ESAIM: COCV 25 (2019) 61
https://doi.org/10.1051/cocv/2018066

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

OPTIMAL RELAXED CONTROL OF STOCHASTIC HEREDITARY EVOLUTION EQUATIONS WITH LÉVY NOISE

Utpal Manna[1,*] and Debopriya Mukherjee[1,2]

Abstract. Existence theory of optimal relaxed control problem for a class of stochastic hereditary evolution equations driven by Lévy noise has been studied. We formulate the problem in the martingale sense of Stroock and Varadhan to establish existence of optimal controls. The construction of the solution is based on the classical Faedo–Galerkin approximation, the compactness method and the Jakubowski version of the Skorokhod theorem for nonmetric spaces, and certain compactness properties of the class of Young measures on Suslin metrizable control sets. As application of the abstract theory, Oldroyd and Jeffreys fluids have been studied and existence of optimal relaxed control is established. Existence and uniqueness of a strong solution and uniqueness in law for the two-dimensional Oldroyd and Jeffreys fluids are also shown.

Mathematics Subject Classification. 93E20, 60H30, 76A10, 49J20.

Received February 6, 2017. Accepted November 22, 2018.

1. Introduction

In this paper, our goal is to study an optimal control problem (OCP) associated to a class of abstract stochastic hereditary evolution equations (SHEEs) driven by Lévy noise, in an infinite-dimensional separable Hilbert space H, of the general form

$$\begin{aligned} \mathrm{d}X(t) + \Big[AX(t) + B(X(t)) + \int_0^t \beta(t-s)AX(s)\,\mathrm{d}s\Big]\mathrm{d}t \\ = \mathsf{L}(U(t))\mathrm{d}t + \sigma(t, X(t))\mathrm{d}W(t) + \int_Z g(X(t-), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z), \end{aligned} \tag{1.1}$$

$$\text{with initial condition} \qquad X(0) = x_0, \tag{1.2}$$

Keywords and phrases: Relaxed controls, Young measure, hereditary evolution equations, martingale solution, Oldroyd fluid, Jeffreys fluid.

[1] School of Mathematics, Indian Institute of Science Education and Research Thiruvananthapuram, Thiruvananthapuram 695551, Kerala, India.
[2] School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.

[*] Corresponding author: manna.utpal@iisertvm.ac.in

Article published by EDP Sciences

associated with a cost functional of the form

$$\bar{J}(X,U) := \mathbb{E}[\mathcal{J}(X,U)] := \mathbb{E}\Big[\int_0^T F(t,X(t),U(t))\mathrm{d}t + \phi(X(T))\Big]. \tag{1.3}$$

Here, $x_0 \in H$ and X is a state variable. The operators A, B satisfy certain conditions which are framed in Section 2. U is a metrizable Suslin space $\mathbb{U}$-valued control depending on the possible nonlinearities of the mechanism of the system, $\mathsf{L} : \mathbb{U} \to H$ can be linear (or possibly nonlinear) operator. $\{W(t)\}_{t\geq 0}$ is a cylindrical H-valued Wiener process and $\tilde{N}(\mathrm{d}t, \mathrm{d}z)$ is a time homogeneous compensated Poisson random measure with Z being the measurable subset of H where the solutions of the above system have its paths and N is a time-homogeneous Poisson random measure on $(Z, \mathcal{B}(Z))$. The processes W and N are assumed to be independent. $F : [0,T] \times H \times \mathbb{U} \to [0,\infty]$ is measurable in $t \in [0,T]$ and lower semicontinuous with respect to $(X,U) \in H \times \mathbb{U}$ and the final cost function $\phi : H \to \mathbb{R}$ is lower semicontinuous.

In the context of the natural generalization of OCP governed by SHEE, when no special condition on the dependence of the non-linear operator L with respect to the control variable are assumed and when the cost functional fail to satisfy the usual convexity condition, introduction of measure valued controls (known as relaxed controls) becomes necessary. By the method of relaxation (due to L. C. Young and J. Warga), the space of admissible controls can be extended to a bigger space known as the space of admissible relaxed controls such that control system and the cost functional are convexified in the usual sense. This gives us a passage to seek the existence of optimal control for the relaxed system.

We disintegrate (see Appendix A.3 and Lem. A.16) a Borel probability measure μ on the Borel σ-algebra $\mathcal{B}(\mathbb{U} \times [0,T])$ for the projection $\mathbb{U} \times [0,T] \to \mathbb{U}$ as

$$\mu(\mathrm{d}U, \mathrm{d}t) = \mu_t(\mathrm{d}U)\mathrm{d}t, \quad \mu \in \mathsf{P}(\mathbb{U} \times [0,T]),$$

where $\mathsf{P}(\mathbb{U} \times [0,T])$ denotes the space of Borel probability measure on $\mathcal{B}(\mathbb{U} \times [0,T])$.

With the original control system (1.1)–(1.3), we associate a relaxed control system by defining the relaxed coefficient as

$$\mathcal{N}\mu_t := \int_{\mathbb{U}} \mathsf{L}(U)\mu_t(\mathrm{d}U), \quad \mu_t \in \mathsf{P}(\mathbb{U}), \quad t \in [0,T], \tag{1.4}$$

where $\mathsf{P}(\mathbb{U})$ denotes the space of Borel probability measure on $\mathcal{B}(\mathbb{U})$.

Hence, the relaxed version of the original control system (1.1)–(1.2) become:

$$\begin{aligned}&\mathrm{d}X(t) + \Big[AX(t) + B(X(t)) + \int_0^t \beta(t-s)AX(s)\,\mathrm{d}s\Big]\mathrm{d}t\\ &= \mathcal{N}\mu_t\mathrm{d}t + \sigma(t,X(t))\mathrm{d}W(t) + \int_Z g(X(t-),z)\tilde{N}(\mathrm{d}t,\mathrm{d}z),\end{aligned} \tag{1.5}$$

$$\text{with initial condition} \qquad X(0) = x_0, \tag{1.6}$$

associated with the relaxed cost functional as

$$\bar{J}(X,\mu) := \mathbb{E}[\mathcal{J}(X,\mu)] := \mathbb{E}\Big[\int_0^T \bar{F}(t,X(t),\mu_t)\mathrm{d}t + \phi(X(T))\Big], \tag{1.7}$$

where the relaxed running cost $\bar{F}$ is defined by

$$\bar{F}(t,X,\mu_t) := \int_{\mathbb{U}} F(t,X,U)\mu_t(dU), \quad \mu_t \in \mathsf{P}(\mathbb{U}), \quad (t,X) \in [0,T]\times H,$$

whenever the above expression is well defined, *i.e.*, the map $\mathbb{U} \ni U \mapsto F(t,X,U) \in H$ is Bochner integrable with respect to $\mu_t \in \mathsf{P}(\mathbb{U})$.

Note that the original control system (1.1)–(1.2) controlled by a $\mathbb{U}$-valued process $U = \{U(t)\}_{t\in[0,T]}$ coincides with the relaxed system controlled by the process of Dirac measures $\mu_t = \delta_{U(t)}$, $t \in [0,T]$.

1.1. Motivation of the study and relevant literatures

It is well-known that certain physical and biological phenomena can be better understood if past history of the system is taken into consideration in the model equations. Such equations are known as hereditary equations or equations with memory. Examples of such models may include (but not limited to)

(a) non-Newtonian fluid flow problems, *e.g.* in viscoelastic and polymeric fluids (see Part I in [68], [74]),
(b) Newtonian averaged field dynamics [24, 52, 66],
(c) heat-flow problems with memory [7, 46],
(d) time-delayed reaction-diffusion systems [31, 35],
(e) hyperbolic heat equation with memory [34], hyperbolic phase-field model with memory [3, 36],
(f) plasmas turbulence [33],
(g) bacteria growth [43, 71],
(h) financial market model with memory [1].

It is needless to say that the hereditary systems become more physically realistic if some kind of uncertainty is also considered in the model equations formulation. Noise can enter the system in a number of ways, *e.g.* in the state equation, forcing due to structural vibration or environmental effects can be incorporated either as a random boundary forcing or as a random distributed forcing. The abstract hereditary model (1.5)–(1.7) considered in this paper is motivated by parabolic partial differential equations arising in fluid dynamics (see Sect. 6 for more details) and in heat flows and their application in OCPs. It is worth to mention here that the applications to OCPs quite naturally arise in the study of control of fluid flows, control of stationary flows, glueing in polymeric materials, heating processes *e.g.* in modelling heating with radiation boundary condition, optimal portfolio problem in a financial market model with memory etc. Interested readers may look into Lions [45] and Tröltzsch [73] (see also Chang [20] and references therein) for more details and applications.

There are considerable amount of literatures available concerning deterministic system of integro-differential equations (see, for instance, [64]) and relatively less amount of work on stochastic systems (*e.g.* [2],[7], [8], to name a few). Semigroup approaches of the deterministic theory (*e.g.* history function approach due to Miller [51]) have recently been introduced in stochastic models by Bonaccorsi *et al.* [10, 11]. When the infinitesimal generator is a family of linear unbounded time dependent operator, it is difficult to describe the semigroup and to study the corresponding stochastic convolution problem associated with the system. Under such set-up, resolvent theory approach has been developed (see, for instance, [21], [63]).

Study of OCP is quite delicate when both noise and control influence dynamics of a system in the presence of memory term, and amount of literature in this direction is quite limited and very recent. Optimal control for stochastic Volterra equations with monotone kernels and that of stochastic heat equation with memory have recently been studied in [10] and [22] with the aid of semigroup and resolvent approaches, respectively. Taking a different approach, Chang, in his monograph [20], has studied hereditary systems influenced by standard Brownian motion with bounded fading memory over finite time horizon via dynamic programming principle focusing on development of the value function as the unique viscosity solution of the corresponding infinite-dimensional equations. Very recently, noting the fact that solutions of SHEE are not usually Markov processes, Agram *et al.* in [1] have used tools from Malliavin calculus to tackle OCP for such equations with jumps.

The approach adopted in this paper is different from the above literatures. We draw motivation from the celebrated works of Young [79, 80] and Warga [77, 78] on measure-valued convexification technique of nonlinear controlled system and relaxation technique of cost functional on the space of admissible controls. It is worth to mention here that this approach was first introduced for finite-dimensional stochastic systems by Fleming *et al.* [29, 30], and this area then gradually expanded due to [28, 32, 38, 56, 81], to name a few. Sritharan in [69] studied optimal relaxed control of stochastic Navier–Stokes equations with linear and nonlinear constitutive relations and with Lusin metrizable control sets. Cutland *et al.* in [25, 26] combined relaxed controls and non-standard analysis techniques to study existence of an optimal control for stochastic 2D and 3D Navier–Stokes equations in a bounded domain with multiplicative Gaussian noise. Brzeźniak and Serrano in [18] have analyzed the existence of a weak optimal relaxed control for semilinear dissipative stochastic evolution equations influenced by cylindrical Wiener process in unconditional martingale difference (UMD) type-2 Banach spaces implementing semigroup approach.

Apart from the application, the main mathematical motivation of this work is to combine the relaxed control techniques with martingale problem formulation in the sense of Stroock and Varadhan. It appears to authors that the approach adopted in this work is new to stochastic Volterra integro-differential equations (or to SHEE). In addition, we consider very general control set, control operator, noise and cost functional (see Sects. 1.2 and 2 for more details) to include a large class of problems under this unified framework.

1.2. Technical challenges and contribution of the paper

This work contributes to the existing literatures on optimal control of SHEEs or of stochastic Volterra integro-differential equations in a number of ways listed below.

(i) The control set in our work is assumed to be only metrizable and Suslin in contrast with [1, 22] and [10], where the control set is real-valued and real separable Hilbert space-valued, respectively. In our case, due to lack of boundedness or compactness of the control set, we follow the strategy due to Fleming [29] to work in the bigger space of admissible relaxed control where we obtain proper compact structure of the control set.

(ii) The admissible control operator assumed in this work is infinite-dimensional Hilbert space-valued, possibly non-linear, continuous and is dominated by an inf-compact function in contrast with usual square integrable control processes (*e.g.* see [10]).

(iii) The running cost function assumed in this paper is measurable, lower semicontinuous and coercive in comparison to [22], where running cost is locally Lipschitz and to [10], where running cost is having quadratic growth in both state and control variables. Since, in our case, there is no special assumption on the convexity or growth condition of cost functional, it is intuitive that the optimal solutions may not even exist leading the control process to be of infinite cost. However, by the method of relaxation for the class of weak admissible controls, the cost functional is convexified in its usual sense and it is written in terms of the admissible control system, not as a function of control process alone.

(iv) In this work, we consider multiplicative Lévy noise consisting of the Wiener processes and compensated time homogeneous Poisson random measures taking values in some Hilbert space. Our choice of noise is more general than [10, 20, 22] and [1], where in the former literatures noise is either a standard real-valued Brownian motion or a Hilbert space-valued cylindrical Wiener process, whereas in the later literature noise is a real-valued multiplicative Lévy process. The key technical issue we face in our martingale solution approach is in proving the tightness (and identification of the limit), which does not follow any standard classical methods. The essential ingredients in this approach are generalization of the Skorokhod representation theorem for non-metric spaces introduced recently in [13], and partial generalization from the case of weakly continuous (see Lem. 4.2 in [16]) to weakly cádlag functions with respect to application of Kuratowski's theorem for non-Polish spaces.

We are now ready to state the main results of this article. To this end, let us assume that $\big(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P}\big)$ is a filtered probability space, where $\big(\mathscr{F}_t\big)_{t\geq 0}$ is the filtration, satisfying the usual conditions, *i.e.*, (i) $\mathbb{P}$ is complete on $(\Omega, \mathscr{F})$, (ii) for each $t \geq 0$, $\mathscr{F}_t$ contains all $(\mathscr{F}, \mathbb{P})$-null sets, and (iii) the filtration $\mathscr{F}_t$ is right-continuous.

The main result of this article is summarized below and is proven in Theorems 4.7 and 5.4 with precise statements. Suppose $x_0 \in H$ and a system $\pi := (\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P}, \{X(t)\}_{t\geq 0}, \{N(t,\cdot)\}_{t\geq 0}, \{W(t)\}_{t\geq 0}, \{\mu_t\}_{t\geq 0})$ is a weak admissible relaxed control (see Def. 5.1) with time horizon $[0,T]$ such that $\bar{J}(\pi) < +\infty$, where $\bar{J}$ is the corresponding associated relaxed cost functional given by (5.3). Suppose Assumption 2.1, Assumption 2.2, Hypothesis 2.4, Assumption 2.7, and Assumption 2.8 are satisfied. Then the relaxed control problems (RCPs) (1.5)–(1.6) associated with the cost functional (1.7) admit a weak optimal relaxed control with time horizon $[0,T]$.

Let us brief here the outline of the proof of the above result. We first consider the classical Faedo–Galerkin approximation of the uncontrolled system of equation corresponding to (1.5)–(1.6), *i.e.*

$$
\begin{aligned}
\mathrm{d}X_n(t) &= -\big[P_nAX_n(t) + B_n(X_n(t)) + \int_0^t \beta(t-\tau)P_nAX_n(\tau)\mathrm{d}\tau \\
&\qquad + P_n\sigma(t, X_n(t))\mathrm{d}W(t) + \int_Z P_ng(X_n(t-), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z), \\
X_n(0) &= P_nX(0),
\end{aligned}
$$

and prove suitable uniform bounds for X_n in Section 3, namely,

$$
\sup_n \mathbb{E}\left(\sup_{0\leq t\leq T} |X_n(t)|_H^2\right) < \infty, \quad \sup_n \mathbb{E}\left(\int_0^T \|X_n(t)\|_V^2 \,\mathrm{d}t\right) < \infty.
$$

The spaces H and V are defined in Section 2. Using certain deterministic compactness results (see Thm. A.7) we are able to establish, in Lemma 4.3, the Aldous condition to obtain tightness criterion for the laws $\{\mathcal{L}(X_n), n \in \mathbb{N}\}$ on an appropriate functional space denoted as $\mathcal{Z}$, which is not metrizable. We further employ Jakubowski's version of the Skorokhod representation theorem for nonmetric spaces, which allows us to construct a stochastic process $\bar{X}$ with trajectories in the space $\mathcal{Z}$, a Wiener process $\bar{W}$, a time homogeneous Poisson random measure $\bar{N}$ defined on some filtered probability space $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathscr{F}}_t, \bar{\mathbb{P}})$, such that the system $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathscr{F}}_t, \bar{\mathbb{P}}, \bar{X}, \bar{N}, \bar{W})$ is a martingale solution of the uncontrolled problem corresponding to (1.5)–(1.6) (see Thm. 4.7 for details).

Next, we move to seek weak optimal solution for the finite horizon RCP *i.e.*, (1.5)–(1.6) subject to associated cost functional (1.7). In Theorem 5.4, we first construct a minimizing sequence of admissible relaxed controls $\pi_n := (\Omega^n, \mathscr{F}^n, \mathscr{F}^n_t, \mathbb{P}^n, \{X^n(t)\}_{t\geq 0}, \{N^n(t,\cdot)\}_{t\geq 0}, \{W^n(t)\}_{t\geq 0}, \{\mu^n_t\}_{t\geq 0})$ and prove $\{\mathcal{L}(\mu_n), n \in \mathbb{N}\}$ is tight on the space of Young measures on $\mathbb{U}$. Further we prove that π_n converges to $\tilde{\pi} := (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathscr{F}}_t, \tilde{\mathbb{P}}, \{\tilde{X}(t)\}_{t\geq 0}, \{\tilde{N}(t,\cdot)\}_{t\geq 0}, \{\tilde{W}(t)\}_{t\geq 0}, \{\tilde{\mu}_t\}_{t\geq 0})$ in appropriate topology which is also a weak martingale solution of (1.5)–(1.6) associated with the cost functional (1.7). Finally, in Theorem 5.4, we prove that $\tilde{\pi}$ is indeed a weak admissible optimal relaxed control.

In the second part of this article, in Section 6, we provide some applications of the abstract theory developed in this paper. We concentrate particularly on certain viscoelastic fluids, *e.g.* Oldroyd and Jeffreys fluid, and obtain some interesting results.

The content of this paper is as follows. In Section 2, we provide some basic definitions and list all essential assumptions. Section 3 is devoted to derive energy estimates of the approximating sequence obtained by the Faedo–Galerkin method for the uncontrolled SHEEs (3.1)–(3.2). In the next section, we prove existence of martingale solution to this uncontrolled system. In Section 5, we study the existence of weak optimal relaxed control for the RCP (1.5)–(1.6) associated with the cost functional (1.7). Section 6 is covered with applications to viscoelastic fluid flow problems such as Oldroyd fluid and Jeffreys fluid. We conclude this section by providing an example and proving the existence and uniqueness of a strong solution and uniqueness in law for the two-dimensional Oldroyd fluid. Finally the Appendix is split into 3 parts, which contain certain important results about compactness and tightness criterion for càdlàg functions, predictable processes and random young measure, positivity of the kernel, quadratic variation processes and some of its properties.

2. Assumptions

In this section, we specify the standing assumptions on the linear and nonlinear operators, the kernel, the noises, the control set, the control operator, and the running and final cost functions.

Let $(H,(\cdot,\cdot)_H)$ and $(V,(\cdot,\cdot)_V)$ be two separable Hilbert spaces and the norms in H and V induced by the inner products be denoted by $|\cdot|$ and $\|\cdot\|$, respectively. We assume that $V \hookrightarrow H$, the embedding being dense and continuous. Identifying H with its dual H', we have the following continuous embeddings

$$V \hookrightarrow H \cong H' \hookrightarrow V'.$$

2.1. Properties of the operators A and B

Assumption 2.1.

(A) Let $A : V \to V'$ be a non-negative, linear, self-adjoint, unbounded operator such that

$$\langle AX, Y\rangle = (X,Y)_V, \ X, Y \in V, \tag{2.1}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between V and V'.

(B)(i) Let $B : V \times V \to V'$ be a bilinear map and there exists a constant $C_1 > 0$ such that

$$\|B(X,Y)\|_{V'} \leq C_1\|X\|\|Y\|, \quad X, Y \in V. \tag{2.2}$$

(ii) For $X_i \in V, i = 1,2,3,$ we have $\langle B(X_1,X_2),X_3\rangle = -\langle B(X_1,X_3),X_2\rangle$. Let us denote $B(X) = B(X,X),\ \forall X \in H$.

(iii) $B : V \to V'$ is locally Lipschitz continuous, *i.e.*, for every $r > 0$ there exists a constant C_r such that

$$\|B(X) - B(Y)\|_{V'} \leq C_r\|X - Y\|, \ X, Y \in V, \|X\|, \|Y\| \leq r. \tag{2.3}$$

(iv) There exists a separable Hilbert space $V_1 \subset V$, the embedding being dense and continuous, such that B can be extended to a bilinear map from $H \times H$ into V_1'. Moreover, there exists a constant $C_2 > 0$ such that

$$\|B(X,Y)\|_{V_1'} \leq C_2|X||Y|, \quad X, Y \in H. \tag{2.4}$$

2.1.1. Auxiliary result

Let us note that we have the following three separable Hilbert spaces such that $V_1 \subset V \subset H$, the embedding being dense and continuous. Since V_1 is a separable Hilbert space, there exists a Hilbert space $\mathcal{U}$ such that $\mathcal{U} \hookrightarrow V_1$ and the embedding $\mathcal{U} \hookrightarrow V_1$ is dense and compact (see Lem. C.1 in Appendix C of [15]). In particular, $\mathcal{U}$ is compactly embedded into the space H. Hence, we have

$$\mathcal{U} \hookrightarrow V \hookrightarrow H \cong H' \hookrightarrow \mathcal{U}', \tag{2.5}$$

where the embedding $\mathcal{U} \hookrightarrow V$ is dense and compact and the embedding $V \hookrightarrow H$ is continuous.

2.2. Properties of the kernel

Here, we formulate appropriate properties of kernel β which will be useful in the later sections.

Assumption 2.2. The kernel $\beta : [0,\infty) \to \mathbb{R}$ satisfies the following assumptions:

(i) β is twice differentiable strictly convex function and $\beta \in C([0,\infty))$.
(ii) $\beta(t) \to 0 \quad \text{as} \quad t \to \infty$.

Lemma 2.3. *Let $t > 0$ and $\varphi \in \mathrm{L}^2(0,t)$, then $\int_0^t \int_0^s \beta(s-\tau)\varphi(s)\,\varphi(\tau)\mathrm{d}s\, d\tau \geqslant 0$.*

For proof see Appendix B.

2.3. Assumptions on the random forces

2.3.1. Some notations

Let $\mathcal{L}_Q(H,H)$ denote the space of all Hilbert–Schmidt operators from H to H where Q is symmetric, positive, trace class operator on H.
Let $\mathcal{P}$ be a predictable $\sigma-$field on $[0,T]\times\Omega$. Let $\mathfrak{L}^2_{\lambda,T}(\mathcal{P}\otimes\mathcal{B}(Z), l\otimes\mathbb{P}\otimes\lambda; H)$ be a space of all H-valued, $\mathcal{P}\otimes\mathcal{B}(Z)-$ measurable processes such that

$$\mathbb{E}\Big[\int_0^T \int_Z |\xi(s,\cdot,z)|^2 \mathrm{d}s\, d\lambda(z)\Big] < \infty.$$

Let $\mathbb{S}$ be a complete separable metric space with a metric ρ. We denote by $D([0,T];\mathbb{S})$, the set of all $\mathbb{S}$-valued functions defined on $[0,T]$, which are right continuous and have left limits (càdlàg functions) for every $t\in[0,T]$. More details about this space and its topology are put forward in Appendix A.

Let us assume that σ and g satisfy the following hypothesis of joint continuity, Lipschitz condition and linear growth.

Hypothesis 2.4. The main hypothesis is the following:

H.1 The function $\sigma \in C([0,T]\times H; \mathcal{L}_Q(H;H))$ and $g \in \mathfrak{L}^2_{\lambda,T}(\mathcal{P}\otimes\mathcal{B}(Z), l\otimes\mathbb{P}\otimes\lambda; H)$.
H.2 Let $p \geq 2$. Then for all $t \in (0,T)$, there exists positive constant K_p such that for all $X \in H$,

$$\|\sigma(t,X(t))\|^p_{\mathcal{L}_Q} + \int_Z |g(X(t),z)|^p \lambda(\mathrm{d}z) \leq K_p\,(1+|X(t)|^p).$$

H.3 For all $t\in(0,T)$, there exists positive constant L such that for all X, $Y \in H$,

$$\|\sigma(t,X)-\sigma(t,Y)\|^2_{\mathcal{L}_Q} + \int_Z |g(X,z)-g(Y,z)|^2\lambda(\mathrm{d}z) \leq L\,|X-Y|^2.$$

2.4. Assumptions on the control operator

Definition 2.5. A Hausdorff topological space $\mathbb{U}$ is said to be Suslin if there exist a Polish space (*i.e.*, separable and completely metrizable) S and a continuous mapping $\varphi : S \to \mathbb{U}$ such that $\varphi(S) = \mathbb{U}$.

Definition 2.6. A function $\kappa : \mathbb{U}$ (Hausdorff topological space, control set) $\to [0,+\infty]$ is called inf-compact iff for every $R \geq 0$ the level set $\{\kappa \leq R\}$ is compact.
For $T > 0$ fixed, we will denote by $\mathcal{IC}([0,T],\mathbb{U})$ the class of measurable functions $\kappa : [0,T]\times\mathbb{U} \to [0,+\infty]$ such that for all $t\in[0,T]$, the function $\kappa(t,\cdot)$ is inf-compact.

We impose the following assumptions on the control set $\mathbb{U}$, control operator L and the Young measure μ. We denote the class of weak admissible relaxed controls (with time horizon $[0,T]$) by $\mathfrak{U}^w_{ad}(X_0,T)$ and is defined in Definition 5.1.

Assumption 2.7.

1. The control set $\mathbb{U}$ is a metrizable and Suslin space.
2. L is an admissible control operator (possibly nonlinear) satisfying the following
   i. $\mathsf{L}:\mathbb{U}\to H$ is continuous and
   ii. there exists $\kappa\in\mathcal{IC}([0,T],\mathbb{U})$ such that $|\mathsf{L}(U)|\leq\kappa(\cdot,U)+C$, $U\in\mathbb{U}$.
3. For any random Young measure $\mu\in\mathfrak{U}^w_{ad}(X_0,T)$

$$\mathbb{E}\left[\int_0^T\int_{\mathbb{U}}\kappa^2(t,U)\mu(\mathrm{d}U,\mathrm{d}t)\right]<+\infty.$$

Note that the constant C does not have any significance in the subsequent estimates concerning L and hence will be ignored.

2.5. Assumptions on the running and final cost functions

Assumption 2.8.

1. The running cost $F:[0,T]\times H\times\mathbb{U}\to[0,\infty]$ is measurable in $t\in[0,T]$ and lower semicontinuous with respect to $(X,U)\in H\times\mathbb{U}$.
2. There exists $\kappa\in\mathcal{IC}([0,T],\mathbb{U})$ such that F satisfies the following coercivity condition:

$$F(t,X,U)\geq c_1\kappa^\gamma(t,U),\quad\text{with}\quad\gamma\geq 2,\ \text{for some constant } c_1>0.$$

3. The final cost function $\phi:H\to\mathbb{R}$ is lower semicontinuous.

Remark 2.9. The coercivity condition on F is important since the nonlinearity is not necessarily bounded with respect to the control variable.

3. *A priori* energy estimates for the uncontrolled equation

In this section, we consider the following uncontrolled system of SHEEs:

$$\mathrm{d}X(t)+\Big[AX(t)+B(X(t))+\int_0^t\beta(t-s)AX(s)\,\mathrm{d}s\Big]\mathrm{d}t=\sigma(t,X(t))\mathrm{d}W(t)+\int_Z g(X(t-),z)\tilde{N}(\mathrm{d}t,\mathrm{d}z),\tag{3.1}$$

$$\text{with initial condition}\qquad X(0)=x_0,\tag{3.2}$$

Let $\{e_j:j\in\mathbb{N}\}\subset D(A)$ be an orthonormal basis in H such that $\mathrm{span}\{e_j:j\in\mathbb{N}\}$ is dense in V. Let us fix $n\in\mathbb{N}$. Let P_n denote the projection operator from V' to $H_n:=\mathrm{span}\,\{e_1,e_2,\cdots,e_n\}$ defined by

$$P_nx^*:=\sum_{i=1}^n\langle x^*,e_i\rangle e_i,\quad x^*\in V',$$

where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between the spaces V and V'. Note that the restriction of P_n to H, still denoted by P_n, is given by

$$P_n x := \sum_{i=1}^{n}(x, e_i)_H e_i, \quad x \in H,$$

and thus it is the $(\cdot,\cdot)_H-$orthogonal projection onto H_n. Observe that

$$(P_n x^*, y)_H = \langle x^*, P_n y\rangle, \quad x^* \in V', \quad y \in V.$$

Let us denote $\tilde{e_i} := \dfrac{e_i}{\|e_i\|_{\mathcal{U}}}, \quad i \in \mathbb{N}.$

Lemma 3.1.

