The Laplace transform of the integrated Volterra Wishart process

Eduardo Abi Jaber

The Laplace transform of the integrated Volterra Wishart process

Eduardo Abi Jaber*

July 14, 2021

Abstract

We establish an explicit expression for the conditional Laplace transform of the integrated Volterra Wishart process in terms of a certain resolvent of the covariance function. The core ingredient is the derivation of the conditional Laplace transform of general Gaussian processes in terms of Fredholm's determinant and resolvent. Furthermore, we link the characteristic exponents to a system of non-standard infinite dimensional matrix Riccati equations. This leads to a second representation of the Laplace transform for a special case of convolution kernel. In practice, we show that both representations can be approximated by either closed form solutions of conventional Wishart distributions or finite dimensional matrix Riccati equations stemming from conventional linear-quadratic models. This allows fast pricing in a variety of highly flexible models, ranging from bond pricing in quadratic short rate models with rich autocorrelation structures, long range dependence and possible default risk, to pricing basket options with covariance risk in multivariate rough volatility models.

Keywords: Gaussian processes, Wishart processes, Fredholm's determinant, quadratic short rate models, rough volatility models.

MSC2010 Classification: 60G15, 60G22, 45B05, 91G20.

*Université Paris 1 Panthéon-Sorbonne, Centre d'Economie de la Sorbonne, 106, Boulevard de l'Hôpital, 75013 Paris, eduardo.abi-jaber@univ-paris1.fr. My work was supported by grants from Région Ile-de-France. I would like to thank Aurélien Alfonsi, Martin Larsson, Sergio Pulido and Mathieu Rosenbaum for helpful comments and fruitful discussions.

1 Introduction

We are interested in the $d \times d$ Volterra Wishart process $XX^\top$ where X is the $d \times m$-matrix valued Volterra Gaussian process

$$X_t = g_0(t) + \int_0^t K(t,s) dW_s,$$

for some given input curve $g_0 : [0,T] \to \mathbb{R}^{d\times m}$, suitable kernel $K : [0,T]^2 \to \mathbb{R}^{d\times d}$ and $d \times m$-matrix Brownian motion W, for a fixed time horizon $T > 0$.

The introduction of the kernel K allows for flexibility in financial modeling as illustrated in the two following examples. First, one can consider asymmetric (possibly negative) quadratic short rates of the form

$$r_t = \mathrm{tr}\left(X_t^\top Q X_t\right) + \xi(t)$$

where $Q \in \mathbb{S}_+^d$, ξ is an input curve used for matching market term structures and tr stands for the trace operator. The kernel K allows for richer autocorrelation structures than the one generated with the conventional Hull and White (1990) and Cox, Ingersoll, and Ross (2005) models. Second, for $d = m$, one can build stochastic covariance models for d–assets $S = (S^1, \dots, S^d)$ by considering the following dynamics for the stock prices:

$$dS_t = \mathrm{diag}(S_t) X_t dB_t$$

where B is d-dimensional and correlated with W. Then, the instantaneous covariance between the assets is stochastic and given by $\frac{d\langle \log S\rangle_t}{dt} = X_t X_t^\top \in \mathbb{S}_+^d$. When $d = m = 1$, one recovers the Volterra version of the Stein and Stein (1991) or Schöbel and Zhu (1999) model. Here, singular kernels K satisfying $\lim_{s\uparrow t} |K(t,s)| = \infty$, allow to take into account roughness of the sample paths of the volatility, as documented in Bennedsen et al. (2016); Gatheral et al. (2018). As an illustrative example for $d = m = 1$, one could consider the Riemann-Liouville fractional Brownian motion

$$X_t = \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s,$$

either with $H \in (0, 1/2)$ to reproduce roughness when modeling the variance process, or with $H \in (1/2, 1)$ to account for long memory in short rate models.

In both cases, integrated quantities of the form $\int_0^\cdot X_s X_s^\top ds$ play a key role for pricing zero-coupon bonds and options on covariance risk. In order to keep the model tractable, one needs to come up with fast pricing and calibration techniques. The main objective of the paper is to show that these models remain highly tractable, despite the inherent non-markovianity and non-semimartingality due to the introduction of the kernel K. For

$w \in \mathbb{S}_+^d$, our main result (Theorem 3.3) provides the explicit expression for the conditional Laplace transform:

$$\mathbb{E}\left[\exp\left(-\int_t^T \operatorname{tr}\left(wX_sX_s^\top\right)ds\right) \;\middle|\; \mathcal{F}_t\right] = \exp\left(\phi_{t,T} + \int_{(t,T]^2} \operatorname{tr}\left(g_t(s)^\top \Psi_{t,T}(ds,du)g_t(u)\right)\right),$$

where (ϕ, Ψ) are defined by

$$\begin{aligned}\partial_t \phi_{t,T} &= -m\int_{(t,T]^2} \operatorname{tr}\left(\Psi_{t,T}(ds,du)K(u,t)K(s,t)^\top\right), \quad \phi_{T,T} = 0,\\ \Psi_{t,T}(ds,du) &= -w\delta_{\{s=u\}}(ds,du) - \sqrt{w}R^w_{t,T}(s,u)\sqrt{w}dsdu,\end{aligned}$$

with $g_t(s) = \mathbb{E}[X_s|\mathcal{F}_t]$ the forward process, $C_t(s,u) = \mathbb{E}[(X_s - g_t(s))(X_u - g_t(u))^\top|\mathcal{F}_t]$ the conditional covariance function, and $R^w_{t,T} : [0,T]^2 \to \mathbb{R}^{d\times d}$ the Fredholm resolvent of $(-2\sqrt{w}C_t\sqrt{w})$ on $[0,T]$ given by

$$R^w_{t,T}(s,u) = -2\sqrt{w}C_t(s,u)\sqrt{w} - \int_t^T 2\sqrt{w}C_t(s,z)\sqrt{w}R^w_{t,T}(z,u)dz, \quad t \le s,u \le T.$$

Using the integral operator $\mathbf{C}_t$ induced by the covariance kernel C_t, i.e. $(\mathbf{C}_t f)(s) = \int_0^T C_t(s,u)f(u)du$ for $f \in L^2\left([0,T],\mathbb{R}^{d\times m}\right)$, the Laplace transform can be re-expressed in analytic form

$$\mathbb{E}\left[\exp\left(-\int_t^T \operatorname{tr}\left(wX_sX_s^\top\right)ds\right) \;\middle|\; \mathcal{F}_t\right] = \frac{\exp\left(-\langle g_t, \sqrt{w}\left(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}\right)^{-1}\sqrt{w}g_t\rangle_{L^2_t}\right)}{\det\left(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}\right)^{m/2}}$$

where $\langle f,g\rangle_{L^2_t} = \int_t^T \operatorname{tr}\left(f(s)^\top g(s)\right)ds$ and det stands for the Fredholm determinant.

The Laplace transform is exponentially quadratic in the forward process $(g_t)_{t\le T}$, and cannot in general be recovered from that of finite dimensional affine Volterra processes introduced in Abi Jaber et al. (2019a), see Remark 2.1. We also mention that the models studied here are quadratic constructions of Gaussian processes and do not pose any difficulty regarding existence and uniqueness, in contrast for instance with conventional Wishart processes that go beyond squares of Gaussians, see Bru (1991).

Furthermore, we link Ψ to a system of non-standard infinite dimensional backward Riccati equations in the general case of non-convolution kernels. This allows us to deduce a second representation of the Laplace transform for a special case of convolution kernels in the form

$$K(t,s) = k(t-s)\mathbf{1}_{s\le t} \quad \text{such that} \quad k(t) = \int_{\mathbb{R}_+} e^{-xt}\mu(dx), \quad t > 0,$$

for some suitable signed measure μ, showing, similarly to Cuchiero and Teichmann (2019); Harms and Stefanovits (2019), that the Volterra Wishart process can be seen as a superposition of possibly infinitely many conventional linear-quadratic models written on the infinite dimensional process

$$Y_t(x) = \int_0^t e^{-x(t-s)} dW_s, \quad t \geq 0, \quad x \in \mathbb{R}_+.$$

In particular, this second representation not only allows us to recover the expressions for the Laplace transform derived in the aforementioned articles but, most importantly, provides an explicit solution for the corresponding infinite dimensional Riccati equations.

Although explicit, the expression for the Laplace transform is not known in closed form, except for certain cases. We provide two approximation procedures either by closed form solutions of conventional Wishart distributions (Section 2.2) or finite dimensional matrix Riccati equations stemming from conventional linear-quadratic models (Section 3.3). These approximations can then be used to price bonds with possible default risk, or options on covariance in multivariate (rough) volatility models by Laplace transform techniques (Section 4).

Literature Conventional Wishart processes initiated by Bru (1991) and introduced in finance by Gourieroux and Sufana (2003) have been intensively applied, together with their variants, in term structure and stochastic covariance modeling, see for instance Alfonsi (2015); Buraschi et al. (2010); Cuchiero et al. (2011, 2016); Da Fonseca et al. (2007, 2008); Gouriéroux et al. (2009); Muhle-Karbe et al. (2012). Conventional linear quadratic models have been characterized in Chen et al. (2004); Cheng and Scaillet (2007). Volterra Wishart processes have been recently studied in Cuchiero and Teichmann (2019); Yue and Huang (2018). Applications of certain quadratic Gaussian processes can be found in Benth and Rohde (2018); Corcuera et al. (2013); Harms and Stefanovits (2019); Kleptsyna et al. (2002). Gaussian stochastic volatility models have been treated in Gulisashvili (2018); Gulisashvili et al. (2019).

Outline In Section 2 we derive the Laplace transform of general quadratic Gaussian processes in $\mathbb{R}^N$, we provide a first approximation procedure by closed form expressions and link the characteristic exponent to non-standard Riccati equations. These results are then used in Section 3 to deduce the Laplace transforms of Volterra Wishart processes. We also provide a second representation formula for the Laplace transform together with an approximation scheme for a special class of convolution kernels. Section 4 presents applications to pricing: (i) bonds in quadratic Volterra short rate models with possible default risk; (ii) options on volatility for basket products in Volterra Wishart (rough) covariance models. Some technical results are collected in the appendices.

Notations For $T > 0$, we define $L^2([0,T]^2, \mathbb{R}^{N\times N})$ to be the space of measurable

functions $F : [0,T]^2 \to \mathbb{R}^{N\times N}$ such that

$$\int_0^T \int_0^T |F(t,s)|^2 dtds < \infty.$$

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

$$(F \star G)(s,u) = \int_0^T F(s,z)G(z,u)dz, \quad s,u \leq T, \tag{1.1}$$

which is well-defined in $L^2([0,T]^2, \mathbb{R}^{N\times N})$ due to the Cauchy-Schwarz inequality. We denote by F^* the adjoint kernel of F in $L^2([0,T], \mathbb{R}^{N\times N})$, that is

$$F^*(s,u) = F(u,s)^\top, \quad s,u \leq T.$$

For any kernel $F \in L^2([0,T]^2, \mathbb{R}^{N\times N})$, we denote by $\mathbf{F}$ the integral operator from $L^2([0,T], \mathbb{R}^N)$ into itself induced by the kernel F that is

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

If $\mathbf{F}$ and $\mathbf{G}$ are two integral operators induced by the kernels F and G in $L^2([0,T]^2, \mathbb{R}^{N\times N})$, then $\mathbf{FG}$ is an integral operator induced by the kernel $F \star G$.

$\mathbb{S}_+^N$ stands for the cone of symmetric non-negative semidefinite $N \times N$-matrices, tr denotes the trace of a matrix and I_N is the $N \times N$ identity matrix. The vectorization operator is denoted by vec and the Kronecker product by $\otimes$, we refer to Appendix B for more details.

2 Quadratic Gaussian processes

Throughout this section, we fix $T > 0$, $N \geq 1$ and let Z denote a $\mathbb{R}^N$-valued square-integrable Gaussian process on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{P})$ with mean function $g_0(s) = \mathbb{E}[Z_s]$ and covariance kernel given by $C_0(s,u) = \mathbb{E}[(Z_s - g_0(s))(Z_u - g_0(u))^\top]$, for each $s,u \in [0,T]$. We note that g_0 and C_0 may depend on T, but we do not make this dependence explicit to ease notations.

2.1 Fredholm's representation and first properties

Assume that C_0 is continuous in both variables. Then, there exists a kernel $K_T \in L^2([0,T]^2, \mathbb{R}^{N\times N})$ and a N-dimensional Brownian motion W such that

$$Z_t = g_0(t) + \int_0^T K_T(t,s)dW_s, \tag{2.1}$$

for all $t \leq T$, see Sottinen and Viitasaari (2016, Theorem 12 and Example 2). In particular, $C_0 = K_T \star K_T^*$, that is

$$C_0(s,u) = \int_0^T K_T(s,z) K_T(u,z)^\top dz, \quad s,u \leq T.$$

For any $t \leq s$, Z_s admits the following decomposition

$$Z_s = g_0(s) + \int_0^t K_T(s,u) dW_u + \int_t^T K_T(s,u) dW_u, \tag{2.2}$$

showing that conditional on $\mathcal{F}_t$, Z_s is again a Gaussian process with conditional mean

$$g_t(s) = \mathbb{E}[Z_s|\mathcal{F}_t] = g_0(s) + \int_0^t K_T(s,u) dW_u, \quad t \leq s \leq T,$$

and conditional covariance function

$$\begin{aligned} C_t(s,u) &= \mathbb{E}[(Z_s - g_t(s))(Z_u - g_t(u))^\top |\mathcal{F}_t] \\ &= \int_t^T K_T(s,z) K_T(u,z)^\top dz, \qquad t \leq s,u \leq T. \end{aligned} \tag{2.3}$$

Again we drop the possible dependence of g_t and C_t on T, and we note in particular that for each $s,u \leq T$, $t \to C_t(s,u)$ is absolutely continuous on $[0, s \wedge u]$ with density

$$\dot{C}_t(s,u) = -K(s,t) K(u,t)^\top, \tag{2.4}$$

and that the process $t \mapsto g_t(s)$ is a semimartingale on $[0,s)$ with dynamics

$$dg_t(s) = K_T(s,t) dW_t, \quad t < s.$$

We are chiefly interested in the $\mathbb{S}_+^N$-valued process $ZZ^\top$. The following remark shows that, in general, $ZZ^\top$ cannot be recast as an affine Volterra process as studied in Abi Jaber et al. (2019a).

Remark 2.1. *To fix ideas, we set $g_0 \equiv Z_0 \in \mathbb{R}^N$. An application of Itô's formula yields*

$$\begin{aligned} g_t(s) g_t(s)^\top = Z_0 Z_0^\top &+ \int_0^t K_T(s,u) K_T(s,u)^\top du \\ &+ \int_0^t K_T(s,u) dW_u g_u(s)^\top + \int_0^t g_u(s) dW_u^\top K_T(s,u)^\top, \quad t < s. \end{aligned}$$

Taking the limit $s \to t$ leads to the dynamics

$$\begin{aligned} Z_t Z_t^\top = Z_0 Z_0^\top &+ \int_0^t K_T(t,u) K_T(t,u)^\top du \\ &+ \int_0^t K_T(t,u) dW_u g_u(t)^\top + \int_0^t g_u(t) dW_u^\top K_T(t,u)^\top. \end{aligned} \tag{2.5}$$

This shows, that in general, because of the presence of the infinite dimensional process $t \mapsto g_t$ in the dynamics, $ZZ^\top$ does not satisfy a stochastic Volterra equation in the form

$$Y_t = Y_0 + \int_0^t K(t,s)b(Y_s)ds + \int_0^t \sigma(Y_s)dW_s^\top K(t,s)^\top + \int_0^t K(t,s)dW_s\sigma(Y_s)^\top,$$

where $b, \sigma : \mathbb{R}^N \mapsto \mathbb{R}^{N\times N}$. For this reason, $ZZ^\top$ falls beyond the scope of the processes studied in Abi Jaber et al. (2019a). Except for very specific cases, for instance, when $K_T \equiv I_N$, we have $g_u(s) = Z_u$ for all $u < s$, and (2.5) reduces to the well-known dynamics of Wishart processes as introduced by Bru (1991).

