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* (see, e.g., [12]), as a part of the "pricing rule". As the main technical tool in solving the Kyle-Back equilibrium in such a setting, 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.

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

AMS subject classifications. 60H10, 93E11, 91G15, 91G80

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 [24] and Back ([4, 5]) (see also [1, 9, 11, 16, 23] 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 value 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 ξ_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 $z_t = \int_0^t \sigma^z_t dB^z_t$, 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., [24], 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,

*Submitted to the editors DATE.

Funding: This work was funded by US NSF grants #1908665 and #2205972.

† Department of Mathematics, University of Southern California, Los Angeles, CA 90089 (jinma@usc.edu).

‡Department of Mathematics, University of Southern California, Los Angeles, CA 90089 (yingtan@usc.edu).

for $t \in [0,T]$,

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

where $\mathcal{F}_t^Y := \sigma\{Y_s, s \le t\}$. An *equilibrium* of the Kyle-Back problem consists of an insider's strategy ξ^* that maximizes her expected wealth at terminal time T, together with the market price P in either form of (1.1) (known 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, 15] 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:

(i) The evolution of the dynamics of the underlying asset can depend on the market price $P = \{P_t\}$ (hence depending on the market information $\mathbb{F}^Y = \{\mathcal{F}_t^Y\}$).

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

(iii) the market maker's pricing rule is in general an "*optional projection*" of the underlying asset, rather than a martingale (note the two different forms in (1.1)).

We note that the feature (i) above, although reasonable (see, e.g., [27]), 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:

$$dV_t = b(t, V_{\cdot\wedge t}, P_{\cdot\wedge t})dt + \sigma(t, V_{\cdot\wedge t}, P_{\cdot\wedge t})dB_t^1, \qquad V_0 = v, \tag{1.2}$$

where b, σ are given measurable functions. We shall also assume, as commonly seen in the literature, that the insider's strategy is of the form $\xi_t = \int_0^t \alpha_s ds$, $t \ge 0$, where the "rate" α can depend on both V and P in an *nonanticipative* way, so that the dynamics the market maker observes is:

$$dY_t = d\xi_t + dz_t = \alpha(t, V_{\cdot\wedge t}, P_{\cdot\wedge t})dt + dB_t^2, \qquad t \ge 0. \tag{1.3}$$

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, 27]). 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

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

Such a factor process X resembles the so-called *weighted total process* (see, e.g., [12]), which was assumed to be a diffusion process driven by the observation process Y (see

§2 for detail). 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 the strategy α^* being optimal is that the corresponding solution (V, X) satisfies:

$$V_T = P_T = H(T, X_T) := g(X_T). \tag{1.4}$$

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 find the Kyle-Back equilibrium (see, e.g., [21, 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 [7]) 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., [6, 30, 33, 32]) 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. [28]) 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., [26]). We hope to be able to address such issues in our future publications.

The rest of the paper is organized a 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 $\mathbb{X}$ be a generic Euclidean space and regardless of its dimension, $(\cdot, \cdot)$ and $|\cdot|$ be its inner product and norm, respectively. We denote the space of $\mathbb{X}$-valued continuous functions defined on $[0,T]$ with the usual sup-norm by $\mathbb{C}([0,T];\mathbb{X})$. In particular, we denote $\mathbb{C}^2_T := \mathbb{C}([0,T];\mathbb{R}^2)$, and let $\mathscr{B}(\mathbb{C}^2_T)$ 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 $(\Omega, \mathcal{F}) := (\mathbb{C}^2_T, \mathscr{B}(\mathbb{C}^2_T))$; $\mathbb{P} \in \mathscr{P}(\Omega)$; and $B = (B^1, B^2)$ is a $\mathbb{P}$-Brownian motion. Moreover, we shall assume that $\mathbb{F}^i = \{\mathcal{F}^{B^i}_t\}_{t\ge 0}$, $i = 1, 2$, 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., [29]). Finally, we denote $\mathbb{Q}^0 \in \mathscr{P}(\Omega)$ to be the Wiener measure on $(\Omega, \mathcal{F})$; $B^0_t(\omega) = \omega(t)$, $\omega \in \Omega$, the canonical process; and $\mathbb{F}^0 := \{\mathcal{F}^0_t\}_{t\in[0,T]}$, where $\mathcal{F}^0_t := \mathscr{B}_t(\mathbb{C}^2_T) := \sigma\{\omega(\cdot \wedge t) : \omega \in \mathbb{C}^2_T\}$, $t \in [0,T]$. In what follows we shall make use of the following notations:

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

• For $1 \le p < \infty$, $\mathbb{G} \subseteq \mathbb{F}$, $L^p_{\mathbb{G}}([0,T];\mathbb{X})$ denotes the space of all $\mathbb{X}$-valued, $\mathbb{G}$-progressively measurable processes ξ satisfying $\mathbb{E}(\int_0^T |\xi_t|^p dt) < \infty$. The meaning of $L^\infty_{\mathbb{G}}([0,T];\mathbb{X})$ is defined similarly. For simplicity, we will often drop $\mathbb{X}(= \mathbb{R})$ from the notation, and denote all "L^p-norms" by $\|\cdot\|_p$, regardless it is for $L^p(\mathcal{G})$, or for $L^p_{\mathbb{F}}([0,T])$, 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 $m^* \in \mathscr{P}(\mathbb{R})$. Specifically, we assume that the asset process V is governed by the following SDE:

$$dV_t = b(t, V_t, P_t)dt + \sigma(t, V_t, P_t)dB^1_t, \qquad V_0 = v, \tag{2.1}$$

where b, σ are measurable functions, and $P = \{P_t\}$ is the market price. We should note that allowing (b, σ) 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 $\mathbb{F}^Y$ (see (2.2) 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 $\xi_t = \xi^\alpha_t := \int_0^t \alpha_s ds$, where the process $\alpha = \{\alpha_t\}$, often referred to as the *trading strategy*, is assumed to have the form $\alpha_t = u(t, V_{\cdot\wedge t}, P_{\cdot\wedge t})$, $t \in [0,T]$, for some function u to be determined (see, e.g., [5, 27]).

(ii) *The noise traders.* For simplicity, in this paper we shall assume that the (collective) order submitted by the noise traders is simply the $z_t = B^2$, for some Brownian motion $B^2 \perp\!\!\!\perp B^1$. 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., [24]), 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 $\mathcal{F}^Y_t$-measurable random variables. That is, $P_t = \mathbb{E}[V_t|\mathcal{F}^Y_t]$,

$t \in [0,T]$, where Y is the total trading volume:

$$Y_t = \xi_t^\alpha + B_t^2 = \int_0^t \alpha_s ds + B_t^2, \qquad t \in [0,T]. \tag{2.2}$$

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 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 describe the equilbrium, let us specify the set of *admissible strategies*:

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

where $L^\alpha_t := \exp\big\{\int_0^t \alpha_s dB^2_s - \frac{1}{2}\int_0^t |\alpha_s|^2 ds\big\}$, $t \in [0,T)$. A (generalized) Kyle-Back equilibrium consists of a "pricing rule", under which $P_t = \mathbb{E}[V_t|\mathcal{F}^Y_t]$, $t \in [0,T]$; and an optimal strategy $\alpha^* \in \mathscr{U}_{ad}$, 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 $\mathbb{E}^{\mathbb{P}}[W_T^{\alpha^*}] = \sup_{\alpha\in\mathscr{U}_{ad}} \mathbb{E}^{\mathbb{P}}[W^\alpha]$.

Remark 2.1. (i) In (2.3) the process L^α is defined only on $[0,T)$. In fact, it has been noted that the optimal strategy α^*_t often explodes when $t \nearrow T$, because the insider will try to use all the information advantage before it ends. (ii) From (2.2) we see that Y depends on α, thus so do the market price P and the asset price V. Therefore, a more precise definition of the admissible control set should be all $\alpha \in \mathscr{U}_{ad}$ such that $V_T = V_T^\alpha \sim m^* \in \mathscr{P}(\mathbb{R})$, the law that is known to the insider. 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 §4). ■

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 $\mathbb{F}^Y = \{\mathcal{F}_t\}$, but not necessarily an $\mathbb{F}^Y$-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 $P_t = \Phi(t, Y_{\cdot\wedge t})$, $t \geq 0$, for some measurable function Φ defined on $[0,T] \times \mathbb{C}([0,T])$. 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, 27]). In this paper we shall follow the idea of [12] to first *Markovzie* the system (2.1)-(2.2) by introducing a *factor* process X, which satisfies an auxiliary SDE of the form:

$$dX_t = \mu(t, X_t)dt + \rho(t, X_t)dY_t, \qquad X_0 = x, \tag{2.4}$$

where the coefficients (μ, ρ) 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 $\mathbb{Q} \in \mathscr{P}(\Omega)$ 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 Ψ, and consequently, we have

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

We note that the factor process X resembles the *weighted total order process* proposed in [12]), and the function H (together with the coefficients (μ, ρ)) 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:

$$\begin{cases} dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB^1_t, & V_0 = v; \\ dX_t = \mu(t, X_t)dt + \rho(t, X_t)dY_t, & X_0 = x. \end{cases} \tag{2.5}$$

Considering (2.5) as a controlled system with the control $\alpha \in \mathscr{U}_{ad}$. 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:

$$\mathbb{E}[W^\alpha_T] = \mathbb{E}\Big[(V_T - P_T)\xi^\alpha_T + \int_0^T \xi^\alpha_t dP_t\Big] = \mathbb{E}\Big[\int_0^T [V_T - P_t]\alpha_t dt\Big]. \tag{2.6}$$

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

$$\mathbb{E}[W^\alpha_T] = \mathbb{E}\Big[\textstyle\int_0^T [\mathbb{E}[V_T|\mathcal{F}^{V,X}_t] - P_t]\alpha_t dt\Big] = \mathbb{E}\Big[\textstyle\int_0^T [F(t, V_t, X_t) - H(t, X_t)]\alpha_t dt\Big],$$

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 functional*:

$$J(t, v, x; u) := \mathbb{E}_{t,v,x}\Big[\int_t^T (F(s, V_s, X_s) - H(s, X_s))u(s, V_s, X_s)ds\Big], \tag{2.7}$$

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

$$\begin{aligned} 0 &= \mathbf{v}_t(t, v, x) + b(t, v, x)\mathbf{v}_v + \mu(t, x)\mathbf{v}_x + \frac{1}{2}\sigma^2(t, v, x)\mathbf{v}_{vv} + \frac{1}{2}\rho^2(t, x)\mathbf{v}_{xx} \\ &\quad + \sup_{u\in\mathbb{R}} \big\{[\rho(t, x)\mathbf{v}_x + F(t, v, x) - H(t, x)]u\big\}. \end{aligned} \tag{2.8}$$

Clearly, a necessary condition for the "sup"-term in (2.8) to be finite is:

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

In particular, noting that $F(T, v, x) = v$, and $\mathbf{v}(T, v, x) \equiv 0$ by (2.7), we deduce that

$$0 \equiv \rho(T, x)\mathbf{v}_x(T, v, x) = H(T, x) - F(T, v, x) =: g(x) - v, \;\; (v, x) \in \mathbb{R}^2, \tag{2.9}$$

where $g(x) = H(T, x)$. In other words, 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, (2.9) indicates an important fact: a necessary condition for $\alpha \in \mathscr{U}_{ad}$ being an equilibrium is that the following condition holds at the terminal time T:

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

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

$$\begin{cases} dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB^1_t, \\ dX_t = [\mu(t, X_t) + \alpha_t\rho(t, X_t)]dt + \rho(t, X_t)dB^2_t, \\ V_0 = v, \quad X_0 = x, \quad V_T = g(X_T). \end{cases} \tag{2.10}$$

In what follows we shall refer to (2.10) as a *Stochastic Two-Point Boundary Value Problem*, whose solvability will be studied in details in the next section. In particular, we are interested in the case when α 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 (2.10) is closely related to the so-called *dynamic Markov bridge* studied in, e.g., [12, 13, 21]. In fact, if $b = \mu = 0$, $\sigma = \rho = 1$, and $g(x) = x$, the problem (2.10) was first studied, as the Brownian bridge, in the context of insider trading in [21]. The more general cases were considered recently in [12, 13, 14], 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 (2.10) 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 (2.10) is based on the notion of the so-called *conditioned SDE* (cf. [7]), which we now briefly describe. Recall the canonical probabilistic set-up $(\Omega, \mathcal{F}, \mathbb{Q}^0; \mathbb{F}, B^0)$ defined in the beginning of §2. In particular, we denote the canonical process by $B^0 = (B^1, Y)$ so that it is a $(\mathbb{Q}^0, \mathbb{F})$-Brownian motion. Consider the SDE on canonical space $(\Omega, \mathcal{F}, \mathbb{Q}^0, B^0)$, for $t \in [0, T]$:

$$(3.1)\qquad \begin{cases} dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB^1_t, & V_0 = v; \\ dX_t = \mu(t, X_t)dt + \rho(t, X_t)dY_t, & X_0 = x. \end{cases}$$

