ON RICE'S FORMULA FOR STATIONARY MULTIVARIATE PIECEWISE SMOOTH PROCESSES

K. BOROVKOV,* *University of Melbourne*

G. LAST,** *Karlsruher Institut für Technologie*

Abstract

Let $X = \{X_t : t \geq 0\}$ be a stationary piecewise continuous $\mathbb{R}^d$-valued process that moves between jumps along the integral curves of a given continuous vector field, and let $S \subset \mathbb{R}^d$ be a smooth surface. The aim of this paper is to derive a multivariate version of Rice's formula, relating the intensity of the point process of (localized) continuous crossings of S by X to the distribution of X_0. Our result is illustrated by examples relating to queueing networks and stress release network models.

Keywords: Level crossing; Rice's formula; stationarity; Palm probability; piecewise-deterministic process; stochastic network

2010 Mathematics Subject Classification: Primary 60J75
Secondary 60G55

1. Introduction

The classical Rice formula going back to [11] gives the intensity $\nu(u)$ of crossings (originally, upcrossings) of a given level u by a univariate continuous stationary Gaussian process X_t in terms of the joint distribution of $(X_t, X_t') \overset{\mathrm{D}}{=} (X_0, X_0')$, the process' value and its derivative at a fixed time (provided that the derivative exists in some suitable sense, e.g. in mean quadratic):

$$\nu(u) = \int |z| p(u, z) \, \mathrm{d}z, \tag{1.1}$$

where $p(\cdot, \cdot)$ is the joint density (X_0, X_0') which is assumed to exist. Later, the result was extended to more general classes of differentiable (in some suitable sense) stationary processes, covering not only the first moments but also higher-order factorial moments of the numbers of crossing, and even to more general settings for continuous random processes and fields. The formula proved to be quite useful in a number of applied areas, including signal processing, reliability, and sea waves. For detailed accounts of the history of results of this kind and further bibliography, the interested reader is referred to [12], [10], and Chapter 3 of [2].

The case of processes with jumps and degenerate finite-dimensional distributions has received far less attention, although, from an applications viewpoint, it is scarcely less interesting than the case of continuous processes. However, the heuristics behind the formula based on 'Kac's counting formula' giving the number of crossings of a level u by a C^1-function

f on [0, 1] as

$$\lim_{\delta \to 0+} \frac{1}{2\delta} \int_0^1 |f'(t)| \mathbf{1}\{|f(t) - u| < \delta\} \, \mathrm{d}t \tag{1.2}$$

(under a couple of further technical assumptions and denoting by $\mathbf{1}A \equiv \mathbf{1}_A$ the indicator function of the set A) seems to be applicable in that case as well, provided that the process jumps at finite intensity and is smooth between the jump times. We note that (1.2) is a consequence of Federer's coarea theorem (see, e.g. (7.4.15) of [1])

$$\int_0^1 g(t) |f'(t)| \, \mathrm{d}t = \int_{-\infty}^{\infty} \sum_{s \in [0,1]} g(s) \mathbf{1}\{f(s) = t\} \, \mathrm{d}t,$$

applied to the function $g(s) := \mathbf{1}\{|f(s) - u| < \delta\}$.

An analogue of (1.1) for the intensity $\nu_c(u)$ of continuous level crossings (i.e. crossings that are not induced by a jump) by general univariate piecewise-deterministic Markov processes that has the form

$$\nu_c(u) = |\mu(u)| p(u),$$

where $\mu(\cdot)$ is the drift coefficient of the process and $p(\cdot)$ is the density of X_0, was established in [6] (see also Theorem 3.2 below; it is worth mentioning here the earlier paper by Bar-David and Nemirovsky [4], who considered the case of Poisson shot noise processes). The result was used in [6] to obtain the asymptotic behaviour of the point processes of high-level crossings (i.e. as $u \to \infty$) in a number of interesting and special cases important for applications.

The proof in [6] relied on the Markov structure of the process and in fact did not assume the existence of the density p. The natural question of whether Rice's formula for piecewise-smooth processes can be extended to the multivariate and non-Markovian cases remained open. In the present paper we answer this question in the affirmative.

The paper is organized as follows. In Section 2 we describe the main class of processes we will be working with. In Section 3 we present the main result together with its proof. Section 4 provides examples to illustrate our main result.

2. Model description

First we will describe the main model of multivariate random processes $X = \{X_t \colon t \geq 0\}$ dealt with in this paper. The two key elements of the model are a point process $N = \{N(B) \colon B \in \mathcal{B}(\mathbb{R}_+)\}$ of jumps in our process X (here and in what follows, $\mathcal{B}(\cdot)$ denotes the class of Borel subsets of $\cdot$), and a vector field $\mu \colon D \to \mathbb{R}^d$ defined on an open domain $D \subset \mathbb{R}^d$ and specifying the dynamics of X between the jumps.

The point process N can be identified with the counting process $N_t := N([0, t])$, $t \geq 0$. We assume that N is stationary, simple (i.e. $N_t - N_{t-} \in \{0, 1\}$ for all $t > 0$ and $N_0 = 0$), and nonnull. Then $N(\mathbb{R}_+) = \infty$ almost surely and we can enumerate the (jump) points of N in increasing order. We denote them by $0 < T_1 < T_2 < \cdots$ and set $T_0 := 0$ for convenience (this is not a point of N). We also assume that N has a finite *intensity* $\lambda_N := \mathrm{E}\, N((0, 1])$.