Whence, the conditional Laplace transform of $ZZ^\top$ cannot be deduced from Abi Jaber et al. (2019a, Theorem 4.3). Nonetheless, it can be directly computed from Wishart distributions that we recall in Appendix A.

Theorem 2.2. *Fix $t \le s \le T$. Conditional on $\mathcal{F}_t$, $Z_sZ_s^\top$ follows a Wishart distribution*

$$Z_sZ_s^\top \sim_{|\mathcal{F}_t} WIS_N\left(1/2, g_t(s)g_t(s)^\top, 2C_t(s,s)\right).$$

Further, for any $u \in \mathbb{S}_+^N$, the conditional Laplace transform reads

$$\mathbb{E}\left[\exp\left(-Z_s^\top uZ_s\right) \,\middle|\, \mathcal{F}_t\right] = \frac{\exp\left(-g_t(s)^\top u\left(I_N + 2C_t(s,s)u\right)^{-1} g_t(s)\right)}{\det\left(I_N + 2C_t(s,s)u\right)^{1/2}}.$$

Proof. Fix $t \le s \le T$, conditional on $\mathcal{F}_t$, it follows from (2.2) that Z_s is a Gaussian vector in $\mathbb{R}^N$ with mean vector $g_t(s) \in \mathbb{R}^N$ and covariance matrix $C_t(s,s) \in \mathbb{R}^{N\times N}$. The claimed result now follows from Proposition A.1. □

In particular, if $N = 1$, $t = 0$ and $s = T$, one obtains the well-known chi-square distribution

$$\mathbb{E}\left[\exp\left(-uZ_T^2\right)\right] = \frac{\exp\left(\frac{-ug_0(T)^2}{1+2uC_0(T,T)}\right)}{\left(1+2uC_0(T,T)\right)^{1/2}}, \quad u \ge 0.$$

The computation of the Laplace transform for the integrated squared process is more involved and is treated in the next subsection.

2.2 Conditional Laplace transform of the integrated quadratic process

We are interested in computing the conditional Laplace transform

$$\mathbb{E}\left[\exp\left(-\int_t^T Z_s^\top wZ_sds\right) \,\middle|\, \mathcal{F}_t\right], \quad w \in \mathbb{S}_+^N, \quad t \le T. \tag{2.6}$$

For $t = 0$ and for centered processes, such computations appeared several times in the literature showing that the quantity of interest can be decomposed as an infinite product of independent chi-square distributions, see for instance Anderson and Darling (1952); Cameron and Donsker (1959); Varberg (1966). The same methodology can be readily adapted to our dynamical case and makes use of the celebrated Kac–Siegert/Karhunen–Loève representation of the process $Y = \sqrt{w}Z$ whose conditional covariance function is $C_t^w = \sqrt{w}C_t\sqrt{w}$, see Kac and Siegert (1947); Karhunen (1946); Loeve (1955). For this, we fix $t \leq T$, we consider the inner product on $L^2([t,T],\mathbb{R}^N)$ given by

$$\langle f, g\rangle_{L_t^2} = \int_t^T f(s)^\top g(s)ds, \qquad f, g \in L^2([t,T],\mathbb{R}^N),$$

and we assume that C_t is continuous in both variables[1]. By definition, the covariance kernel C_t^w is symmetric and nonnegative in the sense that

$$C_t^w(s,u) = C_t^w(u,s)^\top, \quad s,t \leq T,$$

and

$$\int_t^T \int_t^T f(s)^\top C_t^w(s,u)f(u)duds \geq 0, \quad f \in L^2([t,T],\mathbb{R}^N).$$

An application of Mercer's theorem, see Shorack and Wellner (2009, Theorem 1 p.208), yields the existence of a countable orthonormal basis $(e_{t,T}^n)_{n\geq 1}$ in $L^2([t,T],\mathbb{R}^N)$ and a sequence of nonnegative real numbers $(\lambda_{t,T}^n)_{n\geq 1}$ with $\sum_{n\geq 1}\lambda_{t,T}^n < \infty$ such that

$$C_t^w(s,u) = \sum_{n\geq 1} \lambda_{t,T}^n e_{t,T}^n(s)e_{t,T}^n(u)^\top, \quad t \leq s,u \leq T, \tag{2.7}$$

and

$$\int_t^T C_t^w(s,u)e_{t,T}^n(u)du = \lambda_{t,T}^n e_{t,T}^n(s), \quad t \leq s \leq T, \quad n \geq 1, \tag{2.8}$$

where the dependence of $(e_{t,T}^n, \lambda_{t,T}^n)$ on w is dropped to ease notations. This means that $(\lambda_{t,T}^n, e_{t,T}^n)_{n\geq 1}$ are the eigenvalues and the eigenvectors of the integral operator $\sqrt{w}\mathbf{C}_t\sqrt{w}$ from $L^2([t,T],\mathbb{R}^N)$ into itself induced by C_t^w:

$$(\sqrt{w}\mathbf{C}_t\sqrt{w}f)(s) = \int_t^T C_t^w(s,u)f(u)du, \quad t \leq s \leq T, \quad f \in L^2([t,T],\mathbb{R}^N).$$

As a consequence of Mercer's theorem, conditional on $\mathcal{F}_t$, the process Y admits the Kac–Siegert representation

$$Y_s = \sqrt{w}g_t(s) + \sum_{n\geq 1}\sqrt{\lambda_{t,T}^n}\xi^n e_{t,T}^n(s), \quad t \leq s \leq T, \tag{2.9}$$

[1] This is equivalent to assuming that the centered Gaussian process $(Z_\cdot - \mathbb{E}[Z_\cdot])$ is mean-square continuous.

where, conditional on $\mathcal{F}_t$, $(\xi_n)_{n\geq 1}$ is a sequence of independent standard Gaussian random variables, see Shorack and Wellner (2009, Theorem 2 p.210 and the comment below (14) on p.212). We now introduce the quantities needed for the computation of (2.6) in Theorem 2.3 below. We denote by id the identity operator on $L^2([t,T],\mathbb{R}^N)$, i.e. $(\mathrm{id}f)(s) = f(s)$, by $(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w})^{-1}$ the integral operator generated by the kernel

$$\sum_{n\geq 1} \frac{1}{1+2\lambda^n_{t,T}} e^n_{t,T}(s) e^n_{t,T}(u)^\top, \tag{2.10}$$

and we set

$$\det(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}) := \prod_{n\geq 1} \left(1+2\lambda^n_{t,T}\right). \tag{2.11}$$

The last expression is well defined due to the convergence of the series $(\sum_{n=1}^m \lambda^n_{t,T})_{m\geq 1}$ and the inequality

$$1 + 2\sum_{n=1}^m \lambda^n_{t,T} \leq \prod_{n=1}^m \left(1+2\lambda^n_{t,T}\right) \leq \exp\left(2\sum_{n=1}^m \lambda^n_{t,T}\right), \quad m \geq 1.$$

Theorem 2.3. *Fix $w \in \mathbb{S}^N_+$ and $t \leq T$. Assume that the function $(s,u) \mapsto C_t(s,u)$ is continuous. Then,*

$$\mathbb{E}\left[\exp\left(-\int_t^T Z_s^\top w Z_s ds\right) \,\middle|\, \mathcal{F}_t\right] = \frac{\exp\left(-\langle g_t, \sqrt{w}\left(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}\right)^{-1}\sqrt{w} g_t\rangle_{L^2_t}\right)}{\det\left(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}\right)^{1/2}}. \tag{2.12}$$

Proof. Fix $t \leq T$. Parseval's identity gives $\langle \sqrt{w}g_t, \sqrt{w}g_t\rangle_{L^2_t} = \sum_{n\geq 1}\langle \sqrt{w}g_t, e^n_{t,T}\rangle^2_{L^2_t}$ so that

$$\int_t^T Z_s^\top w Z_s ds = \langle Y, Y\rangle_{L^2_t} = \sum_{n\geq 1}\left(\sqrt{\lambda^n_{t,T}}\xi^n + \langle \sqrt{w}g_t, e^n_{t,T}\rangle_{L^2_t}\right)^2,$$

where the first equality follows from the definition $Y := \sqrt{w}Z$ and the second equality is a consequence of (2.9). By the independence of the sequence $(\xi^n)_{n\geq 1}$ and the dominated convergence theorem we can compute

$$\begin{aligned}
\mathbb{E}\left[\exp\left(-\int_t^T Z_s^\top w Z_s ds\right) \,\middle|\, \mathcal{F}_t\right] &= \prod_{n\geq 1} \mathbb{E}\left[\exp\left(-\left(\sqrt{\lambda^n_{t,T}}\xi^n + \langle \sqrt{w}g_t, e^n_{t,T}\rangle_{L^2_t}\right)^2\right) \,\middle|\, \mathcal{F}_t\right] \\
&= \prod_{n\geq 1} \frac{1}{\sqrt{1+2\lambda^n_{t,T}}} \exp\left(-\frac{1}{1+2\lambda^n_{t,T}}\langle \sqrt{w}g_t, e^n_{t,T}\rangle^2_{L^2_t}\right) \\
&= \det(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w})^{-1/2} \\
&\qquad \times \exp\left(-\sum_{n\geq 1}\frac{1}{1+2\lambda^n_{t,T}}\langle \sqrt{w}g_t, e^n_{t,T}\rangle^2_{L^2_t}\right),
\end{aligned}$$

where the second equality is obtained from the chi-square distribution, since the random variable $\left(\lambda^n_{t,T}\xi^n + \langle\sqrt{w}g_t, e^n_{t,T}\rangle_{L^2_t}\right)$ is Gaussian with mean $\langle\sqrt{w}g_t, e^n_{t,T}\rangle_{L^2_t}$ and variance $\lambda^n_{t,T}$, for each $n \geq 1$, see Proposition A.1. The claimed expression now follows upon observing that, thanks to (2.10),

$$\langle g_t, \sqrt{w}\left(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}\right)^{-1}\sqrt{w}g_t\rangle_{L^2_t} = \sum_{n\geq 1}\frac{1}{1+2\lambda^n_{t,T}}\langle\sqrt{w}g_t, e^n_{t,T}\rangle^2_{L^2_t}.$$

□

Remark 2.4. *The determinant (2.11) is named after Fredholm (1903) who defined it for the first time through the following expansion*

$$\det(\mathrm{id} + \mathbf{C}) = \sum_{n\geq 0}\frac{1}{n!}\int_t^T \dots \int_t^T \det\left[(C(s_i, s_j))_{1\leq i,j\leq n}\right] ds_1 \dots ds_n,$$

where $\mathbf{C}$ *is a generic integral operator of trace class with continuous kernel* C*. Lidskii's theorem ensures that Fredholm's definition is equivalent to*

$$\det(\mathrm{id} + \mathbf{C}) = \exp\left(\mathrm{Tr}\left(\log\left(\mathrm{id} + \mathbf{C}\right)\right)\right),$$

where $\mathrm{Tr}(\mathbf{C}) = \int_t^T \mathrm{tr}(C(s,s))ds$*, and consequently equivalent to the infinite product expression as in (2.11), refer to Simon (1977) for more details.*

Closed form solutions are known in some standard cases.

Example 2.5. Set $N = 1$, $t = 0$, $T = 1$ and $Z = W$, where W is a standard Brownian motion and $Z_0 \in \mathbb{R}$. Then, $g_0(s) = 0$ and $C_0(s,u) = s \wedge u$ and the eigenvalues and eigenvectors of the eigenproblem (2.8) are well-known and given by

$$\lambda^n_{0,1} = \frac{w}{(n-1/2)^2\pi^2} \quad \text{and} \quad e^n_{0,1}(s) = \sqrt{2}\sin\left(\left(n - \frac{1}{2}\right)\pi s\right), \quad n \geq 1.$$

Using the identity

$$\prod_{n\geq 1}\left(1 + 2\lambda^n_{0,1}\right) = \prod_{n\geq 1}\left(1 + \frac{2w}{(n-1/2)^2\pi^2}\right) = \cosh\sqrt{2w},$$

(2.12) reads

$$\mathbb{E}\left[\exp\left(-w\int_0^1 W_s^2 ds\right)\right] = \left(\cosh\sqrt{2w}\right)^{-1/2}. \tag{2.13}$$

For arbitrary kernels C, the eigenpairs $(\lambda^n, e^n)_{n\geq 1}$ are, in general, not known in closed form. This is the case for instance for the fractional Brownian motion. We provide in the next subsection an approximation by closed form formulas.

2.3 Approximation by closed form expressions

A natural idea to approximate (2.12) is to discretize the time-integral. Fix $t \leq T$ and let (s_i^n, α_i^n), $i = 1, \dots, n$ be a quadrature rule on $[t, T]$, i.e.

$$\int_t^T f(s)ds = \lim_{n\to\infty} \sum_{i=1}^n \alpha_i^n f(s_i^n).$$

By the dominated convergence theorem it follows that

$$\mathbb{E}\left[\exp\left(-\int_t^T Z_s^\top w Z_s ds\right) \;\middle|\; \mathcal{F}_t\right] = \lim_{n\to\infty} \mathbb{E}\left[\exp\left(-\sum_{i=1}^n \alpha_i^n Z_{s_i^n}^\top w Z_{s_i^n}\right) \;\middle|\; \mathcal{F}_t\right],$$

for all $w \in \mathbb{S}_+^N$. For each n, $(Z_{s_1^n}, \dots, Z_{s_n^n})^\top$ being Gaussian, the right hand side is known in closed form. This is the object of the next proposition which will make use of the Kronecker product $\otimes$ and the vectorization operator vec, we refer to Appendix B for more details.

Proposition 2.6. *Fix $w \in \mathbb{S}_+^N$ and $t \leq T$.*

$$\mathbb{E}\left[\exp\left(-\int_t^T Z_s^\top w Z_s ds\right) \;\middle|\; \mathcal{F}_t\right] = \lim_{n\to\infty} \frac{\exp\left(-g_t^{n\top} w_n \left(I_{nN} + 2C_t^n w_n\right)^{-1} g_t^n\right)}{\det\left(I_{nN} + 2C_t^n w_n\right)^{1/2}}, \tag{2.14}$$

where $w_n = (\mathrm{diag}(\alpha_1^n, \dots, \alpha_n^n) \otimes w) \in \mathbb{R}^{nN\times nN}$, g_t^n is the nN-vector

$$g_t^n = \begin{pmatrix} g_t(s_1^n) \\ \vdots \\ g_t(s_n^n) \end{pmatrix}, \tag{2.15}$$

and C_t^n is the $nN \times nN$-matrix with entries

$$(C_t^n)^{p,q} = C_t(s_i^n, s_k^n)^{jl}, \quad p = (i-1)N + j,\ q = (k-1)N + l, \tag{2.16}$$

for all $i, k = 1, \dots, n$, and $j, l = 1, \dots, N$.

