

# Correlators of Polynomial Processes

Fred Espen Benth\*      Silvia Lavagnini†

July 22, 2021

## Abstract

In the setting of **one-dimensional** polynomial jump-diffusion dynamics, we provide an explicit formula for computing correlators, namely, cross-moments of the process at different time points along its path. The formula appears as a linear combination of exponentials of the generator matrix, extending the well-known moment formula for polynomial processes. The developed framework can, for example, be applied in financial pricing, such as for path-dependent options and in a stochastic volatility models context. In applications to options, having closed and compact formulations is attractive for sensitivity analysis and risk management, since Greeks can be derived explicitly.

**Keywords** Polynomial jump-diffusion process; Correlators; Eliminating and duplicating matrices; Generator matrix; Hankel matrix; Stochastic volatility; Path-dependent option; Greeks.

## 1 Introduction

A jump-diffusion process is called *polynomial* if its extended generator maps any polynomial function to a polynomial function of equal or lower degree. As a consequence, expectations of any polynomial in the future state of the process, conditioned on the information up to the current state, are given by a polynomial in the current state. Conditional moments can thus be calculated in closed form without any knowledge of the probability distribution nor of the characteristic function, up to the computation of the exponential of the generator matrix. The class of polynomial processes includes exponential Lévy processes and affine processes, with the Ornstein–Uhlenbeck processes as a canonical example. Moreover, polynomial jump-diffusions have been studied both in a Markovian [9, 10] and non-Markovian [18] contexts. We refer to [16] for a mathematical analysis on polynomial diffusions.

Because of their closed moment formula, polynomial processes have many applications in finance and one of the first is addressed in [36]. In the literature, we find examples on interest rates [13, 15], stochastic volatility models [1, 2, 16], option pricing [3, 18] and energy modelling [26, 33]. In [10] the properties of jump-diffusion processes are exploited to improve the performance of computational and statistical methods, such as the generalized method of moments, and for variance reduction techniques in Monte Carlo methods. Further examples cover stochastic portfolio theory [11].

We consider a stochastic basis  $(\Omega, \mathcal{F}, \mathbb{P})$  with a filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  and we work in a **one-dimensional setting**. Let  $Y$  be a polynomial jump-diffusion real-valued process  $Y$  and  $H_n(x) \in \mathbb{R}^{n+1}$  a vector basis of polynomials,  $n \geq 1$ . For any polynomial function  $p$  of degree  $n$  with vector of coefficients  $\vec{p} \in \mathbb{R}^{n+1}$  with respect to  $H_n(x)$ , the moment formula gives

$$\mathbb{E}[p(Y(T)) | \mathcal{F}_t] = \vec{p}^\top e^{G_n(T-t)} H_n(Y(t)), \quad 0 \leq t \leq T,$$

with  $G_n \in \mathbb{R}^{(n+1) \times (n+1)}$  the corresponding generator matrix and  $\top$  the transpose operator. In this article, we extend the framework to  $m + 1$  polynomial functions and study conditional expectations of the form

$$\mathbb{E}[p_m(Y(s_0)) p_{m-1}(Y(s_1)) \cdots p_0(Y(s_m)) | \mathcal{F}_t] \tag{1.1}$$

which we call  $(m + 1)$ -point correlators. Here  $t < s_0 < s_1 < \cdots < s_m < T < \infty$  and  $p_k$  are polynomial functions of degree  $n_k$ ,  $k = 0, \dots, m$ . We denote by  $n := \max\{n_0, \dots, n_m\}$  the maximal degree.

\*Department of Mathematics, University of Oslo, 0316 Blindern, Norway; fredb@math.uio.no.

†Department of Mathematics, University of Oslo, 0316 Blindern, Norway; silval@math.uio.no.

For  $m = 0$  equation (1.1) corresponds to computing moments of  $Y$ , which are given by the moment formula. Hence the  $(m + 1)$ -point correlators can in principle be obtained for any  $m > 0$  by iterating the moment formula. For example, for  $m = 1$  one applies the tower rule for  $\mathcal{F}_t \subseteq \mathcal{F}_{s_0}$  to get

$$\mathbb{E}[p_1(Y(s_0))p_0(Y(s_1)) | \mathcal{F}_t] = \mathbb{E}[p_1(Y(s_0))q_0(Y(s_0); s_1 - s_0) | \mathcal{F}_t] \quad (1.2)$$

where

$$q_0(Y(s_0); s_1 - s_0) := \mathbb{E}[p_0(Y(s_1)) | \mathcal{F}_{s_0}] = \vec{p}_0^\top e^{G_{n_0}(s_1 - s_0)} H_{n_0}(Y(s_0))$$

is the polynomial obtained by applying the moment formula to  $p_0(x)$ . In particular,  $q_0(x; s)$  has time dependent coefficients  $q_{0,k}^s$ ,  $s \geq 0$ ,  $k = 0, \dots, n_0$ . The product  $\tilde{p}_1(x; s) := p_1(x)q_0(x; s)$  is then a polynomial function of degree  $n_0 + n_1$  with time dependent coefficients given by

$$\tilde{p}_{1,j}^s = \sum_{k+i=j} p_{1,i} q_{0,k}^s \quad \text{for } j = 0, \dots, n_0 + n_1 \text{ and } s \geq 0.$$

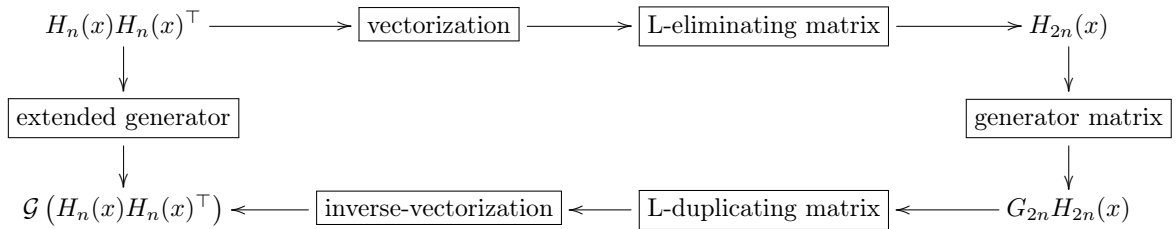
Another application of the moment formula, this time to  $\tilde{p}_1(x; s)$ , produces an expression for (1.2) of the form

$$\mathbb{E}[p_1(Y(s_0))p_0(Y(s_1)) | \mathcal{F}_t] = \vec{p}_1^{\tilde{p}_1^{s_1 - s_0} \top} e^{G_{n_0+n_1}(s_0 - t)} H_{n_0+n_1}(Y(t)).$$

This procedure can then be iterated for larger values of  $m$ . However, performing the calculations is non-trivial because of the algebraic complexity of manipulating the expressions involved. With this article we make headway on this issue by providing a fully explicit closed formula for correlators.

The key for proving the moment formula lies in the existence of the generator matrix  $G_n$ : for a fixed  $n$  and a fixed basis vector of polynomials  $H_n(x)$ , this is the linear representation of the action of the extended generator on  $H_n(x)$ . However, for  $m = 1$  we must deal with the product of two basis vectors, which is an object of the form  $H_n(x)H_n(x)^\top \in \mathbb{R}^{(n+1) \times (n+1)}$  and for which a generator matrix cannot be constructed. We then consider the vectorization of  $H_n(x)H_n(x)^\top$ , namely we stack the columns of  $H_n(x)H_n(x)^\top$  into a single column vector. The matrix  $H_n(x)H_n(x)^\top$  contains however redundant terms and so does its vectorization. For  $H_n(x) := (1, x, x^2, \dots, x^n)^\top$ , which is the case we consider here, redundant terms means repeated powers of  $x$ . This implies that the corresponding generator matrix contains equal rows and/or zero columns, making it impossible to generalize the framework to  $m > 1$ .

We resolve this issue by introducing two linear operators, the first of which we call the *L-eliminating matrix*. This eliminates from the vectorization of  $H_n(x)H_n(x)^\top$  the redundant powers of  $x$  and returns a vector that coincides with  $H_{2n}(x)$ , for which there exists the generator matrix  $G_{2n}$ . Using the inverse operator, called the *L-duplicating matrix*, we then recover the full-dimensional vector, and, finally, via inverse-vectorization we obtain the linear operator required, which allows to compute the correlator formula for  $m = 1$ . We summarize these steps in the following graph:



These steps work also when increasing further the number of polynomials. For  $m + 1 > 1$ , we must deal with  $m + 1 > 1$  basis vectors  $H_n(x)$ . This leads to an object whose structure is more complex and requires the appropriate eliminating and duplicating matrices, for which we prove a recursion formula in the number of polynomials  $m \geq 1$ . With these, we compute the general correlator formula.

As we shall see, for  $H_n(x) = (1, x, x^2, \dots, x^n)^\top$ , the matrix  $H_n(x)H_n(x)^\top \in \mathbb{R}^{(n+1) \times (n+1)}$  is a so-called *Hankel matrix*, for which the elements on the same skew-diagonals coincide. Hankel matrices constitute an important family of matrices that play a fundamental role in diverse fields, from computer science to engineering, mathematics and statistics [31]. They are indeed applied, for example, in theory of moments [14, 30], time series analysis [19, 20], signal analysis [23, 24], and in theory of orthogonal polynomials [32]. This means that (part of) our analysis can be applied in many different fields, going

beyond the polynomial jump-diffusion theory studied here. We also mention that a Hankel matrix is a "row-reversed" Toeplitz matrix, so that some of the results proved in the current article can be adapted to this other class of matrices for possibly further applications.

We point out that our correlator formula is not really an alternative to applying iteratively the moment formula in the sense discussed above. Indeed, it strongly relies on it combined with the tower rule for  $\mathcal{F}_t \subseteq \mathcal{F}_{s_0} \subseteq \dots \subseteq \mathcal{F}_{s_m}$ . However, it provides a solution to the algebraic burden that arises when applying the moment formula directly. The correlator formula is indeed fully explicit, while getting an explicit expression is not straightforward when iterating the moment formula directly. Since having closed formulas is an advantage for example in those applications that require to differentiate, such as for computing Greeks, our approach is thus more convenient. Not surprisingly, numerical experiments show that the correlator values obtained with our formula coincide with the values obtained by iterating the moment formula. Moreover, the time costs for the two approaches is comparable up to around  $m = 10$  polynomials. We compare the results with a Monte Carlo approach, showing that this latter one is outperformed from a time cost point of view, in addition to exhibiting low degrees of accuracy. We stress that the correlator formula only involves linear combinations of the matrix exponential of the generator matrix. Assuming these exponential matrices to be exact, we thus have a formula for correlators which in practice is exact.

We finally provide two recursion formulas for the generator matrix and its matrix exponential. Despite several approaches have been studied for calculating efficiently the matrix exponential of a block triangular matrix [21, 27], up to our knowledge, no rigorous study in terms of the building blocks has been developed yet concerning the generator matrix and its exponential. These results can then be applied for analytical purposes as mentioned before. We point out that our framework is based on the monomial basis, which appears convenient for obtaining formulas more easily and explicitly. However, it can be extended to any other polynomial basis, provided the matrix for the change of basis. For practical applications, orthogonal basis are indeed more convenient, but analytically more challenging.

The rest of the paper is organized as follows. In Section 1.1 we clarify the name correlators and give some financial motivations for studying them. In Section 2 we introduce rigorously polynomial processes and the generator matrix. In Section 3 we solve the two-point correlators problem, presenting the main tools and framework which allows to solve the  $(m + 1)$ -point correlators problem in Section 4. In Section 5 we provide two recursions for the generator matrix and its matrix exponential, together with the formula for the change of basis. Finally, in Section 6 we consider some applications and numerical aspects and Section 7 concludes with some remarks. Appendix A contains some combinatorial properties of the operators introduced in the paper and Appendix B the proofs of the main results.

## 1.1 Motivations

In [5, Section 9.3] the authors define the concept of *correlator*, a standard tool in turbulence theory. For  $t < s_0 < s_1 < T < \infty$  and  $k_0, k_1 \in \mathbb{N}$ , the correlator of order  $(k_0, k_1)$  between  $Y(s_0)$  and  $Y(s_1)$  is a generalization of the autocorrelation defined by

$$\text{Corr}_{k_0, k_1}(s_0, s_1; t) = \frac{\mathbb{E} [Y(s_0)^{k_0} Y(s_1)^{k_1} | \mathcal{F}_t]}{\mathbb{E} [Y(s_0)^{k_0} | \mathcal{F}_t] \mathbb{E} [Y(s_1)^{k_1} | \mathcal{F}_t]}.$$

In this article we extend this definition of correlator to any expectation like the one in equation (1.1).

We introduce now two possible applications: Asian option pricing and pricing in the context of stochastic volatility models. We intend to motivate our analysis, leaving details aside for future work.

### 1.1.1 Path-dependent options

We consider path-dependent options, such as Asian options, for which the entire path of the price process within the settlement period  $[t, T]$  is taken into account by the payoff function [25]. If  $Y$  is the risk-neutral price dynamics of the underlying asset,  $r > 0$  the risk-free interest rate and  $\varphi$  the payoff function, the discounted price at time  $t$  for an Asian-style option settled against the discrete arithmetic

average of the spot price  $Y$  in the settlement period is given by

$$\Pi(t) = e^{-r(T-t)} \mathbb{E} \left[ \varphi \left( \frac{1}{m+1} \sum_{j=0}^m Y(s_j) \right) \middle| \mathcal{F}_t \right] \quad \text{for } t < s_0 < s_1 < \dots < s_m = T \text{ and } m \geq 0. \quad (1.3)$$

This kind of options was traded a decade ago at Nord Pool, the Nordic commodity market for electricity [34]. Other classes of derivatives of similar kind are calendar spread options and options on baskets of assets evaluated at different times, as well as Asian options with continuous averaging.

For  $\varphi$  a real-valued continuous function on a bounded interval, we consider  $\hat{\varphi}$  as the polynomial approximation of  $\varphi$ , e.g., by Hermite polynomials or Taylor expansions, depending on the nature of  $\varphi$  itself. Then the price for the Asian option in equation (1.3) is found by

$$\Pi(t) \approx e^{-r(T-t)} \mathbb{E} \left[ \hat{\varphi} \left( \frac{1}{m+1} \sum_{j=0}^m Y(s_j) \right) \middle| \mathcal{F}_t \right] = e^{-r(T-t)} \sum_{\mathbf{k}} \alpha_{\mathbf{k}} \mathbb{E} [Y(s_0)^{k_0} Y(s_1)^{k_1} \dots Y(s_m)^{k_m} \middle| \mathcal{F}_t]$$

for certain coefficients  $\{\alpha_{\mathbf{k}}\}_{\mathbf{k}}$  and the multi-index  $\mathbf{k} = (k_1, \dots, k_m)$ . This leads to study conditional expectations of the form  $\mathbb{E} [Y(s_0)^{k_0} Y(s_1)^{k_1} \dots Y(s_m)^{k_m} \middle| \mathcal{F}_t]$ , which is a particular instance of equation (1.1) obtained with  $p_j(x) = x^{k_j}$ ,  $j = 0, \dots, m$ . In [28] the author derives explicit price formulas for call-style discrete-average arithmetic Asian options by following this approach, namely by approximating the payoff function with orthogonal polynomials and by the correlator formula developed in this article.

### 1.1.2 Stochastic volatility models

For  $0 \leq t \leq T$  we consider the process  $X$  defined by  $X(T) = \int_t^T \sigma(s) dB(s)$ , with  $B$  a standard Brownian motion and  $\sigma$  a volatility process which we assume to be independent from  $B$ . If  $\varphi$  is the payoff function and  $r > 0$  the risk-free interest rate, we want to price a financial derivative like follows:

$$\Pi(t) = e^{-r(T-t)} \mathbb{E} [\varphi(X(T)) \middle| \mathcal{F}_t].$$

A possible approach suggested in [8] is to consider the Fourier transform  $\hat{\varphi}$  of  $\varphi$ . Under appropriate integrability conditions on  $\varphi$ , we then write that  $\varphi(x) = \int_{-\infty}^{\infty} \hat{\varphi}(z) e^{2\pi i x z} dz$ , and the option price becomes

$$\Pi(t) = e^{-r(T-t)} \mathbb{E} \left[ \int_{-\infty}^{\infty} \hat{\varphi}(z) e^{2\pi i X(T) z} dz \middle| \mathcal{F}_t \right].$$

For  $\sigma$  and  $B$  independent, by the tower rule, we now condition with respect to the filtration  $\{\mathcal{F}_t^\sigma\}_{t \geq 0}$  generated by  $\sigma$  up to time  $T$ . The process  $X(T)$  has then a Gaussian distribution with mean 0 and variance  $\int_t^T \sigma^2(s) ds$ , hence

$$\Pi(t) = e^{-r(T-t)} \mathbb{E} \left[ \int_{-\infty}^{\infty} \hat{\varphi}(z) e^{-2\pi^2 z^2 \int_t^T \sigma^2(s) ds} dz \middle| \mathcal{F}_t \right] = e^{-r(T-t)} \int_{-\infty}^{\infty} \hat{\varphi}(z) \mathbb{E} \left[ e^{\lambda \int_t^T \sigma^2(s) ds} \middle| \mathcal{F}_t \right] dz$$

for  $\lambda \leq 0$ . By considering the Taylor expansion for the exponential function, the expectation becomes

$$\mathbb{E} \left[ e^{\lambda \int_t^T \sigma^2(s) ds} \middle| \mathcal{F}_t \right] = \mathbb{E} \left[ \sum_{k=0}^{\infty} \frac{1}{k!} \left( \lambda \int_t^T \sigma^2(s) ds \right)^k \middle| \mathcal{F}_t \right] = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \mathbb{E} \left[ \left( \int_t^T \sigma^2(s) ds \right)^k \middle| \mathcal{F}_t \right], \quad (1.4)$$

that is, we need to find the moments of the integrated volatility,  $\int_t^T \sigma^2(s) ds$ . For  $Y(s) := \sigma^2(s)$  modelled by a polynomial process, we notice that the bivariate process  $(Y(T), \int_t^T Y(s) ds)$  is also polynomial. Hence the moments of  $\int_t^T \sigma^2(s) ds$  can be computed with the moment formula applied to this bivariate polynomial process. As an alternative approach, using iteratively the fundamental theorem of calculus,

it can be proved that

$$\left( \int_t^T \sigma^2(s) ds \right)^k = \frac{1}{k!} \int_t^T \int_t^{s_k} \cdots \int_t^{s_2} Y(s_1) Y(s_2) \cdots Y(s_k) ds_1 \cdots ds_k, \quad k \geq 1, \quad (1.5)$$

where  $t < s_1 < s_2 < \cdots < s_k < T$  is a partition of  $[t, T]$ . Combining equations (1.4) and (1.5), we get

$$\mathbb{E} \left[ e^{\lambda \int_t^T Y(s) ds} \middle| \mathcal{F}_t \right] = \sum_{k=0}^{\infty} \frac{\lambda^k}{(k!)^2} \int_t^T \int_t^{s_k} \cdots \int_t^{s_2} \mathbb{E} [Y(s_1) Y(s_2) \cdots Y(s_k) | \mathcal{F}_t] ds_1 \cdots ds_k, \quad (1.6)$$

so that for every  $k \geq 1$  we need to study expectations of the form of  $\mathbb{E} [Y(s_1) Y(s_2) \cdots Y(s_k) | \mathcal{F}_t]$ .

