The Annals of Probability
1996, Vol. 24, No. 3, 1563–1588

STABILITY OF NONLINEAR HAWKES PROCESSES

BY PIERRE BRÉMAUD AND LAURENT MASSOULIÉ

CNRS and CNET

We address the problem of the convergence to equilibrium of a general class of point processes, containing, in particular, the nonlinear mutually exciting point processes, an extension of the linear Hawkes processes, and give general conditions guaranteeing the existence of a stationary version and the convergence to equilibrium of a nonstationary version, both in distribution and in variation. We also give a new proof of a result of Kerstan concerning point processes with bounded intensity and general nonlinear dynamics satisfying a Lipschitz condition.

1. Nonlinear Hawkes processes. Let N be a simple point process on $\mathbb{R}$, that is, a family $\{N(C)\}_{C\in\mathscr{B}(\mathbb{R})}$ of random variables with values in $\bar{\mathbb{N}} = \mathbb{N}\cup\{+\infty\}$ indexed by the Borel σ-algebra $\mathscr{B}(\mathbb{R})$ of the real line $\mathbb{R}$, where $N(C) = \sum_{n\in\mathbb{Z}} \mathbf{1}_C(T_n)$ and $\{T_n\}_{n\in\mathbb{Z}}$ is a sequence of extended real-valued random variables such that, almost surely $T_0 \le 0 < T_1$, $T_n < T_{n+1}$ on $\{T_n < +\infty\}\cap\{T_{n+1} > -\infty\}$. Let $\{\mathscr{F}_t\}_{t\in\mathbb{R}}$ be a history of N, that is, a nondecreasing family of σ-fields such that, for all $t\in\mathbb{R}$, $\mathscr{F}_t^N = \sigma(N(C),\ C\in\mathscr{B}(\mathbb{R}),\ C\subset(-\infty,t])\subset\mathscr{F}_t$. The history $\{\mathscr{F}_t^N\}_{t\in\mathbb{R}}$ is called the internal history of N. Any nonnegative $\mathscr{F}_t$-progressively measurable process $\{\lambda(t)\}_{t\in\mathbb{R}}$ such that

$$\mathbf{E}[N((a,b])|\mathscr{F}_a] = \mathbf{E}\left[\int_a^b \lambda(s)\,ds\big|\mathscr{F}_a\right] \quad \text{a.s.}$$

for all intervals $(a,b]$, is called an $\mathscr{F}_t$-intensity of N. Note that $N(C) < \infty$ $\mathbf{P}$-a.s. if and only if $\int_C \lambda(s)\,ds < \infty$ $\mathbf{P}$-a.s., and therefore N is a.s. a Radon measure if and only if $\{\lambda(t)\}_{t\in\mathbb{R}}$ is a.s. locally integrable. Stochastic intensity is a generalization of the notion of the hazard rate (see [3], [8] and [15]). We recall a result and reproduce a proof due to Jacod [8].

LEMMA 1 (Hazard rate). *Let N be a point process admitting an $\mathscr{F}_t^N$-predictable intensity $\lambda(t) = v(t,N)$ on $\mathbb{R}^+$. Then, for all $t\in\mathbb{R}^+\cup\{+\infty\}$, the conditional probability $\mathbf{P}(N((0,t]) = 0|\mathscr{F}_0^N)$ equals $\exp-\int_0^t v(s,N^-)\,ds$, where N^- is, by definition, the restriction of N to $\mathbb{R}^-$ $[N^-(C) = N(C\cap\mathbb{R}^-),\ C\in\mathscr{B}(\mathbb{R})]$.*

PROOF. For all $t > 0$, it holds that

$$\mathbf{1}_{N(0,\,t]=0} = 1 - \int_{(0,\,t]} \mathbf{1}_{N(0,\,s)=0} N(ds).$$

Received September 1994; revised September 1995.
AMS 1991 *subject classifications*. 60G55, 60H20.
Key words and phrases. Stochastic processes, point processes, stochastic intensity, stationary point processes, mutually exciting point processes, Hawkes processes.

For all events $A \in \mathcal{F}_0^N$, the process $\mathbf{1}_A(\omega)\mathbf{1}_{(0,t]}(s)\mathbf{1}_{N(0,s)=0}$ is $\mathcal{F}_s^N$-predictable, so that, multiplying both sides of the previous inequality by $\mathbf{1}_A$, one can replace the integral with respect to $N(ds)$ by an integral with respect to $\lambda(s)\,ds$ when taking expectations, which yields

$$\mathbf{P}(N(0,t]=0;\ A) = \mathbf{P}(A) - \mathbf{E}\left[\mathbf{1}_A \int_0^t \mathbf{1}_{N(0,s)=0}\lambda(s)\,ds\right].$$

The processes $\mathbf{1}_{N(0,s)=0}$ and $\mathbf{1}_{N(0,s]=0}$ differ on a set of zero Lebesgue measure so that the first may be replaced by the second in the above equality; observing that $\mathbf{1}_{N(0,s]=0}\lambda(s) = \mathbf{1}_{N(0,s]=0}v(s, N^-)$, where $v(s, N^-)$ is $\mathcal{F}_0^N$-measurable, implies that

$$\mathbf{P}(N(0,t]=0|\mathcal{F}_0^N) = 1 - \int_0^t \mathbf{P}(N(0,s]=0|\mathcal{F}_0^N)v(s,N^-)\,ds.$$

Iterating this equality gives the announced result. □

REMARK 1. The fact that the $\mathcal{F}_t^N$-intensity $\lambda(t)$ is of the form $v(t, N)$ is not an assumption: it holds since $\lambda(t)$ is $\mathcal{F}_t^N$-predictable. See, for instance, [3], Appendix 2.

The present work is concerned with simple point processes N admitting an $\mathcal{F}_t^N$-intensity of the form

$$\lambda(t) = \phi\left(\int_{(-\infty,t)} h(t-s)N(ds)\right), \tag{1}$$

where ϕ: $\mathbb{R} \to \mathbb{R}^+$ and h: $\mathbb{R}^+ \to \mathbb{R}$. A particular case is Hawkes' self-exciting point process, for which h is nonnegative and $\phi(x) = \nu + x$ where $\nu > 0$ (see [6]; see also Daley and Vere-Jones [5], Chapter 10, page 367), and we therefore call such processes nonlinear Hawkes processes.

The notion of multivariate Hawkes processes, or mutually exciting point processes, can be extended to multivariate nonlinear Hawkes processes, that is, a family N_i, $1 \le i \le K$, of simple point processes without common points, of respective $\mathcal{F}_t$-intensities

$$\lambda_i(t) = \phi_i\left(\sum_{j=1}^K \int_{(-\infty,t)} h_{ji}(t-s)N_j(ds)\right), \tag{2}$$

where $\mathcal{F}_t = \vee_{i=1}^K \mathcal{F}_t^{N_i}$, ϕ_i: $\mathbb{R} \to \mathbb{R}^+$, h_{ji}: $\mathbb{R}^+ \to \mathbb{R}$.

Nonlinear Hawkes processes model a variety of situations.

EXAMPLE 1. In the univariate case, for instance, taking

$$\phi(x) = \lambda \mathbf{1}_{[0,K-1/2]}(x), \qquad h(t) = \mathbf{1}_{[0,a]}(t),$$

N is the input process to an $M/D/K/0$ queue, that is, a queue with Poisson arrivals of intensity $\lambda > 0$, service time $a > 0$, no waiting room and K servers (see [2], page 81, for instance).

EXAMPLE 2. Multivariate Hawkes processes also model neuronal activity and we shall therefore occasionally call them neural networks. In this context,

$$X_i(t) = \sum_{j=1}^{K} \int_{(-\infty,\, t)} h_{ji}(t-s) N_j(ds)$$

is the potential of neuron i at time t, ϕ_i is its excitation function and h_{ij} is the transfer function from neuron i to neuron j. If $1/\lambda_i \phi_i(x) = 1 - \mathbf{1}_{[0,\, \sigma_i]}(x)$, $\lambda_i > 0$, σ_i is called the excitation threshold of neuron i, and this neuron is said to be excited at time t if $X_i(t) \geq \sigma_i$, and at rest (or inhibited) at time t if $X_i(t) < \sigma_i$.

EXAMPLE 3. Consider a network of K neurons, where neuron i, $i \in \{1, \ldots, K\}$, fires at rate λ_i if for all $j \in \{1, \ldots, K\}$ no firing of neuron j occurred during the last θ_{ji} time units for some constant $\theta_{ji} \geq 0$, and is inhibited otherwise (see Example 3 in [4]). Calling N_i the point process of spikes of neuron i, the dynamics of this network are of the general type above, with

$$\phi_i(x) = \lambda_i \mathbf{1}_{[0,\, 1)}(x)$$

and

$$h_{ij}(t) = \mathbf{1}_{[0,\, \theta_{ij}]}(t).$$

Our goal in the present work is to find conditions bearing on the functions ϕ_i and h_{ji}, guaranteeing the existence and uniqueness of a stationary version of $N = (N_i,\ 1 \leq i \leq K)$, as well as the stability of the stationary solution (a concept that we shall explain soon). Our results extend and/or are related to the previous results of Hawkes and Oakes [7], Kerstan [10] and Lindvall [13] (see the discussion after each theorem below).

For any $t \in \mathbb{R}$ and any stochastic process X, $S_t X^{\pm}$ will denote the restriction to $\mathbb{R}^{\pm}$ of the process X shifted t time units to the left: for instance, if $X = N$ is a multivariate point process $(N_1, \ldots, N_K)$, then $S_t N^{\pm} = (S_t N_i^{\pm},\ 1 \leq i \leq K)$, where $S_t N_i^{\pm} = (N_i((t + C) \cap \mathbb{R}^{\pm}),\ C \in \mathscr{B}(\mathbb{R}))$. One says that N has an initial condition $(\mathscr{P}_{-})$ if its restriction to $\mathbb{R}^{-}$, $S_0 N^{-}$, is fixed and verifies condition $(\mathscr{P}_{-})$. An example of initial condition $(\mathscr{P}_{-})$ is: $\lim_{t \to +\infty} t^{-1} N_i((-t, 0]) = \alpha_i$ a.s., $1 \leq i \leq K$, for fixed $\alpha_i > 0$, $1 \leq i \leq K$.

The space of integer-valued measures is endowed with the topology of vague convergence; that is, a sequence m^n of point measures converges to a limit m if and only if, for any continuous function f with compact support, $\int f(x) m^n(dx) \to \int f(x) m(dx)$. The notion of weak convergence for point processes that we consider is the one derived from this topology. We shall also consider the stronger notion of convergence in variation for (distributions of) point processes: the sequence $\{N^n\}$ converges in variation to a limit N if and only if

$$\lim_{n \to \infty} \sup_{C} |\mathbf{P}(N^n \in C) - \mathbf{P}(N \in C)| = 0,$$

where the supremum is taken over the sets C in the Borel σ-field associated with the vague topology.

DEFINITION 1 (Stability). One says that the dynamics (2) are stable in distribution (resp. in variation) with respect to an initial condition ($\mathscr{P}_-$) if for all point processes N' with initial condition ($\mathscr{P}_-$) and following the dynamics (2) on $\mathbb{R}^+$ one can exhibit a point process N such that:

(i) N follows the dynamics (2) on $\mathbb{R}$ and is stationary;
(ii) $S_t N'^+ \to_{\mathscr{D}}$ (resp. $\to_{\mathrm{var}}$) N^+ as $t \to +\infty$.