For the vector field, we assume that $\mu \in C^1(D)$. This implies that there exist continuous functions $t_- \colon D \to (-\infty, 0)$ and $t_+ \colon D \to (0, \infty)$ such that, for any $x \in D$, there exists a unique C^1-function $q(x, \cdot) \colon (t_-(x), t_+(x)) \to D$ satisfying the integral equation

$$q(x, t) = x + \int_0^t \mu(q(x, s)) \, \mathrm{d}s, \qquad t \in (t_-(x), t_+(x)) \tag{2.1}$$

(Picard–Lindelöf theorem; see, e.g. [8, p. 8]). Moreover, for any fixed $x \in D$, there is a neighbourhood of $(x, 0) \in \mathbb{R}^{d+1}$ in which $q(\cdot, \cdot)$ will also be continuously differentiable (Peano's theorem on the dependence of initial conditions; see, e.g. [8, p. 95]). We also mention the *flow* property

$$q(x, s + t) = q(q(x, s), t),$$

which holds whenever one (and then also the other) side is defined.

Using the integral curves q, we can now specify the dynamics of the process X between its jumps. We assume that, for any $n \geq 0$, we have $X_{T_n} \in D$, $T_{n+1} < T_n + t_+(X_{T_n})$, and

$$X_t = q(X_{T_n}, t - T_n), \qquad T_n \leq t < T_{n+1}.$$

This means that, on any interval $[T_n, T_{n+1})$, $n \geq 0$, X follows the smooth integral curve generated by the (time-independent!) drift μ. In particular, a jump epoch $t > 0$ of X (that is, $X_{t-} \neq X_t$) must be a point of N. It is not assumed, however, that any T_n is a jump epoch of X.

Our key probabilistic assumption is that X and N are *jointly stationary*, i.e. the distribution of the bivariate process $\{(X_{s+t}, N((s, s+t])): t \geq 0\}$ does not depend on $s \geq 0$. This means that (X, N) can be extended from $[0, \infty)$ to a stationary pair on $\mathbb{R}$ in the sense of [9, Chapter 11]. Also equivalent is the existence of a stationary marked point process $((T_n^*, Y_n^*))_{n \in \mathbb{Z}}$ (see, e.g. [3] for this notion and note that $T_0^* \leq 0 < T_1^*$) such that the sequence $q(Y_0^*, -T_0^*), T_1^*, Y_1^*, T_2^*, Y_2^*, \ldots$ has the same distribution as the sequence $X_0, T_1, X_{T_1}, T_2, X_{T_1}, \ldots$. Apart from this stationarity, the joint distribution of X and N may exhibit any sort of dependence. In particular, we stress that we do not make Markovian assumptions of any kind.

Our next aim is to introduce continuous crossings of a surface. We assume that $S \subset D$ is the relative interior of a $(d-1)$-dimensional (not necessarily connected) C^1-*manifold* with or without a boundary, and that $\{n(x): x \in S\}$ is a continuous field of unit normals to S. Denoting by $\langle \cdot, \cdot \rangle$ the Euclidean scalar product in $\mathbb{R}^d$, we further assume that the trajectories of the flow are not tangent to S, i.e.

$$\langle n(x), \mu(x) \rangle \neq 0, \qquad x \in S. \tag{2.2}$$

We say that X has a *continuous crossing* of S at time $s > 0$ if $X_{s-} = X_s \in S$ and there is a $\delta > 0$ such that $X_t \notin S$ for $t \in (s - \delta, s + \delta) \setminus \{s\}$.

Remark 2.1. Let $\tau_x := \inf\{t > 0: q(x, t) \in S\}$ be the first positive time the integral curve of μ leaving from x at time 0 hits the surface S. It is not hard to see that if S' is a compact subset of S then, from (2.2) and the C^1-assumption on the drift μ, it follows that

$$\inf\{\tau_x: x \in S'\} > 0. \tag{2.3}$$

This (together with the fact that X jumps only finitely often in finite time intervals) implies that the times of continuous crossings of S through a compact subset of S cannot accumulate in finite time.

The times of continuous crossings of S by X form an at most countable set N_c that will be identified with a random counting measure on $[0, \infty)$. Then

$$\Phi_c(C) := \sum_{s \in N_c} \mathbf{1}\{(s, X_s) \in C\} = \int \mathbf{1}\{(s, X_s) \in C\} N_c(\mathrm{d}s), \qquad C \in \mathcal{B}([0, \infty) \times \mathbb{R}^d),$$

defines a random (integer-valued) measure Φ_c on $[0, \infty) \times \mathbb{R}^d$. For $t \geq 0$ and $B \in \mathcal{B}(\mathbb{R}^d)$, the random variable $\Phi_c([0, t] \times B)$ does not need to be finite. However, if B is a compact subset

of S then (2.3) implies that $\Phi_c([0, t] \times B) < \infty$. Moreover, since N has a finite intensity,

$$\nu_c(B) := \mathrm{E}\, \Phi_c((0, 1] \times B), \qquad B \in \mathcal{B}(\mathbb{R}^d),$$

is finite, whenever B is a compact subset of S. Therefore, $\nu_c(\cdot)$ is a σ-finite measure on $\mathcal{B}(\mathbb{R}^d)$. For any compact $B \subset S$, the point process $\Phi_c(\cdot \times B)$ is stationary. This is enough to derive the (refined) *Campbell theorem*, which states that

$$\mathrm{E} \int g(s, X_s) N_c(\mathrm{d}s) = \iint g(s, x)\, \mathrm{d}s\, \nu_c(\mathrm{d}x) \tag{2.4}$$

for any measurable function $g \colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$; cf., e.g. Equation (1.2.19) of [3].

Remark 2.2. Assuming that A is a small enough open set to ensure that $\nu_c(S \cap A) < \infty$, observe that $\nu_c(S \cap A)^{-1}\nu_c(\cdot \cap S \cap A)$ can be interpreted as the distribution of the value of X at a *typical* time of a continuous crossing of $S \cap A$. This is a particular instance of a Palm distribution. (Ergodicity is not required here.) We refer the reader to [3] for more details on this standard concept of applied probability.

3. The main result

In this section we consider the process $X = \{X_t \colon t \geq 0\}$ and the surface S as introduced in the previous section. Our main result requires the following assumption.

