Stable Multivariate Rational Interpolation for Parameter-dependent Aerospace Models

Pranay Seshadri*
University of Cambridge
Cambridge, CB2 1PZ, U.K.

Paul Constantine†
Stanford University
Stanford, CA 94305, U.S.A

Pedro Gonnet‡
Durham University
Durham, DH1 3LE, U.K.

Geoffrey Parks§
University of Cambridge
Cambridge, CB2 1PZ, U.K.

Shahrokh Shahpar¶
Rolls-Royce plc
Derby, DE24 8BJ, U.K

A multivariate, robust, rational interpolation method for propagating uncertainties in several dimensions is presented. The algorithm for selecting numerator and denominator polynomial orders is based on recent work that uses a singular value decomposition approach. In this paper we extend this algorithm to higher dimensions and demonstrate its efficacy in terms of convergence and accuracy, both as a method for response suface generation and interpolation. To obtain stable approximants for continuous functions, we use an L_2 error norm indicator to rank optimal numerator and denominator solutions. For discontinous functions, a second criterion setting an upper limit on the approximant value is employed. Analytical examples demonstrate that, for the same stencil, rational methods can yield more rapid convergence compared to pseudospectral or collocation approaches for certain problems.

Nomenclature

$a_{i,j}$	Output 2D numerator coefficient matrix
$b_{i,j}$	Output 2D denominator coefficient matrix
$f(x_i, y_i)$	Sampled function at 2D interpolation points
$\mathbf{f}$	Vector of sampled function values, $f(x_i, y_i)$
$m = (m_x, m_y)$	Input 2D numerator coefficient matrix (boolean)
$n = (n_x, n_y)$	Input 2D denominator coefficient matrix (boolean)
$p_m(x, y)$	Rational approximant numerator
$q_n(x, y)$	Rational approximant denominator
$T_i(x)$	Orthogonal polynomial in x of order $i - 1$

I. Introduction

The introduction of rational methods as a suitable platform for uncertainty quantification is a new concept that has recently demonstrated considerable promise.[1,2] As with polynomial methods, the basic idea of expressing stochastic parameters as finite dimensional, functional expansions of random variables is still valid. The difference lies in the replacement of polynomials with rational expressions. One of the reasons

*PhD Research Student, Inglis Building IN2-17, Trumpington Street, AIAA Student Member
†Postdoctoral Researcher, Department of Mechanical Engineering, Building 500
‡Lecturer, School of Engineering and Computing Sciences, E113
§Senior University Lecturer, Engineering Design Centre, Inglis Building IN2-11, Trumpington Street
¶Royal Aeronautical Society Fellow, Rolls-Royce Engineering Associate Fellow, AIAA Associate Fellow

rational methods have not been widely implemented is because they are known to be numerically fragile. Should the numerator and denominator polynomials be less than a desired order, the resulting approximation will degenerate. This desired order is based on the whether the input numerator and denominator orders are sufficient to accurately interpolate the function at sample points. The resulting approximant must also have a denominator that is conditioned not to change sign in its domain. Ascertaining this desired order from a priori information on the function is far from trivial. For univariate problems, the desired order may be reached by iteratively reducing the degree of the denominator. However, in multivariate problems, the denominator will have different degrees in each dimension, and thus merely reducing its degree in each dimension by the same order may not yield a stable approximant. The second complication with rational methods is the appearance of spurious pole-zero pairs, known as Froissart doublets, that prevent pointwise convergence.[3]

Numerous algorithms have been proposed to compute numerator and denominator coefficients for univariate rationals (see Pachón,[4] Hesthaven[5] and van Deun[6]), however, very few have been extended to two dimensions. Two algorithms that do offer such an extension are by Chantrasmi et al.[1] and Gonnet et al.[7] Both approaches effectively attempt to regularize the ill-posed problem of obtaining numerator and denominator coefficients, and do so by resorting to a singular value decomposition.

In this paper, we extend the algorithm proposed in Gonnet et al. to higher dimensions ($N > 2$). The implementation provided here is simple and straightforward, ideal for use in a host of uncertainty quantification problems. Through numerical experiments, we demonstrate that rational methods can yield remarkable convergence properties if properly conditioned. The test functions explored cover both continuous and discontinuous functions and range from two-dimensional to four-dimensional problems.

II. Two-dimensional Rational Interpolation

A. Algorithm

The two-dimensional rational interpolation methodology presented here is based on the work of Gonnet et al.,[8] which in turn is an outgrowth of the work in Van Deun.[4] While outlining the methodology, we will be using two seemingly similar, but different, terminologies: interpolation and approximation. We define these in terms of the minimization of the L_2 error norm, which can be computed as:

$$\min |L_2(x,y)| = \min \left| \int_X \int_Y \left(f(x,y) - R(x,y)\right)^2 dydx \right|^{1/2} \tag{1}$$

Here $f(x,y)$ is the sampled function and $R(x,y)$ is the rational approximant/interpolant. $R(x,y)$ is computed from function values at interpolation points. The interpolation problem is thus a minimization of the L_2 error norm of both $f(x,y)$ and $R(x,y)$ at interpolation points. The approximation problem, on the other hand, is a minimization of the L_2 error norm for all points (interpolation and non-interpolation ones) within the support. Thus, even though an $R(x,y)$ can be obtained that accurately interpolates the provided data, that does not necessarily guarantee that it is a stable approximant to the original function, $f(x,y)$. In fact, one of the major causes of concern with rational approximants is that between interpolation nodes, the denominator polynomial might change sign, which will lead to a singularity in the approximant. This point will be revisited in the ensuing sections. For now we review the two-dimensional algorithm

For data sampled from a two-dimensional anisotropic tensor grid of order $[N_1, N_2]$, we define the rational interpolant to be:

$$f(x_i, y_i) = \frac{p_m(x_i, y_i)}{q_n(x_i, y_i)}; i = 1, \ldots, N \tag{2}$$

where the numerator and denominator are defined to be the weighted sum of multivariate orthogonal polynomials:

$$p_m(x,y) = \sum_{i,j} a_{i,j} T_i(x) T_j(y) \tag{3}$$

$$q_n(x,y) = \sum_{i,j} b_{i,j} T_i(x) T_j(y) \tag{4}$$

These multivariate orthogonal polynomials are simply tensor products of one-dimensional polynomials with respect to an index set. The degrees of the numerator and denominator polynomials are expressed not as

integers but as matrices that describe which coefficients may be zero or non-zero. This may be in the form of anistropic, isotropic tensor grids, or random entries.