Interestingly,  $Y(T) = \int_t^T \sigma^2(s) ds$  appears also in the pricing of VIX-derivatives, that is, derivatives on the realized variance and volatility. For derivatives paying  $\psi(Y(T))$ , we can use the Fourier approach above as long as  $\psi$  is integrable with an integrable Fourier transform  $\widehat{\psi}$ , to end up again with a conditional expectation as in equation (1.6). The volatility swap price, i.e., the swap price on the realized volatility, is defined as the conditional expected value  $\mathbb{E}[Y^{1/2}(T) | \mathcal{F}_t]$  for  $t \leq T$ . Expanding  $x \mapsto \sqrt{x} 1_{x \geq 0}$  in the Hermite functions, being a basis for the space  $L^2(\mathbb{R}, \gamma(x) dx)$  with  $\gamma$  the standard normal density function, we obtain a series representation of the swap price in terms of conditional moments of  $Y(T)$ . For more details on VIX-derivatives with numerical examples based on Fourier methods, we refer to [7]. Here the Barndorff-Nielsen & Shephard stochastic volatility model is considered for  $\sigma^2(t)$ , which is a polynomial jump process as will be defined in the next section. **Moreover, we refer to [12] for an application of the moment formula to price VIX-derivatives in the setting of stochastic volatility modelled with a polynomial process lifted to infinite dimension.**

## 2 Polynomial processes

Following [18], we consider a jump-diffusion operator on  $\mathbb{R}$  of the form

$$\mathcal{G}f(x) = b(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x) + \int_{\mathbb{R}} (f(x+z) - f(x) - f'(x)z) \ell(x, dz), \quad (2.1)$$

for some measurable maps  $b : \mathbb{R} \rightarrow \mathbb{R}$  and  $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ , and a transition kernel  $\ell : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  such that  $\ell(x, \{0\}) = 0$  and  $\int_{\mathbb{R}} |z| \wedge |z|^2 \ell(x, dz) < \infty$  for all  $x \in \mathbb{R}$ . We then let  $Y$  be the jump-diffusion stochastic process having  $\mathcal{G}$  as extended generator. This means that for every bounded function  $f : \mathbb{R} \rightarrow \mathbb{R}$  with continuous second derivative and  $y \in \mathbb{R}$ ,  $f(Y(t)) - f(y) - \int_0^t \mathcal{G}f(Y(s)) ds$  is a  $(\mathcal{F}_t, \mathbb{P}_y)$ -local martingale.

We now denote with  $\text{Pol}(\mathbb{R})$  the space of polynomials on  $\mathbb{R}$  and with  $\text{Pol}_n(\mathbb{R})$  the subspace of all polynomials of degree less than or equal to  $n$  on  $\mathbb{R}$ . We say that  $\mathcal{G}$  is *well defined* on  $\text{Pol}(\mathbb{R})$  if  $\int_{\mathbb{R}} |z|^n \ell(x, dz) < \infty$  for all  $x \in \mathbb{R}$  and  $n \geq 2$ , and  $\mathcal{G}f(x) = 0$  for  $f(x) \equiv 0$  on  $\mathbb{R}$ . We then give the following definition of a polynomial jump-diffusion process.

**Definition 2.1** (Polynomial jump-diffusion process). We call the operator  $\mathcal{G}$  *polynomial* if it is well defined on  $\text{Pol}(\mathbb{R})$  and it maps  $\text{Pol}_n(\mathbb{R})$  to itself for each  $n \in \mathbb{N}$ . In this case, we call  $Y$  a *polynomial jump-diffusion process*.

Assuming  $\mathcal{G}$  to be polynomial, from [18] Lemma D.4], the process

$$p(Y(t)) - p(y) - \int_0^t \mathcal{G}p(Y(s)) ds \quad \text{is a } (\mathcal{F}_t, \mathbb{P}_y)\text{-martingale} \quad (2.2)$$

for all  $p \in \text{Pol}_n(\mathbb{R})$ ,  $y \in \mathbb{R}$  and  $t \geq 0$ . This basically means that all increments of (2.2) have vanishing expectation. Moreover, from [18] Lemma 1], the polynomial property of  $\mathcal{G}$  can be characterized in terms of its coefficients: it must hold that

$$b \in \text{Pol}_1(\mathbb{R}), \quad \sigma^2 + \int_{\mathbb{R}} z^2 \ell(\cdot, dz) \in \text{Pol}_2(\mathbb{R}) \quad \text{and} \quad \int_{\mathbb{R}} z^m \ell(\cdot, dz) \in \text{Pol}_m(\mathbb{R}) \quad \text{for all } m \geq 3.$$

To fulfil these conditions, we shall assume that for every  $m \geq 2$  there exist  $b_0, b_1, \sigma_0, \sigma_1, \sigma_2, \xi_0^m, \dots, \xi_m^m$

real constants such that

$$b(x) = b_0 + b_1x, \quad \sigma^2(x) = \sigma_0 + \sigma_1x + \sigma_2x^2 \quad \text{and} \quad \int_{\mathbb{R}} z^m \ell(x, dz) = \sum_{i=0}^m \xi_i^m x^i. \quad (2.3)$$

We consider an example.

*Example 2.1.* Let  $B$  a standard one-dimensional Brownian motion and  $\tilde{N}(dt, dz)$  a compensated Poisson random measure with compensator  $\nu(dz)dt$ . We consider the jump-diffusion SDE given by

$$dY(t) = b(Y(t))dt + \sigma(Y(t))dB(t) + \int_{\mathbb{R}} \delta(Y(t^-), z)\tilde{N}(dt, dz)$$

with drift, volatility and jump size functions of the form

$$b(x) := b_0 + b_1x, \quad \sigma^2(x) := \sigma_0 + \sigma_1x + \sigma_2x^2, \quad \delta(x, z) := \delta_0(z) + \delta_1(z)x,$$

for  $b_0, b_1, \sigma_0, \sigma_1, \sigma_2 \in \mathbb{R}$ , and  $\delta_0, \delta_1 : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\int_{\mathbb{R}} |\delta_i(z)|^m \nu(dz) < \infty$  for all  $m \geq 2$  and  $i = 0, 1$ . The SDE has a unique strong solution  $Y(t)$  for each initial condition  $Y(0) = y \in \mathbb{R}$ . Moreover,  $Y(t)$  is a polynomial jump-diffusion with linear drift  $b \in \text{Pol}_1(\mathbb{R})$ , quadratic diffusion  $\sigma^2 \in \text{Pol}_2(\mathbb{R})$ , and jump measure  $\ell(x, dz)$  given by  $\int_{\mathbb{R}} f(z)\ell(x, dz) = \int_{\mathbb{R}} f(\delta(x, z))\nu(dz)$ . In particular, for  $m \geq 2$ , by the binomial theorem we find that

$$\int_{\mathbb{R}} (\delta_0(z) + \delta_1(z)x)^m \nu(dz) = \sum_{i=0}^m \binom{m}{i} \int_{\mathbb{R}} \delta_0(z)^{m-i} \delta_1(z)^i \nu(dz) x^i,$$

so that in this case the constants  $\xi_i^m$  introduced in equation (2.3) are  $\xi_i^m = \binom{m}{i} \int_{\mathbb{R}} \delta_0(z)^{m-i} \delta_1(z)^i \nu(dz)$ , for  $i = 0, \dots, m$ .

## 2.1 The generator matrix

We consider the set  $\{1, x, x^2, \dots, x^n\}$  as basis for  $\text{Pol}_n(\mathbb{R})$ , and we introduce the vector valued function

$$H_n : \mathbb{R} \longrightarrow \mathbb{R}^{n+1}, \quad H_n(x) = (1, x, x^2, \dots, x^n)^\top.$$

We shall now report rigorously the moment formula for polynomial processes from [18 Theorem 2.5], for which we include the proof that will be useful for the analysis in Section 3

**Theorem 2.1** (Moment formula). *For  $n \geq 1$  and  $Y$  polynomial process with extended generator  $\mathcal{G}$ :*

1. *There exists a generator matrix  $G_n \in \mathbb{R}^{(n+1) \times (n+1)}$  such that*

$$\mathcal{G}H_n(x) = G_n H_n(x).$$

2. *For every  $p \in \text{Pol}_n(\mathbb{R})$  with vector of coefficients  $\vec{p} \in \mathbb{R}^{n+1}$ , the moment formula holds*

$$\mathbb{E}[p(Y(T)) | \mathcal{F}_t] = \vec{p}^\top e^{G_n(T-t)} H_n(Y(t)), \quad 0 \leq t \leq T.$$

*Proof.* We take  $f(x) = x^k$  for  $0 \leq k \leq n$ . Since  $Y$  is a polynomial process, there exists  $\vec{q}_k \in \mathbb{R}^{n+1}$  such that  $\mathcal{G}x^k = \vec{q}_k^\top H_n(x)$ ,  $0 \leq k \leq n$ . With these vectors we can then construct a matrix  $G_n \in \mathbb{R}^{(n+1) \times (n+1)}$  such that  $\mathcal{G}H_n(x) = G_n H_n(x)$ . This proves claim 1. Next, by equation (2.2) and claim 1., we write

$$\begin{aligned} \mathbb{E}[p(Y(T)) | \mathcal{F}_t] &= \vec{p}^\top \mathbb{E}[H_n(Y(T)) | \mathcal{F}_t] = \vec{p}^\top H_n(Y(t)) + \int_t^T \vec{p}^\top \mathbb{E}[\mathcal{G}H_n(Y(s)) | \mathcal{F}_t] ds \\ &= \vec{p}^\top H_n(Y(t)) + \vec{p}^\top G_n \int_t^T \mathbb{E}[H_n(Y(s)) | \mathcal{F}_t] ds. \end{aligned} \quad (2.4)$$

We focus on  $H_n(Y(T))$ . For  $Z(s) := \mathbb{E}[H_n(Y(s)) | \mathcal{F}_t]$ , equation (2.4) can be written in differential form as  $dZ(s) = G_n Z(s)ds$ , whose solution, by separation of variables, is  $Z(T) = e^{G_n(T-t)} Z(t)$ . From the

definition of  $Z$ , multiplying by the vector  $\bar{p}^\top$ , we conclude the proof.  $\square$

Theorem 2.1 tells us that  $\mathbb{E}[p(Y(T)) | \mathcal{F}_t]$  is a polynomial function in  $Y(t)$  for every  $p \in \text{Pol}_n(\mathbb{R})$ . We point out that this holds for every choice of the vector basis of polynomials, despite in this paper we focus on the vector basis of monomials. Moreover, we stress the fact that the moment formula strongly relies on the existence of the generator matrix  $G_n$  and on the martingale property of the process in equation (2.2). These two elements will be the key for all our framework.

*Example 2.2.* Let  $n = 2$ . Then  $H_2(x) = (1, x, x^2)^\top$  and  $\mathcal{G}H_2(x) = (\mathcal{G}1, \mathcal{G}x, \mathcal{G}x^2)^\top$ . In particular, from equations (2.1) and (2.3), we get that  $\mathcal{G}1 = 0$ ,  $\mathcal{G}x = b_0 + b_1x$  and

$$\mathcal{G}x^2 = (\sigma_0 + \xi_0^2) + (\sigma_1 + 2b_0 + \xi_1^2)x + (\sigma_2 + 2b_1 + \xi_2^2)x^2.$$

One finds that the generator matrix  $G_2 \in \mathbb{R}^{3 \times 3}$  satisfying Theorem 2.1 is then

$$G_2 = \begin{pmatrix} 0 & 0 & 0 \\ b_0 & b_1 & 0 \\ \sigma_0 + \xi_0^2 & \sigma_1 + 2b_0 + \xi_1^2 & \sigma_2 + 2b_1 + \xi_2^2 \end{pmatrix}.$$

### 3 Two-point correlators

Aiming at solving the  $(m + 1)$ -point correlators problem in equation (1.1), we start the analysis for  $m = 1$  because the tools and ideas developed to solve this case are crucial to understand the framework that will be generalized to  $m + 1$  polynomials in Section 4. For  $m = 1$ , equation (1.1) reads like

$$C_{p_0, p_1}(s_0, s_1; t) := \mathbb{E}[p_1(Y(s_0))p_0(Y(s_1)) | \mathcal{F}_t], \quad 0 \leq t < s_0 < s_1,$$

with  $p_0 \in \text{Pol}_{n_0}(\mathbb{R})$  and  $p_1 \in \text{Pol}_{n_1}(\mathbb{R})$ . In particular, for  $n := \max\{n_0, n_1\}$ , we can represent the two polynomial functions  $p_0$  and  $p_1$  respectively by  $p_0(x) = \bar{p}_0^\top H_n(x)$  and  $p_1(x) = \bar{p}_1^\top H_n(x)$ . By the tower rule for  $\mathcal{F}_t \subseteq \mathcal{F}_{s_0}$  and the moment formula in Theorem 2.1,  $C_{p_0, p_1}(s_0, s_1; t)$  can be rewritten by

$$C_{p_0, p_1}(s_0, s_1; t) = \bar{p}_1^\top \mathbb{E}[H_n(Y(s_0))H_n(Y(s_1))^\top | \mathcal{F}_t] e^{G_n^\top (s_1 - s_0)} \bar{p}_0. \quad (3.1)$$

This means that the conditional expectation of the product of two polynomial functions reduces to the conditional expectation of the outer product of the basis function  $H_n(x)$  with itself, which is a matrix of monomial functions of the form

$$X_n(x) := H_n(x)H_n(x)^\top = \begin{pmatrix} 1 & x & x^2 & \dots & x^n \\ x & x^2 & x^3 & \dots & x^{n+1} \\ x^2 & x^3 & x^4 & \dots & x^{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x^n & x^{n+1} & x^{n+2} & \dots & x^{2n} \end{pmatrix}. \quad (3.2)$$

By equation (2.2) we notice that

$$\mathbb{E}[X_n(Y(s_0)) | \mathcal{F}_t] = X_n(Y(t)) + \int_t^{s_0} \mathbb{E}[\mathcal{G}X_n(Y(s)) | \mathcal{F}_t] ds. \quad (3.3)$$

Thus, in the same spirit of the proof of Theorem 2.1 we seek a linear operator  $G_n^{(1)}$  such that

$$G_n^{(1)} : \mathbb{R}^{(n+1) \times (n+1)} \longrightarrow \mathbb{R}^{(n+1) \times (n+1)}, \quad \mathcal{G}X_n(x) = G_n^{(1)}X_n(x), \quad (3.4)$$

which is the equivalent linear operator in the two-polynomial setting to the generator matrix  $G_n$ . However,  $G_n^{(1)}$  cannot be represented with a matrix. We notice that  $G_n$  maps a vector to a vector, while  $G_n^{(1)}$  maps a matrix to a matrix. The idea is then to transform the matrix-matrix problem into a vector-vector problem and to construct the linear operator  $G_n^{(1)}$  in terms of the generator matrix  $G_n$ . We start by introducing the following operators for a general matrix  $A \in \mathbb{R}^{n \times m}$ .

**Definition 3.1** (Vectorization and inverse-vectorization). Given a matrix  $A \in \mathbb{R}^{n \times m}$  whose  $j$ -th column we denote by  $A_{:j}$ , we define  $vec : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{nm}$  as the operator that associates to  $A$  the  $nm$ -column vector

$$vec(A) = (A_{:1}^\top \ A_{:2}^\top \cdots \cdots \ A_{:m}^\top)^\top$$

which is called the *vectorization* of  $A$ . For  $v = vec(A)$ , we then define  $vec^{-1} : \mathbb{R}^{nm} \rightarrow \mathbb{R}^{n \times m}$  as the operator that associates to the vector  $v$  the  $n \times m$  matrix  $B = vec^{-1}(v)$  such that  $[B]_{i,j} = v_{n(j-1)+i}$ , for  $i = 1, \dots, n$  and  $j = 1, \dots, m$ . We say that  $B$  is the *inverse-vectorization* of  $v$ . Moreover,  $B$  and  $A$  coincide.

We then address the problem of finding a linear operator  $G_n^{(1)}$  transforming  $X_n(x)$  into  $\mathcal{G}X_n(x)$ , to the problem of finding a matrix  $\tilde{G}_n^{(1)} \in \mathbb{R}^{(n+1)^2 \times (n+1)^2}$  such that

$$\mathcal{G}vec(X_n(x)) = \tilde{G}_n^{(1)}vec(X_n(x)), \quad (3.5)$$

where  $\mathcal{G}vec(X_n(x)) = vec(\mathcal{G}X_n(x))$ . The operator  $G_n^{(1)}$  satisfying equation (3.4) is then obtained by composing the matrix  $\tilde{G}_n^{(1)}$  with the  $vec$  and  $vec^{-1}$  operators, namely

$$G_n^{(1)} = vec^{-1} \circ \tilde{G}_n^{(1)} \circ vec. \quad (3.6)$$

We consider an example.

*Example 3.1.* Let  $n = 1$ . We seek  $\tilde{G}_1^{(1)} \in \mathbb{R}^{4 \times 4}$  such that  $(\mathcal{G}1, \mathcal{G}x, \mathcal{G}x, \mathcal{G}x^2)^\top = \tilde{G}_1^{(1)}(1, x, x, x^2)^\top$ . Two suitable choices of  $\tilde{G}_1^{(1)}$  are

$$\tilde{G}_1^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ b_0 & b_1 & 0 & 0 \\ b_0 & b_1 & 0 & 0 \\ \sigma_0 + \xi_0^2 & \sigma_1 + 2b_0 + \xi_1^2 & 0 & \sigma_2 + 2b_1 + \xi_2^2 \end{pmatrix} \quad \text{and}$$

$$\tilde{G}_1^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ b_0 & b_1/2 & b_1/2 & 0 \\ b_0 & b_1/2 & b_1/2 & 0 \\ \sigma_0 + \xi_0^2 & (\sigma_1 + 2b_0 + \xi_1^2)/2 & (\sigma_1 + 2b_0 + \xi_1^2)/2 & \sigma_2 + 2b_1 + \xi_2^2 \end{pmatrix}.$$

We notice from Example 3.1 that the first  $\tilde{G}_1^{(1)}$  has two identical rows and a null column, while the second  $\tilde{G}_1^{(1)}$  has both two identical rows and two identical columns. This is due to the double presence of the term  $\mathcal{G}x$  in  $vec(\mathcal{G}X_1(x))$ , or, analogously, the double presence of the term  $x$  in  $vec(X_1(x))$ . Increasing the value of  $n$ , the number of redundant terms in  $vec(\mathcal{G}X_n(x))$  and  $vec(X_n(x))$  increases, hence to find a recursion for the matrix  $\tilde{G}_n^{(1)}$  seems not an easy task. Moreover, we would like to write the matrix  $\tilde{G}_n^{(1)}$  in terms of the generator matrix  $G_n$ . We shall solve this issue in the next section.

### 3.1 The L-vectorization

Looking at equation (3.2), we notice that a possible way, among others, to get from the matrix  $X_n(x)$  all the elements without repetition (that is equivalent to get all the powers of  $x$  from 0 to  $2n$  without repetition) is to select the first column and the last row. For this, we introduce the following operator.

**Definition 3.2** (L-vectorization). Given a matrix  $A \in \mathbb{R}^{n \times m}$  with elements  $[A]_{i,j} = a_{i,j}$ ,  $1 \leq i \leq n$  and  $1 \leq j \leq m$ , we define the *L-vectorization* of  $A$  as the operator  $vecL : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}^{n+m-1}$  that associates to  $A$  the  $(n+m-1)$ -column vector obtained by selecting the first column and the last row of  $A$ , namely

$$vecL(A) = (a_{1,1} \ a_{2,1} \cdots \cdots \ a_{n,1} \ a_{n,2} \cdots \cdots \ a_{n,m})^\top.$$

Intuitively, the  $vecL$  operator is a linear operator selecting from the matrix  $A$  the elements that together form the biggest "L" inscribed in  $A$ . In [29], the authors introduce the half-vectorization operator, which, starting from a matrix  $A$ , returns the column vector obtained by stacking together the columns of the lower-triangular matrix contained in  $A$ . Moreover, they provide two matrices, the



eliminating matrix and the duplicating matrix, that, respectively, transform the vectorization of  $A$  into the half-vectorization, and vice-versa. We aim at the same kind of results for the L-vectorization. The existence of such matrices tells us that there exist a linear transformation to remove the duplicates from  $\text{vec}(X_n(x))$  (what we call the *L-eliminating matrix*) and the corresponding inverse linear transformation (the *L-duplicating matrix*).

From now on, we shall denote with  $\vec{e}_{k,j}$  the  $j$ -th canonical basis vector in  $\mathbb{R}^k$ , with  $I_k$  the identity matrix in  $\mathbb{R}^{k \times k}$ , and with  $\otimes$  the Kronecker product, for which we recall the definition.

**Definition 3.3** (Kronecker product). The *Kronecker product* of a matrix  $A \in \mathbb{R}^{n \times m}$  with elements  $[A]_{i,j} = a_{i,j}$ ,  $1 \leq i \leq n$  and  $1 \leq j \leq m$ , and a matrix  $B \in \mathbb{R}^{r \times s}$ , is the matrix  $A \otimes B \in \mathbb{R}^{nr \times ms}$  given by

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots & a_{1,m}B \\ \vdots & \ddots & \vdots \\ a_{n,1}B & \cdots & a_{n,m}B \end{pmatrix}.$$

We define now the L-eliminating matrix.

