KYLE'S MODEL WITH STOCHASTIC LIQUIDITY

IBRAHIM EKREN, BRAD MOSTOWSKI, AND GORDAN ŽITKOVIĆ

Abstract. We construct an equilibrium for the continuous-time Kyle model with stochastic liquidity, a general distribution of the fundamental price, and correlated stock and volatility dynamics. For distributions with positive support, our equilibrium allows us to study the impact of the stochastic volatility of noise trading on the volatility of the asset. In particular, when the fundamental price is log-normally distributed, informed trading forces the log-return up to maturity to be Gaussian for any choice of noise-trading volatility even though the price process itself comes with stochastic volatility. Surprisingly, we find that in equilibrium both Kyle's Lambda and its inverse (the market depth) are submartingales.

1. Introduction

Kyle's model, introduced in Kyle (1985), is one of the most influential models in the market microstructure literature. The equilibrium constructed in Kyle (1985) shows how the information about an asset is incorporated in its price and how the liquidity in the market and the volatility of the asset price are impacted by noise trading. In its original formulation, at the initial time, an informed trader (insider) learns the fundamental value $\tilde{v}$ of an asset, where $\tilde{v}$ is assumed to have a normal prior distribution. She then trades against a market maker in order to optimize her expected profit. The objective of the market maker, on the other hand, is to filter the fundamental price $\tilde{v}$ by observing the totality of the demand from the informed trader and one or more noise traders. To achieve that, the agents choose a mechanism which continuously transforms the observed demand into a price quote. Such a mechanism, when it satisfies an additional assumption related to market efficiency, is called an equilibrium if neither the insider nor the market maker have an incentive to deviate from their pre-announced strategies.

The equilibrium constructed in Kyle (1985) is linear and the informed trader's trading rate is proportional to the current price mismatch (the difference between the quoted and the fundamental price) of the asset, and inversely proportional to the time to maturity. In equilibrium the increment dP_t of the price is given by

$$dP_t = \lambda \, dY_t,$$

where dY_t is the increment of the total demand received by the market maker (from the informed trader and the noise traders) and the constant λ is the so-called Kyle's Lambda which is the sensitivity of the price to the total demand. This constant is proportional to the standard deviation of the fundamental price and inversely proportional to the standard deviation of noise trading. Thus, Kyle's model is a mathematical expression of the idea that the market liquidity is inversely proportional to the average flow of new information and proportional to the volume of liquidity-motivated transactions (see Bagehot (1971) written by Jack Treynor under the pseudonym Walter Bagehot).

1.1. Literature review.

An impressive number of extensions of Kyle's model have been considered in the literature. In discrete time, Subrahmanyam (1991) allows for risk aversion of the informed trader, while Caballe and Krishnan (1994); Garcia del Molino et al. (2020) work with multiple assets. In continuous time, Back (1992) removes the normality assumption of the fundamental price and proves the existence of an equilibrium using a PDE-based approach. Kyle's model with dynamic information is studied in Back and Pedersen (1998) and Campi et al. (2011). We also mention Back (1993); Back et al. (2018, 2020); Biagini et al. (2012); Bose and Ekren (2023, 2024); Çetin and Danilova

Date: September 18, 2025.

During the preparation of this work the first named author was supported by the National Science Foundation under Grant No. DMS-2007826 the third named author by the National Science Foundation under Grants No. DMS-1815017 and No. DMS-230779. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

Keywords: Kyle model, asymmetric information, liquidity, price impact, market depth, stochastic volatility

AMS classification (2020): Primary 60H30, 60J60; secondary 91B44 .

(2016); Çetin (2018); Cho (2003); Choi et al. (2019); Corcuera and Di Nunno (2020); Corcuera et al. (2019); Aase et al. (2012); Barger and Donnelly (2021); Lasserre (2004) and Ying (2020), among many others.

Recently, Collin-Dufresne and Fos (2016) and Collin-Dufresne et al. (2021) proposed an extension of Kyle's model which allows for stochastic volatility in the definition of the noise traders' cumulative demand process Z_t. In the seminal work of Kyle (1985), the instantaneous demand dZ_t of the noise trades has a deterministic variance $\sigma^2 dt$, where σ is either a constant or a deterministic function of time. In Collin-Dufresne and Fos (2016) and Collin-Dufresne et al. (2021), however, this variance is given by $\sigma_t^2 dt$ for some stochastic process σ_t. Thus, the total demand $Z_T = \int_0^T dZ_t$ is no longer necessarily Gaussian (as in the classical case) and it is not clear how the PDE-based approach of Back (1992), or the optimal-transport methodology of Back et al. (2020), needs to be modified in order to find an equilibrium.

Assuming that the fundamental price $\tilde{v}$ is normally distributed, an equilibrium is constructed in Collin-Dufresne and Fos (2016). Relying on normality, these authors conjecture that the trading rate and the expected wealth of the informed trader are a linear and a quadratic function, respectively, of the price mismatch. A crucial step in their existence proof of the equilibrium is the construction a martingale whose inverse λ_t has the property that the relation

$$dP_t = \lambda_t \, dY_t$$

implies that the conditional variance of $\tilde{v}$ decreases at the rate $\lambda_t^2 \sigma_t^2$ (as in equation (8) on p. 1447 of Collin-Dufresne and Fos (2016)). This is accomplished by introducing a decomposition of λ_t which reduces the problem to a Backward Stochastic Differential Equation (BSDE).

1.2. **Our contributions.** Even though it enhances tractability, the assumption that the fundamental price is Gaussian in Collin-Dufresne and Fos (2016); Collin-Dufresne et al. (2021) permits the equilibrium price to be negative with positive probability. One of the goals of the present paper is to relax this assumption and allow a general distribution for the fundamental price. The same relaxation in Back (1992) renders the pricing rule non-linear, so we cannot expect the linear-quadratic structure of Collin-Dufresne and Fos (2016) to apply to our framework either. In fact, in many aspects the relation between our work and Collin-Dufresne and Fos (2016) is similar to the relation between Back (1992) and Kyle (1985).

From the control-theoretic perspective, the main contribution of this paper is the identification of a relevant state process ξ_t that allows us to compute the price of the asset and the strategy of the informed trader. It turns out that in equilibrium this state process is given by $\xi_t = \int_0^t \lambda_s dY_s$, the pricing rule is a function of ξ_t, and the trading rate of the informed trader is proportional to the deviation of ξ_t from its final value. As such, (ξ_t) carries all the path-dependence of the components of the equilibrium in Y.

Our proof of existence of equilibrium relies on the crucial observation that, in equilibrium, and conditionally on the information available to the market maker at time t, the random variable $\xi_T - \xi_t = \int_t^T \lambda_s dY_s$ is centered Gaussian with variance $\Sigma_t = \int_t^T \lambda_s^2 \sigma_s^2 ds$ (a quantity which turns out to be known to the market maker at time t). That permits us to use the PDE-based construction of Back (1992) and Gaussian filtering to find an equilibrium pricing rule. The equilibrium price is, then, of the form $P_t = H^*(t, \xi_t)$, where ξ_t is the state process mentioned above and the pricing rule $\xi \mapsto H^*(t, \xi)$ is a random field adapted to the filtration of σ_t. It has the property that its section at maturity, i.e., the map $\xi \mapsto H^*(T, \xi) = h(\xi)$, pushes the Gaussian distribution of ξ_T to the distribution of $\tilde{v}$. Moreover, the random field $H^*(t, \xi)$ (and its path-dependence in σ_t) admits a further simplification as $H^*(t, \xi) = R_\xi(\Sigma_t, \xi)$ where R is a deterministic function which solves the heat equation, and Σ_t is the remaining uncertainty in the final value of ξ_T.

The trading rate of the informed trader is proportional to $h^{-1}(\tilde{v}) - \xi_t$ where the constant of proportionality is adapted to the filtration of the market maker and matches the parallel term in Collin-Dufresne and Fos (2016). From the point of view of the market maker, the process ξ_t is a martingale, while for the informed trader it is a bridge process of type constructed in Collin-Dufresne and Fos (2016). In fact, our process ξ_t has many features of the price process of Collin-Dufresne and Fos (2016) and reduces to the price process if $\tilde{v}$ is Gaussian.

In addition to the relaxation of the Gaussian property of $\tilde{v}$, our construction extends the results in Collin-Dufresne and Fos (2016) in several other ways. First of all, we prove that the strategy of the informed trader is not only

optimal among all absolutely continuous strategies but also among all strategies with jumps or diffusive components. The optimality of an absolutely continuous strategy is related to the positivity of the price impact as mentioned in Corcuera and Di Nunno (2020); Bose and Ekren (2023, 2024), which holds in our framework, too.

Just like Collin-Dufresne and Fos (2016) do in Section 4., we allow the cumulative demand process Z_t of the noise trades and its stochastic volatility σ_t to be driven by correlated Brownian motions. As mentioned in Collin-Dufresne and Fos (2016) in the Gaussian case, the fact that σ_t is observable by both agents implies that $d\xi_t$ is not driven by dY_t but by the innovation process for the filtering problem of the market maker. This innovation process is orthogonal to the increments of σ_t even if σ_t and Z_t are driven by correlated Brownian motions. Therefore, quite surprisingly, σ_t and ξ_t (and therefore P_t) are driven by independent Brownian motions in equilibrium.

Another interesting finding concerns Kyle's Lambda, i.e., the sensitivity of the price to the total demand Y_t (or its informative part $\hat{Y}$ when σ_t and Z_t are driven by correlated Brownian motions). In equilibrium it is given by $\frac{R_{\xi\xi}(\Sigma_t,\xi_t)}{\frac{1}{\lambda_t}}$ which is a ratio of two positive orthogonal martingales for the filtration of the market maker. Thus, Kyle's Lambda is a submartingale. Trivially, but also surprisingly, the market depth which is the inverse of Kyle's Lambda is also a submartingale as the ratio of two orthogonal positive martingales. With Gaussian distributions as in Collin-Dufresne and Fos (2016), the function $R_{\xi\xi}$ is constant and Kyle's Lambda is a submartingale, but the market depth - being equal to $\frac{1}{\lambda_t}$ up to a multiplicative constant - is a martingale.

As mentioned in Collin-Dufresne and Fos (2016), the market maker anticipates more informed trading when there is more noise trading. Thus, the rate of injection of the information into the asset price is stochastic. This effect imposes distributional constraints on the dynamics of the price process. For example, if $\tilde{v}$ has a log-normal distribution ξ_t can be identified as the log-return of the asset price up to a multiplicative constant. Thus, the fact that Σ_t is measurable with respect to the information of the market maker at time t means that independently of the stochasticity of future noise trading, the presence of an informed trader renders the log-return of the asset from t to T Gaussian. Note also that the distribution of $\tilde{v}$ does not impact λ_t so that the (conditional) Gaussianity of ξ_T still holds for general distributions. If the fundamental price is Gaussian, ξ_t is the price process up to a multiplicative constant and the adaptedness of Σ_t to the information of the market maker imposes a centered Gaussian conditional distribution onto the price increment $P_T - P_t$. For general distributions, ξ_t cannot be interpreted as either the return of the asset or its price, and the dependence between price process and ξ is nonlinear in general.

The conditional Gaussianity of ξ_T has implications in terms of completeness, as well. Although, in equilibrium, the price process has stochastic (realized) volatility, the conditional Gaussianity of ξ_T implies that every European claim is redundant in the sense that it can be hedged by investing in the asset. The hedging strategy depends on ξ_t and is path-dependent in σ_t. This result is independent of the choice of σ_t and the distribution of $\tilde{v}$, and stands in contrast with Back (1993); Back et al. (2021) where the informed trader can also trade in options on the stock and the options are generically non-redundant.

Since the physical probability is risk neutral, we can use the full set of call option prices to gain information about the distribution of $\tilde{v}$. Indeed, given a choice of dynamics for σ_t, we can use our model to predict the dynamics of the implied volatility curve for a given maturity as a function of ξ_t and Σ_t. With observed call option prices used as input, this leads to an inverse problem for the distribution of $\tilde{v}$. For example, if the distribution of $\tilde{v}$ is log-normal, i.e., with a flat IV (implied volatility) curve, then its IV curve remains flat on $[0, T]$. However, the level of this flat curve moves stochastically depending on the value σ_t, ξ_t, and Σ_t. For general distributions of $\tilde{v}$, the shape of the IV curve might change depending on time, ξ_t, σ_t, and Σ_t in a nonlinear way up to the computation of R_ξ by solving a heat equation and inverting the Black-Scholes formula as a function of the volatility. Qualitatively, we observe that the shape of the IV curve is mainly influenced by the distribution of $\tilde{v}$ whereas its level mainly depends on the dynamics of σ_t.

1.3. **Organization of the paper.** The rest of the paper is organized as follows. In Section 2, we first state the problem and define the concept of equilibrium. Then we introduce its most important building blocks and state our main existence result. Section 3 provides examples and Section 4 contains the proofs of the main theorem and other results.

2. Problem setup and the main result

2.1. **The probabilistic setup.** Given $T > 0$, let $(\Omega, \bar{\mathcal{F}}, \mathbb{P})$ be a probability space which supports a two-dimensional Brownian motion $(W, B) = (W_t, B_t)_{t\in[0,T]}$ on $[0, T]$, as well as an independent random variable $\tilde{v}$. The augmentation of the natural filtration of (W, B) by $\mathbb{P}$-trivial events in $\bar{\mathcal{F}}$ is denoted by $\mathcal{F}^{W,B} = (\mathcal{F}^{W,B}_t)_{t\in[0,T]}$; it satisfies the usual conditions of right continuity and completeness. Moreover, let the filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$, which also satisfies the usual conditions, be defined by $\mathcal{F}_t = \mathcal{F}^{W,B}_t \vee \sigma(\tilde{v})$ for $t \in [0, T]$. In general, for any process S, we denote by $\mathcal{F}^S$ the $\mathbb{P}$-augmentation of the filtration generated by S; in this context, we interpret $\tilde{v}$ as a constant process so that, e.g., $\mathcal{F}^{\tilde{v},W,B} = \mathcal{F}$.

For an $\mathcal{F}^W$-martingale M, we define its BMO (bounded mean oscillation) norm by

$$||M||^2_{BMO} = \sup_\tau ||\mathbb{E}[|M_T - M_\tau|^2|\mathcal{F}^W_\tau]||_\infty$$

where the supremum is taken over all $\mathcal{F}^W$-stopping times τ, and $||\cdot||_\infty$ denotes the essential supremum of a random variable (see Kazamaki (2006)). We call M a BMO-martingale if $||M||_{BMO} < \infty$. For any $\mathcal{F}^W$-adapted and W-integrable process α, we write $\alpha \in \mathrm{bmo}$ if $\int \alpha_s dW_s$ is a BMO-martingale. For a continuous process M and $\gamma \in (0, 1)$, we denote by $|M|_\gamma := \sup_{0\le s<t\le T} \frac{|M_s - M_t|}{|t-s|^\gamma}$, its pathwise γ-Holder semi-norm.

