An approximate factorisation of three bivariate Bernstein basis polynomials defined in a triangular domain

Martin Bourne, [a,⋆] Joab R. Winkler, [a] Yi Su [b]

[a] *The University of Sheffield, Department of Computer Science, Regent Court, 211 Portobello, Sheffield S1 4DP, United Kingdom*

[b] *Institute of High Performance Computing, A*STAR, 1 Fusionopolis Way, 16-16 Connexis North, Singapore 138632*

`mbourne1@sheffield.ac.uk`, `j.r.winkler@sheffield.ac.uk`, `suyi@ihpc.a-star.edu.sg`

Abstract

This paper considers an approximate factorisation of three bivariate Bernstein basis polynomials that are defined in a triangular domain. This problem is important for the computation of the intersection points and curves of surfaces in computer-aided design systems, and it reduces to the determination of an approximate greatest common divisor (AGCD) $d(y)$ of the polynomials. The Sylvester matrix and its subresultant matrices of these three polynomials are formed and it is shown that there are four forms of these matrices. The most difficult part of the computation is the determination of the degree of $d(y)$ because it reduces to the determination of the rank loss of these matrices. This computation is made harder by the presence of trinomial terms in the Bernstein basis functions because they cause the entries of the matrices to span many orders of magnitude. The adverse numerical effects of this wide range of magnitudes of the entries of the four forms of the Sylvester matrix and its subresultant matrices are mitigated by processing the polynomials before these matrices are formed. It is shown that significantly improved results are obtained if the polynomials are processed before computations are performed on their Sylvester matrices and subresultant matrices.

Key words: Bivariate Bernstein basis polynomial; Sylvester resultant matrix and subresultant matrices; approximate greatest common divisor

⋆ Martin Bourne was supported by a studentship from The Agency for Science, Technology and Research (A*STAR), Singapore, and The University of Sheffield.

1 Introduction

The computation of the greatest common divisor (GCD) of two univariate polynomials arises in several applications, including signal processing, image processing and control theory. More recent work has considered the computation of the GCD of two multivariate polynomials [4–9,11,13] and these problems are extended in this paper by consideration of an approximate factorisation of three bivariate Bernstein basis polynomials that are defined in a triangular domain. This computation arises in the determination of the points and curves of intersection of three surfaces in computer-aided design systems [4]. These intersections of the surfaces

$$\hat{f}(x,y)=0, \qquad \hat{g}(x,y)=0 \qquad \text{and} \qquad \hat{h}(x,y)=0,$$

where $\hat{f}(x,y), \hat{g}(x,y)$ and $\hat{h}(x,y)$ are bivariate Bernstein polynomials, are given by the irreducible factors of the GCD of these polynomials. For brevity, the independent variables (x,y) will be deleted from the notation used for polynomials, and these polynomials will therefore be denoted $\hat{f}, \hat{g}$ and $\hat{h}$ respectively, that is, $\hat{f}=\hat{f}(x,y), \hat{g}=\hat{g}(x,y)$ and $\hat{h}=\hat{h}(x,y)$.

A bivariate Bernstein polynomial in a triangular domain is considered in Section 2, and the Sylvester matrix of the polynomials $\hat{f}, \hat{g}$ and $\hat{h}$ is considered in Section 3. The degree t of their GCD $\hat{d}_t = \hat{d}_t(x,y)$ is equal to the rank loss of this matrix, but it is shown that this matrix is not unique because it, and its subresultant matrices, have four variants. The structures of these variants are considered in Sections 3 and 4, and it is shown that t cannot be computed reliably, especially in the presence of noise. Considerably improved results for the degree and coefficients of the GCD are obtained when the polynomials are processed before computations are performed on them, and these preprocessing operations are considered in Section 5. The computation of the coefficients of the GCD and coprime polynomials is considered in Section 6, and the complexity of the algorithms is discussed in Section 7. Examples of an approximate factorisation of $\hat{f}, \hat{g}$ and $\hat{h}$ in the presence of noise are in Section 8 and the paper is summarised in Section 9.

2 Bivariate Bernstein polynomials in a triangular domain

A bivariate Bernstein basis polynomial $\hat{f}$, of total degree m, in the triangular domain Δ,

$$\Delta: \qquad 0 \le x \le 1, \quad 0 \le y \le 1, \quad 0 \le 1-x-y \le 1,$$

is given by

$$\hat{f} = \sum_{i_1+i_2=0}^{m} \hat{a}_{i_1,i_2} B^m_{i_1,i_2} = \sum_{i_1+i_2=0}^{m} \hat{a}_{i_1,i_2} \binom{m}{i_1, i_2} (1-x-y)^{m-i_1-i_2} x^{i_1} y^{i_2}, \qquad (1)$$

where the bivariate Bernstein basis functions $B^m_{i_1,i_2} = B^m_{i_1,i_2}(x,y)$ are

$$B^m_{i_1,i_2} = \binom{m}{i_1, i_2} (1-x-y)^{m-i_1-i_2} x^{i_1} y^{i_2}, \qquad (2)$$

and

$$\binom{m}{i_1, i_2} = \binom{m}{i_2, i_1} = \frac{m!}{i_1! i_2! (m-i_1-i_2)!}.$$

The polynomial $\hat{f}$ has $\binom{m+2}{2}$ coefficients and it follows from (1) that it can be written as the sum of $m+1$ polynomials $\hat{f}_k = \hat{f}_k(x,y), k = 0, \ldots, m$, each of which is of degree m,

$$\begin{aligned} \hat{f} &= \sum_{i_1+i_2=0} \hat{a}_{i_1,i_2} B^m_{i_1,i_2} + \sum_{i_1+i_2=1} \hat{a}_{i_1,i_2} B^m_{i_1,i_2} + \cdots + \sum_{i_1+i_2=m} \hat{a}_{i_1,i_2} B^m_{i_1,i_2} \\ &= \sum_{k=0}^{m} \sum_{i_1+i_2=k} \hat{a}_{i_1,i_2} B^m_{i_1,i_2} \\ &= \sum_{k=0}^{m} \sum_{j=0}^{k} \hat{a}_{k-j,j} \binom{m}{k-j, j} (1-x-y)^{m-k} x^{k-j} y^j \\ &= \sum_{k=0}^{m} \hat{f}_k, \end{aligned} \qquad (3)$$

where

$$\hat{f}_k = \hat{f}_k(x,y) = \sum_{j=0}^{k} \hat{a}_{k-j,j} \binom{m}{k-j, j} (1-x-y)^{m-k} x^{k-j} y^j, \quad k = 0, \ldots, m. \qquad (4)$$

2.1 Vector representation and multiplication

This section considers vector and matrix representations of the polynomial $\hat{f}(x,y)$ because they are required for the development of the Sylvester matrices and their subresultant matrices of three bivariate Bernstein basis polynomials. The polynomial $\hat{f}(x,y)$ is represented by the vector $\hat{\mathbf{f}} \in \mathbb{R}^{\binom{m+2}{2}}$,

$$\hat{\mathbf{f}} = \left[\hat{\mathbf{f}}_0^T \;\; \hat{\mathbf{f}}_1^T \;\; \cdots \;\; \hat{\mathbf{f}}_m^T \right]^T,$$

where $\hat{\mathbf{f}}_k \in \mathbb{R}^{k+1}$, $k = 0, \ldots, m$, contains the coefficients of $\hat{f}_k$, which is defined in (4),

$$\hat{\mathbf{f}}_k = \left[\hat{a}_{k,0} \;\; \hat{a}_{k-1,1} \;\; \cdots \;\; \hat{a}_{0,k} \right]^T. \tag{5}$$

Consider the bivariate polynomials $\hat{f}$ and $\hat{g}$, of total degrees m and n respectively, that are defined in the triangular domain Δ. The polynomial $\hat{f}$ is defined in (1) and the polynomial $\hat{g} = \hat{g}(x,y)$ is

$$\hat{g} = \sum_{i+j=0}^{n} \hat{b}_{i,j} B_{i,j}^{n},$$

where $\hat{b}_{i,j}$ are the coefficients and the basis functions $B_{i,j}^{n}$ are defined in (2). The polynomial $\hat{g}$ can be written as the sum of the polynomials $\hat{g}_k = \hat{g}_k(x,y), k = 0, \ldots, n$,

$$\hat{g} = \hat{g}_0 + \hat{g}_1 + \cdots + \hat{g}_n,$$

where the basis functions of each polynomial $\hat{g}_k$, which is of degree n, are $B_{k-j,j}^{n}$ and its $k+1$ coefficients are $\hat{b}_{k-j,j}, j = 0 \ldots, k$,

$$\hat{g}_k = \sum_{j=0}^{k} \hat{b}_{k-j,j} B_{k-j,j}^{n} = \sum_{j=0}^{k} \hat{b}_{j,k-j} B_{j,k-j}^{n}, \qquad k = 0, \ldots, n.$$

The polynomial $\hat{h} = \hat{h}(x,y) = \hat{f}\hat{g}$ is of degree $m+n$ and it can be written as the sum of the polynomials $\hat{h}_j = \hat{h}_j(x,y), j = 0, \ldots, m+n$, as shown in (3) for $\hat{f}$ and the $m+1$ polynomials $\hat{f}_k, k = 0, \ldots, m$,

$$\hat{h} = \sum_{i+j=0}^{m+n} \hat{c}_{i,j} \binom{m+n}{i,j} (1-x-y)^{m+n-i-j} x^i y^j = \sum_{j=0}^{m+n} \hat{h}_j,$$

where each polynomial $\hat{h}_j, j = 0, \ldots, m+n$, is of degree $m+n$ and $\hat{h}_j$ has $j+1$ coefficients,

$$\hat{h}_j = \sum_{i=0}^{j} \hat{c}_{i,j-i} \binom{m+n}{i,j-i} (1-x-y)^{m+n-j} x^i y^{j-i}. \tag{6}$$

It follows from the definition of $\hat{h}$ that

$$\begin{aligned}\hat{h} = &\left(\hat{f}_0\hat{g}_0\right) + \left(\hat{f}_0\hat{g}_1 + \hat{f}_1\hat{g}_0\right) + \left(\hat{f}_2\hat{g}_0 + \hat{f}_1\hat{g}_1 + \hat{f}_0\hat{g}_2\right) + \cdots + \\ &\left(\hat{f}_m(x,y)\hat{g}_n(x,y)\right),\end{aligned} \tag{7}$$

where a typical polynomial $\hat{f}_s\hat{g}_t = \hat{f}_s(x,y)\hat{g}_t(x,y)$ in this sum is written as

$$\begin{aligned}\hat{f}_s\hat{g}_t = &\sum_{i=0}^{s}\sum_{j=0}^{t} \hat{a}_{i,s-i}\hat{b}_{j,t-j} \binom{m}{i,s-i}\binom{n}{j,t-j} \times \\ &(1-x-y)^{m+n-s-t} x^{i+j} y^{s+t-i-j} \\ = &\sum_{i=0}^{s}\sum_{j=0}^{t} \frac{\hat{a}_{i,s-i}\hat{b}_{j,t-j}\binom{m}{i,s-i}\binom{n}{j,t-j}}{\binom{m+n}{i+j,s+t-i-j}} B^{m+n}_{i+j,s+t-i-j},\end{aligned} \tag{8}$$

and the basis functions $B^{m+n}_{i+j,s+t-i-j}$ are defined in (2). The vector of the coefficients of this polynomial can be written as the matrix-vector product,

$$\dot{C}_t(\hat{f}_s)\hat{\mathbf{g}}_t = \left(\dot{D}^{-1}_{m+n,s+t}\dot{T}_t(\hat{f}_s)\dot{Q}_{n,t}\right)\hat{\mathbf{g}}_t, \tag{9}$$

where $\dot{C}_t(\hat{f}_s) \in \mathbb{R}^{(s+t+1)\times(t+1)}$, the convolution matrix of $\hat{f}_s(x,y)$, is equal to the product of three matrices, $\dot{D}^{-1}_{m+n,s+t}$, $\dot{T}_t(\hat{f}_s)$ and $\dot{Q}_{n,t}$. The diagonal matrix $\dot{D}^{-1}_{m+n,s+t} \in \mathbb{R}^{(s+t+1)\times(s+t+1)}$ is

$$\dot{D}^{-1}_{m+n,s+t} = \operatorname{diag}\left[\begin{array}{cccc} \frac{1}{\binom{m+n}{s+t,0}} & \frac{1}{\binom{m+n}{s+t-1,1}} & \cdots & \frac{1}{\binom{m+n}{0,s+t}}\end{array}\right], \tag{10}$$

the Tœplitz matrix $\dot{T}_t(\hat{f}_s) \in \mathbb{R}^{(s+t+1)\times(t+1)}$ is

$$
\left[\begin{array}{cccc}
\hat{a}_{s,0}\binom{m}{s,0} & & & \\
\hat{a}_{s-1,1}\binom{m}{s-1,1} & \hat{a}_{s,0}\binom{m}{s,0} & & \\
\vdots & \hat{a}_{s-1,1}\binom{m}{s-1,1} & \ddots & \\
\vdots & \vdots & \ddots & \hat{a}_{s,0}\binom{m}{s,0} \\
\hat{a}_{1,s-1}\binom{m}{1,s-1} & \vdots & \ddots & \hat{a}_{s-1,1}\binom{m}{s-1,1} \\
\hat{a}_{0,s}\binom{m}{0,s} & \hat{a}_{1,s-1}\binom{m}{1,s-1} & \ddots & \vdots \\
 & \hat{a}_{0,s}\binom{m}{0,s} & \ddots & \vdots \\
 & & \ddots & \hat{a}_{1,s-1}\binom{m}{1,s-1} \\
 & & & \hat{a}_{0,s}\binom{m}{0,s}
\end{array}\right], \tag{11}
$$

and the diagonal matrix $\dot{Q}_{n,t} \in \mathbb{R}^{(t+1)\times(t+1)}$ is

$$
\dot{Q}_{n,t} = \operatorname{diag}\left[\begin{array}{cccc} \binom{n}{t,0} & \binom{n}{t-1,1} & \cdots & \binom{n}{0,t} \end{array}\right]. \tag{12}
$$

The structure of $\hat{\mathbf{g}}_t \in \mathbb{R}^{t+1}$ in (9) is similar to the structure of the vector of the coefficients of $\hat{f}$ in (5),

$$
\hat{\mathbf{g}}_t = \left[\begin{array}{cccc} \hat{b}_{t,0} & \hat{b}_{t-1,1} & \cdots & \hat{b}_{0,t} \end{array}\right]^T. \tag{13}
$$

Example 2.1 considers the form of a typical term in the polynomial $\hat{h}$.

