A Generalized Kyle-Back Strategic Insider Trading Model with Dynamic Information

Jin Ma* and Ying Tan†

Abstract

In this paper we consider a class of generalized Kyle-Back strategic insider trading models in which the insider is able to use the dynamic information obtained by observing the instantaneous movement of an underlying asset that is allowed to be influenced by its market price. Since such a model will be largely outside the Gaussian paradigm, we shall try to Markovize it by introducing an auxiliary diffusion process, in the spirit of the weighted total order process of, e.g., [12], as a part of the “pricing rule”. As the main technical tool in solving the Kyle-Back equilibrium, we study a class of Stochastic Two-Point Boundary Value Problem (STPBVP), which resembles the dynamic Markov bridge in the literature, but without insisting on its local martingale requirement. In the case when the solution of the STPBVP has an affine structure, we show that the pricing rule functions, whence the Kyle-Back equilibrium, can be determined by the decoupling field of a forward-backward SDE obtained via a non-linear filtering approach, along with a set of compatibility conditions.

Keywords. Kyle-Back equilibrium, strategic insider trading problem, conditioned SDE, two-point boundary value problem, FKK equation, forward-backward SDE, stochastic optimal control.

2020 AMS Mathematics subject classification: 60H10; 93E11; 91G15, 80.

*Department of Mathematics, University of Southern California, Los Angeles, 90089; email: jinma@usc.edu. This author is supported in part by US NSF grants #1106853.
†Department of Mathematics, University of Southern California, Los Angeles, 90089; email: yingtan@usc.edu
1 Introduction

In this paper we are interested in an asset pricing problem with asymmetric information, known as the Kyle-Back strategic insider trading equilibrium problem initiated by Kyle [26] and Back ([4, 5]) (see also [1, 9, 11, 17, 25] and the references therein for various generalizations of such models, along with different approaches). In particular, we will focus on the cases of dynamic information, in which the insider is allowed to use the dynamically observed information on the underlying asset, rather than the information at a fixed terminal time, as it was originally suggested. We shall carry out the analysis in a general Markovian, hence non-Gaussian framework.

The Kyle-Back strategic insider trading problem can be briefly described as follows. Consider a market that involves three types of agents: (i) The insider, who possesses some information of a given asset $V = \{V_t\}_{t \in [0,T]}$ that is not observable in the market. The information can be either the law of $V_T$, or the instantaneous observation of the state $V_t$, $t \in [0,T]$, or both. In the literature, they are often referred to as the “long-lived information” and the “dynamic information”, respectively. The insider will then submit her order, denoted by $\xi_t, t \in [0,T]$. (ii) The noise traders, who have no direct information of the asset $V$, and (collectively) submit an order $z_t$ at time $t \in [0,T]$. It is commonly assumed, by virtue of the central limit theorem, that $z_t = \int_0^t \sigma_z dB_z$, where $B^z$ is a Brownian motion. (iii) Finally, the marked maker, who observes the total traded volume in the market, $Y_t = \xi_t + z_t, t \in [0,T]$, and sets the price for $V_t$. It is standard to assume (see, e.g., [26], by a Bertrand competition argument) that the market price $P_t, t \geq 0$, is the $L^2$-projection of the true value $V$ to the space of $\mathbb{F}^Y$-measurable random variables. In other words, one assumes that, for $t \in [0,T]$,

$$P_t = \begin{cases} E[V_T|\mathcal{F}^Y_t] & \text{(long-lived information)} \\ E[V_t|\mathcal{F}^Y_t] & \text{(dynamic information),} \end{cases} \quad (1.1)$$

where $\mathcal{F}^Y_t \triangleq \sigma\{Y_s, s \leq t\}$. An equilibrium of the Kyle-Back problem consists of an insider’s strategy $\xi^*$ that maximizes her expected wealth at the terminal time $T$, together with a specified market price $P$ in either form of (1.1) (often referred to as the market efficiency).

Strong efforts have been made in recent years to extend the Kyle-Back problem to more general settings beyond the traditional Gaussian framework, and some deeper mathematical tools have been introduced to deal with the solvability issues accompanied by the generality of the modeling (see, for example, [12, 13, 14, 16] and the references cited therein). It is thus always interesting to identify methodologies that are easily accessible and at the same time efficient for solving more general models. This paper is an effort in this general direction.

We are interested in a Kyle-Back equilibrium problem with the following features:
The evolution of the dynamics of the underlying asset can depend on the market price $P = \{P_t\}$ (hence not independent of the market information $\mathbb{F}^Y = \{\mathbb{F}^Y_t\}$).

(ii) The insider can observe both the movements of the underlying asset and the market price, and uses the information when decides her optimal strategy; and

(iii) the market maker’s pricing rule is of the form an “optional projection” of the underlying asset (i.e. the second form in (1.1)), rather than a martingale (the first form in (1.1)).

We note that the feature (i) above, although reasonable (see, e.g., [29]), would put our problem outside most of the cases studied in the literature, due to various technical reasons which will become clear when our analysis proceeds, especially when the idea of “dynamic Markov bridge” is adopted. The requirement (iii), however, will be a natural connecting point to the nonlinear filtering, given the reasonable structure of the asymmetric information. More precisely, in this paper we shall assume that the underlying asset $V$ is governed by the following general SDE:

$$dV_t = b(t, V_{\cdot \wedge t}, P_{\cdot \wedge t}) dt + \sigma(t, V_{\cdot \wedge t}, P_{\cdot \wedge t}) dB^1_t, \quad V_0 = v, \quad (1.2)$$

where $b, \sigma$ are given measurable functions. We shall also assume, as commonly seen in the literature, that the insider’s strategy is of the form $\xi_t = \int_0^t \alpha_s ds$, $t \geq 0$, where the “rate” $\alpha$ can depend on both $V$ and $P$ in a nonanticipative way, so that the dynamics the market maker observes is:

$$dY_t = d\xi_t + dz_t = \alpha(t, V_{\cdot \wedge t}, P_{\cdot \wedge t}) dt + dB^2_t, \quad t \geq 0. \quad (1.3)$$

We remark that under the market efficiency requirement (1.1), the SDEs (1.2) and (1.3) in general form a so-called conditional mean-field SDE (CMFSDE) (or more generally, conditional McKean-Vlasov SDE (CMVSDE), whose well-posedness is not trivial (cf., e.g, [10, 29]). In this paper we shall take a different route, and follow the idea of [12] and introduce a factor model which in a sense Markovizes the “path-dependent” SDEs (1.2) and (1.3) completely. To be more precise, we are looking for a factor process $X$ that is determined completely by the observation Itô process $Y$, in the sense that $X_t = \Psi(t, Y_{\cdot \wedge t})$, such that the market price $P$ is determined by

$$P_t = H(t, X_t) = H(t, \Psi(t, Y_{\cdot \wedge t})) = \Phi(t, Y_{\cdot \wedge t}), \quad t \in [0, T].$$

Such a factor process $X$ resembles the so-called weighted total process (see, e.g., [12]), which was often assumed to be a diffusion process driven by the observation process $Y$ (see §2 for detailed discussion). With such a Markovization, we shall recast the equilibrium problem as a stochastic control problem and show that, by a dynamic programming argument, a necessary condition for $\alpha^* = \tilde{u}^*(\cdot \cdot \cdot)$ being optimal is that the corresponding solution $(V, X)$ satisfies:

$$V_T = P_T = H(T, X_T) := g(X_T). \quad (1.4)$$

We note that the relationship (1.4) naturally leads to a two-point boundary value problem structure, or a “bridge”. In fact, there has been a tremendous effort to use the notion of dynamic
Markov bridge to help finding the Kyle-Back equilibrium (see, e.g., [22, 12, 13]), and the methodology works well when some technical and structural assumptions are made to ensure the solvability. However, these assumptions excludes the more convoluted situations such as (1.2).

The main motivation of this paper is based on the following observation: although dynamic Markov bridge is a powerful tool in solving the problem, it can be slightly relaxed for the purpose for this particular problem. In other words, a slightly generalized version, which we shall refer to as the stochastic two-point boundary value problem (STPBVP), would be sufficient, if not more effective, for our purpose. Our main idea is to simply use the so-called “conditioned” SDE (see, Baudoin [6]) and design a specific minimal probability measure for the two-dimensional Markovian process \((V, X)\), and construct a weak solution to the STPBVP. Some fundamental tools in the study of dynamic Markov bridge should be sufficient for the resolution of TPBVP, whence the desired Kyle-Back equilibrium problem. We should note that the choice of the coefficients of the factor process \(X\) is somewhat ad hoc, and we can and will impose some structural assumptions that would lead to explicit “compatibility conditions” among coefficients of \(V\) and \(X\). In particular, in this paper we shall assume an affine structure, motivated in part by the well-known Widder’s Theorem (cf. e.g., [7, 24, 34, 35]) and the solution of the STPBVP. We shall first argue that, given the affine structure, some analysis similar to affine term structure of interest rates can be used to derive the compatibility conditions; and the optional projection \(P_t = \mathbb{E}[V_t | \mathcal{F}_t^Y]\) can be rigorously put into a nonlinear filtering framework with \((V, X)\) being the state signal process, and \(Y\) being the observation process. Furthermore, the terminal condition (1.4) will lead to a coupled Forward-backward SDE (FBSDE), with the factor process \(X\) being the forward SDE, and the Fujisaki-Kallianpur-Kunita (FKK) equation of the filtering problem being the backward SDE, both driven by the process \(Y\). We then show that the corresponding decoupling field (cf. [30]) is exactly the pricing rule \(H\) (see, e.g., [12]). Note that such a connection opens the door to a potentially much more general framework in which the decoupling field \(H\) is allowed to be a random field, determined by a backward stochastic PDE (BSPDE), as is often seen in the FBSDE literature (cf. e.g., [31]). We hope to be able to address such issues in our future publications.

The rest of the paper is organized as follows. In §2 we formulate the problem and introduce the notations and definitions. In §3 we revisit the conditioned SDE; and in §4 we formulate the stochastic two-point boundary value problem (STPBVP) and investigate its well-posedness and fundamental properties. In §5 we introduce the notion of affine structure for the solution to the STPBVP and associated insider strategies. In §6 we discuss the filtering problem and derive the FKK equation and the corresponding FBSDE under the affine structure. Finally, in §7 we discuss the sufficient conditions for optimality, and determine the equilibrium strategies.
2 Preliminaries and Problem Formulation

Throughout this paper, let $X$ be a generic Euclidean space and regardless of its dimension, $\langle \cdot, \cdot \rangle$ and $| \cdot |$ be its inner product and norm, respectively. We denote the space of $X$-valued continuous functions defined on $[0, T]$ with the usual sup-norm by $C([0, T]; X)$. In particular, we denote $C^2_T := C([0, T]; \mathbb{R}^2)$, and let $\mathcal{B}(C^2_T)$ be its topological Borel field. We shall assume that all randomness in this paper is characterized by a canonical probabilistic set-up: $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, B)$, where $(\Omega, \mathcal{F}) := (C^2_T, \mathcal{B}(C^2_T)); \mathbb{P} \in \mathcal{P}(\Omega)$; and $B = (B^1, B^2)$ is a $\mathbb{P}$-Brownian motion. Moreover, we shall assume that $\mathbb{F}^i = \{\mathcal{F}_t^B\}_{t \geq 0}, i = 1, 2$, is the natural filtration generated by $B^1$ and $B^2$, respectively, and $\mathbb{F} = \mathbb{F}^1 \vee \mathbb{F}^2$, with the usual $\mathbb{P}$-augmentation so that it satisfies the usual hypotheses (cf. e.g., [32]). Finally, we denote $Q^0 \in \mathcal{P}(\Omega)$ to be the Wiener measure on $(\Omega, \mathcal{F})$; $B^0(\omega) = \omega(t), \omega \in \Omega$, the canonical process; and $\mathbb{F}^0 := \{\mathcal{F}_t^0\}_{t \in [0, T]}$, where $\mathcal{F}_t^0 := \mathcal{B}_t(C^2_T) := \sigma(\omega(\cdot \wedge t) : \omega \in C^2_T), t \in [0, T]$. In what follows we shall make use of the following notations:

- For any sub-$\sigma$-field $\mathcal{G} \subseteq \mathcal{F}_T$ and $1 \leq p < \infty$, $L^p(\mathcal{G}; X)$ denotes the space of all $X$-valued, $\mathcal{G}$-measurable random variables $\xi$ such that $\mathbb{E}|\xi|^p < \infty$. As usual, $\xi \in L^\infty(\mathcal{G}; X)$ means that it is $\mathcal{G}$-measurable and bounded.

- For $1 \leq p < \infty$, $\mathcal{G} \subseteq \mathbb{F}$, $L^p_\mathcal{G}([0, T]; X)$ denotes the space of all $X$-valued, $\mathcal{G}$-progressively measurable processes $\xi$ satisfying $\mathbb{E}\int_0^T |\xi_t|^p dt < \infty$. The meaning of $L^\infty_\mathcal{G}([0, T]; X)$ is defined similarly. For simplicity, we will often drop $X(= \mathbb{R})$ from the notation, and denote all “$L^p$-norms” by $\|\cdot\|_p$, regardless it is for $L^p(\mathcal{G})$, or for $L^p_\mathcal{G}([0, T])$, when the context is clear.

The Problem Formulation. As we indicated in before, there are three types of agents in the market: the insider; the noise trader; and the market maker, which we now specify in details.

(i) The insider. In this paper we shall assume that the insider can both dynamically observe the liquidation value of the underlying asset $V = \{V_t\}$, and have some information of $V_T$, in particular, the law of $V_T$, denoted by $m^* \in \mathcal{P}(\mathbb{R})$. Specifically, we assume that the asset process $V$ is governed by the following SDE:

$$dV_t = b(t, V_t, P_t)dt + \sigma(t, V_t, P_t)dB^1_t, \quad V_0 = v,$$

where $b, \sigma$ are measurable functions, and $P = \{P_t\}$ is the market price. We should note that allowing $(b, \sigma)$ to depend on the market price $P$ is one of the main features of this paper, which amounts to saying that the fundamental price $V$ is convoluted with the market information $\mathbb{P}^V$ (see (22) below), which leads to some fundamental difficulties that distinguishes this paper from most of the existing literature, especially in terms of the dynamic Markov bridge.

We should note that although the insider has more information of the underlying asset, even it’s law at a future time, we shall insist that its strategy is in the non-anticipating manner. More
precisely, we shall assume that the order process \( \xi_t, t \in [0,T] \), takes the form \( \xi_t = \xi_t^\alpha := \int_0^t \alpha_s ds \), where the process \( \alpha = \{\alpha_t\} \) is called the intensity of the trading strategy, and is assumed to have the form \( \alpha_t = u(t, V_t, P_t) \), \( t \in [0,T] \), for some function \( u \) to be determined (see, e.g., [5, 29]).

(ii) The noise traders. For simplicity, in this paper we shall assume that the (collective) order submitted by the noise traders is simply the \( z_t = B_t^2 \), for some Brownian motion \( B^2 \perp B^1 \). In other words, we assume that \( B^z = B^2 \), and \( \sigma^z \equiv 1 \).

(iii) The market maker. By virtue of the so-called Bertrand competition argument (see, e.g., [26]), we assume that at each time \( t \in [0,T] \), the market maker sets the (market) price \( P_t \) to be the \((L^2-)\)projection of the (unobservable) underlying price \( V_t \) onto the space of all \( \mathcal{F}_t^Y \)-measurable random variables. That is, \( P_t = \mathbb{E}[V_t|\mathcal{F}_t^Y], \ t \in [0,T] \), where \( Y \) is the total trading volume:

\[
Y_t = \xi_t^\alpha + B_t^2 = \int_0^t \alpha_s ds + B_t^2, \quad t \in [0,T].
\] (2.2)

Furthermore, we require that the asymmetry of information ends at the terminal time \( T \). That is, at terminal \( T > 0 \) the value of the underlying asset \( V_T \) will be revealed (by, e.g., an announcement, cf. [5]) and the market price will be set as \( P_T = V_T \), so that the insider does not have any information advantage by the time \( T \). We should note that such a requirement is not a natural consequence given the market parameters (i.e., the coefficients of SDEs involved), but rather one of the conditions the equilibrium strategy must satisfy.

