Polynomial Volterra processes

Eduardo Abi Jaber[1], Christa Cuchiero[2], Luca Pelizzari[3], Sergio Pulido[4], Sara Svaluto-Ferro[5]

submitted: March 26, 2024

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

[2] Department of Statistics and Operations Research
    Vienna University
    Kolingasse 14-16
    1090 Wien
    Austria
    E-Mail: christa.cuchiero@univie.ac.at

[3] Weierstrass Institute
    Mohrenstr. 39
    10117 Berlin
    Germany
    E-Mail: luca.pelizzari@wias-berlin.de

[4] École Nationale Supérieure d'Informatique
    pour l'Industrie et l'Entreprise (ENSIIE)
    1 square de la Résistance
    91000 Évry-Courcouronnes
    France
    E-Mail: sergio.pulidonino@ensiie.fr

[5] Department of Economics
    University of Verona
    Via Cantarane 24
    37129 Verona
    Italy
    E-Mail: sara.svalutoferro@univr.it

No. 3098

Berlin 2024

Polynomial Volterra processes

Eduardo Abi Jaber, Christa Cuchiero, Luca Pelizzari, Sergio Pulido, Sara Svaluto-Ferro

Abstract

We study the class of continuous polynomial Volterra processes, which we define as solutions to stochastic Volterra equations driven by a continuous semimartingale with affine drift and quadratic diffusion matrix in the state of the Volterra pro- cess. To demonstrate the versatility of possible state spaces within our framework, we construct polynomial Volterra processes on the unit ball. This construction is based on a stochastic invariance principle for stochastic Volterra equations with possibly singular kernels. Similarly to classical polynomial processes, polynomial Volterra processes allow for tractable expressions of the moments in terms of the unique solution to a system of deterministic integral equations, which reduce to a system of ODEs in the classical case. By applying this observation to the moments of the finite-dimensional distributions we derive a uniqueness result for polynomial Volterra processes. Moreover, we prove that the moments are polynomials with respect to the initial condition, another crucial property shared by classical polyno- mial processes. The corresponding coefficients can be interpreted as a deterministic dual process and solve integral equations dual to those verified by the moments themselves. Additionally, we obtain a representation of the moments in terms of a pure jump process with killing, which corresponds to another non-deterministic dual process.

1 Introduction

Polynomial processes in finite dimensions, introduced in [28] and [40], constitute a class of time-homogeneous Itô-semimartingales which are inherently tractable: conditional moments can be expressed through a *deterministic dual process* which is the solution of a linear ODE. This is the so-called *moment formula*. They form a rich class that includes Wright-Fisher diffusions ([48]) from population genetics, Wishart correlation matrices ([9]), and affine processes ([35]), just to name a few. Notably, polynomial diffusions offer greater flexibility than affine diffusions, accommodating more general semialgebraic state spaces, including in particular bounded state spaces; see [40] and [49] for a systematic analysis. The computational advantages due to the moment formula in the polynomial setting (see, e.g., [8] and [41]) have led to a wide range of applications, in particular in population genetics and mathematical finance. Indeed, in population genetics dual processes associated to moments and their interpretation in view of coalescent theory play an important role: the Wright-Fisher diffusion with seed-bank component (see, e.g., [23] and the references therein) is for instance an important example of a recently investigated two-dimensional polynomial process in this field. In mathematical finance, polynomial pro- cesses comprise a plethora of highly popular models, ranging from the famous Black Scholes model over certain jump-diffusions to Jacobi-type processes, which have been used for stochastic volatility models, life insurance liability modeling, variance swaps, and stochastic portfolio theory (see, e.g., [8, 22, 26, 39]).

All these models share a finite dimensional Markov property which sometimes may not be adequate, for in- stance, for modeling volatility where path-dependence is crucial (see, e.g., [25] and [45]). This has motivated the emergence of numerous models in the literature based on stochastic Volterra equations, where the speci- fication of the kernel offers greater flexibility to align with market data ([5, 7, 34, 42, 44, 45, 52]). In particular, the singular fractional kernel is important in view of rough volatility models [42]. To obtain models such as the *rough Heston model* [36], the important class of affine processes has been extended to the Volterra framework. In particular, existence and uniqueness of solutions to the associated equations, invariance over certain do- mains, and formulas for the Fourier-Laplace transform have been established within the affine paradigm, see e.g., [6, 32, 24]. Note that Volterra-type processes are not only used in the realm of volatility modeling but also to model phenomena exhibiting short and long range dependence and self-similarity. For instance, they have

been applied in web-traffic [51, 57] and energy markets (see, e.g. [11, 12]). In the latter, so-called Brownian semistationary processes, introduced in [16], as well as volatility modulated Volterra processes, first considered in [15], play an important role. All these processes can be embedded into the large class of ambit processes, pioneered in [14] to model turbulence and tumour growth. We also refer to the monograph [13] for a far reaching analysis of these processes.

In analogy to continuous *affine* Volterra processes as studied in [6], we shall define continuous polynomial Volterra processes as solutions X of stochastic Volterra equations driven by a continuous semimartingale Z depending on X in a way that resembles the structure appearing in the classical framework. More precisely, Z has an affine drift and a quadratic diffusion matrix in X. To derive moment formulas in this setting, we draw inspiration from recent works on infinite-dimensional polynomial processes. While there is already a vast literature on finite dimensional polynomial processes, a systematic analysis of the infinite dimensional case was only recently provided in [29, 30, 27, 18]. The articles [29, 27] treat probability and non-negative measure-valued processes, which include the famous Fleming-Viot and the Dawson-Watanabe superprocess (see [37] for an introduction to superprocesses and [50] for measure-valued branching Markov processes). To accommodate these and also function space valued processes, a common unifying framework that establishes in particular the moment formula in a generic infinite dimensional setting has been built in [30]. Related concepts have also been in considered [18] and in [17].

Our work is the first systematic study extending the theory of polynomial processes to the Volterra setting and contributes to the existing literature on Volterra processes in multiple ways. In what follows, we describe the organization of the paper and our main contributions.

In Section 2, we set the stage by defining continuous polynomial Volterra processes and by recalling moment estimates and existence of solutions to the associated stochastic Volterra equations from [6]. To illustrate the versatility of possible state spaces, even in the Volterra case, we construct in Theorem 2.5 the first non-trivial example of a polynomial Volterra process with possibly singular kernels that remains confined to the unit ball. Notably, when restricted to the one dimensional case, our results provide the construction of Jacobi Volterra processes, see Corollary 2.6. This construction relies on a more general result in Theorem 5.1 showing existence of general Volterra processes confined to the unit ball under structural assumption on the drift and diffusion coefficient of the driving semimartingale Z. Note that solutions to Volterra equations that remain within a convex set have so far been constructed and studied for domains with no curvature, such as the non-negative orthant $\mathbb{R}_+^d$, see [6, Theorem 3.6] and the extensions in [2, 10]. The primary challenge in constructing a solution X that remains within a given convex set with curvature arises from the potential singularity of the kernel. This singularity could push the process X outside the domain, if the volatility does not vanish on the boundary. This difficulty has already been observed in the construction of Volterra Wishart processes [1, 32].

Section 3 is dedicated to an analysis of the moments of polynomial Volterra processes. In particular, in Section 3.1 we establish an extension of the moment formula for polynomial Volterra processes. This is the main moment formula in our work. It shows that the moments are the unique solutions to a system of deterministic integral equations, which reduce to a system of ODEs for classical polynomial diffusions; see Theorem 3.3. In the terminology of [30] this formula corresponds to the bidual moment formula. In contrast to the generic infinite dimensional framework of [30] we can here actually prove existence and uniqueness of the system of deterministic integral equations. Using a variation of constants technique, our arguments can be applied to deduce more explicit expressions for the first and second-order moments and for all moments in the affine case, see Section 3.2. Moreover, we elucidate in Section 3.3 a crucial structural property, shared also by classical polynomial diffusions, namely that the moments are polynomials with respect to the initial condition. The corresponding coefficients can be interpreted as deterministic dual process and solve integral equations dual to those verified by the moments themselves; see Theorem 3.8. In the terminology of [30] this corresponds to the dual moment formula, where we can again prove existence and uniqueness of the corresponding equations (which had to be assumed in the general framework of [30]). Additionally, in Section 3.4, we illustrate how our results and arguments can be applied to the moments of the finite-dimensional distributions. Our considerations then also lead to a novel result regarding the uniqueness in law for solutions to stochastic Volterra equations in the polynomial

framework, as proved in Theorem 3.12.

In Section 4, we get inspiration from the work on classical Flemming Viot processes and general infinite dimensional polynomial processes as considered in [30] to show that the moments of a polynomial Volterra process can be expressed in terms of expectations of a functional of a (finite dimensional) pure jump process with killing. Indeed, we consider a function valued lift to the so-called Filipović space [38], denoted by B, and then apply the 'dual process' approach. Denote the infinitesimal generator of the function valued lift by $\mathcal{A}$ and consider polynomials $f : B \times \mathbb{R}^k \to \mathbb{R}$, $(\lambda, x) \mapsto \lambda(x_1) \cdots \lambda(x_k)$. Then we show that in the case of bounded kernels there exists a k-dimensional pure jump process with killing, denoted by U, with generator $\mathcal{L}$ such that

$$\mathcal{A}f(\cdot, x)(\lambda) = \mathcal{L}f(\lambda, \cdot)(x). \tag{1}$$

Modulo several technical conditions, e.g., stated in [31, Lemma A.1], it then holds that

$$\mathbb{E}_{\lambda_0}[f(\lambda_t, x)] = \mathbb{E}_{Y_0=x}[f(\lambda_0, U_t)]. \tag{2}$$

As the evaluation of λ at 0 corresponds to the Volterra process, we get a representation of the kth moment by setting $x = Y_0 = 0 \in \mathbb{R}^k$. For the homogeneous case with linear drift and volatility this formula is rigorously proved in Proposition 4.3 and Proposition 4.8 for the multivariate case. Since for $\lambda_0 \equiv X_0$ we have $f(\lambda_0, x) \equiv X_0^k$, we also see that the kth moment of the Volterra process is a monomial of degree k in the initial value, which is thus a special case of Theorem 3.8. For the general non-homogeneous case we retrieve also a similar formula, see Remark 4.6. In this context, let us also mention that the results of Section 3.3 can be seen from a similar duality point of view, here with $\mathcal{L}$ corresponding to the dual operator in the terminology of [30], giving rise to a system of deterministic PDEs. From a numerical perspective the jump representation can sometimes have advantages as it is easy to simulate from a pure jump process and then compute the right hand side of (2) via Monte Carlo.

Section 5 provides the proof of Theorem 5.1. Appendix A contains the main results necessary to guarantee existence and uniqueness of solutions to the equations presented in Section 3, namely the integral equations verified by the moments of polynomial Volterra processes and by the coefficients in the expression as polynomial with respect to the initial condition.

Notation: We denote polynomials on $\mathbb{R}^d$ of degree less than or equal to n by $\mathrm{Pol}_n(\mathbb{R}^d)$. $\mathbb{N}$ is the set of natural numbers and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$, with $\alpha_i \in \mathbb{N}_0$, we write $|\alpha| = \sum_{i=1}^d \alpha_i$ for the sum of its components. For a vector $x \in \mathbb{R}^m$, x_i is its i-th coordinate. For a matrix $A \in \mathbb{R}^{d \times d}$, we denote by $|A| = \sqrt{\mathrm{Tr}(A^\top A)}$ its Frobenius norm. I_d is the $d \times d$ identity matrix. For any $t \geq 0$, we use the symbol $\mathbb{E}_t$ to denote the conditional expectation given the σ-algebra $\mathcal{F}_t$. We sometimes use the convolution notation $(f * g)(t) = \int_0^t f(t-s)g(s)ds$ for functions f and g, and $(f * L)(t) = \int_0^t f(t-s)L(ds)$ for a measure L.

2 Definition and existence of polynomial Volterra processes

Fix a dimension $d \in \mathbb{N}$ and consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, where $(\mathcal{F}_t)_{t \geq 0}$ satisfies the usual conditions and $\mathcal{F}_0$ is the trivial σ-algebra on Ω. A continuous polynomial Volterra process of convolution type is a d-dimensional adapted process X with continuous trajectories solving a stochastic Volterra equation of the form

$$X_t = g_0(t) + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s, \quad t \geq 0, \tag{3}$$

where

- W is a d-dimensional Brownian motion,

- the *initial condition* $g_0 \colon \mathbb{R}_+ \to \mathbb{R}^d$ is in $C(\mathbb{R}_+, \mathbb{R}^d)$,

- the *convolution kernel* $K \colon \mathbb{R}_+ \to \mathbb{R}^{d \times d}$ is in $L^2_{\mathrm{loc}}(\mathbb{R}_+, \mathbb{R}^{d \times d})$,

- the map $b \colon \mathbb{R}^d \to \mathbb{R}^d$ has components in $\mathrm{Pol}_1(\mathbb{R}^d)$, and $\sigma \colon \mathbb{R}^d \to \mathbb{R}^{d \times d}$ is a continuous map such that $a(x) = \sigma(x)\sigma(x)^\top$ has entries in $\mathrm{Pol}_2(\mathbb{R}^d)$. More precisely,

$$b(x) = b_0 + \sum_{i=1}^{d} b_i x_i, \quad a(x) = A_0 + \sum_{i=1}^{d} A_i x_i + \sum_{i,j=1}^{d} A_{ij} x_i x_j \tag{4}$$

for some $b_i \in \mathbb{R}^d$ and $A_i, A_{ij} \in \mathbb{R}^{d \times d}$.

Observe that if all $A_{ij} = 0$, X is an affine Volterra process as in [6]. As for stochastic differential equations, we speak of weak solutions to (3) whenever the filtered probability space and the underlying Brownian motion are not fixed a priori and they are part of the solution. In this case, with a slight abuse of terminology, we say that X is a weak solution to (3).

If we define the d-dimensional semimartingale Z as

$$Z_t = \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s, \tag{5}$$

then the stochastic Volterra equation (3) can be recast as

$$X_t = g_0(t) + \int_0^t K(t-s)dZ_s.$$

The following proposition provides a priori estimates on the moments of any solution to (3).

Proposition 2.1 *Let X be a continuous solution to* (3)*. Then, for any $p \in \mathbb{N}$ and $T \geq 0$,*

$$\sup_{0 \leq t \leq T} \mathbb{E}[|X_t|^p] \leq c \tag{6}$$

for some constant c which depends only on $\sup_{0 \leq t \leq T} |g_0(t)|, p, K|_{[0,T]}, b_i, A_i, A_{ij}$, *and T.*

The initial curve g_0 is continuous and hence bounded on compacts. In addition, by (4), the drift function b and the volatility function σ are continuous and have linear growth. Therefore, to prove this result we can follow the same argument as in the proof of [6, Lemma 3.1].

The following theorem guarantees the existence of solutions to (3). The next assumption is needed to state the result.

Assumption 1 *There exists a constant $\gamma \in (0, 2]$ such that $\int_0^h |K(t)|^2 dt = O(h^\gamma)$ and $\int_0^T |K(t+h) - K(t)|^2 dt = O(h^\gamma)$ for every $T < \infty$.*

Theorem 2.2 (Existence of polynomial Volterra processes) *Suppose that Assumption 1 holds. Then, for any $\alpha < \gamma/2$, (3) admits a weak solution X such that $X - g_0$ has α-Hölder continuous trajectories.*

This result is a consequence of [6, Theorem 3.4] and its proof, which can be adapted to the framework of an initial continuous deterministic curve g_0 instead of a constant initial condition $X_0 \in \mathbb{R}^d$. The main difference is that the proof presented in [6] relies on the existence of a resolvent of the first kind for the kernel K. This hypothesis is not necessary because the same tightness argument to construct the weak solution in [6] can be adapted by considering the integrated form of the equation (3)

$$\int_0^t X_s ds = \int_0^t g_0(s)ds + \int_0^t K(t-s)Z_s ds \tag{7}$$

with Z as in (5) instead of the form $L * (X - g_0) = Z$ with L the resolvent of the first kind of K. The integrated form (7) of the stochastic Volterra equation (3), which is suitable for stability results as illustrated in [2, Section 3], can be obtained using an argument based on the stochastic Fubini theorem as shown in [2, Lemma 3.2].

Example 2.3 *Assumption 1 is satisfied for the (possibly singular) fractional kernel $K(t) = t^{H-1/2}$ with $H \in (0,1)$. In particular, the corresponding polynomial Volterra process (3) fails to be a semimartingale whenever $H \neq 1/2$.*

2.1 Polynomial Volterra processes on the unit ball

In this section, we construct polynomial Volterra processes (3), that remain in the unit ball of $\mathbb{R}^d$ defined by

$$\mathcal{B} = \{x \in \mathbb{R}^d : x^\top x \leq 1\}.$$

For the rest of this section, we will assume that the kernel K is scalar $K : \mathbb{R}_+ \to \mathbb{R}$, and we use the so-called *resolvent of the first kind*, which is a measure L on $\mathbb{R}_+$ of locally bounded variation such that

$$K * L = L * K \equiv 1, \tag{8}$$

see [43, Definition 5.5.1]. Some examples of resolvents of the first kind are given in [6, Table 1]. A resolvent of the first kind does not always exist. For the main result of this section, we need an additional assumption for the kernel.

