



# Markowitz portfolio selection for multivariate affine and quadratic Volterra models

Eduardo Abi Jaber, Enzo Miller, Huy  n Pham

## ► To cite this version:

Eduardo Abi Jaber, Enzo Miller, Huy  n Pham. Markowitz portfolio selection for multivariate affine and quadratic Volterra models. 2020. hal-02877569v3

**HAL Id: hal-02877569**

**<https://hal.science/hal-02877569v3>**

Preprint submitted on 29 Oct 2020 (v3), last revised 23 Dec 2020 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin  e au d  p  t et    la diffusion de documents scientifiques de niveau recherche, publi  s ou non,   manant des   tablissements d'enseignement et de recherche fran  ais ou   trangers, des laboratoires publics ou priv  s.

# Markowitz portfolio selection for multivariate affine and quadratic Volterra models

Eduardo ABI JABER\*      Enzo MILLER†      Huyên PHAM‡

October 29, 2020

## Abstract

This paper concerns portfolio selection with multiple assets under rough covariance matrix. We investigate the continuous-time Markowitz mean-variance problem for a multivariate class of affine and quadratic Volterra models. In this incomplete non-Markovian and non-semimartingale market framework with unbounded random coefficients, the optimal portfolio strategy is expressed by means of a Riccati backward stochastic differential equation (BSDE). In the case of affine Volterra models, we derive explicit solutions to this BSDE in terms of multi-dimensional Riccati-Volterra equations. This framework includes multivariate rough Heston models and extends the results of [Han and Wong \(2020a\)](#). In the quadratic case, we obtain new analytic formulae for the Riccati BSDE and we establish their link with infinite dimensional Riccati equations. This covers rough Stein-Stein and Wishart type covariance models. Numerical results on a two dimensional rough Stein-Stein model illustrate the impact of rough volatilities and stochastic correlations on the optimal Markowitz strategy. In particular for positively correlated assets, we find that the optimal strategy in our model is a ‘buy rough sell smooth’ one.

**Keywords:** Mean-variance portfolio theory; rough volatility; correlation matrices; multi-dimensional Volterra process; Riccati equations; non-Markovian Heston, Stein–Stein and Wishart models.

**MSC Classification:** 93E20, 60G22, 60H10.

---

\*Université Paris 1 Panthéon-Sorbonne, Centre d’Economie de la Sorbonne, 106, Boulevard de l’Hôpital, 75013 Paris, [eduardo.abi-jaber@univ-paris1.fr](mailto:eduardo.abi-jaber@univ-paris1.fr).

†Université de Paris and Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation (LPSM, UMR CNRS 8001), Building Sophie Germain, Avenue de France, 75013 Paris, [enzo.miller@polytechnique.org](mailto:enzo.miller@polytechnique.org)

‡Université de Paris and Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation (LPSM, UMR CNRS 8001), Building Sophie Germain, Avenue de France, 75013 Paris, [pham@lpsm.paris](mailto:pham@lpsm.paris)

# 1 Introduction

The [Markowitz \(1952\)](#) mean-variance portfolio selection problem is the cornerstone of modern portfolio allocation theory. Investment decisions rules are made according to a trade-off between return and risk, and the use of Markowitz efficient portfolio strategies in the financial industry has become quite popular mainly due to its natural and intuitive formulation. A vast volume of research has been devoted over the last decades to extend Markowitz problem from static to continuous-time setting, first in Black-Scholes and complete markets ([Zhou and Li \(2000\)](#)), and then to consider more general frameworks with random coefficients and multiple assets, see e.g. [Lim \(2004\)](#), [Chiu and Wong \(2014\)](#), or more recently [Ismail and Pham \(2019\)](#) for taking into account model uncertainty on the assets correlation.

In the direction of more realistic modeling of asset prices, it is now well-established that volatility is rough ([Gatheral et al., 2018](#)), modeled by fractional Brownian motion with small Hurst parameter, which captures empirical facts of times series of realized volatility and key features of implied volatility surface, see [Alòs et al. \(2007\)](#); [Fukasawa \(2011\)](#). Subsequently, an important literature has focused on option pricing and asymptotics in rough volatility models. In comparison, the research on portfolio optimization in fractional and rough models is still little developed but has gained an increasing attention with the recent papers of [Fouque and Hu \(2018\)](#); [Bäuerle and Desmettre \(2020\)](#); [Han and Wong \(2020b\)](#), which consider fractional Ornstein-Uhlenbeck and Heston stochastic volatility models for power utility function criterion, and the work by [Han and Wong \(2020a\)](#) where the authors study the Markowitz problem in a Volterra Heston model, which covers the rough Heston model of [El Euch and Rosenbaum \(2018\)](#).

Most of the developments in rough volatility literature for asset modeling, option pricing or portfolio selection have been carried out in the mono-asset case. However, investment in multi-assets by taking into account the correlation risk is an importance feature in portfolio choice in financial markets, see [Buraschi et al. \(2010\)](#). Inspired by the recent papers [Abi Jaber \(2019c\)](#); [Abi Jaber et al. \(2019\)](#); [Cuchiero and Teichmann \(2019\)](#); [Rosenbaum and Thomas \(2019\)](#) that consider multivariate versions of rough Volterra volatility models, the basic goal of this paper is to enrich the literature on portfolio selection:

- (i) by introducing a class of multivariate Volterra models, which captures stylized facts of financial assets, namely various rough volatility patterns across assets, (possibly random) correlation between stocks, and leverage effects, i.e., correlation between a stock and its volatility.
- (ii) by keeping the model tractable for explicit computations of the optimal Markowitz portfolio strategy, which can be a quite challenging task in multivariate non-Markovian settings.

**Main contributions.** In this paper, we study the continuous-time Markowitz problem in a multivariate setting with a focus on two classes: (i) affine Volterra models as in [Abi Jaber](#)

et al. (2019) that include multivariate rough Heston models, (ii) quadratic Volterra models, which are new class of Volterra models, and embrace multivariate rough Stein-Stein models, and rough Wishart type covariance matrix models, in the spirit of [Abi Jaber \(2019c\)](#); [Cuchiero and Teichmann \(2019\)](#). We provide:

- **A generic verification result** for the corresponding mean-variance problem, which is formulated in an incomplete non-Markovian and non-semimartingale framework with unbounded random coefficients of the volatility and market price of risk, and under general filtration. This result expresses the solution to the Markowitz problem in terms of a Riccati backward stochastic differential equation (BSDE) by checking in particular the admissibility condition of the optimal control. We stress that related existing verification results in the literature (see [Lim \(2004\)](#), [Jeanblanc et al. \(2012\)](#), [Chiu and Wong \(2014\)](#), [Shen \(2015\)](#)) cannot be applied directly to our setting, and we shall discuss more in detail this point in Section 3.
- **Explicit solutions** to the Riccati BSDE in two concrete specifications of multivariate Volterra models exploiting the representation of the solution in terms of a Laplace transform:
  - (i) **the affine case:** the optimal Markowitz strategy is expressed in terms of multivariate Riccati-Volterra equations which naturally extends the one obtained in [Han and Wong \(2020a\)](#). We point out that the martingale distortion arguments used in [Han and Wong \(2020a\)](#) for the univariate Volterra Heston model, do not apply in higher dimensions, unless the correlation structure is highly degenerate.
  - (ii) **the quadratic case:** our major result is to derive analytic expressions for the optimal investment strategy by explicitly solving operator Riccati equations. This gives new explicit formulae for rough Stein-Stein and Wishart type covariance models. These analytic expressions can be efficiently implemented: the integral operators can be approximated by closed form expressions involving finite dimensional matrices and the underlying processes can be simulated by the celebrated Cholesky decomposition algorithm.
- **Numerical simulations** of the optimal Markowitz strategy in a two-asset rough Stein-Stein model to illustrate our results.<sup>1</sup> We depict the impact of some parameters onto the optimal investment when one asset is rough, and the other smooth (in the sense of the Hurst index of their volatility), and show in particular that for positively correlated assets, the optimal strategy is to “buy rough, sell smooth”, which is consistent with the empirical backtesting in [Glasserman and He \(2020\)](#).

**Outline of the paper.** The rest of the paper is organized as follows: Section 2 formulates the financial market model and the mean-variance problem in a multivariate setting with

---

<sup>1</sup>The code of our implementation can be found at the following [link](#).

random covariance matrix and market price of risk, and defines the general correlation structure. We state in Section 3 our generic verification result, which can be seen as unifying framework for previous results obtained in related literature. Section 4 is devoted to affine Volterra models where we derive an explicit expression for the optimal Markowitz strategy. In Section 5, we consider the class of quadratic Volterra models, and we show how to solve the infinite-dimensional Riccati equations that appear in the closed-form expressions of the optimal portfolio. Numerical illustrations on the behavior of the optimal investment in a two-asset rough Stein-Stein model are given in Section 6. Finally, the proof of the verification result and other technical lemmas are postponed to the Appendices.

**Notations.** Given a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and a filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  satisfying the usual conditions, we denote by

$$\begin{aligned} L_{\mathbb{F}}^{\infty}([0, T], \mathbb{R}^d) &= \left\{ Y : \Omega \times [0, T] \mapsto \mathbb{R}^d, \mathbb{F} - \text{prog. measurable and bounded a.s.} \right\} \\ L_{\mathbb{F}}^p([0, T], \mathbb{R}^d) &= \left\{ Y : \Omega \times [0, T] \mapsto \mathbb{R}^d, \mathbb{F} - \text{prog. measurable s.t. } \mathbb{E} \left[ \int_0^T |Y_s|^p ds \right] < \infty \right\} \\ S_{\mathbb{F}}^{\infty}([0, T], \mathbb{R}^d) &= \left\{ Y : \Omega \times [0, T] \mapsto \mathbb{R}^d, \mathbb{F} - \text{prog. measurable s.t. } \sup_{t \leq T} |Y_t(w)| < \infty \text{ a.s.} \right\}. \end{aligned}$$

Here  $|\cdot|$  denotes the Euclidian norm on  $\mathbb{R}^d$ . Classically, for  $p \in \llbracket 1, \infty \rrbracket$ , we define  $L_{\mathbb{F}}^{p,loc}([0, T], \mathbb{R}^d)$  as the set of progressive processes  $Y$  for which there exists a sequence of increasing stopping times  $\tau_n \uparrow \infty$  such that the stopped processes  $Y^{\tau_n}$  are in  $L_{\mathbb{F}}^p([0, T], \mathbb{R}^d)$  for every  $n \geq 1$ , and we recall that it consists of all progressive processes  $Y$  s.t.  $\int_0^T |Y_t|^p dt < \infty$ , a.s. To unclutter notation, we write  $L_{\mathbb{F}}^{p,loc}([0, T])$  instead of  $L_{\mathbb{F}}^{p,loc}([0, T], \mathbb{R}^d)$  when the context is clear.

## 2 Formulation of the problem

Fix  $T > 0$ ,  $d, N \in \mathbb{N}$ . We consider a financial market on  $[0, T]$  on some filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  with a non-risky asset  $S^0$

$$dS_t^0 = S_t^0 r(t) dt,$$

with a deterministic short rate  $r : \mathbb{R}_+ \rightarrow \mathbb{R}$ , and  $d$  risky assets with dynamics

$$dS_t = \text{diag}(S_t) \left[ (r(t) \mathbf{1}_d + \sigma_t \lambda_t) dt + \sigma_t dB_t \right], \quad (2.1)$$

driven by a  $d$ -dimensional Brownian motion  $B$ , with a  $d \times d$ -matrix valued stochastic volatility process  $\sigma$  and a  $\mathbb{R}^d$ -valued continuous stochastic process  $\lambda$ , called *market price of risk*. Here  $\mathbf{1}_d$  denotes the vector in  $\mathbb{R}^d$  with all components equal to 1. The market is

typically incomplete, in the sense that the dynamics of the continuous volatility process  $\sigma$  is driven by an  $N$ -dimensional process  $W = (W^1, \dots, W^N)^\top$  defined by:

$$W_t^k = C_k^\top B_t + \sqrt{1 - C_k^\top C_k} B_t^{\perp, k}, \quad k = 1, \dots, N, \quad (2.2)$$

where  $C_k \in \mathbb{R}^d$  s.t.  $C_k^\top C_k \leq 1$ , and  $B^\perp = (B^{\perp, 1}, \dots, B^{\perp, N})^\top$  is an  $N$ -dimensional Brownian motion independent of  $B$ . Note that  $d\langle W^k \rangle_t = dt$  but  $W^k$  and  $W^j$  can be correlated, hence  $W$  is not necessarily a Brownian motion. Observe that processes  $\lambda$  and  $\sigma$  are  $\mathbb{F}$ -adapted, possibly unbounded, but not necessarily adapted to the filtration generated by  $W$ . We point out that  $\mathbb{F}$  may be strictly larger than the augmented filtration generated by  $B$  and  $B^\perp$  as we shall deal with weak solutions to stochastic Volterra equations.

**Remark 2.1.** *In our applications, we will be chiefly interested in the case where  $\lambda_t$  is linear in  $\sigma_t$ , and where the dynamics of the matrix-valued process  $\sigma$  is governed by a Volterra equation of the form*

$$\sigma_t = g_0(t) + \int_0^t \mu(t, s, \omega) ds + \int_0^t \chi(t, s, \omega) dW_s. \quad (2.3)$$

The class of models that we shall develop in Sections 4 and 5 includes in particular the case of Volterra Heston model when  $d = 1$  with  $\lambda_t = \theta \sigma_t$ , for some constant  $\theta$ , as studied in Han and Wong (2020a), and the case of Wishart process for the covariance matrix process  $V_t = \sigma_t \sigma_t^\top$ , as studied in Chiu and Wong (2014). The class of models that we will develop in Sections 4 and 5 includes in particular the case of

- (i) multivariate Volterra Heston models based on Volterra square-root processes, see [Abi Jaber et al. \(2019, Section 6\)](#), we refer to [Rosenbaum and Thomas \(2019\)](#) for a microstructural foundation. When  $d = 1$ , we recover the results of [Han and Wong \(2020a\)](#), which cover the case of the rough Heston model of [El Euch and Rosenbaum \(2019\)](#).
- (ii) multivariate Volterra Stein-Stein and Wishart type in the sense of [Abi Jaber \(2019c\)](#), where the instantaneous covariance is given by squares of Gaussians. Under the Markovian setting, we recover a similar structure as in the results of [Chiu and Wong \(2014\)](#).

**Mean-variance optimization problem.** Let  $\pi_t$  denote the vector of the amounts invested in the risky assets  $S$  at time  $t$  in a self-financing strategy and set  $\alpha = \sigma^\top \pi$ . Then, the dynamics of the wealth  $X^\alpha$  of the portfolio we seek to optimize is given by

$$dX_t^\alpha = (r(t)X_t^\alpha + \alpha_t^\top \lambda_t)dt + \alpha_t^\top dB_t, \quad t \geq 0, \quad X_0^\alpha = x_0 \in \mathbb{R}. \quad (2.4)$$

By a solution to (2.4), we mean an  $\mathbb{F}$ -adapted continuous process  $X^\alpha$  satisfying (2.4) on  $[0, T]$   $\mathbb{P}$ -a.s. and such that

$$\mathbb{E}\left[\sup_{t \leq T} |X_t^\alpha|^2\right] < \infty. \quad (2.5)$$

The set of admissible investment strategies is naturally defined by

$$\mathcal{A} = \{\alpha \in L_{\mathbb{F}}^{2,loc}([0, T], \mathbb{R}^d) \text{ such that (2.4) has a solution satisfying (2.5)}\}.$$

The Markowitz portfolio selection problem in continuous-time consists in solving the following constrained problem

$$V(m) := \inf_{\alpha \in \mathcal{A}} \{\text{Var}(X_T) : \text{s.t. } \mathbb{E}[X_T] = m\}. \quad (2.6)$$

given some expected return value  $m \in \mathbb{R}$ , where  $\text{Var}(X_T) = \mathbb{E}[(X_T - \mathbb{E}[X_T])^2]$  stands for the variance.

### 3 A generic verification result

In this section, we establish a generic verification result for the optimization problem (2.6) given the solution of a certain Riccati BSDE. We stress that our mean-variance problem deals with incomplete markets with unbounded random coefficients  $\sigma$  and  $\lambda$ , so that existing results cannot be applied directly to our setting: Lim (2004) presents a general methodology to solve the MV problem for the wealth process (2.4) in an incomplete market without assuming any particular dynamics on  $\sigma$  nor that the excess return is proportional to  $\sigma$ . However, a nondegeneracy assumption is made on  $\sigma\sigma^\top$ , see Lim (2004, Assumption (A.1)). The main verification result in Lim (2004, Proposition 3.3), based on a completion of squares argument, states that if a solution to a certain (nonlinear) Riccati BSDE exists, then the MV is solvable. The difficulty resides in proving the existence of solutions to such nonlinear BSDEs (see also Lim and Zhou (2002) for similar results in complete markets).

Here, we assume that the excess return is proportional to  $\sigma$  (instead of the nondegeneracy condition) and state a verification result in terms of solutions of Riccati BSDEs (completion of squares, ie LQ problem with random coefficients). A verification result depending on the solution of a Riccati BSDE is also stated in Chiu and Wong (2014), but the admissibility of the optimal candidate control is not proved. We also mention the paper of Jeanblanc et al. (2012) where the authors adopt a BSDE approach for general semimartingales, but focusing on situations in which the existence of an optimal strategy is assumed. In our case, the existence of an admissible optimal control is obtained under a suitable exponential integrability assumption involving the market price of risk and the  $Z$  components of the BSDE, which extends the condition in Shen (2015).

Our main result of this section, Theorem 3.1 below, can be seen as unifying framework for the aforementioned results, refer to Table 1. For the sake of presentation, we postpone its proof to Appendix A.

	Random coef.	Unbounded coef.	degenerate $\sigma$	Incomplete market
Lim and Zhou (2002)	✓	✗	✗	✗
Lim (2004)	✓	✗	✗	✓
Shen (2015)	✓	✓	✗	✗
Theorem 3.1	✓	✓	✓	✓

Table 1: Comparison to existing verification results for mean-variance problems.

We define  $C \in \mathbb{R}^{N \times d}$  by

$$C = (C_1, \dots, C_N)^\top, \quad (3.1)$$

where we recall that the vectors  $C_i \in \mathbb{R}^d$  come from the correlation structure (2.2). We will use the matrix norm  $|A| = \text{tr}(A^\top A)$  in the subsequent theorem.