Before we formulate the equilibrium problem, let us specify the set of admissible strategies:

\[
\mathcal{U}_{ad} := \{ \alpha \in \mathbb{L}^2_F([0,T]) : L^\alpha \text{ is a local martingale on } [0,T] \}.
\] (2.3)

where \( L^\alpha \) is the Doléans-Dade stochastic exponential: \( L^\alpha_t := \exp \left\{ \int_0^t \alpha_s dB_s^2 - \frac{1}{2} \int_0^t |\alpha_s|^2 ds \right\} \), \( t \in [0,T] \). A (generalized) Kyle-Back equilibrium consists of a “pricing rule” \( P_t = \mathbb{E}[V_t|\mathcal{F}_t^Y], \ t \in [0,T] \); and an optimal strategy \( \alpha^* \in \mathcal{U}_{ad} \), such that the terminal wealth, defined by

\[
W_T = W_T^{\alpha^*} := \int_0^T \xi_t^{\alpha^*} dP_t
\]

has a maximum expected value \( \mathbb{E}^P[W_T^{\alpha^*}] = \sup_{\alpha \in \mathcal{U}_{ad}} \mathbb{E}^P[W^\alpha] \).

Remark 2.1. (i) We note that, in (2.3) the process \( L^\alpha \) is defined only on \([0,T)\). Indeed, in light of the existing results, the optimal strategy \( \alpha_t \) may very well explode when \( t \nearrow T \), because the insider will try to use all the information advantage before it ends.

(ii) From (2.2) we observe that \( Y \) depends on the choice of \( \alpha \), thus so does the process \( P \), whence the asset price \( V \). Denoting \( V = V^\alpha \), a more precise definition of the admissible control set should be all \( \alpha \in \mathcal{U}_{ad} \) such that \( V^\alpha_T \sim m^* \in \mathcal{P}(\mathbb{R}) \), the law that the insider was expecting.
We prefer not to impose such a restriction in order to avoid unnecessary technical subtlety, but will emphasize this issue when it is needed in our discussion (e.g., in §7).

The Markovization. We note that the market price $P_t = \mathbb{E}[V_t | \mathcal{F}^Y_t], t \in [0, T]$, is in general an optional projection of $V$ onto the filtration $\mathcal{F}^Y = \{\mathcal{F}_t\}$, but not necessarily an $\mathcal{F}^Y$-martingale as the “long-lived information” case (see (1.1)) considered in most of the existing literature. In general the market price $P$ can be written as $P_t = \Phi(t, Y_{\Lambda t}), t \geq 0$, for some measurable function $\Phi$ defined on $\mathbb{C}([0, T])$. Therefore (2.1)–(2.2) is by nature a system of “path-dependent” Conditional McKean-Vlasov SDEs (CMVSDEs) or Conditional Mean-field SDEs (CMFSDEs) (see [10, 29]). In this paper we shall follow the idea of [12] to first Markovize the system (2.1)-(2.2) by introducing a factor process $X$, which satisfies an auxiliary SDE of the form:

$$dX_t = \mu(t, X_t)dt + \rho(t, X_t)dY_t, \quad X_0 = x,$$

where the coefficients $(\mu, \rho)$ are to be determined, so that the market price $P$ can be written as $P_t = H(t, X_t)$ for some function $H$. We note that, if on some probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, where $\mathbb{Q} \in \mathcal{P}(\Omega)$ under which $Y$ is a Brownian motion, then, as the strong solution to SDE (2.4), $X$ can be written as $X_t = \Psi(t, Y_{\cdot \wedge t})$, for some measurable function $\Psi$, and consequently, we have

$$P_t = \mathbb{E}[V_t | \mathcal{F}^Y_t] = H(t, X_t) = H(t, \Psi(t, Y_{\Lambda t})) = \Phi(t, Y_{\Lambda t}), \quad t \in [0, T].$$

We note that the factor process $X$ is similar to the weighted total order process, proposed in [12], and the function $H$ (together with the coefficients $(\mu, \rho)$) can be considered as the “pricing rule” (see [12, 13]). They will be the main subject of this paper.

We should remark here that a direct consequence of the Markovization is that we can now put the problem of finding the equilibrium into a standard stochastic control framework. More specifically, since $P_t = H(t, X_t)$, by a slight abuse of notation, we shall assume from now on that the underlying asset $V$ and the factor process $X$ follow a system of SDEs:

$$\begin{align*}
&\left\{ \begin{array}{ll}
dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB^1_t, & V_0 = v; \\
dX_t = \mu(t, X_t)dt + \rho(t, X_t)dY_t = [\mu(t, X_t) + \alpha(t, X_t)]dt + \rho(t, X_t)dB^2_t, & X_0 = x.
\end{array} \right.
\end{align*}$$

(2.5)

Considering (2.5) as a controlled system with controls $\alpha \in \mathcal{U}_{ad}$. Following the argument of [4] by allowing a market clearing jump at terminal time, then a simple integration by parts shows that the expected terminal wealth can be written as:

$$\mathbb{E}[W^\alpha_T] = \mathbb{E}[(V_T - P_T)\xi^\alpha_T + \int_0^T \xi^\alpha_t dP_t] = \mathbb{E}[\xi^\alpha_T V_T - \int_0^T \alpha_t P_t dt] = \mathbb{E}[-\int_0^T [V_T - P_t] \alpha_t dt].$$

(2.6)
Assuming now the process $\alpha$ takes the feedback form: $\alpha_t = u(t, V_t, X_t)$, then $(V, X)$ becomes Markovian, and we deduce from (2.6) that

$$
\mathbb{E}[W_T^\alpha] = \mathbb{E} \left[ \int_0^T \mathbb{E}[V_T|F_t^{V,X}] - P_t|\alpha_t dt \right] = \mathbb{E} \left[ \int_0^T [F(s, V_s, X_s) - H(s, X_s)]\alpha_t dt \right],
$$

(2.7)

where $F$ is a continuous function satisfying $F(T, v, x) = v$, and can be determined by the Kolmogorov backward equation or Feynman-Kac formula (see §7 for details). Consequently, we can define a stochastic control problem with $(V, X)$ as the controlled dynamics, and the cost function $J(t, v, x; u) := \mathbb{E}_{t, v, x} \left[ \int_t^T (F(s, V_s, X_s) - H(s, X_s))u(s, V_s, X_s)ds \right]$, (2.8)

so that the value function $v(t, v, x) := \sup_{\alpha \in \mathcal{U}_{ad}} J(t, v, x; u)$ satisfies the following HJB equation:

$$
0 = \partial_t v(t, v, x) + b(t, v, x)\partial_v v + \mu(t, x)\partial_x v + \frac{1}{2}\sigma^2(t, v, x)\partial_{vv} v + \frac{1}{2}\rho^2(t, x)\partial_{xx} v
+ \sup_{u \in \mathbb{R}} \left\{ \left[ \rho(t, x)\partial_x v + F(t, v, x) - H(t, x) \right]u \right\}. 
$$

(2.9)

Clearly, a necessary condition for the “sup”-term in (2.9) to be finite is

$$
\rho(t, x)\partial_x v + F(t, v, x) - H(t, x) = 0, \quad (t, v, x) \in [0, T] \times \mathbb{R}^2.
$$

In particular, noting that $F(T, v, x) = v$, and $v(T, v, x) \equiv 0$ by definition (2.8), we deduce that

$$
0 \equiv \rho(T, x)\partial_x v(T, v, x) = H(T, x) - F(T, v, x) =: g(x) - v, \quad (v, x) \in \mathbb{R}^2,
$$

(2.10)

where $g(x) = H(T, x)$. In other words, at the terminal time $T$, it holds that $V_T = g(X_T)$ for some function $g$. In fact, similar to [12], we shall assume from now on that the function $g$ is increasing. Consequently, (2.10) indicates an important fact: a necessary condition for $\alpha \in \mathcal{U}_{ad}$ being an equilibrium is that the following condition holds at the terminal time $T$:

$$
V_T = P_T = H(T, X_T) = g(X_T).
$$

(2.11)

A Stochastic Two-Point Boundary Valued Problem (STPBVP). Summarizing the discussion above we see that we should look for $\alpha \in \mathcal{U}_{ad}$ and coefficients $(\mu, \rho)$ so that the following system of SDEs with initial-terminal conditions is solvable:

$$
\begin{cases}
\begin{align*}
dV_t &= b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB_t^1, \\
\alpha_t &= \mu(t, X_t)dt + \rho(t, X_t)dY_t = [\mu(t, X_t) + \alpha_t\rho(t, X_t)]dt + \rho(t, X_t)dB_t^2,
\end{align*}
\end{cases}
\right.
$$

(2.12)

$$
V_0 = v, \quad X_0 = x, \quad V_T = g(X_T).
$$
In what follows we shall refer to (2.12) as a *Stochastic Two-Point Boundary Value Problem*, which will be studied in details in the next section. The solvability of the STPBVP depends on the choice of the process \( \alpha = \{ \alpha_t \} \). In particular, we are particularly interested in the case when \( \alpha \) takes the form \( \alpha_t = u(t, V_t, X_t) \), which will render the solution \((V, X)\) a Markov process.

We remark that the TPBVP (2.12) is closely related to the so-called *dynamic Markov bridge* studied in, e.g., [22, 12, 13]. In fact, if \( b = \mu = 0, \sigma = \rho = 1, \) and \( g(x) = x \), the problem (2.12) was first studied, as the Brownian bridge, in the context of insider trading in [22]. The more general cases were considered recently in [12, 13, 15], also in the bridge context. But on the other hand, we note that in the description of the problem above we see that the TPBVP (2.12) does not actually require that the solution \( X \) to be a local martingale under its own filtration, a key requirement to be a Markovian bridge (see §3 for a more detailed discussion). Thus, the main point of this paper is to show that such a relaxation enables us to solve the Kyle-Back equilibrium problem in a much more general setting.

### 3 The Conditioned SDE Revisited

Our construction of the (weak) solution to TPBVP (2.12) is based on the notion of the so-called *conditioned SDE* (cf. [6]), which we now briefly describe. Recall the canonical probabilistic set-up \((\Omega, \mathcal{F}, Q^0; \mathcal{F}, B^0)\) defined in the beginning of §2. In particular, we denote the canonical process by \( B^0 = (B_1, Y) \) so that it is a \((Q^0, \mathcal{F})\)-Brownian motion. Consider the SDE on canonical space \((\Omega, \mathcal{F}, Q^0, B^0)\), for \( t \in [0, T] \):

\[
\begin{align*}
    dV_t &= b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB^1_t, \quad V_0 = v; \\
    dX_t &= \mu(t, X_t)dt + \rho(t, X_t)dY_t, \quad X_0 = x.
\end{align*}
\]

We assume that the coefficients \( b, \sigma, \mu, \rho \) and \( g \) satisfy the following *Standing Assumptions*:

**Assumption 3.1.** (i) The functions \( b, \sigma : [0, T] \times \mathbb{R}^2 \mapsto \mathbb{R} \) and \( \mu, \rho : [0, T] \times \mathbb{R} \mapsto \mathbb{R} \) are measurable, and continuous in \( t \in [0, T] \);

(ii) There exists \( L > 0 \), such that, for any \( t \in [0, T], v, v', x, x' \in \mathbb{R} \), it holds that,

\[
\begin{align*}
    &|b(t, 0, 0)| + |\sigma(t, 0, 0)| + |\mu(t, 0)| + |\rho(t, 0)| \leq L, \\
    &|\phi(t, v, x) - \phi(t, v', x')| \leq L(|v - v'| + |x - x'|), \quad \phi = b, \sigma, \\
    &|\psi(t, x) - \psi(t, x')| \leq L|x - x'|, \quad \psi = \mu, \rho;
\end{align*}
\]

(iii) There exists a constant \( \lambda_0 > 0 \), such that \( \sigma(t, v, x) \geq \lambda_0 \), \( (t, v, x) \in [0, T] \times \mathbb{R}^2 \);

(iv) The function \( g \) is uniformly Lipschitz continuous and strictly monotone increasing.
Clearly, under Assumption 3.1, SDE 3.1 has a unique strong solution over \([0, T]\), on \((\Omega, \mathcal{F}, \mathbb{Q}^0)\), denoted by \(\xi := (V^0, X^0)\). Moreover, \(\xi\) is a Markov process, and we denote its transition density by \(p(s, x; t, y)\), \(0 \leq s < t \leq T, x, y \in \mathbb{R}^2\). Now, for any \(\nu \in \mathcal{P}(\mathbb{R}^2)\), consider the triplet \((T, \xi_T, \nu)\), which we shall refer to as a “conditioning” below. Let us now define
\[
\eta_t^y := \frac{p(t, \xi_t; T, y)}{p(0, \xi_0; T, y)}, \quad t < T, \mathbb{Q}^0\text{-a.s.},
\]
where \(p(\cdot, \cdot; \cdot, \cdot)\) is the transition density of \(\xi\) under \(\mathbb{Q}^0\), and the process
\[
L_t^\nu := \int_{\mathbb{R}^2} \eta_t^y \nu(dy), \quad t \in [0, T).
\]

**Definition 3.2.** The conditioning triplet \((T, \xi_T, \nu)\) is called “proper” if

(i) \(\text{supp}(\nu) \subseteq \text{supp}(\mathbb{Q}^0 \circ \xi_T^{-1})\); and

(ii) there exist constants \(C, \lambda > 0\), such that
\[
0 < \sup_{t \in [0, T]} (T - t)\eta_t^y \leq CTe^{\frac{\lambda(\|y\|^2 - \|x\|^2)}{t - s}}, y \in \mathbb{R}^2; \quad \text{and} \quad \int_{\mathbb{R}^2} e^{\frac{\lambda(\|y\|^2 - \|x\|^2)}{4(t-s)}} \nu(dy) < \infty. \tag{3.2}
\]

We note that the condition (i) above is relatively easier to verify. In particular, it would be trivial when the diffusion \(\xi\) has positive density at time \(T\). For condition (ii), we note that \(p(s, y; t, x)\) is the fundamental solution to the Kolmogorov backward (parabolic) PDE, then it is well-known that (see, e.g., [2, 3]), for some constant \(c_1, c_2, \lambda, \Lambda > 0\), it holds that
\[
0 < \frac{c_1}{t - s} e^{-\frac{\lambda(\|y\|^2 - \|x\|^2)}{t - s}} \leq p(s, y; t, x) \leq \frac{c_2}{t - s} e^{-\frac{\lambda(\|y\|^2 - \|x\|^2)}{4(t-s)}}, \quad 0 \leq s < t < T, \quad x, y \in \mathbb{R}^2,
\]
Consequently we see that,
\[
0 < \eta_t \leq \frac{c_2T}{c_1(T-t)} e^{-\frac{\lambda(\|y\|^2 - \|x\|^2)}{t - s} + \frac{\lambda(\|y\|^2 - \|x\|^2)}{4(t-s)}} \leq \frac{c_2T}{c_1(T-t)} e^{\frac{\lambda(\|y\|^2 - \|x\|^2)}{t}}, \quad t \in [0, T), \tag{3.3}
\]
which leads to the first inequality in (3.2). Thus the requirement for the conditioning being “proper” means that \(L_t^\nu < \infty\) for all \(t \in [0, T)\), \(\mathbb{Q}^0\)-a.s..

The following proposition contains some results similar to those in [6], extended to the 2-dimensional case but with slightly different assumptions (see also, [19, 21]). Although some proofs are essentially the same, we give a detailed sketch for completeness.

**Proposition 3.3.** Assume Assumption 3.1. Let \((T, \xi_T, \nu)\) be a given conditioning. Then,

(i) there exists a unique \(\mathbb{P}^\nu \in \mathcal{P}(\Omega)\), such that \(\mathbb{P}^\nu \circ \xi_T^{-1} = \nu\), and for any \(t < T\), any bounded \(X \in L^0(\mathcal{F}_t; \mathbb{R}^2)\), it holds that
\[
\mathbb{E}^{\mathbb{Q}^0}[X|\xi_T = y] = \mathbb{E}^{\mathbb{Q}^0}[\eta_t^y X], \quad t < T, \quad y \in \mathbb{R}^2, \mathbb{Q}^0_{\xi_T}\text{-a.s.}; \tag{3.4}
\]
(ii) assuming further that \((T, \xi_T, \nu)\) is proper, then for any \(t < T\), it holds that
\[
\frac{d\mathbb{P}^{\nu}}{d\mathbb{Q}^0}\big|_{\mathcal{F}_t} = \int_{\mathbb{R}^d} \eta_{\nu}^t \nu(dy); \tag{3.5}
\]

(iii) \(L^{\nu}\) is a non-negative \(Q^0\)-martingale on \([0, T)\), and \(L^\nu_T := \lim_{t \to T} L^\nu_t\) exists, with \(\mathbb{E}^{Q^0}[L^\nu_T] \leq 1\).