Throughout the paper we shall make use of the following *Standing Assumptions*:

Assumption 3.1. (i) The functions $b, \sigma : [0, T] \times \mathbb{R}^2 \to \mathbb{R}$ and $\mu, \rho : [0, T] \times \mathbb{R} \to \mathbb{R}$ 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,

$$\begin{cases} |b(t, 0, 0)| + |\sigma(t, 0, 0)| + |\mu(t, 0)| + |\rho(t, 0)| \le L, \\ |\phi(t, v, x) - \phi(t, v', x')| \le L(|v - v'| + |x - x'|), & \phi = b, \sigma, \\ |\psi(t, x) - \psi(t, x')| \le L|x - x'|, & \psi = \mu, \rho; \end{cases}$$

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

(iv) The functions $g : \mathbb{R} \to \mathbb{R}$ is strictly increasing, and both g and g^{-1} are uniformly Lipschitz continuous. ∎

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, ξ is a Markov process, and we denote its transition density by $p(s, x; t, y)$, $0 \le s < t \le T$, $x, y \in \mathbb{R}^2$. For $\nu \in \mathscr{P}(\mathbb{R}^2)$, we shall refer to the triplet (T, ξ_T, ν) as a "conditioning" below. Define

$$L^\nu_t := \int_{\mathbb{R}^2} \eta^y_t \nu(dy), \quad \text{where} \quad \eta^y_t := \frac{p(t, \xi_t; T, y)}{p(0, \xi_0; T, y)}, \qquad t < T, \quad \mathbb{Q}^0\text{-a.s.}$$

Definition 3.2. *The conditioning triplet (T, ξ_T, ν) is called "proper" if*

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

(ii) there exist constants $C, \lambda > 0$, such that

$$(3.2)\qquad 0 < \sup_{t\in[0,T)} (T - t)\eta^y_t \le CTe^{\frac{\lambda|\xi_0 - y|^2}{T}}, y \in \mathbb{R}^2; \quad \int_{\mathbb{R}^2} e^{\frac{\lambda|\xi_0 - y|^2}{T}} \nu(dy) < \infty.$$

We note that the condition (i) above is relatively easier to verify. In particular, it would be trivial when the diffusion ξ 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, λ, $\Lambda>0$, it holds that

$$0<\frac{c_1}{t-s}e^{-\frac{\lambda|y-x|^2}{t-s}}\le p(s,y;t,x)\le\frac{c_2}{t-s}e^{-\frac{\Lambda|y-x|^2}{4(t-s)}},\quad 0\le s<t<T,\quad x,y\in\mathbb{R}^2,$$

Consequently we see that,

$$0<\eta_t\le\frac{c_2T}{c_1(T-t)}e^{-\frac{\Lambda|\xi_t-y|^2}{4(T-t)}+\frac{\lambda|\xi_0-y|^2}{T}}\le\frac{c_2T}{c_1(T-t)}e^{\frac{\lambda|\xi_0-y|^2}{T}},\ t\in[0,T),$$

which leads to the first inequality in (3.2). Thus the requirement for the conditioning being "proper" means that $L^\nu_t<\infty$ for all $t\in[0,T)$, $\mathbb{Q}^0$-a.s..

The following proposition contains some results similar to those in [7], extended to the 2-dimensional case but with slightly different assumptions (see also, [18, 20]). Although some proofs are quite similar, we give a detailed sketch for completeness.

Proposition 3.3. *Assume Assumption 3.1. Let (T,ξ_T,ν) be a given conditioning. Then,*

(i) there exists a unique $\mathbb{P}^\nu\in\mathscr{P}(\Omega)$, such that $\mathbb{P}^\nu\circ\xi_T^{-1}=\nu$, and for any $t<T$, any bounded $X\in\mathbb{L}^0(\mathcal{F}_t;\mathbb{R}^2)$, it holds that

$$\mathbb{E}^{\mathbb{Q}^0}\big[X|\xi_T=y\big]=\mathbb{E}^{\mathbb{Q}^0}\big[\eta^y_tX\big],\qquad t<T,\ \mathbb{Q}^0\circ\xi_T^{-1}\text{-a.e. }y\in\mathbb{R}^2;\tag{3.3}$$

(ii) assuming further that (T,ξ_T,ν) is proper, then for any $t<T$, it holds that

$$\frac{d\mathbb{P}^\nu}{d\mathbb{Q}^0}\bigg|_{\mathcal{F}_t}=\int_{\mathbb{R}^2}\eta^y_t\nu(dy);\tag{3.4}$$

(iii) L^ν is a $\mathbb{Q}^0$-martingale on $[0,T)$, and $L^\nu_T:=\lim_{t\to T}L^\nu_t$ exists, with $\mathbb{E}^{\mathbb{Q}^0}[L^\nu_T]\le1$.

Proof. Given conditioning (T,ξ_T,ν), let $\mathbb{Q}^y(\cdot)\in\mathscr{P}(\Omega)$ be the regular conditional probability defined by $\mathbb{Q}^y(A):=\mathbb{Q}^0(A|\xi_T=y)$, $A\in\mathcal{F}_T$, $y\in\mathbb{R}^2$, and define

$$\mathbb{P}^\nu(A):=\int_{\mathbb{R}^2}\mathbb{Q}^y(A)\nu(dy),\qquad A\in\mathcal{F}_T.\tag{3.5}$$

We now check (i). That $\mathbb{P}^\nu\circ\xi_T^{-1}=\nu$ is obvious. To see (3.3), we define a finite measure on $(\mathbb{R}^2,\mathscr{B}(\mathbb{R}^2))$ by $\mu^{X|\xi_T}(A):=\int_{\xi_T\in A}X(\omega)\mathbb{Q}^0(d\omega)$, $A\in\mathscr{B}(\mathbb{R}^2)$. Then, by definition we can write, for $A\in\mathscr{B}(\mathbb{R}^2)$,

$$\mu^{X|\xi_T}(A)=\int_A\mathbb{E}^{\mathbb{Q}^0}[X|\xi_T=y]\mathbb{Q}^0\circ\xi_T(dy)=\int_A\mathbb{E}^{\mathbb{Q}^0}[X|\xi_T=y]p(0,z_0;T,y)dy.$$

Since $X\in\mathbb{L}^0(\mathcal{F}_t;\mathbb{R}^2)$, using the Markov property on ξ and Fubini theorem we have

$$\begin{aligned}\mu^{X|\xi_T}(A)&=\int_\Omega\mathbb{E}^{\mathbb{Q}^0}[\mathbf{1}_{\{\xi_T\in A\}}X|\mathcal{F}_t](\omega)\mathbb{Q}^0(d\omega)=\int_\Omega\Big[\int_Ap(t,\xi_t(\omega);T,y)dy\Big]X(\omega)\mathbb{Q}^0(d\omega)\\&=\int_A\mathbb{E}^{\mathbb{Q}^0}[p(t,\xi_t;T,y)X]dy,\quad A\in\mathscr{B}(\mathbb{R}^2).\end{aligned}$$

Comparing the two equations above, we deduce (3.3).

(ii) To see (3.4), it suffices to show that if $Z \in \mathbb{L}^1_{\mathcal{F}_t}(\mathbb{R}^2, \mathbb{Q}^0)$, $t \in [0, T)$, then

$$\mathbb{E}^{\mathbb{P}^\nu}[Z] = \mathbb{E}^{\mathbb{Q}^0}[L^\nu_t Z] = \mathbb{E}^{\mathbb{Q}^0}\Big[\int_{\mathbb{R}^2} \eta^y_t \nu(dy) Z\Big]. \tag{3.6}$$

By a standard truncation, we may assume that Z is bounded. Then by (3.3), we have $\mathbb{E}^{\mathbb{Q}^0}[\eta^y_t Z] = \mathbb{E}^{\mathbb{Q}^0}[Z|\xi_T = y] = \mathbb{E}^{\mathbb{P}^\nu}[Z|\xi_T = y]$, thanks to definition (3.5), thus

$$\mathbb{E}^{\mathbb{P}^\nu}[Z] = \int_{\mathbb{R}^2} \mathbb{E}^{\mathbb{Q}^0}[Z|\xi_T = y]\nu(dy) = \int_{\mathbb{R}^2} \mathbb{E}^{\mathbb{Q}^0}[\eta^y_t Z]\nu(dy).$$

Comparing this to (3.6), we see that it suffices to show that $\int_{\mathbb{R}^2} \mathbb{E}^{\mathbb{Q}^0}[|\eta^y_t Z|]\nu(dy) < \infty$, 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 (3.4), $\frac{d\mathbb{P}^\nu}{d\mathbb{Q}^0}\Big|_{\mathscr{F}_t} := L^\nu_t$, $t < T$. Thus, L^ν is a $\mathbb{Q}^0$-martingale on $[0, T)$. 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 $\mathbb{E}[L^\nu_T] \le \lim_{t\to T} \mathbb{E}[L^\nu_t] = 1$. □

Remark 3.4. (1) The probability $\mathbb{P}^\nu$ in Proposition 3.3 is called the *minimal probability* given the proper conditioning (T, ξ_T, ν). Moreover, Proposition 3.3 shows that the assumption (A1) in [7] is automatically satisfied in our setting.

(2) Proposition 3.3-(ii) only indicates that $\mathbb{P}^\nu \ll \mathbb{Q}^0$ on each $\mathcal{F}_t$, $0 \le t < T$, with the Radon-Nikodým derivative defined by (3.4). But it does not imply that $\mathbb{P}^\nu$ and $\mathbb{Q}^0$ are equivalent on $\mathcal{F}_t$, for $t < T$, neither does it imply that $\mathbb{P}^\nu \ll \mathbb{Q}^0$ on $\mathcal{F}_T$. ■

We now turn our attention to a specific conditioning (T, ξ_T, ν) that will lead to the solution to an STPBVP (2.10). For notational convenience we shall now simply denote $\xi = (V, X)$, when there is no danger of confusion. Let $m^* \in \mathscr{P}(\mathbb{R})$ 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

$$\int_{\mathbb{R}} e^{\lambda_0 v^2} m^*(dv) < \infty. \tag{3.7}$$

We remark that the Assumption 3.5 is actually not over restrictive. In fact, in light of the well-known Fernique Theorem (cf. [17]) (3.7) covers a large class of normal random variables. Now let us define a probability measure $\nu \in \mathscr{P}(\mathbb{R}^2)$ by

$$\nu(A) = \int_{\mathbb{R}} \mathbf{1}_A(v, g^{-1}(v)) m^*(dv) = \int_{(v, g^{-1}(v)) \in A} m^*(dv). \tag{3.8}$$

That is, the measure ν concentrates on the graph of the function $v = g(x)$ (or $x = g^{-1}(v)$), 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 λ_0 in (3.7) being sufficiently large. Let ξ be the solution to (3.1), and $\nu \in \mathscr{P}(\mathbb{R}^2)$ be defined by (3.8). Then, (T, ξ_T, ν) is a proper conditioning. Furthermore, if $\mathbb{P}^\nu$ is the minimum probability given (T, ξ_T, ν), then it holds that*

$$\mathbb{P}^\nu\{V_T = g(X_T)\} = 1. \tag{3.9}$$

Proof. Since under Assumption 3.1 ξ is a diffusion process with positive transition density function (cf. e.g., [19]), we have $\mathrm{supp}(\mathbb{Q}^0 \circ \xi_T^{-1}) = \mathbb{R}^2$. Furthermore, by definition of ν (3.8), for the constants $\lambda > 0$ in (3.2) we deduce from (3.7) that

$$\int_{\mathbb{R}^2} e^{\frac{\lambda|\xi_0-y|^2}{T}}\nu(dy) = \int_{\mathbb{R}} e^{\frac{\lambda[(v_0-v)^2+(x_0-g^{-1}(v))^2]}{T}} m^*(dv) < \infty,$$

provided $\lambda_0 \geq \frac{2\lambda}{T}$, where λ_0 is the constant in (3.7). That is, (T, ξ_T, ν) 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 $\mathbb{R}^2$, are identical. Let us denote $\Gamma := \{(g(x), x) : x \in \mathbb{R}\} = \{(v, g^{-1}(v)) : v \in \mathbb{R}\} \subseteq \mathbb{R}^2$. Then, by definition (3.8) 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 (3.9). □

Remark 3.7. (1) Note that $\xi = (V, X)$ has continuous paths under $\mathbb{Q}^0$, thus the random variable ξ_{T-} and ξ_T have the same law under $\mathbb{Q}^0$. Then by definitions of the measures m^*, ν, and consequently $\mathbb{P}^\nu$, we see that (3.9) can also be written as

$$\mathbb{P}^\nu\{\lim_{t \nearrow T} V_t = V_{T-} = g(X_{T-}) = \lim_{t \nearrow T} g(X_t)\} = 1. \tag{3.10}$$

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 (3.15) would be a desirable candidate, except for a slight difference on the drift coefficients.

