The Annals of Applied Probability
2024, Vol. 34, No. 1A, 318–356
https://doi.org/10.1214/23-AAP1965

VOLTERRA SQUARE-ROOT PROCESS: STATIONARITY AND REGULARITY OF THE LAW

BY MARTIN FRIESEN[1,a] AND PENG JIN[2,b]

[1]*School of Mathematical Sciences, Dublin City University,* [a]*martin.friesen@dcu.ie*
[2]*Guangdong Provincial Key Laboratory of IRADS, BNU-HKBU United International College,* [b]*pengjin@uic.edu.cn*

The Volterra square-root process on $\mathbb{R}_+^m$ is an affine Volterra process with continuous sample paths. Under a suitable integrability condition on the resolvent of the second kind associated with the Volterra convolution kernel, we establish the existence of limiting distributions. In contrast to the classical square-root diffusion process, here the limiting distributions may depend on the initial state of the process. Our result shows that the nonuniqueness of limiting distributions is closely related to the integrability of the Volterra convolution kernel. Using an extension of the exponential-affine transformation formula, we also give the construction of stationary processes associated with the limiting distributions. Finally, we prove that the time marginals as well as the limiting distributions, when restricted to the interior of the state space $\mathbb{R}_+^m$, are absolutely continuous with respect to the Lebesgue measure and their densities belong to some weighted Besov space of type $B_{1,\infty}^{\lambda}$.

1. Introduction.

1.1. *General introduction.* The analysis performed in [28] on intraday stock market data suggests that the volatility, seen as a stochastic process, has sample paths of very low regularity, and hence is not adequately captured by classical models such as the Heston model or the SABR model. Moreover, classical well-established Markovian models are often not able to capture the observed term structure of at-the-money volatility skew. To accommodate for these features, the authors propose to work with rough analogues of stochastic volatility models. The most prominent examples are the *rough Bergomi model* and the *rough Heston model*, where the latter is studied in [14–16]. We also refer interested readers to [4, 5, 20, 29, 32, 33] for some recent developments of this model. Extensions to multifactor settings and general Volterra kernels have been studied in [3], where general affine Volterra processes with continuous sample paths have been constructed. Other extensions to multiasset settings with rough correlations are recently studied in [1, 11]. While the rough sample path behavior observed in [28] still remains controversial (see, e.g., [10, 26, 27]), the newly emerged rough volatility models have proven to fit the empirical data remarkably well.

In this work, we study the Volterra square-root process, which provides the most general example of a continuous affine Volterra process on $\mathbb{R}_+^m$. We investigate their limiting distributions, construction of the stationary process, and absolute continuity of the law. Below we first recall the definition of the Volterra square-root process, and then discuss our results and the related literature. The m-dimensional *Volterra square-root process* $X = (X_t)_{t\geq 0}$ is obtained from the stochastic Volterra equation

$$X_t = x_0 + \int_0^t K(t-s)(b+\beta X_s)\,ds + \int_0^t K(t-s)\sigma(X_s)\,dB_s, \tag{1.1}$$

Received March 2022; revised March 2023.
MSC2020 subject classifications. Primary 60G22; secondary 45D05, 91G20.
Key words and phrases. Affine Volterra processes, square-root process, Volterra integral equations, mean-reversion, limiting distributions, absolute continuity.

where $\sigma(x) = \mathrm{diag}(\sigma_1\sqrt{x_1}, \dots, \sigma_m\sqrt{x_m})$, $x_0 \in \mathbb{R}_+^m$, $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}^{m\times m})$ and $(B_t)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}^m$. Here and below, we call (b, β, σ, K) *admissible*, if they satisfy:

(i) $b \in \mathbb{R}_+^m$;
(ii) $\beta = (\beta_{ij})_{i,j=1,\dots,m} \in \mathbb{R}^{m\times m}$ is such that $\beta_{ij} \geq 0$ for all $i \neq j$;
(iii) $\sigma = (\sigma_1, \dots, \sigma_m)^\top \in \mathbb{R}_+^m$;
(iv) the kernel K is diagonal with $K = \mathrm{diag}(K_1, \dots, K_m)$, where the scalar kernels $K_i \in L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R})$, $i = 1, \dots, m$;
(v) There exist constants $\gamma \in (0, 2]$ and $C_1 > 0$ such that

$$\int_0^h |K_i(r)|^2\,dr \leq C_1 h^\gamma, \quad h \in [0, 1], i = 1, \dots, m,$$

and for each $T > 0$ there exists $C_2(T) > 0$ such that

$$\int_0^T |K_i(r+h) - K_i(r)|^2\,dr \leq C_2(T)h^\gamma, \quad h \in [0, 1], i = 1, \dots, m;$$

(vi) For each $i = 1, \dots, m$ and each $h \in [0, 1]$, the shifted kernel $t \longmapsto K_i(t + h)$ is nonnegative, not identically zero, nonincreasing and continuous on $(0, \infty)$, and its resolvent of the first kind L_i is nonnegative and nonincreasing in the sense that $s \longmapsto L_i([s, s + t])$ is nonincreasing for all $t \geq 0$.

Note that for nonnegative, nonincreasing and not identically zero functions $K_1, \dots, K_m$, the resolvent of the first kind always exists; see [31], Chapter 5, Theorem 5.5. That is, there exist measures $L_1, \dots, L_m$ of locally bounded variation on $\mathbb{R}_+$ such that $K_i * L_i = L_i * K_i = 1$, $i = 1, \dots, m$, where $*$ denotes the usual convolution of functions or measures on $\mathbb{R}_+$.

REMARK 1.1. If $K_1, \dots, K_m$ are completely monotone and not identically zero, then condition (vi) holds; see [31], Chapter 5, Theorem 5.4.

For a discussion of condition (v), we refer to [3]. Given admissible parameters (b, β, σ, K), it follows from [3] that for each $x_0 \in \mathbb{R}_+^m$ there exists a unique (in law) $\mathbb{R}_+^m$-valued weak solution $X = (X_t)_{t\geq 0}$ of (1.1), Theorem 6.1, which is defined on some stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ supporting an m-dimensional Brownian motion $(B_t)_{t\geq 0}$. Moreover, for each $\eta \in (0, \gamma/2)$, X has a modification with η-Hölder continuous sample paths and satisfies $\sup_{t\in[0,T]} \mathbb{E}[|X_t|^p] < \infty$ for each $p \geq 2$ and $T > 0$.

EXAMPLE 1.2. In dimension $m = 1$ with $K(t) = t^{H-1/2}/\Gamma(H + 1/2)$, we recover the *rough Cox–Ingersoll–Ross process*, which reads as

$$X_t = x_0 + \int_0^t \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})}(b + \beta X_s)\,ds + \sigma \int_0^t \frac{(t-s)^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})}\sqrt{X_s}\,dB_s.$$

Here, $(B_t)_{t\geq 0}$ is a one-dimensional Brownian motion, $x_0, b, \sigma \geq 0$, $\beta \in \mathbb{R}$ and $H \in (0, 1/2)$. For this kernel K, one can choose $\gamma = 2H$ in the above definition of admissible parameters; see [3], Example 2.3. The process $(X_t)_{t\geq 0}$ is neither a finite-dimensional Markov process nor a semimartingale, which makes its mathematical study an interesting task.

Similar to the classical square-root process, there is a semiexplicit form for the Fourier–Laplace transform of the Volterra square-root process, that is, it is an affine process on $\mathbb{R}_+^m$. To state this formula in a compact form, let us define

$$R_i(u) = \langle u, \beta^i \rangle + \frac{\sigma_i^2}{2} u_i^2, \quad i = 1, \dots, m, u \in \mathbb{C}^m,$$

where $\beta^i = (\beta_{1i}, \dots, \beta_{mi})^\top$ denotes the ith column of the matrix β. Let $R = (R_1, \dots, R_m)^\top$ and set $\mathbb{C}_-^m = \{u \in \mathbb{C}^m : \mathrm{Re}(u) \le 0\}$. It follows from [3], Theorem 6.1, that for each $u \in \mathbb{C}_-^m$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}_-^m)$ the system of Riccati–Volterra equations

$$\psi(t) = K(t)u + \int_0^t K(t-s) f(s)\, ds + \int_0^t K(t-s) R\big(\psi(s)\big)\, ds \tag{1.2}$$

has a unique global solution $\psi = \psi(\cdot, u, f) \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}_-^m)$. The Volterra square-root process satisfies the exponential-affine transformation formula for the Fourier–Laplace transform

$$\begin{aligned} &\mathbb{E}\big[\mathrm{e}^{\langle u, X_t\rangle + \int_0^t \langle X_{t-s}, f(s)\rangle\, ds}\big] \\ &\quad = \exp\left\{\langle u, x_0\rangle + \int_0^t \langle f(s), x_0\rangle\, ds + \int_0^t \langle x_0, R\big(\psi(s)\big)\rangle\, ds + \int_0^t \langle b, \psi(s)\rangle\, ds\right\}. \end{aligned} \tag{1.3}$$

Here and below, we let, by slight abuse of notation, $\langle z, w\rangle := \sum_{i=1}^m z_i w_i$ for $z, w \in \mathbb{C}^m$. Note that $\langle \cdot, \cdot \rangle$ does not correspond to the usual inner product on $\mathbb{C}^m$.

1.2. *Long-time behavior.* Mean-reversion is a commonly accepted stylized fact in stochastic volatility modeling. Mathematically, this feature may be captured by the notion of ergodicity for the volatility process, that is, by limit distributions and stationarity. Even more so, stationarity also plays an important role for statistical inference. Namely, if one makes discrete time observations, then stationarity guarantees that these samples can be drawn from the same invariant distribution. Hence, one may estimate the parameters through the invariant distribution; see, for example, [8] where this was done for the subcritical Heston model. For other related results and implications for applications, we refer to [6, 7, 30].

It is well known that the classical CIR process (i.e., Example 1.2 with $H = 1/2$) is mean-reverting with long-term mean $-b\beta^{-1}$ and speed of mean-reversion β. Mathematically, this can be justified by studying ergodicity of the process. More precisely, if $\beta < 0$, the expected value satisfies $\mathbb{E}[X_t] \to b|\beta|^{-1}$ as $t \to \infty$ and the process has a unique limiting distribution π, which is stationary; see also [25, 35] where ergodicity of more general affine processes on the canonical state space were studied. Convergence results in the stronger total variation distance, and hence the law-of-large numbers have been studied in [23, 38] for general classes of affine processes.

In this work, we provide a sufficient condition for the existence of limiting distributions of the Volterra square-root process. Moreover, we characterize all limiting distributions and show that each limiting distribution gives rise to a stationary process, hence showing that the process is indeed mean-reverting. In contrast to the classical CIR process, the limiting distributions of the Volterra square-root process may depend on the initial state x_0 (even if $m = 1$ and β is negative or $b = 0$). We also characterize the case where the limiting distribution π_{x_0} actually depends on the initial state x_0, and prove that all limiting distributions satisfy $\pi_{x_0} = \pi_{Px_0} = \pi_0 * \pi_{Px_0}^{b=0}$. Here, P is a certain projection operator, and $\pi_{Px_0}^{b=0}$ denotes the limiting distribution of the Volterra square-root process with initial state Px_0 and $b = 0$. Our proof is essentially based on an analysis of the Riccati–Volterra equation (1.2) reformulated for the Laplace transform. The latter one is more suitable for the cone structure of $\mathbb{R}_+^m$; see Theorem 5.3 and Theorem 5.7.

For the existence of limiting distributions, it suffices to show that the limit $t \to \infty$ in (1.3) exists and is continuous at $u = 0$ and then to apply Lévy's continuity theorem as done in [35]. This requires to show that ψ obtained from (1.2) has additionally global integrability in time (e.g., $\psi \in L^1(\mathbb{R}_+; \mathbb{C}^m) \cap L^2(\mathbb{R}_+; \mathbb{C}^m)$), which is studied in Section 3. However, to prove the existence of an associated stationary process we cannot rely (in contrast to the literature) on the use of the Feller semigroup and classical Markovian techniques. In this work, we propose an alternative approach based on an extension of the exponential-affine transformation formula. Namely, we prove in Section 3 that formulas (1.2) and (1.3) can be extended from (u, f) to a general class of vector-valued measures μ. This allows us to effectively track the finite-dimensional distributions of the process. Hence, we can show that as $h \to \infty$, the process $(X_t^h)_{t\geq 0}$ defined by $X_t^h = X_{t+h}$ converges in law to a continuous process X^{stat}, which is the desired stationary process. Let us mention that this approach seems to be new even for classical affine processes.

At this point, it is worthwhile to mention that some of the Markovian methods can be recovered once the process is lifted to an infinite-dimensional Markov process described by a generalized Feller semigroup in the sense of [12]. Such an approach was recently used in [34] to tacke a similar problem to the one studied in this work. There the authors study a class of continuous affine Volterra processes on the canonical state-space $D = \mathbb{R}^n \times \mathbb{R}_+^m$ where the first n-components correspond to the log-asset prices while the last m-components are the stochastic volatility factors. In this framework, the authors prove under certain subcriticality conditions on the drift and for fractional kernels of type $K(t) = t^{\alpha-1}/\Gamma(\alpha)$ with $\alpha \in (0, 1)$ the existence of a stationary process for the Markovian lift. Uniqueness and the characterization of stationary processes is left open. Finally, applying their result to our setting, that is $n = 0$, we see that the volatility factors are essentially one-dimensional rough CIR processes as given in Example 1.2 with the additional restriction $b = 0$. Their key method is based on the observation that uniform boundedness of the first moment translates to uniform boundedness of the operator norm of the generalized Feller semigroup, which itself is shown to be sufficient for the existence of an invariant measure. In contrast, our methods are based on a detailed study of the Riccati–Volterra equation and allow us to study the multidimensional case for a general class of Volterra kernels K, allowing for $b \neq 0$, and also prove the uniqueness of stationary processes by providing a characterization of their finite-dimensional distributions.

Our results are in line with the existing literature on limiting distributions for stochastic Volterra equations. In [9], the authors studied limiting distributions for Lévy driven Volterra SPDEs. Using the Markovian lift onto the Filipovic space, they have shown that for kernels K being elements of the Filipovic space multiple limiting distributions may occur. At the same time, in [17] an abstract SPDE framework was provided to deal with SPDEs having multiple limiting distributions. This framework covers the Markovian lift onto the Filipovic space as well as stochastic delay equations. While both works cover rather general classes of Volterra stochastic equations, they require that the kernel is sufficiently regular, which excludes, for example, $K(t) = \frac{t^{H-1/2}}{\Gamma(H+1/2)}$ with $H \in (0, 1)$, and hence cannot be applied to the Volterra square-root process.

1.3. *Regularity of the law.* In the second part of this work, we turn to the study of regularity of the law of X_t for fixed $t > 0$. This includes absolute continuity of the law as well as regularity of the density. In the case of classical affine processes, such results can lead to the strong Feller property of the process (see [22]). We also want to point out that better regularity of the density are known for classical affine processes (see, e.g., (see [18])), where the authors studied density approximations and their applications in mathematical finance, and obtained C^k-regularity of the density up to the boundary. It is still unclear whether similar results can be obtained for affine Volterra processes.

In this work, we prove that, when $\sigma_1, \dots, \sigma_m > 0$ and K satisfies a suitable lower bound, the distribution of X_t in (1.1) is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}_{++}^m := \mathbb{R}_+^m \backslash \partial\mathbb{R}_+^m$, the interior of its state space. Moreover, our proof shows that the density $p_t(x)$ satisfies $\min\{1, x_1^{1/2}, \dots, x_m^{1/2}\} p_t(x) \in B_{1,\infty}^{\lambda}(\mathbb{R}^m)$, where $B_{1,\infty}^{\lambda}(\mathbb{R}^m)$ denotes the Besov space of order $(1, \infty)$ and some $\lambda \in (0, 1)$ denoting the regularity of the function. In particular, by Sobolev embeddings, the density has some low $L^p(\mathbb{R}^m)$-regularity. Moreover, assuming that the Hölder increments of K satisfy a global estimate in the spirit of condition (v) of the admissible parameters (see condition (K)), we also show that the limiting distributions π_{x_0} are absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m_{++}$ and have the same regularity as the law of X_t with $t > 0$. For a precise version of this results, we refer to Section 6.

Our proof is based on a method that was first introduced in [13] and subsequently applied in [22, 24] to continuous-state branching processes with immigration. While the aforementioned works aimed to prove existence of a transition density for a Markovian SDE, in this work we extend this method to Volterra stochastic equations and, additionally, demonstrate that this method also can be applied for the limiting distributions. Note that our method here is different from existing methods to study densities of classical affine processes, which are often based on estimates of the characteristic function; see [18] and [23].

1.4. *Application to the Volterra CIR process with a Gamma kernel.* In this section, we briefly state our results when applied to the *Volterra Cox–Ingersoll–Ross process with a Gamma kernel* obtained from

$$X_t = x_0 + \int_0^t K(t-s)(b+\beta X_s)\,ds + \sigma \int_0^t K(t-s)\sqrt{X_s}\,dB_s, \tag{1.4}$$

where $K(t) = \frac{t^{H-\frac{1}{2}}}{\Gamma(H+\frac{1}{2})}e^{-\lambda t}$, $H \in (0, 1/2)$, $\lambda, \sigma, b, x_0 \geq 0$, $\beta \in \mathbb{R}$ and $(B_t)_{t\geq 0}$ is a one-dimensional Brownian motion. The following is our main result on limiting distributions and stationarity of the process.

THEOREM 1.3. *Let X be obtained from (1.4) and $\beta < \lambda^{H+1/2}$. Then the process $(X_{t+h})_{t\geq 0}$ converges in law to a continuous stationary process $(X_t^{\text{stat}})_{t\geq 0}$ when $h \to \infty$. Moreover, the finite-dimensional distributions of X^{stat} have the characteristic function*

$$\mathbb{E}[e^{\sum_{j=1}^n \langle u_j, X_{t_j}^{\text{stat}}\rangle}] = \exp\left(\frac{\lambda^{H+1/2}x_0 + b}{\lambda^{H+1/2} - \beta}\left(\sum_{k=1}^n u_k + \frac{\sigma^2}{2}\int_0^\infty \psi(s)^2\,ds\right)\right),$$

where $0 \leq t_1 < \cdots < t_n$, $u_1, \ldots, u_n \in \mathbb{C}_-$ and ψ is the unique solution of

$$\begin{aligned}\psi(t) = &\sum_{j=1}^n \mathbb{1}_{\{t > t_n - t_j\}} K\big(t - (t_n - t_j)\big)u_j \\ &+ \int_0^t K(t-s)\left(\beta\psi(s) + \frac{\sigma^2}{2}\psi(s)^2\right)ds.\end{aligned}$$

Moreover, the first moment and the autocovariance function of the stationary process satisfy

$$\mathbb{E}[X_t^{\text{stat}}] = \frac{\lambda^{H+1/2}x_0 + b}{\lambda^{H+1/2} - \beta}$$

and if additionally $\beta < 0$ and $\sigma > 0$, then for $0 \leq s \leq t$,

$$\operatorname{cov}\big(X_t^{\text{stat}}, X_s^{\text{stat}}\big) \asymp (t-s)^{-(H+3/2)}e^{-\lambda(t-s)}, \quad t-s \to \infty. \tag{1.5}$$

Here and after, if f, g are positive functions the notation $f \asymp g$ means that there is a positive constant c such that $c^{-1}g \leq f \leq cg$. As a consequence of our results, we see that the stationary process X^{stat} is independent of the initial state x_0 if and only if $\lambda = 0$. Moreover, since for $H = 1/2$ the autocovariance function satisfies $\operatorname{cov}(X_t^{\text{stat}}, X_s^{\text{stat}}) \asymp e^{-(\lambda+|\beta|)(t-s)}$, which can be seen by direct computation for the classical CIR process, we find that for $\lambda = 0$ the autocovariance function has a phase transition from power-law to exponential decay when $H \nearrow 1/2$. Our result implies, in particular, that X_t converges weakly to some limiting distribution π_{x_0} when $t \to \infty$, and that its characteristic function is given by the expression in Theorem 1.3 with $n = 1$ and ψ being determined from (1.2). Note that in contrast to the classical CIR case where $H = 1/2$ and $\lambda = 0$, the limiting distribution may satisfy $\pi_{x_0} \neq \delta_0$ even when $b = 0$. At this point, we would also like to mention the recent work [21], which provides a heuristic argument on the existence of limit distributions and the form of its Fourier-transform for the Volterra CIR process.

The following is our main result on the regularity of the law when applied to the *Volterra CIR process with a Gamma kernel*.

THEOREM 1.4. *Let X be the Volterra CIR process with a Gamma kernel obtained from* (1.4). *Suppose that $\sigma > 0$. Then the following assertions hold*:

(a) *There exists some nonnegative function $p_t \in L^1(\mathbb{R}_+)$ such that*

$$\mathcal{L}(X_t)(dx) = \mathbb{P}[X_t = 0]\delta_0(dx) + p_t(x)\,dx, \quad \forall t > 0.$$

Let $p_t^*(x) = \mathbb{1}_{\mathbb{R}_+}(x)\min\{1, \sqrt{x}\}p_t(x)$, then there exists $\lambda \in (0, 1)$ and another constant $C > 0$ independent of t such that for $h \in [-1, 1]$ and $t > 0$,

$$\int_0^\infty \left|p_t^*(x+h) - p_t^*(x)\right| dx \le C|h|^\lambda (1 \wedge t)^{-H}.$$

(b) *Let π_{x_0} be the limiting distribution of X. Then*

$$\pi_{x_0}(dx) = \pi_{x_0}(\{0\})\delta_0(dx) + \rho_{x_0}(x)\,dx$$

for some $0 \le \rho_{x_0} \in L^1(\mathbb{R}_+)$. Letting $\rho_{x_0}^*(x) = \mathbb{1}_{\mathbb{R}_+}(x)\min\{1, \sqrt{x}\}\rho_{x_0}(x)$, then

$$\int_0^\infty \left|\rho_{x_0}^*(x+h) - \rho_{x_0}^*(x)\right| dx \le C|h|^\lambda, \quad h \in [-1, 1],$$

for the same λ as in part (*a*) *and some constant $C > 0$.*

As a consequence, apart from the origin the process X has a density in the interior of the state space. The above is a special case of our results from Section 6, which are applicable to arbitrary dimensions and a large class of kernels K. One implication of this regularity is related with convergence in total variation to the limiting distribution. Indeed, noting that $p_t^* \longrightarrow \min\{1, \sqrt{x}\}\pi_{x_0}(x)$ weakly as $t \to \infty$ and that $(p_t^*)_{t\ge 1} \subset L^1(\mathbb{R}_+)$ is relatively compact due to $\sup_{t\ge 1}\|p_t^*\|_{B^\lambda_{1,\infty}} < \infty$, we conclude that $p_t^* \longrightarrow \min\{1, \sqrt{x}\}\pi_{x_0}(x)$ in $L^1(\mathbb{R}_+)$, and hence in total variation.

