Section 3 and 4 are devoted to the Malliavin calculus for marked point processes and to its applications to perturbation analysis. Precisely, in Section 3, we define a Malliavin derivative by a variational approach and in Section 4, we define a difference operator using the chaos decomposition of some random measures. Both the Malliavin derivative and the difference operator share the so-called formula of *stochastic integration by parts*, which is central to our work. In Section 5, we mention two results which can be connected to the theory developed in this paper.

2. Methods of the perturbation analysis. Let E be a Lusin space (for practical purposes $E = \mathbf{R}^d$ is sufficient) and let Ω be the space of simple (i.e., there is at most one jump at a time), locally finite (i.e., the number of jumps in each compact time interval is almost surely finite), integer-valued measures on $[0, T] \times E$, where T can be a fixed deterministic time or $T = +\infty$; for details on marked point processes, see, for instance, Jacod (1979). A generic sample path $\omega \in \Omega$ is thus of the form $\sum_{n>0} \delta_{(t_n, z_n)}$, where $\{t_n, n > 0\}$ is a strictly increasing sequence of nonnegative reals (t_n represents the nth jump time) and z_n belongs to E for any n (z_n is the mark associated to the nth jump). θ_0 is fixed and $\mathbf{P}_0$ is called the nominal probability. $\{\mathscr{F}_t, t > 0\}$ is the canonical filtration,

$$\mathscr{F}_0 = \{\emptyset, \Omega\}, \qquad \mathscr{F}_t = \sigma\left\{\int_0^t \int_B \omega(ds, dz),\, s \le t,\, B \in \mathscr{B}(E)\right\}$$

and $\mathscr{P}$ is the predictable σ-field on $\Omega \times [0, T] \times E$. We recall that $A \in \Omega \times [0, T]$ is said to be evanescent whenever its projection $\pi(A) = \{\omega; \exists t \in [0, T], (w, t) \in A\}$ is $\mathbf{P}_0$ negligible. For any $\theta \in \Theta$, we denote by ν_θ the $\mathbf{P}_\theta$ compensating measure of the canonical process, that is, the predictable random measure such that for any nonnegative and $\mathscr{P}$-measurable process Y, the process

$$\int_0^t \int_E Y(\omega, s, z)(\omega - \nu_\theta)(ds, dz)$$

is a $\mathbf{P}_\theta$ local martingale. For technical reasons (see Remark 2.1), we hereafter assume the following hypothesis.

HYPOTHESIS 1. *The process*

$$N_t \stackrel{\text{def}}{=} (\omega, t) \mapsto \int_0^t \int_E \omega(ds, dz)$$

is square-integrable,

$$\sup_{t \le T} \mathbf{E}_0\big[N_t^2\big] < +\infty,$$

and quasi-left-continuous; *for any given time* t, $\nu_{\theta_0}(\{t\} \times E) = 0$.

Going back to the roots of differential calculus, computing a derivative boils down to computing

$$\lim_{\theta \to \theta_0} (\theta - \theta_0)^{-1}\big(\mathbf{E}_\theta[F] - \mathbf{E}_{\theta_0}[F]\big). \tag{1}$$

Both IPA and RPA are based on suitable alterations of the nominal sample path in order to obtain a modified process of law $\mathbf{P}_\theta$. The second step consists of exactly expressing the right-hand-side difference in (1) and then being able to pass to the limit. Let us illustrate these two methods by some examples.

EXAMPLE 1 (Continued). Despite its simplicity, let us have another look at the Poisson process. The rare perturbation analysis [see Brémaud (1992); Brémaud and Vázquez-Abad (1992)] originates from the remark that we can obtain a Poisson process of intensity as close as we want to θ_0 by "decreasing" thinning. Namely, the derivative of the expectation of a functional F with respect to the mean intensity θ of the underlying process at $\theta = \theta_0$ is obtained by considering

$$\theta_0^{-1} \lim_{p \to 1} \frac{1}{1-p} \mathbf{E}_{\theta_0, p}\big[F(\omega) - F(\omega_p)\big],$$

where ω_p is a p-thinning of ω: each jump of ω is kept, independently from the others, with probability p and $\mathbf{E}_{\theta_0, p}$ is the expectation taken under the probability measure $d\mathbf{P}_0 \otimes (p\delta_1 + (1-p)\delta_0)^{\otimes \mathbf{N}}$. If we apply this to $F = N_t$, it is intuitively clear and easy to show that given $N_t(\omega)$, $N_t(\omega) - N_t(\omega_p)$ is distributed as a binomial law of parameters $N_t(\omega)$ and $\theta_0(1-p)$, so that we clearly have

$$\left(\frac{d}{d\theta} \mathbf{E}_\theta[N_t]\right)_{\theta = \theta_0} = t.$$

We see that the principle itself induces that RPA is essentially meaningful for the Poisson process and for the so-called light traffic analysis—the part of perturbation analysis which is dedicated to the analysis of the sensitivity of F when the mean intensity of the underlying process goes to 0.

EXAMPLE 2. Consider a G/GI/1 queue with mean service time θ and distribution function G_θ. Let F be the average waiting time for the first K customers, that is, $F = K^{-1}\sum_{i=1}^K W_i$, whose mean value we want to differentiate with respect to θ at $\theta = \theta_0$. The evolution of the queue is fully described by $\{(T_n, Z_n),\ n \ge 1\}$, where T_n is the arrival time of the nth customer and Z_n is its service time. In a simulation of this queue, when $\theta = \theta_0$, the sequence

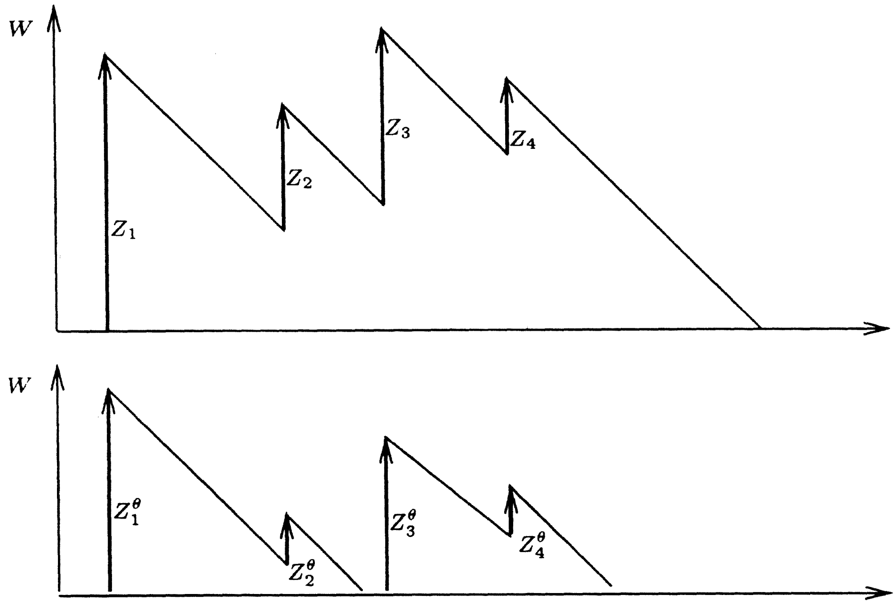

FIG. 1. *The IPA principle. Top: The nominal path; bottom: the altered path with the same jump times but different values of jumps.*

$\{Z_n,\ n \geq 0\}$ can be generated by taking $Z_n = G_{\theta_0}^{-1}(U_n)$, where $U \stackrel{\text{def}}{=} \{U_n,\ n \geq 0\}$ is a sequence of independent random variables, uniformly distributed over $[0,1]$. Following IPA principle, a perturbed path is generated with the same sequence U, but with nth service time Z_n^θ equal to $G_\theta^{-1}(U_n)$ (see Figure 1). In our example, IPA is known to work for the G/M/1 queue (i.e., the limit can be computed) and we have [see L'Ecuyer (1990); Suri and Zazanis (1988)]

$$\left(\frac{d}{d\theta} E_\theta[F]\right)_{\theta=\theta_0} = \frac{1}{K\theta_0} \mathbf{E}_{\theta_0}\left[\sum_{i=1}^{K} \sum_{j \in B_i} Z_j\right], \tag{2}$$

where B_i is the set containing customer i and all the customers that precede him in the same busy period.

It appears from the two previous examples that these two approaches require a very fine knowledge of the sample paths of the underlying process for the difference of expectations which appeared in (1) to be calculated. The LR method does not present a priori this default, but on the other hand it implies some restrictions on the "possible" differentiations.

In order for the LR method to be applicable, it is necessary that for each θ, $\mathbf{P}_\theta$ is locally absolutely continuous with respect to $\mathbf{P}_{\theta_0}$, that is, for any $t \geq 0$, the restriction of $\mathbf{P}_\theta$ to $\mathscr{F}_t$ is absolutely continuous with respect to the restriction of $\mathbf{P}_{\theta_0}$ to $\mathscr{F}_t$. From Jacod [(1979), pages 265–273], it is necessary and sufficient that the following conditions hold:

C1. $\mathbf{P}_\theta \ll \mathbf{P}_0$ on $\mathscr{F}_0$.

C2. There exists a nonnegative, $\mathscr{P}$-measurable process Y_θ such that

$$\nu_\theta(ds, dz) = Y_\theta(\omega, s, z)\nu_{\theta_0}(ds, dz).$$

C3. The process

$$t \mapsto C_t = \int_0^t \int_E \left(1 - \sqrt{Y_\theta}\right)^2 \nu_{\theta_0}(ds, dz)$$

satisfies $\mathbf{P}_\theta(C_t < +\infty) = 1$ for any $t \geq 0$.

Set $S_n = \inf\{t, C_t \geq n\}$ (with the convention that $S_n = +\infty$ if $C_t < n$ for any t). Whenever C1–C3 hold, we have

$$(3) \qquad \left.\frac{d\mathbf{P}_\theta}{d\mathbf{P}_0}\right|_{\mathscr{F}_t} = Z_0^\theta \cdot Z_t^\theta$$
$$\text{where} \quad Z_t^\theta \overset{\text{def}}{=} \mathscr{E}\left(\int_0^t \int_E (Y_\theta(\omega, s, z) - 1)(\omega - \nu_{\theta_0})(ds, dz)\right),$$

for $t \leq \limsup_n S_n$ and $Z_t = 0$ otherwise, where

$$Z_0^\theta \overset{\text{def}}{=} \left.\frac{d\mathbf{P}_\theta}{d\mathbf{P}_0}\right|_{\mathscr{F}_0}$$

and for any local martingale $M = \{M_t, t \geq 0\}$, $\{\mathscr{E}(M_t), t \geq 0\}$ is the solution of the stochastic differential equation

$$R_t = 1 + \int_0^t R_{s^-}\, dM_s.$$

REMARK 2.1. In the preamble, we assumed once and for all that N is quasi-left-continuous under $\mathbf{P}_0$. The main reason for this is that without this hypothesis, the conditions required to have local absolute continuity are too intricate.