(2) By Proposition 3.3-(iii), L^ν 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, ξ_T, ν), a function

$$\varphi(t, z) = \int_{\mathbb{R}^2} \frac{p(t, z; T, y)}{p(0, z_0; T, y)}\nu(dy), \qquad z = (v, x), \tag{3.11}$$

where p is the transition density of ξ under $\mathbb{Q}^0$ (hence $p(\cdot, \cdot, T, y) \in C^{1,2}$). Clearly, $\varphi(0, z_0) = 1$ and $L_t = L_t^\nu = \varphi(t, \xi_t)$, $t \in [0, T)$. Now, applying Itô's formula we have

$$L_t = \varphi = 1 + \int_0^t [\varphi_t + \mathscr{L}[\varphi]]ds + \int_0^t \big(\nabla\varphi, \bar\sigma dB_s^0\big), \tag{3.12}$$

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

$$\begin{cases} \varphi_t + b\varphi_v + \mu(t, x)\varphi_x + \dfrac{1}{2}\sigma^2\varphi_{vv} + \dfrac{1}{2}\rho^2(t, x)\varphi_{xx} = 0; \\ \varphi(0, v_0, x_0) = 1. \end{cases} \tag{3.13}$$

Consequently, it follows from (3.12) that

$$dL_t = d\varphi = \big(\nabla\varphi, \bar\sigma dB_t^0\big) = L_t(\theta_t, dB_t^0), \qquad L_0 = 1, \quad t \in [0, T), \tag{3.14}$$

where $\theta_t := \bar\sigma^T(t, \xi_t)\frac{\nabla\varphi(t,\xi_t)}{\varphi(t,\xi_t)} = \bar\sigma^T(t, \xi_t)\nabla[\ln\varphi(t, \xi_t)]$, $t \in [0, T)$. Denote $W_t = B_t^0 - \int_0^t \theta_s ds$, then by Girsanov's theorem, $\{W_t\}$ is a 2-dimensional $\mathbb{P}^\nu$-Brownian motion on $[0, T)$. We have thus proved the following 2-dimensional extension of a result in [7].

Proposition 3.8 ([7, Proposition 37]). *Assume Assumption 3.1, and let* $\mathbb{P}^\nu$ *be the minimal probability corresponding to the conditioning* (T, ξ_T, ν)*, where* $\xi = (V, X)$ *is the strong solution to* (3.1)*. Then, under* $\mathbb{P}^\nu$*,* ξ *solves the following SDE:*

$$d\xi_t = [\bar{b} + \bar{\sigma}\theta_t]dt + \bar{\sigma}dW_t = \hat{b}dt + \bar{\sigma}dW_t, \quad \xi_0 = z, \quad 0 \le t < T, \tag{3.15}$$

where $(\bar{b}, \bar{\sigma})$ *are the same as those in (3.12),* $\hat{b}(t,z) := \bar{b}(t,z) + \bar{\sigma}\bar{\sigma}^T(t,z)\nabla[\ln\varphi(t,z)]$*, and* $\theta_t := (\theta^1_t, \theta^2_t)^T = \bar{\sigma}^T(t,\xi_t)\nabla[\ln\varphi(t,\xi_t)] = \frac{1}{\varphi}(\varphi_v\sigma, \varphi_x\rho)^T(t,\xi_t)$*;* φ *is defined by (3.11); and* $W = (W^1, W^2)$ *is a* $\mathbb{P}^\nu$*-Brownian motion.*

4. A Stochastic Two-Point Boundary Value Problem. We are now ready to study the STPBVP (2.10) 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)* $\mathbb{P} \in \mathscr{P}(\Omega)$ *and* $B = (B^1, B^2)$ *is a* $\mathbb{P}$*-Brownian motion on* $[0,T]$*; (ii)* $\alpha \in \mathscr{U}_{ad}$*, and* (V, X, α) *satisfies the SDE on* $(\Omega, \mathcal{F}, \mathbb{P})$*:*

$$\begin{cases} dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB^1_t, & V_0 = v; \\ dX_t = \big(\mu(t, X_t) + \alpha_t\rho(t, X_t)\big)dt + \rho(t, X_t)dB^2_t, & X_0 = x, \end{cases} \tag{4.1}$$

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

In particular, (V, X, α) *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* $\mathbb{F}^{V,X}$*-Markov process on* $[0,T)$*.* ■

Remark 4.2. (i) For notational clarity, when necessary we shall often refer to (4.1) as a "STPBVP(b, σ, μ, ρ)", and write the solution (V, X, α) to a STPBVP as (V^α, X^α) 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 σ 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 $\mathbb{F}^X$. Consequently, the results of this paper and those in the existing literature mutually exclusive. ■

To construct a weak solution, we first recall (3.14) and the $\mathbb{P}^\nu$-Brownian motion $W_t = B^0_t - \int_0^t \theta_s ds$; $t \in [0,T)$, where $\theta_t := (\theta^1_t, \theta^2_t)^T = \bar{\sigma}^T(t,\xi_t)\nabla[\ln\varphi(t,\xi_t)] = \frac{1}{\varphi}(\varphi_v\sigma, \varphi_v\rho)^T(t,\xi_t)$, $t \in [0,T)$, and under $\mathbb{P}^\nu$ the process $\xi_t := (V_t, X_t)^T$ satisfies the SDE (3.15). We note that although the coefficient $\hat{b}$ in (3.15) is explicitly defined, it depends on the solution of an ill-posed parabolic PDE (3.13), its behavior is a bit hard to analyze. The following lemma is useful to note.

Lemma 4.3. *Let* (T, ξ_T, ν) *be the conditioning in Lemma 3.6, and* $\mathbb{P}^\nu$ *the corresponding minimum probability. Then, it holds that* $\mathbb{L}^p_{\mathcal{F}_t}(\mathbb{R}^d; \mathbb{Q}^0) \subset \mathbb{L}^p_{\mathcal{F}_t}(\mathbb{R}^d; \mathbb{P}^\nu)$*,* $t < T$*. Specifically, for any* $T_0 < T$*, there exists a constant* $C_{T_0} > 0$*, that depends only on the coefficients* (b, σ, μ, ρ)*, and* T_0*, such that, for any* $X \in \mathcal{F}_t$*,* $t \in [0, T_0]$*, it holds that*

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

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

Proof. We first note that given $T_0 < T$, and $X \in \mathcal{F}_t$, $t \le T_0$, by Lemma 3.6-(ii), $\mathbb{E}^{\mathbb{P}^\nu}[|X|^p] = \mathbb{E}^{\mathbb{Q}^0}[L_{T_0}|X|^p] \le C_{T_0}\mathbb{E}^{\mathbb{Q}^0}[|X|^p]$, where $C_{T_0} = \frac{\widetilde{C}T}{T-T_0}\int_{\mathbb{R}^2} e^{\frac{\lambda|\xi_0-y|^2}{T}}\nu(dy)$, proving (4.2). The rest of the proof is obvious. ∎

Now for $n \in \mathbb{N}$, define $\theta_t^{(n)} := \theta_{t\wedge\tau_n}$, $t \in [0,T]$, where $\tau_n := \inf\{t > 0 : |\theta_t| \ge n\} \wedge T$. Clearly, under probability $\mathbb{P}^\nu$, for each $n \in \mathbb{N}$, the SDE

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

is (strongly) well-posed on $[0,T]$. Now recall from Remark 3.7 we know that under $\mathbb{P}^\nu$, the process $\xi = (V,X)$ has continuous paths on $[0,T]$ and solves (3.15) on $[0,T)$. Thus by pathwise uniqueness, it is readily seen that $\xi_t^{(n)} \equiv \xi_t$, $t \in [0,\tau_n]$, for any n.

We now write $\theta_t^{(n)} = (\theta_t^{1,n}, \theta_t^{2,n})$, $t \in [0,T]$. Since $\theta_t^{1,n}$ is bounded by n, and $\theta_t^{1,n} = \theta_t^{1,n+1}$, on $[0,\tau_n]$. By Girsanov's theorem, there exists a family of probabilities $\{\bar{\mathbb{P}}^{(n)}\}_{n\ge1}$ on $(\Omega,\mathcal{F})$ by

$$\frac{d\bar{\mathbb{P}}^{(n)}}{d\mathbb{P}^\nu}\Big|_{\mathscr{F}_T} = \mathscr{E}(\theta_T^{1,n}) := \exp\Big\{\int_0^T \theta_s^{1,n}dW_s^1 - \frac{1}{2}\int_0^T |\theta_s^{1,n}|^2 ds\Big\}.$$

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

$$\frac{d\bar{\mathbb{P}}^{(n+1)}}{d\mathbb{P}^\nu}\Big|_{\mathcal{F}_{\tau_n}} = \mathscr{E}(\theta_{\tau_n}^{1,n+1}) = \mathscr{E}(\theta_{\tau_n}^{1,n}) = \frac{d\bar{\mathbb{P}}^{(n)}}{d\mathbb{P}^\nu}\Big|_{\mathcal{F}_{\tau_n}}. \tag{4.4}$$

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

$$\bar{\mathbb{P}}|_{\mathcal{F}_{\tau_n}} := \bar{\mathbb{P}}^{(n)}|_{\mathcal{F}_{\tau_n}}, \qquad n \in \mathbb{N}, \tag{4.5}$$

then $\bar{\mathbb{P}} \ll \mathbb{P}^\nu$ 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 $\bar{\mathbb{P}}$-Brownian motion on $[0,T)$, whence on $[0,T]$, thanks to the Martingale Convergence Theorem. Further, under $\bar{\mathbb{P}}$, the process $\xi = (V,X)$ satisfies the SDE:

$$\begin{cases} dV_t = b(t,V_t,X_t)dt + \sigma(t,V_t,X_t)d\bar{B}_t^1, & V_0 = v; \\ dX_t = \big(\mu(t,X_t) + \rho(t,X_t)\theta_t^2\big)dt + \rho(t,X_t)dW_t^2, & X_0 = x; \end{cases} \quad t \in [0,T). \tag{4.6}$$

Comparing (4.6) and (4.1) and noting the facts (3.10) and $\bar{\mathbb{P}}|_{\mathcal{F}_t} \ll \mathbb{P}^\nu|_{\mathcal{F}_t}$, $t \in [0,T)$, we see that $(\bar{\mathbb{P}}, \bar{B}, V, X, \theta^2)$ is a weak solution to (4.1). We have the following result.

PROPOSITION 4.4. *Assume Assumption 3.1. Then there exists a weak solution* $(\mathbb{P}, B, V, X, \alpha)$ *to STPBVP (4.1). Furthermore,* $\mathbb{P}$ *can be chosen so that* $\mathbb{P}|_{\mathcal{F}_t} \ll \mathbb{Q}^0|_{\mathcal{F}_t}$, $t < T$, *and denoting* $V_T := V_{T-} = \lim_{t\nearrow T} V_t$, *it holds that* $\mathbb{P} \circ (V_T)^{-1} = m^*$.

Proof. Consider the probability $\bar{\mathbb{P}}$ defined by (4.4), (4.5) and SDE (4.6). We first claim $\bar{\mathbb{P}} \ll \mathbb{P}^\nu$ on $\mathcal{F}_{T-}$. Indeed, let $\mathscr{A} := \{\mathcal{G} \subset \mathcal{F} : \bar{\mathbb{P}} \ll \mathbb{P}^\nu \text{ on } \mathcal{G}\}$, then $\mathcal{F}_{\tau_n} \in \mathscr{A}$, $n \in \mathbb{N}$. Since $\tau_n \nearrow T$, we have $\mathcal{F}_{T-} = \bigvee_n \mathcal{F}_{\tau_n}$ (see, e.g., [29, Exercise 1.27 or Theorem 3.6]), and thus $\mathcal{F}_{T-} \in \mathscr{A}$, thanks to the Monotone Class Theorem.

Next, since $\{\lim_{t\nearrow T} V_t \ne \lim_{t\nearrow T} g(X_t)\} = \bigcup_m \bigcap_N \bigcup_{r\in\mathbf{Q}(T-\frac{1}{N},T)} \{|V_r - g(X_r)| \ge \frac{1}{m}\} \in \mathcal{F}_{T-}$, where $\mathbf{Q}$ is the rationals in $\mathbb{R}_+$, and $\mathbf{Q}(A) := \mathbf{Q} \cap A$, $A \in \mathcal{B}(\mathbb{R})$, and

$\bar{\mathbb{P}} \ll \mathbb{P}^\nu$ on $\mathcal{F}_{T-}$, we have $\bar{\mathbb{P}}\{\lim_{t\nearrow T} V_t \neq \lim_{t\nearrow T} g(X_t)\} = 0$, thanks to (3.10). That is, $\bar{\mathbb{P}}\{\lim_{t\nearrow T} V_t = \lim_{t\nearrow T} g(X_t)\} = 1$. Now let $\alpha = \theta^2$ in SDE (4.6), we see that $(\bar{\mathbb{P}}, \bar{B}, V, X, \alpha)$ is a weak solution to STPBVP (4.1).

