Pricing Asian Options with Correlators

Silvia Lavagnini
Department of Mathematics, University of Oslo
0316 Blindern, Norway, silval@math.uio.no

February 17, 2022

Abstract

We derive a series expansion by Hermite polynomials for the price of an arithmetic Asian option. This requires the computation of moments and correlators of the underlying asset price which for a polynomial jump-diffusion process are given analytically, hence no numerical simulation is required to evaluate the series. This allows to derive analytical expressions for the option Greeks. The weight function defining the Hermite polynomials is a Gaussian density with scale b. We find that the rate of convergence of the series depends on b, for which we prove a lower bound to guarantee convergence. Numerical examples show that the series expansion is accurate but unstable for initial values of the underlying process far from zero, mainly due to rounding errors.

Keywords Asian option; Option pricing; Greeks; Orthogonal polynomials; Hermite polynomials; Polynomial jump-diffusion process; Correlators.

1 Introduction

Asian options are path-dependent options whose payoff is a function of the (discrete or continuous) average underlying price. This kind of derivatives is bought and sold in, e.g., currency, interest rate, energy and insurance markets. Within the energy markets, for example, Asian options were traded a decade ago at Nord Pool, the Nordic commodity market for electricity [26]. However, because of their path-dependent nature, their valuation is not straightforward, and, in particular, no closed pricing formula is available in general, neither in the standard Black and Scholes setting. We focus here on discrete averaging which is the normal specification in real contracts. In this case, the payoff depends on the average of the underlying asset price over some prespecified period of time, usually a low number of trading days. In particular, we derive a series representation for the option price functional with polynomials which are orthogonal with respect to a Gaussian density function. By modeling the underlying asset price with a polynomial jump-diffusion process, all of the terms in the series expansion can be computed analytically thanks to the moment and correlator formulas, hence no numerical simulation is required.

In fact, several different approaches have been proposed for solving the Asian option pricing problem. In the Black and Scholes setting, the payoff function is a sum of correlated lognormal distributions, for which there is no recognizable density function. Some authors have then proposed to approximate the unknown distribution of the average price with the maximum entropy approach [13] or via the Edgeworth series expansion approach [18, 21], by exploiting that the density law of the logarithm of the arithmetic average is uniquely determined by its moments. Other approaches to evaluate Asian options are via Monte Carlo simulations with variance reduction techniques [16, 17] or via Fourier transform [11]. In some other cases, the authors have derived exact representations for the pricing functional, for example as a triple integral to be evaluated numerically [27], by the Laplace transform [14] or Taylor expansion [15]. Another part of the literature has worked instead on providing lower and upper bounds for the prices of arithmetic Asian options [2, 12, 22].

The part of literature that we are contributing to concerns with obtaining exact series expansions for distributions or option values by orthogonal polynomials, such as the Laguerre polynomials [7]. In particular, a very recent branch of literature has combined this technique with the choice of an

affine or polynomial stochastic process for the underlying asset price in order to obtain analytic price representations [1, 9, 10, 25]. The general idea is to define an auxiliary probability density and introduce the corresponding weighted Hilbert space. One then finds an orthonormal basis of polynomials for the Hilbert space and compute the series expansion for the option price by projecting the price functional on the space of polynomials. The auxiliary density could be, for example, a finite Gaussian mixture density [1], a mixture of log-normal densities [25] or a gamma density [10]. Some of the desirable properties for the auxiliary density are listed in [9, Sec. 7].

Our approach is partly similar to the one in [25], in the sense that we derive a series representation for the option price functional with orthogonal polynomials and we work with polynomial processes. However, there are at least two big fundamental differences. In [25] the underlying spot price is considered to follow a geometric Brownian motion, and, by the time-reversal property of Brownian motions (see [5]), they derive a stochastic differential equation (SDE) whose solution process has the same distribution of the average price process (scaled by the terminal time). Specifically, the SDE defines a polynomial diffusion, so that the moments of the average price process can be computed in closed form by the moment formula for polynomial processes. Since the SDE is driven by a Brownian motion, this framework excludes the possibility of discontinuities in the asset price paths. We thus model the asset price with a polynomial jump-diffusion directly. Hence, on the one hand, we allow for discontinuities, and, on the other hand, the moment formula for polynomial processes is used for computing the moments of the underlying spot price. Since we still need to compute the moments of the average price process, we achieve this by the multinomial theorem and the correlator formula for polynomial processes derived in [4].

We fix a stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$, with $\mathbb{Q}$ a risk-neural measure, and we work in a one-dimensional setting. We want to price the fixed-strike call-style Asian option defined by

$$\Pi_K(t) := e^{-r(T-t)}\mathbb{E}\left[\varphi_K(X(T))\middle|\,\mathcal{F}_t\right] \qquad \text{with} \qquad \varphi_K(x) := \max(x-K,0), \tag{1.1}$$

where $K > 0$ is the strike price, $r \geq 0$ the risk-free interest rate and X is the discrete average of a stochastic process Y over the period $(t, T]$, namely

$$X(T) = \frac{1}{m+1}\sum_{j=0}^{m} Y(s_j) \qquad \text{for } t < s_0 < s_1 < \cdots < s_m = T \text{ and } m \geq 0. \tag{1.2}$$

We point out that one can similarly consider X to be the continuous average $X(T) = \frac{1}{T-t}\int_t^T Y(s)ds$. Then, for a discrete sampling $t < s_0 < s_1 < \cdots < s_m = T$ with constant time step $\Delta := \frac{T-t}{m+1}$ small enough, the integral $\int_t^T Y(s)ds$ can be reasonably well approximated with the sum $\sum_{j=1}^m Y(s_j)\Delta$. This indeed coincides with Eq. (1.2) for $\Delta = \frac{T-t}{m+1}$.

We consider Y to be a polynomial process in the sense introduced by [9]. The idea is to derive the series representation of the payoff function φ_K in terms of Hermite polynomials. More precisely, we shall introduce the *generalized Hermite polynomials* that form a basis for the space

$$L^2\left(\mathbb{R}, \omega_{a,b}(x)dx\right) \qquad \text{with} \qquad \omega_{a,b}(x) = \exp\left(-\frac{(x-a)^2}{2b^2}\right),\ a, b \in \mathbb{R}, b > 0. \tag{1.3}$$

After evaluating the series at $x = X(T)$ in Eq. (1.2), we obtain an infinite sum of polynomial functions in $X(T)$. The price of the Asian option is then given by the discounted expected value of this infinite sum. By the multinomial theorem, we rewrite the terms of the sum as a linear combination of correlator-type terms in the sense of [4], that is, terms of the form

$$\mathbb{E}\left[Y(s_0)^{k_0}Y(s_1)^{k_1}\cdots Y(s_m)^{k_m}\middle|\,\mathcal{F}_t\right], \tag{1.4}$$

which we compute by the closed formula for correlators [4, Theorem 4.5].

This procedure gives an exact formula for pricing discrete Asian options. For numerical purposes, we truncate the infinite summation to a certain $N > 0$ and we study the approximation error with respect to the three parameters involved, namely N, a and b. In particular, we find a lower bound for the scale b in order to achieve convergence. This bound is proved under the strong hypothesis that the

tails of the distribution of $X(T)$ are flatter than the ones of the weight function $\omega_{a,b}$, and it depends on the variance of $X(T)$. We confirm our findings with numerical examples and we compare the results with a Monte-Carlo-simulation approach. This shows that the Hermite series can reach much higher accuracy than Monte Carlo. However, numerical instabilities are observed, mainly due to the intrinsic exploding nature of polynomial functions of high order and to rounding errors. In particular, these are more likely to happen when the initial point of the underlying spot price process Y is far out from 0.

The models considered for the numerical experiments in Sec. 5 are a simple Brownian motion, an Ornstein-Uhlenbeck (OU) process, and a jump-diffusion process whose SDE is basically given by the SDE of the OU plus a normal inverse Gaussian (NIG) jump term. In the first two cases, the asset price follows a Gaussian distribution, and in the last case the asset follows a distribution (that we could define) as of a Gaussian plus a NIG term. In particular, we find that for these three cases the behavior of the approximating series is similar, even when adding the jump component. On the other hand, in Sec. 6 we consider the Black and Scholes setting, hence the process Y follows a log-normal distribution. In this case, the approximation works well only if the parameters are chosen in such a way that the corresponding log-normal distribution can be well approximated with a Gaussian one. This, in particular, is related to the coefficient of variation of the process. In all other cases, a different weight function should be considered.

Due to computational constraints, we can not run experiments for $m > 2$. This means that in the case of the continuous averaging that we described above, approximating the integral with a discrete sum would not be efficient. An alternative approach adopted in Sec. 6 is to consider the bivariate process (Y, X), with $X(T) = \frac{1}{T-t}\int_t^T Y(s)ds$. As this turns out to be a bivariate jump-diffusion process, the moment formula in $\mathbb{R}^2$ can be used to compute the terms in the approximating series with Hermite polynomials. In particular, this requires the generator matrix associated with the bivariate process (Y, X), for which we derive a recursion formula in Appendix A. This extends the formula obtained in [4] for one-dimensional processes, although omitting the jump-component.

The rest of the paper is organized as follows. In Sec. 2 we briefly introduce polynomial processes and recall the moment and correlator formulas. In Sec. 3 we introduce the family of generalized Hermite polynomials and we derive the series expansion for a call-payoff function, also studying the approximation error as a function of the truncation number. In Sec. 4 we derive the option price approximation, first for a European-style option and then for an Asian option with discrete sampling. We also analyze the approximation error and derive explicit representations for two of the option Greeks. Finally in Sec. 5 we show some numerical examples and in Sec. 6 we compare our approach with some others from the literature. Sec. 7 summarizes the findings. In Appendix A we construct the generator matrix for a two-dimensional diffusion process, while Appendix B contains the proofs of the principal results and Appendix C some definitions for understanding the correlator formula. The code for the experiments of Sec. 5 and 6 is implemented in Python and is available at https://github.com/silvialava/Pricing_options_with_correlators.git.

2 Polynomial Processes and Correlator Formula

Let $\mathrm{Pol}_n(\mathbb{R})$ be the space of all polynomials on $\mathbb{R}$ with degree less than or equal to n. Following [9], we consider a jump-diffusion operator on $\mathbb{R}$ of the form

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

for some measurable maps $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$, and a transition kernel $\ell : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, such that the conditions in [9, Lemma 1] are satisfied, namely

$$b \in \mathrm{Pol}_1(\mathbb{R}), \qquad \sigma^2 + \int_{\mathbb{R}} z^2\ell(\cdot, dz) \in \mathrm{Pol}_2(\mathbb{R}) \quad \text{and} \quad \int_{\mathbb{R}} z^m\ell(\cdot, dz) \in \mathrm{Pol}_m(\mathbb{R}) \text{ for all } m \geq 3. \tag{2.2}$$

Under conditions (2.2), the operator $\mathcal{G}$ is called *polynomial* in the sense of [9, Definition 1], and the process Y having $\mathcal{G}$ as extended generator is a *polynomial jump-diffusion process*. Moreover, a polynomial generator $\mathcal{G}$ can be expressed in matrix form. Its matrix representation is called the *generator matrix* and it strictly depends on the polynomial basis of choice.

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

$$H_n : \mathbb{R} \longrightarrow \mathbb{R}^{n+1}, \quad H_n(x) = (1, x, x^2, \dots, x^n)^\top. \tag{2.3}$$

For every $n \geq 1$, the generator matrix associated with $\mathcal{G}$ is thus the matrix $G_n \in \mathbb{R}^{(n+1)\times(n+1)}$ satisfying $\mathcal{G}H_n(x) = G_n H_n(x)$. For Y polynomial process and $p \in \mathrm{Pol}_n(\mathbb{R})$, the conditional expectation $\mathbb{E}[p(Y(T))|\mathcal{F}_t]$ is then a polynomial function in $Y(t)$, $0 \leq t \leq T$, and is given in closed form in [9, Theorem 1] and in the following theorem for completeness.

Theorem 2.1. *For Y polynomial process and $p \in \mathrm{Pol}_n(\mathbb{R})$, the following moment formula holds:*

$$\mathbb{E}[p(Y(T))|\mathcal{F}_t] = \vec{p}_n^{\top} e^{G_n(T-t)} H_n(Y(t)), \qquad 0 \leq t \leq T, \tag{2.4}$$

with $\vec{p}_n \in \mathbb{R}^{n+1}$ the vector of coefficients of p with respect to $H_n(x)$.

Let now $\vec{e}_{n,j}$ denote the j-th euclidean basis vector in $\mathbb{R}^n$. As a consequence of Theorem 2.1, the n-th conditional moment of $Y(T)$ can be computed explicitly as a linear combination of powers of $Y(t)$.

Corollary 2.2. *For every $n \geq 0$, we get that $\mathbb{E}[Y(T)^n|\mathcal{F}_t] = \vec{e}_{n+1,n+1}^{\top}\, e^{G_n(T-t)} H_n(Y(t))$.*

In [4, Theorem 4.8], the moment formula is extended to correlators, namely conditional expectations of products of polynomial functions in the polynomial process Y evaluated at different time points. We shall not discuss here the meaning of the symbols appearing in the formula. However the main details can be found in Appendix C. For more interested readers, we refer to [4].

Theorem 2.3. *For $m \geq 1$, we consider $m+1$ polynomial functions $p_k \in \mathrm{Pol}_{n_k}(\mathbb{R})$, $k = 0, \dots, m$, in the polynomial process Y, evaluated at different time points, $t < s_0 < s_1 < \cdots < s_m$, and with vector of coefficients $\vec{p}_{k,n} \in \mathbb{R}^{n+1}$, for $n = \max\{n_0, \dots, n_m\}$. 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 the following expectation formula holds:*

$$\begin{aligned}\mathbb{E}[p_m(Y(s_0))\, p_{m-1}(Y(s_1)) \cdot \cdots \cdot p_0(Y(s_m))|\mathcal{F}_t] \\ = \vec{p}_{m,n}^{\top} \left\{ vec^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0-t)} \circ vec\left(H_n(Y(t))^\top \otimes^m H_n(Y(t))\right)\right\} \\ \cdot \prod_{k=1}^{m} e^{\tilde{G}_n^{(m-k)\top}(s_k - s_{k-1})} \left\{ I_{n+1} \otimes^{m-k} \vec{p}_{m-k,n} \right\},\end{aligned} \tag{2.5}$$

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 $\tilde{G}_n^{(0)} = G_n$. Here $I_{n+1} \in \mathbb{R}^{(n+1)\times(n+1)}$ denotes the identity matrix.

Corollary 2.4. *For every $n \geq 0$ and $0 \leq k_0, k_1, \dots, k_m \leq n$, the following formula holds:*

$$\begin{aligned}\mathbb{E}\left[Y(s_0)^{k_0} Y(s_1)^{k_1} \cdots Y(s_m)^{k_m} \middle| \mathcal{F}_t\right] \\ = \vec{e}_{n+1,k_0+1}^{\top} \left\{ vec^{-1} \circ e^{\tilde{G}_n^{(m)}(s_0-t)} \circ vec\left(H_n(Y(t))^\top \otimes^m H_n(Y(t))\right)\right\} \\ \cdot \prod_{j=1}^{m} e^{\tilde{G}_n^{(m-j)\top}(s_j - s_{j-1})} \left\{ I_{n+1} \otimes^{m-j} \vec{e}_{n+1,k_j+1} \right\}.\end{aligned}$$

3 Payoff Representation with Hermite Polynomials

We shall construct a polynomial approximation for the payoff function φ_K in Eq. (1.1). In order to do that, we first introduce an orthogonal polynomial basis for $\mathrm{Pol}_n(\mathbb{R})$. Let $q_0(x), q_1(x), \dots$ be orthogonal polynomial functions with values in $\mathbb{R}$ such that the family $\{q_0(x), q_1(x), \dots, q_n(x)\}$ forms a basis for $\mathrm{Pol}_n(\mathbb{R})$, and let

$$Q_n : \mathbb{R} \longrightarrow \mathbb{R}^{n+1}, \quad Q_n(x) = (q_0(x), q_1(x), \dots, q_n(x))^\top. \tag{3.1}$$

From classical linear algebra, there exists a matrix

$$M_n \in \mathbb{R}^{(n+1)\times(n+1)} \quad \text{such that} \quad M_n H_n(x) = Q_n(x) \quad \text{and} \quad M_n^{-1} Q_n(x) = H_n(x). \tag{3.2}$$

By Eq. (3.2) we can exploit both the readability of $H_n(x)$ and the orthogonality of $Q_n(x)$.

3.1 Generalized Hermite polynomials

We restrict our attention to the *probabilistic Hermite polynomials* (which we shall refer to simply as *Hermite polynomials*) defined by

$$q_n(x) := (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2}}, \quad n \geq 0. \tag{3.3}$$

The family $\{q_n\}_{n\geq 0}$ forms an orthogonal basis for the Hilbert space $L^2(\mathbb{R}, w(x)dx)$ with weight function $w(x) := e^{-\frac{x^2}{2}}$. The norm of q_n in $L^2(\mathbb{R}, w(x)dx)$ is

$$\|q_n\|^2_{L^2(\mathbb{R},w(x)dx)} = \int_{-\infty}^{\infty} q_n^2(x) w(x) dx = \sqrt{2\pi} n! \,. \tag{3.4}$$

It is easy to check that $\varphi_K \in L^2(\mathbb{R}, w(x)dx)$. However, the weight function w is centered in $x = 0$, which means that an approximation with Hermite polynomials will have the main focus in a neighborhood of $x = 0$ and will potentially not be good for points far from it. Since we want to approximate the payoff function φ_K whose most interesting point is $x = K$, we thus need a weight function possibly centered in $x = K$. Alternatively, in view of option pricing where φ_K is evaluated in a random variable X, one might want to focus around the mean of X. We should then consider a weight function that allows to shift the focus of the approximation to the area of greatest interest.

To keep it general, for $a, b \in \mathbb{R}$, $b > 0$, we introduce a family of weight functions and the corresponding orthogonal polynomials by

$$w_{a,b}(x) := e^{-\frac{(x-a)^2}{2b^2}} \qquad \text{and} \qquad q_n^{a,b}(x) := (-1)^n e^{\frac{(x-a)^2}{2b^2}} \frac{d^n}{dx^n} e^{-\frac{(x-a)^2}{2b^2}}. \tag{3.5}$$

The family $\{q_n^{a,b}\}_{n\geq 0}$ forms an orthogonal basis for the Hilbert space $L^2(\mathbb{R}, w_{a,b}(x)dx)$ with norm

$$\|f\|^2_{L^2(\mathbb{R},w_{a,b}(x)dx)} := \int_{\mathbb{R}} f(x)^2 w_{a,b}(x) dx \quad \text{for} \quad f \in L^2(\mathbb{R}, w_{a,b}(x)dx). \tag{3.6}$$

The norm of $q_n^{a,b}$ in $L^2(\mathbb{R}, w_{a,b}(x)dx)$ is given in the following lemma.