**Theorem 3.1** (L-eliminating matrix). *For every  $n, m \geq 1$  and for every matrix  $A \in \mathbb{R}^{n \times m}$ , there exists an L-eliminating matrix  $E_{n,m} \in \mathbb{R}^{(n+m-1) \times nm}$  such that*

$$E_{n,m} \text{vec}(A) = \text{vec}L(A) \quad (3.7)$$

$$E_{n,m} = \sum_{i=1}^n \vec{e}_{n+m-1,i} \otimes \vec{e}_{m,1}^\top \otimes \vec{e}_{n,i}^\top + \sum_{i=2}^m \vec{e}_{n+m-1,n+i-1} \otimes \vec{e}_{m,i}^\top \otimes \vec{e}_{n,n}^\top. \quad (3.8)$$

**Corollary 3.2.** *For every  $n \geq 1$ , the L-eliminating matrix  $E_{n+1} \in \mathbb{R}^{(2n+1) \times (n+1)^2}$  transforming  $\text{vec}(X_n(x))$  into  $\text{vec}L(X_n(x))$  is given by  $E_{n+1} := E_{n+1,n+1}$ .*

*Example 3.2.* Let  $n = m = 2$ . Then equation (3.8) becomes

$$E_{2,2} = \sum_{i=1}^2 \vec{e}_{3,i} \otimes \vec{e}_{2,1}^\top \otimes \vec{e}_{2,i}^\top + \sum_{i=2}^2 \vec{e}_{3,2+i} \otimes \vec{e}_{2,i}^\top \otimes \vec{e}_{2,2}^\top = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

For  $A \in \mathbb{R}^{2 \times 2}$  of the form  $A = \begin{pmatrix} a_1 & a_3 \\ a_2 & a_4 \end{pmatrix}$ , we can verify that  $E_{2,2}$  satisfies the definition of an L-eliminating matrix in equation (3.7), as indeed

$$E_{2,2} \text{vec}(A) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_4 \end{pmatrix} = \text{vec}L(A).$$

Moreover, when applied to  $\text{vec}(X_1(x)) = (1, x, x, x^2)^\top$ , it eliminates the double value  $x$ .

Next, we want to define an inverse operator to  $E_{n,m}$ , namely a linear mapping transforming the L-vectorization of a matrix  $A$  into its vectorization. However, this inverse operation is not well defined in the space of matrices in  $\mathbb{R}^{n \times m}$ . Indeed, when applying  $E_{n,m}$  to  $\text{vec}(A)$ , we go from a space of dimension  $nm$  to a space of dimension  $n + m - 1 < nm$ . Then the inverse transformation in general does not exist. It is necessary to find a suitable subspace of  $\mathbb{R}^{n \times m}$  of dimension of  $n + m - 1$  so that image space dimension and domain space dimension coincide. In [29], the authors face a similar issue which they solve by restricting the domain to the space of symmetric matrices.

Looking at the matrix of functions  $X_n(x)$ , we notice that each ascending skew-diagonal from left to right is constant, which is a property of the *Hankel matrices*. These are usually defined in the square case; we however consider an extended definition to rectangular matrices as, for example, in [17].

**Definition 3.4** (Hankel matrix). We define  $\mathcal{A}_{n,m} \subset \mathbb{R}^{n \times m}$  as the space of matrices whose elements on the same skew-diagonal coincide. We distinguish two cases corresponding to whether  $n \geq m$  or  $m \geq n$ ,

so that a matrix  $A \in \mathcal{A}_{n,m}$  takes one the following two forms:

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_m \\ a_2 & \dots & \dots & \dots \\ \vdots & \dots & \dots & \dots \\ a_m & \dots & \dots & a_n \\ \vdots & \dots & \dots & \dots \\ a_n & \dots & \dots & a_{n+m-1} \end{pmatrix} \quad \text{or} \quad A = \begin{pmatrix} a_1 & a_2 & \dots & a_n & \dots & a_m \\ a_2 & \dots & \dots & \dots & \dots & \dots \\ \vdots & \dots & \dots & \dots & \dots & \dots \\ a_n & \dots & \dots & a_m & \dots & a_{n+m-1} \end{pmatrix},$$

for  $a_1, \dots, a_{n+m-1} \in \mathbb{R}$ . We call  $A \in \mathcal{A}_{n,m}$  an *Hankel matrix* and write  $\mathcal{A}_n := \mathcal{A}_{n,n}$ .

We see that  $X_n(x) \in \mathcal{A}_{n+1}$ , and we can now define the inverse operator to  $E_{n,m}$  on  $\mathcal{A}_{n,m}$ .

**Theorem 3.3** (L-duplicating matrix). *For every  $n, m \geq 1$  and for every matrix  $A \in \mathcal{A}_{n,m}$ , there exists an L-duplicating matrix  $D_{n,m} \in \mathbb{R}^{nm \times (n+m-1)}$  such that*

$$D_{n,m} \text{vec}L(A) = \text{vec}(A) \quad (3.9)$$

$$D_{n,m} = \sum_{i=1}^n \sum_{j=1}^m \vec{e}_{n+m-1, i+j-1}^\top \otimes \vec{e}_{m,j} \otimes \vec{e}_{n,i}. \quad (3.10)$$

**Corollary 3.4.** *For every  $n \geq 1$ , the L-duplicating matrix  $D_{n+1} \in \mathbb{R}^{(n+1)^2 \times (2n+1)}$  transforming  $\text{vec}L(X_n(x))$  into  $\text{vec}(X_n(x))$  is given by  $D_{n+1} := D_{n+1, n+1}$ .*

*Example 3.3.* Let  $n = m = 2$ . Then equation (3.10) becomes

$$D_{2,2} = \sum_{i=1}^2 \sum_{j=1}^2 \vec{e}_{3, i+j-1}^\top \otimes \vec{e}_{2,j} \otimes \vec{e}_{2,i} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

For  $A \in \mathcal{A}_2$  of the form  $A = \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}$ , we can verify that

$$D_{2,2} \text{vec}L(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_2 \\ a_3 \end{pmatrix} = \text{vec}(A).$$

Moreover, when applied to  $\text{vec}L(X_1(x)) = (1, x, x^2)^\top$ , it duplicates the missing value  $x$ .

We conclude this section with an important property for the matrices  $E_{n,m}$  and  $D_{n,m}$ .

**Proposition 3.5.** *For every  $n, m \geq 1$ ,  $D_{n,m}$  is the right-inverse of  $E_{n,m}$  and, for every  $A \in \mathcal{A}_{n,m}$ , the product  $D_{n,m}E_{n,m} \in \mathbb{R}^{nm \times nm}$  acts on  $\text{vec}(A)$  like an identity operator:  $D_{n,m}E_{n,m} \text{vec}(A) = \text{vec}(A)$ .*

*Example 3.4.* Let  $n = m = 2$ . From Example 3.2 and 3.3 we get

$$D_{2,2}E_{2,2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \neq I_4.$$

However, for  $A \in \mathcal{A}_2$  with  $\text{vec}(A) = (a_1, a_2, a_2, a_3)^\top$ , we notice that  $D_{2,2}E_{2,2} \text{vec}(A) = \text{vec}(A)$ , hence the product  $D_{2,2}E_{2,2}$  behaves like an identity operator, despite not coinciding with the identity matrix.

## 3.2 The generator for correlators

We focus now on the original problem: by equation (3.6) we seek a linear operator  $\tilde{G}_n^{(1)}$  transforming  $\text{vec}(X_n(x))$  into  $\mathcal{G} \text{vec}(X_n(x))$ . From equation (3.2), we notice that the elements of  $X_n(x)$  lying on the left-bottom "L" coincide with the powers of  $x$  from 0 to  $2n$ .

**Lemma 3.6.** For every  $n \geq 1$ , the  $L$ -vectorization of  $X_n(x)$  coincides with the vectors basis of monomials of order  $2n$ , namely  $\text{vec}L(X_n(x)) = H_{2n}(x)$ .

Hence, by transforming  $\text{vec}(X_n(x))$  into  $\text{vec}L(X_n(x))$ , we address the problem of finding the generator matrix for  $\text{vec}(X_n(x))$  to the problem of finding the generator matrix for  $H_{2n}(x)$ , which was solved in Section 2.1. We can then prove the following result.

**Proposition 3.7.** For every  $t \geq 0$  and  $n \geq 1$ , the matrix  $\tilde{G}_n^{(1)}$  satisfying equation (3.5) and its matrix exponential are respectively given by

$$\tilde{G}_n^{(1)} = D_{n+1}G_{2n}E_{n+1} \quad \text{and} \quad e^{\tilde{G}_n^{(1)}t} = D_{n+1}e^{G_{2n}t}E_{n+1},$$

where  $G_{2n}$  is the generator matrix of order  $2n$ .

We are now able to provide a solution to the two-point correlator problem.

**Theorem 3.8** (Two-point correlator formula). The expression for the two-point correlator is given by

$$C_{p_0, p_1}(s_0, s_1; t) = \vec{p}_1^\top \left\{ \text{vec}^{-1} \circ D_{n+1}e^{G_{2n}(s_0-t)}E_{n+1} \circ \text{vec}(X_n(Y(t))) \right\} e^{G_n^\top(s_1-s_0)}\vec{p}_0,$$

with  $\vec{p}_0, \vec{p}_1 \in \mathbb{R}^{n+1}$  the vectors of coefficients for the polynomial functions  $p_0 \in \text{Pol}_{n_0}(\mathbb{R})$  and  $p_1 \in \text{Pol}_{n_1}(\mathbb{R})$ ,  $n = \max\{n_0, n_1\}$  and  $t < s_0 < s_1$ .

## 4 Higher-order correlators

We prove in this section a correlator formula holding for every  $m \geq 1$  by following similar steps to the ones performed in Section 3 for  $m = 1$ . We recall that we seek an explicit expression for

$$C_{p_0, \dots, p_m}(s_0, \dots, s_m; t) := \mathbb{E}[p_m(Y(s_0))p_{m-1}(Y(s_1)) \cdots p_0(Y(s_m)) | \mathcal{F}_t],$$

with  $p_k \in \text{Pol}_{n_k}(\mathbb{R})$ ,  $k = 0, \dots, m$ , and  $t < s_0 < s_1 < \dots < s_m$ . We start with the following operator.

**Definition 4.1** ( $d$ -Kronecker product). We define the  $d$ -Kronecker product of a matrix  $A \in \mathbb{R}^{n \times m}$  and a matrix  $B \in \mathbb{R}^{r \times s}$ , as the  $d$ -th Kronecker power of  $A$  multiplied in the Kronecker sense with  $B$ , for  $d \geq 1$ , or equal to  $B$ , for  $d = 0$ , namely

$$\begin{cases} A \otimes^d B = A^{\otimes d} \otimes B & d \geq 1 \\ A \otimes^0 B = B & d = 0 \end{cases}.$$

Then for  $n \geq 1$  and  $r \geq 0$ , we introduce the matrix of functions

$$X_n^{(r)}(x) := H_n(x)^\top \otimes^r H_n(x), \quad (4.1)$$

for which we can make the following considerations:

- for  $r = 0$ : we get  $X_n^{(0)}(x) = H_n(x)$ ;
- for  $r = 1$ : we get

$$X_n^{(1)}(x) = H_n(x)^\top \otimes H_n(x) = H_n(x)H_n(x)^\top = X_n(x) \in \mathcal{A}_{n+1}; \quad (4.2)$$

- for  $r = 2$ : by the associativity property of the Kronecker product

$$X_n^{(2)}(x) = H_n(x)^\top \otimes X_n^{(1)}(x) = \left( X_n^{(1)}(x) \vdots xX_n^{(1)}(x) \vdots x^2X_n^{(1)}(x) \vdots \cdots \vdots x^nX_n^{(1)}(x) \right)$$

is composed by  $n + 1$  blocks of the form  $B_{n,2}^{(k)} = x^{k-1}X_n^{(1)}(x) \in \mathcal{A}_{n+1}$ ,  $k = 1, \dots, (n + 1)$ ;

- for  $r = 3$ : we write  $X_n^{(3)}(x) = (H_n(x)^\top)^{\otimes 2} \otimes X_n^{(1)}(x)$ , where

$$(H_n(x)^\top)^{\otimes 2} = (1 \ x \dots \dots x^n \ ; \ x \ x^2 \dots \dots x^{n+1} \ ; \ \dots \ ; \ x^n \ x^{n+1} \dots \dots x^{2n}) \in \mathbb{R}^{(n+1)^2}$$

so that  $X_n^{(3)}(x)$  is composed by  $(n+1)^2$  blocks. For each block  $B_{n,3}^{(k)}$  there exists a  $j_k \in \{0, \dots, 2n\}$  such that  $B_{n,3}^{(k)} = x^{j_k} X_n^{(1)}(x)$  and  $B_{n,3}^{(k)} \in \mathcal{A}_{n+1}$ ,  $k = 1, \dots, (n+1)^2$ . The difference from the previous case is that, now, some of the blocks are repeated since in  $(H_n(x)^\top)^{\otimes 2}$  some powers of  $x$  are repeated.

Generalizing, we can state the following result.

**Proposition 4.1.** *For every  $n, r \geq 1$ ,  $X_n^{(r)}(x)$  is a rectangular block matrix in  $\mathbb{R}^{(n+1) \times (n+1)^r}$ , which is composed by  $(n+1)^{r-1}$  blocks  $B_{n,r}^{(k)}(x) \in \mathcal{A}_{n+1}$ , for which there exists an index  $j_k \in \{0, \dots, (r-1)n\}$  such that  $B_{n,r}^{(k)}(x) = x^{j_k} X_n^{(1)}(x)$ ,  $k = 1, \dots, (n+1)^{r-1}$ .*

One can also notice that  $X_n^{(r)}(x)$  contains all the powers of  $x$  from 0 to  $(r+1)n$ . Thus, after removing the redundant powers, we are left with  $H_{n(r+1)}(x)$ . Moreover, from Proposition 4.1,  $X_n^{(r)}(x)$  is a block matrix whose blocks belong to  $\mathcal{A}_{n+1}$ . However, the matrix itself does not belong to  $\mathcal{A}_{n+1, (n+1)^r}$ .

*Example 4.1.* Let  $n = r = 2$ . Then we get the following matrix

$$X_2^{(2)}(x) = \begin{pmatrix} 1 & x & x^2 & | & x & x^2 & x^3 & | & x^2 & x^3 & x^4 \\ x & x^2 & x^3 & | & x^2 & x^3 & x^4 & | & x^3 & x^4 & x^5 \\ x^2 & x^3 & x^4 & | & x^3 & x^4 & x^5 & | & x^4 & x^5 & x^6 \end{pmatrix}$$

whose blocks belong to  $\mathcal{A}_3$ , but  $X_2^{(2)}(x) \notin \mathcal{A}_{3,9}$  hence it is not a Hankel matrix.

This means, in particular, that we cannot use the L-eliminating and L-duplicating matrices as defined in Section 3. We need instead two new tailor-made operators, given in the following two propositions.

**Proposition 4.2.** *For  $n, m \geq 1$ , there exists an  $m$ -th L-eliminating matrix  $E_{n+1}^{(m)} \in \mathbb{R}^{(n(m+1)+1) \times (n+1)^{m+1}}$  such that  $E_{n+1}^{(m)} \text{vec}(X_n^{(m)}(x)) = H_{n(m+1)}(x)$ . In particular,  $E_{n+1}^{(m)}$  is given by the recursion formula*

$$\begin{cases} E_{n+1}^{(1)} = E_{n+1} & m = 1 \\ E_{n+1}^{(m)} = E_{nm+1, n+1} \left( I_{n+1} \otimes E_{n+1}^{(m-1)} \right) & m \geq 2 \end{cases}.$$

**Proposition 4.3.** *For  $n, m \geq 1$ , there exists an  $m$ -th L-duplicating matrix  $D_{n+1}^{(m)} \in \mathbb{R}^{(n+1)^{m+1} \times (n(m+1)+1)}$  such that  $D_{n+1}^{(m)} H_{n(m+1)}(x) = \text{vec}(X_n^{(m)}(x))$ . In particular,  $D_{n+1}^{(m)}$  is given by the recursion formula*

$$\begin{cases} D_{n+1}^{(1)} = D_{n+1} & m = 1 \\ D_{n+1}^{(m)} = \left( I_{n+1} \otimes D_{n+1}^{(m-1)} \right) D_{nm+1, n+1} & m \geq 2 \end{cases}.$$

We can also prove the following property for the matrices  $E_{n+1}^{(m)}$  and  $D_{n+1}^{(m)}$ .

**Proposition 4.4.** *For every  $n, m \geq 1$ , the matrix  $D_{n+1}^{(m)}$  is the right-inverse of  $E_{n+1}^{(m)}$ , and the product  $D_{n+1}^{(m)} E_{n+1}^{(m)}$  acts on  $\text{vec}(X_n^{(m)}(x))$  like an identity operator:  $D_{n+1}^{(m)} E_{n+1}^{(m)} \text{vec}(X_n^{(m)}(x)) = \text{vec}(X_n^{(m)}(x))$ .*

*Example 4.2.* Let  $n = m = 2$ . By Proposition 4.2, we find that  $E_3^{(1)} = E_3$  and

$$E_3^{(2)} = E_{5,3} \left( I_3 \otimes E_3^{(1)} \right) = E_{5,3} \begin{pmatrix} -E_3^{(1)} & | & \mathbf{0} & | & \mathbf{0} \\ \mathbf{0} & | & E_3^{(1)} & | & \mathbf{0} \\ \mathbf{0} & | & \mathbf{0} & | & E_3^{(1)} \end{pmatrix}.$$

To understand better, we rewrite  $\text{vec}(X_2^{(2)}(x))$  in Example 4.1 as follows:

$$\text{vec}(X_2^{(2)}(x)) = \text{vec} \left( \text{vec} \begin{pmatrix} \textcircled{1} & x & x^2 \\ \textcircled{x} & x^2 & x^3 \\ \textcircled{x^2} & x^3 & x^4 \end{pmatrix} \text{vec} \begin{pmatrix} \textcircled{x} & x^2 & x^3 \\ \textcircled{x^2} & x^3 & x^4 \\ \textcircled{x^3} & x^4 & x^5 \end{pmatrix} \text{vec} \begin{pmatrix} \textcircled{x^3} & x^3 & x^4 \\ \textcircled{x^3} & x^4 & x^5 \\ \textcircled{x^3} & x^5 & x^6 \end{pmatrix} \right),$$

so that, by applying  $I_3 \otimes E_3^{(1)}$ , we select from each of the three blocks of  $X_2^{(2)}(x)$  their L-vectorization (remember that the L-eliminating matrix acts on the vectorization of a matrix and returns the L-vectorization of the matrix itself), elements which are marked with a circle. We are left with

$$(I_3 \otimes E_3^{(1)}) \text{vec}(X_2^{(2)}(x)) = \text{vec} \begin{pmatrix} \textcircled{1} & x & x^2 \\ \textcircled{x} & x^2 & x^3 \\ \textcircled{x^2} & x^3 & x^4 \\ \textcircled{x^3} & x^4 & x^5 \\ \textcircled{x^4} & x^5 & x^6 \end{pmatrix}. \quad (4.3)$$

We notice that the elements we need are on the left-bottom "L" (the ones marked with a circle). Applying  $E_{5,3}$  gives  $H_6(x)$ . Moreover, the matrix on the right hand side of equation (4.3) belongs to  $\mathcal{A}_{5,3}$ . Then the corresponding L-duplicating matrix is given by Proposition 4.3 and is

$$D_3^{(2)} = (I_3 \otimes D_3^{(1)}) D_{5,3} = \begin{pmatrix} D_3^{(1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & D_3^{(1)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & D_3^{(1)} \end{pmatrix} D_{5,3}.$$

In particular,  $D_{5,3}$  acting on  $H_6(x)$  returns the matrix on the right hand side of equation (4.3) while  $D_3^{(1)}$ , acting singularly on each column because of the multiplication with  $I_3$  (namely,  $I_3 \otimes D_3^{(1)}$ ) returns  $X_2^{(2)}(x)$ , showing that  $D_3^{(2)}$  is the inverse operator of  $E_3^{(2)}$ .

We now derive the closed formula for  $(m+1)$ -point correlators.