REMARK 2.2. Note that a simpler but sufficient condition for C3 to hold is that [cf. Jacod (1979), page 273]

$$(\text{C3}') \qquad \mathbf{E}_0\left[\exp\left(\tfrac{1}{2}\int_0^T \int_E Y_\theta^2(s, z)\nu_{\theta_0}(ds, dz)\right)\right] < +\infty.$$

THEOREM 1 (LRM principle). *Assume that conditions* C1–C3 *hold and assume that the following two hypotheses are satisfied*:

HYPOTHESIS 2. *There exists a neighborhood* $\mathscr{V}(\theta_0) \subset \Theta$ *of* θ_0, *a predictable process* h *independent of* θ, *a constant* $c > 0$ *and a family of* *predictable processes* R_θ *satisfying*

$$Y_\theta(\omega, t, z) = 1 + (\theta - \theta_0)h(\omega, t, z)(1 + R_\theta(\omega, t, z)) \tag{4}$$

for any $\theta \in \mathscr{V}(\theta_0)$, *for any* (t, z), $\mathbf{P}_0$-*a.e.*,

$$\frac{d}{dt}\left(\int_0^t \int_E h(\omega, s, z)^2(1 + R_\theta(\omega, s, z))^2 \nu_{\theta_0}(ds, dz)\right) \le c, \tag{5}$$

for any $\theta \in \mathscr{V}(\theta_0)$, *for any* t, $\mathbf{P}_0$-*a.e.*,

$$R_\theta(\omega, s, z) > -1, \qquad d\mathbf{P}_0 \otimes \nu_{\theta_0}(ds, dz) \quad \text{a.e., for any } \theta \in \mathscr{V}(\theta_0), \tag{6}$$

R_θ *tends to* 0 *when* θ *goes to* θ_0, *in the sense that*

$$\lim_{\theta \to \theta_0} \mathbf{E}_0\left[\int_0^T \int_E h(\omega, s, z)^2 R_\theta(\omega, s, z)^2 \nu_{\theta_0}(ds, dz)\right] = 0. \tag{7}$$

HYPOTHESIS 3. *For any* $\theta \in \mathscr{V}(\theta_0)$, $Z_0^\theta = 1$; *see Remark* 2.4.

Then, for any square integrable, $\mathscr{F}_t$-*measurable functional* F, *we have*

$$\left(\frac{d}{d\theta} E_\theta[F]\right)_{\theta = \theta_0} = \mathbf{E}_0\left[F \int_0^t \int_E h(\omega, s, z)(\omega - \nu_{\theta_0})(ds, dz)\right]. \tag{8}$$

PROOF. Since F is $\mathscr{F}_t$-measurable, by the Girsanov theorem [see (3)] and Hypothesis 3,

$$\begin{aligned}\left(\frac{d}{d\theta}\mathbf{E}_\theta[F]\right)_{\theta=\theta_0} &= \lim_{\theta \to \theta_0} (\theta - \theta_0)^{-1}\mathbf{E}_0\left[F(Z_0^\theta - 1)\right] \\ &\quad + \lim_{\theta \to \theta_0} (\theta - \theta_0)^{-1}\mathbf{E}_0\left[FZ_0^\theta(Z_t^\theta - 1)\right] \\ &= \lim_{\theta \to \theta_0} (\theta - \theta_0)^{-1}\mathbf{E}_0\left[F(Z_t^\theta - 1)\right].\end{aligned} \tag{9}$$

For any θ, the process $\{M_t^\theta \overset{\text{def}}{=} \int_0^t \int_E (Y_\theta - 1)(\omega - \nu_{\theta_0})(ds, dz), t \ge 0\}$ is a martingale whose $\mathbf{P}_0$-Doob–Meyer process is

$$\langle M^\theta, M^\theta \rangle_t = (\theta - \theta_0)^2 \int_0^t \int_E h(\omega, s, z)^2(1 + R_\theta(\omega, s, z))^2 \nu_{\theta_0}(ds, dz).$$

Hence, using Ruiz de Chavez (1983) and condition (5) of Hypothesis 2, we have the $L^2(\mathbf{P}_0)$ expansion

$$\begin{aligned} Z_t^\theta &\overset{\text{def}}{=} \mathscr{E}\left(\int_0^t \int_E (Y_\theta(\omega, s, z) - 1)(\omega - \nu_{\theta_0})(ds, dz)\right) \\ &= 1 + \sum_{n > 0} \left(M_t^\theta\right)^{(n)}, \end{aligned}$$

where

$$\left(M_t^\theta\right)^{(n)} = \int_0^t \left(M_{s-}^\theta\right)^{(n-1)} dM_s^\theta$$

and

$$\left(M_t^\theta\right)^{(1)} = \int_0^t \int_E (Y_\theta(\omega, s, z) - 1)(\omega - \nu_{\theta_0})(ds, dz).$$

Hence,

$$\begin{aligned}
&(\theta - \theta_0)^{-1}\mathbf{E}_0\left[F\left(Z_t^\theta - 1\right)\right] \\
&\quad = \mathbf{E}_0\left[F \int_0^t \int_E h(\omega, s, z)(\omega - \nu_{\theta_0})(ds, dz)\right] \\
&\qquad + \mathbf{E}_0\left[F \int_0^t \int_E h(\omega, s, z) R_\theta(\omega, s, z)(\omega - \nu_{\theta_0})(ds, dz)\right] \\
&\qquad + (\theta - \theta_0) \sum_{n=2}^{+\infty} (\theta - \theta_0)^{n-2} \mathbf{E}_0\left[F \cdot \left(M_t^\theta\right)^{(n)}\right].
\end{aligned}$$

By condition (7), the second summand of the last equation tends to 0 when θ goes to θ_0. Moreover, by condition (5), for any $n \geq 2$,

$$\begin{aligned}
\mathbf{E}_0\left[\left|\left(M_t^\theta\right)^{(n)}\right|^2\right] &= \mathbf{E}_0\left[\left|\int_0^t \left(M_s^\theta\right)^{(n-1)} dM_s^\theta\right|^2\right] \\
&= \mathbf{E}_0\left[\int_0^t \left|\left(M_s^\theta\right)^{(n-1)}\right|^2 d\langle M^\theta, M^\theta\rangle_s\right] \\
&\leq c \int_0^t \mathbf{E}_0\left[\left|\left(M_s^\theta\right)^{(n-1)}\right|^2\right] ds.
\end{aligned}$$

Thus by induction,

$$\mathbf{E}_0\left[\left|\left(M_t^\theta\right)^{(n)}\right|^2\right] \leq \frac{(ct)^n}{n!}$$

and

$$\sup_{\theta \in \mathscr{V}(\theta_0)} \sum_{n=2}^{+\infty} (\theta - \theta_0)^{n-2} \mathbf{E}_0\left[F \cdot \left(M_t^\theta\right)^{(n)}\right] < +\infty.$$

Hence,

$$\lim_{\theta \to \theta_0} (\theta - \theta_0)^{-1}\mathbf{E}_0\left[F\left(Z_t^\theta - 1\right)\right] = \mathbf{E}_0\left[F \int_0^t \int_E h(\omega, s, z)(\omega - \nu_{\theta_0})(ds, dz)\right]$$

and the proof is complete. □

Remark 2.3. The proof can also be done when the time interval is random. Let S be a stopping time and let F be $\mathscr{F}_S$-measurable. Then we have

$$\left(\frac{d}{d\theta} E_\theta[F]\right)_{\theta = \theta_0} = \mathbf{E}_0\left[F \int_0^S \int_E h(\omega, s, z)(\omega - \nu_{\theta_0})(ds, dz)\right],$$

provided that

$$\mathbf{E}_0\left[\exp\left(\tfrac{1}{2}\int_0^S\int_E Y_\theta^2(s,z)\,\nu_{\theta_0}(ds,dz)\right)\right] < +\infty.$$

REMARK 2.4. In the perturbation analysis literature, one often aims to work under stationary regime. This would a priori prevent us from taking $Z_0^\theta = 1$ since there is no reason for $\mathbf{P}_\theta$ to coincide with $\mathbf{P}_{\theta_0}$ on $\mathscr{F}_0$ when these probabilities need to be stationary. Actually, another way to define "perturbation analysis in stationary regime" consists of assuming that the system has reached its equilibrium and that we analyze the sensitivity after a sudden change in the driving parameters. In that case, $\mathbf{P}_{\theta_0}$ is still the equilibrium distribution, but $\mathbf{P}_\theta$ is not and we can fix Z_0^θ to be equal to 1. This is implicitly the situation in the current literature on sensitivity analysis because of the difficulty arising when handling the first expectation on the right-hand side of (9).

EXAMPLE 1 (Continued). Consider again Example 1. E is reduced to a singleton and the compensating measure under $\mathbf{P}_\theta$ is given by

$$\nu_\theta(ds) = \theta\,ds.$$

All the conditions imposed in Theorem 1 are satisfied with $h \equiv \theta_0^{-1}$ provided that we work on a fixed time interval $[0,T]$. Note that Theorem 1 is thus an extension of Theorem 1 in Reiman and Weiss (1989b).

EXAMPLE 2 (Continued). As usual in the representation of queues by marked point processes, the marks represent the service times, in particular $E = \mathbf{R}^+$. The $\mathbf{P}_\theta$ compensating measure is given here by

$$\nu_\theta(ds,dz) = f(\omega,s)\,ds\,G_\theta(dz).$$

Assume furthermore that $G_\theta(dz) = g(\theta,z)\,ds$, where $g(\theta,z) > 0$ for any (θ,z). We get

$$\nu_\theta(ds,dz) = \frac{g(\theta,z)}{g(\theta_0,z)}\,\nu_{\theta_0}(ds,dz).$$

Hence, in view of Hypothesis 2, we have to assume that g is twice differentiable with respect to θ, that $\partial g/\partial\theta$ belongs to $L^2(\mathbf{R}, G_{\theta_0}(dz))$ and that $\partial^2 g/\partial\theta^2$ is bounded. In this case, one should take

$$h(s,z) = g(\theta_0,z)^{-1}\,\frac{\partial g}{\partial\theta}(\theta_0,z).$$

In order for condition C3′ to be satisfied on the random interval $[0,T_K]$, one should assume (as we do hereafter) that there exists $\varepsilon > 0$ such that

$$\mathbf{E}_0\left[\exp\left(\varepsilon\int_0^{T_K}\int_E g(\theta_0,z)^{-1}\,\frac{\partial g}{\partial\theta}(\theta_0,z)^2 f(\omega,s)^2\,ds\,dz\right)\right] < +\infty,$$

where T_K is the arrival time of the Kth customer.

