



On stochastic control for time changed Lévy dynamics

Giulia di Nunno^{1,2}

Received: 11 April 2022 / Accepted: 4 May 2022 / Published online: 29 June 2022
© The Author(s) 2022

Abstract

Controlled stochastic differential equations driven by time changed Lévy noises do not enjoy the Markov property in general, but can be treated in the framework of general martingales. From the modelling point of view, time changed noises constitute a feasible way to include time dependencies at noise level and still keep a reasonably simple structure. Furthermore, they are easy to simulate, with the result that time change Lévy dynamics attract attention in various fields of application. In this work we survey an approach to stochastic control via maximum principle for time changed Lévy dynamics. We emphasise the role and use of different information flows in tackling the various control problems. We show how these techniques can be extended to include Volterra type dynamics and the control of forward–backward systems of equations. Our techniques make use of the stochastic non-anticipating (NA) derivative in a general martingale framework.

Keywords Time change · Volterra equations · Stochastic backward differential equations · Forward–backward systems · Optimal control · Partial information · Maximum principle · Stochastic derivative

Mathematics Subject Classification 49N70 · 91A30 · 93E20 · 60H07 · 60H20 · 60G60

1 Introduction

From the modelling point of view, the main idea of time changing a stochastic process is to represent a stochastic function $X = X_t, t \geq 0$, by a process $L = L_t, t \geq 0$, with well-known structure serving as a *base*, and a (possibly stochastic) perturbation $\Lambda = \Lambda_t, t \geq 0$, of the time line

$$X_t = L_{\Lambda_t}.$$

This is a way to change the speed of motion along the trajectories of L , like treating the time line as an elastic band. This modelling potential as well as the easiness in simulation have

✉ Giulia di Nunno
giulian@math.uio.no

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

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

fascinated many researchers. See, e.g., [10, 48]. The concept of time perturbation is formally defined as follows.

Definition 1 A general *time change process* $\Lambda = \Lambda_t, t \geq 0$, is a stochastically continuous, non-decreasing càdlàg stochastic process starting at 0.

In our work the time changed processes are used as models for the *noise* driving the dynamics. As we shall see such noises can be embedded in the framework of general martingales when dealing with the calculus perspective. However, they also benefit of more explicit statistical properties that constitute a good trade-off between generality and specification, which is in itself interesting from the modelling perspective. Indeed we have examples both from the point of view of stochastic volatility in finance (see, e.g., [8, 14]) and turbulence in physics (see, e.g., [7, 11]). Using time change noises, one can have both Markovian and non-Markovian structures and also include *clustering effects* as discussed in [47, Chapter IV, 3e] and confirmed in the study [44, Chapter 3] on market microstructure.

The potential of time change in modelling raises from the pivotal results of Dambis [17] and Dubins and Schwarz [30], in which it is clarified that, *for any continuous local martingale* $M = M_t, t \geq 0$ with $M_0 = 0$ and $\langle M \rangle_\infty = +\infty$, *there exists a Brownian motion* $W = W_t, t \geq 0$, *such that* $M_t = W_{\langle M \rangle_t}$, *for every* $t \geq 0$. As a particular case we can consider martingales arising from classical Itô stochastic integrals

$$M_t = \int_0^t \sigma_s dw_s, \quad t \geq 0,$$

with respect to a Brownian motion w . Then, the result above leads to the representation

$$M_t = W_{\Lambda_t}, \quad \Lambda_t = \langle M \rangle_t = \int_0^t \sigma_s^2 ds \quad (1)$$

where W is another Brownian motion. These considerations reflect the *scaling property* of Brownian motion, namely, $cW_t \stackrel{d}{=} W_{c^2t}$ ($c > 0$). In this respect, it is also interesting to remark that, among all Lévy processes, besides Brownian motion, only α -stable processes share a property similar to (1). Indeed Rosinski and Woyczynski [43, Theorem 3.1] proved that *stochastic integrals* $X_t = \int_0^t \sigma_s d\mathcal{L}_s$, *with respect to the* α -*stable Lévy process* \mathcal{L} ($\alpha \in (0, 2]$) *can be represented as*

$$X_t = L_{\Lambda_t}, \quad \Lambda_t = \int_0^t \sigma_s^\alpha ds \quad (2)$$

where L is another α -stable Lévy process. Here again we observe that α -stable processes enjoy the *scaling property* $\sigma L_t \stackrel{d}{=} L_{\sigma^\alpha t}$. Representations (1) and (2) triggered the use of time change in stochastic volatility modelling, which started with the works [5, 8, 13, 14] and it is still of large interest nowadays. See, e.g., [11, 15, 34, 48].

For a given stochastic process X , the problem of the very existence of a representation in the form $X_t \stackrel{d}{=} W_{\Lambda_t}$ in terms of a Brownian motion and a family of stopping times $\Lambda_t, t \geq 0$, is the so-called *Skorokhod embedding problem*, which has driven researchers to discuss different solutions and to provide extensions beyond W as base, eventually to include also jumps. For a survey on the subject we can refer to [38].

On the other side, the problem of determining the time change Λ given the observations of X and fixed the base process (as a Brownian motion W , or even a Lévy process L) is known as the *recovery problem* and it was first studied in [50], see also [45].

From the probabilistic perspective, two classes of time change processes have turned to be successful: the *subordinators* (i.e. non-decreasing Lévy processes) and the *absolutely continuous time change processes*, characterised by

$$\Lambda_t = \int_0^t \lambda_s ds,$$

with λ being a non-negative integrable process representing the *time change rate*. While with the case of subordinators, the time changed Lévy process $X := L_\Lambda$ is still a Lévy process, with the case of absolutely continuous time changes (with non-constant rate), the process X exits the Lévy family and even the Markov family, giving in this way chances of describing forms of time dependences within the dynamics. Time changed Lévy processes with absolutely continuous time change process are widely used in applications to finance and beyond. As illustration we can propose the representation of affine processes in terms of these time changed processes, see [36]. We recall that affine processes are popular in the modelling, e.g., of interest rate and energy futures.

In the present work, we consider time change processes of the absolutely continuous type. Those noises are the stochastic driving forces of the controlled dynamics we study. The goal of this article is to discuss some techniques of stochastic optimisation for such dynamics with emphasis on the role of information and the probabilistic structure of the noise. In Sect. 2 we give a description of the framework, while in the rest of the paper we summarise an approach via maximum principle to stochastic control, providing sufficient conditions to verify optimality. This is based on [23, 24, 29]. In Sect. 3 we lay out the ideas and justify the technical use of different information flows. We achieve the result exploiting backward stochastic differential equations (BSDEs) driven by time changed Lévy noises. In Sect. 4 we see how to adapt these ideas to the case of Volterra type controlled dynamics. For this, a substantially modified Hamiltonian functional has to be used together with the NA-derivative [20, 21]. The role of stochastic differentiation will be discussed. In Sect. 5 we extend the approach to treat forward–backward systems of equations.

As a side note, we remark that there are extensions of the concept of time change to infinite dimensions such as Hilbert space valued time changes, which first appeared in [37], then also used in, e.g., [4] and recently connected to infinite dimensional stochastic volatility models as, e.g., in [12]. Also further extensions are explored in the direction of the so-called *meta time changes*, which are general forms of distortions of the space-time, in the sense that not only time is perturbed, but also the spacial component. These are substantially non-negative random measures, so far, extending the concept of subordination. See [9].

2 The noise: time-space random fields

The stochastic elements refer to the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The noise is represented as a random field on the time-space, which is here given by the field $\mathbb{X} := [0, T] \times \mathbb{R}$ ($T < \infty$), where the spacial component represents the jump size in the paths. We can define two disjoint complementary sets

$$\mathfrak{X}_B := [0, T] \times \{0\}, \quad \mathfrak{X}_H := [0, T] \times \mathbb{R}_0 \quad (\mathbb{R}_0 := \mathbb{R} \setminus \{0\}),$$

where the first set accommodates the random field of continuous nature (no jumps corresponds to having jump size null) and the second accommodates the random fields with jumps. We remark that our work can deal with further spacial components, for example of the type $\mathbb{X} := [0, T] \times \mathbb{R} \times \mathbb{R}^d$. In this case the extra spacial component may assume different

meaning: in physics, this could be the actual space, in finance, this can represent the space of available assets as, e.g., [18, 22].

The noise is modelled by the *random field* μ , which is defined as the mixture

$$\mu(\Delta, \omega) := B(\Delta \cap \mathfrak{X}_B, \omega) + \tilde{H}(\Delta \cap \mathfrak{X}_H, \omega), \quad \Delta \subseteq \mathbb{X}, \omega \in \Omega, \tag{3}$$

of two components on the Borel sets $\mathcal{B}_{\mathbb{X}}$ of \mathbb{X} described here below as *conditional Gaussian measure* B and *conditional centered Poisson measure* \tilde{H} .

