Mathematical Control and Related Fields
Vol. 14, No. 4, December 2024, pp. 1452-1476
doi:10.3934/mcrf.2024038

STOCHASTIC CONTROL FOR DIFFUSIONS WITH SELF-EXCITING JUMPS: AN OVERVIEW

Alain Bensoussan[1] and Benoît Chevalier-Roignant[*2]

[1]International Center for Decision and Risk Analysis
Naveen Jindal School of Management, University of Texas at Dallas
United States of America

[2]emlyon business school, France

Abstract. We provide an overview of continuous-time processes subject to jumps that do not originate from a compound Poisson process, but from a compound Hawkes process. Such stochastic processes allow for a clustering of self-exciting jumps, a phenomenon for which empirical evidence is strong. Our presentation, which omits certain technical details, focuses on the main ideas to facilitate applications to stochastic control theory. Among other things, we identify the appropriate infinitesimal generators for a set of problems involving various (possibly degenerate) cases of diffusions with self-exciting jumps. Compared to higher-dimensional diffusions, we note a degeneracy of the second-order infinitesimal generator. We derive a Feynman-Kac Theorem for a dynamic system driven by such a jump diffusion and also discuss a problem of continuous control of such a system and provide a verification theorem establishing a link between the value function and a novel type of Hamilton-Jacobi-Bellman (HJB) equation.

1. **Introduction.** [15] introduced jump diffusions to model the impact of rare events on security prices. This pioneering work has led to a rich literature in economics and finance studying Lévy processes as a more general model for rare events. [3] provide an overview of Lévy processes stressing how classical techniques of stochastic control theory must be adjusted to account for their specificities, while [8], [14], and [5] stress the applications of Lévy processes in mathematical finance. In these models, the jumps are driven by a Poisson process (or Poisson random measures). The underlying assumption about the Poisson process does not properly account for the possibility of a self-excitement across generations of jumps, i.e., when the arrival rate of jumps is amplified by the existence of recent jumps.

Hawkes processes were introduced by [12] in a context outside the realm of economics and finance. These processes have the ability to capture cascading disasters and are well suited to model time series that exhibit a clustering effect. Clustering effects can also be observed in financial time series. Initially used for earthquake modeling, Hawkes processes have recently been used for applications in insurance [11] and finance [2, 13] including portfolio management [1, 4, 10]. These works have focused primarily on describing key properties of these Hawkes processes, stressing

2020 *Mathematics Subject Classification.* Primary: 49K20, 93E20.
*Corresponding author: Benoît Chevalier-Roignant.

why these probabilistic properties make Hawkes processes well suited for applications in economics and finance. Some of these works [4, 10] comprise elements of stochastic calculus (e.g., a Dynkin formula in [10]) or of stochastic control theory (e.g., the dynamic programming equation of a specific problem in [4]). But, to the best of our knowledge, no extant work discusses in general how techniques of stochastic calculus and stochastic control must be adjusted to account for the specificities of Hawkes processes.

Our manuscript defines this novel class of continuous-time stochastic processes characterized by jumps driven not by a compound Poisson process but by a compound Hawkes process. The primary objective is to define this mathematical concept and emphasize crucial aspects from the standpoint of stochastic calculus (e.g., stochastic integrals, Itô's lemma) and stochastic control theory (e.g., verification theorem). A significant finding is that the Hawkes diffusion strictly sensu is not a Markov process, but only in connection with the stochastic process characterizing the arrivals of jumps. For a large class of Hawkes diffusions, we determine the relevant Markov semigroup and specify its infinitesimal generator. This finding is a key step in establishing a verification theorem linking the value function of a continuous control problem and the solution of a novel type of dynamic programming equation due to the dynamics of the intensity process. We leave out some technical details to streamline the presentation of concepts, instead focusing on key outcomes.

2. **Counting processes – Review.**

This section recalls key results on counting processes, with the ultimate intent to define a subclass, namely Hawkes processes. All proofs are provided in the Appendix. Let $\big(\Omega, \mathcal{A}, \mathbb{P}\big)$ denote a probability space and $\mathbb{E}$ be the expectation operator in this space. Let $\mathcal{F} := (\mathcal{F}_t)_t$ denote a filtration satisfying the usual conditions and on which we consider semimartingales of dimension one.

Square-integrable martingales. Consider a $\mathcal{F}$-martingale $\mu := (\mu_t)_t$ that satisfies $\sup_{t\geq 0} \mathbb{E}|\mu_t|^2 < \infty$. From the Doob-Meyer decomposition theorem, we can write in a unique way $\mu^2 = \nu + \langle \mu, \mu\rangle$, where $\nu := (\nu_t)_t$ is an $\mathcal{F}$-martingale and $\langle \mu, \mu\rangle$ is an adapted increasing process. If $\langle \mu, \mu\rangle$ is continuous, it can be proven that $\langle \mu, \mu\rangle(t) = \lim_{\delta\downarrow 0} \sum_{k=1}^{K} \mathbb{E}\big[\big(\mu_{t_k} - \mu_{t_k-1}\big)^2 \big| \mathcal{F}_{t_{k-1}}\big]$ in the sense of L^1 convergence, where

$$t_0 = 0 < t_1 < \cdots < t_k < \cdots < t_K = t \text{ is a partition of } [0, t] \text{ with } \sup_k \{t_k - t_{k-1}\} \leq \delta. \quad (1)$$

This notation allows us to define, for two square-integrable martingales μ_1 and μ_2, a *polarization* identity, namely

$$\langle \mu_1, \mu_2\rangle := \frac{1}{2}\Big(\langle \mu_1 + \mu_2, \mu_1 + \mu_2\rangle - \langle \mu_1, \mu_1\rangle - \langle \mu_2, \mu_2\rangle\Big). \quad (2)$$

The process $\langle \mu_1, \mu_2\rangle$ is thus the difference between two increasing processes, while

$$\mu_1\mu_2 - \langle \mu_1, \mu_2\rangle \quad \text{is a } \mathcal{F}\text{-martingale}. \quad (3)$$

Equation (2) defines a bilinear form on the space of square-integrable martingales: μ_1 and μ_2 are orthogonal if $\mu_1\mu_2$ is a martingale or $\langle \mu_1, \mu_2\rangle \equiv 0$.

Counting processes. A *counting process* is defined as an adapted process $N := (N_t)_t$ with $N_0 = 0$ almost surely, that is integer valued, increasing and has jumps of amplitude 1. Because N is a right-continuous submartingale, from the Doob-Meyer's decomposition theorem, it can be written uniquely as $N = \mu + \pi$, where μ

is an $\mathcal{F}$-martingale and $\pi := (\pi_t)_t$ is an adapted increasing process. The processes μ and π are called respectively the *compensated process* and the *compensator* of N [16]. For the partition in eq. (1), one can characterize π via the limit

$$\pi_t = \lim_{\delta \downarrow 0} \sum_{k=1}^{K} \mathbb{E}\Big[\Delta N_{t_k} \Big| \mathcal{F}_{t_{k-1}}\Big], \qquad \text{with} \quad \Delta N_{t_k} := N_{t_k} - N_{t_{k-1}}, \tag{4}$$

in the sense of L^1-convergence. One can establish:

Lemma 1. *The increasing process $\langle \mu, \mu \rangle$ associated with the martingale μ is the compensator π.*

Hereafter, we assume the counting process's compensator to be of the form $(\pi_t = \int_0^t \lambda_s \mathrm{d}s)_t$, where the process $\lambda := (\lambda_t)_t$ is adapted, positive and right-continuous with left limits ("càdlàg"). The term λ_t can be interpreted as a hazard or *arrival rate*. It follows from Lemma 1 and eq. (3) that $(\mu_t^2 - \int_0^t \lambda_s \mathrm{d}s)_t$ is a $\mathcal{F}$-martingale.

The Poisson process, which corresponds to the choice of arrival rate $\lambda(\cdot) \equiv \lambda > 0$ and hence of compensator $(\pi_t = \lambda t)_t$ [16], is a well-known counting process. In this context, the martingale μ is often called the compensated Poisson process [16]. The Poisson process has found many applications in economics and finance, e.g., to model firm bankruptcies or other less frequent dramatic events [15]. The Hawkes process (discussed below) proposes an alternative to Poisson processes, again with statistical properties that may make them more consistent with known empirical facts (e.g., the clustering of rare events).

Stochastic integrals. We now want to define stochastic integrals with respect to a counting process, not specifically to a Poisson process. For this purpose, we denote by $L^2_{\mathcal{F},\lambda}(0,T)$ the set of $\mathcal{F}$-adapted stochastic processes $\varphi := (\varphi_t)_t$ such that $\mathbb{E} \int_0^T \varphi_s^2 \lambda_s \mathrm{d}s < \infty$ The space $L^2_{\mathcal{F},\lambda}(0,T)$ is a Hilbert space that is equipped with a norm $||\varphi||^2 = \mathbb{E} \int_0^T \varphi_s^2 \lambda_s \mathrm{d}s$. We first note that:

Lemma 2 (Isometry between $L^2_{\mathcal{F},\lambda}(0,T)$ and $L^2(\Omega, \mathcal{A}, \mathbb{P})$)**.** *For $\varphi \in L^2_{\mathcal{F},\lambda}(0,T)$, it holds*

$$\mathbb{E}\Big(\int_0^T \varphi_s \mathrm{d}\mu_s \Big)^2 = \mathbb{E} \int_0^T \varphi_s^2 \lambda_s \mathrm{d}s. \tag{5}$$

Lemma 2 establishes an isometry between two spaces, namely $L^2_{\mathcal{F},\lambda}(0,T)$ and $L^2(\Omega, \mathbb{A}, \mathbb{P})$. From the decomposition $N = \mu + \int_0^{\cdot} \lambda_s \mathrm{d}s$ and Lemma 2, we can now define a stochastic integral $\int_0^T \varphi_s \mathrm{d}N_s$ by isometry:

$$\int_0^T \varphi_s \mathrm{d}N_s = \int_0^T \varphi_s \mathrm{d}\mu_s + \int_0^T \varphi_s \lambda_s \mathrm{d}s. \tag{6}$$

We note that, for $\varphi \in L^2_{\mathcal{F},\lambda}(0,T)$, it holds $\mathbb{E} \int_0^T |\varphi_s| \lambda_s \mathrm{d}s < \infty$. Consequently, $\int_0^T \varphi_s \mathrm{d}N_s$ is an element of the space $L^1(\Omega, \mathcal{A}, \mathbb{P})$, which contrasts with $\int_0^T \varphi_s \mathrm{d}\mu_s$, which is square integrable (i.e., belongs to $L^2(\Omega, \mathcal{A}, \mathbb{P})$). Clearly, we must have $\mathbb{E} \int_0^T |\varphi_s| \lambda_s \mathrm{d}s < \infty$ for the integral in eq. (6) to be well defined.

Let $\mathbf{1}$. denote the indicator function. We can also define a process $(\xi_t := \int_0^t \varphi_s \mathrm{d}N_s)_t$. This is because ξ_t can be written $\int_0^T \mathbf{1}_{\{s<t\}} \varphi_s \mathrm{d}N_s$, which is a particular case of the stochastic integral in eq. (6). Using differential notation, the dynamics of ξ is given by $\mathrm{d}\xi_t = \varphi_t \mathrm{d}N_t$. We now want to derive a rule similar to the chain rule applicable to such counting processes.

Itô formula. For applications in stochastic calculus and control, determining the dynamics of the process $(\Psi(\xi_t))_t$ is an important issue one wants to address:

Proposition 1 (Itô formula for counting processes)**.** *Let the function* $\Psi(\cdot)$ *be continuous and such that* $\mathbb{E}\int_0^T \left|\Psi(\xi_t+\varphi_t)-\Psi(\xi_t)\right|^2\lambda_t \mathrm{d}t<\infty$*. Then,*

$$\Psi(\xi_t)=\Psi(0)+\int_0^t \Big\{\Psi(\xi_s+\varphi_s)-\Psi(\xi_s)\Big\}\mathrm{d}N_s, \quad \forall t\in[0,T]. \tag{7}$$

Using differential notations, eq. (7) reads $\mathrm{d}\Psi(\xi_t)=\big[\Psi(\xi_t+\varphi_t)-\Psi(\xi_t)\big]\mathrm{d}N_t$. It follows from Lemma 1 that $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\Psi(\xi_t)=\mathbb{E}\big[\{\Psi(\xi_t+\varphi_t)-\Psi(\xi_t)\}\lambda_t\big]$.

3. **Hawkes processes and Hawkes random measures.**

We now define more carefully a new class of counting processes, namely Hawkes processes, and because we are also interested in modeling the magnitude of jumps if jumps occur, we also introduce the notion of Hawkes random measures.