Assumption 2 *K is nonnegative, not identically zero, non-increasing and continuous on $(0, \infty)$, and its resolvent of the first kind L is nonnegative and non-increasing in the sense that $s \mapsto L([s, s+t])$ is non-increasing for all $t \geq 0$.*

Example 2.4 *If K is completely monotone on $(0, \infty)$ and not identically zero, then Assumption 2 holds due to [43, Theorem 5.5.4]. Recall that a function f is called completely monotone on $(0, \infty)$ if it is infinitely differentiable with $(-1)^k f^{(k)}(t) \geq 0$ for all $t > 0$ and $k \geq 0$. This covers, for instance, any constant positive kernel, the fractional kernel $t^{H-1/2}$ with $H \in (0, 1/2]$, and the exponentially decaying kernel $\mathrm{e}^{-\beta t}$ with $\beta > 0$. Moreover, sums and products of completely monotone functions are completely monotone.*

The next theorem provides the weak existence and uniqueness of a $\mathcal{B}$-valued polynomial Volterra process. The construction follows from a more general result for $\mathcal{B}$-valued stochastic Volterra equations given in Theorem 5.1.

Theorem 2.5 (Existence and uniqueness of polynomial Volterra processes in the unit ball) *Fix a scalar kernel $K : [0, T] \to \mathbb{R}$ that satisfies Assumptions 1 and 2. Assume that b and σ are such that*

$$b(x) = b_0 + Bx, \quad \sigma(x) = c\sqrt{1 - x^\top x}\,\mathrm{I}_d 1_{\{x \in \mathcal{B}\}}, \quad x \in \mathbb{R}^d,$$

where I_d is the $d \times d$ identity matrix, $c \in \mathbb{R}$, and $b_0 \in \mathbb{R}^d$ and $B \in \mathbb{R}^{d \times d}$ are such that

$$x^\top (b_0 + Bx) \leq 0 \quad x \in \partial\mathcal{B}. \tag{9}$$

For any $X_0 \in \mathcal{B}$, there exists a unique continuous weak solution X to (3) such that $X_t \in \mathcal{B}$ a.s. for every $t \in [0, T]$.

Straightfoward application of Theorem 5.1 and Theorem 3.12 below.

By restricting to the one-dimensional case, Theorem 2.5 allows the construction of Jacobi Volterra processes on general compact intervals as shown in the next corollary.

Corollary 2.6 *Let $\alpha_1 \leq \alpha_2, b \in [\alpha_1, \alpha_2], \lambda \geq 0$ and $c > 0$. Fix a scalar kernel $K : [0, T] \to \mathbb{R}$ that satisfies Assumptions 1 and 2. Then, there exists a unique weak $[\alpha_1, \alpha_2]$-valued solution to the equation*

$$Y_t = Y_0 + \lambda \int_0^t K(t-s)(b - Y_s)ds + c \int_0^t K(t-s)\sqrt{(Y_s - \alpha_1)(\alpha_2 - Y_s)}dW_s, \tag{10}$$

$$Y_0 \in [\alpha_1, \alpha_2], \tag{11}$$

that we call Jacobi Volterra process on $[\alpha_1, \alpha_2]$.

We note that in dimension $d = 1$ we have that $\mathcal{B} = [-1, 1]$. In this case, the specification of Example 50 yields an $[-1, 1]$-valued Volterra Jacobi process in the form

$$X_t = X_0 + \int_0^t K(t-s)\lambda(\tilde{b} - X_s)ds + \int_0^t K(t-s)c\sqrt{(1-X_s)(1+X_s)}dW_s,$$
$$X_0 \in [-1, 1],$$

for any $\tilde{b} \in [-1, 1]$. Indeed, in this case it is straightforward to check that (9) is satisfied since

$$x\lambda(\tilde{b} - x) = \begin{cases} \lambda(\tilde{b} - 1) \leq 0 & x = 1 \\ -1\lambda(\tilde{b} + 1) \leq 0 & x = -1. \end{cases} \tag{12}$$

Taking $\tilde{b}$ such that $b = \frac{\alpha_2 - \alpha_1}{2}\tilde{b} + \frac{\alpha_1 + \alpha_2}{2}$ we readily get that

$$Y = \frac{\alpha_2 - \alpha_1}{2}X + \frac{\alpha_1 + \alpha_2}{2}$$

is a Jacobi Volterra process on $[\alpha_1, \alpha_2]$ that satisfies (10).

Remark 2.7 *In [40] the authors study stochastic invariance of polynomial diffusions, that is solutions to* (3) *for the trivial kernel $K \equiv 1$, for more general state-spaces $E \subseteq \mathbb{R}^d$. In the case of $E = \mathcal{B}$, choosing $Q = \mathrm{I}_d$ and $c = 0$ in [40, Proposition 6.1] essentially corresponds to Theorem 2.5 for the trivial kernel. In contrast with the diffusion case, where a tagential diffusive behaviour of the volatility component σ is possible on the boundary, here we restrict to vanishing volatility at the boundary. Because of the possible singularity of the kernel at 0, we expect that any tangential diffusive behavior of the volatility σ at the boundary will push the process outside the ball $\mathcal{B}$.*

3   Moments of polynomial Volterra processes

Our aim is to find moment formulas for a continuous Volterra process X solving (3). More precisely, we want to obtain formulas for expressions of the form

$$\mathbb{E}[X_t^\alpha] = \mathbb{E}[X_{1,t}^{\alpha_1} \cdots X_{d,t}^{\alpha_d}], \quad t \geq 0, \tag{13}$$

with $\alpha = (\alpha_i)_{i=1}^d \in \mathbb{N}_0^d$ a multi-index. One of the difficulties to characterize these moments, compared to the classical framework when the kernel K is equal to I_d, stems from the fact that X is not necessarily a Markovian semimartingale. To circumvent this complication, and inspired by previous works such as [3, 24, 47, 55], for each $T \geq 0$, we consider the process (indexed in time by t)

$$g_t(T) = g_0(T) + \int_0^t K(T-s)dZ_s, \quad t \leq T, \tag{14}$$

for Z as in (5). The following lemma shows that we can control the moments of the process g.

Lemma 3.1 *Suppose that X is a continuous process solving* (3) *and define the processes g as in* (14)*. Then, for any $p \in \mathbb{N}$ and $0 \leq T \leq T'$,*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T}|g_t(T)|^p\right] \leq c \tag{15}$$

for some constant c which depends only on $\sup_{0 \leq t \leq T'}|g_0(t)|$, p, $K|_{[0,T']}$, b_i, A_i, A_{ij}, and T'.

It is sufficient to prove the inequality (15) for $p \geq 2$. Given $T \geq 0$, thanks to the Burkholder-Davis-Gundy and Jensen's inequalities

$$\mathbb{E}\left[\sup_{0 \leq t \leq T}\left|\int_0^t K(T-s)\sigma(X_s)dW_s\right|^p\right] \leq C\,\mathbb{E}\left[\left(\int_0^T |K(T-s)|^2|a(X_s)|ds\right)^{\frac{p}{2}}\right]$$

$$\leq C\left(\int_0^T |K(t)|^2 dt\right)^{\frac{p}{2}} \sup_{0 \leq t \leq T}\mathbb{E}[|a(X_t)|^{\frac{p}{2}}] \qquad (16)$$

$$\leq C\left(\int_0^{T'} |K(t)|^2 dt\right)^{\frac{p}{2}} \sup_{0 \leq t \leq T'}\mathbb{E}[|a(X_t)|^{\frac{p}{2}}]$$

for some constant $C > 0$. Similarly, multiple applications of Jensen's inequality yield

$$\mathbb{E}\left[\left(\sup_{0 \leq t \leq T}\left|\int_0^t K(T-s)b(X_s)ds\right|\right)^p\right] \leq (T')^{\frac{p}{2}}\left(\int_0^{T'} |K(t)|^2 dt\right)^{\frac{p}{2}} \sup_{0 \leq t \leq T'}\mathbb{E}[|b(X_t)|^p]. \qquad (17)$$

Since the functions b and a have the form (4), (16) and (17) together with Proposition 2.1 yield (15).

Notice that $g_T(T) = X_T$ and more generally

$$g_t(T) = \mathbb{E}_t\left[X_T - \int_t^T K(T-s)b(X_s)ds\right], \quad t \leq T, \qquad (18)$$

because – as a result of the proof of Lemma 3.1 above and in particular (16) – the process $M_t = \int_0^t K(T-s)\sigma(X_s)dW_s$, $t \leq T$, is a martingale. Moreover, for each $T \geq 0$, $g_t(T)$ is a semimartingale with dynamics

$$dg_t(T) = K(T-t)dZ_t = K(T-t)b(X_t)dt + K(T-t)\sigma(X_t)dW_t, \quad t < T. \qquad (19)$$

To study the moments of X in (13), we need to understand the behavior of more general moments of the (infinite dimensional) processes g defined in (14). To this end, we consider expressions of the form

$$\mathfrak{m}^{(p)}(t, T_1, \ldots, T_p; w) = \mathbb{E}\left[\prod_{n=1}^p g_{i_n, t}(T_n)\right] \qquad (20)$$

where $p \in \mathbb{N}$, $0 \leq t \leq \min\{T_1, \ldots, T_p\}$, $w = (i_n)_{n=1}^p \in \{1, \ldots, d\}^p$, and $g_{i_n, t}(T_n)$ is the i_n-th coordinate of $g_t(T_n)$. If $d = 1$, we can omit the argument w and write $\mathfrak{m}^{(p)}(t, T_1, \ldots, T_p)$ for $p \in \mathbb{N}$. We shall use the convention $\mathfrak{m}^{(0)} \equiv 1$.

Notice that

$$\mathfrak{m}^{(p)}(t, t, \ldots, t; w) = \mathbb{E}[X_t^{\alpha(w)}] \qquad (21)$$

where $\alpha(w)$ is the multi-index given by $\alpha_k(w) = \#\{n : i_n = k\}$, $k = 1, \ldots, d$. In particular, $|\alpha(w)| = p$.

3.1 The main moment formula

The main result of this section is a characterization for the functions $\mathfrak{m}$ defined in (20), which in view of (21) determine the moments. More precisely, we will see in Theorem 3.3 below, that for some fixed level N, the vector-valued function

$$(t, T_1, \ldots, T_N) \mapsto \{\mathfrak{m}^{(p)}(t, T_1, \ldots, T_p; w) : p \in \{0, \ldots, N\} \text{ and } w \in \{1, \ldots, d\}^p\} \qquad (22)$$

is the unique continuous solution to a specific integral equation. Referring to the terminology of [30], this thus gives an existence and uniqueness result for the bidual moment formula (which had to be assumed in the

generic infinite dimensional setting of [30]). In order to give a precise statement, we have to fix some notations. Given $N \in \mathbb{N}$, we define the set

$$\mathcal{I}^{(N)} = \{(p, w) : p \in \{0, 1, \dots, N\} \text{ and } w \in \{1, \dots, d\}^p\}. \tag{23}$$

Let D_N be the cardinality of the set $\mathcal{I}^{(N)}$. Notice that $D_N = N + 1$ for $d = 1$ and $D_N = \sum_{p=0}^{N} d^p = (d^{N+1} - 1)/(d - 1)$ for $d > 1$. For $T \geq 0$, let

$$\mathcal{D}_T^{(N)} = \{(t, T_1, \dots, T_N) \in [0, T]^{N+1} : t \leq \min\{T_1, \dots, T_N\}\}, \tag{24}$$

and let π be an enumeration of $\mathcal{I}^{(N)}$. As we will see in Theorem 3.3 below, the function in (22) belongs to the following space.

Definition 3.2 Let $\mathcal{X}_T^{(N)}$ be the space of $\mathbb{R}^{D_N}$-valued bounded functions $\mathbf{f}$ on $\mathcal{D}_T^{(N)}$ such that the $\pi(p, w)$-th component $\mathbf{f}_{\pi(p,w)}$ only depends on the variables $(t, T_1, \dots, T_p)$ for any $(p, w) \in \mathcal{I}^{(N)}$, and $\mathbf{f}_{\pi(0,\emptyset)} \equiv 1$.

Given $\mathbf{f} \in \mathcal{X}_T^{(N)}$, define the function $\mathcal{M}_T^{(N)}\mathbf{f}$ on $\mathcal{D}_T^{(N)}$ as follows: $(\mathcal{M}_T^{(N)}\mathbf{f})_{\pi(0,\emptyset)} \equiv 0$, and for $(p, w) \in \mathcal{I}^{(N)}$ with $1 \leq p \leq N$ and $w = (i_n)_{n=1}^p$,

$$
\begin{aligned}
(\mathcal{M}_T^{(N)}\mathbf{f})_{\pi(p,w)}(t, T_1, \dots, T_N) &= \sum_{n=1}^{p} \int_0^t e_{i_n}^\top K(T_n - r) b_0 \mathbf{f}_{\pi(p-1, w_{-n})}(r, (T_m)_{m \neq n}) dr \\
&+ \sum_{n=1}^{p} \sum_{j=1}^{d} \int_0^t e_{i_n}^\top K(T_n - r) b_j \mathbf{f}_{\pi(p, w_{-n}^j)}(r, r, (T_m)_{m \neq n}) dr \\
&+ \sum_{1 \leq n < m \leq p} \int_0^t e_{i_n}^\top K(T_n - r) A_0 K(T_m - r)^\top e_{i_m} \mathbf{f}_{\pi(p-2, w_{-n,-m})}(r, (T_l)_{l \neq m,n}) dr \\
&+ \sum_{1 \leq n < m \leq p} \sum_{j=1}^{d} \int_0^t e_{i_n}^\top K(T_n - r) A_j K(T_m - r)^\top e_{i_m} \mathbf{f}_{\pi(p-1, w_{-n,-m}^j)}(r, r, (T_l)_{l \neq m,n}) dr \\
&+ \sum_{1 \leq n < m \leq p} \sum_{j,k=1}^{d} \int_0^t e_{i_n}^\top K(T_n - r) A_{jk} K(T_m - r)^\top e_{i_m} \mathbf{f}_{\pi(p, w_{-n,-m}^{j,k})}(r, r, r, (T_l)_{l \neq m,n}) dr.
\end{aligned}
\tag{25}
$$

In (25), e_i is the i-th canonical vector in $\mathbb{R}^d$, w_{-n} is the vector obtained by erasing the n-th coordinate of w, $w_{-n,-m}$ is the vector obtained by erasing the n-th and m-th coordinates of w, w_{-n}^j is the vector whose first coordinate is j and the other coordinates are given by w_{-n}, and $w_{-n,-m}^{j,k}$ is the vector whose first two coordinates are j, k and the other coordinates are given by $w_{-n,-m}$. Notice that thanks to the local square-integrability of K the right side of (25) is well-defined for $\mathbf{f} \in \mathcal{X}_T^{(N)}$ and, moreover, $\mathcal{M}_T^{(N)}\mathbf{f} \in \mathcal{X}_T^{(N)}$.

The following theorem provides formulas for the moments defined in (20). In view of (21), this result also establishes relations between the moments of X in (13) and the moments of the process g.

Theorem 3.3 (Main moment formula) Fix $N \in \mathbb{N}$ and $T \geq 0$. Define $\mathbf{m} : \mathcal{D}_T^{(N)} \to \mathbb{R}^{D_N}$ by $\mathbf{m}_{\pi(0,\emptyset)} = \mathbf{m}^{(0)} \equiv 1$, and for $(p, w) \in \mathcal{I}^{(N)}$ with $1 \leq p \leq N$, $\mathbf{m}_{\pi(p,w)}(t, T_1, \dots, T_N) = \mathfrak{m}^{(p)}(t, T_1, \dots, T_p; w)$ as in (20). Then $\mathbf{m} \in \mathcal{X}_T^{(N)}$ (see Definition 3.2) and $\mathbf{m}$ solves the integral equation

$$\mathbf{m}(t, T_1, \dots, T_N) = \mathbf{m}(0, T_1, \dots, T_N) + (\mathcal{M}_T^{(N)}\mathbf{m})(t, T_1, \dots, T_N), \quad (t, T_1, \dots, T_N) \in \mathcal{D}_T^{(N)}, \tag{26}$$

with $\mathcal{M}_T^{(N)}$ as in (25). Furthermore, $\mathbf{m}$ is the unique solution in $\mathcal{X}_T^{(N)}$ of (26) with initial condition $\mathbf{m}(0, T_1, \dots, T_N)$, and $\mathbf{m}$ is continuous on $\mathcal{D}_T^{(N)}$.

