BIT

CrossMark

Multivariate data fitting with error control

Annie Cuyt[1] · Oliver Salazar Celis[1]

Abstract
We show how a recently developed multivariate data fitting technique enables to solve a variety of scientific computing problems in filtering, queueing, networks, metamodelling, computational finance, graphics, and more. We can capture linear as well as nonlinear phenomena because the method uses a generalized multivariate rational model. The technique is a refinement of the basic ideas developed in Salazar et al. (Numer Algorithms 45:375–388, 2007. https://doi.org/10.1007/s11075-007-9077-3) and interpolates interval data. Intervals allow to take the inherent data error in measurements and simulation into consideration, whilst guaranteeing an upper bound on the tolerated range of uncertainty. The latter is the main difference with a best approximation or least squares technique which does as well as it can, but without respecting an a priori imposed threshold on the approximation error. Compared to the best approximations, the interval interpolant is relatively easy to compute. In applications where industry standards need to be guaranteed, the interval interpolation technique may be a valuable alternative.

Keywords Interval interpolation · Multivariate · Computational science · Rational function

Mathematics Subject Classification 41A20 · 65D05 · 65G20

✉ Annie Cuyt
  annie.cuyt@uantwerpen.be

  Oliver Salazar Celis
  oliver.salazarcelis@uantwerpen.be

[1] Department of Mathematics and Computer Science, University of Antwerp (CMI), Middelheimlaan 1, 2020 Antwerp, Belgium

 Springer

1 Multivariate interval interpolation

We work in the d-dimensional space $\mathbb{R}^d$ and denote the d-tuple of variables by $v = (x_1, \ldots, x_d)$. Let $\alpha = (\alpha_1, \ldots, \alpha_d)$ denote a multi-index in $\mathbb{N}^d$. Using multi-index notation, a generalized polynomial is expressed as

$$p(v) = \sum_{\alpha \in N} a_\alpha g_\alpha(v), \quad N \subset \mathbb{N}^d$$

where N is a finite set. Easy choices for $g_\alpha(v)$ are the monomials $g_\alpha(v) = x_1^{\alpha_1} \cdots x_d^{\alpha_d}$ or a family of multivariate orthogonal polynomials, but also trigonometic polynomials and more general functions are possible. We decide upon an ordering $\prec$ of the multi-indices that respects the rule

$$\forall \alpha \in [0, \beta_1] \times \ldots \times [0, \beta_d] \backslash \{\beta\} : \alpha \prec \beta, \qquad \beta = (\beta_1, \ldots, \beta_d).$$

The problem statement now is the following. Given are real-valued intervals $F_i = [f_{i1}, f_{i2}]$ at $s + 1$ datapoints $v_i, i = 0 \ldots, s$. We want to find a generalized rational function

$$r_{N,D}(v) = \frac{p_{N,D}(v)}{q_{N,D}(v)} = \frac{\sum_{\alpha \in N} a_\alpha g_\alpha(v)}{\sum_{\alpha \in D} b_\alpha g_\alpha(v)}, \quad N, D \subset \mathbb{N}^d$$

that satisfies

$$r_{N,D}(v_i) \in F_i, \quad i = 0, \ldots, s, \tag{1.1}$$

where the sum of the cardinalities of N and D is at most $s + 2$. One of the coefficients a_α or b_α is determined by a normalization of $r_{N,D}(v)$ and hence the $s + 1$ given datapoints are sufficient to determine the remaining coefficients a_α and b_α, even in the case of point intervals F_i where $f_{i1} = f_{i2}$. When $f_{i1} \neq f_{i2}$, as is mostly the case when working with measured or simulated data, then often

$$\#N + \#D \ll s + 2.$$

Rational functions are not only adequate in the neighbourhood of singularities or to model asymptotic behaviour. They offer the advantage that they can better model steep gradients and large curvature.

In practice, N and D, and consequently $\#N$ and $\#D$, are determined iteratively. Of the given $s + 1$ data intervals, a small number $s_0 + 1$ of intervals is selected and $r_{N_0,D_0}^{(0)}(v)$ is computed that satisfies (1.1) for these $s_0 + 1$ data. These $s_0 + 1$ intervals are called the training data. Then it is checked how many of the original $s + 1$ interval interpolation conditions are unintentionally satisfied by $r_{N_0,D_0}^{(0)}(v)$ besides the $s_0 + 1$ imposed ones. Usually this is quite a lot more. For illustrations of this we refer to the case studies in the subsequent sections. These $s - s_0$ intervals are called the verification data. Among the interval interpolation bounds violated by $r_{N_0,D_0}^{(0)}(v)$, we select $s_1 - s_0$ additional data points for the computation of $r_{N_1,D_1}^{(1)}(v)$ satifying (1.1) for these $s_1 + 1$ data, with $N_0 \subset N_1, D_0 \subset D_1$. In other words, we update the set of training data

and the rational interval interpolant. These $s_1 - s_0$ additional training data are placed where $r^{(0)}_{N_0, D_0}(v_i)$ deviates most from the given intervals F_i. With $r^{(1)}_{N_1, D_1}(v)$ we then check the remaining $s - s_1$ verification data again. And so on till the rational model satisfies all verification data.

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 in the end entails interval interpolation at all given points. Using this approach we easily achieve in some of the case studies, a reduction of $s+1$ by a factor of several thousands when large datasets were modeled.

Let N and D satisfy the inclusion property, by which we mean that

$$\alpha \in N \Rightarrow [0, \alpha_1] \times \cdots \times [0, \alpha_d] \cap \mathbb{N}^d \subset N$$

and likewise for D. We are actually interested in the smallest $\#N + \#D$ for which (1.1) holds. Therefore, let $\alpha^{(i)}$ denote the i-th element in the ordering of the multi-indices, where $\alpha^{(0)} = (0, \ldots, 0)$ and let $\#N = n + 1$ and $\#D = m + 1$. Then our notation for $g_\alpha(v), a_\alpha, b_\alpha, p_{N,D}(v), q_{N,D}(v)$ and $r_{N,D}(v)$ can be simplified to

$$r_{n,m}(v) = \frac{p_{n,m}(v)}{q_{n,m}(v)} = \frac{\sum_{i=0}^{n} a_i g_i(v)}{\sum_{i=0}^{m} b_i g_i(v)}$$

where $g_i(v) = g_{\alpha^{(i)}}(v)$. So, given the basis functions $g_i(v), i \geq 0$ and the intervals $F_i, i = 0, \ldots, s$ we want to find the smallest $n + m$ for which (1.1) holds. In addition, we assume that $q_{n,m}(v_i) > 0$ and we linearize the interpolation conditions (1.1) to

$$f_{i1} q_{n,m}(v_i) \leq p_{n,m}(v_i) \leq f_{i2} q_{n,m}(v_i). \tag{1.2}$$

Whether for specific n and m a solution of (1.2) exists, depends on the relation between the width of the intervals F_i and the values for n, m and s. As mentioned before $n+m \leq s$ for any width of the given F_i. Let us denote the set of all (possibly) reducible representations $r_{n,m}(v)$ of solutions $p_{n,m}(v)$ and $q_{n,m}(v)$ of (1.2), by $R_{n,m}(v)$. In Sect. 3 we find out when $R_{n,m}(v)$ is empty and which element to select when it contains multiple solutions that do not represent equivalent rational functions.

Making use of the chosen multi-index ordering $\prec$, we find the smallest $n + m$ by starting with $n+m = 0$ and increasing $n+m$ by 1 at every step. This way we build up a table of generalized rational interval interpolation problems diagonal by diagonal:

$$
\begin{array}{llll}
R_{0,0}(v) & R_{0,1}(v) & R_{0,2}(v) & \ldots \\
\\
R_{1,0}(v) & R_{1,1}(v) & R_{1,2}(v) & \ldots \\
\\
R_{2,0}(v) & R_{2,1}(v) & \ddots \\
\\
\vdots & \vdots
\end{array}
$$

Along some k-th diagonal in the table all of the $R_{n,m}(v)$, $n + m = k$ may be empty, while along another ℓ-th diagonal there may be multiple $R_{n,m}(v)$, $n + m = \ell > k$ that are nonempty. In the latter case, without the availability of additional information, we are inclined to prefer the rational interval interpolants with $n \approx m$ [3,6,10].

