Towards More Accurate Separation Bounds of Empirical Polynomials II

Nagasaka, Kosaku

(Citation)
Lecture Notes in Computer Science,3718:318-329

(Issue Date)
2005

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90001899

Towards More Accurate Separation Bounds of Empirical Polynomials II

Kosaku Nagasaka

Faculty of Human Development, Kobe University, Japan
nagasaka@main.h.kobe-u.ac.jp

Abstract. We study the problem of bounding a polynomial which is absolutely irreducible, away from polynomials which are not absolutely irreducible. These separation bounds are useful for testing whether an empirical polynomial is absolutely irreducible or not, for the given tolerance or error bound of its coefficients. In the former paper, we studied some improvements on Kaltofen and May's method which finds applicable separation bounds using an absolute irreducibility criterion due to Ruppert. In this paper, we study the similar improvements on the method using the criterion due to Gao and Rodrigues for sparse polynomials satisfying Newton polytope conditions, by which we are able to find more accurate separation bounds, for such bivariate polynomials. We also discuss a concept of separation bound continuations for both dense and sparse polynomials.

1 Introduction

We consider numerical polynomials with certain tolerances, including empirical polynomials with error bounds on its coefficients, which are useful for applied computations of polynomials. We have to use completely different algorithms from the conventional algorithms since we have to take care of their errors on coefficients and have to guarantee the results within the given tolerances.

In this paper and the former paper [1], we focus on testing absolute irreducibilities of such polynomials, hence we consider the following problem.

Problem 1. For the given polynomial $f \in \mathbb{C}[x, y]$ which is absolutely irreducible, compute the largest value $B(f) \in \mathbb{R}_{>0}$ such that all $\tilde{f} \in \mathbb{C}[x, y]$ with $\|f - \tilde{f}\|_2 < B(f)$ (and $\deg(\tilde{f}) \leq \deg(f)$) must remain absolutely irreducible. $\triangleleft$

This problem is studied by Kaltofen [2], however its separation bound is too small. The first applicable bound is given by the author [3], using an absolute irreducibility criterion due to Sasaki [4], and slightly improved by the author [5]. In ISSAC'03, Kaltofen and May [6] studied an efficient method using an absolute irreducibility criterion due to Ruppert [7], and a similar criterion due to Gao and Rodrigues [8] for sparse polynomials. The former paper [1] gave some improvements on Kaltofen and May's method due to Ruppert. Similar improvements on

[1] This research is partly helped by Grants-in-Aid of MEXT, JAPAN, #16700016.

their method due to Gao and Rodrigues can be available partly. This is one of main topics in this paper. Hence, the problem becomes the following.

Problem 2. For the given polynomial $f \in \mathbb{C}[x,y]$ which is absolutely irreducible, compute the largest value $\bar{B}(f) \in \mathbb{R}_{>0}$ such that all $\tilde{f} \in \mathbb{C}[x,y]$ satisfying $\mathcal{P}(\tilde{f}) \subseteq \mathcal{P}(f)$ with $\|f - \tilde{f}\|_2 < \bar{B}(f)$ must remain absolutely irreducible, where $\mathcal{P}(p)$ means the Newton polytope of a polynomial p. ◁

This is better for the case where we limit the changeable terms to being in the polytope. We note that the Newton polytope of a polynomial $p = \sum_{i,j} a_{i,j} x^i y^j$ is defined as the convex hull in the Euclidean plane $\mathbb{R}^2$ of the exponent vectors (i,j) of all the nonzero terms of p.

Example 1. Let $f(x,y)$ be the following irreducible polynomial in x and y.

$$f(x,y) = (x^2 + yx + 2y - 1)(x^3 + y^2x - y + 7) + 0.2x.$$

We have $B(f)/\|f\|_2 = 3.867 \times 10^{-5}$, by Kaltofen and May's algorithm. Hence, any polynomial which is included in ε-neighborhood of $f(x,y)$ in 2-norm, is still absolutely irreducible, where $\varepsilon = 3.867 \times 10^{-5}$. This bound can be optimized to 4.247×10^{-5} by the improved method [1]. We note that this polynomial can be factored approximately with the backward errors 7.531×10^{-4} [3] and 1.025×10^{-3} [9]. For the problem 2, we have $\bar{B}(f)/\|f\|_2 = 1.349 \times 10^{-4}$. We note that we have $\bar{B}(f) \leq B(f)$ for any polynomial f, since the all changeable terms in the sense of Problem 2 are included in those terms of Problem 1. ◁

The contribution of this paper is the following two points; 1) refining the Kaltofen and May's algorithm due to Gao and Rodrigues and finding more accurate separation bounds, 2) a discussion about a concept of separation bound continuations for both dense and sparse polynomials.

