



Stochastic Volterra equations with time-changed Lévy noise and maximum principles

Giulia di Nunno^{1,2} · Michele Giordano¹

Accepted: 17 March 2023
© The Author(s) 2023

Abstract

Motivated by a problem of optimal harvesting of natural resources, we study a control problem for Volterra type dynamics driven by time-changed Lévy noises, which are in general not Markovian. To exploit the nature of the noise, we make use of different kind of information flows within a maximum principle approach. For this we work with backward stochastic differential equations (BSDE) with time-change and exploit the *non-anticipating* stochastic derivative introduced in Di Nunno and Eide (Stoch Anal Appl 28:54-85, 2009). We prove both a sufficient and necessary stochastic maximum principle.

Keywords Time-change · Conditionally independent increments · Backward stochastic Volterra integral equation · Maximum principle · Stochastic Volterra equations · Non-anticipating stochastic derivative

Mathematics Subject Classification 60H10 · 60H20 · 93E20 · 60G60 · 91B70

1 Introduction

Optimal harvesting is a fairly classical problem in control theory and it is still a timely question to address when thinking of sustainability in the management of natural resources. In this work we deal with a problem of optimal harvesting from a population, the growth of which is modelled by Volterra time dynamics of the type

$$X(t) = X_0 + \int_0^t (r(t, s) - Ku(s)) X(s) ds + \int_0^t \sigma(s) X(s) dB(s), \quad t \in [0, T]. \quad (1.1)$$

The term r represents the growth rate, the constant K is the catchability coefficient, and the control u is the fishing effort. The Volterra structure is inherited from the deterministic

✉ Michele Giordano
michelgi@math.uio.no

Giulia di Nunno
giulian@math.uio.no

¹ Department of Mathematics, University of Oslo, P. O. Box 1053, 0316 Blindern, Oslo, Norway

² Department of Business and Management Science, NHH Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway

analogous models that can be found, e.g., in Gushing (1979), Brauer (1976), Gripenberg et al. (1990). As we can see, this form of time dependence is often used in the description of fish populations. When considering fish as a commodity, the modelling of fish population is representing the possible dynamics of offer, in the interplay between offer and demand. In our work, however, we consider Volterra *stochastic* integral equations, which represent a natural extension including the uncertainty of the environment influencing the population growth. For this we are motivated by Brites and Braumann (2017), Aóvarez and Shepp (1998).

Our model has an element of novelty with respect to the others presented. This is given by the nature of the noise B which is associated to a time-changed Brownian motion. This is well motivated by the clustering effects that such noises can be described. For the description on how time-change helps to describe clustering, we can refer to a first discussion in (Shiryayev, 1999, Chapter IV, 3e) and a more recent study (Saakvitne, 2019, Chapter 3) in the context of market microstructure. Within population dynamics the evidence of clustering is largely discussed in the recent literature in biology and ecology. See just as example (Jokarab et al., 2021).

We remark that in the literature of mathematical finance, dynamics of the form (1.1), but with Lévy type noises were used in models (Agram et al., 2019). On the other side, time-change has been suggested in the study of volatility modelling, e.g. Barndorff-Nielsen and Shiryaev (2010), Carr and Wu (2004), Madan Geman and Yor (2001), Swishchuk (2016), Veraart and Winkel (2010), energy markets, e.g. Borovkova and Schmeck (2017), and default models, e.g. Mai and Scherer (2009). Also it is used in kinetic theory, see e.g. Menon (2017).

Keeping our motivation in mind, we treat here stochastic control for general Volterra type dynamics, allowing also for jumps:

$$X^u(t) = X_0 + \int_0^t b(t, s, \lambda_s, u(s), X^u(s-))ds + \int_0^t \int_{\mathbb{R}} \kappa(t, s, z, \lambda_s, u(s), X^u(s-))\mu(dsdz), \quad (1.2)$$

where the driving noise μ is given by the random measure

$$\mu(\Delta) = B(\Delta \cap [0, T] \times \{0\}) + \tilde{H}(\Delta \cap [0, T] \times \mathbb{R}_0), \quad \Delta \in \mathcal{B}([0, T] \times \mathbb{R}), \quad (1.3)$$

which is the mixture of a *conditional Gaussian measure* B on $[0, T] \times \{0\}$ and a *conditional centered Poisson measure* \tilde{H} on $[0, T] \times \mathbb{R}_0$. Here $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$ and \mathcal{B} represents the Borel σ -algebra. Both B and \tilde{H} are set in relationship with a time-changed Brownian motion and time-changed Poisson measure, respectively, via Theorem 3.1 in Serfozo (1972) [see also Grigelionis (1975)]. Note that the coefficients in (1.2) may also depend on the time-change via the process λ .

The time-change processes involved are of the form

$$\Lambda_t(\omega) = \int_0^t \lambda_s(\omega)ds, \quad (t, \omega) \in [0, T] \times \Omega,$$

($T > 0$). Thus the driving noises (which include jumps) are actually beyond the Brownian and the pure Lévy framework. We abandon noises with independent increments and effectively deal with quite general but still treatable martingales.

Our goal is to find the optimal control \hat{u} such that

$$J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathbb{F}}} J(u) = \sup_{u \in \mathcal{A}^{\mathbb{F}}} \mathbb{E} \left[\int_0^T F(t, \lambda_t, u(t), X^u(t))dt + G(X^u(T)) \right], \quad (1.4)$$

among the set $\mathcal{A}^{\mathbb{F}}$ of admissible \mathbb{F} -adapted controls, where $\mathbb{F} = \{\mathcal{F}_t, t \in [0, T]\}$ represents the smallest right-continuous filtration generated by μ .

Optimization problems such as (1.2), (1.4) are studied, e.g. in Bonaccorsi and Confortola (2020), Agram et al. (2018), Agram et al. (2019). In Agram et al. (2018), Agram et al. (2019) the authors present also a sufficient maximum principle and the dynamics include jumps making use of Malliavin calculus. However, being the restrictions on the domain of the Malliavin derivative extremely serious in the context of optimal control, the authors have lifted the study into the white noise framework and work with the Hida-Malliavin calculus on the space of stochastic distributions. The Hida-Malliavin calculus is tailored for Brownian and for centered Poisson random noises, hence this approach cannot be taken in our work since our driving noises are *not* of the required nature. On the other hand, in Bonaccorsi and Confortola (2020), the authors propose a backward SDE approach to solve (1.4). This is possible due to the introduction of memory in (1.2) by means of convolution with a completely monotone kernel which allows for a Markovian representation of the solution of (1.2).

Note that a Malliavin/Skorokhod calculus extension to noises with conditional independent increments, is proposed in Di Nunno and Sjursen (2013) and Yablonsky (2005). By this, however, we cannot solve the critical issue of the natural restriction of the domains of the involved operators and a Hida-Malliavin type extension is yet not available in the literature. Our approach is then to make use of the *non-anticipating* (NA) derivative. The NA derivative, introduced in Di Nunno (2002) for general martingales and then extended to martingale random fields in Di Nunno and Eide (2009), is the dual of the Itô integral and has an explicit representation in terms of limit of simple integrands in the Itô framework. Also, the NA derivative provides explicit stochastic integral representations. We stress that, contrarily to the Malliavin derivative, the domain of the NA derivative is the whole $L^2(dP)$, thus not creating problems in the context of optimal controls. To the best of our knowledge this is the first time that the non-anticipating derivative is used in optimal control problems such as (1.4).

Our approach to the optimization problem (1.4) is based on the analysis of the noise and the information flows associated. Indeed, we observe that there are two filtrations of interest. The first one is the already mentioned \mathbb{F} and the second is the filtration $\mathbb{G} := \{\mathcal{G}_t, t \in [0, T]\}$, where $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{F}^\Lambda$ generated by μ and the entire history \mathcal{F}^Λ of the time-change processes. Note that while \mathcal{F}_0 is substantially trivial, $\mathcal{G}_0 = \mathcal{F}^\Lambda$. We can regard \mathbb{G} as the initial enlargement of \mathbb{F} or, we can see \mathbb{F} as partial information with respect to \mathbb{G} . With this observation in hands, we work out the solution to problem (1.4) as an optimization problem under partial information. In this we have taken inspiration from Øksendal and Zhang (2010), where the concept of partial information is however not associated to the properties of the noise, and from Di Nunno and Sjursen (2014), where the dynamics are however not of Volterra type. Also, for completeness, we show that our techniques provide necessary and sufficient conditions for the optimization problem

$$J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathbb{G}}} J(u) = \sup_{u \in \mathcal{A}^{\mathbb{G}}} \mathbb{E} \left[\int_0^T F(t, \lambda_t, u(t), X^u(t)) dt + G(X^u(T)) \right] \quad (1.5)$$

on the set $\mathcal{A}^{\mathbb{G}}$ of admissible \mathbb{G} -adapted controls, where $\mathcal{A}^{\mathbb{F}} \subset \mathcal{A}^{\mathbb{G}}$.

The study of maximum principles is associated to a stochastic Hamiltonian map of the so-called dual variables, which in turn are obtained from the solution of a backward stochastic equation. In the sequel, we deal with backward stochastic differential equations (BSDEs) of type

$$p(t) = \xi(t) + \int_t^T g(s, \lambda_s, p(s-), q(s, \cdot)) ds - \int_t^T \int_{\mathbb{R}} q(s, z) \mu(ds dz), \quad (1.6)$$

under the filtration \mathbb{G} . Notice that, these backward equations are not of Volterra type. This is because our Hamiltonian functional is going to involve also the NA-derivatives of the adjoint process p . A different approach could have been to follow the work in Agram et al. (2019), where the authors deal with a backward stochastic Volterra integral equation (BSVIE) of the form:

$$p(t) = \xi(t) + \int_t^T g(s, \lambda_s, p(s^-), q(t, s, \cdot)) ds - \int_t^T \int_{\mathbb{R}} q(t, s, z) \mu(ds dz).$$

Even though this approach would allow us to work with simpler Hamiltonian functionals (in the sense that the NA-derivative of $p(t)$ would not be involved) we would need to assume smoothness conditions with respect to t on $q(t, s, z)$ and, to the best of our knowledge, is not clear to what extent those properties are satisfied.

Existence and uniqueness of (1.6) can be retrieved from Di Nunno and Sjursen (2014). The study of the BSDE under \mathbb{G} is in itself critically based on the stochastic integral representation in the form

$$\xi = \xi^0 + \int_0^T \int_{\mathbb{R}} \phi(s, z) \mu(ds dz), \quad (1.7)$$

where ξ^0 is \mathcal{G}_0 -measurable and the integrand ϕ is \mathbb{G} -predictable. These results are readily available in terms of their existence in the classical Kunita–Watanabe Theorem, while the *explicit* form of ϕ is given by means of the NA derivative in Di Nunno (2002) Theorem 3.1 and Di Nunno and Eide (2009) Theorem 3.1.

The paper is organized as follows. In the next section we give a presentation of the framework providing the necessary details for the random measure μ and the information flows that we are going to use. In Sect. 3 we prove a sufficient maximum principle and in Sect. 4 the corresponding necessary maximum principle. Lastly, we show how the results obtained can be applied to characterise the solution in the optimal harvesting problem associated to the dynamics (1.1).

