Discrete-Time Approximation of Risk-Averse Control Problems for Diffusion Processes

Andrzej Ruszczyński∗ Jianing Yao†

August 22, 2016

Abstract

We consider optimal control problems for diffusion processes, where the objective functional is defined by a time-consistent dynamic risk measure. We focus on coherent risk measures defined by \( g \)-evaluations. For such problems, we construct a family of time and space perturbed systems with piecewise-constant control functions. We obtain a regularized optimal value function by a special mollification procedure. This allows us to establish a bound on the difference between the optimal value functions of the original problem and of the problem with piecewise-constant controls.

1 Introduction

The first introduction of a coherent (static) risk measure, by Artzner et al. \cite{2,3}, was motivated by the capital adequacy rules of the Basel Accord. Large volume of research were devoted to this area, Föllmer, Schied \cite{17} and Frittelli and Rosazza Gianin \cite{20} generalized it to convex risk measure, Ruszczyński and Shapiro \cite{41} studied it from the perspective of optimization. Several classical references concerning static risk measures are \cite{18,19,40,42}.

Further development of the theory of risk measures lead to a dynamic setting, in which the risk is measured at each time instance based on the updated information. The key condition of time-consistency allows for dynamic programming formulations. The discrete time case was extensively explored by Detlfsen and Scandolo \cite{13}, Bion-Nadal \cite{6}, Cheridito et al. \cite{9,10}, Föllmer and Penner \cite{16}, Frittelli and Scandolo \cite{21}, Riedel \cite{37}, and Ruszczyński and Shapiro \cite{40}. For the continuous-time case, Coquet, Hu, Mémin and Peng \cite{11} discovered that time-consistent dynamic risk measures can be represented as solutions of Backward Stochastic Differential Equations (BSDE) (see also \cite{33,38}). Inspired by that, Barrieu and El Karoui provided a comprehensive study in \cite{4,5}; further contributions being made by Delbean, Peng, and Rosazza Gianin \cite{12}, and Quenez and Sulem \cite{36} (for a more general model with Levy processes). In addition, application to finance

∗Rutgers University, Department of Management Science and Information Systems, Piscataway, NJ 08854, USA, Email: rusz@rutgers.edu
†Rutgers University, Department of Management Science and Information Systems, Piscataway, NJ 08854, USA, Email: jy346@rutgers.edu
was considered, for example, in [24]. Using the convergence results of Briand, Delyon and Mémin [7], Stadje [44] finds the drivers of BSDE corresponding to discrete-time risk measures.

As for control with risk aversion, in discrete time setting, Ruszczyński [39], Çavuş and Ruszczyński [8] and Fan and Ruszczyński [14] developed the concept of a Markov risk measure and proposed risk-averse dynamic programming equations as well as computation methods. Our intention is to use continuous-time dynamic risk measures as objective functionals in optimal control problems for diffusion processes. While the traditional continuous stochastic control is well developed and discussed in numerous books (see, e.g., [15, 23, 35, 47]), the risk-averse case appears to be largely unexplored. In the present paper, we consider the risk-averse case with coherent risk measures given by $g$-evaluations. Such control problems are closely related to forward–backward systems of stochastic differential equations (FBSDE) (see, [26, 34]). For controlled fully coupled FBSDEs, Li and Wei [25] obtained the dynamic programming equation and derived the corresponding Hamilton–Jacobi–Bellman equation. Maximum principle for forward–backward systems and corresponding games was derived in [27, 28], including models with Levy processes.

The contribution of this paper is the study of accuracy of discrete-time approximations of risk-averse optimal control problems with coherent risk measures given by $g$-evaluations. For the purpose of the study, we construct a family of perturbed systems with two types of perturbations: of the initial time and the initial state. For such a family, we integrate the value functions of a piecewise-constant control with respect to the said initial time and state values. This yields regularized functions for which Itô calculus can be applied. Using the earlier results on the Hamilton–Jacobi–Bellman equation for risk-averse problems, we establish an error bound of order $\Delta^{1/6}$, between the optimal values of the original system and a system with piecewise constant controls with time step $\Delta$.

Section §2 has a synthetic character. We review in it the concept of $\mathbb{F}$-consistent evaluations and the connections to backward stochastic differential equations and dynamic risk measures. In §3 we formulate the risk-averse optimal control problem and study its basic properties. In the meanwhile, we recall the dynamic programming equation and the risk-averse analog of the Hamilton-Jacobi-Bellman equation. In section §4 we construct a family of time and space perturbed problems. They are used in a specially designed mollification procedure in §5, which yields sufficiently smooth close approximations of the optimal value function. In §6 we prove that the accuracy of the control policies restricted to piecewise-constant controls is of the order $h^{1/3}$, where $h^2$ is the time discretization step.

2 Foundations

2.1 Nonlinear Expectations and Dynamic Risk Measures

We establish a suitable framework and briefly review the concept of $\mathbb{F}$-consistent nonlinear expectations (for an extensive treatment, see [33]). For $0 < T < \infty$, let $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ be a probability space, where $\mathbb{F} = \{\mathcal{F}_t\}_{0 \leq t \leq T}$ is a filtration. A vector-valued stochastic process $\{X_t\}_{0 \leq t \leq T}$ is said to be adapted to $\mathbb{F}$ if $X_t$ is an $\mathcal{F}_t$-measurable random variable for any $t \in [0, T]$.

We introduce the following notation.
• $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot | F_t]$;

• $\mathbb{P}^m[t, T]$: the set of $\mathbb{R}^m$-valued adapted processes on $[t, T] \times \Omega$;

• $L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^m)$: the set of $\mathbb{R}^m$-valued $\mathcal{F}_t$-measurable random variables $\xi$ such that $\|\xi\|^2 := \mathbb{E}[|\xi|^2] < \infty$; for $m = 1$, we write it $L^2(\Omega, \mathcal{F}, \mathbb{P})$;

• $S^{2, m}[t, T]$: the set of elements $Y \in \mathbb{P}^m[t, T]$ such that $\|Y\|_{S^{2, m}[t, T]}^2 := \mathbb{E}\left[\sup_{t \leq s \leq T} |Y_s|^2 \right] < \infty$; for $m = 1$, we write it $S^2[t, T]$;

• $H^{2, m}[t, T]$: the set of elements $Y \in \mathbb{P}^m[t, T]$, such that $\|Y\|_{H^{2, m}[t, T]}^2 := \mathbb{E}\left[\int_t^T |Y_s|^2 \, ds \right] < \infty$; for $m = 1$ we write it $H^2[t, T]$; \footnote{When the norm is clear from the context, the subscripts are skipped.}

• $C^{1,2}([t, T] \times \mathbb{R}^m)$ the space of functions $f : [t, T] \times \mathbb{R}^m \rightarrow \mathbb{R}$, which are differentiable with respect to the first argument and twice differentiable with respect to the second argument, with all these derivatives continuous with respect to both arguments;

• $C^{1,2}_b([t, T] \times \mathbb{R}^m)$ the space of functions $f \in C^{1,2}([t, T] \times \mathbb{R}^m)$ with all derivatives bounded and continuous with respect to both arguments;

• $C^\infty(B)$: the space of functions $f : B \rightarrow \mathbb{R}$ that are infinitely continuously differentiable with respect to all arguments and have compact support on $B \subset \mathbb{R}^n$.

With this notation, we can introduce the concept of a nonlinear expectation.

**Definition 2.1.** For $0 \leq T < \infty$, a nonlinear expectation is a functional $\rho_{0,T} : L^2(\Omega, F_T, \mathbb{P}) \rightarrow \mathbb{R}$ satisfying the strict monotonicity property:

if $\xi_1 \geq \xi_2$ a.s., then $\rho_{0,T}[\xi_1] \geq \rho_{0,T}[\xi_2]$;

if $\xi_1 \geq \xi_2$ a.s., then $\rho_{0,T}[\xi_1] = \rho_{0,T}[\xi_2]$ if and only if $\xi_1 = \xi_2$ a.s.;

and the constant preservation property:

$\rho_{0,T}[c 1_\Omega] = c, \quad \forall c \in \mathbb{R}$,

where $1_A$ is the characteristic function of the event $A \in \mathcal{F}_T$.

Based on that, the $\mathcal{F}$-consistent nonlinear expectation is defined as follows.

**Definition 2.2.** For a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}, \mathbb{F})$, a nonlinear expectation $\rho_{0,T}[\cdot]$ is $\mathcal{F}$-consistent if for every $\xi \in L^2(\Omega, F_T, \mathbb{P})$ and every $t \in [0, T]$ a random variable $\eta \in L^2(\Omega, F_t, \mathbb{P})$ exists such that

$\rho_{0,T}[\xi 1_A] = \rho_{0,T}[\eta 1_A] \quad \forall A \in \mathcal{F}_t$.

The variable $\eta$ in Definition 2.2 is uniquely defined, we denote it by $\rho_{t,T}[\xi]$. It can be interpreted as a nonlinear conditional expectation of $\xi$ at time $t$. We can now define for every $t \in [0, T]$ the corresponding nonlinear expectation $\rho_{0,t} : L^2(\Omega, F_t, \mathbb{P}) \rightarrow \mathbb{R}$ as follows: $\rho_{0,t}[\xi] = \rho_{0,T}[\xi]$, for all $\xi \in L^2(\Omega, F_t, \mathbb{P})$. In this way, a whole system of $\mathcal{F}$-consistent nonlinear expectations \{\rho_{s,t}\}_{0 \leq s \leq t \leq T} is defined.
Proposition 2.3. If $\rho_{0,T} [\cdot]$ is an $\mathbb{F}$-consistent nonlinear expectation, then for all $0 \leq t \leq T$ and all $\xi, \xi' \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, it has the following properties:

(i) **Generalized constant preservation**: If $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, then $\rho_{t,T}[\xi] = \xi$;

(ii) **Time consistency**: $\rho_{s,T}[\xi] = \rho_{s,t}[\rho_{t,T}[\xi]]$, for all $0 \leq s \leq t$;

(iii) **Local property**: $\rho_{t,T}[\xi \mathbb{1}_A + \xi' \mathbb{1}_{A'}] = \mathbb{1}_{A} \rho_{t,T}[\xi] + \mathbb{1}_{A'} \rho_{t,T}[\xi']$, for all $A \in \mathcal{F}_t$.

