Chebyshev Interpolation for Parametric Option Pricing

Maximilian Gaß[1,*], Kathrin Glau[2],
Mirco Mahlstedt[1,*], Maximilian Mair[1]

[1]Technical University of Munich, Germany,
[2]Queen Mary University of London

November 23, 2017

Abstract

Recurrent tasks such as pricing, calibration and risk assessment need to be executed accurately and in real time. We concentrate on Parametric Option Pricing (POP) as a generic instance of parametric conditional expectations and show that polynomial interpolation in the parameter space promises to considerably reduce run-times while maintaining accuracy. The attractive properties of Chebyshev interpolation and its tensorized extension enable us to identify broadly applicable criteria for (sub)exponential convergence and explicit error bounds. The method is most promising when the computation of the prices is most challenging. We therefore investigate its combination with Monte Carlo simulation and analyze the effect of (stochastic) approximations of the interpolation. For a wide and important range of problems, the Chebyshev method turns out to be more efficient than parametric multilevel Monte-Carlo. We conclude with a numerical efficiency study.

Keywords: Multivariate Option Pricing Complexity Reduction (Tensorized) Chebyshev Polynomials Polynomial Interpolation Fourier Transform Methods Monte-Carlo Parametric Monte-Carlo online-offline decomposition

2000 MSC: 91G60 41A10

1 Introduction

The development of fast and accurate computational methods for parametric models is one of the central issues in computational finance. Financial institutions dedicated to the trading or assessment of financial derivatives have to

*We thank the KPMG Center of Excellence in Risk Management for their support.

cope with the daily tasks of computing numerous characteristic financial quantities. Examples of interest include prices, sensitivities and risk measures for products on different models and for varying parameter constellations. With regard to the ever growing market activities, more and more of these evaluations need to be delivered in real time. In addition we face constantly rising model sophistication since the original work of [1] and [35]. From the early nineties onwards stochastic volatility and Lévy models as well as models based on further classes of stochastic processes have been developed that reflect the observed market data in a more appropriate way. For asset models, see e.g. [27], [12], [10], [8]. In the case of fixed income models, see e.g. [13], [29], [15]. The aftermath of the financial crisis 2007–2009, moreover, has lead to a new generation of more complex models, for instance by incorporating more risk factors. The usefulness of a pricing model critically depends on how well its numerical implementation captures the relevant aspects of market reality. Exploiting new ways to deal with the rising computational complexity therefore supports the evolution of pricing models and touches a core concern of present-day mathematical finance.

A large body of computational tasks in finance needs to be repeatedly performed in real time for a varying set of parameters. Prominent examples include option pricing and hedging of different option sensitivities, e.g. delta and vega, that also need to be calculated in real time. In particular for optimization routines arising in model calibration, large parameter sets come into play. Further examples arise in the context of risk control and assessment, such as for the quantification and monitoring of risk measures. The following question serves as a starting point of our investigations: *How can we systemically exploit the recurrent nature of parametric computational problems in finance with the objective of gaining efficiency?* Looking for answers to this question, we focus in the sequel on Parametric Option Pricing (POP) that we identify as a generic instance.

In the present literature on computational methods in finance, complexity reduction for parametric problems has largely been addressed by applying Fourier techniques following the seminal works of [6] and [41]. See also the monograph [2]. Research in this area concentrates on adopting fast Fourier transform (FFT) methods and variants for option pricing. [33] accurately describes pricing European options with FFT. Further developments are for instance provided by [34] for early exercise options and by [14] and [31] for barrier options. Another path towards efficiently handling large parameter sets that has been pursued in finance relies on reduced basis methods. These are techniques for solving parametrized partial differential equations. [44], [7], [38] and [23] and [4] applied this approach to price European, American plain vanilla options and European baskets. FFT methods on the one hand can be highly beneficial when the prices are required in a large number of Fourier variables, e.g. for a large set of strikes of European plain vanillas. A method that tailors Fourier pricing to the whole parametric family of integrands has recently been developed in [17]. On the other hand numerical experiments have shown a promising gain in efficiency of reduced basis methods when an accurate PDE solver is readily available. In essence, all these approaches reveal an immense potential of complexity reduction by targeting parameter dependence. To do this, they exploit the functional architecture of the underlying pricing technique for varying parameters.

Financial institutions have to deal simultaneously with a diversity of models, a multitude of option types, and—as a consequence—a wide variety of underlying pricing techniques. It is therefore worthwhile to explore the possibility of a *generic* complexity reduction method that is independent of the specific pricing technique. To do so, we focus on the set of option prices and the set of parameters of interest, deliberately disregard the pricing technology and view the option price as a function of the parameters. The core idea is now to *introduce interpolation of option prices in the parameter space as a complexity reduction technique for POP*.

The resulting procedure naturally splits into two phases: Pre-computation and real-time evaluation. The first one is also called the *offline phase* while the second is also called the *online phase*. In the pre-computation phase, the interpolation is made available. In the case of polynomial interpolation, this steps amounts to the computation of the coefficients for the basis functions. The actual procedure depends on the choice of the interpolation method. In all cases, however, the prices have to be computed for some fixed parameter configurations. Here, any appropriate pricing method—for instance based on Fourier, PDE or even Monte-Carlo techniques—can be chosen. The real-time evaluation phase then consists in evaluating the interpolation. Provided that the evaluation of the interpolation is faster than the benchmark tool, the scheme permits a gain in efficiency in all cases where accuracy can be maintained. Then, we distinguish several use cases:

- In comparison to the benchmark pricing routine, the fast evaluation of the interpolation will eventually outweigh the expensive pre-computation phase, if pricing is a task which is repeatedly employed. Optimization procedures are an obvious instance where this feature becomes advantageous.

- The interpolation can simultaneously deliver a multitude of outputs. For instance as we will see in the sequel the interpolation can be setup such that it delivers sensitivities as well.

- Even if the number of price computations is limited, we can still benefit from separating the procedure into its two phases. In this way, e.g., idle times in the financial industry can be put to good use by preparing the interpolation for whenever real-time pricing is needed during business activities.

The question arising at this stage is: *Under what circumstances can we hope* *to find an interpolation method that delivers both reliable results and a considerable gain in efficiency?*

One could now be tempted to proceed in a naive manner and first define an equidistant grid and then interpolate piecewise linearly in the parameter space. Numerical experiments for Black&Scholes call prices as function of the volatility, for instance, would then provide convincing evidence that the number of nodes needed for a given accuracy is considerably high. Increasing the polynomial degree might lead to better results. However, convergence might not be guaranteed. [43] showed that polynomial interpolation on equispaced grids may diverge—even for analytic functions. Second, the evaluation of the polynomial interpolants may be numerically problematic, as it is shown in [43] that "the interpolation problem for polynomial interpolation on an equidistant grid is exponentially ill-conditioned", a formulation we borrow from [47]. For these reasons, we abstain from polynomial interpolation with equidistant grids. Rather, we take a step back and ask: *Which methods of interpolating prices as functions of model and payoff parameters are numerically promising in terms of convergence, stability and implementational ease?*

Regarding this research question, we need to take into consideration both the set of interpolation methods as such and the specific features of the functions we investigate. It is well-known that the efficiency of interpolation methods critically depends on the degree of regularity of the approximated function. For the core problem of our study—European (basket) options—we investigate the regularity of the option prices as functions of the parameters. We find that these functions are indeed *analytic* for a large set of option types, models and parameters. Taking the perspective of approximation theory, this inspires the hope that suitable interpolation methods can be found. In particular, it is well-known that orthogonal polynomial interpolation yields (sub)exponential convergence in this case.

In this article we propose and investigate the interpolation of financial quantities in the parameter space by *Chebyshev polynomials*. This has various reasons. We empirically observe that parameters of interest often range within bounded intervals, and Chebyshev polynomial interpolation is well-known for its excellent numerical properties in approximating analytic functions on bounded intervals. Their following key properties are of particular interest for our purposes:

- For univariate functions that are several times differentiable, the method converges polynomially and, for univariate analytic functions, *convergence is exponential*—in stark contrast to polynomial interpolation on equally spaced nodal points. Even more, Chebyshev polynomials appear as an optimal choice when minimizing the error in a certain way among the nodal polynomial interpolations, see Appendix A.

- The method can be implemented in a *numerically stable* way. This is crucial for its actual performance.

- The interpolation nodes are explicitly available and thus *the coefficients are explicitly given* as a linear transformation of the function values at the nodel points. On the one hand this makes the implementation straightforward, a feature that is valuable for both the application of the method in complex IT infrastructures of financial industry and for further developments of the method. On the other hand, since the interpolation nodes are explicitly given we can avoid a significant approximation step, typically a regression, and therewith a major source of inaccuracy.

- The derivatives are trivial for the interpolation and known to converge as well with a rate that is determined by the regularity of the function that is interpolated. Thus *sensitivities are additional outputs* of high accuracy.

- Chebyshev interpolation can be easily concatinated yielding *Chebyshev-spline approximation*, which is extremely appealing when the function exhibits discontinuities for instance.

- Chebyshev interpolation can be highly efficiently *extended to higher dimensionality*, for instance by low-rank tensor and sparse grid techniques.

In a remarkable monograph, [48] gives a comprehensive review of Chebyshev interpolation. Its appealing theoretical properties are indeed of practical use as the software tool Chebfun[1] demonstrates. In this implementation, [40] aim "to combine the feel of symbolics with the speed of numerics". Exploring the potential of interpolation methods for more than one single free parameter, we choose a tensorized version of Chebyshev interpolation:

For parameters $p \in [-1,1]^D$, where $D \in \mathbb{N}$ denotes the dimensionality of the parameter space, the price $Price^p$ is approximated by tensorized Chebyshev polynomials T_j with pre-computed coefficients c_j, $j \in J$, as follows,

$$Price^p \approx \sum_{j\in J} c_j T_j(p).$$

Chebyshev interpolation is a standard numerical method that has proven to be extremely useful for applications in such diverse fields as physics, engineering, statistics and economics. Nevertheless, for pricing tasks in mathematical finance, Chebyshev interpolation still seems to be rarely used and its potential is yet to be unfolded. [39] use Chebyshev interpolation of Black&Scholes prices in the volatility as an intermediate step to derive a pricing methodology for a time-changed model. Independently from us, [37] recently proposed Chebyshev interpolation as a quadrature rule for the computation of option prices with a Fourier-type representation, which is comparable to the cosine method.

Our main results are the following:

[1] Chebfun is an open-source software system, see `http://www.chebfun.org`

- Proposition 2.1 refines the known result that analyticity guarantees an asymptotic error decay of order $O\big(\varrho^{-\sqrt[D]{N}}\big)$ in the total number N of interpolation nodes where $\varrho > 1$ is given by the domain of analyticity and D is the number of varying parameters.

- Proposition 2.3 and 2.4 show convergence results for the related sensitivities.

The method is promises the highest gain in efficiency for the most challenging and therefore most computationally extensive problems. In these cases, the computation of the values at the nodal points cannot be delivered at machine precision but is affected by an approximation error. This approximation error in turn affects the accuracy of the interpolation.

- Theorem 2.5 therefore provides an error bound including distorsions of the nodal points. We consider two types of distorsions, those bounded by a common deterministic threshold and those that are normally distributed. The first case is tailored to an underlying pricing method that is accurate up to a pre-specified deterministic accuracy, the latter to the computation of the values at the nodal points by Monte Carlo.

Qualitatively, we illustrate the relation to advanced Monte-Carlo techniques and compare our approach with the parametric multilevel Monte-Carlo approach of [24] and [26]. We derive a theoretical result of the "offline efficiency", i.e. the asymptotic rate of convergence in terms of the offline cost. This is a measure for the accuracy versus the offline cost:

- In Theorem 2.6, we show that for each $\beta > D/2$ there exist constants $\bar{c}_1, \bar{c}_2 > 0$ such that the offline cost is bounded by $\bar{c}_1\mathcal{M}$ and the expected error of the Chebyshev method is bounded by $\bar{c}_2 \log(\mathcal{M})^\beta \mathcal{M}^{-1/2}$.

In Section 3, we introduce the general framework of POP (parametric option pricing).

- Theorem 3.2 provides accessible sufficient conditions on options and models that guarantee analyticity in the parameters. Moreover it establishes a method to access the domain of analyticity. In combination with Proposition 2.1 this allows to conclude for (sub)exponential convergence rate, as well as to access the constants in the exact error bound.

This motivates us to further explore the potential of the Chebyshev method for multivariate options. Here we also deliberately go beyond the scope of our theoretical results and consider additional features like path-dependency. We present empirical results demonstrating the efficiency of the Chebyshev method:

- The explicit gains in efficiency in comparison to standard Monte-Carlo methods are shown in Section 4.2, taking multivariate lookback options in the Heston model as examples.

