Affine Volterra processes

Eduardo Abi Jaber, Martin Larsson, Sergio Pulido

Affine Volterra processes*

Eduardo Abi Jaber† Martin Larsson‡ Sergio Pulido§

August 29, 2017

Abstract

We introduce affine Volterra processes, defined as solutions of certain stochastic convolution equations with affine coefficients. Classical affine diffusions constitute a special case, but affine Volterra processes are neither semimartingales, nor Markov processes in general. We provide explicit exponential-affine representations of the Fourier–Laplace functional in terms of the solution of an associated system of deterministic integral equations, extending well-known formulas for classical affine diffusions. For specific state spaces, we prove existence, uniqueness, and invariance properties of solutions of the corresponding stochastic convolution equations. Our arguments avoid infinite-dimensional stochastic analysis as well as stochastic integration with respect to non-semimartingales, relying instead on tools from the theory of finite-dimensional deterministic convolution equations. Our findings generalize and simplify recent results in the literature on rough volatility models in finance.

1 Introduction

We study a class of d-dimensional stochastic convolution equations of the form

$$X_t = X_0 + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s, \tag{1.1}$$

*Acknowledgments: The authors wish to thank Bruno Bouchard, Omar El Euch, Camille Illand, and Mathieu Rosenbaum for useful comments and fruitful discussions. Martin Larsson gratefully acknowledges financial support by the Swiss National Science Foundation (SNF) under grant 205121_163425. The research of Sergio Pulido benefited from the support of the Chair Markets in Transition (Fédération Bancaire Française) and the project ANR 11-LABX-0019.

†Université Paris-Dauphine, PSL Research University, CNRS, UMR [7534], CEREMADE, 75016 Paris, France and AXA Investment Managers, Multi Asset Client Solutions, Quantitative Research, 6 place de la Pyramide, 92908 Paris - La Défense, France, abijaber@ceremade.dauphine.fr.

‡Department of Mathematics, ETH Zurich, Rämistrasse 101, CH-8092, Zurich, Switzerland, martin.larsson@math.ethz.ch.

§Laboratoire de Mathématiques et Modélisation d'Évry (LaMME), Université d'Évry-Val-d'Essonne, ENSIIE, Université Paris-Saclay, UMR CNRS 8071, IBGBI 23 Boulevard de France, 91037 Évry Cedex, France, sergio.pulidonino@ensiie.fr.

where W is a multi-dimensional Brownian motion, and the convolution kernel K and coefficients b and σ satisfy regularity and integrability conditions that are discussed in detail after this introduction. We refer to equations of the form (1.1) as *stochastic Volterra equations* (of convolution type), and their solutions are always understood to be adapted processes defined on some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions. Stochastic Volterra equations have been studied by numerous authors; see e.g. Berger and Mizel (1980a,b); Protter (1985); Pardoux and Protter (1990); Zhang (2010) among many others.

We are interested in the situation where $a(x) = \sigma(x)\sigma(x)^\top$ and $b(x)$ are affine maps given by

$$\begin{aligned} a(x) &= A^0 + x_1 A^1 + \cdots + x_d A^d \\ b(x) &= b^0 + x_1 b^1 + \cdots + x_d b^d, \end{aligned} \tag{1.2}$$

for some d-dimensional symmetric matrices A^i and vectors b^i. In this case we refer to solutions of (1.1) as *affine Volterra processes*. *Affine diffusions*, as studied in Duffie et al. (2003), are particular examples of affine Volterra processes of the form (1.1) where the convolution kernel $K \equiv \mathrm{id}$ is constant and equal to the d-dimensional identity matrix. In this paper we do not consider processes with jumps.

Stochastic models using classical affine diffusions are tractable because their Fourier–Laplace transform has a simple form. It can be written as an exponential-affine function of the initial state, in terms of the solution of a system of ordinary differential equations, known as the Riccati equations, determined by the affine maps (1.2). More precisely, let X be an affine diffusion of the form (1.1) with $K \equiv \mathrm{id}$. Then, given a d-dimensional row vector u and under suitable integrability conditions, we have

$$\mathbb{E}\left[\exp\left(uX_T\right) \;\middle|\; \mathcal{F}_t\right] = \exp\left(\phi(T-t) + \psi(T-t)X_t\right), \tag{1.3}$$

where the real-valued function ϕ and row-vector-valued function ψ satisfy the Riccati equations

$$\begin{aligned} \phi(t) &= \int_0^t \left(\psi(s)b_0 + \frac{1}{2}\psi(s)A_0\psi(s)^\top\right) ds \\ \psi(t) &= u + \int_0^t \left(\psi(s)B + \frac{1}{2}A(\psi(s))\right) ds, \end{aligned}$$

with $A(u) = (uA^1u^\top, \ldots, uA^du^\top)$ and $B = (b^1 \;\cdots\; b^d)$. Alternatively, using the variation of constants formula on X and ψ, one can write the Fourier-Laplace transform as

$$\mathbb{E}\left[\exp\left(uX_T\right) \;\middle|\; \mathcal{F}_t\right] = \exp\left(\mathbb{E}[uX_T \mid \mathcal{F}_t] + \frac{1}{2}\int_t^T \psi(T-s)a(\mathbb{E}[X_s \mid \mathcal{F}_t])\psi(T-s)^\top ds\right). \tag{1.4}$$

For more general kernels K, affine Volterra processes are typically neither semimartingales, nor Markov processes. Therefore one cannot expect a formula like (1.3) to hold in

general. However, we show in Theorem 4.3 below that, remarkably, (1.4) does continue to hold, where now the function ψ solves the *Riccati–Volterra equation*

$$\psi(t) = uK(t) + \int_0^t \left(\psi(s)B + \frac{1}{2}A(\psi(s))\right) K(t-s)\, ds. \tag{1.5}$$

Furthermore, it is possible to obtain an explicit expression that is exponential-affine in the past trajectory $\{X_s,\, s \le t\}$.

For the state spaces $\mathbb{R}^d$, $\mathbb{R}^d_+$, and $\mathbb{R} \times \mathbb{R}_+$, corresponding to the *Volterra Ornstein–Uhlenbeck*, *Volterra square-root*, and *Volterra Heston* models, we establish existence and uniqueness of global solutions of both the stochastic equation (1.1) and the associated Riccati–Volterra equation (1.5), under general parameter restrictions. For the state spaces $\mathbb{R}^d_+$ and $\mathbb{R} \times \mathbb{R}_+$ this involves rather delicate invariance properties for these equations. While standard martingale and stochastic calculus arguments play an important role in several places, the key tools that allow us to handle the lack of Markov and semimartingale structure are the *resolvents of first and second kind* associated with the convolution kernel K. Let us emphasize in particular that no stochastic integration with respect to non-semimartingales is needed. Furthermore, by performing the analysis on the level of finite-dimensional integral equations, we avoid the infinite-dimensional analysis used, for instance, by Mytnik and Salisbury (2015). We also circumvent the need to study scaling limits of Hawkes processes as in El Euch and Rosenbaum (2016); El Euch et al. (2016); El Euch and Rosenbaum (2017).

Our motivation for considering affine Volterra processes comes from applications in financial modeling. Classical affine processes arguably constitute the most popular framework for building tractable multi-factor models in finance. They have been used to model a vast range of risk factors such as credit and liquidity factors, inflation and other macroeconomic factors, equity factors, and factors driving the evolution of interest rates. In particular, affine stochastic volatility models, such as the Heston (1993) model, are very popular.

However, a growing body of empirical research indicates that volatility fluctuates more rapidly than Brownian motion, which is inconsistent with standard semimartingale affine models. Fractional volatility models such as those by Guennoun et al. (2017); Gatheral et al. (2014); Bayer et al. (2016); El Euch and Rosenbaum (2016); Bennedsen et al. (2016) have emerged as compelling alternatives, although tractability can be a challenge for these non-Markovian, non-semimartingales models. Nonetheless, Guennoun et al. (2017) and El Euch and Rosenbaum (2016, 2017) show that there exist fractional adaptations of the Heston model where the Fourier–Laplace transform can be found explicitly, modulo the solution of a certain *fractional Riccati equation*. These models are of the affine Volterra type (1.1) involving singular kernels proportional to $t^{\alpha-1}$ for $\alpha \in (\frac{1}{2}, 1)$. Our framework subsumes and extends these examples.

The paper is structured as follows. Section 2 covers preliminaries on convolutions and their resolvents, and in particular develops the necessary stochastic calculus. Section 3

gives existence theorems for stochastic Volterra equations on $\mathbb{R}^d$ and $\mathbb{R}^d_+$. Section 4 introduces affine Volterra processes on general state spaces and develops the exponential-affine transform formula. Sections 5 through 7 contain detailed discussions for the state spaces $\mathbb{R}^d$, $\mathbb{R}^d_+$, and $\mathbb{R} \times \mathbb{R}_+$, which correspond to the Volterra Ornstein–Uhlenbeck, Volterra square-root, and Volterra Heston models, respectively. Additional proofs and supporting results are presented in the appendices. Our basic reference for the deterministic theory of Volterra equations is the excellent book by Gripenberg et al. (1990).

Notation

Throughout the paper we view elements of $\mathbb{R}^m$ and $\mathbb{C}^m = \mathbb{R}^m + \mathrm{i}\mathbb{R}^m$ as column vectors, while elements of the dual spaces $(\mathbb{R}^m)^*$ and $(\mathbb{C}^m)^*$ are viewed as row vectors. For any matrix A with complex entries, $A^\top$ denotes the (ordinary, not conjugate) transpose of A. The symbol $|\cdot|$ is used to denote the Euclidean norm on $\mathbb{C}^m$ and $(\mathbb{C}^m)^*$, as well as the operator norm on $\mathbb{R}^{m\times n}$. We write $\mathbb{S}^m$ for the symmetric $m \times m$ matrices. The shift operator Δ_h, where $h \geq 0$, maps any function f on $\mathbb{R}_+$ to the function $\Delta_h f$ given by

$$\Delta_h f(t) = f(t+h).$$

If the function f on $\mathbb{R}_+$ is of bounded variation, its measure derivative is denoted by df, so that $f(t) = f(0) + \int_{[0,t]} df(s)$.

2 Stochastic calculus of convolutions and resolvents

For a measurable function K on $\mathbb{R}_+$ and a measure L on $\mathbb{R}_+$ of locally bounded variation, the convolutions $K * L$ and $L * K$ are defined by

$$(K * L)(t) = \int_{[0,t]} K(t-s)L(ds), \qquad (L * K)(t) = \int_{[0,t]} L(ds)K(t-s) \tag{2.1}$$

whenever these expressions are well-defined. We allow K and L to be matrix-valued, in which case $K * L$ and $L * K$ may not both be defined (e.g. due to incompatible matrix dimensions), or differ from each other even if they are defined (e.g. if K and L take values among non-commuting square matrices). If F is a function on $\mathbb{R}_+$, we write $K * F = K * (F\,dt)$, that is,

$$(K * F)(t) = \int_0^t K(t-s)F(s)ds. \tag{2.2}$$

Further details can be found in Gripenberg et al. (1990), see in particular Definitions 2.2.1 and 3.2.1, as well as Theorems 2.2.2 and 3.6.1 for a number of properties of convolutions. In particular, if $K \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$ and F is continuous, then $K * F$ is again continuous.

Fix $d \in \mathbb{N}$ and let M be a d-dimensional continuous local martingale. If K is $\mathbb{R}^{m\times d}$-valued for some $m \in \mathbb{N}$, the convolution

$$(K * dM)_t = \int_0^t K(t-s)dM_s \tag{2.3}$$

is well-defined as an Itô integral for every $t \geq 0$ such that $\int_0^t |K(t-s)|^2 d\,\mathrm{Tr}\langle M\rangle_s < \infty$. In particular, if $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$ and $\langle M\rangle_s = \int_0^s a_u du$ for some locally bounded process a, then (2.3) is well-defined for every $t \geq 0$. We always choose a version that is jointly measurable in (t,ω). Just like (2.1)–(2.2), the convolution (2.3) is associative, as the following result shows.

Lemma 2.1. *Let $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{m\times d})$ and let L be an $\mathbb{R}^{n\times m}$-valued measure on $\mathbb{R}_+$ of locally bounded variation. Let M be a d-dimensional continuous local martingale with $\langle M\rangle_t = \int_0^t a_s ds$, $t \geq 0$, for some locally bounded process a. Then*

$$(L * (K * dM))_t = ((L * K) * dM)_t \tag{2.4}$$

for every $t \geq 0$. In particular, taking $F \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$ we may apply (2.4) with $L(dt) = F dt$ to obtain $(F * (K * dM))_t = ((F * K) * dM)_t$.

Proof. By linearity it suffices to take $d = m = n = 1$ and L a locally finite positive measure. In this case,

$$(L * (K * dM))_t = \int_0^t \left(\int_0^t \mathbf{1}_{\{u<t-s\}} K(t-s-u) dM_u \right) L(ds).$$

Since

$$\int_0^t \left(\int_0^t \mathbf{1}_{\{u<t-s\}} K(t-s-u)^2 d\langle M\rangle_u \right)^{1/2} L(ds) \leq \max_{0\leq s\leq t} |a_s|^{1/2} \|K\|_{L^2(0,t)} L([0,t]),$$

which is finite almost surely, the stochastic Fubini theorem, see Veraar (2012, Theorem 2.2), yields

$$(L * (K * dM))_t = \int_0^t \left(\int_0^t \mathbf{1}_{\{u<t-s\}} K(t-s-u) L(ds) \right) dM_u = ((L * K) * dM)_t,$$

as required. □

Under additional assumptions on the kernel K one can find a version of the convolution (2.3) that is continuous in t. We will use the following condition:

$$\begin{gathered} K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}) \text{ and there is } \gamma \in (0,2] \text{ such that } \int_0^h K(t)^2 dt = O(h^\gamma) \\ \text{and } \int_0^T (K(t+h) - K(t))^2 dt = O(h^\gamma) \text{ for every } T < \infty. \end{gathered} \tag{2.5}$$

Lemma 2.2. *Assume K satisfies (2.5) and consider a continuous local martingale M with $\langle M\rangle_t = \int_0^t a_s ds$ for some locally bounded predictable process a. Then the process $\{\int_0^t K(t-s)dM_s,\, t \geq 0\}$ admits a continuous version, denoted $K * dM$, which is Hölder continuous of any order $\alpha < \gamma/2$.*

Proof. The proof is an application of Kolmogorov's continuity theorem. Since a is continuous, we may assume upon localizing M that $a \leq \kappa$ for some constant κ. It is enough to prove continuity up to an arbitrary time $T < \infty$. For any $s < t \leq T$ and $p \geq 1$, we have

$$\begin{aligned}\mathbb{E}\left[|(K*dM)_t - (K*dM)_s|^p\right] \leq 2^{p-1}\mathbb{E}\left[\left|\int_s^t K(t-u)dM_u\right|^p\right] \\ + 2^{p-1}\mathbb{E}\left[\left|\int_0^s (K(t-u)-K(s-u))dM_u\right|^p\right].\end{aligned}$$

The BDG inequalities applied to the local martingale $r \mapsto \int_0^r K(t-u)dM_u$ for $0 \leq r \leq t$, the bound on a, and a change of variables yield

$$\begin{aligned}\mathbb{E}\left[\left|\int_s^t K(t-u)dM_u\right|^p\right] &\leq C_p\,\mathbb{E}\left[\left(\int_s^t K(t-u)^2 a_u du\right)^{p/2}\right] \\ &\leq C_p\,\kappa^{p/2}\left(\int_s^t K(t-u)^2 du\right)^{p/2} \\ &= C_p\,\kappa^{p/2}\left(\int_0^{t-s} K(u)^2 du\right)^{p/2}.\end{aligned}$$