To proceed with the formal definition, we first introduce yet another random measure, which will be directly connected with the concept of time change. We adopt the framework introduced in [29].

Definition 2 We define the random measure Λ on $\mathcal{B}_{\mathbb{X}}$ by

$$\begin{aligned} \Lambda(\Delta) &:= \Lambda^B(\Delta \cap \mathfrak{X}_B) + \Lambda^H(\Delta \cap \mathfrak{X}_H) \\ &= \int_0^T \mathbb{1}_{\{(t,0) \in \Delta\}}(t) \lambda_t^B dt + \int_0^T \int_{\mathbb{R}_0} \mathbb{1}_{\Delta}(t, z) \nu(dz) \lambda_t^H dt, \quad \Delta \subseteq \mathbb{X}, \end{aligned}$$

where the stochastic process $\lambda = (\lambda^B, \lambda^H)$ has non-negative, stochastically continuous components in $L^1(P \times dt)$. The set of these processes is denoted \mathcal{L} . The \mathbb{P} -augmented filtration generated by Λ is denoted $\mathbb{F}^\Lambda = \{\mathcal{F}_t^\Lambda, t \in [0, T]\}$. Set $\mathcal{F}^\Lambda := \mathcal{F}_T^\Lambda$.

Here above ν is a σ -finite measure on the Borel sets $\mathcal{B}_{\mathbb{R}_0}$ of \mathbb{R}_0 satisfying $\int_{\mathbb{R}_0} z^2 \nu(dz) < \infty$. The structure of the noise (3) is linked to the two following components.

Definition 3 The *conditional Gaussian measure* B is a random measure on $\mathcal{B}_{\mathfrak{X}_B}$ such that

- (A1) $\mathbb{P}(B(\Delta) \leq x \mid \mathcal{F}^\Lambda) = \mathbb{P}(B(\Delta) \leq x \mid \Lambda^B(\Delta)) = \Phi\left(\frac{x}{\sqrt{\Lambda^B(\Delta)}}\right)$, $x \in \mathbb{R}, \Delta \subseteq \mathfrak{X}_B$, where Φ is the standard normal cumulative distribution function.
- (A2) $B(\Delta_1)$ and $B(\Delta_2)$ are conditionally independent given \mathcal{F}^Λ whenever Δ_1 and Δ_2 are disjoint sets.

The *conditional Poisson measure* H is a random measure on $\mathcal{B}_{\mathfrak{X}_H}$ such that

- (A3) $\mathbb{P}(H(\Delta) = k \mid \mathcal{F}^\Lambda) = \mathbb{P}(H(\Delta) = k \mid \Lambda^H(\Delta)) = \frac{\Lambda^H(\Delta)^k}{k!} e^{-\Lambda^H(\Delta)}$, $k \in \mathbb{N}, \Delta \subseteq \mathfrak{X}_H$,
- (A4) $H(\Delta_1)$ and $H(\Delta_2)$ are conditionally independent given \mathcal{F}^Λ whenever Δ_1 and Δ_2 are disjoint sets.

Furthermore we assume that

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

We refer to [32] or [35] for the existence of conditional distributions in Definition 3. In (3) we use the *centred conditional Poisson measure*, which is the signed random measure \tilde{H} given by

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

It is easy to verify that the random measure μ in (3) is a *martingale random field*, see e.g. [22, Definition 2.1], with respect to both the following right-continuous information flows \mathbb{F} and \mathbb{G} . The first filtration is

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

and $\mathbb{F}^\mu := \{\mathcal{F}_t^\mu, t \in [0, T]\}$ is the \mathbb{P} -augmented filtration generated by the values $\mu(\Delta)$, $\Delta \subset [0, t] \times \mathbb{R}, t \in [0, T]$. The second filtration is

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

Notice that we have $\mathcal{G}_T = \mathcal{F}_T$ at the horizon T , while the initial information is different in the two flows: $\mathcal{G}_0 = \mathcal{F}^\Lambda$ and \mathcal{F}_0 is trivial. We can understand \mathbb{G} as an *enlargement* of the information \mathbb{F} by the future values of Λ , else we can regard \mathbb{F} as *partial* with respect to \mathbb{G} . From now on, we set $\mathcal{F} := \mathcal{F}_T = \mathcal{G}_T$.

The filtration \mathbb{G} has little real world application since it includes some information of the future values of Λ . However, it will be used technically in the sequel to exploit the distributional nature of the noise from the martingale representation theorem point of view. Indeed we recall that *given a square integrable martingale \mathcal{M} with respect to some reference filtration*

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

any square integrable $\mathcal{F}_T^\mathcal{M}$ -measurable random variable ξ admits representation

$$\xi = \xi^\perp + \int \varphi d\mathcal{M} \tag{4}$$

by means of a unique stochastic integrand φ . Here ξ^\perp is a stochastic remainder orthogonal to all the stochastic integrals with respect to \mathcal{M} (Kunita–Watanabe integral representation). It is well known that ξ^\perp is a constant whenever \mathcal{M} is a Gaussian or a centered Poisson random measure, or a mixture of the two, and the reference filtration $\mathbb{F}^\mathcal{M}$ is generated by \mathcal{M} , see [16, 19] and also [21, 27]. In addition we can show that

$$\xi^\perp = \mathbb{E}[\xi | \mathcal{F}^\Lambda]$$

in the case when the martingale \mathcal{M} is given by μ in (3) with respect to the filtration \mathbb{G} . See [29, Theorem 3.5]. We stress that this stochastic integral representations stand at the foundation of the solutions of BSDEs. Specifically, the following *explicit* stochastic representation theorem is considered in the sequel. See the original version in [20] and the versions for random fields appearing in [22, 28].

Theorem 1 *For any $\xi \in L^2(d\mathbb{P})$ we have*

$$\begin{aligned} \xi &= \mathbb{E}[\xi | \mathcal{F}^\Lambda] + \int_{\mathbb{X}} \mathfrak{D}_{t,z} \xi \mu(dt dz) \\ &= \mathbb{E}[\xi | \mathcal{F}^\Lambda] + \int_{\mathcal{X}_B} \mathfrak{D}_{t,0} \xi B(dt) + \int_{\mathcal{X}_H} \mathfrak{D}_{t,z} \xi \tilde{H}(dt dz), \end{aligned} \tag{5}$$

by means of the NA-derivative $\mathfrak{D}\xi$ which is defined as the $L^2(\Lambda \times d\mathbb{P})$ limit

$$\mathfrak{D}\xi = \lim_{n \rightarrow \infty} \varphi_n \tag{6}$$

of the simple predictable random fields $(\varphi_n)_{n \in \mathbb{N}}$ in the form

$$\varphi_n(t, z) := \sum_{k=1}^{K_n} \mathbb{E} \left[\xi \frac{\mu(\Delta_{nk})}{\mathbb{E}[\Lambda(\Delta_{nk}) | \mathcal{G}_{s_{nk}}]} \Big| \mathcal{G}_{s_{nk}} \right] \mathbf{1}_{\Delta_{nk}}(t, z), \quad (t, z) \in \mathbb{X}, \tag{7}$$

defined on some intersecting system $(\Delta_{nk})_{k=1, \dots, K_n; n \in \mathbb{N}}$ with $\Delta_{nk} = (s_{nk}, u_{nk}) \times B_{nk}$ and B_{nk} belonging to a countable semiring generating $\mathcal{B}_{\mathbb{R}}$.

From (6), (7) we see that $\mathfrak{D}\xi \in \mathcal{I}^{\mathbb{G}}$, which is the space of Itô integrands, i.e. \mathbb{G} -predictable processes in $L^2(d\Lambda \times d\mathbb{P})$. Furthermore, by (5) it is easy to see that *the NA-derivative is the dual of the Itô integral*. For later use, we recall that any $\varphi \in \mathcal{I}^{\mathbb{G}}$ can be regarded as a stochastic process with values in \mathcal{Z} , which is the space of functions $\phi : \mathbb{R} \rightarrow \mathbb{R}$ such that

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

Representation (5) is *explicit*, compared to (4), because it indicates, besides existence, the nature of the integrand. In fact formulae (6), (7) depend only on the random variable ξ to be represented and the framework $(\Omega, \mathbb{G}, \mathbb{P}, \mu)$. This type of stochastic integral representations is in line with the well-known *Clark–Ocone formula*, in which the integrand is given in terms of the projections of the Malliavin derivative on the information flow. We remark that the NA-derivative coincides with the Clark–Ocone formula whenever the integrator $\mathcal{M} = \mu$ is a Brownian motion or a centred Poisson random measure, or mixtures of the two, and when ξ is Malliavin differentiable. The domain of the Malliavin derivative is contained, but not equal, to $L^2(d\mathbb{P})$.