**Theorem 3.1.** *Assume that there exists a solution triplet  $(\Gamma, Z^1, Z^2) \in \mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R}) \times L_{\mathbb{F}}^{2,loc}([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^{2,loc}([0, T], \mathbb{R}^N)$  to the Riccati BSDE*

$$\begin{cases} d\Gamma_t &= \Gamma_t \left[ (-2r(t) + |\lambda_t + Z_t^1 + CZ_t^2|^2) dt + (Z_t^1)^\top dB_t + (Z_t^2)^\top dW_t \right], \\ \Gamma_T &= 1, \end{cases} \quad (3.2)$$

such that

(H1)  $0 < \Gamma_0 < e^{2 \int_0^T r(s) ds}$ , and  $\Gamma_t > 0$ , for all  $t \leq T$ ,

(H2)

$$\mathbb{E} \left[ \exp \left( a(p) \int_0^T (|\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2) ds \right) \right] < \infty, \quad (3.3)$$

for some  $p > 2$  and a constant  $a(p)$  given by

$$a(p) = \max \left[ p(3 + |C|), 3(8p^2 - 2p)(1 + |C|^2) \right]. \quad (3.4)$$

Then, the optimal investment strategy for the Markowitz problem (2.6) is given by the admissible control

$$\alpha_t^* = -(\lambda_t + Z_t^1 + CZ_t^2)(X_t^{\alpha^*} - \xi^* e^{-\int_t^T r(s) ds}), \quad (3.5)$$



where

$$\xi^* = \frac{m - \Gamma_0 e^{-\int_0^T r(t)dt} x_0}{1 - \Gamma_0 e^{-2\int_0^T r(t)dt}}. \quad (3.6)$$

Furthermore, the value of (2.6) for the optimal wealth process  $X^* = X^{\alpha^*}$  is

$$V(m) = \text{Var}(X_T^*) = \Gamma_0 \frac{|x_0 - m e^{-\int_0^T r(t)dt}|^2}{1 - \Gamma_0 e^{-2\int_0^T r(t)dt}}. \quad (3.7)$$

*Proof.* We refer to Appendix A. □

**Remark 3.2.** By setting  $\tilde{Z}_t^i = \Gamma_t Z_t^i$ ,  $i = 1, 2$ , the BSDE (3.2) agrees with the one in Chiu and Wong (2014, Theorem 3.1):

$$d\Gamma_t = \Gamma_t \left[ \left( -2r(t) + \left| \lambda_t + \frac{\tilde{Z}_t^1 + C\tilde{Z}_t^2}{\Gamma_t} \right|^2 \right) dt + (\tilde{Z}_t^1)^\top dB_t + (\tilde{Z}_t^2)^\top dW_t, \right.$$

and justifies the terminology Riccati BSDE.

In the sequel, we will provide concrete specifications of multivariate stochastic Volterra models for which the solution to the non-linear Riccati BSDE (3.2) can be computed in closed and semi-closed forms, while satisfying conditions (H1) and (H2). The key idea is to observe that, first, if such solution exists, then, it admits the following representation as a Laplace transform:

$$\Gamma_t = \mathbb{E} \left[ \exp \left( \int_t^T (2r(s) - |\lambda_s + Z_s^1 + CZ_s^2|^2) ds \right) \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T.$$

In the special case where  $\lambda$  is deterministic, then the solution to (3.2) trivially exists with  $Z^1 = Z^2 = 0$ , and condition (H1) and (H2) are obviously satisfied when  $\lambda$  is nonzero and bounded. In the general case where  $\lambda$  is an (unbounded) stochastic process, the admissibility of the optimal control is obtained under finiteness of a certain exponential moment of the solution triplet  $(\Gamma, Z^1, Z^2)$  and the risk premium  $\lambda$  as precised in (H2). Such estimate is crucial to deal with the unbounded random coefficients in (2.4), see for instance Han and Wong (2020a); Shen et al. (2014); Shen (2015) where similar conditions appear. If the coefficients are bounded, such condition is not needed, see Lim (2004, Lemma 3.1).

Our main interest is to find specific dynamics for the volatility  $\sigma$  and for the market price of risk  $\lambda$  such that the Laplace transform can be computed in (semi)-explicit form. We shall consider models as mentioned in Remark 2.1, where all the randomness in  $\lambda$  comes from the process  $W$  driving  $\sigma$ , and for which we naturally expect that  $Z^1 = 0$ . We solve more specifically this problem for two classes of models:

- (i) Multivariate affine Volterra models of Heston type in Section 4. This extends the results of Han and Wong (2020a) to the multi dimensional case and provides semi-closed formulas.
- (ii) Multivariate quadratic Volterra models of Stein-Stein and Wishart type in Section 5 for which we derive new closed-form solutions.

## 4 Multivariate affine Volterra models

We let  $K = \text{diag}(K_1, \dots, K_d)$  be diagonal with scalar kernels  $K_i \in L^2([0, T], \mathbb{R})$  on the diagonal,  $\nu = \text{diag}(\nu_1, \dots, \nu_d)$  and  $D \in \mathbb{R}^{d \times d}$  such that

$$D_{ij} \geq 0, \quad i \neq j.$$

We assume that  $\sigma$  in (2.3) is given by  $\sigma = \sqrt{\text{diag}(V)}$ , where  $V = (V^1, \dots, V^d)^\top$  is the following  $\mathbb{R}_+^d$ -valued Volterra square-root process

$$V_t = g_0(t) + \int_0^t K(t-s)DV_s ds + \int_0^t K(t-s)\nu\sqrt{\text{diag}(V_s)}dW_s. \quad (4.1)$$

Here  $g_0 : \mathbb{R}_+ \rightarrow \mathbb{R}_+^d$ ,  $W$  is a  $d$ -dimensional Brownian motion and the correlation structure with  $B$  is given by

$$W^i = \rho_i B^i + \sqrt{1 - \rho_i^2} B^{\perp, i}, \quad i = 1, \dots, d, \quad (4.2)$$

for some  $(\rho_1, \dots, \rho_d) \in [-1, 1]^d$ . This corresponds to a particular case of the correlation structure in (2.2) with  $N = d$ , and  $C_i = (0, \dots, \rho_i, \dots, 0)^\top$ . Furthermore, the risk premium is assumed to be in the form  $\lambda = (\theta_1 \sqrt{V^1}, \dots, \theta_d \sqrt{V^d})^\top$ , for some  $\theta_i \geq 0$ , so that the dynamics for the stock prices (2.1) reads

$$dS_t^i = S_t^i (r(t) + \theta_i V_t^i) dt + S_t^i \sqrt{V_t^i} dB_t^i, \quad i = 1, \dots, d. \quad (4.3)$$

We assume that there exists a continuous  $\mathbb{R}_+^{2d}$ -valued weak solution  $(V, S)$  to (4.1)-(4.3) on some filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F})_{t \geq 0}, \mathbb{P})$  such that

$$\sup_{t \leq T} \mathbb{E}[|V_t|^p] < \infty, \quad p \geq 1. \quad (4.4)$$

For instance, weak existence of  $V$  such that (4.4) holds is established under suitable assumptions on the kernel  $K$  and specifications  $g_0$  as shown in the following remark. The existence of  $S$  readily follows from that of  $V$ .

**Remark 4.1.** Assume that, for each  $i = 1, \dots, d$ ,  $K_i$  is completely monotone on  $(0, \infty)$ <sup>2</sup>, and that there exists  $\gamma_i \in (0, 2]$  and  $k_i > 0$  such that

$$\int_0^h K_i^2(t)dt + \int_0^T (K_i(t+h) - K_i(t))^2 dt \leq k_i h^{\gamma_i}, \quad h > 0. \quad (4.5)$$

This covers, for instance, constant non-negative kernels, fractional kernels of the form  $t^{H-1/2}/\Gamma(H+1/2)$  with  $H \in (0, \frac{1}{2}]$ , and exponentially decaying kernels  $e^{-\beta t}$  with  $\beta > 0$ . Moreover, sums and products of completely monotone functions are completely monotone, refer to [Abi Jaber et al. \(2019\)](#) for more details.

- If  $g_0(t) = V_0 + \int_0^t K(t-s)b^0 ds$ , for some  $V_0, b^0 \in \mathbb{R}_+^d$ , then [Abi Jaber et al. \(2019, Theorem 6.1\)](#) ensures the existence of  $V$  such that (4.4) holds,
- In [Abi Jaber and El Euch \(2019a\)](#), the existence is obtained for more general input curves  $g_0$  for the case  $d = 1$ , the extension to the multi-dimensional setting is straightforward.

Exploiting the affine structure of (4.1)-(4.3), see [Abi Jaber et al. \(2019\)](#), we provide an explicit solution to the Riccati BSDE (3.2) in terms of the Riccati-Volterra equation

$$\psi^i(t) = \int_0^t K_i(t-s)F_i(\psi(s))ds, \quad (4.6)$$

$$F_i(\psi) = -\theta_i^2 - 2\theta_i\rho_i\nu_i\psi^i + (D^\top \psi)_i + \frac{\nu_i^2}{2}(1 - 2\rho_i^2)(\psi^i)^2, \quad i = 1, \dots, d, \quad (4.7)$$

and the  $\mathbb{R}^d$ -valued process

$$g_t(s) = g_0(s) + \int_0^t K(s-u)DV_u du + \int_0^t K(s-u)\nu\sqrt{\text{diag}(V_u)}dW_u, \quad s \geq t. \quad (4.8)$$

One notes that for each,  $s \leq T$ ,  $(g_t(s))_{t \leq s}$  is the adjusted forward process

$$g_t(s) = \mathbb{E}\left[V_s - \int_t^s K(s-u)DV_u du \mid \mathcal{F}_t\right].$$

**Lemma 4.2.** Assume that there exists a solution  $\psi \in C([0, T], \mathbb{R}^d)$  to the Riccati-Volterra equation (4.6)-(4.7). Let  $(\Gamma, Z^1, Z^2)$  be defined as

$$\begin{cases} \Gamma_t &= \exp\left(2 \int_t^T r(s)ds + \sum_{i=1}^d \int_t^T F_i(\psi(T-s))g_t^i(s)ds\right), \\ Z_t^1 &= 0, \\ Z_t^{2,i} &= \psi^i(T-t)\nu_i\sqrt{V_t^i}, \quad i = 1, \dots, d, \quad 0 \leq t \leq T, \end{cases} \quad (4.9)$$

---

<sup>2</sup>A function  $f$  is completely monotone on  $(0, \infty)$  if it is infinitely differentiable on  $(0, \infty)$  such that  $(-1)^n f^{(n)}(t) \geq 0$ , for all  $n \geq 1$  and  $t > 0$ .

where  $g = (g^1, \dots, g^d)^\top$  is given by (4.8). Then,  $(\Gamma, Z^1, Z^2)$  is a  $\mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R}) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d)$ -valued solution to (3.2).

*Proof.* We first observe that the correlation structure (4.2) implies that  $C$  in (3.1) is given by  $C = \text{diag}(\rho_1, \dots, \rho_d)$ . Set

$$G_t = 2 \int_t^T r(s) ds + \sum_{i=1}^d \int_t^T F_i(\psi(T-s)) g_t^i(s) ds, \quad t \leq T.$$

Then,  $\Gamma = \exp(G)$  and

$$d\Gamma_t = \Gamma_t \left( dG_t + \frac{1}{2} d\langle G \rangle_t \right). \quad (4.10)$$

Using (4.8), and by stochastic Fubini's theorem, see Veraar (2012, Theorem 2.2), the dynamics of  $G$  reads as

$$\begin{aligned} dG_t &= \left( -2r(t) - \sum_{i=1}^d F_i(\psi(T-t)) V_t^i + \sum_{j=1}^d \int_t^T F_j(\psi(T-s)) K_j(s-t) ds \sum_{i=1}^d D_{ji} V_t^i \right) dt \\ &\quad + \sum_{i=1}^d \int_t^T F_i(\psi(T-s)) K_i(s-t) ds \nu_i \sqrt{V_t^i} dW_t^i \\ &= \left( -2r(t) - \sum_{i=1}^d F_i(\psi(T-t)) V_t^i + \sum_{j=1}^d \psi^j(T-t) \sum_{i=1}^d D_{ji} V_t^i \right) dt \\ &\quad + \sum_{i=1}^d \psi^i(T-t) \nu_i \sqrt{V_t^i} dW_t^i, \end{aligned}$$

where we changed variables and used the Riccati–Volterra equation (4.6) for  $\psi$  for the last equality. This yields that the dynamics of  $\Gamma$  in (4.10) is given by

$$\begin{aligned} d\Gamma_t &= \Gamma_t \left( -2r(t) + \sum_{i=1}^d V_t^i \left( -F_i(\psi(T-t)) + \sum_{j=1}^d D_{ji} \psi^j(T-t) + \frac{\nu_i^2}{2} (\psi^i(T-t))^2 \right) \right) dt \\ &\quad + \Gamma_t \sum_{i=1}^d \psi^i(T-t) \nu_i \sqrt{V_t^i} dW_t^i \\ &= \Gamma_t \left[ \left( -2r(t) + \sum_{i=1}^d V_t^i (\theta_i + \rho_i \nu_i \psi^i(T-t))^2 \right) dt + (Z_t^2)^\top dW_t \right], \end{aligned} \quad (4.11)$$

where we used (4.7) for the last identity. Finally, observing that

$$|\lambda_t + Z_t^1 + C Z_t^2|^2 = \sum_{i=1}^d (\theta_i + \rho_i \nu_i \psi^i(T-t))^2 V_t^i,$$

together with  $\Gamma_T = 1$ , we get that  $(\Gamma, Z^1, Z^2)$  as defined in (4.9) solves the BSDE (3.2).

It remains to show that  $(\Gamma, Z^1, Z^2) \in \mathbb{S}_{\mathbb{F}}^{\infty}([0, T], \mathbb{R}) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d)$ . For this, define the process

$$M_t = \Gamma_t \exp \left( \int_t^T (-2r(s) + \sum_{i=1}^d V_s^i (\theta_i + \rho_i \nu_i \psi^i(T-s))^2) ds \right), \quad t \leq T.$$

An application of Itô's formula combined with the dynamics (4.11) shows that  $dM_t = M_t (Z_t^2)^\top dW_t$ , and so  $M$  is a local martingale of the form

$$M_t = \mathcal{E} \left( \int_t^T \sum_{i=1}^d \psi^i(T-s) \nu_i \sqrt{V_s^i} dW_s^i \right).$$

Since  $\psi$  is continuous, it is bounded so that a straightforward adaptation of [Abi Jaber et al. \(2019, Lemma 7.3\)](#) to the multi-dimensional setting, recall (4.4), yields that  $M$  is a true martingale. Since  $M_T = 1$ , writing  $\mathbb{E}[M_T | \mathcal{F}_t] = M_t$ , we obtain

$$\Gamma_t = \mathbb{E} \left[ \exp \left( \int_t^T (2r(s) - \sum_{i=1}^d V_s^i (\theta_i + \rho_i \nu_i \psi^i(T-s))^2) ds \right) \mid \mathcal{F}_t \right], \quad t \leq T, \quad (4.12)$$

which ensures that  $0 < \Gamma_t \leq e^{2 \int_t^T r(s) ds}$ ,  $\mathbb{P} - a.s.$ , since  $V \in \mathbb{R}_+^d$ . As for  $Z^2$ , it is clear that it belongs to  $L_{\mathbb{F}}^2([0, T], \mathbb{R}^d)$  since  $\Gamma$  and  $\psi$  are bounded and  $\mathbb{E} \left[ \int_0^T \sum_{i=1}^d V_s^i ds \right] < \infty$  by (4.4).  $\square$

The following remark makes precise the existence of a continuous solution to the Riccati-Volterra equation (4.6)-(4.7).

**Remark 4.3.** Assume that  $K$  satisfies the assumptions of Remark 4.1.

- If  $1 - 2\rho_i^2 \geq 0$ , then [Abi Jaber et al. \(2019, Lemma 6.3\)](#) provides the existence of a unique solution  $\psi \in L^2([0, T], \mathbb{R}_+^d)$ . Continuity of such solution can then be easily established, since as opposed to [Abi Jaber et al. \(2019, Lemma 6.3\)](#), (4.6) starts from 0.
- If  $d = 1$  and  $1 - 2\rho_1^2 < 0$ , [Han and Wong \(2020a, Lemma A.4\)](#) establishes the existence of a continuous solution  $\psi$ .

Using Theorem 3.1, we can now explicitly solve the Markowitz problem (2.6) in the multivariate Volterra Heston model (4.1)-(4.2)-(4.3). The next theorem extends ([Han and Wong, 2020a, Theorem 4.2](#)) to the multivariate case. Notice that the martingale distortion argument in this cited paper is specific to the dimension  $d = 1$ , and here, instead, we rely on the generic verification result in Theorem 3.1.

**Theorem 4.4.** Assume that there exists a solution  $\psi \in C([0, T], \mathbb{R}^d)$  to the Riccati-Volterra equation (4.6)-(4.7) such that

$$\max_{1 \leq i \leq d} \max_{t \in [0, T]} (\theta_i^2 + \nu_i^2 \psi^i(t)^2) \leq \frac{a}{a(p)}, \quad \text{for some } p > 2, \quad (4.13)$$

where  $a(p)$  is given by (3.4) and the constant  $a > 0$  is such that  $\mathbb{E} \left[ \exp \left( a \int_0^T \sum_{i=1}^d V_s^i ds \right) \right] < \infty$ . Assume that  $g_0^i(0) > 0$  for some  $i \leq d$ . Then, the optimal investment strategy for the maximization problem (2.6) in the multivariate Volterra Heston model (4.1)-(4.2)-(4.3) is given by the admissible control

$$\alpha_t^{*i} = -(\theta_i + \rho_i \nu_i \psi^i(T-t)) \sqrt{V_t^i} \left( X_t^{\alpha^*} - \xi^* e^{-\int_t^T r(s) ds} \right), \quad 1 \leq i \leq d, \quad (4.14)$$

where  $\xi^*$  is defined as in (3.6), the wealth process  $X^* = X^{\alpha^*}$  by (2.4) with  $\lambda = (\theta_1 \sqrt{V^1}, \dots, \theta_d \sqrt{V^d})^\top$ , and the optimal value is given by (3.7) with  $\Gamma_0$  as in (4.12).

*Proof.* First note that under the specification (4.9), the candidate for the optimal feedback control defined in (3.5) takes the form

$$\begin{aligned} \alpha_t^* &= -(\lambda_t + Z_t^1 + C Z_t^2) (X_t^* - \xi^* e^{-\int_t^T r(s) ds}) \\ &= \left( -(\theta_i + \rho_i \nu_i \psi^i(T-t)) \sqrt{V_t^i} (X_t^* - \xi^* e^{-\int_t^T r(s) ds}) \right)_{1 \leq i \leq d}. \end{aligned}$$

It then suffices to check that the assumptions of Theorem 3.1 are verified to ensure that such  $\alpha^*$  is optimal and to get that (3.7) is the optimal value. The existence of a solution triplet  $(\Gamma, Z^1, Z^2) \in \mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R}) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^N)$  to the stochastic backward Riccati equation (3.2) is ensured by Lemma 4.2. In addition, (4.12) implies that  $\Gamma_0 < e^{2 \int_0^T r(s) ds}$  since  $g_0^i(0) > 0$  for some  $i \leq d$  by assumption and  $V^i$  is continuous. Thus condition (H1) of Theorem 3.1 is verified. As for condition (H2) of Theorem 3.1, note that

$$a(p) \left( |\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2 \right) = a(p) \sum_{i=1}^d V_s^i (\theta_i^2 + \nu_i^2 \psi^i(t)^2) \leq a \sum_{i=1}^d V_s^i,$$

which implies that  $\mathbb{E} \left[ \exp \left( a(p) \int_0^T \left( |\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2 \right) ds \right) \right] < \infty$  and ends the proof.  $\square$

**Remark 4.5.** Condition (4.13) concerns the risk premium constants  $(\theta_1, \dots, \theta_d)$ . For  $a > 0$ , a sufficient condition ensuring  $\mathbb{E} \left[ \exp \left( a \int_0^T \sum_{i=1}^d V_s^i ds \right) \right] < \infty$  is the existence of a continuous solution  $\tilde{\psi}$  to the Riccati-Volterra

$$\tilde{\psi}^i(t) = \int_0^t K_i(t-s) \left( a + (D\tilde{\psi}(s))_i + \frac{\nu_i^2}{2} \tilde{\psi}^i(s) \right) ds,$$

see [Abi Jaber et al. \(2019, Theorem 4.3\)](#). In the one dimensional case  $d = 1$ , such existence is established in [Han and Wong \(2020a, Lemma A.2\)](#) for the case where  $g_0(t) = V_0 + \kappa \int_0^t K(t-s)\phi ds$ ,  $\phi \geq 0$ ,  $D = -\kappa$  and  $a < \frac{\kappa^2}{2\nu^2}$ .