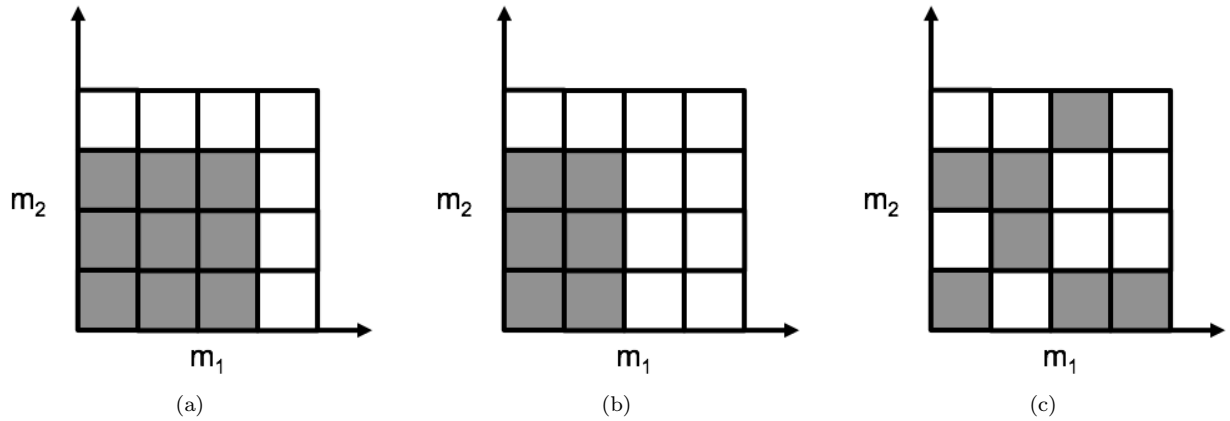


Figure 1. Input matrix forms of m, where entries can be in the form of (a) an isotropic tensor, (b) an anisotropic tensor or (c) random elements

Figure 1 illustrates possible entries for the m matrix, which corresponds to coefficient entries for $p_m(x, y)$. Shaded entries in Figure 1 correspond to non-zero values. The dimensions of the matrix here are $[N_1, N_2]$. Thus, for both m and n matrices, we have the constraints

$$N_1 \geq \max\{m_1, n_1\} \tag{5}$$

$$N_2 \geq \max\{m_2, n_2\}$$

$$m_1 m_2 + n_1 n_2 \leq N_1 N_2$$

Our main objective in this section is to compute the numerator and denominator coefficients $a_{i,j}$ and $b_{i,j}$, in order to obtain the rational approximation. We begin by considering the linearized form of Equation 2, and expressing it as a matrix vector product:

$$\begin{bmatrix} f(x_1, y_1) & & & \\ & f(x_2, y_2) & & \\ & & \ddots & \\ & & & f(x_N, y_N) \end{bmatrix} \begin{bmatrix} q_n(x_1, y_1) \\ q_n(x_2, y_2) \\ \vdots \\ q_n(x_N, y_N) \end{bmatrix} = \begin{bmatrix} p_m(x_1, y_1) \\ p_m(x_2, y_2) \\ \vdots \\ p_m(x_N, y_N) \end{bmatrix} \tag{6}$$

$$\Rightarrow \text{diag}(\mathbf{f})\mathbf{q} = \mathbf{p}$$

Here N is an arbitrary order, that represents the number of equations in the system for $|m|+|n|-1$ unknows. Focusing on the left-hand side of the above equation, we find that the coefficient form of the denominator $\mathbf{q}$, can be obtained by evaluating the multivariate orthogonal polynomial at all tensor grid data points and multiplying it with all the non-zero entries of the n matrix. This is expressed as:

$$\begin{bmatrix} T_1(x_1)T_1(y_1) & T_1(x_1)T_2(y_1) & \dots & T_?(x_1)T_?(y_1) \\ T_1(x_2)T_1(y_2) & T_1(x_1)T_2(y_1) & & \\ \vdots & & \ddots & \\ T_1(x_N)T_2(y_N) & & & T_?(x_N)T_?(y_N) \end{bmatrix} \begin{bmatrix} b_{1,1} \\ b_{1,2} \\ \vdots \\ b_{?,?} \end{bmatrix} = \begin{bmatrix} q_n(x_1, y_1) \\ q_n(x_2, y_2) \\ \vdots \\ q_n(x_N, y_N) \end{bmatrix} \tag{7}$$

$$\Rightarrow W_q \mathbf{b} = \mathbf{q}$$

Here, $\mathbf{b}$ has length $|n|$ and, as mentioned earlier, can be ordered arbitrarily. The right-hand side of equation 6 is also expressed in a similar manner. However, here two such matrices are constructed. First, we construct a matrix W_p similar to W_q, but with all the coefficients in m. Then we construct a matrix E_p which is

similar to W_p, but with all the $N - |m|$ coefficients not in m. The resulting system of equations can now be written as:

$$\text{diag}(\mathbf{f})W_q\mathbf{b} = [W_p \,|E_p\,] \begin{bmatrix} \mathbf{a} \\ \hat{\mathbf{a}} \end{bmatrix} \tag{8}$$

Here the vector $\mathbf{a}$ contains the coefficients in m, while $\hat{\mathbf{a}}$ contains the coefficients not in m. Multiplying by $[W_p \,|E_p\,]^{-1}$ from the left:

$$[W_p \,|E_p\,]^{-1} \,\text{diag}(\mathbf{f})W_q\mathbf{b} = \begin{bmatrix} \mathbf{a} \\ \hat{\mathbf{a}} \end{bmatrix} \tag{9}$$

$$\Rightarrow F\mathbf{b} = \begin{bmatrix} \mathbf{a} \\ \hat{\mathbf{a}} \end{bmatrix}$$

We are interested in finding a set of coefficients $\mathbf{b}$ such that the excess coefficients $\hat{\mathbf{a}}$ are zero. We therefore only consider the last $N - |m|$ rows of F, that is $\hat{F}$ such that:

$$\hat{F}\mathbf{b} = 0 \tag{10}$$

The $\hat{F}$ matrix has dimensions $(N - |m|) \times |n|$. To obtain the null-vector $\mathbf{b}$, we compute the singular value decomposition (SVD) of $\hat{F}$. Once the coefficients $\mathbf{b}$ have been computed, we use equation 8 to compute the coefficients $\mathbf{a}$. Multiplying the coefficients with their respective multivariate orthogonal polynomials yields the numerator and denominator polynomials.