In the rational interpolation of point data where $f_{i1} = f_{i2}$ so-called unattainable interpolation points can occur, meaning points at which $q_{n,m}(v_i) = 0$ and consequently, because of the linearized interpolation conditions, also $p_{n,m}(v_i) = 0$. When $d = 1$ the result is a common factor in $p_{n,m}(v)$ and $q_{n,m}(v)$ which is simplified in $r_{n,m}(v)$ leaving $r_{n,m}(v_i)$ unrelated to F_i [15]. This phenomenon has also been described in the multivariate rational interpolation of point data in [2]. Fortunately, as we shall see, unattainability does not occur when the F_i are true intervals with $f_{i1} < f_{i2}$. And in case of point intervals F_i, going through the table of interpolation problems diagonal by diagonal, one encounters irreducible rational interpolants before their reducible equivalents [1].

We remark that, since $N_i \supset N_{i-1}$ and $D_i \supset D_{i-1}$ in $r_{N_i, D_i}^{(i)}(v)$, the update from $i - 1$ to i only has to start at diagonal $(\#N_{i-1} - 1) + (\#D_{i-1} - 1)$ with $R_{0,0}(v)$ belonging to diagonal zero.

For given F_i, $i = 0, \ldots, s$ we denote the solution set for the vector of coefficients $(a_0, \ldots, a_n, b_0, \ldots, b_m)^t$ satisfying (1.2) by $L_{n,m}(F_0, \ldots, F_s)$. Geometrically this set is the intersection of a finite number of closed half spaces of which the boundary hyperplanes pass through the origin. Let $c = (c_0, \ldots, c_{n+m+1})^t = (a_0, \ldots, a_n, b_0, \ldots, b_m)^t$ and

$$
A_{n,m} = \begin{pmatrix}
g_0(v_0) & \ldots & g_n(v_0) & -f_{01}g_0(v_0) & \ldots & -f_{01}g_m(v_0) \\
\vdots & & \vdots & \vdots & & \vdots \\
g_0(v_s) & \ldots & g_n(v_s) & -f_{s1}g_0(v_s) & \ldots & -f_{s1}g_m(v_s) \\
-g_0(v_0) & \ldots & -g_n(v_0) & f_{02}g_0(v_0) & \ldots & f_{02}g_m(v_0) \\
\vdots & & \vdots & \vdots & & \vdots \\
-g_0(v_s) & \ldots & -g_n(v_s) & f_{s2}g_0(v_s) & \ldots & f_{s2}g_m(v_s)
\end{pmatrix}. \quad (1.3)
$$

So the conditions (1.2) are equivalent with the componentwise inequalities $A_{n,m}c \geq 0$. If the matrix $A_{n,m}$ has full rank, then $L_{n,m}(F_0, \ldots, F_s)$ is a pointed polyhedral convex cone with apex in the origin [7]. This polyhedron is full-dimensional, meaning that it has nonempty interior, if and only if it contains a point c for which the inequalities (1.2) are strict, meaning $A_{n,m}c > 0$. Solutions in the interior of $L_{n,m}(F_0, \ldots, F_s)$ do not suffer from unattainable interpolation points, because then the inequalities in (1.2) are strict and consequently $q_{n,m}(v_i) \neq 0$.

A possible normalization for $r_{n,m}(v)$ is

$$
\|c\|_\infty = \max(|a_0|, \ldots, |a_n|, |b_0|, \ldots, |b_m|) = 1.
$$

Together with the requirement that $q_{n,m}(v_i) > 0$ this normalization yields a unique representation for each $r_{n,m}(v) \in R_{n,m}(v)$. Adding the normalization to the interpolation conditions implies that we look at the intersection of $L_{n,m}(F_0, \ldots, F_s)$ with the

faces of the $(n + m + 2)$-dimensional hypercube $[-1, 1]^{n+m+2}$. However, it is easier to relax the normalization to

$$||c||_\infty = \max \left(|a_0|, \ldots, |a_n|, |b_0|, \ldots, |b_m|\right) \leq 1$$

and study the intersection of the set $L_{n,m}(F_0, \ldots, F_s)$ with the full hypercube $[-1, 1]^{n+m+2}$, despite the fact that the solution set now contains redundant representations of $r_{n,m}(v)$. Another possible normalization is the choice $b_0 = 1$ in combination with scaling the variables to a region that contains the origin. This choice reduces the dimensionality of the problem and avoids the redundancy issue. However, it may leave the thus normalized solution set for the coefficient vector c unbounded which is undesirable. We illustrate that in Sect. 2.

For the sake of robustness we are only interested in interior points of the set $L_{n,m}(F_0, \ldots, F_s)$. We do not want to return a solution for c that violates (1.2) with the slightest perturbation in the coefficients. In Sect. 3 we also indicate how the most robust solution can be computed. First, in Sect. 2, we compare the interval interpolant to the more traditional best approximant. For the latter we consider the best discrete rational approximant in the Euclidean (ℓ_2) as well as the Chebyshev (ℓ_∞) norm.

2 Interval interpolation versus best approximation

We need to find the answer to the question whether for chosen n and m the set $L_{n,m}(F_0, \ldots, F_s)$ is nonempty and then return an interior point of the intersection of $L_{n,m}(F_0, \ldots, F_s)$ with one or other normalization, assuming that it is full-dimensional. For simplicity, let for a moment the interval widths $f_{i2} - f_{i1}$ of F_i be a constant 2ϵ for all $i = 0, \ldots, s$. In general this is rather the exception than the rule though.

Before we answer the first question, we introduce $r_{n,m,\infty}(v) = p_{n,m,\infty}(v)/q_{n,m,\infty}(v)$, satisfying

$$r_{n,m,\infty}(v) = \arg \min_{r_{n,m}(v)} \left(\max_{i=0,\ldots,s} \left| \frac{f_{i2} + f_{i1}}{2} - r_{n,m}(v_i) \right| \right), \quad q_{n,m,\infty}(v_i) > 0.$$

Here we assume, for simplicity, that $r_{n,m,\infty}(v)$ exists, which is not guaranteed [4] because with $q_{n,m,\infty}(v_i) > 0$ the set is not compact. It is clear that $L_{n,m}(F_0, \ldots, F_s)$ is empty if

$$\epsilon < \max_{i=0,\ldots,s} \left| \frac{f_{i2} + f_{i1}}{2} - r_{n,m,\infty}(v_i) \right|,$$

because the interval interpolant $r_{n,m}(v)$ cannot approximate the values $(f_{i1} + f_{i2})/2$ better than the best approximant $r_{n,m,\infty}(v)$ in the Chebyshev sense. Of course this is not the way to determine the feasibility of the interval interpolation conditions for given n and m. The practical computation is discussed later: it will turn out that the generalized rational interval interpolant is much easier to determine than its best approximation counterparts. What we want to point out here is that the coefficient vector of the best discrete Chebyshev rational approximant $r_{n,m,\infty}(v)$ always belongs to a nonempty $L_{n,m}(F_0, \ldots, F_s)$.

On the other hand, the coefficient vector of the best discrete Euclidean rational approximant $r_{n,m,2}(v) = p_{n,m,2}(v)/q_{n,m,2}(v)$ satisfying

$$r_{n,m,2}(v) = \arg \min_{r_{n,m}(v)} \sum_{i=0}^{s} \left(\frac{f_{i2} + f_{i1}}{2} - r_{n,m}(v_i) \right)^2, \quad q_{n,m,2}(v_i) > 0,$$

if it exists, does not necessarily belong to $L_{n,m}(F_0, \ldots, F_s)$. It need not be unique either.