To conclude on this section we recall the following result [46, Theorem 3.1] (see also [32]), which explicitly connects the noise μ in (3) with the concept of time-change.

Theorem 2 *Let $W_t, t \in t \in [0, T]$ be a Brownian motion and $N_t, t \in t \in [0, T]$ be a centred pure jump Lévy process with Lévy measure ν . Assume that both W and N are independent of Λ . Then B satisfies (A1)–(A2) if and only if, for any t ,*

$$B_t \stackrel{d}{=} W_{\Lambda_t^B},$$

and $\eta_t := \int_0^t \int_{\mathbb{R}_0} z \tilde{H}(ds, dz), t \in [0, T]$, satisfies (A3)–(A4) if and only if, for any t ,

$$\eta_t \stackrel{d}{=} N_{\hat{\Lambda}_t^H}$$

with $\hat{\Lambda}_t^H := \int_0^t \lambda_s^H ds$, for $t \in [0, T]$.

3 A sufficient maximum principle

Hereafter we highlight the use of information to tackle optimal control problems for dynamics driven by time-changed Lévy noises. We remark immediately that these problems cannot be solved with dynamic programming methods since the state process is, in general, not Markovian. The approach we present here is via the *maximum principle*. We also make use of the enlarged filtration \mathbb{G} to exploit the structure of the noise. We refer to [29] for the original presentation of the main ideas.

We aim to find the optimal control \hat{u} providing the maximal value over all admissible controls $\mathcal{A}^{\mathbb{F}}$,

$$J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathbb{F}}} J(u), \tag{8}$$

where the performance functional is

$$J(u) = \mathbb{E} \left[\int_0^T F_t(\lambda_t, u_t, X_t) dt + G(X_T) \right]. \tag{9}$$

Here $G(x, \omega)$, $\omega \in \Omega$, is a concave random variable for all $x \in \mathbb{R}$, differentiable in x a.s. and $F_t(\lambda, u, x, \omega)$, $t \in [0, T]$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$, $x \in \mathbb{R}$, $\omega \in \Omega$ is a \mathbb{F} -adapted stochastic process differentiable in x a.s. Here $\mathcal{U} \subseteq \mathbb{R}$ is a closed convex set.

Problem (8), (9) is apparently standard in stochastic control, however, it is the state process X that presents elements of novelty given by the driving noise μ , which provides a non-Markovian framework in general. The state process $X = X_t, t \in [0, T]$, follows the dynamics

$$dX_t = b_t(\lambda_t, u_t, X_t) dt + \int_{\mathbb{R}} \kappa_t(z, \lambda_t, u_t, X_t) \mu(dt dz), \quad X_0 \in \mathbb{R}, \tag{10}$$

where $b_t(\lambda, u, x)$ and $\kappa_t(z, \lambda, u, x)$, $t \in [0, T]$, $\lambda \in [0, \infty)^2$, $z \in \mathbb{R}$, $u \in \mathcal{U}$, $x \in \mathbb{R}$ are \mathbb{F} -predictable stochastic processes differentiable in x a.s. All partial derivatives with respect to some variable x are denoted by ∂_x here and in the sequel. Sufficient conditions for the existence of an \mathbb{F} -adapted strong solution to (10) can be found in [33].

In line with the maximum principle (MP) approach, we introduce a *Hamiltonian functional* $\mathcal{H} : [0, T] \times [0, \infty)^2 \times \mathcal{U} \times \mathbb{R} \times \mathbb{R} \times \mathcal{Z} \times \Omega \rightarrow \mathbb{R}$ given by the stochastic function

$$\begin{aligned} \mathcal{H}_t(\lambda, u, x, p, q) &= F_t(\lambda, u, x) + b_t(\lambda, u, x)p + \kappa_t(0, \lambda, u, x)q(0)\lambda^B \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(z, \lambda, u, x)q(z)\lambda^H v(dz). \end{aligned} \tag{11}$$

Corresponding to an admissible pair (u, X) we have the couple $(p, q) \in L^2(dt \times d\mathbb{P}) \times \mathcal{I}^{\mathbb{G}}$, which is the solution to the BSDE

$$p_t = \partial_x G(X_T) + \int_t^T \partial_x \mathcal{H}_s(\lambda, u_s, X_s, p_s, q_s) ds - \int_t^T \int_{\mathbb{R}} q_s(z) \mu(ds, dz). \tag{12}$$

Here $\partial_x \mathcal{H}_t = \frac{\partial}{\partial x} \mathcal{H}_t(\lambda, u, x, p, q)$ and we note that \mathcal{H} is differentiable in x by the assumptions on F, b and κ . A solution to Eq. (12) is studied in [29, Theorem 4.5]. There sufficient conditions to ensure that the pair $(\partial_x \mathcal{H}, \partial_x G(X_T))$ are *standard parameters* (see [29, Definition 4.1]) are given. In our context these would read: There exists $K_1 > 0$

$$\left| \partial_x \kappa_t(0, \lambda_t, u_t, X_{t-}) \right| \sqrt{\lambda_t^B} \leq K_1 \quad dt \times d\mathbb{P}\text{-a.e.}, \tag{13}$$

$$\int_{\mathbb{R}_0} (\partial_x \kappa_t(z, \lambda_t, u_t, X_{t-}))^2 v(dz) \sqrt{\lambda_t^H} \leq K_1 \quad dt \times d\mathbb{P}\text{-a.e.}, \tag{14}$$

$$\left| \partial_x b_t(\lambda_t, u_t, X_{t-}) \right| \leq K_1 \quad dt \times d\mathbb{P}\text{-a.e.} \tag{15}$$

Remark 1 We have to be aware that to find a solution to the BSDE (12), we work under filtration \mathbb{G} as it is the representation theorem in the form (5) (see Theorem 1) that provides the existence of the solution as shown in [29, Theorem 4.5]. Note that if another filtration were to be used, then the BSDE would present an additional term (an orthogonal martingale) as intrinsic part of the solution. This would not be suitable to pursue a maximum principle type result (as it would involve the explicit use of the quadratic variation of the orthogonal martingale, which is not explicitly known).

Definition 4 The admissible controls in (8), (9) are \mathbb{F} -predictable stochastic processes $u : [0, T] \times \Omega \rightarrow \mathcal{U}$, such that the corresponding state process X in (10) has a unique strong

solution in $L^2(dt \times d\mathbb{P})$, the adjoint equation (12) has a unique solution in $L^2(dt \times d\mathbb{P}) \times \mathcal{I}^{\mathbb{G}}$, and

$$\mathbb{E} \left[\int_0^T |F_t(\lambda_t, u_t, X_{t-})|^2 dt + |G(X_T)| + |\partial_x G(X_T)|^2 \right] < \infty.$$

The couple (u, X) is called an admissible pair. The set of admissible controls is denoted by $\mathcal{A}^{\mathbb{F}}$.

Once evaluating the Hamiltonian functional (11) in the solution (p, q) to (12), we obtain a \mathbb{G} -adapted functional. Thus we cannot apply directly the maximum principle to \mathcal{H} to solve (8) (where actually solutions should be \mathbb{F} -adapted) as it could be done in the case of controlled jump-diffusions, see e.g. [39]. In our context, we need then to "project" the Hamiltonian onto the information flow \mathbb{F} regarded as "partial" information with respect to \mathbb{G} , here we are inspired by [25]. Then we have the following result.

