Weierstraß-Institut
für Angewandte Analysis und Stochastik

Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

State spaces of multifactor approximations of nonnegative Volterra processes

Eduardo Abi Jaber[1], Christian Bayer[2], Simon Breneis[2]

submitted: January 14, 2025

[1] CMAP
Ecole Polytechnique
Route de Saclay
91128 Palaiseau Cedex
France
E-Mail: eduardo.abi-jaber@polytechnique.edu

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: christian.bayer@wias-berlin.de
simon.breneis@wias-berlin.de

No. 3162
Berlin 2025

2020 *Mathematics Subject Classification.* 60H20, 91-10, 91-08.

Key words and phrases. Stochastic Volterra equations, Markovian approximations, rough Heston model, viability and invariance, simulation, PDEs.

EAJ is grateful for the financial support from the Chaires FiME-FDD, Financial Risks, Deep Finance & Statistics and Machine Learning and systematic methods in finance at Ecole Polytechnique. CB and SB gratefully acknowledge the support by the IRTG 2544 "Stochastic Analysis in Interaction". CB also acknowledges support from DFG CRC/TRR 388 "Rough Analysis, Stochastic Dynamics and Related Fields", Project B02.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

State spaces of multifactor approximations of nonnegative Volterra processes

Eduardo Abi Jaber, Christian Bayer, Simon Breneis

Abstract

We show that the state spaces of multifactor Markovian processes, coming from approximations of nonnegative Volterra processes, are given by explicit linear transformation of the nonnegative orthant. We demonstrate the usefulness of this result for applications, including simulation schemes and PDE methods for nonnegative Volterra processes.

1 Introduction

Multifactor Markovian approximations for fractional Brownian motion were initially introduced by Carmona and Coutin [14] and revisited more recently by Abi Jaber and El Euch [3] in the context of nonnegative stochastic Volterra equations, motivated by rough and Volterra Heston models of El Euch and Rosenbaum [19] and Abi Jaber, Larsson, and Pulido [5]. Since then, substantial literature has emerged on such multifactor processes for numerical approximation methods (Alfonsi and Kebaier [7], Bayer and Breneis [11, 10, 12], Chevalier, Pulido, and Zúñiga [15], Harms [20]), deep learning approaches (Papapantoleon and Rou [22]), modeling (Abi Jaber [1]), and optimal control (Abi Jaber, Miller, and Pham [6]). It should be noted that such approximations are also heavily used in physics, chemistry and other fields, see, e.g., Baczewski and Bond [8], Bochud and Challet [13].

The starting point is a nonnegative solution to the stochastic Volterra equation

$$Y_t = Y_0 + \int_0^t K(t-s)b(Y_s)\,\mathrm{d}s + \int_0^t K(t-s)\sigma(Y_s)\,\mathrm{d}W_s, \tag{1}$$

where the kernel K is (approximated by) a weighted sum of exponentials of the form

$$K(t) = \sum_{i=1}^N w_i e^{-x_i t} \tag{2}$$

with positive nodes $\boldsymbol{x} = (x_i)_{i=1,\dots,N}$ and weights $\boldsymbol{w} = (w_i)_{i=1,\dots,N}$. Such series in terms of exponential functions are sometimes known as *Prony series*. The coefficients $b, \sigma : \mathbb{R} \to \mathbb{R}$ are continuous and satisfy the boundary conditions

$$b(0) \geq 0 \quad \text{and} \quad \sigma(0) = 0, \tag{3}$$

to ensure that the process Y remains nonnegative for any $Y_0 \geq 0$.

Then, the nonnegative Volterra process Y can be written in the form $Y_t = \sum_{i=1}^N w_i Y_t^{(i)}$, where $\boldsymbol{Y} = (Y^{(i)})_{i=1,\dots,N}$ is the solution to the N-dimensional stochastic differential equation

$$\mathrm{d}Y_t^{(i)} = -x_i \left(Y_t^{(i)} - y_0^{(i)}\right)\,\mathrm{d}t + b(Y_t)\,\mathrm{d}t + \sigma(Y_t)\,\mathrm{d}W_t, \tag{4}$$

with initial values $Y_0^{(i)}$, which are often, but not necessarily chosen to coincide with $y_0^{(i)}$, for $i = 1, \dots, N$.

The aim of the paper is to determine a state space of the multifactor Markovian process $\boldsymbol{Y}$. That is, we want to determine a set $\mathcal{D} \subseteq \mathbb{R}^N$ such that for every starting value $\boldsymbol{Y}_0 \in \mathcal{D}$, there exists a $\mathcal{D}$-valued solution $\boldsymbol{Y}$ to (4), that is $\boldsymbol{Y}_t \in \mathcal{D}$ for all $t \geq 0$ almost surely.

Beyond the mathematical importance of defining the state space of the Markovian process $\boldsymbol{Y}$, the knowledge of the state space is crucial for several practical applications, some of which have been considered so far:

DOI 10.20347/WIAS.PREPRINT.3162 Berlin 2025

- **Modeling and Calibration:** The multifactor model (4) can serve as a model in its own right (and not solely as an approximation of Volterra models), usually for stochastic volatility factors, as seen in the lifted Heston model of Abi Jaber [1]. Here, the knowledge of the state space is crucial for calibrating the initial values of $\boldsymbol{Y}_0$ directly to market data.
- **Simulation accuracy:** The identification of a valid state space allows for more precise simulation schemes for $\boldsymbol{Y}$. For instance, Bayer and Breneis [12] generalized a simulation scheme for the square-root process of Lileika and Mackevičius [21] to simulate paths from the dynamics in (4) with $\sigma(z) = \sqrt{z}$. However, the authors were not able to show that their simulation scheme is well-defined because they did not know the state space. Indeed, when simulating from (4), great care has to be taken to ensure that the aggregated process Y does not become negative, or if it does, one has to determine how to proceed with the square root term $\sqrt{Y}$.
- **Efficient Domain Meshing for PDE Solutions:** In Papapantoleon and Rou [22], for example, the authors use a rejection algorithm that discards simulations failing to satisfy the condition $\sum_i w_i Y_0^{(i)} \geq 0$, this approach is not precise and becomes inefficient in high dimensions.

For all these reasons, the geometry of the state space for multifactor processes $\boldsymbol{Y}$ quickly became a central focus in the associated literature for nonnegative Volterra processes. For the lifted Heston model of Abi Jaber [1], simulations demonstrated that while individual processes $Y^{(i)}$ might take negative values, their aggregated sum Y remains nonnegative. In this context, two key papers provided abstract characterizations of possible state spaces: one based on the resolvent of the first kind of the kernel by Abi Jaber and El Euch [2], and the other using the resolvent of the second kind of the kernel by Cuchiero and Teichmann [16] (we refer to Appendix B for further details on these state spaces and their connections). While these spaces are valid for a wide range of kernels, they remain somewhat abstract and challenging to make explicit, making it almost impossible to determine the good conditions on the initial values $\boldsymbol{Y} \geq 0$ of the process $\boldsymbol{Y}$. Finally, using the simulation algorithm for the rough Heston model due to Bayer and Breneis [12], we visually represent the process's support by a sample plot, see Figure 1 – replicating a similar plot already presented in that paper. One can clearly recognize that the sample paths do not only seem to lie in the half-plane $Y = w_1 Y^{(1)} + w_2 Y^{(2)} \geq 0$ marked with the down-ward oriented black line, but in an even smaller cone seemingly below the upward oriented line $\{\boldsymbol{y} \in \mathbb{R}^2 : y_1 \geq y_2\}$.

Main contributions. The main question we are interested in can be summarized as follows:

What constitutes a suitable state space for the multifactor Markovian process $\boldsymbol{Y}$ in (4)*?*

Our main results in Theorem 2.3 and Corollary 2.7 establish that this state space can be represented as a linear transformation of $\mathbb{R}_+^N$, and we provide an explicit form for this transformation.

As a first application of this result, we prove that the weak simulation scheme for the rough Heston process proposed by Bayer and Breneis [12] is well-defined in the sense that the variance process always stays non-negative, see Section 3. In Section 4, we derive the corresponding pricing PDE on the transformed domain $\mathbb{R}_+^N$, and solve it numerically by the finite element method after truncation of the domain. Naturally, knowing the PDE's precise domain is crucial for accurate numerical approximations. Finally, in Section B, we show how the explicit formula for the domain compares with general, abstract characterizations given in the literature by Abi Jaber and El Euch [2] and Cuchiero and Teichmann [16].

Notation and Conventions. We denote by $\boldsymbol{e}_i$ is the i-th unit vector, which has a 1 in the i-th component, and 0 in every other component, $\boldsymbol{1} := (1, 1, \ldots, 1)^\top$ is the vector with 1 in every component, Id is the identity matrix, diag$(\boldsymbol{a})$ for a vector $\boldsymbol{a} \in \mathbb{R}^N$ is the diagonal matrix with entries $\boldsymbol{a}$ in the diagonal, and $\overline{w} := \boldsymbol{1}^\top \boldsymbol{w} = \sum_{i=1}^N w_i$. Throughout, italic letters a denote real numbers and bold letters $\boldsymbol{a}$ denote vectors, where we write $\boldsymbol{a} = (a_i)_{i=1}^N$ for the components of $\boldsymbol{a}$. An exception are stochastic processes, where components are denoted by $\boldsymbol{Y}_t = (Y_t^{(i)})_{i=1}^N$ (due to the time variable in the subscript).

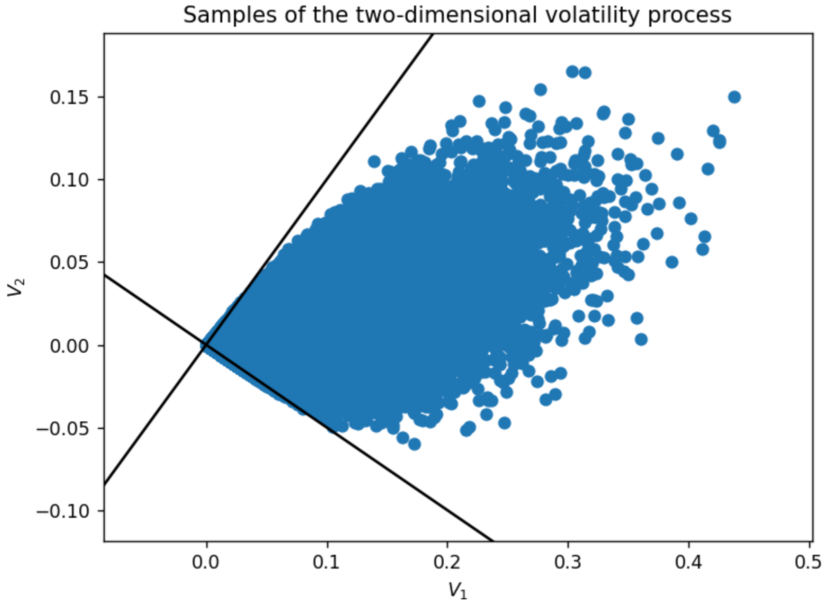

Figure 1: Samples of the two-dimensional process $(Y^{(1)}, Y^{(2)})$ using 10^5 sample paths on a time grid with $M = 1000$ time steps. Plotted are all the points $\boldsymbol{Y}_{t_i}$ for every time step t_i, $i = 0, \ldots, 1000$, and all the 10^5 samples. The decreasing black line is the line where the aggregated process $\boldsymbol{w}_1 Y^{(1)} + w_2 Y^{(2)} = 0$, and the aggregated process is positive above that line. The nearly orthogonal second line cuts out a cone, which seems to give the actual support of the process.

