Analytic models for parameter dependency in option price modelling

Analytic models for Parameter dependency in Option price modelling

A. Cuyt, O. Salazar Celis, M. Lukach, K. In't Hout

November 13, 2014

Abstract

Options are a type of financial instrument classed as derivatives, as they derive their value from an underlying asset.

The equations used to model the option price are often expressed as partial differential equations (PDEs). Once expressed in this form, a discretization method on a finite grid can be applied and the numerical valuation obtained. Remains the problem of writing down an (approximate) closed-form analytic model for the option price in function of all the variables and parameters, which is the main objective of this paper.

At the same time we also consider the Greeks, which are the quantities representing the sensitivities of the price to a change in the underlying variables or parameters. Discrete values for these Greeks can again be derived, either directly from the differentiation matrices occurring in the option price PDE or by solving new but similar PDEs. Next, analytic models for the Greeks are computed in the same way as for the option price.

As a prototype case, the Black-Scholes PDE for European call options is considered.

1 Introduction.

The basic product traded in financial option markets is a European call option. It gives its holder the right, but not the obligation, to purchase from the writer of the option a prescribed asset for a prescribed price E at a prescribed time T in the future. The quantity E is called the exercise price or strike price and T is called the maturity time. In return for granting the option, the writer collects an up-front payment from the buyer.

Black-Scholes. In contemporary financial option pricing theory, the fair value C of a European call option is evaluated according to a given stochastic model for the evolution of the underlying asset price. The standard model for the asset price evolution assumes a

geometric Brownian motion. In their seminal work, Black and Scholes [2] derived from this and additional assumptions on the market, a partial differential equation (PDE) that must be satisfied:

$$\frac{\partial C}{\partial t} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC, \qquad S > 0,\ t > 0. \tag{1}$$

Here $C = C(S,t)$ denotes the fair value of the European call option if at t time units before maturity the asset price equals S. Thus, C is a deterministic function of two independent real variables, S and t. The parameters r and σ are real and denote the risk-neutral interest rate and the volatility, respectively. The Black–Scholes PDE (1) is complemented with initial and boundary conditions:

$$\begin{aligned} C(S,0) &= \max(0, S-E), \qquad S \geq 0, \\ C(0,t) &= 0, \qquad t \geq 0, \\ C(S,t) &\approx S, \qquad t \geq 0 \text{ and } S \text{ large}. \end{aligned}$$

Note that $C(S,0)$ is the payoff of the option at maturity.

It is well recognized in the literature that the original assumptions in the Black–Scholes framework are not fulfilled in the contemporary markets [7]. More advanced asset price models are being considered, for example with stochastic volatility. Nevertheless, the Black–Scholes model remains an important tool in financial practice. Furthermore, it constitutes the germ of a variety of more advanced models that are being considered in the literature today. In view of this, it serves very well to illustrate our proof of principle in this paper. Also, since a semi-closed form formula is available for European call option prices (see [2, 10] and Section 4), it enables a study of the approximation errors.

The approach we follow starts with existing techniques for the numerical solution of the option pricing PDE with the property that they generate pointwise approximations. First the parameters E, r, σ are specified and an approximation $\widetilde{C}$ to C is obtained on a (uniform or non-uniform) grid of (S,t) values. For the numerical experiments in this paper we have applied a standard numerical method which employs a second-order central finite difference discretization in the S-domain followed by a Crank-Nicolson discretization (implicit trapezoidal rule) in the t-domain. This procedure can be repeated several times, which becomes expensive when considering some advanced (multidimensional) generalizations of the Black-Scholes PDE. For a few dozen (E, r, σ) triplets and a moderate (S,t) grid, for instance with 100×50 points, one quickly obtains hundreds of thousands of values $\widetilde{C}(S,t;E,r,\sigma)$. Our objective in this paper is to construct, after obtaining pointwise approximate option values for several parameter sets, an accurate, multivariate continuous model for $\widetilde{C}$ that allows fast evaluation of the option value function in terms of all independent variables as well as parameters. Consequently, option prices can be computed for new values of the variables and parameters S, t, E, r, σ with minor effort, compared to numerically solving the PDE. This

approach is highly promising for the large variety of advanced (multidimensional) PDEs, with multiple parameters, that arise in financial mathematics today, for which no solutions in semi-closed form are available. The subject is treated in more detail in Sections 4 and 5.

Greeks. Besides the option price we also consider the Greeks, which are the quantities representing the sensitivities of the option price to a change in the variables and parameters. In particular, values for the partial derivatives delta, gamma and vega

$$\Delta := \frac{\partial C}{\partial S}, \qquad \Gamma := \frac{\partial^2 C}{\partial S^2}, \qquad \text{vega} := \frac{\partial C}{\partial \sigma}$$

are of interest. Since delta and gamma appear in the Black–Scholes PDE (1), finite difference approximations for these two quantities can directly be obtained during the numerical solution of (1). Approximations for vega can be computed by simultaneously solving along with (1) the PDE obtained by differentiating (1) with respect to σ,

$$\frac{\partial V}{\partial t} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \sigma S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV, \qquad S > 0,\ t > 0.$$

For the latter PDE homogeneous initial and boundary conditions apply.

2 Domain, dimensionality reduction and duality

Before we apply the rational approximation method described in Section 3 to the data obtained from the PDE solver, we need to specify the approximation domain. Some regions are inherently more interesting than others. Moreover, although in theory nothing prevents us from working with the current 5 variables, the proposed method benefits from some dimension reduction techniques and exploiting both homogeneity and symmetries. These issues are described next. Their use also facilitates the graphical illustration of the results.

Domain. Since we know that $C \approx 0$ for very small S and $C \approx S$ for sufficiently large S, we are especially interested in the region where C/S makes the transition from 0 to 1. Our domain is $]0, 200]$ for S, $]0, 100]$ for E, $[0, 0.2]$ for r, $[0, 1]$ for σ and $[0, 3]$ for time t to expiration (in years). In Figure 1 we show the (E, r, σ) triplets for which we solve the PDE (they were randomly generated in the cube $[0, 100] \times [0, 0.2] \times [0, 1]$). Since a static 3-d view is difficult to interpret, we have chosen to give (from left to right) the projections on the (E, r)–, (r, σ)– and (E, σ)–planes.