**Theorem 4.5** (Correlator formula). *For every  $m \geq 1$ , let  $p_k \in \text{Pol}_{n_k}(\mathbb{R})$  with vector of coefficients  $\vec{p}_k \in \mathbb{R}^{n+1}$ ,  $k = 0, \dots, m$ ,  $n = \max\{n_0, \dots, n_m\}$  and  $t < s_0 < s_1 < \dots < s_m$ . There exist  $m+1$  matrices  $\tilde{G}_n^{(r)} \in \mathbb{R}^{(n+1)^{r+1} \times (n+1)^{r+1}}$ ,  $r = 0, \dots, m$ , such that*

$$C_{p_0, \dots, p_m}(s_0, \dots, s_m; t) = \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0-t)} \circ \text{vec} \left( X_n^{(m)}(Y(t)) \right) \right\} \prod_{k=1}^m e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \}$$

where  $\prod_{k=1}^m$  is the product obtained starting with the matrix corresponding to  $k=1$  and multiplying on the right by the following matrices until the matrix corresponding to  $k=m$ . In particular,  $\tilde{G}_n^{(r)} = D_{n+1}^{(r)} G_{n(r+1)} E_{n+1}^{(r)}$  and  $e^{\tilde{G}_n^{(r)}t} = D_{n+1}^{(r)} e^{G_{n(r+1)}t} E_{n+1}^{(r)}$ , with  $\tilde{G}_n^{(0)} = G_n$ .

## 5 Recursion formulas for the generator matrix

We focus in this section on the generator matrix  $G_n$  defined in Theorem 2.1. In particular, we provide a recursion formula for it and one for its matrix exponential. These formulas are referred to the basis vector of monomials, but they can also be generalized to a different polynomial basis vector. In this case, one needs the matrix for the change of basis.

### 5.1 The generator matrix

We provide a recursion formula for the generator matrix.

**Theorem 5.1** (Recursion formula for the generator matrix). *For  $n \geq 2$ , the generator matrix  $G_n \in \mathbb{R}^{(n+1) \times (n+1)}$  satisfying Theorem 2.1 with respect to the vector basis of monomials  $H_n(x)$  is given by*

$$G_n = \begin{pmatrix} G_{n-1} & \vec{0}_n \\ \vec{a}_n^\top & c_n \end{pmatrix} \quad \text{with} \quad G_1 = \begin{pmatrix} 0 & 0 \\ b_0 & b_1 \end{pmatrix}.$$

Here  $\vec{0}_n$  is a  $n$ -dimensional vector of 0's,  $\vec{a}_n = (a_n^n, a_n^{n-1}, \dots, a_n^1)^\top \in \mathbb{R}^n$  with

$$\begin{aligned} a_n^1 &= nb_0 + \frac{1}{2}n(n-1)\sigma_1 + \sum_{k=2}^n \binom{n}{k} \xi_{k-1}^k, & a_n^2 &= \frac{1}{2}n(n-1)\sigma_0 + \sum_{k=2}^n \binom{n}{k} \xi_{k-2}^k, \\ a_n^i &= \sum_{k=i}^n \binom{n}{k} \xi_{k-i}^k \quad \text{for } i = 3, \dots, n, & \text{and } c_n &= nb_1 + \frac{1}{2}n(n-1)\sigma_2 + \sum_{k=2}^n \binom{n}{k} \xi_k^k. \end{aligned} \quad (5.1)$$

*Remark 5.1.* From Theorem [5.1](#) we notice that for  $n \geq 1$  the generator matrix  $G_n$  is lower triangular. Moreover, for  $n \geq 2$  the main diagonal of  $G_n$  is of the form

$$\text{diag}(G_n) = (0, b_1, c_2, c_3, \dots, c_n)^\top, \quad (5.2)$$

so that, in particular, the matrix  $G_n$  is not invertible.

**Lemma 5.2.** *If  $\ell(x, dz) \equiv 0$  on  $\mathbb{R}$ , then  $G_n$  is a (lower) tri-diagonal matrix.*

We now provide a recursion formula for the matrix exponential of  $G_n$ .

**Theorem 5.3** (Recursion formula for the exponential generator matrix). *For a fixed  $n \geq 2$ , if the following conditions are satisfied*

$$\begin{cases} c_j \neq 0 & \text{for every } 2 \leq j \leq n \\ c_j \neq c_i & \text{for every } 1 \leq j < i \leq n \end{cases} \quad (5.3)$$

then the recursion formula holds:

$$e^{G_n t} = \begin{pmatrix} e^{G_{n-1} t} & \vec{0}_n \\ \vec{a}_n^\top \Lambda_n^{-1} (e^{c_n t} I_n - e^{G_{n-1} t}) & e^{c_n t} \end{pmatrix} \quad \text{with}$$

$$e^{G_1 t} = \begin{pmatrix} 1 & 0 \\ \frac{b_0}{b_1} (e^{b_1 t} - 1) & e^{b_1 t} \end{pmatrix} \quad \text{if } b_1 \neq 0 \quad \text{and} \quad e^{G_1 t} = \begin{pmatrix} 1 & 0 \\ b_0 t & 1 \end{pmatrix} \quad \text{if } b_1 = 0.$$

**Lemma 5.4.** *If  $\ell(x, dz) \equiv 0$  on  $\mathbb{R}$ , then condition [\(5.3\)](#) is equivalent to  $b_1 \neq -\frac{k}{2}\sigma_2$  for every  $1 \leq k \leq 2(n-1)$ . In particular, the coefficients  $b_1$  and  $\sigma_2$  cannot be simultaneously equal to 0.*

## 5.2 Change of basis

The vector basis of monomials  $H_n(x)$  is intuitive and allows to write down computations easily and explicitly. However, when it comes to applications, one chooses in practice an orthogonal basis, such as, for example, the Hermite polynomials or the Legendre polynomials. Combining the properties of an orthogonal basis with the properties of polynomial processes leads to improvements, e.g., in option pricing [\[3\]](#) [\[18\]](#) [\[35\]](#). Motivated by this fact, we present a result which allows to obtain the generator matrix and its matrix exponential with respect to any basis of polynomials.

For  $n \geq 1$ , we consider a set of polynomial functions  $\{q_0(x), q_1(x), \dots, q_n(x)\}$  with values in  $\mathbb{R}$  which forms a basis for  $\text{Pol}_n(\mathbb{R})$ . We then introduce the vector valued function

$$Q_n : \mathbb{R} \longrightarrow \mathbb{R}^{n+1}, \quad Q_n(x) = (q_0(x), q_1(x), \dots, q_n(x))^\top.$$

From classical linear algebra, there exists an invertible matrix  $M_n \in \mathbb{R}^{(n+1) \times (n+1)}$  such that

$$M_n H_n(x) = Q_n(x) \quad \text{and} \quad M_n^{-1} Q_n(x) = H_n(x). \quad (5.4)$$

We further indicate with  $J_n \in \mathbb{R}^{(n+1) \times (n+1)}$  the generator matrix in the sense of Theorem [2.1](#) with respect to the basis vector  $Q_n(x)$ , namely such that

$$\mathcal{G}Q_n(x) = J_n Q_n(x). \quad (5.5)$$

We can then prove the following result concerning  $J_n$ .

**Proposition 5.5.** *For every  $n \geq 1$  and  $t \geq 0$ , the generator matrix  $J_n$  and its matrix exponential are given by the following matrix products*

$$J_n = M_n G_n M_n^{-1} \quad \text{and} \quad e^{J_n t} = M_n e^{G_n t} M_n^{-1}.$$

*Example 5.1.* Let  $Q_n(x)$  be the vector basis given by Hermite polynomials. For  $n = 4$  we get

$$Q_4(x) = (1, x, x^2 - 1, x^3 - 3x, x^4 - 6x^2 + 3)$$

while  $M_4$  and  $M_4^{-1}$  are respectively given by

$$M_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 \\ 3 & 0 & -6 & 0 & 1 \end{pmatrix} \quad \text{and} \quad M_4^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 1 & 0 \\ 3 & 0 & 6 & 0 & 1 \end{pmatrix}.$$

We consider  $\ell(x, dz) \equiv 0$ . By direct computation, one finds that

$$G_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ b_0 & b_1 & 0 & 0 & 0 \\ \sigma_0 & 2b_0 + \sigma_1 & 2b_1 + \sigma_2 & 0 & 0 \\ 0 & 3\sigma_0 & 3(b_0 + \sigma_1) & 3(b_1 + \sigma_2) & 0 \\ 0 & 0 & 6\sigma_0 & 2(2b_0 + 3\sigma_1) & 2(2b_1 + 3\sigma_2) \end{pmatrix} \quad \text{and}$$

$$J_4 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ b_0 & b_1 & 0 & 0 & 0 \\ \sigma_0 + 2b_1 + \sigma_2 & 2b_0 + \sigma_1 & 2b_1 + \sigma_2 & 0 & 0 \\ 3\sigma_1 & 3(2b_1 + 3\sigma_2 + \sigma_0) & 3(b_0 + \sigma_1) & 3(b_1 + \sigma_2) & 0 \\ 12\sigma_2 & 12\sigma_1 & 6(2b_1 + \sigma_0 + 5\sigma_2) & 2(2b_0 + 3\sigma_1) & 2(2b_1 + 3\sigma_2) \end{pmatrix}.$$

By matrix multiplication Proposition [5.5](#) can be verified.

## 6 Applications and numerical aspects

Having a closed and compact formula for correlators like in Theorem [4.5](#) is attractive in sensitivity analysis and risk management. For example, in applications to options, the Greeks play an important role in hedging. The Greeks for options are defined as the derivatives of the price functional with respect to various parameters. In the context of path-dependent options introduced in Section [1.1.1](#) we shall derive in this section the expression for two of the Greeks, namely the Delta and the Theta. We shall then analyse our correlator formula from a numerical point of view, also in relation with the iterated moment formula and with a Monte Carlo approach.

### 6.1 Computation of Greeks

Two of the most common Greeks are the Delta and the Theta. The first measures the change in the option price with respect to a change in the underlying asset price. The second measures the sensitivity of the option to time to exercise.

From Section [1.1.1](#), the price of an Asian option can be approximated by a linear combination of conditional expectations of the form  $C^{k_0, \dots, k_m}(s_0, \dots, s_m; t) := \mathbb{E}[Y(s_0)^{k_0} Y(s_1)^{k_1} \dots Y(s_m)^{k_m} | \mathcal{F}_t]$ , corresponding to correlators  $C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)$  with  $p_j(x) = x^{k_j}$ ,  $j = 0, \dots, m$ . To calculate the Delta of the Asian option, which is the partial derivative of the price functional  $\Pi(t)$  with respect to the initial condition  $Y(t)$ , we then need the derivative of  $C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)$  with respect to  $Y(t)$ .

**Proposition 6.1.** *For every  $m \geq 1$ , in the same notation of Theorem [4.5](#) we have that*

$$\frac{\partial C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)}{\partial Y(t)} = \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0 - t)} \circ \text{vec} \left( \frac{\partial X_n^{(m)}(Y(t))}{\partial Y(t)} \right) \right\} \prod_{k=1}^m e^{\tilde{G}_n^{(m-k)\top}(s_k - s_{k-1})} \{I_{n+1} \otimes^{m-k} \vec{p}_{m-k}\}$$

where

$$\frac{\partial X_n^{(m)}(Y(t))}{\partial Y(t)} = \left( \frac{\partial H_n(Y(t))}{\partial Y(t)} \right)^\top \otimes X_n^{(m-1)}(Y(t)) + H_n(Y(t))^\top \otimes \frac{\partial X_n^{(m-1)}(Y(t))}{\partial Y(t)}$$

with  $\frac{\partial H_n(Y(t))}{\partial Y(t)} = \frac{\partial X_n^{(0)}(Y(t))}{\partial Y(t)} = (0, 1, \dots, n) (0, H_{n-1}(Y(t))^\top)^\top$ .

Similarly, to compute the Theta of the Asian option, we need first to compute the derivative of  $C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)$  with respect to the  $m + 1$  time points involved, namely  $s_0 < s_1 < \dots < s_m$ .

**Proposition 6.2.** For every  $m \geq 1$ , in the same notation of Theorem [4.5](#) we have that

$$\begin{aligned}
\Theta_0 &= \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0-t)} \circ \text{vec} \left( X_n^{(m)}(Y(t)) \right) \right\} \prod_{k=1}^m e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \} + \\
&\quad - \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0-t)} \circ \text{vec} \left( X_n^{(m)}(Y(t)) \right) \right\} \tilde{G}_n^{(m-1)\top} \prod_{k=1}^m e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \}; \\
\Theta_j &= \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0-t)} \circ \text{vec} \left( X_n^{(m)}(Y(t)) \right) \right\} \cdot \\
&\quad \cdot \left( \prod_{k=1}^{j-1} e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \} \tilde{G}_n^{(m-j)\top} e^{\tilde{G}_n^{(m-j)\top}(s_j-s_{j-1})} \{ I_{n+1} \otimes^{m-j} \vec{p}_{m-j} \} \cdot \right. \\
&\quad \cdot \prod_{k=j+1}^m e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \} + \\
&\quad \left. - \prod_{k=1}^j e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \} \tilde{G}_n^{(m-j-1)\top} e^{\tilde{G}_n^{(m-j-1)\top}(s_{j+1}-s_j)} \{ I_{n+1} \otimes^{m-j-1} \vec{p}_{m-j-1} \} \cdot \right. \\
&\quad \left. \cdot \prod_{k=j+2}^m e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \} \right) \quad \text{for } 1 \leq j < m; \\
\Theta_m &= \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0-t)} \circ \text{vec} \left( X_n^{(m)}(Y(t)) \right) \right\} \cdot \\
&\quad \cdot \prod_{k=1}^{m-1} e^{\tilde{G}_n^{(m-k)\top}(s_k-s_{k-1})} \{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \} \tilde{G}_n^{(0)\top} e^{\tilde{G}_n^{(0)\top}(s_m-s_{m-1})} \{ I_{n+1} \otimes^0 \vec{p}_0 \},
\end{aligned}$$

where we have introduced the compact notation  $\Theta_j := \frac{\partial C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)}{\partial s_j}$  for  $0 \leq j \leq m$ .

Then, the Delta of the Asian option is obtained by Proposition [6.1](#)

$$\frac{\partial \Pi(t)}{\partial Y(t)} \approx e^{-r(T-t)} \sum_{\mathbf{k}} \alpha_{\mathbf{k}} \frac{\partial C^{k_0, \dots, k_m}(s_0, \dots, s_m; t)}{\partial Y(t)}$$

and, similarly, the Theta is obtained by Proposition [6.2](#)

$$\frac{\partial \Pi(t)}{\partial s_j} \approx e^{-r(T-t)} \sum_{\mathbf{k}} \alpha_{\mathbf{k}} \frac{\partial C^{k_0, \dots, k_m}(s_0, \dots, s_m; t)}{\partial s_j} \quad \text{for } 0 \leq j \leq m,$$

for certain coefficients  $\{\alpha_{\mathbf{k}}\}_{\mathbf{k}}$  and  $\mathbf{k} = (k_1, \dots, k_m)$  a multi-index. A more detailed analysis for Greeks of discrete average arithmetic Asian options can be found in [28](#), where the coefficients  $\{\alpha_{\mathbf{k}}\}_{\mathbf{k}}$  are computed explicitly with respect to the basis of Hermite polynomials.

## 6.2 Numerical performances

We analyse numerically the correlator formula in Theorem [4.5](#) which gives explicitly the value for correlators up to the computation of the exponential of the  $m+1$  matrices  $\tilde{G}_n^{(r)}$ ,  $r = 0, \dots, m$ . This means that, assuming these matrix exponentials to be exact, the correlator formula provides *the* correlator value. The same exact value can be also obtained by applying iteratively the moment formula, as we discussed in Section [1](#). We then compare from a time cost point of view these two procedures. In particular, we consider both an implementation with dense matrices and an implementation with sparse matrices. Finally, we consider a Monte Carlo approach which is compared with our correlator formula both from a time cost and an accuracy point of views.

For the experiments, we consider an Ornstein–Uhlenbeck process  $Y$  defined by

$$dY(t) = (b_0 + b_1 Y(t)) dt + \sqrt{\sigma_0} dB(t) \tag{6.1}$$



and with model specifications

$$\begin{cases} b_0 = +0.75 \\ b_1 = -5.00 \\ \sigma_0 = +0.01 \\ Y(t=0) = +0.15 \end{cases}$$

which corresponds to a polynomial diffusion process with  $\sigma_1 = \sigma_2 = 0$  and  $\ell(x, dz) \equiv 0$  (see equation (2.3)). Moreover, for  $n \geq 1$ , we consider a particular case of equation (1.1) of the form

$$C^n(s_0, \dots, s_m; t) := \mathbb{E}[Y(s_0)^n Y(s_1)^n \cdots Y(s_m)^n | \mathcal{F}_t]$$

corresponding to  $p_k(x) = x^n = \bar{e}_{n+1, n+1}^\top H_n(x)$ ,  $k = 0, \dots, m$ . We found terms of this form in Section 1.1.1 when motivating the study of correlators for pricing a financial derivative in a stochastic volatility model context. The two cases indeed coincide for  $n = 1$ . We also mention that Ornstein–Uhlenbeck processes are common models in finance. Among others, we find examples in modelling the electricity spot price, a non-Gaussian example being treated in [6].

The Monte Carlo simulations are based on the conditional solution of equation (6.1) given by

$$Y(s) = Y(t) e^{b_1(s-t)} + \frac{b_0}{b_1} \left( e^{b_1(s-t)} - 1 \right) + \sqrt{\sigma_0} \int_t^s e^{b_1(s-v)} dB(v), \quad \text{for } s \geq t \geq 0.$$

We define  $\Delta_k := s_k - s_{k-1}$ ,  $k = 0, \dots, m$ , with  $s_{-1} := t$ , and rewrite it for  $s = s_k$  and  $t = s_{k-1}$ , namely

$$Y(s_k) = Y(s_{k-1}) e^{b_1 \Delta_k} + \frac{b_0}{b_1} \left( e^{b_1 \Delta_k} - 1 \right) + \sqrt{\sigma_0} \int_{s_{k-1}}^{s_k} e^{b_1(s_k-v)} dB(v). \quad (6.2)$$

For a fixed number of time points  $m+1 \geq 1$ , the idea is then to sequentially simulate samples from  $Y(s_k)$  according to (6.2), using  $Y(s_{k-1})$  as starting point and the fact that, by the tower rule, it holds

$$C^m(s_0, \dots, s_m; t) = \mathbb{E} \left[ Y(s_0)^n \mathbb{E} \left[ Y(s_1)^n \mathbb{E} \left[ Y(s_2)^n \cdots \mathbb{E} \left[ Y(s_m)^n | \mathcal{F}_{s_{m-1}} \right] \cdots | \mathcal{F}_{s_1} \right] | \mathcal{F}_{s_0} \right] | \mathcal{F}_t \right].$$

When increasing the complexity of the problem, namely, increasing  $n$  and/or  $m$ , the Monte Carlo approach needs more simulations to gain accuracy, requiring also more time for computations. However, in order to compare different experiments, we fix the number of simulations to  $N = 10^4$ , each of which is repeated  $10^2$  times so that to get multiple values. Among these values, we select the worst one in terms of highest relative error with respect to the correlator formula, and we use it as representative of the set. For the time cost assessment, the representative is obtained by averaging the time cost required for the  $10^2$  simulations. Finally, we set a tolerance to  $1 \cdot 10^{-3}$  and claim that Monte Carlo *fails* if the relative error is bigger than that, counting the number of failures over the  $10^2$  simulations.

In Table 1 we report the outcomes of the time-cost experiments. Here we compare the correlator formula, both with dense (Dense) and sparse (Sparse) matrices, with the iterative application of the moment formula, both with dense (Iter. dense) and sparse (Iter. sparse) matrices, and with the Monte Carlo approach (MC average). We notice that the correlator formula with dense matrices is almost comparable with the iteration with dense matrices, and the same holds with sparse matrices. However, for higher complexities, the correlator formula tends to be a bit slower. Despite what one might expect, using sparse matrices makes both the approaches slower. The main reason lies in the fact that the matrix exponential of a sparse matrix is likely not to be sparse. Hence, using a sparse matrix for a dense matrix slows down computations. However, we stress the fact that sparse matrices are crucial when increasing further the complexity of the problem: for a given  $n$  and  $m$ , the generator matrix has dimension  $(n(m+1)+1) \times (n(m+1)+1)$ , so that it is not feasible to store it with a dense matrix.