(a) The system $\{\tilde{e_i}\}_{i\in\mathbb{N}}$ is the orthonormal basis in the space $(\mathcal{U}, (\cdot,\cdot)_{\mathcal{U}})$.
(b) For every $n \in \mathbb{N}$ and $x \in \mathcal{U}$

$$P_n x = \sum_{i=1}^{n}(x, \tilde{e_i})_{\mathcal{U}}\tilde{e_i},$$

i.e., the restriction of P_n to the space $\mathcal{U}$ is the $(\cdot,\cdot)_{\mathcal{U}}$-projection onto $\mathbb{H}_n$.
(c) For every $x \in \mathcal{U}$
1. $\lim_{n\to\infty}\|P_n x - x\|_{\mathcal{U}} = 0$.
2. $\lim_{n\to\infty}|P_n x - x|_{V_1} = 0$.
3. $\lim_{n\to\infty}\|P_n x - x\| = 0$.

For proof see Brzeźniak and Motyl [15].

For every $n \in \mathbb{N}$, we consider the finite-dimensional system of stochastic differential equations on H_n given by

$$\mathrm{d}X_n(t) = P_n\tilde{F}(X_n(t))\mathrm{d}t + \sigma_n(t, X_n(t))dW_n(t) + \int_Z g^n(X_n(t-), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \tag{3.3}$$

$$X_n(0) = P_n X(0) \tag{3.4}$$

where the nonlinear operator $\tilde{F}$ on V is given by $\tilde{F}(X(t)) = -AX(t) - B(X(t)) - \int_0^t \beta(t-\tau)AX(\tau)\mathrm{d}\tau$ and $W_n = P_n W$, $\sigma_n = P_n\sigma$, $g^n = P_n g$. We denote $P_n B(\cdot) = B_n(\cdot)$.

Remark 3.2. Since P_n is a contraction of V', taking account of properties of B in Section 2.1, we assert that $P_n\tilde{F}(\cdot)$ is locally Lipschitz. We note that Hypothesis (2.4) ensures that σ_n, $\int_Z g^n(\cdot, z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) = P_n\int_Z g(., z)\tilde{N}(\mathrm{d}t, \mathrm{d}z)$ are globally Lipschitz. Hence, we infer from Theorem 3.1 of Albeverio *et al.* [4] that for all $n \geq 1$, there exists an adapted process $X_n \in D([0,T]; H_n)$ a.s. such that X_n satisfies (3.3)–(3.4).

Theorem 3.3. *Under the above mathematical setting, let $X(0)$ be $\mathscr{F}_0-$measurable such that $\mathbb{E}|X(0)|^2 < \infty$, $\sigma_n \in C([0,T]\times H; \mathcal{L}_Q(H,H))$, $g^n \in \mathfrak{L}^2_{\lambda,T}(\mathcal{P}\otimes\mathcal{B}(Z), l\otimes\mathbb{P}\otimes\lambda; H)$. Let $X_n \in D([0,T]; H_n)$ a.s. be such that X_n satisfies the finite system of equations (3.3)–(3.4) for all $n \geq 1$. Then with K_2 as in Hypothesis 2.4 H.2 the following estimates hold for all $0 \leq t \leq T$ and $\forall\, n \geq 1$,*

$$\mathbb{E}\left[|X_n(t)|^2 + 2\int_0^t \|X_n(s)\|^2 ds\right] \leq C(\mathbb{E}|X(0)|^2, K_2, T), \tag{3.5}$$

$$\mathbb{E}\left(\sup_{0\le t\le T}|X_n(t)|^2\right)+4\mathbb{E}\left(\int_0^T\|X_n(t)\|^2\,\mathrm{d}t\right)\le C(\mathbb{E}|X(0)|^2,K_2,K_4,T). \tag{3.6}$$

Proof. Let $n\in\mathbb{N}$ be fixed. Let us define

$$\tau_N=\inf\{t:|X_n(t)|^2\ +\ 2\int_0^t\|X_n(s)\|^2ds>N\}$$

as the stopping time. Let us take the function $f(x)=|x|^2$ and apply the Itô's formula (see Appendix of [14]) to the process $X_n(t)$ to obtain

$$\begin{aligned}d|X_n(t)|^2+2\|X_n(t)\|^2\mathrm{d}t&\le -2\left(\int_0^t\beta(t-\tau)(A^{\frac12}X_n(\tau),A^{\frac12}X_n(t))_Hd\tau\right)\mathrm{d}t+I_1\\&+Tr(\sigma_n(t,X_n(t))Q\sigma_n^*(t,X_n(t)))\mathrm{d}t+2(\sigma_n(t,X_n(t))dW_n(t),X_n(t))_H.\end{aligned}$$

where

$$\begin{aligned}I_1:=&\int_Z\left[|X_n(t-)+g^n(X_n(t-),z)|^2-|X_n(t-)|^2\right]\tilde N(\mathrm{d}t,\mathrm{d}z)\\&+\int_Z\left[|X_n(t)+g^n(X_n(t),z)|^2-|X_n(t)|^2\right]\lambda(\mathrm{d}z)\mathrm{d}t\\&-2\int_Z\left[(g^n(X_n(t),z),X_n(t))_H\right]\lambda(\mathrm{d}z)\mathrm{d}t.\\=&\int_Z\left[|X_n(t-)+g^n(X_n(t-),z)|^2-|X_n(t-)|^2\right]\tilde N(\mathrm{d}t,\mathrm{d}z)+\int_Z|g^n(X_n(t),z)|^2\lambda(\mathrm{d}z)\mathrm{d}t.\end{aligned}$$

Integrating from 0 to $t\wedge\tau_N$, we have

$$\begin{aligned}|X_n(t\wedge\tau_N)|^2&+2\int_0^{t\wedge\tau_N}\|X_n(s)\|^2ds\le|X(0)|^2+2M_1(t\wedge\tau_N)+M_2(t\wedge\tau_N)\\&-2\int_0^{t\wedge\tau_N}\left(\int_0^s\beta(s-\tau)(A^{\frac12}X_n(\tau),A^{\frac12}X_n(s))_Hd\tau\right)\mathrm{d}s+\int_0^{t\wedge\tau_n}|\sigma_n(s,X_n(s)|^2_{\mathcal{L}_Q}\mathrm{d}s\\&+\int_0^{t\wedge\tau_n}\int_Z|g^n(X_n(s),z)|^2\lambda(\mathrm{d}z)\mathrm{d}s,\end{aligned}\tag{3.7}$$

$$\begin{aligned}\text{where}\quad M_1(t)&:=\int_0^t(\sigma_n(s,X_n(s))dW_n(s),X_n(s))_H\\ \text{and}\quad M_2(t)&:=\int_0^t\int_Z\left[|X_n(s-)+g^n(X_n(s-),z)|^2-|X_n(s-)|^2\right]\tilde N(\mathrm{d}s,\mathrm{d}z).\end{aligned}$$

We note from the properties of P_n that $|\sigma_n(s,X_n(s))|_{\mathcal{L}_Q}\le|\sigma(s,X_n(s))|_{\mathcal{L}_Q}$ and $|g^n(X_n(s),z)|\le|g(X_n(s),z)|$ for each $n\in\mathbb{N}$. Hence using Lemma 2.3, and Hypothesis 2.4 H.2 we have

$$\begin{aligned}|X_n(t\wedge\tau_N)|^2+2\int_0^{t\wedge\tau_N}\|X_n(s)\|^2ds&\le|X(0)|^2+2M_1(t\wedge\tau_N)\\&+M_2(t\wedge\tau_N)+K_2T+(K_2+1)\int_0^{t\wedge\tau_N}|X_n(s)|^2ds.\end{aligned}\tag{3.8}$$

Now, M_1, M_2 are martingales having zero averages. Hence, we have

$$\mathbb{E}\left[|X_n(t\wedge\tau_N)|^2+2\int_0^{t\wedge\tau_N}\|X_n(s)\|^2ds\right]\leq \mathbb{E}|X(0)|^2+(K_2+1)\mathbb{E}\left[\int_0^{t\wedge\tau_N}|X_n(s)|^2\mathrm{d}s\right]+K_2T$$
$$\leq \mathbb{E}|X(0)|^2+(K_2+1)\int_0^t\left[\mathbb{E}|X_n(s\wedge\tau_N)|^2\right]\mathrm{d}s+K_2T. \tag{3.9}$$

Application of Gronwall's inequality produces

$$\mathbb{E}\left[|X_n(t\wedge\tau_N)|^2+2\int_0^{t\wedge\tau_N}\|X_n(s)\|^2ds\right]\leq\left(\mathbb{E}|X(0)|^2+K_2T\right)e^{(K_2+1)T}.$$

Note as $N\to\infty, t\wedge\tau_N\to t$ a.s. Thus employing the above estimate we achieve

$$\mathbb{E}\left[|X_n(t)|^2+2\int_0^t\|X_n(s)\|^2ds\right]\leq C(\mathbb{E}|X(0)|^2,K_2,T). \tag{3.10}$$

Thus, inequality (3.5) is established.
Let us now take supremum from $0\leq t\leq T\wedge\tau_N$, followed by expectation on the inequality (3.8) to get

$$\mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_N}|X_n(t)|^2+2\int_0^{T\wedge\tau_N}\|X_n(t)\|^2\mathrm{d}t\right]\leq\mathbb{E}|X(0)|^2+2\mathbb{E}\left(\sup_{0\leq t\leq T\wedge\tau_N}M_1(t)\right)$$
$$+(K_2+1)\mathbb{E}\left[\int_0^{T\wedge\tau_N}\sup_{0\leq s\leq t}|X_n(s)|^2\mathrm{d}t\right]+K_2T+\mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_N}M_2(t)\right]. \tag{3.11}$$

Now due to Burkholder–Davis–Gundy inequality, $\|\sigma_n(s,X_n(s))\|_{\mathcal{L}_Q}\leq\|\sigma(s,X_n(s))\|_{\mathcal{L}_Q}$, Hypothesis 2.4 and Young's inequality, we have

$$2\mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_N}|M_1(t)|\right]\leq C\,\mathbb{E}\left[\left(\int_0^{T\wedge\tau_N}\|\sigma_n(s,X_n(s))\|^2_{\mathcal{L}_Q}|X_n(s)|^2\mathrm{d}s\right)\right]^{1/2}$$
$$\leq C\,K_2\mathbb{E}\left[\left(\int_0^{T\wedge\tau_N}(1+|X_n(s)|^2)|X_n(s)|^2\mathrm{d}s\right)\right]^{1/2}$$
$$\leq C\,K_2\mathbb{E}\left[\left(\sup_{0\leq t\leq T\wedge\tau_N}|X_n(s)|^2\right)\left(\int_0^{T\wedge\tau_N}(1+|X_n(s)|^2)\mathrm{d}s\right)\right]^{1/2}$$
$$\leq\frac{1}{4}\mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_N}|X_n(s)|^2\right]+(C\,K_2)^2\mathbb{E}\left[\int_0^{T\wedge\tau_N}|X_n(s)|^2\mathrm{d}s\right]+(CK_2)^2T. \tag{3.12}$$

Applying Burkholder–Davis–Gundy inequality for exponent $p=1$, $|g^n(X_n(s),z)|\leq|g(X_n(s),z)|$ and Hypothesis 2.4, we have

$$\mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_N}|M_2(t)|\right]$$
$$=\mathbb{E}\Big[\sup_{0\leq t\leq T\wedge\tau_N}\Big|\int_0^t\int_Z\Big[2(X_n(s-),g^n(X_n(s-),z))+|g^n(X_n(s-),z)|^2\Big]\tilde{N}(\mathrm{d}s,\mathrm{d}z)\Big|\Big]$$

$$
\begin{aligned}
&\leq C_1\mathbb{E}\Big[\int_0^{T\wedge\tau_N}\int_Z\Big[2|(X_n(s)||g^n(X_n(s),z)|+|g^n(X_n(s),z)|^2\Big]^2\lambda(\mathrm{d}z)\mathrm{d}s\Big]^{\frac{1}{2}}\\
&\leq \sqrt{2}C_1\mathbb{E}\Big[\int_0^{T\wedge\tau_N}\int_Z\Big[4|X_n(s)|^2|g(X_n(s),z)|^2+|g(X_n(s),z)|^4\Big]\lambda(\mathrm{d}z)\mathrm{d}s\Big]^{\frac{1}{2}}\\
&\leq \sqrt{2}C_1\mathbb{E}\left[\int_0^{T\wedge\tau_N}\left[4K_2|X_n(s)|^2\left(1+|X_n(s)|^2\right)+K_4\left(1+|X_n(s)|^4\right)\right]\mathrm{d}s\right]^{\frac{1}{2}}\\
&\leq \sqrt{2}C_1\mathbb{E}\left[(K_4T)^{\frac{1}{2}}+\left(4K_2\int_0^{T\wedge\tau_N}|X_n(s)|^2\mathrm{d}s\right)^{\frac{1}{2}}\right]\\
&\quad+\sqrt{2}C_1\mathbb{E}\left[(4K_2+K_4)^{\frac{1}{2}}\left(\sup_{0\leq t\leq T\wedge\tau_N}|X_n(t)|^2\right)^{\frac{1}{2}}\left(\int_0^{T\wedge\tau_N}|X_n(s)|^2\mathrm{d}s\right)^{\frac{1}{2}}\right]\\
&\leq \sqrt{2}C_1\mathbb{E}\left[\frac{1}{2}+\frac{K_4T}{2}+\frac{1}{2}+2K_2\int_0^{T\wedge\tau_N}|X_n(s)|^2\mathrm{d}s\right]\\
&\quad+\sqrt{2}C_1\mathbb{E}\left[(4K_2+K_4)^{\frac{1}{2}}\left(\sup_{0\leq t\leq T\wedge\tau_N}|X_n(t)|^2\right)^{\frac{1}{2}}\left(\int_0^{T\wedge\tau_N}|X_n(s)|^2\mathrm{d}s\right)^{\frac{1}{2}}\right]\\
&\leq \sqrt{2}C_1\mathbb{E}\left[1+\frac{K_4T}{2}+\epsilon_1\left(4K_2+K_4\right)^{\frac{1}{2}}\left(\sup_{0\leq t\leq T\wedge\tau_N}|X_n(t)|^2\right)\right]\\
&\quad+\sqrt{2}C_1\mathbb{E}\left[\left(2K_2+\frac{1}{4\epsilon_1}\left(4K_2+K_4\right)^{\frac{1}{2}}\right)\left(\int_0^{T\wedge\tau_N}|X_n(s)|^2\mathrm{d}s\right)\right]\\
&\leq \frac{1}{4}\mathbb{E}\left(\sup_{0\leq t\leq T\wedge\tau_N}|X_n(t)|^2\right)+\sqrt{2}C_1\Big[1+\frac{K_4T}{2}+\left(2K_2+C_1\sqrt{2}\left(4K_2+K_4\right)\right)\\
&\qquad\times\mathbb{E}\left(\int_0^{T\wedge\tau_N}\sup_{0\leq s\leq t}|X_n(s)|^2\mathrm{d}t\right)\Big]
\end{aligned}\tag{3.13}
$$

where $\epsilon_1=\frac{1}{4\sqrt{2}C_1(4K_2+K_4)^{\frac{1}{2}}}$.
Using (3.12) and (3.13) from (3.11) we have

$$
\begin{aligned}
&\frac{1}{2}\mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_N}|X_n(t)|^2+2\int_0^{T\wedge\tau_N}\|X_n(t)\|^2\mathrm{d}t\right]\\
&\quad\leq \mathbb{E}|X(0)|^2+C(K_2,K_4)\left(T+1+\mathbb{E}\left[\int_0^{T\wedge\tau_N}\sup_{0\leq s\leq t}|X_n(s)|^2\mathrm{d}t\right]\right).
\end{aligned}\tag{3.14}
$$

Applying Gronwall's inequality we infer

$$
\mathbb{E}\left[\sup_{0\leq t\leq T\wedge\tau_N}|X_n(t)|^2+4\int_0^{T\wedge\tau_N}\|X_n(t)\|^2\mathrm{d}t\right]\leq\left(2\mathbb{E}|X(0)|^2+C(K_2,K_4)(T+1)\right)e^{C(K_2,K_4)T}.\tag{3.15}
$$

Letting $N\to\infty, T\wedge\tau_N\to T$ a.s. we have (3.6). □

Theorem 3.4. *Under the above mathematical setting, let* $p \geq 2, X(0)$ *be* $\mathscr{F}_0-$*measurable such that* $\mathbb{E}|X(0)|^p < \infty$, $\sigma_n \in C([0,T]\times H; \mathcal{L}_Q(H,H))$, $g^n \in \mathfrak{L}^p_{\lambda,T}(\mathcal{P}\otimes\mathcal{B}(Z), l\otimes\mathbb{P}\otimes\lambda; H)$*. Let* $X_n \in D([0,T];H_n)$ *a.s. be such that* X_n *satisfies finite system of equations* (3.3)–(3.4) $\forall\, n \geq 1$*. Then with Assumption 2.7 and with* K_p *as in Hypothesis 2.4 H.2, the following estimates hold for all* $0 \leq t \leq T$ *and* $\forall\, n \geq 1$,

$$\mathbb{E}\left[|X_n(t)|^p + p\int_0^t \|X_n(s)\|^2 |X_n(s)|^{p-2}\mathrm{d}t\right] \leq C(\mathbb{E}|X(0)|^p, K_p, T), \tag{3.16}$$

$$\mathbb{E}\left(\sup_{0\leq t\leq T}|X_n(t)|^p\right) + 2p\mathbb{E}\left(\int_0^T \|X_n(t)\|^2|X_n(t)|^{p-2}\,\mathrm{d}t\right) \leq C(\mathbb{E}|X(0)|^p, K_p, C_F, T). \tag{3.17}$$

Proof. Applying Itô's formula (see Appendix of [14]) to the process $|X_n(t)|^p$ for $p \geq 2$, the proof follows similarly as in Theorem 3.3. We omit the details for the sake of brevity. □

4. Existence of martingale solution for the uncontrolled equation

4.1. Statement of the martingale problem

We first introduce the following functional spaces endowed with the respective topologies:

$$D([0,T];\mathcal{U}')_J := \text{ the space of càdlàg functions } X:[0,T]\to\mathcal{U}'$$

with the extended Skorokhod topology $\mathcal{T}_1$,

$$\mathrm{L}^2_w(0,T;V) := \text{ the space } \mathrm{L}^2(0,T;V) \text{ with the weak topology } \mathcal{T}_2,$$

$$H_w := \text{ the Hilbert space } H \text{ endowed with the weak topology},$$

$$D([0,T];H_w) := \text{ the the space of all weakly càdlàg functions } X:[0,T]\to H$$

with the weakest topology $\mathcal{T}_3$ such that for all $h \in H$ the mappings

$$D([0,T];H_w) \ni X \mapsto (X(\cdot),h)_H \in D([0,T];\mathbb{R})$$

are continuous. In particular, $X_n \to X$ in $D([0,T];H_w)$ iff for all $h\in H$

$$(X_n(\cdot),h)_H \to (X(\cdot),h)_H \text{ in } D([0,T];\mathbb{R}).$$

$$\mathrm{L}^2(0,T;H) \text{ is endowed with its strong topology } \mathcal{T}_4.$$

Let $\mathcal{Z} = D([0,T];\mathcal{U}')_J \cap D([0,T];H_w) \cap \mathrm{L}^2_w(0,T;V) \cap \mathrm{L}^2(0,T;H)$ and let $\mathscr{T}$ be the supremum of the corresponding topologies.

Definition 4.1. A martingale solution of (3.1)–(3.2) is a system $(\bar\Omega, \bar{\mathscr{F}}, \bar{\mathscr{F}}_t, \bar{\mathbb{P}}, \{\bar X(t)\}_{t\geq 0}, \{\bar N(t,\cdot)\}_{t\geq 0}, \{\bar W(t)\}_{t\geq 0},)$ where

(i) $(\bar\Omega, \bar{\mathscr{F}}, \bar{\mathscr{F}}_t, \bar{\mathbb{P}})$ is a filtered probability space with a filtration $\{\bar{\mathscr{F}}_t\}_{t\geq 0}$,
(ii) $\bar N$ is a time homogeneous Poisson random measure over $(\bar\Omega, \bar{\mathscr{F}}, \bar{\mathscr{F}}_t, \bar{\mathbb{P}})$ with the intensity measure λ,

(iii) $\bar{W}$ is a cylindrical Wiener process over $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathscr{F}}_t, \bar{\mathbb{P}})$,
(iv) $\bar{X} : [0,T] \times \bar{\Omega} \to H$ is a predictable process with $\bar{\mathbb{P}}$ − a.e. paths

$$\bar{X}(\cdot,\omega) \in D([0,T];H_w) \cap \mathrm{L}^2_w(0,T;V) \cap \mathrm{L}^2(0,T;H) \cap D([0,T];\mathcal{U}')$$

such that for all $t \in [0,T]$ and all $\varphi \in \mathcal{U}$ the following identity holds $\bar{\mathbb{P}}$ − a.s.

$$\begin{aligned}(\bar{X}(t),\varphi) &+ \int_0^t (A\bar{X}(s),\varphi)\mathrm{d}s + \int_0^t (B(\bar{X}(s)),\varphi)\mathrm{d}s + \int_0^t\int_0^s \beta(s-r)(A\bar{X}(r),\varphi)\mathrm{d}r\,\mathrm{d}s \\ &= (\bar{X}(0),\varphi) + \int_0^t (\sigma(s,\bar{X}(s))d\bar{W}(s),\varphi) + \int_0^t\int_Z (g(\bar{X}(s-),z),\varphi)\tilde{\bar{N}}(\mathrm{d}s,\mathrm{d}z). \qquad (4.1)\end{aligned}$$

4.2. Tightness of the laws of approximating sequence

We consider the classical Faedo–Galerkin approximation (3.3)–(3.4) which generates a sequence of probability measures on suitable functional space. Taking into account Theorem 3.4, we directly ensure the validity of the following lemma.

Lemma 4.2. *The processes $(X_n)_{n\in\mathbb{N}}$ satisfy the following estimates.*

(a) For every $p \geq 2$, there exists a positive constant $C_1(p)$ such that

$$\sup_{n\in\mathbb{N}} \mathbb{E}[\sup_{s\in[0,T]} |X_n(s)|^p] \leq C_1(p), \qquad (4.2)$$

(b) there exists a positive constant C_2 such that

$$\sup_{n\in\mathbb{N}} \mathbb{E}[\int_0^T \|X_n(s)\|^2 \mathrm{d}s] \leq C_2. \qquad (4.3)$$

Lemma 4.3. *The set of measures $\{\mathcal{L}(X_n), n \in \mathbb{N}\}$ is tight on $(\mathcal{Z}, \mathscr{T})$.*

Proof. Taking into account Lemma 4.2, we directly ensure *(a)* and *(b)* of Theorem A.11. Hence, it suffices to prove that the sequence X_n satisfies the Aldous condition $[A]$ in the space $\mathcal{U}'$. Let $\theta > 0$. Let $(\tau_n)_{n\in\mathbb{N}}$ be a sequence of stopping times where $0 \leq \tau_n < \tau_n + \theta \leq T$. Then using Galerkin approximation we have

$$\begin{aligned}X_n(t) &= X_n(0) - \int_0^t AX_n(s)\mathrm{d}s - \int_0^t B_n(X_n(s))\mathrm{d}s - \int_0^t\int_0^s \beta(s-r)AX_n(r)\mathrm{d}r\mathrm{d}s \\ &\quad + \int_0^t (\sigma_n(s,X_n(s))dW_n(s) + \int_0^t\int_Z g^n(X_n(s-),z)\tilde{N}(\mathrm{d}s,\mathrm{d}z) \\ &:= J_1^n + J_2^n(t) + J_3^n(t) + J_4^n(t) + J_5^n(t) + J_6^n(t), \qquad t \in [0,T]. \qquad (4.4)\end{aligned}$$

We analyze each of these $J_i^n,\ i \in \{1,2,\dots,6\}$ such that they satisfy condition (A.4) in Theorem A.9 for suitable choices of α and ζ. J_1^n being independent of time clearly satisfies (A.4).
Now exploiting the fact $A : V \to V'$ given by $\|AX\|_{V'} \leq \|X\|$, the embedding $V' \hookrightarrow \mathcal{U}'$ is continuous, then by Hölder's inequality and (4.3), J_2^n can be estimated as

$$\mathbb{E}\Big[\|J_2^n(\tau_n+\theta) - J_2^n(\tau_n)\|_{\mathcal{U}'}\Big] = \mathbb{E}\Big[\|\int_{\tau_n}^{\tau_n+\theta} AX_n(s)\mathrm{d}s\|_{\mathcal{U}'}\Big] \leq C\mathbb{E}\Big[\|\int_{\tau_n}^{\tau_n+\theta} AX_n(s)\mathrm{d}s\|_{V'}\Big]$$

$$\leq C_1\mathbb{E}\Big[\int_{\tau_n}^{\tau_n+\theta}\|X_n(s)\|\mathrm{d}s\Big]\leq C_1\mathbb{E}\Big[\theta^{1/2}\left(\int_0^T\|X_n(s)\|^2\mathrm{d}s\right)^{1/2}\Big]\leq c_2\theta^{1/2}.$$

Thus, J_2^n satisfies (A.4) with $\alpha=1$ and $\zeta=\frac{1}{2}$.
Using (2.4) and the embedding $V_1'\hookrightarrow\mathcal{U}'$ is continuous, then by using (4.2) we estimate J_3^n as

$$\mathbb{E}\Big[\|J_3^n(\tau_n+\theta)-J_3^n(\tau_n)\|_{\mathcal{U}'}\Big]\leq C\mathbb{E}\Big[\int_{\tau_n}^{\tau_n+\theta}|B(X_n(s))|_{V_1'}\mathrm{d}s\Big]$$
$$\leq C\mathbb{E}\Big[\int_{\tau_n}^{\tau_n+\theta}|X_n(s)|^2\mathrm{d}s\Big]\leq C\mathbb{E}\Big[\sup_{0\leq s\leq T}|X_n(s)|^2\mathrm{d}s\Big]\cdot\theta=:c_3\theta.$$

Thus, J_3^n satisfies (A.4) with $\alpha=1$ and $\zeta=1$.
In order to estimate J_4^n, we follow J_2^n and further using Assumption 2.2, (4.3), we achieve

$$\begin{aligned}&\mathbb{E}\Big[\|J_4^n(\tau_n+\theta)-J_4^n(\tau_n)\|_{\mathcal{U}'}\Big]\\&\leq\mathbb{E}\Big[\int_{\tau=\tau_n}^{\tau_n+\theta}\Big(\int_{s=0}^{\tau}|\beta(\tau-s)|^2\mathrm{d}s\Big)^{1/2}\Big(\int_{s=0}^{\tau}\|X_n(s)\|^2\mathrm{d}s\Big)^{1/2}\mathrm{d}\tau\Big]\\&\leq\mathbb{E}\Big[\int_{\tau=\tau_n}^{\tau_n+\theta}\Big(\int_{s=0}^{\tau}|\beta(\tau-s)|^2\mathrm{d}s\Big)\mathrm{d}\tau\Big]+\mathbb{E}\Big[\int_{\tau=\tau_n}^{\tau_n+\theta}\Big(\int_{s=0}^{\tau}\|X_n(s)\|^2\mathrm{d}s\Big)\mathrm{d}\tau\Big]\leq c_4\theta.\end{aligned}$$

Thus, J_4^n satisfies (A.4) with $\alpha=1$ and $\zeta=\frac{1}{2}$. Exploiting $\|\sigma_n(s,X_n(s))\|_{\mathcal{L}_Q}\leq\|\sigma(s,X_n(s))\|_{\mathcal{L}_Q}$, Hypothesis 2.4 H.2, (4.2) and Itô isometry,

$$\mathbb{E}\Big[\|J_5^n(\tau_n+\theta)-J_5^n(\tau_n)\|_{\mathcal{U}'}^2\Big]\leq C\mathbb{E}\Big[\int_{\tau_n}^{\tau_n+\theta}\|\sigma(s,X_n(s))\|_{\mathcal{L}_Q}^2\mathrm{d}s\Big]$$
$$\leq C\,K_2\mathbb{E}\Big[\int_{\tau_n}^{\tau_n+\theta}(1+|X_n(s)|^2)\mathrm{d}s\Big]\leq C\,K_2\theta\Big(1+\mathbb{E}\Big[\sup_{0\leq s\leq T}|X_n(s)|^2\Big]\Big)\leq c_6\theta.$$

Thus, J_5^n satisfies (A.4) with $\alpha=2$ and $\zeta=1$.
Using $|g^n(X_n(s),z)|\leq|g(X_n(s),z)|$, Hypothesis 2.4 H.2 and by Itô–Lévy isometry, we get

$$\mathbb{E}[\|J_6^n(\tau_n+\theta)-J_6^n(\tau_n)\|_{\mathcal{U}'}^2]\leq C\mathbb{E}[|\int_{\tau_n}^{\tau_n+\theta}\int_Z g^n(X_n(s-),z)\tilde{N}(\mathrm{d}s,\mathrm{d}z)|^2]$$
$$\leq C\mathbb{E}[\int_{\tau_n}^{\tau_n+\theta}\int_Z|g(X_n(s),z)|^2\lambda(\mathrm{d}z)\mathrm{d}s]\leq C\,K_2\mathbb{E}[\int_{\tau_n}^{\tau_n+\theta}(1+|X_n(s)|^2)\mathrm{d}s]\leq c_6\theta.$$

Thus, J_6^n satisfies (A.4) with $\alpha=2$ and $\zeta=1$.
Thus, Lemma A.9 ensures the Aldous condition $[A]$ is satisfied by the sequence $(X_n)_{n\in\mathbb{N}}$ in the space $\mathcal{U}'$. Thus the proof is complete. □

4.3. Construction of new probability space and processes

Lemma 4.3 guaranties the tightness of $\{\mathcal{L}(X_n),n\in\mathbb{N}\}$ on the space $\mathcal{Z}$. Let $N_n:=N,n\in\mathbb{N}$. The set of measures $\{\mathcal{L}(N_n),n\in\mathbb{N}\}$ is tight on the space $M_{\bar{\mathbb{N}}}([0,T]\times Z)$, where $\bar{\mathbb{N}}:=\mathbb{N}\cup\{\infty\}$ and $M_{\bar{\mathbb{N}}}(S)$ denotes the set of all $\bar{\mathbb{N}}$-valued measures on the measurable space $(S,\mathscr{B}(S))$. Let $W_n:=W,n\in\mathbb{N}$. The set of measures $\{\mathcal{L}(W_n),n\in\mathbb{N}\}$ is tight on the space $C([0,T];\mathbb{R})$ of continuous function from $[0,T]$ to $\mathbb{R}$ with standard supremum

norm. Thus, the set $\{\mathcal{L}(X_n, N_n, W_n), n \in \mathbb{N}\}$ is tight on $\mathcal{Z} \times M_{\bar{\mathbb{N}}}([0,T] \times Z) \times C([0,T];\mathbb{R})$.
Note that the space of all weakly càdlàg functions $X : [0,T] \to H$ with the weakest topology, *i.e.* the space $D([0,T]; H_w)$ is not metrizable (see Appendix B of [54]). Thus, the space $\mathcal{Z}$ is a non-metrizable locally convex space. By the Jakubowski's version of Skorokhod representation theorem (see Prop. A.12 and Cor. A.13), there exists a subsequence $(n_k)_{k\in\mathbb{N}}$, a probability space $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathbb{P}})$, and on this space random variables (X_*, N_*, W_*), $(\bar{X}_k, \bar{N}_k, \bar{W}_k)$, $k \in \mathbb{N}$ such that

- (i) $\mathcal{L}((\bar{X}_k, \bar{N}_k, \bar{W}_k) = \mathcal{L}((X_{n_k}, N_{n_k}, W_{n_k}))$ for all $k \in \mathbb{N}$,
- (ii) $(\bar{X}_k, \bar{N}_k, \bar{W}_k) \to (X_*, N_*, W_*)$ in $\mathcal{Z} \times M_{\bar{\mathbb{N}}}([0,T] \times Z) \times C([0,T];\mathbb{R})$ with probability 1 on $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathbb{P}})$ as $k \to \infty$,
- (iii) $(\bar{N}_k(\bar{\omega}), \bar{W}_k(\bar{\omega})) = (N_*(\bar{\omega}), W_*(\bar{\omega}))$ for all $\bar{\omega} \in \bar{\Omega}$.