Lemma 3.1. *For every $n \geq 0$, the norm of $q_n^{a,b}$ in $L^2(\mathbb{R}, w_{a,b}(x)dx)$ is $\left\|q_n^{a,b}\right\|^2_{L^2(\mathbb{R},w_{a,b}(x)dx)} = \frac{\sqrt{2\pi}n!}{b^{2n-1}}$.*

We shall from now on refer to a as the *drift* and to b as the *scale*, while to $\{q_n^{a,b}\}_{n\geq 0}$ as *generalized Hermite polynomials* (GHPs). We also introduce the notation $L^2_{a,b} := L^2(\mathbb{R}, w_{a,b}(x)dx)$, where $L^2_{0,1} = L^2(\mathbb{R}, w(x)dx)$. We point out that, while we use the terminology "weight function", "approximating series", etc., we deal in practice with a family of weight functions, a family of approximating series, etc., depending on the choice of the parameters a and b.

3.2 Series expansion for the call-payoff function

We introduce $\varphi_K^{a,b}$ as the series representation of φ_K in terms of the GHPs $\{q_n^{a,b}\}_{n\geq 0}$, namely

$$\varphi_K^{a,b}(x) := \sum_{n=0}^{\infty} \frac{\left\langle \varphi_K, q_n^{a,b} \right\rangle_{L^2_{a,b}}}{\left\| q_n^{a,b} \right\|^2_{L^2_{a,b}}} q_n^{a,b}(x) = \sum_{n=0}^{\infty} \frac{b^{2n-1}}{\sqrt{2\pi} n!} \int_{-\infty}^{\infty} \varphi_K(y) q_n^{a,b}(y) w_{a,b}(y) dy \; q_n^{a,b}(x), \tag{3.7}$$

which we shall compute explicitly. From now on, we denote with ϕ and Φ, respectively, the probability density function and the cumulative distribution function of a standard Gaussian random variable.

Proposition 3.2. *The series $\varphi_K^{a,b}$ can be written in terms of the Hermite polynomials $\{q_n\}_{n\geq 0}$ by*

$$\varphi_K^{a,b}(x) = \sum_{n=0}^{\infty} \beta_n^{a,b} q_n\left(\frac{x-a}{b}\right) \text{ with } \beta_n^{a,b} := \begin{cases} b\,\phi\left(\frac{K-a}{b}\right) + (a-K)\left(1-\Phi\left(\frac{K-a}{b}\right)\right) & \text{for } n=0 \\ b\left(1-\Phi\left(\frac{K-a}{b}\right)\right) & \text{for } n=1 \\ \frac{b}{n!}\phi\left(\frac{K-a}{b}\right) q_{n-2}\left(\frac{K-a}{b}\right) & \text{for } n\geq 2 \end{cases}. \quad (3.8)$$

Example 3.1. Let X be a random variable with mean and variance denoted respectively with μ and σ^2. We then consider the drift $a=\mu$ and the scale $b=\sigma$. From Proposition 3.2 we get

$$\varphi_K^{\mu,\sigma}(x) = \sigma\,\phi\left(\frac{K-\mu}{\sigma}\right) + \left(1-\Phi\left(\frac{K-\mu}{\sigma}\right)\right)(x-K) + \sum_{n=2}^{\infty} \beta_n^{\mu,\sigma} q_n\left(\frac{x-\mu}{\sigma}\right). \quad (3.9)$$

If X follows a Gaussian distribution, then by computing the expectation of $\varphi_K^{\mu,\sigma}(X)$ we get that

$$\begin{aligned} \mathbb{E}\left[\varphi_K^{\mu,\sigma}(X)\right] &= \sigma\,\phi\left(\frac{K-\mu}{\sigma}\right) + \left(1-\Phi\left(\frac{K-\mu}{\sigma}\right)\right)(\mu-K) + \mathbb{E}\left[\sum_{n=2}^{\infty} \beta_n^{\mu,\sigma} q_n\left(\frac{X-\mu}{\sigma}\right)\right] \\ &= \mathbb{E}\left[\varphi_K(X)\right] + \mathbb{E}\left[\sum_{n=2}^{\infty} \beta_n^{\mu,\sigma} q_n\left(\frac{X-\mu}{\sigma}\right)\right]. \end{aligned} \quad (3.10)$$

In this case the weight function $\omega_{\mu,\sigma}$ coincides with the density function of X. Hence, calculating the expectation $\mathbb{E}\left[\varphi_K(X)\right]$ by computing $\mathbb{E}\left[\varphi_K^{\mu,\sigma}(X)\right]$ might add uncertainty to the result, unless the coefficients $\beta_n^{\mu,\sigma}$ are non significant for $n\geq 2$.

Example 3.2. For X as in Example 3.1, we let the drift be $a=K$ and the scale be $b=\sigma$. Then

$$\varphi_K^{K,\sigma}(x) = \frac{\sigma}{\sqrt{2\pi}} + \frac{x-K}{2} + \sum_{n=2}^{\infty} \beta_n^{K,\sigma} q_n\left(\frac{x-K}{\sigma}\right) \qquad \text{with } \beta_n^{K,\sigma} = \frac{\sigma}{\sqrt{2\pi n!}} q_{n-2}(0). \quad (3.11)$$

If n is an odd number then $\beta_n^{K,\sigma}=0$ because the Hermite polynomials of odd orders have no constant term. More precisely, for every $n\geq 2$, we introduce $k\geq 1$ as the integer such that either $n=2k$ or $n=2k+1$. Then, the coefficients $\beta_n^{K,\sigma}$ in Eq. (3.11) can be rewritten as

$$\beta_n^{K,\sigma} = \begin{cases} \frac{(-1)^{k-1}\sigma}{\sqrt{2\pi k!}(2k-1)2^k} & \text{for } n=2k \\ 0 & \text{for } n=2k+1 \end{cases}. \quad (3.12)$$

This is obtained by observing that $q_n(0) = (-1)^{\frac{n}{2}} \frac{n!}{2^{n/2}k!}$ for $n=2k$ even, and $q_n(0)=0$ for n odd, where !! denotes the double factorial. Then $\beta_n^{K,\sigma}=0$ for n odd, while for n even we write that

$$\beta_n^{K,\sigma} = \frac{\sigma}{\sqrt{2\pi n!}} q_{n-2}(0) = \frac{(-1)^{\frac{n-2}{2}}(n-3)!!\,\sigma}{\sqrt{2\pi n!}} = \frac{(-1)^{\frac{2k-2}{2}}(2k-3)!!\,\sigma}{\sqrt{2\pi}(2k)!} = \frac{(-1)^{k-1}(2k-1)!!\,\sigma}{\sqrt{2\pi}(2k)!(2k-1)}. \quad (3.13)$$

In particular, $(2k-1)!! = \frac{(2k-1)!}{2^{k-1}(k-1)!}$, so that, after simplification, we obtain (3.12).

3.3 Error analysis

For computational purposes, we truncate the infinite sum defining $\varphi_K^{a,b}$ to a certain N big enough so that the resulting series well approximates the payoff function φ_K. This leads to introduce

$$\varphi_{K,N}^{a,b}(x) := \sum_{n=0}^{N} \beta_n^{a,b} q_n\left(\frac{x-a}{b}\right) = \boldsymbol{\beta}_N^{a,b\,\top} Q_N\left(\frac{x-a}{b}\right) = \boldsymbol{\beta}_N^{a,b\,\top} M_N H_N\left(\frac{x-a}{b}\right), \quad (3.14)$$

where $\boldsymbol{\beta}_N^{a,b} := (\beta_0^{a,b}, \beta_1^{a,b}, \ldots, \beta_N^{a,b})^\top$ and M_N is the matrix for the change of basis with respect to $H_N(x)$ in Eq. (3.2). We point out that $\varphi_{K,N}^{a,b}$ can be computed for any choice of the orthogonal basis $\{q_n\}_{n\geq 0}$. Then Eq. (3.14) holds with the obvious modifications for M_N and $\boldsymbol{\beta}_N^{a,b}$.

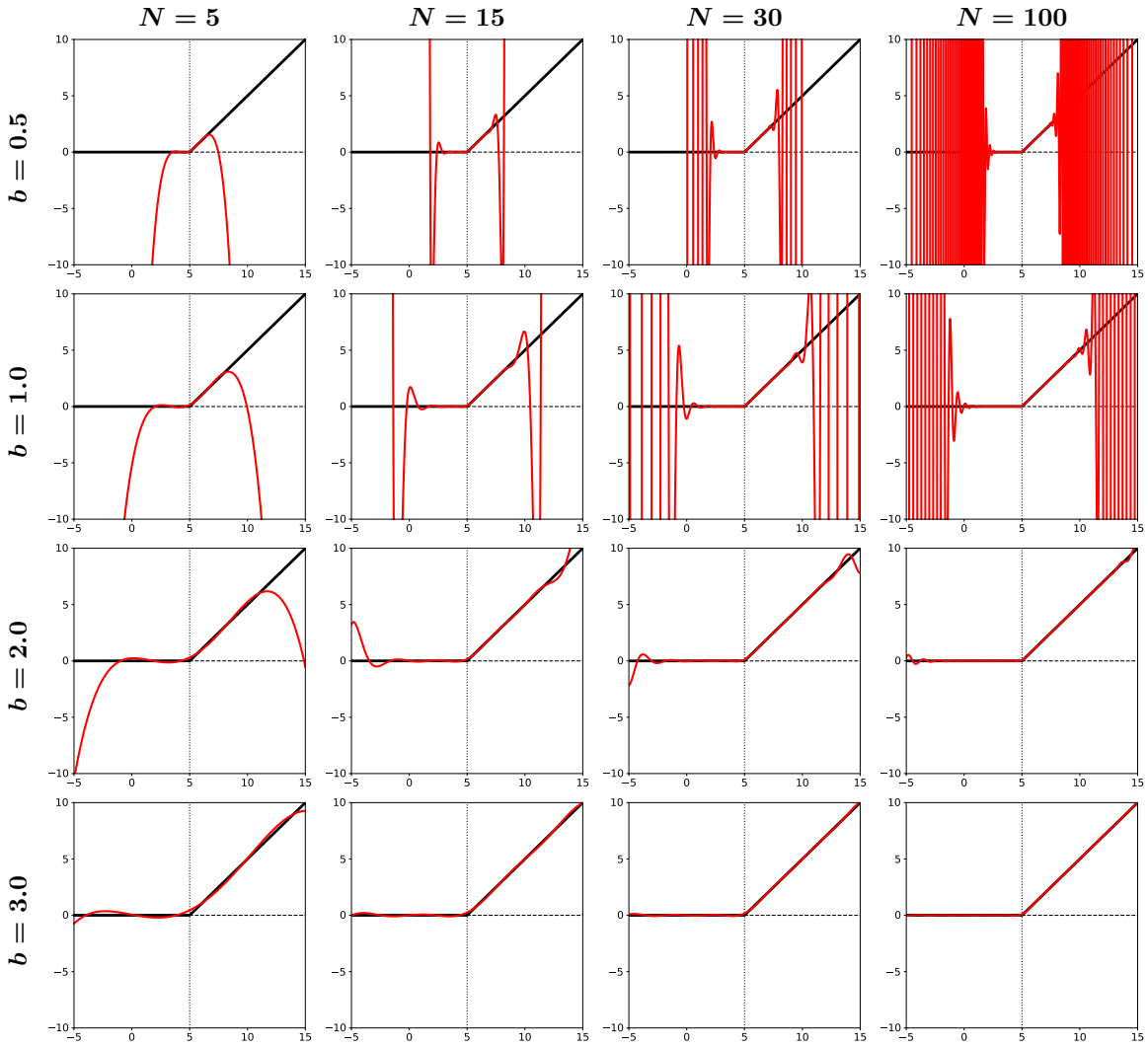

Figure 1: Approximation of the call payoff function φ_K with strike price $K = 5.0$ (black line) by generalized Hermite polynomials of different order N and scale b (red line). The drift is fixed to $a = K$.

In Figure 1 we observe the behavior of $\varphi^{a,b}_{K,N}$ for different values of $N \in \{5, 15, 30, 100\}$ and $b \in \{0.5, 1.0, 2.0, 3.0\}$. In particular, we fix the drift to $a = K$ so that the approximations are centered around the strike price value $K = 5.0$. The area where the Hermite series well approximates the payoff function gets wider when increasing the scale b. Similarly, increasing the truncation number N gives better performances, however this is more evident for higher values of b. On the other hand, including higher order polynomials in the series adds oscillations to the approximation. We point out that the value of the drift a is kept constant since the only effect of changing the drift is a shift in the center of the approximation, which is not particularly interesting. We stress the fact that $\varphi^{a,b}_{K,N}$ converges to φ_K in the norm of $L^2_{a,b}$, so that we cannot expect convergence in the supreme norm, as it can also be observed in Figure 1. The next proposition gives a semi-explicit formula for the $L^2_{a,b}$-norm of the approximation error.

Proposition 3.3. *By the Parseval identity, the norm in $L^2_{a,b}$ of the approximation error is*

$$\left\| \varphi_K - \varphi^{a,b}_{K,N} \right\|_{L^2_{a,b}} = b\,\phi\left(\frac{K-a}{b}\right) \sqrt{\sum_{n=N+1}^{\infty} \left(\frac{1}{n!} q_{n-2}\left(\frac{K-a}{b}\right)\right)^2}. \tag{3.15}$$

Proposition 3.3 does not give a precise intuition on the behavior of the approximation error as a function of N. We however observe that the quantity

$$h_K^{a,b} := b\,\phi\left(\frac{K-a}{b}\right) = \frac{b}{\sqrt{2\pi}}\exp\left(-\frac{(K-a)^2}{2b^2}\right) \tag{3.16}$$

does not depend on N, but it does depend on b. More precisely, if ignoring the dependence of the squared root in Eq. (3.15) on b (which is not straightforward), the approximation error is an increasing function of b. Hence, despite Figure 1 shows an improving in the approximation for larger values of the scale, the $L^2_{a,b}$-norm of the approximation error might grow with the scale b.

To analyze this further, in Figure 2 we plot the $L^2_{a,b}$-norm of the approximation error for different values of N, b and a, and $K = 5.0$. In the first row, the norm is a function of N for three different cases, namely $a = 5.0$, $a = 7.0$ and $a = 10.0$. Here $b \in \{0.5, 1.0, 2.0, 3.0, 6.0, 10.0\}$, however the lines are not distinguishable. In the second row we report a zoom of the three previous plots, where we focus on $10 \leq N \leq 30$. Here we distinguish the six different lines and observe in particular that the approximation error is smaller for smaller values of b. Finally, in the last row we plot the error together with $h_K^{a,b}$ as functions of b. Here we fix $N = 20$ and consider many values for the scale in the interval $0.5 \leq b \leq 10.0$. To calculate the infinite sum in Eq. (3.15), we truncate at $n = 160$ since for bigger values the factorial cannot be converted to a floating-point number.

For the plots in the first column, since $a = K = 5.0$, the coefficient $h_K^{a,b} = \frac{b}{\sqrt{2\pi}}$ is proportional to b and the squared root in Eq. (3.15) does not depend on b. We see indeed in the last plot of the first column that the approximation error is proportional to $h_K^{a,b}$ (proportional to b in fact). In the second and third plots of the last row, since $a \neq K$, the behavior of the approximation error diverges from the one of $h_K^{a,b}$. In particular, it is not monotone in b. However, since the main interest for the polynomial approximation is around K, we shall mostly deal with configurations that we can approximately consider increasing functions of the scale parameter.

4 Pricing Options with Correlators

We focus on the pricing of options with call-style payoff function. If the underlying spot process is evaluated in a single time point $T \geq t$, then the option is of European type. If the underlying spot process is averaged over the settlement period $(t, T]$, then the option is of Asian type. We shall see that the price formula for this second kind of contracts can be derived starting from the price formula for the first kind. We shall also derive explicit formulations for two of the Greeks of the Asian option.

4.1 European options

We consider $X = Y$, where Y is an $\mathcal{F}_t$-adapted stochastic process. For the payoff function φ_K in Eq. (1.1), we compute the conditional expectation $\Pi_K(t) = \mathbb{E}\left[\varphi_K(X(T)) \,|\, \mathcal{F}_t\right]$ with respect to the risk-neutral measure $\mathbb{Q}$. Starting from the Hermite series constructed in Sec. 3, we have a family of approximations depending on a, b and N, namely

$$\Pi^{a,b}_{K,N}(t) := \mathbb{E}\left[\varphi^{a,b}_{K,N}(X(T))\middle|\,\mathcal{F}_t\right] \approx \mathbb{E}\left[\varphi_K(X(T))|\,\mathcal{F}_t\right] = \Pi_K(t). \tag{4.1}$$

These give an approximation for the price of a European-style call option with underlying process X, which we state in the following theorem. Notice that, at the current stage, X is a generic $\mathcal{F}_t$-adapted stochastic process, and not the discrete average defined in Eq. (1.2). Thus Theorem 4.1 gives an approximation for the price of a European-style call option with underlying process X.

Theorem 4.1. *The price approximation by Hermite polynomials for a European-style call option is*

$$\Pi^{a,b}_{K,N}(t) = \sum_{k=0}^{N} \hat{\boldsymbol{\beta}}^{a,b}_{N,k+1}\frac{1}{b^k}\sum_{i=0}^{k}\binom{k}{i}(-a)^{k-i}\,\mathbb{E}\left[\,X(T)^i\middle|\,\mathcal{F}_t\right], \tag{4.2}$$

where $\hat{\boldsymbol{\beta}}^{a,b}_N := \boldsymbol{\beta}^{a,b\,\top}_N M_N$ *with components* $\hat{\boldsymbol{\beta}}^{a,b}_{N,k+1} = \boldsymbol{\beta}^{a,b\,\top}_N (M_N)_{:,(k+1)}$, $k = 0, \ldots, N$.