It remains to check the last statement. To this end, let $\xi = (V, X)$. Since $\bar{\mathbb{P}} \ll \mathbb{P}^\nu \ll \mathbb{Q}^0$ on $\mathcal{F}_{T-}$ and $\mathbb{Q}^0\{\xi \in \mathbb{C}([0,T];\mathbb{R}^2)\} = 1$, we can naturally extend ξ to $[0,T]$ by setting $\xi_T = \lim_{t\nearrow T} \xi_t$ so that $\mathbb{P}^\nu\{\xi \in \mathbb{C}([0,T];\mathbb{R}^2)\} = \bar{\mathbb{P}}\{\xi \in \mathbb{C}([0,T];\mathbb{R}^2)\} = 1$ as well. We first claim that $\mathbb{P}^\nu \circ V_T^{-1} = m^*$. Indeed, let $B \in \mathscr{B}(\mathbb{R})$ and $A := B \times \mathbb{R} \in \mathscr{B}(\mathbb{R}^2)$. By (3.8) 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 $\bar{\mathbb{P}} \circ V_T^{-1} = m^*$, 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 $\xi_t(\omega) = \Phi(t, B^0_{\cdot\wedge t}(\omega)) = \Phi(t,\omega)$, $(t,\omega) \in [0,T]\times\Omega$, for some (progressively) measurable function $\Phi : [0,T] \times \Omega \mapsto \mathbb{R}^2$. Consequently, we can write $\theta_t^2(\omega) = (\ln \varphi(t, \xi_t(\omega)))_x = (\ln \varphi(t, \Phi(t,\omega)))_x$, $(t,\omega) \in [0,T] \times \Omega$. By virtue of Lemma 4.3, the process θ^2 is well-defined on $[0,T) \times \Omega$, $\bar{\mathbb{P}}$-a.s. and $\theta_t^2 \in \mathbb{L}^2(\bar{\mathbb{P}})$, for $t \in [0,T)$.

Now let us denote the solutions to (3.15) and (4.6) as $(\tilde{V}_t, \tilde{X}_t)$ and $(\bar{V}_t, \bar{X}_t)$ respectively. Then we see that $((\tilde{X}_t, W_t^2), \mathbb{P}^\nu)$ and $((\bar{X}_t, W_t^2), \bar{\mathbb{P}})$ are two weak solutions to the same SDE, well-defined on any $[0,T_0] \subset [0,T)$. Consequently, we have $\mathbb{P}^\nu \circ \tilde{X}^{-1} = \bar{\mathbb{P}} \circ \bar{X}^{-1}$ on $[0,T_0]$ for any $T_0 < T$. Extending the solution to $[0,T]$, we have $\mathbb{P}^\nu \circ \tilde{X}_T^{-1} = \bar{\mathbb{P}} \circ \bar{X}_T^{-1}$. Since $V_T = g(X_T)$, both $\bar{\mathbb{P}}$-a.s. and $\mathbb{P}^\nu$-a.s., we obtain that $\bar{\mathbb{P}} \circ V_T^{-1} = \mathbb{P}^\nu \circ V_T^{-1} = m^*$, 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 $(\bar{\mathbb{P}}, \bar{B}, V, X, \alpha)$ that we constructed has the following properties:

(i) there exists a sequence of $\bar{\mathbb{P}}$-stopping times $\{\tau_n\}$, and a sequence of probabilities $\bar{\mathbb{P}}^{(n)}$ on $(\Omega, \mathcal{F})$, such that $\tau_n \nearrow T$, $\bar{\mathbb{P}}$-a.s., and $\bar{\mathbb{P}}|_{\mathcal{F}_{\tau_n}} = \bar{\mathbb{P}}^{(n)}|_{\mathcal{F}_{\tau_n}}$, $n \in \mathbb{N}$;

(ii) for each $n \in \mathbb{N}$, $\bar{B} = \bar{B}^{(n)}$ on $[0, \tau_n]$, where $\bar{B}^{(n)} = (\bar{B}^{(n,1)}, \bar{B}^{(n,2)})$ is a $\mathbb{P}^{(n)}$-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 $[0,T]$:

$$\text{(4.7)} \qquad \begin{cases} dV_t = b(t, V_t, X_t)dt + \sigma(t, V_t, X_t)dB_t^{(n,1)}, & V_0 = v; \\ dX_t = \big(\mu(t, X_t) + \rho(t, X_t)\alpha_t^{(n)}\big)dt + \rho(t, X_t)dB_t^{(n,2)}, & X_0 = x, \end{cases}$$

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

(iv) $\bar{\mathbb{P}}|_{\mathcal{F}_t} \ll \mathbb{P}^\nu|_{\mathcal{F}_t} \ll \mathbb{Q}^0|_{\mathcal{F}_t}$, $t \in [0,T)$.

In what follows we shall denote $(\bar{\mathbb{P}}, \{\tau_n\})$ to specify that $\bar{\mathbb{P}}$ 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 [27].

Definition 4.5. *We call a weak solution $(\bar{\mathbb{P}}, \bar{V}, \bar{X}, \bar{B}, \alpha)$ of STPBVP (4.1) satisfying (i)–(iii) above a "nested weak solution" and the corresponding family of stopping times $\{\tau_n\}$ the "announcing sequence" of probability $\bar{\mathbb{P}}$. We call $(\{\tau_n\}, \alpha)$ the characteristic pair of the weak solution.*

Furthermore, a nested weak solution is called a $\mathbb{P}^\nu$-weak solution if (iv) holds. ■

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 $\{\tau_n^1\}, \{\tau_n^2\}$ are two announcing sequences of stopping times, then so is $\{\tau_n^1 \wedge \tau_n^2\}$. Thus the weak solution is independent of the choice of the announcing sequence $\{\tau_n\}$. Since the process α determines the coefficient of SDE (4.6), whence the solution, we often specify its role by calling $(\bar{\mathbb{P}}, \bar{V}, \bar{X}, \bar{B}, \alpha)$ the α-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 (4.1) on* $[0, T)$, *such that* $\mathbb{P}^1 = \mathbb{P}^2 = \mathbb{P}$, $\xi_0^1 = \xi_0^2$, *and* $\mathbb{P}\{\alpha_t^1 = \alpha_t^2,\ B_t^1 = B_t^2,\ t \in [0, T)\} = 1$, *then* $\mathbb{P}\{\xi_t^1 = \xi_t^2,\ t \in [0, T_0]\} = 1$, *for any* $T_0 < T$.

Remark 4.8. The time T_0 in Definition 4.7 can be changed to any stopping time τ 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 $\Delta\xi := \xi_t^1 - \xi_t^2$, then we obtain

$$\mathbb{E}\big[|\Delta\xi|_{T_0}^*\big] \leq \mathbb{E}\big[|\Delta\xi|_{\tau_n}^* \mathbf{1}_{\{T_0 \leq \tau_n\}}\big] + \mathbb{E}\big[|\Delta\xi|_{T_0}^* \mathbf{1}_{\{T_0 > \tau_n\}}\big] \leq \mathbb{E}\big[|\Delta\xi|_{T_0}^* \mathbf{1}_{\{T_0 > \tau_n\}}\big];$$

where $|\eta|_\tau^* := \sup_{t\in[0,\tau]} |\eta_t|$, for $\tau > 0$ and $\eta \in \mathbb{C}([0, \tau])$. Similarly, for any $T_0 < T$,

$$\mathbb{E}\big[|\Delta\xi|_\tau^*\big] \leq \mathbb{E}\big[|\Delta\xi|_{T_0}^* \mathbf{1}_{\{\tau \leq T_0\}}\big] + \mathbb{E}\big[|\Delta\xi|_\tau^* \mathbf{1}_{\{\tau > T_0\}}\big] \leq \mathbb{E}\big[|\Delta\xi|_\tau^* \mathbf{1}_{\{\tau > T_0\}}\big].$$

Since $\lim_{n\to\infty} \mathbb{P}\{T_0 > \tau_n\} = 0$ and $\lim_{T_0 \nearrow T} \mathbb{P}\{\tau > T_0\} = 0$, 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 stopping time τ, with $\mathbb{P}\{\tau < T\} = 1$. ■

The definition of the uniqueness in law for the STPBVP is a bit more involved. First note that the component "α" 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 α is properly fixed. To this end, denote $\mathscr{A} := \{A \in \mathscr{B}([0, T]) \otimes \mathcal{F} : A_t \in \mathcal{F}_t,\ t \in [0, T]\}$, where A_t is the t-section of A; and denote all $\mathscr{A}$-measurable functions by $\mathbb{L}^0_{\mathscr{A}}([0, T] \times \Omega)$. We should note that the space $\mathbb{L}^0_{\mathscr{A}}([0, T] \times \Omega)$ is independent of any probability measure, and we can therefore use it to identify the α-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* α*-weak solutions* $(\bar{\mathbb{P}}^i, \bar{V}^i, \bar{X}^i, \bar{B}^i, \bar{\alpha}^i)$, $i = 1, 2$ *of (4.1) on* $[0, T)$, *such that* $(v^1, x^1) = (v^2, x^2)$; $\bar{\mathbb{P}}^1 \circ (\tau_n^1)^{-1} = \bar{\mathbb{P}}^2 \circ (\tau_n^2)^{-1}$, $n \in \mathbb{N}$; *and* $\bar{\mathbb{P}}^i\{\bar{\alpha}_t^i = \alpha_t,\ t \in [0, T)\} = 1$, $i = 1, 2$, *for some* $\alpha \in \mathbb{L}^0_{\mathcal{A}}([0, T] \times \Omega)$, *then for any cylindrical set* $E_{t_1,\cdots,t_n}^{A_1,\ldots,A_n} := \{(\mathbf{v}, \mathbf{x}) \in \mathbb{C}([0, T]; \mathbb{R}^2) : (\mathbf{v}, \mathbf{x})(t_i) \in A_i,\ i = 1, \cdots, n\}$, *where* $0 \leq t_1 < t_2 < \cdots < t_n < T$ *and* $A_i \in \mathscr{B}(\mathbb{R}^2)$, $i = 1, \cdots, n$, *it holds that*

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

We now give the main theorem of this subsection.

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

The proof of Proposition 4.10 is based on a lemma that is interesting in its own right.

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

Proof. Let $(\bar{\mathbb{P}}, \bar{\xi}, \bar{\alpha})$ be the nested Markovian weak solution. Then $\bar{\alpha}_t = u(t, \bar{\xi}_t)$, $t \in [0,T]$, for some $u \in \mathbb{L}^0([0,T] \times \mathbb{R}^2)$. 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 $T_0 < T$, $\bar{\xi}$ is the pathwisely unique strong solution on $[0, T_0]$, and there exists a $\Phi^{T_0} \in \mathbb{L}^0_{\mathscr{A}}([0, T_0] \times \Omega)$, such that $\bar{\xi}_t = \Phi^{T_0}(t, \cdot)$, $t \in [0, T_0]$, $\bar{\mathbb{P}}$-a.s.. 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 $\bar{\xi}_t = \Phi(t, \cdot)$, $t \in [0, T)$, $\bar{\mathbb{P}}$-a.s.. Since $\bar{\alpha}_t = u(t, \bar{\xi}_t) = u(t, \Phi(t, \cdot))$ by assumption, we have $\alpha_t = \bar{\alpha}_t = u(t, \Phi(t, \cdot))$, $dt \otimes d\mathbb{P}$-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 (4.1) 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 $\alpha^i = u(t, \bar{\xi}^i)$, $i = 1, 2$, from the interval $[-m, m]$.