We observe from Table 1 that the time cost for these four experiments increases when the complexity of the problem increases, as one would expect due to the dimensions of the matrices involved. On the other hand, for a fixed  $m$ , the time cost for the Monte Carlo approach (MC average) is almost invariant reflecting the fact that the number of time points is fixed. However, to keep the approach as general as possible, instead of computing the power  $x^n$  directly we do the vector multiplication  $\bar{e}_{n+1, n+1}^\top H_n(x)$  to reflect the situation that we would get if considering general polynomial function instead of monomials. This explains why, for a fixed  $m$ , when increasing the power  $n$  the time cost also increases slightly.

## Time-cost performances

$m$	$n$	Dense	Sparse	Iter. dense	Iter. sparse	MC average
0	1	$3.372e^{-04}$	$5.254e^{-03} (\approx 16x)$	$3.051e^{-04} (\approx 1x)$	$4.904e^{-03} (\approx 15x)$	$1.212e^{-01} (\approx 359x)$
0	2	$3.451e^{-04}$	$6.700e^{-03} (\approx 19x)$	$3.384e^{-04} (\approx 1x)$	$6.199e^{-03} (\approx 18x)$	$1.233e^{-01} (\approx 357x)$
0	3	$3.612e^{-04}$	$7.879e^{-03} (\approx 22x)$	$3.340e^{-04} (\approx 1x)$	$7.459e^{-03} (\approx 21x)$	$1.281e^{-01} (\approx 355x)$
0	4	$3.819e^{-04}$	$9.242e^{-03} (\approx 24x)$	$3.747e^{-04} (\approx 1x)$	$8.932e^{-03} (\approx 23x)$	$1.317e^{-01} (\approx 345x)$
0	5	$4.053e^{-04}$	$1.072e^{-02} (\approx 26x)$	$3.919e^{-04} (\approx 1x)$	$1.028e^{-02} (\approx 25x)$	$1.360e^{-01} (\approx 336x)$
0	10	$4.916e^{-04}$	$1.676e^{-02} (\approx 34x)$	$4.883e^{-04} (\approx 1x)$	$1.651e^{-02} (\approx 34x)$	$1.570e^{-01} (\approx 319x)$
1	1	$6.383e^{-04}$	$1.424e^{-02} (\approx 22x)$	$6.296e^{-04} (\approx 1x)$	$1.246e^{-02} (\approx 20x)$	$2.186e^{-01} (\approx 342x)$
1	2	$7.363e^{-04}$	$1.952e^{-02} (\approx 27x)$	$7.058e^{-04} (\approx 1x)$	$1.824e^{-02} (\approx 25x)$	$2.277e^{-01} (\approx 309x)$
1	3	$7.926e^{-04}$	$2.449e^{-02} (\approx 31x)$	$8.393e^{-04} (\approx 1x)$	$2.451e^{-02} (\approx 31x)$	$2.371e^{-01} (\approx 299x)$
1	4	$8.737e^{-04}$	$2.993e^{-02} (\approx 34x)$	$8.893e^{-04} (\approx 1x)$	$3.064e^{-02} (\approx 35x)$	$2.489e^{-01} (\approx 285x)$
1	5	$9.278e^{-04}$	$3.428e^{-02} (\approx 37x)$	$9.001e^{-04} (\approx 1x)$	$3.759e^{-02} (\approx 41x)$	$2.572e^{-01} (\approx 277x)$
1	10	$1.354e^{-03}$	$6.013e^{-02} (\approx 44x)$	$1.975e^{-03} (\approx 1x)$	$7.980e^{-02} (\approx 59x)$	$3.002e^{-01} (\approx 222x)$
2	1	$1.064e^{-03}$	$2.549e^{-02} (\approx 24x)$	$9.527e^{-04} (\approx 1x)$	$2.105e^{-02} (\approx 20x)$	$3.213e^{-01} (\approx 302x)$
2	2	$1.226e^{-03}$	$3.622e^{-02} (\approx 30x)$	$1.116e^{-03} (\approx 1x)$	$3.424e^{-02} (\approx 28x)$	$3.319e^{-01} (\approx 271x)$
2	3	$1.385e^{-03}$	$4.746e^{-02} (\approx 34x)$	$1.122e^{-03} (\approx 1x)$	$4.756e^{-02} (\approx 34x)$	$3.488e^{-01} (\approx 252x)$
2	4	$1.588e^{-03}$	$5.943e^{-02} (\approx 37x)$	$1.377e^{-03} (\approx 1x)$	$6.340e^{-02} (\approx 40x)$	$3.658e^{-01} (\approx 230x)$
2	5	$1.787e^{-03}$	$7.125e^{-02} (\approx 40x)$	$1.493e^{-03} (\approx 1x)$	$7.975e^{-02} (\approx 45x)$	$3.727e^{-01} (\approx 209x)$
2	10	$7.786e^{-03}$	$1.283e^{-01} (\approx 16x)$	$3.979e^{-03} (\approx 1x)$	$1.870e^{-01} (\approx 24x)$	$4.380e^{-01} (\approx 56x)$
3	1	$1.611e^{-03}$	$3.930e^{-02} (\approx 24x)$	$1.238e^{-03} (\approx 1x)$	$3.261e^{-02} (\approx 20x)$	$4.216e^{-01} (\approx 262x)$
4	1	$2.101e^{-03}$	$5.545e^{-02} (\approx 26x)$	$1.474e^{-03} (\approx 1x)$	$4.468e^{-02} (\approx 21x)$	$5.259e^{-01} (\approx 250x)$
5	1	$2.910e^{-03}$	$7.554e^{-02} (\approx 26x)$	$1.887e^{-03} (\approx 1x)$	$6.005e^{-02} (\approx 21x)$	$6.175e^{-01} (\approx 212x)$
10	1	$6.423e^{-02}$	$2.302e^{-01} (\approx 4x)$	$4.005e^{-03} (\approx 0x)$	$1.634e^{-01} (\approx 3x)$	$1.112e^{+00} (\approx 17x)$

Table 1: Time-cost performances for the correlator formula with dense (**Dense**) and sparse (**Sparse**) matrices, for the iterative application of the moment formula with dense (**Iter. dense**) and sparse (**Iter. sparse**) matrices, and for the the Monte Carlo approach (**MC average**). In parenthesis, how many times the different approaches are slower than the correlator formula with dense matrices.

In Table 2 we report the outcomes from the accuracy tests between the correlator formula and the Monte Carlo approach. We mention indeed that the four experiments previously discussed lead to the exact same values, as one would expect since they are basically four different representations for the same exact computation. Comparing the values from the correlator formula (Formula value) with the worst over the  $10^2$  outcomes of the Monte Carlo simulations (MC worst value), we notice that, when increasing the complexity, the second method becomes worse and worse in terms of relative error (in parenthesis). Of course, by increasing the number  $N$  of simulations, one can aim at better results, but this would also mean higher time costs. In the last column (MC fails) we also notice that the number of failures in terms of the tolerance defined above increases up to almost all the calls of the method.

*Remark 6.1.* As observed above, the correlator  $C^n(s_0, \dots, s_m; t)$  coincides for  $n = 1$  with the terms encountered in Section 1.1.1 in the setting of a stochastic volatility model. We then chose the parameters for the model in equation (6.1) so that to make it relevant in view of that application. However, the choice of a Gaussian Ornstein–Uhlenbeck process is for simplicity and illustration, and not intended as a precise stochastic volatility model. We also point out that in equation (1.6) one has in practise to truncate the summation to a certain  $\bar{k} \in \mathbb{N}$  to get an approximated option price. From Table 2, we notice that for  $n = 1$  the correlator values are very small in this configuration. This means that  $k$  can be chosen small, possibly not bigger than 5. For small values of  $k$  (that means, small values of  $m$ ), the iterated integral of the correlator formula (see the terms in equation (1.6)) can be computed by hand, avoiding additional error due to some multidimensional integration technique necessary otherwise.

## Accuracy performances

$m$	$n$	Formula value	MC worst value		MC fails
0	1	1.500e <sup>-01</sup>	1.490e <sup>-01</sup>	(0.66%)	66/100
0	2	2.350e <sup>-02</sup>	2.378e <sup>-02</sup>	(1.18%)	77/100
0	3	3.825e <sup>-03</sup>	3.751e <sup>-03</sup>	(1.93%)	91/100
0	4	6.442e <sup>-04</sup>	6.567e <sup>-04</sup>	(1.94%)	93/100
0	5	1.119e <sup>-04</sup>	1.092e <sup>-04</sup>	(2.43%)	92/100
0	10	2.618e <sup>-08</sup>	2.828e <sup>-08</sup>	(8.02%)	98/100
1	1	2.258e <sup>-02</sup>	2.278e <sup>-02</sup>	(0.88%)	79/100
1	2	5.594e <sup>-04</sup>	5.505e <sup>-04</sup>	(1.60%)	94/100
1	3	1.503e <sup>-05</sup>	1.460e <sup>-05</sup>	(2.82%)	91/100
1	4	4.338e <sup>-07</sup>	4.164e <sup>-07</sup>	(4.00%)	95/100
1	5	1.337e <sup>-08</sup>	1.268e <sup>-08</sup>	(5.09%)	97/100
1	10	8.366e <sup>-16</sup>	6.916e <sup>-16</sup>	(17.3%)	99/100
2	1	3.436e <sup>-03</sup>	3.467e <sup>-03</sup>	(0.92%)	81/100
2	2	1.383e <sup>-05</sup>	1.346e <sup>-05</sup>	(2.62%)	92/100
2	3	6.372e <sup>-08</sup>	6.662e <sup>-08</sup>	(4.55%)	95/100
2	4	3.310e <sup>-10</sup>	3.529e <sup>-10</sup>	(6.60%)	98/100
2	5	1.914e <sup>-12</sup>	1.742e <sup>-12</sup>	(9.00%)	99/100
2	10	4.638e <sup>-23</sup>	8.030e <sup>-23</sup>	(73.2%)	100/100
3	1	5.291e <sup>-04</sup>	5.352e <sup>-04</sup>	(1.15%)	86/100
4	1	8.245e <sup>-05</sup>	8.097e <sup>-05</sup>	(1.80%)	87/100
5	1	1.300e <sup>-05</sup>	1.326e <sup>-05</sup>	(2.02%)	88/100
10	1	1.477e <sup>-09</sup>	1.413e <sup>-09</sup>	(4.36%)	91/100

Table 2: Accuracy performances for the correlator formula (**Formula value**) compared with a Monte Carlo approach (**MC worst value**). In parenthesis, the relative error with respect to the correlator formula. In the last column (**MC fails**), the ratio of values with relative error exceeding the tolerance.

## 7 Conclusions

We have derived an explicit formula for computing correlators of **one-dimensional** polynomial processes consisting of linear combinations of exponentials of the generator matrix associated with the polynomial process. Our analysis is based on a recursive use of the moment formula for conditional expectations of polynomial processes along with the introduction of two new linear operators, respectively, the L-eliminating and the L-duplicating matrices. The closed-form expression of the correlators of polynomial processes is attractive in studies of options and risk management. The connection to Hankel matrices open our studies of correlators of polynomial processes to other areas as well.

We want to stress that a closed formula allows to make analysis with respect to the variables and parameters involved. The correlator formula in Theorem 4.5 depends on the polynomial jump-diffusion coefficients,  $b$ ,  $\sigma$  and  $\xi$ , as introduced in equation (2.3). But it also depends on the time points  $t < s_0 < s_1 < \dots < s_m < T < +\infty$ . This fact may be exploited for option price analysis as we demonstrated. Indeed, the correlator formula allows for explicit computation of derivatives, and lend themselves to the calculation of option Greeks.

We finally highlight that the formulas obtained in this paper are restricted to **one-dimensional** models. When increasing the dimension of the problem, calculations get more challenging and further considerations are required. Indeed, for any  $d \geq 2$ , the monomial basis for  $\mathbb{R}^d$  is more complex than the one used here, and, when multiplying two or more basis vectors, the object obtained requires more analysis and potentially different eliminating and duplicating matrices.

**Acknowledgements** Christa Cuchiero is thanked for interesting discussions. The authors are also grateful for the careful reading and the suggestions of three anonymous referees, which have led to a significant improvement on the presentation of the paper.

## A Some combinatorial properties

We report in this section some combinatorial properties concerning the L-eliminating and the L-duplicating matrices and the matrix of functions  $X_n^{(r)}(x)$ . However, these are not necessary to understand the main part of the article.

### A.1 The L-eliminating and L-duplicating matrices

The total number of elements of  $E_{n,m}$  is equal to  $nm(n+m-1)$ . When multiplying  $E_{n,m}$  with a matrix  $A \in \mathbb{R}^{n \times m}$ , we select the elements of  $A$  which are in the first column and last row, that account for exactly  $n+m-1$  terms. That means that  $E_{n,m}$  must have exactly  $n+m-1$  elements equal to 1 and the rest must be zeros. Moreover, for  $A \in \mathcal{A}_{n,m}$ , the matrix  $D_{n,m}$  duplicates each element  $a_k$  in  $\text{vec}L(A)$  as many times as the number of elements in the  $k$ -th skew-diagonal of  $A$ ,  $k = 1, \dots, n+m-1$ . Then in the  $k$ -th column of  $D_{n,m}$  there are as many 1's as the number of elements in the  $k$ -th skew-diagonal of  $A$ , while the remaining elements are all zeros. We shall be more precise in the following lemmas.

**Lemma A.1.** *For  $E_{n,m} \in \mathbb{R}^{(n+m-1) \times nm}$ , the number of non-zero elements is  $n+m-1$ .*

**Lemma A.2.** *If  $n \neq m$ , then the number of 1's in the  $k$ -th column of  $D_{n,m}$  is*

$$\begin{cases} k & \text{for } 1 \leq k \leq \min(n, m) - 1 \\ \min(n, m) & \text{for } \min(n, m) \leq k \leq \max(n, m) - 1 \\ n + m - k & \text{for } \max(n, m) \leq k \leq n + m - 1 \end{cases} \quad (\text{A.1})$$

If  $n = m$ , then the following formula holds instead:

$$\begin{cases} k & \text{for } 1 \leq k \leq n - 1 \\ 2n - k & \text{for } n \leq k \leq 2n - 1 \end{cases} \quad (\text{A.2})$$

In particular, if  $n = m$  then the number of 1's in the  $k$ -th column corresponds to the coefficient of the  $(k-1)$ -th power of  $x$  in the power expansion  $\left(\sum_{\alpha=0}^{n-1} x^\alpha\right)^2$ .

*Proof.* Let  $n \neq m$  and  $A \in \mathcal{A}_{n,m}$ . In the  $k$ -th column of  $D_{n,m}$  there are as many 1's as the number of elements in the  $k$ -th skew-diagonal of  $A$ . Denoting with  $a_k$  the value of the elements on the  $k$ -th skew-diagonal of  $A$ ,  $k = 1, \dots, n+m-1$ , equation (A.1) gives the cardinality of each  $a_k$ . Then the sum of the elements in equation (A.1) should give the total number of elements in  $A$ , that is  $nm$ :

$$\begin{aligned} & \sum_{k=1}^{\min(n,m)-1} k + \sum_{k=\min(n,m)}^{\max(n,m)-1} \min(n, m) + \sum_{k=\max(n,m)}^{n+m-1} (m+n-k) \\ &= \frac{\min(n, m) (\min(n, m) - 1)}{2} + \min(n, m) (\max(n, m) - \min(n, m)) + \frac{\min(n, m) (\min(n, m) + 1)}{2} \\ &= \max(n, m) \min(n, m) = nm, \end{aligned}$$

where for the third sum we used the change of variables  $k' = m+n-k$  so that  $\sum_{k=\max(n,m)}^{n+m-1} (m+n-k) = \sum_{k'=1}^{n+m-\max(n,m)} k'$ , and the fact that  $m+n-\max(n,m) = \min(n,m)$ . The case  $n = m$  is similar, so that the first part of the lemma is proved.

We need now to prove that the numbers in equation (A.2) correspond to the coefficients in the power expansion  $\left(\sum_{\alpha=0}^{n-1} x^\alpha\right)^2$ . We proceed by induction on the matrix dimension  $n \geq 2$  ( $n = 1$  is trivial).

- $n = 2$ : a matrix  $A \in \mathcal{A}_{2,2}$  is of the form  $A = \begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix}$  and the cardinality of  $a_k$ ,  $k = 1, 2, 3$ , is  $1-2-1$ , which corresponds to the coefficients of the polynomial  $(\sum_{\alpha=0}^1 x^\alpha)^2 = (1+x)^2 = 1+2x+x^2$ .
- $n \rightarrow n+1$ : we denote with  $A_n$  a general matrix in  $\mathcal{A}_n$ , and with  $A_{n+1}$  a general matrix in  $\mathcal{A}_{n+1}$ . Then  $A_n$  and  $A_{n+1}$  can be represented as

$$A_n = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ a_2 & a_3 & \cdots & a_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_n & a_{n+1} & \cdots & a_{2n-1} \end{pmatrix} \quad \text{and} \quad A_{n+1} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n & a_{n+1} \\ a_2 & a_3 & \cdots & a_{n+1} & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_n & a_{n+1} & \cdots & a_{2n-1} & a_{2n} \\ a_{n+1} & \cdots & a_{2n-1} & a_{2n} & a_{2n+1} \end{pmatrix}.$$

In particular,  $A_n$  contains the entries  $a_k$  from  $k = 1$  to  $k = 2n - 1$ , whose cardinality, by induction hypothesis, corresponds to the coefficients of the polynomial  $(\sum_{\alpha=0}^{n-1} x^\alpha)^2$ . Moreover, the entries  $a_k$  from  $k = n + 1$  to  $k = 2n - 1$  appear two extra times in  $A_{n+1}$ : once in the last row and once in the last column. Finally, in  $A_{n+1}$  we have two additional entries,  $a_{2n}$  and  $a_{2n+1}$ , that are not in  $A_n$ , and whose cardinality is, respectively, 2 and 1. To summarize, the cardinality of the entry  $a_k$  in  $A_{n+1}$ ,  $k = 1, \dots, 2n + 1$ , corresponds to the  $(k - 1)$ -th power of  $x$  in the following polynomial:

$$\left( \sum_{\alpha=0}^{n-1} x^\alpha \right)^2 + 2(x^n + \cdots + x^{2n-2}) + 2x^{2n-1} + x^{2n} = \left( \sum_{\alpha=0}^{n-1} x^\alpha \right)^2 + 2x^n \left( \sum_{\alpha=0}^{n-1} x^\alpha \right) + x^{2n} = \left( \sum_{\alpha=0}^n x^\alpha \right)^2,$$

which concludes the proof. □

*Example A.1.* Let  $n = m = 3$ . After some technical calculations, we get

$$E_{3,3} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad D_{3,3} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The number of non-zero elements in  $E_{3,3}$  is  $n + m - 1 = 5$  as stated in Lemma [A.1](#) and the number of non-zero elements in each column of  $D_{3,3}$  is  $1 - 2 - 3 - 2 - 1$ , like the coefficients in the polynomial expansion

$$\left( \sum_{\alpha=0}^2 x^\alpha \right)^2 = (1 + x + x^2)^2 = 1 + 2x + 3x^2 + 2x^3 + x^4,$$

according to Lemma [A.2](#)

## A.2 The matrix of functions

From Proposition [4.1](#)  $X_n^{(r)}(x)$  is a rectangular block matrix composed by  $(n + 1)^{r-1}$  blocks  $B_{n,r}^{(k)}(x) \in \mathcal{A}_{n+1}$ , which can be expressed by  $B_{n,r}^{(k)}(x) = x^{j_k} X_n^{(1)}(x)$  for a certain  $j_k \in \{0, \dots, (r - 1)n\}$ . We shall now give the cardinality of each block and an explicit formula for  $j_k$ . We denote by  $\text{mod}$  and  $\%$  the operators which, respectively, return the remainder and the quotient of the division between two natural numbers, namely for  $a, b, c, d \in \mathbb{N}$ ,  $c = a \text{ mod } b$  and  $d = a \% b$  means that  $a = bd + c$ .

**Lemma A.3.** *For every  $n, r \geq 1$ , each block of the form  $x^j X_n^{(1)}(x)$ , for  $j = 0, \dots, (r - 1)n$ , is repeated with cardinality  $\beta_{n,r}^{(j)} := \#\{k : j_k = j\}$ , that is equal to the coefficient of the  $j$ -th power of  $x$  in the polynomial expansion  $(\sum_{\alpha=0}^n x^\alpha)^{r-1}$ .*

*Proof.* We proceed by induction on  $r \geq 1$ .