Similarly,

$$\mathbb{E}\left[\left|\int_0^s (K(t-u)-K(s-u))dM_u\right|^p\right] \leq C_p\,\kappa^{p/2}\left(\int_0^s (K(u+t-s)-K(u))^2 du\right)^{p/2}.$$

Therefore, due to (2.5),

$$\mathbb{E}\left[|(K*dM)_t - (K*dM)_s|^p\right] \leq c\,(t-s)^{\gamma p/2}$$

for some constant c that depends on T, κ, and p, but not on s, t. For any $p > 2/\gamma$, Kolmogorov's continuity theorem, see Revuz and Yor (1999, Theorem I.2.1), then implies that $\{\int_0^t K(t-s)dM_s, t \geq 0\}$ has a version Y^p that is Hölder continuous of any order $\alpha < \gamma/2 - 1/p$. Since any two continuous versions agree except on a nullset of paths, we may find a nullset N such that $Y^p = Y^q$ outside N for all integers $p, q > 2/\gamma$. In particular, outside N, the paths of Y^p are actually Hölder continuous of any order $\alpha < \gamma/2$. Setting $K * dM = Y^p$ outside N and $K * dM = 0$ on N therefore gives the stated properties. □

Remark 2.3. *Other conditions than (2.5) have appeared in the literature. For non-convolution kernels, Decreusefond (2002) considers $dM = \sigma dW$ defined on the Wiener space with (one-dimensional) coordinate process W, and requires $F \mapsto K * F$ to be continuous from certain L^p spaces to appropriate Besov spaces. Whereas in Mytnik and Neuman (2011), K is assumed to be a function of smooth variation and M is a semimartingale. See also (Wang, 2008, Theorem 1.3).*

Example 2.4. *Let us list some examples of kernels that satisfy (2.5):*

(i) *Locally Lipschitz kernels K clearly satisfy (2.5) with $\gamma = 1$.*

(ii) *The fractional kernel $K(t) = t^{\alpha-1}$ with $\alpha \in (\frac{1}{2}, 1)$ satisfies (2.5) with $\gamma = 2\alpha - 1$. Indeed, it is locally square integrable, and we have $\int_0^h K(t)dt = h^{2\alpha-1}/(2\alpha-1)$ as well as*

$$\int_0^T (K(t+h) - K(t))^2 dt \leq h^{2\alpha-1} \int_0^\infty \left((t+1)^{\alpha-1} - t^{\alpha-1}\right)^2 dt,$$

where the constant on the right-hand side is bounded by $\frac{1}{2\alpha-1} + \frac{1}{5-2\alpha}$. Note that the case $\alpha \geq 1$ falls in the locally Lipschitz category mentioned previously.

(iii) *If K_1 and K_2 satisfy (2.5), then so does $K_1 + K_2$.*

(iv) *If K_1 satisfies (2.5) for some $\gamma \in (0, 2]$ and K_2 is locally Lipschitz, then $K = K_1 K_2$ satisfies (2.5). Indeed, letting $\|K_2^2\|_{\infty,T}$ denote the maximum of K_2^2 over $[0, T]$ and $\mathrm{Lip}_T(K_2)$ the best Lipschitz constant on $[0, T]$, we have*

$$\int_0^h K(t)^2 dt \leq \|K_2^2\|_{\infty,h} \int_0^h K_1(t)dt = O(h^\gamma)$$

and

$$\begin{aligned}\int_0^T (K(t+h) - K(t))^2 dt \leq{}& 2\|K_2\|_{\infty,T+h}^2 \int_0^T (K_1(t+h) - K_1(t))^2 dt \\ &+ 2\|K_1\|_{L^2(0,T)}^2 \mathrm{Lip}_T(K_2)^2 h^2 \\ ={}& O(h^\gamma).\end{aligned}$$

In particular, K again satisfies (2.5) with the same γ.

(v) *By combining the above examples we find that, for instance, exponentially damped and possibly singular kernels like the Gamma kernel $K(t) = t^{\alpha-1}e^{-\beta t}$ for $\alpha > \frac{1}{2}$ and $\beta \geq 0$ satisfy (2.5).*

Under suitable moment bounds, the qualitative Hölder continuity statement in Lemma 2.2 can be improved to a quantitative estimate of the corresponding Hölder seminorms.

Lemma 2.5. *Assume K satisfies (2.5) and consider a continuous process X of the form $X = X_0 + K * (bdt + dM)$, where b is a predictable process and M is a continuous local martingale with $\langle M\rangle_t = \int_0^t a_s ds$ for some predictable process a. Then for $T < \infty$ and $p \geq 2$ one has*

$$\mathbb{E}\left[\left(\sup_{0\leq s<t\leq T}\frac{|X_t - X_s|}{|t-s|^\alpha}\right)^p\right] \leq c \sup_{t\leq T}\mathbb{E}[|a_t|^p + |b_t|^p]$$

for all $\alpha \in [0, \gamma/2 - 1/p)$, where c is a constant that only depends on p, K, and T.

Proof. We assume that the right-hand side is finite since otherwise the statement is trivial. For any $p \geq 2$ and any $s < t \leq T < \infty$ we have

$$\begin{aligned}
|X_t - X_s|^p &\leq 4^{p-1}\left|\int_s^t K(t-u)b_u du\right|^p + 4^{p-1}\left|\int_0^s (K(t-u) - K(s-u))\, b_u du\right|^p \\
&\quad + 4^{p-1}\left|\int_s^t K(t-u) dM_u\right|^p + 4^{p-1}\left|\int_0^s (K(t-u) - K(s-u))\, dM_u\right|^p \\
&= 4^{p-1}\left(\mathbf{I} + \mathbf{II} + \mathbf{III} + \mathbf{IV}\right).
\end{aligned}$$

Jensen's inequality applied twice yields $\mathbf{I} \leq (t-s)^{p/2}\,\|K\|^{p-2}_{L^2(s,t)}\int_s^t |b_u|^p\, K(t-u)^2 du$. Taking expectations and changing variables we obtain

$$\mathbb{E}[\,\mathbf{I}\,] \leq (t-s)^{p/2}\left(\int_0^{t-s} K(u)^2 du\right)^{p-1}\sup_{u\leq T}\mathbb{E}[|b_u|^p]. \tag{2.6}$$

In a similar manner,

$$\mathbb{E}[\,\mathbf{II}\,] \leq (t-s)^{p/2}\left(\int_0^{s} (K(u+t-s) - K(u))^2 du\right)^{p-1}\sup_{u\leq T}\mathbb{E}[|b_u|^p]. \tag{2.7}$$

Using in addition the BDG inequalities we get

$$\begin{aligned}
\mathbb{E}\,[\,\mathbf{III}\,] &\leq C_p\,\mathbb{E}\left[\left(\int_s^t K(t-u)^2\, a_u\, du\right)^{p/2}\right] \\
&\leq C_p\left(\int_0^{t-s} K(u)^2 du\right)^{p/2}\sup_{u\leq T}\mathbb{E}[|a_u|^p]
\end{aligned} \tag{2.8}$$

and

$$\mathbb{E}\,[\,\mathbf{IV}\,] \leq C_p\left(\int_0^{s} (K(u+t-s) - K(u))^2 du\right)^{p/2}\sup_{u\leq T}\mathbb{E}[|a_u|^p]. \tag{2.9}$$

Combining (2.6)–(2.9) with (2.5) yields

$$\mathbb{E}\,[|X_t - X_s|^p] \leq c' \sup_{u\leq T}\mathbb{E}[|a_u|^p + |b_u|^p]\,(t-s)^{\gamma p/2},$$

where c' is a constant that only depends on p, K, and T, but not on s or t. The statement now follows from the Komogorov continuity theorem; see Revuz and Yor (1999, Theorem I.2.1). □

Consider a kernel $K \in L^1_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{m\times m})$. The *resolvent*, or *resolvent of the second kind*, corresponding to K is the kernel $R \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}^{m\times m})$ such that

$$K * R = R * K = K - R. \tag{2.10}$$

The resolvent always exists and is unique, and a number of properties such as (local) square integrability and continuity of the original kernel K are inherited by its resolvent; see Gripenberg et al. (1990, Theorems 2.3.1 and 2.3.5).[1] Another object related to K is its *resolvent of the first kind*, which is an $\mathbb{R}^{d\times d}$-valued measure L on $\mathbb{R}_+$ of locally bounded variation such that

$$K * L = L * K \equiv \mathrm{id}, \tag{2.11}$$

see Gripenberg et al. (1990, Definition 5.5.1). Some examples of resolvents of the first and second kind are presented in Table 1. A resolvent of the first kind does not always exist. When it does, it has the following properties, which play a key role in several of our arguments.

Lemma 2.6. *Let X be a continuous process and $Z = \int b\,dt + \int \sigma\,dW$ a continuous semimartingale with b, σ, and $K * dZ$ continuous. Assume that K admits a resolvent of the first kind L. Then*

$$X - X_0 = K * dZ \qquad \Longleftrightarrow \qquad L * (X - X_0) = Z. \tag{2.12}$$

In this case, for any $F \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{C}^{m\times d})$ such that $F * L$ is of locally bounded variation, one has

$$F * dZ = (F * L)(0)X - (F * L)X_0 + d(F * L) * X. \tag{2.13}$$

Proof. Assume $X - X_0 = K * dZ$. Apply L to both sides to get

$$L * (X - X_0) = L * (K * dZ) = (L * K) * dZ = \mathrm{id} * dZ = Z,$$

where the second equality follows from Lemma 2.1. This proves the forward implication in (2.12). Conversely, assume $L * (X - X_0) = Z$. Then,

$$\begin{aligned} \mathrm{id} * (X - X_0) &= (K * L) * (X - X_0) \\ &= K * (L * (X - X_0)) \\ &= K * Z \\ &= K * (\mathrm{id} * dZ) \\ &= \mathrm{id} * (K * dZ), \end{aligned}$$

[1] Rather than (2.10), it is common to require $K * R = R * K = R - K$ in the definition of resolvent. We use (2.10) to remain consistent with Gripenberg et al. (1990).

using Gripenberg et al. (1990, Theorem 3.6.1(ix)) for the second equality and Lemma 2.1 for the last equality. Since both $X - X_0$ and $K * dZ$ are continuous, they must be equal.

To prove (2.13), observe that the assumption of locally bounded variation entails that

$$F * L = (F * L)(0) + d(F * L) * \mathrm{id}.$$

Convolving this with K, using (2.11), and inspecting the densities of the resulting absolutely continuous functions, we get

$$F = (F * L)(0)K + d(F * L) * K.$$

Using (2.4) and the fact that $K * dZ = X - X_0$ by assumption, it follows that

$$\begin{aligned} F * dZ &= (F * L)(0)K * dZ + d(F * L) * (K * dZ) \\ &= (F * L)(0)X - (F * L)X_0 + d(F * L) * X, \end{aligned}$$

as claimed. □

	$K(t)$	$R(t)$	$L(dt)$
Constant	c	ce^{-ct}	$c^{-1}\delta_0(dt)$
Fractional	$c\,\frac{t^{\alpha-1}}{\Gamma(\alpha)}$	$ct^{\alpha-1}E_{\alpha,\alpha}(-ct^\alpha)$	$c^{-1}\,\frac{t^{-\alpha}}{\Gamma(1-\alpha)}dt$
Exponential	$ce^{-\lambda t}$	$ce^{-(c+\lambda)t}$	$c^{-1}(\delta_0(dt) + \lambda\,dt)$
Gamma	$ce^{-\lambda t}\frac{t^{\alpha-1}}{\Gamma(\alpha)}$	$ce^{-\lambda t}t^{\alpha-1}E_{\alpha,\alpha}(-ct^\alpha)$	$c^{-1}\,\frac{1}{\Gamma(1-\alpha)}e^{-\lambda t}\frac{d}{dt}(t^{-\alpha} * e^{\lambda t})(t)dt$

Table 1: Some kernels K and their resolvents R and L of the second and first kind. Here $E_{\alpha,\beta}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n+\beta)}$ denotes the Mittag–Leffler function, and the constant c may be an invertible matrix.

3 Stochastic Volterra equations

Fix $d \in \mathbb{N}$ and consider the stochastic Volterra equation (1.1) for a given kernel $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$, initial condition $X_0 \in \mathbb{R}^d$, and coefficients $b\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$, where W is m-dimensional Brownian motion. The equation (1.1) can be written more compactly as

$$X = X_0 + K * (b(X)dt + \sigma(X)dW).$$

We will always require the coefficients b and σ as well as solutions of (1.1) to be continuous in order to avoid problems with the meaning of the stochastic integral term. As for stochastic (ordinary) differential equations, we call X a strong solution if it is adapted to the filtration generated by W, and a weak solution otherwise.

The following moment bound holds for any solution of (1.1) under linear growth conditions on the coefficients.

Lemma 3.1. *Assume b and σ are continuous and satisfy the linear growth condition*

$$|b(x)| \vee |\sigma(x)| \leq c_{\mathrm{LG}}(1 + |x|), \qquad x \in \mathbb{R}^d, \tag{3.1}$$

for some constant c_{LG}. Let X be a continuous solution of (1.1) with initial condition $X_0 \in \mathbb{R}^d$. Then for any $p \geq 2$ and $T < \infty$ one has

$$\sup_{t \leq T} \mathbb{E}[|X_t|^p] \leq c$$

for some constant c that only depends on $|X_0|$, $K|_{[0,T]}$, c_{LG}, p and T.

Proof. Let $\tau_n = \inf\{t \geq 0 \colon |X_t| \geq n\} \wedge T$, and observe that

$$|X_t|^p \mathbf{1}_{\{t < \tau_n\}} \leq \left| X_0 + \int_0^t K(t-s) \Big(b(X_s \mathbf{1}_{\{s < \tau_n\}}) ds + \sigma(X_s \mathbf{1}_{\{s < \tau_n\}}) dW_s \Big) \right|^p .$$

Routine application of the Jensen and BDG inequalities as well as the linear growth condition (3.1) show that the expectations $\mathbb{E}[|X_t|^p \mathbf{1}_{\{t < \tau_n\}}]$ satisfy the inequality

$$f_n \leq c' + c' |K|^2 * f_n$$

on $[0, T]$ for some constant c' that only depends on $|X_0|$, $\|K\|_{L^2(0,T)}$, c_{LG}, p and T. Consider now the scalar non-convolution Volterra kernel $K'(t,s) = |K(t-s)|^2 \mathbf{1}_{s \leq t}$. For any interval $[u, v] \subset \mathbb{R}_+$, Young's inequality implies that

$$\big\| |K'| \big\|_{L^1(u,v)} \leq \|K\|_{L^2(0, v-u)},$$

where $\| |\cdot| \|_{L^1(u,v)}$ is defined in Gripenberg et al. (1990, Definition 9.2.2); see also Lemma C.1 below for related calculations. Thus $-c'|K|^2$ is a nonpositive Volterra kernel of type L^1 on $[0, T]$ and, due to Gripenberg et al. (1990, Proposition 9.8.1), has a nonpositive resolvent R'. The Gronwall type inequality in Gripenberg et al. (1990, Lemma 9.8.2) then yields $f_n \leq c'(1 - R' * 1) \leq c'(1 - R' * 1)(T)$ on $[0, T]$. Sending n to infinity and using Fatou's lemma completes the proof. □

Remark 3.2. *It is clear from the proof that the conclusion of Lemma 3.1 holds also for state- and time-dependent predictable coefficients $b(x, t, \omega)$ and $\sigma(\omega, t, x)$, provided they satisfy a linear growth condition uniformly in (t, ω), that is,*

$$|b(x,t,\omega)| \vee |\sigma(x,t,\omega)| \leq c_{\mathrm{LG}}(1 + |x|), \qquad x \in \mathbb{R}^d,\ t \in \mathbb{R}_+,\ \omega \in \Omega,$$

for some constant c_{LG}.