Let $\mathcal{S}^\infty$ denote the set of continuous, $\mathcal{F}^W$-adapted and uniformly bounded processes. The set $\mathcal{S}^+_0$ consists of all continuous $\mathcal{F}^W$-adapted processes G, strictly positive on $[0, T)$, with $G_T = 0$, and $\mathcal{P}^2$ denotes the set of all $\mathcal{F}^W$-progressively measurable processes z with $\int_0^T z_u^2\, du < \infty$, a.s.

2.2. **The model.** As in Collin-Dufresne and Fos (2016), we consider an interaction among an informed trader, a market maker, and noise traders during the time period $[0, T]$:

- At time $t = 0$, the informed trader (insider) learns the value of $\tilde{v}$, a random variable that represents the fundamental (or liquidation) value of an asset at maturity $t = T$. She trades in the market using a strategy X, where X denotes her total cumulative demand. We allow X to depend on $\tilde{v}$, as well as both W and B in an adapted way, i.e., the insider's filtration is $\mathcal{F}$.
- Noise traders place their trades at random without any regard to the actions of the other participants. Their trading intensity is not constant; instead, the cumulative order process (Z_t) of the noise traders is given by

$$Z_t = \int_0^t \sigma_s\Big(\hat{\rho}\, dB_s + \rho\, dW_s\Big), t \in [0, T], \text{ where } \rho \in (-1, 1), \hat{\rho} = \sqrt{1-\rho^2}, \tag{2.1}$$

  and σ is a nonnegative stochastic process adapted to $\mathcal{F}^W$.
- The market maker has no access to the value of $\tilde{v}$ or the demand X of the insider. On the other hand, he knows the distribution ν of $\tilde{v}$ and observes the total order flow $Y = X + Z$, as well as the value of W; in other words, his filtration is given by $\mathcal{F}^m = \mathcal{F}^{Y,W}$. He precommits to a pricing functional $\boldsymbol{P}$ which transforms the entire observed path of Y and W, as in Collin-Dufresne and Fos (2016), to a price process P.
- The equilibrium is achieved when the insider has no incentive to alter her trading strategy X, given the pricing functional $\boldsymbol{P}$, and the market maker's price $P = \boldsymbol{P}(X + Z)$ is rational (i.e., the $\mathbb{L}^2$-optimal estimate of $\tilde{v}$ based on his information) given the insider's strategy X.

The two Brownian motions, B and W, have distinct economic interpretations. On the one hand, B is idiosyncratic in that it drives a portion of the noise trades and is known to the informed trader, but is not observed directly by the market maker. On the other hand, W plays the role of "public noise" and will be observable by all market participants - the market maker in particular. This is the context in which the assumption that σ is $\mathcal{F}^W$-adapted should be understood. While this is a non-trivial assumption which will have important mathematical consequences in the sequel and seems to be hard to relax, it is automatically satisfied in many pertinent cases. Indeed, as we will learn later (see (2.17) below), the market maker will have direct access to σ in equilibrium (through the quadratic variation of the total demand process $\hat{Y}$). From there, he will be able to infer the value of W in a variety of diffusive models for σ.

We proceed by giving rigorous definitions for the concepts introduced informally above with the set of all $\mathcal{F}$-semimartingales denoted by $\mathcal{S}(\mathcal{F})$.

Definition 2.1. *A **pricing rule** (or a **pricing functional**) is a map $\boldsymbol{P} : \mathcal{S}(\mathcal{F}) \to \mathcal{S}(\mathcal{F})$, with the following properties:*

(1) $P(S)$ is an $\mathcal{F}^{W,S}$-semimartingale for each $S \in \mathcal{S}(\mathcal{F})$, and

(2) P is non-anticipative in the sense that for all $S, S' \in \mathcal{S}(\mathcal{F})$ and each $t \in [0,T]$ we have

$$S_s = S'_s \text{ for all } s \leq t, \text{ a.s. } \Rightarrow \boldsymbol{P}(S)_s = \boldsymbol{P}(S')_s \text{ for all } s \leq t, \text{ a.s..}$$

Remark 2.1. Our equilibrium price functional $\boldsymbol{P}^*$ will be built in two steps. First, a state process ξ^* will be constructed as the solution of a stochastic differential equation driven by B and whose data is $\mathcal{F}^W$-adapted and final condition is given by $\tilde{v}$ (see (2.12)). Then, we define the candidate equilibrium strategy X^* of the informed trader and the candidate equilibrium total demand Y^*. One can check that ξ^* can be obtained by applying a non-anticipative functional $\boldsymbol{\xi}^*$ of the paths of Y^* and W. Next, a random field H^* adapted to $\mathcal{F}^W$ will be applied to it: compared to Back (1992), we interpret ξ^* as a path-dependent generalization of the total demand process Y, while H^* adds W-dependence to Back's H. In this regard, ξ^* is a novel state variable allowing us to formulate the equilibrium as a one-dimensional Markov control problem (of ξ) from the perspective of the informed trader and the pricing rule as a functional of (ξ_t). We refer to Cho (2003); Campi et al. (2011); Bose and Ekren (2023, 2024) for the introduction of auxiliary state processes in Kyle's model.

Once the price functional $\boldsymbol{P}$ is given, the insider's goal is to maximize the expected gains from investing in the market. To rule out doubling strategies and other pathologies, we impose an admissibility constraint in the standard way. Consider an adapted RCLL process X, whose value X_{t-} is interpreted as the number of shares of the asset held by informed the agent at time t. We recall from Back (1992), that the total profit/loss from trading accumulated by the informed trader who uses the strategy X against the price process P is given by $(\tilde{v} - P_T)X_T + \int_0^T X_{t-}\, dP_t$. Under the assumption that X is, additionally, a semimartingale - so that the integration-by-parts formula applies - the same author casts this expression into an equivalent, but more convenient form $\int_0^T (\tilde{v} - P_s)\, dX_s - [X,P]_T$, which we use in the definition of admissibility below:

Definition 2.2. *Any $(\mathcal{F}_t)$ semimartingale X with $X_0 = 0$ is called a **trading strategy**. Given a semimartingale P, and a trading strategy X, the random variable*

$$\Pi(X,P)_T = (\tilde{v} - P_T)X_T + \int_0^T X_{t-}\, dP_t = \int_0^T (\tilde{v} - P_t)\, dX_t - [X,P]_T \tag{2.2}$$

*is called the **realized wealth** of the strategy X, with respect to P.*

*Given a semimartingale P, a trading strategy X is said to be P**-admissible** if the process $\Pi(X,P)_t = \int_0^t (\tilde{v} - P_s)\, dX_s - [X,P]_t$ is uniformly bounded from below by an integrable random variable.*

*Given a pricing functional $\boldsymbol{P}$, a trading strategy X is said to be $\boldsymbol{P}$**-admissible** if it is $\boldsymbol{P}(X+Z)$-admissible.*

Remark 2.2. Note that unlike Collin-Dufresne and Fos (2016), we allow the informed trader to use both diffusive and jump strategies. However, we prove in the sequel that these strategies are not profitable for the informed trader, and in equilibrium it is optimal for the informed trader to use an absolutely continuous strategy. As noted in Back et al. (2020); Corcuera and Di Nunno (2020), this point is inherited from the fact that the final pricing rule of the market maker is an increasing function of the underlying state variable.

Finally, we introduce the standard notion of equilibrium:

Definition 2.3. *A pair $(\boldsymbol{P}^*, X^*)$ consisting of a pricing rule $\boldsymbol{P}^*$ and a trading strategy X^*, is an **equilibrium** if*

(i) X^* is $\boldsymbol{P}^*$-admissible and

$$\mathbb{E}\Big[\Pi\Big(X; \boldsymbol{P}^*(X+Z)_T\Big)\Big] \leq \mathbb{E}\Big[\Pi\Big(X^*; \boldsymbol{P}^*(X^*+Z)_T\Big)\Big],$$

whenever X is an $\boldsymbol{P}^*$-admissible trading strategy.

(ii) $\boldsymbol{P}^*$ is rational i.e.,

$$\boldsymbol{P}^*(X^*+Z)_t = \mathbb{E}\Big[\tilde{v} \,\Big|\, \mathcal{F}_t^{W, X^*+Z}\Big], \text{ a.s., for all } t \in [0,T]. \tag{2.3}$$

Remark 2.3. Notationally, we distinguish between functionals (bold, like $\boldsymbol{P}$) and processes (light, like P). Similarly, starred quantities (like X^*) will refer to the (candidate) equilibrium, while their non-starred versions (like X) denote their generic analogues. The two conventions are often used together (as in $\boldsymbol{\xi}^*$).

2.3. **Regularity assumptions.** Before we state our main result, we discuss the regularity assumptions imposed on its inputs. Examples of processes which satisfy part (3) will be provided in Section 3 below.

Assumption 2.1.

(1) $|\rho| < 1$;

(2) $\tilde{v}$ *is square-integrable, independent of* (W, B), *and its distribution* ν *is absolutely continuous;*

(3) σ *admits a decomposition of the form*

$$\sigma = LJ, \tag{2.4}$$

where L *is a stochastic exponential of a BMO martingale adapted to* $\mathcal{F}^W$, *and* J *is an* $\mathcal{F}^W$*-adapted, bounded and bounded-away-from-*0 *continuous process with the property that*

$$\mathbb{E}[e^{r|J|_\gamma}] < \infty \text{ for all } r > 0. \tag{2.5}$$

for some Hölder exponent $\gamma > 0$.

Remark 2.4. It is readily seen that the requirements of Assumption 2.1, (3) are satisfied for a volatility process with the decomposition

$$d\sigma_t = \sigma_t(b_t\, dt + \psi_t\, dW_t),\ \sigma_0 > 0$$

where b and ψ are $\mathcal{F}^W$-adapted, b is bounded and $\psi \in \mathrm{bmo}$ (or, more restrictively, bounded as well). In general, this condition is more stringent than the assumptions imposed on σ in Collin-Dufresne and Fos (2016). This is due the fact that we aim to fix a few minor mistakes in Collin-Dufresne and Fos (2016). Indeed, Lemma 8., p. 1471 in Collin-Dufresne and Fos (2016) states the martingality of two processes for any admissible strategy of the informed trader. This claim is proven by using Lemma 4., p. 1468 in Collin-Dufresne and Fos (2016). Unfortunately, this Lemma only applies to the candidate optimal strategy of the informed trader and not to all admissible strategies. Additionally, due to $\Sigma_T = 0$, the martingality of the process in (Collin-Dufresne and Fos, 2016, Equation (66)) requires additional arguments. Fixing these minor mistakes for a reasonable class of admissible strategies turns out to be a somewhat challenging problem that requires the more stringent Assumption 2.1.

2.4. **Building blocks of the equilibrium.** The main goal of this paper is to show that under Assumption 2.1, an equilibrium exists, and to describe its structure. We outline its construction here, with proofs relegated to Section 4.

2.4.1. *The function* h. Let h be the unique nondecreasing function that pushes the standard normal distribution forward to the distribution ν of $\tilde{v}$. More precisely, let F_ν be the cdf (cumulative distribution function) of ν with F_ν^{-1} its generalized inverse, let Φ be the cdf of the standard normal, and let

$$h(x) = F_\nu^{-1}(\Phi(x)) \text{ for } x \in \mathbb{R}.$$

It is clear that h is nondecreasing and right-continuous, and unique (in the class of nondecreasing functions) with the property that $h_\sharp N(0,1) = \nu$, where $h_\sharp$ denotes the push-forward and $N(\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2. Moreover, since ν is absolutely continuous, the inverse h^{-1} of the transport map h is well-defined and the random variable $h^{-1}(\tilde{v})$ has the standard normal distribution and satisfies $h(h^{-1}(\tilde{v})) = \tilde{v}$, a.s.

2.4.2. *The function* R. Let $p(u, \cdot)$ denote the probability density function (pdf) of $N(0, u)$ for $u \in (0, 1]$; recall that p is the fundamental solution of the heat equation on $\mathbb{R}$. Lemma 4.2, (1), guarantees that the function $R : [0,1] \times \mathbb{R} \to \mathbb{R}$, given by

$$R(t, \xi) = \int \Big(\int_0^{\xi+\zeta} h(x)\, dx \Big) p(t, \zeta)\, d\zeta, \tag{2.6}$$

is well defined. Moreover, it belongs to the class $C^{1,2}((0,1]\times\mathbb{R})\cap C([0,1]\times\mathbb{R})$ and solves the following initial-value problem (see, e.g., (Karatzas and Shreve, 1991, Section 4.3, p. 254) for details)

$$\begin{cases} R_t = \frac{1}{2}R_{\xi\xi},\ (t,\xi)\in(0,1)\times\mathbb{R}, \\ R(0,\xi) = \int_0^\xi h(x)\,dx,\ \xi\in\mathbb{R}. \end{cases} \tag{2.7}$$

The dominated convergence theorem implies that the equation (2.6) can be differentiated under the integral sign and that the derivative R_ξ solves an initial problem for the heat equation, too, but with the terminal condition $R_\xi(0,\xi) = h(\xi)$. Note that the monotonicity of h implies that $R(t,\cdot)$ is convex and $R_\xi(t,\cdot)$ nondecreasing.

2.4.3. *Processes G, Σ and λ.* The core of the argument needed for the construction of processes Σ and λ is given in the following proposition, whose proof is postponed until section 4.

Proposition 2.1. *Under Assumption 2.1, the backward stochastic differential equation (BSDE)*

$$G_t = \int_t^T \Big(\hat\rho^2\sigma_s^2 - \frac{U_s^2}{4G_s}\Big)\,ds - \int_t^T U_s\,dW_s \tag{2.8}$$

admits a solution (G,U), unique in the class $\mathcal{S}_0^+\times\mathcal{P}^2$. This solution has the following properties:

$$\frac{U}{G}\in\mathrm{bmo}\ \textit{and}\ \int_0^T \frac{\sigma_s^2}{G_s}\,ds = +\infty,\ \textit{a.s.} \tag{2.9}$$

Using the process G of Proposition 2.1 above, we define two more $\mathcal{F}^W$-adapted processes

$$\Sigma_t = \exp\Big(-\int_0^t \frac{\hat\rho^2\sigma_s^2}{G_s}\,ds\Big)\ \text{for } t<T,\ \text{and } \lambda_t = \frac{\sqrt{\Sigma_t}}{\sqrt{G_t}}, \tag{2.10}$$

noting that, by the second statement in (2.9), the extension $\Sigma_T = 0$ makes Σ continuous. By Itô's formula we have

$$d\Big(\frac{1}{\lambda_t}\Big) = \frac{1}{\lambda_t}\frac{U_t}{2G_t}\,dW_t, \tag{2.11}$$

and, by the first statement in (2.9), $\frac{1}{\lambda}$ is an $\mathcal{F}^W$-martingale.