- $r = 1$ :  $X_n^{(1)}(x) = x^0 X_n^{(1)}(x)$  is composed by only one block, and the polynomial  $(\sum_{\alpha=0}^n x^\alpha)^0 = 1$  has the only coefficient 1, that is the cardinality of the unique block.
- $r \rightarrow r + 1$ : we assume the statement holds for  $r$ . We need to prove that for each  $j = 0, \dots, rn$ , the cardinality of the block of the form  $x^j X_n^{(1)}(x)$  corresponds to the coefficient of the  $j$ -th power of  $x$  in the polynomial expansion  $(\sum_{\alpha=0}^n x^\alpha)^r$ . Since the Kronecker product is associative, we write

$$X_n^{(r+1)}(x) = H_n(x)^\top \otimes X_n^{(r)}(x) = \left( X_n^{(r)}(x) \vdots x X_n^{(r)}(x) \vdots x^2 X_n^{(r)}(x) \vdots \dots \vdots x^n X_n^{(r)}(x) \right). \quad (\text{A.3})$$

We know by induction hypothesis that each block of  $X_n^{(r)}(x)$  of the form  $x^j X_n^{(1)}(x)$ ,  $j \in \{0, \dots, (r-1)n\}$ , has cardinality according to the  $j$ -th coefficient of the polynomial expansion  $(\sum_{\alpha=0}^n x^\alpha)^{r-1}$ . It is clear that such cardinality shifts to the upper power when we multiply  $X_n^{(r)}(x)$  by  $x$ , and it shifts by two positions when we multiply  $X_n^{(r)}(x)$  by  $x^2$ , and so on (for example, if the block  $x^j X_n^{(1)}(x)$  has cardinality  $\beta_{n,r}^{(j)}$  in  $X_n^{(r)}(x)$ , then  $\beta_{n,r}^{(j)}$  is the cardinality of  $x^{j+1} X_n^{(1)}(x)$  in  $x X_n^{(r)}(x)$ , of  $x^{j+2} X_n^{(1)}(x)$  in  $x^2 X_n^{(r)}(x)$ , and so on). To summarize, we say that in  $X_n^{(r+1)}(x)$  the cardinality for the block of the form  $x^j X_n^{(1)}(x)$  corresponds to the coefficient of the  $j$ -th power in the following polynomial:

$$\left( \sum_{\alpha=0}^n x^\alpha \right)^{r-1} + x \left( \sum_{\alpha=0}^n x^\alpha \right)^{r-1} + \dots + x^n \left( \sum_{\alpha=0}^n x^\alpha \right)^{r-1} = \left( \sum_{\alpha=0}^n x^\alpha \right)^r,$$

which concludes the proof. □

**Lemma A.4.** For every  $n, r \geq 1$ , the blocks  $B_{n,r}^{(k)}(x)$  composing the matrix of functions  $X_n^{(r)}(x)$  are of the form  $B_{n,r}^{(k)}(x) = x^{\gamma_{n,r}^{(k)}} X_n^{(1)}(x)$ , for  $\gamma_{n,r}^{(k)} \in \{0, \dots, (r-1)n\}$  given by the formula

$$\gamma_{n,r}^{(k)} = \sum_{j=0}^{r-1} \{(k-1) \bmod (n+1)^{r-j}\} \% (n+1)^{r-1-j} \quad \text{for } k = 1, \dots, (n+1)^{r-1}.$$

*Proof.* By associativity property of the Kronecker product,  $X_n^{(r)}(x) = (H_n(x)^\top)^{\otimes(r-1)} \otimes X_n^{(1)}(x)$ , where  $(H_n(x)^\top)^{\otimes(r-1)}$  is a row vector in  $\mathbb{R}^{(n+1)^{r-1}}$  with elements the monomials  $x^{\gamma_{n,r}^{(k)}}$  whose exponents we want to study,  $k = 1, \dots, (n+1)^{r-1}$ . We focus on  $(H_n(x)^\top)^{\otimes r}$  for simplicity: we need then to prove that the  $k$ -th element of  $(H_n(x)^\top)^{\otimes r}$  is a monomial with exponent given by

$$P_{n,r}^{(k)} = \sum_{j=0}^r \{(k-1) \bmod (n+1)^{r-j+1}\} \% (n+1)^{r-j} \quad \text{for } k = 1, \dots, (n+1)^r. \quad (\text{A.4})$$

The result will follow noticing that  $\gamma_{n,r}^{(k)} = P_{n,r-1}^{(k)}$ . We proceed by induction on  $r \geq 1$ .

- $r = 1$ : for  $H_n(x)$  we easily notice that the exponent of the  $k$ -th term equals  $k-1$ ,  $k = 1, \dots, (n+1)$ . We now look at equation [\(A.4\)](#):

$$\begin{aligned} P_{n,1}^{(k)} &= \sum_{j=0}^1 \{(k-1) \bmod (n+1)^{2-j}\} \% (n+1)^{1-j} \\ &= \{(k-1) \bmod (n+1)^2\} \% (n+1) + \{(k-1) \bmod (n+1)\} \% (n+1)^0 \end{aligned}$$

where the first term is 0 and the second one is  $k-1$ , since  $0 \leq k-1 \leq n$ .

- $r \rightarrow r + 1$ : we now assume formula [\(A.4\)](#) holds for  $r$ . By associativity property,  $(H_n(x)^\top)^{\otimes(r+1)} =$

$H_n(x)^\top \otimes (H_n(x)^\top)^{\otimes r}$ , so that  $(H_n(x)^\top)^{\otimes(r+1)}$  is the block vector

$$(H_n(x)^\top)^{\otimes(r+1)} = \left( (H_n(x)^\top)^{\otimes r} : x (H_n(x)^\top)^{\otimes r} : x^2 (H_n(x)^\top)^{\otimes r} : \dots : x^n (H_n(x)^\top)^{\otimes r} \right).$$

For each element in  $(H_n(x)^\top)^{\otimes(r+1)}$ , the exponent is  $P_{n,r}^{(k)} + \alpha$  with  $\alpha \in \{0, \dots, n\}$ . However, we notice that  $(H_n(x)^\top)^{\otimes r}$  has index  $k = 1, \dots, (n+1)^r$ , while  $(H_n(x)^\top)^{\otimes(r+1)}$  has index  $\hat{k} = 1, \dots, (n+1)^{r+1}$ . Then in formula [\(A.4\)](#) we must substitute  $(k-1) = (\hat{k}-1) \bmod (n+1)^r$ . Moreover,  $\alpha = (\hat{k}-1) \% (n+1)^r$ . Putting all these considerations together, we write that (we omit the superscript  $\wedge$  on the index  $k$ ):

$$P_{n,r+1}^{(k)} = (k-1) \% (n+1)^r + \sum_{j=0}^r \left\{ (k-1) \bmod (n+1)^r \bmod (n+1)^{r-j+1} \right\} \% (n+1)^{r-j}. \quad (\text{A.5})$$

In particular, it is easy to see that

$$(k-1) \bmod (n+1)^r \bmod (n+1)^{r-j+1} = \begin{cases} (k-1) \bmod (n+1)^{r-j+1} & j \geq 1 \\ (k-1) \bmod (n+1)^r & j = 0 \end{cases},$$

so that equation [\(A.5\)](#) becomes

$$P_{n,r+1}^{(k)} = (k-1) \% (n+1)^r + (k-1) \bmod (n+1)^r \% (n+1)^r + \sum_{j=1}^r \left\{ (k-1) \bmod (n+1)^{r-j+1} \right\} \% (n+1)^{r-j}. \quad (\text{A.6})$$

Moreover, since  $1 \leq k \leq (n+1)^{r+1}$ , we also notice that

$$(k-1) \% (n+1)^r = (k-1) \bmod (n+1)^{r+1} \% (n+1)^r \quad \text{and} \\ (k-1) \bmod (n+1)^r \% (n+1)^r (= 0) = (k-1) \bmod (n+1)^{r+2} \% (n+1)^{r+1},$$

and equation [\(A.6\)](#) can be rewritten as

$$P_{n,r+1}^{(k)} = \sum_{j=-1}^r \left\{ (k-1) \bmod (n+1)^{r-j+1} \right\} \% (n+1)^{r-j} \\ = \sum_{j=0}^{r+1} \left\{ (k-1) \bmod (n+1)^{r-j+2} \right\} \% (n+1)^{r-j+1},$$

which concludes the proof. □

## B Proofs

We report in this section the proofs of the main results of the paper together with some additional results that are needed for the proofs.

**Proposition B.1.** *For the matrices  $A, B \in \mathbb{R}^{n \times m}$  with elements, respectively,  $[A]_{i,j} = a_{i,j}$  and  $[B]_{i,j} = b_{i,j}$ ,  $1 \leq i \leq n$  and  $1 \leq j \leq m$ , and  $\vec{x}, \vec{y}$  vectors of any order, we have the following properties:*

$$A = \sum_{i=1}^n \sum_{j=1}^m a_{i,j} \vec{e}_{n,i} \vec{e}_{m,j}^\top \quad (\text{B.1a})$$

$$(\text{vec}(A))^\top \text{vec}(B) = \text{tr}(A^\top B) \quad (\text{B.1b})$$

$$\vec{x} \otimes \vec{y} = \text{vec}(\vec{y} \vec{x}^\top) \quad (\text{B.1c})$$

$$\vec{x} \otimes \vec{y}^\top = \vec{x} \vec{y}^\top = \vec{y}^\top \otimes \vec{x} \quad (\text{B.1d})$$

where  $\text{tr}$  denotes the trace operator. Moreover, for every  $A \in \mathbb{R}^{p \times q}$ ,  $B \in \mathbb{R}^{r \times s}$ ,  $C \in \mathbb{R}^{q \times k}$  and  $D \in \mathbb{R}^{s \times l}$ , the mixed-product property holds:

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD). \quad (\text{B.2})$$

*Proof.* We refer to [29] Section 2] and [22] Lemma 4.2.10].

### Proof of Theorem 3.1

*Proof.* We will prove the existence of the matrix  $E_{n,m}$  by proving its explicit definition. We first give a characterization of  $\text{vec}L(A)$  in terms of the unitary vectors. By equation (B.1a), one easily sees that

$$\text{vec}L(A) = \sum_{i=1}^n a_{i,1} \vec{e}_{n+m-1,i} + \sum_{i=2}^m a_{n,i} \vec{e}_{n+m-1,n+i-1}. \quad (\text{B.3})$$

In particular,  $a_{i,1} = \vec{e}_{n,i}^\top A \vec{e}_{m,1} = \text{tr}(\vec{e}_{m,1} \vec{e}_{n,i}^\top A)$  and  $a_{n,i} = \vec{e}_{n,n}^\top A \vec{e}_{m,i} = \text{tr}(\vec{e}_{m,i} \vec{e}_{n,n}^\top A)$ . Moreover, by property (B.1b) we write that

$$\begin{aligned} \text{tr}(\vec{e}_{m,1} \vec{e}_{n,i}^\top A) &= \text{tr}((\vec{e}_{n,i} \vec{e}_{m,1}^\top)^\top A) = \text{vec}(\vec{e}_{n,i} \vec{e}_{m,1}^\top)^\top \text{vec}(A), \\ \text{tr}(\vec{e}_{m,i} \vec{e}_{n,n}^\top A) &= \text{tr}((\vec{e}_{n,n} \vec{e}_{m,i}^\top)^\top A) = \text{vec}(\vec{e}_{n,n} \vec{e}_{m,i}^\top)^\top \text{vec}(A). \end{aligned}$$

By combining these results with equations (B.1c), (B.1d) and (B.3), we get that

$$\begin{aligned} \text{vec}L(A) &= \left( \sum_{i=1}^n \vec{e}_{n+m-1,i} \text{vec}(\vec{e}_{n,i} \vec{e}_{m,1}^\top)^\top + \sum_{i=2}^m \vec{e}_{n+m-1,n+i-1} \text{vec}(\vec{e}_{n,n} \vec{e}_{m,i}^\top)^\top \right) \text{vec}(A) \\ &= \left( \sum_{i=1}^n \vec{e}_{n+m-1,i} \otimes \vec{e}_{m,1}^\top \otimes \vec{e}_{n,i}^\top + \sum_{i=2}^m \vec{e}_{n+m-1,n+i-1} \otimes \vec{e}_{n,n}^\top \otimes \vec{e}_{m,i}^\top \right) \text{vec}(A). \end{aligned}$$

Then the L-eliminating matrix  $E_{n,m}$  satisfying the implicit definition in equation (3.7) is the one in equation (3.8). This concludes the proof.  $\square$

### Proof of Theorem 3.3

*Proof.* We construct the matrix  $D_{n,m}$  explicitly. Since  $A \in \mathcal{A}_{n,m}$ , the elements of  $A$  along the skew-diagonals coincide and  $A$  has exactly  $n + m - 1$  skew-diagonals, leading to at most  $n + m - 1$  different values. In the notation of Definition 3.4 let  $a_k$ ,  $k = 1, \dots, (n + m - 1)$ , such that  $\text{vec}L(A) = (a_1, a_2, \dots, a_{n+m-1})^\top \in \mathbb{R}^{n+m-1}$ . For  $1 \leq i \leq n$  and  $1 \leq j \leq m$ , it holds that

$$[A]_{i,j} = a_{i+j-1} = [\text{vec}L(A)]_{i+j-1} = \text{vec}L(A)^\top \vec{e}_{n+m-1,i+j-1} = \vec{e}_{n+m-1,i+j-1}^\top \text{vec}L(A). \quad (\text{B.4})$$

We notice that  $\vec{e}_{m,j} \otimes \vec{e}_{n,i}$  is the unitary vector in  $\mathbb{R}^{nm}$  with 1 in position  $n(j-1) + i$  and 0 elsewhere. We use this fact together with equation (B.4) to express the vectorization of  $A$  as follows:

$$\text{vec}(A) = \sum_{i=1}^n \sum_{j=1}^m a_{i+j-1} \vec{e}_{m,j} \otimes \vec{e}_{n,i} = \left( \sum_{i=1}^n \sum_{j=1}^m (\vec{e}_{m,j} \otimes \vec{e}_{n,i}) \vec{e}_{n+m-1,i+j-1}^\top \right) \text{vec}L(A).$$

By equation (B.1d), we define  $D_{n,m}$  as in equation (3.10), which proves the theorem.  $\square$



### Proof of Proposition 3.5

*Proof.* For the second part of the statement, the proof is straightforward from the definitions of  $E_{n,m}$  and  $D_{n,m}$ , namely combining equations (3.7) and (3.9). We resume the situation as follows:

$$\text{vec}(A) \xrightarrow{E_{n,m}} \text{vec}L(A) \xrightarrow{D_{n,m}} \text{vec}(A)$$

so that  $D_{n,m}E_{n,m} \in \mathbb{R}^{nm \times nm}$  acts like an identity operator on  $\text{vec}(A)$  for each  $A \in \mathcal{A}_{n,m}$ .

We prove now the first part of the statement. By equations (3.8) and (3.10), we write that

$$\begin{aligned} E_{n,m}D_{n,m} &= \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^n (\vec{e}_{n+m-1,k} \otimes \vec{e}_{m,1}^\top \otimes \vec{e}_{n,k}^\top) (\vec{e}_{n+m-1,i+j-1}^\top \otimes \vec{e}_{m,j} \otimes \vec{e}_{n,i}) + \\ &\quad + \sum_{i=1}^n \sum_{j=1}^m \sum_{k=2}^m (\vec{e}_{n+m-1,n+k-1} \otimes \vec{e}_{m,k}^\top \otimes \vec{e}_{n,n}^\top) (\vec{e}_{n+m-1,i+j-1}^\top \otimes \vec{e}_{m,j} \otimes \vec{e}_{n,i}). \end{aligned} \quad (\text{B.5})$$

We denote with  $\mathbb{I}_i \in \mathbb{R}^{(n+m-1) \times (n+m-1)}$  the matrix with 1 in position  $(i, i)$  and 0 elsewhere, and with  $\mathbb{I}^{i,j} \in \mathbb{R}^{(n+m-1) \times (n+m-1)}$  the matrix with 1 in position  $(k, k)$  for  $i \leq k \leq j$ . We focus on the first sum in equation (B.5). We notice that:

- $\vec{e}_{m,1}^\top \otimes \vec{e}_{n,k}^\top = \vec{e}_{nm,k}^\top$  and  $A_k := \vec{e}_{n+m-1,k} \otimes \vec{e}_{m,1}^\top \otimes \vec{e}_{n,k}^\top \in \mathbb{R}^{(n+m-1) \times nm}$  is the matrix with 1 in position  $(k, k)$  and 0 elsewhere;
- $\vec{e}_{m,j} \otimes \vec{e}_{n,i} = \vec{e}_{nm,n(j-1)+i}$  and  $B_{i,j} := \vec{e}_{n+m-1,i+j-1}^\top \otimes \vec{e}_{m,j} \otimes \vec{e}_{n,i} \in \mathbb{R}^{nm \times (n+m-1)}$  is the matrix with 1 in position  $(n(j-1) + i, i + j - 1)$  and 0 elsewhere.

Then  $A_k B_{i,j} = \mathbb{I}_i$ . Similarly, looking at the second sum in equation (B.5), we notice that  $\vec{e}_{m,k}^\top \otimes \vec{e}_{n,n}^\top = \vec{e}_{nm,kn}^\top$  and  $\tilde{A}_k := \vec{e}_{n+m-1,n+k-1} \otimes \vec{e}_{m,k}^\top \otimes \vec{e}_{n,n}^\top \in \mathbb{R}^{(n+m-1) \times nm}$  is the matrix with 1 in position  $(n+k-1, kn)$  and 0 elsewhere. Then  $\tilde{A}_k B_{i,j} = \mathbb{I}_{n+k-1}$ . Combining these results into equation (B.5) we get

$$\begin{aligned} E_{n,m}D_{n,m} &= \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^n A_k B_{i,j} + \sum_{i=1}^n \sum_{j=1}^m \sum_{k=2}^m \tilde{A}_k B_{i,j} \\ &= \sum_{k=1}^n \mathbb{I}_k + \sum_{k=2}^m \mathbb{I}_{n+k-1} = \mathbb{I}^{1,n} + \mathbb{I}^{n+1,n+m-1} = I_{n+m-1}, \end{aligned}$$

that concludes the proof.  $\square$

### Proof of Proposition 3.7

*Proof.* By combining  $\mathcal{G}H_{2n}(x) = G_{2n}H_{2n}(x)$  with Lemma 3.6 we get  $\text{vec}L(\mathcal{G}X_n(x)) = G_{2n}\text{vec}L(X_n(x))$ , which, by definition of  $E_{n+1}$  and multiplying both sides by  $D_{n+1}$  from the left, becomes

$$D_{n+1}E_{n+1}\text{vec}(\mathcal{G}X_n(x)) = D_{n+1}G_{2n}E_{n+1}\text{vec}(X_n(x)).$$

In particular,  $D_{n+1}E_{n+1}\text{vec}(\mathcal{G}X_n(x)) = \text{vec}(\mathcal{G}X_n(x))$  by Proposition 3.5 so that  $\tilde{G}_n^{(1)} = D_{n+1}G_{2n}E_{n+1}$ . We consider now the definition of exponential function as infinite sum of powers and write that

$$e^{\tilde{G}_n^{(1)}t} = e^{D_{n+1}G_{2n}E_{n+1}t} = \sum_{k=0}^{\infty} \frac{t^k}{k!} (D_{n+1}G_{2n}E_{n+1})^k.$$

The result follows from Proposition 3.5 since  $(D_{n+1}G_{2n}E_{n+1})^k = D_{n+1}(G_{2n})^k E_{n+1}$ ,  $k \geq 0$ .  $\square$

**Proof of Theorem 3.8**

*Proof.* Starting from equation (3.3) and applying the operator  $vec$  on both sides, we get

$$\mathbb{E} [vec(X_n(Y(s_0))) | \mathcal{F}_t] = vec(X_n(Y(t))) + \int_t^{s_0} \mathbb{E} [\mathcal{G}vec(X_n(Y(s))) | \mathcal{F}_t] ds,$$

which, by equation (3.5) and Proposition 3.7 becomes

$$\mathbb{E} [vec(X_n(Y(s_0))) | \mathcal{F}_t] = vec(X_n(Y(t))) + D_{n+1}G_{2n}E_{n+1} \int_t^{s_0} \mathbb{E} [vec(X_n(Y(s))) | \mathcal{F}_t] ds.$$

For  $Z(s) := \mathbb{E}[vec(X_n(Y(s))) | \mathcal{F}_t]$ , the proof then proceeds in a similar way to the proof of Theorem 2.1 and the statement is proved by combining the result with equation (3.1).  $\square$

**Proof of Proposition 4.1**

*Proof.* By definition of d-Kronecker product, since  $H_n(x) \in \mathbb{R}^{n+1}$ , one can verify that  $X_n^{(r)}(x) \in \mathbb{R}^{(n+1) \times (n+1)^r}$ . We then proceed by induction on  $r \geq 1$ .