The following basic existence result can be proved using techniques based on classical methods for stochastic differential equations; the proof is given in Section A.

Theorem 3.3. *Assume that K admits a resolvent of the first kind, that the components of K satisfy (2.5), and that b and σ are continuous and satisfy the linear growth condition (3.1). Then (1.1) admits a continuous weak solution for any initial condition $X_0 \in \mathbb{R}^d$.*

Remark 3.4. *At the cost of increasing the dimension, (1.1) also covers the superficially different equation $X = X_0 + K_1 * (b(X)dt) + K_2 * (\sigma(X)dW)$ where the drift and diffusion terms are convolved with different kernels K_1 and K_2. Indeed, if one defines*

$$\widetilde{K} = \begin{pmatrix} K_1 & K_2 \\ 0 & K_2 \end{pmatrix}, \qquad \widetilde{b}(x,y) = \begin{pmatrix} b(x) \\ 0 \end{pmatrix}, \qquad \widetilde{\sigma}(x,y) = \begin{pmatrix} 0 & \sigma(x) \\ 0 & 0 \end{pmatrix},$$

and obtains a solution $Z = (X, Y)$ of the equation $Z = Z_0 + \widetilde{K} * (\widetilde{b}(Z)dt + \widetilde{\sigma}(Z)d\widetilde{W})$ in $\mathbb{R}^{2d}$, where $Z_0 = (X_0, 0)$ and $\widetilde{W} = (W', W)$ is a $2d$-dimensional Brownian motion, then X is a solution of the original equation of interest. If K_1 and K_2 admit resolvents of the first kind L_1 and L_2, then

$$\widetilde{L} = \begin{pmatrix} L_1 & -L_1 \\ 0 & L_2 \end{pmatrix}$$

is a resolvent of the first kind of $\widetilde{K}$, so that Theorem 3.3 is applicable.

Our next existence result is more delicate, as it involves an assertion about stochastic invariance of the nonnegative orthant $\mathbb{R}^d_+$. This forces us to impose stronger conditions on the kernel K along with suitable boundary conditions on the coefficients b and σ. We note that any nonnegative and non-increasing kernel that is not identically zero admits a resolvent of the first kind; see Gripenberg et al. (1990, Theorem 5.5.5).

Theorem 3.5. *Assume that K is diagonal with scalar kernels K_i on the diagonal that satisfy (2.5) as well as*

$$\begin{aligned} &K_i \text{ is nonnegative, non-increasing, not identically zero, and its} \\ &\text{resolvent of the first kind } L_i \text{ is nonnegative and non-increasing} \\ &\text{in the sense that } 0 \le L_i([s, s+t]) \le L_i([0,t]) \text{ for all } s, t \ge 0. \end{aligned} \tag{3.2}$$

Assume also that b and σ are continuous and satisfy the linear growth condition (3.1) along with the boundary conditions

$$x_i = 0 \text{ implies } b_i(x) \ge 0 \text{ and } \sigma_i(x) = 0,$$

where $\sigma_i(x)$ is the ith row of $\sigma(x)$. Then (1.1) admits an $\mathbb{R}^d_+$-valued continuous weak solution for any initial condition $X_0 \in \mathbb{R}^d_+$.

Example 3.6. *If K_i is completely monotone on $(0,\infty)$ and not identically zero, then (3.2) holds due to Gripenberg et al. (1990, Theorem 5.5.4). Recall that a function f is called completely monotone on $(0,\infty)$ if it is infinitely differentiable there with $(-1)^k f^{(k)}(t) \geq 0$ for all $t > 0$ and $k = 0, 1, \ldots$. This covers, for instance, any constant positive kernel, the fractional kernel $t^{\alpha-1}$ with $\alpha \in (\frac{1}{2}, 1)$, and the exponentially decaying kernel $e^{-\beta t}$ with $\beta > 0$. Moreover, sums and products of completely monotone functions are completely monotone.*

Proof of Theorem 3.5. Define coefficients b^n and σ^n by

$$b^n(x) = b\left((x - n^{-1})^+\right), \qquad \sigma^n(x) = \sigma\left((x - n^{-1})^+\right),$$

and let X^n be the solution of (1.1) given by Theorem 3.3, with b and σ replaced by b^n and σ^n. Note that b^n and σ^n are continuous, satisfy (3.1) with a common constant, and converge to $b(x^+)$ and $\sigma(x^+)$ locally uniformly. Lemmas A.2 and A.3 therefore imply that, along a subsequence, X^n converges weakly to a solution X of the stochastic Volterra equation

$$X_t = X_0 + \int_0^t K(t-s) b(X_s^+) ds + \int_0^t K(t-s)\sigma(X_s^+) dW_s.$$

It remains to prove that X is $\mathbb{R}^d_+$-valued and hence a solution of (1.1). For this it suffices to prove that each X^n is $\mathbb{R}^d_+$-valued.

Dropping the superscript n, we are thus left with the task of proving the theorem under the stronger condition that, for some fixed $n \in \mathbb{N}$,

$$x_i \leq n^{-1} \text{ implies } b_i(x) \geq 0 \text{ and } \sigma_i(x) = 0. \tag{3.3}$$

Define $Z = \int b(X)dt + \int \sigma(X) dW$ and $\tau = \inf\{t \geq 0 \colon X_t \notin \mathbb{R}^d_+\}$. On $\{\tau < \infty\}$ we have

$$X_{\tau+h} = X_0 + (K * dZ)_{\tau+h} = X_0 + (\Delta_h K * dZ)_\tau + \int_0^h K(h-s) dZ_{\tau+s} \tag{3.4}$$

for all $h \geq 0$. We claim that

$$(\Delta_h K_i * L_i)(t) \text{ is nondecreasing in } t \text{ for any } h \geq 0. \tag{3.5}$$

Indeed, using that $K_i * L_i \equiv 1$ we have

$$\begin{aligned}(\Delta_h K_i * L_i)(t) &= \int_0^t K_i(t+h-u) L_i(du) \\ &= 1 - \int_t^{t+h} K_i(t+h-u) L_i(du) \\ &= 1 - \int_0^h K_i(h-u) L_i(t+du),\end{aligned}$$

and therefore, for any $s \leq t$,

$$(\Delta_h K_i * L_i)(t) - (\Delta_h K_i * L_i)(s) = \int_0^h K_i(h-u)\left(L_i(s+du) - L_i(t+du)\right).$$

This is nonnegative since K_i is nonnegative and L_i non-increasing, proving (3.5). Furthermore, since K_i is non-increasing and L_i nonnegative we obtain

$$0 \leq (\Delta_h K_i * L_i)(t) \leq (K_i * L_i)(t) = 1. \tag{3.6}$$

It follows from (3.5) that $\Delta_h K_i * L_i$ is of locally bounded variation. Thus (2.13) in Lemma 2.6, along with (3.5)–(3.6) and the fact that X_t is $\mathbb{R}^d_+$-valued for $t \leq \tau$, yield

$$\begin{aligned} X_{i,0} + (\Delta_h K_i * dZ_i)_\tau &= \left(1 - (\Delta_h K_i * L_i)(\tau)\right) X_{i,0} \\ &\quad + (\Delta_h K_i * L_i)(0) X_{i,\tau} \\ &\quad + (d(\Delta_h K_i * L_i) * X_i)_\tau \\ &\geq 0. \end{aligned}$$

In view of (3.4) it follows that

$$X_{i,\tau+h} \geq \int_0^h K_i(h-s)\left(b_i(X_{\tau+s})ds + \sigma_i(X_{\tau+s})dW_{\tau+s}\right) \tag{3.7}$$

on $\{\tau < \infty\}$ for all i and all $h \geq 0$.

Now, on $\{\tau < \infty\}$ there is an index i (depending on ω) such that $X_{i,\tau} = 0$ and $X_{i,\tau+h} < 0$ for arbitrarily small but positive values of h. On the other hand, by continuity there is some $\varepsilon > 0$ (again depending on ω) such that $X_{i,\tau+h} \leq n^{-1}$ for all $h \in [0, \varepsilon)$. Thus (3.3) and (3.7) yield $X_{i,\tau+h} \geq 0$ for all $h \in [0, \varepsilon)$. This contradiction shows that $\tau = \infty$, as desired. The proof is complete. □

4 Affine Volterra processes

Fix a dimension $d \in \mathbb{N}$ and a kernel $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$. Let $a\colon \mathbb{R}^d \to \mathbb{S}^d$ and $b\colon \mathbb{R}^d \to \mathbb{R}^d$ be affine maps given by

$$\begin{aligned} a(x) &= A^0 + x_1 A^1 + \cdots + x_d A^d \\ b(x) &= b^0 + x_1 b^1 + \cdots + x_d b^d \end{aligned} \tag{4.1}$$

for some $A^i \in \mathbb{S}^d$ and $b^i \in \mathbb{R}^d$, $i = 0, \ldots, d$. To simplify notation we introduce the $d \times d$ matrix

$$B = \begin{pmatrix} b^1 & \cdots & b^d \end{pmatrix},$$

and for any row vector $u \in (\mathbb{C}^d)^*$ we define the row vector

$$A(u) = (uA^1u^\top, \ldots, uA^du^\top).$$

Let E be a subset of $\mathbb{R}^d$, which will play the role of state space for the process defined below, and assume that $a(x)$ is positive semidefinite for every $x \in E$. Let $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times d}$ be continuous and satisfy $\sigma(x)\sigma(x)^\top = a(x)$ for every $x \in E$. For instance, one can take $\sigma(x) = \sqrt{\pi(a(x))}$, where π denotes the orthogonal projection onto the positive semidefinite cone, and the positive semidefinite square root is understood.

Definition 4.1. *An* affine Volterra process (with state space E) *is a continuous E-valued solution X of (1.1) with $a = \sigma\sigma^\top$ and b as in (4.1). In this paper we always take X_0 deterministic.*

Setting $K \equiv \mathrm{id}$ we recover the usual notion of an affine diffusion with state space E; see e.g. Filipović (2009). Even in this case, existence and uniqueness is often approached by first fixing a state space E of interest, and then studying conditions on (a, b) under which existence and uniqueness can be proved; see e.g. Duffie et al. (2003); Cuchiero et al. (2011); Spreij and Veerman (2012); Larsson and Krühner (2017). A key goal is then to obtain explicit parameterizations that can be used in applications. In later sections we carry out this analysis for affine Volterra processes with state space $\mathbb{R}^d$, $\mathbb{R}^d_+$, and $\mathbb{R} \times \mathbb{R}_+$. In the standard affine case more general results are available. Spreij and Veerman (2010) characterize existence and uniqueness of affine jump-diffusions on closed convex state spaces, while Abi Jaber et al. (2016) provide necessary and sufficient first order geometric conditions for existence of affine diffusions on general closed state spaces. We do not pursue such generality here for affine Volterra processes.

Assuming that an affine Volterra process is given, one can however make statements about its law. In the present section we develop general results in this direction. We start with a formula for the conditional mean.

Lemma 4.2. *Let X be an affine Volterra process. Then for all $t \le T$,*

$$\mathbb{E}[X_T \mid \mathcal{F}_t] = \left(\mathrm{id} - \int_0^T R_B(s)ds\right) X_0 + \left(\int_0^T E_B(s)ds\right) b^0 + \int_0^t E_B(T-s)\sigma(X_s)dW_s,$$

where R_B is the resolvent of $-KB$ and $E_B = K - R_B * K$. In particular,

$$\mathbb{E}[X_T] = \left(\mathrm{id} - \int_0^T R_B(s)ds\right) X_0 + \left(\int_0^T E_B(s)ds\right) b^0.$$

Proof. We have $X = X_0 + (KB) * X + K * (b^0 dt + \sigma(X)dW)$, and therefore

$$\begin{aligned} X - R_B * X = {} & \big(\mathrm{id} - R_B * 1\big) X_0 \\ & + \big((KB) - R_B * (KB)\big) * X \\ & + E_B * (b^0 dt + \sigma(X)dW). \end{aligned}$$

The resolvent equation (2.10) states that $KB - R_B * (KB) = -R_B$, so that

$$X = (\mathrm{id} - R_B * 1)\, X_0 + E_B * (b^0 dt + \sigma(X)dW).$$

Consider the local martingale $M_t = \int_0^t E_B(T-s)\sigma(X_s)dW_s$, $t \in [0,T]$. Its quadratic variation satisfies

$$\mathbb{E}[\langle M\rangle_T] \leq \int_0^T |E_B(T-s)|^2\, \mathbb{E}[|\sigma(X_s)|^2]ds \leq \|E_B\|_{L^2(0,T)} \max_{s\leq T} \mathbb{E}[|\sigma(X_s)|^2],$$

which is finite by Lemma 3.1. Thus M is a martingale, so taking $\mathcal{F}_t$-conditional expectations completes the proof. □

The main result of this section is the following theorem, which expresses the conditional Fourier–Laplace functional of an affine Volterra process in terms of the conditional mean in Lemma 4.2 and the solution of a quadratic Volterra integral equation, which we call a *Riccati–Volterra equation*.

Theorem 4.3. *Let X be an affine Volterra process and fix some $T < \infty$, $u \in (\mathbb{C}^d)^*$, and $f \in L^1([0,T],(\mathbb{C}^d)^*)$. Assume $\psi \in L^2([0,T],(\mathbb{C}^d)^*)$ solves the Riccati–Volterra equation*

$$\psi = uK + \left(f + \psi B + \frac{1}{2}A(\psi)\right) * K. \tag{4.2}$$

Then the process $\{Y_t,\, 0 \leq t \leq T\}$ defined by

$$Y_t = Y_0 + \int_0^t \psi(T-s)\sigma(X_s)dW_s - \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds, \tag{4.3}$$

$$Y_0 = uX_0 + \int_0^T \left(f(s)X_0 + \psi(s)b(X_0) + \frac{1}{2}\psi(s)a(X_0)\psi(s)^\top\right) ds \tag{4.4}$$

satisfies

$$Y_t = \mathbb{E}[uX_T + (f*X)_T \mid \mathcal{F}_t] + \frac{1}{2}\int_t^T \psi(T-s)a(\mathbb{E}[X_s \mid \mathcal{F}_t])\psi(T-s)^\top ds \tag{4.5}$$

for all $0 \leq t \leq T$. The process $\{\exp(Y_t),\, 0 \leq t \leq T\}$ is a local martingale and, if it is a true martingale, one has the exponential-affine transform formula

$$\mathbb{E}\left[\exp\left(uX_T + (f*X)_T\right) \,\middle|\, \mathcal{F}_t\right] = \exp(Y_t), \quad t \leq T. \tag{4.6}$$

Referring to (4.6) as an exponential-affine transform formula is motivated by the fact that Y_t depends affinely on the conditional expectations $\mathbb{E}[X_s \mid \mathcal{F}_t]$. We show in Lemma 4.5 below that under suitable assumptions, Y_t is actually an affine function of the past trajectory $\{X_s,\, s \leq t\}$. Before proving Theorem 4.3 we give the following lemma.

Lemma 4.4. *The Riccati–Volterra equation (4.2) is equivalent to*

$$\psi = uE_B + \left(f + \frac{1}{2}A(\psi)\right) * E_B, \tag{4.7}$$

where $E_B = K - R_B * K$ and R_B is the resolvent of $-KB$.

