ESAIM: COCV 31 (2025) 42
https://doi.org/10.1051/cocv/2025030

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

SINGULAR LIMIT OF BSDES AND OPTIMAL CONTROL OF TWO SCALE SYSTEMS WITH JUMPS IN INFINITE DIMENSIONAL SPACES

Elena Bandini[1,*], Giuseppina Guatteri[2] and Gianmario Tessitore[3]

Abstract. The paper is devoted to a stochastic optimal control problem for a two scale, infinite dimensional, stochastic system. The state of the system consists of "slow" and "fast" component and its evolution is driven by both continuous Wiener noises and discontinuous Poisson-type noises. The presence of discontinuous noises is the main feature of the present work. We use the theory of backward stochastic differential equations (BSDEs) to prove that, as the speed of the fast component diverges, the value function of the control problem converges to the solution of a reduced forward-backward system that, in turn, is related to a reduced stochastic optimal control problem. The results of this paper generalize to the case of discontinuous noise the ones in [Guatteri and Tessitore, *Appl. Math. Optim.* **83** (2021) 1025–1051] and [Święch, *ESAIM Control Optim. Calc. Var.* **27** (2021) Paper No. 6, 34].

Mathematics Subject Classification. 60H15, 60G51, 60G57, 93C70, 93E03.

Received January 15, 2024. Accepted March 3, 2025.

1. Introduction

Stochastic dynamic models that evolve based on distinct and divergent time scales constitute a well-established area of research, encompassing both controlled and uncontrolled scenarios. This field has found extensive applications, as evidenced by classical results detailed in [1, 2], and more recent applications, such as those explored in [3, 4], particularly in the realms of neural networks and climate models.

Our specific focus lies in examining two scales controlled systems featuring both Wiener and discontinuous (jump) noise:

Keywords and phrases: Discontinuous noise, Poisson random measures, optimal control of two scales systems, backward stochastic differential equations, stochastic evolution equations in Hilbert spaces.

[1] University of Bologna, Department of Mathematics, piazza di Porta San Donato 5, 40126 Bologna, Italy.
[2] Politecnico di Milano, Department of Mathematics, *via* Bonardi 9, 20133 Milan, Italy.
[3] University of Milano-Bicocca, Department of Mathematics and its Applications, *via* Roberto Cozzi 55, 20125 Milan, Italy.

[*] Corresponding author: elena.bandini7@unibo.it

$$\begin{cases} dX_t^\varepsilon = AX_t^\varepsilon \mathrm{d}t + b(X_t^\varepsilon, Q_t^\varepsilon, \alpha_t)\mathrm{d}t + RdW_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - r(X_t^\varepsilon, Q_t^\varepsilon, \alpha_t, w)\nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0 \in H, \\ \varepsilon dQ_t^\varepsilon = [BQ_t^\varepsilon + F(X_t^\varepsilon, Q_t^\varepsilon) + G\rho(\alpha_t)]\mathrm{d}t + \sqrt{\varepsilon}GdW_t^2 \\ \qquad + \int_{K\setminus 0} w[\varepsilon N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \gamma(\alpha_t, w)\nu_2(\mathrm{d}w)\mathrm{d}t], \\ Q_0^\varepsilon = q_0 \in K. \end{cases}$$

In the above, process (α) is the control, (X^ε) is the slowly evolving state variable and (Q^ε) is the fast state variable. Notice that (α), (X^ε) and (Q^ε) all take values in infinite dimensional spaces. (W^1) and (W^2) are independent cylindrical Wiener processes. $N_1(\mathrm{d}t\,\mathrm{d}w)$ and $N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) = N_2(d(t/\varepsilon)\,\mathrm{d}w)$ are Poisson random measures, independent of each other and from (W^1) and (W^2), with compensated random measures

$$\begin{aligned} \tilde{N}_1(\mathrm{d}t\,\mathrm{d}w) &= N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t, \\ \tilde{N}_2^\varepsilon(d(t/\varepsilon)\,\mathrm{d}w) &= N_2(\mathrm{d}t\,\mathrm{d}w) - \varepsilon^{-1}\nu_2(\mathrm{d}w)\mathrm{d}t, \end{aligned}$$

while r and γ represent their controlled intensity. Finally, A and B are, possibly unbounded, linear operators, we assume that $B + F$ is dissipative and that operator R is invertible; in other words, it is crucial to suppose that, if we forget the controlled terms, the fast variable has an "ergodic" behaviour while the slow variable is perturbed by non degenerate gaussian noise.

The above state equation is coupled with a standard cost functional as:

$$J^\varepsilon(x_0, q_0, \alpha) = \mathbb{E}\left[\int_0^1 l(X_t^\varepsilon, Q_t^\varepsilon, \alpha_t)\mathrm{d}t + h(X_1^\varepsilon)\right].$$

The value function of the optimal control problem is then classically defined by:

$$V^\varepsilon(x_0, q_0) = \inf_\alpha J^\varepsilon(x_0, q_0, \alpha). \tag{1.1}$$

Our purpose is to characterize the limiting value function $\lim_{\varepsilon\to 0} V^\varepsilon(x_0, q_0)$.

As first we will need to represent the value functions of both the problems, the approximated and the limit one, using the Dynamic Programming Principle. In literature there are two approaches. One is purely analytic and consists in proving the existence and uniqueness of the viscosity solution (classical when the coefficients are regular enough) to the Hamilton-Jacobi-Bellman (HJB) PDEs associated to each problem. The alternative approach is probabilistic, as introduced in the pioneering paper by Pardoux and Peng [5]. This method replaces the direct analysis of the PDE with the study of a forward-backward stochastic differential equation (FBSDE) system, which generalizes the stochastic differential equations used in the Feynman-Kac formula. Once this system is solved, the backward component of the FBSDE solution yields the value function and, in some cases, even the optimal control. These two approaches are quite used in the finite-dimensional case; however, the theory of viscosity solutions is generally preferred for fully nonlinear problems. In the infinite dimensional case the BDSE approach becomes more competitive, see [6], Chapter 6, since it allows to treat more general noise terms, like the cylindrical Wiener process, thus including the white noise case. On the other hand, the HJB techniques provide more flexibility regarding the structure of the model (possibility of considering control-dependent noise) but show an obstruction for the inclusion of cylindrical noise in the model and seem to be limited to the treatment of regular (B-continuous) value functions (see [6], Chap. 3).

Once the representations of the value functions are at hand then we can proceed to study the convergence as the rate ε tends to 0.

Numerous papers have addressed the aforementioned convergence problem within the finite-dimensional framework, specifically considering continuous noise. If, in very specific cases, the result can be derived by direct computations on the control problem (see for instance [2, 7]), in the largest part of the literature, see [8–12], the problem is approached by examining the convergence of viscosity solutions to the HJB equation, which describes the evolution of the approximating value functions V^ε.

As far as the infinite-dimensional scenario the earliest results, to the best of our knowledge, can be traced back to [13]. In this work, the cylindrical Wiener process case is addressed, assuming non-degeneracy of the noise in the slow evolution. The critical technical aspect lies in proving that the solutions to the resulting singular sequence of forward-backward stochastic differential equations converge towards a reduced system existing in a smaller space. When the two-scale system is controlled, we cannot expect the reduced system to be the original system with the averaged coefficients, as in the free case [1], but we need to introduce the so-called *effective Hamiltonian*, see [10]. The convergence argument becomes considerably more complex, as the coefficients of the approximated problems completely differ from those of the limit (reduced) problem. In [13], we exploited the fact that the effective Hamiltonian can be expressed in terms of the solution of an "ergodic" BSDE, a class of equations discussed in works such as [14–16]. Subsequently, in [17], the author obtains similar results successfully applying viscosity solution techniques in infinite dimensional spaces with, on the one hand, the restriction of considering only a gaussian trace class noise, but, on the other hand, including the case of control dependent noise.

It is worth noticing that in both approaches the problem of rapidity of convergence or the one of convergence of optimal controls seems out of reach. The primary challenge is that, as already pointed out, the limit system here does not correspond to the averaged system, thus the techniques used in the case without control (see [18–21]) seem inapplicable.

The main novelty in the present paper lies in the inclusion of discontinuous noise, that is here present in both the slow and fast equations beside the cylindrical Wiener process.

Building on the approach introduced in [13], and being interested in the generality of the noises, in particular wanting to allow both cylindrical Wiener process and discontinuous noises, we opt for the forward-backward stochastic system representation of the value function. To be slightly more detailed about this choice: the inclusion of Poisson noises in the state equation induces the appearance of complicated, non local terms, in the HJB equation, see [22, 23]. Moreover, as we have already mentioned, existence and uniqueness of a viscosity solution to the corresponding integro-differential HJB equation has only been obtained when noise is trace class and under the extra continuity assumptions on the solutions for uniqueness (see again [22, 23] and [24]). So a proof of convergence of value functions, based on the convergence of the viscosity solutions of such non trivial equations, seems technically tricky, beside being limited to trace class diffusions. On the other side, the crucial tool for the forward-backward approach to control problems, the Girsanov transform, is still available, albeit in a non-trivial form reported in Appendix A. Therefore, an approach based on the stochastic representation of the value function seemed more feasible, although not easy, in the case of non-Gaussian noises.

In comparison with [13], we notice that integral terms of the form $\int_K \vartheta(w)N(\mathrm{d}t,\mathrm{d}w)$ appear in the BSDEs under consideration (see (4.1), (4.2), (4.6)). While the treatment of such terms is well understood for a single BSDE (see *e.g.* [25–28]) and poses minimal challenges regarding the existence and uniqueness of ergodic BSDEs (see Sect. 4.1, where we have omitted details that closely mirror those in [13]), their interaction with the key discretization argument (see the proof of Thm. 4.7) was previously unexplored. We note that identifying the appropriate framework for embedding our problem and the necessary assumptions for our argument was not initially straightforward (see Sect. 2).

Let us now describe the procedure we will follow.

We introduce the following forward-backward system to represent $V^\varepsilon(x_0,q_0)$ (see [29] for the extension of the classical results to the discontinuous case):

$$\begin{cases} dX_t = AX_t\mathrm{d}t + RdW_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0 \in H, \\ \varepsilon dQ_t^\varepsilon = [BQ_t^\varepsilon + F(X_t,Q_t^\varepsilon)]\mathrm{d}t + \sqrt{\varepsilon}GdW_t^2 + \int_{K\setminus 0} w[\varepsilon N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}t], \\ Q_0^\varepsilon = q_0 \in K, \\ -dY_t^\varepsilon = \psi\Big(X_t, Q_t^\varepsilon, Z_t^\varepsilon, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, U_t^\varepsilon(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big)\mathrm{d}t - Z_t^\varepsilon dW_t^1 - \Xi_t^\varepsilon dW_t^2 \\ - \int_{H\setminus 0} U_t^\varepsilon(w)(N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t) - \int_{K\setminus 0}\Theta_t^\varepsilon(w)\Big(N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big), \\ Y_1^\varepsilon = h(X_1). \end{cases}$$

Then the main technical result, see Theorem 4.7, consists in the proof that the solution of the forward-backward system above (depending on ε in a singular way) converges towards the solution of the following "reduced" system

$$\begin{cases} dX_t = AX_t \mathrm{d}t + R dW^1_t + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0, \\ -d\bar{Y}_t = \lambda(X_t, \bar{Z}_t, \bar{U}_t(\cdot))\mathrm{d}t - \bar{Z}_t dW^1_t - \int_{H\setminus 0} \bar{U}_t(w)[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ \bar{Y}_1 = h(X_1). \end{cases}$$

The term λ is the effective Hamiltonian, that is characterized as the value function of a suitable ergodic control problem that we find in Section 4.1 solving the associated ergodic BSDE. The proof relies on a discretization strategy for the slow variable. In particular, the result is achieved through several applications of the Girsanov transformation, which involve managing probability changes dependent on the parameter ε (see Steps 3 and 4 of the proof of Theorem 4.7), and that become delicate due to the presence of Poisson random measures N_1 and N_2. As a matter of fact, in order to deal with the new terms arising in the main freezing argument (see for instance the terms involving U in (4.29)), one needs a general version of the Girsanov theorem, holding for general random measures, as it is provided in Appendix A. To this end, a key tool is Hypothesis $(\mathbf{H}\psi)$-d) (see also Rem. 4.3), which we prove to be verified in our control problem, see Proposition 5.2. Notice that the application of the correct transformation is quite tricky because uniform estimates have to hold at each step, see (4.23), (4.34) and (4.35).

After establishing the result concerning the singular forward-backward system, we can finalize the program by expressing the limit $\lim_{\varepsilon\to 0} V^\varepsilon(x_0, q_0)$ through the solution of an appropriate reduced forward-backward system, see Theorem 5.4. Additionally, this limit can be further characterized as the value function of a suitable "reduced" control problem, see Theorem 6.2.

The paper is structured as follows. Section 2 establishes notations. In Section 3 we point out some results on the forward system. In Section 4, we present and prove the main result concerning the convergence of singular forward-backward systems of stochastic equations, also discussing the well-posedness of the ergodic equation required to characterize the effective Hamiltonian. Section 5 applies the aforementioned convergence result to the original stochastic two-scale control problem featuring jump noise. Finally, in Section 6, we provide an interpretation of the main result in the context of a "reduced" optimal control problem. We eventually provide some results on the Girsanov transformation for general semimartingales in the Appendix.

2. Notation

For any topological space Λ, $\mathcal{B}(\Lambda)$ denotes the corresponding Borel σ-algebra. Let E be a Banach space norm $|\cdot|_E$ or simply $|\cdot|$ when no confusion is possible; for any other Banach space F, we denote by $L(E, F)$ the space of bounded linear operators from E to F, endowed with the usual norm. If $F = \mathbb{R}$, the dual space $L(E, \mathbb{R})$ is denoted by E^*.

Ξ, H, K are Hilbert spaces with scalar product denoted by $\langle \cdot, \cdot \rangle$. They are assumed to be real and separable and the dual of Hilbert spaces will never be identified with the space itself. By $L_2(\Xi, H)$ (resp. $L_2(\Xi, K)$) we will denote the space of Hilbert-Schmidt operators from Ξ to H (resp. to K).

We consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and a filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. T will be a fixed horizon, and $\mathcal{P}$ will denote the σ-algebra of predictable sets on $\Omega \times [0, T]$.

We consider two cylindrical $(\mathcal{F}_t)$-Wiener processes $(W^1_t)_{t\geq 0}$ and $(W^2_t)_{t\geq 0})$ with values in Ξ. We assume $(W^1_t)_{t\geq 0}$ and $(W^2_t)_{t\geq 0}$ to be independent. $N_1(\mathrm{d}t\,\mathrm{d}w)$, $N_2(\mathrm{d}t\,\mathrm{d}w)$ will be independent $(\mathcal{F}_t)$-Poisson random measures on H and on K respectively, with compensated random measures

$$\tilde{N}_1(\mathrm{d}t\,\mathrm{d}w) = N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t,$$
$$\tilde{N}_2(\mathrm{d}t\,\mathrm{d}w) = N_2(\mathrm{d}t\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}t,$$

where $\nu_1(\mathrm{d}w)$ and $\nu_2(\mathrm{d}w)$ are σ-finite measure on H and on K respectively, such that

$$\int_{H\setminus 0} |w|^2 \nu_1(\mathrm{d}w) < \infty, \quad \int_{K\setminus 0} |w|^2 \nu_2(\mathrm{d}w) < \infty.$$

$N_1(\mathrm{d}t\,\mathrm{d}w)$ and $N_2(\mathrm{d}t\,\mathrm{d}w)$ are also independent of $(W_t^1)_{t\geq 0}$ and $(W_t^2)_{t\geq 0}$. Moreover for every $\varepsilon \in (0,1]$ we set

$$N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) = N_2(d(t/\varepsilon)\,\mathrm{d}w)$$

that is a random measure with compensator $\frac{1}{\varepsilon}\nu_2(\mathrm{d}w)\mathrm{d}t$.

We also introduce

$$L^2(\nu_1) := \{f : H \to \mathbb{R} \text{ measurable} : \int_{H\setminus 0} |f(w)|^2 \nu_1(\mathrm{d}w) < \infty\},$$
$$L^2(\nu_2) := \{f : K \to \mathbb{R} \text{ measurable} : \int_{K\setminus 0} |f(w)|^2 \nu_2(\mathrm{d}w) < \infty\},$$

and

$$\begin{aligned} \mathbb{L}^2(\tilde N_1) := \{&\mathcal{P}\otimes\mathcal{B}(H)\text{-measurable processes}\, U \text{ such that the norm} \\ &||U||_{\mathbb{L}^2(\tilde N_1)} := \Big(\mathbb{E}\Big[\int_0^T \int_{H\setminus 0} |U_s(w)|^2 \nu_1(\mathrm{d}w)\mathrm{d}s\Big]\Big)^{1/2} \text{ is finite}\}, \\ \mathbb{L}^2(\tilde N_2) := \{&\mathcal{P}\otimes\mathcal{B}(K)\text{-measurable processes}\, U \text{ such that the norm} \\ &||U||_{\mathbb{L}^2(\tilde N_2)} := \Big(\mathbb{E}\Big[\int_0^T \int_{K\setminus 0} |U_s(w)|^2 \nu_2(\mathrm{d}w)\mathrm{d}s\Big]\Big)^{1/2} \text{ is finite}\}. \end{aligned}$$

$\mathbb{L}^{2,\mathrm{loc}}(\tilde N_1)$ (resp. $\mathbb{L}^{2,\mathrm{loc}}(\tilde N_2)$) denotes the set of processes defined on $\mathbb{R}_+$ such that their restriction to $[0,\,T]$ belongs to $\mathbb{L}^{2,\mathrm{loc}}(\tilde N_1)$ (resp. $\mathbb{L}^2(\tilde N_2)$).

We also define the following spaces of equivalence classes of processes with values in a Hilbert space V. For all $\tau \in [0,T]$:

$$\begin{aligned} S^2(\mathbb{D}([\tau,\,T];V)) := \{&\mathcal{P}\text{-measurable processes}\, X \text{ with càdlàg paths in } V, \text{ such that the norm} \\ &|X|_{S^2} := \Big(\mathbb{E}\Big[\sup_{t\in[\tau,\,T]} |X_s|_V^2 \mathrm{d}s\Big]\Big)^{1/2} \text{ is finite}\}, \\ \mathbb{L}^2_V := \{&\mathcal{P}\text{-measurable processes}\, X \text{ with values in } V, \text{ such that the norm} \\ &|X|_{\mathbb{L}^2_V} := \Big(\mathbb{E}\Big[\int_0^T |X_s|_V^2 \mathrm{d}s\Big]\Big)^{1/2} \text{is finite}\}. \end{aligned}$$

$S^{2,\mathrm{loc}}(\mathbb{D}([0,\,+\infty);V))$ denotes the set of processes defined on $\mathbb{R}_+$ such that their restriction to $[0,\,T]$ belongs to $S^2(\mathbb{D}([0,\,T];V)$. $\mathbb{L}^{2,\mathrm{loc}}_V$ denotes the set of processes defined on $\mathbb{R}_+$ such that their restriction to $[0,\,T]$ belongs to $\mathbb{L}^2_V$.

