Journal of Optimization Theory and Applications (2024) 202:1027–1059
https://doi.org/10.1007/s10957-024-02495-2

# Optimal Relaxed Control for a Decoupled *G*-FBSDE

**Hafida Bouanani[1] · Omar Kebiri[2,3] · Carsten Hartmann[3] · Amel Redjil[4]**

Received: 29 September 2022 / Accepted: 9 July 2024 / Published online: 24 July 2024


**Abstract**

In this paper we study a system of decoupled forward-backward stochastic differential equations driven by a $G$-Brownian motion ($G$-FBSDEs) with non-degenerate diffusion. Our objective is to establish the existence of a relaxed optimal control for a non-smooth stochastic optimal control problem. The latter is given in terms of a decoupled $G$-FBSDE. The cost functional is the solution of the backward stochastic differential equation at the initial time. The key idea to establish existence of a relaxed optimal control is to replace the original control problem by a suitably regularised problem with mollified coefficients, prove the existence of a relaxed control, and then pass to the limit.

**Keywords** Decoupled forward–backward stochastic differential equations · $G$-Brownian motion · Relaxed optimal control · Hamilton–Jacobi–Bellman equation

**Mathematics Subject Classification** 60H05 · 60H20 · 60J75 · 93E20 · 91G80 · 91B70

## 1 Introduction

In this paper we study systems of decoupled controlled forward-backward stochastic differential equations (FBSDEs) driven by $G$-Brownian motion. These FBSDEs appear in connection with optimal control problems for diffusion with uncertain drift and diffusivity, e.g. as representations of the associated dynamic programming equations [4]. We call these FBSDEs in the framework of $G$-Brownian motion and the associated nonlinear expectation space $G$-FBSDEs, so as to distinguish them from standard FBSDEs that are driven by a standard Brownian motion.

✉ Omar Kebiri
Omar.Kebiri@b-tu.de

1 Laboratory of Stochastic Models, Statistics and Applications, University of Saida Dr Moulay Tahar, Saida, Algeria

2 Institute of Mathematics, The Free University of Berlin, Berlin, Germany

3 Department of Stochastics and its Applications, University of Cottbus, Cottbus, Germany

4 Department of Mathematics, Faculty of Sciences, LaPS Laboratory, Badji-Mokhtar University, Annaba, Algeria

In the standard framework, (F)BSDEs were studied by many authors, starting from the seminal work on linear BSDEs by Bismut [3] in 1973, as equation for the adjoint process in the stochastic version of Pontryagin maximum principle. Nonlinear backward stochastic differential equations (BSDEs) and associated decoupled FBSDEs driven by standard Brownian motion have been introduced by Pardoux and Peng [17] in 1990. They proved that, in a Markovian framework, the solution of a BSDE describes the viscosity solution of the associated semi-linear PDE. We refer for more details to the paper by El-Karoui et al. [5] the references therein. The relation between viscosity solutions to a certain class of nonlinear Hamilton–Jacobi–Bellman (HJB) equation for optimal control and the solutions to BSDEs has been studied in [14], using that the associated cost function is described by an adapted solution of a BSDE. Early examples of applications of BSDEs to stochastic control, especially in finance, are found in [5, 12, 13], to mention just a few examples. More recent examples of fully-coupled FBSDEs with non-smooth coefficients and their applications to stochastic control problems include existence results for optimal controls that have been established in [1, 11].

In the last years, aspects of model ambiguity, such as volatility uncertainty, have been studied by Peng [15, 16] who introduced the $G$-expectation space as the canonical nonlinear expectation space associated with the $G$-Brownian motion. Denis and Martini [4] suggested a structure based on quasi-sure analysis from abstract potential theory to construct a similar structure using a tight family $P$ of possibly mutually singular probability measures and established a corresponding Itô type stochastic calculus for the $G$-framework. Existence and uniqueness of solutions of $G$-SDEs were studied by Peng [16] and, later on, by Bai and Lin [2] under slightly weaker integral-Lipschitz conditions on the SDE coefficients.

BSDEs driven by a $G$-Brownian motion ($G$-BSDEs) were studied by Hu and al. [10] who showed that, unlike classical BSDEs that do not admit a direct probabilistic interpretation of fully nonlinear PDEs, $G$-BSDEs provide the missing link for fully nonlinear PDEs that appear for example, in mathematical finance in connection with the pricing of path-dependent contingent claims in uncertain volatility model; see also [7] for a related existence and uniqueness result for $G$-BSDEs.

### 1.1 Own Contribution and Related Work

In this paper, we consider decoupled controlled $G$-FBSDEs of the form

$$dX_s^u = b(s, X_s^u, u_s)ds + \sigma(s, X_s^u)dW_s + h(s, X_s^u, u_s)d\langle W\rangle_s \tag{1a}$$

$$dY_s^u = -f(s, X_s^u, Y_s^u, Z_s^u, u_s)ds - g(s, X_s^u, Y_s^u, Z_s^u, u_s)d\langle W\rangle_s + Z_s^u dW_s + dM_s^u\,, \tag{1b}$$

where $s \in [t, T]$ for some $T \in (0, \infty)$, and $h(s, X_s^u, u_s)d\langle W\rangle_s = \sum_{i,j} h_{ij} d\langle W_i, W_j\rangle_s$ and likewise for $g = (g_{ij})$. Equation (1) is endowed with the boundary data

$$X_t^u = x, \quad Y_T^u = \Phi(X_T^u), \quad M_t^u = 0. \tag{2}$$

Precise assumptions on the coefficients are given below. Our cost function if the solution of the BSDE at the initial time $t$. We call the system *decoupled*, because the controlled forward SDE (1a) is independent of the solution to the BSDE (1b). Existence and uniqueness of *coupled* $G$-FBSDEs has been studied by Wang and Yuan [20] under some monotonicity assumptions of the coefficients conditions.

Our objective is to establish the existence of a relaxed optimal control allowing the coefficients to be non-smooth. The key idea to establish existence of a relaxed optimal control is to replace the coefficients in the original control problem by a suitably regularised problem with mollified coefficients, prove the existence of a relaxed control and then pass to the limit. Existence of solutions for *uncontrolled* $G$-BSDEs with discontinuous drift coefficient was proved by [21], in our case the jump coefficient is zero.

Relaxed controls for stochastic differential equations driven by a $G$-Brownian motion ($G$-SDEs) were studied by Redjil and Choutri [18] who introduced the notion of $G$-relaxed controls and proved the so called $G$-Chattering Lemma to establish existence of an relaxed optimal control.

In the present work, one of the key steps in proving existence of a relaxed optimal control is to prove that the gradient of the solution of the approximate HJB equation is bounded, which allows us to define the second component of the approximate solution of the BSDE as a control in a bounded set before passing to the limit.

The paper is organized as follows: in the remainder of this section, we will record the main technical assumptions and definitions. In Sect. 2, we define the mollified control problem and the corresponding dynamic programming (HJB) equation, and prove various stability results, including convergence of the value function in the limit of vanishing mollification parameter. The existence of the relaxed control is shown in Sect. 3 by proving convergence of the relaxed controls. Conclusions are given in Sect. 4. The Appendix contains various technical Lemmas for Sects. 2 and 3.

### 1.2 Preliminaries

Let $U$ be any compact subset of $\mathbb{R}^l$, and call $\mathcal{U}(t)$ the set of adapted controls $u\colon [t, T] \to U$ for any $0 \le t < T$.

**Assumption 0** The coefficients of our control problem have the form

$$\begin{aligned}
&b\colon [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n, && h_{ij} : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n,\\
&f : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}, && g_{ij} : [0, T] \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \times U \to \mathbb{R},\\
&\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times n};\\
&\Phi : \mathbb{R}^n \to \mathbb{R}.
\end{aligned}$$

In what follows, we suppose that the diffusion coefficient $\sigma$ is non-degenerate, with uniformly bounded inverse. The analysis in this paper is based on the sublinear $G$-expectation space framework of [16] (called "$G-$framework" in what follows). The corresponding canonical process on the $G$-expectation space is called "$G$-Brownian motion" and is characterized as follows:

**Definition 1** A d-dimensional process $(W_t)_{t\geq 0}$ on a sublinear expectation space $(\Omega, \mathcal{H}, \hat{\mathbb{E}})$ is called a *G-Brownian motion* if it has the following properties:

(i) $W_0(\omega) = 0$,
(ii) For every $t, s \geq 0$, the increment $W_{t+s} - W_t$ is $N(\{0\}, s\Sigma)$[1] is distributed and is independent of $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$, for any $0 \leq t_1 \leq \ldots \leq t_n \leq t, n \in \mathbb{N}$,

where $\Sigma$ is any positive semidefinite symmetric $d \times d$ matrix.

We define

$M_G^p([0, T]; \mathbb{R}^n)$, $p \in [1, \infty)$ to be the completion of the space of simple processes of the form

$$\xi_t(\omega) = \sum_{k=1}^{N} \eta_{k-1}(\omega)\mathbf{1}_{[t_{k-1},t_k)}(t)$$

with $0 = t_0 < t_1 < \ldots < t_N = T$ being a partition of $[0, T]$ and $\eta_k \in L_G^p(\Omega_{t_k})$, where the completion is with respect to the norm

$$\|\xi\|_{G,p} := \left(\hat{\mathbb{E}}\left(\int_0^T |\xi_t|^p\, dt\right)\right)^{1/p}$$

$H_G^p(0, T)$ to be the completion of $M_G^0(0, T)$ under the norm $\|\eta\|_{\mathbb{H}^p} = \left\{\hat{\mathbb{E}}\left[\left(\int_0^T |\eta_s|^2 ds\right)^{\frac{p}{2}}\right]\right\}^{\frac{1}{p}}$;
$S_G^0(0, T) := \left\{h(W_{t_1\wedge t}, \ldots, W_{t_n\wedge t}) : t_1, \ldots, t_n \in [0, T], h \in C_{b.lip}(\mathbb{R}^{n+1})\right\}$;
$S_G^p(0, T)$ is the completion of $S_G^0(0, T)$ under the norm $\|\eta\|_{S_G^p} = \left\{\hat{\mathbb{E}}\left(\sup_{s\in[0,T]} |\eta_s|^p\right)\right\}^{\frac{1}{p}}$;
$\mathbb{L}_G^p(\Omega_T)$ is the space of decreasing $G$-martingales with $K_0 = 0$ and $K_T \in L_G^p(\Omega_T)$;

We also need the following BDG-type inequalities in the $G$-framework:

[1] $N(\{0\}, s\Sigma)$ is the $G$-normal distribution that is defined in terms of the associated $G$-heat equation; see [16].

**Proposition 2** (from [16]) *Let $\beta \in M_G^p(0,T)$ with $p \geq 2$. Then we have $\int_0^T \beta_t dB_t \in L_G^p(\Omega_T)$ and*

$$\hat{\mathbb{E}}\left(\left|\int_0^T \beta_t dW_t\right|^p\right) \leq C_p \hat{\mathbb{E}}\left(\left|\int_0^T \beta_t^2 d\langle W\rangle_t\right|^{\frac{p}{2}}\right). \tag{3}$$

**Proposition 3** (from [7]) *For each $\eta \in H_G^\alpha(0,T)$ with $\alpha \geq 1$ and $p \in (0,\alpha]$, it holds*

$$\underline{l}^p c_p \hat{\mathbb{E}}\left(\left|\int_0^T \eta_s^2 ds\right|^{\frac{p}{2}}\right) \leq \hat{\mathbb{E}}\left(\sup_{t\in[0,T]}\left|\int_0^t \eta_s dB_s\right|^p\right) \leq \bar{l}^p C_p \hat{\mathbb{E}}\left(\left|\int_0^T \eta_s^2 ds\right|^{\frac{p}{2}}\right),$$

*where $0 < c_p < C_p < \infty$ are constants, and $\underline{l}$ and $\bar{l}$ are upper and lower bounds for the component-wise variance of $W_1$.*

**Assumption 1** (a) For every fixed $(x,\upsilon) \in \mathbb{R}^n \times U$, $b(\cdot,x,\upsilon)$, $h_{ij}(\cdot,x,\upsilon)$, and $\sigma_j(\cdot,x)$ are continuous in $t$, where $\sigma_j$ denotes the $j$-th column of the matrix $\sigma$;
(b) $b, h_{ij}, \sigma_j$ are given functions satisfying $b(\cdot,x,\upsilon), h_{ij}(\cdot,x,\upsilon), \sigma_j(\cdot,x) \in L_G^2([0,T],\mathbb{R}^n)$;
(c) There exists a constant $L > 0$, such that for every $t \in [0,T]$, $x,\ x' \in \mathbb{R}^n$, and every $u \in U$,

$$\left|\phi(t,x,u) - \phi(t,x',u)\right| \leq L\left|x - x'\right|,$$

where $\phi$ is a placeholder for $b, h_{ij}$. We further suppose that $\sigma_j$ is Lipschitz in $x$.

**Assumption 2** (a) There exists some $\beta > 2$ such that for any $y, z, u$, we have

$$f(\cdot,\cdot,y,z,u), g_{ij}(\cdot,\cdot,y,z,u) \in L_G^\beta([0,T],\mathbb{R}^n).$$

(b) There exists $L > 0$, such that for every $t \in [0,T]$, $u \in U$, $y,\ y' \in \mathbb{R}$, and $z,\ z' \in \mathbb{R}^d$, it holds

$$\begin{aligned}|f(t,x,y,z,u) - f(t,x,y',z',u)| + \sum_{i,j=1}^{d} |g_{ij}(t,x,y,z,u) - g_{ij}(t,x,y',z',u)| \\ \leq L(|y-y'| + |z-z'|).\end{aligned}$$

(c) $f(\cdot,x,y,z,u), g_{ij}(\cdot,x,y,z,u)$ are continuous in $t \in [0,T]$, for every fixed $(x,y,z,u)$
(d) There exist a constant $L > 0$, such that

$$|\Phi(x) - \Phi(x')| \leq L(|x-x'|).$$

**Assumption 3** The functions $b, \sigma, f, h = (h_{ij}), g = (g_{ij})$, and $\Phi$ are bounded.
For every fixed initial time $t \in [0, T]$, and initial state $x \in \mathbb{R}^n$, and for every fixed control $u \in \mathcal{U}(t)$ under Assumption 1 (see [16]), the forward stochastic differential equation in (1) has a unique solution $X^{t,x,u}$. Moreover, under Assumption 3 (see [7]) the backward stochastic differential equation of (1) has a unique solution denoted by $(Y, Z, M) = (Y^{t,x,u}, Z^{t,x,u}, M^{t,x,u})$. Note that, the first component, $Y$, of the solution of the backward stochastic differential equation in (1) is deterministic in that $Y_t^{t,x,u}$ is a deterministic function of $(t, x)$; see [10].