Proof. Assume (4.7) holds. Using the identity $E_B * (BK) = -R_B * K$ we get

$$\psi - \psi * (BK) = u(E_B + R_B * K) + \left(f + \frac{1}{2}A(\psi)\right) * (E_B + R_B * K),$$

which is (4.2). Conversely, assume (4.2) holds. With $\widetilde{R}_B$ being the resolvent of $-BK$, we obtain

$$\psi - \psi * \widetilde{R}_B = u(K - K * \widetilde{R}_B) + \left(f + \frac{1}{2}A(\psi)\right) * (K - K * \widetilde{R}_B) - \psi * \widetilde{R}_B.$$

To deduce (4.7) it suffices to prove $K * \widetilde{R}_B = R_B * K$. Equivalently, we show that for each $T < \infty$, there is some $\sigma > 0$ such that

$$(e^{-\sigma t}K) * (e^{-\sigma t}\widetilde{R}_B) = (e^{-\sigma t}R_B) * (e^{-\sigma t}K) \text{ on } [0, T], \tag{4.8}$$

where $e^{-\sigma t}$ is shorthand for the function $t \mapsto e^{-\sigma t}$. It follows from the definitions that $e^{-\sigma t}R_B$ is the resolvent of $-e^{-\sigma t}KB$, and that $e^{-\sigma t}\widetilde{R}_B$ is the resolvent of $-e^{-\sigma t}BK$; see Gripenberg et al. (1990, Lemma 2.3.3). Choosing σ large enough that $\|e^{-\sigma t}KB\|_{L^1(0,T)} < 1$ we get, as in the proof of Gripenberg et al. (1990, Theorem 2.3.1),

$$e^{-\sigma t}R_B = -\sum_{k\geq 1}(e^{-\sigma t}KB)^{*k} \quad \text{and} \quad e^{-\sigma t}\widetilde{R}_B = -\sum_{k\geq 1}(e^{-\sigma t}BK)^{*k}$$

on $[0, T]$. This readily implies (4.8), as required. □

Proof of Theorem 4.3. Let $\widetilde{Y}_t$ be defined by the right-hand side of (4.5) for $0 \leq t \leq T$. We first prove that $\widetilde{Y}_0 = Y_0$. A calculation using the identity $va(x)v^\top = vA^0v^\top + A(v)x$ yields

$$\begin{aligned}\widetilde{Y}_0 - Y_0 = {} & u\,\mathbb{E}[X_T - X_0] + (f * \mathbb{E}[X - X_0])(T) \\ & + \left(\frac{1}{2}A(\psi) * \mathbb{E}[X - X_0]\right)(T) - \left(\psi * (b^0 + BX_0)\right)(T),\end{aligned} \tag{4.9}$$

where $\mathbb{E}[X - X_0]$ denotes the function $t \mapsto \mathbb{E}[X_t - X_0]$. This function satisfies

$$\mathbb{E}[X - X_0] = K * \left(b^0 + B\,\mathbb{E}[X]\right),$$

as can be seen by taking expectations in (1.1) and using Lemma 3.1. Consequently,

$$\begin{aligned}\frac{1}{2}A(\psi) * \mathbb{E}[X - X_0] &= \frac{1}{2}A(\psi) * K * \left(b^0 + B\,\mathbb{E}[X]\right) \\ &= (\psi - uK - (f + \psi B) * K) * \left(b^0 + B\,\mathbb{E}[X]\right) \\ &= \psi * \left(b^0 + B\,\mathbb{E}[X]\right) - u\,\mathbb{E}[X - X_0] - (f + \psi B) * \mathbb{E}[X - X_0].\end{aligned}$$

Substituting this into (4.9) yields $\widetilde{Y}_0 - Y_0 = 0$, as required.

We now prove that $\widetilde{Y} = Y$. In the following calculations we let C denote an $\mathcal{F}_0$-measurable quantity that does not depend on t, and may change from line to line. Using again the identity $va(x)v^\top = vA^0v^\top + A(v)x$ we get

$$\begin{aligned}\widetilde{Y}_t &= C + u\,\mathbb{E}[X_T \mid \mathcal{F}_t] + \int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\,\mathbb{E}[X_s \mid \mathcal{F}_t]\,ds \\ &\quad - \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds.\end{aligned}$$

Lemma 4.2, the stochastic Fubini theorem, see Veraar (2012, Theorem 2.2), and a change of variables yield

$$\begin{aligned}&\int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\,\mathbb{E}[X_s \mid \mathcal{F}_t]\,ds \\ &= C + \int_0^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)\int_0^t \mathbf{1}_{\{r<s\}}E_B(s-r)\sigma(X_r)dW_r\,ds \\ &= C + \int_0^t \left(\int_r^T \left(f + \frac{1}{2}A(\psi)\right)(T-s)E_B(s-r)ds\right)\sigma(X_r)dW_r \\ &= C + \int_0^t \left(\left(f + \frac{1}{2}A(\psi)\right) * E_B\right)(T-r)\sigma(X_r)dW_r,\end{aligned}$$

where the application of the stochastic Fubini theorem in the second equality is justified by the fact that

$$\begin{aligned}&\int_0^T \left(\int_0^t \left|\left(f + \frac{1}{2}A(\psi)\right)(T-s)\,\mathbf{1}_{\{r<s\}}E_B(s-r)\sigma(X_r)\right|^2 dr\right)^{1/2} ds \\ &\leq \max_{0\leq s\leq T}|\sigma(X_s)|\,\|E_B\|_{L^2(0,T)}\|f + \frac{1}{2}A(\psi)\|_{L^1(0,T)} < \infty.\end{aligned}$$

Since $\mathbb{E}[X_T \mid \mathcal{F}_t] = C + \int_0^t E_B(T-r)\sigma(X_r)dW_r$ by Lemma 4.2, we arrive at

$$\begin{aligned}\widetilde{Y}_t &= \widetilde{Y}_0 + \int_0^t \left(uE_B + \left(f + \frac{1}{2}A(\psi)\right) * E_B\right)(T-r)\sigma(X_r)dW_r \\ &\quad - \frac{1}{2}\int_0^t \psi(T-s)a(X_s)\psi(T-s)^\top ds.\end{aligned}$$

Due to Lemma 4.4 and (4.4) we then arrive at (4.3).

The final statements are now straightforward. Indeed, (4.3) shows that $Y + \frac{1}{2}\langle Y\rangle$ is a local martingale, so that $\exp(Y)$ is a local martingale by Itô's formula. In the true martingale situation, the exponential-affine transform formula then follows upon observing that $Y_T = uX_T + (f * X)_T$ by (4.5). □

In the particular case $f \equiv 0$ and $t = 0$, Theorem 4.3 yields two different expressions for the Fourier–Laplace transform of X,

$$\mathbb{E}[e^{uX_T}] = \exp\left(\mathbb{E}[uX_T] + \frac{1}{2}\int_0^T \psi(T-t)a(\mathbb{E}[X_t])\psi(T-t)^\top dt\right) \tag{4.10}$$

$$= \exp\left(\phi(T) + \chi(T)X_0\right), \tag{4.11}$$

where ϕ and χ are defined by

$$\phi'(t) = \psi(t)b^0 + \frac{1}{2}\psi(t)A^0\psi(t)^\top, \qquad \phi(0) = 0, \tag{4.12}$$

$$\chi'(t) = \psi(t)B + \frac{1}{2}A(\psi(t)), \qquad \chi(0) = u. \tag{4.13}$$

If K admits a resolvent of the first kind L, one sees upon convolving (4.2) by L and using (2.11) that $\chi = \psi * L$; see also Example 4.7 below. Note that (4.12)–(4.13) reduce to the classical Riccati equations when $K \equiv \mathrm{id}$, since in this case $L = \delta_0 \mathrm{id}$ and hence $\psi = \chi$. While the first expression (4.10) does exist in the literature on affine diffusions in the classical case $K \equiv \mathrm{id}$, see Spreij and Veerman (2010, Proposition 4.2), the second expression (4.11) is much more common.

In the classical case one has a conditional version of (4.11), namely

$$\mathbb{E}[e^{uX_T} \mid \mathcal{F}_t] = \exp\left(\phi(T-t) + \psi(T-t)X_t\right).$$

This formulation has the advantage of showing clearly that the right-hand side depends on X_t in an exponential-affine manner. In the general Volterra case the lack of Markovianity precludes such a simple form, but using the resolvent of the first kind it is still possible to obtain an explicit expression that is exponential-affine in the past trajectory $\{X_s,\, s \le t\}$. Note that this property is not at all obvious neither from (4.6), nor from the expression

$$\mathbb{E}[e^{uX_T} \mid \mathcal{F}_t] = \mathcal{E}\Big(Y_0 + \int \psi(T-s)\sigma(X_s)dW_s\Big)_t \tag{4.14}$$

which follows directly from (4.3)–(4.4). Here $\mathcal{E}$ denotes stochastic exponential.

Lemma 4.5. *Consider the setting of Theorem 4.3, and assume in addition that K admits a resolvent of the first kind L. Define*

$$\pi_h = \Delta_h\psi * L - \Delta_h(\psi * L),$$

and assume π_h is of locally bounded variation for every $h \ge 0$. The process Y in (4.5) is then given by

$$Y_t = \phi(h) + (\Delta_h f * X)_t + (\Delta_h\psi * L)(0)X_t - \pi_h(t)X_0 + (d\pi_h * X)_t \tag{4.15}$$

with $h = T - t$ *and*

$$\phi(h) = \int_0^h \left(\psi(s) b_0 + \frac{1}{2} \psi(s) A_0 \psi(s)^\top \right) ds.$$

If instead $\operatorname{Re} \pi_h$ *is of locally bounded variation for every* $h \geq 0$*, then* $\operatorname{Re} Y_t$ *is given by* (4.15) *with* $\phi(h)$*,* f*, and* π_h *replaced by* $\operatorname{Re} \phi(h)$*,* $\operatorname{Re} f$*, and* $\operatorname{Re} \pi_h$*. An analogous statement holds for* $\operatorname{Im} Y_t$*.*

Proof. Define $Z = \int b(X)dt + \int \sigma(X)dW$. By (4.3) and the definition of Δ_h we have

$$Y_t = Y_0 + (\Delta_h \psi * dZ)_t - \int_0^t \left(\psi(T-s) b(X_s) + \frac{1}{2} \psi(T-s) a(X_s) \psi(T-s)^\top \right) ds, \quad (4.16)$$

where $h = T - t$. Note that $\psi * L = u + (f + \psi B + \frac{1}{2} A(\psi)) * \mathrm{id}$, so that $\psi * L$ is of locally bounded variation (even absolutely continuous). Thus $\Delta_h(\psi * L)$ is also of locally bounded variation, and hence so is $\Delta_h \psi * L = \pi_h + \Delta_h(\psi * L)$. Applying (2.13) in Lemma 2.6 with $F = \Delta_h \psi$ now yields

$$\Delta_h \psi * dZ = (\Delta_h \psi * L)(0) X - (\Delta_h \psi * L) X_0 + d(\Delta_h \psi * L) * X.$$

Substituting this into (4.16) and rearranging terms yields the claimed expression. □

Remark 4.6. *If* $K \equiv \mathrm{id}$ *so that* $L = \mathrm{id}\, \delta_0$*, then the correction* π_h *vanishes and the expression for* Y_t *reduces to the classical form* $\phi(T-t) + \psi(T-t) X_t$*, where* ϕ *satisfies* $\phi' = \psi b^0 + \frac{1}{2} \psi A^0 \psi^\top$ *and* $\phi(0) = 0$*.*

Example 4.7 (Fractional affine processes)**.** *Let* $K = \operatorname{diag}(K_1, \ldots, K_d)$*, where*

$$K_i(t) = \frac{t^{\alpha_i - 1}}{\Gamma(\alpha_i)}$$

for some $\alpha_i \in (\frac{1}{2}, 1]$*. Then* $L = \operatorname{diag}(L_1, \ldots, L_d)$ *with* $L_i(dt) = \frac{t^{-\alpha_i}}{\Gamma(1-\alpha_i)} dt$ *if* $\alpha_i < 1$*, and* $L_i(dt) = \delta_0(dt)$ *if* $\alpha_i = 1$*. It follows that* $\chi_i = \psi_i * L_i = I^{1-\alpha_i} \psi_i$*, where* $I^{1-\alpha_i}$ *denotes the Riemann-Liouville fractional integral operator. Hence,* (4.2) *and* (4.12) *reduce to the following system of fractional Riccati equations,*

$$\begin{aligned} \phi' &= \psi b^0 + \frac{1}{2} \psi A^0 \psi^\top, & \phi(0) &= 0, \\ D^{\alpha_i} \psi_i &= f_i + \psi b^i + \frac{1}{2} \psi A^i \psi^\top, \quad i = 1, \ldots, d, & I^{1-\alpha} \psi(0) &= u, \end{aligned}$$

where $D^{\alpha_i} = \frac{d}{dt} I^{1-\alpha_i}$ *is the Riemann-Liouville fractional differential operator. Moreover, for* $t = 0$*,* (4.6) *reads*

$$\mathbb{E}\left[e^{u X_T + (f * X)_T} \right] = \exp\left(\phi(T) + I^{1-\alpha} \psi(T) X_0 \right)$$

where we write $I^{1-\alpha} \psi = (I^{1-\alpha_1} \psi_1, \ldots, I^{1-\alpha_d} \psi_d)$*. This generalizes the expressions in El Euch and Rosenbaum (2016, 2017). Notice that the identity* $L_{\alpha_i} * K_{\alpha_i} \equiv 1$ *is equivalent to the identity* $D^{\alpha_i}(I^{\alpha_i} f) = f$*.*

5 The Volterra Ornstein–Uhlenbeck process

The particular specification of (4.1) where $A^1 = \cdots = A^d = 0$, so that $a \equiv A^0$ is a constant symmetric positive semidefinite matrix, yields an affine Volterra process with state space $E = \mathbb{R}^d$ that we call the *Volterra Ornstein–Uhlenbeck process*. It is the solution of the equation

$$X_t = X_0 + \int_0^t K(t-s)(b^0 + BX_s)ds + \int_0^t K(t-s)\sigma dW_s,$$

where $\sigma \in \mathbb{R}^{d\times d}$ is a constant matrix with $\sigma\sigma^\top = A^0$. Here existence and uniqueness is no issue. Indeed, Lemma 4.2 with $T = t$ yields the explicit formula

$$X_t = \left(\mathrm{id} - \int_0^t R_B(s)ds\right) X_0 + \left(\int_0^t E_B(s)ds\right) b^0 + \int_0^t E_B(t-s)\sigma dW_s,$$

where R_B is the resolvent of $-KB$ and $E_B = K - R_B * K$. In particular X_t is Gaussian. Furthermore, the solution of the Riccati–Volterra equation (4.2) is obtained explicitly via Lemma 4.4 as

$$\psi = uE_B + f * E_B.$$

The quadratic variation of the process Y in (4.3) is given by

$$\langle Y\rangle_t = \int_0^t \psi(T-s)\sigma\sigma^\top\psi(T-s)^\top ds,$$

and is in particular deterministic. The martingale condition in Theorem 4.3 is thus clearly satisfied, and the exponential-affine transform formula (4.6) holds for any $T < \infty$, $u \in (\mathbb{C}^d)^*$, and $f \in L^1([0,T],(\mathbb{C}^d)^*)$.