In Example 1, when applying (8) to the Poisson process, we get for any F $\mathscr{F}_t$-measurable,

$$\left(\frac{d}{d\theta}\mathbf{E}_\theta[F]\right)_{\theta=\theta_0} = \mathbf{E}_0\left[F\cdot\int_0^t h(s)(\omega-\nu_{\theta_0})(ds)\right]$$
$$= \mathbf{E}_0\left[F\cdot\int_0^t h(s)\,d\tilde{N}_s\right],$$

where $h\equiv\theta_0^{-1}$ and $\tilde{N}$ is the compensated Poisson process, that is, $\tilde{N}_t = N_t - \theta_0 t$. Written this way, the latter expectation is nothing but one of the terms appearing in the so-called *stochastic integration by parts formula*. It turns out that integration by parts formulas are the core of the Malliavin calculus and that is why perturbation analysis can be naturally seen as a part of this latter theory. For instance, in case of the Poisson process, define, for any "nice" functional F and any $h\in L^2[0,1]$, the random variable $DF(h)$ by

$$DF(h) = \int_0^1 (F(\omega+\delta_s) - F(\omega))h(s)\,ds.$$

We then have

$$\mathbf{E}_0\left[F\cdot\int_0^t h(s)\,d\tilde{N}_s\right] = \mathbf{E}_0[DF(h)]. \tag{10}$$

This formula is the key point of this work: there exist at least three sensible ways to define $DF(h)$ in the sense that all of them are such that (10) holds; hence, all three of them give new expressions of the derivative we aim to compute. The rest of this paper is devoted to showing how $DF(h)$ can be defined and how these definitions are related and can be applied to sensitivity analysis.

3. A variational approach. In this section, we assume that ν_{θ_0} still satisfies Hypothesis 1, but also the following hypothesis:

HYPOTHESIS 4. *We have*

$$\nu_{\theta_0}(ds,dz) = q(\omega,s,z)\eta(dz)\,ds,$$

where η is a Radon measure on E, q is a predictable process and there exists $m>0$ and $Q(s,z)\in L^1(ds\otimes\eta(dz))$ such that

$$m\le q(\omega,s,z)\le Q(s,z),$$

for any s, z and $\mathbf{P}_0$-almost everywhere.

By $L^2_p(\mathbf{P}_0\otimes\nu_{\theta_0})$, we mean the set of predictable processes such that

$$\mathbf{E}_0\left[\int_0^T\int_E h(\omega,s,z)^2\nu_{\theta_0}(ds,dz)\right] < +\infty.$$

Remember that for such a process. By the Cauchy–Schwarz inequality,

$$
\begin{aligned}
&\mathbf{E}_0\left[\left|\int_0^T\int_E h(\omega,s,z)\,\omega(ds,dz)\right|^2\right]\\
&\quad\le 2\mathbf{E}_0\left[\left|\int_0^T\int_E h(\omega,s,z)(\omega-\nu_{\theta_0})(ds,dz)\right|^2\right]\\
&\qquad + 2\mathbf{E}_0\left[\left|\int_0^T\int_E h(\omega,s,z)\,\nu_{\theta_0}(ds,dz)\right|^2\right] \qquad (11)\\
&\quad\le 2\big(1+\nu_{\theta_0}([0,T]\times E)\big)\mathbf{E}_0\left[\int_0^T\int_E h(\omega,s,z)^2\,\nu_{\theta_0}(ds,dz)\right]\\
&\quad\le 2\big(1+\mathbf{E}_0[N_T^2]\big)\mathbf{E}_0\left[\int_0^T\int_E h(\omega,s,z)^2\,\nu_{\theta_0}(ds,dz)\right].
\end{aligned}
$$

DEFINITION 3.1. By $\mathscr{H}$ we denote the Hilbert space of deterministic processes $h(s,z)$ such that

$$\|h\|_{\mathscr{H}}^2 \overset{\text{def}}{=} \int_0^T\int_E\big(h(s,z)^2+\hat{h}(s,z)^2\big)Q(s,z)\,\eta(dz)\,ds < +\infty,$$

where

$$\hat{h}(t,z) = \int_0^t h(s,z)Q(s,z)\,ds.$$

LEMMA 1. *Step functions, that is, functions of the form*

$$\sum_{i=1}^{n-1}\alpha_i\mathbf{1}_{[t_i,t_{i+1})}(s)\mathbf{1}_{B_i}(z),$$

where for any i, $t_i<t_{i+1}$, $B_i\in\mathscr{B}(E)$ *and* $\alpha_i\in\mathbf{R}$ *are dense in* $\mathscr{H}$.

PROOF. Let f be orthogonal to all step functions, for any t_0, t_1 and any $B\in\mathscr{B}(E)$,

$$
\begin{aligned}
&\int_0^T\int_E \hat{f}(t,z)\int_{t_0}^{t_1}Q(r,z)\,dr\,\mathbf{1}_B(z)Q(t,z)\,\eta(dz)\,dt\\
&\qquad = -\int_{t_0}^{t_1}\int_E \mathbf{1}_B(z)f(t,z)Q(t,z)\,\eta(dz)\,dt.
\end{aligned}
$$

Hence $d\eta$ almost surely,

$$\int_0^T\hat{f}(t,z)Q(t,z)\,dt\cdot\int_{t_0}^{t_1}Q(r,z)\,dr = -\int_{t_0}^{t_1}f(t,z)Q(t,z)\,dt.$$

Taking $t_0=0$ and $t_1=T$, one gets

$$\int_0^T\hat{f}(t,z)Q(t,z)\,dt\cdot\left(\int_0^T Q(r,z)\,dr+1\right)=0.$$

Since Q is positive, $\hat{f}$ is identically zero (dt-a.e.) which in turn implies that $\|f\|_{\mathscr{H}}=0$. □

Denote by $\mathscr{S}$ the set of functionals of the form

$$F = f\left(\int_0^T \int_E f_1(s) g_1(z)\, \omega(ds, dz), \dots, \int_0^T \int_E f_n(s) g_n(z)\, \omega(ds, dz)\right),$$

where f is a bounded twice differentiable function with bounded derivatives, $f_i g_i$ belongs to $\mathscr{H}$ and f_i is continuously differentiable with bounded derivative for each $i = 1, \dots, n$.

DEFINITION 3.2. For any functional $F \in \mathscr{S}$ and any $h \in \mathscr{H}$, $DF(h)$ is defined by

$$\begin{aligned} DF(h) = -\sum_{i=1}^n \frac{\partial f}{\partial x_i}\bigg(&\int_0^T \int_E f_1(s) g_1(z)\, \omega(ds, dz), \dots, \\ &\int_0^T \int_E f_n(s) g_n(z)\, \omega(ds, dz)\bigg) \\ &\times \int_0^T \int_E f_i'(s) g_i(z) \left(\frac{1}{q(\omega, s, z)} \int_0^s h(r, z) q(\omega, r, z)\, dr\right) \\ &\quad \times \omega(ds, dz). \end{aligned} \tag{12}$$

THEOREM 2. *For any $F \in \mathscr{S}$, there exists $d > 0$ such that*

$$\mathbf{E}_0\left[|DF(h)|^2\right] \le d\mathbf{E}_0\left[\|h\|_{\mathscr{H}}^2\right],$$

for any predictable process h such that for any ω, $h(\omega, \cdot, \cdot)$ belongs to $\mathscr{H}$.

PROOF. Since F belongs to $\mathscr{S}$, q is lower bounded, and using (11), there exists $c > 0$,

$$\begin{aligned} &\mathbf{E}_0\left[|DF(h)|^2\right] \\ &\quad \le n \cdot c\mathbf{E}_0\left[\left|\int_0^T \int_E \frac{1}{q(\omega, s, z)} \int_0^s h(r, z) q(\omega, r, z)\, dr\, \omega(ds, dz)\right|^2\right] \\ &\quad \le \frac{n \cdot c}{m^2} \mathbf{E}_0\left[\int_0^T \int_E \left|\hat{h}(r, z)\right|^2 \nu_{\theta_0}(ds, dz)\right] \\ &\quad \le \frac{n \cdot c}{m^2} \|h\|_H^2, \end{aligned}$$

where c is a generic constant. □

As an easy consequence of the definition, we have the following lemma.

LEMMA 2. *For any $F, G \in \mathscr{S}$, FG belongs to $\mathscr{S}$ and*

$$D(FG)(h) = F \cdot DG(h) + G \cdot DF(h).$$

We now aim to prove the stochastic integration by parts formula, that is,

$$\mathbf{E}_0[DF(h)] = \mathbf{E}_0\left[F \int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right].$$

Consider

$$v(\omega, t, z) = \int_0^t q(\omega, s, z)\, ds.$$

For (ω, z) fixed, the map $t \mapsto v(\omega, t, z)$ is an increasing function. Hence we can consider its right inverse defined by

$$v^{-1}(w, t, z) = \inf\{r \geq 0, v(\omega, r, z) = t\}.$$

Define

$$v_\theta^h(\omega, t, z) = v^{-1}\left(\omega, \int_0^t (1 + (\theta - \theta_0) h(s, z)) q(\omega, s, z)\, ds, z\right) - t.$$

For any $h \in \mathscr{H}$ and nonpositive, consider the map τ_θ^h from Ω into itself, where $\tau_\theta^h \omega$ is the random measure defined by

$$\begin{aligned} &\iint \mathbf{1}_{[0,t)}(s) \mathbf{1}_B(z) \tau_\theta^h \omega(ds, dz) \\ &\quad = \iint \mathbf{1}_{[0,t)}\big(s + v_\theta^h(s, z)\big) \mathbf{1}_B(z)\, \omega(ds, dz), \end{aligned} \tag{13}$$

for any B in $\mathscr{B}(E)$ and any $t \in [0, T]$. Actually, $\tau_\theta^h \omega$ is the process which jumps at time $s + v_\theta^h(\omega, s, z)$ with mark z if and only if ω jumps at time s with mark z.

Let $\mathbf{Q}_\theta$ be the probability measure defined by $d\mathbf{Q}_\theta = \bar{Z}_T^\theta\, d\mathbf{P}_{\theta_0}$ on $\mathscr{F}_T$, where

$$\bar{Z}_T^\theta \overset{\text{det}}{=} \mathscr{E}\left((\theta - \theta_0) \int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right).$$

THEOREM 3. *For any $\theta \geq \theta_0$ and any $h \in \mathscr{H}$, h nonpositive, the law of the marked point process $\tau_\theta^h \omega$ under $\mathbf{Q}_\theta$ is the same as the law of ω under $\mathbf{P}_0$.*

PROOF. Since $h \in L_p^2(\mathbf{P}_0 \otimes \nu_{\theta_0})$, by condition C3′, the probability measure $\mathbf{Q}_\theta$ is well defined and the process

$$\begin{aligned} \bar{Z}_t^\theta &= \mathbf{E}_0\big[\bar{Z}_T^\theta \mid \mathscr{F}_t\big] \\ &= \mathscr{E}\left((\theta - \theta_0) \int_0^t \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right) \end{aligned}$$

is a square-integrable martingale.