**Remark 4.6.** Note that in the one dimensional case the condition (4.13) can be made more explicit by bounding  $\psi$  with respect to  $\theta$ . Indeed since  $-\theta^2 < 0$  we get from [Abi Jaber and El Euch \(2019b, Theorem C.1\)](#) that  $\psi$  is non-positive. Furthermore, the fact that  $\psi$  is solution to the following linear Volterra equation

$$\chi(t) = \int_0^t K(t-s) \left( -\theta^2 + ((D - 2\theta\rho\nu) + \frac{\nu^2}{2}(1 - 2\rho^2)\psi(s))\chi(s) \right) ds,$$

leads to, see [Abi Jaber and El Euch \(2019b, Corollary C.4\)](#),

$$\sup_{t \in [0, T]} |\psi_t| \leq |\theta|^2 \int_0^T R_D(s) ds,$$

where  $R_D$  is the resolvent of  $KD$ . Consequently, a sufficient condition on  $\theta$  to ensure (4.13) would be

$$\theta^2 \left( 1 + (\theta\nu)^2 \int_0^T R_D(s) ds \right) \leq \frac{a}{a(p)}.$$

**Remark 4.7.** In order to numerically implement the optimal strategy (4.14), one needs to simulate the possibly non-Markovian process  $V$  and to discretize the Riccati-Volterra equation for  $\psi$ . [Abi Jaber \(2019a\)](#); [Abi Jaber and El Euch \(2019b\)](#) develop a Taylor-made approximating procedure for the stochastic Volterra equation (4.1) (resp. the Riccati-Volterra equation (4.6)), using finite-dimensional Markovian semimartingales (resp. finite-dimensional Riccati ODE's). An illustration of such procedure on the mean-variance problem in the univariate Volterra Heston model for the fractional kernel is given in [Han and Wong \(2020a, Section 5\)](#).

## 5 Multivariate quadratic Volterra models

Before we introduce the class of multivariate quadratic Volterra models, we need to define and introduce some notations on integral operators.

### 5.1 Integral operators

Fix  $T > 0$ . We denote by  $\langle \cdot, \cdot \rangle_{L^2}$  the inner product on  $L^2([0, T], \mathbb{R}^N)$  that is

$$\langle f, g \rangle_{L^2} = \int_0^T f(s)^\top g(s) ds, \quad f, g \in L^2([0, T], \mathbb{R}^N).$$

We define  $L^2([0, T]^2, \mathbb{R}^{N \times N})$  to be the space of measurable kernels  $K : [0, T]^2 \rightarrow \mathbb{R}^{N \times N}$  such that

$$\int_0^T \int_0^T |K(t, s)|^2 dt ds < \infty.$$

For any  $K, L \in L^2([0, T]^2, \mathbb{R}^{N \times N})$  we define the  $\star$ -product by

$$(K \star L)(s, u) = \int_0^T K(s, z) L(z, u) dz, \quad (s, u) \in [0, T]^2,$$

which is well-defined in  $L^2([0, T]^2, \mathbb{R}^{N \times N})$  due to the Cauchy-Schwarz inequality. For any kernel  $K \in L^2([0, T]^2, \mathbb{R}^{N \times N})$ , we denote by  $\mathbf{K}$  the integral operator induced by the kernel  $K$  that is

$$(\mathbf{K}g)(s) = \int_0^T K(s, u) g(u) du, \quad g \in L^2([0, T], \mathbb{R}^N).$$

$\mathbf{K}$  is a linear bounded operator from  $L^2([0, T], \mathbb{R}^N)$  into itself. If  $\mathbf{K}$  and  $\mathbf{L}$  are two integral operators induced by the kernels  $K$  and  $L$  in  $L^2([0, T]^2, \mathbb{R}^{N \times N})$ , then  $\mathbf{K}\mathbf{L}$  is the integral operator induced by the kernel  $K \star L$ .

We denote by  $K^*$  the adjoint kernel of  $K$  for  $\langle \cdot, \cdot \rangle_{L^2}$ , that is

$$K^*(s, u) = K(u, s)^\top, \quad (s, u) \in [0, T]^2,$$

and by  $\mathbf{K}^*$  the corresponding adjoint integral operator.

**Definition 5.1.** A kernel  $K \in L^2([0, T]^2, \mathbb{R}^{N \times N})$  is symmetric nonnegative if  $K = K^*$  and

$$\int_0^T \int_0^T f(s)^\top K(s, u) f(u) du ds \geq 0, \quad \forall f \in L^2([0, T], \mathbb{R}^N).$$

In this case, the integral operator  $\mathbf{K}$  is said to be symmetric nonnegative and  $\mathbf{K} = \mathbf{K}^*$  and  $\langle f, \mathbf{K}f \rangle_{L^2} \geq 0$ .  $\mathbf{K}$  is said to be symmetric nonpositive, if  $(-\mathbf{K})$  is symmetric nonnegative.

We recall the definition of Volterra kernels of continuous and bounded type in the terminology of [Gripenberg et al. \(1990, Definitions 9.2.1, 9.5.1 and 9.5.2\)](#).

**Definition 5.2.** A kernel  $K : \mathbb{R}_+^2 \rightarrow \mathbb{R}^{N \times N}$  is a Volterra kernel of continuous and bounded type in  $L^2$  if  $K(t, s) = 0$  whenever  $s > t$  and

$$\sup_{t \in [0, T]} \int_0^T |K(t, s)|^2 ds < \infty, \quad \text{and} \quad \lim_{h \rightarrow 0} \int_0^T |K(u+h, s) - K(u, s)|^2 ds = 0, \quad u \leq T. \quad (5.1)$$

Any convolution kernel of the form  $K(t, s) = k(t-s)\mathbf{1}_{s \leq t}$  with  $k \in L^2([0, T], \mathbb{R}^{N \times N})$  satisfies (5.1), we refer to [Abi Jaber \(2019c, Example 3.1\)](#) for additional examples. Note that  $(s, t) \mapsto K(s, t)$  is not necessarily continuous nor bounded.

For completeness, we collect in [Appendix B.1](#) below standard results for integral operators and their resolvents.



## 5.2 The model

In this section, we assume that the components of the stochastic volatility matrix  $\sigma$  in (2.1) are given by  $\sigma^{ij} = \gamma_{ij}^\top Y$ , where  $\gamma_{ij} \in \mathbb{R}^N$  and  $Y = (Y^1, \dots, Y^N)^\top$  is the following  $N$ -dimensional Volterra Ornstein–Uhlenbeck process

$$Y_t = g_0(t) + \int_0^t K(t, s) DY_s ds + \int_0^t K(t, s) \eta dW_s, \quad (5.2)$$

where  $D, \eta \in \mathbb{R}^{N \times N}$ ,  $g_0 : \mathbb{R}_+ \rightarrow \mathbb{R}^N$  is locally bounded,  $W$  is a  $N$ -dimensional process as in (2.2), i.e.,

$$W_t^k = C_k^\top B_t + \sqrt{1 - C_k^\top C_k} B_t^{\perp, k}, \quad (5.3)$$

where  $C_k \in \mathbb{R}^d$ , such that  $C_k^\top C_k \leq 1$ ,  $k = 1, \dots, N$ , and  $K : [0, T]^2 \rightarrow \mathbb{R}^{N \times N}$  is a Volterra kernel of continuous and bounded type in  $L^2$  as in Definition 5.2. We stress that the process  $W$  is not necessarily a  $N$ -dimensional Brownian motion due to the possible correlations.

Furthermore, the risk premium is assumed to be in the form

$$\lambda_t = \Theta Y_t, \quad t \leq T,$$

for some  $\Theta \in \mathbb{R}^{d \times N}$ , so that the dynamics for the stock prices (2.1) reads as

$$dS_t^i = S_t^i \left( r(t) + \sum_{k, \ell=1}^N \sum_{j=1}^d \gamma_{ij}^\ell \Theta^{jk} Y_t^\ell Y_t^k \right) dt + S_t^i \sum_{j=1}^d \gamma_{ij}^\top Y_t dB_t^j, \quad i = 1, \dots, d. \quad (5.4)$$

The appellation quadratic reflects the quadratic dependence of the drift and the covariance matrix of  $\log S$  in  $Y$ . Such models nest as special cases the Volterra extensions of the celebrated Stein and Stein (1991) or Schöbel and Zhu (1999) model and certain Wishart models of Bru (1991) as shown in the following example.

**Example 5.3.** (i) *The multivariate Volterra Stein–Stein model:*

For  $N = d$ ,  $K = \text{diag}(K^1, \dots, K^d)$  and  $\gamma_{ij} = \beta_{ij} e_i$  with  $\beta_{ij} \in \mathbb{R}$  such that  $\sum_{j=1}^d \beta_{ij}^2 = 1$  and  $(e_1, \dots, e_d)$  the canonical basis of  $\mathbb{R}^d$ , we recover the multivariate Volterra Stein–Stein model defined by

$$\begin{cases} dS_t^i &= S_t^i \left( r(t) + \sum_{j,k=1}^d \beta_{ij} \Theta^{jk} Y_t^j Y_t^k \right) dt + S_t^i Y_t^i \sum_{j=1}^d \beta_{ij} dB_t^j, \\ Y_t^i &= g_0^i(t) + \int_0^t K^i(t, s) \sum_{j=1}^d D^{ij} Y_s^j ds + \int_0^t K^i(t, s) \eta^i dW_s^i, \end{cases} \quad i = 1, \dots, d,$$

and  $C_i = \rho_i(\beta_{i1}, \dots, \beta_{id})^\top$  to take into account the leverage effect. Recall that  $W$  is possibly correlated and is not necessarily a Brownian motion.

(ii) *The Volterra Wishart covariance model:*

Using the vectorization operator, which stacks the columns of a matrix one underneath another in a vector, see [Abi Jaber \(2019c, Section 3.1\)](#), one can recover the Volterra Wishart covariance model for  $N = d^2$ :

$$\begin{cases} dS_t &= \text{diag}(S_t)[r(t)\mathbf{1}_d dt + \tilde{Y}_t dB_t], \quad S_0 \in \mathbb{R}_+^d, \\ \tilde{Y}_t &= \tilde{g}_0(t) + \int_0^t \tilde{K}(t, s) DY_s ds + \int_0^t \tilde{K}(t, s) \eta dW_s, \end{cases}$$

with  $\tilde{g}_0 : [0, T] \rightarrow \mathbb{R}^{d \times d}$ , a suitable measurable kernel  $\tilde{K} : [0, T]^2 \rightarrow \mathbb{R}^{d \times d}$ , a  $d \times d$  Brownian motion  $W$  and

$$W^{ij} = \rho_{ij}^\top B + \sqrt{1 - \rho_{ij}^\top \rho_{ij}} B^{\perp, ij}, \quad i, j = 1, \dots, d,$$

for some  $\rho_{ij} \in \mathbb{R}^{d \times d}$  such that  $\rho_{ij}^\top \rho_{ij} \leq 1$ , for  $i, j = 1, \dots, d$ , where  $B^\perp$  is a  $d \times d$ -dimensional Brownian motion independent of  $B$ . Here the process  $\tilde{Y}$  is  $d \times d$ -matrix valued.

**Remark 5.4.** *Note that with (5.3), there are no restrictions on the correlations between  $Y^i$  and the stocks  $S^i$  in (5.2) and (5.4), in contrast with the correlation structure (4.1) of the multivariate Volterra Heston model. Moreover, the models in Example 5.3 allow us to deal with correlated stocks in contrast with the multivariate Heston model in (4.3) where no correlation between the driving Brownian motion of the assets  $S^i$  and  $S^j$  is allowed in order to keep the affine structure.*

Since  $K$  is a Volterra kernel of continuous and bounded type in  $L^2$ , there exists a progressively measurable  $\mathbb{R}^N \times \mathbb{R}_+^d$ -valued strong solution  $(Y, S)$  to (5.2) and (5.4) such that

$$\sup_{t \leq T} \mathbb{E}[|Y_t|^p] < \infty, \quad p \geq 1.$$

Indeed, the solution for (5.2) is given in the following closed form

$$Y_t = g_0(t) + \int_0^t R_D(t, s) g_0(s) ds + \int_0^t (K(t, s) + R_D(t, s)) \eta dW_s, \quad (5.5)$$

where  $R_D$  is the resolvent of  $KD$ , whose existence is ensured by Lemma B.2-(i) below, we refer to Appendix B.1 for more details on the resolvents. The existence of  $S$  readily follows from that of  $Y$  and is given as a stochastic exponential. In the sequel, we will assume that the solution  $Y$  is continuous. Additional conditions on  $K$ , in the spirit of (4.5), are needed to ensure the existence of continuous modification, by an application of the Kolmogorov-Chentsov continuity criterion, for instance, as shown in the following remark.

**Remark 5.5.** For  $s \leq t$  and  $p \geq 2$ , an application of Jensen and Burkholder-Davis-Gundy's inequalities yield

$$\begin{aligned} \mathbb{E}[|(Y_t - g_0(t)) - (Y_s - g_0(s))|^p] &\leq c \left(1 + \sup_{r \leq T} \mathbb{E}[|Y_s|^p]\right) \\ &\quad \times \left(\int_s^t |K(t, r)|^2 dr + \int_0^T |K(t, r) - K(s, r)|^2 dr\right)^{p/2}. \end{aligned}$$

This shows that  $(Y - g_0)$  admits a continuous modification, by the Kolmogorov-Chentsov continuity criterion, provided that

$$\int_s^t |K(t, r)|^2 dr + \int_0^T |K(t, r) - K(s, r)|^2 dr \leq c|t - s|^\gamma,$$

for some  $\gamma > 0$ .

### 5.3 The explicit solution

In this section, we provide an explicit solution for the Markowitz problem for quadratic Volterra models, and our main result is stated in Theorem 5.9 below.

Exploiting the quadratic structure of (5.2)-(5.4), see [Abi Jaber \(2019c\)](#), we provide an explicit solution to the Riccati BSDE in Lemma 5.7 below, in terms of the following family of linear operators  $(\Psi_t)_{0 \leq t \leq T}$  acting on  $L^2([0, T], \mathbb{R}^N)$ :

$$\Psi_t = -\left(\text{Id} - \hat{K}\right)^{-*} \Theta^\top \left(\text{Id} + 2\Theta \tilde{\Sigma}_t \Theta^\top\right)^{-1} \Theta \left(\text{Id} - \hat{K}\right)^{-1}, \quad 0 \leq t \leq T, \quad (5.6)$$

where  $\mathbf{F}^{-*} := (\mathbf{F}^{-1})^*$ , and  $\hat{K}$  is the integral operator induced by the kernel  $\hat{K} = K(D - 2\eta C^\top \Theta)$  and  $\tilde{\Sigma}_t$  the integral operator defined by

$$\tilde{\Sigma}_t = (\text{Id} - \hat{K})^{-1} \Sigma_t (\text{Id} - \hat{K})^{-*}, \quad t \in [0, T], \quad (5.7)$$

with  $\Sigma_t$  defined as the integral operator associated to the kernel

$$\Sigma_t(s, u) = \int_t^{s \wedge u} K(s, z) \eta(U - 2C^\top C) \eta^\top K(u, z)^\top dz, \quad t \in [0, T], \quad (5.8)$$

where  $U = \frac{d\langle W \rangle_t}{dt} = (1_{i=j} + 1_{i \neq j} (C_i)^\top C_j)_{1 \leq i, j \leq N}$ .

We start by deriving some first properties of  $t \mapsto \Psi_t$ , namely that it is well-defined, strongly differentiable and satisfies an operator Riccati equation under the following additional assumption on the kernel:

$$\sup_{t \leq T} \int_0^T |K(s, t)|^2 ds < \infty. \quad (5.9)$$

We recall that  $t \mapsto \Psi_t$  is said to be strongly differentiable at time  $t \geq 0$ , if there exists a bounded linear operator  $\dot{\Psi}_t$  from  $L^2([0, T], \mathbb{R}^N)$  into itself such that

$$\lim_{h \rightarrow 0} \frac{1}{h} \|\Psi_{t+h} - \Psi_t - h\dot{\Psi}_t\|_{\text{op}} = 0, \quad \text{where } \|\mathbf{G}\|_{\text{op}} = \sup_{f \in L^2([0, T], \mathbb{R}^N)} \frac{\|\mathbf{G}f\|_{L^2}}{\|f\|_{L^2}}.$$

**Lemma 5.6.** *Fix a kernel  $K$  as in Definition 5.2 satisfying (5.9). Assume that  $(U - 2C^\top C) \in \mathbb{S}_+^N$ . Then, for each  $t \leq T$ ,  $\Psi_t$  given by (5.6) is well-defined and is a bounded linear operator from  $L^2([0, T], \mathbb{R}^N)$  into itself. Furthermore,*

(i)  $(\Theta^\top \Theta \text{Id} + \Psi_t)$  is an integral operator induced by a kernel  $\psi_t(s, u)$  such that

$$\sup_{t \leq T} \int_{[0, T]^2} |\psi_t(s, u)|^2 ds du < \infty. \quad (5.10)$$

(ii) For any  $f \in L^2([0, T], \mathbb{R}^N)$ ,

$$(\Psi_t f 1_t)(t) = (-\Theta^\top \Theta \text{Id} + \hat{K}^* \Psi_t)(f 1_t)(t),$$

where  $1_t : s \mapsto \mathbf{1}_{t \leq s}$ .

(iii)  $t \mapsto \Psi_t$  is strongly differentiable and satisfies the operator Riccati equation

$$\begin{aligned} \dot{\Psi}_t &= 2\Psi_t \dot{\Sigma}_t \Psi_t, \quad t \in [0, T] \\ \Psi_T &= -\left(\text{Id} - \hat{K}\right)^{-*} \Theta^\top \Theta \left(\text{Id} - \hat{K}\right)^{-1} \end{aligned} \quad (5.11)$$

where  $\dot{\Sigma}_t$  is the strong derivative of  $t \mapsto \Sigma_t$  induced by the kernel

$$\dot{\Sigma}_t(s, u) = -K(s, t) \eta (U - 2C^\top C) \eta^\top K(u, t)^\top, \quad a.e. \quad (5.12)$$

*Proof.* The proof is given in Appendix B.2. □

We are now ready to provide a solution for the Riccati-BSDE (3.2). For this, denote by  $g$  the process

$$g_t(s) = \mathbf{1}_{t \leq s} \left( g_0(s) + \int_0^t K(s, u) DY_u du + \int_0^t K(s, u) \eta dW_u \right). \quad (5.13)$$

One notes that for each,  $s \leq T$ ,  $(g_t(s))_{t \leq s}$  is the adjusted forward process

$$g_t(s) = \mathbb{E} \left[ Y_s - \int_t^s K(s, u) DY_u du \mid \mathcal{F}_t \right], \quad s \geq t.$$

We also denote the trace of an integral operator  $\mathbf{F}$  by  $\text{Tr}(\mathbf{F}) = \int_0^T \text{tr}(F(s, s))ds$ , where  $\text{tr}$  is the usual trace of a matrix, and we define the function  $\phi$  by

$$\begin{cases} \dot{\phi}_t &= \text{Tr}(\Psi_t \dot{\mathbf{A}}_t) - 2r(t) \\ &= \int_{(t, T]} \text{tr}(\Theta^\top \Theta K(s, t) \eta U \eta^\top K(s, t)^\top) ds \\ &\quad - \int_{(t, T]^2} \text{tr}(\psi_t(s, u) K(u, t) \eta U \eta^\top K(s, t)^\top) dsdu - 2r(t), \\ \phi_T &= 0, \end{cases} \quad (5.14)$$

where  $\dot{\mathbf{A}}_t$  is the integral operator induced by the kernel given by

$$\dot{\mathbf{A}}_t(s, u) = -K(s, t) \eta U \eta^\top K(u, t)^\top, \quad u, s \leq T.$$

**Lemma 5.7.** *Fix a kernel  $K$  as in Definition 5.2 satisfying (5.9). Assume that  $(U - 2C^\top C) \in \mathbb{S}_+^N$ . Let  $\Psi$  be the operator defined in (5.6). Then, the process  $(\Gamma, Z^1, Z^2)$  defined by*

$$\begin{cases} \Gamma_t &= \exp(\phi_t + \langle g_t, \Psi_t g_t \rangle_{L^2}), \\ Z_t^1 &= 0, \\ Z_t^2 &= 2((\Psi_t K \eta)^* g_t)(t), \end{cases} \quad (5.15)$$

where  $g$  and  $\phi$  are respectively given by (5.13) and (5.14), is a  $\mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R}) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^N)$ -valued solution to the Riccati-BSDE (3.2).