We denote these sequences again by $((X_n, N_n, W_n))_{n\in\mathbb{N}}$ and $((\bar{X}_n, \bar{N}_n, \bar{W}_n)_{n\in\mathbb{N}}$. Using the definition of the space $\mathcal{Z}$, we have

$$\bar{X}_n \to X_* \text{ in } \mathrm{L}^2_w(0,T;V) \cap \mathrm{L}^2(0,T;H) \cap D([0,T];\mathcal{U}') \cap D([0,T];H_w), \quad \bar{\mathbb{P}} - a.s. \tag{4.5}$$

4.4. Properties of new processes and limiting processes

Since the space $\mathcal{Z}$ is a non-metrizable locally convex space, it is not a Polish space. So, the following result cannot be deduced directly from the Kuratowski theorem [44]. We adopt here a method from Lemma 4.2 in [16], and generalize for Skorokhod spaces.

Proposition 4.4. *The set $D([0,T]; H_n) \cap \mathcal{Z}$ is a Borel subset of $\mathcal{Z}$ and the corresponding embedding transforms Borel sets into Borel subsets.*

Proof. The space $D([0,T];\mathcal{U}') \cap L^2(0,T;H)$ is a Polish space. Then by Kuratowski theorem, $D([0,T];H_n)$ is a Borel subset of $D([0,T];\mathcal{U}') \cap L^2(0,T;H)$. Hence $D([0,T];H_n) \cap \mathcal{Z}$ is a Borel subset of $D([0,T];\mathcal{U}') \cap L^2(0,T;H) \cap \mathcal{Z}$, which happens to be equal to $\mathcal{Z}$. □

Since the laws of $(X_n)_{n\in\mathbb{N}}$ and $(\bar{X}_n)_{n\in\mathbb{N}}$ are same in $\mathcal{Z}$, sequence $(\bar{X}_n)_{n\in\mathbb{N}}$ satisfies the same estimates as the original sequence $(X_n)_{n\in\mathbb{N}}$. In particular, for any $p \geq 2$, we have

$$\sup_{n\in\mathbb{N}} \bar{\mathbb{E}}[\sup_{s\in[0,T]} |\bar{X}_n(s)|^p] \leq C_1(p), \tag{4.6}$$

$$\sup_{n\in\mathbb{N}} \bar{\mathbb{E}}[\int_0^T \|\bar{X}_n(s)\|^2 \mathrm{d}s] \leq C_2. \tag{4.7}$$

By (4.7), $\bar{X}_n$ is uniformly bounded in the space $\mathrm{L}^2(\bar{\Omega};\mathrm{L}^2(0,T;V))$, hence there exists a weak convergent subsequence, still denoted by $\bar{X}_n$ such that $\bar{X}_n \xrightarrow{w} X_*$ in $\mathrm{L}^2([0,T] \times \bar{\Omega}; V)$ (using (4.5)) and $X_* \in \mathrm{L}^2([0,T] \times \bar{\Omega}; V)$, *i.e.*, $\bar{\mathbb{E}}[\int_0^T \|X_*(s)\|^2 \mathrm{d}s] < \infty$. Arguing as above, inequality (4.6) with $p := 2$ we can further extract a subsequence of $\bar{X}_n$ such that $\bar{X}_n \xrightarrow{w^*} X_*$ in $\mathrm{L}^2(\bar{\Omega};\mathrm{L}^\infty(0,T;H))$ (using (4.5)) and $\bar{\mathbb{E}}[\sup_{s\in[0,T]} |X_*(s)|^2] < \infty$.

Proposition 4.5. *Let X_* be the limiting process defined above. Then for $r \geq 2$,*

$$\bar{\mathbb{E}}\Big[\sup_{s\in[0,T]} |X_*(s)|^r\Big] \leq C_r.$$

Proof. By (4.6) we have that $\{\bar{X}_n\}_{n\geq 1}$ is uniformly bounded in $\mathrm{L}^r(\bar{\Omega};\mathrm{L}^\infty(0,T;H))$ for $r \geq 2$. Since $\mathrm{L}^r(\bar{\Omega};\mathrm{L}^\infty(0,T;H))$ is isomorphic to the space $(\mathrm{L}^{\frac{r}{r-1}}(\bar{\Omega};\mathrm{L}^1(0,T;H)))^*$, by Banach–Alaoglu theorem, there exists a subsequence, still denoted by $(\bar{X}_n)_{n\in\mathbb{N}}$, and $Y \in \mathrm{L}^r(\bar{\Omega};\mathrm{L}^\infty(0,T;H))$ such that $\bar{X}_n \xrightarrow{w^*} Y$ in $\mathrm{L}^r(\bar{\Omega};\mathrm{L}^\infty(0,T;H))$, *i.e.*,

$$\bar{\mathbb{E}}\Big[\int_0^T (\bar{X}_n(t,\omega),\varphi(t,\omega))_H \mathrm{d}t\Big] \to \bar{\mathbb{E}}\Big[\int_0^T (Y(t,\omega),\varphi(t,\omega))_H \mathrm{d}t\Big] \quad \forall\, \varphi \in \mathrm{L}^{\frac{r}{r-1}}(\bar{\Omega};\mathrm{L}^1(0,T;H)).$$

Also as $\bar{X}_n \xrightarrow{w} X_*$ in $\mathrm{L}^2(\bar{\Omega};\mathrm{L}^2(0,T;V))$, we have

$$\bar{\mathbb{E}}\Big[\int_0^T \langle \bar{X}_n(t,\omega),\varphi(t,\omega)\rangle \mathrm{d}t\Big] \to \bar{\mathbb{E}}\Big[\int_0^T \langle X_*(t,\omega),\varphi(t,\omega)\rangle \mathrm{d}t\Big] \quad \forall\, \varphi \in \mathrm{L}^2(\bar{\Omega};\mathrm{L}^2(0,T;V')).$$

Exploiting the Gelfand triple $V \subset H \subset V'$ we therefore have,

$$\bar{\mathbb{E}}\Big[\int_0^T (\bar{X}_n(t,\omega),\varphi(t,\omega))_H \mathrm{d}t\Big] \to \bar{\mathbb{E}}\Big[\int_0^T (X_*(t,\omega),\varphi(t,\omega))_H \mathrm{d}t\Big] \quad \forall\, \varphi \in \mathrm{L}^2(\bar{\Omega};\mathrm{L}^2(0,T;H)).$$

For $r \geq 2$, $\mathrm{L}^2(\bar{\Omega};\mathrm{L}^2(0,T;H))$ is a dense subspace of $\mathrm{L}^{\frac{r}{r-1}}(\bar{\Omega};\mathrm{L}^1(0,T;H))$. Hence, we have

$$\bar{\mathbb{E}}\Big[\int_0^T (Y(t,\omega),\varphi(t,\omega))_H \mathrm{d}t\Big] = \bar{\mathbb{E}}\Big[\int_0^T (X_*(t,\omega),\varphi(t,\omega))_H \mathrm{d}t\Big] \quad \forall\, \varphi \in \mathrm{L}^2(\bar{\Omega};\mathrm{L}^2(0,T;H)).$$

Thus we have, $X_*(t,\omega) = Y(t,\omega)$ for l-almost every $t \in [0,T]$ and $\bar{\mathbb{P}}$-almost all $\omega \in \bar{\Omega}$. Since $Y \in \mathrm{L}^r(\bar{\Omega};\mathrm{L}^\infty(0,T;H))$, we infer $X_* \in \mathrm{L}^r(\bar{\Omega};\mathrm{L}^\infty(0,T;H))$, *i.e.*, $\bar{\mathbb{E}}\Big[\sup_{s\in[0,T]} |X_*(s)|^r\Big] \leq C_r$, for some constant C_r (depending on r).

□

Lemma 4.6. *For all* $\varphi \in \mathcal{U}$,

(i) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^T |(\bar{X}_n(t) - X_*(t),\varphi)_H|\mathrm{d}t\Big] = 0.$
(ii) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[|(\bar{X}_n(0) - X_*(0),\varphi)_H|^2\Big] = 0.$
(iii) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^T \Big|\int_0^t \langle A\bar{X}_n(s) - AX_*(s),\varphi\rangle \mathrm{d}s\Big|\mathrm{d}t\Big] = 0.$
(iv) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^T \Big|\int_0^t \langle B_n(\bar{X}_n(s)) - B(X_*(s)),\varphi\rangle \mathrm{d}s\Big|\mathrm{d}t\Big] = 0.$
(v) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^T \Big|(\int_0^t [\sigma_n(s,\bar{X}_n(s)) - \sigma(s,X_*(s))]dW_*(s),\varphi)_H\Big|^2 \mathrm{d}t\Big] = 0.$
(vi) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^T \Big|\int_0^t \int_Z (g^n(\bar{X}_n(s),z) - g(X_*(s),z),\varphi)_H \lambda(\mathrm{d}z)\mathrm{d}s\Big|^2 \mathrm{d}t\Big] = 0.$
(vii) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^T \Big|\int_0^t \int_Z (g^n(\bar{X}_n(s-),z) - g(X_*(s-),z),\varphi)_H \tilde{N}_*(\mathrm{d}s,\mathrm{d}z)\Big|^2 \mathrm{d}t\Big] = 0.$
(viii) $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^T \Big|\int_0^t \int_0^\tau \beta(\tau - s)\langle A\bar{X}_n(s) - AX_*(s),\varphi\rangle \mathrm{d}s\,\mathrm{d}\tau\Big|\mathrm{d}t\Big] = 0.$

Proof.

(i) Let $\varphi \in \mathcal{U}$ be fixed. Owing to (4.5) we have $\bar{X}_n \to X_*$ in $D([0,T];H_w)$, $\bar{\mathbb{P}}$ − a.s. *i.e.*, $(\bar{X}_n(t) - X_*(t),\varphi)_H \to 0$ in $D([0,T];\mathbb{R})$, $\bar{\mathbb{P}}$ − a.s. Hence, in particular for almost all $t \in [0,T]$, $\lim_{n\to\infty}(\bar{X}_n(t),\varphi)_H = (X_*(t),\varphi)_H$ $\bar{\mathbb{P}}$ − a.s. Also by (4.6), $\sup_{t\in[0,T]} |\bar{X}_n(t)|^2 < \infty$, so $\bar{\mathbb{E}}\Big[\int_0^T |(\bar{X}_n(t) -$

$X_*(t),\varphi)_H|^2 dt\Big] \leq C$ for some constant $C > 0$. Hence, by employing Vitali theorem we have

$$\lim_{n\to\infty} \|(\bar{X}_n,\varphi)_H - (X_*,\varphi)_H\|_{\mathrm{L}^1([0,T]\times\bar{\Omega})} = \lim_{n\to\infty} \bar{\mathbb{E}}[\int_0^T |(\bar{X}_n(t) - X_*(t),\varphi)_H|\mathrm{d}t] = 0.$$

(ii) By (4.5) we have $\bar{X}_n \to X_*$ in $D([0,T];H_w)$, $\bar{\mathbb{P}}$ − a.s. and X_* is right continuous at $t = 0$, we infer that $(\bar{X}_n(0),\varphi)_H \to (X_*(0),\varphi)_H$ $\bar{\mathbb{P}}$ − a.s. (4.6) and Vitali theorem gives $\lim_{n\to\infty} \bar{\mathbb{E}}[|(\bar{X}_n(0) - X_*(0),\varphi)_H|^2] = 0$. Hence $\lim_{n\to\infty} \|(\bar{X}_n(0) - X_*(0),\varphi)_H\|^2_{L^2([0,T]\times\bar{\Omega})} = 0$.

(iii) Since $\bar{X}_n \to X_*$ in $\mathrm{L}^2_w(0,T;V)$, $\bar{\mathbb{P}}$ − a.s.,(by (4.5)) so for any $\tilde{\varphi} \in \mathrm{L}^2(0,T;V)$ we have

$$\lim_{n\to\infty} \int_0^T \langle A\bar{X}_n(s) - AX_*(s), \tilde{\varphi}(s)\rangle \mathrm{d}s = 0. \tag{4.8}$$

Let $\varphi \in \mathcal{U}$. Let $t \in [0,T]$ be fixed. We choose $\tilde{\varphi}(s) = \chi_{(0,t)}(s)\varphi$ and note that $\tilde{\varphi} \in \mathrm{L}^2(0,T;V)$, indeed,

$$\int_0^T \|\tilde{\varphi}(s)\|^2 \mathrm{d}s \leq C\int_0^T \chi_{(0,t)}(s)\|\varphi\|^2_{\mathcal{U}} \mathrm{d}s \leq C\|\varphi\|^2_{\mathcal{U}} T < \infty, \quad \bar{\mathbb{P}} - \text{a.s.}$$

Hence, using (4.8) we have

$$0 = \lim_{n\to\infty} \int_0^T \langle A\bar{X}_n(s) - AX_*(s), \tilde{\varphi}(s)\rangle \mathrm{d}s = \lim_{n\to\infty} \int_0^t \langle A\bar{X}_n(s) - AX_*(s), \varphi\rangle \mathrm{d}s. \tag{4.9}$$

Hence by Hölder's inequality and (4.7) we achieve for all $t \in [0,T]$, and $n \in \mathbb{N}$,

$$\bar{\mathbb{E}}\Big[|\int_0^t \langle A\bar{X}_n(s),\varphi\rangle \mathrm{d}s|^2\Big] \leq \bar{\mathbb{E}}\Big[\int_0^t |(\bar{X}_n(s),\varphi)_V|^2 \mathrm{d}s\Big] \leq C\|\varphi\|^2_{\mathcal{U}}\,\bar{\mathbb{E}}\Big[\int_0^t \|\bar{X}_n(s)\|^2 \mathrm{d}s\Big] \leq c \tag{4.10}$$

for some constant $c > 0$. Therefore, using (4.9) and (4.10) and by Vitali theorem, we conclude for all $t \in [0,T]$ $\lim_{n\to\infty} \bar{\mathbb{E}}\Big[|\int_0^t \langle A\bar{X}_n(s) - AX_*(s),\varphi\rangle \mathrm{d}s|\Big] = 0$. Hence (4.7) and the dominated convergence theorem gives (iii). (iv) Let $\varphi \in \mathcal{U}$. We have, $B(\bar{X}_n,\bar{X}_n) - B(X_*,X_*) = B(\bar{X}_n - X_*,\bar{X}_n) + B(X_*,\bar{X}_n - X_*)$. Now, using Hölder's inequality, we obtain

$$\begin{aligned}
&|\int_0^t \langle B(\bar{X}_n(s),\bar{X}_n(s)),\varphi\rangle \mathrm{d}s - \int_0^t \langle B(X_*(s),X_*(s)),\varphi\rangle \mathrm{d}s| \\
&\leq \int_0^t \|B(\bar{X}_n(s) - X_*(s),\bar{X}_n(s))\|_{V_1'}\|\varphi\|_{V_1}\mathrm{d}s + \int_0^t \|B(X_*(s),\bar{X}_n(s) - X_*(s))\|_{V_1'}\|\varphi\|_{V_1}\mathrm{d}s \\
&\leq C\|\varphi\|_{\mathcal{U}}\Big(\int_0^T |\bar{X}_n(s) - X_*(s)||\bar{X}_n(s)|\mathrm{d}s + \int_0^T |\bar{X}_n(s) - X_*(s)||X_*(s)|\mathrm{d}s\Big) \\
&\leq C\|\bar{X}_n - X_*\|_{\mathrm{L}^2(0,T;H)}\Big(\|\bar{X}_n\|_{\mathrm{L}^2(0,T;H)} + \|X_*\|_{\mathrm{L}^2(0,T;H)}\Big)\|\varphi\|_{\mathcal{U}},
\end{aligned}$$

where $C > 0$ is a constant. Since $\bar{X}_n \to X_*$ in $L^2(0,T;H)$, we have

$$\lim_{n\to\infty} \int_0^t \langle B(\bar{X}_n(s)) - B(X_*(s)),\varphi\rangle \mathrm{d}s = 0 \quad \bar{\mathbb{P}} - \text{a.s.} \tag{4.11}$$

For every $\varphi \in \mathcal{U}$, by Lemma 3.1 we have $P_n\varphi \to \varphi$ in V_1. Since $\mathcal{U} \subset V_1$, we conclude that for all $\varphi \in \mathcal{U}$ and all $t \in [0,T]$,

$$\lim_{n\to\infty} \int_0^t \langle B_n(\bar{X}_n(s)) - B(X_*(s)), \varphi\rangle \mathrm{d}s = 0 \quad \bar{\mathbb{P}} - \text{a.s.} \tag{4.12}$$

Using Hölder's inequality, (2.4) and (4.6) we obtain for all $t \in [0,T]$, $r > 1$ and $n \in \mathbb{N}$,

$$\begin{aligned}\bar{\mathbb{E}}\Big[\Big|\int_0^t \langle B_n(\bar{X}_n(s)), \varphi\rangle \mathrm{d}s\Big|^r\Big] &\le \bar{\mathbb{E}}\Big[\Big(\int_0^t \|B_n(\bar{X}_n(s))\|_{V_1'} \|\varphi\|_{V_1} \mathrm{d}s\Big)^r\Big] \\ \le C\,\|\varphi\|_{\mathcal{U}}^r\, t^{r-1}\, \bar{\mathbb{E}}\Big[\int_0^t |\bar{X}_n(s)|^{2r} \mathrm{d}s\Big] &\le C\,\|\varphi\|_{\mathcal{U}}^r\, t^r\, \bar{\mathbb{E}}\Big[\sup_{s\in[0,T]} |\bar{X}_n(s)|^{2r}\Big] \le CC_1(r)\end{aligned} \tag{4.13}$$

for some constant $C > 0$. Considering (4.12) and (4.13) and by the Vitali theorem we achieve for all $t \in [0,T]$

$$\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\Big|\int_0^t \langle B_n(\bar{X}_n(s)) - B(X_*(s)), \varphi\rangle \mathrm{d}s\Big|\Big] = 0. \tag{4.14}$$

Hence, in view of (4.13), (4.14) and the dominated convergence theorem, we infer that

$$\lim_{n\to\infty} \int_0^T \bar{\mathbb{E}}\Big[\Big|\int_0^t \langle B_n(\bar{X}_n(s)) - B(X_*(s)), \varphi\rangle \mathrm{d}s\Big|\Big] = 0. \tag{4.15}$$

Hence we have (iv).

(v) Let $\varphi \in \mathcal{U}$. Then using Hypothesis 2.4 and that $\bar{X}_n \to X_*$ in $\mathrm{L}^2(0,T;H)$, $\bar{\mathbb{P}}$ – a.s.,

$$\begin{aligned}&\lim_{n\to\infty} \int_0^t \|(\sigma(s,\bar{X}_n(s)) - \sigma(s,X_*(s)), \varphi)_H\|_{\mathcal{L}_Q}^2 \mathrm{d}s \le C\,L\|\varphi\|_{\mathcal{U}}^2 \lim_{n\to\infty} \int_0^t |\bar{X}_n(s) - X_*(s)|^2 \mathrm{d}s \\ &= 0.\end{aligned} \tag{4.16}$$

Using Hypothesis 2.4, (4.6) and Proposition 4.5, we observe that for every $t \in [0,T]$ and $r > 1$ and for every $n \in \mathbb{N}$,

$$\begin{aligned}&\bar{\mathbb{E}}\Big[\Big|\int_0^t \|(\sigma(s,\bar{X}_n(s)) - \sigma(s,X_*(s)), \varphi)_H\|_{\mathcal{L}_Q}^2 \mathrm{d}s\Big|^r\Big] \\ &\le |\varphi|^{2r}\, T^{r-1}\, 2^{r-1}\, \bar{\mathbb{E}}\Big[\int_0^t \Big(\|\sigma(s,\bar{X}_n(s))\|_{\mathcal{L}_Q}^{2r} + \|\sigma(s,X_*(s))\|_{\mathcal{L}_Q}^{2r}\Big)\mathrm{d}s\Big] \\ &\le CK_2\, T^r\, 2^{r-1}\, \|\varphi\|_{\mathcal{U}}^{2r}\, \bar{\mathbb{E}}\Big[(2 + \sup_{0\le s\le T} |\bar{X}_n(s)|^{2r} + \sup_{0\le s\le T} |X_*(s)|^{2r})\Big] \le \tilde{C}_r\end{aligned} \tag{4.17}$$

for some positive constant $\tilde{C}_r$. Thus, employing (4.16) and (4.17) and by Vitali's theorem we conclude

$$\lim_{n\to\infty} \bar{\mathbb{E}}\Big[\int_0^t \|(\sigma(s,\bar{X}_n(s)) - \sigma(s,X_*(s)), \varphi)_H\|_{\mathcal{L}_Q}^2 \mathrm{d}s\Big] = 0 \quad \forall\, \varphi \in \mathcal{U}. \tag{4.18}$$

For every $\varphi \in \mathcal{U}$ and every $s \in [0,T]$ we have

$$(\sigma_n(s,\bar{X}_n(s)) - \sigma(s,X_*(s)), \varphi)_H = (\sigma(s,\bar{X}_n(s)), P_n\varphi - \varphi)_H + (\sigma(s,\bar{X}_n(s)) - \sigma(s,X_*(s)), \varphi)_H$$

$$\leq C\|\sigma(s,\bar{X}_n(s))\|_{\mathcal{L}_Q}\|P_n\varphi-\varphi\|_{\mathcal{U}}+(\sigma(s,\bar{X}_n(s))-\sigma(s,X_*(s)),\varphi)_H.$$

Then by Hypothesis 2.4 H.2 and by (4.6) we obtain

$$\begin{aligned}\bar{\mathbb{E}}\Big[\int_0^t \|(\sigma_n(s,\bar{X}_n(s))-\sigma(s,X_*(s)),\varphi)_H\|^2_{\mathcal{L}_Q}\mathrm{d}s\Big] &\leq 2C\,K_2\|P_n\varphi-\varphi\|^2_{\mathcal{U}}\,\bar{\mathbb{E}}\Big[\int_0^T(1+|\bar{X}_n(s)|^2)\mathrm{d}s\Big]\\ &+\bar{\mathbb{E}}\Big[\int_0^t \|(\sigma(s,\bar{X}_n(s))-(\sigma(s,X_*(s)),\varphi)_H\|^2_{\mathcal{L}_Q}\mathrm{d}s\Big]\\ &\leq 2\tilde{c}\|P_n\varphi-\varphi\|^2_{\mathcal{U}}+\bar{\mathbb{E}}\Big[\int_0^t \|(\sigma(s,\bar{X}_n(s))-(\sigma(s,X_*(s)),\varphi)_H\|^2_{\mathcal{L}_Q}\mathrm{d}s\Big].\end{aligned}$$

Thus, by Lemma 3.1 and by (4.18) we conclude that

$$\lim_{n\to\infty}\bar{\mathbb{E}}\Big[\int_0^t \|(\sigma_n(s,\bar{X}_n(s))-\sigma(s,X_*(s)),\varphi)_H\|^2_{\mathcal{L}_Q}\mathrm{d}s\Big]=0\quad \forall\,\varphi\in\mathcal{U}. \tag{4.19}$$

Using the Itô isometry, (4.17) and $\bar{W}_n=W_*$,

$$\begin{aligned}&\lim_{n\to\infty}\bar{\mathbb{E}}[\|(\int_0^t[\sigma_n(s,\bar{X}_n(s))-\sigma(s,X_*(s))]dW_*(s),\varphi)_H\|^2_{\mathcal{L}_Q}]\\ &=\lim_{n\to\infty}\bar{\mathbb{E}}[\int_0^t \|(\sigma_n(s,\bar{X}_n(s))-\sigma(s,X_*(s)),\varphi)_H\|^2_{\mathcal{L}_Q}\mathrm{d}s]=0.\end{aligned} \tag{4.20}$$

Again by the Itô isometry, Hypothesis 2.4 and (4.17) for $r=2$ we have for all $t\in[0,T]$ and all $n\in\mathbb{N}$

$$\bar{\mathbb{E}}\Big[\int_0^t |([\sigma_n(s,\bar{X}_n(s))-\sigma(s,X_*(s))]dW_*(s),\varphi)_H|^2\mathrm{d}s\Big]\leq\tilde{C}_2. \tag{4.21}$$

Hence by (4.20) and (4.21) and the dominated convergence theorem, we have the assertion (v).