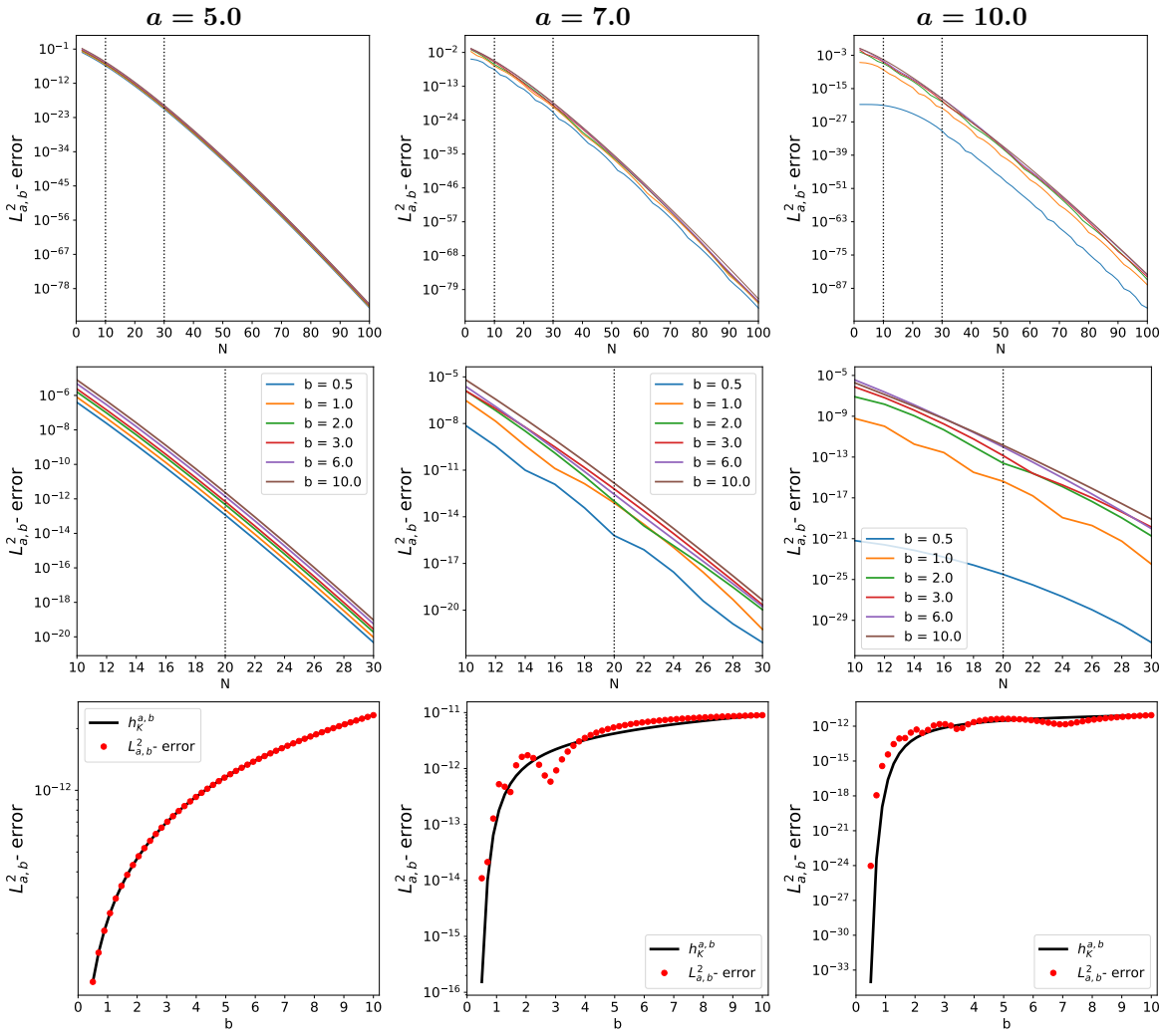

Figure 2: $L^2_{a,b}$-norm of the approximation error of $\varphi^{a,b}_{K,N}$ with $K = 5.0$. In the first and second row, the error is a function of the truncation number N. In the third row, the error is a function of b.

In Theorem 4.1 the vector $\boldsymbol{\beta}_N^{a,b}$ and the matrix M_N depend on the choice of the orthogonal basis, the GHPs in our case, while the moments $\mathbb{E}\left[X(T)^i \middle| \mathcal{F}_t\right]$ only depend on the distribution of the random variable $X(T)$. Then, once the orthogonal basis is chosen, the approximation of the expected payoff $\Pi_{K,N}^{a,b}(t)$ is fully determined by the conditional moments of $X(T)$. In the next section we extend the result in Theorem 4.1 to Asian options (hence for X as in Eq. (1.2)) by the multinomial theorem.

4.2 Asian options

We now consider Asian-style options as introduced at the beginning of Sec. 1. For $m \geq 0$, X is the discrete average of an $\mathcal{F}_t$-adapted stochastic process Y over the period $(t, T]$, namely

$$X(T) = \frac{1}{m+1} \sum_{j=0}^{m} Y(s_j) \qquad \text{for} \qquad t < s_0 < s_1 < \cdots < s_m = T. \tag{4.3}$$

From Theorem 4.1, we need to compute the conditional moments $\mathbb{E}\left[X(T)^i \middle| \mathcal{F}_t\right]$, for $i = 1, \ldots, N$, where N is the truncation number for the Hermite series (3.14). These can be rewritten in terms of correlator-type expectations (as defined in Eq. (1.4)) by means of the multinomial theorem.

Proposition 4.2. *For every $1 \leq i \leq N$, the i-th conditional moment of the average process X in Eq. (4.3) can be rewritten as a linear combination of correlator terms for the process Y, namely*

$$\mathbb{E}\left[X(T)^i \middle| \mathcal{F}_t\right] = \frac{1}{(m+1)^i} \sum_{|\boldsymbol{k}|=i} \frac{i!}{k_0! k_1! \cdots k_m!} \mathbb{E}\left[Y(s_0)^{k_0} Y(s_1)^{k_1} \cdots Y(s_m)^{k_m} \middle| \mathcal{F}_t\right], \tag{4.4}$$

where the summation is over the multi-indexes $\boldsymbol{k} = (k_0, \ldots, k_m)$ with $|\boldsymbol{k}| = k_0 + k_1 + \cdots + k_m$.

By means of Proposition 4.2, we obtain the pricing formula for Asian options.

Theorem 4.3. *For every $N \geq 0$, the price of a discretely sampled arithmetic Asian option can be approximated with generalized Hermite polynomials by*

$$\Pi_{K,N}^{a,b}(t) = \sum_{k=0}^{N} \sum_{i=0}^{k} \sum_{|\boldsymbol{k}|=i} \binom{k}{i} \frac{\hat{\boldsymbol{\beta}}_{N,k+1}^{a,b} (-a)^{k-i}}{(m+1)^i b^k} \frac{i!}{k_0! k_1! \cdots k_m!} \mathbb{E}\left[Y(s_0)^{k_0} Y(s_1)^{k_1} \cdots Y(s_m)^{k_m} \middle| \mathcal{F}_t\right]. \tag{4.5}$$

By Theorem 4.1 and 4.3, we have explicit approximation formulas for the price of European and Asian options which require the computation of conditional moments and correlators. As we have seen in Sec. 2, for a jump-diffusion polynomial process both conditional moments and correlators admit closed formulation. In all other cases, one must rely on Monte Carlo simulations, loosing the advantages of an explicit price functional representation.

Remark 1. In the continuous case, the payoff of an arithmetic Asian option is a function of the integrated underlying process, namely of $X(T) := \frac{1}{T-t} \int_t^T Y(s) ds$. If Y is a polynomial jump-diffusion process, then the pair (Y, X) is a bivariate polynomial jump-diffusion process. Hence one can reduce the Asian option pricing problem for the process Y to a European option pricing problem for the process $Z := (Y, X)$. On the one hand, this only requires the conditional moments of $X(T)$ ($Z(T)$), and not the joint moments, namely the correlators, of Y at the discrete times $t < s_0 < s_1 < \cdots < s_m = T$. On the other hand, one needs to deal with a two-dimensional process. We shall analyze this alternative approach in Sec. 6.

4.3 Error analysis and scale criterion

Let $\psi_{X(T)}$ be the density function of $X(T)$. We estimate the error in approximating $\Pi_K(t)$ with GHPs.

Theorem 4.4. *If the density function $\psi_{X(T)}$ satisfies the condition*

$$\int_{\mathbb{R}} \psi_{X(T)}^2(x) \omega_{a,b}^{-1}(x) dx < \infty, \tag{4.6}$$

then the absolute error in approximating the option price $\Pi_K(t)$ *with* $\Pi^{a,b}_{K,N}(t)$ *is bounded by the* $L^2_{a,b}$*-norm of the error in approximating the payoff function* φ_K *with* $\varphi^{a,b}_{K,N}$*, namely*

$$\left|\Pi_K(t) - \Pi^{a,b}_{K,N}(t)\right| \leq C_{a,b} \left\|\varphi_K - \varphi^{a,b}_{K,N}\right\|_{L^2_{a,b}}, \tag{4.7}$$

where $C_{a,b} := \left(\int_{\mathbb{R}} \psi^2_{X(T)}(x)\omega^{-1}_{a,b}(x)dx\right)^{\frac{1}{2}}$.

Remark 2. To have convergence for the price approximation, by Theorem 4.4 we need $\psi_{X(T)}$ to satisfy condition (4.6) which is quite restrictive. Indeed, it basically asks the tails of $\psi_{X(T)}$ to vanish faster than the tails of a Gaussian density function. In general, this does not hold for Lévy processes, which are characterized by heavy tails. However, condition (4.6) is only sufficient for proving the error bound in Theorem 4.4, and not really necessary for convergence. Indeed, we shall see from the numerical examples that the convergence is reached also when considering a jump process, such as a jump-diffusion process with normal inverse Gaussian Lévy measure as in Sec. 5.3 and 5.4.

Remark 3. Since condition (4.6) is obtained from a weight function being the density of a Gaussian random variable, one can aim at theoretical convergence criteria for other families of distributions by considering a different weight function. For example, the Laguerre polynomials are orthogonal polynomials with respect to the weight function $w(x) = e^{-x}$, which would allow to prove convergence for the Gamma distribution. Similarly, [25] obtained convergence for a log-normal distribution by considering a log-normal density for the weight function. Other examples are in [20, Appendix B].

If $X(T)$ follows a Gaussian distribution with mean μ and variance σ^2, by direct computation we get

$$C^2_{a,b} = \int_{\mathbb{R}} \psi^2_{X(T)}(x)\omega^{-1}_{a,b}(x)dx = \frac{b}{\sqrt{2\pi\sigma^2}} \frac{e^{\frac{(a-\mu)^2}{(2b^2-\sigma^2)}}}{\sqrt{2b^2-\sigma^2}}, \tag{4.8}$$

which leads to a more explicit formulation for condition (4.6).

Proposition 4.5. *If the random variable* $X(T)$ *follows a Gaussian distribution with mean* μ *and variance* σ^2*, then condition* (4.6) *is equivalent to*

$$b > \frac{\sigma}{\sqrt{2}} =: \underline{b}_\sigma \tag{4.9}$$

where b *is the scale for the GHPs. For* $b = \underline{b}_\sigma$ *we expect instabilities in the approximation.*

Remark 4. Condition (4.6) extends [25, Proposition 3.1], since it basically coincides with asking that the likelihood ratio function η defined by $\psi_{X(T)}(x) = \eta(x)\omega_{a,b}(x)$ is such that $\eta \in L^2_{a,b}$ for any generic density function $\psi_{X(T)}$ (and not only for a log-normal density function like in [25]). Indeed, if considering $\psi_{X(T)}$ and $\omega_{a,b}$ to be the density functions of two log-normal distributions with mean μ and variance σ^2, respectively, with mean a and variance b^2, then $C_{a,b}$ takes a form similar to Eq. (4.8). In particular, the squared root $\sqrt{2b^2-\sigma^2}$ still appears. Then, with abuse of notation, by setting $\sigma^2 = \sigma^2 T$, we recover [25, Proposition 3.1] which now coincides with condition (4.9)[1].

Corollary 4.6. *In the same setting of Proposition 4.5, if* $a = \mu$ *then* $C_{a,b}$ *is a monotone decreasing function of the scale* b *with limit* $\frac{1}{\sqrt[4]{4\pi\sigma^2}}$.

Remark 5. Theorem 4.4 shows that the absolute error in approximating the option price $\Pi_K(t)$ with generalized Hermite polynomials is bounded by the product of $C_{a,b}$ with the $L^2_{a,b}$-norm of the error in approximating the payoff function φ_K. In particular, due to Proposition 3.3, this last term is an increasing function of b (if a is in a neighborhood of K), while, according to Corollary 4.6, $C_{a,b}$ is a decreasing function of b, at least in the Gaussian case.

[1] To be more precise, the setting in [25] defines a likelihood ratio function ℓ in terms of a the weight ω and a density g, where the latter one is the density function of the average price process defined in a continuous manner starting from a log-normally distributed underlying spot price. Then g does not define a log-normal distribution. However, its tails are dominated by the tails of a log-normal density function, as proved in [25].

4.4 Greeks for path-dependent options

The pricing formulas of Theorem 4.1 and 4.3, together with the moment and correlator formulas of Theorem 2.1 and 2.3, allow for sensitivity analysis and risk management. Indeed, thanks to the compact and closed formulation, it is possible to differentiate the price functional with respect to various parameters and obtain the option *Greeks*. In [4, Section 6] the authors derive the expressions for the Delta and the Theta for correlators, namely for the partial derivatives

$$\begin{aligned} \Delta^{k_0,\dots,k_m}_{(s_0,\dots,s_m;t)} &:= \frac{\partial C^{k_0,\dots,k_m}(s_0,\dots,s_m;t)}{\partial Y(t)} \qquad \text{and} \\ \Theta j^{k_0,\dots,k_m}_{(s_0,\dots,s_m;t)} &:= \frac{\partial C^{k_0,\dots,k_m}(s_0,\dots,s_m;t)}{\partial s_j}, \end{aligned} \tag{4.10}$$

for $0 \le j \le m$, with $C^{k_0,\dots,k_m}(s_0,\dots,s_m;t) := \mathbb{E}\left[Y(s_0)^{k_0}Y(s_1)^{k_1}\cdots Y(s_m)^{k_m}\middle|\mathcal{F}_t\right]$. Starting from their results, we compute Delta and Theta for discretely sampled arithmetic Asian options.

Proposition 4.7. *For every $N \ge 0$, the Delta of a discretely sampled arithmetic Asian option can be approximated with generalized Hermite polynomials by*

$$\frac{\partial \Pi^{a,b}_{K,N}(t)}{\partial Y(t)} = \sum_{k=0}^{N}\sum_{i=0}^{k}\sum_{|\boldsymbol{k}|=i}\binom{k}{i}\frac{\hat{\boldsymbol{\beta}}^{a,b}_{N,k+1}(-a)^{k-i}}{(m+1)^i b^k}\,\frac{i!}{k_0!k_1!\cdots k_m!}\,\Delta^{k_0,\dots,k_m}_{(s_0,\dots,s_m;t)} \tag{4.11}$$

with $\Delta^{k_0,\dots,k_m}_{(s_0,\dots,s_m;t)}$ given in [4, Proposition 6.1].

Proposition 4.8. *For every $N \ge 0$, the Theta of a discretely sampled arithmetic Asian option can be approximated with generalized Hermite polynomials by*

$$\frac{\partial \Pi^{a,b}_{K,N}(t)}{\partial s_j} = \sum_{k=0}^{N}\sum_{i=0}^{k}\sum_{|\boldsymbol{k}|=i}\binom{k}{i}\frac{\hat{\boldsymbol{\beta}}^{a,b}_{N,k+1}(-a)^{k-i}}{(m+1)^i b^k}\,\frac{i!}{k_0!k_1!\cdots k_m!}\,\Theta j^{k_0,\dots,k_m}_{(s_0,\dots,s_m;t)} \quad \text{for } 0 \le j \le m, \tag{4.12}$$

with $\Theta j^{k_0,\dots,k_m}_{(s_0,\dots,s_m;t)}$ given in [4, Proposition 6.2].

One similarly computes Delta and Theta for European options as dealt in Sec. 4.1.

5 Numerical Examples

We shall implement numerically the pricing formulas in Theorem 4.1 and 4.3. We first test the pricing formula with moments (Theorem 4.1) for a Brownian motion, a Gaussian Ornstein–Uhlenbeck process and a jump-diffusion process, all of these being polynomial processes as described in Sec. 2. We then test the pricing formula with correlators (Theorem 4.3) for the Gaussian Ornstein–Uhlenbeck process and the jump-diffusion process. The code for reproducing the experiments is available at https://github.com/silvialava/Pricing_options_with_correlators.git.

5.1 Brownian motion

We consider $X = B$, where B is a Brownian motion. Then $X(T)|\,\mathcal{F}_t \sim \mathcal{N}(0, T-t)$ and the price functional $\Pi_K(t)$ is given in closed form by

$$\Pi_K(t) = \sigma_X(T;t)\,\phi\left(\frac{K-\mu_X(T;t)}{\sigma_X(T;t)}\right) - (K-\mu_X(T;t))\left(1-\Phi\left(\frac{K-\mu_X(T;t)}{\sigma_X(T;t)}\right)\right), \tag{5.1}$$

with $\sigma_X(T;t) = \sqrt{T-t}$ and $\mu_X(T;t) = 0$, so that we can benchmark the price approximation. To do that, we introduce the quantity

$$\gamma_{a,b}^N := -\log\left(\frac{\left|\Pi_K(t) - \Pi_{K,N}^{a,b}(t)\right|}{\Pi_K(t)}\right) \tag{5.2}$$

which measures the accuracy of $\Pi_{K,N}^{a,b}(t)$, namely, the accuracy is of order $10^{-\gamma_{a,b}^N}$. We also compare $\gamma_{a,b}^N$ with the accuracy of a Monte-Carlo-simulation (MC) approach calculated in the same manner. We report $\gamma_{a,b}^N$ in Figure 3 and 4 as a function of N for different values of K and b. Here we draw with a red horizontal dashed line the MC accuracy, and with a red vertical bar the value of N for which the Hermite series reaches the same accuracy of the MC method. This latter one is obtained by averaging over 10^2 outcomes, each of them resulting from $2 \cdot 10^4$ simulations.

In Figure 3 we have $T = 1/2$ and $t = 0$, hence $\sigma_X(T;t) \approx 0.707$. We observe that for $b = 0.5$ the Hermite series is barely able to reach the accuracy of the Monte Carlo simulations, and it is not clear if we can actually consider it converging. This coincides indeed with the case $b = \underline{b}_\sigma$ in Proposition 4.5 which is expected to show instabilities. Things get better for $b = 0.6$ and $b = 1.0$: here we observe convergence of the Hermite approximation, reaching a level of accuracy of order 10^{-10}. This convergence is slower for $b = 1.0$ than for $b = 0.6$. For $b = 2.0$ and $b = 3.0$ the convergence is even slower and to reach the same accuracy of Monte Carlo we need more than 50 terms in the first case and more than 100 terms in the second case, meaning that a big scale parameter slows down the convergence.

Very similar comments hold for Figure 4, where $T = 2$, $t = 0$ and $\sigma_X(T;t) \approx 1.41$. Here the plots show a similar behavior to the ones in Figure 3, even if the values of the scale b considered are different. More precisely, for Figure 4 we consider values for b which are two times (i.e. the double) the values used in Figure 3. Since from Eq. (4.8) $C_{a,b}$ is proportional to $\frac{b}{\sigma}$, it seems reasonable to think that this phenomenon is related to the fact that the standard deviation of the process is exactly two times the standard deviation of the process in Figure 3. In other words, because of the ratio $\frac{b}{\sigma}$ that somehow controls the approximation error, in order to get the same accuracy we need to keep this ratio constant. Hence, if σ doubles, also b must double. Finally, we notice that the singularity in the sense of Proposition 4.5 is here expected for $b = 1.0$, as indeed Figure 4 shows.

Another phenomenon observable in both Figure 3 and 4 is that, after reaching the best accuracy, the bars in the plots decrease. Moreover, some parts of the plots are empty, as, for example, in the plot corresponding to $K = 0.0$ and $b = 0.6$ of Figure 3 after $N = 60$. This is because, after that, the values of $\gamma_{a,b}^N$ become negative thus they don't appear in the plot. A negative $\gamma_{a,b}^N$ means in particular that the value of $\Pi_{K,N}^{a,b}(t)$ is completely far away from the true price value. We believe that this is due to numerical instabilities. In computing the approximation $\Pi_{K,N}^{a,b}(t)$ in Theorem 4.1 we need indeed the conditional moments of $X(T)$. It is clear that for high values of the truncation number N, we need to calculate high-order moments of X, which create numerical instabilities due to rounding errors.