Next, let the cylindrical set $E^{A_1, \ldots, A_n}_{t_1, \ldots, t_n}$ be given, with $t_n < T$. Since $\tau^i_m \nearrow T$, we can write $(\bar{\xi}^i)^{-1}(E^{A_1, \ldots, A_n}_{t_1, \ldots, t_n}) = \cap_{j=1}^n (\bar{\xi}^i_{t_j})^{-1}(A_j) = \cup_{m=1}^\infty \cap_{j=1}^n \{\tau^i_m \geq t_j\} \cap (\bar{\xi}^i_{t_j})^{-1}(A_j)$, $i = 1, 2$. Denoting $E^i_{j,m} := \{\tau^i_m \geq t_j\} \cap (\bar{\xi}^i_{t_j})^{-1}(A_j) = \{\tau^i_m \geq t_j\} \cap (\bar{\xi}^{i,(m)}_{t_j})^{-1}(A_j)$, 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, one has

$$\{\tau^i_m \leq t\} \cap E^i_{j,m} = \{t_j \leq \tau^i_m \leq t\} \cap (\bar{\xi}^{i,(m)}_{t_j})^{-1}(A_j) \in \mathcal{F}_t, \quad t \in [0, T), i = 1, 2.$$

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 $\hat{E}_m$ is increasing in m, thanks to the extension nature of solutions $\bar{\xi}^{i,(m)}$. 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

$$\bar{\mathbb{P}}^i \circ (\bar{\xi}^i)^{-1}(E^{A_1, \ldots, A_n}_{t_1, \ldots, t_n}) = \bar{\mathbb{P}}^i\big\{\cup_{m=1}^\infty \hat{E}^i_m\big\} = \lim_{m \to \infty} \bar{\mathbb{P}}^i\big\{\hat{E}^i_m\big\} = \lim_{m \to \infty} \bar{\mathbb{P}}^{i,(m)}\big\{\hat{E}^i_m\big\} \tag{4.8}$$

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 $\Phi \in \mathbb{L}^0_{\mathscr{A}}([0,T] \times \Omega)$. In other words, $(\bar{\mathbb{P}}^{i,(m)}, \bar{\xi}^{i,(m)})$, $i = 1, 2$, satisfy the same SDE (4.7) on $[0, \tau_m]$ 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 $\bar{\mathbb{P}}^{1,(m)}\big\{\hat{E}^1_m\big\} = \bar{\mathbb{P}}^{2,(m)}\big\{\hat{E}^2_m\big\}$, $m \in \mathbb{N}$, 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" (μ, ρ), we first argue that (μ, ρ) can be chosen so that the equilibrium strategy takes a particular form. Specifically, from Propositions 3.8 and 4.6 we see that the α-component in a weak solution is closely related to an ill-posed parabolic PDE (3.13), and in light of the well-known Widder's Theorem and its extensions (cf. e.g., [6, 30, 33, 32]), 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

$$u(t, v, x) = \rho(t, x)(\ln \varphi)_x = u_0(t, x) + u_1(t, x)v, \quad (t, v, x) \in [0, T) \times \mathbb{R}^2, \tag{5.1}$$

for some functions $u_0, u_1 : [0, T] \times \mathbb{R} \to \mathbb{R}$ to be determined later. In what follows we call a function u of the form (5.1) as having an *Affine Structure*.

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

$$\alpha_t = \beta_t(V_t - P_t), \qquad t \in [0, T), \tag{5.2}$$

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, 24]), as well as dynamic information case (see, e.g., [27]). 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 φ must have the form $\varphi(t, v, x) = \exp\{I(t, v, x)\}$, where

$$I(t, v, x) = h(t, v) + A(t, x) + B(t, x)v, \tag{5.3}$$

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

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

Now assume that φ satisfies the PDE (3.13), then we derive a PDE for function I:

$$\begin{cases} I_t + bI_v + \mu(t, x)I_x + \frac{1}{2}\sigma^2[(I_v)^2 + I_{vv}] + \frac{1}{2}\rho^2(t, x)[(I_x)^2 + I_{xx}] = 0; \\ I(0, v, x) = h(0, v) + A(0, x) + B(0, x)v. \end{cases} \tag{5.4}$$

Plugging (5.3) into (5.4) we obtain

$$\begin{aligned} 0 =& \frac{1}{2}\rho^2(t, x)B_x^2 v^2 + \left\{B_t + \mu(t, x)B_x + \frac{1}{2}\rho^2(t, x)[B_{xx} + A_x B_x]\right\}v + A_t \\ &+ \mu(t, x)A_x + \frac{1}{2}\rho^2(t, x)[A_{xx} + A_x^2] + h_t + b[h_v + B] + \frac{1}{2}\sigma^2\{h_{vv} + [h_v + B]^2\}. \end{aligned} \tag{5.5}$$

For notational simplicity, for given coefficients b, σ, μ, ρ, we define

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

Then, (5.5) becomes

$$I_2(t, x)v^2 + I_1(t, x)v + I_0(t, x) + G(t, v, x) = 0, \quad (t, v, x) \in [0, T] \times \mathbb{R}^2. \tag{5.7}$$

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

PROPOSITION 5.1. *The function $u(t, v, x) = \rho(t, x)(\ln \varphi(t, v, x))_x$ has an affine structure (5.1), where φ solves (3.13), if and only if the coefficients b, σ, μ, ρ satisfy the compatibility conditions (5.7) with I_0-I_2 and G being defined respectively by (5.6).*
Furthermore, it holds that $G_{vvv}(t, v, x) \equiv 0$, $(t, v, x) \in [0, T] \times \mathbb{R}^2$. ■

We should note that the compatibility condition (5.7) 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" (μ, ρ). In what follows we impose some specific structures on the functions h, b and σ, and try to find the conditions under which the function $u(t, v, x)$ is of an affine structure.

We begin with an example of a Kyle-Back problem 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 [27]. Namely, we assume $b(t,v,x) = f_t v + g_t x + k_t$, $\sigma(t,v,x) = 1$. Denote $X_t = P_t = \mathbb{E}^{\mathbb{P}}[V_t|\mathcal{F}_t^Y]$. Then, by [27, Theorem 3.6], we have $\mu(t,x) = (f_t + g_t)x + k_t$, and $\rho(t,x) = \rho(t) = S_t\beta_t$, where S_t satisfies a (deterministic) Riccati equation. Furthermore, in [27] it was shown that the equilibrium strategy takes the form (5.2). That is, the equilibrium α 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

$$\begin{cases} A(t,x) = \displaystyle\int_0^x \frac{u_0(t,y)}{\rho(t,y)}dy = -\frac{1}{S_t}\int_0^x ydy = -\frac{x^2}{2S_t}; \\ B(t,x) = \displaystyle\int_0^x \frac{u_1(t,y)}{\rho(t,y)}dy = \int_0^x \frac{1}{S_t}dy = \frac{x}{S_t}. \end{cases}$$

Plugging these into (5.6) and noting that S satisfies the Riccati equation $\frac{dS_t}{dt} = 2f_tS_t - \beta_t^2 S_t^2 + 1$, $t \in [0,T)$, we see that the compatibility condition (5.7) holds. ■

In general nonlinear cases, the analysis becomes too complicated to have a generic result. We therefore consider some special cases that might be useful in practice.

Case 1. $h = h(t)$, $b(t,x,v) = b(t,x)$ and $\sigma(t,v,x) = \sigma(t,x)$. Then, (5.5) becomes

$$I_0(t,x) + I_1(t,x)v + I_2(t,x)v^2 = 0, \tag{5.8}$$

where $I_0 = h_t + bB + \frac{1}{2}\sigma^2 B^2 + A_t + \mu A_x + \frac{1}{2}\rho^2(A_{xx} + A_x^2)$, $I_1 = B_t + \mu B_x + \frac{1}{2}\rho^2(B_{xx} + A_x B_x)$, and $I_2 = \frac{1}{2}\rho^2 B_x^2$. Clearly, (5.8) implies that $I_0 = I_1 = I_2 = 0$. Then, by definition $B_x = \frac{u_1}{\rho} = 0$, which implies $u_1(t,x) \equiv 0$, and $B(t,x) \equiv 0$. It then follows

$$h_t + A_t + \mu A_x + \frac{1}{2}\rho^2(A_{xx} + A_x^2) = 0. \tag{5.9}$$

That is, a necessary condition for affine structure is that $u_1 \equiv 0$ and (5.9) holds.

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 (5.5) and denote

$$\begin{cases} I_0 = h_t + b_0 B + \frac{1}{2}\sigma_0^2 B^2 + A_t + \mu A_x + \frac{1}{2}\rho^2(A_{xx} + A_x^2); \\ I_1 = b_1 B + \sigma_0\sigma_1 B^2 + B_t + \mu B_x + \frac{1}{2}\rho^2(B_{xx} + A_x B_x); \\ I_2 = \frac{1}{2}\rho^2 B_x^2 + \frac{1}{2}\sigma_1^2 B^2. \end{cases}$$

We see from $I_2 = 0$ that $u_1 \equiv 0$, which again leads to (5.9).

Case 3. $h = h(t)$, $b = b_0(t,x) + b_1(t,x)v + b_2(t,x)v^2$, $\sigma = \sigma_0(t,x) + \sigma_1(t,x)v$. Then,

$$\begin{cases} I_0 = h_t + b_0 B + \frac{1}{2}\sigma_0^2 B^2 + A_t + \mu A_x + \frac{1}{2}\rho^2(A_{xx} + A_x^2) = 0; \\ I_1 = b_1 B + \sigma_0\sigma_1 B^2 + B_t + \mu B_x + \frac{1}{2}\rho^2(B_{xx} + A_x B_x) = 0; \\ I_2 = \frac{1}{2}\rho^2 B_x^2 + \frac{1}{2}\sigma_1^2 B^2 + b_2 B = 0. \end{cases} \tag{5.10}$$

In particular, $I_2 = 0$ if and only if

$$u_1^2(t,x) = -\sigma_1^2\Big(\int_{x_0}^x \frac{u_1(t,y)}{\rho(t,y)}dy\Big)^2 - 2b_2\int_{x_0}^x \frac{u_1(t,y)}{\rho(t,y)}dy. \tag{5.11}$$

If we choose $u_1 = \rho$, then (5.11) implies $\rho^2 = -\sigma_1^2(x-x_0)^2 - 2b_2(x-x_0)$. Using $I_1 = 0$ in (5.10), we can write u_0 as

$$u_0 = \frac{2}{u_1}\big[-B_t - \mu B_x - \frac{1}{2}\rho^2 B_{xx} - b_1 B - \sigma_0\sigma_1 B^2\big]. \tag{5.12}$$

Therefore, (5.10), (5.11), and (5.12) guarantee the affine structure in this case.

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

$$\begin{cases} I_0 = (h_0)_t + b_0(h_1+B) + \frac{1}{2}\sigma_0^2(h_1+B)^2 + A_t + \mu A_x + \frac{1}{2}\rho^2(A_{xx} + A_x^2) = 0; \\ I_1 = (h_1)_t + b_1(h_1+B) + \sigma_0\sigma_1(h_1+B)^2 + B_t + \mu B_x + \frac{1}{2}\rho^2(B_{xx} + A_x B_x) = 0; \\ I_2 = \frac{1}{2}\rho^2 B_x^2 + \frac{1}{2}\sigma_1^2(h_1+B)^2 + b_2(h_1+B) = 0. \end{cases}$$

Case 5. $h = h_0(t)+h_1(t)v+h_2(t)v^2$. Since there are the terms bh_v, $\sigma^2 h_v^2$ in $G(t,v,x)$, and h is quadratic, we see that $\sigma(t,v,x)$ must be independent of v, and b is linear in v. We thus assume that $b = b_0(t,x) + b_1(t,x)v$, $\sigma = \sigma(t,x)$, in other words,

$$\begin{cases} I_0 = (h_0)_t + b_0(h_1+B) + \frac{1}{2}\sigma^2[2h_2 + (h_1+B)^2] + A_t + \mu A_x + \frac{1}{2}\rho^2(A_{xx} + A_x^2) = 0; \\ I_1 = (h_1)_t + 2b_0h_2 + b_1(h_1+B) + 2\sigma^2(h_1+B)h_2 + B_t + \mu B_x + \frac{1}{2}\rho^2(B_{xx} + A_x B_x) = 0; \\ I_2 = (h_2)_t + 2b_1h_2 + 2\sigma^2 h_2^2 + \frac{1}{2}\rho^2 B_x^2 = 0. \end{cases}$$

6. The Filtering Problem and FBSDE under Affine Structure. A popular approach in studying Kyle-Back equilibrium problem is nonlinear filtering (cf. e.g., [1, 16, 27]). In fact, when the market price is in the form of an *optional projection*: $P_t = \mathbb{E}[V_t|\mathcal{F}_t^Y]$, $t \in [0,T]$, 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. Let $(\bar{\mathbb{P}}, V, X, B, \alpha)$ be a (Markovian) weak solution, with $\alpha_t = u(t,V_t,X_t)$, and under $\bar{\mathbb{P}}$,

$$\begin{cases} dV_t = b(t,V_t,X_t)dt + \sigma(t,V_t,X_t)dB_t^1, & V_0 = v_0; \\ dX_t = [\mu(t,X_t) + \rho(t,X_t)u(t,V_t,X_t)]dt + \rho(t,X_t)dB_t^2, & X_0 = x_0; \\ dY_t = u(t,V_t,X_t)dt + dB_t^2, & Y_0 = 0. \end{cases} \tag{6.1}$$

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 Fujisaki-Kallianpur-Kunita (FKK) equation ([22, Theorem 4.1]) holds for $P_t = \mathbb{E}[V_t|\mathcal{F}_t^Y]$. To this end, we assume that u has the *affine structure*: $u = u_0(t,x) + u_1(t,x)v$. Denoting $\alpha_t = u(t,V_t,X_t)$, and consider the SDE:

$$dM_t = -\alpha_t M_t dB_t^2, \quad M_0 = 1, \qquad t \in [0,T]. \tag{6.2}$$

The following result is a modification of [8, Lemma 4.1.1] to the current case.