2 The noise and the non-anticipating derivative

Let us consider a complete probability space (Ω, \mathcal{F}, P) and a time horizon $T < \infty$. We shall consider the noise on the time-space

$$\mathbb{X} := [0, T] \times \mathbb{R} := ([0, T] \times \{0\}) \cup ([0, T] \times \mathbb{R}_0),$$

where $\mathbb{R}_0 = \mathbb{R} \setminus \{0\}$. The Borel σ -algebra on \mathbb{X} is denoted $\mathcal{B}_{\mathbb{X}}$. Let \mathcal{L} be the space of the two dimensional stochastic processes $\lambda = (\lambda^B, \lambda^H)$ such that, for each component $k = B, H$, we have that

1. $\lambda_t^k \geq 0$ P - a.s. for all $t \in [0, T]$,
2. $\lim_{h \rightarrow 0} P(|\lambda_{t+h}^k - \lambda_t^k| \geq \epsilon) = 0$ for all $\epsilon > 0$ and almost all $t \in [0, T]$,
3. $\mathbb{E} \left[\int_0^T \lambda_t^k dt \right] < \infty$.

The processes $\lambda \in \mathcal{L}$ represent the *stochastic time-change rate*. Let ν be a σ -finite measure on the Borel sets of \mathbb{R}_0 satisfying $\int_{\mathbb{R}_0} z^2 \nu(dz) < \infty$. We define the random measure Δ on $\mathcal{B}_{\mathbb{X}}$ by

$$\Delta(\Delta) := \int_0^T \mathbb{1}_{\{(t,0) \in \Delta\}}(t) \lambda_t^B dt + \int_0^T \int_{\mathbb{R}_0} \mathbb{1}_{\Delta}(t, z) \nu(dz) \lambda_t^H dt, \quad \Delta \subseteq \mathbb{X}. \quad (2.1)$$

Furthermore, denote the restrictions of Λ to $[0, T] \times \{0\}$ and $[0, T] \times \mathbb{R}_0$ by Λ^B and Λ^H , respectively. For later use we also introduce the filtration

$$\mathbb{F}^\Lambda = \{\mathcal{F}_t^\Lambda, t \in [0, T]\},$$

where \mathcal{F}_t^Λ is generated by the values of Λ on the Borelian sets of $[0, t] \times \mathbb{R}$. Set $\mathcal{F}^\Lambda := \mathcal{F}_T^\Lambda$. We recall the following definitions.

Definition 2.1 The *conditional Gaussian measure* (given \mathcal{F}^Λ) B is a signed random measure on the Borel sets of $[0, T] \times \{0\}$ satisfying

- A1. $P(B(\Delta) \leq x | \mathcal{F}^\Lambda) = P(B(\Delta) \leq x | \Lambda^B(\Delta)) = \Phi\left(\frac{x}{\sqrt{\Lambda^B(\Delta)}}\right)$,
 $x \in \mathbb{R}, \Delta \subseteq [0, T] \times \{0\}$. Here Φ is the cumulative probability distribution function of a standard normal random variable.
- A2. For all disjoint $\Delta_1, \Delta_2 \subseteq [0, T] \times \{0\}$, $B(\Delta_1)$ and $B(\Delta_2)$ are conditionally independent given \mathcal{F}^Λ .

The *conditional Poisson measure* (given \mathcal{F}^Λ) H is a random measure on the Borel sets of $[0, T] \times \mathbb{R}_0$ satisfying

- A3. $P(H(\Delta) = k | \mathcal{F}^\Lambda) = P(H(\Delta) = k | \Lambda^H(\Delta)) = \frac{\Lambda^H(\Delta)^k}{k!} e^{-\Lambda^H(\Delta)}$, $k \in \mathbb{N}, \Delta \subseteq [0, T] \times \mathbb{R}_0$.
- A4. For all disjoint $\Delta_1, \Delta_2 \subseteq [0, T] \times \{\mathbb{R}_0\}$, $H(\Delta_1)$ and $H(\Delta_2)$ are conditionally independent given \mathcal{F}^Λ .

Moreover,

- A5. B and H are conditionally independent given \mathcal{F}^Λ .

Also the *conditional centered Poisson random measure* is defined as

$$\tilde{H}(\Delta) := H(\Delta) - \Lambda^H(\Delta), \quad \Delta \subset \mathbb{X}.$$

Observe that if λ^B and λ^H were deterministic, then B would be a Gaussian process and H a Poisson random measure. Furthermore, B would be a Wiener process if $\lambda^B \equiv 1$ and H a homogeneous Poisson random measure for $\lambda^H \equiv 1$.

Definition 2.2 We define the signed random measure μ on the Borel sets $\Delta \subseteq \mathbb{X}$ by

$$\mu(\Delta) := B(\Delta \cap [0, T] \times \{0\}) + \tilde{H}(\Delta \cap [0, T] \times \mathbb{R}_0).$$

The random measure μ has *conditionally independent values*, see Grigelionis (1975), Serfozo (1972). Observe that (A1) and (A3) yield

$$\mathbb{E}[\mu(\Delta) | \mathcal{F}^\Lambda] = 0, \quad \mathbb{E}[\mu(\Delta)^2 | \mathcal{F}^\Lambda] = \Lambda(\Delta), \quad \Delta \subseteq \mathbb{X}. \tag{2.2}$$

The random measures B and H are related to a time-changed Brownian motion and time-changed pure jump Lévy process. To illustrate, consider the processes on $[0, T]$:

$$\begin{aligned} B_t &:= B([0, t] \times \{0\}), & \Lambda_t^B &:= \int_0^t \lambda_s^B ds, \\ \eta_t &:= \int_0^t \int_{\mathbb{R}_0} z \tilde{H}(dsdz), & \Lambda_t^H &:= \int_0^t \lambda_s^H ds, \end{aligned}$$

and compute the characteristic functions of B and η . From (A1) and (A3) we have that

$$\mathbb{E} \left[e^{icB_t} \right] = \int_{\mathbb{R}} \mathbb{E} \left[e^{icB_t} | \Lambda_t^B = x \right] P_{\Lambda_t^B}(dx) = \int_{\mathbb{R}} e^{\frac{1}{2}c^2x} P_{\Lambda_t^B}(dx), \quad c \in \mathbb{R},$$

where $P_{\Lambda_t^B}$ is the probability distribution of the time-change Λ_t^B . Correspondingly, we have that

$$\mathbb{E} \left[e^{ic\eta_t} \right] = \int_{\mathbb{R}} \exp \left\{ \int_{\mathbb{R}_0} [e^{iczx} - 1 - iczx] \nu(dz) \right\} P_{\Lambda_t^H}(dx), \quad c \in \mathbb{R},$$

where $P_{\Lambda_t^H}$ is the probability distribution of the time-change Λ_t^H . Indeed we recall the following characterization (Serfozo 1972, Theorem 3.1):

Theorem 2.3 *Let $W_t, t \in [0, T]$, be a Brownian motion independent of Λ^B and $N_t, t \in [0, T]$, be a centered pure jump Lévy process with Lévy measure ν independent of Λ^H . Then B satisfies (A1)-(A2) if and only if, for any $t \geq 0$, $B_t \stackrel{d}{=} W_{\Lambda_t^B}$ and η satisfies (A3)-(A4) if and only if, for any $t \geq 0$, $\eta_t \stackrel{d}{=} N_{\Lambda_t^H}$.*

In the sequel we shall consider two types of information flows. The first one is represented by the filtration

$$\mathbb{F} := \{\mathcal{F}_t, t \in [0, T]\}, \quad \mathcal{F}_t := \bigcap_{r>t} \mathcal{F}_r^\mu,$$

where $\mathbb{F}^\mu := \{\mathcal{F}_t^\mu, t \in [0, T]\}$ is generated by the values $\mu(\Delta)$, $\Delta \subset [0, t] \times \mathbb{R}$, $t \in [0, T]$. Correspondingly, let $\mathbb{F}^B := \{\mathcal{F}_t^B, t \in [0, T]\}$ denote the filtration generated by $B(\Delta \cap [0, t] \times \{0\})$, and $\mathbb{F}^H := \{\mathcal{F}_t^H, t \in [0, T]\}$ the filtration generated by $H(\Delta \cap [0, t] \times \mathbb{R}_0)$. We remark that, for any $t \in [0, T]$, $\mathcal{F}_t^\mu = \mathcal{F}_t^B \vee \mathcal{F}_t^H \vee \mathcal{F}_t^\Lambda$. See Di Nunno and Sjursen (2013).

The second information flow of interest is

$$\mathbb{G} := \{\mathcal{G}_t, t \in [0, T]\}, \quad \mathcal{G}_t := \mathcal{F}_t^\mu \vee \mathcal{F}^\Lambda.$$

The filtration \mathbb{G} is right-continuous, see Di Nunno and Sjursen (2014). Moreover we note that $\mathcal{G}_T = \mathcal{F}_T$, $\mathcal{G}_0 = \mathcal{F}^\Lambda$, and \mathcal{F}_0 is substantially trivial. Namely, \mathbb{G} includes information on the future values of Λ^B and Λ^H . In the sequel we shall technically exploit the interplay between the two filtrations.

For $\Delta \subseteq (t, T] \times \mathbb{R}$, the conditional independence in (A2) and (A4), together with (2.2) yield

$$\mathbb{E}[\mu(\Delta) | \mathcal{G}_t] = \mathbb{E}[\mu(\Delta) | \mathcal{F}_t \vee \mathcal{F}^\Lambda] = \mathbb{E}[\mu(\Delta) | \mathcal{F}^\Lambda] = 0. \quad (2.3)$$

Moreover, (A5) gives us

$$\mathbb{E}[\mu(\Delta_1)\mu(\Delta_2) | \mathcal{G}_t] = \mathbb{E}[\mu(\Delta_1) | \mathcal{F}^\Lambda] \mathbb{E}[\mu(\Delta_2) | \mathcal{F}^\Lambda] = 0,$$

for disjoint $\Delta_1, \Delta_2 \subseteq (t, T] \times \mathbb{R}$. Hence, μ is a martingale random field with respect to \mathbb{G} , see e.g. Di Nunno and Eide (2009) Definition 2.1:

Definition 2.4 A square integrable martingale random field μ with conditionally orthogonal values is a stochastic set function $\mu(\Delta)$, $\Delta \subseteq \mathbb{X}$ such that

- $m(\Delta) := \mathbb{E}[\mu(\Delta)^2] = \mathbb{E}[\Lambda(\Delta)]$, $\Delta \subseteq \mathbb{X}$, defines a variance measure
- μ is \mathbb{G} -adapted
- μ satisfies the martingale property (2.3)

- μ has conditionally orthogonal values: $\mathbb{E}[\mu(\Delta_1)\mu(\Delta_2)|\mathcal{G}_t] = 0$, for every disjoint $\Delta_1, \Delta_2 \in (t, T] \times \mathbb{R}$.

It is immediate to see that μ is also a martingale random field with respect to \mathbb{F} .

With the above structures, we access the framework of Itô stochastic integration. For this we introduce $\mathcal{I}^{\mathbb{G}} \subseteq L^2(d\Lambda \times dP)$ representing the subspace of the random fields admitting a \mathbb{G} -predictable modification and $\mathcal{I}^{\mathbb{F}} \subset \mathcal{I}^{\mathbb{G}}$, the one of \mathbb{F} -predictable random fields. Observe that, for all $\phi \in \mathcal{I}^{\mathbb{G}}$, we have that

$$\mathbb{E} \left[\left(\iint_{\mathbb{X}} \phi(s, z) \mu(ds dz) \right)^2 \right] = \mathbb{E} \left[\iint_{\mathbb{X}} \phi(s, z)^2 \Lambda(ds dz) \right] \tag{2.4}$$

thanks to (A5) and the martingale property of μ .