To prove that $\mathbf{m} \in \mathcal{X}_T^{(N)}$ it is enough to show that $\mathbf{m}$ is bounded on $\mathcal{D}_T^{(N)}$. This is a consequence of Hölder's inequality and Lemma 3.1. Indeed, by Lemma 3.1, for any $i_n \in \{1, \ldots, d\}$, $p \in \mathbb{N}$, and $t \leq T_n \leq T$,

$$\mathbb{E}[|g_{i_n,t}(T_n)|^p] \leq \mathbb{E}\left[\sup_{0 \leq t \leq T_n} |g_{i_n,t}(T_n)|^p\right] \leq c$$

where c is constant depending only on $\sup_{0 \leq t \leq T} |g_0(t)|$, p, $K|_{[0,T]}$, b_i, A_i, A_{ij}, and T.

We now prove (26). Clearly this equation holds by definition over the coordinate $\pi(0, \emptyset)$. We consider then $(p, w) \in \mathcal{I}^{(N)}$ such that $1 \leq p \leq N$ and $w = (i_n)_{n=1}^p$. Itô's formula, together with (19), yields

$$d\left(\prod_{n=1}^p g_{i_n,r}(T_n)\right) = \sum_{n=1}^p \left(\prod_{m \neq n} g_{i_m,r}(T_m)\right) dg_{i_n,r}(T_n)$$

$$+ \sum_{1 \leq n < m \leq p} \left(\prod_{l \neq n,m} g_{i_l,r}(T_l)\right) d\langle g_{i_n,\cdot}(T_n), g_{i_m,\cdot}(T_m)\rangle_r. \tag{27}$$

The local martingale part in (27) is a true martingale thanks to (19), (4), the fact that $X_r = g_r(r)$, and Lemma 3.1. The finite variation part can be written as

$$\sum_{n=1}^p e_{i_n}^\top K(T_n - r) b_0 \left(\prod_{m \neq n} g_{i_m,r}(T_m)\right) dr$$

$$+ \sum_{n=1}^p \sum_{j=1}^d e_{i_n}^\top K(T_n - r) b_j X_{j,r} \left(\prod_{m \neq n} g_{i_m,r}(T_m)\right) dr$$

$$+ \sum_{1 \leq n < m \leq p} e_{i_n}^\top K(T_n - r) A_0 K(T_m - r)^\top e_{i_m} \left(\prod_{l \neq m,n} g_{i_l,r}(T_l)\right) dr$$

$$+ \sum_{1 \leq n < m \leq p} \sum_{j=1}^d e_{i_n}^\top K(T_n - r) A_j K(T_m - r)^\top e_{i_m} X_{j,r} \left(\prod_{l \neq m,n} g_{i_l,r}(T_l)\right) dr$$

$$+ \sum_{1 \leq n < m \leq p} \sum_{j,k=1}^d e_{i_n}^\top K(T_n - r) A_{jk} K(T_m - r)^\top e_{i_m} X_{j,r} X_{k,r} \left(\prod_{l \neq m,n} g_{i_l,r}(T_l)\right) dr.$$

Since $g_{i_n,r}(r) = X_{i_n,r}$, integrating on $[0, t]$, taking expectation, using Fubini's theorem – which can be applied

thanks to Lemma 3.1 – and by the definition of $\mathfrak{m}$ in (20), we obtain

$$
\mathfrak{m}^{(p)}(t, T_1, \ldots, T_p; w) = \prod_{n=1}^{p} g_{i_n, 0}(T_n)
$$

$$
+ \sum_{n=1}^{p} \int_0^t e_{i_n}^\top K(T_n - r) b_0 \mathfrak{m}^{(p-1)}(r, (T_m)_{m \neq n}; w_{-n}) dr
$$

$$
+ \sum_{n=1}^{p} \sum_{j=1}^{d} \int_0^t e_{i_n}^\top K(T_n - r) b_j \mathfrak{m}^{(p)}(r, r, (T_m)_{m \neq n}; w_{-n}^j) dr
$$

$$
+ \sum_{1 \leq n < m \leq p} \int_0^t e_{i_n}^\top K(T_n - r) A_0 K(T_m - r)^\top e_{i_m} \mathfrak{m}^{(p-2)}(r, (T_l)_{l \neq m,n}; w_{-n,-m}) dr \qquad (28)
$$

$$
+ \sum_{1 \leq n < m \leq p} \sum_{j=1}^{d} \int_0^t e_{i_n}^\top K(T_n - r) A_j K(T_m - r)^\top e_{i_m} \mathfrak{m}^{(p-1)}(r, r, (T_l)_{l \neq m,n}; w_{-n,-m}^j) dr
$$

$$
+ \sum_{1 \leq n < m \leq p} \sum_{j,k=1}^{d} \int_0^t e_{i_n}^\top K(T_n - r) A_{jk} K(T_m - r)^\top e_{i_m} \mathfrak{m}^{(p)}(r, r, r, (T_l)_{l \neq m,n}; w_{-n,-m}^{j,k}) dr.
$$

This is precisely the coordinate $\pi(p, w)$ of the identity (26), which concludes the proof of (26). The uniqueness of solutions in $\mathcal{X}_T^{(N)}$ of the integral equation (26) and the fact that $\mathbf{m}$ is continuous on $\mathcal{D}_T^{(N)}$ are a consequence of Corollary A.2 in Appendix A.

Remark 3.4 *If we define the function*

$$
\overline{\mathbf{m}}(t, x_1, \ldots, x_N) = \mathbf{m}(t, t + x_1, \ldots, t + x_N), \qquad (29)
$$

for $t, x_1, \ldots, x_N \geq 0$ such that $t + x_i \leq T$, $i = 1, \ldots, N$. Then, Theorem 3.3 implies that the function $\overline{\mathbf{m}}$ is a mild-solution to the following non local PDE

$$
\partial_t \overline{\mathbf{m}}_{\pi(p,w)}(t, x_1, \ldots, x_p) = (\partial_{x_1} + \cdots + \partial_{x_p}) \overline{\mathbf{m}}_{\pi(p,w)}(t, x_1, \ldots, x_p)
$$

$$
+ \sum_{n=1}^{p} e_{i_n}^\top K(x_n) b_0 \overline{\mathbf{m}}_{\pi(p-1,w_{-n})}(t, (x_m)_{m \neq n})
$$

$$
+ \sum_{n=1}^{p} \sum_{j=1}^{d} e_{i_n}^\top K(x_n) b_j \overline{\mathbf{m}}_{\pi(p,w_{-n}^j)}(t, 0, (x_m)_{m \neq n})
$$

$$
+ \sum_{1 \leq n < m \leq p} e_{i_n}^\top K(x_n) A_0 K(x_m)^\top e_{i_m} \overline{\mathbf{m}}_{\pi(p-2,w_{-n,-m})}(t, (x_l)_{l \neq m,n}) \qquad (30)
$$

$$
+ \sum_{1 \leq n < m \leq p} \sum_{j=1}^{d} e_{i_n}^\top K(x_n) A_j K(x_m)^\top e_{i_m} \overline{\mathbf{m}}_{\pi(p-1,w_{-n,-m}^j)}(t, 0, (x_l)_{l \neq m,n})
$$

$$
+ \sum_{1 \leq n < m \leq p} \sum_{j,k=1}^{d} e_{i_n}^\top K(x_n) A_{jk} K(x_m)^\top e_{i_m} \overline{\mathbf{m}}_{\pi(p,w_{-n,-m}^{j,k})}(t, 0, 0, (x_l)_{l \neq m,n}).
$$

Indeed, this can be deduced using the change of variables $T_n = t + x_n$ in (28) together with the definition of the function $\overline{\mathbf{m}}$ in (29).

Remark 3.5 *In the classical polynomial processes framework, where $K = \mathrm{I}_d$ and $g_0 \equiv X_0 \in \mathbb{R}^d$, the process g defined in (14) coincides with X, i.e. $g_t(T) = X_t$. Hence, in this case, the function $\mathbf{m}$ in Theorem 3.3 does not depend on $T_1, \ldots, T_N$. Consequently, the function $\overline{\mathbf{m}}$ defined in (29) does not depend on $x_1, \ldots, x_N$ and the PDE (30) reduces to a linear ODE with constant coefficients. When $d > 1$, the dimension of this linear*

ODE – which is $D_N = (d^{N+1} - d)/(d-1)$ in the framework of this study – can be reduced to $\binom{N+d}{N}$. This dimension reduction is possible because, as it can be seen from (21), multiple $w \in \{1,\dots,d\}^p$, $1 \le p \le N$, yield the same moments of X. The solution to this linear ODE can be expressed in terms of an exponential matrix. This observation establishes a relation, when $K = \mathrm{I}_d$, between Theorem 3.3 and the classical moment formula for polynomial processes as stated in [40, Theorem 3.1].

In the next subsection, using Theorem 3.3, we establish an alternative moment formula using a variation of constants technique. This moment formula is useful to characterize the first and second order moments of a polynomial Volterra process, and the moments of affine Volterra processes, namely when $A_{ij} = 0$ for all i, j in (4).

3.2   Moment formula using a variation of constants technique

Let B be a matrix with columns equal to $b_1, \dots, b_d$ and let R_B be the resolvent of $-KB$, i.e. the solution to the linear equation $KB * R_B = KB + R_B$. Thanks to [6, Lemma 2.5], (3) is equivalent to the integral equation

$$X_t = g_0(t) - \int_0^t R_B(t-s)g_0(s)ds + \left(\int_0^t E_B(s)ds \right) b_0 + \int_0^t E_B(t-s)\sigma(X_s)dW_s$$

where $E_B = K - R_B * K$. Applying Theorem 3.3 to this reformulation of the stochastic Volterra equation yields alternative moment formulas. Indeed, let

$$\widetilde{g}_t(T) = \widetilde{g}_0(T) + \int_0^t E_B(T-s)\sigma(X_s)dW_s = \mathbb{E}_t[X_T], \quad t \le T, \tag{31}$$

where $\widetilde{g}_0(t) = g_0(t) - \int_0^t R_B(t-s)g_0(s)ds + \left(\int_0^t E_B(s)ds \right) b_0$. Define further, for $(p, w) \in \mathcal{I}^{(N)}$, with $1 \le p \le N$ and $w = (i_n)_{n=1}^p$,

$$\widetilde{\mathfrak{m}}^{(p)}(t, T_1, \dots, T_p; w) = \mathbb{E}\left[\prod_{n=1}^p \widetilde{g}_{i_n,t}(T_n) \right].$$

Adopt as before the convention $\widetilde{\mathfrak{m}}^{(0)} \equiv 1$. Theorem 3.3 shows that for $(p, w) \in \mathcal{I}^{(N)}$, with $1 \le p \le N$ and $w = (i_n)_{n=1}^p$,

$$\widetilde{\mathfrak{m}}^{(p)}(t, T_1, \dots, T_p; w) = \prod_{n=1}^p \widetilde{g}_{i_n,0}(T_n)$$

$$+ \sum_{1 \le n < m \le p} \int_0^t e_{i_n}^\top E_B(T_n - r) A_0 E_B(T_m - r)^\top e_{i_m} \widetilde{\mathfrak{m}}^{(p-2)}(r, (T_l)_{l \ne m,n}; w_{-n,-m}) dr$$

$$+ \sum_{1 \le n < m \le p} \sum_{j=1}^d \int_0^t e_{i_n}^\top E_B(T_n - r) A_j E_B(T_m - r)^\top e_{i_m} \widetilde{\mathfrak{m}}^{(p-1)}(r, r, (T_l)_{l \ne m,n}; w_{-n,-m}^j) dr \tag{32}$$

$$+ \sum_{1 \le n < m \le p} \sum_{j,k=1}^d \int_0^t e_{i_n}^\top E_B(T_n - r) A_{jk} E_B(T_m - r)^\top e_{i_m} \widetilde{\mathfrak{m}}^{(p)}(r, r, r, (T_l)_{l \ne m,n}; w_{-n,-m}^{j,k}) dr.$$

The following two remarks present two important consequences of (32).

Remark 3.6 *This formulation of the moment formula yields more explicit expressions for the first and second order moments. Indeed, the definition of $\widetilde{g}$ in (31) provides directly a formula for the first order moments $\widetilde{\mathfrak{m}}^{(1)}$. Regarding the second order moments, observe that by taking $p = 2$ and $t = T_1 = T_2$ in (32), we obtain a*

linear system of integral convolution equations for the functions $\mathbf{f}_{\pi(2,w)}(t) = \widetilde{\mathfrak{m}}^{(2)}(t,t,t;w)$, $w = (i_n)_{n=1}^2 \in \{1,\dots,d\}^2$. *More precisely,*

$$\mathbf{f}_{\pi(2,w)}(t) = \prod_{n=1}^{2} \widetilde{g}_{i_n,0}(t) + \int_0^t e_{i_1}^\top E_B(t-r) A_0 E_B(t-r)^\top e_{i_2} dr$$

$$+ \sum_{j=1}^{d} \int_0^t e_{i_1}^\top E_B(t-r) A_j E_B(t-r)^\top e_{i_2} \widetilde{\mathfrak{m}}^{(1)}(r,r;w_{-n,-m}^j) dr$$

$$+ \sum_{j,k=1}^{d} \int_0^t e_{i_1}^\top E_B(t-r) A_{jk} E_B(t-r)^\top e_{i_2} \mathbf{f}_{\pi(2,w_{-n,-m}^{j,k})}(r) dr.$$

This linear system of convolution equations can be solved using the resolvent of the associated (matrix) kernel. In addition, thanks to (26), all the second order moments $\widetilde{\mathfrak{m}}^{(2)}$ *can be expressed in terms of* $\widetilde{\mathfrak{m}}^{(1)}$ *and the functions* $\mathbf{f}_{\pi(2,w)}$, $w \in \{1,\dots,d\}^2$.

Remark 3.7 *If* X *is an affine Volterra process, i.e.* $A_{jk} = 0$, *then (32) provides a recursive algorithm to find the moments* $\widetilde{\mathfrak{m}}^{(p)}$ *of any order* $p \in \mathbb{N}$.

Remark 3.5 explained how the moment formula in Theorem 3.3 extends the moment formula from the classical to the Volterra framework. The next subsection elucidates that there is one important structural property that is common to the classical and Volterra settings.

3.3 Moments of polynomial Volterra processes are polynomials

For a polynomial diffusion X starting at X_0, i.e. when $K \equiv \mathrm{I}_d$ and $g_0 \equiv X_0$ in our framework, [40, Theorem 3.1] gives the following explicit formula for the moments of X

$$\mathbb{E}[p(X_t)] = H(X_0)^T e^{tG} \vec{p}, \quad p \in \mathrm{Pol}_n(\mathbb{R}^d), \, t \geq 0. \tag{33}$$

In (33), $H(x)$ is a vector whose components are elements of a basis for $\mathrm{Pol}_n(\mathbb{R}^d)$, G is the matrix of the infinitesimal generator of X restricted to $\mathrm{Pol}_n(\mathbb{R}^d)$, and $\vec{p}$ are the coordinates of the polynomial p with respect to the basis in $H(x)$. As a consequence, the moments of a polynomial diffusion X with initial value X_0 are again polynomials in the variable X_0. More precisely,

$$\mathbb{E}[X_t^\alpha] = \sum_{\beta \in \mathbb{N}_0^d, |\beta| \leq |\alpha|} c_\beta(t) X_0^\beta, \tag{34}$$

for some deterministic and time-dependent family of coefficients $\{c_\beta : \beta \in \mathbb{N}_0^d, |\beta| \leq |\alpha|\}$, which can be computed explicitly from (33). It turns out that this structural property still holds for the moments $\mathfrak{m}^{(p)}$ in (20) of a polynomial Volterra process X starting at X_0, i.e. when $g_0 \equiv X_0$ in (3). Furthermore, by an application of Theorem 3.3, the coefficients can be obtained by solving an integral equation similar to (26). Indeed, for $p = |\alpha| = 0$ this trivially holds true. For $p = 1$ and $w = i_1 \in \{1,\dots,d\}$, we have

$$\mathfrak{m}^{(1)}(t,T_1;w) = \mathbb{E}[X_{i_i,T_1}] - e_{i_1}^\top \int_t^{T_1} K(T_1 - s) b(\mathbb{E}[X_s]) ds, \quad 0 \leq t \leq T_1, \tag{35}$$

where we used (18) and the fact that $b \in \mathrm{Pol}_1(\mathbb{R}^d)$. Plugging the identity (31) into (35) yields

$$\mathfrak{m}^{(1)}(t,T_1;w) = e_{i_1}^\top \left(\mathrm{I}_d - \int_0^{T_1} R_B(s) ds - \int_t^{T_1} K(T_1-s) B \left(\mathrm{I}_d - \int_0^s R_B(u) du \right) ds \right) X_0$$

$$+ e_{i_1}^\top \left(\int_0^{T_1} E_B(s) ds - \int_t^{T_1} K(T_1-s) \left(\mathrm{I}_d + B \int_0^s E_B(u) du \right) ds \right) b_0.$$

This readily shows the representation of the form (34) for first-order moments. In the following theorem we exploit the moment-formula in Theorem 3.3, to generalize this result for higher order moments.