(A1) The distribution π of X_0 has a continuous density p in a neighbourhood of S.

To introduce a second assumption, we define the following measure π_0 on $\mathbb{R}^d \times \mathbb{R}^d$:

$$\pi_0(B) := \mathrm{E} \sum_{n=1}^{\infty} \mathbf{1}\{T_n \leq 1,\, X_{T_n-} \neq X_{T_n},\, (X_{T_n-}, X_{T_n}) \in B\}, \qquad B \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d).$$

Note that $\pi_0(\mathbb{R}^d \times \mathbb{R}^d) \leq \lambda_N < \infty$. The normalized version of π_0 can be interpreted as the distribution of the values of X just before and after a 'typical jump' of X; see again [3] for more details on Palm distributions.

(A2) The measure π_0 satisfies

$$\min\{\pi_0((\mathbb{R}^d \setminus S) \times S),\, \pi_0(S \times (\mathbb{R}^d \setminus S))\} = 0. \tag{3.1}$$

If (3.1) holds then at least one of the following two conditions is met: (i) the process does not jump to the surface S from elsewhere, (ii) the process does not jump from S to its complement (almost surely in both cases).

We now present the main result of our paper.

Theorem 3.1. *Assume that (A1) and (A2) hold. Then*

$$\nu_c(B) = \int_{S \cap B} |\langle n(x), \mu(x) \rangle| p(x) \mathcal{H}^{d-1}(\mathrm{d}x), \qquad B \in \mathcal{B}(\mathbb{R}^d), \tag{3.2}$$

where $\mathcal{H}^{d-1}$ is the normalized $(d-1)$-dimensional Hausdorff measure on $\mathbb{R}^d$.

Remark 3.1. As will be seen from the first half of the proof of Theorem 3.1, the full continuity assumption (A1) on p can actually be somewhat weakened to the boundedness of p in a neighbourhood of S and its right continuity (in case the first number in (3.1) vanishes) on S along the flow, meaning that $p(q(x, 0+)) = p(x)$, $x \in S$.

In the one-dimensional case the above theorem simplifies to the following assertion.

Theorem 3.2. *In the case $d = 1$, assuming that $S = \{u\}$ for some $u \in D$ such that $\mu(u) \neq 0$, and that (A1) and (A2) are satisfied, we have*

$$\nu_c(\{u\}) = |\mu(u)| p(u). \tag{3.3}$$

Remark 3.2. In the case when X is a stationary univariate Markovian process, representation (3.3) was established under broad conditions in [6]. More precisely, it was actually shown that the stationary distribution π of X has a density p that satisfies (3.3). Note that, due to the Markovian structure of the process, it was possible to derive the result under weaker technical assumptions than those imposed in the general setting of the present paper.

In the case when

$$S = S^u := \{x \in \mathbb{R}^d : x_1 = u\}$$

for some $u \in \mathbb{R}$, a continuous crossing of S is a continuous crossing of the level u by the first component of X. In this case Theorem 3.1 takes the following form.

Theorem 3.3. *Let $u \in \mathbb{R}$ be such that $S^u \subset D$, and let $\mu_1(x) \neq 0$ for all $x \in S^u$, where μ_1 is the first component of μ. Assume that (A1) and (A2) hold with $S = S^u$. Then, for $B \in \mathcal{B}(\mathbb{R}^d)$,*

$$\nu_c(B) = \int \cdots \int \mathbf{1}_B(u, x_2, \ldots, x_d) |\mu_1(u, x_2, \ldots, x_d)| p(u, x_2, \ldots, x_d) \, dx_2 \cdots dx_d.$$

Remark 3.3. Theorem 3.3 is another, more straightforward, generalization of (3.3). Assume now that $0 < \nu_c(S^u) < \infty$ and consider a 'typical time' of a continuous crossing of the level u by the first component of X. Then the measure $\mathbb{Q}_u(\cdot) := \nu_c(S^u)^{-1} \nu_c(\{u\} \times \cdot)$ describes the distribution of the other components of X at this time. This distribution can be interpreted in terms of the drift-modulated density p_1 proportional to $|\mu_1(x)| p(x)$ (assuming that $\mathrm{E}|\mu_1(X_0)| < \infty$). If $(Y_1, \ldots, Y_d)$ is a random vector with density p_1 then $\mathbb{Q}_u$ is the conditional distribution of $(Y_2, \ldots, Y_d)$ given that $Y_1 = u$.

Remark 3.4. Let $k \in \{1, \ldots, d\}$, and assume that $S = \tilde{S} \times \mathbb{R}^{d-k}$, where $\tilde{S} \subset \mathbb{R}^k$ is a $(k-1)$-dimensional smooth surface. Let $\{\tilde{n}(x) : x \in \tilde{S}\}$ be a continuous field of unit normals to $\tilde{S}$. Let $\tilde{X} := (X^{(1)}, \ldots, X^{(k)})$ and $Y := (X^{(k+1)}, \ldots, X^{(d)})$, where $X = (X^{(1)}, \ldots, X^{(d)})$. There is a one-to-one correspondence between the continuous crossings of S by the process X and the continuous crossings of $\tilde{S}$ by the process $\tilde{X}$. Equation (3.2) can be written as

$$\nu_c(B) = \int_{\mathbb{Y}} \int_{\tilde{S}} |\langle \tilde{n}(x), \tilde{\mu}(x, y) \rangle| \mathbf{1}_B(x, y) p(x \mid y)$$
$$\times \mathcal{H}^{k-1}(dx) \, \mathrm{P}(Y_0 \in dy), \qquad B \in \mathcal{B}(\mathbb{R}^k \times \mathbb{Y}),$$

where $\mathbb{Y} := \mathbb{R}^{d-k}$, $\tilde{\mu}(x, y)$ is the vector of the first k components of $\mu(x, y)$, and $x \mapsto p(x \mid y)$ is the conditional density of $\tilde{X}_0$ given that $Y_0 = y$. In this form the result might be generalizable to other stationary pairs $(\tilde{X}, Y)$. The process $\tilde{X}$ should remain piecewise deterministic for given Y. But the process Y might take values in a more general space $\mathbb{Y}$. In this paper we will make no attempt to establish such an extension of our results.