Dimensionality reduction. It is convenient to introduce the new variables [7, p.110],[8]

$$u = \ln(S) - \ln(E) + rt, \tag{2}$$

$$w = \sigma\sqrt{t}, \tag{3}$$

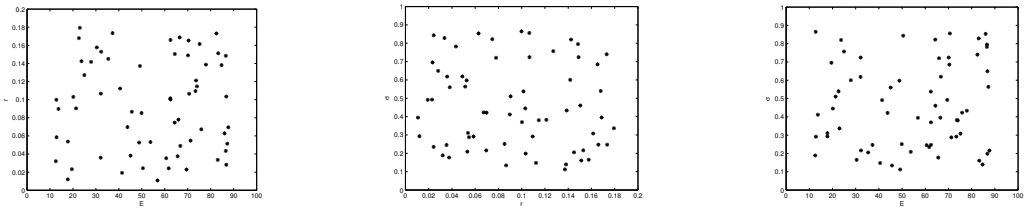


Figure 1: Randomly sampled (E, r, σ) in $]0, 100] \times [0, 0.2] \times [0, 1]$ (64 triplets).

and consider $c(u, w) = C/S$. This is allowed for specific types of options. For instance, it is well-known that European call options are homogeneous of degree one in S and E [2, 10].

Hence each quintuple $(S, t; E, r, \sigma)$ is mapped to some vector $v = (u, w)$. We use the computed values for C at $(S_i, t_i; E_i, r_i, \sigma_i)$ to obtain

$$c(u_i, w_i) = C(S_i, t_i; E_i, r_i, \sigma_i)/S_i. \tag{4}$$

Application of Leibniz's rule gives for the Greeks

$$\Delta := \frac{\partial C}{\partial S} = c(u, w) + \frac{\partial c(u, w)}{\partial u} = \Delta(u, w), \tag{5}$$

$$\Gamma := \frac{\partial^2 C}{\partial S^2} = \frac{1}{S}\left(\frac{\partial c(u, w)}{\partial u} + \frac{\partial^2 c(u, w)}{\partial u^2}\right) = \frac{1}{S}\gamma(u, w), \tag{6}$$

$$\text{vega} := \frac{\partial C}{\partial \sigma} = S\sqrt{t}\ \frac{\partial c(u, w)}{\partial w} = \left(S\sqrt{t}\right)\ \kappa(u, w), \tag{7}$$

where the newly introduced γ and κ are functions of (u, w).

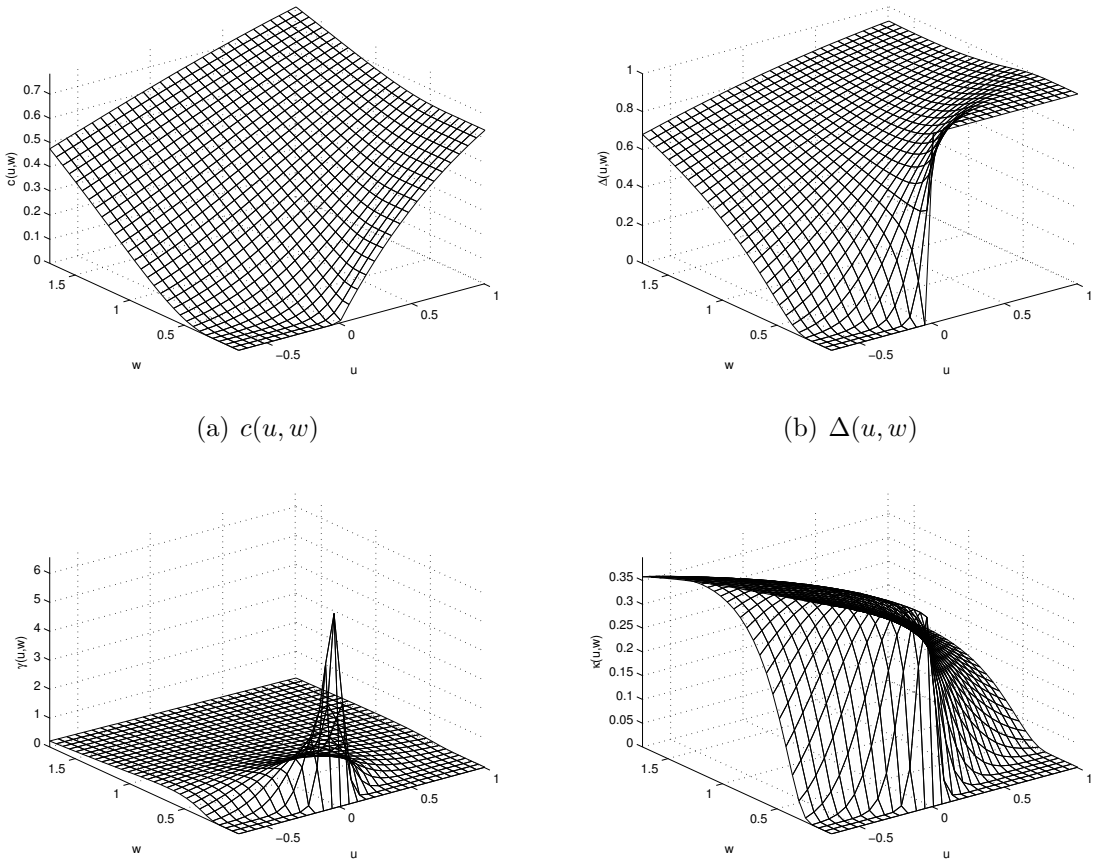
For the interesting region where $1/2 \le S/E \le 3/2$ and the ranges of the remaining variables as above, the ranges of u and w are respectively

$$\ln(1/2) \le u \le \ln(3/2) + 0.6, \qquad 0 \le w \le \sqrt{3}. \tag{8}$$

Plots of c, Δ, γ and κ on this domain are shown in Figure 2.

At this point, it is worth noting that given a sufficiently differentiable approximation for the option price C, approximations for the Greeks can obviously be obtained directly from differentiating this approximation with respect to S or σ. Equations (5) to (7) indicate that such approximations can also be derived from differentiating (with respect to u or w) an approximation for $c(u, w)$. However, such a direct approach is not advised because typically one order of magnitude in accuracy is lost in each differentiation step. Hence, instead of differentiating an approximation for $c(u, w)$ we use the computed values for Δ, Γ and vega at $(S_i, t_i; E_i, r_i, \sigma_i)$ to obtain

$$\begin{aligned} &\Delta(u_i, w_i), \\ &\gamma(u_i, w_i) = \Gamma(u_i, w_i)S_i, \\ &\kappa(u_i, w_i) = \text{vega}(u_i, w_i)/(S_i\sqrt{t_i}). \end{aligned}$$



(a) $c(u, w)$ (b) $\Delta(u, w)$

(c) $\gamma(u, w)$ (d) $\kappa(u, w)$

Figure 2: The exact functions expressed in the variables u and w.

From these values we construct independent approximations for each of the Greeks.