It follows that $\mathbb{F}$-consistent nonlinear expectations are special cases of dynamic time-consistent measures of risk, enjoying a number of useful properties. They do not, however, have the properties of convexity, translation invariance, or positive homogeneity, unless additional assumptions are made. We shall return to this issue in the next subsection.

2.2 Backward Stochastic Differential Equations and $g$-Evaluations

Close relation exists between $\mathbb{F}$-consistent nonlinear expectations on the space $L^2(\Omega, \mathcal{F}, \mathbb{P})$, with the natural filtration of the Brownian motion, and backward stochastic differential equations (BSDE) [29, 30, 32]. We equip $(\Omega, \mathcal{F}, \mathbb{P})$ with a $d$-dimensional Brownian filtration, i.e., $\mathcal{F}_t = \sigma\{W_s; 0 \leq s \leq T\} \cup \mathcal{N}$, where $\mathcal{N}$ is the collection of $\mathbb{P}$-null sets in $\Omega$. In this paper we consider the following one-dimensional BSDE:

$$- dY_t = g(t, Y_t, Z_t) \, dt - Z_t \, dW_t, \quad Y_T = \xi,$$

where the data is the pair $(\xi, g)$, called the terminal condition and the generator (or driver), respectively. Here, $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, and $g : [0, T] \times \mathbb{R} \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ is a measurable function (with respect to the product $\sigma$-algebra), which is nonanticipative, that is, $g(t, Y_t, Z_t)$ is $\mathcal{F}_t$-measurable for all $t \in [0, T]$.

The solution of the BSDE is a pair of processes $(Y, Z) \in S^2[0, T] \times H^{2,d}[0, T]$ such that

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s) \, ds - \int_t^T Z_s \, dW_s, \quad t \in [0, T].$$

The existence and uniqueness of the solution of (1) can be guaranteed under the following assumption.

Assumption 2.4 (Peng and Pardou [29]).

(i) $g$ is jointly Lipschitz in $(y, z)$, i.e., a constant $K > 0$ exists such that for all $t \in [0, T]$, all $y_1, y_2 \in \mathbb{R}$ and all $z_1, z_2 \in \mathbb{R}^d$ we have

$$|g(t, y_1, z_1) - g(t, y_2, z_2)| \leq K (|y_1 - y_2| + |z_1 - z_2|) \quad a.s.;$$

(ii) the process $g(\cdot, 0, 0) \in H^2[0, T]$.

Under Assumption 2.4, we can define the $g$-evaluation.

Definition 2.5. For each $0 \leq t \leq T$ and $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, the $g$-evaluation at time $t$ is the operator $\rho_{0,T}^{d} : L^2(\Omega, \mathcal{F}_T, \mathbb{P}) \to L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ defined as follows:

$$\rho_{t,T}^{d}[\xi] = Y_t,$$

where $(Y, Z) \in S^{2,d}[t, T] \times H^{2}[t, T]$ is the unique solution of (1).
The following theorem reveals the relationship between $g$-evaluation and $\mathbb{F}$-consistent nonlinear expectation.

**Theorem 2.6.** Let the driver $g$ satisfy Assumption 2.4 and the condition: $g(\cdot,0,0) \equiv 0$ a.s. Then the system of $g$-evaluations $(\rho^g_{t,T})_{0 \leq t \leq T}$ defined in (3) is a system of $\mathbb{F}$-consistent nonlinear expectations. Furthermore, we have

$$\lim_{s \uparrow t} \rho^g_{s,t}[\xi] = \xi, \quad \forall \xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}), \ t \in [0,T].$$

Surprisingly, Coquet, Hu, Mémin, and Peng proved in (11) that every $\mathbb{F}$-consistent nonlinear expectation which is “dominated” by $\rho^g_{0,T}$ (a $g$-evaluation with the driver $\mu|y| + \nu|z|$ with some $\nu, \mu > 0$) is in fact a $g$-evaluation with some $g$. The domination is understood as follows: $\rho_{0,T}[Y + \eta] - \rho_{0,T}[Y] \leq \rho^\mu_0[\eta]$, for all $Y, \eta \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

From now on we shall use only $g$-evaluations as time-consistent dynamic measures of risk. To ensure desirable properties of the resulting measures of risk, we shall impose additional conditions on the driver $g$.

**Assumption 2.7.** The driver $g$ satisfies for almost all $t \in [0,T]$ the following conditions:

(i) $g$ is deterministic and independent of $y$, that is, $g : [0,T] \times \mathbb{R}^d \to \mathbb{R}$, and $g(\cdot,0) \equiv 0$;

(ii) $g(t, \cdot)$ is convex for all $t \in [0,T]$;

(iii) $g(t, \cdot)$ is positively homogeneous for all $t \in [0,T]$.

Under these conditions, one can derive new properties of the evaluations $\rho^g_{t,T}$, $t \in [0,T]$, in addition to the general properties of $\mathbb{F}$-consistent nonlinear expectations stated in Proposition 2.3.

**Theorem 2.8.** Suppose $g$ satisfies Assumption 2.4 and condition (i) of Assumption 2.7. Then the system of $g$-evaluations $\rho^g_{t,r}$, $0 \leq t \leq r \leq T$ has the following properties:

(i) **Normalization:** $\rho^g_{t,r}(0) = 0$;

(ii) **Translation Property:** for all $\xi \in L^2(\Omega, \mathcal{F}_r, \mathbb{P})$ and $\eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$,

$$\rho^g_{t,r}(\xi + \eta) = \rho^g_{t,r}(\xi) + \eta, \quad \text{a.s.;}$$

If, additionally, condition (ii) of Assumption 2.7 is satisfied, then $\rho^g_{t,r}$ has the following property:

(iii) **Convexity:** for all $\xi, \xi' \in L^2(\Omega, \mathcal{F}_r, \mathbb{P})$ and all $\lambda \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $0 \leq \lambda \leq 1$,

$$\rho^g_{t,r}(\lambda \xi + (1 - \lambda) \xi') \leq \lambda \rho^g_{t,r}(\xi) + (1 - \lambda) \rho^g_{t,r}(\xi'), \quad \text{a.s..}$$

Moreover, if $g$ also satisfies condition (iii) of Assumption 2.7, then $\rho^g_{t,r}$ has also the following property:

(iv) **Positive Homogeneity:** for all $\xi \in L^2(\Omega, \mathcal{F}_r, \mathbb{P})$ and all $\beta \in L^\infty(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $\beta \geq 0$, we have

$$\rho^g_{t,r}(\beta \xi) = \beta \rho^g_{t,r}(\xi), \quad \text{a.s.}$$

It follows that under Assumptions 2.4 and 2.7, the $g$-evaluations $\rho^g_{t,r}$ are convex or coherent conditional measures of risk (depending on whether (iii) is assumed or not).

Finally, we can derive their dual representation, by specializing the general results of [5].
**Theorem 2.9.** Suppose $g$ satisfies Assumption 2.4 and 2.7. Then for all $0 \leq t \leq r \leq T$ and all $\xi \in L^2(\Omega, F_r, \mathbb{P})$ we have
\[
\rho^0_{t,r}(\xi) = \sup_{\Gamma \in \mathcal{A}_{t,r}} E \left[ \Gamma \xi \mid F_t \right]
\]
where $\mathcal{A}_{t,r} = \partial \rho^0_{t,r}(0)$ is defined as follows:
\[
\mathcal{A}_{t,r} = \left\{ \exp \left( \int_t^r \gamma_s dW_s - \frac{1}{2} \int_t^r |\gamma_s|^2 ds \right) : \gamma \in H^2[t,r], \gamma_s \in \partial g(s,0), s \in [t,r] \right\}.
\]

**Corollary 2.10.** A constant $C$ exists, such that for all $0 \leq t \leq r \leq T$ and all $\Gamma_{t,r} \in \mathcal{A}_{t,r}$ we have
\[
\|\Gamma_{t,r} - 1\| \leq \frac{r - t}{T} e^{CT}.
\]

**Proof.** It follows from the definition of $\mathcal{A}_{t,r}$ that $\Gamma_{t,r}$ is the solution of the SDE
\[
d\Gamma_{t,s} = \gamma_s \Gamma_{t,s} dW_s, \quad \gamma_s \in \partial g(s,0), \quad s \in [t,r], \quad \Gamma_{t,t} = 1.
\]
Using Itô isometry, we obtain the chain of relations
\[
\|\Gamma_{t,r} - 1\| = \int_t^r \|\gamma_s \Gamma_{t,s}\|^2 ds \leq \int_t^r \|\gamma_s\|^2 \|\Gamma_{t,s}\|^2 ds \leq \int_t^r \|\gamma_s\|^2 \left(1 + \|\Gamma_{t,s} - 1\|^2\right) ds.
\]
If $u$ is a uniform upper bound on the norm of the subgradients of $g(s,0)$ we deduce that $\|\Gamma_{t,r} - 1\|^2 \leq \Delta_s, s \in [t,r]$, where $\Delta$ satisfies the ODE: $\frac{d}{ds} \Delta_s = u(1 + \Delta_s)$, with $\Delta_t = 0$. Consequently,
\[
\|\Gamma_{t,r} - 1\|^2 \leq \Delta_r = e^{u^2(r-t)} - 1.
\]
The convexity of the exponential function yields the postulated bound.

\[\square\]

**3 The Risk-Averse Control Problem**

**3.1 Problem Formulation**

Our objective is to evaluate and optimize the risk of the cumulative cost generated by a diffusion process.