6 The Volterra square-root process

We now consider affine Volterra processes whose state space is the nonnegative orthant $E = \mathbb{R}^d_+$. We let K be diagonal with scalar kernels $K_i \in L^2_{\mathrm{loc}}(\mathbb{R}_+,\mathbb{R})$ on the diagonal. The coefficients a and b in (4.1) are chosen so that $A^0 = 0$, A^i is zero except for the (i,i) element which is equal to σ_i^2 for some $\sigma_i > 0$, and

$$b^0 \in \mathbb{R}^d_+ \text{ and } B_{ij} \geq 0 \text{ for } i \neq j. \tag{6.1}$$

The conditions on a and b are the same as in the classical situation $K \equiv \mathrm{id}$, in which case they are necessary and sufficient for (1.1) to admit an $\mathbb{R}^d_+$-valued solution for every initial condition $X_0 \in \mathbb{R}^d_+$. With this setup, we obtain an affine Volterra process that we call the *Volterra square-root process*. It is the solution of the equation

$$X_{i,t} = X_{i,0} + \int_0^t K_i(t-s)b_i(X_s)ds + \int_0^t K_i(t-s)\sigma_i\sqrt{X_{i,s}}dW_{i,s}, \quad i = 1,\dots,d. \tag{6.2}$$

The Riccati–Volterra equation (4.2) becomes

$$\psi_i(t) = u_i K_i(t) + \int_0^t K_i(t-s)\left(f_i(s) + \psi(s)b^i + \frac{\sigma_i^2}{2}\psi_i(s)^2\right) ds, \quad i = 1, \dots, d. \tag{6.3}$$

The following theorem is our main result on Volterra square-root processes.

Theorem 6.1. *Assume each K_i is continuous on $(0, \infty)$ and the shifted kernels $\Delta_h K_i$ satisfy (2.5) and (3.2) for all $h \in [0, 1]$. Assume also that (6.1) holds.*

(i) *The stochastic Volterra equation (6.2) has a unique in law $\mathbb{R}^d_+$-valued continuous weak solution X for any initial condition $X_0 \in \mathbb{R}^d_+$. For each i, the paths of X_i are Hölder continuous of any order less than $\gamma_i/2$, where γ_i is the constant associated with K_i in (2.5).*

(ii) *For any $u \in (\mathbb{C}^d)^*$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{C}^d)^*)$ such that*

$$\operatorname{Re} u_i \le 0 \text{ and } \operatorname{Re} f_i \le 0 \text{ for all } i = 1, \dots, d,$$

the Riccati–Volterra equation (6.3) has a unique global solution $\psi \in L^2_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{C}^d)^*)$, which satisfies $\operatorname{Re}\psi_i \le 0$, $i = 1, \dots, d$. Moreover, the exponential-affine transform formula (4.6) holds with Y given by (4.3)–(4.5).

Example 6.2. *A sufficient condition for K_i to satisfy the assumptions of Theorem 6.1 is that it satisfies (2.5) and is completely monotone and not identically zero; see Example 3.6. This covers in particular the gamma kernel $t^{\alpha-1}e^{-\beta t}$ with $\alpha \in (\frac{1}{2}, 1]$ and $\beta \ge 0$.*

Proof. Thanks to (6.1) and the form of $\sigma(x)$, Theorem 3.5 yields an $\mathbb{R}^d_+$-valued continuous weak solution X of (6.2) for any initial condition $X_0 \in \mathbb{R}^d_+$. The stated path regularity then follows from Lemma 2.2.

Next, the existence, uniqueness, and non-positivity statement for the Riccati–Volterra equation (6.3) is proved in Lemma 6.3 below. Thus in order to apply Theorem 4.3 to obtain the exponential-affine transform formula, it suffices to argue that $\operatorname{Re} Y_t$ is bounded above on $[0, T]$, since $\exp(Y)$ is then bounded and hence a martingale. This is done using Lemma 4.5, and we start by observing that

$$\pi^{\mathbf{r}}_{h,i}(r) = -\int_0^h \psi^{\mathbf{r}}_i(h-s) L_i(r+ds), \quad r \ge 0,$$

where $\pi_h = \Delta_h\psi * L - \Delta_h(\psi * L)$ and we write $\pi^{\mathbf{r}}_h = \operatorname{Re}\pi_h$ and $\psi^{\mathbf{r}} = \operatorname{Re}\psi$. Due to the assumption (3.2) on L_i and since $-\psi^{\mathbf{r}}_i \ge 0$, it follows that $\pi^{\mathbf{r}}_{h,i}$ is nonnegative and non-increasing. In particular, $\pi^{\mathbf{r}}_h$ is of locally bounded variation. Lemma 4.5 then yields

$$\operatorname{Re} Y_t = \operatorname{Re}\phi(h) + (\operatorname{Re}\Delta_h f * X)_t + (\Delta_h\psi^{\mathbf{r}} * L)(0)X_t - \pi^{\mathbf{r}}_h(t)X_0 + (d\pi^{\mathbf{r}}_h * X)_t$$

where $h = T - t$ and, since $A^0 = 0$,

$$\phi(h) = \int_0^h \psi(s) b^0 ds.$$

Observe that $\psi^{\mathbf{r}}$, $(\Delta_h \psi^{\mathbf{r}} * L)(0)$, $\operatorname{Re} \Delta_h f$, $-\pi_h^{\mathbf{r}}$, and $d\pi_h^{\mathbf{r}}$ all have nonpositive components. Since b^0 and X take values in $\mathbb{R}_+^d$ we thus get

$$\operatorname{Re} Y_t \leq 0.$$

Thus $\exp(Y)$ is bounded, whence Theorem 4.3 is applicable and the exponential-affine transform formula holds.

It remains to prove uniqueness in law for X. This follows since the law of X is determined by the Laplace functionals $\mathbb{E}[\exp((f*X)_T)]$ as f ranges through, say, all $(\mathbb{R}^d)^*$-valued continuous functions f with nonpositive components, and T ranges through $\mathbb{R}_+$. □

Lemma 6.3. *Assume K is as in Theorem 6.1. Let $u \in (\mathbb{C}^d)^*$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{C}^d)^*)$ satisfy*

$$\operatorname{Re} u_i \leq 0 \text{ and } \operatorname{Re} f_i \leq 0 \text{ for all } i = 1, \ldots, d.$$

Then the Riccati–Volterra equation (6.3) has a unique global solution $\psi \in L^2_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{C}^d)^*)$, and this solution satisfies $\operatorname{Re} \psi_i \leq 0$, $i = 1, \ldots, d$.

Proof. By Theorem B.1 there exists a unique non-continuable solution $(\psi, T_{\max})$ of (6.3). Let $\psi^{\mathbf{r}}$ and $\psi^{\mathbf{i}}$ denote the real and imaginary parts of ψ. They satisfy the equations

$$\begin{aligned}
\psi_i^{\mathbf{r}} &= (\operatorname{Re} u_i) K_i + K_i * \left(\operatorname{Re} f_i + \psi^{\mathbf{r}} b^i + \frac{\sigma_i^2}{2} \left((\psi_i^{\mathbf{r}})^2 - (\psi_i^{\mathbf{i}})^2 \right) \right) \\
\psi_i^{\mathbf{i}} &= (\operatorname{Im} u_i) K_i + K_i * \left(\operatorname{Im} f_i + \psi^{\mathbf{i}} b^i + \sigma_i^2 \psi_i^{\mathbf{r}} \psi_i^{\mathbf{i}} \right)
\end{aligned}$$

on $[0, T_{\max})$. Moreover, on this interval, $-\psi_i^{\mathbf{r}}$ satisfies the linear equation

$$\chi_i = -(\operatorname{Re} u_i) K_i + K_i * \left(-\operatorname{Re} f_i + \chi b^i + \frac{\sigma_i^2}{2} \left((\psi_i^{\mathbf{i}})^2 + \chi_i \psi_i^{\mathbf{r}} \right) \right).$$

Due to (6.1) and since $\operatorname{Re} u$ and $\operatorname{Re} f$ both have nonpositive components, Theorem C.2 yields $\psi_i^{\mathbf{r}} \leq 0$, $i = 1, \ldots, d$. Next, let $g \in L^2_{\mathrm{loc}}([0, T_{\max}), (\mathbb{R}^d)^*)$ and $h, \ell \in L^2_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{R}^d)^*)$ be the unique solutions of the linear equations

$$\begin{aligned}
g_i &= |\operatorname{Im} u_i| K_i + K_i * \left(|\operatorname{Im} f_i| + g b^i + \sigma_i^2 \psi_i^{\mathbf{r}} g_i \right) \\
h_i &= |\operatorname{Im} u_i| K_i + K_i * \left(|\operatorname{Im} f_i| + h b^i \right) \\
\ell_i &= (\operatorname{Re} u_i) K_i + K_i * \left(\operatorname{Re} f_i + \ell b^i - \frac{\sigma_i^2}{2} h_i^2 \right).
\end{aligned}$$

These solutions exist on $[0, T_{max})$ thanks to Corollary B.3. We now perform multiple applications of Theorem C.2. The functions $g \pm \psi^{\mathbf{i}}$ satisfy the equations

$$\chi_i = 2(\operatorname{Im} u_i)^{\pm} K_i + K_i * \left(2(\operatorname{Im} f_i)^{\pm} + \chi b^i + \sigma_i^2 \psi_i^{\mathbf{r}} \chi_i\right)$$

on $[0, T_{\max})$, so $|\psi_i^{\mathbf{i}}| \leq g_i$ on $[0, T_{\max})$ for all i. Similarly, $h - g$ satisfies the equation

$$\chi_i = K_i * \left(\chi b^i - \sigma_i^2 \psi_i^{\mathbf{r}} g_i\right)$$

on $[0, T_{\max})$, so $g_i \leq h_i$ on $[0, T_{\max})$. Finally, $\psi^{\mathbf{r}} - \ell$ satisfies the equation

$$\chi_i = K_i * \left(\chi b^i + \frac{\sigma_i^2}{2}\left((\psi_i^{\mathbf{r}})^2 + h_i^2 - (\psi_i^{\mathbf{i}})^2\right)\right),$$

on $[0, T_{\max})$, so $\ell_i \leq \psi_i^{\mathbf{r}}$ on $[0, T_{\max})$. In summary, we have shown that

$$\ell_i \leq \psi_i^{\mathbf{r}} \leq 0 \text{ and } |\psi_i^{\mathbf{i}}| \leq h_i \text{ on } [0, T_{\max}) \text{ for } i = 1, \ldots, d.$$

Since ℓ and h are global solutions and thus have finite norm on any bounded interval, this implies that $T_{\max} = \infty$ and completes the proof of the lemma. □

7 The Volterra Heston model

We now consider an affine Volterra process with state space $\mathbb{R} \times \mathbb{R}_+$, which can be viewed as a generalization of the classical Heston stochastic volatility model in finance, and which we refer to as the *Volterra Heston model*. We thus take $d = 2$ and consider the process $X = (\log S, V)$, where price process S and its variance process V are given by

$$\frac{dS_t}{S_t} = \sqrt{V_t}\left(\sqrt{1-\rho^2}\, dW_{1,s} + \rho\, dW_{2,s}\right), \qquad S_0 \in (0, \infty), \tag{7.1}$$

and

$$V_t = V_0 + \int_0^t K(t-s)\left(\kappa(\theta - V_s)ds + \sigma\sqrt{V_s}\, dW_{2,t}\right), \tag{7.2}$$

with kernel $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R})$, a standard Brownian motion $W = (W_1, W_2)$, and parameters $V_0, \kappa, \theta, \sigma \in \mathbb{R}_+$ and $\rho \in [-1, 1]$. Here the notation has been adapted to comply with established conventions in finance. Weak existence and uniqueness of V follows from Theorem 6.1 under suitable conditions on K. This in turn determines S. Moreover, observe that the log-price satisfies

$$\log S_t = \log S_0 - \int_0^t \frac{V_s}{2}\, ds + \int_0^t \sqrt{V_s}\left(\sqrt{1-\rho^2}\, dW_{1,s} + \rho\, dW_{2,s}\right).$$

Therefore the process $X = (\log S, V)$ is indeed an affine Volterra process with diagonal kernel $\operatorname{diag}(1, K)$ and coefficients a and b in (4.1) given by

$$A^0 = A^1 = 0, \quad A^2 = \begin{pmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{pmatrix},$$
$$b^0 = \begin{pmatrix} 0 \\ \kappa\theta \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -\frac{1}{2} \\ 0 & -\kappa \end{pmatrix}. \tag{7.3}$$

The Riccati–Volterra equation (4.2) takes the form

$$\psi_1 = u_1 + 1 * f_1, \tag{7.4}$$
$$\psi_2 = u_2 K + K * \left(f_2 + \frac{1}{2}\left(\psi_1^2 - \psi_1\right) - \kappa\psi_2 + \frac{1}{2}\left(\sigma^2\psi_2^2 + 2\rho\sigma\psi_1\psi_2\right)\right). \tag{7.5}$$

Theorem 7.1. *Assume K is continuous on $(0, \infty)$ and the shifted kernel $\Delta_h K$ satisfies (2.5) and (3.2) for all $h \in [0, 1]$.*

(i) *The stochastic Volterra equation (7.1)-(7.2) has a unique in law $\mathbb{R} \times \mathbb{R}_+$-valued continuous weak solution $(\log S, V)$ for any initial condition $(\log S_0, V_0) \in \mathbb{R} \times \mathbb{R}_+$. The paths of V are Hölder continuous of any order less than $\gamma/2$, where γ is the constant associated with K in (2.5).*

(ii) *Let $u \in (\mathbb{C}^2)^*$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{C}^2)^*)$ be such that*

$$\operatorname{Re} u_1 \in [0, 1],\ \operatorname{Re} u_2 \le 0,\ f_1 = 0, \text{ and } \operatorname{Re} f_2 \le 0.$$

Then, if ψ_1 is given by (7.4), the Riccati–Volterra equation-(7.5) has a unique global solution $\psi_2 \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{C}^*)$, which satisfies $\operatorname{Re}\psi_2 \le 0$. Moreover, the exponential-affine transform formula (4.6) holds with Y given by (4.3)–(4.5).

(iii) *The process S is a martingale.*

Proof. As already mentioned above, part (i) follows directly from Theorem 6.1 along with the fact that S is determined by V. Part (iii) is proved in Lemma 7.3 below. The existence, uniqueness, and non-positivity statement for the Riccati–Volterra equation (7.5) is proved in Lemma 7.4 below. Thus in order to apply Theorem 4.3 to obtain the exponential-affine transform formula, it suffices to argue that $\exp(Y)$ is a martingale. This is done using Lemma 4.5 and part (iii). As the argument closely parallels that of the proof of Theorem 6.1, we only provide an outline. We use the notation of Lemma 4.5 and Theorem 6.1, in particular π_h and $\pi_h^{\mathbf{r}} = \operatorname{Re}\pi_h$, and let L be the resolvent of the first kind of K. After verifying that $\pi_{h,2}^{\mathbf{r}}$ is of locally bounded variation, we apply Lemma 4.5 to get

$$\begin{aligned}\operatorname{Re} Y_t = {} & \operatorname{Re} u_1 \log S_t + \operatorname{Re}\phi(h) + (\Delta_h\psi_2^{\mathbf{r}} * L)(0)V_t \\ & + (\operatorname{Re}\Delta_h f_2 * V)_t - \pi_{2,h}^{\mathbf{r}}(t)V_0 + (d\pi_{2,h}^{\mathbf{r}} * V)_t\end{aligned}$$

where $h = T - t$ and

$$\phi(h) = \kappa\theta \int_0^h \psi_2(s)\, ds.$$

An inspection of signs and monotonicity properties gives

$$\operatorname{Re} Y_t \leq \operatorname{Re} u_1 \log S_t.$$

Since $\operatorname{Re} u_1 \in [0,1]$ we get

$$|\exp(Y_t)| \leq S_t^{\operatorname{Re} u_1} \leq 1 + S_t,$$

where the right-hand side is a true martingale by Part (iii), and hence of Class (DL). Thus $\exp(Y)$ is a true martingale, Theorem 4.3 is applicable, and the exponential-affine transform formula holds. □

Example 7.2 (Rough Heston model)**.** *In the fractional case $K(t) = \frac{t^{1-\alpha}}{\Gamma(\alpha)}$ with $\alpha \in (\frac{1}{2}, 1)$ we recover the rough Heston model introduced and studied by El Euch and Rosenbaum (2016, 2017). Theorem 7.1 generalizes some of their main results. For instance, in the notation of Example 4.7 and using that $L(dt) = \frac{t^{-\alpha}}{\Gamma(1-\alpha)} dt$, we have*

$$\chi = (u_1, I^{1-\alpha}\psi_2),$$

which yields the full Fourier–Laplace functional with integrated variance,

$$\mathbb{E}\left[e^{u_1 \log S_T + u_2 V_T + (f_2 * V)_T}\right] = \exp\left(\phi(T) + u_1 \log S_0 + I^{1-\alpha}\psi_2(T) V_0\right),$$

where ϕ and ψ_2 solve the fractional Riccati equations

$$\begin{aligned} \phi' &= \kappa\theta\psi_2, & \phi(0) &= 0, \\ D^\alpha \psi_2 &= f_2 + \frac{1}{2}\left(u_1^2 - u_1\right) + (2\rho\sigma u_1 - \kappa)\psi_2 + \frac{\sigma^2}{2}\psi_2^2, & I^{1-\alpha}\psi_2(0) &= u_2. \end{aligned}$$

This extends some of the main results of El Euch and Rosenbaum (2016, 2017).