**Definition** The cost functional which will be minimized, is defined for $u \in \mathcal{U}(t)$ by:

$$J(t, x; u) := Y_t^{t,x,u}.$$

A control $\widehat{u}$ is called optimal if it minimizes $J$ that is:

$$Y_t^{t,x,\widehat{u}} = \operatorname*{essinf}_{u \in \mathcal{U}(t)} Y_t^{t,x,u}.$$

If $\widehat{u} \in \mathcal{U}(t)$, we say that $\widehat{u}$ is an optimal strict control. We define the value function by:

$$V(t, x) = \operatorname*{essinf}_{u \in \mathcal{U}(t)} J(t, x, u).$$

$V(t, x)$ is the unique viscosity solution of the $G$-HJB equation
In what follows, we will suppress the dependence of $V$ on the variables $t$ and $x$ and simply write $V = V(t, x)$

$$\begin{aligned} &\partial_t V + \inf_{u \in \mathcal{U}(t)} H(\nabla^2 V, \nabla V, V, x, t, u) = 0, \\ &V(T, x) = \Phi(x), \end{aligned} \tag{4}$$

where

$$\begin{aligned} H(\nabla^2 V, \nabla V, V, x, t, u) =& G(F(\nabla^2 V, \nabla V, V, x, t, u)) + \langle b(t, x, u), \nabla V \rangle \\ &+ f(t, x, V, \sigma^T \nabla V, u), \end{aligned}$$

with $G$ denoting the nonlinear infinitesimal generator of the $G$-Brownian motion, and

$$\begin{aligned} F_{ij}(\nabla^2 V, \nabla V, V, x, t, u) =& \langle \nabla^2 V \sigma_i(t, x), \sigma_j(t, x) \rangle + 2\langle \nabla V, h_{ij}(t, x, u) \rangle \\ &+ 2g_{ij}(t, x, V, \sigma^T \nabla V, u), \end{aligned}$$

where $(\sigma^T \nabla V)(t, x) = (\langle \sigma_1, \nabla V \rangle, \ldots, \langle \sigma_d, \nabla V \rangle)^T(t, x) = Z_t^{\delta}$.

We give now the definition of a relaxed stochastic control:

**Definition 4** (*Relaxed stochastic control*) A relaxed stochastic control (or simply a relaxed control) on $(\Omega_T, \mathrm{Lip}(\Omega_T), \widehat{\mathbb{E}})$ is a random measure $q(\omega, dt, da) = \mu_t(\omega, da)dt$ such that for each subset $A \in \mathcal{B}(U)$, the process $(\mu_t(A))_{t\in[0,T]}$ is $\mathbb{F}^P$-progressively measurable i.e. for every $t \in [0, T]$, the mapping $[0, t] \times \Omega \to [0, 1]$ defined by $(s, \omega) \mapsto \mu_s(\omega, A)$ is $\mathcal{B}([0, t]) \otimes \widehat{\mathcal{F}}_t^P$-measurable. In particular, the process $(\mu_t(A))_{t\in[0,T]}$ is adapted to the *universal* filtration $\mathbb{F}^P$. We denote by $R$ the class of relaxed stochastic controls.

Our coefficients are not smooth enough to ensure the existence of a strong solution of the HJB equation (4). Therefore, the next step will be to define an approximate HJB equation.

## 2 The Mollified Hamilton–Jacobi–Bellman Equation

In this section we aim to determine explicitly an optimal feedback control process from a sequence of stochastic control problems which value functions converge to the value function of our original problem. The coefficients of the original control problem are not smooth enough to get a smooth solution, therefore we replace them by mollified coefficients.

To this end, we define the mollification of a given function as follows: for any integer $m \geq 1$, let $\varphi : \mathbb{R}^m \to \mathbb{R}$ have the following properties:

- $\varphi \geq 0$ is non-negative and smooth function.
- The support of $\varphi$ is contained in the unit ball in $\mathbb{R}^m$, i.e. $\mathrm{supp}(\varphi) \subset B_0(1)$ where $B_0(1)$ denotes the unit ball in $\mathbb{R}^m$.
- $\varphi$ is normalised, i.e. $\int_{\mathbb{R}^m} \varphi(\xi)\, d\xi = 1$.

The function $\varphi$ is called a "mollifier". Then, for any Lipschitz function $l : \mathbb{R}^m \to \mathbb{R}$, we define its mollification by

$$l_\delta(\xi) := \delta^{-m} \int_{\mathbb{R}^m} l\left(\xi - \xi'\right) \varphi\left(\delta^{-1}\xi'\right) d\xi', \quad \xi \in \mathbb{R}^m,\ \delta > 0.$$

**Properties 5** *The mollification $l_\delta$ of a function $l$ enjoys the following properties:*

*(1)* $|l_\delta(\xi) - l(\xi)| \leq C_l \delta$
*(2)* $|l_\delta(\xi) - l_{\delta'}(\xi)| \leq C_l |\delta - \delta'|$
*(3)* $\left|l_\delta(\xi) - l_\delta\left(\xi'\right)\right| \leq C_l |\xi - \xi'|$ *for all* $\xi, \xi' \in \mathbb{R}^m$ *and* $\delta, \delta' > 0$ *where* $C_l$ *denotes the Lipschitz constant of* $l$ *that is independently of* $\delta$.

***Proof*** We prove only the last statement. It holds

$$\begin{aligned}|l_\delta(\xi) - l_\delta(\xi')| &= \delta^{-m} \int_{\mathbb{R}^m} |l(\xi - x)\varphi(\delta^{-1}x) - l(\xi' - x)\varphi(\delta^{-1}x)| dx \\ &\leq \delta^{-m} \int_{\mathbb{R}^m} \varphi(\delta^{-1}x) C_l |\xi - \xi'| dx\end{aligned}$$

$$\leq \delta^{-m} C_l |\xi - \xi^{'}| \int_{\mathbb{R}^m} \varphi(\delta^{-1}x)dx\,.$$

Substituting $M = \delta^{-1}x$, it follows that

$$|l_\delta(\xi) - l_\delta(\xi')| \leq C_l |\xi - \xi^{'}| \int_{\mathbb{R}^m} \varphi(M)dM \leq C_l|\xi - \xi'|.$$

Then $\left|l_\delta\left(\xi\right) - l_\delta\left(\xi'\right)\right| \leq C_l|\xi - \xi'|$. □

**Definition 6** For each $\delta \in (0, 1]$ we denote by $b_\delta, \sigma_\delta, f_\delta$ and $\Phi_\delta$ the mollification of the functions $b, \sigma, f$ and $\Phi$, respectively, introduced in Sect. 1, with $l = b\left(\cdot, v\right)$, $\sigma\left(\cdot, v\right)$, $f\left(\cdot, v\right)$ and $\Phi\left(.\right)$.

Now, let Assumptions 1–3 hold, and let $\delta \in (0, 1]$ be an arbitrary fixed number. We define the function $F^\delta = (F^\delta_{ij})$ by:

$$\begin{aligned} F^\delta_{ij}(t, x, V, \nabla V, \nabla^2 V, v) = &\langle \nabla^2 V \sigma^\delta_i(t, x), \sigma^\delta_j(t, x)\rangle \\ &+2\langle \nabla V, h^\delta_{ij}(t, x, V, v)\rangle + 2g^\delta_{ij}(t, x, V, (\sigma^\delta)^T \nabla V). \end{aligned}$$

The mollification $H^\delta(t, x, V, \nabla V, \nabla^2 V, u)$ of the Hamiltonian is defined as

$$\begin{aligned} H^\delta(t, x, V, \nabla V, \nabla^2 V, v) = &G^\delta(F^\delta(t, x, V, \nabla V, \nabla^2 V, v)) + \langle b_\delta, \nabla V\rangle \\ &+ f_\delta(t, x, V, \nabla V, \nabla^2 V, v). \end{aligned}$$

Now, since $H^\delta$ is a smooth function and $G^\delta$ is uniformly elliptic, the corresponding $G$-HJB equation

$$\begin{aligned} &\partial_t V^\delta + \inf_{v \in U} H^\delta(t, x, V^\delta, \nabla V^\delta, \nabla^2 V^\delta, v) = 0 \\ &V^\delta(T, x) = \Phi_\delta(x),\ \ x \in \mathbb{R}^n. \end{aligned} \tag{5}$$

has a unique bounded and continuous viscosity solution $V^\delta$. It moreover follows from the regularity result of Krylov [8], that (5) even has a classical $C^{1+\frac{l}{2}, 2+l}([0, T] \times \mathbb{R}^n)$ solution. The regularity of $V^\delta$ and the compactness of the control set $U$ allows us to find a measurable function $v^\delta : [0, T] \times \mathbb{R}^n \mapsto U$ such that for all $(t, x) \in (0, T] \times \mathbb{R}^n$,

$$H^\delta(x, (V, \nabla V, \nabla^2 V)(t, x), v^\delta) = \inf_{v \in U} H^\delta(x, (V^\delta, \nabla V^\delta, \nabla^2 V^\delta)(t, x), v). \tag{6}$$

**Lemma 7** *Assume that the Assumptions 1–3 are satisfied, then*

$$J^\delta(t, x; u^\delta) = V^\delta(t, x) = \operatorname*{ess\,inf}_{u \in \mathcal{U}(t)} J^\delta(u).$$

*Moreover* $u^\delta_s := v^\delta_s(s, X^\delta_s)$, $s \in [0, T]$ *is an admissible control.*

***Proof*** Let $(t, x) \in [0, T] \times \mathbb{R}^n$ a fixed arbitrary initial datum and $\delta \in (0, 1]$. Further let $V^\delta$ be the solution of (5) and $v^\delta$ the function defined by (6). We consider the following $G$-SDE

$$\begin{aligned} dX_s^\delta &= b_\delta(s, X_s^\delta, v^\delta(s, X_s^\delta))ds + \sigma_\delta(s, X_s^u)dW_s \\ &\quad + h_\delta(s, X_s^\delta, v^\delta(s, X_s^\delta))d\langle W\rangle_s, \; s \in [t, T] \\ X_t^\delta &= x. \end{aligned} \tag{7}$$

Since $b_\delta, h_\delta$ are bounded measurable functions in $(t; x)$ and $\sigma_\delta$ is Lipschitz in $x$, and as $b$ and $h$ are in $L^2$, then, due to a result in [6], see also [21], there exists a unique solution $X_t^\delta \in M_G^2([0, T]; \mathbb{R}^n)$. We define

$$Y_s^\delta = V^\delta(s, X_s^\delta), \quad \text{and} \quad Z_s^\delta = \nabla V^\delta(s, X_s^\delta)\sigma_\delta(s, X_s^\delta). \tag{8}$$

and apply the $G$-Itô formula to $V^\delta(s, X_s^\delta)$:

$$\begin{aligned} V^\delta(s, X_T^\delta) - V^\delta(s, X_t^\delta) = &\int_t^T \nabla V^\delta(s, X_s^\delta)\sigma_\delta(s, X_s^\delta)dW_s \\ &+ \int_t^T \nabla V^\delta(s, X_s^\delta b_\delta(s, X_s^\delta, v^\delta(s, X_s^\delta))ds \\ &+ \int_t^T [\nabla V^\delta(s, X_s^\delta)h_\delta(s, X_s^\delta v^\delta(s, X_s^\delta)), v^\delta((s, X_s^\delta)) \\ &+ \int_t^T \frac{1}{2}\nabla^2 V^\delta(s, X_s^\delta)\sigma_\delta(s, X_s^\delta)]d\langle W\rangle_s. \end{aligned}$$

Taking the partial derivative with respect to $s$ and time $s = t$ yields together with the $G$-HJB equation (5) the following system of coupled equations for $(X_s^\delta, Y_s^\delta, Z_s^\delta, M_s^\delta)$:

$$dX_s^{\delta,u} = b_\delta(s, X_s^{\delta,u}, u_s^\delta)ds + \sigma_\delta(s, X_s^{\delta,u})dW_s + h_\delta(s, X_s^{\delta,u}, u_s^\delta)d\langle W\rangle_s, \tag{9a}$$

$$dY_s^{\delta,u} = -f_\delta(s, X_s^{\delta,u}, Y_s^{\delta,u}, Z_s^{\delta,u}, v^\delta(s, X_s^\delta))ds - g_\delta(s, X_s^{\delta,u}, Y_s^{\delta,u}, Z_s^{\delta,u})d\langle W\rangle_s + Z_s^{\delta,u}dW_s + dM_s^u \tag{9b}$$

$$X_t^{\delta,u} = x, \quad Y_T^{\delta,u} = \Phi_\delta(X_T^{\delta,u}), \quad M_t^u = 0. \tag{9c}$$

By [7], the backward equation (9b) has a unique solution $(Y^\delta, Z^\delta, M^\delta) \in S_G^2(0, T) \times H_G^2(0, T)$, and we can identify the solution $(X^\delta, Y^\delta, Z^\delta)$ of (9) with the solution $(X^{\delta,t,x,u^\delta}, Y^{\delta,t,x,u^\delta}, Z^{\delta,t,x,u^\delta}, M^u)$ of (7)–(8). In particular, $Y_t^{\delta,t,x,u^\delta} = V^\delta(t, x)$.