Before stating the theorem, we recall the notation at the beginning of Subsection 3.1 and observe that for $p \in \mathbb{N}$ and $w = (i_n)_{n=1}^p \in \{1, \ldots, d\}^p$

$$\mathfrak{m}^{(p)}(0, T_1, \ldots, T_p; w) = X_0^{\alpha(w)}, \tag{36}$$

where $\alpha(w)$ is the multi-index given by $\alpha_k(w) = \#\{n : i_n = k\}$, $k = 1, \ldots, d$.

Theorem 3.8 (Moments are polynomials) *Suppose that X is a continuous solution to* (3) *with $g_0 \equiv X_0$ and fix $T \geq 0$. Then, for all $p \in \mathbb{N}_0$ and all $w \in \{1, \ldots, d\}^p$, we can express the functions $\mathfrak{m}^{(p)}$ in* (20) *as*

$$\mathfrak{m}^{(p)}(t, T_1, \ldots, T_p; w) = \sum_{\beta \in \mathbb{N}_0^d, |\beta| \leq p} C_\beta^{(p)}(t, T_1, \ldots, T_p; w) X_0^\beta, \quad (t, T_1, \ldots, T_p) \in \mathcal{D}_T^{(p)}, \tag{37}$$

where the functions $C_\beta^{(p)}(\cdot, \cdots, \cdot; w) \in C(\mathcal{D}_T^{(p)}, \mathbb{R})$ are independent of X_0.
In particular, for any $\alpha \in \mathbb{N}_0^d$, there exists a family $\{c_\beta\}_{\beta \in \mathbb{N}_0^d, |\beta| \leq |\alpha|}$ of real-valued continuous functions on $[0, T]$, independent of X_0, such that

$$\mathbb{E}[X_t^\alpha] = \sum_{\beta \in \mathbb{N}_0^d, |\beta| \leq |\alpha|} c_\beta(t) X_0^\beta, \quad t \geq 0. \tag{38}$$

Furthermore – for any $p \in \mathbb{N}$, $w = (i_n)_{n=1}^p \in \{1, \ldots, d\}^p$, and $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq p$ – we have

$$\begin{aligned}
C_\beta^{(p)}(t, T_1, \ldots, T_p; w) &= \mathbf{f}_\beta^{(p)}(t, T_1, \ldots, T_p; w) \\
&+ \sum_{n=1}^p \sum_{j=1}^d \int_0^t e_{i_n}^\top K(T_n - r) b_j C_\beta^{(p)}(r, r, (T_m)_{m \neq n}; w_{-n}^j) dr \\
&+ \sum_{1 \leq n < m \leq p} \sum_{j,k=1}^d \int_0^t e_{i_n}^\top K(T_n - r) A_{jk} K(T_m - r)^\top e_{i_m} C_\beta^{(p)}(r, r, r, (T_l)_{l \neq m,n}; w_{-n,-m}^{j,k}) dr.
\end{aligned} \tag{39}$$

where

$$\begin{aligned}
\mathbf{f}_\beta^{(p)}(t, T_1, \ldots, T_p; w) &= 1_{\{\beta = \alpha(w)\}} + \sum_{n=1}^p \int_0^t e_{i_n}^\top K(T_n - r) b_0 C_\beta^{(p-1)}(r, (T_m)_{m \neq n}; w_{-n}) dr 1_{\{|\beta| < p\}} \\
&+ \sum_{1 \leq n < m \leq p} \int_0^t e_{i_n}^\top K(T_n - r) A_0 K(T_m - r)^\top e_{i_m} C_\beta^{(p-2)}(r, (T_l)_{l \neq m,n}; w_{-n,-m}) dr 1_{\{|\beta| < p-1\}} \\
&+ \sum_{1 \leq n < m \leq p} \sum_{j=1}^d \int_0^t e_{i_n}^\top K(T_n - r) A_j K(T_m - r)^\top e_{i_m} C_\beta^{(p-1)}(r, r, (T_l)_{l \neq m,n}; w_{-n,-m}^j) dr 1_{\{|\beta| < p\}}.
\end{aligned} \tag{40}$$

Remark 3.9 *Referring to the terminology used in [30], (39) provides a unique solution to the dual moment formula and $(C_\beta^{(p)}(t, T_1, \ldots, T_p; w))_{t \in [0,T]}$ can be interpreted as deterministic dual process.*

Notice that, thanks to (21), (38) directly follows from (37) after taking $t = T_1 = \cdots = T_p$. The arguments before the statement of the theorem show that (37) holds for $p \in \{0, 1\}$. Reasoning by induction, assume that for each $1 \leq q < p$ we have constructed continuous functions $C_\beta^{(q)}(\cdot, \ldots, \cdot; w)$ on $\mathcal{D}_T^{(q)}$ such that (37) holds with p replaced by q. Corollary A.3 in Appendix A shows that, for any $\beta \in \mathbb{N}_0^d$ such that $|\beta| \leq p$, the system of

equations (39) – seen as a system indexed over the elements $w \in \{1, \ldots, d\}^p$ – with initial condition (40), has a unique solution such that $C_\beta^{(p)}(\cdot, \ldots, \cdot; w) \in C(\mathcal{D}_T^{(p)})$.

It is straightforward to check that the family of functions on $\mathcal{D}_T^{(p)}$ defined as

$$\mathfrak{n}^{(q)}(t, T_1, \ldots, T_p; w) = \sum_{\beta \in \mathbb{N}_0^d, |\beta| \leq q} C_\beta^{(q)}(t, T_1, \ldots, T_p; w) X_0^\beta, \quad (q, w) \in \mathcal{I}^{(p)},$$

are a continuous solution (in $\mathcal{X}_T^{(p)}$) to the moment-equation (26), with N replaced by p and with initial conditions $X_0^{\alpha(w)}$ for each $(q, w) \in \mathcal{I}^{(p)}$. Notice that the initial conditions are the same for the moments $\mathfrak{m}^{(q)}$ by (36). Since by Theorem 3.3 the solution to this moment equation is unique, we conclude that $\mathbf{m} = \mathbf{n}$ and (37) holds.

Remark 3.10 *Observe that the integral equations verified by the coefficients C_β in (39) resemble the equations satisfied by the moments in (26). This structural property is well-known for classical polynomial processes. Indeed, in the classical case the coefficients of the moments in (33) are given in terms of the exponential matrix e^{tG} which solves a linear ODE as well. We could have vectorized the equations for the coefficients (39) as we did it for the moments in (26). We opted, however, to not use the vectorization at this point for clarity of exposition.*

3.4 Moments of the finite dimensional distributions and uniqueness in law

In this section we explain how the previous ideas can be extended to characterize moments of the finite dimensional distributions of a weak solution to (3). These considerations will allow us to prove a result concerning the uniqueness in law for solutions to (3).

Fix $T \geq 0$, $0 \leq t_1 \leq T$ and α^1 a multi-index. Define

$$\mathcal{D}_{t_1, T}^{(N)} = \{(t_2, T_1, \ldots, T_N) \in [t_1, T]^{N+1} : t_2 \leq \min\{T_1, \ldots, T_N\}\}.$$

Consider the $\mathbb{R}^{D_N}$-valued function $\mathbf{f}$ on $\mathcal{D}_{t_1, T}^{(N)}$ given by $\mathbf{f}_{\pi(0, \emptyset)} \equiv 1$ and, for $1 \leq p \leq N$ and $w = (i_n)_{n=1}^p$,

$$\mathbf{f}_{\pi(p, w)}(t_2, T_1, \ldots, T_N) = \mathbf{f}_{\pi(p, w)}(t_2, T_1, \ldots, T_p) = \mathbb{E}[X_{t_1}^{\alpha^1} g_{i_1, t_2}(T_1) \cdots g_{i_p, t_2}(T_p)].$$

Then - by the same considerations as in the proof of Theorem 3.3 – $\mathbf{f}$ is a bounded function on $\mathcal{D}_{t_1, T}^{(N)}$ satisfying the integral equation

$$\begin{aligned}
\mathbf{f}_{\pi(p, w)}(t_2, T_1, \ldots, T_p) &= \mathbb{E}[X_{t_1}^{\alpha^1} g_{i_1, t_1}(T_1) \cdots g_{i_p, t_1}(T_p)] \\
&+ \sum_{n=1}^p \int_{t_1}^{t_2} e_{i_n}^\top K(T_n - r) b_0 \mathbf{f}_{\pi(p-1, w_{-n})}(r, (T_m)_{m \neq n}) dr \\
&+ \sum_{n=1}^p \sum_{j=1}^d \int_{t_1}^{t_2} e_{i_n}^\top K(T_n - r) b_j \mathbf{f}_{\pi(p, w_{-n}^j)}(r, r, (T_m)_{m \neq n}) dr \\
&+ \sum_{1 \leq n < m \leq p} \int_{t_1}^{t_2} e_{i_n}^\top K(T_n - r) A_0 K(T_m - r)^\top e_{i_m} \mathbf{f}_{\pi(p-2, w_{-n, -m})}(r, (T_l)_{l \neq m, n}) dr \\
&+ \sum_{1 \leq n < m \leq p} \sum_{j=1}^d \int_{t_1}^{t_2} e_{i_n}^\top K(T_n - r) A_j K(T_m - r)^\top e_{i_m} \mathbf{f}_{\pi(p-1, w_{-n, -m}^j)}(r, r, (T_l)_{l \neq m, n}) dr \\
&+ \sum_{1 \leq n < m \leq p} \sum_{j, k=1}^d \int_{t_1}^{t_2} e_{i_n}^\top K(T_n - r) A_{jk} K(T_m - r)^\top e_{i_m} \mathbf{f}_{\pi(p, w_{-n, -m}^{j,k})}(r, r, r, (T_l)_{l \neq m, n}) dr
\end{aligned}$$

where $1 \leq p \leq N$ and $w = (i_n)_{n=1}^p$. The same arguments as in Appendix A show that, for given inputs (K, b, a), this equation has a unique bounded solution with initial condition $\mathbb{E}[X_{t_1}^{\alpha^1} g_{i_1, t_1}(T_1) \cdots g_{i_p, t_1}(T_p)]$. Since we already know that this initial condition only depends on (K, b, a), i.e. it is the same for any weak solution X to (3), we conclude that $\mathbf{f}$ only depends on (K, b, a). In particular, the mixed moments of the form $\mathbb{E}[X_{t_1}^{\alpha^1} X_{t_2}^{\alpha^2}]$ are the same for any weak solution X to (3).

Remark 3.11 *It is not surprising that the function $\mathbf{f}$ defined above solves an integral equation similar to* (26)*, with initial condition expressed in terms of simple moments at time t_1. Indeed, consider the classical case where $K \equiv \mathrm{I}_d$. Then, the tower property of the conditional expectation and the classical conditional moment formula yield*

$$\mathbb{E}[X_{t_1}^{\alpha^1} X_{t_2}^{\alpha^2}] = \mathbb{E}[X_{t_1}^{\alpha^1} H(X_{t_1})] \mathrm{e}^{(t_2 - t_1)G} \vec{p}$$

where we use the same notation as in (33) *with $n \geq |\alpha^2|$, and where $\vec{p}$ are the coordinates of the monomial x^{α^2} with respect to the chosen basis. Observe that $\mathbb{E}[X_{t_1}^{\alpha^1} H(X_{t_1})] \mathrm{e}^{(t_2 - t_1)G}$ solves a linear ODE with initial condition at t_1 given by $\mathbb{E}[X_{t_1}^{\alpha^1} H(X_{t_1})]$.*

A recursive argument, following the same lines as the above mentioned considerations, proves that all weak solutions X of (3) have the same moments of the form

$$\mathbb{E}[X_{t_1}^{\alpha^1} \cdots X_{t_l}^{\alpha^l}]$$

for $l \in \mathbb{N}$, times $t_1 \leq \cdots \leq t_l$ and multi-indices $\alpha^1, \ldots, \alpha^l$. Using this observation and arguing as in the proof of [40, Lemma 4.1], we deduce the following result.

Theorem 3.12 (Uniqueness in law for polynomial Volterra processes) *Suppose that for any weak solution X to* (3) *and for any $t \geq 0$, the law of X_t is determined by its moments. Then uniqueness in law holds for* (3)*.*

Remark 3.13 *As explained in the proof of [40, Lemma 4.1], the hypothesis of Theorem 3.12 holds for instance if for any weak solution X to* (3) *and $t \geq 0$, there is $\epsilon > 0$ such that $\mathbb{E}[\exp(\epsilon |X_t|)] < \infty$. For example, this is the case if any such solution X is bounded.*

Remark 3.14 *Theorem 3.12 constitutes a new result regarding uniqueness in law for solutions to stochastic Volterra equations. For affine Volterra processes [6] provide uniqueness in law via the Fourier-Laplace transform. Pathwise uniqueness for stochastic Volterra equations has been established for Lipschitz coefficients e.g., in [21] and, for certain coefficients in the one-dimensional case, when the kernel K is regular [4, 53] and when the kernel is singular in [46, 56].*

4 Jump representation of polynomial Volterra processes

Let X be the Volterra process introduced in (3) whose dynamics are given by

$$X_t = \lambda_0(t) + \int_0^t K(t-s)b(X_s)ds + \int_0^t K(t-s)\sigma(X_s)dW_s, \quad t \in [0, T].$$

In this section we consider a different lift, denoted by λ that will provide an alternative representation of the moments of X, namely in terms of a dual process which is a pure jump process with killing. For this reason we also call the initial condition λ_0 instead of g_0.

4.1   One-dimensional case

For the reader's convenience we consider first the one dimensional case setting $d = 1$. Let $(\lambda_t)_{t \in [0,T]}$ be a solution (in a sense made precise in Proposition 4.1 below) of

$$\lambda_t(x) = \lambda_0(x) + \int_0^t \partial_x \lambda_u(x) du + \int_0^t K(x) \Big[b(\lambda_u(0)) du + \sigma(\lambda_u(0)) dW_u \Big], \quad \lambda_0 \equiv X_0 \in \mathbb{R}, \quad (41)$$

for $x \in \mathbb{R}_+$. Similarly as in [32, Section 5] and as argued in Remark 4.2 below it holds that $\lambda_t(0) = X_t$ and that λ actually corresponds to the *Musiela parameterization* of the processes $(g_t(T))_{t \in [0,T]}$ considered in Section 3.

As a first step we introduce an appropriate space for the function-valued process $(\lambda_t)_{t \in [0,T]}$. As in [19, 20], we let $\alpha : \mathbb{R}_+ \to [1, \infty)$ be a nondecreasing C^1-function such that $\alpha^{-1} \in L^1(\mathbb{R}_+)$. The so-called Filipović space is then defined by

$$B = \{ y \in AC(\mathbb{R}_+, \mathbb{R}) : \|y\|_\alpha < \infty \},$$

where $AC(\mathbb{R}_+, \mathbb{R})$ denotes the space of absolutely continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$ and

$$\|y\|_\alpha^2 = |y(0)|^2 + \int_0^\infty |y'(x)|^2 \alpha(x) dx.$$

By [19, Lemma 3.2] we also know that $B \subseteq \mathbb{R} + C_0(\mathbb{R}_+)$, namely the space of bounded continuous functions with continuous continuation to infinity. Furthermore, let $(S_t)_{t \in [0,T]}$ denote the left-shift semigroup on B, i.e., $S_t y = y(t+\cdot)$. Then $(S_t)_{t \in [0,T]}$ is the C_0-semigroup generated by the operator ∂_x (see Filipović [38, Theorem 5.1.1]).

To ease technicalities and the exposition we make the following assumptions throughout this section (unless otherwise stated).

Assumption 3   *1 We assume that the kernel K satisfies $K \in B$.*

2 Both b and σ in (41) *are linear maps. We thus set $b(y) = b_1 y$ and $\sigma(y)^2 = A_{11} y^2$ for constants $b_1, A_{11} \in \mathbb{R}$.*

With the shift semigroup at hand we can now define the notion of a mild solution to (41) as in [33, Section 7.1]:

A predictable B-valued process $(\lambda_t)_{t \in [0,T]}$, is said to be a *mild solution* to (41) if for every $t \in [0,T]$

$$\mathbb{P}\left[\int_0^t \|\lambda_s\|_\alpha^2 ds < \infty \right] = 1,$$

and

$$\lambda_t = S_t \lambda_0 + \int_0^t S_{t-u} K b_1 \lambda_u(0) du + \int_0^t S_{t-u} K \sqrt{A_{11}} \lambda_u(0) dW_u, \quad \mathbb{P}\text{-a.s.} \quad (42)$$

Proposition 4.1 (Existence and moment bounds for solutions of (41)**)** *Under Assumption 3 there exists a unique mild solution to* (41) *which has a continuous modification. Moreover, for every $k \geq 1$, we have*

$$\mathbb{E}[\sup_{t \in [0,T]} \|\lambda_t\|_\alpha^k] \leq C_{T,k,\lambda_0}, \quad (43)$$

where C_{T,k,λ_0} is a constant that depends on T, k and λ_0. This solution is also an (analytically) weak solution in the sense of [33, page 161].