We now proceed with the lemmas used in the proof of Theorem 7.1.

Lemma 7.3. *The process S is a martingale.*

Proof. Since S is a nonnegative local martingale, it is a supermartingale by Fatou's lemma, and it suffices to show that $\mathbb{E}[S_T] \geq S_0$ for any $T \in \mathbb{R}_+$. To this end, define stopping times $\tau_n = \inf\{t \geq 0 \colon V_t \geq n\}$. Then S^{τ_n} is a uniformly integrable martingale for each n, and we may define probability measures $\mathbb{Q}^n$ by

$$\frac{d\mathbb{Q}^n}{d\mathbb{P}} = \frac{S_{\tau_n}}{S_0}.$$

By Girsanov's theorem, the process $dW_t^n = dW_{2,t} - \mathbf{1}_{\{t\le\tau_n\}}\rho\sqrt{V_t}dt$ is a Brownian motion under $\mathbb{Q}^n$, and we have

$$V = V_0 + K * ((\kappa\theta - (\kappa - \rho\sigma\mathbf{1}_{[\![0,\tau_n]\!]})V)dt + \sigma\sqrt{V}dW^n).$$

Let γ be the constant appearing in (2.5) and choose p sufficiently large that $\gamma/2 - 1/p > 0$. Due to Lemma 3.1 and Remark 3.2 we have the moment bound

$$\sup_{t\le T}\mathbb{E}_{\mathbb{Q}^n}[|V_t|^p] \le c$$

for some constant c that does not depend on n. For any real-valued function f, write

$$|f|_{C^{0,\alpha}(0,T)} = \sup_{0\le s<t\le T}\frac{|f(t)-f(s)|}{|t-s|^\alpha}$$

for its α-Hölder seminorm. An application of Lemma 2.5 with $\alpha = 0$ now yields

$$\begin{aligned}
\mathbb{Q}^n(\tau_n \le T) &= \mathbb{Q}^n\Big(\sup_{t\le T}|V_t| \ge n\Big)\\
&\le \mathbb{Q}^n\Big(|V_0| + |V|_{C^{0,0}(0,T)} \ge n\Big)\\
&\le \left(\frac{1}{n-|V_0|}\right)^p \mathbb{E}_{\mathbb{Q}^n}\left[|V|^p_{C^{0,0}(0,T)}\right]\\
&\le \left(\frac{1}{n-|V_0|}\right)^p c'
\end{aligned}$$

for a constant c' that does not depend on n. We deduce that

$$\mathbb{E}_{\mathbb{P}}\left[\frac{S_T}{S_0}\right] \ge \mathbb{E}_{\mathbb{P}}\left[\frac{S_T}{S_0}\mathbf{1}_{\{\tau_n>T\}}\right] = \mathbb{Q}^n(\tau_n > T) \ge 1 - \left(\frac{1}{n-|X_0|}\right)^p c',$$

and sending n to infinity we obtain $\mathbb{E}_{\mathbb{P}}[S_T] \ge S_0$. This completes the proof. □

Lemma 7.4. *Assume K is as in Theorem 7.1. Let $u \in (\mathbb{C}^2)^*$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{C}^2)^*))$ be such that*

$$\operatorname{Re} u_1 \in [0,1],\ \operatorname{Re} u_2 \le 0,\ f_1 = 0,\ \text{and } \operatorname{Re} f_2 \le 0.$$

Then, with $\psi_1 \equiv u_1$ given by (7.4), the Riccati–Volterra equation (7.5) has a unique global solution $\psi_2 \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{C}^*)$, which satisfies $\operatorname{Re}\psi_2 \le 0$.

Proof. The proof parallels that of Lemma 6.3. For any complex number z, we denote by $z^{\mathbf{r}}$ and $z^{\mathbf{i}}$ the real and imaginary parts of z. Since $f_1 = 0$, ψ_1 is constant and equal to u_1. Equation (7.5) for ψ_2 thus reads

$$\psi_2 = u_2 K + K * \left(f_2 + \frac{1}{2}(u_1^2 - u_1) + (\rho\sigma u_1 - \kappa)\psi_2 + \frac{\sigma^2}{2}\psi_2^2\right). \tag{7.6}$$

By Theorem B.1 there exists a unique non-continuable solution $(\psi_2, T_{\max})$ of (7.6). The functions $\psi_2^{\mathbf{r}}$ and $\psi_2^{\mathbf{i}}$ satisfy the equations

$$\begin{aligned}\psi_2^{\mathbf{r}} &= u_2^{\mathbf{r}} K + K * \Big(f_2^{\mathbf{r}} + \frac{1}{2}((u_1^{\mathbf{r}})^2 - u_1^{\mathbf{r}} - (u_1^{\mathbf{i}})^2) - \rho\sigma u_1^{\mathbf{i}} \psi_2^{\mathbf{i}} \\ &\qquad - \frac{\sigma^2}{2}(\psi_2^{\mathbf{i}})^2 + (\rho\sigma u_1^{\mathbf{r}} - \kappa)\psi_2^{\mathbf{r}} + \frac{\sigma^2}{2}(\psi_2^{\mathbf{r}})^2 \Big) \\ \psi_2^{\mathbf{i}} &= u_2^{\mathbf{i}} K + K * \Big(f_2^{\mathbf{i}} + \frac{1}{2}\left(2u_1^{\mathbf{r}} u_1^{\mathbf{i}} - u_1^{\mathbf{i}}\right) + \rho\sigma u_1^{\mathbf{i}} \psi_2^{\mathbf{r}} + (\rho\sigma u_1^{\mathbf{r}} - \kappa + \sigma^2 \psi_2^{\mathbf{r}})\psi_2^{\mathbf{i}} \Big)\end{aligned}$$

on $[0, T_{\max})$. After some rewriting, we find that on $[0, T_{\max})$, $-\psi_2^{\mathbf{r}}$ satisfies the linear equation

$$\begin{aligned}\chi = -u_2^{\mathbf{r}} K + K * \Big(&- f_2^{\mathbf{r}} + \frac{1}{2}(u_1^{\mathbf{r}} - (u_1^{\mathbf{r}})^2 + (1-\rho^2)(u_1^{\mathbf{i}})^2) \\ &+ \frac{(\sigma\psi_2^{\mathbf{i}} + \rho u_1^{\mathbf{i}})^2}{2} - \left(\rho\sigma u_1^{\mathbf{r}} - \kappa + \frac{\sigma^2}{2}\psi_2^{\mathbf{r}}\right)\chi \Big).\end{aligned}$$

Due to (6.1) and since $u_1^{\mathbf{r}}, |\rho| \in [0,1]$, and $f_2^{\mathbf{r}}$ and $u_2^{\mathbf{r}}$ are nonnpositive, Theorem C.2 yields $\psi_2^{\mathbf{r}} \leq 0$ on $[0, T_{\max})$.

Now, if $\sigma = 0$, then (7.6) is a linear Volterra equation and thus admits a unique global solution $\psi_2 \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{C}^*)$ by Corollary B.3. Therefore it suffices to consider the case $\sigma > 0$. Following the proof of Lemma 6.3, we let $g \in L^2_{\mathrm{loc}}([0, T_{\max}), (\mathbb{R})^*)$ and $h, \ell \in L^2_{\mathrm{loc}}(\mathbb{R}_+, (\mathbb{R})^*)$ be the unique solutions of the linear equations

$$\begin{aligned}g &= |u_2^{\mathbf{i}}| K + \left|\rho u_1^{\mathbf{i}} \sigma^{-1}\right| + K * \Big(\left| f_2^{\mathbf{i}} + \frac{u_1^{\mathbf{i}}}{2}\left(2(1-\rho^2)u_1^{\mathbf{r}} - 1 + \frac{2\kappa\rho}{\sigma}\right)\right| \\ &\qquad + (\rho\sigma u_1^{\mathbf{r}} - \kappa + \sigma^2 \psi_2^{\mathbf{r}})g \Big) \\ h &= |u_2^{\mathbf{i}}| K + \left|\rho u_1^{\mathbf{i}} \sigma^{-1}\right| + K * \Big(\left| f_2^{\mathbf{i}} + \frac{u_1^{\mathbf{i}}}{2}\left(2(1-\rho^2)u_1^{\mathbf{r}} - 1 + \frac{2\kappa\rho}{\sigma}\right)\right| + (\rho\sigma u_1^{\mathbf{r}} - \kappa)h \Big) \\ \ell &= u_2^{\mathbf{r}} K + K * \Big(f_2^{\mathbf{r}} + \frac{1}{2}((u_1^{\mathbf{r}})^2 - u_1^{\mathbf{r}} - (u_1^{\mathbf{i}})^2) - |\rho\sigma u_1^{\mathbf{i}}| \left(h + \left|\rho u_1^{\mathbf{i}} \sigma^{-1}\right|\right) \\ &\qquad - \frac{\sigma^2}{2}\left(h + \left|\rho u_1^{\mathbf{i}} \sigma^{-1}\right|\right)^2 + (\rho\sigma u_1^{\mathbf{r}} - \kappa)\ell \Big).\end{aligned}$$

These solutions exist on $[0, T_{\max})$ thanks to Corollary B.3. We now perform multiple applications of Theorem C.2. The functions $g \pm (\psi_2^{\mathbf{i}} + (\rho u_1^{\mathbf{i}} \sigma^{-1}))$ satisfy the equations

$$\begin{aligned}\chi = (u_2^{\mathbf{i}})^{\pm} K + \left(\rho u_1^{\mathbf{i}} \sigma^{-1}\right)^{\pm} + K * \Big(&\left(f_2^{\mathbf{i}} + \frac{u_1^{\mathbf{i}}}{2}\left(2(1-\rho^2)u_1^{\mathbf{r}} - 1 + \frac{2\kappa\rho}{\sigma}\right)\right)^{\pm} \\ &+ (\rho\sigma u_1^{\mathbf{r}} - \kappa + \sigma^2 \psi_2^{\mathbf{r}})\chi \Big)\end{aligned}$$

on $[0, T_{max})$, so that $0 \le |\psi_2^{\mathbf{i}} + (\rho u_1^{\mathbf{i}} \sigma^{-1})| \le g$ on $[0, T_{max})$. Similarly, the function $h - g$ satisfies the equation

$$\chi = K * \left(-\sigma^2 \psi_2^{\mathbf{r}} g + (\rho\sigma u_1^{\mathbf{r}} - \kappa)\chi\right)$$

on $[0, T_{\max})$, so that $g \le h$ on $[0, T_{\max})$. This yields $|\psi_2^{\mathbf{i}}| \le h + \left|\rho u_1^{\mathbf{i}} \sigma^{-1}\right|$ on $[0, T_{max})$. Finally, the function $\psi_2^{\mathbf{r}} - \ell$ satisfies the linear equation

$$\begin{aligned}\chi = K * \Bigg(& |\rho\sigma u_1^{\mathbf{i}}| \left(h + \left|\rho u_1^{\mathbf{i}} \sigma^{-1}\right|\right) - \rho\sigma u_1^{\mathbf{i}} \psi_2^{\mathbf{i}} \\ & + \frac{\sigma^2}{2}\left(\left(h + \left|\rho u_1^{\mathbf{i}} \sigma^{-1}\right|\right)^2 - \left(\psi_2^{\mathbf{i}}\right)^2 + (\psi_2^{\mathbf{r}})^2\right) + (\rho\sigma u_1^{\mathbf{r}} - \kappa)\chi\Bigg)\end{aligned}$$

on $[0, T_{\max})$, so that $\ell \le \psi_2^{\mathbf{r}} \le 0$ on $[0, T_{\max})$. Since h and ℓ are global solutions and thus have finite norm on any bounded interval, this implies that $T_{\max} = \infty$ and completes the proof of the lemma. □

We conclude this section with a remark on an alternative variant of the Volterra Heston model in the spirit of Guennoun et al. (2017).

Example 7.5. *Let $\widetilde{K}$ denotes a scalar locally square integrable non-negative kernel. Consider the following variant of the Volterra Heston model*

$$\begin{aligned} dS_t &= S_t \sqrt{\widetilde{V}_t} dB_t, & S_0 &= 1, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dB_t^{\perp}, & V_0 &\ge 0, \\ \widetilde{V}_t &= \widetilde{V}_0 + (\widetilde{K} * V)_t, \end{aligned}$$

where B and $B^{\perp}$ are independent Brownian motions. Since $\widetilde{K}$ is nonnegative, one readily sees that there exists a unique strong solution taking values in $\mathbb{R} \times \mathbb{R}_+^2$. The 3-dimensional process $X = (\log S, V, \widetilde{V})$ is an affine Volterra process with

$$K = \operatorname{diag}(1, 1, \widetilde{K}), \quad b^0 = \begin{pmatrix} 0 \\ \kappa\theta \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 0 & -\frac{1}{2} \\ 0 & -\kappa & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$A^0 = 0, \quad A^1 = 0, \quad A^2 = \operatorname{diag}(0, \sigma^2, 0), \quad A^3 = \operatorname{diag}(1, 0, 0).$$

The Riccati–Volterra equation (4.2) reads

$$\begin{aligned} \psi_1' &= f_1, & \psi_1(0) &= u_1 \\ \psi_2' &= f_2 + \psi_3 - \kappa\psi_2 + \frac{\sigma^2}{2}\psi_2^2, & \psi_2(0) &= u_2, \qquad (7.7) \\ \psi_3 &= u_3 \widetilde{K} + \widetilde{K} * \left(f_3 + \frac{1}{2}\psi_1(\psi_1 - 1)\right). \qquad (7.8) \end{aligned}$$

Under suitable conditions the solution exists and is unique, and the process e^Y with Y given by (4.3)–(4.5) is a true martingale. Hence by Theorem 4.3 the exponential-affine transform formula (4.6) holds. We omit the details. In particular, for $f \equiv 0$ we get, using Example 4.7,

$$\chi(t) = (\psi * L)(t) = \left(u_1, \psi_2(t), u_3 + \frac{(u_1^2 - u_1)t}{2} \right)$$

and

$$\mathbb{E}\left[e^{u_1 \log S_T + u_2 V_T + u_3 \widetilde{V}_T} \right] = \exp\left(\phi(T) + u_1 \log S_0 + \psi_2(T) V_0 + \left(u_3 + \frac{(u_1^2 - u_1)T}{2} \right) \widetilde{V}_0 \right),$$

where ϕ and ψ_2 solve

$$\begin{aligned} \phi' &= \kappa\theta\psi_2, & \phi(0) &= 0, \\ \psi_2' &= u_3 \widetilde{K} + \widetilde{K} * \frac{(u_1^2 - u_1)}{2} - \kappa\psi_2 + \frac{\sigma^2}{2}\psi_2^2, & \psi_2(0) &= u_2. \end{aligned}$$

Setting $\widetilde{K} = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$ and $u_2 = 0$, this formula agrees with Guennoun et al. (2017, Theorem 2.1). If B and $B^\perp$ are correlated one loses the affine property, as highlighted in Guennoun et al. (2017, Remark 2.2).