5.2 Gaussian Ornstein–Uhlenbeck process

We consider $X = Y$, where Y is the Gaussian Ornstein–Uhlenbeck (OU) process defined by

$$dY(t) = (b_0 + b_1 Y(t))\, dt + \sqrt{\sigma_0}\, dB(t), \tag{5.3}$$

for $b_0, b_1, \sigma_0 \in \mathbb{R}$ and $\sigma_0 > 0$. Then $X(T)|\,\mathcal{F}_t \sim \mathcal{N}(\mu_X(T;t), \sigma_X^2(T;t))$ with

$$\mu_X(T;t) = X(t)\, e^{b_1(T-t)} + \frac{b_0}{b_1}\left(e^{b_1(T-t)} - 1\right) \quad \text{and} \quad \sigma_X^2(T;t) := \frac{\sigma_0}{2b_1}\left(e^{2b_1(T-t)} - 1\right). \tag{5.4}$$

Moreover, since Y is a polynomial process (thus X is a polynomial process), the moments of $X(T)$ are given by Corollary 2.2 and the price functional $\Pi_K(t)$ is again given in closed form by Eq. (5.1).

In Figure 5 and 6 we report the numerical results for $(b_0, b_1, \sigma_0) = (-0.02, 0.01, 0.98)$, $T = 2$ and $t = 0$, which have been chosen so to get $\mu_X(T;t) = X(0) = 2.0$ for Figure 5 and $\mu_X(T;t) = X(0) = 20.0$ for Figure 6. Moreover $\sigma_X(T;t) \approx 1.41$ for both figures as for the process in Figure 4. Indeed, Figure 5 looks very similar to Figure 4 and the behavior of the approximation with respect to the scale b

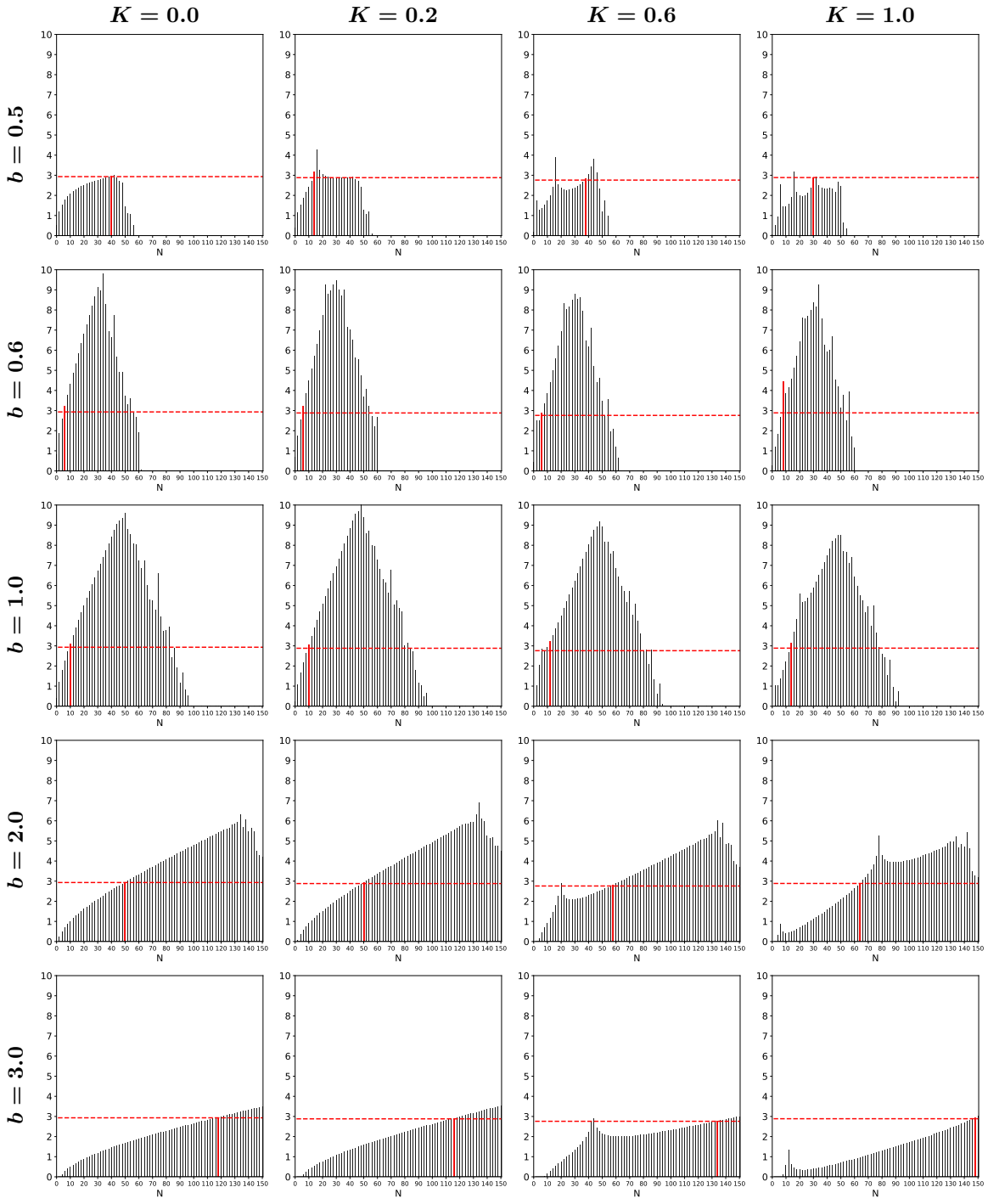

Figure 3: $\gamma_{a,b}^{N}$ as a function of N when the underlying process is a BM with $\sigma_X(T;t) \approx 0.707$. The drift is $a = \mu_X(T;t) = 0$. The dashed red horizontal lines indicate the accuracy of the MC method. The red vertical bars indicate when the Hermite series reaches the MC accuracy.

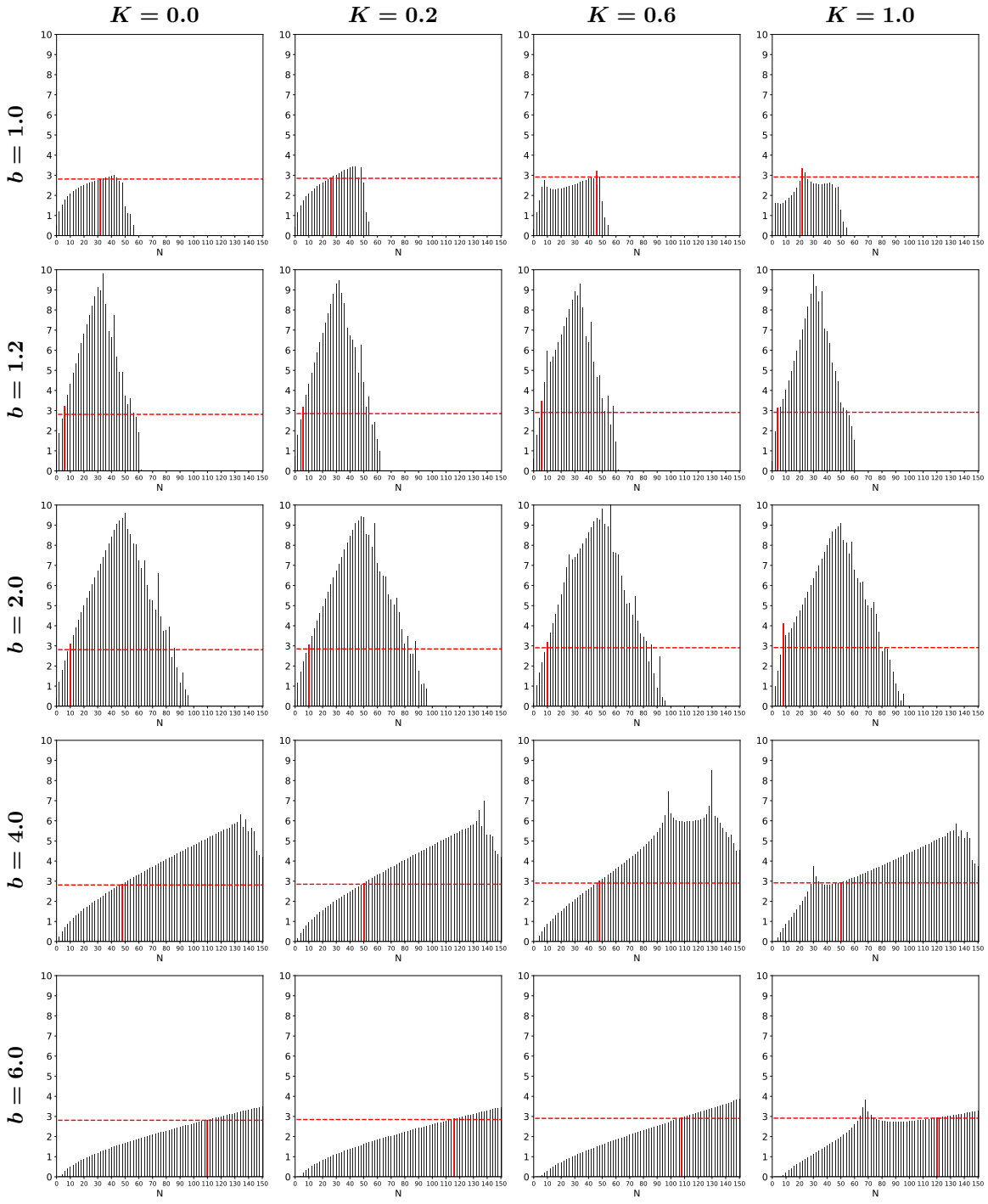

Figure 4: $\gamma_{a,b}^{N}$ as a function of N when the underlying process is a BM with $\sigma_X(T;t) \approx 1.41$. The drift is $a = \mu_X(T;t) = 0$. The dashed red horizontal lines indicate the accuracy of the MC method. The red vertical bars indicate when the Hermite series reaches the MC accuracy.

is similar, as we would expect due to the fact the the volatility is the same. However, the maximum accuracy reached is lower than the one for the Brownian motion (10^{-8} at best). Moreover, the numerical instabilities in the sense discussed above appear earlier, namely, for a smaller N. We believe that both these phenomena are related to the fact that $X(0) = 2.0 > 0$: the high-order moments of $X(T)$ that we need to calculate for approximating the price reach here larger values than for the Brownian motion, which has mean zero. Hence the instabilities occur at an earlier stage. This phenomenon is even more emphasized in Figure 6 where $X(0) = 20.0 >> 0$. Here indeed the numerical instabilities start around $N = 20$, which in some cases is not even enough for reaching a reasonable accuracy.

5.3 Polynomial jump-diffusion process

We consider $X = Y$, where Y is the polynomial jump-diffusion process following the dynamics

$$dY(t) = (b_0 + b_1 Y(t))\, dt + \sqrt{\sigma_0}\, dB(t) + \int_{\mathbb{R}} z\tilde{N}(dt, dz), \tag{5.5}$$

where $\tilde{N}(dt, dz)$ is a compensated Poisson random measure with compensator $\nu(dz)dt$. In this case, the jump measure $\ell(x, dz)$ is given by $\int_{\mathbb{R}} f(z)\ell(x, dz) = \int_{\mathbb{R}} f(\delta(x, z))\nu(dz)$, see [4, Example 2.1] for details. Moreover, we consider ν to be the Lévy measure of a normal inverse Gaussian (NIG) process with location parameter $\mu \in \mathbb{R}$, scale $\delta > 0$, asymmetry parameter β and steepness parameter α (see [3]). Since Y is a polynomial process, the moments of $X(T)$ are given in closed form by Corollary 2.2. However, there is no explicit price functional in this case.

In Figure 7 we report the results for a jump-diffusion process with $(b_0, b_1, \sigma_0) = (-0.02, 0.01, 0.49)$, $(\alpha, \beta, \mu, \delta) = (1.0, 0, 0, 0.05)$, $T = 2$, $t = 0$ and initial condition $X(0) = 2.0$. The mean and standard deviation, calculated with the moment formula in Corollary 2.2, are $\mu_X(T; t) = 2.00$ and $\sigma_X(T; t) = 1.00$ respectively. Since no closed price formula is available, we cannot quantify the level of accuracy. The plots of Figure 7 show then the approximated price with generalized Hermite polynomials compared with the approximated price via Monte Carlo. For each experiment, we print two plots: the first one showing the results from the whole experiment and the second one showing a selected subset of it.

In particular, we use values for b which have the same proportion with respect to $\underline{b}_\sigma = \frac{\sigma}{\sqrt{2}}$ (for $\sigma = \sigma_X(T; t)$ calculated with the moment formula (4.2)) of Figure 3. Even if $\underline{b}_\sigma$ has been introduced in Proposition 4.5 for a Gaussian random variable, the results in Figure 7 are in line with the previous ones. Specifically, we see that for higher values of the scale the convergence is slower but more stable at the same time. For example, for $b = 2.0\underline{b}_\sigma$, the convergence is reached around $N = 17$, but after $N = 47$, due to numerical instabilities, the series starts to diverge. On the other hand, for $b = 6.0\underline{b}_\sigma$, we see that $N = 60$ terms are not enough to reach convergence. Last, we notice that for $b = \underline{b}_\sigma$ there is no convergence, as expected in line with Proposition 4.5. This is noticeable both for $K = 1.0$ and $K = 2.0$ in the zoomed plots: we see that the option price oscillates around the MC price, without reaching convergence. After a certain number of iterations (around $N = 30$), because of numerical instabilities, the series starts to diverge.

5.4 Asian options

We shall now test the pricing formula with correlators in Theorem 4.3 for discretely sampled arithmetic Asian options. In particular, making use of the insight learnt from the previous experiments, we shall deal only with the Gaussian OU process and the polynomial-jump diffusion process introduced above. In the first case, we benchmark the approximation with a closed pricing formula, while for the second process no closed formula is available.

We first consider the OU process Y introduced in Eq. (5.3). Then the average process X is

$$\begin{aligned} X(T) &= \frac{1}{m+1}\sum_{j=0}^{m} Y(s_j) \quad \text{with} \\ Y(s_j) &= Y(t)e^{b_1(s_j - t)} + \frac{b_0}{b_1}\left(e^{b_1(s_j-t)} - 1\right) + \sqrt{\sigma_0}\int_t^{s_j} e^{b_1(s_j - v)} dB(v). \end{aligned} \tag{5.6}$$

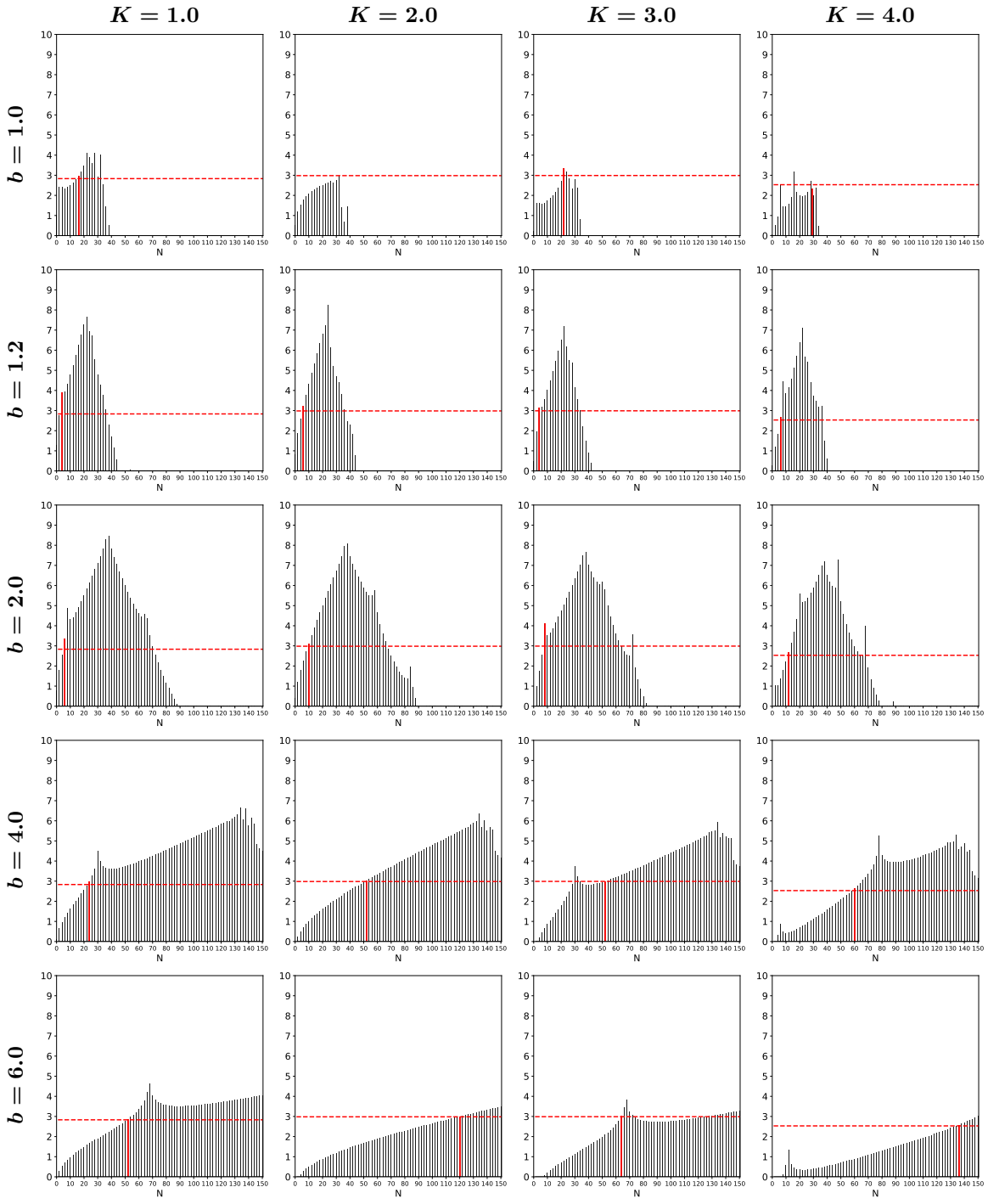

Figure 5: $\gamma_{a,b}^{N}$ as a function of N when the underlying process is an OU with $\sigma_X(T;t) \approx 1.41$. The drift is $a = \mu_X(T;t) = 2.0$. The dashed red horizontal lines indicate the accuracy of the MC method. The red vertical bars indicate when the Hermite series reaches the MC accuracy.