We apply [33, Theorem 7.2] for the Hilbert space $H = B$ and $U_0 = \mathbb{R}$. We thus need to verify the conditions of [33, Hypotheses 7.1]. As stated above ∂_x generates the strongly continuous right shift semigroup $(S_t)_{t \in [0,T]}$. Moreover, all measurability criteria are satisfied. Furthermore, since the point evaluations at 0 are bounded linear functionals and the fact that $K \in B$ we have

$$\|Kb_1 g(0) - Kb_1 h(0)\|_\alpha + \|K\sqrt{A_{11}}g(0) - K\sqrt{A_{11}}h(0)\|_\alpha \leq C\|g - h\|_\alpha, \quad \forall g, h \in B,$$

and of course also the linear growth condition. This implies the existence and uniqueness of a mild solution (42). Concerning the moment estimate we apply [33, Equation 7.7] stating that for $p > 2$

$$\mathbb{E}[\sup_{t \in [0,T]} \|\lambda_t\|_\alpha^p] \leq \widehat{C}_{T,p}(1 + \mathbb{E}[\|\lambda_0\|_\alpha^p]),$$

By the initial condition $\lambda_0 \equiv X_0 \in \mathbb{R}$ we get a constant on the right hand side. Moreover, since for $1 \leq k \leq 2$ and $p > 2$

$$\|\lambda_t\|_\alpha^k \leq 1 + \|\lambda_t\|_\alpha^p$$

we obtain

$$\mathbb{E}[\sup_{t \in [0,T]} \|\lambda_t\|_\alpha^k] \leq 1 + \mathbb{E}[\sup_{t \in [0,T]} \|\lambda_t\|_\alpha^p] \leq C_{T,k,\lambda_0},$$

and we get the assertion for all $k \geq 1$. The last assertion concerning the (analytically) weak solution follows from the same arguments as in [33, Theorem 6.5] since

$$\mathbb{E}\left[\int_0^T \|K\sqrt{A_{11}}\lambda_t(0)\|_\alpha^2\right] < \infty.$$

Remark 4.2 *Since under Assumption 3 $(\lambda_t)_{t \in [0,T]}$ is a mild solution to (41) with values in B we immediately get from (42) by evaluating at some fixed x that*

$$\lambda_t(x) = \lambda_0(t + x) + \int_0^t K(t + x - u)b_1\lambda_u(0)du + \int_0^t K(t + x - u)\sqrt{A_{11}}\lambda_u(0)dW_u.$$

From this we see that $\lambda_t(x) = g_t(t + x)$ where the process g was the defined in (14), implying that λ just corresponds to the Musiela parameterization of the processes g. Moreover, $(\lambda_t(0))_{t \in [0,T]}$ is the unique solution to the Volterra equation

$$X_t = X_0 + \int_0^t K(t - u)(b_1 X_u du + \sqrt{A_{11}}X_u dW_u),$$

where uniqueness is a consequence of the linearity and thus Lipschitz property of the coefficients.

We are now ready to state in Proposition 4.3 the expression of the moments of the polynomial Volterra process in terms of a functional of a pure jump process with killing. In the proof, we use the notation $\lambda^{\otimes k}(x)$ to denote the product $\lambda(x_1) \cdots \lambda(x_k)$ for $x = (x_1, \ldots, x_k) \in \mathbb{R}_+^k$.

Proposition 4.3 (Jump representation of one-dimensional polynomial Volterra processes) *Fix $k \geq 1$ and let $(Y_t)_{t \in [0,T]}$ be the $[0,T]^k$-valued process generated by $\mathcal{G}_k$ with domain $\mathcal{D}(\mathcal{G}_k) = AC([0,T]^k, \mathbb{R})$ given by*

$$\mathcal{G}_k f(x) = 1^\top \nabla f(x) + \int (f(\xi) - f(x))\nu(x, d\xi),$$

for

$$\nu(x, \cdot) = b_1 \sum_{i=1}^k K(x_i)\delta_{(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_k)}$$

$$+ A_{11} \sum_{i=1}^k \sum_{i<j} K(x_i)K(x_j)\delta_{(x_1,\ldots,x_{i-1},0,x_{i+1},\ldots,x_{j-1},0,x_{j+1},\ldots,x_k)},$$

and $Y_0 = (0, \ldots, 0)$. *Suppose that Assumption 3 is in force and assume that the mild solution* $(\lambda_t)_{t \in [0,T]}$ *to*
(41) *additionally satisfies*

$$\mathbb{E}[\sup_{x \in [0,T]^k} |\lambda_t'(x_1)\lambda_t(x_2)\cdots\lambda_t(x_k)|] < \infty. \tag{44}$$

Then for each $t \in [0, T]$ *it holds*

$$\mathbb{E}[X_t^k] = X_0^k \mathbb{E}\left[\exp\left(\int_0^t \kappa(Y_\tau)d\tau\right)\right],$$

where

$$\kappa(x) = b_1 \sum_{i=1}^k K(x_i) + A_{11} \sum_{i=1}^k \sum_{j<i} K(x_i)K(x_j).$$

Set

$$\mathcal{M}_k f(x) = \mathcal{G}_k f(x) + \kappa(x)f(x),$$

for each $f \in \mathcal{D}(\mathcal{G}_k)$. Observe that by Itô's formula and (41) (since $(\lambda_t)_{t \in [0,T]}$ is also an analytically weak solution), we have

$$\lambda_t^{\otimes k}(x) = X_0^k + \int_0^t \mathcal{M}_k \lambda_u^{\otimes k}(x)du + \int_0^t \mathcal{Q}\lambda_u^{\otimes k}(x)dW_u$$

for

$$\mathcal{Q}f(x) = \sqrt{A_{11}} \sum_{i=1}^k K(x_i)f(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_k).$$

Since

$$\mathbb{E}[\sup_{t \in [0,T]} |\lambda_t(x_1)\lambda_t(x_2)\cdots\lambda_t(x_k)|] < \infty$$

due to (43) and as $K \in B$ the third term is a true martingale and we thus have

$$\mathbb{E}[\lambda_t^{\otimes k}(x)] = X_0^k + \int_0^t \mathbb{E}[\mathcal{M}_k \lambda_u^{\otimes k}(x)]du.$$

Set then $m_t(x) = \mathbb{E}[\lambda_t^{\otimes k}(x)]$ for each $x \in [0,T]^k$. Since (44) holds, Leibniz rule and Fubini yield

$$\mathcal{M}_k m_t(x) = 1^\top \nabla \mathbb{E}[\lambda_t^{\otimes k}(\cdot)](x) + \int \mathbb{E}[\lambda_t^{\otimes k}(\xi)]\nu(x, d\xi)$$

$$= \mathbb{E}[1^\top \nabla \lambda_t^{\otimes k}(x)] + \mathbb{E}\left[\int \lambda_t^{\otimes k}(\xi)\nu(x, d\xi)\right],$$

proving that $\mathcal{M}_k m_t(x) = \mathbb{E}[\mathcal{M}_k \lambda_t^{\otimes k}(x)]$ and thus that

$$m_t(x) = X_0^k + \int_0^t \mathcal{M}_k m_u(x)du.$$

Note that since $m_s \in \mathcal{D}(\mathcal{G}_k)$ for each $s \in [0,T]$, by Itô's formula we get

$$dm_s(Y_t) = \mathcal{G}_k m_s(Y_t)dt + dM_t(m_s),$$

for

$$M_t(f) = \int_0^t \int \left(f(\xi) - f(Y_t)\right)(\mu_t(dt, d\xi) - \nu(Y_t, d\xi)dt),$$

where μ_t is the jump measure corresponding to Y. Fixing now $t \in [0,T]$ we thus get that

$$dm_{t-s}(Y_s) = -\partial_s m_{t-s}(Y_s)ds + \mathcal{G}_k m_{t-s}(Y_s)ds + dM_s(m_{t-s})$$
$$= -\mathcal{M}_k m_{t-s}(Y_s)ds + \mathcal{G}_k m_{t-s}(Y_s)ds + dM_s(m_{t-s}),$$

and similarly for $Z_s = \exp(\int_0^s \kappa(Y_\tau)d\tau)m_{t-s}(Y_s)$ we get

$$
\begin{aligned}
dZ_s = &- \exp\left(\int_0^s \kappa(Y_\tau)d\tau\right)\mathcal{M}_k m_{t-s}(Y_s)ds \\
&+ \exp\left(\int_0^s \kappa(Y_\tau)d\tau\right)\mathcal{G}_k m_{t-s}(Y_s)ds + \exp\left(\int_0^s \kappa(Y_\tau)d\tau\right)dM_s(m_{t-s}) \\
&+ \kappa(Y_s)\exp\left(\int_0^s \kappa(Y_\tau)d\tau\right)m_{t-s}(Y_s)ds \\
= &\exp\left(\int_0^s \kappa(Y_\tau)d\tau\right)dM_s(m_{t-s}),
\end{aligned}
$$

showing that $(Z_s)_{s\in[0,t]}$ is a local martingale. Since $\sup_{t,x\in[0,T]}|m_t(x)| < \infty$ and K is bounded we can conclude that

$$
\mathbb{E}\left[\exp\left(\int_0^t \kappa(Y_\tau)d\tau\right)X_0^k\right] = \mathbb{E}[Z_t] = \mathbb{E}[Z_0] = m_t(0,\dots,0) = \mathbb{E}[\lambda_t^{\otimes k}(0,\dots,0)] = \mathbb{E}[X_t^k].
$$

Example 4.4 *Note that for $K \equiv \lambda_0 \equiv 1$ we get that $X = X_0\mathcal{E}((b_1 t + \sqrt{A_{11}}W_t)_{t\in[0,T]})$. The corresponding representation yields*

$$
\mathbb{E}[X_t^k] = X_0^k \exp\left(\int_0^t b_1 k + A_{11}k(k-1)d\tau\right) = X_0^k \exp((b_1 k + \frac{1}{2}A_{11}k(k-1))t),
$$

as expected.

Remark 4.5 *Observe that Proposition 4.3 shows that in the current homogeneous case where both b and σ are linear functions, the kth moment of the Volterra process is just a monomial in X_0^k with coefficient $\mathbb{E}[\exp(\int_0^t \kappa(Y_\tau)d\tau)]$. This is thus a special case of Theorem 3.8. Note that the proof of Proposition 4.3 also shows that*

$$
\mathbb{E}[\lambda_t(x_1)\cdots\lambda_t(x_k)] = \mathbb{E}[\lambda_t^{\otimes k}(x)] = \mathbb{E}_{Y_0=x}\left[\exp\left(\int_0^t \kappa(Y_\tau)d\tau\right)\lambda_0(Y_{1,t})\cdots\lambda_0(Y_{k,t})\right],
$$

so that we also get a jump representation for $\mathbf{m}(t, t+x_1,\dots,t+x_k)$ (where we apply the notation of Section 3.1). If $\lambda_0 \equiv X_0$, we obtain again a representation that involves only X_0^k and thus again a special case of Theorem 3.8.

Let us also explicitly draw the connection to the dual process approach outlined in the introduction via (1). *The infinitesimal generator $\mathcal{A}$ of the process λ applied to cylindrical functions $f(\lambda) = g(\langle a_1, \lambda\rangle,\dots,\langle a_k, \lambda\rangle)$ where $g \in C^2(\mathbb{R}^k)$ and $a_1,\dots,a_k$ are continuous linear functionals is given by*

$$
\begin{aligned}
\mathcal{A}f(\lambda) = &\sum_{i=1}^k \partial_i g(\langle a_1, \lambda\rangle,\dots,\langle a_k, \lambda\rangle)\langle \partial_x \lambda + Kb_1\lambda(0), a_i\rangle \\
&+ \frac{1}{2}\sum_{i,j=1}^k \partial_{ij} g(\langle a_1, \lambda\rangle,\dots,\langle a_k, \lambda\rangle)\langle K\sqrt{A_{11}}\lambda(0), a_i\rangle\langle K\sqrt{A_{11}}\lambda(0), a_j\rangle.
\end{aligned}
$$

Letting $g(y) = \prod_{i=1}^k y_i$ and $\langle a_i, \lambda\rangle = \lambda(x_i)$, i.e. the point evaluations for some fixed $x = (x_1,\dots,x_k)$, we thus obtain

$$
\begin{aligned}
&\mathcal{A}\lambda(x_1)\cdots\lambda(x_k) \\
&= \sum_{i=1}^k \lambda(x_1)\cdots\lambda(x_{i-1})\lambda(x_{i+1})\cdots\lambda(x_k)(\partial_x\lambda(x_i) + K(x_i)b_1\lambda(0)) \\
&+ \frac{1}{2}\sum_{i,j=1}^k \lambda(x_1)\cdots\lambda(x_{i-1})\lambda(x_{i+1})\cdots\lambda(x_{j-1})\lambda(x_{j+1})\cdots\lambda(x_k)K(x_i)K(x_j)A_{11}\lambda^2(0).
\end{aligned}
$$

This is exactly the same as $\mathcal{L}$ in (1) with $\mathcal{L}\lambda(x_1)\cdots\lambda(x_k) = \mathcal{G}_k\lambda(x_1)\cdots\lambda(x_k) + \kappa(x)\lambda(x_1)\cdots\lambda(x_k)$ and the process U from the introduction corresponds to the process Y killed at rate κ.

Remark 4.6 *We illustrate now how the proposed method works without supposing the homogeneity condition of Assumption 32. Consider coefficients $b_0, b_1, A_0, A_1, A_{11}$ such that $b(x) = b_0 + b_1 x$ and $\sigma(x)^2 = A_0 + A_1 x + A_{11}x^2$. Fix $k \geq 1$ and let $(Y_t)_{t\in[0,T]}$ be the $([0,T]\cup\{\dagger\})^k$-valued process generated by the linear operator $\mathcal{G}_k$ with domain*

$$\overline{\mathcal{D}(\mathcal{G}_k)} = \{f : ([0,T]\cup\{\dagger\})^k \to \mathbb{R} \colon f \text{ is symmetric,}$$
$$f(\underbrace{\cdot,\ldots,\cdot}_{i},\underbrace{\dagger,\ldots,\dagger}_{k-i})|_{[0,T]^i} \in AC([0,T]^i,\mathbb{R}), \ i \in \{1,\ldots,k\}, \ f(\dagger,\ldots,\dagger) = 1\}$$

and given by

$$\mathcal{G}_k f(y) = \mathbf{1}^\top \nabla f(y) + \int (f(\xi) - f(y))\nu(y, d\xi),$$

for $\nabla_{x_i} f(\dagger) = 0$, $K(\dagger) = 0$,

$$\nu(y,\cdot) = b_0 \sum_{i=1}^{k} K(y_i)\delta_{(y_1,\ldots,y_{i-1},\dagger,y_{i+1},\ldots,y_k)} + b_1 \sum_{i=1}^{k} K(y_i)\delta_{(y_1,\ldots,y_{i-1},0,y_{i+1},\ldots,y_k)}$$

$$+ A_0 \sum_{i=1}^{k}\sum_{j<i} K(y_i)K(y_j)\delta_{(y_1,\ldots,y_{i-1},\dagger,y_{i+1},\ldots,y_{j-1},\dagger,y_{j+1},\ldots,y_k)}$$

$$+ \frac{1}{2}A_1 \sum_{i=1}^{k}\sum_{j\neq i} K(y_i)K(y_j)\delta_{(y_1,\ldots,y_{i-1},0,y_{i+1},\ldots,y_{j-1},\dagger,y_{j+1},\ldots,y_k)}$$

$$+ A_{11} \sum_{i=1}^{k}\sum_{j<i} K(y_i)K(y_j)\delta_{(y_1,\ldots,y_{i-1},0,y_{i+1},\ldots,y_{j-1},0,y_{j+1},\ldots,y_k)},$$

and $Y_0 = (0,\ldots,0)$. Suppose that Assumption 31 holds true and assume that the mild solution $(\lambda_t)_{t\in[0,T]}$ to (41) additionally satisfies

$$\mathbb{E}[\sup_{y\in[0,T]^k} |\lambda_t'(y_1)\lambda_t(y_2)\cdots\lambda_t(y_k)|] < \infty.$$

Then for each $t \in [0,T]$ it holds

$$\mathbb{E}[X_t^k] = \mathbb{E}\left[\exp\left(\int_0^t \kappa(Y_\tau)d\tau\right) X_0^{\sum_{i=1}^k \mathbf{1}_{\{Y_{i,t}\neq\dagger\}}}\right],$$

where

$$\kappa(y) = b(1)\sum_{i=1}^{k} K(y_i) + \sigma(1)^2 \sum_{i=1}^{k}\sum_{j<i} K(y_i)K(y_j).$$

Observe that there is a direct connection between this representation and Remark 3.4. Indeed, given a solution $(f_t)_{t\in[0,T]}$ of

$$\partial_t f_t(y) = \mathcal{G}_k f_t(y)$$

we obtain that $\overline{\mathbf{m}}_{\pi(p,w)}(t, x_1,\ldots,x_p) = f_t(x_1,\ldots,x_p,\dagger,\ldots,\dagger)$, $p = 0,\ldots,k$ solves (30). Moreover, noting that

$$\mathbb{E}[X_t^k] = \sum_{j=0}^{k} \mathbb{E}\left[\exp\left(\int_0^t \kappa(Y_\tau)d\tau\right) \Big| \sum_{i=1}^{k}\mathbf{1}_{\{Y_{i,t}\neq\dagger\}} = j\right]\mathbb{P}\left(\sum_{i=1}^{k}\mathbf{1}_{\{Y_{i,t}\neq\dagger\}} = j\right)X_0^j,$$

we can also establish a direct connection with Theorem 3.8 and in particular equation (38).

