Stochastic Optimal Control Subject to Variational Norm Uncertainty: Viscosity Subsolution for Generalized HJB Inequality

Farzad Rezaei, Charalampos D. Charalambous and Nasir U. Ahmed

Abstract—This paper is concerned with optimization of stochastic uncertain systems, when systems are described by measures and the pay-off by a linear functional on the space of measures, on general abstract spaces. Robustness is formulated as a minimax game, in which the control seeks to minimize the pay-off over the admissible controls while the measure aims at maximizing the pay-off over the total variational distance uncertainty constraint between the uncertain and nominal measures. This paper is a continuation of the abstract results in [1], where existence of the maximizing measure over the total variational distance constraint is established, while the maximizing payoff is shown to be equivalent to an optimization of a payoff which is a linear combination of $L_1$ and $L_{\infty}$ norms. The maximizing measure is constructed from a convex combination of a sequence of tilted measures and the nominal measure. Here emphasis is geared towards the application of the abstract results to uncertain continuous-time controlled stochastic differential equations, in which the control seeks to minimize the pay-off while the measure seeks to maximize it over the total variational distance constraint. The maximization over the total variational distance constraint is resolved resulting in an equivalent payoff which is a non-linear functional of the nominal measure of non-standard form. The minimization over the admissible controls of the non-linear functional is addressed by deriving a HJB inequality and viscosity subsolution. Throughout the paper the formulation and conclusions are related to previous work found in the literature.

I. INTRODUCTION

In the parlance of robustness minimax strategies are employed to deal with signal processing, estimation, decision and control subject to parametric and non-parametric uncertainty of system models. In an abstract measure theoretic setting in which systems are described by measures, non-parametric uncertainty models attempt to quantify the set of all measures induced by the uncertain system. A general quantitative definition of an uncertainty model presupposes a certain smoothness or continuity with respect to small perturbations, while it is appropriate to describe uncertainty on general abstract spaces. Both requirements are met when distance metrics are employed to codify the uncertainty set. In applications of statistics and applied probability in which systems are described by measures and pay-off’s by linear functionals on the space of measures (e.g., error or energy criteria), a natural description of uncertainty is the total variational distance between measures.

The objectives of this paper are the following.

Employ the definition of uncertainty via the total variational distance between measures on general abstract spaces introduced in [1] and the results therein to

- Investigate stochastic optimal control of systems governed by stochastic differential equations subject to total variational distance uncertainty.

At the abstract level, a general framework of uncertainty description via the total variation distance is put forward, in which the fundamental results are described. At this level, systems are represented by measures on abstract spaces, pay-off’s by linear functionals on the space of measures, and uncertainty by a set described by the total variation distance (uncertainty ball with respect to total variational norm) centered at the nominal measure having a pre-specified radius. At the abstract level robustness is formulated via the minimax value of the pay-off, where the maximization is over the measures which satisfy the total variational distance constraint, while the minimization is over the admissible controls. It is shown in [1] that the maximization of the pay-off over the total variational distance uncertainty exists and it is equivalent to a pay-off which is a linear combination of $L_1$ and $L_{\infty}$ norms. Further, the maximizing measure is constructed as a convex combination of a sequence of tilted measures and the nominal measure.

In this paper, the theory developed in [1] is applied to formulate and solve minimax stochastic uncertain controlled systems described by continuous-time nonlinear stochastic controlled differential equations, in which the control seeks to minimize the pay-off while the measure seeks to maximize it over the total variational distance constraint, described by a ball with respect to total variation norm centered at the nominal measure. The main objective of this part is to provide a solution to the minimax game via dynamic programming and HJB inequality. It turns out that once the maximizing measure is found and substituted into the pay-off the equivalent optimization problem to be solved is a stochastic optimal control problem. There is however, a fundamental difference from the classical pay-off of stochastic control problems treated in the literature in that the pay-off is a non-linear functional of measure induced by the stochastic system, contrary to the classical which is a linear functional. The main contributions of this part is to show that the value function satisfies a HJB inequality which admits a viscosity subsolution. The point to be made here is that while stochastic optimal control theory is well developed when the

This work was supported by the Cyprus Research Promotion Foundation project ARTEMIS-0506/20 and a University of Cyprus Research grant.