Example 2.1 Let $s = t = 2$, and thus (8) becomes

$$
\hat{f}_2\hat{g}_2 = \sum_{i=0}^{2}\sum_{j=0}^{2} \frac{\hat{a}_{i,2-i}\hat{b}_{j,2-j}\binom{m}{i,2-i}\binom{n}{j,2-j}}{\binom{m+n}{i+j,4-i-j}} B^{m+n}_{i+j,4-i-j},
$$

and the coefficients of this polynomial are

$$\frac{1}{\binom{m+n}{0,4}}\left(\hat{a}_{0,2}\binom{m}{0,2}\hat{b}_{0,2}\binom{n}{0,2}\right),$$
$$\frac{1}{\binom{m+n}{1,3}}\left(\hat{a}_{0,2}\binom{m}{0,2}\hat{b}_{1,1}\binom{n}{1,1}+\hat{a}_{1,1}\binom{m}{1,1}\hat{b}_{0,2}\binom{n}{0,2}\right),$$
$$\frac{1}{\binom{m+n}{2,2}}\left(\hat{a}_{0,2}\binom{m}{0,2}\hat{b}_{2,0}\binom{n}{2,0}+\hat{a}_{1,1}\binom{m}{1,1}\hat{b}_{1,1}\binom{n}{1,1}+\hat{a}_{2,0}\binom{m}{2,0}\hat{b}_{0,2}\binom{n}{0,2}\right),$$
$$\frac{1}{\binom{m+n}{3,1}}\left(\hat{a}_{1,1}\binom{m}{1,1}\hat{b}_{2,0}\binom{n}{2,0}+\hat{a}_{2,0}\binom{m}{2,0}\hat{b}_{1,1}\binom{n}{1,1}\right),$$
$$\frac{1}{\binom{m+n}{4,0}}\left(\hat{a}_{2,0}\binom{m}{2,0}\hat{b}_{2,0}\binom{n}{2,0}\right).$$

The vector of these coefficients can be written as the matrix-vector product (9),

$$\operatorname{diag}\left[\begin{array}{ccccc}\frac{1}{\binom{m+n}{4,0}} & \frac{1}{\binom{m+n}{3,1}} & \frac{1}{\binom{m+n}{2,2}} & \frac{1}{\binom{m+n}{1,3}} & \frac{1}{\binom{m+n}{0,4}}\end{array}\right]\times$$
$$\begin{bmatrix} \hat{a}_{2,0}\binom{m}{2,0} & & \\ \hat{a}_{1,1}\binom{m}{1,1} & \hat{a}_{2,0}\binom{m}{2,0} & \\ \hat{a}_{0,2}\binom{m}{0,2} & \hat{a}_{1,1}\binom{m}{1,1} & \hat{a}_{2,0}\binom{m}{2,0} \\ & \hat{a}_{0,2}\binom{m}{0,2} & \hat{a}_{1,1}\binom{m}{1,1} \\ & & \hat{a}_{0,2}\binom{m}{0,2} \end{bmatrix}\begin{bmatrix}\binom{n}{2,0} & & \\ & \binom{n}{1,1} & \\ & & \binom{n}{0,2}\end{bmatrix}\begin{bmatrix}\hat{b}_{2,0}\\ \hat{b}_{1,1}\\ \hat{b}_{0,2}\end{bmatrix}.$$

$\square$

Each of the terms in (7) can be written in the form in (9), and thus the sum of all the terms can be written as the product of a block matrix and a vector,

$$C_n(\hat{f})\hat{\mathbf{g}} = \hat{\mathbf{h}}, \tag{14}$$

where the convolution matrix $C_n(\hat{f}) \in \mathbb{R}^{\binom{m+n+2}{2}\times\binom{n+2}{2}}$ is

$$C_n(\hat{f}) = \begin{bmatrix} \dot{C}_0(\hat{f}_0) & & & \\ \dot{C}_0(\hat{f}_1) & \dot{C}_1(\hat{f}_0) & & \\ \vdots & \dot{C}_1(\hat{f}_1) & \ddots & \\ \vdots & \vdots & \ddots & \dot{C}_n(\hat{f}_0) \\ \dot{C}_0(\hat{f}_m) & \vdots & \ddots & \dot{C}_n(\hat{f}_1) \\ & \dot{C}_1(\hat{f}_m) & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & \dot{C}_n(\hat{f}_m) \end{bmatrix},$$

and each matrix $\dot{C}_t(\hat{f}_s)$, $s = 0, \ldots, m; t = 0, \ldots, n$, is the convolution matrix (9). The matrix $C_n(\hat{f})$ can be written in a form that is the extension of the product $\dot{D}^{-1}_{m+n,s+t}\dot{T}_t(\hat{f}_s)\dot{Q}_{n,t}$ (9) to its block form,

$$C_n(\hat{f}) = D^{-1}_{m+n} T_n(\hat{f}) Q_n, \tag{15}$$

where the block diagonal matrix $D^{-1}_{m+n} \in \mathbb{R}^{\binom{m+n+2}{2} \times \binom{m+n+2}{2}}$ is

$$D^{-1}_{m+n} = \text{diag}\left[\, \dot{D}^{-1}_{m+n,0} \;\; \dot{D}^{-1}_{m+n,1} \;\; \cdots \;\; \dot{D}^{-1}_{m+n,m+n} \,\right],$$

and the structure of each matrix $\dot{D}^{-1}_{m+n,i}$, $i = 0, \ldots, m+n$, is shown in (10). The matrix $T_n(\hat{f})$ is

$$T_n(\hat{f}) = \begin{bmatrix} \dot{T}_0(\hat{f}_0) & & & \\ \dot{T}_0(\hat{f}_1) & \dot{T}_1(\hat{f}_0) & & \\ \vdots & \dot{T}_1(\hat{f}_1) & \ddots & \\ \vdots & \vdots & \ddots & \dot{T}_n(\hat{f}_0) \\ \dot{T}_0(\hat{f}_m) & \vdots & \ddots & \dot{T}_n(\hat{f}_1) \\ & \dot{T}_1(\hat{f}_m) & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & \dot{T}_n(\hat{f}_m) \end{bmatrix}, \tag{16}$$

where each of the matrices $\dot{T}_j(\hat{f}_i) \in \mathbb{R}^{(i+j+1)\times(j+1)}$ has the same structure as the Tœplitz matrix in (11). The block diagonal matrix $Q_n \in \mathbb{R}^{\binom{n+2}{2} \times \binom{n+2}{2}}$ is

$$Q_n = \mathrm{diag}\left[\begin{array}{cccc} \dot{Q}_{n,0} & \dot{Q}_{n,1} & \cdots & \dot{Q}_{n,n} \end{array}\right], \tag{17}$$

where $\dot{Q}_{n,t}$ is defined in (12). The vector $\hat{\mathbf{g}}$ in (14) is

$$\hat{\mathbf{g}} = \left[\begin{array}{cccc} \hat{\mathbf{g}}_0^T & \hat{\mathbf{g}}_1^T & \cdots & \hat{\mathbf{g}}_n^T \end{array}\right]^T \in \mathbb{R}^{\binom{n+2}{2}},$$

where $\hat{\mathbf{g}}_j \in \mathbb{R}^{j+1}$ is defined in (13), and the vector $\hat{\mathbf{h}}$ contains the coefficients of $\hat{h}(x,y)$,

$$\hat{\mathbf{h}} = \left[\begin{array}{cccc} \hat{\mathbf{h}}_0^T & \hat{\mathbf{h}}_1^T & \cdots & \hat{\mathbf{h}}_{m+n}^T \end{array}\right]^T \in \mathbb{R}^{\binom{m+n+2}{2}},$$

where $\hat{\mathbf{h}}_j$ contains the coefficients of $\hat{h}_j(x,y)$, which is defined in (6).

Example 2.2 considers the form of (14) for $m = n = 2$.

Example 2.2 The polynomial $\hat{f}$ of degree $m = 2$ is

$$\begin{aligned} \hat{f} &= \sum_{i+j=0}^{2} \hat{a}_{i,j} \binom{2}{i,j} (1-x-y)^{2-i-j} x^i y^j \\ &= \hat{a}_{0,0} B_{0,0}^2 + \hat{a}_{1,0} B_{1,0}^2 + \hat{a}_{0,1} B_{0,1}^2 + \hat{a}_{2,0} B_{2,0}^2 + \hat{a}_{1,1} B_{1,1}^2 + \hat{a}_{0,2} B_{0,2}^2. \end{aligned}$$

It follows from (3) that $\hat{f}$ can also be written as the sum of the polynomials $\hat{f}_0 = \hat{f}_0(x,y)$, $\hat{f}_1 = \hat{f}_1(x,y)$ and $\hat{f}_2 = \hat{f}_2(x,y)$,

$$\begin{aligned} \hat{f}_0 &= \hat{a}_{0,0} \binom{2}{0,0} (1-x-y)^2, \\ \hat{f}_1 &= \left(\hat{a}_{1,0} \binom{2}{1,0} x + \hat{a}_{0,1} \binom{2}{0,1} y \right) (1-x-y), \\ \hat{f}_2 &= \hat{a}_{2,0} \binom{2}{2,0} x^2 + \hat{a}_{1,1} \binom{2}{1,1} xy + \hat{a}_{0,2} \binom{2}{0,2} y^2. \end{aligned}$$

The polynomial $\hat{g}$ of degree $n = 2$ is

$$\hat{g} = \sum_{i+j=0}^{2} \hat{b}_{i,j} \binom{2}{i,j} (1-x-y)^{2-i-j} x^i y^j,$$

and it is equal to the sum of the polynomials $\hat{g}_0 = \hat{g}_0(x,y)$, $\hat{g}_1 = \hat{g}_1(x,y)$ and

$\hat{g}_2 = \hat{g}_2(x, y)$,

$$\hat{g}_0 = \hat{b}_{0,0} \binom{2}{0, 0} (1 - x - y)^2,$$
$$\hat{g}_1 = \left(\hat{b}_{1,0} \binom{2}{1, 0} x + \hat{b}_{0,1} \binom{2}{0, 1} y \right) (1 - x - y),$$
$$\hat{g}_2 = \hat{b}_{2,0} \binom{2}{1, 0} x^2 + \hat{b}_{1,1} \binom{2}{1, 1} xy + \hat{b}_{0,2} \binom{2}{0, 2} y^2.$$

The polynomial $\hat{h}$ is equal to the product of $\hat{f}$ and $\hat{g}$,

$$\begin{aligned}
\hat{h} &= \hat{f}\hat{g} \\
&= \left(\hat{f}_0 + \hat{f}_1 + \hat{f}_2 \right) (\hat{g}_0 + \hat{g}_1 + \hat{g}_2) \\
&= \hat{f}_0\hat{g}_0 + \left(\hat{f}_0\hat{g}_1 + \hat{f}_1\hat{g}_0 \right) + \left(\hat{f}_0\hat{g}_2 + \hat{f}_1\hat{g}_1 + \hat{f}_2\hat{g}_0 \right) + \left(\hat{f}_1\hat{g}_2 + \hat{f}_2\hat{g}_1 \right) + \hat{f}_2\hat{g}_2 \\
&= \sum_{j=0}^{4} \hat{h}_j.
\end{aligned}$$

This polynomial multiplication can be written as the matrix-vector product $C_2(\hat{f})\hat{\mathbf{g}} = \hat{\mathbf{h}}$,

$$\begin{bmatrix}
\dot{C}_0(\hat{f}_0) & & \\
\dot{C}_0(\hat{f}_1) & \dot{C}_1(\hat{f}_0) & \\
\dot{C}_0(\hat{f}_2) & \dot{C}_1(\hat{f}_1) & \dot{C}_2(\hat{f}_0) \\
 & \dot{C}_1(\hat{f}_2) & \dot{C}_2(\hat{f}_1) \\
 & & \dot{C}_2(\hat{f}_2)
\end{bmatrix}
\begin{bmatrix} \hat{\boldsymbol{g}}_0 \\ \hat{\boldsymbol{g}}_1 \\ \hat{\boldsymbol{g}}_2 \end{bmatrix}
=
\begin{bmatrix} \hat{\boldsymbol{h}}_0 \\ \hat{\boldsymbol{h}}_1 \\ \hat{\boldsymbol{h}}_2 \\ \hat{\boldsymbol{h}}_3 \\ \hat{\boldsymbol{h}}_4 \end{bmatrix}.$$

□

The next section uses the matrix-vector form of the product of two bivariate Bernstein basis polynomials to form the Sylvester matrices and subresultant matrices of three bivariate Bernstein basis polynomials.