In this work we shall make use of the non-anticipating derivative introduced in Di Nunno and Eide (2009) for martingale random fields.

Definition 2.5 The *non-anticipating derivative* (NA-derivative) \mathcal{D} is a linear operator defined for all the elements $\zeta \in L^2(dP)$ as the limit in $L^2(d\Lambda \times dP)$

$$\mathcal{D}\zeta := \lim_{n \rightarrow \infty} \varphi_n, \tag{2.5}$$

of simple \mathbb{G} -predictable random fields $\varphi_n, n \in \mathbb{N}$, defined as:

$$\varphi_n(t, x) := \sum_{k=1}^{K_n} \mathbb{E} \left[\zeta \frac{\mu(\Delta_{nk})}{\mathbb{E}[\Lambda(\Delta_{nk})|\mathcal{G}_{s_{nk}}]} \Big| \mathcal{G}_{s_{nk}} \right] \mathbb{1}_{\Delta_{nk}}(t, x), \quad (t, x) \in \mathbb{X}.$$

Here the Borel sets Δ_{nk} take the form $\Delta_{nk} := (s_{nk}, u_{nk}] \times B_{nk}, k = 1, \dots, K_n$, with $0 \leq s_{nk} \leq u_{nk} \leq T$, and $B_{nk} \in \mathfrak{B}$ where \mathfrak{B} is any countable semi-ring that generates the Borel σ -algebra $\mathcal{B}(\mathbb{R})$. Then $\bigcup_{n \in \mathbb{N}} \bigcup_{k=1}^{K_n} \Delta_{nk} = \mathbb{X}$. With a slight abuse of terminology we call the sets $\Delta_{nk}, k = 1, \dots, K_n$, a *partition* of \mathbb{X} with refinement n . Clearly all the sets $\Delta_{nk}, k = 1, \dots, K_n, n \in \mathbb{N}$ constitute a semiring generating $\mathcal{B}(\mathbb{X})$ see, e.g. Di Nunno and Eide (2009) and the references therein.

The NA-derivative allows for an explicit integral representation. Namely the integrand is characterized in terms of the inputs: the very random variable to represent, the integrator, and the filtration. See Theorem 3.1 in Di Nunno and Eide (2009).

Theorem 2.6 For any $\xi \in L^2(dP)$ the NA-derivative $\mathcal{D}\xi$ is well defined and the following stochastic integral representation holds

$$\xi = \xi^0 + \iint_{\mathbb{X}} \mathcal{D}_{t,z}\xi \mu(dt dz), \tag{2.6}$$

where $\xi^0 = \mathbb{E}[\xi | \mathcal{F}^\Lambda]$ satisfies $\mathcal{D}\xi^0 \equiv 0$.

The existence and unicity of a stochastic integral representation is well-known from the Kunita–Watanabe Theorem. Theorem 2.6 provides an explicit representation to the integrand. The spirit of this result is in line with representations à la Clark-Haussman-Ocone (CHO), see, e.g. Di Nunno et al. (2008). However in that case the noise is either a Brownian motion or a centered Poisson random measure and the integrand is characterized in terms of the Malliavin derivative. We remark that an extension of the Malliavin calculus and CHO representations to the conditional Brownian and the conditional Poisson cases is provided in Yablonsky (2005)

and Di Nunno and Sjursen (2014). When applying Malliavin calculus to optimal control, the domain of the Malliavin derivative constitutes a serious restriction as the variables depend on a control yet to be found. In Agram and Øksendal (2015) this was overcome for the Brownian and centered Poisson cases by using the Hida-Malliavin extension which is an extension of Malliavin calculus to the white noise framework (stochastic distributions), see Di Nunno et al. (2008). At present there is no such an extension for time-changed noises hence the method cannot be used. In this paper we suggest to use the NA-derivative, which has no restrictions on the domain and it is well defined for all martingales in $L^2(dP)$ as integrators. Furthermore, from Theorem 2.6 we can see that \mathcal{D} is actually the dual of the Itô integral:

Proposition 2.7 For all ϕ in $\mathcal{I}^{\mathbb{G}}$ and all ξ in $L^2(dP)$, we have

$$\mathbb{E} \left[\xi \iint_{\mathbb{X}} \phi(t, z) \mu(dt dz) \right] = \mathbb{E} \left[\iint_{\mathbb{X}} q(t, z) \mathcal{D}_{t,z} \xi \wedge(dt dz) \right].$$

Also we have the martingale representation theorem:

Theorem 2.8 For any square integrable \mathbb{G} martingale, $M(t)$, $t \in [0, T]$, the following representation holds true

$$M(t) = \mathbb{E}[M(T)|\mathcal{F}^{\Lambda}] + \int_0^t \int_{\mathbb{R}} \mathcal{D}_{s,z} M(T) \mu(ds dz).$$

For future use we also introduce the space S of the \mathbb{G} -adapted stochastic processes $p(t, \omega)$, $t \in [0, T]$, $\omega \in \Omega$ such that

$$\|p\|_S := \mathbb{E} \left[\sup_{0 \leq t \leq T} |p(t)|^2 \right]^{1/2} < \infty.$$

3 A sufficient maximum principle with time-change

We are now ready to study the optimization problem (1.4) with performance functional

$$J(u) = \mathbb{E} \left[\int_0^T F(t, \lambda_t, u(t), X^u(t)) dt + G(X^u(T)) \right], \quad (3.1)$$

where

$$\begin{aligned} F &: [0, T] \times [0, \infty)^2 \times \mathcal{U} \times \mathbb{R} \times \Omega \longrightarrow \mathbb{R}, \\ G &: \mathbb{R} \times \Omega \longrightarrow \mathbb{R}, \end{aligned}$$

with \mathcal{U} a closed convex subset of \mathbb{R} . For all $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$, $x \in \mathbb{R}$ the process $F(\cdot, \lambda, u, x, \cdot)$ is \mathbb{F} -adapted and the mapping $F(t, \lambda, u, x)$ is C^1 in x P -a.s. uniformly w.r.t. $t \in [0, T]$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$. Also for all $x \in \mathbb{R}$, $G(x, \cdot)$ is \mathcal{F}_T -measurable and G is C^1 in x P -a.s. uniformly w.r.t. $t \in [0, T]$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$. The controlled dynamics of X are given by the equation

$$\begin{aligned} X^u(t) &= X_0 + \int_0^t b(t, s, \lambda_s, u(s), X^u(s-)) ds \\ &\quad + \int_0^t \int_{\mathbb{R}} \kappa(t, s, z, \lambda_s, u(s), X^u(s-)) \mu(ds dz), \end{aligned} \quad (3.2)$$

where $X_0 \in \mathbb{R}$ and the coefficients are given by the mappings

$$\begin{aligned}
 b &: [0, T] \times [0, T] \times [0, \infty)^2 \times \mathcal{U} \times \mathbb{R} \times \Omega \longrightarrow \mathbb{R}, \\
 \kappa &: [0, T] \times [0, T] \times \mathbb{R} \times [0, \infty)^2 \times \mathcal{U} \times \mathbb{R} \times \Omega \longrightarrow \mathbb{R}.
 \end{aligned}$$

We assume $b(t, \cdot, \lambda, u, x, \cdot)$ and $\kappa(t, \cdot, z, \lambda, u, x, \cdot)$ to be \mathbb{F} -predictable for all $t \in [0, T], \lambda \in [0, \infty)^2, u \in \mathcal{U}, x \in \mathbb{R}$ and $z \in \mathbb{R}$. We also require them to be C^2 with respect to t and to x with partial derivatives L^2 -integrable with respect to $dt \times dP$ and $d\Lambda \times dP$, respectively. Notice also that we will often drop the superscript u when it is clear the dependence of X on u .

Later on we can see the coefficients b and κ in a functional setup:

$$\begin{aligned}
 b &: [0, T] \times [0, T] \times \Xi_{\mathbb{R}_+^2} \times \Xi_{\mathcal{U}} \times \Xi_{\mathbb{R}} \times \Omega \longrightarrow \mathbb{R}, \\
 \kappa &: [0, T] \times [0, T] \times \mathbb{R} \times \Xi_{\mathbb{R}_+^2} \times \Xi_{\mathcal{U}} \times \Xi_{\mathbb{R}} \times \Omega \longrightarrow \mathbb{R},
 \end{aligned}$$

where we denoted by Ξ_S the space of measurable function on $[0, T]$ with values in S . Then we can interpret the coefficients in (3.2) via the evaluation at the point $s \in [0, T]$:

$$\begin{aligned}
 b(t, \cdot, \lambda, u, X^u(\cdot))(s) &= b(t, s, \lambda_s, u(s), X^u(s-)) \\
 \kappa(t, \cdot, z, \lambda, u, X^u(\cdot))(s) &= \kappa(t, s, z, \lambda_s, u(s), X^u(s-)).
 \end{aligned}$$

We assume that b and κ are Fréchet differentiable (in the standard topology of càdlàg paths) with C^2 regularity in t, x and u (with the corresponding derivatives).

In the sequel we assume existence and uniqueness of a solution for (3.2). Sufficient conditions for this are provided in the next result, which is in line with the study in Agram et al. (2019), though there the driving noises are the Brownian motion and Poisson random measure.

Theorem 3.1 *Assume that:*

1. $b(t, \cdot, \lambda, u, x, \cdot)$ and $\kappa(t, \cdot, z, \lambda, u, x, \cdot)$ are \mathbb{F} -predictable for all $t \in [0, T], z \in \mathbb{R}, \lambda \in [0, \infty)^2, u \in \mathcal{U}$ and $x \in \mathbb{R}$.
2. $b(t, s, \lambda, u, \cdot)$ and $\kappa(t, s, \cdot, \lambda, u, \cdot)$ are Lipschitz continuous with respect to x , uniformly in $t, s \in [0, T]^2, u \in \mathcal{U}, \lambda \in [0, \infty)^2$, i.e., for all $x_1, x_2 \in \mathbb{R}$,

$$\begin{aligned}
 &|b(t, s, \lambda, u, x_1) - b(t, s, \lambda, u, x_2)| + |\kappa(t, s, 0, \lambda, u, x_1) - \kappa(t, s, 0, \lambda, u, x_2)|\sqrt{\lambda^B} \\
 &+ \left(\int_{\mathbb{R}_0} |\kappa(t, s, z, \lambda, u, x_1) - \kappa(t, s, z, \lambda, u, x_2)|^2 \nu(dz) \right)^{1/2} \sqrt{\lambda^H} \\
 &\leq C|x_1 - x_2|, \quad P - a.s.
 \end{aligned}$$

3. $b(t, s, \lambda, u, \cdot)$ and $\kappa(t, s, z, \lambda, u, \cdot)$ have linear growth with respect to x , i.e., for all $t, s \in [0, T]^2, u \in \mathcal{U}, \lambda \in [0, \infty)^2, x \in \mathbb{R}$, we have

$$\begin{aligned}
 &|b(t, s, \lambda, u, x)| + |\kappa(t, s, 0, \lambda, u, x)|\sqrt{\lambda^B} \\
 &+ \left(\int_{\mathbb{R}_0} |\kappa(t, s, z, \lambda, u, x)|^2 \nu(dz) \right)^{1/2} \sqrt{\lambda^H} \leq C(1 + |x|) \quad P - a.s.
 \end{aligned}$$

Then there exists a unique \mathbb{F} -adapted solution to (3.2) in $L^2(dt \times dP)$.