To conclude this introduction we mention some of the areas, where we expect the application of this approach to be especially fruitful, namely the

- approximation of the implied volatility, see [20],
- development of efficient solvers for stochastic dynamic programming problems, see [21],
- acceleration of calibration procedures,
- acceleration of nested Monte Carlo simulations, for instance in risk scenario evaluation.

The remainder of the article is organized as follows. In Section 2 we introduce Chebyshev interpolation in detail and present the general error estimates and convergence results. Section 3 establishes a convergence analysis of Chebyshev interpolation for POP. The numerical experiments in Section 5 confirm these findings using Fourier techniques. The gain in efficiency when pricing basket options is numerically investigated. Experiments based on Monte-Carlo and finite differences moreover suggest to further explore the potential of the approach beyond the scope of the theoretical investigations from the previous sections. We conclude the section with complexity considerations and by discussing the relation to advanced Monte-Carlo techniques. The resulting conclusion and outlook are presented in Section 6. Finally, the appendix provides the proof of the multivariate convergence result.

2 Chebyshev Polynomial Interpolation

Let us introduce the notation for Chebyshev interpolation along with its tensor-based extension to the multivariate case, see e.g. [45]. In order to obtain convenient notation, consider the interpolation of prices

$$Price^p, \qquad p \in [-1,1]^D. \tag{1}$$

For a more general hyperrectangular parameter space $\mathcal{P} = [\underline{p}_1, \overline{p}_1] \times \ldots \times [\underline{p}_D, \overline{p}_D]$, appropriate linear transformations need to be applied. Let $\overline{N} := (N_1, \ldots, N_D)$ with $N_i \in \mathbb{N}_0$ for $i = 1, \ldots, D$. The interpolation, which has $\prod_{i=1}^{D}(N_i + 1)$ summands, is given by

$$I_{\overline{N}}(Price^{(\cdot)})(p) := \sum_{j \in J} c_j T_j(p), \tag{2}$$

where the summation index j is a multi-index ranging over $J := \{(j_1, \ldots, j_D) \in \mathbb{N}_0^D : j_i \leq N_i \text{ for } i = 1, \ldots, D\}$, the basis functions

$$T_j(p_1, \ldots, p_D) = \prod_{i=1}^{D} T_{j_i}(p_i), \qquad T_{j_i}(p_i) := \cos\big(j_i \arccos(p_i)\big), \tag{3}$$

the coefficients

$$c_j = \Big(\prod_{i=1}^{D} \frac{2^{\mathbb{1}_{\{0<j_i<N_i\}}}}{N_i} \Big) \sum_{k_1=0}^{N_1}{}'' \cdots \sum_{k_D=0}^{N_D}{}'' \, Price^{p^{(k_1,\ldots,k_D)}} \prod_{i=1}^{D} \cos\left(j_i \pi \frac{k_i}{N_i} \right), \tag{4}$$

where $\sum''$ indicates that the first and last summands are halved, and the Chebyshev nodes p^k for the multi-index $k = (k_1, \ldots, k_D) \in J$ given by

$$p^k = (p_{k_1}, \ldots, p_{k_D}), \qquad p_{k_i} = \cos\left(\pi \frac{k_i}{N_i} \right). \tag{5}$$

At the $N_i + 1$ points p_{k_i}, the Chebyshev polynomial $T_{N_i}(x)$ reaches its extreme values. These points are also referred to as Chebyshev-Lobatto points, Chebyshev extreme points, or Chebyshev points of the second kind and satisfy a discrete orthogonality property, see (26).[2]

2.1 Exponential Convergence

The following proposition is a slight improvement of [45, Lemma 7.3.3]. Namely, the error bound is given in terms of the constants $\varrho_1, \ldots, \varrho_D$ rather than by $\underline{\varrho} := \min_{i=1}^{D} \varrho_i$. For our purpose, it is worth to provide this generalization since the different Bernstein radii ϱ_i are accessible in examples, compare Section (3). This allows to derive stricter error bounds and improved assertions on the convergence rates.

Proposition 2.1. *Let $\mathcal{P} \ni p \mapsto Price^p$ be a real-valued function that has an analytic extension to some generalized Bernstein ellipse $B(\mathcal{P}, \varrho)$ for some parameter vector $\varrho \in (1, \infty)^D$ and suppose that $\sup_{p \in B(\mathcal{P},\varrho)} |Price^p| \leq V$. Then*

$$\max_{p \in \mathcal{P}} \left| Price^p - I_{\overline{N}}(Price^{(\cdot)})(p) \right| \leq 2^{\frac{D}{2}+1} \cdot V \cdot \left(\sum_{i=1}^{D} \varrho_i^{-2N_i} \prod_{j=1}^{D} \frac{1}{1 - \varrho_j^{-2}} \right)^{\frac{1}{2}}.$$

The proof of the proposition is provided in Appendix B and is based on the proof of [45, Lemma 7.3.3]. A further improved error bound for tensorized Chebyshev interpolation is presented in [22]. Sharper error bounds reduce the computational cost, because they allow us to use fewer nodal points to achieve a given accuracy.

Corollary 2.2. *Under the assumptions of Proposition 2.1, there exists a constant $C > 0$ such that*

$$\max_{p \in \mathcal{P}} \left| Price^p - I_{\overline{N}}(Price^{(\cdot)})(p) \right| \leq C \underline{\varrho}^{-\underline{N}}, \tag{6}$$

where $\underline{\varrho} = \min_{1 \leq i \leq D} \varrho_i$ *and* $\underline{N} = \min_{1 \leq i \leq D} N_i$.

[2] According to [48], these points are more often applied than the $N_i + 1$ zeros of $T_{N_i+1}(x)$ as nodal points, which are the Chebyshev points of the first kind.

In particular, under the assumptions of Proposition 2.1, where $N = \prod_{i=1}^{D}(N_i + 1)$ denotes the total number of nodes, Corollary 2.2 shows that the error decay has (sub)exponential order $O\big(\varrho^{-\sqrt[D]{N}}\big)$ for some $\varrho > 1$.

2.2 Convergence of the Sensitivities

The sensitivities play a crucial role in financial applications. We therefore state convergence results for the partial derivatives as well. The one-dimensional result is shown in [46] and a multivariate result is derived in [5] for functions in Sobolev spaces. These results allow us to obtain the Chebyshev approximation of derivatives at no additional cost. To state the corresponding convergence results, we follow the approach of [5] and introduce the weighted Sobolev spaces for $\sigma \in \mathbb{N}$ as

$$W_2^{\sigma,\omega}(\mathcal{P}) = \left\{\phi \in L^2(\mathcal{P}) : \|\phi\|_{W_2^{\sigma,\omega}(\mathcal{P})} < \infty\right\}, \tag{7}$$

with norm

$$\|\phi\|^2_{W_2^{\sigma,\omega}(\mathcal{P})} = \sum_{|\alpha|\leq\sigma} \int_{\mathcal{P}} |\partial^\alpha \phi(p)|^2 \omega(p)\, \mathrm{d}\, p, \tag{8}$$

where $\alpha = (\alpha_1, \dots, \alpha_D) \in \mathbb{N}_0^D$ is a multi-index, $\partial^\alpha = \partial^{\alpha_1} \cdots \partial^{\alpha_D}$, and the weight function ω on $\mathcal{P}$ is given by

$$\omega(x) := \prod_{j=1}^{D} \omega(\tau^{-1}_{[\underline{p}_j,\overline{p}_j]}(x_j)), \qquad \omega(\tau^{-1}_{[\underline{p}_j,\overline{p}_j]}(x_j)) := (1 - \tau^{-1}_{[\underline{p}_j,\overline{p}_j]}(x_j)^2)^{-\frac{1}{2}}$$

with $\tau_{[\underline{p}_j,\overline{p}_j]}(p) = \overline{p}_j + \frac{\underline{p}_j - \overline{p}}{2}\big(1-p\big)$. We are now in a position to present the following result.

Proposition 2.3. *Let $\mathcal{P} \ni p \mapsto Price^p \in W_2^{\sigma,\omega}(\mathcal{P})$ and set $N_i = N$, $i = 1, \dots, D$, i.e. the same number of nodal points in each dimension. Then for any $\frac{D}{2} < \sigma \in \mathbb{N}$ and any $\sigma \geq \mu \in \mathbb{N}_0$ there exists a constant $C > 0$ such that*

$$\|Price^{(\cdot)} - I_{\overline{N}}(Price^{(\cdot)})(\cdot)\|_{W_2^{\mu,\omega}(\mathcal{P})} \leq C N^{2\mu-\sigma} \|Price^{(\cdot)}\|_{W_2^{\sigma,\omega}(\mathcal{P})},$$

The proof of the proposition is provided in Appendix C and applies [5, Theorem 3.1] together with the transformation τ.

The result in Proposition 2.3 is given in terms of weighted Sobolev norms. In the following proposition, we relate the approximation error in the weighted Sobolev norm to the $C^l(\mathcal{P})$ norm, where $C^l(\mathcal{P})$ is the Banach space of all functions u in $\mathcal{P}$ such that u and $\partial^\alpha u$ with $|\alpha| \leq l$ are uniformly continuous on $\mathcal{P}$ and the norm

$$\|u\|_{C^l(\mathcal{P})} = \max_{|\alpha|\leq l} \max_{p\in\mathcal{P}} |\partial^\alpha u(p)|$$

is finite.

Proposition 2.4. *Let $\mathcal{P} \ni p \mapsto Price^p \in W_2^{\sigma,\omega}(\mathcal{P})$ and set $N_i = N$, $i = 1, \ldots, D$, i.e. the same number of nodal points in each dimension. Then, for any $\frac{D}{2} < \sigma \in \mathbb{N}$ and any $\sigma \geq \mu \in \mathbb{N}_0$ and $l \in \mathbb{N}_0$ with $\mu - l > \frac{D}{2}$, there exists a constant $\bar{C}(\sigma) > 0$ depending on σ such that*

$$\|Price^{(\cdot)} - I_{\overline{N}}(Price^{(\cdot)})(\cdot)\|_{C^l(\mathcal{P})} \leq \bar{C}(\sigma) N^{2\mu-\sigma} \max_{|\alpha|\leq\sigma} \sup_{p\in\mathcal{P}} |\partial^\alpha Price^p|.$$

The proof of the proposition is elementary and combines [5, Theorem 3.1] and [49, Corollary 6.2].

2.3 Interaction of Approximation Errors at Nodal Points and Interpolation Errors

The Chebyshev method is most promising for use cases where computationally intensive pricing methods are required. In such cases, the issue of distorted prices and their consequences arises naturally when computing the prices at the Chebyshev nodes. The observed noisy prices at the Chebyshev nodes are

$$Price_\varepsilon^{p^{(k_1,\ldots,k_D)}} = Price^{p^{(k_1,\ldots,k_D)}} + \varepsilon^{p^{(k_1,\ldots,k_D)}},$$

where $\varepsilon^{p^{(k_1,\ldots,k_D)}}$ is the approximation error introduced by the underlying numerical technique at the Chebyshev nodes. By linearity, the resulting interpolation is of the form

$$I_{\overline{N}}(Price_\varepsilon^{(\cdot)})(p) = I_{\overline{N}}(Price^{(\cdot)})(p) + I_{\overline{N}}(\varepsilon^{(\cdot)})(\cdot) \tag{9}$$

with error function

$$\varepsilon^{(p)} = \sum_{j_D=0}^{N_D} \cdots \sum_{j_1=0}^{N_1} c^\varepsilon_{j_1,\ldots,j_D} T_{j_1,\ldots,j_D}(p), \tag{10}$$

where the coefficients c_j^ε for $j = (j_1, \ldots, j_D) \in J$ are given by

$$c_j^\varepsilon = \Big(\prod_{i=1}^{D} \frac{2^{\mathbb{1}_{\{0<j_i<N_i\}}}}{N_i}\Big) \sum_{k_1=0}^{N_1}{}'' \cdots \sum_{k_D=0}^{N_D}{}'' \varepsilon^{p^{(k_1,\ldots,k_D)}} \prod_{i=1}^{D} \cos\left(j_i \pi \frac{k_i}{N_i}\right). \tag{11}$$

We are interested in two types of distorsions ε^{p^j} for the multiindices $j = (k_1, \ldots, k_D) \in J$. First, analysing the case that all distorsions are bounded by a fixed constant $\overline{\varepsilon}$ will give a stability result. Second, when computing the values at the nodal points independently with a crude Monte Carlo method, the distorsions will be i.i.d. and asympotically normally distributed, $\varepsilon^{p^j} \cong N(0, \sigma_M)$ with $\sigma_M = \sigma/\sqrt{M}$. Third, in practice it will often turn out to be considerably more efficient to compute the values at the nodal points in the offline phase in a stochastically dependent way. For instance it is often advantageous to sample the driving stochastic process once and reuse it for the whole set of parameters p^j.