Remember that the conditions $q_{n,m}(v_i) \neq 0$ are automatically fulfilled for coefficient vectors in the interior of $L_{n,m}(F_0, \ldots, F_s)$. The most robust interior points of the set $L_{n,m}(F_0, \ldots, F_s) \cap [-1, 1]^{n+m+2}$ are the ones that maximize the distance to the complement

$$\mathbb{R}^{n+m+2} \backslash \left(L_{n,m}(F_0, \ldots, F_s) \cap [-1, 1]^{n+m+2} \right).$$

Such points are called Chebyshev centers [5, p. 148] of $L_{n,m}(F_0, \ldots, F_s) \cap [-1, 1]^{n+m+2}$. A Chebyshev center need not be unique but in a nonempty full-dimensional polyhedron it always exists. In Sect. 3 we also indicate how the generalized rational interval interpolation problem can be tuned such that a unique solution is returned, although this is not a major problem.

First we perform a small experiment to illustrate the difference between $r_{n,m,\infty}(v)$ and the solution of (1.2) built using a Chebyshev center. The latter is denoted by $r_{n,m,\diamond}(v)$. Also, we illustrate that the normalization $b_0 = 1$ may leave the solution set unbounded, in which case a Chebyshev center cannot be returned.

Consider the one-dimensional function $f(v) = (1/2)/(1 + v)$. We sample it in $v_0 = -1/2, v_1 = 0, v_2 = 1/2$ and add some uncertainty to the sampled values to obtain $F_0 = [1 - \epsilon_0, 1 + \epsilon_0], F_1 = [1/2 - \epsilon_1, 1/2 + \epsilon_1], F_2 = [1/3 - \epsilon_2, 1/3 + \epsilon_2]$. Then we want to find a rational function of the form

$$r_{0,1}(v) = \frac{a_0}{b_0 + b_1 v}, \quad (c_0, c_1, c_2) = (a_0, b_0, b_1),$$

that interpolates these intervals, in other words reconstructs the given $f(v)$ up to the added uncertainty.

We first choose the normalization $b_0 = 1$. The properly normalized best discrete Chebyshev rational approximant is $r_{0,1,\infty}(v) = f(v)$, the function itself. For this rational function the residues with respect to the midpoint of the intervals are zero. Hence for whatever values of $\epsilon_i, i = 0, 1, 2$ the coefficient vector $(1/2, 1, 1)$ belongs to $L_{0,1}(F_0, \ldots, F_s) \cap \{c_1 = b_0 = 1\}$. With the normalization $c_1 = b_0 = 1$ the solutions for $c_0 = a_0$ and $c_2 = b_1$ from (1.2) lie in the polygon defined by:

- 2 straight lines through the point $(0, 2)$ with respective slopes $-2/(1 + \epsilon_0)$ and $-2/(1 - \epsilon_0)$,
- 2 vertical lines through $1/2 - \epsilon_1$ and $1/2 + \epsilon_1$,
- 2 straight lines through the point $(0, -2)$ with respective slopes $6/(1 + 3\epsilon_2)$ and $6/(1 - 3\epsilon_2)$.

Fig. 1 Feasible region for (1.2) with $s = 2$, $\epsilon_i = 0.4$, $i = 0, 1, 2$ and $c_1 = b_0 = 1$
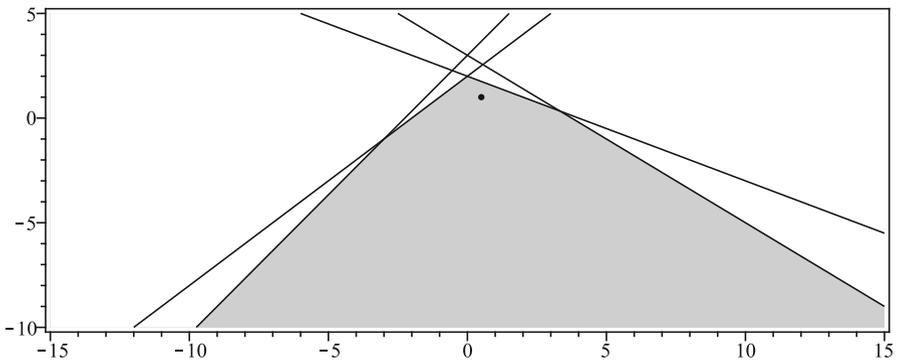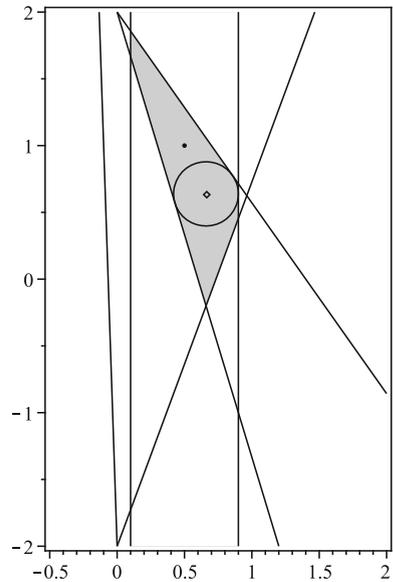

Fig. 2 Feasible region for (1.2) with $s = 1$, $\epsilon_0 = \epsilon_1 = 3$ and $c_1 = b_0 = 1$

This region is shown in Fig. 1. The coefficient vector for $r_{0,1,\infty}(v)$ is marked with ● and the coefficient vector for $r_{0,1,\diamond}(v)$ with ⬦. Note that as the ϵ_i increase, the Chebyshev center moves away from the best discrete Chebyshev approximant, as it tries to maximize the robustness. As the ϵ_i shrink the Chebyshev center moves toward the best discrete Chebyshev approximant.

Now let us change the interpolation conditions. We take $s = 1$, $v_0 = -1/2$ and $v_1 = -1/3$, which is the minimal number of datapoints to consider when $n = 0, m = 1$. The given intervals are $F_0 = [1 - \epsilon_0, 1 + \epsilon_0]$, $F_1 = [3/4 - \epsilon_1, 3/4 + \epsilon_1]$. Then with the normalization $b_0 = 1$ the solution set for (a_0, b_1) is unbounded as shown in Fig. 2. Let us switch to the alternative normalization $\max_{i=0,1,2} |c_i| \leq 1$. The set $L_{0,1}(F_0, \ldots, F_s) \cap [-1, 1]^3$ then contains multiple representations of the same

rational functions, but this need not be a problem. Scalar multiples of the vector $(1/2, 1, 1)$ deliver $r_{0,1,\infty}(v)$, and multiples of $c_\diamond = (0.1799777\ldots, 1, 1)$, which is the Chebyshev center of $L_{0,1}(F_0, \ldots, F_s) \cap [-1, 1]^3$, deliver $r_{0,1,\diamond}(v)$. So the appropriate normalization for use with (1.2) is

$$\max_{i=0,\ldots,n+m+1} |c_i| \leq 1$$

and not $b_0 = 1$. In the next section we discuss the computation of a Chebyshev center of $L_{n,m}(F_0, \ldots, F_s)$ which is the coefficient vector of $r_{n,m,\diamond}(v)$, a most robust solution of (1.2).

3 Existence, uniqueness and computational issues

The center c of a largest Euclidean ball in the polyhedron $L_{n,m}(F_0, \ldots, F_s) \cap [-1, 1]^{n+m+2}$, described by the linear inequalities (1.2) combined with the normalization $\max_{i=0,\ldots,n+m+1} |c_i| \leq 1$, is called a Chebyshev center of the polyhedron. Note that a Chebyshev center need not be unique. We write the points in the ball as $c + u$ with $||u||_2 \leq \rho$ and so we maximize ρ subject to the constraint that the Euclidean ball remains in the polyhedron. Let $A_{n,m}^{(j)}$ denote the j-th row of the matrix $A_{n,m}$ given by (1.3). The constraint that the Euclidean ball lies in the halfspace $A_{n,m}^{(j)} c \geq 0$ implies that

$$A_{n,m}^{(j)}(c + u) \geq 0, \quad ||u||_2 \leq \rho$$

or

$$A_{n,m}^{(j)} c - \rho ||A_{n,m}^{(j)}||_2 \geq 0, \quad j = 1, \ldots, 2s + 2. \tag{3.1}$$

Hence the constraint that the Euclidean ball lies in the halfspace is a linear inequality in c and ρ. Maximizing ρ under the linear constraints (3.1) for $j = 1, \ldots, 2s + 2$ is a well-understood linear programming problem.