REMARK 2. Stability in variation of the dynamics (2) with respect to an initial condition ($\mathscr{P}_-$) will hold if for all N' with initial condition ($\mathscr{P}_-$) and following the dynamics (2) on $\mathbb{R}^+$ one can construct N and N' on the same probability space, satisfying (i) above, and such that they couple, that is, such that:

(ii′) $S_t N^+ \equiv S_t N'^+$ for all $t \geq T$, where $\mathbf{P}(T < +\infty) = 1$.

The reader is referred to Lindvall [14] for more insight into the notion of coupling.

REMARK 3. Assume that one can exhibit some initial condition ($\mathscr{P}_-$) for which the dynamics (2) are stable and such that the distribution of the stationary point process N in (i) is the same for any N' satisfying ($\mathscr{P}_-$). Then any stationary solution N' satisfying ($\mathscr{P}_-$) is distributed as this N: indeed, for any $t > 0$, $S_t N'^+$ is distributed as $S_0 N'^+$, so that the laws of $S_0 N'^+$ and $S_0 N^+$ are identical; since the two processes are stationary, the laws of $S_{-t} N^+$ and $S_{-t} N'^+$ also coincide for any $t > 0$. Letting $t \to +\infty$, one sees that $N' =_{\mathscr{D}} N$. Assume further that any stationary solution necessarily satisfies ($\mathscr{P}_-$). One then obtains uniqueness of the stationary solution.

Remark 3 will actually be used to prove the uniqueness of the stationary distribution in some cases. The situation in which all the transfer functions h_{ij} have compact support and the excitation functions ϕ_i are bounded is an easy consequence, for instance, of the more general results of Lindvall [13] (see Theorem 3). The difficulty lies in the unbounded case for the excitation function and/or the absence of a compact support assumption for the transfer functions (see Theorems 1 and 2).

REMARK 4. Our use of the term stability in the title of this article does not seem to be standard in the theory of stochastic processes. The parallel with the notion of stability in the theory of ODE is the following: the initial condition is (the law of) $S_0 N^-$, the past of the point process at time 0. Considering (the law of) $S_t N^-$ as the state at time t, this state converges (weakly) to some limit as $t \to \infty$ if $S_t N^+ \to_{\mathscr{D}}$ some limit as $t \to \infty$. The parallel, in general, does not exist for stability in variation. Indeed, consider the following dynamics:

whatever the past $S_t N^-$ of the process at time t, $S_t N^+$ is independent of $S_t N^-$ and Poisson with intensity $\lambda > 0$ (this corresponds to the choice $\phi \equiv \lambda$). For the empty initial condition $S_0 N^- = \varnothing$ a.s., $S_t N^+$ converges in variation as $t \to \infty$ to the Poisson process with intensity $\lambda > 0$, and one therefore has stability in variation as in Definition 1. However, $S_t N^-$ fails to converge in variation, since it has almost surely finitely many points for all $t > 0$, which ensures that its law and the law of the Poisson process with intensity $\lambda > 0$ (under which there are a.s. infinitely many points) are mutually singular.

The article is structured as follows. In Section 2, we state the basic stability results for the single neuron case. In Section 3, we give the necessary propositions on Poisson imbedding, since all the constructions are based on them. In Section 4 we prove the results announced in Section 2, and in Section 5 we give extensions to the multivariate case.

2. Stability results for the single neuron. Hawkes and Oakes [7] have studied the stability of linear mutually exciting point processes. We shall state the two main generalizations of their results. Theorem 1 is a strict generalization, and Theorem 2 is more relevant to neuronal activity modeling.

THEOREM 1 (Unbounded Lipschitz dynamics). *Let ϕ be α-Lipschitz for some $\alpha > 0$, and let h be such that*

$$\alpha \int_{\mathbb{R}^+} |h(t)|\, dt < 1. \tag{3}$$

(a) *There exists a unique stationary distribution of N with finite average intensity $\mathbf{E}N((0,1])$ and with dynamics* (1).

(b) *The dynamics* (1) *are stable in distribution with respect to either initial condition* (i) *or* (ii) *below*:

(i) $\sup_{t\geq 0} \varepsilon_a(t) < +\infty$ *a.s. and* $\lim_{t\to+\infty} \varepsilon_a(t) = 0$ *a.s. for all* $a > 0$,

(ii) $\sup_{t\geq 0} \mathbf{E}(\varepsilon_a(t)) < +\infty$ *and* $\lim_{t\to+\infty} \mathbf{E}\varepsilon_a(t) = 0$ *for all* $a > 0$,

where $\varepsilon_a(t) := \int_{t-a}^{t} \int_{\mathbb{R}^-} |h(s-u)| N(du)\, ds$.

(c) *The dynamics* (1) *are stable in variation with respect to the initial condition*:

$$\text{(iii)} \qquad \int_{\mathbb{R}^+} |h(t)| N([-t,0))\, dt = \int_{\mathbb{R}^+} dt \int_{-\infty}^{0} |h(t-s)| N(ds) < +\infty \quad a.s.$$

if, *moreover*,

$$\int_{\mathbb{R}^+} t|h(t)|\, dt < +\infty. \tag{4}$$

REMARK 5. Condition (iii) says that the direct influence of the initial condition N^- vanishes as $t \to \infty$: indeed, in the linear case, each point $T_n < 0$ of N^- generates on $\mathbb{R}^+$ a Poisson process of intensity $h(t-T_n)$ with, conditionally on N^-, an average number of points equal to $\int_{\mathbb{R}^+} h(t-T_n)\, dt$. Summing up the (direct) contributions of all $T_n < 0$, one obtains a total number of points on $\mathbb{R}^+$ equal to $\int_{\mathbb{R}^+} dt \int_{-\infty}^{0} h(t-s)N(ds)$. Since this number is finite, the aftereffect of N^- vanishes in finite time. In the Lipschitz case, the direct influence of N^- produces a total number of points on $\mathbb{R}^+$ less than $\alpha \int_{\mathbb{R}^+} dt \int_{-\infty}^{0} |h(t-s)|N(ds)$, and when this number is finite, the aftereffect of N^- also vanishes in finite time. Condition (ii) has an analogous interpretation.

REMARK 6. Observe that if $\phi(0) = 0$ the empty point process is a solution, and therefore the solution with finite average intensity of the existence problem. This shows, in particular, that there is no linear stationary Hawkes process with stochastic intensity $\int_{(-\infty,\, t)} h(t-s)N(ds)$ (and finite average intensity) when $\int_{\mathbb{R}^+} h(t)\,dt < 1$, and that any transient process with these dynamics and an initial condition as in Theorem 1 eventually dies out in distribution or in variation.

We shall now relate Theorem 1 to the result of Hawkes and Oakes [7], taking $\alpha = 1$, since the general case $\alpha > 0$ is easily obtainable from the case $\alpha = 1$. In [7], existence is proven in the linear case $[\phi(x) = \nu + x$, where $\nu > 0]$ for nonnegative h verifying the condition $\int_{\mathbb{R}^+} h(t)\,dt < 1$, by constructing the corresponding stationary point process as a Poisson branching process or cluster Poisson process and using the elementary theory of branching processes. Hawkes and Oakes then prove uniqueness and stability in distribution with respect to an empty initial the condition, that is, $(\mathscr{P}_-) \equiv \{N(\mathbb{R}^-) = 0 \text{ a.s.}\}$. Their results are presented in Daley and Vere-Jones [5], where it is stated that the condition $\int_{\mathbb{R}^+} th(t)\,dt < +\infty$ guarantees stability in variation with respect to the empty initial condition. In this article, this initial condition will be replaced by more general ones. From a systems theory point of view, such general conditions are needed to take into account the previous behavior of the system before a change in the parameters, for instance.

The proofs of stability in Hawkes and Oakes [7] are based on arguments where the linearity of ϕ is crucial. More direct proofs can be devised using only purely branching arguments. In any case, both types of arguments (linear or branching) are not available in the generality we have placed the problem, and we shall have to resort to the theory of stochastic intensity.

The Lipschitz condition allows a treatment of existence analogous in some way to Picard's existence proof for ordinary differential equations. Another feature of our approach is the use of imbedded representations of point processes with stochastic intensities which permit coupling arguments for proving convergence in distribution or in variation. Since imbedding is realized in a homogeneous Poisson process on $\mathbb{R}^2$ which is mixing with respect to translations along the time axis, mixing of the constructed point process follows immediately if this point process is compatible with the translations of the homogeneous Poisson process on $\mathbb{R}^2$ (by this we mean that the first one is shifted by t whenever the second one is shifted by t along the first coordinate axis; see [2], Chapter 1, for formal definitions).

THEOREM 2 (Bounded Lipschitz dynamics). *Let ϕ be α-Lipschitz for some $\alpha > 0$ and bounded by $\Lambda > 0$, and let h be such that $\int_{\mathbb{R}^+} |h(t)|\,dt < \infty$ and that* (4) *holds. Then there exists a unique stationary distribution of N with dynamics* (1). *Moreover, the dynamics* (1) *are stable in variation with respect to the initial condition*

$$\lim_{t\to+\infty} \int_t^{+\infty} ds \int_{\mathbb{R}^-} |h(s-u)| N(du) = 0 \quad a.s. \tag{5}$$

The difference with Theorem 1 is that, at the expense of a boundedness condition, we can avoid fixing a maximal value of the $\mathbb{L}^1$-norm of h which depends on the Lipschitz coefficient. Consider, for instance, the case where

$$\phi(x) = \begin{cases} \nu, & \text{if } x \le A - \varepsilon, \\ \nu + \dfrac{\Lambda - \nu}{\varepsilon}(x - A + \varepsilon), & \text{if } A - \varepsilon \le x \le A, \\ \Lambda, & \text{if } x \ge A, \end{cases}$$

where $A > \varepsilon > 0, \Lambda > \nu > 0$. Theorem 2 shows that if h is in $\mathbb{L}^1$ and if (4) holds, there exists a unique stationary distribution of N with the dynamics (1), and that the dynamics (1) are stable with respect to the initial condition (5), and this for all ε, $0 < \varepsilon < A$. Choosing ε arbitrarily small yields an approximation of the pure threshold case $\phi(x) = \nu$ if $x < A$, $\phi(x) = \Lambda$ if $x \ge A$. Strictly speaking, however, this limiting case is not in the scope of Theorem 2. However, we can still say something about this case if h has a compact support (see Theorem 3).

Kerstan [10] obtained Theorem 2 in 1964, in an apparently more general form, since this author does not require the special form of intensity (1), but only a direct Lipschitz condition

$$|\lambda(t, N_1) - \lambda(t, N_2)| \le \alpha \int_{(-\infty, t)} |h(t-s)|\, |N_1 - N_2|(ds).$$

The proofs of Theorems 1 and 2 in the present article only use this condition. Our Theorem 2 is therefore of the same generality as the corresponding result of Kerstan, but does not contain more. However, our proof seems more easy, due to the fact that in his pioneering article Kerstan did not have at his disposal the martingale theory of stochastic intensity.

In order to state Theorem 3, some notation is needed. Call $(M, \mathscr{M})$ the measurable space of Radon measures on $\mathbb{R}$, where $\mathscr{M}$ is generated by the mappings $m \to m(C)$, $C \in \mathscr{B}(\mathbb{R})$. Let S_t denote the usual shift operator on the space $(M, \mathscr{M})$: $S_t m(C) = m(C + t)$.