3. The forward system

Given four operators $A : D(A) \subseteq H \to H$, $B : D(B) \subseteq K \to K$, $R : \Xi \to H$, $G : \Xi \to K$ and $x_0 \in H$, $q_0 \in K$, we consider the following system of SDEs:

$$\begin{cases} dX_t = AX_t \mathrm{d}t + RdW^1_t + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0, \\ \varepsilon dQ^\varepsilon_t = [BQ^\varepsilon_t + F(X_t, Q^\varepsilon_t)]\mathrm{d}t + \sqrt{\varepsilon} G dW^2_t + \int_{K\setminus 0} w[\varepsilon N^\varepsilon_2(\mathrm{d}t\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}t], \\ Q^\varepsilon_0 = q_0, \end{cases} \tag{3.1}$$

where the "slow" variable X takes values in H and the "fast" variable Q^ε takes values in K.

We will consider the following assumptions on the coefficients of the system (3.1).

(HAB) $A : D(A) \subseteq H \to H$ and $B : D(B) \subseteq K \to K$ are two linear unbounded operators generating the C_0 semigroups $\{\mathrm{e}^{tA}\}_{t\geq 0}$, $\{\mathrm{e}^{tB}\}_{t\geq 0}$ over H and K respectively, such that

$$|\mathrm{e}^{tA}|_{L(H)} \leq M_A \mathrm{e}^{w_A t}, \quad t \geq 0, M_A > 0, w_A \in \mathbb{R},$$
$$|\mathrm{e}^{tB}|_{L(K)} \leq M_B \mathrm{e}^{w_B t}, \quad t \geq 0, M_B > 0, w_B \in \mathbb{R}.$$

Moreover, for $t > 0$, e^{tA} and e^{tB} are Hilbert-Schmidt operators with

$$|\mathrm{e}^{tA}|_{L_2(H,H)} + |\mathrm{e}^{tB}|_{L_2(K,K)} \leq M t^{-\gamma t},$$

for all $t \in (0, 1)$, suitable $\gamma \in (0, 1/4)$ and $M > 0$.

(HRG) $R \in L(\Xi, H)$, $G \in L(\Xi, K)$.

(HF) $F : H \times K \to K$ is bounded, and there exists $L_F > 0$:

$$|F(x, y) - F(x', y')|_K \leq L_F(|x - x'|_H + |y - y'|_K), \quad x, x' \in H, y, y' \in K.$$

(HF+B) $B + F$ is dissipative: there exists $\mu > 0$ such that

$$\langle Bq + F(x, q) - [Bq' + F(x, q')], q - q'\rangle \leq -\mu |q - q'|^2_K, \quad x \in H, q, q' \in D(B).$$

Let $(\gamma_t)_{t\geq 0}$ be a cylindrical Wiener process with values in Ξ and $p(\omega, \mathrm{d}t\,\mathrm{d}v)$ a random measure on K with compensator $\kappa_t(\omega, v)\nu(\mathrm{d}v)\mathrm{d}t$, where ν is a σ-finite measure on K and κ is a random field on K such that

$$\int_{K\setminus 0} |v|^2 |\kappa_s(\omega, v)|^2 \nu(\mathrm{d}v) < \infty, \quad \text{for every } (\omega, s) \in \Omega \times [0, T]. \tag{3.2}$$

We denote by $(\beta^B_t)_{t\geq 0}$ the stochastic convolution

$$\beta^B_t := \int_0^t \mathrm{e}^{(t-s)B}\Big(G\, d\gamma_s + \int_{K\setminus 0} v[p(\mathrm{d}s\,\mathrm{d}v) - \kappa_s(v)\nu(\mathrm{d}v)\mathrm{d}s]\Big).$$

We will assume the following:

(Hβ^B) $\sup_{t\geq 0} \mathbb{E}[|\beta^B_t|^2_K] < \infty$.

Remark 3.1. We notice that **(Hβ^B)** is automatically satisfied when B is dissipative and consequently $w_B < 0$ in **(HAB).**

3.1. Well-posedness results

The well-posedness of the slow equation in system (3.1) comes from the following theorem, that is a direct corollary of Theorem 9.29 in [30].

Lemma 3.2. *Let assumptions* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, *and* **(Hβ^B)** *hold. For any $T < \infty$, $\tau \in [0, T]$ and any $\mathcal{F}_\tau$-measurable square integrable random variable $\tilde{X}_\tau \in H$,*

$$\begin{cases} dX_t = AX_t\mathrm{d}t + RdW_t^1 + \int_{H\setminus 0} z[N_1(\mathrm{d}t\,\mathrm{d}z) - \nu_1(\mathrm{d}z)\mathrm{d}t], \quad t \in [\tau, T], \\ X_\tau = \tilde{X}_\tau \end{cases}$$

has a unique (up to modification) mild solution $(X_t^{\tau,\tilde{X}_\tau})_{t\in[\tau,T]}$ with a càdlàg version. More precisely, there exists a unique process $(X_t)_{t\in[\tau,T]} \in S^2(\mathbb{D}([\tau, T]; H))$ such that

$$X_t = \mathrm{e}^{(t-\tau)A}\tilde{X}_\tau + \int_\tau^t \mathrm{e}^{(s-\tau)A}R\,dW_s^1 + \int_\tau^t \mathrm{e}^{(s-\tau)A}\int_{H\setminus 0} z[N_1(\mathrm{d}s\,\mathrm{d}z) - \nu_1(\mathrm{d}z)\mathrm{d}s]$$

is satisfied $\mathbb{P}$-a.s. Moreover, for every T there exists $c_T < +\infty$ such that

$$\sup_{t\in[\tau,T]} \mathbb{E}[|X_t^{\tau,x} - X_t^{\tau,y}|_H^2] \le c_T|x-y|_H^2, \quad x, y \in H,$$

and, for all $p \ge 1$, there exists a positive constant c_p depending only on p and on the quantities introduced in the hypotheses, such that

$$\mathbb{E}\Big[\sup_{t\in[\tau,T]} |X_t^{\tau,x}|^p\Big] \le c_p(1 + |x|^p), \quad x \in H. \tag{3.3}$$

Concerning the fast equation in system (3.1), we start by noticing that, if we make the change of time $s \mapsto s\varepsilon$, it reads

$$\begin{cases} \varepsilon dQ_{\varepsilon t}^\varepsilon = [BQ_{\varepsilon t}^\varepsilon + F(X_{\varepsilon t}, Q_{\varepsilon t}^\varepsilon)]\varepsilon\mathrm{d}t + \sqrt{\varepsilon}GdW_{\varepsilon t}^2 + \int_{K\setminus 0} w[\varepsilon N_2^\varepsilon(d(\varepsilon t)\,\mathrm{d}w) - \varepsilon\nu_2(\mathrm{d}w)\mathrm{d}t], \quad t \in [0, T], \\ Q_0 = q_0, \end{cases}$$

that in turn gives

$$\begin{cases} d\hat{Q}_t = [B\hat{Q}_t + F(X_{\varepsilon t}, \hat{Q}_t)]\mathrm{d}t + Gd\,\hat{W}_t^2 + \int_{K\setminus 0} w[N_2(\mathrm{d}t\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}t], \quad t \in [0, T], \\ \hat{Q}_0 = q_0, \end{cases} \tag{3.4}$$

where we have set

$$\hat{Q}_t := Q_{\varepsilon t}^\varepsilon, \qquad \hat{W}_t^2 := \frac{1}{\sqrt{\varepsilon}}dW_{\varepsilon t}^2.$$

We have the following.

Lemma 3.3. *Let $(\Gamma_s)_{s\ge 0}$ be a given, H-valued, predictable process in $\mathbb{L}_H^{2,\mathrm{loc}}$. Let $(g_s)_{s\ge 0}$ be a given, K-valued, process with $g \in \mathbb{L}_K^{2,\mathrm{loc}}$. Let $(\gamma_t)_{t\ge 0}$ be a Q-Wiener process with values in Ξ and $p(\omega, \mathrm{d}t\,\mathrm{d}v)$ a random measure on K with compensator $\kappa_t(\omega, v)\nu(\mathrm{d}v)\mathrm{d}t$, where ν is a σ-finite measure on K and κ is a random field on K such* *that* (3.2) *holds true. Let us consider the following equation:*

$$\begin{cases} dQ_t = [BQ_t + F(\Gamma_t, Q_t)]\mathrm{d}t + g_t\mathrm{d}t + Gd\gamma_t + \int_{K\setminus 0} v[p(\mathrm{d}t\,\mathrm{d}v) - \kappa_t(v)\nu(\mathrm{d}v)\mathrm{d}t], \quad t \in [0,T], \\ Q_0 = q_0. \end{cases} \tag{3.5}$$

Then equation (3.5) *admits a unique mild solution* $Q \in S^{2,\mathrm{loc}}(\mathbb{D}([0,\infty);K))$*. Moreover, for all* $T > 0$*, there exists* $k < +\infty$*, independent of* T*, such that*

$$\sup_{t\in[0,T]} \mathbb{E}[|Q_t|_K^2] \le k\Big(1 + |q_0|_K^2 + \sup_{t\in[0,T]} \mathbb{E}[|\Gamma_t|_H^2] + \sup_{t\in[0,T]} \mathbb{E}[|\beta_t^B|_K^2] + \sup_{t\in[0,T]} \mathbb{E}[|g_t|_K^2]\Big). \tag{3.6}$$

Finally, if $(\Gamma'_s)_{s\ge 0}$ *is another process in* $\mathbb{L}_H^{2,\mathrm{loc}}$*, and* Q' *is a the mild solution of equation*

$$\begin{cases} dQ'_t = [BQ'_t + F(\Gamma'_t, Q'_t)]\mathrm{d}t + g_t\mathrm{d}t + Gd\gamma_t + \int_{K\setminus 0} v[p(\mathrm{d}t\,\mathrm{d}v) - \kappa_t(v)\nu(\mathrm{d}v)\mathrm{d}t], \quad t \in [0,T], \\ Q'_0 = q_0, \end{cases}$$

then, for all $T > 0$*,* $\mathbb{P}$*-a.s.,*

$$|Q_T - Q'_T| \le \kappa \int_0^T \mathrm{e}^{-\mu(T-s)} |\Gamma_s - \Gamma'_s| \,\mathrm{d}s$$

where μ *is the dissipativity constant in* **(HF+B)***, and* κ *does not depend on* T*.*

Proof. The proof follows the same arguments of [30] as far as existence and uniqueness of the solution and estimate (3.6) is concerned, while the last estimate follows as in [31] (also see [13][Lem. 3.10]) noticing that

$$\begin{cases} d(Q_t - Q'_t) = [B(Q_t - Q'_t) + F(\Gamma'_t, Q'_t) - F(\Gamma_t, Q_t)]\mathrm{d}t, \quad t \in [0,T], \\ Q'_0 = q_0. \end{cases}$$

□

For fixed $\Gamma \equiv x \in H$, $q_0 \in K$ and $g \equiv 0$, equation (3.4) is a special case of equation (3.5) and Lemma 3.3 applies to it. We will denote by $\hat{Q}^{x,q_0}$ the mild solution to equation (3.4).

4. Limit equation and convergence of singular BSDEs

We introduce a function $h : H \to \mathbb{R}$ satisfying the following assumption.

(H*h***)** $h : H \to \mathbb{R}$ is bounded, and there exist positive constants L_h and M:

$$|h(x) - h(x')| \le L_h|x - x'|_H, \quad |h(x)| \le M, \quad x, x' \in H, q, q' \in K.$$

We consider the following forward-backward system for $t \in [0,1]$:

$$\begin{cases} dX_t = AX_t\mathrm{d}t + RdW_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0 \in H, \\ \varepsilon dQ_t^\varepsilon = [BQ_t^\varepsilon + F(X_t, Q_t^\varepsilon)]\mathrm{d}t + \sqrt{\varepsilon}GdW_t^2 + \int_{K\setminus 0} w[\varepsilon N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}t], \\ Q_0^\varepsilon = q_0 \in K, \\ -dY_t^\varepsilon = \psi\Big(X_t, Q_t^\varepsilon, Z_t^\varepsilon, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, U_t^\varepsilon(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big)\mathrm{d}t - Z_t^\varepsilon dW_t^1 - \Xi_t^\varepsilon dW_t^2 \\ - \int_{H\setminus 0} U_t^\varepsilon(w)(N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t) - \int_{K\setminus 0} \Theta_t^\varepsilon(w)\Big(N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big), \\ Y_1^\varepsilon = h(X_1). \end{cases} \tag{4.1}$$

Theorem 4.1. *Assume* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, **(Hβ^B)**, **(Hψ)** *and* **(Hh)**. *Then, for every* $\varepsilon > 0$, *there exists a unique 7-uple of processes* $(X, Q^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Xi^\varepsilon, U^\varepsilon(\cdot), \Theta^\varepsilon(\cdot))$ *in* $S^2(\mathbb{D}([0,\, T]; H)) \times S^2(\mathbb{D}([0,\, T]; K)) \times S^2(\mathbb{D}([0,\, +\infty); \mathbb{R})) \times \mathbb{L}^2_{\Xi^*} \times \mathbb{L}^2_{\Xi^*} \times \mathbb{L}^2(\tilde{N}_1) \times \mathbb{L}^2(\tilde{N}_2)$ *such that* $\mathbb{P}$*-a.s. the system* (4.1) *is verified for all* $t \in [0, 1]$.

Proof. System (4.1) is a decoupled forward-backward (that is, the forward equation can be solved independently, see for instance [32]). The proof of existence and uniqueness for the finite dimensional case can be found in [33] and can be extended easily to the infinite dimensional setting following [29]. □

We aim at studying the limit behaviour of Y^ε as ε goes to 0.

4.1. The ergodic parametrized BSDE

To formulate an equation for the limiting system we first have to introduce ergodic BSDEs in the context of discontinuous noise extending the results of [13], Section 4 and [14], Section 4. We remark that ergodic BSDEs with jumps in infinite dimension have also been studied in a different contest in [34].

We assume that we are given a function $\psi : H \times K \times \Xi^* \times \Xi^* \times L^2(\nu_1) \times L^2(\nu_2) \to \mathbb{R}$, on which we ask the following.

(Hψ)

a) ψ measurable;
b) $\sup_{x \in H, q \in K} \psi(x, q, 0, 0, 0, 0) =: M_\psi < \infty$;
c) there exist positive constants $L_x, L_q, L_z, L_\zeta, L_u, L_\vartheta$ such that

$$
\begin{aligned}
&|\psi(x, q, z, \zeta, u, \vartheta) - \psi(x', q', z', \zeta', u', \vartheta')| \\
&\leq L_z |z - z'|_{\Xi^*} + L_\zeta |\zeta - \zeta'|_{\Xi^*} \\
&+ L_u \Big(\int_{H \setminus 0} |u_s(w) - u'_s(w)|^2 \nu_1(\mathrm{d}w) \Big)^{\frac{1}{2}} + L_\vartheta \Big(\int_{K \setminus 0} |\vartheta_s(w) - \vartheta'_s(w)|^2 \nu_2(\mathrm{d}w) \Big)^{\frac{1}{2}} \\
&+ L_x \Big(1 + |z|_{\Xi^*} + \Big(\int_{H \setminus 0} |u_s(w)|^2 \nu_1(\mathrm{d}w) \Big)^{\frac{1}{2}} \Big) |x - x'|_H \\
&+ L_q \Big(1 + |z|_{\Xi^*} + \Big(\int_{H \setminus 0} |u_s(w)|^2 \nu_1(\mathrm{d}w) \Big)^{\frac{1}{2}} \Big) |q - q'|_K;
\end{aligned}
$$

d) for every $x \in H,\ q \in K,\ z, \zeta \in \Xi^*,\ u, u' \in L^2(\nu_1),\ \vartheta, \vartheta' \in L^2(\nu_2)$, there exist measurable functions $\gamma_1 : H \to \mathbb{R}$, $\gamma_2 : K \to \mathbb{R}$, (that may depend on $x,\ q,\ z, \zeta,\ u,\ u',\ \vartheta, \vartheta'$) such that

$$
\begin{aligned}
\psi(x, q, z, \zeta, u, \vartheta) - \psi(x, q, z, \zeta, u', \vartheta) &\leq \int_{H \setminus 0} (u(w) - u'(w)) \gamma_1(w) \nu_1(\mathrm{d}w), \\
\psi(x, q, z, \zeta, u, \vartheta) - \psi(x, q, z, \zeta, u, \vartheta') &\leq \int_{K \setminus 0} (\vartheta(w) - \vartheta'(w)) \gamma_2(w) \nu_2(\mathrm{d}w),
\end{aligned}
$$

and satisfying

$$
\begin{aligned}
&C_1(1 \wedge |w|_H) \leq \gamma_1(w) \leq C_2(1 \wedge |w|_H), \quad C_1 \in (-1, 0], C_2 \geq 0, \\
&\bar{C}_1(1 \wedge |w|_K) \leq \gamma_2(w) \leq \bar{C}_2(1 \wedge |w|_K), \quad \bar{C}_1 \in (-1, 0], \bar{C}_2 \geq 0.
\end{aligned}
$$

Remark 4.2. Notice that the lipschitzianity of f in the two latter components stated in **(Hf)**-c) is a consequence of **(Hf)**-d).