On the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, we consider control processes $u : [0, T] \times \Omega \to U$ such that $u(\cdot)$ is $\mathbb{F}$-adapted, where $U \subset \mathbb{R}^m$ is a compact set, and a diffusion process under any such control with initial time $t \in [0, T]$ and state $x \in \mathbb{R}^n$:
\[
\begin{aligned}
&dX^t_{s,x,u} = b(s, X^t_{s,x,u}, u_s) ds + \sigma(s, X^t_{s,x,u}, u_s) dW_s, \quad s \in [t, T], \\
&X^t_{t,x,u} = x.
\end{aligned}
\]

Here, $b : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n \times d}$ are Borel measurable functions. We also introduce the *cost rate* function: a measurable map $c : [0, T] \times \mathbb{R}^n \times U \to \mathbb{R}$, and the *final
state cost: a measurable function $\Psi : \mathbb{R}^n \to \mathbb{R}$. Therefore, the random cost accumulated on the interval $[t, T]$ for any $t \in [0, T]$ can be expressed as follows:

$$
\xi_{t,T}(u, x) := \int_t^T c(s, X_s^{t,x;u}, u_s) \, ds + \Psi(X_T^{t,x;u}), \quad \text{a.s.}
$$

(7)

**Assumption 3.1.** A constant $K > 0$ exists such that, for any $s \in [t, T]$ and $(x_1, u_1), (x_2, u_2) \in \mathbb{R}^n \times U$, the functions $b, \sigma, c$, and $\Psi$ satisfy the following conditions:

$$
|b(s, x_1, u_1) - b(s, x_2, u_2)| + |\sigma(s, x_1, u_1) - \sigma(s, x_2, u_2)| + |c(s, x_1, u_1) - c(s, x_2, u_2)| \\
\leq K(|x_1 - x_2| + |u_1 - u_2|),
$$

$$
|b(s, x_1, u_1)| + |\sigma(s, x_1, u_1)| + |c(s, x_1, u_1)| + |\Psi(x_1)| \leq K(1 + |x_1| + |u_1|).
$$

Under Assumption 3.1, the controlled diffusion process (6) has a strong solution and the cost functional is square integrable.

We define the control value function as follows:

$$
V^u(t, x) := \rho_{t,T}^u[\xi_{t,T}(u, x)], \quad \text{a.s.}
$$

(8)

where $\{\rho_{t,T}^u\}_{t \in [0, T]}$, is a system of $g$-evaluations discussed in section 2.2. Using Definition 2.5 we can express the control value function as follows:

$$
V^u(t, x) = \xi_{t,T}(u, x) + \int_t^T g(s, Z_s^{t,x;u}) \, ds - \int_t^T Z_s^{t,x;u} \, dW_s
$$

$$
= \Psi(X_T^{t,x;u}) + \int_t^T [c(s, X_s^{t,x;u}, u_s) + g(s, Z_s^{t,x;u})] \, ds - \int_t^T Z_s^{t,x;u} \, dW_s,
$$

where $(Y^{t,x;u}, Z^{t,x;u})$ solve the following BSDE:

$$
\left\{ 
\begin{array}{l}
- dY_s^{t,x;u} = [c(s, X_s^{t,x;u}, u_s) + g(s, Z_s^{t,x;u})] \, ds - Z_s^{t,x;u} \, dW_s, \quad s \in [t, T], \\
Y_T^{t,x;u} = \Psi(X_T^{t,x;u}).
\end{array}
\right.
$$

(9)

Equivalently, $V^u(t, x) = Y_T^{t,x;u}$.

If Assumptions 3.1, 2.4, and 2.7 are satisfied, then for every $(t, x) \in [0, T] \times \mathbb{R}^n$, the BSDE (9) has a unique solution $(Y^{t,x;u}, Z^{t,x;u}) \in S^2[t, T] \times H^2[t, T]$ (see, Peng [33]), and, therefore, the control value function is well-defined.

In this way, the study of a risk-averse controlled system has been reduced to the study of controlled forward-backward stochastic differential equations (FBSDE). Such systems were extensively studied by Ma and Yong in [26]; other important references are [1, 31, 46, 45]. In our case, the FBSDE is decoupled, that is, the solution of the backward equation does not affect the forward equation, which substantially simplifies the analysis and allows for further advances.

Notice, when the driver $g \equiv 0$, the control value function (8) reduces to the expected value of (7). The risk-aversion is incorporated if other other drivers satisfying Assumption 2.7 are considered. By the comparison theorem of Peng [30], if $g_1$ is dominated by $g_2$, i.e., $g_1 \leq g_2$, then $\rho_{t,T}^{g_1}(\xi_{t,T}(u, x)) \leq \rho_{t,T}^{g_2}(\xi_{t,T}(u, x))$ almost surely; the larger the driver, the more risk aversion in the objective functional. For example, if we use $g_1(t, z) = \kappa|z|$, and $g_2(t, z) = \kappa|z_1|$, with $\kappa > 0$, then $g_1$ dominates $g_2$. 

7
3.2 Risk-Averse Dynamic Programming and Hamilton–Jacobi–Bellman Equations

We now proceed to the control problem. We define the admissible control system as in Yong and Zhu [47, p. 177].

**Definition 3.2.** \( \mathcal{U}[t, T] \) is called an admissible control system if it satisfies the following conditions:

(i) \((\Omega, \mathcal{F}, P)\) is a complete probability space;
(ii) \(\{W(s)\}_{s \geq t}\) is an \(d\)-dimensional standard Brownian motion defined on \( (\Omega, \mathcal{F}, P) \) over \( [t, T] \) and \( \mathbb{F}^t = (\mathcal{F}^t_s)_{s \in [t,T]} \), where \( \mathcal{F}^t_s = \sigma\{W_s; t \leq s \leq T\} \cup \mathcal{N} \) and \( \mathcal{N} \) is the collection of all \( P \)-null sets in \( \mathcal{F}^t \);
(iii) \( u : [s, T] \times \Omega \to U \) is an \( \{\mathcal{F}^t_s\}_{s \geq t} \) adapted process with \( \mathbb{E} \int_t^T |u_s|^2 \, ds < +\infty \);
(iv) For any \( x \in \mathbb{R}^n \) the system (6)–(9) admits a unique solution \((X, Y, Z)\) on \((\Omega, \mathcal{F}, P, \mathbb{F}^t)\).

The optimal value function \( V : [0, T] \times \mathbb{R}^n \to \mathbb{R} \) is defined as follows:

\[
V(t, x) = \inf_{u \in \mathcal{U}[t, T]} V^u(t, x). \tag{10}
\]

The weak formulation of a risk-averse control problem is the following: given \( (t, x) \in [0, T] \times \mathbb{R}^n \), find \( u^* \in \mathcal{U}[t, T] \) such that

\[
V^{u^*}(t, x) = \inf_{u \in \mathcal{U}[t, T]} V^u(t, x). \tag{11}
\]

We can now formulate the dynamic programming equation for our control problem.

**Theorem 3.3.** ([43]). Suppose Assumptions 3.1, 2.4, and 2.7 are satisfied. Then, for any \( (t, x) \in [0, T] \times \mathbb{R}^n \) and all \( r \in [t, T] \), we have

\[
V(t, x) = \inf_{u(\cdot) \in \mathcal{U}} \rho^u_{t, r} \left[ \int_t^r c(s, X^t_{s; u}, u_s) \, ds + V(r, X^t_{r; u}) \right]. \tag{12}
\]

For \( \alpha \in U \) we define the Laplacian operator \( \mathbb{L}^\alpha \) as follows: for \( w \in C^{1,2}_b([0, T] \times \mathbb{R}^n) \) and \((t, x) \in [0, T] \times \mathbb{R}^n\),

\[
[\mathbb{L}^\alpha w](t, x) = \partial_t w(t, x) + \sum_{i,j=1}^n \frac{1}{2} \left( \sigma(t, x, \alpha) \sigma(t, x, \alpha)^\top \right)_{ij} \partial_{x_ix_j} w(t, x) + \sum_{i=1}^n b_i(t, x, \alpha) \partial_{x_i} w(t, x).
\]

On the space \( C^{1,2}_b([0, T] \times \mathbb{R}^n) \), we consider the following equation

\[
\min_{\alpha \in \mathcal{U}} \left\{ c(t, x, \alpha) + [\mathbb{L}^\alpha v](t, x) + g(t, [\mathcal{D}x \cdot \sigma^\alpha](t, x)) \right\} = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^n, \tag{13}
\]

with the boundary condition

\[
v(T, x) = \Psi(T, x), \quad x \in \mathbb{R}^n. \tag{14}
\]

We call (13)–(14) the risk-averse Hamilton–Jacobi–Bellman equation associated with the controlled system (6) and the risk functional (7). It is a generalization of the classical Hamilton–Jacobi–Bellman Equation with the extra term \( g(\cdot, \cdot) \) responsible for risk aversion. In the special case, when \( g \equiv 0 \), we obtain the standard equation.

The following two theorems can be derived from general results on fully coupled forward–backward systems in [25]. For decoupled systems, a direct proof is provided in [43].
Theorem 3.4. Suppose Assumptions 3.1, 2.4, and 2.7 are satisfied; in addition, the functions $b$ and $\sigma$ are bounded in $x$. Then the value function $V(\cdot, \cdot)$ is a viscosity solution of the equation (13)–(14).

It is clear that if $V \in C_{b}^{1,2}([t, T] \times \mathbb{R}^{n})$ then it satisfies (13)–(14). We can also prove the converse relation (verification theorem).