To prove Theorem 3.1, we will need an auxiliary result that requires some further notation. First of all, note that the assertion of the theorem will immediately follow if we prove representation (3.2) in a 'local setting', i.e. with S replaced by $S \cap A$, where A is a small enough open subset of $\mathbb{R}^d$. As can easily be seen from the observation that we made after (2.1) and from the manifold properties of S, if we understand by S a 'small piece' of the original surface then the following assumption will be satisfied.

(A3) The surface S is connected and relatively compact, (2.3) holds with $S' = S$, and $\nu_c := \nu_c(S) < \infty$. Furthermore, there exists a $u_0 > 0$ such that $t_+(x) \geq u_0$ for all $x \in S$ and, for any $u \in [0, u_0]$,

$$S_u := \{q(x, u) \colon x \in S\}$$

is a C^1-surface with a continuous field $\{n_u(x) \colon x \in S_u\}$ of unit normals to it satisfying

$$\inf\{\langle n_u(x), \mu(x) \rangle \colon x \in S_u, \ u \in [0, u_0]\} > 0. \tag{3.4}$$

Moreover, π has a density p in a neighbourhood of $S_{(0,u_0)}$, where

$$S_I := \bigcup_{u \in I} S_u, \qquad I \subset \mathbb{R}.$$

Now denote by N_c^u the stationary point process of the times of all continuous crossings of S_u by X. For any $C \in \mathcal{B}([0, \infty) \times \mathbb{R}^d)$, let $\Phi_c^u(C)$ be the number of all $s \in N_c^u$ such that $(s, X_s) \in C$ and $\nu_c^u(B) := \mathrm{E} \, \Phi_c^u([0, 1] \times B)$, $B \in \mathcal{B}(\mathbb{R}^d)$.

Proposition 3.1. *Assume that (A3) holds. Then, for any measurable $g \colon \mathbb{R}^d \to \mathbb{R}_+$,*

$$\int g(x) \nu_c^u(\mathrm{d}x) = \int_{S_u} |\langle n_u(x), \mu(x) \rangle| g(x) p(x) \mathcal{H}^{d-1}(\mathrm{d}x) \tag{3.5}$$

for $\mathcal{H}^1$-almost all $u \in [0, u_0]$.

Proof. For any $j \geq 0$, set $T_j' := T_j \wedge 1$ and, in particular, $T_0' := 0$. For $j \geq 1$, we define

$$I_j := (T_{j-1}', T_j'), \qquad L_j := \{X_t \colon t \in I_j\} = \{q(X_{T_{j-1}'}, t - T_{j-1}') \colon t \in I_j\}.$$

Fix a $B \in \mathcal{B}(\mathbb{R}^d)$, and assume that $u \in (0, u_0)$. By definition, $\Phi_c^u(I_j \times B) > 0$ if and only if $L_j \cap S_u \cap B \neq \varnothing$. On the other hand, (3.4) implies that $\Phi_c^u(I_j \times B) \leq 1$, so that

$$\Phi_c^u(I_j \times B) = \mathbf{1}\{L_j \cap S_u \cap B \neq \varnothing\}.$$

Therefore,

$$\Phi_c^u((0, 1) \times B) = \sum_{j=1}^{\infty} \mathbf{1}\{L_j \cap S_u \cap B \neq \varnothing\}$$

and, for any $v \in (0, u_0)$,

$$\int_0^v \Phi_c^u((0, 1) \times B) \, \mathrm{d}u = \sum_{j=1}^{\infty} \int_0^v \mathbf{1}\{L_j \cap S_u \cap B \neq \varnothing\} \, \mathrm{d}u. \tag{3.6}$$

Now set

$$J_j(v) := \{t \in I_j : X_t \in S_{(0,v)}\}, \qquad U_j(v) := \{u \in (0, v) : L_j \cap S_u \neq \varnothing\}.$$

Clearly, the last two sets are either simultaneously empty or are open intervals of the same length; in the latter case, put $u_j(v) := \inf U_j(v)$. Therefore,

$$\begin{aligned}
\int_0^v \mathbf{1}\{L_j \cap S_u \cap B \neq \varnothing\}\, \mathrm{d}u &= \int_{U_j(v)} \mathbf{1}\{q(X_{u_j(v)}, u - u_j(v)) \in B\}\, \mathrm{d}u \\
&= \int_{J_j(v)} \mathbf{1}\{X_t \in B\}\, \mathrm{d}t \\
&= \int_{I_j(v)} \mathbf{1}\{X_t \in S_{(0,v)} \cap B\}\, \mathrm{d}t,
\end{aligned} \tag{3.7}$$

so that (3.6) becomes

$$\int_0^v \Phi_c^u((0,1) \times B)\, \mathrm{d}u = \int_0^1 \mathbf{1}\{X_t \in S_{(0,v)} \cap B\}\, \mathrm{d}t.$$

Taking expectations on both sides of the last relation and using Fubini's theorem and the stationarity of X, we obtain

$$\int_0^v v_c^u(B)\, \mathrm{d}u = \mathrm{E} \int_0^1 \mathbf{1}\{X_t \in S_{(0,v)} \cap B\}\, \mathrm{d}t = \mathrm{P}(X_0 \in S_{(0,v)} \cap B).$$

As functions of $B \in \mathcal{B}(\mathbb{R}^d)$, both sides specify a measure, and so the standard argument shows that, for any measurable function $g : \mathbb{R}^d \to \mathbb{R}_+$,

$$\int_0^v \mathrm{d}u \int g(x) v_c^u(\mathrm{d}x) = \int_{S_{(0,v)}} g(x) p(x) \mathcal{H}^d(\mathrm{d}x). \tag{3.8}$$

Now we can assume without loss of generality that S admits a C^1-parametrization $(w_1, \dots, w_{d-1}) \mapsto z(w_1, \dots, w_{d-1})$, where $(w_1, \dots, w_{d-1})$ varies in an open set $W \subset \mathbb{R}^{d-1}$. For $(w_1, \dots, w_{d-1}) \in W$ and $u \in [0, u_0]$, define

$$\psi(w_1, \dots, w_{d-1}, u) := q(z(w_1, \dots, w_{d-1}), u),$$

which, for a fixed $u \in [0, u_0]$, will be a C^1-parametrization of the 'parallel' surface S_u.