Proof. We simply observe that

$$\sum_{i=1}^n \alpha_i^n Z_{s_i^n}^\top w Z_{s_i^n} = \mathbf{Z}^{n\top} \left(\mathrm{diag}(\alpha_1^n, \dots, \alpha_n^n) \otimes w\right) \mathbf{Z}^n,$$

where $\mathbf{Z}^n = \mathrm{vec}(Z^n)$ and $Z^n = (Z_{s_1^n}, \dots, Z_{s_n^n})$. Conditional on $\mathcal{F}_t$, $\mathbf{Z}^n$ being a Gaussian vector in $\mathbb{R}^{nN}$ with mean vector (2.15) and covariance matrix (2.16), the claimed result readily follows from Proposition A.1 combined with the dominated convergence theorem. □

We now illustrate the approximation procedure in practice for $N = 1$. Consider a one dimensional fractional Brownian motion W^H with Hurst index $H \in (0, 1)$ and set

$$I(H) = \mathbb{E}\left[\exp\left(-\int_0^1 \left(W_s^H\right)^2 ds\right)\right]. \tag{2.17}$$

The (unconditional) covariance function of the fractional Brownian motion is given by

$$C_0^H(s, u) = \frac{1}{2}\left(|s|^{2H} + |u|^{2H} - |s - u|^{2H}\right). \tag{2.18}$$

Fix $n \geq 1$ we consider two quadrature rules on $[0, 1]$: the left Riemann sum with $s_i^n = i/n$ and $\alpha_i^n = 1/n$ and the Gauss-Legendre rule advocated in Bornemann (2010). Since W^H is centered, (2.15) reads $g_0^n = 0$ and the right hand side in (2.14) reduces to

$$I^n(H) = \det\left(I_n + 2C_0^{H,n}\,\mathrm{diag}(\alpha_1^n, \ldots, \alpha_n^n)\right)^{-\frac{1}{2}}, \tag{2.19}$$

where $C_0^{H,n}(i, j) = C_0^H(s_i^n, s_j^n)$, $i, j = 1, \ldots, n$. We proceed as follows. First, we determine the reference value of (2.17) for several values of H. For $H = 1/2$, the exact value is $I(1/2) = \cosh(\sqrt{2})^{-1/2}$, recall (2.13). For $H \in \{0.1, 0.3, 0.7, 0.9\}$, we run a Monte–Carlo simulation to estimate $I(H)$ with the trapezoidal rule with a 95% confidence interval and 10^6 sample paths with 10^3 time steps for each sample path. Second, for each value of H, we compute $I^n(H)$ as in (2.19), for several values of n with the left Riemann sum and the Gauss–Legendre quadrature. The results are collected in Tables 1–2 and Figure 1 below. We observe that the Gauss–Legendre quadrature performs better than the left Riemann sum rule, especially for higher values of H. When $H \geq 0.5$, even with $n = 10$, $I^n(H)$ with the Gauss–Legendre rule falls already within the 95% confidence interval of the Monte–Carlo simulation. Other quadrature rules can be used in Proposition 2.6, see for instance Bornemann (2010). We refer to Remark 3.4 below for a numerical illustration in higher dimensions.

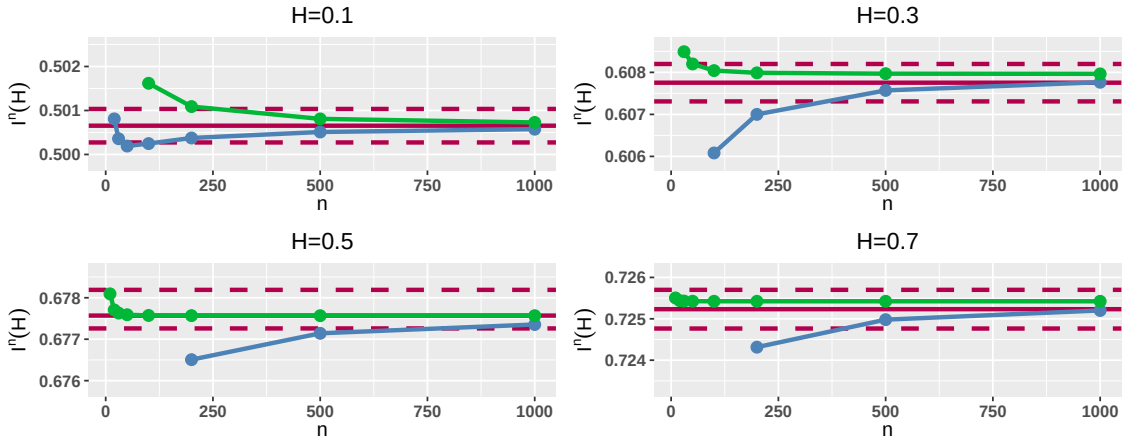

Figure 1: Convergence of $I^n(H)$ with the Riemann sum (blue) and the Gauss–Legendre quadrature (green) towards the benchmark MC value $I(H)$ (red) for different values of (H, n) from Table 1. The dashed lines delimit the 90% confidence interval of the Monte–Carlo simulation.

H	0.1	0.3	0.5	0.7	0.9
ref. $I(H)$	0.50065	0.60775	0.67757*	0.72523	0.76023
$n \backslash I^n(H)$					
10	0.50310	0.59301	0.65763	0.70376	0.73779
20	0.50081	0.59961	0.66727	0.71445	0.74810
30	0.50027	0.60291	0.67160	0.71912	0.75386
50	0.50019	0.60433	0.67337	0.72101	0.75581
100	0.50025	0.60608	0.67545	0.72321	0.75801
200	0.50037	0.60701	0.67650	0.72431	0.75924
500	0.50051	0.60757	0.67714	0.72498	0.75992
1000	0.50058	0.60776	0.67735	0.72520	0.76015

Table 1: Approximation of $I(H)$ by $I^n(H)$ with the left Riemann sum for several values of H with n ranging between 10 and 1000. *exact value for $I(1/2)$.

H	0.1	0.3	0.5	0.7	0.9
ref. $I(H)$	0.50065	0.60775	0.67757*	0.72523	0.76023
$n \backslash I^n(H)$					
10	0.51331	0.61075	0.67810	0.72550	0.76039
20	0.50665	0.60895	0.67771	0.72544	0.76038
30	0.50447	0.60850	0.67763	0.72543	0.76038
50	0.50279	0.60820	0.67759	0.72543	0.76038
100	0.50162	0.60804	0.67758	0.72542	0.76038
200	0.50109	0.60799	0.67757	0.72542	0.76038
500	0.50081	0.60797	0.67757	0.72542	0.76038
1000	0.50072	0.60797	0.67757	0.72542	0.76038

Table 2: Approximation of $I(H)$ by $I^n(H)$ with the Gauss–Legendre quadrature for several values of H with n ranging between 10 and 1000. *exact value for $I(1/2)$.

2.4 Connection to Riccati equations

The expression (2.12) is reminiscent of the formula obtained for finite dimensional Wishart processes in Bru (1991) and more generally that of linear quadratic diffusions, see Cheng and Scaillet (2007), suggesting a connection with infinite dimensional Riccati equations. Indeed, setting

$$\phi_{t,T} = -\frac{1}{2}\operatorname{Tr}\left(\log\left(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}\right)\right),$$
$$\mathbf{\Psi}_{t,T} = -\sqrt{w}\left(\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w}\right)^{-1}\sqrt{w},$$

it follows from Remark 2.4 that (2.12) can be rewritten as

$$\mathbb{E}\left[\exp\left(-\int_t^T Z_s^\top w Z_s ds\right) \;\middle|\; \mathcal{F}_t\right] = \exp\left(\phi_{t,T} + \langle g_t, \mathbf{\Psi}_{t,T} g_t\rangle_{L^2_t}\right), \quad t \le T. \tag{2.20}$$

Since $t \to C_t(s,u)$ is absolutely continuous with density $\dot{C}_t(s,u)$ given by (2.4), one would expect $t \to \mathbf{C}_t$ to be strongly differentiable[2] with derivative $\dot{\mathbf{C}}_t$ given by the integral operator

$$(\dot{\mathbf{C}}_t f)(s) = \int_t^T \dot{C}_t(s,u) f(u) du, \quad f \in L^2([0,T],\mathbb{R}^N), \quad s \le T. \tag{2.21}$$

By taking the derivatives we get that $(\phi, \mathbf{\Psi})$ solves the following system of operator Riccati equations

$$\dot{\phi}_{t,T} = \operatorname{Tr}\left(\mathbf{\Psi}_{t,T}\dot{\mathbf{C}}_t\right), \qquad \phi_{T,T} = 0, \tag{2.22}$$
$$\dot{\mathbf{\Psi}}_{t,T} = 2\mathbf{\Psi}_{t,T}\sqrt{w}\dot{\mathbf{C}}_t\sqrt{w}\mathbf{\Psi}_{t,T}, \qquad \mathbf{\Psi}_{T,T} = -w\mathrm{id}, \tag{2.23}$$

where $\dot{\mathbf{F}}_t$ denotes the derivative of $\mathbf{F}_t$ with respect to t.

This induces a system of Riccati equations for the kernels. To see this, we introduce the concept of resolvent. Fix $t \le T$ and define the kernel

$$R^w_{t,T}(s,u) = \sum_{n\ge 1}\left(\frac{1}{1+2\lambda^n_{t,T}} - 1\right) e^n_{t,T}(s) e^n_{t,T}(u)^\top, \quad t \le s,u \le T. \tag{2.24}$$

It is straightforward to check, using (2.7), that for all $t \le s,u \le T$,

$$2\int_t^T R^w_{t,T}(s,z)C^w_t(z,u)dz = 2\int_t^T C^w_t(s,z)R^w_{t,T}(z,u)dz = -R^w_{t,T}(s,u) - 2C^w_t(s,u). \tag{2.25}$$

[2] We recall that $t \mapsto \mathbf{C}_t$ is strongly differentiable at time $t \ge 0$, if there exists a bounded linear operator $\dot{\boldsymbol{C}}_t$ from $L^2\left([0,T],\mathbb{R}^N\right)$ into itself such that

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

$R^w_{t,T}$ is called the resolvent kernel of $(-2C^w_t)$ and the integral operator $\mathbf{R}^w_{t,T}$ induced by $R^w_{t,T}$ satisfies the relation

$$\mathbf{R}^w_{t,T} = (\mathrm{id} + 2\sqrt{w}\mathbf{C}_t\sqrt{w})^{-1} - \mathrm{id}, \tag{2.26}$$

so that $\mathbf{\Psi}_{t,T}$ can be re-expressed in terms of the resolvent

$$\mathbf{\Psi}_{t,T} = -w\mathrm{id} - \sqrt{w}\mathbf{R}^w_{t,T}\sqrt{w}.$$

The next theorem, whose proof is postponed to Appendix C, establishes the representation of the Laplace transform together with the Riccati equations (2.22)-(2.23) in terms of the induced kernel

$$\Psi_{t,T}(ds, du) = -w\delta_{s=u}(ds, du) + \psi_{t,T}(s, u)dsdu, \tag{2.27}$$

where $\psi_{t,T} = -\sqrt{w}R^w_{t,T}\sqrt{w}$ is the density of $\Psi_{t,T}$ with respect to the Lebesgue measure. We recall the $\star$-product defined in (1.1).

Theorem 2.7. *Fix $w \in \mathbb{S}^N_+$ and $T > 0$. Assume that the function $(s, u) \mapsto C_t(s, u)$ is continuous, for each $t \leq T$, such that*

$$\sup_{t\leq T}\sup_{t\leq s,u\leq T} |C_t(s, u)| < \infty. \tag{2.28}$$

Assume that $t \mapsto \mathbf{C}_t$ is strongly differentiable on $[0, T]$ with derivative (2.21). Then,

$$\mathbb{E}\left[\exp\left(-\int_t^T Z_s^\top w Z_s ds\right) \;\middle|\; \mathcal{F}_t\right] = \exp\left(\phi_{t,T} + \int_{(t,T]^2} g_t(s)^\top \Psi_{t,T}(ds, du) g_t(u)\right), \quad t \leq T,$$

where $t \mapsto \Psi_{t,T}$ is given by (2.27) and $\phi_{t,T}$ by

$$\dot\phi_{t,T} = -\int_{(t,T]^2} \mathrm{tr}\left(\Psi_{t,T}(ds, du)K_T(u, t)K_T(s, t)^\top\right), \quad \phi_{T,T} = 0.$$

In particular, $t \mapsto \Psi_{t,T}$ solves the Riccati equation with moving boundary

$$\begin{aligned} \dot\psi_{t,T} &= 2\Psi_{t,T} \star \dot C_t \star \Psi_{t,T} \qquad \text{on } (t, T]^2 \quad a.e., \\ \psi_{t,T}(t, \cdot) &= \psi_{t,T}(\cdot, t)^\top = 0 \qquad \text{on } [t, T] \quad a.e. \end{aligned} \tag{2.29}$$

We note that, since $\psi_{t,T}(s,u) = 0$ whenever $s \wedge u \leq t$, equation (2.29) is the compact form of

$$\begin{aligned}\dot{\psi}_{t,T}(s,u) = &-2wK_T(s,t)K_T(u,t)^\top w \\ &- 2wK_T(s,t)\int_t^T K_T(z,t)^\top \psi_{t,T}(z,u)dz \\ &- 2\int_t^T \psi_{t,T}(s,z)K_T(z,t)dzK_T(u,t)^\top w \\ &- 2\int_t^T \psi_{t,T}(s,z)K_T(z,t)dz\int_t^T K_T(z',t)^\top \psi_{t,T}(z',u)dz', \quad t < s,u \leq T \text{ a.e.}\end{aligned}$$

and the expanded form of ϕ is given by

$$\dot{\phi}_{t,T} = \int_t^T \operatorname{tr}\left(wK_T(s,t)K_T(s,t)^\top\right)ds - \int_t^T\int_t^T \operatorname{tr}\left(\psi_{t,T}(s,u)K_T(u,t)K_T(s,t)^\top\right)dsdu.$$

Remark 2.8. *The Riccati equation (2.29) can be compared to the Bellman (1957) and Krein (1955) variation formula for Fredholm's resolvent, see also Golberg (1973); Schumitzky (1968).*

3 The Volterra Wishart process and its Laplace transforms

Fix $T > 0$ and a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\leq T}, \mathbb{P})$ supporting a $d \times m$–matrix valued Brownian motion W. In this section, we consider the special case of the matrix-valued Volterra Gaussian process

$$X_t = g_0(t) + \int_0^t K(t,s)dW_s, \tag{3.1}$$

where $g_0 : [0,T] \to \mathbb{R}^{d\times m}$ is continuous and $K : [0,T] \to \mathbb{R}^{d\times d}$ is a $d \times d$–measurable kernel of Volterra type, that is $K(t,s) = 0$ for $s > t$. Compared to (2.1), since the kernel K is of Volterra type, the integration in (3.1) goes up to time t rather than T.

Under the assumption

$$\sup_{t\leq T}\int_0^T |K(t,s)|^2 ds < \infty \quad \text{and} \quad \lim_{h\to 0}\int_0^T |K(u+h,s) - K(u,s)|^2 ds = 0, \quad u \leq T, \tag{3.2}$$

the stochastic convolution

$$N_t = \int_0^t K(t,s)dW_s,$$

is well defined as an Itô integral, for each $t \in [0,T]$. Furthermore, Itô's isometry leads to

$$\mathbb{E}\left[|N_t - N_s|^2\right] \leq 2\int_s^t |K(t,u)|^2 du + 2\int_0^T |K(t,u) - K(s,u)|^2 du \tag{3.3}$$

which goes to 0 as $s \to t$ showing that N is mean-square continuous, and by virtue of Peszat and Zabczyk (2007, Proposition 3.21), the process N admits a predictable version. Furthermore, by the Burkholder-Davis-Gundy inequality applied on the local martingale $\left(\int_0^r K(t,s)dW_s\right)_{r\in[0,t)}$, it holds that

$$\sup_{t\leq T}\mathbb{E}\left[\left|\int_0^t K(t,s)dW_s\right|^p\right] \leq c_{p,T}\left(\sup_{t\leq T}\int_0^T |K(t,s)|^2 ds\right)^{p/2} < \infty, \quad p \geq 2, \tag{3.4}$$

where $c_{p,T}$ is a positive constant only depending on T and p. Kernels satisfying (3.2) are known as Volterra kernels of continuous and bounded type in L^2 in the terminology of Gripenberg et al. (1990, Definitions 9.2.1, 9.5.1 and 9.5.2).