Proof The proof follows a classical Picard iteration scheme. Here we provide the main ideas. Fix $u \in \mathcal{A}^{\mathbb{R}}$ and define inductively

$$\begin{aligned} X^0(t) &:= X_0 \\ X^n(t) &:= X_0 + \int_0^t b(t, s, \lambda_s, u(s), X^{n-1}(s)) ds \\ &\quad + \int_0^t \int_{\mathbb{R}} \kappa(t, s, z, \lambda_s, u(s), X^{n-1}(s-)) \mu(ds dz), \quad t \in [0, T], \quad n \geq 1 \end{aligned}$$

Then, for all $t \in [0, T]$ and for all $n \geq 1$, we have the following estimate

$$\begin{aligned} \mathbb{E} [|X^{n+1}(t) - X^n(t)|^2] &\leq 2\mathbb{E} \left[t \int_0^t |b(t, s, \lambda_s, u(s), X^n(s-)) \right. \\ &\quad \left. - b(t, s, \lambda_s, u(s), X^{n-1}(s-))|^2 ds \right] \\ &\quad + 2\mathbb{E} \left[\int_0^t \int_{\mathbb{R}} |\kappa(t, s, z, \lambda_s, u(s), X^n(s-)) - \kappa(t, s, z, \lambda_s, u(s), X^{n-1}(s-))|^2 \Lambda(ds dz) \right]. \end{aligned}$$

By (2.1) and using the Lipschitz condition on b and κ , we get

$$\mathbb{E} [|X^{n+1}(t) - X^n(t)|^2] \leq 2C^2 \mathbb{E} \left[t \int_0^t 2|X^n(s-) - X^{n-1}(s-)| ds \right],$$

which leads to

$$\mathbb{E} [|X^{n+1}(t) - X^n(t)|^2] \leq K \mathbb{E} \left[\int_0^t |X^n(s-) - X^{n-1}(s-)|^2 ds \right], \quad (3.3)$$

for $K := 4TC^2$. Also, by the linear growth condition 3. on b and κ , we get that

$$\mathbb{E} [|X^1(t) - X^0(t)|] \leq Kt(1 + X_0)^2. \quad (3.4)$$

Combining now (3.3) and (3.4), we have that

$$\mathbb{E} [|X^{n+1}(t) - X^n(t)|^2] \leq \frac{2(1 + X_0)^2 (Kt)^{n+1}}{(n+1)!}.$$

Thus we have that $\{X(t)^n\}_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(dP)$ and $\{X(t)^n\}_{n=1}^{\infty}$ is in $L^2(dP \times dt)$. Taking the limit on $n \rightarrow \infty$ gives the solution to (3.2). The uniqueness is obtained by standard arguments and estimates similar to the ones above. \square

Before moving forward, we need to state a fundamental result that will allow us to rewrite X in (3.2) in differential form. This is due to Protter (1993) and it is known as transformation rule. Hereafter we state the result within our setting.

Lemma 3.2 (Transformation rule) *Assume that for all $z \in \mathbb{R}$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$, $x \in \mathbb{R}$ the partial derivative of κ with respect to t (denoted with $\partial_t \kappa(t, s, z, \lambda, u, x)$) is locally bounded (uniformly in t) and satisfies*

$$|\partial_t \kappa(t_1, s, z, \lambda, u, x) - \partial_t \kappa(t_2, s, z, \lambda, u, x)| \leq K|t_1 - t_2|, \quad (3.5)$$

for some $K > 0$ and for each fixed $s \leq t$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$, $x \in \mathbb{R}$.

Then, the forward equation (3.2) can be rewritten in differential notation as

$$dX(t) = \left(b(t, t, \lambda_t, u(t), X(t)) + \int_0^t \partial_t b(t, s, \lambda_s, u(s), X(s)) ds \right)$$

$$\begin{aligned}
 & + \int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z, \lambda_s, u(s), X(s)) \mu(dsdz) dt \\
 & + \int_{\mathbb{R}} \kappa(t, t, z, \lambda_t, u(t), X(t)) \mu(dt dz). \tag{3.6}
 \end{aligned}$$

Proof The proof follows the one in Protter (1993). We report it here for completeness. Observe that

$$\begin{aligned}
 X(t) &= \int_0^t b(t, s, \lambda_s, u(s), X(s)) ds + \int_0^t \int_{\mathbb{R}} \kappa(t, s, z, \lambda_s, u(s), X(s)) \mu(dsdz) \\
 &= \int_0^t b(t, s, \lambda_s, u(s), X(s)) ds + \int_0^t \int_{\mathbb{R}} \kappa(s, s, z, \lambda_s, u(s), X(s)) \mu(dsdz) \\
 &\quad + \int_0^t \int_{\mathbb{R}} \kappa(t, s, z, \lambda_s, u(s), X(s)) - \kappa(s, s, z, \lambda_s, u(s), X(s)) \mu(dsdz)
 \end{aligned}$$

Note that

$$\begin{aligned}
 \kappa(t, s, z, \lambda_s, u(s), X(s)) - \kappa(s, s, z, \lambda_s, u(s), X(s)) &= \int_s^t \partial_r \kappa(r, s, z, \lambda_s, u(s), X(s)) dr \\
 &= \int_0^t \mathbb{1}_{s \leq r} \partial_r \kappa(r, s, z, \lambda_s, u(s), X(s)) dr
 \end{aligned}$$

Then we can apply the Fubini theorem for stochastic integration as in Jacod (1979) and we obtain that

$$\begin{aligned}
 & \int_0^t \int_{\mathbb{R}} \kappa(t, s, z, \lambda_s, u(s), X(s)) - \kappa(s, s, z, \lambda_s, u(s), X(s)) \mu(dsdz) \\
 &= \int_0^t \int_{\mathbb{R}} \left\{ \int_0^t \mathbb{1}_{s \leq r} \partial_r \kappa(r, s, z, \lambda_s, u(s), X(s)) dr \right\} \mu(dsdz) \\
 &= \int_0^t \left\{ \int_0^r \int_{\mathbb{R}} \partial_r \kappa(r, s, z, \lambda_s, u(s), X(s)) \mu(dsdz) \right\} dr.
 \end{aligned}$$

The well posedness and the Lebesgue integrability of $\int_0^r \int_{\mathbb{R}} \partial_r \kappa(r, s, z, \lambda_s, u(s), X(s)) \mu(dsdz)$, $r \in [0, t]$ is achieved in Theorem 3.2 Protter (1993) thanks to (3.5). \square

Remark 3.3 (A link with functional SDEs) Lemma 3.2, suggests a link between the Volterra integral equations of the kind (3.2) and functional SDEs (FSDEs). It is in fact clear that, by defining

$$B(t, \lambda., u., X., Z.) := \left(b(t, t, \lambda_t, u(t), X(t)) + \int_0^t \partial_t b(t, s, \lambda_s, u(s), X(s)) ds + Z(t) \right),$$

where

$$Z(t) = \int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z, \lambda_s, u(s), X(s)) \mu(dsdz),$$

we have that (3.2) can be rewritten as

$$X(t) = X_0 + \int_0^t B(t, \lambda., u., X., Z.) dt + \int_0^t \int_{\mathbb{R}} \kappa(t, t, z, \lambda_t, u(t), X(t)) \mu(dt dz). \tag{3.7}$$

We notice that (3.7) is a functional SDE, so we could have tried to state an existence result for functional SDEs instead of using Theorem 3.1. Some existence results for SDEs such as

(3.7) are available [see e.g. Cont and Kalinin (2020), Possamai et al. (2020), Kromer et al. (2015), Baños et al. (2019), Dahl et al. (2016)], but no one of those deals with noises such as μ . While some of those results [e.g. Cont and Kalinin (2020), Possamai et al. (2020), Kromer et al. (2015)] present condition that would be too restrictive for the current setting, we also point out that the results presented in Dahl et al. (2016), Baños et al. (2019) could possibly be extended to the current framework. Nonetheless, this would require to impose some Lipschitz and linear growth conditions on b and κ (like in Theorem 3.1) and, additionally, to impose a Lipschitzianity condition on $\partial_t b$, not required in the hypothesis of Theorem 3.1.

Having discussed the existence of a solution for (3.2), we are finally ready to proceed to our optimization results. We start by introducing the notion of admissible controls:

Definition 3.4 The admissible controls for (3.2) in the optimization problems (1.4) and (1.5) are predictable stochastic processes $u : [0, T] \times \Omega \mapsto \mathcal{U}$ such that X in (3.2) has a unique strong solution and

$$\mathbb{E} \left[\int_0^T F(t, \lambda_t, u(t), X(t)) dt + G(X(T)) + |\partial_x G(X(T))|^2 \right] < \infty$$

We denote $\mathcal{A}^{\mathbb{F}}$ and $\mathcal{A}^{\mathbb{G}}$ the sets of \mathbb{F} - or \mathbb{G} -predictable controls, respectively. We say that (\hat{u}, \hat{X}) is an *optimal pair* if

$$J(\hat{u}) = \sup_{u \in \mathcal{A}} \mathbb{E} \left[\int_0^T F(t, \lambda_t, u(t), X(t)) dt + G(X(T)) \right], \quad (3.8)$$

where $\hat{X} := X^{\hat{u}}$ is as in (3.2), and \mathcal{A} is either the set $\mathcal{A}^{\mathbb{F}}$ or $\mathcal{A}^{\mathbb{G}}$.

Define $\mathcal{R}^{\mathbb{G}}$ to be the space of \mathbb{G} -predictable processes with values in $L^2(dP)$. We remark that, if $y \in \mathcal{R}^{\mathbb{G}}$, then the NA-derivative (2.5) is also in $\mathcal{R}^{\mathbb{G}}$ i.e. for all t, z $\mathcal{D}_{t,z} y(\cdot) \in \mathcal{R}^{\mathbb{G}}$. In the sequel, when no confusion arises, we will denote with $\mathcal{D}_{t,0} y(\cdot)$ the NA-derivative with respect to the conditional Brownian motion, and with $\mathcal{D}_{t,z} y(\cdot)$, $z \in \mathbb{R}_0$, the NA-derivative with respect to the conditional Poisson random measure.