4.2 Multivariate case

We move now to the d-dimensional setting, letting again X be the process defined in (3). In this case we consider a solution $(\lambda_t)_{t\in[0,T]}$ of the SPDE given by

$$
\lambda_t(x,i) = X_{0,i} + \int_0^t \frac{d}{dx}\lambda_u(x,i)du + \sum_{j=1}^d e_i^\top K(x)b_j\lambda_u(0,j)du
$$
$$
+ \sum_{j_1,j_2=1}^d \sqrt{\lambda_u(0,j_1)\lambda_u(0,j_2)}e_i^\top K(x)\sqrt{A_{j_1 j_2}}dW_u,
$$

(45)

for $x \in \mathbb{R}_+$ and $i \in \{1,\ldots,d\}$ and where e_i denotes the canonical basis vectors. The Hilbert space for the corresponding function-valued process $(\lambda_t)_{t\in[0,T]}$ is given by

$$
B_d = \{f : \mathbb{R}_+ \times \{1,\ldots d\} \to \mathbb{R} : f(\cdot,i) \in B\},
$$

with

$$
\|y\|_{\alpha,d}^2 = \sum_{i=1}^d \left(|y(0,i)|^2 + \int_0^\infty |y'(x,i)|^2\alpha(x)dx \right).
$$

We first start by establishing the analog of Proposition 4.1. Even though we believe that existence of an analytically weak solution can be proved via the martingale problem approach we here just focus on the case $A_{j_1 j_2} = 0$ whenever $j_1 \neq j_2$ so that we only have linear terms in the diffusion part and can apply similar arguments as in Proposition 4.1.

For completeness we recall the notion of a mild solution in the current setting when $A_{j_1 j_2} = 0$ for $j_1 \neq j_2$. A predictable B_d-valued process $(\lambda_t)_{t\in[0,T]}$, is said to be a *mild solution* to (45) if for every $t \in [0,T]$

$$
\mathbb{P}\left[\int_0^t \|\lambda_s\|_{\alpha,d}^2 ds < \infty \right] = 1,
$$

and for every $i \in \{1,\ldots,d\}$

$$
\lambda_t(\cdot,i) = \underbrace{S_t\lambda_0(\cdot,i)}_{\equiv X_{0,i}} + \sum_{j=1}^d S_{t-u}e_i^\top Kb_j\lambda_u(0,j)du + \sum_{j=1}^d S_{t-u}e_i^\top K\sqrt{A_{jj}}\lambda_u(0,j)dW_u, \quad \mathbb{P}\text{-a.s.}
$$

(46)

Throughout this section we shall assume the following condition on K.

Assumption 4 *We assume that K satisfies $K_{ij} \in B$ for all $i,j \in \{1,\ldots,d\}$.*

Proposition 4.7 (Existence and moment bounds for solutions of (45)**)** *Let Assumption 4 hold true and suppose that $A_{j_1 j_2} = 0$ for $j_1 \neq j_2$. Then there exists a unique mild solution to (45) which has a continuous modification. Moreover, for every $k \geq 1$, we have*

$$
\mathbb{E}[\sup_{t\in[0,T]}\|\lambda_t\|_{\alpha,d}^k] \leq C_{T,k,X_0},
$$

(47)

where C_{T,k,X_0} is a constant that depends on T, k and X_0. This solution is also an (analytically) weak solution in the sense of [33, page 161].

We again apply Theorem [33, Theorem 7.2] for the Hilbert space $H = B_d$ and $U_0 = \mathbb{R}^d$. Verifying the conditions of [33, Hypotheses 7.1] is completely analogous to the one dimensional case and we obtain existence

and uniqueness of a mild solution to (46). The moments estimates and the assertion concerning the weak solution also follow analogously.

In the next proposition we establish the jump representation in the multivariate case. Here, we also allow for $A_{j_1 j_2} \neq 0$ when $j_1 \neq j_2$. Note that all the assumptions in Proposition 4.8 except of (48) are satisfied under the conditions of Proposition 4.7.

Proposition 4.8 (Jump representation of d-dimensional polynomial Volterra processes) *Fix $k \geq 1$ and let $(Y_t)_{t \in [0,T]}$ be the $([0,T] \times \{1, \ldots, d\})^k$-valued process generated by $\mathcal{G}_k$ with domain $\mathcal{D}(\mathcal{G}_k) = AC(([0,T] \times \{1, \ldots, d\})^k, \mathbb{R})$ given by*

$$\mathcal{G}_k f(x, i) = 1^\top \nabla_x f(x, i) + \int (f(\xi, \ell) - f(x, i)) \nu((x, i), d(\xi, \ell)),$$

for

$$\nu((x, i), \cdot) = \sum_{m=1}^k \sum_{j=1}^d e_{i_m}^\top K(x_m) b_j \delta_{\{((x_1, i_1), \ldots, (0, j), \ldots, (x_k, i_k))\}}$$

$$+ \sum_{m=1}^k \sum_{n<m} \sum_{j_1, j_2=1}^d e_{i_m}^\top K(x_m) A_{j_1 j_2} K(x_n)^\top e_{i_n} \delta_{\{((x_1, i_1), \ldots, (0, j_1), \ldots, (0, j_2), \ldots, (x_k, i_k))\}}$$

and

$$Y_0 = (\underbrace{(0, 1) \ldots, (0, 1)}_{k_1\text{-times}}, \underbrace{(0, 2), \ldots, (0, 2)}_{k_2\text{-times}}, \ldots, \underbrace{(0, d), \ldots, (0, d)}_{k_d\text{-times}}).$$

Suppose that Assumption 4 holds true and assume that $(\lambda_t)_{t \in [0,T]}$ is a B_d-valued (analytically) weak solution to (45) satisfying

$$\mathbb{E}[\sup_{t \in [0,T], i} |\lambda_t(x_1, i_1) \cdots \lambda_t(x_k, i_k)|] < \infty \quad \text{and}$$

$$\mathbb{E}[\sup_{x \in [0,T]^k, i} |(\partial_x \lambda_t(x_1, i_1)) \lambda_t(x_2, i_2) \cdots \lambda_t(x_k, i_k)|] < \infty. \tag{48}$$

Then for each $t \in [0,T]$ and each multi-index $\mathbf{k} = (k_1, \ldots k_d)$ with $|\mathbf{k}| = k$, it holds

$$\mathbb{E}[X_t^{\mathbf{k}}] = \mathbb{E}[X_{t,1}^{k_1} \cdots X_{t,d}^{k_d}] = X_0^{\mathbf{k}} \mathbb{E}\left[\exp\left(\int_0^t \kappa(Y_\tau) d\tau\right)\right],$$

where

$$\kappa(x, i) = \sum_{m=1}^k \sum_{j=1}^d e_{i_m}^\top K(x_m) b_j + \sum_{m=1}^k \sum_{n<m} \sum_{j_1, j_2=1}^d e_{i_m}^\top K(x_m) A_{j_1 j_2} K(x_n)^\top e_{i_n}.$$

Set

$$\mathcal{M}_k f(x, i) = \mathcal{G}_k f(x, i) + \kappa(x, i) f(x, i),$$

for each $f \in \mathcal{D}(\mathcal{G}_k)$ and observe that by Itô's formula it holds

$$\lambda_t^{\otimes k}(x, i) = X_{0,i_1} \cdots X_{0,i_k} + \int_0^t \mathcal{M}_k \lambda_u^{\otimes k}(x, i) du + \int_0^t \mathcal{Q} \lambda_u^{\otimes k}(x, i) dW_u.$$

for

$$\mathcal{Q} f(x, i) = \sum_{m=1}^k \sum_{j_1, j_2=1}^d e_{i_m}^\top K(x_m) \sqrt{A_{j_1 j_2}}$$

$$\times \sqrt{f((x_1, i_1), \ldots, (0, j_1), \ldots, (x_d, i_d)) f((x_1, i_1), \ldots, (0, j_2), \ldots, (x_d, i_d))}.$$

Since the third term is a true martingale, we have

$$\mathbb{E}[\lambda_t^{\otimes k}(x,i)] = X_{0,i_1} \cdots X_{0,i_k} + \int_0^t \mathbb{E}[\mathcal{M}_k \lambda_u^{\otimes k}(x,i)] du.$$

Set then $m_t(x,i) = \mathbb{E}[\lambda_t^{\otimes k}(x,i)]$. Since (48) holds, Leibniz rule and Fubini yield

$$\mathcal{M}_k m_t(x,i) = 1^\top \nabla_x \mathbb{E}[\lambda_t^{\otimes k}(\cdot,i)](x) + \int \mathbb{E}[\lambda_t^{\otimes k}(\xi,\ell)] \nu((x,i), d(\xi,\ell))$$

$$= \mathbb{E}[1^\top \nabla_x \lambda_t^{\otimes k}(x,i)] + \mathbb{E}\left[\int \lambda_t^{\otimes k}(\xi,\ell) \nu((x,i), d(\xi,\ell))\right],$$

proving that $\mathcal{M}_k m_t(x,i) = \mathbb{E}[\mathcal{M}_k \lambda_t^{\otimes k}(x,i)]$ and thus that

$$m_t(x,i) = X_{0,i_1} \cdots X_{0,i_k} + \int_0^t \mathcal{M}_k m_u(x,i) du.$$

Since m_s lies in the domain of $\mathcal{G}_k$ for each $s \in [0,T]$ proceeding as in the one dimensional case we get that the process $(Z_s)_{s \in [0,t]}$ for

$$Z_s = \exp\left(\int_0^s \kappa(Y_\tau) d\tau\right) m_{t-s}(Y_s)$$

is a true martingale and thus that

$$X_0^{\mathbf{k}} \mathbb{E}\left[\exp\left(\int_0^t \kappa(Y_\tau) d\tau\right)\right] = m_t(Y_0) = \mathbb{E}[\lambda_t^{\otimes k}(Y_0)] = \mathbb{E}[X_t^{\mathbf{k}}].$$

Example 4.9 *Observe that for $K \equiv \mathrm{I}_d$, $\lambda_0 \equiv (1,\ldots,1)$, $b_i = \beta_i e_i$, $A_{ij} = \sigma_i^\top \sigma_j e_i e_j^\top$ for some $\beta_i \in \mathbb{R}$, $\sigma_i \in \mathbb{R}^d$ we get that*

$$X_i = \mathcal{E}((\beta_i t + \sum_{j=1}^d \sigma_{ij} W_t^j)_{t \in [0,T]}).$$

In this setting

$$\nu((x,i),\cdot) = \sum_{m=1}^k \sum_{j=1}^d e_{i_m}^\top \beta_j e_j \delta_{\{\ldots,(0,j),\ldots\}} + \sum_{m=1}^k \sum_{n<m} \sum_{j_1,j_2=1}^d \sigma_{j_1}^\top \sigma_{j_2} e_{i_m}^\top e_{j_1} e_{j_2}^\top e_{i_n} \delta_{\{\ldots,(0,j_1),\ldots,(0,j_2),\ldots\}}$$

$$= \sum_{m=1}^k \beta_{i_m} \delta_{\{\ldots,(0,i_m),\ldots\}} + \sum_{m=1}^k \sum_{n<m} \sigma_{i_m}^\top \sigma_{i_n} \delta_{\{\ldots,(0,i_m),\ldots,(0,i_n),\ldots\}},$$

showing that the second component of each Y_i is constant over time. Since

$$\kappa(x,i) = \sum_{m=1}^k \beta_{i_m} + \sum_{m=1}^k \sum_{n<m} \sigma_{i_m}^\top \sigma_{i_n} = \sum_{m=1}^k \beta_{i_m} + \frac{1}{2} \sum_{m,n=1}^k \sigma_{i_m}^\top \sigma_{i_n} - \sum_{m=1}^k \sigma_{i_m}^\top \sigma_{i_m},$$

the corresponding representation yields

$$\mathbb{E}[X_t^{\mathbf{k}}|\lambda_0] = \exp\left(\left(\sum_{j=1}^d k_j \beta_j + \frac{1}{2} \sum_{j_1,j_2=1}^d k_{j_1} k_{j_2} \sigma_{j_1}^\top \sigma_{j_2} - \sum_{j=1}^d k_j \sigma_j^\top \sigma_j\right) t\right),$$

as expected.

Remark 4.10 *Note that in analogy to Remark 4.6 also in the multivariate setting the non-homogeneous case can be treated. This can be done by adding zero indices in the definition of κ and ν, i.e. including b_0, and $A_{0i} = A_{i0} = A_i$ for $i = 0, 1, \ldots, d$, and changing the corresponding jumps to $(\dagger, 0)$ (instead of $(0, j)$ as it is the case for $j = 1, \ldots, d$).*

5 Volterra processes in the unit ball

In this section, we are interested in constructing general Volterra processes (3), that remain in the unit ball $\mathcal{B}$ of $\mathbb{R}^d$. To achieve this, structural conditions on the coefficients $b : \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ and the kernel $K \in L^2_{\text{loc}}$ are needed. In a first step, we consider more general continuous coefficients with linear growth

$$|b(x)| \vee |\sigma(x)| \leq c(1 + |x|), \quad x \in \mathbb{R}^d. \tag{49}$$

It follows from [6, Theorem 3.4] that under Assumption 1 and (49), there exists a weak solution to (3) with values in $\mathbb{R}^d$. For the rest of this section, we will assume that the kernel K is scalar $K : \mathbb{R}_+ \to \mathbb{R}$.

The next theorem provides the existence of a $\mathcal{B}$-valued solution to the stochastic Volterra equation (3) under structural conditions on the coefficients (b, σ) and the kernel K. We denote by $\partial \mathcal{B}$ the boundary of $\mathcal{B}$, that is

$$\partial \mathcal{B} = \{x \in \mathbb{R}^d : x^\top x = 1\}.$$

Theorem 5.1 (Existence of Volterra processes in the unit ball) *Fix a scalar kernel $K : [0, T] \to \mathbb{R}$ that satisfies Assumptions 1 and 2. Assume that b and σ are continuous, with linear growth* (49) *and*

$$x^\top b(x) \leq 0 \text{ and } \sigma(x) = 0, \quad x \in \partial \mathcal{B}. \tag{50}$$

For any $X_0 \in \mathcal{B}$, there exists a continuous weak solution X to (3) *such that $X_t \in \mathcal{B}$ a.s. for every $t \in [0, T]$.*

The proof is given in Section 5.1 below.

Remark 5.2 *Theorem 5.1 allows the construction of Volterra processes living in the more general domain $\{x \in \mathbb{R}^d : x^\top Q x \leq 1\}$, for some positive definite $d \times d$ matrix Q. Indeed, it suffices to construct a weak solution X for the unit ball and set $Y = Q^{-1/2}X$.*

5.1 Proof of Theorem 5.1

[Proof of Theorem 5.1]For any $n \in \mathbb{N}$ consider the coefficients b^n and σ^n by

$$b^n(x) = \begin{cases} b(\frac{x}{1-1/n}) - \frac{x}{n(1-1/n)}, & |x| \leq 1 - 1/n \\ b(\frac{x}{|x|}) - \frac{x}{n|x|}, & else \end{cases}$$

$$\sigma^n(x) = \begin{cases} \sigma\left(\frac{x}{1-1/n}\right), & |x| \leq 1 - 1/n \\ 0, & else. \end{cases}$$

By the assumptions for b and σ, it follows that b^n and σ^n are continuous, and they satisfy the linear growth condition (49) uniformly in n, with common constant $\widehat{C}$. Denote by X^n a continuous weak solution to (3) with respect to (b^n, σ^n), obtained from [6, Theorem 3.4]. Applying tightness and stability results for stochastic Volterra equations, see for instance [6, Lemma A.1 and A.2], we have $X^n \implies \hat{X}$, where $\hat{X}$ solves (3) with coefficients $\hat{b}(x) = 1_{\mathcal{B}}(x)b(x) + 1_{\mathcal{B}^c}(x)b(x/|x|)$ and $\hat{\sigma}(x) = 1_{\mathcal{B}}(x)\sigma(x)$. If we can prove that $X^n \in \mathcal{B}$ almost surely for all $n \in \mathbb{N}$, then $\hat{X} \in \mathcal{B}$ almost surely and since $\hat{b}|_{\mathcal{B}} = b$ and $\hat{\sigma}|_{\mathcal{B}} = \sigma$, the claim follows.