2.4.4. *The functional $\boldsymbol{\xi}^*$ and the pricing rule $\boldsymbol{P}^*$.* With G, U, and λ at our disposal, we are ready to define the candidate pricing rule. First, we define the functional $\boldsymbol{\xi}^*$ which acts on an $\mathcal{F}$-semimartingale S as follows:

$$\boldsymbol{\xi}^*(S)_t = \int_0^t \lambda_s\Big(d\hat S_s - \frac{U_s}{2G_s}d[\hat S, W]_s\Big)\ \text{where } \hat S_t = S_t - \int_0^t \rho\sigma_s\,dW_s. \tag{2.12}$$

Since λ is continuous and $U/G\in\mathrm{bmo}\subseteq\mathcal{P}^2$ the integral in (2.12) exists a.s., and defines an $\mathcal{F}^{W,S}$-semimartingale for any semimartingale S, in a nonanticipating way.

With $\boldsymbol{\xi}^*$ at hand, we define the candidate pricing rule $\boldsymbol{P}^*(S)$ as a composition

$$\boldsymbol{P}^*(S)_t = H^*\big(t, \boldsymbol{\xi}^*(S)_t\big)\ \text{where } H^*(t,\xi) = R_\xi(\Sigma_t,\xi)\ \text{for } (t,\xi)\in[0,T]\times\mathbb{R}. \tag{2.13}$$

The adaptivity properties of H^* and $\boldsymbol{\xi}^*$ imply that $\boldsymbol{P}^*$ is, indeed, a pricing rule in the sense of Definition 2.1 above.

2.4.5. *Processes ξ^* and X^*.* We prove in Subsection 4.2 below that there exists a unique process ξ_t^* which is continuous on $[0,T]$ and satisfies

$$\xi_t^* = \int_0^t \frac{\lambda_s^2\hat\rho^2\sigma_s^2}{\Sigma_s}\big(h^{-1}(\tilde v) - \xi_s^*\big)\,ds + \hat\rho\int_0^t \lambda_s\sigma_s dB_s\ \text{for } t\in[0,T). \tag{2.14}$$

The process ξ^* is, in turn, used to define the candidate equilibrium trading strategy of the informed trader as

$$X_t^* := \int_0^t \frac{\lambda_s\hat\rho^2\sigma_s^2}{\Sigma_s}\big(h^{-1}(\tilde v) - \xi_s^*\big)\,ds,\ t\in[0,T], \tag{2.15}$$

so that by definition of X^* and ξ^* and denoting

$$Y_t^* = X_t^* + Z_t \text{ and } \hat{Y}_t^* = X_t^* + \int_0^t \hat{\rho}\sigma_s \, dB_s = Y_t^* - \int_0^t \rho\sigma_s \, dW_s,$$

we also have

$$\xi_t^* = \boldsymbol{\xi}^*(Y^*)_t.$$

2.5. **The main theorem and some properties of the equilibrium.** With all the main building blocks defined and the notation introduced in Subsection 2.4 above, we are ready to state our main result. We remind the reader that $\mathcal{F}^{m*} = \mathcal{F}^{Y^*,W}$ corresponds to the information available to the market maker, that $\overset{(d)}{=}$ denotes the equality in distribution, and that $h_\sharp\mu$ denotes the push-forward of the measure μ by the function h. The convex conjugate R^c of R is defined by

$$R^c(u,v) = \sup_{\xi\in\mathbb{R}}\Big(\xi v - R(u,\xi)\Big), \text{ for } u \in [0,1] \text{ and } v \in \mathbb{R}. \tag{2.16}$$

Theorem 2.1. *Under Assumption 2.1, the pair $(\boldsymbol{P}^*, X^*)$ constructed in subsection 2.4 above, is an equilibrium.*

Additionally, in that equilibrium:

(1) The $\tilde{v}$-conditional (i.e., time $0+$) expected profit/loss of the informed trader is

$$\frac{R(1,0) + R^c(0,\tilde{v})}{\lambda_0}$$

(2) There exists an $\mathcal{F}^{m*}$-Brownian motion $\hat{B}$ orthogonal to W so that

$$d\hat{Y}_t^* = \hat{\rho}\sigma_t \, d\hat{B}_t, \tag{2.17}$$

$$d\xi_t^* = \lambda_t \, d\hat{Y}_t^*, \text{ and} \tag{2.18}$$

$$dP_t^* = \frac{R_{\xi\xi}(\Sigma_t\, \xi_t^*)}{\frac{1}{\lambda_t}} \, d\hat{Y}_t^*. \tag{2.19}$$

Therefore, the processes ξ^*, Y^*, and $\hat{Y}^*$ are $\mathcal{F}^{m*}$-martingales, with ξ^* and $\hat{Y}^*$ orthogonal to W.

(3) $\xi_T^* = h^{-1}(\tilde{v})$ a.s., and, conditionally on $\mathcal{F}_t^{m*}$, we have $\xi_T^* \overset{(d)}{=} N(\xi_t^*, \Sigma_t)$. Thus, the $\mathcal{F}_t^{m*}$-conditional distribution of $\tilde{v}$ is $h_\sharp N(\xi_t^*, \Sigma_t)$.

Remark 2.5.

(1) The $\mathcal{F}^{m*}$-martingale property of ξ^*, the fact that $\xi_T^* = h^{-1}(\tilde{v})$, and the definition of the optimal strategy X^* imply that

$$\mathbb{E}\left[\frac{dX_t^*}{dt} \,\middle|\, \mathcal{F}_t^{m*}\right] = 0 \text{ for almost every } t \in [0,T]. \tag{2.20}$$

This is the "inconspicuous informed trading" property in Cho (2003); at each time, the trading intensity of the insider has zero expectation from the point of view of the market maker.

(2) On the filtration $\mathcal{F}$, ξ^* has the same dynamics as the price process in Collin-Dufresne and Fos (2016) and therefore it is of the new class of bridges constructed in Collin-Dufresne and Fos (2016). The final condition of this bridge is $h^{-1}(\tilde{v})$ and the price at maturity is $\tilde{v}$. Thus, all information is incorporated to the price at time $T-$. In fact, all qualitative properties of the price process in Collin-Dufresne and Fos (2016) hold for ξ^* in our context.

(3) The trading rate of the informed trader is proportional to mismatch $(h^{-1}(\tilde{v}) - \xi_t^*)$ between the value of $\xi_T^* = h^{-1}(\tilde{v})$ which is determined by the private information of the informed trader and the current value of the state process ξ_t^*. The proportionality term $\hat{\rho}^2\lambda_t\sigma_t^2\Sigma_t^{-1}$ is $\mathcal{F}_t^W-$measurable and explodes as $t \to T$ due to the equality $\Sigma_T = 0$. This term has the property that if the market maker solves a filtering problem, the $\mathcal{F}_t^{m*}$-conditional distribution of $h^{-1}(\tilde{v})$ is $N(\xi_t^*, \Sigma_t)$.

(4) It is standard to define Kyle's Lambda as the sensitivity of the price to the total-demand process Y^*. In our context, the natural quantity is not Y^* but the adjusted order flow $\hat{Y}^*_t = Y^*_t - \int_0^t \rho\sigma_s dW_s$ which is the innovation process from the perspective of the market maker (i.e., the informative part of total demand). Thus, in our context, Kyle's Lambda is given by

$$\frac{R_{\xi\xi}(\Sigma_t, \xi^*_t)}{\frac{1}{\lambda_t}}$$

which is positive thanks to the convexity of R in ξ.

(5) Similarly to the computation leading to (2.19) above, we can show that

$$dR_{\xi\xi}(\Sigma_t, \xi^*_t) = \frac{R_{\xi\xi\xi}(\Sigma_t, \xi^*_t)}{\frac{1}{\lambda_t}}\, d\hat{Y}^*_t$$

In equilibrium, Y^* is a martingale in the filtration $\mathcal{F}^{m*}$ of the market maker and Kyle's Lambda becomes the ratio of two $\mathcal{F}^{m*}$-local martingales, $R_{\xi\xi}(\Sigma_t, \xi^*_t)$ and $\frac{1}{\lambda_t}$. As is shown in Lemma 4.2, (8) below, these two positive local martingales are orthogonal to each other. Thus, when $R_{\xi\xi}(\Sigma_t, \xi^*_t)$ is a true martingale, similarly to Collin-Dufresne and Fos (2016), Kyle's Lambda is, in fact, a submartingale and therefore increasing on average. As noted in Collin-Dufresne and Fos (2016), the submartingality of Kyle's Lambda in our framework is in contrast with Back (1992); Baruch (2002); Bose and Ekren (2023); Cho (2003); Caldentey and Stacchetti (2010) where Kyle's Lambda decreases with time and the informed trader suffers less from adverse selection from her trades close to the maturity. In the context of Collin-Dufresne and Fos (2016) and ours, at the beginning of time, the informed trader has the possibility to wait for higher liquidity moments to inject her private information. Thus, in equilibrium, the Kyle's Lambda has to increase in average to incentivize the informed trader not to wait for these potential future high liquidity moments.

(6) It is also standard to define the market depth as the inverse of Kyle's Lambda. Given the submartingality of Kyle's Lambda one would expect its inverse

$$\frac{1}{\lambda_t R_{\xi\xi}(\Sigma_t, \xi^*_t)}$$

to be non-increasing in average. This is true in the case of Collin-Dufresne and Fos (2016) since this term is proportional to $\frac{1}{\lambda_t}$ which is a martingale. But in our context, as ratio of two orthogonal martingales, the market depth is again a submartingale and increases in average as long as $R_{\xi\xi}(\Sigma_t, \xi^*_t)$ is not constant.

(7) The introduction of processes G, Σ, and λ is an important contribution of Collin-Dufresne and Fos (2016) in the understanding of Kyle's models. In particular, we interpret the $\mathcal{F}^W$-measurable process λ as a way of changing the conditional distribution of the underlying state process. Indeed, although we are unable to describe explicitly the equilibrium distribution of the original state process Y_T conditional to the information of the market maker at time t, the random variable $\int_t^T \hat{\rho}^2\lambda_s^2\sigma_s^2\, ds = \Sigma_t$ is known to the market maker at time t due to the choice of λ. Thus, in equilibrium, ξ^*_T has a Gaussian distribution with mean ξ^*_t and variance Σ_t conditionally on the information of the market maker. Thus, by integrating Y against λ in the definition of the novel state process ξ^* in (2.12), we render the novel state process conditionally Gaussian which allows us to use the optimal transport tools of Back et al. (2021) and postulate an expression for h and H^*.

2.6. **Completeness of equilbrium dynamics.** Although the dynamics of the price obtained in (2.19) has stochastic volatility due to the stochasticity of Σ_t, the fact that $\int_t^T \sigma_s^2\lambda_s^2 ds$ is $\mathcal{F}^W_t$-measurable has implications in terms of completeness. For the purpose of stating this property, we take the perspective of a price taking agent who neglects the impact of trades on the price and considers the price dynamics (2.19) as given. Thus, we define the weak completeness in the sense of the following proposition.

Proposition 2.2. *The equilibrium price dynamics* (2.19) *yields to a weakly complete market in the sense that for any bounded measurable function F, there exists an $\mathcal{F}^{m*}$-adapted square integrable process α, so that*

$$F(P^*_T) = F(h(\xi^*_T)) = \mathbb{E}[F(h(\xi^*_T))] + \int_0^T \alpha_t d\xi^*_t = \mathbb{E}[F(P^*_T)] + \int_0^T \frac{\alpha_t}{R_{\xi\xi}(\Sigma_t\, \xi^*_t)} dP^*_t.$$

Remark 2.6. The Proposition 2.2 implies that there exists no choice of σ_t allowing us to obtain any incomplete stochastic volatility model as an equilibrium dynamics in our context.

3. Examples

We split our examples into two groups. In Subsection 3.1 we treat several different distributions ν of the fundamental value $\tilde{v}$ and investigate the shape of the corresponding implied volatility (IV) curve. Subsection 3.2 contains a description of a rough volatility model based on a stochastic Volterra equation that satisfies our regularity assumption, Assumption 2.1, (3).

3.1. **Distributions of the fundamental value and implied-volatility curves.**

3.1.1. *Normal belief of the market maker.* If ν is Gaussian $N(m_\nu, \sigma_\nu^2)$, then h defined in Assumption 2.1 is

$$h(\xi) = \sigma_\nu \xi + m_\nu \text{ and } R_{\xi\xi}(s,\xi) = \sigma_\nu.$$

Thus,

$$P_t^* = \sigma_\nu \xi_t^* + m_\nu$$

and we recover the equilibrium in Collin-Dufresne and Fos (2016).

In equilibrium, $dP_t^* = \sigma_\nu \hat{\rho} \lambda_t \sigma_t d\hat{B}_t$ has a stochastic diffusion term. However, (3) of Theorem 2.1 implies that, conditionally on $\mathcal{F}_t^{m*}$,

$$P_T^* - P_t^* = \sigma_\nu \hat{\rho} \int_t^T \lambda_s \sigma_s d\hat{B}_s$$

is Gaussian with mean 0 and variance $\sigma_\nu^2 \Sigma_t$.

3.1.2. *Log-normal belief of the market maker.* Let ν be the distribution of the random variable $m \exp\left(-\frac{\sigma_v^2}{2} + \sigma_v G\right)$, where $G \sim N(0,1)$. Then,

$$h(x) = m \exp\left(-\frac{\sigma_v^2}{2} + \sigma_v x\right) \text{ and } R_\xi(t,\xi) = m \exp\left(\frac{\sigma_v^2}{2}(t-1) + \sigma_v x\right).$$

We can compute the price as

$$P_t^* = R_\xi(\Sigma_t, \xi_t^*) = m \exp\left(\frac{\sigma_v^2}{2}(\Sigma_t - 1) + \sigma_v \xi_t^*\right). \tag{3.1}$$

Differentiating (3.1), we obtain the dynamics

$$\frac{dP_t^*}{P_t^*} = \sigma_v d\xi_t^* = \sigma_v \hat{\rho} \lambda_t \sigma_t d\hat{B}_t,$$

where $\hat{B}$ is defined in (2.17). Thus, $\frac{dP_t^*}{P_t^*}$ defines a martingale orthogonal to W with quadratic variation $\sigma_v^2(1-\Sigma_t)$. Note that in equilibrium, we obtain a stochastic (realized) volatility, $t \mapsto \sigma_v \hat{\rho} \lambda_t \sigma_t$, for the price process, where the increments of the price and volatility are driven by orthogonal Brownian motions.

We obtain the log-return $\ln \frac{P_T^*}{P_t^*} = -\frac{\sigma_v^2}{2}\Sigma_t + \sigma_v(\xi_T^* - \xi_t^*)$. Note that Σ_t and $\int_t^T \lambda_s^2 \sigma_s^2 ds$ are $\mathcal{F}_t^{m*}-$measurable. Thus, conditionally on $\mathcal{F}_t^{m*}$, the log-return is Gaussian with mean $-\frac{\sigma_v^2}{2}\Sigma_t$ and variance $\sigma_v^2(1-\rho^2)\int_t^T \lambda_s^2\sigma_s^2 ds = \sigma_v^2 \Sigma_t$ which means that informed trading imposes distributional constraints on the distribution of the log-returns.