In view of the Volterra structure of the dynamics (3.2), the system is not Markovian. We tackle the problem (3.8) by the maximum principle approach, better suited in this case, see e.g. Yong and Zhou (1999). We introduce the Hamiltonian function:

$$\mathcal{H} : [0, T] \times \mathfrak{E}_{\mathbb{R}_+^2} \times \mathfrak{E}_{\mathcal{U}} \times \mathfrak{E}_{\mathbb{R}} \times \mathcal{R}^{\mathbb{G}} \times \mathfrak{E}_{\mathcal{Z}} \times \Omega \longrightarrow \mathbb{R},$$

as the mapping given by the sum

$$\mathcal{H}(t, \lambda, u, x, p, q) := H_0(t, \lambda, u, x, p, q) + H_1(t, \lambda, u, x, p, q) \quad (3.9)$$

of the two components

$$\begin{aligned} H_0(t, \lambda, u, x, p, q) &:= F(t, \lambda_t, u_t, x_t) + b(t, t, \lambda_t, u_t, x_t) p(t) + \kappa(t, t, 0, \lambda_t, u_t, x_t) q_t(0) \lambda_t^B \\ &\quad + \int_{\mathbb{R}_0} \kappa(t, t, z, \lambda_t, u_t, x_t) q_t(z) \lambda_t^H v(dz) \\ H_1(t, \lambda, u, x, p, q) &:= \int_0^t \partial_t b(t, s, \lambda_s, u_s, x_s) ds p(t) \\ &\quad + \int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z, \lambda_s, u_s, x_s) \mathcal{D}_{s,z} p(t) \Lambda(ds dz), \end{aligned}$$

where \mathcal{Z} is the space of functions $q : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$|q(0)|^2 + \int_{\mathbb{R}_0} q(z)^2 \nu(dz) < \infty.$$

Remark 3.5 Following up on Remark 3.3, instead of considering (3.2) as a Volterra equation, we could have taken the FSDE (3.7) and, following e.g. Dahl et al. (2016), write the Hamiltonian functional for the functional SDE. We notice that, regardless of the chosen approach, we would still end up with the Hamiltonian functional (3.9).

Associated to \mathcal{H} (3.9), we introduce a BSDE of the type (1.6), which we study under \mathbb{G} :

$$p(t) = \partial_x G(X(T)) + \int_t^T \partial_x \mathcal{H}(s, \lambda_s, u(s), X(s^-), p(s^-), q(s, \cdot)) ds - \int_t^T \int_{\mathbb{R}} q(s, z) \mu(ds dz), \quad t \in [0, T], \tag{3.10}$$

where the derivative $\partial_x \mathcal{H}$ is meant in the Fréchet sense.

Sufficient conditions to guarantee the existence of (3.10) on $\mathcal{R}^{\mathbb{G}} \times \mathcal{I}^{\mathbb{G}}$ can be found in Di Nunno and Sjursen (2014).

Remark 3.6 Notice that (3.10) is actually a BSDE and not a Volterra-type backward SDE. In fact, the term $\partial_x H_1(t, \lambda, u, X, p, q)$ in the driver $\partial_x \mathcal{H}(t, \lambda, u, X, p, q)$, corresponds to

$$\begin{aligned} \partial_x H_1(s, \lambda_s, u, X, p, q) &= \partial_x \int_0^s \partial_s b(s, r, \lambda_r, u(r), X(r)) dr p(s) \\ &+ \partial_x \int_0^s \int_{\mathbb{R}} \partial_s \kappa(s, r, z, \lambda_r, u(r), X(r)) \mathcal{D}_{r,z} p(s) \Lambda(dr dz), \end{aligned}$$

which is a function of time s , after integration.

The optimal control problem (1.4):

$$J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathbb{F}}} J(u) = \sup_{u \in \mathcal{A}^{\mathbb{F}}} \mathbb{E} \left[\int_0^T F(t, \lambda_t, u(t), X^u(t)) dt + G(X^u(T)) \right], \tag{3.11}$$

associated to the performance functional (3.1) is treated in the framework of optimization under partial information. This is inspired by Di Nunno and Sjursen (2014), where this approach is taken for standard time-changed dynamics. In the Volterra case treated in the present work, the functionals stemming out of (3.9) are very different from the ones in Di Nunno and Sjursen (2014). Indeed we introduce the mapping $\mathcal{H}^{\mathbb{F}}$ defined for $t \in [0, T]$, $\lambda \in \mathfrak{E}_{\mathbb{R}^2_+}$, $u \in \mathfrak{E}_{\mathcal{U}}$, $x \in \mathfrak{E}_{\mathbb{R}}$, $p \in \mathcal{R}^{\mathbb{G}}$ and $q \in \mathcal{I}^{\mathbb{G}}$ as

$$\begin{aligned} \mathcal{H}^{\mathbb{F}}(t, \lambda, u, x, p, q) &:= H_0^{\mathbb{F}}(t, \lambda, u, x, p, q) + H_1^{\mathbb{F}}(t, \lambda, u, x, p, q) \\ &:= \mathbb{E} [\mathcal{H}(t, \lambda, u, x, p, q) | \mathcal{F}_t], \end{aligned} \tag{3.12}$$

where

$$\begin{aligned} H_0^{\mathbb{F}}(t, \lambda, u, x, p, q) &:= F(t, \lambda_t, u_t, x_t) + b(t, t, \lambda_t, u_t, x_t) \mathbb{E}[p(t) | \mathcal{F}_t] \\ &+ \kappa(t, t, 0, \lambda_t, u_t, x_t) \mathbb{E}[q(t, 0) | \mathcal{F}_t] \lambda_t^B \\ &+ \int_{\mathbb{R}_0} \kappa(t, t, z, \lambda_t, u_t, x_t) \mathbb{E}[q(t, z) | \mathcal{F}_t] \lambda_t^H \nu(dz) \end{aligned}$$

$$H_1^{\mathbb{F}}(t, \lambda, u, x, p, q) := \int_0^t \partial_t b(t, s, \lambda_s, u_s, x_s) ds \mathbb{E}[p(t)|\mathcal{F}_t] \\ + \int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z, \lambda_s, u_s, x_s) \mathbb{E}[\mathcal{D}_{s,z} p(t)|\mathcal{F}_t] \Lambda(ds dz)$$

Notation 1 Given $u, \hat{u} \in \mathcal{A}$, X, \hat{X} represent the associated controlled dynamics of (3.2) and $(p, q), (\hat{p}, \hat{q})$ are the corresponding solutions of (3.10). From now on, if no confusion arises, we will use the compact notation:

$$b(t, s) := b(t, s, \lambda_s, u(s), X(s)), \quad \hat{b}(t, s) := b(t, s, \lambda_s, \hat{u}(s), \hat{X}(s)).$$

Similarly, for $\kappa, \hat{\kappa}, F, \hat{F}, G, \hat{G}$, we will also write:

$$\mathcal{H}^u(s) := \mathcal{H}(s, \lambda, u, X, \hat{p}, \hat{q}), \quad \mathcal{H}^{\hat{u}}(s) := \mathcal{H}(s, \lambda, \hat{u}, \hat{X}, \hat{p}, \hat{q})$$

and similarly for $\mathcal{H}^{\mathbb{F},u}, \mathcal{H}^{\mathbb{F},\hat{u}}, H_0^u, H_0^{\hat{u}}, H_0^{\mathbb{F},u}, H_0^{\mathbb{F},\hat{u}}, H_1^u, H_1^{\hat{u}}, H_1^{\mathbb{F},u}, H_1^{\mathbb{F},\hat{u}}$.

Theorem 3.7 (Sufficient maximum principle with respect to \mathbb{F}). *Let $\lambda \in \mathcal{L}$. Let $\hat{u} \in \mathcal{A}^{\mathbb{F}}$ and assume that the corresponding solutions $\hat{X}, (\hat{p}, \hat{q})$ of (3.2) and (3.10) exist. Assume that*

- $x \mapsto G(x)$ is concave.
- For any t , the map

$$x \mapsto \text{ess sup}_{u \in \Xi_{\mathcal{U}}} \mathcal{H}^{\mathbb{F}}(t, \lambda, u, x, \hat{p}, \hat{q}), \quad x \in \Xi_{\mathbb{R}}, \quad (3.13)$$

is concave.

- For all $t \in [0, T]$,

$$\text{ess sup}_{u \in \Xi_{\mathcal{U}}} \mathcal{H}^{\mathbb{F}}(t, \lambda, u, x, \hat{p}, \hat{q}) = \mathcal{H}^{\mathbb{F}}(t, \lambda, \hat{u}, \hat{X}, \hat{p}, \hat{q}). \quad (3.14)$$

Then \hat{u} is an optimal control for problem (3.1) and (\hat{u}, \hat{X}) is an optimal pair.

Proof This proof is inspired by both the proof of Agram et al. (2018) Theorem 4.1 and Di Nunno and Sjursen (2014) Theorem 6.2. The main difference with Agram et al. (2018) is the use of the random measure μ instead of a Brownian motion and a compensated random Poisson measure, which requires to abandon the framework of Malliavin calculus. The main difference with Di Nunno and Sjursen (2014) is the Volterra structure of the dynamics for the forward equation (3.2), which lead to more involved stochastic calculus. Recall that $\hat{u} \in \mathcal{A}^{\mathbb{F}}$ is a candidate to be optimal and $X^{\hat{u}}$ is the corresponding solution of (3.2). Choose an arbitrary other $u \in \mathcal{A}^{\mathbb{F}}$ with corresponding controlled dynamics X and consider $J(u) - J(\hat{u}) = I_1 + I_2$, where

$$I_1 := \mathbb{E} \left[\int_0^T F(t, \lambda_t, u(t), X(t)) - F(t, \lambda_t, \hat{u}(t), \hat{X}(t)) dt \right], \quad (3.15)$$

$$I_2 := \mathbb{E} \left[G(X(T)) - G(\hat{X}(T)) \right]. \quad (3.16)$$

Considering now I_1 , from the definition of $H_0^{\mathbb{F}}$ we get that,

$$I_1 = \mathbb{E} \left[\int_0^T \left\{ H_0^{\mathbb{F},u}(t) - H_0^{\mathbb{F},\hat{u}}(t) - [b(t, t) - \hat{b}(t, t)] \mathbb{E}[\hat{p}(t)|\mathcal{F}_t] \right\} dt \right. \\ \left. - \int_0^T \int_{\mathbb{R}} [\kappa(t, t, z) - \hat{\kappa}(t, t, z)] \mathbb{E}[\hat{q}(t, z)|\mathcal{F}_t] \Lambda(dt dz) \right].$$

By the concavity of G , we have

$$I_2 \leq \mathbb{E} \left[\partial_x G(\hat{X}(T)) \left(X(T) - \hat{X}(T) \right) \right] = \mathbb{E} \left[\hat{p}(T) \left(X(T) - \hat{X}(T) \right) \right].$$

We apply the transformation rule (Lemma 3.2) to rewrite the Volterra forward dynamics of X as

$$dX(t) = \left(b(t, t) + \int_0^t \partial_t b(t, s) ds + \int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z) \mu(ds dz) \right) dt + \int_{\mathbb{R}} \kappa(t, t, z) \mu(dt dz).$$

Also the BSDE \hat{p} (3.10) associated to the optimal pair (\hat{u}, \hat{X}) in differential notation is

$$d\hat{p}(t) = -\partial_x \mathcal{H}^{\hat{u}}(t) dt + \int_{\mathbb{R}} \hat{q}(t, z) \mu(dt dz).$$

Using the Itô formula for the product we obtain

$$\begin{aligned} I_2 \leq & \mathbb{E} \left[\int_0^T \left\{ \hat{p}(t) \left((b(t, t) - \hat{b}(t, t)) + \int_0^t (\partial_t b(t, s) - \partial_t \hat{b}(t, s)) ds \right. \right. \right. \\ & \left. \left. + \int_0^t \int_{\mathbb{R}} (\partial_t \kappa(t, s, z) - \partial_t \hat{\kappa}(t, s, z)) \mu(ds dz) \right) \right\} dt \\ & - \int_0^T \partial_x \mathcal{H}^{\hat{u}}(t) \left(X(t) - \hat{X}(t) \right) dt + \int_0^T \left\{ [\kappa(t, t, 0) - \hat{\kappa}(t, t, 0)] \hat{q}(t, 0) \lambda_t^B \right. \\ & \left. + \int_{\mathbb{R}_0} [\kappa(t, t, z) - \hat{\kappa}(t, t, z)] \hat{q}(t, z) \nu(dz) \lambda_t^H \right\} dt \right]. \end{aligned} \tag{3.17}$$