The process v_θ^h is left-continuous and since h is negative, it is adapted (and nonpositive); hence it is predictable so that $\tau_\theta^h \omega$ is well defined. Consider the martingale M:

$$M_t = \iint (1 + (\theta - \theta_0) h(s, z)) \mathbf{1}_{[0,t) \times B}\big(s + v_\theta^h(s, z), z\big)(\omega - \nu_{\theta_0})(ds, dz).$$

We have

$$\int_0^t (\bar{Z}_s^\theta)^{-1} d[M^\theta, Z^\theta]_s$$
$$= \iint (1 + (\theta - \theta_0)h(s,z))\mathbf{1}_B(z)\mathbf{1}_{[0,t)}\big(s + v_\theta^h(\omega,s,z)\big)$$
$$\times\Big(1 - (1 + (\theta-\theta_0)h(s,z))^{-1}\Big)\omega(ds,dz)$$
$$= \iint (1 + (\theta - \theta_0)h(s,z))\mathbf{1}_B(z)\mathbf{1}_{[0,t)}\big(s + v_\theta^h(\omega,s,z)\big)\omega(ds,dz)$$
$$- \tau_\theta^h\omega([0,t]\times B).$$

Thus

$$M_t^\theta - \int_0^t (Z_s^\theta)^{-1} d[M^\theta, Z^\theta]_s$$
$$= \tau_\theta^h\omega([0,t]\times B) - \iint (1 + (\theta-\theta_0)h(s,z))\mathbf{1}_B(z)$$
$$\times \mathbf{1}_{[0,t)}\big(s + v_\theta^h(\omega,s,z)\big)\nu_{\theta_0}(ds,dz) \tag{14}$$
$$= \tau_\theta^h\omega([0,t]\times B) - \iint (1 + (\theta-\theta_0)h(s,z))\mathbf{1}_B(z)$$
$$\times \mathbf{1}_{[0,t)}\big(s + v_\theta^h(\omega,s,z)\big)q(\omega,s,z)\,ds\,dz.$$

By the definition of v_θ^h,

$$V(t) \stackrel{\text{def}}{=} v\big(t + v_\theta^h(\omega,t,z)\big) = \int_0^t (1 + (\theta-\theta_0)h(s,u))q(\omega,s,z)\,ds.$$

Hence, on one hand,

$$\frac{\partial V}{\partial t}(t) = (1 + (\theta-\theta_0)h(t,z))q(\omega,t,z)$$

and on the other hand,

$$\frac{\partial V}{\partial t}(t) = q\big(\omega, t + v_\theta^h(\omega,t,z), z\big)\left(1 + \frac{\partial v_\theta^h}{\partial t}(\omega,t,z)\right).$$

Thus, using the change of variable $u = s + v_\theta^h(\omega,s,z)$ in (14), we get

$$M_t^\theta - \int_0^t (\bar{Z}_s^\theta)^{-1} d[M^\theta, Z^\theta]_s = \tau_\theta^h\omega([0,t]\times B) - \nu_{\theta_0}([0,t]\times B).$$

Then $M_t^\theta - \int_0^t (Z_s^\theta)^{-1} d[M^\theta, Z^\theta]_s$ is a $\mathbf{Q}_\theta$ martingale, so ν_{θ_0} is the $\mathbf{Q}_\theta$ compensating measure of $\tau_\theta^h\omega$. This means that the $\mathbf{Q}_\theta$-law of $\tau_\theta^h\omega$ is the same as the law of ω under $\mathbf{P}_0$; cf. Jacod (1979), page 86. □

DEFINITION 3.3. A functional $F\colon \Omega \to \mathbf{R}$ is smooth whenever, for any $h \in \mathscr{H}$, there exists a square-integrable random variable, denoted by $\overline{D}F(h)$ such that

$$\lim_{\substack{\theta \to \theta_0 \\ \theta > \theta_0}} (\theta - \theta_0)^{-1}\big(F(\omega) - F(\tau_\theta^h \omega) - (\theta - \theta_0)\overline{D}F(h)\big) = 0,$$

where the limit is taken in $L^2(\mathbf{P}_0)$.

THEOREM 4. *Any $F \in \mathscr{S}$ is smooth and for any $h \in \mathscr{H}$.*

$$DF(h) = \overline{D}F(h).$$

PROOF. Consider the case when $f \equiv x$ and $n = 1$. Using the Taylor expansion at order 1 and 2,

$$\begin{aligned}
&\mathbf{E}_{\theta_0}\Bigg[\bigg|\int_0^T \int_E f_1(s)\big(g_1(z + v_\theta^h(\omega, s, z)) - g_1(z)\big)\omega(ds, dz) \\
&\qquad - (\theta - \theta_0)\int_0^T \int_E f_1(s) g_1'(z) \frac{\partial v_{\theta_0}^h}{\partial \theta}(\omega, s, z)\,\omega(ds, dz)\bigg|^2\Bigg] \\
&\quad \le 2\mathbf{E}_0\Bigg[\bigg|\int_0^T \int_E \big(g_1(z + v_\theta^h(\omega, s, z)) - g_1(z) \\
&\qquad\qquad - g_1'(z) v_{\theta_0}^h(\omega, s, z)\big) f_1(s)\,\omega(ds, dz)\bigg|^2\Bigg] \\
&\qquad + 2\mathbf{E}_0\Bigg[\bigg|\int_0^T \int_E f_1(s) g_1'(z) \\
&\qquad\qquad \times \left(v_\theta^h(\omega, s, z) - (\theta - \theta_0)\frac{\partial v_{\theta_0}^h}{\partial \theta}(\omega, s, z)\right)\omega(ds, dz)\bigg|^2\Bigg] \qquad (15)\\
&\quad \le 2\,\| f_1 g_1' \|_\infty^2\, \mathbf{E}_0\Bigg[\bigg|\int_0^T \int_E v_\theta^h(\omega, s, z)\,\omega(ds, dz)\bigg|^2; \\
&\qquad\qquad \bigg|\int_0^T \int_E v_\theta^h(\omega, s, z)\,\omega(ds, dz)\bigg| \ge 1\Bigg] \\
&\qquad + 2\,\| f_1 g_1'' \|_\infty^2\, \mathbf{E}_0\Bigg[\bigg|\int_0^T \int_E v_\theta^h(s, z)\,\omega(ds, dz)\bigg|^4; \\
&\qquad\qquad \bigg|\int_0^T \int_E v_\theta^h(\omega, s, z)\,\omega(ds, dz)\bigg| \le 1\Bigg]
\end{aligned}$$

$$+\|f_1 g_1'\|_\infty^2 \mathbf{E}_0\left[\left|\int_0^T\int_E\left|v_\theta^h(\omega,s,z)\right.\right.\right.$$

$$\left.\left.\left.-(\theta-\theta_0)\frac{\partial v_{\theta_0}^h}{\partial\theta}(\omega,s,z)\right|\omega(ds,dz)\right|^2\right].$$

By the usual derivation rules,

$$\begin{aligned}\left(\frac{\partial v_{\theta_0}^h}{\partial\theta}(\omega,t,z)\right)_{\theta=\theta_0} &= \left\{\frac{\partial v}{\partial\theta}\left(\omega,\int_0^t q(\omega,s,z)\,ds,z\right)\right\}^{-1}\\ &\qquad\times\int_0^t h(s,z)q(\omega,s,z)\,ds\\ &= q(\omega,t,z)^{-1}\int_0^t h(s,z)q(\omega,s,z)\,ds.\end{aligned}$$

Hence, by the Taylor expansion again,

$$\begin{aligned}&\left|v_\theta^h(\omega,s,z)-(\theta-\theta_0)\frac{\partial v_{\theta_0}^h}{\partial\theta}(\omega,s,z)\right|\\ &\quad\le\left|v_\theta^h(\omega,s,z)-v_{\theta_0}^h(\omega,s,z)\right|+\left|(\theta-\theta_0)\frac{\partial v_{\theta_0}^h}{\partial\theta}(\omega,s,z)\right|\\ &\quad\le\frac{2}{m}\,|\theta-\theta_0|\int_0^t\int_E|h(s,z)|\,Q(s,z)\,ds,\end{aligned}$$

since $v_{\theta_0}^h\equiv 0$. Thus, by dominated convergence, it follows that

$$\lim_{\theta\to\theta_0}(\theta-\theta_0)^{-1}\left(v_\theta^h(s,z)-(\theta-\theta_0)\frac{\partial v_{\theta_0}^h}{\partial\theta}(s,z)\right)=0$$

in $L^2(\mathbf{P}_0\otimes\nu_{\theta_0})$ and thus the three expectations on the right-hand side of (15), when divided by $(\theta-\theta_0)^2$, tend to 0 as θ goes to θ_0. Hence in this case, $\overline{D}F(h)=DF(h)$.

In the general case, denote by $\sum_{i=1}^n F_i'(h)$ the right-hand side of (12). We have, by the second order Taylor expansion,

$$\begin{aligned}&\mathbf{E}_0\left[\left|F(\tau_\theta^h\omega)-F(\omega)-(\theta-\theta_0)\sum_{i=1}^n F_i'(h)\right|^2\right]\\ &\quad\le d\,\sup_i\left\|\frac{\partial f}{\partial x_i}\right\|_\infty^2\\ &\qquad\times\sum_{i=1}^n\mathbf{E}_0\left[\left|\int_0^T\int_E f_i(s)g_i(z)(\tau_\theta^h\omega-\omega)(ds,dz)-F_i'(h)\right|^2\right]\end{aligned}$$

$$+d \sup_{i,j} \left\| \frac{\partial^2 f}{\partial x_i \, \partial x_j} \right\|_\infty^2$$
$$\times \sum_{i=1}^{n} \mathbf{E}_0\left[\left|\int_0^T \int_E f_i(s) g_i(z)\left(\tau_\theta^h\omega - \omega\right)(ds, dz) - F_i'(h)\right|^4\right],$$

where d is a constant. Using the first part of this proof, the result follows. □

THEOREM 5 (Stochastic integration by parts). *For any $F \in \mathscr{S}$ and for any $h \in \mathscr{H}$,*

$$\mathbf{E}_0[DF(h)] = \mathbf{E}_0\left[F \int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right]. \tag{16}$$