Theorem 3.5. Suppose the assumptions of Theorem 3.4 are fulfilled and let the function $K \in C_{b}^{1,2}([t, T] \times \mathbb{R}^{n})$ satisfy (13)–(14). Then $K(t, x) \leq V(t, x)$ for any control $u(\cdot) \in \mathcal{U}$ and all $(t, x) \in [0, T] \times \mathbb{R}^{n}$. Furthermore, if a control process $u^*(\cdot) \in \mathcal{U}$ exists, satisfying for almost all $(s, \omega) \in [0, T] \times \Omega$, together with the corresponding trajectory $X_{s}^{0, x; u^*}$, the relation

$$u_{s}^* \in \arg \min_{\alpha \in \mathcal{U}} \left\{ c(s, X_{s}^{0, x; u^*}, \alpha) + \mathbb{I}^{\alpha} K(s, X_{s}^{0, x; u^*}) + g(t, [\mathcal{D}_{s} K \cdot \sigma^{\alpha}](t, X_{s}^{0, x; u^*})) \right\},$$

then $K(t, x) = V(t, x) = V^{u^*}(t, x)$ for all $(t, x) \in [0, T] \times \mathbb{R}^{n}$.

4 Piecewise-Constant Control Policies and the Perturbed Problem

Let $h^{2} \in (0, 1]$ be a time discretization step. We use the square to simplify further analysis.

Definition 4.1. For any $h \in (0, 1]$ and $t \in [0, T)$, let $\mathcal{U}^{2}_{h}$ be the subset of $\mathcal{U}$ consisting of all $\mathbb{F}$-adapted processes $u_{t}$ which are constant on intervals $[t, t + h^{2})$, $[t + h^{2}, t + 2h^{2})$, ..., $[t + kh^{2}, T]$, where $T - h^{2} \leq t + kh^{2} \leq T$.

We define the corresponding value function $V_{h} : [0, T] \times \mathbb{R}^{n} \rightarrow \mathbb{R}$ as follows:

$$V_{h}(t, x) = \inf_{u(\cdot) \in \mathcal{U}^{2}_{h}} V^{u}(t, x).$$

We assume a stronger condition than Assumptions 2.4 and 3.1

Assumption 4.2. Let $\mu(t, x, z, \alpha)$ stand for $\sigma(t, x, \alpha)$, $b(t, x, \alpha)$, $c(t, x, \alpha)$, and $\Phi(x)$ We assume that a constant $K$ exists such that

(i) For all $\alpha, \alpha_1, \alpha_2 \in U$, $x, x_1, x_2 \in \mathbb{R}^n$, $z, z_1, z_2 \in \mathbb{R}^d$ we have

$$|\mu(t, x, z, \alpha)| \leq K,$$

$$|\mu(t, x_1, z_1, \alpha_1) - \mu(t, x_2, z_2, \alpha_2)| \leq K(|x_1 - x_2| + |z_1 - z_2| + |\alpha_1 - \alpha_2|);$$

(ii) For all $\alpha \in U$, $s, t \in [0, T]$, $x \in \mathbb{R}^n$, $z \in \mathbb{R}^d$, we have

$$|\mu(t, x, z, \alpha) - \mu(s, x, z, \alpha)| \leq K|t - s|^{1/2}.$$

\footnote{We sometimes write $\mu^\alpha$ instead of $\mu(\cdot, \cdot, \cdot, \alpha)$}
By general results on forward–backward systems, the system (6), (7) and (9) has a unique solution and thus both functions: $V$ in (10) and $V_h$ in (15), are well-defined. In particular, they are both deterministic. We focus on the difference between the value functions $V$ and $V_h$.

The idea is to embed the original control problem into a family of time and space perturbed problems, and then obtain a smooth approximation of the value function by means of an integral regularization (mollification).

Let $B = \{(\tau, \zeta) \in \mathbb{R} \times \mathbb{R}^n : \tau \in (-1, 0), |\zeta| < 1\}$. Consider a time $t \in [0,T]$ and time instants $t_i = t + ih^2$, $i = 0, 1, \ldots, k$ and $t_{k+1} = T$. For a piecewise-constant control $u_s = \alpha_i$, $s \in [t_i, t_{i+1})$, $i = 0, 1, \ldots, k$, and perturbations $\beta_i \in B$, $i = 0, 1, \ldots, k$, we define the perturbed controlled FBSDE system:

\[
\begin{align*}
    d\tilde{X}_s &= b(s + \varepsilon^2 \tau_i, \tilde{X}_s + \varepsilon \zeta_i, \alpha_i) \, ds + \sigma(s + \varepsilon^2 \tau_i, \tilde{X}_s + \varepsilon \zeta_i, \alpha_i) \, d\tilde{W}_s, \quad (17) \\
    d\tilde{Y}_s &= \left[c(s + \varepsilon^2 \tau_i, \tilde{X}_s + \varepsilon \zeta_i, \alpha_i) + g(s + \varepsilon^2 \tau_i, \tilde{Z}_s)\right] \, ds - \tilde{Z}_s \, d\tilde{W}_s, \quad (18)
\end{align*}
\]

with a fixed $\varepsilon > 0$, with the initial condition $\tilde{X}_t = x$, and with the final condition $\tilde{Y}_T = \Phi(\tilde{X}_T)$.

By general results on forward–backward systems, the system (6), (7) and (9) has a unique solution and thus both functions: $V$ in (10) and $V_h$ in (15), are well-defined. In particular, they are both deterministic. We focus on the difference between the value functions $V$ and $V_h$.

The process $\tilde{W}$ is a Brownian motion. We assume here that $b(t, x, \alpha) = b(0, x, \alpha)$, $\sigma(t, x, \alpha) = \sigma(0, x, \alpha)$, $c(t, x, \alpha) = c(0, x, \alpha)$, and $g(t, z) = g(0, z)$ for all $t \in [-\varepsilon^2, 0]$. We consider the following discrete-time optimal control problem associated with the system (17)–(18). At each time $t_i$, we select a control value $\alpha_i \in U$ and a perturbation $\beta_i \in B$. The system evolves to time $t_{i+1}$, when new controls $\alpha_{i+1}$ and $\beta_{i+1}$ are selected. The objective of the controller is to make $\tilde{Y}_i$ the smallest possible. From now on, we use $\tilde{\alpha}$ and $\tilde{\beta}$ to represent the random sequences $\alpha_i$ and $\beta_i$, for $i = 0, 1, \ldots, k$.

**Lemma 4.3.** Functions $\tilde{V}^{\tilde{\alpha}, \tilde{\beta}} : \{t_0, t_1, \ldots, t_k\} \times \mathbb{R}^n \rightarrow \mathbb{R}$ exist, such that for all $x_i \in \mathbb{R}^n$, if the system (17)–(18) starts at time $t_i$ from $\tilde{X}_t = x_i$, then $\tilde{Y}_i = \tilde{V}^{\tilde{\alpha}, \tilde{\beta}}(t_i, x_i)$. Moreover,

\[
\tilde{V}^{\tilde{\alpha}, \tilde{\beta}}(t_i, x_i) = \rho_{t_i + \varepsilon^2 \tau_i, t_{i+1} + \varepsilon^2 \tau_i}^g \left[ \int_{t_i + \varepsilon^2 \tau_i}^{t_{i+1} + \varepsilon^2 \tau_i} c(s, X_s^{t_i + \varepsilon^2 \tau_i, x_i + \varepsilon \zeta_i; \alpha_i, \alpha_i) \, ds \\
+ \tilde{V}^{\tilde{\alpha}, \tilde{\beta}}(t_{i+1}, X_{t_{i+1} + \varepsilon^2 \tau_i}^{t_i + \varepsilon^2 \tau_i, x_i + \varepsilon \zeta_i; \alpha_i - \varepsilon \zeta_i) \right]. \quad (19)
\]

**Proof.** With $\tilde{W}_s \sim W_{s+\varepsilon^2 \tau_i}$ for $s \in [t_i, t_{i+1})$ directly from the equations (17) and (6) we obtain:

\[
\tilde{X}_s^{t_i, x_i; \alpha_i} = X_{s+\varepsilon^2 \tau_i}^{t_i+\varepsilon^2 \tau_i, x_i+\varepsilon \zeta_i; \alpha_i} - \varepsilon \zeta_i, \quad s \in [t_i, t_{i+1}], \quad a.s. \quad (20)
\]

With this substitution, the BSDE (18) at $s = t_i$ is equivalent to:

\[
\begin{align*}
    \tilde{Y}_i &= \tilde{Y}_{i+1} + \int_{t_i + \varepsilon^2 \tau_i}^{t_{i+1} + \varepsilon^2 \tau_i} \left[c(s, X_s^{t_i + \varepsilon^2 \tau_i, x_i + \varepsilon \zeta_i; \alpha_i, \alpha_i) + g(s, Z_s)\right] \, ds - \int_{t_i + \varepsilon^2 \tau_i}^{t_{i+1} + \varepsilon^2 \tau_i} Z_s \, dW_s \\
    &= \rho_{t_i + \varepsilon^2 \tau_i, t_{i+1} + \varepsilon^2 \tau_i}^g \left[ \int_{t_i + \varepsilon^2 \tau_i}^{t_{i+1} + \varepsilon^2 \tau_i} c(s, X_s^{t_i + \varepsilon^2 \tau_i, x_i + \varepsilon \zeta_i; \alpha_i, \alpha_i) \, ds + \tilde{Y}_{i+1} \right].
\end{align*}
\]
By definition, \( \overline{Y}_T = \Phi(\overline{X}_T) \). Supposing \( \overline{Y}_{t_{j+1}} = \overline{V}^\alpha,\beta(t_{j+1}, \overline{X}_{t_{j+1}}) \) for some \( j \), proceeding backwards in time we conclude from the last equation that we can write \( \overline{Y}_{t_i} = \overline{V}^\alpha,\beta(t_i, x_i) \) for some function \( \overline{V}^\alpha,\beta(\cdot, \cdot) \). We can thus write the recursive relation

\[
\overline{V}^\alpha,\beta(t_i, x_i) = \rho^\beta_{t_i + \varepsilon^2 \tau, t_{i+1} + \varepsilon^2 \tau_i} \int_{t_i + \varepsilon^2 \tau_i}^{t_{i+1} + \varepsilon^2 \tau_i} c(s, X_{s}^t + \varepsilon^2 \tau, x_i + \varepsilon \zeta_i ; \alpha_i, \alpha_i) \, ds + \overline{V}^\alpha,\beta(t_{i+1}, \overline{X}_{t_{i+1}}) .
\]

Substitution of (20) proves (19).