3 The Sylvester matrix of three bivariate Bernstein basis polynomials

The Sylvester matrix of three univariate Bernstein basis polynomials is considered in [3] and it is shown that it can take four forms, one 3×3 block matrix form and three 2×3 block matrix forms. The extension of these forms from univariate polynomials to bivariate polynomials follows easily and it will therefore be considered briefly.

Let the degrees of the Bernstein basis polynomials $\hat{f}, \hat{g}$ and $\hat{h}$ be m, n and p respectively. The polynomials have common divisors $\hat{d}_k = \hat{d}_k(x, y)$, where the degree of $\hat{d}_k$ is k, the degree of their GCD is t, $\hat{d}_k$ is not unique for $k = 1, \dots, t-1$, and $\hat{d}_t$ is unique up to a non-zero constant. It follows that

$$\frac{\hat{f}}{\hat{u}_k} = \frac{\hat{g}}{\hat{v}_k} = \frac{\hat{h}}{\hat{w}_k} = \hat{d}_k,$$

where $\hat{u}_k$, $\hat{v}_k$ and $\hat{w}_k$ are cofactor polynomials of degrees $m-k$, $n-k$ and $p-k$, respectively. This yields three equations,

$$\hat{f}\hat{v}_k - \hat{g}\hat{u}_k = 0, \tag{18}$$
$$\hat{f}\hat{w}_k - \hat{h}\hat{u}_k = 0, \tag{19}$$
$$\hat{h}\hat{v}_k - \hat{g}\hat{w}_k = 0, \tag{20}$$

which can be written in matrix form,

$$\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})\mathbf{x}_k = \mathbf{0}. \tag{21}$$

This equation has non-zero solutions for $k = 1, \dots, t$, and the coefficient matrix $\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})$ is the 3×3 block matrix of the kth subresultant matrix of $\hat{f}, \hat{g}$ and $\hat{h}$. Each block in this matrix arises from the product of two polynomials whose matrix form is given in (15), and thus

$$\tilde{S}_k(\hat{f}, \hat{g}, \hat{h}) = \tilde{D}_k^{-1}\tilde{T}_k(\hat{f}, \hat{g}, \hat{h})\tilde{Q}_k, \tag{22}$$

which is equal to

$$\begin{aligned} &\operatorname{diag}\left[\, D_{m+n-k}^{-1} \;\; D_{m+p-k}^{-1} \;\; D_{n+p-k}^{-1} \,\right] \begin{bmatrix} T_{n-k}(\hat{f}) & & T_{m-k}(\hat{g}) \\ & T_{p-k}(\hat{f}) & T_{m-k}(\hat{h}) \\ T_{n-k}(\hat{h}) & -T_{p-k}(\hat{g}) & \end{bmatrix} \times \\ &\operatorname{diag}\left[\, Q_{n-k} \;\; Q_{p-k} \;\; Q_{m-k} \,\right]. \end{aligned} \tag{23}$$

The block diagonal matrix

$$\tilde{D}_k^{-1} = \operatorname{diag}\left[\, D_{m+n-k}^{-1} \;\; D_{m+p-k}^{-1} \;\; D_{n+p-k}^{-1} \,\right], \tag{24}$$

is of order

$$\binom{m+n-k+2}{2}+\binom{m+p-k+2}{2}+\binom{n+p-k+2}{2},$$

D_{m+n-k}^{-1} is equal to

$$\operatorname{diag}\left[\binom{m+n-k}{0,0}\middle|\binom{m+n-k}{1,0}\ \binom{m+n-k}{0,1}\middle|\cdots\middle|\binom{m+n-k}{m+n-k,0}\ \cdots\ \binom{m+n-k}{0,m+n-k}\right],$$

and D_{m+p-k}^{-1} and D_{n+p-k}^{-1} have the same structure as D_{m+n-k}^{-1}, and $\mathbf{x}_k$ contains the coefficients of the cofactor polynomials $\hat{u}_k$, $\hat{v}_k$ and $\hat{w}_k$,

$$\mathbf{x}_k=\begin{bmatrix}\hat{\mathbf{v}}_k^T & \hat{\mathbf{w}}_k^T & -\hat{\mathbf{u}}_k^T\end{bmatrix}^T, \qquad k=1,\ldots,t. \tag{25}$$

The matrix $\tilde{T}_k(\hat{f},\hat{g},\hat{h})$ is

$$\begin{bmatrix} T_{n-k}(\hat{f}) & & T_{m-k}(\hat{g}) \\ & T_{p-k}(\hat{f}) & T_{m-k}(\hat{h}) \\ T_{n-k}(\hat{h}) & -T_{p-k}(\hat{g}) & \end{bmatrix}, \tag{26}$$

where the matrices $T_{q-k}(\hat{f})$, $q\in\{m,n,p\}$, have the same structure as the matrix $T_n(\hat{f})$ in (16). The block diagonal matrix $\tilde{Q}_k$ is of order

$$\binom{n-k+2}{2}+\binom{p-k+2}{2}+\binom{m-k+2}{2},$$

and given by

$$\tilde{Q}_k=\operatorname{diag}\begin{bmatrix}Q_{n-k} & Q_{p-k} & Q_{m-k}\end{bmatrix}, \tag{27}$$

where the diagonal matrices Q_{n-k}, Q_{p-k} and Q_{m-k} have the same structure as the matrix Q_n that is defined in (17).

It follows from (15) that

$$C_{n-k}(\hat{f})=D_{m+n-k}^{-1}T_{n-k}(\hat{f})Q_{n-k},$$

and thus from (22), (23), (24), (26) and (27), $\tilde{S}_k(\hat{f},\hat{g},\hat{h})$ can be written as

$$\tilde{S}_k(\hat{f},\hat{g},\hat{h}) = \begin{bmatrix} \mathcal{R}_{a,k} \\ \mathcal{R}_{b,k} \\ \mathcal{R}_{c,k} \end{bmatrix} = \begin{bmatrix} C_{n-k}(\hat{f}) & & C_{m-k}(\hat{g}) \\ & C_{p-k}(\hat{f}) & C_{m-k}(\hat{h}) \\ C_{n-k}(\hat{h}) & -C_{p-k}(\hat{g}) & \end{bmatrix},$$

where

$$\mathcal{R}_{a,k} = \begin{bmatrix} C_{n-k}(\hat{f}) & 0_{\binom{m+n-k+2}{2},\binom{p-k+2}{2}} & C_{m-k}(\hat{g}) \end{bmatrix}, \tag{28}$$

$$\mathcal{R}_{b,k} = \begin{bmatrix} 0_{\binom{m+p-k+2}{2},\binom{n-k+2}{2}} & C_{p-k}(\hat{f}) & C_{m-k}(\hat{h}) \end{bmatrix}, \tag{29}$$

$$\mathcal{R}_{c,k} = \begin{bmatrix} C_{n-k}(\hat{h}) & -C_{p-k}(\hat{g}) & 0_{\binom{n+p-k+2}{2},\binom{m-k+2}{2}} \end{bmatrix}. \tag{30}$$

Alternatively, any two of the three equations (18), (19) and (20) are sufficient to describe the system completely because the third equation can be derived from these two equations. This gives rise to three variants of the 2×3 block subresultant matrices.

Variant 1: Equations (18) and (19) are written in matrix form as

$$\hat{S}_k(\hat{f},\hat{g},\hat{h})\mathbf{x}_{k,1} = \mathbf{0}, \tag{31}$$

which has non-zero solutions for $k = 1, \ldots, t$, and the solution vector $\mathbf{x}_{k,1}$ has the same structure as $\mathbf{x}_k$, which is defined in (25). The matrix $\hat{S}_k(\hat{f},\hat{g},\hat{h})$ is

$$\begin{bmatrix} \mathcal{R}_{a,k} \\ \mathcal{R}_{b,k} \end{bmatrix} = \begin{bmatrix} C_{n-k}(\hat{f}) & & C_{m-k}(\hat{g}) \\ & C_{p-k}(\hat{f}) & C_{m-k}(\hat{h}) \end{bmatrix},$$

where $\mathcal{R}_{a,k}$ and $\mathcal{R}_{b,k}$ are defined in (28) and (29) respectively.

Variant 2: Equations (18) and (20) are written in matrix form as

$$\hat{S}_k(\hat{g},\hat{f},\hat{h})\mathbf{x}_{k,2} = \mathbf{0}, \tag{32}$$

which has non-zero solutions for $k = 1, \ldots, t$, and $\mathbf{x}_{k,2}$ is obtained by reordering the entries of $\mathbf{x}_{k,1}$,

$$\mathbf{x}_{k,2} = \begin{bmatrix} \hat{\mathbf{u}}_k^T & \hat{\mathbf{w}}_k^T & -\hat{\mathbf{v}}_k^T \end{bmatrix}^T.$$

The matrix $\hat{S}_k(\hat{g},\hat{f},\hat{h})$ is

$$\begin{bmatrix} \tilde{\mathcal{R}}_{a,k} \\ \tilde{\mathcal{R}}_{c,k} \end{bmatrix} = \begin{bmatrix} C_{m-k}(\hat{g}) & & C_{n-k}(\hat{f}) \\ & C_{p-k}(\hat{g}) & C_{n-k}(\hat{h}) \end{bmatrix},$$

where $\tilde{\mathcal{R}}_{a,k}$ and $\tilde{\mathcal{R}}_{c,k}$ are variations of $\mathcal{R}_{a,k}$ and $\mathcal{R}_{c,k}$, which are defined in (28) and (30) respectively.

Variant 3: Equations (19) and (20) are written in matrix form as

$$\hat{S}_k(\hat{h}, \hat{g}, \hat{f})\mathbf{x}_{k,3} = \mathbf{0}, \tag{33}$$

which has non-trivial solutions for $k = 1, \ldots, t$, and $\mathbf{x}_{k,3}$ is obtained by reordering the entries of $\mathbf{x}_{k,1}$,

$$\mathbf{x}_{k,3} = \begin{bmatrix} \hat{\mathbf{u}}_k^T & \hat{\mathbf{v}}_k^T & -\hat{\mathbf{w}}_k^T \end{bmatrix}^T.$$

The matrix $\hat{S}_k(\hat{h}, \hat{g}, \hat{f})$ is the third variant of the 2×3 block subresultant matrix and it is given by

$$\begin{bmatrix} \tilde{\mathcal{R}}_{b,k} \\ \tilde{\mathcal{R}}_{c,k} \end{bmatrix} = \begin{bmatrix} C_{m-k}(\hat{h}) & & C_{p-k}(\hat{f}) \\ & C_{n-k}(\hat{h}) & C_{p-k}(\hat{g}) \end{bmatrix},$$

where $\tilde{\mathcal{R}}_{a,k}$ and $\tilde{\mathcal{R}}_{c,k}$ are variations of $\mathcal{R}_{a,k}$ and $\mathcal{R}_{c,k}$, respectively.