F. Rezaei and N. U. Ahmed are with School of Information Technology and Engineering, University of Ottawa, 800 King Edward Ave., Ottawa, Canada. E-mail: frezaei@site.uottawa.ca, ahmed@site.uottawa.ca

C.D. Charalambous is with the Department of Electrical Engineering, University of Cyprus, Nicosia, Cyprus. E-mail: chadcha@ucy.ac.cy
pay-off is a linear functional of the measure induced by the system, such as, integral or exponential-of-integral payoff functionals \[10, 11, 12\], the resulting pay-off emerging from the maximization solution is nonlinear with respect to the measure induced by the control.

The rest of the paper is organized as follows. In Section II various models of uncertainty are introduced and their relation to total variational distance is described. In Sections III the abstract formulation is introduced while in Section III-A the solution of the maximization problem over the variational norm constraint set is presented. In section IV, the abstract setup is applied to stochastic fully observable continuous-time uncertain controlled systems. In Section IV-B, the new type of principle of optimality associated with the non-linear pay-off is derived.

II. MOTIVATION

Throughout the paper the mathematical formulation as well as the results obtained are related to other existing models of uncertainty as well as to other optimization problems found in the literature. Below, a summary of the uncertainty models and their relations to total variation distance model is discussed.

Let \((\Sigma, d_\Sigma)\) denote a complete separable metric space (a Polish space), and \((\Sigma, B(\Sigma))\) the corresponding measurable space, in which \(B(\Sigma)\) is the \(\sigma\)-algebra generated by open sets in \(\Sigma\). Let \(M_1(\Sigma)\) denote space of countably additive probability measures on \((\Sigma, B(\Sigma))\).

The relative entropy of \(\nu \in M_1(\Sigma)\) with respect to \(\mu \in M_1(\Sigma)\) is a mapping \(H(\cdot|\cdot) : M_1(\Sigma) \times M_1(\Sigma) \to [0, \infty]\) defined by

\[
H(\nu|\mu) = \begin{cases} 
\int_\Sigma \log\left(\frac{d\nu}{d\mu}\right)d\mu, & \text{if } \nu << \mu \text{ and } \log\frac{d\nu}{d\mu} \in L_1(\nu) \\
\infty & \text{otherwise}
\end{cases}
\]

Here \(\nu << \mu\) is the notation often used to denote that measure \(\nu \in M_1(\Sigma)\) is absolutely continuous with respect to measure \(\mu \in M_1(\Sigma)\).\(^1\)

Relative Entropy Uncertainty. Given a known or nominal probability measure \(\mu \in M_1(\Sigma)\) the uncertainty set based on relative entropy is defined by

\[
A_R(\mu) \triangleq \left\{ \nu \in M_1(\Sigma) : H(\nu|\mu) \leq \bar{R} \right\}
\]

where \(\bar{R} \in (0, \infty)\) (when \(\bar{R} = 0\) then \(\nu = \mu, \nu - a.s\)). Over the last few years, the relative entropy uncertainty model has received particular attention due to its simplicity and its connection to risk sensitive pay-off, minimax games, and large deviations \[3, 4, 5, 6, 7, 8, 9\]. Unfortunately, relative entropy uncertainty modeling has two disadvantages. 1) it does not define a true metric on the space of measures; 2) relative entropy between two measures is not defined if the measures are absolutely continuous. The latter rules out the possibility of measures \(\nu \in M_1(\Sigma)\) and \(\mu \in M_1(\Sigma)\) to be initially defined on different spaces (e.g., one being defined on a higher dimension space than the other measure). It is one of the main disadvantages of employing relative entropy in the context of uncertainty modelling for stochastic controlled diffusions (or stochastic differential equations) \[3, 4, 5, 7, 8, 9\]. Specifically, by invoking a change of measure it can be shown that relative entropy allows uncertainty in the drift coefficient of stochastic controlled diffusions, but not in the diffusion coefficient, because the latter kind of uncertainty leads to measures which are not absolutely continuous with respect to the nominal measure \[3\].

Motivated by the above issues, in \[1\] an uncertainty model based on the total variational distance defined on the space of measures is introduced. The uncertainty set is described by a ball with respect to the total variational norm, centered at the nominal measure having positive radius.