PROOF. Theorems 3 and 4 induce that

$$\mathbf{E}_0[F] = \mathbf{E}_0\left[F(\tau_\theta^h\omega)\mathscr{E}\left((\theta - \theta_0)\int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right)\right], \tag{17}$$

for any $F \in \mathscr{S}$ and any negative element h of $\mathscr{H}$. Formula (16) follows by differentiating (17) with respect to θ. By linearity, (16) holds for the step element h of $\mathscr{H}$. Since, by Theorem 2,

$$\mathbf{E}_0\left[|DF(h)|^2\right] \le c\mathbf{E}_0\left[\|h\|_{\mathscr{H}}^2\right],$$

and by the Cauchy–Schwarz inequality,

$$\begin{aligned}
&\left|\mathbf{E}_0\left[F \int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right]\right|^2 \\
&\quad \le \mathbf{E}_0[F^2]\mathbf{E}_0\left[\int_0^T \int_E h(s, z)^2 \nu_{\theta_0}(ds, dz)\right] \\
&\quad \le \mathbf{E}_0[F^2]\mathbf{E}_0\left[\|h\|_{\mathscr{H}}^2\right],
\end{aligned}$$

it follows by the density of step functions in $\mathscr{H}$ (see Lemma 1) that (16) holds true for any $h \in \mathscr{H}$. □

THEOREM 6. *The set $\mathscr{S}$ is dense in $L^2(\mathbf{P}_0)$.*

PROOF. There exists $\{B_n,\ n \ge 0\}$ a sequence of compact sets of E such that $\bigcup_n B_n = E$ and $\eta(B_n) < +\infty$, for any n. Let $\{t_n,\ n \ge 0\}$ be an enumeration of $[0, T] \cap \mathbb{Q}$. The canonical filtration is generated by the set

$$\left\{\int_0^T \int_E \mathbf{1}_{[0, t_i)}(s)\mathbf{1}_{B_j}(z)\,\omega(ds, dz),\ i, j \ge 0\right\}.$$

Let ψ be a bijection from $\mathbf{N} \times \mathbf{N}$ onto $\mathbf{N}$ and

$$\mathscr{F}_n \overset{\text{def}}{=} \sigma\left\{\int_0^T \int_E \mathbf{1}_{[0, t_i)}(s)\mathbf{1}_{B_j}(z)\,\omega(ds, dz),\ i, j \text{ such that } \psi(i, j) \le n\right\}.$$

Since $\mathscr{F}_T = \bigvee_n \mathscr{F}_n$, by the martingale convergence theorem, for any $F \in L^2(\mathbf{P}_0)$, the sequence $\{\mathbf{E}_0[F \mid \mathscr{F}_n],\ n \geq 0\}$ converges to F in $L^2(\mathbf{P}_0)$. Moreover, by Doob's lemma, there exists f measurable from $\mathbf{R}^n$ into $\mathbf{R}$ such that

$$\mathbf{E}_0[F \mid \mathscr{F}_n] = f\left(\int_0^T \int_E \mathbf{1}_{[0,t_i)}(s)\mathbf{1}_{B_j}(z)\,\omega(ds,dz),\ i,j \text{ s.t. } \psi(i,j) \leq n\right).$$

It is then classical to approximate $\mathbf{E}_0[F \mid \mathscr{F}_n]$ and thus F by a sequence of elements of $\mathscr{S}$. □

COROLLARY 1. *For any $h \in \mathscr{H}$, the map $F \mapsto DF(h)$ is closable.*

PROOF. Let $\{F_n,\ n \geq 1\}$ be a sequence of $\mathscr{S}$ such that F_n converges to 0 in $L^2(\mathbf{P}_0)$ and for any $h \in \mathscr{H}$, $DF_n(h)$ converges to a limit denoted by $\zeta(h)$. For any $G \in \mathscr{S}$,

$$\begin{aligned}
\mathbf{E}_0[\zeta(h) \cdot G] &= \lim_{n \to +\infty} \mathbf{E}_0[DF_n(h)G] \\
&= \lim_{n \to +\infty} \left(\mathbf{E}_0[D(F_nG)(h)] - \mathbf{E}_0[F_n DG(h)]\right) \\
&= \lim_{n \to +\infty} \left(\mathbf{E}_0\left[F_n G \int_0^T \int_E h(s,z)\,\nu_{\theta_0}(ds,dz)\right] - \mathbf{E}_0[F_n DG(h)]\right) \\
&= 0.
\end{aligned}$$

Since $\mathscr{S}$ is dense in $L^2(\mathbf{P}_0)$, $\zeta(h) = 0$ $\mathbf{P}_0$-a.e. □

DEFINITION 3.4. The set $\mathbb{D}_{2,1}$ is the closure of $\mathscr{S}$ for the τ-topology defined by its converging sequences as the sequence $\{F_n,\ n \geq 0\}$ of elements of $\mathscr{S}$ converges for the τ-topology to F whenever $\{F_n,\ n \geq 0\}$ tends to F in L^2 and $DF_n(h)$ converges weakly in L^2 for any $h \in \mathscr{H}$.

PROPOSITION 1. *For any $F \in \mathbb{D}_{2,1}$ and any $h \in \mathscr{H}$,*

$$\mathbf{E}_0[DF(h)] = \mathbf{E}_0\left[F \cdot \int_0^T \int_E h(s,z)(\omega - \nu_{\theta_0})(ds,dz)\right].$$

PROOF. Formula (16) holds for $F \in \mathscr{S}$ (which is known to be smooth); hence by a limiting procedure, it still holds for F in $\mathbb{D}_{2,1}$. □

PROPOSITION 2. *For any functionals F in $\mathbb{D}_{2,1}$ and any function φ in $\mathscr{C}^2(\mathbf{R})$ with bounded derivatives, we have*

$$D_\varphi(F)(h) = \varphi'(F)DF(h). \tag{18}$$

PROOF. For $F \in \mathscr{S}$, it is clear that $\varphi(F)$ still belongs to $\mathscr{S}$ and by the usual derivation rules,

$$D_\varphi(F)(h) = \varphi'(F)DF(h).$$

Let $F \in \mathbb{D}_{2,1}$ and $\{F_n, n \geq 0\}$ a sequence of elements of $\mathscr{S}$ converging to F in $\mathbb{D}_{2,1}$. We have

$$\mathbf{E}_0\Big[|\varphi(F_n) - \varphi(F)|^2\Big] \leq \|\varphi'\|_\infty^2 \mathbf{E}_0\Big[|F_n - F|^2\Big]$$

and

$$\begin{aligned}&\big|\mathbf{E}_0[G\varphi'(F)DF(h) - G\varphi'(F_n)DF_n(h)]\big|\\ &\qquad \leq \|\varphi'\|_\infty \big|\mathbf{E}_0[G(DF(h) - DF_n(h))]\big|;\end{aligned}$$

hence $\varphi(F)$ is the limit in $\mathbb{D}_{2,1}$ of $(\varphi_n(F), n \geq 1)$ and (18) is true. □

Formula (18) and a limiting procedure yield the following proposition.

PROPOSITION 3. *If F is in $\mathbb{D}_{2,1}$, then $|F|$ and $F^+ = \max(0, F)$ belong to $\mathbb{D}_{2,1}$ and*

$$\begin{aligned}D|F|(h) &= DF(h)\big(\mathbf{1}_{\{F>0\}} - \mathbf{1}_{\{F<0\}}\big),\\ DF^+(h) &= DF(h)\mathbf{1}_{\{F>0\}},\end{aligned} \tag{19}$$

$$DF(h)\mathbf{1}_{\{F=0\}} = 0. \tag{20}$$

PROOF. Let $F_n = \sqrt{F^2 + 1/n}$. Then $(F_n)_n$ converges a.s. to $|F|$ as n goes to $+\infty$ and from (18), F_n belongs to $\mathbb{D}_{2,1}$ and

$$DF_n(h) = \frac{F}{\sqrt{F^2 + 1/n}} DF(h) = \frac{F}{\sqrt{F^2 + 1/n}} DF(h)\mathbf{1}_{\{F \neq 0\}}.$$

Now we see that $\|DF_n(h)\|_{L^2(\mathbf{P}_0)}$ are bounded uniformly with respect to n; hence there exists a weakly convergent subsequence $(DF_{n_k}(h))_k$ in $L^2(\mathbf{P}_0)$. Since $DF_{n_k}(h)$ converges almost surely to $DF(h)(\mathbf{1}_{\{F>0\}} - \mathbf{1}_{\{F<0\}})$, it follows that $|F|$ belongs to $\mathbb{D}_{2,1}$ with

$$D|F| = DF(h)\big(\mathbf{1}_{\{F>0\}} - \mathbf{1}_{\{F<0\}}\big).$$

Since $F^+ = (F + |F|)/2$, it follows that

$$\begin{aligned}DF^+(h) &= \tfrac{1}{2}\Big(DF(h)\big(\mathbf{1}_{\{F>0\}} + \mathbf{1}_{\{F<0\}} + \mathbf{1}_{\{F=0\}}\big) + DF(h)\big(\mathbf{1}_{\{F>0\}} - \mathbf{1}_{\{F<0\}}\big)\Big)\\ &= DF(h)\big(\mathbf{1}_{\{F>0\}} + \tfrac{1}{2}\mathbf{1}_{\{F=0\}}\big).\end{aligned}$$

If F is nonnegative, $F = F^+$ almost surely. Thus

$$DF(h)\mathbf{1}_{\{F=0\}} = \tfrac{1}{2}DF(h)\mathbf{1}_{\{F=0\}};$$

hence $DF(h)\mathbf{1}_{\{F=0\}} = 0$. In general,

$$DF(h)\mathbf{1}_{\{F=0\}} = (DF^+(h) - DF^-(h))\mathbf{1}_{\{F^+=0\}}\mathbf{1}_{\{F^-=0\}} = 0.$$

Reporting this result in the current expression of $DF^+(h)$ yields (19). □

REMARK 3.1. The key result of this part is in fact Theorem 3. Indeed, thanks to it, we are able to find the convenient expression of $DF(h)$ for regular functionals. By *convenient* expression, we mean here that (16) should hold. By looking deeper in the previous construction, one should realize that the main idea [which comes from Bismut (1983)] is the construction of a family of perturbations $\{\tau_\theta^h,\ \theta \in \Theta\}$ of the sample paths and a new probability measure $\mathbf{P}_\theta$ such that, for any θ, the modified process $\tau_\theta^h \omega$ under the new law $\mathbf{P}_\theta$ has the same law as the original process under the reference probability $\mathbf{P}_{\theta_0}$; see (17). We have in fact two main possibilities to find τ_θ^h: either it is obtained by modifying the jumps magnitude (see Bismut (1983); Bass and Cranston (1986); Bichteler and Jacod (1983); Norris (1987); Privault (1994)] or by changing the jumps times [see Decreusefond (1994); Privault (1994)].

We have worked here with the transformations obtained by changing the jump times, but the same lines can be followed for the other approach. Modifying the jump magnitudes is meaningful only when η is the Lebesgue measure (this implies that $E = \mathbf{R}^d$). The unique change is the definition of $\mathscr{S}$ and of the derivative of an element of $\mathscr{S}$.