We now provide several kernels of interest that satisfy (3.2). In particular, we stress that (3.2) does not exclude a singularity of the kernel at $s = t$.

Example 3.1. (i) For $H \in (0,1)$, the fractional Brownian motion with covariance function (2.18) admits a Volterra representation of the form (3.1) on $[0,T]$ with the kernel

$$K_H(t,s) = \frac{(t-s)^{H-1/2}}{\Gamma(H+\frac{1}{2})}\,{}_2F_1\left(H-\frac{1}{2};\frac{1}{2}-H;H+\frac{1}{2};1-\frac{t}{s}\right), \quad s \leq t,$$

where ${}_2F_1$ is the Gauss hypergeometric integral, see Decreusefond and Ustunel (1999).

(ii) If K is continuous on $[0,T]^2$, then (3.2) is satisfied by boundedness and the dominated convergence theorem. This is the case for instance for the Brownian Bridge W^{T_1} conditioned to be equal to $W_0^{T_1}$ at a time T_1: for all $T < T_1$, W^{T_1} admits the Volterra representation (3.1) on $[0,T]$ with the continuous kernel $K(t,s) = (T_1 - t)/(T_1 - s)$, for all $s, t \leq T$.

(iii) If K_1 an K_2 satisfy (3.2) then so does $K_1 \star K_2$ by an application of Cauchy-Schwarz inequality.

(iv) Any convolution kernel of the form $K(t,s) = k(t-s)\mathbf{1}_{s\leq t}$ with $k \in L^2([0,T],\mathbb{R}^{d\times d})$ satisfies (3.2). Indeed, for any $t \leq T$,

$$\int_0^T |K(t,s)|^2 ds = \int_0^t |k(t-s)|^2 ds = \int_0^t |k(s)|^2 ds \leq \int_0^T |k(s)|^2 ds,$$

yielding the first part of (3.2). The second part follows from the L^2-continuity of k, see (Brezis, 2010, Lemma 4.3).

We denote the conditional expectation of X by

$$g_t(s) = \mathbb{E}\left[X_s|\mathcal{F}_t\right], \quad t \leq s \leq T, \tag{3.5}$$

which is well-defined thanks to (3.4). For each $t \geq 0$, we denote by C_t the conditional covariance function of X with respect to $\mathcal{F}_t$, that is

$$C_t(s,u) = \int_t^{s\wedge u} K(s,r)K(u,r)^\top dr, \quad t \leq s,u \leq T. \tag{3.6}$$

C satisfies the assumption of Theorem 2.7 as shown in the next lemma. The expression of the strong derivative of $\boldsymbol{C}_t$ is given in terms of the density $\dot{C}_t$ of the kernel C_t given by (2.4) under the following additional assumption on the kernel:

$$\sup_{t\leq T} \int_0^T |K(s,t)|^2 ds < \infty. \tag{3.7}$$

Lemma 3.2. *Under (3.2), $(s,u) \mapsto C_t(s,u)$ is continuous, for all $t \leq T$ and (2.28) holds. Furthermore, under (3.7), $t \to \boldsymbol{C}_t$ is strongly differentiable on $[0,T]$ with derivative $\dot{\boldsymbol{C}}_t$ at $t \leq T$ given by the integral operator induced by the kernel $\boldsymbol{C_t}$ given by (2.4), that is*

$$(\dot{\boldsymbol{C}}_t f)(s) = \int_0^T \dot{C}_t(s,u)f(u)du = \int_t^T \dot{C}_t(s,u)f(u)du, \quad f \in L^2\left([0,T],\mathbb{R}^N\right).$$

Proof. First, it follows from (3.3) that the process X is mean-square continuous, which implies the continuity of $(s,u) \mapsto C_t(s,u)$. Second, an application of the Cauchy-Schwarz inequality on (3.6) yields

$$|C_t(s,u)|^2 \leq \left(\sup_{s'\leq T} \int_0^T |K(s',r)|^2 dr\right)^2$$

which proves (2.28). Finally, to prove the differentiability statement, we fix $t \leq T$ and first observe that

$$\int_0^T \int_0^T |\dot{C}_t(s,u)|^2 ds du = \left(\int_0^T |K(s,t)|^2 ds\right)^2$$

which is finite by virtue of (3.7). Whence, the kernel $\dot{C}_t$ belongs to $L^2\left([0,T]^2,\mathbb{R}^{N\times N}\right)$ so that it induces a linear bounded integral operator $\dot{\mathbf{C}}_t$ from $L^2\left([0,T],\mathbb{R}^N\right)$ into istelf. We now prove that $r \mapsto \mathbf{C}_r$ is differentiable at t with derivative given by $\dot{\mathbf{C}}_t$. For this, fix $f \in L^2\left([0,T],\mathbb{R}^N\right)$, $s \leq T$ and h such that $t+h \leq T$. Using the fact that, for all $u,s \leq T$, $t \mapsto C_t(s,u)$ is absolutely continuous with density $\dot{C}_t(s,u)$, we get that

$$(\mathbf{C}_{t+h}f)(s) - (\mathbf{C}_t f)(s) - h(\dot{\mathbf{C}}_t f)(s) = \int_0^T \int_t^{t+h} \left(\dot{C}_r(s,u) - \dot{C}_t(s,u)\right) dr f(u) du := A(s).$$

We now bound the right hand side in $L^2\left([0,T],\mathbb{R}^N\right)$. Successive applications of the Cauchy-Schwarz inequality together with the Fubini-Tonelli theorem yield

$$\begin{aligned}\|A\|_{L^2}^2 &= \int_0^T \left| \int_0^T \int_t^{t+h} \left(\dot{C}_r(s,u) - \dot{C}_t(s,u)\right) dr f(u) du \right|^2 ds \\ &\leq h\|f\|_B^2 \int_t^{t+h} \int_0^T \int_t^T \left|\dot{C}_r(s,u) - \dot{C}_t(s,u)\right|^2 du ds dr\end{aligned}$$

Therefore,

$$\frac{1}{h}\|\mathbf{C}_{t+h} - \mathbf{C}_t - h\dot{\mathbf{C}}_t\|_{\text{op}} \leq \int_t^{t+h} \int_0^T \int_0^T \left|\dot{C}_r(s,u) - \dot{C}_t(s,u)\right|^2 du ds dr.$$

The right hand side goes to 0 by virtue of (3.7), which ends the proof. □

3.1 A first representation

By construction the process $XX^\top$ is $\mathbb{S}_+^d$–valued and its Laplace transforms can be deduced from Theorems 2.2 and 2.7. Indeed, using the vectorization operator vec, which stacks the column of a $d \times m$–matrix A one underneath another in a vector of dimension $N = dm$, see Appendix B, the study of the matrix valued process X reduces to that of the $\mathbb{R}^{dm}$-valued Gaussian process $Z = \text{vec}(X)$ as done in Section 2.

The following theorem represents the main result of the paper.

Theorem 3.3. *Let X be the $d \times m$–matrix valued process defined in (3.1) for some Volterra kernel K satisfying (3.2) and (3.7). Fix $t \leq T$. For any $u \in \mathbb{S}_+^d$,*

$$\mathbb{E}\left[\exp\left(-\operatorname{tr}\left(uX_T X_T^\top\right)\right) \;\middle|\; \mathcal{F}_t\right] = \frac{\exp\left(-\operatorname{tr}\left(u\left(I_d + 2C_t(T,T)u\right)^{-1} g_t(T) g_t(T)^\top\right)\right)}{\det\left(I_d + 2C_t(T,T)u\right)^{m/2}}. \quad (3.8)$$

For any $w \in \mathbb{S}_+^d$, the Laplace transform

$$\mathcal{L}_{t,T}(w) = \mathbb{E}\left[\exp\left(-\int_t^T \operatorname{tr}\left(wX_s X_s^\top\right) ds\right) \;\middle|\; \mathcal{F}_t\right],$$

is given by

$$\mathcal{L}_{t,T}(w) = \exp\left(\phi_{t,T} + \int_{(t,T]^2} \operatorname{tr}\left(g_t(s)^\top \Psi_{t,T}(ds,du) g_t(u)\right)\right), \quad (3.9)$$

where (ϕ, Ψ) are defined by

$$\dot{\phi}_{t,T} = -m \int_{(t,T]^2} \operatorname{tr}\left(\Psi_{t,T}(ds,du) K(u,t) K(s,t)^\top\right), \quad \phi_{T,T} = 0, \quad (3.10)$$

$$\Psi_{t,T}(ds,du) = -w\delta_{\{s=u\}}(ds,du) - \sqrt{w} R_{t,T}^w(s,u)\sqrt{w} ds du, \quad \text{on} \quad [t,T], \quad (3.11)$$

where $R^w_{t,T}$ is the $d \times d$–matrix valued resolvent of $(-2\sqrt{w}C_t\sqrt{w})$, with C_t the conditional covariance function (3.6) and g_t the conditional mean given by (3.5). In particular, $t \mapsto \Psi_{t,T}$ solves the Riccati equation with moving boundary

$$\dot{\psi}_{t,T} = 2\Psi_{t,T} \star \dot{C}_t \star \Psi_{t,T} \quad \text{on } (t,T]^2 \quad a.e., \tag{3.12}$$

$$\psi_{t,T}(t,\cdot) = \psi_{t,T}(\cdot,t)^\top = 0 \quad \text{on } [t,T] \quad a.e., \tag{3.13}$$

where $\psi_{t,T}(s,u) = \sqrt{w}R^w_{t,T}(s,u)\sqrt{w}$.

Proof. Setting $Z = \mathrm{vec}(X)$ and $\boldsymbol{W} = \mathrm{vec}(W)$, an application of the vectorization operator vec on both sides of the $d\times m$–matrix valued equation (3.1) yields the $N := dm$ dimensional vector valued Gaussian process

$$Z_s = \mathrm{vec}\,(g_0(s)) + \int_0^s \mathcal{K}(s,u)d\boldsymbol{W}_u. \tag{3.14}$$

where $\mathcal{K}$ is the $\mathbb{R}^{N\times N}$ kernel

$$\mathcal{K} : (s,u) \mapsto (I_m \otimes K(s,u))$$

coming from the relation (B.1), with $\otimes$ the Kronecker product. Whence, the conditional mean and covariance functions of Z are given respectively by $\mathrm{vec}(g_t)$ and

$$\mathcal{C}_t(s,u) = (I_m \otimes C_t(s,u)), \quad u,s \le T. \tag{3.15}$$

In addition, due to (B.2),

$$\mathrm{tr}(wXX^\top) = Z^\top(I_m \otimes w)Z, \quad w \in \mathbb{S}^N.$$

• We first prove (3.8). Fix $t \le T$ and $u \in \mathbb{S}^d_+$. An application of Theorem 2.2 yields

$$\mathbb{E}\left[\exp\left(-Z_T^\top(I_m \otimes u)Z_T\right) \;\middle|\; \mathcal{F}_t\right] = \frac{\exp\left(-H_t(T)\right)}{\det\left(I_N + 2\mathcal{C}_t(T,T)(I_m \otimes u)\right)^{1/2}},$$

with

$$H_t(T) = \mathrm{vec}(g_t(T))^\top(I_m \otimes u)\left(I_N + 2\mathcal{C}_t(T,T)(I_m \otimes u)\right)^{-1}\mathrm{vec}(g_t(T)).$$

We observe that by (3.15) and successive applications of the product rule (B.3)

$$\begin{aligned}
\left(I_N + 2\mathcal{C}_t(T,T)(I_m \otimes u)\right)^{-1} &= \left((I_m \otimes I_d) + 2(I_m \otimes C_t(T,T))(I_m \otimes u)\right)^{-1} \\
&= \left(I_m \otimes (I_d + 2C_t(T,T)u)\right)^{-1} \\
&= \left(I_m \otimes (I_d + 2C_t(T,T)u)^{-1}\right)
\end{aligned}$$

where the last equality follows from (B.5). Another application of (B.3) combined with (B.2) yields that

$$H_t(T) = \operatorname{tr}\left(u\left(I_d + 2C_t(T,T)u\right)^{-1} g_t(T) g_t(T)^\top\right).$$

Similarly,

$$\begin{aligned}\det\left(I_N + 2\mathcal{C}_t(T,T)(I_m \otimes u)\right) &= \det\left(I_m \otimes (I_d + 2C_t(T,T)u)\right) \\ &= \det\left(I_m \otimes (I_d + 2C_t(T,T)u)\right) \\ &= \det\left(I_d + 2C_t(T,T)u\right)^m\end{aligned}$$

where we used (B.6) for the last identity. Combining the above proves (3.8).

• We now prove (3.9). Fix $t \le T$ and $w \in \mathbb{S}_+^d$. An application of Theorem 2.7, justified by Lemma 3.2, yields that

$$\mathcal{L}_{t,T}(w) = \exp\left(\phi_{t,T} + \int_{(t,T]^2} \operatorname{vec}(g_t(s))^\top \widetilde{\Psi}_{t,T}(ds,du)\operatorname{vec}(g_t(u))\right), \tag{3.16}$$

where

$$\begin{aligned}\dot{\phi}_{t,T} &= -\int_{(t,T]^2} \operatorname{tr}\left(\widetilde{\Psi}_{t,T}(ds,du)\mathcal{K}(u,t)\mathcal{K}(s,t)^\top\right), \quad \phi_{T,T} = 0, \\ \widetilde{\Psi}_{t,T}(ds,du) &= -(I_m \otimes w)\delta_{s=u}(ds,du) - (I_m \otimes \sqrt{w})\widetilde{\mathcal{R}}^w_{t,T}(s,u)(I_m \otimes \sqrt{w})dsdu,\end{aligned} \tag{3.17}$$

and $\widetilde{\mathcal{R}}^w_{t,T}$ is the resolvent of $2\mathcal{C}^w_t(s,u)$. The claimed expressions now follows provided we prove that

$$\widetilde{\mathcal{R}}^w_{t,T} = \left(I_m \otimes R^w_{t,T}\right), \tag{3.18}$$

where $R^w_{t,T}$ is the resolvent kernel of $2C^w_t(s,u)$. Indeed, if this is the case, then, using the the product rule (B.3) we get that

$$\widetilde{\Psi}_{t,T} = \left(I_m \otimes \Psi_{t,T}\right), \tag{3.19}$$

where $\Psi_{t,T}$ is given by (3.11), so that, by (B.2),

$$\operatorname{vec}(g_t(s))^\top \widetilde{\Psi}_{t,T}(ds,du)\operatorname{vec}(g_t(u)) = \operatorname{tr}\left(g_t(s)^\top \Psi_{t,T}(ds,du) g_t(u)\right).$$

Plugging (3.19) back in (3.17) and using the identity (B.4) yields (3.10). Combining the above shows that (3.16) is equal to (3.9). We now prove (3.18). For this, we define $\mathcal{R}^w_{t,T} = \left(I_m \otimes R^w_{t,T}\right)$. Then, it follows from the resolvent equation (2.25) of $R^w_{t,T}$ and the product rule (B.3) that $\mathcal{R}^w_{t,T}$ solves

$$\mathcal{R}^w_{t,T} = -2\mathcal{C}_t - 2\mathcal{R}^w_{t,T} \star \mathcal{C}_t, \quad \mathcal{R}^w_{t,T} \star \mathcal{C}_t = \mathcal{C}_t \star \mathcal{R}^w_{t,T},$$

showing that $\mathcal{R}^w_{t,T}$ is a resolvent of $(-2\mathcal{C}_t)$. By uniqueness of the resolvent, see Gripenberg et al. (1990, Lemma 9.3.3), (3.18) holds.
• Finally, the Riccati equations (3.12)–(3.13) follow along the same lines by invoking Theorem 2.7. □