**Proof.** For the given conditioning \((T, \xi_T, \nu)\), let \(Q^\nu(\cdot) \in \mathcal{P}(\Omega)\) be the regular conditional probability defined by \(Q^\nu(A) := Q^0(A|\xi_T = y), A \in \mathcal{F}_T, y \in \mathbb{R}^2\), and define
\[
\mathbb{P}^{\nu}(A) := \int_{\mathbb{R}^2} Q^\nu(A) \nu(dy), \quad A \in \mathcal{F}_T. \tag{3.6}
\]

We now check (i). That \(\mathbb{P}^{\nu} \circ \xi_T^{-1} = \nu\) is obvious. To see (3.4), we define a finite measure on \((\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))\) by \(\mu^{X|\xi_T}(A) := \int_{\xi_T \in A} X(\omega)Q^0(\omega)\), \(A \in \mathcal{B}(\mathbb{R}^2)\). Then, by definition we can write
\[
\mu^{X|\xi_T}(A) = \int_A E^{Q^0}[X|\xi_T = y]Q^0_{\xi_T}(dy) = \int_A E^{Q^0}[X|\xi_T = y]p(0, z_0; T, y)dy, \quad A \in \mathcal{B}(\mathbb{R}^2). \tag{3.7}
\]

But since \(X \in L^0(\mathcal{F}_T; \mathbb{R}^2)\), using the Markov property on \(\xi\) and Fubini theorem we also have
\[
\mu^{X|\xi_T}(A) = \int_\Omega \mathbb{E}^{Q^0}[1_{\{\xi_T \in A\}}X|\mathcal{F}_t](\omega)Q^0(d\omega) = \int_\Omega \left[ \int_A p(t, \xi_T(\omega); T, y)dy \right] X(\omega)Q^0(d\omega)
= \int_A \mathbb{E}^{Q^0}[p(t, \xi_T; T, y)X]dy, \quad A \in \mathcal{B}(\mathbb{R}^2). \tag{3.8}
\]

Comparing (3.7) and (3.8), we deduce (3.4).

(ii) To see (3.5), it suffices to show that for any \(Z \in L^1_{\mathcal{F}_T}(\mathbb{R}^2, Q^0)\), \(t \in [0, T)\), it holds that
\[
\mathbb{E}^{P^{\nu}}[Z] = \mathbb{E}^{Q^0}[L^\nu_t Z] = \mathbb{E}^{Q^0}\left[\int_{\mathbb{R}^2} \eta_{\nu}^t \nu(dy)Z \right]. \tag{3.9}
\]

By a standard truncation argument, we may assume that \(Z\) is bounded. Then by (3.4), we have \(\mathbb{E}^{Q^0}[\eta_{\nu}^t Z] = \mathbb{E}^{Q^0}[Z|\xi_T = y] = \mathbb{E}^{P^{\nu}}[Z|\xi_T = y]\), thanks to definition (3.6), we can write
\[
\mathbb{E}^{P^{\nu}}[Z] = \int_{\mathbb{R}^2} \mathbb{E}^{Q^0}[Z|\xi_T = y]\nu(dy) = \int_{\mathbb{R}^2} \mathbb{E}^{Q^0}[\eta_{\nu}^t Z]\nu(dy).
\]

Comparing above to (3.9), we see that it suffices to show that \(\int_{\mathbb{R}^2} \mathbb{E}^{Q^0}[|\eta_{\nu}^t Z|]\nu(dy) < \infty\), so that the Fubini theorem can be applied. But this clearly follows from the boundedness of \(Z\) and the assumption that the conditioning is proper.

(iii) Finally, by (3.5), \(\frac{d\mathbb{P}^{\nu}}{d\mathbb{Q}^0}\big|_{\mathcal{F}_t} := L^\nu_t, t < T\). Thus, \(L^{\nu}\) is a non-negative \(Q\)-martingale on \([0, T)\). Furthermore, since \(L^\nu_t > 0, t \in [0, T)\), by martingale convergence theorem, \(L_T := \lim_{t \to T} L_t\) exists, and by Fatou’s lemma, one easily shows that \(\mathbb{E}[L^\nu_T] \leq \lim_{t \to T} \mathbb{E}[L^\nu_t] = 1\).
Remark 3.4. (1) The probability $P\nu$ in Proposition 3.3 is called the minimal probability given the proper conditioning $(T, \xi_T, \nu)$. Moreover, Proposition 3.3 shows that the assumption (A1) in [6] is automatically satisfied in our setting.

(2) Proposition 3.3-(ii) indicates that $P\nu$ is absolutely continuous with respect to $Q^0$ on each $F_t$, $0 \leq t < T$, with the Radon-Nikodým derivative defined by (3.5). But it does not imply that $P\nu$ and $Q$ are equivalent on $F_t$, for $t < T$, neither does it imply that $P\nu < < Q$ on $F_T$.

We now turn our attention to a specific conditioning $(T, \xi_T, \nu)$ that will lead to the solution to an STPBVP (2.12). For notational convenience we shall now simply denote $\xi = (V, X)$, when there is no danger of confusion. Let $m^* \in P(R)$ be a law of the underlying asset $V_T$ that is known to the insider. For technical reasons we shall assume that $m^*$ satisfies the following condition:

Assumption 3.5. There exists $\lambda_0 > 0$ sufficiently large, such that

$$\int_R e^{\lambda_0 v^2} m^*(dv) < \infty.$$  \hspace{1cm} (3.10)

We remark that the Assumption 3.5 is actually not over restrictive. In fact, in light of the well-known Fernique Theorem (cf. [18]) (3.10) covers a large class of normal random variables.

Now let us define a probability measure $\nu \in P(R^2)$ by

$$\nu(A) = \int_R 1_A(v, g^{-1}(v))m^*(dv) = \int_{(v, g^{-1}(v)) \in A} m^*(dv).$$  \hspace{1cm} (3.11)

That is, the measure $\nu$ concentrates on the graph of the function $v = g^{-1}(x)$, or equivalently $x = g(v)$, thanks to Assumption 3.1-(iii). Furthermore, we have the following lemma.

Lemma 3.6. Assume Assumptions 3.1, 3.5 are in force, with $\lambda_0$ in (3.10) being sufficiently large. Let $\xi$ be the solution to (3.1), and $\nu \in P(R^2)$ be defined by (3.11). Then, $(T, \xi_T, \nu)$ is a proper conditioning. Furthermore, if $P\nu$ is the minimum probability given $(T, \xi_T, \nu)$, then it holds that

$$P\nu\{V_T = g(X_T)\} = 1.$$  \hspace{1cm} (3.12)

Proof. Since under Assumption 3.1 $\xi$ is a diffusion process with positive transition density function (cf. e.g., [20]), we have supp$(Q \circ \xi^{-1}_T) = R^2$. Furthermore, by definition of $\nu$ (3.11), for the constants $\lambda > 0$ in (3.3) we deduce from (3.10) that

$$\int_{R^2} e^{\lambda (v_0 - y)^2} \nu(dy) = \int_{R} e^{\lambda (v_0 - v)^2 + (g(x_0) - v)^2} m^*(dv) < \infty,$$

provided that $\lambda_0 \geq \frac{2\lambda}{\lambda_0}$, where $\lambda_0$ is the constant in Assumption 3.5. That is, $(T, \xi_T, \nu)$ is proper.

To show the second assertion, first note that $g$ is strictly increasing, the graphs of $g$ and $g^{-1}$, as the subset of $R^2$, are identical. Let us denote $\Gamma := \{(g(x), x) : x \in R\} = \{(v, g^{-1}(v)) : v \in
$\mathbb{R} \subseteq \mathbb{R}^2$. Then, by definition (3.11) we see that $\nu(A) = 1$ if and only if $\Gamma \subseteq A$. In particular, $\nu(\Gamma) = 1$. Consequently, by definition of the minimum probability, we have

$$\mathbb{P}^\nu(V_T = g(X_T)) = \mathbb{P}^\nu \circ \xi_T^{-1}(\Gamma) = \nu(\Gamma) = 1,$$

proving (3.12).

**Remark 3.7.** (1) Given (3.12), and the fact that $\xi = (V,X)$ has continuous paths under $\mathbb{Q}^0$, it is readily seen that $\mathbb{P}^\nu\{\lim_{t \to T} V_t = V_T = g(X_T) = \lim_{t \to T} g(X_t)\} = 1$. Consequently, we have

$$\mathbb{P}^\nu\{\lim_{t \to T} V_t = \lim_{t \to T} g(X_t)\} = 1. \quad (3.13)$$

This, together with Proposition 3.8, indicates that as far as the solution to the two-point boundary value problem is concerned, without the specific requirement of Markovian bridge, the SDE (3.18) would be a desirable candidate, except for a slight difference on the drift coefficients.

(2) By Proposition 3.3-(iii), $L'$ is a closeable supermartingale on $[0,T]$. But it cannot be a martingale, unless $\mathbb{Q}^0\{V_T = g(X_T)\} = 1$, which is obviously not true in general. Thus $\mathbb{P}^\nu$ cannot be absolutely continuous with respect to $\mathbb{Q}^0$ on $\mathcal{F}_T$, as we pointed out in Remark 3.4.

To end this section, let us define, for any proper conditioning $(T,\xi_T,\nu)$, a function

$$\varphi(t,z) = \int_{\mathbb{R}^2} \frac{p(t,z;T,y)}{p(0,z_0;T,y)} \nu(dy), \quad z = (v,x), \quad (3.14)$$

where $p$ is the transition density of $\xi$ under $\mathbb{Q}^0$. Then, clearly, $\varphi(0,z_0) = 1$, and $L_t = L'_t = \varphi(t,\xi_t)$, $t \in [0,T)$. Now, applying Itô’s formula we have

$$L_t = \varphi(t,\xi_t) = 1 + \int_0^t [\partial_t \varphi(s,\xi_s) + \mathcal{L}[\varphi](s,\xi_s)]ds + \int_0^t \left(\nabla \varphi(s,\xi_s), \tilde{\sigma}(s,\xi_s)dB^0_s\right), \quad (3.15)$$

where $\mathcal{L}[\varphi](t,z) := (\tilde{b}, \nabla \varphi)(t,z) + \text{tr} [D^2 \varphi \tilde{\sigma} \tilde{\sigma}^T](t,z)$, and $\tilde{b} := (b,\mu)^T$, $\tilde{\sigma} := \text{diag}[\sigma,\rho]$. Since by Proposition 3.3-(iii), $L$ is a $\mathbb{Q}^0$-martingale for $t \in [0,T)$, we conclude that $\varphi$ must satisfy the following PDE (noting the definition of $\tilde{b}$ and $\tilde{\sigma}$) for $t \in [0,T)$ and $z = (v,x) \in \mathbb{R}^2$,

$$\begin{cases}
\frac{\partial \varphi(t,z)}{\partial t} + b(t,z) \frac{\partial \varphi(t,z)}{\partial v} + \mu(t,x) \frac{\partial \varphi(t,z)}{\partial x} + \frac{1}{2} \sigma^2(t,z) \frac{\partial^2 \varphi(t,z)}{\partial v^2} + \frac{1}{2} \rho^2(t,x) \frac{\partial^2 \varphi(t,z)}{\partial x^2} = 0; \\
\varphi(0,v_0,\xi_0) = 1.
\end{cases} \quad (3.16)$$

Consequently, it follows from (3.15) that

$$dL_t = d\varphi(t,\xi_t) = (\nabla \varphi(t,\xi_t), \tilde{\sigma}(t,\xi_t)dB^0_t) = L_t(\theta_t, dB^0_t), \quad L_0 = 1, \quad t \in [0,T), \quad (3.17)$$

where $\theta_t := \tilde{\sigma}^T(t,\xi_t) \nabla \varphi(t,\xi_t) = \tilde{\sigma}^T(t,\xi_t) \nabla [\ln \varphi(t,\xi_t)]$, $t \in [0,T)$. Denote $W_t = B^0_t - \int_0^t \theta_sds$, then by Girsanov theorem, $\{W_t\}$ is a 2-dimensional $\mathbb{P}^\nu$-Brownian motion on $[0,T)$. In other words, we have proved the following 2-dimensional extension of a result in [6].
that, if the coefficients $b$ when necessary, and write the solution $(V,X,\alpha)$ (3.1) to its own filtration essentially a dynamic Markov bridge without requiring that $X$ be a local martingale with respect to its own filtration $\mathbb{F}^X$. Such a difference will make the results of this paper and those in the existing literature mutually exclusive.

4 A Stochastic Two-Point Boundary Value Problem

We are now ready to study the STPBVP and compare it to the well-known dynamic Markov bridge in the literature. We begin by giving the precise definition of the STPBVP.

Definition 4.1. A six-tuple $(\mathbb{P},B^1,B^2,V,X,\alpha)$ is called a (weak) solution of a stochastic Two-Point Boundary Value Problem (STPBVP) on $[0,T]$ if

(i) $\mathbb{P} \in \mathcal{P}(\Omega)$ and $B = (B^1,B^2)$ is a $\mathbb{P}$-Brownian motion on $[0,T]$;

(ii) $\alpha \in \mathcal{A}_{ad}$, and $(V,X,\alpha)$ satisfies the SDE on $(\Omega,\mathcal{F},\mathbb{P})$:

\[
\begin{align*}
&dV_t = b(t,V_t,X_t)dt + \sigma(t,V_t,X_t)dB^1_t, \quad V_0 = v; \\
&dX_t = (\mu(t,X_t) + \alpha_t\rho(t,X_t))dt + \rho(t,X_t)dB^2_t, \quad X_0 = x,
\end{align*}
\]

(iii) $\lim_{t \to T} [V_t - g(X_t)] = 0$, $\mathbb{P}$-a.s.;

In particular, $(V,X,\alpha)$ is called the solution to a Markovian STPBVP, if $\alpha_t = u(t,V_t,X_t)$, $t \in [0,T)$, for some measurable function $u$, and $(V,X)$ is an $\mathbb{F}^V,\mathbb{F}^X$-Markov process on $[0,T)$.

Remark 4.2. (i) For notational clarity we shall often refer to (4.1) as a “STPBVP($b,\sigma,\mu,\rho$)” when necessary, and write the solution $(V,X,\alpha)$ to a STPBVP as $(V^\alpha,X^\alpha)$ for convenience.

(ii) Comparing Definition 4.1 to that of a dynamic Markov bridge (see, e.g., [12]), we see that, if the coefficients $b$ and $\sigma$ are independent of $X$ and $\mu \equiv 0$, then a Markovian TPBVP is essentially a dynamic Markov bridge without requiring that $X$ be a local martingale with respect to its own filtration $\mathbb{F}^X$. Such a difference will make the results of this paper and those in the existing literature mutually exclusive.

To construct a weak solution, we first recall (3.17) and the $\mathbb{P}^\nu$-Brownian motion $W_t = B^0_t - \int_0^t \theta_s ds; t \in [0,T)$, where $\theta_t := (\theta^1_t,\theta^2_t)^T =\sigma^T(t,\xi_t)\nabla \ln \varphi(t,\xi_t) = \frac{1}{\varphi}(\partial_\nu\varphi\sigma,\partial_x\varphi\rho)^T(t,\xi_t)$, $t \in [0,T)$.
\[0, T\), and under \(\mathbb{P}^\nu\) the process \(\xi_t := (V_t, X_t)^T\) satisfies the SDE (3.18). We note that although the coefficient \(\hat{b}\) in (3.18) is explicitly defined, it depends on the solution of an ill-posed parabolic PDE (3.16), its behavior is a bit hard to analyze. The following lemma is useful to note.

**Lemma 4.3.** Let \((T, \xi_T, \nu)\) be the conditioning in Lemma 3.6, and \(\mathbb{P}^\nu\) the corresponding minimum probability. Then, it holds that \(L^p_t \subset L^p_T\), \(t < T\). Specifically, for any \(T_0 < T\), there exists a constant \(C_{T_0} > 0\), that depends only on the coefficients \((b, \sigma, \mu, \rho)_n\) and \(T_0\), such that, for any \(X \in \mathcal{F}_t\), \(t \in [0, T_0]\), it holds that

\[
\mathbb{E}^{\mathbb{P}^\nu}[|X|^p] \leq C_{T_0} \mathbb{E}^{Q^0}[|X|^p]. \tag{4.2}
\]

In particular, the \(Q^0\)-diffusion process \(\xi\) is well-defined for \(t \in [0, T]\) on the probability space \((\Omega, \mathcal{F}, \mathbb{P}^\nu)\), and \(\mathbb{P}^\nu\{\int_0^T |\xi_t|^2 dt < \infty\} = 1\), for any \(T_0 < T\).