DEFINITION 3.5 (Perturbing jump magnitudes). Assume that $\eta(dz) = dz$, where dz is the Lebesgue measure on $\mathbf{R}$. Denote by $\mathscr{S}$ the set of functionals of the form

$$F = f\left(\int_0^T \int_E f_1(s) g_1(z)\, \omega(ds, dz), \dots, \int_0^T \int_E f_n(s) g_n(z)\, \omega(ds, dz)\right),$$

where f is a bounded twice differentiable function with bounded derivatives, $f_i g_i$ belongs to $\mathscr{H}$ and g_i is continuously differentiable with bounded derivative for each $i = 1, \dots, n$.

For any functional $F \in \mathscr{S}$ and any $h \in \mathscr{H}$, $DF(h)$ is defined by

$$\begin{aligned} DF(h) &= -\sum_{i=1}^n \frac{\partial f}{\partial x_i}\left(\int_0^T \int_E f_1(s) g_1(z)\, \omega(ds, dz), \dots,\right. \\ &\qquad\qquad \left.\int_0^T \int_E f_n(s) g_n(z)\, \omega(ds, dz)\right) \\ &\qquad \times \int_0^T \int_E f_i(s) g_i'(z) \left(\frac{1}{q(\omega, s, z)} \int_{-\infty}^z h(s, u) q(\omega, s, u)\, du\right) \omega(ds, dz). \end{aligned}$$

The sequel follows without any difference. When we have the choice between the two possibilities, the main difference between the two approaches lies in the set of differentiable functionals. For instance, when altering the jump times, the functional $\omega \mapsto N_t(\omega)$ (i.e., the number of jumps up to time t) does not belong to $\mathbb{D}_{2,1}$. Conversely, when changing the jump magnitudes, this functional belongs to the associated space $\mathbb{D}_{2,1}$.

EXAMPLE 2 (Continued). Recall that F is the average waiting time of the first K customers and assume enough regularity for g. The waiting time of the ith customer is given by the well known formula $W_i = (W_{i-1} + Z_i - (T_i - T_{i-1}))^+$. For any i, it is clear that the functionals $Z_i(\omega)$ and $T_i(\omega) - T_{i-1}(\omega)$ belong to $\mathbb{D}_{2,1}$ and then that W_i is also in $\mathbb{D}_{2,1}$ with

$$DW_i(h) = \{DW_{i-1}(h) + DZ_i(h) - D(T_i - T_{i-1})(h)\} \cdot \mathbf{1}_{\{W_i > 0\}}. \tag{21}$$

Let $\kappa_i = \sup_{j \le i}\{j,\ W_j = 0\}$, that is, κ_i is the index of the customer who initiates the busy period the ith customer belongs to. Since for $\kappa_i < l \le i$, $W_l > 0$ and $W_{\kappa_i} = 0$, by iteration of (21), we get

$$DW_i(h) = \sum_{j \in B_i} DZ_j(h) - D(T_j - T_{j-1})(h).$$

Moreover, for the perturbation we are considering, $D(T_j - T_{j-1})(h) = 0$ because we only modify the jump magnitudes and

$$\begin{aligned} DZ_j(h) &= -\lim_{\theta \to \theta_0} (\theta - \theta_0)^{-1} v_\theta^h(T_i, Z_i) = -\left(\frac{\partial v_\theta^h}{\partial \theta}(T_i, Z_i)\right)_{\theta = \theta_0} \\ &= g(\theta_0, Z_i)^{-1} \int_0^{Z_i} \frac{\partial g}{\partial \theta}(\theta_0, u) g(\theta_0, u)\, du. \end{aligned}$$

Hence, we obtain

$$\left(\frac{d}{d\theta} E_\theta[F]\right)_{\theta = \theta_0} = \frac{1}{K} \sum_{i=1}^{K} \mathbf{E}_0\left[\sum_{j \in B_i} g(\theta_0, Z_j)^{-1} \int_0^{Z_j} \frac{\partial g}{d\theta}(\theta_0, u) g(\theta_0, u)\, du\right].$$

Note that when Z_i is exponentially distributed with parameter θ^{-1}, we have

$$g(\theta_0, Z_j)^{-1} \int_0^{Z_j} \frac{\partial g}{d\theta}(\theta_0, u) g(\theta_0, u)\, du = \frac{Z_i}{\theta_0},$$

so that we obtain a generalization of (2).

Viewing IPA as a part of the stochastic analysis enables us to answer the following conjecture: experimental data tend to prove that estimates deduced from IPA have a lower variance than those obtained with LRM. For a given perturbation h, we know that IPA works for smooth functionals and that

$$\left(\frac{d}{d\theta} \mathbf{E}_\theta[F]\right)_{\theta = \theta_0} = \mathbf{E}_0[DF(h)].$$

On the other hand, by LRM, we get

$$\left(\frac{d}{d\theta} \mathbf{E}_\theta[F]\right)_{\theta = \theta_0} = \mathbf{E}_0\left[F \int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right].$$

From a statistical point of view, we can estimate the derivative $(d/d\theta)\mathbf{E}_\theta[F]$ by averaging either $DF(h)$ or the product of $F \int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)$ over a large number of sample paths. Comparing the variances of these two estimates is thus comparing $\mathbf{E}_0[DF(h)^2]$ and

$$\mathbf{E}_0\left[F^2 \left|\int_0^T \int_E h(s, z)(\omega - \nu_{\theta_0})(ds, dz)\right|^2\right].$$

For F constant, DF is null and then if one method yields estimates with lower variances, it has to be IPA. Nevertheless, for $F = \varphi(F_1)$, F_1 smooth and φ in $\mathscr{C}_b^2$, we have

$$\mathbf{E}_0\left[F^2\left|\int_0^T\int_E h(s,z)(\omega-\nu_{\theta_0})(ds,dz)\right|^2\right]$$
$$\leq \|\varphi\|_\infty^2\,\mathbf{E}_0\left[\int_0^T\int_E h(s,z)^2\nu_{\theta_0}(ds,dz)\right],$$

whereas

$$\mathbf{E}_0\left[DF(h)^2\right] = \mathbf{E}_0\left[|\varphi'(F_1)|^2\,|DF_1(h)|^2\right].$$

Since we can choose φ bounded but with an arbitrary large derivative, it is easy to see that we can achieve a lower variance for some estimates originating from LRM.

4. Chaos decomposition and applications. A rather different approach to define $DF(h)$ consists of using the so-called chaos decomposition. To distinguish this object from that previously defined, the new object will be denoted by $\hat{D}F(h)$.

We denote by $L_d^2(\mathbf{P}_0 \otimes \nu_{\theta_0}^{\otimes n})$ [respectively, $L_p^2(\mathbf{P}_0 \otimes \nu_{\theta_0}^{\otimes n})$] the Hilbert space of deterministic real-valued functions (respectively, real-valued predictable processes) defined on $(\mathbf{R}^+ \otimes E)^n$ [respectively, $\Omega \times (\mathbf{R}^+ \times E)^n$] which are square-integrable for the measure $\mathbf{P}_0 \otimes \otimes_{i=1}^n \nu_{\theta_0}(ds_i, dz_i)$; that is,

$$\mathbf{E}_0\left[\iint_{(\mathbf{R}^+\times E)^n} f_n(s_1,z_1,\dots,s_n z_n)^2 \bigotimes_{i=1}^n \nu_{\theta_0}(ds_i,dz_i)\right] < +\infty.$$

Let $S_n \overset{\text{def}}{=} \{(s_1,\dots,s_n) \in [0,T]^n,\ 0 < s_n < \cdots < s_1 < T\}$. The nth order integral $I_n(f_n)$ of a deterministic function $f_n \in L_s^2(\mathbf{P}_0 \otimes \nu_{\theta_0}^{\otimes n})$ [i.e., f_n belongs to $L_d^2(\mathbf{P}_0 \otimes \nu_{\theta_0}^{\otimes n})$ and is symmetrical] is defined by

$$I_n(f_n) \overset{\text{def}}{=} n! \iint_{S_n\times E^n} f_n(s_1,z_1,\dots,s_n,z_n) \bigotimes_{i=1}^n (\omega-\nu_{\theta_0})(ds_i,dz_i).$$

When f_n belongs only to $L_d^2(\mathbf{P}_0 \otimes \nu_{\theta_0}^{\otimes n})$, we set $I_n(f_n) = I_n(\hat{f}_n)$, where $\hat{f}_n$ is the symmetrization of f_n defined by

$$\hat{f}_n(s_1,z_1,\dots,s_n,z_n) = \frac{1}{n!}\sum_{\sigma\in\Sigma_n} f_n(s_{\sigma(1)},z_{\sigma(1)},\dots,s_{\sigma(n)},z_{\sigma(n)})$$

and Σ_n is the group of the permutations of $\{1,\dots,n\}$. Define $C_0 = \mathbf{R}$ and

$$C_n = \overline{\operatorname{span}}\left\{I_n(h),\ h \in L_d^2\left(\mathbf{P}_0 \otimes \nu_{\theta_0}^{\otimes n}\right)\right\},$$

where $\overline{\operatorname{span}}\{\cdots\}$ represents the $L^2(\mathbf{P}_{\theta_0})$ closure of the vector space spanned by $\{\cdots\}$. Define also $\mathscr{C}_0 = \mathbf{R}$ and

$$\mathscr{C}_n = C_n \ominus (\mathscr{C}_0 \oplus \cdots \oplus \mathscr{C}_{n-1}),$$

where $\ominus$ denotes the orthogonal complementation with respect to the canonical scalar product on $L^2(\mathbf{P}_{\theta_0})$.

DEFINITION 4.1. A marked point process of $\mathbf{P}_{\theta_0}$ compensating measure ν_{θ_0} admits a chaos decomposition if and only if

$$L^2(\mathbf{P}_{\theta_0}) = \bigoplus_{n \geq 0} \mathscr{C}_n.$$

It is a challenging (and open) question to characterize processes which admit a chaos decomposition. Known results indicate that this property holds for Poisson processes and for Markov chains whose state space is a discrete group such that the jumps (differences between two consecutive states) can take only a finite number of values; see Biane (1989). We no longer need Hypotheses 2 or 4, but we now need another hypotheses:

HYPOTHESIS 5. *We have*

$$L^2(\mathbf{P}_{\theta_0}) = \bigoplus_{n \geq 0} \mathscr{C}_n.$$

In this case, each square-integrable functional F can be written

$$F = \mathbf{E}_0[F] + \sum_{n \geq 1} I_n(f_n), \tag{22}$$

where $f_n \in L^2_d(\mathbf{P}_0 \otimes \nu_{\theta_0}^{\otimes n})$ and the series converges in $L^2(\mathbf{P}_0)$.