- $r = 1$ : we get that  $X_n^{(1)}(x) = H_n(x)^\top \otimes H_n(x) = x^0 X_n^{(1)}(x) \in \mathcal{A}_{n+1}$ .
- $r \rightarrow r + 1$ : assuming the statement holds for  $r$ , from equation (A.3), we see that we need to multiply the row vector  $H_n(x)^\top = (1, x, \dots, x^n)$  in the Kronecker sense with the matrix  $X_n^{(r)}(x) = H_n(x)^\top \otimes^r H_n(x)$ , which we know satisfies the statement of the proposition. That means that each of the  $(n+1)^{r-1}$  blocks  $B_{n,r}^{(k)}(x) = x^{j_k} X_n^{(1)}(x)$ , for  $j_k \in \{0, \dots, (r-1)n\}$ , must be multiplied with each of the elements of the vector  $H_n(x)$ , namely with each power  $x^\alpha$ ,  $\alpha = 0, \dots, n$ . We can then say that there are  $(n+1)^r$  blocks  $B_{n,r+1}^{(k)}(x) \in \mathcal{A}_{n+1}$  and that for each block there exists an index  $\gamma_k \in \{0, \dots, rn\}$  such that  $B_{n,r+1}^{(k)}(x) = x^{\gamma_k} X_n^{(1)}(x)$ . This concludes the proof.  $\square$

**Lemma B.2.** *It holds that  $vec(X_n^{(m)}(x)) = H_n(x)^{\otimes m+1}$ . Moreover, after removing all the duplicates from  $vec(X_n^{(m)}(x))$ , we are left with  $H_{n(m+1)}(x)$ .*

*Proof.* The result follows from a direct verification.  $\square$

**Lemma B.3.** *There exist an L-eliminating matrix  $E_{nm+1,n+1}$  and an L-duplicating matrix  $D_{nm+1,n+1}$  such that*

$$E_{nm+1,n+1}(H_n(x) \otimes H_{nm}(x)) = H_{n(m+1)}(x) \quad \text{and} \quad D_{nm+1,n+1}H_{n(m+1)}(x) = H_n(x) \otimes H_{nm}(x). \quad (\text{B.6})$$

*Proof.* From a direct verification, it can be seen that

$$vec(H_n(x)^\top \otimes H_{nm}(x)) = H_n(x) \otimes H_{nm}(x) \quad \text{and} \quad vecL(H_n(x)^\top \otimes H_{nm}(x)) = H_{n(m+1)}(x). \quad (\text{B.7})$$

Then, from Theorem 3.1 there exists an L-eliminating matrix  $E_{nm+1,n+1}$  transforming the vectorization of  $H_n(x)^\top \otimes H_{nm}(x)$  into its L-vectorization. By equation (B.7), this is equivalent to saying that  $E_{nm+1,n+1}$  maps  $H_n(x) \otimes H_{nm}(x)$  to  $H_{n(m+1)}(x)$ , which is what claimed in equation (B.6). Similarly, by Theorem 3.3 there exists an L-duplicating matrix  $D_{nm+1,n+1}$  satisfying equation (B.6).  $\square$

**Proof of Proposition 4.2**

*Proof.* We proceed by induction on  $m \geq 1$ .

- $m = 1$ : see Corollary 3.2 and equation (4.2).

- $m - 1 \rightarrow m$ : we assume the statement holds for  $m - 1$ , namely there exists a matrix  $E_{n+1}^{(m-1)}$  that applied to  $\text{vec}(X_n^{(m-1)}(x))$  removes all the duplicates. By Lemma [B.2](#) this means that

$$E_{n+1}^{(m-1)} H_n(x)^{\otimes m} = H_{nm}(x).$$

We now multiply both sides in the Kronecker sense by  $H_n(x)$ , and successively apply on the left the matrix  $E_{nm+1, n+1}$ , obtaining that

$$E_{nm+1, n+1} \left[ H_n(x) \otimes \left( E_{n+1}^{(m-1)} H_n(x)^{\otimes m} \right) \right] = E_{nm+1, n+1} [H_n(x) \otimes H_{nm}(x)].$$

From the identity  $H_n(x) = I_{n+1} H_n(x)$ , and applying the mixed-product property of the Kronecker product (equation [B.2](#)) on the left hand side, and equation [B.6](#) on the right hand side, we get

$$E_{nm+1, n+1} \left[ \left( I_{n+1} \otimes E_{n+1}^{(m-1)} \right) \left( H_n(x) \otimes H_n(x)^{\otimes m} \right) \right] = H_{n(m+1)}(x).$$

Since  $H_n(x) \otimes H_n(x)^{\otimes m} = H_n(x)^{\otimes(m+1)} = \text{vec}(X_n^{(m)}(x))$  by Lemma [B.2](#) the matrix  $E_{n+1}^{(m)} = E_{nm+1, n+1} \left( I_{n+1} \otimes E_{n+1}^{(m-1)} \right)$  is exactly the one removing all the duplicates from  $\text{vec}(X_n^{(m)}(x))$ . □

### Proof of Proposition [4.3](#)

*Proof.* We proceed by induction on  $m \geq 1$ .

- $m = 1$ : see Corollary [3.4](#) and equation [4.2](#).
- $m - 1 \rightarrow m$ : we assume the statement holds for  $m - 1$ . Then, starting from equation [B.6](#) and multiplying both sides with  $I_{n+1} \otimes D_{n+1}^{(m-1)}$  we get

$$\left( I_{n+1} \otimes D_{n+1}^{(m-1)} \right) \left( D_{nm+1, n+1} H_{n(m+1)}(x) \right) = \left( I_{n+1} \otimes D_{n+1}^{(m-1)} \right) \left( H_n(x) \otimes H_{nm}(x) \right). \quad (\text{B.8})$$

By the mixed-product property of the Kronecker product (equation [B.2](#)), the right hand side is

$$\begin{aligned} \left( I_{n+1} \otimes D_{n+1}^{(m-1)} \right) \left( H_n(x) \otimes H_{nm}(x) \right) &= \left( I_{n+1} H_n(x) \right) \otimes \left( D_{n+1}^{(m-1)} H_{nm}(x) \right) \\ &= H_n(x) \otimes \left( D_{n+1}^{(m-1)} H_{nm}(x) \right). \end{aligned}$$

From the induction hypothesis,  $D_{n+1}^{(m-1)}$  satisfies  $D_{n+1}^{(m-1)} H_{nm}(x) = \text{vec}(X_n^{(m-1)}(x))$ , and, by Lemma [B.2](#) we also have  $\text{vec}(X_n^{(m-1)}(x)) = H_n(x)^{\otimes m}$ . Then equation [B.8](#) becomes

$$\left( I_{n+1} \otimes D_{n+1}^{(m-1)} \right) D_{nm+1, n+1} H_{n(m+1)}(x) = H_n(x) \otimes H_n(x)^{\otimes m} = H_n(x)^{\otimes(m+1)},$$

and the matrix  $D_{n+1}^{(m)} = \left( I_{n+1} \otimes D_{n+1}^{(m-1)} \right) D_{nm+1, n+1}$  is exactly the one required. □

**Proposition B.4.** For every  $n, m \geq 1$ ,  $\vec{v}_n \in \mathbb{R}^{n+1}$  and  $M_n^{(m-1)} \in \mathbb{R}^{(n+1)^m \times (n+1)^m}$ , the identity holds:

$$H_n(x) \vec{v}_n^\top \left\{ \text{vec}^{-1} \circ M_n^{(m-1)} \circ \text{vec} \left( X_n^{(m-1)}(x) \right) \right\} = \left\{ X_n^{(m)}(x) \right\} M_n^{(m-1)\top} \left\{ I_{n+1} \otimes^{m-1} \vec{v}_n \right\}. \quad (\text{B.9})$$

*Proof.* We proceed by induction on the order  $m \geq 1$ .

- $m = 1$ : starting from the left hand side of identity [B.9](#), we get

$$H_n(x) \vec{v}_n^\top \left\{ \text{vec}^{-1} \circ M_n^{(0)} \circ \text{vec} \left( X_n^{(0)}(x) \right) \right\} = H_n(x) \vec{v}_n^\top M_n^{(0)} H_n(x) = H_n(x) H_n(x)^\top M_n^{(0)\top} \vec{v}_n$$

$$= \left\{ X_n^{(1)}(x) \right\} M_n^{(0)\top} \{ I_{n+1} \otimes^0 \vec{v}_n \}.$$

Remember indeed that the  $vec^{-1}$  operator transforms a vector into an object with the same dimension as the argument of the operator  $vec$  previously applied. But in this case the argument of  $vec$  is a vector already, hence both  $vec$  and  $vec^{-1}$  coincide in practise with the identity operator. Moreover,  $\vec{v}_n^\top M_n^{(0)} H_n(x) \in \mathbb{R}$  and it equals its transpose. This proves the base case.

- $m \rightarrow m + 1$ : we assume identity [\(B.9\)](#) holds for  $m$  and consider

$$H_n(x) \vec{v}_n^\top \left\{ vec^{-1} \circ M_n^{(m)} \circ vec \left( X_n^{(m)}(x) \right) \right\}.$$

In particular,  $M_n^{(m)} \in \mathbb{R}^{(n+1)^{m+1} \times (n+1)^{m+1}}$  can be seen as made up of  $(n+1)^2$  matrices of the form  $M_{i,j}^{(m-1)} \in \mathbb{R}^{(n+1)^m \times (n+1)^m}$ ,  $1 \leq i, j \leq n+1$ , so that  $M_n^{(m)}$  looks like

$$M_n^{(m)} = \begin{pmatrix} M_{1,1}^{(m-1)} & \cdots & \cdots & M_{1,n+1}^{(m-1)} \\ \vdots & \ddots & \ddots & \vdots \\ M_{n+1,1}^{(m-1)} & \cdots & \cdots & M_{n+1,n+1}^{(m-1)} \end{pmatrix}.$$

The idea is then to break up the matrix  $M_n^{(m)}$  into sub-matrices, for which we know the statement holds by induction hypothesis. In what follows, starting from  $X_n^{(m)}(x)$ , we will apply in the following order: the  $vec$  operator, the matrix  $M_n^{(m)}$ , the  $vec^{-1}$  operator and finally the matrix  $H_n(x) \vec{v}_n^\top$ . At this point we will be able to apply the induction hypothesis, and prove the statement.

By Lemma [B.2](#) and associativity property of the Kronecker product, we get

$$vec \left( X_n^{(m)}(x) \right) = H_n(x) \otimes H_n(x)^{\otimes m} = \begin{pmatrix} H_n(x)^{\otimes m} \\ x H_n(x)^{\otimes m} \\ \vdots \\ x^n H_n(x)^{\otimes m} \end{pmatrix},$$

where  $x^k H_n(x)^{\otimes m} \in \mathbb{R}^{(n+1)^m}$ ,  $k = 0, \dots, n$ , thus

$$M_n^{(m)} vec \left( X_n^{(m)}(x) \right) = \begin{pmatrix} M_{1,1}^{(m-1)} H_n(x)^{\otimes m} + M_{1,2}^{(m-1)} x H_n(x)^{\otimes m} + \cdots + M_{1,n+1}^{(m-1)} x^n H_n(x)^{\otimes m} \\ \vdots \\ M_{n+1,1}^{(m-1)} H_n(x)^{\otimes m} + M_{n+1,2}^{(m-1)} x H_n(x)^{\otimes m} + \cdots + M_{n+1,n+1}^{(m-1)} x^n H_n(x)^{\otimes m} \end{pmatrix}.$$

Applying the  $vec^{-1}$  operator to the last matrix obtained, by linearity we get

$$\begin{aligned} & \left( vec^{-1} \left( M_{1,1}^{(m-1)} H_n(x)^{\otimes m} \right) + \cdots + x^n vec^{-1} \left( M_{1,n+1}^{(m-1)} H_n(x)^{\otimes m} \right), \right. \\ & \quad \dots, \\ & \left. vec^{-1} \left( M_{n+1,1}^{(m-1)} H_n(x)^{\otimes m} \right) + \cdots + x^n vec^{-1} \left( M_{n+1,n+1}^{(m-1)} H_n(x)^{\otimes m} \right) \right), \end{aligned}$$

where  $vec^{-1} \left( M_{i,j}^{(m-1)} H_n(x)^{\otimes m} \right) \in \mathbb{R}^{(n+1) \times (n+1)^{m-1}}$ ,  $1 \leq i, j \leq n+1$ . Multiplying the above equation by  $H_n(x) \vec{v}_n^\top$ , we obtain that

$$\begin{aligned} & H_n(x) \vec{v}_n^\top \left\{ vec^{-1} \circ M_n^{(m)} \circ vec \left( X_n^{(m)}(x) \right) \right\} \\ &= \left( H_n(x) \vec{v}_n^\top vec^{-1} \left( M_{1,1}^{(m-1)} H_n(x)^{\otimes m} \right) + \cdots + x^n H_n(x) \vec{v}_n^\top vec^{-1} \left( M_{1,n+1}^{(m-1)} H_n(x)^{\otimes m} \right), \right. \\ & \quad \dots, \\ & \left. H_n(x) \vec{v}_n^\top vec^{-1} \left( M_{n+1,1}^{(m-1)} H_n(x)^{\otimes m} \right) + \cdots + x^n H_n(x) \vec{v}_n^\top vec^{-1} \left( M_{n+1,n+1}^{(m-1)} H_n(x)^{\otimes m} \right) \right). \end{aligned}$$

Since  $H_n(x)^{\otimes m} = vec \left( X_n^{(m-1)}(x) \right)$ , we apply the induction hypothesis to each term

$$H_n(x) \vec{v}_n^\top vec^{-1} \left( M_{i,j}^{(m-1)} H_n(x)^{\otimes m} \right) = \left\{ X_n^{(m)}(x) \right\} M_{i,j}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \}$$

for  $1 \leq i, j \leq n+1$ , so that

$$\begin{aligned} H_n(x) \vec{v}_n^\top \left\{ \text{vec}^{-1} \circ M_n^{(m)} \circ \text{vec} \left( X_n^{(m)}(x) \right) \right\} = \\ \left( \left\{ X_n^{(m)}(x) \right\} M_{1,1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} + \dots + x^n \left\{ X_n^{(m)}(x) \right\} M_{1,n+1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \}, \right. \\ \dots, \\ \left. \left\{ X_n^{(m)}(x) \right\} M_{n+1,1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} + \dots + x^n \left\{ X_n^{(m)}(x) \right\} M_{n+1,n+1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} \right) \end{aligned}$$

which can also be seen as the following matrix product:

$$\left( \left\{ X_n^{(m)}(x) \right\}, \dots, x^n \left\{ X_n^{(m)}(x) \right\} \right) \begin{pmatrix} M_{1,1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} & \dots & M_{n+1,1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} \\ \vdots & \ddots & \vdots \\ M_{1,n+1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} & \dots & M_{n+1,n+1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} \end{pmatrix}.$$

In particular, the following identity also holds

$$\left( \left\{ X_n^{(m)}(x) \right\}, \dots, x^n \left\{ X_n^{(m)}(x) \right\} \right) = H_n(x)^\top \otimes \left\{ X_n^{(m)}(x) \right\} = X_n^{(m+1)}(x),$$

so that we can conclude with the expression

$$\begin{aligned} & \begin{pmatrix} M_{1,1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} & \dots & M_{n+1,1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} \\ \vdots & \ddots & \vdots \\ M_{1,n+1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} & \dots & M_{n+1,n+1}^{(m-1)\top} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} \end{pmatrix} = \\ & \begin{pmatrix} M_{1,1}^{(m-1)\top} & \dots & M_{n+1,1}^{(m-1)\top} \\ \vdots & \ddots & \vdots \\ M_{1,n+1}^{(m-1)\top} & \dots & M_{n+1,n+1}^{(m-1)\top} \end{pmatrix} \begin{pmatrix} \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} \end{pmatrix} \end{aligned}$$

where the first matrix on the right hand side coincides with  $M_n^{(m)\top}$ , while the second is  $I_{n+1} \otimes \{ I_{n+1} \otimes^{m-1} \vec{v}_n \} = I_{n+1} \otimes^m \vec{v}_n$ . This means we proved

$$H_n(x) \vec{v}_n^\top \left\{ \text{vec}^{-1} \circ M_n^{(m)} \circ \text{vec} \left( X_n^{(m)}(x) \right) \right\} = \left\{ X_n^{(m+1)}(x) \right\} M_n^{(m)\top} \{ I_{n+1} \otimes^m \vec{v}_n \},$$

and therefore reached the claim.  $\square$

### Proof of Theorem 4.5

*Proof.* Following the same idea as in the proof of Theorem 3.8, we start by proving that

$$\mathbb{E} \left[ X_n^{(r)}(Y(s)) \middle| \mathcal{F}_t \right] = \text{vec}^{-1} \circ e^{\tilde{G}_n^{(r)}(s-t)} \circ \text{vec} \left( X_n^{(r)}(Y(t)) \right) \quad (\text{B.10})$$

for  $\tilde{G}_n^{(r)} = D_{n+1}^{(r)} G_{n(r+1)} E_{n+1}^{(r)}$ . By equation 2.2 we write

$$\mathbb{E} \left[ X_n^{(r)}(Y(s)) \middle| \mathcal{F}_t \right] = X_n^{(r)}(Y(t)) + \int_t^s \mathbb{E} \left[ \mathcal{G} \left( X_n^{(r)}(Y(u)) \right) \middle| \mathcal{F}_t \right] du,$$

and applying the  $\text{vec}$  operator on both sides we get

$$\mathbb{E} \left[ \text{vec} \left( X_n^{(r)}(Y(s)) \right) \middle| \mathcal{F}_t \right] = \text{vec} \left( X_n^{(r)}(Y(t)) \right) + \int_t^s \mathbb{E} \left[ \mathcal{G} \text{vec} \left( X_n^{(r)}(Y(u)) \right) \middle| \mathcal{F}_t \right] du. \quad (\text{B.11})$$

By Proposition [4.2](#), there exists an  $r$ -th L-eliminating matrix  $E_{n+1}^{(r)}$  such that

$$E_{n+1}^{(r)} \text{vec} \left( X_n^{(r)}(x) \right) = H_{n(r+1)}(x). \quad (\text{B.12})$$

From Theorem [2.1](#), there also exists a generator matrix  $G_{n(r+1)}$  such that

$$\mathcal{G}H_{n(r+1)}(x) = G_{n(r+1)}H_{n(r+1)}(x), \quad (\text{B.13})$$

and by Proposition [4.3](#) an  $r$ -th L-duplicating matrix  $D_{n+1}^{(r)}$  such that

$$D_{n+1}^{(r)}H_{n(r+1)}(x) = \text{vec} \left( X_n^{(r)}(x) \right). \quad (\text{B.14})$$

Combining equations [\(B.12\)](#), [\(B.13\)](#), [\(B.14\)](#) with equation [\(B.11\)](#) we get

$$\mathbb{E} \left[ \text{vec} \left( X_n^{(r)}(Y(s)) \right) \middle| \mathcal{F}_t \right] = \text{vec} \left( X_n^{(r)}(Y(t)) \right) + D_{n+1}^{(r)}G_{n(r+1)}E_{n+1}^{(r)} \int_t^s \mathbb{E} \left[ \text{vec} \left( X_n^{(r)}(Y(u)) \right) \middle| \mathcal{F}_t \right] du,$$

and, proceeding the proof as in Theorem [3.8](#) we obtain that for every  $n \geq 1$  and  $r \geq 0$ , the matrix  $\tilde{G}_n^{(r)} = D_{n+1}^{(r)}G_{n(r+1)}E_{n+1}^{(r)} \in \mathbb{R}^{(n+1)^{r+1} \times (n+1)^{r+1}}$  is such that the expectation formula [\(B.10\)](#) holds.

We now proceed by induction on the number of polynomials  $m \geq 1$  to prove the correlator formula.