The pairwise GCDs of $\hat{f}, \hat{g}$ and $\hat{h}$ are

$$\hat{d}_a = \hat{d}_a(x,y) = \text{GCD}\left(\hat{f}, \hat{g}\right), \tag{34}$$
$$\hat{d}_b = \hat{d}_b(x,y) = \text{GCD}\left(\hat{f}, \hat{h}\right), \tag{35}$$
$$\hat{d}_c = \hat{d}_c(x,y) = \text{GCD}\left(\hat{g}, \hat{h}\right), \tag{36}$$

and thus (31), (32) and (33) are obtained by considering two of the three pairwise two-polynomial GCD problems.

The computation of the degree t of the GCD of $\hat{f}, \hat{g}$ and $\hat{h}$ from the subresultant matrices $\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})$ and $\hat{S}_k(\hat{f}^*, \hat{g}^*, \hat{h}^*)$, $k = 1, \ldots, \min(m, n, p)$, where $\hat{S}_k(\hat{f}^*, \hat{g}^*, \hat{h}^*)$ denotes that the order of $\hat{f}, \hat{g}$ and $\hat{h}$ in the arguments of $\hat{S}_k(\cdot)$ is arbitrary, is considered in Theorem 3.1. Each ordering yields one of the three variants of the subresultant matrices discussed above.

Theorem 3.1 The degree of the GCD of $\hat{f}, \hat{g}$ and $\hat{h}$ is equal to the largest index k such that the subresultant matrices $\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})$ and $\hat{S}_k(\hat{f}^*, \hat{g}^*, \hat{h}^*)$ are rank deficient, $k = 1, \ldots, \min(m, n, p)$.

The proof of the theorem follows easily from the same result for univariate polynomials [1–3]. In particular, if the degree of the GCD of $\hat{f}, \hat{g}$ and $\hat{h}$ is t, then these polynomials have more than one common divisor of degree $k = 1, \ldots, t-1$, and one common divisor (the GCD) of degree t, and thus $\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})$ and $\hat{S}_k(\hat{f}^*, \hat{g}^*, \hat{h}^*)$ are rank deficient for $k = 1, \ldots, t$. The polynomials do not, however, have a common divisor of degree $k = t+1, \ldots, \min(m, n, p)$, and thus $\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})$ and $\hat{S}_k(\hat{f}^*, \hat{g}^*, \hat{h}^*)$ have full rank for $k = t+1, \ldots, \min(m, n, p)$.

Theorem 3.1 considers the GCD of three polynomials, but their coefficients are corrupted by noise in practical problems. It is therefore assumed that the given polynomials are coprime, but that they posses an approximate greatest common divisor (AGCD) that is near the GCD of the exact forms of the polynomials. An AGCD of three polynomials is defined in Definition 3.1.

Definition 3.1 (An AGCD) A polynomial $d = d(x, y)$ of degree t is an AGCD of $\hat{f}, \hat{g}$ and $\hat{h}$ if it is the polynomial of maximum degree that is an exact divisor of $\hat{f} + \delta\hat{f}, \hat{g} + \delta\hat{g}$ and $\hat{h} + \delta\hat{h}$ for perturbations $\left\|\delta\hat{f}\right\| \leq \varepsilon_f$, $\|\delta\hat{g}\| \leq \varepsilon_g$ and $\left\|\delta\hat{h}\right\| \leq \varepsilon_h$, and $\left\|\delta\hat{f}\right\|^2 + \|\delta\hat{g}\|^2 + \left\|\delta\hat{h}\right\|^2$ is minimised over all polynomials of degree t.

Example 3.1 shows that numerical problems may arise if the singular values of the Sylvester matrices and their subresultant matrices are used for the computation of the degree of the GCD of three polynomials. The cause of the problems is identified and the subsequent sections discuss the preprocessing operations on the polynomials that are required to guarantee a computationally reliable solution.

Example 3.1 Consider the Bernstein forms of the polynomials $\hat{f}, \hat{g}$ and $\hat{h}$, which are of degrees $m = 14, n = 14$ and $p = 14$ respectively,

$$\begin{aligned}
\hat{f} &= (x - 0.72)(x - 0.52)^2(x + 0.75)(y - 0.75)^2(y - 0.15)(y^2 - 1.7) \times \\
&\quad (x + y - 0.5)^5, \\
\hat{g} &= (x - 0.72)(x - 0.52)^2(x - 0.192)(y - 0.15)(x + y - 0.5)^5 \times \\
&\quad (y^2 - 1.7)(x^2 + y^2 + 0.7), \\
\hat{h} &= (x - 1.91987)^4(x - 0.72)(y - 0.15)(y^2 - 1.7)(x^2 + y^2 - 0.34)^3,
\end{aligned}$$

whose GCD $\hat{d}_t$ is of degree $t = 4$,

$$\hat{d}_t = (x - 0.72)(y - 0.15)(y^2 - 1.7).$$

The polynomials $\hat{f}$ and $\hat{g}$ have a GCD $\hat{d}_a$ of degree $t_a = 11$,

$$\hat{d}_a = (x - 0.72)(y - 0.15)(y^2 - 1.7)(x - 0.52)^2(x + y - 0.5)^5,$$

and $\hat{d}_b = \hat{d}_c = \hat{d}_t$, where d_a, d_b and d_c are defined in (34), (35) and (36) respectively.

The coefficients of $\hat{f}, \hat{g}$ and $\hat{h}$ were multiplied by 10^5, 10^5 and 10^{-5} respectively, and noise was then added, such that the coefficients of the inexact polynomials f, g and h were

$$\begin{aligned} a_{i_1,i_2} &= \hat{a}_{i_1,i_2} + \hat{a}_{i_1,i_2}\,\epsilon_{f,i_1,i_2}\,r_{f,i_1,i_2}, & i_1 + i_2 &= 0, \ldots, m, \\ b_{j_1,j_2} &= \hat{b}_{j_1,j_2} + \hat{b}_{j_1,j_2}\,\epsilon_{g,j_1,j_2}\,r_{g,j_1,j_2}, & j_1 + j_2 &= 0, \ldots, n, \\ c_{l_1,l_2} &= \hat{c}_{l_1,l_2} + \hat{c}_{l_1,l_2}\,\epsilon_{h,l_1,l_2}\,r_{h,l_1,l_2}, & l_1 + l_2 &= 0, \ldots, p, \end{aligned} \tag{37}$$

where $\{\epsilon_{f,i_1,i_2}\}$, $\{\epsilon_{g,j_1,j_2}\}$ and $\{\epsilon_{h,l_1,l_2}\}$ are uniformly distributed random variables in the interval $[10^{-8}, 10^{-6}]$, and $\{r_{f,i_1,i_2}\}$, $\{r_{g,j_1,j_2}\}$ and $\{r_{h,l_1,l_2}\}$ are uniformly distributed random variables in the interval $[-1, 1]$. The coefficients of the inexact polynomials are plotted in Figure 1 and it is seen that the coefficients of h are much smaller than the coefficients of f and g.

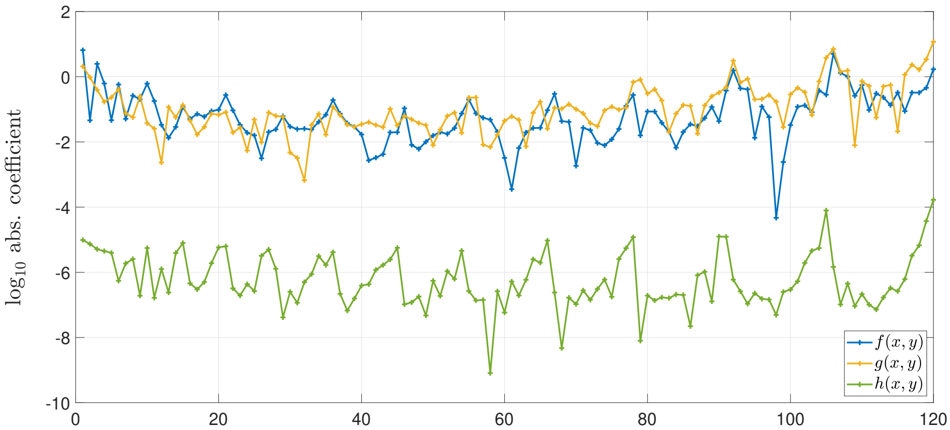


Fig. 1. The coefficients of f, g and h in Example 3.1.

The singular values of the 2×3 block subresultant matrix $\hat{S}_k(f, g, h)$ and the 3×3 block subresultant matrix $\tilde{S}_k(f, g, h)$ are shown in Figure 2. The figures suggest that the 11th subresultant matrix is the largest rank deficient subresultant matrix, and therefore the degree of an AGCD is $t = 11$. This is incorrect and is due to the coefficients of f and g being significantly larger than the coefficients of h, such that the block matrices $C_{m-k}(h)$ and $C_{n-k}(h)$ are approximately zero with respect to the block matrices that contain the coefficients of f and g. The results for the other 2×3 block subresultant matrices, $\hat{S}_k(g, f, h)$ and $\hat{S}_k(h, g, f)$, are very similar to the results in Figure 2. $\square$

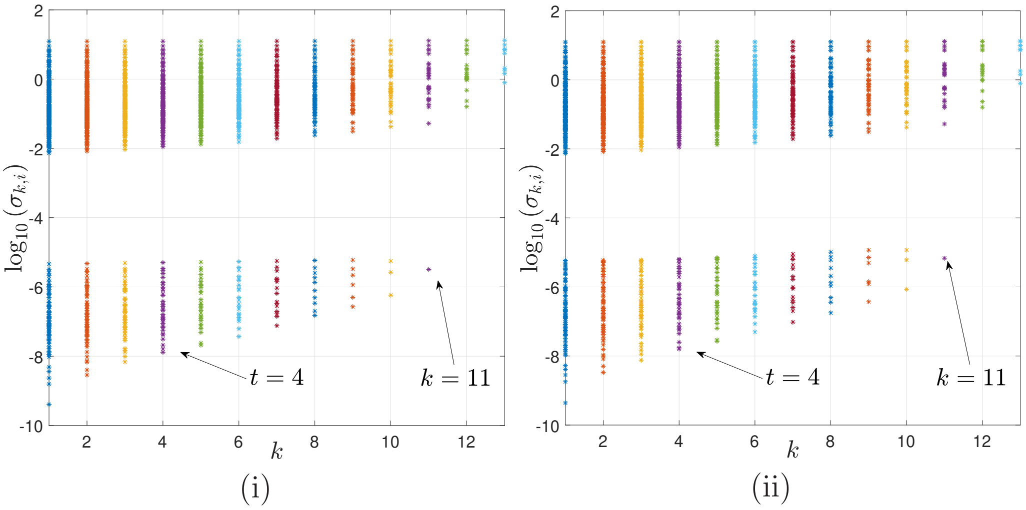


Fig. 2. The singular values of (i) $\hat{S}_k(f, g, h)$ and (ii) $\tilde{S}_k(f, g, h)$ in Example 3.1.

Example 3.1 shows that numerical problems arise when the entries of the subresultant matrices span many orders of magnitude [1–3]. The adverse numerical effects of these problems can be mitigated by processing the polynomials before computations are performed on them. The preprocessing operations are described in Section 5, and Example 3.1 is reproduced with the inclusion of the preprocessing operations. It is shown that significantly improved results are obtained when the polynomials are processed before computations are performed on their Sylvester matrices and subresultant matrices.

4 Entries of the variants of the Sylvester matrix and its subresultant matrices

It was shown in Section 3 that the GCD of $\hat{f}, \hat{g}$ and $\hat{h}$ can be calculated from the 3×3 block form and the three 2×3 block forms of their Sylvester matrix and its subresultant matrices.