Total Variational Distance Uncertainty. Given a known or nominal probability measure \(\mu \in M_1(\Sigma)\) the uncertainty set based on total variational distance is defined by

\[
B_R(\mu) = \left\{ \nu \in M_1(\Sigma) : ||\nu - \mu|| \leq R \right\}
\]

where \(R \in (0, \infty)\). The total variational distance \(2\) on \(M_1(\Sigma) \times M_1(\Sigma)\) is defined by

\[
||\alpha - \beta|| = \sup_{P \in \mathcal{P}(\Sigma)} \sum_{F_i \in P} |\alpha(F_i) - \beta(F_i)|, \quad \alpha, \beta \in M_1(\Sigma)
\]

where \(\mathcal{P}(\Sigma)\) denotes the collection of all finite partitions of \(\Sigma\). Note that the metric induced by the total variational norm satisfied the properties of a metric, and does not require absolute continuity of measures when defining the uncertainty ball (i.e., singular measures are admissible), that is, the measures need not be defined on the same space. It can very well be the case that \(\tilde{\mu} \in M_1(\Sigma), \tilde{\Sigma} \subset \Sigma\) and \(\mu \in M_1(\Sigma)\) is the extension of \(\tilde{\mu}\) on \(\Sigma\).

Additionally, since \(M_1(\Sigma)\) are probability measures then it follows that the radius of uncertainty belongs to the restricted set \(R \in (0, 2]\). Clearly, the total variational distance uncertainty set larger than the relative entropy uncertainty set. This can be concluded from Pinsker's inequality \[2\] as well.

\[
||\nu - \mu||^2 \leq 2H(\nu|\mu),
\]

\[
\nu, \mu \in M_1(\Sigma), \quad \text{if } \nu << \mu, \quad \log\frac{d\nu}{d\mu} \in L_1(\nu)
\]

Hence, even for those measures \(\nu << \mu, \log\frac{d\nu}{d\mu} \in L_1(\nu)\) the uncertainty set described by relative entropy is a subset of the much larger total variation distance uncertainty set, that is, \(A_{R^2}(\mu) \subset B_R(\mu)\). In the parlance of stochastic differential equations the total variational distance covers the case when both drift and diffusion coefficients of the stochastic differential equations are uncertain.


\[1\] If \(\mu(A) = 0\) for some \(A \in B(\Sigma)\) then \(\nu(A) = 0\).
are absolutely continuous with respect to a fixed measure \( \sigma \in \mathcal{M}_1(\Sigma) \) (e.g., \( \mu \ll \sigma, \nu \ll \sigma \)). Under these conditions it can be shown that total variation distance reduces to \( L_1(\sigma) \) distance as follows.

\[
C_R(\mu) \triangleq \left\{ \nu \in \mathcal{M}_1(\Sigma); ||\nu - \mu|| \leq R \right\} = \left\{ \varphi \in L_1(\sigma); \int_\Sigma |\varphi(x) - \psi(x)|\sigma(dx) \leq R \right\} \equiv C_R(\varphi)
\]

where existence of \( \varphi = \frac{d\nu}{d\sigma} \in L_1(\sigma), \psi = \frac{d\mu}{d\sigma} \in L_1(\sigma) \) follows from the absolute continuity \( \mu \ll \sigma, \nu \ll \sigma \). Hence, the \( L_1(\sigma) \) distance set \( C_R(\varphi) \) is a smaller set that the total variation distance set \( B_R(\mu) \). Robustness via \( L_1 \) distance uncertainty on the space of spectral densities is investigated in the context of Wiener -Kolmogorov theory in an estimation and decision framework in [17], [18]. Although in the current paper deals with stochastic control applications, a similar formulation applies to estimation problems, and to minimax control and estimation when the nominal system and uncertainty set are described by spectral measures.

### III. Abstract Formulation

Let \((\Sigma, d_\Sigma)\) denote a complete separable metric space (a Polish space), and \((\Sigma, \mathcal{B}(\Sigma))\) the corresponding measurable space, in which \( \mathcal{B}(\Sigma) \) is the \( \sigma \)-algebra generated by open sets in \( \Sigma \).

Let \( \mathcal{X} \) denote \( BC(\Sigma) \) denote the Banach space of bounded continuous functions on \( \Sigma \) equipped with the sup-norm. It is known that the dual space \([16]\) \( \mathcal{X}^\ast \) is isometrically isomorphic to \( \mathcal{M}_{rba}(\Sigma) \), the Banach space of finitely additive finite signed regular measures on \((\Sigma, \mathcal{B}(\Sigma))\). Note that when \( \Sigma \) is compact then \( \mathcal{X}^\ast \) is isometrically isomorphic to \( \mathcal{M}_1(\Sigma) \), the Banach space of countably additive probability measures on \((\Sigma, \mathcal{B}(\Sigma))\).