Using this recursive relation, we define the value function of the optimally perturbed problem \( \overline{V}_{h,\varepsilon}(t_i, x_i) \) at each time \( t_i \) and the corresponding state \( x_i \) as follows. At \( t_{k+1} = T \) we set \( \overline{V}_{h,\varepsilon}(T, x_T) = \Phi(x_T) \), and then, proceeding backwards in time,

\[
\overline{V}_{h,\varepsilon}(t_i, x_i) = \inf_{\alpha_i \in U} \inf_{\beta_i \in B} \rho^\beta_{t_i + \varepsilon^2 \tau, t_{i+1} + \varepsilon^2 \tau_i} \left[ \int_{t_i + \varepsilon^2 \tau_i}^{t_{i+1} + \varepsilon^2 \tau_i} c(s, X_{s}^t + \varepsilon^2 \tau, x_i + \varepsilon \zeta_i ; \alpha_i, \alpha_i) \, ds + \overline{V}_{h,\varepsilon}(t_{i+1}, X_{t_{i+1}} + \varepsilon^2 \tau_i, x_i + \varepsilon \zeta_i ; \alpha_i - \varepsilon \zeta_i) \right].
\]

This construction can be carried out for every \( t \in [0, T] \) and the resulting points \( t_i = t + \varepsilon^2 h^2 \), thus defining a function \( \overline{V}_{h,\varepsilon} : [0, T] \times \mathbb{R}^n \to \mathbb{R} \) which satisfies the relation

\[
\overline{V}_{h,\varepsilon}(t, x) = \inf_{\alpha_i \in U} \inf_{\beta_i \in B} \rho^\beta_{t + \varepsilon^2 \tau, t + h^2 + \varepsilon^2 \tau} \left[ \int_{t + \varepsilon^2 \tau}^{t + h^2 + \varepsilon^2 \tau} c(s, X_{s}^t + \varepsilon^2 \tau, x + \varepsilon \zeta_i ; \alpha_i, \alpha_i) \, ds + \overline{V}_{h,\varepsilon}(t + h^2, X_{t + h^2} + \varepsilon^2 \tau, x + \varepsilon \zeta_i ; \alpha_i - \varepsilon \zeta_i) \right].
\]

If \( t \in (T - h^2, T) \) we replace \( t + h^2 \) with \( T \) in the above equation. The function \( \overline{V}_{h,\varepsilon}(t, x) \) represents the optimal value of the perturbed problem starting at time \( t \) from the state \( x \) and proceeding with piecewise constant controls and perturbations on intervals of length \( h^2 \) (except, perhaps, the last one, which ends at \( T \)). Let us stress that the perturbations are treated as additional controls in this construction.

We now present a number of useful estimates from Krylov [22].

**Lemma 4.4.** For \( t \in [0, T] \), \( x, y \in \mathbb{R}^n \), and \( \alpha \in U^t_h \), denote by \( \overline{X}^{t, x; \alpha, \beta} \) the solution of (17) and by \( X^{t, x; \alpha} \) the solution of (6), with the initial state \( x \in \mathbb{R}^n \) at time \( t \). Then

\[
\mathbb{E}\left[ \sup_{t \leq s \leq T} |\overline{X}^{t, x; \alpha, \beta} - X^{s, x; \alpha}|^2 \right] \leq N e^{N T} \varepsilon^2 , \quad (22)
\]

\[
\mathbb{E}\left[ \sup_{t \leq s \leq T} |\overline{X}^{t, x; \alpha, \beta} - \overline{X}^{t, y; \alpha, \beta}|^2 \right] \leq N e^{N T} |x - y|^2 , \quad (23)
\]

\[
\mathbb{E}\left[ \sup_{t \leq r \leq T} |\overline{X}^{t, x; \alpha, \beta} - \overline{X}^{t, y; \alpha, \beta}|^2 \right] \leq N e^{N T} |t - r| , \quad (24)
\]

for \( N > 0 \) depending on \( (K, d, n) \) only.
The proof is by using Theorem 2.5.9 in [23]. These estimates can be used to derive the following bounds.

**Lemma 4.5.** A constant \( N \) exists, depending on \((K, d, n)\) only, such that:

(i) For \( t \in [0, T] \) and \( x \in \mathbb{R}^d \), we have

\[
|V^{\alpha}(t, x) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, x)| \leq N e^{NT} \varepsilon.
\]

(ii) For \( t, r \in [0, T] \) and \( x, y \in \mathbb{R}^n \), we have

\[
|\bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, x) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(r, y)| \leq N e^{NT} (|x - y| + |t - r|^\frac{1}{2}).
\]

**Proof.** For fixed \( \alpha, \beta \), recall that

\[
V^{\alpha}(t, x) = \rho^0_{t, T} \left[ \int_t^T c(r, X^s, x; \alpha) \, dr + \Phi(X^T_t \alpha) \right],
\]

\[
\bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, x) = \rho^0_{t, T} \left[ \int_t^T c(r, \bar{X}^s_r, \beta; \bar{\alpha}_r) \, dr + \Phi(X^T_T \beta) \right].
\]

By standard estimates for BSDE and Lemma 4.4, we have with some \( \gamma > 0 \) depending on \((K, d, n)\),

\[
|V^{\alpha}(t, x) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, x)| \leq e \left[ |\Phi(X^t_T \alpha) - \Phi(\bar{X}^T_T \beta)|^2 e^{\gamma(T-t)} \right] + e \left[ \int_t^T |c(r, X^s_r, \alpha) - c(r, \bar{X}^s_r, \beta; \bar{\alpha}_r)|^2 e^{\gamma(r-t)} \, dr \right]
\]

\[
\leq K^2 e^{\gamma(T-t)} e \left[ \sup_{t \leq T} |X^t_T \alpha - \bar{X}^T_T \beta|^2 \right] + KT^2 e^{\gamma(T-t)} e \left[ \sup_{t \leq T} |X^t_T \alpha - \bar{X}^T_T \beta|^2 \right] \leq Ne^{NT} \varepsilon.
\]

The first assertion follows. For (ii), we observe that

\[
|\bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, x) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, y)| \leq |\bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, x) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, y)| + |\bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, y) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(r, y)|.
\]

Similar to the proof for (i), by applying the second and third inequalities of Lemma 4.4 we have

\[
|\bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, x) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, y)| \leq Ne^{NT} |x - y|^2;
\]

\[
|\bar{V}^{\alpha, \beta}_{h, \varepsilon}(t, y) - \bar{V}^{\alpha, \beta}_{h, \varepsilon}(r, y)| \leq Ne^{NT} |t - r|^\frac{1}{2},
\]

which implies the postulated estimates. \( \square \)

## 5 Mollification of the Value Function

We now introduce an integral transformation of the value function. We take a non-negative function \( \varphi \in C^\infty(B) \) with \( \int_B \varphi(\tau, \zeta) \, d\tau \, d\zeta = 1 \), called a mollifier. For \( \varepsilon > 0 \), we re-scale the mollifier as \( \varphi_\varepsilon(\tau, \zeta) = e^{-n-2} \varphi(\tau/\varepsilon^2, \zeta/\varepsilon) \), and we introduce the following notation of the convolution of the function \( \bar{V}^{\alpha, \beta}_{h, \varepsilon} \) with the re-scaled mollifier:

\[
\tilde{V}^{\alpha, \beta}_{h, \varepsilon}(t, x) = [\bar{V}^{\alpha, \beta}_{h, \varepsilon} \ast \varphi_\varepsilon](t, x) = \int_B \bar{V}^{\alpha, \beta}_{h, \varepsilon}(t - \varepsilon^2 \tau, x - \varepsilon \zeta) \varphi_\varepsilon(\tau, \zeta) \, d\tau \, d\zeta,
\]
where $t \in [0, T - \varepsilon^2]$ and $x \in \mathbb{R}^n$. We shall need an estimate of the seminorm $\|\hat{V}_{h,\varepsilon}\|_{2,1}$ defined as follows:

$$
\|w\|_{2,1} = \sup_{(t,x)} |w(t,x)| + \sup_{(t,x)} \|D_x w(t,x)\| + \sup_{(t,x)} \|D_{xx} w(t,x)\| + \sup_{(t,x)} |\partial_t w(t,x)|
$$

$$
+ \sup_{(t,x),(s,y)} \frac{\|D_{xx} w(t,x) - D_{xx} w(s,y)\|}{|t-s| + |x-y|} + \sup_{(t,x),(s,y)} \frac{|\partial_t w(t,x) - \partial_t w(s,y)|}{|t-s| + |x-y|}.
$$

In the formula above, we use $D_x$ and $D_{xx}$ to denote the gradient and the Hessian matrix, and the supremum is always over $(t,x),(s,y) \in [0, T - \varepsilon^2] \times \mathbb{R}^2$.

**Lemma 5.1.** If $\varepsilon \geq h$, then $\|\hat{V}_{h,\varepsilon}\|_{2,1} \leq Ne^{NT}\varepsilon^{-2}$ and $\|\hat{V}_{h,\varepsilon} - \hat{V}_{h,\varepsilon}\|_0 \leq Ne^{NT}\varepsilon$.