Consider the 3×3 block matrix $\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})$, which is defined in (22) and for which the diagonal matrices $\tilde{D}_k^{-1}$ and $\tilde{Q}_k$ have full rank for all values of k. It follows that the degree t of the GCD of $\hat{f}, \hat{g}$ and $\hat{h}$ is equal to the rank loss of the following four matrices, which are obtained from $\tilde{S}_k(\hat{f}, \hat{g}, \hat{h})$:

$$\left\{ \tilde{T}_k(\hat{f}, \hat{g}, \hat{h}),\ \tilde{D}_k^{-1}\tilde{T}_k(\hat{f}, \hat{g}, \hat{h}),\ \tilde{T}_k(\hat{f}, \hat{g}, \hat{h})\tilde{Q}_k,\ \tilde{D}_k^{-1}\tilde{T}_k(\hat{f}, \hat{g}, \hat{h})\tilde{Q}_k \right\}. \tag{38}$$

The entries of these four forms are now considered:

(1) The first form is $\tilde{T}_k(\hat{f}, \hat{g}, \hat{h})$ and the first block matrix in this matrix is, from (26), $T_{n-k}(\hat{f})$, the non-zero entries of which are

$$\hat{a}_{i_1,i_2}\binom{m}{i_1, i_2}, \qquad i_1+i_2 = 0,\dots,m.$$

(2) The second form is $\tilde{D}_k^{-1}\tilde{T}_k(\hat{f},\hat{g},\hat{h})$ and the first block matrix in this matrix is, from (24) and (26), $D_{m+n-k}^{-1}T_{n-k}(\hat{f})$, the non-zero entries of which are

$$\frac{\hat{a}_{i_1,i_2}\binom{m}{i_1,i_2}}{\binom{m+n-k}{i_1+j_1,i_2+j_2}}, \qquad i_1+i_2=0,\dots,m, \qquad j_1+j_2=0,\dots,n-k.$$

(3) The third form is $\tilde{T}_k(\hat{f},\hat{g},\hat{h})\tilde{Q}_k$ and the first block matrix in this matrix is, from (26) and (27), $T_{n-k}(\hat{f})Q_{n-k}$, the non-zero entries of which are

$$\hat{a}_{i_1,i_2}\binom{m}{i_1, i_2}\binom{n-k}{j_1, j_2}, \quad i_1+i_2=0,\dots,m, \quad j_1+j_2=0,\dots,n-k.$$

(4) The fourth form is $\tilde{D}_k^{-1}\tilde{T}_k(\hat{f},\hat{g},\hat{h})\tilde{Q}_k$ and the first block matrix in this form is, from (24), (26) and (27), $D_{m+n-k}^{-1}T_{n-k}(\hat{f})Q_{n-k}$, the non-zero entries of which are

$$\frac{\hat{a}_{i_1,i_2}\binom{m}{i_1,i_2}\binom{n-k}{j_1,j_2}}{\binom{m+n-k}{i_1+j_1,i_2+j_2}}, \qquad i_1+i_2=0,\dots,m, \qquad j_1+j_2=0,\dots,n-k.$$

The entries of the three variants of the 2×3 block subresultant matrices follow identically, and the presence of potentially large trinomial terms may cause numerical issues because the entries of the matrices may differ by many orders of magnitude. It is shown in [3] that the form $\tilde{D}_k^{-1}\tilde{T}_k(\hat{f},\hat{g},\hat{h})\tilde{Q}_k$ in (38) yields the best results for the 3×3 block subresultant matrices because the ratio r of the maximum trinomial term to the minimum trinomial term assumes its minimum value from the set of four matrices (38), and similarly, this form also yields the best results for the 2×3 block subresultant matrices. Although these forms yield the best results, the ratio r may still be large, which may cause numerical problems. These numerical problems are mitigated by processing $\hat{f},\hat{g}$ and $\hat{h}$ before computations are performed on their Sylvester matrices and subresultant matrices. These preprocessing operations are considered in Section 5.

5 Preprocessing operations

This section describes the operations that are implemented on $\hat{f},\hat{g}$ and $\hat{h}$ before computations are performed on their Sylvester matrices and subresultant matrices. The preprocessing operations for univariate Bernstein basis polynomials are considered in [1–3] and the examples in these references show that the inclusion of these preprocessing operations yields significantly improved results because the rank deficiency of their Sylvester matrices and subresultant matrices is clearly defined.

The three preprocessing operations are now considered.

Normalisation The coefficients of $\hat{f}, \hat{g}$ and $\hat{h}$ are normalised by their geometric means. This normalisation is performed for each subresultant matrix because the entries of the kth subresultant matrix are functions of k due to the trinomial terms. The normalising constant depends on the Sylvester matrix and its subresultant matrices considered, that is, the 2×3 block matrices or the 3×3 block matrix.

Consider initially the 3×3 block matrix form. Since each of the three polynomials appears in two blocks, the normalised polynomials are

$$\bar{f}_k = \frac{\hat{f}}{\hat{\mathcal{G}}_k(\hat{f})}, \qquad \bar{g}_k = \frac{\hat{g}}{\hat{\mathcal{G}}_k(\hat{g})} \qquad \text{and} \qquad \bar{h}_k = \frac{\hat{h}}{\hat{\mathcal{G}}_k(\hat{h})},$$

where $\hat{\mathcal{G}}_k(\hat{f})$, $\hat{\mathcal{G}}_k(\hat{g})$ and $\hat{\mathcal{G}}_k(\hat{h})$ are the geometric means of the coefficients of $\hat{f}, \hat{g}$ and $\hat{h}$ respectively. The polynomial $\hat{f}$ appears in the block matrices $C_{n-k}(\hat{f})$ and $C_{p-k}(\hat{f})$, which are of dimensions

$$\binom{m+n-k+2}{2} \times \binom{n-k+2}{2},$$

and

$$\binom{m+p-k+2}{2} \times \binom{p-k+2}{2},$$

respectively, and thus the geometric mean of the coefficients of $\hat{f}$ in the 3×3 block subresultant matrix is

$$\hat{\mathcal{G}}_k(\hat{f}) = \left(\prod_{i_1+i_2=0}^{m} \prod_{j_1+j_2=0}^{n-k} \frac{\hat{a}_{i_1,i_2} \binom{m}{i_1,i_2}\binom{n-k}{j_1,j_2}}{\binom{m+n-k}{i_1+j_1,i_2+j_2}} \times \prod_{i_1+i_2=0}^{m} \prod_{j_1+j_2=0}^{p-k} \frac{\hat{a}_{i_1,i_2} \binom{m}{i_1,i_2}\binom{p-k}{j_1,j_2}}{\binom{m+p-k}{i_1+j_1,i_2+j_2}} \right)^{\frac{1}{r+c}}, \tag{39}$$

where

$$r = \binom{m+n-k+2}{2}\binom{n-k+2}{2},$$
$$c = \binom{m+p-k+2}{2}\binom{p-k+2}{2}.$$

The formulae for $\hat{\mathcal{G}}_k(\hat{g})$ and $\hat{\mathcal{G}}_k(\hat{h})$ are derived similarly.

Consider now the 2×3 block matrix forms of the Sylvester matrices and their subresultant matrices. For example, the polynomial $\hat{f}$ occurs twice in $\hat{S}_k(\hat{f}, \hat{g}, \hat{h})$ and thus the geometric mean of its coefficients is given in (39). The polynomials $\hat{g}$ and $\hat{h}$ occur in the block matrices $C_{m-k}(\hat{g})$ and $C_{m-k}(\hat{h})$ respectively, and thus the geometric means of their coefficients assume simpler forms.

The normalisation operation yields the polynomials

$$\begin{aligned}
\bar{f}_k &= \frac{\hat{f}}{\hat{\mathcal{G}}_k(\hat{f})} = \sum_{i_1+i_2=0}^{m} \bar{a}_{i_1,i_2} B^m_{i_1,i_2}, & \bar{a}_{i_1,i_2} &= \frac{\hat{a}_{i_1,i_2}}{\hat{\mathcal{G}}_k(\hat{f})}, \\
\bar{g}_k &= \frac{\hat{g}}{\hat{\mathcal{G}}_k(\hat{g})} = \sum_{i_1+i_2=0}^{n} \bar{b}_{i_1,i_2} B^n_{i_1,i_2}, & \bar{b}_{i_1,i_2} &= \frac{\hat{b}_{i_1,i_2}}{\hat{\mathcal{G}}_k(\hat{g})}, \\
\bar{h}_k &= \frac{\hat{h}}{\hat{\mathcal{G}}_k(\hat{h})} = \sum_{i_1+i_2=0}^{p} \bar{c}_{i_1,i_2} B^p_{i_1,i_2}, & \bar{c}_{i_1,i_2} &= \frac{\hat{c}_{i_1,i_2}}{\hat{\mathcal{G}}_k(\hat{h})}.
\end{aligned}$$

A parameter substitution The independent variables x and y are replaced by the independent variables ω_1 and ω_2,

$$x = \theta_1\omega_1 \qquad \text{and} \qquad y = \theta_2\omega_2, \tag{40}$$

where θ_1 and θ_2 are constants to be determined. It is shown in the sequel that the optimal values of these constants are computed for each subresultant matrix, that is, for each value of k. The criterion for the computation of these optimal values is discussed after the third preprocessing operation is considered.

Non-uniqueness of the GCD The GCD of two or more polynomials is defined to within an arbitrary non-zero scalar multiplier, and thus

$$\text{GCD}\,(\hat{f}, \hat{g}, \hat{h}) \sim \text{GCD}\,(\lambda\hat{f}, \hat{g}, \rho\hat{h}), \qquad \lambda, \rho \in \mathbb{R} \setminus \{0\},$$

where $\sim$ denotes equivalence to within an arbitrary non-zero scalar multiplier. As with the constants θ_1 and θ_2, the optimal values of λ and ρ are calculated for each value of k.

These three preprocessing operations yield the polynomials

$$
\begin{aligned}
\lambda \ddot{f}_k(\theta_1,\theta_2;\omega_1,\omega_2) &= \lambda \sum_{i_1+i_2=0}^{m} \bar{a}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{m}{i_1,\, i_2} \times \\
&\qquad (1-\theta_1\omega_1-\theta_2\omega_2)^{m-i_1-i_2}\,\omega_1^{i_1}\omega_2^{i_2}, \\
\ddot{g}_k(\theta_1,\theta_2;\omega_1,\omega_2) &= \sum_{i_1+i_2=0}^{n} \bar{b}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{n}{i_1,\, i_2} \times \\
&\qquad (1-\theta_1\omega_1-\theta_2\omega_2)^{n-i_1-i_2}\,\omega_1^{i_1}\omega_2^{i_2}, \\
\rho \ddot{h}_k(\theta_1,\theta_2;\omega_1,\omega_2) &= \rho \sum_{i_1+i_2=0}^{p} \bar{c}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{p}{i_1,\, i_2} \times \\
&\qquad (1-\theta_1\omega_1-\theta_2\omega_2)^{p-i_1-i_2}\,\omega_1^{i_1}\omega_2^{i_2}.
\end{aligned}
$$

The preprocessing operations introduce the constants $\theta_1, \theta_2, \lambda$ and ρ, and their optimal values must be computed for each value of k. The same criterion is used for the 2×3 and 3×3 block forms of the Sylvester matrix and its subresultant matrices, and it is therefore adequate to consider the computation of the optimal values of the constants for the 3×3 block forms of these matrices.

The sets of non-zero entries in the six matrices in the 3×3 block matrix $\tilde{S}_k(\lambda\ddot{f}_k, \ddot{g}_k, \rho\ddot{h}_k)$ (23) are

$$
D^{-1}_{m+n-k}T_{n-k}(\lambda\ddot{f}_k)Q_{n-k}:\ \mathcal{P}_{1,k}\left(\lambda,\theta_1,\theta_2\right) = \left\{ \frac{\left|\lambda\bar{a}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{m}{i_1,i_2}\binom{n-k}{j_1,j_2}\right|}{\binom{m+n-k}{i_1+j_1,i_2+j_2}} \right\},
$$

for $i_1+i_2 = 0,\ldots,m$, and $j_1+j_2 = 0,\ldots,n-k$,

$$
D^{-1}_{m+n-k}T_{m-k}(\ddot{g}_k)Q_{m-k}:\ \mathcal{P}_{2,k}\left(\theta_1,\theta_2\right) = \left\{ \frac{\left|\bar{b}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{n}{i_1,i_2}\binom{m-k}{j_1,j_2}\right|}{\binom{m+n-k}{i_1+j_1,i_2+j_2}} \right\},
$$

for $i_1+i_2 = 0,\ldots,n$, and $j_1+j_2 = 0,\ldots,m-k$,

$$
D^{-1}_{m+p-k}T_{p-k}(\lambda\ddot{f}_k)Q_{p-k}:\ \mathcal{P}_{3,k}\left(\lambda,\theta_1,\theta_2\right) = \left\{ \frac{\left|\lambda\bar{a}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{m}{i_1,i_2}\binom{p-k}{j_1,j_2}\right|}{\binom{m+p-k}{i_1+j_1,i_2+j_2}} \right\},
$$

for $i_1+i_2 = 0,\ldots,m$, and $j_1+j_2 = 0,\ldots,p-k$,

$$
D^{-1}_{m+p-k}T_{m-k}(\rho\ddot{h}_k)Q_{m-k}:\ \mathcal{P}_{4,k}\left(\rho,\theta_1,\theta_2\right) = \left\{ \frac{\left|\rho\bar{c}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{p}{i_1,i_2}\binom{m-k}{j_1,j_2}\right|}{\binom{m+p-k}{i_1+j_1,i_2+j_2}} \right\},
$$

for $i_1+i_2 = 0,\ldots,p$, and $j_1+j_2 = 0,\ldots,m-k$,

$$D_{n+p-k}^{-1}T_{n-k}(\rho\ddot{h}_k)Q_{n-k}: \ \mathcal{P}_{5,k}\left(\rho,\theta_1,\theta_2\right)=\left\{\frac{\left|\rho\bar{c}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{p}{i_1,i_2}\binom{n-k}{j_1,j_2}\right|}{\binom{n+p-k}{i_1+j_1,i_2+j_2}}\right\},$$

for $i_1+i_2=0,\dots,p$, and $j_1+j_2=0,\dots,n-k$, and

$$D_{n+p-k}^{-1}T_{p-k}(\ddot{g}_k)Q_{p-k}: \ \mathcal{P}_{6,k}\left(\theta_1,\theta_2\right)=\left\{\frac{\left|\bar{b}_{i_1,i_2}\theta_1^{i_1}\theta_2^{i_2}\binom{n}{i_1,i_2}\binom{p-k}{j_1,j_2}\right|}{\binom{n+p-k}{i_1+j_1,i_2+j_2}}\right\},$$

for $i_1+i_2=0,\dots,n$, and $j_1+j_2=0,\dots,p-k$.