At the abstract level, systems are represented by measures \( \theta \in \mathcal{M}_1(\Sigma) \) induced by the underlying random processes, which are defined on an appropriate Polish space. Similarly, controls denoted by \( u \), are defined on a subset \( U_0 \) of an appropriate Polish space \((U, d_U)\), while we choose a suitable subset \( U_{ad} \subset U_0 \) for the class of admissible controls. The pay-off is represented by a linear functional on the space of probability measures \( \mathcal{M}_1(\Sigma) \).

**Nominal System.** The nominal system is defined as follows. By choosing a control policy \( u \in U_{ad} \) for the nominal system (which is perfectly known), then the nominal system induces a nominal probability measure \( \mu_n \in \mathcal{M}_1(\Sigma) \).

**Uncertain System.** For a given \( u \in U_{ad} \), let \( M(u) \subset \mathcal{M}_1(\Sigma) \) denote the set of probability measures induced by the perturbed system while control \( u \in U_{ad} \) is applied. We assume that in the absence of perturbation the set \( M(u) \) reduces to the singleton \( \{\mu_n\} \). The perturbed system or uncertain system \( \nu^u \in M(u) \) is further restricted to the following constraint described by the variational norm.

\[
M_r(u) \triangleq B_R^\mu(\mu) = \left\{ \nu \in M(u); ||\nu - \mu|| \leq R \right\}
\]

**Mini-Max Optimization.** Let \( \ell^u: \Sigma \to \mathbb{R} \) be a real-valued bounded non-negative measurable function. The uncertainty tries to maximize the average pay-off functional denoted by \( \int_\Sigma \ell^u(x)\nu(dx) \) over the set \( M_r(u) \) for a given \( u \in U_{ad} \).

The effect of uncertainty leads to the following maximization problem:

\[
\sup_{\nu \in M_r(u)} \int_\Sigma \ell^u(x)\nu(dx) = \sup_{\xi \in \mathcal{M}^r_\mathcal{rb}(\Sigma)} \int_\Sigma \ell(x)\xi(dx) + E_\mu(\ell) \leq R||\ell||_\infty + E_\mu(\ell) \quad (III.4)
\]

for every control \( u \in U_{ad} \).

The designer on the other hand, tries to choose a control policy to minimize the worst case average pay-off. This gives rise to the min-max problem

\[
\inf_{u \in U_{ad}} \sup_{\nu \in M_r(u)} \int_\Sigma \ell^u(x)\nu(dx) \quad (III.1)
\]

As a first step, we present the existence of a \( \nu^\ast \in M_r(u) \) at which the supremum in (III.1) is attained. Subsequently, we present an explicit characterization of this measure.

**A. Characterization of the Maximizing Measure**

In this section we drop the dependence on the control \( u \) of the various measures and functions. Suppose \( \ell \) is a non negative element in \( BC(\Sigma) \). Clearly, \( \mathcal{M}_1(\Sigma) \subset \mathcal{M}_{rba}(\Sigma) \). Let \( \mu \in \mathcal{M}_1(\Sigma) \) be a given probability measure referred to as the nominal measure. Define the uncertainty set by

\[
B_R(\mu) \triangleq \left\{ \nu \in \mathcal{M}_1(\Sigma); ||\nu - \mu|| \leq R \right\}
\]

The objective is to find the worst case (supremum) of average pay-off over the uncertain set \( B_R(\mu) \). The average pay-off is defined as a linear functional acting on \( \ell \in BC(\Sigma) \), i.e., \( \int_\Sigma \ell(x)\nu(dx) \), where \( \nu \in B_R(\mu) \). Hence the problem is the following

\[
J(\ell, R) = \sup_{\nu \in B_R(\mu)} \int_\Sigma \ell(x)\nu(dx), \quad \mu \in \mathcal{M}_1(\Sigma) \quad (III.2)
\]