These findings motivate us to study boundary nonattainment for the Volterra CIR process with a Gamma kernel, that is, conditions for $\mathbb{P}[X_t = 0, \forall t > 0] = 1$. In the classical case when $H = 1/2$ and $\lambda = 0$, such problem is related with the Feller condition to be imposed on σ and b. Its rough analogue is left for future research.

1.5. *Structure of the work.* This work is organized as follows. In Section 2, we collect some preliminaries for the study of Volterra integral equations. In Section 3, we study the Riccati–Volterra equation (1.2) and establish some regularity results for its solution $\psi(t, u, f)$. In Section 4, we first prove global bounds for the moments of the Volterra square-root process. Moreover, we show that the Hölder increments are uniformly bounded in L^2. Based on the moment bounds, we prove in Section 5 the weak convergence of the law of X_t to π_{x_0} when $t \to \infty$, construct the stationary process and finally provide a characterization when π_{x_0} actually depends on x_0. In Section 6, we prove for a large class of kernels K that the distribution of X_t as well as π_{x_0} is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^m_{++}$.

2. Preliminaries on Volterra integral equations.

2.1. *Convolution on* $\mathbb{R}_+$. For $p \in [1, \infty]$, we let $L^p([0,T]; \mathbb{C}^m)$ be the Banach space of equivalence classes of functions $f : [0,T] \longrightarrow \mathbb{C}^m$ with finite norm $\|f\|_{L^p([0,T])}$. Similarly, we define $L^p([0,T]; \mathbb{C}^{n\times k})$ with $n, k \in \mathbb{N}$ as the Banach space of matrix-valued functions $f : [0,T] \longrightarrow \mathbb{C}^{n\times k}$, where $\mathbb{C}^{n\times k}$ is equipped with the operator norm $\|A\|_2 = \sup_{|v|=1} |Av|$ with respect to the Euclidean distances on $\mathbb{C}^n$ and $\mathbb{C}^k$. We denote by $L^p_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m) = L^p_{\mathrm{loc}}$ and $L^p_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^{n\times k})$ the spaces of locally p-integrable functions.

The $m \times m$ identity matrix is denoted by I_m. Let $\|\cdot\|_{\mathrm{HS}}$ be defined by $\|A\|_{\mathrm{HS}} = \sqrt{\operatorname{tr}(A^*A)}$ be the Hilbert–Schmidt norm on $\mathbb{C}^{m\times m}$. Note that $\sqrt{m}^{-1}\|A\|_{\mathrm{HS}} \le \|A\|_2 \le \|A\|_{\mathrm{HS}}$ and both norms are submultiplicative in the sense that $\|AB\|_2 \le \|A\|_2\|B\|_2$ and $\|AB\|_{\mathrm{HS}} \le \|A\|_{\mathrm{HS}}\|B\|_{\mathrm{HS}}$. Let $T > 0$. The convolution of two functions $f : [0,T] \longrightarrow \mathbb{C}^{n\times k}$, $g : [0,T] \longrightarrow \mathbb{C}^{k\times d}$ is defined by $f * g(t) = \int_0^t f(t-s)g(s)\,ds$ for $t \in [0,T]$ with the matrix multiplication under the integral. We frequently use Young's inequality, which states that $\|f*g\|_{L^r([0,T])} \le \|f\|_{L^p([0,T])}\|g\|_{L^q([0,T])}$ whenever for $p, q, r \in [1,\infty]$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$ the right-hand side is finite.

REMARK 2.1. If $f \in L^p([0,T])$ and $g \in L^q([0,T])$ with $p, q \in [1,\infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$, then $f * g$ is continuous. To see this, set $\tilde{f}(x) = f(x)\mathbf{1}_{[0,T]}(x)$, $\tilde{g}(x) = g(x)\mathbf{1}_{[0,T]}(x)$ and then apply [37], Lemma 2.20.

We also use the convolution of a function and a measure. Namely, let $\mathcal{M}_{lf}$ be the space of all $\mathbb{C}^m$-valued set functions μ on $\mathbb{R}_+$ for which the restriction $\mu|_{[0,T]}$ with $T > 0$ is a $\mathbb{C}^m$-valued finite measure. For $\mu \in \mathcal{M}_{lf}$ and a compact set $E \subset \mathbb{R}_+$ recall that

$$|\mu|(E) := \sup\left\{\sum_{j=1}^N |\mu(E_j)| : \{E_j\}_{j=1}^N \text{ is a measurable partition of } E\right\}.$$

Given $f \in L^p([0,T]; \mathbb{C}^{m\times m})$ for some $p \in [1,\infty]$, we define the convolution with $\mu \in \mathcal{M}_{lf}$ by $f * \mu(t) = \int_{[0,t]} f(t-s)\mu(ds)$ where $t \in [0,T]$. It is easy to see that for each $p \in [1,\infty]$, $\|f * \mu\|_{L^p([0,T])} \le \|f\|_{L^p([0,T])}|\mu|([0,T])$. Moreover, if f is continuous on $[0,T]$ with $f(0) = 0$, then $f * \mu$ is also continuous on $[0,T]$.

2.2. *Volterra integral equations*. Let $K : \mathbb{R}_+ \longrightarrow \mathbb{C}^{m\times m}$ be locally integrable, that is, $K \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^{m\times m})$. Consider for given $h \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m)$ and $B \in \mathbb{C}^{m\times m}$ the Volterra convolution equation

$$x(t) = h(t) + \int_0^t K(t-s)Bx(s)\,ds.$$

Note that this equation is equivalent to $x + K_B * x = h$, where $K_B(t) = -K(t)B$. According to [31], Chapter 2, Theorem 3.5, it has a unique solution $x \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m)$ given by $x = h - R_B * h$, where $R_B \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m)$ is the *resolvent of the second kind* of the kernel K_B defined by the relation

$$R_B * K_B = K_B * R_B = K_B - R_B. \tag{2.1}$$

Note that $K_B \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m)$ guarantees that such a function R_B always exists; see [31], Chapter 2, Theorem 3.1. If the function h is of the form $h = K * \mu$ with $\mu \in \mathcal{M}_{lf}$, then the unique solution takes the form $x = K * \mu - R_B * K * \mu = E_B * \mu$, where we have set

$$E_B = K - R_B * K. \tag{2.2}$$

Note that $E_B(-B) = R_B$.

REMARK 2.2. By (2.1), (2.2) and Young's inequality one has $R_B, E_B \in L^p_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^{m\times m})$, whenever $K \in L^p_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^{m\times m})$ with $p \geq 1$. Moreover, if $p \geq 2$ and K is continuous on $(0,\infty)$, then R_B and E_B are also continuous on $(0,\infty)$.

REMARK 2.3. If $K = I_m$, then $E_B(t) = e^{Bt}$ and $R_B(t) = (-B)e^{Bt}$. In this case, integrability of E_B on $\mathbb{R}_+$ is equivalent to B having only eigenvalues with strictly negative real parts.

3. Analysis of the Riccati–Volterra equation.

3.1. *Regularity in time.* Here and below, we let $\psi \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m_-)$ benote the unique solution of (1.2) with $u \in \mathbb{C}^m_-$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m_-)$. We use the notation $R_{\beta^\top}$ and $E_{\beta^\top,}$ which are respectively defined by (2.1) and (2.2) with $B = \beta^\top$. Using different methods, we will later on see that $R_\beta = (R_{\beta^\top})^\top$ and $E_\beta = (E_{\beta^\top})^\top$.

LEMMA 3.1. *If $K \in L^p_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{R}^{m\times m})$ for some $p \in [1,\infty]$, then*

$$\begin{aligned}\|\psi(\cdot,u,f)\|_{L^p([0,T])} &\leq 2\big(|u| + \|f\|_{L^1([0,T])}\big)\|E_{\beta^\top}\|_{L^p([0,T])} \\ &\quad + \left(\sum_{i=1}^m \frac{\sigma_i^2}{2}\right)\big(|u| + \|f\|_{L^1([0,T])}\big)^2 \|E_{\beta^\top}\|_{L^p([0,T])}\|E_{\beta^\top}\|^2_{L^2([0,T])}\end{aligned}$$

for each $T > 0$. Moreover, if $\mathrm{Im}(u) = \mathrm{Im}(f) = 0$, *then*

$$\|\psi(\cdot,u,f)\|_{L^p([0,T])} \leq \big(|u| + \|f\|_{L^1([0,T])}\big)\|E_{\beta^\top}\|_{L^p([0,T])}.$$

PROOF. It follows from the proof of [3], Lemma 6.3, that the real and imaginary parts of ψ satisfy for each $i = 1,\dots,m$ the inequalities $\ell_i(t) \leq \mathrm{Re}(\psi_i(t,u,f)) \leq 0$ and $|\mathrm{Im}(\psi_i(t,u,f))| \leq h_i(t)$, where the functions ℓ_i and h_i are the unique global solutions of

$$\begin{aligned} h_i(t) &= K_i(t)\big|\mathrm{Im}(u_i)\big| + \int_0^t K_i(t-s)\big(\big|\mathrm{Im}\big(f_i(s)\big)\big| + \langle h(s), \beta^i\rangle\big)\,ds, \\ \ell_i(t) &= K_i(t)\,\mathrm{Re}(u_i) + \int_0^t K_i(t-s)\left(\mathrm{Re}\big(f_i(s)\big) + \langle \ell(s), \beta^i\rangle - \frac{\sigma_i^2}{2}h_i(s)^2\right)ds \end{aligned}$$

and $\beta^i = (\beta_{1i},\dots,\beta_{mi})^\top$ denotes the ith column of β. Hence, we obtain $\|\psi(\cdot,u,f)\|_{L^p([0,T])} \leq \|\ell\|_{L^p([0,T])} + \|h\|_{L^p([0,T])}$, where $\ell = (l_1,\dots,l_m)^\top$ and $h = (h_1,\dots,h_m)^\top$. To estimate the right-hand side of the previous inequality, we define for $i = 1,\dots,m$,

$$\begin{aligned} h_{0,i}(t) &= K_i(t)\big|\mathrm{Im}(u_i)\big| + \int_0^t K_i(t-s)\big|\mathrm{Im}\big(f_i(s)\big)\big|\,ds, \\ \ell_{0,i}(t) &= K_i(t)\,\mathrm{Re}(u_i) + \int_0^t K_i(t-s)\,\mathrm{Re}\big(f_i(s)\big)\,ds - \frac{\sigma_i^2}{2}\int_0^t K_i(t-s)h_i(s)^2\,ds. \end{aligned}$$

Then $h(t) = h_0(t) + \int_0^t K(t-s)\beta^\top h(s)\,ds$ and $\ell(t) = \ell_0(t) + \int_0^t K(t-s)\beta^\top \ell(s)\,ds$. This gives $h(t) = h_0(t) - (R_{\beta^\top} * h_0)(t)$, and hence

$$\begin{aligned} h(t) &= E_{\beta^\top}(t)\big(\big|\mathrm{Im}(u_1)\big|,\dots,\big|\mathrm{Im}(u_m)\big|\big)^\top \\ &\quad + E_{\beta^\top} * \big(\big|\mathrm{Im}(f_1)\big|,\dots,\big|\mathrm{Im}(f_m)\big|\big)^\top(t). \end{aligned} \tag{3.1}$$

Likewise we obtain $\ell(t) = \ell_0(t) - (R_{\beta^\top} * \ell_0)(t)$, and hence

$$\ell(t) = E_{\beta^\top}(t)\,\mathrm{Re}(u) + E_{\beta^\top} * \big(\mathrm{Re}(f) - \varkappa\big)(t), \tag{3.2}$$

where $\varkappa_i(t) = \frac{\sigma_i^2}{2} h_i(t)^2$. Young's inequality yields

$$\|h\|_{L^p([0,T])} \le \big(|u| + \|f\|_{L^1([0,T])}\big)\|E_{\beta^\top}\|_{L^p([0,T])}$$

and

$$\|\ell\|_{L^p([0,T])} \le \big(|u| + \|f\|_{L^1([0,T])} + \|\varkappa\|_{L^1([0,T])}\big)\|E_{\beta^\top}\|_{L^p([0,T])}. \tag{3.3}$$

Estimating

$$\begin{aligned}\|\varkappa\|_{L^1([0,T])} &\le \left(\sum_{i=1}^m \frac{\sigma_i^2}{2}\right)\|h\|^2_{L^2([0,T])} \\ &\le \left(\sum_{i=1}^m \frac{\sigma_i^2}{2}\right)\big(|u| + \|f\|_{L^1([0,T])}\big)^2\|E_{\beta^\top}\|^2_{L^2([0,T])}\end{aligned}$$

readily yields the first assertion. For the second assertion, note that $\mathrm{Im}(u) = \mathrm{Im}(f) = 0$ implies $\mathrm{Im}(\psi) = h = 0$, and hence $\ell(t) = E_{\beta^\top}(t)u + E_{\beta^\top} * f(t)$. The second estimate is now a consequence of Young's inequality. $\square$

The next result proves the continuity of ψ under the condition $f \in L^2_{\mathrm{loc}}$.

THEOREM 3.2. *If $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m_-)$, then $\psi \in C((0,\infty); \mathbb{C}^m_-)$. Moreover, if for each $T > 0$ there exists a nonnegative and nonincreasing function $v_T \in L^1([0,T]) \cap C((0,\infty))$, and $\alpha \in (0,1]$ such that*

$$\big\|K(t) - K(s)\big\|_{\mathrm{HS}} \le v_T(s)|t-s|^\alpha, \quad 0 < s < t \le T,$$

then $\psi \in C^{\alpha\wedge(\gamma/2)}_{\mathrm{loc}}((0,\infty); \mathbb{C}^m_-)$, where γ is given in assumption (v) of admissible parameters.

PROOF. Using representation (1.2), it suffices to show that each term on the right-hand side of (1.2) is continuous on $(0,\infty)$. Since $f, K \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$, Remark 2.1 yields that $K * f \in C([0,\infty))$, and since K is continuous on $(0,\infty)$, it suffices to show that $K * R(\psi)$ is also continuous on $(0,\infty)$. The latter one follows from Lemma A.1 applied to $k = K_i$ and $g = R_i(\psi)$ with $i = 1,\dots,m$ once we have shown that ψ is essentially bounded on $[t_0, T]$ for all $0 < t_0 < T$. Using the proof of Lemma 3.1, we get $|\psi(t)| \le |\mathrm{Re}(\psi(t))| + |\mathrm{Im}(\psi(t))| \le |\ell(t)| + |h(t)|$ where h is given by (3.1) and ℓ is the unique solution of (3.2). It suffices to prove that ℓ, h are continuous on $(0,\infty)$. Since $E_{\beta^\top}$ is continuous on $(0,\infty)$ and since $E_{\beta^\top}, (|\mathrm{Im}(f_1(s))|, \dots, |\mathrm{Im}(f_m(s))|)^\top \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$, formula (3.1) combined with Remark 2.1 yields that h is continuous on $(0,\infty)$. Likewise, since K is continuous on $(0,\infty)$ and since $K, f \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$, we see that the first two terms in (3.2) are continuous on $(0,\infty)$. Since $\ell \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$ by (3.3) and $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$, Remark 2.1 implies that the third term in (3.2) is continuous on $[0,\infty)$. Finally, $t \longmapsto \int_0^t K_i(t-s)h_i(s)^2\,ds$ is continuous on $(0,\infty)$ by Lemma A.1 from the Appendix. This proves that ψ is continuous on $(0,\infty)$.

Next, we prove local Hölder continuity. Let $s, t \in (0,T]$ be such that $s < t$. Then $|\psi(t) - \psi(s)| \le I_1 + I_2 + I_3 + I_4$ with $I_1 = \|K(t) - K(s)\|_{\mathrm{HS}}|u| \le v_T(s)|u||t-s|^\alpha$,

$$\begin{aligned}I_2 &= \int_0^s \big\|K(t-r) - K(s-r)\big\|_{\mathrm{HS}}\big|f(r)\big|\,dr + \int_s^t \big\|K(t-r)\big\|_{\mathrm{HS}}\big|f(r)\big|\,dr, \\ I_3 &= \int_0^s \big\|K(t-r) - K(s-r)\big\|_{\mathrm{HS}}\big|R\big(\psi(r)\big)\big|\,dr, \\ I_4 &= \int_s^t \big\|K(t-r)\big\|_{\mathrm{HS}}\big|R\big(\psi(r)\big)\big|\,dr.\end{aligned}$$

For I_2, we use assumption (v) from the admissible parameters to find that $I_2 \le (C_1 + C_2(T))\|f\|_{L^2([0,T])}(t-s)^{\gamma/2}$. For I_3, we use $|R(\psi)| \le C(1+|\psi|^2)$ to find that $I_3 \le C(t-s)^\alpha \int_0^s v_T(s-r)(1+|\psi(r)|^2)\,dr$. Since $1+|\psi|^2 \in L^1_{\mathrm{loc}}(\mathbb{R}_+) \cap L^\infty_{\mathrm{loc}}((0,\infty))$, it follows from Lemma A.1 that the convolution on the right-hand side defines a continuous function in $s \in (0,\infty)$. Finally, for I_4 we obtain $I_4 \le C(1+\|\psi\|^2_{L^\infty([s,T])}) \int_s^t \|K(t-r)\|_{\mathrm{HS}}\,ds \le C(1+\|\psi\|^2_{L^\infty([s,T])})(t-s)^{\gamma/2}$. Collecting all estimates proves the assertion. $\square$

REMARK 3.3. Suppose that $f \in L^{3/2}_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{C}^m_-)$, $K \in L^3_{\mathrm{loc}}(\mathbb{R}_+)$, and there exists $\gamma' \in (0,1]$ such that for each $T>0$ there exists $C(T)>0$ with

$$\int_0^T \|K(r+h)-K(r)\|^3_{\mathrm{HS}}\,dr + \int_0^h \|K(r)\|^3_{\mathrm{HS}}\,dr \le C(T)h^{3\gamma'} \tag{3.4}$$

for $h \in (0,1]$. Then for $u=0$, we have $\psi \in C^{\gamma'}_{\mathrm{loc}}([0,\infty);\mathbb{C}^m_-)$, while for $u \ne 0$ and $K \in C^\alpha_{\mathrm{loc}}((0,\infty))$ we have $\psi \in C^{\alpha\wedge\gamma'}_{\mathrm{loc}}((0,\infty);\mathbb{C}^m_-)$.

PROOF. For $0 \le s < t$, we find that $|\psi(t)-\psi(s)| \le I_1+I_2+I_3+I_4$ with $I_1,\dots,I_4$ as above. Now $I_1 = 0$, $I_2 \le C(T)\|f\|_{L^{3/2}([0,T])}(t-s)^{\gamma'}$, and using (3.4) we have $I_3+I_4 \le 2(t-s)^{\gamma'}\|R(\psi)\|_{L^{3/2}([0,T])}$, where $R(\psi) \in L^{3/2}_{\mathrm{loc}}$ due to $|R(\psi)| \le C(1+|\psi|^2)$ and $\psi \in L^3_{\mathrm{loc}}$ by Lemma 3.1 with $p=3$. This proves $\psi \in C^{\gamma'}_{\mathrm{loc}}([0,\infty);\mathbb{C}^m_-)$. Moreover, if $u \ne 0$ and $K \in C^\alpha_{\mathrm{loc}}((0,\infty))$, then $I_1 \le C_T(t-s)^\alpha$ for $0 < s < t \le T$, which proves the second assertion. $\square$

EXAMPLE 3.4. Let $m=1$ and $K(t) = \frac{t^{H-1/2}}{\Gamma(H)}$ with $H \in (0,1/2)$. Then the conditions of Theorem 3.2 are satisfied for $v_T(s) = C_T \min\{1, s^{2H-1}\}$ and $\alpha = \frac{1}{2}-H$. Hence, $\psi \in C^{H\wedge(1/2-H)}_{\mathrm{loc}}((0,\infty);\mathbb{C}_-)$, provided that $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{C}_-)$. Moreover, if $H \in (1/6,1/2)$, then $K \in L^3_{\mathrm{loc}}(\mathbb{R}_+)$ and (3.4) holds for $\gamma' = H-1/6$, and hence $\psi \in C^{H-1/6}_{\mathrm{loc}}([0,\infty),\mathbb{C}_-)$ if $u=0$.

Repeating the above proofs for the components ψ_j, $j=1,\dots,m$, we may also obtain Hölder continuity for ψ with kernels with different $H_1,\dots,H_m$.

REMARK 3.5. Let $K(t) = \mathrm{diag}(\frac{t^{H_1-1/2}}{\Gamma(H_1)},\dots,\frac{t^{H_m-1/2}}{\Gamma(H_m)})$ with $H_1,\dots,H_m \in (0,1/2)$, $u \in \mathbb{C}^m_-$ and $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{C}^m_-)$. Then $\psi_j \in C^{H_j\wedge(1/2-H_j)}_{\mathrm{loc}}((0,\infty);\mathbb{C}^m_-)$, and if $H_j > 1/6$ and $u=0$ then $\psi_j \in C^{H_j-1/6}_{\mathrm{loc}}([0,\infty),\mathbb{C}^m_-)$. Details are left for the reader.