DEFINITION 2. The mapping ψ: $(M, \mathscr{M}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ is causal if, whenever $m \equiv m'$ on $(-\infty, 0)$, $\psi(m) = \psi(m')$. Furthermore, ψ has a bounded memory if there exists A, $0 < A < \infty$, such that, whenever $m \equiv m'$ on $[-A, 0)$, $\psi(m) = \psi(m')$ (in particular, a mapping ψ with a bounded memory is a fortiori causal). A point process N is said to have dynamics with bounded memory if it admits an $\mathscr{F}_t^N$-intensity of the form

$$\lambda(t) = \psi(S_t N), \tag{6}$$

where ψ: $(M, \mathscr{M}) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ has a bounded memory.

For instance, if the support of the transfer function h is compact, the dynamics (1) have bounded memory.

THEOREM 3 (Bounded memory dynamics). *There exists a unique stationary distribution of a point process N with bounded memory dynamics* (6) *if*

$$\sup_{m \in M} \psi(m) = \Lambda < \infty. \tag{7}$$

Moreover, in this case, irrespective of the initial condition, the dynamics are stable in variation, and convergence in variation is exponentially fast.

Stationary point processes with bounded memory dynamics have been studied by Lindvall [13]. The class of processes studied by this author are the (A, m)-processes, for which the intensity $\lambda(t)$ depends only on the m last events before t and the events in the interval $[t - A, t)$. An extension to consistent finite random memory was made by Brémaud and Massoulié [4], to which the reader is referred for the exact definitions. Whether in [13] or [4], construction of the stationary point processes with the prescribed dynamics was made by first finding suitable regenerative events. Lindvall's technique in [13] was based on Harris recurrence theory, whereas the technique of [4] was based on imbedding (see, however, [16], pages 12–16, where [4] was revisited, for a shorter proof relying on a Doeblin-like minoration). A simple corollary of the main result of [13], or of [4], which contains Theorem 3 is given in [2] (Example 4.2.3, Chapter 2). Note that an unbounded support for the function h makes the memory always infinite, and regenerative arguments do not come up naturally, although they may exist.

Theorem 4 ensures the existence of stationary distributions for N with dynamics (1) without requiring finite memory or Lipschitz assumptions on ϕ. Instead, some form of monotonicity is required from the dynamics (see [13] for a discussion on the corresponding monotonic property of stochastic intensity kernels).

THEOREM 4 (Increasing kernel). *Let ϕ be a nonnegative, nondecreasing and left-continuous function, satisfying*

$$\phi(x) \leq \lambda + \alpha x, \qquad x \in \mathbb{R}, \tag{8}$$

for some $\lambda > 0$ and some $\alpha \geq 0$. Let h: $\mathbb{R}^+ \to \mathbb{R}^+$ be such that

$$\alpha \int_{\mathbb{R}^+} h(t)\, dt < 1. \tag{9}$$

Then there exists a stationary p.p. N with dynamics (1).

If $\alpha = 0$, the result of the theorem holds even for nonintegrable h.

The proofs of Theorems 1–4 will be given in Section 4, after we give some basic results concerning imbedding.

3. Poisson imbedding. We need to introduce some notation. Let $N = \{T_n\}_{n\in\mathbb{Z}}$ be a simple, nonexplosive point process, and let $\{Z_n\}_{n\in\mathbb{Z}}$ be a sequence of E-valued variables, where $(E, \mathscr{E})$ is an arbitrary measurable space. The double sequence $\{(T_n, Z_n)\}_{n\in\mathbb{Z}}$ is denoted N_Z and called a marked, simple, nonexplosive point process with marks in E. For any function H: $\mathbb{R}\times E \to \mathbb{R}^+$, one denotes

$$\int_{\mathbb{R}\times E} H(t, z) N_Z(dt \times dz) = \sum_{n\in\mathbb{Z}} H(T_n, Z_n) \tag{10}$$

and, for $C \subset \mathbb{R} \times E$,

$$N_Z(C) = \sum_{n\in\mathbb{Z}} \mathbf{1}_C(T_n, Z_n). \tag{11}$$

For $t \in \mathbb{R}$, denote by $\mathscr{F}_t^{N_Z}$ the σ-field generated by the random variables $N_Z(C)$, $C \in \mathscr{B}((-\infty, t]) \otimes \mathscr{E}$. The next two results are well known and were used by Kerstan [10] in a similar context as ours, and by Lewis and Shedler [12] and Ogata [18] in a simulation context (see the discussion in Daley and Vere-Jones [5]; see also Last [11]).

LEMMA 2. *Let $\bar{N} = \{(T_n, U_n)\}_{n\in\mathbb{Z}}$ be a marked point process for which $\{T_n\}_{n\in\mathbb{Z}}$ is a Poisson process with intensity Λ, $\{U_n\}_{n\in\mathbb{Z}}$ is an i.i.d. sequence of $[0, 1]$-uniform random variables, independent of $\{T_n\}_{n\in\mathbb{Z}}$. Let $\{\mathscr{F}_t\}_{t\in\mathbb{R}}$ be a history of $\bar{N}$ such that $\mathscr{F}_s$ and $S_t\bar{N}^+$ are independent for all $s \le t$. Let $\{\lambda(t)\}_{t\in\mathbb{R}}$ be a nonnegative $\mathscr{F}_t$-predictable process bounded uniformly in (t, ω) by Λ. The point process N defined by*

$$N(C) = \sum_{n\in\mathbb{Z}} \mathbf{1}_C(T_n)\mathbf{1}_{[0,\, \lambda(T_n)/\Lambda]}(U_n) = \int_C \bar{N}\left(dt \times \left[0, \frac{\lambda(t)}{\Lambda}\right]\right) \tag{12}$$

for all $C \in \mathscr{B}(\mathbb{R})$ admits $\{\lambda(t)\}_{t\in\mathbb{R}}$ as an $\mathscr{F}_t^{\bar{N}}$-intensity.

A possible candidate for $\mathscr{F}_t$ is $\mathscr{F}_t^{\bar{N}} \vee \mathscr{F}_t^{Z}$, where the process Z is independent of $\bar{N}$. For any simple point process $\bar{N}$ on $\mathbb{R}^2$ and any $t \in \mathbb{R}$, denote by $\mathscr{F}_t^{\bar{N}}$ the σ-field generated by the random variables $\bar{N}(C)$, $C \in \mathscr{B}((-\infty, t] \times \mathbb{R})$. The next lemma is a variant of Lemma 2.

LEMMA 3. *Let $\bar{N}$ be a Poisson process of intensity 1 on $\mathbb{R}^2$. Let $\mathscr{F}_t$ be a history of $\bar{N}$ (i.e., $\mathscr{F}_t^{\bar{N}} \subset \mathscr{F}_t$, $t \in \mathbb{R}$) such that $\mathscr{F}_s$ and $S_t\bar{N}^+$ are independent for all $s < t$. Let $\{\lambda(t)\}_{t\in\mathbb{R}}$ be a nonnegative $\mathscr{F}_t^{\bar{N}}$-predictable process and define the point process N by*

$$N(C) = \int_{C\times\mathbb{R}} \mathbf{1}_{[0,\, \lambda(t)]}(z)\bar{N}(dt \times dz) \tag{13}$$

for all $C \in \mathscr{B}(\mathbb{R})$. Then N admits the $\mathscr{F}_t^{\bar{N}}$-intensity $\{\lambda(t)\}_{t\in\mathbb{R}}$.

EXAMPLE 4. Let N' be some $\mathcal{F}_t^{\bar{N}}$-adapted p.p. Let ϕ be a nonnegative function on $\mathbb{R}$, and let h be a function on $\mathbb{R}^+$. If the process

$$\lambda(t) = \phi\left[\int_{(-\infty,\, t)} h(t-s)N'(ds)\right], \qquad t \in \mathbb{R},$$

is a.s. locally integrable, then the p.p. N defined by (13) is $\mathcal{F}_t^{\bar{N}}$-adapted, and admits $\{\lambda(t)\}$ as an $\mathcal{F}_t^{\bar{N}}$-intensity. Indeed, this follows from Lemma 3 if $\{\lambda(t)\}$ is $\mathcal{F}_t^{\bar{N}}$-predictable, which can be shown by means of [3], Theorem 24, Appendix A2, page 304 [the fact that the integral defining $\lambda(t)$ is taken on $(-\infty, t)$ rather than on $(-\infty, t]$ is crucial to ensure predictability].

The result to follow is a kind of converse of Lemma 3, and is a special case of the results of Jacod [9], Chapter 14, pages 469–478.

LEMMA 4 (Poisson inversion). *Let $N = \{T_n\}_{n\in\mathbb{Z}}$ be a simple, nonexplosive point process on $\mathbb{R}$ with $\mathcal{F}_t$-intensity $\{\lambda(t)\}_{t\in\mathbb{R}}$, and assume that this intensity is $\mathcal{F}_t$-predictable. Let $\{U_n\}_{n\in\mathbb{Z}}$ be a sequence of i.i.d. random variables, uniformly distributed on* $[0, 1]$ *and independent of $\mathcal{F}_\infty$. Let $\hat{N}$ be a homogeneous Poisson process on $\mathbb{R}^2$ with intensity* 1, *independent of $\mathcal{F}_\infty \vee \mathcal{F}_\infty^U$. Define a point process $\bar{N}$ on $\mathbb{R}^2$ by*

$$\bar{N}((a, b] \times L) = \sum_{n\in\mathbb{Z}} \mathbf{1}_{(a,\, b]}(T_n)\mathbf{1}_L(\lambda(T_n)U_n) + \int_{(a,\, b]}\int_{L-(0,\, \lambda(t)]} \hat{N}(dt \times dz).$$

Then $\bar{N}$ is a homogeneous Poisson process on $\mathbb{R}^2$ with intensity 1 *and such that $S_t\bar{N}^+$ is independent of $\mathcal{F}_s \vee \mathcal{F}_s^{\bar{N}}$ for all $s < t$; that is, it is an $\mathcal{F}_t \vee \mathcal{F}_t^{\bar{N}}$-Poisson process.*

The physical meaning of Lemma 4 is the following: it shows that any point process with a stochastic intensity can be constructed as in Lemma 3. The corresponding bivariate homogeneous Poisson process $\bar{N}$ is any such process outside the strip $\{(t, z)\colon\ t \in \mathbb{R},\ 0 < z \le \lambda(t)\}$ (a "random" strip), and inside this strip $\bar{N}$ is obtained by marking the events T_n of N by the marks $Z_n = \lambda(T_n)U_n$ or, equivalently, by placing a point at random in the segment $[(T_n, 0), (T_n, \lambda(T_n))]$.

4. Proofs. Let us start with the proof of Theorem 3, which is simpler than the other proofs, since the construction of the stationary solution is done in only one step, whereas in the proofs of Theorems 1, 2 and 4 the stationary solution is identified as the limit of an iterative construction scheme.