The set \( B_R(\mu) \) is weak*-compact, while the pay-off is weak* continuous. Hence, there exists a maximizing measure in \( B_R(\mu) \). The optimization in (III.2) is solved by appealing to the Hahn-Banach theorem [13]. Since \( \ell \in BC(\Sigma) \) is fixed, then there exists \( \eta \in (BC(\Sigma))^\ast = \mathcal{M}_{rba}(\Sigma) \) such that

\[
\eta(\ell) = \int_\Sigma \ell d\eta = ||\ell||_\infty, \quad \text{with} \quad ||\eta|| = 1 \quad (III.3)
\]

Define \( \xi \triangleq \nu - \mu \in \mathcal{M}_{rba}(\Sigma) \). Then from (III.2) and the above result we have the following.

\[
\sup_{\nu \in B_R(\mu)} \int_\Sigma \ell(x)\nu(dx) = \sup_{\xi \in \mathcal{M}(\mathcal{M}_{rba}(\Sigma))} \int_\Sigma \ell(x)\xi(dx) + E_\mu(\ell) \leq R||\ell||_\infty + E_\mu(\ell) \quad (III.4)
\]
where $B_R (\mathcal{M}_{rba}(\Sigma)) = \{ \eta \in \mathcal{M}_{rba}(\Sigma); ||\eta|| \leq R \}$. The supremum on the right hand side of (III.4) is attained by a signed measure $\xi^* \in \mathcal{M}_{rba}(\Sigma)$, having the property $||\xi^*|| = R$. Clearly, if $\nu^* = \xi^* + \mu$ is a probability measure, then the upperbound in (III.4) is attained by $\nu^* \in \mathcal{M}_1(\Sigma)$. Therefore, it remains to establish that $\nu^*$ is a probability measure and $\nu^* \in B_R(\mu)$.

As the first step, it is shown that $\nu^*$ is non-negative.

**Lemma 3.1:** Suppose $\ell \in BC(\Sigma)$ is non-negative. The maximizing measure $\nu^* = \xi^* + \mu$, where $\xi^* \in \mathcal{M}_{rba}(\Sigma)$, $\mu \in \mathcal{M}_{rba}(\Sigma)$, $||\xi^*|| = R$ in (III.4) is a non-negative measure.

**Proof.** See [1].

The maximizing measure $\nu^*$ is not unique. The next lemma is crucial in the characterization of the class of maximizing measures. Assume that an appropriate metric such as $d(\cdot, \cdot)$ is defined on the space $\Sigma$.

**Lemma 3.2:** Suppose $\ell : \Sigma \to \mathbb{R}$ is a bounded non-negative measurable function, and $\eta$ is a finitely additive non-negative finite measure defined on $(\Sigma, B(\Sigma))$.

Assume
\[ \text{i) The total weight of the measure } \eta \text{ is not concentrated on any bounded measurable subset of } \Sigma; \]
\[ \text{ii) The support of } \eta \text{ contains the point at which } \ell \text{ attains its maximum.} \]

Then
\[
\sup_{s>0} \int_{\Sigma} \ell(x)e^{s\ell(x)}\eta(dx) = ||\ell||_{\infty}
\]

**Proof.** See [1].

**Remark 3.3:** In Lemma 3.2, $\int_{\Sigma} \ell(x)e^{s\ell(x)}\eta(dx)$ is a monotone non-decreasing function of $s$, hence
\[
\lim_{s\to+\infty} \frac{\int_{\Sigma} \ell(x)e^{s\ell(x)}\eta(dx)}{\int_{\Sigma} e^{s\ell(x)}\eta(dx)} = ||\ell||_{\infty}
\]

Next, we state the main theorem which characterizes the maximizing measure as a convex combination of two probability measures.

**Theorem 3.4:** Under the assumptions of Lemma 3.2, there exists a family of probability measures which attain the supremum in (III.2) given by
\[
\nu^*(E) = \frac{\beta}{\beta + 1} \int_{\Sigma} e^{s\ell(t)}\eta(dx) + \frac{1}{\beta + 1} \mu(E)
\]

where $E \in B(\Sigma), \beta \in (2, \infty)$ is arbitrary, and $\eta$ is an arbitrary finite non-negative finitely additive measure defined on $(\Sigma, B(\Sigma))$. Moreover, $\beta$ and $\eta$ are chosen such that $||\nu^* - \mu|| = R$.