Now notice that,

$$\begin{aligned} \mathbb{E} \left[\int_0^T \left(\int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z) \mu(ds dz) \right) \hat{p}(t) dt \right] &= \int_0^T \mathbb{E} \left[\left(\int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z) \mu(ds dz) \right) \hat{p}(t) \right] dt \\ &= \int_0^T \mathbb{E} \left[\int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z) \mathcal{D}_{s,z} \hat{p}(t) \Lambda(ds dz) \right] dt \\ &= \mathbb{E} \left[\int_0^T \int_0^t \int_{\mathbb{R}} \partial_t \kappa(t, s, z) \mathcal{D}_{s,z} \hat{p}(t) \Lambda(ds dz) dt \right] \end{aligned} \tag{3.18}$$

where we have used Fubini’s theorem and the duality formula (Proposition 2.7). By substituting (3.18) into (3.17), and taking the conditional expectation given \mathcal{F}_t we get that

$$\begin{aligned} I_2 \leq & \mathbb{E} \left[\int_0^T \left\{ (b(t, t) - \hat{b}(t, t)) \mathbb{E}[\hat{p}(t) | \mathcal{F}_t] + \int_0^t (\partial_t b(t, s) - \partial_t \hat{b}(t, s)) ds \mathbb{E}[\hat{p}(t) | \mathcal{F}_t] \right. \right. \\ & \left. \left. + \int_0^t \int_{\mathbb{R}_0} (\partial_t \kappa(t, s, z) - \partial_t \hat{\kappa}(t, s, z)) \mathbb{E}[\mathcal{D}_{s,z} \hat{p}(t) | \mathcal{F}_t] \Lambda(ds dz) \right\} dt \right. \\ & \left. - \int_0^T \partial_x \mathcal{H}^{\hat{u}}(t) \left(X(t) - \hat{X}(t) \right) dt + \int_0^T \int_{\mathbb{R}} \mathbb{E}[\hat{q}(t, z) | \mathcal{F}_t] [\kappa(t, t, z) - \hat{\kappa}(t, t, z)] \Lambda(dt dz) \right]. \end{aligned}$$

Hence

$$\begin{aligned} I_1 + I_2 \leq & \mathbb{E} \left[\int_0^T \left(H_0^{\mathbb{F}, \hat{u}}(t) - H_0^{\mathbb{F}, \hat{u}}(t) + H_1^{\mathbb{F}, \hat{u}}(t) \right. \right. \\ & \left. \left. - H_1^{\mathbb{F}, \hat{u}}(t) - \partial_x \mathcal{H}^{\mathbb{F}, \hat{u}}(t) \left(X(t) - \hat{X}(t) \right) \right) dt \right] \\ & = \mathbb{E} \left[\int_0^T \left(\mathcal{H}^{\mathbb{F}, \hat{u}}(t) - \mathcal{H}^{\mathbb{F}, \hat{u}}(t) - \partial_x \mathcal{H}^{\mathbb{F}, \hat{u}}(t) \left(X(t) - \hat{X}(t) \right) \right) dt \right] \leq 0, \end{aligned} \tag{3.19}$$

$dt \times dP$ a.e. by the maximality of \hat{u} in (3.14) and the concavity condition (3.13). Hence $J(u) \leq J(\hat{u})$ and \hat{u} is an optimal control for (3.1). This conclusion is reached applying a separating hyperplane argument to the concave map (3.13). \square

Notice that a result analogous to Theorem 3.7 can also be obtained when working under the initially enlarged filtration \mathbb{G} . Though the next result might not be of direct applicability in view of the anticipated information included in \mathbb{G} , the study has mathematical validity.

Remark 3.8 The transformation rule under (3.5) allows for the use of an Itô-type formula in the context of Volterra dynamics. If the equation would not present Volterra structure in the stochastic integral part (i.e. in the coefficient κ), then the requirement (3.5) is clearly lifted.

Proposition 3.9 (Sufficient maximum principle with respect to \mathbb{G}). *Let $\lambda \in \mathcal{L}$. Let $\hat{u} \in \mathcal{A}^{\mathbb{G}}$ and assume that the corresponding solutions $\hat{X}(t)$, (\hat{p}, \hat{q}) of (3.2) and (3.10) exist. Assume that:*

- $x \mapsto G(x)$ is concave.
- For any t , \hat{p} , \hat{q} , the function

$$x \mapsto \text{ess sup}_{u \in \Xi_{\mathcal{U}}} \mathcal{H}(t, \lambda, u, x, \hat{p}, \hat{q}), \quad x \in \xi_{\mathbb{R}} \quad (3.20)$$

is concave in x .

- For all $t \in [0, T]$,

$$\text{ess sup}_{v \in \Xi_{\mathcal{U}}} \mathcal{H}(t, \lambda, v, \hat{X}, \hat{p}, \hat{q}) = \mathcal{H}(t, \lambda, \hat{u}, \hat{X}, \hat{p}, \hat{q}). \quad (3.21)$$

Then \hat{u} is an optimal control for the problem (1.5).

Proof Once considering the filtration \mathbb{G} , the arguments in the proof of Theorem 3.7 leading to

$$J(u) - J(\hat{u}) \leq \mathbb{E} \left[\int_0^T \left(\mathcal{H}^u(t) - \mathcal{H}^{\hat{u}}(t) - \partial_x \mathcal{H}^{\hat{u}}(t) \left(X(t) - \hat{X}(t) \right) \right) dt \right] \leq 0$$

apply directly without conditioning. \square

4 Necessary maximum principles with time-change

Hereafter we study necessary conditions to identify the possible candidates for optimal controls. This can be a useful starting point before applying a verification theorem to ensure optimality. We remark that our results relax the condition of concavity present in Theorems 3.7 and 3.9. However, we introduce some other assumptions on the set of admissible controls and the first variation process of the forward dynamics (3.2).

In the literature we find a first version of necessary maximum principle for Volterra dynamics in Agram and Øksendal (2015). There the driving noises were the Gaussian and the centered Poisson random measure. Our work goes beyond these noises.

For any $t \in [0, T]$, we consider a random perturbation of the type

$$\beta(s) := \alpha_t \mathbb{1}_{[t, t+h]}(s), \quad s \in [0, T], \quad (4.1)$$

where α_t is a bounded \mathcal{F}_t measurable random variable and $h \in [0, T - t]$. We make the following assumptions:

1. The set of admissible controls $\mathcal{A}^{\mathbb{F}}$ is such that, for all $u \in \mathcal{A}^{\mathbb{F}}$,

$$u + \varepsilon\beta \in \mathcal{A}^{\mathbb{F}},$$

for all perturbations β as in (4.1) and all $\varepsilon > 0$ sufficiently small.

2. The first variation process $\chi(t), t \in [0, T]$, given by the derivative

$$\chi(t) := \partial_{\varepsilon} X^{(u+\varepsilon\beta)}|_{\varepsilon=0} \tag{4.2}$$

(see (3.2)) exists and is well defined.

3. $\partial_x b(t, s)$ and $\partial_u b(t, s)$ are well defined and C^1 with respect to t with partial derivatives L^2 -integrable with respect to $dt \times dP$. $\partial_x \kappa(t, s, \cdot)$ and $\partial_u \kappa(t, s, \cdot)$ are well defined and C^1 with respect to t with partial derivatives L^2 -integrable with respect to $d\Lambda \times dP$.

4. $\partial_x \kappa(t, s, z)$ and $\partial_u \kappa(t, s, z)$ are such that, for all $z \in \mathbb{R} \lambda \in [0, \infty)^2, u \in \mathcal{U}, x \in \mathbb{R}$, the partial derivative of $\partial_x \kappa + \partial_u \kappa$ with respect to t is locally bounded (uniformly in t) and satisfies

$$|\partial_t(\partial_x \kappa(t_1, s, z) + \partial_u \kappa(t_1, s, z)) - \partial_t(\partial_x \kappa(t_2, s, z) + \partial_u \kappa(t_2, s, z))| \leq K|t_1 - t_2|,$$

for some $K > 0$ and for each fixed $s \leq t, \lambda \in [0, \infty)^2, u \in \mathcal{U}, x \in \mathbb{R}$.

Assumption 2. above implies that

$$\begin{aligned} \chi(t) &= \int_0^t \left(\partial_x b(t, s)\chi(s) + \partial_u b(t, s)\beta(s) \right) ds \\ &+ \int_0^t \int_{\mathbb{R}} \left(\partial_x \kappa(t, s, z)\chi(s) + \partial_u \kappa(t, s, z)\beta(s) \right) \mu(dsdz), \end{aligned}$$

exists and is well defined, whereas assumption 4. ensure us to be able to apply the transformation rule for χ .

Remark that sufficient conditions that ensure the existence of the first variation process are that b and κ are in $C^1(\mathcal{U})$ uniformly for all $s, t \in [0, T] \lambda \in [0, \infty)^2, x \in \mathbb{R}$ and that $(\partial_x b(t, s)\chi(s) + \partial_u b(t, s)\beta(s))$ and $(\partial_x \kappa(t, s, z)\chi(s) + \partial_u \kappa(t, s, z)\beta(s))$ satisfy the linear growth and Lipschitzianity conditions of Theorem 3.1.

As above we consider the performance functional (3.1) with the related conditions on F and G as in Sect. 3. We also continue using the compact notation there introduced, see Notation 1.

Theorem 4.1 (Necessary maximum principle with respect to \mathbb{F}). *Let $\lambda \in \mathcal{L}$. Suppose that $\hat{u} \in \mathcal{A}^{\mathbb{F}}$ and the corresponding solutions $\hat{X}, (\hat{p}, \hat{q})$ of (3.2) and (3.10) exist. Assume also that P -a.s. $F \in C^1(\mathcal{U})$ for all $t \in [0, T] \lambda \in [0, \infty)^2, x \in \mathbb{R}$. If, for all perturbations β as in (4.1), we have that*

$$\partial_{\varepsilon} J(\hat{u} + \varepsilon\beta)|_{\varepsilon=0} = 0, \tag{4.3}$$

then

$$\partial_u \mathcal{H}^{\mathbb{F}, \hat{u}}(t) = 0. \tag{4.4}$$

The converse also holds true.