Formally, we distinguish the following cases

(i) $|\varepsilon^{p^j}| \le \overline{\varepsilon}$ for all $j \in J$,

(ii) ε^{p^j} is normally distributed with distribution $\mathcal{N}(0, \sigma_{j,M})$ for all multi-indices $j \in J$ (not necessarily independent).

In order to express the error bounds for these different cases, let

$$\varepsilon^*(N_1, \dots, N_D) := \begin{cases} \overline{\varepsilon}, & \text{in case (i)} \\ \sqrt{2log\Big(2\prod_{i=1}^D (N_i+1)\Big)} max_{j\in J}\sigma_{j,M}, & \text{in case (ii).} \end{cases}$$

Theorem 2.5. *Let $\mathcal{P} \ni p \mapsto Price^p$ be given as in Proposition 2.1 and asume one of the conditions (i), respectively (ii), for the distorsions ε^{p^j}, $j \in J$. Then the interpolation including the distorsions, $I_{\overline{N}}(Price_{\varepsilon}^{(\cdot)})$ satisfies*

$$E\left(\max_{p\in\mathcal{P}} \left|Price^p - I_{\overline{N}}(Price_{\varepsilon}^{(\cdot)})(p)\right|\right)$$
$$\le 2^{\frac{D}{2}+1} \cdot V \cdot \left(\sum_{i=1}^D \varrho_i^{-2N_i} \prod_{j=1}^D \frac{1}{1-\varrho_j^{-2}}\right)^{\frac{1}{2}} + \Lambda_{\overline{N}} \cdot \varepsilon^*(N_1, \dots, N_D), \quad (12)$$

where the Lebesgue constant $\Lambda_{\overline{N}}$ is bounded by $\Lambda_{\overline{N}} \le \frac{2^D}{\pi^D}\prod_{i=1}^D \big(\log(N_i+1)+1\big)$.

Proof. In order to derive a significant estimate, we rewrite the interpolation in Lagrangian form, i.e.

$$I_{\overline{N}}(f)(p) = \sum_{k_1=0}^{N_1} \dots \sum_{k_d=0}^{N_D} f(p^{(k_1,\dots,k_D)})\lambda^{(k_1,\dots,k_D)}(p) \quad (13)$$

with the appropriate polynomials $\lambda^{(k_1,\dots,k_D)}$. From Proposition 2.1 we deduce

$$E\left(\max_{p\in\mathcal{P}} \left|Price^p - I_{\overline{N}}(Price_{\varepsilon}^{(\cdot)})(p)\right|\right) \le c + E\left(\max_{p\in\mathcal{P}} \sum_{k_1=0}^{N_1} \dots \sum_{k_d=0}^{N_D} |\varepsilon^{p^{(k_1,\dots,k_D)}}||\lambda^{(k_1,\dots,k_D)}(p)|\right)$$

with constant c as defined in the assertion of the theorem. Invoking the well-known estimate for the Lebesgue constant, which is defined by $\Lambda_{\overline{N}} := \max_{p\in\mathcal{P}} \sum_{k_d=0}^{N_D} \dots \sum_{k_d=0}^{N_D} \left|\lambda^{(k_1,\dots,k_D)}(p)\right|$, immediately yields the result for case (i).

Assuming (ii), for arbitrary $t > 0$ we estimate using Jensen's inequality and the symmetry of the centered normal distribution

$$\mathrm{e}^{tE(\max_{j\in J}|p^j|)} < E\,\mathrm{e}^{t\max_{j\in J}|p^j|} = E\max_{j\in J}\mathrm{e}^{t|p^j|} < \sum_{j\in J} E\big(\mathrm{e}^{t|p^j|}\big) < 2\sum_{j\in J} E\big(\mathrm{e}^{tp^j}\big). \quad (14)$$

Inserting the characteristic function of the normal distribution and setting $n := \prod_{i=1}^D (N_i+1)$ and $\overline{\sigma} := \max_{j\in J}\sigma_{j,M}$ we obtain $E(\max_{j\in J}|p^j|) \le \log(2n)/t + \overline{\sigma}^2 t/2$. Finally, minimizing over $t > 0$ yields

$$E\big(\max_{j\in J}|p^j|\big) < \sqrt{2\log(2n)}\overline{\sigma}.$$

□ □

2.4 Relation to parametric multilevel Monte-Carlo

There is an interesting relation between the Chebyshev interpolation approach that we proposed in this section and the parametric multilevel Monte-Carlo method suggested by S. Heinrich in [24] and [26]. To be more precise, as concisely explained in Section 2.1 in [25], the starting point of [24] is the interpolation of the function

$$p^1 \mapsto E\big(f^{p^1}(X)\big) \tag{15}$$

and the computation of $E[f^{p_k}(X)]$ at the nodes p_k with Monte-Carlo. Note that in this setting, the random variable X is not parametric. Next, he introduces the multilevel Monte-Carlo method. This is a hierarchical procedure based on nested grids. In each step, the estimator of the coarser grid serves as a control variate. The grids then are chosen optimally to balance cost and accuracy. [26] shows that the resulting algorithm is optimal for a certain class of problems. This class of problems is characterized by the regularity of the function $(p^1, x) \mapsto f^{p^1}(x)$, namely that it belongs to a Sobolev space of a certain order r. The order r for which partial derivatives in (p^1, x) are assumed is the determining factor for the efficiency. In particular, the weak partial derivatives in both the parameters $p^1 \in \mathbb{R}^D$ and in $x \in \mathbb{R}^d$ need to exist in order to apply the approach of [26].

In contrast, our error analysis is based on the regularity of the mapping

$$(p^1, p^2) \mapsto E\big(f^{p^1}(X^{p^2})\big). \tag{16}$$

The resulting problem class is significantly larger than the setting of [24] and [26]. This is essential for applications in finance, as the examples of a European call and digital option prove: The payoff function of a European call has a kink. According to the ansatz of [24] this yields a very poor convergence rate. The call option prices as a function of the parameters, however, are in many cases even analytic as we will prove in the following Section 3. The situation is even more severe for digital options, whose payoffs are not even weakly differentiable. Here, the approach of [24] does not even yield convergence. The Chebyshev method as proposed in this article, however, can be shown to converge with an exponential rate for a wide range of applications. We again refer to Section 3.

We relate the error analysis presented in Section 2.3 with the results of [26]. [26] presents a bound of the expected error in the L^2-norm as a function of the cost to obtain an asymptotic analysis of the efficiency. In [25], it is shown that there exist constants c_1 and c_2 such that for each integer M the cost of the parametric Monte-Carlo method is bounded by $c_1 M$ and the error is bounded by $c_2 M^{-\alpha}$, where α depends on the regularity of the function f and $\alpha \in (0, \frac{1}{2})$. The index α depends on the dimension of the parameter space and the Sobolev order of the function space to which f belongs.

To present an error analysis in the same spirit, we observe that the cost for estimating $\mathit{Price}^{(p)}$ for a fixed p is bounded by $\bar{c}_1 M$ for M Monte-Carlo simulations with a constant $\bar{c}_1 > 0$. It follows directly that the cost of deriving the interpolation $I_{\overline{N}}(Price_\varepsilon^{(\cdot)})$ is bounded by $\bar{c}_1\mathcal{M} = \bar{c}_1 \prod_{i=1}^{D}(N_i+1)M$, where N_i is the number of nodal points in dimension i and M is the number of sample paths at each nodal point. In order to estimate the error, according by the central limit theorem it is reasonable to assume $\sigma_{j,M} = \sigma_j/\sqrt{M}$ for large M and σ_M from case (ii) in Theorem 2.5.

Departing from the framework of [25], we estimate the expectation of the $L^\infty-$norm of the error, instead of the weaker L^2-norm. The maximum norm is more suitable for quantifying mispricing and it is available without additional cost, since the Chebyshev interpolation is tailored to minimize the maximum error.

Theorem 2.6. *Let the assumptions of Theorem 2.5 with condition (ii) on the distorsions hold and $\sigma_{j,M} = \sigma_j/\sqrt{M}$. For each $\beta > D/2$ there exist constants $\bar{c}_1, \bar{c}_2 > 0$ such that for each integer $\mathcal{M} > 1$ there is a choice of M, N such that the offline cost of the Chebyshev method is bounded by $\bar{c}_1\mathcal{M}$ and*

$$E\left(\max_{p\in\mathcal{P}} \left|Price^p - I_{\overline{N}}(Price_\varepsilon^{(\cdot)})(p)\right|\right) \le \bar{c}_2 \log(\mathcal{M})^\beta \mathcal{M}^{-1/2}.$$

Proof. Combining (12) and Corollary 6 results in

$$err := E\left(\max_{p\in\mathcal{P}} \left|Price^p - I_{\overline{N}}(Price_\varepsilon^{(\cdot)})(p)\right|\right) \le C_1 \underline{\varrho}^{-\underline{N}} + \Lambda_{\overline{N}} \cdot \varepsilon^*(N_1, \dots, N_D), \tag{17}$$

where $\underline{\varrho} = \min_{1\le i\le D} \varrho_i$ and $\underline{N} = \min_{1\le i\le D} N_i$. Setting $N_i(M) = N(M) = \alpha \log(M)$ for some $\alpha > 1/(2\log(\underline{\varrho}))$ and $\mathcal{M} = cN(M)^D M$ the result follows from elementary estimations. □ □

Remark 2.7. *In Theorem 2.6, the error of the resulting Chebyshev interpolation is put in relation to the cost of the offline phase. This is in the spirit of [25]. The following two observations show that our approach is in this regard competitive:*

(i) In contrast to [25, Theorem 1], the payoff function $(p^1, x) \mapsto f^{p^1}(x)$ is not *required to be weakly differentiable to a specific order. Moreover, Theorem 2.6 allows a parametrized random variable X^{p^2}.*

(ii) The error of the multilevel Monte-Carlo estimate of [25, Theorem 1] decays with $\sqrt{\mathcal{M}}$, if the function f is of high regularity. This is the only case in which the asymptotic order of convergence in [25, Theorem 1] is slightly better than the rate of Theorem 2.6, where the logarithmic term appears additionally. Note that the error in Theorem 2.6 is measured in the stronger $L^\infty-$norm, however.

We emphasize that the analysis according to [25] considers efficiency in terms of accuracy versus the *cost of the offline phase* and ignores the online phase. From an application point of view, however, the *cost of the online phase* is crucial. This is especially the case where real-time evaluation is required. In some applications, it is even rather the offline cost that can be disregarded. This is for instance the case if the offline phase can be executed in idle times.

To make the implications clear, let us consider a concrete example. Following the reasoning of efficiency as accuracy versus "offline cost", the number of nodal points of the interpolation is of minor importance. So is the choice of the interpolation method. This is in line with [26] choosing piecewise linear interpolation to illustrate the multilevel Monte-Carlo method, that they originally described for an arbitrary nodal interpolation. Whereas this choice of interpolation can be appropriate for a one-dimensional parameter space, a simple calculation makes clear how crucial it becomes for multivariate parameter spaces to require as few nodal points as possible to achieve a pre-specified accuracy. For instance, when interpolating piecewise linearly on an equidistant grid in the multilevel Monte-Carlo method of [25] with L levels, 2^L nodal points in each direction are applied. For a D-dimensional parameter space, this results in 2^{LD} nodal points. For $L = 10$ and $D = 2$, this results in more than 1 million nodal points. In this case, the "online cost" is in the range of the cost of a Monte-Carlo simulation, which makes the interpolation redundant. Applying Chebyshev polynomial interpolation, a small number of nodal points such as 7, as shown in Section 4.2, suffices for the Chebyshev interpolation method to obtain an appropriate accuracy. In this case, the total number of nodes is 49 for the tensorized Chebyshev interpolation in two dimensions. Thus, the online cost outperforms Monte-Carlo significantly.

This highlights the fact that the choice of the interpolation method is crucial. Let us now turn to an algorithmic structure to balance accuracy and online cost in the Chebshev interpolation approach.

Remark 2.8. *The online cost is proportional to the number of nodal points. If the highest priority is given to the efficiency of the online phase, one can proceed as follows to achieve a pre-specified accuracy ε: First, choose the number of nodal points such that the first summand of the error bound in* (12) *is smaller than $\varepsilon/2$. Then, choose the number of samples M of the selected Monte-Carlo technique such that also the second summand of the error bound in* (12) *is smaller than $\varepsilon/2$.*

3 Exponential Convergence of Chebyshev Interpolation for POP

In this section, we embed the multivariate Chebyshev interpolation into the option pricing framework. We provide sufficient conditions under which option prices depend analytically on their parameters. We keep the option pricing framework sufficiently abstract so that it also comprises various different applications such as the computation of risk quantities on basis of parametric random variables.