- $m = 1$ : the formula coincides with the one given in Theorem [3.8](#)
- $m \rightarrow m + 1$ : we suppose the correlator formula holds for  $m$  and consider  $m + 1$  polynomial functions. By the tower rule and the induction hypothesis, we write that

$$\begin{aligned} C_{p_0, \dots, p_{m+1}}(s_0, \dots, s_{m+1}; t) &= \mathbb{E} \left[ p_{m+1}(Y(s_0)) C_{p_0, \dots, p_m}(s_1, \dots, s_{m+1}; s_0) \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[ \vec{p}_{m+1}^\top H_n(Y(s_0)) \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_1-s_0)} \circ \text{vec} \left( X_n^{(m)}(Y(s_0)) \right) \right\} \right. \\ &\quad \left. \cdot \prod_{k=1}^m e^{\tilde{G}_n^{(m-k)\top}(s_{k+1}-s_k)} \left\{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \right\} \middle| \mathcal{F}_t \right]. \end{aligned}$$

From Proposition [B.4](#) we also have the following crucial equality:

$$\begin{aligned} H_n(Y(s_0)) \vec{p}_m^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m)}(s_1-s_0)} \circ \text{vec} \left( X_n^{(m)}(Y(s_0)) \right) \right\} \\ = \left\{ X_n^{(m+1)}(Y(s_0)) \right\} e^{\tilde{G}_n^{(m)\top}(s_1-s_0)} \left\{ I_{n+1} \otimes^m \vec{p}_m \right\}. \end{aligned}$$

Combining the previous results and equation [\(B.10\)](#), we can write that

$$\begin{aligned} C_{p_0, \dots, p_{m+1}}(s_0, \dots, s_{m+1}; t) &= \mathbb{E} \left[ \vec{p}_{m+1}^\top \left\{ X_n^{(m+1)}(Y(s_0)) \right\} e^{\tilde{G}_n^{(m)\top}(s_1-s_0)} \left\{ I_{n+1} \otimes^m \vec{p}_m \right\} \right. \\ &\quad \left. \cdot \prod_{k=1}^m e^{\tilde{G}_n^{(m-k)\top}(s_{k+1}-s_k)} \left\{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \right\} \middle| \mathcal{F}_t \right] \\ &= \vec{p}_{m+1}^\top \mathbb{E} \left[ X_n^{(m+1)}(Y(s_0)) \middle| \mathcal{F}_t \right] \prod_{k=0}^m e^{\tilde{G}_n^{(m-k)\top}(s_{k+1}-s_k)} \left\{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \right\} \\ &= \vec{p}_{m+1}^\top \left\{ \text{vec}^{-1} \circ e^{\tilde{G}_n^{(m+1)}(s_0-t)} \circ \text{vec} \left( X_n^{(m+1)}(Y(t)) \right) \right\} \prod_{k=0}^m e^{\tilde{G}_n^{(m-k)\top}(s_{k+1}-s_k)} \left\{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k} \right\}. \end{aligned}$$

Rearranging the index in the product of the last equation, we get the formula for  $m + 1$  polynomial functions and conclude the proof.  $\square$

**Proof of Theorem 5.1**

*Proof.* We proceed by induction on the dimension  $n \geq 2$ .

- $n = 2$ : see Example 2.2
- $n - 1 \rightarrow n$ : we assume the recursion formula holds for  $n - 1$ . We then need to compute  $\mathcal{G}x^n$ . By equations (2.1) and (2.3), we write that

$$\mathcal{G}x^n = n(b_0 + b_1x)x^{n-1} + \frac{1}{2}n(n-1)(\sigma_0 + \sigma_1x + \sigma_2x^2)x^{n-2} + \int_{\mathbb{R}} ((x+z)^n - x^n - nx^{n-1}z) \ell(x, dz).$$

In particular, by binomial expansion

$$\begin{aligned} \int_{\mathbb{R}} ((x+z)^n - x^n - nx^{n-1}z) \ell(x, dz) &= \sum_{k=2}^n \binom{n}{k} x^{n-k} \int_{\mathbb{R}} z^k \ell(x, dz) \\ &= \sum_{k=2}^n \sum_{i=0}^k \binom{n}{k} \xi_i^k x^{n-k+i} = \sum_{i=0}^n \left( \sum_{k=\max(2,i)}^n \binom{n}{k} \xi_{k-i}^k \right) x^{n-i}, \end{aligned}$$

so that

$$\mathcal{G}x^n = \left( nb_1 + \frac{1}{2}n(n-1)\sigma_2 \right) x^n + \left( nb_0 + \frac{1}{2}n(n-1)\sigma_1 \right) x^{n-1} + \sum_{i=0}^n \left( \sum_{k=\max(2,i)}^n \binom{n}{k} \xi_{k-i}^k \right) x^{n-i},$$

which must be rearranged to collect the coefficients of  $x^k$ ,  $k = 0, \dots, n$ , to be inserted in the last row of  $G_n$ . This leads to  $(a_n^n, a_n^{n-1}, \dots, a_n^1, c_n)^\top$  as defined in the theorem and concludes the proof.  $\square$

**Proof of Theorem 5.3**

*Proof.* We first prove the base case  $n = 1$ . We consider the definition of matrix exponential as infinite sum of powers. If  $b_1 \neq 0$  then for every  $k \geq 1$ ,  $G_1^k = \begin{pmatrix} 0 & 0 \\ b_0 b_1^{k-1} & b_1^k \end{pmatrix}$ ; if  $b_1 = 0$  then for every  $k \geq 2$ ,  $G_1^k = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ . With  $G_1^0 = I_2$ , for  $b_1 \neq 0$  we get:

$$\begin{aligned} e^{G_1 t} &= \sum_{k=0}^{\infty} \frac{(G_1 t)^k}{k!} = I_2 + \sum_{k=1}^{\infty} \frac{t^k}{k!} \begin{pmatrix} 0 & 0 \\ b_0 b_1^{k-1} & b_1^k \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ b_0 \sum_{k=1}^{\infty} \frac{t^k b_1^{k-1}}{k!} & 1 + \sum_{k=1}^{\infty} \frac{t^k b_1^k}{k!} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{b_0}{b_1} (e^{b_1 t} - 1) & e^{b_1 t} \end{pmatrix}, \end{aligned}$$

and similarly for  $b_1 = 0$ , so that the base case is proved.

We now set  $n > 1$ . For  $\Lambda_n := c_n I_n - G_{n-1}$ , it is easy to verify that for every  $k \geq 1$  the powers of  $G_n$  are given by

$$G_n^k = \begin{pmatrix} G_{n-1}^k & \vec{0}_n \\ \vec{a}_n^\top \Lambda_n^{-1} (c_n^k I_n - G_{n-1}^k) & c_n^k \end{pmatrix},$$

provided that  $\Lambda_n$  is invertible. More precisely, as a consequence of the recursion formula for  $G_n$ , the power  $G_n^k$  involves  $\Lambda_n^{-1}$ , but it also involves  $G_{n-1}^k$ , that means it involves  $\Lambda_{n-1}^{-1}$ , and so on. Thus the matrices  $\Lambda_r^{-1}$  are all involved in  $G_n^k$  for every  $2 \leq r \leq n$ . In particular, we need all these to be invertible. For a fixed  $r$ , the determinant of  $\Lambda_r$  must then be different from zero for every  $2 \leq r \leq n$ . In particular, since  $G_{r-1}$  is a (lower) triangular matrix,  $\Lambda_r$  is a (lower) triangular matrix with determinant given by

the product of the elements on the main diagonal. By Theorem 5.1 and equation 5.2 we then get:

$$\det(c_r I_r - G_{r-1}) = c_r \prod_{j=1}^{r-1} (c_r - c_j) \neq 0 \quad (\text{B.15})$$

where  $c_1 = b_1$ . Condition 5.15 is equivalent to ask that  $c_r \neq 0$  and  $c_r \neq c_j$  for every  $1 \leq j \leq r-1$ . One easily notice that asking these conditions to hold for every  $2 \leq r \leq n$  means to ask that all the coefficients  $\{c_j\}_{j=2}^n$  are not null, and that  $\{c_j\}_{j=1}^n$  are all different among each others. This coincides with condition 5.3.

Then, with  $G_n^0 = I_{n+1}$ , we get:

$$\begin{aligned} e^{G_n t} &= \sum_{k=0}^{\infty} \frac{(G_n t)^k}{k!} = I_{n+1} + \sum_{k=1}^{\infty} \frac{t^k}{k!} \begin{pmatrix} G_{n-1}^k & \vec{0}_n \\ c_n^k I_n - G_{n-1}^k & c_n^k \end{pmatrix} \\ &= \begin{pmatrix} I_n + \sum_{k=1}^{\infty} \frac{(G_{n-1} t)^k}{k!} & 0 \\ \vec{a}_n^\top \Lambda_n^{-1} \sum_{k=1}^{\infty} \frac{t^k}{k!} (c_n^k I_n - G_{n-1}^k) & 1 + \sum_{k=1}^{\infty} \frac{t^k c_n^k}{k!} \end{pmatrix} \\ &= \begin{pmatrix} e^{G_{n-1} t} & \vec{0}_n \\ \vec{a}_n^\top \Lambda_n^{-1} \left( I_n \sum_{k=1}^{\infty} \frac{(c_n t)^k}{k!} - \sum_{k=1}^{\infty} \frac{(G_{n-1} t)^k}{k!} \right) & e^{c_n t} \end{pmatrix} = \begin{pmatrix} e^{G_{n-1} t} & \vec{0}_n \\ \vec{a}_n^\top \Lambda_n^{-1} (e^{c_n t} I_n - e^{G_{n-1} t}) & e^{c_n t} \end{pmatrix} \end{aligned}$$

which proves the matrix exponential formula and concludes the proof.  $\square$

#### Proof of Lemma 5.4

*Proof.* We consider the definition of  $c_j$  in equation 5.1: if  $\ell(x, dz) \equiv 0$ , then for every  $j = 2, \dots, n$ ,  $c_j = j b_1 + \frac{1}{2} j(j-1) \sigma_2$ , so that the first condition in 5.3 equals  $b_1 \neq -\frac{(j-1)}{2} \sigma_2$ , equivalent also to

$$b_1 \neq -\frac{k}{2} \sigma_2, \text{ for every } 1 \leq k \leq n-1. \quad (\text{B.16})$$

In the second condition in 5.3, we require  $c_j \neq c_i$ ,  $1 \leq j < i \leq n$ , that is

$$j b_1 + \frac{1}{2} j(j-1) \sigma_2 \neq i b_1 + \frac{1}{2} i(i-1) \sigma_2, \text{ for every } 1 \leq j < i \leq n,$$

which, after some simplifications, can be rewritten as  $b_1 \neq -\frac{(j+i-1)}{2} \sigma_2$  for every  $1 \leq j < i \leq n$ . In particular, since  $3 \leq j+i \leq 2n-1$ , this is also equivalent to

$$b_1 \neq -\frac{k}{2} \sigma_2, \text{ for every } 2 \leq k \leq 2(n-1).$$

Adding this to the condition previously found in 5.16, we conclude the proof.  $\square$

#### Proof of Proposition 5.5

*Proof.* By equations 5.4 and 5.5 and linearity of the extended generator  $\mathcal{G}$ , we write the equalities

$$J_n M_n H_n(x) = J_n Q_n(x) = \mathcal{G} Q_n(x) = \mathcal{G} (M_n H_n(x)) = M_n \mathcal{G} H_n(x) = M_n G_n H_n(x).$$

By comparing the first and the last terms we get  $J_n M_n = M_n G_n$ , which, rearranged, gives the first equality of the proposition. As a direct consequence, from the definition of exponential function as infinite sum of powers and by the identity  $M_n^{-1} M_n = I_n$ , we get also the second equality.  $\square$

#### Proof of Proposition 6.1

*Proof.* From Theorem 4.5 we easily notice that the only dependence of  $C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)$  on the initial condition  $Y(t)$  is inside the matrix of function  $X_n^{(m)}(Y(t))$ , which is defined by equation 4.1 as the  $m$ -th Kronecker product  $X_n^{(m)}(x) = H_n(x)^\top \otimes^m H_n(x)$ . By associativity property of the



Kronecker product, we write that  $X_n^{(m)}(x) = H_n(x)^\top \otimes X_n^{(m-1)}(x)$ . By the product rule applied to  $H_n(Y(t))^\top \otimes X_n^{(m-1)}(Y(t))$ , we obtain the recursive formula for  $\frac{\partial X_n^{(m)}(Y(t))}{\partial Y(t)}$  as in the statement, where

$$\frac{\partial H_n(Y(t))}{\partial Y(t)} = \frac{\partial X_n^{(0)}(Y(t))}{\partial Y(t)} = (0, 1, 2Y(t), \dots, nY(t)^{n-1})$$

can be seen as the vector product of  $(0, 1, \dots, n)$  with  $(0, H_{n-1}(Y(t))^\top)^\top$ . This concludes the proof.  $\square$

### Proof of Proposition [6.2](#)

*Proof.* From the correlator formula in Theorem [4.5](#), we distinguish three different cases, namely  $j = 0$ ,  $1 \leq j < m$ , and  $j = m$ , which we shall analyse separately.

For  $j = 0$ , the time point  $s_0$  appears in  $C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)$  two times: once inside the curl parenthesis in the matrix exponential  $e^{\tilde{G}_n^{(m)}(s_0-t)}$ , and once in the product  $\prod_{k=1}^m$  for  $k = 1$  in the matrix exponential  $e^{\tilde{G}_n^{(m-1)\top}(s_1-s_0)}$ . By the product rule, one gets  $\Theta_0$ .

For  $1 \leq j < m$ , the time point  $s_j$  also appears in  $C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)$  two times, both in the product  $\prod_{k=1}^m$ , first for  $k = j$  in the matrix exponential  $e^{\tilde{G}_n^{(m-j)\top}(s_j-s_{j-1})}$ , and then for  $k = j + 1$  in the matrix exponential  $e^{\tilde{G}_n^{(m-j-1)\top}(s_{j+1}-s_j)}$ . By the product rule, one gets  $\Theta_j$ .

For  $j = m$ , the time point  $s_m$  appears in  $C_{p_0, \dots, p_m}(s_0, \dots, s_m; t)$  only one time, that is in the product  $\prod_{k=1}^m$  for  $k = m$  in the matrix exponential  $e^{\tilde{G}_n^{(0)\top}(s_m-s_{m-1})}$ . By differentiation, one gets  $\Theta_m$ .

This concludes the proof.  $\square$

## References

- [1] Akerer, Damien and Damir Filipovic (2020). Linear Credit Risk Models. *Finance and Stochastics* 24(1), 169-214.
- [2] Akerer, Damien, Damir Filipovic and Sergio Pulido (2018). The Jacobi stochastic volatility model. *Finance and Stochastics* 22(3), 667-700.
- [3] Akerer, Damien and Damir Filipovic (2020). Option Pricing with Orthogonal Polynomial Expansions. *Mathematical Finance* 30(1), 47-84.
- [4] Applebaum, David (2009). *Lévy Processes and Stochastic Calculus*. Second edition. Cambridge University Press.
- [5] Barndorff-Nielsen, Ole E., Fred E. Benth and Almut E.D. Veraart (2018). *Ambit Stochastics*. Springer International Publishing.
- [6] Benth, Fred E., Jan Kallsen, and Thilo Meyer-Brandis (2007). A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance* 14(2), 153-169.
- [7] Benth, Fred E., Martin Groth and Rodwell Kufakunesu (2007). Valuing volatility and variance swaps for a non-Gaussian Ornstein-Uhlenbeck stochastic volatility model. *Applied Mathematical Finance* 14(4), 347-363.
- [8] Carr, Peter and Dilip B. Madan (1999). Option valuation using the fast Fourier transform. *The Journal of Computational Finance* 2(4), 61-73.
- [9] Cuchiero, Christa (2011). *Affine and polynomial processes*. PhD Thesis, ETH Zurich.
- [10] Cuchiero, Christa, Martin Keller-Ressel and Josef Teichmann (2012). Polynomial processes and their applications to mathematical finance. *Finance and Stochastics* 16(4), 711-740.
- [11] Cuchiero, Christa (2018). Polynomial processes in stochastic portfolio theory. *Stochastic processes and their applications* 129(5), 1829-1872.

- [12] Cuchiero, Christa and Sara Svaluto-Ferro (2021). Infinite-dimensional polynomial processes. *Finance and Stochastics* 25(2), 383-426.
- [13] Delbaen, Freddy and Hiroshi Shirakawa (2002). An interest rate model with upper and lower bounds. *Asia-Pacific Financial Markets* 9(3-4), 191-209.
- [14] Fasino, Dario (1995). Spectral properties of Hankel matrices and numerical solutions of finite moment problems. *Journal of Computational and Applied Mathematics* 65(1-3), 145-155.
- [15] Filipović, Damir, Martin Larsson and Anders B. Trolle (2016). Linear-rational term structure models. *The Journal of Finance* 72(2), 655-704.
- [16] Filipović, Damir and Martin Larsson (2016). Polynomial diffusions and applications in finance. *Finance and Stochastics* 20(4), 931-972.
- [17] Fiedler, Miroslav (1985). Polynomials and Hankel matrices. *Linear Algebra and its Applications* 66, 235-248.
- [18] Filipović, Damir and Martin Larsson (2020). Polynomial jump-diffusion models. *Stochastic Systems* 10(1), 71-97.
- [19] Golyandina, Nina, Vladimir Nekrutkin and Anatoly A. Zhigljavsky (2001). *Analysis of Time Series Structure: SSA and Related Techniques*. CRC press.
- [20] Hassani, Hossein and Dimitrios Thomakos (2010). A review on singular spectrum analysis for economic and financial time series. *Statistics and its Interface* 3(3), 377-397.
- [21] Higham, Nicholas J. (2005). The scaling and squaring method for the matrix exponential revisited. *SIAM Journal on Matrix Analysis and Applications* 26(4), 1179-1193.
- [22] Horn, Roger A. and Charles R. Johnson (1991). *Topics in Matrix Analysis*. Cambridge University Press.
- [23] Jain, Pooja and Ram B. Pachori (2014). Event-based method for instantaneous fundamental frequency estimation from voiced speech based on eigenvalue decomposition of the Hankel matrix. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22(10), 1467-1482.
- [24] Jain, Pooja and Ram B. Pachori (2015). An iterative approach for decomposition of multi-component non-stationary signals based on eigenvalue decomposition of the Hankel matrix. *Journal of the Franklin Institute* 352(10), 4017-4044.
- [25] Kemna, Angelien G.Z. and Ton A.C.F. Vorst (1990). A pricing method for options based on average asset values. *Journal of Banking & Finance* 14(1), 113-129.
- [26] Kleisinger-Yu, Xi, Vlatka Komaric, Martin Larsson and Markus Regez (2020). A multifactor polynomial framework for long-term electricity forwards with delivery period. *SIAM Journal on Financial Mathematics* 11(3), 928-957.
- [27] Kressner, Daniel, Robert Luce and Francesco Statti (2017). Incremental computation of block triangular matrix exponentials with application to option pricing. *Electronic Transactions on Numerical Analysis* 47, 57-72.
- [28] Lavagnini, Silvia (2021). Pricing Asian Options with Correlators. arXiv preprint arXiv:2104.11684.
- [29] Magnus, Jan R. and Heinz Neudecker (1980). The eliminating matrix: some lemmas and applications. *SIAM. Journal of Algebraic Discrete Methods* 1(4), 422-449.
- [30] Munkhammar, Joakim, Lars Mattsson and Jesper Rydén (2017). Polynomial probability distribution estimation using the method of moments. *PLOS ONE* 12(4), e0174573.
- [31] Peller, Vladimir (2012). *Hankel Operators and their Applications*. Springer Science & Business Media.

- [32] Townsend, Alex, Marcus Webb and Sheehan Olver (2018). Fast polynomial transforms based on Toeplitz and Hankel matrices. *Mathematics of Computation* 87(312), 1913-1934.
- [33] Ware, Tony (2019). Polynomial processes for power prices. *Applied Mathematical Finance* 26(5), 453-474.
- [34] Weron, Rafal (2007). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. Vol. 403. John Wiley & Sons.
- [35] Willems, Sander (2019). Asian option pricing with orthogonal polynomials. *Quantitative Finance* 19(4), 605-618.
- [36] Zhou, Hao (2003). Itô conditional moment generator and the estimation of short-rate processes. *Journal of Financial Econometrics* 1(2), 250-271.