To show that $u^\delta$ is optimal, consider any $\delta' \in (0, 1]$ different from $\delta$. Then dropping the dependence on the initial conditions, and denoting $X^{\delta',u^\delta} := X^{\delta',t,x,u^\delta}$ the solution of the forward stochastic differential equation

$$\begin{aligned} dX_s^{\delta',u^\delta} &= b_{\delta'}(s, X_s^{\delta',u^\delta}, u^\delta)ds + \sigma_{\delta'}(s, X_s^{\delta',u^\delta})dW_s \\ &\quad + h_{\delta'}(s, X_s^{\delta',u^\delta}, u^\delta)d\langle W\rangle_s, \; s \in [t, T] \end{aligned}$$

$$X_t^{\delta',u^\delta} = x.$$

Here $u^\delta$ is a shorthand for $u_s^\delta := v_s^\delta(s, X_s^{\delta',u^\delta})$ for any $s \in [t, T]$. Applying $G$-Itô's formula to

$$Y^{\delta',u^\delta} := V^{\delta'}(s, X_s^{\delta',u^\delta}) \quad \text{and} \quad Z^{\delta',u^\delta} := \sigma_{\delta'}(s, X_s^{\delta',u^\delta})\nabla V^{\delta'}(s, X_s^{\delta',u^\delta})$$

and introducing the shorthand

$$\tilde{f}_{\delta',u} := G^{\delta'}(F^{\delta'}(t, x, V^{\delta'}, \nabla V^{\delta'}, \nabla^2 V^{\delta'}, u)) + \langle b_{\delta'}(t, x, V^{\delta'}, u), \nabla V^{\delta'}\rangle,$$

it follows that the process $(Y^{\delta',u^\delta}, Z^{\delta',u^\delta}, M')$ is the unique solution of the system of equations

$$\begin{aligned} dY_s^{\delta',u^\delta} &= -\tilde{f}_{\delta',u^\delta} ds - g_\delta(s, X_s^{\delta',u^\delta}, Y_s^{\delta',u^\delta}, Z_s^{\delta',u^\delta}) d\langle W\rangle_s + Z_s^{\delta',u^\delta} dW_s + dM'_s \\ Y_T^{\delta',u^\delta} &= \Phi_{\delta'}(X_T^{\delta',u^\delta}), \end{aligned}$$

where $M'$ is a $G$-decreasing martingale [6]. The mollified $G$-HJB equation with classical solution $V^{\delta'}$ implies $\tilde{f}_{\delta',u^\delta} \le \tilde{f}_{\delta',u^{\delta'}}$. Then, using the comparison theorem for BSDEs, [10, Thm 3.6, p. 1183], and observing that $\tilde{f}_{\delta',u^{\delta'}} = f_{\delta'}$, we can conclude that

$$Y_t^{\delta',u^{\delta'}} \le Y_t^{\delta',u^\delta},$$

where the left hand side is the solution of (9b) for $\delta = \delta'$. Since $\delta'$ was arbitrary, the last inequality shows that $u^\delta = v^\delta(s, X_s^{\delta,u^\delta}) =: v^\delta(s, X_s^\delta)$ is the optimal control. □

**Lemma 8** *Suppose that Assumptions 1–3 hold. Then there exists a non-negative constant $\bar{C}$, only depending on the Lipschitz constants of the coefficients and the terminal time $T$, such that:*

$$|V^{\delta'}(t, x) - V^\delta(t, x)|^2 \le \bar{C}|\delta' - \delta|^2, \quad t \in [0, T],\ x \in \mathbb{R}^n. \tag{10}$$

For the proof of this Lemma, we refer to Appendix A.1.

**Lemma 9** *As $f_\delta, b_\delta, \Phi_\delta, \sigma_\delta, h_\delta$ and $g_\delta$ are bounded $\mathbb{C}^\infty$ functions, all their derivatives are bounded. Then*

$$\begin{aligned} &\partial_t V^\delta(t, x) + H^\delta(x, (V^\delta, \nabla V^\delta, \nabla^2 V^\delta)(t, x), v^\delta(t, x)) = 0 \quad (t, x) \in [0, T] \times \mathbb{R}^n \\ &V^\delta(T, x) = \Phi(x),\ x \in \mathbb{R}^n. \end{aligned} \tag{11}$$

*admits a unique solution $V^\delta \in C_b^{1,2}([0, T] \times \mathbb{R}^n)$, where*

$$\nabla V^\delta \quad \text{and} \quad \nabla^2 V^\delta \quad \text{are uniformly bounded on} \quad [0, T] \times \mathbb{R}^n. \tag{12}$$

*Moreover, there exists a constant $\bar{C}$ only depending on $T$ and constants $\bar{\Gamma}$, $\bar{\kappa}$ only depending on T and the Lipschitz constant L of the FBSDE coefficients (cf. Assumption 2), such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} |V^{\delta}(t,x)| \le \bar{C} \tag{13}$$

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^n} |\nabla V^{\delta}(t,x)| \le \bar{\Gamma} \tag{14}$$

$$\forall t, t' \in [0,T], \quad |V^{\delta}(t',x) - V^{\delta}(t,x)| \le \bar{\kappa}|t'-t|^{\frac{1}{2}}. \tag{15}$$

Similar result of the estimations given in the above lemma were already proved in [9] even when the diffusion is depend on the control. Here our diffusion is independent on the control because then when it is the case we need to prove existence of the solution of G-SDE with measurable diffusion which we leave for a future work.

For the proof of the lemma, see Appendix A.2.

We briefly discuss the implications of Lemmas 7–9: To this end, note that

$$|V^{\delta'}(t',x') - V^{\delta}(t,x)| \le |V^{\delta'}(t',x') - V^{\delta}(t',x')| + |V^{\delta}(t',x') - V^{\delta}(s,x)|,$$

as a consequence of which, (14) and (15) imply that

$$|V^{\delta'}(t',x') - V^{\delta}(t,x)| \le \bar{\kappa}|t-t'|^{\frac{1}{2}} + \bar{\Gamma}|x-x'|,$$

Using (10) and adapting the constants whenever necessary, we obtain

$$|V^{\delta'}(t',x') - V^{\delta}(t,x)| \le C\left(|t-t'|^{\frac{1}{2}} + |x-x'| + |\delta-\delta'|\right)$$

for some generic constant $C > 0$ depending on $L$ and $T$. As $V^{\delta}$ is bounded in $(t,x)$, we conclude that it converges to a function $\bar{V}$ as $\delta \to 0$, moreover the Hamiltonian $H^{\delta}$ converges to $H$ because of the stability of the viscosity solution, in fact $\bar{V}$ is also solution of (5). The uniqueness of the solution of equation (5) entails that $\bar{V} = V$. This prove that

$$V^{\delta} \to V \quad \text{as} \quad \delta \to 0.$$

More precisely,

$$|V^{\delta}(t,x) - V(t,x)| \le C\delta, \quad \text{for all} \quad \delta \in [0,1) \quad \text{and} \quad (t,x) \in [0,T]\times\mathbb{R}^n,$$

for a constant $C$ that depends on $x$ and $t$.

## 3 Convergence of the Optimal Control

In this section we show that the solution to the mollified problem (5) converges to the value function of our original problem, and show that the optimal control of our original stochastic optimal control is the limit of the sequence of the optimal control of the mollified systems. The result is stated in the next theorem.

**Theorem 10** *Assume that the Assumptions 1–3 are satisfied. Let $(t, x) \in [0, T] \times \mathbb{R}^n$ and $(\delta_n)_{n\in\mathbb{N}}$ be a sequence of positive real numbers which tends to 0. Then, there exists a process $(\bar{X}, \bar{Y}, \bar{Z}, \bar{M}) \in M_G^2([0, T]) \times \mathfrak{S}_G^2(0, T)$, with $\bar{M}$ is a decreasing martingale and an admissible control $\bar{u} \in \mathcal{U}(t)$, such that:*

1. *There is a subsequence of $(X^{\delta_n}, Y^{\delta_n})_{n\in\mathbb{N}}$ which converges in distribution to $(\bar{X}, \bar{Y})$,*
2. *$(\bar{X}, \bar{Y}, \bar{Z}, \bar{M})$ is a solution of the FBSDE system*

$$\begin{aligned} d\bar{X}_s &= b(s, \bar{X}_s, \bar{u}_s)ds + \sigma(s, \bar{X}_s)d\bar{W}_s + h(s, \bar{X}_s, u_s)d\langle W\rangle_s, \\ d\bar{Y}_s &= -f(s, \bar{X}_s, \bar{Y}_s, \bar{Z}_s, \bar{u}_s)ds - g(s, \bar{X}_s, \bar{Y}_s, \bar{Z}_s)d\langle W\rangle_s + \bar{Z}_s dW_s + d\bar{M}_s, \\ \bar{X}_t &= x \quad \bar{Y}_T = \xi = \Phi(\bar{X}_T), \quad \bar{M}_t = 0. \end{aligned} \tag{16}$$

3. *For every $(t, x) \in [0, T] \times \mathbb{R}^n$, it holds that*

$$\bar{Y}_t = V(t, x) = \operatorname*{ess\,inf}_{u\in\mathcal{R}(t)} J(t, x; u) \,,$$

*i.e. there is a relaxed control, such that $\bar{u} \in \mathcal{R}$ is optimal for (16) and hence for our original SOC problem.*

To prove the Theorem, we need the following Lemma:

**Lemma 11** *For all $n \in \mathbb{N}$. There exists a constant $L$ such that*

$$\hat{\mathbb{E}}\left(\sup_{s\in[t,T]} |X_s^{\delta_n} - X_s^n|^2\right) \leq L\delta_n^2 \tag{17}$$

$$\hat{\mathbb{E}}\left(\sup_{s\in[t,T]} |Y_s^{\delta_n} - Y_s^n|^2\right) \leq L\delta_n^2. \tag{18}$$

For the proof see Appendix B.

***Proof of Theorem 10*** We will prove that the limit of the sequence $(X^{\delta_n}, Y^{\delta_n})$ coincides with the limit of a subsequence of an auxiliary sequence of forward SDEs whose solutions converge in law to $(\bar{X}, \bar{Y})$. To this end, we define the sequence of auxiliary processes $(X_s^n, Y_s^n)$ as the unique solution of the following controlled forward system:

$$\begin{aligned} dX_s^n &= b(s, X_s^n, u_s^{\delta_n})ds + \sigma(s, X_s^n)dW_s + h(s, X_s^n, u_s^{\delta_n})d\langle W\rangle_s, \\ dY_s^n &= -f(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^n)ds - g(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n})d\langle W\rangle_s \\ &\qquad + \sigma^T w_s^{\delta_n} dW_s + \theta_s^n d\langle B\rangle_s - 2G(\theta_s^n)ds, \\ X_t^n &= x, \qquad Y_t^n = V^{\delta_n}(t, x), \end{aligned} \tag{19}$$

where $u_s^{\delta_n} := v^{\delta_n}(s, X_s^{\delta_n})$ and $w_s^{\delta_n} = \nabla V^{\delta_n}(s, X_s^{\delta_n})$. We moreover use the shorthand $\sigma^T w_s^{\delta_n} = (\sigma(s, X_s^n))^T w_s^{\delta_n}$ here and $\sigma_{\delta_n}^T w_s^{\delta_n} = (\sigma_{\delta_n}(s, X_s^{\delta_n}))^T w_s^{\delta_n}$ below for the mollified $Z$-process. Here, $(\theta^n)_{n\in\mathbb{N}}$ is some sequence of stochastic processes $\theta^n = (\theta^n_s)_{s\in[0,T]}$ with $\theta^n_0 = 0$ for all $n \in \mathbb{N}$. We further define the process $(X^{\delta_n}_s, Y^{\delta_n}_s)$ as the solution of the following controlled forward SDE

$$
\begin{aligned}
dX^{\delta_n}_s &= b_{\delta_n}(s, X^{\delta_n}_s, u^{\delta_n}_s)ds + \sigma_{\delta_n}(s, X^{\delta_n}_s)dW_s + h_{\delta_n}(s, X^{\delta_n}_s, u^{\delta_n}_s)d\langle W\rangle_s,\\
dY^{\delta_n}_s &= -f_{\delta_n}(s, X^{\delta_n}_s, Y^{\delta_n}_s, \sigma^T_{\delta_n} w^{\delta_n}_s, u^{\delta_n}_s)ds - g(s, X^{\delta_n}_s, Y^{\delta_n}_s, \sigma^T_{\delta_n} w^{\delta_n}_s)d\langle W\rangle_s\\
&\quad +\sigma^T_{\delta_n} w^{\delta_n}_s\, dW_s,\\
X^{\delta_n}_t &= x \quad Y^{\delta_n}_t = V^{\delta_n}(t, x)
\end{aligned}
$$

From (8) we have, for $t \le s \le T$, that $Y^{\delta_n}_s = V^{\delta_n}(s, X^{\delta_n})$. Since $(s, x) \mapsto V^{\delta_n}(s, x)$ is a $C^{1,2}$ function that satisfies equation (11) with $\delta = \delta_n$, it follows from $G$-Itô's formula that

$$
\begin{aligned}
Y^{\delta_n}_t = \Phi_{\delta_n}(X^{\delta_n}_T) &+ \int_t^T f_{\delta_n}(s, X^{\delta_n}_s, Y^{\delta_n}_s, \sigma^T_{\delta_n} w^{\delta_n}_s, u^{\delta_n}_s)ds\\
&+ \int_t^T g(s, X^{\delta_n}_s, Y^{\delta_n}_s, \sigma^T_{\delta_n} w^{\delta_n}_s)d\langle W\rangle_s - \int_t^T \sigma^T_{\delta_n} w^{\delta_n}_s\, dW_s.
\end{aligned}
$$

Setting

$$
\chi^n_s := \begin{pmatrix} X^n_s \\ Y^n_s \end{pmatrix}, \quad r^n_s := (\sigma^T w^{\delta_n}_s, 0, u^{\delta_n}_s) \quad \mathcal{W} := \begin{pmatrix} W \\ W \end{pmatrix}, \quad \text{and} \quad d\langle\mathcal{W}\rangle := \begin{pmatrix} d\langle W\rangle \\ d\langle W\rangle \end{pmatrix}
$$

then (19) can be recast as

$$
\begin{aligned}
d\chi^n_s &= \beta(\chi^n_s, r^n_s)ds + \Pi(\chi^n_s, r^n_s)d\langle\mathcal{W}\rangle_s + \Sigma(\chi^n_s, r^n_s)d\mathcal{W}_s, \quad s \in [t, T],\\
\chi^n_t &= \left(x^T,\ V^{\delta_n}(t, x)\right)^T,
\end{aligned}
$$

with

$$
\begin{aligned}
\beta(\chi^n_s, r^n_s) &= \begin{pmatrix} b(s, X^n_s, u^{\delta_n}_s) \\ -f(s, X^n_s, Y^n_s, \sigma^T w^{\delta_n}, u^{\delta_n}_s) - 2G(\theta_n) \end{pmatrix}\\
\Pi(\chi^n_s, r^n_s) &= \begin{pmatrix} h(s, X^n_s, u^{\delta_n}_s) \\ -g(s, X^n_s, Y^n_s, w^{\delta_n}\sigma(X^n_s)) + \theta_n \end{pmatrix}\\
\Sigma(\chi^n_s, r^n_s) &= \begin{pmatrix} \sigma(s, X^n_s) \\ w^{\delta_n}_s\sigma(X^n_s) \end{pmatrix}.
\end{aligned}
$$

From (9), we can conclude that $w^{\delta_n}_s = \nabla V^{\delta_n}(s, X^{\delta_n}_s)$ is uniformly bounded. As a consequence, we can regard $(r^n_s)_{s\in[t,T]}$ as a control with values in the compact set $A \subset \mathbb{R}^{d+1} \times U$.