**Proof.** We first note that under the conditioning in Lemma 3.6 (ii),

\[
\mathbb{E}^{\mathbb{P}^\nu}[|X|^p] = \mathbb{E}^{Q^0}[|X|^p] \leq C_{T_0} \mathbb{E}^{Q^0}[|X|^p],
\]

where \(C_{T_0} := \tilde{C}_{T_0} \int_{\mathbb{R}^d} e^{\frac{\lambda_0(y-y)^2}{2}} \nu(dy)\), proving (4.2). The rest of the proof is obvious. \(\blacksquare\)

Now for any \(n \in \mathbb{N}\), define the stopping time \(\tau_n := \inf\{t > 0 : |\xi_t| \geq n\} \wedge T\), and then denote \(\theta_t^{(n)} := \theta_{\tau_n \wedge \tau} t \in [0, T]\). Clearly, under probability \(\mathbb{P}^\nu\), for each \(n \in \mathbb{N}\), the SDE

\[
d\xi_t^{(n)} = [\hat{b}(t, \xi_t^{(n)}) + \sigma(t, \xi_t^{(n)})\theta_t^{(n)}]dt + \sigma(t, \xi_t^{(n)})dW_t, \quad \xi_0^{(n)} = z, \tag{4.3}
\]

is well-posed on \([0, T]\), and \(\xi_t^{(n)} \equiv \xi_t, t < \tau_n\). Furthermore, since \(\tau_n \nearrow T\) as \(n \to \infty\), we can simply define \(\xi_t = \xi_t^{(n)}, t \in [0, \tau_n]\), \(n = 1, 2, \ldots\). Then it holds that \(\lim_{t \nearrow T} \xi_t^{(n)} = \lim_{t \nearrow T} \xi_t = \xi_T\).

In particular, we have \(\mathbb{P}^\nu\{\lim_{t \nearrow T} V_t = \lim_{t \nearrow T} g(X_t)\} = 1\).

We now write \(\theta_t^{(n)} = (\theta_t^{(1,n)}, \theta_t^{(2,n)}), t \in [0, T]\). Since \(\theta_t^{(n)}\) is bounded by \(n\), and \(\theta_t^{(1,n)} = \theta_t^{(1,n+1)}, n \in [0, \tau_n]\). By Girsanov’s theorem, there exists a family of probabilities \(\{\mathbb{P}^{(n)}\}_{n \geq 1}\) on \((\Omega, \mathcal{F})\) by

\[
\frac{d\mathbb{P}^{(n)}}{d\mathbb{P}^\nu}|_{\mathcal{F}_{T_n}} = \mathcal{E}(\theta_T^{(1,n)}) := \exp \left\{ \int_0^T \theta_s^{(1,n)} dW_s - \frac{1}{2} \int_0^T |\theta_s^{(1,n)}|^2 ds \right\}.
\]

Then for each \(n \in \mathbb{N}\), the process \(\tilde{B}_t^{(n)} = (\tilde{B}_t^{(1,n)}, W_t^2) := (W_t^1 - \int_0^t \theta_s^{(1,n)} ds, W_t^2), t \in [0, T]\), is a 2-dim \(\mathbb{P}^{(n)}\)-Brownian motion. Moreover, by the property of \(\{\theta^n\}\), we must have

\[
\frac{d\mathbb{P}^{(n+1)}}{d\mathbb{P}^\nu}|_{\mathcal{F}_{T_n}} = \mathcal{E}(\theta_T^{(1,n+1)}) = \mathcal{E}(\theta_T^{(1,n)}) = \frac{d\mathbb{P}^{(n)}}{d\mathbb{P}^\nu}|_{\mathcal{F}_{T_n}}. \tag{4.4}
\]

Consequently, we have \(\mathbb{P}^{(n+1)}|_{\mathcal{F}_{T_n}} = \mathbb{P}^{(n)}|_{\mathcal{F}_{T_n}}\), and \(\tilde{B}_t^{(n+1)} = \tilde{B}_t^{(n)}, t \in [0, \tau_n]\), for each \(n \in \mathbb{N}\). Observing that \(\tau_n \nearrow T\) as \(n \to \infty\), we can define a new probability measure \(\mathbb{P}\) on \((\Omega, \mathcal{F}_{T-})\) by

\[
\mathbb{P}|_{\mathcal{F}_{T_n}} := \mathbb{P}^{(n)}|_{\mathcal{F}_{T_n}}, \quad n \in \mathbb{N}, \tag{4.5}
\]
then $\bar{P} \ll \mathbb{P}^\nu$ on $\mathcal{F}_t$, $t \in [0, T)$. Furthermore, if we define $\bar{B}_t = \bar{B}_t^{(n)}$, $t \in [0, \tau_n]$, $n \in \mathbb{N}$, then $\bar{B}$ is a $\bar{P}$-Brownian motion on $[0, T)$, and under $\bar{P}$, the process $\xi = (V, X)$ satisfies the SDE:

$$
\begin{align*}
&dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)d\bar{B}_t^1,
&dX_t = (\mu(t, X_t) + \rho(t, X_t)\theta_2^2)dt + \rho(t, X_t)dW_t^2, \\
&V_0 = v; \quad t \in [0, T).
\end{align*}
$$

(4.6)

Comparing (4.6) and (4.1) and noting the facts (3.13) and $\bar{P}|_{\mathcal{F}_T} \ll \mathbb{P}^\nu|_{\mathcal{F}_T}$, $t \in [0, T)$, we see that $(\bar{P}, \bar{B}, V, X, \theta^2)$ should be a weak solution to STPBPV (1.1). We have the following result.

**Proposition 4.4.** Assume Assumption 3.1. Then there exists a weak solution to STPBPV (4.7). Furthermore, if $(\mathbb{P}, B, V, X, \alpha)$ denotes the weak solution, then $\mathbb{P}$ can be chosen so that $\mathbb{P}|_{\mathcal{F}_T} \ll \mathbb{Q}^0|_{\mathcal{F}_T}$, $t < T$, and denoting $V_T := V_{T-} = \lim_{t \uparrow T} V_t$, it holds that $\mathbb{P} \circ (V_T)^{-1} = m^*$.

**Proof.** Consider the probability $\bar{P}$ defined by (4.3), (4.5) and SDE (4.6). We first claim $\bar{P} \ll \mathbb{P}^\nu$ on $\mathcal{F}_{T-}$. Indeed, let $\mathcal{A} := \{G \subset F : \mathbb{P} \ll \mathbb{P}^\nu \text{ on } G\}$, then $\mathcal{F}_\tau \in \mathcal{A}$, $n \in \mathbb{N}$. Since $\tau_n \not\uparrow T$, we have $\mathcal{F}_{T-} = \bigvee_n \mathcal{F}_\tau$ (see, e.g., [32]), and thus $\mathcal{F}_{T-} \in \mathcal{A}$, thanks to the Monotone Class Theorem.

Next, since $\{\lim_{t \uparrow T} V_t \neq \lim_{t \uparrow T} g(X_t)\} = \bigcup_m \bigcap_N \bigcup_{r \in \mathbb{R}_+} \{V_r - g(X_r) \geq \frac{1}{m}\} \in \mathcal{F}_{T-}$, where $\mathbb{Q}$ is the rationals in $\mathbb{R}_+$, and $\mathbb{Q}(A) := Q \cap A$, $A \in \mathcal{B}(\mathbb{R})$, and $\bar{P} \ll \mathbb{P}^\nu$ on $\mathcal{F}_{T-}$, we have $\bar{P}\{\lim_{t \uparrow T} V_t \neq \lim_{t \uparrow T} g(X_t)\} = 0$, thanks to (3.13). That is, $\bar{P}\{\lim_{t \uparrow T} V_t = \lim_{t \uparrow T} g(X_t)\} = 1$. Now let $\alpha = \theta^2$ in SDE (4.6), we see that $(\bar{P}, \bar{B}, V, X, \alpha)$ is a weak solution to STPBPV (1.1).

It remains to check the last statement. To this end, let $\xi = (V, X)$. Since $\bar{P} \ll \mathbb{P}^\nu \ll \mathbb{Q}^0$ on $\mathcal{F}_{T-}$ and $\mathbb{Q}^0\{\xi \in \mathbb{C}([0, T]; \mathbb{R}^2)\} = 1$, we can naturally extend $\xi$ to $[0, T]$ by setting $\xi_{T} = \lim_{t \uparrow T} \xi_t$ so that $\mathbb{P}^\nu\{\xi \in \mathbb{C}([0, T]; \mathbb{R}^2)\} = \bar{P}\{\xi \in \mathbb{C}([0, T]; \mathbb{R}^2)\} = 1$ as well. We first claim that $\mathbb{P}^\nu \circ V_T^{-1} = m^*$. Indeed, let $B \in \mathcal{B}(\mathbb{R})$ and $A := B \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$ by (3.11) we have $B = \{v : (v, g^{-1}(v)) \in A\}$, and $\mathbb{P}^\nu\{V_T \in B\} = \mathbb{P}^\nu\{(V_T, X_T) \in A\} = \nu\{A\} = m^*(B)$. That is, $\mathbb{P}^\nu \circ V_T^{-1} = m^*$.

To see $\bar{P} \circ V_T^{-1} = m^*$, we note that $\xi = (V, X)$ is the unique strong solution to SDE (3.1) under $\mathbb{Q}^0$ with canonical process $B^0 = (B^1, Y)$. Therefore we can write $\xi_t(\omega) = \Phi(t, B^0_t(\omega)) = \Phi(t, \omega)$, $(t, \omega) \in [0, T] \times \Omega$, for some (progressively) measurable function $\Phi : [0, T] \times \Omega \rightarrow \mathbb{R}^2$. Consequently, we can write $\theta_t^2(\omega) = \partial_x \ln \varphi(t, \xi_t(\omega)) = \partial_x \ln \varphi(t, \Phi(t, \omega))$, $(t, \omega) \in [0, T] \times \Omega$. By virtue of Lemma 4.3 the process $\theta_t^2$ is well-defined on $[0, T] \times \Omega$, $\bar{P}$-a.s. and $\theta_t^2 \in L^2(\bar{P})$, for $t \in [0, T)$.

Now let us denote the solutions to (3.18) and (4.6) as $(\bar{V}_t, \bar{X}_t)$ and $(\bar{V}_t, \bar{X}_t)$ respectively. Then we see that $(\bar{X}_t, W_t^2)$ and $(\bar{X}_t, W_t^2)$ are two weak solutions to the same SDE, well-defined on any $[0, T_0] \subset [0, T)$. Consequently, we have $\mathbb{P}^\nu \circ \bar{X}_T^{-1} = \bar{P} \circ \bar{X}_T^{-1}$ on $[0, T_0]$ for any $T_0 < T$. Extending the solution to $[0, T)$, we have $\mathbb{P}^\nu \circ \bar{X}_T^{-1} = \bar{P} \circ \bar{X}_T^{-1}$. Since $V_T = g(X_T)$, both $\bar{P}$-a.s. and $\mathbb{P}^\nu$-a.s., we obtain that $\bar{P} \circ V_T^{-1} = \mathbb{P}^\nu \circ V_T^{-1} = m^*$, proving the proposition. 

**Uniqueness in law.** Let us now turn to the issue of uniqueness. To begin with let us recall that the weak solution $(\bar{P}, \bar{B}, V, X, \alpha)$ that we constructed has the following properties:
Definition 4.5. We call a weak solution with

\[ \bar{P} \text{-a.s.} \]

where \( \bar{P} \) is an \( \bar{P} \)-announcing sequence of probability \( \bar{P} \) for an STPBVP (4.1), for each \( n \in \mathbb{N} \), where \( \bar{B}^{(n)} = (\bar{B}^{(n,1)}, \bar{B}^{(n,2)}) \) is a \( \bar{P} \)-Brownian motion on \( [0, T] \);

(iii) the solution \( (\bar{V}, \bar{X}) = (V^{(n)}, X^{(n)}) \) on \( [0, \tau_n] \), where \( (V^{(n)}, X^{(n)}) \) is a (pathwisely) unique solution to the following SDE, defined on \( [0, T] \):

\[
\begin{align*}
\left\{ 
\begin{array}{ll}
\ dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB^{(n,1)}_t, & V_0 = v; \\
\ dX_t = (\mu(t, X_t) + \rho(t, X_t)\alpha^{(n)}_t)dt + \rho(t, X_t)dB^{(n,2)}_t, & X_0 = x,
\end{array}
\right.
\end{align*}
\]  

(4.7)

where \( |\alpha^{(n)}_t| \leq M_n, t \in [0, T] \), for some \( M_n > 0 \); and \( \alpha^{(n+1)}_t = \alpha^{(n)}_t, t \in [0, \tau_n], \bar{P} \)-a.s.;

(iv) \( \bar{P}|_{\mathcal{F}_t} \ll \bar{P}'|_{\mathcal{F}_t} \ll \mathbb{Q}^0|_{\mathcal{F}_t}, t \in [0, T) \).

In what follows we shall denote \( (\bar{P}, \{\tau_n\}) \) to specify that \( \bar{P} \) is "announced" by \( \{\tau_n\} \), and make use of the following definitions in the spirit of the so-called "\( \mathbb{Q}^0 \)-weak solutions" in [29].

Definition 4.7. We call a weak solution \( (\bar{P}, \bar{V}, \bar{X}, \bar{B}, \alpha) \) of STPBVP (4.7) satisfying (i)–(iii) above a "nested weak solution" and the corresponding family of stopping times \( \{\tau_n\} \) the "announcing sequence" of probability \( \bar{P} \). We call \( \{\tau_n\}, \alpha \) the characteristic pair of the weak solution.

Furthermore, a nested weak solution is called a \( \mathbb{P}' \)-weak solution if it also satisfies (iv). 

Remark 4.6. Comparing to the usual SDEs, the characteristic pair \( \{\tau_n\}, \alpha \) is important in determining a solution to an STPBVP. Note that if \( \{\tau^1_n\}, \{\tau^2_n\} \) are two announcing sequences of stopping times, then so is \( \{\tau^1_n \wedge \tau^2_n\} \). Thus the weak solution is independent of the choice of the announcing sequence \( \{\tau_n\} \). Since the process \( \alpha \) determines the coefficient of SDE (4.6), whence the solution, we often specify its role by calling \( (\bar{P}, \bar{V}, \bar{X}, \bar{B}, \alpha) \) the \( \alpha \)-weak solution.

Definition 4.7. We say that the pathwise uniqueness holds for STPBVP (4.1), if for two nested solutions \( (\mathbb{P}^i, \xi^i = (V^i, X^i), B^i, \alpha^i), i = 1, 2 \) of (4.7) on \( [0, T) \), such that \( \mathbb{P}^1 = \mathbb{P}^2 = \mathbb{P}, \xi^1_0 = \xi^2_0 \), and \( \mathbb{P}\{\alpha^1_t = \alpha^2_t, B^1_t = B^2_t, t \in [0, T)\} = 1 \), then \( \mathbb{P}\{\xi^1_t = \xi^2_t, t \in [0, T_0)\} = 1 \), for any \( T_0 < T \).

Remark 4.8. We note that the time \( T_0 \) in Definition 4.7 can be changed to any stopping time \( \tau \) with \( \mathbb{P}\{\tau < T\} = 1 \). In fact, the following two statements are equivalent: (i) the pathwise uniqueness holds on \( [0, T_0] \), for any \( T_0 < T \); and (ii) there exists a sequence of stopping time \( \{\tau_n, n \geq 1\}, \lim_{n \to \infty} \tau_n = T \) almost surely, such that the pathwise uniqueness holds on \( [0, \tau_n] \), for each \( n \geq 1 \). Indeed, let \( (\mathbb{P}^i, \xi^i = (V^i, X^i), i = 1, 2 \) be two nested solutions as in Definition 4.7 and denote \( \Delta \xi := \xi^1_t - \xi^2_t \), then we obtain

\[
\mathbb{E}\left[ \sup_{t \in [0, T_0]} |\Delta \xi| \right] \leq \mathbb{E}\left[ \sup_{t \in [0, \tau_n]} |\Delta \xi| 1_{(\tau_0 \leq \tau_n)} \right] + \mathbb{E}\left[ \sup_{t \in [0, T_0]} |\Delta \xi| 1_{(\tau_0 > \tau_n)} \right] \leq \mathbb{E}\left[ \sup_{t \in [0, T_0]} |\Delta \xi| 1_{(\tau_0 > \tau_n)} \right];
\]

17
and similarly, for any $T_0 < T$,

$$
\mathbb{E} \left[ \sup_{t \in [0, \tau]} |\Delta \xi| \right] \leq \mathbb{E} \left[ \sup_{t \in [0, T_0]} |\Delta \xi| 1_{\{\tau \leq T_0\}} \right] + \mathbb{E} \left[ \sup_{t \in [0, \tau]} |\Delta \xi| 1_{\{\tau > T_0\}} \right] \leq \mathbb{E} \left[ \sup_{t \in [0, \tau]} |\Delta \xi| 1_{\{\tau > T_0\}} \right].
$$

Since $\lim_{n \to \infty} \mathbb{P}\{T_0 > \tau_n\} = 0$ and $\lim_{T_0 \to T} \mathbb{P}\{\tau > T_0\} = 0$, it is readily seen that the statements (i) and (ii) above are equivalent, and $T_0$ in Definition 4.7 can be replaced by any $\tau < T$. 