*Proof.* Set  $G_t = \phi_t + \langle g_t, \Psi_t g_t \rangle_{L^2}$ , so that  $\Gamma_t = \exp(G_t)$  and

$$d\Gamma_t = \Gamma_t(dG_t + \frac{1}{2}d\langle G \rangle_t). \quad (5.16)$$

To obtain the dynamics of  $G$  it suffices to determine the dynamics of the process  $t \mapsto \langle g_t, \Psi_t g_t \rangle_{L^2}$ .

*Step 1.* In this step we prove that the dynamics of  $t \mapsto \langle g_t, \Psi_t g_t \rangle_{L^2}$  is given by

$$d\langle g_t, \Psi_t g_t \rangle_{L^2} = \left( \langle g_t, \dot{\Psi}_t g_t \rangle_{L^2} + \lambda_t^\top \lambda_t + 2\lambda_t^\top C Z_t^2 + \text{Tr}(\Psi_t \dot{\mathbf{A}}_t) \right) dt + (Z_t^2)^\top dW_t. \quad (5.17)$$

We first note that

$$\langle g_t, \Psi_t g_t \rangle_{L^2} = \int_0^T g_t(s)^\top (\Psi_t g_t)(s) ds,$$

and compute the dynamics of  $t \mapsto g_t(s)^\top (\Psi_t g_t)(s)$ . For fixed  $s \leq T$ , it follows from (5.13) and the fact that  $Y_t = g_t(t)$ , that

$$dg_t(s) = -\delta_{t=s} g_t(t) dt + K(s, t) Dg_t(t) dt + K(s, t) \eta dW_t.$$

Together with Lemma 5.6-(iii), we deduce that  $t \mapsto (\Psi_t g_t)(s)$  is a semimartingale with the following dynamics

$$\begin{aligned} d(\Psi_t g_t)(s) &= (\dot{\Psi}_t g_t)(s)dt + (\Psi_t dg_t)(s) \\ &= (\dot{\Psi}_t g_t)(s)dt - \psi_t(s, t)g_t(t)dt + (\Psi_t K(\cdot, t)Dg_t(t))(s)dt + (\Psi_t K(\cdot, t)\eta dW_t)(s). \end{aligned}$$

Here, we used the fact that  $\text{Id}\delta_t = 0$ : indeed, for every  $f \in L^2([0, T], \mathbb{R}^d)$  we have  $(\text{Id}\delta_t)(f) = (f(\cdot)\mathbf{1}_{t=}) = 0_{L^2}$ . Moreover,

$$\begin{aligned} d\langle g(\cdot)(s), (\Psi \cdot g)(s) \rangle_t &= -\text{tr}(\Theta^\top \Theta K(s, t)\eta U \eta^\top K(s, t)^\top)dt \\ &\quad + \int_t^T \text{tr}(\psi_t(s, u)K(u, t)\eta U \eta^\top K(s, t)^\top)dudt \\ &= \text{tr}(\Theta^\top \Theta \dot{\Lambda}_t(s, s))dt - \int_t^T \text{tr}(\psi_t(s, u)\dot{\Lambda}_t(u, s))dudt \\ &= -\text{tr}\left((\Psi_t \dot{\Lambda}_t(\cdot, s))(s)\right). \end{aligned}$$

Whence, combining the previous three identities, we get

$$\begin{aligned} d\left(g_t(s)^\top (\Psi_t g_t)(s)\right) &= dg_t(s)^\top (\Psi_t g_t)(s) + g_t(s)^\top d(\Psi_t g_t)(s) + d\langle g(\cdot)(s), (\Psi \cdot g)(s) \rangle_t \\ &= -\delta_{t=s}g_t(t)^\top (\Psi_t g_t)(s)dt + g_t(t)^\top D^\top K(s, t)^\top (\Psi_t g_t)(s)dt \\ &\quad + g_t(s)^\top (\dot{\Psi}_t g_t)(s)dt - g_t(s)^\top \psi_t(s, t)g_t(t)dt + g_t(s)^\top (\Psi_t K(\cdot, t)Dg_t(t))(s)dt \\ &\quad - \text{tr}\left((\Psi_t \dot{\Lambda}_t(\cdot, s))(s)\right) \\ &\quad + dW_t^\top \eta^\top K(s, t)^\top (\Psi_t g_t)(s) + g_t(s)^\top (\Psi_t K(\cdot, t)\eta dW_t)(s) \\ &= [\mathbf{I}(s) + \mathbf{II}(s) + \mathbf{III}(s) + \mathbf{IV}(s) + \mathbf{V}(s) + \mathbf{VI}(s)]dt + \mathbf{VII}(s) + \mathbf{VIII}(s). \end{aligned}$$

We now integrate in  $s$ . First, using Lemma 5.6-(i) we get that

$$\begin{aligned} \int_0^T [\mathbf{I}(s) + \mathbf{IV}(s)]ds &= -g_t(t)^\top (\Psi_t g_t)(t) - g_t(t)^\top \int_t^T \psi_t(t, u)g_t(u)du \\ &= \lambda_t^\top \lambda_t - 2g_t(t)^\top \int_t^T \psi_t(t, u)g_t(u)du \\ &= \lambda_t^\top \lambda_t - 2g_t(t)^\top ((\Psi_t + \Theta^\top \Theta \text{Id})g_t)(t). \end{aligned}$$

On the other hand, since,  $\Psi^* = \Psi$ , we have

$$\int_0^T [\mathbf{II}(s) + \mathbf{V}(s)]ds = 2g_t(t)^\top \left( ((KD)^* \Psi_t)g_t \right)(t).$$

Therefore, summing the above, using Lemma 5.6-(ii), and the definition of  $\hat{\mathbf{K}}$ , we get

$$\begin{aligned} \int_0^T [\mathbf{I}(s) + \mathbf{IV}(s) + \mathbf{II}(s) + \mathbf{V}(s)] ds &= \lambda_t^\top \lambda_t - 2g_t(t)^\top \left( (\Psi_t + \Theta^\top \Theta \text{Id} - ((\mathbf{K}D)^*) \Psi_t) g_t \right)(t) \\ &= \lambda_t^\top \lambda_t + 4g_t(t)^\top ((\mathbf{K}\eta C^\top \Theta)^* \Psi_t) g_t(t) \\ &= \lambda_t^\top \lambda_t + 2\lambda_t^\top C Z_t^2. \end{aligned}$$

Finally, observing that

$$\begin{aligned} \int_0^T \mathbf{III}(s) ds &= \langle g_t, \dot{\Psi}_t g_t \rangle_{L^2}, \quad \int_0^T \mathbf{VI}(s) ds = \text{Tr}(\Psi_t \dot{\Lambda}_t), \\ \int_0^T [\mathbf{VII}(s) + \mathbf{VIII}(s)] ds &= (Z_t^2)^\top dW_t, \end{aligned}$$

we obtain the claimed dynamics (5.17).

*Step 2.* Plugging the dynamics (5.17) in (5.16) yields

$$\begin{aligned} \frac{d\Gamma_t}{\Gamma_t} &= \left[ \underbrace{\dot{\phi}_{t,T} - \text{Tr}(\Psi_t \dot{\Lambda}_t)}_{\mathbf{1}} + \underbrace{\langle g_t, \dot{\Psi}_t g_t \rangle_{L^2} + \frac{(Z_t^2)^\top U Z_t^2}{2}}_{\mathbf{2}} + \underbrace{\lambda_t^\top \lambda_t + 2\lambda_t^\top C Z_t^2}_{\mathbf{3}} \right] dt \\ &\quad + (Z_t^2)^\top dW_t. \end{aligned}$$

By (5.14), we have:  $\mathbf{1} = -2r(t)$ . From the definition of  $Z^2$ , we have

$$\begin{aligned} \frac{(Z_t^2)^\top U Z_t^2}{2} &= 2 \left[ \left( (\Psi_t \mathbf{K} \eta)^* g_t \right)(t) \right]^\top U \left( (\Psi_t \mathbf{K} \eta)^* g_t \right)(t) \\ &= -2 \langle g_t, (\Psi_t \dot{\Lambda}_t \Psi_t) g_t \rangle_{L^2}. \end{aligned}$$

Thus, using the Riccati relation (5.11), we get

$$\begin{aligned} \mathbf{2} &= \langle g_t, (\dot{\Psi}_t - \Psi_t \dot{\Lambda}_t \Psi_t) g_t \rangle_{L^2} = 4 \left[ \left( (\Psi_t \mathbf{K} \eta)^* g_t \right)(t) \right]^\top C^\top C \left( (\Psi_t \mathbf{K} \eta)^* g_t \right)(t) \\ &= (Z_t^2)^\top C C^\top Z_t^2. \end{aligned}$$

Combining  $\mathbf{1}$ ,  $\mathbf{2}$  and  $\mathbf{3}$  yields

$$\frac{d\Gamma_t}{\Gamma_t} = \left( -2r(t) + |\lambda_t + Z_t^1 + C Z_t^2|^2 \right) dt + (Z_t^2)^\top dW_t.$$

This shows that  $(\Gamma, Z^1, Z^2)$  solves (3.2).

*Step 3.* It remains to check that  $(\Gamma, Z^1, Z^2) \in \mathbb{S}_{\mathbb{F}}^{\infty}([0, T], \mathbb{R}) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^N)$ . For this, observe that since  $\Psi$  is a nonpositive operator over  $[0, T]$ , we have the bound  $0 < \Gamma_t \leq e^{\int_t^T 2r(s)ds}$ . Finally, to show that  $Z^2 \in L_{\mathbb{F}}^2([0, T], \mathbb{R}^d)$ , it is enough to show that

$$\mathbb{E} \left[ \int_0^T \left| \int_t^T K(s, t)^{\top} g_t(s) ds \right|^2 dt \right] < \infty,$$

and  $\mathbb{E} \left[ \int_0^T \left| \int_{(t, T]^2} K(v, t)^{\top} \psi_t(v, s) g_t(s) dv ds \right|^2 dt \right] < \infty.$

This follows from the fact that  $K$  and  $\psi$  satisfy (5.1)-(5.10) respectively, and

$$\sup_{0 \leq t \leq s \leq T} \mathbb{E} [|g_t(s)|^2] \leq \sup_{s \leq T} |g_0(s)|^2 \left( 1 + \sup_{s \leq T} \int_0^T |R_D(s, u)|^2 du \right) < \infty,$$

where  $R_D$  is the resolvent of  $KD$ . □

From Theorem 3.1, we can now explicitly solve the Markowitz problem (2.6) in the quadratic Volterra model (5.2), (5.3) and (5.4), see Theorem 5.9 below. In order to verify condition (H2) of Theorem 3.1, we will first need the following lemma whose proof is postponed to Appendix B.3.

**Lemma 5.8.** *Let the assumptions of Lemma 5.7 be in force. Assume  $|D - 2\eta C^{\top} \Theta| \times \|K\|_{L^2([0, T]^2)}^2 < 1$ , then*

$$|\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2 \leq \kappa(\Theta) \left( |g_s(s)|^2 + \int_0^T |g_s(u)|^2 du \right), \quad s \leq T, \quad \Theta \in \mathbb{R}^{d \times N} \quad (5.18)$$

where  $\kappa(\Theta) = c|\Theta|^2(1 + |\Theta|^4 \hat{\kappa}(\Theta))$  with  $c > 0$  independent of  $\Theta$  and

$$\hat{\kappa}(\Theta) = \left( \frac{|f(\Theta)| \times \|K\|_{L^2([0, T]^2)}^2}{1 - |f(\Theta)| \times \|K\|_{L^2([0, T]^2)}^2} \right)^4.$$

*Proof.* See Appendix B.3. □

We now arrive to the main result of this section.

**Theorem 5.9.** *Fix a kernel  $K$  as in Definition 5.2 satisfying (5.9) and assume that  $(U - 2C^{\top}C) \in \mathbb{S}_+^N$ . Let  $a(p)$  be as in (3.4) and  $\kappa$  the function defined in Lemma 5.8. Assume that there exists  $\Theta \in \mathbb{R}^{d \times N}$  such that*

$$\mathbb{E} \left[ \exp \left( a(p) \kappa(\Theta) \int_0^T (|g_s(s)|^2 + \int_0^T |g_s(u)|^2 du) ds \right) \right] < \infty, \quad (5.19)$$



for some  $p > 2$ . Assume that  $g_0^i(0) > 0$  for some  $i \leq d$ . Then, the optimal investment strategy for the Markowitz problem (2.6) is given by the admissible control

$$\alpha_t^* = -\left((\Theta + 2C[\Psi_t \mathbf{K} \eta]^*)g_t\right)(t) \left(X_t^{\alpha^*} - \xi^* e^{-\int_t^T r(s)ds}\right), \quad (5.20)$$

where  $\xi^*$  is defined in (3.6), and the optimal value is given by (3.7) with  $\Gamma_0$  as in (5.15).

*Proof.* First note that under the specification (5.15), and  $\lambda_t = \Theta Y_t = \Theta g_t(t)$ , the candidate for the optimal feedback control defined in (3.5) takes the form

$$\begin{aligned} \alpha_t^* &= -(\lambda_t + Z_t^1 + CZ_t^2) (X_t^{\alpha^*} - \xi e^{-\int_t^T r(s)ds}) \\ &= \left((\Theta + 2C[\Psi_t \mathbf{K} \eta]^*)g_t\right)(t) (X_t^{\alpha^*} - \xi e^{-\int_t^T r(s)ds}). \end{aligned}$$

It thus suffices to check that the assumptions of Theorem 3.1 are verified to ensure that  $\alpha^*(\xi^*)$  is optimal and to get that (3.7) is the optimal value. The existence of a solution triplet  $(\Gamma, Z^1, Z^2) \in \mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R}) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^2([0, T], \mathbb{R}^N)$  to the stochastic backward Riccati equation (3.2) is ensured by Lemma 5.7. In addition, we have

$$\Gamma_0 = \mathbb{E} \left[ e^{\int_0^T (2r(s) - |\lambda_s + Z_s^1 + CZ_s^2|^2) ds} \right] = \mathbb{E} \left[ e^{\int_0^T [2r(s) - |((\Theta + 2C[\Psi_s \mathbf{K} \eta]^*)g_s)(s)|^2] ds} \right],$$

which implies that  $\Gamma_0 < e^{2 \int_0^T r(s)ds}$  since  $g_0^i(0) > 0$  for some  $i \leq d$  by assumption. Thus condition (H1) of Theorem 3.1 is verified. Condition (H2) follows directly from Lemma 5.8 and (5.19). The proof is complete.  $\square$

The following lemma provides a general sufficient condition for the existence of  $\Theta$  satisfying (5.19). Without loss of generality, we assume that  $D = 0$  in (5.2).<sup>3</sup> Define  $Z(s, u) = (\frac{1}{T}g_s(s), g_s(u))^\top$  for any  $s, u \in [0, T]$ , which we view as a random variable in  $L^2([0, T]^2, \mathbb{R}^{2N})$ . Its mean is given by  $\mu(s, u) = \mathbb{E}[Z(s, u)] = (\frac{1}{T}g_0(s), g_0(u))^\top$  and its covariance kernel by

$$\bar{\Sigma}((s, u), (t, r)) = \mathbb{E} \left[ \left( Z(s, u) - \mathbb{E}(Z(s, u)) \right) \left( Z(t, r) - \mathbb{E}(Z(t, r)) \right)^\top \right], \quad s, u, t, r \in [0, T],$$

which is symmetric and nonnegative. It follows from assumption (5.1) that  $\bar{\Sigma}$  is continuous on  $[0, T]^4$  so that an application of Mercer's theorem, see [Shorack and Wellner \(2009, Theorem 1 p.208\)](#), yields the existence of a countable orthonormal basis  $(e^n)_{n \geq 1}$

<sup>3</sup>If  $D \neq 0$ , then making use of the resolvent kernel  $R_D$  of  $KD$ , we reduce to the case  $D = 0$  as illustrated on (5.5) by working on the kernel  $(K + R_D)$  instead of  $K$ .

in  $L^2([0, T]^2, \mathbb{R}^{2N})$  and a non increasing sequence of nonnegative numbers  $(\lambda^n)_{n \geq 1}$ , with  $\lambda^n \rightarrow 0$ , as  $n \rightarrow \infty$ , such that

$$\bar{\Sigma}((s, u), (t, r)) = \sum_{n \geq 1} \lambda^n e^n(s, u) e^n(t, r)^\top. \quad (5.21)$$

In addition, we observe by virtue of (5.1) that

$$\begin{aligned} \sum_{n \geq 1} \lambda^n = \text{tr}(\bar{\Sigma}) &= \frac{1}{T} \int_0^T \left( \int_0^s \text{tr} \left( K(s, z) \eta U \eta^\top K(s, z)^\top \right) dz \right) ds \\ &+ \int_0^T \left( \int_0^T \left( \int_0^u \text{tr} \left( K(s, z) \eta U \eta^\top K(s, z)^\top \right) dz \right) ds \right) du < \infty. \end{aligned} \quad (5.22)$$

**Lemma 5.10.** *Set  $D = 0$ . Let  $a > 0$  be such that  $2a < \frac{1}{\lambda^1}$ . Then,*

$$\mathbb{E} \left[ \exp \left( a \int_0^T (|g_s(s)|^2 + \int_0^T |g_s(u)|^2 du) ds \right) \right] < \infty.$$

*In particular, (5.19) holds if  $2a(p)\kappa(\Theta) < \frac{1}{\lambda^1}$  for some  $p > 2$ .*

*Proof.* We refer to Appendix B.4. □

**Remark 5.11.** *In practice, as  $\lambda_1 \leq \text{tr}(\bar{\Sigma})$ , it follows from Lemma 5.10 and (5.22), that a sufficient condition for the existence of  $\Theta$  satisfying (5.19) would be*

$$2a(p)\kappa(\Theta) < \frac{1}{\text{tr}(\bar{\Sigma})}.$$

*For instance, for the fractional convolution kernel  $K(t, s) = \mathbf{1}_{s \leq t} (t - s)^{H-1/2}$ , we have  $\int_0^T \int_0^T |K(t, s)|^2 ds dt = T^{2H+1}$ . Consequently  $\text{tr}(\bar{\Sigma}) \geq \eta^2 (T^{2H} + T^{2(H+1)})$  and the condition on  $\Theta$  reads*

$$\kappa(\Theta) \leq (2a(p)\eta^2 (T^{2H} + T^{2(H+1)}))^{-1}.$$

The following corollary treats the standard Markovian and semimartingale case for  $K = I_N$  and shows how to recover the well-known formulae in the spirit of [Chiu and Wong \(2014\)](#).