Additionally, conditionally on $\mathcal{F}_t^{m*}$, the distribution of P_T^* is log-normal and the price of a call option with maturity T and strike K can be computed as

$$\begin{aligned}\mathbb{E}[(h(\xi_T^*) - K)^+ | \mathcal{F}_t^{m*}] &= \mathbb{E}\left[\left(P_t^* \exp\left(-\frac{\sigma_v^2}{2}\Sigma_t + \sigma_v(\xi_T^* - \xi_t^*)\right) - K\right)^+ | \mathcal{F}_t^{m*}\right] \\ &= BS\left(t, P_t^*, T, K, \sigma_v \sqrt{\frac{\Sigma_t}{T-t}}\right) = BS\left(t, P_t^*, T, K, \sigma_v \hat{\rho} \sqrt{\frac{\int_t^T \lambda_s^2 \sigma_s^2 ds}{T-t}}\right)\end{aligned}$$

where $BS(t,p,T,K,\hat{\sigma})$ is the Black-Scholes price of the call option where the volatility between t and T is $\hat{\sigma}$ and $\hat{\rho}^2\int_t^T \lambda_s^2\sigma_s^2 ds = \Sigma_t$ is $\mathcal{F}_t^W-$measurable. Thus, in our model, with a log-normal fundamental price, the implied stochastic volatility curve remains flat at each time $t\in[0,T]$ and equal to $\sigma_v\hat{\rho}\sqrt{\frac{\int_t^T \lambda_s^2\sigma_s^2 ds}{T-t}}$, and the realized stochastic volatility is $\sigma_v\hat{\rho}\lambda_t\sigma_t$.

Remark 3.1. The weak completeness result in Proposition 2.2 is in contrast with Back (1993); Back et al. (2021) where the authors consider a Kyle's model where the informed trader and the noise traders can also trade an option on the stock. Note that in our context with log-normal prices, $\sigma_v\hat{\rho}\sqrt{\frac{\int_t^T \lambda_s^2\sigma_s^2 ds}{T-t}}$ is the stochastic (implied) volatility of the stock price which can be seen as an exogenously given (uncontrolled) option price whereas in Back (1993); Back et al. (2021) the option is also priced via the equilibrium and its demand is controlled by the informed trader.

This distinction makes that in our context (or the context of Collin-Dufresne and Fos (2016)) the option still has stochastic volatility but as shown in Proposition 2.2, it is redundant in the sense that it can be created by trading in the stock whereas in Back (1993); Back et al. (2021) it is not redundant.

3.1.3. *Non-flat IV curve at initial time.* For general distributions of the fundamental prices, it is not possible to solve the heat equation (2.7) explicitly. However, we can still rely on the Feynman-Kac representation to express the equilibrium price as

$$P_t^* = \mathbb{E}[h(\xi_T^*)|\mathcal{F}_t^{m*}] = \frac{1}{\sqrt{2\pi}}\int h(\xi_t^* + \sqrt{\Sigma_t}y)e^{-\frac{y^2}{2}}dy,$$

where h is not necessarily exponential or linear as in the previous sections.

In fact, the distribution of ξ_T^* conditional on $\mathcal{F}_t^{m*}$ is $N(\xi_t^*,\Sigma_t)$, we can price any option with payoff $p\mapsto F(p)$ as

$$\mathbb{E}[F(h(\xi_T^*))|\mathcal{F}_t^{m*}] = \frac{1}{\sqrt{2\pi}}\int F(h(\xi_t^* + \sqrt{\Sigma_t}y))e^{-\frac{y^2}{2}}dy, \tag{3.2}$$

and in particular obtain the dynamics of the IV curve at maturity T for any t,ξ_t^*,Σ_t.

3.1.4. *Gaussian mixtures for returns.* We suppose that the distribution ν of $\tilde{v}$ is the mixture of log-normal distributions given by $\{X_{m_i,\sigma_i};w_i\}_{i=1}^N$, where

$$X_{m_i,\sigma_i} = m_i\exp\left(-\frac{1}{2}\sigma_i^2 + \sigma_i Z\right)$$

with Z being a standard normal random variable and $w_i\geq 0$ being weights satisfying $\sum_{i=1}^N w_i = 1$. This is equivalent to assuming that, conditionally on the choice of an index i with probability w_i, the log-price $\log(\tilde{v})$ will be given by $\ln(m_i)-\frac{1}{2}\sigma_i^2+\sigma_i Z$, i.e., the log-price is a Gaussian mixture. Then, the cumulative distribution function of ν is given by

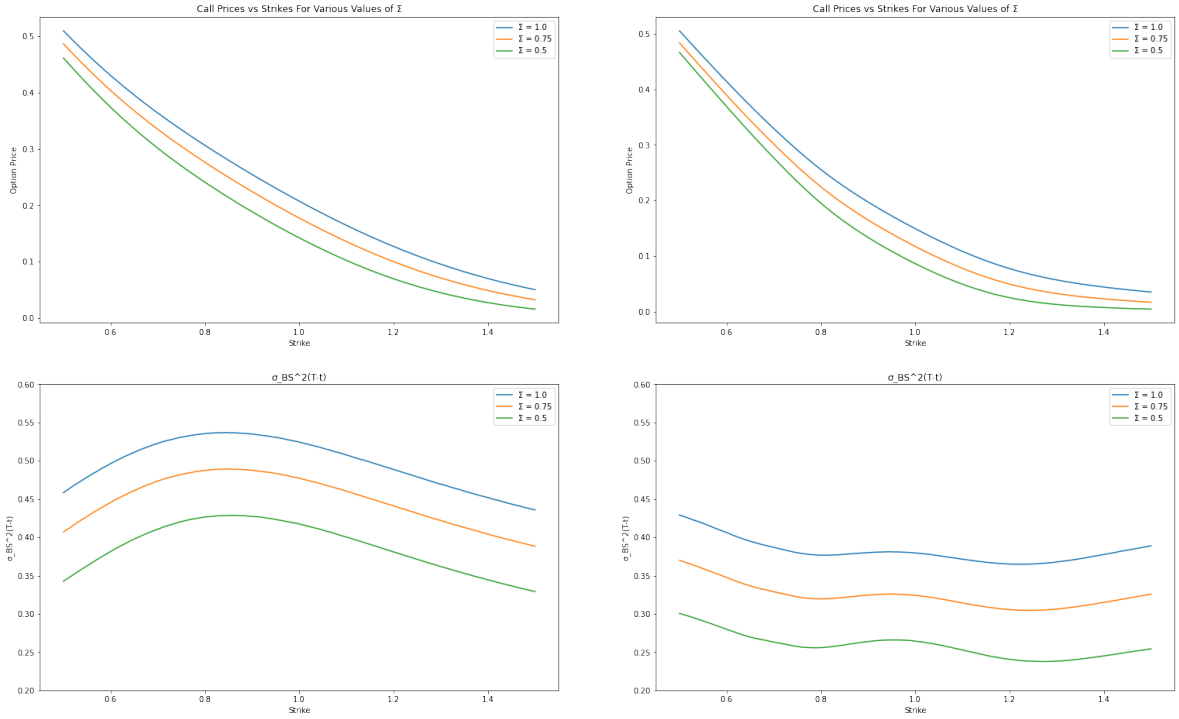
$$F_\nu(x) = \sum_{i=1}^N w_i\Phi\Big(\frac{\ln(\frac{x}{m_i})+\frac{1}{2}\sigma_i^2}{\sigma_i}\Big)$$

where Φ is the cumulative distribution function of the standard normal distribution. This means that the transport map $h(x)$ satisfies

$$\Phi(x) = F_\nu(h(x)) = \sum_{i=1}^N w_i\Phi\Big(\frac{\ln(\frac{h(x)}{m_i})+\frac{1}{2}\sigma_i^2}{\sigma_i}\Big)$$

and we have a simple way of numerically computing $h(x)$.

Since, conditionally on $\mathcal{F}_t^m$, ξ_T^* is normal with mean ξ_t^* and variance Σ_t, we can price options as in (3.2). Thus, we can compute numerically call option prices for any strike price K given values ξ of ξ_t^* and Σ of Σ_t. If the time to maturity $T-t$ is given, from the stock and option price, we can obtain the implied volatility σ_{BS} predicted from the model by inverting the Black-Scholes formula. In fact, in this inversion, $\sigma_{BS}^2(T-t)$ is only a function of ξ and Σ and not a function of $T-t$. Thus, in order to eliminate a parameter, instead of σ_{BS}, we compute $\sigma_{BS}^2(T-t)$. In figure 1, we plot the option prices and $\sigma_{BS}^2(T-t)$ as a function of the strike K for 3 different values of Σ and the number of mixtures $N=2$ and $N=3$. Corollary 4.1 of Glasserman and Pirjol (2023) applies to our model and



(A) $N = 2$ components

(B) $N = 3$ components

FIGURE 1. $\sigma^2_{BS}(T-t)$ of Call Options with Gaussian Mixture as log price distribution

shows that W-shaped curves are not possible for $N = 2$ components, as seen with the graph on the bottom left. On the other hand, for $N = 3$ components, we obtain W-shaped curves. The implied volatilities and option prices are increasing in Σ.

3.2. **Examples of volatility dynamics.** Our goal in this subsection is to establish a general sufficient condition for existence of exponential moments in Assumption 2.1, (3), and to apply it to two specific examples of commonly used volatility models, namely, the classical CIR model and its fractional counterpart. Without loss of generality, we assume that $T = 1$ throughout this subsection.

3.2.1. *A general sufficient condition.* Our sufficient condition, stated in Proposition 3.1 below, is based on the classical Garsia-Rodemich-Rumsey inquality (see Garsia et al. (1970)) whose statement we reproduce here for the convenience of the reader:

Theorem 3.1 (Garsia, Rodemich and Rumsey inequality)**.** *Suppose that Ψ and p are continuous and strictly increasing, $\Psi(0) = p(0) = 0$ and $\Psi(x) \to \infty$, as $x \to \infty$. For a continuous function $f : [0,1] \to \mathbb{R}$ let*

$$c = \int_0^1 \int_0^1 \Psi\Big(\frac{|f(t) - f(s)|}{p(|t-s|)}\Big)\, ds\, dt.$$

If $c < \infty$, then

$$|f(t) - f(s)| \le 8 \int_0^{t-s} \Psi^{-1}\Big(\frac{4c}{u^2}\Big)\, dp(u) \text{ for all } 0 \le s < t \le 1.$$

Lemma 3.1. *Let X be a continuous process on $[0,1]$. For $\gamma_0 \in (0,1)$ and $M > 0$, let*

$$F_M = \int_0^1 \int_0^1 \exp\left(M \frac{|X_t - X_s|}{|t-s|^{\gamma_0}}\right) dt\, ds. \tag{3.3}$$

If $\mathbb{E}[F_M] < \infty$ for all $M > 0$, then the pathwise γ-Hölder seminorm

$$|X|_\gamma = \operatorname*{esssup}_{0\le s<t\le 1} \frac{|X_t - X_s|}{|t-s|^\gamma}$$

admits all exponential moments, for each $\gamma \in [0, \gamma_0)$.

Proof. It follows directly from Theorem 3.1 with $p(u) = u^{\gamma_0}$ and $\Psi_M = \exp(Mx) - 1$ that, for each $M > 0$, with F_M is as in the statement, we have

$$|X|_\gamma \le L_M(4F_M) \text{ where } L_M(x) = \sup_{r\in(0,1]} 8r^{-\gamma} \int_0^r \Psi_M^{-1}\Big(\frac{x}{u^2}\Big)\, dp(u).$$

For $x \ge 2$ and $r \in (0,1)$ we have

$$M\int_0^r \Psi_M^{-1}\Big(\frac{x}{u^2}\Big)\, dp(u) = \gamma_0 \int_0^r \log\Big(1+\frac{x}{u^2}\Big) u^{\gamma_0-1}\, du \le \gamma_0 \int_0^r \log\left(\frac{3}{2}\frac{x}{u^2}\right) u^{\gamma_0-1}\, du \le L(r),$$

where

$$L(r) = 2\gamma_0 \int_0^r \log\Big(\frac{x}{u^2}\Big) u^{\gamma_0-1}\, du = \frac{2r^{\gamma_0}}{\gamma_0}\Big(2 + \gamma_0 \log\Big(\frac{x}{r^2}\Big)\Big).$$

We pick $\gamma \in [0, \gamma_0)$ and observe that for x large enough, namely $x \ge x_0 := \exp\Big(\frac{2\gamma}{\gamma_0(\gamma_0-\gamma)}\Big)$, the function

$$r \mapsto 8r^{-\gamma} L(r) = 8r^{\gamma_0-\gamma}(2 + \gamma_0 \log(x/r^2))$$

is non-decreasing on $(0,1]$. Therefore,

$$L_M(x) \le \frac{2}{M}\left(\frac{2}{\gamma_0} + \log(\max(x, x_0, 2))\right),$$

so that, for each $k > 0$ and each $M > 0$, we have

$$\mathbb{E}\Big[e^{k|X|_\gamma}\Big] \le e^{4k/(M\gamma_0)}\mathbb{E}\Big[e^{2k/M \log(\max(4F_M, x_0, 2))}\Big] = e^{4k/(M\gamma_0)}\mathbb{E}\Big[(\max(4F_M, x_0, 2))^{2k/M}\Big].$$

Given $k > 0$, it remains to take $M = 2k$ and use the assumption that $\mathbb{E}[F_M] < \infty$. □ □

Proposition 3.1. *Suppose that there exist constants $\varepsilon > 0$ and $\gamma_0 > 0$ such that*

$$\sup_{s\ne t\in[0,1]} \mathbb{E}\left[\exp\left(\varepsilon \frac{|\sigma_t^2 - \sigma_s^2|}{|t-s|^{\gamma_0}}\right)\right] < \infty.$$

Then the γ-Hölder semi-norm of σ admits all exponential moments for any $\gamma \in [0, \gamma_0/2)$.