The definition of the uniqueness in law for the STPBVP is a bit more involved. First note that the component “$\alpha$” of the solution is part of the drift coefficient of the SDE (4.6), and in general it is not unique. Thus the uniqueness of the solution, even in the weak sense, depends on how the process $\alpha$ is properly fixed. To this end, denote $A := \{ A \in \mathcal{B}([0, T]) \otimes F : A_t \in \mathcal{F}_t, \ t \in [0, T]\}$, where $A_t$ is the $t$-section of $A$; and denote all $A$-measurable functions by $L^0_{A}(\{0, T] \times \Omega)$. We should note that the space $L^0_{A}(\{0, T] \times \Omega)$ is independent of any probability measure, and we can therefore use it to identify the $\alpha$-component of the solution in an “universal” way.

**Definition 4.9.** We say that the nested weak solution to the STPBVP (4.1) is unique in law, if for any two $\alpha$-weak solutions $(\bar{P}^1, V^1, \bar{X}^1, B^1, \bar{\alpha}^1), \ i = 1, 2$ of (4.1) on $[0, T]$, such that $(v^1, x^1) = (v^2, x^2); \bar{P}^1 \circ (\tau_n)^{-1} = \bar{P}^2 \circ (\tau_n)^{-1}, n \in \mathbb{N};$ and $\bar{P}^i\{\bar{\alpha}^i_t = \alpha_t, \ t \in [0, T]\} = 1, \ i = 1, 2$, for some $\alpha \in L^0_{A^i}(\{0, T] \times \Omega)$, then for any cylindrical set $E^{A_1, \cdots, A_n}_{t_1, \cdots, t_n} := \{(v, x) \in C([0, T]; \mathbb{R}^2) : (v, x)(t) \in A_t, \ i = 1, \cdots, n\}, \ 0 \leq t_1 < t_2 < \cdots < t_n < T$ and $A_i \in \mathcal{B}(\mathbb{R}^2), \ i = 1, \cdots, n$, it holds that

$$
\bar{P}^1 \circ (\bar{V}^1, \bar{X}^1)^{-1}\{E^{A_1, \cdots, A_n}_{t_1, \cdots, t_n}\} = \bar{P}^2 \circ (\bar{V}^2, \bar{X}^2)^{-1}\{E^{A_1, \cdots, A_n}_{t_1, \cdots, t_n}\}.
$$

We now give the main theorem of this subsection.

**Proposition 4.10.** Assume Assumption [3.1]. Then, the Markovian $P^\nu$-weak solution to STPBVP (4.1) is unique in law.

Before we prove Proposition 4.10, we first prove a lemma that is interesting in its own right.

**Lemma 4.11.** Assume Assumption [3.1], and let $(\bar{P}, \bar{\xi}, \bar{\alpha})$ be a nested Markovian weak solution with $\bar{\alpha}_t = u(t, \bar{\xi}_t), \ u \in L^0([0, T] \times \mathbb{R}^2), \ such that \mathbb{P}\{\bar{\alpha}_t = \alpha_t, \ t \in [0, T]\} = 1$ for some $\alpha \in L^0_{A^i}([0, T] \times \Omega)$. Then $\alpha_t(\omega) = u(t, \Phi(t, \omega)), \ dt \otimes d\mathbb{P}$-a.e.-$\omega \in (0, T] \times \Omega$, for some $\Phi \in L^0_{A^i}([0, T] \times \Omega)$.

**Proof.** Let $(\bar{P}, \bar{\xi}, \bar{\alpha})$ be the nested Markovian weak solution. Then $\bar{\alpha}_t = u(t, \bar{\xi}_t), \ t \in [0, T], \ for some u \in L^0([0, T] \times \mathbb{R}^2)$. By Definition 4.5, the solution $\bar{\xi}$ is the pathwisely unique weak solution of SDE (4.7) on any $[0, \tau_n], n \geq 1$, whence on $[0, T_0]$, for any $T_0 < T$, thanks to Remark 4.8. Thus, by Yamada-Watanabe theorem, for any $T_0 < T$, $\xi$ is the pathwisely unique strong solution on $[0, T_0]$, and there exists a $\Phi^0 \in L^0_{A^i}([0, T_0] \times \Omega), \ such that \bar{\xi}_t = \Phi^0(t, \cdot), \ t \in [0, T_0]$. 18
As before, we can define a $\Phi \in \mathbb{L}^0([0, T] \times \Omega)$ so that $\Phi(t, \cdot) = \Phi_T^n(t, \cdot)$, $t \in [0, T_n]$, for any sequence $T_n \nearrow T$, and $\xi_t = \Phi(t, \cdot)$, $t \in [0, T)$, $\bar{\mathbb{P}}$-a.s.. Since $\bar{\alpha}_t = u(t, \xi_t) = u(t, \Phi(t, \cdot))$, $t \in [0, T)$, by assumption, we have $\alpha_t(\omega) = \bar{\alpha}_t(\omega) = u(t, \Phi(t, \omega))$, $dt \otimes d\bar{\mathbb{P}}$-a.e., proving the lemma.

[Proof of Proposition 4.10] Let $(\bar{\mathbb{P}}^i, \bar{\xi}^i_t = (\bar{V}^i, \bar{X}^i), \bar{B}^i, \alpha^i)$, $i = 1, 2$, be two Markovian weak solutions of (4.11) on $[0, T)$, with characteristic pair $\{\tau^i_m\}, \alpha^i$, $i = 1, 2$. Without loss of generality, we assume that $\{\tau^i_m\}$ is the exit time of $\alpha^i = u(t, \bar{\xi}^i_t)$, $i = 1, 2$, from the interval $[-m, m]$.

Next, let the cylindrical set $E^A_{t_1, \ldots, t_n}$ be given, with $t_n < T$. Since $\tau^i_m \nearrow T$, we can write

$$
(\bar{\xi}^i_0)^{-1}(E^A_{t_1, \ldots, t_n}) = \bigcap_{j=1}^n (\xi^i_j)^{-1}(A_j) = \bigcup_{m=1}^n \bigcap_{j=1}^n \{\tau^i_m \geq t_j\} \cap (\xi^i_j)^{-1}(A_j), \quad i = 1, 2.
$$

Denoting $E^i_{j,m} := \{\tau^i_m \geq t_j\} \cap (\xi^i_j)^{-1}(A_j) = \{\tau^i_m \geq t_j\} \cap (\xi^i_j)^{-1}(A_j)$, $i=1,2$, we claim that $E^i_{j,m} \in \mathcal{F}_{\tau^i_m}$, for each $i, j, m$. Indeed, fix $i, j$, and $m$, for $t \in [0, T)$, one has

$$
\{\tau^i_m \leq t\} \cap E^i_{j,m} = \{t_j \leq \tau^i_m \leq t\} \cap (\xi^i_j)^{-1}(A_j) \in \mathcal{F}_t, \quad i = 1, 2.
$$

That is, $E^i_{j,m} \in \mathcal{F}_{\tau^i_m}$, whence $\hat{E}^i_m := \bigcap_{j=1}^n E^i_{j,m} \in \mathcal{F}_{\tau^i_m}$, $i = 1, 2$.

On the other hand, note that the set $\hat{E}^i_m$ is increasing in $m$, thanks to the extension nature of solutions $\hat{\xi}^i(m)$. Thus, noting that $\bar{\mathbb{P}}^i|_{\mathcal{F}_{\tau^i_m}} = \bar{\mathbb{P}}^i(m)|_{\mathcal{F}_{\tau^i_m}}$, for $i = 1, 2$, we have

$$
\bar{\mathbb{P}}^i \circ (\bar{\xi}^i)^{-1}(E^A_{t_1, \ldots, t_n}) = \bar{\mathbb{P}}^i \{ \bigcup_{m=1}^n \bigcap_{j=1}^n E^i_{j,m} \} = \bar{\mathbb{P}}^i \{ \bigcup_{m=1}^n \hat{E}^i_m \} = \lim_{m \to \infty} \bar{\mathbb{P}}^i \{ \hat{E}^i_m \} = \lim_{m \to \infty} \bar{\mathbb{P}}^i(m) \{ \hat{E}^i_m \}. \tag{4.8}
$$

Now, by Lemma 4.11, for two Markovian weak solutions satisfying $\bar{\mathbb{P}}^i \{ \bar{\alpha}^i_t = \alpha_t, \ t \in [0, T) \} = 1$, $i = 1, 2$, we must have $\bar{\alpha}^1_t = \bar{\alpha}^2_t = \alpha_t = u(t, \Phi(t, \cdot))$, $t \in [0, T)$, $\bar{\mathbb{P}}^1, \bar{\mathbb{P}}^2$-a.s. for some functions $u \in \mathbb{L}^0([0, T] \times \mathbb{R}^2)$ and $\Phi \in \mathbb{L}^0_d([0, T] \times \Omega)$. In other words, $\bar{\mathbb{P}}^i(m), \bar{\xi}^i(m), i = 1, 2$, satisfy the same SDE (4.11) on $[0, \tau^i_m]$ with the same coefficients induced by a (bounded) process $\alpha(m)$, for which the pathwise uniqueness holds. We conclude that $\bar{\mathbb{P}}^1(m) \circ (\bar{\xi}^1(m))^{-1} = \bar{\mathbb{P}}^2(m) \circ (\bar{\xi}^2(m))^{-1}$. Note that $\{\tau^i_m \geq t_j\} = \{u(t_j, \bar{\xi}^i(m)) \leq m\}$, we see that $\bar{\mathbb{P}}^1(m) \{ \hat{E}^1_m \} = \bar{\mathbb{P}}^2(m) \{ \hat{E}^2_m \}$, $m \in \mathbb{N}$, and the result follows from (4.8).

5 Affine Structure of Insider Strategy

In the rest of the paper we shall use the STPBVP to construct the equilibrium strategy. Note that the solution to STPBVP (4.1) depends on the “pricing rule” $(\mu, \rho)$, we first argue that $(\mu, \rho)$ can be chosen so that the equilibrium strategy takes a particular form. Specifically, from Propositions 3.5 and 4.6 we see that the $\alpha$-component in a weak solution is closely related to an ill-posed parabolic PDE (3.16), and in light of the well-known Widder’s Theorem and its
Plugging (5.3) into (5.5) we obtain extensions (cf. e.g., [4, 24, 34, 35]), we may assume that \( \varphi(t, v, x) = \exp\{I(t, v, x)\} \), where \( I(t, \cdot, \cdot) \) is quadratic in \( (v, x) \). Thus, if a Markovian strategy \( \bar{\alpha}_t = u(t, \Phi(t, \cdot)) \) (see Remark 4.11), then

\[
\begin{align*}
    u(t, v, x) &= \rho(t, x)\partial_x \ln \varphi = u_0(t, x) + u_1(t, x)v, \quad (t, v, x) \in [0, T) \times \mathbb{R}^2, \\
    &\text{for some functions } u_0, u_1 : [0, T] \times \mathbb{R} \to \mathbb{R} \text{ to be determined later. In what follows we call a function } u \text{ of the form (5.1) as having an Affine Structure.}
\end{align*}
\]

We should note that the affine structure of the insider strategy has been widely observed in the literature. In particular, the equilibrium strategy of the form

\[
\alpha_t = \beta_t(V_t - P_t), \quad t \in [0, T),
\]

where \( \beta = \{\beta_t\} \) is a deterministic function known as the “trading intensity”, can be found in many static information case (see, e.g., [1, 26]), as well as dynamic information case (see, e.g., [29]). The general form in (5.1) can also be found in [4, 5]. In order to validate the affine structure, let us begin with some simple analysis.

Assume, for example, that a solution to the STPBVP (4.6) is such that \( \bar{\alpha}_t = u(t, \bar{V}_t, \bar{X}_t) \), where \( u(t, v, x) \) satisfies (5.1), then the function \( \varphi \) must have the form \( \varphi(t, v, x) = \exp\{I(t, v, x)\} \), where

\[
I(t, v, x) = h(t, v) + A(t, x) + B(t, x)v,
\]

and \( A(t, x) \) and \( B(t, x) \) are defined respectively by

\[
A(t, x) := \int_0^x \frac{u_0(t, y)}{\rho(t, y)} dy; \quad B(t, x) := \int_0^x \frac{u_1(t, y)}{\rho(t, y)} dy, \quad h(t, v) := \ln \varphi(0, v).
\]

Now assume that \( \varphi \) satisfies the PDE (3.16), then we can derive a PDE for the function \( I \):

\[
\begin{align*}
    &\left\{ \begin{array}{l}
    I_t + b(t, v, x)I_v + \mu(t, x)I_x + \frac{1}{2}\sigma^2(t, v, x)\{[I_v]^2 + I_{vv}\} + \frac{1}{2}\rho^2(t, x)\{I_x^2 + I_{xx}\} = 0; \\
    I(0, v, x) = h(0, v) + A(0, x) + B(0, x)v.
    \end{array} \right.
\end{align*}
\]

Plugging (5.3) into (5.5) we obtain

\[
0 = \frac{1}{2}\rho^2(t, x)B_x^2v^2 + \{B_t + \mu(t, x)B_x + \frac{1}{2}\rho^2(t, x)\{B_{xx} + A_xB_x\}\}v + A_t + \mu(t, x)A_x \\
+ \frac{1}{2}\rho^2(t, x)\{A_{xx} + A_x^2\} + h_t + b(t, v, x)[h_v + B] + \frac{1}{2}\sigma^2(t, v, x)\{h_{vv} + [h_v + B]^2\}.
\]

For notational simplicity, for given coefficients \( b, \sigma, \mu, \rho \), we define

\[
\begin{align*}
    I_0(t, x) &= I_0(t, x; \mu, \rho) = A_t + \mu(t, x)A_x + \frac{1}{2}\rho^2(t, x)\{A_{xx} + A_x^2\}; \\
    I_1(t, x) &= I_1(t, x; \mu, \rho) = B_t + \mu(t, x)B_x + \frac{1}{2}\rho^2(t, x)\{B_{xx} + A_xB_x\}; \\
    I_2(t, x) &= I_2(t, x; \mu, \rho) = \frac{1}{2}\rho^2(t, x)B_x^2; \\
    G(t, v, x) &= h_t(t, v) + b(t, v, x)[h_v(t, v) + B] + \frac{1}{2}\sigma^2(t, v, x)\{h_{vv}(t, v) + [h_v(t, v) + B]^2\}.
\end{align*}
\]
Then, (5.6) becomes

\[ I_2(t,x)v^2 + I_1(t,x)v + I_0(t,x) + G(t,v,x) = 0, \quad (t,v,x) \in [0,T] \times \mathbb{R}^2. \]

(5.8)

We thus obtained the following result for affine structure of function \( u \).

**Proposition 5.1.** The function \( u(t,v,x) = \rho(t,x) \partial_x \ln \varphi(t,v,x) \) has an affine structure (5.1), where \( \varphi \) solves (3.16), if and only if the coefficients \( b, \sigma, \mu, \rho \) satisfy the compatibility conditions (5.8) with \( I_0 - I_2 \) and \( G \) being defined respectively by (5.7).

Furthermore, it holds that \( \partial_{vvv} G(t,v,x) \equiv 0, \quad (t,v,x) \in [0,T] \times \mathbb{R}^2 \).

We should note that the compatibility condition (5.8) is technically difficult to verify in general, as it involves not only a fairly complicated systems of differential equations, but also the selection of the “pricing rule” \( (\mu, \rho) \). In what follows we impose some specific structures on the functions \( h, b, \sigma \), and try to find the conditions under which the function \( u(t,v,x) \) is of an affine structure.

Let us begin with an example of a Kyle-Back problem with dynamic information that fits the generality considered in this paper, and justifies the validity of the compatibility condition.