A Proof of Theorem 3.3

Lemma A.1. *Assume b and σ are Lipschitz continuous. Then (1.1) admits a unique continuous strong solution X for any initial condition $X_0 \in \mathbb{R}^d$.*

Proof. The proof parallels that of Mytnik and Salisbury (2015, Proposition 2.1), which uses a standard Picard iteration scheme; we refrain from repeating the argument here. Let us only mention that the proof hinges on the extended Gronwall's lemma given in Dalang (1999, Lemma 15), as well as routine applications of the BDG inequalities and Jensen's inequality. □

Lemma A.2. *Fix an initial condition $X_0 \in \mathbb{R}^d$ and a constant c_{LG}. Let $\mathcal{X}$ denote the set of all continuous processes X that solve (1.1) for some continuous coefficients b and σ satisfying the linear growth bound (3.1) with the given constant c_{LG}. Then $\mathcal{X}$ is tight, meaning that the family $\{\text{law of } X \colon X \in \mathcal{X}\}$ of laws on $C(\mathbb{R}_+; \mathbb{R}^d)$ is tight.*

Proof. Let $X \in \mathcal{X}$ be any solution of (1.1) for some continuous b and σ satisfying the linear growth bound (3.1). Lemma 3.1 implies that $\sup_{u \le T} \mathbb{E}[|b(X_u)|^p]$ and $\sup_{u \le T} \mathbb{E}[|\sigma(X_u)|^p]$

are bounded above by a constant that only depends on $|X_0|$, $\|K\|_{L^2(0,T)}$, c_{LG}, p, and T. Therefore, since the components of K satisfy (2.5), we may apply Lemma 2.5 to obtain

$$\mathbb{E}\left[\left(\sup_{0\le s<t\le T}\frac{|X_t - X_s|}{|t-s|^\alpha}\right)^p\right] \le c$$

for all $\alpha \in [0, \bar{\gamma}/2 - 1/p)$, where $\bar{\gamma}$ is the smallest of the constants γ appearing in (2.5) for the components of K, and where c is a constant that only depends on $|X_0|$, $\|K\|_{L^2(0,T)}$, c_{LG}, p, and T, but not on s or t, nor on the specific choice of $X \in \mathcal{X}$. Choosing p so that $\bar{\gamma}p/2 > 1$, and using that closed Hölder balls are compact in $C(\mathbb{R}_+;\mathbb{R}^d)$, it follows that $\mathcal{X}$ is tight. □

Lemma A.3. *Assume that K admits a resolvent of the first kind L. For each $n \in \mathbb{N}$, let X^n be a weak solution of (1.1) with b and σ replaced by some continuous coefficients b^n and σ^n that satisfy (3.1) with a common constant c_{LG}. Assume that $b^n \to b$ and $\sigma^n \to \sigma$ locally uniformly for some coefficients b and σ, and that $X^n \Rightarrow X$ for some continuous process X. Then X is a weak solution of (1.1).*

Proof. Lemma 2.6 yields the identity

$$L * (X^n - X_0) = \int b^n(X^n)dt + \int \sigma^n(X^n)dW.$$

Moreover, Gripenberg et al. (1990, Theorem 3.6.1(ii) and Corollary 3.6.2(iii)) imply that the map

$$F \mapsto L * (F - F(0))$$

is continuous from $C(\mathbb{R}_+;\mathbb{R}^d)$ to itself. Using also the locally uniform convergence of b^n and σ^n, the continuous mapping theorem shows that the martingales

$$M^n = \int \sigma^n(X^n)dW = L * (X^n - X_0) - \int b^n(X^n)dt$$

converge weakly to some limit M, that the quadratic variations $\langle M^n\rangle = \int \sigma^n\sigma^{n\top}(X^n)dt$ converge weakly to $\int \sigma\sigma^\top(X)dt$, and that $\int b^n(X^n)dt$ converge weakly to $\int b(X)dt$.

Consider any $s < t$, $m \in \mathbb{N}$, any bounded continuous function $f\colon \mathbb{R}^m \to \mathbb{R}$, and any $0 \le t_1 \le \cdots \le t_m \le s$. Observe that the moment bound in Lemma 3.1 is uniform in n since the X^n satisfy the linear growth condition (3.1) with a common constant. Using Billingsley (1999, Theorem 3.5), one then readily shows that

$$\mathbb{E}[f(X_{t_1},\dots,X_{t_m})(M_t - M_s)] = \lim_{n\to\infty}\mathbb{E}[f(X^n_{t_1},\dots,X^n_{t_m})(M^n_t - M^n_s)] = 0,$$

and similarly for the increments of $M^n_i M^n_j - \langle M^n_i, M^n_j\rangle$. It follows that M is a martingale with respect to the filtration generated by X with quadratic variation $\langle M\rangle = \int \sigma\sigma^\top(X)dt$.

This carries over to the usual augmentation. Enlarging the probability space if necessary, we may now construct a d-dimensional Brownian motion $\overline{W}$ such that $M = \int \sigma(X) d\overline{W}$.

The above shows that $L * (X - X_0) = \int b(X)dt + \int \sigma(X)d\overline{W}$. The converse direction of Lemma 2.6 then yields $X = X_0 + K * (b(X)dt + \sigma(X)d\overline{W})$, that is, X solves (1.1) with the Brownian motion $\overline{W}$. □

Proof of Theorem 3.3. Using Hofmanová and Seidler (2012, Proposition 1.1) we choose Lipschitz coefficients b^n and σ^n that satisfy the linear growth bound (3.1) with c_{LG} replaced by $2c_{\mathrm{LG}}$, and converge locally uniformly to b and σ as $n \to \infty$. Let X^n be the unique continuous strong solution of (1.1) with b and σ replaced by b^n and σ^n; see Lemma A.1. Due to Lemma A.2 the sequence $\{X^n\}$ is tight, so after passing to a subsequence we have $X^n \Rightarrow X$ for some continuous process X. The result now follows from Lemma A.3. □

B Local solutions of Volterra integral equations

Fix a kernel $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$ along with functions $g\colon \mathbb{R}_+ \to \mathbb{C}^d$ and $p\colon \mathbb{R}_+ \times \mathbb{C}^d \to \mathbb{C}^d$, and consider the Volterra integral equation

$$\psi = g + K * p(\,\cdot\,, \psi). \tag{B.1}$$

A *non-continuable solution* of (B.1) is a pair $(\psi, T_{\max})$ with $T_{\max} \in (0, \infty]$ and $\psi \in L^2_{\mathrm{loc}}([0, T_{\max}), \mathbb{C}^d)$, such that ψ satisfies (B.1) on $[0, T_{\max})$ and $\|\psi\|_{L^2(0,T_{\max})} = \infty$ if $T_{\max} < \infty$. If $T_{\max} = \infty$ we call ψ a *global solution* of (B.1). With some abuse of terminology we call a non-continuable solution $(\psi, T_{\max})$ *unique* if for any $T \in \mathbb{R}_+$ and $\widetilde{\psi} \in L^2([0, T], \mathbb{C}^d)$ satisfying (B.1) on $[0, T]$, we have $T < T_{\max}$ and $\widetilde{\psi} = \psi$ on $[0, T]$.

Theorem B.1. *Assume that $g \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{C}^d)$, $p(\,\cdot\,, 0) \in L^1_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{C}^d)$, and that for all $T \in \mathbb{R}_+$ there exist a positive constant Θ_T and a function $\Pi_T \in L^2([0, T], \mathbb{R}_+)$ such that*

$$|p(t, x) - p(t, y)| \le \Pi_T(t)|x - y| + \Theta_T|x - y|(|x| + |y|), \qquad x, y \in \mathbb{C}^d,\ t \le T. \tag{B.2}$$

The Volterra integral equation (B.1) has a unique non-continuable solution $(\psi, T_{\max})$. If g and p are real-valued, then so is ψ.

Remark B.2. *If $K \in L^{2+\varepsilon}_{\mathrm{loc}}$ for some $\varepsilon > 0$, then it is possible to apply Gripenberg et al. (1990, Theorem 12.4.4) with $p = 2 + \varepsilon$ to get existence.*

Proof. We focus on the complex-valued case; for the real-valued case, simply replace $\mathbb{C}^d$ by $\mathbb{R}^d$ below. We first prove that a solution exists for small times. Let $\rho \in (0, 1]$ and $\varepsilon > 0$ be constants to be specified later, and define

$$B_{\rho,\varepsilon} = \{\psi \in L^2([0, \rho], \mathbb{C}^d)\colon \|\psi\|_{L^2(0,\rho)} \le \varepsilon\}.$$

Consider the map F acting on elements $\psi \in B_{\rho,\varepsilon}$ by

$$F(\psi) = g + K * p(\,\cdot\,, \psi).$$

We write $\|\cdot\|_q = \|\cdot\|_{L^q(0,\rho)}$ for brevity in the following computations. The growth condition (B.2) along with the Young, Cauchy–Schwarz, and triangle inequalities yield for $\psi, \widetilde{\psi} \in B_{\rho,\varepsilon}$

$$\begin{aligned} \|F(\psi)\|_2 &\le \|g\|_2 + \|K\|_2 \|p(\,\cdot\,, \psi)\|_1 \\ &\le \|g\|_2 + \|K\|_2 \left(\|p(\,\cdot\,, 0)\|_1 + \|\Pi_1\|_2 \|\psi\|_2 + \Theta_1 \|\psi\|_2^2 \right) \\ &\le \|g\|_2 + \|K\|_2 \left(\|p(\,\cdot\,, 0)\|_1 + \|\Pi_1\|_{L^2(0,1)} \varepsilon + \Theta_1 \varepsilon^2 \right) \end{aligned} \tag{B.3}$$

and

$$\begin{aligned} \|F(\psi) - F(\widetilde{\psi})\|_2 &\le \|K\|_2 \left(\|\Pi_1\|_2 + \Theta_1 \left(\|\psi\|_2 + \|\widetilde{\psi}\|_2 \right) \right) \|\psi - \widetilde{\psi}\|_2 \\ &\le \|K\|_2 \left(\|\Pi_1\|_{L^2(0,1)} + 2\Theta_1 \varepsilon \right) \|\psi - \widetilde{\psi}\|_2. \end{aligned}$$

Choose $\varepsilon > 0$ so that $1 + \frac{\varepsilon}{2} + \|\Pi_1\|_{L^2(0,1)} \varepsilon + \Theta_1 \varepsilon^2 < 2$ and $\varepsilon \left(\|\Pi_1\|_{L^2(0,1)} + 2\Theta_1 \varepsilon \right) < 2$. Then choose $\rho > 0$ so that $\|g\|_2 \vee \|K\|_2 \vee \|p(\,\cdot\,, 0)\|_1 \le \varepsilon/2$. This yields

$$\|F(\psi)\|_2 \le \frac{\varepsilon}{2} \left(1 + \frac{\varepsilon}{2} + \|\Pi_1\|_{L^2(0,1)} \varepsilon + \Theta_1 \varepsilon^2 \right) \le \varepsilon$$

and

$$\|F(\psi) - F(\widetilde{\psi})\|_2 \le \kappa \|\psi - \widetilde{\psi}\|_2, \qquad \kappa = \frac{\varepsilon}{2} \left(\|\Pi_1\|_{L^2(0,1)} + 2\Theta_1 \varepsilon \right) < 1.$$

Thus F maps $B_{\rho,\varepsilon}$ to itself and is a contraction there, so Banach's fixed point theorem implies that F has a unique fixed point $\psi \in B_{\rho,\varepsilon}$, which is a solution of (B.1).

We now extend this to a unique non-continuable solution of (B.1). Define the set

$$J = \{T \in \mathbb{R}_+ \colon \text{(B.1) has a solution } \psi \in L^2([0,T], \mathbb{C}^d) \text{ on } [0,T]\}.$$

Then $0 \in J$, and if $T \in J$ and $0 \le S \le T$, then $S \in J$. Thus J is a nonempty interval. Moreover, J is open in $\mathbb{R}_+$. Indeed, pick $T \in J$, let ψ be a solution on $[0,T]$, and set

$$h(t) = g(T + t) + \int_0^T K(T + t - s) p(s, \psi(s)) ds, \quad t \ge 0,$$

which lies in $L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{C}^d)$ by a calculation similar to (B.3). By what we already proved, the equation

$$\chi = h + K * p(\,\cdot\, + T, \chi)$$

admits a solution $\chi \in L^2([0,\rho], \mathbb{C}^d)$ on $[0,\rho]$ for some $\rho > 0$. Defining $\psi(t) = \chi(t - T)$ for $t \in (T, T + \rho]$, one verifies that ψ solves (B.1) on $[0, T + \rho]$. Thus $T + \rho \in J$, so J is open

in $\mathbb{R}_+$ and hence of the form $J = [0, T_{\max})$ for some $0 < T_{\max} \le \infty$ with $T_{\max} \notin J$. This yields a non-continuable solution $(\psi, T_{\max})$.

It remains to argue uniqueness. Pick $T \in \mathbb{R}_+$ and $\widetilde{\psi} \in L^2([0,T], \mathbb{C}^d)$ satisfying (B.1) on $[0,T]$. Then $T \in J$, so $T < T_{\max}$. Let S be the supremum of all $S' \le T$ such that $\widetilde{\psi} = \psi$ on $[0, S']$. Then $\widetilde{\psi} = \psi$ on $[0, S]$ (almost everywhere, as elements of L^2). If $S < T$, then for $\rho > 0$ sufficiently small we have $0 < \|\psi - \widetilde{\psi}\|_{L^2(0,S+\rho)} \le \frac{1}{2}\|\psi - \widetilde{\psi}\|_{L^2(0,S+\rho)}$, a contradiction. Thus $S = T$, and uniqueness is proved. □

Corollary B.3. *Let $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$, $u \in \mathbb{R}^d$, $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$ and $G \in L^2(\mathbb{R}_+, \mathbb{R}^{d\times d})$. Suppose that $p\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ is a Lispschitz continuous function in the second argument such that $p(\cdot, 0) \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$. Then the equation*

$$\chi = Ku + K * (f + Gp(\cdot, \chi)) \tag{B.4}$$

has a unique global solution $\psi \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$. Moreover, if K is continuous on $[0,\infty)$ then ψ is also continuous on $[0,\infty)$ and $\psi(0) = uK(0)$.