The next step is to take the limit $n \to \infty$, for the purpose of which we consider the random measure

$$q^n(\omega, ds, da) = \delta_{r_s^n(\omega)}(da)ds, \ (s,a) \in [0,T] \times A, \ \omega \in \Omega.$$

We can identify the control process $r^n$ with the measure $q^n$, which amounts to saying that the controls $r^n$ are in the set of relaxed controls. Specifically, we consider $r^n$ as random variable with values in the space $V$ of all Borel measures $q^n$ on $[0, T] \times A$, whose projection $q^n(\cdot \times A)$ coincides with the Lebesgue measure on $A$, endowed with the usual $\sigma$-algebra of Borel sets, $\mathcal{B}(A)$.

From the boundedness of our coefficients and by the compactness of $V$ with respect to the topology induced by the weak convergence of measures, we get the tightness of the laws of $(\chi^n, q^n)$ on this space, and then, from this and the use of the $G$-Chattering Lemma [18] we can extract a subsequence that converges in law to $(\chi, \bar{r})$, with $\bar{r}$ having values in $\mathcal{R}$. The limit process satisfies

$$\begin{aligned} d\chi_s &= \beta(\chi, \bar{r}_s)ds + \Pi(\chi_s, \bar{r}_s)d\langle \mathcal{W}\rangle_s + \Sigma(\chi_s, \bar{r}_s)d\mathcal{W}_s, \quad s \in [t,T], \\ \chi_t &= \left(x^T, \ V^{\delta_n}(t,x)\right)^T, \end{aligned} \tag{20}$$

with the straightforward definitions of $\Sigma$, $\prod$ and $\beta$, and

$$\chi := \begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix}, \mathcal{W} := \begin{pmatrix} W \\ W \end{pmatrix}, \quad \text{and} \quad d\langle \mathcal{W}\rangle := \begin{pmatrix} d\langle W\rangle \\ d\langle W\rangle \end{pmatrix}.$$

Noting that $\bar{r} := (\bar{w}, \bar{\theta}, \bar{u})$, the system (20) can be again recast as

$$\begin{aligned} &d\bar{X}_s = b(s, \bar{X}_s, \bar{u}_s)ds + \sigma(s, \bar{X}_s)dW_s + h(s, \bar{X}_s, \bar{u}_s)d\langle W\rangle_s, \\ &d\bar{Y}_s = -f(s, \bar{X}_s, \bar{Y}_s, \bar{Z}_s, \bar{u}_s)ds - g(s, \bar{X}_s, \bar{Y}_s, \bar{Z}_s)d\langle W\rangle_s + \bar{Z}_s dW_s + d\bar{M}_s, \\ &\bar{X}_t = x \quad \bar{Y}_t = V(t,x) \end{aligned}$$

with

$$(\bar{M}_s)_{s\in[t,T]}, \quad \bar{M}_s := \int_t^s \bar{\theta}_r d\langle B\rangle_r - 2\int_t^s G(\bar{\theta}_r)dr$$

being a decreasing $G$-martingale. This proves the first part of the Theorem (Note that by tightness, we can choose an arbitrary subsequence).

The second statement follows from Lemma 11 that shows that, if the sequence $(X^n, Y^n)_{n\in\mathbb{N}}$ converges in law, the same holds true for $(X^{\delta_n}, Y^{\delta_n})_{n\in\mathbb{N}}$, and the limits have the same distribution. Further, we deduce from (17)–(18) and Proposition 9, that $\bar{Y}_s = V(s, \bar{X}_s)$ for each $s \in [t, T]$ quasi-surly (q.-s.). In particular, $Y_T = \Phi(X_T)$ q.-s. So, the second part of the assertion is proved.

To prove the last part, note that $\bar{Y}_s = V(s, \bar{X}_s)$ for all $s \in [t, T]$ q.-s.. On the other hand, it is well known that, for the unique bounded viscosity solution $V$ of the HJB equation (4), it holds

$$V(t,x) = \operatorname*{ess\,inf}_{u \in \mathcal{R}(t)} J(t,x;u), \quad \text{q.-s.}$$

Hence the Theorem is proved. □

## 4 Conclusions

In this paper, we have proved existence of a relaxed optimal control for a stochastic control problem that involves a system of decoupled forward-backward SDEs, with non-smooth coefficients and driven by a $G$-Brownian motion ($G$-FBSDEs). From a high level perspective, the proof is based on the $G$-Chattering Lemma that has been proved by one of the authors of this paper. The key idea is to replace all coefficients in the original $G$-FBSDE by mollified coefficients, for which existence of an optimal control can be proved, and then send the mollification parameter to zero. From a technical perspective, the proof relies on uniform gradient bounds for the corresponding value function that implies that the associated backward SDE is driven by a control that is bounded uniformly in time. The latter then allows for proving existence of a relaxed optimal control by extracting appropriate subsequences. We believe that the general approach of the $G$-Chattering Lemma can even be applied to optimal control problems for smooth $G$-FBSDEs, such as $G$-FBSDEs with coefficients that have multiscale features. We leave this investigation, however, for future work.

## Appendices

## A Approximation of the HJB Equation

### A.1 Proof of Lemma 8

***Proof*** We start by introducing some notation:

$$\begin{aligned} X &:= X^{\delta,t,x,u^{\delta}}, \qquad Y := Y^{\delta,t,x,u^{\delta}}, \qquad Z := Z^{\delta,t,x,u^{\delta}}, \\ X' &:= X^{\delta',t,x,u^{\delta'}}, \qquad Y' := Y^{\delta',t,x,u^{\delta'}}, \qquad Z' := Z^{\delta',t,x,u^{\delta'}} \end{aligned}$$

and

$$\begin{aligned} & f_\delta(s) := f_\delta(s, X_s, Y_s, Z_s, u^\delta), \quad f_{\delta'}(s) := f_{\delta'}(s, X'_s, Y'_s, Z'_s, u^{\delta'}) \\ & g_\delta(s) := g_\delta(s, X_s, Y_s), \qquad g_{\delta'}(s) = g_{\delta'}(s, X'_s, Y'_s) \\ & h_\delta(s) := h_\delta(s, X_s, u^\delta_s), \qquad h_{\delta'}(s) = h_{\delta'}(s, X'_s, u^{\delta'}_s)\,. \end{aligned}$$

Applying Itô's formula to $|Y'_s - Y_s|^2$, we obtain

$$
\begin{aligned}
|Y'_t - Y_t|^2 + \int_t^T |Z'_s - Z_s|^2 d\langle W\rangle_s =& - \int_t^T 2|Y'_s - Y_s||Z'_s - Z_s| dW_s \\
&- \int_t^T |Y'_s - Y_s|^2 dM_s + |\Phi_\delta(X_T) - \Phi_{\delta'}(X'_T)|^2 \\
&+ \int_t^T 2|Y'_s - Y_s|(f_\delta(s) - f_{\delta'}(s)) ds \\
&+ \int_t^T 2|Y'_s - Y_s|(g_\delta(s) - g_{\delta'}(s)) d\langle W\rangle_s .
\end{aligned}
$$

Let $J_t = \int_t^T 2|Y'_s - Y_s||Z'_s - Z_s| dW_s + \int_t^T |Y'_s - Y_s|^2 dM_s$; then

$$
\begin{aligned}
|Y'_t - Y_t|^2 + J_t =& |\Phi_\delta(X_T) - \Phi_{\delta'}(X'_T)|^2 + \int_t^T 2|Y'_s - Y_s|(f_\delta(s) - f_{\delta'}(s)) ds \\
&+ \int_t^T 2|Y'_s - Y_s|(g_\delta(s) - g_{\delta'}(s)) d\langle W\rangle_s .
\end{aligned}
$$

According to [7], the process $J$ is a $G$-martingale, therefore

$$
\begin{aligned}
\hat{\mathbb{E}}(|Y'_t - Y_t|^2) \leq& \hat{\mathbb{E}}(|\Phi_{\delta'}(X'_T) - \Phi_\delta(X_T)|^2) + \hat{\mathbb{E}}\left(\int_t^T 2|Y'_s - Y_s|(f_\delta(s) - f_{\delta'}(s)) ds\right) \\
&+ \hat{\mathbb{E}}\left(\int_t^T 2|Y'_s - Y_s|(g_\delta(s) - g_{\delta'}(s)) d\langle W\rangle_s\right),
\end{aligned}
$$

and so, by Young's inequality,

$$
\begin{aligned}
\hat{\mathbb{E}}(|Y'_t - Y_t|^2) \leq& \hat{\mathbb{E}}(|\Phi_{\delta'}(X'_T) - \Phi_\delta(X'_T) + \Phi_\delta(X'_T) - \Phi_\delta(X_T)|^2) \\
&+ \hat{\mathbb{E}}\left(\int_t^T \frac{1}{\epsilon}|Y'_s - Y_s|^2 + \epsilon|f_\delta(s) - f_{\delta'}(s)|^2 ds\right) \\
&+ \hat{\mathbb{E}}\left(\int_t^T \frac{1}{\epsilon_1}|Y'_s - Y_s|^2 + \epsilon_1|g_\delta(s) - g_{\delta'}(s)|^2 d\langle W\rangle_s\right).
\end{aligned}
$$

Using the BDG inequality under $G$-expectations, [2, Lemma 2.18], with $p = 1$, we have

$$
\begin{aligned}
\hat{\mathbb{E}}(|Y'_t - Y_t|^2) \leq& 2\hat{\mathbb{E}}(|\Phi_{\delta'}(X'_T) - \Phi_\delta(X'_T)|^2) + 2\hat{\mathbb{E}}(|\Phi_\delta(X'_T) - \Phi_\delta(X_T)|^2) \\
&+ \frac{1}{\epsilon}\int_t^T \hat{\mathbb{E}}(|Y'_s - Y_s|^2) \\
&+ \epsilon\hat{\mathbb{E}}\left(\int_t^T |f_\delta(s) - f_{\delta'}(s)|^2 ds\right) + \frac{\underline{l} + \bar{l}}{4\epsilon_1}\int_t^T \hat{\mathbb{E}}(|Y'_s - Y_s|^2) ds
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{\epsilon_1(\underline{l}+\bar{l})}{4}\hat{\mathbb{E}}\left(\int_t^T |g_\delta(s)-g_{\delta'}(s)|^2 ds\right)\\
\leq & 2\hat{\mathbb{E}}(|\Phi_{\delta'}(X'_T)-\Phi_\delta(X'_T)|^2)+2\hat{\mathbb{E}}(|\Phi_\delta(X'_T)-\Phi_\delta(X_T)|^2)\\
&+\frac{1}{\epsilon}\int_t^T \hat{\mathbb{E}}(|Y'_s-Y_s|^2)ds\\
&+\epsilon\hat{\mathbb{E}}\bigg(\int_0^T |f_\delta(s)-f_\delta(s,X'_s,Y'_s,Z'_s,u^\delta)+f_\delta(s,X'_s,Y'_s,Z'_s,u^\delta)\\
&-f_{\delta'}(s,X'_s,Y'_s,Z'_s,u^\delta)+f_{\delta'}(s,X'_s,Y'_s,Z'_s,u^\delta)-f_{\delta'}(s)|^2 ds\bigg)\\
&+\frac{\underline{l}+\bar{l}}{4\epsilon_1}\int_t^T \hat{\mathbb{E}}(|Y'_s-Y_s|^2)ds+\frac{\epsilon_1(\underline{l}+\bar{l})}{4}\hat{\mathbb{E}}\left(\int_0^T |g_\delta(s)-g_{\delta'}(s)|^2 ds\right).
\end{aligned}
$$

Since the function $f$ is Lipschitz, with Lipschitz constant $L$, and bounded by $b_f$, the properties of the mollifier function imply that

$$
\begin{aligned}
\hat{\mathbb{E}}(|Y'_t-Y_t|^2)\leq & 2L^2|\delta'-\delta|^2+2L^2\hat{\mathbb{E}}(|X'_T-X_T|^2)+\frac{1}{\epsilon}\int_t^T \hat{\mathbb{E}}(|Y'_s-Y_s|^2)ds\\
&+6L^2\epsilon\hat{\mathbb{E}}\left(\int_t^T |(|X_s-X'_s|^2+|Y_s-Y'_s|^2+|Z_s-Z'_s|^2)ds\right)\\
&+2L^2(T-t)\epsilon|\delta'-\delta|^2+2L^2(T-t)\epsilon b_f\\
&+\frac{\underline{l}+\bar{l}}{4\epsilon_1}\int_t^T \hat{\mathbb{E}}(|Y'_s-Y_s|^2)ds\\
&+\frac{\epsilon_1(\underline{l}+\bar{l})}{4}\hat{\mathbb{E}}\left(\int_t^T |g_\delta(s,X_s,Y_s)-g_{\delta'}(s,X_s,Y_s)|^2 ds\right)\\
&+\frac{\epsilon_1(\underline{l}+\bar{l})}{4}\hat{\mathbb{E}}(\int_t^T |g_{\delta'}(s)-g_{\delta'}(s)|^2 ds)\\
\leq & 2L^2|\delta'-\delta|^2+2L^2\hat{\mathbb{E}}(|X'_T-X_T|^2)\\
&+\frac{1}{\epsilon}\int_t^T \hat{\mathbb{E}}(|Y'_s-Y_s|^2)ds\\
&+6L^2\epsilon\hat{\mathbb{E}}\left(\int_t^T |X_s-X'_s|^2+|Y_s-Y'_s|^2+|Z_s-Z'_s|^2 ds\right)\\
&+2L^2(T-t)\epsilon|\delta'-\delta|^2+2L^2(T-t)\epsilon b_f\\
&+\frac{\underline{l}+\bar{l}}{4\epsilon_1}\int_t^T \hat{\mathbb{E}}(|Y'_s-Y_s|^2)ds+(T-t)\frac{\epsilon_1(\underline{l}+\bar{l})L^2}{4}|\delta-\delta'|^2\\
&+\frac{2L^2\epsilon_1(\underline{l}+\bar{l})}{4}\hat{\mathbb{E}}\left(\int_0^T |X_s-X'_s|^2+|Y_s-Y'_s)|^2 ds\right).
\end{aligned}
$$