Proof. Using the elementary fact that $|b-a| \le \sqrt{|b^2 - a^2|}$ for all $a, b \ge 0$, followed by Young's inequality, we obtain that for all $s \ne t \in [0,1]$, $M \ge 0$ and $\varepsilon \in (0,1)$ we have

$$M\frac{|\sigma_t - \sigma_s|}{|t-s|^{\gamma_0/2}} \le M\sqrt{\frac{|\sigma_t^2 - \sigma_s^2|}{|t-s|^{\gamma_0}}} \le \varepsilon \frac{|\sigma_t^2 - \sigma_s^2|}{|t-s|^{\gamma_0}} + \frac{M^2}{4\varepsilon}.$$

We use Lemma 3.2 to conclude that

$$\mathbb{E}\left[\exp\left(M\frac{|\sigma_t - \sigma_s|}{|t-s|^{\gamma_0/2}}\right)\right] \le C\exp\left(C\varepsilon\kappa + \frac{M^2}{\varepsilon}\right) < \infty. \tag{3.4}$$

Since the right-hand side of (3.4) is independent of s, t, we conclude that, for each M,

$$\sup_{s\ne t\in[0,1]} \mathbb{E}\left[\exp\left(M\frac{|\sigma_t - \sigma_s|}{|t-s|^{\gamma_0/2}}\right)\right] < \infty.$$

It remains to apply Lemma 3.1. □ □

3.2.2. *CIR-based volatility.* Let (V_t) be a nonnegative CIR process on $[0,1]$ started at $x>0$ at $t=0$. More precisely, we assume that V admits the dynamics of the form

$$V_t = x + \int_0^t (a - kV_s)\,ds + \int_0^t \eta\sqrt{V_t}\,dW_t, \tag{3.5}$$

where $a,k,\eta>0$ are such that the Feller condition $2a\geq\eta^2$ is satisfied; consequently, $V_t\geq 0$ for all $t\in[0,1]$, a.s. An explicit formula for the moment-generating function of V_t (see Alfonsi (2015), Proposition 1.2.4, p. 7) is given by

$$\mathbb{E}[\exp(uV_t)] = F(t,u)^{2a/\eta^2} e^{u\exp(-kt)F(t,u)x}, \text{ where } F(t,u) = \left(1 - \frac{\eta^2}{2k}\big(1-e^{-kt}\big)u\right)^{-1}, \tag{3.6}$$

for $u<u_0 = 2k\eta^{-2}(1-\exp(-k))^{-1}$.

Since $(1-e^{-1})\leq \frac{1}{s}(1-\exp(-s))\leq 1$, for all $s\in(0,1]$, there exist constants $0<c_1<c_2$, that depend only on a,k and η, such that

$$(1-c_1\varepsilon\sqrt{t})^{-1}\leq F\left(t,\frac{\varepsilon}{\sqrt{t}}\right)\leq\left(1-c_2\varepsilon\sqrt{t}\right)^{-1} \text{ for } t\in(0,1] \text{ and } \varepsilon\in\left(-\infty,\tfrac{1}{2}\max(c_1,c_2)^{-1}\right).$$

Therefore, there exists a constant C such that $\left|F(t,\varepsilon/\sqrt{t})-1\right|\leq C\sqrt{t}$ for $t\in(0,1]$ and $\varepsilon<\frac{1}{2}\max(c_1,c_2)^{-1}$. Perhaps with a different constant C we then also have

$$\left|\exp(-kt)F(t,\varepsilon/\sqrt{t})-1\right|\leq C\sqrt{t} \text{ for } \varepsilon<\tfrac{1}{2}\max(c_1,c_2)^{-1}. \tag{3.7}$$

This estimate implies, in particular, that $F(t,\varepsilon/\sqrt{t})$ admits an upper bound, uniform in t and all sufficiently small ε. Therefore, for $\varepsilon>0$ small enough, (3.6) and (3.7) imply that

$$\begin{aligned}\mathbb{E}\left[\exp\left(\varepsilon\frac{|V_t-x|}{\sqrt{t}}\right)\right] &\leq \mathbb{E}\left[\exp\left(\varepsilon\frac{V_t-x}{\sqrt{t}}\right)\right] + \mathbb{E}\left[\exp\left(-\varepsilon\frac{V_t-x}{\sqrt{t}}\right)\right]\\ &= F^{2a/\eta^2}\left(t,\frac{\varepsilon}{\sqrt{t}}\right)\left(\exp\left(\frac{\varepsilon}{\sqrt{t}}x\left(\exp(-kt)F\left(t,\frac{\varepsilon}{\sqrt{t}}\right)-1\right)\right)\right.\\ &\quad\left.+\exp\left(\frac{\varepsilon}{\sqrt{t}}x\left(1-\exp(-kt)F\left(t,-\frac{\varepsilon}{\sqrt{t}}\right)\right)\right)\right)\leq C\exp(C\varepsilon x),\end{aligned}$$

for some constant C which depends only on a,k and η. Since V is a homogeneous Markov process, we deduce that for all $0\leq s<t\leq 1$ and all $x\geq 0$, we have

$$\mathbb{E}\left[\exp\left(\varepsilon\frac{|V_t-V_s|}{\sqrt{t-s}}\right)\right]\leq C\mathbb{E}[\exp(C\varepsilon V_{t-s})].$$

It follows directly from (3.6) that, for small enough ε, we have $\sup_{\delta\in[0,1]}\mathbb{E}[\exp(C\varepsilon V_\delta)]<\infty$. Thus, the conditions of Proposition 3.1 are met, and so for $0<\underline{\sigma}<\overline{\sigma}$, the process $\sigma_t=\underline{\sigma}+\min(\sqrt{V_t},\overline{\sigma})$ satisfies Assumption 2.1, (3).

3.2.3. *The rough CIR model.* Next, we consider a class of models driven by Volterra stochastic differential equations, including a truncation of the volatility process in the rough Heston model, as described in El Euch and Rosenbaum (2019); Abi Jaber et al. (2019). We mention also that in Biagini et al. (2012) a Kyle model where the noise trading is assumed to be a fractional process is studied.

We fix $H\in(0,1/2)$, set $\alpha=H+1/2$, and let

$$K(t) = t^{\alpha-1}/\Gamma(\alpha) \tag{3.8}$$

be the α-fractional kernel, where Γ denotes the Gamma function. We choose two bounded continuous functions $b,s:\mathbb{R}\to\mathbb{R}$ and a constant κ such that

$$|b(x)|, s^2(x)\leq\kappa, \text{ for all } x\in\mathbb{R}.$$

We also assume that the stochastic Volterra equation

$$V_t = V_0 + \int_0^t K(t-u)b(V_u)\,du + \int_0^t K(t-u)s(V_u)\,dW_u,\ t\in[0,1], \tag{3.9}$$

admits a strong solution. One sufficient condition is the Lipschitz continuity of the coefficients b and s (see Abi Jaber et al. (2019), Theorem 3.3., p. 3167). A (financially) more relevant case occurs when b and s define a **truncated Volterra square-root process**, i.e., when

$$V_t = V_0 + \int_0^t K(t-u)(b^0 - b^1 V_u)\,du + \int_0^t K(t-u)\sqrt{A^1 \min(V_t, \overline{V})}\,dW_u, \tag{3.10}$$

where b^0, b^1, A^1 and $\overline{V}$ are nonnegative constants. The existence, uniqueness and positivity theory of (3.10), parallels closely that of its version presented in Lemma 6.3, p. 3182 of Abi Jaber et al. (2019), so we do not go into details here.

Fix $0 < \underline{\sigma} < \overline{\sigma}$. To show that the process $\sigma_t = \underline{\sigma} + \min(\sqrt{V_t}, \overline{\sigma})$ satisfies Assumption 2.1, (3), we employ Proposition 3.1 above to $t \mapsto \sqrt{V_t}$. Lemma 3.2 below ensures that its conditions are satisified.

Lemma 3.2. *There exists a constant $C > 0$, which depends only on α and κ such that*

$$\mathbb{E}\left[\exp\left(\varepsilon \frac{|V_t - V_s|}{|t-s|^H}\right)\right] \leq C\exp(C\varepsilon\kappa) \text{ for all } s \neq t \text{ and } \varepsilon \in [0,1). \tag{3.11}$$

Proof. Given $0 \leq s < t \leq 1$, we decompose

$$V_t - V_s = A + M + D + N,$$

where

$$\begin{aligned} A &= \int_0^s (K(t-u) - K(s-u))b(V_u)\,du \\ M &= \int_0^s (K(t-u) - K(s-u))s(V_u)\,dW_u \\ D &= \int_s^t K(t-u)b(V_u)\,du \\ N &= \int_s^t K(t-u)s(V_u)\,dW_u. \end{aligned}$$

For $r \geq 0$, the Cauchy-Schwarz inequality implies that

$$\begin{aligned} \mathbb{E}[\exp(r|V_t - V_s|)] &\leq \mathbb{E}[\exp(r(|A| + |D|))\exp(r|M+N|)] \\ &\leq \mathbb{E}[\exp(2r(|A| + |D|)]^{1/2}\mathbb{E}[\exp(2r|M+N|)]^{1/2}. \end{aligned}$$

With C denoting a generic constant, which may change from use to use and is allowed to depend only on α and κ, we have

$$\int_s^t K(t-u)\,du = C(t-s)^\alpha \text{ and } 0 \leq \int_0^s K(s-u) - K(t-u)\,du = C(s^\alpha - t^\alpha + (t-s)^\alpha) \leq C(t-s)^\alpha.$$

Combining the equalities above with fact that $(t-s)^\alpha \leq (t-s)^H$, we obtain that

$$\begin{aligned} |A| &\leq C\kappa \int_0^s (K(s-u) - K(t-u))\,du \leq C\kappa(t-s)^\alpha \leq C\kappa(t-s)^H \text{ and} \\ |D| &\leq C\kappa \int_s^t K(t-u)\,du \leq C\kappa(t-s)^\alpha \leq C\kappa(t-s)^H. \end{aligned}$$

Consequently,

$$\mathbb{E}[\exp(2r(|A| + |D|))] \leq \exp\big(Cr\kappa(t-s)^H\big). \tag{3.12}$$

Turning to the $M+N$-term, we note that $M + N = \int_0^t H_u\,dW_u$ where

$$H_u = \begin{cases} (K(t-u) - K(s-u))s(V_u) & u \in [0,s], \\ K(t-u)s(V_u) & u \in [s,t]. \end{cases}$$

Using the inequality $\mathbb{E}[\exp(\pm L_t)] \leq \mathbb{E}[\exp(2\langle L\rangle_t)]^{1/2}$, valid for any continuous local martingale L, we conclude that

$$\mathbb{E}[\exp(2r|M+N|)] \leq \mathbb{E}[\exp(2r(M+N))] + \mathbb{E}[\exp(-2r(M+N))] \leq 2\mathbb{E}\left[\exp\left(8r^2\int_0^t H_u^2\,du\right)\right]^{1/2}.$$

For $0 \leq u \leq s \leq t$, we have that

$$\begin{aligned} 0 &\leq ((s-u)^{\alpha-1} - (t-u)^{\alpha-1})^2 \leq (s-u)^{2\alpha-2} - 2(s-u)^{\alpha-1}(t-u)^{\alpha-1} + (t-u)^{2\alpha-2} \\ &\leq (s-u)^{2\alpha-2} - (t-u)^{2\alpha-2}. \end{aligned}$$

The (in)equalities

$$\int_s^t K(t-u)^2\,du = C(t-s)^{2\alpha-1} \text{ and } \int_0^s (K(t-u)-K(s-u))^2\,du \leq C(t-s)^{2\alpha-1}$$

further yield

$$\begin{aligned} \mathbb{E}[\exp(2r|M+N|)] &\leq 2\mathbb{E}\left[\exp\left(8r^2\Big(\int_0^s (K(s-u)-K(t-u))^2\eta(V_u)^2\,du + \int_s^t K(t-u)^2(\eta(V_u))^2\,du\Big)\right)\right]^{1/2} \\ &\leq 2\exp\big(Cr^2\kappa(t-s)^{2\alpha-1}\big)^{1/2}. \end{aligned}$$

Since $(t-s)^{2\alpha-1} = (t-s)^{2H}$, this implies that

$$\mathbb{E}[\exp(2r|M+N|)] \leq 2\exp\left(C\Big(r(t-s)^H\Big)^2\kappa\right)^{1/2}. \tag{3.13}$$

If we replace r by $\varepsilon/(t-s)^H$, combine the estimates (3.12) and (3.13) above, and use that $\varepsilon^2 < \varepsilon$ for $\varepsilon \in [0,1)$, we get

$$\mathbb{E}\left[\exp\left(\varepsilon\tfrac{|V_t-V_s|}{|t-s|^H}\right)\right] \leq \sqrt{2}\exp(C\varepsilon\kappa)^{1/2}\exp\big(C\varepsilon^2\kappa\big)^{1/4} \leq C\exp(C\varepsilon\kappa)$$

□ □

4. Proofs

We divide this section into two parts. In the first part we work towards the proof of the Proposition 2.1, while the second one focuses on the proof of the main Theorem 2.1. We refer the reader to subsection 2.4 above for all unexplained notation.

4.1. Proof of Proposition 2.1.

We start with a modest generalization of the standard existence and comparison result for Lipschitz BSDE in a special case.

Lemma 4.1.

(1) (Existence) Suppose that the random field $f : [0,T]\times\Omega\times\mathbb{R}\to\mathbb{R}$ is $\mathcal{F}^W$-progressively measurable, $f(t,0) \in \mathcal{S}^\infty$,

$$|f(t,y') - f(t,y)| \leq C|y'-y| \text{ for all } y, y' \in \mathbb{R} \text{ and all } t, \text{ a.s.} \tag{4.1}$$

and that b is a bmo *process. Then the BSDE*

$$y_t = \int_t^T \Big(f(u,y_u) + z_u b_u\Big)\,du + \int_t^T z_u\,dW_u \tag{4.2}$$

admits a solution $(y,z) \in \mathcal{S}^\infty \times$ bmo.

(2) (Comparison) Suppose that $b \in$ bmo, f^1 and f^2 are $\mathcal{F}^W$-progressive random fields which satisfy (4.1) above and $f^1(t,y) \leq f^2(t,y)$ for all t, y, a.s. If (y^1,z^1) and (y^2,z^2) in $\mathcal{S}^\infty \times$ bmo satisfy

$$y^i = \int_t^T (f^i(u,y_u^i) + z_u^i\,b_u)\,du + \int_t^T z_u^i\,dW_u,$$

then $y_t^1 \geq y_t^2$, for all t, a.s. In particular, the solution in 1. above is unique.

Proof. (1) Let $\Gamma_t = \mathcal{E}(\int_0^t b_u\,dW_u)$, and let $\mathbb{E}^b$ denote the expectation under the probability measure $\mathbb{P}^b$ defined by $d\mathbb{P}^b = \Gamma_T\,d\mathbb{P}$. We define $\Phi : \mathcal{S}^\infty \to \mathcal{S}^\infty$ by

$$\Phi(y)_t = \mathbb{E}^b\left[\int_t^T f(u, y_u)\,du \,\middle|\, \mathcal{F}_t^W\right]$$

and note that the Lipschitz property of f implies that

$$\|\Phi(y^2)_t - \Phi(y^1)_t\|_{L^\infty} \le C\int_t^T \|y_u^2 - y_u^1\|_{L^\infty}\,du.$$

From there, it follows immediately that Φ is a contraction on $\mathcal{S}^\infty$ equipped with the (Banach) norm $\|y\|_w = \sup_t e^{2Ct}\|y_t\|_{L^\infty}$. The fixed point $\hat{y}$ of Φ has the property that $\hat{y}_T = 0$ and $\Gamma_t\hat{y}_t + \int_0^t \Gamma_u f(u, \hat{y}_u)\,du$ is a $\mathbb{P}^b$-martingale satisfying

$$\sup_{t\in[0,T]}|\Gamma_t\hat{y}_t| + \int_0^t |\Gamma_u f(u,\hat{y}_u)|\,du \le C\sup_{0\le t\le T}|\Gamma_t|.$$

Theorem 3.1., p. 57, in Kazamaki (2006) applied to Γ and the martingale representation theorem imply that (4.2) holds for some progressive process $\hat{z}$ with $\mathbb{E}^b\left[\sqrt{\int_0^T \hat{z}_u^2\,du}\right] < \infty$. Since $\hat{y}_t$ and $f(t,\hat{y}_t)$ are both bounded processes, the process $\int_0^t \hat{z}_u(-b_u\,du + dW_u)$ is a bounded $\mathbb{P}^b$-local martingale, and, therefore, a $\mathbb{P}^b$-BMO martingale. Thanks to the invariance of bmo spaces under equivalent measure changes (see (Kazamaki, 2006, Theorem 3.3., p. 57)), $\hat{z}$ is a bmo process with respect to $\mathbb{P}$, as well.