Remark 3.4. *Proposition 2.6 can be applied to the vectorized Gaussian process $Z = \mathrm{vec}(X)$ given by (3.14) to get an approximation formula for*

$$\mathbb{E}\left[\exp\left(-\int_t^T \mathrm{tr}\left(wX_sX_s^\top\right)ds\right)\;\middle|\;\mathcal{F}_t\right] = \mathbb{E}\left[\exp\left(-\int_t^T \mathrm{tr}\left(Z_s^\top(I_m\otimes w)Z_s\right)ds\right)\;\middle|\;\mathcal{F}_t\right].$$

To illustrate the convergence, we consider $d = m$, $K \equiv I_d$ and $g_0 \equiv X_0 \in \mathbb{R}^{d\times d}$. In this case, the Laplace transform of the integrated process $XX^\top$ is given in the following closed form, see Gnoatto and Grasselli (2014, Theorem 1),

$$\mathbb{E}\left[\exp\left(-\int_0^T \mathrm{tr}\left(wX_sX_s^\top\right)ds\right)\right] = \exp\left(-\phi(T) - \mathrm{tr}\left(\psi(T)X_0X_0^\top\right)\right),$$

with

$$\begin{aligned}\phi(T) &= \frac{d}{2}\,\mathrm{tr}(\log(\cosh(\sqrt{2w}T)))\\ \psi(T) &= \frac{1}{2}\left(\cosh(\sqrt{2w}T)\right)^{-1}\left(\sqrt{2w}\sinh(\sqrt{2w}T)\right).\end{aligned}$$

We set $d = 2$,

$$X_0 = \begin{pmatrix} 0.1 & 0.3\\ 0.2 & 0.4\end{pmatrix}, \quad w = \begin{pmatrix} 1 & 0.5\\ 0.5 & 1\end{pmatrix},$$

and we test the approximation of Proposition 2.6 with the left Riemann sum and the Gauss-Legendre quadrature applied to the vectorized Gaussian process $Z = \mathrm{vec}(X)$. The convergence is illustrated in Table 3 as the discretization step n varies.

3.2 A second representation for certain convolution kernels

The aim of this section is to link the Volterra Wishart distribution with conventional linear-quadratic processes (Chen et al., 2004; Cheng and Scaillet, 2007) for the special case of convolution kernels:

$$K(t,s) = k(t-s)\mathbf{1}_{s\le t} \quad \text{such that} \quad k(t) = \int_{\mathbb{R}_+} e^{-xt}\mu(dx), \quad t > 0, \tag{3.20}$$

where μ is a $d\times m$–measure of locally bounded variation satisfying

$$\int_{\mathbb{R}_+}\left(1\wedge x^{-1/2}\right)|\mu|(dx) < \infty, \tag{3.21}$$

Reference value	0.1749568	
n	Riemann	Gauss-Legendre
10	0.1592312	0.1756507
20	0.1666600	0.1751393
30	0.1693255	0.1750393
50	0.1715292	0.1749869
100	0.1732245	0.1749645
200	0.1740860	0.1749588
500	0.1746074	0.1749572
1000	0.1747819	0.1749569

Table 3: Approximation with n ranging between 10 and 1000.

and $|\mu|$ is the total variation of the measure, as defined in Gripenberg et al. (1990, Definition 3.5.1). The condition (3.21) ensures that k is locally square integrable, see Abi Jaber et al. (2019b, Lemma A.1). This is inspired by the approach initiated in Carmona et al. (2000) and generalized to stochastic Volterra equations in Abi Jaber and El Euch (2019b); Cuchiero and Teichmann (2019); Harms and Stefanovits (2019).

Several kernels of interest satisfy (3.20)-(3.21) such as weighted sums of exponentials and the Riemann-Liouville fractional kernel $K_{RL}(t) = \frac{t^{H-1/2}}{\Gamma(H+1/2)}$, for $H \in (0, 1/2)$. We refer to Abi Jaber et al. (2019b, Example 2.2) for more examples.

A straightforward application of stochastic Fubini's theorem provides the representation of $(X_t, g_t)_{t\geq 0}$ in terms of μ and the possibly infinite system of $d\times m$-matrix-valued Ornstein-Uhlenbeck processes

$$Y_t(x) = \int_0^t e^{-x(t-s)} dW_s, \quad t \geq 0, \quad x \in \mathbb{R}_+,$$

see for instance Abi Jaber et al. (2019b, Theorem 2.3).

Lemma 3.5. *Assume that K is of the form (3.20) with μ satisfying (3.21), then*

$$X_t = g_0(t) + \int_{\mathbb{R}_+} \mu(dx) Y_t(x), \qquad t \leq T,$$
$$g_t(s) = g_0(s) + \int_{\mathbb{R}_+} e^{-x(s-t)} \mu(dx) Y_t(x), \qquad t \leq s \leq T.$$

Combined with (3.9), we get an exponentially quadratic representations of the characteristic function of $XX^\top$ in terms of the process Y.

Theorem 3.6. *Assume that K is of the form (3.20) with μ satisfying (3.21) and fix $w \in \mathbb{S}_+^d$. Then,*

$$\mathcal{L}_{t,T}(w) = \exp\left(\Theta_{t,T} + 2\,\mathrm{tr}\left(\int_{\mathbb{R}_+} \Lambda_{t,T}(x)^\top \mu(dx) Y_t(x)\right) + \mathrm{tr}\left(\int_{\mathbb{R}_+^2} Y_t(x)^\top \mu(dx)^\top \Gamma_{t,T}(x,y)\mu(dy) Y_t(y)\right)\right), \quad (3.22)$$

where $t \mapsto (\Theta_{t,T}, \Lambda_{t,T}, \Gamma_{t,T})$ are given by

$$\Theta_{t,T} = \int_{(t,T]^2} \mathrm{tr}\left(g_0(s)^\top \Psi_{t,T}(ds,du) g_0(s)\right) + \phi_{t,T}, \quad (3.23)$$

$$\Lambda_{t,T}(x) = \int_{(t,T]^2} e^{-x(s-t)} \Psi_{t,T}(ds,du) g_0(u), \quad (3.24)$$

$$\Gamma_{t,T}(x,y) = \int_{(t,T]^2} e^{-x(s-t)} \Psi_{t,T}(ds,du) e^{-y(u-t)}, \quad (3.25)$$

with (ϕ, Ψ) as in (3.10)-(3.11).

A direct differentiation of $(\Theta, \Lambda, \Gamma)$ combined with the Riccati equations (3.12)–(3.13) for (ϕ, Ψ) yield a system of Riccati equation for $(\Theta, \Lambda, \Gamma)$.

Proposition 3.7. *The functions $t \mapsto (\Theta_{t,T}, \Lambda_{t,T}, \Gamma_{t,T})$ given by (3.23), (3.24) and (3.25) solve the system of backward Riccati equations*

$$\dot{\Theta}_{t,T} = -\mathcal{R}_0(t, \Lambda_{t,T}, \Gamma_{t,T}), \qquad \Theta_{T,T} = 0, \quad (3.26)$$

$$\dot{\Lambda}_{t,T}(x) = x\Lambda_{t,T}(x) - \mathcal{R}_1(t, \Lambda_{t,T}, \Gamma_{t,T})(x), \qquad \Lambda_{T,T}(x) = 0, \quad (3.27)$$

$$\dot{\Gamma}_{t,T}(x,y) = (x+y)\Gamma_{t,T}(x,y) - \mathcal{R}_2(\Gamma_{t,T})(x,y), \qquad \Gamma_{T,T}(x,y) = 0, \quad (3.28)$$

where

$$\mathcal{R}_0(t,\Lambda,\Gamma) = -\,\mathrm{tr}\left(g_0(t)^\top w g_0(t)\right) + m\,\mathrm{tr}\left(\int_{\mathbb{R}_+^2} \Gamma(x,y)\mu(dy)\mu(dx)^\top\right) + 2\,\mathrm{tr}\left(\left(\int_{\mathbb{R}_+} \Lambda(x)^\top \mu(dx)\right)\left(\int_{\mathbb{R}_+} \Lambda(y)^\top \mu(dy)\right)^\top\right),$$

$$\mathcal{R}_1(t,\Lambda,\Gamma)(x) = -w g_0(t) + 2\left(\int_{\mathbb{R}_+} \Gamma(x,x')\mu(dx')\right)\left(\int_{\mathbb{R}_+} \Lambda(y)^\top \mu(dy)\right)^\top,$$

$$\mathcal{R}_2(\Gamma)(x,y) = -w + 2\left(\int_{\mathbb{R}_+} \Gamma(x,x')\mu(dx')\right)\left(\int_{\mathbb{R}_+} \Gamma(y,y')\mu(dy')\right)^\top.$$

Similar Riccati equations to that of Γ have appeared in the literature when dealing with convolution kernels of the form (3.20) in the presence of a quadratic structure, see Abi Jaber et al. (2019b, Theorem 3.7), Alfonsi and Schied (2013, Theorem 1), Harms and Stefanovits (2019, Lemma 5.4), Cuchiero and Teichmann (2019, Corollary 6.1). A general existence and uniqueness result for more general equations has been recently obtained in Abi Jaber et al. (2019c).

Remark 3.8. *The expression (3.22) can be re-written in the following compact form*

$$\mathcal{L}_{t,T}(w) = \exp\left(\Theta_{t,T} + 2\langle \Lambda_{t,T}, Y_t\rangle_\mu + \langle Y_t, \mathbf{\Gamma}_{t,T} Y_t\rangle_\mu\right).$$

where $\mathbf{\Gamma}_{t,T}$ is the integral operator acting on $L^1(\mu, \mathbb{R}^{d\times m})$ induced by the kernel $\Gamma_{t,T}$:

$$(\mathbf{\Gamma}_{t,T} f)(x) = \int_{\mathbb{R}_+} \Gamma_{t,T}(x,y)\mu(dy)f(y), \quad f \in L^1(\mu, \mathbb{R}^{d\times m})$$

and $\langle \cdot, \cdot \rangle_\mu$ is the dual pairing

$$\langle f, g\rangle_\mu = \operatorname{tr}\left(\int_{\mathbb{R}_+} f(x)^\top \mu(dx)^\top g(x)\right), \quad (f,g) \in L^1(\mu, \mathbb{R}^{d\times m}) \times L^\infty(\mu^\top, \mathbb{R}^{d\times m}).$$

We end this subsection with two examples establishing the connection with conventional quadratic models.

Example 3.9. Fix $\Sigma \in \mathbb{R}^{d\times d}$. For the constant case $k \equiv \Sigma$ we have $\mu(dx) = \Sigma\delta_0(dx)$, $\operatorname{supp}\mu = \{0\}$ and $Y_t(0) = W_t \in \mathbb{R}^{d\times m}$. For $g_0(t) \equiv 0$, $\Lambda \equiv 0$ and (3.26) and (3.28) read

$$\begin{aligned} \dot\Theta_{t,T} &= -m \operatorname{tr}\left(\Gamma_{t,T}(0,0)\Sigma\Sigma^\top\right), & \Theta_{T,T} &= 0, \\ \dot\Gamma_{t,T}(0,0) &= w - 2\Gamma_{t,T}(0,0)\Sigma\Sigma^\top\Gamma_{t,T}(0,0), & \Gamma_{T,T}(0,0) &= 0. \end{aligned}$$

These are precisely the conventional backward matrix Riccati equations encountered for conventional Wishart processes, see Alfonsi (2015, Equation (5.15)). In this case, we recover the well-known Markovian expression for the conditional Laplace transform (3.22):

$$\mathbb{E}\left[\exp\left(\int_t^T \operatorname{tr}\left(-wW_sW_s^\top\right) ds\right) \,\middle|\, \mathcal{F}_t\right] = \exp\left(\Theta_{t,T} + \operatorname{tr}\left(\Gamma_{t,T}(0,0)W_tW_t^\top\right)\right).$$

Example 3.10. Fix $n \geq 1$, $x_i^n \in \mathbb{R}_+$ and $c_i^n \in \mathbb{R}^{d\times d}$, $i = 1, \dots, n$. Consider the kernel

$$k^n(t) = \sum_{i=1}^n c_i^n e^{-x_i^n t}, \quad t \geq 0, \tag{3.29}$$

which corresponds to the measure $\mu^n(dx) = \sum_{i=1}^n c_i^n \delta_{x_i^n}(dx)$. The system of Riccati equations (3.26), (3.27) (3.28) is reduced to a system of finite dimensional matrix Riccati equations for with values in $\mathbb{R} \times \mathbb{R}^{nd\times m} \times \mathbb{R}^{nd\times nd}$ given by:

$$\dot{\Theta}^n_{t,T} = \operatorname{tr}\left(g_0(t)^\top w g_0(t)\right) - m \operatorname{tr}\left(\Gamma^n_{t,T} C^n\right) - 2\operatorname{tr}\left(\Lambda^{n\top}_{t,T} C^n \Lambda^n_{t,T}\right), \qquad \Theta^n_{T,T} = 0, \quad (3.30)$$

$$\dot{\Lambda}^n_{t,T} = D^n_t + B^n \Lambda^n_{t,T} - 2\Gamma^n_{t,T} C^n \Lambda^n_{t,T}, \qquad \Lambda^n_{T,T} = 0, \quad (3.31)$$

$$\dot{\Gamma}^n_{t,T} = A^n + B^n \Gamma^n_{t,T} + \Gamma^n_{t,T} B^{n\top} - 2\Gamma^n_{t,T} C^n \Gamma^n_{t,T}, \qquad \Gamma^n_{T,T} = 0, \quad (3.32)$$

where for all $r = 1, \dots, m$ $i, j, = 1, \dots, n$ and $k, l = 1, \dots, d$, $p = (i-1)d+k, q = (j-1)d+l$,

$$(D^n_t)^{pr} = (w g_0(t))^{kr}, \qquad (\Lambda^n_{t,T})^{pr} = \Lambda_{t,T}(x_i)^{kr},$$
$$(C^n)^{pq} = (c^n_i c^{n\top}_j)^{kl}, \qquad (\Gamma^n_{t,T})^{pq} = \Gamma_{t,T}(x^n_i, x^n_j)^{kl},$$

and A^n and B^n are the $nd \times nd$ defined by

$$A^n = (\mathbb{1}_n \otimes w), \quad B^n = (\operatorname{diag}(x^n_1, \dots, x^n_n) \otimes I_d)$$

with $\mathbb{1}_n$ the $n \times n$ matrix with all components equal to 1. The Riccati equation (3.32) can be linearized by doubling the dimension and its solution is given explicitly by

$$\Gamma^n_{t,T} = G_2(T-t) G_4(T-t)^{-1}, \quad t \le T,$$

where

$$\begin{pmatrix} G_1(t) & G_2(t) \\ G_3(t) & G_4(t) \end{pmatrix} = \exp\left(t \begin{pmatrix} -B^n & -A^n \\ -2C^n & B^n \end{pmatrix}\right), \quad t \le T,$$

see Levin (1959). Furthermore, we recover the well-known Markovian expression for the conditional Laplace transform (3.22):

$$\mathcal{L}^n_{t,T}(w) = \exp\left(\Theta^n_{t,T} + 2\operatorname{tr}\left(\Lambda^{n\top}_{t,T} \widetilde{Y}^n_t\right) + \operatorname{tr}\left(\Gamma^n_{t,T} \widetilde{Y}^n_t \widetilde{Y}^{n\top}_t\right)\right), \quad (3.33)$$

where

$$(\widetilde{Y}^n_t)^{pr} = (c^n_i Y_t(x^n_i))^{kr}, \quad (3.34)$$

$p = (i-1)d + k$, for each $i = 1, \dots, n$, $k = 1, \dots, d$, $r = 1, \dots, m$.