The last inequality implies

$$
\begin{aligned}
\hat{\mathbb{E}}(|Y'_t - Y_t|^2) \leq & \left(2L^2 + 6L^2(T-t)\epsilon + (T-t)\frac{2\epsilon_1(\underline{l}+\bar{l})L^2}{4}\right)|\delta' - \delta|^2 \\
& + 2L^2\hat{\mathbb{E}}(|X'_T - X_T|^2) \\
& + \left(\frac{1}{\epsilon} + 2L^2\epsilon + \frac{\underline{l}+\bar{l}}{4\epsilon_1} + \frac{2L^2\epsilon_1(\underline{l}+\bar{l})}{4}\right)\int_t^T \hat{\mathbb{E}}(|Y'_s - Y_s|^2)ds \\
& + 2L^2(T-t)\epsilon b_f \\
& + 6L^2\epsilon\hat{\mathbb{E}}\left(\int_0^T |Z_s - Z'_s|^2 ds\right) \\
& + \left(6L^2\epsilon + \frac{2L^2\epsilon_1(\underline{l}+\bar{l})}{4}\right)\hat{\mathbb{E}}\left(\int_0^T |X_s - X'_s|^2 ds\right).
\end{aligned}
\tag{21}
$$

On the other hand,

$$
\begin{aligned}
X_t - X'_t = & \int_0^t (b_\delta(s, X_s, u^\delta) - b_{\delta'}(s, X'_s, u^{\delta'}))ds + \int_0^t (\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X'_s))dW_s \\
& + \int_0^t (h_\delta(s) - h_{\delta'}(s))d\langle W\rangle_s.
\end{aligned}
$$

We apply Itô's formula to $|X_s - X'_s|^2$ where $X_t$ and $X'_t$ have the same initial condition, so $|X_t - X'_t| = 0$, then

$$
\begin{aligned}
|X_t - X'_t|^2 = & \int_0^t 2|X_s - X'_s|(b_\delta(s, X_s, u^\delta) - b_{\delta'}(s, X'_s, u^{\delta'}))ds \\
& + \int_0^t 2|X_s - X'_s|(\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X'_s))dW_s \\
& + \int_0^t 2|X_s - X'_s|(h_\delta(s) - h_{\delta'}(s)) + (\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X'_s))^2 d\langle W\rangle_s.
\end{aligned}
$$

By Young's inequality, we have for any $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$

$$
\begin{aligned}
|X_t - X'_t|^2 \leq & \int_0^t \left(\frac{1}{\varepsilon_1}|X_s - X'_s|^2 + \varepsilon_1|b_\delta(s, X_s, u^\delta) - b_{\delta'}(s, X'_s, u^{\delta'})|^2\right)ds \\
& + \int_0^t 2|X_s - X'_s|(\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X'_s))dW_s \\
& + \int_0^t \varepsilon_2|X_s - X'_s|^2 + |h_\delta(s) - h_{\delta'}(s)|^2 \\
& + (\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X'_s))^2 d\langle W\rangle_s.
\end{aligned}
$$

Since

$$\hat{\mathbb{E}}\left(\int_0^t \varepsilon_3|X_s - X_s'|^2 + \frac{1}{\varepsilon_3}|\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X_s')|^2 dW_s\right) = 0,$$

we obtain after taking expectations

$$\begin{aligned}\hat{\mathbb{E}}(|X_t - X_t'|^2) \leq &\hat{\mathbb{E}}\left(\int_0^t \left(\frac{1}{\varepsilon_1}|X_s - X_s'|^2 + \varepsilon_1|b_\delta(s, X_s, u^\delta) - b_{\delta'}(s, X_s', u^{\delta'})|^2\right) ds\right)\\ &+ \hat{\mathbb{E}}\left(\int_0^t \frac{1}{\varepsilon_2}|X_s - X_s'|^2 + \varepsilon_2(|h_\delta(s) - h_{\delta'}(s)|^2\right.\\ &\left.+(\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X_s'))^2)d\langle W\rangle_s\right).\end{aligned}$$

It then follows from the BDG inequality in the $G$-framework for $p = 1$ that

$$\begin{aligned}\hat{\mathbb{E}}(|X_t - X_t'|^2) \leq &\left(\frac{1}{\varepsilon_1} + \frac{(\bar{l} + \underline{l})}{4\varepsilon_2}\right)\int_0^t \hat{\mathbb{E}}(|X_s - X_s'|^2)\\ &+ \varepsilon_1\hat{\mathbb{E}}\left(\int_0^t |b_\delta(s, X_s, u^\delta) - b_{\delta'}(s, X_s', u^{\delta'})|^2 ds\right)\\ &+ \frac{(\bar{l} + \underline{l})\varepsilon_2}{4}\hat{\mathbb{E}}\left(\int_0^t |h_\delta(s) - h_{\delta'}(s)|^2 ds\right)\\ &+ \frac{(\bar{l} + \underline{l})\varepsilon_2}{4}\hat{\mathbb{E}}\left(\int_0^t |\sigma_\delta(s, X_s) - \sigma_{\delta'}(s, X_s')|^2 ds\right)\\ \leq &\left(\frac{1}{\varepsilon_1} + \frac{(\bar{l} + \underline{l})}{4\varepsilon_2}\right)\int_0^t \hat{\mathbb{E}}(|X_s - X_s'|^2)ds\\ &+ \varepsilon_1\hat{\mathbb{E}}\left(\int_0^t |b_\delta(s, X_s, u^\delta) - b_{\delta'}(s, X_s, u^\delta)|^2 ds + \varepsilon_1\int_0^t |b_{\delta'}(s, X_s, u^\delta)\right.\\ &\left.-b_{\delta'}(s, X_s', u^{\delta'})|^2 ds\right)\\ &+ \frac{(\bar{l} + \underline{l})\varepsilon_2}{4}\hat{\mathbb{E}}\left(\int_0^t |h_\delta(s) - h_\delta(s, X_s', u^{\delta'})|^2 ds\right)\\ &+ \frac{(\bar{l} + \underline{l})\varepsilon_2}{4}\hat{\mathbb{E}}\left(\int_0^t |h_\delta(s, X_s', u^{\delta'}) - h_{\delta'}(s)|^2 ds\right)\\ &+ \frac{(\bar{l} + \underline{l})\varepsilon_2}{4}\hat{\mathbb{E}}\left(\int_0^t |\sigma_\delta(s, X_s) - \sigma_\delta(s, X_s')|^2 ds\right)\\ &+ \frac{(\bar{l} + \underline{l})\varepsilon_2}{4}\hat{\mathbb{E}}\left(\int_0^t |\sigma_\delta(s, X_s') - \sigma_{\delta'}(s, X_s')|^2 ds\right),\end{aligned}$$

which by Assumptions 2 and 3 implies

$$\hat{\mathbb{E}}(|X_t - X_t'|^2) \leq \left(\frac{1}{\varepsilon_1} + \frac{(\bar{l} + \underline{l})}{4\varepsilon_2}\right)\int_0^t \hat{\mathbb{E}}(|X_s - X_s'|^2)ds$$

$$+ \varepsilon_1 \left( tL^2|\delta - \delta'|^2 + 2L^2 \int_0^t |X_s - X_s'|^2 ds \right)$$
$$+ \frac{4(\bar{l} + \underline{l})L^2\varepsilon_2}{4} \int_0^t \hat{\mathbb{E}}(|X_s - X_s'|^2)ds + \frac{2(\bar{l} + \underline{l})\varepsilon_2 t}{4} b_h^2$$
$$+ \frac{(\bar{l} + \underline{l})L^2\varepsilon_2}{4} \int_0^t \hat{\mathbb{E}}(|X_s - X_s'|^2)ds + \frac{(\bar{l} + \underline{l})\varepsilon_2 t}{4} b_\sigma^2$$

Setting

$$b_{h\sigma} = \frac{2(\bar{l} + \underline{l})b_h^2}{4} + \frac{(\bar{l} + \underline{l})}{4} b_\sigma^2, \quad C_1 = \frac{1}{\varepsilon_1} + \frac{(\bar{l} + \underline{l})}{\varepsilon_2} + \frac{(\bar{l} + \underline{l})\, L^2}{\varepsilon_2 4} + \varepsilon_1 L^2 + \frac{(\bar{l} + \underline{l})L^2}{\varepsilon_3 4}.$$

the last inequality can be brought in the form

$$\hat{\mathbb{E}}(|X_t - X_t'|^2) \leq C_1 \int_0^t \hat{\mathbb{E}}(|X_s - X_s'|^2)ds + \varepsilon_1 t L^2|\delta - \delta'|^2 + b_{h\sigma}\varepsilon_2 t.$$

We now choose $\varepsilon_1$ and $\varepsilon_2$ sufficiently small, such that

$$\varepsilon_1 \leq \frac{L^2}{b_{h\sigma}}, \quad b_{h\sigma}\varepsilon_2 t \leq \varepsilon_1 t L^2|\delta - \delta'|^2.$$

and therefore

$$\hat{\mathbb{E}}(|X_t - X_t'|^2) \leq C_1 \int_0^t \hat{\mathbb{E}}(|X_s - X_s'|^2)ds + 3\varepsilon_1 t L^2|\delta - \delta'|^2.$$

If we now apply Gronwall's Lemma, the last inequality turns into

$$\hat{\mathbb{E}}(|X_t - X_t'|^2) \leq e^{C_1 t}(3\varepsilon_1 t L^2|\delta - \delta'|^2). \tag{22}$$

By the same steps done above, we get

$$\hat{\mathbb{E}}(|X_T - X_T'|^2) \leq e^{C_1' T}(3\varepsilon_1' T L^2|\delta - \delta'|^2). \tag{23}$$

Now, by applying Itô's formula to $|Y_s' - Y_s|^2$, we find the following inequality

$$\hat{\mathbb{E}}\left(\int_t^T |Z_s' - Z_s|^2 d\langle W\rangle_s\right) \leq \left(2L^2 + 6L^2(T-t)\rho + (T-t)\frac{2\rho_1(\underline{l} + \bar{l})L^2}{4}\right)|\delta' - \delta|^2$$
$$+ 2L^2\hat{\mathbb{E}}(|X_T' - X_T|^2)$$
$$+ \left(\frac{1}{\rho} + 2L^2\rho + \frac{\underline{l} + \bar{l}}{4\rho_1} + \frac{L^2\rho_1(\underline{l} + \bar{l})}{4}\right)\int_t^T \hat{\mathbb{E}}(|Y_s' - Y_s|^2)ds$$
$$+ 2L^2(T-t)\rho b_f + 2L^2\rho\hat{\mathbb{E}}\left(\int_0^T |Z_s - Z_s'|^2 ds\right)$$

$$+\left(2L^2\rho+\frac{L^2\rho_1(\underline{l}+\bar{l})}{4}\right)\hat{\mathbb{E}}\left(\int_0^T|X_s-X_s'|^2ds\right).$$

that holds for any $\rho, \rho_1 > 0$. Using the $G$-Itô isometry

$$\hat{\mathbb{E}}\left(\left|\int_t^T\xi dW_s\right|^2\right)=\hat{\mathbb{E}}\left(\int_t^T|\xi|^2d\langle W\rangle_s\right)$$

and the BDG inequality under $G$-expectations for $p=2$ (see [2]), we obtain

$$\begin{aligned}\hat{\mathbb{E}}\left(\int_0^T|Z_s'-Z_s|^2ds\right)\leq&\frac{1}{\underline{l}c_2}\left(2L^2+2L^2(T-t)\rho+(T-t)\frac{\rho_1(\underline{l}+\bar{l})L^2}{4}\right)|\delta'-\delta|^2\\&+\frac{2L^2}{\underline{l}c_2}\hat{\mathbb{E}}(|X_T'-X_T|^2)\\&+\frac{2L^2}{\underline{l}c_2}\left(\frac{1}{\rho}+2L^2\rho+\frac{\underline{l}+\bar{l}}{4\rho_1}+\frac{L^2\rho_1(\underline{l}+\bar{l})}{4}\right)\int_0^T\hat{\mathbb{E}}(|Y_s'-Y_s|^2)ds\\&+\frac{2L^2}{\underline{l}c_2}(T-t)\rho b_f+\frac{2L^2}{\underline{l}c_2}\rho\hat{\mathbb{E}}\left(\int_0^T|Z_s-Z_s'|^2ds\right)\\&+\frac{1}{\underline{l}c_2}\left(2L^2\rho+\frac{L^2\rho_1(\underline{l}+\bar{l})}{4}\right)\hat{\mathbb{E}}\left(\int_0^T|X_s-X_s'|^2ds\right),\end{aligned}$$

for some sufficiently small constant $c_2 > 0$. Choosing

$$\rho\leq\min\left\{\frac{\underline{l}c_2}{8L^2},\frac{2L^2+2L^2}{2L^2b_f}\right\},$$

then yields

$$\begin{aligned}\left(1-\frac{2L^2}{\underline{l}c_2}\right)\hat{\mathbb{E}}\left(\int_0^T|Z_s'-Z_s|^2ds\right)\leq&C_5|\delta'-\delta|^2+\frac{2L^2}{\underline{l}c_2}\hat{\mathbb{E}}(|X_T'-X_T|^2)\\&+C_y\int_0^T\hat{\mathbb{E}}(|Y_s'-Y_s|^2)ds+\frac{2L^2}{\underline{l}c_2}(T-t)\rho b_f\\&+C_x\hat{\mathbb{E}}\left(\int_0^T|X_s-X_s'|^2ds\right),\end{aligned}$$

where

$$\begin{aligned}C_5&=\frac{1}{\underline{l}c_2}\left(2L^2+2L^2(T-t)\rho+(T-t)\frac{\rho_1(\underline{l}+\bar{l})K^2}{4}\right)\\C_x&=\frac{1}{\underline{l}c_2}\left(2L^2\rho+\frac{L^2\rho_1(\underline{l}+\bar{l})}{4}\right)\\C_y&=\frac{2L^2}{\underline{l}c_2}\left(\frac{1}{\rho}+2L^2\rho+\frac{\underline{l}+\bar{l}}{4\rho_1}+\frac{L^2\rho_1(\underline{l}+\bar{l})}{4}\right).\end{aligned}$$