Duality. We can further restrict the domain of approximation to either $u \geq 0$ or $u \leq 0$ due to the following duality [8],[6]:

$$c(-u, w) = e^u c(u, w) + 1 - e^u. \tag{9}$$

Applying this relation to the Greeks gives for $(u, w) \neq (0, 0)$:

$$\Delta(-u, w) = e^u c(u, w) + 1 - e^u \Delta(u, w), \tag{10}$$

$$\gamma(-u, w) = e^u \gamma(u, w), \tag{11}$$

$$\kappa(-u, w) = e^u \kappa(u, w). \tag{12}$$

Let D^+ and D^- denote the part of the domain where $u \geq 0$ and $u \leq 0$ respectively.

When extending the dualities (9–12) to approximations for c, Δ, γ and κ, we have to take into account that we may be introducing discontinuities at $u = 0$. For instance, if $c(u, w)$ on D^- is approximated by $r^-(u, w)$ and if we define

$$r^+(u, w) := e^{-u} r^-(-u, w) + 1 - e^{-u}, \qquad (u, w) \in D^+,$$

then $r^+(0, w) = r^-(0, w)$ but the approximation is not differentiable anymore at $u = 0$. With Δ, γ and κ the approximation itself even becomes discontinuous so that we need to define it on the line $u = 0$ as the average of the approximations on D^+ and D^-. Note that this discontinuity is not visible in the displayed figures, such as Figure 4, because the approximation is sufficiently accurate (the threshold on the relative error is too small to be visible in the graphics; the discontinuity is visible with a coarser approximation though).

We remark that we compute the approximation $r(u, w)$ on D^- (and not on D^+) and extend it from there to the rest of the domain because then the error is not magnified by a factor e^u with $u \geq 0$. It is merely multiplied by a factor less than 1. A similar remark was pointed out in [8].

3 Rational interval interpolation

We aim to capture the underlying parameter dependency with a generalized rational function

$$r_{\ell,m}(u, w) = \frac{p_{\ell,m}(u, w)}{q_{\ell,m}(u, w)} = \frac{\sum_{i=0}^{\ell} p_i \, b_i(u, w)}{\sum_{i=0}^{m} q_i \, b_i(u, w)},$$

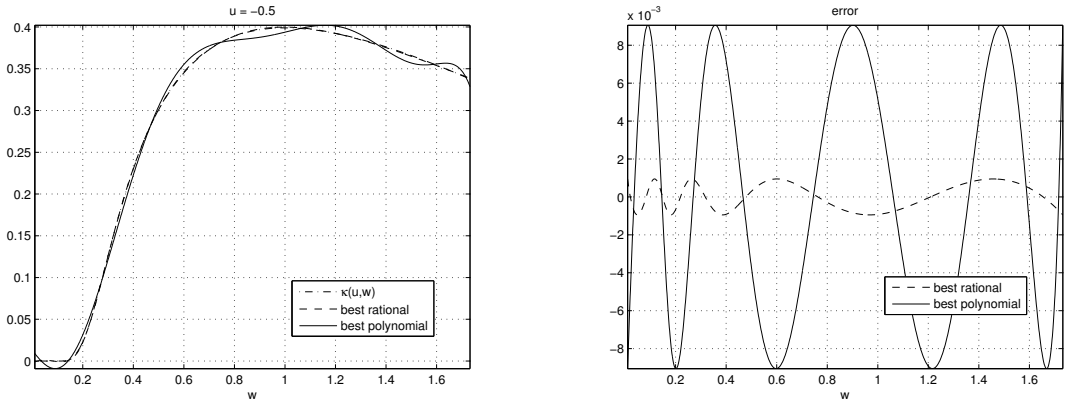
where the multivariate basis functions $b_i(u, w)$ can be multinomials $u^{k1} w^{k2}$ $(k1, k2) \in \mathbb{N}^2$, or some multivariate orthogonal polynomials, trigonometric or other basis functions. In what follows we simply take the multinomials as basis functions, ordered by increasing total degree, with $u^k w^\ell < u^{k-i} w^{\ell+i}$ for $i > 0$.

The rational function $r_{\ell,m}(u, w)$ essentially has only $\ell + m + 1$ degrees of freedom because one of the coefficients p_i and q_i can always be fixed to normalize the representation (a common choice is $q_0 = 1$). Typically, fairly low degree rational functions (i.e. $\ell + m$ small) can accurately approximate functions like the ones shown in Figure 2. This is mainly due to their outstanding ability to display flat behavior followed by a sudden and steep increase. As an example, consider different approximations for the univariate slice $\kappa(-0.5, w)$, where $u = -0.5$ is fixed and $0 \leq w \leq \sqrt{3}$. It is not difficult to find, using the Black-Scholes formula,

that an analytical expression for $\kappa(u, w)$ is

$$\kappa(u, w) = \frac{e^{-\frac{(2u+w^2)^2}{8w^2}}}{\sqrt{2\pi}}.$$

For the univariate function $\kappa(-0.5, w)$, we compare a low degree best rational approximation to the best polynomial approximation with the same number of unknowns. The distance is measured using the uniform (Chebyshev) norm $||f(w)||_\infty = \sup\{|f(w)| : 0 \leq w \leq \sqrt{3}\}$. Figure 3(a) shows the best rational approximation of degree 4 in numerator and denominator as well as the best polynomial approximation of degree 8. The polynomial approximation is clearly a lot worse than the rational approximation, which is visually indistinguishable from the original function. How much worse is shown in Figure 3(b), where the difference between $\kappa(-0.5, w)$ and the rational, respectively the polynomial approximation is shown. Note the characteristic equi-oscillating behavior of the error curves. The maximum error of the polynomial is 0.0091, while that of the rational function is 0.00095.



(a) Best rational and polynomial approximations.

(b) Corresponding approximation errors.

Figure 3: Univariate approximations for the slice $\kappa(-0.5, w)$ on $0 \leq w \leq \sqrt{3}$.

Although best rational approximations are optimal, their possible non-existence [3] or non-uniqueness [14] are profound and subtle topics which require special attention. Besides these theoretical considerations, they are usually computationally more expensive to obtain and in general require iterative techniques. For instance, for fixed ℓ and m, the differential correction algorithm [1] for best discrete rational approximations requires solving a sequence of linear programming (LP) problems. In the current paper, we propose to solve the involved approximation problem by a different and very natural method of which the details are given next.