B. Singular Values

As stated above, $\hat{F}$ has dimensons $(N - |m|) \times |n|$, which lead to three possible scenarios regarding its null vector:

1. If $N = |m| + |n| - 1$, $\hat{F}$ has dimensions $(|n| - 1) \times |n|$, and thus at least one zero singular value.
2. If $N > |m| + |n| - 1$, $\hat{F}$ is not guaranteed to have a zero singular value vector. However, in this case the right singular vector associated with the smallest singular value will provide the least-squares approximation. It should be noted that this is the discrete least squares over all the data points and not the continuous one defined in equation 1.
3. If $\hat{F}$ has more than one singular value, then this means that there is more than one solution that solves the interpolation exactly.

Mathematically, the last scenario can be represented as:

$$\frac{p_m^{(1)}(x_i, y_i)}{q_n^{(1)}(x_i, y_i)} = \frac{p_m^{(2)}(x_i, y_i)}{q_n^{(2)}(x_i, y_i)} \tag{11}$$

where the numerical superscripts denote different numerator and denominator polynomials. Given the equality condition above, this implies that both solutions carry some common sub-factor $s(x, y)$ that can be eliminated, that is:

$$\frac{p_m(x_i, y_i) \cdot s^{(1)}(x_i, y_i)}{q_n(x_i, y_i) \cdot s^{(1)}(x_i, y_i)} = \frac{p_m(x_i, y_i) \cdot s^{(2)}(x_i, y_i)}{q_n(x_i, y_i) \cdot s^{(2)}(x_i, y_i)} \tag{12}$$

Although both solutions here are correct, that is they solve the interpolation problem, the roots of $s(x, y)$ introduce spurious pole-zero pairs in the solution. In order to remove these spurious poles, we can iteratively reduce the degree of $q_n(x, y)$, re-assemble $\hat{F}$ and compute the singular values until we get one single solution.

III. N-dimensional Extension

A. Algorithm

The N-dimensional rational interpolation problem can be expressed as:

$$f(x_{1i}, x_{2i}, \ldots, x_{Di}) = \frac{p_m(x_{1i}, x_{2i}, \ldots, x_{Di})}{q_n(x_{1i}, x_{2i}, \ldots, x_{Di})}; i = 1, \ldots, N \tag{13}$$

where, as before, the numerator and denominator polynomials are defined to be:

$$p_m(x_1, x_2, \ldots, x_D) = \sum_{i,j,\ldots,d} a_{i,j,\ldots,d} T_i(x_1) T_j(x_2) \ldots T_d(x_D) \tag{14}$$

$$q_n(x_1, x_2, \ldots, x_D) = \sum_{i,j,\ldots,d} b_{i,j,\ldots,d} T_i(x_1) T_j(x_2) \ldots T_d(x_D) \tag{15}$$

Here D represents the total number of dimensions, and $x_1, ..., x_D$ are the variables. As the number of dimensions increases, the matrices m, n and N become D-dimensional structured matrices. For example, when $D = 3$, the anisotropic m matrix can be represented as shown in Figure 2. With each additional

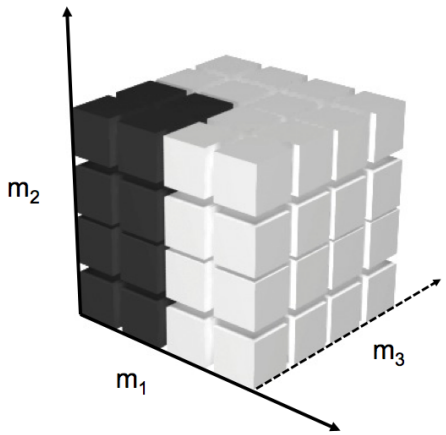


Figure 2. Illustration of boolean m matrix in three dimensions. The size of cube is given by $[N_x, N_y, N_z]$

dimension, the possible combinations of index set choices increases exponentially. Thus, the total number of possible ordered sets for the numerator and denominator also rises. It should be noted that the constraints for these matrices are the same as for the two-dimensional case, shown in Equation 5. Equations 6–9 are still valid, and the SVD of $\hat{F}$ is still computed to determine the coefficients.

IV. Coefficient Selection

A. Overview

As stated earlier, there are two main difficulties in obtaining stable rational interpolants: sign-changing denominators and spurious pole-zero pairs. Both the behavior of the denominator and the number of spurious solutions are governed by the selected numerator and denominators orders (given by the input entries of the m and n matrices). Thus, the coefficient selection problem is to select the best set of m and n matrices that mitigates both issues.

Chantrasmi et al.[1] approach this problem (although their algorithm was different and did not consider spurious pole-zero pairs), by wrapping the coefficient selection problem within a multi-objective optimizer. Their first objective was to minimize the L_2 norm error over the interpolation points, and the second was to minimize a smoothness factor based on the total variation of the rational interpolant.

In this work, no explicit optimization routine is utilized. Instead we use a simple ranking approach to obtain optimal numerator and denominator coefficients.

B. Ranking Algorithm

The coefficient selection algorithm presented in this paper is outlined below:

```
1  Input N, the total number of samples in each dimension
2  Input function values at appropriate quadrature nodes
3  Generate the tensor order index set
```

```
4  For each index, compute SVD
5  Eliminate indices that yield excess singular values
6  Compute L2 norms via a quadrature rule (note: the accuracy of this rule depends on the stencil and order)
7  Find the index with least L2 norm
8  Output m and n corresponding to the lowest L2 norm candidate
```

To begin with, the user inputs the number of samples (N) in each dimension along with the number of dimensions. Based on this data an anisotropic/isotropic tensor grid of points is generated. The program then expects function values at the grid points.