PROOF OF THEOREM 3. Let $(\Omega, \mathcal{F})$ be the canonical space of marked point processes on $\mathbb{R}$, with $[0, 1]$-valued marks. Endow it with the probability $\mathbf{P}$ under which $\omega = \{T_n, U_n\}$ is distributed as the marked point process $\bar{N}$ of Lemma 2; that is, $\{T_n\}$ is Poisson with intensity $\Lambda = \sup_{m\in M}\psi(m)$ [which is finite according to (7)] and independent of the marks $\{U_n\}$ which are i.i.d. and uniform on $[0, 1]$. Note $\bar{N}(\omega) = \omega$ for convenience. Let $\{\theta_t\}$ denote the usual shift operator on $(\Omega, \mathscr{F})$ (although the shift $\{\theta_t\}$ is of the same nature as the shift $\{S_t\}$ previously defined, we use a different notation because of the special role played by the point process ω, as the stochastic basis of the following construction). As is well known, $(\Omega, \mathscr{F}, \mathbf{P})$ is ergodic (and even mixing; see [5]) for the shift $\{\theta_t\}$. Assume that one can construct a p.p. N that is θ_t-compatible [i.e., for all $t \in \mathbb{R}$, $\omega \in \Omega$, $S_t N(\omega) = N(\theta_t \omega)$: a translation of the basis ω upon which N is constructed yields a translation of the same length for the resulting N] and such that

$$N(C) = \int_C \bar{N}\left(dt \times \left[0, \frac{\psi(S_t N)}{\Lambda}\right]\right), \qquad C \in \mathscr{B}(\mathbb{R}).$$

If N is $\mathscr{F}_t^{\bar{N}}$-adapted, then, by the causality of ψ, $\lambda(t) = \psi(S_t N)$ is $\mathscr{F}_t^{\bar{N}}$-predictable (see Example 4), and according to Lemma 2, it is an $\mathscr{F}_t^{\bar{N}}$-intensity of N. Since it is also $\mathscr{F}_t^{N}$-adapted, it is an $\mathscr{F}_t^{N}$-intensity of N as well. The θ_t-compatibility of N then ensures that it is a stationary (and even mixing) p.p. with the expected dynamics. Indeed, the stationarity of N means that, for every bounded, measurable functional f, one has

$$\mathbf{E}[f(N(\omega))] = \mathbf{E}[f(S_t N(\omega))].$$

By θ_t-compatibility, the right-hand side of the above equation equals $\mathbf{E}[f(N(\theta_t \omega))]$, which equals $\mathbf{E}[f(N(\omega))]$ because $\mathbf{P}$ is θ_t-invariant, and thus the above equality holds; that is, N is stationary. Using θ_t-compatibility, one shows in a similar fashion that N inherits properties such as ergodicity, weak mixing and mixing if they are true for the underlying probability space $(\Omega, \mathscr{F}, \mathbf{P})$ endowed with the shift $\{\theta_t\}$.

The announced construction is feasible indeed: let $\{R_k\}$ be the point process counting those points T_n of $\bar{N}$ such that $T_n - T_{n-1} > A$. The p.p. $\{R_k\}$ is, by definition, θ_t-compatible, and has finite, nonnull (average) intensity $\Lambda e^{-\Lambda A}$. Since $\psi(m)$ depends on the p.p. $m \in M$ through its behavior on $[-A, 0)$ only, the p.p. N may be constructed as announced on $[R_k, +\infty)$ without knowledge of N on $(-\infty, R_k)$ for all $k \in \mathbb{Z}$. Since R_k tends to $-\infty$ with k, N can thus be constructed on $\mathbb{R}$, and it is θ_t-compatible since $\bar{N}$ and $\{R_k\}$ are so. Also, it is $\mathscr{F}_t^{\bar{N}}$-adapted, since, for all $s < t$, $N(s, t]$ can be constructed from the restriction of $\bar{N}$ to $[R^-(s), t]$, where $R^-(s) = \sup\{R_k \mid R_k \le s\}$ is clearly $\mathscr{F}_t^{\bar{N}}$-measurable.

Let N' be some p.p. admitting the $\mathscr{F}_t^{N'}$-intensity $\lambda'(t) = \psi(S_t N')$ on $\mathbb{R}^+$. Use Lemma 4 to define (on the probability space where N' lives, possibly enlarged) a marked p.p. $\bar{N}$ distributed as above and such that

$$N'(C) = \int_C \bar{N}\left(dt \times \left[0, \frac{\lambda'(t)}{\Lambda}\right]\right)$$

for all $C \subset \mathbb{R}^+$. Construct the stationary p.p. N from $\bar{N}$ as above, and define

$$T = \mathbf{1}_{\{T_1 \le A\}}\left[T_1 + \sum_{n \ge 2} (T_n - T_{n-1}) \prod_{k=2}^{n} \mathbf{1}_{\{T_k - T_{k-1} \le A\}}\right].$$

That is, $T = 0$ if $T_1 > A$; otherwise it is the first point T_k such that $T_{k+1} - T_k > A$. Clearly, T is a coupling time for N and N', so that the variation distance between the distributions of $S_t N^+$ and $S_t N'^+$ is less than $\mathbf{P}(T > t)$, which, by Chebyshev's inequality, is less than $(\mathbf{E}e^{\alpha T})e^{-\alpha t}$ for all $\alpha > 0$ (see Lindvall [14]). Writing

$$e^{\alpha T} = \mathbf{1}_{\{T_1 > A\}} + \mathbf{1}_{\{T_1 \le A\}} e^{\alpha T_1} \sum_{n \ge 2} \mathbf{1}_{\{T_n - T_{n-1} > A\}} \prod_{k=2}^{n-1} e^{\alpha(T_k - T_{k-1})} \mathbf{1}_{\{T_k - T_{k-1} \le A\}},$$

one sees that the expectation $\mathbf{E}(e^{\alpha T})$ is finite iff

$$\frac{\Lambda}{\alpha - \Lambda}[e^{(\alpha - \Lambda)A} - 1]$$

is strictly less than 1, which holds for α small enough; this ensures that the convergence in variation of the law of $S_t N'^+$ as $t \to \infty$ takes place at an exponential speed. □

One could have taken R_1, which is the first point $T_n > 0$ such that $T_n - T_{n-1} > A$, instead of T as a coupling time for N and N'. However, the bounds thus obtained are not as sharp as those obtained with T: indeed, since $T \le R_1$ a.s., if $e^{\alpha R_1}$ has finite expectation, then so has $e^{\alpha T}$. Also, because $R_1 \ge T_1$, which is exponential with parameter Λ, in order to have $\mathbf{E}e^{\alpha R_1} < \infty$, it is necessary that $\alpha < \Lambda$, whereas

$$\frac{\Lambda}{\alpha - \Lambda}[e^{(\alpha - \Lambda)A} - 1]$$

can be less than 1 with $\alpha > \Lambda$ (for fixed α and Λ, when A decreases to 0 this quantity is equivalent to ΛA, which eventually becomes less than 1 no matter how large α is).

PROOF OF THEOREM 4. Let $(\Omega, \mathscr{F})$ be the canonical space of a p.p. on $\mathbb{R}^2$, which we endow with the probability $\mathbf{P}$ under which $\bar{N}(\omega) = \omega$ is Poisson with intensity 1 (and is therefore distributed as in Lemma 3). The shift $\{\theta_t\}$ corresponds to translations along the x-axis. Let $\lambda^0(t) = 0$, $t \in \mathbb{R}$, and let N^0 be the p.p. counting the points of $\bar{N}$ below the curve $t \to \lambda^0(t)$ (i.e., $N^0 = \varnothing$). Construct recursively the processes $\{\lambda^n(t)\}$ and N^n, $n \ge 0$, as follows:

$$\lambda^{n+1}(t) = \phi\left[\int_{(-\infty, t)} h(t-s) N^n(ds)\right], \qquad t \in \mathbb{R},$$

(14)

$$N^{n+1}(C) = \int_C \bar{N}(dt \times [0, \lambda^{n+1}(t)]), \qquad C \in \mathscr{B}(\mathbb{R}).$$

These processes are θ_t-compatible; also, it is easily seen by induction on n, using Lemma 3 and Example 4, that each N^n is $\mathscr{F}_t^{\bar{N}}$-adapted, each $\{\lambda^n(t)\}$ is $\mathscr{F}_t^{\bar{N}}$-predictable and is an $\mathscr{F}_t^{\bar{N}}$-intensity of N^n. Moreover, $h \ge 0$ and ϕ nondecreasing imply that $\lambda^n(t)$ and $N^n(C)$ increase with n for all $\omega \in \Omega$, $t \in \mathbb{R}$, $C \in \mathscr{B}(\mathbb{R})$, so that the limiting processes $\{\lambda(t)\}$ and N (which are

θ_t-compatible by construction) are defined for all $\omega \in \Omega$. Stationarity of the processes $\{\lambda^n(t)\}$ implies, together with assumption (8),

$$\mathbf{E}\,\lambda^{n+1}(0) \le \lambda + [\mathbf{E}\,\lambda^n(0)]\,\alpha \int_{\mathbb{R}^+} h(s)\,ds, \qquad n \ge 0,$$

so that, recalling (9), $\mathbf{E}\lambda(0) \le \lambda[1 - \alpha \int h(s)\,ds]^{-1} < \infty$. Here $\{\lambda(t)\}$ is $\mathscr{F}_t^{\bar{N}}$-predictable as a limit of such processes; the p.p. N, which counts the points of $\bar{N}$ below the curve $t \to \lambda(t)$, therefore admits (by Lemma 3) $\{\lambda(t)\}$ as an $\mathscr{F}_t^{\bar{N}}$-intensity. The proof will be complete if

$$\lambda(t) = \phi\left[\int_{(-\infty,\,t)} h(t-s)N(ds)\right], \qquad t \in \mathbb{R}. \tag{15}$$

The monotonicity properties of N^n and $\{\lambda^n(t)\}$ ensure that, for all $n \ge 0$, $t \in \mathbb{R}$,

$$\lambda^n(t) \le \phi\left[\int_{(-\infty,\,t)} h(t-s)N(ds)\right]$$

and

$$\lambda(t) \ge \phi\left[\int_{(-\infty,\,t)} h(t-s)N^n(ds)\right].$$

Letting n go to ∞ in both inequalities (this is valid in the second one because ϕ is left-continuous) then yields (15). □

REMARK 7. Under the hypotheses of Theorem 4, it is not clear whether the stationary law for N is unique (nothing is known, a fortiori, on the asymptotic stationarity of transient processes with such dynamics). Still, one can note that the process N constructed in the proof is the smallest stationary solution, in the sense that, given another stationary solution $\tilde{N}$, one can construct (using Lemma 4) a version of N on the space where $\tilde{N}$ lives, so that $N(C) \le \tilde{N}(C)$ for all $C \in \mathscr{B}(\mathbb{R})$. If ϕ is also right-continuous (and therefore continuous), one can construct a largest stationary solution (among those with finite mean intensity) to the problem: initiate the recursive construction procedure in the proof by taking for N^0 a stationary p.p. with stochastic intensity $\lambda^0(t) = \lambda + \int_{(-\infty,\,t)} h(t-s)N^0(ds)$; (8) implies then that the scheme is decreasing and that its limit is the largest solution to the problem considered.