Let us first observe that an analysis along the lines of [24] and [26] would start from the regularity of the function $p^1 \mapsto f^{p^1}(X)$. Yet, we have to recognize that basic functions such as the payoff of a plain vanilla option, $K \mapsto (S_T \pm K)^+$, a digital option, $B \mapsto \mathbb{1}_{(-\infty,B]}(S)$, $B \mapsto \mathbb{1}_{[B,\infty)}(S)$ (the latter also underlies the computation of the quantile function and thus the Var for instance) are not differentiable respectively not even continuous. We therefore conclude that the approach of [24] and [26] is too restrictive for financial applications.

Invoking that–although the payoff of a plain vanilla option has a kink–its price function is smooth in virtually all models, we exploit the smoothing property of the distribution in order to derive analyticity of the price as function of its parameters. For linear problems, this can be conveniently studied in terms of Fourier transforms. First, Fourier representations of option prices are explicitly available for a large class of both option types and asset models. Second, Fourier transformation unveils the analytic properties of both the payoff structure and the distribution of the underlying stochastic quantity in a beautiful way. The Fourier transform of the damped call or digital payoff function for instance evidently is analytic in the strike.

To introduce a general option pricing framework we consider option prices of the form

$$Price^{p=(p^1,p^2)} = E\big(f^{p^1}(X^{p^2})\big) \tag{18}$$

where f^{p^1} is a parametrized family of measurable payoff functions $f^{p^1} : \mathbb{R}^d \to \mathbb{R}_+$ with payoff parameters $p^1 \in \mathcal{P}^1$ and X^{p^2} is a family of $\mathbb{R}^d$-valued random variables with model parameters $p^2 \in \mathcal{P}^2$. The parameter set

$$p = (p^1, p^2) \in \mathcal{P} = \mathcal{P}^1 \times \mathcal{P}^2 \subset \mathbb{R}^D \tag{19}$$

is again of hyperrectangular structure, i.e. $\mathcal{P}^1 = [\underline{p}_1, \overline{p}_1] \times \ldots \times [\underline{p}_m, \overline{p}_m]$ and $\mathcal{P}^2 = [\underline{p}_{m+1}, \overline{p}_{m+1}] \times \ldots \times [\underline{p}_D, \overline{p}_D]$ for some $1 \leq m \leq D$ and real $\underline{p}_i \leq \overline{p}_i$ for all $i = 1, \ldots, D$.

Typically, we are given a parametrized $\mathbb{R}^d$-valued driving stochastic process $H^{p'}$ where $S^{p'}$ is the vector of asset price processes modelled as an exponential of $H^{p'}$, i.e.

$$S_t^{p',i} = S_0^{p',i} \exp(H_t^{p',i}), \qquad 0 \leq t \leq T, \quad 1 \leq i \leq d, \tag{20}$$

and X^{p^2} is an $\mathcal{F}_T$-measurable $\mathbb{R}^d$-valued random variable, possibly depending on the history of the d driving processes, i.e. $p^2 = (T, p')$ and

$$X^{p^2} := \Psi\big(H_t^{p'}, 0 \leq t \leq T\big),$$

where Ψ is an $\mathbb{R}^d$-valued measurable functional.

We now focus on the case where the price (18) is given in terms of Fourier transforms. This enables us to provide sufficient conditions under which the parametrized prices have an analytic extension to an appropriate generalized Bernstein ellipse. For most relevant options, the payoff profile f^{p^1} is not integrable and its Fourier transform over the real axis is not well-defined. Instead, there exists an exponential damping factor $\eta \in \mathbb{R}^d$ such that $\mathrm{e}^{\langle \eta, \cdot \rangle} f^{p^1} \in L^1(\mathbb{R}^d)$. We therefore introduce exponential weights in our set of conditions and denote the Fourier transform of $g \in L^1(\mathbb{R}^d)$ by

$$\hat{g}(z) := \int_{\mathbb{R}^d} \mathrm{e}^{i\langle z, x \rangle} g(x) \,\mathrm{d}\, x$$

and we denote the Fourier transform of $\mathrm{e}^{\langle \eta, \cdot \rangle} f \in L^1(\mathbb{R}^d)$ by $\hat{f}(\cdot - i\eta)$. The exponential weight of the payoff will be compensated by exponentially weighting the distribution of X^{p^2} and that weight will reappear in the argument of φ^{p^2}, the characteristic function of X^{p^2}.

Conditions 3.1. Let $\mathcal{P} = \mathcal{P}^1 \times \mathcal{P}^2 \subset \mathbb{R}^D$ be a parameter set with hyperrectangular structure as in (19). Let $\varrho \in (1, \infty)^D$ and denote $\varrho^1 := (\varrho_1, \dots, \varrho_m)$ and $\varrho^2 := (\varrho_{m+1}, \dots, \varrho_D)$, and consider a weight $\eta \in \mathbb{R}^d$.

(A1) For every $p^1 \in \mathcal{P}^1$ the mapping $x \mapsto \mathrm{e}^{\langle \eta, x \rangle} f^{p^1}(x)$ is in $L^1(\mathbb{R}^d)$.

(A2) For every $z \in \mathbb{R}^d$ the mapping $p^1 \mapsto \widehat{f^{p^1}}(z - i\eta)$ is analytic in the generalized Bernstein ellipse $B(\mathcal{P}^1, \varrho^1)$ and there are constants $c_1, c_2 > 0$ such that $\sup_{p^1 \in B(\mathcal{P}^1, \varrho^1)} |\widehat{f^{p^1}}(-z - i\eta)| \leq c_1 e^{c_2 |z|}$ for all $z \in \mathbb{R}^d$.

(A3) For every $p^2 \in \mathcal{P}^2$ the exponential moment condition $E\big(\mathrm{e}^{-\langle \eta, X^{p^2} \rangle}\big) < \infty$ holds.

(A4) For every $z \in \mathbb{R}^d$ the mapping $p^2 \mapsto \varphi^{p^2}(z + i\eta)$ is analytic in the generalized Bernstein ellipse $B(\mathcal{P}^2, \varrho^2)$ and there are constants $\alpha \in (1, 2]$ and $c_1, c_2 > 0$ such that $\sup_{p^2 \in B(\mathcal{P}^2, \varrho^2)} |\varphi^{p^2}(z + i\eta)| \leq c_1 \,\mathrm{e}^{-c_2 |z|^\alpha}$ for all $z \in \mathbb{R}^d$.

Conditions (A1)–(A4) are satisfied for a large class of payoff functions and asset models, see [16]

Theorem 3.2. *Let $\varrho \in (1, \infty)^D$ and consider a weight $\eta \in \mathbb{R}^d$. Under conditions (A1)–(A4), $\mathcal{P} \ni p \mapsto Price^p$ has an analytic extension to the generalized Bernstein ellipse $B(\mathcal{P}, \varrho)$ and $\sup_{p \in B(\mathcal{P}, \varrho)} |Price^p| \leq V$, and therefore,*

$$\max_{p \in \mathcal{P}} \big| Price^p - I_{\overline{N}}(Price^{(\cdot)})(p) \big| \leq 2^{\frac{D}{2}+1} \cdot V \cdot \left(\sum_{i=1}^{D} \varrho_i^{-2N_i} \prod_{j=1}^{D} \frac{1}{1 - \varrho_j^{-2}} \right)^{\frac{1}{2}}.$$

Proof. Thanks to assumptions (A2) and (A4), the mapping $z \mapsto \widehat{f^{p^1}}(-z-i\eta)\varphi^{p^2}(z+i\eta)$ belongs to $L^1(\mathbb{R}^d)$ for every $p=(p^1,p^2)\in\mathcal{P}$. Together with conditions (A1) and (A3), we therefore can apply [11, Theorem 3.2]. This gives the following Fourier representation of the option prices,

$$Price^p = \frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d+i\eta} \widehat{f^{p^1}}(-z)\varphi^{p^2}(z)\,\mathrm{d}\,z.$$

By assumptions (A2) and (A4), the mapping

$$p=(p^1,p^2)\mapsto \widehat{f^{p^1}}(-z)\varphi^{p^2}(z)$$

has an analytic extension to $B(\mathcal{P},\varrho)$.

Let γ be the contour of a compact triangle in the interior of $B([\underline{p}_i,\overline{p}_i],\varrho_i)$ for arbitrary $i=1,\dots,D$. Then, by assumptions (A2) and (A4) we may apply Fubini's theorem to obtain

$$\begin{aligned}\int\limits_\gamma Price^{(p_1,\dots,p_D)}(z)\,\mathrm{d}\,p_i &= \frac{1}{(2\pi)^d}\int\limits_\gamma\int\limits_{\mathbb{R}^d+i\eta}\widehat{f^{p^1}}(-z)\varphi^{p^2}(z)\,\mathrm{d}\,z\,\mathrm{d}\,p_i\\ &= \frac{1}{(2\pi)^d}\int\limits_{\mathbb{R}^d+i\eta}\int\limits_\gamma\widehat{f^{p^1}}(-z)\varphi^{p^2}(z)\,\mathrm{d}\,p_i\,\mathrm{d}\,z = 0.\end{aligned}$$

Moreover, thanks to assumptions (A2) and (A4), dominated convergence shows continuity of $p\mapsto Price^p$ in $B(\mathcal{P},\varrho)$ which yields the analyticity of $p\mapsto Price^p$ in $B(\mathcal{P},\varrho)$ thanks to a version of Morera's theorem provided in [28, Satz 8]. □ □

Similar to Proposition 2.3, if Conditions 3.1 are satisfied, the Chebyshev interpolation also allows the corresponding derivatives to be well approximated. One very interesting application of this result in finance is the computation of sensitivities like delta or vega of an option price for risk assessment purposes. Theorem 3.2 together with Proposition 2.3 yield the following corollary.

Corollary 3.3. *Set $N_i=N$, $i=1,\dots,D$, i.e. the same number of nodal points in each dimension. Under Conditions 3.1, $\mathcal{P}\ni p\mapsto Price^p\in W_2^{\sigma,\omega}(\mathcal{P})$ for all $\sigma\in\mathbb{N}$, and therefore for all $l\in\mathbb{N}$, μ and σ with $\sigma>\frac{D}{2}$, $0\le\mu\le\sigma$ and $\mu-l>\frac{D}{2}$ there exists a constant C such that*

$$\|Price^p - I_{\overline{N}}(Price^{(\cdot)})(p)\|_{C^l(\mathcal{P})}\le CN^{2\mu-\sigma}\|Price^p\|_{W_2^\sigma(\mathcal{P})}.$$

4 Numerical Experiments

In this section, we use the Chebyshev method to price basket and path-dependent options. First, we apply the method to interpolate Monte-Carlo estimates of prices of financial products and check the resulting accuracy. To this end, we choose example basket, barrier and lookback options in 5-dimensional Black&Scholes, Heston and Merton models. Second, we combine the Chebyshev method with a Crank-Nicolson finite difference solver using the Brennan Schwartz approximation, see [3], in order to price a univariate American put option in the Black&Scholes model.

In our Monte-Carlo simulation, we use 10^6 sample paths, antithetic variates as a variance reduction technique, and 400 time steps per year. The error of the Monte-Carlo method cannot be computed directly. We thus turn to statistical error analysis and use 95% confidence bounds to determine the accuracy. These bounds are derived from the assumption of a normally distributed Monte-Carlo estimator with mean equal to the estimator's value and variance equal to the empirical variance of the payoff on the Monte-Carlo samples. The confidence bounds then yield a range around the mean that includes the true price with 95% probability. We pick two free parameters $p_{i_1},\ p_{i_2}$ out of (19), $1 \leq i_1 < i_2 \leq D$, in each model setup and fix all other parameters at reasonable constant values. In this section, we define the discrete parameter grid $\overline{\mathcal{P}} \subseteq [\underline{p}_{i_1}, \overline{p}_{i_1}] \times [\underline{p}_{i_2}, \overline{p}_{i_2}]$ by

$$\begin{aligned} \overline{\mathcal{P}} &= \left\{ \left(p_{i_1}^{k_{i_1}}, p_{i_2}^{k_{i_2}} \right),\ k_{i_1}, k_{i_2} \in \{0, \dots, 40\} \right\}, \\ p_{i_j}^{k_{i_j}} &= \underline{p}_{i_j} + \frac{k_{i_j}}{40} \left(\overline{p}_{i_j} - \underline{p}_{i_j} \right),\ k_{i_j} \in \{0, \dots, 40\},\ j \in \{1, 2\}, \end{aligned} \tag{21}$$

and call $\overline{\mathcal{P}}$ the test grid. On this test grid, the largest confidence bound is 0.025, and is less than 0.013 on average. For the finite difference method, we find that the absolute error between the numerical approximation and the option price is below 0.005 on all computed parameter tuples in $\overline{\mathcal{P}}$. This error bound was computed by comparing each approximation to the limit of the sequence of finite difference approximations as the grid size is increased. In our calculations, we work with a grid size in time as well as in space (log-moneyness) of $50 \cdot \max\{1, T\}$ and compare the result to the prices obtained using grid sizes of $1000 \cdot \max\{1, T\}$. This grid size was determined to be sufficient for approximating the limit, since it was observed that a grid size of $500 \cdot \max\{1, T\}$ produces nearly identical prices.