Next, a tensor order index set based on the number of samples in each dimension is generated, subject to the constraints in Equation 5. For each index set entry, the vector $\mathbf{b}$ is computed through the SVD of matrix F. The singular values are given by the diagonal matrix of the SVD.

We define excess singular values to be singular values that are below 10^{-16}. This represents machine double-precision epsilon and noise level for a vector with $\|.\|$ one.

Numerator and denominator combinations that have excess singular values are over-fitted, and introduce spurious pole-zero pairs as alluded to earlier. Hence, these numerator and denominator combinations are not candidates for approximation/interpolation.

For candidates that do not have excess singular values, L_2 error norms are computed, using a quadrature rule. It is because of this computation that input tensor grids must correspond to tensor products of univariate quadrature rules. In order to ensure that the error norm is computed accurately, sample stencils are typically of the form of Gauss-Legendre or Chebyshev points. This quadrature rule has a certain error associated with it, and varies depending on the complexity of the function and the order of the stencils. Once the error norms of all permissible index set values (without excess singular value candidates) are computed, they are ranked from the lowest to the highest. The lowest value is our optimal candidate.

This computation for all feasibile indices makes the algorithm particularly expensive – as the numbers of dimensions and orders increase, the size of the index set grows exponentially. Work on reducing this curse of dimensionality, by choosing index sets corresponding to particular numerator-denominator pairs is still ongoing.

C. Addressing Discontinuities

While the above architecture is stable for continuous functions, it tends to yield degenerate solutions for lower-order discontinuous functions. To address this, a limit on the maximum absolute value of the function at any point in its domain is set. This is based on either a known function evaluation or an estimate of function response. In any case, the following is added between steps 7 and 8:

```
1  Compute response surfaces of the top few L2 norm candidates
2  If the function limit criterion is violated, eliminate that candidate
3  From the remaining candidates, the one with the lowest L2 norm is selected
```

This criterion ensures that for a given function, discontinuous or otherwise, a stable rational interpolant and approximant can be computed. Tests carried out on a range of analytical functions will now be presented.

V. Numerical Experiments

A. Continuous Functions

We begin with various two-dimensional test functions, shown in Table 1. Even for polynomials, these test functions are challenging as they feature functions with exponential growth, sharp peaks and oscillations.

Figure 3 shows the convergence behavior of the rational method compared to traditional pseudospectral and collocation approaches for the bi-variate test functions. The $\log_{10}$ of the absolute value of the L_2 approximation error is shown on the y-axis, with the corresponding orders on the x-axis. Here the orders refer to the number of interpolation points used to obtain both polynomial and rational approximations. For instance, order 7 uses $7^2 = 49$ points, order 8 uses $8^2 = 64$ points and so on.

The pseudospectral approach utilizes a Legendre polynomial basis where the coefficients are computed using a Fourier-series type approach, by integrating function values at Gauss-Legendre quadrature points with Legendre polynomials. Details on the pseudospectral method can be found in Constantine et al.[9] The collocation approach uses a Lagrange cardinal function on a grid of Chebyshev points, while the rational

Table 1. Bi-variate test functions defined on $x = [-1, 1]$ and $y = [-1, 1]$

No.	Function
1	$f(x, y) = \exp(x + y)$
2	$f(x, y) = 1/(2 + 16(x - 0.1)^2 + 25(y - 0.1)^2)$
3	$f(x, y) = 3x^2 - 2xy + 5y^3$
4	$f(x, y) = \sin(3x - 1.5) + \cos(3y - 3)$

method uses Chebyshev polynomials in both the numerator and denominator on the same grid of Chebyshev points. During the generation of the polynomial and rational approximants, both methods are given the same function information, that is, the same number of points and corresponding function values.

As can be seen from the convergence diagrams, rationals either match the convergence rates of polynomials (Figure 3(c)) and (d)) or provide an ameliorated convergence profile. For the case of a fraction, the method exactly approximates the function, and, for the exponential function, a deviation from the polynomial convergence at orders $[N, N] > 8$ is observed.

The convergence of the approximation for both polynomial and rational formulations can also be gauged by an inspection of the expansion coefficients. Figure 4 plots the polynomial (a, d, g, j) coefficients along with those of the rational numerator (b, e, h, k) and rational denominator (c, f, i, l) for the bi-variate functions in Figure 3. We note that, for the exponential function, the rational coefficients exhibit a more rapid decay. By $N = [7, 8]$, the rational coefficients reach 10^{-15}, while, even for $N = [10, 10]$, the pseudospectral coefficients have values of the order 10^{-5}.

The greatest difference in coefficient values is for the $f(x, y) = 1/(2+16(x-0.1)^2+25(y-0.1)^2)$ fraction. Here the L_2 norm error for the polynomials is of the order 10^{-3}, while for the rationals it is 10^{-16}. Surface plots of the two approximants reveal that the rational method is able to accurately capture the peak at $(x, y) = (0.1, 0.1)$, while polynomials are not.

Another point should be noted from Figure 3(b). At $N = [3, 3]$ the L_2 norm error for the computed rational approximant is higher than that of the polynomials. This occurs because of a difference between the approximation and interpolation problem statements, as alluded to in Section II. For $N = [3, 3]$, the ranking algorithm computes numerator and denominator coefficients that yield the best possible interpolant. However, in this case it turns out that the interpolant obtained is not a good candidate for approximation. Figure 5 explains this in more detail.

As discussed earlier, the ranking algorithm sorts the set of all possible numerator and denominator coefficients according to their L_2 interpolation error norm. Combinations with the lowest L_2 error are ranked higher than those with larger errors. From this sorted set, the first entry, that is the one with the lowest L_2 error, is used as the rational approximant. In Figure 5, the approximation and interpolation errors of the top 20 entries are plotted. Here, the interpolation problem errors are sorted, such that an index set value of 1 corresponds to the lowest L_2 norm interpolation error. Ideally, this should also yield the lowest L_2 norm approximation error. However, as shown in the close up in Figure 5(b), this is not always the case. This discrepancy tends to fade away as the order of the interpolant increases, as seen in Figure 3(a). The reason for the error is that the singular value decomposition will yield the discretized least-squares solution over the interpolation nodes, and not the least-squares in equation 1.