2 Original Method

Kaltofen and May's method mainly uses the following absolute irreducibility criterion due to Ruppert [7]. For the given polynomial, consider the following differential equation w.r.t. unknown polynomials g and h.

$$f\frac{\partial g}{\partial y} - g\frac{\partial f}{\partial y} + h\frac{\partial f}{\partial x} - f\frac{\partial h}{\partial x} = 0, \quad g,h \in \mathbb{C}[x,y], \tag{1}$$

$$\deg_x g \leq \deg_x f - 1,\ \deg_y g \leq \deg_y f,\ \deg_x h \leq \deg_x f,\ \deg_y h \leq \deg_y f - 2.$$

The criterion is that $f(x,y)$ is absolutely irreducible if and only if this differential equation (1) does not have any non-trivial solutions.

Their method uses matrix representations of absolute irreducibility criteria, and check whether those matrices are of certain ranks or not. They use the following matrix, for the above criterion, considering the above differential equation w.r.t. g and h as a linear system w.r.t. unknown coefficients of g and h.

Fig. 1. Ruppert matrix $R(f)$

$$
\begin{pmatrix}
G_n & 0 & \cdots & 0 & 0\cdot H_n & 0 & \cdots & 0 & 0 \\
G_{n-1} & G_n & \ddots & \vdots & -H_{n-1} & H_n & \ddots & \vdots & \vdots \\
\vdots & G_{n-1} & \ddots & 0 & \vdots & 0\cdot H_{n-1} & \ddots & 0 & \vdots \\
G_1 & \vdots & \ddots & G_n & (1-n)H_1 & \vdots & \ddots & (n-1)H_n & 0 \\
G_0 & G_1 & \ddots & G_{n-1} & -nH_0 & (2-n)H_1 & \ddots & (n-2)H_{n-1} & nH_n \\
0 & G_0 & \ddots & \vdots & 0 & (1-n)H_0 & \ddots & \vdots & (n-1)H_{n-1} \\
\vdots & \ddots & \ddots & G_1 & \vdots & 0 & \ddots & 0\cdot H_1 & \vdots \\
0 & \cdots & 0 & G_0 & 0 & \cdots & 0 & -H_0 & H_1
\end{pmatrix}
$$

$$
G_i = \begin{pmatrix}
0 & 0 & \cdots & 0 & 0 & 0 \\
c_{i,m-1} & -c_{i,m} & \ddots & \vdots & \vdots & 0 \\
2c_{i,m-2} & 0 & \ddots & 0 & \vdots & \vdots \\
\vdots & c_{i,m-2} & \ddots & (2-m)\,c_{i,m} & 0 & \vdots \\
\vdots & \vdots & \ddots & \vdots & (1-m)\,c_{i,m} & 0 \\
m\,c_{i,0} & \vdots & \ddots & \vdots & \vdots & -m\,c_{i,m} \\
0 & (m-1)\,c_{i,0} & \ddots & 0 & \vdots & \vdots \\
\vdots & 0 & \ddots & c_{i,1} & -c_{i,2} & \vdots \\
0 & \vdots & \ddots & 2c_{i,0} & 0 & -2c_{i,2} \\
0 & 0 & \cdots & 0 & c_{i,0} & -c_{i,1}
\end{pmatrix},
\quad
H_i = \begin{pmatrix}
0 & \cdots & 0 \\
c_{i,m} & \ddots & \vdots \\
c_{i,m-1} & \ddots & 0 \\
\vdots & \ddots & c_{i,m} \\
c_{i,1} & \ddots & c_{i,m-1} \\
c_{i,0} & \ddots & \vdots \\
 & \ddots & c_{i,1} \\
0 & & c_{i,0}
\end{pmatrix}
$$

Let $R(f)$ be the coefficient matrix of the linear system as in the figure 1, where the block matrices G_i and H_i are the matrices of sizes $2m \times (m+1)$ and $2m \times (m-1)$, respectively, where the given polynomial be

$$f = \sum_{i=0}^{n} \sum_{j=0}^{m} c_{i,j} x^i y^j, \quad c_{i,j} \in \mathbb{C}.$$

We call $R(f)$ the Ruppert matrix. The size of Ruppert matrix $R(f)$ is $(4nm) \times (2nm + m - 1)$ where $n = \deg_x(f)$ and $m = \deg_y(f)$.