Here, our main concern is the accuracy of the Chebyshev interpolation as we vary the strike and maturity parameters of each option analogously to the previous section. For $N \in \{5, 10, 30\}$, we precompute the Chebyshev coefficients as defined in (4) with $D = 2$ while always keeping $N_1 = N_2 = N$. An overview of the fixed and free parameters in our model selection is given in Table 1. For computational simplicity in the Monte-Carlo simulation, we assume that the underlyings are uncorrelated.

Let us briefly define the payoffs of the multivariate basket and path-dependent

Model	**fixed parameters**		**free parameters**	
	p^1	p^2	p^1	p^2
BS	$S_0^j = 100$, $r = 0.005$	$\sigma_j = 0.2$	$K \in [83.33,\ 125]$	$T \in [0.5, 2]$
Heston	$S_0^j = 100$, $r = 0.005$	$\kappa_j = 2$, $\theta_j = 0.2^2$, $\sigma_j = 0.3$, $\rho_j = -0.5$, $v_{j,0} = 0.2^2$	$K \in [83.33,\ 125]$	$T \in [0.5, 2]$
Merton	$S_0^j = 100$, $r = 0.005$	$\sigma_j = 0.2$, $\alpha_j = -0.1$, $\beta_j = 0.45$, $\lambda_j = 0.1$	$K \in [83.33,\ 125]$	$T \in [0.5, 2]$

Table 1: Parametrization of models, basket and path-dependent options. The model parameters are given for $j = 1, \ldots, d$ to reflect the multivariate setting with free parameters given by the strike K and the maturity T. Note that, in contrast to the two-dimensional Heston model described in Section **??**, in the numerical experiments here we use a multivariate Heston model in which the volatility of each underlying is driven by its own volatility process.

options. The payoff profile of a basket option for d underlyings is given as

$$f^K\big(S_T^1, \ldots, S_T^d\big) = \left(\left(\frac{1}{d}\sum_{j=1}^{d} S_T^j\right) - K\right)^+ .$$

We write $S_t = (S_t^1, \ldots, S_t^d)$, $\underline{S}_T^j := \min_{0\leq t\leq T} S_t^j$ and $\overline{S}_T^j := \max_{0\leq t\leq T} S_t^j$. A lookback option for d underlyings is defined as

$$f^K\left(\overline{S}_T^1, \ldots, \overline{S}_T^d\right) = \left(\left(\frac{1}{d}\sum_{j=1}^{d} \overline{S}_T^j\right) - K\right)^+ .$$

As an example of a multivariate barrier option on d underlyings, we define the payoff

$$f^K\big(\{S(t)\}_{0\leq t\leq T}\big) = \left(\left(\frac{1}{d}\sum_{j=1}^{d} S_T^j\right) - K\right)^+ \cdot \mathbb{1}_{\{\underline{S}_T^j \geq 80,\ j=1,\ldots,d\}}.$$

The payoff of an American put option is the same as that of a European put,

$$f^K\big(S_t\big) = (K - S_t)^+ ,$$

but the option holder has the right to exercise the option at any time t up to maturity T.

4.1 Accuracy study

We now turn to the results of our numerical experiments. In order to evaluate the accuracy of the Chebyshev interpolation. We investigate the worst-case error ε_{L^∞}. The absolute error of the Chebyshev interpolation method can be directly computed by comparing the interpolated option prices with those obtained by the reference numerical algorithm i.e. either the Monte-Carlo or the Finite Difference method. Since the Chebyshev interpolation matches the reference method on the Chebyshev nodes, we will use the out-of-sample test grid as in (21). Table 2 shows the numerical results for the basket and path-dependent options for $N = 5$, Table 3 shows $N = 10$, and Table 4 shows $N = 30$. In addition to the L^∞ errors, the tables display the Monte-Carlo (MC) prices, the Monte-Carlo confidence bounds, and the Chebyshev Interpolation (CI) prices for the parameters at which the L^∞ error is realized.

Model	**Option**	ε_{L^∞}	**MC price**	**MC conf. bound**	**CI price**
BS	Basket	$1.338 \cdot 10^{-1}$	8.6073	$1.171 \cdot 10^{-2}$	8.4735
Heston	Basket	$9.238 \cdot 10^{-2}$	0.0009	$1.036 \cdot 10^{-4}$	0.0933
Merton	Basket	$9.815 \cdot 10^{-2}$	8.8491	$1.552 \cdot 10^{-2}$	8.7510
BS	Lookback	$2.409 \cdot 10^{-1}$	9.4623	$9.861 \cdot 10^{-3}$	9.2213
Heston	Lookback	$5.134 \cdot 10^{-1}$	0.0314	$6.472 \cdot 10^{-4}$	-0.4820
Merton	Lookback	$2.074 \cdot 10^{-1}$	1.0919	$9.568 \cdot 10^{-3}$	0.8844
BS	Barrier	$1.299 \cdot 10^{-1}$	1.0587	$5.092 \cdot 10^{-3}$	1.1887
Heston	Barrier	$1.073 \cdot 10^{-1}$	2.7670	$9.137 \cdot 10^{-3}$	2.6597
Merton	Barrier	$9.916 \cdot 10^{-2}$	1.3810	$1.102 \cdot 10^{-2}$	1.4802

Table 2: Interpolation of exotic options with Chebyshev interpolation. $N = 5$ and $d = 5$ in all cases. In addition to the L^∞ errors, the table displays the Monte-Carlo (MC) prices, the Monte-Carlo confidence bounds, and the Chebyshev Interpolation (CI) prices for the parameters at which the L^∞ error is realized.

The results show that for $N = 30$ the accuracy of all selected options is 10^{-3}. We see that the Chebyshev interpolation error is dominated by the Monte-Carlo confidence bounds to the extent that the interpolation error becomes negligible when comparing the two. For basket and barrier options, the L^∞ error already reaches satisfying levels of order 10^{-3} at $N = 10$. Again, the Chebyshev approximation falls within the confidence bounds of the Monte-Carlo approximation. Thus, Chebyshev interpolation with only $121 = (10+1)^2$ nodes suffices to mimic the Monte-Carlo pricing results. This statement does not hold for lookback options, where the L^∞ error still differs noticeably when comparing $N = 10$ to $N = 30$. As can be seen from Table 2, Chebyshev interpolation with $N = 5$ may yield unreliable pricing results. For lookback options in the Heston model, we even observe negative prices in individual cases.

We conclude that the Chebyshev interpolation is highly promising for the evaluation of multivariate basket and path-dependent options. However, the

Model	**Option**	ε_{L^∞}	**MC price**	**MC conf. bound**	**CI price**
BS	Basket	$2.368 \cdot 10^{-3}$	2.4543	$7.493 \cdot 10^{-3}$	2.4566
Heston	Basket	$2.134 \cdot 10^{-3}$	3.1946	$1.073 \cdot 10^{-2}$	3.1925
Merton	Basket	$3.521 \cdot 10^{-3}$	6.1929	$2.231 \cdot 10^{-2}$	6.1894
BS	Lookback	$2.861 \cdot 10^{-2}$	0.9827	$4.197 \cdot 10^{-3}$	0.9541
Heston	Lookback	$1.098 \cdot 10^{-1}$	2.0559	$4.826 \cdot 10^{-3}$	2.1656
Merton	Lookback	$3.221 \cdot 10^{-2}$	4.7072	$1.264 \cdot 10^{-2}$	4.7394
BS	Barrier	$4.414 \cdot 10^{-3}$	5.3173	$1.725 \cdot 10^{-2}$	5.3129
Heston	Barrier	$5.393 \cdot 10^{-3}$	0.7158	$5.879 \cdot 10^{-3}$	0.7212
Merton	Barrier	$3.376 \cdot 10^{-3}$	9.2688	$2.302 \cdot 10^{-2}$	9.2722

Table 3: Interpolation of exotic options with Chebyshev interpolation. $N = 10$ and $d = 5$ in all cases. In addition to the L^∞ errors, the table displays the Monte-Carlo (MC) prices, the Monte-Carlo confidence bounds, and the Chebyshev Interpolation (CI) prices for the parameters at which the L^∞ error is realized.

Model	**Option**	ε_{L^∞}	**MC price**	**MC conf. bound**	**CI price**
BS	Basket	$1.452 \cdot 10^{-3}$	5.1149	$1.200 \cdot 10^{-2}$	5.1163
Heston	Basket	$1.047 \cdot 10^{-3}$	7.6555	$1.371 \cdot 10^{-2}$	7.6545
Merton	Basket	$3.765 \cdot 10^{-3}$	7.2449	$2.359 \cdot 10^{-2}$	7.2412
BS	Lookback	$3.766 \cdot 10^{-3}$	25.9007	$1.032 \cdot 10^{-2}$	25.9045
Heston	Lookback	$1.914 \cdot 10^{-3}$	16.4972	$9.754 \cdot 10^{-3}$	16.4991
Merton	Lookback	$3.646 \cdot 10^{-3}$	27.1018	$1.623 \cdot 10^{-2}$	27.1054
BS	Barrier	$5.331 \cdot 10^{-3}$	5.6029	$1.730 \cdot 10^{-2}$	5.6082
Heston	Barrier	$2.486 \cdot 10^{-3}$	3.6997	$1.353 \cdot 10^{-2}$	3.6972
Merton	Barrier	$4.298 \cdot 10^{-3}$	6.6358	$2.309 \cdot 10^{-2}$	6.6315

Table 4: Interpolation of exotic options with Chebyshev interpolation. $N = 30$ and $d = 5$ in all cases. In addition to the L^∞ errors, the table displays the Monte-Carlo (MC) prices, the Monte-Carlo confidence bounds, and the Chebyshev Interpolation (CI) prices for the parameters at which the L^∞ error is realized.

accuracy of the interpolation critically depends on the accuracy of the reference method at the nodal points, which motivates further analysis that we perform in the subsequent subsection.

4.2 Study of the gain in efficiency

We compute the results on a standard PC with an Intel i5 CPU, 2.50 GHz with cache size 3 MB. In Section 4.2, we used a PC with Intel Xeon CPU with 3.10 GHz with 20 MB SmartCache. All codes are written in Matlab R2014a. In this section, we choose a multivariate lookback option in the Heston model, based on 5 underlyings, as an example. For the efficiency study, we first vary

Varying	ε_{L^∞}	MC price	MC conf. bound	CI price
σ, ρ	$5.260 \cdot 10^{-2}$	5.239	$1.428 \cdot 10^{-2}$	5.292

Table 5: Interpolation of multivariate lookback options with Chebyshev interpolation for $N = 6$ based on an enriched Monte-Carlo setting with $5 \cdot 10^6$ sample paths, antithetic variates, and 400 time steps per year. In addition to the L^∞ error on the test grid, we also report the Monte-Carlo (MC) price, the Monte-Carlo confidence bound, and the Chebyshev Interpolation (CI) price for the parameters at which the L^∞ error is realized. We observe that the accuracy of the Chebyshev interpolation $N = 6$ is roughly in the same range as the accuracy of the benchmark Monte-Carlo setting (worst-case confidence bound of $6.783 \cdot 10^{-2}$ and worst-case error of $2.791 \cdot 10^{-2}$).

one parameter, then we vary two.

Variation of two model parameters
We choose $\rho_j = \rho$, $j = 1, \ldots, 5$, and vary

$$\begin{aligned} &\rho \in [\rho_{\min},\ \rho_{\max}], && \rho_{\min} = -1, \quad \rho_{\max} = 1, \\ &\sigma \in [\sigma_{\min},\ \sigma_{\max}], && \sigma_{\min} = 0.1, \quad \sigma_{\max} = 0.5, \end{aligned} \tag{22}$$

fixing all other parameters to the values of setting (**??**). In order to guarantee a roughly comparable accuracy between the Chebyshev interpolation method and the benchmark Monte-Carlo pricing, we use the following test grid $\overline{\mathcal{P}} \subseteq [\sigma_{\min}, \sigma_{\max}] \times [\rho_{\min}, \rho_{\max}]$,

$$\begin{aligned} \overline{\mathcal{P}} &= \left\{ \left(\sigma^{k_1}, \rho^{k_2}\right),\ k_1, k_2 \in \{0, \ldots, 20\} \right\}, \\ \sigma^{k_1} &= \sigma_{\min} + \frac{k_1}{20} \left(\sigma_{\max} - \sigma_{\min}\right),\ k_1 \in \{0, \ldots, 20\}, \\ \rho^{k_2} &= \rho_{\min} + \frac{k_2}{20} \left(\rho_{\max} - \rho_{\min}\right),\ k_2 \in \{0, \ldots, 20\}. \end{aligned}$$

In Table 5, we present the accuracy results for the Chebyshev interpolation with $N^{\text{Heston}}_{\text{Cheby}} = 6$ based on the enriched Monte-Carlo setting. Comparing the benchmark Monte-Carlo setting and the enriched Monte-Carlo setting on this test grid, we observe that the maximal absolute error is $2.791 \cdot 10^{-2}$ and the confidence bounds of the benchmark Monte-Carlo setting do not exceed $6.783 \cdot 10^{-2}$.