The last inequality can be recast as

$$\hat{\mathbb{E}}\left(\int_0^T |Z_s' - Z_s|^2 ds\right) \leq \frac{2C_5}{C_z}|\delta' - \delta|^2 + \frac{2C_z L^2}{\underline{l}c_2}\hat{\mathbb{E}}(|X_T' - X_T|^2) + \frac{C_y}{C_z}\int_0^T \hat{\mathbb{E}}(|Y_s' - Y_s|^2)ds + \frac{C_x}{C_z}\hat{\mathbb{E}}\left(\int_0^T |X_s - X_s'|^2 ds\right), \tag{24}$$

with the shorthands $C_z = 1 - \frac{2L^2}{\underline{l}c_2}$ We can choose $\rho > 0$ such that

$$\frac{2L^2}{\underline{l}c_2}(T-t)\rho b_f \leq C_5|\delta' - \delta|^2 \quad \text{and} \quad \rho \leq \min\left\{\frac{\underline{l}c_2}{8L^2}, \frac{2L^2 + 2L^2}{2L^2 b_f}\right\}.$$

and substitute (24) in (21). Furthermore, we define

$$\bar{C}_1 = 2\left(2L^2 + 2L^2(T-t)\epsilon + (T-t)\frac{3\epsilon_1(\underline{l}+\bar{l})L^2}{4}\right) + 6L^2\epsilon + \frac{2C_5}{C_z}$$
$$\bar{C}_2 = \frac{1}{\epsilon} + 6L^2\epsilon + \frac{\underline{l}+\bar{l}}{4\epsilon_1} + \frac{3L^2\epsilon_1(\underline{l}+\bar{l})}{4} + \frac{C_y 2L^2\epsilon}{C_z}$$
$$\bar{C}_3 = 2L^2\epsilon + \frac{3L^2\epsilon_1(\underline{l}+\bar{l})}{4} + \frac{C_x 2L^2\epsilon}{C_z},$$

which allows us to simplify the following expression:

$$\hat{\mathbb{E}}(|Y_t' - Y_t|^2) \leq \bar{C}_1|\delta' - \delta|^2 + \left(2L^2 + \frac{24C_z L^4\epsilon}{\underline{l}c_2}\right)\hat{\mathbb{E}}(|X_T' - X_T|^2) + \bar{C}_2\int_0^T \hat{\mathbb{E}}(|Y_s' - Y_s|^2)ds + \bar{C}_3\hat{\mathbb{E}}\left(\int_0^T |X_s - X_s'|^2 ds\right). \tag{25}$$

If we now substitute (22) and (23) in the (25), it follows after further simplifications

$$\begin{aligned}\hat{\mathbb{E}}(|Y_t' - Y_t|^2) \leq& \bar{C}_1|\delta' - \delta|^2 + \left(2L^2 + \frac{24C_z L^2 L^2\epsilon}{\underline{l}c_2}\right)e^{C_1' T}(3\varepsilon_1' T L^2|\delta - \delta'|^2) \\ &+ \bar{C}_2\int_0^T \hat{\mathbb{E}}(|Y_s' - Y_s|^2)ds + \bar{C}_3 e^{C_1 t}(3\varepsilon_1 t L^2|\delta - \delta'|^2), \hat{\mathbb{E}}(|Y_s' - Y_s|^2) \\ \leq& \left(\bar{C}_1 + \left(2L^2 + \frac{24C_z L^2 L^2\epsilon}{\underline{l}c_2}\right)e^{C_1' T}3\varepsilon_1' T L^2 + \bar{C}_3 e^{C_1 t}(3\varepsilon_1 t L^2)\right)|\delta - \delta'|^2 \\ &+ \bar{C}_2\int_0^T \hat{\mathbb{E}}(|Y_s' - Y_s|^2)ds,\end{aligned}$$

Finally, setting

$$\bar{C} = \bar{C}_1 + \left(2L^2 + \frac{24C_z L^2 L^2\epsilon}{\underline{l}c_2}\right)e^{C_1' T}(3\varepsilon_1' T L^2) + \bar{C}_3 e^{C_1 t}(3\varepsilon_1 t L^2)$$

and using Gronwall's Lemma, we find

$$\hat{\mathbb{E}}(|Y_t' - Y_t|^2) \le \bar{C} e^{\bar{C}_2 T} |\delta' - \delta|^2.$$

This proves Lemma 8. □

### A.2 Proof of Lemma 9

***Proof*** Since $G$ satisfies a uniform ellipticity condition, the unique bounded continuous viscosity solution $V^\delta$ of equation (11) is smooth with regularity $C^{1,2}([0,T] \times \mathbb{R}^n)$, and thus the regularity results of Krylov [8, Theorems 6.4.3 and 6.4.4] apply. As a consequence, $V^\delta$ satisfies (12).

Let $(t, x) \in [0, T] \times \mathbb{R}^n$ be given. We introduce the shorthands

$$\begin{aligned} B(t,x) &:= b_\delta(s, X^{\delta,t,x,u^\delta}, v^\delta(s, X^{\delta,t,x,u^\delta})) \\ \Xi(t,x) &:= \sigma_\delta(s, X^{\delta,t,x,u^\delta}) \\ \Theta(t,x) &:= h_\delta(s, X^{\delta,t,x,u^\delta}, v^\delta(s, X^{\delta,t,x,u^\delta}))\,. \end{aligned}$$

For every $s \in [t, T]$, the SDE

$$X_s^{t,x,\delta} = x + \int_t^s B(r, X_r^{t,x,\delta}) dr + \int_t^s \Xi(r, X_r^{t,x,\delta}) dW_r + \int_t^s \Theta(r, X_r^{t,x,\delta}) d\langle W \rangle_r,$$

has a unique solution. We define $Y^{t,x,\delta}$ and $Z^{t,x,\delta}$, $t \le s \le T$ by

$$Y_s^\delta = V^\delta(s, X_s^\delta), \quad \text{and} \quad Z_s^\delta = \sigma_\delta(s, X_s^\delta) \nabla V^\delta(s, X_s^\delta).$$

By applying Itô's formula to the function $(t, x) \mapsto V^\delta(t, x)$ using that

$$\begin{aligned} &\partial_t V(t,x)^\delta + \bar{H}^\delta(t, x, V^\delta, \nabla V^\delta, \nabla^2 V^\delta, u) = 0 \\ &V^\delta(T, x) = \Phi^\delta(x), \mathbb{R}^n, \end{aligned}$$

it follows that

$$\begin{aligned} Y_t^{t,x,\delta} = \Phi^\delta(X_T^{t,x,\delta}) &- \textstyle\int_t^T f_\delta(s, X_s^{t,x,\delta}, Y_s^{t,x,\delta}, Z_s^{t,x,\delta}, v_s^\delta) ds \\ &- \textstyle\int_t^T g_\delta(s, X_s^{t,x,\delta}, Y_s^{t,x,\delta}, Z_s^{t,x,\delta}) d\langle W^\delta \rangle_s + \int_t^T Z_s^{t,x,\delta} dW_s^\delta - (M_T - M_t), \end{aligned} \tag{26}$$

which shows that the process $(X_s^{t,x,\delta}, Y_s^{t,x,\delta}, Z_s^{t,x,\delta})$ is the solution of the $G$-FBSDE associated with the coefficients $\Phi_\delta, b_\delta, \sigma_\delta, h_\delta, f_\delta, g_\delta$ and initial data $(t, x)$.

To prove the bounds (13)–(15), we apply Itô's formula to the function $(t, x) \mapsto |Y_s^{t,x,\delta}|$, $t \le s \le T$. To this end, we drop all superscripts whenever there is no risk of confusion and write $X_s := X_s^{t,x,\delta}$, $Y_s := Y_s^{t,x,\delta}$, etc., and we define

$$J_t' = \int_t^T 2|Y_s||Z_s| dW_s + \int_t^T |Y_s|^2 dM_s.$$

Then

$$|Y_t|^2 + \int_t^T |Z_s|^2 d\langle W\rangle_s + J_t' = |\Phi_\delta(X_T)|^2 + \int_t^T 2|Y_s| f_\delta(s) ds + \int_t^T 2|Y_s| g_\delta(s) d\langle W\rangle_s,$$

which implies

$$|Y_t|^2 + J_t' \leq |\Phi_\delta(X_T)|^2 + \int_t^T 2|Y_s| f_\delta(s) ds + \int_t^T 2|Y_s| g_\delta(s) d\langle W\rangle_s,$$

and thus

$$\begin{aligned}
&\hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_t|^2\right) \\
&\leq \hat{\mathbb{E}}\left(|\Phi_\delta(X_T)|^2 + \sup_{t\in[0,T]} \int_t^T 2|Y_s| f_\delta(s) ds + \sup_{t\in[0,T]} \int_t^T 2|Y_s| g_\delta(s) d\langle W\rangle_s\right).
\end{aligned}$$

We then apply Young's and BDG inequalities to obtain for any $\epsilon - \epsilon_1 > 0$:

$$\begin{aligned}
\hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_t|^2\right) \leq & \hat{\mathbb{E}}\left(C_\Phi^2 + \sup_{t\in[0,T]}\left\{\int_t^T \epsilon|Y_s|^2 + \frac{1}{\epsilon}|f_\delta(s)|^2 ds\right\}\right) \\
& + C_2\bar{l}\hat{\mathbb{E}}\left(\sup_{t\in[0,T]}\left\{\int_t^T \epsilon_1|Y_s|^2 + \frac{1}{\epsilon_1}|g_\delta(s)|^2 ds\right\}\right) \\
\leq & C_\Phi^2 + T\epsilon\hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_s|^2\right) + \frac{TC_f^2}{\epsilon} \\
& + C_2\bar{l}T\epsilon_1\hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_s|^2\right) + \frac{C_2\bar{l}C_g^2 T}{\epsilon_1}.
\end{aligned}$$

Here $C_\Phi > 0$, $C_2 > 0$ and $C_f, C_g > 0$ are appropriate constants. Now, setting $\epsilon = (8T)^{-1}$ and $\epsilon_1 = (8C_2\bar{l}T)^{-1}$, it follows that

$$\hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_t|^2\right) \leq \frac{4C_\Phi^2}{3} + \frac{32T^2C_f^2}{3} + \frac{32C_2^2\bar{l}^2C_g^2T^2}{3}. \tag{27}$$

We conclude that there exist a constant $\bar{C} = \bar{C}(T) > 0$, such that (13) is satisfied.

Based on the a priori estimate of the supremum norm of $|\nabla V(t,x)|^2$ in the work of Ladyzhenskaya et al. [19, Theorem 6.1 & Chapter VII], we can estimate the gradient on every compact subset of $[0, T] \times \mathbb{R}^n$, moreover, we can extend this result to the cylinder $[0, T] \times \{x \in \mathbb{R}^n, |x| \leq n\}$ and $[0, T] \times \{x \in \mathbb{R}^n, |x| \leq n+1\}$. Exploiting the Lipschitz continuity of the FBSDE coefficients, $\sup\{|\nabla V(t,x)|^2 : t \in [0, T], \ |x| \leq n\}$

can be bounded from above by a constant that depends on $\bar{C}$, $T$ and $L$, but not on $n$. As a consequence, there exist a constant $\bar{\Gamma} = \bar{\Gamma}(L, T)$, such that

$$|\nabla V^\delta(t, x)| \leq \bar{\Gamma}, \quad (t, x) \in [0, T] \times \mathbb{R}^n.$$

This proves (14). Finally, it is easy to check that there exists a constant $\bar{\kappa} > 0$, such that the following bounds hold (see Appendix A.2.1 below):

$$\hat{\mathbb{E}}(|X_s - X_r|^2) \leq \bar{\kappa}(s - r), \quad \hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_s^{t,x,\delta} - Y_r^{t,x,\delta}|^2\right) \leq \bar{\kappa}(s - r)^2.$$

Further using the fact that $Y_r^{s,x,\delta} = V^\delta(r, X_r^{s,x,\delta})$, it follows that

$$\begin{aligned}
\hat{\mathbb{E}}(|V^\delta(s, x) - V^\delta(r, x)|^2) &\leq 2\hat{\mathbb{E}}(|V^\delta(s, x) - Y_r^{s,x,\delta}|^2) + 2\hat{\mathbb{E}}(|Y_r^{s,x,\delta} - V^\delta(r, x)|^2) \\
&\leq 2\bar{\kappa}(s - r) + 2\bar{\Gamma}\hat{\mathbb{E}}(|X_s - X_r|^2) \\
&\leq 2\bar{\kappa}(s - r) + 2\bar{\kappa}(s - r), \qquad \text{modifying} \quad \bar{\kappa} \\
&\leq 4\bar{\kappa}(s - r).
\end{aligned}$$