Remark 4.3. One can easily deduce (see *e.g.* the proof of Lemma 19.3.8 in [35]) from Hypothesis **(H**ψ**)**-d) the existence of two functions γ_s^1 and γ_s^2 (that may depend on x, q, z, ζ, u, u' ϑ, ϑ') satisfying

$$C_1(1 \wedge |w|_H) \leq \gamma^1(w) \leq C_2(1 \wedge |w|_H), \quad C_1 \in (-1, 0], C_2 \geq 0,$$
$$\bar{C}_1(1 \wedge |w|_K) \leq \gamma^2(w) \leq \bar{C}_2(1 \wedge |w|_K), \quad \bar{C}_1 \in (-1, 0], \bar{C}_2 \geq 0,$$

and such that

$$\psi(x, q, z, \zeta, u, \vartheta) - \psi(x, q, z, \zeta, u', \vartheta) = \int_{H \setminus 0} \gamma^1(w)(u(w) - u'(w))\nu_1(\mathrm{d}w)$$

and

$$\psi(x, q, z, \zeta, u, \vartheta) - \psi(x, q, z, \zeta, u(\cdot), \vartheta'(\cdot)) = \int_{K \setminus 0} \gamma^2(w)(\vartheta(w) - \vartheta'(w))\nu_2(\mathrm{d}w).$$

We are interested in ergodic BSDEs of the form: $\mathbb{P}$-a.s., for all $x \in H$, $q \in K$, $z \in \Xi^*$, $u \in L^2(\nu_1)$,

$$\begin{aligned} \hat{Y}_t = \hat{Y}_T &+ \int_t^T [\psi(x, \hat{Q}_s^{x,q}, z, \hat{\Xi}_s, u(\cdot), \hat{\Theta}_s(\cdot)) - \lambda(x, z, u(\cdot))]\mathrm{d}s \\ &- \int_t^T \hat{\Xi}_s d\hat{W}_s^2 - \int_t^T \int_{\Xi \setminus 0} \hat{\Theta}_s(w)[N_2(\mathrm{d}s\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}s], \quad 0 \leq t \leq T, \end{aligned} \tag{4.2}$$

where $\hat{Q}_s^{x,q}$ is the unique solution to

$$\begin{cases} d\hat{Q}_t = [B\hat{Q}_t + F(x, \hat{Q}_t)]\mathrm{d}t + Gd\hat{W}_t^2 + \int_{K \setminus 0} w[N_2(\mathrm{d}t\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}t], \ t \in [0, T], \\ \hat{Q}_0 = q. \end{cases} \tag{4.3}$$

Remark 4.4. Under assumptions **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, and **(H**β^B**)**, the solution $\hat{Q}$ to (4.3) satisfies

$$|\hat{Q}_t^{x,q} - \hat{Q}_t^{x,q'}|_K \leq \mathrm{e}^{-\mu t}|q - q'|_K, \quad t \in [0, T], \ q, q' \in K, x \in H, \tag{4.4}$$

$$\mathbb{E}\Big[\sup_{t \in [0, T]} |\hat{Q}_t^{x,q}|_K^2\Big] \leq C_T(1 + |q|_K^2), \quad q \in K, x \in H, \tag{4.5}$$

for some constant C_T only depending on T (estimate (4.4) is a direct consequence of assumption **(HF+B)**, while the proof of (4.5) follows the same lines of the one of (3.6)).

The theorem below is the analogue of the result in [13], Section 4. We omit the proof since, on the one side it is similar to the one in [13], on the other side the technical issues related to the presence of discontinuous noise will be treated, in all details, in the much more complicated proof of Theorem 4.7.

Theorem 4.5. *Under assumptions* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, **(H**β^B**)**, *and* **(H**ψ**)**, *there exist measurable functions*

$$\bar{v} : H \times K \times \Xi^* \times L^2(\nu_1) \to \mathbb{R},$$
$$\bar{\vartheta} : H \times K \times \Xi^* \times L^2(\nu_1) \to L^2(\nu_2),$$
$$\bar{\zeta} : H \times K \times \Xi^* \times L^2(\nu_1) \to \Xi^*,$$
$$\lambda : H \times \Xi^* \times L^2(\nu_1) \to \mathbb{R},$$

such that

1. *for all fixed $x \in H$, $z \in \Xi^*$, $u \in L^2(\nu_1)$, $\bar{v}$ is Lipschitz with respect to q and verifies, for some $c > 0$,*

$$|\bar{v}(x,q,z,u)| \leq c\Big(1 + |z| + \Big(\int_{H\setminus 0} |u(\ell)|^2 \nu_1(d\ell)\Big)^{\frac{1}{2}}\Big)|q|_K;$$

2. *if we set, for fixed $x \in H$, $q \in K$, $z \in \Xi^*$, $u \in L^2(\nu_1)$,*

$$\begin{cases} \bar{Y}_t^{x,q,z,u} := \bar{v}(x, \hat{Q}_t^{x,q}, z, u), \\ \bar{\Xi}_t^{x,q,z,u} := \bar{\zeta}(x, \hat{Q}_t^{x,q}, z, u), \\ \bar{\Theta}_t^{x,q,z,u}(\cdot) := \bar{\vartheta}(x, \hat{Q}_t^{x,q}, z, u)(\cdot), \end{cases}$$

then $\bar{\Xi}^{x,q,z,u} \in \mathbb{L}^{2,\mathrm{loc}}_{\Xi^*}$, $\bar{\Theta}^{x,q,z,u}(\cdot) \in \mathbb{L}^{2,\mathrm{loc}}(\tilde{N}_2)$, and the ergodic BSDE (4.2) is satisfied by $(\bar{Y}^{x,q,z,u}, \bar{\Xi}^{x,q,z,u}, \bar{\Theta}^{x,q,z,u}(\cdot), \lambda(x,z,u))$.
Moreover, if in addition, we impose that there exists a constant $c > 0$ such that $|\bar{Y}_t^{x,q,z,u}| \leq c(1 + |Q_t^{q_0}|)$, then $\lambda(x,z,u)$ is uniquely determined.

3. *for fixed $x, x' \in H$, $z, z' \in \Xi^*$, $u, u' \in L^2(\nu_1)$,*

$$\begin{aligned} |\lambda(x,z,u) - \lambda(x',z',u')| \leq\ & L'_x\Big(1 + |z|_{\Xi^*} + \Big(\int_{H\setminus 0} |u(\ell)|^2 \nu_1(d\ell)\Big)^{\frac{1}{2}}\Big)|x - x'|_H \\ & + L'_z|z - z'|_{\Xi^*} + L'_u\Big(\int_{H\setminus 0} |u(\ell) - u'(\ell)|^2 \nu_1(d\ell)\Big)^{\frac{1}{2}} \end{aligned}$$

for some positive constants L'_x, L'_z, L'_u.

4.2. Main convergence result

We can now come back to our main object of investigation, namely the characterization of the limit, as ε goes to 0, of the sequence Y_0^ε (where Y^ε is the third component of the solution to equation (4.1).)

Our candidate limit is the component $\bar{Y}$ of the solution of the following forward-backward system on the finite horizon $[0,1]$ and on the reduced space H:

$$\begin{cases} dX_t = AX_t \mathrm{d}t + R dW_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0, \\ -d\bar{Y}_t = \lambda(X_t, \bar{Z}_t, \bar{U}_t(\cdot))\mathrm{d}t - \bar{Z}_t dW_t^1 - \int_{H\setminus 0} \bar{U}_t(w)[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ \bar{Y}_1 = h(X_1), \end{cases} \tag{4.6}$$

where λ is the function in Theorem 4.5 and h is the function in $(\mathbf{H}h)$.

Theorem 4.6. *Let assumptions* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, $(\mathbf{H}\beta^B)$, $(\mathbf{H}\psi)$ *and* $(\mathbf{H}h)$ *hold true. Then, for every $\varepsilon > 0$, there exists a unique 4-tuple of processes $(X, \bar{Y}, \bar{Z}, \bar{U}(\cdot))$ in $S^2(\mathbb{D}([0,T];H)) \times S^2(\mathbb{D}([0,+\infty);\mathbb{R})) \times \mathbb{L}^2_{\Xi^*} \times \mathbb{L}^2(\tilde{N}_1)$ such that $\mathbb{P}$-a.s. the system (4.6) is verified for all $t \in [0,1]$. Moreover, there exist measurable functions*

$$\bar{v} : H \to \mathbb{R}, \quad \bar{\zeta} : H \to \Xi^*, \quad \bar{\vartheta} : H \to L^2(\nu_1),$$

such that

$$\bar{Y}_t = \bar{v}(X_t), \quad \bar{Z}_t = \bar{\zeta}(X_t), \quad \bar{U}_t(\cdot) = \bar{\vartheta}(X_t)(\cdot).$$

Proof. See the proof of Theorem 4.1. □

The following is the main technical result of the paper. The proof has the same structure of the one of [13], Theorem 5.4. The involvement of discontinuous noises induces nontrivial technical difficulties. The main differences lie in the application of the Girsanov transform (see Appendix A) and in the estimates of the residual process $R^{\varepsilon,N}$, see (4.21).

Theorem 4.7. *Let assumptions* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, **(H**β^B**)**, **(H**ψ**)** *and* **(H**h**)** *hold true. For every $\varepsilon > 0$, let $(X, Q^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Xi^\varepsilon, U^\varepsilon(\cdot))$ and $(X, \bar{Y}, \bar{Z}, \bar{U}(\cdot))$ be respectively the unique solution to the forward-backward systems* (4.1) *and* (4.6) *obtained in Theorems 4.1 and 4.6. Then*

$$\lim_{\varepsilon\to 0} Y_0^\varepsilon = \bar{Y}_0.$$

Proof. Adding and subtracting the term $\int_0^1 \psi\Big(X_t, Q_t^\varepsilon, \bar{Z}_t, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, \bar{U}_t(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big)\mathrm{d}t$ in the equation for Y^ε, we have

$$\begin{aligned}
Y_0^\varepsilon - \bar{Y}_0 = &\int_0^1 \Big(\psi\Big(X_t, Q_t^\varepsilon, Z_t^\varepsilon, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, U_t^\varepsilon(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big) - \psi\Big(X_t, Q_t^\varepsilon, \bar{Z}_t, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, \bar{U}_t(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big)\Big)\mathrm{d}t \\
&+ \int_0^1 \Big(\psi\Big(X_t, Q_t^\varepsilon, \bar{Z}_t, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, \bar{U}_t(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big) - \lambda(X_t, \bar{Z}_t, \bar{U}_t(\cdot))\Big)\mathrm{d}t \\
&- \int_0^1 (Z_t^\varepsilon - \bar{Z}_t) dW_t^1 - \int_0^1 \Xi_t^\varepsilon dW_t^2 \\
&- \int_0^1 \int_{H\backslash 0} (U_t^\varepsilon(w) - \bar{U}_t(w))(N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t) \\
&- \int_0^1 \int_{K\backslash 0} \Theta_t^\varepsilon(w)\Big(N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big).
\end{aligned} \tag{4.7}$$

The rest of the, rather lengthy, proof consists in estimating the right hand side in the above formula. For the reader's convenience we try to divide it in several steps.

Step 1. The proof is based on a discretization argument that allows to consider the "slow" variables as "frozen" constants. We start by introducing the discretized processes.

Let $N \in \mathbb{N}$. We introduce a partition of the interval $[0,1]$ of the form $t_k = k2^{-N}$, $k = 0, 1, ..., 2^N$, and we define a triple of step processes $(X^N, \tilde{Z}^N, \tilde{U}^N(\cdot))$ as follows:

$$X_t^N = X_{t_k}, \quad t \in [t_k, t_{k+1}),\ k = 0, ..., 2^N - 1, \tag{4.8}$$

$$\tilde{Z}_t^N = 0, \quad t \in [0, t_1),\ \text{and } \tilde{Z}_t^N = 2^N \int_{t_{k-1}}^{t_k} \bar{Z}_\ell d\ell, \quad t \in [t_k, t_{k+1}),\ k = 1, ..., 2^N - 1, \tag{4.9}$$

$$\tilde{U}_t^N(\cdot) = 0, \quad t \in [0, t_1),\ \text{and } \tilde{U}_t^N(\cdot) = 2^N \int_{t_{k-1}}^{t_k} \bar{U}_\ell^N(\cdot) d\ell, \quad t \in [t_k, t_{k+1}),\ k = 1, ..., 2^N - 1. \tag{4.10}$$

By construction one has

$$\lim_{N\to\infty} \mathbb{E}\Big[\int_0^1 |\tilde{Z}_t^N - \bar{Z}_t|^2 \mathrm{d}t\Big] = 0, \tag{4.11}$$

$$\lim_{N\to\infty} \mathbb{E}\Big[\int_0^1 \int_{H\backslash 0} |\tilde{U}_t^N(w) - \bar{U}_t(w)|^2 \nu_1(\mathrm{d}w)\mathrm{d}t\Big] = 0. \tag{4.12}$$

For $k = 0, ..., 2^N - 1$ we consider the following class of forward SDEs of the type in (3.4):

$$\begin{cases} d\hat{\mathcal{Q}}_t^{N,k} = [B\hat{\mathcal{Q}}_t^{N,k} + F(X_{t_k}, \hat{\mathcal{Q}}_t^{N,k})]\mathrm{d}t + Gd\hat{W}_t^2 + \int_{K\setminus 0} w[N_2(\mathrm{d}t\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}t], \quad t \geq \frac{t_k}{\varepsilon}, \\ \hat{\mathcal{Q}}_{\frac{t_k}{\varepsilon}}^{N,k} = \hat{\mathcal{Q}}_{\frac{t_k}{\varepsilon}}^{N,k-1}, \end{cases} \tag{4.13}$$

with $\hat{W}_t^2 = \frac{1}{\sqrt{\varepsilon}} dW_{\varepsilon t}^2$. At this point, let $\bar{v} : H \times K \times \Xi^* \times L^2(\nu_1) \to \mathbb{R}$, $\bar{\zeta} : H \times K \times \Xi^* \times L^2(\nu_1) \to \Xi^*$, and $\bar{\vartheta} : H \times K \times \Xi^* \times L^2(\nu_1) \to L^2(\nu_2)$ be the functions introduced in Theorem 4.5, and set, for $t \geq \frac{t_k}{\varepsilon}$,

$$\begin{cases} \check{Y}_t^{N,k} := \bar{v}(X_{t_k}, \hat{\mathcal{Q}}_t^{N,k}, \tilde{Z}_{t_k}^N, \tilde{U}_{t_k}^N(\cdot)), \\ \check{\Xi}_t^{N,k} := \bar{\zeta}(X_{t_k}, \hat{\mathcal{Q}}_t^{N,k}, \tilde{Z}_{t_k}^N, \tilde{U}_{t_k}^N(\cdot)), \\ \check{\Theta}_t^{N,k}(\cdot) := \bar{\vartheta}(X_{t_k}, \hat{\mathcal{Q}}_t^{N,k}, \tilde{Z}_{t_k}^N, \tilde{U}_{t_k}^N(\cdot))(\cdot). \end{cases}$$

By Theorem 4.5, $(\check{Y}^{N,K}, \check{\Xi}^{N,k}, \check{\Theta}^{N,k}(\cdot), \lambda(X_{t_k}, \tilde{Z}_{t_k}^N, \tilde{U}_{t_k}^N(\cdot)))$ verifies, for all $k = 0, ..., 2^N - 1$,

$$\begin{aligned} \check{Y}_{\frac{t_k}{\varepsilon}}^{N,k} - \check{Y}_s^{N,k} - \int_{\frac{t_k}{\varepsilon}}^s [\psi(X_{t_k}, \hat{\mathcal{Q}}_s^{N,k}, \tilde{Z}_{t_k}^N, \check{\Xi}_s^{N,K}, \tilde{U}_{t_k}^N(\cdot), \check{\Theta}_s^{N,k}(\cdot)) - \lambda(X_{t_k}, \tilde{Z}_{t_k}^N, \tilde{U}_{t_k}^N(\cdot))]\mathrm{d}s \\ + \int_{\frac{t_k}{\varepsilon}}^s \check{\Xi}_s^{N,k} d\hat{W}_s^2 + \int_{\frac{t_k}{\varepsilon}}^s \int_{K\setminus 0} \check{\Theta}_s^{N,k}(w)[N_2(\mathrm{d}s\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}s] = 0, \quad s \geq \frac{t_k}{\varepsilon}. \end{aligned} \tag{4.14}$$

Moreover, there exists $c > 0$ independent of k and N, such that

$$|\check{Y}_s^{N,k}| \leq c\Big(1 + |\tilde{Z}_{t_k}^N| + \Big(\int_{H\setminus 0} |\tilde{U}_{t_k}^N(\cdot)|^2 \nu_1(d\ell)\Big)^{\frac{1}{2}}\Big)|\hat{\mathcal{Q}}_s^{N,k}|_K, \quad s \geq \frac{t_k}{\varepsilon}. \tag{4.15}$$

Let us set, for $s \in [0, \frac{1}{\varepsilon}]$,

$$\hat{\mathcal{Q}}_s^N = \sum_{k=0}^{2^N-1} \hat{\mathcal{Q}}_s^{N,k} \, 1_{\left[\frac{t_k}{\varepsilon}, \frac{t_{k+1}}{\varepsilon}\right)}(s), \tag{4.16}$$

and

$$\check{\Xi}_s^N = \sum_{k=0}^{2^N-1} \check{\Xi}_s^{N,k} \, 1_{\left[\frac{t_k}{\varepsilon}, \frac{t_{k+1}}{\varepsilon}\right)}(s), \quad \check{\Theta}_s^N(\cdot) = \sum_{k=0}^{2^N-1} \check{\Theta}_s^{N,k}(\cdot) \, 1_{\left[\frac{t_k}{\varepsilon}, \frac{t_{k+1}}{\varepsilon}\right)}(s).$$

Taking $s = \frac{t_{k+1}}{\varepsilon}$ in (4.14), we get that, for all $k = 0, ..., 2^N - 1$,

$$\begin{aligned} \check{Y}_{\frac{t_k}{\varepsilon}}^{N,k} - \check{Y}_{\frac{t_{k+1}}{\varepsilon}}^{N,k} - \int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}} \Big(\psi(X_{\varepsilon s}^N, \hat{\mathcal{Q}}_s^N, \tilde{Z}_{\varepsilon s}^N, \check{\Xi}_s^N, \tilde{U}_{\varepsilon s}^N(\cdot), \check{\Theta}_s^N(\cdot)) - \lambda(X_{\varepsilon s}^N, \tilde{Z}_{\varepsilon s}^N, \tilde{U}_{\varepsilon s}^N(\cdot))\Big)\mathrm{d}s \\ + \int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}} \check{\Xi}_s^N d\hat{W}_s^2 + \int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}} \int_{K\setminus 0} \check{\Theta}_s^N(w)[N_2(\mathrm{d}s\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}s] = 0, \end{aligned} \tag{4.17}$$

where we have used that, for $s \in [\frac{t_k}{\varepsilon}, \frac{t_{k+1}}{\varepsilon}]$, one has $X_{t_k} = X_{\varepsilon s}^N$, $\tilde{Z}_{t_k}^N = \tilde{Z}_{\varepsilon s}^N$ and $\tilde{U}_{t_k}^N(\cdot) = \tilde{U}_{\varepsilon s}^N(\cdot)$ (recall (4.8)–(4.9)–(4.10)).

Step 2. Here we take advantage of relations (4.17) to rearrange equation (4.7) and prepare it to be estimated by a suitable change of probability (see Step 3 below).