**Example 5.2.** Consider the Kyle-Back problem studied in [29]. More precisely, we assume that

\[ b(t,v,x) = f(t,v)x + g(t,v)x + k(t), \quad \sigma(t,v,x) = 1. \]

Denote now \( X_t = P_t = \mathbb{E}^{\mathbb{P}}[V_t|\mathcal{F}_t^Y] \). Then, by [29, Theorem 3.6], we have

\[ \mu(t,x) = (f(t,v)x + k(t), \quad \rho(t,x) = \rho(t) = S_t^\beta_t, \]

where \( S_t \) satisfies a (deterministic) Riccati equation. Furthermore, in [29] it was shown that the equilibrium strategy takes the form (5.2). That is, the equilibrium \( \alpha \) has an affine structure (5.1) with \( u_0(t,x) = -\beta_t x, u_1(t,x) = \beta_t \). By definition (5.4) we then have

\[
\begin{align*}
A(t,x) &= \int_0^x \frac{u_0(t,y)}{\rho(t,y)} dy = -\frac{1}{S_t} \int_0^x ydy = -\frac{x^2}{2S_t}; \\
B(t,x) &= \int_0^x \frac{u_1(t,y)}{\rho(t,y)} dy = \int_0^x \frac{1}{S_t} dy = \frac{x}{S_t}.
\end{align*}
\]

Plugging these into (5.7) and using the fact that \( S \) satisfies the Riccati equation \( \frac{dS_t}{dt} = 2f(t,S_t - \beta_t^2 S_t^2 + 1, t \in [0,T), \) one can check that the compatibility condition (5.8) holds.

In the general nonlinear case, the analysis becomes too complicated to have a generic result. We therefore consider several special cases that might be useful in practice.
Case 1. $h = h(t), b(t,x,v) = b(t,x) \text{ and } \sigma(t,v,x) = \sigma(t,x)$. In this case, \((5.6)\) is reduced to
\[
I_0(t,x) + I_1(t,x)v + I_2(t,x)v^2 = 0, \tag{5.9}
\]
where (suppressing variables)
\[
I_0 = \partial_t h + bB + \frac{1}{2} \sigma^2 B^2 + \partial_t A + \mu \partial_x A + \frac{1}{2} \rho^2 (\partial_{xx} A + (\partial_x A)^2),
\]
\[
I_1 = \partial_t B + \mu \partial_x B + \frac{1}{2} \rho^2 (\partial_{xx} B + \partial_x A \partial_x B), \quad I_2 = \frac{1}{2} \rho^2 (\partial_x B)^2.
\]
Clearly, \((5.9)\) implies that $I_0 = I_1 = I_2 = 0$. Then, by definition we have $\partial_t B = \frac{u_1(t,x)}{\rho(t,x)} = 0$, which implies that $u_1(t,x) \equiv 0$. Consequently, $B(t,x) \equiv 0$. It then follows that
\[
\partial_t h + \partial_t A + \mu \partial_x A + \frac{1}{2} \rho^2 (\partial_{xx} A + (\partial_x A)^2) = 0. \tag{5.10}
\]
That is, a necessary condition for affine structure is that $u_1 \equiv 0$ and $h, u_0, b, \sigma, \mu, \rho$ satisfy \((5.10)\).

Case 2. $h = h(t), b(t,v,x) = b_0(t,x) + b_1(t,x)v, \sigma(t,v,x) = \sigma_0(t,x) + \sigma_1(t,x)v$. Then, similar to Case 1, we simplify the equation \((5.6)\) and denote the coefficients as $I_0, I_1, I_2$, where
\[
\begin{cases}
I_0 = \partial_t h(t) + b_0 B + \frac{1}{2} \sigma_0^2 B^2 + \partial_t A + \mu \partial_x A + \frac{1}{2} \rho^2 (\partial_{xx} A + (\partial_x A)^2); \\
I_1 = b_1 B + \sigma_0 \sigma_1 B^2 + \partial_t B + \mu \partial_x B + \frac{1}{2} \rho^2 (\partial_{xx} B + \partial_x A \partial_x B); \\
I_2 = \frac{1}{2} \rho^2 (\partial_x B)^2 + \frac{1}{2} \sigma_1^2 B^2.
\end{cases}
\]
We see from $I_2 = 0$ that $u_1 \equiv 0$, which again leads to \((5.10)\).

Case 3. $h = h(t), b(t,v,x) = b_0(t,x) + b_1(t,x)v + b_2(t,x)v^2, \sigma(t,v,x) = \sigma_0(t,x) + \sigma_1(t,x)v$. In this case we need
\[
\begin{cases}
I_0 = \partial_t h(t) + b_0 B + \frac{1}{2} \sigma_0^2 B^2 + \partial_t A + \mu \partial_x A + \frac{1}{2} \rho^2 (\partial_{xx} A + (\partial_x A)^2) = 0; \\
I_1 = b_1 B + \sigma_0 \sigma_1 B^2 + \partial_t B + \mu \partial_x B + \frac{1}{2} \rho^2 (\partial_{xx} B + \partial_x A \partial_x B) = 0; \\
I_2 = \frac{1}{2} \rho^2 (\partial_x B)^2 + \frac{1}{2} \sigma_1^2 B^2 + b_2 B = 0. \tag{5.11}
\end{cases}
\]
In particular, $I_2 = 0$ if and only if
\[
u_1^2(t,x) = -\sigma_1^2 \left( \int_{x_0}^{x} \frac{u_1(t,y)}{\rho(t,y)} dy \right)^2 - 2b_2 \int_{x_0}^{x} \frac{u_1(t,y)}{\rho(t,y)} dy. \tag{5.12}
\]
If we choose $u_1 = \rho$, then \((5.12)\) implies $\rho^2 = -\sigma_1^2 (x-x_0)^2 - 2b_2(x-x_0)$. Using $I_1 = 0$ in \((5.11)\), we can write $u_0$ as
\[
u_0 = \frac{2}{u_1} \left[ - \partial_t B - \mu \partial_x B - \frac{1}{2} \rho^2 \partial_{xx} B - b_1 B - \sigma_0 \sigma_1 B^2 \right]. \tag{5.13}
\]
h is quadratic, we conclude that

Since the function

optional projection

In fact, in the dynamic information case where the market price is in the form of an affine structure

particularly effective in determining the equilibrium strategy, which we now explain.

Therefore, (5.11), together with (5.12), (5.13), guarantees the affine structure in this case.

Case 4. \( h(t,v) = h_0(t) + h_1(t)v, b, \sigma \) same as Case 3. In this case,

\[
\begin{align*}
I_0 & = \partial_t h_0 + b_0(h_1 + B) + \frac{1}{2}\sigma_0^2(h_1 + B)^2 + \partial_t A + \mu \partial_x A + \frac{1}{2}\rho^2(\partial_{xx} A + (\partial_x A)^2) = 0; \\
I_1 & = \partial_t h_1 + b_1(h_1 + B) + \sigma_0 \sigma_1(h_1 + B)^2 + \partial_t B + \mu \partial_x B + \frac{1}{2}\rho^2(\partial_{xx} B + \partial_x A \partial_x B) = 0; \\
I_2 & = \frac{1}{2}\sigma^2(\partial_x B)^2 + \frac{1}{2}\sigma_1^2(h_1 + B)^2 + b_2(h_1 + B) = 0.
\end{align*}
\]

Case 5. \( h = h_0(t) + h_1(t)v + h_2(t)v^2 \). Since there are the terms \( bh_v, \sigma^2 \partial_v h^2 \) in \( G(t,v,x) \), and \( h \) is quadratic, we conclude that \( \sigma(t,v,x) \) must be independent of \( v \), and \( b \) has to be linear in \( v \).

We thus assume that \( b = b_0(t,x) + b_1(t,x)v, \sigma = \sigma(t,x) \), in other words,

\[
\begin{align*}
I_0 & = \partial_t h_0 + b_0(h_1 + B) + \frac{1}{2}\sigma^2[2h_2 + (h_1 + B)^2] + \partial_t A + \mu \partial_x A + \frac{1}{2}\rho^2(\partial_{xx} A + (\partial_x A)^2) = 0; \\
I_1 & = \partial_t h_1 + 2b_0h_2 + b_1(h_1 + B) + 2\sigma^2(h_1 + B)h_2 + \partial_t B + \mu \partial_x B + \frac{1}{2}\rho^2(\partial_{xx} B + \partial_x A \partial_x B) = 0; \\
I_2 & = \partial_t h_2 + 2b_1h_2 + 2\sigma^2h_2^2 + \frac{1}{2}\rho^2(\partial_x B)^2 = 0.
\end{align*}
\]

6 The Filtering Problem and FBSDE under Affine Structure

A popular approach in studying Kyle-Back equilibrium problem is nonlinear filtering (cf. e.g., [17,17,29]). In fact, in the dynamic information case where the market price is in the form of an optional projection: \( P_t = \mathbb{E}[V_t|F^Y_t], t \in [0,T] \), we believe that the filtering approach should be particularly effective in determining the equilibrium strategy, which we now explain.

We begin by recasting the STPBVP (4.1) as a nonlinear filtering problem. Consider a (Markovian) weak solution \((\mathbb{P}, V, X, B, \alpha)\) to STPBVP (4.1), where \( \alpha_t = u(t,V_t,X_t) \), and under \( \mathbb{P} \),

\[
\begin{align*}
\text{d}V_t &= b(t,V_t,X_t)\text{d}t + \sigma(t,V_t,X_t)\text{d}B^1_t, \quad V_0 = v_0; \\
\text{d}X_t &= [\mu(t,X_t) + \rho(t,X_t)u(t,V_t,X_t)]\text{d}t + \rho(t,X_t)\text{d}B^2_t, \quad X_0 = x_0; \quad (6.1) \\
\text{d}Y_t &= u(t,V_t,X_t)\text{d}t + \text{d}B^2_t, \quad Y_0 = 0.
\end{align*}
\]

Since the function \( u \) is now fixed, (6.1) can be thought of as a nonlinear filtering problem with correlated noises, in which \( (V,X) \) is the signal process and \( Y \) is the observation process. The only technical problem, however, is whether the function \( u \) satisfies usual technical requirements so that the SDE for \( P_t = \mathbb{E}[V_t|F^Y_t] \) (known as the Fujisaki-Kallianpur-Kunita (FKK) equation [23]) holds.

To this end, we assume that \( u(t,v,x) \) has the affine structure: \( u(t,v,x) = u_0(t,x) + u_1(t,x)v \).

Denoting \( \alpha_t = u(t,V_t,X_t) \), and consider the SDE:

\[
\text{d}M_t = -\alpha_t M_t \text{d}B^2_t, \quad M_0 = 1, \quad t \in [0,T]. \tag{6.2}
\]

The following result is a modification of [8, Lemma 4.1.1] to the current case.
Proposition 6.1. Assume Assumptions \(\textbf{A}\) and assume further that the function \(u\) in (6.1) satisfies \(|u(t, v, x)| \leq K(t)(1 + |v| + |x|),\) for \((t, v, x) \in [0, T) \times \mathbb{R}^2,\) for some function \(K \in L^2([0, T]; \mathbb{R}_+).\) Then, the solution \(M\) to (6.2) is a true martingale on \([0, T].\)

Proof. Clearly, \(M\) is a local martingale. Then, by Fatou’s lemma, for any \(t \in [0, T],\) we have
\[
E[M_t] \leq \lim_{n \to \infty} E[M_{t \wedge \tau_n}] = E[M_0] = 1,
\]
where \(\{\tau_n\}\) is any announcing sequence for \(M,\) and \(M\) is a true martingale iff \(E[M_t] = 1, 0 \leq t \leq T,\) which we now prove. For any \(\varepsilon > 0,\) define \(f_\varepsilon := \frac{x}{1 + \varepsilon x},\) and \(M_t^\varepsilon := f_\varepsilon(M_t), t \in [0, T].\) Clearly, by bounded convergence theorem, we have
\[
\lim_{\varepsilon \to 0} E[M_t^\varepsilon] = E[M_t].
\]
On the other hand, by a simple application of Itô’s formula and then taking expectation one has
\[
E[M_t^\varepsilon] := \frac{1}{1 + \varepsilon} - \mathbb{E}\left[\int_0^t G^\varepsilon(\alpha_s, M_s)ds\right], t \in [0, T],
\]
where \(G^\varepsilon(\alpha, x) := \frac{\varepsilon^2 x^2}{(1 + \varepsilon x)^2}.\) It is easy to check that there exists \(C > 0,\) such that \(|G^\varepsilon(\alpha, x)| \leq C\alpha^2 x,\) for all \(\varepsilon, x > 0.\) Denoting \(U_t := M_t(V_t^2 + X_t^2),\) then the linear growth assumption for \(\alpha_t\) gives
\[
E[G^\varepsilon(\alpha_t, M_t)] \leq CE[\alpha^2 M_t] \leq CK^2(t)[1 + E[U_t]].
\]
We claim that \(\sup_{t \in [0, T]} E[U_t] < \infty.\) The result then follows easily from the Dominated Convergence theorem. Applying Itô’s formula to \(U_t\) and \(f_\varepsilon(U_t),\) we have (denoting \(|\xi_0|^2 = v_0^2 + x_0^2)\)
\[
f_\varepsilon(U_t) = \frac{|\xi_0|^2}{1 + \varepsilon|\xi_0|^2} + \int_0^t \frac{2M_s[V_s b_s + X_s \mu_s + \frac{1}{2}(\sigma_s^2 + \rho_s^2)]ds}{(1 + \varepsilon U_s)^2} + \int_0^t \frac{2M_s V_s \sigma_s}{(1 + \varepsilon U_s)^2}dB_s^1
\]
\[+ \int_0^t \frac{-4V_s^2 \sigma_s^2 M_s^2 + (2M_s X_s \rho_s - U_s \alpha_s)^2}{(1 + \varepsilon U_s)^3}ds + \int_0^t \frac{-U_s \alpha_s + 2M_s X_s \rho_s}{(1 + \varepsilon U_s)^2}dB_s^2.
\]
Taking expectation on both sides, and by the linear growth of \(b, \sigma, \mu\) and \(\rho,\) we obtain
\[
E[f_\varepsilon(U_t)] \leq |\xi_0|^2 + \int_0^t \mathbb{E}\left[\frac{2M_s[V_s b_s + X_s \mu_s + \frac{1}{2}(\sigma_s^2 + \rho_s^2)]}{(1 + \varepsilon U_s)^2}\right]ds \leq |\xi_0|^2 + \int_0^t L(E[f_\varepsilon(U_t)] + 1)ds.
\]
Now, first applying Gronwall’s inequality and then applying Fatou’s lemma (sending \(\varepsilon \to 0),\) we deduce that \(\sup_{t \in [0, T]} E[U_t] < \infty,\) proving the claim.

We should note that with Proposition 6.1 and the affine structure assumption on \(u\) the SDE (6.1) can be naturally extended to \([0, T],\) and we can follow the same argument of [23 Theorem 4.1] to derive the FKK equation for \(P_t = \mathbb{E}^t[U_t|\mathcal{F}_t^\varepsilon],\) which takes the following form:
\[
\begin{cases}
\quad dP_t = \mathbb{E}^t[b(t, V_t, X_t)] - \mathbb{E}^t[u(t, V_t, X_t)]Z_t dt + Z_t dY_t, \\
Z_t := \mathbb{E}^t[U_t u(t, V_t, X_t)] - P_t \mathbb{E}^t[u(t, V_t, X_t)],
\end{cases} t \in [0, T],
\]
(6.4)
where \( \mathbb{E}[^{·}]:=\mathbb{E}_{\tilde{P}}[^{·}|\mathcal{F}_t^\gamma] \), \( t \in [0,T] \). Now if we assume that the coefficient \( b(\cdots) \) is also of affine structure: \( b(t,v,x) = b_0(t,x) + b_1(t,x)v \), and \( X \) is \( \mathbb{P}^Y \)-adapted, then (6.4) can be rewritten as

\[
dP_t = \{b_0(t,X_t) + b_1(t,X_t)P_t - (u_0(t,X_t) + u_1(t,X_t)P_t)Z_t\}dt + Z_t dY_t, \quad t \in [0,T]. \tag{6.5}
\]

Let us now choose \( \alpha_t = u(t,V_t,X_t), \ t \in [0,T], \) to be the \( \alpha \)-component of a Markovian weak solution to the STPBVP (4.1), and assume that it has the affine structure. By Proposition 6.1, the process \( M \) defined by (6.2) is a martingale on \([0,T]\), so we can define a new probability measure \( \bar{Q} \) on the canonical space \((\Omega, \mathcal{F})\) by \( \frac{d\bar{Q}}{d\bar{P}}|_{\mathcal{F}_T} = M_T \), then under \( \bar{Q} \), the process \( Y \) (for the given \( \alpha \)) is a Brownian motion, and \( \bar{Q}\{V_T = g(X_T)\} = \bar{P}\{V_T = g(X_T)\} = 1 \). In other words, under \( \bar{Q} \), we can rewrite (6.5) and the SDE (4.1) for \( X \) as the following forward-backward SDE (FBSDE):

\[
\begin{aligned}
\begin{cases}
\quad dX_t = \mu(t,X_t)dt + \rho(t,X_t)dY_t, \quad X_0 = x; \\
\quad dP_t = [\beta_0(t,X_t,P_t) + \beta_1(t,X_t,P_t)Z_t]dt + Z_t dY_t, \quad P_T = g(X_T),
\end{cases}
\end{aligned}
\tag{6.6}
\]

where \( \beta_0(t,x,y) = -b_0(t,x) - b_1(t,x)y, \ \beta_1(t,x,y) = u_0(t,x) + u_1(t,x)y \).