3.1. **Hawkes processes.**

Definition. The Hawkes process, which is presented next, captures a clustering effect for jumps, which makes it suitable for numerous applications. In its simplest form, a *Hawkes process* is defined as a counting process $N=(N_t)_t$ characterized by a compensator which takes the form $(\pi_t=\int_0^t\lambda_s\mathrm{d}s)_t$ with $\lambda:=(\lambda_t)_t$ being the solution of a stochastic differential equation (SDE), namely

$$\begin{cases} \mathrm{d}\lambda_t = & \alpha\big(\lambda_\infty-\lambda_t\big)\mathrm{d}t+\beta\mathrm{d}N_t, \\ \lambda_0 = & \lambda. \end{cases} \tag{8}$$

Note that eq. (8) can be understood as a fixed-point equation because the submartingale N on the right-hand side (RHS) has a Doob-Meyer decomposition which involves $(\int_0^t\lambda_s\mathrm{d}s)_t$ as its compensator. Provided a solution exists, the SDE (8) has an explicit solution given by

$$\lambda_t=e^{-\alpha t}\lambda+\big(1-e^{-\alpha t}\big)\lambda_\infty+\beta\int_0^t e^{-\alpha(t-s)}\mathrm{d}N_s, \quad \forall t>0. \tag{8'}$$

[Indeed, from eq. (8'), $\mathrm{d}\big[e^{\alpha t}\lambda_t\big]=\mathrm{d}\big[\lambda+\big(e^{\alpha t}-1\big)\lambda_\infty+\beta\int_0^t e^{\alpha s}\mathrm{d}N_s\big]$, which is equivalent to $\alpha e^{\alpha t}\lambda_t+e^{\alpha t}\mathrm{d}\lambda_t=e^{\alpha t}\big(\alpha\lambda_\infty\mathrm{d}t+\beta\mathrm{d}N_t\big)$ and yields eq. (8).] The jump intensity in eq. (8') is a convex combination of an initial intensity $\lambda\geq 0$ and a baseline intensity $\lambda_\infty\geq 0$ plus a term that is a linear function of past jumps of the counting process N. The parameter $\alpha>0$ is an exponential decay rate driving the jump intensity back to the long-run average, while $\beta\geq 0$ captures the sensitivity of the jumps on the intensity dynamics. This model is appealing because the accumulation of rare events self-excites new events via the integral term on the RHS of eq. (8'). Two special cases are of interest: (a) if $\alpha\to\infty$ (resp., $\alpha\to 0$), the process $(\lambda_t)_t$ is constant, equal to λ_∞ (resp., λ) between two events, so the counting process becomes a birth process and (b) if $\beta\to 0$, then λ is a deterministic function of time t. We assume $\alpha>\beta$ throughout the paper.

Itô formulas. For the stochastic process λ in eq. (8), it follows from Proposition 1 that for any function $\Psi\in C^{1,1}(\mathbb{R}_+^2)$, we have

$$\mathrm{d}\Psi(\lambda_t,t)=\Big\{\frac{\partial\Psi}{\partial t}(\lambda_t,t)+\frac{\partial\Psi}{\partial\lambda}(\lambda_t,t)\alpha(\lambda_\infty-\lambda_t)\Big\}\mathrm{d}t+\Big\{\Psi(\lambda_t+\beta,t)-\Psi(\lambda_t,t)\Big\}\mathrm{d}N_t \tag{9}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\Psi(\lambda_t,t) = \mathbb{E}\left[\frac{\partial\Psi}{\partial t}(\lambda_t,t) + \frac{\partial\Psi}{\partial\lambda}(\lambda_t,t)\alpha(\lambda_\infty - \lambda_t) + \{\Psi(\lambda_t+\beta,t) - \Psi(\lambda_t,t)\}\,\lambda_t\right].$$

Similarly, we can establish that, for any function $\Psi : \mathbb{R}_+ \times \mathbb{N} \times (0,T) \to \mathbb{R}$ that is continuously differentiable in λ and t, we have

$$\begin{aligned}\mathrm{d}\Psi(\lambda_t,N_t,t) =& \Big\{\frac{\partial\Psi}{\partial t}(\lambda_t,N_t,t) + \frac{\partial\Psi}{\partial\lambda}(\lambda_t,N_t,t)\alpha(\lambda_\infty-\lambda_t)\Big\}\mathrm{d}t \\ &+ \Big\{\Psi(\lambda_t+\beta,N_t+1,t) - \Psi(\lambda_t,N_t,t)\Big\}\mathrm{d}N_t. \end{aligned} \tag{10}$$

Markov process. Proposition 2 establishes that the stochastic process λ can be defined uniquely as a stationary Markov process with an infinitesimal generator given by

$$\mathbb{A}\varphi(\lambda) = \varphi'(\lambda)\alpha\big(\lambda_\infty - \lambda\big) + \big[\varphi(\lambda+\beta) - \varphi(\lambda)\big]\lambda. \tag{11}$$

Note that the operator in eq. (11) is not a local operator because of the second RHS term. For some specific problems, this can pose a particular challenge.

Proposition 2 (Markov property of the intensity λ in eq. (8))**.** *The map $\varphi \mapsto \mathrm{T}_{s,t}\varphi := \mathbb{E}\big[\varphi(\lambda_t)\big|\lambda_s = \cdot\big]$ is a linear map from the space of measurable bounded functions on $\mathbb{R}_+$ onto itself. The family of such maps $(\mathrm{T}_{s,t})_{t\geq s}$ is a Markov semigroup characterized by an infinitesimal operator. In a subset of the infinitesimal generator's domain, namely the functional space $C^{1,1}(\mathbb{R}^2_+)$, the infinitesimal generator coincides with the operator $\mathbb{A}$ given in eq. (11).*

We recall Dynkin's formula [9]: for a Markov process $X := (X_t)_t$ characterized by an infinitesimal operator $\mathfrak{A}$, it holds for suitable functions Ψ that

$$\mathbb{E}\Psi(X_t,t) = \Psi(X,0) + \mathbb{E}\int_0^t \big(\tfrac{\partial}{\partial t} + \mathfrak{A}\big)\Psi(X_t,s)\mathrm{d}s. \tag{12}$$

In our case, the Markov process is the solution of eq. (8) and has an infinitesimal operator given by $\mathbb{A}$ in eq. (11) for sufficiently smooth functions. It thus holds $\mathbb{E}\Psi(\lambda_t,t) = \Psi(\lambda,0) + \mathbb{E}\int_0^t \big(\frac{\partial}{\partial t} + \mathbb{A}\big)\Psi(\lambda_t,s)\mathrm{d}s$.

When both α and β are 0, the process λ in eq. (8) is constant, equal to the initial value $\lambda \in \mathbb{R}_+$. The process N is then a Poisson process with a jump intensity λ. The Poisson process is known to be a Markov process, but this property does not hold for the counting process N of a Hawkes process (for $\alpha, \beta \neq 0$). To recover the Markov property, one has to consider the pair $\{N,\lambda\}$ jointly. To state this result, we first introduce the operator:

$$\bar{\mathbb{A}}\varphi(k,\lambda) = \frac{\partial\varphi}{\partial\lambda}(k,\lambda)\alpha(\lambda_\infty - \lambda) + \Big\{\varphi(k+1,\lambda+\beta) - \varphi(k,\lambda)\Big\}\lambda. \tag{13}$$

Proposition 3 (Markov property of the pair process $\{N,\lambda\}$)**.** *The map*

$$\varphi \mapsto \bar{\mathrm{T}}_{s,t}\varphi := \mathbb{E}\big[\varphi(N_t,\lambda_t)\big|(N_s,\lambda_s) = \cdot\big]$$

is a linear map from the space of countable measurable bounded functions onto itself. The family $(\bar{\mathrm{T}}_{s,t})_{t\geq s}$ is a Markov semigroup, for which the infinitesimal generator in the case of sufficiently smooth functions is given in eq. (13).

It thus follows from eq. (12) and Proposition 3 that

$$\mathbb{E}\Psi(N_t,\lambda_t,t) = \Psi(0,\lambda,0) + \mathbb{E}\int_0^t \big(\tfrac{\partial}{\partial t} + \bar{\mathbb{A}}\big)\Psi(N_s,\lambda_s,s)\mathrm{d}s.$$

For suitable choices of functions (e.g., $\Psi(k,\lambda,t)=\lambda^m$ for $m\in\mathbb{N}$), we can use the above formula to derive several important probabilistic properties of Hawkes processes $\{N,\lambda\}$ including their moments (see, e.g., Theorem 3.1 in [7] or Corollary 1 in [6]). (See also Appendix K for estimates of the moments of the random variable λ_t.)

3.2. **Hawkes random measures.** Counting processes can be used to model the arrival of shocks in dynamic systems. To model the magnitude of these shocks requires an extension. As usual, we make the assumption that the intensity is independent of the arrival process, for which we assume a Hawkes process.

The magnitude is a vector in $\mathbb{R}^p$. We consider the Borel σ-algebra $\mathcal{B}$ on $\mathbb{R}^p$ and $\mathcal{B}_0$ the Borel σ-algebra on $\mathbb{R}^p\setminus\{0\}$. We define a *random measure* N as a measurable map on $[0,T]\times\mathcal{B}$ such that $N(\cdot,A)$ is a counting process for any event A in $\mathcal{B}$ and the function $A\mapsto N(t,A)$ is additive with $N(t,\{0\})=0$ for any $t\geq 0$. We use the shorthand notation $N_A(t)$ for $N(t,A)$ and note that $N_t=N_{\mathbb{R}^p}(t)$. We define a probability measure $\mathrm{m}(\mathrm{d}z)$ on $\{\mathbb{R}^p,\mathcal{B}\}$ with $\mathrm{m}(\{0\})=0$. The probability space $(\mathbb{R}^p,\mathcal{B},\mathrm{m})$ is independent of $(\Omega,\mathcal{A},\mathbb{P})$ and is used to model the magnitude of the shocks.

For any $A\in\mathcal{B}$, N_A has the unique Doob-Meyer decomposition, namely $N_A=\mu_A+\pi_A$, where μ_A is an $\mathcal{F}$-martingale—called a *martingale measure*—and π_A is an adapted increasing process—called *Hawkes random measure* and defined by $\pi_A(t)=\int_0^t\lambda_s\mathrm{d}s\,\mathrm{m}(A)$. As in Lemma 1, we associate an increasing process with the martingale μ_A, namely

$$\langle\mu_A,\mu_A\rangle(t)=\int_0^t\lambda_s\mathrm{d}s\,\mathrm{m}(A). \tag{14}$$

We obtain the following properties for martingale measures:

Lemma 3. *If two events A and B in $\mathcal{B}_0$ are such that $A\cap B=\emptyset$, then the martingale measure $\mu_{A\cup B}$ satisfies $\mu_{A\cup B}=\mu_A+\mu_B$ and the measures μ_A and μ_B are orthogonal.*

Stochastic integrals. We are now interested in defining stochastic integrals with respect to Hawkes random measures. This integral differs from stochastic integrals with respect to Poisson random measures [5]. For our purposes, let $L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T)\times\mathbb{R}^p)$ denote the set of $\mathcal{F}$-adapted random fields φ such that

$$\mathbb{E}\int_0^T\int_{\mathbb{R}^p}\varphi^2(s,z)\lambda_s\mathrm{d}s\,\mathrm{m}(\mathrm{d}z)<\infty. \tag{15}$$

This space is a Hilbert space is equipped with a norm $||\varphi||^2=\mathbb{E}\int_0^T\int_{\mathbb{R}^p}\varphi^2(s,z)\lambda_s\mathrm{d}s\,\mathrm{m}(\mathrm{d}z)$. To define the stochastic integral $\int_0^T\int_{\mathbb{R}^p}\varphi(s,z)\mathrm{d}N(s,z)$, we proceed as before by establishing an isometry:

Lemma 4 (Isometry between $L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T)\times\mathbb{R}^p)$ and $L^2(\Omega,\mathcal{A},\mathbb{P})$)**.** *For any $\varphi\in L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T)\times\mathbb{R}^p)$, we have*

$$\mathbb{E}\left(\int_0^T\int_{\mathbb{R}^p}\varphi(s,z)\mathrm{d}\mu(s,z)\right)^2=\mathbb{E}\int_0^T\int_{\mathbb{R}^p}\varphi^2(s,z)\lambda_s\mathrm{d}s\,\mathrm{m}(\mathrm{d}z). \tag{16}$$

Lemma 4 generalizes a result known for Poisson random measures (see Theorem 38 in [16] or Theorem 8.7 in [5]) to a case with a Hawkes random measure. Thanks to Lemma 4, we can define a stochastic integral $\int_0^T \int_{\mathbb{R}^p} \varphi(s,z) \mathrm{d}N(s,z)$ as

$$\int_0^T \int_{\mathbb{R}^p} \varphi(s,z) \mathrm{d}N(s,z) = \int_0^T \int_{\mathbb{R}^p} \varphi(s,z) \lambda_s \mathrm{d}s\, \mathrm{m}(\mathrm{d}z) + \int_0^T \int_{\mathbb{R}^p} \varphi(s,z) \mathrm{d}\mu(s,z), \quad (17)$$

where the first term in eq. (17) is an ordinary integral, while the second term is defined thanks to the isometry established in Lemma 4. For a random field φ in $L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T) \times \mathbb{R}^p)$, we define the stochastic process $\xi = (\xi_t)_t$ by

$$\xi_t := \int_0^T \int_{\mathbb{R}^p} \mathbf{1}_{[0,t)}(s) \varphi(s,z) \mathrm{d}N(s,z), \quad 0 \leq t \leq T, \quad (18)$$

which is a special case of eq. (17).