In order to extend the affine transformation formula toward measures, it is convenient to use compactness arguments in $L^p([0,T])$. For this purpose, we prove explicit bounds on the fractional Sobolev norm of $\psi(\cdot,u,f)$. Given $p \ge 2$ and $\eta \in (0,1)$, let $W^{\eta,p}([0,T])$ be the Banach space of equivalence classes of functions $g:[0,T] \longrightarrow \mathbb{C}^m$ with finite norm

$$\|g\|_{W^{\eta,p}([0,T])} = \left(\int_0^T |g(t)|^p\,dt + \int_0^T\int_0^T \frac{|g(t)-g(s)|^p}{|t-s|^{1+\eta p}}\,ds\,dt\right)^{1/p}.$$

Finally, define

$$[K]_{\eta,p,T} = \left(\int_0^T t^{-\eta p}\|K(t)\|_2^p\,dt + \int_0^T\int_0^T \frac{\|K(t)-K(s)\|_2^p}{|t-s|^{1+\eta p}}\,ds\,dt\right)^{1/p}.$$

EXAMPLE 3.6. The kernel $K(t) = \frac{t^{H-1/2}}{\Gamma(H+1/2)} e^{-\lambda t} I_m$ with $H \in (0, 1/2)$ and $\lambda \geq 0$ satisfies $[K]_{\eta,p,T} < \infty$ for each $T > 0$, $p = 2$, and $\eta \in (0, H)$, [2].

The following is our second regularity result in t for the solution of (1.2).

THEOREM 3.7. *If $[K]_{\eta,p,T} < \infty$ for $p \geq 2$, $\eta \in (0, 1)$ and $T > 0$, then*

$$\begin{aligned}
&\|\psi(\cdot, u, f)\|_{W^{\eta,p}([0,T])} \\
&\quad \leq \|\psi(\cdot, u, f)\|_{L^p([0,T])} \\
&\qquad + C\bigl(1 + [K]_{\eta,p,T}\bigr)\bigl(1 + |u| + \|f\|_{L^1([0,T])} + \|\psi(\cdot, u, f)\|^2_{L^2([0,T])}\bigr),
\end{aligned}$$

where the constant C only depends on T, p, m, β, σ.

PROOF. Here and below, we let C be a generic constant independent of u and f. Let I_1, I_2, I_3 be the same as in the proof of Theorem 3.2. Then we obtain $|\psi(t, u, f) - \psi(s, u, f)|^p \leq C I_1^p + C I_2^p + C I_3^p$ for $0 \leq s < t \leq T$. It is easy to see that $I_1^p \leq \|K(t) - K(s)\|_2^p |u|^p$. Let $g(r) := |R(\psi(r, u, f)|$. Then

$$\begin{aligned}
I_3^p &\leq C\biggl(\int_s^t \|K(t-r)\|_2 g(r)\,dr\biggr)^p \\
&\quad + C\biggl(\int_0^s \|K(t-r) - K(s-r)\|_2 g(r)\,dr\biggr)^p \\
&\leq C\bigl(1 + \|\psi\|^{2p-2}_{L^2([0,T])}\bigr)\biggl(\int_s^t \|K(t-r)\|_2^p g(r)\,dr\biggr) \\
&\quad + C\bigl(1 + \|\psi\|^{2p-2}_{L^2([0,T])}\bigr)\biggl(\int_0^s \|K(t-r) - K(s-r)\|_2^p g(r)\,dr\biggr),
\end{aligned}$$

where we have used $\int_0^T g(r)\,dr \leq C(1 + \|\psi\|^2_{L^2([0,T])})$ due to $|R(u)| \leq C(1 + |u|^2)$, and the inequality $(\int_a^b h(r) g(r)\,dr)^p \leq (\int_a^b g(r)\,dr)^{p-1} \int_a^b h(r)^p g(r)\,dr$ for $h \geq 0$ and $0 \leq a < b$. Thus, we obtain

$$\begin{aligned}
&\int_0^T \int_0^T \frac{I_3^p}{|t-s|^{1+\eta p}}\,ds\,dt \\
&\quad \leq C\bigl(1 + \|\psi\|^{2p-2}_{L^2([0,T])}\bigr) \int_0^T \int_0^t \int_s^t \frac{\|K(t-r)\|_2^p}{|t-s|^{1+\eta p}} g(r)\,dr\,ds\,dt \\
&\qquad + C\bigl(1 + \|\psi\|^{2p-2}_{L^2([0,T])}\bigr) \int_0^T \int_0^t \int_0^s \frac{\|K(t-r) - K(s-r)\|_2^p}{|t-s|^{1+\eta p}} g(r)\,dr\,ds\,dt \qquad (3.5) \\
&\quad \leq C\bigl(1 + \|\psi\|^{2p-2}_{L^2([0,T])}\bigr)[K]_{\eta,p,T}\biggl(\int_0^T g(r)\,dr\biggr) \\
&\quad \leq C\bigl(1 + [K]^p_{\eta,p,T}\bigr)\bigl(1 + \|\psi\|^{2p}_{L^2([0,T])}\bigr),
\end{aligned}$$

where in (3.5) we have used Fubini's theorem to get

$$\int_0^T \int_0^t \int_s^t \frac{\|K(t-r)\|_2^p}{|t-s|^{1+\eta p}} g(r)\,dr\,ds\,dt \leq (\eta p)^{-1}[K]_{\eta,p,T}\biggl(\int_0^T g(r)\,dr\biggr).$$

Repeating the above arguments for I_2, we obtain

$$\int_0^T \int_0^T \frac{I_2^p}{|t-s|^{1+\eta p}}\,ds\,dt \leq C\bigl(1 + [K]^p_{\eta,p,T}\bigr)\bigl(1 + \|f\|^p_{L^1([0,T])}\bigr).$$

In view of these estimates, we obtain

$$\begin{aligned}
&\int_0^T \int_0^T \frac{|\psi(t,u,f)-\psi(s,u,f)|^p}{|t-s|^{1+\eta p}}\,ds\,dt \\
&\quad\le C|u|^p \int_0^T \int_0^T \frac{\|K(t)-K(s)\|_2^p}{|t-s|^{1+\eta p}}\,ds\,dt \\
&\qquad + C\big(1+[K]_{\eta,p,T}^p\big)\big(1+\|f\|_{L^1([0,T])}^p+\|\psi\|_{L^2([0,T])}^{2p}\big) \\
&\quad\le C\big(1+[K]_{\eta,p,T}^p\big)\big(1+|u|^p+\|f\|_{L^1([0,T])}^p+\|\psi\|_{L^2([0,T])}^{2p}\big).
\end{aligned}$$

In view of Lemma 3.1, the assertion is proved. □

3.2. *Extension to measure convolutions.* Let $\mathcal{M}_{lf}^- \subset \mathcal{M}_{lf}$ be the subset of $\mathbb{C}^m$-valued set functions $\mu \in \mathcal{M}_{lf}$, which satisfy $\mathrm{Re}(\mu) \le 0$. Below we extend the affine formula from $(u,f) \in \mathbb{C}_-^m \times L_{\mathrm{loc}}^1(\mathbb{R}_+;\mathbb{C}_-^m)$ to $(u,\mu) \in \mathbb{C}_-^m \times \mathcal{M}_{lf}^-$. The latter one provides us the key tool to explicitly identify the finite-dimensional distributions of the stationary process via a Volterra–Riccati-type equation (see Section 5).

LEMMA 3.8. *For each $\mu \in \mathcal{M}_{lf}^-$, there exists $(f_n)_{n\ge 1} \subset L_{\mathrm{loc}}^1(\mathbb{R}_+;\mathbb{C}_-^m)$ such that*:

(i) $\|f_n\|_{L^1([0,T])} \le |\mu|([0,T])$ *for each $T>0$*;

(ii) *For each $T>0$, $p\ge 1$ and $g \in L^p([0,T];\mathbb{C}^m)$, one has $g^\top * f_n \to g^\top * \mu$ in $L^p([0,T])$*;

(iii) *For each $T>0$ and each $g \in C([0,T];\mathbb{C}^m)$ with $g(0)=0$, one has*

$$\lim_{n\to\infty} \int_0^t \langle g(t-s), f_n(s)\rangle\,ds = \int_{[0,t]} \langle g(t-s), \mu(ds)\rangle$$

pointwise for each $t \in [0,T]$.

PROOF. Let $\rho_n(t) = ne^{-nt}$, $t \ge 0$ and define $f_n(t) = \int_{[0,t]} \rho_n(t-s)\mu(ds)$. Then $\mathrm{Re}(f_n) \le 0$ and $\|f_n\|_{L^1([0,T])} \le |\mu|([0,T])$. Classical results from Fourier analysis (see [37], Theorem 2.16) shows that $\rho_n * h \to h$ in L^p if $h \in L^p$. The assertion (ii) now follows from $g_i * f_{n,i} = g_i * \rho_n * \mu_i = \rho_n * (g_i * \mu_i)$, $i=1,\dots,m$. Let us turn to (iii). First, for $t=0$, the desired convergence is true due to $g(0)=0$. Suppose $t>0$. We note that $\int_0^t g_i(t-s) f_{n,i}(s)\,ds = \int_{[0,t)} g_i * \rho_n(t-s)\mu_i(ds)$, where we have used $(g_i * \rho_n)(0)\mu_i(\{t\}) = 0$. By dominated convergence, the right-hand side of the last equality converges to $\int_{[0,t)} g_i(t-s)\mu_i(ds)$, since $g_i * \rho_n(t-s) \to g_i(t-s)$ for each $s \in [0,t)$ as $n \to \infty$. Hence,

$$\lim_{n\to\infty} \int_0^t g_i(t-s) f_{n,i}(s)\,ds = \int_{[0,t)} g_i(t-s)\mu_i(ds) = \int_{[0,t]} g_i(t-s)\mu_i(ds),$$

which proves the assertion (iii). □

REMARK 3.9. If μ is nonatomic, that is, $\mu(\{t\})=0$ holds for each $t\ge 0$, then the additional assumption $g(0)=0$ can be omitted.

The next result extends (1.2) to measures $\mu \in \mathcal{M}_{lf}^-$.

THEOREM 3.10. *Suppose there exist $p\ge 2$ and $\eta \in (0,1)$ such that $[K]_{\eta,p,T} < \infty$ for each $T>0$. Then the following assertions hold*:

(a) *For each $\mu \in \mathcal{M}_{lf}^-$, there exists a unique $L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m_-)$-valued solution*

$$\psi(t, \mu) = \int_{[0,t]} K(t-s)\mu(ds) + \int_0^t K(t-s) R\big(\psi(s, \mu)\big)\, ds. \tag{3.6}$$

(b) *For each $q \in [1, p]$, this unique solution satisfies*

$$\begin{aligned}\|\psi(\cdot, \mu)\|_{L^q([0,T])} &\le 2|\mu|\big([0, T]\big)\|E_{\beta^\top}\|_{L^q([0,T])} \\ &\quad + \left(\sum_{i=1}^m \frac{\sigma_i^2}{2}\right)|\mu|\big([0, T]\big)^2 \|E_{\beta^\top}\|_{L^q([0,T])}\|E_{\beta^\top}\|^2_{L^2([0,T])}\end{aligned}$$

and if $\mathrm{Im}(\mu) = 0$, *then even*

$$\|\psi(\cdot, \mu)\|_{L^q([0,T])} \le |\mu|\big([0, T]\big)\|E_{\beta^\top}\|_{L^q([0,T])}.$$

Finally, there exists a constant C independent of μ such that

$$\begin{aligned}\|\psi(\cdot, \mu)\|_{W^{\eta,p}([0,T])} &\le \|\psi(\cdot, \mu)\|_{L^p([0,T])} \\ &\quad + C\big(1 + [K]_{\eta,p,T}\big)\big(1 + |\mu|([0, T]) + \|\psi(\cdot, \mu)\|^2_{L^2([0,T])}\big).\end{aligned}$$

(c) *If $p \ge 3$, then the function $\psi(\cdot, \mu)$ is continuous at each $t_0 \ge 0$ for which the convolution $K * \mu(\cdot)$ is continuous at t_0.*

PROOF. Let $(f_n)_{n\ge1} \subset L^1_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m_-)$ be a sequence of functions as given in Lemma 3.8. Let $\psi_n = \psi(\cdot, 0, f_n)$ be the sequence of unique solutions of (1.2). Fix $T > 0$. Then using Lemma 3.1 and Lemma 3.8(i), we obtain for each $q \in [1, p]$,

$$\begin{aligned}\|\psi(\cdot, 0, f_n)\|_{L^q([0,T])} &\le 2|\mu|\big([0, T]\big)\|E_{\beta^\top}\|_{L^q([0,T])} \\ &\quad + \left(\sum_{i=1}^m \frac{\sigma_i^2}{2}\right)|\mu|\big([0, T]\big)^2 \|E_{\beta^\top}\|_{L^q([0,T])}\|E_{\beta^\top}\|^2_{L^2([0,T])},\end{aligned}$$

and if $\mathrm{Im}(\mu) = 0$, then $\|\psi(\cdot, 0, f_n)\|_{L^q([0,T])} \le |\mu|([0, T])\|E_{\beta^\top}\|_{L^q([0,T])}$. Hence, Theorem 3.7 combined with Lemma 3.8(i) implies that

$$\begin{aligned}\|\psi(\cdot, 0, f_n)\|_{W^{\eta,p}([0,T])} &\le \|\psi(\cdot, 0, f_n)\|_{L^p([0,T])} \\ &\quad + C\big(1 + [K]_{\eta,p,T}\big)\big(1 + |\mu|([0, T]) + \|\psi(\cdot, 0, f_n)\|^2_{L^2([0,T])}\big).\end{aligned}$$

In view of the L^q-estimates on $\psi(\cdot, 0, f_n)$ and Remark 2.2, the right-hand side is bounded in n. Since the ball $\{g \in L^p([0, T]) : \|g\|_{W^{\eta,p}([0,T])} \le R\}$ with $R > 0$ is relatively compact in $L^p([0, T]; \mathbb{C}^m)$ (see [19], Theorem 2.1), we find a subsequence $(f_{n_k})_{k\ge1}$ such that $\psi(\cdot, 0, f_{n_k}) \longrightarrow \psi$ in $L^p([0, T]; \mathbb{C}^m)$. Further, we can choose a subsubsequence, still denoted by (f_{n_k}), such that $\psi(\cdot, 0, f_{n_k})$ converges almost surely to ψ on $[0, T]$. Taking the limit $k \to \infty$ and using the lemma of Fatou proves the estimates from part (b).

Next, we show that $\psi = \psi(\cdot, \mu)$ is a solution of (3.6) on $[0, T]$. Since $\psi_{n_k} \to \psi$ and $K * f_{n_k} \longrightarrow K * \mu$ in $L^p([0, T]; \mathbb{C}^m)$, it suffices to show that $K * R(\psi_{n_k}) \longrightarrow K * R(\psi)$ holds in $L^p([0, T]; \mathbb{C}^m)$. For this purpose, we first use Young's inequality, then

$$\big|R(u) - R(v)\big| \le C\big(1 + |v| + |u|\big)|u - v|, \tag{3.7}$$

and finally the Cauchy–Schwarz inequality to find that

$$\begin{aligned}&\|K * R(\psi_{n_k}) - K * R(\psi)\|_{L^p([0,T])} \\ &\quad \le C\|K\|_{L^p([0,T])}\big(1 + \|\psi_{n_k}\|_{L^2([0,T])} + \|\psi\|_{L^2([0,T])}\big)\|\psi_{n_k} - \psi\|_{L^2([0,T])}.\end{aligned}$$

Since the right-hand side converges to zero, we find that ψ is a global solution of (3.6). Noting that (3.7) holds and that $K * \mu \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^m)$, [3], Theorem B.1, implies that this equation has a unique maximal solution. Since ψ is a global solution, the unique maximal solution is defined on all $\mathbb{R}_+$ and coincides with ψ. This proves part (a).

To prove part (c), in view of (3.6), it suffices to show that $K * R(\psi)$ is continuous on $\mathbb{R}_+$. The latter one is true, if $K \in L^3_{\mathrm{loc}}$ and $R(\psi) \in L^{3/2}_{\mathrm{loc}}$, which holds true due to $|R(\psi)| \leq C(1 + |\psi|^2)$ and $\psi \in L^{3/2}_{\mathrm{loc}}$. This proves part (c). $\square$

Finally, we extend the exponential-affine transformation formula.

COROLLARY 3.11. *Let (b, β, σ, K) be admissible parameters and suppose there exist $p \geq 2$ and $\eta \in (0, 1)$ such that $[K]_{\eta,p,T} < \infty$ for each $T > 0$. Then*

$$\begin{aligned}
&\mathbb{E}\big[e^{\int_{[0,t]}\langle X_{t-s}, \mu(ds)\rangle}\big] \\
&\quad = \exp\left\{\langle x_0, \mu([0,t])\rangle + \int_0^t \langle x_0, R(\psi(s,\mu))\rangle\, ds + \int_0^t \langle b, \psi(s,\mu)\rangle\, ds\right\} \\
&\quad = \exp\left\{\int_{[0,t]} \langle \mathbb{E}[X_{t-s}], \mu(ds)\rangle + \sum_{i=1}^m \frac{\sigma_i^2}{2} \int_0^t \mathbb{E}[X_{i,t-s}]\psi_i(s,\mu)^2\, ds\right\}
\end{aligned}$$

hold for each $\mu \in \mathcal{M}_{lf}^-$, where ψ denotes the unique solution of (3.6).

PROOF. For the first equality, we let f_{n_k}, ψ_{n_k} be the same as in the proof of Theorem 3.10. Then

$$\mathbb{E}\big[e^{\int_0^t \langle X_s - x_0, f_{n_k}(t-s)\rangle\, ds}\big] = \exp\left\{\int_0^t \langle x_0, R(\psi_{n_k}(s))\rangle\, ds + \int_0^t \langle b, \psi_{n_k}(s)\rangle\, ds\right\}. \tag{3.8}$$

Using Lemma 3.8(iii) for $g(s) = X_s - x_0$ gives $\int_0^t \langle X_s - x_0, f_{n_k}(t-s)\rangle\, ds \longrightarrow \int_{[0,t]} \langle X_{t-s} - x_0, \mu(ds)\rangle$ pointwise, and hence dominated convergence combined with the L^2-convergence of ψ_{n_k} implies that we can pass to the limit in (3.8). For the second identity, first observe that

$$\begin{aligned}
&\mathbb{E}\big[e^{\int_0^t \langle X_s - x_0, f_{n_k}(t-s)\rangle\, ds}\big] \\
&\quad = \exp\left\{\int_0^t \langle \mathbb{E}[X_{t-s}] - x_0, f_{n_k}(s)\rangle\, ds + \sum_{i=1}^m \frac{\sigma_i^2}{2} \int_0^t \mathbb{E}[X_{i,t-s}]\psi_{i,n}(s)^2\, ds\right\}
\end{aligned}$$

holds by [3], equations (4.5), (4.7). It suffices to show that the right-hand side converges to the desired limit as $n \to \infty$. Since $g(r) = \mathbb{E}[X_r] - x_0$ is continuous (see (4.3) below) and $g(0) = 0$, by Lemma 3.8(iii), we find $\lim_{k\to\infty} \int_0^t \langle \mathbb{E}[X_{t-s}] - x_0, f_{n_k}(s)\rangle\, ds = \int_{[0,t]} \langle \mathbb{E}[X_{t-s}] - x_0, \mu(ds)\rangle$. For the second term, we note that

$$\begin{aligned}
&\left|\int_0^t \mathbb{E}[X_{i,t-s}]\psi_{i,n}(s)^2\, ds - \int_0^t \mathbb{E}[X_{i,t-s}]\psi_i(s)^2\, ds\right| \\
&\quad \leq \sup_{s\in[0,t]} \mathbb{E}[X_{i,s}]\Big(\sup_{k\geq 1} \|\psi_{n_k}\|_{L^2([0,T])} + \|\psi\|_{L^2([0,T])}\Big)\|\psi_{n_k} - \psi\|_{L^2([0,T])},
\end{aligned}$$

which tends to zero as $k \to \infty$. Here, we used the fact that $\sup_{s\in[0,t]} \mathbb{E}[X_{i,s}] < \infty$ due to [3]. $\square$

3.3. *Stability in the admissible parameters.* As a particular application of our results, we show that the Volterra square-root process depends continuously on the admissible parameters (b, β, σ, K).

THEOREM 3.12. *Let* $(b, \beta, \sigma, K), (b, \beta^{(n)}, \sigma^{(n)}, K^{(n)})$ *be admissible parameters with the properties* (i) $\beta^{(n)} \longrightarrow \beta$; (ii) $\sigma^{(n)} \longrightarrow \sigma$; (iii) *there exists* $p \geq 2$ *such that* $\|K^{(n)} - K\|_{L^p([0,T])} \longrightarrow 0$ *for each* $T > 0$; (iv) *there exists* $\eta \in (0,1)$ *with* $[K]_{\eta,p,T} + \sup_{n\geq 1}[K^{(n)}]_{\eta,p,T} < \infty$ *for each* $T > 0$. *For* $\mu, \mu^{(n)} \in \mathcal{M}_{lf}^-$, *let* ψ *and* ψ_n *be the corresponding unique solutions of* (3.6). *Suppose that* $|\mu^{(n)} - \mu|([0,T]) \longrightarrow 0$ *for each* $T > 0$. *Then*

$$\lim_{n\to\infty} \|\psi_n - \psi\|_{L^p([0,T])} = 0.$$

PROOF. Since $\beta^{(n)}, \sigma^{(n)}$ converge, they are bounded. Similarly, we have $\sup_{n\geq 1}\|K^{(n)}\|_{L^2([0,t])} < \infty$ for each $T > 0$. Noting that $K^{(n)}(\beta^{(n)})^\top \to K\beta^\top$ in $L^p([0,T])$ for each $T > 0$, by [31], Theorem 2.3.1, we also obtain $\sup_{n\geq 1}\|R^{(n)}_{(\beta^{(n)})^\top}\|_{L^1([0,T])} < \infty$. In view of the definition in (2.2), we thus get $\sup_{n\geq 1}\|E^{(n)}_{(\beta^{(n)})^\top}\|_{L^p([0,T])} < \infty$ by Young's inequality. Moreover, the particular form of the inequalities in Theorem 3.10(b) as well as the constant C imply $\sup_{n\geq 1}\|\psi_n\|_{L^p([0,T])} < \infty$ and subsequently $\sup_{n\geq 1}\|\psi_n\|_{W^{\eta,p}([0,T])} < \infty$. Hence $(\psi_n)_{n\in\mathbb{N}} \subset L^p([0,T]; \mathbb{C}^m_-)$ is relatively compact. Let $\tilde{\psi}$ be the limit for some subsequence ψ_{n_k}. If we can show that $\tilde{\psi}$ is a solution of (3.6), by uniqueness, we must have $\psi = \tilde{\psi}$, and thus the convergence $\|\psi_n - \psi\|_{L^p([0,T])} = 0$ as $n \to \infty$, since the convergent subsequence is arbitrarily chosen.