**Corollary 5.12.** Set  $K(t, s) = I_N 1_{s \leq t}$  and  $g_0(t) \equiv Y_0$  for some  $Y_0 \in \mathbb{R}^N$ . Then, the solution to the Riccati BSDE can be re-written in the form

$$\Gamma_t = \exp(\phi_t + Y_t^\top P_t Y_t), \quad \text{and} \quad Z_t^2 = 2\eta^\top P_t Y_t, \quad (5.23)$$

where  $P : [0, T] \mapsto \mathbb{R}^{N \times N}$  and  $\phi$  solve the conventional system of  $N \times N$ -matrix Riccati equations

$$\begin{cases} \dot{P}_t &= \Theta^\top \Theta + P_t(2\eta C^\top \Theta - D) + (2\eta C^\top \Theta - D)^\top P_t + 2P_t(\eta(U - 2C^\top C)\eta^\top)P_t, \\ P_T &= 0, \\ \dot{\phi}_t &= -2r(t) - \text{tr}(P_t \eta U \eta^\top), \quad t \in [0, T], \\ \phi_T &= 0. \end{cases}$$

Furthermore, the optimal control reads

$$\alpha_t^* = -\left(\Theta + 2C(D\eta)^\top P_t Y_t\right)\left(X_t^{\alpha^*} - \xi^* e^{-\int_t^T r(s)ds}\right). \quad (5.24)$$

*Proof.* For  $K(t, s) = I_N 1_{s \leq t}$ ,

$$Y_s = Y_t + \int_t^s DY_u du + \int_t^s \eta dW_u, \quad s \geq t,$$

so that the adjusted forward process reads

$$g_t(s) = \mathbb{E}\left[Y_s - \int_t^s DY_u du \mid \mathcal{F}_t\right] = \mathbf{1}_{t \leq s} Y_t,$$

and the solution to the Riccati BSDE can be re-written in the form

$$\Gamma_t = \exp(\phi_t + \langle g_t, \Psi_t g_t \rangle_{L^2}) = \exp(\phi_t + Y_t^\top P_t Y_t),$$

where  $P_t = \int_t^T (\Psi_t \mathbf{1}_t)(s) ds$  with the  $\mathbb{R}^N$ -valued indicator function  $\mathbf{1}_t : (s) \mapsto (1_{t \leq s}, \dots, 1_{t \leq s})^\top$ . We now derive the equations satisfied by  $P$  and  $\phi$ . First we have  $K_T = 0$  and

$$\begin{aligned} \dot{P}_t &= -(\Psi_t \mathbf{1}_t)(t) + \int_t^T \frac{d(\Psi_t \mathbf{1}_t)(s)}{dt} ds \\ &= -(\Psi_t \mathbf{1}_t)(t) + \int_t^T (\dot{\Psi}_t \mathbf{1}_t)(s) ds - \int_t^T \psi_t(s, t) ds \\ &= \mathbf{1} + \mathbf{2} + \mathbf{3}. \end{aligned}$$

Using Lemma 5.6–(ii) and the expression  $\hat{K}(s, u) = 1_{u \leq s}(D - 2\eta C^\top \Theta)$  we get

$$\mathbf{1} = (-\Theta^\top \Theta \text{Id} + \hat{K}^* \Psi_t)(1)(t) = -\Theta^\top \Theta + (D - 2\eta C^\top \Theta)^\top P_t.$$

Furthermore, Lemma 5.6–(iii) and  $\dot{\Sigma}_t(s, u) = 1_{t \leq s \wedge u} \eta(U - 2C^\top C) \eta^\top$  yield

$$\begin{aligned} \mathbf{2} &= \int_t^T (\dot{\Psi}_t 1_t)(s) ds = \int_t^T (\Psi_t \dot{\Sigma}_t \Psi_t 1_t)(s) ds \\ &= \left( \int_t^T (\Psi_t 1_t)(s) ds \right) \eta(U - 2C^\top C) \eta^\top \left( \int_t^T (\Psi_t 1_t)(s) ds \right) \\ &= P_t(\eta(U - 2C^\top C) \eta^\top) P_t. \end{aligned}$$

Moreover, by using Lemma 5.6–(i)–(ii), we obtain

$$\mathbf{3} = - \int_t^T \psi_t(s, t) ds = -(\Psi_t + \Theta^\top \Theta id)^*(1_t) = -(\hat{K}^* \Psi_t)^*(1)(t) = -P_t(D - 2\eta C^\top \Theta).$$

This proves the equation for  $P$ , and that of  $\phi$  is immediate. Finally to prove the formula of  $Z^2$  in (5.23) and  $\alpha^*$  in (5.24) it suffices to observe the following identity

$$((\Psi_t K \eta)^* g_t)(t) = \eta^\top P_t Y_t.$$

□

## 6 Numerical experiment: rough Stein-Stein for two assets

We illustrate the results of Section 5 on a special case of the two dimensional rough Stein-Stein model as described in Example 5.3. We consider a four dimensional Brownian motion  $(B^1, B^2, B^{1,\perp}, B^{2,\perp})$ , and define

$$\tilde{B}^1 = B^1, \quad \tilde{B}^2 = \rho B^1 + \sqrt{1 - \rho^2} B^2, \quad W^i = c_i \tilde{B}^i + \sqrt{1 - c_i^2} \tilde{B}^{\perp, i},$$

for some  $\rho \in [-1, 1]$ , and  $c_i \in [-1, 1]$ ,  $i = 1, 2$ .

For simplicity we set  $r \equiv 0$ , and consider two stocks of price process  $S^1$  and  $S^2$  with the following dynamics<sup>4</sup>

$$\begin{cases} dS_t^i &= S_t^i \theta_i (Y_t^i)^2 dt + S_t^i Y_t^i d\tilde{B}_t^i, \\ Y_t^i &= Y_0^i + \frac{1}{\Gamma(H_i + 1/2)} \int_0^t (t-s)^{H_i-1/2} \eta_i dW_s^i, \quad i = 1, 2, \end{cases}$$

with  $H^i > 0$ ,  $\eta_i, \theta_i \geq 0$  and  $Y_0^i \in \mathbb{R}$ .

Although the framework of Section 5 allows for a more general correlation structure for the Brownian motion, the model is already rich enough to capture the following stylized facts:

---

<sup>4</sup>This corresponds to Example 5.3–(i) with  $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (1, 0, \rho, \sqrt{1 - \rho^2})$  and  $\Theta = \beta^{-1} \text{diag}(\theta_1, \theta_2)$ .

- the two stocks  $S^i$ ,  $i = 1, 2$ , are correlated through  $\rho$ ,
- each stock  $S^i$  has a stochastic rough volatility  $|Y^i|$  with possibly different Hurst indices  $H_i$ ,
- each stock  $S^i$  is correlated with its own volatility process through  $c_i$  to take into account the leverage effect.

Our main motivation for considering the multivariate rough Stein-Stein model is to study the ‘buy rough sell smooth’ strategy of [Glasserman and He \(2020\)](#) that was back-tested empirically: this strategy consisting in buying the roughest assets while shorting on the smoothest ones was shown to be profitable. We point out that the numerical simulations for the one dimensional rough Heston model carried in [Han and Wong \(2020a\)](#) by varying the Hurst index  $H$  could not provide much insight on such strategy, apart from suggesting that the vol-of-vol has a possible impact on the ‘buy rough sell smooth strategy’. Our quadratic multivariate framework allows for more flexible simulations, with a richer correlation structure compared to multivariate extensions of the rough Heston model, recall [Remark 5.4](#). Our results below provide new insights on the strategy by showing that the correlation between stocks plays a key role.

Our present goal is to illustrate the influence of some parameters, namely the horizon  $T$ , the vol-of-vol  $\eta$  and the correlation  $\rho$  between the stocks, onto the optimal investment strategy when two assets, one rough and one smooth with  $H_1 < H_2$ , are at stakes. To ease comparison, we set  $c_1 = c_2 = -0.7$  for the leverage effects,  $Y_0^1 = Y_0^2$  and we normalize the vol-of-vols by setting  $\eta_1 = \eta_2$ . We consider the evolution of optimal vector of amount invested into each stock, i.e.,  $t \mapsto \pi_t^*$  (recall that  $\alpha_t^* = \sigma_t^\top \pi_t^*$  with  $\sigma = \text{diag}(Y^1, Y^2)\beta$  and  $\alpha^*$  is given by [\(5.20\)](#)).  $\pi$  being a stochastic process, we also consider the deterministic function  $t \mapsto ((\Theta + 2C[\Psi_t \mathbf{K}\eta]^*)Y_0)(t)(\xi^*)$ , where  $\xi^*$  is defined in [\(3.6\)](#), to help us in our analysis.

For our implementation of  $\alpha^*$  given by [\(5.20\)](#), we discretize in time the operators acting on  $L^2$ , so that the kernel of the operator  $\Psi$  in [\(5.6\)](#) is approximated by a finite dimensional matrix (see for instance [Abi Jaber \(2019c, Section 2.3\)](#) for a similar procedure) and the Gaussian process  $(g_t(s))_{t \leq s \leq T}$  defined in [\(5.13\)](#) is simulated by Cholesky’s decomposition algorithm. We refer to the following [url](#) for the full code and additional simulations.

Our observations from the simulations are the following.

**1. Horizon  $T$ :** With the goal of understanding the effect of the horizon  $T$  on the investment strategy, we fix all parameters but  $T$  with  $\rho = 0$ . The results are illustrated on [Figures 1a-1b-1c](#) and [2a-2b-2c](#). We can distinguish 3 regimes:

- $T \ll 1$  : When the investment horizon is close to the end, the rough asset is over-weighted over the smooth one.

- $T \approx 1$  : A transition appears, as the smooth asset is first overweighted and then the rough asset becomes overweighted as we approach the final horizon.
- $T \gg 1$  : The smooth asset is overweighted all along the experiment, letting its first position only when the maturity is close, suggesting that the transition point becomes closer to  $T$  as  $T$  grows.

One possible interpretation of this transition is the following. Rough processes are more volatile than smooth processes in the short term but less volatile in the long term, since their variances evolve approximately as  $t^{2H}$ . Thus, when there is not much time left, it seems natural to look for rough processes to obtain some performance. Conversely, the more time we have, the more we favor the smooth asset.

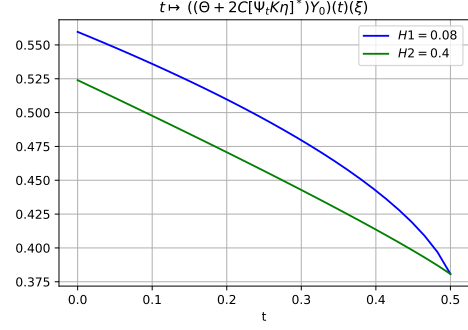
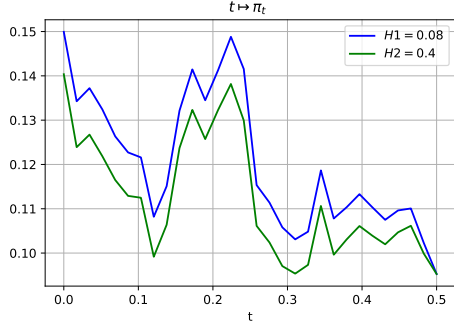
**2. Vol-of-vol  $\eta$ :** The volatility of volatility seems to have the opposite effect of the horizon  $T$  over the investment strategy as shown on Figures 3a-3b.

- $\eta \ll 1$  : The smooth asset and then the rough asset are successively overweighted.
- $\eta \gg 1$  : The rough asset is overweighted.

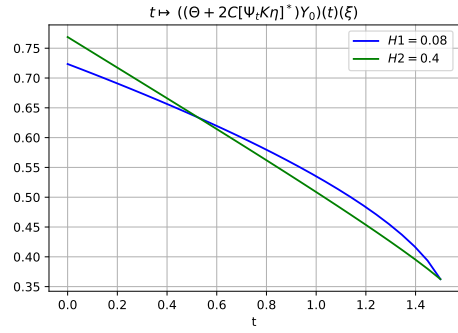
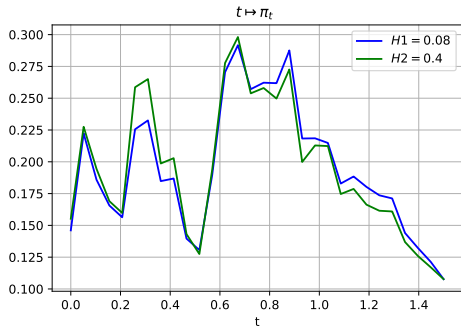
It is quite natural to expect the vol-of-vol to have an inverse effect when compared to the horizon  $T$ , since increasing the vol-of-vol is similar to accelerating the time scale at a certain rate depending on  $H$  (think of the self-similarity property of fractional Brownian motion).

### 3. Correlation $\rho$ :

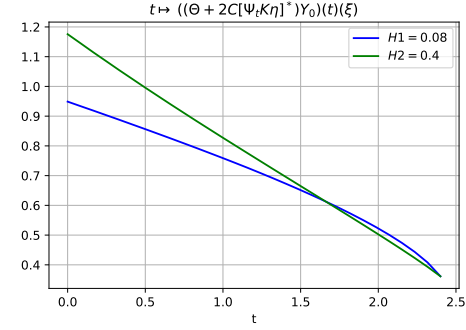
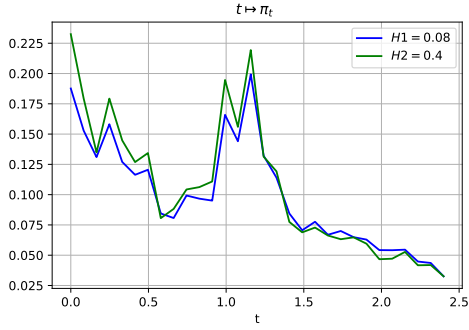
- $\rho < 0$  : In the case of negatively correlated assets it is natural to expect the following strategy : pick both assets in order to be protected from volatility and benefit from the drift. So we expect the case  $\rho < 0$  to be similar from  $\rho = 0$  except that the transition from  $T \ll 1$  to  $T \gg 1$  should appear at a greater  $T$ . This is what we observe on Figures 5a-5b-5c. We interpret this evolution towards the equally weighted portfolio as the possibility to be protected from volatility by holding both assets.
- $\rho > 0$  : when the two stocks are positively correlated with  $\rho > 0$ , there is no minimization of variance through diversification by going long in both assets. Thus in the case a positively correlated assets, it is natural to expect the emergence of a starker choice between the assets. In the  $\rho > 0$  case, see Figures 4a-4b, we observe a *buy rough sell smooth* strategy as the one empirically found in Glasserman and He (2020).



(a)  $T = 0.5$



(b)  $T = 1.5$



(c)  $T = 2.4$

Figure 1: Effect of the horizon  $T$  on the optimal allocation strategy. When the horizon  $T$  approaches, the rough stock in blue is preferred. When  $T$  is big enough and the horizon far enough the smooth stock in green is preferred. (The parameters are:  $H_1 = 0.08$ ,  $H_2 = 0.4$ ,  $\rho = 0$ ,  $\eta_1 = \eta_2 = 1$ ,  $c_i = -0.7$ .)

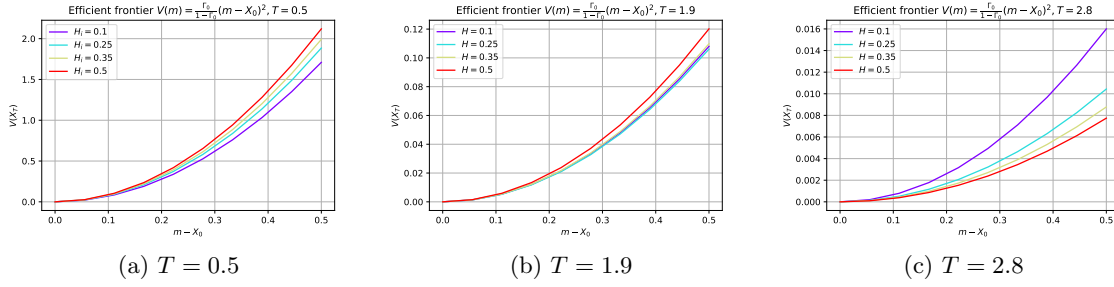


Figure 2: The efficient frontier in the case where both assets have the same roughness  $H_1 = H_2 = H$ . When the horizon  $T$  is small, the rough stocks allows for lower variance. When  $T$  increases we observe a transition and an inversion of the relation order. Indeed, when  $T$  increases, it is the smoothest stocks that allow for a lower variance.

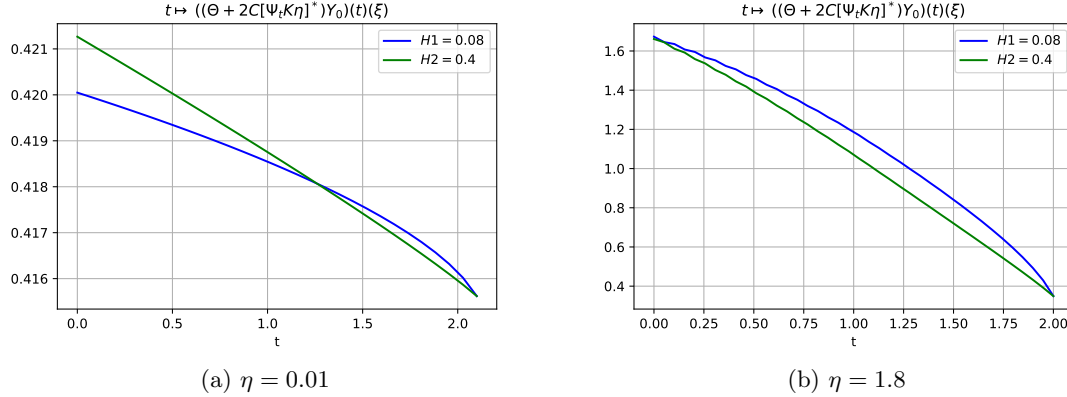


Figure 3: As the vol-of-vol  $\eta$  increases, it is as if the horizon  $T$  was decreasing and the rough stock in blue begins to be preferred.  $H_1 = 0.08$ ,  $H_2 = 0.4$ ,  $T = 2.1$ ,  $\rho = 0$ ,  $c_i = -0.7$ .

As a further line of research, we see two interesting paths :

- A theoretical study of influence of the parameters onto the investments strategies.
- An empirical study testing the different conjectures made about the influence of some parameters such as  $T, \eta, \rho, H$ , etc.

Our numerical results extend to larger horizon  $T$ . For instance, in Figure 6, we took a maturity of  $T = 20$  years, although we noted that a smaller  $\eta = 0.1$  had to be chosen to avoid any blow-up, in accordance with Remark 5.11.