**Proof.** By elementary properties of the convolution,

$$
\frac{\partial}{\partial t} \hat{V}_{h,\varepsilon}(t,x) = \frac{\partial}{\partial t} \left( \hat{V}_{h,\varepsilon} * \varphi_{\varepsilon} \right)(t,x) = \left( \hat{V}_{h,\varepsilon} * \frac{\partial}{\partial t} \varphi(\varepsilon, \zeta) \right)(t,x)
$$

$$
= \varepsilon^{-2} \int_B \hat{V}_{h,\varepsilon}(t - \varepsilon^2 \tau, x - \varepsilon \zeta) \frac{\partial}{\partial \tau} \varphi(\tau, \zeta) \, d\tau \, d\zeta.
$$

Thus, due to Lemma 3.5(ii),

$$
\left| \frac{\partial}{\partial t} \hat{V}_{h,\varepsilon}(t,x) \right| = \varepsilon^{-2} \left| \int_B \hat{V}_{h,\varepsilon}(t - \varepsilon^2 \tau, x - \varepsilon \zeta) \frac{\partial}{\partial \tau} \varphi(\tau, \zeta) \, d\tau \, d\zeta \right|
$$

$$
= \varepsilon^{-2} \left| \int_B \left[ \hat{V}_{h,\varepsilon}(t - \varepsilon^2 \tau, x - \varepsilon \zeta) - \hat{V}_{h,\varepsilon}(t,x) \right] \frac{\partial}{\partial \tau} \varphi(\tau, \zeta) \, d\tau \, d\zeta \right|
$$

$$
\leq 2Ne^{NT} \varepsilon^{-1} \int_B \left| \frac{\partial}{\partial \tau} \varphi(\tau, \zeta) \right| \, d\tau \, d\zeta.
$$

We can thus increase $N$ to write for all $\varepsilon > 0$ the inequality

$$
\left| \frac{\partial}{\partial t} \hat{V}_{h,\varepsilon}(t,x) \right| \leq Ne^{NT} \varepsilon^{-1}.
$$

Similarly, after redefining $N$ in an appropriate way,

$$
\left| \frac{\partial}{\partial x^1} \hat{V}_{h,\varepsilon} \right|_0 \leq Ne^{NT} \varepsilon^{-1}, \quad \left| \frac{\partial^2}{\partial x^1 \partial x^2} \hat{V}_{h,\varepsilon} \right|_0 \leq Ne^{NT} \varepsilon^{-2},
$$

$$
\left| \frac{\partial^2}{\partial t \partial x^1} \hat{V}_{h,\varepsilon} \right|_0 + \left| \frac{\partial^3}{\partial x^1 \partial x^2 \partial \tau} \hat{V}_{h,\varepsilon} \right|_0 \leq Ne^{NT} \varepsilon^{-3},
$$

$$
\left| \frac{\partial^2}{\partial t \partial x^1} \hat{V}_{h,\varepsilon} \right|_0 + \left| \frac{\partial^3}{\partial x^1 \partial x^2 \partial x^k} \hat{V}_{h,\varepsilon} \right|_0 \leq Ne^{NT} \varepsilon^{-2}.
$$

It follows that

$$
\left| \frac{\partial}{\partial t} \hat{V}_{h,\varepsilon}(t,x) - \frac{\partial}{\partial t} \hat{V}_{h,\varepsilon}(s,y) \right| \leq |t-s| \left| \frac{\partial^2}{\partial t^2} \hat{V}_{h,\varepsilon} \right|_0 + N|x-y| \left| \frac{\partial^2}{\partial x^1 \partial t} \hat{V}_{h,\varepsilon} \right|_0
$$

$$
\leq Ne^{NT} |t-s| \varepsilon^{-3} + Ne^{NT} |x-y| \varepsilon^{-2}
$$

$$
= Ne^{NT} \varepsilon^{-2} (|t-s| \varepsilon^{-1} + |x-y|).
$$
The last expression is less than $Ne^{NT}e^{-2(|t-s|^\frac{1}{2} + |x-y|)}$ if $|t-s| \geq \varepsilon^2$. On the other hand, if $|t-s| \leq \varepsilon^2$, then
\[
\left| \frac{\partial}{\partial t} \tilde{V}_{h,\varepsilon}(t, x) - \frac{\partial}{\partial t} \tilde{V}_{h,\varepsilon}(s, y) \right| \leq 2 \left| \frac{\partial}{\partial t} \tilde{V}_{h,\varepsilon} \right|_0 \leq Ne^{NT}e^{-1} \leq Ne^{NT}e^{-2(|t-s|^\frac{1}{2} + |x-y|)}.
\]

Hence the inequality in question holds for all $t, s \in [0, T]$. In the same way one gets that
\[
\left\| \frac{\partial^2}{\partial x^i \partial x^j} \tilde{V}_{h,\varepsilon}(t, x) - \frac{\partial^2}{\partial x^i \partial x^j} \tilde{V}_{h,\varepsilon}(s, y) \right\| \leq Ne^{NT}e^{-2(|t-s|^\frac{1}{2} + |x-y|)}.
\]

This proves the first inequality in the assertions.

To prove the second one, we notice that Lemma 4.5 yields
\[
\left| \tilde{V}_{h,\varepsilon}(t, x) - \tilde{V}_{h,\varepsilon}(t, x) \right| \leq \int_B \left| \tilde{V}_{h,\varepsilon}(t, x) - \tilde{V}_{h,\varepsilon}(t - \varepsilon^2 \tau, x - \varepsilon \zeta) \right| \varphi(\tau, \zeta) d\tau d\zeta \leq 2Ne^{NT}\varepsilon.
\]

We can thus adjust $N$, if needed, to establish the second estimate for all $\varepsilon > 0$.

We can now establish a dynamic programming bound for the mollified value function.

**Lemma 5.2.** Suppose Assumptions 2.7 and 4.2 are satisfied. Then for all $x \in \mathbb{R}^n, t \in [0, T - \varepsilon^2 - h^2]$, and all $\alpha \in U$ we have
\[
\tilde{V}_{h,\varepsilon}(t, x) \leq \rho_{t,t+h^2}^\vartheta \left[ \int_t^{t+h^2} c(s, X_s^{t,x;u,\alpha}) ds + \tilde{V}_{h,\varepsilon}(t + h^2, X_{t+h^2}^{t,x;\alpha}) \right] + Ne^{NT}h^2\varepsilon,
\]
where $N$ is a constant independent on $h, \varepsilon, and T$.

**Proof.** Fixing $\beta = (\tau, \zeta)$ on the right hand side of (21), for every $\alpha \in U$ we obtain the inequality
\[
\tilde{V}_{h,\varepsilon}(t, x) \leq \rho_{t+\varepsilon^2\tau,t+h^2+\varepsilon^2\tau}^\vartheta \left[ \int_{t+\varepsilon^2\tau}^{t+h^2+\varepsilon^2\tau} c(s, X_s^{t+\varepsilon^2\tau,x+\varepsilon\zeta;\alpha}) ds + \tilde{V}_{h,\varepsilon}(t + h^2, X_{t+h^2+\varepsilon^2\tau}^{t+\varepsilon^2\tau,x+\varepsilon\zeta;\alpha} - \varepsilon\zeta) \right].
\]
Since $t \leq T - \varepsilon^2 - h^2$, we can substitute $t - \varepsilon^2 \tau$ for $t$ and $x - \varepsilon \zeta$ for $x$. We obtain
\[
\tilde{V}_{h,\varepsilon}(t - \varepsilon^2 \tau, x - \varepsilon \zeta) \leq \rho_{t,t+h^2}^\vartheta \left[ \int_t^{t+h^2} c(s, X_s^{t,x;u,\alpha}) ds + \tilde{V}_{h,\varepsilon}(t + h^2 - \varepsilon^2 \tau, X_{t+h^2}^{t,x;\alpha} - \varepsilon \zeta) \right].
\]
By virtue of Theorem 2.8 the risk measure $\rho_{t,t+h^2}[\cdot]$ is subadditive, and thus
\[
\tilde{V}_{h,\varepsilon}(t - \varepsilon^2 \tau, x - \varepsilon \zeta) \leq \rho_{t,t+h^2}^\vartheta \left[ \int_t^{t+h^2} c(s, X_s^{t,x;u,\alpha}) ds + \tilde{V}_{h,\varepsilon}(t + h^2, X_{t+h^2}^{t,x;\alpha}) \right] + \rho_{t,t+h^2}^\vartheta \left[ \tilde{V}_{h,\varepsilon}(t + h^2 - \varepsilon^2 \tau, X_{t+h^2}^{t,x;\alpha} - \varepsilon \zeta) - \tilde{V}_{h,\varepsilon}(t + h^2, X_{t+h^2}^{t,x;\alpha}) \right].
\]
(26)
The last term on the right hand side of (26) can be equivalently bounded by using the dual representation of the risk measure $\rho_{t,t+h^2}$. We can thus write the following chain of relations:

\[
\rho_{t,t+h^2}^\rho \left[ \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right] \\
= \sup_{\Gamma \in A_{t,t+h^2}} \mathbb{E}_t \left[ \Gamma \left( \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right) \right] \\
\leq \mathbb{E}_t \left[ \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right] \\
+ \sup_{\Gamma \in A_{t,t+h^2}} \|\Gamma - 1\| \left\| \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right\|.
\]

Owing to Corollary 2.10 and Lemma 5.1(ii), we obtain the estimate

\[
\rho_{t,t+h^2}^\rho \left[ \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right] \\
\leq \mathbb{E}_t \left[ \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right] + N\varepsilon^{NT}h^2\varepsilon.
\]

Substitution for the last term in (26) yields

\[
\hat{V}_{h,\varepsilon}(t - \varepsilon^2\tau, x, \varepsilon) \leq \rho_{t,t+h^2}^\rho \left[ \int_t^{t+h^2} c(s, X_{s+h}^\tau, x, \alpha) \, ds + \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right] \\
+ \mathbb{E}_t \left[ \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right] + N\varepsilon^{NT}h^2\varepsilon. \tag{27}
\]

We now multiply both sides of (27) by $\varphi(\tau, \varepsilon)$ and integrate over $B$. By changing the order of integration in the expected value term of (27) we observe that

\[
\int_B \mathbb{E}_t \left[ \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right] \varphi(\tau, \varepsilon) \, d\tau \, d\varepsilon \\
= \mathbb{E}_t \left[ \int_B \left( \hat{V}_{h,\varepsilon}(t+h^2 - \varepsilon^2\tau, X_{t+h}^\tau, x, \alpha) - \hat{V}_{h,\varepsilon}(t+h^2, X_{t+h}^\tau) \right) \varphi(\tau, \varepsilon) \, d\tau \, d\varepsilon \right] = 0.
\]

Other terms on the right hand side of (27) do not depend on $(\tau, \varepsilon)$ and thus (25) follows.