(vi) Let us consider $\varphi\in\mathcal{U}$. Using Hypothesis 2.4 and $\bar{X}_n\to X_*$ in $\mathrm{L}^2(0,T;H)$, $\bar{\mathbb{P}}-$a.s. we have

$$\int_0^t\int_Z\Big|(g(\bar{X}_n(s),z)-g(X_*(s),z),\varphi)_H\Big|^2\lambda(\mathrm{d}z)\mathrm{d}s\leq CL\,\|\varphi\|^2_{\mathcal{U}}\int_0^t|\bar{X}_n(s)-X_*(s)|^2\mathrm{d}s\to 0\quad\text{as}\quad n\to\infty. \tag{4.22}$$

Furthermore, using Hypothesis 2.4 and (4.6) for every $t\in[0,T]$, $r>1$ and $n\in\mathbb{N}$, we have the following inequality

$$\begin{aligned}&\bar{\mathbb{E}}\Big[\int_0^t\int_Z\Big|(g(\bar{X}_n(s),z)-g(X_*(s),z),\varphi)_H\Big|^2\lambda(\mathrm{d}z)\mathrm{d}s|^r\Big]\\ &\leq c2^rK_2^r\|\varphi\|^{2r}_{\mathcal{U}}\bar{\mathbb{E}}\Big[\Big|\int_0^t\Big(2+|\bar{X}_n(s)|^2+|X_*(s)|^2\Big)\mathrm{d}s\Big|^r\Big]\leq\tilde{C}_r,\end{aligned} \tag{4.23}$$

for some constant $\tilde{C}_r>0$. Hence by (4.22), (4.23) and by the Vitali theorem we infer that for all $t\in[0,T]$, $\forall\varphi\in\mathcal{U}$.

$$\lim_{n\to\infty}\bar{\mathbb{E}}\Big[\int_0^t\int_Z\Big|(g(\bar{X}_n(s),z)-g(X_*(s),z),\varphi)_H\Big|^2\lambda(\mathrm{d}z)\mathrm{d}s\Big]=0. \tag{4.24}$$

Since the restriction of P_n to the space H is the $(\cdot,\cdot)_H$-projection onto H_n, we conclude that

$$\lim_{n\to\infty}\bar{\mathbb{E}}\Big[\int_0^t\int_Z\Big|(g^n(\bar{X}_n(s),z)-g(X_*(s),z),\varphi)_H\Big|^2\lambda(\mathrm{d}z)\mathrm{d}s\Big]=0\quad\forall\varphi\in H. \tag{4.25}$$

Since $\mathcal{U}\subset H$, (4.25) holds for all $\varphi\in\mathcal{U}$.
Moreover, Hypotheses 2.4 and (4.6) yield the following inequality

$$\bar{\mathbb{E}}\Big[\int_0^t\int_Z\Big|(g^n(\bar{X}_n(s),z)-g(X_*(s),z),\varphi)_H\Big|^2\lambda(\mathrm{d}z)\mathrm{d}s\Big]\le\tilde{C}_2. \tag{4.26}$$

Now (4.25), (4.26) and the dominated convergence theorem assures assertion (vi).
(vii) Using Itô–Lévy isometry, and the fact that $\bar{N}_n=N_*$ we have

$$\begin{aligned}&\bar{\mathbb{E}}\Big[\Big|\int_0^t\int_Z(g^n(\bar{X}_n(s-),z)-g(X_*(s-),z),\varphi)_H\tilde{N}_*(\mathrm{d}s,\mathrm{d}z)\Big|^2\Big]\\&=\bar{\mathbb{E}}\Big[\int_0^t\int_Z\Big|(g^n(\bar{X}_n(s),z)-g(X_*(s),z),\varphi)_H\Big|^2\lambda(\mathrm{d}z)\mathrm{d}s\Big].\end{aligned} \tag{4.27}$$

Thus $\forall\,\varphi\in\mathcal{U}$, combining (4.25), (4.26), (4.27) and then exploiting dominated convergence theorem we ensure assertion (vii). Thus the dominated convergence theorem ensures assertion (vii).
(viii) Let $\varphi\in\mathcal{U}$. Since $\bar{X}_n\to X_*$ in $\mathrm{L}^2_w(0,T;V)$, $\bar{\mathbb{P}}-$a.s. so for any $\tilde{\varphi}\in\mathrm{L}^2(0,T;V)$ we have

$$\lim_{n\to\infty}\int_0^T\langle A\bar{X}_n(s)-AX_*(s),\tilde{\varphi}(s)\rangle\mathrm{d}s=0. \tag{4.28}$$

Let $t\in[0,T]$ be fixed. Let us choose $\tilde{\varphi}(s)=\chi_{(0,t)}(s)\Big(\int_s^t\beta(\tau-s)\mathrm{d}\tau\Big)\varphi$ and note that $\tilde{\varphi}\in\mathrm{L}^2(0,T;V)$, indeed, $\int_0^T\|\tilde{\varphi}(s)\|^2\mathrm{d}s\le c\|\varphi\|^2_{\mathcal{U}}\int_0^t\Big(\int_s^t|\beta(\tau-s)|\mathrm{d}\tau\Big)^2\mathrm{d}s\le C\|\varphi\|^2_{\mathcal{U}}\|\beta\|^2_{\mathrm{L}^2(0,T)}\int_0^t s\,\mathrm{d}s<\infty,\quad\bar{\mathbb{P}}-$a.s.
Hence, using (4.28) we have

$$\begin{aligned}0&=\lim_{n\to\infty}\int_0^T\langle A\bar{X}_n(s)-AX_*(s),\tilde{\varphi}(s)\rangle\mathrm{d}s\\&=\lim_{n\to\infty}\int_0^t\int_s^t\beta(\tau-s)(\bar{X}_n(s)-X_*(s),\varphi)_V\mathrm{d}\tau\mathrm{d}s\\&=\lim_{n\to\infty}\int_0^t\int_0^\tau\beta(\tau-s)(\bar{X}_n(s)-X_*(s),\varphi)_V\mathrm{d}s\mathrm{d}\tau.\end{aligned} \tag{4.29}$$

Furthermore, by Hölder's inequality and (4.7) for every $t\in[0,T]$ and $n\in\mathbb{N}$, we achieve the following inequality

$$\begin{aligned}&\bar{\mathbb{E}}\Big[\Big|\int_0^t\int_0^\tau\beta(\tau-s)\langle A\bar{X}_n(s),\varphi\rangle\mathrm{d}s\,\mathrm{d}\tau\Big|^2\Big]\le\bar{\mathbb{E}}\Big[\Big|\int_0^t\|\varphi\|\|\bar{X}_n(s)\|\int_s^t\Big|\beta(\tau-s)\Big|\mathrm{d}\tau\,\mathrm{d}s\Big|^2\Big]\\&\le C\|\varphi\|^2_{\mathcal{U}}\|\beta\|^2_{\mathrm{L}^2(0,T)}\,T\,\bar{\mathbb{E}}\Big|\int_0^t\|\bar{X}_n(s)\|\mathrm{d}s\Big|^2\le C\,C_2,\end{aligned} \tag{4.30}$$

for some constant $C > 0$. Hence using (4.29) and (4.30) and the Vitali theorem we have for all $t \in [0,T]$, $\bar{\mathbb{E}}\Big[\Big|\int_0^t\int_0^\tau \beta(\tau-s)\langle A\bar{X}_n(s) - AX_*(s), \varphi\rangle \mathrm{d}s\,\mathrm{d}\tau\Big|\Big] = 0$. Again applying dominated convergence theorem we have (viii).

□

4.5. Existence of weak martingale solution

Theorem 4.7. *Let us assume that A, B satisfy Assumption 2.1, β satisfies Assumption 2.2, and σ, g meet Hypothesis 2.4. Then there exists a martingale solution to the system (3.1)–(3.2).*

Proof.

Step I: Let us define two functionals for all $\varphi \in \mathcal{U}$,

$$\begin{aligned}\mathcal{K}_n(\bar{X}_n, \bar{N}_n, \bar{W}_n, \varphi)(t) &:= (\bar{X}_n(0), \varphi)_H - \int_0^t \langle A\bar{X}_n(s), \varphi\rangle \,\mathrm{d}s - \int_0^t \langle B(\bar{X}_n(s)), \varphi\rangle \,\mathrm{d}s \\ &- \int_0^t\int_0^s \beta(s-r)\langle A\bar{X}_n(r), \varphi\rangle \,\mathrm{d}r\,\mathrm{d}s + \int_0^t \langle \sigma_n(s, \bar{X}_n(s))\mathrm{d}\bar{W}_n(s), \varphi\rangle \\ &+ \int_0^t\int_Z \langle g^n(s, \bar{X}_n(s-), z), \varphi\rangle \,\tilde{\bar{N}}_n(\mathrm{d}s, \mathrm{d}z), \qquad t \in [0,T] \quad \text{and}\end{aligned} \tag{4.31}$$

$$\begin{aligned}\mathcal{K}(X_*, N_*, W_*, \varphi)(t) &:= (X_*(0), \varphi)_H - \int_0^t \langle AX_*(s), \varphi\rangle \,\mathrm{d}s - \int_0^t \langle B(X_*(s)), \varphi\rangle \,\mathrm{d}s \\ &- \int_0^t\int_0^s \beta(s-r)\big\langle AX_*(r), \varphi\big\rangle \,\mathrm{d}r\,\mathrm{d}s + \int_0^t \langle \sigma(s, X_*(s))\mathrm{d}W_*(s), \varphi\rangle \\ &+ \int_0^t\int_Z \langle g(s, X_*(s-), z), \varphi\rangle \,\tilde{N}_*(\mathrm{d}s, \mathrm{d}z), \qquad t \in [0,T].\end{aligned} \tag{4.32}$$

Lemma 4.8.

$$\lim_{n\to\infty} \|(\bar{X}_n(\cdot), \varphi)_H - (X_*(\cdot), \varphi)_H\|_{\mathrm{L}^1([0,T]\times\bar{\Omega})} = 0. \tag{4.33}$$

$$\lim_{n\to\infty} \|\mathcal{K}_n(\bar{X}_n, \bar{N}_n, \bar{W}_n, \varphi) - \mathcal{K}(X_*, N_*, W_*, \varphi)\|_{\mathrm{L}^1([0,T]\times\bar{\Omega})} = 0. \tag{4.34}$$

Proof. (4.33) is subsequent from Lemma 4.6 (i), since as $n \to \infty$ we have

$$\|(\bar{X}_n(\cdot), \varphi)_H - (X_*(\cdot), \varphi)_H\|_{\mathrm{L}^1([0,T]\times\bar{\Omega})} = \bar{\mathbb{E}}\Big[\int_0^T |(\bar{X}_n(t) - X_*(t), \varphi)_H| \mathrm{d}t\Big] \to 0.$$

We now propose to prove (4.34). By the Fubini theorem we note that

$$\begin{aligned}&\|\mathcal{K}_n(\bar{X}_n, \bar{N}_n, \bar{W}_n, \varphi) - \mathcal{K}(X_*, N_*, W_*, \varphi)\|_{\mathrm{L}^1([0,T]\times\bar{\Omega})} \\ &= \int_0^T \bar{\mathbb{E}}\Big[|\mathcal{K}_n(\bar{X}_n, \bar{N}_n, \bar{W}_n, \varphi)(t) - \mathcal{K}(X_*, N_*, W_*, \varphi)(t)|\Big]\mathrm{d}t.\end{aligned}$$

Lemma 4.6 ensures that each term on the right hand side of (4.31) converges to the corresponding term in (4.32). This completes the proof. □

Step II: Now, X_n being an approximate solution of the Galerkin equation, for all $t \in [0,T]$ we see $(X_n(t),\varphi)_H = \mathcal{K}_n(X_n, N_n, W_n, \varphi)(t), \quad \mathbb{P}-$ a.s. More precisely,

$$\int_0^T \mathbb{E}\Big[|(X_n(t),\varphi)_H - \mathcal{K}_n(X_n, N_n, W_n, \varphi)(t)|\Big]\mathrm{d}t = 0.$$

Now, $\mathcal{L}(X_n, N_n, W_n) = \mathcal{L}(\bar{X}_n, \bar{N}_n, \bar{W}_n)$ we directly have,

$$\int_0^T \bar{\mathbb{E}}\Big[|(\bar{X}_n(t),\varphi)_H - \mathcal{K}_n(\bar{X}_n, \bar{N}_n, \bar{W}_n, \varphi)(t)|\Big]\mathrm{d}t = 0.$$

Moreover, (4.33) and (4.34) produce

$$\int_0^T \bar{\mathbb{E}}\Big[|(X_*(t),\varphi)_H - \mathcal{K}(X_*, N_*, W_*, \varphi)(t)|\Big]\mathrm{d}t = 0.$$

Hence, for l (lebesgue measure of $\mathbb{R}$)- almost all $t \in [0,T]$ and $\bar{\mathbb{P}}$ almost all $\omega \in \bar{\Omega}$

$$(X_*(t),\varphi)_H - \mathcal{K}(X_*, N_*, W_*, \varphi)(t) = 0.$$

Hence, for $l-$almost all $t \in [0,T]$, $\varphi \in \mathcal{U}$ and $\bar{\mathbb{P}}$ almost all $\omega \in \bar{\Omega}$,

$$\begin{aligned}(X_*(t),\varphi)_H = \big(X_*(0),\varphi\big)_H - \int_0^t \langle AX_*(s),\varphi\rangle\,\mathrm{d}s - \int_0^t \langle B(X_*(s)),\varphi\rangle\,\mathrm{d}s - \int_0^t\int_0^s \beta(s-r)\langle AX_*(r),\varphi\rangle\,\mathrm{d}r\mathrm{d}s \\ + \int_0^t \langle \sigma(s, X_*(s))\mathrm{d}W_*(s),\varphi\rangle + \int_0^t\int_Z \langle g(s, X_*(s-), z),\varphi\rangle\,\tilde{N}_*(\mathrm{d}s,\mathrm{d}z). \qquad (4.35)\end{aligned}$$

Now X_* being $\mathcal{Z}$-valued random variable, X_* is an element in $D([0,T];H_w)$ *i.e.*, X_* is weakly càdlàg. This asserts that (4.35) is true for all $t \in [0,T]$ and all $\varphi \in \mathcal{U}$. Substituting $\bar{X} := X_*$, $\bar{N} := N_*$, $\bar{W} := W_*$, we conclude that the system $(\bar{\Omega}, \bar{\mathscr{F}}, \bar{\mathscr{F}}_t, \bar{\mathbb{P}}, \{\bar{X}(t)\}_{t\geq 0}, \{\bar{N}(t,\cdot)\}_{t\geq 0}, \{\bar{W}(t)\}_{t\geq 0})$ is a martingale solution of (3.1)–(3.2). ☐

5. Existence of optimal relaxed control

We now return to the original control problem (1.5)–(1.6) associated with the relaxed cost functional given by (1.7).

Definition 5.1. Let $X_0 \in H$, and let $T > 0$ be fixed. A weak admissible relaxed control (with time horizon $[0,T]$) is a system

$$\pi := (\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P}, \{X(t)\}_{t\geq 0}, \{N(t,\cdot)\}_{t\geq 0}, \{W(t)\}_{t\geq 0}, \{\mu_t\}_{t\geq 0}) \qquad (5.1)$$

such that

(i) $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ is a filtered probability space with a filtration $\{\mathscr{F}_t\}_{t\geq 0}$,
(ii) N is a time homogeneous Poisson random measure over $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ with the intensity measure λ,
(iii) W is a cylindrical Wiener process over $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$,
(iv) $\{\mu_t\}_{t\geq 0}$ is an $\mathscr{F}_t$ adapted $\mathsf{P}(\mathbb{U})$-valued relaxed control process,
(v) $X : [0,T]\times\Omega \to H$ is a predictable process with $\mathbb{P}-$ a.e. paths

$$X(\cdot,\omega) \in D([0,T];H_w) \cap \mathrm{L}^2_w(0,T;V) \cap \mathrm{L}^2(0,T;H) \cap D([0,T];\mathcal{U}')$$

such that for all $t \in [0,T]$ and all $\varphi \in \mathcal{U}$ the following identity holds $\mathbb{P}$ − a.s.

$$\begin{aligned}&(X(t),\varphi)+\int_0^t (AX(s),\varphi)\mathrm{d}s+\int_0^t (B(X(s)),\varphi)\mathrm{d}s+\int_0^t\int_0^s \beta(s-r)(AX(r),\varphi)\mathrm{d}r\,\mathrm{d}s\\&=(X(0),\varphi)+\int_0^t(\mathcal{N}\mu_s,\varphi)\mathrm{d}s+\int_0^t(\sigma(s,X(s))dW(s),\varphi)+\int_0^t\int_Z (g(X(s-),z),\varphi)\tilde{N}(\mathrm{d}s,\mathrm{d}z). \qquad (5.2)\end{aligned}$$

(vi) the mapping

$$[0,T]\times\Omega \ni (t,\omega)\mapsto \bar{F}(t,X(t,\omega),\mu_t(\omega))\in\mathbb{R}$$

belongs to $\mathrm{L}^1([0,T]\times\Omega;\mathbb{R})$ and $\phi(X(T))\in \mathrm{L}^1(\Omega;\mathbb{R})$.

The set of weak admissible relaxed controls (with time horizon $[0,T]$) will be denoted by $\mathfrak{U}^w_{ad}(X_0,T)$.

Remark 5.2. In view of Definition 4.1, from (i) to (v) of Definition 5.1, we note that π is a martingale solution of (1.5)–(1.6) with the associated relaxed cost functional given by (1.7).

In this context, under this weak formulation, the relaxed cost functional is defined as

$$\bar{J}(\pi):=\mathbb{E}\Big[\int_0^T \bar{F}(t,X(t),\mu_t)\mathrm{d}t+\phi(X(T))\Big],\quad \pi\in\mathfrak{U}^w_{ad}(X_0,T). \qquad (5.3)$$

where π has the form (5.1).

Remark 5.3. The RCP is to minimize $\bar{J}$ over $\mathfrak{U}^w_{ad}(X_0,T)$ for $X_0\in H$ and $T>0$ be fixed. Namely, we seek $\tilde{\pi}\in\mathfrak{U}^w_{ad}(X_0,T)$ such that

$$\bar{J}(\tilde{\pi})=\inf_{\pi\in\mathfrak{U}^w_{ad}(X_0,T)}\bar{J}(\pi).$$

We now state the second main result of this paper which guaranties the existence of weak ORC.

Theorem 5.4. *Let $X_0\in H$ and $\pi\in\mathfrak{U}^w_{ad}(X_0,T)$ be such that $\bar{J}(\pi)<+\infty$, where $\bar{J}$ is defined by (5.3). Let all the assumptions of Theorem 4.7, Assumptions 2.7 and 2.8 be satisfied. Then the RCP, i.e. the system (1.5)–(1.6) associated with the cost functional (1.7), admits a weak optimal relaxed control with time horizon $[0,T]$.*

Proof. From hypothesis we have $X_0\in H$ and $\bar{J}(\pi)<+\infty$. Also $\bar{J}\geq 0$ implies $\bar{J}$ is bounded below on $\mathfrak{U}^w_{ad}(X_0,T)$. Hence there exists a minimizing sequence $\{\pi_n\}_{n\geq1}$ such that $\bar{J}(\pi_n)\to\inf_{\pi\in\mathfrak{U}^w_{ad}(X_0,T)}\bar{J}(\pi)$. In other words, let us define $\pi_n=(\Omega^n,\mathscr{F}^n,\mathscr{F}^n_t,\mathbb{P}^n,\{W_n(t)\}_{t\geq0},\{\mu^n_t\}_{t\geq0},\{X_n(t)\}_{t\geq0},\{N_n(t,\cdot)\}_{t\geq0})$ be minimizing sequence of weak admissible relaxed controls, that is,

$$\lim_{n\to\infty}\bar{J}(\pi_n)=\inf_{\pi\in\mathfrak{U}^w_{ad}(X_0,T)}\bar{J}(\pi). \qquad (5.4)$$

Note that the filtration $\{\mathscr{F}^n_t\}_{t\in[0,T]}$ is defined by $\mathscr{F}^n_t:=\sigma\left\{(X_n(s),\mu^n_s):0\leq s<t\right\}, t\in[0,T]$.

Since for each $n\in\mathbb{N}$, $\pi_n\in\mathfrak{U}^w_{ad}(u_0,T)$, each π_n satisfies

$$\begin{aligned}&(X_n(t),\varphi)+\int_0^t (AX_n(s),\varphi)\mathrm{d}s+\int_0^t (B(X_n(s)),\varphi)\mathrm{d}s+\int_0^t\int_0^s \beta(s-r)(AX_n(r),\varphi)\mathrm{d}r\,\mathrm{d}s\\&=(X(0),\varphi)+\int_0^t(\mathcal{N}\mu^n_s,\varphi)\mathrm{d}s+\int_0^t(\sigma(s,X_n(s))dW(s),\varphi)+\int_0^t\int_Z (g(X_n(s-),z),\varphi)\tilde{N}(\mathrm{d}s,\mathrm{d}z). \qquad (5.5)\end{aligned}$$

This holds $\mathbb{P}^n$ − a.s. for almost all $t \in [0,T]$, for all $\varphi \in \mathcal{U}$.

Now taking into account Assumption 2.7, proceeding as in similar lines as in Theorems 3.3–3.4, and using

$$\sup_{n\in\mathbb{N}} \mathbb{E}^n[|\mu_n|^2_\kappa] := \sup_{n\in\mathbb{N}} \mathbb{E}^n\Big[\int_0^T \int_{\mathbb{U}} \kappa^2(t,U)\,\mu_n(\mathrm{d}U,\mathrm{d}t)\Big] = k < +\infty,$$

we have the following *a priori* estimates (uniformly in n)

$$\mathbb{E}^n\left[|X_n(t)|^2 + 2\int_0^t \|X_n(s)\|^2 ds\right] \le C, \quad \mathbb{E}^n\left(\sup_{0\le t\le T}|X_n(t)|^2\right) + 4\mathbb{E}^n\left(\int_0^T \|X_n(t)\|^2\,\mathrm{d}t\right) \le C, \tag{5.6}$$

where C depends on $\mathbb{E}|X(0)|^2, k, K_2, K_4, T$ and $\mathbb{E}^n$ denotes the expectation with respect to $\mathbb{P}^n$.

Due to the uniform *a priori* bounds (5.6), one can establish that $\{\mathcal{L}(X_n), n \in \mathbb{N}\}$ is tight on $(\mathcal{Z}, \mathscr{T})$ by first proving the Aldous condition (as in Lem. 4.3 and taking into account Rem. A.20 in Appendix A) and then employing Theorem A.11.

For each $n \in \mathbb{N}$, we define the random Young measure μ_n on $(\Omega^n, \mathscr{F}^n, \mathbb{P}^n)$ by the formula:

$$\mu_n(\mathrm{d}U,\mathrm{d}t) := \mu^n_t(\mathrm{d}U)\mathrm{d}t.$$

The set of measures $\{\mathcal{L}(\mu_n), n \in \mathbb{N}\}$ is tight on the space of Young measures on $\mathbb{U}$, denoted as $\mathcal{Y}(0,T;\mathbb{U})$, due to part 3 of the Assumption 2.7 and Theorem A.24. Thus, the set $\{\mathcal{L}(X_n, N_n, W_n, \mu_n), n \in \mathbb{N}\}$ is tight on $\mathcal{Z} \times M_{\bar{\mathbb{N}}}([0,T]\times Z) \times C([0,T];\mathbb{R}) \times \mathcal{Y}(0,T;\mathbb{U})$, where $W_n := W$ and $N_n := N$. Now arguing in the similar fashion as in Section 4.3, by the Skorokhod–Jakubowski theorem, we ensure the existence of a new probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathscr{F}}_t, \tilde{\mathbb{P}})$, and on this space random variables $(\tilde{X}_n, \tilde{N}_n, \tilde{W}_n, \tilde{\mu}_n)_{n\in\mathbb{N}}$, $(\tilde{X}, \tilde{N}, \tilde{W}, \tilde{\mu})$, such that

$$\mathcal{L}((\tilde{X}_n, \tilde{N}_n, \tilde{W}_n, \tilde{\mu}_n)) = \mathcal{L}((X_n, N_n, W_n, \mu_n)) \ \forall n \in \mathbb{N} \quad \text{and}$$

$$(\tilde{X}_n, \tilde{N}_n, \tilde{W}_n) \to (\tilde{X}, \tilde{N}, \tilde{W}) \quad \text{in} \quad \mathcal{Z} \times M_{\bar{\mathbb{N}}}([0,T]\times Z) \times C([0,T];\mathbb{R}) \quad \text{and}$$

$$\tilde{\mu}_n \to \tilde{\mu} \quad \text{stably in} \quad \mathcal{Y}(0,T;\mathbb{U}),\ \tilde{\mathbb{P}} - \text{a.s. as } n \to \infty.$$

Moreover, by Lemma A.16, for each $n \in \mathbb{N}$ there exists relaxed control processes $\{\tilde{\mu}^n_t\}_{t\ge 0}$ and $\{\tilde{\mu}_t\}_{t\ge 0}$ defined on $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ such that

$$\tilde{\mu}_n(\mathrm{d}U,\mathrm{d}t) := \tilde{\mu}^n_t(\mathrm{d}U)\mathrm{d}t, \quad \tilde{\mathbb{P}} - \text{a.s.}, \quad \text{and} \quad \tilde{\mu}(\mathrm{d}U,\mathrm{d}t) := \tilde{\mu}_t(\mathrm{d}U)\mathrm{d}t, \quad \tilde{\mathbb{P}} - \text{a.s.}$$

We now assert that for every $\varphi \in \mathcal{U}$,

$$\lim_{n\to\infty} \tilde{\mathbb{E}}\Big[\int_0^T \Big|\int_0^t \langle \mathcal{N}\tilde{\mu}^n_s - \mathcal{N}\tilde{\mu}_s, \varphi\rangle\,\mathrm{d}s\Big|\mathrm{d}t\Big] = 0. \tag{5.7}$$

Indeed, by (B.2) in Lemma B.1 in Appendix B.1, we have for every $\varphi \in \mathcal{U}$,

$$\lim_{n\to\infty} \tilde{\mathbb{E}}\Big|\int_0^t \int_{\mathbb{U}} (\mathsf{L}(U),\varphi)\,\tilde{\mu}_n(\mathrm{d}U,\mathrm{d}s) - \int_0^t \int_{\mathbb{U}} (\mathsf{L}(U),\varphi)\,\tilde{\mu}(\mathrm{d}U,\mathrm{d}s)\Big| = 0.$$

Hence, by dominated convergence theorem, we obtain for every $\varphi \in \mathcal{U}$,

$$\lim_{n\to\infty} \tilde{\mathbb{E}}\Big[\int_0^T \Big|\int_0^t \int_{\mathbb{U}} (\mathsf{L}(U),\varphi)\,\tilde{\mu}_n(\mathrm{d}U,\mathrm{d}s) - \int_0^t \int_{\mathbb{U}} (\mathsf{L}(U),\varphi)\,\tilde{\mu}(\mathrm{d}U,\mathrm{d}s)\Big|\mathrm{d}t\Big] = 0,$$

and finally using $\mathcal{N}\tilde{\mu}_t^n = \int_{\mathbb{U}} \mathsf{L}(U)\tilde{\mu}_t^n(\mathrm{d}U)$, and $\mathcal{N}\tilde{\mu}_t := \int_{\mathbb{U}} \mathsf{L}(U)\tilde{\mu}_t(\mathrm{d}U)$ we prove the above assertion.

Thus, parts (i)–(iv) of Definition 5.1 are satisfied. Identification of the limit and thereby part (v) of Definition 5.1 follow using Lemma 4.6, (5.7) and imitating the steps as in proof of Theorem 4.7. In other words, in view of Remark 5.2,

$$\tilde{\pi} = (\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathscr{F}}_t, \tilde{\mathbb{P}}, \{\tilde{X}(t)\}_{t\geq 0}, \{\tilde{N}(t,\cdot)\}_{t\geq 0}, \{\tilde{W}(t)\}_{t\geq 0}, \{\tilde{\mu}_t\}_{t\geq 0})$$

is a martingale solution of (1.5)–(1.6) associated with (1.7).
Moreover using Lemma A.25(b), Fatou's lemma and using $\mathcal{L}(\tilde{X}_n, \tilde{\mu}_n) = \mathcal{L}(X_n, \mu_n)$, we have

$$\begin{aligned}\tilde{\mathbb{E}}\Big[\int_0^T \int_{\mathbb{U}} F(t,\tilde{X}(t),U)\tilde{\mu}(\mathrm{d}U,\mathrm{d}t)\Big] &\leq \liminf_{n\to\infty} \tilde{\mathbb{E}}\Big[\int_0^T \int_{\mathbb{U}} F(t,\tilde{X}_n(t),U)\tilde{\mu}_n(\mathrm{d}U,\mathrm{d}t)\Big] \\ &= \liminf_{n\to\infty} \mathbb{E}^n\Big[\int_0^T \int_{\mathbb{U}} F(t,X_n(t),U)\mu_n(\mathrm{d}U,\mathrm{d}t)\Big]\end{aligned} \tag{5.8}$$

which is finite because of (5.4). Since $\int_0^T \bar{F}(t,\tilde{X}(t),\tilde{\mu}_t)\mathrm{d}t = \int_0^T \int_{\mathbb{U}} F(t,\tilde{X}(t),U)\tilde{\mu}(\mathrm{d}U,\mathrm{d}t)$, hence, the mapping

$$[0,T]\times\tilde{\Omega} \ni (t,\omega) \mapsto \bar{F}(t,\tilde{X}(t,\omega),\tilde{\mu}_t(\omega)) \in \mathbb{R}$$

belongs to $\mathrm{L}^1([0,T]\times\tilde{\Omega};\mathbb{R})$. For the sake of simplicity, we assume that T is a point of continuity of both $\tilde{X}_n$ and $\tilde{X}$. We note that $\phi(\tilde{X}(T)) \in \mathrm{L}^1(\tilde{\Omega};\mathbb{R})$. Thus, part (vi) of Definition 5.1 is also satisfied. Hence, $\tilde{\pi}$ is a weak admissible relaxed control.