Theorem 3 *Let $\lambda \in \mathcal{L}$ be fixed. Let $\hat{u} \in \mathcal{A}^{\mathbb{F}}$. Denote the corresponding state process as \hat{X} with solution (\hat{p}, \hat{q}) of the adjoint equation (12). Set*

$$\begin{aligned} \mathcal{H}_t^{\mathbb{F}}(\lambda_t, u, x, \hat{p}_t, \hat{q}_t) &:= \mathbb{E} [\mathcal{H}_t(\lambda_t, u, x, \hat{p}_t, \hat{q}_t) | \mathcal{F}_t] \\ &= F_t(\lambda_t, u, x) + b_t(\lambda_t, u, x) \mathbb{E} [\hat{p}_t | \mathcal{F}_t] + \kappa_t(0, \lambda_t, u, x) \mathbb{E} [\hat{q}_t(0) | \mathcal{F}_t] \lambda_t^B \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(z, \lambda_t, u, x) \mathbb{E} [q_t(z) | \mathcal{F}_t] \lambda_t^H \nu(dz) \end{aligned}$$

for all $t \in [0, T]$. If

$$h_t^{\mathbb{F}}(x) := \operatorname{ess\,sup}_{u \in \mathcal{U}} \mathcal{H}_t^{\mathbb{F}}(\lambda_t, u, x, \hat{p}_t, \hat{q}_t) \tag{16}$$

exists and is a concave function in x for all $t \in [0, T]$ and

$$\mathcal{H}_t^{\mathbb{F}}(\lambda_t, \hat{u}_t, \hat{X}_t, \hat{p}_t, \hat{q}_t) = h_t^{\mathbb{F}}(\hat{X}_t), \tag{17}$$

then (\hat{u}, \hat{X}) is an optimal pair for (8), (9).

In the sequel we set $\hat{b}_s = b_s(\lambda_s, \hat{u}_s, \hat{X}_s)$, etc. for the coefficients associated with the admissible pair (\hat{u}, \hat{X}) with solution (\hat{p}, \hat{q}) of the adjoint equation (12). Set $b_s = b_s(\lambda_s, u_s, X_s)$ and so forth for the coefficients associated to another arbitrary admissible pair (u, X) . In addition $\hat{\mathcal{H}}_s(u, x) = \mathcal{H}_s(\lambda_s, u, x, \hat{p}_s, \hat{q}_s)$.

In the proof we assume that the integrals are well defined. Some sufficient conditions can be given to ensure this. For instance we can assume (13)–(15) and that there exist a $K_2 > 0$ such that

$$|b_t(\lambda, u, x) - b_t(\lambda, u, x')| \leq K_2|x - x'| \tag{18}$$

$$|\kappa_t(0, \lambda, u, x) - \kappa_t(0, \lambda, u, x')| \leq K_2|x - x'| \tag{19}$$

$$\int_{\mathbb{R}_0} |\kappa_t(z, \lambda, u, x) - \kappa_t(z, \lambda, u, x')|^2 \nu(dz) \leq K_2|x - x'|^2 \tag{20}$$

\mathbb{P} -a.s., for all $t \in [0, T]$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$ and all $x, x' \in \mathbb{R}$.

Proof Observe that

$$J(\hat{u}) - J(u) = \mathbb{E} [G(\hat{X}_T) - G(X_T)] + \mathbb{E} \left[\int_0^T \{ \hat{F}_s - F_s \} ds \right] = I_1 + I_2.$$

Since $\hat{X}_0 - X_0 = 0$ and G is concave, we have

$$I_1 = \mathbb{E} \left[G(\hat{X}_T) - G(X_T) \right] \geq \mathbb{E} \left[\partial_x G(\hat{X}_T) (\hat{X}_T - X_T) \right] = \mathbb{E} \left[\hat{p}_T (\hat{X}_T - X_T) \right].$$

By the Itô's formula, we have

$$\begin{aligned} I_1 &\geq \mathbb{E} \left[\int_0^T \left\{ -(\hat{X}_s - X_s) \partial_x \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) + \hat{p}_s (\hat{b}_s - b_s) \right\} ds \right. \\ &\quad + \int_0^T \int_{\mathbb{R}} \left\{ \hat{p}_{s-} (\hat{\kappa}_s(z) - \kappa_s(z)) + (\hat{X}_{s-} - X_{s-}) \hat{q}_s(z) \right\} \mu(ds, dz) \\ &\quad + \int_0^T (\hat{\kappa}_s(0) - \kappa_s(0)) \hat{q}_s(0) \lambda_s^B ds + \int_0^T \int_{\mathbb{R}_0} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{q}_s(z) \nu(dz) \lambda_s^H ds \Big] \\ &= \mathbb{E} \left[\int_0^T \left\{ -(\hat{X}_s - X_s) \partial_x \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) + \hat{p}_{s-} (\hat{b}_s - b_s) \right\} ds \right. \\ &\quad \left. + \int_0^T \int_{\mathbb{R}} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{q}_s(z) \Lambda(ds dz) \right]. \end{aligned}$$

Furthermore, from the Hamiltonian functional (11), we have

$$\begin{aligned} I_2 &= \mathbb{E} \left[\int_0^T \{ \hat{F}_s - F_s \} ds \right] = \mathbb{E} \left[\int_0^T \left\{ \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s(u_s, X_s) - (\hat{b}_s - b_s) \hat{p}_s \right. \right. \\ &\quad \left. \left. - (\hat{\kappa}_s(0) - \kappa_s(0)) \hat{q}_s(0) \lambda_s^B - \int_{\mathbb{R}_0} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{q}_s(z) \nu(dz) \lambda_s^H \right\} ds \right] \\ &= \mathbb{E} \left[\int_0^T \left\{ \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s(u_s, X_s) - (\hat{b}_s - b_s) \hat{p}_s \right\} ds \right. \\ &\quad \left. - \int_0^T \int_{\mathbb{R}} (\hat{\kappa}_s(z) - \kappa_s(z)) \hat{q}_s(z) \Lambda(ds dz) \right]. \end{aligned}$$

Hence

$$\begin{aligned} &J(\hat{u}) - J(u) \\ &\geq \mathbb{E} \left[\int_0^T \mathbb{E} \left[\hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s(u_s, X_s) - (\hat{X}_s - X_s) \partial_x \hat{\mathcal{H}}_s(\hat{u}_s, \hat{X}_s) \mid \mathcal{F}_s \right] ds \right] \\ &= \mathbb{E} \left[\int_0^T \hat{F}_s - F_s + (\hat{b}_s - b_s) \mathbb{E}[\hat{p}_s \mid \mathcal{F}_s] + (\hat{X}_s - X_s) \partial_x \hat{F}_s + \partial_x \hat{b}_s \mathbb{E}[\hat{p}_s \mid \mathcal{F}_s] ds \right] \end{aligned}$$

$$\begin{aligned}
 & + \mathbb{E} \left[\int_0^T \int_{\mathbb{R}} \left\{ (\hat{\kappa}_s(z) - \kappa_s(z)) \mathbb{E}[\hat{q}_s(z) | \mathcal{F}_s] \right. \right. \\
 & \left. \left. + (\hat{X}_s - X_s) \partial_x \hat{\kappa}_s(z) \mathbb{E}[\hat{q}_s(z) | \mathcal{F}_s] \right\} \Lambda(dsdz) \right] \\
 & = \mathbb{E} \left[\int_0^T \hat{\mathcal{H}}_s^{\mathbb{F}}(\hat{u}_s, \hat{X}_s) - \hat{\mathcal{H}}_s^{\mathbb{F}}(u_s, X_s) - \partial_x \hat{\mathcal{H}}_s^{\mathbb{F}}(\hat{u}_s, \hat{X}_s)(\hat{X}_s - X_s) ds \right]. \tag{21}
 \end{aligned}$$

We can show that the integrand here above is non-negative $dt \times d\mathbb{P}$ -a.e. applying a separating hyperplane argument (see [42, Chapter 5, Section 23]) to the concave function (16) and exploiting the maximality (17). \square

4 MP for Volterra dynamics

In a fashion similar to Sect. 3 we consider the control problem

$$J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathbb{F}}} J(u), \tag{22}$$

associated to the performance functional

$$J(u) = \mathbb{E} \left[\int_0^T F_t(\lambda_t, u_t, X_t) dt + G(X_T) \right]. \tag{23}$$

cf. (8), (9). However, we now study a state process X with Volterra type dynamics:

$$X_t = X_0 + \int_0^t b_s(t, \lambda_s, u_s, X_s) ds + \int_0^t \int_{\mathbb{R}} \kappa_s(t, z, \lambda_s, u_s, X_{s-}) \mu(dsdz), \tag{24}$$

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 $b(t, \lambda, u, x, \cdot)$ and $\kappa(t, z, \lambda, u, x, \cdot)$ 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 that they are C^1 with respect to t and to x . For later use, in order to apply the transformation rule (see [41] Theorem 3.3), we 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 $\partial_t \kappa_s(t, z, \lambda, u, x)$) is locally bounded (uniformly in t) and satisfies

$$|\partial_t \kappa_s(t_1, z, \lambda, u, x) - \partial_t \kappa_s(t_2, z, \lambda, u, x)| \leq K_3 |t_1 - t_2|, \tag{25}$$

for some $K_3 > 0$ and for each fixed $s \leq t$, $\lambda \in [0, \infty)^2$, $u \in \mathcal{U}$, $x \in \mathbb{R}$. Sufficient conditions for the existence of a strong \mathbb{F} -adapted solution to (24) in $L^2(dt \times d\mathbb{P})$ are studied in [23, Theorem 3.2].