In the same spirit as Figure 3, convergence histories of three-dimensional extensions of the bi-variate functions are presented in Figure 6. Compared to Figure 3, rapid convergence to numerical precision is observed when approximating polynomial and fractions. For the exponential function (Figure 6(a)), at orders below than $N = [6, 6, 6]$ the pseudospectral method provides lower errors. At higher orders, the methods seem to be aligned. The mild non-linearity of the convergence plot for the rational approximation method here is attributed to the offset between the best interpolation and best approximation solutions.

It should be noted that the discrepancy between the pseudospectral and rational methods may be attributed to the fact that we are restricting ourselves to anisotropic tensors for the numerator and denominator coefficients, as opposed to using random elements as was shown in Figure 1. This may be viewed as a constraint on the overall degree of freedom of the polynomials. This will be investigated further in future work.

Four-dimensional convergence tests on a variety of functions were also carried out. The $\log_{10}$ errors for

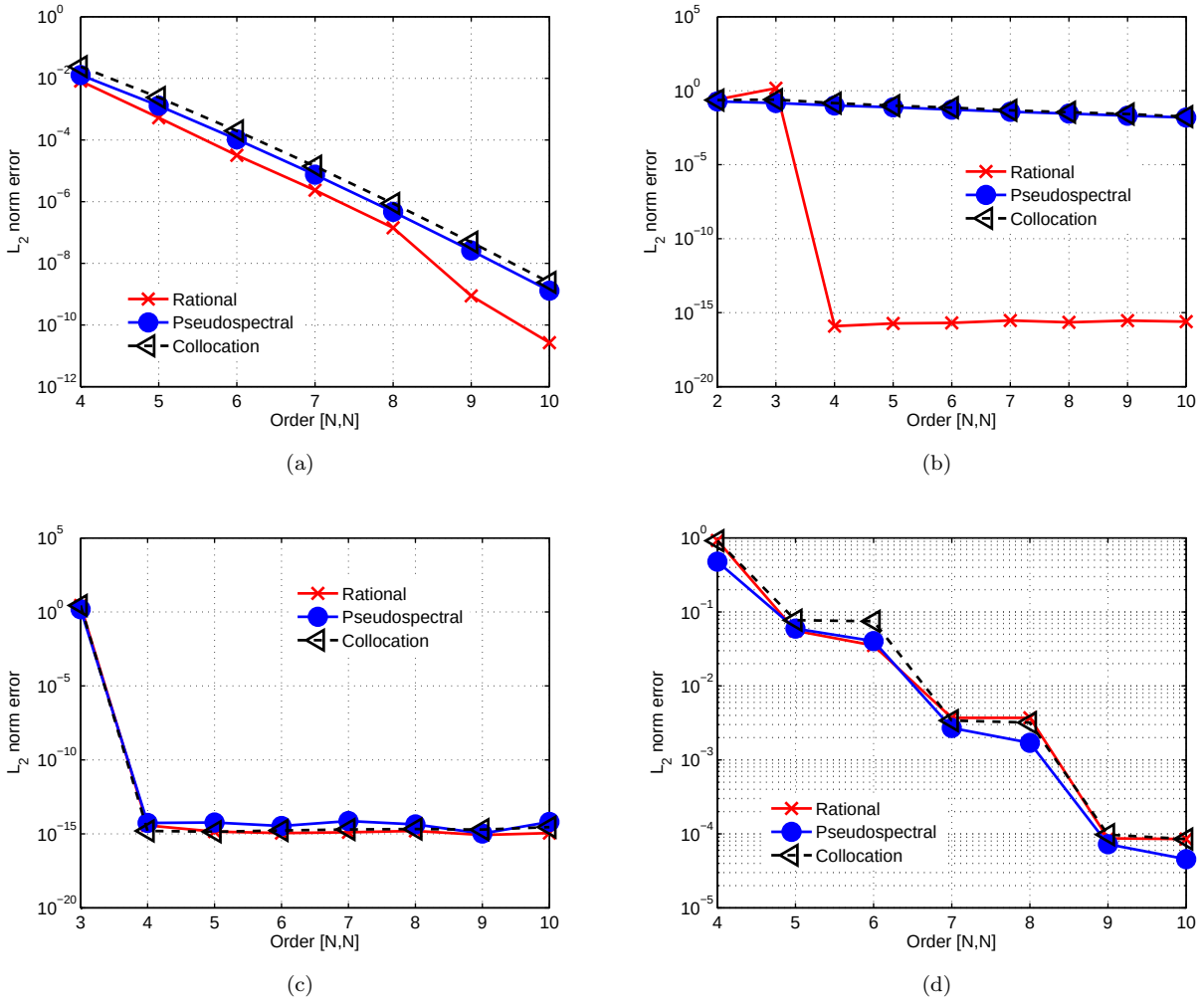


Figure 3. L_2 **norm convergence for rational, pseudospectral and collocation methods:** **(a)** $f(x,y) = \exp(x+y)$; **(b)** $f(x,y) = 1/(2+16(x-0.1)^2+25(y-0.1)^2)$; **(c)** $f(x,y) = 3x^2 - 2xy + 5y^3$; **(d)** $f(x,y) = \sin(3x-1.5) + \cos(3y-3)$