The original expressions of separation bounds in Kaltofen and May's algorithm [6] are the following $B_\alpha(f)$ and $B_\beta(f)$. We note that $B_\alpha(f)$ is a lower bound of $B_\beta(f)$ by bounding the largest coefficient of $\|R(\varphi)\|_F^2$ in $B_\beta(f)$, where $\|A\|_F$ denotes the Frobenius norm (the square root of the sum of squares of all the elements) of matrix A.

$$B_\alpha(f) = \frac{\sigma(R(f))}{\max\{n,m\}\sqrt{2nm-n}}, \; B_\beta(f) = \frac{\sigma(R(f))}{\sqrt{(\text{the largest coef. of } \|R(\varphi)\|_F^2)}},$$

where $R(\varphi)$ denotes $R(f)$ calculated by treating $c_{i,j}$ as variables and $\sigma(A)$ denotes the $(2nm + m - 1)$-th largest singular value of matrix A.

3 Previous Work

In the former article [1], we decomposed $R(f)$ to integer matrices and complex coefficients parts, and gave some improvements using those matrices.

We refer the former results, briefly. The Ruppert matrix can be written as

$$R(f) = \sum_{i=0}^{n}\sum_{j=0}^{m} R_{i,j}c_{i,j}, \quad R_{i,j} \in \mathbb{Z}^{(4nm)\times(2nm+m-1)}, \tag{2}$$

where each elements of $R_{i,j}$ is an integer coefficient generated by differentiating polynomials, and $R_{i,j}$ has the same shape as $R(f)$ but whose elements are different. Then, the expressions of separation bounds can be refined as the following expression, by Lemma 1 in the former paper.

$$B(f) = \sqrt{6}\,\sigma(R(f))/\sqrt{n(m(m+1)(2m+1)+(m-1)(n+1)(2n+1))}. \tag{3}$$

3.1 Improvement Strategy

We refer the strategy of the former paper [1], improving the original method of Kaltofen and May due to the Ruppert.

The method uses the absolute irreducibility criteria as a necessary condition which the given polynomial is absolutely irreducible. In the Kaltofen and May's algorithm, $\sigma(R(f))$ is considered as a threshold whether the differential equation (or the linear system) (1) has non-trivial solutions or not. In this point of view, to determine that the differential equation does not have non-trivial solutions, corresponding to that the given polynomial is absolutely irreducible, we do not need to use all the constraint equations w.r.t. unknown coefficients of polynomials g and h, since the corresponding linear system is over-determined. We can lessen the number of constraint equations appeared in the Ruppert matrix $R(f)$, without decreasing its matrix rank.

We note that removing rows (constraint equations) may decrease the numerator of the expression (3) and may decrease the denominator depending on the elements of $R_{i,j}$. Hence, depending on variations of the numerator and denominator, $R(f)$ changes and it can be larger if we choose suitable rows.

As in the former paper, we define the following "drop" notations for removing rows from a matrix, which are corresponding to removing constraint equations.

$$\mathrm{drop}_i(A) = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{i-1}, \boldsymbol{0}, \boldsymbol{a}_{i+1}, \ldots, \boldsymbol{a}_{k_1})^t, \ A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{k_1})^t \in \mathbb{C}^{k_1\times k_2},$$

$$R^{(k)}(f) = \mathrm{drop}_{d_k}(\cdots(\mathrm{drop}_{d_1}(R(f)))), \quad R^{(k)}_{i,j} = \mathrm{drop}_{d_k}(\cdots(\mathrm{drop}_{d_1}(R_{i,j}))),$$

where $d_1, \ldots, d_k$ are indices of rows removed from the given matrix.

Improving the original method now becomes the following problem.