Fix $n \in \mathbb{N}$ and denote $X = X^n$. Note that the coefficients (b^n, σ^n) satisfy the stronger conditions

$$\sigma^n(x) = 0 \quad \text{and} \quad x^\top b^n(y) \leq 0, \quad |y|, |x| > 1 - 1/n \text{ and } |x - y| < \epsilon, \tag{51}$$

for some ϵ depending on n and the constant from the linear growth condition. Indeed let y such that $|y| > 1 - 1/n$ and denote by $\hat{y} = y/|y| \in \partial \mathcal{B}$, then

$$
\begin{aligned}
x^\top b^n(y) = y^\top b^n(y) + (x-y)^\top b^n(y) &= |y|(\hat{y}^\top b(\hat{y}) - (1/n)\hat{y}^\top \hat{y}) + (x-y)^\top b^n(y) \\
&\leq |y|(\hat{y}^\top b(\hat{y}) - (1/n)\hat{y}^\top \hat{y}) + |x-y|\,(|b(\hat{y})| + 1/n) \\
&\leq |y|(\hat{y}^\top b(\hat{y}) - (1/n)\hat{y}^\top \hat{y}) + |x-y|(2C+1) \\
&\leq -\frac{(n-1)}{n^2} + (2C+1)|x-y| \\
&\leq 0,
\end{aligned}
$$

whenever $|x-y| \leq \frac{n-1}{(2C+1)n^2} = \epsilon$, where we used Cauchy-Schwarz for the second inequality with C denoting the linear growth constant of b, and the assumption (50) for the third inequality.

Now consider the stopping time $\tau = \inf\{t \geq 0 : X_t \notin \mathcal{B}\}$, we aim to prove that $\mathbb{P}(\tau < \infty) = 0$. Applying similar reasoning to [6, Theorem 3.6], under Assumption 2, for any stopping time τ and $h > 0$ we have

$$
\begin{aligned}
X_{\tau+h} = {}& (1 - (\Delta_h K * L)(\tau))\, X_0 + (\Delta_h K * L)(0)X_\tau + (d(\Delta_h K * L) * (X))(\tau) \\
& + \int_0^\infty 1_{(\tau,\tau+h]}(s)K(\tau+h-s)b^n(X_s)ds \\
& + \left(\int_0^\infty 1_{(\tau,\tau+h]}(s)K(t+h-s)\sigma^n(X_s)dW_s \right)\bigg|_{t=\tau} \\
=: {}& a_0(h) + a_1(h) + a_2(h).
\end{aligned}
\tag{52}
$$

Now for $\epsilon' > 0$, we define the events

$$
\Omega^{\epsilon'} = \left\{ \tau < \infty, X_s^\top b^n(X_u) \leq 0 \text{ and } \sigma^n(X_u) = 0 \quad \forall u, s \in [\tau, \tau + \epsilon'[\right\}.
$$

On the event $\Omega^{\epsilon'}$, for all $h \in (0, \epsilon')$, we have $a_2(h) = 0$, and for $\psi(t) = |X_{\tau+t}|$

$$
\psi(h)^2 = X_{\tau+h}^\top \left(a_0(h) + a_1(h) \right) \leq \psi(h)|a_0(h)| + X_{\tau+h}^\top a_1(h),
\tag{53}
$$

where we simply applied Cauchy-Schwarz for the inequality. An application of Lemma 5.4 below yields that $|a_0(h)| \leq 1$. Moreover, by the definition of $\Omega^{\epsilon'}$ and the fact that K is nonnegative, it follows that

$$
\int_\tau^{\tau+h} K(\tau+h-s)X_{\tau+h}^\top b^n(X_s)ds \leq 0.
$$

Therefore we can conclude

$$
\psi(h)^2 \leq \psi(h)|a_0(h)| + X_{\tau+h}^\top a_1(h) \leq \psi(h),
$$

which readily yields $\psi(h) \leq 1$, and thus $X_{\tau+h} \in \mathcal{B}$ for all $h \in (0, \epsilon')$ on $\Omega^{\epsilon'}$. On the other hand, by the definition of τ and continuity of X, there exists $h \in (0, \epsilon')$ such that $X_{\tau+h} \notin \mathcal{B}$ on $\Omega^{\epsilon'}$. But this readily shows $\mathbb{P}(\Omega^{\epsilon'}) = 0$ for all ϵ'. On the other hand, by (51) and continuity of X, it follows that

$$
\mathbb{P}(\tau < \infty) = \mathbb{P}\left(\bigcup_{\epsilon' \in (0,\epsilon) \cap \mathbb{Q}} \Omega^{\epsilon'} \right) = 0.
$$

The two following lemmas where used in the proof.

Lemma 5.3 *Let K satisfy* (2). *Then,*

$$(\Delta_h K * L)(t) \text{ is nondecreasing in } t, \tag{54}$$

as well as

$$0 \leq (\Delta_h K * L)(t) \leq (K * L)(t) = 1. \tag{55}$$

The proof of (54) and (55) can be found in the proof of [6, Theorem 3.6], see (3.9) and (3.10) therein.

Lemma 5.4 *Let K satisfy* (2) *and $f : [0,t] \to \mathcal{B}$ be continuous such that $f(t) \in \partial\mathcal{B}$ for some t. Then, the quantity*

$$a_0(h) = (1 - (\Delta_h K * L)(t)) f(0) + (\Delta_h K * L)(0)f(t) + (d(\Delta_h K * L) * f)(t), \quad h \geq 0,$$

satisfies

$$a_0(h)^\top a_0(h) \leq 1, \quad h \geq 0.$$

We first notice that, thanks to the Cauchy-Schwarz inequality, we have

$$x^\top y \leq \sqrt{x^\top x}\sqrt{y^\top y} \leq 1, \quad x, y \in \mathcal{B}. \tag{56}$$

We expand

$$\begin{aligned}
a_0(h)^\top a_0(h) &= (1 - (\Delta_h K * L)(t))^2 f(0)^\top f(0) + (\Delta_h K * L)(0)^2 f(t)^\top f(t) + \\
&\quad + (d(\Delta_h K * L) * f)(t)^\top (d(\Delta_h K * L) * f)(t) \\
&\quad + 2(1 - (\Delta_h K * L)(t))(\Delta_h K * L)(0)f(0)^\top f(t) \\
&\quad + 2(1 - (\Delta_h K * L)(t)) f(0)^\top (d(\Delta_h K * L) * f)(t) \\
&\quad + 2(\Delta_h K * L)(0)f(t)^\top (d(\Delta_h K * L) * f)(t) \\
&\leq (1 - (\Delta_h K * L)(t))^2 + (\Delta_h K * L)(0)^2 + \\
&\quad + (d(\Delta_h K * L) * 1)^2(t) \\
&\quad + 2(1 - (\Delta_h K * L)(t))(\Delta_h K * L)(0) \\
&\quad + 2(1 - (\Delta_h K * L)(t))(d(\Delta_h K * L) * 1)(t) \\
&\quad + 2(\Delta_h K * L)(0)(d(\Delta_h K * L) * 1)(t) \\
&= ((1 - (\Delta_h K * L)(t)) + (\Delta_h K * L)(0) + (d(\Delta_h K * L) * 1)(t))^2
\end{aligned}$$

where for the second inequality, we used (56) combined with (54) and (55). Finally, observing that

$$(d(\Delta_h K * L) * 1)(t) = (\Delta_h K * L)(t) - (\Delta_h K * L)(0)$$

yields the desired claim.

A Well-posedness for Volterra-type integral equations

In this section we study existence and uniqueness of solutions to a certain class of integral equations, including all the equations appearing in Section 3 for the characterization of moments. Fix $N, D \in \mathbb{N}$ and $T \geq 0$, and consider the domain $\mathcal{D}_T^{(N)}$ as in (24). Denote by $C(\mathcal{D}_T^{(N)}, \mathbb{R}^D)$ the space of continuous, vector-valued functions on the compact set $\mathcal{D}_T^{(N)}$. Moreover, consider the families of kernels $\{\mathcal{A}_1^{n,i}\}_{1 \leq n \leq N; 1 \leq i \leq 2}$, and $\{\mathcal{A}_2^{n,m,j}\}_{1 \leq n,m \leq N; 1 \leq j \leq 3}$, such that

$$\mathcal{A}_1^{n,i} : [0,T] \to \mathbb{R}^{D \times D}, \mathcal{A}_2^{n,m,j} : [0,T] \times [0,T] \to \mathbb{R}^{D \times D}, 1 \leq n, m \leq N, 1 \leq i \leq 2, 1 \leq j \leq 3.$$

We will make the following integrability assumptions for the matrix-norm of the kernels:

$$|\mathcal{A}_1^{n,i}| \in L^1([0,T]), \quad |\mathcal{A}_2^{n,m,j}(t,s)| \leq |\mathcal{G}_1^{n,m}(t)| \cdot |\mathcal{G}_2^{n,m}(s)|, \tag{57}$$

for some $\mathcal{G}_1^{n,m}, \mathcal{G}_2^{n,m} \in L^2([0,T])$, for all $1 \leq n, m \leq N$, $1 \leq i \leq 2$, $1 \leq j \leq 3$. Finally, for all $\mathbf{f} \in C(\mathcal{D}_T^{(N)}, \mathbb{R}^D)$ we define

$$
\begin{aligned}
(\Psi\mathbf{f})(t, T_1, \dots, T_N) &= \sum_{1 \leq n \leq N} \int_0^t \mathcal{A}_1^{n,1}(T_n - r)\mathbf{f}(r, r, (T_k)_{k \neq n})dr \\
&\quad \sum_{1 \leq n \leq N} \int_0^t \mathcal{A}_1^{n,2}(T_n - r)\mathbf{f}(r, (T_k)_{k \neq n}, r)dr \\
&\quad + \sum_{1 \leq n < m \leq N} \int_0^t \mathcal{A}_2^{n,m,1}(T_n - r, T_m - r)\mathbf{f}(r, r, r, (T_k)_{k \neq n,m})dr \\
&\quad + \sum_{1 \leq n < m \leq N} \int_0^t \mathcal{A}_2^{n,m,2}(T_n - r, T_m - r)\mathbf{f}(r, r, (T_k)_{k \neq n,m}, r)dr \\
&\quad + \sum_{1 \leq n < m \leq N} \int_0^t \mathcal{A}_2^{n,m,3}(T_n - r, T_m - r)\mathbf{f}(r, (T_k)_{k \neq n,m}, r, r)dr.
\end{aligned}
\tag{58}
$$

Proposition A.1 *Suppose that* (57) *holds. Then, for any function* $\mathbf{f}_0 \in C([0,T]^N, \mathbb{R}^D)$ *there exists a unique solution* $\mathbf{f} \in C(\mathcal{D}_T^{(N)}, \mathbb{R}^D)$ *to the integral equation*

$$\mathbf{f}(t, T_1, \dots, T_N) = \mathbf{f}_0(T_1, \dots, T_N) + (\Psi\mathbf{f})(t, T_1, \dots, T_N), \quad (t, T_1, \dots, T_N) \in \mathcal{D}_T^{(N)}. \tag{59}$$

For $\lambda \geq 0$ we introduce the norm $\|\mathbf{f}\|_{\lambda,\infty} = \sup_{(t,T_1,\dots,T_N) \in \mathcal{D}_T^{(N)}} e^{-\lambda t}|\mathbf{f}(t, T_1, \dots, T_N)|$, where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^D$. One can check that $\|\cdot\|_{\lambda,\infty}$ is equivalent to the sup-norm, and $(C(\mathcal{D}_T^{(N)}, \mathbb{R}^D), \|\cdot\|_{\lambda,\infty})$ is a Banach space. Following a standard proof pattern, we wish to prove that the operator

$$T\mathbf{f} = \mathbf{f}_0(\cdot) + (\Psi\mathbf{f})(\cdot), \quad \mathbf{f} \in C(\mathcal{D}_T^{(N)}, \mathbb{R}^D),$$

is invariant and contracts. For the invariance, we only need to show that $(t, T_1, \dots, T_N) \mapsto (T\mathbf{f})(t, T_1, \dots, T_N)$ is continuous. By assumption, $\mathbf{f}_0$ is continuous, and for $x = (t, T_1, \dots, T_N)$ and $x' = (t', T_1', \dots, T_N')$, with $t \leq t'$, we have

$$
\begin{aligned}
\left|(\Psi\mathbf{f})(t, T_1, \dots, T_N) - (\Psi\mathbf{f})(t', T_1', \dots, T_N')\right| &\leq \left|(\Psi\mathbf{f})(t, T_1, \dots, T_N) - (\Psi\mathbf{f})(t, T_1', \dots, T_N')\right| \\
&\quad + \left|(\Psi\mathbf{f})(t', T_1', \dots, T_N') - (\Psi\mathbf{f})(t, T_1', \dots, T_N')\right|.
\end{aligned}
$$

The second term converges to zero as $t \to t'$ by dominated convergence and (57). For the first term we only analyze the first and third components of $\Psi\mathbf{f}$ as the others behave similarly. Notice that for any $1 \leq n, m \leq N$

$$
\begin{aligned}
&\int_0^t |\mathcal{A}_1^{n,1}(T_n - r)\mathbf{f}(r, r, (T_k)_{k \neq n}) - \mathcal{A}_1^{n,1}(T_n' - r)\mathbf{f}(r, r, (T_k')_{k \neq n})|dr \\
&\quad \leq \int_0^t |\mathcal{A}_1^n(T_n - r)| \cdot |\mathbf{f}(r, r, (T_k)_{k \neq n}) - \mathbf{f}(r, r, (T_k')_{k \neq n})|dr \\
&\quad + \int_0^t |\mathcal{A}_1^n(T_n - r) - \mathcal{A}_1^n(T_n' - r)| \cdot |\mathbf{f}(r, r, (T_k')_{k \neq n})|dr.
\end{aligned}
$$

and

$$\int_0^t |\mathcal{A}_2^{n,m,1}(T_n - r, T_m - r)\mathbf{f}(r, r, r, (T_k)_{k\neq n,m}) - \mathcal{A}_2^{n,m,1}(T_n' - r, T_m' - r)\mathbf{f}(r, r, r, (T_k')_{k\neq n,m})|dr$$

$$\leq \int_0^t |\mathcal{A}_2^{n,m,1}(T_n - r, T_m - r)| \cdot |\mathbf{f}(r, r, r, (T_k)_{k\neq n,m}) - \mathbf{f}(r, r, r, (T_k')_{k\neq n,m})|dr$$

$$+ \int_0^t |\mathcal{A}_2^{n,m,1}(T_n - r, T_m - r) - \mathcal{A}_2^{n,m}(T_n' - r, T_m' - r)| \cdot |\mathbf{f}(r, r, r, (T_k')_{k\neq n,m})|dr.$$

In the expressions on the right of the two previous inequalities, the first integrals converge to zero as $x \to x'$ by (57), dominated convergence and the fact that $\mathbf{f}$ is continuous. For the second integrals, we can take $\|\mathbf{f}\|_\infty$ out of the integral, and then use (57) and the fact that the translation of L^p-functions is continuous, see [54, Proposition 1.6.13], to show that both terms converge to zero as $x' \to x$. Applying similar arguments to the other components of $\Psi\mathbf{f}$, we conclude that $T\mathbf{f} \in C(\mathcal{D}_T^{(N)}, \mathbb{R}^D)$.