(2) This can be proved using the standard linearization method (see, e.g., the proof of (Zhang, 2017, Theorem 4.4.1, p. 87)) while keeping in mind the fact that $\mathcal{E}(\int_0^\cdot b_u\,dW_u)$ is a uniformly integrable martingale as soon as b is a bmo-process (see (Kazamaki, 2006, Theorem 2.3, p. 31)). □ □

Proof of Proposition 2.1. Since we are dealing only with $\mathcal{F}^W$-adapted processes in this proof, we assume, without loss of generality, that we are working on the canonical space $C[0,T]$, equipped with the Wiener measure. In particular, for $\omega, \omega' \in C[0,T]$ and $t \in [0,T]$, we define the concatenated trajectory $\omega \oplus_t \tilde{\omega} \in C[0,T]$ by

$$(\omega \oplus_t \tilde{\omega})_s = \begin{cases} \omega_s, & s \in [0,t) \\ \omega_t + \tilde{\omega}_{s-t} - \tilde{\omega}_0, & s \in [t,T]. \end{cases}$$

The first step is to solve the auxiliary BSDE

$$y_t = \int_t^T \Big(\frac{\tilde{J}_u^2}{2y_u} - z_u\psi_u\Big)\,du + \int_t^T z_u\,dW_u, \tag{4.3}$$

where $\tilde{J}_t = \hat{\rho}J_t$, ψ is given by $L = L_0\mathcal{E}(\int_0^\cdot \psi_t\,dW_t)$, and J, L are as in (2.4). To do that, we pose a sequence of BSDEs

$$y_t^n = \int_t^T \Big(\frac{\tilde{J}_u^2}{2(\frac{1}{n}\tilde{J}_0 + |y_u^n|)} - z_u^n\psi_u\Big)\,du + \int_t^T z_u^n\,dW_u, \tag{4.4}$$

each of which has a unique solution $(y^n, z^n) \in \mathcal{S}^\infty \times \mathrm{bmo}$ by Lemma 4.1, (1) above. The second part of the same Lemma implies that

$$n_1 \le n_2 \to y_t^{n_1} \le y_t^{n_2} \text{ for all } t \in [0,T], \text{ a.s.}$$

Assuming, without loss of generality, that we are working on the canonical space $C[0,T]$, we define the adapted processes

$$\tilde{J}^+(t,\omega) = \operatorname*{esssup}_{s\ge t,\,\tilde{\omega}\in C[0,T]} |\tilde{J}_s(\omega\oplus_t\tilde{\omega})|,\ \tilde{J}^-(t,\omega) = \operatorname*{essinf}_{s\ge t,\,\tilde{\omega}\in C[0,T]} |\tilde{J}_s(\omega\oplus_t\tilde{\omega})|,$$

and for fixed $t \in [0,T]$ consider two additional families of BSDEs

$$y_s^{t,\pm,n} = \int_s^T \Big(\frac{(\tilde{J}_t^\pm)^2}{2(\frac{1}{n}\tilde{J}_0 + |y_u^{t,\pm,n}|)} - z_u^{t,\pm,n}\psi_u\Big)\,du + \int_s^T z_u^{t,\pm,n}\,dW_u. \tag{4.5}$$

For each $n \in \mathbb{N}$, these equations admit positive deterministic (conditionally on $\mathcal{F}_t^W$) solutions:

$$y_s^{t,\pm,n} = \sqrt{\tfrac{\tilde{J}_0^2}{n^2} + (\tilde{J}_t^{\pm})^2(T-s)} - \frac{\tilde{J}_0}{n},\ z_s^{t,\pm,n} = 0.$$

Moreover, since Lemma 4.1, (2) applies to these equations as well, we have

$$y_t^n \le y_t^{t,+,n} \le \tilde{J}_t^+ \sqrt{T-t}$$

as well as

$$\tfrac{1}{n}\tilde{J}_0 + y_t^n \ge \tfrac{1}{n}\tilde{J}_0 + y_t^{t,-,n} \ge \tilde{J}_t^- \sqrt{T-t}. \tag{4.6}$$

Combining both inequalities, we obtain the following estimates:

$$\tilde{J}_t^- \sqrt{T-t} - \tfrac{1}{n}\tilde{J}_0 \le y_t^n \le \tilde{J}_t^+ \sqrt{T-t}. \tag{4.7}$$

Let the process y be defined by $y_t = \liminf_{n\to\infty} y_t^n$. It is progressively measurable and we have $y_t = \lim_n y_t^n$, a.s., for each t. Thanks to (4.7) the sequence $(\tilde{J}_0/n + y_t^n)^{-1}$ is dominated by the integrable function $(T-t)^{-1/2}$ (up to a multiplicative constant), so that the dominated convergence theorem can be applied to pass to the limit $n \to \infty$ on both sides of the equality $y_t^n = \mathbb{E}_t^{\psi}[\int_t^T \tilde{J}_u^2/(2\tilde{J}_0/n + 2y_u^n)\,du]$ to conclude that

$$y_t = \mathbb{E}_t^{\psi}\left[\int_t^T \frac{\tilde{J}_u^2}{2y_u}\,du\right],\ \text{a.s., for each } t.$$

The fact the augmented filtration of $W^{\psi} = W - \int_0^{\cdot} \psi_u\,du$ is Brownian allows us to show that y admits a continuous modification, and, then, as in the proof of Lemma 4.1 above, that (y,z) satisfy the BSDE (4.3) for some $z \in \mathrm{bmo}$.

Passing to the limit in (4.7) we obtain

$$\tilde{J}_t^- \le h_t \le \tilde{J}_t^+ \text{ where } h_t = \frac{y_t}{\sqrt{T-t}}. \tag{4.8}$$

The two bounds above imply that h is positive, bounded, bounded away from 0 and $\lim_{t\to T} h_t = \tilde{J}_T$. Recalling that $|\tilde{J}|_\gamma$ denotes the γ-Holder norm of $\tilde{J}$, Itô's formula yields the following dynamics for h on $[0,T)$

$$dh_t = \Big(a_t + K_t\,\psi_t\Big)\,dt - K_t\,dW_t,$$

where $a_t = \frac{\tilde{J}_t + h_t}{2h_t}\frac{h_t - \tilde{J}_t}{T-t}$ is bounded by $\left(\frac{\tilde{J}_t}{2h_t} + \frac{1}{2}\right)|\tilde{J}|_\gamma |T-t|^{1-\gamma}$ and $K_t = \frac{z_t}{\sqrt{T-t}} = \frac{h_t z_t}{y_t}$. We apply Itô's formula a second time to obtain

$$d\ln h_t = \left(\frac{a_t}{h_t} + \frac{1}{2}\psi_t^2 - \frac{1}{2}\left(\frac{z_t}{y_t} - \psi_t\right)^2\right)dt - \frac{z_t}{y_t}\,dW_t$$

which, thanks to the identity $\sigma_t = L_t J_t = J_t e^{\int_0^t \psi_s dW_s - \frac{1}{2}\int_0^t \psi_s^2 ds}$, leads to

$$\frac{h_t}{J_t e^{\int_0^t \frac{a_s}{h_s}\,ds}}\sigma_t = h_0 e^{\int_0^t \psi_s - \frac{z_s}{y_s}\,dW_s - \frac{1}{2}\int_0^t \left(\psi_s - \frac{z_s}{y_s}\right)^2 ds}. \tag{4.9}$$

Assumption 2.1, (3) and the bounds of a imply that the term $e^{\int_0^t \frac{a_s}{h_s}\,ds}$ admits all moments. The right hand side of (4.9) is proportional to $1/\lambda_t$. In fact a direct computation shows that (4.9) leads to

$$\hat{\rho}^2\sigma_t = \sqrt{T} h_0 \sigma_0 \gamma_t e^{\int_0^t \psi_s - \frac{z_s}{y_s}\,dW_s - \frac{1}{2}\int_0^t \left(\psi_s - \frac{z_s}{y_s}\right)^2 ds}. \tag{4.10}$$

where $\gamma_t := \tilde{J}_t \frac{e^{\int_0^t \frac{a_s}{h_s}\,ds}}{h_t\sqrt{T}}$. We can compute

$$\begin{aligned}\int_0^T \gamma_t^2 dt &= \int_0^T \frac{\tilde{J}_t^2}{T h_t^2} e^{\int_0^t 2\frac{a_s}{h_s}\,ds} dt = \int_0^T \frac{\tilde{J}_t^2}{T h_t^2} e^{\int_0^t \frac{1 - \tilde{J}_s^2/h_s^2}{(T-s)}\,ds} dt \\ &= \int_0^T \frac{\tilde{J}_t^2}{h_t^2 (T-t)} e^{-\int_0^t \frac{\tilde{J}_s^2}{(T-s)h_s^2}\,ds} dt = 1\end{aligned}$$

which is deterministic.

As the last step in the existence proof, we define $G_t = y_t^2 L_t^2$ where $L_t = \mathcal{E}(\int_0^\cdot \psi_u \, dW_u)$. A direct computation implies that G satisfies the original BSDE (2.8) with

$$U_t = 2G_t\Big(\psi_t - \frac{z_t}{y_t}\Big). \tag{4.11}$$

Since $\psi \in \mathrm{bmo}$, $y_t \geq C\sqrt{T-t}$ for some constant C and $z_t/\sqrt{T-t} \in \mathrm{bmo}$, the equality (4.11) above implies that $\frac{U}{G} \in \mathrm{bmo}$. On the other hand, the relation (4.8) implies that $G_t \leq \eta(T-t)$ for some strictly positive random variable η. With σ being continuous and bounded away from 0, this implies that $\int_0^T \frac{\sigma_s^2}{G_s}\, ds = \infty$, a.s.

To prove uniqueness, let $(G^i, U^i) \in \mathcal{S}_0^+ \times \mathcal{P}^2$ for $i = 1, 2$ be two solutions of (2.8). By a direct computation, $y^i = \frac{\sqrt{G^i}}{L}$ solves (4.3). We define $\delta y = y^2 - y^1$ and $\delta z = z^2 - z^1$, and observe that

$$d\delta y_t = r_t \delta y_t \, dt + \delta z_t (dW_t + \psi \, dt) \text{ where } r_t = \frac{\tilde{J}_t^2}{2y_t^1 y_t^2}$$

on $[0, T)$. Since $r_t \geq 0$, the process R defined by $R_t = \exp\Big(-\int_0^t r_u \, du\Big)$ for $t < T$ is positive and bounded, and admits a limit as $t \to T$. Itô's formula implies that $R_t \delta y_t$ is a bounded local martingale on $[0, T)$ with $R_t \delta y_t \to 0$ as $t \to T$. It follows that $R_t \delta y_t$ is a uniformly integrable martingale with the last element 0 so that $R_t \delta y_t = 0$, for all t. since $R_t > 0$ for $t < T$, we conclude that $y_t^1 = y_t^2$ for all t, a.s., and $z^1 = z^2$ Leb $\times \mathbb{P}$-a.e. □ □

4.2. **Proof of Theorem 2.1.** The proof starts with the introduction of several processes used in the construction of the novel state process ξ^* of (2.14). After that, Lemma 4.2 collects some of their essential properties and establishes the existence and uniqueness of ξ^*; it also features three other facts needed in the sequel. The rest of this subsection lays out the details of the proof and is divided into several subsections. We refer the reader to subsection 2.4 above for all unexplained notation, and note that Assumption 2.1 is in force throughout.

We start with the auxiliary $\mathcal{F}^{W,B}$-martingale $\hat{Z}$ given by

$$\hat{Z}_t = \int_0^t \hat{\rho} \lambda_s \sigma_s \, dB_s, \tag{4.12}$$

which, by the Dambis-Dubins-Schwarz theorem, admits the representation

$$\hat{Z}_t = \beta_{\langle \hat{Z} \rangle_t} \text{ where } \langle \hat{Z} \rangle_t = \int_0^t \hat{\rho}^2 \lambda_s^2 \sigma_s^2 \, ds = 1 - \Sigma_t \in \mathcal{F}_t^W,$$

and where β is an $\mathcal{F}^{W,B}$-Brownian motion given by

$$\beta_u = \hat{Z}_{\Gamma_u} \text{ where } \Gamma_u = \inf\{t \geq 0 \,:\, 1 - \Sigma_t = u\}, \tag{4.13}$$

which satisfy $\Gamma_{1-\Sigma_t} = t$ and $1 - \Sigma_{\Gamma_u} = u$, thanks to continuity and strict monotonicity of Σ_t. With $\hat{\xi}$ defined on $[0, 1]$ by

$$\hat{\xi}_t = t h^{-1}(\tilde{v}) + (1-t) \int_0^t \frac{1}{1-s} \, d\beta_s, \text{ for } t \in [0, 1) \text{ and } \hat{\xi}_1 = h^{-1}(\tilde{v}), \tag{4.14}$$

we set

$$\xi_t^* = \hat{\xi}_{1-\Sigma_t}.$$

Lemma 4.2.

(1) For all $\xi \in \mathbb{R}$ and $u \in (0, 1]$, we have

$$\int \left(\int_0^{\xi+\zeta} |h(x)| \, dx \right) p(u, \zeta) \, d\zeta < \infty.$$

(2) The Brownian motion β is independent of W and $\tilde{v}$.

(3) *Conditionally on $\tilde{v}$, $\hat{\xi}$ is a Brownian bridge from 0 to $h^{-1}(\tilde{v})$ and admits the following dynamics:*

$$d\hat{\xi}_t = \frac{h^{-1}(\tilde{v}) - \hat{\xi}_t}{1-t}\,dt + d\beta_t \text{ for } t \in [0,1),\ \lim_{t\to 1}\hat{\xi}_t = \hat{\xi}_1 = h^{-1}(\tilde{v}),\ \text{a.s.} \tag{4.15}$$

(4) *$\hat{\xi}$ is a Brownian motion independent of W.*

(5) *For $u \in [0,1]$ and $p \in [1,2]$ we have*

$$\mathbb{E}\Big[\Big|R_\xi(1-u,\hat{\xi}_u)\Big|^p\Big] \le \mathbb{E}[|\tilde{v}|^p] < \infty. \tag{4.16}$$

(6) *$\mathbb{E}[R^c(0,\tilde{v})] < \infty$.*

(7) *The random variable ξ^*_T is $\mathcal{F}^{\xi^*,W}_t$-conditionally normal with mean ξ^*_t and variance Σ_t. In particular, ξ^* is an $\mathcal{F}^{\xi^*,W}$-martingale.*

(8) *The process ξ^* is a continuous $\mathcal{F}$-semimartingale on $[0,T]$, orthogonal to W. Moreover, it is the unique continuous process that satisfies (2.14) on $[0,T)$.*

(9) *We have*

$$d\xi^*_t = \lambda_t\, d\hat{Y}^*_t, \tag{4.17}$$

and the pairs (ξ^*,W), $(\hat{Y}^*,W)$, and (Y,W) all generate the same filtration.