DEFINITION 4.2. We denote by Dom $\hat{D}$ the subset of $L^2(\mathbf{P}_0)$ of functionals $F = \sum_{n=0}^{+\infty} I_n(f_n)$ such that the series

$$\sum_n n^2 \mathbf{E}_0 \left[\iint_{[0, +\infty]^n \times E^n} f_n^2(s_1, z_1, \dots, s_n, z_n) \bigotimes_{i=1}^{n} \nu_{\theta_0}(\omega, ds_i, dz_i) \right] \tag{23}$$

converges. For $F \in$ Dom $\hat{D}$, we define $\hat{D}F(\omega)$, the $L^2_p(\mathbf{P}_0 \otimes \nu_{\theta_0})$ process:

$$\hat{D}F(\omega): (s, z) \mapsto \hat{D}_{s,z} F(\omega) = \sum_{n \geq 1} n I_{n-1}(f_n(\cdot, s, z)).$$

PROPOSITION 4. *Let $\mathscr{M}$ be defined by*

$$\begin{aligned} \mathscr{M} \stackrel{\text{def}}{=} \bigg\{ &\mathscr{E}\left(\int_0^T \int_E h(s,z)(\omega - \nu_{\theta_0})(ds, dz) \right), \\ &h \text{ such that } \frac{d}{dt} \int_0^t \int_E h(s,z)^2 \nu_{\theta_0}(ds, dz) < c, \mathbf{P}_0\text{-}a.e. \bigg\} \end{aligned} \tag{24}$$

Every element F of $\mathscr{M}$ satisfies (23) *and*

$$\hat{D}_{s,z} F(\omega) = F(\omega) h(s,z). \tag{25}$$

PROOF. From Ruiz de Chavez (1983), we have

$$\mathscr{E}\left(\int_0^T \int_E h(s,z)(\omega - \nu_{\theta_0})(ds, dz) \right) = 1 + \sum_{n \geq 1} \frac{1}{n!} I_n(h^{\otimes n});$$

hence,

$$\hat{D}_{s,z}\mathscr{E}\left(\int_0^T\int_E h(s,z)(\omega-\nu_{\theta_0})(ds,dz)\right)$$
$$=\sum_{n\geq 1}\frac{n}{n!}I_{n-1}(h^{\otimes n-1})h(s,z)$$
$$=\mathscr{E}\left(\int_0^T\int_E h(s,z)(\omega-\nu_{\theta_0})(ds,dz)\right)\cdot h(s,z). \qquad \square$$

DEFINITION 4.3. For any $F\in\operatorname{Dom}\hat{D}$ and any predictable process $h\in L^2_p(\mathbf{P}_0\otimes\nu_{\theta_0})$, let

$$\hat{D}F(h)(\omega)\overset{\text{def}}{=}\int_0^T\int_E\hat{D}_{s,z}F(\omega)h(\omega,s,z)\nu_{\theta_0}(ds,dz).$$

Then we have the following theorem:

THEOREM 7. *For any* $F\in\operatorname{Dom}\hat{D}$ *and any predictable process* $h\in L^2_p(\mathbf{P}_0\otimes\nu_{\theta_0})$,

$$\mathbf{E}_0\big[\hat{D}F(h)\big]=\mathbf{E}_0\left[F\int_0^T\int_E h(\omega,s,z)(\omega-\nu_{\theta_0})(ds,dz)\right].$$

PROOF. For any $t\geq 0$, set $S_n^t=S_n\cap[0,t]^n$ and define

$$I_n^t(f^{\otimes(n)})\overset{\text{def}}{=}\iint_{S_n^t\times E^n}\prod_{i=1}^n f(s_i,z_i)\bigotimes_{i=1}^n(\omega-\nu_{\theta_0})(ds_i,dz_i).$$

For any F given by

$$F=\mathscr{E}\left(\int_0^T\int_E f(s,z)(\omega-\nu_{\theta_0})(ds,dz)\right),$$

we have

$$\mathbf{E}_0\left[\int_0^T\int_E\hat{D}_{s,z}F(\omega)h(\omega,s,z)\nu_{\theta_0}(ds,dz)\right]$$
$$=\sum_{n=1}^{+\infty}\frac{n}{n!}\mathbf{E}_0\left[\int_0^T\int_E f(s,z)h(\omega,s,z)I_n^s(f^{\otimes(n-1)})\nu_{\theta_0}(ds,dz)\right]$$
$$=\sum_{n=1}^{+\infty}\frac{1}{(n-1)!}\mathbf{E}_{\theta_0}\left[\int_0^T\int_E f(s,z)I_n^s(f^{\otimes(n-1)})(\omega-\nu_{\theta_0})(ds,dz)\right.$$
$$\left.\times\int_0^T\int_E h(\omega,s,z)(\omega-\nu_{\theta_0})(ds,dz)\right]$$
$$=\sum_{n=1}^{+\infty}\mathbf{E}_0\left[I_n(f^{\otimes n})\cdot\int_0^T\int_E h(\omega,s,z)(\omega-\nu_{\theta_0})(ds,dz)\right]$$
$$=\mathbf{E}_0\left[F\int_0^T\int_E h(\omega,s,z)(\omega-\nu_{\theta_0})(ds,dz)\right].$$

Hence the result holds for F in $\mathscr{M}$ and by density it also holds for $F \in \operatorname{Dom} \hat{D}$. □

DEFINITION 4.4. For any square-integrable functional F, define the difference operator Δ by

$$\Delta_{s,z} F(\omega) \stackrel{\text{def}}{=} F(\omega + \omega_{s,z}) - F(\omega),$$

for any s, z, where $\omega + \delta_{s,z}$ is the measure ω plus a jump at time s of mark z.

THEOREM 8. *Let $\mathbf{P}_0$ satisfy Hypothesis 5. The relation*

$$\mathbf{E}_0\left[\int_0^T \int_E (\Delta_{s,z} F) h(\omega, s, z) \nu_{\theta_0}(ds, dz)\right] = \mathbf{E}_0\left[F \cdot \int_0^T \int_E h(\omega, s, z)(\omega - \nu_{\theta_0})(ds, dz)\right]$$

holds for any square-integrable F and any predictable $h \in L_p^2(\mathbf{P}_0 \otimes \nu_{\theta_0})$ if and only if ν_{θ_0} is deterministic; that is, ω is a compound Poisson process.

PROOF. If ν_{θ_0} is deterministic, the result follows from Nualart and Vives (1988). In the converse direction, let f be in $L_d^2(\mathbf{P}_0 \otimes \nu_{\theta_0})$. Then we have

$$\begin{aligned}
&\mathbf{E}_0\left[\int_0^T \int_E \Delta_{s,z}\left(\int_0^T \int_E f(t,v)(\omega - \nu_{\theta_0})(dt, dv)\right) g(s,z) \nu_{\theta_0}(ds, dz)\right] \\
&\quad = \mathbf{E}_0\left[\int_0^T \int_E f(t,v)(\omega - \nu_{\theta_0})(dt, dv) \cdot \int_0^T \int_E g(s,z)(\omega - \nu_{\theta_0})(ds, dz)\right] \\
&\quad = \mathbf{E}_0\left[\int_0^T \int_E f(s,z) g(s,z) \nu_{\theta_0}(ds, dz)\right]
\end{aligned}$$

for any $g \in L_d^2(\mathbf{P}_0 \otimes \nu_{\theta_0})$. It follows by identification that

$$\Delta_{s,z} \int_0^T \int_E f(t,v)(\omega - \nu_{\theta_0})(dt, dv) = f(s,z).$$

Since

$$\Delta_{s,z} \int_0^T \int_E f(t,v) \omega(dt, dv) = f(s,z),$$

we obtain

$$\Delta_{s,z}\left(\int_0^T \int_E f(t,v) \nu_{\theta_0}(dt, dv)\right) = 0 \quad \text{for any } s, z.$$

As a consequence, for any g predictable in $L_p^2(\mathbf{P}_0 \otimes \nu_{\theta_0})$, we have

$$\begin{aligned}
0 &= \mathbf{E}_0\left[\int_0^T \int_E \Delta_{s,z} \int_0^T \int_E f(t,v) \nu_{\theta_0}(dt, dv) \cdot g(s,z) \nu_{\theta_0}(ds, dz)\right] \\
&= \mathbf{E}_0\left[\int_0^T \int_E f(t,v) \nu_{\theta_0}(dt, dv) \int_0^T \int_E g(s,z)(\omega - \nu_{\theta_0})(ds, dz)\right].
\end{aligned}$$

By the chaos decomposition property, we know that any square-integrable functional F can be written

$$F = \mathbf{E}_0[F] + \int_0^T \int_E g(s,z)(\omega - \nu_{\theta_0})(dt, dv),$$

where g is predictable and belongs to $L^2_p(\mathbf{P}_0 \otimes \nu_{\theta_0})$. Hence, for any $F \in L^2(\mathbf{P}_0)$,

$$\mathbf{E}_0\left[F \int_0^T \int_E f(t,v)\,\nu_{\theta_0}(dt,dv)\right] = \mathbf{E}_0[F]\mathbf{E}_0\left[\int_0^T \int_E f(t,v)\,\nu_{\theta_0}(dt,dv)\right].$$

Since f belongs to $L^2_d(\mathbf{P}_0 \otimes \nu_{\theta_0})$, $\int_0^T \int_E f(t,v)\nu_{\theta_0}(dt,dv)$ belongs to $L^2(\mathbf{P}_0)$ and it follows that the variance of $\int_0^T \int_E f(t,v)\nu_{\theta_0}(dt,dv)$ is zero for any $f \in L^2_p(\nu_{\theta_0})$; thus ν_{θ_0} is deterministic. □

EXAMPLE 2 (Continued). We keep the framework of Example 2 except that we now work on $[0,t]$ and the functional F is the virtual waiting time at time t denoted by W_t. We also assume that the $\mathbf{P}_0$ compensating measure is deterministic so that we can apply all the previous considerations. By (8) and Theorem 8, we know that

$$\left(\frac{d}{d\theta}\mathbf{E}_0[W_t]\right)_{\theta=\theta_0} = \mathbf{E}_0\left[\int_0^T \int_E \Delta_{s,z} W_t\, f(s) g(\theta_0, z)\, ds\, dz\right].$$

When we add a jump at time s and mark z to the nominal path, $W_t(\omega)$ is increased by $(z - \int_s^t \mathbf{1}_{\{W_u(\omega)=0\}}\, du)^+$; hence,

$$\left(\frac{d}{d\theta}\mathbf{E}_0[W_t]\right)_{\theta=\theta_0} = \mathbf{E}_0\left[\int_0^T \int_E \left(z - \int_s^t \mathbf{1}_{\{W_u=0\}}\, du\right)^+ \frac{\partial g}{\partial \theta}(\theta_0, z)\, ds\, dz\right].$$