As explained in the Sections 1 and 2, from a PDE solver, approximate values $\tilde{f}_i$ are obtained at distinct locations (u_i, w_i), $i = 0, ..., s$, after the dimensionality reduction. Here the $\tilde{f}_i$ denote approximate values for either the option price or one of the aforementioned Greeks. A useful property of some modern discretization methods, in which the time steps and the grid sizes are determined adaptively, is their ability to estimate the errors of the computed values (see for instance [9]). That is, the computed $\tilde{f}_i$ obtained from such methods are expected to approximate the (essentially unknown) exact f_i within a user-specified (absolute or relative) tolerance $\epsilon > 0$. Consequently, the upper and lower bounds

$$\underline{f_i} = \tilde{f}_i - \epsilon, \qquad \overline{f_i} = \tilde{f}_i + \epsilon$$

or

$$\underline{f_i} = \tilde{f}_i(1 - \epsilon), \qquad \overline{f_i} = \tilde{f}_i(1 + \epsilon), \tag{13}$$

readily define real-valued intervals $F_i = [\underline{f_i}, \overline{f_i}]$ which encapsulate both the approximate value $\tilde{f}_i$ and the exact (but unknown) value f_i. In the sequel we work with

$$\underline{f_i} = \tilde{f}_i - \epsilon(1 + |\tilde{f}_i|), \qquad \overline{f_i} = \tilde{f}_i - \epsilon(1 + |\tilde{f}_i|),$$

to provide an easy transition between absolute errors in case of small values and relative errors in case of regular values.

Rather than approximating the point-values $\tilde{f}_i$ we look for a rational function $r_{\ell,m}(u, w)$ which satisfies the interval interpolation conditions

$$r_{\ell,m}(u_i, w_i) \in F_i \quad \Leftrightarrow \quad \underline{f_i} \leq r_{\ell,m}(u_i, w_i) \leq \overline{f_i}, \qquad i = 0, \ldots, s, \tag{14}$$

and this for the smallest possible $\ell + m$ with $\ell + m \ll s$. For fixed ℓ and m and provided that $q_{\ell,m}(u_i, w_i) > 0$, it follows that the numerator $p(u, w)$ and denominator $q(u, w)$ of $r_{\ell,m}(u, w)$ have to satisfy the linear inequalities

$$\underline{f_i}\, q_{\ell,m}(u_i, w_i) \leq p_{\ell,m}(u_i, w_i) \leq \overline{f_i}\, q_{\ell,m}(u_i, w_i), \qquad i = 0, .., s. \tag{15}$$

Denote the vector of unknown coefficients by

$$\mathbf{c} = (p_0, \ldots, p_\ell, q_0, \ldots, q_m)^T \in \mathbb{R}^{\ell+m+2}$$

and denote by $\mathbf{A}$ the $(2s+2) \times (\ell+m+2)$ constraint matrix implied by the inequalities (15). In order to obtain a nontrivial vector $\mathbf{c} \neq \mathbf{0}$ which strictly satisfies the component wise inequalities $\mathbf{Ac} \leq 0$, we propose the computation of a Chebyshev direction [12] of the corresponding unbounded polyhedral cone described by $\mathbf{Ac} \leq 0$ by solving the strictly convex quadratic programming (QP) problem:

$$\begin{aligned} &\arg \min_{\mathbf{c} \in \mathbb{R}^{\ell+m+2}} \left(\|\mathbf{c}\|_2\right)^2 \\ &\text{subject to } \mathbf{A}_j \mathbf{c} \leq -\delta \left\|\mathbf{A}_j\right\|_2, \qquad j = 1, \ldots, 2s+2. \end{aligned} \tag{16}$$

Here $\delta > 0$ is an arbitrary robustness margin, $\mathbf{A}_j$ denotes the j-th row of the matrix $\mathbf{A}$ and $\|\cdot\|_2$ is the Euclidean norm. For an in-depth discussion of the geometrical interpretation of this QP formulation and possible alternatives, we refer to the forthcoming [5]. The above QP problem can for instance be solved using the freely available MATLAB interface `qpas` [13].

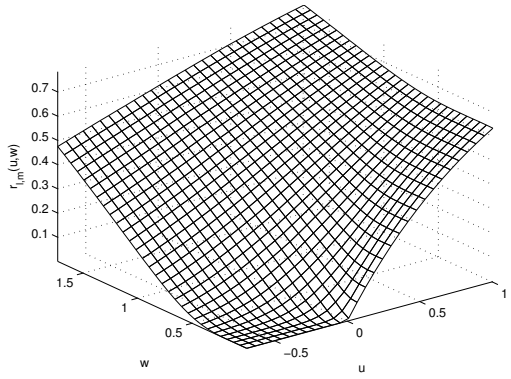
The smallest possible value $\ell + m$ is determined by solving the QP problem (16) for each combination of increasing $\ell + m = 0, 1, 2, \ldots$ until a feasible solution is found. Although modern QP solvers can efficiently handle several thousands of constraints, we propose a selection procedure among the given $s + 1$ data to reduce the total number of constraints and speed up this process.

4 Modelling large data clouds

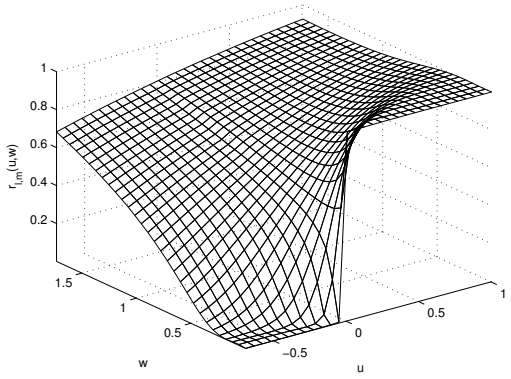
Remember that the total number $s + 1$ of given data can easily reach the hundreds of thousands. For instance, if for each of the 64 random triplets (E, r, σ) shown in Figure 1, a 200×100 (uniform or non-uniform) (S, t)-grid is constructed, a total of 1.28×10^6 5-dimensional data points is generated. Now let us transform these points $(S, t; E, r, \sigma)$ to 2-dimensional data (u, w), without any overlap occuring. Taking also into account the reduction to the domain of interest (8), the duality and the extension from negative to positive u, the number of data that remain in the ranges $-\ln(3/2) - 0.6 \leq u \leq 0$ and $0 \leq w \leq \sqrt{3}$ is $s + 1 = 292353$. For the interval widths, we allow a relative deviation in (13) of $\epsilon = 0.005$ for the Greeks and only $\epsilon = 0.001$ on the option price.