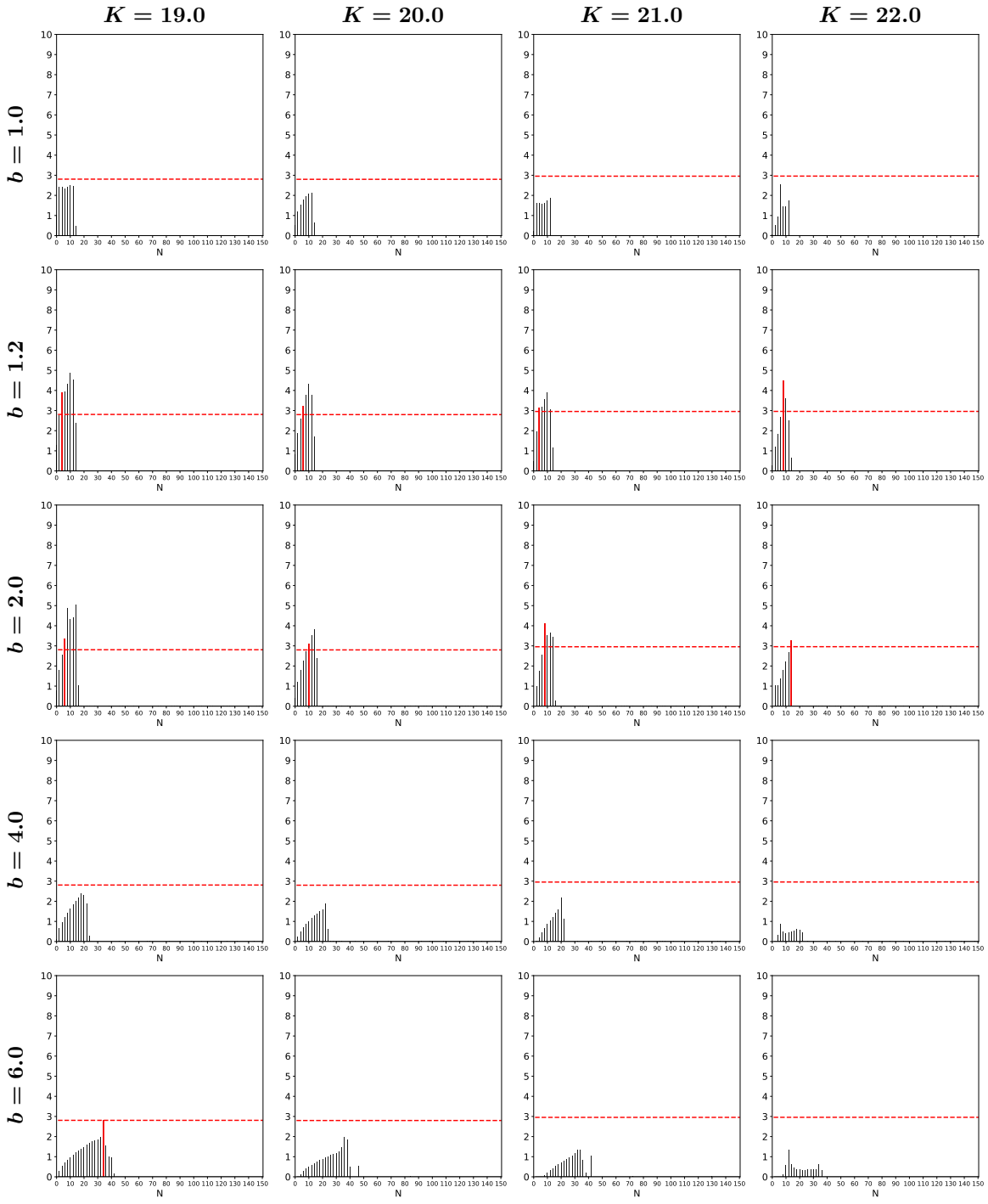

Figure 6: $\gamma_{a,b}^{N}$ as a function of N when the underlying process is an OU with $\sigma_X(T;t) \approx 1.41$. The drift is $a = \mu_X(T;t) = 20.0$. The dashed red horizontal lines indicate the accuracy of the MC method. The red vertical bars indicate when the Hermite series reaches the MC accuracy.

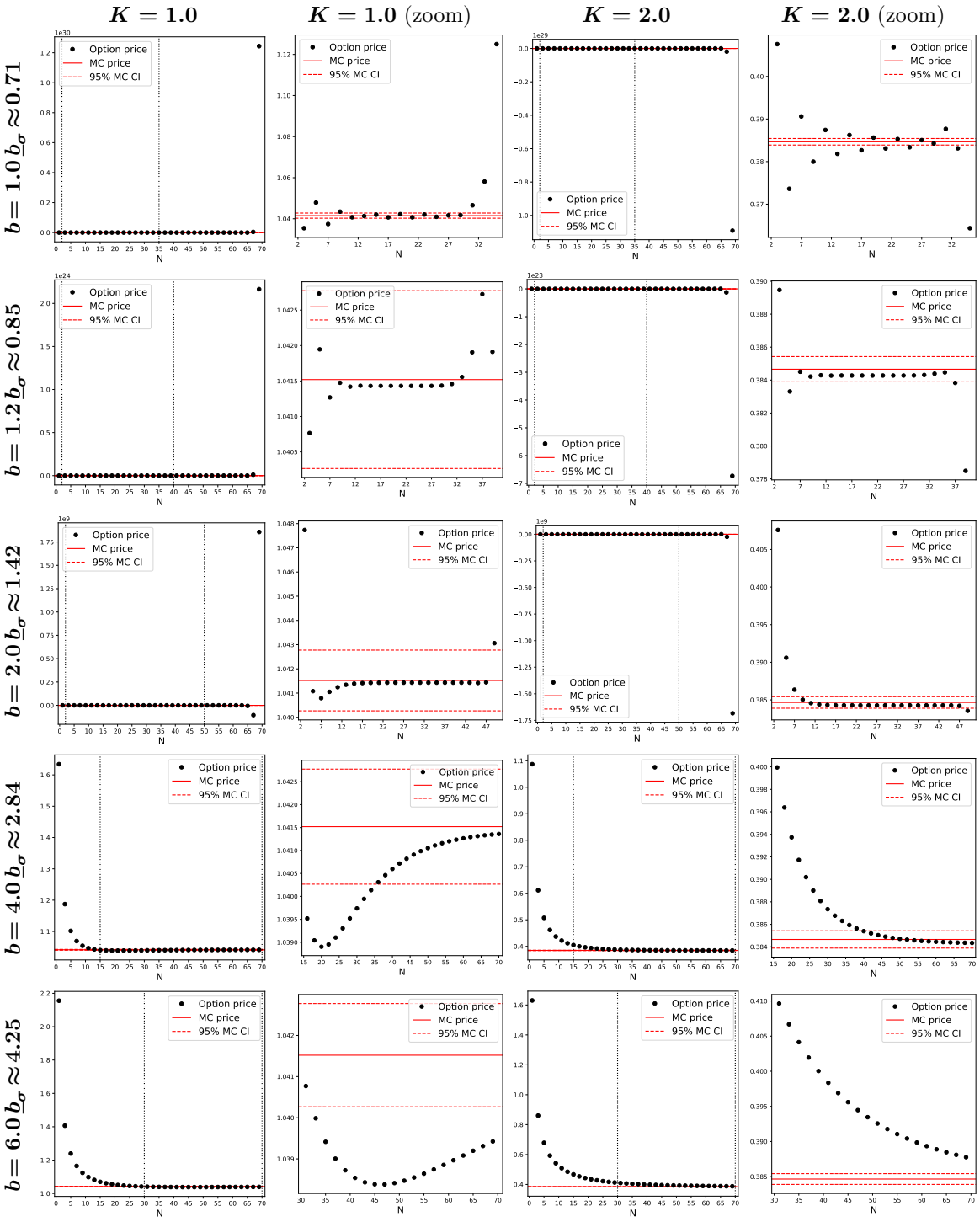

Figure 7: $\Pi^{a,b}_{K,N}(t)$ as a function of N when the underlying process is a jump-diffusion with NIG measure and $\sigma_X(T;t) \approx 1.41$. The drift is $a = \mu_X(T;t) = 2.0$. The black dots represent the option price, the solid red line the MC price, the two red dashed lines are the 95% confidence interval for MC. The plots in the second and fourth columns are a zoomed subplot of the plots in the first and third columns.

In particular, the random variables $\{Y(s_j)\}_{j=0}^m$ are not independent. We can however rewrite their sum as the sum of some other independent random variables $\{Z_j\}_{j=0}^m$.

Proposition 5.1. *For $s_{-1} := t$, the random variable $X(T)$ equals in distribution the weighted sum of $m+1$ independent random variables $\{Z_j\}_{j=0}^m$, namely $X(T) \overset{d}{=} \frac{1}{m+1}\sum_{j=0}^m Z_j$, where Z_j is defined by*

$$Z_j := Y(t)e^{b_1(s_j-t)} + \frac{b_0}{b_1}\left(e^{b_1(s_j-t)}-1\right) + \sqrt{\sigma_0}\int_{s_{j-1}}^{s_j}\left(\sum_{k=j}^m e^{b_1(s_k-v)}\right)dB(v) \qquad \text{for } j=0,\dots,m. \tag{5.7}$$

As a direct consequence of Proposition 5.1, we find that $X(T)|\,\mathcal{F}_t \sim \mathcal{N}(\mu_X(T;t), \sigma_X^2(T;t))$ with

$$\begin{aligned}
\mu_X(T;t) &= \frac{1}{m+1}\sum_{j=0}^m\left(Y(t)\,e^{b_1(s_j-t)} + \frac{b_0}{b_1}\left(e^{b_1(s_j-t)}-1\right)\right) \quad \text{and} \\
\sigma_X^2(T;t) &= \frac{\sigma_0}{(m+1)^2}\sum_{j=0}^m\sum_{k_1=j}^m\sum_{k_2=j}^m \frac{e^{b_1(s_{k_1}+s_{k_2}-2s_{j-1})} - e^{b_1(s_{k_1}+s_{k_2}-2s_j)}}{2b_1},
\end{aligned} \tag{5.8}$$

which, together with Eq. (5.1), gives a benchmark for the experiments.

In Figure 8 and 9 we report the results for, respectively, the Gaussian OU and the polynomials jump-diffusion process with $a = \mu_X(T;t)$ and $b = 2.0\underline{b}_\sigma$, since in the previous experiments this was the value of the scale performing at best. All the experiments are in line with the previous ones: the accuracy of the approximation increases with N increasing, until a certain value after which it starts decreasing. We see that the Hermite approximation performs well also for a path-dependent option, whose evaluation requires the correlator formula instead of the moment formula for polynomial processes. Indeed, there is no significant difference between $m=0$, $m=1$ and $m=2$. From the zoomed plots in Figure 9 we also notice some discrepancy between the price values estimated with the Hermite series and the ones estimated with MC. Based on the previous experiments, we think that the Hermite series is more accurate than the MC approach. The difference is indeed in the fourth decimal, while in the previous experiments the accuracy of MC was not much better than 10^{-3}.

For practical purposes, one needs a way to understand when to truncate the series, i.e., how to choose the value for N. We then propose the following stopping criterion: for each N one calculates

$$\tilde{\gamma}_{a,b}^N := -\log\left(\frac{\left|\Pi_{K,N-1}^{a,b}(t) - \Pi_{K,N}^{a,b}(t)\right|}{\Pi_{K,N-1}^{a,b}(t)}\right). \tag{5.9}$$

For a selected level of accuracy $\bar{\gamma} > 0$, if $\tilde{\gamma}_{a,b}^N > \bar{\gamma}$, then the contribution of the N-th term is smaller than $10^{-\bar{\gamma}}$ and one truncates the series. In Figure 9 the prices obtained with $\bar{\gamma} = 4$ are marked with a red star. Unfortunately, due to computational constraints, it is not possible to test the approximation for $m > 2$.

6 A Comparison Case Study in Continuous Time

In Sec. 1 we introduced discretely-sampled arithmetic Asian options, which we have priced analytically in Sec. 4 and numerically in Sec. 5. From a realistic point of view, the underlying price process (and the option price process) is observed at discrete time intervals. However, Asian options are often defined with continuous sampling. This means that the process X is the continuous average

$$X(T) := \frac{1}{T}\int_0^T Y(s)ds \tag{6.1}$$

of the underlying process Y. As outlined in Sec. 1, the integral defining X can be approximated with a sum by considering a discrete sampling with time step small enough, so that to fall again in the discrete-time setting analyzed so far. However, to have a small time step means that the number

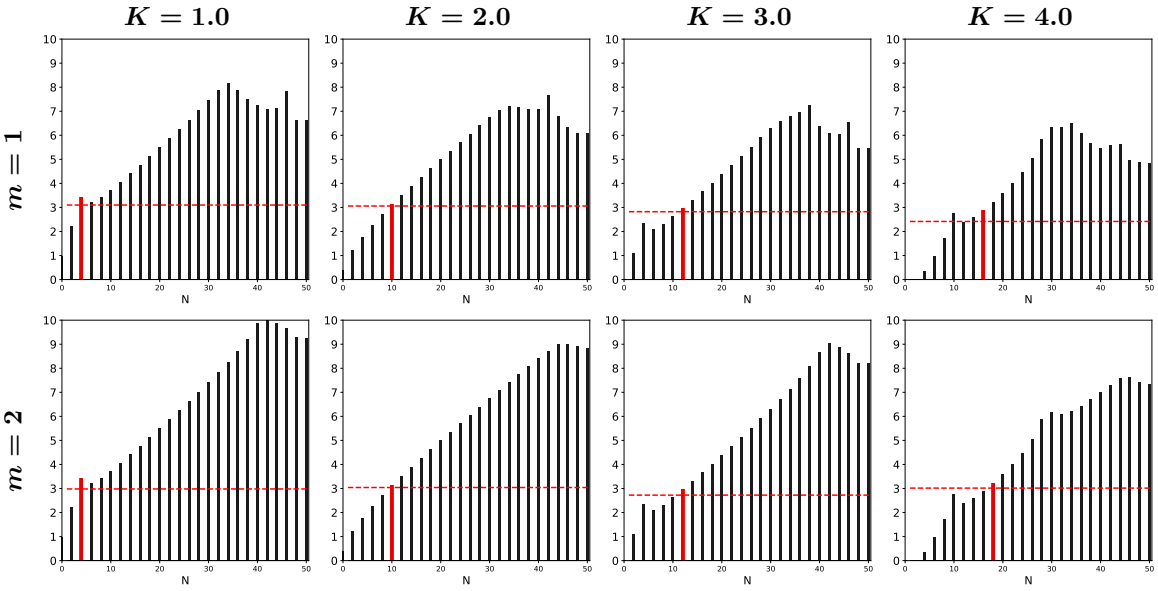

Figure 8: $\gamma_{a,b}^N$ as a function of N when the underlying process is an $(m+1)$-point weighted OU. The drift is $a = \mu_X(T;t) = 2.0$ and the scale is $b = 2.0\underline{b}_\sigma$. The dashed red horizontal lines indicate the accuracy of MC. The red vertical bars indicate when the Hermite series reaches the MC accuracy.

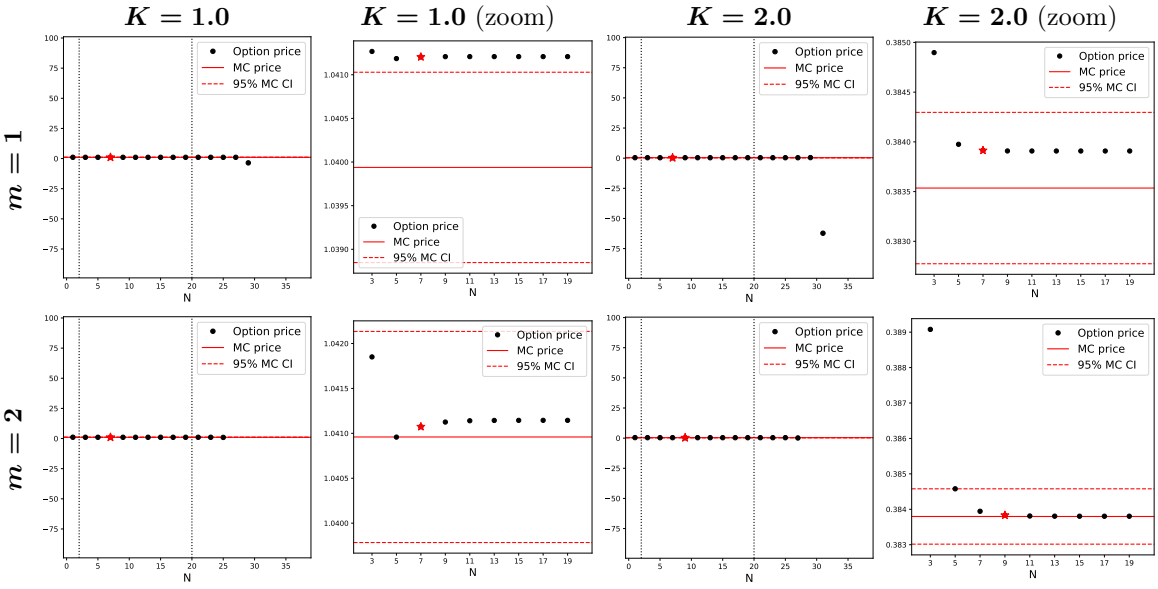

Figure 9: $\gamma_{a,b}^N$ as a function of N when the underlying process is an $(m+1)$-point weighted jump-diffusion with NIG measure. The drift is $a = \mu_X(T;t) = 2.0$ and the scale is $b = 1.2\underline{b}_\sigma$. The black dots represent the option price, the solid red line the MC price, the two red dashed lines are the 95% MC confidence interval. The red stars denote the prices chosen with the stopping criterion (5.9). The plots in the second and fourth columns are a zoomed subplot of the plots in the first and third columns.

Case	r	σ	T	Y_0	GHP	LNS10	LNS15	LNS20	LS	EE	VEC	MC 95% CI
1	.02	.10	1	2.0	**.05604**	.05601	.05600	.05599	.0197	.05599	.05595	[.05598, .05599]
2	.18	.30	1	2.0	**.2188**	.2185	.2184	.2184	.2184	.2184	.2184	[.2183, .2185]
3	.0125	.25	2	2.0	**.1736***	.1723	.1722	.1722	.1723	.1723	.1723	[.1722, .1724]
4	.05	.50	1	1.9	**.1949***	.1930	.1927	.1928	.1932	.1932	.1932	[.1929, .1933]
5	.05	.50	1	2.0	**.2524***	.2466	.2461	.2461	.2464	.2464	.2464	[.2461, .2466]
6	.05	.50	1	2.1	**.3162***	.3068	.3062	.3061	.3062	.3062	.3062	[.3060, .3065]
7	.05	.50	2	2.0	**.3574****	.3501	.3499	.3499	.3501	.3501	.3500	[.3494, .3504]

Table 1: Asian option price approximation for different sets of parameters and different methods. The column **GHP** refers to generalized Hermite polynomials with drift $a = \mu_X(T;t)$ and scale $b = 2\underline{b}_\sigma$ as discussed in this paper; **LNSX** refers to the method presented in [25] with $1 + X$ terms; **LS**, **EE** and **VEC** to the methods presented in [7], [19] and [23, 24], respectively; the last column **MC 95% CI** refers to the 95% confidence interval of the Monte-Carlo simulation.

of points in the discretization grid, namely m, is reasonably high. On the other hand, the numerical experiments in Sec. 5 revealed that, due to computational constraints, we cannot deal with $m > 2$.

An alternative approach to introducing a discretization grid and approximating the integral in the definition of X with a sum, is to look at the pair $Z := (Y, X)$ as a bivariate process. If Y is a polynomial jump-diffusion whose coefficients satisfy (2.2), then Z is a bivariate jump-diffusion process. In this framework, the Asian option pricing problem for the process Y reduces to a European option pricing problem for the bivariate process Z. In particular, since Z is a polynomial jump-diffusion, we can use the price approximation with generalized Hermite polynomials as given in Theorem 4.1, together with the moment formula for (two-dimensional) polynomial processes in order to compute the price approximation for an arithmetic Asian option with continuous sampling.