To compare the run-times, we show the run-times necessary to compute the prices for M^2 parameter tuples for different values of M. Again, the run-times are measured for $M = 1$ and extrapolated for other values of M. Table 6 presents the results. In Figure 1, for each $M = 1, \ldots, 100$, the run-times of the Chebyshev interpolation method, including the offline phase, are presented and compared to the Monte-Carlo method. We observe that for $M = 15$ both lines intersect and for $M > 15$ the Chebyshev method outperforms its benchmark. Contrary to the case where only one parameter is varied, the intersection of both lines occurs at a significantly lower value of M due to the fact that for each M pricing must be performed for M^2 parameter tupels.

	Heston			
M	1	10	50	100
$T^{\text{Cheby}}_{\text{online}}$ (s)	$7.1{\cdot}10^{-4}$	$7.1{\cdot}10^{-2}$	1.8	7.1
$T^{\text{Cheby}}_{\text{offline+online}}$ (s)	$8.2{\cdot}10^{4}$	$8.2{\cdot}10^{4}$	$8.2{\cdot}10^{4}$	$8.2{\cdot}10^{4}$
$T^{\text{Monte-Carlo}}$ (s)	$3.4{\cdot}10^{2}$	$3.4{\cdot}10^{4}$	$8.4{\cdot}10^{5}$	$3.4{\cdot}10^{6}$
$\frac{T^{\text{Cheby}}_{\text{offline+online}}}{T^{\text{Monte-Carlo}}}$	24313.9%	243.1%	9.7%	2.4%

Table 6: Efficiency study for a multivariate lookback option in the Heston model based on 5 underlyings. Here, we vary two model parameters and compare the Chebyshev results to Monte-Carlo. Both methods have been set up to deliver comparable accuracies. As the number of computed prices increases, the Chebyshev algorithm increasingly profits from the initial investment of the offline phase.

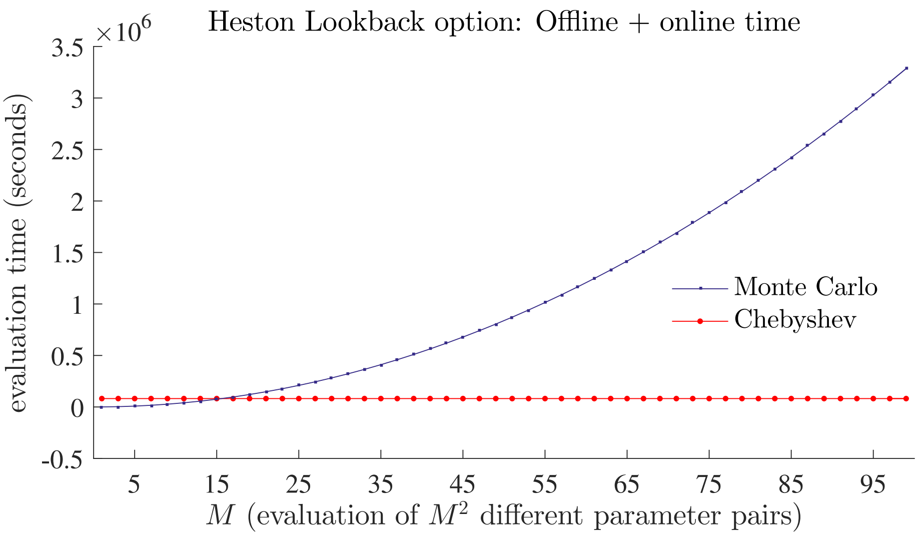

Figure 1: Effiency study for a multivariate lookback option in the Heston model based on 5 underlyings, varying the two model parameters σ and ρ. Comparison of run-times for Monte-Carlo pricing and Chebyshev pricing including the offline phase. Both methods have been set up to deliver comparable accuracies. We observe that the Monte-Carlo and the Chebyshev curves intersect at roughly $M = 15$.

Additionally, Table 6 highlights that, in the case of a total number of 50^2 parameter tuples, the Chebyshev method exhibits a significant decrease in (total) pricing run-times. For the maximal number of 100^2 parameter tuples that we investigated, pricing in either model resulted in more than 97% of run-time savings in our implementation. While computating 100^2 Heston prices using the Monte-Carlo method requires up to 39 days, the Chebyshev method computes the very same prices in 23 hours only. Note that only 7 seconds of this time span are consumed by actual pricing during the online phase.

5 Conclusion and Outlook

This article introduces the famous Chebyshev interpolation method to the problem of parametric option pricing and more generally of parametric conditional expectations. The introduction explains the advantage of tackling the complexity by of Chebyshev interpolation in this context and We analysed the resulting online-offline numerical scheme. The main convergence results are established in Sections 2, special care is taken of the error resulting from deriving the prices at the nodal point by Monte Carlo-Simulation. A comparison of the efficiency in terms of accuracy versus offline costs shows significant improvement over the existing approaches in literature. We emphasize again that moreover, this type of efficiency needs to be accomplished by efficiency as online cost versus accuracy. The "online efficiency" is more signifigant than the "offline scheme" in many situations, although typically the "offline efficiency" still matters on a lower scale. In a numerical case study, we investigated the gain in "online efficiency". The results reveal that the method has a high potential for a variety of applications and further developments.

The most urgent and challenging problems in finance are of high dimensionality. For multivariate polynomial interpolation, the introduction of sparsity techniques promises higher efficiency, for instance by using compression techniques for tensors as reviewed by [30]. The high potential of low-rank tensor methods is illustrated in a numerical example for evaluating spread options in the bivariate Black&Scholes model, which is available online, see [19]. These types of techniques have to beat the curse of dimensionality for both the online as well as the offline complexity.

Addressing further the offline complexity, we note that up to this point, we have compared the Chebyshev interpolation method with a standard Monte-Carlo technique. Since the invention of Monte-Carlo methods in the 1940s, see [36], Monte-Carlo techniques have been further developed. In particular, quasi Monte-Carlo and multilevel Monte-Carlo methods have proved to be significantly more efficient in a variety of examples in mathematical finance, [32] and [18]. Thus, by employing these techniques in the offline phase, the Chebyshev interpolation method can be enhanced. In terms of efficiency, we expect Figure 1 to change only by rescaling the time axis: The run-time for the computation of the Monte-Carlo prices on the test grid is reduced proportionally.

Obviously, the offline phase of the Chebyshev interpolation scales in the same way. As a first improvement of our implementation of the offline phase, in which, for each nodal point we produce a new independent set of samples, one can reuse a once drawn sample set to compute the prices at all nodal points. Furthermore, the run-time of the offline phase can be reduced significantly by parallelisation and computations with the help of technical devices such as graphics processing unit.

A Remark on Chebyshev polynomials

Following [9] the Chebyshev polynomials appear as an optimal choice when minimizing the error in a certain way among the nodal polynomial interpolation. Namely, let f be a function that is n times continuously differentiable on $[-1,1]$, and for which $f^{(n+1)}$ exists and is bounded on $(-1,1)$. Let p_n be a polynomial that coincides with f at the nodal points $x_0,\dots,x_n$. Then there exists $\zeta \in (-1,1)$ such that the error is given by

$$\big|f(x)-p_n(x)\big| = \left|\frac{\prod_{i=0}^{n}(x-x_i)}{(n+1)!} f^{(n+1)}(\zeta)\right| \leq \frac{\prod_{i=0}^{n}|x-x_i|}{(n+1)!} \sup_{z\in(-1,1)} \big|f^{(n+1)}(z)\big|.$$

Now minimizing the expression on the right side of the inequality yields the Chebyshev points of first kind, and the Chebyshev polynomial interpolation as the resulting minimizing polynomial. We point out that we decide to implement the Chebyshev points of first kind, which are the extrema of the Chebyshev polynomials. The advantage of this choice will become clear when generalizing the method presented in this article to the case of piecewise polynomial interpolation: The Chebyshev points of second kind contain the two end points of the interval, and thus it is straightforward to concatenate interpolations on adjacent intervals.

B Proof of Proposition 2.1

The basic structure of the proof is the same as in [45, Proof of Lemma 7.3.3]. To provide a complete, understandable proof, we first show the same steps as in [45, Proof of Lemma 7.3.3] and state explicitly at which point the proof changes.

Proof. In [45, Proof of Lemma 7.3.3], the proof is given for the following error bound:

$$\max_{p\in\mathcal{P}} \big|f - I_{\overline{N}}(f)\big| \leq \sqrt{D}2^{\frac{D}{2}+1} V \varrho_{\min}^{-N}(1-\varrho_{\min}^{-2})^{-\frac{D}{2}},$$

where N is the number of interpolation points in each of the D dimensions, $\varrho_{\min} := \min_{i=1}^{D} \varrho_i$ and V is the bound of f on $B(\mathcal{P},\varrho)$ with $\mathcal{P}=[-1,1]^D$. Here, we extend [45, Proof of Lemma 7.3.3] by incorporating the different values of N_i, $i = 1, \dots, D$, as well as expressing the error bound with the different ϱ_i, $i = 1, \dots, D$.

In general, we work with a parameter space $\mathcal{P}$ of hyperrectangular structure, $\mathcal{P} = [\underline{p}_1, \overline{p}_1] \times \ldots \times [\underline{p}_D, \overline{p}_D]$. The linear transformation introduced in Section 2 gives a transformation $\tau_{\mathcal{P}} : [-1,1]^D \to \mathcal{P}$ defined by

$$\tau_{\mathcal{P}}(p) = \left(\overline{p}_i + \frac{\underline{p}_i - \overline{p}_i}{2}(1-p) \right)_{i=1}^{D}.$$

Let $p \mapsto \mathit{Price}^p$ be a function on $\mathcal{P}$. We set $\widehat{\mathit{Price}^p} = \mathit{Price}^p \circ \tau_{\mathcal{P}}(p)$. Furthermore, let $\widehat{I}_{\overline{N}}(\widehat{\mathit{Price}}^{(\cdot)})(p)$ be the Chebyshev interpolation of $\widehat{\mathit{Price}^p}$ on $[-1,1]^D$. Then it holds that

$$I_{\overline{N}}(\mathit{Price}^{(\cdot)})(p) = \hat{I}_{\overline{N}}(\widehat{\mathit{Price}}^{(\cdot)})(\cdot) \circ \tau_{\mathcal{P}}^{-1}(p).$$

From this, it directly follows that

$$\mathit{Price}^p - I_{\overline{N}}(\mathit{Price}^{(\cdot)})(p) = \left(\widehat{\mathit{Price}} - \widehat{I}_{\overline{N}}(\widehat{\mathit{Price}}^{(\cdot)})(\cdot) \right) \circ \tau_{\mathcal{P}}^{-1}(p).$$

Applying the error estimate from [45, Lemma 7.3.3] results in

$$\begin{aligned} \left|\mathit{Price} - I_{\overline{N}}(\mathit{Price}^{(\cdot)})(\cdot)\right|_{C^0(\mathcal{P})} &= \left|\mathit{Price} - I_{\overline{N}}(\mathit{Price}^{(\cdot)})(\cdot)\right|_{C^0([-1,1]^D)} \\ &\leq \sqrt{D} 2^{\frac{D}{2}+1} \widehat{V} \varrho_{\min}^{-N} (1-\varrho_{\min}^{-2})^{-\frac{D}{2}} \\ &= \sqrt{D} 2^{\frac{D}{2}+1} V \varrho_{\min}^{-N} (1-\varrho_{\min}^{-2})^{-\frac{D}{2}}, \end{aligned}$$

where $\widehat{V} = \sup_{p \in B([-1,1]^D, \varrho)} \widehat{\mathit{Price}}^p$, $V = \sup_{p \in B(\mathcal{P}, \varrho)} \mathit{Price}^p$. Summarizing, the transformation $\tau_{\mathcal{P}} : [-1,1]^D \to \mathcal{P}$ does not affect the error analysis, simply by applying the transformation as described in Section 2,

$$B(\mathcal{P}, \varrho) := B([\underline{p}_1, \overline{p}_1], \varrho_1) \times \ldots \times B([\underline{p}_D, \overline{p}_D], \varrho_D),$$

with $B([\underline{p}, \overline{p}], \varrho) := \tau_{[\underline{p}, \overline{p}]} \circ B([-1,1], \varrho)$. Note that ϱ_i is not the radius of the ellipse $B([\underline{p}_i, \overline{p}_i], \varrho_i)$ but of the normed ellipse $B([-1,1], \varrho_i)$. Therefore, in the following it suffices to show the proof for $\mathcal{P} = [-1,1]^D$.