PROPOSITION 6.1. *Assume Assumptions 3.1, and that the function u in (6.1) satisfies $|u(t,v,x)| \le K(t)(1+|v|+|x|)$, $(t,v,x) \in [0,T) \times \mathbb{R}^2$, for some function $K \in \mathbb{L}^2([0,T];\mathbb{R}_+)$. 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 $\mathbb{E}[M_t] \le \lim_{n\to\infty} \mathbb{E}[M_{t\wedge\tau_n}] = \mathbb{E}[M_0] = 1$, where $\{\tau_n\}$ is any announcing sequence for M, and M is a true martingale iff $\mathbb{E}[M_t] = 1, 0 \le t \le T$, 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 $\lim_{\varepsilon\to 0} \mathbb{E}[M_t^\varepsilon] = \mathbb{E}[M_t]$. On the other hand, by a simple application of Itô's formula and then taking expectation one has

$$\mathbb{E}[M_t^\varepsilon] := \frac{1}{1+\varepsilon} - \mathbb{E}\Big[\int_0^t G^\varepsilon(\alpha_s, M_s)ds\Big],\ t \in [0,T],$$

where $G^\varepsilon(\alpha, x) := \frac{\varepsilon\alpha^2x^2}{(1+\varepsilon x)^3}$. It is easy to check that there exists $C > 0$, such that $|G^\varepsilon(\alpha,x)| \le C\alpha^2 x$, for all $\varepsilon, x > 0$. Denoting $U_t := M_t(V_t^2 + X_t^2)$, then the linear growth assumption for α_t gives $\mathbb{E}[G^\varepsilon(\alpha_t, M_t)] \le C\mathbb{E}[\alpha_t^2 M_t] \le CK^2(t)\big[1 + \mathbb{E}[U_t]\big]$. We claim that $\sup_{t\in[0,T]} \mathbb{E}[U_t] < \infty$. 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 $|\xi_0|^2 = v_0^2 + x_0^2$)

$$f_\varepsilon(U_t) = \frac{|\xi_0|^2}{1+\varepsilon|\xi_0|^2} + \int_0^t \frac{2M_s\big[V_s b_s + X_s\mu_s + \frac{1}{2}(\sigma_s^2+\rho_s^2)\big]}{(1+\varepsilon U_s)^2}ds + \int_0^t \frac{2M_sV_s\sigma_s}{(1+\varepsilon U_s)^2}dB_t^1$$
$$+ \int_0^t \frac{-\varepsilon\big[4V_s^2\sigma_s^2M_s^2 + \big(2M_sX_s\rho_s - U_s\alpha_s\big)^2\big]}{(1+\varepsilon U_s)^3}ds + \int_0^t \frac{-U_s\alpha_s + 2M_sX_s\rho_s}{(1+\varepsilon U_s)^2}dB_s^2.$$

Taking expectation on both sides, and by the linear growth of b, σ, μ and ρ, we obtain

$$\begin{aligned}\mathbb{E}[f_\varepsilon(U_t)] &\le |\xi_0|^2 + \int_0^t \mathbb{E}\Big[\frac{2M_s\big[V_sb_s + X_s\mu_s + \dfrac{1}{2}(\sigma_s^2+\rho_s^2)\big]}{(1+\varepsilon U_s)^2}\Big]ds \\ &\le |\xi_0|^2 + \int_0^t L(\mathbb{E}[f_\varepsilon(U_t)] + 1)ds.\end{aligned}$$

Now, first applying Gronwall's inequality and then applying Fatou's lemma (sending $\varepsilon \to 0$), we deduce that $\sup_{t\in[0,T]} \mathbb{E}[U_t] < \infty$, 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 [22, Theorem 4.1] to derive the FKK equation for $P_t = \mathbb{E}^{\bar{\mathbb{P}}}[V_t|\mathcal{F}_t^Y]$, which takes the following form:

$$(6.3)\quad \begin{cases} dP_t = [\mathbb{E}^t[b(t,V_t,X_t)] - \mathbb{E}^t[u(t,V_t,X_t)]Z_t]dt + Z_t dY_t, \\ Z_t := \mathbb{E}^t[V_t u(t,V_t,X_t)] - P_t\mathbb{E}^t[u(t,V_t,X_t)], \end{cases} \qquad t \in [0,T],$$

where $\mathbb{E}^t[\cdot] := \mathbb{E}^{\bar{\mathbb{P}}}[\cdot|\mathcal{F}_t^Y]$, $t \in [0,T]$. 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 $\mathbb{F}^Y$-adapted, then for $t \in [0,T]$,(6.3) can be rewritten as

$$(6.4)\quad dP_t = \{b_0(t,X_t) + b_1(t,X_t)P_t - (u_0(t,X_t) + u_1(t,X_t)P_t)Z_t\}dt + Z_t dY_t,$$

Let us now choose $\alpha_t = u(t,V_t,X_t)$, $t \in [0,T]$, to be the α-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 $\bar{\mathbb{Q}}$ on the canonical space $(\Omega, \mathcal{F})$ by

$\frac{d\bar{\mathbb{Q}}}{d\bar{\mathbb{P}}}|_{\mathcal{F}_T} = M_T$, then under $\bar{\mathbb{Q}}$, the process Y (for the given α) is a Brownian motion, and $\bar{\mathbb{Q}}\{V_T = g(X_T)\} = \bar{\mathbb{P}}\{V_T = g(X_T)\} = 1$. In other words, under $\bar{\mathbb{Q}}$, we can rewrite (6.4) and the SDE (4.1) for X as the following forward-backward SDE (FBSDE):

$$\text{(6.5)} \qquad \begin{cases} dX_t = \mu(t, X_t)dt + \rho(t, X_t)dY_t, \quad X_0 = x; \\ dP_t = [\beta_0(t, X_t, P_t) + \beta_1(t, X_t, P_t)Z_t]dt + Z_t dY_t, \quad P_T = g(X_T), \end{cases}$$

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

(ii) Under Assumption 3.1, X is a diffusion driven by the $\bar{\mathbb{Q}}$-Brownian motion Y, hence it is $\mathbb{F}^Y$-adapted, which justifies (6.4), whence (6.5). ■

We should note that the FBSDE (6.5) is actually "decoupled", in the sense that the forward SDE is independent of the backward components (Y, Z). But the BSDE in (6.5) is somewhat non-standard in that the coefficients are neither Lipschitz nor of linear growth. Specifically, the fact that $|\beta_1(t, x, y)z| \le K(1 + |y||z|)$ makes it super-linear in (y, z), and is beyond the usual "quadratic BSDE" framework. Nevertheless, the well-posedness of (6.5) can be argued via a more or less standard localization argument following the idea of [25]. Since this is not the main purpose of the paper, we shall only state the following result, but omit the proof (see [31] for details).

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

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

(ii) denoting $Y_t = B_t^2 + \int_0^t \alpha_s ds$ and $P_t = \mathbb{E}^{\bar{\mathbb{P}}}[V_t|\mathcal{F}_t^Y]$, $t \in [0, T]$, then Y is a $\bar{\mathbb{Q}}$-Brownian motion, and under $\bar{\mathbb{Q}}$, (X, P) satisfies the FBSDE (6.5). ■

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., [26, Chapter 4], [28, Section 2]) that, if (X, P, Z) is the solution to the FBSDE (6.5), 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):

$$\text{(6.6)} \qquad \begin{cases} H_t + \frac{1}{2}\rho^2(t, x)H_{xx} + \mu(t, x)H_x + h(t, x, H, \rho(t, x)H_x) = 0; \\ H(T, x) = g(x), \end{cases}$$

where $h = -\beta_0(t, x, y) - \beta_1(t, x, y)z$, and it 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 Example 5.2, in which the coefficients b, σ, μ 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 (6.6) now reads (suppressing variables):

$$\text{(6.7)} \begin{cases} H_t + \big((f_t + g_t)x + k_t + \rho(-\beta_t x + \beta_t H)\big)H_x + \frac{1}{2}\rho^2 H_{xx} = g_t x + k_t + f_t H; \\ H(T, x) = x, \end{cases}$$

We can easily check that $H(T,x) = x$ is the (unique) solution to (6.7), 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 $b \equiv 0$, then (6.6) is reduced to

$$\text{(6.8)} \qquad \begin{cases} H_t(t,x) + \mu(t,x)H_x(t,x) + \frac{1}{2}\rho^2(t,x)H_{xx}(t,x) = 0; \\ H(T,x) = g(x). \end{cases}$$

We should note that the PDE (6.8) is well-posed with properly chosen (μ,ρ), as part of the pricing rule. The determination of (μ,ρ), 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 α^* for the insider, which maximizes her expected terminal wealth W_T, given the pricing rule $P_t = \mathbb{E}[V_t|\mathcal{F}_t]$, $t \in [0,T]$.

In light of the analysis in the previous sections, we recast the problem of finding the Kyle-Back equilibrium as follows. First recall the Markovized system (2.5):

$$\text{(7.1)} \qquad \begin{cases} dV_t = b(t,V_t,X_t)dt + \sigma(t,V_t,X_t)dB^1_t, & V_0 = v; \\ dX_t = [\mu(t,X_t) + \rho(t,X_t)\alpha_t]dt + \rho(t,X_t)dB^2_t, & X_0 = x. \end{cases}$$

where $\alpha \in \mathscr{U}_{ad}$ (see (2.3) for definition). Assume that the process α 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 (2.7). Moreover, a necessary condition for $\alpha \in \mathscr{U}_{ad}$ being an equilibrium is that $V_T = P_T = H(T,X_T) = g(X_T)$ (see (1.4)). Therefore, We shall consider only the (weak) solution $(\bar{\mathbb{P}},V,X,\alpha)$ to STPBVP (4.1), and by Proposition 4.4, we shall assume that $\bar{\mathbb{P}}|_{\mathcal{F}_t} \ll \mathbb{Q}^0|_{\mathcal{F}_t}$, $t < T$, and $\bar{\mathbb{P}} \circ (V_T)^{-1} = m^*$.

It is worth noting that the solution to STPBVT (4.1) or SDE (7.1), depends on the coefficients (μ,ρ). We shall argue that the equilibrium can be determined by properly choosing (μ,ρ) through some "compatibility conditions".

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

$$P_s = \mathbb{E}^{\mathbb{P}}[V_s|\mathcal{F}^Y_s] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[V_t|\mathcal{F}_s]|\mathcal{F}^Y_s] = \mathbb{E}^{\mathbb{P}}[V_t|\mathscr{F}^Y_s] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[V_t|\mathcal{F}^Y_t]|\mathcal{F}^Y_s] = \mathbb{E}^{\mathbb{P}}[P_t|\mathcal{F}^Y_s].$$

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

$$\text{(7.2)} \qquad \begin{cases} H_t + \mu(t,x)H_x + \frac{1}{2}\rho^2(t,x)H_{xx} = 0; \qquad t \in [0,T) \\ H(T,x) = g(x). \end{cases}$$

Comparing to (6.6) and recalling (6.5) we see that under affine structure we have

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

Therefore we have $u_0(t,X_t) = -\beta_t H(t,X_t)$, where $\beta_t = u_1(t,X_t)$. Consequently, we see that $\alpha_t = u_0(t,X_t)+u_1(t,X_t)V_t = u_1(t,X_t)(V_t - H(t,X_t)) = \beta_t(V_t - P_t)$, which is exactly the form commonly seen in the literature (see, e.g., [1, 24, 27]), except that β_t is no longer deterministic. Our first main result of this section is the following.