2 State spaces of the multifactor Markovian process

Fix $N \geq 1$. We consider the N-dimensional stochastic differential equation

$$\mathrm{d}\boldsymbol{Y}_t = -\text{diag}(\boldsymbol{x})\left(\boldsymbol{Y}_t - \boldsymbol{y}_0\right)\mathrm{d}t + b(\boldsymbol{w}^\top \boldsymbol{Y}_t)\mathbf{1}\,\mathrm{d}t + \sigma(\boldsymbol{w}^\top \boldsymbol{Y}_t)\mathbf{1}\,\mathrm{d}W_t, \tag{5}$$

where $b, \sigma : \mathbb{R} \to \mathbb{R}$ are continuous, satisfy the linear growth condition

$$|b(y)| \vee |\sigma(y)| \leq C\left(1 + |y|\right), \quad y \in \mathbb{R}, \tag{6}$$

and the boundary conditions (3). The speeds of mean-reversion $\boldsymbol{x} = (x_i)_{i=1,\ldots,N}$ are positive and ordered, i.e. $0 < x_1 \leq x_2 \leq \cdots \leq x_N$, the weights $\boldsymbol{w} = (w_i)_{i=1,\ldots,N}$ are positive, W is a one-dimensional Brownian motion, and where $\boldsymbol{Y}_0, \boldsymbol{y}_0 \in \mathbb{R}^N$ may be different. This corresponds to (4) written in vector form.

The aim of this section is to determine a state space of the multifactor process $\boldsymbol{Y}$. That is, we want to determine a set $\mathcal{D} \subseteq \mathbb{R}^N$ such that for every starting value $\boldsymbol{Y}_0 \in \mathcal{D}$, there exists a $\mathcal{D}$-valued weak solution $\boldsymbol{Y}$ to (5), that is $\boldsymbol{Y}_t \in \mathcal{D}$ for all $t \geq 0$ almost surely. In particular, the domain $\mathcal{D}$ should be a subset of the half-plane $\{\boldsymbol{y} \in \mathbb{R}^N : \boldsymbol{w}^\top \boldsymbol{y} \geq 0\}$ to ensure non-negativity of the aggregated weighted process $Y := \boldsymbol{w}^\top \boldsymbol{Y}$, which for the specific case $\boldsymbol{y}_0 = \boldsymbol{Y}_0$ would correspond the Volterra process (1) for the weighted sum of exponential kernel K given in (2). As illustrated in Figure 1, and following the abstract characterizations of domains of (possibly infinite-dimensional) lifts of nonnegative Volterra processes in Abi Jaber and El Euch [2] and Cuchiero and Teichmann [16], we suspect the domain $\mathcal{D}$ to be a cone.

2.1 Main result

We prove that the domain $\mathcal{D}$ is a cone characterized by the set $\mathcal{Q}$ of admissible matrices, which is defined as follows.

Definition 2.1. A matrix $Q \in \mathbb{R}^{N\times N}$ is called admissible if it satisfies the following assumptions:

1 Q is invertible,

2 $\boldsymbol{e}_N^\top Q = \boldsymbol{w}^\top$,

3 $Q\mathbf{1} = \overline{w}\boldsymbol{e}_N$, where $\overline{w} := \boldsymbol{w}^\top \mathbf{1}$,

4 $(Q\text{diag}(\boldsymbol{x})Q^{-1})_{i,j} \le 0$ for $i, j \in \{1, \dots, N\}$ with $i \neq j$.

We denote by $\mathcal{Q}$ the set of all admissible matrices.

Before stating our main theorem, we first show that the set of admissible matrices $\mathcal{Q}$ is nonempty by providing an explicit example of an admissible matrix. However, $\mathcal{Q}$ is not reduced to a singleton as shown in Example 2.10 below.

Theorem 2.2. *The matrix $Q = (q_{i,j})_{i,j=1,\dots N} \in \mathbb{R}^{N\times N}$ given by*

$$q_{i,j} = w_j, \quad j \le i, \qquad q_{i,i+1} = -\sum_{j=1}^{i} w_j, \quad i = 1, \dots, N-1,$$

and zeros elsewhere is admissible. In particular, $\mathcal{Q}$ is nonempty.

Proof. The proof is given in Appendix C.1. □

We are now ready to state our main theorem that gives state spaces of the multifactor Markovian process (5).

Theorem 2.3. *Let $b, \sigma : \mathbb{R} \to \mathbb{R}$ be continuous functions satisfying the linear growth condition* (6) *and the boundary conditions* (3). *Let $Q \in \mathcal{Q}$ be an admissible matrix in the sense of Definition 2.1 and suppose that*

$$\boldsymbol{y}_0 = \mu\text{diag}(\boldsymbol{x})^{-1}\mathbf{1} \quad \textit{for some } \mu \ge 0. \tag{7}$$

Set $\mathcal{D} = Q^{-1}\mathbb{R}_+^N$. Then, for each $\boldsymbol{Y}_0 \in \mathcal{D}$, there exists a $\mathcal{D}$-valued weak solution $\boldsymbol{Y}$ to (5).

Proof. The proof is given in Section 2.4. □

Example 2.4. For the admissible matrix Q given in Theorem 2.2, the set $\mathcal{D} = Q^{-1}\mathbb{R}_+^N$ corresponds to the set of $\boldsymbol{y} \in \mathbb{R}^N$ such that $\boldsymbol{w}^\top \boldsymbol{y} \ge 0$ and

$$\sum_{j=1}^{i} w_j y_j \ge \sum_{j=1}^{i} w_j y_{i+1} \qquad \text{for} \qquad i = 1, \dots, N-1.$$

Remark 2.5. The domain $Q^{-1}\mathbb{R}_+^N$ is not unique, see Appendix A.

In practice, for instance for the multifactor approximations of Volterra processes, we often have $\boldsymbol{Y}_0 = \boldsymbol{y}_0$. Hence, it may be interesting to know whether $\boldsymbol{y}_0$ as given in (7) is in $Q^{-1}\mathbb{R}_+^N$. We treat this question in a slightly more general context in the following lemma.

Lemma 2.6. *Let Q be an admissible matrix and let $\boldsymbol{y}_0 \in \mathbb{R}^N$ satisfy* (7). *Then $\boldsymbol{y}_0 \in Q^{-1}\mathbb{R}_+^N$.*

Proof. The proof is given in Appendix C.2. □

We remark that condition (7) on $\boldsymbol{y}_0$ can easily be dropped by using affine transformations of $\mathbb{R}_+^N$ instead of linear ones. We give the specific details in the next corollary.

Corollary 2.7. *Let $b, \sigma : \mathbb{R} \to \mathbb{R}$ be continuous functions satisfying the linear growth condition* (6) *and the boundary conditions* (3)*. Let $Q \in \mathcal{Q}$ be an admissible matrix in the sense of Definition 2.1, and assume that $\boldsymbol{w}^\top \boldsymbol{y}_0 \geq 0$. Let $\widetilde{\boldsymbol{y}}_0$ be chosen according to* (7) *with $\boldsymbol{w}^\top \widetilde{\boldsymbol{y}}_0 = \boldsymbol{w}^\top \boldsymbol{y}_0$. Define the set*

$$\mathcal{D} = Q^{-1}\mathbb{R}_+^N + (\boldsymbol{y}_0 - \widetilde{\boldsymbol{y}}_0). \tag{8}$$

Then, for any $\boldsymbol{Y}_0 \in \mathcal{D}$, there exists a $\mathcal{D}$-valued weak solution $\boldsymbol{Y}$ to (5)*.*

Proof. Denote by $\widetilde{\boldsymbol{Y}}$ a weak solution to

$$\mathrm{d}\widetilde{\boldsymbol{Y}}_t = -\text{diag}(\boldsymbol{x}) \left(\widetilde{\boldsymbol{Y}}_t - \widetilde{\boldsymbol{y}}_0\right) \mathrm{d}t + b(\boldsymbol{w}^\top \widetilde{\boldsymbol{Y}}_t)\mathbf{1}\,\mathrm{d}t + \sigma(\boldsymbol{w}^\top \widetilde{\boldsymbol{Y}}_t)\mathbf{1}\,\mathrm{d}W_t$$

with initial condition $\widetilde{\boldsymbol{Y}}_0 := \boldsymbol{Y}_0 + \widetilde{\boldsymbol{y}}_0 - \boldsymbol{y}_0$. Note that $\widetilde{\boldsymbol{Y}}_0 \in Q^{-1}\mathbb{R}_+^N$, so Theorem 2.3 imply the existence of such a solution $\widetilde{\boldsymbol{Y}}$ that stays in $Q^{-1}\mathbb{R}_+^N$. Define the process $\boldsymbol{R} := \widetilde{\boldsymbol{Y}} + \boldsymbol{y}_0 - \widetilde{\boldsymbol{y}}_0$ and note that $\boldsymbol{w}^\top \boldsymbol{R} = \boldsymbol{w}^\top \widetilde{\boldsymbol{Y}}$. Thus, $\boldsymbol{R}$ satisfies $\boldsymbol{R}_0 = \boldsymbol{Y}_0$ and

$$\mathrm{d}\boldsymbol{R}_t = -\text{diag}(\boldsymbol{x}) \left(\boldsymbol{R}_t - \boldsymbol{y}_0\right) \mathrm{d}t + b(\boldsymbol{w}^\top \boldsymbol{R}_t)\mathbf{1}\,\mathrm{d}t + \sigma(\boldsymbol{w}^\top \boldsymbol{R}_t)\mathbf{1}\,\mathrm{d}W_t.$$

But this means that $\boldsymbol{R}$ is a solution to (5) that stays in $\mathcal{D}$. The corollary follows immediately. ☐

We now give the specific result for the multifactor square-root process.