Finally, employing (5.4) and (5.8), it follows that

$$\begin{aligned}\bar{J}(\tilde{\pi}) &= \tilde{\mathbb{E}}\Big[\int_0^T \int_{\mathbb{U}} F(t,\tilde{X}(t),U)\tilde{\mu}(\mathrm{d}U,\mathrm{d}t)\Big] + \tilde{\mathbb{E}}\,\phi(\tilde{X}(T)) \\ &\leq \liminf_{n\to\infty}\Big[\mathbb{E}^n \int_0^T \int_{\mathbb{U}} F(t,X_n(t),U)\mu_n(\mathrm{d}U,\mathrm{d}t) + \mathbb{E}^n\,\phi(X_n(T))\Big] = \inf_{\pi\in\mathfrak{U}_{ad}^w(X_0,T)} \bar{J}(\pi).\end{aligned}$$

This proves $\tilde{\pi}$ is a weak optimal relaxed control for the RCP. Thus, the proof is now complete. □

6. Applications

In the following section, we will provide applications to viscoelastic fluid models of the abstract theory developed earlier in this paper.

6.1. Oldroyd fluid

One of the most well-known linear viscoelastic fluid model was proposed by Oldroyd [57] and is known as Oldroyd-B fluid. The focus of this subsection is to concentrate on the two and three-dimensional stochastic controlled Oldroyd model influenced by Lévy noise in an open connected subset O in $\mathbb{R}^d$ (possibly unbounded) with smooth boundary ∂O. We denote the velocity field by $\mathbf{u}$ and the pressure field by p. Let $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ be a filtered probability space satisfying usual hypotheses and let $\mathbb{U}$ be the control set. The stochastic controlled system of equations of motion arising in the Oldroyd fluids of order one is:

$$
\begin{aligned}
&d\mathbf{u}(t) + [(\mathbf{u}\cdot\nabla)\mathbf{u} - \theta\Delta\mathbf{u} - \int_0^t \beta(t-\tau)\Delta\mathbf{u}(x,\tau)\,d\tau + \nabla p]\mathrm{d}t \\
&= \mathsf{L}(U)\mathrm{d}t + \sigma(t,\mathbf{u})\mathrm{d}W(t) + \int_Z g(\mathbf{u},z)\tilde{N}(\mathrm{d}t,\mathrm{d}z) \qquad \text{in} \quad O\times(0,T).
\end{aligned} \tag{6.1}
$$

$$
\nabla\cdot\mathbf{u} = 0 \quad \text{in} \quad O\times(0,T) \tag{6.2}
$$

with initial and boundary conditions,

$$
\mathbf{u} = 0 \qquad \text{in} \quad \partial O\times(0,T) \qquad \text{and} \qquad \mathbf{u} = \mathbf{u}_0 \qquad \text{in} \quad O\times\{0\} \tag{6.3}
$$

with the associated cost functional as

$$
\bar{J}(\mathbf{u},U) := \mathbb{E}[\mathcal{J}(\mathbf{u},U)] := \mathbb{E}\Big[\int_0^T F(t,\mathbf{u}(t),U(t))\mathrm{d}t\Big] \tag{6.4}
$$

where $U\in\mathbb{U}$. Here, $\theta = \frac{2\kappa}{\rho} > 0$ and the kernel $\beta(t) = \gamma e^{-\delta t}$ with $\gamma = \frac{2}{\rho}(\nu - \frac{\kappa}{\rho}) > 0$ and $\delta = \frac{1}{\rho} > 0$. For further details we refer Oldroyd [57] and Joseph [42]. We note that the kernel β satisfies Assumption 2.2 and assume that the running cost function F satisfies Assumption 2.8 (1) and (2).
Here $\{W(t)\}_{t\geq 0}$ is an L^2-valued cylindrical Wiener process and $\tilde{N}(\mathrm{d}t,\mathrm{d}z)$ is a time homogeneous compensated Poisson random measure with Z being the measurable subset of $\mathrm{L}^2(O)$ where the solutions of the above system have its paths and N is a time-homogeneous Poisson random measure on $(Z,\mathcal{B}(Z))$ over the filtered probability space $(\Omega,\mathscr{F},\mathscr{F}_t,\mathbb{P})$. The processes W and N are assumed to be independent. Furthermore, we assume that the noise coefficients satisfy Hypothesis 2.4. Associating the relaxed control by defining the relaxed coefficient as in Section A.3 we define

$$
\mathcal{N}\mu_t := \int_{\mathbb{U}} \mathsf{L}(U)\mu_t(\mathrm{d}U), \quad t\in[0,T]. \tag{6.5}
$$

where $\{\mu_t\}_{t\geq 0}$ is an $\mathscr{F}_t-$adapted relaxed control process.

6.2. Jeffreys fluid

Let $O\subset\mathbb{R}^d\,; d = 2,3$ be an open connected subset in $\mathbb{R}^d$ (possibly unbounded) with $\partial O\in C^2$. We consider the following system for $T\in(0,\infty]$ for the velocity vector $\mathbf{u}$, pressure p and the fluid stress tensor τ of a deterministic viscoelastic Jeffreys fluid model:

$$
\begin{aligned}
\partial_t\mathbf{u} - \theta\Delta\mathbf{u} + \nabla p = \nabla\cdot\tau \quad &\text{in} \quad O\times(0,T) &(6.6)\\
\nabla\cdot\mathbf{u} = 0 \quad &\text{in} \quad O\times(0,T) &(6.7)\\
\partial_t\tau + \rho\tau = 2\lambda D\mathbf{u} \quad &\text{in} \quad O\times(0,T) &(6.8)\\
\mathbf{u} = 0 \quad &\text{on} \quad \partial O\times(0,T) &(6.9)\\
\mathbf{u}(\cdot,0) = \mathbf{u}_0 \qquad \tau(\cdot,0) = \tau_0 \quad &\text{in} \quad O. &(6.10)
\end{aligned}
$$

where θ,ρ and λ are positive constants and $D\mathbf{u}$ is the symmetrized gradient tensor defined by $D\mathbf{u} := \frac{1}{2}\left(\nabla\mathbf{u} + \nabla^t\mathbf{u}\right)$. For additional information on the physical meaning of these parameters, see for instance Renardy *et al.* [65], Joseph [42]. For the notational convenience we will assume throughout that $\theta = 1$. Note that from the above equation τ can be written as

$$\tau(t) = e^{-\rho t}\tau_0 + 2\lambda \int_0^t e^{-\rho(t-s)} D\mathbf{u}(s)\mathrm{d}s \quad \forall\, t > 0. \tag{6.11}$$

Using (6.11), the stochastic controlled system of equations of motion perturbed by Lévy noise arising from (6.6)–(6.10) becomes

$$\begin{aligned} d\mathbf{u}(t) - [\mathbf{u}(t) + \lambda \int_0^t e^{-\rho(t-s)}\mathbf{u}(s)\mathrm{d}s + \nabla p]\mathrm{d}t &= [e^{-\rho t}\nabla\cdot\tau_0 + \mathsf{L}(U)]\mathrm{d}t \\ &+ \sigma(t,\mathbf{u})\mathrm{d}W(t) + \int_Z g(\mathbf{u},z)\tilde{N}(\mathrm{d}t,\mathrm{d}z) \quad \text{for all } t > 0, \end{aligned} \tag{6.12}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in} \quad O\times(0,T) \qquad \text{and} \qquad \mathbf{u} = \mathbf{u}_0 \quad \text{in} \quad O\times\{0\}. \tag{6.13}$$

We associate a cost functional with the above system in the similar manner following Oldroyd model and is given by (6.4). We also assume that Hypothesis 2.4 is satisfied by the noise coefficients.

The viscoelastic fluids of the Jeffreys kind, can be used as first approximations (taking into account that $\mathbf{u}, \tau$ are small) of the nonlinear system, (6.1)–(6.3), see Doubova *et al.* [27], Joseph [42].

6.3. Functional set-up of Oldroyd and Jeffreys fluid and compactness result

Below we provide the functional set-up of Oldroyd and Jeffreys fluid in both bounded and unbounded domain in $\mathbb{R}^d$ to fit into the abstract model considered in (1.5)–(1.6). We also elaborate the modifications required in connection with unbounded domain due to lack of compact embeddings in the Gelfand triple. Note that the Section 6.3.2 is not relevant to Jeffreys fluid because of the absence of the nonlinear term.

6.3.1. Functional spaces

Let $\mathcal{V} := \{\mathbf{u} \in \mathcal{C}_c^\infty(O,\mathbb{R}^d) : \operatorname{div}\mathbf{u} = 0\}$.

$$\mathbb{H} := \text{the closure of } \mathcal{V} \text{ in } \mathrm{L}^2(O,\mathbb{R}^d), \quad \text{and} \quad \mathbb{V} := \text{the closure of } \mathcal{V} \text{ in } H^1(O,\mathbb{R}^d). \tag{6.14}$$

The scalar product and the norm inherited from $\mathrm{L}^2(O,\mathbb{R}^d)$ in the space $\mathbb{H}$ are denoted by $(\cdot,\cdot)_{\mathbb{H}}$ and $|\cdot|$, respectively, *i.e.*, $(\mathbf{u},\mathbf{v})_{\mathbb{H}} := (\mathbf{u},\mathbf{v})_{\mathrm{L}^2}, \quad |\mathbf{u}| := \|\mathbf{u}\|_{\mathrm{L}^2}, \quad \mathbf{u},\mathbf{v}\in\mathbb{H}$.
In the space $\mathbb{V}$ we consider the scalar product inherited from the Sobolev space $H^1(O,\mathbb{R}^d)$, *i.e.*, $(\mathbf{u},\mathbf{v})_{\mathbb{V}} := (\mathbf{u},\mathbf{v})_{\mathrm{L}^2} + (\nabla\mathbf{u},\nabla\mathbf{v})_{\mathrm{L}^2}, \quad \mathbf{u},\mathbf{v}\in\mathbb{V}$ and the norm $\|\mathbf{u}\|^2 := |\mathbf{u}|^2 + \|\nabla\mathbf{u}\|^2_{\mathrm{L}^2}$.

6.3.2. Trilinear operator b

Let us define the trilinear operator $b : \mathrm{L}^p(O,\mathbb{R}^d)\times W^{1,q}(O,\mathbb{R}^d)\times\mathrm{L}^r(O,\mathbb{R}^d)\to\mathbb{R}$ defined by

$$b(\mathbf{u},\mathbf{v},\mathbf{w}) := \int_O (\mathbf{u}\cdot\nabla\mathbf{w})\mathbf{v}dx, \quad \mathbf{u}\in\mathrm{L}^p(O,\mathbb{R}^d),\, \mathbf{v}\in W^{1,q}(O,\mathbb{R}^d),\, \mathbf{w}\in\mathrm{L}^r(O,\mathbb{R}^d)$$

with $\dfrac{1}{p}+\dfrac{1}{q}+\dfrac{1}{r}\le 1$ and $p,q,r\in[1,\infty]$.
Let us recall some fundamental properties of the form b which are valid for both bounded and unbounded domains (see Temam [70]).

(i) $b(\mathbf{u},\mathbf{v},\mathbf{w}) = -b(\mathbf{u},\mathbf{w},\mathbf{v}),\ \mathbf{u},\mathbf{v},\mathbf{w}\in\mathbb{V}$. In particular $b(\mathbf{u},\mathbf{v},\mathbf{v}) = 0,\ \mathbf{u},\mathbf{v}\in\mathbb{V}$.

(ii) Using Sobolev embedding theorem and Hölder's inequality, we obtain $|b(\mathbf{u},\mathbf{v},\mathbf{w})| \le c\|\mathbf{u}\|\|\mathbf{w}\|\|\mathbf{v}\|,\ \mathbf{u},\mathbf{w},\mathbf{v} \in \mathbb{V}$ for some positive constant c. Hence, the form b is continuous on $\mathbb{V}$.

(iii) Let us define a bilinear map $B : \mathbb{V}\times\mathbb{V} \to \mathbb{V}'$ by $B(\mathbf{u},\mathbf{v}) = b(\mathbf{u},\mathbf{v},\cdot)$. Then we have $B(\mathbf{u},\mathbf{v}) \in \mathbb{V}'$ and $\|B(\mathbf{u},\mathbf{v})\|_{\mathbb{V}'} \le c\|\mathbf{u}\|\|\mathbf{v}\|$ for $\mathbf{u},\mathbf{v}\in\mathbb{V}$. Hence $B : \mathbb{V}\times\mathbb{V}\to\mathbb{V}'$ is linear and continuous.

(iv) For any $s > 0$, let us define $\mathbb{V}_s :=$ the closure of $\mathcal{V}$ in $H^s(O,\mathbb{R}^d)$. If $s > \frac{d}{2}+1$ then by the Sobolev embedding theorem, $H^{s-1}(O,\mathbb{R}^d) \hookrightarrow C_b(O,\mathbb{R}^d) \hookrightarrow \mathrm{L}^\infty(O,\mathbb{R}^d)$. If $\mathbf{u},\mathbf{w}\in\mathbb{V}$ and $\mathbf{v}\in\mathbb{V}_s$ with $s > \frac{d}{2}+1$ then $|b(\mathbf{u},\mathbf{w},\mathbf{v})| \le c\|\mathbf{u}\|_{\mathrm{L}^2}\|\mathbf{w}\|_{\mathrm{L}^2}\|\mathbf{v}\|_{\mathbb{V}_s}$, for some constant $c > 0$.

6.3.3. Some operators

Consider the natural embedding $j : \mathbb{V} \hookrightarrow \mathbb{H}$ and its adjoint $j^* : \mathbb{H}\to\mathbb{V}$. Since the range of j is dense in $\mathbb{H}$, the map j^* is one-to-one. Let us consider

$$D(\mathcal{A}) := j^*(\mathbb{H}) \subset \mathbb{V},\ \ \mathcal{A}\mathbf{u} := (j^*)^{-1}\mathbf{u}, \quad \mathbf{u}\in D(A).$$

Also let us now put $A\mathbf{u} := (\nabla\mathbf{u},\nabla\cdot)_{\mathrm{L}^2},\ \mathbf{u}\in\mathbb{V}$. We note that

$$\begin{aligned}&(\mathcal{A}\mathbf{u},\mathbf{v})_{\mathbb{H}} = (\mathbf{u},\mathbf{v})_{\mathbb{V}} \quad \forall\,\mathbf{u}\in D(\mathcal{A}) \text{ and } \mathbf{v}\in\mathbb{V},\ \text{since}\\ &(\mathcal{A}\mathbf{u},\mathbf{v})_{\mathbb{H}} = ((j^*)^{-1}\mathbf{u},\mathbf{v})_{\mathbb{H}} = ((j^*)^{-1}\mathbf{u}, j\mathbf{v})_{\mathbb{H}} = (j^*(j^*)^{-1}\mathbf{u},\mathbf{v})_{\mathbb{V}} = (\mathbf{u},\mathbf{v})_{\mathbb{V}}.\end{aligned}$$

Hence, for any $\mathbf{u}\in D(\mathcal{A})$ and $\mathbf{v}\in\mathbb{V}$ we have, $((\mathcal{A}-I)\mathbf{u},\mathbf{v})_{\mathbb{H}} = (\nabla\mathbf{u},\nabla\mathbf{v})_{\mathrm{L}^2} = \langle A\mathbf{u},\mathbf{v}\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the duality pairing between $\mathbb{V}$ and $\mathbb{V}'$ and I stands for the identity operator on $\mathbb{H}$. In particular, $\|A\mathbf{u}\|_{\mathbb{V}'} \le |(\mathcal{A}-I)\mathbf{u}|_{\mathbb{H}}$. Indeed this follows immediately from the following equalities

$$(\mathcal{A}\mathbf{u},\mathbf{v})_{\mathbb{H}} = (\mathbf{u},\mathbf{v})_{\mathbb{V}} = (\mathbf{u},\mathbf{v})_{\mathrm{L}^2} + (\nabla\mathbf{u},\nabla\mathbf{v})_{\mathrm{L}^2} = (I\mathbf{u},\mathbf{v})_{\mathbb{H}} + \langle A\mathbf{u},\mathbf{v}\rangle$$

Additionally, we have $D(\mathcal{A})$ is dense in $\mathbb{H}$. Also let us notice that if $\mathbf{u}\in\mathbb{V}$, then $A\mathbf{u}\in\mathbb{V}'$ and we have the following inequalities

$$|(A\mathbf{u},\mathbf{v})| \le \|\nabla\mathbf{u}\|_{\mathrm{L}^2}\cdot\|\nabla\mathbf{v}\|_{\mathrm{L}^2} \le \|\nabla\mathbf{u}\|_{\mathrm{L}^2}(\|\nabla\mathbf{v}\|^2_{\mathrm{L}^2} + |\mathbf{v}|^2)^{\frac{1}{2}} = \|\nabla\mathbf{u}\|_{\mathrm{L}^2}\cdot\|\mathbf{v}\|_{\mathbb{V}},$$

$\mathbf{v}\in\mathbb{V}$, so we infer$\|A\mathbf{u}\|_{\mathbb{V}'} \le \|\nabla\mathbf{u}\|_{\mathrm{L}^2}$.

6.3.4. Compactness results

Note that if O is unbounded, then the embedding $\mathbb{V}\hookrightarrow\mathbb{H}$ is not compact. Let us assume that $s > 1$. It is obvious that $\mathbb{V}_s$ is continuously embedded in $\mathbb{V}$ and the embedding is dense. Then by Lemma C.1 in Appendix C of [15], there exists a Hilbert space $\mathcal{U}$ such that $\mathcal{U}\subset\mathbb{V}_s$, $\mathcal{U}$ is dense in $\mathbb{V}_s$ and the natural embedding $\mathcal{U}\hookrightarrow\mathbb{V}_s$ is compact. Hence the space $\mathcal{Z}$, in unbounded domain, is defined as $D([0,T];\mathcal{U}')_J \cap D([0,T];\mathbb{H}_w)\cap \mathrm{L}^2_w(0,T;\mathbb{V})\cap \mathrm{L}^2(0,T;\mathbb{H}_{loc})$.

Remark 6.1.

1. We apply the abstract framework identifying the spaces $\mathbb{H},\mathbb{V}$ (defined in 6.14) as H and V respectively, and identify V_1 (in Asm. 2.1 B(iv)) as $\mathbb{V}_s$ for $s > \frac{d}{2}+1$, as defined above. Hence for $s > \frac{d}{2}+1$, the operator B, defined in Section 6.3.2, satisfies all the Assumptions 2.1 (B).
2. The kernels in Oldroyd and Jeffreys fluid satisfy the Assumption 2.2.

3. With the above functional set up, the system of equations (6.1)–(6.3) for Oldroyd fluid and the system of equations (6.12)–(6.13) for Jeffreys fluid, combined with (6.5) take the abstract form (1.5)–(1.6), while the cost functional (6.4) becomes (1.7).

6.4. Existence of optimal relaxed control for Oldroyd and Jeffreys fluids

In this section, we provide some important results as an application to the theory developed in this paper for Oldroyd and Jeffreys fluids in both two and three dimensions.

Theorem 6.2. *Under the above mathematical settings, let us assume* $\mathbf{u}_0 \in \mathrm{L}^2(\Omega;\mathbb{H})$ *(where* $\mathbb{H}$ *is defined as in* (6.14)*),* σ *and* g *meet Hypothesis 2.4. Then there exists a martingale solution to the uncontrolled system associated to* (6.1)–(6.3) *(resp.* (6.12)–(6.13)*)* .

Theorem 6.3. *Let assumptions of the Theorem 6.2, Assumptions 2.7 and 2.8 (1) and (2) be satisfied. Let* $\pi \in \mathfrak{U}^w_{ad}(\mathbf{u}_0, T)$ *be such that* $\bar{J}(\pi) < +\infty$*, where* $\bar{J}$ *is defined by* (6.4)*. Then the RCP admits a weak optimal relaxed control with time horizon* $[0,T]$.

The proofs of Theorems 6.2 and 6.3 are direct applications of Theorems 4.7 and 5.4, respectively.

6.5. An example

Example 6.4. Consider the Oldroyd model (6.1)–(6.3) in a bounded domain in two dimension. For the sake of simplicity and practical understanding, consider a very particular case of the general control problem, where the control input operator L is continuous linear with the control set $\mathbb{U} = \mathrm{L}^2$ (which is a Suslin space, see p. 556 in [72]) *i.e.*, $\mathsf{L} \in \mathcal{L}(\mathrm{L}^2;\mathrm{L}^2)$ and $|\mathsf{L}(U)| \le C|U|$, $\forall\, U \in \mathrm{L}^2$ and for some $C > 0$. We work under the same functional set-up as defined in Section 6.3.1. We associate a convex cost functional, in its simplest form, as

$$J(\mathbf{u}, U) := \mathbb{E}\Big[\int_0^T (\|\mathbf{u}(s)\|^2 + |U(s)|^2_{\mathrm{L}^2})\mathrm{d}s\Big]. \tag{6.15}$$

Define $\mathfrak{U}^w_{ad}(X_0, T)$ as the set of weak admissible controls (with time horizon $[0,T]$) and is given by: $\mathfrak{U}^w_{ad}(X_0,T) =: \{(\mathbf{u}, U) \in \Big(\mathrm{L}^2(0,T;\mathbb{V}) \cap \mathrm{L}^\infty(0,T;\mathbb{H})\Big) \times \mathrm{L}^2(0,T;\mathrm{L}^2) : (\mathbf{u}, U) \text{ solves } (6.1)\text{–}(6.3) \text{ and } J(\mathbf{u}, U) < +\infty\}$.

We treat the control problem of finding $(\tilde{\mathbf{u}}, \tilde{U}) \in \mathfrak{U}^w_{ad}(X_0,T)$ and $J(\tilde{\mathbf{u}}, \tilde{U}) = \inf\{J(\mathbf{u}, U) : (\mathbf{u}, U) \in \mathfrak{U}^w_{ad}(X_0,T)\}$. It is clear that J is bounded below, infact $J \ge 0$. By definition of $\mathfrak{U}^w_{ad}(X_0,T)$ we can say that

$$0 \le \inf\{J(\mathbf{u}, U) : (\mathbf{u}, U) \in \mathfrak{U}^w_{ad}(X_0,T)\} < +\infty.$$

Hence, there exists a minimizing sequence $(\mathbf{u}_n, U_n) \in \mathfrak{U}^w_{ad}(X_0,T)$ such that $J(\mathbf{u}_n, U_n) \to \inf\{J(\mathbf{u}, U) : (\mathbf{u}, U) \in \mathfrak{U}^w_{ad}(X_0,T)\}$ as $n \to \infty$. So, there exists $R > 0$ large such that $\sup_{n\ge1} J(\mathbf{u}_n, U_n) \le R$, *i.e.*, $\sup_{n\ge1} \mathbb{E}\big[\int_0^T (\|\mathbf{u}_n(s)\|^2 + |U_n(s)|^2_{\mathrm{L}^2})\mathrm{d}s\big] \le R$, *i.e.*, $\mathbb{E}\big[\int_0^T |U_n(s)|^2_{\mathrm{L}^2}\mathrm{d}s\big] \le R$, $\forall n \ge 1$. It can be shown that $(\mathbf{u}_n, U_n)$ satisfies the energy inequality $\mathbb{E}\left(\sup_{0\le t\le T} |\mathbf{u}_n(t)|^2\right) + 2\mathbb{E}\left(\int_0^T \|\mathbf{u}_n(t)\|^2\,\mathrm{d}t\right) \le c(\mathbb{E}|\mathbf{u}(0)|^2, R, K_2, K_4, T)$.

Now proceeding in the similar manner as in the proof of Theorem 5.4, we ensure the existence of a probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathscr{F}}_t, \tilde{\mathbb{P}})$, and on this space, a sequence of random variables $(\tilde{\mathbf{u}}_k, \tilde{U}_k)_{k\in\mathbb{N}}$ and a subsequence $(\mathbf{u}_{n_k}, U_{n_k})$ such that $\mathcal{L}((\tilde{\mathbf{u}}_k, \tilde{U}_k)) = \mathcal{L}((\mathbf{u}_{n_k}, U_{n_k}))\ \forall k \in \mathbb{N}$ and $(\tilde{\mathbf{u}}_k, \tilde{U}_k) \to (\tilde{\mathbf{u}}, \tilde{U})$ in $\Big(\mathrm{L}^2(0,T;\mathbb{V}) \cap \mathrm{L}^\infty(0,T;\mathbb{H})\Big) \times \mathrm{L}^2(0,T;\mathrm{L}^2)$ $\tilde{\mathbb{P}}$ – a.s. Furthermore, we have $(\tilde{\mathbf{u}}, \tilde{U}) \in \mathfrak{U}^w_{ad}(X_0,T)$ and $J(\tilde{\mathbf{u}}, \tilde{U}) = \inf\{J(\mathbf{u}, U) : (\mathbf{u}, U) \in \mathfrak{U}^w_{ad}(X_0,T)\}$. Hence the result.

6.6. Uniqueness and strong solutions

In the two-dimensional case, we will prove that $\mathbb{P}$ – a.s. the trajectories are equal almost everywhere to an $\mathbb{H}$-valued càdlàg functions defined on $[0, T]$. We will also prove that the solutions are pathwise unique and show the existence of strong solutions and uniqueness in law. We provide proofs for the Oldroyd fluid, and proofs for the Jeffreys fluid follow similarly.

Definition 6.5. The problems (6.1)–(6.3) have a strong solution iff for every stochastic basis $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ and for every $\mathbb{H}$-valued cylindrical Wiener process $W(t)$ and time homogeneous Poisson random measure N with intensity measure λ, there exists a progressively measurable process $\mathbf{u} : [0, T] \times \Omega \to \mathbb{H}$ with $\mathbb{P}$ – a.e. paths

$$\mathbf{u}(\cdot, \omega) \in \mathrm{L}^2(0, T; \mathbb{V}) \cap D([0, T]; \mathbb{H})$$

such that for all $t \in [0, T]$ and all $\varphi \in \mathcal{V}$:

$$\begin{aligned}(\mathbf{u}(t), \varphi)_H = (\mathbf{u}(0), \varphi)_H &- \int_0^t \langle A\mathbf{u}(s), \varphi\rangle \, \mathrm{d}s - \int_0^t \langle B(\mathbf{u}(s)), \varphi\rangle \, \mathrm{d}s \\ &- \int_0^t \int_0^s \beta(s-r)\langle A\mathbf{u}(r), \varphi\rangle \, \mathrm{d}r \, \mathrm{d}s + \int_0^t (\mathcal{N}\mu_s, \varphi)_H \mathrm{d}s \\ &+ \int_0^t (\sigma(s, \mathbf{u}(s))\mathrm{d}W(s), \varphi)_H + \int_0^t \int_Z (g(s, \mathbf{u}(s-), z), \varphi)_H \, \tilde{N}(\mathrm{d}s, \mathrm{d}z),\end{aligned} \tag{6.16}$$

the identity holds $\mathbb{P}$ – a.s.

Definition 6.6. Let $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P}, \{\mathbf{u}_i(t)\}_{t\geq 0}, \{N(t, \cdot)\}_{t\geq 0}, \{W(t)\}_{t\geq 0}, \{\mu_t\}_{t\geq 0})$ for $i = 1, 2$ be two martingale solutions of (6.1)–(6.3) with $\mathbb{P}(\mathbf{u}_1(0) = \mathbf{u}_2(0)) = 1$. Then we say they are pathwise unique if $\mathbb{P}(\mathbf{u}_1(t) = \mathbf{u}_2(t)) = 1$ for every $t \in [0, T]$.

Definition 6.7. Let $(\Omega_i, \mathscr{F}_i, \mathscr{F}_t^i, \mathbb{P}_i, \{\mathbf{u}_i(t)\}_{t\geq 0}, \{N_i(t, \cdot)\}_{t\geq 0}, \{W_i(t)\}_{t\geq 0}, \{\mu_t^i\}_{t\geq 0})$ for $i = 1, 2$ be two martingale solutions of (6.1)–(6.3) with $\mathbf{u}_i(0) = \mathbf{u}_0$; $i = 1, 2$. Then we say the solutions are unique in law if

$$\mathcal{L}_{\mathbb{P}_1}(\mathbf{u}_1) = \mathcal{L}_{\mathbb{P}_2}(\mathbf{u}_2) \quad \text{on} \quad \mathrm{L}^2(0, T; \mathbb{V}) \cap D([0, T]; \mathbb{H}),$$

where $\mathcal{L}_{\mathbb{P}_i}(\mathbf{u}_i)$; $i = 1, 2$ are probability measures on $\mathrm{L}^2(0, T; \mathbb{V}) \cap D([0, T]; \mathbb{H})$.