We shall test this second approach numerically. In particular, we consider the Black and Scholes framework and define the underlying process Y to follow a geometric Brownian motion of the form

$$dY(t) = rY(t)dt + \sigma Y(t)dB(t), \tag{6.2}$$

where $r \in \mathbb{R}$ is the short rate and $\sigma > 0$ is the volatility of the asset. Then X is given by Eq. (6.1). It is easy to see that Y is a polynomial diffusion process with $b_1 = r$, $\sigma_2 = \sigma^2$ and $b_0 = \sigma_0 = \sigma_1 = 0$. Moreover, the pair (Y, X) is a bivariate polynomial diffusion process. Indeed

$$d\begin{pmatrix} Y(t) \\ X(t) \end{pmatrix} = \begin{pmatrix} rY(t) \\ Y(t) \end{pmatrix} dt + \begin{pmatrix} \sigma Y(t) \\ 0 \end{pmatrix} dB(t) \tag{6.3}$$

satisfies the necessary conditions given in [9, Lemma 1] for polynomial diffusion processes in $\mathbb{R}^2$.

The aim of this experiment is twofold. On the one hand, we test the option price approximation with GHPs on one of the most popular models, that, differently from the models considered in Sec. 5, is exponential and log-normally distributed. On the other hand, we compare our results with other existing methods from the literature. In particular, we shall consider the same set of parameters as in [25], so that to be able to use their numerical results.

In Table 1 we report the price results for seven different sets of parameters, namely different r's, different σ's, different T's, different initial condition Y_0's and strike $K = 2.0$. Here **GHP** denotes the price obtained with generalized Hermite polynomials with drift $a = \mu_X(T;t)$ and scale $b = 2\underline{b}_\sigma = 2\frac{\sigma_X(T;t)}{\sqrt{2}}$ (where $\mu_X(T;t)$ and $\sigma_X(T;t)$ were computed with the moment formula for bivariate polynomial process, see Appendix A); **LNSX** denotes the price approximated with the method presented in [25] with $1+X$ terms; **LS**, **EE** and **VEC** are the prices obtained with the method presented in [7], [19] and [23, 24], respectively; finally **MC 95% CI** is the 95% confidence interval of the Monte-Carlo simulation. We point out that all the values (except the GHP column) have been simply copied from [25, Table 1].

Since there is no exact price formula for these experiments, we use the stopping criterion defined in Eq. (5.9). However, for the prices marked with an asterisk * we had to reduce the tolerance to 10^{-2} and the price marked with a double asterisk ** was obtained with a tolerance equal to 10^{-1}. This means that in these cases the approximating series could not converge, unless we reduced the tolerance level. The truncation number N for the Hermite series is reported in the last column of Table 2. In particular,

Case	r	σ	T	Y_0	$\mu_X(T;t)$	$\sigma_X(T;t)$	**cv**	**N**
1	.02	.10	1	2.0	2.020	.1170	.05795	11
2	.18	.30	1	2.0	2.191	.3926	.1792	12
3	.0125	.25	2	2.0	2.025	.4213	.2080	7
4	.05	.50	1	1.9	1.948	.5843	.3000	5
5	.05	.50	1	2.0	2.051	.6152	.3000	5
6	.05	.50	1	2.1	2.153	.6459	.3000	5
7	.05	.50	2	2.0	2.103	.9287	.4415	6

Table 2: Characteristics for the seven sets of parameters considered. Here $\boldsymbol{\mu_X(T;t)}$ and $\boldsymbol{\sigma_X(T;t)}$ have been computed with the moment formula for bivariate polynomial process in Eq. (A.5). The column **cv** denotes the coefficient of variation and $\boldsymbol{N}$ is the truncation number for the Hermite series approximations which led to the price results in Table 1.

the GHP prices are less accurate than the corresponding values obtained with the other methods.

The explanation for the phenomenon we are observing is the following. The weight function used to define the generalized Hermite polynomials corresponds to the density function of a Gaussian random variable. However, the underlying process considered here follows a log-normal distribution. It is known that, if the coefficient of variation of the underlying process, namely the ratio between $\sigma_X(T;t)$ and $\mu_X(T;t)$ is indicatively below 0.18, then the log-normal distribution $\text{Log}\mathcal{N}(\mu_X(T;t), \sigma_X^2(T;t))$ can be approximated with a normal distribution $\mathcal{N}(\mu_X(T;t), \sigma_X^2(T;t))$ with the same mean and variance. In Table 2 we report for each of the seven cases the corresponding value for the mean $\mu_X(T;t)$, variance $\sigma_X(T;t)$ and their ratio, namely the coefficient of variation (**cv**). We see that only in the first two cases the coefficient-of-variance criterion described above is satisfied. This explains the results of Table 1.

To confirm this statement, we consider one last numerical example which is reported in Figure 10. Here we consider a European option (similarly to Sec. 4.1) with $X = Y$ and Y being the geometric Brownian motion defined in Eq. (6.2). Here the drift is set to $a = \mu_X(T;t)$ and the scale to $b = 2\underline{b}_\sigma$, where mean $\mu_X(T;t)$ and standard deviation $\sigma_X(T;t)$ vary for each row. More precisely, in each row we consider a different value for the coefficient of variation (**cv**). We notice that, as this latter one increases, the accuracy of the approximation gets worse. In particular, we can talk about convergence of the series only for the experiments in the first two rows.

7 Conclusions

We derive explicit pricing formulas for discrete-average arithmetic Asian options in the context of one-dimensional polynomial jump-diffusion processes. This can be extended to continuous-average Asian options either by approximating the integral with a discrete sum and appropriate small time step, or, alternatively, by dealing with a two-dimensional process with components the underlying asset and its integral. The proposed approach is based on approximating the payoff function with generalized Hermite polynomials. In particular, the generalized Hermite polynomials are defined in relation to a weight function which is the density of a Gaussian random variable with mean a (the drift) and standard deviation b (the scale). Hence we get a family of approximations depending on the parameters a and b. By modeling the underlying asset price with a jump-diffusion polynomial process, we obtain a fully explicit expression for the price functional thanks to the well-known moment formula for polynomial processes and to the correlator formula derived in [4]. This allows for sensitivity analysis, since Greeks are within reach, as we show. From the numerical point of view, the most time consuming part is the computation of moments and correlators. However, these do not depend on the strike price of the option, hence one can compute moments and correlators for the underlying asset price process and use these values to evaluate different options with different strike prices.

We summarize the following findings:

1. We provide a lower bound for the scale b which is proportional to the standard deviation, say σ, of the underlying process. Values for b smaller than this threshold do not guarantee convergence. On the other hand, big values of b slow down the convergence rate. Despite the lower bound is proved

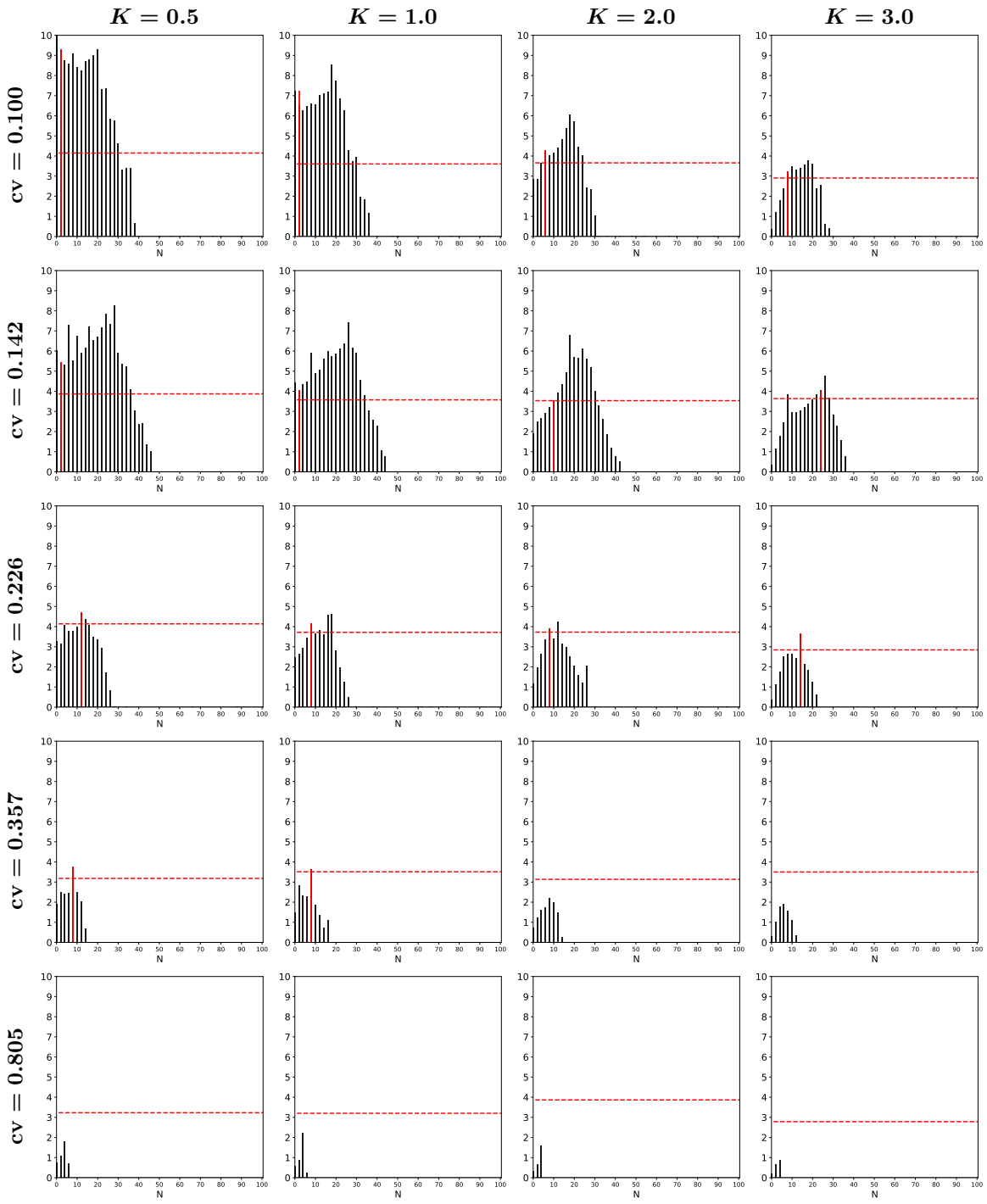

Figure 10: $\gamma_{a,b}^{N}$ as a function of N when the underlying process is a geometric BM with different coefficient of variation (**cv**) values. The drift for the Hermite series is fixed to $a = \mu_X(T;t)$ and the scale is fixed to $b = 2\underline{b}_{\sigma}$. The dashed red horizontal lines indicate the accuracy of the MC method. The red vertical bars indicate when the Hermite series reaches the MC accuracy.

in the case of a Gaussian underlying process, numerical experiments show similar behaviors also for processes composed by a Gaussian term plus a jump term.

2. We find analytically a relation between the behavior of the series and the ratio $\frac{b}{\sigma}$ which is also confirmed by numerical results. For example, when doubling the value of σ, also b must be doubled to get the same behavior for the approximation. Indeed, in the experiments with the Brownian motion and the Gaussian Ornstein–Uhlenbeck process, by choosing the parameters in such a way that the two processes have the same standard deviation, we obtain two series with very similar behaviors. Moreover, the numerical experiments for the jump process are in line with this result.
3. Working with polynomial functions require the evaluation of high-order powers. In our context, this means the evaluation of high-order moments or correlators in the underlying process. The bigger is the initial value of the process, the higher is the value of its moments and correlators. High initial values coupled with high-order powers create numerical instabilities due to rounding errors. This is the main reason which makes the series to diverge after convergence. If the initial value of the process is high enough, then numerical instabilities might start even before reaching convergence.
4. Numerical experiments with Gaussian underlying processes, for which closed price formulas are available for benchmarking, show that the Hermite price approximation can reach higher accuracy than the MC approach, namely 10^{-10} against 10^{-3} for MC. However, for big values of the scale b the convergence might be so slow that to reach such an accuracy level one needs many more terms than what numerically feasible. Despite we do not have closed price formulas for the jump-diffusion case, the plots obtained show that the Hermite series converges to a value very closed to the price value approximated via MC. Based on the experiments in the Gaussian case, we actually believe the Hermite series to be more accurate than the MC approximation.
5. When considering a geometric model, such as a geometric Brownian motion, the underlying distribution is a log-normal one. We find that our price approximation still works whenever the log-normal distribution is well approximated with a Gaussian one. This, in particular, depends on the coefficient of variation of the underlying process, as we have seen in Sec. 6.

The generalized Hermite polynomials can be replaced with any other family of orthogonal polynomials. In particular, Theorem 4.4 sets a very strong condition on the tails for the distribution of the underlying process, which are required to vanish faster than the tails of a Gaussian density function. Despite this is usually not true for jump processes, such a condition is only sufficient, and not necessary, for convergence. The numerical results on the jump-diffusion process reveal indeed that convergence is reached even with an additional NIG-jump-measure term. As pointed out in Remark 3, one may obtain theoretical convergence results for other families of distributions (other than the Gaussian one) by considering a different weight function. For example, the weight function $w(x) = e^{-x}$ is used to construct the Laguerre polynomials, and it would give convergence for a Gamma distribution. Other examples can be found in [20, Appendix B].

We also point out that working with the class of polynomial jump-diffusion processes is the key for getting fully explicit price formulas. However, the price approximation with Hermite polynomials can be applied to other kind of processes. In these cases, one must rely on Monte Carlo simulations for computing moments and correlators. Similarly, remaining in the class of polynomial jump-diffusions, it is also possible to avoid the use of the correlator formula. Indeed, as pointed out in [4], this can be replaced by the iteratively applying the moment formula combined with the tower rule. However, in both these cases, one looses the advantages of an explicit price functional.

We finally remind that the formulas obtained in this paper are restricted to one-dimensional models. The correlator formula, which is the key for obtaining an explicit price functional for Asian-style options, has indeed been derived in [4] only for the one-dimensional case. Increasing the dimension would make calculations more challenging and would require further considerations, since for any $d \geq 2$ the monomial basis for $\mathbb{R}^d$ is more complex than the univariate one. On the other hand, the moment formula for polynomial jump-diffusion processes holds for processes in any (finite) dimension. Hence one can still price European options via Hermite polynomials approximation as in Sec. 4.1.

A Generator Matrix for Two-Dimensional Polynomial Processes

We derive a recursion formula for the generator matrix associated with two-dimensional polynomial processes. This extends the formula obtained in [4] for one-dimensional processes, although omitting

the jump component for simplicity. We implemented the formula in Python and the code is available at https://github.com/silvialava/Pricing_options_with_correlators.git. We now start by defining a polynomial diffusion operator and by stating the moment formula for processes in $\mathbb{R}^d$, for $d > 0$.

Following [8], we denote by $\mathbb{S}^d$ the set of real symmetric $d \times d$ matrices and with $\mathrm{Pol}_n(\mathbb{R}^d)$ the space of all polynomials on $\mathbb{R}^d$ with degree less than or equal to n. We then consider the maps $\boldsymbol{a} : \mathbb{R}^d \to \mathbb{S}^d$ and $\boldsymbol{b} : \mathbb{R}^d \to \mathbb{R}^d$ satisfying, respectively,

$$\boldsymbol{a}_{ij} \in \mathrm{Pol}_2(\mathbb{R}^d) \qquad \text{and} \qquad \boldsymbol{b}_i \in \mathrm{Pol}_1(\mathbb{R}^d) \qquad \text{for all } i, j, \tag{A.1}$$

and we introduce the partial differential operator of the form

$$\mathcal{G}f(z) = \boldsymbol{b}(z)^\top \nabla f(z) + \frac{1}{2}\mathrm{Tr}\left(\boldsymbol{a}(z)\nabla^2 f(x)\right). \tag{A.2}$$

Under conditions (A.1), the operator $\mathcal{G}$ defines the polynomials diffusion

$$dZ(t) = \boldsymbol{b}(Z(t))dt + \boldsymbol{\sigma}(Z(t))dW(t), \tag{A.3}$$

for a d-dimensional Brownian motion W and a continuous function $\boldsymbol{\sigma} : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ with $\boldsymbol{a} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\top$.

For every n and d, the dimension of $\mathrm{Pol}_n(\mathbb{R}^d)$ is $D := \binom{d+n}{d}$. We then consider a family of polynomial functions $\{h_1(z), \dots, h_D(z)\}$ that forms a basis for $\mathrm{Pol}_n(\mathbb{R}^d)$, and introduce the vector valued function

$$H_{n,d} : \mathbb{R}^d \longrightarrow \mathbb{R}^D, \quad H_{n,d}(z) = (h_1(z), \cdots, h_D(z))^\top. \tag{A.4}$$

By the polynomial property, the operator $\mathcal{G}$ can be expressed in matrix form by introducing $G_{n,d} \in \mathbb{R}^{D\times D}$ such that $\mathcal{G}H_{n,d}(z) = G_{n,d}H_{n,d}(z)$. As a consequence, we obtain the following moment formula for the d-dimensional polynomial diffusion Z defined in Eq. (A.3):

$$\mathbb{E}\left[\boldsymbol{p}(Z(T))|\,\mathcal{F}_t\right] = \vec{p}_D^\top e^{G_{n,d}(T-t)}H_{n,d}(Z(t)), \qquad 0 \le t \le T, \tag{A.5}$$

where $\boldsymbol{p} \in \mathrm{Pol}_n(\mathbb{R}^d)$ and $\vec{p}_D \in \mathbb{R}^D$ is the vector of coefficients of $\boldsymbol{p}$ with respect to $H_{n,d}(z)$.