PROOF OF THEOREM 1. (a) Without loss of generality, assume that $\alpha = 1$ [this amounts to replacing $\phi(\cdot)$ by $\phi(\alpha^{-1}\cdot)$ and h by αh]. Construct exactly as in the proof of Theorem 4 the sequence of θ_t-compatible processes N^n and $\{\lambda^n(t)\}$ [see (14)]. The Lipschitz property of ϕ ensures that

$$\mathbf{E}|\lambda^{n+1}(0) - \lambda^n(0)| \le \mathbf{E}\int_{(-\infty,\,0)} |h(-s)|\,|N^n - N^{n-1}|(ds), \qquad n \ge 1,$$

where the p.p. $|N^n - N^{n-1}|$ is defined by $|N^n - N^{n-1}|(\{t\}) = |N^n(\{t\}) - N^{n-1}(\{t\})|$, $t \in \mathbb{R}$ [it counts the points of $\bar{N}$ that fall between the two curves

$t \to \lambda^n(t)$ and $t \to \lambda^{n-1}(t)$]. By an easy modification of Lemma 3, it admits $\{|\lambda^n(t) - \lambda^{n-1}(t)|\}$ as an $\mathscr{F}_t^{\bar{N}}$-intensity, so that

$$\mathbf{E}|\lambda^{n+1}(0) - \lambda^n(0)| \le \left[\int_{\mathbb{R}^+} |h(s)|\, ds\right] \mathbf{E}|\lambda^n(0) - \lambda^{n-1}(0)|. \tag{16}$$

The sum $\sum_{n\ge 0} \mathbf{E}|\lambda^{n+1}(0) - \lambda^n(0)|$ is therefore finite according to hypothesis (3), so that $\lambda^n(0)$ converges in $\mathbb{L}^1$ to a limit $\lambda(0)$. This convergence is a.s.: indeed, it follows from Chebyshev's inequality and (16) that

$$\mathbf{P}\left[|\lambda^{n+1}(0) - \lambda^n(0)| \ge \left(\int_0^\infty |h(s)|\, ds\right)^{n/2}\right] \le \phi(0)\left(\int_0^\infty |h(s)|\, ds\right)^{n/2},$$

the sum of which is finite, and a.s. convergence is then a consequence of Neveu [17], Proposition 2.4.2, page 45. Also, the calculations

$$\begin{aligned}\sum_{n\ge 0} \mathbf{P}\left(\int_C |N^{n+1} - N^n|(ds) \ne 0\right) &\le \sum_{n\ge 0} \mathbf{E}\left(\int_C |N^{n+1} - N^n|(ds)\right)\\ &= \left(\int_C ds\right) \sum_{n\ge 0} \mathbf{E}|\lambda^{n+1}(0) - \lambda^n(0)|\\ &< +\infty,\end{aligned}$$

where C is any bounded set in $\mathscr{B}(\mathbb{R})$, imply, by the Borel–Cantelli lemma, that the processes N^n remain eventually constant on any bounded set as n increases; they therefore converge to a limiting, θ_t-compatible p.p. N as $n \to +\infty$. This limiting process counts the points of $\bar{N}$ below $t \to \lambda(t)$: indeed, for all bounded C, by Fatou's lemma,

$$\begin{aligned}&\mathbf{E}\int_C |N(ds) - \bar{N}(ds \times [0, \lambda(s)])|\\ &\quad \le \lim_{n\to\infty} \mathbf{E}\int_C |\bar{N}(ds \times [0, \lambda^n(s)]) - \bar{N}(ds \times [0, \lambda(s)])|\\ &\quad = \left(\int_C ds\right) \lim_{n\to\infty} \mathbf{E}|\lambda^n(0) - \lambda(0)|\\ &\quad = 0.\end{aligned}$$

Let us check that $\{\lambda(t)\}$ is a modification of $\{\phi[\int_{(-\infty,\, t)} h(t-s)N(ds)]\}$, which will establish the existence part of Theorem 1. Use again the Lipschitz property of ϕ to check that

$$\begin{aligned}&\mathbf{E}\left|\lambda(0) - \phi\left[\int_{(-\infty,\, 0)} h(-s)N(ds)\right]\right|\\ &\quad \le \mathbf{E}|\lambda(0) - \lambda^n(0)| + \mathbf{E}\int_{(-\infty,\, 0)} |h(-s)|\, |N - N^{n-1}|(ds)\\ &\quad = \mathbf{E}|\lambda(0) - \lambda^n(0)| + \left[\int_{\mathbb{R}^+} |h(s)|\, ds\right] \mathbf{E}|\lambda(0) - \lambda^{n-1}(0)|,\end{aligned}$$

and let n tend to $+\infty$ in the above expression to conclude.

We shall show in (b) below that any transient $\tilde{N}$ with the expected dynamics on $\mathbb{R}^+$ is such that $S_t\tilde{N}$ converges in law to the stationary process N as $t \to \infty$, provided $\tilde{N}$ satisfies any of the two initial conditions (i) or (ii). Uniqueness in law of a stationary solution with finite mean intensity will therefore hold if any such solution satisfies (ii). Let us show that this is indeed the case: for stationary $\tilde{N}$ with finite mean intensity $\tilde{\lambda}$, the corresponding quantity $\varepsilon_a(t)$ is such that

$$\begin{aligned}\mathbf{E}(\varepsilon_a(t)) &= \tilde{\lambda}\int_{t-a}^{t} ds \int_{\mathbb{R}^-} |h(s-u)|\,du \\ &= \tilde{\lambda}\int_{t-a}^{t} ds \int_{s}^{+\infty} |h(u)|\,du \\ &\le \tilde{\lambda}a \int_{t-a}^{+\infty} |h(u)|\,du.\end{aligned}$$

It follows from this bound that the mean $\mathbf{E}(\varepsilon_a(t))$ is bounded in t [take $\tilde{\lambda}a\int_{\mathbb{R}^+} |h(u)|\,du$ as the upper bound] and goes to 0 as $t \to \infty$, these two properties being true for all $a > 0$. That is to say, $\tilde{N}$ satisfies initial condition (ii).

(b) Now let N' be some transient p.p. with dynamics (1) on $\mathbb{R}^+$ and initial condition (i). We will show later that the $\mathscr{F}_t^{N'}$-intensity $\lambda'(t)$ of N', which equals $\phi[\int_{(-\infty,t)} h(t-s)N'(ds)]$, is such that $t \to \mathbf{E}(\lambda'(t)|\mathscr{F}_0^{N'})$ is a.s. locally integrable. This ensures that N' is nonexplosive. One can then use Lemma 4 to construct $\bar{N}$ distributed as above and such that N' counts the points of $\bar{N}$ below $t \to \lambda'(t)$ on $\mathbb{R}^+$; then construct the stationary p.p. N from the same $\bar{N}$, as above. Set

$$f(t) = \begin{cases} \mathbf{E}[|\lambda(t) - \lambda'(t)| \mid \mathscr{F}_0^{N'}], & \text{if } t \ge 0, \\ 0, & \text{otherwise.}\end{cases} \tag{17}$$

Since $\mathbf{E}(\lambda'(t)|\mathscr{F}_0^{N'})$ is a.s. locally integrable, then so is f. Calling λ the (average) intensity of N, the Lipschitz property of ϕ ensures

$$\begin{aligned} f(t) &\le \int_{\mathbb{R}^-} |h(t-s)|N'(ds) + \lambda \int_{\mathbb{R}^-} |h(t-s)|\,ds \\ &\quad + \int_0^t |h(t-s)|f(s)\,ds, \qquad t \ge 0.\end{aligned} \tag{18}$$

Fix $a > 0$, and set $F(t) = \int_{t-a}^{t} f(u)\,du$, $t \in \mathbb{R}$. Integrating (18) between t and $t+a$, one gets after rearrangement (and further bounding for the second term on the right-hand side)

$$F(t) \le \varepsilon_a(t) + \lambda a \int_{t-a}^{+\infty} |h(s)|\,ds + \int_0^t F(t-s)|h(s)|\,ds.$$

Noting $\varepsilon'_a(t)$ the sum of the first two terms on the right-hand side of the above inequality,

$$F(t) \le \varepsilon'_a(t) + \int_0^t F(t-s)|h(s)|\,ds, \qquad t \in \mathbb{R}.$$

Iterating this inequality gives, for all n and t,

$$F(t) \le \sum_{i=0}^{n-1} \varepsilon'_a * |h|^{*i}(t) + F * |h|^{*n}(t), \tag{19}$$

where F is bounded on any finite interval since f is locally integrable; also, the $\mathbb{L}^1$-norm of h is less than 1. Therefore, the last term on the right-hand side of the above inequality vanishes as $n \to +\infty$, and

$$F(t) \le \int_{\mathbb{R}^+} \varepsilon'_a(t-s) H(s)\, ds, \tag{20}$$

where $H = \sum_{n \ge 0} |h|^{*n}$ is the density of the defective renewal measure associated with $|\ h\ |$, which is of finite total mass $(1 - \int |h(s)|\, ds)^{-1}$. Under (i), $\varepsilon'_a(t)$ is bounded and tends to 0 as $t \to \infty$, so that, by dominated convergence, $\lim_{t\to\infty} F(t) = 0$. Observing that

$$F(t) = \mathbf{E}\left[\int_{(t-a,\,t]} |N - N'|(ds) \big| \mathscr{F}_0^{N'}\right] \ge 1 - \mathbf{P}(N' \equiv N \text{ on } [t-a, t) | \mathscr{F}_0^{N'}),$$

this implies that the finite-dimensional distributions of $S_t N'$ converge in variation, as $t \to \infty$, to those of N, and therefore $S_t N' \to_{\mathscr{D}} N$ as $t \to +\infty$ [see Theorem 9.1.6, page 274 in [5]]. Under (ii), take expectations in (18) and apply the same arguments to obtain the same results.

Uniqueness of the stationary law of a p.p. with such dynamics and finite average intensity follows, since condition (ii) is verified for such a process.

(c) Now integrate (18) with respect to t between 0 and $T > 0$, to obtain

$$\int_0^T f(t)\, dt \le \left[1 - \int_{\mathbb{R}^+} |h(t)|\, dt\right]^{-1} \left[\int_{\mathbb{R}^+} |h(t)| N'(-t, 0]\, dt + \lambda \int_{\mathbb{R}^+} t|h(t)|\, dt\right].$$

Under (4) and (i′), the right-hand side of the above inequality is a.s. finite and independent of T, so that $\int_0^{+\infty} f(t)\, dt$ is a.s. finite. Since $\int_0^{+\infty} f(t)\, dt$ is, by the definition of stochastic intensity, the mean number of points of $|N - N'|$ on $\mathbb{R}^+$ conditionally on $\mathscr{F}_0^{N'}$, there are a.s. finitely many points of $|N - N'|$ on $\mathbb{R}^+$; that is, N and N' couple.