Next, we show that $\tilde{\psi}$ is a solution of (3.6), that is, $\tilde{\psi} = K * \mu + K * R(\tilde{\psi})$. Noting that $\psi_n = K^{(n)} * \mu^{(n)} + K^{(n)} * R^{(n)}(\psi_n)$ with $R_i^{(n)}(u) = \langle u, \beta^{i,(n)}\rangle + \frac{(\sigma_i^{(n)})^2}{2}u_i^2$ and $\beta^{i,(n)} = (\beta_{1i}^{(n)}, \dots, \beta_{mi}^{(n)})^\top$, it suffices to show that

$$\lim_{n\to\infty}\|K^{(n)} * \mu^{(n)} + K^{(n)} * R^{(n)}(\psi_n) - K * \mu - K * R(\tilde{\psi})\|_{L^p([0,T])} = 0.$$

Using the properties (i)–(iv) combined to similar estimates to the proofs of previous sections, it is not difficult to see that this convergence is satisfied. The details are left to the reader. □

Consequently, we can now prove that the law of the Volterra square-root process depends continuously on the parameters.

COROLLARY 3.13. *Let* $(b, \beta, \sigma, K), (b^{(n)}, \beta^{(n)}, \sigma^{(n)}, K^{(n)})$ *be admissible parameters with the properties* (i)–(iv) *from Theorem* 3.12, *and* $b^{(n)} \longrightarrow b$. *Let* X *and* X^n *be the Volterra square-root processes with admissible parameters* (b, β, σ, K) *and* $(b^{(n)}, \beta^{(n)}, \sigma^{(n)}, K^{(n)})$ *starting from the same initial state. Then the law of* $X^{(n)}$ *on* $C(\mathbb{R}_+; \mathbb{R}^m_+)$ *converges weakly to that of* X.

PROOF. For $\mu \in \mathcal{M}_{lf}^-$, let ψ and ψ_n be the corresponding unique solutions of (3.6). Let $R^{(n)}$ be the same as in the proof of Theorem 3.12, where we implicitly showed that $R^{(n)}(\psi_n) \to R(\psi)$ in $L^1([0,T])$. Then using $\psi_n \to \psi$ in $L^p([0,T])$ and the first identity in Corollary 3.11, we find that

$$\lim_{n\to\infty}\mathbb{E}\big[e^{\int_{[0,T]}\langle X^n_{t-s}, \mu(ds)\rangle}\big] = \mathbb{E}\big[e^{\int_{[0,T]}\langle X_{t-s}, \mu(ds)\rangle}\big].$$

In particular, letting $\mu(ds) = \sum_{j=1}^n u_j \delta_{t_j}(ds)$ with $u_1, \dots, u_n \in \mathbb{C}^m_-$, and $0 \leq t_1 < \cdots < t_n$ shows that the finite-dimensional distributions of X^n are convergent to those of X. Arguing as in the proof of [3], Lemma A.1, we also know that X^n is tight in $C(\mathbb{R}_+;\mathbb{R}_+^m)$. So, for any subsequence of X^n, it has a subsequence converging in law to X. This proves that the law of $X^{(n)}$ on $C(\mathbb{R}_+;\mathbb{R}_+^m)$ converges weakly to that of X. $\square$

3.4. *Differentiability in the initial condition.* In this section, we study continuity and differentiability of ψ with respect to the initial condition μ.

THEOREM 3.14. *Suppose there exist $p \geq 2$ and $\eta \in (0,1)$ such that $[K]_{\eta,p,T} < \infty$ for each $T > 0$. Then for each pair of $\mu, \nu \in \mathcal{M}_{lf}^-$, there exists a constant $C(T,p) > 0$ such that*

$$\|\psi(\cdot,\mu+\varepsilon\nu) - \psi(\cdot,\mu)\|_{L^p([0,T])} \leq C(T,p)\varepsilon.$$

PROOF. Note that $\Delta_\varepsilon(t) := \psi(t,\mu+\varepsilon\nu) - \psi(t,\mu)$ satisfies

$$\Delta_\varepsilon(t) = \varepsilon(K*\nu)(t) + \int_0^t K(t-s)\big(R\big(\psi(s,\mu+\varepsilon\nu)\big) - R\big(\psi(s,\mu)\big)\big)\,ds.$$

Hence, we obtain from (3.7) that

$$\big|\Delta_\varepsilon(t)\big|^2 \leq 2\varepsilon^2\big|K*\nu(t)\big|^2 + C(\mu,\nu)\int_0^t \|K(t-s)\|_2^2\big|\Delta_\varepsilon(s)\big|^2\,ds,$$

where

$$C(\mu,\nu) = 6C^2\Big(1 + \sup_{\varepsilon\in(0,1)}\|\psi(\cdot,\mu+\varepsilon\nu)\|_{L^2([0,T])}^2 + \|\psi(\cdot,\mu)\|_{L^2([0,T])}^2\Big)$$

is finite due to Theorem 3.10. Now let $\widetilde{R} \in L^{p/2}_{\mathrm{loc}}([0,T];\mathbb{R})$ be the resolvent of the second kind of $-C(\mu,\nu)\|\widetilde{K}(\cdot)\|_2^2 \in L^{p/2}_{\mathrm{loc}}([0,T];\mathbb{R})$. Using [31], Proposition 9.8.1, we find $\widetilde{R} \leq 0$. Hence, a Volterra analogue of the Gronwall inequality (see [2], Theorem A.2) gives

$$\big|\Delta_\varepsilon(t)\big|^2 \leq 2\varepsilon^2\big|K*\nu(t)\big|^2 + 2\varepsilon^2\int_0^t \big|\widetilde{R}(t-s)\big|\big|K*\nu(s)\big|^2\,ds.$$

This gives

$$\begin{aligned}
&\|\Delta_\varepsilon\|_{L^p([0,T])} \\
&\quad\leq 2\varepsilon\|K\|_{L^p([0,T])}|\nu|\big([0,T]\big) + 2\varepsilon\|\widetilde{R}\|_{L^{p/2}([0,T])}^{1/2}\|K\|_{L^2([0,T])}|\nu|\big([0,T]\big),
\end{aligned}$$

and hence proves the assertion. $\square$

Next, we prove differentiability in μ.

THEOREM 3.15. *Suppose there exist $p \geq 2$ and $\eta \in (0,1)$ such that $[K]_{\eta,p,T} < \infty$ for each $T > 0$. Then for all $\mu, \nu \in \mathcal{M}_{lf}^-$ the limit*

$$\lim_{\varepsilon\to 0}\frac{\psi(\cdot,\mu+\varepsilon\nu) - \psi(\cdot,\mu)}{\varepsilon} = D_\nu\psi(\cdot,\mu)$$

exists in $L^p([0,T];\mathbb{C}^m)$ for each $T > 0$. This limit satisfies

$$D_\nu\psi(t,\mu) = \int_{[0,t]} K(t-s)\nu(ds) + \int_0^t K(t-s)(DR)\big(\psi(s,\mu)\big)D_\nu\psi(s,\mu)\,ds, \tag{3.9}$$

where $DR(x) = \beta^\top + \frac{1}{2}\operatorname{diag}(\sigma_1^2 x_1,\ldots,\sigma_m^2 x_m)$.

PROOF. Note that $G(s) = (DR)(\psi(s,\mu))$ satisfies $G \in L^2_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{C}^{m\times m})$. Letting $p(t,x) = x$ and noting that $K * \nu \in L^2_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{C}^{m\times m})$, we may apply [3], Theorem B, to find a unique solution $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{C}^m)$ of $f = K * \nu + K * (Gp(\cdot, f))$. By definition, this solution is precisely the unique solution of (3.9), that is, $f = D_\nu\psi(\cdot,\mu)$.

It suffices to show that $\lim_{\varepsilon\to 0}\|\Delta_\varepsilon(\cdot) - D_\nu\psi(\cdot,\mu)\|_{L^p([0,T])} = 0$, where $\Delta_\varepsilon(t) := \varepsilon^{-1}(\psi(t,\mu+\varepsilon\nu) - \psi(t,\mu))$. To prove this, first note that

$$\begin{aligned}\Delta_\varepsilon(t) &= (K*\nu)(t)\\ &\quad + \int_0^t K(t-s)\biggl(\int_0^1 (DR)\bigl(\tau\psi(s,\mu+\varepsilon\nu) + (1-\tau)\psi(s,\mu)\bigr)\,d\tau\biggr)\Delta_\varepsilon(s)\,ds.\end{aligned}$$

Hence, by using $\|DR(x)\|_2 \le C(1+|x|)$ and $\|DR(u) - DR(v)\|_2 \le C|u-v|$, we obtain

$$\begin{aligned}&\bigl|\Delta_\varepsilon(t) - D_\nu\psi(t,\mu)\bigr|\\ &\quad\le \int_0^t \|K(t-s)\|_2 \int_0^1 \bigl\|(DR)\bigl(\tau\psi(s,\mu+\varepsilon\nu) + (1-\tau)\psi(s,\mu)\bigr)\\ &\qquad - (DR)\bigl(\psi(s,\mu)\bigr)\bigr\|_2\,d\tau\bigl|D_\nu\psi(s,\mu)\bigr|\,ds\\ &\qquad + \int_0^t \|K(t-s)\|_2 \int_0^1 \bigl\|(DR)\bigl(\tau\psi(s,\mu+\varepsilon\nu)\\ &\qquad + (1-\tau)\psi(s,\mu)\bigr)\bigr\|_2\,d\tau\bigl|\Delta_\varepsilon(s) - D_\nu\psi(s,\mu)\bigr|\,ds\\ &\quad\le C\int_0^t \|K(t-s)\|_2\bigl|\psi(s,\mu+\varepsilon\nu) - \psi(s,\mu)\bigr|\bigl|D_\nu\psi(s,\mu)\bigr|\,ds\\ &\qquad + C\int_0^t \|K(t-s)\|_2\bigl(1 + \bigl|\psi(s,\mu+\varepsilon\nu)\bigr| + \bigl|\psi(s,\mu)\bigr|\bigr)\bigl|\Delta_\varepsilon(s) - D_\nu\psi(s,\mu)\bigr|\,ds.\end{aligned}$$

After a short computation, we obtain

$$\begin{aligned}&\bigl|\Delta_\varepsilon(t) - D_\nu\psi(t,\mu)\bigr|^2\\ &\quad\le \|\psi(\cdot,\mu+\varepsilon\nu) - \psi(\cdot,\mu)\|^2_{L^2([0,T])} f(t) + \int_0^t k(t-s)\bigl|\Delta_\varepsilon(s) - D_\nu\psi(s,\mu)\bigr|^2\,ds\end{aligned}$$

with

$$k(t) = 8C\Bigl(1 + \|\psi(\cdot,\mu)\|^2_{L^2([0,T])} + \sup_{\varepsilon\in(0,1)}\|\psi(\cdot,\mu+\varepsilon\nu)\|^2_{L^2([0,T])}\Bigr)\|K(t)\|_2^2$$

and $f(t) = 2C\int_0^t \|K(t-s)\|_2^2|D_\nu\psi(s,\mu)|^2\,ds$. Let $\widetilde{R} \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{R})$ be the resolvent of the second kind for k. Arguing as in the proof of Theorem 3.14 gives

$$\begin{aligned}\bigl|\Delta_\varepsilon(t) - D_\nu\psi(t,\mu)\bigr|^2 &\le \|\psi(\cdot,\mu+\varepsilon\nu) - \psi(\cdot,\mu)\|^2_{L^2([0,T])} f(t)\\ &\quad + \|\psi(\cdot,\mu+\varepsilon\nu) - \psi(\cdot,\mu)\|^2_{L^2([0,T])}\int_0^t \bigl|\widetilde{R}(t-s)\bigr| f(s)\,ds.\end{aligned}$$

Hence, we obtain

$$\begin{aligned}&\|\Delta_\varepsilon - D_\nu\psi(\cdot,\mu)\|_{L^p([0,T])}\\ &\quad\le C\|\psi(\cdot,\mu+\varepsilon\nu) - \psi(\cdot,\mu)\|_{L^2([0,T])}\\ &\qquad\cdot\bigl(\|K\|_{L^p([0,T])}\|D_\nu\psi(\cdot,\mu)\|_{L^2([0,T])} + \|\widetilde{R}\|^{1/2}_{L^1([0,T])}\|f\|^{1/2}_{L^{p/2}([0,T])}\bigr).\end{aligned}$$

The assertion now follows from Theorem 3.14. $\square$

REMARK 3.16. By inspection of the proof, one can see that the above results hold true for $\mu = u\delta_0$ with $u \in \mathbb{C}_-^m$ without assuming $[K]_{\eta,p,T} < \infty$.

COROLLARY 3.17. *Let X be the Volterra square-root process with admissible parameters (b, β, σ, K) and initial state $x_0 \in \mathbb{R}_+^m$. Then*

$$\mathbb{E}[X_t] = \left(I_m + \int_0^t (E_{\beta^\top}(s))^\top \beta\, ds\right)x_0 + \left(\int_0^t (E_{\beta^\top}(s))^\top\, ds\right)b.$$

PROOF. Taking $\mu = u\delta_0$ with $u \in \mathbb{C}_-^m$ and then noting $\psi(s, 0) = 0$, we find for $D_u\psi(t, 0) := D_\mu\psi(t, 0)$ that $D_u\psi(t, 0) = K(t)u + \int_0^t K(t-s)\beta^\top D_u\psi(s, 0)\, ds$, where we have used $DR(0) = \beta^\top$. Solving this linear Volterra equation gives $D_\mu\psi(t, 0) = (E_{\beta^\top} * \mu)(t) = E_{\beta^\top}(t)u$. Hence, we obtain

$$\begin{aligned}
\mathbb{E}[\langle u, X_t\rangle] &= \frac{d}{d\varepsilon}\bigg|_{\varepsilon=0} \mathbb{E}[e^{\langle X_t, \varepsilon u\rangle}] \\
&= \langle u, x_0\rangle + \int_0^t \langle x_0, (DR)(0)D_u\psi(s, 0)\rangle\, ds + \int_0^t \langle b, D_u\psi(s, 0)\rangle\, ds \\
&= \langle u, x_0\rangle + \int_0^t \langle (E_{\beta^\top}(s))^\top \beta x_0, u\rangle\, ds + \int_0^t \langle (E_{\beta^\top}(s))^\top b, u\rangle\, ds.
\end{aligned}$$

Since u is arbitrary, the assertion is proved. □

4. Uniform moment and Hölder bounds.

4.1. *Uniform moment bounds.* Let X be the Volterra square-root process with admissible parameters (b, β, σ, K) and initial state x_0. In this section, we prove uniform in time moment bounds on the process X, which extend [2], Lemma 3.1, where similar bounds have been obtained on finite-time intervals $[0, T]$.

First, observe that after taking expectations in (1.1) we arrive at a convolution equation for $\mathbb{E}[X_t]$, which has the unique solution

$$\mathbb{E}[X_t] = \left(I_m - \int_0^t R_\beta(s)\, ds\right)x_0 + \left(\int_0^t E_\beta(s)\, ds\right)b, \tag{4.1}$$

where R_β, E_β are respectively defined by (2.1) and (2.2) with $B = \beta$, compare with [2], Lemma 4.1.

LEMMA 4.1. *It holds that $R_\beta = (E_{\beta^\top})^\top(-\beta)$ and $E_\beta = (E_{\beta^\top})^\top$.*

PROOF. Using Corollary 3.17 and (4.1) for $x_0 = 0$ yields $(\int_0^t E_\beta(s)\, ds)b = (\int_0^t (E_{\beta^\top}(s))^\top\, ds)b$ for each $b \in \mathbb{R}_+^m$. Since b is arbitrary and the cone $\mathbb{R}_+^m$ is generating (i.e., $\mathbb{R}^m = \mathbb{R}_+^m - \mathbb{R}_+^m$), we conclude $\int_0^t E_\beta(s)\, ds = \int_0^t (E_{\beta^\top}(s))^\top\, ds$. Taking now the derivative in t and noting that the integrands are continuous on $(0, \infty)$ because K is continuous, yields $E_\beta(t) = (E_{\beta^\top}(t))^\top$ for all $t > 0$. This proves the second identity. The first identity follows from $(E_{\beta^\top}(t))^\top(-\beta) = E_\beta(t)(-\beta) = R_\beta(t)$. This proves the assertion. □

It seems natural that the above relations may also be derived directly from the definition of R_β, E_β. However, we have not succeeded in this way. The above relations provide the following observation used throughout this section.

REMARK 4.2. Let $p \in [1, \infty]$. One has $E_\beta \in L^p(\mathbb{R}_+; \mathbb{R}^{m\times m})$ if and only if $E_{\beta^\top} \in L^p(\mathbb{R}_+; \mathbb{R}^{m\times m})$.

The next lemma shows that integrability of E_β is sufficient for the boundedness of the first moment.

LEMMA 4.3. *For each $v \in \mathbb{R}^m_+$, one has*

$$\left(I_m - \int_0^t R_\beta(s)\,ds\right)v \in \mathbb{R}^m_+ \quad \text{and} \quad \left(\int_0^t E_\beta(s)\,ds\right)v \in \mathbb{R}^m_+. \tag{4.2}$$

Moreover, if $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m})$, then

$$\lim_{t\to\infty} \mathbb{E}[X_t] = \left(I_m - \int_0^\infty R_\beta(s)\,ds\right)x_0 + \left(\int_0^\infty E_\beta(s)\,ds\right)b$$

and $\sup_{t\ge 0} \mathbb{E}[|X_t|] \le C_\beta \max\{|x_0|, |b|\}$ holds for some constant C_β.

PROOF. Note that (4.1) holds for all $x_0, b \in \mathbb{R}^m_+$. Taking $b = 0$ shows that $I_m - \int_0^t R_\beta(s)\,ds$ leaves $\mathbb{R}^m_+$ invariant. Taking $x_0 = 0$ shows that $\int_0^t E_\beta(s)\,ds$ leaves $\mathbb{R}^m_+$ invariant. If E_β is integrable, then using $E_\beta(-\beta) = R_\beta$, we find that also R_β is integrable, and hence we can pass to the limit $t \to \infty$ in (4.1). This proves the desired convergence of the first moment. The last assertion then follows from $\mathbb{E}[|X_t|] \le \mathbb{E}[\sum_{i=1}^m X_{i,t}] \le \sqrt{m}|\mathbb{E}[X_t]| \le \sqrt{m}(1 + \|R_\beta\|_{L^1})|x_0| + \|E_\beta\|_{L^1}|b|$. $\square$

REMARK 4.4. If $m = 1$ and $\beta < 0$, then using $E_\beta = R_\beta(-\beta)^{-1}$ combined with (4.2) implies that $0 \le \int_0^t R_\beta(s)\,ds \le 1$, and hence $0 \le \mathbb{E}[X_t] \le x_0 + \frac{b}{|\beta|}$.

To prove the uniform boundedness of higher order moments, we use [3], Lemma 2.5, to see that (1.1) is equivalent to

$$X_t = \left(I_m - \int_0^t R_\beta(s)\,ds\right)x_0 + \left(\int_0^t E_\beta(s)\,ds\right)b + \int_0^t E_\beta(t-s)\sigma(X_s)\,dB_s. \tag{4.3}$$

Based on this equivalent representation for the Volterra square-root process, we obtain the following simple observation.

LEMMA 4.5. *Suppose that $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$. Then*

$$\sup_{t\ge 0} \mathbb{E}\big[|X_t|^p\big] < \infty$$

holds for each $p \ge 2$.

The proof of this lemma is postponed to the Appendix.

4.2. *Uniform bound on Hölder increments.* It follows from [3], Lemma 2.4, that the Volterra square-root process has Hölder continuous sample paths. Below we recall their key estimate.

PROPOSITION 4.6. *For each $T > 0$ and $p \ge 2$, there exists a constant $C(T, p) > 0$ such that for all $s, t \in [0, T]$ with $0 \le t - s \le 1$ one has $\mathbb{E}[|X_t - X_s|^p] \le C(T, p)(t - s)^{\gamma p/2}$.*

To prove the existence of a stationary process and absolute continuity of the limiting distribution, we need a similar estimate but with a constant $C(T)$ independent of T. For this purpose, we need a stronger assumption then condition (v) in the definition of admissible parameters. Namely, we impose the following condition on K:

(K) There exists a constant $C_3 > 0$ such that

$$\int_0^\infty \left|K_i(r+h) - K_i(r)\right|^2 dr \le C_3 h^\gamma$$

holds for all $i = 1, \dots, m$, $h \in (0, 1]$.

Note that this condition is satisfied for $K_i(t) = t^{H-1/2}/\Gamma(H+1/2)e^{-\lambda t}$ with $\lambda \ge 0$, $H \in (0, 1/2)$, where $\gamma = 2H$. Also, under additional conditions on β we may obtain $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$ as required in the subsequent proposition (compare with Section 5.3).

The next proposition gives the desired global Hölder estimate for the process.

PROPOSITION 4.7. *Assume $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$ and condition (K). Then there exists for each $p \ge 2$ a constant $C(p) > 0$ such that $\mathbb{E}[|X_t - X_s|^p] \le C(p)(t-s)^{\gamma p/2}$ holds for all $s, t \ge 0$ with $0 \le t - s \le 1$.*

The proof of this proposition is postponed to the Appendix.

5. Limiting distributions.

5.1. *Existence of limiting distributions.* Below under an integrability condition on E_β, we prove the existence of limiting distributions for the Volterra square-root process and, therefore, provide a mathematical justification of the mean-reversion property. As a first step, we prove the convergence of the Fourier–Laplace transform.