Finally, denoting by d_λ the metric induced by the norm $\|\cdot\|_{\lambda,\infty}$, we want to show that T contracts for λ large enough, that is

$$d_\lambda(T\mathbf{f}, T\mathbf{g}) \leq q d_\lambda(\mathbf{f}, \mathbf{g}), \quad \forall \mathbf{f}, \mathbf{g} \in C(\mathcal{D}_T^{(N)}, \mathbb{R}^D) \tag{60}$$

for some $q < 1$ and $\lambda > 0$. By definition

$$e^{-\lambda t} |(T\mathbf{f})(t, T_1, \ldots, T_N) - (T\mathbf{g})(t, T_1, \ldots, T_N)|$$

$$\leq \sum_{1 \leq n \leq N} \int_0^t e^{-\lambda t} |\mathcal{A}_1^{n,1}(T_n - r)||\mathbf{f}(r, r, (T_k)_{k\neq n}) - \mathbf{g}(r, r, (T_k)_{k\neq n})|dr$$

$$+ \sum_{1 \leq n \leq N} \int_0^t e^{-\lambda t} |\mathcal{A}_1^{n,2}(T_n - r)||\mathbf{f}(r, (T_k)_{k\neq n}, r) - \mathbf{g}(r, (T_k)_{k\neq n}, r)|dr$$

$$+ \sum_{1 \leq n < m \leq N} \int_0^t e^{-\lambda t} |\mathcal{A}_2^{n,m,1}(T_n - r, T_m - r)||\mathbf{f}(r, r, r, (T_k)_{k\neq n,m}) - \mathbf{g}(r, r, r, (T_k)_{k\neq n,m})|dr$$

$$+ \sum_{1 \leq n < m \leq N} \int_0^t e^{-\lambda t} |\mathcal{A}_2^{n,m,2}(T_n - r, T_m - r)||\mathbf{f}(r, r, (T_k)_{k\neq n,m}, r) - \mathbf{g}(r, r, (T_k)_{k\neq n,m}, r)|dr$$

$$+ \sum_{1 \leq n < m \leq N} \int_0^t e^{-\lambda t} |\mathcal{A}_2^{n,m,3}(T_n - r, T_m - r)||\mathbf{f}(r, (T_k)_{k\neq n,m}, r, r) - \mathbf{g}(r, (T_k)_{k\neq n,m}, r, r)|dr$$

$$\leq d_\lambda(\mathbf{f}, \mathbf{g}) \times$$

$$\underbrace{\left(\sum_{\substack{1 \leq n \leq N \\ i=1,2}} \int_0^t e^{-\lambda(t-r)} |\mathcal{A}_1^{n,i}(T_n - r)|dr + \sum_{\substack{1 \leq n < m \leq N \\ j=1,2,3}} \int_0^t e^{-\lambda(t-r)} |\mathcal{A}_2^{n,m,j}(T_n - r, T_m - r)|dr \right)}_{C_\lambda(t,T_1,\ldots,T_N)}.$$

$$\tag{61}$$

We claim that for all $1 \leq n, m \leq N$, $1 \leq i \leq 2$, and $1 \leq j \leq 3$, we have

$$\sup_{0 \leq t \leq T_1 \leq T} \int_0^t e^{-\lambda(t-r)} |\mathcal{A}_1^{n,i}(T_1 - r)|dr, \quad \sup_{0 \leq t \leq T_1, T_2 \leq T} \int_0^t e^{-\lambda(t-r)} |\mathcal{A}_2^{n,m,j}(T_1 - r, T_2 - r)|dr \xrightarrow{\lambda \to \infty} 0. \tag{62}$$

Indeed, using a change of variable $s = t - r$, for any $\delta > 0$ we can write the first integral as

$$\int_0^t e^{-\lambda s} |\mathcal{A}_1^{n,i}(T_1 - t + s)|ds \leq \int_0^\delta e^{-\lambda s} |\mathcal{A}_1^{n,i}(T_1 - t + s)|ds + \int_\delta^{t\vee\delta} e^{-\lambda s} |\mathcal{A}_1^{n,i}(T_1 - t + s)|ds$$

$$\leq \|\mathcal{A}_1^{n,i}\|_{L^1([T_1-t, T_1-t+\delta])} + e^{-\lambda\delta} \|\mathcal{A}_1^{n,i}\|_{L^1([0,T])}.$$

Since $\mathcal{A}_1^{n,i} \in L^1([0,T])$ by assumption (57), for any $\epsilon > 0$, we can choose δ small enough, such that $\|\mathcal{A}_1^{n,i}\|_{L^1([T_1-t,T_1-t+\delta])} \leq \epsilon/2$, uniformly in (t,T_1). Choosing $\lambda > 0$ such that $e^{-\lambda\delta}\|\mathcal{A}_1^{n,i}\|_{L^1([0,T])} \leq \epsilon/2$, we obtain

$$\sup_{0 \leq t \leq T_1 \leq T} \int_0^t e^{-\lambda s} |\mathcal{A}_1^{n,i}(T_1 - t + s)| ds \leq \epsilon.$$

Similar considerations can be applied for the second term using the assumptions in (57). Indeed, for some $\mathcal{G}_1^{n,m}, \mathcal{G}_2^{n,m} \in L^2([0,T])$ we have

$$\int_0^t e^{-\lambda s} |\mathcal{A}_2^{n,m,j}(T_1 - t + s, T_2 - t + s)| ds \leq \max_{i \in \{1,2\}} \|\mathcal{G}_i^{n,m}\|_{L^2([T_i - t, T_i - t + \delta])}^2 + e^{-\lambda\delta} \max_{i \in \{1,2\}} \|\mathcal{G}_i^{n,m}\|_{L^2([0,T])}^2,$$

and we conclude with the same arguments as before, that we can find λ large enough such that

$$\sup_{0 \leq t \leq T_1, T_2 \leq T} \int_0^t e^{-\lambda(t-r)} |\mathcal{A}_2^{n,m,j}(T_1 - r, T_2 - r)| dr \leq \epsilon.$$

Hence, (62) holds and we can choose λ large enough, such that

$$\sup_{(t,T_1,\ldots,T_N) \in \mathcal{D}_T^{(N)}} C_\lambda(t, T_1, \ldots, T_N) < 1,$$

with C_λ as in (61). This implies (60) for some $q < 1$ and $\lambda > 0$. Banach fixed point theorem yields the conclusion.

Corollary A.2 *The function $\mathbf{m}$ defined in Theorem 3.3 is the unique solution in $\mathcal{X}_T^{(N)}$ of (26) with initial condition $\mathbf{m}(0, T_1, \ldots, T_N)$, and $\mathbf{m}$ is continuous on $\mathcal{D}_T^{(N)}$.*

We start by observing that $\mathcal{M}_T^{(N)}$ can be defined for bounded functions f on $\mathcal{D}_T^{(N)}$ by inserting r (one or multiple times) as the last arguments of the function $\mathbf{f}$ in the first, third and fourth addends in (25). Furthermore, this extension of the operator $\mathcal{M}_T^{(N)}$ restricted to the subspace $C(\mathcal{D}_T^{(N)}, \mathbb{R}^D)$ has the structure of the operator Ψ in (58) with $D = D_N$ (the cardinality of the set $\mathcal{I}^{(N)}$ in (23)), and the hypothesis (57) holds. In addition, since g_0 is continuous then $\mathbf{m}(0, T_1, \ldots, T_N)$ is continuous as well. By Proposition A.1 there exists a solution $\mathbf{m}' \in C(\mathcal{D}_T^{(N)}, \mathbb{R}^D)$ to the equation (26) with initial condition $\mathbf{m}(0, T_1, \ldots, T_N)$, where instead of $\mathcal{M}_T^{(N)}$ we consider the above mentioned extension of the operator. Furthermore, the proof of (61) in Proposition A.1 also holds for bounded functions $\mathbf{f}, \mathbf{g}$. Hence, uniqueness also holds over the space of bounded functions on $\mathcal{D}_T^{(N)}$. These observations imply uniqueness of the solutions in $\mathcal{X}_T^{(N)}$ of (26) with initial condition $\mathbf{m}(0, T_1, \ldots, T_N)$, and that the function $\mathbf{m}$ defined in Theorem 3.3 has to coincide with $\mathbf{m}'$ and in particular it is continuous.

Corollary A.3 *Fix $\beta \in \mathbb{N}_0^d$ and $p \in \mathbb{N}$ such that $|\beta| \leq p$. Suppose that that for each $1 \leq q < p$ and any $w \in \{1, \ldots, d\}^q$, $C_\beta^{(q)}(\cdot, \ldots, \cdot; w)$ is continuous on $\mathcal{D}_T^{(q)}$. Consider the system of equations (39) in Theorem 3.8 – seen as a system indexed over the elements $w \in \{1, \ldots, d\}^p$ – with initial condition $\mathbf{f}_\beta^{(p)}(\cdots; w)$, $w \in \{1, \ldots, d\}^p$, as in (40). Then for each $w \in \{1, \ldots, d\}^p$, $\mathbf{f}_\beta^{(p)}(\cdots; w)$ is in $C(\mathcal{D}_T^{(p)}, \mathbb{R})$ and (39) has a unique solution such that $C_\beta^{(p)}(\cdot, \ldots, \cdot; w) \in C(\mathcal{D}_T^{(p)}, \mathbb{R})$.*

The continuity of $\mathbf{f}_\beta^{(p)}(\cdots; w)$ can be shown in the same way that we showed in the proof of Proposition A.1 that $T\mathbf{f}$ was continuous for $\mathbf{f}$ continuous. By taking $D = d^p$, i.e. the cardinality of $\{1, \ldots, d\}^p$, and $N = p$, equation (39) has the structure of the equation in Proposition A.1 and hence the claim follows.

References

[1] Eduardo Abi Jaber. The Laplace transform of the integrated Volterra Wishart process. *Mathematical Finance*, 32(1):309–348, 2022.

[2] Eduardo Abi Jaber, Christa Cuchiero, Martin Larsson, and Sergio Pulido. A weak solution theory for stochastic Volterra equations of convolution type. *The Annals of Applied Probability*, 31(6):2924–2952, 2021.

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

[4] Eduardo Abi Jaber and Omar El Euch. Multifactor approximation of rough volatility models. *SIAM journal on financial mathematics*, 10(2):309–349, 2019.

[5] Eduardo Abi Jaber, Camille Illand, and Shaun Li. Joint SPX-VIX calibration with Gaussian polynomial volatility models: deep pricing with quantization hints. *arXiv preprint arXiv:2212.08297*, 2022.

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

[7] Eduardo Abi Jaber and Shaun Li. Volatility models in practice: Rough, path-dependent or markovian? *Path-Dependent or Markovian*, 2024.

[8] Damien Ackerer, Damir Filipović, and Sergio Pulido. The Jacobi stochastic volatility model. *Finance and Stochastics*, 22:667–700, 2018.

[9] Abdelkoddousse Ahdida and Aurélien Alfonsi. A mean-reverting SDE on correlation matrices. *Stochastic Processes and their Applications*, 123(4):1472–1520, 2013.

[10] Aurélien Alfonsi. Nonnegativity preserving convolution kernels. Application to Stochastic Volterra Equations in closed convex domains and their approximation. *arXiv preprint arXiv:2302.07758*, 2023.

[11] Ole E. Barndorff-Nielsen, Fred Espen Benth, and Almut E. D. Veraart. Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. *Bernoulli*, 19(3):803 – 845, 2013.

[12] Ole E. Barndorff-Nielsen, Fred Espen Benth, and Almut E. D. Veraart. Modelling electricity futures by ambit fields. *Advances in Applied Probability*, 46(3):719 – 745, 2014.

[13] Ole E Barndorff-Nielsen, Fred Espen Benth, and Almut ED Veraart. Ambit processes and stochastic partial differential equations. *Advanced mathematical methods for finance*, pages 35–74, 2011.

[14] Ole E Barndorff-Nielsen and Jürgen Schmiegel. Ambit processes; with applications to turbulence and tumour growth. In *Stochastic Analysis and Applications: The Abel Symposium 2005*, pages 93–124. Springer, 2007.

[15] Ole E Barndorff-Nielsen and Jürgen Schmiegel. Time change, volatility, and turbulence. In *Mathematical Control Theory and Finance*, pages 29–53. Springer, 2008.

[16] Ole E Barndorff-Nielsen and Jürgen Schmiegel. Brownian semistationary processes and volatility/intermittency. *Advanced financial modelling*, 8:1–26, 2009.

[17] Fred Espen Benth, Nils Detering, and Paul Krühner. Abstract polynomial processes. *arXiv preprint arXiv:2010.02483*, 2020.

[18] Fred Espen Benth, Nils Detering, and Paul Krühner. Independent increment processes: a multilinearity preserving property. *Stochastics*, 93(6):803–832, 2021.

[19] Fred Espen Benth and Paul Krühner. Representation of infinite-dimensional forward price models in commodity markets. *Communications in Mathematics and Statistics*, 2(1):47–106, 2014.

[20] Fred Espen Benth and Paul Krühner. Derivatives pricing in energy markets: an infinite-dimensional approach. *SIAM Journal on Financial Mathematics*, 6(1):825–869, 2015.

[21] Marc A Berger and Victor J Mizel. Volterra equations with Itô integrals—I. *The Journal of Integral Equations*, pages 187–245, 1980.

[22] Francesca Biagini and Yinglin Zhang. Polynomial diffusion models for life insurance liabilities. *Insurance: Mathematics and Economics*, 71:114–129, 2016.

[23] Jochen Blath, Adrián González Casanova, Noemi Kurt, and Maite Wilke-Berenguer. A new coalescent for seed-bank models. *The Annals of Applied Probability*, 26(2):857 – 891, 2016.

[24] Alessandro Bondi, Giulia Livieri, and Sergio Pulido. Affine Volterra processes with jumps. *Stochastic Processes and their Applications*, 168:104264, 2024.

[25] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance*, 1(2):223, 2001.

[26] Christa Cuchiero. Polynomial processes in stochastic portfolio theory. *Stochastic processes and their applications*, 129(5):1829–1872, 2019.

[27] Christa Cuchiero, Francesco Guida, Luca di Persio, and Sara Svaluto-Ferro. Measure-valued affine and polynomial diffusions. *Preprint ArXiv:2112.15129*, 2021.

[28] Christa Cuchiero, Martin Keller-Ressel, and Josef Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16:711–740, 2012.

[29] Christa Cuchiero, Martin Larsson, and Sara Svaluto-Ferro. Probability measure-valued polynomial diffusions. *Electronic Journal of Probability*, 24:1 – 32, 2019.

[30] Christa Cuchiero and Sara Svaluto-Ferro. Infinite-dimensional polynomial processes. *Finance and Stochastics*, 25(2):383–426, 2021.

[31] Christa Cuchiero, Sara Svaluto-Ferro, and Josef Teichmann. Signature SDEs from an affine and polynomial perspective. *Preprint arXiv:2302.01362*, 2023.

[32] 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.

[33] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge university press, 2014.

[34] Jules Delemotte, Stefano De Marco, and Florent Segonne. Yet another analysis of the sp500 at-the-money skew: Crossover of different power-law behaviours. *Available at SSRN 4428407*, 2023.

[35] Darrell Duffie, Damir Filipović, and Walter Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053, 2003.

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

[37] Alison Etheridge. *An introduction to superprocesses*. Number 20. American Mathematical Soc., 2000.

[38] Damir Filipovic. *Consistency problems for Heath-Jarrow-Morton interest rate models*. Springer Science & Business Media, 2001.

[39] Damir Filipović, Elise Gourier, and Loriano Mancini. Quadratic variance swap models. *Journal of Financial Economics*, 119(1):44–68, 2016.

[40] Damir Filipović and Martin Larsson. Polynomial diffusions and applications in finance. *Finance and Stochastics*, 20(4):931–972, 2016.

[41] Damir Filipović, Martin Larsson, and Sergio Pulido. Markov cubature rules for polynomial processes. *Stochastic Processes and their Applications*, 130(4):1947–1971, 2020.

[42] Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. In *Commodities*, pages 659–690. Chapman and Hall/CRC, 2022.

[43] Gustaf Gripenberg, Stig-Olof Londen, and Olof Staffans. *Volterra integral and functional equations*, volume 34 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1990.

[44] Julien Guyon and Mehdi El Amrani. Does the term-structure of equity at-the-money skew really follow a power law? *Available at SSRN 4174538*, 2022.

[45] Julien Guyon and Jordan Lekeufack. Volatility is (mostly) path-dependent. *Quantitative Finance*, 23(9):1221–1258, 2023.

[46] Yushi Hamaguchi. Weak well-posedness of stochastic Volterra equations with completely monotone kernels and non-degenerate noise. *arXiv preprint arXiv:2310.16030*, 2023.

[47] Antoine Jack Jacquier and Mugad Oumgari. Deep PPDEs for rough local stochastic volatility. *Available at SSRN 3400035*, 2019.

[48] Motoo Kimura. Diffusion models in population genetics. *Journal of Applied Probability*, 1(2):177–232, 1964.

[49] Martin Larsson and Sergio Pulido. Polynomial diffusions on compact quadric sets. *Stochastic Processes and their Applications*, 127(3):901–926, 2017.

[50] Zenghu Li. *Measure-Valued Branching Markov Processes*. Springer, 2022.

[51] Ilkka Norros. A storage model with self-similar input. *Queueing systems*, 16:387–396, 1994.

[52] Léo Parent. The rough path-dependent volatility model. *Available at SSRN 4270481*, 2022.

[53] David J Prömel and David Scheffels. Stochastic Volterra equations with Hölder diffusion coefficients. *Stochastic Processes and their Applications*, 161:291–315, 2023.

[54] Terence Tao. *An introduction to measure theory*, volume 126. American Mathematical Soc., 2011.

[55] Frederi Viens and Jianfeng Zhang. A martingale approach for fractional Brownian motions and related path dependent PDEs. *The Annals of Applied Probability*, 29(6):3489–3540, 2019.

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

[57] Walter Willinger, Murad S Taqqu, Will E Leland, and Daniel V Wilson. Self-similarity in high-speed packet traffic: analysis and modeling of ethernet traffic measurements. *Statistical science*, pages 67–85, 1995.