Example 2.8. Consider the multifactor square-root model

$$\mathrm{d}\boldsymbol{V}_t^N = -\text{diag}(\boldsymbol{x}) \left(\boldsymbol{V}_t^N - \boldsymbol{v}_0\right) \mathrm{d}t + \left(\theta - \lambda \boldsymbol{w}^\top \boldsymbol{V}_t^N\right)\mathbf{1}\,\mathrm{d}t + \nu\sqrt{\boldsymbol{w}^\top \boldsymbol{V}_t^N}\mathbf{1}\,\mathrm{d}W_t. \tag{9}$$

Let Q be an admissible matrix, and assume that $\boldsymbol{w}^\top \boldsymbol{v}_0 \geq 0$. Then,

$$\mathcal{D} = Q^{-1}\mathbb{R}_+^N + \left(\boldsymbol{v}_0 - \frac{\boldsymbol{w}^\top \boldsymbol{v}_0}{\boldsymbol{w}^\top \text{diag}(\boldsymbol{x})^{-1}\mathbf{1}} \text{diag}(\boldsymbol{x})^{-1}\mathbf{1}\right).$$

2.2 Link with nonnegative Volterra processes

As an application of our result, one can obtain the existence of nonnegative solutions to Volterra equations with kernels of the form (1). We note that such existence can be obtained by working directly on the level of the Volterra equation as done in Abi Jaber, Larsson, and Pulido [5, Theorem 3.6 and Example 3.7]. Here, our result provides another alternative as illustrated in the following corollary.

Corollary 2.9. *Let $b, \sigma : \mathbb{R} \to \mathbb{R}$ be continuous functions satisfying the linear growth condition* (6) *and the boundary conditions* (3)*. Let the kernel K be given by a weighted sum of exponentials as in* (2)*. Then, for each $Y_0 \geq 0$, the stochastic Volterra equation* (1) *admits a nonnegative weak solution Y.*

Proof. Fix $Y_0 \geq 0$ and let $\boldsymbol{y}_0 \in \mathbb{R}^N$ be such that $\boldsymbol{w}^\top \boldsymbol{y}_0 = Y_0$. Let $\widetilde{\boldsymbol{y}}_0$ be chosen according to (7) with $\boldsymbol{w}^\top \widetilde{\boldsymbol{y}}_0 = \boldsymbol{w}^\top \boldsymbol{y}_0$. Let $Q \in \mathcal{Q}$ be an admissible matrix, for instance given by Theorem 2.2. Then, it follows from Lemma 2.6 that $\widetilde{\boldsymbol{y}}_0 \in Q^{-1}\mathbb{R}$. Hence, $\boldsymbol{y}_0 = \widetilde{\boldsymbol{y}}_0 + (\boldsymbol{y}_0 - \widetilde{\boldsymbol{y}}_0) \in \mathcal{D}$, with $\mathcal{D}$ given by (8). An application of Corollary 2.7, with the starting value $Y_0 = \boldsymbol{y}_0 \in \mathcal{D}$, yields the existence of a $\mathcal{D}$-valued weak solution $\boldsymbol{Y}$ to the equation (5). Thanks to the variation of constants formula, we can re-write the equation in the form

$$\boldsymbol{Y}_t = \boldsymbol{y}_0 + \int_0^t \exp(-\text{diag}(\boldsymbol{x})(t-s))\mathbf{1}\left(b(\boldsymbol{w}^\top \boldsymbol{Y}_s)\,\mathrm{d}s + \sigma(\boldsymbol{w}^\top \boldsymbol{Y}_s)\,\mathrm{d}W_s\right), \tag{10}$$

so that the process Y defined by $Y = \boldsymbol{w}^\top \boldsymbol{Y}$ solves the equation

$$Y_t = Y_0 + \int_0^t \sum_{i=1}^N w_i e^{-x_i(t-s)} \left(b(Y_s)\,\mathrm{d}s + \sigma(Y_s)\,\mathrm{d}W_s\right),$$

which is precisely the Volterra equation (1) with the kernel K given by (2). It remains to argue that, for all $t \geq 0$, Y_t remains nonnegative by using the fact that $\boldsymbol{Y}_t \in \mathcal{D}$. Indeed, using Condition 2 of Definition 2.1 and the fact that $\boldsymbol{w}^\top \widetilde{\boldsymbol{y}}_0 = \boldsymbol{w}^\top \boldsymbol{y}_0$, we obtain that

$$\boldsymbol{w}^\top \mathcal{D} = \boldsymbol{e}_N^\top Q Q^{-1} \mathbb{R}_+^N + (\boldsymbol{w}^\top \boldsymbol{y}_0 - \boldsymbol{w}^\top \widetilde{\boldsymbol{y}}_0) = \boldsymbol{e}_N^\top \mathbb{R}_+^N = \mathbb{R}_+.$$

Hence, for all $t \geq 0$, $Y_t = \boldsymbol{w}^\top \boldsymbol{Y}_t \in \boldsymbol{w}^\top \mathcal{D} = \mathbb{R}_+$, which ends the proof. □

2.3 On admissible matrices for $N \in \{2, 3\}$

In this section we give examples of admissible matrices.

Example 2.10. In the case $N = 2$, the Conditions 2 and 3 of Definition 2.1 imply that we are looking for a matrix of the form

$$Q = \begin{pmatrix} q & -q \\ w_1 & w_2 \end{pmatrix},$$

for some $q \neq 0$ to ensure invertibility. Then,

$$Q\text{diag}(\boldsymbol{x})Q^{-1} = \frac{1}{\overline{w}} \begin{pmatrix} w_1 x_2 + w_2 x_1 & (x_1 - x_2) q \\ w_1 w_2 (x_1 - x_2) q^{-1} & w_1 x_1 + w_2 x_2 \end{pmatrix}.$$

Since $x_1 \leq x_2$, the last condition in Definition 2.1 is satisfied for any $q > 0$, and indeed, the domain $Q^{-1}\mathbb{R}_+^2$ is independently of the precise choice of q given by

$$\mathcal{D} = \left\{ \boldsymbol{y} \in \mathbb{R}_+^2 : \boldsymbol{w}^\top \boldsymbol{y} \geq 0,\ y_1 \geq y_2 \right\}. \tag{11}$$

For the case of the multifactor square-root process (9), the resulting sample paths of $\boldsymbol{V}^2$ and $\boldsymbol{U} := Q\boldsymbol{V}^2$ are illustrated in Figure 2. Note that we chose the large maturity $T = 100$ to give the process more time to explore its domain. Thereby, it is more clearly visible that the domain of $\boldsymbol{U}$ is indeed $\mathbb{R}_+^2$, than if we had set $T = 1$.

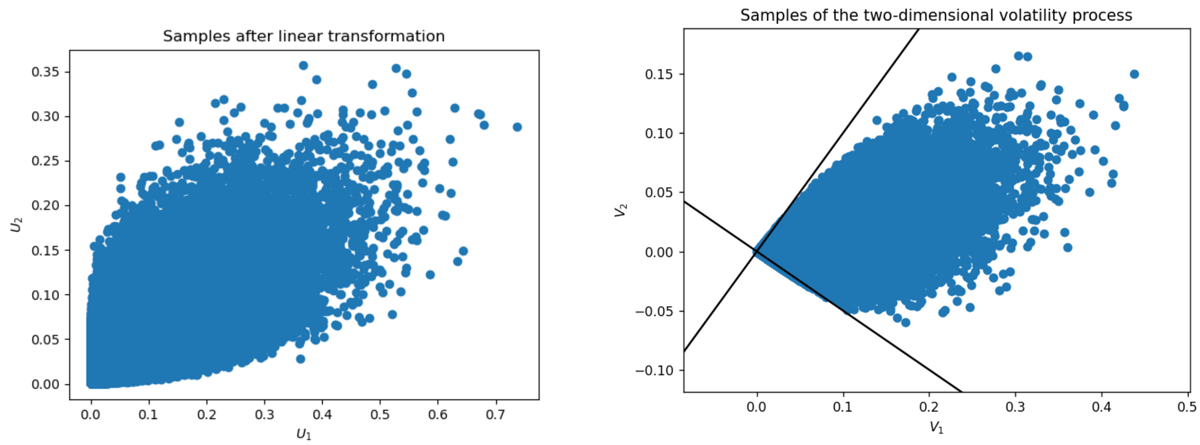

Figure 2: Samples of $\boldsymbol{V}^2$ (right) and $\boldsymbol{U}$ (left) using 10^3 sample paths on a time grid with $M = 10^5$ time steps. The black lines correspond to the hyperplanes in (11). The parameters used are $\boldsymbol{x} = (1, 10), \boldsymbol{w} = (1, 2), \lambda = 0.3, \nu = 0.3, V_0 = 0.02, \theta = 0.02, T = 100$, and $\boldsymbol{v}_0 = V_0/(2\boldsymbol{x})(w_1/x_1 + w_2/x_2)$, i.e. $\boldsymbol{v}_0$ is chosen to be proportional to $\boldsymbol{x}^{-1}$.

For $N = 3$, a similar computation – relegated to the appendix due to its length – gives multiple choices of domains. See Appendix A for details.

2.4 Proof of Theorem 2.3

Fix an admissible matrix $Q \in \mathcal{Q}$. The main idea of the proof is to reduce the study to the process $\boldsymbol{Z} = Q\boldsymbol{Y}$ and prove that its associated stochastic differential equation admits an $\mathbb{R}_+^N$-valued solution.

We start by writing the stochastic differential equation for $\boldsymbol{Z}$. For this we first observe that due to (7), $\boldsymbol{Y}$ satisfies

$$\mathrm{d}\boldsymbol{Y}_t = -\text{diag}(\boldsymbol{x})\boldsymbol{Y}_t\,\mathrm{d}t + b_\mu(\boldsymbol{w}^\top \boldsymbol{Y}_t)\mathbf{1}\,\mathrm{d}t + \sigma(\boldsymbol{w}^\top \boldsymbol{Y}_t)\mathbf{1}\,\mathrm{d}W_t$$

where $b_\mu(z) = b(z) + \mu$. Using Q as a transformation of basis (in the sense $\boldsymbol{Z} = Q\boldsymbol{Y}$), we get, thanks to the invertibility of Q, the following stochastic differential equation

$$\mathrm{d}\boldsymbol{Z}_t = -Q\text{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{Z}_t\,\mathrm{d}t + b_\mu(\boldsymbol{w}^\top Q^{-1}\boldsymbol{Z}_t)Q\mathbf{1}\,\mathrm{d}t + \sigma(\boldsymbol{w}^\top Q^{-1}\boldsymbol{Z}_t)Q\mathbf{1}\,\mathrm{d}W_t.$$

Using the admissibility conditions 2 and 3 in Definition 2.1, we have that $\boldsymbol{w}^\top Q^{-1} = \boldsymbol{e}_N^\top QQ^{-1} = \boldsymbol{e}_N^\top$ and $Q\mathbf{1} = \overline{w}\boldsymbol{e}_N$, which simplifies the equation to

$$\mathrm{d}\boldsymbol{Z}_t = -Q\text{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{Z}_t\,\mathrm{d}t + \overline{w}b_\mu(Z_t^{(N)})\boldsymbol{e}_N\,\mathrm{d}t + \overline{w}\sigma(Z_t^{(N)})\boldsymbol{e}_N\,\mathrm{d}W_t. \tag{12}$$

Recall that $Z^{(N)}$ is the N-th component of $\boldsymbol{Z}$.