The previous example shows that Volterra Wishart processes can be seen as a superposition of possibly infinitely many conventional linear-quadratic processes in the sense of Chen et al. (2004); Cheng and Scaillet (2007). This idea is used to build another approximation procedure in the next subsection.

3.3 Another approximation procedure

An application of the Burkholder-Davis-Gundy inequality yields the following stability result for the sequence

$$X_t^n = g_0^n(t) + \int_0^t k^n(t-s)dW_s, \quad n \geq 1,$$

where $g_0^n : [0,T] \to \mathbb{R}^{d\times m}$ and $k^n \in L^2([0,T],\mathbb{R}^{d\times d})$, for each $n \geq 1$.

Lemma 3.11. *Fix $k \in L^2([0,T],\mathbb{R}^{d\times d})$ and $g_0 : [0,T] \to \mathbb{R}^{d\times m}$ measurable and bounded. If*

$$\int_0^T |k^n(s) - k(s)|^2 ds \to 0 \quad \textit{and} \quad \sup_{t\leq T} |g_0^n(t) - g_0(t)| \to 0, \quad \textit{as} \quad n \to \infty, \tag{3.35}$$

then,

$$\sup_{t\leq T} \mathbb{E}\left[|X_t^n - X_t|^p\right] \to 0, \quad \textit{as} \quad n \to \infty, \quad p \geq 2.$$

Combined with Example 3.10, we obtain another approximation scheme for the Laplace transform based on finite-dimensional matrix Riccati equations (compare with Remark 3.4).

Proposition 3.12. *Fix $w \in \mathbb{S}_+^d$ and $t \leq T$. For each n, let k^n be as in (3.29) for some $x_i^n \in \mathbb{R}_+$ and $c_i^n \in \mathbb{R}^{d\times d}$. Assume that (3.35) holds. Then,*

$$\mathcal{L}_{t,T}(w) = \lim_{n\to\infty} \exp\left(\Theta_{t,T}^n + 2\,\mathrm{tr}\left(\Lambda_{t,T}^{n\top}\widetilde{Y}_t^n\right) + \mathrm{tr}\left(\Gamma_{t,T}^n \widetilde{Y}_t^n \widetilde{Y}_t^{n\top}\right)\right)$$

where $(\Theta^n, \Lambda^n, \Gamma^n)$ solve (3.30), (3.31) and (3.32) and $\widetilde{Y}^n$ is given by (3.34).

Proof. Fix $t \leq s \leq T$. Writing $X_s^{n\top} w X_s^n - X_s^\top w X_s = (X_s^n + X_s)^\top w (X_s^n - X_s)$, we get by the Cauchy-Schwarz inequality that

$$\mathbb{E}\left[\left|\int_t^T (X_s^{n\top} w X_s^n - X_s^\top w X_s) ds\right|^2\right] \leq c \sup_{s\leq T}\left(\mathbb{E}\left[|X_s|^2\right] + \mathbb{E}\left[|X_s^n|^2\right]\right) \sup_{s\leq T} \mathbb{E}\left[|X_s^n - X_s|^2\right],$$

for some constant c independent of n. It follows from Lemma 3.11 that $(\sup_{s\leq T} \mathbb{E}\left[|X_s^n|^2\right])_{n\geq 1}$ is uniformly bounded in n, so that the right hand side converges to 0 as $n \to \infty$. Whence, $\int_t^T X_s^{n\top} w X_s^n ds \to \int_t^T X_s^\top w X_s ds$ a.s. along a subsequence and the claimed convergence follows from the dominated convergence theorem combined with (3.33). □

For $d = m = 1$ and k of the form (3.20) for some measure μ, for suitable partitions $(\eta_i^n)_{0\leq i\leq n}$ of $\mathbb{R}_+$, the choice

$$c_i^n = \int_{\eta_{i-1}^n}^{\eta_i^n} \mu(dx) \quad \text{and} \quad x_i^n = \frac{1}{c_i^n} \int_{\eta_{i-1}^n}^{\eta_i^n} x\mu(dx), \quad i = 1, \ldots, n,$$

ensures the L^2-convergence of the kernels k^n in (3.29), we refer to Abi Jaber (2019); Abi Jaber and El Euch (2019a) for such constructions, see also Harms (2019) for other choices of quadratures and for a detailed study of strong convergence rates.

4 Applications

4.1 Bond pricing in quadratic Volterra short rate models with default risk

We consider a quadratic short rate model of the form

$$r_t = \operatorname{tr}\left(X_t^\top Q X_t\right) + \xi(t), \quad t \leq T,$$

where X is the $d \times m$ Volterra process as in (3.1), $Q \in \mathbb{S}_+^d$ and $\xi : [0,T] \to \mathbb{R}$ is an input curve used to match today's yield curve and/or control the negativity level of the short rate. The model replicates the asymmetrical distribution of interest rates, allows for rich auto-correlation structures, and the possibility to account for long range dependence, see for instance Benth and Rohde (2018); Corcuera et al. (2013).

An application of Theorem 3.3 yields the price $P(\cdot, T)$ of a zero-coupon bond with maturity T:

$$P(t,T) = \mathbb{E}\left[\exp\left(-\int_t^T r_s ds\right) \,\middle|\, \mathcal{F}_t\right] = \exp\left(-\int_t^T \xi(s)ds\right)\mathcal{L}_{t,T}(Q), \quad t \leq T,$$

where $\mathcal{L}$ is given by (3.9). In this case, the zero-coupon yield with time to maturity $\tau = T - t$ is quadratic in g and reads

$$\begin{aligned} y_t(\tau) &= -\frac{1}{T-t}\log P(t,T) \\ &= \frac{1}{\tau}\int_t^{t+\tau} \xi(s)ds + \frac{1}{\tau}\langle g_t, \mathbf{\Psi}_{t,t+\tau} g_t\rangle_{L_t^2} + \frac{m}{2\tau}\log\det\left(\mathrm{id} + 2\sqrt{Q}\mathbf{C}_{t,t+\tau}\sqrt{w}\right), \end{aligned}$$

with $\mathbf{\Psi}_{t,t+\tau} = \sqrt{Q}\left(\mathrm{id} + 2\sqrt{Q}\mathbf{C}_{t,t+\tau}\sqrt{Q}\right)^{-1}\sqrt{Q}$ and $\langle f, g\rangle_{L_t} = \int_t^{t+\tau} \operatorname{tr}(f(s)^\top g(s))ds$. The role of the input curve ξ becomes apparent: it allows to perfectly match any given yield curve and/or possibly push the yields into negative territory (observe that $\mathbf{\Psi}_{t,t+\tau}$ is a non-negative operator). Furthermore, various shapes of the deformation of the yield curve can be replicated. For instance, the left graph of Figure 2 shows that in the one dimensional setting, when $X = W^H$ with W^H a fractional Brownian motion with Hurst index $H \in (0,1)$, the variation of the Hurst index H can produce inverse, hump-shaped and normal yield curves. The combination of two independent fractional Brownian motion with different Hurst indices lead to richer deformations as displayed on the right graph of Figure 2.

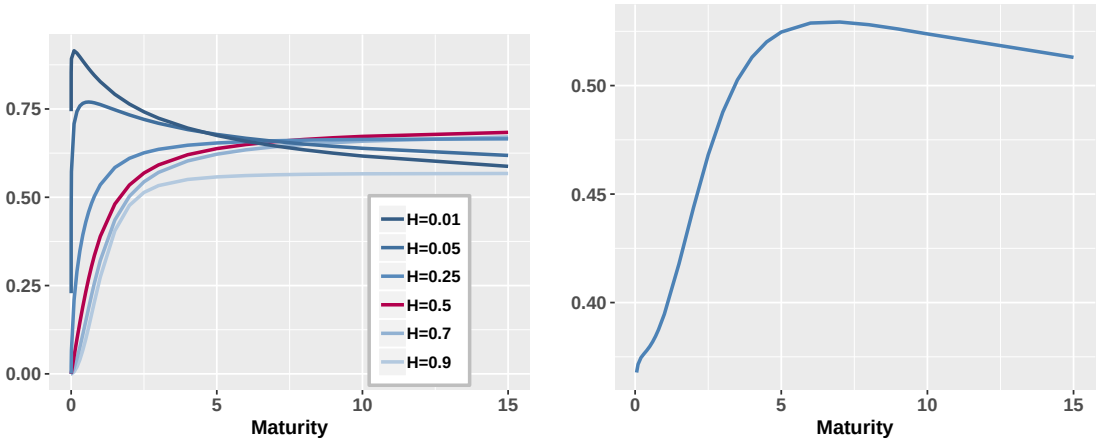

Figure 2: Left: Sensitivity of the yield curve $T \mapsto y_0(T)$ with respect to the Hurst index for $d = m = 1$, $X = W^H$ and $\xi \equiv 0$. Right: yield curve $T \mapsto y_0(T)$ for $d = 2$, $m = 1$, $X = (W^{H_1}, W^{H_2})^\top$, $(H_1, H_2) = (0.05, 0.9)$, $(Q_{11}, Q_{12}, Q_{22}) = (0.4, 0.05, 0.1)$ and $\xi \equiv 0$

From the dynamical perspective, for two maturities τ_1, τ_2, the instantaneous covariance of the variation of the yields over time is given by

$$d\langle y_\cdot(\tau_1), y_\cdot(\tau_2)\rangle_t = \frac{4}{\tau_1\tau_2} \operatorname{tr}\left((\boldsymbol{K}^* \boldsymbol{\Psi}_{t,t+\tau_1})(\mathbf{1}_t g_t)(t) \left((\boldsymbol{K}^* \boldsymbol{\Psi}_{t,t+\tau_2})(\mathbf{1}_t g_t)(t)\right)^\top \right) dt,$$

which is stochastic, non-trivial and allow sign changing across time. For instance, for the standard case $K(t,s) = e^{-B(t-s)}\eta$ and $g_0(t) \equiv X_0$, with $B, \eta, X_0 \in \mathbb{R}^{d\times d}$, we have that $g_t(s) = X_t$ for all $s \geq t$, one recovers the expression of the instantaneous covariance in a Wishart short rate model (see Buraschi et al. (2010)):

$$d\langle y_\cdot(\tau_1), y_\cdot(\tau_2)\rangle_t = \frac{4}{\tau_1\tau_2} \operatorname{tr}\left(\bar{\Psi}_{t,\tau_1} X_t X_t^\top \bar{\Psi}_{t,\tau_2}^\top \right) dt$$

with $\bar{\Psi}_{t,\tau_i} = \int_{(t,t+\tau_i]^2} e^{-B(s-t)} \eta \Psi_{t,t+\tau_i}(ds, du)$.

Compared to the standard case, more general kernels allow to capture both "time series" and "cross section" features of interest rates even with one single factor, see for instance Backus and Zin (1993); Dai and Singleton (2003); Ritchken and Chuang (2000):

- Figure 3 highlights the auto-correlation structure of short rates: for $H = 0.9$ the rates are highly persistent as observed in practice;
- Figure 4 shows the term structure of the variance of the yields: the case $H = 0.9$ allows to reproduce a humped term structure decaying at a slower rate than exponential (i.e. for the standard case $H = 0.5$) in agreement with the empirical observations.

The impact is amplified in a multifactor setting where all the factors share the same Hurst index, as shown by the principal component analysis on Figure 5. It might be interesting to consider a mixture of several factors with different Hurst indices to better capture the behavior of yields across several maturities.

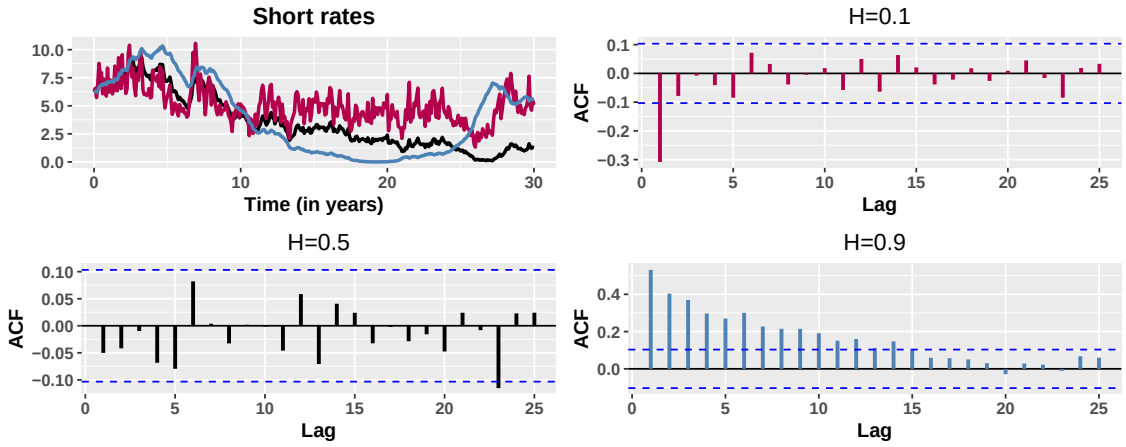

Figure 3: Simulation of monthly short rates (top left) with $X_t = 2.5 + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s$ and varying H index: $H = 0.1$ (red), $H = 0.5$ (black) and $H = 0.9$ (blue) with the corresponding autocorrelation plots.

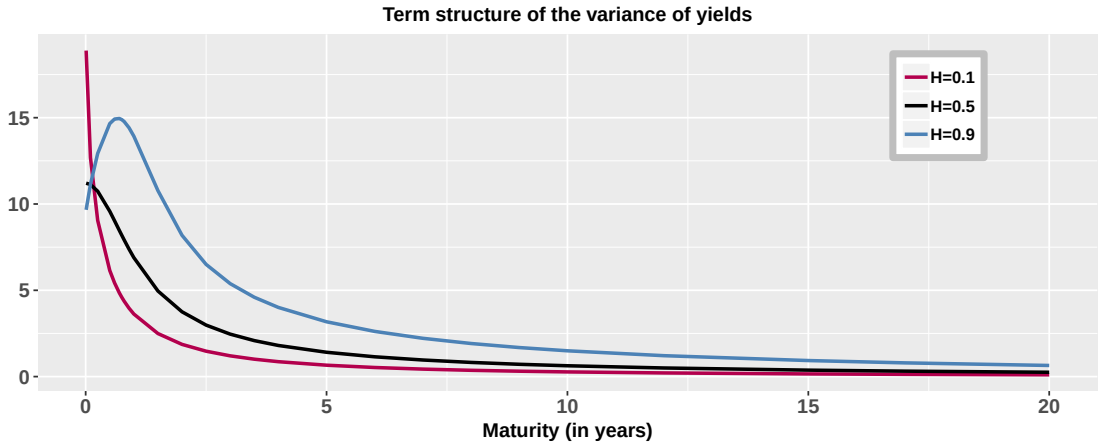

Figure 4: Impact of the Hurst index on the term structure of the variance of yields with $X_t = 2.5 + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s$.