We go back to (4.7). The second term of it's right hand side, namely:

$$\int_0^1 \Big(\psi\Big(X_t, Q_t^\varepsilon, \bar{Z}_t, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, \bar{U}_t(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big) - \lambda(X_t, \bar{Z}_t, \bar{U}_t(\cdot))\Big)\mathrm{d}t$$

can be rewritten as

$$\varepsilon\sum_{k=0}^{2^N-1}\int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}}\Big(\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}_s^\varepsilon(\cdot)) - \lambda(X_{\varepsilon s}, \bar{Z}_{\varepsilon s}, \bar{U}_{\varepsilon s}(\cdot))\Big)\mathrm{d}s,$$

where we have denoted $\hat{Q}_s^\varepsilon = Q_{\varepsilon t}^\varepsilon$, $\hat{\Xi}_s^\varepsilon = \frac{1}{\sqrt{\varepsilon}}\Xi_{\varepsilon s}^\varepsilon$ and $\hat{\Theta}_s^\varepsilon(\cdot) = \frac{1}{\varepsilon}\Theta_{\varepsilon s}^\varepsilon(\cdot)$. Adding in the previous expression the null terms in (4.17) for $k = 0, ..., 2^N - 1$, we get

$$\begin{aligned}
&\int_0^1 \Big(\psi\Big(X_t, Q_t^\varepsilon, \bar{Z}_t, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, \bar{U}_t(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big) - \lambda(X_t, \bar{Z}_t, \bar{U}_t(\cdot))\Big)\Big)\mathrm{d}t \\
&= \varepsilon\sum_{k=0}^{2^N-1}\Big(\check{Y}^{N,k}_{\frac{t_k}{\varepsilon}} - \check{Y}^{N,k}_{\frac{t_{k+1}}{\varepsilon}}\Big) \\
&+ \varepsilon\sum_{k=0}^{2^N-1}\int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}}\Big(\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}_s^\varepsilon(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}_s^N, \tilde{Z}^N_{\varepsilon s}, \check{\Xi}_s^N, \tilde{U}^N_{\varepsilon s}(\cdot), \check{\Theta}_s^N(\cdot))\Big)\mathrm{d}s \\
&- \varepsilon\sum_{k=0}^{2^N-1}\int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}}\Big(\lambda(X_{\varepsilon s}, \bar{Z}_{\varepsilon s}, \bar{U}_{\varepsilon s}(\cdot)) - \lambda(X^N_{\varepsilon s}, \tilde{Z}^N_{\varepsilon s}, \tilde{U}^N_{\varepsilon s}(\cdot))\Big)\mathrm{d}s \\
&+ \varepsilon\sum_{k=0}^{2^N-1}\int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}}\check{\Xi}_s^N d\hat{W}_s^2 + \varepsilon\sum_{k=0}^{2^N-1}\int_{\frac{t_k}{\varepsilon}}^{\frac{t_{k+1}}{\varepsilon}}\int_{K\setminus 0}\check{\Theta}_s^N(w)(N_2(\mathrm{d}s\,\mathrm{d}w) - \nu_2(\mathrm{d}w)\mathrm{d}s).
\end{aligned} \tag{4.18}$$

On the other hand, the first term in the right-hand side of (4.7) reads

$$\begin{aligned}
&\int_0^1 \Big(\psi\Big(X_t, Q_t^\varepsilon, Z_t^\varepsilon, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, U_t^\varepsilon(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big) - \psi\Big(X_t, Q_t^\varepsilon, \bar{Z}_t, \frac{\Xi_t^\varepsilon}{\sqrt{\varepsilon}}, \bar{U}_t(\cdot), \frac{\Theta_t^\varepsilon(\cdot)}{\varepsilon}\Big)\Big)\mathrm{d}t \\
&= \varepsilon\int_0^{\frac{1}{\varepsilon}}\Big(\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, Z_{\varepsilon s}^\varepsilon, \hat{\Xi}_s^\varepsilon, U_{\varepsilon s}^\varepsilon(\cdot), \hat{\Theta}_s^\varepsilon(\cdot)) - \psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}_s^\varepsilon(\cdot))\Big)\mathrm{d}s.
\end{aligned} \tag{4.19}$$

Plugging (4.19) and (4.18) in (4.7) we obtain

$$\begin{aligned}
Y_0^\varepsilon - \bar{Y}_0 = {}& \varepsilon\sum_{k=0}^{2^N-1}\Big(\check{Y}^{N,k}_{\frac{t_k}{\varepsilon}} - \check{Y}^{N,k}_{\frac{t_{k+1}}{\varepsilon}}\Big) \\
&+ \varepsilon\int_0^{\frac{1}{\varepsilon}}\Big(\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, Z_{\varepsilon s}^\varepsilon, \hat{\Xi}_s^\varepsilon, U_{\varepsilon s}^\varepsilon(\cdot), \hat{\Theta}_s^\varepsilon(\cdot)) - \psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}_s^\varepsilon(\cdot))\Big)\mathrm{d}s \\
&+ \varepsilon\int_0^{\frac{1}{\varepsilon}}\Big(\psi(X_{\varepsilon s}, \hat{Q}_s^\varepsilon, \bar{Z}_{\varepsilon s}, \hat{\Xi}_s^\varepsilon, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}_s^\varepsilon(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}_s^N, \tilde{Z}^N_{\varepsilon s}, \check{\Xi}_s^N, \tilde{U}^N_{\varepsilon s}(\cdot), \check{\Theta}_s^N(\cdot))\Big)\mathrm{d}s
\end{aligned}$$

$$
\begin{aligned}
&- \varepsilon \int_0^{\frac{1}{\varepsilon}} \Big(\lambda(X_{\varepsilon s}, \bar{Z}_{\varepsilon s}, \bar{U}_{\varepsilon s}(\cdot)) - \lambda(X^N_{\varepsilon s}, \tilde{Z}^N_{\varepsilon s}, \tilde{U}^N_{\varepsilon s}(\cdot))\Big) \mathrm{d}s \\
&- \int_0^1 (Z^\varepsilon_t - \bar{Z}_t) dW^1_t - \sqrt{\varepsilon} \int_0^1 \Big(\check{\Xi}^N_{\frac{t}{\varepsilon}} - \hat{\Xi}^\varepsilon_{\frac{t}{\varepsilon}}\Big) dW^2_t \\
&- \int_0^1 \int_{H\setminus 0} (U^\varepsilon_t(w) - \bar{U}_t(w))(N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t) \\
&- \varepsilon \int_0^1 \int_{K\setminus 0} (\check{\Theta}^N_{\frac{t}{\varepsilon}}(w) - \hat{\Theta}^\varepsilon_{\frac{t}{\varepsilon}}(w)) \Big(N^\varepsilon_2(\mathrm{d}t\,\mathrm{d}w) - \frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big).
\end{aligned} \tag{4.20}
$$

Now we notice that

$$
\begin{aligned}
&\psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, Z^\varepsilon_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, U^\varepsilon_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, \bar{Z}_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) \\
&+ \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, \bar{Z}_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \check{\Xi}^N_s, \tilde{U}^N_{\varepsilon s}(\cdot), \check{\Theta}^N_s(\cdot)) \\
&= \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, Z^\varepsilon_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, U^\varepsilon_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, \bar{Z}_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, U^\varepsilon_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) \\
&+ \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, \bar{Z}_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, U^\varepsilon_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, \bar{Z}_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) \\
&+ R^{\varepsilon,N}_s \\
&+ \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, \tilde{U}^N_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \check{\Xi}^N_s, \tilde{U}^N_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) \\
&+ \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \check{\Xi}^N_s, \tilde{U}^N_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \check{\Xi}^N_s, \tilde{U}^N_{\varepsilon s}(\cdot), \check{\Theta}^N_s(\cdot)),
\end{aligned}
$$

where we have set

$$
R^{\varepsilon,N}_s := \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, \bar{Z}_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, \bar{U}_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, \tilde{U}^N_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)). \tag{4.21}
$$

Recalling Hypothesis **(H**ψ**)**-c), there exist $L_x, L_q, L_z, L_u > 0$ such that

$$
\begin{aligned}
|R^{\varepsilon,N}_s| \le\ & L_z |\bar{Z}_{\varepsilon s} - \tilde{Z}^N_{\varepsilon s}|_{\Xi^*} + L_u \Big(\int_{H\setminus 0} |\bar{U}_{\varepsilon s}(w) - \tilde{U}^N_{\varepsilon s}(w)|^2 \nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}} \\
&+ L_x \Big(1 + |\bar{Z}_{\varepsilon s}|_{\Xi^*} + \Big(\int_{H\setminus 0} |\bar{U}_{\varepsilon s}(w)|^2 \nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big) |X_{\varepsilon s} - X^N_{\varepsilon s}|_H \\
&+ L_q \Big(1 + |\bar{Z}_{\varepsilon s}|_{\Xi^*} + \Big(\int_{H\setminus 0} |\bar{U}_{\varepsilon s}(w)|^2 \nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big) |\hat{Q}^\varepsilon_s - \hat{\mathcal{Q}}^N_s|_K.
\end{aligned} \tag{4.22}
$$

At this point we introduce

$$
\delta^{1,\varepsilon}(s) = \begin{cases} \frac{|\psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, Z^\varepsilon_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, U^\varepsilon_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X_{\varepsilon s}, \hat{Q}^\varepsilon_s, \bar{Z}_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, U^\varepsilon_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot))|}{|Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s}|^2} (Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s})^*, & \text{if } |Z^\varepsilon_{\varepsilon s} - \bar{Z}_{\varepsilon s}| \neq 0, \\ 0, & \text{otherwise,} \end{cases}
$$

$$
\delta^{2,\varepsilon,N}(s) = \begin{cases} \frac{|\psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \hat{\Xi}^\varepsilon_s, \tilde{U}^N_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde{Z}^N_{\varepsilon s}, \check{\Xi}^N_s, \tilde{U}^N_{\varepsilon s}(\cdot), \hat{\Theta}^\varepsilon_s(\cdot))|}{|\hat{\Xi}^\varepsilon_s - \check{\Xi}^N_s|^2} (\hat{\Xi}^\varepsilon_s - \check{\Xi}^N_s)^*, & \text{if } |\hat{\Xi}^\varepsilon_s - \check{\Xi}^N_s| \neq 0, \\ 0, & \text{otherwise.} \end{cases}
$$

We notice that, by Hypothesis **(H**ψ**)-c)**, we have the uniform bounds

$$
|\delta^{1,\varepsilon}(s)| \le L_z, \quad |\delta^{2,\varepsilon,N}(s)| \le L_\zeta, \quad s \in \big[0, 1/\varepsilon\big]. \tag{4.23}
$$

Moreover, by Hypothesis **(Hψ)**-d), Remark 4.3 insures the existence of two functions $\gamma_s^{1,\varepsilon}$ and $\gamma_s^{2,\varepsilon,N}$, depending also on $U^\varepsilon, \bar U, \hat\Theta^\varepsilon, \check\Theta^N$, such that

$$\begin{aligned} &C_1(1\wedge |w|_H) \le \gamma^{1,\varepsilon}(w) \le C_2(1\wedge |w|_\Xi), \quad C_1 \in (-1,0], C_2 \ge 0,\\ &\bar C_1(1\wedge |w|_K) \le \gamma^{2,\varepsilon,N}(w) \le \bar C_2(1\wedge |w|_K), \quad \bar C_1 \in (-1,0], \bar C_2 \ge 0, \end{aligned} \tag{4.24}$$

and such that

$$\begin{aligned} &\psi(X_{\varepsilon s}, \hat Q^\varepsilon_s, \bar Z_{\varepsilon s}, \hat\Xi^\varepsilon_s, U^\varepsilon_{\varepsilon s}(\cdot), \hat\Theta^\varepsilon_s(\cdot)) - \psi(X_{\varepsilon s}, \hat Q^\varepsilon_s, \bar Z_{\varepsilon s}, \hat\Xi^\varepsilon_s, \bar U_{\varepsilon s}(\cdot), \hat\Theta^\varepsilon_s(\cdot))\\ &= \int_{H\setminus 0} \gamma_s^{1,\varepsilon}(w)(U^\varepsilon_{\varepsilon s}(w) - \bar U_{\varepsilon s}(w))\nu_1(\mathrm{d}w) \end{aligned}$$

and

$$\begin{aligned} &\psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde Z^N_{\varepsilon s}, \check\Xi^N_s, \tilde U^N_{\varepsilon s}(\cdot), \hat\Theta^\varepsilon_s(\cdot)) - \psi(X^N_{\varepsilon s}, \hat{\mathcal{Q}}^N_s, \tilde Z^N_{\varepsilon s}, \check\Xi^N_s, \tilde U^N_{\varepsilon s}(\cdot), \check\Theta^N_s(\cdot))\\ &= \int_{K\setminus 0} \gamma_s^{2,\varepsilon,N}(w)(\hat\Theta^\varepsilon_s(w) - \check\Theta^N_s(w))\nu_2(\mathrm{d}w). \end{aligned}$$

Identity (4.20) reads

$$\begin{aligned} Y^\varepsilon_0 - \bar Y_0 &= \varepsilon \sum_{k=0}^{2^N-1}\Big(\check Y^{N,k}_{\frac{t_k}{\varepsilon}} - \check Y^{N,k}_{\frac{t_{k+1}}{\varepsilon}}\Big) + \varepsilon\int_0^{\frac{1}{\varepsilon}} R^{\varepsilon,N}_s \mathrm{d}s\\ &\quad + \varepsilon\int_0^{\frac{1}{\varepsilon}} \delta^{1,\varepsilon}_s (Z^\varepsilon_{\varepsilon s} - \bar Z_{\varepsilon s})\mathrm{d}s + \varepsilon\int_0^{\frac{1}{\varepsilon}} \delta^{2,\varepsilon,N}_s (\hat\Xi^\varepsilon_s - \check\Xi^N_s)\mathrm{d}s\\ &\quad + \varepsilon\int_0^{\frac{1}{\varepsilon}}\int_{H\setminus 0} \gamma_s^{1,\varepsilon}(w)(U^\varepsilon_{\varepsilon s}(w) - \bar U_{\varepsilon s}(w))\nu_1(\mathrm{d}w)\mathrm{d}s\\ &\quad + \varepsilon\int_0^{\frac{1}{\varepsilon}}\int_{K\setminus 0} \gamma_s^{2,\varepsilon,N}(w)(\hat\Theta^\varepsilon_s(w) - \check\Theta^N_s(w))\nu_2(\mathrm{d}w)\mathrm{d}s\\ &\quad - \varepsilon\int_0^{\frac{1}{\varepsilon}}\Big(\lambda(X_{\varepsilon s}, \bar Z_{\varepsilon s}, \bar U_{\varepsilon s}(\cdot)) - \lambda(X^N_{\varepsilon s}, \tilde Z^N_{\varepsilon s}, \tilde U^N_{\varepsilon s}(\cdot))\Big)\mathrm{d}s\\ &\quad - \int_0^1 (Z^\varepsilon_t - \bar Z_t)dW^1_t - \sqrt{\varepsilon}\int_0^1 \Big(\check\Xi^N_{\frac{t}{\varepsilon}} - \hat\Xi^\varepsilon_{\frac{t}{\varepsilon}}\Big)dW^2_t\\ &\quad - \int_0^1\int_{H\setminus 0} (U^\varepsilon_t(w) - \bar U_t(w))(N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t)\\ &\quad - \varepsilon\int_0^1\int_{K\setminus 0} (\check\Theta^N_{\frac{t}{\varepsilon}}(w) - \hat\Theta^\varepsilon_{\frac{t}{\varepsilon}}(w))\Big(N^\varepsilon_2(\mathrm{d}t\,\mathrm{d}w) - \frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big), \end{aligned}$$

and, rescaling the time, it gives

$$\begin{aligned} Y^\varepsilon_0 - \bar Y_0 &= \varepsilon \sum_{k=0}^{2^N-1}\Big(\check Y^{N,k}_{\frac{t_k}{\varepsilon}} - \check Y^{N,k}_{\frac{t_{k+1}}{\varepsilon}}\Big) + \int_0^1 R^{\varepsilon,N}_{\frac{t}{\varepsilon}} \mathrm{d}t\\ &\quad - \int_0^1 \Big(\lambda(X_t, \bar Z_t, \bar U_t(\cdot)) - \lambda(X^N_t, \tilde Z^N_t, \tilde U^N_t(\cdot))\Big)\mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&+\int_0^1 \delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}(Z^\varepsilon_t-\bar Z_t)\mathrm{d}t-\int_0^1(Z^\varepsilon_t-\bar Z_t)dW^1_t\\
&+\int_0^1 \delta^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}\Big(\hat\Xi^\varepsilon_{\frac{t}{\varepsilon}}-\check\Xi^N_{\frac{t}{\varepsilon}}\Big)\mathrm{d}t-\sqrt{\varepsilon}\int_0^1\Big(\check\Xi^N_{\frac{t}{\varepsilon}}-\hat\Xi^\varepsilon_{\frac{t}{\varepsilon}}\Big)dW^2_t\\
&-\int_0^1\int_{H\setminus 0}(U^\varepsilon_t(w)-\bar U_t(w))\Big(N_1(\mathrm{d}t\,\mathrm{d}w)-\big(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1\big)\nu_1(\mathrm{d}w)\mathrm{d}t\Big)\\
&-\varepsilon\int_0^1\int_{K\setminus 0}\Big(\check\Theta^N_{\frac{t}{\varepsilon}}(w)-\hat\Theta^\varepsilon_{\frac{t}{\varepsilon}}(w)\Big)\Big(N^\varepsilon_2(\mathrm{d}t\,\mathrm{d}w)-\big(\gamma^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}(w)+1\big)\frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big).
\end{aligned}
\tag{4.25}
$$

Step 3. We can now simplify equation (4.25) by applying an appropriate change in probability (see Eq. (4.28)). For further purposes we also carefully study integrability of the negative powers of the Girsanov density (see relations (4.37) and (4.38) below). It is worth noting that the argument here differs from the purely diffusive case.