Lemma 6.8. *Let $d = 2$ and Assumptions 2.1, 2.2, 2.7 and Hypothesis 2.4 be satisfied. Let $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P}, \{\mathbf{u}(t)\}_{t\geq 0}, \{N(t, \cdot)\}_{t\geq 0}, \{W(t)\}_{t\geq 0}, \{\mu_t\}_{t\geq 0})$ be a martingale solutions of (6.1)–(6.3) such that*

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |\mathbf{u}(t)|^2\right) + 2\mathbb{E}\left(\int_0^T \|\mathbf{u}(t)\|^2 \, \mathrm{d}t\right) < \infty. \tag{6.17}$$

Then the trajectory $\mathbf{u}(\cdot, \omega)$ is $\mathbb{P}$ almost everywhere equal to a càdlàg $\mathbb{H}$-valued function defined on $[0, T]$. Moreover, for every $t \in [0, T]$, $\mathbb{P}$ – a.s.

$$\begin{aligned}\mathbf{u}(t) = \mathbf{u}(0) &- \int_0^t [A\mathbf{u}(s) + B(\mathbf{u}(s)) + \int_0^s \beta(s-r)A\mathbf{u}(r)\mathrm{d}r - \mathcal{N}\mu_s]\mathrm{d}s \\ &+ \int_0^t (\sigma(s, \mathbf{u}(s))\mathrm{d}W(s) + \int_0^t \int_Z (g(s, \mathbf{u}(s-), z) \, \tilde{N}(\mathrm{d}s, \mathrm{d}z).\end{aligned} \tag{6.18}$$

Proof. As $\mathbf{u}$ is a martingale solution of (6.1)–(6.3),

$$\mathbf{u} \in \mathrm{L}^2(\Omega; \mathrm{L}^2_w(0,T;\mathbb{V}) \cap \mathrm{L}^2(0,T;\mathbb{H}) \cap D([0,T];\mathcal{U}') \cap D([0,T];\mathbb{H}_w)),$$

$\mathbb{P}$ – a.s. (replace $\mathrm{L}^2(0,T;\mathbb{H})$ by $\mathrm{L}^2(0,T;\mathbb{H}_{loc})$ in case the domain is unbounded). Let us first show that right hand side (R.H.S.) of (6.18) makes sense, *i.e.*, we will show that each term on the R.H.S is well-defined in the space $\mathbb{V}'$.

Let us start with the linear term. Exploiting the fact $A : \mathbb{V} \to \mathbb{V}'$ given by $|A\mathbf{u}|_{\mathbb{V}'} \leq \|\mathbf{u}\|$ and then by Hölder's inequality and (6.17) we have the following bound:

$$\mathbb{E}\Big(\| \int_0^t A\mathbf{u}(s)\mathrm{d}s\|^2_{\mathbb{V}'}\Big) \leq C\,\mathbb{E}\Big(\int_0^T \|A\mathbf{u}(s)\|^2_{\mathbb{V}'}\mathrm{d}s\Big) \leq C\,\mathbb{E}\Big(\int_0^T \|\mathbf{u}(s)\|^2\mathrm{d}s\Big) < \infty.$$

Now we move to the nonlinear term. Repeating arguments as done above for the linear term and using (2.3) we have

$$\mathbb{E}\Big(\| \int_0^t B(\mathbf{u}(s))\mathrm{d}s\|^2_{\mathbb{V}'}\Big) \leq C\,\mathbb{E}\Big(\int_0^T \|B(\mathbf{u}(s))\|^2_{\mathbb{V}'}\mathrm{d}s\Big) \leq C\,\mathbb{E}\Big(\int_0^T \|\mathbf{u}(s)\|^2\mathrm{d}s\Big) < \infty.$$

For the memory term again using Hölder's inequality and (6.17) we have

$$\mathbb{E}\Big(\| \int_0^t \int_0^s \beta(s-r)A\mathbf{u}(r)\mathrm{d}r\,\mathrm{d}s\|^2_{\mathbb{V}'}\Big) \leq C\,T\,\mathbb{E}\Big(\int_0^T \|A\mathbf{u}(r)\|^2_{\mathbb{V}'}dr\Big) < \infty.$$

Next we deal with the control term. Using Hölder's inequality and proceeding as in Remark A.20 we have

$$\mathbb{E}\Big[\int_0^T |\mathcal{N}\mu_t|^2\mathrm{d}t\Big] \leq \mathbb{E}\Big[\int_0^T \Big(\int_{\mathbb{U}} |\mathsf{L}(U)|\mu_t(\mathrm{d}U)\Big)^2\mathrm{d}t\Big] \leq \mathbb{E}\Big[\int_0^T \int_{\mathbb{U}} \kappa^2(t,U)\mu_t(\mathrm{d}U)\mathrm{d}t\Big] < +\infty.$$

Due to Itô isometry, Hypotheses (2.4) and (6.17),

$$\mathbb{E}\Big(|\int_0^t \sigma(s,\mathbf{u}(s))\mathrm{d}W(s)|^2\Big) \leq C\,T + C\,T\,\mathbb{E}\Big(\sup_{0\leq s\leq T}|\mathbf{u}(s)|^2\Big) < \infty.$$

and by Itô–Lévy isometry, Hypothesis 2.4 and (6.17),

$$\mathbb{E}\Big(|\int_0^t \int_Z g(\mathbf{u}(s-),z)\tilde{N}(\mathrm{d}s,\mathrm{d}z)|^2\Big) \leq C\,T + C\,T\,\mathbb{E}\Big(\sup_{0\leq s\leq T}|\mathbf{u}(s)|^2\Big) < \infty.$$

Thus we have shown that each term in (6.18) is well defined. Now we will show that the equality in (6.18) holds. Since $\mathbf{u}$ is a martingale solution of (6.1)–(6.3), for every $\varphi \in \mathbb{V}, t \in [0,T]$ we have

$$\begin{aligned}\langle \mathbf{u}(t), \varphi\rangle = \langle \mathbf{u}_0, \varphi\rangle &- \int_0^t \langle A\mathbf{u}(s), \varphi\rangle\,\mathrm{d}s - \int_0^t \langle B(u(s)), \varphi\rangle\,\mathrm{d}s \\ &- \int_0^t\int_0^s \beta(s-r)\big\langle A\mathbf{u}(r), \varphi\big\rangle\,\mathrm{d}r\,\mathrm{d}s + \int_0^t \langle \mathcal{N}\mu_s, \varphi\rangle \mathrm{d}s \\ &+ \int_0^t \langle \sigma(s,\mathbf{u}(s))\mathrm{d}W(s), \varphi\rangle + \int_0^t\int_Z \langle g(s,\mathbf{u}(s-),z), \varphi\rangle\,\tilde{N}(\mathrm{d}s,\mathrm{d}z).\end{aligned} \tag{6.19}$$

Note that the above equation holds true for $\varphi \in \mathcal{V}$ and hence (6.18) holds in the distribution sense. Since $\mathcal{V}$ is dense in $\mathbb{V}$, (6.18) holds almost everywhere. Let us write $\mathbf{u}(t) = \mathbf{u}_0 + \int_0^t G(s)\mathrm{d}s + M(t)$ where G contains all the deterministic terms and M corresponds to the noise term. We have already shown above that $M \in \mathrm{L}^2(\Omega; L^2(0,T;\mathbb{H}))$. The proof for $G \in \mathrm{L}^2(\Omega; L^2(0,T;\mathbb{H}))$ is not so straightforward and omit the details of the proof and refer the readers to Lemma 7.2. and Theorem D.2 of Brzeźniak and Motyl [15] (see also Gyöngy [37], Pardoux [60] for similar ideas). We only give a sketch of the proof here.
Let us consider the following shifted stochastic hereditary Stokes equation

$$\mathbf{y}(t) = -\int_0^t \mathcal{A}\mathbf{y}(s)\mathrm{d}s - \int_0^t \int_0^s \beta(s-r)\mathcal{A}\mathbf{y}(r)\,\mathrm{d}r\,\mathrm{d}s + M(t) + \mathcal{N}\mu_t, \tag{6.20}$$

where $\mathcal{A}$ is the operator defined in Section 6.3.3. Since $\mathcal{A}$ satisfies $(\mathcal{A}\mathbf{y}, \mathbf{w})_{\mathbb{H}} = (\mathbf{y}, \mathbf{w})_{\mathbb{V}}$ for all $\mathbf{y} \in D(\mathcal{A})$ and $\mathbf{w} \in \mathbb{V}$, and $\mathbf{u}$ satisfies inequality (6.17), by Theorem 1.3 in [59] (with suitable modification for the càdlàg martingale), we infer that equation (6.20) has a unique progressively measurable solution $\mathbf{y}$ such that $\mathbb{P}$ – a.s. $\mathbf{y} \in D([0,T];\mathbb{H})$ and

$$\mathbb{E}\left(\sup_{0\le t\le T} |\mathbf{y}(t)|^2\right) + 2\mathbb{E}\left(\int_0^T \|\mathbf{y}(t)\|^2\,\mathrm{d}t\right) < \infty.$$

Let $\mathbf{v}(t) := \mathbf{u}(t) - \mathbf{y}(t),\ t \in [0,T]$. For $\mathbb{P}$-almost all $\omega \in \Omega$, the function $\mathbf{v} = \mathbf{v}(\cdot,\omega)$ is a weak solution of the following deterministic equation

$$\begin{aligned}\frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t} &= -\mathcal{A}\mathbf{v}(t) + \mathbf{v}(t) + \mathbf{y}(t) - \int_0^t \beta(t-r)\mathcal{A}\mathbf{v}(r)\mathrm{d}r + \int_0^t \beta(t-r)\big(\mathbf{v}(r) + \mathbf{y}(r)\big)\mathrm{d}r \\ &\quad - B\big(\mathbf{v}(t) + \mathbf{y}(t)\big),\end{aligned} \tag{6.21}$$

with $\mathbf{v}(0) = \mathbf{u}_0$. Let $\omega \in \Omega$ be such that $\mathbf{u}(\cdot,\omega) \in \mathrm{L}^2(0,T;\mathbb{V}) \cap D([0,T];\mathbb{H}_w)$ and $\mathbf{y}(\cdot,\omega) \in \mathrm{L}^2(0,T;\mathbb{V}) \cap D([0,T];\mathbb{H})$. Proceeding in the similar way as in Theorem D.2 of Brzeźniak and Motyl [15], one can ensure the existence of a unique weak solution of problem (6.21), and the weak solution is almost everywhere equal to a càdlàg $\mathbb{H}$-valued function defined on $[0,T]$. Let $\tilde{\mathbf{v}}(\cdot,\omega) \in \mathrm{L}^2(0,T;\mathbb{V}) \cap D([0,T];\mathbb{H})$ be the unique solution of (6.21) with the initial condition $\tilde{\mathbf{v}}(0) = \mathbf{u}_0$. Due to the uniqueness, we obtain for almost all $t \in [0,T]$, $\tilde{\mathbf{v}}(t) = \mathbf{u}(t) - \mathbf{y}(t)$. Put $\hat{\mathbf{u}}(t) := \tilde{\mathbf{v}}(t) + \mathbf{y}(t),\ t \in [0,T]$. Then $\hat{\mathbf{u}} \in D([0,T];\mathbb{H})$ and $\mathbf{u}(t) = \hat{\mathbf{u}}(t)$ for almost all $t \in [0,T]$. The sketch of the proof is now complete. □

In the following lemma, we will prove that the solutions of (6.1)–(6.3) are pathwise unique in $\mathbb{R}^d; d = 2,3$. The result in $d = 3$ is under a very restrictive moment condition. For the sake of simplicity, we restrict the proof for bounded domain only.

Theorem 6.9. *Let O be a bounded domain in $\mathbb{R}^d; d = 2,3$. Let Assumptions 2.1, 2.2, 2.7 and Hypothesis 2.4 be satisfied. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be two martingale solutions defined on the same filtered probability space $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ satisfying system of equations (6.1)–(6.3)with the same initial condition $\mathbf{u}_1(0) = \mathbf{u}_2(0) = \mathbf{u}_0, \mathbb{P} - a.s.$ where $\mathbb{E}[|\mathbf{u}_0|^2] < \infty$. Then*

- *(i) for $d = 2$, $\mathbb{P}$ – a.s. for all $t \in [0,T]$ $\mathbf{u}_1(t) = \mathbf{u}_2(t)$,*
- *(ii) for $d = 3$, $\mathbb{P}$ – a.s. for all $t \in [0,T]$ $\mathbf{u}_1(t) = \mathbf{u}_2(t)$, provided $\mathbb{E}\Big[\int_0^T \|\mathbf{u}(t)\|^4\mathrm{d}t\Big] < +\infty$.*

Remark 6.10. In comparison to the proof of Theorem 6.9, proof of pathwise uniqueness theorem for Jeffreys fluid ((6.12)–(6.13)) in both two and three dimensions in bounded domain are straight forward because of the absence of nonlinear term. So unlike Oldroyd model, even for three dimension, no additional assumption is required for the pathwise uniqueness of the weak solution of Jeffreys model in bounded domain.

Proof. We begin the proof in both two and three dimensions when $\mathbf{u}_1, \mathbf{u}_2$ satisfies (6.1)–(6.3). Let $\tilde{\mathbf{u}} = \mathbf{u}_1 - \mathbf{u}_2$. Then $\tilde{\mathbf{u}}$ solves the stochastic differential equation

$$\begin{aligned} d\tilde{\mathbf{u}}(t) &+ \Big[A\tilde{\mathbf{u}}(t) + \big(B(\mathbf{u}_1(t)) - B(\mathbf{u}_2(t))\big) + \int_0^t \beta(t-s)A\tilde{\mathbf{u}}(s)\,\mathrm{d}s\Big]\mathrm{d}t \\ &= \big(\sigma(t,\mathbf{u}_1) - \sigma(t,\mathbf{u}_2)\big)\mathrm{d}W(t) + \int_Z \big(g(\mathbf{u}_1,z) - g(\mathbf{u}_2,z)\big)\tilde{N}(\mathrm{d}t,\mathrm{d}z), \end{aligned} \tag{6.22}$$

We denote $\sigma_d(t) = \sigma(t,\mathbf{u}_1(t)) - \sigma(t,\mathbf{u}_2(t))$ and $g_d(t) = g(\mathbf{u}_1(t-),z) - g(\mathbf{u}_2(t-),z)$.
Let us define

$$\tau_N = \inf\{t : |\tilde{\mathbf{u}}_n(t)|^2 + \int_0^t \|\tilde{\mathbf{u}}_n(s)\|^2 ds > N\}$$

as the stopping time. We have to find the 2nd moment estimate for the system (6.22). For this let us take the function $f(x) = |x|^2$ and apply the Itô's formula to the process $\tilde{\mathbf{u}}(t)$ to obtain

$$\begin{aligned} d|\tilde{\mathbf{u}}(t)|^2 + 2\|\tilde{\mathbf{u}}(t)\|^2\mathrm{d}t &\le -2\left(\int_0^t \beta(t-s)\left(A^{\frac12}\tilde{\mathbf{u}}(s), A^{\frac12}\tilde{\mathbf{u}}(t)\right)\mathrm{d}s\right)\mathrm{d}t + 2(\sigma_d(t)\mathrm{d}W(t),\tilde{\mathbf{u}}(t)) \\ &+ \big(B(\mathbf{u}_1(t)) - B(\mathbf{u}_2(t)),\tilde{\mathbf{u}}(t)\big)\mathrm{d}t + \|\sigma_d(t)\|^2_{\mathcal{L}_Q}\mathrm{d}t + 2\int_Z (\tilde{\mathbf{u}}(t), g_d(t,z))\tilde{N}(\mathrm{d}t,\mathrm{d}z) \\ &+ \int_Z |g_d(t,z)|^2\lambda(\mathrm{d}t)\mathrm{d}z. \end{aligned}$$

Integrating from 0 to $t\wedge\tau_N$, and using Lemma 2.3 we have

$$\begin{aligned} |\tilde{\mathbf{u}}(t\wedge\tau_N)|^2 + 2\int_0^{t\wedge\tau_N} \|\tilde{\mathbf{u}}(s)\|^2 ds &\le |\tilde{\mathbf{u}}(0)|^2 + \int_0^{t\wedge\tau_N} \big(B(\mathbf{u}_1(s)) - B(\mathbf{u}_2(s)),\tilde{\mathbf{u}}(s)\big)\mathrm{d}s \\ &+ \int_0^{t\wedge\tau_N} \|\sigma_d(s)\|^2_{\mathcal{L}_Q}\mathrm{d}s + 2\int_0^{t\wedge\tau_N} (\sigma_d(s)\mathrm{d}W(s),\tilde{\mathbf{u}}(s)) + \int_0^{t\wedge\tau_N}\int_Z |g_d(s,z)|^2\lambda(\mathrm{d}s)\mathrm{d}z \\ &+ 2\int_0^{t\wedge\tau_N}\int_Z (\tilde{\mathbf{u}}(s), g_d(s,z))\tilde{N}(\mathrm{d}s,\mathrm{d}z). \end{aligned} \tag{6.23}$$

Case 1. In two dimensions, using energy inequality it can be shown that for $0 \le t \le T$, $\mathbb{E}\left[|\tilde{\mathbf{u}}(t)|^2 + 2\int_0^t \|\tilde{\mathbf{u}}(s)\|^2 ds\right] < \infty$. We also have $|\langle B(\tilde{\mathbf{u}},\mathbf{u}_1),\tilde{\mathbf{u}}\rangle| \le C|\tilde{\mathbf{u}}|\|\tilde{\mathbf{u}}\|\|\mathbf{u}_1\| \le \epsilon\|\tilde{\mathbf{u}}\|^2 + C_\epsilon|\tilde{\mathbf{u}}|^2\|\mathbf{u}_1\|^2$. Hence from (6.23) we have

$$\begin{aligned} |\tilde{\mathbf{u}}(t\wedge\tau_N)|^2 + 2\int_0^{t\wedge\tau_N} \|\tilde{\mathbf{u}}(s)\|^2 ds &\le |\tilde{\mathbf{u}}(0)|^2 + \int_0^{t\wedge\tau_N} \|\tilde{\mathbf{u}}(s)\|^2\mathrm{d}s + \int_0^{t\wedge\tau_N} \|\sigma_d(s)\|^2_{\mathcal{L}_Q}\mathrm{d}s \\ &+ C_1\int_0^{t\wedge\tau_N} |\tilde{\mathbf{u}}(s)|^2\|\mathbf{u}_1(s)\|^2\mathrm{d}s + 2\int_0^{t\wedge\tau_N} (\sigma_d(s)\mathrm{d}W(s),\tilde{\mathbf{u}}(s)) \\ &+ 2\int_0^{t\wedge\tau_N}\int_Z (\tilde{\mathbf{u}}(s), g_d(s,z))\tilde{N}(\mathrm{d}s,\mathrm{d}z) + \int_0^{t\wedge\tau_N}\int_Z |g_d(s,z)|^2\lambda(\mathrm{d}s)\mathrm{d}z. \end{aligned} \tag{6.24}$$

Using Hypothesis 2.4 and then taking expectation we have,

$$\mathbb{E}\big[|\tilde{\mathbf{u}}(t\wedge\tau_N)|^2\big] + \mathbb{E}\big[\int_0^{t\wedge\tau_N} \|\tilde{\mathbf{u}}(s)\|^2 ds\big] \le \mathbb{E}\big[|\tilde{\mathbf{u}}(0)|^2\big] + \mathbb{E}\big[\int_0^{t\wedge\tau_N} (C_1\|\mathbf{u}_1(s)\|^2 + L)|\tilde{\mathbf{u}}(s)|^2\mathrm{d}s\big], \tag{6.25}$$

since $\int_0^{t\wedge\tau_N}(\sigma_d(s)\mathrm{d}W(s),\tilde{\mathbf{u}}(s))$ and $\int_0^{t\wedge\tau_N}\int_Z(\tilde{\mathbf{u}}(s),g_d(s,z))\tilde{N}(\mathrm{d}s,\mathrm{d}z)$ are local martingales having zero averages. Therefore, application of Gronwall's inequality produces

$$\mathbb{E}\big[|\tilde{\mathbf{u}}(t\wedge\tau_N)|^2\big]+\mathbb{E}\big[\int_0^{t\wedge\tau_N}\|\tilde{\mathbf{u}}(s)\|^2ds\big]\le\mathbb{E}\big[|\tilde{\mathbf{u}}(0)|^2\big]\exp\big(C_1\int_0^{t\wedge\tau_N}(\|\mathbf{u}_1(s)\|^2+L(t\wedge\tau_N)\big).\tag{6.26}$$

Letting $N\to\infty$, $t\wedge\tau_N\to t$ a.s. and we have

$$\mathbb{E}\big[|\tilde{\mathbf{u}}(t)|^2\big]+\mathbb{E}\big[\int_0^{t}\|\tilde{\mathbf{u}}(s)\|^2ds\big]\le\mathbb{E}\big[|\tilde{\mathbf{u}}(0)|^2\big]\exp\big(C_1\int_0^{t}(\|\mathbf{u}_1(s)\|^2+LT)\big).$$

As we have $\int_0^T\|\mathbf{u}_1(s)\|^2\mathrm{d}s<+\infty$, $\mathbb{P}-$ a.s. and $\tilde{\mathbf{u}}(0)=0$, $\mathbb{P}-$ a.s., we have $\mathbb{E}|\tilde{\mathbf{u}}(t)|^2=0$, *i.e.*, $\tilde{\mathbf{u}}(t)=0$, $\mathbb{P}-$ a.s. and this assures the pathwise uniqueness of the solution in two dimensions.

Case 2. In three dimensions we have the estimate

$$|\langle B(\tilde{\mathbf{u}},\mathbf{u}_1),\tilde{\mathbf{u}}\rangle|\le C|\tilde{\mathbf{u}}|^{1/2}\|\tilde{\mathbf{u}}\|^{3/2}\|\mathbf{u}_1\|\le\epsilon\|\tilde{\mathbf{u}}\|^2+C_\epsilon|\tilde{\mathbf{u}}|^2\|\mathbf{u}_1\|^4.$$

We have from (6.23)

$$\begin{aligned}|\tilde{\mathbf{u}}(t\wedge\tau_N)|^2&+2\int_0^{t\wedge\tau_N}\|\tilde{\mathbf{u}}(s)\|^2ds\le|\tilde{\mathbf{u}}(0)|^2+\int_0^{t\wedge\tau_N}\|\tilde{\mathbf{u}}(s)\|^2\mathrm{d}s+\int_0^{t\wedge\tau_N}\|\sigma_d(s)\|^2_{\mathcal{L}_Q}\mathrm{d}s\\&+C_1\int_0^{t\wedge\tau_N}|\tilde{\mathbf{u}}(s)|^2\|\mathbf{u}_1(s)\|^4\mathrm{d}s+2\int_0^{t\wedge\tau_N}(\sigma_d(s)\mathrm{d}W(s),\tilde{\mathbf{u}}(s))\\&+2\int_0^{t\wedge\tau_N}\int_Z(\tilde{\mathbf{u}}(s),g_d(s,z))\tilde{N}(\mathrm{d}s,\mathrm{d}z)+\int_0^{t\wedge\tau_N}\int_Z|g_d(s,z)|^2\lambda(\mathrm{d}s)\mathrm{d}z.\end{aligned}\tag{6.27}$$

Similarly, as in Case 1, we have

$$\mathbb{E}\big[|\tilde{\mathbf{u}}(t)|^2\big]+\mathbb{E}\big[\int_0^{t}\|\tilde{\mathbf{u}}(s)\|^2ds\big]\le\mathbb{E}\big[|\tilde{\mathbf{u}}(0)|^2\big]\exp\big(C_1\int_0^{t}(\|\mathbf{u}_1(s)\|^4+LT)\big).$$

Since we assume $\mathbb{E}\big[\int_0^T\|\mathbf{u}_1(s)\|^4\mathrm{d}s\big]<+\infty$, we have $\int_0^T\|\mathbf{u}_1(s)\|^4\mathrm{d}s<+\infty$, $\mathbb{P}-$a.s. As $\tilde{\mathbf{u}}(0)=0$, $\mathbb{P}-$ a.s., we have $\mathbb{E}|\tilde{\mathbf{u}}(t)|^2=0$, *i.e.*, $\tilde{\mathbf{u}}(t)=0$, $\mathbb{P}-$ a.s. and this assures the pathwise uniqueness of the solution in three dimensions. ☐

Theorem 6.11. *Let $d=2$ and Assumptions 2.1, 2.2, 2.7 and Hypothesis 2.4 be satisfied.*

(1) There exists a pathwise unique strong solution of (6.1)–(6.3).
(2) Moreover, if $(\Omega,\mathscr{F},\mathscr{F}_t,\mathbb{P},\{\mathbf{u}(t)\}_{t\ge0},\{N(t,\cdot)\}_{t\ge0},\{W(t)\}_{t\ge0},\{\mu_t\}_{t\ge0})$ is a strong solution of (6.1)–(6.3) then for $\mathbb{P}-$ almost all $\omega\in\Omega$ the trajectory $\mathbf{u}(\cdot,\omega)$ is equal almost everywhere to a càdlàg $\mathbb{H}$-valued function defined on $[0,T]$.
(3) The martingale solution of (6.1)–(6.3) is unique in law.

Proof. By Theorem 6.2 there exists a martingale solution to the system (6.1)–(6.3) and by Theorem 6.9 the solution is pathwise unique. Assertions (1) and (3) follow from Theorems 2 and 11 in Ondreját [58]. Assertion (2) is a direct consequence of Lemma 6.8. ☐

Corollary 6.12. *In view of Remark 6.10, it is clear that Theorem 6.11 holds true for Jeffreys fluid (6.12)–(6.13) in dimensions $d=2,3$.*

Appendix A. Compactness and tightness criteria

A.1 Certain compactness results

In this subsection, we provide some definitions and known results borrowed mostly from Aldous [5], Billingsley [9], Métivier [49] and Motyl [54]. Let $\mathbb{S}$ be a complete separable metric space with a metric ρ. Let us denote by $D([0,T];\mathbb{S})$, the set of all $\mathbb{S}$-valued functions defined on $[0,T]$, which are right continuous and have left limits (càdlàg functions) for every $t \in [0,T]$. The space $D([0,T];\mathbb{S})$ is endowed with the Skorokhod J-topology, which is defined below.

Definition A.1. The J-topology on $D([0,T];\mathbb{S})$ is generated by the following metric $\delta_{T,\mathbb{S}}$, see (formulae (12.13) and (12.16) of [9])

$$\delta_{T,\mathbb{S}}(x,y) := \inf_{\lambda \in \Lambda_T} \Big[\sup_{t\in[0,T]} \rho(x(t), y(\lambda(t))) + \sup_{t\in[0,T]} |t-\lambda(t)| + \sup_{s<t} \left| \log\left(\frac{\lambda(t)-\lambda(s)}{t-s} \right) \right| \Big], \tag{A.1}$$

where Λ_T is the set of all increasing homeomorphisms of $[0,T]$.

Remark A.2. Hence, it follows from Definition A.1, see also top of page 112 in [9], that sequence $(x_n)_{n\in\mathbb{N}}$ converges in $D([0,T];\mathbb{S})$ to x if and only if there exists a sequence (λ_n) in Λ_T such that with id being the identity map of $[0,T]$,

$$\lambda_n \to \text{ id and } x_n \circ \lambda_n \to x, \text{ uniformly on } [0,T]. \tag{A.2}$$

In particular, if an $D([0,T];\mathbb{S})$-valued sequence $(x_n)_{n\in\mathbb{N}}$ converges to x uniformly, then it converges to x in $D([0,T];\mathbb{S})$ as well.

For more details see Métivier (Chap. II of [49]) and Billingsley (Chap. 3 of [9]).

Definition A.3. Let $(\mathbb{S},\rho)$ be a separable and complete metric space. Let $u \in D([0,T];\mathbb{S})$ and let $\delta > 0$ be given. A modulus of continuity $w_{[0,T],\mathbb{S}}(u,\delta)$ of u is defined by

$$w_{[0,T],\mathbb{S}}(u,\delta) := \inf_{P\in\Pi_\delta} \max_{t_i\in P} \sup_{t_i\le s<t<t_{i+1}} \rho(u(t),u(s)), \tag{A.3}$$

where Π_δ is the set of all partitions $P = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ of $[0,T]$ with such that

$$t_{i+1} - t_i \ge \delta, \qquad i = 0,\ldots,n-1.$$

We have the following deterministic (càdlàg version of the Arzelà–Ascoli theorem) criterion for relative compactness in the space $D([0,T];\mathbb{S})$, see [41], (Chap. II of [49]) and (Chap. 3 of [9]).