Itô formula. Again, determining the dynamics of $(\Psi(\xi_t))_t$ is an interesting issue:

Lemma 5 (Ito's formula for Hawkes random measures)**.** *Let ξ be the stochastic process defined in eq.* (18)*. If the function Ψ is continuous and such that*

$$\int_0^T \int_{\mathbb{R}^p} \left[\Psi(\xi_s + \varphi(s,z)) - \Psi(\xi_s) \right]^2 \lambda_s \mathrm{d}s\, \mathrm{m}(\mathrm{d}z) < \infty, \quad (19)$$

then

$$\Psi(\xi_t) = \Psi(0) + \int_0^t \int_{\mathbb{R}^p} \Big\{ \Psi(\xi_s + \varphi(s,z)) - \Psi(\xi_s) \Big\} \mathrm{d}N(s,z).$$

It follows from Lemma 5 that, for Ψ continuously differentiable in λ and t, we have

$$\begin{aligned} \mathrm{d}\Psi\left(\lambda_t, N_t^A, t\right) = & \Big\{ \frac{\partial \Psi}{\partial t}(\lambda_t, N_t^A, t) + \frac{\partial \Psi}{\partial \lambda}(\lambda_t, N_t^A, t) \alpha(\lambda_\infty - \lambda_t) \Big\} \mathrm{d}t \\ & + \Big\{ \Psi(\lambda_t + \beta, N_t^A + 1, t) - \Psi(\lambda_t, N_t^A, t) \Big\} \mathrm{d}N_t^A. \end{aligned} \quad (20)$$

Markov process. Similarly to Proposition 3, we want to prove the Markov property of the pair $\{N_A, \lambda\}$. We define the operator

$$\hat{\mathbb{A}}\varphi(N^A, \lambda) = \frac{\partial \varphi}{\partial \lambda}(N^A, \lambda) \alpha(\lambda_\infty - \lambda) + \Big\{ \varphi(N^A + 1, \lambda + \beta) - \varphi(N^A, \lambda) \Big\} \lambda \mathrm{m}(A) \quad (21)$$

Proposition 4 (Markov property of the pair process $\{\lambda, N_A\}$)**.** *The map*

$$\varphi \mapsto \hat{\mathrm{T}}^A_{s,t}\varphi := \mathbb{E}\Big[\varphi\left(N_t^A, \lambda_t\right) \Big| (N_s^A, \lambda_s) = \cdot \Big]$$

is a linear map from the set of measurable bounded random fields onto itself. The family $(\hat{\mathrm{T}}_{s,t})_{t \geq s}$ is a Markov semigroup with an infinitesimal operator given in eq. (21)*.*

4. **Continuous-time processes with self-exciting jumps.** For the sake of applications, one often wants to consider a mixture of a stochastic process with continuous sample paths and a process that fully models the arrival and magnitude of jumps. This is the task that we are now turning to.

4.1. **Hawkes-Itô processes.** We consider a probability space $(\Omega, \mathcal{A}, P)$ equipped with a filtration $\mathcal{F}$ and on which a standard Wiener process $W = (W_t)_t$ taking value in $\mathbb{R}^n$ and a Hawkes random measure $\{N_A, \lambda\}$ are defined. Furthermore, consider stochastic processes α, β, and φ satisfying

$$\alpha \in L^2_{\mathcal{F}}(0,T), \quad \beta \in L^2_{\mathcal{F}}(0,T;\mathbb{R}^n), \quad \text{and} \quad \varphi \in L^2_{\mathcal{F},\lambda m}((0,T)\times\mathbb{R}^p). \tag{22}$$

We consider a stochastic process ξ characterized by a SDE of the form

$$\xi_t = \xi_0 + \int_0^t \alpha_s \mathrm{d}s + \int_0^t \beta_s \cdot \mathrm{d}W_s + \int_0^t \int_{\mathbb{R}^p} \varphi(s,z)\mathrm{d}N(s,z). \tag{23}$$

Equation (23) does not define a Itô-Lévy process because the third integral on the RHS is a stochastic integral defined thanks to an Hawkes random measure, not a Poisson random measure. We call it a *Itô-Hawkes process* by analogy. Furthermore, this equation does not define a jump-diffusion because the RHS term does not depend on the process itself. The process in eq. (18) is a degenerate case of the process in eq. (23) (with $\alpha \equiv 0$ and $\beta \equiv 0$).

We have

Proposition 5 (Itô formula for Hawkes-Itô processes)**.** *Consider the stochastic process ξ in eq. (23) and let $\Psi \in C^{2,1}(\mathbb{R}\times[0,T])$ be such that*

$$\mathbb{E}\int_0^T \left[\frac{\partial \Psi}{\partial x}(\xi_s,s)\right]^2 |\beta_s|^2 \mathrm{d}s < \infty \text{ and } \mathbb{E}\int_0^T \int_{\mathbb{R}^p} \Big|\Psi(\xi_s+\varphi(s,z),s) - \Psi(\xi_s,s)\Big| \lambda_s \mathrm{d}s\, \mathrm{m}(\mathrm{d}z) < \infty. \tag{24}$$

Then,

$$\begin{aligned}\Psi(\xi_t,t) =& \Psi(\xi_0,0) + \int_0^t \Big\{\frac{\partial \Psi}{\partial s}(\xi_s,s) + \alpha_s \frac{\partial \Psi}{\partial x}(\xi_s,s) + \frac{1}{2}\frac{\partial^2 \Psi}{\partial x^2}(\xi_s,s)|\beta_s|^2\Big\}\mathrm{d}s + \\ &+ \int_0^t \frac{\partial \Psi}{\partial x}(\xi_s,s)\beta_s \cdot \mathrm{d}W_s + \int_0^t \int_{\mathbb{R}^p} \Big\{\Psi(\xi_s+\varphi(s,z),s) - \Psi(\xi_s,s)\Big\}\mathrm{d}N(s,z).\end{aligned} \tag{25}$$

Proposition 5 states an Itô-type formula in case of a Hawkes-Itô process of dimension one. The formula can be extended to an arbitrary larger dimension $n \in \mathbb{N}$.

Finally, one can generalize eq. (25) to sufficiently regular functions of the form $(\xi, \lambda, t) \mapsto \Psi(\xi, \lambda, t)$ obtaining

$$\begin{aligned}\Psi(\xi_t,\lambda_t,t) =& \Psi(\xi_0,\lambda_0,0) \\ &+ \int_0^t \Big\{\frac{\partial \Psi}{\partial s}(\xi_s,\lambda_s,s) + \alpha_s \frac{\partial \Psi}{\partial x}(\xi_s,\lambda_s,s) + \frac{1}{2}\frac{\partial^2 \Psi}{\partial x^2}(\xi_s,\lambda_s,s)\big|\beta_s\big|^2\Big\}\mathrm{d}s \\ &+ \int_0^t \frac{\partial \Psi}{\partial \lambda}(\xi_s,\lambda_s,s)\alpha(\lambda_\infty - \lambda_s)\mathrm{d}s + \int_0^t \frac{\partial \Psi}{\partial x}(\xi_s,s)\beta_s \cdot \mathrm{d}W_s \\ &+ \int_0^t \int_{\mathbb{R}^p} \Big\{\Psi(\xi_s+\varphi(s,z),\lambda_s+\beta,s) - \Psi(\xi_s,\lambda_s,s)\Big\}\mathrm{d}N(s,z).\end{aligned} \tag{26}$$

Compared to the Itô formula in case of Lévy processes [5], the formula in eq. (26) comprises a component related to the dynamics of the intensity process $(\lambda_s)_{0\le s\le t}$ in eq. (8).

4.2. **Hawkes-diffusion processes.** Because our ultimate goal is to adjust functional techniques used for the study of stochastic control problems driven by dynamic systems involving a Hawkes process, we introduce further structure on the underlying dynamics, defining a new class of processes, which we call *Hawkes-diffusion processes*. We define Hawkes diffusions as diffusions which have jumps driven by a Hawkes process. Let * denote the transpose and tr stand for the trace operator, i.e., the sum of the elements on the main diagonal. Furthermore, we introduce the notations $a(x,s) := \sigma(x,s)\sigma^*(x,s)$ and

$$D_x := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \quad \text{and} \quad D_x^2 := \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} & \cdots & \frac{\partial}{\partial x_1 \partial x_n} \\ \vdots & & \\ \frac{\partial^2}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2}{\partial x_n^2} \end{pmatrix}.$$

We make usual assumptions about the coefficients of our dynamical system:

Assumption 1 (Jump-diffusion coefficients)**.** *The drift g, diffusion σ, and jump sizes γ are maps such that $g : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \times [0,T] \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, and $\gamma : \mathbb{R}^n \times [0,T] \times \mathbb{R}^p \to \mathbb{R}^n$. Furthermore, there exists a positive constant C such that*

$$\left|g(x,t)\right|^2 + \mathrm{tr}\left\{\sigma(x,t)\sigma^*(x,t)\right\} \le C\left(1+|x|^2\right), \tag{27a}$$

$$\int_{\mathbb{R}^p} |\gamma(x,t,z)|^2 \mathrm{m}(\mathrm{d}z) \le C\left(1+|x|\right), \tag{27b}$$

$$\left|g(x,t)-g(y,t)\right|^2 + \mathrm{tr}\left\{\left[\sigma(x,t)-\sigma(y,t)\right]\left[\sigma(x,t)-\sigma(y,t)\right]^*\right\} \le C|x-y|^2, \tag{27c}$$

$$\int_{\mathbb{R}^p} |\gamma(x,t,z)-\gamma(y,t,z)|^2 \mathrm{m}(\mathrm{d}z) \le C|x-y|, \tag{27d}$$

for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$.

Assumption 1 is relatively classical and remiscent of assumptions also made for jump diffusions. The first condition on the drift and diffusion coefficients can be interpret as a condition of "at most linear growth," while the second inequality prescribes Lipschitz continuity. This assumption allows us to define a diffusion that has jumps that are driven by a Hawkes process:

Proposition 6 (Hawkes-diffusion process)**.** *Under Assumption 1, there exists one and only one solution $X := (X_s)_s$ of the stochastic (integro)differential equation (SDE) given by*

$$X_s = x + \int_t^s g(X_\tau,\tau)\mathrm{d}\tau + \int_t^s \sigma(X_\tau,\tau)\mathrm{d}W_\tau + \int_t^s \int_{\mathbb{R}^p} \gamma(X_\tau,\tau,z)\mathrm{d}N(\tau,z). \tag{28}$$

Equation (28) differs from the Lévy SDE because of the way the stochastic integral is defined, i.e., thanks to an Hawkes random measure, not a Poisson random measure. From the decomposition $N(t,A) = \mu(t,A) + \int_0^t \lambda_s \mathrm{d}s \mathrm{m}(A)$, we can write eq. (28) as

$$\begin{aligned} X_s =& x + \int_t^s g(X_\tau,\tau)\mathrm{d}\tau + \int_t^s \sigma(X_\tau,\tau)\mathrm{d}W_\tau + \int_t^s \int_{\mathbb{R}^p} \gamma(X_\tau,\tau,z)\lambda_\tau \mathrm{d}\tau\,\mathrm{dm}(z) \\ &+ \int_t^s \int_{\mathbb{R}^p} \gamma(X_\tau,\tau,z)\mathrm{d}\mu(\tau,z). \end{aligned} \tag{28'}$$

Markov process. The solution X of eq. (28) is not a Markov process, but Proposition 7 proves that the pair process $\{X,\lambda\}$ turns out to be Markovian. Define the operator

$$\begin{aligned}\tilde{\mathbb{A}}f(x,\lambda,s) :=& \frac{1}{2}\mathrm{tr}\left\{D_x^2 f(x,\lambda,s)a(x,s)\right\}\\ &+ D_x f(x,\lambda,s)\cdot g(x,s) + \alpha(\lambda_\infty-\lambda)\frac{\partial f}{\partial\lambda}(x,\lambda,s)\\ &+ \lambda\int_{\mathbb{R}^p}\Big\{f\Big(x+\gamma(x,s,z),\lambda+\beta,s\Big) - f(x,\lambda,s)\Big\}\,\mathrm{m}(\mathrm{d}z). \qquad (29)\end{aligned}$$

Proposition 7 (Markov property of the pair process $\{X,\lambda\}$)**.** *The map*

$$\varphi \mapsto \tilde{\mathrm{T}}_{s,t}\varphi := \mathbb{E}\Big[\varphi(X_t,\lambda_t)\Big|(x,\lambda)=\cdot\Big]$$

is a linear map from the set of measurable and bounded functions onto itself. The family $(\tilde{\mathrm{T}}_{s,t})_{t\geq s}$ defines a Markov semigroup with an infinitesimal operator $\tilde{\mathbb{A}}$ given in eq. (29)*.*

In Proposition 3, the time-dependence of the infinitesimal generator comes from the time dependence of the functions for the drift μ, diffusion σ, and jump magnitude γ. Again, Proposition 7 allows us to derive a Dynkin formula, using the integro-differential operator $\tilde{\mathbb{A}}$ for suitable functions.