Proof. Theorem B.1 implies the existence and uniqueness of a non-continuable solution $(\chi, T_{\max})$. If K is continuous on $[0,\infty)$, then this solution is continuous on $[0, T_{\max})$ with $\chi(0) = K(0)u$. To prove that $T_{\max} = \infty$, observe that

$$|\chi| \le |K||u| + |K| * (|f| + |G|(|p(\cdot, 0)| + \Theta|\chi|)) \tag{B.5}$$

for some positive constant Θ. Define the scalar non-convolution Volterra kernel $K'(t,s) = \Theta|K(t-s)||G(s)|\mathbf{1}_{s\le t}$. For any interval $[u,v] \subset \mathbb{R}_+$, Young's inequality implies that

$$\big\|\!\big|K'\big|\!\big\|_{L^2(u,v)} \le \Theta\|K\|_{L^2(0,v-u)}\|G\|_{L^2(0,v-u)},$$

where $\|\!|\cdot|\!\|_{L^1(u,v)}$ is defined in Gripenberg et al. (1990, Definition 9.2.2); see also Lemma C.1 below for related calculations. Thus $-K'$ is a nonpositive kernel of type L^2_{loc} and, due to Gripenberg et al. (1990, Proposition 9.8.1), has a nonpositive resolvent R' of type L^2_{loc}. The Gronwall type inequality in Gripenberg et al. (1990, Lemma 9.8.2) and (B.5) then yield

$$|\chi| \le f' - \int_0^{\cdot} R'(\cdot, s) f'(s)\, ds \tag{B.6}$$

on $[0, T_{\max}]$, where

$$f' = |K||u| + |K| * (|f| + |G|(|p(\cdot, 0)|).$$

Since the function on the right-hand side of (B.6) is in $L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R})$ due to Gripenberg et al. (1990, Theorem 9.3.6), we conclude that $T_{\max} = \infty$. □

C Invariance results for Volterra integral equations

Lemma C.1. *Fix $T < \infty$. Let $u \in \mathbb{C}^d$, $G \in L^2([0,T], \mathbb{C}^{d\times d})$, as well as $F^n \in L^2([0,T], \mathbb{C}^d)$ and $K^n \in L^2([0,T], \mathbb{C}^{d\times d})$ for $n = 0, 1, 2, \ldots$. For each n, there exists a unique element $\chi^n \in L^2([0,T], \mathbb{C}^{d\times d})$ such that*

$$\chi^n = F^n + K^n * (G\chi^n).$$

Moreover, if $F^n \to F^0$ and $K^n \to K^0$ in $L^2(0,T)$, then $\chi^n \to \chi^0$ in $L^2(0,T)$.

Proof. For any $K \in L^2([0,T], \mathbb{C}^{d\times d})$, define $\widehat{K}(t,s) = K(t-s)G(s)\mathbf{1}_{s\le t}$. This is a Volterra kernel in the sense of Gripenberg et al. (1990, Definition 9.2.1), and we have

$$\int_0^T \int_0^T \mathbf{1}_{s\le t} |K(t-s)|^2 |G(s)|^2 ds\, dt \le \| K\|^2_{L^2(0,T)} \| G\|^2_{L^2(0,T)}$$

by Young's inequality. Thus by Gripenberg et al. (1990, Proposition 9.2.7(iii)), $\widehat{K}$ is of type L^2 on $(0,T)$ with norm

$$|||\widehat{K}|||_{L^2(0,T)} \le \| K\|_{L^2(0,T)} \| G\|_{L^2(0,T)}, \tag{C.1}$$

see Gripenberg et al. (1990, Definition 9.2.2). Next, we show that $-\widehat{K}$ admits a resolvent of type L^2 on $(0,T)$ in the sense of Gripenberg et al. (1990, Definition 9.3.1). For any $0 \le u \le v \le T$, a change of variables and Young's inequality give

$$\begin{aligned}\int_u^v \int_u^v \mathbf{1}_{s\le t} |K(t-s)|^2 |G(s)|^2 ds\, dt &= \int_0^{v-u} \int_0^t |K(t-s)|^2 |G(u+s)|^2 ds\, dt \\ &\le \| K\|^2_{L^2(0,v-u)} \| G\|^2_{L^2(0,T)}.\end{aligned}$$

For $v-u$ sufficiently small, the right-hand side is smaller than 1, whence $|||\widehat{K}|||_{L^2(u,v)} < 1$ by Gripenberg et al. (1990, Proposition 9.2.7(iii)). We now apply Gripenberg et al. (1990, Corollary 9.3.14) to obtain a resolvent of type L^2 on $(0,T)$ of $-\widehat{K}$, which we denote by $\widehat{R}$. Due to Gripenberg et al. (1990, Theorem 9.3.6), the unique solution in $L^2(0,T)$ of the equation

$$\chi(t) = F(t) + \int_0^t \widehat{K}(t,s)\chi(s) ds, \quad t \in [0,T],$$

for a given $F \in L^2([0,T], \mathbb{C}^d)$, is

$$\chi(t) = F(t) - \int_0^t \widehat{R}(t,s) F(s) ds, \quad t \in [0,T].$$

This proves the existence and uniqueness statement for the χ^n. Next, assume $F^n \to F^0$ and $K^n \to K^0$ in $L^2(0,T)$. Applying (C.1) with $K = K^n - K^0$ shows that $\widehat{K}^n \to \widehat{K}^0$ with respect to the norm $|||\cdot|||_{L^2(0,T)}$. An application of Gripenberg et al. (1990, Corollary 9.3.12) now shows that $\chi^n \to \chi^0$ in $L^2(0,T)$ as claimed. □

Theorem C.2. *Assume $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$ is diagonal with scalar kernels K_i on the diagonal. Assume each K_i is continuous on $(0,\infty)$ and the shifted kernels $\Delta_h K_i$ satisfy (2.5) and (3.2) for all $h \in [0,1]$. Let $u \in \mathbb{R}^d$, $F \in L^1_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$ and $G \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{d\times d})$ be such that $u_i \geq 0$, $F_i \geq 0$, and $G_{ij} \geq 0$ for all $i, j = 1, \dots, d$ and $i \neq j$. Then the linear Volterra equation*

$$\chi = Ku + K * (F + G\chi) \tag{C.2}$$

has a unique solution $\chi \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$ with $\chi_i \geq 0$ on $[0,T]$ for $i = 1, \dots, d$.

Proof. Define kernels $K^n = K(\cdot + n^{-1})$ for $n \in \mathbb{N}$, which are diagonal with scalar kernels on the diagonal that satisfy (2.5) and (3.2). Lemma C.1 shows that (C.2) (respectively (C.2) with K replaced by K^n) has a unique solution χ (respectively χ^n), and that $\chi^n \to \chi$ in $L^2(\mathbb{R}_+, \mathbb{R}^d)$. Therefore, we can suppose without loss of generality that K is continuous on $[0,\infty)$ with $K_i(0) \geq 0$. To shows that χ takes values in $\mathbb{R}^d_+$, it is therefore enough to consider the case where K is continuous on $[0,\infty)$ with $K_i(0) \geq 0$ for all i.

For $x \in \mathbb{R}^d$ define $b(x) = F + Gx$. For all positive n, Corollary B.3 implies that there exists a unique solution $\chi^n \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^d)$ of the equation

$$\chi^n = Ku + K * b((\chi^n - n^{-1})^+),$$

and that χ^n is continuous on $[0,\infty)$ with $\chi^n_i(0) = K_i(0)u_i \geq 0$ for $i = 1, \dots, d$. We claim that χ^n is $\mathbb{R}^d_+$ valued for all n. Indeed, arguing as in the proof of Theorem 3.5, we can show that if L_i denotes the resolvent of the first kind of K_i, then $(\Delta_h K_i * L_i)(t)$ is nonnegative, bounded and nondecreasing in t for any $h \geq 0$. Fix n and define $Z = \int b((\chi^n - n^{-1})^+ \, dt$. The argument of Lemma 2.6 shows that for all $h \geq 0$ and $i = 1, \dots, d$,

$$\begin{aligned} \Delta_h K_i * dZ_i &= (\Delta_h K_i * L_i)(0) K_i * dZ_i + d(\Delta_h K_i * L_i) * K_i * dZ_i \\ &= (\Delta_h K_i * L_i)(0)\chi^n_i + d(\Delta_h K_i * L_i) * \chi^n_i \\ &\quad - u_i \left((\Delta_h K_i * L_i)(0) K_i + d(\Delta_h K_i * L_i) * K_i\right). \end{aligned} \tag{C.3}$$

Convolving the quantity $d(\Delta_h K_i * L_i) * K_i$ first by L_i, then by K_i, and comparing densities of the resulting absolutely continuous functions, we deduce that

$$d(\Delta_h K_i * L_i) * K_i = \Delta_h K_i - (\Delta_h K_i * L_i)(0) K_i.$$

Plugging this identity into (C.3) yields

$$\Delta_h K_i * dZ_i = (\Delta_h K_i * L_i)(0)\chi^n_i + d(\Delta_h K_i * L_i) * \chi^n_i - u_i \Delta_h K_i. \tag{C.4}$$

Define $\tau = \inf\{t \geq 0 \colon X_t \notin \mathbb{R}^d_+\}$ and assume for contradiction that $\tau < \infty$. Then

$$\chi^n(\tau + h) = \Delta_h K(\tau)u + (K * dZ)_{\tau+h} = \Delta_h K(\tau)u + (\Delta_h K * dZ)_\tau + \int_0^h K(h-s) dZ_{\tau+s} \tag{C.5}$$

for any $h \geq 0$. By definition of τ, the identities (C.4) and (C.5) imply

$$\chi_i^n(\tau + h) \geq \int_0^h K_i(h-s) b_i((\chi^n(\tau+s) - n^{-1})^+)\, ds \quad i = 1, \dots, d.$$

As in the proof of Theorem 3.5, these inequalities lead to a contradiction. Hence $\tau = \infty$ and χ^n is $\mathbb{R}^d_+$-valued for all n.

To conclude that χ is $\mathbb{R}^d_+$-valued it suffices to prove that χ^n converges to χ in $L^2([0,T], \mathbb{R}^d)$ for all $T \in \mathbb{R}_+$. To this end we write

$$\chi - \chi^n = K * \left(G(\chi^n - (\chi^n - n^{-1})^+) + G(\chi - \chi^n) \right),$$

from which we infer

$$|\chi - \chi^n| \leq \frac{\sqrt{d}}{n} |K| * |G| + |K| * (|G||\chi - \chi^n|).$$

The same argument as in the proof of Corollary B.3 shows that

$$|\chi - \chi^n| \leq \frac{\sqrt{d}}{n} \left(F' - \int_0^{\cdot} R'(\,\cdot\,, s) F'(s)\, ds \right), \tag{C.6}$$

where R' is the nonpositive resolvent of type L^2_{loc} of $K'(t,s) = |K(t-s)||G(s)|$, and $F' = |K| * |G|$. Since the right-hand side of (C.6) is in $L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R})$ in view of Gripenberg et al. (1990, Theorem 9.3.6), we conclude that χ^n converges to χ in $L^2([0,T], \mathbb{R}^d)$ for all $T \in \mathbb{R}_+$. □

References

Eduardo Abi Jaber, Bruno Bouchard, and Camille Illand. Stochastic invariance of closed sets with non-Lipschitz coefficients. *arXiv preprint arXiv:1607.08717.*, 2016.

Christian Bayer, Peter Friz, and Jim Gatheral. Pricing under rough volatility. *Quantitative Finance*, 16(6):887–904, 2016. doi: 10.1080/14697688.2015.1099717. URL http://dx.doi.org/10.1080/14697688.2015.1099717.

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

Marc A. Berger and Victor J. Mizel. Volterra equations with Itô integrals. I. *J. Integral Equations*, 2(3):187–245, 1980a. ISSN 0163-5549.

Marc A. Berger and Victor J. Mizel. Volterra equations with Itô integrals. II. *J. Integral Equations*, 2(4):319–337, 1980b. ISSN 0163-5549.

Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. ISBN 0-471-19745-9. doi: 10.1002/9780470316962. URL http://dx.doi.org/10.1002/9780470316962. A Wiley-Interscience Publication.

Christa Cuchiero, Damir Filipović, Eberhard Mayerhofer, and Josef Teichmann. Affine processes on positive semidefinite matrices. *Ann. Appl. Probab.*, 21(2):397–463, 2011. ISSN 1050-5164. doi: 10.1214/10-AAP710. URL http://dx.doi.org/10.1214/10-AAP710.

Robert C. Dalang. Extending the martingale measure stochastic integral with applications to spatially homogeneous S.P.D.E.'s. *Electron. J. Probab.*, 4:no. 6, 29 pp. 1999. ISSN 1083-6489. doi: 10.1214/EJP.v4-43. URL http://dx.doi.org/10.1214/EJP.v4-43.

Laurent Decreusefond. Regularity properties of some stochastic Volterra integrals with singular kernel. *Potential Analysis*, 16(2):139–149, 2002.

D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053, 2003. ISSN 1050-5164. doi: 10.1214/aoap/1060202833. URL http://dx.doi.org/10.1214/aoap/1060202833.

Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *arXiv preprint arXiv:1609.02108*, 2016.

Omar El Euch and Mathieu Rosenbaum. Perfect hedging in rough Heston models. *arXiv preprint arXiv:1703.05049*, 2017.

Omar El Euch, Masaaki Fukasawa, and Mathieu Rosenbaum. The microstructural foundations of leverage effect and rough volatility. *arXiv preprint arXiv:1609.05177*, 2016.

Damir Filipović. *Term-structure models*. Springer Finance. Springer-Verlag, Berlin, 2009. ISBN 978-3-540-09726-6. doi: 10.1007/978-3-540-68015-4. URL http://dx.doi.org/10.1007/978-3-540-68015-4. A graduate course.

Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Available at SSRN 2509457*, 2014.

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. doi: 10.1017/CBO9780511662805. URL http://dx.doi.org/10.1017/CBO9780511662805.

Hamza Guennoun, Antoine Jacquier, and Patrick Roome. Asymptotic behaviour of the fractional Heston model. *arXiv preprint arXiv:1411.7653*, 2017.

Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993. ISSN 08939454, 14657368. URL http://www.jstor.org/stable/2962057.

Martina Hofmanová and Jan Seidler. On weak solutions of stochastic differential equations. *Stoch. Anal. Appl.*, 30(1):100–121, 2012. ISSN 0736-2994. doi: 10.1080/07362994.2012.628916. URL http://dx.doi.org/10.1080/07362994.2012.628916.

Martin Larsson and Paul Krühner. Affine processes with compact state space. *arXiv preprint arXiv:1706.07579*, 2017.

Leonid Mytnik and Eyal Neuman. Sample path properties of Volterra processes. *arXiv preprint arXiv:1101.4969*, 2011.

Leonid Mytnik and Thomas S. Salisbury. Uniqueness for Volterra-type stochastic integral equations. *arXiv preprint arXiv:1502.05513*, 2015.

Étienne Pardoux and Philip Protter. Stochastic Volterra equations with anticipating coefficients. *Ann. Probab.*, 18(4):1635–1655, 1990. ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(199010)18:4<1635:SVEWAC>2.0.CO;2-9&origin=MSN.

Philip Protter. Volterra equations driven by semimartingales. *Ann. Probab.*, 13 (2):519–530, 1985. ISSN 0091-1798. URL http://links.jstor.org/sici?sici=0091-1798(198505)13:2<519:VEDBS>2.0.CO;2-3&origin=MSN.

Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. ISBN 3-540-64325-7. doi: 10.1007/978-3-662-06400-9. URL http://dx.doi.org/10.1007/978-3-662-06400-9.

Peter Spreij and Enno Veerman. The affine transform formula for affine jump-diffusions with a general closed convex state space. *arXiv preprint arXiv:1005.1099*, 2010.

Peter Spreij and Enno Veerman. Affine diffusions with non-canonical state space. *Stoch. Anal. Appl.*, 30(4):605–641, 2012. ISSN 0736-2994. doi: 10.1080/07362994.2012.684322. URL http://dx.doi.org/10.1080/07362994.2012.684322.

Mark Veraar. The stochastic Fubini theorem revisited. *Stochastics*, 84(4):543–551, 2012. ISSN 1744-2508. doi: 10.1080/17442508.2011.618883. URL http://dx.doi.org/10.1080/17442508.2011.618883.

Zhidong Wang. Existence and uniqueness of solutions to stochastic Volterra equations with singular kernels and non-Lipschitz coefficients. *Statistics & Probability Letters*, 78 (9):1062–1071, 2008.

Xicheng Zhang. Stochastic Volterra equations in Banach spaces and stochastic partial differential equation. *J. Funct. Anal.*, 258(4):1361–1425, 2010. ISSN 0022-1236. doi: 10.1016/j.jfa.2009.11.006. URL http://dx.doi.org/10.1016/j.jfa.2009.11.006.