Hence (15) is proved. □

#### A.2.1 Estimating Growth in Time

Let $0 \leq t \leq r \leq s \leq T$. Then

$$\begin{aligned}
Y_s^{t,x,\delta} - Y_r^{t,x,\delta} = &- \int_s^T f_\delta(\tau)d\tau + \int_r^T f_\delta(\tau)d\tau - \int_s^T g_\delta(\tau)d\langle W\rangle_\tau + \int_r^T g_\delta(\tau)d\langle W\rangle_\tau \\
&+ \int_s^T Z_\tau^{t,x,\delta}dW_\tau - \int_r^T Z_\tau^{t,x,\delta}dW_\tau - (M_T - M_s) + (M_T - M_r) \\
&- \int_r^s f_\delta(\tau)d\tau - \int_r^s g_\delta(\tau)d\langle W\rangle_\tau \\
&+ \int_r^s Z_\tau^{t,x,\delta}dW_\tau + (M_s - M_r).
\end{aligned}$$

Comparing (26), and using the result (27), we conclude

$$\hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_s^{t,x,\delta} - Y_r^{t,x,\delta}|^2\right) \leq \frac{32(s - r)^2 C_f^2}{3} + \frac{32C_2^2\bar{l}^2 C_g^2(s - r)^2}{3},$$

and thus

$$\hat{\mathbb{E}}\left(\sup_{t\in[0,T]} |Y_s^{t,x,\delta} - Y_r^{t,x,\delta}|^2\right) \leq \bar{\kappa}(s - r)^2.$$

On the other hand, it follows from the gradient bound on $V^\delta$ that

$$\begin{aligned}\hat{\mathbb{E}}(|Y_r^{t,x,\delta} - V(r,x)|^2) &= \hat{\mathbb{E}}(|V(r, X_r^{t,x,\delta}) - V(r,x)|^2)\\ &\le \bar{\Gamma}\hat{\mathbb{E}}(|X_r^{t,x,\delta} - x|^2)\\ &\le \bar{\Gamma}\hat{\mathbb{E}}(|X_r^{t,x,\delta} - X_s^{t,x,\delta}|^2),\end{aligned}$$

where we have redefined $\bar{\Gamma}$ in the last step. For $X_s - X_r$, it readily follows that

$$X_s - X_r = \int_r^s (b_\delta(\tau, X_\tau, u_\tau)d\tau + \int_r^s \sigma_\delta(\tau, X_\tau)dW_\tau + \int_r^s h_\delta(\tau)d\langle W\rangle_\tau,$$

which, together with the boundedness of $b$, $\sigma$, $h$ and the BDG inequality yields

$$\hat{\mathbb{E}}(|X_s - X_r|^2) \le 2(s-r)^2 C_b^2 + 2C_2\bar{l}(s-r)C_\sigma^2 + \frac{(\underline{l}+\bar{l})^2}{8}(s-r)^2 C_h^2.$$

As a consequence

$$\hat{\mathbb{E}}(|X_s - X_r|^2) \le \bar{\kappa}(s-r).$$

## B Convergence of the Auxiliary Control Problem

***Proof of Lemma 11*** Let the sequence of processes $(X_s^n, Y_s^n)$ satisfy the following controlled $G$-SDE

$$\begin{aligned}&dX_s^n = b(s, X_s^n, Y_s^n, u_s^{\delta_n})ds + \sigma(s, X_s^n)dW_s + h(s, X_s^n, Y_s^n, u_s^{\delta_n})d\langle W\rangle_s,\\ &dY_s^n = -f(s, X_s^n, Y_s^n, (\sigma(s, X_s^n))^T w_s^{\delta_n}, u_s^{\delta_n})ds - g(s, X_s^n, Y_s^n, (\sigma(s, X_s^n))^T w_s^{\delta_n})d\langle W\rangle_s\\ &\qquad +(\sigma(s, X_s^{\delta_n}))^T w_s^{\delta_n})dW_s + dM_s^n,\\ &X_t^n = x, \qquad Y_t^n = V^{\delta_n}(t,x), \quad M_t^n = 0.\end{aligned}$$

Here $w_s^{\delta_n} = \nabla V^{\delta_n}(s, X_s^{\delta_n})$ and $u_s^{\delta_n} = v^{\delta^n}(s, X_s^{\delta_n})$ being the optimal control for the mollified process. In the following, we will use the shorthands

$$\sigma^T w_s^{\delta_n} := (\sigma(s, X_s^n))^T w_s^{\delta_n} \quad \text{and} \quad \sigma_{\delta_n}^T w_s^{\delta_n} := (\sigma_{\delta_n}(s, X_s^{\delta_n}))^T w_s^{\delta_n}.$$

We now consider a subsequence $(X_s^{\delta_n}, Y_s^{\delta_n})$, that for convenience we denote with the same index $n$ as the full sequence, and that satisfies the following controlled SDE

$$\begin{aligned}&dX_s^{\delta_n} = b_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, u_s^{\delta_n})ds + \sigma_{\delta_n}(s, X_s^{\delta_n})dW_s + h_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, u_s^{\delta_n})d\langle W\rangle_s,\\ &dY_s^{\delta_n} = -f_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n}, u_s^{\delta_n})ds - g_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n})d\langle W\rangle_s\\ &\qquad +\sigma_{\delta_n}^T w_s^{\delta_n}\, dW_s + dM_s^{\delta_n},\\ &X_t^{\delta_n} = x \quad Y_t^{\delta_n} = V^{\delta_n}(t,x), \quad M_t^{\delta_n} = 0\end{aligned}$$

We apply Itô's formula to $|X_t^{\delta_n} - X_t^n|^2$, which gives

$$
\begin{aligned}
&|X_t^n - X_t^{\delta_n}|^2 \\
&= \int_0^t 2|X_s^n - X_s^{\delta_n}|(b(s, X_s^n, Y_s^n, u_s^{\delta_n}) - b_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, u_s^{\delta_n}))ds \\
&+ \int_0^t 2|X_s^n - X_s^{\delta_n}|(\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n}))dW_s \\
&+ \int_0^t 2|X_s^n + X_s^{\delta_n}|(h(s, X_s^n, Y_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, Y_s, u_s^{\delta_n})) + (\sigma(s, X_s^n) \\
&\quad - \sigma_{\delta_n}(s, X_s^{\delta_n}))^2 d\langle W\rangle_s .
\end{aligned}
$$

As a consequence,

$$
\begin{aligned}
&\hat{\mathbb{E}}(|X_t^n - X_t^{\delta_n}|^2) \\
&= \hat{\mathbb{E}}\left(\int_0^t 2|X_s^n - X_s^{\delta_n}|(b(s, X_s^n, Y_s^n, u_s^{\delta_n}) - b_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, u_s^{\delta_n}))ds\right) \\
&\quad + \hat{\mathbb{E}}\left(\int_0^t 2|X_s^n - X_s^{\delta_n}(|\sigma(s, X_s^n) - (\sigma_{\delta_n}(s, X_s^{\delta_n}))dW_s)\right) \\
&\quad + \hat{\mathbb{E}}\left(\int_0^t 2|X_s^n - X_s^{\delta_n}|(h(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})) + (\sigma(s, X_s^n)\right. \\
&\quad \left. -\sigma_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n}))^2 d\langle W\rangle_s\right).
\end{aligned}
$$

Now using BDG inequalities (see [2, Lemma 2.18] and [7, Proposition2. 6]) for $p = 1$,

$$
\begin{aligned}
\hat{\mathbb{E}}(|X_t^n - X_t^{\delta_n}|^2) \leq & \hat{\mathbb{E}}\left(\int_0^t 2|X_s^n - X_s^{\delta_n}|(b(s, X_s^n, u_s^{\delta_n}) - b^{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n}))ds\right) \\
& + \frac{(\underline{l} + \bar{l})}{4}\hat{\mathbb{E}}\left(\int_0^t 2|X_s^n - X_s^{\delta_n}|h(s, X_s^n, u_s^{\delta_n}) - (h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})\right. \\
& \left. + (\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n}))ds\right),
\end{aligned}
$$

and (extensively) Young's inequality,

$$
\begin{aligned}
\hat{\mathbb{E}}(|X_t^n - X_t^{\delta_n}|^2) \leq & \hat{\mathbb{E}}\left(\int_0^t \frac{1}{\epsilon}|X_s^n - X_s^{\delta_n}|^2 + \epsilon|b(s, X_s^n, u_s^{\delta_n}) - b^{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|2ds\right) \\
& + \frac{(\underline{l} + \bar{l})}{4}\hat{\mathbb{E}}\left(\int_0^t \frac{1}{\epsilon_1}|X_s^n - X_s^{\delta_n}|^2 + 2\epsilon_1|h(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2\right. \\
& \left. + 2\epsilon_1|\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n})|^2 ds\right) \\
\leq & \hat{\mathbb{E}}\left(\int_0^t \frac{1}{\epsilon}|X_s^n - X_s^{\delta_n}|^2 + \epsilon|b(s, X_s^n, u_s^{\delta_n}) - b_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|2ds\right)
\end{aligned}
$$

$$
\begin{aligned}
&+ \frac{(\underline{l}+\bar{l})}{4}\hat{\mathbb{E}}\bigg( \int_0^t \frac{1}{\epsilon_1}|X_s^n - X_s^{\delta_n}|^2 + 2\epsilon_1 |h(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2 \\
&+ 2\epsilon_1 |\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n})|^2 ds \bigg) \\
\le& \left( \frac{1}{\epsilon} + \frac{(\underline{l}+\bar{l})}{4\epsilon_1} \right) \int_0^t \hat{\mathbb{E}}(|X_s^n - X_s^{\delta_n}|^2) ds \\
&+ \epsilon \hat{\mathbb{E}} \left( \int_0^t |b(s, X_s^n, u_s^{\delta_n}) - b_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|2 ds \right) \\
&+ \frac{\epsilon_1(\underline{l}+\bar{l})}{2} \hat{\mathbb{E}} \left( \int_0^t |h(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2 ds \right) \\
&\frac{(\underline{l}+\bar{l})\epsilon_1}{2} \hat{\mathbb{E}} \left( \int_0^t (\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n}))^2 ds \right).
\end{aligned}
$$

and

$$
\begin{aligned}
&\hat{\mathbb{E}} \left( \int_0^t |b(s, X_s^n, u_s^{\delta_n}) - b_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2 ds \right) \\
&\quad \le 2\hat{\mathbb{E}} \left( \int_0^t |b(s, X_s^n, u_s^{\delta_n}) - b^{\delta_n}(s, X_s^n, u_s^{\delta_n})|^2 ds \right) \\
&\quad + 2\hat{\mathbb{E}} \left( \int_0^t |b^{\delta_n}(s, X_s^n, u_s^{\delta_n}) - b^{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2 ds \right) \\
&\hat{\mathbb{E}} \left( \int_0^T |h(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2 ds \right) \\
&\quad \le 2\hat{\mathbb{E}} \left( \int_0^T |h(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^n, u_s^{\delta_n})|^2 ds \right) \\
&\quad + 2\hat{\mathbb{E}} \left( \int_0^T |h_{\delta_n}(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2 ds \right) \\
&\hat{\mathbb{E}} \left( \int_0^T |\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n})|^2 ds \right) \\
&\quad \le 2\hat{\mathbb{E}} \left( \int_0^T |\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^n)|^2 ds \right) \\
&\quad + 2\hat{\mathbb{E}} \left( \int_0^T |\sigma_{\delta_n}(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n}) ds \right)
\end{aligned}
$$

together with the mollification properties (cf. Properties 5))

$$
\begin{aligned}
&\hat{\mathbb{E}} \left( \int_0^t |b(s, X_s^n, u_s^{\delta_n}) - b_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|^2 ds \right) \\
&\quad \le 2TL_{\delta_n}^2 \delta_n^2 + 4L^2 \hat{\mathbb{E}} \left( \int_0^T |X_s^n - X_s^{\delta_n}|^2 ds \right)
\end{aligned}
$$

$$\hat{\mathbb{E}}\left(\int_0^t |h(s, X_s^n, u_s^{\delta_n}) - h_{\delta_n}(s, X_s^{\delta_n}, u_s^{\delta_n})|2ds\right)$$
$$\le 2TL_{\delta_n}^2\delta_n^2 + 2L^2\hat{\mathbb{E}}\left(\int_0^T |X_s^n - X_s^{\delta_n}|^2 ds\right)$$
$$\hat{\mathbb{E}}\left(\int_0^t |\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n})|2ds\right)$$
$$\le 2TL_{\delta_n}^2\delta_n^2 + 2L^2\hat{\mathbb{E}}\left(\int_0^T |X_s^n - X_s^{\delta_n}|^2 ds\right),$$

we obtain the following bound

$$\begin{aligned}\hat{\mathbb{E}}(|X_t^n - X_t^{\delta_n}|^2) \le \left(\frac{1}{\epsilon} + \frac{(\underline{l}+\bar{l})}{4\epsilon_1}\right)\int_0^t \hat{\mathbb{E}}(|X_s^n - X_s^{\delta_n}|^2)ds \\ + \epsilon 2TL_{\delta_n}^2\delta_n^2 + 2L^2\epsilon\hat{\mathbb{E}}\left(\int_0^T |X_s^n - X_s^{\delta_n}|^2 ds\right) \\ + \epsilon_1(\underline{l}+\bar{l})TL_{\delta_n}^2\delta_n^2 + \epsilon_1(\underline{l}+\bar{l})L^2\hat{\mathbb{E}}\left(\int_0^T |X_s^n - X_s^{\delta_n}|^2 ds\right) \\ + \epsilon_1(\underline{l}+\bar{l})TL_{\delta_n}^2\delta_n^2 + \epsilon_1(\underline{l}+\bar{l})L^2\hat{\mathbb{E}}\left(\int_0^T |X_s^n - X_s^{\delta_n}|^2 ds\right).\end{aligned}$$

Applying Gronwall's lemma, it follows that there exists a constant $K$ independent of $\delta_n$, such that

$$\hat{\mathbb{E}}(|X_t^n - X_t^{\delta_n}|^2) \le K\delta_n^2, \tag{28}$$

which proves the first part of the assertion.