In the Volterra framework the dynamics produce a memory effect that, together with the noise μ , makes the state process X clearly non-Markovian. So, as in Sect. 3, we propose a maximum principle approach. However, we see that we have to modify the Hamiltonian functional to accommodate the Volterra features.

Hereafter, the general space Φ_S is the space of measurable functions on $[0, T]$ with values in S . We introduce the Hamiltonian functional :

$$\mathcal{H} : [0, T] \times \Phi_{[0,\infty)^2} \times \Phi_{\mathcal{U}} \times \Phi_{\mathbb{R}} \times \Phi_{L^2(d\mathbb{P})} \times \Phi_{\mathcal{Z}} \times \Omega \longrightarrow \mathbf{R},$$

as the mapping given by the sum

$$\mathcal{H}_t(\lambda, u, x, p, q) := H_0(t) + H_1(t) \tag{26}$$

of the two components

$$\begin{aligned} H_0(t) &:= F_t(\lambda_t, u_t, x_t) + b_t(t, \lambda_t, u_t, x_t)p_t + \kappa_t(t, 0, \lambda_t, u_t, x_t)q_t(0)\lambda_t^B \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(t, z, \lambda_t, u_t, x_t)q_t(z)\lambda_t^H v(dz) \\ H_1(t) &:= \int_0^t \partial_t b_s(t, \lambda_s, u_s, x_s)ds p_t + \int_0^t \partial_t \kappa_s(t, z, \lambda_s, u_s, x_s)\mathfrak{D}_{s,0} p_t \lambda_s^B ds \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \partial_t \kappa_s(t, z, \lambda_s, u_s, x_s)\mathfrak{D}_{s,z} p_t \lambda_s^H v(dz)ds. \end{aligned}$$

As we see here above we have used the NA-derivative (6), (7). Stochastic differentiation was first used in the maximum principle approach in [1], where the controlled dynamics were jump-diffusions driven by a standard Brownian motion and centred Poisson random measure. There the Malliavin derivative was used. We recall that Malliavin calculus is tailored on the type of noise and it does not extend easily to general martingales as integrators. See, e.g., [26]. We find an extension to the case of processes with conditional independent increment in [49]. However, the very use of Malliavin calculus has the drawback that the domains of the operators involved are strict subspaces of $L^2(d\mathbb{P})$ and it is difficult to verify whether the random variables of interests belong to such domains (particularly thinking of the dependence of the variables on different controls u). To circumvent this problem [2] proposes to use an extension of the Malliavin calculus to the white noise setting. See [26] and see [3] for the case of Volterra type dynamics. The white noise framework is a generalisation of Malliavin calculus to distribution spaces and it is again strongly depending on the nature of the noise. At present there is no extension of such framework to the case of time-change Lévy noises and this is topic of current research.

For those reasons, we use here the NA-derivative (6), (7) and its duality with the Itô integral, see Theorem 1. The NA-derivative is well-defined with respect to all square integrable martingales and its domain is the whole $L^2(d\mathbb{P})$. All our variables belong to such domain.

For $\lambda \in \mathcal{L}$ and for (u, X) in (24), we introduce the adjoint equation as the BSDE

$$p_t = \partial_x G(X_T) + \int_t^T \partial_x \mathcal{H}_s(\lambda, u, X, p, q)ds - \int_t^T \int_{\mathbb{R}} q_s(z)\mu(dsdz). \tag{27}$$

Observe that (27) is a true BSDE as the terms $\partial_x H_1$ in the driver $\partial_x \mathcal{H}$ are integrated (see [23, Remark 3.3]). Such equation is admitting \mathbb{G} -adapted solution following [29]. Hence we can implement the same form of "projection" on the filtration \mathbb{F} to solve the optimal control problem (22), which is studied over the following set of admissible controls.

Definition 5 The set $\mathcal{A}^{\mathbb{F}}$ of admissible controls in problem (22), (23) consists of \mathbb{F} -predictable stochastic processes $u : [0, T] \times \Omega \longmapsto \mathcal{U}$ such that X in (24) and (p, q) in (27) have a unique strong solution in $L^2(dt \times d\mathbb{P})$ and $L^2(dt \times d\mathbb{P}) \times \mathcal{T}^{\mathbb{G}}$, respectively, and

$$\mathbb{E} \left[\int_0^T |F_s(\lambda_s, u_s, X_s)|^2 ds + |G(X_T)| + |\partial_x G(X_T)|^2 \right] < \infty.$$

The couple (u, X) is an *admissible pair*.

Theorem 4 *Set $\lambda \in \mathcal{L}$. Let $\hat{u} \in \mathcal{A}^{\mathbb{F}}$ and assume that the corresponding solutions $\hat{X}, (\hat{p}, \hat{q})$ of (24) and (27) exist. Define the mapping $\mathcal{H}^{\mathbb{F}}$ as*

$$\mathcal{H}_t^{\mathbb{F}}(\lambda, u, x, \hat{p}, \hat{q}) := \mathbb{E} [\mathcal{H}_t(\lambda, u, x, \hat{p}, \hat{q}) | \mathcal{F}_t] := H_0^{\mathbb{F}}(t) + H_1^{\mathbb{F}}(t), \tag{28}$$

for $t \in [0, T], u \in \Phi_{\mathcal{U}}, x \in \Phi_{\mathbb{R}}$, where

$$\begin{aligned} H_0^{\mathbb{F}}(t) &:= F_t(\lambda_t, u_t, x_t) + b_t(t, \lambda_t, u_t, x_t)\mathbb{E}[\hat{p}_t | \mathcal{F}_t] + \kappa_t(t, 0, \lambda_t, u_t, x_t)\mathbb{E}[q_t(0) | \mathcal{F}_t] \lambda_t^B \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(t, z, \lambda_t, u_t, x_t)\mathbb{E}[\hat{q}_t(z) | \mathcal{F}_t] \lambda_t^H \nu(dz) \\ H_1^{\mathbb{F}}(t) &:= \int_0^t \partial_t b_s(t, \lambda_s, u_s, x_s) ds \mathbb{E}[\hat{p}_t | \mathcal{F}_t] + \int_0^t \partial_t \kappa_s(t, 0, \lambda_s, u_s, x_s)\mathbb{E}[\mathfrak{D}_{s,0} \hat{p}_t | \mathcal{F}_t] \lambda_s^B ds \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \partial_t \kappa_s(t, z, \lambda_s, u_s, x_s)\mathbb{E}[\mathfrak{D}_{s,z} \hat{p}_t | \mathcal{F}_t] \lambda_s^H \nu(dz) ds \end{aligned}$$

Assume that, for fixed t , the function

$$h_t^{\mathbb{F}}(x) := \text{ess sup}_{u \in \mathcal{U}} \mathcal{H}_t^{\mathbb{F}}(\lambda, u, x, \hat{p}, \hat{q}) \tag{29}$$

exists and is concave. Also assume that, for all $t \in [0, T]$,

$$h_t^{\mathbb{F}}(\hat{X}) = \mathcal{H}_t^{\mathbb{F}}(\lambda, \hat{u}, \hat{X}, \hat{p}, q). \tag{30}$$

Then \hat{u} is an optimal control and (\hat{u}, \hat{X}) is an optimal pair for problem (22).

To ease the reading, we introduce some notation in the same style as in Sect. 3. For a given $\hat{u} \in \mathcal{A}^{\mathbb{F}}$, \hat{X} represents the associated controlled dynamics. Also, $b_s(t) := b_s(t, \lambda_t, u_t, X_t)$, $\hat{b}_s(t) := b_s(t, \lambda_t, \hat{u}_t, \hat{X}_t)$, and similarly for $\kappa, \hat{\kappa}, F, \hat{F}, G, \hat{G}$. We will also write $\mathcal{H}_s^u := \mathcal{H}_s(\lambda_s, u, X_s, \hat{Y}_s, \hat{\phi}_s(\cdot))$, $\mathcal{H}_s^{\hat{u}} := \mathcal{H}_s(\lambda_s, \hat{u}_s, \hat{X}_s, \hat{Y}_s, \hat{\phi}_s(\cdot))$ and similarly for the other functionals in (26), (28).

Proof The arguments proceed in a fashion similar to those in Theorem 3, however the presence of the Volterra structure needs to be considered with care. Here below we present only those specific arguments treating the Volterra feature of the dynamics (24).