PROPOSITION 5.1. *Suppose there exists $\eta \in (0, 1)$ such that $[K]_{\eta,2,T} < \infty$ for each $T > 0$. Then the following assertions hold*:

(a) *Suppose that $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m})$. Let $\mu \in \mathcal{M}_{lf}^-$ be such that $|\mu|(\mathbb{R}_+) < \infty$ and $\mathrm{Im}(\mu) = 0$. Then $\psi \in L^1(\mathbb{R}_+; \mathbb{C}^m) \cap L^2(\mathbb{R}_+; \mathbb{C}^m)$, and*

$$\lim_{t\to\infty} \mathbb{E}\big[e^{\int_{[0,t]}\langle X_{t-s}, \mu(ds)\rangle}\big]$$

$$= \exp\left\{\langle x_0, \mu(\mathbb{R}_+)\rangle + \int_0^\infty \langle x_0, R(\psi(s,\mu))\rangle ds + \int_0^\infty \langle b, \psi(s,\mu)\rangle ds\right\} \tag{5.1}$$

$$= \exp\left\{\langle A(\beta, x_0, b), \mu(\mathbb{R}_+)\rangle + \sum_{i=1}^m \frac{\sigma_i^2}{2} A_i(\beta, x_0, b) \int_0^\infty \psi_i(s,\mu)^2 ds\right\}, \tag{5.2}$$

where $\psi(\cdot, \mu)$ denotes the unique solution of (3.6),

$$A(\beta, x_0, b) = \left(I_m - \int_0^\infty R_\beta(s)\, ds\right)x_0 + \left(\int_0^\infty E_\beta(s)\, ds\right)b, \tag{5.3}$$

and $A_i(\beta, x_0, b)$ denotes the ith component of the vector $A(\beta, x_0, b)$.

(b) *Suppose that $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$. Let $\mu \in \mathcal{M}_{lf}^-$ with $|\mu|(\mathbb{R}_+) < \infty$. Then $\psi \in L^1(\mathbb{R}_+; \mathbb{C}_-^m) \cap L^2(\mathbb{R}_+; \mathbb{C}_-^m)$, and the identities* (5.1) *and* (5.2) *still hold.*

PROOF. (a) Instead of Fourier–Laplace transform, let us first reformulate the affine formula for the Laplace transform. The latter formulation is more natural to exploit the nonnegativity of the process. Noting $-\mu$ is an $\mathbb{R}_+^m$-valued measure, we let $V(t,\mu) = -\psi(t,\mu) \in \mathbb{R}_+^m$ and $\widetilde{R}_i(x) := -R_i(-x) = \langle x, \beta^i\rangle - \frac{\sigma_i^2}{2}x_i^2$, where $i = 1, \dots, m$ and $x \in \mathbb{R}_+^m$. Then $V(t,\mu)$ satisfies the Riccati–Volterra equation

$$V_i(t,\mu) = -\int_{[0,t]} K_i(t-s)\mu_i(ds) + \int_0^t K_i(t-s)\widetilde{R}_i\big(V(s,\mu)\big)\,ds,$$

and, by Corollary 3.11, it holds

$$\begin{aligned}&\mathbb{E}\big[\mathrm{e}^{-\int_{[0,t]}\langle X_{t-s}, -\mu(ds)\rangle}\big]\\ &\quad= \exp\Big\{-\big\langle x_0, -\mu([0,t])\big\rangle - \int_0^t \big\langle x_0, \widetilde{R}\big(V(s,\mu)\big)\big\rangle\,ds - \int_0^t \big\langle b, V(s,\mu)\big\rangle\,ds\Big\},\end{aligned} \tag{5.4}$$

where $\widetilde{R} = (\widetilde{R}_1, \dots, \widetilde{R}_m)^\top$. Applying Jensen's inequality, we have

$$e^{-\int_{[0,t]}\langle \mathbb{E}[X_{t-s}], -\mu(ds)\rangle} \le \mathbb{E}\big[\mathrm{e}^{-\int_{[0,t]}\langle X_{t-s}, -\mu(ds)\rangle}\big],$$

and hence

$$\begin{aligned}&\big\langle x_0, -\mu([0,t])\big\rangle + \int_0^t \big\langle x_0, \widetilde{R}\big(V(s,\mu)\big)\big\rangle\,ds + \int_0^t \big\langle b, V(s,\mu)\big\rangle\,ds\\ &\quad\le \int_{[0,t]} \big\langle \mathbb{E}[X_{t-s}], -\mu(ds)\big\rangle\\ &\quad\le |\mu|(\mathbb{R}_+)\sup_{t\ge 0}\mathbb{E}\big[|X_t|\big]\\ &\quad\le |\mu|(\mathbb{R}_+)C_\beta \max\big\{|x_0|, |b|\big\}.\end{aligned}$$

Note that this inequality holds for all choices of $b, x_0 \in \mathbb{R}_+^m$. In particular, choosing $b = (1, \dots, 1)^\top$ and $x_0 = 0$ gives

$$\sum_{i=1}^m \int_0^t V_i(s,\mu)\,ds \le \sqrt{m}|\mu|(\mathbb{R}_+)C_\beta. \tag{5.5}$$

To estimate the integral involving $\widetilde{R}$, let us first note that the left-hand side of (5.4) is bounded by 1, which gives

$$\big\langle x_0, -\mu([0,t])\big\rangle + \int_0^t \big\langle x_0, \widetilde{R}\big(V(s,\mu)\big)\big\rangle\,ds + \int_0^t \big\langle b, V(s,\mu)\big\rangle\,ds \ge 0$$

for all $x_0, b \in \mathbb{R}_+^m$. For $b = 0$ and $x_0 = (1, \dots, 1)^\top$, we obtain

$$\begin{aligned}\sum_{i=1}^m \frac{\sigma_i^2}{2}\int_0^t V_i(s,\mu)^2\,ds &\le \sum_{i=1}^m \big(-\mu_i([0,t])\big) + \sum_{i=1}^m \int_0^t \big\langle V(s,\mu), \beta^i\big\rangle\,ds\\ &\le \sqrt{m}|\mu|(\mathbb{R}_+) + \sum_{i=1}^m |\beta^i| \int_0^t \big|V(s,\mu)\big|\,ds\\ &\le \sqrt{m}|\mu|(\mathbb{R}_+) + m\|\beta\|_{\mathrm{HS}}|\mu|(\mathbb{R}_+)C_\beta,\end{aligned}$$

where we have used (5.5). In view of (5.5) and the particular form of $\tilde{R}$, we obtain

$$\begin{aligned}\int_0^t \big|\widetilde{R}\big(V(s,\mu)\big)\big|\,ds &\le \sum_{i=1}^m \int_0^t \big|\big\langle V(s,\mu), \beta^i\big\rangle\big|\,ds + \sum_{i=1}^m \frac{\sigma_i^2}{2}\int_0^t V_i(s,\mu)^2\,ds\\ &\le \big(\sqrt{m} + 2m\|\beta\|_{\mathrm{HS}}C_\beta\big)|\mu|(\mathbb{R}_+).\end{aligned} \tag{5.6}$$

This estimate combined with (5.5) proves the convergence in (5.1). For the second identity in part (a), use the second identity from Corollary 3.11 to pass to the limit $t \to \infty$, that is, we show that:

(i) $\lim_{t\to\infty} \int_{[0,t]} \langle \mathbb{E}[X_{t-s}], \mu(ds)\rangle = \langle A(\beta, x_0, b), \mu(\mathbb{R}_+)\rangle$;
(ii) $\lim_{t\to\infty} \int_0^t \mathbb{E}[X_{i,t-s}] V_i(s,\mu)^2\, ds = A_i(\beta, x_0, b) \int_0^\infty V_i(s,\mu)^2\, ds$.

So, let $\varepsilon > 0$. Since $A(\beta, x_0, b) = \lim_{t\to\infty} \mathbb{E}[X_t]$ by Lemma 4.3, we find $t_0 > 0$ such that $|\mathbb{E}[X_{i,r}] - A_i(\beta, x_0, b)| < \varepsilon$ for all $r \geq t_0$ and $i = 1, \dots, m$. Then for all $t \geq 2t_0$ we obtain

$$
\begin{aligned}
&\left|\int_{[0,t]} \langle \mathbb{E}[X_{t-s}], \mu(ds)\rangle - \langle A(\beta, x_0, b), \mu(\mathbb{R}_+)\rangle\right| \\
&\quad \leq \left|\int_{[0,t/2]} \langle \mathbb{E}[X_{t-s}] - A(\beta, x_0, b), \mu(ds)\rangle\right| \\
&\qquad + \left|\int_{(t/2,t]} \langle \mathbb{E}[X_{t-s}] - A(\beta, x_0, b), \mu(ds)\rangle\right| + \left|\langle A(\beta, x_0, b), \mu\big((t,\infty)\big)\rangle\right| \\
&\quad \leq \varepsilon\sqrt{m}|\mu|(\mathbb{R}_+) + \Big(\sup_{s\geq 0} \mathbb{E}\big[|X_s|\big]\Big)|\mu|((t/2,t]) + 2\big|A(\beta, x_0, b)\big||\mu|\big((t/2,\infty)\big).
\end{aligned}
$$

Since $|\mu|(\mathbb{R}_+) < \infty$, we have $|\mu|((t/2,\infty)) \to 0$ as $t \to \infty$, which proves (i). For (ii), we have

$$
\begin{aligned}
&\left|\int_0^t \mathbb{E}[X_{i,t-s}] V_i(s,\mu)^2\, ds - A_i(\beta, x_0, b) \int_0^\infty V_i(s,\mu)^2\, ds\right| \\
&\quad \leq \int_0^t \big|\mathbb{E}[X_{i,t-s}] - A_i(\beta, x_0, b)\big| V_i(s,\mu)^2\, ds + A_i(\beta, x_0, b) \int_t^\infty V_i(s,\mu)^2\, ds.
\end{aligned}
$$

The second term tends to zero due to $\int_0^\infty V_i(s,\mu)^2\, ds < \infty$. For the first term, we have for all $t \geq 2t_0$,

$$
\begin{aligned}
&\int_0^t \big|\mathbb{E}[X_{i,t-s}] - A_i(\beta, x_0, b)\big| V_i(s,\mu)^2\, ds \\
&\quad = \int_0^{t/2} \big|\mathbb{E}[X_{i,t-s}] - A_i(\beta, x_0, b)\big| V_i(s,\mu)^2\, ds \\
&\qquad + \int_{t/2}^t \big|\mathbb{E}[X_{i,t-s}] - A_i(\beta, x_0, b)\big| V_i(s,\mu)^2\, ds \\
&\quad \leq \varepsilon \int_0^\infty V_i(s,\mu)^2\, ds + \Big(\sup_{s\geq 0} \mathbb{E}\big[|X_s|\big] + \big|A(\beta, x_0, b)\big|\Big) \int_0^\infty \mathbb{1}_{[t/2,t]}(s) V_i(s,\mu)^2\, ds.
\end{aligned}
$$

The dominated convergence theorem implies that the second term tends to zero as $t \to \infty$. Since ε is arbitrary, this proves (ii), and thus completes the proof of part (a).

Part (b) can be shown in a similar way, since Theorem 3.10 applied for $T = \infty$ still provides the desired integrability $\int_0^\infty (|\psi(t,\mu)| + |\psi(t,\mu)|^2)\, dt < \infty$. □

REMARK 5.2. If we choose $\mu(ds) = u\delta_0(ds)$, then the statements of Proposition 5.1 and the estimates established in the above proof still hold even if we drop the condition $[K]_{\eta,2,T} < \infty$. Essentially, this condition was used to ensure the existence of $\psi(t,\mu)$ and the applicability of Theorem 3.10 and Corollary 3.11. However, for the particular choice of $\mu(ds) = u\delta_0(ds)$, we can work directly with (1.2) and $\psi(\cdot, u, 0)$ instead of the extension (3.6) and $\psi(\cdot, u\delta_0)$, and then apply Lemma 3.1. The above proof still works in this case with some obvious adaptations.

From the convergence of the Fourier–Laplace transform, we can now deduce convergence toward limiting distributions. The following is our main result on limiting distributions for the Volterra square-root process. In contrast to the classical case, the limiting distribution now also involves the initial state of the process. For this purpose, we define

$$\mathcal{N} = \left\{ v \in \mathbb{R}^m : \int_0^\infty R_\beta(t) v \, dt = v \right\}, \tag{5.7}$$

and let $\mathcal{N}^\perp$ be the orthogonal complement of $\mathcal{N}$. Denote by P the orthogonal projection operator onto $\mathcal{N}^\perp$, that is, $\ker(P) = \mathcal{N}$ and $\operatorname{ran}(P) = \mathcal{N}^\perp$.

THEOREM 5.3. *Suppose that $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m})$. Then the law of X_t converges for $t \to \infty$ weakly to a limiting distribution π_{x_0}, whose Fourier–Laplace transform is for $u \in \mathbb{C}_-^m$ with* $\operatorname{Im}(u) = 0$ *given by*

$$\begin{aligned}
&\int_{\mathbb{R}_+^m} e^{\langle u, y\rangle} \pi_{x_0}(dy) \\
&\quad = \exp\left\{ \langle x_0, u\rangle + \int_0^\infty \langle x_0, R(\psi(s, u\delta_0))\rangle \, ds + \int_0^\infty \langle b, \psi(s, u\delta_0)\rangle \, ds \right\}.
\end{aligned}$$

Moreover, π_{x_0} has finite first moment and satisfies

$$\pi_{x_0} = \pi_{Px_0} = \pi_0 * \pi_{Px_0}^{b=0}, \tag{5.8}$$

where $*$ denotes the usual convolution of probability measures on $\mathbb{R}_+^m$, and $\pi_{Px_0}^{b=0}$ is the limiting distribution of the Volterra square-root process with admissible parameters $(b = 0, \beta, \sigma, K)$ and initial state Px_0. Finally, if, in addition, $E_\beta \in L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$, then the Fourier–Laplace transform representation for π_{x_0} can be extended to all $u \in \mathbb{C}_-^m$.

PROOF. Consider $u \in \mathbb{C}_-^m$ with $\operatorname{Im}(u) = 0$. According to Remark 5.2, if we take $\mu(ds) = u\delta_0(ds)$, then it holds

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}[e^{\langle X_t, u\rangle}] \\
&\quad = \exp\left\{ \langle x_0, u\rangle + \int_0^\infty \langle x_0, R(\psi(s, u\delta_0))\rangle \, ds + \int_0^\infty \langle b, \psi(s, u\delta_0)\rangle \, ds \right\}.
\end{aligned}$$

Moreover, the estimates (5.5) and (5.6) hold with $|\mu|(\mathbb{R}_+) = |u|$, showing that the right-hand side is continuous at $u = 0$. Hence, using Lévy's continuity theorem for Laplace transforms proves that X_t converges weakly to some distribution π_{x_0} and that the desired formula for the Laplace transform of π_{x_0} holds. The extension to the Fourier–Laplace transform with $u \in \mathbb{C}_-^m$ follows now from Proposition 5.1(b). An application of the lemma of Fatou shows that the limit distribution π_{x_0} has finite first moment, that is, $\int_{\mathbb{R}_+^m} |x| \pi_{x_0}(dx) \le \sup_{t\ge 0} \mathbb{E}[|X_t|] < \infty$. It remains to prove (5.8). For this purpose, we use the second identity from Proposition 5.1, that is,

$$\begin{aligned}
&\lim_{t\to\infty} \mathbb{E}[e^{\langle X_t, u\rangle}] \\
&\quad = \exp\left\{ \langle A(\beta, x_0, b), u\rangle + \sum_{i=1}^m \frac{\sigma_i^2}{2} A_i(\beta, x_0, b) \int_0^\infty \psi_i(s, u\delta_0)^2 \, ds \right\},
\end{aligned}$$

where $A(\beta, x_0, b)$ is defined in (5.3). Then noting that

$$A(\beta, x_0, b) = A(\beta, Px_0, b) = A(\beta, 0, b) + A(\beta, Px_0, 0)$$

readily yields (5.8) due to uniqueness of the Laplace transform. This completes the proof. $\square$

5.2. *Stationary process.* Next, we construct the stationary process.

THEOREM 5.4. *Suppose that $E_\beta \in L^1(\mathbb{R}_+;\mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+;\mathbb{R}^{m\times m})$, condition (K) holds and there exists $\eta \in (0,1)$ such that $[K]_{\eta,2,T} < \infty$ for each $T > 0$. Then there exists a stationary process X^{stat} with continuous sample paths such that the following assertions hold*:

(a) *It holds that $(X_{t+h})_{t\ge 0} \Rightarrow (X_t^{\text{stat}})_{t\ge 0}$ weakly on $C(\mathbb{R}_+;\mathbb{R}_+^m)$ as $h\to\infty$.*
(b) *The finite-dimensional distributions of X^{stat} are determined by*

$$\begin{aligned}
&\mathbb{E}\big[e^{\sum_{j=1}^n \langle X_{t_j}^{\text{stat}}, u_j\rangle}\big] \\
&\quad = \exp\Bigg\{\sum_{j=1}^n \langle A(\beta,x_0,b),u_j\rangle + \sum_{i=1}^m \frac{\sigma_i^2}{2} A_i(\beta,x_0,b)\int_0^\infty \psi_i(s)^2\,ds\Bigg\},
\end{aligned}$$

where $A(\beta,x_0,b)$ is defined by (5.3), $\psi(\cdot)=\psi(\cdot,\mu_{t_1,\dots,t_n})$ denotes the unique solution of (3.6) with $\mu_{t_1,\dots,t_n}(ds)=\sum_{j=1}^n u_j\delta_{t_n-t_j}(ds)$, $n\in\mathbb{N}$, $u_1,\dots,u_n\in\mathbb{C}_-^m$ and $0\le t_1<\cdots<t_n$.

PROOF. Choose $p\ge 2$ sufficiently large so that $\gamma p > 2$. By Proposition 4.7, we find some constant $C(p)>0$ such that $\mathbb{E}[|X_t - X_s|^p] \le C(p)(t-s)^{\gamma p/2}$ holds for all $t,s\ge 0$ with $0\le t-s\le 1$. Define for $h\ge 0$ the process X^h by $X_t^h = X_{h+t}$, where $t\ge 0$. Then X^h has continuous sample paths and satisfies $\sup_{h\ge 0}\mathbb{E}[|X_t^h - X_s^h|^p]\le C(p)(t-s)^{\gamma p/2}$ for $0\le t-s\le 1$. Applying the Kolmogorov tightness criterion (see, e.g., [36], Corollary 16.9) shows that $(X^h)_{h\ge 0}$ is tight on $C(\mathbb{R}_+;\mathbb{R}_+^m)$.

Hence, we conclude that along a sequence $h_k\uparrow\infty$, X^{h_k} converges in law to some continuous process X^{stat}. Take $n\in\mathbb{N}$ and let $0\le t_1<\cdots<t_n$. Applying Proposition 5.1 for the particular choice $\mu_{t_1,\dots,t_n}(ds)=\sum_{j=1}^n u_j\delta_{t_n-t_j}(ds)$, where $u_1,\dots,u_m\in\mathbb{C}_-^m$, we find that, for all $h\ge 0$,

$$\begin{aligned}
\mathbb{E}\big[e^{\sum_{j=1}^n\langle X_{t_j+h}^{\text{stat}},u_j\rangle}\big] &= \lim_{k\to\infty}\mathbb{E}\big[e^{\sum_{j=1}^n\langle X_{t_j+h}^{h_k},u_j\rangle}\big] \\
&= \lim_{k\to\infty}\mathbb{E}\big[e^{\sum_{j=1}^n\langle X_{h_k+h+t_j},u_j\rangle}\big] \qquad (5.9)\\
&= \lim_{k\to\infty}\mathbb{E}\big[e^{\int_{[0,h_k+h+t_n]}\langle X_{h_k+h+t_n-s},\mu_{t_1,\dots,t_n}(ds)\rangle}\big] \\
&= \exp\Bigg\{\sum_{j=1}^n\langle x_0,u_j\rangle + \int_0^\infty \big\langle x_0, R\big(\psi(s,\mu_{t_1,\dots,t_n})\big)\big\rangle\,ds \\
&\qquad + \int_0^\infty \big\langle b,\psi(s,\mu_{t_1,\dots,t_n})\big\rangle\,ds\Bigg\}. \qquad (5.10)
\end{aligned}$$

In view of (5.2), (5.9) and (5.10), the desired formula of the Fourier transform is proved. Since $\{h_k\}$ is arbitrary and (5.10) is independent of $\{h_k\}$, it is standard to verify the weak convergence in (a). The assertion is proved. □

A direct consequence of Theorem 5.4 is that $X_t \longrightarrow \pi_{x_0}$ weakly as $t\to\infty$, and X_t^{stat} has distribution π_{x_0} for each $t\ge 0$. In the next statement, we compute the moments, covariance structure and autocovariance function of the stationary process.

COROLLARY 5.5. *Under the same conditions as in Theorem 5.4, the stationary process X^{stat} satisfies $\mathbb{E}[|X_t^{\text{stat}}|^p]=\int_{\mathbb{R}_+^m}|x|^p\pi_{x_0}(dx)<\infty$ for each $p\ge 2$. Moreover, its first moment* *is given by*

$$\mathbb{E}[X_t^{\mathrm{stat}}] = \Big(I_m - \int_0^\infty R_\beta(s)\,ds\Big)x_0 + \Big(\int_0^\infty E_\beta(s)\,ds\Big)b,$$

while its autocovariance function is, for $0 \le s \le t$, *given by*

$$\begin{aligned}&\operatorname{cov}\big(X_t^{\mathrm{stat}}, X_s^{\mathrm{stat}}\big)\\ &\quad= \int_0^\infty E_\beta(t-s+u)\sigma\big(A(\beta, x_0, b)\big)\sigma\big(A(\beta, x_0, b)\big)^\top E_\beta(u)^\top\,du.\end{aligned}$$

PROOF. Since $\sup_{t\ge 0}\mathbb{E}[|X_t|^p] < \infty$ holds for each $p \ge 2$ and $X_t \longrightarrow \pi_{x_0}$ weakly, the lemma of Fatou implies that π_{x_0} has all finite moments. Since X^{stat} is stationary, we conclude the first assertion. For the first moment formula, we note that $\mathbb{E}[X_t] \longrightarrow A(\beta, x_0, b)$ as $t \to \infty$. Since $\sup_{t\ge 0}\mathbb{E}[|X_t|^2] < \infty$, we easily conclude that $\lim_{t\to\infty}\mathbb{E}[X_t] = \int_{\mathbb{R}_+^m} x\pi_{x_0}(dx) = \mathbb{E}[X_t^{\mathrm{stat}}]$. This proves the desired first moment formula for the stationary process. Noting that $X_t - \mathbb{E}[X_t] = \int_0^s E_\beta(t-u)\sigma(X_u)\,dB_u + \int_s^t E_\beta(t-u)\sigma(X_u)\,dB_u$ and $X_s - \mathbb{E}[X_s] = \int_0^s E_\beta(s-u)\sigma(X_u)\,dB_u$, we find that the autocovariance function for X is given by

$$\operatorname{cov}(X_t, X_s) = \int_0^s E_\beta(t-u)\sigma\big(\mathbb{E}[X_u]\big)\sigma\big(\mathbb{E}[X_u]\big)^\top E_\beta(s-u)^\top\,du,$$

where we have used the particular form of $\sigma(x)$ so that $\mathbb{E}[\sigma(X_u)\sigma(X_u)^\top] = \sigma(\mathbb{E}[X_u]) \times \sigma(\mathbb{E}[X_u])^\top$. Thus, the autocovariance function of the stationary process is given by

$$\begin{aligned}&\operatorname{cov}\big(X_t^{\mathrm{stat}}, X_s^{\mathrm{stat}}\big)\\ &\quad= \lim_{h\to\infty}\operatorname{cov}(X_{t+h}, X_{s+h})\\ &\quad= \lim_{h\to\infty}\int_0^{s+h} E_\beta(t+h-u)\sigma\big(\mathbb{E}[X_u]\big)\sigma\big(\mathbb{E}[X_u]\big)^\top E_\beta(s+h-u)^\top\,du\\ &\quad= \lim_{h\to\infty}\int_{-h}^{s} E_\beta(t-u)\sigma\big(\mathbb{E}[X_{u+h}]\big)\sigma\big(\mathbb{E}[X_{u+h}]\big)^\top E_\beta(s-u)^\top\,du\\ &\quad= \int_{-\infty}^{s} E_\beta(t-u)\sigma\big(A(\beta, x_0, b)\big)\sigma\big(A(\beta, x_0, b)\big)^\top E_\beta(s-u)^\top\,du\\ &\quad= \int_0^\infty E_\beta(t-s+u))\sigma(A(\beta, x_0, b)\sigma\big(A(\beta, x_0, b)\big)^\top E_\beta(u)^\top\,du,\end{aligned}$$

which proves the assertion. □

The particular form of the Laplace transform for the limiting distribution and the stationary process X^{stat} give the following characterization for the independence on the initial condition x_0.