Problem 3. Find an integer k, row indices d_1, ..., d_k to be removed, and the following separation bound $B^{(k)}(f) > B(f)$.

$$B^{(k)}(f) = \sigma(R^{(k)}(f))/\max_{i,j} \|R^{(k)}_{i,j}\|_F\,. \qquad \triangleleft$$

Lemma 1 (Lemma 2 in [1]). *We have to remove at least 2 rows ($k = 2$) from the Ruppert matrix for finding more accurate separation bounds satisfying $B^{(k)}(f) > B(f)$. For $k = 2$, rows to be removed from the matrix, must satisfy*

$$\begin{cases} d_1 = 2md_x + d_y \ (0 \le d_x \le n-1 \ \wedge \ d_y = m+1), \\ d_2 = 2md_x + d_y \ (n \le d_x \le 2n-1 \ \wedge \ d_y = m+1) \end{cases}$$
$$or \begin{cases} d_1 = 2md_x + d_y \ (d_x = n \ \wedge \ 2 \le d_y \le m), \\ d_2 = 2md_x + d_y \ (d_x = n \ \wedge \ m+2 \le d_y \le 2m). \end{cases} \quad \triangleleft$$

By Lemma 1, the simple algorithm was introduced, which give us about 1.6% more accurate separation bounds, according to the experimental result in the former paper. We note that "removing multiple rows" versions of the algorithm were also introduced in the paper.

4 Newton Polytope Version

Kaltofen and May also argued briefly the method using the following criterion due to Gao and Rodrigues [8] which is effective for factoring sparse polynomials. For the given polynomial, consider the following differential equation w.r.t. unknown polynomials g and h in $\mathbb{C}[x, y]$.

$$f\frac{\partial g}{\partial y} - g\frac{\partial f}{\partial y} + h\frac{\partial f}{\partial x} - f\frac{\partial h}{\partial x} = 0, \ \ \mathcal{P}(xg) \subseteq \mathcal{P}(f) \text{ and } \mathcal{P}(yh) \subseteq \mathcal{P}(f). \quad (4)$$

The criterion that the given polynomial is absolutely irreducible is a little bit different from the Ruppert criterion. Let $\mathcal{R}(f)$ be the coefficient matrix of the linear system of the above differential equation (4) w.r.t. unknown coefficients of polynomials g and h. We call $\mathcal{R}(f)$ the sparse Ruppert matrix. Polynomials g and h do not have the same forms as in the differential equation (1) by Ruppert, hence, for sparse polynomials, the size of sparse Ruppert matrix $\mathcal{R}(f)$ is less than the size of Ruppert matrix $R(f)$. The figure of the sparse Ruppert matrix is depending on the Newton polytope of the given polynomial and we can not show its general form. For easiness of discussions, we define the skeleton of the sparse Ruppert matrix $\bar{R}(f)$, with full terms of g and h, as in the figure 1, where the block matrices G_i and H_i are the matrices of sizes $2m \times (m+1)$ and $2m \times m$, respectively, as in the figure 2. The size of the skeleton matrix $R(f)$ is $(4nm) \times (2nm + n + m)$. We note that 1) the only difference between $R(f)$ and $\bar{R}(f)$ is on the block matrix H_i, 2) an actual sparse Ruppert matrix $\mathcal{R}(f)$ can be generated by replacing all elements with zeros, on some columns corresponding to unnecessary terms of polynomials g and h by the condition due to the Newton polytope of $f(x, y)$, or by removing such columns.

The criterion is that $f(x, y)$ is absolutely irreducible if and only if the sparse Ruppert matrix $\mathcal{R}(f)$ has the rank $\rho - 1$, where ρ denotes the number of unknown coefficients of polynomials g and h. We note that Problem 2 is corresponding to this criterion, and contributions of this paper are mainly for this problem. We

Fig. 2. Block matrices of skeleton matrix $\bar{R}(f)$

$$G_i = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ c_{i,m-1} & -c_{i,m} & \ddots & \vdots & \vdots & 0 \\ 2c_{i,m-2} & 0 & \ddots & 0 & \vdots & \vdots \\ \vdots & c_{i,m-2} & \ddots & (2-m)\times c_{i,m} & 0 & \vdots \\ \vdots & \vdots & \ddots & \vdots & (1-m)\, c_{i,m} & 0 \\ m\, c_{i,0} & \vdots & \ddots & \vdots & \vdots & -m\, c_{i,m} \\ 0 & (m-1)\, c_{i,0} & \ddots & 0 & \vdots & \vdots \\ \vdots & 0 & \ddots & c_{i,1} & -c_{i,2} & \vdots \\ 0 & \vdots & \ddots & 2c_{i,0} & 0 & -2c_{i,2} \\ 0 & 0 & \cdots & 0 & c_{i,0} & -c_{i,1} \end{pmatrix}, \quad H_i = \begin{pmatrix} c_{i,m} & 0 & \cdots & 0 \\ c_{i,m-1} & c_{i,m} & \ddots & \vdots \\ c_{i,m-2} & c_{i,m-1} & \ddots & 0 \\ \vdots & \vdots & \ddots & c_{i,m} \\ c_{i,0} & c_{i,1} & \ddots & c_{i,m-1} \\ 0 & c_{i,0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{i,1} \\ 0 & \cdots & & c_{i,0} \end{pmatrix}$$