In order to further reduce the total number of inequalities in (16) we propose the following selection procedure. Of the given $s + 1$ data intervals, initially only a small number s_0 of intervals is selected (for instance, according to a Latin hypercube design) and a rational function $r_{\ell_0,m_0}(u, w)$ is computed that satisfies (14) for these s_0 data. These s_0 intervals are called the training data. Then it is checked how many of the original $s + 1$ interval interpolation conditions are automatically satisfied by $r_{\ell_0,m_0}(u, w)$ in addition to the s_0 imposed ones. Usually this is quite a lot more. These $s + 1 - s_0$ intervals are called the verification data.

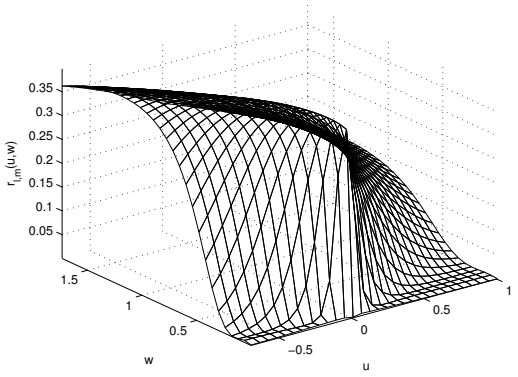
Among the violated interval verification data, we select $s_1 - s_0$ additional data points to compute $r_{\ell_1,m_1}(u, w)$ that satisfies (14) for these $s_0 + (s_1 - s_0) = s_1$ data. In other words, we update the set of training data. These $s_1 - s_0$ additional training data are placed where $r_{\ell_0,m_0}(u_i, w_i)$ deviates most from the given intervals F_i. Because the previous s_0 training data are obviously a subset of the new s_1 training data, necessarily $\ell_1 + m_1 \geq \ell_0 + m_0$. That is, the number of coefficients $\ell_1 + m_1 + 2$ needed to interpolate the updated s_1 training data cannot be less than the previously needed $\ell_0 + m_0 + 2$ coefficients. Hence the search for ℓ_1 and m_1 can be continued from the diagonal $\ell_0 + m_0$ rather than from scratch each time. With $r_{\ell_1,m_1}(u, w)$ we then check the new $s + 1 - s_1$ verification data again. And so on till the rational model satisfies all verification data.



(a) $c^-(u,w) \approx r^{c^-}_{9,13}(u,w)$, $c^+(u,w) \approx$ from (9)



(b) $\Delta^-(u,w) \approx r^{\Delta^-}_{13,8}(u,w)$, $\Delta^+(u,w) \approx$ from (10)



(c) $\kappa^-(u,w) \approx r^{\kappa^-}_{16,11}(u,w)$, $\kappa^+(u,w) \approx$ from (12)

Figure 4: Models based on approximations on the domain D^- and extended to D^+ by application of (9–12).

This procedure keeps the computational complexity as low as possible: the solution is not computed from imposing all $s+1$ interval data but rather from a carefully selected subset which then entails all interpolation conditions. Using this approach we easily achieve a reduction of $s+1$ by a factor of several thousands of the given large dataset.

Modelling $c(u,w), \Delta(u,w)$ and $\kappa(u,w)$. For the rational approximations shown in Figures 4(a) till 4(c), respectively only 74, 74 and 133 samples are used of the 292353 available ones. With these small training sets, rational interval interpolants for $c^-(u,w), \Delta^-(u,w)$

and $\kappa^-(u,w)$ of the form

$$r_{\ell,m}(u,w) = \frac{\sum_{(h,k)\in L} p_{hk} u^h w^k}{\sum_{(h,k)\in M} q_{hk} u^h w^k}, \qquad \#L = \ell + 1, \#M = m + 1$$

are computed with L and M respectively given by

$$\{(h,k) \in \mathbb{N} : 0 \le h + k \le 3\}, \quad \{(h,k) \in \mathbb{N} : 0 \le h + k \le 4\} \setminus \{(3,1)\}$$

for $r^{c^-}_{9,13}(u,w)$,

$$\{(h,k) \in \mathbb{N} : 0 \le h + k \le 4\} \setminus \{(3,1)\}, \quad \{(h,k) \in \mathbb{N} : 0 \le h + k \le 3\} \setminus \{(1,2)\}$$

for $r^{\Delta^-}_{13,8}(u,w)$ and

$$\{(h,k) \in \mathbb{N} : 0 \le h + k \le 4\} \cup \{(5,0),(0,5)\}, \quad \{(h,k) \in \mathbb{N} : 0 \le h + k \le 3\} \cup \{(4,0),(0,4)\}$$

for $r^{\kappa^-}_{16,11}(u,w)$. An analytic model as a function of the original variables $(S,t;E,r,\sigma)$ is obtained by substituting the change of variables (2) in the final function $r(u,w)$. The modelling of $\gamma(u,w)$ is discussed separately.

Modelling $\gamma(u,w)$. Let $N(x)$ denote the cumulative distribution function (CDF) of the standard normal distribution,

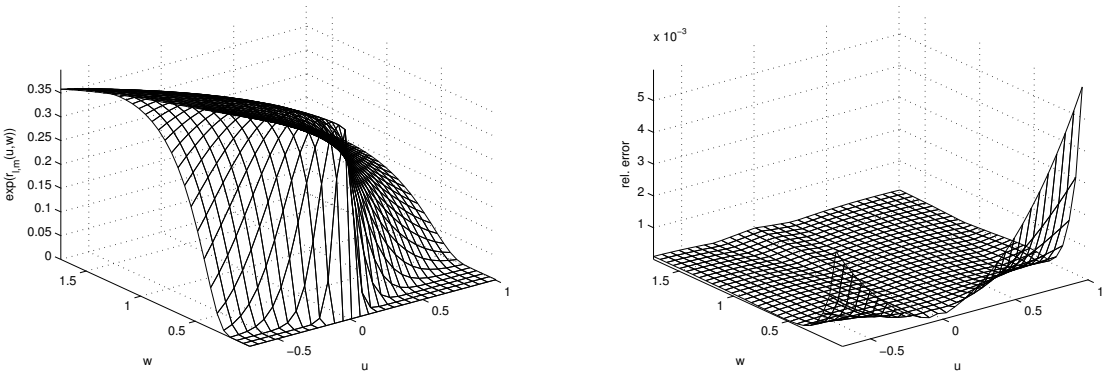
$$N(x) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right).$$