COROLLARY 5.6. *Suppose* $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$, *then the following are equivalent*:

(i) *The stationary process* X^{stat} *is independent of* x_0;
(ii) *The limiting distribution* π_{x_0} *is independent of* x_0;
(iii) *The function* $x_0 \longmapsto \int_{\mathbb{R}_+^m} x\pi_{x_0}(dx)$ *is constant*;
(iv) $\int_0^\infty R_\beta(t)\,dt = I_m$.

PROOF. Since π_{x_0} is the law of X_t^{stat}, clearly (i) implies (ii), and (ii) implies (iii). Suppose that (iii) holds. Using the first moment for the stationary process, we have $A(\beta, x_0, b) = \int_{\mathbb{R}_+^m} y\pi_{x_0}(dy) = \int_{\mathbb{R}_+^m} y\pi_0(dy) = A(\beta, 0, b)$. The particular form of $A(\beta, x_0, b)$ readily yields (iv). Finally, suppose that (iv) is satisfied. Then $A(\beta, x_0, b) = \int_0^\infty E_\beta(t)b\,dt$ is independent of x_0, and hence the Laplace transform for the stationary process implies that X^{stat} is independent of x_0, that is, (i) holds. $\square$

Finally, we discuss implications and also a sufficient condition for (iv).

THEOREM 5.7. *Suppose that $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m})$. Then the following assertions hold*:

(a) *If $\int_0^\infty R_\beta(t)\,dt = I_m$, then $K\beta$ is not integrable on $\mathbb{R}_+$*;
(b) *If $\beta^\top\beta$ has only strictly positive eigenvalues and the function*

$$K_*(t) = \min_{j=1,\dots,m} K_j(t)$$

is not integrable over $\mathbb{R}_+$, then $\int_0^\infty R_\beta(t)\,dt = I_m$.

PROOF. (a) Assume $\int_0^\infty R_\beta(t)\,dt = I_m$. Suppose $K_\beta := K(-\beta)$ is integrable on $\mathbb{R}_+$. The integrability of R_β and K_β on $\mathbb{R}_+$ implies that $\widehat{R}_\beta(z)$, $\widehat{K}_\beta(z)$ are well-defined for all $\operatorname{Re}(z) \geq 0$. Using the Paley–Wiener theorem (see [31], Chapter 2) shows that $\det(I_m + \widehat{K}_\beta(z)) \neq 0$. Solving (2.1) (for the Laplace transforms) yields

$$\widehat{R}_\beta(z) = \widehat{K_\beta}(z)\big(I_m + \widehat{K_\beta}(z)\big)^{-1}, \qquad \operatorname{Re}(z) \geq 0. \tag{5.11}$$

Evaluating at $z = 0$ gives

$$I_m = \int_0^\infty R_\beta(t)\,dt = \int_0^\infty K_\beta(t)\,dt\left(I_m + \int_0^\infty K_\beta(t)\,dt\right)^{-1}.$$

Hence, $I_m + \int_0^\infty K_\beta(t)\,dt = \int_0^\infty K_\beta(t)\,dt$, which is impossible. So, $K\beta$ is not integrable on $\mathbb{R}_+$.

(b) Define $K_\beta^\lambda(t) = e^{-\lambda t}K_\beta(t)$ for $\lambda > 0$. Since $K_1, \dots, K_m$ are nonnegative and nonincreasing, it follows that $e^{-\lambda t}K_1(t), \dots, e^{-\lambda t}K_m(t)$ are integrable on $\mathbb{R}_+$. Hence, also K_β^λ is integrable on $\mathbb{R}_+$. Let $R_\beta^\lambda(t) := e^{-\lambda t}R_\beta(t)$. By the definition in (2.1), it is easy to verify that R_β^λ is the resolvent of the second kind of K_β^λ. As mentioned in Remark 2.2, we have $E_\beta(-\beta) = R_\beta$, so $R_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m})$. Similar to (5.11), it holds

$$\begin{aligned}\int_0^\infty e^{-\lambda t}R_\beta(t)\,dt &= \int_0^\infty e^{-\lambda t}K_\beta(t)\,dt\left(I_m + \int_0^\infty e^{-\lambda t}K_\beta(t)\,dt\right)^{-1}\\ &= I_m - \left(I_m + \int_0^\infty e^{-\lambda t}K_\beta(t)\,dt\right)^{-1}.\end{aligned}$$

Since R_β is integrable, the left-hand side converges to $\int_0^\infty R_\beta(t)\,dt$ when $\lambda \searrow 0$. Thus, it suffices to show that

$$\lim_{\lambda\searrow 0}\left(I_m + \int_0^\infty e^{-\lambda t}K_\beta(t)\,dt\right)^{-1} = 0. \tag{5.12}$$

Note that

$$\begin{aligned}&\left\|\left(\int_0^\infty e^{-\lambda t}K_\beta(t)\,dt\right)^{-1}\right\|_2^{-2}\\ &\quad= \inf_{|v|=1}\left\langle\left(\int_0^\infty e^{-\lambda t}K_\beta(t)\,dt\right)v, \left(\int_0^\infty e^{-\lambda t}K_\beta(t)\,dt\right)v\right\rangle\end{aligned}$$

$$= \inf_{|v|=1} \int_0^\infty \int_0^\infty e^{-\lambda(t+s)} \langle K(t)\beta v, K(s)\beta v\rangle \, ds\, dt$$

$$\geq \inf_{|v|=1} \int_0^\infty \int_0^\infty e^{-\lambda(t+s)} K_*(t) K_*(s) \langle v, \beta^\top \beta v\rangle \, ds\, dt$$

$$\geq \lambda_{\min}(\beta^\top\beta) \left(\int_0^\infty e^{-\lambda t} K_*(t)\, dt \right)^2,$$

where $\lambda_{\min}(\beta^\top\beta) > 0$ denotes the smallest eigenvalue of $\beta^\top\beta$. We thus obtain $\|(\int_0^\infty e^{-\lambda t} \times K_\beta(t)\, dt)^{-1}\|_2 \to 0$ as $\lambda \searrow 0$. So,

$$\lim_{\lambda \searrow 0} \left\| \int_0^\infty e^{-\lambda t} K_\beta(t)\, dt \right\|_2 \geq \lim_{\lambda \searrow 0} \left\| \left(\int_0^\infty e^{-\lambda t} K_\beta(t)\, dt \right)^{-1} \right\|_2^{-1} = \infty,$$

which easily implies (5.12). The assertion is proved. $\square$

5.3. *Sufficient conditions and examples.* Next, we provide some examples for admissible parameters (b, β, σ, K) for which our results are applicable. Here, we focus first on completely monotone and then integrable kernels.

By an abuse of notation, let $\langle \cdot, \cdot \rangle$ denote *the inner product* on $\mathbb{C}^m$, that is, $\langle v, w\rangle = \sum_{j=1}^m v_i \overline{w}_j$. For a matrix $A \in \mathbb{C}^{m\times m}$, we write $A \succeq 0$ if $\langle v, Av\rangle \geq 0$ for all $v \in \mathbb{C}^m$, and write $A \succ 0$ if $\langle v, Av\rangle > 0$ for all nonzero $v \in \mathbb{C}^m$. For another matrix $B \in \mathbb{C}^{m\times m}$, we write $A \succeq B$ if $A - B \succeq 0$, and $A \succ B$ if $A - B \succ 0$. The notation "$\preceq$" and "$\prec$" are similarly defined. A kernel $k \in L^1_{\text{loc}}(\mathbb{R}_+; \mathbb{C}^{m\times m})$ is called *completely monotone* with respect to the order induced by $\langle \cdot, \cdot \rangle$, if it is smooth on $(0, \infty)$ and satisfies

$$(-1)^n \left(\frac{d}{dt} \right)^n k(t) \succeq 0, \quad \forall n \in \mathbb{N}.$$

Note that in the one-dimensional case the above definition of complete monotonicity reduces to the classical one. For additional results on completely monotone functions on $\mathbb{C}^{m\times m}$, we refer to [31], Chapter 5. We are now ready to provide a sufficient condition on how examples based on complete monotonicity can be constructed.

PROPOSITION 5.8. *Let $K = \widetilde{K} I_m$ with $\widetilde{K} \in L^2_{\text{loc}}(\mathbb{R}_+; \mathbb{R})$ being completely monotone. Let $\beta \in \mathbb{R}^{m\times m}$ be symmetric with only strictly negative eigenvalues. Define R_β, E_β by* (2.1) *and* (2.2), *respectively. Then $R_\beta, E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$ are completely monotone.*

PROOF. It is clear that K satisfies conditions (iv)–(vi) from the definition of admissible parameters given in the Introduction. Observe that $\langle v, (-\beta)v\rangle \geq 0$ for all $v \in \mathbb{C}^m$. Hence,

$$\langle v, (-1)^n K^{(n)}(t)(-\beta)v\rangle = (-1)^n \widetilde{K}^{(n)}(t) \langle v, (-\beta)v\rangle \geq 0$$

shows that also $-K\beta$ is completely monotone. Thus, by [31], Chapter 5, Theorem 3.1, R_β is completely monotone and integrable on $\mathbb{R}_+$. Hence, $0 \preceq R_\beta(t) \preceq R_\beta(t_0)$ for $t \geq t_0 > 0$. Using $\beta^\top = \beta$ so that $R_\beta^\top = R_\beta$, we find that if $t \geq t_0 > 0$, then for $v \in \mathbb{C}^m$,

$$\begin{aligned}
|R_\beta(t)v|^2 &= \langle R_\beta(t)^{1/2} v, R_\beta(t) R_\beta(t)^{1/2} v\rangle \\
&\leq \langle R_\beta(t)^{1/2} v, R_\beta(t_0) R_\beta(t)^{1/2} v\rangle \\
&\leq |v|^2 \|R_\beta(t_0)\|_2 \|R_\beta(t)^{1/2}\|_2^2 \\
&\leq |v|^2 \|R_\beta(t_0)\|_2 \|R_\beta(t)\|_2,
\end{aligned}$$

where we have used $\|R_\beta(t)^{1/2}\|_2^2 = \sup_{|v|=1}\langle v, R_\beta(t)v\rangle \le \|R_\beta\|_2$. This shows $\|R_\beta(t)\|_2 \le \|R_\beta(t_0)\|_2$ for $t \ge t_0$. So,

$$\int_0^\infty \|R_\beta(t)\|_2^2\,dt \le \int_0^{t_0} \|R_\beta(t)\|_2^2\,dt + \|R_\beta(t_0)\|_2 \int_{t_0}^\infty \|R_\beta(t)\|_2\,dt.$$

Since R_β is locally square integrable, the right-hand side is finite, and hence $R_\beta \in L^2(\mathbb{R}_+;\mathbb{R}^{m\times m})$. Since $E_\beta = R_\beta(-\beta)^{-1}$, it follows that also E_β is completely monotone and belongs to $L^1(\mathbb{R}_+;\mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+;\mathbb{R}^{m\times m})$. This proves the assertion. □

Below we apply our results to the fractional kernel.

EXAMPLE 5.9. Let $\beta \in \mathbb{R}^{m\times m}$ be symmetric with only strictly negative eigenvalues and $K(t) = \frac{t^{H-1/2}}{\Gamma(H+1/2)} I_m$ with $H \in (0, 1/2)$. Then $R_\beta, E_\beta \in L^1(\mathbb{R}_+;\mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+;\mathbb{R}^{m\times m})$ are completely monotone and satisfy

$$\int_0^\infty R_\beta(t)\,dt = I_m, \qquad \int_0^\infty E_\beta(t)\,dt = (-\beta)^{-1}.$$

PROOF. In view of the previous proposition, it is clear that R_β, E_β are completely monotone and satisfy the integrability conditions. The explicit formulas for the integrals can be obtained from the Laplace transforms. Namely,

$$\begin{aligned}\int_0^\infty R_\beta(t)\,dt &= \lim_{z\downarrow 0} \int_0^\infty e^{-zt} R_\beta(t)\,dt \\ &= \lim_{z\downarrow 0} \widehat{K_\beta}(z)\big(I_m + \widehat{K_\beta}(z)\big)^{-1} \\ &= \lim_{z\downarrow 0} I_m - \big(I_m + \widehat{K_\beta}(z)\big)^{-1} \\ &= \lim_{z\downarrow 0} I_m - \big(I_m - z^{-H-1/2}\beta\big)^{-1} = I_m,\end{aligned}$$

and hence $\int_0^\infty E_\beta(t)\,dt = \int_0^\infty R_\beta(t)(-\beta)^{-1}\,dt = (-\beta)^{-1}$. □

The next example also covers the case where β is not symmetric but the fractional kernel is instead replaced by an integrable Gamma kernel.

EXAMPLE 5.10. Let $K(t) = \frac{t^{H-1/2}}{\Gamma(H+1/2)} e^{-\lambda t} I_m$ with $H \in (0, 1/2)$, $\lambda > 0$, and let $\beta \in \mathbb{R}^{m\times m}$ be invertible. Then R_β, E_β are integrable if and only if

$$\sigma(\beta) \cap \{(z+\lambda)^{H+1/2} : z \in \mathbb{C}, \operatorname{Re}(z) \ge 0\} = \varnothing. \tag{5.13}$$

In such a case one has, by direct computations using Laplace transforms,

$$\int_0^\infty R_\beta(t)\,dt = (-\beta)\big(\lambda^{H+1/2} - \beta\big)^{-1}, \qquad \int_0^\infty E_\beta(t)\,dt = \big(\lambda^{H+1/2} - \beta\big)^{-1}. \tag{5.14}$$

PROOF. Since $-K\beta$ is integrable, the Paley–Wiener theorem (see [31], Chapter 2, Theorem 4.1) states that R_β is integrable if and only if $\det(I_m - \widehat{K}(z)\beta) \neq 0$ holds for all $z \in \mathbb{C}$ with $\operatorname{Re}(z) \ge 0$. Computing the Laplace transform of K gives $\widehat{K}(z) = (z+\lambda)^{-H-1/2} I_m$. Hence, $\det(I_m - \widehat{K}(z)\beta) \neq 0$ is equivalent to $\det((z+\lambda)^{H+1/2} I_m - \beta) \neq 0$, that is, R_β is integrable if and only if (5.13) holds. Since $E_\beta = R_\beta(-\beta)^{-1}$, we see that $E_\beta \in L^1(\mathbb{R}_+)$ is also equivalent to (5.13). To prove (5.14), we can use (5.11) to get

$$\int_0^\infty R_\beta(t)\,dt = \widehat{K_\beta}(0)\big(I_m + \widehat{K_\beta}(0)\big)^{-1} = (-\beta)\big(\lambda^{H+1/2} - \beta\big)^{-1}.$$

The second inequality in (5.14) now follows again from $E_\beta = R_\beta(-\beta)^{-1}$. □

Note that condition (5.13) is satisfied, if $\sigma(\beta) \subset \{z \in \mathbb{C} : \mathrm{Re}(z) < \lambda^{H+1/2}\}$. Finally, for the Gamma kernel in dimension $m = 1$, we provide the asymptotics for $\mathrm{cov}(X_t^{\mathrm{stat}}, X_s^{\mathrm{stat}})$ (see Corollary 5.5 for its explicit formula) as $t - s \to \infty$, and hence prove the formula (1.5) given in Theorem 1.3.

EXAMPLE 5.11. Let $m = 1$ and (b, β, σ, K) be admissible parameters with $\beta < 0, \sigma > 0$ and $K(t) = \frac{t^{H-1/2}}{\Gamma(H+1/2)} e^{-\lambda t}$, where $H \in (0, 1/2)$ and $\lambda \geq 0$. Then there exist positive constants $h_0, c(h_0), C(h_0)$ such that for all $h \geq h_0$,

$$c(h_0) h^{-(H+3/2)} e^{-\lambda h} \leq \int_0^\infty E_\beta(h+u) E_\beta(u)\, du \leq C(h_0) h^{-(H+3/2)} e^{-\lambda h}.$$

PROOF. For $\alpha \in (0, 1)$, define $e_\alpha(t) = t^{\alpha-1} M_\alpha(-t^\alpha)$, where $M_\alpha(z) = \sum_{n=0}^\infty z^n / \Gamma(\alpha n + \alpha)$ denotes the Mittag–Leffler function. It follows from [16], Section A.1, that $e_\alpha(t) \asymp t^{-1-\alpha}$ as $t \to \infty$. By [3], Table 1, we have $E_\beta(t) = |\beta|^{-1+\alpha^{-1}} e^{-\lambda t} e_\alpha(|\beta|^{1/\alpha} t)$ with $\alpha = H + 1/2$. Hence, we obtain $E_\beta(t) \asymp t^{-H-3/2} e^{-\lambda t}$ as $t \to \infty$. Thus, we find $h_0 \geq 1$ large enough and constants $c(h_0), C(h_0) > 0$ such that

$$\begin{aligned}\int_0^\infty E_\beta(h+u) E_\beta(u)\, du &\leq C(h_0) \int_0^\infty (h+u)^{-H-3/2} e^{-\lambda(h+u)} E_\beta(u)\, du \\ &\leq h^{-(H+3/2)} e^{-\lambda h} C(h_0) \int_0^\infty E_\beta(u)\, du\end{aligned}$$

holds for $h \geq h_0$. Similarly, for $h \geq h_0$, we obtain

$$\begin{aligned}\int_0^\infty E_\beta(h+u) E_\beta(u)\, du &\geq c(h_0) \int_0^\infty (h+u)^{-H-3/2} e^{-\lambda(h+u)} E_\beta(u)\, du \\ &\geq c(h_0)(2h)^{-(H+3/2)} e^{-\lambda h} \int_0^h e^{-\lambda u} E_\beta(u)\, du \\ &\geq c(h_0) 2^{-(H+3/2)} h^{-(H+3/2)} e^{-\lambda h} \int_0^{h_0} e^{-\lambda u} E_\beta(u)\, du.\end{aligned}$$

Combining both estimates proves the assertion. □

REMARK 5.12. Under same conditions as in the previous example, if $H = 1/2$, then $E_\beta(t) = e^{-(\lambda+|\beta|)t}$, and hence

$$\int_0^\infty E_\beta(h+u) E_\beta(u)\, du = \frac{e^{-(\lambda+|\beta|)h}}{2(\lambda + |\beta|)}.$$

Thus, when $\lambda = 0$, we observe a phase transition from power-law to exponential decay in the asymptotics of the autocovariance function as $H \nearrow 1/2$.

Finally, below we provide a method with which examples for nonscalar-valued kernels K can be constructed.

PROPOSITION 5.13. *Let $K = \mathrm{diag}(K_1, \ldots, K_m) \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m})$ and let $\beta \in \mathbb{R}^{m\times m}$. If $\|\beta\|_2 \sum_{j=1}^m \|K_j\|_{L^1} < 1$, then R_β is integrable.*

PROOF. Since $-K\beta$ is integrable, we can use again the Paley–Wiener theorem to find that that R_β is integrable if and only if $\det(I_m - \widehat{K}(z)\beta) \neq 0$, for all $z \in \mathbb{C}$ with $\mathrm{Re}(z) \geq 0$.

Thus, it suffices to show that $I_m - \widehat{K}(z)\beta$ is invertible for all z with $\mathrm{Re}(z) \geq 0$. This is indeed the case, if $\|\widehat{K}(z)\beta\|_2 < 1$. Estimating $\|\widehat{K}(z)\beta\|_2 \leq \|\beta\|_2 \sum_{j=1}^m \|K_j\|_{L^1}$ proves the assertion. $\square$

The proof also provides a necessary condition for the integrability of R_β.

REMARK 5.14. Let $K = \mathrm{diag}(K_1, \dots, K_m) \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m})$ and let $\beta \in \mathbb{R}^{m\times m}$. Let $l \in \mathbb{N}$ be the number of distinct eigenvalues $\rho_1, \dots, \rho_l$ of β. If R_β is integrable, then $\rho_k(\widehat{K}_1(z), \dots, \widehat{K}_m(z))^\top \neq (1, \dots, 1)^\top$ holds for each $k = 1, \dots, m$ and $z \in \mathbb{C}$ with $\mathrm{Re}(z) \geq 0$.