When we find $\rho = 0$ then $L_{n,m}(F_0, \ldots, F_s)$ has empty interior. When also $c = 0$ then (1.1) has no solution. By definition a Chebyshev center is not to be found at the apex of the polyhedral cone. When a solution is returned with $\rho \neq 0$ then automatically $q_{n,m,\diamond}(v_i) \neq 0$ for the denominator $q_{n,m,\diamond}(v)$ built with the accompanying coefficient vector c.

Although a Chebyshev center need not be unique, it is always a global optimum. Also $r_{n,m,\infty}(v)$ delivered by the iterative differential correction algorithm which solves a sequence of linear programming problems, is a global optimum. This is in sharp contrast with the iterative algorithms that compute $r_{n,m,2}(v)$ which is the solution of a possibly nonconvex optimization problem. These may end up in a local optimum, depending on the provided starting point. Moreover, the global optimum $r_{n,m,2}(v)$ may also be nonunique.

So compared to the best approximations $r_{n,m,\infty}(v)$ and $r_{n,m,2}(v)$, the interval interpolant $r_{n,m,\diamond}(v)$ is relatively easy to compute: solving a single linear programming problem suffices. In addition, we have indicated how to check for the existence of a solution (numerical issues in this respect are discussed later in this section), which

poses a problem with the other approaches. Remains the problem of uniqueness for which we return to the question of the normalization of $r_{n,m}(v)$.

Non-uniqueness is not a real issue as long as the computed solution is a global optimum, which it is in case of a Chebyshev center providing the coefficients of $r_{n,m,\diamond}(v)$. Nevertheless uniqueness can be obtained. When looking at the unbounded polyhedral cone $L_{n,m}(F_0, \ldots, F_s)$, disregarding any normalization, a unique coefficient vector c in that cone can be identified as follows. The price one pays is the switch from a linear programming problem to a stricly convex quadratic programming problem. In the subsequent sections we have computed the generalized rational interval interpolants both with the LP approach and the QP approach and found no difference in the numerical outcome, as long as a Chebyshev center c and a nonzero radius ρ exist.

It is proved in [16] that if, for a chosen $\delta > 0$, the quadratic programming problem

$$\arg \min_{c \in \mathbb{R}^{n+m+2}} ||c||_2^2$$
$$A_{n,m}^{(j)}c - \delta ||A_{n,m}^{(j)}||_2 \geq 0, \quad j = 1, \ldots, 2s + 2 \tag{3.2}$$

has a solution, then $L_{n,m}(F_0, \ldots, F_s)$ has nonempty interior and the solution vector c is unique. This vector points in the direction of the axis of the largest cone that can be inscribed in the unbounded polyhedral $L_{n,m}(F_0, \ldots, F_s)$ both with the apex in the origin. So throwing the normalization overboard leads to an unbounded search area and results in a halfline instead of a point. But the QP problem has a unique solution while the LP problem does not.

The QP formulation offers another advantage. If the Chebyshev center c returned by the LP problem is very close to the origin and, at the same time, the radius ρ of the largest Euclidean ball with center c contained in $L_{n,m}(F_0, \ldots, F_s) \cap [-1, 1]^{n+m+2}$ is terribly small, then it may be difficult to conclude that the returned solution does not coincide with the origin and that the radius is nonzero. However, the value of δ in the QP problem formulation can be made arbitrarily large because $L_{n,m}(F_0, \ldots, F_s)$ is unbounded. While we cannot control the value of ρ in the LP problem, we have complete freedom in the choice of δ in the QP problem. Both play a similar role, namely they express the robustness of the solution vector c because in both the LP and QP formulations a Euclidean ball of radius ρ (in LP) and δ (in QP) and centered at c is guaranteed to be included in the solution set. The choice of δ is discussed in detail in [16]. We conclude with a lemma that connects the two problem statements.

Lemma 1 *If for some $\delta > 0$ the vector c^* solves $\arg \min_{c \in \mathbb{R}^{n+m+2}} ||c||_2^2$ under the constraint* (3.2), *then for $L = ||c^*||_\infty + \delta$ the vector c^* is a Chebyshev center of the polyhedron $L_{n,m}(F_0, \ldots, F_s) \cap [-L, L]^{n+m+2}$.*

Proof So we assume that for some $\delta > 0$ a solution c^* of the QP problem satisfying constraint (3.2) exists. Hence the interior of $L_{n,m}(F_0, \ldots, F_s)$ is nonempty. With $L = ||c^*||_\infty + \delta$ the Euclidean ball $B(c^*, \delta)$ is entirely contained in the polyhedron $L_{n,m}(F_0, \ldots, F_s) \cap [-L, L]^{n+m+2}$. We now show that there are no Euclidean balls with radius larger than δ inside this polyhedron. The proof of this fact is by contradiction.

A Euclidean ball with center c and radius $\Delta > \delta$ always exists in the interior of the unbounded $L_{n,m}(F_0, \ldots, F_s)$, since this interior is nonempty. Assume that such a Euclidean ball also lies within $[-L, L]^{n+m+2}$. Then necessarily

$$\sup_{||u||_2 \leq \Delta} ||c + u||_2 \leq L,$$

from which we find that

$$||c||_2 + \Delta \leq ||c^*||_2 + \delta.$$

Because $\Delta > \delta$ we have $||c||_2 < ||c^*||_2$. The latter contradicts the minimality of $||c^*||_2$ because $B(c, \Delta) \subset L_{n,m}(F_0, \ldots, F_s)$ implies that c also satisfies constraint (3.2):

$$A_{n,m}^{(j)} c \geq \Delta ||A_{n,m}^{(j)}||_2 \geq \delta ||A_{n,m}^{(j)}||_2, \quad j = 1, \ldots, 2s + 2.$$

Hence $\Delta \leq \delta$ and c^* is a Chebyshev center of $L_{n,m}(F_0, \ldots, F_s) \cap [-L, L]^{n+m+2}$.

□

Let us now apply the new data fitting technique to a variety of problems which we collected from projects with industrial partners.

4 Designing multidimensional recursive filters

The design of multidimensional recursive filters is of considerable interest, but it is inherently a nonlinear complex approximation problem. The technology is required in many diverse areas, including image processing, video signal filtering, tomography and different grid-based methods in scientific computing. For instance, an ideal lowpass filter is specified by its frequency response

$$H(\exp(\mathtt{i}t), \exp(\mathtt{i}u)) = \begin{cases} 1, & (t, u) \in P \subset [-\pi, \pi] \times [-\pi, \pi] \\ 0, & (t, u) \notin P \end{cases} \tag{4.1}$$

where the domain P can take different forms and be composed of squares, triangles, horizontal and vertical strips. We illustrate our technique for the computation of the centro symmetric filter [13] where

$$P = \{\pi/4 \leq t \leq 3\pi/4, -3\pi/4 \leq u \leq -\pi/4\}$$
$$\cup \{-3\pi/4 \leq t \leq -\pi/4, \pi/4 \leq u \leq 3\pi/4\}$$

and the more difficult 90 degree quadrantally symmetric fan filter [9] where

$$P = \{|t| \leq u\} \cup \{-|t| \geq u\}.$$

For the former the base functions

$$g_{\alpha_1, \alpha_2}(t, u) = \cos(\alpha_1 t + \alpha_2 u), \quad \alpha = (\alpha_1, \alpha_2) \in \mathbb{N} \times \mathbb{N} \cup \mathbb{N}_0 \times (-\mathbb{N}_0)$$
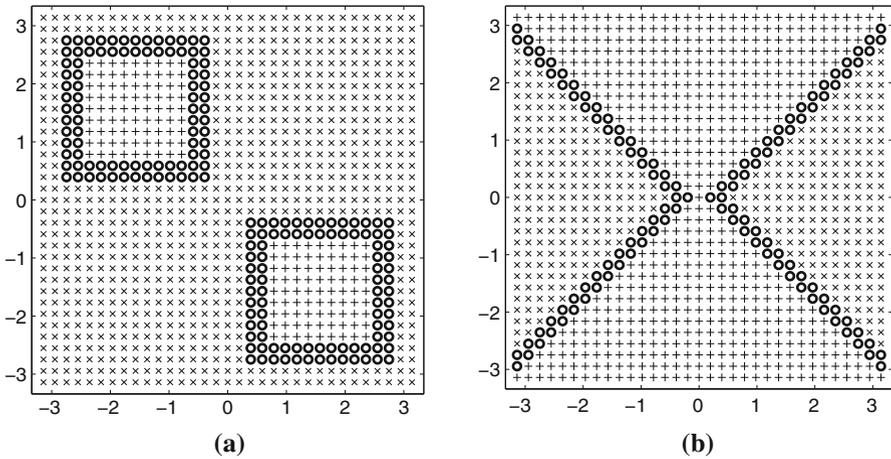
Fig. 3 Centro-symmetric filter (**a**) and quadrantically symmetric filter (**b**)

are used (we need to adapt our ordering $\prec$ to include the negative α_2). For the latter

$$g_{\alpha_1,\alpha_2}(t, u) = \cos(\alpha_1 t) \cos(\alpha_2 u), \quad \alpha = (\alpha_1, \alpha_2) \in \mathbb{N}^2.$$