Next we denote by ∂_i the operator of partial differentiation with respect to w_i, $i = 1, \dots, d-1$, and let $\partial_d \psi := \partial \psi / \partial u$. A simple linear algebra calculation shows that the Jacobian $J\psi$ of $\psi = \psi(w_1, \dots, w_{d-1}, u)$ satisfies

$$|J\psi| = |\langle n_u(\psi), \partial_d \psi \rangle| \sqrt{H} \equiv |\langle n_u(\psi), \mu(\psi) \rangle| \sqrt{H},$$

where $H = H(w_1, \dots, w_{d-1}, u)$ is the determinant of the matrix $(\langle \partial_i \psi, \partial_j \psi \rangle)_{i,j=1,\dots,d-1}$. However, for any fixed $u \in [0, u_0]$, $\sqrt{H(w_1, \dots, w_{d-1}, u)}\, \mathrm{d}w_1 \cdots \mathrm{d}w_{d-1}$ is the surface element of S_u in the coordinates $(w_1, \dots, w_{d-1})$, so that changing coordinates on the right-hand

side of (3.8) yields

$$\int_{S_{(0,v)}} g(x)p(x)\mathcal{H}^d(\mathrm{d}x) = \int_{W\times(0,v)} g(\psi)p(\psi)|J\psi|\,\mathrm{d}w_1\cdots\mathrm{d}w_{d-1}\,\mathrm{d}u$$

$$= \int_{W\times(0,v)} g(\psi)p(\psi)|\langle n_u(\psi),\mu(\psi)\rangle|\sqrt{H}g\,\mathrm{d}w_1\cdots\mathrm{d}w_{d-1}\,\mathrm{d}u$$

$$= \int_0^v \mathrm{d}u \int_{S_u} g(x)p(x)|\langle n_u(x),\mu(x)\rangle|\mathcal{H}^{d-1}(\mathrm{d}x), \qquad (3.9)$$

which immediately implies the assertion of Proposition 3.1.

Remark 3.5. Assume that f is a real-valued C^1-function defined on an open domain $\tilde{D} \subset R^d$, with nonvanishing gradient and such that $S_u = \{x \in \tilde{D}: f(x) = u\}$ for all small enough u. Such a function exists, at least for suitably small pieces of S. We may then apply Federer's coarea theorem (see, e.g. Equation (7.4.15) of [1]) on each open interval (T'_{j-1}, T'_j) to the level sets of the function $t \mapsto f(X_t)$. While this would provide an alternative way for deriving (3.7), we opted to give the direct argument presented in the proof of Proposition 3.1. In a quite similar spirit the coarea theorem can be used to derive Rice's formula for smooth processes; see Section 11.4 of [1]. It was actually Zähle [13] who first used the coarea theorem to prove Rice's formula for certain continuous processes. Observe also that the coarea formula could be used to establish (3.9) as well. However, our more explicit argument yields additional information on the continuity properties of the surface element of S_u, which is needed in the proof of Theorem 3.1.

Proof of Theorem 3.1. Since both sides of (3.2) are σ-additive in B, it is no restriction of generality to assume that assumption (A3) is satisfied. Moreover, we can assume that S admits a smooth parametrization as in the proof of Proposition 3.1. This is due to the fact that the surface S can be represented as a 'mosaic' of 'small pieces' for which the assumption will be satisfied owing to our assumptions on the original S.

Furthermore, it is not hard to see that, to prove the theorem, it suffices to demonstrate that (3.5) holds at $u = 0$ for continuous and bounded g. We will show this by proving that, under the assumption that $\pi_0((\mathbb{R}^d \setminus S) \times S) = 0$, both sides of (3.5) are right continuous at $u = 0$, as Proposition 3.1 will then imply the desired result. The case when only the second term on the left-hand side of (3.1) is zero (i.e. $\pi_0(S \times (\mathbb{R}^d \setminus S)) = 0$) can be dealt with in exactly the same way by establishing the left continuity of both sides of (3.5) at $u = 0$ (essentially via a time-reversal argument).

Using the notation from the proof of Proposition 3.1 and setting

$$h_u(w) := \langle n_u(\psi), \mu(\psi)\rangle g(\psi)p(\psi), \qquad w = (w_1, \ldots, w_{d-1}), \qquad \psi = \psi(w, u),$$

we have, for $u \in [0, u_0]$,

$$\int_{S_u} \langle n_u(x), \mu(x)\rangle g(x)p(x)\mathcal{H}^{d-1}(\mathrm{d}x) = \int h_u(w)\sqrt{H(w,u)}\mathcal{H}^{d-1}(\mathrm{d}w). \qquad (3.10)$$

As noted above, we have $q \in C^1$, and so $n_u(\psi(y, u))$ is a continuous function of u, leading to

$$h_{0+}(w) = \langle n(z(w)), \mu(z(w))\rangle g(z(w))p(z(w)).$$

Similarly, as $u \to 0+$, $H(w, u)$ converges to the value of the determinant of the matrix $(\langle \partial_i z(w), \partial_j z(w) \rangle)_{i,j=1,\dots,d-1}$. Now the dominated convergence theorem implies that (3.10) converges to the right-hand side of (3.5) at $u = 0$.