5. Related works.

5.1. *Palm–Khinchin expansions.* When ν_{θ_0} is deterministic, it follows from Theorem 8 that, for any $F \in \hat{\mathbb{D}}_{2,1}$ and $\mathscr{F}_1$-measurable,

$$\left(\frac{d}{d\theta}\mathbf{E}_\theta[F]\right)_{\theta=\theta_0} = \mathbf{E}_0\left[\int_0^1 \int_E (F(\omega + \delta_{s,z}) - F(\omega)) h(s,z)\, \nu_{\theta_0}(ds,dz)\right]. \tag{26}$$

This formula was obtained in Baccelli, Klein and Zuyev (1995), Moller and Zuyev (1996) and Zuyev (1993). We denote by $\omega_{|s}$ the measure coinciding with ω up to time s and with no atoms after. Whenever F belongs to span $\mathscr{M}$, F is continuous in 0 in the sense that $\lim_{s\to 0} F(\omega_{|s})$ exists and is independent of the particular representation of F: we denote by $F(0)$ this limit. Formally, take ω to be the null path in the chaos expansion of an element of span $\mathscr{M}$. The random part in each multidimensional integral vanishes and we only keep

$$\mathbf{E}_0[F] = F(0) + \sum_{n \ge 1} (-1)^{n-1} \iint_{S_n \times E^n} \left(\hat{D}_{s_1,z_1} \cdots \hat{D}_{s_n,z_n} F\right)(0) \bigotimes_{i=1}^n \nu_{\theta_0}(ds_i, dz_i).$$

This formula is very similar to the factorial moment expansion of Blaszczyszyn (1995), which is itself an extension of the Palm–Khinchin formula. More precisely, the terms in both formulas are the same, but our approach does not give the convergence of the series—we have here only an L^2 convergence where Blaszczyszyn obtains a pointwise convergence (with the additional hypothesis of stationarity).

5.2. *Rare perturbation analysis*. On the other hand, (26) is also the common point to rare perturbation analysis and to the other methods of perturbation analysis. As mentioned previously, RPA consists of perturbing the nominal path by a sequence of decreasing thinning:

$$\left(\frac{d}{d\theta}\mathbf{E}_\theta[F]\right)_{\theta=\theta_0} = \theta_0^{-1}\lim_{p\to 1}\frac{1}{1-p}\mathbf{E}_{\theta_0,p}\big[F(\omega)-F(\omega_p)\big].$$

We now limit our considerations to the case of a Poisson process on the time interval $[0,1]$, of intensity θ_0, with independent and identically distributed marks independent of the jump times. We have

$$\mathbf{E}_{\theta_0,p}\big[F(\omega)-F(\omega_p)\big] = \sum_{j=0}^{\infty}\mathbf{E}_{\theta_0,p}\Big[\big(F(\omega)-F(\omega_p)\big)\mathbf{1}_{\{|\omega-\omega_p|=j\}}\Big].$$

Conditionally to $|\omega|$, the random variable $|\omega-\omega_p|$ is binomially distributed with parameters $(|\omega|, 1-p)$, so that, by the Cauchy–Schwarz inequality, we see that all the terms with j greater than 2 vanish when we take the limit. Hence,

$$\left(\frac{d}{d\theta}\mathbf{E}_\theta[F]\right)_{\theta=\theta_0} = \theta_0^{-1}\lim_{p\to 1}\frac{1}{1-p}\mathbf{E}_{\theta_0,p}\Big[\big(F(\omega)-F(\omega_p)\big)\mathbf{1}_{\{|\omega|=|\omega_p|+1\}}\Big].$$

Moreover, a compound Poission process of the type we are dealing with can be written as the superposition of two independent compound Poisson processes of the same type with respective intensities $\theta_0 p$ and $\theta_0(1-p)$. Hence, we can write

$$\begin{aligned}&\mathbf{E}_{\theta_0,p}\Big[\big(F(\omega)-F(\omega_p)\big)\mathbf{1}_{\{|\omega|=|\omega_p|+1\}}\Big]\\ &\qquad=\int_{\Omega\times\Omega}\big(F(\omega_1+\omega_2)-F(\omega_1)\big)\mathbf{1}_{\{|\omega_2|=1\}}\,d\mathbf{P}_{\theta_0 p}(\omega_1)\,d\mathbf{P}_{\theta_0(1-p)}(\omega_2).\end{aligned}$$

It is known that conditionally to $\{|\omega_2|=1\}$, the distribution of the jump time of ω_2 is uniform over $[0,1]$; hence,

$$\begin{aligned}&\mathbf{E}_{\theta_0,p}\Big[\big(F(\omega)-F(\omega_p)\big)\mathbf{1}_{\{|\omega|=|\omega_p|+1\}}\Big]\\ &\qquad=\theta_0^{-1}\mathbf{P}_{(1-p)\theta_0}(|\omega_2|=1)\\ &\qquad\quad\times\int_\Omega\left(\int_0^1\int_E\big(F(\omega_1+\delta_{s,z})-F(\omega_1)\big)\nu_{\theta_0}(ds,dz)\right)d\mathbf{P}_{\theta_0 p}(\omega_1).\end{aligned}$$

Thus, when we take the limit, we get

$$\theta_0^{-1} \lim_{p \to 1} \frac{1}{1-p} \mathbf{E}_{\theta_0, p}\big[F(\omega) - F(\omega_p)\big]$$
$$= \theta_0^{-1} \mathbf{E}_{\theta_0, p}\left[\int_0^1 \int_E (F(\omega_1 + \delta_{s,z}) - F(\omega_1)) \nu_{\theta_0}(ds, dz)\right].$$

We then observe that this latter term is nothing but the right-hand side of formula (26), since here $h \equiv \theta_0^{-1}$.

Acknowledgments. The author thanks Laure Coutin and an anonymous referee who helped improve an earlier version of this paper. We also benefited from useful discussions with A. S. Üstünel.

REFERENCES

BACCELLI, F., KLEIN, M. and ZUYEV, S. (1995). Perturbation analysis of functionals of random measures. *Adv. in Appl. Probab.* **27** 306–325.

BASS, R. F. and CRANSTON, M. (1986). The Malliavin calculus for pure jump processes and applications to local time. *Ann. Probab.* **14** 490–532.

BIANE, P. (1989). Chaotic representation for finite Markov chains. *Stochastics Stochastics Rep.* **30** 61–68.

BICHTELER, K. and JACOD, J. (1983). Calcul de Malliavin pour les diffusions avec sauts: existence d'une densité dans le cas unidimensionnel. *Séminaire de Probabilités XVII* 132–157. Springer, New York.

BISMUT, J.-M. (1983). Calcul des variations stochastiques et processus de sauts. *Zeitschrift für Wahrscheinlichkeits Theory* **63** 147–235.

BLASZCZYSZYN, B. (1995). Factorial moment expansion for stochastic systems. *Stochastic Process. Appl.* **56**(2) 321–336.

BRÉMAUD, P. (1992). Maximal coupling and rare perturbation sensitivity analysis. *Queuing Systems Theory Appl.* **11**(4) 307–333.

BRÉMAUD, P. and VÁZQUEZ-ABAD, F. J. (1992). On the pathwise computation of derivatives with respect to the rate of a point process: the phantom RPA method. *Queuing Systems Theory Appl.* **10**(3) 249–269.

DECREUSEFOND, L. (1994). Méthode de perturbation pour les réseaux de files d'attente. Ph.D. dissertation, École Nationale Supérieure des Télécommunications.

DEVETSIKIOTIS, M., WAEL, A. A.-Q., FREEBERSYSER, J. A. and TOWNSEND, J. K. (1993). Stochastic gradient techniques for the efficient simulation of high-speed networks using importance sampling. *IEEE Trans. Comm.* **XX** 751–756.

GLASSERMAN, P. (1990). *Gradient Estimation via Perturbation Analysis*. Kluwer, Dordrecht.

GLYNN, P. W. (1987). Likelihood ratio gradient simulation: an overview. In *Proceedings of the 1987 Winter Simulation Conference* (A. Thesen, H. Grant and W. Kelton, eds.).

HEIDELBERGER, P. (1987). Limitations of infinitesimal perturbation analysis. Technical Report, Bell Labs, Yorktown Heights, NY.

HO, Y. C. and CAO, X. R. (1983). Perturbation analysis and optimization of queueing networks. *J. Optim. Theory Appl.* **40** 559–582.

HO, Y. C., CAO, X. R. and CASSANDRAS, C. (1983). Infinitesimal and finite perturbation analysis for queueing networks. *Automatica* **19** 439–445.

JACOD, J. (1979). *Calcul Stochastique et Problèmes de Martingales*. Springer, Berlin.

L'ECUYER, P. (1990). A unified view of the IPA, SF and LR gradient estimation techniques. *Management Science* **36** 1364–1383.

MALLIAVIN, P. (1978). Stochastic calculus of variations and hypoelliptic operators. In *Proceedings of the International Symposium on S.D.E.* (K. Itô, ed.).

MOLLER, J. and ZUYEV, S. A. (1996). Gamma-type results and other related properties of Poisson processes. *Adv. in Appl. Probab.* **28** 662–673.

NORRIS, J. R. (1987). Integration for jump processes. *Séminaire de Probabilités XXII* 271–315. Springer, Berlin.

NUALART, D. and VIVES, J. (1988). Anticipative calculus for the Poisson process based on the Fock space. *Séminaire de Probabilités XXIV* 154–165. Springer, Berlin.

PRIVAULT, N. (1994). Chaotic and variational calculus in discrete and continuous time for the Poisson process. *Stochastics Stochastics Rep.* **51** 83–109.

REIMAN, M. I. and WEISS, A. (1989a). Light traffic derivatives via likelihood ratios. *IEEE Trans. Inform. Theory* **35**(3).

REIMAN, M. I. and WEISS, A. (1989b). Sensitivity analysis via likelihood ratios. *Oper. Res.* **37** 830–844.

RUIZ DE CHAVEZ, J. (1983). Sur les intégrales stochastiques multiples. *Séminaire de Probabilités XIX* 248–261. Springer, Berlin.

SURI, R. (1989). Perturbation analysis: The state of the art and research issues explained via the GI/G/1 queue. *Proc. IEEE* **77** 114–137.

SURI, R. and ZAZANIS, M. (1988). Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queues. *Management Science* **34** 39–64.

ÜSTÜNEL, A. S. (1995). *An Introduction to Analysis on Wiener Space. Lecture Notes in Math.* **1610**. Springer, Berlin.

ZUYEV, S. A. (1993). Russo's formula for Poisson point fields and its applications. *Discrete Math. Appl.* **3** 63–73.

NETWORKS DEPARTMENT
ÉCOLE NATIONALE SUPÉRIEURE DES TÉLÉCOMMUNICATIONS
46, RUE BARRAULT
75 634 PARIS CEDEX 13
FRANCE
E-MAIL: laurent.decreusefond@enst.fr