We set, for $s\in[0,1]$,

$$
M^1_s:=\int_{[0,\,s]}\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}dW^1_t+\int_{]0,\,s]}\int_{H\setminus 0}\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)(N_1(\mathrm{d}t\,\mathrm{d}w)-\nu_1(\mathrm{d}w)\mathrm{d}t),\tag{4.26}
$$

$$
M^2_s:=\int_{[0,\,s]}\frac{1}{\sqrt{\varepsilon}}\delta^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}dW^2_t+\int_{]0,\,s]}\int_{K\setminus 0}\gamma^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}(w)\Big(N^\varepsilon_2(\mathrm{d}t\,\mathrm{d}w)-\frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big).\tag{4.27}
$$

Notice that the jumps of M^1 and M^2 in (4.26)-(4.27) are completely described by the discontinuities of the processes X and Q (see system (4.1)), namely

$$
\Delta M^1_s=\gamma^{1,\varepsilon}_{\frac{s}{\varepsilon}}(\Delta X_s),\qquad \Delta M^2_s:=\gamma^{2,\varepsilon,N}_{\frac{s}{\varepsilon}}(\Delta Q^\varepsilon_{\frac{s}{\varepsilon}}),\quad s\in[0,1],
$$

where for any càdlàg process L we have set $\Delta L_s:=L_s-L_{s^-}$. We define (see (A.1))

$$
\begin{aligned}
\frac{d\mathbb{P}^{\varepsilon,N}}{d\mathbb{P}}:=\mathcal{E}_1(M^1+M^2)=\,&\mathrm{e}^{\int_0^1\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}dW^1_t-\frac12\int_0^1\left|\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}\right|^2\mathrm{d}t}\,\mathrm{e}^{-\int_{]0,\,1]}\int_{H\setminus 0}\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(\ell)\nu_1(d\ell)\mathrm{d}t}\\
&\cdot\mathrm{e}^{\int_0^1\frac{1}{\sqrt{\varepsilon}}\delta^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}dW^2_t-\frac12\int_0^1\frac{1}{\varepsilon}\left|\delta^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}\right|^2\mathrm{d}t}\,\mathrm{e}^{-\int_{]0,\,1]}\int_{K\setminus 0}\gamma^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}(\ell)\frac{1}{\varepsilon}\nu_2(d\ell)\mathrm{d}t}\\
&\cdot\prod_{n\geq 1:T_n,S_n\leq 1}\Big(\gamma^{1,\varepsilon}_{\frac{T_n}{\varepsilon}}(\Delta X_{T_n})+\gamma^{2,\varepsilon,N}_{\frac{S_n}{\varepsilon}}(\Delta Q^\varepsilon_{\frac{S_n}{\varepsilon}})+1\Big),
\end{aligned}
$$

where T_n (resp. S_n) denote the jump times of $N_1(\mathrm{d}t\,\mathrm{d}w)$ (resp. $N_2(\mathrm{d}t\,\mathrm{d}w)$).

By the Girsanov Theorem A.4 (notice that, by (4.24), $\gamma^{i,\varepsilon}$ is uniformly bounded and $\gamma^{i,\varepsilon}(\omega,t,\cdot)\in L^2(\nu_i)$, $i=1,2$, for all $(\omega,t)\in\Omega\times[0,\,T]$), the compensator of $(N_1(\mathrm{d}t\,\mathrm{d}z),N^\varepsilon_2(\mathrm{d}t\,\mathrm{d}z))$ under $\mathbb{P}^\varepsilon$ is

$$
\Big(\big(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1\big)\nu_1(\mathrm{d}w)\mathrm{d}t,\big(\gamma^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}(w)+1\big)\frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big),
$$

and the process

$$
(\tilde W^1,\tilde W^2):=\Big(W^1-\int_{[0,\,\cdot]}\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}\mathrm{d}t,W^2-\int_{[0,\,\cdot]}\frac{1}{\sqrt{\varepsilon}}\delta^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}\mathrm{d}t\Big)
$$

is such that $\tilde{W}^1$ and $\tilde{W}^2$ are cylindrical Wiener processes under $\mathbb{P}^{\varepsilon,N}$. Taking in (4.25) the expectation $\mathbb{E}^{\varepsilon,N}$ under the new probability $\mathbb{P}^{\varepsilon,N}$, we get

$$\begin{aligned}
Y_0^\varepsilon - \bar{Y}_0 = \varepsilon \sum_{k=0}^{2^N-1} \mathbb{E}^{\varepsilon,N}\Big[\check{Y}^{N,k}_{\frac{t_k}{\varepsilon}} - \check{Y}^{N,k}_{\frac{t_{k+1}}{\varepsilon}}\Big] + \mathbb{E}^{\varepsilon,N}\Big[\int_0^1 R^{\varepsilon,N}_{\frac{t}{\varepsilon}}\mathrm{d}t\Big] \\
- \mathbb{E}^{\varepsilon,N}\Big[\int_0^1 \Big(\lambda(X_t,\bar{Z}_t,\bar{U}_t(\cdot)) - \lambda(X_t^N,\tilde{Z}_t^N,\tilde{U}_t^N(\cdot))\Big)\mathrm{d}t\Big]. \qquad (4.28)
\end{aligned}$$

Recalling estimate (4.22), and the fact that, by Theorem 4.5, there exist some positive constants L'_x, L'_z, L'_u such that

$$\begin{aligned}
&|\lambda(X_t,\bar{Z}_t,\bar{U}_t(\cdot)) - \lambda(X_t^N,\tilde{Z}_t^N,\tilde{U}_t^N(\cdot))| \\
&\le L'_x\Big(1+|\bar{Z}_t|_{\Xi^*} + \Big(\int_{H\setminus 0}|\bar{U}_t(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)|X_t - X_t^N|_H \\
&+ L'_z|\bar{Z}_t - \tilde{Z}_t^N|_{\Xi^*} + L'_u\Big(\int_{H\setminus 0}|\bar{U}_t(w) - \tilde{U}_t^N(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}},
\end{aligned}$$

previous identity yields

$$\begin{aligned}
Y_0^\varepsilon - \bar{Y}_0 \le{}& \varepsilon \sum_{k=0}^{2^N-1} \mathbb{E}^{\varepsilon,N}\Big[\check{Y}^{N,k}_{\frac{t_k}{\varepsilon}} - \check{Y}^{N,k}_{\frac{t_{k+1}}{\varepsilon}}\Big] + (L_z + L'_z)\mathbb{E}^{\varepsilon,N}\Big[\int_0^1 |\bar{Z}_t - \tilde{Z}_t^N|_{\Xi^*}\mathrm{d}t\Big] \\
&+ (L_u + L'_u)\mathbb{E}^{\varepsilon,N}\Big[\int_0^1 \Big(\int_{H\setminus 0}|\bar{U}_t(w) - \tilde{U}_t^N(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\mathrm{d}t\Big] \\
&+ (L_x + L'_x)\mathbb{E}^{\varepsilon,N}\Big[\int_0^1 \Big(1+|\bar{Z}_t|_{\Xi^*} + \Big(\int_{H\setminus 0}|\bar{U}_t(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)|X_t - X_t^N|_H\mathrm{d}t\Big] \\
&+ L_q\mathbb{E}^{\varepsilon,N}\Big[\int_0^1 \Big(1+|\bar{Z}_t|_{\Xi^*} + \Big(\int_{H\setminus 0}|\bar{U}_t(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)\big|\hat{Q}^\varepsilon_{\frac{t}{\varepsilon}} - \hat{\mathcal{Q}}^N_{\frac{t}{\varepsilon}}\big|_K\mathrm{d}t\Big] \\
&=: I_1 + I_2 + I_3 + I_4 + I_5. \qquad (4.29)
\end{aligned}$$

We notice that

$$\left\{\begin{array}{l}
dX_t = AX_t\mathrm{d}t - R\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}\mathrm{d}t + Rd\tilde{W}_t^1 + R\int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\
X_0 = x_0, \\
-d\bar{Y}_t = \lambda(X_t,\bar{Z}_t,\bar{U}_t(\cdot))\mathrm{d}t - \bar{Z}_t\big[d\tilde{W}_t^1 - \delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}\mathrm{d}t\big] - \int_{H\setminus 0}\bar{U}_t(w)[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\
\bar{Y}_1 = h(X_1).
\end{array}\right. \qquad (4.30)$$

We apply the Girsanov Theorem A.1 with $\nu(\omega,\mathrm{d}t\,d\ell) = \big(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(\omega,\ell)+1\big)\nu_1(d\ell)\mathrm{d}t$, $\beta(\omega,t) = \delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}(\omega)$ and $\Gamma(\omega,t,\ell) = \big(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(\omega,\ell)+1\big)^{-1}$. If follows that, setting

$$\tilde{\mathbb{P}}^\varepsilon := \rho\,\mathbb{P}^{\varepsilon,N}$$

with

$$\rho = \mathrm{e}^{\int_0^1 \delta^{1,\varepsilon}_{\frac{t}{\varepsilon}} d\tilde W^1_t - \frac{1}{2}\int_0^1 |\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}|^2 \mathrm{d}t} \mathrm{e}^{\int_{]0,\,1]}\int_{H\setminus 0} \gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)\nu_1(\mathrm{d}w)\mathrm{d}t} \prod_{n\geq 1: T_n\leq 1} \big(\gamma^{1,\varepsilon}_{\frac{T_n}{\varepsilon}}(\Delta X_{T_n}) + 1\big)^{-1}, \tag{4.31}$$

then

$$\tilde W^1_s - \int_{[0,\,s]} \delta^{1,\varepsilon}_{\frac{t}{\varepsilon}} \mathrm{d}t, \quad s\in[0,1],$$

is a cylindrical Wiener process, and the compensator of $N_1(\omega, \mathrm{d}t\, d\ell)$ is $\nu_1(d\ell)\mathrm{d}t$, while $\tilde W^2$ remains a Wiener process and the compensator of $N_2^\varepsilon(\omega, \mathrm{d}t\, d\ell)$ is $(\gamma^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}(\omega,\ell)+1)\dfrac{\nu_2(d\ell)}{\varepsilon}\,\mathrm{d}t$.

In particular, by the uniqueness of the solution of the forward-backward system (4.30) (see Thm. 4.6), the laws of the process (X) under $\tilde{\mathbb{P}}^\varepsilon$ and $\mathbb{P}$ coincide. Moreover, also the laws of $(\bar Z)$ and $(\tilde Z^N)$ and of $(\bar U(\cdot))$ and $(\tilde U^N(\cdot))$ under $\tilde{\mathbb{P}}^\varepsilon$ coincide with the corresponding laws under $\mathbb{P}$. As a matter of fact, again by Theorem 4.6, $\bar Z_t = \bar\zeta(X_t)$ (resp. $\bar U_t(\cdot) = \bar\vartheta(X_t)(\cdot)$) where $\bar\zeta$ (resp. $\bar\vartheta$) is a deterministic Borel-measurable function from H to Ξ^* (resp. to $L^2(\nu_1)$), so the laws of $(\bar Z)$ and $(\tilde Z^N)$ (resp. of $(\bar U(\cdot))$ and $(\tilde U^N(\cdot))$) depend only on the law of (X) in a non-anticipating way. From these considerations it follows that

$$\mathbb{E}^{\varepsilon,N}\Big[\rho\int_0^1 |\bar Z_t|^2_{\Xi^*}\mathrm{d}t\Big] = \mathbb{E}\Big[\int_0^1 |\bar Z_t|^2_{\Xi^*}\mathrm{d}t\Big] < +\infty, \tag{4.32}$$

$$\mathbb{E}^{\varepsilon,N}\Big[\rho\int_0^1\int_{H\setminus 0} |\bar U_t(w)|^2\nu_1(\mathrm{d}w)\mathrm{d}t\Big] = \mathbb{E}\Big[\int_0^1\int_{H\setminus 0} |\bar U_t(w)|^2\nu_1(\mathrm{d}w)\mathrm{d}t\Big] < +\infty. \tag{4.33}$$

Now we notice that, by (4.23) and (4.24), there exist constants η, C and M such that

$$|\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)| \leq C_1\vee C_2, \qquad 0<\eta\leq (\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1)^{-1}\leq C, \quad |\delta^{1,\varepsilon}_{t/\varepsilon}|\leq M. \tag{4.34}$$

In particular , for every $p\geq 1$,

$$(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1)^p \leq \frac{1}{\eta^p}. \tag{4.35}$$

Recalling (4.31), for every $p\geq 1$ we have

$$\begin{aligned}
\rho^{-p} &= \mathrm{e}^{-\int_0^1 p\,\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}} d\tilde W_t + \frac{1}{2}\int_0^1 p|\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}|^2\mathrm{d}t} \mathrm{e}^{-\int_{]0,\,1]}\int_{H\setminus 0} p\,\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)\nu_1(\mathrm{d}w)\mathrm{d}t} \prod_{n\geq 1:T_n\leq 1}\big(\gamma^{1,\varepsilon}_{\frac{T_n}{\varepsilon}}(\Delta X_{T_n})+1\big)^p \\
&= \mathrm{e}^{-\int_0^1 p\,\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}} d\tilde W_t + \frac{1}{2}\int_0^1 p|\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}|^2\mathrm{d}t} \mathrm{e}^{-\int_{]0,\,1]}\int_{H\setminus 0} p\,\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)\nu_1(\mathrm{d}w)\mathrm{d}t} \\
&\quad \mathrm{e}^{\pm\int_{]0,\,1]}\int_{H\setminus 0}\big[1-\big(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1\big)^p\big]\big(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1\big)\nu_1(dw)\mathrm{d}t} \mathrm{e}^{\pm\frac{1}{2}\int_0^1 p^2|\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}|^2\mathrm{d}t} \prod_{n\geq 1:T_n\leq 1}\big(\gamma^{1,\varepsilon}_{\frac{T_n}{\varepsilon}}(\Delta X_{T_n})+1\big)^p \\
&= \mathrm{e}^{\frac{1}{2}\int_0^1 (p+p^2)|\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}|^2\mathrm{d}t} \mathrm{e}^{\int_{]0,\,1]}\int_{H\setminus 0}\big(-1-\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)(p+1)+\big(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1\big)^{p+1}\big)\nu_1(\mathrm{d}w)\mathrm{d}t}\,\tilde\rho_p(1),
\end{aligned}$$

with $\tilde{\rho}_p$ the exponential martingale defined as

$$\tilde{\rho}_p(r) := \mathrm{e}^{-\int_0^r p\, \delta^{1,\varepsilon}_{\frac{t}{\varepsilon}} d\tilde{W}_t - \frac{1}{2}\int_0^r p^2 |\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}|^2 \mathrm{d}t}\, \mathrm{e}^{\int_{]0,\, r]} \int_{H \setminus 0} \left[1 - \left(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1\right)^p\right] \left(\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1\right) \nu_1(\mathrm{d}w)\mathrm{d}t} \cdot \prod_{n \geq r : T_n \leq 1} \left(\gamma^{1,\varepsilon}_{\frac{T_n}{\varepsilon}}(\Delta X_{T_n}) + 1\right)^p. \tag{4.36}$$

It follows that for every $p \geq 1$

$$\rho^{-p} \leq \kappa_p\, \tilde{\rho}_p(1), \tag{4.37}$$

where we have set

$$\kappa_p := \mathrm{e}^{\frac{1}{2}(p-p^2)M^2} \mathrm{e}^{(p\left(1-\frac{1}{C}\right) - \frac{1}{C} + \frac{1}{\eta^{p+1}})\nu_1(H)}.$$

Since $\mathbb{E}^{\varepsilon,N}[\tilde{\rho}_p(1)] = 1$, we get that, for every $p \geq 1$,

$$\mathbb{E}^{\varepsilon,N}[\rho^{-p}] \leq \kappa_p. \tag{4.38}$$

Step 4. We are in a position to complete the proof by estimating the the terms $I_1 - I_5$ in (4.29). We set

$$\Delta_{X,N} := \sup_{t \in [0,1]} |X_t - X_t^N|_H, \quad \Delta_{Z,N} := \int_0^1 |\bar{Z}_t - \tilde{Z}_t^N|_{\Xi}^2 \mathrm{d}t,$$
$$\Delta_{U,N} := \int_0^1 \int_{H \setminus 0} |\bar{U}_t(w) - \tilde{U}_t^N(w)|^2 \nu_1(\mathrm{d}w)\mathrm{d}t,$$

where X^N, $\tilde{Z}^N$ and $\tilde{U}^N$ are the processes defined in (4.8)–(4.9)–(4.10).