To establish the desired right continuity of the left-hand side of (3.5), we assume that (3.4) holds. Introduce the following point process Φ_d on $\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d$:

$$\Phi_d(\cdot) := \sum_{n=1}^{\infty} \mathbf{1}\{X_{T_n-} \neq X_{T_n}\} \mathbf{1}\{(T_n, X_{T_n-}, X_{T_n}) \in \cdot\}.$$

Let $u \in [0, u_0]$ and $t \geq 0$. A continuous crossing of S_u can only occur on a trajectory of X that arrives at the surface from the *inside* of $S_{(0,u)}$ along an integral curve of μ (cf. (3.4)). Therefore, each such crossing of S_u should be preceded by an entry to $S_{[0,u)}$, either along a drift line or by a jump. Taking into account the possibility of having $X_0 \in S_{[0,u)}$, we obtain the bound

$$N_c^u([u, t+u]) \leq N_c([0, t+u]) + \Phi_d([0, t+u] \times (\mathbb{R}^d \setminus S_{[0,u]}) \times S_{[0,u]}) + 1.$$

Therefore,

$$\int_u^{t+u} g(X_s) N_c^u(\mathrm{d}s) \leq \int_0^{t+u} (g(X_s) + \varepsilon(u)) N_c(\mathrm{d}s)$$
$$+ g^*[\Phi_d([0, t+u] \times (\mathbb{R}^d \setminus S_{[0,u]}) \times S_{[0,u]}) + 1],$$

where $g^* := \sup_x g(x)$ and

$$\varepsilon(u) := \sup\{|g(q(x, v)) - g(x)| : x \in S, \, 0 \leq v \leq u\} \to 0 \quad \text{as } u \to 0+, \tag{3.11}$$

owing to the uniform continuity of the mapping $(x, u) \mapsto g(q(x, u))$ on $\bar{S} \times [0, u_0]$, $\bar{S}$ denoting the closure of S.

Now taking expectations on both sides of the obtained inequality and using Campbell's formula (2.4) yields

$$t \int g(x) \nu_c^u(\mathrm{d}x) \leq (t+u) \int (g(x) + \varepsilon(u)) \nu_c(\mathrm{d}x)$$
$$+ g^*((t+u)\pi_0((\mathbb{R}^d \setminus S_{[0,u]}) \times S_{[0,u]}) + 1),$$

where we also used Campbell's theorem for $\pi_0(\cdot) = \mathrm{E} \, \Phi_d([0, 1] \times \cdot)$. After dividing by t and letting $t \to \infty$, we obtain

$$\int g(x) \nu_c^u(\mathrm{d}x) \leq \int g(x) \nu_c(\mathrm{d}x) + \varepsilon(u)\nu_c(S) + g^* \pi_0((\mathbb{R}^d \setminus S_{[0,u]}) \times S_{[0,u]}).$$

In view of (3.11) and the fact that the assumption that $\pi_0((\mathbb{R}^d \setminus S) \times S) = 0$ implies that $\pi_0((\mathbb{R}^d \setminus S_{[0,u]}) \times S_{[0,u]}) \to 0$ as $u \to 0+$, this leads to

$$\limsup_{u \to 0+} \int g(x) \nu_c^u(\mathrm{d}x) \leq \int g(x) \nu_c(\mathrm{d}x).$$

To derive the converse inequality, we start with the observation that any continuous crossing of S in $[0, t]$ is followed either by a continuous crossing of S_u or by a jump from $S_{(0,u]}$ to its complement within the time interval $[0, t+u]$, so that

$$N_c^u([0, t+u]) \geq N_c([0, t]) - \Phi_d([0, t+u] \times S_{(0,u]} \times (\mathbb{R}^d \setminus S_{(0,u]})) - 1.$$

Next, similarly to our argument above, we obtain

$$\int g(x)\nu_c^u(\mathrm{d}x) \geq \int g(x)\nu_c(\mathrm{d}x) - \varepsilon(u)\nu_c(S) - g^*\pi_0(S_{(0,u]} \times (\mathbb{R}^d \setminus S_{(0,u]})).$$

Since $\lim_{u\to 0+} S_{(0,u]} = \varnothing$, it is clear that the continuity of π_0 implies now that

$$\liminf_{u\to 0+} \int g(x)\nu_c^u(\mathrm{d}x) \geq \int g(x)\nu_c(\mathrm{d}x),$$

which completes the proof of the theorem.

4. Examples

In this section we present a couple of examples showing possible uses of our main result. Deriving closed-form expressions for stationary distributions for nontrivial processes of the type we are dealing with in the paper is notoriously difficult. To give an analytically computable result, we will begin with a rather simple example. The other two examples are less trivial and, accordingly, require numerical evaluation of the quantities of interest.