For the second part, we have to estimate the solution of the corresponding $G$-BSDE. We start by applying Itô's formula to $|Y_s^n - Y_s^{\delta_n}|^2$, with the shorthand

$$\begin{aligned}J_t = \int_t^T 2|Y_s^n - Y_s^{\delta_n}||((\sigma(s, X_s^n))^T - (\sigma_{\delta_n}(s, X_s^{\delta^n}))^T)w_s^{\delta_n}|dW_s \\ + \int_t^T |Y_s^n - Y_s^{\delta_n}|^2 dM_s,\end{aligned}$$

where $M = M^n - M^{\delta_n}$. This yields

$$\begin{aligned}|Y_t^n - Y_t^{\delta_n}|^2 + J_t \le |\Phi(X_T^n) - \Phi^{\delta^n}(X_T^{\delta_n})|^2 \\ + \int_t^T 2|Y_s^n - Y_s^{\delta_n}|(f(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^{\delta_n}) - f_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n}, u_s^{\delta_n}))ds \\ + \int_t^T 2|Y_s^n - Y_s^{\delta_n}|(g(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}) - g_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n})d\langle W\rangle_s,\end{aligned}$$

Taking expectations, we obtain

$$
\begin{aligned}
&\hat{\mathbb{E}}(|Y_t^n - Y_t^{\delta_n}|^2 + J_t) \le \hat{\mathbb{E}}(|\Phi(X_T^n) - \Phi^{\delta^n}(X_T^{\delta_n})|^2) \\
&\quad + \hat{\mathbb{E}}\left(\int_t^T 2|Y_s^n - Y_s^{\delta_n}|(f(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^{\delta_n}) - f_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n}, u_s^{\delta_n}))ds\right) \\
&\quad + \hat{\mathbb{E}}\left(\int_t^T 2|Y_s^n - Y_s^{\delta_n}|(g(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}) - g_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n})d\langle W\rangle_s\right).
\end{aligned}
$$

As before, we apply BDG inequalities for $p = 1$,

$$
\begin{aligned}
&\hat{\mathbb{E}}(|Y_t^n - Y_t^{\delta_n}|^2) \le \hat{\mathbb{E}}(|\Phi(X_T^n) - \Phi^{\delta^n}(X_T^{\delta_n})|^2) \\
&\quad + \hat{\mathbb{E}}\left(\int_t^T 2|Y_s^n - Y_s^{\delta_n}|(f(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^{\delta_n}) - f_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n}, u_s^{\delta_n}))ds\right) \\
&\quad + \frac{(\underline{l} + \bar{l})}{4}\hat{\mathbb{E}}\left(\int_t^T 2|Y_s^n - Y_s^{\delta_n}|(g(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}) - g_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n})ds\right),
\end{aligned}
$$

and Young's inequality for any $\varepsilon, \varepsilon_1 > 0$,

$$
\begin{aligned}
\hat{\mathbb{E}}(|Y_t^n - Y_t^{\delta_n}|^2) \le \hat{\mathbb{E}}(|\Phi(X_T^n) - \Phi^{\delta^n}(X_T^{\delta_n})|^2) &+ \left(\frac{1}{\varepsilon} + \frac{(\underline{l} + \bar{l})}{4\varepsilon_1}\right)\hat{\mathbb{E}}\left(\int_t^T |Y_s^n - Y_s^{\delta_n}|^2 ds\right) \\
&+ \varepsilon\hat{\mathbb{E}}\left(\int_t^T |f(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^{\delta_n}) - f_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n}, u_s^{\delta_n})|^2 ds\right) \\
&+ \frac{(\underline{l} + \bar{l})}{4\varepsilon_1}\hat{\mathbb{E}}\left(\int_t^T |g(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}) - g_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n})|^2 ds\right).
\end{aligned}
$$

To simplify the notation, we define

$$
\begin{aligned}
f - f_{\delta_n} &:= f(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^{\delta_n}) - f_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n}, u_s^{\delta_n}) \\
g - g_{\delta_n} &:= g(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}) - g_{\delta_n}(s, X_s^{\delta_n}, Y_s^{\delta_n}, \sigma_{\delta_n}^T w_s^{\delta_n}).
\end{aligned}
$$

Then,

$$
\begin{aligned}
&\hat{\mathbb{E}}\left(\int_t^T |f - f_{\delta_n}|^2 ds\right) \\
&\quad \le 2\hat{\mathbb{E}}\left(\int_t^T |f - f_{\delta_n}(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^{\delta_n})|^2 ds\right) \\
&\qquad + 2\hat{\mathbb{E}}\left(\int_t^T |f_{\delta_n}(s, X_s^n, Y_s^n, \sigma^T w_s^{\delta_n}, u_s^{\delta_n}) - f_{\delta_n}|^2 ds\right) \\
&\quad \le 2(T - t)L_{\delta_n}^2 \delta_n^2 \\
&\qquad + 6L\hat{\mathbb{E}}\left(\int_t^T (|X_s^n - X_s^{\delta_n}|^2 + |Y_s^n - Y_s^{\delta_n}|^2 + |(\sigma^T - \sigma_{\delta_n})w_s^{\delta_n}|^2)ds\right).
\end{aligned}
$$

and, using the same argument,

$$\hat{\mathbb{E}}\left(\int_t^T |g - g_{\delta_n}|^2 ds\right) \leq 2(T-t)L_{\delta_n}^2 \delta_n^2 + 6L\hat{\mathbb{E}}(\int_t^T (|X_s^n - X_s^{\delta_n}|^2 + |Y_s^n - Y_s^{\delta_n}|^2 + |(\sigma^T - \sigma_{\delta_n}^T)w_s^{\delta_n}|^2)ds).$$

Hence,

$$\begin{aligned}
\hat{\mathbb{E}}(|Y_t^n - Y_t^{\delta_n}|^2) &\leq \left(\frac{1}{\varepsilon} + \frac{(\underline{l}+\bar{l})}{4\varepsilon_1}\right)\hat{\mathbb{E}}\left(\int_t^T |Y_s^n - Y_s^{\delta_n}|^2 ds\right) + \varepsilon 2(T-t)L_{\delta_n}^2 \delta_n^2 \\
&+ 6L\varepsilon\hat{\mathbb{E}}\left(\int_t^T (|X_s^n - X_s^{\delta_n}|^2 + |Y_s^n - Y_s^{\delta_n}|^2 + |(\sigma^T - \sigma_{\delta_n}^T)w_s^{\delta_n}|^2)ds\right) \\
&+ \frac{(T-t)L_{\delta_n}^2(\underline{l}+\bar{l})}{2\varepsilon_1}\delta_n^2 \\
&+ \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1}\hat{\mathbb{E}}\left(\int_t^T (|X_s^n - X_s^{\delta_n}|^2 + |Y_s^n - Y_s^{\delta_n}|^2 + |(\sigma^T - \sigma_{\delta_n}^T)w_s^{\delta_n}|^2)ds\right) \\
&\leq \left(\frac{1}{\varepsilon} + \frac{(\underline{l}+\bar{l})}{4\varepsilon_1} + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1} + 6L\varepsilon\right)\hat{\mathbb{E}}\left(\int_t^T |Y_s^n - Y_s^{\delta_n}|^2 ds\right) \\
&+ \varepsilon 2(T-t)L_{\delta_n}^2 \delta_n^2 + \frac{(T-t)L_{\delta_n}^2(\underline{l}+\bar{l})}{2\varepsilon_1}\delta_n^2 \\
&+ \left(6L\varepsilon + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1}\right)\hat{\mathbb{E}}\left(\int_t^T |X_s^n - X_s^{\delta_n}|^2 ds\right) \\
&+ \left(6L\varepsilon + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1}\right)\hat{\mathbb{E}}\left(\int_t^T |(\sigma^T - \sigma_{\delta_n}^T)w_s^{\delta_n}|^2 ds\right) \\
&\leq \left(\frac{1}{\varepsilon} + \frac{(\underline{l}+\bar{l})}{4\varepsilon_1} + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1} + 6L\varepsilon\right)\hat{\mathbb{E}}\left(\int_t^T |Y_s^n - Y_s^{\delta_n}|^2 ds\right) \\
&+ 2\varepsilon(T-t)L_{\delta_n}^2 \delta_n^2 + \frac{(T-t)L_{\delta_n}^2(\underline{l}+\bar{l})}{2\varepsilon_1}\delta_n^2 \\
&+ \left(6L\varepsilon + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1}\right)\hat{\mathbb{E}}(\int_t^T |X_s^n - X_s^{\delta_n}|^2 ds\right) \\
&+ \left(6L\varepsilon + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1}\right)\hat{\mathbb{E}}\left(\int_t^T \|\sigma - \sigma_{\delta_n}\|_F^2 |w_s^{\delta_n}|^2 ds\right),
\end{aligned}$$

where $\|\cdot\|_F$ in the last line denotes the Frobenius norm of a matrix. From the Lipschitz property

$$\hat{\mathbb{E}}\left(\int_0^t \|\sigma(s, X_s^n) - \sigma_{\delta_n}(s, X_s^{\delta_n})\|_F^2 ds\right) \leq 2TL_{\delta_n}^2 \delta_n^2 + 2L^2\hat{\mathbb{E}}\left(\int_0^T |X_s^n - X_s^{\delta_n}|^2 ds\right)$$

and the fact that $w^{\delta_n}$ is bounded due to Lemma 9, it follows that

$\hat{\mathbb{E}}(|Y_t^n - Y_t^{\delta_n}|^2)$

$$\leq \left(\frac{1}{\varepsilon} + \frac{(\underline{l}+\bar{l})}{4\varepsilon_1} + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1} + 6L\varepsilon\right)\hat{\mathbb{E}}\left(\int_t^T |Y_s^n - Y_s^{\delta_n}|^2 ds\right)$$
$$+ 2\varepsilon(T-t)L_{\delta_n}^2\delta_n^2 + \frac{(T-t)L_{\delta_n}^2(\underline{l}+\bar{l})}{2\varepsilon_1}\delta_n^2$$
$$+ \left(6L\varepsilon + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1}\right)C_w^2 2TL_{\delta_n}^2\delta_n^2$$
$$+ \left(6L\varepsilon + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1} + 2L^2(6L\varepsilon + \frac{3L(\underline{l}+\bar{l})}{2\varepsilon_1}\right)C_w^2)\hat{\mathbb{E}}\left(\int_t^T |X_s^n - X_s^{\delta_n}|^2 ds\right).$$

Hence, from (28) and Gronwall's Lemma,

$$\hat{\mathbb{E}}(|Y_t^n - Y_t^{\delta_n}|^2) \leq e^{C_1(T-t)}C\delta_n^2.$$

for some $C, C_1 > 0$. This concludes the proof. □

**Funding** Open Access funding enabled and organized by Projekt DEAL.



## References


1. Bahlali, K., Kebiri, O., Mezerdi, B., Mtiraoui, A.: Existence of an optimal control for a coupled FBSDE with a non degenerate diffusion coefficient. Stochastics **90**(6), 861–875 (2018)
2. Bai, X., Lin, Y.: On the existence and uniqueness of solutions to stochastic differential equations driven by G-Brownian motion with integral-Lipschitz coefficients. Acta Mathematicae Applicatae Sinica English Ser. **30**(3), 589–610 (2014)
3. Bismut, J.-M.: Conjugate convex functions in optimal stochastic control. J. Math. Anal. Appl. **44**(2), 384–404 (1973)
4. Denis, L., Martini, C., et al.: A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. Ann. Appl. Probab. **16**(2), 827–852 (2006)
5. El Karoui, N., Peng, S., Quenez, M.C.: Backward stochastic differential equations in finance. Math. Financ. **7**(1), 1–71 (1997)
6. Faizullah, F.: Existence of solutions for stochastic differential equations under G-Brownian motion with discontinuous coefficients. Zeitschrift für Naturforschung A **67a**, 692–698, 05 (2012)
7. Hu, M., Ji, S., Peng, S., Song, Y.: Backward stochastic differential equations driven by G-Brownian motion. Stoch. Proc. Appl. **124**(1), 759–784 (2014)
8. Krylov, N.V.: Nonlinear elliptic and parabolic equations of the second order, vol. 7. Springer, Nikolaj Vladimirovič Krylov (1987)
9. Mingshang, H., Ji, S.: Dynamic programming principle for stochastic recursive optimal control problem driven by a G-Brownian motion. Stoch. Process. Appl. **127**(1), 107–134 (2017)
10. Mingshang, H., Ji, S., Peng, S., Song, Y.: Comparison theorem, Feynman–Kac formula and Girsanov transformation for BSDEs driven by G-Brownian motion. Stoch. Proc. Appl. **124**(2), 1170–1195 (2014)

11. Mtiraoui, A., Bahlali, K., Kebiri, O.: Existence of an optimal control for a system driven by a degenerate coupled forward-backward stochastic differential equations. C. R. Math. **355**, 84–89 (2017)
12. Peng, S.: A general stochastic maximum principle for optimal control problems. SIAM J. Control. Optim. **28**(4), 966–979 (1990)
13. Peng, S.: Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. Stoch. Stoch. Rep. **37**(1–2), 61–74 (1991)
14. Peng, S.: A generalized dynamic programming principle and Hamilton–Jacobi–Bellman equation. Stochastics **38**(2), 119–134 (1992)
15. Peng, S.: G-Brownian motion and dynamic risk measure under volatility uncertainty. arXiv e-prints arXiv:0711.2834 (2007)
16. Peng, S.: Nonlinear Expectations and Stochastic Calculus Under Uncertainty: With Robust CLT and G-Brownian Motion, volume 95. Springer (2019)
17. Pardoux, E., Peng, S.G.: Adapted solution of a backward stochastic differential equation. Syst. Control Lett. **14**(1), 55–61 (1990)
18. Redjil, A., Choutri, S.E.: On relaxed stochastic optimal control for stochastic differential equations driven by G-Brownian motion. ALEA Lat. Am. J. Probab. **15**(1), 201–212 (2018)
19. Solonnikov, V.A., Ladyzhenskaya, O.A., Ural'tseva, N.N.: Linear and Quasilinear Equations of Parabolic Type. AMS, Rhode Island (1968)
20. Wang, B., Yuan, M.: Forward-backward stochastic differential equations driven by G-Brownian motion. Appl. Math. Comput. **349**, 39–47 (2019)
21. Wang, B., Yuan, M.: Existence of solution for stochastic differential equations driven by G-Lévy process with discontinuous coefficients. Adv. Differ. Equ. **2017**(1), 1–13 (2017)