Below C will denote a constant independent of N and ε, that may vary from line to line. Let us start by considering I_4. Applying the Hölder inequality and using (4.32)–(4.33) and (4.38), we have

$$\begin{aligned}
I_4 &= (L_x + L_x')\mathbb{E}^{\varepsilon,N}\Big[\int_0^1 \Big(1 + |\bar{Z}_t|_{\Xi^*} + \Big(\int_{H\setminus 0} |\bar{U}_t(w)|^2 \nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)|X_t - X_t^N|_H \mathrm{d}t\Big] \\
&\leq (L_x + L_x')\mathbb{E}^{\varepsilon,N}\Big[\Delta_{X,N} \int_0^1 \Big(1 + |\bar{Z}_t|_{\Xi^*} + \Big(\int_{H\setminus 0} |\bar{U}_t(w)|^2 \nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)\mathrm{d}t\Big] \\
&= (L_x + L_x')\mathbb{E}^{\varepsilon,N}\Big[\rho^{-3/4}(\rho^{1/4}\Delta_{X,N})\rho^{1/2} \int_0^1 \Big(1 + |\bar{Z}_t|_{\Xi^*} + \Big(\int_{H\setminus 0} |\bar{U}_t(w)|^2 \nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)\mathrm{d}t\Big] \\
&\leq (L_x + L_x')(\mathbb{E}^{\varepsilon,N}[\rho^{-3}])^{1/4}(\mathbb{E}^{\varepsilon,N}[\rho\, \Delta_{X,N}^4])^{1/4} \cdot \\
&\quad \cdot \Big(2\mathbb{E}^{\varepsilon,N}\Big[\rho \int_0^1 \Big(1 + |\bar{Z}_t|_{\Xi^*}^2 + \int_{H\setminus 0} |\bar{U}_t(w)|^2 \nu_1(\mathrm{d}w)\Big)\mathrm{d}t\Big]\Big)^{1/4} \\
&\leq C\, (\mathbb{E}[\Delta_{X,N}^4])^{1/4}.
\end{aligned} \tag{4.39}$$

Concerning I_2, applying the Hölder inequality, (4.32) for $\bar{Z}_t - \tilde{Z}_t^N$ in place of $\bar{Z}_t$, and (4.38), we get

$$\begin{aligned}
I_2 &= (L_z + L_z')\mathbb{E}^{\varepsilon,N}\Big[\int_0^1 |\bar{Z}_t - \tilde{Z}_t^N|_{\Xi^*}\mathrm{d}t\Big] \\
&= (L_z + L_z')\mathbb{E}^{\varepsilon,N}\Big[\rho^{-1/2}\rho^{1/2}\int_0^1 |\bar{Z}_t - \tilde{Z}_t^N|_{\Xi^*}\mathrm{d}t\Big] \\
&\le (L_z + L_z')(\mathbb{E}^{\varepsilon,N}[\rho^{-1}])^{1/2}\Big(\mathbb{E}^{\varepsilon,N}\Big[\rho\int_0^1 |\bar{Z}_t - \tilde{Z}_t^N|^2_{\Xi^*}\mathrm{d}t\Big]\Big)^{1/2} \\
&\le C(\mathbb{E}[\Delta_{Z,N}])^{1/2}. \qquad (4.40)
\end{aligned}$$

Analogously, for I_3 (using (4.33) for $\bar{U}_t - \tilde{U}_t^N$ in place of $\bar{U}_t$) we have

$$\begin{aligned}
I_3 &= (L_u + L_u')\mathbb{E}^{\varepsilon,N}\Big[\int_0^1 \Big(\int_{H\setminus 0} |\bar{U}_t(w) - \tilde{U}_t^N(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\mathrm{d}t\Big] \\
&\le (L_z + L_z')(\mathbb{E}^{\varepsilon,N}[\rho^{-1}])^{1/2}\Big(\mathbb{E}^{\varepsilon,N}\Big[\rho\int_0^1\int_{H\setminus 0} |\bar{U}_t(w) - \tilde{U}_t^N(w)|^2\nu_1(\mathrm{d}w)\mathrm{d}t\Big]\Big)^{1/2} \\
&\le C(\mathbb{E}[\Delta_{U,N}])^{1/2}. \qquad (4.41)
\end{aligned}$$

Moreover

$$\begin{aligned}
I_1 &= \varepsilon\sum_{k=0}^{2^N-1}\mathbb{E}^{\varepsilon,N}\Big[\check{Y}^{N,k}_{\frac{t_k}{\varepsilon}} - \check{Y}^{N,k}_{\frac{t_{k+1}}{\varepsilon}}\Big] \qquad (4.42) \\
&\le \varepsilon c\sum_{k=0}^{2^N-1}\mathbb{E}^{\varepsilon,N}\Big[\Big(1 + |\tilde{Z}^N_{t_k}| + \Big(\int_{H\setminus 0}|\tilde{U}^N_{t_k}(\cdot)|^2\nu_1(d\ell)\Big)^{\frac{1}{2}}\Big)(1 + |\hat{\mathcal{Q}}^N_{\frac{t_k}{\varepsilon}}|_K + |\hat{\mathcal{Q}}^N_{\frac{t_k+1}{\varepsilon}}|_K)\Big], \\
&\le \varepsilon c\sum_{k=0}^{2^N-1}\Big(\mathbb{E}^{\varepsilon,N}\Big[\rho\Big(1 + |\tilde{Z}^N_{t_k}|^2 + \Big(\int_{H\setminus 0}|\tilde{U}^N_{t_k}(\cdot)|^2\nu_1(d\ell)\Big)\Big]\Big)^{1/2}\cdot \\
&\cdot\Big(\mathbb{E}^{\varepsilon,N}\Big[\rho^{-1}\Big(1 + |\hat{\mathcal{Q}}^N_{\frac{t_k}{\varepsilon}}|_K + |\hat{\mathcal{Q}}^N_{\frac{t_k+1}{\varepsilon}}|_K\Big)^2\Big]\Big)^{1/2}, \\
&\le \varepsilon c\,\kappa_1^{\frac{1}{2}}\sum_{k=0}^{2^N-1}\Big(\mathbb{E}^{\varepsilon,N}\Big[\rho\Big(1 + |\tilde{Z}^N_{t_k}|^2 + \Big(\int_{H\setminus 0}|\tilde{U}^N_{t_k}(\cdot)|^2\nu_1(d\ell)\Big)\Big]\Big)^{1/2}\cdot \\
&\cdot\Big(\mathbb{E}^{\varepsilon,N}\Big[\tilde{\rho}_1(1)\Big(1 + |\hat{\mathcal{Q}}^N_{\frac{t_k}{\varepsilon}}|_K + |\hat{\mathcal{Q}}^N_{\frac{t_k+1}{\varepsilon}}|_K\Big)^2\Big]\Big)^{1/2},
\end{aligned}$$

where $\tilde{\rho}_1(1)$ is the process defined in (4.36) with $p = r = 1$, κ_1 is the constant defined in (4.37) with $p = 1$, and c is the constant (independent of k and N) appearing in (4.15).

We define a new probability $\tilde{\mathbb{P}}^{\varepsilon,N}$ by

$$d\tilde{\mathbb{P}}^{\varepsilon,N} := \tilde{\rho}_1(1)d\,\mathbb{P}^{\varepsilon,N}.$$

Then last term in the previous inequality can be written as:

$$\mathbb{E}^{\varepsilon,N}\Big[\tilde{\rho}_1(1)\Big(1 + |\hat{\mathcal{Q}}^N_{\frac{t_k}{\varepsilon}}|_K + |\hat{\mathcal{Q}}^N_{\frac{t_k+1}{\varepsilon}}|_K\Big)^2\Big] = \tilde{\mathbb{E}}^{\varepsilon,N}\Big[\Big(1 + |\hat{\mathcal{Q}}^N_{\frac{t_k}{\varepsilon}}|_K + |\hat{\mathcal{Q}}^N_{\frac{t_k+1}{\varepsilon}}|_K\Big)^2\Big]. \qquad (4.43)$$

Again by the Girsanov Theorem A.1, under $\tilde{\mathbb{P}}^{\varepsilon,N}$,

$$\check{W}^1_s = \tilde{W}^1_s + \int_0^s \delta^{1,\varepsilon}_r \, dr, \quad s \in [0,1],$$

is a cylindrical Wiener process, and the compensator of $N_1(\omega, \mathrm{d}t\, d\ell)$ is $(\gamma^{1,\varepsilon,N}_t(\omega,\ell)+1)^2\nu_1(d\ell)\mathrm{d}t$ on $[0,1]$, while $\tilde{W}^2$ remains a Wiener process and $N^\varepsilon_2(\omega, \mathrm{d}t\, d\ell)$ has compensator $(\gamma^{2,\varepsilon,N}_{\frac{t}{\varepsilon}}(\omega,\ell)+1)\dfrac{\nu_2(d\ell)}{\varepsilon}\,\mathrm{d}t$.

Therefore the forward equation (4.30) reads:

$$\begin{cases} dX_t &= AX_t\mathrm{d}t - 2R\delta^{1,\varepsilon}_{\frac{t}{\varepsilon}}\mathrm{d}t + \int_{H\setminus 0} w(2\,\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w) + \gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)^2)\nu_1(\mathrm{d}w)\mathrm{d}t + Rd\check{W}^1_t \\ &\quad + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - (\gamma^{1,\varepsilon}_{\frac{t}{\varepsilon}}(w)+1)^2\nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 &= x_0. \end{cases}$$

Moreover, by (4.16) and (4.13), $(\hat{\mathcal{Q}}^N_t)_{t\in[0,1/\varepsilon]}$ satisfies

$$\begin{cases} d\hat{\mathcal{Q}}^N_t &= [B\hat{\mathcal{Q}}^N_t + F(X_{\varepsilon t}, \hat{\mathcal{Q}}^N_t)]\mathrm{d}t - G\delta^{2,\varepsilon,N}_t\,\mathrm{d}t - \int_{K\setminus 0} w\,\gamma^{2,\varepsilon}_t(w)\nu_2(\mathrm{d}w)\mathrm{d}t + Gd\hat{\tilde{W}}^2_t \\ &\quad + \int_{K\setminus 0} w[N_2(\mathrm{d}t\,\mathrm{d}w) - (1+\gamma^{2,\varepsilon}_t(\omega))\nu_2(\mathrm{d}w)\mathrm{d}t], \\ \hat{\mathcal{Q}}^N_0 &= q_0, \end{cases}$$

with $\hat{\tilde{W}}^2_t = \frac{1}{\sqrt{\varepsilon}}d\tilde{W}^2_{\varepsilon t}$. On the other hand, according to (4.1), $(\hat{Q}^\varepsilon_t = Q^\varepsilon_{\varepsilon t})_{t\in[0,1/\varepsilon]}$ satisfies

$$\begin{cases} d\hat{Q}^\varepsilon_t &= [B\hat{Q}^\varepsilon_t + F(X_{\varepsilon t}, \hat{Q}^\varepsilon_t)]\mathrm{d}t - G\delta^{2,\varepsilon,N}_t\,\mathrm{d}t - \int_{K\setminus 0} w\,\gamma^{2,\varepsilon}_t(w)\nu_2(\mathrm{d}w)\mathrm{d}t + Gd\hat{\tilde{W}}^2_t \\ &\quad + \int_{K\setminus 0} w[N_2(\mathrm{d}t\,\mathrm{d}w) - (1+\gamma^{2,\varepsilon}_t(\omega))\nu_2(\mathrm{d}w)\mathrm{d}t], \\ \hat{Q}^\varepsilon_0 &= q_0. \end{cases}$$

Then by Lemma 3.3 with $g_t = -G\delta^{2,\varepsilon,N}_t - \int_{K\setminus 0} w\,\gamma^{2,\varepsilon}_t(w)\nu_2(\mathrm{d}w)$, and recalling that (4.23), there exists $k_2 < \infty$, independent of ε and N such that,

$$\sup_{t\in[0,\frac{1}{\varepsilon}]} \tilde{\mathbb{E}}^{\varepsilon,N}[|\hat{\mathcal{Q}}^N_t|^2_K] \le k_2\Big(1 + |q_0|^2_K + \sup_{t\in[0,\frac{1}{\varepsilon}]} \tilde{\mathbb{E}}^{\varepsilon,N}[|X_{\varepsilon t}|^2_H] + |G|^2_{L(\Xi,K)}\Big). \tag{4.44}$$

Recalling estimate (3.3),

$$\sup_{t\in[0,\frac{1}{\varepsilon}]} \tilde{\mathbb{E}}^{\varepsilon,N}[|X_{\varepsilon t}|^2_H] \le c\Big(1 + |x_0|^2]\Big)$$

(where the constant c, in view of (4.34), can be chosen independently of ε and N), estimate (4.44) reads

$$\sup_{t\in[0,1/\varepsilon]} \mathbb{E}^{\varepsilon,N}[|\hat{\mathcal{Q}}^N_t|^2_K] \le c. \tag{4.45}$$

Then, plugging (4.45) in (4.42) (taking into account (4.43)) we get:

$$I_1 \le 3\varepsilon c \sum_{k=0}^{2^N-1} \Big(\mathbb{E}^{\varepsilon,N}\Big[\rho\Big(1 + |\tilde{Z}^N_{t_k}|^2 + \int_{H\setminus 0} |\tilde{U}^N_{t_k}(\cdot)|^2\nu_1(d\ell)\Big)\Big]\Big)^{1/2}. \tag{4.46}$$

Proceeding as above, recalling (4.32)–(4.33) and (4.9)–(4.10), we get

$$\begin{aligned}&\mathbb{E}^{\varepsilon,N}\Big[\rho\Big(1+|\tilde{Z}^N_{t_k}|^2+\int_{H\setminus 0}|\tilde{U}^N_{t_k}(\cdot)|^2\nu_1(d\ell)\Big)\Big]\\&\le c\Big(\mathbb{E}\Big[1+2^N\int_{t_{k-1}}^{t_k}|\bar{Z}_t|^2\mathrm{d}t+2^N\int_{t_{k-1}}^{t_k}\int_{H\setminus 0}|\bar{U}^N_t(\cdot)|^2\nu_1(d\ell)\Big]\Big),\end{aligned}\tag{4.47}$$

with c doesn't depend neither on ε nor on N. Plugging in turn (4.47) in (4.46) we achieve

$$\begin{aligned}I_1&\le\varepsilon C\sum_{k=0}^{2^N-1}\Big(\mathbb{E}\Big[1+2^N\int_{t_{k-1}}^{t_k}|\bar{Z}_t|^2\mathrm{d}t+2^N\int_{t_{k-1}}^{t_k}\int_{H\setminus 0}|\bar{U}^N_t(\cdot)|^2\nu_1(d\ell)\Big]\Big)^{1/2}\\&\le\varepsilon C\sum_{k=0}^{2^N-1}\Big(1+2^{\frac{N}{2}}\Big(\mathbb{E}\Big[\int_{t_{k-1}}^{t_k}|\bar{Z}_t|^2\mathrm{d}t\Big]\Big)^{1/2}+2^{\frac{N}{2}}\Big(\mathbb{E}\Big[\int_{t_{k-1}}^{t_k}\int_{H\setminus 0}|\bar{U}^N_t(\cdot)|^2\nu_1(d\ell)\Big]\Big)^{1/2}\Big)\\&\le\varepsilon C\Big(2^N+2^{\frac{3}{2}N}\Big(\mathbb{E}\Big[\int_0^1|\bar{Z}_t|^2\mathrm{d}t\Big]\Big)^{1/2}+2^{\frac{3}{2}N}\Big(\mathbb{E}\Big[\int_0^1\int_{H\setminus 0}|\bar{U}^N_t(\cdot)|^2\nu_1(d\ell)\Big]\Big)^{1/2}\Big)\\&=C\varepsilon(2^N+2^{\frac{3}{2}N}).\end{aligned}\tag{4.48}$$

Concerning now I_5, we notice that $\mathbb{P}$-a.s., for all $s>0$,

$$|\hat{Q}^\varepsilon_s-\hat{\mathcal{Q}}^N_s|\le K\int_0^s\mathrm{e}^{-\mu(s-t)}|X_{\varepsilon t}-X^N_{\varepsilon t}|\,\mathrm{d}t,\tag{4.49}$$

where μ is the dissipative constant in **(HF+B)**, and K does not depend of ε and N. Thus, thanks to (4.49) and proceeding as for I_4, we deduce that

$$\begin{aligned}I_5&\le L_qK\mathbb{E}^{\varepsilon,N}\Big[\int_0^1\Big(1+|\bar{Z}_t|^2_{\Xi^*}+\int_{H\setminus 0}|\bar{U}_t(w)|^2\nu_1(\mathrm{d}w)\Big)\Big(\int_0^{\frac{t}{\varepsilon}}\mathrm{e}^{-\mu(s-r)}|X_{\varepsilon r}-X^N_{\varepsilon r}|\,dr\Big)\mathrm{d}t\Big]\\&\le C(\mathbb{E}^{\varepsilon,N}[\Delta^4_{X,N}])^{1/4}.\end{aligned}\tag{4.50}$$

Collecting (4.39)–(4.40)–(4.41)–(4.48)–(4.50), we get from (4.29) that

$$Y^\varepsilon_0-\bar{Y}_0\le C(\varepsilon(2^N+2^{\frac{3}{2}N})+(\mathbb{E}[\Delta_{Z,N}])^{1/2}+(\mathbb{E}[\Delta_{U,N}])^{1/2}+(\mathbb{E}[\Delta^4_{X,N}])^{1/4}).\tag{4.51}$$

Finally we notice that, having X càdlàg paths and since $\mathbb{E}[\sup_{t\in[0,1]}|X_t|^4]<\infty$ by (3.3),

$$\mathbb{E}[\Delta^4_{X,N}]\to 0\text{ as }N\to\infty.$$

Moreover, by (4.11)–(4.12),

$$\mathbb{E}[\Delta_{Z,N}]\to 0\quad\text{and}\quad\mathbb{E}[\Delta_{U,N}]\to 0\text{ as }N\to\infty.$$

Therefore, letting first $\varepsilon\to 0$ and then $N\to\infty$ in (4.51) we get the result. □

Remark 4.8. Similar techniques would apply to the more general systems where the slow equation has multiplicative noise term. In particular, the forward equation in (4.1) and (4.6) can be replaced by

$$\begin{cases} dX_t = AX_t \mathrm{d}t + R(X_t)dW_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0, \end{cases}$$

where R is Lipschitz with values in $L(\Xi, H)$, also see [36]. To avoid technicalities and improve readability, we restrict our analysis to the additive case.

On the other hand, the dependence of R on Y_ε would transform our system (4.1) into a fully coupled forward-backward system that is much more difficult to treat and doesn't have such a natural control interpretation.

5. The two scale control problem

One of the purposes of the article is to give a representation of the limit of the value functions of a sequence of optimal control problems for a singularly perturbed infinite dimensional state equation.

5.1. Formulation of the problem

We denote by $\mathcal{A}$ the set of predictable processes $(\alpha_t)_{t\in[0,1]}$ taking values in a complete metric space U. Given a solution (X, Q^ε) of (3.1) and a control $\alpha \in \mathcal{A}$, we associate to them a probability measure $\mathbb{P}^{\varepsilon,\alpha}$ on $(\Omega, \mathcal{F})$ that we are going to describe. Let $b : H \times K \times U \to H$, $\rho : U \to K$, $r : H \times K \times U \times H \to \mathbb{R}$ and $\gamma : U \times K \to \mathbb{R}$ be measurable functions satisfying the following assumptions.

(H$b\rho$) There exist positive constants L_b and M' such that, for every $x, x' \in H$, $q, q' \in K$, $a \in U$,

$$\begin{aligned} |b(x,q,a) - b(x',q',u)|_H &\le L_b(|x - x'|_H + |q - q'|_K), \\ |b(x,q,a)|_H + |\rho(a)|_K &\le M'. \end{aligned}$$

(H$r\gamma$) There exist constants $C_r > 1$, $L_r > 0$ and $C_\gamma > 1$ such that, for every $x \in H$, $q \in K$, $a \in U$, $w \in \Xi$,

$$\begin{gathered} 0 < \eta \le r(x,q,a,w) \le C_r, \\ |r(x,q,a,w) - r(x',q',a,w)| \le L_r(|x - x'|_H + |q - q'|_K), \\ 0 < \eta' \le \gamma(a,w) \le C_\gamma. \end{gathered}$$

We also need to consider the following additional assumption on the operator R.

(H') R admits a bounded right inverse $R^{-1} \in L(H, \Xi)$.