**Proof.** See [1].

Clearly, $\nu^*(dx)$ is a convex combination of the tilted measure $e^{\sigma \ell(t)}\eta(dx)$ and $\mu(dx)$. Moreover, the initial optimization problem is also a convex combination of $L_1$ and $L_\infty$ optimization problems in view of
\[
\sup_{\nu \in B_R(\mu)} \int_{\Sigma} \ell(x)\nu(dx) = R ||\ell||_{\infty} + E_u(\ell)
\]
\[
= \frac{\beta}{\beta + 1} \int_{\Sigma} \ell(x)e^{\sigma \ell(x)}\eta(dx) + \frac{1}{\beta + 1} \int_{\Sigma} \ell(x)\mu(dx)
\]

The rest of the paper deals with the application of the above results to stochastic systems described by controlled stochastic differential equations.

IV. FULLY OBSERVED UNCERTAIN CONTROL SYSTEMS

Let $M(u) \subset \mathcal{M}_{rba}(\Sigma)$ denote the set of probability measures induced by the perturbed system under the action of control $u$. In the subsequent sections, the measurable space $(\Sigma, B(\Sigma))$ is defined on the space of continuous sample paths.

A. Problem Formulation

Let $\{x(t)\}_{t \geq 0}$ denote the state process subject to control, $\{u(t)\}_{t \geq 0}$ denote the control process, both defined for a fixed finite time interval $[0, T]$. For each $u \in U_{ad}$ (set of admissible controls to be defined shortly) the nominal state process, giving rise to a nominal probability measure $P \in M(u)$, is governed by the following stochastic differential equations.

\begin{equation}
\begin{aligned}
(\Omega, \mathcal{F}, P) : \quad & dx(t) = b(t, x(t), u(t))dt \\
& + \sigma(t, x(t), u(t))dW(t), \\
x(s) = \xi
\end{aligned}
\end{equation}

(IV.5)

where $W(\cdot)$ is an $m$-dimensional Standard Brownian motion.

The set of admissible controls is defined as follows.

**Definition 4.1:** The set of admissible controls denoted by $U_w[s, T]$ is the set of all $5$-tuples $(\Omega, \mathcal{F}, P, w(\cdot), u(\cdot))$ which satisfy the following conditions.

1) $(\Omega, \mathcal{F}, P)$ is a complete probability space.
2) $\{w(t)\}_{t \geq s}$ as an $\mathbb{R}^m$-dimensional standard Brownian motion defined on $(\Omega, \mathcal{F}, P)$ over the time interval $[s, T]$, with $w(s) = 0, P - a.s.$, and $\mathcal{F}_t$ is generated by $\sigma(\cdot; s \leq r \leq t)$ augmented by the $P$-null sets in $\mathcal{F}$.
3) $u : [s, T] \times \Omega \to U$ is an $\{\mathcal{F}_s, t \geq s\}$-adapted process on $(\Omega, \mathcal{F}, P)$.
4) For each $u(\cdot)$, and any $y \in \mathbb{R}^n$, (IV.5) admits a unique weak solution $x(\cdot)$ on $(\Omega, \mathcal{F}, \{\mathcal{F}_s, t \geq s\})$ such that $x(s) = y$.
5) $f(\cdot, x(\cdot), u(\cdot)) \in L^1_T(0, T; \mathbb{R})$ and $h(x(T)) \in L^1_T(\Omega; \mathbb{R})$.

**Assumptions 4.2:** The nominal system satisfies the following assumptions:
1) $(U, d)$ is Polish space, where $U \subset \mathbb{R}^k$. The control \(\{u(t); t \in [0, T]\}\) is non anticipative and takes values in $U$ which is compact and convex.

2) The maps $b: [0, T] \times U \times \mathbb{R}^n \to \mathbb{R}^n$, $\sigma: [0, T] \times U \times \mathbb{R}^n \to \mathbb{R}^n$ and $h: [0, T] \times \mathbb{R}^n \to \mathbb{R}$ are uniformly continuous. Also $\sigma$ and $b$ are assumed to be bounded. There exists a constant $L$ such that $\phi(t, x, u) = b(t, x, u), f(t, x, u), h(x), \sigma(t, x, u)$ satisfy

\[
|\phi(t, x, u) - \phi(t, x, u_0)| \leq L|x - x_0|, \forall x, x_0 \in \mathbb{R}^n,
\]

\[
u U, \forall t \in [0, T]
\]

$|\phi(t, 0, u)| \leq L, \forall t \in [0, T] \times U$

Under the above two assumptions, for any $(s, y) \in [0, T] \times \mathbb{R}^n$ and $u(\cdot) \in \mathcal{U}_u[s, T]$, the SDE in (IV.5) admits a unique weak solution $x(\cdot) \equiv x(s, y, u(\cdot))$. The precise problem statement associated with (IV.7) is the following.