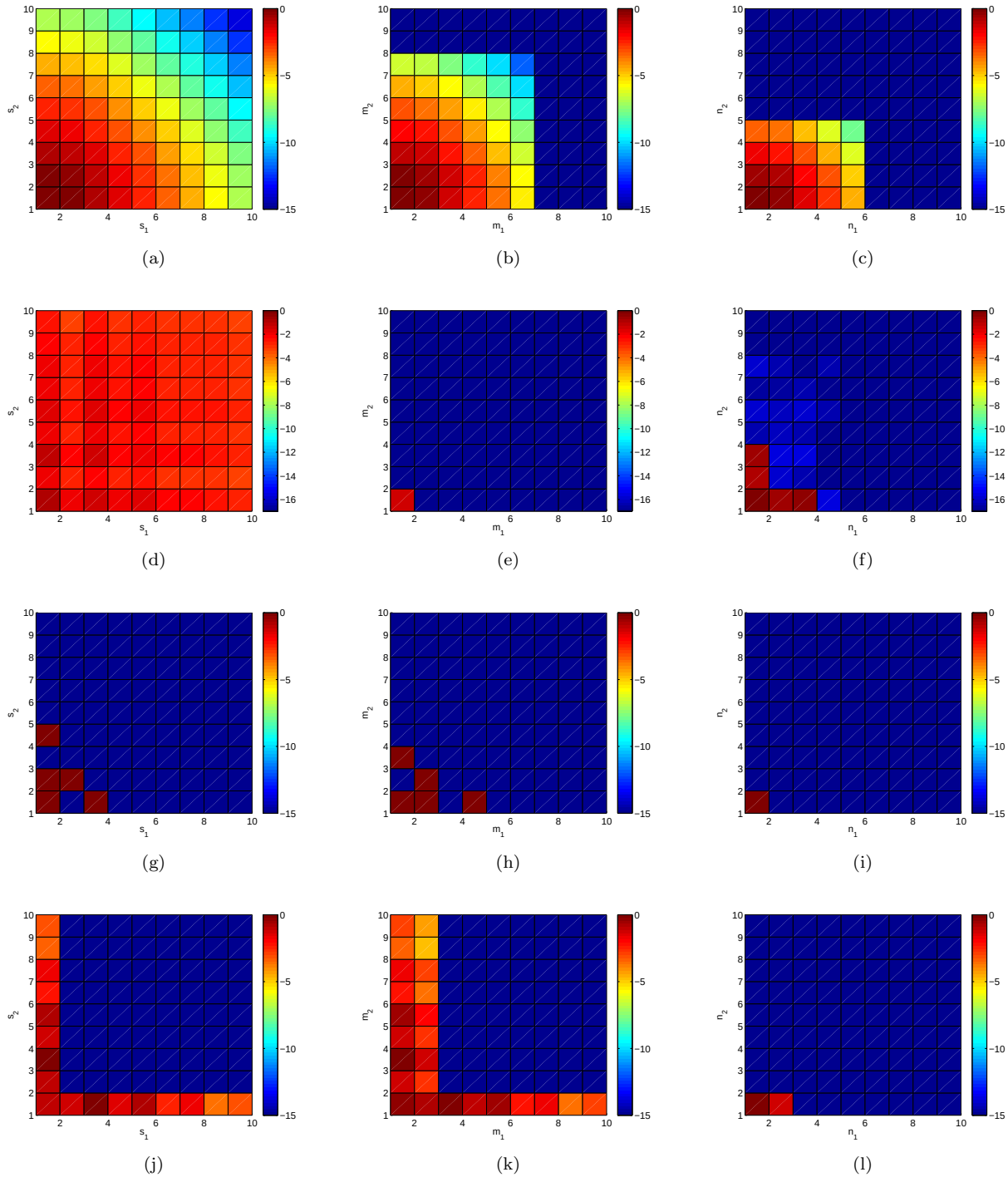


Figure 4. Pseudospectral polynomial coefficients (a, d, g, j) along with those of the rational numerator (b, e, h, k) and rational denominator (c, f, i, l) for: (a–c) $f(x,y) = \exp(x+y)$**; (d–f)** $f(x,y) = 1/(2+16(x-0.1)^2+25(y-0.1)^2)$**; (g–i)** $f(x,y) = 3x^2 - 2xy + 5y^3$**; (j–l)** $f(x,y) = \sin(3x - 1.5) + \cos(3y - 3)$