Define the processes

$$\begin{aligned} M^1 &:= \int_{[0,\cdot]} R^{-1} b(X_t, Q_t^\varepsilon, \alpha_t) dW_t^1 + \int_{]0,\cdot]} \int_{H\setminus 0} [r(X_t, Q_t^\varepsilon, \alpha_t, w) - 1](N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t), \\ M^2 &:= \int_{[0,\cdot]} \frac{\rho(\alpha_t)}{\sqrt{\varepsilon}} dW_t^2 + \int_{]0,\cdot]} \int_{K\setminus 0} [\gamma(\alpha_t, w) - 1]\Big(N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}t\Big), \end{aligned}$$

and set

$$\frac{d\mathbb{P}^{\varepsilon,\alpha}}{d\mathbb{P}} := \mathcal{E}_1(M^1 + M^2).$$

By the Girsanov Theorem A.4, the compensator of $(N_1(\mathrm{d}t\,\mathrm{d}z), N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}z))$ under $\mathbb{P}^{\varepsilon,\alpha}$ is $(r(X_t, Q_t^\varepsilon, \alpha_t, w)\nu_1(\mathrm{d}w)\mathrm{d}t, \frac{\gamma(\alpha_t,w)}{\varepsilon}\nu_2(\mathrm{d}w)\mathrm{d}t)$, and the process

$$(\tilde{W}^1, \tilde{W}^2) := \left(W^1 - \int_{[0,\cdot]} R^{-1}b(X_t, Q_t^\varepsilon, \alpha_t)\mathrm{d}t, W^2 - \int_{[0,\cdot]} \frac{\rho(\alpha_t)}{\sqrt{\varepsilon}}\mathrm{d}t\right)$$

is such that $\tilde{W}^1$ and $\tilde{W}^2$ are independent cylindrical Wiener processes under $\mathbb{P}^{\varepsilon,\alpha}$. We can therefore conclude that the solution (X, Q^ε) of (3.1) satisfies as well the system of controlled SDEs:

$$\begin{cases} dX_t = AX_t\mathrm{d}t + b(X_t, Q_t^\varepsilon, \alpha_t)\mathrm{d}t + Rd\tilde{W}_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - r(X_t, Q_t^\varepsilon, \alpha_t, w)\nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0 \in H, \\ \varepsilon dQ_t^\varepsilon = [BQ_t^\varepsilon + F(X_t, Q_t^\varepsilon)]\mathrm{d}t + G\rho(\alpha_t)\mathrm{d}t + \sqrt{\varepsilon}Gd\tilde{W}_t^2 \\ \qquad\qquad + \int_{K\setminus 0} w[\varepsilon N_2^\varepsilon(\mathrm{d}t\,\mathrm{d}w) - \gamma(\alpha_t, w)\nu_2(\mathrm{d}w)\mathrm{d}t], \\ Q_0^\varepsilon = q_0 \in K. \end{cases}$$

Remark 5.1. Similarly to Remark 4.8, see also [36], we could consider a more general controlled slow equation with a multiplicative noise, namely:

$$\begin{cases} dX_t = AX_t\mathrm{d}t + b(X_t, Q_t^\varepsilon, \alpha_t)\mathrm{d}t + R(X_t)d\tilde{W}_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - r(X_t, Q_t^\varepsilon, \alpha_t, w)\nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0 \in H, \end{cases}$$

with R Lipschitz and admitting a uniformly bounded right inverse. Nonetheless, we decided to remain with the additive case to avoid further technicalities. On the other side, allowing R to depend on Q would completely change the technical data of the problem.

5.2. Solution of the problem: the BSDE approach

Let us now consider a measurable function $l : H \times K \times U \to \mathbb{R}$ satisfying the following assumption.

(Hl) There exist positive constants L_l and M_l such that, for every $x, x' \in H$, $q, q' \in K$, $a \in U$,

$$|l(x,q,a) - l(x',q',a)| \leq L_l(|x - x'|_H + |q - q'|_K),$$
$$|l(x,q,a)| \leq M_l.$$

We define, for $x \in H$, $q \in K$, $z, \zeta \in \Xi^*$, $u \in L^2(\nu_1)$, $\vartheta \in L^2(\nu_2)$, the function

$$\psi(x,q,z,\zeta,u,\vartheta) := \inf_{a\in U}\Big\{l(x,q,a) + z[R^{-1}b(x,q,a)] + \zeta\rho(a) + \int_{H\setminus 0} u(w)(r(x,q,a,w)-1)\nu_1(\mathrm{d}w) + \int_{K\setminus 0}\vartheta(w)(\gamma(a,w)-1)\nu_2(\mathrm{d}w)\Big\}. \tag{5.1}$$

Proposition 5.2. *Let assumptions* **(H$b\rho$)**, **(H$r\gamma$)**, **(H')** *and* **(Hl)** *hold true. Then the function ψ in (5.1) verifies hypothesis* **(Hψ)**.

Proof. By standard estimates

$$|\psi(x,q,z,\zeta,u,\vartheta) - \psi(x',q',z',\zeta',u',\vartheta')|$$
$$\leq L_u\Big(\int_{\Xi\setminus 0}|u_s(w) - u_s'(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}} + L_\vartheta\Big(\int_{\Xi\setminus 0}|\vartheta_s(w) - \vartheta_s'(w)|^2\nu_2(\mathrm{d}w)\Big)^{\frac{1}{2}}$$

$$+ L_x\Big(1+|z|_{\Xi^*}+\Big(\int_{H\setminus 0}|u_s(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)|x-x'|_H$$
$$+ L_q\Big(1+|z|_{\Xi^*}+\Big(\int_{H\setminus 0}|u_s(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}\Big)|q-q'|_K,$$

for some constants $L_x, L_q, L_z, L_\zeta, L_u, L_\vartheta > 0$ defined opportunely.

Finally, we have to verify that for every $x \in H$, $q \in K$, $z, \zeta \in \Xi^*$, $u, u' \in L^2(\nu_1)$, $\vartheta, \vartheta' \in L^2(\nu_2)$, there exist measurable functions $\gamma_1 : H \to \mathbb{R}$, $\gamma_2 : K \to \mathbb{R}$ (depending on x, q, z, ζ, u, ϑ, ϑ') such that

$$\psi(x,q,z,\zeta,u,\vartheta)-\psi(x,q,z,\zeta,u',\vartheta) \le \int_{H\setminus 0}(u(w)-u'(w))\gamma_1(w)\nu_1(\mathrm{d}w) \tag{5.2}$$
$$\psi(x,q,z,\zeta,u,\vartheta)-\psi(x,q,z,\zeta,u,\vartheta') \le \int_{K\setminus 0}(\vartheta(w)-\vartheta'(w))\gamma_2(w)\nu_2(\mathrm{d}w), \tag{5.3}$$

and satisfying

$$\begin{aligned}&C_1(1\wedge|w|_H)\le\gamma_1(w)\le C_2(1\wedge|w|_H), \quad C_1\in(-1,0], C_2\ge 0,\\ &\bar C_1(1\wedge|w|_K)\le\gamma_2(w)\le \bar C_2(1\wedge|w|_K), \quad \bar C_1\in(-1,0], \bar C_2\ge 0.\end{aligned}$$

We have

$$\begin{aligned}&\int_{H\setminus 0}u(w)(r(x,q,a,w)-1)\nu_1(\mathrm{d}w)\\ &=\int_{H\setminus 0}u'(w)(r(x,q,a,w)-1)\nu_1(\mathrm{d}w)+\int_{H\setminus 0}(u(w)-u'(w))(r(x,q,a,w)-1)\nu_1(\mathrm{d}w)\\ &\le\int_{H\setminus 0}u'(w)(r(x,q,a,w)-1)\nu_1(\mathrm{d}w)+\int_{H\setminus 0}(u(w)-u'(w))\gamma_1(w)\nu_1(\mathrm{d}w)\end{aligned}$$

with

$$\gamma_1(w):=\sup_{a\in U}(r(x,q,a,w)-1)I_{[0,+\infty)}(u(w)-u'(w))+\inf_{a\in U}(r(x,q,a,w)-1)I_{(0,+\infty)}(u'(w)-u(w)).$$

Notice that, by assumption **(H**$r\gamma$**)**,

$$-1<\eta-1\le\gamma_1(w)\le C_r-1, \quad C_r>1.$$

Adding and subtracting in both sides

$$l(x,q,a)+z[R^{-1}b(x,q,a)]+\zeta\rho(a)+\int_{H\setminus 0}u(w)(r(x,q,a,w)-1)\nu_1(\mathrm{d}w)$$

and taking the infimum over $a \in U$, we get (5.2).

Analogously, one observes that

$$\begin{aligned}&\int_{K\setminus 0}\vartheta(w)(\gamma(a,w)-1)\nu_2(\mathrm{d}w)\le\int_{K\setminus 0}\vartheta'(w)(\gamma(a,w)-1)\nu_2(\mathrm{d}w)\\ &+\int_{K\setminus 0}(\vartheta(w)-\vartheta'(w))\gamma_2(w)\nu_2(\mathrm{d}w)\end{aligned}$$

with

$$\gamma_2(w) := \sup_{a\in U}(\gamma(a,w)-1)I_{[0,+\infty)}(\vartheta(w)-\vartheta'(w)) + \inf_{a\in U}(\gamma(a,w)-1)I_{(0,+\infty)}(\vartheta'(w)-\vartheta(w)).$$

In particular, by assumption $(\mathbf{H}r\gamma)$,

$$-1 < \eta' - 1 \le \gamma_2(w) \le C_\gamma - 1, \quad C_\gamma > 1.$$

Proceeding as before (5.3) follows. □

Let $h : H \to \mathbb{R}$ be a measurable function satisfying $(\mathbf{H}h)$. We consider the following BSDE: $\mathbb{P}$-a.s., for $t \in [0,1]$,

$$\begin{aligned} Y_t^\varepsilon = h(X_1) &+ \int_t^1 \psi\Big(X_s, Q_s^\varepsilon, Z_s^\varepsilon, \frac{\Xi_s^\varepsilon}{\sqrt{\varepsilon}}, U_s^\varepsilon(\cdot), \frac{\Theta_s^\varepsilon(\cdot)}{\varepsilon}\Big)\mathrm{d}s - \int_t^1 Z_s^\varepsilon dW_s^1 - \int_t^1 \Xi_s^\varepsilon dW_s^2 \\ &- \int_t^1\int_{H\setminus 0} U_s^\varepsilon(w)(N_1(\mathrm{d}s\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}s) - \int_t^1\int_{K\setminus 0}\Theta_s^\varepsilon(w)\Big(N_2^\varepsilon(\mathrm{d}s\,\mathrm{d}w) - \frac{\nu_2(\mathrm{d}w)}{\varepsilon}\mathrm{d}s\Big). \end{aligned} \tag{5.4}$$

Lemma 5.3. *Let assumptions* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, $(\mathbf{H}\beta^B)$, $(\mathbf{H}b\rho)$, $(\mathbf{H}r\gamma)$, **(H')**, $(\mathbf{H}l)$ *and* $(\mathbf{H}h)$ *hold true. Then for every* $\varepsilon > 0$, *there exists a unique solution* $(X, Q^\varepsilon, Y^\varepsilon, Z^\varepsilon, \Xi^\varepsilon, U^\varepsilon(\cdot))$ *to the forward-backward system* (5.4), *and*

$$Y_0^\varepsilon = V^\varepsilon(x_0, q_0), \quad x_0 \in H, q_0 \in K,$$

where V^ε *is the value function in* (1.1).

Proof. The well-posedness of the forward-backward system (5.4) directly follows from Theorem 4.1 together with Lemma 5.2. Then the control interpretation of Y_0^ε follows as in [29], that extends to the infinite dimensional framework the classical representation in finite dimension, see *e.g.* [25]. □

The following result is a direct consequence of Theorem 4.7, together with Lemma 5.3.

Theorem 5.4. *Let assumptions* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, $(\mathbf{H}\beta^B)$, $(\mathbf{H}b\rho)$, $(\mathbf{H}r\gamma)$, **(H')**, $(\mathbf{H}l)$ *and* $(\mathbf{H}h)$ *hold true, and let* $(X, \bar{Y}, \bar{Z}, \bar{U}(\cdot))$ *be the unique solution to the forward-backward system* (4.6). *Then*

$$\lim_{\varepsilon\to 0} V^\varepsilon(x_0, q_0) = \bar{Y}_0, \quad x_0 \in H, q_0 \in K,$$

where V^ε, $\varepsilon > 0$, *is the value function in* (1.1).

6. Control interpretation of the limit forward-backward system

Let $(X, \bar{Y}, \bar{Z}, \bar{U}(\cdot))$ be the unique solution to the forward-backward system (4.6). We aim now at interpreting $\bar{Y}_0$ as the value function of a corresponding reduced control problem. To this end, we start by noticing that, for every $t \in [0,1]$,

$$\bar{Y}_t = h(X_1) + \int_t^1 \lambda(X_s, \bar{Z}_s, \bar{U}_s(\cdot))\mathrm{d}s - \int_t^1 \bar{Z}_s dW_s^1 - \int_t^1\int_{H\setminus 0}\bar{U}_s(w)[N_1(\mathrm{d}s\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}s], \tag{6.1}$$

where $(X_t)_{t\geq 0}$ is the mild solution of

$$\begin{cases} dX_t = AX_t\mathrm{d}t + RdW_t^1 + \int_{H\setminus 0} w[N_1(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t], \\ X_0 = x_0, \end{cases}$$

and (see Thm. 4.5) $\lambda : H \times \Xi^* \times L^2(\nu_1) \to \mathbb{R}$ is uniquely determined and satisfies, for fixed $x \in H$, $z, z' \in \Xi^*$, $u, u' \in L^2(\nu_1)$,

$$|\lambda(x,z,u) - \lambda(x,z',u')| \leq L_z|z-z'|_{\Xi^*} + L_u\Big(\int_{H\setminus 0} |u(w)-u'(w)|^2\nu_1(\mathrm{d}w)\Big)^{\frac{1}{2}}, \tag{6.2}$$

$$|\lambda(x,0,0)| \leq c(1+|x|_H), \tag{6.3}$$

for some positive constants L_z, L_u and c. Moreover, $\lambda(\cdot,z,u)$ is concave with respect to z and u, so by the Fenchel-Moreau theorem one can write $\lambda = \lambda_{**}$, where, for all $x \in H$,

$$\lambda_*(x,p,v) = \inf_{z\in\Xi^*, u\in L^2(\nu_1)} (-zp - \langle u, 1-v\rangle_{L^2(\nu_1)} - \lambda(x,z,u)), \quad p \in \Xi, v \in L^2(\nu_1). \tag{6.4}$$

It follows that, for all $x \in H$, $z \in \Xi$ and $u \in L^2(\nu_1)$,

$$\lambda(x,z,u) = \inf_{p\in\Xi,\, v\in L^2(\nu_1)} (-zp - \langle u, 1-v\rangle_{L^2(\nu_1)} - \lambda_*(x,p,v)).$$

Moreover, by the Lipschitzianity of λ we can restrict the infimum to a bounded set, namely the following holds.

Proposition 6.1. *We have*

$$\lambda(x,z,u) = \inf_{\substack{p\in\Xi,\, v\in L^2(\nu_1):\\ |p|\leq L_z,\, |1-v|_{L^2(\nu_1)}\leq L_u}} (-zp - \langle u, 1-v\rangle_{L^2(\nu_1)} - \lambda_*(x,p,v)), \tag{6.5}$$

where L_z and L_u are the Lipschitz constants in (6.2).

Proof. By (6.4) together with (6.3) we get

$$\lambda_*(x,p,v) \leq -\lambda(x,0,0) \leq c(1+|x|), \quad p \in \Xi, v \in L^2(\nu_1).$$

Then, again by (6.4) together with (6.2), we obtain that, for all $p \in \Xi, v \in L^2(\nu_1)$,

$$\lambda_*(x,p,v) \leq \inf_{z\in\Xi^*, u\in L^2(\nu_1)} (-zp - \langle u, 1-v\rangle_{L^2(\nu_1)} + L_z|z|_{\Xi^*} + L_u|u|_{L^2(\nu_1)}) + c.$$

It follows that $\lambda_*(x,p,v) = -\infty$ if $|p|_\Xi > L_z$ or $|1-v|_{L^2(\nu_1)} > L_u$, which implies (6.5). ☐

We now define

$$\begin{aligned}\mathcal{S} := \{(p_s, v_s(\cdot))_{s\in[0,1]} :\, & p \text{ progressively measurable s.t. } |p_s|_\Xi \leq L_z, \\ & v(\cdot) \text{ predictable s.t. } |1-v_s|_{L^2(\nu_1)} \leq L_u\}.\end{aligned}$$

For every $(p, v(\cdot)) \in \mathcal{S}$, let us set

$$M := -\int_{[0,\,\cdot]} p_t^* dW_t^1 + \int_{]0,\,\cdot]} \int_{H\setminus 0} (v_t(w) - 1)[N(\mathrm{d}t\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}t].$$

Being $(1 - v_t(\omega, \cdot)) \in L^2(\nu_1)$ for all $(\omega, t) \in \Omega \times [0,\,1]$, and since, for all $t \in [0, 1]$,

$$\mathbb{E}\Big[\mathrm{e}^{\frac{1}{2}\langle M^c\rangle_t + \langle M^d\rangle_t}\Big] = \mathbb{E}\Big[\mathrm{e}^{\frac{1}{2}\int_0^t |p|_s^2 \mathrm{d}s + \int_0^t \int_{H\setminus 0}(1-v_s(w))^2 \nu_1(\mathrm{d}w)\mathrm{d}s}\Big] \leq \mathbb{E}\Big[\mathrm{e}^{(\frac{1}{2}L_z^2 + L_u)t}\Big] < \infty,$$

by Theorem A.1 together with Corollary A.3, under $\mathbb{P}^{p,v}$ defined by

$$\frac{d\mathbb{P}^{p,v}}{d\mathbb{P}} := \mathcal{E}_1\left(M\right),$$

the process

$$W_t^p := \int_0^t p_s \mathrm{d}s + W_t^1, \quad t \in [0, 1],$$

is a Q_1-Wiener process, and $N_1(\mathrm{d}t\,\mathrm{d}w)$ has compensator $v_t(w)\nu_1(\mathrm{d}w)\mathrm{d}t$.

We can finally give the following result that follows the classical argument relating the BDSEs to control problems. We sketch the proof for the reader convenience.