Proof.

(1) Using the monotonicity of h to justify the inequality $|h(\xi+\sqrt{u}y)| \le |h(\xi)| + |h(y+\xi)|$, for $u \in [0,1]$, and a simple change of variables, we obtain

$$\begin{aligned}
&\int\left(\int_0^{\xi+\zeta}|h(x)|\,dx\right)p(u,\zeta)\,d\zeta = \int\left(\int_0^{\xi+y\sqrt{u}}|h(x)|\,dx\right)p(1,y)\,dy \le\\
&\le \int \max\Big(\big|h(\xi+\sqrt{u}y)\big|,|h(0)|\Big)\big|\xi+y\sqrt{u}\big|p(1,y)\,dy\\
&\le \int\Big(|h(0)|+|h(\xi)|+|h(\xi+y)|\Big)\big|\xi+y\sqrt{u}\big|p(1,y)\,dy\\
&\le C + C\int|h(r)|(1+|r|)p(1,r-\xi)\,dr = C + C\int|h(r)|(1+|r|)e^{-\xi^2/2}e^{\xi r}p(1,r)\,dr,
\end{aligned}$$

where C is a finite constant which depends on h and ξ. The last integral is finite by the Cauchy-Schwarz inequality because both $h(r)$ and $(1+|r|)e^{\xi r}$ are square-integrable with respect to the Gaussian measure with density $p(1,r)$.

(2) Since $\hat{Z}$ is a Brownian motion independent of W and $\tilde{v}$, and Γ is a time change with respect to $\mathcal{F}^W$, the process β is a continuous martingale, conditionally on W. Its quadratic variation is Brownian, so, by Lévy's criterion, β is a Brownian motion independent of W and $\tilde{v}$.

(3) By its definition, β is independent of $\tilde{v}$, so $\hat{\xi}$ is a Gaussian process conditionally on $\tilde{v}$. Moreover, its conditional mean and covariance functions, namely $th^{-1}(\tilde{v})$ and $\min(s,t)-st$, match those of the Brownian bridge from 0 to $h^{-1}(\tilde{v})$. The dynamics (4.15) is a direct consequence of Itô's formula.

(4) The process $\hat{\xi}$ is $\mathcal{F}^{\beta,\tilde{v}}$-adapted and both β and $\tilde{v}$ are independent of W, as established in (2) above, so $\hat{\xi}$ is independent of $\mathcal{F}^W$. To see that it is a Brownian motion, it suffices to observe that it is a Brownian bridge from 0 to an independent standard normal (or simply compute its mean and covariance functions as above).

(5) Since $R_\xi(1-\cdot,\cdot)$ is a space-time harmonic function and $\hat{\xi}$ is a Brownian motion, $V_u = R_\xi(1-u,\hat{\xi}_u)$ is a martingale with the terminal condition $R_\xi(0,\hat{\xi}_1) = \tilde{v}$ and the inequality (4.16) is a direct consequence of Jensen's inequality.

(6) Using the fact that the supremum in

$$R^c(0,\tilde{v}) = \sup_{\xi\in\mathbb{R}}\Big(\xi\tilde{v} - \int_0^\xi h(x)\,dx\Big)$$

is attained at $\xi = h^{-1}(\tilde v)$, we obtain

$$R^c(0,\tilde v) = \left|\tilde v h^{-1}(\tilde v) - \int_0^{h^{-1}(\tilde v)} h(x)\,dx\right| = \left|\int_0^{h^{-1}(\tilde v)} (\tilde v - h(x))\,dx\right|$$
$$\le \left|h^{-1}(\tilde v)\right| \max_{r\in[0,1]}\left|\tilde v - h(rh^{-1}(\tilde v))\right| \le \left|h^{-1}(\tilde v)\right|\Big(|\tilde v| + |h(0)|\Big),$$

where the last inequality follows from the monotonicity of h. The random variable $h^{-1}(\tilde v)$ is normally distributed and therefore in $\mathbb{L}^2$, and so, by Assumption 2.1, (2), both $h(0)h^{-1}(\tilde v)$ and $h^{-1}(\tilde v)\tilde v$ belong to $\mathbb{L}^1$.

(7) $\hat\xi$ is a Brownian motion independent of W and Σ is $\mathcal{F}^W$-adapted. Therefore, conditionally on $\mathcal{F}_T^W$, $\xi_t^* = \hat\xi_{1-\Sigma_t}$ is a centered Gaussian process with independent increments and the deterministic variance function $1-\Sigma_t$. Therefore, conditionally on $\mathcal{F}_T^W \vee \mathcal{F}_t^{\xi^*}$, the random variable ξ_T^* is normally distributed with mean ξ_t^* and variance Σ_t. The statement now follows by further conditioning on $\mathcal{F}_t^W$.

(8) To show that ξ^* solves (2.14), we use (4.15) to get the following representation

$$\xi_t^* = \int_0^{1-\Sigma_t} \frac{h^{-1}(\tilde v) - \hat\xi_u}{1-u}\,du + \beta_{1-\Sigma_t} = \int_0^t \frac{h^{-1}(\tilde v) - \xi_s^*}{\Sigma_s}\,\hat\rho^2\sigma_s^2\lambda_s^2\,ds, + \hat Z_t,\ t\in[0,T),$$

where we used the change of variable $s = \Sigma_{1-u}^{-1}$. Since $d\hat Z = \hat\rho\lambda\sigma\,dB$, ξ^* indeed satisfies (2.14).

For uniqueness, it suffices to note that the equation for the difference of two solutions is a linear ODE with random coefficients. These coefficients are bounded a.s., on $[0,t]$ for each $t<T$, and our claim follows from Gronwall's inequality.

By (4.15) above, $\hat\xi$ is a Brownian bridge, conditionally on $\tilde v$. Therefore, we have

$$\int_0^1 \frac{|h^{-1}(\tilde v) - \hat\xi_u|}{1-u}\,du < \infty,\ \text{a.s.}$$

The same change of variable as above, namely $s = \Sigma_{1-u}^{-1}$, allows us to conclude that (2.14) provides an $\mathcal{F}$-semimartingale decomposition of ξ^* on the entire $[0,T]$. The orthogonality with W is the direct consequence of the fact that W and B are orthogonal.

(9) The identity (4.17) follows directly from (2.14). Since λ is bounded and bounded away from 0, (ξ^*, W) and $(\hat Y^*, W)$ generate the same filtration. The pairs $(\hat Y^*, W)$ and (Y, W) generate the same filtration because the difference $Y^* - \hat Y^*$ is $\mathcal{F}^W$-adapted.

□ □

4.2.1. *An expression for* $\Pi(X, \boldsymbol{P}^*(X+Z))$. Let X be a trading strategy, $Y = X + Z$, $\hat Y = X - \int_0^{\cdot}\rho\sigma_s\,dW_s$, $\xi = \boldsymbol{\xi}^*(Y)$, and $P = \boldsymbol{P}^*(Y)$, with the functionals $\boldsymbol{\xi}^*$ and $\boldsymbol{P}^*$ defined in (2.12) and (2.13) above. We also introduce the following shortcuts for a generic function F:

$$(\Delta F)_t = F(\Sigma_t,\xi_t) - F(\Sigma_t,\xi_{t-}),\ \text{and}\ (\Delta^{2,\pm}F)_t = (\Delta F)_t - F_\xi(\Sigma_t,\xi_{t\pm})\Delta\xi_t.$$

Thanks to Itô's lemma, we have the following expressions for the dynamics of P and $[P,X]$:

$$dP_t = -\hat\rho^2\lambda_t^2\sigma_t^2 R_{\xi t}(\Sigma_t,\xi_{t-})\,dt + R_{\xi\xi}(\Sigma_t,\xi_{t-})\,d\xi_t + \frac{1}{2}R_{\xi\xi\xi}(\Sigma_t,\xi_{t-})d[\xi,\xi]_t^c + (\Delta^{2,-}R_\xi)_t,$$
$$d[P,X]_t = \lambda_t R_{\xi\xi}(\Sigma_t,\xi_{t-})\Big(d[X,X]^c + \hat\rho\sigma_t\,d[X,B]^c\Big) + (\Delta R_\xi)_t\Delta X_t.$$

Since $(\Delta^{2-}R)_t + (\Delta R_\xi)_t\Delta\xi_t = (\Delta^{2+}R)_t$, applying Ito's formula to $\frac{\tilde v\xi_t - R(\Sigma_t,\xi_t)}{\lambda_t}$ and rearranging terms, it follows that

$$\Pi(X,P)_t = \int_0^t (\tilde v - P_{s-})dX_s - [P,X]_t = (I) + (II) + (III) + (IV), \tag{4.18}$$

where

$$\begin{aligned}(I) &= \frac{\tilde{v}\xi_t - R(\Sigma_t,\xi_t)}{\lambda_t} + \frac{R(1,0)}{\lambda_0},\\ (II) &= -\int_0^t \lambda_s \frac{R_{\xi\xi}(\Sigma_s,\xi_{s-})}{2}\, d[X]^c + \sum_{s\le t}(\Delta^{2,+}R)_s,\\ (III) &= \int_0^t (R(\Sigma_s,\xi_{s-}) - \tilde{v}\xi_{s-})\, d\frac{1}{\lambda_s} \text{ and}\\ (IV) &= \hat{\rho}\int_0^t (P_{s-}-\tilde{v})\sigma_s dB_s.\end{aligned} \tag{4.19}$$

4.2.2. *An upper bound for $\boldsymbol{P}^*$-admissible strategies.* Suppose now that X is $\boldsymbol{P}^*$-admissible and let $P = \boldsymbol{P}^*(X+Z)$. Moreover, let (τ_n) be a common reducing sequence for the $(\mathcal{F}_t)$-local martingales in (III) and (IV) of (4.19). By the convexity of $R(\Sigma_t,\cdot)$, we have $\Delta^{2,+}R_t \le 0 \le \Delta^{2,-}R_t$. Therefore, part (II) is non-positive for each t, and the admissibility of X implies, via Fatou's lemma, that

$$\mathbb{E}[\Pi(X,P)_T \mid \tilde{v}] = \mathbb{E}\Big[\liminf_n \Pi(X,P)_{\tau_n} \,\Big|\, \tilde{v}\Big] \le \frac{1}{\lambda_0}R(1,0) + \liminf_n \mathbb{E}\Big[\frac{1}{\lambda_{\tau_n}}\Big(\tilde{v}\xi_{\tau_n} - R(\Sigma_{\tau_n},\xi_{\tau_n})\Big) \,\Big|\, \tilde{v}\Big]. \tag{4.20}$$

For all ξ and u, we have

$$\begin{aligned}\xi\tilde{v} - R(u,\xi) &\le \sup_\xi\Big(\xi\tilde{v} - \int R(0,\xi+\zeta)p(u,\zeta)\,d\zeta\Big)\\ &= \sup_\xi\Big(\int\big((\xi+\zeta)\tilde{v} - R(0,\xi+\zeta)\big)p(u,\zeta)\,d\zeta\Big)\\ &\le \int \sup_\xi\big((\xi+\zeta)\tilde{v} - R(0,\xi+\zeta)\big)p(u,\zeta)\,d\zeta = R^c(0,\tilde{v}).\end{aligned}$$

Since $0 \le R^c(0,\tilde{v})$, by the $\mathcal{F}$-martingale property of λ, we have

$$\mathbb{E}[\Pi(X,P)_T \mid \tilde{v}] \le \frac{1}{\lambda_0}R(1,0) + \liminf_n \mathbb{E}\Big[\frac{1}{\lambda_{\tau_n}}R^c(0,\tilde{v}) \,\Big|\, \tilde{v}\Big] = \frac{R(1,0)+R^c(0,\tilde{v})}{\lambda_0}. \tag{4.21}$$

4.2.3. *$\boldsymbol{P}^*$-admissibility of X^*.* The chain rule implies that

$$d\xi^*_t = \frac{\hat{\xi}_1 - \hat{\xi}_{1-\Sigma_t}}{\Sigma_t}\,d(1-\Sigma_t) + d\hat{Z}_t, \tag{4.22}$$

as well as

$$X^*_t = \int_0^{1-\Sigma_t} \frac{1}{\lambda_{\Gamma_u}}\frac{\hat{\xi}_1 - \hat{\xi}_u}{1-u}\,du. \tag{4.23}$$

We define $P^*_t = \boldsymbol{P}^*(X^*+Z)_t$ and use the identities (4.22) and (4.23), together with (2.2) above, to obtain the following expression for $\Pi(X^*,P^*)_t$:

$$\begin{aligned}\Pi(X^*,P^*)_t &= \int_0^t \frac{1}{\lambda_s}\Big(\tilde{v} - R_\xi(\Sigma_s,\xi^*_s)\Big)\Big(h^{-1}(\tilde{v}) - \xi^*_s\Big)\frac{d(1-\Sigma_s)}{\Sigma_s},\\ &= \int_0^{1-\Sigma_t} \frac{1}{\lambda_{\Gamma_u}}\Big(\tilde{v} - R_\xi(1-u,\hat{\xi}_u)\Big)\frac{\hat{\xi}_1 - \hat{\xi}_u}{1-u}\,du.\end{aligned} \tag{4.24}$$

Therefore,

$$\sup_{t\in[0,T]}\Big|\Pi(X^*,P^*)_t\Big| \le \int_0^1 \frac{1}{\lambda_{\Gamma_u}}\Big|h(\hat{\xi}_1)\Big|\tfrac{|\hat{\xi}_1-\hat{\xi}_u|}{1-u}\,du + \int_0^1 \frac{1}{\lambda_{\Gamma_u}}\Big|R_\xi(1-u,\hat{\xi}_u)\Big|\tfrac{|\hat{\xi}_1-\hat{\xi}_u|}{1-u}\,du.$$

Since $\frac{1}{\lambda_{\Gamma_u}}$ is a martingale in its own filtration and measurable with respect to $\mathcal{F}_T^W$, and $\mathcal{F}^W$ and $\hat{\xi}$ are independent by Lemma 2.4, (3), we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\Pi(X^*,P^*)_t\right|\right] \le \frac{1}{\lambda_0}\mathbb{E}\left[\int_0^1 |h(\hat{\xi}_1)|\tfrac{|\hat{\xi}_1-\hat{\xi}_u|}{1-u}\,du\right]+ \\ + \frac{1}{\lambda_0}\mathbb{E}\left[\int_0^1 \left|R_\xi(1-u,\hat{\xi}_u)\right|\tfrac{|\hat{\xi}_1-\hat{\xi}_u|}{1-u}\,du\right]. \tag{4.25}$$