\section{Accuracy of the Approximation}

We can now investigate the effect of the size of the discretization interval, $h^2$, on the accuracy of the value function approximation. For simplicity of presentation, we write $\sigma^\alpha(s, x)$ for $\sigma(s, x, \alpha)$ and $c^\alpha(s, x)$ for $c(s, x, \alpha)$.
Lemma 6.1. For any $w \in C^1_b([0, T] \times \mathbb{R}^n)$, any $0 \leq t \leq \theta \leq T$, and all $u(\cdot) \in \mathcal{U}$ we have:

$$w(t, x) = \rho^2_{t, \theta} \left[ \int_t^\theta c(s, X_s^{t,x;u}, u_s) \, ds + w(\theta, X_{\theta}^{t,x;u}) - \tilde{\zeta} \right],$$

(28)

where

$$\tilde{\zeta} = \int_t^\theta \left\{ [c^{u_s} + \mathbb{L} u_s w] (s, X_s^{t,x;u}) + g(s, [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u})) \right\} \, ds. \quad (29)$$

Proof. For any $u(\cdot) \in \mathcal{U}$, we apply Itô formula to $w(s, X_s^{t,x;u})$:

$$w(\theta, X_{\theta}^{t,x;u}) - w(t, x) - \int_t^\theta [\mathbb{L} u_s w] (s, X_s^{t,x;u}) \, ds = \int_t^\theta [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u}) \, dW_s.$$  

Subtraction of $\int_t^\theta g(s, [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u})) \, ds$ from both sides and evaluation of the risk on both sides yields

$$\rho^2_{t, \theta} \left[ w(\theta, X_{\theta}^{t,x;u}) - w(t, x) - \int_t^\theta \left( [\mathbb{L} u_s w] (s, X_s^{t,x;u}) + g(s, [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u})) \right) \, ds \right] = \rho^2_{t, \theta} \left[ \int_t^\theta [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u}) \, dW_s - \int_t^\theta g(s, [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u})) \, ds \right]. \quad (30)$$

The risk measure on the right hand side of (30) is the solution of the following backward stochastic differential equation:

$$Y_{t}^{t,x;u} = \int_t^\theta [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u}) \, dW_s - \int_t^\theta g(s, [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u})) \, ds$$

$$+ \int_t^\theta g(s, Z_s^{t,x;u}) \, ds - \int_t^\theta Z_s^{t,x;u} \, dW_s.$$  

Substitution of $Z_s^{t,x;u} = [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u})$ yields $Y_{t}^{t,x;u} = 0$. By the uniqueness of the solution of BSDE, the right hand side of (30) is zero. Using the translation property on the left hand side of (30), we obtain

$$w(t, x) = \rho^2_{t, \theta} \left[ - \int_t^\theta \left( [\mathbb{L} u_s w] (s, X_s^{t,x;u}) + g(s, [D_x w \cdot \sigma^{u_s}](s, X_s^{t,x;u})) \right) \, ds + w(\theta, X_{\theta}^{t,x;u}) \right].$$

This is the same as (28). \hfill \square

The integral in (29) can be bounded by the following lemma.

Lemma 6.2. For all $t \in [0, T - h^2 - \varepsilon^2]$, $x \in \mathbb{R}^n$, and all $\alpha \in \mathcal{U}$, we have

$$[c^\alpha + \mathbb{L} \hat{V}_h \alpha](t, x) + g(t, [D_x \hat{V}_h \alpha \cdot \sigma^\alpha](t, x)) \geq -N e^{NT} \left( \varepsilon + \frac{h}{\varepsilon^2} \right), \quad (31)$$

where the constant $N$ does not depend on $h$, $\varepsilon$, and $T$.  

16
Proof. By Lemma 5.2 for every $\alpha \in U$ we have

$$
\hat{V}_{h,\varepsilon}(t, x) \leq \rho_{t,t+h^2}^2 \left( \int_t^{t+h^2} c(s, X^t_{s}, \alpha) \, ds + \hat{V}_{h,\varepsilon}(t + h^2, X^t_{t+h^2}) \right) + Ne^{NT}h^2\varepsilon.
$$

Using the translation property of $\rho_{t,t+h^2}$, we obtain the inequality:

$$
\rho_{t,t+h^2}^2 \left( \int_t^{t+h^2} c(s, X^t_{s}, \alpha) \, ds + \hat{V}_{h,\varepsilon}(t + h^2, X^t_{t+h^2}) - \hat{V}_{h,\varepsilon}(t, x) \right) \geq -Ne^{NT}h^2\varepsilon.
$$

Since $\hat{V}_{h,\varepsilon} \in C_b^{1,2}([t, T - \varepsilon^2] \times \mathbb{R}^n)$, we can evaluate the difference $\hat{V}_{h,\varepsilon}(t + h^2, X^t_{t+h^2}) - \hat{V}_{h,\varepsilon}(t, x)$ by Itô formula between $t$ and $t + h^2$:

$$
\hat{V}_{h,\varepsilon}(t+h^2, X^t_{t+h^2}) - \hat{V}_{h,\varepsilon}(t, x) = \int_t^{t+h^2} \left[ \mathbb{L}^\alpha \hat{V}_{h,\varepsilon}(s, X^t_{s}) \right] \, ds + \int_t^{t+h^2} \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](s, X^t_{s}) \, dW_s.
$$

Substitution into the previous inequality yields:

$$
\rho_{t,t+h^2}^2 \left( \int_t^{t+h^2} \left[ c^\alpha + \mathbb{L}^\alpha \hat{V}_{h,\varepsilon} \right](s, X^t_{s}) \, ds + \int_t^{t+h^2} \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](s, X^t_{s}) \, dW_s \right) \geq -Ne^{NT}h^2\varepsilon.
$$

The evaluation of the risk measure amounts to solving the following backward stochastic differential equation:

$$
Y_t = \int_t^{t+h^2} \left[ c^\alpha + \mathbb{L}^\alpha \hat{V}_{h,\varepsilon} \right](s, X^t_{s}) \, ds + \int_t^{t+h^2} \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](s, X^t_{s}) \, dW_s
$$

$$
+ \int_t^{t+h^2} g(s, Z_s) \, ds - \int_t^{t+h^2} Z_s \, dW_s.
$$

The equation has a unique solution:

$$
Z_s = \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](s, X^t_{s}), \quad t \leq s \leq t + h^2,
$$

$$
Y_t = \int_t^{t+h^2} \left\{ \left[ c^\alpha + \mathbb{L}^\alpha \hat{V}_{h,\varepsilon} \right](s, X^t_{s}) + g(s, \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](s, X^t_{s})) \right\} \, ds.
$$

We can thus write the inequality

$$
Y_t \leq h^2 \left( \left[ c^\alpha + \mathbb{L}^\alpha \hat{V}_{h,\varepsilon} \right](t, x) + g(t, \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](t, x)) \right)
$$

$$
+ h^2 \max_{t \leq s \leq t+h^2} E_t \left\{ \left[ c^\alpha + \mathbb{L}^\alpha \hat{V}_{h,\varepsilon} \right](s, X^t_{s}) - \left[ c^\alpha + \mathbb{L}^\alpha \hat{V}_{h,\varepsilon} \right](t, x) \right\}
$$

$$
+ h^2 \max_{t \leq s \leq t+h^2} E_t \left\{ g(s, \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](s, X^t_{s})) - g(t, \left[ D_x \hat{V}_{h,\varepsilon} \cdot \sigma \right](t, x)) \right\}.
$$

The last two terms can be bounded by $Ne^{NT}h^2/\varepsilon^2$, owing to Assumption 4.2 and Lemma 5.1. Combining this inequality with (32) and dividing by $h^2$, we conclude that for all $\alpha \in U$ the estimate (31) is true. 

17
We are now ready to prove the main theorem of this section.

**Theorem 6.3.** Suppose Assumptions 2.7 and 4.2 are satisfied. Then for any $t \in [0, T]$, $x \in \mathbb{R}^n$, and $h \in (0, 1]$, we have

$$|V(t, x) - V_h(t, x)| \leq Ne^{NT}h^\frac{1}{3},$$

where the constant $N$ depends only on $(K, n, d)$.

**Proof.** We set $\varepsilon = h^\frac{1}{3}$ and organize the proof in three steps.

**Step 1:** If $t \in [T - h^2 - \varepsilon^2, T]$, then for any $u(\cdot)$ and some constant $C$ we have,

$$\left| V^u(t, x) - \Phi(x) \right| \leq \rho_{t, T} \left( \int_t^T |c(s, X_s^{t,x,u}, u_s)| ds + \left| \Phi(X_{T-s}^{t,x,u}) - \Phi(x) \right| \right) \leq K(h^2 + \varepsilon^2) + K\mathbb{E}_{t,x} \left[ |X_{T-s}^{t,x,u} - x| \right] \leq K(1 + C)(h^2 + \varepsilon^2) \leq 2K(1 + C)h^\frac{7}{3}.$$

In the above estimate we also used the fact that the solution of the forward–backward system (6)–(9) is Lipschitz in the initial condition [26]. The same reasoning works for $V^u_h$, and thus

$$\left| V^u_h(s, x) - \Phi(x) \right| \leq 2K(1 + C)h^\frac{7}{3}.$$

We can, therefore, for some constant $N$ write the inequality

$$\left| V^u(t, x) - V^u_h(t, x) \right| \leq Ne^{NT}h^\frac{1}{3}.$$

The optimization over $u$ will not make it worse, and thus our assertion is true for these $t$.