Theorem A.4. *A set $A \subset D([0,T];\mathbb{S})$ has compact closure (in $D([0,T];\mathbb{S})$) iff it satisfies the following two conditions:*

(a) there exists a dense subset $J \subset [0,T]$ such that for every $t \in J$ the set $\{u(t),\ u\in A\}$ has compact closure in $\mathbb{S}$,
(b) $\lim_{\delta\to 0} \sup_{u\in A} w_{[0,T]}(u,\delta) = 0$.

A simple consequence of the above compactness criterion is the following.

Proposition A.5. *If $\mathbb{S}$ a compact metric space, then the set $D([0,T];\mathbb{S})$ (endowed with the Skorokhod J-topology) is compact.*

Let us consider, for a fixed $r > 0$, the closed ball in H centered at 0 of radius r: $\mathbb{B}_H(0,r) := \{x \in H : |x|_H \le r\}$. Let $\mathbb{B}_{H_w}$ denote the ball $\mathbb{B}_H$ endowed with the weak topology which, according to [12], is metrizable. Let q_r denote the metric on $\mathbb{B}_H$ that is compatible with the weak topology on $\mathbb{B}_H$. Consider the following functional space

$$D([0,T];\mathbb{B}_{H_w}(0,r)) := \left\{ u \in D([0,T];H_w) : \sup_{t\in[0,T]} |u(t)|_H \le r \right\}.$$

We endow the space $D([0,T];\mathbb{B}_{H_w})$ with the topology induced by the topology $\mathcal{T}_3$ on $D([0,T];H_w)$. Then $D([0,T];\mathbb{B}_{H_w})$ is metrizable with metric $\delta_{T,r}$ defined by

$$\delta_{T,r}(u,v) := \inf_{\lambda\in\Lambda_T} \left[\sup_{t\in[0,T]} q_r(u(t), v\circ\lambda(t)) + \sup_{t\in[0,T]} |t-\lambda(t)| + \sup_{s\neq t} \left| \log\left(\frac{\lambda(t)-\lambda(s)}{t-s} \right) \right| \right].$$

Since by the Banach–Alaoglu theorem, the topological space $\mathbb{B}_{H_w}$ is compact, in view of Proposition A.5, the space $(D([0,T];\mathbb{B}_{H_w}), \delta_{T,r})$ is a compact metric space.

The following result (see Lem. 4.3 in Motyl [53]) provides a criterion for convergence of a sequence in $D([0,T];\mathbb{B}_H(0,r))$.

Lemma A.6. *Let $X_n : [0,T] \to H,\ n\in\mathbb{N}$ be functions such that*

1. $\sup_{n\in\mathbb{N}} \sup_{s\in[0,T]} |X_n(s)|_H \le r$,
2. $X_n \to X$ *in* $D([0,T];\mathcal{U}')$.

Then $X, X_n \in D([0,T];\mathbb{B}_{H_w})$ and $X_n \to X$ in $D([0,T];\mathbb{B}_{H_w})$ as $n\to\infty$.

Theorem A.7. *Let $\mathcal{Z} = D([0,T];\mathcal{U}')_J \cap D([0,T];H_w) \cap \mathrm{L}^2_w(0,T;V) \cap \mathrm{L}^2(0,T;H)$ and let $\mathscr{T}$ be the supremum of the corresponding topologies. Then a set $\mathcal{K}\subset\mathcal{Z}$ is $\mathscr{T}-$relatively compact if the following three conditions are satisfied :*

(a) $\forall X\in\mathcal{K}$ *and all* $t\in[0,T], X(t)\in H$ *and* $\sup_{X\in\mathcal{K}} \sup_{s\in[0,T]} |X(s)| < \infty$,
(b) $\sup_{X\in\mathcal{K}} \int_0^T \|X(s)\|^2 \mathrm{d}s < \infty$, *i.e.* $\mathcal{K}$ *is bounded in* $\mathrm{L}^2(0,T;V)$,
(c) $\lim_{\delta\to 0} \sup_{X\in\mathcal{K}} w_{[0,T],\mathcal{U}'}(X,\delta) = 0$.

For proof see Lemma 3.3 in Brzeźniak and Motyl [15], Lemma 4.1 in Motyl [55], Theorem 2 of Motyl [54], Lemma 2.7 in Mikulevicius and Rozovskii [50].

A.2 Tightness criterion for predictable process

Let $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$ be a filtered probability space with a filtration $\{\mathscr{F}_t\}_{t\ge 0}$ satisfying the usual hypotheses.

Definition A.8. Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of càdlàg, $\mathcal{F}_t$-adapted stochastic processes in a Banach space E. Assume that for every $\varepsilon > 0$ and $\eta > 0$ there is $\delta > 0$ such that for every sequence $(\tau_n)_{n\in\mathbb{N}}$ of $\mathcal{F}_t$-stopping times with $\tau_n \le T$ one has

$$\sup_{n\in\mathbb{N}} \sup_{0<\theta\le\delta} \mathbb{P}\{\|X_n(\tau_n+\theta) - X_n(\tau_n)\|_E \ge \eta\} \le \varepsilon.$$

In this case, we say that $(X_n)_{n\in\mathbb{N}}$ satisfies the Aldous condition $[A]$.

We now quote the following lemma given in Métivier [49] and Motyl [54] which ensures the Aldous condition $[A]$ in a separable Banach space for the sequence $(X_n)_{n\in\mathbb{N}}$.

Lemma A.9. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of càdlàg, $\mathcal{F}_t$-adapted stochastic processes in a separable Banach space E. Assume that for every sequence $(\tau_n)_{n\in\mathbb{N}}$ of $\mathcal{F}_t$-stopping times with $\tau_n \le T$ and for every $n \in \mathbb{N}$ and $\theta \ge 0$ the following condition holds*

$$\mathbb{E}\left[\|X_n(\tau_n+\theta)-X_n(\tau_n)\|_E^{\alpha}\right] \le C\theta^{\zeta}, \tag{A.4}$$

for some $\alpha, \zeta > 0$ and some constant $C > 0$. Then the sequence $(X_n)_{n\in\mathbb{N}}$ satisfies the Aldous condition $[A]$ in E.

The following lemma (see [54], Lem. A.7) gives us a useful consequence of the Aldous condition $[A]$.

Lemma A.10. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of càdlàg, $\mathcal{F}_t$-adapted stochastic processes in a separable Banach space E, which satisfies the Aldous condition $[A]$. Then, for every $\varepsilon > 0$ there is a measurable subset $A_\varepsilon \subset D([0,T];E)$ with*

$$\mathbb{P}^{X_n}(A_\varepsilon) \ge 1-\varepsilon, \qquad \lim_{\delta\to 0}\sup_{u\in A_\varepsilon}\sup_{|t-s|\le\delta}\|u(t)-u(s)\|_E = 0.$$

The deterministic compactness result in Theorem *A.7* and the last lemma can be used to get the following criterion for tightness in $\mathcal{Z}$.

Theorem A.11. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of càdlàg $\mathscr{F}_t$-adapted $\mathcal{U}'$-valued processes such that*

(a) there exists a positive constant C_1 such that $\sup_{n\in\mathbb{N}}\mathbb{E}[\sup_{s\in[0,T]}|X_n(s)|] \le C_1$,
(b) there exists a positive constant C_2 such that $\sup_{n\in\mathbb{N}}\mathbb{E}[\int_0^T \|X_n(s)\|^2 \mathrm{d}s] \le C_2$,
(c) $(X_n)_{n\in\mathbb{N}}$ satisfies the Aldous condition $[A]$ in $\mathcal{U}'$.

Let $\mathbb{P}_n$ be the law of X_n on $\mathcal{Z}$. Then for every $\epsilon > 0$ there exists a compact subset K_ϵ of $\mathcal{Z}$ such that $\mathbb{P}_n(K_\epsilon) \ge 1-\epsilon$, and the sequence of measures $\{\mathbb{P}_n, n\in\mathbb{N}\}$ is said to be tight on $(\mathcal{Z},\mathscr{T})$.

For a proof see Corollary 1, Motyl [54]. In metric spaces, one can apply Prokhorov theorem (see [61], Thm. II.6.7) and Skorokhod theorem (see [9], Thm. 6.7) to obtain convergence from tightness. Since the space $\mathcal{Z}$ is a locally convex space, we use the following generalization of Skorokhod's theorem to nonmetric spaces.

Proposition A.12 (Skorokhod–Jakubowski)**.** *Let $\mathcal{X}$ be a topological space such that there is a sequence of continuous functions $f_m : \mathcal{X} \to \mathbb{C}$ that separates points of $\mathcal{X}$. Let $\mathcal{A}$ be the σ-algebra generated by $(f_m)_m$. Then, we have the following assertions:*

(a) Every compact set $K \subset \mathcal{X}$ is metrizable.
(b) Let $(\zeta_n)_{n\in\mathbb{N}}$ be a tight sequence of probability measures on $(\mathcal{X},\mathcal{A})$. Then, there are a subsequence $(\zeta_{n_k})_{k\in\mathbb{N}}$, random variables ξ_k, ξ for $k \in \mathbb{N}$ on a common probability space $(\bar\Omega,\bar{\mathcal{F}},\bar{\mathcal{F}}_t,\bar{\mathbb{P}})$ with $\bar{\mathbb{P}}^{\xi_k} = \zeta_{n_k}$ for $k\in\mathbb{N}$, and $\xi_k \to \xi$ $\bar{\mathbb{P}}$-almost surely for $k\to\infty$.

We stated Proposition *A.12* in the form of [17] (see also [40]) where it was first used to construct martingale solutions for stochastic evolution equations. We apply this result to the concrete situation and obtain the final result of this subsection.

Corollary A.13. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of adapted $\mathcal{U}'$-valued processes satisfying the Aldous condition $[A]$ in $\mathcal{U}'$ and*

$$\sup_{n\in\mathbb{N}}\mathbb{E}\left[\|X_n\|^2_{\mathrm{L}^\infty(0,T;H)}\right] < \infty, \quad \sup_{n\in\mathbb{N}}\mathbb{E}\left[\|X_n\|^2_{\mathrm{L}^2(0,T;V)}\right] < \infty.$$

Then, there are a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ and random variables $\bar X_k$, $\bar X$ for $k\in\mathbb{N}$ on a common probability space $(\bar\Omega,\bar{\mathcal{F}},\bar{\mathcal{F}}_t,\bar{\mathbb{P}})$with $\bar{\mathbb{P}}^{\bar X_k} = \mathbb{P}^{X_{n_k}}$ for $k\in\mathbb{N}$, and $\bar X_k \to \bar X$ $\bar{\mathbb{P}}$-almost surely in $\mathcal{Z}$ for $k\to\infty$.

Proof. We recall that

$$\mathcal{Z} = D([0,T];\mathcal{U}')_J \cap D([0,T];H_w) \cap \mathrm{L}^2_w(0,T;V) \cap \mathrm{L}^2(0,T;H)$$

is a locally convex space. Therefore, the assertion follows by an application of the Theorem *A*.11 and Proposition *A*.12 if for each of the spaces in the definition of $\mathcal{Z}$, we find a sequence $f_m : \mathcal{Z} \to \mathbb{R}$ of continuous functions separating points which generates the Borel σ-algebra.

Since $D([0,T];\mathcal{U}')$ and $\mathrm{L}^2(0,T;H)$ are separable and completely metrizable spaces, we infer that each of these spaces have this property.

For the space $\mathrm{L}^2_w(0,T;V)$, define

$$f_m(u) := \int_0^T (u(t), v_m(t))_V \mathrm{d}t \in \mathbb{R}, \quad u \in \mathrm{L}^2(0,T;V), \quad m \in \mathbb{N},$$

where $\{v_m, m \in \mathbb{N}\}$ is a dense subset of $\mathrm{L}^2(0,T;V)$. Then $(f_m)_{m\in\mathbb{N}}$ is a sequence of continuous real valued mappings separating points of the space $\mathrm{L}^2_w(0,T;V)$.

Let $\{h_m : m \in \mathbb{N}\}$ be a dense subset of H. We define the countable set $F := \{f_{m,t} : m \in \mathbb{N}, t \in [0,T] \cap \mathbb{Q}\}$ of functionals on $D([0,T];H_w)$ by

$$f_{m,t}(u) := (u(t), h_m)_H$$

for $m \in \mathbb{N}$, $t \in [0,T] \cap \mathbb{Q}$ and $u \in D([0,T];H_w)$.
The set F separates points, since for $u, v \in D([0,T];H_w)$ with $f_{m,t}(u) = f_{m,t}(v)$ for all $m \in \mathbb{N}$ and $t \in [0,T] \cap \mathbb{Q}$, we get $(u, h_m)_H = (v, h_m)_H$ on $[0,T]$ for all $m \in \mathbb{N}$ by continuous continuation and therefore $u = v$ on $[0,T]$.

Furthermore, the density of $\{h_m : m \in \mathbb{N}\}$ and the definition of the locally convex topology yield that $(f_{m,t})_{m\in\mathbb{N}, t\in[0,T]\cap\mathbb{Q}}$ generate the Borel σ-algebra on $D([0,T];H_w)$. □

A.3 Tightness criterion of random Young measure

This section is devoted to the class of Young measures on metrizable Suslin control sets and the ideas are borrowed mainly from Brzeźniak and Serrano [18], Castaing *et al.* [19] and Sritharan [69].

A.3.1 Preliminaries

In this paper, $\mathbb{U}$ denotes a Hausdorff topological space (the control set), $\mathcal{B}(\mathbb{U})$ denotes the Borel $\sigma-$algebra on $\mathbb{U}$ and $\mathsf{P}(\mathbb{U})$ denotes the set of all probability measures on $\mathcal{B}(\mathbb{U})$.

Definition A.14. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. A stochastic process $\{\mu_t\}_{t\geq 0}$ with values in $\mathsf{P}(\mathbb{U})$ is called a stochastic relaxed control (or relaxed control process) on $\mathbb{U}$ if the map $[0,T] \times \Omega \ni (t,\omega) \mapsto \mu_t(\omega,\cdot) \in \mathsf{P}(\mathbb{U})$ is measurable.

Definition A.15. Let $L(\cdot)$ denote the Lebesgue measure on $[0,T]$ and μ be a bounded nonnegative $\sigma-$additive measure on $\mathcal{B}(\mathbb{U} \times [0,T])$. μ is a Young measure on $\mathbb{U}$ iff μ satisfies $\mu(\mathbb{U} \times E) = L(E) \quad \forall E \in \mathcal{B}[0,T]$; *i.e.*, the marginal of μ on $\mathcal{B}[0,T]$ is equal to the Lebesgue measure L. We denote by $\mathcal{Y}(0,T;\mathbb{U})$ the set of Young measures on $\mathbb{U}$.

The following lemma elaborates disintegration of random Young measure which explicitly tells that to every Young measure, there is a stochastic relaxed control associated with it.

Lemma A.16. *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $\mathbb{U}$ be a Radon space. Let $\mu : \Omega \to \mathcal{Y}(0,T;\mathbb{U})$ be such that, for every $M \in \mathcal{B}(\mathbb{U} \times [0,T])$, the mapping $\Omega \ni \omega \mapsto \mu(\omega)(M) = \mu(\omega, M) \in [0,T]$ is measurable. Then there* *exists a stochastic relaxed control $\{\mu_t\}_{t\in[0,T]}$ on $\mathbb{U}$ such that for $\mathbb{P}-a.e$ $\omega\in\Omega$ we have*

$$\mu(\omega, C\times E) = \int_E \mu_t(\omega, C)\mathrm{d}t \quad \forall C\in\mathcal{B}(\mathbb{U}), E\in\mathcal{B}[0,T]. \tag{A.5}$$

Direct application of the previous lemma yields $\mu(\mathrm{d}U,\mathrm{d}t)=\mu_t(\mathrm{d}U)\mathrm{d}t$.

Definition A.17. The stable topology on $\mathcal{Y}(0,T;\mathbb{U})$ is the weakest topology on $\mathcal{Y}(0,T;\mathbb{U})$ for which the mappings $\mathcal{Y}(0,T;\mathbb{U})\ni\mu\mapsto\int_E\int_{\mathbb{U}} f(U)\,\mu(\mathrm{d}U,\mathrm{d}t)\in\mathbb{R}$ are continuous for every $E\in\mathcal{B}[0,T]$ and $f\in C_b(\mathbb{U})$.

Remark A.18. If $\mathbb{U}$ is metrizable (resp., metrizable Suslin), then the space $\mathcal{Y}(0,T;\mathbb{U})$ endowed with the stable topology is also metrizable (resp., metrizable Suslin).

Remark A.19. If $\mathbb{U}$ is Hausdorff, the level sets of every inf-compact function are closed, hence Borel measurable. Every inf-compact function is lower semicontinuous and the converse is also true if $\mathbb{U}$ is compact.

Remark A.20. We note that $\mathcal{N}\mu_t := \int_{\mathbb{U}}\mathsf{L}(U)\mu_t(\mathrm{d}U)$ is meaningful using Lemma A.16 and

$$\begin{aligned}\int_0^T|\mathcal{N}\mu_t|^2\mathrm{d}t &\le \int_0^T\Big(\int_{\mathbb{U}}|\mathsf{L}(U)|\mu_t(\mathrm{d}U)\Big)^2\mathrm{d}t \le \int_0^T\Big(\int_{\mathbb{U}}\kappa(t,U)\mu_t(\mathrm{d}U)\Big)^2\mathrm{d}t\\ &\le \int_0^T\Big(\int_{\mathbb{U}}\kappa^2(t,U)\mu_t(\mathrm{d}U)\Big)\Big(\int_{\mathbb{U}}\mu_t(\mathrm{d}U)\Big)\mathrm{d}t = \int_0^T\int_{\mathbb{U}}\kappa^2(t,U)\mu_t(\mathrm{d}U)\mathrm{d}t<+\infty.\end{aligned} \tag{A.6}$$

Further, if μ is a random Young measure such that $\mu\in\mathfrak{U}_{ad}(X_0,T)$, then,

$$\mathbb{E}\Big[\int_0^T|\mathcal{N}\mu_t|^2\mathrm{d}t\Big]\le\mathbb{E}\Big[\int_0^T\int_{\mathbb{U}}\kappa^2(t,U)\mu_t(\mathrm{d}U)\mathrm{d}t\Big]<+\infty.$$

A.3.2 Tightness criterion

This subsection briefly explains flexibility criterion for relative compactness in stable topology on a metrizable Suslin control set along with tightness criterion of sequence of random Young measures. The notion of tightness for Young measures that has been used in this work was introduced by Valadier [76] (see also Brzeźniak and Serrano [18], Castaing *et al.* [19], Crauel [23]).

Definition A.21. We say that a set $\mathcal{G}\subset\mathcal{Y}(0,T;\mathbb{U})$ is flexibly tight if, for each $\epsilon>0$, there exists a measurable set-valued mapping $[0,T]\ni t\mapsto K_t\subset\mathbb{U}$ such that K_t is compact for all $t\in[0,T]$ and $\sup_{\mu\in\mathcal{G}}\int_0^T\int_{\mathbb{U}}\mathbf{1}_{K_t^c}(U)\mu(\mathrm{d}U,\mathrm{d}t)<\epsilon$.

Below we state an equivalence theorem for flexible tightness. For proof see Balder [6].

Theorem A.22. *The following two conditions are equivalent for any $\mathcal{G}\subset\mathcal{Y}(0,T;\mathbb{U})$:*

1. *$\mathcal{G}$ is flexibly tight.*
2. *There exists $\kappa\in\mathcal{IC}([0,T],\mathbb{U})$ such that $\sup_{\mu\in\mathcal{G}}\int_0^T\int_{\mathbb{U}}\kappa(t,U)\mu(\mathrm{d}U,\mathrm{d}t)<+\infty$.*

Next we state the Prohorov criterion for relative compactness (see Thm. 4.3.5 of Castaing *et al.* [19]).

Theorem A.23. *Let $\mathbb{U}$ be a metrizable and Suslin space. Then every flexibly tight subset of $\mathcal{Y}(0,T;\mathbb{U})$ is sequentially relatively compact in the stable topology.*

Now we proceed to theorem which guaranties tightness criterion of random Young measures.

Theorem A.24. *Assume that $\mathbb{U}$ is metrizable and Suslin. For each $n \in \mathbb{N}$, let μ_n be a random Young measure on $\mathbb{U}$ defined on $(\Omega^n, \mathscr{F}^n, \mathbb{P}^n)$. Assume there exists $\kappa \in \mathcal{IC}([0,T], \mathbb{U})$ with*

$$\mathbb{E}^n \int_0^T \int_{\mathbb{U}} \kappa(t,U)\, \mu_n(\mathrm{d}U, \mathrm{d}t) \le R \quad \forall n \in \mathbb{N}$$

for some $R > 0$, where $\mathbb{E}^n$ denotes expectation with respect to $\mathbb{P}^n$. Then the family of laws of $\{\mu_n\}_{n\in\mathbb{N}}$ is tight on $\mathcal{Y}(0,T;\mathbb{U})$.

Proof. The proof follows from Lemma 2.18 of Brzeźniak and Serrano [18]. Let us provide a sketch of the proof. For each $\epsilon > 0$, define the set

$$K_\epsilon := \Big\{\mu \in \mathcal{Y}(0,T;\mathbb{U}) : \int_0^T \int_{\mathbb{U}} \kappa(t,U)\mu(\mathrm{d}U,\mathrm{d}t) \le \frac{R}{\epsilon}\Big\}.$$

By Theorems A.22 and A.23, K_ϵ is relatively compact in the stable topology of $\mathcal{Y}(0,T;\mathbb{U})$. Therefore by Chebyshev's inequality we have

$$\mathbb{P}^n(\mu_n \in \mathcal{Y} \setminus \bar{K}_\epsilon) \le \mathbb{P}^n(\mu_n \in \mathcal{Y} \setminus K_\epsilon) \le \frac{\epsilon}{R}\mathbb{E}^n \int_0^T \int_{\mathbb{U}} \kappa(t,U)\, \mu_n(\mathrm{d}U,\mathrm{d}t) \le \epsilon,$$

where $\bar{K}_\epsilon$ denotes the closure of K_ϵ. Hence, the tightness of the laws of $\{\mu_n\}_{n\ge 1}$ follows. □

The following lemma states that the relaxed cost functional is lower semicontinuous.

Lemma A.25. *Let $\mathbb{U}$ be a metrizable Suslin space, and let $\gamma \in \mathrm{L}^1(0,T;\mathbb{R})$. Let us assume that $F : [0,T] \times H \times \mathbb{U} \to [0,\infty]$ is measurable in $t \in [0,T]$ and lower semicontinuous for every $t \in [0,T]$ with respect to $(X,U) \in H \times \mathbb{U}$ and satisfies one of the following two conditions:*

1. *$|F(t,X,U)| \le \gamma(t)$, a.e. $t \in [0,T]$.*
2. *$F \ge 0$. Then we have*

(a) If $\mu_n \to \mu$ stably in $\mathcal{Y}(0,T;\mathbb{U})$, then

$$\int_0^T \int_{\mathbb{U}} F(t,X,U)\mu(\mathrm{d}U,\mathrm{d}t) \le \liminf_{n\to\infty} \int_0^T \int_{\mathbb{U}} F(t,X,U)\mu_n(\mathrm{d}U,\mathrm{d}t).$$

(b) If in addition $X_n \to X$ in $\mathscr{T}-$topology, then

$$\int_0^T \int_{\mathbb{U}} F(t,X,U)\mu(\mathrm{d}U,\mathrm{d}t) \le \liminf_{n\to\infty} \int_0^T \int_{\mathbb{U}} F(t,X_n,U)\mu_n(\mathrm{d}U,\mathrm{d}t).$$

Proof. For (a) if condition 1 holds, the result follows from Theorem 2.1.3-Part G in Castaing *et al.* [19] and for condition 2, the result follows using Proposition 2.1.12-Part (d) in Castaing *et al.* [19]. Also see Brzeźniak and Serrano [18]. Proof of (b) is from Lemma 10 of Sritharan [69], Haussmann and Lepeltier [38], Jacod and Memin [39]. □

Appendix B. Some proofs

Proof of Lemma 2.3. Let $I = \int_0^t \int_0^s \beta(s-\tau)\varphi(s)\,\varphi(\tau)\mathrm{d}s\,d\tau$. Extend β by $\tilde{\beta}$ on the negative semi-axis by an even function, *i.e.*,

$$\tilde{\beta}(t) = \begin{cases} \beta(t), & \text{if} \quad t \geq 0, \\ \beta(-t), & \text{if} \quad t \leq 0. \end{cases}$$

Then, $I = \frac{1}{2}\int_0^t \int_0^t \tilde{\beta}(s-\tau)\,\varphi(s)\,\varphi(\tau)\,\mathrm{d}s\,d\tau$. Again, define $\tilde{\varphi}$ by φ on $(0,t)$ and zero elsewhere. Hence, by using the properties of convolution of two functions (denoted by $\star$) and Plancherel's theorem, one gets

$$\begin{aligned} I &= \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \tilde{\beta}(x-y)\,\tilde{\varphi}(x)\,\tilde{\varphi}(y)\,dx\,dy = \frac{1}{2}\int_{-\infty}^{\infty} \tilde{\varphi}(x)\left(\tilde{\beta}\star\tilde{\varphi}(x)\right)dx \\ &= \frac{1}{2}\left(\tilde{\varphi}(\cdot), (\tilde{\beta}\star\tilde{\varphi})(\cdot)\right)_{\mathrm{L}^2} = \frac{1}{2}\left(\hat{\tilde{\varphi}}(\cdot), \widehat{(\tilde{\beta}\star\tilde{\varphi})}(\cdot)\right)_{\mathrm{L}^2} = \frac{1}{2}\int_{-\infty}^{\infty} |\hat{\tilde{\varphi}}(\xi)|^2\hat{\tilde{\beta}}(\xi)\,d\xi. \end{aligned}$$

Using the definition of Fourier transform,

$$\begin{aligned} \hat{\tilde{\beta}}(\xi) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ix\xi}\tilde{\beta}(x)dx = \frac{1}{\sqrt{2\pi}}\int_0^{\infty}(e^{-ix\xi}+e^{ix\xi})\beta(x)dx = \frac{2}{\sqrt{2\pi}}\int_0^{\infty}\cos x\xi\,\beta(x)dx \\ &= \frac{2}{\sqrt{2\pi}}\Big(\left[\frac{\beta(x)\,\sin x\xi}{\xi}\right]_0^{\infty} - \int_0^{\infty}\frac{\beta'(x)\,\sin x\xi}{\xi}\,dx\,\Big) = -\frac{2}{\sqrt{2\pi}}\int_0^{\infty}\frac{\beta'(x)\,\sin x\xi}{\xi}\,dx \end{aligned}$$

and the last step is true as $\left[\frac{\beta(x)\,\sin x\xi}{\xi}\right]_0^{\infty} = 0$ because of Assumption 2.2 (ii). Exploiting the fact that Fourier-sine transform of a decreasing function, if it exists, is always non-negative (see Tuck [75]) and using Assumption 2.2 (i), we have $\hat{\tilde{\beta}} \geq 0$, which directly imlpies $I \geq 0$. □

B.1 A result on random Young measure

The following lemma lists certain properties of the limiting random Young measure in the stable topology, which has been used to prove Theorem 5.4.