As in [45, Proof of Lemma 7.3.3], we introduce the scalar product

$$\langle f, g \rangle_{\varrho} := \int_{B(\mathcal{P}, \varrho)} \frac{f(z)\overline{g(z)}}{\prod_{i=1}^{D} \sqrt{|1-z_i^2|}} \, \mathrm{d}z$$

and the Hilbert space

$$L^2(B(\mathcal{P}, \varrho)) := \{ f : \ f \text{ is analytic in } B(\mathcal{P}, \varrho) \text{ and } ||f||_{\varrho}^2 := \langle f, f \rangle_{\varrho} < \infty \}.$$

Following the approach in [45, Proof of Lemma 7.3.3], we define a complete orthonormal system for $L^2(B(\mathcal{P},\varrho))$ w.r.t. the scalar product $\langle\cdot,\cdot\rangle_\varrho$ by the scaled Chebyshev polynomials

$$\tilde{T}_\mu(z) := c_\mu T_\mu(z) \text{ with } c_\mu := \left(\frac{2}{\pi}\right)^{\frac{D}{2}} \prod_{i=1}^{D} (\varrho_i^{2\mu_i} + \varrho_i^{-2\mu_i})^{-\frac{1}{2}}, \quad \text{for all } \mu \in \mathbb{N}_0^D.$$

Then, for any arbitrary bounded linear functional E on $L^2(B(\mathcal{P},\varrho))$, we have

$$|E(f)| \leq ||E||_\varrho ||f||_\varrho, \tag{23}$$

where $||E||_\varrho$ denotes the operator norm. By the orthonormality of $\left(\tilde{T}_\mu\right)_{\mu\in\mathbb{N}_0^D}$, it follows that

$$||E||_\varrho = \sup_{f\in L^2(B(\mathcal{P},\varrho))\setminus\{0\}} \frac{|E(f)|}{\|f\|_\varrho} = \sqrt{\sum_{\mu\in\mathbb{N}_0^D} |E(\tilde{T}_\mu)|^2}.$$

In the following, let E be the error of the Chebyshev polynomial interpolation at a fixed $p \in \mathcal{P}$,

$$E(f) := f(p) - I_{\overline{N}}(f(\cdot))(p).$$

Starting with (23), we first focus on $||E||_\varrho$,

$$\|E\|_\varrho^2 = \sum_{\mu\in\mathbb{N}_0^D} |E(\tilde{T}_\mu)|^2 = \sum_{\mu\in\mathbb{N}_0^D} c_\mu^2 |E(T_\mu)|^2.$$

From now on the proof differs from [45, Proof of Lemma 7.3.3], since we use the values of N_i, $i = 1,\dots,D$ and ϱ_i, $i = 1,\dots,D$. Since we chose Chebyshev points of the second kind instead of Chebyshev points of the first kind in the Chebyshev interpolation, we cannot apply [45, Corollary 7.3.1], but adjust this in Lemma B.1 to the Chebyshev points of the second kind. At this step, we apply Lemma B.1 to obtain

$$\sum_{\mu\in\mathbb{N}_0^D} c_\mu^2 |E(T_\mu)|^2 = \sum_{\mu\in\mathbb{N}_0^D,\exists i:\mu_i>N_i} c_\mu^2 |E(T_\mu)|^2 \leq \sum_{\mu\in\mathbb{N}_0^D,\exists i:\mu_i>N_i} 4c_\mu^2.$$

Overall, using $\left(\prod_{j=1}^D \varrho_j^{2\mu_j} + x\right)^{-1} \leq \left(\prod_{j=1}^D \varrho_j^{2\mu_j}\right)^{-1} = \prod_{j=1}^D \varrho_j^{-2\mu_j}$ for $x > 0$,

$\mu_j \in \mathbb{N}_0$ and $j = 1, \ldots, D$ and this leads to

$$
\begin{aligned}
\|E\|_\varrho^2 &\le 4 \sum_{\mu \in \mathbb{N}_0^D, \exists i: \mu_i > N_i} c_\mu^2 \le 4 \left(\frac{2}{\pi}\right)^D \sum_{i=1}^D \left(\sum_{\mu \in \mathbb{N}_0^D, \mu_i > N_i} \prod_{j=1}^D \varrho_j^{-2\mu_j} \right) \\
&\le 4 \left(\frac{2}{\pi}\right)^D \sum_{i=1}^D \varrho_i^{-2N_i} \left(\sum_{\mu \in \mathbb{N}_0^D, \mu_i > N_i} \varrho_i^{-2(\mu_i - N_i)} \prod_{j=1, j \neq i}^D \varrho_j^{-2\mu_j} \right) \\
&\le 4 \left(\frac{2}{\pi}\right)^D \sum_{i=1}^D \varrho_i^{-2N_i} \left(\sum_{\mu \in \mathbb{N}_0^D} \prod_{j=1}^D \varrho_j^{-2\mu_j} \right).
\end{aligned}
$$

From this point on, we use the convergence of the geometric series, since $|\varrho_j^{-2}| < 1,\ j = 1, \ldots, D$,

$$
\begin{aligned}
\|E\|_\varrho^2 &\le 4 \left(\frac{2}{\pi}\right)^D \sum_{i=1}^D \varrho_i^{-2N_i} \left(\sum_{\mu_1=0}^\infty \cdots \sum_{\mu_D=0}^\infty \prod_{j=1}^D \varrho_j^{-2\mu_j} \right) \\
&= 4 \left(\frac{2}{\pi}\right)^D \sum_{i=1}^D \varrho_i^{-2N_i} \prod_{j=1}^D \frac{1}{1 - \varrho_j^{-2}}.
\end{aligned}
$$

Recalling (23), we have to estimate $\|f\|_\varrho$,

$$
\|f\|_\varrho^2 = \int_{B(\mathcal{P},\varrho)} \frac{f(z)\overline{f(z)}}{\prod_{i=1}^D \sqrt{|1 - z_i^2|}} \,\mathrm{d}z \le \left(\sup_{z \in B(\mathcal{P},\varrho)} |f(z)| \right)^2 \|1\|_\varrho^2.
$$

From $\pi^{\frac{D}{2}} \tilde{T}_0 = 1$ it directly follows that $\|1\|_\varrho^2 = \left(\pi^{\frac{D}{2}}\right)^2 \|\tilde{T}_0\|_\varrho^2 = \pi^D$ and hence

$$
\|f\|_\varrho^2 \le \pi^D \cdot V^2.
$$

Combining the results leads to

$$
\begin{aligned}
|E(f)| = \left|f(p) - I_{\overline{N}}(f(\cdot))(p)\right| &\le \left(\pi^D \cdot V^2 \cdot 4 \left(\frac{2}{\pi}\right)^D \sum_{i=1}^D \varrho_i^{-2N_i} \prod_{j=1}^D \frac{1}{1 - \varrho_j^{-2}} \right)^{\frac{1}{2}} \\
&= 2^{\frac{D}{2}+1} V \left(\sum_{i=1}^D \varrho_i^{-2N_i} \prod_{j=1}^D \frac{1}{1 - \varrho_j^{-2}} \right)^{\frac{1}{2}}. \qquad \square
\end{aligned}
$$

$\square$

The following lemma shows that the Chebyshev interpolation of a polynomial with a degree at most as high as the degree of the interpolating Chebyshev polynomial is exact and furthermore determines an upper bound for interpolating Chebyshev polynomials with higher degrees.

Lemma B.1. *For $x \in [-1,1]^D$, it holds that*

$$|T_\mu(x) - I_{\overline{N}}(T_\mu(\cdot))(x)| = 0 \quad \forall \mu \in \mathbb{N}_0^D : \mu_i \leq N_i, i = 1, \dots, D, \tag{24}$$
$$|T_\mu(x) - I_{\overline{N}}(T_\mu(\cdot))(x)| \leq 2 \quad \forall \mu \in \mathbb{N}_0^D : \exists i \in \{1, \dots, D\} : \mu_i > N_i. \tag{25}$$

Proof. The uniqueness properties of the Chebyshev interpolation directly imply (24). The proof of (25) is similar to [45, Proof of Hilfssatz 7.3.1]. They use the zeros of the Chebyshev polynomial as interpolation points, whereas we use the extreme points and, therefore, we use a different orthogonality property in this proof. We first focus on the one-dimensional case. Recalling (2), the Chebyshev interpolation of T_μ, $\mu > N$, is given as

$$I_N(T_\mu)(x) = \sum_{j=0}^{N} c_j T_j(x) \quad \text{with} \quad c_j = \frac{2^{\mathbb{1}_{0<j<N}}}{N} {\sum_{k=0}^{N}}'' T_\mu(x_k) T_j(x_k), \quad j \leq N,$$

where x_k denotes the k-th extremum of T_N. Here, we can apply the following orthogonality property ([42, p.54]),

$$
{\sum_{k=0}^{N}}'' T_\mu(x_k) T_j(x_k) = \begin{cases} 0, & \mu + j \neq 0 \bmod (2N) \text{ and } |\mu - j| \neq 0 \bmod (2N), \\ N, & \mu + j = 0 \bmod (2N) \text{ and } |\mu - j| = 0 \bmod (2N), \\ \frac{N}{2}, & \mu + j = 0 \bmod (2N) \text{ and } |\mu - j| \neq 0 \bmod (2N), \\ \frac{N}{2}, & \mu + j \neq 0 \bmod (2N) \text{ and } |\mu - j| = 0 \bmod (2N). \end{cases} \tag{26}
$$

For $j \leq N$ and $\mu > N$, this yields the existence of $\gamma \leq N$ such that

$$I_N(T_\mu) = T_\gamma. \tag{27}$$

(27) follows elementarily from the case where for any $\mu > N$ the orthogonality leads to a coefficient $c_j > 0$ for exactly one $0 \leq j \leq N$.

To prove the claim, we distinguish several cases. In all of these cases, we assume that there exists $0 \leq j \leq N$ such that ${\sum_{k=0}^{N}}'' T_\mu(x_k) T_j(x_k) \neq 0$. We will then show that for all other $0 \leq i \leq N, i \neq j$ it follows that ${\sum_{k=0}^{N}}'' T_\mu(x_k) T_j(x_k) = 0$.

Firstly, assume there exists j such that $\mu + j = 0 \bmod (2N)$ and $\mu - j = 0 \bmod (2N)$. Then it directly follows for all $0 \leq i \leq N$, $i \neq j$ that $\mu + i \neq 0 \bmod (2N)$ and $\mu - i \neq 0 \bmod (2N)$.

Secondly, assume there exists j such that $\mu + j = 0 \bmod (2N)$ and $\mu - j \neq 0 \bmod (2N)$. Analogously, for all $0 \leq i \leq N$, $i \neq j$, we have $\mu + i \neq 0 \bmod (2N)$ and additionally from $\mu + j = 0 \bmod (2N)$ it follows that $\mu + j - 2N = 0 \bmod (2N)$ and hence for all $0 \leq i \leq N$, $i \neq j$, we have $\mu - i > \mu + j - 2N$, which is equivalent to $\mu - i \neq 0 \bmod (2N)$.

Similar argumentation holds for the third case $\mu + j \neq 0 \bmod (2N)$ and $|\mu - j| = 0 \bmod (2N)$.