The frequency domain $[-\pi, \pi] \times [-\pi, \pi]$ for $v = (t, u)$ is covered by a grid of $s+1 = 33 \times 33 = 1089$ points by sampling in both directions equidistantly with distance $\pi/16$. The uncertainty intervals F_i offer the advantage that the filter's performance specifications are incorporated in the characterization. The practical specifications of such a filter then take the form of a tolerance scheme, graphically illustrated in Fig. 3a (centro symmetric filter) and Fig. 3b (fan filter), in which:

- There is a passband wherein the frequency response must approximate 1 with an error of $\pm\delta_1$ (plotted with plus signs),

$$r_{n,m}(t_i, u_i) \in [1 - \delta_1, 1 + \delta_1], \quad (t_i, u_i) \in P,$$

- There is a stopband in which the response must approximate zero with an error less than $\pm\delta_2$ (plotted with crosses),

$$r_{n,m}(t_i, u_i) \in [-\delta_2, \delta_2], \quad (t_i, u_i) \notin P \cup T,$$

- There is a transition band of a certain width in which the response drops smoothly from the passband to the stopband (plotted with circles),

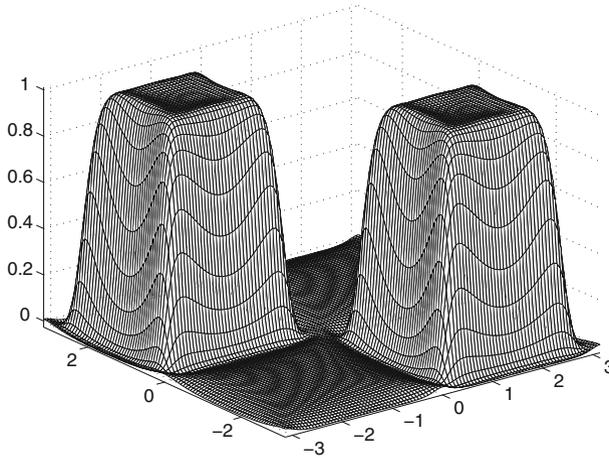$$r_{n,m}(t_i, u_i) \in [-\delta_2, 1 + \delta_1], \quad (t_i, u_i) \in T.$$

Fig. 4 Interval interpolant $r_{19,20,\diamond}(t, u)$

In Fig. 3a, b the transition band is respectively given by

$$
\begin{aligned}
T &= (\{\pi/8 \leq t \leq 7\pi/8, -7\pi/8 \leq u \leq -\pi/8\} \\
&\quad \cup \{-7\pi/8 \leq t \leq -\pi/8, \pi/8 \leq u \leq 7\pi/8\}) \setminus P, \\
T &= \{-|u| - \pi/8 \leq t \leq -|u|\} \cup \{|u| \leq t \leq |u| + \pi/8\}.
\end{aligned}
$$

For the choice $\delta_1 = 0.01$ the passband ripple

$$
20 \log_{10} \frac{1 + \delta_1}{1 - \delta_1} = 0.17 \text{ dB}
$$

and for $\delta_2 = 0.02$ the stopband attenuation is

$$
-20 \log_{10} \delta_2 = 34 \text{ dB}.
$$

With these parameters we obtain for the centrosymmetric filter a model with

$$
N = D \setminus \{(1, -3)\}
$$
$$
D = \{(\alpha_1, \alpha_2) \in \mathbb{N}^2 \mid 0 \leq \alpha_1 + \alpha_2 \leq 4\} \cup \{(\alpha_1, \alpha_2) \in \mathbb{N}^2 \mid 2 \leq \alpha_1 - \alpha_2 \leq 4\}.
$$

The rational function $r_{19,20,\diamond}(t, u)$ is plotted in Fig. 4. As envisaged $n + m = 39 \ll s = 1088$.

The second model, obtained for the fanfilter, is of much less complexity than the one presented in [9]: its numerator and denominator only contain 13 and 15 terms respectively with the sets N and D given by
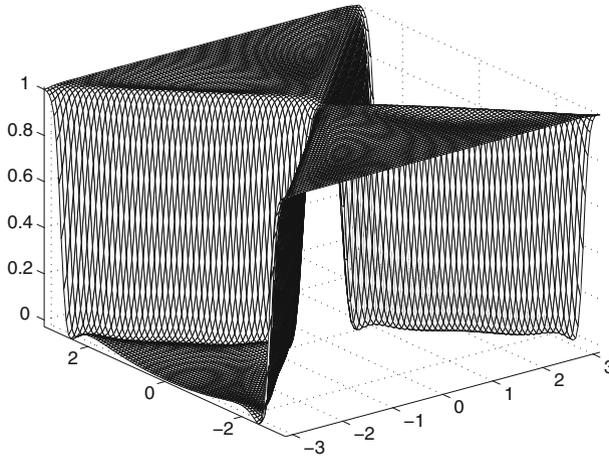
Fig. 5 Interval approximant $r_{12,14,\diamond}(t, u)$

$$N = \{(\alpha_1, \alpha_2) \in \mathbb{N}^2 \mid 0 \leq \alpha_1, \alpha_2 \leq 3\} \backslash \{(3, 2), (2, 3), (3, 3)\}$$
$$D = \{(\alpha_1, \alpha_2) \in \mathbb{N}^2 \mid 0 \leq \alpha_1, \alpha_2 \leq 3\} \backslash \{(3, 3)\}.$$

The resulting rational function $r_{n,m,\diamond}(t, u) = r_{12,14,\diamond}(t, u)$ is shown in Fig. 5. Here again $n + m = 26 \ll s = 1088$.

When the parameters are relaxed to $\delta_1 = 0.05$ and $\delta_2 = 0.1$ then, as expected, the obtained model has even lower complexity:

$$N = \{(\alpha_1, \alpha_2) \in \mathbb{N}^2 \mid 0 \leq \alpha_1 + \alpha_2 \leq 3\} \cup \{(2, 2)\}$$
$$D = \{(\alpha_1, \alpha_2) \in \mathbb{N}^2 \mid 0 \leq \alpha_1 + \alpha_2 \leq 2\}.$$

The passband ripple becomes 0.87 dB while the stopband attenuation now equals 20 dB. This approximation $r_{n,m,\diamond}(t, u) = r_{10,5,\diamond}(t, u)$ is graphed in Fig. 6.

Instead of relaxing the tolerances δ_1 and δ_2, the dataset in Fig. 3b can be reduced to about one quarter because of the symmetric nature of the basis functions. We have

$$q_{N,D}(t, u) = q_{N,D}(-t, u) = q_{N,D}(t, -u) = q_{N,D}(-t, -u)$$

and similarly for $p_{N,D}(t, u)$. This may lead to a subtle difference in the neighbourhood of the axes $t = 0$ and $u = 0$ as can be seen in Fig. 7 compared to Fig. 5.

Similar experiments were made for the circularly symmetric and diamond-shaped lowpass filters given in [8].