Problem 4.3: Given the nominal measure $P \in \mathcal{M}(u)$, find a $u^* \in \mathcal{U}_u[s, T]$ and a probability measure $Q^{u^*,*} \in \mathcal{M}(\Omega)$ which solve the following constrained optimization problem.

\[
J(u^*, Q^{u^*,*}) = \inf_{u(\cdot) \in \mathcal{U}_u[s, T]} \sup_{Q \in \mathcal{M}(u)} \left( E_Q \left\{ \int_s^T f(t, x(t), u(t))dt + h_x(T) \right\} \right)
\]

subject to fidelity

\[
|Q - P| \leq R, \quad R \in (0, 2]
\]

(IV.6)

Using Lemma 3.2, the above problem can be re-written in an exponential form as follows.

\[
J(u^*, Q^{u^*,*}) = \inf_{u(\cdot) \in \mathcal{U}_u[s, T]} \left( R. \left\| \int_s^T f(t, x(t), u(t))dt + h(x(T)) \right\| \right)
\]

\[
+ E_P \left\{ \int_s^T f(t, x(t), u(t))dt + h(x(T)) \right\}
\]

\[
= (1 + \beta) \inf_{u(\cdot) \in \mathcal{U}_u[s, T]} E_{Q^{u^*,*}} \left( \int_s^T f(t, x(t), u(t))dt + h(x(T)) \right)
\]

(IV.7)

where above $\| \cdot \|_\infty$ is defined over the space $\Omega$, and $Q^{u^*,*}$ is given by the following measure.

\[
Q^{u^*,*}(E) = \gamma E \left\{ \left( 1 + \frac{1}{\gamma} V(s, x(s), y, u(\cdot)) \right) \right\}
\]

\[
+ (1 - \gamma) P(E)
\]

(IV.8)

where $E \in \mathcal{F}$ and $\alpha$ depends on the choice of control $u(\cdot)$. Notice that $\eta \in \mathcal{M}_1(\Omega)$ is an arbitrary measure which is not unique. Clearly, the minimization problem (IV.7) is not standard. Further investigation is needed to derive a principle of optimality and a dynamic programming equation of Hamilton-Jacobi-Bellman type.

In the rest of the paper we consider the case when $\eta = P$, and we derive a Hamilton-Jacobi-Bellman equation for minimization problem (IV.7).

B. Principle of Optimality

In this section we employ (verify) Bellman’s principle of optimality for the non-standard minimization problem (IV.7) to derive a dynamic programming equation.

The precise problem statement associated with (IV.7) is the following.

Problem 4.4: Given $(s, y) \in [0, T] \times \mathbb{R}^n$, find a 5-tuple $(\Omega, \mathcal{F}, P, \mathcal{U}(\cdot), \pi(\cdot)) \in \mathcal{U}_u[s, T]$ such that

\[
J(s, \alpha, y; \pi(\cdot)) = \inf_{u(\cdot) \in \mathcal{U}_u[s, T]} J(s, \alpha, y; u(\cdot))
\]

where

\[
J(s, \alpha, y; u(\cdot)) = \gamma E \left\{ \left( \frac{1}{\gamma} V(s, x(s), y, u(\cdot)) \right) \right\}
\]

\[
+ (1 - \gamma) E \left( \int_s^T f(t, x(t), u(t))dt + h(x(T)) \right)
\]

(IV.9)

Associated with Problem 4.4, define the following value function or cost-to-go from $s$ to $T$.

\[
\left\{ \begin{array}{ll}
V(s, y) & = \inf_{u(\cdot) \in \mathcal{U}_u[s, T]} J(s, \alpha, y; u(\cdot)), \\
\psi(s, y) & \in [0, T] \times \mathbb{R}^n, \\
V(T, y) & = h(y), \quad \forall y \in \mathbb{R}^n
\end{array} \right.
\]

(IV.10)

Since $\alpha$ depends on the initial time $s$, from now on we use the notations $\alpha_s$ and $J(s, \alpha_s, y; u(\cdot))$. However, since $\alpha$ also depends on the choice of control $u(\cdot)$, the value function $V(s, \cdot)$ will be independent of $\alpha_s$.