It remains to show that $\mathbf{E}[\lambda'(t) | \mathscr{F}_0^{N'}]$ is a.s. locally integrable. Set $\lambda'^0(t) = \phi[\int_{\mathbb{R}^-} h(t-s) N'(ds)]$, and let N'^0 be the point process that coincides with N' on $\mathbb{R}^-$ and that counts the points of $\bar{N}$ below $\lambda'^0(t)$ on $\mathbb{R}^+$. Similar to the construction of N, construct recursively the processes $\lambda'^n(t) = \phi[\int_{(-\infty,\,t)} h(t - s) N'^{n-1}(ds)]$ and N'^n to be the p.p. that coincides with N' on $\mathbb{R}^-$ and that counts the points of $\bar{N}$ below $\lambda'^n(t)$ on $\mathbb{R}^+$. Fix $T > 0$; using the Lipschitz property of ϕ, one easily derives the inequality

$$\begin{aligned} &\int_0^T \mathbf{E}[|\lambda'^{n+1}(t) - \lambda'^n(t)|\, | \mathscr{F}_0^{N'}]\, dt \\ &\qquad \le \left[\int_{\mathbb{R}^+} |h(t)|\, dt\right] \int_0^T \mathbf{E}[|\lambda'^n(t) - \lambda'^{n-1}(t)|\, | \mathscr{F}_0^{N'}]\, dt. \end{aligned}$$

It also ensures that $\lambda'^0(t)$ is less than $\phi(0)+\int_{\mathbb{R}^-}|h(t-s)|N'(ds)$; since N' has initial condition (i), it follows that $\lambda'^0(t)$ is a.s. locally integrable. The above inequality then implies that every $\lambda'^n(t)$, $n>0$, is a.s. locally integrable. Recalling (3), it also ensures the existence of limiting processes $\lambda'^\infty(t)$, N'^∞ such that N'^∞ coincides with N' on $\mathbb{R}^-$ and counts the points of $\bar{N}$ below $\lambda'^\infty(t)$ on $\mathbb{R}^+$, and $\lambda'^\infty(t)$ equals $\phi[\int_{(-\infty,t)}h(t-s)N'^\infty(ds)]$ (the detailed arguments are the same as for the existence of N). This also implies that the integral $\int_0^T \mathbf{E}[\lambda'^\infty(t)|\mathscr{F}_0^{N'}]\,dt$ is a.s. finite, so that $\mathbf{E}[\lambda'^\infty(t)|\mathscr{F}_0^{N'}]$ is a.s. locally integrable. Let us finally check that N'^∞, λ'^∞ coincide with N', λ'. This will follow if the point process $N''=|N'^\infty-N'|$ is such that $N''((0,+\infty))=0$ a.s. Its stochastic intensity $\lambda''(t)$ equals $|\lambda'^\infty(t)-\lambda'(t)|$, and satisfies (by the Lipschitz property of ϕ and since N'^∞ coincides with N' on $\mathbb{R}^-$)

$$\lambda''(t)\le\int_{(0,t)}|h(t-s)|N''(ds).$$

This inequality and Lemma 1 then ensure that $N''((0,+\infty))=0$ a.s. □

PROOF OF THEOREM 2. We first establish the existence of a stationary solution. The proof is then concluded by using Lemma 5 below.

Let $(\Omega,\mathscr{F},\mathbf{P})$ and θ_t be as in the proof of Theorem 3. Construct recursively the $\mathscr{F}_t^{\bar{N}}$-predictable processes $\{\lambda^n(t)\}$ and the p.p. N^n according to

$$N^n(C)=\int_C\bar{N}\left(ds\times\left[0,\frac{\lambda^n(t)}{\Lambda}\right]\right),\qquad C\in\mathscr{B}(\mathbb{R}),\tag{21}$$

$$\lambda^{n+1}(t)=\phi\left[\int_{(-\infty,t)}h(t-s)N^n(ds)\right],\qquad t\in\mathbb{R},\tag{22}$$

the procedure being initialized by $\lambda^0(t)\equiv 0$.

Suppose (as will be proved later) that, for all bounded $C\subset\mathbb{R}^+$, with probability 1 the processes N^n remain eventually constant on C as n increases, and let N denote the limiting process. This process, which is θ_t-compatible, is therefore stationary (and mixing), and will have the expected dynamics if it counts the points of $\bar{N}$ below the curve $t\to\lambda(t)$, where $\lambda(t)$ is given by (1). Let C be some bounded set in $\mathbb{R}$. By Fatou's lemma and an extension of Lemma 3,

$$\begin{aligned}&\mathbf{E}\int_C\left|N(ds)-\bar{N}\left(ds\times\left[0,\frac{\lambda(s)}{\Lambda}\right]\right)\right|\\&\quad\le\lim_{n\to\infty}\mathbf{E}\int_C\left|N^n(ds)-\bar{N}\left(ds\times\left[0,\frac{\lambda(s)}{\Lambda}\right]\right)\right|\\&\quad=\left(\int_C ds\right)\lim_{n\to\infty}\mathbf{E}|\lambda^n(0)-\lambda(0)|.\end{aligned}\tag{23}$$

The Lipschitz property of ϕ ensures

$$|\lambda^n(0) - \lambda(0)| \le \alpha \int_{\mathbb{R}^-} |h(-s)|\,|N^n - N|(ds).$$

The right-hand side of the above inequality tends to 0 as $n \to \infty$; therefore, by Lebesgue's dominated convergence [the dominating variable is $\int_{\mathbb{R}^-} |h(-s)|\bar{N}(ds \times [0,1])$], the first term in (23) is null, so that N has its intensity given by (1). Convergence of the p.p. N^n will hold if the p.p. $\tilde{N}$ defined by

$$\tilde{N}(\{t\}) = \limsup_{n\to\infty} N^n(\{t\}) - \liminf_{n\to\infty} N^n(\{t\}), \qquad t \in \mathbb{R},$$

is a.s. equal to the null measure $\varnothing$. Note that an $\mathscr{F}_t^{\bar{N}}$-intensity of $\tilde{N}$ is

$$\tilde{\lambda}(t) = \limsup_{n\to\infty} \lambda^n(t) - \liminf_{n\to\infty} \lambda^n(t), \qquad t \in \mathbb{R}.$$

Indeed, $\tilde{N}$ counts exactly the points of $\bar{N}$ between the two predictable curves $\{\limsup_{n\to\infty} \lambda^n(t)\}$ and $\{\liminf_{n\to\infty} \lambda^n(t)\}$. Writing $\tilde{\lambda}(t)$ as

$$\lim_{n\to\infty} \sup_{i,\, j\ge n} [\lambda^i(t) - \lambda^j(t)],$$

it follows from the Lipschitz property of ϕ that

$$\begin{aligned}\tilde{\lambda}(t) &\le \alpha \lim_{n\to\infty}\left[\sup_{i\ge n} \int_{(-\infty,\, t)} |h(t-s)| N^i(ds) - \inf_{j\ge n} \int_{(-\infty,\, t)} |h(t-s)| N^j(ds)\right] \\ &= \alpha A + \alpha B,\end{aligned}$$

where

$$A = \lim_{n\to\infty}\left[\sup_{i\ge n} \int_{(-\infty,\, t-a)} |h(t-s)| N^i(ds) - \inf_{j\ge n} \int_{(-\infty,\, t-a)} |h(t-s)| N^j(ds)\right]$$

and B is the similar quantity with the interval $[t-a, t)$ instead of $(-\infty, t-a)$ and where $a > 0$ is arbitrary. Since there are only a finite number of points of $\bar{N}$ involved in B, we have

$$B \le \int_{[t-a,\, t)} |h(t-s)| \tilde{N}(ds),$$

whereas for A we have the immediate bound

$$A \le \int_{(-\infty,\, t-a)} |h(t-s)| \bar{N}(ds \times [0,1]).$$

Therefore, for arbitrary $a > 0$,

$$\tilde{\lambda}(t) \le \alpha\left[\int_{(-\infty,\, t-a)} |h(t-s)| \bar{N}(ds \times [0,1]) + \int_{[t-a,\, t)} |h(t-s)| \tilde{N}(ds)\right].$$

Letting a tend to $+\infty$, we obtain the majoration

$$\tilde{\lambda}(t) \le \alpha \int_{(-\infty,\, t)} |h(t-s)| \tilde{N}(ds). \tag{24}$$

The process $\tilde{N}$ will be a.s. equal to the null measure if $\mathbf{P}(\tilde{N}(0,+\infty)=0)=1$. Since $\{\tilde{N}(t,+\infty)=0\}\subset\{\tilde{N}(0,+\infty)=0\}$, by the ergodicity of $\mathbf{P}$, this will hold if $\mathbf{P}(\tilde{N}(0,+\infty)=0)>0$, which will hold in turn if, with positive probability, $\mathbf{P}(\tilde{N}((0,+\infty)|\mathcal{F}_0^{\tilde{N}})=0)>0$. The $\mathcal{F}_t^{\tilde{N}}$-intensity of $\tilde{N}$, which is of the form $v(t,\tilde{N})$, equals $\mathbf{E}(\tilde{\lambda}(t)|\mathcal{F}_t^{\tilde{N}})$ and therefore satisfies majoration (24); it follows that $v(t,\tilde{N}^-)$ is less than $\int_{\mathbb{R}^-}|h(t-s)|\bar{N}(ds\times[0,1])$, so that, according to Lemma 1,

$$\mathbf{P}(\tilde{N}((0,+\infty)|\mathcal{F}_0^{\tilde{N}})=0)\geq\exp-\left[\int_0^{+\infty}dt\int_{\mathbb{R}^-}|h(t-s)|\bar{N}(ds\times[0,1])\right].$$

The argument of the exponential in the above expression is a.s. finite in view of (4), so that $\mathbf{P}(\tilde{N}(0,+\infty)=0|\mathcal{F}_0^{\tilde{N}})>0$ a.s., which concludes the first part of the proof (existence of the stationary solution).

Let N' be a p.p. with initial condition (5). Use Lemma 4 (this is valid since N', having bounded intensity, is nonexplosive on $\mathbb{R}^+$) to construct a homogeneous Poisson process $\bar{N}$ such that N' counts, on $\mathbb{R}^+$, the points of $\bar{N}$ below

$$t\to\lambda'(t)=\phi\left[\int_{(-\infty,t]}h(t-s)N'(ds)\right].$$

Construct the stationary p.p. N from $\bar{N}$ as in the first part of the proof. Let $\mathcal{F}_t=\mathcal{F}_t^{N'}\vee\mathcal{F}_t^{\bar{N}}$, $t\in\mathbb{R}$. Here $\{\mathcal{F}_t\}$ is a history of both processes N and N'. For all $s,t>0$, note $f_s(t)=\mathbf{P}(|N-N'|(s,s+t)=0|\mathcal{F}_s)$. Note that the p.p. $|N-N'|$ admits $\{|\lambda(t)-\lambda'(t)|\}$ as an $\mathcal{F}_t$-intensity on $\mathbb{R}^+$. The Lipschitz property of ϕ can be used to derive the majoration

$$|\lambda(t)-\lambda'(t)|\leq\alpha\left\{\int_{\mathbb{R}^-}|h(t-s)|N'(ds)+\int_{\mathbb{R}^-}|h(t-s)|\bar{N}(ds)\right.$$
$$\left.+\int_{(0,t)}|h(t-s)|\,|N-N'|(ds)\right\}.$$

Then, arguing exactly as in the proof of Lemma 1, one can get the minoration

$$f_s(+\infty)\geq\exp-\left[\int_s^{+\infty}g_s(u)\,du\right],$$

where

$$g_s(u)=\alpha\left\{\int_{(-\infty,s]}|h(u-v)|\bar{N}(dv)+\int_{\mathbb{R}^-}|h(u-v)|N'(dv)\right\}.$$

Setting

$$Z(s)=\exp-\alpha\int_s^{+\infty}du\int_{(-\infty,s]}|h(u-v)|\bar{N}(dv),$$
$$\varepsilon(s)=Z(s)-\exp-\left[\int_s^{+\infty}g_s(u)\,du\right],$$

the last inequality can be written as

$$\mathbf{P}(N \equiv N' \text{ on } (s, +\infty) | \mathscr{F}_s) \geq Z(s) - \varepsilon(s).$$

Clearly, Z is ergodic since $\bar{N}$ is so; also, it satisfies (26) below because the argument of the exponential defining $Z(s)$ is a.s. finite when (4) is in force. According to (5), it holds that $\int_s^{+\infty} du \int_{\mathbb{R}^-} |h(u-v)| N'(dv)$ tends to 0 a.s. as $s \to \infty$, so that $\varepsilon(s)$ tends to 0 a.s. as $s \to +\infty$. Coupling of N and N' then follows from Lemma 5 below. □

LEMMA 5 (Coupling). *Let $(\Omega, \mathscr{F}, \mathbf{P})$ be some probability space endowed with a filtration $\{\mathscr{F}_t\}_{t \in \mathbb{R}}$, and let X, Y be two $\mathscr{F}_t$-adapted processes. Assume that, for all $s \geq 0$,*

$$\mathbf{P}(X \equiv Y \text{ on } (s, +\infty) | \mathscr{F}_s) \geq Z(s) - \varepsilon(s) \tag{25}$$

for some real-valued process ε, such that $\varepsilon(s)$ tends to 0 a.s. as $s \to \infty$, and some real-valued, ergodic process Z, satisfying

$$\mathbf{P}(Z(s) > 0) > 0. \tag{26}$$

Then the processes X and Y couple in a.s. finite time.