5. **Stochastic control of Hawkes-diffusion processes.** Again, our ultimate goal is to study how to adjust classical techniques for the study of stochastic control problems in cases in which the underlying dynamic system is subject to self-exciting jumps. For that purpose, we consider a problem of continuous control, which is quite standard except for the inclusion of a Hawkes process driving self-exciting jumps. Specifically, we consider the dynamic system $\{X,\lambda\}$ where X is the Hawkes diffusion in eq. (28) and λ solves eq. (8). We recall from Proposition 7 that $\{X,\lambda\}$ is a Markov process with an infinitesimal generator $\tilde{\mathbb{A}}$ given in eq. (29). We start by establishing a Feynman-Kac theorem applicable in this context before we derive an appropriate verification theorem for a problem of continuous control.

5.1. **Feynman-Kac theorem for Hawkes diffusions.** Consider two deterministic functions, namely $f:\mathbb{R}^n\times[0,T]\to\mathbb{R}$ and $h:\mathbb{R}^n\to\mathbb{R}$. The stochastic process $\big(f(X_t,t)\big)_t$ denotes the running cost of an economic agent, while the random variable $h(X_T)$ denotes the unknown terminal value (e.g., decommissioning cost) accruing to that agent at the end of a planning horizon of length $T>0$. For now, we assume that the economic agent does not control the evolution of the dynamic system. We make the following restrictions on these two functions:

Assumption 2. *There exists scalars $\bar{f}>0$ and $\bar{g}>0$ such that*

$$0\leq f(x,s)\leq\overline{f} \quad \textit{and} \quad 0\leq h(x)\leq\overline{h} \quad \textit{for all } x\in\mathbb{R} \textit{ and } s\in[0,T].$$

This assumption sets bounds on the cost incurred by the economic agent. Making this assumption also allows us to derive a key theorem, namely:

Theorem 1 (Feynman-Kac Theorem for Hawkes diffusions)**.** *We make Assumptions 1 and 2 and assume that there exists a scalar $\bar{a}>0$ such that $a(x,s)\geq\bar{a}I$ for all x and s. The function*

$$(x,\lambda,t)\mapsto V(x,\lambda,t) := \mathbb{E}\left[\int_t^T f(X_s,s)\mathrm{d}s + h(X_T)\middle| X_t=x,\lambda_t=\lambda\right] \qquad (30)$$

solves the PDE

$$\begin{cases} \tilde{\mathbb{A}}V(x,\lambda,s) &= f(x,s), \\ V(x,\lambda,T) &= h(x). \end{cases} \tag{31}$$

The Feyman-Kac Theorem, which we here generalized to a new type of Markov processes, is very important and useful in stochastic analysis because it allows one to interpret an expected value $V(x,\lambda,t)$ as the solution of a partial (integro)differential equation and to leverage the realm of techniques in functional analysis to solve the related PDE. In the context of mathematical finance, this theorem would permit to represent the fair price of a European option in two manners, namely as the expected option payoff under a suitable probability measure or as the solution to a partial differential equation, which in essence is a generalization of the Black-Scholes-Merton PDE. (In our context, as for Levy processes, the market is incomplete because the underlying uncertainty is driven by two stochastic factors. One would need to specify a mechanism to determine the appropriate measure among a large set of equivalent martingale measures if need be.) Naturally, time preferences can be captured by considering a running cost of the form $f(x,s) = e^{\int_0^s r(\tau)\mathrm{d}\tau}F(x)$, where $r(\cdot)$ is a given deterministic function of time and $F(x)$ is a known deterministic function of the process state x.

5.2. **Continuous control of Hawkes diffusions.** We now consider a continuous/regular control understood as a stochastic process $(v_s)_{s\geq 0}$ adapted to the filtration $\mathcal{F}$ and which take values in $\mathbb{R}^d$. This control variable $v \in \mathbb{R}^d$ is assumed to affect the drift of the dynamic system $g(x,v,s)$ and the magnitude of the self-exciting jumps $\gamma(x,v,s,z)$. In other words, we consider the controlled dynamic system

$$\begin{cases} \mathrm{d}X_s = & g(X_s,v_s,s)\mathrm{d}s + \sigma(X_s,s)\mathrm{d}W_s + \int_{\mathbb{R}^p} \gamma(X_s,v_s,s,z)\mathrm{d}N(s,z) \\ \mathrm{d}\lambda_s = & \alpha[\lambda_\infty - \lambda_s]\mathrm{d}s + \beta \mathrm{d}N_s, \; s > t \\ X_t = & x, \\ \lambda_t = & \lambda. \end{cases} \tag{32}$$

Another departure from the setting considered in Section 5.1 is that the economic agent's choice about the control variable $v \in \mathbb{R}^d$ affects its running cost $f(x,v,s)$. The control v also affects the terminal value, but only indirectly by affecting the distribution of the random variable X_T.

Our setting for the control problem makes simplifying assumptions for us to focus on key results. For instance,

Remark 1 (Control on the diffusion term)**.** The diffusion term $\sigma(x,s)$ in eq. (32) is not affected by the control variable. By doing so, we avoid at a later point having to deal with a fully nonlinear PDE. Relaxing this assumption would require introducing notions from the more advanced viscosity theory of stochastic control.

Remark 2 (Strong solutions of SDEs)**.** The stochastic differential eq. (32) is understood in a strong sense. For applications in stochastic control theory, this interpretation makes us take strong assumptions about the model primitives, e.g., about the regularity of the solution of the HJB equation. We need a weak-sense theory to avoid such an assumption. In the case of Poisson processes instead of Hawkes processes, it is partly done by [3].

In the spirit of Assumptions 1 and 2, we introduce some restrictions

Assumption 3 (Assumptions on primitives). *There exist $\bar{f} > 0$, $\bar{g} > 0$, and $C > 0$ are such that:*

$$0 \le f(x,v,s) \le \overline{f} \quad \textit{and} \quad \big|g(x,v,s)\big| \le \overline{g} \tag{33}$$

$$\int_{\mathbb{R}^p} |\gamma(x,v,s;z)|^2 \mathrm{m}(\mathrm{d}z) \le C(1+|x|) \tag{34}$$

$$\int_{\mathbb{R}^p} |\gamma(x,v,s,z) - \gamma(y,v,s,z)|^2 \mathrm{m}(\mathrm{d}z) \le C|x-y|. \tag{35}$$

We now consider an objective functional given by

$$v = (v_s)_{s\in(t,T)} \mapsto J^{x,\lambda,t}(v) = \mathbb{E}\left[\int_t^T f(X_s, v_s, s)\mathrm{d}s + h(X_T)\middle| X_t = x, \lambda_t = \lambda\right], \tag{36}$$

which the economic agent wants to minimize. The *value function* is given by

$$V(x,\lambda,t) = \inf_{v=(v_s)_{s\in(t,T)}} J^{x,\lambda,t}(v). \tag{37}$$

(Maximization problems could be treated using the method described below after some adjustments.)

In the spirit of dynamic programming, our goal is to relate the value function in eq. (37) to a dynamic programming equation. Standard heuristics arguments leveraging the infinitesimal generator of our Markov process would lead us to consider the HJB equation given by

$$\begin{cases} \frac{\partial V}{\partial s}(x,\lambda,s) + \frac{1}{2}\mathrm{tr}\big\{D_x^2 V(x,\lambda,s)\, a(x)\big\} + \alpha \frac{\partial V}{\partial \lambda}(x,\lambda,s)[\lambda_\infty - \lambda] + \inf_v \mathcal{H}(x,\lambda,v,s) & = 0, \\ V(x,\lambda,T) & = h(x) \end{cases} \tag{38a}$$

where

$$\begin{aligned} \mathcal{H}(x,\lambda,v,s) :=& f(x,v,s) + D_x V(x,\lambda,s)\cdot g(x,v,s) \\ &+ \lambda \int_{\mathbb{R}^p} \big\{V(x+\gamma(x,v,s;z), \lambda+\beta, s) - V(x,\lambda,s)\big\}\mathrm{m}(\mathrm{d}z). \end{aligned} \tag{38b}$$

Equation (38) is a novel type of HJB equation, specific to our setting with Hawkes diffusions. A rigorous connection between the HJB equation and the value function can be established once we prove the existence of a solution of the dynamic-programming equation (including regularity properties), possibly by construction, and provide an appropriate verification theorem.

To obtain such said regularity, we assume:

Assumption 4 (Regularity.). *If the function V on the RHS of eq. (38b) is continuously differentiable in x and Lipschitz continuous in λ, then there exists a $\hat{v}(x,\lambda,s)$ such that*

$$\mathcal{H}(x,\lambda,\hat{v}(x,\lambda,s),s) = \inf_v \mathcal{H}(x,\lambda,v,s). \tag{39}$$

Furthermore, assume that $\hat{v}(x,\lambda,s)$ is Lipschitz continuous in x and continuous in λ.

We prove the following result:

Theorem 2 (Verification theorem in case of Hawkes diffusions). *We make Assumptions 3 and 4 and assume that $a(x,s) \ge \overline{a}I$ for all (x,s). There exists one and only one solution of eq. (38), which coincides with the value function in eq. (37).*

Theorem 2 is useful and critical for applications. In fact, it says that, to solve the dynamic optimization problem in eq. (37), it suffices to solve the static optimization problem in eq. (39). Furthermore, if the solution $\hat{v}(x,\lambda,t)$ of this static optimization problem satisfies certain regularity properties specified in Assumption 4, then the HJB eq. (38) has a classical solution, which coincides with the value function in eq. (37). Obviously, we made restrictions on model primitives to obtain this key result:

Remark 3 (Boundedness of cost functions)**.** The key result in Theorem 2 rests on the boundedness of the functions f and g as per Assumption 3. In many economic and financial applications, this assumption does not hold. To accommodate these cases would require a theory for non-bounded functions, similarly to what has been done for diffusions.

6. Conclusion.

Hawkes diffusions have been introduced initially in the context of earthquake modeling but have been increasingly used in economics and finance. Such processes have neat properties, validated by empirical evidence (especially a clustering of jump events). This paper summarizes key properties and theorems (including Dynkin's formulas, Itô's lemmas, Feynman-Kac theorem, and a verification theorem for continuous controls) with the view of allowing applications in the context of stochastic control theory. We lay out the mathematical foundations that would prove useful as this stream of literature is emerging.

Many extensions are left for future research. For instance, for convenience, we made restrictions on the regularity (and boundedness) of certain functions. A generalization thanks to viscosity techniques may prove useful. We also just discussed one type of dynamic programming equation, but other stochastic control problems may be characterized by other types of dynamic programming equations (e.g., variational inequalities, quasivariational inequalities).

Acknowledgments. The authors would like to congratulate Helene Frankowska on the occasion of her 70th birthday. Alain Bensoussan acknowledges the support of the National Science Foundation under grant NSF-DMS 220 47 95. The authors thank Alexey Minabutdinov for raising questions about Hawkes processes.

Appendix A. **Proof of Lemma 1.** Denote $\Delta\mu_{t_k} = \Delta N_{t_k} - \Delta\pi_{t_k}$. We must check that $\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\mathbb{E}\big[\Delta\mu_{t_k}^2\big|\mathcal{F}_{t_{k-1}}\big] = \pi_t$ in L^1.

Step 1. The jumps of $(N_t)_t$ have an amplitude 1, so necessarily

$$\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\Delta N_{t_k}^2 \to N_t, \text{ a.s.} \tag{40}$$

Because, also,

$$\sum_{k=1}^{K}\Delta N_{t_k}^2\mathbf{1}_{\{\Delta N_{t_k}\leq 1\}} = \sum_{k=1}^{K}\Delta N_{t_k}\mathbf{1}_{\{\Delta N_{t_k}\leq 1\}}$$

we can state that

$$\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\Delta N_{t_k}^2\mathbf{1}_{\{\Delta N_{t_k}\leq 1\}} = \lim_{\delta\downarrow 0}\sum_{k=1}^{K}\Delta N_{t_k}\mathbf{1}_{\{\Delta N_{t_k}\leq 1\}} = \lim_{\delta\to 0}\sum_{k=1}^{K}\Delta N_{t_k} = N_t, \text{ a.s.}$$

Therefore,

$$\lim_{\delta\to 0}\sum_{k=1}^{K}\Delta N_{t_k}^2\mathbf{1}_{\{\Delta N_{t_k}\geq 2\}}=0,\ \text{a.s.} \tag{41}$$

The convergences in eqs. (40) and (41) are also valid in L^1 because $\sum_{k=1}^K \Delta N_{t_k}^2 \leq N_t^2$, which is integrable.

Step 2. We use successively (understanding all convergences in the L^1 sense):

$$\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\mathbb{E}[\Delta\pi_{t_k}^2|\mathcal{F}_{t_{k-1}}]=0,\quad \lim_{\delta\downarrow 0}\sum_{k=1}^{K}\mathbb{E}[\Delta N_{t_k}\Delta\pi_{t_k}|\mathcal{F}_{t_{k-1}}]=0,\quad \lim_{\delta\downarrow 0}\sum_{k=1}^{K}\mathbb{E}[(\Delta N_{t_k}^2\mathbf{1}_{\{\Delta N_{t_k}\geq 2\}}|\mathcal{F}_{t_{k-1}}]=0$$

and

$$\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\mathbb{E}[(\Delta N_{t_k}^2\mathbf{1}_{\{\Delta N_{t_k}\leq 1\}}|\mathcal{F}_{t_{k-1}}]=\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\mathbb{E}[\Delta N_{t_k}\mathbf{1}_{\{\Delta N_{t_k}\leq 1\}}|\mathcal{F}_{t_{k-1}}]=\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\mathbb{E}[\Delta N_{t_k}|\mathcal{F}_{t_{k-1}}]=\pi_t,$$

where the last line obtains by eq. (4) and concludes the proof of Lemma 1.