Denote by $x(\cdot; s, \xi, u(\cdot))$ the solution of (IV.5) starting at the initial condition $x(s) = \xi$.

The following theorem gives the non-linear stochastic version of Bellman’s principle of optimality, also known as the dynamic programming equation.

Theorem 4.5: Let assumption 4.2 hold. Then for any $(s, y) \in [0, T] \times \mathbb{R}^n$ we have

\[
V(s, y) \leq \inf_{u(\cdot) \in \mathcal{U}_u[s, T]} \left\{ \frac{1}{\gamma} V(s, x(s), y, u(\cdot)) \right\}
\]

\[
+ \int_s^T f(t, x(t; s, y, u(\cdot)), u(t))dt
\]

\[
x E_{\eta} \left( \int_s^T f(t, x(t; s, y, u(\cdot)), u(t))dt + \frac{\partial}{\partial y} V(s, x(s), y, u(\cdot)) \right)
\]

\[
\times E \left( \int_s^T f(t, x(t; s, y, u(\cdot)), u(t))dt + \frac{\partial}{\partial y} V(s, x(s), y, u(\cdot)) \right)
\]

\[
+ (1 - \gamma) E \left( \int_s^T f(t, x(t; s, y, u(\cdot)), u(t))dt \right)
\]
Proof. Omitted due to space limitation.

C. Generalized Hamilton-Jacobi-Bellman Equation

Next we use the dynamic programming equation of the previous theorem to derive a generalized partial differential equation satisfied by the value function \( V(\cdot, \cdot) \) known as HJB equation. The derivation assumes that \( V \in C^{1,2}(0, T \times \mathbb{R}^n) \) (the set of all continuous functions \( V : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R} \) such that the partial derivatives \( V_t, V_x, V_{xx} \) are continuous in \((t, x)\)).

Theorem 4.6: Suppose Assumptions 4.2 hold and the value function \( V \in C^{1,2}(0, T \times \mathbb{R}^n) \).

Define

\[
\begin{align*}
G(t, x, u, q, r) &\triangleq (1 - \frac{\alpha}{\gamma})\left(<p, b(t, x, u)>\right) \\
&+ \frac{1}{2}tr(q\sigma^T) + \frac{\alpha^2}{2\gamma^2}tr(p^2\sigma^T) - f(t, x, u),
\end{align*}
\]

\[\forall(t, x, u, q, r) \in [0, T] \times \mathbb{R}^n \times U \times \mathbb{R}^n \times S^n \times \mathbb{R}\]

where \( S^n \) is the space of \( n \) by \( n \) symmetric matrices.

Then \( V \) is the solution of the following second-order partial differential inequality of generalized Hamilton-Jacobi-Bellman type.

\[
\begin{aligned}
&-V_t(s, y) + \sup_{u \in U} \{ G(s, y, u, -V_x(s, y), -V_{xx}(s, y), -V(s, y)) \} \\
&\leq 0, \quad (s, x) \in [0, T] \times \mathbb{R}^n
\end{aligned}
\]

Theorem 4.9: Let assumption 4.2 hold. Then value function is a viscosity subsolution of (IV.11).

Proof. Omitted due to space limitations.

D. Viscosity Solutions

In this section, we discuss the case in which value function is not smooth enough and may not have first or second derivatives. First, the definition of viscosity solutions is provided. For further details see [15].

Definition 4.8: A function \( V \in C([0, T] \times \mathbb{R}^n) \) is called a viscosity subsolution of the stochastic HJB equation IV.11, if

\[
V(T, x) \leq h(x) \quad \forall x \in \mathbb{R}^n \quad (IV.11)
\]

and for any \( \phi \in C^{1,2}([0, T] \times \mathbb{R}^n) \), whenever \( V - \phi \) attains a local maximum at \((t, x) \in [0, T] \times \mathbb{R}^n\) we have

\[
\begin{aligned}
&\phi_t(s, y) + \sup_{u \in U} \{ G(s, y, u, -\phi_x(s, y), -\phi_{xx}(s, y), -\phi(s, y)) \} \\
&\leq 0
\end{aligned}
\]

\[C^{1,2}([0, T] \times \mathbb{R}^n) \] is the space of all continuous functions \( g : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R} \) such that \( g_t, g_x \) and \( g_{xx} \) are all continuous in \((t, x)\).