For $u \in \mathcal{A}^{\mathbb{F}}$ with corresponding controlled dynamics X , we consider $J(u) - J(\hat{u}) = I_1 + I_2$, where

$$\begin{aligned} I_1 &:= \mathbb{E} \left[\int_0^T F_t(\lambda_t, u_t, X_t) - F_t(\lambda_t, \hat{u}_t, \hat{X}_t) dt \right], \\ I_2 &:= \mathbb{E} \left[G(X_T) - G(\hat{X}_T) \right]. \end{aligned}$$

From the definition of $H_0^{\mathbb{F}}(t)$ we have

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

In the study of I_2 we exploit the concavity of G , use the Itô formula for the product and the duality between the NA-derivative and the Itô integral. Finally we obtain

$$\begin{aligned}
 I_2 \leq & \mathbb{E} \left[\int_0^T \left\{ \hat{p}_t \left((b_t(t) - \hat{b}_t(t)) + \int_0^t (\partial_t b_s(t) - \partial_t \hat{b}_s(t)) ds \right. \right. \right. \\
 & + \left. \left. \int_0^t \int_{\mathbb{R}} (\partial_t \kappa_s(t, z) - \partial_t \hat{\kappa}_s(t, z)) \mu(ds dz) \right) \right\} dt \\
 & - \int_0^T \partial_x \mathcal{H}_t^{\hat{u}} (X_t - \hat{X}_t) dt + \int_0^T \left\{ [\kappa_t(t, 0) - \hat{\kappa}_t(t, 0)] \hat{q}_t(0) \lambda_t^B \right. \\
 & \left. + \int_{\mathbb{R}_0} [\kappa_t(t, z) - \hat{\kappa}_t(t, z)] \hat{q}_t(z) \nu(dz) \lambda_t^H \right\} dt \Big]. \tag{31}
 \end{aligned}$$

Now notice that,

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

The use of Itô formula is justified by the transformation rule [41, Theorem 3.3] and assumption (25). Then we can substitute the above into (31).

The claim is then achieved thanks to (30) following a separation hyperplane argument applied to the concave function (29). □

In the arguments above we have assumed that the integrals are well defined. Sufficient conditions on the coefficients of (24) can be imposed in the same spirit as the ones proposed in Sect. 3.

5 MP for forward–backward systems

Inspired by the concept of recursive utility introduced in [31], we see that the performance functional in problem (22)

$$J(\hat{u}) = \sup_{u \in \mathcal{A}^{\mathbb{F}}} J(u),$$

can include evaluations involving a backward differential equation in the form

$$J(u) := \mathbb{E} \left[\int_0^T F_t(\lambda_t, u_t, X_t, Y_t) dt + G(X_T) + \Gamma(Y_0) \right] \tag{32}$$

where

$$\begin{aligned}
 Y_t = & h(X_T) + \int_t^T g_s(\lambda_s, u_s, X_s, Y_s, \Theta_s) ds \\
 & - \int_t^T \int_{\mathbb{R}} \Theta_s(z) \mu(ds dz) - \int_t^T dM_s
 \end{aligned} \tag{33}$$

is a BSDE driven by the noise μ considered under the filtration \mathbb{F} . The driver $g(\lambda, u, x, y, \theta)$ is an \mathbb{F} -adapted process in $L^2(dt \times d\mathbb{P})$ for all $\lambda \in [0, \infty)^2, u \in \mathbb{U}, x \in \mathbb{R}, y \in \mathbb{R}, \theta \in \mathcal{Z}, C^1$ both with respect to s and x , and with partial derivatives in $L^2(d\mathbb{P})$. Also we assume that $h(x)$ is \mathbb{F}_T -measurable for all $x \in \mathbb{R}$ and C^1 with respect to x , a.s. For simplicity we assume that is bounded and with bounded derivative. These conditions are sufficient to technically guarantee the well definition of some integrals that occur later in the proofs. They can be weekend in specific models. The BSDE (33) is studied under the filtration \mathbb{F} within the setting of general martingale noise. This yields the presence of the orthogonal martingale M as intrinsic part of the solution, which derives from the stochastic integral representations of type (4).

The forward dynamics of X are given in (24) with Volterra features and satisfying the same assumptions presented in Sect. 4. Observe that (24)–(33) is a *partially coupled* system, the solution of which is not a challenge. First one solves (24), then the value of X_T is used in the final value of (33). For the existence of a strong solution (Y, Θ, M) to (33) we can refer, e.g., to [40].

Here above $\Gamma(x)$ is a random variable for all $x \in \mathbb{R}$, concave and differentiable in x a.s., while F and G satisfy the same conditions introduced earlier in Sect. 3.

In [6] we see one of the first works in which a couple forward–backward system considers the backward equation with respect to a general filtration (thus presenting an orthogonal martingale part). In this work the driving noise is however a Brownian motion only. The goal of that paper is to prove existence of an optimal control.

In our context we see that a maximum principle can be obtained by an adequate modification of the Hamiltonian functional, which will in turn be coupled with a forward–backward system of equations. Indeed we have

$$\mathcal{H} : [0, T] \times \Phi_{[0, \infty)^2} \times \Phi_{\mathbb{R}} \times \Phi_{\mathbb{R}} \times \Phi_{\mathbb{R}} \times \Phi_{\mathcal{Z}} \times \mathbb{R} \times \Phi_{\mathbb{R}} \times \Phi_{L^2(\mathbb{P})} \times \Phi_{\mathcal{Z}} \times \Omega \longrightarrow \mathbf{R},$$

given by

$$\mathcal{H}_t(\lambda, u, x, y, \theta, y^0, \zeta, p, q) := H_0(t) + H_1(t), \tag{34}$$

where

$$\begin{aligned} H_0(t) &:= F_t(\lambda_t, u_t, x_t, y_t) + b_t(t, \lambda_t, u_t, x_t)p_t + \kappa_t(t, 0, \lambda_t, u_t, x_t)q_t(0)\lambda_t^B \\ &\quad + \int_{\mathbb{R}_0} \kappa_t(t, z, \lambda_t, u_t, x_t)q_t(z)\lambda_t^H v(dz) + g_t(\lambda_t, u_t, x_t, y_t, \theta_t)\zeta_t \\ H_1(t) &:= \int_0^t \partial_t b_s(t, \lambda_s, u_s, x_s)ds p_t + \int_0^t \partial_t \kappa_s(t, 0, \lambda_s, u_s, x_s)\mathcal{D}_{s,0} p_t \lambda_s^B ds \\ &\quad + \int_0^t \int_{\mathbb{R}} \partial_t \kappa_s(t, z, \lambda_s, u_s, x_s)\mathcal{D}_{s,z} p_t \lambda_s^H v(dz)dt. \end{aligned}$$

As in Sect. 4, the NA-derivative (6), (7) appears in presence of Volterra structures. Also we see that the functional H_0 contains elements from the backward differential equation (33) in the performance functional. For $\lambda \in \mathcal{L}$ and in association to the coupled system (24)–(33), we then consider the functional

$$\mathcal{H}(t) := \mathcal{H}_t(\lambda, u, X, Y, \Theta, y^0, \zeta, p, q),$$

and the adjoint forward–backward system $(\zeta, (p, q))$ here below.

For $\theta \in \mathcal{Z}$, we denote with ∂_{θ_0} the partial derivative with respect to $\theta(0)$ and ∇_{θ_z} denotes the Fréchet derivative with respect to $\theta(z), z \neq 0$. We also denote with $\frac{d}{dv} \nabla_{\theta_z} \mathcal{H}_i(t)$ the Radon-Nikodym derivative of $\nabla_{\theta_z} \mathcal{H}(t)$ with respect to $v(dz)$.

The process ζ satisfies stochastic forward equation

$$\begin{cases} d\zeta_t = \partial_y \mathcal{H}(t)dt + \partial_{\theta_0} \mathcal{H}(t)B(dt) + \int_{\mathbb{R}_0} \frac{d}{dv} \nabla_{\theta_z} \mathcal{H}(t) \tilde{H}(dt dz) \\ \zeta_0 = \partial_y \psi(y^0). \end{cases} \tag{35}$$

Later on the value $\zeta_0 = y^0 = Y_0$ will be associated to the initial value of the backward dynamics Y .

The stochastic backward equation is

$$\begin{cases} dp_t = -\partial_x \mathcal{H}(t)dt + \int_{\mathbb{R}} q_t(z)\mu(dt dz) \\ p_T = \partial_x \varphi(X_T) + h(X_T)\zeta_T, \end{cases} \tag{36}$$

For the sequel we assume:

- $\mathcal{H}(t)$ is well defined, differentiable with respect to $x, y, \theta(0), u$ and Fréchet differentiable with respect to $\theta(z), z \neq 0$.
- $\frac{d}{dv} \nabla_{\theta_z} \mathcal{H}(t)$ is well defined.