Lemma B.1. *Assume that* $\mathbb{U}$ *is metrizable and Suslin. Let* $(\tilde{\mu}_n)_{n\in\mathbb{N}}$ *and* $\tilde{\mu}$ *be random Young measures on* $\mathbb{U}$ *defined on* $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathscr{F}}_t, \tilde{\mathbb{P}})$. *Let* $\tilde{\mu}_n \to \tilde{\mu}$ *in the stable topology* $\tilde{\mathbb{P}}$ *– a.s. and assume there exists* $\kappa \in \mathcal{IC}([0,T],\mathbb{U})$ *such that*

$$\tilde{\mathbb{E}}\int_0^T\int_{\mathbb{U}}\kappa^2(t,U)\,\tilde{\mu}_n(\mathrm{d}U,\mathrm{d}t) \leq C, \quad \forall n\in\mathbb{N} \tag{B.1}$$

for some $C>0$. *Assume* L *as the admissible control operator. Then we have the following convergences, for all* $0<t<T$, *and* $\forall z\in \mathrm{L}^4(\tilde{\Omega};H)$,

$$\lim_{n\to\infty}\tilde{\mathbb{E}}\left|\int_0^t\int_{\mathbb{U}}(\mathsf{L}(U),z)\,\tilde{\mu}_n(\mathrm{d}U,\mathrm{d}s) - \int_0^t\int_{\mathbb{U}}(\mathsf{L}(U),z)\,\tilde{\mu}(\mathrm{d}U,\mathrm{d}s)\right| = 0. \tag{B.2}$$

If in addition, we assume $\boldsymbol{y}_n \to \boldsymbol{y}$ strongly in $\mathrm{L}^4(\tilde{\Omega}; \mathrm{L}^2(0,T;H))$, *then*

$$\lim_{n\to\infty} \tilde{\mathbb{E}}\left[\int_0^T \int_{\mathbb{U}} (\boldsymbol{y}_n(t), \mathsf{L}(U))\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t) - \int_0^T \int_{\mathbb{U}} (\boldsymbol{y}(t), \mathsf{L}(U))\tilde{\mu}(\mathrm{d}U, \mathrm{d}t)\right] = 0. \tag{B.3}$$

Proof. Choose $\psi_R \in C_b(\mathbb{U})$ such that

$$\psi_R(U) = \begin{cases} 1 & \text{if} \quad \kappa(\cdot, U) \leq R, \\ 0 & \text{if} \quad \kappa(\cdot, U) \geq 2R. \end{cases}$$

$$\begin{aligned}
\text{Now} \quad & \Big|\int_0^t \int_{\mathbb{U}} (\mathsf{L}(U), z)\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s) - \int_0^t \int_{\mathbb{U}} (\mathsf{L}(U), z)\, \tilde{\mu}(\mathrm{d}U, \mathrm{d}s)\Big| \\
&\leq \int_0^t \int_{\mathbb{U}} |(1-\psi_R(U))\, (\mathsf{L}(U), z)\, |\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s) + \int_0^t \int_{\mathbb{U}} |(1-\psi_R(U))\, (\mathsf{L}(U), z)\, |\tilde{\mu}(\mathrm{d}U, \mathrm{d}s) \\
&\quad + \Big|\int_0^t \int_{\mathbb{U}} \psi_R(U)\, (\mathsf{L}(U), z)\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s) - \int_0^t \int_{\mathbb{U}} \psi_R(U)\, (\mathsf{L}(U), z)\, \tilde{\mu}(\mathrm{d}U, \mathrm{d}s)\Big| \\
&=: I_1 + I_2 + I_3.
\end{aligned}$$

For each $s \in [0,T]$, let us define the level set $\mathbb{K}_R := \{U \in \mathbb{U} : \kappa(\cdot, U) \geq R\}$. We then have, using Remark A.20,

$$\begin{aligned}
I_1 &\leq \Big[\int_0^t \int_{\mathbb{K}_R} |z|\kappa^{-1}(s,U)\kappa^2(s,U)\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s)\Big] \leq \frac{1}{R}\Big[\int_0^t \int_{\mathbb{K}_R} |z|\kappa^2(s,U)\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s)\Big] \\
&\leq \frac{c}{R}|z| \to 0 \quad \text{as} \quad R \to \infty \quad \text{for some constant} \quad c > 0.
\end{aligned}$$

Similar calculation yields $I_2 \to 0$. Stable convergence of $\tilde{\mu}_n$ yields $\int_0^t \int_{\mathbb{U}} \psi_R(U)\, (\mathsf{L}(U), z)\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s) \to \int_0^t \int_{\mathbb{U}} \psi_R(U)\, (\mathsf{L}(U), z)\, \tilde{\mu}(\mathrm{d}U, \mathrm{d}s)$. Combining all the above we have $\int_0^t \int_{\mathbb{U}} (\mathsf{L}(U), z)\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s) \to \int_0^t \int_{\mathbb{U}} (\mathsf{L}(U), z)$ $\tilde{\mu}(\mathrm{d}U, \mathrm{d}s)$. Further using Hölder's inequality, Assumption 2.7 and $\tilde{\mu}_n$ being the probability measure, we have for $1 < r < 2$

$$\begin{aligned}
\tilde{\mathbb{E}}\Big|\int_0^t \int_{\mathbb{U}} (\mathsf{L}(U), z)\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s)\Big|^r &\leq \tilde{\mathbb{E}}\Big|\int_0^T \int_{\mathbb{U}} |\mathsf{L}(U)|\, |z|\, \tilde{\mu}_t^n(\mathrm{d}U)\mathrm{d}t\Big|^r \\
&\leq \tilde{\mathbb{E}}\Big[|z|^r \Big\{\int_0^T \int_{\mathbb{U}} \kappa^2(t,U)\tilde{\mu}_t^n(\mathrm{d}U)\mathrm{d}t\Big\}^{r/2}\Big\{\int_0^T \int_{\mathbb{U}} \tilde{\mu}_t^n(\mathrm{d}U)\mathrm{d}t\Big\}^{r/2}\Big] \\
&\leq T^{\frac{r}{2}}\frac{r}{2}\tilde{\mathbb{E}}\Big[\int_0^T \int_{\mathbb{U}} \kappa^2(t,U)\tilde{\mu}_t^n(\mathrm{d}U)\mathrm{d}t\Big] + T^{\frac{r}{2}}\frac{2-r}{2}\tilde{\mathbb{E}}\Big[|z|^{\frac{2r}{2-r}}\Big].
\end{aligned}$$

In particular, choosing $r = \frac{4}{3}$ we have $\tilde{\mathbb{E}}\big[\big|\int_0^t \int_{\mathbb{U}} (\mathsf{L}(U), z)\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s)\big|^{\frac{4}{3}}\big] < C_r$ which ensures the uniform integrability of $\Big|\int_0^t \int_{\mathbb{U}} (\mathsf{L}(U), z)\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}s)\Big|$.
Hence, by Vitali's theorem, we have (B.2).
In addition, let us assume $\mathbf{y}_n \to \mathbf{y}$ strongly in $\mathrm{L}^4(\tilde{\Omega}; \mathrm{L}^2(0,T;H))$. We have,

$$I_n := \int_0^T \int_{\mathbb{U}} (\mathbf{y}_n(t), \mathsf{L}(U))\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t) - \int_0^T \int_{\mathbb{U}} (\mathbf{y}(t), \mathsf{L}(U))\, \tilde{\mu}(\mathrm{d}U, \mathrm{d}t)$$

$$= \underbrace{\int_0^T \int_{\mathbb{U}} (\mathbf{y}(t) - \mathbf{y}_n(t), \mathsf{L}(U))\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t)}_{I_n^1} - \underbrace{\int_0^T \int_{\mathbb{U}} (\mathbf{y}(t), \mathsf{L}(U))\, (\tilde{\mu}(\mathrm{d}U, \mathrm{d}t) - \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t))}_{I_n^2}.$$

$$|I_n^1| \le \int_0^T \int_{\mathbb{U}} |\mathbf{y}_n(t) - \mathbf{y}(t)||\mathsf{L}(U)|\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t) = \int_0^T |\mathbf{y}_n(t) - \mathbf{y}(t)|\Big(\int_{\mathbb{U}} \kappa(t, U)\tilde{\mu}_t^n(\mathrm{d}U)\Big)\mathrm{d}t$$
$$\le c\|\mathbf{y}_n - \mathbf{y}\|_{\mathrm{L}^2(0,T;H)} \to 0 \quad \text{as} \quad n \to \infty \quad \tilde{\mathbb{P}} - \text{a.s.}$$

We now move to estimate I_n^2. Since, $\mathbf{y}(\cdot) \in \mathrm{L}^4(\tilde{\Omega}; \mathrm{L}^2(0, T; H))$ is strongly measurable, there exists a sequence of H-valued simple functions, $\mathbf{y}_k(t) := \sum_{l=1}^m c_{k,l}\chi_{k,l}(t)$, $c_{k,l} \in H$, such that $\mathbf{y}_k(\cdot) \to \mathbf{y}(\cdot)$ in $\mathrm{L}^4(\tilde{\Omega}, \mathrm{L}^2(0, T; H))$, and also for t in $[0, T]$ a.s., $\tilde{\mathbb{P}}$ – a.s., $\mathbf{y}_k(t) \to \mathbf{y}(t)$ in H. Hence arguing as before, we conclude that $I_{n,2,k} \to I_{n,2}$ as $k \to \infty$ where

$$\begin{aligned} I_{n,2,k} :&= \int_0^T \int_{\mathbb{U}} (\mathbf{y}_k(t), \mathsf{L}(U))\, (\tilde{\mu}(\mathrm{d}U, \mathrm{d}t) - \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t)) \\ &= \sum_{l=1}^{m_k} \left\{ \int_{\chi_{k,l}} \int_{\mathbb{U}} (c_{k,l}, \mathsf{L}(U))\tilde{\mu}(\mathrm{d}U, \mathrm{d}t) - \int_{\chi_{k,l}} \int_{\mathbb{U}} (c_{k,l}, \mathsf{L}(U))\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t) \right\}. \end{aligned} \tag{B.4}$$

Also by (B.2) we have $I_{n,2,k} \to 0$ as $n \to \infty$. Hence we proved that $I_{n,2} \to 0$ as $n \to \infty$. Therefore, we have $I_n \to 0$ as $n \to \infty$. We note that using

$$|\int_0^T \int_{\mathbb{U}} (\mathbf{y}(t), \mathsf{L}(U))\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t)| \le \left(\int_0^T |\mathbf{y}(t)|^2 \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_0^T \int_{\mathbb{U}} \kappa^2(t, U)\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t)\right)^{\frac{1}{2}} \tag{B.5}$$

and exploiting Young's inequality (with exponents $\frac{3}{2}$ and 3), (B.1), and the hypothesis $\mathbf{y}_n \in \mathrm{L}^4(\Omega; \mathrm{L}^2(0, T; H))$, is uniformly bounded (as convergent) the above inequality reduces to

$$\tilde{\mathbb{E}}\Big[|\int_0^T \int_{\mathbb{U}} (\mathbf{y}_n(t), \mathsf{L}(U))\, \tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t)|^{\frac{4}{3}}\Big] \le \tilde{\mathbb{E}}\Big[\left(\int_0^T |\mathbf{y}_n(t)|^2 \mathrm{d}t\right)^{\frac{2}{3}} \left(\int_0^T \int_{\mathbb{U}} \kappa^2(t, U)\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t)\right)^{\frac{2}{3}}\Big]$$
$$\le \frac{1}{3}\tilde{\mathbb{E}}\Big[\left(\int_0^T |\mathbf{y}_n(t)|^2 \mathrm{d}t\right)^2\Big] + \frac{2}{3}\mathbb{E}\Big[\left(\int_0^T \int_{\mathbb{U}} \kappa^2(t, U)\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t)\right)\Big] \le \frac{2C}{3} + \frac{1}{3}\tilde{\mathbb{E}}\Big[\left(\int_0^T |\mathbf{y}_n(t)|^2 \mathrm{d}t\right)^2\Big] < \infty.$$

Hence, by Vitali Theorem,

$$\lim_{n\to\infty} \tilde{\mathbb{E}} I_n = \lim_{n\to\infty} \tilde{\mathbb{E}} \left[\int_0^T \int_{\mathbb{U}} (\mathbf{y}_n(t), \mathsf{L}(U))\tilde{\mu}_n(\mathrm{d}U, \mathrm{d}t) - \int_0^T \int_{\mathbb{U}} (\mathbf{y}(t), \mathsf{L}(U))\tilde{\mu}(\mathrm{d}U, \mathrm{d}t)\right] = 0.$$

□

Appendix C. Quadratic variation process

Definition C.1 (Quadratic variation process and Meyer process)**.** Let M be a square integrable martingale with right continuous paths with values in a separable Hilbert space H on $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$. Then there exists two real right continuous increasing processes $[M]$ and $\lhd M \rhd$ with $0 = [M]_0 = \lhd M \rhd_0$ such that

$$\|M_t\|_H^2 = \|M_0\|_H^2 + 2\int_0^t (M_{s-}, \mathrm{d}M_s)_H + [M]_t. \tag{C.1}$$

$\lhd M \rhd$ is the unique real right continuous increasing predictable process such that

$$\|M_t\|_H^2 - \|M_0\|_H^2 - \lhd M \rhd_t \text{ is a martingale.} \tag{C.2}$$

Here $[M]$ is called the quadratic variation of M and $\lhd M \rhd$ the Meyer process of M.

Remark C.2. Let H be Hilbert spaces and let $Q : H \to H$ be a trace class operator. Let

$$\mathrm{d}X(t) = \sigma(t, X)\mathrm{d}W(t) + \int_Z g(X(t-), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z),$$

where $W(\cdot)$ is an H-valued Wiener process, $\sigma(\cdot,\cdot) : [0,t) \times H \to \mathcal{L}_Q(H,H)$, Z is a measurable subspace of H, $g(\cdot,\cdot) : H \times Z \to H$ and $\tilde{N}(\cdot,\cdot)$ is the compensated Poisson random measure. From Métivier [49]), Sakthivel and Sritharan [67] and Manna, Manil and Sritharan [47], we observe that the quadratic variation process of X is

$$[X]_t = \int_0^t \|\sigma(s, X)\|^2_{\mathcal{L}_Q(H,H)} \mathrm{d}s + \int_0^t \int_Z \|g(X, z)\|_H^2 N(\mathrm{d}s, \mathrm{d}z)$$

and the Meyer process of X is

$$\lhd X \rhd_t = \int_0^t \|\sigma(s, X)\|^2_{\mathcal{L}_Q(H,H)} \mathrm{d}s + \int_0^t \int_Z \|g(X, z)\|_H^2 \lambda(\mathrm{d}z)\mathrm{d}s,$$

and finally by the martingale property one obtains that the expectation of the quadratic variation process and that of Meyer process are same. Thus we have,

$$\begin{aligned} &\mathbb{E}\left[\int_0^t \|\sigma(s, X)\|^2_{\mathcal{L}_Q(H,H)} \mathrm{d}s + \int_0^t \int_Z \|g(X, z)\|_H^2 N(\mathrm{d}s, \mathrm{d}z)\right] \\ &= \mathbb{E}\left[\int_0^t \|\sigma(s, X)\|^2_{\mathcal{L}_Q(H,H)} \mathrm{d}s + \int_0^t \int_Z \|g(X, z)\|_H^2 \lambda(\mathrm{d}z)\mathrm{d}s\right]. \end{aligned} \tag{C.3}$$

Lemma C.3. *(Burkholder–Davis–Gundy inequality) Let H be a Hilbert space valued càdlàg martingale with $M_0 = 0$ and let $p \geq 1$ be fixed. Then for any $\mathscr{F}$-stopping time τ, there exists constants c_p and C_p such that*

$$\mathbb{E}\left\{[M]_\tau^{p/2}\right\} \leq c_p \mathbb{E}\left\{\sup_{0\leq t\leq\tau} \|M_t\|_H^p\right\} \leq C_p \mathbb{E}\left\{[M]_\tau^{p/2}\right\}$$

for all τ, $0 \leq \tau \leq \infty$, where $[M]$ is the quadratic variation of process M. The constants are universal (independent of M).

For proof see Theorem 1.1 of Marinelli and Röckner [48]. For real-valued càdlàg martingales see Theorem 3.50 of Peszat and Zabczyk [62].

Acknowledgements. Authors would like to thank Professor Zdzisław Brzeźniak from University of York for useful comments and discussions. Also the authors would like to sincerely thank the anonymous referees for their valuable comments and criticisms, which led to the improvement of this paper.

References

[1] N. Agram and B. Oksendal, Malliavin calculus and optimal control of stochastic Volterra equations. *J. Optim. Theory Appl.* **167** (2015) 1070–1094.

[2] S. Aizicovici and V. Barbu, Existence and asymptotic results for a system of integro-partial differential equations. *NoDEA Nonlin. Differ. Equ. Appl.* **3** (1996) 1–18.

[3] S. Aizicovici, M. Grasselli and M. McKibben, A hyperbolic integrodifferential system related to phase-field models. *Adv. Math. Sci. Appl.* **10** (2000) 601–625.

[4] S. Albeverio, Z. Brzeźniak and J.L. Wu, Existence of global solutions and invariant measures for stochastic differential equations driven by Poisson type noise with non-Lipschitz coefficients. *J. Math. Anal. Appl.* **371** (2010) 309–322.

[5] D. Aldous, Stopping times and tightness. *Ann. Probab.* **6** (1978) 335–340.

[6] E.J. Balder, Lectures on Young measure theory and its applications in economics. *Rend. Inst. Mat. Univ. Trieste* **31** (2000) 1–69.

[7] V. Barbu, S. Bonaccorsi and L. Tubaro, Existence and asymptotic behavior for hereditary stochastic evolution equations. *Appl. Math. Optim.* **69** (2014) 273–314.

[8] V. Barbu, Z. Brzeźniak, E. Hausenblas, L. Tubaro, Existence and convergence results for infinite dimensional nonlinear stochastic equations with multiplicative noise. *Stoch. Process. Appl.* **123** (2013) 934–951.

[9] P. Billingsley, Convergence of Probability Measures. Wiley, New York (1999).

[10] S. Bonaccorsi, F. Confortola and E. Mastrogiacomo, Optimal control for stochastic Volterra equations with completely monotone kernels. *SIAM J. Control Optim.* **50** (2012) 748–789.

[11] S. Bonaccorsi and W. Desch, Volterra equations perturbed by noise. *NoDEA Nonlinear Differ. Equ. Appl.* **20** (2013) 557–594.

[12] H. Brezis. Analyse Fonctionnelle. Masson, Paris (1983).

[13] Z. Brzeźniak, E. Hausenblas and P.A. Razafimandimby, Stochastic reaction-diffusion equations driven by jump processes. *Potential Anal.* **49** (2018) 131–201.

[14] Z. Brzeźniak, E. Hausenblas and J. Zhu, 2D stochastic Navier-Stokes equations driven by jump noise. *Nonlinear Anal.* **79** (2013) 122–139.

[15] Z. Brzeźniak and E. Motyl, Existence of a martingale solution of the stochastic Navier-Stokes equations in unbounded 2D and 3D domains. *J. Differ. Equ.* **254** (2013) 1627–1685.

[16] Z. Brzeźniak, E. Motyl and M. Ondrejàt, Invariant measure for the stochastic Navier-Stokes equations in unbounded 2D domains. *Ann. Probab.* **45** (2017) 3145–3201.

[17] Z. Brzeźniak and M. Ondreját, Stochastic geometric wave equations with values in compact Riemannian homogeneous spaces. *Ann. Probab.* **41** (2013) 1938–1977.

[18] Z. Brzeźniak and R. Serrano, Optimal relaxed control of dissipative stochastic partial differential equations in Banach spaces. *SIAM J. Control Optim.* **51** (2013) 2664–2703.

[19] C. Castaing, P. Raynaud de Fitte and M. Valadier, Young Measures on Topological Spaces: With Applications in Control Theory and Probability Theory. Vol. 571 of *Mathematics and Its Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands (2004).

[20] M.-H. Chang, Stochastic control of hereditary systems and applications. In Vol. 59 of *Stochastic Modelling and Applied Probability*. Springer, New York (2008).

[21] Ph. Clément and G. Da Prato, White noise perturbation of the heat equation in materials with memory. *Dyn. Syst. Appl.* 6, (1997) 441–460.

[22] F. Confortola and E. Mastrogiacomo, Optimal control for stochastic heat equation with memory. *Evol. Equ. Control Theory* **3** (2014) 35–58.

[23] H. Crauel, Random Probability Measures on Polish Spaces. Vol. 11 of *Stochastics Monographs*. Taylor & Francis, London (2002).

[24] S.C. Crow, Viscoelastic properties of fine-grained incompressible turbulence. *J. Fluid Mech.* **33** (1968) 1–20.

[25] N.J. Cutland and K. Grzesiak, Optimal control for 3D stochastic Navier-Stokes equations. *Stochastics* **77** (2005) 437–454.

[26] N.J. Cutland and K. Grzesiak, Optimal control for two-dimensional stochastic Navier-Stokes equations. *Appl. Math. Optim.* **55** (2007) 61–91.

[27] A. Doubova and E. Fernández-Cara, On the control of viscoelastic Jeffreys fluids. *Syst. Control Lett.* **61** (2012) 573–579.

[28] N.H. El Karoui, D. Nguyen and M. Jeanblanc-Picqué, Compactification methods in the control of degenerate diffusions: existence of an optimal control. *Stochastics* **20** (1987) 169–219.

[29] W.H. Fleming, Measure-valued processes in the control of partially-observable stochastic systems. *Appl. Math. Optim.* **6** (1980) 271–285.

[30] W.H. Fleming and M. Nisio, On stochastic relaxed control for partially observed diffusions. *Nagoya Math. J.* **93** (1984) 71–108.

[31] J. Fort and V. Méndez, Wavefront in time-delayed reaction-diffusion systems. Theory and comparison to experiments. *Rep. Prog. Phys.* **65** (2002) 895–954.

[32] D. Gatarek and J. Sobczyk, On the existence of optimal controls of Hilbert space-valued diffusions. *SIAM J. Control Optim.* **32** (1994) 170–175.

[33] C.H. Gibson, Turbulence in the ocean, atmosphere, galaxy, and universe. *Appl. Mech. Rev.* **49** (1996) 299–315.

[34] C. Giorgi and V. Pata, Asymptotic behavior of a nonlinear hyperbolic heat equation with memory. *NoDEA Nonlinear Differ. Equ. Appl.* **8** (2001) 157–171.

[35] M. Grasselli and V. Pata, A reaction-diffusion equation with memory. *Discrete Continuous Dyn. Syst.* **15** (2006) 1079–1088.

[36] M. Grasselli and V. Pata, Upper semicontinuous attractor for a hyperbolic phase-field model with memory. *Indiana Univ. Math. J.* **50** (2001) 1281–1308.

[37] I. Gyöngy, On stochastic equations with respect to semimartingales III. *Stochastics* **7** (1982) 231–254.

[38] U.G. Haussmann and J.-P. Lepeltier, On the existence of optimal controls. *SIAM J. Control Optim.* **28** (1990) 851–902.

[39] J. Jacod and J. Memin, Sur un type de convergence intermediaire entre la convergence en loi et la convergence en probabilite. Seminaire de Probabilites XV, edited by J. Azema and M. Yor. Vol. 850 of *Lecture Notes in Mathematics*. Springer-Verlag, New York (1979/1980) 529–546.

[40] A. Jakubowski, The almost sure Skorohod representation for subsequences in nonmetric spaces. *Teor. Veroyatnost. i Primenen.* **42** (1997) 209–216. [Translation in *Theory Probab. Appl.* **42** (1998) 167–174].

[41] A. Joffe and M. Métivier, Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. Appl. Prob.* **18** (1986) 20–65.

[42] D.D. Joseph, Fluid dynamics of Viscoelastic liquids. Springer Verlag, New York (1990).

[43] I. Klapper, C.J. Rupp, R. Cargo, B. Purvedorj and P. Stoodley, Viscoelastic fluid description of bacterial biofilm material properties. *Biotechnol. Bioeng.* **80** (2002) 289–296.

[44] C. Kuratowski, Topologie, Vol. I, 3rd edn, Monografie Matematyczne XX. Polskie Towarzystwo Matematyczne, Warsawa (1952).

[45] J.L. Lions, Optimal control of systems governed by partial differential equations. Translated from the French by S. K. Mitter. In Vol. 170 of *Die Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, Berlin (1971).

[46] A. Lunardi, On the linear heat equation with fading memory. *SIAM J. Math. Anal.* **21** (1990) 1213–1224.

[47] U. Manna, M.T. Mohan and S.S. Sritharan, Stochastic non-resistive magnetohydrodynamic system with Lévy noise. *Random Oper. Stoch. Equ.* **25** (2017) 155–193.

[48] C. Marinelli and M. Röckner, On the Maximal inequalities of Burkholder, Davis and Gundy. *Expo. Math.* **34** (2016) 1–26.

[49] M. Métivier, Stochastic Partial Differential Equations in Infinite Dimensional Spaces. Quaderni, Scuola Normale Superiore, Pisa (1988).

[50] R. Mikulevicius and B.L. Rozovskii, Global L^2-solutions of stochastic Navier-Stokes equations. *Ann. Probab.* **33** (2005) 137–176.

[51] R.K. Miller, Linear Volterra integrodifferntial equations as semigroups. *Funkcial. Ekvac.* **17** (1974) 39–55.

[52] A.S. Monin and A.M. Yaglom, Statistical Fluid Mechanics: Mechanics of Turbulence, Vol. 1. The MIT Press, MA (1971).

[53] E. Motyl, Martingale Solutions to the 2D and 3D Stochastic Navier-Stokes Equations Driven by the Compensated Poisson Random Measure, Preprint 13. Department of Mathematics and Computer Sciences, Lodz University (2011).

[54] E. Motyl, Stochastic Navier–Stokes equations driven by Lévy noise in unbounded 3D domains. *Potential Anal.* **38** (2013) 863–912.

[55] E. Motyl, Stochastic hydrodynamic-type evolution equations driven by Lévy noise in 3D unbounded domains-Abstract framework and applications. *Stoch. Process. Appl.* **124** (2014) 2052–2097.

[56] N. Nagase and M. Nisio, Optimal controls for stochastic partial differential equations. *SIAM J. Control Optim.* **28** (1990) 186–213.

[57] J.G. Oldroyd, Non-Newtonian flow of liquids and solids. In Vol. I of Rheology: Theory and applications, edited by F.R. Eirich. Academic Press Inc., New York, (1956), 653–682.

[58] M. Ondreját, Uniqueness for stochastic evolution equations in Banach spaces. *Diss. Math.* **426** (2004) 1–63.

[59] E. Pardoux, Stochastic partial differential equations and filtering of diffusion processes. *Stochastics* **3** (1979) 127–167.

[60] E. Pardoux, Équations aux dérivées partielles stochastiques de type monotone. *Séminaire Jean Leray* **3** (1975) 1–10.

[61] K.R. Parthasarathy, Probability Measures on Metric Spaces. Academic Press, New York (1967).

[62] S. Peszat and J. Zabczyk, Stochastic Partial Differential Equations with Lévy Noise. Vol. 13 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (2007).

[63] J. Prüss, On linear volterra equations of parabolic type in Banach spaces. *Trans. Am. Math. Soc.* **301** (1987) 691–721.

[64] J. Prüss, Evolutionary Integral Equations and Applications. Vol. 87 of *Monographs in Mathematics*. Birkhaüser Verlag, Basel (1993).

[65] M. Renardy, W.J. Hrusa and J.A. Nohel, Mathematical Problems in Viscoelasticity. Vol. 35 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*. Longman Scientific and Technical, Harlow; John Wiley and Sons, Inc., New York (1987).

[66] R.S. Rivlin, The relation between the flow of non-Newtonian fluids and turbulent Newtonian fluids. *Quart. Appl. Math.* **15** (1957) 212–215.

[67] K. Sakthivel and S.S. Sritharan, Martingale solutions for stochastic Navier-Stokes equations driven by Lévy noise. *Evol. Equ. Control Theory* **1** (2012) 355–392.

[68] S.E. Spagnolie, Complex Fluids in Biological Systems: Experiment, Theory, and Computation. Biological and Medical Physics Biomedical Engineering. Springer, New York (2015).

[69] S.S. Sritharan, Deterministic and stochastic control of Navier-Stokes equation with linear, monotone, and hyperviscosities. *Appl Math. Optim.* **41** (2000) 255–308.

[70] R. Temam, Navier-Stokes Equations, Theory and Numerical Analysis. AMS Chelsea Publishing, North-Holland, Amsterdam (1984).

[71] B.W. Towler, C.J. Rupp, A.B. Cunningham and P. Stoodley, Viscoelastic properties of a mixed culture biofilm from rheometer creep analysis. *Biofouling* **19** (2003) 279–285.

[72] F. Treves, Topological Vector Spaces, Distributions and Kernels. Academic Press, Indiana (1967).

[73] F. Tröltzsch, Optimal Control of Partial Differential Equations: Theory, Methods and Applications. Vol. 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI (2010).
[74] C.A. Truesdell and K.R. Rajagopal, An Introduction to the Mechanics of Fluids. Birkhauser, Boston (1999).
[75] E.O. Tuck, On positivity of Fourier transforms. *Bull. Aust. Math. Soc.* **74** (2006) 133–138.
[76] M. Valadier, Young measures, in Methods of Nonconvex Analysis (Varenna, 1989). Vol. 1446 of *Lecture Notes in Mathematics*. Springer, Berlin (1990) 152–188.
[77] J. Warga, Relaxed variational problems. *J. Math. Anal. Appl.* **4** (1962) 111–128.
[78] J. Warga, Optimal Control of Differential and Functional Equations. Academic Press, New York (1972).
[79] L.C. Young, Generalized surfaces in the calculus of variations. *Ann. Math.* **43** (1942) 84–103.
[80] L.C. Young, Lectures on the Calculus of Variations and Optimal Control Theory. W. B. Saunders, Philadelphia (1969).
[81] X.Y. Zhou, On the existence of optimal relaxed controls of stochastic partial differential equations. *SIAM J. Control Optim.* **30** (1992) 247–261.