Note that the equiripple behaviour around 1 and 0 in the pass- and stopband respectively, is typical for the obtained models. This is in line with its connection with the theory of best approximation, as discussed in Sect. 2.
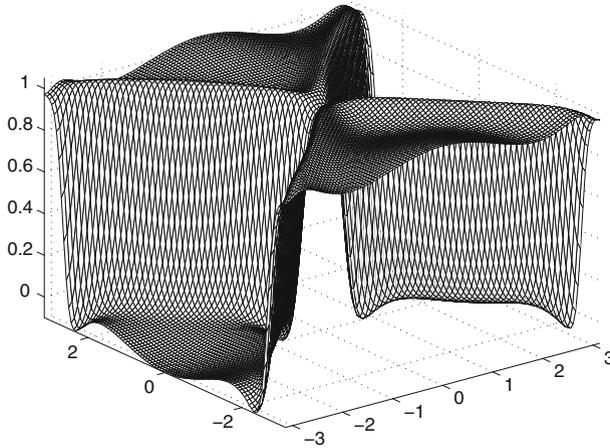
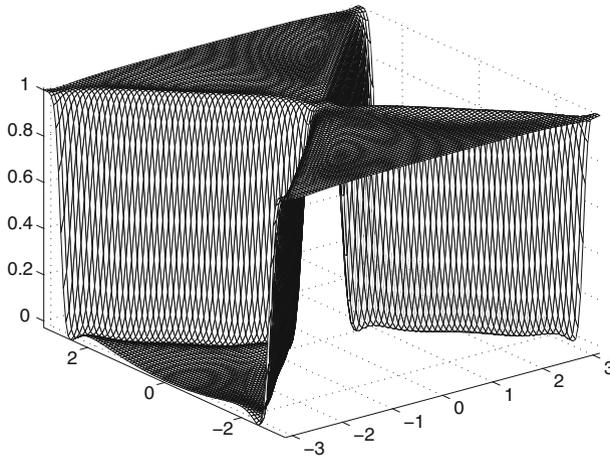**Fig. 6** Interval interpolant $r_{10,5,\diamond}(t, u)$

Fig. 7 Quadrantically symmetric filter with reduced dataset

5 Option price modelling

Options are a type of financial instrument classed as derivatives, as they derive their value from an underlying asset. A European call option 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. In return for granting the option, the writer of the option collects a payment from the buyer. The theoretical value $C(S, t)$ of an option is evaluated according to several models. These models attempt to predict how the value of an option changes in response to changing conditions. Black and Scholes derived the partial differential equation that must be satisfied by the price of any derivative dependent on a non-dividend-paying stock,
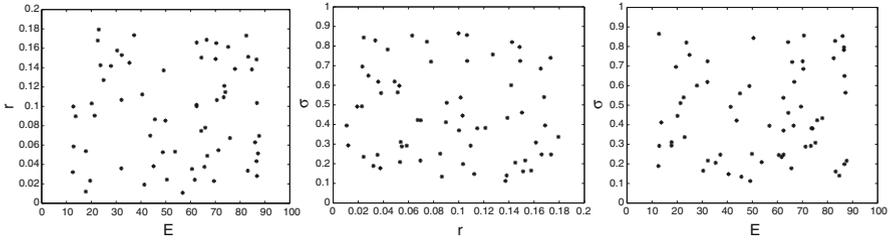
Fig. 8 Sampled (E, r, σ) in $[0, 100] \times [0, 0.2] \times [0, 1]$ (64 triplets)

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

with boundary conditions

$$C(S, T) = \max(S - E, 0), \tag{5.1}$$

$$C(0, t) = 0, \quad 0 \le t \le T, \tag{5.2}$$

$$C(S, t) \approx S - E\exp(-rt), \quad \text{large } S, t \ge 0, \tag{5.3}$$

and produced a closed-form solution for a European option's theoretical price. Here t denotes the elapsed time with the expiry date at T, S is the asset price which depends on t, r is the annual interest rate (assumed constant) and σ is the market volatility. The value E is called the exercise price or strike price. The application of the model in actual options trading is clumsy because of the assumptions on the dividend payment, the constant volatility, and a constant interest rate. But stochastic volatility models require complex numerical methods. Nevertheless, the Black-Scholes model is still one of the most important methods and foundations for the existing financial market. We use it here only to illustrate a proof of principle.

Typically the parameters r and σ are estimated, E is chosen and the equation is solved for $C(S, t)$ on a grid. This procedure can be repeated a few times. For a few dozen of these (E, r, σ) triplets and moderate 200×200 (uniform or non-uniform) grids of (S, t) we quickly obtain a few million values. Remains the problem of writing down an analytical model for the variation of the option price $C(S, t)$ with the parameters E, r and σ. Since we know that $C \approx 0$ for very small S and $C \approx S - E\exp(-rt)$ 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 t (in years). In Fig. 8 we show the (E, r, σ) triplets for which we have solved the partial differential equation (they were randomly generated). 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.

In order to be able to show plots of the rational interval approximant, we introduce the new variables [11]
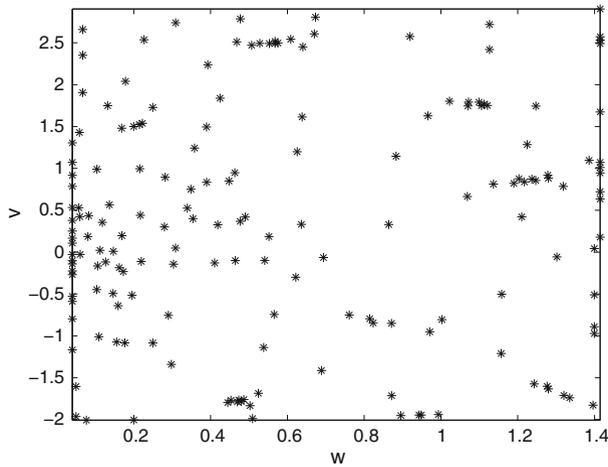
Fig. 9 Datapoints (u_i, w_i) used for $r_{29,11}(u, w)$

$$u = \ln(S) - \ln(E) + rt, \tag{5.4}$$

$$w = \sigma\sqrt{t}, \tag{5.5}$$

and compute a rational interval approximant $r_{n,m,\diamond}(v) = r_{n,m,\diamond}(u, w)$ for $C(u, w)/S$. Each quintuple (S, E, r, σ, t) is mapped to some vector $v = (u, w)$. We use the computed values for $C(S, t)$ at $(S_i, E_i, r_i, \sigma_i, t_i)$ to obtain $f(u_i, w_i) = C(u_i, w_i)/S_i$ and we define

$$f_{i1} = \max(0, f(u_i, w_i) - 0.002), \tag{5.6}$$

$$f_{i2} = \min(1, f(u_i, w_i) + 0.002). \tag{5.7}$$

For the ordered sequence of base functions $g_i(u, w)$ we use

$$1, T_1(u), T_1(w), T_2(u), T_2(w), T_1(u)T_1(w),$$
$$T_3(u), T_3(w), T_2(u)T_1(w), T_1(u)T_2(w), \ldots$$

where T_n denotes the n-th Chebyshev polynomial of the first kind. All variables are also rescaled to the interval $[-1, 1]$. With the subset of 175 points shown in Fig. 9 we obtain $r_{29,11}(u, w)$ given in Fig. 10 where it can also be compared to the explicit formula for $C(u, w)/S$. Note that the exact formula is non-differentiable at the point $(0, 0)$ and hence that the rational approximation method uses more samples in the neighbourhood of that point. The index sets N and D used for $r_{29,11}(v)$ are

$$N = \{(\alpha_1, \alpha_2) \mid 0 \leq \alpha_1 + \alpha_2 \leq 6\} \cup \{(7, 0), (0, 7)\} \tag{5.8}$$

$$D = \{(\alpha_1, \alpha_2) \mid 0 \leq \alpha_1 + \alpha_2 \leq 3\} \cup \{(4, 0), (0, 4)\}. \tag{5.9}$$
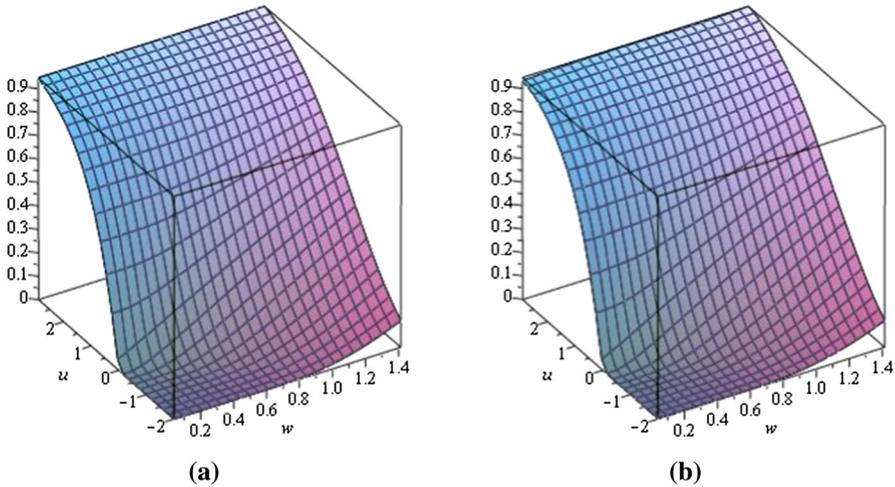
(a)   **(b)**