PROOF. Arguing as before, R_β is integrable if and only if $I_m - \widehat{K}(z)\beta$ is invertible. In particular, one has $(I_m - \widehat{K}(z)\beta)w \neq 0$ for $w \neq 0$. Let $k \in \{1, \dots, m\}$ and w_k be an eigenvector for the eigenvalue ρ_k. Since K is diagonal, $\widehat{K}$ is also diagonal, and hence $\widehat{K}(z)\beta w_k = \rho_k \widehat{K}(z) w_k \neq w_k$. This proves the assertion. $\square$

EXAMPLE 5.15. Let (b, β, σ, K) be admissible parameters with $K_j(t) = \frac{t^{H_j - 1/2}}{\Gamma(H_j + 1/2)} \times e^{-\lambda_j t}$, where $H_j \in (0, 1/2)$ and $\lambda_j > 0$. Then $\int_0^\infty K_j(t)\,dt = \lambda_j^{-H_j - 1/2}$. Therefore, for each β we can find $\lambda_1, \dots, \lambda_m$ large enough such that Proposition 5.13 is applicable.

6. Absolute continuity of the law. In this section, we provide sufficient conditions for the distribution of the Volterra square-root process to be absolutely continuous with respect to the Lebesgue measure. Moreover, we establish also similar results for the limiting distributions π_{x_0}. For $f : \mathbb{R}^m \longrightarrow \mathbb{R}$ and $x, h \in \mathbb{R}^m$ set $\Delta_h f(x) = f(x+h) - f(x)$. For $\lambda \in (0, 1)$, the Besov space $B_{1,\infty}^\lambda(\mathbb{R}^m)$ of order λ and integrability $(1, \infty)$ consists of all equivalence classes of functions f with finite norm

$$\|f\|_{B_{1,\infty}^\lambda} = \|f\|_{L^1} + \sup_{|h|\leq 1} |h|^{-\lambda} \|\Delta_h f\|_{L^1}. \tag{6.1}$$

The following condition guarantees that the contribution from the noise term $\int_0^t K(t-s)\sigma(X_s)\,dB_s$ is nondegenerate on the event where X belongs to the interior of the state space; see (6.8).

(R) There exists $\alpha \in [\gamma, 2]$ and a constant $C_* > 0$ such that

$$\int_0^h K_i(r)^2\,dr \geq C_* h^\alpha, \quad \forall h \in [0, 1] \tag{6.2}$$

holds for all $i = 1, \dots, m$.

For an admissible kernel K satisfying (6.2), one necessarily has $\alpha \geq \gamma$.

EXAMPLE 6.1. Let $K(t) = \frac{t^{H-1/2}}{\Gamma(H+1/2)} e^{-\lambda t} I_m$ with $H \in (0, 1/2)$ and $\lambda \geq 0$. Then condition (R) holds for $\alpha = 2H = \gamma$. Thus, all results below are applicable to this kernel.

The following is our first main result for this section.

THEOREM 6.2. *Let X be the Volterra square-root process with admissible parameters (b, β, σ, K). Suppose that condition (R) holds, and that*

$$\sigma_* := \min\{\sigma_1, \dots, \sigma_m\} > 0, \qquad \frac{\gamma}{\alpha} > \frac{2}{3}. \tag{6.3}$$

Denote by $\mu_t(dx)$ the distribution of X_t at time $t \geq 0$. Then for each $t > 0$ the finite measure $\mu_t^*(dx) = \min\{1, x_1^{1/2}, \dots, x_m^{1/2}\}\mu_t(dx)$ has a density $\mu_t^*(x)$, and there exists $\lambda \in (0,1)$ such that for each $T > 0$,

$$\|\mu_t^*\|_{B^s_{1,\infty}} \leq C(T)(1 \wedge t)^{-\alpha/2}, \quad t \in (0,T], s \in (0,\lambda],$$

where $C(T) > 0$ is a constant.

PROOF. Here and below, we denote by $C(T) > 0$ a constant, which may vary from line to line. The proof relies essentially on an application of [13], Lemma 2.1, to the finite measure μ_t^*.

Step 1. For given $t > 0$ and $\varepsilon \in (0, 1 \wedge t)$, define

$$\begin{aligned} X_t^\varepsilon = {} & X_{t-\varepsilon} + \int_{t-\varepsilon}^t K(t-r)(b + \beta X_{t-\varepsilon})\,dr \\ & + \int_0^{t-\varepsilon} \big(K(t-r) - K(t-\varepsilon-r)\big)(b + \beta X_r)\,dr \\ & + \int_0^{t-\varepsilon} \big(K(t-r) - K(t-\varepsilon-r)\big)\sigma(X_r)\,dB_r \\ & + \int_{t-\varepsilon}^t K(t-r)\sigma(X_{t-\varepsilon})\,dB_r. \end{aligned} \tag{6.4}$$

By direct computation, we find that

$$X_t - X_t^\varepsilon = \int_{t-\varepsilon}^t K(t-r)\beta(X_r - X_{t-\varepsilon})\,dr + \int_{t-\varepsilon}^t K(t-r)\big(\sigma(X_r) - \sigma(X_{t-\varepsilon})\big)\,dB_r.$$

Using Proposition 4.6, one can easily show that for each $T > 0$ there exists a constant $C(T) > 0$ such that

$$\mathbb{E}\big[|X_t - X_t^\varepsilon|^2\big] \leq C(T)\varepsilon^{3\gamma/2}, \quad t \in (0,T], \varepsilon \in (0, 1 \wedge t). \tag{6.5}$$

Step 2. Let $\eta \in (0,1)$ and $0 < t \leq T$. Define $\rho(x) = \min\{1, x_1^{1/2}, \dots, x_m^{1/2}\}$. In this step, we prove that there exists a constant $C(T) > 0$ such that for all $\phi \in C_b^\eta(\mathbb{R}^m)$ and $h \in \mathbb{R}^m$ satisfying $|h| \leq 1$ one has

$$\big|\mathbb{E}\big[\rho(X_t)\Delta_h\phi(X_t)\big]\big| \leq C(T)\|\phi\|_{C_b^\eta}\big(|h|^\eta \varepsilon^{3\gamma/8} + \varepsilon^{3\eta\gamma/4} + |h|\varepsilon^{-\alpha/2}\big), \tag{6.6}$$

for $\varepsilon \in (0, 1 \wedge t]$. Observe that

$$\begin{aligned} \big|\mathbb{E}\big[\rho(X_t)\Delta_h\phi(X_t)\big]\big| & \leq \mathbb{E}\big[|\rho(X_t) - \rho(X_{t-\varepsilon})||\Delta_h\phi(X_t)|\big] \\ & \quad + \mathbb{E}\big[\rho(X_{t-\varepsilon})|\Delta_h\phi(X_t) - \Delta_h\phi(X_t^\varepsilon)|\big] \\ & \quad + \big|\mathbb{E}\big[\rho(X_{t-\varepsilon})\Delta_h\phi(X_t^\varepsilon)\big]\big| \\ & \leq C(T)|h|^\eta\|\phi\|_{C_b^\eta}\varepsilon^{3\gamma/8} + C(T)\|\phi\|_{C_b^\eta}\varepsilon^{3\eta\gamma/4} \\ & \quad + \big|\mathbb{E}\big[\rho(X_{t-\varepsilon})\Delta_h\phi(X_t^\varepsilon)\big]\big|, \end{aligned}$$

where we have used $|\rho(x)-\rho(y)| \le m|x-y|^{1/2}$. To estimate the last term, we first note that $X_t^\varepsilon = U_t^\varepsilon + V_t^\varepsilon$, where U_t^ε and V_t^ε are given by

$$\begin{cases} U_t^\varepsilon = X_{t-\varepsilon} + \displaystyle\int_{t-\varepsilon}^t K(t-r)(b+\beta X_{t-\varepsilon})\,dr \\ \qquad + \displaystyle\int_0^{t-\varepsilon} \big(K(t-r)-K(t-\varepsilon-r)\big)(b+\beta X_r)\,dr \\ \qquad + \displaystyle\int_0^{t-\varepsilon} \big(K(t-r)-K(t-\varepsilon-r)\big)\sigma(X_r)\,dB_r, \\ V_t^\varepsilon = \displaystyle\int_{t-\varepsilon}^t K(t-r)\sigma(X_{t-\varepsilon})\,dB_r. \end{cases} \tag{6.7}$$

Observe that $V_t^\varepsilon = \int_0^\varepsilon K(\varepsilon-r)\sigma(X_{t-\varepsilon})\,d\widetilde{B}_r$ holds almost surely, where $\widetilde{B}_r = B_{t-\varepsilon+r} - B_{t-\varepsilon}$ is a new Brownian motion with respect to a new (shifted) filtration $\widetilde{\mathcal{F}}_r = \mathcal{F}_{t-\varepsilon+r}$. Hence, V_t^ε has, for fixed t and ε, conditionally on $\mathcal{F}_{t-\varepsilon}$ a Gaussian distribution with mean zero and variance satisfying for each $x \in \mathbb{R}^m$,

$$\begin{aligned} \langle x, \mathrm{var}(V_t^\varepsilon | \mathcal{F}_{t-\varepsilon}) x\rangle &= \int_0^\varepsilon \langle K(\varepsilon-r)^\top x, \sigma(X_{t-\varepsilon})\sigma(X_{t-\varepsilon})^\top K(\varepsilon-r)^\top x\rangle\,dr \\ &\ge \sigma_*^2 \rho(X_{t-\varepsilon})^2 \int_0^\varepsilon \langle x, K(r)K(r)^\top x\rangle\,dr \\ &\ge \sigma_*^2 \rho(X_{t-\varepsilon})^2 C_* \varepsilon^\alpha |x|^2. \end{aligned} \tag{6.8}$$

The product in front of $|x|^2$ in (6.8) is strictly positive on $\{\rho(X_{t-\varepsilon})>0\}$. Hence, the law of V_t^ε, when restricted to $\{\rho(X_{t-\varepsilon})>0\}$, has conditionally on $\mathcal{F}_{t-\varepsilon}$ a density $f_t^\varepsilon(z; X_{t-\varepsilon})$. Noting that U_t^ε is $\mathcal{F}_{t-\varepsilon}$ measurable, we find that

$$\begin{aligned} R_3 &= \left|\mathbb{E}\left[\int_{\mathbb{R}^m} \rho(X_{t-\varepsilon})\mathbb{1}_{\{\rho(X_{t-\varepsilon})>0\}} \Delta_h \phi\big(U_t^\varepsilon + z\big) f_t^\varepsilon(z; X_{t-\varepsilon})\,dz\right]\right| \\ &= \left|\mathbb{E}\left[\int_{\mathbb{R}^m} \rho(X_{t-\varepsilon})\mathbb{1}_{\{\rho(X_{t-\varepsilon})>0\}} \phi\big(U_t^\varepsilon + z\big) \Delta_{-h} f_t^\varepsilon(z; X_{t-\varepsilon})\,dz\right]\right| \\ &\le \|\phi\|_{C_b^\eta} |h| \mathbb{E}\left[\rho(X_{t-\varepsilon})\mathbb{1}_{\{\rho(X_{t-\varepsilon})>0\}} \int_0^1 \int_{\mathbb{R}^m} \big|\nabla f_t^\varepsilon(z-rh; X_{t-\varepsilon})\big|\,dz\,dr\right] \\ &\le C(T)\|\phi\|_{C_b^\eta} \mathbb{E}\left[\rho(X_{t-\varepsilon})\mathbb{1}_{\{\rho(X_{t-\varepsilon})>0\}} \frac{|h|}{\sqrt{\lambda_{\min}(\mathrm{var}(V_t^\varepsilon | \mathcal{F}_{t-\varepsilon}))}}\right] \\ &\le C(T)\|\phi\|_{C_b^\eta} |h| \varepsilon^{-\alpha/2}, \end{aligned}$$

where we have used that on $\{\rho(X_{t-\varepsilon})>0\}$,

$$\int_{\mathbb{R}^m} \big|\nabla f_t^\varepsilon(z-rh; X_{t-\varepsilon})\big|\,dz \le \frac{C}{\sqrt{\lambda_{\min}(\mathrm{var}(V_t^\varepsilon | \mathcal{F}_{t-\varepsilon}))}} \le \frac{C(T)}{\rho(X_{t-\varepsilon})} \varepsilon^{-\alpha/2}$$

for $\varepsilon \in (0, 1\wedge t]$. Here, $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue of a symmetric positive definite matrix. This proves (6.6).

Step 3. Since $\frac{\gamma}{\alpha} > \frac{2}{3}$, we find a constant a such that $\frac{4}{3\gamma} < a < \frac{2}{\alpha}$. For this choice of a, we find sufficiently small $\eta > 0$ such that $1-\eta-a\alpha/2 > 0$. Letting $\varepsilon = |h|^a(1\wedge t)$ with $|h| \le 1$, we find for any $\phi \in C_b^\eta(\mathbb{R}^m)$,

$$\begin{aligned} &\big|\mathbb{E}\big[\rho(X_t)\Delta_h\phi(X_t)\big]\big| \\ &\quad \le C(T)\|\phi\|_{C_b^\eta}\big(|h|^{\eta+3a\gamma/8} + |h|^{3a\eta\gamma/4} + |h|^{1-a\alpha/2}\big)(1\wedge t)^{-\alpha/2} \\ &\quad \le C(T)\|\phi\|_{C_b^\eta}|h|^{\eta+\min\{3a\gamma/8, 3a\eta\gamma/4-\eta, 1-\eta-a\alpha/2\}}(1\wedge t)^{-\alpha/2} \end{aligned}$$

$$\leq C(T)\|\phi\|_{C_b^\eta}|h|^{\eta+\lambda}(1\wedge t)^{-\alpha/2},$$

where $\lambda = \min\{3a\gamma/8, \eta(3a\gamma/4-1), 1-\eta-a\alpha/2\} > 0$. This shows that [13], Lemma 2.1, is applicable, and hence proves the assertion. □

Note that μ_t^* is equivalent to μ_t on $\mathbb{R}_{++}^m = \{x \in \mathbb{R}_+^m : x_i > 0, \forall i = 1, \dots, m\}$. Hence, we have the following corollary.

COROLLARY 6.3. *Assume that all assumptions of Theorem 6.2 hold true. Then $\mu_t|_{\mathbb{R}_{++}^m}$ has a density with respect to the Lesbegue measure.*

Note that, without further conditions only μ_t^* but not necessarily the density of $\mu_t|_{\mathbb{R}_{++}^m}$ belong to the Besov space $B_{1,\infty}^\lambda(\mathbb{R}^d)$ with some small $\lambda \in (0,1)$. To see this, let X be the classical CIR process where $K(t) = 1$ and $m = 1$. Assume additionally $\beta = x_0 = 0$ and $\sigma, b > 0$. Then $\frac{2X_t}{\sigma^2 t}$ has χ^2-distribution with $k := \frac{2b}{\sigma^2} > 0$ degrees of freedom (see also [18], remark after Theorem 4), that is, it has distribution $\nu(x)\,dx$ with $\nu(x) = \mathbb{1}_{(0,\infty)}(x)\frac{x^{k/2-1}e^{-x/2}}{2^{k/2}\Gamma(k/2)}$. In particular, there exists $\lambda \in (0,1)$ with $\nu \in B_{1,\infty}^\lambda(\mathbb{R})$ iff $k = \frac{2b}{\sigma^2} > 2$, while $\min\{1, x^{1/2}\}\nu \in B_{1,\infty}^\lambda(\mathbb{R})$ holds whenever $\lambda \in (0, k/2 \wedge 1)$.

COROLLARY 6.4. *Let X be the Volterra square-root process with admissible parameters (b, β, σ, K) and initial state x_0. Suppose that (R) and (K) are satisfied, and that $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m\times m}) \cap L^2(\mathbb{R}_+; \mathbb{R}^{m\times m})$. Assume that (6.3) holds. Let π_{x_0} be the limiting distribution of X_t as $t \to \infty$. Then $\pi_{x_0}^*$ defined by $\pi_{x_0}^*(dy) = \min\{1, y_1^{1/2}, \dots, y_m^{1/2}\}\pi_{x_0}(dy)$ has a density $\pi_{x_0}^*(y)$ with respect to the Lebesgue measure, and there exists $\lambda \in (0,1)$ such that $\|\pi_{x_0}^*\|_{B_{1,\infty}^s} < \infty$ holds for $s \in (0, \lambda]$. In particular, π_{x_0} is absolutely continuous with respect to the Lebesgue measure, when restricted to $\mathbb{R}_{++}^m$.*

PROOF. Following the same steps of the previous proof but now applying Proposition 4.7 instead, we see that all constants can be chosen to be independent of T, that is, we find a constant $C > 0$ such that for each $\phi \in C_b^\eta(\mathbb{R}^m)$ and $h \in \mathbb{R}^m$ satisfying $|h| \leq 1$,

$$\left|\mathbb{E}\left[\rho(X_t)\Delta_h\phi(X_t)\right]\right| \leq C\|\phi\|_{C_b^\eta}\left(|h|^\eta\varepsilon^{3\gamma/8} + \varepsilon^{3\eta\gamma/4} + |h|\varepsilon^{-\alpha/2}\right).$$

Hence, arguing as in Step 3, we can find $a \in (\frac{4}{3}\frac{1}{\gamma}, \frac{2}{\alpha})$ and $\eta \in (0,1)$ with $1 - \eta - a/\alpha > 0$ such that

$$\left|\mathbb{E}\left[\rho(X_t)\Delta_h\phi(X_t)\right]\right| \leq \|\phi\|_{C_b^\eta}|h|^{\eta+\lambda}(1\wedge t)^{-\alpha/2},$$

where $\lambda = \min\{3a\gamma/8, \eta(3a\gamma/4-1), 1-\eta-a\alpha/2\} > 0$ and $\phi \in C_b^\eta(\mathbb{R}^m)$. Hence, we find that

$$\left|\int_{\mathbb{R}_+^m}\Delta_h\phi(x)\pi_{x_0}^*(dx)\right| = \lim_{t\to\infty}\left|\mathbb{E}\left[\rho(X_t)\Delta_h\phi(X_t)\right]\right| \leq C\|\phi\|_{C_b^\eta}|h|^{\eta+\lambda}.$$

This shows that [13], Lemma 2.1, is applicable, and hence proves the assertion. □

APPENDIX A: CONVOLUTIONS AND RESOLVENTS

The following elementary lemma is crucial for the continuity of solutions to the Riccati–Volterra equation.

LEMMA A.1. *Let $k \in L^1_{\mathrm{loc}}(\mathbb{R}_+) \cap C((0,\infty))$ be scalar-valued, nonnegative and nonincreasing. Let $g \in L^1_{\mathrm{loc}}(\mathbb{R}_+) \cap L^\infty_{\mathrm{loc}}((0,\infty))$. Then $k * g \in C((0,\infty))$.*

PROOF. Fix $t > 0$. It suffices to show that $\lim_{n\to\infty} k * g(t_n) = k * g(t)$ holds for any sequence $(t_n)_{n\geq 1}$ such that either $t_n \searrow t$ or $t_n \nearrow t$. Suppose first that $t_n \nearrow t$. Then $|k * g(t_n) - k * g(t)| \leq \int_0^{t_n} |k(t-s) - k(t_n - s)||g(s)|\,ds + \int_{t_n}^t k(t-s)|g(s)|\,ds$. By dominated convergence, the second term converges to zero. To estimate the first term, we suppose without loss of generality that $0 < t_1 < t_2 < t$. Let $\delta \in (0, t_1)$, then for $n \geq 2$

$$
\begin{aligned}
&\int_0^{t_n} \big|k(t-s) - k(t_n - s)\big|\big|g(s)\big|\,ds \\
&\quad = \int_0^{\delta} \big|k(t-s) - k(t_n - s)\big|\big|g(s)\big|\,ds + \int_{\delta}^{t_n} \big|k(t-s) - k(t_n - s)\big|\big|g(s)\big|\,ds \\
&\quad \leq \big(k(t-t_1) + k(t_2 - t_1)\big)\int_0^{\delta} \big|g(s)\big|\,ds \\
&\qquad + \|g\|_{L^\infty([\delta,t])} \int_0^t \big|k\big((t-t_n)+s\big) - k(s)\big|\,ds,
\end{aligned}
$$

where we have used that k is nonnegative and nonincreasing. Note that $|k((t-t_n)+s) - k(s)| \leq 2k(s)$, hence dominated convergence implies that the second term converges to zero. Hence, letting first $n \to \infty$ and then $\delta \searrow 0$ shows that $\int_0^{t_n} |k(t-s) - k(t_n - s)||g(s)|\,ds \to 0$, which proves the assertion for the case $t_n \nearrow t$. Suppose now that $t_n \searrow t$. Then $|k * g(t_n) - k * g(t)| \leq \int_0^t |k(t-s) - k(t_n - s)||g(s)|\,ds + \int_t^{t_n} k(t_n - s)|g(s)|\,ds$. Using that k is nonincreasing and $t_n > t$, we obtain $|k(t-s) - k(t_n - s)||g(s)| \leq k(t-s)|g(s)| \leq \mathbb{1}_{[0,t/2]}(s)k(t/2)|g(s)| + \mathbb{1}_{(t/2,t]}(s)k(t-s)|g(s)|$. Since the right-hand side is integrable in $s \in [0,t]$, dominated convergence implies that $\int_0^t |k(t-s) - k(t_n - s)||g(s)|\,ds \to 0$. For the second term, we note that $\int_t^{t_n} k(t_n - s)|g(s)|\,ds \leq \|g\|_{L^\infty([t,t_1])} \int_0^{t_n - t} k(s)\,ds \to 0$, $n \to \infty$. This proves the assertion for the case $t_n \searrow t$. $\square$

Next, we provide some technical estimates on the resolvent of the second kind.