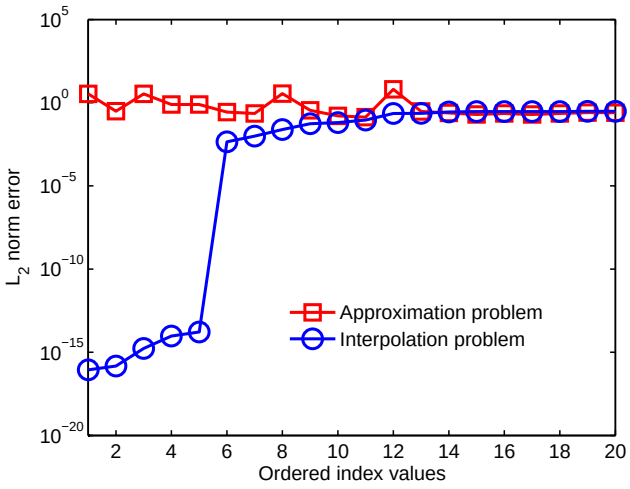


Figure 5. Difference between interpolation and approximation error

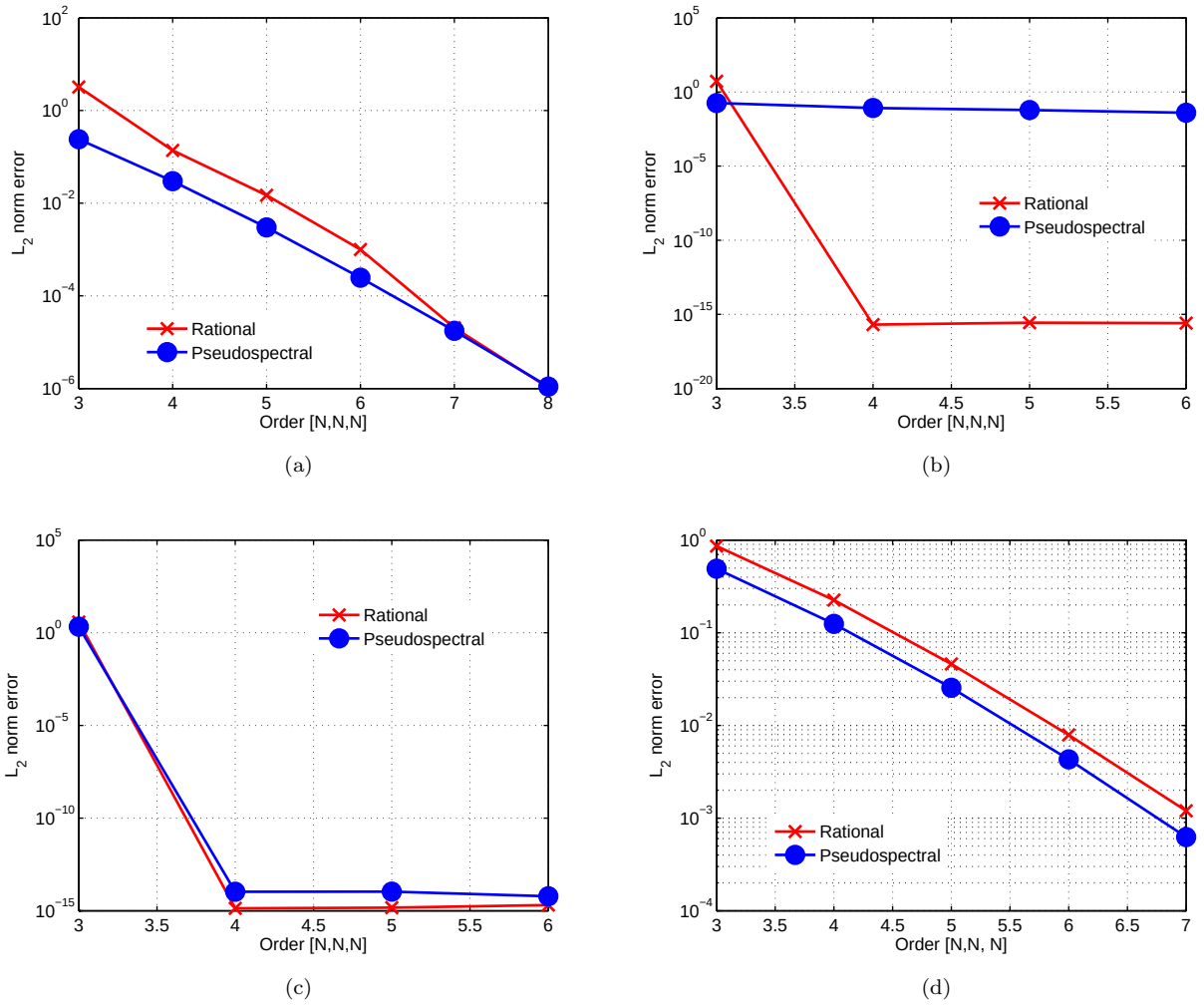
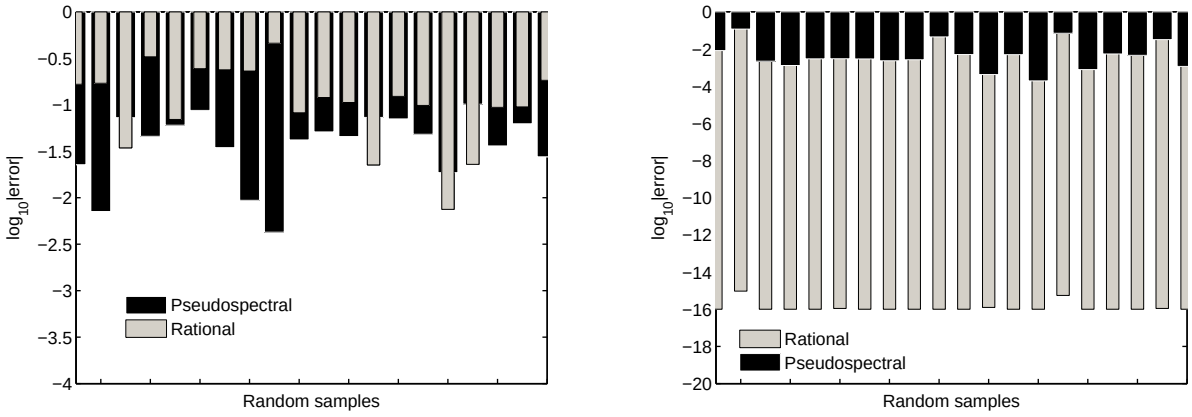


Figure 6. L_2 **norm convergence for rational, pseudospectral and collocation methods: (a)** $f(x,y,z) = \exp(x+y+z)$**; (b)** $f(x,y,z) = 1/(2+16(x-0.1)^2+25(y-0.1)^2+9(z-0.1)^2)$**; (c)** $f(x,y,z) = 3x^2y - 2xyz + 5y^3 + z^2$**; (d)** $f(x,y,z) = \sin(2x+y) - \cos(2z-y)$

random samples within the domain are plotted in Figure 7 for two of the functions explored. Comparable error margins are observed for the trigonometric function, with both methods providing errors in the range of $10^{-0.5}$ to 10^{-2}. For the fraction approximation in Figure 7, we once again observe that the rational method reaches numerical precision while the polynomial method struggles to get errors below 10^{-3}.



(a) (b)

Figure 7. Comparison of $\log_{10}|error|$ for two quad-variate functions using 256 points on the hypercube $[-1, 1]$: (a) $f(x, y, z, w) = \sin(2x + 2y + 2z + 2w)$; (b) $f(x, y, z, w) = 1/(2 + 16(x - 0.1)^2 + 25(y - 0.1)^2 + 9(z - 0.1)^2 + 36(w - 0.1)^2)$

The results in Figure 7 and earlier figures, demonstrate that the current methodology can be readily extended to higher-dimensional problems, while ensuring considerable accuracy for continuous functions. Now we briefly look into discontinuous functions.