Fig. 10 Exact value $C(u, w)/S$ (left) and approximant $r_{29,11}(u, w)$ (right)

We now have to explain how the 175 datapoints were selected from the full dataset. We start with a small number of $s_0 + 1$ samples, for instance 9 in a 3×3 subdivision of the (u, w)-square, one point in each of the 9 small squares. We compute $r_{n_0,m_0}^{(0)}(u, w)$ for these $s_0 + 1$ points. Then we check how many of the intervals F_i in the full dataset are interpolated by this model in addition, without imposing. This is usually more than merely these $s_0 + 1$. Of the (u_i, w_i) where $r_{n_0,m_0}^{(0)}(u, w)$ does not pass through F_i, we select $s_1 - s_0$ ones where the value of the rational interpolant lies furthest from the prescribed interval. We add these points to obtain $s_1 + 1$ samples in total and repeat the process. In other words, we compute $r_{n_1,m_1}^{(1)}(u, w)$ that interpolates these $s_1 + 1$ data. We check again how many additional intervals in the full dataset are interpolated automatically, without imposing. And so on.

When restricting the region of interest to the more interesting

$$\tfrac{1}{2} \le S/E \le \tfrac{3}{2},$$

even while reducing the absolute error from 0.002 to 0.001, then n, m and s in $r_{n,m,\diamond}(u, w)$ decrease to 16, 13 and 101 respectively. The resulting approximation looks no different from the one given earlier. Therefore we show in Fig. 11, for this restricted approximant, the relative error $\left| Sr_{16,13}(u, w) - C(u, w) \right| / C(u, w)$.

6 Modelling bidirectional reflectance distribution functions

The bidirectional reflectance distribution function (BRDF) ρ is a four-dimensional function that defines how light is reflected at an opaque surface and it is the cornerstone of the rendering equation in physics and computer graphics. Therefore building a model that is able to explain, describe and simulate all complex phenomena involved

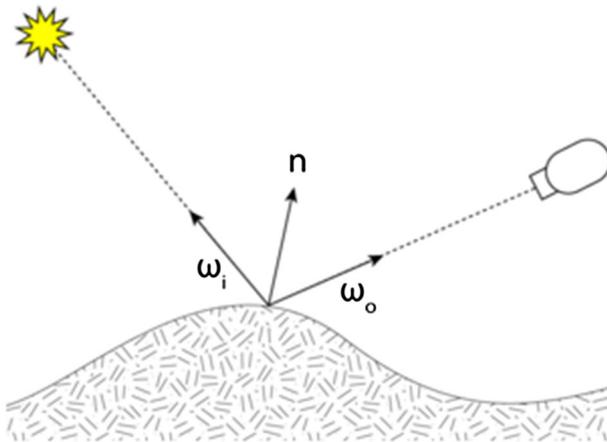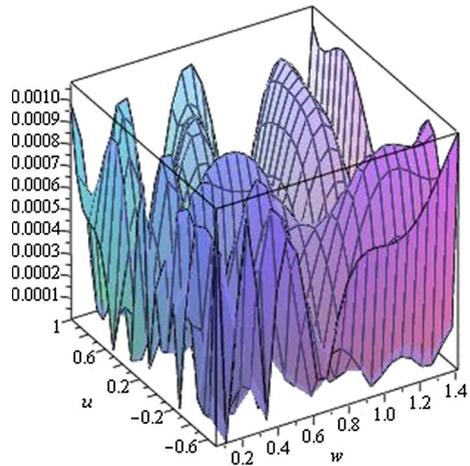**Fig. 11** Relative error $\left|Sr_{16,13}(u, w) - C(u, w)\right|/C(u, w)$

Fig. 12 Vectors $\boldsymbol{\omega}_i$ toward the light source and $\boldsymbol{\omega}_o$ toward the viewer or camera

in this process remains an active topic of research. The function takes an incoming light direction $\boldsymbol{\omega}_i$ and outgoing direction $\boldsymbol{\omega}_o$, both defined with respect to the surface normal **n**, and returns the ratio of reflected radiance exiting along $\boldsymbol{\omega}_o$ to the irradiance incident on the surface from direction $\boldsymbol{\omega}_i$. Note that each direction $\boldsymbol{\omega}$ is itself parameterized by azimuth angle ϕ and zenith angle θ, therefore $\rho(\theta_i, \phi_i, \theta_o, \phi_o)$ as a whole is 4-dimensional. It is convenient to take the vectors $\boldsymbol{\omega}_i$, $\boldsymbol{\omega}_o$ and **n** of unit length (Fig. 12).

The BRDF is a fundamental radiometric concept, and accordingly is used in computer graphics for photorealistic rendering of synthetic scenes as well as in computer vision for many inverse problems such as object recognition.

Many researchers have developed devices for efficiently acquiring BRDFs from real world samples. BRDFs can be measured directly, at different points on the surface and under several point source illuminatons at known positions, from real objects using
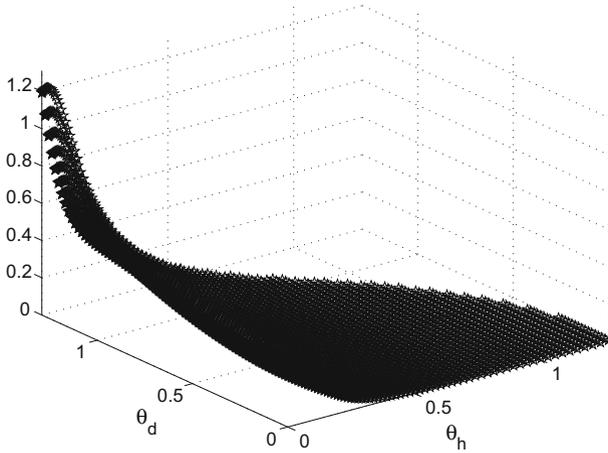
Fig. 13 Blue channel BRDF (blue metallic paint) as a function of θ_h and θ_d (colour figure online)
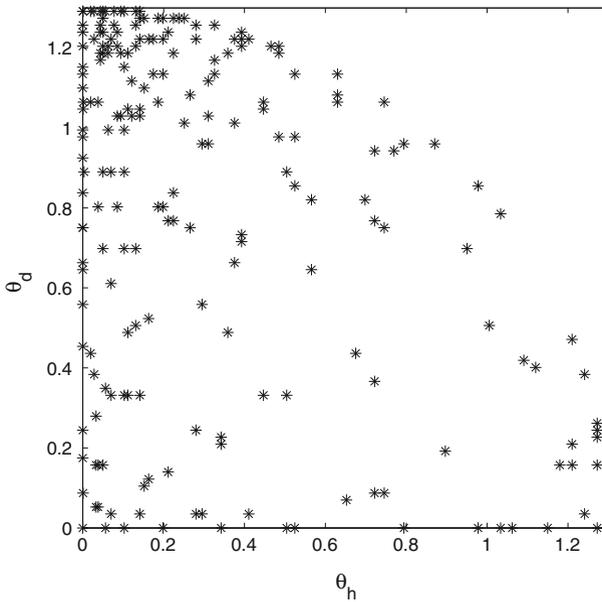

Fig. 14 Subset of 207 training data for $r_{25,18,\diamond}(\theta_h, \theta_d)$

calibrated cameras and lightsources. From this set of images, one then wants to recover the surface normal and BRDF at each point on the objects surface.