In order to prove Theorem 2.3, it suffices to prove that for each $\boldsymbol{Z}_0 \in \mathbb{R}_+^N$, there exists an $\mathbb{R}_+^N$-valued $\boldsymbol{Z}$ weak solution to (12). In particular, this would hold for any initial value of the form $\boldsymbol{Z}_0 = Q\boldsymbol{Y}_0$ with $Y_0 \in \mathcal{D} = Q^{-1}\mathbb{R}_N^+$ and setting $\boldsymbol{Y} = QZ$, one obtains a $\mathcal{D}$-valued weak solution $\boldsymbol{Y}$ to (5) started at $\boldsymbol{Y}_0$.

Hence, this boils down to establish that the set $\mathbb{R}_+^N$ is stochastically viable with respect to the equation (12). Viability and invariance theory for stochastic differential equations have been extensively studied in the literature in various contexts and with different assumptions on the domain and the coefficients, we refer to Abi Jaber, Bouchard, and Illand [4], Da Prato and Frankowska [17, 18] and the references therein.

For the non-negative orthant $\mathbb{R}_+^N$ the characterization in terms of the coefficients is very simple and means that, at boundary points, the diffusive coefficient has to be tangential to the boundary and the drift inward pointing. This is summarized in the following lemma.

Lemma 2.11. *Let $\tilde{b}, \tilde{\sigma} : \mathbb{R}^N \to \mathbb{R}^N$ be continuous satisfying the growth conditions*

$$\|\tilde{b}(\boldsymbol{z})\| + \|\tilde{\sigma}(\boldsymbol{z})\| \leq L(1 + \|\boldsymbol{y}\|), \qquad \boldsymbol{z} \in \mathbb{R}^N,$$

and the boundary conditions, for all $\boldsymbol{z} \in \mathbb{R}_+^N$,

$$z_i = 0 \ \Rightarrow\ \boldsymbol{e}_i^\top \tilde{b}(\boldsymbol{z}) \geq 0 \text{ and } \boldsymbol{e}_i^\top \tilde{\sigma}(\boldsymbol{z}) = 0, \quad i = 1, \ldots, N, \tag{13}$$

then, for each $\widetilde{\boldsymbol{Z}}_0 \in \mathbb{R}_+^N$, there exists a weak $\mathbb{R}_+^N$-valued solution to the following stochastic differential equation

$$\mathrm{d}\widetilde{\boldsymbol{Z}}_t = \tilde{b}\left(\widetilde{\boldsymbol{Z}}_t\right)\,\mathrm{d}t + \tilde{\sigma}\left(\widetilde{\boldsymbol{Z}}_t\right)\,\mathrm{d}W_t.$$

Proof. See for instance Da Prato and Frankowska [18, Example 2.7]. □

We now proceed to the proof of Theorem 2.3.

Proof of Theorem 2.3. It remains to apply Lemma 2.11 on the equation (12). For this, we define

$$\tilde{b}(\boldsymbol{z}) = -Q\text{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{z} + \overline{w}b_\mu(z_N)\boldsymbol{e}_N \quad \text{and} \quad \tilde{\sigma}(\boldsymbol{z}) = \overline{w}\sigma(z_N)\boldsymbol{e}_N, \quad z \in \mathbb{R}^N.$$

Then, it readily follows from the continuity and growth conditions of b and σ that $\tilde{b}, \tilde{\sigma}$ are also continuous with at most linear growth conditions. As for the boundary conditions (13), we fix $\boldsymbol{z} \in \mathbb{R}_+^N$ such that $z_i = 0$ for some $i = 1, \ldots, N$.

• For the diffusion term, we have

$$\boldsymbol{e}_i^\top \tilde{\sigma}(\boldsymbol{z}) = \overline{w}\sigma(z_N)\boldsymbol{e}_i^\top \boldsymbol{e}_N = 0,$$

since $\boldsymbol{e}_i^\top \boldsymbol{e}_N = 0$ if $i < N$ and $\sigma(z_N) = \sigma(0) = 0$ if $i = N$, where we used the boundary condition on σ in (3).

• For the drift term, we first observe that for the same reason $b_\mu(z_N)\boldsymbol{e}_i^\top \boldsymbol{e}_N = (b(z_N) + \mu)\boldsymbol{e}_i^\top \boldsymbol{e}_N \geq 0$, since $b(0) + \mu \geq 0$ thanks to the boundary condition on b in (3) and the fact that $\mu \geq 0$, so that we can write

$$\begin{aligned}
\boldsymbol{e}_i^\top \tilde{b}(\boldsymbol{z}) &= -\boldsymbol{e}_i^\top Q\text{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{z} + \overline{w}b_\mu(z_N)\boldsymbol{e}_i^\top \boldsymbol{e}_N \\
&\geq -\sum_{j\neq i}(Q\text{diag}(\boldsymbol{x})Q^{-1})_{ij}z_j \\
&\geq 0,
\end{aligned}$$

where the first inequality follows from $z_i = 0$ and the second inequality follows from the admissibility condition 4 in Definition 2.1 for the matrix Q and the fact that $z_j \geq 0$.

This shows that the boundary conditions (13) are satisfied by $\tilde{b}, \tilde{\sigma}$, so that an application of Lemma 2.11 yields the existence of an $\mathbb{R}_+^N$-valued solution $\boldsymbol{Z}$ to (12) for any initial condition $\boldsymbol{Z}_0 \in \mathbb{R}_+^N$. In particular, it holds for the initial value $\boldsymbol{Z}_0 = Q\boldsymbol{Y}_0$ with $Y_0 \in \mathcal{D} = Q^{-1}\mathbb{R}_N^+$. Setting $\boldsymbol{Y} = QZ$, one obtains a $\mathcal{D}$-valued weak solution $\boldsymbol{Y}$ to (5) started at $\boldsymbol{Y}_0$ and ends the proof of theorem. □

3 The weak scheme is cone-preserving

In Section 2, we determined the state space $\mathcal{D} \subseteq \mathbb{R}^N$ of the multifactor square-root process $\boldsymbol{V}^N$ given by (9). Assume now that we approximate the process $\boldsymbol{V}^N$ using the weak simulation scheme proposed in Bayer and Breneis [12]. The goal of this section is to prove that the resulting approximation has the same viable domain $\mathcal{D}$ as $\boldsymbol{V}^N$.

Let us first start by recalling the weak simulation scheme of Bayer and Breneis [12]. First, the SDE in (9) is split into two parts, one containing the drift and the other the diffusion. Denote by $D(\boldsymbol{z}, h) := \boldsymbol{Z}_h := (Z_h^{(i)})_{i=1}^N$ the solution at time h of the ordinary differential equation (ODE)

$$\mathrm{d}Z_t^i = -x_i(Z_t^i - v_0^i)\,\mathrm{d}t + (\theta - \lambda Z_t)\,\mathrm{d}t, \quad Z_0^i = z^i, \quad i = 1, \ldots, N, \quad Z_t = \boldsymbol{w}^\top \boldsymbol{Z}_t, \tag{14}$$

and by $S(\boldsymbol{y}, h) := \boldsymbol{Y}_h := Y_h^i$ the solution at time h of the SDE

$$\mathrm{d}Y_t^i = \nu\sqrt{Y_t}\,\mathrm{d}W_t, \qquad Y_0^i = y^i, \qquad i = 1, \ldots, N, \qquad Y_t = \boldsymbol{w}^\top \boldsymbol{Y}_t. \tag{15}$$

Then, the ODE (14) is linear and can hence be solved exactly. Therefore, the simulation scheme $\widehat{D}$ for the ODE is simply given by

$$\widehat{D}(\boldsymbol{z}, h) := D(\boldsymbol{z}, h) := e^{Ah}\boldsymbol{z} + A^{-1}(e^{Ah} - \text{Id})b,$$

where

$$A := -\lambda \boldsymbol{1}\boldsymbol{w}^\top - \text{diag}(\boldsymbol{x}), \quad \text{and} \quad b := \theta\boldsymbol{1} + \text{diag}(\boldsymbol{x})\boldsymbol{v}_0.$$

We now recall the simulation scheme for the SDE (15). Note that the right-hand side of (15) is the same for all i. Thus, after multiplying (15) with $\boldsymbol{w}$, we get

$$dY_t = \nu\overline{w}\sqrt{Y_t}\,\mathrm{d}W_t, \qquad Y_0 = \boldsymbol{w}^\top \boldsymbol{y},$$

where $\overline{w} := \boldsymbol{1}^\top \boldsymbol{w}$. This is now a one-dimensional SDE, which was already studied in Lileika and Mackevičius [21], where a second-order simulation scheme was given. This scheme is based on matching the first 5 moments, while preserving the non-negativity of Y. Define the quantities

$$
\begin{aligned}
& x := \boldsymbol{w}^\top \boldsymbol{y}, \qquad z := \nu^2 \overline{w}^2 h, && (16)\\
& m_1 := x, \qquad m_2 := x^2 + xz, \qquad m_3 := x^3 + 3x^2 z + \frac{3}{2} x z^2, \\
& p_1 := \frac{m_1 x_2 x_3 - m_2(x_2 + x_3) + m_3}{x_1(x_3 - x_1)(x_2 - x_1)}, \\
& p_2 := \frac{m_1 x_1 x_3 - m_2(x_1 + x_3) + m_3}{x_2(x_3 - x_2)(x_1 - x_2)}, \\
& p_3 := \frac{m_1 x_1 x_2 - m_2(x_1 + x_2) + m_3}{x_3(x_1 - x_3)(x_2 - x_3)}, \\
& x_1 := x + \left(A + \frac{3}{4}\right) z - \sqrt{\left(3x + \left(A + \frac{3}{4}\right)^2 z\right) z}, && (17)\\
& x_2 := x + Az, && (18)\\
& x_3 := x + \left(A + \frac{3}{4}\right) z + \sqrt{\left(3x + \left(A + \frac{3}{4}\right)^2 z\right) z}, && (19)\\
& A := \frac{3 + \sqrt{3}}{4}.
\end{aligned}
$$

Then, we define $\widehat{Y}_h$ to be the random variable which is x_i with probability p_i, $i = 1, 2, 3$.