**Remark 6.2.** (i) Although \( \bar{Q} \sim \bar{P} \ll \bar{Q}^0 \) and the process \( Y \) is a Brownian motion under both measures \( \bar{Q} \) and \( \bar{Q}^0 \), \( \bar{Q} \) and \( \bar{Q}^0 \) are not equivalent on \( \mathcal{F}_T \), since \( \bar{Q}^0\{V_T \neq g(X_T)\} > 0 \) in general. In fact, \( \mathbb{L}' \) is local martingale, but \( M \) is a true martingale.

(ii) Under Assumption 3.1, \( X \) is a diffusion driven by the \( \bar{Q} \)-Brownian motion \( Y \), hence it is \( \mathbb{P}^Y \)-adapted, which justifies (6.5), whence (6.6). \( \blacksquare \)

We should note that the FBSDE (6.6) is actually “decoupled”, in the sense that the forward SDE is independent of the backward components \((Y,Z)\). But the BSDE in (6.6) is somewhat non-standard in that the coefficients are neither Lipschitz nor of linear growth. Specifically, the fact that \( |\beta_1(t,x,y)z| \leq K(1 + |y||z|) \) makes it super-linear in \((y,z)\), and is beyond the usual “quadratic BSDE” framework. Nevertheless, the well-posedness of (6.6) can be argued via a more or less standard localization argument following the idea of [28]. Since this is not the main purpose of the paper, we shall only state the following result, but omit the proof (see [33] for details).

**Proposition 6.3.** Assume Assumption 3.1, and let \((\bar{P}, (B^1, B^2), (V, X), \alpha)\) be a Markovian nested solution to STPBVP (4.1), and assume that \( \alpha \) has an affine structure. Then there exists a probability measure \( \bar{Q} \) on the canonical space \((\Omega, \mathcal{F})\), such that

(i) \( \frac{d\bar{Q}}{d\bar{P}}|_{\mathcal{F}_T} = M_T \), where \( M \) satisfies the linear SDE (6.2);

(ii) denoting \( Y_t = B_t^2 + \int_0^t \alpha_s ds \) and \( P_t = \mathbb{E}_{\bar{P}}[V_t|\mathcal{F}_t^\gamma] \), \( t \in [0,T] \), then \( Y \) is a \( \bar{Q} \)-Brownian motion, and under \( \bar{Q} \), \((X, P)\) satisfies the FBSDE (6.6). \( \blacksquare \)
In the rest of this section we try to determine the most important element of the pricing mechanism: the function \( H : [0, T] \times \mathbb{R} \mapsto \mathbb{R} \), so that \( P_t = H(t, X_t) \), \( t \in [0, T] \). To begin with, we recall from the general theory of FBSDE (cf. e.g., [31], [30]) that, if \((X, P, Z)\) is the solution to the FBSDE (6.6), then under appropriate conditions on the coefficients, there is a decoupling field \( H : [0, T] \times \mathbb{R} \mapsto \mathbb{R} \), which satisfies the following semilinear PDE (at least in the viscosity sense):

\[
\begin{cases}
H_t(t, x) + \frac{1}{2} \rho^2(t, x) H_{xx}(t, x) + \mu(t, x) H_x(t, x) + h(t, x, H(t, x), \rho(t, x) H_x(t, x)) = 0; \\
H(T, x) = g(x),
\end{cases}
\tag{6.7}
\]

where \( h(t, x, y, z) = \beta_0(t, x, y) + \beta_1(t, x, y)z \), and the following relation holds: \( P_t = H(t, X_t) \), \( Z_t = \rho(t, X_t) H_x(t, X_t) \), \( t \in [0, T] \). The following extension of Example 5.2 justifies this fact.

**Example 6.4.** Recall that in Example 5.2, in which the coefficients \( b, \sigma, \mu \) and the function \( u \) have the specific form: \( b(t, v, x) = f_t v + g_t x + k_t \), \( \sigma \equiv 1 \), \( \mu(t, x) = (f_t + g_t) x + k_t \), \( u(t, v, x) = \beta_t v - \beta_t x \), and thus the PDE (6.7) now reads (suppressing variables):

\[
\begin{cases}
H_t + ((f_t + g_t)x + k_t + \rho(-\beta_t x + \beta_t H)) H_x + \frac{1}{2} \rho^2 H_{xx} = g_t x + k_t + f_t H; \\
H(T, x) = x,
\end{cases}
\tag{6.8}
\]

We can easily check that \( H(T, x) = x \) is the (unique) solution to (6.8), and hence \( P_t = H(t, X_t) = X_t \), for \( t \in [0, T] \), and \( X_T = H(T, X_T) = P_T = V_T \).

**Remark 6.5.** If we restrict the strategy to the form \( \alpha_t = \beta_t (V_t - P_t) = \beta_t (V_t - H(t, X_t)) \), that is, \( u_0 = -\beta_t H(t, x) \), \( u_1 = \beta_t \), and we assume further that the original asset \( V \) is under the risk neutral probability so that \( b(t, v, x) = 0 \), then (6.7) is reduced to

\[
\begin{cases}
H_t(t, x) + \mu(t, x) H_x(t, x) + \frac{1}{2} \rho^2(t, x) H_{xx}(t, x) = 0; \\
H(T, x) = g(x).
\end{cases}
\tag{6.9}
\]

We should note that the PDE (6.9) is well-posed with properly given coefficients, and \((\mu, \rho)\) can be chosen as part of the pricing rule. The determination of \((\mu, \rho)\), however, is the main task for finding the Kyle-Back equilibrium, which will be discussed in details in the next section.

### 7 Sufficient Conditions for Optimality

We are now ready to investigate the main issue of this paper: finding the equilibrium of the pricing problem. That is, we are to find the optimal strategy \( \alpha^* \) for the insider, which maximizes her expected terminal wealth \( W_T \), given the pricing rule \( P_t = E[V_t|\mathcal{F}_t] \), \( t \in [0, T] \).
In light of the analysis in the previous sections, we can recast the problem of finding the Kyle-Back equilibrium as follows. First recall the Markovized system (7.5):

\[
\begin{align*}
    dV_t &= b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB_t^1; \\
    dX_t &= \mu(t, X_t)dt + \rho(t, X_t)dY_t = [\mu(t, X_t) + \rho(t, X_t)\alpha_t]dt + \rho(t, X_t)dB_t^2, \\
    V_0 &= v; \\
    X_0 &= x.
\end{align*}
\]

(7.1)

where \( \alpha \in \mathcal{U} \) (see (2.3) for definition). Assume that the process \( \alpha \) takes the feedback form \( \alpha_t = u(t, V_t, X_t) \), we have argued in §2 that finding the optimal strategy amounts to solving a stochastic control problem with state equation (7.1) (or (2.5)) and the cost functional (2.8). Moreover, a necessary condition for \( \alpha \in \mathcal{U} \) being an equilibrium is that \( V_T = P_T = H(T, X_T) = g(X_T) \) (see (2.11)). Therefore, We shall consider only the (weak) solution \((\bar{P}, V, X, \alpha)\) to STPBVP (4.1), and by Proposition 4.4 we shall assume that \( \bar{P}|_{\mathcal{F}_t} \ll Q|_{\mathcal{F}_t}, t < T, \) and \( \bar{P} \circ (V_T)^{-1} = m^\ast \).

It is worth noting that the candidates for the equilibrium described above is based on the SDE (7.4), whence the coefficients \((\mu, \rho)\). We shall argue that the equilibrium can be determined by properly choosing \((\mu, \rho)\) through some “compatibility conditions”.

**The case** \( b(t, v, x) \equiv 0. \) For notational simplicity, in what follows we use \( P \) instead of \( \bar{P} \). As we pointed out in Remark 6.5, this could be the case when \( P \) is the risk neutral probability measure, and \( V \) is the discounted asset price, hence a \((P, \mathbb{F})\)-martingale. We note that in this case the market price \( P_t = \mathbb{E}[V_t|\mathcal{F}_t^Y], t \geq 0 \) is a \((P, \mathbb{F}^Y)\)-martingale. Indeed, since \( V = \{V_t\} \) is a \((P, \mathbb{F})\)-martingale, for \( s < t \), we have

\[
P_s = \mathbb{E}^P[V_s|\mathcal{F}_s^Y] = \mathbb{E}^P[\mathbb{E}^P[V_t|\mathcal{F}_s]|\mathcal{F}_s^Y] = \mathbb{E}^P[\mathbb{E}^P[V_t|\mathcal{F}_t]|\mathcal{F}_s^Y] = \mathbb{E}^P[\mathbb{E}^P[V_t|\mathcal{F}_t]|\mathcal{F}_s^Y] = \mathbb{E}^P[P_t|\mathcal{F}_s^Y].
\]

On the other hand, if we assume that \( P_t = H(t, X_t), t \in [0, T], \) where \( X \) satisfies the SDE (7.1), such that \( P_T = g(X_T) \), then a simple application of Itô’s formula shows that \( P = \{P_t\} \) begin an \( \mathbb{F}^Y \)-martingale means that the decoupling field \( H \) must satisfy the PDE:

\[
\begin{align*}
    H_t + \mu(t, x)\partial_x H + \frac{1}{2}\rho^2(t, x)H_{xx} &= 0; \quad t \in [0, T) \\
    H(T, x) &= g(x).
\end{align*}
\]

(7.2)

Comparing to (6.1) and recalling (6.4) we see that under the affine structure we should have

\[ h(t, x, H(t, x), \rho(t, x)H_x(t, x)) = \beta_1(t, x, H)\rho(t, x)H_x = (u_0(t, x) + u_1(t, x)H)\rho(t, x)H_x \equiv 0. \]

Therefore we have \( u_0(t, X_t) = -\beta_1 H(t, X_t) \), where \( \beta_1 = u_1(t, X_t) \). Consequently, we see that \( \alpha_t = u_0(t, X_t) + u_1(t, X_t)V_t = u_1(t, X_t)(V_t - H(t, X_t)) = \beta_1(V_t - P_t) \), which is exactly the form commonly seen in the literature (see, e.g., [1, 20, 29]), except that \( \beta_1 \) is no longer deterministic.

Our first main result of this section is the following theorem.
Theorem 7.1. Assume Assumption [3.7] and that \(b(t,v,x) = 0\). Assume further that \((\bar{P}, \bar{V}, \bar{X}, \bar{\alpha})\) is a weak solution to the SDE (4.1) such that \(\bar{\alpha}\) has an affine structure. Then,

(i) the market price \(P_t = \mathbb{E}_t^\bar{P}[\mathcal{F}_t^\bar{Y}] = H(t, \bar{X}_t), \ t \in [0, T)\) is an \(\mathbb{P}\)-martingale, where \(H\) solves the PDE (7.2);

(ii) the process \(\bar{\alpha}\) is of the form \(\bar{\alpha}_t = \beta(t, \bar{X}_t)(\bar{V}_t - H(t, \bar{X}_t)) = \beta(t, \bar{X}_t)(\bar{V}_t - P_t), \ t \in [0, T)\), where \((V, X)\) is the solution to the SDE (7.1) under some probability measure \(\bar{P}\), such that \(\bar{V}_T = g(\bar{X}_T)\), \(\mathbb{P}\)-a.s.;

(iii) \(\bar{\alpha}\) is an equilibrium strategy if the following “compatibility condition” holds for \((\mu, \rho)\):

\[
\partial_t \rho(t, x) - \partial_x \mu(t, x) \rho(t, x) + \partial_x \rho(t, x) \mu(t, x) + \frac{1}{2} \rho^2(t, x) \partial_{xx} \rho(t, x) = 0. \tag{7.3}
\]

Proof. The parts (i) and (ii) have been argued prior to the theorem. We shall prove only part (iii). To this end, we shall borrow the idea of [30], and look for a function \(J(t, x; a)\) such that for fixed \(a \in \mathbb{R}\), \(J(\cdot, \cdot; a) \in C^{1,2}([0, T] \times \mathbb{R})\), and satisfies the following properties

\[
\begin{align*}
J_t(s, x; a) + J_x(s, x; a) \mu(s, x) + \frac{1}{2} J_{xx}(s, x; a) \rho^2(s, x) &= 0; \\
J_x(s, x; a) \rho(s, x) &= H(t, x) - a; \\
J(T, x; a) &\geq 0, \text{and } J(T, x; a) = 0 \iff a = g(x).
\end{align*}
\tag{7.4}
\]

Assume now that a function \(J\) satisfying (7.4) exists. Then for any \(\alpha \in \mathcal{U}_{ad}\), we let \((\bar{P}, V^\alpha, X^\alpha)\) be a weak solution to the SDE (4.1) with \(
J(\cdot, \cdot; a)\) we have

\[
\begin{align*}
J(t, X^\alpha_t; a) &= J(0, x_0; a) + \int_0^t \left[ J_t(\cdot, \cdot; a) + J_x(\cdot, \cdot; a) \mu + \frac{1}{2} J_{xx}(\cdot, \cdot; a) \rho^2 \right](s, X^\alpha_s) ds \\
&\quad + \int_0^t J_x(s, X^\alpha_s; a) \rho(s, X_s) dY_s = J(0, x_0; a) + \int_0^t (H(s, X^\alpha_s) - a) dY_s \\
&\quad = J(0, x_0; a) + \int_0^t (H(s, X^\alpha_s) - a) \alpha_s ds + \int_0^t H(s, X^\alpha_s) dB^2_s - aB^2_t.
\end{align*}
\tag{7.5}
\]

Denoting \((V, X) = (V^\alpha, X^\alpha)\) and applying the total probability formula and (7.5) we have

\[
\begin{align*}
\mathbb{E}_t^\bar{P}[J(T, X_T; V_T) - J(0, x_0; V_T)] &= \int_\mathbb{R} \mathbb{E}_t^\bar{P}[J(T, X_T; a) - J(0, x_0; a)|V_T = a] \mathbb{P} \circ (V_T)^{-1}(da) \\
&= \int_\mathbb{R} \mathbb{E}_t^\bar{P}\left[ \int_0^T (H(s, X_s) - a) \alpha_t dt + \int_0^T H(t, X_t) dB^2_t - aB^2_t|a = V_T \right] \mathbb{P} \circ (V_T)^{-1}(da) \tag{7.6} \\
&= \mathbb{E}_t^\bar{P}\left[ \int_0^T (H(s, X_s) - V_T) \alpha_t dt \right] + \mathbb{E}_t^\bar{P}\left[ \int_0^T H(t, X_t) dB^2_t \right] - \mathbb{E}_t^\bar{P}[V_T B^2_T] \\
&= \mathbb{E}_t^\bar{P}\left[ \int_0^T (H(s, X_s) - V_T) \alpha_t dt \right] - \mathbb{E}[V_T B^2_T].
\end{align*}
\]
But, since \( \langle B^1, B^2 \rangle \equiv 0 \), we have \( d(V_tB_t^1) = V_tdB_t^1 + B_t^1 \sigma(t, V_t, X_t)dB_t^1 \), \( t \geq 0 \). That is, \( \{V_tB_t^1\} \) is a \( \mathbb{P} \)-martingale, hence \( \mathbb{E}^F[V_TB_T^2] = 0 \). Recalling (2.6) we deduce from (7.6) that
\[
\mathbb{E}^F[W_T^a] = \mathbb{E}^P \left[ \int_0^T (V_t^a - H(s, X_t^a)) \alpha_t \, dt \right] = \mathbb{E}^P \left[ J(0, x_0; V_T^a) - J(T, X_T^a; V_T^a) \right] = \mathbb{E}^P \left[ J(0, x_0; V_T^a) \right].
\]

Here the last inequality is due to property (7.3) of the function \( J \), and furthermore, the equality holds if and only if the terminal condition \( V_T^a = g(X_T^a) \) holds. Consequently, if we let \( (\mathbb{P}, \tilde{V}, \tilde{X}, \tilde{\alpha}) \) be a weak solution to STPBVP (4.1), then Proposition 4.4, together with (7.7), shows that
\[
\mathbb{E}^F[W_T^a] = \sup_{\alpha \in \mathbb{H}_{ad, \mathbb{F}(V_T^a)^{-1} = m^*}} \mathbb{E}^P[W_T^a] = \int_{\mathbb{R}} J(0, x_0; a)m^*(da).
\]

In other words, the solution to the STPBVP leads to the optimal strategy for the insider, among all the strategies satisfying \( \mathbb{P} \circ (V_T^a)^{-1} = m^* \).