Appendix B. **Proof of Lemma 2.** For the partition in eq. (1), consider a **simple process** φ such that $\varphi_t=\varphi_k$ for $t\in[t_{k-1},t_k]$ with φ_k being $\mathcal{F}_{t_{k-1}}$-measurable. Then, $\int_0^T\varphi_s\mathrm{d}\mu_s=\sum_{k=1}^K\varphi_k(\mu_{t_k}-\mu_{t_{k-1}})$. Because μ is a $\mathcal{F}$-martingale,

$$\mathbb{E}\Big(\int_0^T\varphi_s\mathrm{d}\mu_s\Big)^2=\sum_{k=1}^{K}\mathbb{E}\left(\varphi_k^2\mathbb{E}[(\mu_{t_k}-\mu_{t_{k-1}})^2|\mathcal{F}_{t_{k-1}}]\right)=\sum_{k=1}^{K}\mathbb{E}\Big(\varphi_k^2\int_{t_{k-1}}^{t_k}\lambda_s\mathrm{d}s\Big)$$

from Lemma 1. This establishes $\mathbb{E}\Big(\int_0^T\varphi_s\mathrm{d}\mu_s\Big)^2=\mathbb{E}\int_0^T\varphi_s^2\lambda_s\mathrm{d}s$ for a simple process φ.

One can approximate **any stochastic process** in $L^2_{\mathcal{F},\lambda}(0,T)$ by a simple process, namely

$$t\mapsto\varphi_t^\delta:=\begin{cases}\varphi_1=0 & \text{if } t\in[0,t_1)\\ \varphi_{k+1}=\dfrac{\int_{t_{k-1}}^{t_k}\varphi_s\lambda_s\mathrm{d}s}{\int_{t_{k-1}}^{t_k}\lambda_s\mathrm{d}s} & \text{if } t\in[t_k,t_{k+1}),\ k\in\{1,\dots,K-1\}.\end{cases} \tag{42}$$

using a partition with $t_k=k\delta$ and $K\delta=T$. This proves $\mathbb{E}\Big(\int_0^T\varphi_s\mathrm{d}\mu_s\Big)^2=\mathbb{E}\int_0^T\varphi_s^2\lambda_s\mathrm{d}s$ by convergence. This concludes the proof.

Appendix C. **Proof of Proposition 1.** We prove eq. (7) for $t=T$. We can take Ψ bounded. We consider the partition in eq. (1) and associate to the process φ the simple process φ^δ in eq. (42). We further define the process $(\xi_t^\delta:=\int_0^t\varphi_s^\delta\mathrm{d}N_s)_t$ for $t\in[0,T]$. We have $\xi_{t_0}^\delta=0$ and $\xi_{t_k}^\delta=\sum_{j=1}^k\varphi_j\left(N_{t_j}-N_{t_{j-1}}\right)$ for $k=1,\dots,K-1$. It follows that

$$\xi_s^\delta=\xi_{t_{k-1}}+\varphi_k(N_s-N_{t_{k-1}})\quad\text{for}\quad s\in[t_{k-1},t_k]. \tag{43}$$

We now write

$$\Psi(\xi_T^\delta)=\Psi(0)+\sum_{k=1}^{K}\left\{\Psi(\xi_{t_k}^\delta)-\Psi(\xi_{t_{k-1}}^\delta)\right\}=\Psi(0)+\sum_{k=1}^{K}\left\{\Psi(\xi_{t_{k-1}}^\delta+\varphi_k\Delta N_{t_k})-\Psi(\xi_{t_{k-1}}^\delta)\right\}.$$

Using eq. (41), we have

$$\lim_{\delta\downarrow 0}\Psi(\xi_T^\delta)=\Psi(0)+\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\left\{\Psi(\xi_{t_{k-1}}^\delta+\varphi_k\Delta N_{t_k})-\Psi(\xi_{t_{k-1}}^\delta)\right\}\mathbf{1}_{\{\Delta N_{t_k}\leq 1\}}$$

$$
\begin{aligned}
&= \Psi(0) + \lim_{\delta\downarrow 0} \sum_{k=2}^{K} \big\{ \Psi(\xi^\delta_{t_{k-1}} + \varphi_k \Delta N_{t_k}) - \Psi(\xi^\delta_{t_{k-1}}) \big\} \mathbf{1}_{\{\Delta N_{t_k} \le 1\}} \\
&= \Psi(0) + \lim_{\delta\downarrow 0} \sum_{k=1}^{K-1} \big\{ \Psi(\xi^\delta_{t_k} + \varphi_{k+1} \Delta N_{t_{k+1}}) - \Psi(\xi^\delta_{t_k})) \mathbf{1}_{\{\Delta N_{t_{k+1}} \le 1\}} \\
&= \Psi(0) + \lim_{\delta\downarrow 0} \sum_{k=1}^{K-1} \big\{ \Psi(\xi^\delta_{t_k} + \varphi_{k+1}) - \Psi(\xi^\delta_{t_k}) \big\} \Delta N_{t_{k+1}} \mathbf{1}_{\{\Delta N_{t_{k+1}} \le 1\}} \\
&= \Psi(0) + \lim_{\delta\downarrow 0} \sum_{k=1}^{K-1} \big\{ \Psi(\xi^\delta_{t_k} + \varphi_{k+1}) - \Psi(\xi^\delta_{t_k}) \big\} \Delta N_{t_{k+1}}.
\end{aligned}
$$

Therefore, we can assert that

$$
\lim_{\delta\downarrow 0} \Psi(\xi^\delta_T) = \Psi(0) + \lim_{\delta\downarrow 0} \sum_{k=1}^{K} \big[\Psi(\xi^\delta_{t_{k-1}} + \varphi_k) - \Psi(\xi^\delta_{t_{k-1}}) \big] \Delta N_{t_k}. \tag{44}
$$

On the other hand, let us consider

$$
\begin{aligned}
&\int_0^T \big\{ \Psi(\xi^\delta_s + \varphi^\delta_s) - \Psi(\xi^\delta_s) \big\} \mathrm{d}N_s \\
&= \sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \big\{ \Psi(\xi^\delta_s + \varphi^\delta_s) - \Psi(\xi^\delta_s) \big\} \mathrm{d}N_s \\
&= \sum_{k=1}^{K} \int_{t_{k-1}}^{t_k} \Big\{ \Psi(\xi^\delta_{t_{k-1}} + \varphi_k (N_s - N_{t_{k-1}}) + \varphi_k) - \Psi(\xi^\delta_{t_{k-1}} + \varphi_k (N_s - N_{t_{k-1}})) \Big\} \mathrm{d}N_s,
\end{aligned}
$$

where the last line comes from eq. (43). Then,

$$
\lim_{\delta\downarrow 0} \int_0^T \big\{ \Psi(\xi^\delta_s + \varphi^\delta_s) - \Psi(\xi^\delta_s) \big\} \mathrm{d}N_s = \lim_{\delta\downarrow 0} \sum_{k=1}^{K} \Big\{ \Psi(\xi^\delta_{t_{k-1}} + \varphi_k) - \Psi(\xi^\delta_{t_{k-1}}) \Big\} \Delta N_{t_k}.
$$

Comparing with eq. (44), we obtain

$$
\lim_{\delta\downarrow 0} \Psi(\xi^\delta_T) = \Psi(0) + \lim_{\delta\downarrow 0} \int_0^T \big\{ \Psi(\xi^\delta_s + \varphi^\delta_s) - \Psi(\xi^\delta_s) \big\} \mathrm{d}N_s. \tag{45}
$$

From the construction of the stochastic integral, eq. (45) is a special case of eq. (7) for $t = T$. This extends to any t. Equation (7) is valid when $\mathbb{E} \int_0^T \big| \Psi(\xi_t + \varphi_t) - \Psi(\xi_t) \big|^2 \lambda_t \mathrm{d}t < \infty$. This concludes the proof of Proposition 1.

Appendix D. **Proof of Proposition 2.** For the differential operator $\mathbb{A}$ defined in eq. (11), consider the equation

$$
\begin{cases} \frac{\partial z}{\partial s}(\lambda, s) + \mathbb{A} z(\lambda, s) &= 0, \qquad 0 < s < t \\ z(\lambda, t) &= \varphi(\lambda). \end{cases} \tag{46}
$$

where φ is measurable bounded function on $\mathbb{R}_+$. We establish the existence of a solution of eq. (46) by considering the following iteration:

$$
\begin{cases} -\frac{\partial z^{n+1}}{\partial s}(\lambda, s) - \frac{\partial z^{n+1}}{\partial \lambda}(\lambda, s) \alpha(\lambda_\infty - \lambda) + \lambda z^{n+1}(\lambda, s) = & \lambda z^n(\lambda + \beta, s), \qquad s < t \\ z^{n+1}(\lambda, t) = & \varphi(\lambda), \\ z^0(\lambda, s) = & ||\varphi|| := \sup_\lambda |\varphi(\lambda)|. \end{cases} \tag{47}
$$

The solution of eq. (47) is explicit. Indeed, for $s < \tau < t$, define

$$\lambda_\tau^{\lambda,s} := \lambda_\infty - (\lambda_\infty - \lambda)e^{-\alpha(\tau-s)}. \tag{48}$$

Then, we have

$$z^{n+1}(\lambda, s) = \varphi(\lambda_t^{\lambda,s})e^{-\int_s^t \lambda_\tau^{\lambda,s}\mathrm{d}\tau} + \int_s^t \lambda_\tau^{\lambda,s} z^n(\lambda_\tau^{\lambda,s} + \beta)e^{-\int_s^\tau \lambda_\theta^{\lambda,s}\mathrm{d}\theta}\mathrm{d}\tau. \tag{49}$$

The sequence z^n is monotone decreasing and $||z^n|| \le ||\varphi||$. It converges pointwise to a function z that solves eq. (46) and such that $||z|| \le ||\varphi||$.

Having established the existence of a solution of eq. (46), we now turn to probabilistic considerations. Applying the formula in eq. (9) to $z(\lambda_s, s)$ yields

$$\begin{aligned} z(\lambda_t, t) =& z(\lambda_s, s) + \int_s^t \Big\{\frac{\partial z}{\partial \tau}(\lambda_\tau, \tau) + \frac{\partial z}{\partial \lambda}(\lambda_\tau, \tau)\alpha(\lambda_\infty - \lambda_\tau)\Big\}\mathrm{d}\tau \\ &+ \int_s^t \Big\{z(\lambda_\tau + \beta, \tau) - z(\lambda_\tau, \tau)\Big\}\mathrm{d}N_\tau. \end{aligned} \tag{50}$$

Conditioning with respect to $\mathcal{F}_s$ we obtain from Lemma 1 (first line), eq. (11) (second line), and eq. (46) (third line) that

$$\begin{aligned} &\mathbb{E}\Big[z(\lambda_t, t)\Big|\mathcal{F}_s\Big] - z(\lambda_s, s) \\ =&\mathbb{E}\left[\int_s^t \Big\{\frac{\partial z}{\partial \tau}(\lambda_\tau, \tau) + \frac{\partial z}{\partial \lambda}(\lambda_\tau, \tau)\alpha(\lambda_\infty - \lambda_\tau) + \big[z(\lambda_\tau + \beta, \tau) - z(\lambda_\tau, \tau)\big]\lambda_\tau\Big\}\mathrm{d}\tau\Big|\mathcal{F}_s\right] \\ =&\mathbb{E}\left[\int_s^t \Big\{\frac{\partial z}{\partial \tau}(\lambda_\tau, \tau) + \mathbb{A}z(\lambda_\tau, \tau)\Big\}\mathrm{d}\tau\Big|\mathcal{F}_s\right] \\ =&0. \end{aligned}$$

We thus obtain $z(\lambda_s, s) = \mathbb{E}\big[\varphi(\lambda_t)\big|\mathcal{F}_s\big]$, which establishes the result.

Appendix E. **Proof of Proposition 3.** Equation (10) follows from eq. (9). For a $k \in \{0, 1, \dots\}$, let $z_k : \mathbb{R}_+ \times (0, t) \to \mathbb{R}$ denote the solution of

$$\begin{cases} -\frac{\partial z_k}{\partial s}(\lambda, s) = & \frac{\partial z_k}{\partial s}(\lambda, s)\alpha(\lambda_\infty - \lambda) + \lambda\big\{z_{k+1}(\lambda + \beta, s) - z_k(\lambda, s)\big\}, \quad 0 < s < t \\ \quad z_k(\lambda, t) = & \varphi_k(\lambda) \end{cases} \tag{51}$$

To study the problem for each k, we can proceed with an iteration as in eq. (47). We then obtain $z_{N_s}(\lambda_s, s) = \mathbb{E}\big[\varphi_{N_t}(\lambda_t)\big|\mathcal{F}_s\big]$, which completes the proof.