LEMMA A.2. *Let $K \in L^2_{\mathrm{loc}}(\mathbb{R}_+; \mathbb{C}^{m\times m})$ and $B \in \mathbb{C}^{m\times m}$. Let R_B be the resolvent of the second kind for $K_B(t) = -K(t)B$ and let E_B be defined by (2.2). Then for all $T > 0$ and $s, t \in [0, T]$,*

$$
\begin{aligned}
\int_s^t \|E_B(r)\|_2^2\,dr \leq{}& 2\int_s^t \|K(r)\|_2^2\,dr \\
&+ 2\|R_B\|_{L^1([0,T])} \int_0^s \left(\int_{s-u}^{t-u} \|K(r)\|_2^2\,dr\right) \|R_B(u)\|_2\,du \\
&+ 2\|R_B\|_{L^1([0,T])}^2 \int_0^{t-s} \|K(r)\|_2^2\,dr.
\end{aligned}
$$

Moreover, for each $T > 0$ and $h \in (0,1]$ we have

$$
\begin{aligned}
\int_0^T \|E_B(t+h) - E_B(t)\|_2^2\,dt \leq{}& 3\|R_B\|_{L^1([0,T+1])}^2 \int_0^h \|K(r)\|_2^2\,dr \\
&+ 3\big(1 + \|R_B\|_{L^1([0,T])}\big)^2 \int_0^T \|K(t+h) - K(t)\|_2^2\,dt.
\end{aligned}
$$

PROOF. Using (2.2) gives $\|E_B\|_{L^2([s,t])} \leq \|K\|_{L^2([s,t])} + \|R_B * K\|_{L^2([s,t])}$. To estimate the second term, we first use Jensen's inequality and then Fubini's theorem to find that

$$\begin{aligned}\|R_B * K\|^2_{L^2([s,t])} &\leq \|R_B\|_{L^1([0,T])} \int_0^s \left(\int_{s-u}^{t-u} \|K(r)\|_2^2 \, dr \right) \|R_B(u)\|_2 \, du \\ &\quad + \|R_B\|_{L^1([0,T])} \int_s^t \left(\int_0^{t-u} \|K(r)\|_2^2 \, dr \right) \|R_B(u)\|_2 \, du \\ &\leq \|R_B\|_{L^1([0,T])} \int_0^s \left(\int_{s-u}^{t-u} \|K(r)\|_2^2 \, dr \right) \|R_B(u)\|_2 \, du \\ &\quad + \|R_B\|^2_{L^1([0,T])} \int_0^{t-s} \|K(r)\|_2^2 \, dr.\end{aligned}$$

Combining these estimates prove the first inequality. For the second estimate, we first note that $\|E_B(\cdot + h) - E_B(\cdot)\|_{L^2([0,T])} \leq \|I_1\|_{L^2([0,T])} + \|I_2\|_{L^2([0,T])} + \|I_3\|_{L^2([0,T])}$ with $I_1 = K(t+h) - K(t)$, $I_2 = \int_0^t R_B(r)(K(t+h-r) - K(t-r)) \, dr$, and $I_3 = \int_t^{t+h} R_B(r) K(t+h-r) \, dr$. Below we estimate the last two terms separately. For I_2, we obtain from Jensen's inequality

$$\begin{aligned}\|I_2\|^2_{L^2([0,T])} &\leq \|R_B\|_{L^1([0,T])} \int_0^T \int_0^t \|R_B(r)\|_2 \|K(t+h-r) - K(t-r)\|_2^2 \, dr \, dt \\ &\leq \|R_B\|^2_{L^1([0,T])} \int_0^T \|K(t+h) - K(t)\|_2^2 \, dt.\end{aligned}$$

Similarly, we obtain by Fubini's theorem

$$\begin{aligned}\|I_3\|^2_{L^2([0,T])} &\leq \|R_B\|_{L^1([0,T+1])} \int_0^T \int_t^{t+h} \|R_B(r)\|_2 \|K(t+h-r)\|_2^2 \, dr \, dt \\ &\leq \|R_B\|^2_{L^1([0,T+1])} \int_0^h \|K(r)\|_2^2 \, dr.\end{aligned}$$

Combining all estimates proves the assertion. □

Below we specify the above estimates to the case of admissible parameters (b, β, σ, K) with $B = \beta$. There we will also use the following observation: for any $t, s \geq 0$ with $s < t$ and $i = 1, \ldots, m$, it holds the inequality

$$\int_s^t |K_i(r)|^2 \, dr \leq \int_0^{t-s} |K_i(r)|^2 \, dr \leq C_1 (t-s)^\gamma,$$

since K_i is admissible, and thus nonincreasing.

LEMMA A.3. *Let (b, β, σ, K) be admissible parameters and suppose that condition (K) holds. Suppose that $E_\beta \in L^1(\mathbb{R}_+; \mathbb{R}^{m \times m})$. Then there exists a constant $C > 0$ such that for all $t, s \geq 0$ with $0 \leq t - s \leq 1$, the following inequality holds*:

$$\begin{aligned}&\int_s^t \|R_\beta(r)\|_2^2 \, dr + \int_s^t \|E_\beta(r)\|_2^2 \, dr \\ &\quad + \int_0^s \|E_\beta(t-r) - E_\beta(s-r)\|^2_{\mathrm{HS}} \, dr + \int_s^t \|E_\beta(t-r)\|^2_{\mathrm{HS}} \, dr \leq C(t-s)^\gamma.\end{aligned}$$

PROOF. The first two terms can be estimated by Lemma A.2 and condition (v) for admissible parameters, that is,

$$\begin{aligned}\int_s^t \|R_\beta(r)\|_2^2\,dr + \int_s^t \|E_\beta(r)\|_2^2\,dr &\le \big(1+\|\beta\|_2^2\big)\int_s^t \|E_\beta(r)\|_2^2\,dr \\ &\le 4mC_1\big(1+\|\beta\|_2^2\big)\big(1+\|R_\beta\|_{L^1(\mathbb{R}_+)}^2\big)(t-s)^\gamma.\end{aligned}$$

Similarly, we obtain from Lemma A.2,

$$\begin{aligned}&\int_s^t \|E_\beta(t-r)\|_{\mathrm{HS}}^2\,dr \\ &\quad\le 2m\int_0^{t-s}\|K(r)\|_2^2\,dr + 2m\|R_\beta\|_{L^1(\mathbb{R}_+)}^2\int_0^{t-s}\|K(r)\|_2^2\,dr \\ &\quad\le 2m^2C_1(t-s)^\gamma + 2m^2\|R_\beta\|_{L^1(\mathbb{R}_+)}^2 C_1(t-s)^\gamma.\end{aligned}$$

Finally, the last term can be estimated by

$$\begin{aligned}&\int_0^s \|E_\beta(t-r)-E_\beta(s-r)\|_{\mathrm{HS}}^2\,dr \\ &\quad\le 3m\big(1+\|R_\beta\|_{L^1(\mathbb{R}_+)}\big)^2\int_0^s \|K(r+t-s)-K(r)\|_2^2\,dr \\ &\qquad + 3m\|R_\beta\|_{L^1(\mathbb{R}_+)}^2\int_0^{t-s}\|K(r)\|_2^2\,dr \\ &\quad\le 3m^2\big(1+\|R_\beta\|_{L^1(\mathbb{R}_+)}\big)^2 C_3(t-s)^\gamma + 3m^2\|R_\beta\|_{L^1(\mathbb{R}_+)}^2 C_1(t-s)^\gamma.\end{aligned}$$

Combining all estimates proves the assertion. □

APPENDIX B: UNIFORM IN TIME BOUNDS

B.1. Moment estimates.

PROOF OF LEMMA 4.5. Note that

$$\begin{aligned}\mathbb{E}\big[|X_t|^p\big] \le 2^{p-1}\bigg(1+\int_0^\infty \|R_\beta(s)\|_2\,ds\bigg)^p|x_0|^p + 2^{p-1}\bigg(\int_0^\infty \|E_\beta(s)\|_2\,ds\bigg)^p|b|^p \\ + c2^{p-1}\mathbb{E}\bigg[\bigg(\int_0^t \|E_\beta(t-s)\|_{\mathrm{HS}}^2\|\sigma(X_s)\|_{\mathrm{HS}}^2\,ds\bigg)^{p/2}\bigg],\end{aligned}$$

where $c>0$ is a constant only depending on m,p and the first two terms are finite since E_β (and hence R_β) is integrable over $\mathbb{R}_+$. For the last term, we obtain with $\sigma^*=\max\{\sigma_1^2,\dots,\sigma_m^2\}$

$$\begin{aligned}&\mathbb{E}\bigg[\bigg(\int_0^t \|E_\beta(t-s)\|_{\mathrm{HS}}^2\|\sigma(X_s)\|_{\mathrm{HS}}^2\,ds\bigg)^{p/2}\bigg] \\ &\quad\le (\sqrt{m}\sigma^*)^{p/2}\bigg(\int_0^t \|E_\beta(t-s)\|_{\mathrm{HS}}^2\,ds\bigg)^{p/2-1}\int_0^t \|E_\beta(t-s)\|_{\mathrm{HS}}^2\mathbb{E}\big[|X_s|^{p/2}\big]\,ds \\ &\quad\le (\sqrt{m}\sigma^*)^{p/2}\sup_{t\ge0}\mathbb{E}\big[|X_t|^{p/2}\big]\bigg(\int_0^\infty \|E_\beta(s)\|_{\mathrm{HS}}^2\,ds\bigg)^{p/2},\end{aligned}$$

where we have used $\|\sigma(x)\|_{\mathrm{HS}}^2=\sum_{k=1}^m \sigma_k^2 x_k \le \sigma^*\sqrt{m}|x|$ and Jensen's inequality. Thus, we have shown that

$$1+\sup_{t\ge0}\mathbb{E}\big[|X_t|^p\big] \le C(p)\Big(1+\sup_{t\ge0}\mathbb{E}\big[|X_t|^{p/2}\big]\Big),$$

where $C(p)>0$ is a constant. Letting $n\in\mathbb{N}$ be the smallest integer with $p/2^n\le 2$ gives

$$\begin{aligned}1+\sup_{t\ge 0}\mathbb{E}[|X_t|^p]&\le\prod_{k=0}^{n-1}C\Big(\frac{p}{2^k}\Big)\Big(1+\sup_{t\ge 0}\mathbb{E}[|X_t|^{p/2^n}]\Big)\\&\le\prod_{k=0}^{n-1}C\Big(\frac{p}{2^k}\Big)\Big(1+\sup_{t\ge 0}(1+\mathbb{E}[|X_t|^2])\Big).\end{aligned}$$

Thus, we can use the previous estimate to find $\sup_{t\ge 0}\mathbb{E}[|X_t|^2]\le C(2)(1+\sup_{t\ge 0}\mathbb{E}[|X_t|])$. This proves the assertion. $\square$

B.2. Uniform Hölder estimate.

PROOF OF PROPOSITION 4.7. Using (4.3), we find for $t,s\ge 0$ with $0\le t-s\le 1$ that $X_t-X_s=I_1+I_2+I_3+I_4$ with $I_1=-\int_s^t R_\beta(r)x_0\,dr$, $I_2=\int_s^t E_\beta(r)b\,dr$, $I_3=\int_0^s(E_\beta(t-r)-E_\beta(s-r))\sigma(X_r)\,dB_r$, $I_4=\int_s^t E_\beta(t-r)\sigma(X_r)\,dB_r$. For the first term, we obtain from Lemma A.3,

$$\begin{aligned}|I_1|^p&\le|x_0|^p(t-s)^{p/2}\Big(\int_s^t\|R_\beta(r)\|_2^2\,dr\Big)^{p/2}\\&\le|x_0|^pC^{p/2}(t-s)^{p/2+\gamma p/2}.\end{aligned}$$

Likewise, we obtain for the second term,

$$|I_2|^p\le|b|^p(t-s)^{p/2}\Big(\int_s^t\|E_\beta(r)\|_2^2\,dr\Big)^{p/2}\le|b|^pC^{p/2}(t-s)^{p/2+\gamma p/2}.$$

For the third term, we use the Burkholder–Davis–Gundy inequality to find

$$\begin{aligned}\mathbb{E}[|I_3|^p]&\le(\sqrt{m}\sigma^*)^{p/2}c(m,p)\Big(\int_0^s\|E_\beta(t-r)-E_\beta(s-r)\|_{\mathrm{HS}}^2\,dr\Big)^{p/2-1}\\&\quad\cdot\int_0^s\|E_\beta(t-r)-E_\beta(s-r)\|_{\mathrm{HS}}^2\mathbb{E}[|X_r|^{p/2}]\,dr\\&\le(\sqrt{m}\sigma^*)^{p/2}c(m,p)\sup_{t\ge 0}\mathbb{E}[|X_r|^{p/2}]C^{p/2}(t-s)^{\gamma p/2},\end{aligned}$$

where we have used $\|\sigma(x)\|_{\mathrm{HS}}^2\le\sigma^*\sqrt{m}|x|$ with $\sigma^*=\max\{\sigma_1^2,\dots,\sigma_m^2\}$. Similarly, we obtain for the last term

$$\begin{aligned}\mathbb{E}[|I_4|^p]&\le(\sqrt{m}\sigma^*)^{p/2}c(m,p)\sup_{t\ge 0}\mathbb{E}[|X_r|^{p/2}]\Big(\int_s^t\|E_\beta(t-r)\|_{\mathrm{HS}}^2\,dr\Big)^{p/2}\\&\le(\sqrt{m}\sigma^*)^{p/2}c(m,p)\sup_{t\ge 0}\mathbb{E}[|X_r|^{p/2}]C^{p/2}(t-s)^{\gamma p/2}.\end{aligned}$$

Combining all estimates and invoking Lemma 4.5 proves the assertion. $\square$

Acknowledgments. The authors would like to thank the referees for a careful reading of this manuscript, which lead to a great improvement of this work. Additional address of the second author is Department of Mathematics, Shantou University, Shantou, Guangdong 515063, China.

Funding. The research of Peng Jin is supported by the Guangdong Basic and Applied Basic Research Foundation (No. 2020A1515010436), the Guangdong Provincial Key Laboratory of IRADS, BNU-HKBU United International College (2022B1212010006), the Guangdong Higher Education Upgrading Plan (2021–2025) (UIC R0400024-21), the UIC Start-up Research Fund (No. R72021102) and NSFC (Nos. 11861029, 12071499).

REFERENCES

[1] ABI JABER, E. (2022). The Laplace transform of the integrated Volterra Wishart process. *Math. Finance* **32** 309–348. MR4370808 https://doi.org/10.1111/mafi.12334

[2] ABI JABER, E., CUCHIERO, C., LARSSON, M. and PULIDO, S. (2021). A weak solution theory for stochastic Volterra equations of convolution type. *Ann. Appl. Probab.* **31** 2924–2952. MR4350978 https://doi.org/10.1214/21-aap1667

[3] ABI JABER, E., LARSSON, M. and PULIDO, S. (2019). Affine Volterra processes. *Ann. Appl. Probab.* **29** 3155–3200. MR4019885 https://doi.org/10.1214/19-AAP1477

[4] ABI JABER, E., MILLER, E. and PHAM, H. (2021). Markowitz portfolio selection for multivariate affine and quadratic Volterra models. *SIAM J. Financial Math.* **12** 369–409. MR4229250 https://doi.org/10.1137/20M1347449

[5] ABI JABER, E., MILLER, E. and PHAM, H. (2021). Linear-quadratic control for a class of stochastic Volterra equations: Solvability and approximation. *Ann. Appl. Probab.* **31** 2244–2274. MR4332695 https://doi.org/10.1214/20-aap1645

[6] BARCZY, M., DÖRING, L., LI, Z. and PAP, G. (2013). On parameter estimation for critical affine processes. *Electron. J. Stat.* **7** 647–696. MR3035268 https://doi.org/10.1214/13-EJS786

[7] BARCZY, M., DÖRING, L., LI, Z. and PAP, G. (2014). Parameter estimation for a subcritical affine two factor model. *J. Statist. Plann. Inference* **151/152** 37–59. MR3216637 https://doi.org/10.1016/j.jspi.2014.04.001

[8] BARCZY, M., PAP, G. and SZABÓ, T. T. (2016). Parameter estimation for the subcritical Heston model based on discrete time observations. *Acta Sci. Math.* (*Szeged*) **82** 313–338. MR3526353 https://doi.org/10.14232/actasm-015-016-0

[9] BENTH, F. E., DETERING, N. and KRÜHNER, P. (2022). Stochastic Volterra integral equations and a class of first-order stochastic partial differential equations. *Stochastics* **94** 1054–1076. MR4503737 https://doi.org/10.1080/17442508.2021.2019738

[10] CONT, R. and DAS, P. (2022). Rough volatility: Fact or artefact? Available at arXiv:2203.13820.

[11] CUCHIERO, C. and TEICHMANN, J. (2019). Markovian lifts of positive semidefinite affine Volterra-type processes. *Decis. Econ. Finance* **42** 407–448. MR4031333 https://doi.org/10.1007/s10203-019-00268-5

[12] CUCHIERO, C. and TEICHMANN, J. (2020). Generalized Feller processes and Markovian lifts of stochastic Volterra processes: The affine case. *J. Evol. Equ.* **20** 1301–1348. MR4181950 https://doi.org/10.1007/s00028-020-00557-2

[13] DEBUSSCHE, A. and FOURNIER, N. (2013). Existence of densities for stable-like driven SDE's with Hölder continuous coefficients. *J. Funct. Anal.* **264** 1757–1778. MR3022725 https://doi.org/10.1016/j.jfa.2013.01.009

[14] EL EUCH, O., FUKASAWA, M. and ROSENBAUM, M. (2018). The microstructural foundations of leverage effect and rough volatility. *Finance Stoch.* **22** 241–280. MR3778355 https://doi.org/10.1007/s00780-018-0360-z

[15] EL EUCH, O. and ROSENBAUM, M. (2018). Perfect hedging in rough Heston models. *Ann. Appl. Probab.* **28** 3813–3856. MR3861827 https://doi.org/10.1214/18-AAP1408

[16] EL EUCH, O. and ROSENBAUM, M. (2019). The characteristic function of rough Heston models. *Math. Finance* **29** 3–38. MR3905737 https://doi.org/10.1111/mafi.12173

[17] FARKAS, B., FRIESEN, M., RÜDIGER, B. and SCHROERS, D. (2021). On a class of stochastic partial differential equations with multiple invariant measures. *NoDEA Nonlinear Differential Equations Appl.* **28** 28. MR4241464 https://doi.org/10.1007/s00030-021-00691-x

[18] FILIPOVIĆ, D., MAYERHOFER, E. and SCHNEIDER, P. (2013). Density approximations for multivariate affine jump-diffusion processes. *J. Econometrics* **176** 93–111. MR3084047 https://doi.org/10.1016/j.jeconom.2012.12.003

[19] FLANDOLI, F. and GATAREK, D. (1995). Martingale and stationary solutions for stochastic Navier–Stokes equations. *Probab. Theory Related Fields* **102** 367–391. MR1339739 https://doi.org/10.1007/BF01192467

[20] FORDE, M., GERHOLD, S. and SMITH, B. (2021). Small-time, large-time, and $H \to 0$ asymptotics for the rough Heston model. *Math. Finance* **31** 203–241. MR4205882 https://doi.org/10.1111/mafi.12290

[21] FORDE, M., GERHOLD, S. and SMITH, B. (2022). Small-time VIX smile and the stationary distribution for the rough heston model.

[22] FRIESEN, M. and JIN, P. (2020). On the anisotropic stable JCIR process. *ALEA Lat. Am. J. Probab. Math. Stat.* **17** 643–674. MR4130578 https://doi.org/10.30757/alea.v17-25

[23] FRIESEN, M., JIN, P., KREMER, J. and RÜDIGER, B. (2023). Regularity of transition densities and ergodicity for affine jump-diffusions. *Math. Nachr.* **296** 1117–1134. MR4585666 https://doi.org/10.1002/mana.202000299

[24] FRIESEN, M., JIN, P. and RÜDIGER, B. (2020). Existence of densities for multi-type continuous-state branching processes with immigration. *Stochastic Process. Appl.* **130** 5426–5452. MR4127334 https://doi.org/10.1016/j.spa.2020.03.012

[25] FRIESEN, M., JIN, P. and RÜDIGER, B. (2020). Stochastic equation and exponential ergodicity in Wasserstein distances for affine processes. *Ann. Appl. Probab.* **30** 2165–2195. MR4149525 https://doi.org/10.1214/19-AAP1554

[26] FUKASAWA, M. (2021). Volatility has to be rough. *Quant. Finance* **21** 1–8. MR4188876 https://doi.org/10.1080/14697688.2020.1825781

[27] FUKASAWA, M., TAKABATAKE, T. and WESTPHAL, R. (2019). Is volatility rough? Available at arXiv:1905.04852.

[28] GATHERAL, J., JAISSON, T. and ROSENBAUM, M. (2018). Volatility is rough. *Quant. Finance* **18** 933–949. MR3805308 https://doi.org/10.1080/14697688.2017.1393551

[29] GERHOLD, S., GERSTENECKER, C. and PINTER, A. (2019). Moment explosions in the rough Heston model. *Decis. Econ. Finance* **42** 575–608. MR4031338 https://doi.org/10.1007/s10203-019-00267-6

[30] GLASSERMAN, P. and KIM, K.-K. (2010). Moment explosions and stationary distributions in affine diffusion models. *Math. Finance* **20** 1–33. MR2599675 https://doi.org/10.1111/j.1467-9965.2009.00387.x

[31] GRIPENBERG, G., LONDEN, S.-O. and STAFFANS, O. (1990). *Volterra Integral and Functional Equations. Encyclopedia of Mathematics and Its Applications* **34**. Cambridge Univ. Press, Cambridge. MR1050319 https://doi.org/10.1017/CBO9780511662805

[32] HAN, B. and WONG, H. Y. (2021). Mean-variance portfolio selection under Volterra Heston model. *Appl. Math. Optim.* **84** 683–710. MR4283941 https://doi.org/10.1007/s00245-020-09658-3

[33] HAN, B. and WONG, H. Y. (2021). Merton's portfolio problem under Volterra Heston model. *Finance Res. Lett.* **39** 101580.

[34] JACQUIER, A., PANNIER, A. and SPILIOPOULOS, K. (2022). On the ergodic behaviour of affine Volterra processes. Available at arXiv:2204.05270.

[35] JIN, P., KREMER, J. and RÜDIGER, B. (2020). Existence of limiting distribution for affine processes. *J. Math. Anal. Appl.* **486** 123912. MR4060087 https://doi.org/10.1016/j.jmaa.2020.123912

[36] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[37] LIEB, E. H. and LOSS, M. (2001). *Analysis*, 2nd ed. *Graduate Studies in Mathematics* **14**. Amer. Math. Soc., Providence, RI. MR1817225 https://doi.org/10.1090/gsm/014

[38] MAYERHOFER, E., STELZER, R. and VESTWEBER, J. (2020). Geometric ergodicity of affine processes on cones. *Stochastic Process. Appl.* **130** 4141–4173. MR4102262 https://doi.org/10.1016/j.spa.2019.11.012