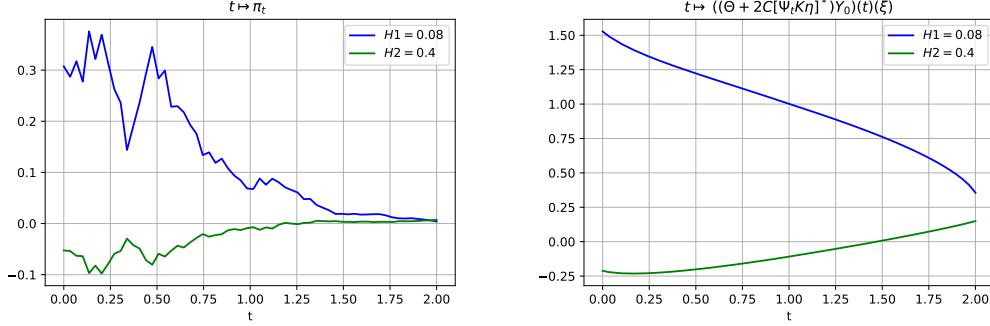


Figure 4:  $\rho = 0.7$ , when the two assets are positively correlated we recover *the buy rough sell smooth* strategy as it is described in [Glasserman and He \(2020\)](#). (the parameters are:  $H_1 = 0.08$ ,  $H_2 = 0.4$ ,  $T = 2.1$ ,  $\eta_1 = \eta_2 = 1$ ,  $c_i = -0.7$ .)

## A Proof of the verification result

In this section, we provide a detailed proof of Theorem 3.1. It is well-known that Markowitz problem (2.6) is equivalent to the following max-min problem, see e.g. ([Pham, 2009](#), Proposition 6.6.5):

$$V(m) = \max_{\eta \in \mathbb{R}} \min_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[ |X_T - (m - \eta)|^2 \right] - \eta^2 \right\}. \quad (\text{A.1})$$

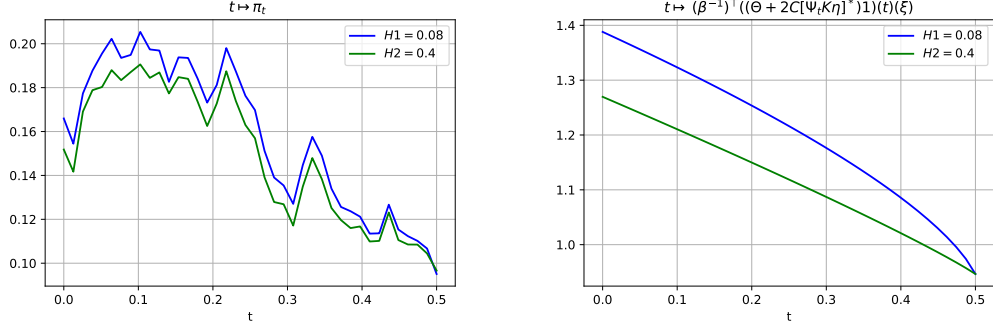
Thus, solving problem (2.6) involves two steps. First, the internal minimization problem in term of the Lagrange multiplier  $\eta$  has to be solved. Second, the optimal value of  $\eta$  for the external maximization problem has to be determined. Let us then introduce the inner optimization problem:

$$\tilde{V}(\xi) := \min_{\alpha \in \mathcal{A}} \mathbb{E} \left[ |X_T^\alpha - \xi|^2 \right], \quad \xi \in \mathbb{R}. \quad (\text{A.2})$$

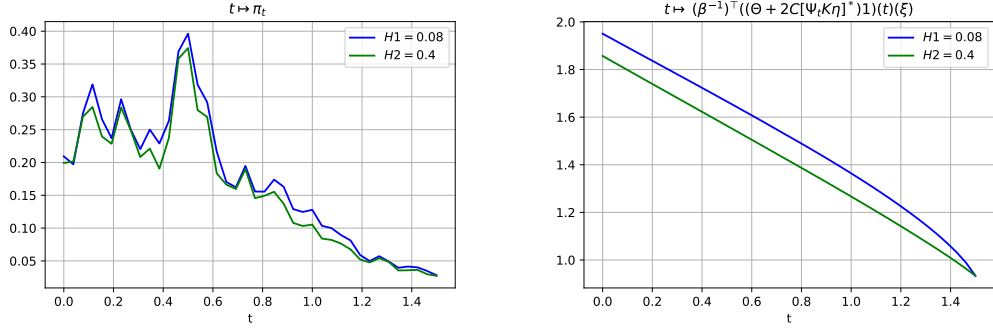
First, we provide a verification result for the inner optimization problem (A.2) via the standard completion of squares technique, see for instance [Lim and Zhou \(2002, Proposition 3.1\)](#), [Lim \(2004, Proposition 3.3\)](#) and [Chiu and Wong \(2014, Theorem 3.1\)](#).

**Lemma A.1.** *Assume there exists a solution triplet  $(\Gamma, Z^1, Z^2) \in \mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R}) \times L_{\mathbb{F}}^{2,loc}([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^{2,loc}([0, T], \mathbb{R}^N)$  to the Riccati BSDE (3.2) such that  $\Gamma_t > 0$ , for all  $t \leq T$ . Fix  $\xi \in \mathbb{R}$ , and assume that there exists an admissible control  $\alpha^*(\xi)$  satisfying*

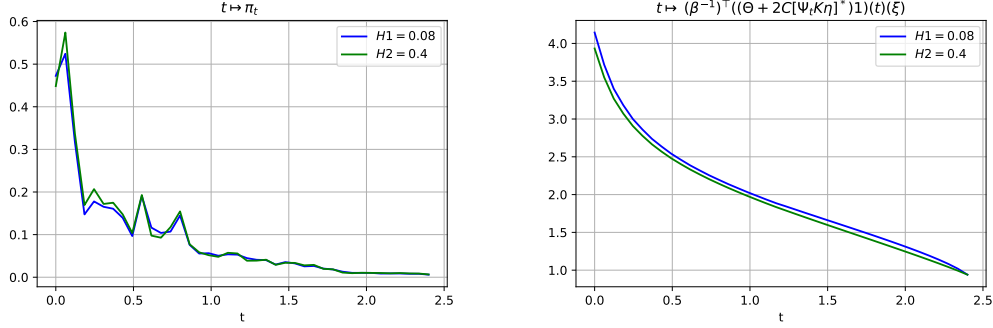
$$\alpha_t^*(\xi) = -(\lambda_t + Z_t^1 + CZ_t^2) \left( X_t^{\alpha^*(\xi)} - \xi e^{-\int_t^T r(s)ds} \right), \quad 0 \leq t \leq T. \quad (\text{A.3})$$



(a)  $T = 0.5$



(b)  $T = 1.5$



(c)  $T = 2.4$

Figure 5: Effect of the horizon  $T$  on the optimal allocation strategy when the two assets are negatively correlated ( $\rho = -0.4$ ),  $H_1 = 0.08$ ,  $H_2 = 0.4$ . As  $T$  increases the smooth stock in green is more and more weighted in comparison to the rough one in blue. But the transition takes more time compared to the case  $\rho = 0$ , see Figures 1a-1c.  $\eta_1 = \eta_2 = 1, c_i = -0.7$ . Note the beginning of the blow-up when  $T$  reaches  $T = 2.4$ , as it could be foreseen by the condition of Lemma 5.10.

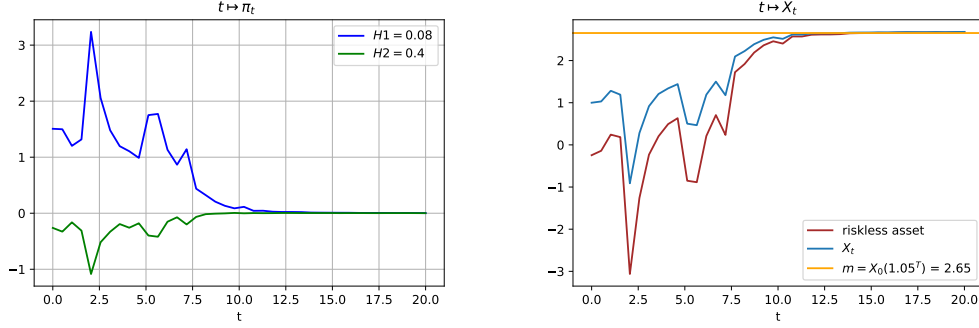


Figure 6: Simulation for a larger horizon  $T = 20$  years. Note that a smaller  $\eta$  had to be chosen in accordance with Remark 5.11. (The parameters are:  $H_1 = 0.08$ ,  $H_2 = 0.4$ ,  $\rho = 0$ ,  $\eta_1 = \eta_2 = 0.1$ ,  $c_i = -0.7$ .)

Then, the inner minimization problem (A.2) admits  $\alpha^*(\xi)$  as an optimal feedback control and the optimal value is

$$\tilde{V}(\xi) = \Gamma_0 \left| x_0 - \xi e^{-\int_0^T r(s)ds} \right|^2. \quad (\text{A.4})$$

*Proof.* Let us first define  $\tilde{X}_t^\alpha = X_t^\alpha - \xi e^{-\int_t^T r(s)ds}$ , for any  $\alpha \in \mathcal{A}$ . Then, by Itô's lemma we have

$$d\tilde{X}_t^\alpha = (r(t)\tilde{X}_t^\alpha + \alpha_t^\top \lambda_t)dt + \alpha_t^\top dB_t, \quad 0 \leq t \leq T, \quad \tilde{X}_0^\alpha = x_0 - \xi e^{-\int_0^T r(s)ds}.$$

As a result,  $\tilde{X}^\alpha$  and  $X^\alpha$  have the same dynamics and  $\tilde{X}_T^\alpha = X_T^\alpha - \xi$  so that problem (A.2) can be alternatively written as

$$\min_{\alpha \in \mathcal{A}} \mathbb{E} \left[ |\tilde{X}_T^\alpha|^2 \right].$$

To ease notations, we set  $h_t = \lambda_t + Z_t^1 + CZ_t^2$ . For any  $\alpha \in \mathcal{A}$ , Itô's lemma combined with (3.2) and a completion of squares in  $\alpha$  yield

$$\begin{aligned} d\left(\Gamma_t |\tilde{X}_t^\alpha|^2\right) &= |\tilde{X}_t^\alpha|^2 \Gamma_t (-2r(t) + h_t^\top h_t)dt + \Gamma_t |\tilde{X}_t^\alpha|^2 \left( (Z_t^1)^\top dB_t + (Z_t^2)^\top dW_t \right) \\ &\quad + \Gamma_t \left( 2\tilde{X}_t^\alpha (r(t)\tilde{X}_t^\alpha + \alpha_t^\top \lambda_t) + \alpha_t^\top \alpha_t \right)dt + 2\Gamma_t \tilde{X}_t^\alpha \alpha_t^\top dB_t \\ &\quad + 2\alpha_t^\top (Z_t^1 + CZ_t^2) \tilde{X}_t^\alpha dt \\ &= (\alpha_t + h_t \tilde{X}_t^\alpha)^\top \Gamma_t (\alpha_t + h_t \tilde{X}_t^\alpha)dt \\ &\quad + 2\Gamma_t \tilde{X}_t^\alpha \alpha_t^\top dB_t + \Gamma_t |\tilde{X}_t^\alpha|^2 \left( (Z_t^1)^\top dB_t + (Z_t^2)^\top dW_t \right). \end{aligned}$$

As a consequence, using  $\Gamma_T = 1$ , we get

$$\begin{aligned} |\tilde{X}_T^\alpha|^2 &= \Gamma_0 |\tilde{X}_0^\alpha|^2 + \int_0^T (\alpha_s + h_s \tilde{X}_s^\alpha)^\top \Gamma_s (\alpha_s + h_s \tilde{X}_s^\alpha) ds \\ &\quad + \int_0^T 2\Gamma_s \tilde{X}_s^\alpha \alpha_s^\top dB_s + \int_0^T 2\Gamma_s |\tilde{X}_s^\alpha|^2 \left( (Z_s^1)^\top dB_s + (Z_s^2)^\top dW_s \right). \end{aligned}$$

Note that the stochastic integrals

$$\int_0^\cdot 2\Gamma_s \tilde{X}_s^\alpha \alpha_s^\top dB_s, \quad \int_0^\cdot \Gamma_s |\tilde{X}_s^\alpha|^2 (Z_s^1)^\top dB_s, \quad \int_0^\cdot \Gamma_s (\tilde{X}_s^\alpha)^2 (Z_s^2)^\top dW_s,$$

are well-defined since  $X^\alpha$  is continuous,  $(\alpha, Z^1, Z^2)$  are in  $L_{\mathbb{F}}^{2,loc}([0, T])$  and  $\Gamma$  in  $\mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R})$ . Furthermore, they are local martingales. Let  $\{\tau_k\}_{k \geq 1}$  be a common localizing increasing sequence of stopping times converging to  $T$ . Then,

$$\mathbb{E} \left[ |\tilde{X}_{T \wedge \tau_k}^\alpha|^2 \right] = \Gamma_0 |\tilde{X}_0^\alpha|^2 + \mathbb{E} \left[ \int_0^{T \wedge \tau_k} (\alpha_s + h_s \tilde{X}_s^\alpha)^\top \Gamma_s (\alpha_s + h_s \tilde{X}_s^\alpha) ds \right].$$

Since  $\alpha \in \mathcal{A}$ ,  $X^\alpha$  satisfies (2.5), and so  $\mathbb{E} \left[ \sup_{t \leq T} |\tilde{X}_t^\alpha|^2 \right] < \infty$ . An application of the dominated convergence theorem on the left term combined with the monotone convergence theorem on the right term, recall that  $\Gamma$  is  $\mathbb{S}_+^d$ -valued, yields, as  $k \rightarrow \infty$ ,

$$\mathbb{E} \left[ |\tilde{X}_T^\alpha|^2 \right] = \Gamma_0 |\tilde{X}_0^\alpha|^2 + \mathbb{E} \left[ \int_0^T (\alpha_s + h_s \tilde{X}_s^\alpha)^\top \Gamma_s (\alpha_s + h_s \tilde{X}_s^\alpha) ds \right].$$

Since  $\Gamma_s$  is positive definite for any  $s \leq T$ , we obtain that the optimal strategy  $\alpha^*(\xi)$  is given by (A.3) and the optimal value of (A.2) is equal to

$$\tilde{V}(\xi) = \Gamma_0 |\tilde{X}_0^{\alpha^*(\xi)}|^2 = \Gamma_0 |X_0 - \xi e^{-\int_0^T r(s) ds}|^2,$$

which gives (A.4).  $\square$

We next address the admissibility of the candidate for the optimal control.

**Lemma A.2.** *Assume that there exists a solution triplet  $(\Gamma, Z^1, Z^2) \in \mathbb{S}_{\mathbb{F}}^\infty([0, T], \mathbb{R}) \times L_{\mathbb{F}}^{2,loc}([0, T], \mathbb{R}^d) \times L_{\mathbb{F}}^{2,loc}([0, T], \mathbb{R}^N)$  to the Riccati BSDE (3.2) such that (3.3) holds for some  $p > 2$  and a constant  $a(p)$  given by (3.4). Then, for any  $\xi \in \mathbb{R}$ , there exists an admissible control process  $\alpha^*(\xi)$  satisfying (A.3).*

*Proof.* Fix  $\xi \in \mathbb{R}$ . We first prove that there exists a control  $\alpha^*(\xi)$  satisfying (A.3). For this, we prove that the corresponding wealth equation (2.4) admits a solution. As in the proof of Lemma A.1, it is enough to consider the modified equation

$$d\tilde{X}_t^* = (r(t)\tilde{X}_t^* + \lambda_t^\top A_t \tilde{X}_t^*) dt + (A_t \tilde{X}_t^*)^\top dB_t, \quad \tilde{X}_0^* = x_0 - \xi e^{-\int_0^T r(s) ds},$$

where  $A_t = -(\lambda_t + Z_t^1 + CZ_t^2)$ , and then set  $X_t^* = \tilde{X}_t^* + \xi e^{-\int_t^T r(s)ds}$ . By virtue of Itô's lemma the unique continuous solution is given by

$$\tilde{X}_t^* = \tilde{X}_0^* \exp \left( \int_0^t (r(s) + \lambda_s^\top A_s - \frac{A_s^\top A_s}{2}) ds + \int_0^t A_s^\top dB_s \right).$$

Setting  $\alpha_t^*(\xi) := A_t \tilde{X}_t^*$ , we obtain that  $\alpha^*(\xi)$  satisfies (A.3) with the controlled wealth  $X^{\alpha(\xi)*} = X^*$ . The crucial step is now to obtain the admissibility condition (2.5). For that purpose, observe by virtue of (3.3), that the Doléans-Dade exponential  $\mathcal{E}(\int_0^t A_s^\top dB_s)$  satisfies Novikov's condition, and is therefore a true martingale. Whence, successive applications of the inequality  $ab \leq (a^2 + b^2)/2$  and Doob's maximal inequality yield, for some constant  $K > 0$  which may vary from line to line,

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, T]} |\tilde{X}_t^*|^p \right] &\leq K \mathbb{E} \left[ \sup_{t \in [0, T]} \left| e^{\int_0^t (r(s) + \lambda_s^\top A_s) ds} \right|^{2p} \right] + K \mathbb{E} \left[ \sup_{t \in [0, T]} \left| e^{-\int_0^t \frac{A_s^\top A_s}{2} ds + \int_0^t A_s^\top dB_s} \right|^{2p} \right] \\ &\leq K \mathbb{E} \left[ e^{\int_0^T 2p |\lambda_s^\top A_s| ds} \right] + K \mathbb{E} \left[ e^{-p \int_0^T A_s^\top A_s ds + 2p \int_0^T A_s^\top dB_s} \right] \\ &= K(1 + 2), \end{aligned}$$

which is finite since

$$1 \leq \mathbb{E} \left[ \exp \left( a(p) \int_0^T (|\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2) ds \right) \right] < \infty,$$

and, by virtue of the Cauchy-Schwarz inequality,

$$\begin{aligned} 2 &\leq \left( \mathbb{E} \left[ e^{(8p^2 - 2p) \int_0^T A_s^\top A_s ds} \right] \right)^{1/2} \left( \mathbb{E} \left[ e^{-8p^2 \int_0^T A_s^\top A_s ds + 4p \int_0^T A_s^\top dB_s} \right] \right)^{1/2} \\ &\leq \left( \mathbb{E} \left[ e^{a(p) \int_0^T (|\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2) ds} \right] \right)^{1/2} \times 1 < \infty, \end{aligned}$$

where we used Jensen's inequality to bound

$$A_s^\top A_s = |\lambda_s + Z_s^1 + CZ_s^2|^2 \leq 3(|\lambda_s|^2 + |Z_s^1|^2 + |CZ_s^2|^2) \leq 3(1 + |C|^2)(|\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2),$$

together with assumption (H2) and Novikov's condition to the Doléans-Dade exponential  $\mathcal{E}(4p \int_0^T A_s^\top dB_s)$ . Finally, to get that  $\alpha^*(\xi)$  is admissible, we are left to prove that  $\alpha^*(\xi) \in$

$L^2_{\mathbb{F}}([0, T], \mathbb{R}^d)$ . Let  $2/p + 1/\hat{q} = 1$ , by Hölder's inequality we obtain

$$\begin{aligned}
\mathbb{E} \left[ \int_0^T |\alpha_s^*(\xi)|^2 ds \right] &= \mathbb{E} \left[ \int_0^T |A_s \tilde{X}_s^*|^2 ds \right] \\
&\leq \mathbb{E} \left[ \sup_{t \in [0, T]} |\tilde{X}_t^*|^2 \int_0^T |A_s|^2 ds \right] \\
&\leq \left( \mathbb{E} \left[ \sup_{t \in [0, T]} |\tilde{X}_t^*|^p \right] \right)^{2/p} \left( \mathbb{E} \left[ \left( \int_0^T |A_s|^2 ds \right)^{\hat{q}} \right] \right)^{1/\hat{q}} \\
&\leq C \left( \mathbb{E} \left[ \sup_{t \in [0, T]} |\tilde{X}_t^*|^p \right] \right)^{2/p} \left( \mathbb{E} \left[ \left( \int_0^T (|\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2) ds \right)^{\hat{q}} \right] \right)^{1/\hat{q}} \\
&< \infty,
\end{aligned}$$

where the last term is finite due to condition (3.3) and the inequality  $|z|^q \leq c_q e^{|z|}$ . The proof is complete.  $\square$

Finally, combining the above, we deduce the solution for the outer optimization problem (2.6) under a non-degeneracy condition on the solution  $\Gamma$  to the Riccati BSDE, yielding Theorem 3.1.