THEOREM 7.1. *Assume Assumption 3.1, and $b \equiv 0$. Let $(\bar{\mathbb{P}}, \bar V, \bar X, \bar\alpha)$ be a weak solution to the STPBVP (4.1) such that $\bar\alpha_t$ has an affine structure. Then,*

(i) the market price $P_t = \mathbb{E}^{\bar{\mathbb{P}}}[\bar V_t|\mathcal{F}^Y_t] = H(t,\bar X_t)$, $t\in[0,T)$ is an $\mathbb{F}^Y$-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 $\bar{\mathbb{P}}$, such that $\bar V_T = g(\bar X_T)$, $\bar{\mathbb{P}}$-a.s.;

(iii) $\bar\alpha$ is an equilibrium strategy if the following "compatibility condition" holds:

$$\rho_t(t,x) - \mu_x(t,x)\rho(t,x) + \rho_x(t,x)\mu(t,x) + \frac{1}{2}\rho^2(t,x)\rho_{xx}(t,x) = 0. \tag{7.3}$$

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 [34], and look for a function $J(t,x;a)$ such that for fixed $a\in\mathbb{R}$, $J(\cdot,\cdot;a)\in\mathbb{C}^{1,2}([0,T]\times\mathbb{R})$, and satisfies the following properties

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

Assume now that a function J satisfying (7.4) exists. Then for any $\alpha\in\mathscr{U}_{ad}$, we let $(\mathbb{P}, V^\alpha, X^\alpha)$ be a weak solution to the SDE (4.1). Given $a\in\mathbb{R}$, applying Itô's formula to $J(\cdot,\cdot;a)$ we have

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

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

$$\begin{aligned} &\mathbb{E}^{\mathbb{P}}\big[J(T,X_T;V_T) - J(0,x_0;V_T)\big] = \int_{\mathbb{R}} \mathbb{E}^{\mathbb{P}}\big[J(T,X_T;a) - J(0,x_0;a)|V_T = a\big]\mathbb{P}_{V_T}(da)\\ &= \int_{\mathbb{R}} \mathbb{E}^{\mathbb{P}}\Big[\int_0^T (H(s,X_s)-a)\alpha_t dt + \int_0^T H(t,X_t)dB^2_t - aB^2_T | a = V_T\Big]\mathbb{P}_{V_T}(da)\\ &= \mathbb{E}^{\mathbb{P}}\Big[\int_0^T (H(s,X_s)-V_T)\alpha_t dt\Big] + \mathbb{E}^{\mathbb{P}}\Big[\int_0^T H(t,X_t)dB^2_t\Big] - \mathbb{E}^{\mathbb{P}}[V_T B^2_T]\\ &= \mathbb{E}^{\mathbb{P}}\Big[\int_0^T (H(s,X_s)-V_T)\alpha_t dt\Big] - \mathbb{E}^{\mathbb{P}}[V_T B^2_T]. \end{aligned}$$

But, since $\langle B^1, B^2 \rangle \equiv 0$, we have $d(V_t B_t^2) = V_t dB_t^2 + B_t^2 \sigma(t, V_t, X_t) dB_t^1$, $t \geq 0$. That is, $\{V_t B_t^2\}$ is a $\mathbb{P}$-martingale, hence $\mathbb{E}^{\mathbb{P}}[V_T B_T^2] = 0$. Recalling (2.6) we deduce from equations above that

$$\text{(7.6)} \quad \mathbb{E}^{\mathbb{P}}[W_T^\alpha] = \mathbb{E}^{\mathbb{P}}\Big[\int_0^T (V_T^\alpha - H(s, X_s^\alpha))\alpha_t dt\Big] = \mathbb{E}^{\mathbb{P}}\big[J(0, x_0; V_T^\alpha) - J(T, X_T^\alpha; V_T^\alpha)\big]$$
$$\leq \mathbb{E}^{\mathbb{P}}\big[J(0, x_0; V_T^\alpha)\big].$$

Here the last inequality is due to property (7.4) of the function J, and furthermore, the equality holds if and only if the terminal condition $V_T^\alpha = g(X_T^\alpha)$ holds. Consequently, if we let $(\bar{\mathbb{P}}, \bar{V}, \bar{X}, \bar{\alpha})$ be a weak solution to STPBVP (4.1), then Proposition 4.4, together with (7.6), shows that

$$\mathbb{E}^{\bar{\mathbb{P}}}[W_T^{\bar{\alpha}}] = \sup_{\alpha \in \mathscr{U}_{ad}, \mathbb{P}\circ(V_T^\alpha)^{-1} = m*} \mathbb{E}^{\mathbb{P}}[W_T^\alpha] = \int_{\mathbb{R}} J(0, x_0; a) m^*(da).$$

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

Our last task is to construct a function J that satisfies all the requirements in (7.4). In light of [34], we consider the following function:

$$\text{(7.7)} \qquad J(t, x; a) = \int_{g^{-1}(a)}^x \frac{H(t, y) - a}{\rho(t, y)} dy + \int_t^T f(s; a) ds,$$

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 μ, ρ, and f, we simply plugging the function J into the PDE in (7.4) to get

$$f(t; a) = \Big[\Big(\frac{\mu}{\rho} - \frac{\rho_x}{2}\Big)(H - a)\Big] + \frac{(H_x \rho)(t, x)}{2} + \int_{g^{-1}(a)}^x \Big[\frac{H_t}{\rho} - \frac{(H - a)\rho_t}{\rho^2}\Big](t, y) dy.$$

In order that $f(\cdot; a)$ is independent of x, we take derivative of the right hand side with respect to x, and multiply it by $\rho^2(t, x)$ 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 ρ is positive, we see that $f_x \equiv 0$ provided (7.3) holds. We note that if the function f in (7.7) 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, and $\rho(T, y) > 0$, we have $g(y) \geq g(g^{-1}(a)) = a$, for $y \geq g^{-1}(a)$. Thus $J(T, x; a) \geq 0$, for $x \geq g^{-1}(a)$, and $J(T, x; a) = 0$ iff $x = g^{-1}(a)$, proving (7.4). □

Remark 7.2. The compatibility condition (7.3) between the coefficients μ, ρ, 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, [34] 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 ρ and H satisfy $\rho_t + \frac{\rho^2}{2}\rho_{xx} = 0$, $H_t + \frac{\rho^2}{2}H_{xx} = 0$, and $f(t; a) = H_x(t, g^{-1}(a))\rho(t, g^{-1}(a))$, justifying (7.2) and (7.3). ∎

The General Case. We now try to apply the same scheme to the general case without assuming that $b(t, v, x) = 0$. We first observe that in this case the market price $P_t = \mathbb{E}[V_t|\mathcal{F}_t^Y]$, $t \geq 0$, is an "optional projection", which is not necessarily an $\mathbb{F}^Y$-martingale. Thus the discussion is more involved, and the final outcome is less explicit. We hope to be able find some more effective approaches in future research.

Let us assume now that both b and α have the general affine structure: $b(t, v, x) = b_0(t, x) + b_1(t, x)v$ and $u(t, v, x) = u_0(t, x) + u_1(t, x)v$. By Proposition 6.3, the decoupling field $H(t, x)$ would satisfy a more general PDE:

$$H_t + \mu H_x + \frac{1}{2}\rho^2 H_{xx} = (b_0 + b_1 H) - (u_0 + u_1 H)\rho H_x, \;\; H(T, x) = g(x). \tag{7.8}$$

So if we still try to construct the function $J(t, x; a)$ as in (7.7), then it may not be possible to find a corresponding function f that is independent of x. We propose to modify (7.7) in the following way. First recall that when α is Markovian, we can write

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

where $F(t, V_t, X_t) := \mathbb{E}^{\mathbb{P}}[V_T|\mathcal{F}_t^{V,X}]$, thanks to the Markovian property of the solution (V^α, X^α). Further, by Feynman-Kac formula, we see that F satisfies the PDE:

$$F_t + \frac{1}{2}F_{vv}\sigma^2 + \frac{1}{2}F_{xx}\rho^2 + F_v b + F_x(\mu + u\rho) = 0; \qquad F(T, v, x) = v. \tag{7.9}$$

In light of (7.7), we now look for the function $J(t, v, x)$ with the following properties:

$$\begin{cases} J_x\rho(t, x) = H(t, x) - F(t, v, x); \\ J_t + b(t, v, x)J_v + \mu(t, x)J_x + \frac{1}{2}\sigma^2(t, v, x)J_{vv} + \frac{1}{2}\rho^2(t, x)J_{xx} = 0; \\ J(T, v, x) \geq 0, \quad \text{and} \quad J(T, v, x) = 0 \text{ iff } v = g(x). \end{cases} \tag{7.10}$$

If such function J exists, then a simple application of Itô's formula shows that

$$\mathbb{E}^{\mathbb{P}}[W_T^\alpha] = \mathbb{E}^{\mathbb{P}}[-J(T, V_T, X_T) + J(0, v_0, x_0)] \leq J(0, v_0, x_0),$$

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 (7.7) as follows. Define

$$J(t, v, x) = \int_{g^{-1}(v)}^x \frac{H(t, y) - F(t, v, y)}{\rho(t, y)}dy + G(t, v) := \bar{J}(t, v, x) + G(t, v), \tag{7.11}$$

where $G(t, v)$ is a function to be determined. We first note that the first identity in (7.10) is trivial by definition of the function J (7.11). Next, we observe that $\bar{J}(T, v, x) = \int_{g^{-1}(v)}^x \frac{g(y) - v}{\rho(t, y)}dy$, which satisfies that $\bar{J}(T, v, x) \geq 0$, 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 (7.11) satisfies the terminal condition in (7.10) provided $G(T, v) \equiv 0$. Let us now look at the PDE in (7.10). Plugging (7.11) into (7.10), we have

$$0 = G_t + bG_v + \frac{1}{2}\sigma^2 G_{vv} + \bar{J}_t + b\bar{J}_v + \frac{1}{2}\sigma^2 \bar{J}_{vv} + \frac{\mu(H-F)}{\rho} + \frac{1}{2}[(H_x - F_x)\rho - (H-F)\rho_x]. \tag{7.12}$$

We see that if we can find the function $G(t, v)$ satisfying the PDE (7.12) with the terminal condition $G(T, v) \equiv 0$, then we will be able to define J as in (7.11). Summarizing the discussions above, we have the following result.

THEOREM 7.3. *Assume Assumption 3.1, Then, a weak solution* $(\bar{\mathbb{P}}, \bar{V}, \bar{X}, \bar{\alpha})$ *to STPBVP (4.1) with* α *having the affine structure is an equilibrium strategy if there exists a function* $G(t, v)$ *satisfying* (7.12) *with* $G(T, v) = 0$. ∎

We remark that by looking at (7.11), it seems that the function J depends on the choice of the strategy α since both PDEs (7.8) and (7.9) (for H and F) do. However, we should also note that the PDE in (7.10) for J, as well as its terminal condition are independent of α. Therefore the function J should depend solely on the choice of coefficients but independent of α. We should also note that Theorem 7.3 is only a sufficient condition for identifying the equilibrium, which is by no means necessary. That is, there could be different ways to find equilibrium, and Theorem 7.3 is only associated to the specific scheme that follows the idea of constructing the function J with the form (7.11). We conclude this section by using Theorem 7.3 to analyze two special cases in which the underlying asset process V is not a martingale.

Example 7.4. Consider the linear model in [27] again. That is, we let $b(t, v, x) = f_t v + g_t x + h_t$, $\sigma(t, v, x) = \sigma_t$, $H(t, x) = x$, and $\alpha(t, v, x) = \beta_t(v - x)$, where f, g, h, σ, β are deterministic functions. Then by [27, Theorem 3.6], we have $\mu(t, x) = (f_t + g_t)x + h_t$ and $\rho(t, x) = S_t\beta_t$, where S_t solves a Riccati equation. In this case, we can check that

$$\bar{J}(t, v, x) = \int_v^x \frac{y - F(t, v, y)}{S_t\beta_t} dy = \frac{1}{S_t\beta_t}\Big[\frac{x^2}{2} - \frac{v^2}{2} - \int_v^x F(t, v, y)dy\Big],$$

and a direct computation shows that (7.12) is now reduced to

$$0 = G_t + (f_t v + h_t)G_v + \frac{1}{2}\sigma_t^2 G_{vv} + \Theta_1(t, v, x), \tag{7.13}$$

where $\Theta_1 := \bar{J}_t + g_t x G_v + (f_t v + g_t x + h_t)\bar{J}_v + \frac{\sigma_t^2 \bar{J}_{vv}}{2} + \frac{[(f_t+g_t)x+h_t](x-F)}{S_t\beta_t} + \frac{[(1-F_x)S_t\beta_t]}{2}$. Since G is independent of x, we deduce from (7.13) that $\partial_x\Theta_1(t, v, x) = 0$, that is

$$0 = G_v g_t + \Theta_2(t, v, x), \tag{7.14}$$

where $\Theta_2(t, v, x) := \bar{J}_{tx} + (f_t v + g_t x + h_t)\bar{J}_{vx} + g_t\bar{J}_v + \frac{1}{2}\sigma_t^2\bar{J}_{vvx} + \frac{1}{S_t\beta_t}[(f_t + g_t)(2x - F) + h_t] - \frac{1}{2}F_{xx}S_t\beta_t$. Similarly, we can conclude that $\partial_x\Theta_2 = 0$, which leads to that

$$(F_x - 1)(S_t\beta_t)_t = S_t\beta_t\big[F_{tx} + (f_t v + g_t x + h_t)F_{vx} + 2g_t F_v + \frac{1}{2}\sigma_t^2 F_{vvx} + (f_t + g_t)(F_x - 2) + \frac{1}{2}(S_t\beta_t)^2 F_{xxx}\big] \tag{7.15}$$

Recall that $F(t,v,x)$ satisfies the PDE (7.9), we deduce from (7.15) that

$$\frac{(F_x-1)(S_t\beta_t)_t}{S_t\beta_t} = g_tF_v - F_{xx}[(f_t+g_t)x + h_t + S_t\beta_t^2(v-x)] + F_xS_t\beta_t^2 - 2(f_t+g_t). \tag{7.16}$$

Therefore, the compatibility conditions become (7.13), (7.14), (7.16), and $G(T,v)=0$.