Then it is easy to verify that [7, p. 110]

$$c(u,w) = N(u/w + w/2) - e^{-u} N(u/w - w/2)$$

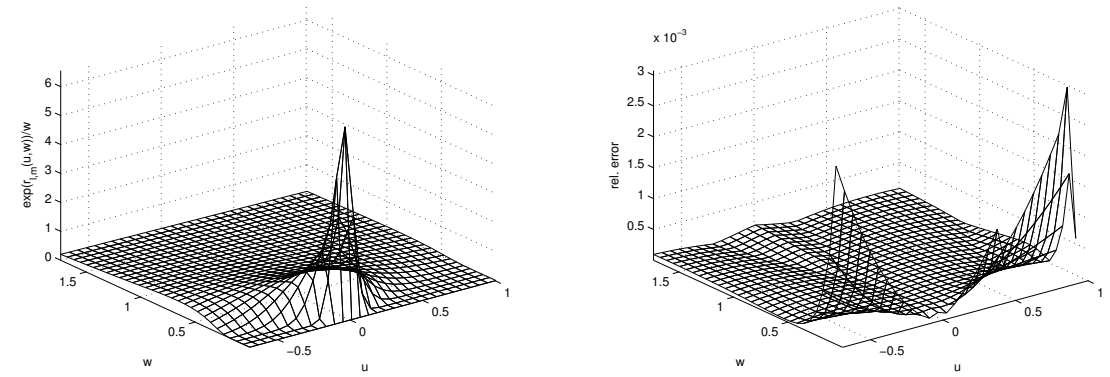
and that $\gamma(u,w) = \kappa(u,w)/w$. However, if we divide $r^{\kappa}_{16,11}(u,w)$ by w, we do not obtain a very good approximation for $\gamma(u,w)$, although the error $|r^{\kappa}_{16,11}(u,w) - \kappa(u,w)|/(1+|\kappa(u,w)|)$ is bounded overall by 0.005. This is a consequence of the fact that we need $r^{\gamma}_{\ell,m}(u,w)$ to equal zero when $w = 0$ except at the point $(u,w) = (0,0)$. To model such a singularity, $r^{\kappa}_{\ell,m}(u,w)$ need have the proper behaviour near $w = 0$: the approximation really needs to decrease very fast for $w \approx 0$, taking values of the order of 10^{-250} and smaller. We therefore follow another approach for the modelling of $\gamma(u,w)$. As explained earlier, fitting $\gamma(u,w)$ by differentiating an approximation for $c(u,w)$,

$$\gamma(u,w) = \frac{\partial c(u,w)}{\partial u} + \frac{\partial^2 c(u,w)}{\partial u^2},$$



(a) $\exp\left(r_{10,3}^{\log\kappa}(u,w)\right) \approx \kappa(u,w)$ (b) true relative error

Figure 5: Models based on fitting $\log(\kappa(u,w))$.



(a) $\exp\left(r_{10,3}^{\log\kappa}(u,w)\right)/w \approx \gamma(u,w)$ (b) true relative error

Figure 6: Approximation of $\kappa(u,w)/w = \gamma(u,w)$ based on fitting $\log(\kappa(u,w))$.

is not advised. We rather focus on $\log(\kappa(u,w))$.

We choose $\epsilon = 0.0001$ and compute $r_{10,3}^{\log\kappa^-}(u,w)$ with the numerator and denominator degree sets L and M respectively given by

$$L = \{(h,k) \in \mathbb{N} : 0 \le h+k \le 3\} \cup \{(0,4)\}, \quad M = \{(h,k) \in \mathbb{N} : 0 \le h+k \le 1\} \cup \{(0,2)\}.$$

Only 53 training data are required. In Figure 5 we show $\exp\left(r_{10,3}^{\log\kappa}(u,w)\right)$ as an approximation for $\kappa(u,w)$ and the true relative error, computed a posteriori. As before, the approximation is constructed on $u \le 0$ and extended to $u > 0$. The true relative error rises only

slightly above 0.005. The approximation

$$\frac{\exp\left(r_{10,3}^{\log\kappa}(u,w)\right)}{w} \approx \gamma(u,w)$$

is also very good. In Figure 6 we show the function and the true relative error, which is of course identical to the true relative error on $\kappa(u,w)$.

5 Modelling sparse data samples

A different challenge is to construct an analytic model $r_{\ell,m}(u,w)$ from only a minimal number of data. We now describe the required adaptations to the method of Section 4 when no large number of verification data is available. In addition we omit to make use of the dualities (9) and (10 – 12), because we want to illustrate use of the modelling technique when not too much about the underlying function is known. We still do restrict our analysis to the region of interest given in (8).

So we start the modelling process with the computation of $r_{\ell_0,m_0}(u,w)$ from a small number of data and we want to refine this model iteratively by sampling as few additional values as possible. How do we proceed? We discuss the matter in terms of the original variables $(S,t;E,r,\sigma)$ instead of the transformed variables (u,w). The latter will still be used in the computation of the analytic model though. For simplicity, we store the transformation between the discrete grid of datapoints $(S,t;E,r,\sigma)$ obtained from the PDE solver and the discrete set of interpolation points (u,w) used in the models $r_{\ell,m}(u,w)$ in a table (for the randomly generated (E,r,σ) given in Figure 1 and uniform 80×80 (S,t) grids the transformation was bijective).

So assume you only have at your disposal the PDE solution at an 80×80 (uniform or non-uniform) (S,t)-grid for 4 (E,r,σ) triplets. In a first step we model these 25600 data intervals by a rational function $r_{\ell_0,m_0}(u,w)$ following the algorithm described in the previous section. As before, all available data are divided into interpolation points (training data) and verification points (verification data). But different from Section 4 is that in the end, we want the model to be (sufficiently) accurate on the 80×80 (S,t) grids associated with each of the 64 randomly selected (E,r,σ) triplets given in Figure 1, while we only want to collect data on 80×80 (S,t) grids for a limited number of (E,r,σ) tuples.

So when iterating, several things need to be taken into consideration:

- When having a data cloud of reference (or validation) material besides the interpolation conditions, the (absolute or relative) model error at these non-interpolation points is easily obtained. Without this reference material at our disposition, a model error can

only be estimated from

$$\epsilon_i(u,w) := \frac{\left|r_{\ell_i,m_i}(u,w) - r_{\ell_{i-1},m_{i-1}}(u,w)\right|}{1+|r_{\ell_i,m_i}(u,w)|}, \qquad (u,w) \in (S,t;E,r,\sigma)\text{-grid of size } 80\times 80\times 64$$

where r_{ℓ_i,m_i} and $r_{\ell_{i-1},m_{i-1}}$ respectively denote the analytic models computed from s_i+1 and $s_{i-1}+1$ training data, in other words the models from the last and the one but last set of training data. The denominator in $\epsilon_i(u,w)$ is again chosen to provide a smooth transition from relative to absolute error in the case of very small function or model values.