*Proof of Theorem 3.1.* From Lemmas A.1 and A.2, we have that the max-min problem (A.1) (which is equivalent to the Markowitz problem (2.6)) is equivalent to

$$\max_{\eta \in \mathbb{R}} J(\eta), \quad \text{with } J(\eta) = \Gamma_0 |X_0 - (m - \eta) e^{-\int_0^T r(s) ds}|^2 - \eta^2.$$

Furthermore, condition (H1):  $\Gamma_0 < e^{2 \int_0^T r(s) ds}$ , ensures that the quadratic function  $J$  is strictly concave. This yields that the maximum is achieved from the first-order condition  $J'(\eta^*) = 0$ , which gives

$$\eta^* = \frac{\Gamma_0 e^{-\int_0^T r(s) ds} \left( x_0 - m e^{-\int_0^T r(s) ds} \right)}{1 - \Gamma_0 e^{-2 \int_0^T r(s) ds}},$$

and thus  $\xi^* = m - \eta^*$  is given by (3.6). We conclude that the optimal control is equal to  $\alpha^* = \alpha^*(\xi^*)$  as in (3.5), and by (A.1), the optimal value of (2.6) is equal to  $V(m) = \tilde{V}(\xi^*) - (\eta^*)^2$ , given by (3.7).  $\square$

## B Proofs of some technical lemmas

### B.1 Reminder on resolvents of integral operators

**Lemma B.1.** *Let  $K$  satisfy (5.1) and  $L \in L^2([0, T]^2, \mathbb{R}^{N \times N})$ . Then,  $K \star L$  satisfies (5.1). Furthermore, if  $L$  satisfies (5.1), then,  $(s, u) \mapsto (K \star L^*)(s, u)$  is continuous.*

*Proof.* An application of the Cauchy-Schwarz inequality yields the first part. The second part follows along the same lines as in the proof of [Abi Jaber \(2019c, Lemma 3.2\)](#).  $\square$

For a kernel  $K \in L^2([0, T]^2, \mathbb{R}^{N \times N})$ , we define its resolvent  $R_T \in L^2([0, T]^2, \mathbb{R}^{N \times N})$  by the unique solution to

$$R_T = K + K \star R_T, \quad K \star R_T = R_T \star K. \quad (\text{B.1})$$

In terms of integral operators, this translates into

$$\mathbf{R}_T = \mathbf{K} + \mathbf{K} \mathbf{R}_T, \quad \mathbf{K} \mathbf{R}_T = \mathbf{R}_T \mathbf{K}.$$

In particular, if  $K$  admits a resolvent,  $(\text{Id} - \mathbf{K})$  is invertible and

$$(\text{Id} - \mathbf{K})^{-1} = \text{Id} + \mathbf{R}_T, \quad (\text{B.2})$$

where  $\text{Id}$  denotes the identity operator, i.e.  $(\text{Id}f) = f$  for all  $f \in L^2([0, T], \mathbb{R}^N)$ .

The following lemma establishes the existence of resolvents for the two classes of kernels introduced above.

**Lemma B.2.** *Let  $K \in L^2([0, T]^2, \mathbb{R}^{N \times N})$ .  $K$  admits a resolvent if either one of the following conditions hold:*

- (i)  *$K$  is a Volterra kernel of continuous and bounded type in  $L^2$  in the sense of Definition 5.2. In this case, the resolvent is again a Volterra kernel of continuous and bounded type.*
- (ii)  *$K$  is symmetric nonpositive in the sense of Definition 5.1 and  $(s, u) \mapsto K(s, u)$  is continuous.*

*Proof.* (i) follows from [Gripenberg et al. \(1990, Lemma 9.3.3, Theorem 9.5.5\(i\)\)](#). (ii) follows from an application of Mercer's theorem, see [Abi Jaber \(2019c, Section 2.1\)](#).  $\square$

### B.2 Proof of Lemma 5.6

Fix  $t \leq T$ . We start by proving that  $\Psi_t$  is well defined and is a bounded linear operator from  $L^2([0, T], \mathbb{R}^N)$  to  $L^2([0, T], \mathbb{R}^N)$ . First, since  $K$  is a Volterra kernel of continuous

and bounded type in  $L^2$ , so is  $\hat{K}$ , and Lemma B.2-(i) yields the existence of its resolvent  $\hat{R}$  such that

$$\sup_{s \leq T} \int_0^T |\hat{R}(s, u)| ds < \infty, \quad \sup_{u \leq T} \int_0^T |\hat{R}(s, u)| du < \infty. \quad (\text{B.3})$$

In particular, denoting by  $\hat{\mathbf{R}}$  the integral operator induced by  $\hat{R}$ , we obtain that  $(\text{Id} - \hat{\mathbf{K}})$  is invertible with an inverse given by  $(\text{Id} - \hat{\mathbf{K}})^{-1} = \text{Id} + \hat{\mathbf{R}}$ , recall (B.2). Next, we prove that  $(\text{Id} + 2\Theta\tilde{\Sigma}_t\Theta^\top)$  is invertible. It follows from (5.7) that

$$\tilde{\Sigma}_t = (\text{Id} + \hat{\mathbf{R}})\Sigma_t(\text{Id} + \hat{\mathbf{R}})^* = \Sigma_t + \Sigma_t\hat{\mathbf{R}}^* + \hat{\mathbf{R}}\Sigma_t + \hat{\mathbf{R}}\Sigma_t\hat{\mathbf{R}}^*.$$

Whence,  $\tilde{\Sigma}_t$  is an integral operator generated by the kernel

$$\tilde{\Sigma}_t = \Sigma_t + \Sigma_t \star \hat{R}^* + \hat{R} \star \Sigma_t + \hat{R} \star \Sigma_t \star \hat{R}^*. \quad (\text{B.4})$$

Since  $K$  satisfies (5.1) and  $(U - 2C^\top C) \in \mathbb{S}_+^N$ ,  $\Sigma_t$  defined in (5.8) is clearly a symmetric nonnegative kernel. Combined with (B.4), we get that  $\tilde{\Sigma}_t$  is symmetric nonnegative. Successive applications of Lemma B.1 yield that  $(s, u) \mapsto \tilde{\Sigma}_t(s, u)$  is continuous. Therefore,  $(-2\Theta\tilde{\Sigma}_t\Theta^\top)$  is symmetric nonpositive and continuous so that an application of Lemma B.2-(ii) yields the existence of its resolvent  $R_t^\Theta$ . In particular,  $(\text{Id} + 2\Theta\tilde{\Sigma}_t\Theta^\top)$  is invertible with an inverse given by  $(\text{Id} + \mathbf{R}_t^\Theta)$ , recall (B.2). Combining the above, we get that  $\Psi_t$  is well-defined, and satisfies

$$\begin{aligned} \Psi_t &= -(\text{Id} + \hat{\mathbf{R}})^*\Theta^\top(\text{Id} + \mathbf{R}_t^\Theta)\Theta(\text{Id} + \hat{\mathbf{R}}) \\ &= -\Theta^\top\Theta\text{Id} - \hat{\mathbf{R}}^*\Theta^\top\Theta - \Theta^\top\Theta\hat{\mathbf{R}} - \hat{\mathbf{R}}^*\Theta^\top\mathbf{R}_t^\Theta\Theta - \Theta^\top\mathbf{R}_t^\Theta\Theta\hat{\mathbf{R}} \\ &\quad - \hat{\mathbf{R}}^*\Theta^\top\mathbf{R}_t^\Theta\Theta\hat{\mathbf{R}} - \hat{\mathbf{R}}^*\Theta^\top\Theta\hat{\mathbf{R}} - \Theta^\top\mathbf{R}_t^\Theta\Theta, \end{aligned} \quad (\text{B.5})$$

showing that  $\Psi_t$  is a bounded operator.

(i): From (B.5), we see that  $(\Theta^\top\Theta\text{Id} + \Psi_t)$  is an integral operator whose kernel is of the form

$$\begin{aligned} \psi_t &= -\hat{R}^*\Theta^\top\Theta - \Theta^\top\Theta\hat{R} - \hat{R}^*\Theta^\top \star R_t^\Theta\Theta - \Theta^\top R_t^\Theta \star \Theta\hat{R} \\ &\quad - \hat{R}^*\Theta^\top \star R_t^\Theta \star \Theta\hat{R} - \hat{R}^*\Theta^\top \star \Theta\hat{R} - \Theta^\top R_t^\Theta\Theta. \end{aligned}$$

Then, from Abi Jaber (2019c, Lemma C.1) we get that

$$\sup_{t \leq T} \int_{[0, T]^2} |R_t^\Theta(s, u)|^2 ds du < \infty,$$

which, combined with (B.3) ensures (5.10).



(ii): Fix  $f \in L^2([0, T], \mathbb{R}^N)$  and  $t \leq T$ . We first argue that

$$R_t^\Theta(t, \cdot) = 0 \text{ and } \hat{R}(s, u) = 0, \quad \text{for any } s < u. \quad (\text{B.6})$$

Indeed, since  $\hat{K}$  is a Volterra kernel, its resolvent  $\hat{R}$  is also a Volterra kernel so that  $R(s, u) = 0$  whenever  $s < u$ . This, combined with the fact that  $\Sigma_t(t, \cdot) = 0$  and (B.4), yields that  $\tilde{\Sigma}_t(t, \cdot) = 0$ , so that  $R_t^\Theta(t, \cdot) = 0$  by virtue of the resolvent equation (B.1). Using the relations (B.6), we compute

$$\begin{aligned} (\Theta^\top \Theta \hat{R})(f1_t)(t) &= \Theta^\top \Theta \int_0^T \hat{R}(t, s) f(s) 1_t(s) ds = 0, \\ (\Theta^\top R_t^\Theta \Theta)(f1_t)(t) &= \Theta^\top \int_0^T R_t^\Theta(t, s) \Theta f(s) 1_t(s) ds = 0, \\ (\Theta^\top R_t^\Theta \Theta \hat{R})(f1_t)(t) &= \Theta^\top \int_0^T \int_0^T R_t^\Theta(t, u) \Theta \hat{R}(u, s) f(s) 1_t(s) du ds = 0. \end{aligned} \quad (\text{B.7})$$

Thus, (B.7) combined with (B.5) and the resolvent's relations  $\hat{R} = \hat{K} + \hat{K} \hat{R}$  and  $\hat{R}^* = \hat{K}^* + \hat{K}^* \hat{R}^*$  yield

$$\begin{aligned} -(\Theta^\top \Theta \text{Id} + \Psi_t)(f1_t)(t) &= (\hat{R}^* \Theta^\top \Theta + \hat{R}^* \Theta^\top R_t^\Theta \Theta + \hat{R}^* \Theta^\top R_t^\Theta \Theta \hat{R} + \hat{R}^* \Theta^\top \Theta \hat{R})(f1_t)(t) \\ &= -(\hat{K}^* \Psi_t)(f1_t)(t) \end{aligned}$$

which proves the second claim (ii).

(iii): Under (5.9), [Abi Jaber \(2019c, Lemma 3.2\)](#) yields that  $t \mapsto \Sigma_t$  is strongly differentiable on  $[0, T]$  with a derivative given by  $t \mapsto \dot{\Sigma}_t$  induced by the kernel (5.12). Whence, it follows from (5.7), that  $t \mapsto \tilde{\Sigma}_t$  is also differentiable such that  $\dot{\tilde{\Sigma}}_t = (\text{Id} - \hat{K})^{-1} \dot{\Sigma}_t (\text{Id} - \hat{K})^{-*}$ . Thus, (5.6) yields that  $t \mapsto \Psi_t$  is strongly differentiable with a derivative given by

$$\begin{aligned} \dot{\Psi}_t &= 2(\text{Id} - \hat{K})^{-*} \Theta^\top (\text{Id} + 2\Theta \tilde{\Sigma}_t \Theta^\top)^{-1} \Theta \dot{\tilde{\Sigma}}_t \Theta^\top (\text{Id} + 2\Theta \tilde{\Sigma}_t \Theta^\top)^{-1} \Theta (\text{Id} - \hat{K})^{-1} \\ &= 2\Psi_t \dot{\tilde{\Sigma}}_t \Psi_t. \end{aligned}$$

Finally, evaluating (5.8) at  $t = T$ , yields that  $\Sigma_T(s, u) = 0$  for all  $s, u \leq T$ , leading to  $\Sigma_T = \mathbf{0}$  so that  $\Psi_T = -(\text{Id} - \hat{K})^{-*} \Theta^\top \Theta (\text{Id} - \hat{K})^{-1}$ . This proves (5.11).

### B.3 Proof of Lemma 5.8

We start with a lemma to bound the kernel  $\tilde{\Sigma}$ .

**Lemma B.3.** *Let  $f(\Theta) = D - 2\eta C^\top \Theta$  and assume that  $|f(\Theta)| \times \|K\|_{L^2([0, T]^2)}^2 < 1$ . Then there exists a constant  $c > 0$  such that*

$$\sup_{t \leq T} \|\tilde{\Sigma}_t\|_{L^2([0, T]^2)}^2 \leq c(1 + \hat{\kappa}(\Theta)), \quad (\text{B.8})$$

where  $\hat{\kappa}$  is defined as

$$\hat{\kappa}(\Theta) = \left( \frac{|f(\Theta)| \times \|K\|_{L^2([0,T]^2)}^2}{1 - |f(\Theta)| \times \|K\|_{L^2([0,T]^2)}^2} \right)^4. \quad (\text{B.9})$$

*Proof.* Let  $\hat{R}$  denote the resolvent kernel of  $\hat{K} = Kf(\Theta)$  as in the proof of Lemma 5.6. First note that the relation  $(\text{Id} - \hat{K})^{-1} = \text{Id} + \hat{R}$  yields

$$\begin{aligned} \|\tilde{\Sigma}_t\|_{L^2([0,T]^2)}^2 &= \|(\text{Id} - \hat{K})^{-1} \star \Sigma_t \star (\text{Id} - \hat{K})^{-*}\|_{L^2([0,T]^2)}^2 \\ &= \|\Sigma_t + \hat{R} \star \Sigma_t + \Sigma_t \star \hat{R} + \hat{R} \star \Sigma_t \star \hat{R}\|_{L^2([0,T]^2)}^2 \\ &\leq 2^3 \left( \|\Sigma_t\|_{L^2([0,T]^2)}^2 + \|\hat{R} \star \Sigma_t\|_{L^2([0,T]^2)}^2 \right. \\ &\quad \left. + \|\Sigma_t \star \hat{R}\|_{L^2([0,T]^2)}^2 + \|\hat{R} \star \Sigma_t \star \hat{R}\|_{L^2([0,T]^2)}^2 \right). \end{aligned}$$

An application of the Cauchy-Schwarz inequality combined with Tonelli's theorem implies that

$$\|K \star H\|_{L^2([0,T]^2)} \leq \|K\|_{L^2([0,T]^2)} \|H\|_{L^2([0,T]^2)}, \quad K, H \in L^2([0,T]^2, \mathbb{R}^{N \times N}), \quad (\text{B.10})$$

so that

$$\begin{aligned} \|\tilde{\Sigma}_t\|_{L^2([0,T]^2)}^2 &\leq 2^3 \left( \|\Sigma_t\|_{L^2([0,T]^2)}^2 + \|\hat{R} \star \Sigma_t\|_{L^2([0,T]^2)}^2 + \|\Sigma_t \star \hat{R}\|_{L^2([0,T]^2)}^2 + \|\hat{R} \star \Sigma_t \star \hat{R}\|_{L^2([0,T]^2)}^2 \right) \\ &\leq 2^3 \|\Sigma_t\|_{L^2([0,T]^2)}^2 \left( 1 + \|\hat{R}\|_{L^2([0,T]^2)}^2 + \|\hat{R}\|_{L^2([0,T]^2)}^4 \right) \\ &\leq c \|\Sigma_t\|_{L^2([0,T]^2)}^2 \left( 1 + \|\hat{R}\|_{L^2([0,T]^2)}^4 \right), \end{aligned}$$

where  $c > 0$  is a constant independent of  $\Sigma$  and  $\hat{R}$ . Thus, to obtain (B.8) it is enough to show that  $\|\hat{R}\|_{L^2([0,T]^2)}^2 \leq \left( \frac{|f(\Theta)| \times \|K\|_{L^2([0,T]^2)}^2}{1 - |f(\Theta)| \times \|K\|_{L^2([0,T]^2)}^2} \right)^2$ . For this, note that applying successive Picard's iteration to  $\hat{R} = \hat{K} + \hat{K} \star R$  yields

$$\hat{R}(s, u) = \sum_{n=1}^{\infty} \hat{K}^{\star n}(s, u) = \sum_{n=1}^{\infty} (Kf(\Theta))^{\star n}(s, u), \quad (\text{B.11})$$

where  $\hat{K}^{\star n}$  is the  $(n)$ -fold  $\star$ -product of  $\hat{K}$  by itself. Combining (B.10) and (B.11) together

with the submultiplicativity of the Frobenius norm yields

$$\begin{aligned}
\|\hat{R}\|_{L^2([0,T]^2)}^2 &\leq \sum_{1 \leq n, m \leq \infty} \int_0^T \int_0^T |(K(s, u)f(\Theta))^{\star n}| |(K(s, u)f(\Theta))^{\star m}| ds du \\
&\leq \sum_{1 \leq n, m \leq \infty} |f(\Theta)|^{n+m} \|K^{\star n}\|_{L^2([0,T]^2)}^2 \|K^{\star m}\|_{L^2([0,T]^2)}^2 \\
&= \left( \sum_{n=1}^{\infty} |f(\Theta)|^n \|K^{\star n}\|_{L^2([0,T]^2)}^2 \right)^2 \\
&\leq \left( \sum_{n=1}^{\infty} |f(\Theta)|^n \|K\|_{L^2([0,T]^2)}^{2n} \right)^2 \\
&\leq \left( \frac{|f(\Theta)| \times \|K\|_{L^2([0,T]^2)}^2}{1 - |f(\Theta)| \times \|K\|_{L^2([0,T]^2)}^2} \right)^2.
\end{aligned}$$

This proves the desired inequality on  $\hat{R}$  and the claimed inequality (B.8) follows.  $\square$

We can now complete the proof of Lemma 5.8.