have the following separation bound $\bar{B}_\beta(f)$, by the same way of the paper [6].

$$\bar{B}_\beta(f) = \bar{\sigma}(\mathcal{R}(f))/\sqrt{(\text{the largest coefficient of } \|\mathcal{R}(\varphi)\|_F^2)},$$

where $\bar{\sigma}(A)$ denotes the $(\rho-1)$-th largest singular value of matrix A. In the rest of this paper, we discuss the similar refining of $\bar{B}_\beta(f)$ as in the former paper [1].

4.1 Integer Matrices

We decompose $\mathcal{R}(f)$ and $\bar{R}(f)$ to integer matrices and complex coefficients parts, as in the previous section. These matrices can be written as

$$\mathcal{R}(f) = \sum_{i=0}^{n}\sum_{j=0}^{m} \mathcal{R}_{i,j} c_{i,j}, \quad \bar{R}(f) = \sum_{i=0}^{n}\sum_{j=0}^{m} \bar{R}_{i,j} c_{i,j},$$

where each elements of $\mathcal{R}_{i,j}$ and $\bar{R}_{i,j}$ is an integer coefficient generated by differentiating polynomials, and $\mathcal{R}_{i,j}$ and $\bar{R}_{i,j}$ have the same shape of $\mathcal{R}(f)$ and $\bar{R}(f)$, respectively, but all the elements are defined as in the figure 3 where $\delta_{i,j}$ denotes Kronecker delta. These integer matrices have the following properties similar to those of the integer matrices of the Ruppert matrix.

Lemma 2. *We have*

$$\max_{i,j} \|\bar{R}_{i,j}\|_F^2 = nm((2n+1)(n+1)+(2m+1)(m+1))/6. \quad \triangleleft$$

Proof. The same way as in Lemma 1 in [1]. □

Corollary 1. *We have the following equality.*

$$\max_{i,j} \|\bar{R}_{i,j}\|_F = \|\bar{R}_{n,m}\|_F = \|\bar{R}_{n,0}\|_F = \|\bar{R}_{0,m}\|_F = \|\bar{R}_{0,0}\|_F \,. \quad \triangleleft$$

Fig. 3. Integer matrix $R_{i,j}$

$$
\begin{pmatrix}
\delta_{i,n}G_n & \cdots & 0 & 0\cdot\delta_{i,n}H_n & 0 & \cdots & 0 & 0 \\
\delta_{i,n-1}G_{n-1} & \ddots & \vdots & -\delta_{i,n-1}H_{n-1} & \delta_{i,n}H_n & \ddots & \vdots & \vdots \\
\vdots & \ddots & 0 & \vdots & 0\cdot\delta_{i,n-1}H_{n-1} & \ddots & 0 & \vdots \\
\delta_{i,1}G_1 & \ddots & \delta_{i,n}G_n & (1-n)\delta_{i,1}H_1 & \vdots & \ddots & (n-1)\delta_{i,n}H_n & 0 \\
\delta_{i,0}G_0 & \ddots & \delta_{i,n-1}G_{n-1} & -n\delta_{i,0}H_0 & (2-n)\delta_{i,1}H_1 & \ddots & (n-2)\delta_{i,n-1}\times H_{n-1} & nH_n \\
0 & \ddots & \vdots & 0 & (1-n)\delta_{i,0}H_0 & \ddots & \vdots & (n-1)\delta_{i,n-1}\times H_{n-1} \\
\vdots & \ddots & \delta_{i,1}G_1 & \vdots & 0 & \ddots & 0\cdot\delta_{i,1}H_1 & \vdots \\
0 & \cdots & \delta_{i,0}G_0 & 0 & \cdots & 0 & -\delta_{i,0}H_0 & \delta_{i,1}H_1
\end{pmatrix}
$$