Consequently, (27) holds and it directly follows that $|T_\mu - I_N(T_\mu)| \leq |T_\mu| + |I_N(T_\mu)| \leq 1 + 1 = 2$. Thus (25) holds in the one-dimensional case. The extension to the $D-$dimensional case follows analogously by applying the triangle inequality $|\prod_{i=1}^D T_{i,\mu_i} - \prod_{i=1}^D I_{N_i}(T_{i,\mu_i})| \leq |\prod_{i=1}^D T_{i,\mu_i}| + |\prod_{i=1}^D I_{N_i}(T_{i,\mu_i})| \leq \prod_{i=1}^D |T_{i,\mu_i}| + \prod_{i=1}^D |I_{N_i}(T_{i,\mu_i})|$ and applying the one-dimensional result to each tensor component. □ □

C Proof of Proposition 2.3

Proof. Before we apply [5, Theorem 3.1], which assumes $\mathcal{P} = [-1,1]^D$, we investigate how the linear transformation $\tau_{\mathcal{P}}$, as introduced in the proof of Proposition 2.1, influences the derivatives. Let $p \mapsto Price^p$ be a function on $\mathcal{P}$. We set $\widehat{h}(p) = Price^p \circ \tau_{\mathcal{P}}(p)$. Furthermore, let $\widehat{I_{\overline{N}}}(\widehat{h})(p)$ be the Chebyshev interpolation of $\widehat{h}(p)$ on $[-1,1]^D$. Then, it directly follows that

$$Price^p - I_{\overline{N}}(Price^{(\cdot)})(p) = \left(\widehat{h}(\cdot) - \widehat{I_{\overline{N}}}(\widehat{h})(\cdot)\right) \circ \tau_{\mathcal{P}}^{-1}(p).$$

First, let us assume $D = 1$, i.e. $\mathcal{P} = [\underline{p}, \overline{p}]$, and let $\alpha \in \mathbb{N}_0$. For the partial derivatives, it holds that

$$\begin{aligned}
\partial^\alpha Price^p - \partial^\alpha I_{\overline{N}}(Price^{(\cdot)})(p) &= \partial^\alpha \left(Price^p - I_{\overline{N}}(Price^{(\cdot)})(p)\right) \\
&= \partial^\alpha \left(\left(\widehat{h}(\cdot) - \widehat{I_{\overline{N}}}(\widehat{h})(\cdot)\right) \circ \tau_{\mathcal{P}}^{-1}(p)\right) \\
&= \partial^{\alpha-1} \left(\partial^1 \widehat{h}(\tau_{\mathcal{P}}^{-1}(p)) - \partial^1 \widehat{I_{\overline{N}}}(\widehat{h}^{(\cdot)})(\tau_{\mathcal{P}}^{-1}(p))\right) \\
&= \partial^{\alpha-1} \frac{2}{\overline{p} - \underline{p}} \left(\left[\partial^1 \widehat{h}\right](\tau_{\mathcal{P}}^{-1}(p)) - \left[\partial^1 \widehat{I_{\overline{N}}}(\widehat{h}^{(\cdot)})\right](\tau_{\mathcal{P}}^{-1}(p))\right).
\end{aligned}$$

Repeating this step iteratively yields

$$\partial^\alpha Price^p - \partial^\alpha I_{\overline{N}}(Price^{(\cdot)})(p) = \frac{2^\alpha}{(\overline{p} - \underline{p})^\alpha} \left(\left[\partial^\alpha \widehat{h}\right](\tau_{\mathcal{P}}^{-1}(p)) - \left[\partial^\alpha \widehat{I_{\overline{N}}}(\widehat{h}^{(\cdot)})\right](\tau_{\mathcal{P}}^{-1}(p))\right).$$

This scales the error in $[-1,1]$ by a factor $\frac{2^\alpha}{(\overline{p} - \underline{p})^\alpha}$. Extending this to the D-variate case where $\alpha = (\alpha_1, \ldots, \alpha_D) \in \mathbb{N}_0^D$ is a multi-index and $\partial^\alpha = \partial^{\alpha_1} \cdots \partial^{\alpha_D}$ results in

$$\partial^\alpha Price^p - \partial^\alpha I_{\overline{N}}(Price^{(\cdot)})(p) = \\ \prod_{i=1}^{D} \frac{2^{|\alpha_i|}}{(\overline{p}_i - \underline{p}_i)^{|\alpha_i|}} \left(\left[\partial^\alpha \widehat{h}\right](\tau_{\mathcal{P}}^{-1}(p)) - \left[\partial^\alpha \widehat{I_{\overline{N}}}(\widehat{h}^{(\cdot)})\right](\tau_{\mathcal{P}}^{-1}(p))\right).$$

From Theorem 3.1 in [5], the assertion follows directly for $\widehat{h}(\cdot)$ on $\mathcal{P} = [-1,1]^D$, i.e. for any $\frac{D}{2} < \sigma \in \mathbb{N}$ and any $\sigma \geq \mu \in \mathbb{N}_0$ there exists a constant $\tilde{C} > 0$ such that

$$\|\widehat{h}(\cdot) - \widehat{I_{\overline{N}}}(\widehat{h})(\cdot)\|_{W_2^{\mu,\omega}(\mathcal{P})} \leq \tilde{C} N^{2\mu - \sigma} \|\widehat{h}(\cdot)\|_{W_2^{\sigma,\omega}(\mathcal{P})}, \tag{28}$$

For arbitrary $\mathcal{P}$, the constant from (28) has to be multiplied with the corresponding factor resulting from the linear transformation $\tau_{\mathcal{P}}$. $\square$ $\square$

acknowledgements

We thank the KPMG Center of Excellence in Risk Management for their support.
We would also like to thank Jonas Ballani, Behnam Hashemi, Daniel Kressner and Nick Trefethen for fruitful discussions on Chebyshev interpolation. Moreover, we gratefully acknowledge valuable feedback from Christian Bayer, Ernst Eberlein, Alexander Gnedin, Emmanuel Gobet, Dilip Madan, Christian Pötz, Peter Tankov, Ralf Werner and two anonymous referees. For further discussions, we thank Paul Harrenstein and Pit Forster. Additionally, we thank the participants of the conferences Advanced Modelling in Mathematical Finance, A conference in honour of Ernst Eberlein, Kiel. May 20–22, 2015, Stochastic Methods in Physics and Finance, 2015 in Heraklion, MoRePas2015: Model Reduction of Parametrized Systems III, held in Trieste and the 12th German Probability and Statistic Days 2016 in Bochum. Furthermore, we thank the participants of the research seminars Seminar Stochastische Analysis und Stochastik der Finanzmärkte, Technical University Berlin. May 28, 2015 and the Groupe de Travail MathfiProNum: Finance mathématique, probabilités numériques et statistique des processus, Université Diderot, Paris. June 11, 2015.

References

[1] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[2] Svetlana I Boyarchenko and Sergei Z Levendorskiĭ. *Non-Gaussian Merton-Black-Scholes Theory*, volume 9. World Scientific, 2002.

[3] M. J. Brennan and E. S. Schwartz. The valuation of American put options. *The Journal of Finance*, 2(32):449–462, 1977.

[4] O. Burkovska, Bernard Haasdonk, J. Salomon, and B. Wohlmuth. Reduced basis methods for pricing options with the Black-Scholes and Heston models. *SIAM Journal on Financial Mathematics*, 6(1):685–712, 2015.

[5] C Canuto and A Quarteroni. Approximation results for orthogonal polynomials in Sobolev spaces. *Mathematics of Computation*, 38(157):67–86, 1982.

[6] P. Carr and D. B. Madan. Option valuation and the fast Fourier transform. *Journal of Computional Finance*, 2(4):61–73, 1999.

[7] R. Cont, N. Lantos, and O. Pironneau. A reduced basis for option pricing. *SIAM Journal on Financial Mathematics*, 2(1):287–316, 2011.

[8] C. Cuchiero, M. Keller-Ressel, and J. Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 4(16), 2012.

[9] Philip J Davis. *Interpolation and approximation.* Courier Corporation, 1975.

[10] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053, 2003.

[11] E. Eberlein, K. Glau, and A. Papapantoleon. Analysis of Fourier transform valuation formulas and applications. *Applied Mathematical Finance*, 17(3):211–240, 2010.

[12] E. Eberlein, U. Keller, and K. Prause. New insights into smile, mispricing and value at risk: the hyperbolic model. *Journal of Business*, 71(3):371–405, 1998.

[13] E. Eberlein and F. Özkan. The Lévy LIBOR model. *Finance and Stochastics*, 9(3):327–348, 2005.

[14] L. Feng and V. Linetsky. Pricing discretely monitored barrier options and defaultable bonds in Lévy process models: A fast Hilbert transform approach. *Mathematical Finance*, 18(3):337–384, 2008.

[15] D. Filipović, M. Larsson, and A. Trolle. Linear-rational term structure models. *The Journal of Finance*, 2016. published online.

[16] M. Gaß, K. Glau, and M. Mahlstedt. Analyticity of parametric expectations: Chebyshev interpolation for the computation of option prices and sensitivities. work in progress, 2017.

[17] M. Gaß, K. Glau, and M. Mair. Magic Points in Finance: Empirical Interpolation for Parametric Option Pricing. Preprint available on `https://arxiv.org/abs/1511.00884`, 2016.

[18] Michael B Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259, 2015.

[19] K. Glau, B. Hashemi, M. Mahlstedt, and C. Pötz. Spread option in 2D Black-Scholes. Example for Chebfun3 in the Chebfun toolbox, `http://www.chebfun.org/examples/applics/BlackScholes2D.html`.

[20] K. Glau, P. Herold, D. B. Madan, and C. Pötz. The chebyshev method for the implied volatility. `https://arxiv.org/abs/1710.01797`.

[21] K. Glau and C. Pötz. The chebyshev method to solve dynamic programming problems. working paper.

[22] Kathrin Glau and Mirco Mahlstedt. Improved error bound for multivariate Chebyshev polynomial interpolation. preprint avalaible on `https://arxiv.org/abs/1611.08706`, 2016.

[23] Bernard Haasdonk, Julien Salomon, and Barbara Wohlmuth. A reduced basis method for the simulation of American options. In *Numerical Mathematics and Advanced Applications 2011*, pages 821–829. Springer, 2013.

[24] S. Heinrich. Monte Carlo complexity of global solution of integral equations. *Journal of Complexity*, 14(2):151–175, 1998.

[25] S. Heinrich. Multilevel Monte Carlo methods. In *International Conference on Large-Scale Scientific Computing*, pages 58–67. Springer, 2001.

[26] S. Heinrich and E. Sindambiwe. Monte Carlo complexity of parametric integration. *Journal of Complexity*, 15(3):317–341, 1999.

[27] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[28] K. Jänich. *Funktionentheorie: eine Einführung*. Springer-Lehrbuch. Springer, 2004.

[29] M. Keller-Ressel, A. Papapantoleon, and J. Teichmann. The affine libor models. *Mathematical Finance*, 23(4):627–658, 2013.

[30] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.

[31] O. Kudryavtsev and S. Z. Levendorskiĭ. Fast and accurate pricing of barrier options under Lévy processes. *Finance and Stochastics*, 13(4):531–562, 2009.

[32] Pierre L'Ecuyer. Quasi-Monte Carlo methods with applications in finance. *Finance and Stochastics*, 13(3):307–349, 2009.

[33] R. W. Lee. Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.

[34] R. Lord, F. Fang, F. Bervoets, and C. W. Oosterlee. A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes. *SIAM Journal on Scientific Computing*, 30(4):1678–1705, 2008.

[35] R. C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

[36] N. Metropolis. The beginning of the Monte Carlo method. pages 125–130. 1987. Los Alamos Science. Special Issue dedicated to Stanislaw Ulam.

[37] R. Pachon. Numerical pricing of European options with arbitrary payoffs. Available at SSRN: `http://ssrn.com/abstract=2712402`, 2016.

[38] O. Pironneau. Reduced basis for vanilla and basket options. *Risk and Decision Analysis*, 2(4):185–194, 2011.

[39] M. Pistorius and J. Stolte. Fast computation of vanilla prices in time-changed models and implied volatilities. *International Journal of Theoretical and Applied Finance*, 15(4):1250031–1250031, 2012.

[40] R. B. Platte and N. L. Trefethen. *Chebfun: A new kind of numerical computing*, pages 69–86. Springer, 2008.

[41] S. Raible. *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*. PhD thesis, Universität Freiburg, 2000.

[42] Theodore-J Rivlin. *Chebyshev polynomials.* John Wiley & Sons, Inc., copublished in the United States with John Wiley &, 1990.

[43] C. Runge. Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. *Zeitschrift für Mathematik und Physik*, 46:224–243, 1901.

[44] E. W. Sachs and M. Schu. Reduced order models in PIDE constrained optimization. *Control and Cybernetics*, 39(3):661–675, 2010.

[45] Stefan Sauter and Christoph Schwab. *Randelementmethoden: Analyse, Numerik und Implementierung schneller Algorithmen.* Vieweg+Teubner Verlag, 2004.

[46] Eitan Tadmor. The exponential accuracy of Fourier and Chebyshev differencing methods. *SIAM Journal on Numerical Analysis*, 23(1):1–10, 1986.

[47] L. N. Trefethen. Talk: Six myths of polynomial interpolation and quadrature. `https://people.maths.ox.ac.uk/trefethen/mythstalk.pdf`, 2011.

[48] L. N. Trefethen. *Approximation Theory and Approximation Practice.* SIAM books, 2013.

[49] J. Wloka. *Partial Differential Equations.* Cambridge University Press, 1987.