We can now reconstruct an approximation $\widehat{\boldsymbol{Y}}$ from $\widehat{Y}$. Indeed, since the right-hand side of (15) is the same for all $i = 1, \dots, N$, the solution of (15) must be of the form

$$
Y_h^i = y^i + R, \qquad i = 1, \dots, N, \tag{20}
$$

for some scalar random variable R. Taking the inner product of (20) with $\boldsymbol{w}$, we get

$$
Y_h = \boldsymbol{w}^\top \boldsymbol{y} + \overline{w} R, \quad \text{implying} \quad R = \frac{Y_h - \boldsymbol{w}^\top \boldsymbol{y}}{\overline{w}}.
$$

Hence, we set

$$
\widehat{S}(\boldsymbol{y}, h) := \widehat{\boldsymbol{Y}}_h := \boldsymbol{y} + \frac{\widehat{Y}_h - \boldsymbol{w}^\top \boldsymbol{y}}{\overline{w}}.
$$

Finally, we use Strang splitting to get the scheme

$$
A^{\mathsf{CIR}}(\boldsymbol{v}, h) := D\left(\widehat{S}\left(D\left(\boldsymbol{v}, \frac{h}{2}\right), h\right), \frac{h}{2}\right)
$$

for approximating $\boldsymbol{V}_h$ given $\boldsymbol{v}$. Therefore, we get a simulation algorithm

$$
\boldsymbol{V}_{t_{j+1}}^{N,M} := A^{\mathsf{CIR}}(\boldsymbol{V}_{t_j}^{N,M}, t_{j+1} - t_j), \qquad j = 0, \dots, M-1,
$$

where $0 = t_0 < t_1 < \cdots < t_M = T$. The only problem that could occur is that the square root in (17) or (19) is not well-defined. However, note that if we can prove that $\boldsymbol{V}^{N,M}$ does not leave $\mathcal{D}$, where $\mathcal{D}$ is the same cone as in Theorem 2.3, then in particular, $x = \boldsymbol{w}^\top \boldsymbol{y}$ in (16) will always be non-negative, and hence the square roots in (17) and (19) are always well-defined. Proving that $\boldsymbol{V}^{N,M}$ stays in $\mathcal{D}$ is the aim of the following theorem.

Theorem 3.1. *Let Q be an admissible matrix and let $\boldsymbol{v}_0$ be chosen according to* (7)*. Then, for all $\boldsymbol{v} \in Q^{-1}\mathbb{R}_+^N$ and $h \geq 0$, the weak simulation algorithm A^{CIR} described above satisfies $A^{\mathsf{CIR}}(\boldsymbol{v}, h) \in Q^{-1}\mathbb{R}_+^N$. In particular, A^{CIR} is well-defined.*

Proof. Given $\boldsymbol{z}, \boldsymbol{y} \in Q^{-1}\mathbb{R}_+^N$ and $h \geq 0$, we want to show that $D(\boldsymbol{z}, h) \in Q^{-1}\mathbb{R}_+^N$, and $\widehat{S}(\boldsymbol{y}, h) \in Q^{-1}\mathbb{R}_+^N$. This will prove the theorem.

Consider first the algorithm $\widehat{S}$. Recall that $\widehat{S}(\boldsymbol{y}, h) = \boldsymbol{y} + R\mathbf{1}$ for some scalar random variable R. We have to verify that $Q\widehat{S}(\boldsymbol{y}, h) = Q\boldsymbol{y} + RQ\mathbf{1} \in \mathbb{R}_+^N$. The last component of this vector is given by

$$(Q\widehat{S}(\boldsymbol{y}, h))_N = \boldsymbol{w}^\top \boldsymbol{y} + R\overline{w},$$

and we recall that this was given by the random variable $\widehat{Y}_h$ in Section 3, which by definition is non-negative, as verified in Lileika and Mackevičius [21]. Conversely, for $i = 1, \dots, N-1$, we have

$$(Q\widehat{S}(\boldsymbol{y}, h))_i = (Q\boldsymbol{y})_i + 0 \geq 0$$

by the assumption that $Q\boldsymbol{y} \in \mathbb{R}_N^+$. Hence, $\widehat{S}$ leaves the domain $Q^{-1}\mathbb{R}_+^N$ invariant.

Next, consider the algorithm D. Recall that $D(\boldsymbol{z}, h)$ was given as the exact solution at time h of the ODE

$$\mathrm{d}\boldsymbol{Z}_t = -\mathrm{diag}(\boldsymbol{x})(\boldsymbol{Z}_t - \boldsymbol{v}_0)\,\mathrm{d}t + (\theta - \lambda \boldsymbol{w}^\top \boldsymbol{Z}_t)\mathbf{1}\,\mathrm{d}t, \qquad \boldsymbol{Z}_0 = \boldsymbol{z}.$$

Note that due to (7), $\mathrm{diag}(\boldsymbol{x})\boldsymbol{v}_0 = \mu\mathbf{1}$ for some $\mu \geq 0$. Defining $\widetilde{\boldsymbol{Z}} \coloneqq Q\boldsymbol{Z}$, we have

$$\mathrm{d}\widetilde{\boldsymbol{Z}}_t = \left(-Q\mathrm{diag}(\boldsymbol{x})Q^{-1}\widetilde{\boldsymbol{Z}}_t + (\theta + \mu - \lambda \widetilde{Z}_t^N)Q\mathbf{1}\right)\mathrm{d}t, \qquad \widetilde{\boldsymbol{Z}}_0 = Q\boldsymbol{z} \in \mathbb{R}_+^N,$$

and we have to show that $\widetilde{\boldsymbol{Z}}_t \in \mathbb{R}_+^N$. We prove this by invoking Lemma 2.11, where we note that

$$b(\boldsymbol{z}) = -Q\mathrm{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{z} + (\theta + \mu - \lambda z^N)Q\mathbf{1}, \qquad \sigma \equiv 0.$$

In particular, we have to show that $b_i(\boldsymbol{z}) \geq 0$ for $\boldsymbol{z} \in \mathbb{R}_+^N$ with $z^i = 0$.

We start with $i = N$. Here, we have

$$b_N(\boldsymbol{z}) = -(Q\mathrm{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{z})_N + (\theta + \mu)\overline{w}.$$

Of course, $(\theta + \mu)\overline{w} \geq 0$. Moreover, due to Definition 2.1, $-(Q\mathrm{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{z})_N$ is a linear combination of z^i where all the coefficients are non-negative, with the exception of the coefficient of z^N. However, since $z^N = 0$, this implies that $b_N(\boldsymbol{z}) \geq 0$.

Next, consider $i = 1, \dots, N-1$. Then,

$$b_i(\boldsymbol{z}) = -(Q\mathrm{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{z})_i.$$

As before, $-(Q\mathrm{diag}(\boldsymbol{x})Q^{-1}\boldsymbol{z})_i$ is again a linear combination of the z^j, where all coefficients are non-negative, with the exception of the coefficient of z^i. But since $z^i = 0$, this implies that $b_i(\boldsymbol{z}) \geq 0$, proving the theorem. □

4 Solving PDEs

As an application of the domain, we want to solve PDEs. Recall that the multifactor square-root process $\boldsymbol{V}^N$ is given by

$$\mathrm{d}\boldsymbol{V}_t^N = -\mathrm{diag}(\boldsymbol{x})\left(\boldsymbol{V}_t^N - \boldsymbol{v}_0\right)\,\mathrm{d}t + \left(\theta - \lambda \boldsymbol{w}^\top \boldsymbol{V}_t^N\right)\mathbf{1}\,\mathrm{d}t + \nu\sqrt{\boldsymbol{w}^\top \boldsymbol{V}_t^N}\mathbf{1}\,\mathrm{d}W_t,$$

see (9). After a transformation of variables using $\boldsymbol{Z} \coloneqq Q\boldsymbol{V}^N$ and $\boldsymbol{z}_0 \coloneqq Q\boldsymbol{v}_0$, where Q is the matrix in Theorem 2.2, we have

$$\mathrm{d}\boldsymbol{Z}_t = -Q\mathrm{diag}(\boldsymbol{x})Q^{-1}\left(\boldsymbol{Z}_t - \boldsymbol{z}_0\right)\,\mathrm{d}t + \overline{w}\left(\theta - \lambda Z_t^{(N)}\right)\boldsymbol{e}_N\,\mathrm{d}t + \nu\overline{w}\sqrt{Z_t^{(N)}}\boldsymbol{e}_N\,\mathrm{d}W_t.$$

N	θ	λ	ν	$\mathbf{x}$	$\mathbf{w}$	$\mathbf{v}_0$
2	0.8	1.2	0.7	(0.1, 3.5)	(0.4, 1.8)	(0.2, 0.3)

Table 1: Parameters of the stochastic volatility of the lifted rough Heston model used for the numerical example.

Let $f : \mathbb{R}_+^N \to \mathbb{R}$ be a "nice" payoff function. Then, we define the value function $u : \mathbb{R}_+^N \times [0,T] \to \mathbb{R}$,

$$u(\boldsymbol{z},t) := \mathbb{E}\left[f(\boldsymbol{Z}_T)\middle|\boldsymbol{Z}_t = \boldsymbol{z}\right].$$

Then, u satisfies the PDE

$$\partial_t u - (\nabla u)^\top Q\text{diag}(\boldsymbol{x})Q^{-1}(\boldsymbol{z}-\boldsymbol{z}_0) + \overline{w}\,(\theta - \lambda z_N)\,\partial_{z_N} u + \frac{1}{2}\nu^2\overline{w}^2 z_N \partial^2_{z_N} u = 0 \tag{21}$$

with the boundary condition $u(\boldsymbol{z},T) = f(\boldsymbol{z})$, $\boldsymbol{z} \in \mathbb{R}_+^N$.

For numerical approximation, we then need to truncate the domain in space, and impose appropriate boundary conditions. For simplicity, we will instead fabricate an appropriate source term such that the PDE has an explicit, given solution, which we then also impose as Dirichlet boundary condition on the boundary of the truncated domain.