B. Discontinuous Functions

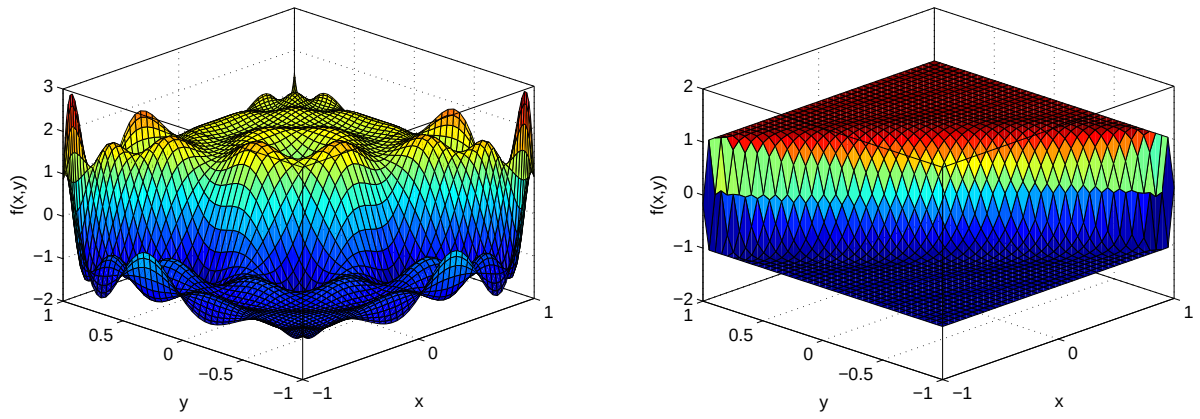
One of the main advantages of rational methods in uncertainty quantification is the ability to capture discontinuities such as shocks in the stochastic space. Chantrasmi et al.[2] demonstrated that rational methods are able to outperform polynomials for problems with shocks as the latter succumb to the Gibbs phenomenon. To demonstrate that our method can handle discontinuities we carry out a brief investigation with discontinuous functions.

Figure 8 plots responses surfaces obtained via the pseudospectral and rational methods for the function $f(x, y) = \text{sign}(x + y)$, using 100 points. An upper limit of $f(x, y)_{\max} = 1.5$ was used for the rational approximation.

From Figure 8, it is clear that the polynomial expansion succumbs to the Gibbs phenomenon, leading to large errors throughout the domain. The rational method, on the other hand, provides an accurate representation of the original function – the discontinuity along the axis is clearly visible. Plots of the coefficients in Figure 9 also reveal the extent of coefficient decay in the expansion.

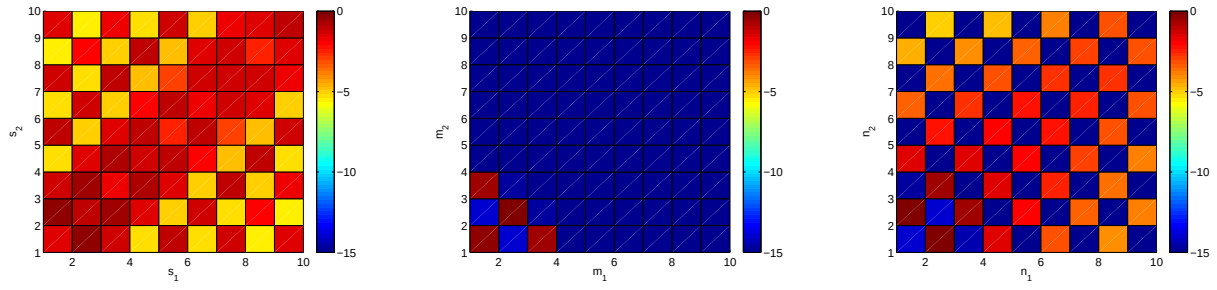
Conclusions

In this work, a high-dimensional, rational interpolant has been presented. Numerical experiments shed light on the remarkable convergence properties of multivariate rationals for both continuous and discontinuous functions. The findings detailed in this study are part of the initial phase of an ongoing investigation into the applicability of rational methods to uncertainty quantification. There still remains work to be done in obtaining a more efficient search algorithm for determining optimal coefficients, as well as extending sample points to sparse grids.



(a) (b)

Figure 8. Approximation of $f(x,y) = \text{sign}(x+y)$ with $N = [10, 10]$ using the (a) Pseudospectral and (b) Rational methods. No filters are used here



(a) (b) (c)

Figure 9. Comparison of expansion coefficients for $f(x,y) = \text{sign}(x+y)$: (a) Pseudospectral; (b) Rational numerator; (c) Rational denominator

Acknowledgements

The authors are grateful to Rolls-Royce plc for their permission to publish this work. This study was funded through a Dorothy Hodgkin Postgraduate Award, under a joint partnership between EPSRC and Rolls-Royce plc. The authors would like to thank St. Edmund's College, Cambridge and the University of Cambridge Department of Engineering for their support.

References

[1]Chantrasmi, T. and Iaccarino, G., "Forward and Backward Uncertainty Propagation for Discontinuous System Response using the Padé-Legendre Method," *International Journal of Uncertainty Quantification*, Vol. 2, 2011, pp. 123–145.

[2]Chantrasmi, T., Doostan, A., and Iaccarino, G., "Padé-Legendre Approximant for Uncertainty Quantification with Discontinuous Response Surfaces," *Journal of Computational Physics*, Vol. 228, 2009, pp. 7159–7180.

[3]Gonnet, P., Guettel, S., and Trefethen, L. N., "Robust Padé Approximation via SVD," *SIAM Review*, Vol. 55, No. 1, 2012, pp. 101–117.

[4]Pachón, R., Gonnet, P., and van Deun, J., "Fast and Stable Rational Interpolation in Roots of Unity and Chebyshev Points," *SIAM Journal of Numerical Analysis*, Vol. 50, 2012, pp. 1713–1734.

[5]Hesthaven, J., Kaber, S., and Lurati, L., "Padé-Legendre Interpolants for Gibbs Reconstruction," *SIAM Journal of Scientific Computing*, Vol. 28, No. 2-3, 2006, pp. 337–359.

[6]van Deun, J. and Trefethen, L. N., "A Robust Implementation of the Caratheodory-Fejer Method for Rational Approximation," *BIT Numerical Math*, Vol. 51, 2011, pp. 1039–1050.

[7]Gonnet, P., Pachón, R., and Trefethen, L. N., "Robust Rational Interpolation and Least-Squares," *Electronic Transactions on Numerical Analysis*, Vol. 38, 2011, pp. 146–167.

[8]Gonnet, P., "Robust Rational Interpolation in Two Dimensions," *3rd Dolomites Workshop on Constructive Approximation and Applications (DWCAA12)*, 2012, p. 82.

[9]Constantine, P. G., Eldred, M. S., and Phipps, E. T., "Sparse Pseudospectral Approximation Method," *Computer Methods in Applied Mechanics and Engineering*, Vol. 229, 2012, pp. 1–12.