- The value $\max_{(u,w)} \epsilon_i(u,w)$ is then checked against the threshold ϵ and here some care must be paid to avoid a false optimistic result. When r_{ℓ_i,m_i} is computed from only a slight update of the data that produced $r_{\ell_{i-1},m_{i-1}}$, then the two analytic models may be so similar that the error estimate $\epsilon_i(u,w)$ is small on the entire domain. As a first precaution we can check both

$$\begin{aligned} \max \epsilon_i^{(1)}(u,w) &\le \epsilon, \\ \max \epsilon_i^{(2)}(u,w) &\le \epsilon, \end{aligned} \tag{17}$$

where

$$\begin{aligned} \epsilon_i^{(1)}(u,w) &:= \frac{\left|r_{\ell_i,m_i}(u,w) - r_{\ell_{i-1},m_{i-1}}(u,w)\right|}{1+|r_{\ell_i,m_i}(u,w)|}, \\ \epsilon_i^{(2)}(u,w) &:= \frac{\left|r_{\ell_i,m_i}(u,w) - r_{\ell_{i-2},m_{i-2}}(u,w)\right|}{1+|r_{\ell_i,m_i}(u,w)|}. \end{aligned}$$

As a second precaution we steer any additional sample points moderately away from the already available data points as described below. This improves at the same time the conditioning of the interpolation problem, which is an added advantage.

- So when adding interpolation data we avoid samples at positions that are very close to previously sampled positions (u,w) because these usually add little or no extra information. Hence we stay some minimal distance δ away from every already sampled position (u,w), measured using one or other distance function, when collecting a new sample. On the one hand the value of δ should not be too small, while on the other hand it should not be so large that it leads us away from an interesting region. A strategy could be to adapt δ as the iteration continues, in other words, impose a step dependent distance δ_i (instead of a constant δ) between the locations of the s_i+1 already available data samples and the $s_{i+1}-s_i$ newly added data samples. Here δ_i can be decreased as the iteration proceeds.

- When adding interpolation points we make a distinction between choosing a new triplet of parameters (E, r, σ) and choosing some new points (S, t). A new triplet is only selected after all verification data associated with the last added 80×80 grid are satisfied. The reason for this is that for each separate triplet (E, r, σ) the PDE solver usually returns a full (uniform or non-uniform) grid of data points in the remaining variables S and t. So, inspecting $\arg\max \epsilon_i^{(1)}(u, w)$ or $\arg\max \epsilon_i^{(2)}(u, w)$, in that order and depending on which one of both violates (17), identifies a new and different triplet $(E_{i+1}, r_{i+1}, \sigma_{i+1})$ for which the PDE solver is then called. Having obtained the solution at an additional grid of (S, t) values, the technique of the previous section is used to compute $r_{\ell_{i+1}, m_{i+1}}(u, w)$. By this we mean that the values obtained at the grid of (S, t) locations need not be added all at once to the interpolation conditions, but can for each new triplet $(E_{i+1}, r_{i+1}, \sigma_{i+1})$ be separated into training and verification data.

- Since $\epsilon_i^{(1)}(u, w)$ and $\epsilon_i^{(2)}(u, w)$ are only estimates of the error and not guaranteed upper bounds, we recommend for the stop criterion to divide the objective ϵ by some safety factor $\phi \geq 1$, in order to obtain a more reliable analytic model. This way one safeguards the technique against underestimating the hardness of the problem. The iteration is terminated only after

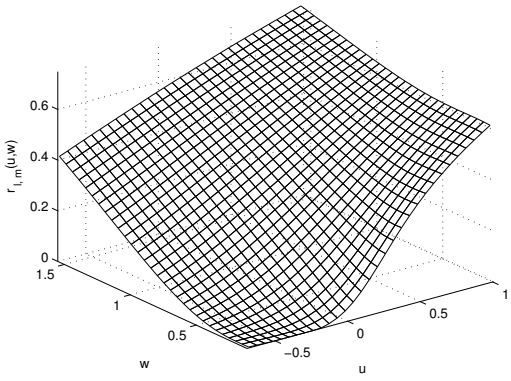
$$\begin{aligned} \max \epsilon_i^{(1)}(u, w) &\leq \epsilon/\phi, \\ \max \epsilon_i^{(2)}(u, w) &\leq \epsilon/\phi. \end{aligned} \tag{18}$$

  When expressing ϵ as $\epsilon = \beta^{-n}$, where usually $\beta = 2$ or 10, we found it useful to choose $\phi \geq \beta$.

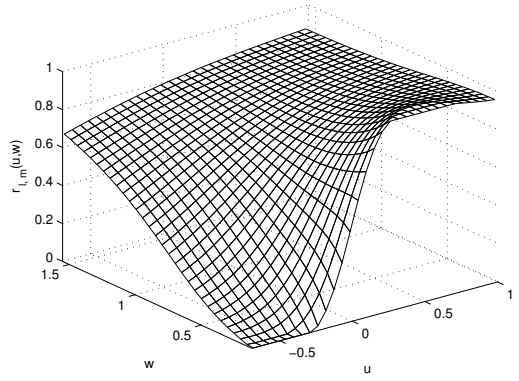
- When ϕ in (18) is too large, then it may happen that the iteration doesn't stop, while (17) is satisfied. As a result additional (E, r, σ) tuples are added at $\arg\max \epsilon_i^{(1)}(u, w)$ or $\arg\max \epsilon_i^{(2)}(u, w)$ and data on (S, t) grids are generated without much use. It may be that the approximate model $r_{\ell_i, m_i}(u, w)$ is accurate enough and automatically satisfies all added interval data. When a whole grid of 6400 data is generated that does not lead to a single additional interpolation point and an update of the model, we suggest to decrease ϕ as the iteration continues, until $\phi = 1$. One runs into this problem for instance, when choosing ϕ for $\kappa(u, w)$ too large.

We illustrate the sparse adaptive sampling technique and compare with the results of Section 4. As mentioned, for each of 4 start triplets (E, r, σ) data are collected at 6400 (S, t) locations. This adds up to a total of 25600 samples. We target a relative deviation of $\epsilon = 0.05$ for the Greeks and $\epsilon = 0.005$ for the price $c(u, w)$. We further choose $\phi = 50$ for all functions except $\kappa(u, w)$ where we take $\phi = 10$. A uniform choice of $\delta = 0.01$ worked well.