PROOF. According to (26), there exists $\beta > 0$ such that $\mathbf{P}(Z(s) \geq \beta) \geq \beta$. The event $A_s = \{X \equiv Y \text{ on } (s, +\infty)\}$ increases as $s \to +\infty$ to the event $A_\infty = \{X \text{ and } Y \text{ couple}\}$, which is $\mathscr{F}_\infty$-measurable since both processes X and Y are $\mathscr{F}_t$-adapted. Equation (25) therefore implies

$$\mathbf{P}(A_\infty \mid \mathscr{F}_s) \geq \frac{\beta}{2} \mathbf{1}_{[\beta, +\infty)}(Z(s)) \mathbf{1}_{(-\infty, \beta/2]}(\varepsilon(s)).$$

Integrating the above between s and $s+t$, one obtains

$$\frac{1}{t} \int_s^{s+t} \mathbf{P}(A_\infty \mid \mathscr{F}_u)\, du \geq \frac{\beta}{2} \mathbf{1}_{(-\infty, \beta/2]}\Big(\sup_{u \geq s} \varepsilon(u)\Big) \frac{1}{t} \int_s^{s+t} \mathbf{1}_{[\beta, +\infty)}(Z(u))\, du. \tag{27}$$

Since $\mathbf{P}(A_\infty | \mathscr{F}_s)$ is a uniformly integrable martingale, it converges almost surely to $\mathbf{1}_{A_\infty}$ as $s \to \infty$ (see, for instance, [17], Proposition 4.5.6, page 134), and therefore the left-hand side of (27) also converges a.s. to $\mathbf{1}_{A_\infty}$; by the ergodicity of Z, the Cesaro mean on the right-hand side of (27) converges a.s. to $\mathbf{P}(Z(s) \geq \beta)$, which is greater than β; we therefore obtain

$$\mathbf{1}_{A_\infty} \geq \frac{\beta^2}{2} \mathbf{1}_{(-\infty, \beta/2]}\Big(\sup_{u \geq s} \varepsilon(u)\Big).$$

As $s \to \infty$ the indicator on the right-hand side of the above equation tends to 1 a.s. so that the indicator of the event A_∞ is a.s. strictly positive, and is therefore a.s. equal to 1; that is, the processes X and Y couple. □

REMARK 8. It follows from the proof of Theorem 1 that we have solved a stochastic integral system, namely,

$$X_t = \int_{(-\infty,\,t)} h(t-s)N(ds),$$
$$N_t = \int_{(-\infty,\,t]}\int_{\mathbb{R}^+} \mathbf{1}_{(0,\,\phi(X_s))}(y)\bar{N}(ds\times dy).$$

This system is "driven" by the Poisson process $\bar{N}$.

5. Neuron networks (mutually exciting nonlinear Hawkes processes). The results of Section 2 shall now be generalized to a network consisting of K neurons, $K \geq 1$. The activity of the network is summarized by K p.p. $N_1,\dots,N_K$ without common points and with $\mathscr{F}_t$-intensity (2), where $\mathscr{F}_t = \vee \mathscr{F}_t^{N_i}$. For given mappings ϕ_i, h_{ji}, $i, j \in \{1,\dots,K\}$, this fully characterizes the distribution of $(N_1,\dots,N_K)$ (see [8]).

A multidimensional version of Theorem 3 is straightforward.

THEOREM 5. *Let $\psi_1,\dots,\psi_K$ be K mappings from $(M^K, \mathscr{M}^K)$ into $(\mathbb{R}^+, \mathscr{B}(\mathbb{R}^+))$, which are bounded from above by $\Lambda > 0$ and which depend on their arguments through their behavior on $[-A, 0)$ only, for some $A > 0$. Then there exists a unique stationary law for the process $N = (N_1,\dots,N_K)$, where N_i admits $\lambda_i(t) = \psi_i(S_tN_1,\dots,S_tN_K)$ as an $\mathscr{F}_t$-intensity and the N_i do not share common points. Moreover, under these assumptions, the dynamics are stable in variation, irrespective of the initial condition, and convergence in variation is exponentially fast.*

The proof is a straightforward adaptation of that of Theorem 3. The extension of Theorem 4 is more delicate. Here is a partial result concerning two neurons.

THEOREM 6. *Let ϕ_1, ϕ_2 be two nonnegative, nondecreasing functions on $\mathbb{R}$, bounded from above by $\Lambda > 0$. Let h_{11}, h_{22} be two nonnegative integrable functions on $\mathbb{R}^+$, and let h_{12}, h_{21} be two integrable functions on $\mathbb{R}^+$ that are either both nonnegative or both nonpositive. Assume that ϕ_1 is left-continuous. Assume also that ϕ_2 is left-continuous if h_{12}, h_{21} are greater than or equal to 0, and right-continuous if h_{12}, h_{21} are less than or equal to 0. Then there exists a stationary law for the process $N = (N_1, N_2)$ with dynamics (2).*

PROOF. Let $\bar{N}_1, \bar{N}_2$ be mutually independent marked point processes distributed as $\bar{N}$ of Lemma 2. For $i = 1, 2$, construct recursively the processes $\{\lambda_i^n\}$, N_i^n by letting, for all measurable $C \subset \mathbb{R}$ and all $t \in \mathbb{R}$,

$$N_i^n(C) = \int_C \bar{N}_i\left(dt \times \left[0, \frac{\lambda_i^n(t)}{\Lambda}\right]\right), \tag{28}$$

$$\lambda_i^{n+1}(t) = \phi_i\left[\sum_{j=1}^{2}\int_{(-\infty,\,t)} h_{ji}(t-s)N_j^n(ds)\right], \tag{29}$$

the procedure being initialized by $\lambda_1^0(t) \equiv \lambda_2^0(t) \equiv 0$ if h_{12} and h_{21} are greater than or equal to 0, and by $\lambda_1^0(t) \equiv 0$, $\lambda_2^0(t) \equiv \Lambda$ otherwise. These processes are stationary and mixing. Also, by Lemma 2 and Example 4, it is easily seen by induction on n that each N_i^n is $\mathscr{F}_t^{\bar{N}}$-adapted and admits $\{\lambda_i^n\}$ as an $\mathscr{F}_t^{\bar{N}}$-intensity. In the case where h_{12} and h_{21} are greater than or equal to 0, the nondecreasingness of the ϕ_i ensures that $\lambda_i^n(t)$ and $N_i^n(C)$ increase with n for all t, C and i. There exist therefore limiting processes N_i and $\{\lambda_i\}$, the N_i counting the points of $\bar{N}_i$ below the curve $t \to \lambda_i(t)/\Lambda$. According to Lemma 2, the stationary, mixing process $N = (N_1, N_2)$ will have the expected dynamics if (2) holds. However, the monotonicity properties of N_i^n and $\{\lambda_i^n\}$ ensure that

$$\lambda_i^n(t) \le \phi_i\left[\sum_{j=1}^{2} \int_{(-\infty,\, t)} h_{ji}(t-s) N_j(ds)\right], \tag{30}$$

$$\lambda_i(t) \ge \phi_i\left[\sum_{j=1}^{2} \int_{(-\infty,\, t)} h_{ji}(t-s) N_j^n(ds)\right]. \tag{31}$$

Letting n go to ∞ in these inequalities (which is feasible by the left continuity of ϕ_i) yields the result. If h_{21} and h_{21} are less than or equal to 0, the nondecreasingness of the ϕ_i implies this time that $\lambda_1^n(t)$ and $N_1^n(C)$ increase with n while $\lambda_2^n(t)$ and $N_2^n(C)$ decrease with n, for all t and C. The limiting processes N_i and $\{\lambda_i\}$ are still well defined, and the stationary, mixing process N will have the expected dynamics if (2) holds. The monotonicity properties of N_i^n, $\{\lambda_i^n\}$ ensure that (30) and (31) hold for $i = 1$, and (30) and (31) hold with the sign of the inequalities reversed for $i = 2$. The left continuity of ϕ_1 and the right continuity of ϕ_2 ensure that one can let go n to ∞ in these inequalities, which yields the result. □

REMARK 9. There are straightforward extensions of Theorem 6 to arbitrary K. First, if each ϕ_i is bounded, left-continuous and nondecreasing and if each h_{ji} is nonnegative, the increasing construction can be done without modification (this corresponds to a network where each synaptic connection is excitatory). Second, if each ϕ_i is nondecreasing and bounded, left-continuous for $i \le k$ and right-continuous for $i > k$, if each h_{ji} is greater than or equal to 0 for $i, j \le k$ or $i, j > k$, and less than or equal to 0 for $i \le k$, $j > k$ or $i > k$, $j \le k$, then a monotonic construction (corresponding to the case $h_{12}, h_{21} \le 0$ in the proof) is available with $N_i^n, \{\lambda_i^n\}$ increasing for $i \le k$ and decreasing for $i > k$. This corresponds to a network partitioned into two subsets (neurons 1–k and neurons $k+1$–K) such that each synaptic connection between two neurons of the same subset is excitatory, and each connection between two neurons from different subsets is inhibitory.

The extension of Theorem 1 to the multivariate case is a strict extension of Hawkes and Oakes [7] results on linear mutually exciting point processes.