A.1 The two-dimensional case

We now restrict to $d = 2$ and introduce the notation $z := (y, x) \in \mathbb{R}^2$. Moreover, we consider $\boldsymbol{\sigma}$ as a vector and not as a matrix, namely $\boldsymbol{\sigma} : \mathbb{R}^2 \to \mathbb{R}^2$. As a consequence, in order to match the dimensions, the Brownian motion W must be unidimensional, namely $W = B$ as, e.g., in Eq. (6.3). In order to fulfill conditions (A.1), we set

$$\boldsymbol{b}(z) := \begin{pmatrix} b_y(z) \\ b_x(z) \end{pmatrix} = \begin{pmatrix} b_0 + b_1 y + b_2 x \\ \beta_0 + \beta_1 y + \beta_2 x \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{\sigma}(z) := \begin{pmatrix} \sigma_y(z) \\ \sigma_x(z) \end{pmatrix} = \begin{pmatrix} \sigma_0 + \sigma_1 y + \sigma_2 x \\ s_0 + s_1 y + s_2 x \end{pmatrix} \tag{A.6}$$

so that

$$\boldsymbol{a}(z) = \boldsymbol{\sigma}(z)\boldsymbol{\sigma}(z)^\top = \begin{pmatrix} \sigma_y^2(z) & \sigma_y(z)\sigma_x(z) \\ \sigma_x(z)\sigma_y(z) & \sigma_x^2(z) \end{pmatrix}, \tag{A.7}$$

where

$$\begin{aligned} \sigma_y^2(z) &= \sigma_0^2 + 2\sigma_0\sigma_1 y + 2\sigma_0\sigma_2 x + \sigma_1^2 y^2 + 2\sigma_1\sigma_2 yx + \sigma_2^2 x^2, \\ \sigma_x(z)\sigma_y(z) &= \sigma_0 s_0 + (\sigma_1 s_0 + \sigma_0 s_1)y + (\sigma_0 s_2 + \sigma_2 s_0)x + \sigma_1 s_1 y^2 + (\sigma_1 s_2 + \sigma_2 s_1)yx + \sigma_2 s_2 x^2, \\ \sigma_x^2(z) &= s_0^2 + 2s_0 s_1 y + 2s_0 s_2 x + s_1^2 y^2 + 2s_1 s_2 yx + s_2^2 x^2. \end{aligned} \tag{A.8}$$

The operator $\mathcal{G}$ in Eq. (A.2) can then be rewritten as

$$\mathcal{G}f(z) = b_y(z)f_y(z) + b_x(z)f_x(z) + \frac{1}{2}\sigma_y^2(z)f_{yy}(z) + \sigma_y(z)\sigma_x(z)f_{yx}(z) + \frac{1}{2}\sigma_x^2(z)f_{xx}(z), \tag{A.9}$$

where we have introduced the notation

$$f_y(z) = \frac{\partial f}{\partial y}(z),\ f_x(z) = \frac{\partial f}{\partial x}(z),\ f_{yy}(z) = \frac{\partial^2 f}{\partial y^2}(z),\ f_{yx}(z) = \frac{\partial^2 f}{\partial y \partial x}(z),\ f_{xx}(z) = \frac{\partial^2 f}{\partial x^2}(z). \tag{A.10}$$

Finally, for a fixed n, we consider as basis for $\mathrm{Pol}_n(\mathbb{R}^2)$ the set of bivariate monomials up to degree n, namely the set of dimension $D = \frac{(n+1)(n+2)}{2}$ of monomials of the form $y^m x^w$ for $0 \leq m + w \leq n$. This is given by the set

$$\mathcal{B}_n = \{y^m x^w : 0 \leq m + w \leq j,\ \text{for } 0 \leq j \leq n\}, \tag{A.11}$$

for which we fix a total order like follows[2]:

$$\left\{1, \quad y, x, \quad y^2, yx, x^2, \quad y^3, y^2x, yx^2, x^3, \quad \ldots\ldots, \quad y^n, y^{n-1}x, y^{n-2}x^2, \ldots, yx^{n-1}, x^n\right\}. \tag{A.12}$$

With this total order, the vector valued function $H_{n,2}(z)$ introduced in Eq. (A.4) is uniquely defined.

In order to obtain a recursion formula for the generator matrix associated to $H_{n,2}(z)$, we need first to compute the action of $\mathcal{G}$ on the elements of $\mathcal{B}_n$. Since $\mathcal{G}$ is a second order differential operator, we distinguish the following eight cases which need to be considered:

$$y, \quad x, \quad yx, \quad y^m, \quad y^m x, \quad y^m x^w, \quad yx^w, \quad x^w, \quad \text{for } m, w \geq 2. \tag{A.13}$$

Moreover, since $G_{n,2}$ satisfies $\mathcal{G}H_{n,2}(z) = G_{n,2}H_{n,2}(z)$, we start by constructing the row vectors $A_{N,w}$ satisfying $\mathcal{G}(y^m x^w) = A_{N,w}H_{N,2}(z)$ for $m, w \geq 0$ and $N = m + w \leq n$. The matrix $G_{n,2}$ will then be obtained by stacking these row vectors in an appropriate manner.

Proposition A.1. *For $m, w \geq 2$, the action of $\mathcal{G}$ on the elements of $\mathcal{B}_n$ can be outlined as follows:*

$$\mathcal{G}y = b_0 + b_1 y + b_2 x; \tag{A.14}$$

$$\mathcal{G}x = \beta_0 + \beta_1 y + \beta_2 x; \tag{A.15}$$

$$\begin{aligned}\mathcal{G}yx = {} & \sigma_0 s_0 + (\beta_0 + \sigma_0 s_1 + \sigma_1 s_0)y + (b_0 + \sigma_0 s_2 + \sigma_2 s_0)x + (\beta_1 + \sigma_1 s_1)y^2 \\ & + (b_1 + \beta_2 + \sigma_1 s_2 + \sigma_2 s_1)yx + (b_2 + \sigma_2 s_2)x^2;\end{aligned} \tag{A.16}$$

$$\begin{aligned}\mathcal{G}y^m = {} & \frac{m(m-1)}{2}\sigma_0^2 y^{m-2} + m(b_0 + (m-1)\sigma_0\sigma_1)y^{m-1} + m(m-1)\sigma_0\sigma_2 y^{m-2}x \\ & + m(b_1 + \frac{(m-1)}{2}\sigma_1^2)y^m + m(b_2 + (m-1)\sigma_1\sigma_2)y^{m-1}x + \frac{m(m-1)}{2}\sigma_2^2 y^{m-2}x^2;\end{aligned} \tag{A.17}$$

$$\begin{aligned}\mathcal{G}y^m x = {} & m\sigma_0 s_0 y^{m-1} + \frac{m(m-1)}{2}\sigma_0^2 y^{m-2}x + (\beta_0 + m(\sigma_0 s_1 + \sigma_1 s_0))y^m \\ & + m(b_0 + (m-1)\sigma_0\sigma_1 + \sigma_0 s_2 + \sigma_2 s_0)y^{m-1}x + m(m-1)\sigma_0\sigma_2 y^{m-2}x^2 \\ & + (\beta_1 + m\sigma_1 s_1)y^{m+1} + (mb_1 + \beta_2 + \frac{m(m-1)}{2}\sigma_1^2 + m(\sigma_1 s_2 + \sigma_2 s_1))y^m x \\ & + m(b_2 + (m-1)\sigma_1\sigma_2 + \sigma_2 s_2)y^{m-1}x^2 + \frac{m(m-1)}{2}\sigma_2^2 y^{m-2}x^3;\end{aligned} \tag{A.18}$$

$$\begin{aligned}\mathcal{G}y^m x^w = {} & \frac{w(w-1)}{2}s_0^2 y^m x^{w-2} + mw\sigma_0 s_0 y^{m-1}x^{w-1} + \frac{m(m-1)}{2}\sigma_0^2 y^{m-2}x^w \\ & + w(w-1)s_0 s_1 y^{m+1}x^{w-2} + w(\beta_0 + (w-1)s_0 s_2 + m(\sigma_0 s_1 + \sigma_1 s_0))y^m x^{w-1} \\ & + m(b_0 + (m-1)\sigma_0\sigma_1 + w(\sigma_0 s_2 + \sigma_2 s_0))y^{m-1}x^w + m(m-1)\sigma_0\sigma_2 y^{m-2}x^{w+1} \\ & + \frac{w(w-1)}{2}s_1^2 y^{m+2}x^{w-2} + w(\beta_1 + (w-1)s_1 s_2 + m\sigma_1 s_1)y^{m+1}x^{w-1} \\ & + (mb_1 + w\beta_2 + \frac{m(m-1)}{2}\sigma_1^2 + mw(\sigma_1 s_2 + \sigma_2 s_1) + \frac{w(w-1)}{2}s_2^2)y^m x^w \\ & + m(b_2 + (m-1)\sigma_1\sigma_2 + w\sigma_2 s_2)y^{m-1}x^{w+1} + \frac{m(m-1)}{2}\sigma_2^2 y^{m-2}x^{w+2};\end{aligned} \tag{A.19}$$

$$\mathcal{G}yx^w = \frac{w(w-1)}{2}s_0^2 yx^{w-2} + w\sigma_0 s_0 x^{w-1} + w(w-1)s_0 s_1 y^2 x^{w-2} \tag{A.20}$$

[2]Intuitively, this is the order that one would naturally use when expanding the binomial $(y + x)^j$, for $0 \leq j \leq n$.

$$
\begin{aligned}
&+ w(\beta_0 + (w-1)s_0s_2 + s_0\sigma_1 + \sigma_0 s_1)yx^{w-1} + (b_0 + w(\sigma_0 s_2 + \sigma_2 s_0))x^w \\
&+ \frac{w(w-1)}{2}s_1^2 y^3 x^{w-2} + w(\beta_1 + (w-1)s_1s_2 + \sigma_1 s_1)y^2 x^{w-1} \\
&+ (b_1 + w\beta_2 + \frac{w(w-1)}{2}s_2^2 + w(\sigma_1 s_2 + \sigma_2 s_1))yx^w + (b_2 + w\sigma_2 s_2)x^{w+1}; \\
\mathcal{G}x^w &= \frac{w(w-1)}{2}s_0^2 x^{w-2} + w(w-1)s_0 s_1 yx^{w-2} + w(\beta_0 + (w-1)s_0 s_2)x^{w-1} + \\
&+ \frac{w(w-1)}{2}s_1^2 y^2 x^{w-2} + w(\beta_1 + (n-1)s_1 s_2)yx^{w-1} + w(\beta_2 + \frac{(w-1)}{2}s_2^2)x^w.
\end{aligned} \tag{A.21}
$$

Proof. The result is obtained by applying the generator in Eq. (A.9) to the elements in Eq. (A.13). □

Let us notice that in each equation of Proposition A.1 the elements have been ordered accordingly to Eq. (A.12). The vectors $A_{N,w}$ are then obtained by rewriting these equations in vector form.

Corollary A.2. *We denote by $\vec{0}[r] \in \mathbb{R}^r$ a row vector of 0's. Then for $m, w \geq 2$, we construct the following row vectors:*

$$A_{1,0} = (b_0,\ b_1,\ b_2)\,; \tag{A.22}$$

$$A_{1,1} = (\beta_0,\ \beta_1,\ \beta_2)\,; \tag{A.23}$$

$$A_{2,1} = (\sigma_0 s_0,\ \beta_0 + \sigma_0 s_1 + \sigma_1 s_0,\ b_0 + \sigma_0 s_2 + \sigma_2 s_0,\ \beta_1 + \sigma_1 s_1, \tag{A.24}$$
$$b_1 + \beta_2 + \sigma_1 s_2 + \sigma_2 s_1,\ b_2 + \sigma_2 s_2)\,;$$

$$A_{m,0} = \Big(\vec{0}\left[\frac{(m-1)(m-2)}{2}\right],\ \frac{m(m-1)}{2}\sigma_0^2,\ \vec{0}\,[m-2]\,,\ m(m-1)\sigma_0\sigma_1 + mb_0, \tag{A.25}$$
$$m(m-1)\sigma_0\sigma_2,\ \vec{0}\,[m-2]\,,\ mb_1 + \frac{m(m-1)}{2}\sigma_1^2,\ mb_2 + m(m-1)\sigma_1\sigma_2,\ \frac{m(m-1)}{2}\sigma_2^2,\ \vec{0}\,[m-2]\Big)\,;$$

$$A_{m+1,1} = \Big(\vec{0}\left[\frac{m(m-1)}{2}\right],\ m\sigma_0 s_0,\ \frac{m(m-1)}{2}\sigma_0^2,\ \vec{0}\,[m-2]\,,\ \beta_0 + m(\sigma_0 s_1 + \sigma_1 s_0), \tag{A.26}$$
$$m(b_0 + (m-1)\sigma_0\sigma_1 + \sigma_0 s_2 + \sigma_2 s_0),\ m(m-1)\sigma_0\sigma_2,\ \vec{0}\,[m-2]\,,\ \beta_1 + m\sigma_1 s_1,$$
$$mb_1 + \beta_2 + \frac{m(m-1)}{2}\sigma_1^2 + m(\sigma_1 s_2 + \sigma_2 s_1),\ m(b_2 + (m-1)\sigma_1\sigma_2 + \sigma_2 s_2),\ \frac{m(m-1)}{2}\sigma_2^2,\ \vec{0}\,[m-2]\Big)\,;$$

$$A_{m+w,w} = \Big(\vec{0}\left[\frac{(m+w-2)(m+w-1)}{2} + w - 2\right],\ \frac{w(w-1)}{2}s_0^2,\ mw\sigma_0 s_0,\ \frac{m(m-1)}{2}\sigma_0^2, \tag{A.27}$$
$$\vec{0}\,[m+w-4]\,,\ w(w-1)s_0 s_1,\ w(\beta_0 + (w-1)s_0 s_2 + m(\sigma_0 s_1 + \sigma_1 s_0)),$$
$$m(b_0 + (m-1)\sigma_0\sigma_1 + w(\sigma_0 s_2 + \sigma_2 s_0)),\ m(m-1)\sigma_0\sigma_2,\ \vec{0}\,[m+w-4]\,,\ \frac{w(w-1)}{2}s_1^2,$$
$$w(\beta_1 + (w-1)s_1 s_2 + m\sigma_1 s_1),\ mb_1 + w\beta_2 + \frac{m(m-1)}{2}\sigma_1^2 + mw(\sigma_1 s_2 + \sigma_2 s_1) + \frac{w(w-1)}{2}s_2^2,$$
$$m(b_2 + (m-1)\sigma_1\sigma_2 + w\sigma_2 s_2),\ \frac{m(m-1)}{2}\sigma_2^2, \vec{0}\,[m-2]\Big)\,;$$

$$A_{w+1,w} = \Big(\vec{0}\left[\frac{w(w+1)}{2} - 2\right],\ \frac{w(w-1)}{2}s_0^2,\ w\sigma_0 s_0,\ \vec{0}\,[w-2]\,,\ w(w-1)s_0 s_1, \tag{A.28}$$
$$w(\beta_0 + (w-1)s_0 s_2 + s_0\sigma_1 + \sigma_0 s_1),\ b_0 + w(\sigma_0 s_2 + \sigma_2 s_0),\ \vec{0}\,[w-2]\,,\ \frac{w(w-1)}{2}s_1^2,$$
$$w(\beta_1 + (w-1)s_1 s_2 + \sigma_1 s_1),\ b_1 + w\beta_2 + \frac{w(w-1)}{2}s_2^2 + w(\sigma_1 s_2 + \sigma_2 s_1),\ b_2 + w\sigma_2 s_2\Big)\,;$$

$$A_{w,w} = \Big(\vec{0}\left[\frac{w(w-1)}{2} - 1\right],\ \frac{w(w-1)}{2}s_0^2,\ \vec{0}\,[w-2]\,,\ w(w-1)s_0 s_1,\ w(\beta_0 + (w-1)s_0 s_2), \tag{A.29}$$
$$\vec{0}\,[w-2]\,,\ \frac{w(w-1)}{2}s_1^2,\ w(\beta_1 + (n-1)s_1 s_2),\ w(\beta_2 + \frac{(w-1)}{2}s_2^2)\Big)\,.$$

Proof. Starting from Proposition A.1, we rewrite each equation in vector form, so that $\mathcal{G}(y^m x^w) = A_{m+w,w} H_{m+w,2}(z)$ for $m, w \geq 0$. In particular, we need to add appropriate vectors of 0's in order to fill the gaps, namely we set to 0 the coefficients of the powers not appearing in Eq. (A.14)–(A.21). To do that, we use the fact that the term of the form $y^m x^w$ takes in the vector $H_{m+w,2}(z)$ (and also in all vectors $H_{N,2}(z)$ for $N \geq m+w$) the position $\mathrm{pos}(y^m x^w) := \frac{(m+w)(m+w+1)}{2} + w + 1$ (for example, the term $y^2 x$ is in position $\mathrm{pos}(y^2 x) = 8$, see Eq. (A.12)). This gives the result. □

We now provide the recursion formula for the generator matrix in $\mathbb{R}^2$.

Theorem A.3. *For every $n \geq 2$, the generator matrix $G_{n,2} \in \mathbb{R}^{D \times D}$ associated with the operator $\mathcal{G}$ in Eq. (A.9) with respect to the monomial basis $H_{n,2}(z)$, can be obtained by the following recursive formula:*

$$G_{n,2} = \begin{pmatrix} G_{n-1,2} & \mathbf{0}_n \\ & A_n \end{pmatrix} \quad \text{with} \quad G_{1,2} = \begin{pmatrix} 0 & 0 & 0 \\ b_0 & b_1 & b_2 \\ \beta_0 & \beta_1 & \beta_2 \end{pmatrix}. \tag{A.30}$$

Here $\mathbf{0}_n \in \mathbb{R}^{\frac{n(n+1)}{2} \times (n+1)}$ is a matrix of 0's, and $A_n \in \mathbb{R}^{(n+1) \times D}$ is a matrix with rows $A_{n,0}, A_{n,1}, \ldots, A_{n,n}$ given in Corollary A.2, where the j-th row $A_{n,j}$ corresponds to the coefficients of $\mathcal{G}(y^{n-j}x^j)$, $0 \leq j \leq n$.

Proof. We proceed by induction on the power $n \geq 1$.
$n = 1$: the base case is easily verified from Eq. (A.22) and (A.23), since $\mathcal{G}f(z) = 0$ for $f(z) = 1$.
$n \to n+1$: we assume the statement holds for $n-1$. This means that $G_{n-1,2} \in \mathbb{R}^{\tilde{D} \times \tilde{D}}$ is the generator matrix for $H_{n-1,2}(z)$, where $\tilde{D} = \frac{n(n+1)}{2}$ is the number of basis elements in $\mathcal{B}_{n-1}$. In particular, by passing from $\mathcal{B}_{n-1}$ to $\mathcal{B}_n$ we need to add $D - \tilde{D} = n+1$ elements of the form $y^{n-j}x^j$ for $0 \leq j \leq n$. From Proposition A.1 and Corollary A.2, the extended generator applied to $y^{n-j}x^j$ is expressed by $\mathcal{G}(y^{n-j}x^j) = A_{n,j} H_{n,2}$, where $A_{n,j} \in \mathbb{R}^D$. By defining $A_n \in \mathbb{R}^{(n+1) \times D}$ as the matrix with rows $A_{n,0}, A_{n,1}, \ldots, A_{n,n}$, we then obtain $G_{n,2}$ by stacking A_n under $G_{n-1,2}$. In order to make dimensions to match, we also add a matrix of 0's of dimensions $\frac{n(n+1)}{2} \times (n+1)$, where $\frac{n(n+1)}{2}$ is the squared dimension of $G_{n-1,2}$, and $n+1$ is the number of basis elements we added to $H_{n-1,2}(z)$ to obtain $H_{n,2}(z)$. □

B Proofs

This section contains the proofs of the main results together with a lemma that is needed for the proofs.