Theorem 6.2. *Assume that* **(HAB)**, **(HRG)**, **(HF)**, **(HF+B)**, **(Hβ^B)**, **(H$b\rho$)**, **(H$r\gamma$)**, **(H')**, **(Hl)** *and* **(Hh)** *hold true. Then*

$$\bar{Y}_0 = \inf_{(p,v(\cdot))\in\mathcal{S}} \mathbb{E}^{p,v}\Big[h(X_1) - \int_0^1 \lambda_*(X_s, p_s, v_s(\cdot))\mathrm{d}s\Big].$$

Proof. We prove the two inequalities separately.
($\Leftarrow$) Let $(p, v(\cdot)) \in \mathcal{S}$. By Proposition 6.1, for all $t \in [0, 1]$, from (6.1) we get

$$\begin{aligned}
\bar{Y}_t &= h(X_1) + \int_t^1 \lambda(X_s, \bar{Z}_s, \bar{U}_s(\cdot))\mathrm{d}s - \int_t^1 \bar{Z}_s dW_s^1 - \int_t^1 \int_{H\setminus 0} \bar{U}_s(w)[N_1(\mathrm{d}s\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}s] \\
&\leq h(X_1) - \int_t^1 \lambda_*(X_s, p_s, v_s(\cdot))\mathrm{d}s - \int_t^1 \bar{Z}_s p_s \mathrm{d}s - \int_t^1 \int_{H\setminus 0} (1 - v_s(w))\bar{U}_s(w)\nu_1(\mathrm{d}w)\mathrm{d}s \\
&\quad - \int_t^1 \bar{Z}_s dW_s^1 - \int_t^1 \int_{H\setminus 0} \bar{U}_s(w)[N_1(\mathrm{d}s\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}s] \\
&= h(X_1) - \int_t^1 \lambda_*(X_s, p_s, v_s(\cdot))\mathrm{d}s - \int_t^1 \bar{Z}_s dW_s^p \\
&\quad - \int_t^1 \int_{H\setminus 0} \bar{U}_s(w)[N_1(\mathrm{d}s\,\mathrm{d}w) - v_s(w)\nu_1(\mathrm{d}w)\mathrm{d}s].
\end{aligned}$$

Applying the conditional expectation with respect to $\mathbb{P}^{p,v}$ we get

$$\bar{Y}_t \leq \mathbb{E}^{p,v}\Big[h(X_1) - \int_t^1 \lambda_*(X_s, p_s, v_s(\cdot))\mathrm{d}s\Big|\mathcal{F}_t\Big], \quad t \in [0, 1].$$

that provides

$$\bar{Y}_t \leq \inf_{(p,v(\cdot))\in\mathcal{S}} \mathbb{E}^{p,v}\Big[h(X_1) - \int_t^1 \lambda_*(X_s,p_s,v_s(\cdot))\mathrm{d}s\Big|\mathcal{F}_t\Big], \quad t\in[0,1],$$

by the arbitrariness of $(p, v(\cdot) \in \mathcal{S}$.

($\Rightarrow$) Again by Proposition 6.1, we may choose, for any $n \geq 1$, $(p_t^n, v_t^n(\cdot))_{t\in[0,1]}$ such that $|p_s^n|_\Xi \leq L_z$, $|1 - v_s^n|_{L^2(\nu_1)}, \leq L_u$, and

$$\lambda(X_t,\bar{Z}_t,\bar{U}_t(\cdot)) \geq -\bar{Z}_t p_t^n - \langle \bar{U}_t, 1-v_t^n\rangle_{L^2(\nu_1)} - \lambda_*(X_t,p_t^n,v_t^n(\cdot)) - \frac{1}{n}.$$

By a measurable selection theorem (see *e.g.* Thms. 8.2.10 and 8.2.11 in [37]) we can assume that $(p_t^n)_{t\in[0,1]}$ are progressively measurable and $(v_t^n(\cdot))_{t\in[0,1]}$ are predictable.

Recalling (6.1), we get

$$\begin{aligned}
\bar{Y}_t &= h(X_1) + \int_t^1 \lambda(X_s,\bar{Z}_s,\bar{U}_s(\cdot))\mathrm{d}s - \int_t^1 \bar{Z}_s dW_s^1 - \int_t^1\int_{H\setminus 0} \bar{U}_s(w)[N_1(\mathrm{d}s\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}s] \\
&\geq h(X_1) - \int_t^1 \bar{Z}_s p_s^n \mathrm{d}s - \int_t^1\int_{H\setminus 0} \bar{U}_s(w)(1-v_s^n(w))\nu_1(\mathrm{d}w)\mathrm{d}s - \int_t^1 \lambda_*(X_s,p_s^n,v_s^n(\cdot))\mathrm{d}s \\
&\quad - \frac{1-t}{n} - \int_t^1 \bar{Z}_s dW_s^1 - \int_t^1\int_{H\setminus 0} \bar{U}_s(w)[N_1(\mathrm{d}s\,\mathrm{d}w) - \nu_1(\mathrm{d}w)\mathrm{d}s] \\
&= h(X_1) - \int_t^1 \lambda_*(X_s,p_s^n,v_s^n(\cdot))\mathrm{d}s - \frac{1-t}{n} \\
&\quad - \int_t^1 \bar{Z}_s dW_s^{p^n} - \int_t^1\int_{H\setminus 0} \bar{U}_s(w)[N_1(\mathrm{d}s\,\mathrm{d}w) - v_s^n(w)\nu_1(\mathrm{d}w)\mathrm{d}s].
\end{aligned}$$

Applying the conditional expectation with respect to $\mathbb{P}^{p^n,v^n}$ we get

$$\bar{Y}_t + \frac{1-t}{n} \geq \mathbb{E}^{p^n,v^n}\Big[h(X_1) - \int_t^1 \lambda_*(X_s,p_s^n,v_s^n(\cdot))\mathrm{d}s\Big|\mathcal{F}_t\Big], \quad t\in[0,1].$$

We can therefore conclude that

$$\begin{aligned}
\bar{Y}_t &\geq \liminf_{n\to\infty} \mathbb{E}^{p^n,v^n}\Big[h(X_1) - \int_t^1 \lambda_*(X_s,p_s^n,v_s^n(\cdot))\mathrm{d}s\Big|\mathcal{F}_t\Big] \\
&\geq \inf_{(p,v(\cdot))\in\mathcal{S}} \mathbb{E}^{p,v}\Big[h(X_1) - \int_t^1 \lambda_*(X_s,p_s,v_s(\cdot))\mathrm{d}s\Big|\mathcal{F}_t\Big], \quad t\in[0,1].
\end{aligned}$$

□

Acknowledgments

The authors would like to thank the anonymous referees for their careful reading of the manuscript and for their valuable suggestions, which have greatly improved the quality of this work.

Funding

This research was partially supported by the 2023 GNAMPA-INdAM project *Riduzione del modello in sistemi dinamici in dimensione infinita a due scale* and by "Stochastic-particle based anomalous reaction-diffusion models with heterogeneous interaction for radiation therapy" - Next Generation EU, Missione 4. Componente 1-CUP:D53D2301897001.

Data availability statement

The research data associated with this article are included in the article.

References

[1] M.I. Freidlin and A.D. Wentzell, Random perturbations of dynamical systems. Vol. 260 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, 3rd edn. Springer, Heidelberg (2012). Translated from the 1979 Russian original by Joseph Szücs.

[2] Yu.M. Kabanov and S.M. Pergamenshchikov, On optimal control of singularly perturbed stochastic differential Equations, in *Modeling, estimation and control of systems with uncertainty (Sopron, 1990)*. Vol. 10 of *Progr. Systems Control Theory*. Birkhäuser Boston, Boston, MA (1991).

[3] M. Bardi and H. Kouhkouh, Singular perturbations in stochastic optimal control with unbounded data. *ESAIM Control Optim. Calc. Var.* **29** (2023) Paper No. 52, 25.

[4] S. Assing, F. Flandoli and U. Pappalettera, Stochastic model reduction: convergence and applications to climate equations. *J. Evol. Equ.* **21** (2021) 3813–3848.

[5] É. Pardoux and S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, in Stochastic partial differential equations and their applications (Charlotte, NC, 1991). Vol. 176 of *Lect. Notes Control Inf. Sci.*. Springer, Berlin (1992) 200–217.

[6] G. Fabbri, F. Gozzi and A. Swiech, Stochastic optimal control in infinite dimension. Vol. 82 of *Probability Theory and Stochastic Modelling*. Springer, Cham (2017). Dynamic programming and HJB equations, With a contribution by Marco Fuhrman and Gianmario Tessitore.

[7] Yu.M. Kabanov and W.J. Runggaldier, On control of two-scale stochastic systems with linear dynamics in the fast variables. *Math. Control Signals Syst.* **9** (1996) 107–122.

[8] O. Alvarez and M. Bardi, Viscosity solutions methods for singular perturbations in deterministic and stochastic control. *SIAM J. Control Optim.* **40** (2001) 1159–1188.

[9] O. Alvarez and M. Bardi, Singular perturbations of nonlinear degenerate parabolic PDEs: a general convergence result. *Arch. Ration. Mech. Anal.* **170** (2003) 17–61.

[10] O. Alvarez, M. Bardi and C. Marchi, Multiscale problems and homogenization for second-order Hamilton–Jacobi equations. *J. Differ. Equ.* **243** (2007) 349–387.

[11] M. Bardi and A. Cesaroni, Optimal control with random parameters: a multiscale approach. *Eur. J. Control* **17** (2011) 30–45.

[12] M. Bardi, A. Cesaroni and L. Manca, Convergence by viscosity methods in multiscale financial models with stochastic volatility. *SIAM J. Financial Math.* **1** (2010) 230–265.

[13] G. Guatteri and G. Tessitore, Singular limit of BSDEs and optimal control of two scale stochastic systems in infinite dimensional spaces. *Appl. Math. Optim.* **83** (2021) 1025–1051.

[14] M. Fuhrman, Y. Hu and G. Tessitore, Ergodic BSDES and optimal ergodic control in Banach spaces. *SIAM J. Control Optim.* **48** (2009) 1542–1566.

[15] P.Y. Madec, Y. Hu and A. Richou, A probabilistic approach to large time behavior of mild solutions of HJB equations in infinite dimension. *SIAM J. Control Optim.* **53** (2015) 378–398.

[16] G. Guatteri and G. Tessitore, Ergodic BSDEs with multiplicative and degenerate noise. *SIAM J. Control Optim.* **58** (2020) 2050–2077.

[17] A. Swiech, Singular perturbations and optimal control of stochastic systems in infinite dimension: HJB equations and viscosity solutions. *ESAIM Control Optim. Calc. Var.* **27** (2021) Paper No. 6, 34.

[18] C.-E. Bréhier, Orders of convergence in the averaging principle for SPDEs: the case of a stochastically forced slow component. *Stochastic Process. Appl.* **130** (2020) 3325–3368.

[19] C.-E. Bréhier, Strong and weak orders in averaging for SPDEs. *Stochastic Process. Appl.* **122** (2012) 2553–2593.

[20] M. Röckner and L. Xie, Averaging principle and normal deviations for multiscale stochastic systems. *Commun. Math. Phys.* **383** (2021) 1889–1937.

[21] F. De Feo, The order of convergence in the averaging principle for slow-fast systems of stochastic evolution equations in Hilbert spaces. *Appl. Math. Optim.* **88** (2023) 1–36.

[22] A. Święch and J. Zabczyk, Uniqueness for integro-PDE in Hilbert spaces. *Potential Anal.* **38** (2013) 233–259.

[23] A. Święch and J. Zabczyk, Integro-PDE in Hilbert spaces: existence of viscosity solutions. *Potential Anal.* **45** (2016) 703–736.

[24] A. Święch, "Unbounded" second order partial differential equations in infinite-dimensional Hilbert spaces. *Commun. Part. Differ. Equ.* **19** (1994) 1999–2036.

[25] G. Barles, R. Buckdahn and E. Pardoux, Backward stochastic differential equations and integral-partial differential equations. *Stoch. Rep.* **60** (1997) 57–83.
[26] R. Carbone, B. Ferrario and M. Santacroce, Backward stochastic differential equations driven by càdlàg martingales. *Theory Probab. Appl.* **52** (2008) 304–314.
[27] F. Confortola, M. Fuhrman and J. Jacod, Backward stochastic differential equations driven by a marked point process: an elementary approach, with an application to optimal control. *Ann. Appl. Probab.* **26** (2016) 1743–1773.
[28] E. Bandini, Existence and uniqueness for backward stochastic differential equations driven by a random measure, possibly non quasi-left continuous. *Electron. Commun. Probab.* **20** (2015) 1–13.
[29] E. Bandini, F. Confortola and A. Cosso, BSDE representation and randomized dynamic programming principle for stochastic control problems of infinite-dimensional jump-diffusions. *Electron. J. Probab.* **24** (2019) 1–37.
[30] S. Peszat and J. Zabczyk, Stochastic partial differential equations with Lévy noise. Vol, 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (2007).
[31] G. Da Prato and J. Zabczyk, *Ergodicity for infinite-dimensional systems*. Vol. 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge (1996).
[32] M. Fuhrman and G. Tessitore, Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. *Ann. Probab.* **30** (2002) 1397–1465.
[33] F. Confortola and M. Fuhrman, Filtering of continuous-time Markov chains with noise-free observation and applications. *Stochastics* **85** (2013) 216–251.
[34] S.N. Cohen and V Fedyashov, Ergodic BSDEs with jumps and time dependence. Preprint arXiv:1406.4329v2 (2015).
[35] S.N. Cohen and R.J. Elliott, *Stochastic calculus and applications*. Probability and its Applications, 2nd edn. Springer, Cham (2015).
[36] G. Guatteri and G. Tessitore, Singular limit of two-scale stochastic optimal control problems in infinite dimensions by vanishing noise regularization. *SIAM J. Control Optim.* **60** (2022) 575–596.
[37] J.-P. Aubin and H. Frankowska, *Set-valued analysis*. Modern Birkhäuser Classics. Birkhäuser Boston, Inc., Boston, MA (2009). Reprint of the 1990 edition.

Appendix A.

A.1 Girsanov theorems

In the sequel we will denote by $\mathcal{E}(S)$ the Doléans-Dade exponential of general real-valued semimartingale S. We recall that

$$\mathcal{E}_T(S) = \mathrm{e}^{S_T - \frac{1}{2}\langle S^c\rangle_T} \prod_{s\le T} (1+\Delta S_s)\mathrm{e}^{-\Delta S_s}. \tag{A.1}$$

We recall the following version of the Girsanov theorem for martingales, see *e.g.* Theorem 15.3.10 in [35].

Theorem A.1. *Let W be an $(\mathcal{F}_t)$ cylindrical Wiener process on Ξ, and $\mu(\omega, \mathrm{d}t\, \mathrm{d}\ell)$ a locally integrable integer-valued $(\mathcal{F}_t)$-random measure on Ξ, with compensator $\nu(\omega, \mathrm{d}t\, \mathrm{d}\ell)$. Assume that $\nu(\{t\}\times\Xi) = 0$ up to indistinguishability (so μ has no predictable jumps).*

Let $\beta : \Omega \times [0, T] \to \Xi^*$ and $\Gamma : \Omega \times [0, T] \times \Xi \to \mathbb{R}_+$ be predictable functions. Suppose that $\Gamma - 1$ is stochastically ν-integrable, and set

$$M := \int_{[0,\,\cdot]} \beta(t) dW_t + \int_{]0,\,\cdot]} \int_{\Xi\setminus 0} [\Gamma(t,\ell) - 1](\mu(\mathrm{d}t\, d\ell) - \nu(\mathrm{d}t\, d\ell)). \tag{A.2}$$

If $\mathcal{E}(M)$ is a uniformly integrable martingale, then under $\tilde{\mathbb{P}}$ defined by

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} := \mathcal{E}_T(M)$$

the process $W - \int_{[0,\,\cdot]} \beta(t)\mathrm{d}t$ is a cylindrical Wiener process, and the compensator of $\mu(\omega, \mathrm{d}t\, d\ell)$ is $\Gamma(\omega, t, \ell)\nu(\omega, \mathrm{d}t\, d\ell)$.

Remark A.2. Noting that by (A.2)

$$-\Delta M_s = \int_{\Xi\setminus 0} (1 - \Gamma(s,\ell)) N(\{s\} \times d\ell),$$

we have

$$\prod_{s \le T} \mathrm{e}^{-\Delta M_s} = \mathrm{e}^{\sum_{s\le T}(-\Delta M_s)} = \mathrm{e}^{\int_{]0,T]}\int_{\Xi\setminus 0}(1-\Gamma(s,\ell))\mu(\mathrm{d}s\, d\ell)},$$

so (A.1) reads

$$\mathcal{E}_T(M) = \mathrm{e}^{\int_0^T \beta(t) dW_t - \frac{1}{2}\int_0^T |\beta(t)|^2 \mathrm{d}t} \mathrm{e}^{\int_{]0,\,T]}\int_{\Xi\setminus 0}[1-\Gamma(t,\ell)]\nu(d\ell\, \mathrm{d}t)} \prod_{n\ge 1: T_n \le T} \Gamma(T_n, \xi_n),$$

where $(T_n, \xi_n)_n$ denotes the sequence of jump times and positions associated to the counting measure μ.

The following result provides a sufficient condition under which a local martingale M is uniformly integrable, see *e.g.* Corollary 15.4.4 in [35].

Corollary A.3. *Let M be a local martingale with $\Delta M \ge -1$, and let*

$$T = \inf\{t : \Delta M_t = -1\} = \inf\{t : \mathcal{E}_t(M) = 0\}.$$

If M is a locally square integrable local martingale, and

$$\mathbb{E}\left[\mathrm{e}^{\frac{1}{2}\langle M^c\rangle_T + \langle M^d\rangle_T}\right] < \infty,$$

then $\mathcal{E}(M)$ is a uniformly integrable martingale and $\{\mathcal{E}_\infty(M) > 0\} = \{T = \infty\}$ almost surely.

With the help of Corollary A.3, when μ is a Poisson random measure one can specialize Theorem A.1 in the following way.

Theorem A.4. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a fixed filtered probability space. Let W be an $(\mathcal{F}_t)$ cylindrical Wiener process on Ξ, and $N(\mathrm{d}t\, d\ell)$ an $(\mathcal{F}_t)$-Poisson random measure on Ξ, with compensator $F(d\ell)\mathrm{d}s$, such that $\int_{\Xi\setminus 0} ||\ell||^2_\Xi F(d\ell) < \infty$. Suppose we have uniformly bounded functions $\beta : \Omega \times [0, T] \to \Xi^*$, $\Gamma : \Omega \times [0, T] \times \Xi \to \mathbb{R}_+$ such that $\Gamma(\omega, t, \cdot) - 1 \in L^2(F)$ for all $(\omega, t) \in \Omega \times [0, T]$, and define*

$$M := \int_{[0,\,\cdot]} \beta(t) dW_t + \int_{]0,\,\cdot]} \int_{\Xi\setminus 0} [\Gamma(t,\ell) - 1](N(\mathrm{d}t\, d\ell) - F(d\ell)\mathrm{d}t).$$

Then, setting

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} := \mathcal{E}_T\left(M\right),$$

we have that $\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t}$ *is a positive square integrable martingale. Moreover, under* $\tilde{\mathbb{P}}$,

$$W^{\tilde{\mathbb{P}}} := W - \int_{[0,\,\cdot]} \beta(t)\mathrm{d}t$$

is a cylindrical Wiener process, and the compensator of $N(\omega, \mathrm{d}t\, d\ell)$ *is* $\Gamma(\omega, t, \ell)F(d\ell)\mathrm{d}t$.