The constants λ,ρ,θ_1 and θ_2 are chosen such that the ratio of the maximum entry to the minimum entry of the kth subresultant matrix is minimised,

$$\begin{aligned}(\lambda_k,\rho_k,\theta_{1,k},\theta_{2,k}) = \arg \min_{\lambda,\rho,\theta_1,\theta_2} & \\ &\left\{\frac{\max\left\{\max\{\mathcal{P}_{1,k}(\lambda,\theta_1,\theta_2)\},\max\{\mathcal{P}_{2,k}(\theta_1,\theta_2)\},\right.}{\min\left\{\min\{\mathcal{P}_{1,k}(\lambda,\theta_1,\theta_2)\},\min\{\mathcal{P}_{2,k}(\theta_1,\theta_2)\},\right.}\right.\\ &\frac{\max\{\mathcal{P}_{3,k}(\lambda,\theta_1,\theta_2)\},\max\{\mathcal{P}_{4,k}(\rho,\theta_1,\theta_2)\},}{\min\{\mathcal{P}_{3,k}(\lambda,\theta_1,\theta_2)\},\min\{\mathcal{P}_{4,k}(\rho,\theta_1,\theta_2)\},}\\ &\left.\frac{\left.\max\{\mathcal{P}_{5,k}(\rho,\theta_1,\theta_2)\},\max\{\mathcal{P}_{6,k}(\theta_1,\theta_2)\}\right\}}{\left.\min\{\mathcal{P}_{5,k}(\rho,\theta_1,\theta_2)\},\min\{\mathcal{P}_{6,k}(\theta_1,\theta_2)\}\right\}}\right\},\end{aligned}$$

for $k=1,\dots,\min(m,n,p)$. This minimisation problem is reduced to a linear programming problem, which must be solved for each subresultant matrix, that is, for each value of k, as shown in [12] for the two-polynomial GCD problem. In particular, the minimisation problem for each value of k follows identically from this problem,

$$\min z^T x \qquad \text{subject to} \qquad Ax \geq b,$$

because the vectors z,x and b, and the matrix A, are formed in the same way, but they are larger. The entries of z are $-1,1$ and 0, the entries of A are integers and the entries of x include the logarithms of the magnitudes of $\lambda_k,\rho_k,\theta_{1,k}$ and $\theta_{2,k}$.

The preprocessed polynomials $\lambda_k\ddot{f}_k(\omega_1,\omega_2)$, $\ddot{g}_k(\omega_1,\omega_2)$ and $\rho_k\ddot{h}_k(\omega_1,\omega_2)$ are therefore given by

$$
\begin{aligned}
\lambda_k \ddot{f}_k(\omega_1,\omega_2) = \lambda_k & \sum_{i_1+i_2=0}^{m} \bar{a}_{i_1,i_2}\theta_{1,k}^{i_1}\theta_{2,k}^{i_2}\binom{m}{i_1,i_2} \times \\
& (1-\theta_{1,k}\omega_1-\theta_{2,k}\omega_2)^{m-i_1-i_2}\omega_1^{i_1}\omega_2^{i_2}, && (41)\\
\ddot{g}_k(\omega_1,\omega_2) = & \sum_{i_1+i_2=0}^{n} \bar{b}_{i_1,i_2}\theta_{1,k}^{i_1}\theta_{2,k}^{i_2}\binom{n}{i_1,i_2} \times \\
& (1-\theta_{1,k}\omega_1-\theta_{2,k}\omega_2)^{n-i_1-i_2}\omega_1^{i_1}\omega_2^{i_2}, && (42)\\
\rho_k \ddot{h}_k(\omega_1,\omega_2) = \rho_k & \sum_{i_1+i_2=0}^{p} \bar{c}_{i_1,i_2}\theta_{1,k}^{i_1}\theta_{2,k}^{i_2}\binom{p}{i_1,i_2} \times \\
& (1-\theta_{1,k}\omega_1-\theta_{2,k}\omega_2)^{p-i_1-i_2}\omega_1^{i_1}\omega_2^{i_2}. && (43)
\end{aligned}
$$

These polynomials are expressed in a basis that is similar to, but distinct from, the basis defined by the functions (2). This basis is, for a bivariate polynomial of total degree m,

$$
\binom{m}{i_1,i_2}(1-\alpha\omega_1-\beta\omega_2)^{m-i_1-i_2}\omega_1^{i_1}\omega_2^{i_2}, \qquad i_1+i_2=0,\ldots,m, \tag{44}
$$

where α and β are constants. The coefficients of the polynomials (41), (42) and (43) define the entries of the Sylvester matrices and their subresultant matrices for the factorisation of $\lambda_1\ddot{f}_1(\omega_1,\omega_2), \ddot{g}_1(\omega_1,\omega_2)$ and $\rho_1\ddot{h}_1(\omega_1,\omega_2)$. It is shown in Section 8 that the preprocessing operations yield matrices whose rank is clearly defined, and in particular, more clearly defined than in the subresultant matrices whose entries are the coefficients of polynomials that have not been preprocessed.

6 The coefficients of the coprime polynomials and an AGCD

This section considers the computation of the coefficients of an AGCD and the coprime polynomials of $\lambda_1\ddot{f}_1(\omega_1,\omega_2), \ddot{g}_1(\omega_1,\omega_2)$ and $\rho_1\ddot{h}_1(\omega_1,\omega_2)$ when their coefficients are perturbed by random noise. The method used for this computation requires that the degree t of the AGCD be known because the coefficients are computed from the subresultant matrices $\tilde{S}_t(\lambda_t\ddot{f}_t,\ddot{g}_t,\rho_t\ddot{h}_t)$, or $\hat{S}_t(\lambda_t\ddot{f}_t,\ddot{g}_t,\rho_t\ddot{h}_t)$, or the other 2×3 block matrices. It follows from (41), (42) and (43) that these subresultant matrices are for polynomials expressed in the basis (44). The value of t is computed using Theorem 3.1, and the details follow from this computation for two univariate Bernstein polynomials [1] and three univariate Bernstein polynomials [3].

The subresultant matrices $\tilde{S}_t(\lambda_t\ddot{f}_t,\ddot{g}_t,\rho_t\ddot{h}_t)$ and $\hat{S}_t(\lambda_t\ddot{f}_t,\ddot{g}_t,\rho_t\ddot{h}_t)$ have unit loss of rank in the absence of noise, and thus from (21) and (31) in the presence of noise

$$\tilde{S}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)\tilde{\mathbf{x}}_t \approx \mathbf{0} \qquad \text{and} \qquad \hat{S}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)\hat{\mathbf{x}}_t \approx \mathbf{0},$$

where $\tilde{\mathbf{x}}_t$ and $\hat{\mathbf{x}}_t$ contain the coefficients of the coprime polynomials. The coefficient matrices in these approximate homogeneous equations have approximate unit loss of rank, and there therefore exists a column $\tilde{\mathbf{c}}_{t,q}$, indexed by q, of $\tilde{S}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)$, and a column $\hat{\mathbf{c}}_{t,r}$, indexed by r, of $\hat{S}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)$, that almost lie in the spaces spanned by the remaining columns of these matrices. If $\tilde{A}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)$ is the matrix formed by the removal of the column $\tilde{\mathbf{c}}_{t,q}$ from $\tilde{S}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)$, and $\hat{A}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)$ and $\hat{\mathbf{c}}_{t,r}$ are defined similarly with respect to $\hat{S}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)$, then

$$\tilde{A}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)\tilde{\mathbf{x}}_{t,q} \approx \tilde{\mathbf{c}}_{t,q} \qquad \text{and} \qquad \hat{A}_t(\lambda_t \ddot{f}_t, \ddot{g}_t, \rho_t \ddot{h}_t)\hat{\mathbf{x}}_{t,r} \approx \hat{\mathbf{c}}_{t,r}, \tag{45}$$

where $\tilde{\mathbf{x}}_{t,q}$ and $\hat{\mathbf{x}}_{t,r}$ are computed by the method of least squares. Each vector yields an estimate of an AGCD of perturbed forms of the polynomials defined in (41), (42) and (43). The vector $\tilde{\mathbf{x}}_t$ is obtained by the insertion of -1 in the qth position of $\tilde{\mathbf{x}}_{t,q}$, and $\hat{\mathbf{x}}_t$ is obtained in an identical manner from $\hat{\mathbf{x}}_{t,r}$.

The least squares solutions of (45) yield two estimates of the coprime polynomials, $(\tilde{u}_t(\omega_1,\omega_2), \tilde{v}_t(\omega_1,\omega_2), \tilde{w}_t(\omega_1,\omega_2))$ and $(\hat{u}_t(\omega_1,\omega_2), \hat{v}_t(\omega_1,\omega_2), \hat{w}_t(\omega_1,\omega_2))$, and they allow two estimates of the AGCD, $\tilde{d}_t(\omega_1,\omega_2)$, and $\hat{d}_t(\omega_1,\omega_2)$, to be computed,

$$\begin{aligned}
&\lambda_t \ddot{f}_t(\omega_1,\omega_2) \approx \tilde{d}_t(\omega_1,\omega_2)\tilde{u}_t(\omega_1,\omega_2),\\
&\ddot{g}_t(\omega_1,\omega_2) \approx \tilde{d}_t(\omega_1,\omega_2)\tilde{v}_t(\omega_1,\omega_2),\\
&\rho_t \ddot{h}_t(\omega_1,\omega_2) \approx \tilde{d}_t(\omega_1,\omega_2)\tilde{w}_t(\omega_1,\omega_2),
\end{aligned}$$

and

$$\begin{aligned}
&\lambda_t \ddot{f}_t(\omega_1,\omega_2) \approx \hat{d}_t(\omega_1,\omega_2)\hat{u}_t(\omega_1,\omega_2),\\
&\ddot{g}_t(\omega_1,\omega_2) \approx \hat{d}_t(\omega_1,\omega_2)\hat{v}_t(\omega_1,\omega_2),\\
&\rho_t \ddot{h}_t(\omega_1,\omega_2) \approx \hat{d}_t(\omega_1,\omega_2)\hat{w}_t(\omega_1,\omega_2).
\end{aligned}$$

These approximate equations are written in matrix form and solved in the least squares sense,

$$\begin{bmatrix} C(\tilde{u}_t) \\ C(\tilde{v}_t) \\ C(\tilde{w}_t) \end{bmatrix} \tilde{\mathbf{d}}_t \approx \begin{bmatrix} \lambda_t \ddot{\mathbf{f}}_t \\ \ddot{\mathbf{g}}_t \\ \rho_t \ddot{\mathbf{h}}_t \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} C(\hat{u}_t) \\ C(\hat{v}_t) \\ C(\hat{w}_t) \end{bmatrix} \hat{\mathbf{d}}_t \approx \begin{bmatrix} \lambda_t \ddot{\mathbf{f}}_t \\ \ddot{\mathbf{g}}_t \\ \rho_t \ddot{\mathbf{h}}_t \end{bmatrix},$$

where $C(s)$ is the convolution matrix whose entries are the coefficients of the bivariate polynomial $s(\omega_1, \omega_2)$ that is expressed in the basis (44). This method is used for the computation of the coefficients of the coprime polynomials and AGCD from the other 2×3 block matrix forms of the Sylvester matrix and its subresultant matrices, but the order in which the coefficients of the coprime polynomials are stored in $\hat{\mathbf{x}}_t$ is different, as shown for Variants 2 and 3 in (32) and (33) respectively.

7 Complexity

The non-zero blocks in the Sylvester matrices and subresultant matrices include Tœplitz matrices, but this Tœplitz structure is not preserved in the blocks. It follows that efficient algorithms that exploit this Tœplitz structure cannot be employed and generic algorithms must therefore be used. This cause of algorithmic complexity does not occur when power basis polynomials are considered because the Sylvester matrix and subresultant matrices of three power basis polynomials are formed by blocks of Tœplitz matrices, which allows efficient algorithms to be used.