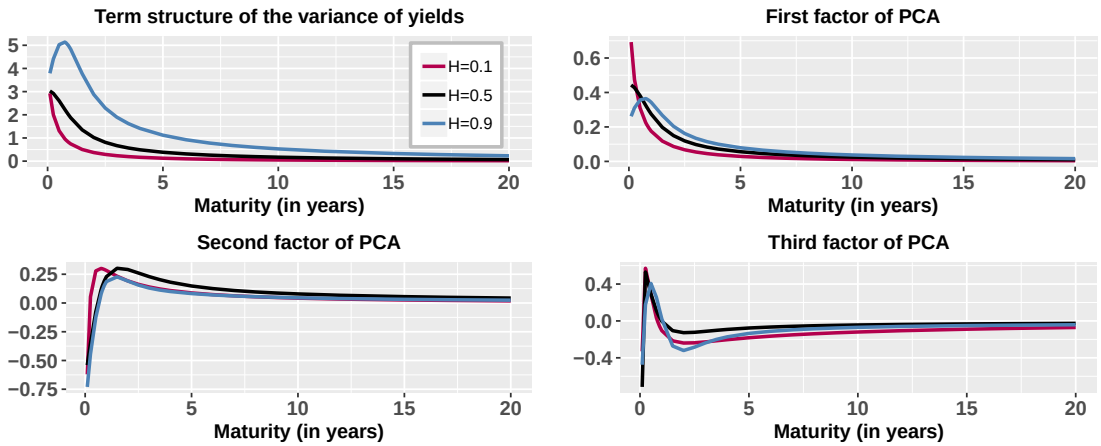

Figure 5: Impact of the Hurst index on the principal component analysis of the covariance of the yields with $d = 3$, $m = 1$, $Q_{ii} = 1$ and $Q_{12} = Q_{23} = 0.5$ and $X_t^i = 0.33 + \frac{1}{\Gamma(H+1/2)} \int_0^t (t-s)^{H-1/2} dW_s^i$, $i = 1, \dots, 3$.

One can also add multiple spreads by considering stochastic processes of the form

$$\lambda_t = \mathrm{tr}\left(X_t^\top \widetilde{Q} X_t\right) + \widetilde{\xi}(t), \quad t \leq T,$$

for some $\widetilde{Q} \in \mathbb{S}_+^d$ and $\widetilde{\xi} : [0,T] \to \mathbb{R}_+$ bounded function. By definition the spread is nonnegative, correlated to the short rate with a possible long range dependence or roughness. The introduction of λ can serve in two ways. Either in a multiple curve modeling framework, to add a risky curve on top of the non-risk one with instantaneous rate $r + \lambda$ or to model default time. In the latter case, λ would correspond to the instantaneous intensity of a Poisson process N such that the default time τ is defined as the first jump time of N. In both cases, we denote by $\widetilde{P}(\cdot, T)$ the price of the risky curve or the price of a defaultable bond paying $\mathbf{1}_{\tau \leq T}$ at maturity T. Then, on $\{t < \tau\}$, the price is given by

$$\begin{aligned}\widetilde{P}(t,T) &= \mathbb{E}\left[\exp\left(-\int_t^T (r_s + \lambda_s) ds\right) \,\middle|\, \mathcal{F}_t\right] \\ &= \exp\left(-\int_t^T (\xi(s) + \widetilde{\xi}(s)) ds\right) \mathcal{L}_{t,T}(Q + \widetilde{Q}),\end{aligned}$$

for all $t \leq T$, we refer to Lando (1998) for more details on the derivation of the defaultable bond price.

4.2 Pricing options on volatility/variance for basket products in Volterra Wishart covariance models

We consider $d \geq 1$ risky assets $S = (S^1, \dots, S^d)$ such that the instantaneous realized covariance is given by

$$d\langle \log S \rangle_t = X_t X_t^\top dt \tag{4.1}$$

where X is the $d \times m$ process as in (3.1). The following specifications for the dynamics of S fall into this framework.

Example 4.1. (i) The Volterra Wishart covariance model for $d = m$:

$$dS_t = \operatorname{diag}(S_t) X_t dB_t, \quad S_0 \in \mathbb{R}^d_+$$
$$X_t = g_0(t) + \int_0^t K(t,s) dW_s,$$

with $g_0 : [0,T] \to \mathbb{R}^{d\times d}$, a suitable measurable kernel $K : [0,T]^2 \to \mathbb{R}^{d\times d}$, a $d \times d$ Brownian motion W and

$$B^j = \operatorname{tr}\left(W_j \rho_j^\top\right) + \sqrt{1 - \operatorname{tr}\left(\rho_j \rho_j^\top\right)}\, W^{\perp,j}, \quad j = 1, \dots, d,$$

for some $\rho_j \in \mathbb{R}^{d\times m}$ such that $\operatorname{tr}\left(\rho_j \rho_j^\top\right) \leq 1$, for $j = 1, \dots, n$, where $W^\perp$ is a d–dimensional Brownian motion independent of W.

(ii) The Volterra Stein-Stein model when $d = m = 1$:

$$dS_t = S_t X_t dB_t, \quad S_0 > 0,$$
$$X_t = g_0(t) + \int_0^t K(t,s) dW_s, \quad d\langle B, W \rangle_t = \rho dt,$$

for some $\rho \in [-1, 1]$.

The approach of Carr and Lee (2008), based on Schürger (2002), can be adapted to price various volatility and variance options on basket products. Indeed, consider a basket product of the form

$$P_t^\alpha = \sum_{j=1}^d \alpha_j \log S_t^j = \alpha^\top \log S_t, \quad t \leq T,$$

for some $\alpha = (\alpha_1, \dots, \alpha_d)^\top \in \mathbb{R}^d$. It follows from (4.1) that the integrated realized variance Σ^α of P^α is given by

$$\Sigma_t^\alpha = \int_0^t \alpha^\top X_s X_s^\top \alpha ds = \int_0^t \operatorname{tr}\left(\alpha\alpha^\top X_s X_s^\top\right) ds, \quad t \leq T.$$

Fix $q \in (0,1]$ and consider the q-th power variance swap whose payoff at maturity T is given by

$$(\Sigma_T^\alpha)^q - F = \left(\int_0^T \operatorname{tr}\left(\alpha\alpha^\top X_s X_s^\top\right) ds \right)^q - F,$$

for some strike $F \geq 0$. In particular, for $q = 1/2$ one recovers a volatility swap and for $q = 1$ a variance swap. The value of the contract being null at $t = 0$, the fair strike F_q^* reads

$$F_q^* = \mathbb{E}\left[\left(\int_0^T \operatorname{tr}\left(\alpha\alpha^\top X_s X_s^\top\right) ds \right)^q \right].$$

The following proposition establishes the expression of the fair strike in terms of the Laplace transform provided by Theorem 3.3.

Proposition 4.2. *Assume that $0 < q < 1$, then the fair strike of the q-th power variance swap is given by*

$$F_q^* = \frac{q}{\Gamma(1-q)} \int_0^\infty \frac{1 - \mathcal{L}_{0,T}\left(z\alpha\alpha^\top\right)}{z^{q+1}} dz,$$

where $\mathcal{L}$ is given by (3.9). If $q = 1$, the fair strike for the variance swap reads

$$F_1^* = \int_0^T \operatorname{tr}\left(\alpha\alpha^\top g_0(s) g_0(s)^\top ds\right) ds + \int_0^T \int_0^s \operatorname{tr}\left(\alpha\alpha^\top K(s,u) K(s,u)^\top\right) du ds.$$

Proof. For $0 < q < 1$, applying the identity

$$v^q = \frac{q}{\Gamma(1-q)} \int_0^\infty \frac{1 - e^{-zv}}{z^{q+1}} dz, \quad 0 < q < 1, \quad v \geq 0,$$

see Schürger (2002), to $v = \int_0^T \operatorname{tr}\left(\alpha\alpha^\top X_s X_s^\top\right) ds$, taking expectation and invoking Tonelli's theorem together with Theorem 3.3 yield the claimed identity. For $q = 1$, one could proceed by differentiating the Laplace transform or more simply by using the dynamics of $XX^\top$ as in Remark 2.1. □

Similarly, one can obtain the following formulas for negative powers

$$\mathbb{E}\left[\left(\int_0^T \operatorname{tr}\left(\alpha\alpha^\top X_s X_s^\top\right) ds + \epsilon \right)^{-q} \right] = \frac{1}{\Gamma(1+q)} \int_0^\infty \mathcal{L}_{0,T}(z^{1/q}\alpha\alpha^\top) e^{-z^{1/q}\epsilon} dz, \quad \epsilon, q > 0,$$

using the integral representation, taken from Schürger (2002),

$$v^{-q} = \frac{1}{q\Gamma(1+q)} \int_0^\infty e^{-z^{1/q} v} dz, \quad q, v > 0.$$

Again, the approximation formulas of Remark 3.4 and Section 3.3 can be applied to compute $\mathcal{L}_{0,T}$.

A Wishart distribution

Proposition A.1. *Let ξ be an $\mathbb{R}^N$ Gaussian vector with mean vector $\mu \in \mathbb{R}^N$ and covariance matrix $\Sigma \in \mathbb{S}_+^N$, then $\xi\xi^\top$ follows a non-central Wishart distributions with shape parameter $1/2$, scale parameter 2Σ and non-centrality parameter $\mu\mu^\top$, written as*

$$\xi\xi^\top \sim WIS_N\left(\frac{1}{2}, \mu\mu^\top, 2\Sigma\right).$$

Furthermore,

$$\mathbb{E}\left[\exp\left(-\operatorname{tr}\left(u\xi\xi^\top\right)\right)\right] = \frac{\exp\left(-\operatorname{tr}\left(u(I_N + 2\Sigma u)^{-1}\mu\mu^\top\right)\right)}{\det\left(I_N + 2\Sigma u\right)^{1/2}}, \quad u \in \mathbb{S}_+^N.$$

B Matrix tools

We recall some definitions and properties of matrix tools used in the proofs throughout the article. For a complete review and proofs we refer to Magnus and Neudecker (2019).

Definition B.1. *The vectorization operator* vec *is defined from $\mathbb{R}^{d\times m}$ to $\mathbb{R}^{dm}$ by stacking the columns of a $d \times m$-matrix A one underneath another in a dm–dimensional vector* $\operatorname{vec}(A)$*, i.e.*

$$\operatorname{vec}(A)_p = A_{ij}, \quad p = (j-1)d + i,$$

for all $i = 1, \ldots, d$ and $j = 1, \ldots, m$.

Definition B.2. *Let $A \in \mathbb{R}^{d_1\times m_1}$ and $B \in \mathbb{R}^{d_2\times m_2}$. The Kronecker product $(A \otimes B)$ is defined as the $d_1d_2 \times m_1m_2$ matrix*

$$A \otimes B = \begin{pmatrix} A_{11}B & \cdots & A_{1m_1}B \\ \vdots & & \vdots \\ A_{d_11}B & \cdots & A_{d_1m_1}B \end{pmatrix}.$$

or equivalently

$$(A \otimes B)_{pq} = A_{ik}B_{jl}, \quad p = (i-1)d_2 + j, \quad q = (k-1)m_2 + l,$$

for all $i = 1, \ldots, d_1$, $j = 1, \ldots, d_2$, $k = 1, \ldots, m_1$ and $l = 1, \ldots, m_2$.

Proposition B.3. *For matrices A, B, C, D, X, w of suitable dimensions, the following relations hold:*

$$\operatorname{vec}(AXB) = \left(B^\top \otimes A\right) \operatorname{vec}(X) \tag{B.1}$$

$$\operatorname{tr}(A^\top wA) = \operatorname{vec}(A)^\top (I_m \otimes w) \operatorname{vec}(A) \tag{B.2}$$

$$(A \otimes B)(C \otimes D) = (AC \otimes BD) \tag{B.3}$$

$$\operatorname{tr}(A \otimes B) = \operatorname{tr}(A) \operatorname{tr}(B) \tag{B.4}$$

$$(A \otimes B)^{-1} = (A^{-1} \otimes B^{-1}) \tag{B.5}$$

$$\det(I_m \otimes A) = \det(A)^m. \tag{B.6}$$

C Proof of Theorem 2.7

Throughout this section we assume that the function $(s,u) \mapsto C_t(s,u)$ is continuous such that (2.28) holds, where C_t is given by (2.3).

For each $t \le T$, we consider the integral operator $\mathbf{C}_t$ induced by the kernel C_t

$$(\mathbf{C}_t f)(s) = \int_0^T C_t(s,u) f(u) du = \int_t^T C_t(s,u) f(u) du, \quad f \in L^2([0,T], \mathbb{R}^N), \quad s \le T,$$

where the last equality follows from the fact that $C_t(s,u) = 0$ for any $u \le t$. We assume that $t \mapsto \mathbf{C}_t$ is differentiable with derivative $\dot{\mathbf{C}}_t$ given by (2.21).

Lemma C.1. *Let $w \in \mathbb{S}_+^N$ and $t \mapsto R_{t,T}^w$ be defined as in (2.24). Then,*

$$\sup_{t \le T} \int_t^T \int_t^T |R_t^w(s,u)|^2 ds du < \infty, \tag{C.1}$$

$$\sup_{t \le T} \sup_{t \le s \le T} \int_t^T |R_t^w(s,u)|^2 du < \infty, \tag{C.2}$$

$$\sup_{t \le T} \sup_{t \le s,u \le T} |R_{t,T}^w(s,u)| < \infty. \tag{C.3}$$

Proof. Fix $t \le T$. It follows from (2.24) that

$$\int_t^T \int_t^T |R_t^w(s,u)|^2 ds du = \sum_{n \ge 1} \frac{4(\lambda_{t,T}^n)^2}{(1 + 2\lambda_{t,T}^n)^2} \le 4 \sum_{n \ge 1} (\lambda_{t,T}^n)^2 = 4|w| \int_t^T \int_t^T |C_t(s,u)|^2 ds du,$$

which, combined with (2.28), proves (C.1). Furthermore, an application of Jensen and Cauchy-Schwarz inequalities on the resolvent equation (2.25) yields

$$|R_t^w(s,u)|^2 \le 8 \sup_{t' \le T} \sup_{t' \le s', u' \le T} |C_{t',T}(s',u')|^2 \left(1 + T \int_t^T |R_t^w(z,u)|^2 dz\right), \quad t \le s, u \le T.$$

Integrating the previous identity with respect to u leads to

$$\int_t^T |R_t^w(s,u)|^2 du \le 8T \sup_{t' \le T} \sup_{t' \le s', u' \le T} |C_{t',T}(s',u')|^2 \left(1 + T \int_t^T \int_t^T |R_t^w(z,u)|^2 dz du\right),$$

for all $s \ge t$. Combined with (2.28) and (C.1), we obtain (C.2). Finally, it follows from the resolvent equation (2.25) together with Jensen and Cauchy-Schwarz inequalities that

$$|R_t(s,u)|^2 \le 8 \sup_{t' \le T} \sup_{t' \le s', u' \le T} |C_{t',T}(s',u')|^2 \left(1 + T \int_t^T |R_t^w(s,z)|^2 dz\right)$$

for all $t \le s, u \le T$. The right hand side is bounded by a finite quantity which does not depend on t, thanks to (2.28) and (C.2), yielding (C.3). □

Lemma C.2. *For each $t \le s \le T$, $u \mapsto R^w_{t,T}(s,u)$ is continuous. For each $s, u \le T$, $t \mapsto R^w_{t,T}(s,u)$ is continuous.*

Proof. The first statement follows directly from the continuity of $(s,u) \mapsto C_t(s,u)$ for all $t \le T$, the resolvent equation (2.25) and the dominated convergence theorem which is justified by (2.28). The second statement is proved as follows. Fix $t \le s, u \le T$ and $h \in \mathbb{R}$ such that $0 \le t + h \le T$. The resolvent equation (2.25) yields

$$\begin{aligned} R^w_{t+h}(s,u) - R^w_t(s,u) &= -2(C^w_{t+h}(s,u) - C^w_t(s,u)) \\ &\quad - 2\int_t^T R^w_{t+h}(s,z)(C^w_{t+h}(z,u) - C^w_t(z,u))dz \\ &\quad - 2\int_t^T (R^w_{t+h}(s,z) - R^w_t(s,z))C^w_t(z,u)dz \\ &\quad + 2\int_t^{t+h} R^w_{t+h}(s,z)C^w_{t+h}(z,u)dz \\ &= \mathbf{I} + \mathbf{II} + \mathbf{III} + \mathbf{IV} \end{aligned}$$