Specifically, suppose that we want the exact solution to have the form

$$u(\mathbf{z},t) = \widetilde{u}(\mathbf{z},t) := 1 + \sum_{i=1}^N \alpha_i (z^i)^2 + \beta t, \quad \mathbf{z} \in \mathbb{R}_+^N,\ t \in [0,T].$$

Plugging this formula into (21), we obtain a source term

$$\phi(\mathbf{z}) = \beta - 2\sum_{i=1}^N \alpha_i z^i \sum_{j=1}^N g_{ij}(z^j - z_0^j) + 2\alpha_N \overline{w}(\theta - \lambda z^N)z^N + \nu^2\overline{w}^2\alpha_N z^N,$$

i.e., u satisfies

$$\partial_t u - (\nabla u)^\top Q\text{diag}(\boldsymbol{x})Q^{-1}(\boldsymbol{z}-\boldsymbol{z}_0) + \overline{w}\,(\theta - \lambda z_N)\,\partial_{z_N} u + \frac{1}{2}\nu^2\overline{w}^2 z_N \partial^2_{z_N} u = \phi,$$

now with the terminal condition $u(\mathbf{z},T) = \widetilde{u}(\mathbf{z},T)$. The precise parameters chosen are summarized in Table 1, with a dimension $N = 2$ and an admissible matrix Q given by Example 2.10. We furthermore choose $\alpha = (3,\ 4)$, $\beta = 1.6$, and the terminal time $T = 2$.

After truncation of the domain, we solve the PDE by the finite element method, using the package FEniCSx, see Baratta et al. [9], and compare against the exact solution $\widetilde{u}$. In Table 2 we present the L^2-errors on the truncated domain for three choices of truncated domains, each of side-length 4:

1 $D = [0,4]^2$, corresponding to a truncation in v-space which respects the cone-shaped actual domain of the process;

2 $D = [-0.5, 3.5]^2$ corresponding to a truncation in v-space, which neither respects the cone-shaped actual domain, nor the non-negativity condition;

3 $D = [-0.5, 3.5] \times [0,4]$ corresponding to a domain truncation, which does not respect the cone-shaped actual domain in v-space, but does respect the non-negativity.

We use first order Lagrange-type finite elements, with n_t time-steps as well as mesh-size $n_x = n_t$ in each space dimension. (We refer to `https://github.com/bayerc2/domain_multifactor_volterra` for more details.)

	L^2-error over the domain D		
n	$D = [0,4]^2$	$D = [-0.5, 3.5]^2$	$D = [-0.5, 3.5] \times [0, 4]$
4	7.3×10^1	3.0×10^3	5.5×10^1
8	1.4×10^1	7.7×10^1	1.5×10^1
16	3.2×10^0	2.2×10^{10}	3.3×10^0
32	7.5×10^{-1}	1.5×10^{80}	8.0×10^{-1}
64	1.8×10^{-1}	1.4×10^{50}	2.0×10^{-1}
128	4.6×10^{-2}	inf	4.9×10^{-2}
256	1.1×10^{-2}	inf	1.2×10^{-2}
512	2.9×10^{-3}	inf	3.0×10^{-3}
1024	7.2×10^{-4}	inf	7.6×10^{-4}

Table 2: L^2 errors over the truncated domain for the approximate FEM solution to the PDE for $n_t = n_x = n$.

We would like to emphasize that, while it might seem trivial to choose $[0,4]^2$ as the domain instead of, say, $[-0.5, 3.5] \times [0, 4]$, this choice is based on correctly identifying the matrix Q and the domain $\mathcal{D} = Q^{-1}\mathbb{R}^2_+$, which is appropriately truncated here to $Q^{-1}[0,4]^2$. Without knowledge of Q, one would need to guess $\mathcal{D}$ to truncate the domain, a task that becomes increasingly nontrivial in higher dimensions.

When non-negativity of the variance process is preserved (cases 1 and 3), the numerical method empirically exhibits second order convergence, with slightly smaller error when the computational domain is a subset of the support of the process (case 1). On the other hand, when non-negativity of the variance process is not preserved on the computational domain (case 2) the error explodes due to the instability of the heat equation backward in time.

A Invariant domains in dimension $N = 3$

We extend the calculations presented for $N = 2$ in Example 2.10 to the three-dimensional case. We are looking for a matrix of the form

$$Q = \begin{pmatrix} a_1 & a_2 & -a_1 - a_2 \\ b_1 & b_2 & -b_1 - b_2 \\ w_1 & w_2 & w_3 \end{pmatrix}. \tag{22}$$

Note that there are some scaling invariances in the equation $Q\boldsymbol{x} \in \mathbb{R}^N_+$. We may multiply rows of Q with positive (!) constants without changing this condition. Hence, we restrict ourselves to

$$Q = \begin{pmatrix} 1 & -a & -1 + a \\ 1 & b & -1 - b \\ w_1 & w_2 & w_3 \end{pmatrix}.$$

Note that this corresponds to the assumption that a_1 and b_1 in (22) are both positive. Indeed, if we chose one of these entries to be 0 or -1, we would fail to find an appropriate matrix Q.

Define the matrix $R := -\overline{w}Q\text{diag}(\boldsymbol{x})Q^{-1}$, where we denote $R = (r_{i,j})_{i,j=1}^N$, and set $y_1 := x_2 - x_1$ and

$y_2 \coloneqq x_3 - x_2$. Then,

$$\begin{aligned}
r_{1,3} &= y_1 + y_2 - a y_2,\\
r_{2,3} &= y_1 + y_2 + b y_2,\\
r_{3,2} &= \frac{(w_1 w_2 y_1 + w_1 w_3 (y_1+y_2))a - w_1 w_2 y_1 + w_2 w_3 y_2}{a+b},\\
r_{3,1} &= \frac{(w_1 w_2 y_1 + w_1 w_3 (y_1+y_2))b + w_1 w_2 y_1 - w_2 w_3 y_2}{a+b},\\
r_{1,2} &= \frac{w_1 y_2 a^2 + (w_3 y_1 + w_2(y_1+y_2) - w_1 y_2)a - w_2(y_1+y_2)}{a+b},\\
r_{2,1} &= \frac{-w_1 y_2 b^2 + (w_3 y_1 + w_2(y_1+y_2) - w_1 y_2)b + w_2(y_1+y_2)}{a+b},
\end{aligned}$$

and all these quantities have to be non-negative. Assume now further that $a, b \geq 0$. Then, $r_{2,3} \geq 0$ is trivially satisfied, and $r_{1,3}, r_{3,2}, r_{3,1}, r_{1,2}, r_{2,1} \geq 0$ simplify to

$$\begin{aligned}
a &\leq \frac{y_1+y_2}{y_2},\\
a &\geq \frac{w_2}{w_1}\frac{w_1 y_1 - w_3 y_2}{w_2 y_1 + w_3(y_1+y_2)},\\
b &\geq \frac{w_2}{w_1}\frac{-w_1 y_1 + w_3 y_2}{w_2 y_1 + w_3(y_1+y_2)},\\
0 &\leq w_1 y_2 a^2 + ca - w_2(y_1+y_2),\\
0 &\geq w_1 y_2 b^2 - cb - w_2(y_1+y_2),
\end{aligned}$$

where $c \coloneqq w_3 y_1 + w_2(y_1+y_2) - w_1 y_2$.

This further implies for a that

$$\frac{-c + \sqrt{c^2 + 4 w_1 w_2 y_2 (y_1+y_2)}}{2 w_1 y_2} \vee \frac{w_2}{w_1}\frac{w_1 y_1 - w_3 y_2}{w_2 y_1 + w_3(y_1+y_2)} \leq a \leq \frac{y_1+y_2}{y_2}.$$

One can verify that the lower bound is always smaller than the upper bound, proving that such an a exists. However, it is slightly simpler and perhaps more illustrative to prove that $a = 1$ satisfies these inequalities. For the upper bound, this is trivial. For the lower bound, note that

$$\frac{w_2}{w_1}\frac{w_1 y_1 - w_3 y_2}{w_2 y_1 + w_3(y_1+y_2)} \leq \frac{w_2}{w_1}\frac{w_1 y_1}{w_2 y_1} = 1,$$

and

$$\begin{aligned}
&\frac{-c + \sqrt{c^2 + 4 w_1 w_2 y_2 (y_1+y_2)}}{2 w_1 y_2} \leq 1\\
\iff& \sqrt{c^2 + 4 w_1 w_2 y_2 (y_1+y_2)} \leq 2 w_1 y_2 + c\\
\iff& c^2 + 4 w_1 w_2 y_2 (y_1+y_2) \leq c^2 + 4 w_1^2 y_2^2 + 4 w_1 y_2 c\\
\iff& w_2(y_1+y_2) \leq w_1 y_2 + c.
\end{aligned}$$

which follows immediately from the definition of c.

Next, for b we get the conditions

$$0 \vee \frac{w_2}{w_1}\frac{-w_1 y_1 + w_3 y_2}{w_2 y_1 + w_3(y_1+y_2)} \leq b \leq \frac{c + \sqrt{c^2 + 4 w_1 w_2 y_2 (y_1+y_2)}}{2 w_1 y_2}.$$

This time, we verify that $b = \frac{w_2}{w_1}$ is admissible. For the lower bound, this is clear. For the upper bound, note that

$$\begin{aligned} \frac{w_2}{w_1} &\leq \frac{c + \sqrt{c^2 + 4w_1w_2y_2(y_1+y_2)}}{2w_1y_2} \\ \iff 2w_2y_2 - c &\leq \sqrt{c^2 + 4w_1w_2y_2(y_1+y_2)} \\ \Longleftarrow c^2 + 4w_2^2y_2^2 - 4w_2y_2c &\leq c^2 + 4w_1w_2y_2(y_1+y_2) \\ \iff w_2y_2 - c &\leq w_1(y_1+y_2). \end{aligned}$$

This again follows from the definition of c.

Hence, we have shown that we can choose $a = 1$ and $b = \frac{w_2}{w_1}$, yielding

$$Q = \begin{pmatrix} 1 & -1 & 0 \\ 1 & \frac{w_2}{w_1} & -1 - \frac{w_2}{w_1} \\ w_1 & w_2 & w_3 \end{pmatrix}.$$

Note that due to scaling invariance, the matrix

$$Q = \begin{pmatrix} w_1 & -w_1 & 0 \\ w_1 & w_2 & -w_1 - w_2 \\ w_1 & w_2 & w_3 \end{pmatrix}$$

would be equivalent.

Consider now the specific example $\boldsymbol{x} := (1, 5, 25)$ and $\boldsymbol{w} := (1, 2, 3)$. Then, we get the conditions

$$\begin{aligned} 0.84 \approx \frac{-5 + \sqrt{85}}{5} \leq a \leq \frac{6}{5} = 1.2, \\ 1.4 = \frac{7}{5} \leq b \leq \frac{5 + \sqrt{85}}{5} \approx 2.84. \end{aligned}$$

Comparing to the previous discussion, we see that indeed, $a = 1$ and $b = \frac{w_2}{w_1} = 2$ are admissible.