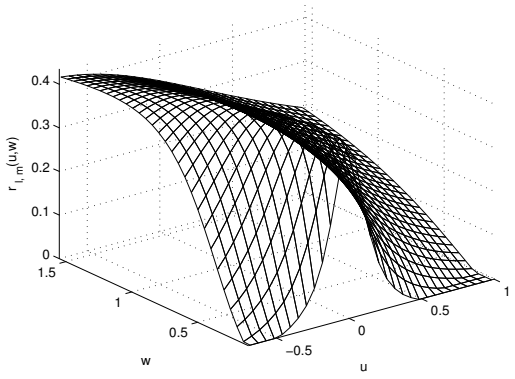
The first 25600 interval data are all interpolated from only 101 samples for $c(u, w)$, 118 samples for $\Delta(u, w)$ and 88 samples for $\kappa(u, w)$. Now that the analytic model passes through



(a) $c(u,w) \approx r^{c}_{11,12}(u,w)$



(b) $\Delta(u,w) \approx r^{\Delta}_{7,7}(u,w)$



(c) $\kappa(u,w) \approx r^{\kappa}_{4,7}(u,w)$

Figure 7: Models based on sparse interval interpolations.

all 25600 initial data intervals, we can add a new (E,r,σ) triplet and call the PDE solver to deliver a new grid of solutions at 6400 (S,t) tuples. The location of the new (E,r,σ) triplet is decided from inspecting $\arg\max_{(u,w)} \epsilon_i^{(1)}(u,w)$, or $\arg\max_{(u,w)} \epsilon_i^{(2)}(u,w)$, and checking the transformation table for the 5-variable vector that is associated with the (u,w) location where a maximum is attained.

Continuing in this way we ultimately obtain the models graphed in Figure 7. The numerator and denominator index sets L and M of the graphed models are respecively given by

$$\{(h,k) \in \mathbb{N} : 0 \le h+k \le 3\} \cup \{(4,0),(0,4)\}, \quad \{(h,k) \in \mathbb{N} : 0 \le h+k \le 4\} \setminus \{(1,3),(2,2)\}$$

for $r^{c}_{11,12}(u,w)$,

$\{(h,k) \in \mathbb{N} : 0 \le h+k \le 2\} \cup \{(3,0),(0,3)\}, \quad \{(h,k) \in \mathbb{N} : 0 \le h+k \le 2\} \cup \{(3,0),(0,3)\}$

for $r^{\Delta}_{7,7}(u,w)$ and

$$\{(h,k) \in \mathbb{N} : 0 \le h+k \le 2\} \setminus \{(1,1)\}, \quad \{(h,k) \in \mathbb{N} : 0 \le h+k \le 2\} \cup \{(3,0),(0,3)\}$$

for $r^{\kappa}_{4,7}(u,w)$. The rational approximations are computed from only 17 (E,r,σ) and a total of 145 (u,w) interpolation points for c, 8 (E,r,σ) and 132 (u,w) interpolation points for Δ and 4 (E,r,σ) and 88 (u,w) interpolation points for κ.

Since the Black-Scholes PDE serves as a benchmark here, and since we know the explicit expressions for $C(S,t)$ and the Greeks in terms of the parameters E, r and σ, we can compute (a posteriori) the true overall error (following the formula used in the definition of ϵ_i in order to accomodate both small and regular values):

- for $r^{c}_{11,12}(u,w)$ it equals 0.0053,
- for $r^{\Delta}_{7,7}(u,w)$ it is 0.0499,
- and for $r^{\kappa}_{4,7}(u,w)$ we obtain 0.0476.

When changing the start data for the computation of the rational models, the results differ of course. But from the many, many runs that we have performed, we are showing results that are very representative. For instance, $\kappa(u,w)$ and $\Delta(u,w)$ were easier to model than $c(u,w)$ requiring less (E,r,σ) tupels, meaning less calls to the PDE solver to generate new data, and less interpolation points overall (of course the accuracy conditions on the Greeks are more relaxed).

6 Conclusions

We describe how to compute analytic models for option prices and Greeks, both from a large data cloud (in Section 4) and from a smaller set of adaptively collected data (in Section 5). We employ rational interval interpolation and, as a prototype case, consider the reputed Black-Scholes PDE for European call options. The rational interval interpolation technique is especially useful for capturing steep increases or decreases, which are common in financial applications, such as in the Greeks.

A relative accuracy of 1‰ for $c(u,w)$ and 5‰ for the functions $\Delta(u,w)$ and $\kappa(u,w)$ is easy to guarantee when working with a large data set. In case of a selective smaller data set we easily achieve respectively 5‰ and 5%. The rational interval interpolation technique can in principle be extended to more advanced option pricing models in finance, with more variables and parameters, for which no exact solutions in semi-closed form are available. If desired, the quadratic programming problem (16) can also be extended with positivity conditions for the denominator, as in [4], or monotonicity, as in [11].

References

[1] I. Barrodale, M. J. D. Powell, and F. D. K. Roberts. The differential correction algorithm for rational ℓ_∞-approximation. *SIAM Journal on Numerical Analysis*, 9(3):pp. 493–504, 1972.

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

[3] B.W. Boehm. *Existence, characterization, and convergence of best rational Tchebycheff approximations*. R (Rand Corporation). Rand, 1964.

[4] A. Cuyt, O. Salazar Celis, and M. Lukach. Multidimensional IIR filters and robust rational interpolation. *Multidimensional Systems and Signal Processing*, 2012. to appear.

[5] Annie Cuyt and Oliver Salazar Celis. Multivariate data fitting with error control. In preparation.

[6] Ernst Eberlein, Antonis Papapantoleon, and Albert N. Shiryaev. On the duality principle in option pricing: semimartingale setting. *Finance and Stochastics*, 12(2):265–292, 2008.

[7] Desmond J. Higham. *An introduction to financial option valuation: mathematics, stochastics and computation*. Cambridge Univ. Press, 2008.

[8] Minqiang Li. Approximate inversion of the Black-Scholes formula using rational functions. *European Journal of Operational Research*, 185(2):743–759, March 2008.

[9] Per Lötstedt, Jonas Persson, Lina von Sydow, and Johan Tysk. Space-time adaptive finite difference method for european multi-asset options. *Computers and Mathematics with Applications*, 53:1159–1180, 2007.

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

[11] Romain Pacanowski, Oliver Salazar Celis, Christophe Schlick, Xavier Granier, Pierre Poulin, and Annie Cuyt. Rational BRDF. *IEEE Transactions on Visualization and Computer Graphics*, 18:1824–1835, 2012.

[12] Oliver Salazar Celis, Annie Cuyt, and Brigitte Verdonk. Rational approximation of vertical segments. *Numerical Algorithms*, 45:375–388, 2007.

[13] Adrian Wills. Quadratic programming in C - MATLAB interface. http://sigpromu.org/quadprog/.

[14] Jerry M. Wolfe. Discrete rational l_p approximation. *Mathematics of Computation*, 29(130):pp. 540–548, 1975.