Since $t \mapsto C_t(s,u)$ is absolutely continuous, we have that $\mathbf{I} \to 0$ as $h \to 0$ and also that $\mathbf{II} \to 0$ by an application of Cauchy–Schwarz inequality, the bound (C.3), and the dominated convergence theorem, which is justified by (2.28). To prove that $\mathbf{III} \to 0$, we fix $q \in \mathbb{R}^N$ and $f_u(s) := C^w_t(s,u)q$. Then,

$$\int_t^T (R^w_{t+h}(s,z) - R^w_t(s,z))C^w_t(z,u)q dz = (\mathbf{R}^w_{t+h} f_u)(s) - (\mathbf{R}^w_t f_u)(s) \to 0, \quad \text{as } h \to 0,$$

where the convergence follows from the continuity of $t \mapsto \mathbf{R}^w_t$ obtained from that of $t \mapsto \mathbf{C}_t$, recall (2.26). By arbitrariness of q, we get $\mathbf{III} \to 0$. Finally, it follows from (2.28) and (C.3), that $\mathbf{IV} \to 0$ as $h \to 0$. Combining the above yields $R^w_{t+h}(s,u) \to R^w_t(s,u)$ as $h \to 0$. □

Lemma C.3. *$t \mapsto R^w_{t,T}(s,u)$ is absolutely continuous for almost every (s,u) such that*

$$\begin{aligned}\dot{R}^w_{t,T}(s,u) = &-2\sqrt{w}\dot{C}_{t,T}(s,u)\sqrt{w} - 2\int_t^T \sqrt{w}\dot{C}_{t,T}(s,z)\sqrt{w}R^w_{t,T}(z,u)dz \\ &- 2\int_t^T R^w_{t,T}(s,z)\sqrt{w}\dot{C}_{t,T}(z,u)\sqrt{w}dz \\ &- 2\int_t^T\int_t^T R^w_{t,T}(s,z)\sqrt{w}\dot{C}_{t,T}(z,z')\sqrt{w}R^w_{t,T}(z',u)dzdz', \quad \text{on } [t,T] \text{ a.e.}\end{aligned}$$

with the boundary condition

$$R^w_{t,T}(\cdot,t) = R^w_{t,T}(t,\cdot)^\top = 0, \quad t \le T. \tag{C.4}$$

Proof. The boundary condition (C.4) follows from the resolvent equation (2.25) and the fact that $C_t(\cdot,t) = C_t(t,\cdot)^\top = 0$, for all $t \le T$.
Step 1. It follows from (2.26) and the fact that $t \mapsto \mathbf{C}_t$ is differentiable, that $t \mapsto \mathbf{R}^w_{t,T}$ is differentiable, so that

$$(\mathbf{R}^w_{t+h,T}f)(s) = (\mathbf{R}^w_{t,T}f)(s) + h(\dot{\mathbf{R}}^w_{t,T}f)(s) + o(|h|), \quad f \in L^2([0,T],\mathbb{R}^N), \quad s \le T, \tag{C.5}$$

for all $h \in \mathbb{R}$ such that $0 \le t+h \le T$, with

$$\dot{\mathbf{R}}^w_{t,T} = -2(\mathrm{id} + \mathbf{R}^w_{t,T})\sqrt{w}\dot{\mathbf{C}}_{t,T}\sqrt{w}(\mathrm{id} + \mathbf{R}^w_{t,T}).$$

The right hand side being a composition of integral operators, $\dot{\mathbf{R}}^w_{t,T}$ is again an integral operator with kernel given by

$$-2(\delta + R^w_{t,T}) \star \sqrt{w}\dot{C}_{t,T}\sqrt{w} \star (\delta + R^w_{t,T}),$$

where by some abuse of notations δ denotes the kernel induced by the identity operator id, that is $(\mathrm{id}f)(s) = \int_t^T \delta_{s=u}(ds,du)f(u) = f(s)$.
Step 2. Fix f a measurable and bounded function, t, h such that $0 \le t+h \le T$, $s \le T$ and write

$$\begin{aligned}(\mathbf{R}^w_{t+h,T}f)(s,u) &= \int_{t+h}^T R^w_{t+h,T}(s,u)f(u)du \\ &= (\mathbf{R}^w_{t,T}f)(s,u) + \int_t^T \left(R^w_{t+h,T}(s,u) - R^w_{t,T}(s,u)\right) f(u)du \\ &\quad - \int_t^{t+h} \left(R^w_{t+h,T}(s,u) - R^w_{t,T}(s,u)\right) f(u)du \\ &\quad + \int_t^{t+h} \left(R^w_{t,T}(s,t) - R^w_{t,T}(s,u)\right) f(u)du \\ &= \mathbf{I} + \mathbf{II} + \mathbf{III} + \mathbf{IV}\end{aligned}$$

where we used the vanishing boundary condition (C.4) to introduce $R^w_{t,T}(s,t)$ in **IV**. Subtracting the previous equation to (C.5) yields

$$\mathbf{II} = h(\dot{\mathbf{R}}^w_{t,T} f)(s) - \mathbf{III} - \mathbf{IV} + o(|h|). \tag{C.6}$$

An application of the Heine–Cantor theorem yields that the continuity statements in Lemma C.2 can be strengthened to uniform continuity. Whence, for an arbitrary $\varepsilon > 0$ and for h small enough,

$$\sup_{u\in[t,t+h]} |R^w_{t,T}(s,t) - R^w_{t,T}(s,u)| + \sup_{u\in[t,t+h]} |R^w_{t+h,T}(s,u) - R^w_{t,T}(s,u)| \leq \varepsilon, \quad t \leq s \leq T.$$

This yields $|\mathbf{III}| + |\mathbf{IV}| \leq ch\varepsilon$, for some constant $c > 0$, so that taking limits in (C.6) gives

$$\lim_{h\to 0} \frac{1}{h}\mathbf{II} = (\dot{\mathbf{R}}^w_{t,T} f)(s).$$

An application of the dominated convergence theorem, which is justified by (C.3), yields that for any $u, s \leq T$ $t \mapsto R_t(s,u)$ is absolutely continuous with

$$\dot{R}^w_t(s,u) = -2(\delta + R^w_{t,T}) \star \sqrt{w}\dot{C}_{t,T}\sqrt{w} \star (\delta + R^w_{t,T}),$$

which is the claimed expression. □

We can now complete the proof of Theorem 2.7.

Proof of Theorem 2.7. The claimed expression for the Laplace transform follows from (2.20), the Riccati equation for Ψ as defined in (2.27) follows from Lemma C.3, and that of ϕ is straightforward from (2.22). □

References

Eduardo Abi Jaber. Lifting the Heston model. *Quantitative Finance*, pages 1–19, 2019.

Eduardo Abi Jaber and Omar El Euch. Multifactor approximation of rough volatility models. *SIAM Journal on Financial Mathematics*, 10(2):309–349, 2019a.

Eduardo Abi Jaber and Omar El Euch. Markovian structure of the Volterra Heston model. *Statistics & Probability Letters*, 149:63–72, 2019b.

Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200, 2019a.

Eduardo Abi Jaber, Enzo Miller, and Huyên Pham. Linear-Quadratic control for a class of stochastic Volterra equations: solvability and approximation. *arXiv:1911.01900, to appear in Annals of Applied Probability*, 2019b.

Eduardo Abi Jaber, Enzo Miller, and Huyên Pham. Integral operator Riccati equations arising in stochastic Volterra control problems. *arXiv:1911.01903, to appear in Siam Journal on Control and Optimization*, 2019c.

Aurélien Alfonsi. *Affine diffusions and related processes: simulation, theory and applications*, volume 6. Springer, 2015.

Aurélien Alfonsi and Alexander Schied. Capacitary measures for completely monotone kernels via singular control. *SIAM Journal on Control and Optimization*, 51(2):1758–1780, 2013.

Theodore Anderson and Donald Darling. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, 23(2):193–212, 1952.

David K Backus and Stanley E Zin. Long-memory inflation uncertainty: Evidence from the term structure of interest rates. Technical report, National Bureau of Economic Research, 1993.

Richard Bellman. Functional equations in the theory of dynamic programming–vii. a partial differential equation for the Fredholm resolvent. *Proceedings of the American Mathematical Society*, 8(3):435–440, 1957.

Mikkel Bennedsen, Asger Lunde, and Mikko S Pakkanen. Decoupling the short-and long-term behavior of stochastic volatility. *arXiv preprint arXiv:1610.00332*, 2016.

Fred Espen Benth and Victor Rohde. On non-negative modeling with CARMA processes. *Journal of Mathematical Analysis and Applications*, 2018.

Folkmar Bornemann. On the numerical evaluation of Fredholm determinants. *Mathematics of Computation*, 79(270):871–915, 2010.

Haim Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, 2010.

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. *The Journal of Finance*, 65(1):393–420, 2010.

RH Cameron and MD Donsker. Inversion formulae for characteristic functionals of stochastic processes. *Annals of Mathematics*, pages 15–36, 1959.

Philippe Carmona, Laure Coutin, and Gérard Montseny. Approximation of some Gaussian processes. *Statistical inference for stochastic processes*, 3(1-2):161–171, 2000.

Peter Carr and Roger Lee. Robust replication of volatility derivatives. In *Prmia award for best paper in derivatives, mfa 2008 annual meeting*. Citeseer, 2008.

Li Chen, Damir Filipović, and H Vincent Poor. Quadratic term structure models for risk-free and defaultable rates. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 14(4):515–536, 2004.

Peng Cheng and Olivier Scaillet. Linear-quadratic jump-diffusion modeling. *Mathematical Finance*, 17(4):575–598, 2007.

José Manuel Corcuera, Gergely Farkas, Wim Schoutens, and Esko Valkeila. A short rate model using ambit processes. In *Malliavin Calculus and Stochastic Analysis*, pages 525–553. Springer, 2013.

John C Cox, Jonathan Jr E Ingersoll, and Stephen A Ross. A theory of the term structure of interest rates. In *Theory of Valuation*, pages 129–164. World Scientific, 2005.

Christa Cuchiero and Josef Teichmann. Markovian lifts of positive semidefinite affine Volterra-type processes. *Decisions in Economics and Finance*, 42(2):407–448, 2019.

Christa Cuchiero, Damir Filipović, Eberhard Mayerhofer, and Josef Teichmann. Affine processes on positive semidefinite matrices. *The Annals of Applied Probability*, 21(2):397–463, 2011.

Christa Cuchiero, Claudio Fontana, and Alessandro Gnoatto. Affine multiple yield curve models. *Mathematical Finance*, 2016.

José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10(2):151–180, 2007.

José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. A multifactor volatility Heston model. *Quantitative Finance*, 8(6):591–604, 2008.

Qiang Dai and Kenneth Singleton. Term structure dynamics in theory and reality. *The Review of financial studies*, 16(3):631–678, 2003.

Laurent Decreusefond and Ali Suleyman Ustunel. Stochastic analysis of the fractional Brownian motion. *Potential analysis*, 10(2):177–214, 1999.

Ivar Fredholm. Sur une classe d'équations fonctionnelles. *Acta mathematica*, 27(1):365–390, 1903.

Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

Alessandro Gnoatto and Martino Grasselli. The explicit Laplace transform for the Wishart process. *Journal of Applied Probability*, 51(3):640–656, 2014.

MA Golberg. A generalization of a formula of Bellman and Krein. *Journal of Mathematical Analysis and Applications*, 42(3):513–521, 1973.

Christian Gourieroux and Razvan Sufana. Wishart quadratic term structure models. *Les Cahiers du CREF of HEC Montreal Working Paper*, (03-10), 2003.

Christian Gouriéroux, Joann Jasiak, and Razvan Sufana. The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics*, 150(2):167–181, 2009.

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. ISBN 0-521-37289-5.

Archil Gulisashvili. Large deviation principle for Volterra type fractional stochastic volatility models. *SIAM Journal on Financial Mathematics*, 9(3):1102–1136, 2018.

Archil Gulisashvili, Frederi Viens, and Xin Zhang. Extreme-strike asymptotics for general gaussian stochastic volatility models. *Annals of Finance*, 15(1):59–101, 2019.

Philipp Harms. Strong convergence rates for numerical approximations of fractional Brownian motion. *arXiv preprint arXiv:1902.01471*, 2019.

Philipp Harms and David Stefanovits. Affine representations of fractional processes with applications in mathematical finance. *Stochastic Processes and their Applications*, 129(4):1185–1228, 2019.

John Hull and Alan White. Pricing interest-rate-derivative securities. *The review of financial studies*, 3(4):573–592, 1990.

Mark Kac and Arnold JF Siegert. On the theory of noise in radio receivers with square law detectors. *Journal of Applied Physics*, 18(4):383–397, 1947.

Kari Karhunen. Zur spektraltheorie stochastischer prozesse. *Ann. Acad. Sci. Fennicae, AI*, 34, 1946.

ML Kleptsyna, A Le Breton, and M Viot. New formulas concerning Laplace transforms of quadratic forms for general Gaussian sequences. *International Journal of Stochastic Analysis*, 15(4):309–325, 2002.

MG Krein. On a new method for solving linear integral equations of the first and second kinds. In *Dokl. Akad. Nauk SSSR*, volume 100, pages 413–414, 1955.

David Lando. On Cox processes and credit risky securities. *Review of Derivatives research*, 2(2-3): 99–120, 1998.

JJ Levin. On the matrix Riccati equation. *Proceedings of the American Mathematical Society*, 10 (4):519–524, 1959.

Michel Loeve. Probability theory: foundations, random sequences. 1955.

Jan R Magnus and Heinz Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. John Wiley & Sons, 2019.

Johannes Muhle-Karbe, Oliver Pfaffel, and Robert Stelzer. Option pricing in multivariate stochastic volatility models of OU type. *SIAM Journal on Financial Mathematics*, 3(1):66–94, 2012.

Szymon Peszat and Jerzy Zabczyk. *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2007.

Peter Ritchken and Iyuan Chuang. Interest rate option pricing with volatility humps. *Review of Derivatives Research*, 3(3):237–262, 2000.

Rainer Schöbel and Jianwei Zhu. Stochastic volatility with an Ornstein–Uhlenbeck process: an extension. *Review of Finance*, 3(1):23–46, 1999.

Alan Schumitzky. On the equivalence between matrix Riccati equations and Fredholm resolvents. *Journal of Computer and System Sciences*, 2(1):76–87, 1968.

Klaus Schürger. Laplace transforms and suprema of stochastic processes. In *Advances in Finance and Stochastics*, pages 285–294. Springer, 2002.

Galen R Shorack and Jon A Wellner. *Empirical processes with applications to statistics*. SIAM, 2009.

Barry Simon. Notes on infinite determinants of Hilbert space operators. *Advances in Mathematics*, 24(3):244–273, 1977.

Tommi Sottinen and Lauri Viitasaari. Stochastic analysis of Gaussian processes via Fredholm representation. *International journal of stochastic analysis*, 2016, 2016.

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.

Dale E Varberg. Convergence of quadratic forms in independent random variables. *The Annals of Mathematical Statistics*, 37(3):567–576, 1966.

Jia Yue and Nan-jing Huang. Fractional Wishart processes and ε-fractional Wishart processes with applications. *Computers & Mathematics with Applications*, 75(8):2955–2977, 2018.