The corresponding plots for the multifactor square-root process are shown in Figure 3. We give projections to two-dimensional planes, as this makes it easier to visually verify that the samples lie in $\mathbb{R}_+^3$. Furthermore, we give three different choices of (a, b), the first two being admissible, and the third not. Indeed, we see for the first two choices that the samples lie in $\mathbb{R}_+^3$, while this is not the case for the third choice.

B Remark on the link between the sets $\mathcal{E}$ and $\mathcal{G}$

In this section, we argue that the two abstract 'invariance' sets that appeared in the literature in Abi Jaber and El Euch [2], Cuchiero and Teichmann [16] are equal. This part is valid for more general locally square-integrable kernels K beyond the weighted sum of exponential case.

We introduce the following notations. For suitable functions f, g and measure L we denote their convolution by $*$:

$$(f * g)(t) = \int_0^t f(t-s)g(s)\mathrm{d}s = \int_0^t f(s)g(t-s)\mathrm{d}s, \quad (f * L)(t) := \int_0^t f(t-s)L(\mathrm{d}s).$$

The shift operator Δ_h with $h \geq 0$, maps any function f on $\mathbb{R}_+$ to the function $\Delta_h f$ given by

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

If the function f on $\mathbb{R}_+$ is right-continuous and of locally bounded variation, the measure induced by its distributional derivative is denoted $\mathrm{d}f$, so that $f(t) = f(0) + \int_{[0,t]} \mathrm{d}f(s)$ for all $t \geq 0$. By convention, df does not

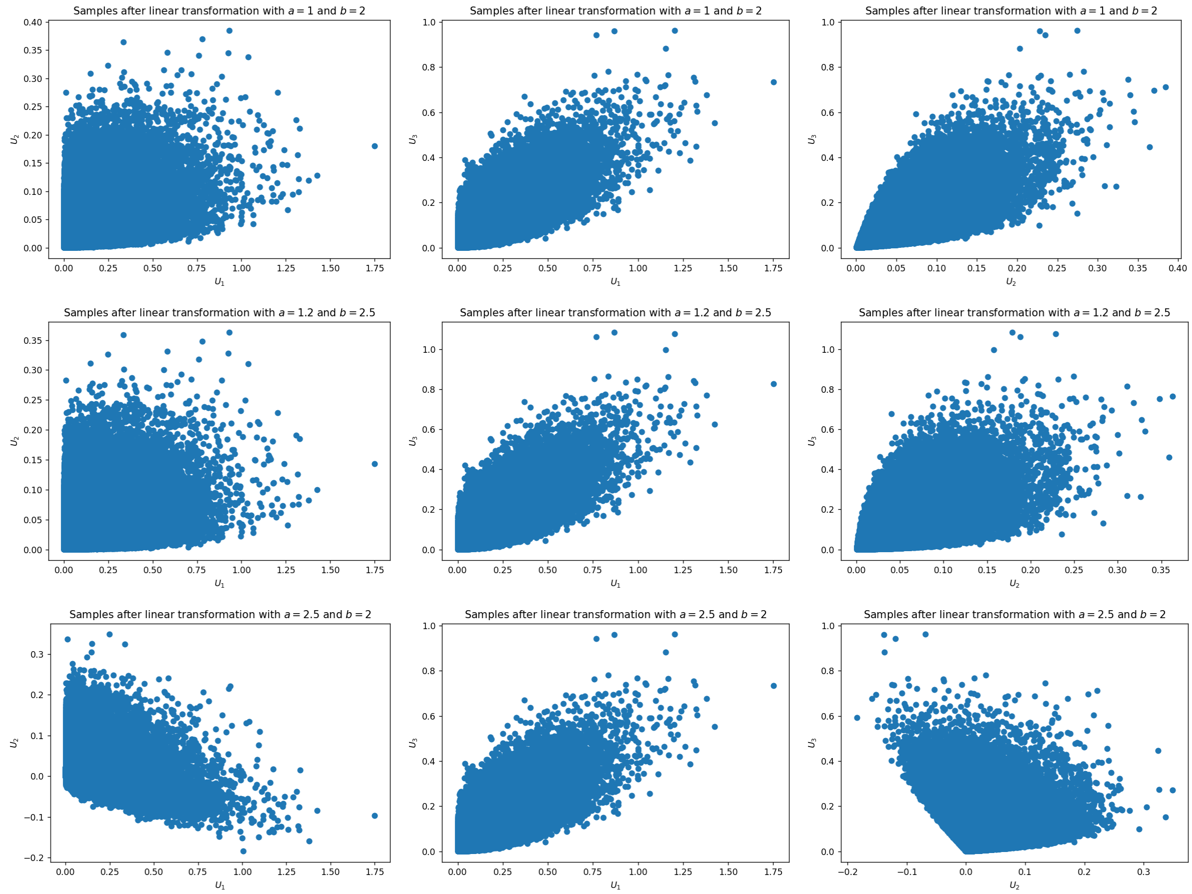

Figure 3: Samples of projections of $\boldsymbol{U}$ using 10^3 sample paths on a time grid with $M = 10^5$ time steps. The parameters used are $\boldsymbol{x} = (1, 5, 25), \boldsymbol{w} = (1, 2, 3), \lambda = 0.3, \nu = 0.3, V_0 = 0.02, \theta = 0.02, T = 100$, and $\boldsymbol{v}_0$ is chosen to be proportional to $\boldsymbol{x}^{-1}$.

charge $\{0\}$.

Two sets appeared so far in the literature to characterize the non-negativity of solutions to stochastic Volterra equations:

1 Set of Cuchiero and Teichmann [16, Equation (4.7), Definition 4.12 and Theorem 4.17(i)]:

$$\mathcal{E} = \bigcap_{w>0} \mathcal{E}^\eta \text{ with } \mathcal{E}^\eta := \{g_0 : [0,T] \to \mathbb{R} \text{ such that } g_0 - R^\eta * g_0 \geq 0\}.$$

Here $R^\eta(t)$ is the resolvent of the second kind of the kernel (ηK) defined by

$$R^\eta = \eta K - \eta K * R^\eta = \eta K - R^\eta * \eta K.$$

2 Set of Abi Jaber and El Euch [2, Equations (2.4)-(2.5) and Theorem 2.1]:

$$\mathcal{G} = \big\{g_0 : [0,T] \to \mathbb{R} \text{ such that } \Delta_h g_0 - (\Delta_h K * L)(0) g_0 - d(\Delta_h K * L) * g_0 \geq 0 \text{ and } g_0(0) \geq 0.\big\}$$

where $L(\mathrm{d}t)$ is the resolvent of the first kind of the kernel[1]

$$K * L = 1 = L * K.$$

[1]Under some suitable assumptions on the kernel, see [2, Assumption (H1)], one can show that K admits a resolvent of the first kind such that $\Delta_h K * L$ is right-continuous and of locally bounded variation, see [2, Remark B.3], thus the associated measure $d(\Delta_h K * L)$ that appears in the set $\mathcal{G}$ is well defined.

One can argue that the two sets are equal:

$$\mathcal{E} = \mathcal{G}$$

since both conditions that appear in the set are necessary and sufficient conditions for the non-negativity of the linear Volterra equation

$$f^\eta = g_0 - \eta K * f^\eta, \tag{23}$$

for $w > 0$. Indeed, on the one hand the solution of (23) can be expressed in terms of the resolvent of the second kind in the form

$$f^\eta = g_0 - R^\eta * g_0,$$

which is exactly the form that appears in $\mathcal{E}$. On the other hand, by relying on the properties of the resolvent of the first kind, see for instance Abi Jaber and El Euch [2, the proof of Theorem A.2], one can write that

$$\begin{aligned} f^\eta(t+h) = {} & \Delta_h g_0(t) - (\Delta_h K * L)(0) g_0(t) - (d(\Delta_h K * L) * g_0)(t) \\ & + (\Delta_h K * L)(0) f^\eta(t) + (d(\Delta_h K * L) * f^\eta)(t) \\ & - w \int_t^{t+h} K(t-s) f^\eta(s) \mathrm{d}s. \end{aligned}$$

We note that the first line is exactly the condition that appears in the set $\mathcal{G}$. Using the two above expressions one can show that the non-negativity of the linear Volterra equation (23), for any $w > 0$, is equivalent to the condition that appear in $\mathcal{E}$ as well as the one that appears in $\mathcal{G}$, which shows that the two sets are equal.

In principle, to establish a link with our cone $\mathcal{D}$, one should restrict to kernels that are weighted sums of exponentials of the form

$$K(t) = \sum_{i=1}^N w_i e^{-x_i t},$$

and input curves of the form

$$g_0(t) = \sum_{i=1}^n w_i e^{-x_i t} Y_0^i.$$

Then, the resolvents of the second kind and first kind for such kernels must be computed and plugged into the conditions defining the sets $\mathcal{E}$ and $\mathcal{G}$. Even in dimension $N = 2$, this leads to highly cumbersome and non-trivial computations, and it is not clear how to explicitly determine a suitable domain, as for instance our cone $\mathcal{D}$, for Y_0^i from $\mathcal{E}$ and $\mathcal{G}$. This makes the approach in the current paper particularly crucial.