All the Sylvester matrices and subresultant matrices considered in this paper are large, even for polynomials of moderate degrees, because their sizes are defined by binomial terms. Consideration of this point and the absence of a simple structure of the matrices, as discussed above, implies that the complexity of the algorithms discussed in this paper is high.

8 Examples

This section contains three examples that illustrate the theory of the previous sections.

Example 8.1 Consider the Bernstein forms of the exact polynomials $\hat{f}(x,y)$ and $\hat{g}(x,y)$, whose factorisations are

$$\begin{aligned}\hat{f}(x,y) &= (x+0.56)(x^2+y^2+0.51)^2(x+y+1.12)^3(x+y+0.0124)^6,\\ \hat{g}(x,y) &= (x+0.56)(x^2+y^2+0.51)^2(x+y+1.12)^3(x+y+0.4512)^3,\end{aligned}$$

and whose GCD $\hat{d}_t(x,y)$ is of degree $t=8$,

$$\hat{d}_t(x,y) = (x+0.56)(x^2+y^2+0.51)^2(x+y+1.12)^3.$$

Noise was added to the coefficients of $\hat{f}(x, y)$ and $\hat{g}(x, y)$ such that the coefficients of the inexact polynomials $f(x, y)$ and $g(x, y)$ were

$$a_{i_1,i_2} = \hat{a}_{i_1,i_2} + \hat{a}_{i_1,i_2}\epsilon_f r_{f,i_1,i_2} \qquad \text{and} \qquad b_{j_1,j_2} = \hat{b}_{j_1,j_2} + \hat{b}_{j_1,j_2}\epsilon_g r_{g,j_1,j_2},$$

where $\{r_{f,i_1,i_2}\}$ and $\{r_{g,j_1,j_2}\}$ are uniformly distributed random variables in the interval $[-1, 1]$ and $\epsilon_f = \epsilon_g = 10^{-10}$.

Figure 3 shows that the coefficients of $f(x, y)$ span approximately 9 orders of magnitude, but the coefficients of the preprocessed polynomial $\lambda_1\ddot{f}_1(\omega_1, \omega_2)$ span 5 orders of magnitude. The coefficients of $g(x, y)$ and $\ddot{g}_1(\omega_1, \omega_2)$ span approximately 4 orders of magnitude, and the coefficients of $\ddot{g}_1(\omega_1, \omega_2)$ are of the same order of magnitude as the coefficients of $\lambda_1\ddot{f}_1(\omega_1, \omega_2)$.

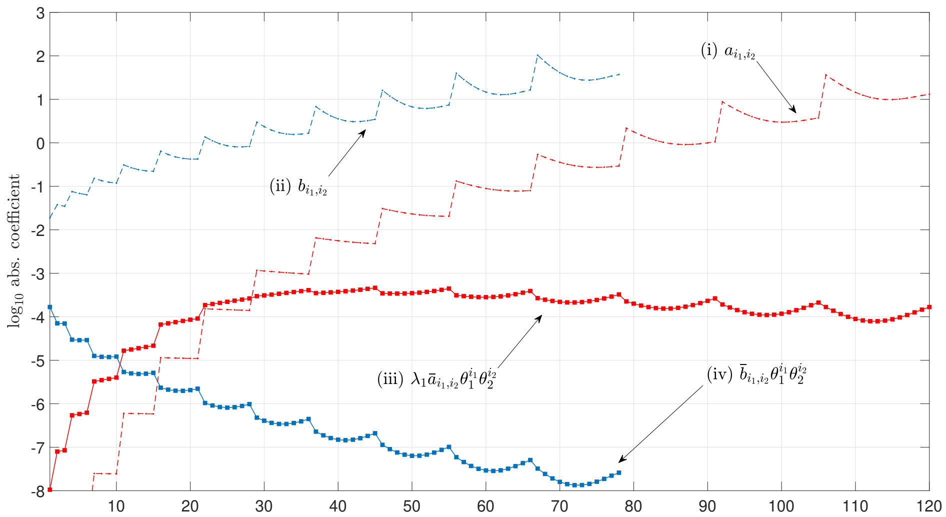
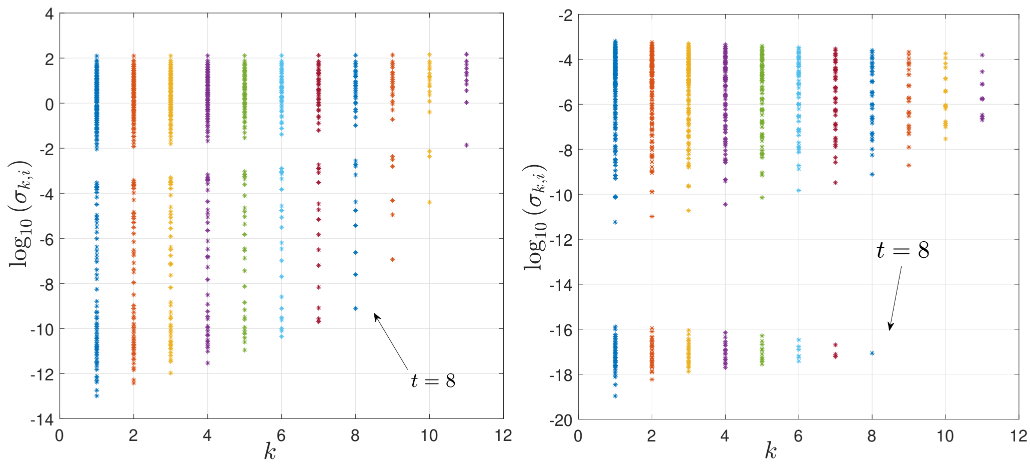


Fig. 3. The coefficients of (i) the unprocessed polynomial $f(x, y)$, (ii) the unprocessed polynomial $g(x, y)$, (iii) the preprocessed polynomial $\lambda_1\ddot{f}_1(\omega_1, \omega_2)$ and (iv) the preprocessed polynomial $\ddot{g}_1(\omega_1, \omega_2)$, in Example 8.1.

The singular values of the subresultant matrices of the unprocessed and preprocessed polynomials are plotted in Figure 4. There does not exist a clear separation between the numerically zero and non-zero singular values of the subresultant matrices of the unprocessed polynomials, but this separation of the zero and non-zero singular values is clearly defined for the subresultant matrices of the preprocessed polynomials.

The calculation of the coprime polynomials and AGCD requires that the variables (ω_1, ω_2) be transformed to the variables (x, y), using the optimal values $\theta = \theta_{1,t}$ and $\theta = \theta_{2,t}$ in (40). The errors in the coefficients of the coprime polynomials and AGCD are shown in Table 1. These coefficients were not computed from the unprocessed polynomials because the degree of an AGCD



(i) (ii)

Fig. 4. The singular values $\{\sigma_{k,i}\}$ of (i) $S_k(f, g)$ and (ii) $S_k(\lambda_k \ddot{f}_k, \ddot{g}_k)$ in Example 8.1.

was not defined.

Error $\hat{u}_t(x, y)$	9.00e-10
Error $\hat{v}_t(x, y)$	7.41e-10
Error $\hat{d}_t(x, y)$	3.09e-10

Table 1
Errors in the coprime polynomials $\hat{u}_t(x, y)$ and $\hat{v}_t(x, y)$, and AGCD $\hat{d}_t(x, y)$, in Example 8.1.

A reduction in the level of noise such that $\epsilon_f = \epsilon_g = 10^{-14}$ yielded a significant improvement in the solution because the errors in $\hat{u}_t(x, y)$, $\hat{v}_t(x, y)$ and $\hat{d}_t(x, y)$ without preprocessing and with preprocessing were, respectively, 10^{-9} and 10^{-14}. The errors in the unprocessed polynomials could be computed because the reduction in the noise level was such that the degree of an AGCD was clearly defined. □

Example 8.2 The 3×3 block matrix $\tilde{S}_k(\lambda_k \ddot{f}_k, \ddot{g}_k, \rho_k \ddot{h}_k)$ and three 2×3 block matrices $\hat{S}_k(\lambda_k \ddot{f}_k, \ddot{g}_k, \rho_k \ddot{h}_k)$, $\hat{S}_k(\ddot{g}_k, \lambda_k \ddot{f}_k, \rho_k \ddot{h}_k)$ and $\hat{S}_k(\rho_k \ddot{h}_k, \ddot{g}_k, \lambda_k \ddot{f}_k)$ of the polynomials in Example 3.1, after preprocessing, were formed and their singular value decomposition computed. The singular values are shown in Figure 5 and it is seen that the correct result, $t = 4$, is achieved by the four sets of matrices. These results are significantly better than the results of Example 3.1, and furthermore, the quality of the result is independent of the form of matrix structure (the 3×3 block form and the 2×3 block forms) of the Sylvester matrix and its subresultant matrices that is used for the computation. □

Example 8.3 Consider the Bernstein forms of the exact polynomials $\hat{f}(x, y)$, $\hat{g}(x, y)$ and $\hat{h}(x, y)$, which are of degrees $m = 17, n = 11$ and $p = 10$ respec-

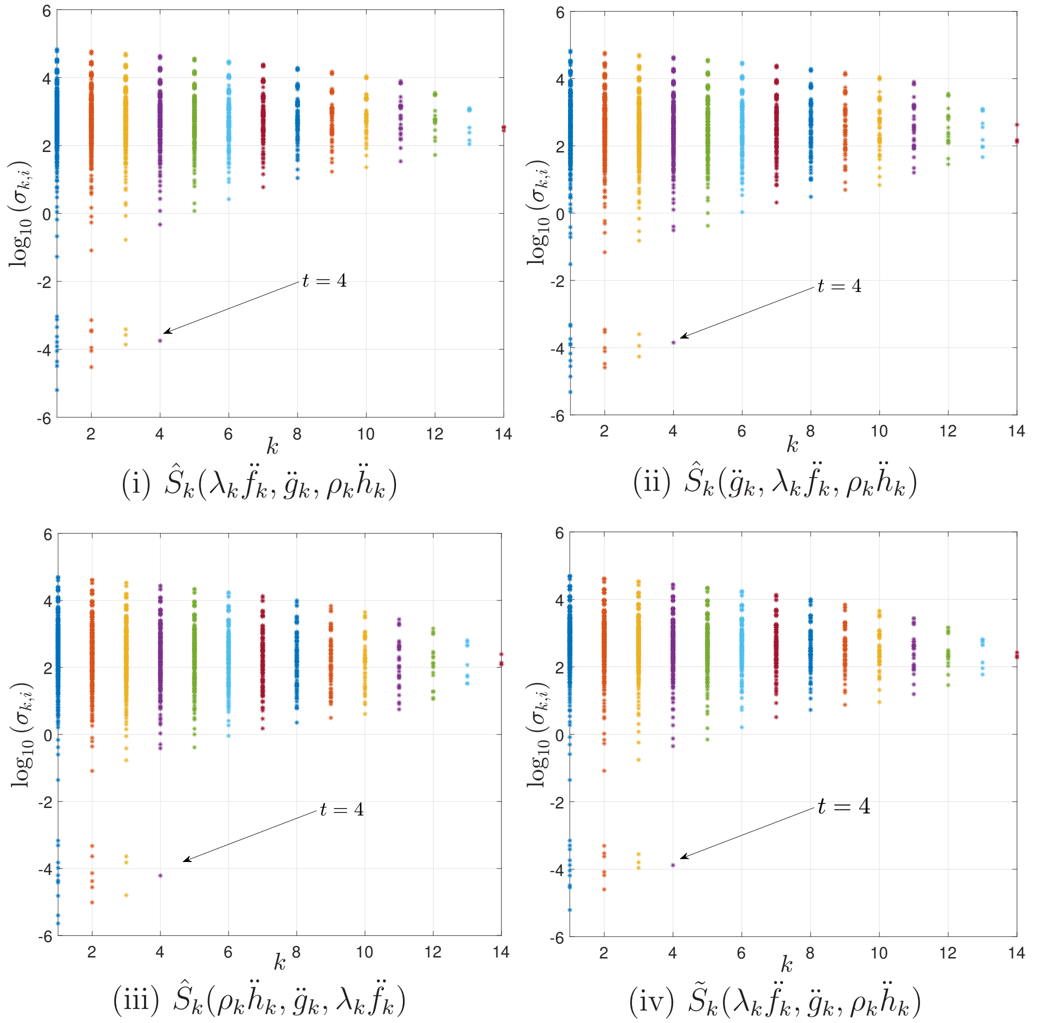


Fig. 5. The singular values $\{\sigma_{k,i}\}$ of (i) $\hat{S}_k(\lambda_k \ddot{f}_k, \ddot{g}_k, \rho_k \ddot{h}_k)$, (ii) $\hat{S}_k(\ddot{g}_k, \lambda_k \ddot{f}_k, \rho_k \ddot{h}_k)$, (iii) $\hat{S}_k(\rho_k \ddot{h}_k, \ddot{g}_k, \lambda_k \ddot{f}_k)$ and (iv) $\tilde{S}_k(\lambda_k \ddot{f}_k, \ddot{g}_k, \rho_k \ddot{h}_k)$ in Example 8.2.