Thus, the finiteness of the expectation on the LHS of (4.25) above boils down to the finiteness of the two terms on the right-hand side of (4.25) above. By Lemma 4.2, (4), $\hat{\xi}$ is a Brownian motion so that $\frac{\hat{\xi}_1-\hat{\xi}_u}{\sqrt{1-u}}$ is standard normal. Therefore, the inequality $ab \le \frac{1}{2}a^2+\frac{1}{2}b^2$ establishes the finiteness of the first term as follows:

$$\mathbb{E}\left[\int_0^1 |h(\hat{\xi}_1)|\frac{|\hat{\xi}_1-\hat{\xi}_u|}{1-u}\,du\right] \le \tfrac{1}{2}\int_0^1 \left(\mathbb{E}[|h(\hat{\xi}_1)|^2]+\mathbb{E}\left[\left|\tfrac{\xi_1-\xi_u}{\sqrt{1-u}}\right|^2\right]\right)\frac{1}{\sqrt{1-u}}\,du = \mathbb{E}\big[\tilde{v}^2\big]+1<\infty.$$

For the second term, we first use the fact that $\frac{\hat{\xi}_1-\hat{\xi}_u}{\sqrt{1-u}}$ is independent of $\hat{\xi}_u$:

$$\mathbb{E}\left[\int_0^1 \left|R_\xi(1-u,\hat{\xi}_u)\right|\tfrac{|\hat{\xi}_1-\hat{\xi}_u|}{1-u}\,du\right] = \int_0^1 \mathbb{E}\left[\tfrac{|R_\xi(1-u,\hat{\xi}_u)|}{\sqrt{1-u}}\right]\mathbb{E}\left[\tfrac{|\hat{\xi}_1-\hat{\xi}_u|}{\sqrt{1-u}}\right]du = \int_0^1 \mathbb{E}\left[\tfrac{|R_\xi(1-u,\hat{\xi}_u)|}{\sqrt{1-u}}\right]du.$$

By the definition of R we have the stochastic representation $R_\xi(1-u,\hat{\xi}_u) = \mathbb{E}\Big[h(\hat{\xi}_1)\,\Big|\,\hat{\xi}_u\Big]$. Thus, we have the finiteness by

$$\int_0^1 \frac{\mathbb{E}\Big[\Big|R_\xi(1-u,\hat{\xi}_u)\Big|\Big]}{\sqrt{1-u}}\,du \le \int_0^1 \frac{\mathbb{E}\Big[\Big|\mathbb{E}\Big[h(\hat{\xi}_1)\,\Big|\,\hat{\xi}_u\Big]\Big|\Big]}{\sqrt{1-u}}\,du \le \int_0^1 \frac{\mathbb{E}[|\tilde{v}|]}{\sqrt{1-u}}\,du<\infty.$$

We can, therefore, conclude that X^* is $\boldsymbol{P}^*$-admissible.

4.2.4. *P^*-optimality of X^*.* By (4.15) and the space-time harmonicity of $R(1-\cdot,\cdot)$, we have

$$\int_0^T \Big(\tilde{v}-R_\xi(1-u,\hat{\xi}_u)\Big)\frac{\hat{\xi}_1-\hat{\xi}_u}{1-u}\,du = \\ = \int_0^T \Big(\tilde{v}-R_\xi(1-u,\hat{\xi}_u)\Big)\,d\hat{\xi}_u - \tilde{v}\beta_T + \int_0^T R_\xi(1-u,\hat{\xi}_u)\,d\beta_u \\ = R^c(0,\tilde{v})+R(1,0)-\tilde{v}\beta_T+\int_0^T R_\xi(1-u,\hat{\xi}_u)\,d\beta_u. \tag{4.26}$$

By Lemma 4.2, (5), the $d\beta$-integral in the last line of (4.26) is a martingale. Since λ, $\tilde{v}$ and β are independent of each other, we have

$$\mathbb{E}\left[\frac{1}{\lambda_T}\left(\beta_T\tilde{v}-\int_0^T R_\xi(1-u,\hat{\xi}_u)\,d\beta_u\right)\,\middle|\,\tilde{v}\right]=0.$$

Therefore, by (4.24) and the martingale property of $\frac{1}{\lambda_\Gamma}$, and, then, by (4.26), we have

$$\mathbb{E}[\Pi(X^*,P^*)_T\mid\tilde{v}] = \mathbb{E}\left[\int_0^T \frac{1}{\lambda_{\Gamma_u}}(\tilde{v}-R_\xi(1-u,\hat{\xi}_u))\frac{\hat{\xi}_1-\hat{\xi}_u}{1-u}\,du\,\middle|\,\tilde{v}\right] \\ = \mathbb{E}\left[\frac{1}{\lambda_T}\int_0^T \Big(\tilde{v}-R_\xi(1-u,\hat{\xi}_u)\Big)\frac{\hat{\xi}_1-\hat{\xi}_u}{1-u}\,du\,\middle|\,\tilde{v}\right] \\ = \mathbb{E}\left[\frac{1}{\lambda_T}(R^c(0,\tilde{v})+R(1,0))\,\middle|\,\tilde{v}\right] = \frac{R^c(0,\tilde{v})+R(1,0)}{\lambda_0}$$

Therefore, since it attains the upper bound (4.21), X^* is an optimal strategy for the insider.

4.2.5. *Rationality of $\boldsymbol{P}^*(Y^*)$.* To show that P^* is a rational pricing rule we use (4.17) of Lemma 4.2 to conclude that $[\hat{Y}^*, W] = 0$. Therefore,

$$\boldsymbol{\xi}^*(Y^*) = \xi^* \text{ and } \boldsymbol{P}^*(Y^*)_t = R_\xi(\Sigma_t, \xi_t^*),$$

which reveals that $\boldsymbol{P}^*(Y^*)$ is a time-changed (by $1-\Sigma_t$) version of the martingale $R_\xi(1-u, \hat{\xi}_u)$. Hence $\boldsymbol{P}^*(Y^*)$ is a martingale itself with respect to $\mathcal{F}^{X^*,W} = \mathcal{F}^{\xi^*,W} = \mathcal{F}^{m*}$, and the rationality condition (2.3) follows from the fact that

$$\boldsymbol{P}^*(Y^*)_T = R_\xi(0, \xi_T^*) = h(h^{-1}(\tilde{v})) = \tilde{v}.$$

Proof of Proposition 2.2. For $\hat{\sigma} > 0$, let $u(t, \xi, \hat{\sigma})$ be the solution to the heat equation

$$u_t(t,\xi,\hat{\sigma}) + \frac{\hat{\sigma}^2}{2} u_{\xi\xi}(t,\xi,\hat{\sigma}) = 0$$
$$u(T,\xi,\hat{\sigma}) = F(h(\xi))$$

which is smooth in all of its variables for $t < T$. We apply Ito's formula to $u\left(t, \xi_t^*, \sqrt{\frac{\Sigma_t}{T-t}}\right)$ to obtain that

$$\begin{aligned} du\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right) &= u_t\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)dt + \frac{\hat{\rho}^2\lambda_t^2\sigma_t^2}{2}u_{\xi\xi}\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)dt \\ &\quad + u_{\hat{\sigma}}\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)\frac{\partial\sqrt{\frac{\Sigma_t}{T-t}}}{\partial t}dt + u_\xi\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)d\xi_t^*. \end{aligned}$$

u admits the stochastic representation $u(t,\xi,\hat{\sigma}) := \mathbb{E}[F(\xi + \hat{\sigma}\sqrt{T-t}B_1)]$ and by differentiation we have that

$$u_{\xi\xi}\hat{\sigma}(T-t) = u_{\hat{\sigma}}.$$

Additionally, (2.10) implies that

$$2(T-t)\sqrt{\frac{\Sigma_t}{T-t}}\frac{\partial\sqrt{\frac{\Sigma_t}{T-t}}}{\partial t} = (T-t)\frac{\partial\frac{\Sigma_t}{T-t}}{\partial t} = -\frac{\Sigma_t\hat{\rho}^2\sigma_t^2}{G_t} + \frac{\Sigma_t}{T-t} = -\lambda_t^2\hat{\rho}^2\sigma_t^2 + \frac{\Sigma_t}{T-t}.$$

Thus, we obtain

$$\begin{aligned} du\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right) &= u_t\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)dt + \frac{\hat{\rho}^2\lambda_t^2\sigma_t^2}{2}u_{\xi\xi}\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)dt \\ &\quad + (T-t)u_{\xi\xi}\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)\sqrt{\frac{\Sigma_t}{T-t}}\frac{\partial\sqrt{\frac{\Sigma_t}{T-t}}}{\partial t}dt + u_\xi\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)d\xi_t^* \\ &= u_t\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)dt + \frac{1}{2}u_{\xi\xi}\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)\frac{\Sigma_t}{T-t}dt + u_\xi\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)d\xi_t^* \\ &= u_\xi\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)d\xi_t^*. \end{aligned}$$

The choice of $\alpha_t = u_\xi\left(t,\xi_t^*,\sqrt{\frac{\Sigma_t}{T-t}}\right)$ proves the claim.

□ □

References

Aase, K. K., Bjuland, T., and Øksendal, B. (2012). Partially informed noise traders. *Mathematics and Financial Economics*, 6(2):93–104.

Abi Jaber, E., Larsson, M., and Pulido, S. (2019). Affine Volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200.

Alfonsi, A. (2015). *Affine diffusions and related processes : simulation, theory and applications*. Springer.

Back, K. (1992). Insider trading in continuous time. *The Review of Financial Studies*, 5(3):387–409.

Back, K. (1993). Asymmetric information and options. *The Review of Financial Studies*, 6(3):435–472.

Back, K., Cocquemas, F., Ekren, I., and Lioui, A. (2020). Optimal transport and risk aversion in Kyle's model of informed trading. *arXiv preprint arXiv:2006.09518*.

Back, K., Cocquemas, F., Ekren, I., and Lioui, A. (2021). Optimal transport and risk aversion in Kyle's model of informed trading. *arXiv preprint arXiv:2006.09518*.

Back, K., Crotty, K., and Li, T. (2018). Identifying information asymmetry in securities markets. *The Review of Financial Studies*, 31(6):2277–2325.

Back, K. and Pedersen, H. (1998). Long-lived information and intraday patterns. *Journal of financial markets*, 1(3-4):385–402.

Bagehot, W. (1971). The only game in town. *Financial Analysts Journal*, 27(2):12–14.

Barger, W. and Donnelly, R. (2021). Insider trading with temporary price impact. *International Journal of Theoretical and Applied Finance*, 24(02):2150006.

Baruch, S. (2002). Insider trading and risk aversion. *Journal of Financial Markets*, 5(4):451–464.

Biagini, F., Hu, Y., Meyer-Brandis, T., and Øksendal, B. (2012). Insider trading equilibrium in a market with memory. *Mathematics and Financial Economics*, 6(3):229–247.

Bose, S. and Ekren, I. (2023). Kyle–Back models with risk aversion and non-Gaussian beliefs. *The Annals of Applied Probability*, 33(6A):4238–4271.

Bose, S. and Ekren, I. (2024). Multidimensional Kyle–Back model with a risk averse informed trader. *SIAM Journal on Financial Mathematics*, 15(1):93–120.

Caballe, J. and Krishnan, M. (1994). Imperfect competition in a multi-security market with risk neutrality. *Econometrica: Journal of the Econometric Society*, pages 695–704.

Caldentey, R. and Stacchetti, E. (2010). Insider trading with a random deadline. *Econometrica*, 78(1):245–283.

Campi, L., Cetin, U., and Danilova, A. (2011). Dynamic Markov bridges motivated by models of insider trading. *Stochastic Processes and their Applications*, 121(3):534–567.

Çetin, U. (2018). Financial equilibrium with asymmetric information and random horizon. *Finance and Stochastics*, 22(1):97–126.

Çetin, U. and Danilova, A. (2016). Markovian Nash equilibrium in financial markets with asymmetric information and related forward–backward systems. *The Annals of Applied Probability*, 26(4):1996–2029.

Cho, K.-H. (2003). Continuous auctions and insider trading: uniqueness and risk aversion. *Finance and Stochastics*, 7(1):47–71.

Choi, J. H., Larsen, K., and Seppi, D. J. (2019). Information and trading targets in a dynamic market equilibrium. *Journal of Financial Economics*, 132(3):22–49.

Collin-Dufresne, P. and Fos, V. (2016). Insider trading, stochastic liquidity, and equilibrium prices. *Econometrica*, 84(4):1441–1475.

Collin-Dufresne, P., Fos, V., and Muravyev, D. (2021). Informed trading in the stock market and option-price discovery. *Journal of Financial and Quantitative Analysis*, 56(6):1945–1984.

Corcuera, J. M. and Di Nunno, G. (2020). Path-dependent Kyle equilibrium model. *arXiv preprint arXiv:2006.06395*.

Corcuera, J. M., Di Nunno, G., and Fajardo, J. (2019). Kyle equilibrium under random price pressure. *Decisions in Economics and Finance*, 42(1):77–101.

El Euch, O. and Rosenbaum, M. (2019). The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38.

Garcia del Molino, L. C., Mastromatteo, I., Benzaquen, M., and Bouchaud, J.-P. (2020). The multivariate Kyle model: More is different. *SIAM Journal on Financial Mathematics*, 11(2):327–357.

Garsia, A. M., Rodemich, E., Rumsey, H., and Rosenblatt, M. (1970). A real variable lemma and the continuity of paths of some Gaussian processes. *Indiana University Mathematics Journal*, 20(6):565–578.

Glasserman, P. and Pirjol, D. (2023). W-shaped implied volatility curves and the Gaussian mixture model. *Quantitative Finance*, 23(4):557–577.

Karatzas, I. and Shreve, S. E. (1991). *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, second edition.

Kazamaki, N. (2006). *Continuous exponential martingales and BMO*. Springer.

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

Lasserre, G. (2004). Asymmetric information and imperfect competition in a continuous time multivariate security model. *Finance and Stochastics*, 8(2):285–309.

Subrahmanyam, A. (1991). Risk aversion, market liquidity, and price efficiency. *The Review of Financial Studies*, 4(3):417–441.

Ying, C. (2020). The pre-FOMC announcement drift and private information: Kyle meets macro-finance. *Available at SSRN 3644386*.

Zhang, J. (2017). *Backward Stochastic Differential Equations*, volume 84 of *Probability Theory and Stochastic Modeling*. Springer.

University of Michigan, Department of Mathematics, 530 Church Street Ann Arbor, MI 48109

Email address: iekren@umich.edu

Florida State University, Department of Mathematics, 1017 Academic Way, Tallahassee, FL 32306

Email address: bm18g@my.fsu.edu

The University of Texas at Austin, Department of Mathematics, 2515 Speedway, Austin, TX 78712

Email address: gordanz@math.utexas.edu