C Some proofs

C.1 Proof of Theorem 2.2

In preparation for the proof, we introduce the matrix $R = (r_{i,j})_{i,j=1,\dots,N}$ defined by

$$\begin{aligned} r_{i,N} &= \frac{1}{w}, && i = 1, \dots, N, \\ r_{i,j} &= \frac{w_{j+1}}{\sum_{\ell=1}^j w_\ell \sum_{\ell=1}^{j+1} w_\ell}, && i \le j < N, \\ r_{i+1,i} &= \frac{-1}{\sum_{\ell=1}^{i+1} w_\ell}, && i = 1, \dots, N-1, \\ r_{i,j} &= 0, && i \ge j+2, \end{aligned} \tag{24}$$

that will turn out to be the inverse of Q. For example, for $N=4$ we have

$$R=\begin{pmatrix}\frac{w_2}{w_1(w_1+w_2)} & \frac{w_3}{(w_1+w_2)(w_1+w_2+w_3)} & \frac{w_4}{(w_1+w_2+w_3)(w_1+w_2+w_3+w_4)} & \frac{1}{w_1+w_2+w_3+w_4}\\ \frac{-1}{w_1+w_2} & \frac{w_3}{(w_1+w_2)(w_1+w_2+w_3)} & \frac{w_4}{(w_1+w_2+w_3)(w_1+w_2+w_3+w_4)} & \frac{1}{w_1+w_2+w_3+w_4}\\ 0 & \frac{-1}{w_1+w_2+w_3} & \frac{w_4}{(w_1+w_2+w_3)(w_1+w_2+w_3+w_4)} & \frac{1}{w_1+w_2+w_3+w_4}\\ 0 & 0 & \frac{-1}{w_1+w_2+w_3+w_4} & \frac{1}{w_1+w_2+w_3+w_4}\end{pmatrix}$$

Proof of Theorem 2.2. Conditions 2 and 3 of Definition 2.1 are readily satisfied by construction. To argue Condition 1, we will prove that R given in (24) is actually the inverse of Q, i.e. $QR=\mathrm{Id}$. Indeed, consider first the diagonal elements. Here, we have

$$(QR)_{NN}=\sum_{k=1}^{N} q_{Nk}r_{kN}=\sum_{k=1}^{N} w_k\frac{1}{w}=1,$$

$$(QR)_{ii}=\sum_{k=1}^{N} q_{ik}r_{ki}=\sum_{k=1}^{i} w_k\frac{w_{i+1}}{\sum_{\ell=1}^{i} w_\ell\sum_{\ell=1}^{i+1} w_\ell}+\left(-\sum_{\ell=1}^{i} w_\ell\right)\frac{-1}{\sum_{\ell=1}^{i+1} w_\ell}=1,$$

for $i=1,\dots,N-1$. Next, consider off-diagonal elements. We have

$$(QR)_{Nj}=\sum_{k=1}^{N} q_{Nk}r_{kj}=\sum_{k=1}^{j} w_k\frac{w_{j+1}}{\sum_{\ell=1}^{j} w_\ell\sum_{\ell=1}^{j+1} w_\ell}+w_{j+1}\frac{-1}{\sum_{\ell=1}^{j+1} w_\ell}=0,$$

for $j\le N-1$,

$$(QR)_{iN}=\sum_{k=1}^{N} q_{ik}r_{kN}=\sum_{k=1}^{i} w_k\frac{1}{w}+\left(-\sum_{\ell=1}^{i} w_\ell\right)\frac{1}{w}=0,$$

for $i\le N-1$,

$$(QR)_{ij}=\sum_{k=1}^{N} q_{ik}r_{kj}=\sum_{k=1}^{i} w_k\frac{w_{j+1}}{\sum_{\ell=1}^{j} w_\ell\sum_{\ell=1}^{j+1} w_\ell}+\left(-\sum_{\ell=1}^{i} w_\ell\right)\frac{w_{j+1}}{\sum_{\ell=1}^{j} w_\ell\sum_{\ell=1}^{j+1} w_\ell}=0,$$

for $i<j\le N-1$, and

$$(QR)_{ij}=\sum_{k=1}^{N} q_{ik}r_{kj}=\sum_{k=1}^{j} w_k\frac{w_{j+1}}{\sum_{\ell=1}^{j} w_\ell\sum_{\ell=1}^{j+1} w_\ell}+w_{j+1}\frac{-1}{\sum_{\ell=1}^{j+1} w_\ell}=0,$$

for $j<i\le N-1$. In particular, this proves that $R=Q^{-1}$.

Finally, we verify Condition 4 of Definition 2.1 by direct computations. First, it is easily verified that we have $Q\mathrm{diag}(\boldsymbol{x})=S:=(s_{i,j})_{i,j=1,\dots,N}$ with

$$s_{ij}=q_{ij}x_j=w_jx_j,\qquad j\le i,\qquad s_{i,i+1}=-x_{i+1}\sum_{\ell=1}^{i} w_\ell,\qquad i=1,\dots,N-1.$$

Now, let us compute $Q\mathrm{diag}(\boldsymbol{x})Q^{-1}=T:=(t_{i,j})_{i,j=1,\dots,N}$.

We have

$$t_{i,N}=\sum_{k=1}^{N} s_{i,k}r_{k,N}=\frac{1}{w}\left(\sum_{k=1}^{i} w_kx_k-x_{i+1}\sum_{\ell=1}^{i} w_\ell\right)\le 0$$

for $i=1,\dots,N-1$, since the x_i are ordered increasingly. Similarly,

$$t_{N,j}=\sum_{k=1}^{N} s_{N,k}r_{k,j}=\sum_{k=1}^{j} x_kw_k\frac{w_{j+1}}{\sum_{\ell=1}^{j} w_\ell\sum_{\ell=1}^{j+1} w_\ell}+x_{j+1}w_{j+1}\frac{-1}{\sum_{\ell=1}^{j+1} w_\ell}\le 0$$

for $j = 1, \dots, N-1$. Next,

$$t_{i,j} = \sum_{k=1}^{N} s_{i,k} r_{k,j} = \left(\sum_{k=1}^{i} w_k x_k - x_{i+1} \sum_{\ell=1}^{i} w_\ell \right) \frac{w_{j+1}}{\sum_{\ell=1}^{j} w_\ell \sum_{\ell=1}^{j+1} w_\ell} \le 0$$

for $i < j \le N-1$. Finally,

$$t_{i,j} = \sum_{k=1}^{N} s_{i,k} r_{k,j} = \sum_{k=1}^{j} w_k x_k \frac{w_{j+1}}{\sum_{\ell=1}^{j} w_\ell \sum_{\ell=1}^{j+1} w_\ell} + w_{j+1} x_{j+1} \frac{-1}{\sum_{\ell=1}^{j+1} w_\ell} \le 0$$

for $j < i \le N-1$. This verifies Condition 4 of Definition 2.1 and proves the theorem. □

C.2 The mean-reversion level is in the domain

Proof of Lemma 2.6. We prove the equivalent statement $Q\boldsymbol{y}_0 \in \mathbb{R}_+^N$. First, note that

$$Q\boldsymbol{y}_0 = \mu Q \text{diag}(\boldsymbol{x})^{-1} \mathbf{1} = \mu Q \text{diag}(\boldsymbol{x})^{-1} Q^{-1} Q \mathbf{1} = \mu \overline{w} Q \text{diag}(\boldsymbol{x})^{-1} Q^{-1} \boldsymbol{e}_N.$$

Recall that in linear algebra, an M-matrix is a square matrix with non-positive off-diagonal entries and with eigenvalues whose real parts are non-negative. Clearly, the matrix $Q\text{diag}(\boldsymbol{x})Q^{-1}$ is an M-matrix: The non-positivity of the off-diagonal entries holds by assumption, and its eigenvalues are (real and) positive, since it is just the matrix $\text{diag}(\boldsymbol{x})$ written in a different basis. Now it is well-known that the inverse of an M-matrix has non-negative entries (in fact, this property characterizes M-matrices). Therefore,

$$\left(Q\text{diag}(\boldsymbol{x})Q^{-1}\right)^{-1} = Q\text{diag}(\boldsymbol{x})^{-1}Q^{-1}$$

has only non-negative entries. In particular,

$$Q\boldsymbol{y}_0 = \mu \overline{w} Q \text{diag}(\boldsymbol{x})^{-1} Q^{-1} \boldsymbol{e}_N \in \mathbb{R}_+^N,$$

proving the lemma. □

References

[1] Eduardo Abi Jaber. Lifting the Heston model. *Quantitative finance*, 19(12):1995–2013, 2019.

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

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

[4] Eduardo Abi Jaber, Bruno Bouchard, and Camille Illand. Stochastic invariance of closed sets with non-Lipschitz coefficients. *Stochastic Processes and their Applications*, 129(5):1726–1748, 2019.

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

[6] Eduardo Abi Jaber, Enzo Miller, and Huyên Pham. Linear-quadratic control for a class of stochastic Volterra equations: solvability and approximation. *The Annals of Applied Probability*, 31(5):2244–2274, 2021.

[7] Aurélien Alfonsi and Ahmed Kebaier. Approximation of stochastic Volterra equations with kernels of completely monotone type. *arXiv preprint arXiv:2102.13505*, 2021.

[8] Andrew D Baczewski and Stephen D Bond. Numerical integration of the extended variable generalized Langevin equation with a positive Prony representable memory kernel. *The Journal of chemical physics*, 139(4), 2013.

[9] Igor A. Baratta, Joseph P. Dean, Jørgen S. Dokken, Michal Habera, Jack S. Hale, Chris N. Richardson, Marie E. Rognes, Matthew W. Scroggs, Nathan Sime, and Garth N. Wells. DOLFINx: the next generation FEniCS problem solving environment. preprint, 2023.

[10] Christian Bayer and Simon Breneis. Weak Markovian approximations of rough Heston. *arXiv preprint arXiv:2309.07023*, 2023.

[11] Christian Bayer and Simon Breneis. Markovian approximations of stochastic Volterra equations with the fractional kernel. *Quantitative Finance*, 23(1):53–70, 2023.

[12] Christian Bayer and Simon Breneis. Efficient option pricing in the rough Heston model using weak simulation schemes. *Quantitative Finance*, 24(9):1247–1261, 2024.

[13] Thierry Bochud and Damien Challet. Optimal approximations of power laws with exponentials: application to volatility models with long memory. *Quantitative Finance*, 7(6):585–589, 2007.

[14] Philippe Carmona and Laure Coutin. Fractional Brownian motion and the Markov property. *Electronic Communications in Probability [electronic only]*, 3:95–107, 1998.

[15] Etienne Chevalier, Sergio Pulido, and Elizabeth Zúñiga. American options in the Volterra Heston model. *SIAM Journal on Financial Mathematics*, 13(2):426–458, 2022.

[16] Christa Cuchiero and Josef Teichmann. Generalized Feller processes and Markovian lifts of stochastic Volterra processes: the affine case. *Journal of evolution equations*, 20(4):1301–1348, 2020.

[17] Giuseppe Da Prato and Hélene Frankowska. Invariance of stochastic control systems with deterministic arguments. *Journal of differential equations*, 200(1):18–52, 2004.

[18] Giuseppe Da Prato and Hélene Frankowska. Stochastic viability of convex sets. *Journal of mathematical analysis and applications*, 333(1):151–163, 2007.

[19] Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38, 2019.

[20] Philipp Harms. Strong convergence rates for Markovian representations of fractional Brownian motion, 2019.

[21] Gytenis Lileika and Vigirdas Mackevičius. Second-order weak approximations of CKLS and CEV processes by discrete random variables. *Mathematics*, 9(12):1337, 2021.

[22] Antonis Papapantoleon and Jasper Rou. A time-stepping deep gradient flow method for option pricing in (rough) diffusion models. *arXiv preprint arXiv:2403.00746*, 2024.