*Proof of Lemma 5.8.* Fix  $s \leq T$  and  $\Theta \in \mathbb{R}^{d \times N}$ . We first note that

$$|\lambda_s|^2 + |Z_s^1|^2 + |Z_s^2|^2 = |\Theta g_s(s)|^2 + 4 |((\Psi_s \mathbf{K} \eta)^* g_s)(s)|^2.$$

Using 5.6-(i), and denoting by  $\psi_s^{\text{op}}$  the operator induced by the kernel  $\psi_s$  there, we write

$$\begin{aligned}
|((\Psi_s \mathbf{K} \eta)^* g_s)(s)|^2 &= |-(\Theta^\top \Theta \mathbf{K} \eta)^* g_s(s) + ((\psi_s^{\text{op}} \mathbf{K} \eta)^* g_s)(s)|^2 \\
&= |\mathbf{1} + \mathbf{2}|^2 \\
&\leq 2(|\mathbf{1}|^2 + |\mathbf{2}|^2).
\end{aligned}$$

An application of the Cauchy-Schwarz inequality combined with (5.9) leads to

$$|\mathbf{1}|^2 = \left| - \int_0^T \eta^\top K(z, s)^\top \Theta^\top \Theta g_s(z) dz \right|^2 \leq |\eta|^2 |\Theta \Theta^\top|^2 \sup_{u' \leq T} \int_0^T |K(z, u')|^2 du' \int_0^T |g_s(u)|^2 du. \quad (\text{B.12})$$

Similarly,

$$\begin{aligned}
|\mathbf{2}|^2 &= \left( \int_0^T \eta^\top \left( \int_0^T K(r, s)^\top \psi_s(r, z) dr \right) g_s(z) dz \right)^2 \\
&\leq |\eta|^2 \left( \int_0^T \int_0^T |K(r, s)|^2 |\psi_s(r, z)|^2 dr dz \right) \left( \int_0^T |g_s(z)|^2 dz \right) \\
&\leq |\eta|^2 \sup_{u' \leq T} \int_0^T |K(r, u')|^2 dr \left( \int_0^T \int_0^T |\psi_s(r, z)|^2 dr dz \right) \left( \int_0^T |g_s(z)|^2 dz \right),
\end{aligned}$$

where we stress that  $\psi_s$  is the only term on the right hand side depending on  $\Theta$ . Let us now show that there exists a constant  $c > 0$  independent of  $\Theta$  such that

$$\sup_{s \in [0, T]} \int_0^T \int_0^T |\psi_s(r, z)|^2 dr dz \leq c|\Theta|^2(1 + |\Theta|^4 \hat{\kappa}(\Theta)), \quad (\text{B.13})$$

where  $\hat{\kappa}$  is defined as in (B.9). Recall from (B.5) that we have

$$\begin{aligned} \psi_t = & -\hat{R}^* \Theta^\top \Theta - \Theta^\top \Theta \hat{R} - \hat{R}^* \Theta^\top \star R_t^\Theta \Theta - \Theta^\top R_t^\Theta \star \Theta \hat{R} \\ & - \hat{R}^* \Theta^\top \star R_t^\Theta \star \Theta \hat{R} - \hat{R}^* \Theta^\top \star \Theta \hat{R} - \Theta^\top R_t^\Theta \Theta. \end{aligned}$$

Thus, recalling (B.3), there exists a constant  $c > 0$  independent of  $\Theta$  such that

$$\sup_{s \in [0, T]} \int_0^T \int_0^T |\psi_s(r, z)|^2 dr dz \leq c|\Theta|^2 \left( 1 + \sup_{t \in [0, T]} \int_0^T \int_0^T |R_t^\Theta(s, u)|^2 ds du \right). \quad (\text{B.14})$$

To obtain (B.13), it is enough to show that

$$\sup_{t \in [0, T]} \int_0^T \int_0^T |R_t^\Theta(s, u)|^2 ds du \leq c|\Theta|^4(1 + \hat{\kappa}(\Theta)), \quad (\text{B.15})$$

for some constant  $c > 0$  not depending on  $\Theta$  and  $\hat{\kappa}$  defined in (B.9). For this recall that  $\mathbf{R}_t^\Theta$  is the resolvent of  $-2\Theta\tilde{\Sigma}_t\Theta^\top$  which implies that  $\mathbf{R}_t^\Theta = (\text{Id} + 2\Theta\tilde{\Sigma}_t\Theta^\top)^{-1} - \text{Id}$ . Since, for each  $t \leq T$ ,  $\Theta\tilde{\Sigma}_t\Theta^\top$  is a positive symmetric operator on  $L^2([0, T], \mathbb{R}^d)$  induced by a continuous kernel, an application of Mercer's theorem, see [Shorack and Wellner \(2009, Theorem 1, p.208\)](#), yields the existence of a countable orthonormal basis  $(e_{t,\Theta}^n)_{n \geq 1}$  of  $L^2([0, T], \mathbb{R}^d)$  such that

$$2\Theta\tilde{\Sigma}_t(s, u)\Theta = \sum_{n \geq 1} \lambda_{t,\Theta}^n e_{t,\Theta}^n(s) e_{t,\Theta}^n(u)^\top,$$

where  $\lambda_{t,\Theta}^n \geq 0$ , for all  $n \geq 1$ . Consequently

$$R_t^\Theta(s, u) = \sum_{n \geq 1} \frac{-\lambda_{t,\Theta}^n}{1 + \lambda_{t,\Theta}^n} e_{t,\Theta}^n(s) e_{t,\Theta}^n(u)^\top,$$

which yields

$$\begin{aligned} \int_0^T \int_0^T |R_t^\Theta(s, u)|^2 ds du &= \sum_{n \geq 1} \frac{(\lambda_{t,\Theta}^n)^2}{(1 + \lambda_{t,\Theta}^n)^2} \\ &\leq \sum_{n \geq 1} (\lambda_{t,\Theta}^n)^2 = \int_0^T \int_0^T |2\Theta\tilde{\Sigma}_t(s, u)\Theta^\top|^2 ds du \\ &\leq 4|\Theta|^4 \sup_{t \leq T} \int_0^T \int_0^T |\tilde{\Sigma}_t(s, u)|^2 ds du \\ &\leq c|\Theta|^4(1 + \hat{\kappa}(\Theta)), \end{aligned}$$

where the last inequality comes from Lemma B.3. Consequently, inequality (B.15) combined with (B.14) yield inequality (B.13). Finally, the claimed bound (5.18) follows by recollecting inequalities (B.13) and (B.12).

#### B.4 Proof of Lemma 5.10

*Proof.* Recalling the decomposition (5.21), the process  $Z$  admits the following Karhunen-Loeve representation

$$Z(s, u) = \sum_{n \geq 1} \xi_n e^n(s, u), \quad s, u \in [0, T]^2, \quad (\text{B.16})$$

where  $(\xi_n)_{n \geq 1}$  is a sequence of independent Gaussian random variables with mean  $\mu_n = \langle \mu, e^n \rangle_{L^2([0, T]^2, \mathbb{R}^{2N})}$  and variance  $\lambda^n$ , for each  $n \in \mathbb{N}$ . Now observe that the representation (B.16) combined with the orthogonality of  $(e_n)_{n \geq 1}$  in  $L^2([0, T]^2, \mathbb{R}^{2N})$  yields

$$a \int_0^T \left( |g_s(s)|^2 + \int_0^T |g_s(u)|^2 du \right) ds = a \|Z\|_{L^2([0, T]^2, \mathbb{R}^{2N})}^2 = \sum_{n \geq 1} a \xi_n^2,$$

so that the independence of  $(\xi_n)_{n \geq 1}$  leads to

$$\begin{aligned} \mathbb{E} \left[ \exp \left( a \int_0^T \left( |g_s(s)|^2 + \int_0^T |g_s(u)|^2 du \right) ds \right) \right] &= \mathbb{E} \left[ \exp \left( \sum_{n \geq 1} a \xi_n^2 \right) \right] \\ &= \prod_{n \geq 1} \mathbb{E} [\exp (a \xi_n^2)] \\ &= \prod_{n \geq 1} \frac{e^{\frac{a \mu_n^2}{1 - 2a \lambda^n}}}{\sqrt{1 - 2a \lambda^n}} \end{aligned} \quad (\text{B.17})$$

where the last equality follows from the fact that  $\xi_n^2$  is chi-squared distributed and  $0 < 1 - 2a \lambda^n < 1 - 2a \lambda^1$  by hypothesis. We now argue that the right hand side of (B.17) is finite. For the denominator, due to the inequality

$$1 - 2a \sum_{n \geq 1} \lambda^n \leq \prod_{n \geq 1} (1 - 2a \lambda^n) \leq \exp \left( 2a \sum_{n \geq 1} \lambda^n \right),$$

and  $\sum_{n \geq 1} \lambda^n < \infty$ , we obtain that  $\prod_{n \geq 1} (1 - 2a \lambda^n)^{-1} < \infty$ . For the numerator, since  $\lambda^n \rightarrow 0$ , as  $n \rightarrow \infty$ ,  $\left( \frac{1}{1 - 2a \lambda^n} \right)_{n \geq 1}$  is uniformly bounded by a constant  $c > 0$  so that an

application of Parseval's identity yields

$$\begin{aligned}
\prod_{n \geq 1} \exp \left( \frac{a\mu_n^2}{1 - 2a\lambda^n} \right) &\leq \exp \left( ca \|\mu\|_{L^2([0,T]^2, \mathbb{R}^{2N})}^2 \right) \\
&= \exp \left( ca \left( \int_0^T \int_0^T \left( \frac{1}{T^2} |g_0(s)|^2 + |g_0(u)|^2 \right) ds du \right) \right) \\
&= \exp \left( ca \left( T + \frac{1}{T} \right) \|g_0\|_{L^2([0,T], \mathbb{R}^N)}^2 \right) < \infty.
\end{aligned}$$

The proof is complete.  $\square$

## C Additional proof for the martingale property

For completeness, we adapt [Abi Jaber et al. \(2019, Lemma 7.3\)](#) to the multi-dimensional setting to prove that the local martingale

$$M_t = M_0 \mathcal{E} \left( - \int_0^t \sum_{i=1}^d \psi^i(T-s) \nu_i \sqrt{V_s^i} dW_s^i \right).$$

is a true martingale. For this we set  $U = \int_0^\cdot V_s ds$  and we observe that, thanks to stochastic Fubini's theorem, integrating (4.1) yields

$$U_t^i = \int_0^t g_0^i(s) ds + \int_0^t K_i(t-s) Z_s^i ds$$

with

$$Z_t^i = \int_0^t (DV_s)_i ds + \int_0^t \nu_i \sqrt{V_s^i} dW_s^i.$$

*Proof.* Since  $M$  is a nonnegative local martingale, it is a supermartingale by Fatou's lemma. Whence to obtain the true martingality it suffices to show that  $\mathbb{E}[M_T] = 1$  for any  $T \in \mathbb{R}_+$ . To this end, fix  $T > 0$  and define the stopping times  $\tau_n = \inf\{t \geq 0: \int_0^t V_s^i ds > n \text{ for some } i \leq d\} \wedge T$ . Novikov's condition, recall that  $\psi$  is bounded on  $[0, T]$  being continuous, yields that  $M^{\tau_n} = M_{\tau_n \wedge \cdot}$  is a uniformly integrable martingale for each  $n$ . Whence,

$$1 = M_0^{\tau_n} = \mathbb{E}_{\mathbb{P}} [M_T^{\tau_n}] = \mathbb{E}_{\mathbb{P}} [M_T \mathbf{1}_{\tau_n \geq T}] + \mathbb{E}_{\mathbb{P}} [M_{\tau_n} \mathbf{1}_{\tau_n < T}],$$

where we made the dependence of the expectation on  $\mathbb{P}$  explicit. Since  $\mathbb{E}_{\mathbb{P}} [M_T \mathbf{1}_{\tau_n \geq T}] \rightarrow \mathbb{E}_{\mathbb{P}} [M_T]$  as  $n \rightarrow \infty$ , by dominated convergence, in order to get that  $\mathbb{E}_{\mathbb{P}} [M_T] = 1$ , it suffices to prove that

$$\mathbb{E}_{\mathbb{P}} [M_{\tau_n} \mathbf{1}_{\tau_n < T}] \rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{C.1}$$

To this end, since  $M^{\tau_n}$  is a martingale, we may define probability measures  $\mathbb{Q}^n$  by

$$\frac{d\mathbb{Q}^n}{d\mathbb{P}} = M_{\tau_n}^{\tau_n}.$$

By Girsanov's theorem, the process  $W^n = (W^{n,1}, \dots, W^{n,d})$  defined by

$$W^{n,i} = W^i + \int_0^{\cdot} \mathbf{1}_{s \leq \tau_n} \psi^i(T-s) \nu_i \sqrt{V_s^i} ds, \quad i = 1, \dots, d,$$

is a Brownian motion under  $\mathbb{Q}^n$ . Furthermore, under  $\mathbb{Q}^n$ , we have

$$\begin{aligned} U_t^i &= \int_0^t g_0^i(s) ds + \int_0^t K_i(t-s) Z_s^{n,i} ds \\ Z_t^{n,i} &= \int_0^t ((DV_s)_i - \mathbf{1}_{s \leq \tau_n} \psi^i(T-s) \nu_i^2 V_s^i) ds + \int_0^t \nu_i \sqrt{V_s^i} dW_s^{n,i}. \end{aligned}$$

and we observe that, due to the boundedness of  $\psi$ , the drift of  $Z^n$  under  $\mathbb{Q}^n$  satisfy a linear growth condition in  $U$  for some constant  $\kappa_L$  independent of  $n$ . An application of the generalized Grönwall inequality for convolution equations would yield the moment bound

$$\mathbb{E}_{\mathbb{Q}^n} [|U_T|^2] \leq \eta(\kappa_L, T, K, g_0),$$

where  $\eta(\kappa_L, T, K, g_0)$  does not depend on  $n$ , see for instance [Abi Jaber \(2019b, Lemma 3.1\)](#). We then get by an application of Chebyshev's inequality

$$\begin{aligned} \mathbb{E}_{\mathbb{P}} [M_{\tau_n} \mathbf{1}_{\tau_n \leq T}] &= \mathbb{Q}^n(\tau_n < T) \\ &\leq \sum_{i=1}^d \mathbb{Q}^n(U_T^i > n) \\ &\leq \sum_{i=1}^d \frac{1}{n^2} \mathbb{E}_{\mathbb{Q}^n} [|U_T^i|^2] \\ &= \frac{1}{n^2} \mathbb{E}_{\mathbb{Q}^n} [|U_T|^2] \\ &\leq \frac{1}{n^2} \eta(\kappa_L, T, K, g_0). \end{aligned}$$

Sending  $n \rightarrow \infty$ , we obtain [\(C.1\)](#), proving that  $M$  is martingale. □

## References

- Eduardo Abi Jaber. Lifting the Heston model. *Quantitative Finance*, 19(12):1995–2013, 2019a.
- Eduardo Abi Jaber. Weak existence and uniqueness for affine stochastic Volterra equations with L1-kernels. *arXiv preprint arXiv:1912.07445*, *Bernoulli Journal*, to appear, 2019b.
- Eduardo Abi Jaber. The Laplace transform of the integrated Volterra Wishart process. *arXiv:1911.07719*, 2019c.
- Eduardo Abi Jaber and Omar El Euch. Markovian structure of the Volterra Heston model. *Statistics & Probability Letters*, 2019a.
- Eduardo Abi Jaber and Omar El Euch. Multifactor approximation of rough volatility models. *SIAM Journal on Financial Mathematics*, 10(2):309–349, 2019b.
- Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200, 2019.
- Elisa Alòs, Jorge A León, and Josep Vives. On the short-time behavior of the implied volatility for jump-diffusion models with stochastic volatility. *Finance and Stochastics*, 11(4):571–589, 2007.
- Nicole Bäuerle and Sascha Desmettre. Portfolio optimization in fractional and rough Heston models. *SIAM Journal on Financial Mathematics*, 11(1):240–273, 2020.
- Marie-France Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751, 1991.
- Andrea Buraschi, Paolo Porchia, and Fabio Trojani. Correlation risk and optimal portfolio choice. *Journal of Finance*, 65(393-420), 2010.
- Mei Choi Chiu and Hoi Ying Wong. Mean–variance portfolio selection with correlation risk. *Journal of Computational and Applied Mathematics*, 263:432–444, 2014.
- Christa Cuchiero and Josef Teichmann. Markovian lifts of positive semidefinite affine Volterra-type processes. *Decisions in Economics and Finance*, 42(2):407–448, 2019.
- Omar El Euch and Mathieu Rosenbaum. Perfect hedging in rough Heston models. *The Annals of Applied Probability*, 28(6):3813–3856, 2018.
- Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 2019.
- Jean-Pierre Fouque and Ruiming Hu. Optimal portfolio under fast mean-reverting fractional stochastic environment. *SIAM Journal on Financial Mathematics*, 6(2):564–601, 2018.
- Masaaki Fukasawa. Asymptotic analysis for stochastic volatility: martingale expansion. *Finance and Stochastics*, 15(4):635–654, 2011.
- Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.



- Paul Glasserman and Pu He. Buy rough, sell smooth. *Quantitative Finance*, 20(3):363–378, 2020.
- Gustaf Gripenberg, Stig-Olof Londen, and Olof Staffans. *Volterra integral and functional equations*, volume 34 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1990.
- Bingyan Han and Hoi Ying Wong. Mean–Variance portfolio selection under Volterra Heston model. *Applied Mathematics & Optimization*, pages 1–28, 2020a.
- Bingyan Han and Hoi Ying Wong. Merton’s portfolio problem under Volterra Heston model. *Finance Research Letters*, page 101580, 2020b.
- Amine Ismail and Huy  n Pham. Robust Markowitz mean-variance portfolio selection under ambiguous covariance matrix. *Mathematical Finance*, 29(1):174–207, 2019.
- Monique Jeanblanc, Martin Schweizer, and Michael Mania. Mean-variance hedging via stochastic control and BSDEs for general semimartingales. *Annals of Applied Probability*, 22(6):2388–2428, 2012.
- Andrew EB Lim. Quadratic hedging and mean-variance portfolio selection with random parameters in an incomplete market. *Mathematics of Operations Research*, 29(1):132–161, 2004.
- Andrew EB Lim and Xun Yu Zhou. Mean-variance portfolio selection with random parameters in a complete market. *Mathematics of Operations Research*, 27(1):101–120, 2002.
- Harry Markowitz. Portfolio selection. *Journal of Finance*, pages 77–91, 1952.
- Huy  n Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.
- Mathieu Rosenbaum and Mehdi Thomas. From microscopic price dynamics to multidimensional rough volatility models. *arXiv:1910.13338*, 2019.
- Rainer Sch  bel and Jianwei Zhu. Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. *Review of Finance*, 3(1):23–46, 1999.
- Yang Shen. Mean–variance portfolio selection in a complete market with unbounded random coefficients. *Automatica*, 55:165–175, 2015.
- Yang Shen, Xin Zhang, and Tak Kuen Siu. Mean–variance portfolio selection under a constant elasticity of variance model. *Operations Research Letters*, 42(5):337–342, 2014.
- Galen R Shorack and Jon A Wellner. *Empirical processes with applications to statistics*. SIAM, 2009.
- Elias M Stein and Jeremy C Stein. Stock price distributions with stochastic volatility: an analytic approach. *The Review of Financial Studies*, 4(4):727–752, 1991.
- Mark Veraar. The stochastic Fubini theorem revisited. *Stochastics*, 84(4):543–551, 2012.
- Xun Yu Zhou and Duan Li. Continuous-time mean-variance portfolio selection: a stochastic LQ framework. *Applied Mathematics & Optimization*, 42:19–33, 2000.