Appendix F. **Proof of Lemma 3.** Suppose for two events A and B in $\mathcal{B}_0$ such that $A \cap B = \emptyset$. Then, $N_{A\cup B} = N_A + N_B$. The property $\mu_{A\cup B} = \mu_A + \mu_B$ follows from the uniqueness of the Doob's decomposition.

Furthermore,

$$\begin{aligned} \langle \mu_A, \mu_B\rangle(t) =& \frac{1}{2}\langle \mu_A + \mu_B, \mu_A + \mu_B\rangle(t) - \frac{1}{2}\langle \mu_A, \mu_A\rangle(t) - \frac{1}{2}\langle \mu_B, \mu_B\rangle(t) \\ =& \frac{1}{2} < \mu_{A\cup B}, \mu_{A\cup B} > (t) - \frac{1}{2}\langle \mu_A, \mu_A\rangle(t) - \frac{1}{2}\langle \mu_B, \mu_B\rangle(t) \\ =& \frac{1}{2}\int_0^T \lambda_s \mathrm{d}s\mathrm{m}(A \cup B) - \frac{1}{2}\int_0^T \lambda_s \mathrm{d}s\mathrm{m}(A) - \frac{1}{2}\int_0^T \lambda_s \mathrm{d}s\mathrm{m}(B). \end{aligned}$$

But

$$\int_0^T \lambda_s \mathrm{d}s\, \mathrm{m}(A \cup B) = \int_0^T \lambda_s \mathrm{d}s \left(\mathrm{m}(A) + \mathrm{m}(B)\right). \tag{52}$$

Hence, $\langle \mu_A, \mu_B \rangle(t) = 0$, which proves orthogonality.

Appendix G. **Proof of Lemma 4.** For the partition in eq. (1), we define $\Delta_k := (t_{k-1}, t_k]$. Further, consider a **simple random fields** of the form

$$\varphi = \sum_{k=1}^{K} \sum_{h=1}^{H} \varphi_{kh} \mathbf{1}_{\Delta_k \times A_h}, \tag{53}$$

where the events $\{A_h\}_{h \in \{1,\ldots,H\}}$ in $\mathcal{B}_0$ are such that $A_h \cap A_{h'} = \emptyset$ if $h \neq h'$, and where φ_{kh} are assumed $\mathcal{F}_{t_k-1}$ measurable. In this case with simple random fields, we can define a stochastic integral with respect to the martingale measure μ as

$$\int_0^T \int_{\mathbb{R}^p} \varphi(s,z) \mathrm{d}\mu(s,z) = \sum_{k=1}^{K} \sum_{h=1}^{H} \varphi_{kh} \mu(\Delta_k, A_h), quad \text{where} \quad \mu(\Delta_k, A_h) = \mu_{A_h}(t_k) - \mu_{A_h}(t_{k-1}).$$

Because

$$\begin{cases} \mathbb{E}\Big[\mu(\Delta_k, A_h)\mu(\Delta_k, A_{h'}) \Big| \mathcal{F}_{t_{k-1}}\Big] = 0, & \text{if} \quad h \neq h', \forall k, \\ \mathbb{E}\Big[\mu(\Delta_k, A_h)\mu(\Delta_{k'}, A_{h'}) \Big| \mathcal{F}_{t_{k'}}\Big] = 0, & \text{if } k > k', \forall h, h', \end{cases} \tag{54}$$

it follows that

$$\begin{aligned} \mathbb{E}\left(\int_0^T \int_{\mathbb{R}^p} \varphi(s,z) \mathrm{d}\mu(s,z)\right)^2 &= \mathbb{E} \sum_{k=1}^{K} \sum_{h=1}^{H} \varphi_{kh}^2 \mathbb{E}\Big[\mu^2(\Delta_k, A_h) \Big| \mathcal{F}_{t_{k-1}}\Big] \\ &= \mathbb{E} \sum_{k=1}^{K} \sum_{h=1}^{H} \varphi_{kh}^2 \int_{t_{k-1}}^{t_k} \lambda_s \mathrm{d}s\, \mathrm{m}(A_h). \end{aligned}$$

Thus, eq. (16) is satisfied for simple random fields.

More generally, we can approximate an **arbitrary random field** $\varphi \in L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T) \times \mathbb{R}^p)$ by a sequence of simple random fields. We first define the sequence

$$\varphi^\delta(s,z) = \begin{cases} \varphi_1(z) = 0, & \text{if } \ 0 \leq s < t_1, \\ \varphi_{k+1}(z) = \frac{\int_{t_{k-1}}^{t_k} \varphi(s,z)\lambda_s \mathrm{d}s}{\int_{t_{k-1}}^{t_k} \lambda_s \mathrm{d}s} & \text{if } \ t_k \leq s < t_{k+1} \text{ for } 1 \leq k \leq K-1. \end{cases} \tag{55}$$

We then consider a subdivision of $\mathbb{R}^p$, with Borelians $A_h \in \mathcal{B}_0$ with $A_h \cap A_{h'} = \emptyset$ if $h \neq h'$ and $\cup_h A_h = \mathbb{R}^p \setminus \{0\}$. We have $\sum_h \mathrm{m}(A_h) = 1$. We can always assume $h \in \{1, \cdots, H\}$. We then define

$$\varphi^{\delta,H}(s,z) = \varphi_{k+1,h} = \frac{\int_{A^h} \varphi_{k+1}(z) \mathrm{d}\mathrm{m}(z)}{\mathrm{m}(A_h)}, \text{if } \ t_k \leq s < t_{k+1}, z \in A_h, \tag{56}$$

for $k \in \{0, \ldots, K-1\}$ and $h \in \{1, \ldots, H\}$. Then, $\varphi^{\delta,H} \in L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T) \times \mathbb{R}^p)$ and converges to φ in $L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T) \times \mathbb{R}^p)$. This completes the proof.

Appendix H. **Proof of Lemma 5.**

Bounded Ψ**.** We approximate $z \mapsto \varphi(s,z)$ by a function which is piecewise constant. Consider a subdivision $\{A_1, \cdots, A_H\}$ of $\mathbb{R}^p$ such that $A_h \cap A_{h'} = \emptyset$ and $\cup_h A_h = \mathbb{R}^p$ and define

$$\varphi^H(s,z) := \sum_{h=1}^{H} \varphi_h(s) \mathbf{1}_{A_h}(z), \quad \text{where} \quad \varphi_h(s) = \frac{\int_{A_h} \varphi(s,z) \mathrm{dm}(z)}{\mathrm{m}(A_h)}.$$

If $\mathrm{m}(A_h) \leq \frac{c}{H}$, we can view φ^H as an approximation of the random field φ because

$$\varphi^H \to \varphi \text{ in } L^2_{\mathcal{F},\lambda\mathrm{m}}((0,T) \times \mathbb{R}^p) \quad \text{as } H \to \infty,$$

We define

$$\xi^H_t := \int_0^t \int_{\mathbb{R}^p} \varphi^H(s,z) \mathrm{d}N(s,z) = \sum_{h=1}^{H} \int_0^t \varphi_h(s) \mathrm{d}N^{A_h}_s.$$

We want to prove

$$\Psi(\xi^H_t) = \Psi(0) + \int_0^t \sum_{h=1}^{H} \Big\{ \Psi(\xi^H_s + \varphi_h(s)) - \Psi(\xi^H_s) \Big\} \mathrm{d}N^{A_h}_s \tag{57}$$

We prove this result for $t = T$ and consider the partition in eq. (1). We set

$$\begin{aligned}
\varphi^\delta_h(t) &= \begin{cases} \varphi_{h1} = 0, & \text{if } 0 \leq t < t_1 \\ \varphi_{hk} = \frac{\int_{t_{k-2}}^{t_{k-1}} \varphi_h(s) \lambda_s \mathrm{d}s}{\int_{t_{k-2}}^{t_{k-1}} \lambda_s \mathrm{d}s}, & \text{if } t_{k-1} \leq t < t_k, \ 2 \leq k < K \end{cases} \\
\xi^{\delta H}_t &= \xi^{\delta H}_{t_{k-1}} \text{ for } t_{k-1} \leq t < t_k, \text{ with } k \in \{1, \cdots, K\} \\
\xi^{\delta H}_{t_k} &= \xi^{\delta H}_{t_{k-1}} + \sum_{h=1}^{H} \varphi_{hk} \Delta N^{A_h}_{t_k}, \text{ with } k \in \{1, \cdots, K\}, \\
\xi^{\delta H}_0 &= 0, \\
\Delta N^{A_h}_{t_k} &= N^{A_h}_{t_k} - N^{A_h}_{t_{k-1}}.
\end{aligned}$$

Therefore,

$$\xi^{\delta H}_T = \sum_{k=1}^{K} \sum_{h=1}^{H} \varphi_{hk} \Delta N^{A_h}_{t_k}$$

and

$$\begin{aligned}
\Psi(\xi^{\delta H}_T) =& \Psi(0) + \sum_{k=1}^{K} \Big\{ \Psi(\xi^{\delta H}_{t_k}) - \Psi(\xi^{\delta H}_{t_{k-1}}) \Big\} \\
=& \Psi(0) + \sum_{k=1}^{K} \left\{ \Psi\left(\xi^{\delta H}_{t_{k-1}} + \sum_{h=1}^{H} \varphi_{hk} \Delta N^{A_h}_{t_k} \right) - \Psi\left(\xi^{\delta H}_{t_{k-1}} \right) \right\} \\
=& \Psi(0) + \sum_{k=1}^{K} \sum_{h=1}^{H} \int_0^1 \Psi' \left(\xi^{\delta H}_{t_{k-1}} + \theta \sum_{h'=1}^{H} \varphi_{h'k} \Delta N^{A_{h'}}_{t_k} \right) \varphi_{hk} \Delta N^{A_h}_{t_k} \mathrm{d}\theta
\end{aligned}$$

by Taylor's Theorem. We use now

$$\lim_{\delta \downarrow 0} \sum_{k=1}^{K} \sum_{h=1}^{H} \int_0^1 \Psi' \left(\xi^{\delta H}_{t_{k-1}} + \theta \sum_{h'=1}^{H} \varphi_{h'k} \Delta N^{A_{h'}}_{t_k} \right) \varphi_{hk} \Delta N^{A_h}_{t_k} \mathbf{1}_{\{\Delta N^{A_h}_{t_k} \geq 2\}} d\theta \to 0,$$

to obtain

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\Psi(\xi_T^{\delta H})\\
=&\Psi(0)+\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\sum_{h=1}^{H}\int_0^1\Psi'\left(\xi_{t_{k-1}}^{\delta H}+\theta\sum_{h'=1}^{H}\varphi_{h'k}\Delta N_{t_k}^{A_{h'}}\right)\varphi_{hk}\Delta N_{t_k}^{A_h}\mathbf{1}_{\{\Delta N_{t_k}^{A_h}\leq 1\}}\mathrm{d}\theta\\
=&\Psi(0)+\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\sum_{h=1}^{H}\int_0^1\Psi'\left(\xi_{t_{k-1}}^{\delta H}+\theta\sum_{h'=1}^{H}\varphi_{h'k}\Delta N_{t_k}^{A_{h'}}\Delta N_{t_k}^{A_h}\mathbf{1}_{\{\Delta N_{t_k}^{A_h}\leq 1\}}\right)\varphi_{hk}\mathbf{1}_{\{\Delta N_{t_k}^{A_h}\leq 1\}}\mathrm{d}\theta
\end{aligned}$$

However,

$$\lim_{\delta\downarrow 0}\Delta N_{t_k}^{A_{h'}}\Delta N_{t_k}^{A_h}\mathbf{1}_{\{\Delta N_{t_k}^{A_{h'}}\geq 1\}}\mathbf{1}_{\{\Delta N_{t_k}^{A_h}\leq 1\}}=0,\text{ if } h'\neq h$$

because

$$\lim_{\delta\downarrow 0}\Delta N_{t_k}^{A_{h'}}\mathbf{1}_{\{\Delta N_{t_k}^{A_{h'}}\geq 1\}}\mathbf{1}_{\{\Delta N_{t_k}^{A_h}=1\}}\leq\lim_{\delta\downarrow 0}\Delta N_{t_k}\mathbf{1}_{\{\Delta N_{t_k}\geq 2\}}=0.$$

Collecting results, we can write

$$\begin{aligned}
\lim_{\delta\downarrow 0}\Psi(\xi_T^{\delta H})=&\Psi(0)+\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\sum_{h=1}^{H}\int_0^1\Psi'\left(\xi_{t_{k-1}}^{\delta H}+\theta\varphi_{hk}\right)\varphi_{hk}\mathbf{1}_{\{\Delta N_{t_k}^{A_h}\leq 1\}}\mathrm{d}\theta\\
=&\Psi(0)+\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\sum_{h=1}^{H}\int_0^1\Psi'\left(\xi_{t_{k-1}}^{\delta H}+\theta\varphi_{hk}\right)\varphi_{hk}\Delta N_{t_k}^{A_h}\mathrm{d}\theta.
\end{aligned}$$

Therefore,

$$\lim_{\delta\downarrow 0}\Psi(\xi_T^{\delta H})=\Psi(0)+\lim_{\delta\downarrow 0}\sum_{k=1}^{K}\sum_{h=1}^{H}\left\{\Psi(\xi_{t_{k-1}}^{\delta H}+\varphi_{hk})-\Psi(\xi_{t_{k-1}}^{\delta H})\right\}\Delta N_{t_k}^{A_h},$$

which means that

$$\lim_{\delta\downarrow 0}\Psi(\xi_T^{\delta H})=\Psi(0)+\lim_{\delta\downarrow 0}\sum_{h=1}^{H}\int_0^T\left\{\Psi(\xi_s^{\delta H}+\varphi_h^{\delta}(s))-\Psi(\xi_s^{\delta H})\right\}\Delta N_s^{A_h}$$

which is eq. (57) and completes the proof for Ψ bounded.