$$
G_i = \begin{pmatrix}
0 & 0 & \cdots & 0 & 0 & 0 \\
\delta_{j,m-1} & -\delta_{j,m} & \ddots & \vdots & \vdots & 0 \\
2\delta_{j,m-2} & 0 & \ddots & 0 & \vdots & \vdots \\
\vdots & \delta_{j,m-2} & \ddots & (2-m)\times\delta_{j,m} & 0 & \vdots \\
\vdots & \vdots & \ddots & \vdots & (1-m)\,\delta_{j,m} & 0 \\
m\,\delta_{j,0} & \vdots & \ddots & \vdots & \vdots & -m\,\delta_{j,m} \\
0 & (m-1)\,\delta_{j,0} & \ddots & 0 & \vdots & \vdots \\
\vdots & 0 & \ddots & \delta_{j,1} & -\delta_{j,2} & \vdots \\
0 & \vdots & \ddots & 2\delta_{j,0} & 0 & -2\delta_{j,2} \\
0 & 0 & \cdots & 0 & \delta_{j,0} & -\delta_{j,1}
\end{pmatrix}
$$

$$
H_i = \begin{pmatrix}
\delta_{j,m} & 0 & \cdots & 0 \\
\delta_{j,m-1} & \delta_{j,m} & \ddots & \vdots \\
\delta_{j,m-2} & \delta_{j,m-1} & \ddots & 0 \\
\vdots & \vdots & \ddots & \delta_{j,m} \\
\delta_{j,0} & \delta_{j,1} & \ddots & \delta_{j,m-1} \\
0 & \delta_{j,0} & \ddots & \vdots \\
\vdots & \ddots & \ddots & \delta_{j,1} \\
0 & \cdots & & \delta_{j,0}
\end{pmatrix}
$$

Remark 1. Using the above lemma and corollary, we can rewrite the expression $\bar{B}_\beta(f)$ as in the former paper, however, it is useless since an actual sparse Ruppert matrix which does not have some columns corresponding to unnecessary terms of polynomials g and h by the Newton polytope of $f(x, y)$, and separation bounds based on $\bar{R}(f)$ may be larger than those of an actual sparse Ruppert matrix $\mathcal{R}(f)$. Hence, we have to use the expression $\bar{B}_\beta(f)$ still. This is different from the Ruppert matrix case. ◁

The following lemma helps us to calculate the largest coefficient of $\|\mathcal{R}(\varphi)\|_F^2$ which is appeared in the denominator of $\bar{B}_\beta(f)$.

Lemma 3. *We have the following equality.*

$$\max_{i,j} \|\mathcal{R}_{i,j}\|_F = \max\{\|\mathcal{R}_{n,m}\|_F, \|\mathcal{R}_{n,0}\|_F, \|\mathcal{R}_{0,m}\|_F, \|\mathcal{R}_{0,0}\|_F\}. \quad \triangleleft$$

Proof. Since we can construct $\mathcal{R}_{i,j}$ by removing some columns from $\bar{R}_{i,j}$ (or replacing them with zeros), we only have to prove that Corollary 1 is still valid after removing columns. We focus only on the index j and consider the left hand

side of $\bar{R}_{i,j}$ formed by block matrices G_i and the the right hand side of $\bar{R}_{i,j}$ formed by block matrices H_i separately.

For the right hand side, each sum of squares of elements corresponding to an index j on each column has the same Frobenius norm. Hence, removing columns on the right hand side does not affect the equality of Corollary 1. Therefore, we only have to show that: the largest coefficients of $c_{i,m}$ and $c_{i,0}$ of the Frobenius norm of G_i is the largest coefficient among $c_{i,j}$ after removing.

Let $\Delta_{k,0}$ and $\Delta_{k,m}$ be differences between the coefficients of $c_{i,0}$ and $c_{i,m}$ and $c_{i,m-\kappa}$ of $\|G_i\|_F^2$ on the $k+1$-th column, respectively. We have

$$\begin{aligned}\Delta_{k,0} &= (m-k)^2 - (m-k-m+\kappa)^2 = (2k-\kappa-m)(\kappa-m),\\ \Delta_{k,m} &= (-k)^2 - (m-k-m+\kappa)^2 \quad = (2k-\kappa)\kappa.\end{aligned}$$