**Step 2:** Consider $t \in [0, T - h^2 - \varepsilon^2]$. By Lemma 6.1, for all $u(\cdot) \in \mathcal{U}$ on $[t, T - h^2 - \varepsilon^2]$, we have

$$\hat{V}_{h, \varepsilon}(t, x) \leq \rho_{t, T - h^2 - \varepsilon^2} \left( \int_t^{T - h^2 - \varepsilon^2} c(s, X_s^{t,x,u}, u_s) ds + \hat{V}_{h, \varepsilon}(T - h^2 - \varepsilon^2, X_{T-h^2-\varepsilon^2}^{t,x,u}) - \zeta \right),$$

where, owing to Lemma 6.2,

$$\zeta = \int_t^{T - h^2 - \varepsilon^2} \left( \left[ c_{us} + L u_s \hat{V}_{h, \varepsilon} \right](s, x) + g(s, [\partial_x \hat{V}_{h, \varepsilon} \cdot \sigma^u]_{s,x}) \right) ds \geq -Ne^{NT} \left( \varepsilon + \frac{h}{\varepsilon^2} \right).$$

These relations, the monotonicity of the risk measure, and Lemmas 4.5 and 5.1 imply the estimate

$$V_h(t, x) \leq \rho_{t, T - h^2 - \varepsilon^2} \left( \int_t^{T - h^2 - \varepsilon^2} c(s, X_s^{t,x,u}, u_s) ds + V_h(T - h^2 - \varepsilon^2, X_{T-h^2-\varepsilon^2}^{t,x,u}) \right) + Ne^{NT} \left( \varepsilon + \frac{h}{\varepsilon^2} \right) + 4Ne^{NT}\varepsilon.$$
In view of the inequality established in Step 1, using $\varepsilon = h^{\frac{1}{3}}$, and redefining $N$ appropriately, we obtain the following inequality:

$$V_h(t, x) \leq \rho^g_{t, T - h^2 - \varepsilon^2} \left( \int_t^{T - h^2 - \varepsilon^2} c(s, X_{s, t}^{\varepsilon}, u_s) \, ds + V(T - h^2 - \varepsilon^2, X_{T - h^2 - \varepsilon^2}^{\varepsilon}) \right) + Ne^{NT} h^{\frac{1}{3}},$$

(33)

Step 3: We apply the dynamic programming equation (12) to the right hand side of (33) to conclude that

$$V_h(t, x) \leq \inf_{u(\cdot) \in U} \rho^g_{t, T - h^2 - \varepsilon^2} \left( \int_t^{T - h^2 - \varepsilon^2} c(s, X_{s, t}^{\varepsilon}, u_s) \, ds 
+ V(T - h^2 - \varepsilon^2, X_{T - h^2 - \varepsilon^2}^{\varepsilon}) \right) + Ne^{NT} h^{\frac{1}{3}} = V(t, x) + Ne^{NT} h^{\frac{1}{3}},$$

as required.

References

[1] F. Antonelli. Backward-forward stochastic differential equations. *Annals of Applied Probability*, 3:777 – 793, 1993.

[2] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Thinking coherently. *RISK*, 10:68–71, 1997.

[3] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9:203–228, 1999.

[4] P. Barrieu and N. El Karoui. Optimal derivative design under dynamic risk measures. *Mathematics of Finance, Contemporary Mathematics*, 351:13–26, 2004.

[5] P. Barrieu and N. El Karoui. Pricing, hedging and optimally designing derivatives via minimization of risk measures. *Volume on Indifference Pricing, Princeton University Press*, 2009.

[6] J. Bion-Nadal. Dynamic risk measures: time consistency and risk measures from bmo martingales. *Finance and Stochastic*, 12:219–244, 2008.

[7] P. Briand, B. Delyon, and J. Mémin. On the robustness of backward stochastic differential equations. *Stochastic Processes and their Applications*, 97:229–253, 2002.

[8] Ö. Çavuş and A. Ruszczyński. Risk-averse control of undiscounted transient Markov models. *SIAM Journal on Control and Optimization*, 52(6):3935–3966, 2014.

[9] P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, 11:57–106, 2006.

[10] P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(1):137–162, 2011.
[11] F. Coquet, Y. Hu, J. Mémin, and S. Peng. Filtration-consistent nonlinear expectations and related g-expectations. *Probability Theory and Related Fields*, 123(1):1–27, 2002.

[12] F. Delbaen, S. Peng, and E. Rosazza Gianin. Representation of the penalty term of dynamic concave utilities. *Finance and Stochastics*, 14(3):449–472, 2010.

[13] K. Detlefsen and G. Scandolo. Conditional and dynamic convex risk measures. *Finance and Stochastic*, 9:539–561, 2005.

[14] J. Fan and A. Ruszczynski. Dynamic risk measures for finite-state partially observable Markov decision problems. In *Proceedings of the Conference on Control and its Applications*, pages 153–158. SIAM, 2015.

[15] W. Flemming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*. Springer, 2006.

[16] H. Föllmer and I. Penner. Convex risk measures and the dynamics of their penalty functions. *Statistics and Decisions*, 24:61–96, 2006.

[17] H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastic*, 6:429–447, 2002.

[18] H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time, 2nd ed*. De Gruyter Berlin, 2004.

[19] M. Frittelli and G. Scandolo. Risk measures and capital requirements for processes. *Math. Finance*, 16:589–612, 2006.

[20] M. Frittelli and E. Rosazza Gianin. Putting order in risk measures. *Journal of Banking and Finance*, 26:1473–1486, 2002.

[21] M. Frittelli and E. Rosazza Gianin. Dynamic convex risk measures. pages 227–248. Wiley, New York, 2004.

[22] N. V. Krylov. Approximating value funtions for controlled degenerate diffusion processes by using piece-wise constant polices. *Electronic Journal Of Probability*, 4:1–19, 1999.

[23] N. V. Krylov. *Controlled Diffusion Process (Stochastic Modelling and Applied Probability)*. Springer, 2008.

[24] R. J. A. Laeven and M. Stadje. Robust portfolio choice and indifference valuation. *Mathematics of Operations Research*, 39(4):1109–1141, 2014.

[25] J. Li and Q. Wei. Optimal control problems of fully coupled FBSDEs and viscosity solutions of Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optimization*, 52(3):1622–1662, 2014.
[26] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their Applications*. Springer, 2007.

[27] B. Øksendal and A. Sulem. Maximum principles for optimal control of forward-backward stochastic differential equations with jumps. *SIAM Journal on Control and Optimization*, 48(5):2945–2976, 2009.

[28] B. Øksendal and A. Sulem. Forward–backward stochastic differential games and stochastic control under model uncertainty. *Journal of Optimization Theory and Applications*, 161(1):22–55, 2014.

[29] E. Pardoux and S. Peng. Adapted solutions of backward stochastic differential equation. *System and Control Letters*, 14:55–61, 1990.

[30] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic Differential Equations and Their Applications*, 176:200–217, 1992.

[31] E. Pardoux and S. Tang. Forward-backward stochastic differential equations and quasilinear parabolic pdes. *Probability Theory and Related Fields*, 114:123 – 150, 1999.

[32] S. Peng. Backward SDE and related $g$-expectation. *Backward Stochastic Differential Equations, Pitman Research Notes Mathematics*, 364:141–159, 1997.

[33] S. Peng. Nonlinear expectations, nonlinear evaluations and risk measures. *Lecture Notes in Mathematics, Springer*, 2004.

[34] S. Peng and Z. Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization*, 37(3):825–843, 1999.

[35] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Application (Stochastic Modelling and Applied Probability)*. Springer, 2010.

[36] M.-C. Quenez and A. Sulem. Bsdes with jumps, optimization and applications to dynamic risk measures. *Stochastic Processes and their Applications*, 123(8):3328–3357, 2013.

[37] F. Riedel. Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112:185–200, 2004.

[38] E. Rosazza Gianin. Risk measures via g-expectations. *Insurance: Mathematics and Economics*, 39(1):19–34, 2006.

[39] A. Ruszczyński. Risk-averse dynamic programming for Markov decision process. *Mathematical Programming Series B*, 125:235–261, 2010.

[40] A. Ruszczyński and A. Shapiro. Conditional risk mappings. *Mathematics of Operations Research*, 31:544–561, 2006.
[41] A. Ruszczyński and A. Shapiro. Optimization of convex risk functions. Mathematics of Operations Research, 31:433–452, 2006.

[42] A. Ruszczyński, A. Shapiro, and D. Dentcheva. Lectures on Stochastic Programming. Modeling and Theory. SIAM-Society for Industrial and Applied Mathematics, 2009.

[43] A. Ruszczyński and J. Yao. A risk-averse analog of the Hamilton-Jacobi-Bellman equation. In Proceedings of the SIAM Conference on Control and its Applications, Paris, pages 462–468, 2015.

[44] M. Stadje. Extending dynamic convex risk measures from discrete time to continuous time: a convergence approach. Insurance: Mathematics and Economics, 47:391–404, 2010.

[45] J. Yong. Finding adapted solutions of forward-backward stochastic differential equations: method of continuation. Probability Theory and Related Fields, Springer, 1997.

[46] J. Zhang. The well posedness of FBSDE. Discrete and Continuous Dynamical Systerms Series B, 6:927 – 940, 2006.

[47] X. Zhou and J. Yong. Stochastic Control - Hamiltonian Systems and HJB Equations. Springer, 1998.