THEOREM 7. *Assume that the function ϕ_i: $\mathbb{R} \to \mathbb{R}^+$ is α_i-Lipschitz, $i = 1, \dots, K$, and that the functions h_{ji}: $\mathbb{R}^+ \to \mathbb{R}$ are such that the $K \times K$ matrix A with entries $a_{ij} = \alpha_i \int_0^{+\infty} |h_{ji}(t)|\, dt$ has a spectral radius strictly less than 1. Then there exists a unique stationary law for a process N with such dynamics and finite average intensity.*

Moreover, the dynamics are stable in distribution with respect to either initial condition (i′) *or* (ii′) *below*:

(i′) $\sup_{t\geq 0} \varepsilon_a(t) < +\infty$ *a.s. and* $\lim_{t\to+\infty} \varepsilon_a(t) = 0$ *a.s. for all* $a > 0$,
(ii′) $\sup_{t\geq 0} \mathbf{E}(\varepsilon_a(t)) < +\infty$ *and* $\lim_{t\to+\infty} \mathbf{E}(\varepsilon_a(t)) = 0$ *for all* $a > 0$,

where

$$\varepsilon_a(t) = \sum_{i,j} \int_{t-a}^{t} ds \int_{\mathbb{R}^-} |h_{ji}(s-u)| N_j(ds).$$

Assume further that, for all $i, j \in \{1, \dots, K\}$,

$$\int_0^{+\infty} t|h_{ji}(t)|\, dt < +\infty. \tag{32}$$

Then the dynamics (2) *are stable in variation with respect to the initial condition*:

$$\text{(iii}'\text{)} \qquad \sum_{i,j} \int_0^{+\infty} dt \int_{\mathbb{R}^-} |h_{ji}(t-s)| N'_j(ds) < +\infty \quad \textit{a.s.}$$

PROOF. Let $\bar{N}_1, \dots, \bar{N}_K$ be K i.i.d. replicates of a bivariate Poisson process on $\mathbb{R}^2$ with intensity 1. Construct recursively the processes $\{\lambda_i^n(t)\}$, N_i^n, $i = 1, \dots, K$, by letting

$$N_i^n(C) = \int_C \bar{N}_i(dt \times [0, \lambda_i^n(t)]), \qquad C \in \mathscr{B}(\mathbb{R}), \tag{33}$$

$$\lambda_i^{n+1}(t) = \phi_i\left[\sum_{j=1}^{K} \int_{(-\infty,t)} h_{ji}(t-s) N_j^n(ds)\right], \qquad t \in \mathbb{R}, \tag{34}$$

the procedure being initialized by setting $\lambda_1^0 \equiv \cdots \equiv \lambda_K^0 \equiv 0$. For all $i \in \{1, \dots, K\}$, the Lipschitz property of ϕ_i yields

$$\mathbf{E}|\lambda_i^{n+1}(0) - \lambda_i^n(0)| \leq \alpha_i \sum_{j=1}^{K} \mathbf{E}|\lambda_j^n(0) - \lambda_j^{n-1}(0)| \int_0^{+\infty} |h_{ji}(t)|\, dt,$$

which we write in vector form as

$$\mathbf{E}\, |\lambda^{n+1}(0) - \lambda^n(0)| \leq A\mathbf{E}|\lambda^n(0) - \lambda^{n-1}(0)|. \tag{35}$$

For some $k \geq 0$, the norm of the matrix A^n is equivalent to $n^k\rho^n$, where ρ is the spectral radius of A (see, e.g., Lemma 1.1, page 223, in [1]). Since, by assumption, $\rho < 1$, (35) implies that $\lambda_i^n(t)$ converges a.s. and in $\mathbb{L}^1$ to a limit

$\lambda_i(t)$. The rest of the proof of the existence of a stationary solution for the dynamics (2) is as in the univariate case.

To prove asymptotic stationarity (stability), let N' be some transient process with the expected dynamics on $\mathbb{R}^+$, and satisfying one of the initial conditions (i′), (ii′) or (iii′). Enlarge the ambient probability space in order to define $\bar{N}_1, \dots, \bar{N}_K$ distributed as above and such that each N'_i counts the points of $\bar{N}_i$ below the curve $t \to \lambda'_i(t) = \phi_i[\sum_j \int_{(-\infty,\,t)} h_{ji}(t-s)N'_j(ds)]$ [the tool for this is Lemma 4; it can be used here since the point processes N'_i are nonexplosive when any of the initial conditions (i′), (ii′) or (iii′) are in force: see the proof of Theorem 1]. Let N denote the stationary solution constructed from $\bar{N}$ as above. For all i, define

$$f_i(t) = \begin{cases} \mathbf{E}|\lambda_i(t) - \lambda'_i(t)|\,|\mathscr{F}_0^{N'}], & \text{if } t \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Use the Lipschitz property of ϕ_i to check that

$$\begin{aligned} f_i(t) &\leq \alpha_i \sum_j \int_{\mathbb{R}^-} |h_{ji}(t-s)|N'_j(ds) \\ &\quad + \lambda_j \int_t^{+\infty} |h_{ji}(s)|\,ds + \int_0^t |h_{ji}(t-s)|f_j(s)\,ds, \end{aligned} \tag{36}$$

where λ_j is the average intensity of N_j. Fix some $a > 0$. Define $F_i(t) = \int_{t-a}^t f_i(s)\,ds$. Integrating the above expression between $t-a$ and t, one gets

$$F_i(t) \leq \varepsilon_i(t) + \alpha_i \sum_j \int_0^t |h_{ji}(t-s)|F_j(s)\,ds,$$

where

$$\varepsilon_i(t) = \alpha_i \sum_j \int_{t-a}^t du \int_{\mathbb{R}^-} |h_{ji}(u-s)|N'_j(ds) + \lambda_j a \int_{t-a}^{+\infty} |h_{ji}(s)|\,ds.$$

When N' satisfies initial condition (i′), $\varepsilon_i(t)$ is a.s. bounded on $\mathbb{R}$, and $F_i(t)$ is bounded on any finite interval (this follows since f_i is a.s. locally integrable, which is shown exactly as in the proof of Theorem 1). It follows then from the last inequality that

$$F_i(t) \leq \sum_{n \geq 0} \sum_j \int_0^t \varepsilon_j(t-s) g_{ij}^n(s)\,ds, \tag{37}$$

where $g_{ij}^0(t) = \mathbf{1}_{\{i=j\}}\delta_0(t)$ and $g_{ij}^{n+1}(t) = \sum_k \alpha_i \int_0^t |h_{ki}(t-s)|g_{kj}^n(s)\,ds$. Equation (37) is the analog of (19), with a Markov renewal measure instead of a renewal measure; see [1], Chapter 10. By the assumption on the spectral radius of A, the right-hand side of (37) is finite; since $\varepsilon_j(t)$ is bounded and tends to 0 a.s. as $t \to +\infty$, it follows by dominated convergence that $F_i(t) \to 0$ as $t \to \infty$. Convergence in distribution and uniqueness of the stationary distribution with finite average intensities are then proven exactly as in Theorem 1.

Assume now that (32) holds. Integrating (36) between 0 and T for some fixed $T > 0$, one obtains the majoration

$$\int_0^T f_i(t)\,dt \le \varepsilon_i + \alpha_i \sum_j \int_0^T f_j(t)\,dt \int_0^{+\infty} |h_{ji}(t)|\,dt,$$

where

$$\varepsilon_i = \alpha_i \sum_j \int_0^{+\infty} dt \int_{\mathbb{R}^-} |h_{ji}(t-s)| N'_j(ds) + \lambda_j \int_0^{+\infty} t|h_{ji}(t)|\,dt.$$

Under (32) and (iii′), $\int_0^T f_i(t)\,dt$ and ε_i are finite for all T. Iterating the above majoration yields (in vector form)

$$\int_0^T f(t)\,dt \le \sum_{n\ge 0} A^n \varepsilon,$$

which implies, letting T go to ∞, that each $\int_0^{+\infty} f_i(t)\,dt$ is finite. One can then conclude as in the proof of Theorem 1 that N' and N couple in a.s. finite time. □

THEOREM 8. *Assume that each function ϕ_i: $\mathbb{R} \to \mathbb{R}^+$ is α_i-Lipschitz and bounded by $\Lambda > 0$. Assume that each h_{ij}: $\mathbb{R}^+ \to \mathbb{R}$ is integrable and such that*

$$\int_0^{+\infty} t|h_{ij}(t)|\,dt < +\infty. \tag{38}$$

Then there exists a unique stationary law for a process $N = \{N_1, \dots, N_K\}$ with dynamics (2), *and these dynamics are stable in variation with respect to the initial condition*

$$\lim_{s\to+\infty} \int_s^{+\infty} du \int_{\mathbb{R}^-} |h_{ji}(u-v)| N'_j(dv) = 0 \quad \textit{a.s.}, \qquad i, j \in \{1, \dots, K\}. \tag{39}$$

The proof is a straightforward adaptation of that of Theorem 2.

EXAMPLE 5 (Example 3 continued). In Example 3, the ϕ_i, $1 \le i \le K$, are not Lipschitz; however, they can be replaced by any Lipschitz functions such that $\phi_i(0) = \lambda_i$ and $\phi_i(x) = 0$ for $x \ge 1$ because the potentials $X_i(t)$ take only integer values. Thus Theorem 7 applies in this case, also.

6. Conclusion. We have studied from the point of view of stability a general model of interacting point processes of the mutually exciting type. The techniques of proof are quite general and may be applied to various extensions of the model, for instance, neuron dynamics with potentials of the type

$$X_i(t) = \sum_{j=1}^K \sum_{n\in\mathbb{Z}} \mathbf{1}_{(-\infty,t]}(T_n^{(j)}) h_{ji}(t - T_n^{(j)}, Z_n^{(j)}),$$

where the $\{Z_n^{(i)}\}_{n\in\mathbb{Z}}$, $1 \le i \le K$, are independent sequences of independent marks. This extension is straightforward and extends the validity of the model of Example 3 to inhibiting spikes θ_{ij} which are random. The details can be found in [16].

Other extensions concern the point process which is thinned, in the basic model a Poisson process. Indeed, Theorems 2 and 4 are relative to thinning, and their extension to non-Poisson processes is presently under investigation by the authors. This extension would include in the model $G/GI/K/0$ queues, for instance, as well as queues with vacations.

Acknowledgment. The authors thank Günter Last for bringing to their attention the work of Kerstan.

REFERENCES

[1] ASMUSSEN, S. (1987). *Applied Probability and Queues*. Wiley, New York.
[2] BACCELLI, F. and BRÉMAUD, P. (1994). *Elements of Queueing Theory*. Springer, New York.
[3] BRÉMAUD, P. (1981). *Point Processes and Queues*. Springer, New York.
[4] BRÉMAUD, P. and MASSOULIÉ, L. (1994). Imbedded construction of stationary point processes and sequences with a random memory. *Queueing Systems* **17** 213–234.
[5] DALEY, D. J. and VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes*. Springer, New York.
[6] HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83–90.
[7] HAWKES, A. G. and OAKES, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Probab.* **11** 493–503.
[8] JACOD, J. (1975). Multivariate point processes: predictable projections, Radon-Nikodym derivatives, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **31** 235–253.
[9] JACOD, J. (1979). *Calcul stochastique et problèmes de martingales*. Springer, Berlin.
[10] KERSTAN, J. (1964). Teilprozesse Poissonscher Prozesse. *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes* 377–403.
[11] LAST, G. (1993). On dependent marking and thinning of point processes. *Stochastic Process. Appl.* **45** 73–94.
[12] LEWIS, P. A. W. and SHEDLER, G. S. (1976). Simulation of nonhomogeneous Poisson processes with log linear rate function. *Biometrika* **63** 501–506.
[13] LINDVALL, T. (1988). Ergodicity and inequalities in a class of point processes. *Stochastic Process. Appl.* **30** 121–131.
[14] LINDVALL, T. (1992). *Lectures on the Coupling Method*. Wiley, New York.
[15] LIPTSER, R. S. and SHIRYAYEV, A. N. (1978). *Statistics of Random Processes, II: Applications*. Springer, New York.
[16] MASSOULIÉ, L. (1995). Stabilité, simulation et optimisation de systèmes à événements discrets. Ph.D. thesis, Univ. Paris-sud.
[17] NEVEU, J. (1965). *Mathematical Foundations of the Calculus of Probability*. Holden-Day, San Francisco.
[18] OGATA, Y. (1981). On Lewis' simulation method for point processes. *IEEE Trans. Inform. Theory* **27** 23–31.

LABORATOIRE DES SIGNAUX
ET SYSTÈMES
CNRS-ESE
PLATEAU DE MOULON
91192 GIF-SUR-YVETTE
FRANCE

CNET
38-40 RUE DU GÉNÉRAL-LECLERC
92131 ISSY-LES-MOULINEAUX
FRANCE