As already noticed in Sect. 4, we remark that (36) is a true BSDE since the term $\partial_x H_1$ in the driver $\partial_x \mathcal{H}$ is integrated (see [23, Remark 3.3]). Also the forward–backward system (35), (36) is *partially coupled* and studied under filtration \mathbb{G} . The solution of (35) leads to the terminal condition in (36). Sufficient conditions for the existence of strong solutions to (35) and (36) are well known. See, e.g., [33], [29].

In this framework the admissible controls are given as follows.

Definition 6 The admissible controls in the stochastic optimal control problem with performance (32) are \mathbb{F} -predictable stochastic processes $u : [0, T] \times \Omega \rightarrow \mathcal{U}$, such that the corresponding forward–backward system $(X, (Y, \Theta, M))$ in (24)–(33) has a unique strong solution and the adjoint forward–backward system $(\zeta, (p, q))$ in (35), (36) has a unique solution, and

$$\mathbb{E} \left[\int_0^T |F_t(\lambda_t, u_t, X_t)|^2 dt + |G(X_T)| + |\partial_x G(X_T)|^2 + |\Gamma(X_T)| + |\partial_x \Gamma(X_T)|^2 \right] < \infty.$$

The set of admissible controls is still denoted by $\mathcal{A}^{\mathbb{F}}$.

Following the approach already suggested in the previous sections, we present the maximum principle exploiting a form of “projection” of the Hamiltonian functional on the filtration \mathbb{F} .

Theorem 5 Fix $\lambda \in \mathcal{L}$. Let $\hat{u} \in \mathcal{A}^{\mathbb{F}}$ and assume that the corresponding solutions $\hat{X}, (\hat{Y}, \hat{\Theta}, \hat{M})$ of (24) and (33) exist together with the solutions $\hat{\zeta}, (\hat{p}, \hat{q})$ of (35), (36), with $y^0 = \hat{Y}_0$ in the initial value. Define the \mathbb{F} -Hamiltonian functional as

$$\mathcal{H}_t^{\mathbb{F}}(\lambda, u, x, y, \theta, \hat{Y}_0, \hat{\zeta}, \hat{p}, \hat{q}) := \mathbb{E} \left[\mathcal{H}_t(\lambda, u, x, y, \theta, \hat{Y}_0, \hat{\zeta}, \hat{p}, \hat{q}) | \mathcal{F}_t \right]$$

for $t \in [0, T]$, $u \in \Phi_{\mathbb{R}}, x, y \in \Phi_{\mathbb{R}}, \theta \in \Phi_{\mathcal{Z}}$, which is naturally given by the sum of the following two functionals

$$\begin{aligned} H_0^{\mathbb{F}}(t) &:= F_t(\lambda_t u_t, x_t, y_t) + b_t(t, \lambda_t, u_t, x_t) \mathbb{E}[\hat{p}_t | \mathcal{F}_t] + \kappa_t(t, 0, \lambda_t, u_t, x_t) \mathbb{E}[\hat{q}_t(0) | \mathcal{F}_t] \lambda_t^B \\ &+ \int_{\mathbb{R}_0} \kappa_t(t, z, \lambda_t, u_t, x_t) \mathbb{E}[\hat{q}_t(z) | \mathcal{F}_t] \lambda_t^H v(dz) + g_t(\lambda_t, u_t, x_t, y_t, \theta_t) \mathbb{E}[\hat{\zeta}_t | \mathcal{F}_t] \end{aligned}$$

$$\begin{aligned}
 H_1^{\mathbb{F}}(t) &:= \int_0^t \partial_t b_s(t, \lambda_s, u_s, x_s) ds \mathbb{E}[\hat{p}_t | \mathcal{F}_t] + \int_0^t \partial_t \kappa_s(t, 0, \lambda_s, u_s, x_s) \mathbb{E}[\mathcal{D}_{s,0} \hat{p}_t | \mathcal{F}_t] \lambda_s^B ds \\
 &+ \int_0^t \int_{\mathbb{R}} \partial_t \kappa_s(t, z, \lambda_s, u_s, x_s) \mathbb{E}[\mathcal{D}_{s,z} \hat{p}_t | \mathcal{F}_t] \lambda_s^H v(dz) ds
 \end{aligned}$$

Assume that

- the functionals

$$x, y, \theta \mapsto \operatorname{ess\,sup}_{u \in \mathcal{U}} \mathcal{H}_t^{\mathbb{F}}(\lambda, u, x, y, \theta, \hat{Y}^0, \hat{\zeta}, \hat{p}, \hat{q}) \tag{37}$$

exist and are concave for all t

- and, for all $t \in [0, T]$,

$$\sup_{u \in \mathcal{U}} \mathcal{H}_t^{\mathbb{F}}(\lambda, u, \hat{X}, \hat{Y}, \hat{\Theta}, \hat{Y}^0, \hat{\zeta}, \hat{p}, \hat{q}) = \mathcal{H}_t^{\mathbb{F}}(\lambda, \hat{u}, \hat{X}, \hat{Y}, \hat{\Theta}, \hat{Y}^0, \hat{\zeta}, \hat{p}, \hat{q}). \tag{38}$$

Then \hat{u} is an optimal control for the forward–backward system (24)–(33) with performance (32).

6 Concluding remarks

In this work we have summarised how stochastic control for time changed Lévy dynamics can be addressed in the framework of maximum principles. To exploit the structure of the time changed noise, we have used different filtrations, a technical one which is anticipating of the information on the time change process and the natural one generated by the noise. Correspondingly, we have used different types of Hamiltonian functionals. The use of different information flows has impact on the possibility to solve backward stochastic differential equations, which in turn depend on different forms of stochastic integral representations.

In our work we have seen that with appropriate modifications of the Hamiltonian functionals we can deal also with Volterra type dynamics, which involve memory in the coefficients and not only in the noise via time change. The work has progressed to include forward–backward systems of equations.

We remark that in Sect. 5, the forward–backward system involved a Volterra type of forward dynamics (24) and a backward stochastic differential equation (33). It is natural to ask whether one could extend this to Volterra type backward equations. It turns out that this extension is rather tricky. Indeed, if the time-change was *deterministic*, then we would be able to progress in our solution. In this case the filtrations involved would coincide $\mathbb{F} = \mathbb{G}$ and we would not have to deal with orthogonal martingale parts in the solution of a Volterra version of Eq. (33). See [24] for details. On the opposite, if the time change is genuinely stochastic, then the Volterra version of equation (33) would present an orthogonal component, which would have to be dealt with. For this some necessary assumptions on the structure of the orthogonal martingale part would be necessary, but there are no result that would guarantee those assumption at the time being.

In this presentation we have discussed only sufficient maximum principles, which are forms of verification that an “educated” guess of possible optimal control is indeed optimal. Necessary maximum principle are useful to identify those guesses. A version of a necessary maximum principle is available in the framework of Sects. 3 and 4, see [23]. The results are not yet polished enough in the case of Volterra type dynamics as presented in Sect. 5. There as well some structure on the orthogonal martingale part of (33) would be necessary.

Acknowledgements This note presents a survey, necessarily partial, of works on calculus and control that have seen many collaborators involved, colleagues and students. In the many travels to Africa I have met many stochastic enthusiastic students and colleagues that it was easy to start collaboration in science and for science. Persisting in the development and diffusion of mathematical knowledge is the challenge. For this, fortunately, there are many individual mathematicians and institutions engaged, which is good as *Mathematics Unites!* On a more formal note I thank the committee for awarding the ICIAM Su-Buchin Prize 2019, it is a recognition that highlights all efforts of promoting excellency in mathematics all around the world. I also thank the Norwegian Research Council for the support in many early and present projects, including the current one *STORM—Stochastics for Time-Space Risk Models* (nr. 274410), sponsoring also research collaborations and with this development.