Proof With (4.2), we consider for $u \in \mathcal{A}^{\mathbb{F}}$ and the perturbation (4.1),

$$\partial_{\varepsilon} J(u + \varepsilon\beta)|_{\varepsilon=0}$$

$$\begin{aligned}
&= \mathbb{E} \left[\int_0^T (\partial_x F(t, \lambda_t, u(t), X(t))) \chi(t) + \partial_u F(t, \lambda_t, u(t), X(t)) \beta(t) dt \right. \\
&\quad \left. + \partial_x G(X(T)) \chi(T) \right]. \tag{4.5}
\end{aligned}$$

By considering a suitable increasing family of stopping times converging to T as in Øksendal and Sulem (2014) Theorem 2.2, we may assume that all the local martingales appearing here are true martingales. From (3.12), the transformation rule (Lemma 3.2) and the Itô formula for the product, we find that

$$\begin{aligned}
\mathbb{E} [\partial_x G(X(T)) \chi(T)] &= \mathbb{E} [p(T) \chi(T)] \\
&= \mathbb{E} \left[\int_0^T p(t) (\partial_x b(t, t) \chi(t) + \partial_u b(t, t) \beta(t)) dt - \int_0^T \chi(t) \partial_x \mathcal{H}(t) dt \right. \\
&\quad \left. + \int_0^T p(t) \left(\int_0^t (\partial_t \partial_x b(t, s) \chi(s) + \partial_t \partial_u b(t, s) \beta(s)) ds \right) dt \right. \\
&\quad \left. + \int_0^T p(t) \left(\int_0^t \int_{\mathbb{R}} (\partial_t \partial_x \kappa(t, s, z) \chi(s) + \partial_t \partial_u \kappa(t, s, z) \beta(s)) \mu(ds dz) \right) dt \right. \\
&\quad \left. + \int_0^T \int_{\mathbb{R}} q(s, z) (\partial_x \kappa(t, t, z) \chi(t) + \partial_u \kappa(t, t, z) \beta(t)) \Lambda(dt dz) \right].
\end{aligned}$$

Now, recalling equality (3.18), and taking the conditional expectation under \mathcal{F}_t we get that

$$\begin{aligned}
&\mathbb{E} [p(T) \chi(T)] \\
&= \mathbb{E} \left[\int_0^T \left\{ \partial_x b(t, t) \mathbb{E} [p(t) | \mathcal{F}_t] + \int_0^t \partial_x \partial_t b(t, s) ds \mathbb{E} [p(t) | \mathcal{F}_t] \right. \right. \\
&\quad \left. \left. + \int_0^t \int_{\mathbb{R}} \partial_x \partial_t \kappa(t, s, z) \mathbb{E} [\mathcal{D}_{s,z} p(t) | \mathcal{F}_t] \Lambda(ds dz) \right\} \chi(t) dt \right. \\
&\quad \left. + \int_0^T \left\{ \partial_u b(t, t) \mathbb{E} [p(t) | \mathcal{F}_t] + \int_0^t \partial_u \partial_t b(t, s) ds \mathbb{E} [p(t) | \mathcal{F}_t] \right. \right. \\
&\quad \left. \left. + \int_0^t \int_{\mathbb{R}} \partial_u \partial_t \kappa(t, s, z) \mathbb{E} [\mathcal{D}_{s,z} p(t) | \mathcal{F}_t] \Lambda(ds dz) \right\} \beta(t) dt - \int_0^T \partial_x \mathcal{H}^{\mathbb{F}, u}(t) \chi(t) dt \right. \\
&\quad \left. + \int_0^T \int_{\mathbb{R}} (\partial_x \kappa(t, t, z) \chi(t) + \partial_u \kappa(t, t, z) \beta(t)) \mathbb{E} [q(t, z) | \mathcal{F}_t] \Lambda(dt dz) \right].
\end{aligned}$$

So that, from (3.12), we can write

$$\begin{aligned}
&\mathbb{E} \left[\int_0^T (\partial_x F(t, \lambda_t, u(t), X(t))) \chi(t) + \partial_u F(t, \lambda_t, u(t), X(t)) \beta(t) dt + \partial_x G(X(T)) \chi(T) \right] \\
&= \mathbb{E} \left[\int_0^T \partial_x \mathcal{H}^{\mathbb{F}, u}(t) \chi(t) dt - \int_0^T \partial_x \mathcal{H}^{\mathbb{F}, u}(t) \chi(t) dt + \int_0^T \partial_u \mathcal{H}^{\mathbb{F}, u}(t) \beta(t) dt \right]. \tag{4.6}
\end{aligned}$$

Summarizing, equation (4.5) together with (4.6) and the perturbations in (4.1) give

$$\partial_\varepsilon J(u + \varepsilon \beta)|_{\varepsilon=0} = \mathbb{E} \left[\int_0^T \partial_u \mathcal{H}^{\mathbb{F}, u}(t) \beta(t) dt \right] = \mathbb{E} \left[\int_t^{t+h} \partial_u \mathcal{H}^{\mathbb{F}, u}(s) ds \alpha_t \right], \tag{4.7}$$

and, for \hat{u} , (4.3) gives

$$\partial_\varepsilon J(\hat{u} + \varepsilon \beta)|_{\varepsilon=0} = 0.$$

Applying the Fubini theorem to the right-hand side of (4.7) and differentiating at $h = 0$ we obtain

$$\mathbb{E} \left[\partial_u \mathcal{H}^{\mathbb{F}, \hat{u}}(t) \alpha_t \right] = 0,$$

for all α_t bounded and \mathcal{F}_t measurable. Hence

$$\mathbb{E} \left[\partial_u \mathcal{H}^{\mathbb{F}, \hat{u}}(t) \middle| \mathcal{F}_t \right] = \partial_u \mathcal{H}^{\mathbb{F}, \hat{u}}(t) = 0. \tag{4.8}$$

Vice versa, if (4.8) holds, we can reverse the argument to obtain (4.3). □

As in Sect. 3, for the sake of completeness, we propose a necessary maximum principle under the information flow \mathbb{G} . This refers to the optimization problem (1.5). In this case we assume that, for all $u \in \mathcal{A}^{\mathbb{G}}, u + \varepsilon\beta \in \mathcal{A}^{\mathbb{G}}$ for all perturbations β as in (4.1) and $\varepsilon > 0$ sufficiently small.

Proposition 4.2 (Necessary maximum principle with respect to \mathbb{G}). *Let $\lambda \in \mathcal{L}$. Suppose that $\hat{u} \in \mathcal{A}^{\mathbb{G}}$ and the corresponding solutions $\hat{X}, (\hat{p}, \hat{q})$ of (3.2) and (3.10) exist. Also assume that $F \in C^1(\mathcal{U})$ for all $t \in [0, T] \lambda \in [0, \infty)^2, x \in \mathbb{R}$. If, for all perturbations β ,*

$$\partial_\varepsilon J(\hat{u} + \varepsilon\beta)|_{\varepsilon=0} = 0, \tag{4.9}$$

then

$$\partial_u \mathcal{H}^{\hat{u}}(t) = 0. \tag{4.10}$$

Conversely, if (4.10) holds, then (4.9) is true.

Proof The argument in the proof of Theorem 4.2 leading to

$$\partial_\varepsilon J(\hat{u} + \varepsilon\beta)|_{\varepsilon=0} = \mathbb{E} \left[\int_0^T \partial_u \mathcal{H}^{\hat{u}}(t) \beta(t) dt \right]$$

still holds with no need to use conditional expectations. Since \hat{u} and $\hat{u} + \varepsilon\beta$ are \mathbb{G} -predictable, we obtain

$$\begin{aligned} & \mathbb{E} \left[\int_0^T \left\{ \partial_u \hat{b}(t, t) \hat{p}(t) + \int_0^t \partial_u \partial_t \hat{b}(t, s) ds \hat{p}(t) + \int_0^t \int_{\mathbb{R}} \partial_u \partial_t \hat{\kappa}(t, s, z) \mathcal{D}_{s, z} \hat{p}(t) \Lambda(ds dz) \right\} \beta(t) dt \right. \\ & \left. + \int_0^T \int_{\mathbb{R}} (\partial_u \hat{\kappa}(t, t, z) \beta(t)) \hat{q}(t, z) \Lambda(dt dz) \right] = \mathbb{E} \left[\int_0^T \partial_u \mathcal{H}^{\hat{u}}(t) \beta(t) dt \right], \end{aligned}$$

where we have used the definition of \mathcal{H} as in (3.9). We conclude as in Theorem 4.1. □

5 A maximum principle approach in optimal harvesting

We now go back to the optimal harvesting problem within fishery, where the population dynamics is given by the dynamics (1.1). We recall that our starting point are Gushing (1979), Brauer (1976), Gripenberg et al. (1990), where the authors consider deterministic Volterra models to model population growth and, following e.g. Brites and Braumann (2017), Aòvarez and Shepp (1998), we introduce some random fluctuations that will affect the population growth. Hence, the dynamics considered are of type (1.1):

$$X^u(t) = X_0 + \int_0^t (r(t, s) - Ku(s)) X^u(s) ds + \int_0^t \sigma(s) X^u(s) dB(s), \quad t \in [0, T], \tag{5.1}$$

where $r(t, s) : [0, T]^2 \rightarrow \mathbb{R}$, $\sigma(s) : [0, T] \rightarrow \mathbb{R}$, $K > 0$, $X_0 > 0$. Here, B is the conditional Gaussian measure. We assume that (5.1) admits a solution, that $r(t, s)$ is C^2 with respect to both t and s , and that σ is C^1 with respect to t and $\sigma(t) > -1$ for all $s \in [0, T]$, $z \in \mathbb{R}$. Lastly we assume $r(t, s)$, $\partial_t r(t, s)$ and $\sigma(t)$ are in $L^2(dt)$. For sufficient conditions that guarantee the existence of a solution of X we refer to Theorem 3.1. In the context of optimal harvesting of fish, r represents the growth rate, K the catchability coefficient, and the control u is the fishing effort.

Let us define

$$\tau := \inf\{t \in [0, T], \text{ such that } X^u(t) = 0\} \wedge T.$$

Then we can see that $X^u(t) = 0$ for all $t \geq \tau$. In fact, for $0 \leq \tau \leq t \leq T$, we have that (1.1) can be rewritten as

$$\begin{aligned} X(t) &= X_0 + \int_0^\tau (r(\tau, s) - Ku(s))X(s)ds + \int_0^\tau \sigma(s)X(s)dB(s) \\ &\quad + \int_\tau^t (r(t, s) - Ku(s))X(s)ds + \int_\tau^t \sigma(s)X(s)dB(s) + \int_0^\tau (r(t, s) - r(\tau, s))X(s)ds \\ &= X(\tau) + \int_\tau^t (r(t, s) - Ku(s))X(s)ds + \int_\tau^t \sigma(s)X(s)dB(s) \\ &\quad + \int_0^\tau (r(t, s) - r(\tau, s))X(s)ds, \end{aligned}$$

Being $X(0) = X_0 > 0$ and the process X continuous, we have that X is strictly positive, up to restricting ourselves to the interval $[0, \tau]$.