tively, and whose factorisations are

$$
\begin{aligned}
\hat{f}(x,y) &= (x + 2.21657951321)(x^2 + y^2 + 0.5679814324687)^2 \times \\
&\quad (x + y + 42.46578784351654)^3 (x + y + 0.0124)^6 \times \\
&\quad (x - 0.554687987932164654)^3, \\
\hat{g}(x,y) &= (x + 2.21657951321)(x^2 + y^2 + 0.5679814324687)^2 \times \\
&\quad (x + y + 42.46578784351654)^3 (x + y + 0.4512)^3, \\
\hat{h}(x,y) &= (x + 2.21657951321)(x^2 + y^2 + 0.5679814324687)^2 \times \\
&\quad (x + y + 42.46578784351654)^3 (12x^2 + y^2 - 52.34).
\end{aligned}
$$

Their GCD $\hat{d}_t(x,y)$ is of degree $t = 8$,

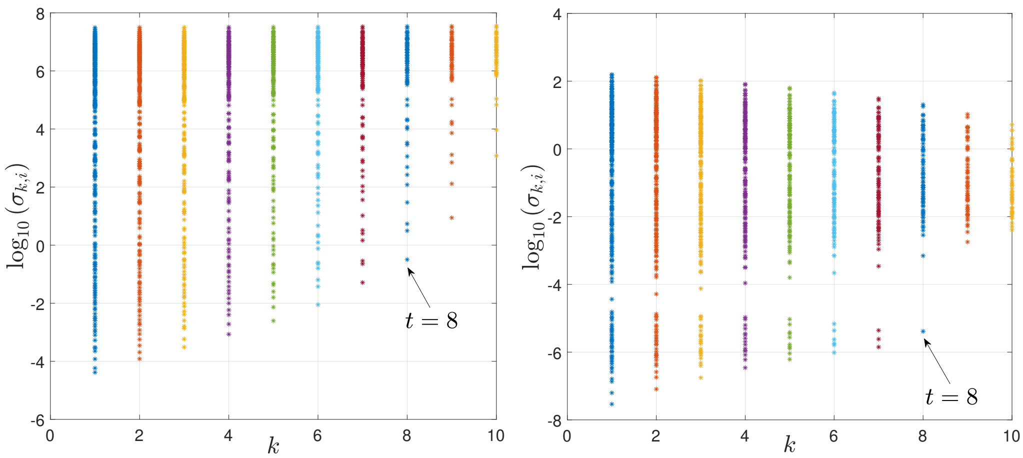
Error $\hat{u}_t(x,y)$	5.89e-04
Error $\hat{v}_t(x,y)$	6.05e-04
Error $\hat{w}_t(x,y)$	4.34e-04
Error $\hat{d}_t(x,y)$	5.72e-04

Table 2
Errors in the coprime polynomials $\hat{u}_t(x,y), \hat{v}_t(x,y)$ and $\hat{w}_t(x,y)$, and AGCD $\hat{d}_t(x,y)$, where $\{\epsilon_{f,i_1,i_2}\}, \{\epsilon_{g,j_1,j_2}\}$ and $\{\epsilon_{h,l_1,l_2}\}$ are uniformly distributed random variables in the interval $[10^{-6}, 10^{-4}]$, in Example 8.3.

$$\begin{aligned}\hat{d}_t(x,y) = {} & (x + 2.21657951321)(x^2 + y^2 + 0.5679814324687)^2 \times \\ & (x + y + 42.46578784351654)^3.\end{aligned}$$

Noise was added to the coefficients of $\hat{f}(x,y), \hat{g}(x,y)$ and $\hat{h}(x,y)$ such that the coefficients of the noisy polynomials $f(x,y), g(x,y)$ and $h(x,y)$ are given by (37) where $\{r_{f,i_1,i_2}\}$, $\{r_{g,j_1,j_2}\}$ and $\{r_{h,l_1,l_2}\}$ are uniformly distributed random variables in the interval $[-1,1]$, and $\{\epsilon_{f,i_1,i_2}\}$, $\{\epsilon_{g,j_1,j_2}\}$ and $\{\epsilon_{h,l_1,l_2}\}$ are uniformly distributed random variables in the interval $[10^{-6}, 10^{-4}]$.

The singular values of the matrices $\hat{S}_k(f,g,h)$, $k = 1, \ldots, \min(m,n,p)$, are plotted in Figure 6(i), and it is seen that the degree of an AGCD cannot be computed from these values because there does not exist a clear separation between the numerically zero and non-zero singular values. Figure 6(ii) shows, however, that the degree of an AGCD can be computed from the singular values of the subresultant matrices formed from the preprocessed polynomials $\hat{S}_k(\lambda_k \ddot{f}_k, \ddot{g}_k, \rho_k \ddot{h}_k)$ because the largest value of k for which a subresultant matrix is numerically rank deficient occurs for $k = 8$.



(i) $\hat{S}_k(f,g,h)$ (ii) $\hat{S}_k(\lambda_k \ddot{f}_k, \ddot{g}_k, \rho_k \ddot{h}_k)$

Fig. 6. The singular values $\{\sigma_{k,i}\}$ of (i) $\hat{S}_k(f,g,h)$, and (ii) $\hat{S}_k(\lambda_k \ddot{f}_k, \ddot{g}_k, \rho_k \ddot{h}_k)$ in Example 8.3.

Figure 6(i) shows that the degree t of an AGCD cannot be computed from the unprocessed polynomials $f(x,y), g(x,y)$ and $h(x,y)$, and thus the coef-

Polynomial	Without preprocessing	With preprocessing
$\hat{u}_t(x,y)$	1.68e-04	5.38e-06
$\hat{v}_t(x,y)$	1.48e-04	5.58e-06
$\hat{w}_t(x,y)$	1.27e-04	3.99e-06
$\hat{d}_t(x,y)$	1.58e-04	5.21e-06

Table 3
Errors in the coprime polynomials $\hat{u}_t(x,y), \hat{v}_t(x,y)$ and $\hat{w}_t(x,y)$, and AGCD $\hat{d}_t(x,y)$, where $\{\epsilon_{f,i_1,i_2}\}, \{\epsilon_{g,j_1,j_2}\}$ and $\{\epsilon_{h,l_1,l_2}\}$ are uniformly distributed random variables in the interval $[10^{-8}, 10^{-6}]$, in Example 8.3.

ficients of the coprime polynomials and AGCD cannot be computed. Figure 6(ii) shows, however, that the value of t is clearly defined when the polynomials are preprocessed, and thus the coprime polynomials $\hat{u}_t(x,y), \hat{v}_t(x,y)$ and $\hat{w}_t(x,y)$, and AGCD $\hat{d}_t(x,y)$, can be computed. The clear distinction between the rank deficient subresultant matrices and full rank subresultant matrices in Figure 6(ii) was also obtained when the other 2×3 block matrix forms, and the 3×3 block matrix form, were used. The errors in the coprime polynomials and AGCD are shown in Table 2 and it is seen that they lie in the interval $[10^{-6}, 10^{-4}]$ of the random variables $\{\epsilon_{f,i_1,i_2}\}$, $\{\epsilon_{g,j_1,j_2}\}$ and $\{\epsilon_{h,l_1,l_2}\}$, from which it follows that the computed solution is acceptable.

The noise level was reduced such that ϵ_{f,i_1,i_2}, ϵ_{g,j_1,j_2} and ϵ_{h,l_1,l_2} were uniformly distributed random variables in the interval $[10^{-8}, 10^{-6}]$. An AGCD of degree $t = 8$ could be computed for this noise level for the unprocessed and preprocessed polynomials, and the errors in the computed coefficients of the coprime polynomials and AGCD are shown in Table 3. Comparison with the errors in Table 2 shows that a reduction in the noise level by two orders of magnitude caused a similar decrease in the errors. Furthermore, the computed solution is acceptable because the errors lie in the interval $[10^{-8}, 10^{-6}]$ of the random variables $\{\epsilon_{f,i_1,i_2}\}$, $\{\epsilon_{g,j_1,j_2}\}$ and $\{\epsilon_{h,l_1,l_2}\}$.

The method of structured non-linear total least norm (SNTLN) [10] is used in [2] for the computation of a structured low rank approximation of the Sylvester matrix of two univariate Bernstein polynomials. This low rank approximation allows accurate results of the coprime polynomials and AGCD to be obtained, and they are more accurate than the results in Tables 2 and 3, which are obtained by the method of least squares. The results obtained by the method of least squares show, however, that processing the polynomials before computations are performed on them yields significantly improved results. The application of the method of SNTLN to the Sylvester matrix and its subresultant matrices of three bivariate Bernstein polynomials requires, however, significantly more work than is required for the Sylvester matrix and subresultant matrices of univariate Bernstein polynomials. $\square$

9 Summary

This paper has considered the computation of the degree and coefficients of an AGCD of three bivariate Bernstein polynomials defined in a triangular domain. The Sylvester matrix and its subresultant matrices were defined for these polynomials and it was shown that these matrices are not unique. In particular, there are three forms of these matrices when the polynomials are considered pairwise, and each of these matrices yields a block matrix of order 2×3. The three polynomials can also be considered simultaneously, which yields a block matrix of order 3×3.

It was shown that the computation of the degree of an AGCD of poorly scaled polynomials may return unsatisfactory results. The problems that arise from this poor scaling were addressed by processing the polynomials by three operations before computations are performed on them. Examples showed the improved results that are obtained when these preprocessing operations are implemented, but the operations increase the computational complexity of the method because it is necessary to solve a linear programming problem for each subresultant matrix.

References

[1] M. Bourne, J. R. Winkler, and Y. Su, *The computation of the degree of an approximate greatest common divisor of two Bernstein polynomials*, Applied Numerical Mathematics, 111 (2017), pp. 17–35.

[2] M. Bourne, J. R. Winkler, and Y. Su, *A non-linear structure-preserving matrix method for the computation of the coefficients of an approximate greatest common divisor of two Bernstein polynomials*, Journal of Computational and Applied Mathematics, 320 (2017), pp. 221–241.

[3] M. Bourne, J. R. Winkler, and S. Yi, *The computation of the degree of the greatest common divisor three Bernstein basis polynomials*, Journal of Computational and Applied Mathematics, 373 (2020), 112373.

[4] E. Chionh, R. Goldman, and J. Miller, *Using multivariate resultants to find the intersection of three quadric surfaces*, ACM Trans. Graphics, 10 (1991), pp. 378–400.

[5] R. M. Corless, P. M. Gianni, B. M. Trager, and S. M. Watt, *The singular value decomposition for polynomial systems*, in Proc. Int. Symp. Symbolic and Algebraic Computation, ACM Press, New York, 1995, pp. 195–207.

[6] R. M. Corless, M. W. Giesbrecht, M. van Hoeij, I. S. Kotsireas, and S. M. Watt, *Towards factoring bivariate approximate polynomials*, in Proc.

Int. Symp. Symbolic and Algebraic Computation, ACM Press, New York, 2001, pp. 85–92.

[7] S. Gao, *Factoring multivariate polynomials via partial differential equations*, Mathematics of Computation, 72 (2003), pp. 801–822.

[8] E. Kaltofen, J. May, Z. Yang, and L. Zhi, *Approximate factorization of multivariate polynomials using singular value decomposition*, Journal of Symbolic Computation, 43 (2008), pp. 359–376.

[9] M. A. Ochi, M. T. Noda, and T. Sasaki, *Approximate greatest common divisor of multivariate polynomials and its application to ill-conditioned systems of algebraic equations*, J. Information Processing, 14 (1991), pp. 292–300.

[10] J. B. Rosen, H. Park, and J. Glick, *Structured total least norm for nonlinear problems*, SIAM J. Mat. Anal. Appl., 20 (1998), pp. 14–30.

[11] C. E. Wee and R. Goldman, *Elimination and resultants: Part 2 Multivariate resultants*, IEEE Computer Graphics and Applications, (1995), pp. 60–69.

[12] J. R. Winkler and N. Yang, *Resultant matrices and the computation of the degree of an approximate greatest common divisor of two inexact Bernstein basis polynomials*, Computer Aided Geometric Design, 30 (4) (2013), pp. 410–429.

[13] L. Zhi and Z. Yang, *Computing approximate GCD of multivariate polynomials by structured total least norm*, tech. report, Institute of Systems Science, AMSS, Academia Sinica, Beijing, China, 2004.