Funding Open access funding provided by University of Oslo (incl Oslo University Hospital)

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

1. Agram, N., Øksendal, B.: Malliavin calculus and optimal control of stochastic Volterra equations. *J. Optim. Theory Appl.* **167**, 1070–1094 (2015)
2. Agram, N., Øksendal, B.: A Hida–Malliavin white noise calculus approach to optimal control. *Infin. Dimens. Anal. Quantum Probab. Relat.* **21**(3), 1850014 (2018)
3. Agram, N., Øksendal, B., Yakhlef, S.: New approach to optimal control of stochastic Volterra integral equations. *Stochastics* **91**, 873–894 (2019)
4. Albeverio, S., Rüdiger, B.: Infinite-dimensional stochastic differential equations obtained by subordination and related Dirichlet forms. *J. Funct. Anal.* **204**, 122–156 (2003)
5. Ané, T., Geman, H.: Order flow, transaction clock and normality of asset returns. *J. Financ.* **55**, 2259–2285 (2000)
6. Baghery, F., Khellallah, N., Mezerdi, B., Turpin, I.: On optimal control of forward–backward stochastic differential equations. *Afr. Mat.* **28**, 1075–1092 (2017)
7. Barndorff-Nielsen, O.E., Blæsild, P., Schmiegel, J.: A parsimonious and universal description of turbulent velocity increments. *Eur. Phys. J. B* **41**, 345–363 (2004)
8. Barndorff-Nielsen, O.E., Nicolato, E., Shephard, N.: Some recent developments in stochastic volatility modelling. *Quant. Financ.* **2**, 11–23 (2002)
9. Barndorff-Nielsen, O.E., Pedersen, J.: Meta-times and extended subordination. *Theory Probab. Appl.* **56**, 319–327 (2012)
10. Barndorff-Nielsen, O.E., Shiryaev, A.: Change of Time and Change of Measure, Advanced Series on Statistical Science and Applied Probability 21, 2nd edn. World Scientific Publishing Co., Hackensack (2015)
11. Barndorff-Nielsen, O.E., Schmiegel, J.: Time change, volatility, and turbulence. In: *Mathematical Control Theory and Finance*. Springer, Berlin, pp. 29–53 (2008)
12. Benth, F.E., Krühner, P.: Integrability of multivariate subordinated Lévy processes in Hilbert space. *Stochastics* **87**, 458–476 (2015)
13. Carr, P., Geman, H., Madan, D.B., Yor, M.: Stochastic volatility for Lévy processes. *Math. Financ.* **13**, 345–382 (2003)
14. Carr, P., Geman, H., Madan, D.B., Yor, M.: The fine structure of asset returns: an empirical investigation. *J. Bus.* **75**, 305–332 (2002)
15. Carr, P., Wu, L.: Time-changed Lévy processes and option pricing. *J. Financ. Econ.* **71**, 113–141 (2004)
16. Chou, C.-S., Meyer, P.A.: Sur la représentation des martingales comme intégrales stochastiques dans les processus ponctuels. *Séminaire de probabilités de Strasbourg* **9**, 226–236 (1975)
17. Dambis, K.É.: On decomposition of continuous submartingales. *Teor. Veroyatnost. i Primenen.* **10**, 438–448 (1965)

18. De Donno, M.: The term structure of interest rates as a random field: a stochastic integration approach. In: *Stochastic Processes and Applications to Mathematical Finance, Proceedings of the Ritsumeikan International Symposium*, pp. 27–51 (2004)
19. Dellacherie, C.: Intégrales stochastiques par rapport aux processus de Wiener ou de Poisson. *Séminaire de probabilités de Strasbourg* **8**, 25–26 (1974)
20. Di Nunno, G.: Stochastic integral representations, stochastic derivatives and minimal variance hedging. *Stoch. Stoch. Rep.* **73**(1–2), 181–198 (2002)
21. Di Nunno, G.: Random fields evolution: non-anticipating integration and differentiation. *Theory Probab. Math. Stat.* **66**, 91–104 (2006)
22. Di Nunno, G., Eide, I.B.: Minimal variance hedging in large financial markets: random fields approach. *Stoch. Anal. Appl.* **28**, 54–85 (2010)
23. Di Nunno, G., Giordano, M.: Backward Volterra equations with time-changed Lévy noise and maximum principles. [ArXiv:2011.07341](https://arxiv.org/abs/2011.07341) (2021)
24. Di Nunno, G., Giordano, M.: Maximum principles for stochastic time-changed Volterra games. Technical report, University of Oslo (2022)
25. Di Nunno, G., Øksendal, B.: Optimal portfolio, partial information and Malliavin calculus. *Stochastics* **81**(3–4), 303–322 (2009)
26. Di Nunno, G., Øksendal, B., Proske, F.: *Malliavin Calculus for Lévy Processes with Applications to Finance*. Springer, Berlin (2009)
27. Di Nunno, G., Rozanov, Y.A.: Stochastic integrals and adjoint derivatives. In: *Stochastic analysis and applications*, pp. 265–307. Abel Symposium, 2, Springer, Berlin (2007)
28. Di Nunno, G., Sjursen, S.: On chaos representation and orthogonal polynomials for the doubly stochastic Poisson process. In: Dalang, R.C., Dozzi, M., Russo, F. (eds.) *Seminar on Stochastic Analysis, Random Fields and Applications VII*, Volume 67 of *Progress in Probability*, pp. 23–54. Springer, Basel (2013)
29. Di Nunno, G., Sjursen, S.: BSDEs driven by time-changed Lévy noises and optimal control. *Stoch. Process. Appl.* **124**, 1679–1709 (2013)
30. Dubins, L.E., Schwarz, G.: On continuous martingales. *Proc. Natl. Acad. Sci. USA* **53**, 913–916 (1965)
31. Duffie, D., Epstein, L.G.: Stochastic differential utility. *Econometrica* **60**, 353–394 (1992)
32. Grigelionis, G.: Characterization of stochastic processes with conditionally independent increments. *Lith. Math. J.* **15**, 562–567 (1975)
33. Jacod, J.: Existence and uniqueness for stochastic differential equations. In: Kohlmann, M., Vogel, W. (eds.) *Stochastic Control Theory and Stochastic Differential Systems*. Lecture Notes in Control and Information Sciences, vol. 16, pp. 435–446. Springer, Berlin (1979)
34. James, L.F., Müller, G., Zhang, Z.: Stochastic volatility models based on OU-Gamma time change: theory and estimation. *J. Bus. Econ. Stat.* **36**, 75–87 (2018)
35. Kallenberg, O.: *Foundations of Modern Probability*. Springer, Cham (2021)
36. Kallsen, J.: A didactic note on affine stochastic volatility models. In: Kabanov, Y., Liptser, R., Stoyanov, J. (eds.) *From Stochastic Calculus to Mathematical Finance*, pp. 343–368. Springer, Berlin (2006)
37. Mandrekar, V., Salehi, H.: Subordination of infinite-dimensional stationary stochastic processes. *Ann. Inst. H. Poincaré Sect. B (N.S.)* **6**, 115–130 (1970)
38. Oblój, J.: The Skorokhod embedding problem and its offspring. *Probab. Surv.* **1**, 321–390 (2004)
39. Øksendal, B., Sulem, A.: *Applied Stochastic Control of Jump Diffusions*. Springer, Cham (2019)
40. Papantoleon, A., Possamaï, D., Saplaouras, A.: Existence and uniqueness results for BSDE with jumps: the whole nine yards. *Electron. J. Probab.* **23**, 1–68 (2018)
41. Protter, P.: Volterra equations driven by semimartingales. *Ann. Probab.* **13**, 519–530 (1993)
42. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton (1970)
43. Rosiński, J., Woyczyński, W.A.: On Itô stochastic integration with respect to p -stable motion: inner clock, integrability of sample paths, double and multiple integrals. *Ann. Probab.* **14**, 271–286 (1986)
44. Saakvitne, J.A.: *Essays on market microstructure*. Ph.D. thesis, BI Norwegian Business School, Oslo, Norway (2019)
45. Sauri, O., Veraart, A.E.D.: On the class of distributions of subordinated Levy processes. *Stoch. Process. Appl.* **127**, 475–496 (2017)
46. Serfozo, R.F.: Processes with conditional stationary independent increments. *J. Appl. Probab.* **9**(2), 303–315 (1972)
47. Shiryaev, A.N.: *Essentials of Stochastic Finance*, Advanced Series on Statistical Science and Applied Probability, 3. World Scientific Publishing Co., River Edge (1999)
48. Swishchuk, A.: *Change of Time Methods in Quantitative Finance*. SpringerBriefs in Mathematics. Springer, Cham (2016)

49. Yablonsky, A.L.: The Malliavin calculus for processes with conditionally independent increments. In: Benth, F.E., Di Nunno, G., Lindstrøm, T., Øksendal, B., Zhang, T. (eds.) *Stochastic Analysis and Applications, The Abel symposium 2005*, pp. 641–678. Springer, Berlin (2007)
50. Winkel, M.: The recovery problem for time-changed Lévy processes. Research Report MaPhySto, Aarhus, Denmark 2001–37 (2001)

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