Unbounded Ψ. If Ψ satisfies eq. (19), we first note that, in view of the linearity in Ψ, it is sufficient to assume $\Psi>0$. It is also sufficient to prove Lemma 5 with Ψ replaced with $\Psi\wedge M$, where M is a constant. This is true since $\Psi\wedge M$ is bounded. This concludes the proof.

Appendix I. **Proof of Proposition 4.** For $k\in\mathbb{N}$, we introduce the partial differential equation

$$\begin{cases}\frac{\partial z_k^A}{\partial s}(\lambda,s)+\frac{\partial z_k^A}{\partial s}(\lambda,s)\alpha(\lambda_\infty-\lambda)+\lambda\mathrm{m}(A)\left[z_{k+1}^A(\lambda+\beta,s)-z_k^A(\lambda,s)\right] &=0, \qquad s<t,\\ \qquad z_k^A(\lambda,t) &=\varphi_k^A(\lambda)\end{cases}\tag{58}$$

and denote by z_{kA} its solution. We can study eq. (58) as we did before and obtain a Feynman-Kac formula:

$$z_k^A(\lambda,s)=\mathbb{E}\Big[\varphi_k^A(\lambda_t)\Big|\lambda_s=\lambda\Big].$$

This allows use to define a semigroup on the set of measurable bounded functions φ_k, namely

$$\hat{\mathrm{T}}_{s,t}^A\varphi=z_\cdot^A(\cdot,s),\tag{59}$$

which corresponds to the Markov process $\{N^A,\lambda\}$. This completes the proof.

Appendix J. Proof of Proposition 5. Our proof is similar to the one given in [3]. We limit the proof to a function Ψ not dependent on t, which is C^2 with bounded derivatives. The result obtains as a consequence of $\int_0^t \int_{R^p} \varphi(s,z) dN(s,z)$ having bounded variations. Let us set

$$\xi_t^c := \xi_0 + \int_0^t \alpha(s)\mathrm{d}s + \int_0^t \beta(s) \cdot \mathrm{d}W_s \quad \text{and} \quad \zeta_t = \int_0^t \int_{R^p} \varphi(s,z) dN(s,z).$$

We establish eq. (25) for $t = T$ and consider a partition of $[0,T]$ of the form in eq. (1). We write $\Psi(\xi_T) = \Psi(\xi_0) + \sum_{j=1}^N \left\{\Psi(\xi_{t_j}) - \Psi(\xi_{t_{j-1}})\right\}$. Defining

$$S_1 = \sum_{j=1}^N \Big\{\Psi(\xi_{t_j}^c + \zeta_{t_{j-1}}) - \Psi(\xi_{t_{j-1}}^c + \zeta_{t_{j-1}})\Big\}$$

$$S_2 = \sum_{j=1}^N \Big\{\Psi(\xi_{t_{j-1}}^c + \zeta_{t_j}) - \Psi(\xi_{t_{j-1}}^c + \zeta_{t_{j-1}})\Big\}$$

$$S_3 = \sum_{j=1}^N \Big\{\Psi(\xi_{t_j}^c + \zeta_{t_j}) - \Psi(\xi_{t_j}^c + \zeta_{t_{j-1}}) - \Psi(\xi_{t_{j-1}}^c + \zeta_{t_j}) + \Psi(\xi_{t_{j-1}}^c + \zeta_{t_{j-1}})\Big\},$$

we can write $\Psi(\xi_T) = S_1 + S_2 + S_3$. We study each term in turn.

Term S_3. We set $\Delta\xi_{t_j}^c := \xi_{t_j}^c - \xi_{t_{j-1}}^c$ and $\Delta\zeta_{t_j} = \zeta_{t_j} - \zeta_{t_{j-1}}$. We recall that $\Psi \in C^2(\mathbb{R})$. We have

$$S_3 = \sum_{j=1}^N \int_0^1 \int_0^1 D_x^2 \Psi\big(\xi_{t_{j-1}}^c + \theta\Delta\xi_{t_j}^c + \theta'\Delta\zeta_{t_j}\big)\Delta\zeta_{t_j}\Delta\xi_{t_j}^c \mathrm{d}\theta\mathrm{d}\theta'.$$

It follows that there exists a $C > 0$ such that

$$\big|S_3\big| \le C \max_{j=1,\ldots,N} \big|\Delta\xi_{t_j}^c\big| \sup_\delta \sum_{j=1}^N \big|\Delta\zeta_{t_j}\big|.$$

Hence, $\big|S_3\big|$ vanishes almost surely because ζ has bounded variations almost surely and $\max_{j=1}^N |\Delta\xi_{t_j}^c| \to 0$, almost surely as $\delta \to 0$.

Term S_1. For S_1 we apply Ito's formula for ordinary diffusions to obtain

$$S_1 = \sum_{j=1}^N \Big\{ \int_{t_{j-1}}^{t_j} D_x\Psi(\xi_s^c + \zeta_{t_{j-1}})\mathrm{d}\xi_s^c + \frac{1}{2}\int_{t_{j-1}}^{t_j} D_x^2\Psi(\xi_s^c + \zeta_{t_{j-1}})|\beta(s)|^2 \mathrm{d}s\Big\},$$

from which it follows that

$$S_1 \to \int_0^T D_x\Psi(\xi_s)\mathrm{d}\xi_s^c + \frac{1}{2}\int_0^T D_x^2\Psi(\xi_s)\big|\beta(s)\big|^2 \mathrm{d}s \text{ almost surely.}$$

Term S_2. We now apply the result in Lemma 5:

$$\begin{aligned} S_2 &= \sum_{j=1}^N \int_{t_{j-1}}^{t_j} \int_{\mathbb{R}^p} \Big\{\Psi(\xi_{t_{j-1}}^c + \zeta_s + \varphi(s,z)) - \Psi(\xi_{t_{j-1}}^c + \zeta_s)\Big\}\mathrm{d}N(s,z) \\ &= \int_0^T \int_{\mathbb{R}^p} \Big\{\Psi(\xi_s^{c,\delta} + \zeta_s + \varphi(s,z)) - \Psi(\xi_s^{c,\delta} + \zeta_s)\Big\}\mathrm{d}N(s,z) \end{aligned}$$

with $\xi_s^{c,\delta} := \xi_{t_{j-1}}^c$ for $s \in (t_{j-1}, t_j]$. We can check that

$$S_2 \to \int_0^T \int_{\mathbb{R}^p} \Big\{\Psi(\xi_s + \varphi(s,z), s) - \Psi(\xi_s, s)\Big\} \mathrm{d}N(s,z) \quad \text{in} \quad L^1$$

which completes the proof.

Appendix K. **Proof of Proposition 6.** We sketch the proof.
Estimates for the moments. We consider the function $\Psi(\lambda) = \frac{\lambda^m}{m}$ and apply the formula in eq. (10). This gives us

$$\mathrm{d}\frac{\lambda_t^m}{m} = \alpha\lambda_t^{m-1}(\lambda_\infty - \lambda(t)\mathrm{d}t + \frac{(\lambda_t+\beta)^m - \lambda_t^m}{m}\mathrm{d}N_t.$$

Consequently,

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathbb{E}\lambda_t^m}{m} = \alpha\Big[\lambda_\infty\mathbb{E}\lambda_t^{m-1} - \mathbb{E}\lambda_t^m\Big] + \frac{\mathbb{E}\Big[\big\{(\lambda_t+\beta)^m - \lambda_t^m\big\}\lambda_t\Big]}{m}. \tag{60}$$

We now note that

$$\begin{aligned}\frac{\mathbb{E}\Big[\big\{(\lambda_t+\beta)^m - \lambda_t^m\big\}\lambda_t\Big]}{m} =& \mathbb{E}\int_0^1 \big(\lambda_t + \theta\beta\big)^{m-1}\lambda_t \mathrm{d}\theta \\ \le& \mathbb{E}\Big[\max\big\{\lambda_t;\beta\big\}^m\Big]\int_0^1 (1+\theta)^{m-1}\mathrm{d}\theta \\ \le& 2^{m-1}\times\mathbb{E}\big[\lambda_t^m + \beta^m\big]\end{aligned}$$

and

$$\mathbb{E}\lambda_t^{m-1} \le \frac{m-1}{m}\mathbb{E}\lambda_t^m + \frac{1}{m}.$$

We now use these estimates in eq. (60) and obtain, by standard techniques, that

$$\mathbb{E}\big[\lambda_t^m\big] \le (\lambda^m + B_m)e^{A_m t} \tag{61}$$

where A_m and B_m are constants depending only on the power m.
Stochastic differential equation. Equation (28') can be viewed as a fixed-point equation. We consider processes $\xi := (\xi_t)_t$ that are adapted, càdlàg, and such that

$$\big|\big|\xi\big|\big|^2 = \sup_{0<t<T} \mathbb{E}|\xi_t|^2 < \infty. \tag{62}$$

This norm defines a Banach space. We then define the process

$$\begin{aligned}\mathrm{I}(s;\xi) :=& x + \int_t^s g(\xi_\tau,\tau)\mathrm{d}\tau + \int_t^s \sigma(\xi_\tau,\tau)\mathrm{d}W_\tau + \int_t^s\int_{\mathbb{R}^p}\gamma(\xi_\tau,\tau,z)\lambda_\tau\mathrm{d}\tau\,\mathrm{m}(\mathrm{d}z) \\ &+ \int_t^s\int_{\mathbb{R}^p}\gamma(\xi_\tau,\tau,z)\mathrm{d}\mu(\tau,z),\end{aligned} \tag{63}$$

which satisfies also the property in eq. (62) because of ineq. (27a) in Assumption 1 and the estimate in eq. (61).

We then check from martingale estimates that

$$\mathbb{E}\sup_{0\le s\le T}\big|\mathrm{I}(s;\xi^1) - \mathrm{I}(s;\xi^2)\big|^2 \le K\int_0^T \mathbb{E}\big|\xi_s^1 - \xi_s^2\big|^2\mathrm{d}s.$$

From this estimate, it is well known that the iteration

$$\xi_s^0 = x \quad \text{and} \quad \xi_s^{n+1} = \mathrm{I}(s;\xi^n) \tag{64}$$

converges almost surely uniformly with respect to s to a process which solves eq. (63) and that the solution is unique. This completes the proof.

Appendix L. **Proof of Proposition 7.** Again, the proof involves constructing the solution z of the PDE

$$\begin{cases} \left(\frac{\partial}{\partial t} + \tilde{\mathbb{A}}\right) z(x,\lambda,s) = & 0, \qquad s < t \\ \qquad\qquad z(x,\lambda,t) = & \varphi(x,\lambda). \end{cases} \tag{65}$$

This solution has the Feynman-Kac representation:

$$z(X_s,\lambda_s,s) = \mathbb{E}\Big[\varphi(X_t,\lambda_t)\Big|\mathcal{F}_s\Big]. \tag{66}$$

This completes the proof.