Instead of the incoming and outgoing directions, one can use the half-angle which is the angle between the surface normal **n** and the bisector vector $\boldsymbol{\omega}_i + \boldsymbol{\omega}_o$, and the difference angle which is the angle between the source vector and the bisector. Both these angles entail azimuth and zenith variables $\theta_h, \phi_h, \theta_d, \phi_d$. For isotropic surface materials $\phi_h = 0$ and so one measurement typically consists of $90 \times 90 \times 180$ samples

Fig. 15 Relative error of $r_{25,18,\diamond}(\theta_h, \theta_d)$ in all datapoints
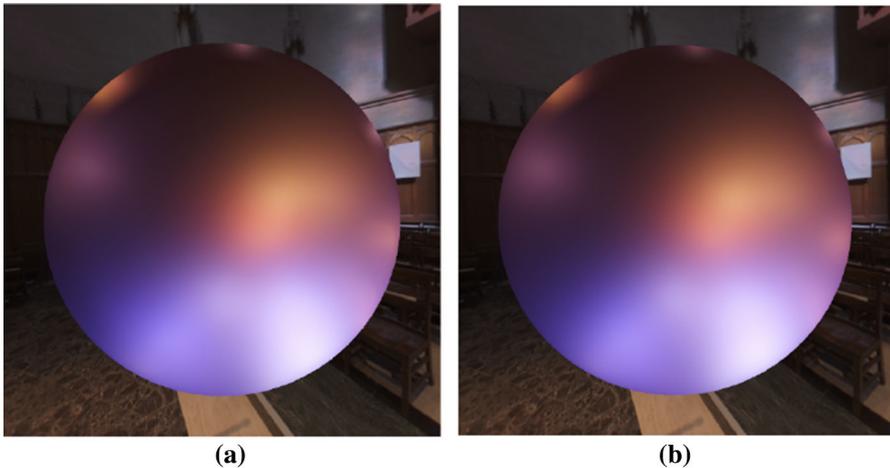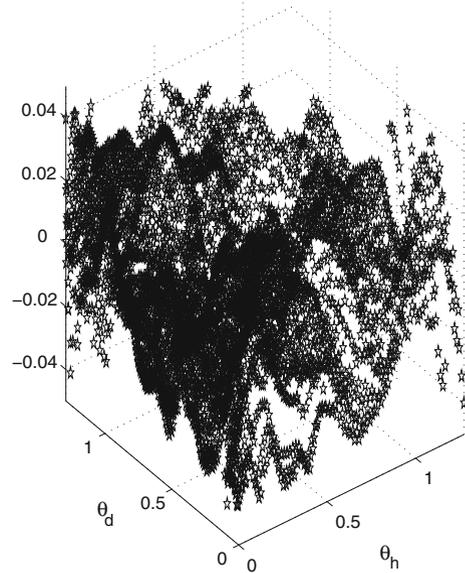

Fig. 16 Rendered example (blue metallic paint): original (267 Mb) left and approximation (8.5 Kb) right (colour figure online)

where a sample handles multiple color channels, for instance the 3 RGB values. In other words a measurement amounts to about 267 Mb.

Because the BRDF satisfies a number of empirically observed physical properties shared by many materials, one can further reduce the domain of the BRDF to two variables [14]. For isotropic materials the 3-d BRDF can be projected on the 2-d (θ_h, θ_d)-space disregarding ϕ_d [12] and hence we compute $r_{n,m,\diamond}(\theta_h, \theta_d)$. For the bivariate representation, the original 267 Mb of data can thus be reduced to some 6000 RGB datapoints, the equivalent of about 1.12 Mb. We remark that data from materials

with high specularity are especially hard to fit. An advantage of our technique is that there is no need to separate the diffuse and the specular components. The nonlinear character of the rational model can handle the challenge.

Defining intervals F_i at each of these 6000 datapoints with a relative accuracy of $\pm\,3\%$ where θ_h is close to zero and $\pm\,5\%$ elsewhere (see Fig. 13), we can start the same procedure as in the previous section to come up with the set of training data (see Fig. 14) required to build an approximant $r_{n,m,\diamond}(\theta_h, \theta_d)$ accurate on the whole set of datapoints (see Fig. 15).

Using a rational interval interpolant $r_{n_C,m_C,\diamond}(\theta_h, \theta_d)$ for each of the RGB colour channels ($C \in \{R, G, B\}$) instead of the full dataset of 267 Mb when rendering the blue metallic painted ball of Fig. 16, does not generate any noticeable difference. Note that the data footprint of the rational functions is only the storage of the rational function coefficients, namely

$$\sum_{C \in \{R,G,B\}} (n_C + m_C + 1)(32 \text{ or } 64 \text{ bits}).$$

References

1. Allouche, H., Cuyt, A.: On the structure of a table of multivariate rational interpolants. Constr. Approx. **8**, 69–86 (1992). https://doi.org/10.1007/BF01208907
2. Allouche, H., Cuyt, A.: Unattainable points in multivariate rational interpolation. J. Approx. Theory **72**, 159–173 (1993). https://doi.org/10.1006/jath.1993.1013
3. Berrut, J.P.: Rational functions for guaranteed and experimentally well-conditioned global interpolation. Comput. Math. Appl. **15**, 1–16 (1988). https://doi.org/10.1016/0898-1221(88)90067-3
4. Boehm, B.: Existence of best rational tchebycheff approximations. Pac. J. Math. **15**(1), 19–28 (1965)
5. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
6. Gilewicz, J., Magnus, A.: Valleys in C-table. LNM **765**, 135–149 (1979)
7. Goldman, A.J., Tucker, A.W.: Polyhedral convex cones. In: Kuhn, H.W., Tucker, A.W. (eds.) Linear Inequalities and Related Systems. Annals of Mathematics Studies, vol. 38, p. 1940. Princeton University Press, Princeton (1956)
8. Gorinevsky, D., Boyd, S.: Optimization-based design and implementation of multi-dimensional zero-phase IIR filters. IEEE Trans. Circuits Syst. I **53**(2), 372–383 (2006). https://doi.org/10.1109/TCSI.2005.856048
9. Harris, D.: Design of 2-D rational digital filters. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '81, pp. 696–699 (1981)
10. Ibrahimoglu, B.A., Cuyt, A.: Sharp bounds for Lebesgue constants of barycentric rational interpolation. Exp. Math. **25**(3), 347–354 (2016). https://doi.org/10.1080/10586458.2015.1072862
11. Li, M.: Approximate inversion of the Black–Scholes formula using rational functions. Eur. J. Oper. Res. **185**(2), 743–759 (2008). https://doi.org/10.1016/j.ejor.2006.12.028
12. Matusik, W., Pfister, H., Brand, M., McMillan, L.: A data-driven reflectance model. ACM Trans. Graph. **22**(3), 759–769 (2003)
13. Reddy, H., Khoo, I.H., Rajan, P.: 2-D symmetry: theory and filter design applications. IEEE Circuits Syst. Mag. **Third Quarter**, 4–33 (2003)
14. Rusinkiewicz, S.M.: A new change of variables for efficient BRDF representation. In: In Eurographics Workshop on Rendering, pp. 11–22 (1998)
15. Salazar Celis, O., Cuyt, A., Van Deun, J.: Symbolic and interval rational interpolation: the problem of unattainable data. In: Simos, T., Psihoyios, G., Tsitouras, C. (eds.) International conference on numerical analysis and applied mathematics, AIP Conference Proceedings, vol. 1048, pp. 466–469 (2008). https://doi.org/10.1063/1.2990963
16. Salazar Celis, O., Cuyt, A., Verdonk, B.: Rational approximation of vertical segments. Numer. Alg. **45**, 375–388 (2007). https://doi.org/10.1007/s11075-007-9077-3