Our goal is to characterise the optimal solution to maximization of the performance functional

$$J(u) = \mathbb{E} \left[\int_0^T e^{-\delta(T-t)} X(t)u(t)dt \right], \quad (5.2)$$

where $u \in \mathcal{A}^{\mathbb{F}}$, $\delta > 0$. In the context of optimal harvesting this can be regarded as the aggregated net discounted revenue, see Belyakov and Veliov (2016). Following the approach given in this work, we consider the Hamiltonian functional (3.9), which can be here rewritten as

$$\begin{aligned} \mathcal{H}^u(t) &= e^{-\delta(T-t)}u(t)X(t) + (r(t, t) - Ku(t))X(t)p(t) + \sigma(t)X(t)q(t)\lambda_t^B \\ &\quad + \int_0^t \partial_t r(t, s)X(s) - Ku(s)ds p(t), \end{aligned}$$

where the backward dynamics for p are given by

$$\begin{aligned} dp(t) &= e^{-\delta(T-t)}u(t)dt + \left(r(t, t) + \int_0^t \partial_t r(t, s)ds \right) p(t)dt + \sigma(t)q(t)\lambda_t^B dt + q(t)dB(t) \\ p(T) &= 0. \end{aligned} \quad (5.3)$$

Also, we consider the mapping $\mathcal{H}^{\mathbb{F}}$ in (3.12):

$$\begin{aligned} \mathcal{H}^{\mathbb{F}, u}(t) &= e^{-\delta(T-t)}u(t)X(t) + (r(t, t) - Ku(t))X(t)\mathbb{E}[p(t)|\mathcal{F}_t] \\ &\quad + \sigma(t)X(t)\mathbb{E}[q(t)|\mathcal{F}_t]\lambda_t^B + \int_0^t \partial_t r(t, s)X(s) - Ku(s)ds \mathbb{E}[p(t)|\mathcal{F}_t]. \end{aligned}$$

From Theorem 4.1 we see that a necessary condition for an admissible control \hat{u} to be optimal is that, for all $t \in [0, T]$, $\partial_u \mathcal{H}^{\mathbb{F}, \hat{u}}(t) = 0$. Furthermore, from Theorem 3.7, being the map (3.13) trivially concave, the condition $\partial_u \mathcal{H}^{\mathbb{F}, \hat{u}}(t) = 0$ is also sufficient for the maximality. In particular this means that an admissible control \hat{u} is optimal if and only if

$$e^{-\delta(T-t)} \hat{X}(t) = K \hat{X}(t) \mathbb{E}[\hat{p}(t) | \mathcal{F}_t]. \tag{5.4}$$

Namely, for all $t \in [0, \tau]$

$$\mathbb{E}[\hat{p}(t) | \mathcal{F}_t] = K^{-1} e^{-\delta(T-t)}. \tag{5.5}$$

To find a solution to (5.3) with respect to the information flow \mathbb{G} , we use a Girsanov change of measure as presented in Di Nunno and Karlsen (2015). Define the measure \mathbb{Q} by $d\mathbb{Q} = \mathcal{M}(T) dP(T)$ on \mathcal{G}_T , where

$$\begin{aligned} d\mathcal{M}(t) &= \mathcal{M}(t) \sigma(t) dB(t) \\ \mathcal{M}(0) &= 1. \end{aligned} \tag{5.6}$$

An explicit solution for (5.6) is obtained by the Itô formula [see Di Nunno and Karlsen (2015)] and is given by

$$\mathcal{M}(t) = \exp \left\{ \int_0^t \sigma(s) dB(s) - \int_0^t \frac{1}{2} \sigma(s)^2 \lambda_s^B ds \right\}, \quad t \in [0, T].$$

We thus have that, under the measure \mathbb{Q} ,

$$dB^\sigma(t) = dB(t) - \sigma(t) \lambda_t^B dt,$$

is a \mathbb{G} -martingale. Equation (5.3) can now be rewritten under \mathbb{Q} as

$$\begin{aligned} d\hat{p}(t) &= e^{-\delta(T-t)} \hat{u}(t) dt + \left(r(t, t) + \int_0^t \partial_t r(t, s) ds \right) \hat{p}(t) dt + \hat{q}(t) dB^\sigma(t) \\ \hat{p}(T) &= 0, \end{aligned} \tag{5.7}$$

Thanks to Di Nunno and Sjursen (2014) we know that (5.7) admits a unique solution (\hat{p}, \hat{q}) and that the process \hat{p} is given by

$$\hat{p}(t) = \mathbb{E}_{\mathbb{Q}} \left[\int_t^T \exp \left\{ \int_t^s \tilde{r}(v) dv \right\} e^{-\delta(T-s)} \hat{u}(s) ds \right],$$

where we defined $\tilde{r}(t) := r(t, t) + \int_0^t \partial_t r(t, s) ds$. We thus obtain that

$$\mathbb{E} \left[\hat{p}(t) | \mathcal{F}_t \right] = \mathbb{E} \left[\frac{1}{\mathcal{M}(T)} \int_t^T \exp \left\{ \int_t^s \tilde{r}(v) dv \right\} e^{-\delta(T-s)} \hat{u}(s) ds \middle| \mathcal{F}_t \right]. \tag{5.8}$$

Substituting (5.8) in (5.5) we obtain a characterization of $\hat{u}(t)$.

Funding Open access funding provided by University of Oslo (incl Oslo University Hospital) The research leading to these results received funding from the Research Council of Norway within the project STORM: Stochastics for Time-Space Risk Models, Grant number: 274410.

Declarations

Conflicts of interest The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agram, N., Øksendal, B., & Yakhlef, S. (2018). Optimal control of forward-backward stochastic Volterra equations. In F. Gesztesy et al., (Ed.), *Non-linear partial differential equations, mathematical physics, and stochastic analysis*, (pp. 3–36). European Mathematical Society Publishing House.
- Agram, N., & Øksendal, B. (2015). Malliavin calculus and optimal control of stochastic Volterra equations. *J. Optim. Theory Appl.*, *167*, 1070–1094.
- Agram, N., Øksendal, B., & Yakhlef, S. (2019). New approach to optimal control of stochastic Volterra integral equations. *Stochastics*, *91*, 873–894.
- Aóvarez, L. H. R., & Shepp, L. A. (1998). Optimal harvesting of stochastically fluctuating populations. *Journal of Mathematical Biology*, *37*.
- Baños, D. R., Cordoni, F., Di Nunno, G., Di Persio, L., & Røse, E. E. (2019). Stochastic systems with memory and jumps. *Journal of Differential Equations*, *266*.
- Barndorff-Nielsen, A., & Shiryaev, O. E. (2010). *Change of time and change of measure*. World Scientific.
- Belyakov, A. O., & Veliov, V. M. (2016). On optimal harvesting in age-structured populations. *Dynamic Perspectives on Managerial Decision Making*, *22*, 149–166.
- Bonaccorsi, S., & Confortola, F. (2020). Optimal control for stochastic Volterra equations with multiplicative Lévy noise. *Nonlinear differential equations and applications*, 1–26.
- Borovkova, S., & Schmeck, M. D. (2017). Electricity price modeling with stochastic time change. *Energy Economics*, *63*, 51–65.
- Brauer, F. (1976). Constant rate harvesting of populations governed by volterra integral equations. *Journal of Mathematical Analysis and Applications*, *56*, 18–27.
- Brites, N. M., & Braumann, C. A. (2017). Fisheries management in random environments: comparison of harvesting policies for the logistic mode. *Fisheries Research*, *195*.
- Carr, P., & Wu, L. (2004). Time-changed Levy processes and option pricing. *Journal of Financial Economics*, *71*, 113–141.
- Cont, R., & Kalinin, A. (2020). On the support of solutions to stochastic differential equations with path-dependent coefficients. *Stochastic Processes and their Applications*, *130*.
- Dahl, K., Mohammed, S.-E.A., Øksendal, B., & Røse, E. E. (2016). Optimal control of systems with noisy memory and BSDEs with Malliavin derivatives. *Journal of Functional Analysis*, *271*.
- Di Nunno, G. & Karlsen, E. H. (2015). Hedging under worst-case-scenario in a market driven by time-changed lévy noises. In *The fascination of probability, statistics and their applications*. Springer.
- Di Nunno, G. & Sjursen, S. (2013). On chaos representation and orthogonal polynomials for the doubly stochastic Poisson process. In R. C. Dalang, M. Dozzi, and F. Russo, (Eds.), *Seminar on stochastic analysis, random fields and applications VII*, (pp. 23–54). Springer Basel.
- Di Nunno, G., Øksendal, B., & Proske, F. (2008). *Malliavin calculus for Lévy processes with applications to finance*. Universitext. Springer-Verlag.
- Di Nunno, G. (2002). Stochastic integral representations, stochastic derivatives and minimal variance hedging. *Stochastics and Stochastic Reports*, *73*(1–2), 181–198.
- Di Nunno, G., & Eide, I. B. (2009). Minimal-variance hedging in large financial markets: random fields approach. *Stochastic Analysis and Applications*, *28*, 54–85.
- Di Nunno, G., & Sjursen, S. (2014). BSDEs driven by time-changed Lévy noises and optimal control. *Stochastic Processes and Their Applications*, *124*, 1679–1709.
- Grigelionis, B. (1975). Characterization of stochastic processes with conditionally independent increments. *Lithuanian Mathematical Journal*, *15*, 562–567.
- Gripenberg, G., Londen, S. O., & Staffans, O. (1990). *Volterra integral and functional equations. Encyclopedia of mathematics and its applications*. Cambridge University Press.
- Gushing, J. M. (1979). Volterra integrodifferential equations in population dynamics. In Iannelli, M. (Ed.), *Mathematics of biology, C.I.M.E. Summer Schools*, (Vol. 80). Springer Berlin Heidelberg.

- Jacod, Jean. (1979). *Calcul Stochastique et Problèmes de Martingales. Lecture Notes in Mathematics* (Vol. 714). Springer-Verlag.
- Jokarab, M., Subbeyba, S., & Gjørseterb, H. (2021). A logistic function to track time-dependent fish population dynamics. *Fisheries Research*, 236.
- Kromer, E., Overbeck, L., & Röder, J. A. L. (2015). Feynman-Kac for functional jump diffusions with an application to credit value adjustment. *Statistics and Probability Letters*, 105.
- Madan Geman, H., & Yor, M. B. D. (2001). Time changes for Lévy processes. *Mathematical Finance*, 11, 79–96.
- Mai, J.-F., & Scherer, M. (2009). A tractable multivariate default model based on a stochastic time-change. *International Journal of Theoretical and Applied Finance*, 12(02), 227–249.
- Menon, G. (2017). A quick introduction to kinetic theory.
- Øksendal, B., & Sulem, A. (2014). Risk minimization in financial markets modeled by Ito-Lévy processes. *Afrika Matematika*, 26, 939–979.
- Øksendal, B., & Zhang, T. (2010). Optimal control with partial information for stochastic Volterra equations. *International Journal of Stochastic Analysis*, 1–25, 2010.
- Possamai, D., Touzi, N., & Zhang, J. (2020). Zero-sum path-dependent stochastic differential games in weak formulation. *Annales of applied probability*, 30.
- Protter, P. (1993). Volterra equations driven by semimartingales. *Annals of Probability*, 13, 519–530.
- Saakvitne, J. A. (2019). *Essays on market microstructure*. PhD Thesis, BI Norwegian Business School, Oslo, Norway.
- Serfozo, R. (1972). Processes with conditional stationary independent increments. *Journal of Applied Probability*, 3, 303–315.
- Shiryayev, A. N. (1999). *Essentials of stochastic finance*. World Scientific Publishing Co.
- Swishchuk, A. (2016). *Change of time methods in quantitative finance*. Springer.
- Veraart, A. E. D., & Winkel, M. (2010). Time change. In *Encyclopedia of quantitative finance* (Vol. 4, pp. 1812–1816). Wiley Hoboken.
- Yablonsky, A. L. (2005). The Malliavin calculus for processes with conditionally independent increments. In F. E. Benth, G. Di Nunno, T. Lindstrøm, B. Øksendal, & T. Zhang (Eds.), *Stochastic analysis and applications. The Abel symposium 2005* (pp. 641–678). Springer.
- Yong, J., & Zhou, X. Y. (1999). *Stochastic controls - Hamiltonian systems and HJB equations*. Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.