It might be interesting to note that (7.16) can be further simplified in the case $g=0$. Indeed, by [27, Theorem 6.1], we see that in this case $S_t\beta_t = \frac{1}{2}\alpha_0\exp\{\int_0^t f_u du\}$, where α_0 is a constant. Then, it is easy to check that (7.16) can be simplifies as

$$-f_t = ([f_tx + h_t + S_t\beta_t^2(v-x)]F_x)_x,$$

which immediately gives $F_x = \frac{-f_tx + C_0(t,v)}{(f_t - S_t\beta_t^2)x + S_t\beta_t^2v + h_t}$, for some function $C_0(t,v)$ to be determined later. It is not hard to check that $F(t,v,x)$ can be written explicitly as

$$F = \frac{-f_tx}{A_t} + \Psi(t,v)\log\Phi(t,v,x) + C_1(t,v), \tag{7.17}$$

where $A_t = f_t - S_t\beta_t^2$, $\Phi = x + \frac{S_t\beta_t^2}{A_t}v + \frac{h_t}{A_t}$, $\Psi = \frac{C_0(t,v)}{A_t} + \frac{f_t(S_t\beta_t^2v+h_t)}{A_t^2}$, and $C_1(t,v)$ is another function to be determined. After calculating F_t, F_v, F_{vv}, F_x and F_{xx} accordingly and plugging them into (7.9), we obtain

$$0 = xF_1(t,v) + \log\Phi F_2(t,v) + F_3(t,v) + \Phi^{-1}F_4(t,v) + \Phi^{-2}F_5(t,v), \tag{7.18}$$

where

$$\begin{aligned}
F_1 =& -\partial_t\Big(\frac{f_t}{A_t}\Big) - f_t;\\
F_2 =& \partial_t\Psi(t,v) + (f_tv+h_t)\partial_v\Psi(t,v) + \frac{\sigma_t^2}{2}\frac{\partial_{vv}C_0}{A_t};\\
F_3 =& (\partial_t + (f_tv+h_t)\partial_v + \frac{1}{2}\sigma_t^2\partial_{vv})C_1 + C_0;\\
F_4 =& \Psi(t,v)\Big(\partial_t\Phi(t,v,x) + (f_tv+h_t)\frac{S_t\beta_t^2}{A_t}\Big) + \frac{\sigma_t^2S_t\beta_t^2}{A_t}\partial_v\Psi(t,v);\\
F_5 =& \Big[-\frac{\sigma_t^2}{2}\Big(\frac{S_t\beta_t}{A_t}\Big)^2 - \frac{(S_t\beta_t^2)^2}{2}\Big]\Psi(t,v).
\end{aligned}$$

Multiplying Φ^2, and denoting $\Lambda(t,v) = \frac{S_t\beta_t^2}{A_t}v + \frac{h_t}{A_t}$, we see that $\Phi = x+\Lambda$, and (7.18) now reads

$$\begin{aligned}
0 =& F_1x^3 + F_2(x+\Lambda)^2\log\Phi + (2\Lambda F_1 + F_3)x^2 + \big(\Lambda^2F_1 + 2\Lambda F_3 + F_4\big)x\\
&+(\Lambda^2F_3 + \Lambda F_4 + F_5).
\end{aligned}$$

Therefore, to show F defined in (7.17) satisfies the PDE (7.9), it is sufficient to show the following equations hold:

$$F_1 = F_2 = 2\Lambda F_1 + F_3 = \Lambda^2F_1 + 2\Lambda F_3 + F_4 = \Lambda^2F_3 + \Lambda F_4 + F_5 = 0.$$

which immediately implies $F_1 = F_2 = F_3 = F_4 = F_5 = 0$. We observe that $F_1 = 0$ is an ODE which determines f_t. Next, setting $C_0 := \frac{-f_t(S_t\beta_t^2v+h_t)}{A_t}$ we have $\Psi(t,v)\equiv 0$, and hence $F_4 = F_5 = 0$. Further, since $\partial_{vv}C_0 = 0$, this implies $F_2 = 0$. Finally, given C_0, we can solve an ODE for C_1 so that $F_3 = 0$. Therefore, with such f_t, C_0, and C_1 the function F defined in (7.17) satisfies (7.9) and (7.16) for arbitrary h_t and σ_t. ■

Example 7.5. We now extend the previous example by adding a slight nonlinearity into the system, but assuming that b and σ do not depend on x. More precisely, we let $b(t,v,x) = f_t v + h_t$, but $\sigma(t,v,x) = \sigma(t,v)$. We note that although in this case $g \equiv 0$, the solution is no longer Gaussian, and the decoupling field H is not explicitly known. To find the desired function $G(t,v)$ in Theorem 7.3, we differentiate both sides of (7.12) with respect to x and multiply by ρ^2 to get (suppressing variables):

$$\begin{aligned} 0 &= (H-F)[-\rho_t + \mu_x\rho - \mu\rho_x - \frac{1}{2}\rho^2\rho_{xx}] + \rho(H_t + \mu H_x + \frac{1}{2}\rho^2 H_{xx}) \\ &\quad -\rho\{F_t + bF_v + \mu F_x + \frac{1}{2}\sigma^2 F_{vv} + \frac{1}{2}\rho^2 F_{xx}\}. \end{aligned} \tag{7.19}$$

Note that H and F satisfy PDEs (7.8) and (7.9), respectively, we deduce that

$$\begin{cases} h_t + f_t H + \rho[(u_0+u_1v)F_x - H_x(u_0+u_1H)] = 0. \\ \rho_t - \mu_x\rho + \rho_x\mu + \dfrac{1}{2}\rho^2\rho_{xx} = 0. \end{cases}$$

We now observe that the function $\phi(t,x) := (u_0+u_1v)F_x$ is independent of v. Thus for $v \neq -u_0/u_1$, we can write $F_x = \frac{\phi(t,x)}{u_0+u_1v}$ and compute $F_{xx}, F_{xt}, F_{xv}, F_{xvv}$ and F_{xxx} accordingly. Differentiating (7.9) with respect to x, plugging the corresponding partial derivatives above, and denoting

$$\begin{aligned} A =& \phi_t + \rho\rho_x\phi_x + \frac{1}{2}\rho^2\phi_{xx} + \phi_x(h+fv+\mu) + \phi\mu_x; \\ B =& \phi[(u_0)_t + (u_1)_t v] + \rho\rho_x\phi[(u_0)_x + (u_1)_x v] + \rho^2\phi_x[(u_0)_x + (u_1)_x v] \\ &+ \frac{1}{2}\rho^2\phi[(u_0)_{xx} + (u_1)_{xx}v] + \phi(h+fv+\mu)[(u_0)_x + (u_1)_x v]; \\ C =& \sigma^2\phi u_1^2 + \rho^2\phi[(u_0)_x + (u_1)_x v]^2, \end{aligned}$$

we obtain the following equation:

$$0 = (\phi\rho)_x + \frac{A}{u_0+u_1v} - \frac{B}{(u_0+u_1v)^2} + \frac{C}{(u_0+u_1v)^3}. \tag{7.20}$$

Now fix (t,x) and let $v \to \infty$, by definitions of A and B, we can easily check that

$$(\phi\rho)_x + \frac{\phi_x f}{u_1} - \frac{\phi f(u_1)_x}{u_1^2} = (\phi\rho)_x + \big(\frac{\phi f}{u_1}\big)_x = 0.$$

This implies $\phi(t,x) = c(t)\big[\rho(t,x) + \frac{f_t}{u_1(t,x)}\big]^{-1}$, for some function $c(t)$. Moreover, setting $v = -\frac{u_0}{u_1} + \varepsilon$, multiplying (7.20) by ε^3, and sending ε to 0 will yield: $\sigma^2\phi u_1^2 + \rho^2\phi\{(u_0)_x - (u_1)_x\frac{u_0}{u_1}\}^2 \equiv 0$, which implies $\phi \equiv 0$, and hence $F_x = F_{xx} \equiv 0$. Consequently, we can rewrite the compatibility conditions from (7.19):

$$\begin{cases} F_t + bF_v + 1/2\sigma^2 F_{vv} = 0; \\ H_t + \mu H_x + 1/2\rho^2 H_{xx} = 0; \\ \rho_t - \mu_x\rho + \rho_x\mu + 1/2\rho^2\rho_{xx} = 0. \end{cases}$$

We note that in the above the first equation is (7.9), the second and the third condition coincide with the ones in Theorem 7.1. Furthermore, the second equation implies $\{P_t\}$ is a martingale, but since $b = fv + h \neq 0$, $\{V_t\}$ is not a martingale. ■

Acknowledgement. We would like to express our sincere gratitude to the anonymous referee for the very careful reading of the original manuscripts and many helpful comments and suggestions. In fact, part of the computation of Example 7.5 is due to referee's report.

REFERENCES

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

[2] D. G. Aronson, *Bounds for the fundamental solution of a parabolic equation*, Bulletin of the American Mathematical society, 73 (1967), pp. 890–896.

[3] D. G. Aronson, *Non-negative solutions of linear parabolic equations*, Annali della Scuola Normale Superiore di Pisa-Classe di Scienze, 22 (1968), pp. 607–694.

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

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

[6] B. Barrios, I. Peral, F. Soria, and E. Valdinoci, *A widder's type theorem for the heat equation with nonlocal diffusion*, arXiv preprint arXiv:1302.1786, (2013).

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

[8] A. Bensoussan, *Stochastic control of partially observable systems*, Cambridge University Press, 1992.

[9] F. Biagini, Y. Hu, T. Meyer-Brandis, and B. Oksendal, *Insider trading equilibrium in a market with memory*, Mathematics and Financial Economics, 6 (2012), pp. 229–247.

[10] R. Buckdahn, J. Li, and J. Ma, *A mean-field stochastic control problem with partial observations*, The Annals of Applied Probability, 27 (2017), pp. 3201–3245.

[11] R. Caldentey and E. Stacchetti, *Insider trading with a random deadline*, Econometrica, 78 (2010), pp. 245–283.

[12] L. Campi, U. Cetin, and A. Danilova, *Dynamic markov bridges motivated by models of insider trading*, Stochastic Processes and their Applications, 121 (2011), pp. 534–567.

[13] L. Campi, U. Cetin, and A. Danilova, *Explicit construction of a dynamic bessel bridge of dimension* 3, Electronic Journal of Probability, 18 (2013), pp. 1–25.

[14] U. Cetin and A. Danilova, *On pricing rules and optimal strategies in general kyle–back models*, SIAM Journal on Control and Optimization, 59 (2021), pp. 3973–3998.

[15] F. Cocquemas, I. Ekren, and A. Lioui, *A general solution method for insider problems*, Available at SSRN 3628726, (2020).

[16] A. Danilova, *Stock market insider trading in continuous time with imperfect dynamic information*, Stochastics An International Journal of Probability and Stochastics Processes, 82 (2010), pp. 111–131.

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

[18] P. Fitzsimmons, J. Pitman, and M. Yor, *Markovian bridges: construction, palm interpretation, and splicing*, Seminar on Stochastic Processes, 1992, (1993), pp. 101–134.

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

[20] H. Föllmer and P. Imkeller, *Anticipation cancelled by a girsanov transformation: a paradox on wiener space*, Annales de l'IHP Probabilités et statistiques, 29 (1993), pp. 569–586.

[21] H. Föllmer, C.-T. Wu, and M. Yor, *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.

[22] M. Fujisaki, G. Kallianpur, and H. Kunita, *Stochastic differential equations for the non linear filtering problem*, Osaka Journal of Mathematics, 9 (1972), pp. 19–40.

[23] C. W. Holden and A. Subrahmanyam, *Long-lived private information and imperfect competition*, The Journal of Finance, 47 (1992), pp. 247–270.

[24] A. S. Kyle, *Continuous auctions and insider trading*, Econometrica: Journal of the Econometric Society, (1985), pp. 1315–1335.

[25] Y. Liu and J. Ma, *Bsde with jumps and multiplicative coefficients*, 2005.

[26] J. Ma, J.-M. Morel, and J. Yong, *Forward-backward stochastic differential equations and their applications*, no. 1702, Springer Science & Business Media, 1999.

[27] J. Ma, R. Sun, and Y. Zhou, *Kyle–back equilibrium models and linear conditional mean-field*

sdes, SIAM Journal on Control and Optimization, 56 (2018), pp. 1154–1180.

[28] J. Ma, Z. Wu, D. Zhang, and J. Zhang, *On well-posedness of forward–backward sdes—a unified approach*, The Annals of Applied Probability, 25 (2015), pp. 2168–2214.

[29] P. Propter, *Stothastic Integration and Differential Equations: A New Approach*, Springer, 1990.

[30] G. R.B, *An extension of widder's theorem*, Master Thesis, Oregon State University.

[31] Y. Tan, *Stochastic two-point boundary value problem and applications*, PhD Thesis, University of Southern California, (2023).

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

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

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