Lemma B.1. *We introduce the map $t_{a,b} : \mathbb{R} \to \mathbb{R}$, $x \mapsto \frac{x-a}{b}$. Then:*

1. *The weight function $w_{a,b}$ is obtained by composing w with $t_{a,b}$, namely $w_{a,b}(x) = w\left(\frac{x-a}{b}\right)$.*
2. *The generalized Hermite polynomial $q_n^{a,b}$ is obtained by composing the Hermite polynomial q_n with $t_{a,b}$ and by scaling with the inverse of b^n, namely $q_n^{a,b}(x) = \frac{1}{b^n} q_n\left(\frac{x-a}{b}\right)$, $n \geq 0$.*

Proof. The first part of the lemma is easily verified. For the second part, we proceed by induction on the order $n \geq 0$. Since for $n = 0$, $q_0^{a,b}(x) = q_0(x) = 1$ is trivial, we start from $n = 1$.
$n = 1$: one finds that $q_1(x) = x$ and $q_1^{a,b}(x) = \frac{x-a}{b^2} = \frac{1}{b} q_1\left(\frac{x-a}{b}\right)$, so the base case holds;
$n \to n+1$: we assume the statement holds for n. This means that

$$(-1)^n e^{\frac{(x-a)^2}{2b^2}} \frac{d^n}{dx^n} e^{-\frac{(x-a)^2}{2b^2}} = \frac{1}{b^n} q_n\left(\frac{x-a}{b}\right), \tag{B.1}$$

hence

$$(-1)^n \frac{d^n}{dx^n} e^{-\frac{(x-a)^2}{2b^2}} = \frac{1}{b^n} e^{-\frac{(x-a)^2}{2b^2}} q_n\left(\frac{x-a}{b}\right). \tag{B.2}$$

We now focus on $q_{n+1}^{a,b}$: by induction hypothesis

$$q_{n+1}^{a,b}(x) = (-1)^{n+1} e^{\frac{(x-a)^2}{2b^2}} \frac{d^{n+1}}{dx^{n+1}} e^{-\frac{(x-a)^2}{2b^2}} = (-1) e^{\frac{(x-a)^2}{2b^2}} \frac{d}{dx}\left(\frac{1}{b^n} e^{-\frac{(x-a)^2}{2b^2}} q_n\left(\frac{x-a}{b}\right)\right)$$

$$= \frac{1}{b^{n+1}} \left(\left(\frac{x-a}{b} \right) q_n \left(\frac{x-a}{b} \right) - q_n' \left(\frac{x-a}{b} \right) \right) = \frac{1}{b^{n+1}} q_{n+1} \left(\frac{x-a}{b} \right), \tag{B.3}$$

where the last equality is due to the recurrence relation for Hermite polynomials, namely $q_{n+1}(y) = yq_n(y) - q_n'(y)$, see [6]. This concludes the proof. □

Proof of Lemma 3.1

Proof. From the definition of norm in $L^2_{a,b}$ and Lemma B.1, we get

$$\left\| q_n^{a,b} \right\|^2_{L^2_{a,b}} = \int_{-\infty}^{\infty} \left(q_n^{a,b}(x) \right)^2 w_{a,b}(x) dx = \frac{b}{b^{2n}} \int_{-\infty}^{\infty} \left(q_n(y) \right)^2 w(y)\, dy = \frac{\sqrt{2\pi} n!}{b^{2n-1}}, \tag{B.4}$$

where we have used the change of variables $y = \frac{x-a}{b}$ and Eq. (3.4). □

Proof of Proposition 3.2

Proof. We start by proving that the value of the integrals in the series (3.7) is

$$\int_{-\infty}^{\infty} \varphi_K(y) q_n^{a,b}(y) w_{a,b}(y) dy = \begin{cases} b\sqrt{2\pi} \left(b\,\phi\left(\frac{K-a}{b}\right) + (a-K)\left(1 - \Phi\left(\frac{K-a}{b}\right)\right)\right) & \text{for } n = 0 \\ b\sqrt{2\pi}\left(1 - \Phi\left(\frac{K-a}{b}\right)\right) & \text{for } n = 1\,. \\ \sqrt{2\pi}\phi\left(\frac{K-a}{b}\right) q_{n-2}^{a,b}(K) & \text{for } n \geq 2 \end{cases} \tag{B.5}$$

The first two equalities are easily verified with $q_0^{a,b}(y) = 1$ and $q_1^{a,b}(y) = \frac{y-a}{b^2}$ by, possibly, the change of variables $z = \frac{y-a}{b}$ and by integrating (twice) by parts. For $n \geq 2$, we integrate by parts twice:

$$\begin{aligned} \int_{-\infty}^{\infty} \varphi_K(y) q_n^{a,b}(y) w_{a,b}(y) dy &= (-1)^n \int_K^{\infty} (y-K) \frac{d^n}{dy^n} e^{-\frac{(y-a)^2}{2b^2}} dy = (-1)^{n-2} \frac{d^{n-2}}{dy^{n-2}} e^{-\frac{(y-a)^2}{2b^2}} \Big|_{y=K} \\ &= e^{-\frac{(K-a)^2}{2b^2}} \left((-1)^{n-2} e^{\frac{(y-a)^2}{2b^2}} \frac{d^{n-2}}{dy^{n-2}} e^{-\frac{(y-a)^2}{2b^2}} \right) \Big|_{y=K} = \sqrt{2\pi}\phi\left(\frac{K-a}{b}\right) q_{n-2}^{a,b}(K)\,. \end{aligned} \tag{B.6}$$

From Eq. (3.7) and (B.5), the function φ_K is then expressed in terms of the GHPs by

$$\varphi_K^{a,b}(x) = \sum_{n=0}^{\infty} \alpha_n^{a,b} q_n^{a,b}(x) \qquad \text{with } \alpha_n^{a,b} := \begin{cases} b\,\phi\left(\frac{K-a}{b}\right) + (a-K)\left(1 - \Phi\left(\frac{K-a}{b}\right)\right) & \text{for } n = 0 \\ b^2\left(1 - \Phi\left(\frac{K-a}{b}\right)\right) & \text{for } n = 1\,. \\ \frac{b^{2n-1}}{n!}\phi\left(\frac{K-a}{b}\right) q_{n-2}^{a,b}(K) & \text{for } n \geq 2 \end{cases} \tag{B.7}$$

The result then follows by Lemma B.1. □

Proof of Theorem 4.1

Proof. Starting from Eq. (3.14), for every $N \geq 0$ we can express $\Pi_{K,N}^{a,b}(t)$ in matrix form by

$$\Pi_{K,N}^{a,b}(t) = \boldsymbol{\beta}_N^{a,b\,\top} M_N \boldsymbol{E}_N^X(T;t) \qquad \text{where} \qquad \boldsymbol{E}_N^X(T;t) := \mathbb{E}\left[H_N\left(\frac{X(T)-a}{b} \right) \middle| \mathcal{F}_t \right]. \tag{B.8}$$

Thus we need to compute the entries of $\boldsymbol{E}_N^X(T;t)$. By the binomial theorem, its k-th component, $k = 1, \ldots, N+1$, is of the form

$$\boldsymbol{E}_N^X(T;t)_k = \frac{1}{b^{k-1}} \mathbb{E}\left[(X(T)-a)^{k-1} \middle| \mathcal{F}_t \right] = \frac{1}{b^{k-1}} \sum_{i=0}^{k-1} \binom{k-1}{i} (-a)^{k-1-i}\, \mathbb{E}\left[X(T)^i \middle| \mathcal{F}_t \right]. \tag{B.9}$$

Then, for $\hat{\boldsymbol{\beta}}_N^{a,b} := \boldsymbol{\beta}_N^{a,b\top} M_N$, we rewrite $\Pi_{K,N}^{a,b}(t)$ in Eq. (B.8) by expanding the matrix multiplication into the following sum

$$\Pi_{K,N}^{a,b}(t) = \sum_{j=1}^{N+1} \hat{\boldsymbol{\beta}}_{N,j}^{a,b} \boldsymbol{E}_N^X(T;t)_j = \sum_{k=0}^{N} \hat{\boldsymbol{\beta}}_{N,k+1}^{a,b} \frac{1}{b^k} \sum_{i=0}^{k} \binom{k}{i} (-a)^{k-i} \mathbb{E}\left[X(T)^i \middle| \mathcal{F}_t\right], \tag{B.10}$$

where the last equality we used the change $k = j - 1$. This concludes the proof. □

Proof of Theorem 4.4

Proof. By definition of $\Pi_K(t)$ and $\Pi_{K,N}^{a,b}(t)$, we write that

$$\begin{aligned}\left|\Pi_K(t) - \Pi_{K,N}^{a,b}(t)\right| &= \left|\mathbb{E}\left[\varphi_K(X(T))\middle|\mathcal{F}_t\right] - \mathbb{E}\left[\varphi_{K,N}^{a,b}(X(T))\middle|\mathcal{F}_t\right]\right| \\ &\leq \mathbb{E}\left[\left|\varphi_K(X(T)) - \varphi_{K,N}^{a,b}(X(T))\right|\middle|\mathcal{F}_t\right] \end{aligned} \tag{B.11}$$

were the last inequality is due to Jensen's inequality. By definition of conditional expectation, and being $\psi_{X(T)}$ the density function of $X(T)$, by the Cauchy–Schwartz inequality, this becomes

$$\begin{aligned}\mathbb{E}\left[\left|\varphi_K(X(T)) - \varphi_{K,N}^{a,b}(X(T))\right|\middle|\mathcal{F}_t\right] &= \int_{\mathbb{R}} \left|\varphi_K(x) - \varphi_{K,N}^{a,b}(x)\right| \psi_{X(T)}(x)dx \\ &\leq \left(\int_{\mathbb{R}} \left|\varphi_K(x) - \varphi_{K,N}^{a,b}(x)\right|^2 \omega_{a,b}(x)dx\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}} \psi_{X(T)}^2(x)\omega_{a,b}^{-1}(x)dx\right)^{\frac{1}{2}}, \end{aligned} \tag{B.12}$$

which concludes the proof. □

Proof of Proposition 4.5

Proof. From Eq. (4.8), we see that for $C_{a,b}$ to make sense condition (4.9) must hold. In particular, it cannot be $b < \underline{b}_\sigma$ because of the squared root, while the expression becomes singular for $b = \underline{b}_\sigma$, hence, in this case, we can expect instabilities in the approximation. □

Proof of Corollary 4.6

Proof. By taking $a = \mu$ in Eq. (4.8), we obtain that

$$C_{a,b}^2 = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{b}{\sqrt{2b^2 - \sigma^2}} \qquad \text{and} \qquad \frac{\partial C_{a,b}^2}{\partial b} = -\frac{1}{\sqrt{2\pi\sigma^2}} \frac{\sigma^2}{(2b^2 - \sigma^2)^{\frac{3}{2}}} < 0, \tag{B.13}$$

hence $C_{a,b}^2$ is decreasing in b. The limit is easily computed. □

Proof of Proposition 5.1

Proof. One can observe that

$$\int_t^{s_j} e^{b_1(s_j - v)} dB(v) = \sum_{k=0}^{j} \int_{s_{k-1}}^{s_k} e^{b_1(s_j - v)} dB(v), \quad \text{for} \quad j = 0, \ldots, m, \tag{B.14}$$

hence, by rearranging the terms in the two summations, we write that

$$\sum_{j=0}^{m} \int_t^{s_j} e^{b_1(s_j - v)} dB(v) = \sum_{j=0}^{m} \sum_{k=0}^{j} \int_{s_{k-1}}^{s_k} e^{b_1(s_j - v)} dB(v) = \sum_{k=0}^{m} \int_{s_{k-1}}^{s_k} \left(\sum_{j=k}^{m} e^{b_1(s_j - v)}\right) dB(v). \tag{B.15}$$

By switching the role of k and j, we get the result. □

C Some Details on the Correlator Formula

We briefly report the definitions needed to understand the correlator formulas in Theorem 2.3. For more details and for the idea behind this construction, we refer the reader to [4].

Vectorization Given a matrix $A \in \mathbb{R}^{n\times m}$ whose j-th column we denote by $A_{:j}$, we define the *vectorization* of A as the operator $vec : \mathbb{R}^{n\times m} \to \mathbb{R}^{nm\times 1}$ that associates to A the nm-column vector $vec(A) = \begin{pmatrix} A_{:1}^\top & A_{:2}^\top \cdots\cdots A_{:m}^\top \end{pmatrix}^\top$.

Inverse-vectorization Given a vector $v \in \mathbb{R}^{nm}$, we define the *inverse-vectorization* of v as the operator $vec^{-1} : \mathbb{R}^{nm} \to \mathbb{R}^{n\times m}$ that associates to v the $n \times m$ matrix $A = vec^{-1}(v)$ such that $[A]_{i,j} = v_{n(j-1)+i}$, for $i = 1, \dots, n$ and $j = 1, \dots, m$.

L-vectorization Given a matrix $A \in \mathbb{R}^{n\times m}$ with elements $[A]_{i,j} = a_{i,j}$, for $1 \le i \le n$ and $1 \le j \le m$, we define the *L-vectorization* of A as the operator $vecL : \mathbb{R}^{n\times m} \to \mathbb{R}^{n+m-1}$ that associates to A the $(n + m - 1)$-column vector obtained by selecting the first column and the last row of A, namely $vecL(A) = \begin{pmatrix} a_{1,1} & a_{2,1} \cdots\cdots a_{n,1} & a_{n,2} \cdots\cdots a_{n,m} \end{pmatrix}^\top$.

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 call $A \in \mathcal{A}_{n,m}$ an *Hankel matrix* and write $\mathcal{A}_n := \mathcal{A}_{n,n}$.

Kronecker product The *Kronecker product* of a matrix $A \in \mathbb{R}^{n\times m}$ with elements $[A]_{i,j} = a_{i,j}$, for $1 \le i \le n$ and $1 \le j \le m$, and a matrix $B \in \mathbb{R}^{r\times s}$, is defined by

$$A \otimes B = \begin{pmatrix} a_{1,1}B & \cdots\cdots & a_{1,m}B \\ \vdots & \ddots\ddots & \vdots \\ a_{n,1}B & \cdots\cdots & a_{n,m}B \end{pmatrix} \in \mathbb{R}^{nr\times ms}. \tag{C.1}$$

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

$$\begin{cases} A \otimes^d B = A^{\otimes d} \otimes B & d \ge 1 \\ A \otimes^0 B = B & d = 0 \end{cases}. \tag{C.2}$$

L-eliminating matrix For $n, m \ge 1$ and $A \in \mathbb{R}^{n\times m}$, we define the L-eliminating matrix as the matrix $E_{n,m} \in \mathbb{R}^{(n+m-1)\times nm}$ such that $E_{n,m}vec(A) = vecL(A)$. We write $E_n := E_{n,n}$.

L-duplicating matrix For $n, m \ge 1$ and $A \in \mathcal{A}_{n,m}$, we define the L-duplicating matrix as the matrix $D_{n,m} \in \mathbb{R}^{nm\times(n+m-1)}$ such that $D_{n,m}vecL(A) = vec(A)$. We write $D_n := D_{n,n}$.

Let now $X_n^{(m)}(x) := H_n(x)^\top \otimes^m H_n(x)$.

m-th L-eliminating matrix For $n, m \ge 1$, we define the *m-th L-eliminating matrix* as the matrix $E_{n+1}^{(m)} \in \mathbb{R}^{(n(m+1)+1)\times(n+1)^{m+1}}$ such that $E_{n+1}^{(m)}vec(X_n^{(m)}(x)) = H_{n(m+1)}(x)$. In particular

$$\begin{cases} E_{n+1}^{(1)} = E_{n+1} & m = 1 \\ E_{n+1}^{(m)} = E_{nm+1,n+1}\left(I_{n+1} \otimes E_{n+1}^{(m-1)}\right) & m \ge 2 \end{cases}. \tag{C.3}$$

m-th L-duplicating matrix For $n, m \ge 1$, we define the *m-th L-duplicating matrix* as the matrix $D_{n+1}^{(m)} \in \mathbb{R}^{(n+1)^{m+1}\times(n(m+1)+1)}$ such that $D_{n+1}^{(m)}H_{n(m+1)}(x) = vec(X_n^{(m)}(x))$. In particular

$$\begin{cases} D_{n+1}^{(1)} = D_{n+1} & m = 1 \\ D_{n+1}^{(m)} = \left(I_{n+1} \otimes D_{n+1}^{(m-1)}\right) D_{nm+1,n+1} & m \ge 2 \end{cases}. \tag{C.4}$$

Acknowledgments The author would like to thank Fred Espen Benth, Vegard Antun and Salvador Ortiz-Latorre for discussions, and two anonymous referees for valuable comments.

References

[1] D. Ackerer & D. Filipović (2020) Option pricing with orthogonal polynomial expansions, *Mathematical Finance* **30** (1), 47–84.

[2] H. Albrecher, P. A. Mayer & W. Schoutens (2008) General lower bounds for arithmetic Asian option prices, *Applied Mathematical Finance* **15** (2), 123-149.

[3] O. E. Barndorff-Nielsen (1997) Processes of normal inverse Gaussian type, *Finance and stochastics* **2** (1), 41–68.

[4] F. E. Benth & S. Lavagnini (2021) Correlators of Polynomial Processes, *SIAM Journal on Financial Mathematics* **12** (4), 1374–1415.

[5] P. Carmona, F. Petit & M. Yor (1997) On the distribution and asymptotic results for exponential functionals of Lévy processes, *Exponential functionals and principal values related to Brownian motion*, 73–121.

[6] G. Djordjevic (1996) On some properties of generalized Hermite polynomials, *Fibonacci Quarterly* **34**, 2–6.

[7] D. Dufresne (2000) Laguerre series for Asian and other options, *Mathematical Finance* **10** (4), 407–428.

[8] D. Filipović & M. Larsson (2016) Polynomial diffusions and applications in finance, *Finance and Stochastics* **20** (4), 931–972.

[9] D. Filipović & M. Larsson (2020) Polynomial Jump-Diffusion Models, *Stochastic Systems* **10** (1), 71–97.

[10] D. Filipović, E. Mayerhofer & P. Schneider (2013) Density approximations for multivariate affine jump-diffusion processes, *Journal of Econometrics* **176** (2), 93–111.

[11] G. Fusai & A. Meucci (2008) Pricing discretely monitored Asian options under Lévy processes, *Journal of banking & finance* **32** (10), 2076–2088.

[12] G. Fusai & I. Kyriakou (2016) General optimized lower and upper bounds for discrete and continuous arithmetic Asian options, *Mathematics of Operations Research* **41** (2), 531–559.

[13] G. Fusai & A. Tagliani (2002) An accurate valuation of Asian options using moments, *International Journal of Theoretical and Applied Finance* **5** (02), 147–169.

[14] H. Geman & M. Yor (1993) Bessel processes, Asian options, and perpetuities, *Mathematical finance* **3** (4), 349–375.

[15] N. Ju (2002) Pricing Asian and basket options via Taylor expansion, *Journal of Computational Finance* **5** (3), 79–103.

[16] A. G. Kemna & A. C. Vorst (1990) A pricing method for options based on average asset values, *Journal of Banking & Finance* **14** (1), 113–129.

[17] B. Lapeyre & E. Temam (2001) Competitive Monte Carlo methods for the pricing of Asian options, *Journal of computational finance* **5** (1), 39–58.

[18] W. Li & S. Chen (2016) Pricing and hedging of arithmetic Asian options via the Edgeworth series expansion approach, *The Journal of Finance and Data Science* **2** (1), 1–25.

[19] V. Linetsky (2004) Spectral expansions for Asian (average price) options, *Operations Research* **52** (6), 856–867.

[20] W. Schoutens (2012) *Stochastic Processes and Orthogonal Polynomials* (Vol. 146) Springer Science & Business Media.

[21] S. M. Turnbull & L. M. Wakeman (1991) A quick algorithm for pricing European average options, *Journal of financial and quantitative analysis* **26** (3), 377–389.

[22] M. Vanmaele, G. Deelstra, J. Liinev, J. Dhaene & M. J. Goovaerts (2006) Bounds for the price of discrete arithmetic Asian options, *Journal of Computational and Applied Mathematics* **185** (1), 51–90.

[23] J. Vecer (2001) A new PDE approach for pricing arithmetic average Asian options, *Journal of computational finance* **4** (4), 105–13.

[24] J. Vecer (2002) Unified pricing of Asian options, *Risk* **15** (6), 113–116.

[25] S. Willems (2019) Asian option pricing with orthogonal polynomials, *Quantitative Finance* **19** (4), 605–618.

[26] R. Weron (2007) *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach* (Vol. 403) John Wiley & Sons.

[27] M. Yor (1992) On some exponential functionals of Brownian motion, *Advances in applied probability* **24** (3), 509–531.