Example 4.1. In this example we set $D := \mathbb{R}^d \setminus \{0\}$ and $\mu(x) := x/|x|$. We consider a divergent sequence $(T_n')_{n\geq 0}$ of strictly increasing times with $T_0' = 0$ and a sequence $(Y_n')_{n\geq 0}$ of random vectors with values on the unit sphere $\mathbb{S}^{d-1} := \{x \in \mathbb{R}^d : |x| = 1\}$. We assume that $((T_n', Y_n'))_{n\geq 0}$ is *cycle stationary* in the sense that $(0, Y_1'), (T_2' - T_1', Y_2'), (T_3' - T_2', Y_3'), \ldots$ has the same distribution as the original sequence; see [9, p. 205]. We let F denote the distribution function of T_1' and M the distribution of Y_0', and we assume that F has a finite mean $a > 0$ and that M has a density m (with respect to $\mathcal{H}^{d-1}$ on $\mathbb{S}^{d-1}$). Our assumptions guarantee the existence of a time-stationary version $((T_n, Y_n))_{n\in\mathbb{Z}}$ of the marked point process $((T_n', Y_n'))_{n\geq 0}$; see [3, Section 1.3.5] or [9, Theorem 11.4] for more detail. As in Section 2, we can then define the process $\{X_t : t \geq 0\}$ piecewise on the interval $[0, T_1)$ by letting the process start in $Y_0 - T_0 Y_0$ (a point that is not in the unit sphere) and then follow the trajectory induced by the drift μ (a ray that has the same direction as Y_0). On the intervals $[T_n, T_{n+1})$, $n \geq 1$, the process starts in Y_n and then again follows the trajectory induced by the drift μ. We consider the surface $S := r\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = r\}$ for some fixed $r > 1$. Clearly, the general model assumptions of Section 2 are all satisfied.

To obtain an explicit formula for the intensities of continuous crossings, we now assume that Y_0' and T_1' are independent. Using Formula (1.3.19) of [3] (or Formula (4) of [9, Chapter 11]), we obtain, for any $B \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned}
\mathrm{P}(X_0 \in B) &= \frac{1}{a}\mathrm{E}\int_0^{T_1'} \mathbf{1}\{Y_0' + tY_0' \in B\}\,\mathrm{d}t \\
&= \frac{1}{a}\int_0^\infty \int_{\mathbb{S}^{d-1}} (1 - F(t))\mathbf{1}\{(1+t)y \in B\}m(y)\mathcal{H}^{d-1}(\mathrm{d}y)\,\mathrm{d}t \\
&= \frac{1}{a}\int_1^\infty \int_{\mathbb{S}^{d-1}} (1 - F(t-1))\mathbf{1}\{ty \in B\}m(y)\mathcal{H}^{d-1}(\mathrm{d}y)\,\mathrm{d}t.
\end{aligned}$$

Using polar coordinates, we obtain

$$\mathrm{P}(X_0 \in B) = \frac{1}{a}\int_{\mathbb{R}^d} \mathbf{1}\{|x| \geq 1\}(1 - F(|x| - 1))\mathbf{1}\{x \in B\}m\left(\frac{x}{|x|}\right)|x|^{-(d-1)}\,\mathrm{d}x.$$

Therefore, X_0 has density $p(x) = a^{-1}|x|^{-(d-1)}m(x/|x|)(1-F(|x|-1))$, $|x| > 1$, with respect to the d-dimensional Lebesgue measure. This density is continuous in a neighbourhood of S provided that F is continuous in a neighbourhood of $r-1$. Assumption (A2) is trivially satisfied. Now it remains to observe that $|\langle n(x), \mu(x) \rangle| \equiv 1$ on S, and so, by Theorem 3.1, the intensity of continuous crossings of $S \cap B$ by X is clearly given by

$$\nu_c(B) = a^{-1}M(r^{-1}B)(1-F(r-1)). \tag{4.1}$$

Note that S and μ are simple in this example. But our probabilistic assumptions are still rather general. We have not assumed that the sequences (T'_n) and (Y'_n) have any special independence or Markov properties. Neither have we assumed that these sequences are independent. What is important for the explicit product form of (4.1) is the independence of Y'_0 and T'_1.

Example 4.2. Consider a general queueing network model with d servers operating in a stationary regime, with arrivals of customers (possibly in batches) to the network being governed by a stationary simple point process. Each customer, upon completion of its service at node $j \in \{1, \ldots, d\}$ of the network, proceeds to another node for further service or leaves the network, according to some routeing mechanism. All the arrival, transition, and departure times form a stationary point process N, and it is at these times that the state of the process $X_t = (X_t^{(1)}, \ldots, X_t^{(d)}) \in \mathbb{R}^d$ describing the residual workloads on the nodes can change by a jump. Between the events, the values of X_t decrease according to the relation $dX_t/dt = \mu(X_t)$ for some C^1-function $\mu \colon \mathbb{R}^d \to \mathbb{R}^d_-$, so that the service rate at node j can depend on the residual workload at the node and, moreover, it can even depend on the workloads at other nodes $i \neq j$ as well. To make this description compatible with the assumptions in Section 2, we allow $X_t^{(j)} < 0$, interpreting $\max\{X_t^{(j)}, 0\}$ as the residual workload at node j at time t, and let $D := \mathbb{R}^d$.

For $i \in \{1, \ldots, d\}$, let $S_i := \{x = (x_1, \ldots, x_d) \in \mathbb{R}^d \colon x_i = 0\}$. Then the continuous crossing of the surface S_i corresponds to server i becoming idle. Let $\nu_i(B)$ denote the intensity of these crossings through a point in $B \in \mathcal{B}(\mathbb{R}^d)$. Provided that the assumptions of Theorem 3.3 are satisfied, we obtain

$$\nu_i(B) = \int \cdots \int \mathbf{1}_B(x^i)|\mu_i(x^i)|p(x^i)\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_d \tag{4.2}$$

for $B \in \mathcal{B}(\mathbb{R}^d)$, where $x^i := (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_d)$ and μ_i is the ith component of μ. The normalization of (4.2) yields the (Palm) distribution of the network at a typical departure time from node i, making the node idle.

Note that our general model assumptions are rather mild and that Theorem 3.3 also requires $\mu_i(x) < 0$ for $x \in S_i$. In assumption (A2) only the condition $\pi_0(S_i \times (\mathbb{R}^d \setminus S_i)) = 0$ is of relevance. This assumption says that if there is a jump at an instant when server i becomes empty then the workload of this server is not allowed to increase by this jump, neither by an internal transition (including feedback) nor by an external arrival. Again, this is a rather weak assumption.