Appendix M. **Proof of Theorem 1.**
Reformulation as a fixed-point equation. Equation (31) reads

$$\begin{cases} -\frac{\partial V}{\partial s}(x,\lambda,s) - \alpha\frac{\partial V}{\partial \lambda}(x,\lambda,s)(\lambda_\infty - \lambda) & -D_x V(x,\lambda,s)\cdot g(x,s) - \frac{1}{2}\mathrm{tr}\{D_x^2 V(x,\lambda,s)\,a(x,s)\} + \lambda V(x,\lambda,s) \\ & = \lambda \int_{\mathbb{R}^p} V(x+\gamma(x,s;z),\lambda+\beta,s)\mathrm{m}(\mathrm{d}z) + f(x,s), \\ V(x,\lambda,T) & = h(x) \end{cases} \tag{31'}$$

Let $(\lambda_s^{\lambda,t})_{s\geq t}$ denote the solution of

$$\begin{cases} \frac{d\lambda}{ds} = & \alpha(\lambda_\infty - \lambda_s), \quad s > t \\ \lambda_t = & \lambda, \end{cases} \tag{67}$$

which is

$$\lambda_s^{\lambda,t} = \lambda e^{-\alpha(s-t)} + \lambda_\infty\left(1 - e^{-\alpha(s-t)}\right). \tag{67'}$$

Define next

$$Z^{\lambda,t}(x,s) := V(x,\lambda_s^{\lambda,t},s). \tag{68}$$

By differentiation,

$$\frac{\partial Z^{\lambda,t}}{\partial s}(x,s) = \frac{\partial V}{\partial s}(x,\lambda_s^{\lambda,t},s) + \alpha\frac{\partial V}{\partial \lambda}(x,\lambda_s^{\lambda,t},s)\big(\lambda_\infty - \lambda_s^{\lambda,t}\big)$$

Hence, from eq. (31'),

$$\begin{cases} -\frac{\partial Z^{\lambda,t}}{\partial s}(x,s) - D_x Z^{\lambda,t}(x,s)\cdot g(x,s) & -\frac{1}{2}\mathrm{tr}\{D_x^2 Z^{\lambda,t}(x,s)a(x,s)\} + \lambda_s^{\lambda,t} Z^{\lambda,t}(x,s) \\ & = f(x,s) + \lambda_s^{\lambda,t}\int_{\mathbb{R}^p} V(x+\gamma(x,s;z),\lambda_s^{\lambda,t}+\beta,s)\mathrm{m}(\mathrm{d}z), \ s > t \\ Z^{\lambda,t}(x,T) & = h(x), \\ V(x,\lambda,t) & = Z^{\lambda,t}(x,t). \end{cases} \tag{69}$$

We also introduce the solution $(X_s^{x,t})_{s\geq t}$ of

$$\begin{cases} \mathrm{d}X_s^{x,t} & = g(X_s^{x,t},s)\mathrm{d}s + \sigma(X_s^{x,t},s)\mathrm{d}W_s, \\ X_t^{x,t} & = x. \end{cases} \tag{70}$$

From Feyman-Kac Theorem, eq. (69) has the probabilistic interpretation:

$$V(x,\lambda,t) = \mathbb{E}\Big[\int_t^T \Big\{f(X_s^{x,t},s) + \lambda_s^{\lambda,t}\int_{\mathbb{R}^p} V(X_s^{x,t}+\gamma(X_s^{x,t},s;z),\lambda_s^{\lambda,t}+\beta,s)\mathrm{m}(\mathrm{d}z)\Big\} e^{-\int_t^s \lambda_\tau^{\lambda,t}\mathrm{d}\tau}\mathrm{d}s + h(X_T^{x,t})e^{-\int_t^T \lambda_\tau^{\lambda,t}\mathrm{d}\tau}\mathrm{d}s\Big]. \tag{71}$$

Equation (71) is a fixed point equation, which reformulates in a probabilistic fashion the fixed-point eq. (31). The RHS defines a monotone increasing affine operator.

A priori estimates. Let us check the a priori estimate

$$0 \le V(x,\lambda,t) \le \overline{f}(T-t) + \overline{h}. \tag{72}$$

Indeed, if the estimate in eq. (72) holds for V on the RHS of eq. (71), then from Assumption 2, the LHS of eq. (72) satisfies

$$\begin{aligned} V(x,\lambda,t) \le & \int_t^T \Big\{ \overline{f} + \lambda_s \big[\overline{f}(T-s) + \overline{h}\big] \Big\} e^{-\int_t^s \lambda_\tau \mathrm{d}\tau} \mathrm{d}s + \overline{h} e^{-\int_t^T \lambda_\tau \mathrm{d}\tau} \\ = & \int_t^T \overline{f} e^{-\int_t^s \lambda_\tau \mathrm{d}\tau} \mathrm{d}s - \int_t^T \big[\overline{f}(T-s) + \overline{h}\big] \mathrm{d}\left(e^{-\int_t^s \lambda_\tau \mathrm{d}\tau}\right) + \overline{h} e^{\int_t^T \lambda_\tau \mathrm{d}\tau} \\ = & \overline{f}(T-t) + \overline{h}. \end{aligned}$$

This implies the RHS estimate in eq. (72) for V on the LHS of eq. (71).

Because f and h are positive, the function V in eq. (71) is positive. Hence, the LHS inequality in eq. (72) is also proven.

We next obtain an a priori estimate for the derivative

$$\overline{V}(x,\lambda,t) := \frac{\partial V}{\partial \lambda}(x,\lambda,t)$$

From eq. (67'), we have

$$\frac{\partial \lambda_s^{\lambda,t}}{\partial \lambda} = e^{-\alpha(s-t)} \quad \text{and} \quad \int_t^s \frac{\partial \lambda_\tau^{\lambda,t}}{\partial \lambda} \mathrm{d}\tau = \frac{1 - e^{-\alpha(s-t)}}{\alpha}. \tag{73}$$

It follows from differentiating eq. (71) in λ that

$$\begin{aligned} \overline{V}(x,\lambda,t) = & \mathbb{E} \int_t^T \Big\{ \int_{\mathbb{R}^p} V(X_s^{x,t} + \gamma(X_s, s; z), \lambda_s + \beta, s) \mathrm{m}(\mathrm{d}z) \\ & + \lambda_s \int_{\mathbb{R}^p} \overline{V}(X_s + \gamma(X_s, s; z), \lambda_s + \beta, s) \mathrm{m}(\mathrm{d}z) e^{-\alpha(s-t)} \Big\} e^{-\int_t^s \lambda_\tau^{\lambda,t} \mathrm{d}\tau} \\ & - \frac{1 - e^{-\alpha(s-t)}}{\alpha} \Big[f(X_s, s) + \lambda_s \int_{\mathbb{R}^p} V(X_s + \gamma(X_s, s; z), \lambda_s + \beta, s) \mathrm{m}(\mathrm{d}z) \Big] \Big\} e^{-\int_t^s \lambda_\tau^{\lambda,t} \mathrm{d}\tau} \mathrm{d}\tau \mathrm{d}s \\ & - \frac{1 - e^{-\alpha(T-t)}}{\alpha} \mathbb{E} h(X_T) e^{-\int_t^T \lambda_\tau \mathrm{d}\tau}. \end{aligned} \tag{74}$$

We want to check the a priori estimate

$$\Big| \frac{\partial V}{\partial \lambda}(x,\lambda,t) \Big| \le \int_t^T \big\{ \overline{f}(T-s) + \overline{h} \big\} e^{-\alpha(s-t)} \mathrm{d}s. \tag{75}$$

Assuming that $\overline{V}$ on the RHS of eq. (74) satisfies eq. (75), we obtain from eq. (72) that

$$\begin{aligned} \overline{V}(x,\lambda,t) \le & \int_t^T \Big\{ e^{-\alpha(s-t)} \{\overline{f}(T-s) + \overline{h}\} + \lambda_s \int_s^T \{\overline{f}(T-\tau) + \overline{h}\} e^{-\alpha(\tau - s) \mathrm{d}\tau} \Big\} e^{-\int_t^s \lambda_\tau \mathrm{d}\tau} \mathrm{d}s \\ \le & \int_t^T \{\overline{f}(T-s) + \overline{h}\} e^{-\alpha(s-t)} \mathrm{d}s \end{aligned}$$

Similarly we check the estimate

$$\overline{V}(x,\lambda,t) \ge - \int_t^T \big\{ \overline{f}(T-s) + \overline{h} \big\} e^{-\alpha(s-t)} \mathrm{d}s.$$

The estimate in eq. (75) is thus obtained.

We can write eq. (31') as

$$\begin{aligned}-\frac{\partial V}{\partial s}(x,\lambda,s) - D_x V(x,\lambda,s)\cdot g(x,s) - \frac{1}{2}\mathrm{tr}\{D_x^2 V(x,\lambda,s)a(x)\}\\ =&\alpha\frac{\partial V}{\partial \lambda}(x,\lambda,s)(\lambda_\infty-\lambda)\\ &+\lambda\int_{\mathbb{R}^p} V(x+\gamma(x,s;z),\lambda+\beta,s)\mathrm{m}(\mathrm{d}z)\\ &-\lambda V(x,\lambda,s)+f(x,s).\end{aligned} \tag{76}$$

We denote by $F(x,\lambda,s)$ the RHS of eq. (76). Thanks to eqs. (72) and (75), we can state the estimate

$$\big|F(x,\lambda,s)\big| \leq \alpha(\lambda_\infty+\lambda)\int_s^T\{\overline{f}(T-\tau)+\overline{h}\}e^{-\alpha(\tau-s)d\tau}+\lambda\{\overline{f}(T-s)+\overline{h}\}+\overline{f}. \tag{77}$$

With this writing, λ is just a parameter in eq. (76). We can apply the classical regularity results of nondegenerate parabolic PDE to obtain a sufficiently smooth solution of eq. (31'). We can then apply Ito's formula in eq. (26) to compute the Ito differential of $(V(X_s,\lambda_s,s))_{s\geq t}$. We then obtain the probabilistic representation in eq. (30) following classical arguments. This concludes the proof.

Appendix N. **Proof of Theorem 2.** We can proceed as in Theorem 1 to check that, if we assume

$$\Big|\bar{V}(x,\lambda,t)\Big| \leq \int_t^T\Big\{\bar{f}(T-s)+\bar{h}\Big\}e^{-\alpha(s-t)}\mathrm{d}s,$$

then it follows that

$$\left|\frac{\partial}{\partial\lambda}K^{x,\lambda,t}(v;V)\right| \leq \int_t^T\Big\{\bar{f}(T-s)+\bar{h}\Big\}e^{-\alpha(s-t)}\mathrm{d}s. \tag{78}$$

Since this estimate does not depend on the argument v, it carries over to $V(x,\lambda,t) = \inf K^{x,\lambda,t}(v;V)$. This reasoning proves the a prori estimate

$$\left|\frac{\partial V}{\partial\lambda}(x,\lambda,t)\right| \leq \int_t^T e^{-\alpha(s-t)}\Big\{\overline{f}(T-s)+\overline{h}\Big\}\mathrm{d}s. \tag{79}$$

With this estimate, we can proceed as in the case of Theorem 1, putting $\alpha\frac{\partial V}{\partial\lambda}(x,\lambda,s)(\lambda_\infty-\lambda)$ to the RHS of eq. (38) and consider it as a given function. So λ is a parameter. We obtain an ordinary HJB equation. This allows us to benefit from the regularity results existing for quasilinear parabolic PDEs.

We can then write the Itô differential of $(V(X_s,\lambda_s,s))_s$, where X solves eq. (32). We check that

$$V(x,\lambda,t) \leq J^{x,\lambda,t}(v), \quad \forall v.$$

We next use the feedback $\widehat{v}(x,\lambda,s)$. We can solve the system

$$\begin{cases}\mathrm{d}\hat{X}_s = & g(\hat{X}_s,\hat{v}(\hat{X}_s,\lambda_s,s),s)\mathrm{d}s+\sigma(\hat{X}_s,s)\mathrm{d}W_s+\int_{\mathbb{R}^p}\gamma(\hat{X}_s,\hat{v}(\hat{X}_s,\lambda_s,s),s,z)\mathrm{d}N(s,z)\\ \mathrm{d}\lambda_s = & \alpha(\lambda_\infty-\lambda_s)\mathrm{d}s+\beta\mathrm{d}N_s, \quad s>t\\ \hat{X}_t = & x,\\ \lambda_t = & \lambda,\end{cases} \tag{80}$$

and define the control

$$\hat{V}_s = \hat{v}(\hat{X}_s,\lambda_s,s), \quad \forall s\in(t,T).$$

We can compute the Ito differential of $(V(\hat{X}_s, \lambda_s, s))_s$ and obtain

$$V(x, \lambda, t) = J^{x,\lambda,t}(\hat{V}).$$

It immediately follows that

$$V(x, \lambda, t) = \inf_v J^{x,\lambda,t}(v), \tag{81}$$

which completes the proof.

REFERENCES

[1] Y. Aït-Sahalia and T. R. Hurd, Portfolio choice in markets with contagion, *Journal of Financial Econometrics*, **14** (2015).

[2] E. Bacry, I. Mastromatteo and J.-F. Muzy, Hawkes processes in finance, *Market Microstructure and Liquidity*, **1** (2015).

[3] A. Bensoussan and J.-L. Lions, *Impulse Control and Quasi-Variational Inequalities*, North Holland, 1984.

[4] B. Bian, X. Chen and X. Zeng, Optimal portfolio choice in a jump-diffusion model with self-exciting, *Journal of Mathematical Finance*, **09** (2019).

[5] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[6] L. Cui, A. Hawkes and H. Yi, An elementary derivation of moments of Hawkes processes, *Advances in Applied Probability*, **52** (2020), 102-137.

[7] A. Dassios and H. Zhao, A dynamic contagion process, *Advanced Applied Probability*, **43** (2011), 814-846.

[8] D. Duffie, J. Pan and K. Singleton, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, **68** (2000), 1343-1376.

[9] E. B. Dynkin, *Markov Processes*, Springer Berlin Heidelberg, 1965.

[10] E. Errais, K. Giesecke and L. R. Goldberg, Affine point processes and portfolio credit risk, *SIAM Journal on Financial Mathematics*, **1** (2010), 642-665.

[11] D. Hainaut, Impact of volatility clustering on equity indexed annuities, *Insurance: Mathematics and Economics*, **71** (2016), 367-381.

[12] A. G. Hawkes, Point spectra of some mutually exciting point processes, *Journal of the Royal Statistical Society: Series B*, **33** (1971), 438-443.

[13] A. G. Hawkes, Hawkes processes and their applications to finance: A review, *Quantitative Finance*, **18** (2018), 193-198.

[14] S. G. Kou, A jump-diffusion model for option pricing, *Management Science*, **48** (2002), 1086-1101.

[15] R. C. Merton, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics*, **3** (1976), 125-144.

[16] P. Protter, *Stochastic Integration and Differential Equations*, Springer Berlin Heidelberg, 2004.

Received April 2024; revised July 2024; early access August 2024.