Our last task is to construct a function \( J \) that satisfies all the requirements in (7.4). In light of [36], we consider the following function:
\[
J(t, x; a) = \int_{g^{-1}(a)}^x \frac{H(t, y) - a}{\rho(t, y)} \, dy + \int_t^T f(s; a) \, ds,
\]
where \( H(\cdot, \cdot) \) satisfies (7.2), and \( f(t; a) \) is a function to be determined and independent of \( x \). To check that such a function is possible for the proper choices of \( \mu, \rho, \) and \( f \), we simply plugging the function \( J \) into the PDE in (7.4) to get
\[
f(t; a) = \left( \frac{H}{\rho} - \frac{\rho_x}{2} \right)(H - a)(t, x) + \frac{H_x \rho(t, x)}{2} + \int_{g^{-1}(a)}^x \frac{H_t}{\rho} - \frac{(H - a) \rho_t}{\rho^2} \, dy.
\]

In order that \( f(\cdot; a) \) is independent of \( x \), we take derivative of the right hand side of (7.9) with respect to \( x \), and multiply it by \( \rho_x^2(t, x) \) to obtain (suppressing variables and rearranging terms)
\[
f_x \rho^2 = \rho[H_t + \mu H_x + \frac{1}{2} \rho^2 H_{xx}] + [(\mu_x \rho - \mu \rho_x) - \frac{1}{2} \rho_{xx} \rho^2 - \rho_t](H - a)
\]
\[
= [(\mu_x \rho - \mu \rho_x) - \frac{1}{2} \rho_{xx} \rho^2 - \rho_t](H - a),
\]
thanks to (7.2). Since \( \rho \) is positive, we see that \( f_x \equiv 0 \) provided (7.3) holds. We note that if the function \( f \) in (7.8) is independent of \( x \), then the second equation in (7.4) is obvious by definition. It thus remains to verify the last requirement of (7.4). To see this we note that
\[
J(T, x; a) = \int_{g^{-1}(a)}^x \frac{H(T, y) - a}{\rho(T, y)} \, dy = \int_{g^{-1}(a)}^x \frac{g(y) - a}{\rho(T, y)} \, dy.
\]

Since \( g \) is increasing, for any \( y \geq g^{-1}(a) \) we have \( g(y) \geq g(g^{-1}(a)) = a \). Thus \( J(T, x; a) \geq 0 \), for \( x \geq g^{-1}(a) \) as \( \rho(T, y) > 0 \), and \( J(T, x; a) = 0 \) if and only if \( x = g^{-1}(a) \). The proof is now complete.
Remark 7.2. The compatibility condition (7.3) between the coefficients \( \mu, \rho \), and the PDE (7.2) for the pricing rule \( H \) are not new. In the so-called “long-lived” information case, for example, the market price \( P_t = \mathbb{E}[V_T | \mathcal{F}_t^Y] \), \( t \geq 0 \), is naturally a martingale, and \( b \equiv 0 \) is by assumption, thus Theorem 7.1 always applies. In this case, \( \mu \) chooses \( \mu = 0 \) and \( \rho = 1 \), which obviously satisfies the compatibility condition (7.3), and (7.2) becomes \( H_t + \frac{1}{2} H_{xx} = 0 \), and \( f(t) = H_x(t, g^{-1}(a)) \).

As another example, in [12] it is derive from a control theoretic argument via HJB equation that \( \mu = 0 \), and \( \rho \) and \( H \) satisfy

\[
\rho_t + \frac{\rho^2}{2} \rho_{xx} = 0, \quad H_t + \frac{\rho^2}{2} H_{xx} = 0,
\]

and \( f(t;a) = H_x(t, g^{-1}(a))\rho(t, g^{-1}(a)) \), justifying our compatibility conditions (7.2) and (7.3).

The Case \( b(t, v, x) \neq 0 \). In this case, the market price \( P_t = \mathbb{E}[V_T | \mathcal{F}_t^Y] \), \( t \geq 0 \), is no longer an \( \mathbb{F}^Y \)-martingale, but rather an “optional projection”, and the discussion is more involved. Assume again that both \( b \) and \( \alpha \) have the affine structure: \( b = b_0(t, x)+b_1(t, x)v \) and \( u = u_0(t, x)+u(t, x)v \). By Proposition 6.3 the decoupling field \( H(t, x) \) would satisfy a more general PDE:

\[
H_t + \mu H_x + \frac{1}{2} \rho^2 H_{xx} = -(b_0 + b_1 H) + (u_0 + u_1 H) \rho H_x, \quad H(T, x) = g(x). \quad (7.10)
\]

So if we still try to construct the function \( J(t, x; a) \) as in (7.3), then it may not be possible to find a corresponding function \( f \) that is independent of \( x \). We now try to modify (7.9) accordingly. First recall (2.7) we can write

\[
\mathbb{E}^F[W_T^a] = \mathbb{E}^F \left[ \int_0^T [F(t, V_t^\alpha, X_t^\alpha) - H(t, X_t^\alpha)]u(t, V_t^\alpha, X_t^\alpha)dt \right],
\]

where \( F(t, V_t, X_t) := \mathbb{E}^F[V_T | \mathcal{F}_t^Y, X] \), thanks to the Markovian nature of the process \( \alpha \), whence the solution \( (V^\alpha, X^\alpha) \). Now by Feynman-Kac formula, we see that \( F \) satisfies the PDE:

\[
F_t + \frac{1}{2} F_{vv} \sigma^2 + \frac{1}{2} F_{xx}\rho^2 + F_x b + F_x (\mu + u \rho) = 0; \quad F(T, v, x) = v. \quad (7.11)
\]

In light of the case \( b = 0 \), we now look for the function \( J(t, v, x) \) with the following properties:

\[
\begin{cases}
J_x \rho(t, x) = H(t, x) - F(t, v, x); \\
J_t + b(t, v, x) J_v + \mu(t, x) J_x + \frac{1}{2} \sigma^2 (t, v, x) J_{vv} + \frac{1}{2} \rho^2 (t, x) J_{xx} = 0; \\
J(T, v, x) \geq 0, \quad \text{and} \quad J(T, v, x) = 0 \text{ iff } v = g(x).
\end{cases}
\quad (7.12)
\]

If such function \( J \) exists, then a simple application of Itô’s formula shows that

\[
\mathbb{E}^F[W_T^a] = \mathbb{E}^F[-J(T, V_T, X_T) + J(0, v_0, x_0)] \leq J(0, v_0, x_0),
\]
and the equality holds when \( V_T = g(X_T) \) \( \mathbb{P} \)-a.s., which would imply the optimality of the solution to STPBVP. To find such a function, we modify (7.8) as follows. Define

\[
J(t, v, x) = \int_{g^{-1}(v)}^{x} \frac{H(t, y) - F(t, v, y)}{\rho(t, y)} dy + G(t, v) := \bar{J}(t, v, x) + G(t, v),
\]

(7.13)

where \( G(t, v) \) is a function to be determined. We have the following result.

**Theorem 7.3.** Assume Assumption 3.1 and assume further that \( b \) and \( \sigma \) take the form:

\[
b(t, v, x) = b_0(t) + b_1(t)v, \quad \sigma(t, v, x) = \sigma(t, v).
\]

Then, a weak solution \((\bar{\mathbb{P}}, \bar{V}, \bar{X}, \bar{\alpha})\) to STPBVP (4.7) with \( \alpha \) having the affine structure is an equilibrium strategy if the following compatibility condition holds:

\[
\left\{ \begin{array}{l}
\rho(b_0(t) + b_1(t)H) + \rho^2 [bu_0 + u_1v]x F - \partial_x H(u_0 + u_1H) = 0; \\
-\partial_x \rho \mu + \partial_x \rho \mu + \partial_t \rho + \frac{1}{2} \rho^2 \partial_x \rho = 0.
\end{array} \right.
\]

(7.14)

where \( H \) and \( F \) satisfy the PDEs (7.10) and (7.11), respectively.

**Proof.** We shall argue that under the compatibility condition (7.14), the function \( J \) defined by (7.13) satisfies the desired properties (7.12). To see this, first note that the first identity in (7.12) is trivial by definition of the function \( J \). Next, we observe that \( J(T, v, x) = \bar{J}(T, v, x) + G(T, v) = \int_{g^{-1}(v)}^{x} g(y) - \frac{v}{\rho(t, y)} dy + G(T, v) \), where \( \bar{J}(T, v, x) \geq 0 \) and \( \bar{J}(T, v, x) = 0 \) if and only if \( x = g^{-1}(v) \), as we argued in Theorem 7.1. Therefore the function \( J \) defined by (7.13) satisfies the terminal condition in (7.12) provided \( G(T, v) \equiv 0 \).

Let us now look at the PDE in (7.12). Plugging (7.13) into the PDE in (7.12), we have

\[
0 = \partial_t G + b(t, v)\partial_v G + \frac{1}{2} \rho^2 (\rho t, v) \partial_v G + \partial_t \bar{J} + b(t, v)\partial_v \bar{J} + \frac{1}{2} \rho^2 (\rho t, v) \partial_v \bar{J} + \mu \frac{H - F}{\rho} + \frac{1}{2} [ (\rho_x H - \rho_x F) \rho - (H - F) \partial_x \rho ].
\]

(7.15)

We shall now find conditions that \( G(t, v) \) is a desired function by showing that the PDE (7.15) with the terminal condition \( G(T, v) \equiv 0 \) is independent of \( x \). To see this we simply take the partial derivative with respect to \( x \) on both sides of (7.15), multiply by \( \rho^2 \), and set it to zero to get (suppressing variables):

\[
0 = (H - F) \left[ - \partial_t \rho + \partial_x \rho \mu - \mu \partial_x \rho - \frac{1}{2} \rho^2 \rho_x \rho \right]
\]

(7.16)

\[
-\rho \left\{ \partial_t F + b\partial_v F + \mu \partial_x F + \frac{1}{2} \rho^2 \partial_v F + \frac{1}{2} \rho^2 \partial_x F \right\} + \rho (\partial_t H + \mu \partial_x H + \frac{1}{2} \rho^2 \partial_x H).
\]

Now, since \( H \) and \( F \) satisfies PDEs (7.10) and (7.11), respectively, we can easily check that a sufficient condition for (7.16) is the compatibility condition (7.14), proving the theorem.
Remark 7.4. By (7.13), it seems that the function $J$ depends on the choice of the strategy $\alpha$ since both PDEs (7.10) and (7.11) (for $H$ and $F$) do. However, the PDE in (7.12) for $J$, as well as its terminal condition are independent of $u$. Consequently, the function $J$ is independent of the choice of $u$. 

References

[1] Aase, K. K., Bjuland T., and Oksendal, B., Strategic insider trading equilibrium: a filter theory approach, Afr. Mat., 23 (2012), pp.145-162.

[2] Aronson, D.G., Bounds for the fundamental solution of a parabolic equation, Bull. Amer. Math. Soc., 73 (1967), pp.890-896.

[3] Aronson, D.G., Non-negative solutions of linear parabolic equations, Ann. Scuola Norm. Sup. Pisa., 22 (1968), pp.607-694.

[4] Back, K., Insider trading in continuous time, The Review of Financial Studies, 5 (1992), pp.387-409.

[5] Back, K., and Pedersen, H., Long-lived information and intraday patterns, Journal of financial markets, 1 (1998), pp.385-402.

[6] Baudoin, F., Conditioned stochastic differential equations: theory, examples and application to finance, Stochastic Processes and their Applications, 100(1-2) (2002), pp.109-145.

[7] Barrios, B., Peral, I., Soria, F., and Valdinoci, E., A Widder’s type theorem for the heat equation with non-local diffusion, (2013), arXiv:1302.1786v1.

[8] Bensoussan, A., Stochastic control of partially observable systems, Cambridge University Press, (2004).

[9] Biagini, F., Hu, Y., Meyer-Brandis, T., and Oksendal, B., Insider trading equilibrium in a market with memory, Math. Finan. Econ., 6 (2012), pp.229-247.

[10] Buckdahn, R., Li, J., and Ma, J., A Mean-field Stochastic Control Problem with Partial Observations, The Annals of Applied Probab., 27(5) (2017), pp.3201-3245.

[11] Caldentey, R., and Stacchetti, E., Insider Trading with a Random Deadline, Econometrica, 1 (2010), pp.245-283.
[12] Campi, L., Çetin, U., and Danilova, A., *Dynamic Markov bridges motivated by models of insider trading*, Stochastic Processes and their Applications, 121 (2011), pp.534-567.

[13] Campi, L., Çetin, U., and Danilova, A., *Explicit construction of a dynamic Bessel bridge of dimension 3*, Electron. J. Probab., 18 (2013), pp.1-25.

[14] Campi, L., Çetin, U., and Danilova, A., *Equilibrium model with default and dynamic insider information*, Finance and Stochastics, 17 (2013), pp.565-585.

[15] Çetin, U., and Danilova, A., *On pricing rules and optimal strategies in general Kyle-Back models*, (2021), arXiv:1812.07529v2.

[16] Cocquemas, F., Ekren, I., and Lioui, A., *A general solution method for insider problems*, (2020), arXiv:2006.09518v1.

[17] Danilova, A., *Stock market insider trading in continuous time with imperfect dynamic information*, Stochastics, 82 (2010), pp.111-131.

[18] Fernique, X., *Intégrabilité des vecteurs gaussiens*, CR Acad. Sci. Paris Serie A 270, (1970), pp.1698-1699.

[19] Fitzsimmons, P., Pitman, J., and Yor, M., *Markovian bridges: construction, Palm interpretation, and splicing*, In Seminar on Stochastic Processes, (1992), pp.101-134.

[20] Fleming, W.H., and Rishel, R.W., *Deterministic and stochastic optimal control*, Springer Science & Business Media, 1 (2012).

[21] Föllmer, H. and Imkeller, P., *Anticipation cancelled by a Girsanov transformation: a paradox on Wiener space*, Annales de l'I. H. P., section B, 29(4) (1993), pp.569-586.

[22] Föllmer, H., Wu, C.T., and Yor, M., *Canonical decomposition of linear transformations of two independent Brownian motions motivated by models of insider trading*, Stochastic Processes and Their Applications, 84 (1999), pp.137-164.

[23] Fujisaki, M., Kallianpur, G., and Kunita, H., *Stochastic differential equations for the nonlinear filtering problem*, Osaka Journal of Mathematics, 9(1) (1972), pp.19-40.

[24] Guenther, R.B., *An Extension of Widder’s Theorem*, Master Thesis, Oregon State University, (1962).

[25] Holden, C.W., and Subrahmanyam, A., *Long-lived private information and imperfect Competition*, Journal of Finance, 1 (1992), pp.247-270.
[26] Kyle, A. S., *Continuous auctions and insider trading*, Econometrica: Journal of the Econometric Society, (1985), pp.1315-1335.

[27] Lepeltier, J.P., and San Martin, J., *Backward stochastic differential equations with continuous coefficient*, Statistics & Probability Letters, 32(4) (1997), pp.425-430.

[28] Liu, Y., and Ma, J., *BSDE with jumps and multiplicative coefficients*, Preprint (2005).

[29] Ma, J., Sun, R., and Zhou, Y., *Kyle–Back Equilibrium Models and Linear Conditional Mean-Field SDEs*, SIAM Journal on Control and Optimization, 56(2) (2018), pp.1154-1180.

[30] Ma, J., Wu, Z., Zhang, D., and Zhang, J., *On Wellposedness of Forward-Backward SDEs — A Unified Approach*, Ann. Appl. Probab, 25(4) (2015), pp.2168-2214.

[31] Ma, J., and Yong, J., *Forward-backward stochastic differential equations and their applications*, Springer Science & Business Media No.1702, (1999).

[32] Protter, P., *Integration and Stochastic Differential Equations—A New Approach*, (1990).

[33] Tan, Y., *Generalized Kyle-Back Equilibrium Models with Dynamic Information and Related Topics*, PhD dissertation, University of Southern California, (2022).

[34] Widder, D., *Positive temperatures on an infinite rod*, Transactions of the American Mathematical Society, 55 (1944), pp.85-95.

[35] Widder, D., *Positive temperatures on a semi-infinite rod*, Transactions of the American Mathematical Society, 75 (1953), pp.510-525.

[36] Wu, C.T., *Construction of Brownian motions in enlarged filtrations and their role in mathematical models of insider trading*, (1999).