We can also consider the 'composite surface' $S := \bigcup_i S_{(i)}$, where $S_{(i)}$ is the set of all $x \in \mathbb{R}^d$ with $x_i = 0$ and $x_j \neq 0$ for $j \neq i$. (Under weak assumptions, any continuous crossing of S_i is also a continuous crossing of $S_{(i)}$.) Theorem 3.3 provides the Palm distribution of the residual workloads at the time when one of the servers becomes idle while all the others are still working. The probability of server i becoming idle given a typical instant when (exactly) one of the servers becomes idle is then given by $\nu_c(S_{(i)})/\nu_c(S)$. We omit further details.

Example 4.3. The classical stress release model in seismology (see, e.g. [7] and the references to earlier works therein) is a piecewise-deterministic Markov process X_t representing the level of 'stress' at a seismic fault at time t. The value X_t continuously increases at a linear rate due to the tectonic loading of the fault and drops by random jumps $\xi_i < 0$ (which may be assumed independent and identically distributed) when the stress discharges by way of earthquakes that occur at random times T_i whose intensity is given by $\psi(X_t)$ for some suitably chosen increasing risk function ψ (e.g. $\psi(x) = e^{\beta x}$ for some $\beta > 0$). In other words, the generator A of the process has the form

$$Ah(x) = \mu h'(x) + \psi(x)(E\, h(x + \xi) - h(x))$$

for a constant $\mu > 0$, 'generic' random jump size ξ, and functions $h \in C^1$ on $\mathbb{R}$.

Note that the remote measuring of stress levels at seismic faults is an extremely difficult problem, so the value X_t is usually not observable. All the information on the process one can have access to is contained in the times, locations, and magnitudes of the jumps.

A more interesting multinode analog of the model was discussed in [5], where it was demonstrated, in particular, that a two-node stress release network can reproduce the famous Omori's law for the intensity of earthquake aftershocks.

In the multinode model, the values of the components of the random process $X_t = (X_t^{(1)}, \ldots, X_t^{(d)}) \in \mathbb{R}^d$ represent the time t stress levels at individual seismic faults $j \in \{1, \ldots, d\}$ constituting a local fault system. Between jumps, the dynamics of the process are given by $dX_t/dt = \mu$ for a constant vector $\mu \in \mathbb{R}^d$. Note that we can have $\mu_j < 0$ which corresponds to the tectonic unloading of stress at node j (of course, we can consider a more general model with a variable μ as well; similar remarks apply to all the other elements of the model construction). The occurrence of jumps ('seismic events') at node j is driven by a Markovian random mechanism with the probability of a jump occurring at the node in the infinitesimal time interval dt given by $\psi_j(X_{t-}^{(j)})\, dt$ for a given risk function $\psi_j(x)$.

When the nth seismic event occurs at node j (say, at time $T_{j,n}$), the value of stress at the node changes by a random quantity $\xi_{j,n}$, $n = 1, 2, \ldots$, which may be assumed to be independent and identically distributed random variables. Moreover, the stress levels at other nodes can also change at that instance: for a given constant transfer matrix $(r_{ij}) \in \mathbb{R}^{d \times d}$, we have $X_{T_{j,n}}^{(i)} = X_{T_{j,n}-}^{(i)} + r_{ij}\xi_{j,n}, i \neq j$ (for more details, see [5]).

One of the main problems one hopes to be able to solve in mathematical seismology is to give advanced earthquake warnings. Within the framework of the multinode stress release model, that warning would have to be given at the time when the cumulative jump intensity $\sum_{j=1}^{d} \psi_j(X_t^{(j)})$ exceeds a given threshold $u > 0$. That is, we are looking at continuous crossings of the surface $S := \{x \colon \sum_{j=1}^{d} \psi_j(x_j) = u\}$ by our process X_t. Our main result allows us to find the distribution of X_t at the (typical) time of such crossings and, hence, for example, to derive the probability for a given fault to trigger the forthcoming seismic event.

References

[1] ADLER, R. J. AND TAYLOR, J. E. (2007). *Random Fields and Geometry*. Springer, New York.

[2] AZAÏS, J.-M. AND WSCHEBOR, M. (2009). *Level Sets and Extrema of Random Processes and Fields*. John Wiley, Hoboken, NJ.

[3] BACCELLI, F. AND BRÉMAUD, P. (1994). *Elements of Queueing Theory*. Springer, Berlin.

[4] BAR-DAVID, I. AND NEMIROVSKY, A. (1972). Level crossings of nondifferentiable shot processes. *IEEE Trans. Inf. Theory* **18,** 27–34.

[5] BOROVKOV, K. AND BEBBINGTON, M. S. (2003). A stochastic two-node stress transfer model reproducing Omori's law. *Pure Appl. Geophys.* **160,** 1429–1445.

[6] BOROVKOV, K. AND LAST, G. (2008). On level crossings for a general class of piecewise-deterministic Markov processes. *Adv. Appl. Prob.* **40,** 815–834.

[7] BOROVKOV, K. AND VERE-JONES, D. (2000). Explicit formulae for stationary distributions of stress release processes. *J. Appl. Prob.* **37,** 315–321.

[8] HARTMAN, P. (2002). *Ordinary Differential Equations*, 2nd edn. SIAM, Philadelphia, PA.

[9] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.

[10] LEADBETTER, M. R. AND SPANIOLO, G. V. (2004). Reflections on Rice's formulae for level crossings—history, extensions and use. *Austral. N. Z. J. Statist.* **46,** 173–180.

[11] RICE, S. O. (1944). Mathematical analysis of random noise. *Bell System Tech. J.* **24,** 282–332.

[12] RYCHLIK, I. (2000). On some reliability applications of Rice's formula for the intensity of level crossings. *Extremes* **3,** 331–348.

[13] ZÄHLE, U. (1984). A general Rice formula, Palm measures, and horizontal-window conditioning for random fields. *Stoch. Process. Appl.* **17,** 265-283.