Let $\mathcal{I}$ be the set of column indices of the rest columns after removing. We suppose that the lemma is not valid and $c_{i,m-\kappa}$ has the largest coefficient. We have

$$\sum_{k\in\mathcal{I}} \Delta_{k,0} = (\kappa-m)\sum_{k\in\mathcal{I}}(2k-\kappa-m) < 0, \quad \sum_{k\in\mathcal{I}} \Delta_{k,m} = \kappa\sum_{k\in\mathcal{I}}(2k-\kappa) < 0.$$

Since $\kappa - m$ is not positive and κ is not negative, we have

$$\sum_{k\in\mathcal{I}}(2k-\kappa-m) = \sum_{k\in\mathcal{I}}(2k-\kappa) - \#\mathcal{I}m > 0, \quad \sum_{k\in\mathcal{I}}(2k-\kappa) < 0,$$

where $\#\mathcal{I}$ denotes the number of elements of the set $\mathcal{I}$. This leads a contradiction. Therefore the lemma is valid. We note that we can prove for the index i by the similar way even if it not necessary for the proof. □

By Lemma 3, we have the following separation bound.

$$\bar{B}(f) = \bar{\sigma}(\mathcal{R}(f))/\max\{\|\mathcal{R}_{n,m}\|_F, \|\mathcal{R}_{n,0}\|_F, \|\mathcal{R}_{0,m}\|_F, \|\mathcal{R}_{0,0}\|_F\}.$$

4.2 Improvement Strategy

For the sparse Ruppert matrix, the improvement strategy of the former paper is still applicable. Hence, the aim of this subsection is the following problem.

Problem 4. Find an integer k, indices $d_1, \ldots, d_k$ to be removed, and the following separation bound $\bar{B}^{(k)}(f) > \bar{B}(f)$.

$$\bar{B}^{(k)}(f) = \bar{\sigma}(\mathcal{R}^{(k)}(f))/\max_{i,j} \|\mathcal{R}^{(k)}_{i,j}\|_F\ . \qquad \triangleleft$$

- Removing Two Rows - For the sparse Ruppert matrix, we still consider "removing two rows from the matrix" even though the important corollary in [1] is not valid and we have only Lemma 3. Because even for such cases, we may have to remove rows providing that $\|\mathcal{R}_{n,m}\|_F$, $\|\mathcal{R}_{n,0}\|_F$, $\|\mathcal{R}_{0,m}\|_F$ and $\|\mathcal{R}_{0,0}\|_F$ become smaller and $\bar{B}^{(k)}(f) > \bar{B}(f)$, depending on $\mathcal{R}(f)$. Therefore, we follow

the same discussion. We consider variations of $\|\bar{R}_{i,j}\|_F$, provided by removing a $(2md_x + d_y)$-th row from $\bar{R}(f)$, satisfying $0 \leq d_x \leq 2n-1$ and $1 \leq d_y \leq 2m$.

Let Δ_G be the square of Frobenius norm of variations of the left hand side part of $\bar{R}_{i,j}$, corresponding to G_i and Δ_H be that of the right hand side part of $\bar{R}_{i,j}$, corresponding to H_i. We have

$$\|\mathrm{drop}_{2md_x+d_y}(\bar{R}_{i,j})\|_F^2 = \|\bar{R}_{i,j}\|_F^2 - \Delta_G - \Delta_H.$$

By the same way in the former paper, we have the following relations that are slightly different from those of the Ruppert matrix.

$$\Delta_G = \begin{cases} 0 & (i < n - d_x) \ \vee \ (2n - d_x - 1 < i) \ \vee \\ & \quad (j < m - d_y + 1) \ \vee \ (2m + 1 - d_y < j) \\ (2m + 1 - d_y - 2j)^2 & \text{otherwise} \end{cases} \quad (5)$$

$$\Delta_H = \begin{cases} 0 & (i < n - d_x) \ \vee \ (2n - d_x < i) \ \vee \\ & \quad (j < m - d_y + 1) \ \vee \ (2m - d_y < j) \\ (2i + d_x - 2n)^2 & \text{otherwise} \end{cases} \quad (6)$$

Lemma 4. *We may have to remove at least 2 rows ($k = 2$) from the sparse Ruppert matrix for finding more accurate separation bound satisfying $\bar{B}^{(k)}(f) > \bar{B}(f)$. For $k = 2$, rows to be removed from the matrix, should satisfy*

$$\textit{or} \begin{cases} d_1 = 2md_x + d_y \ (0 \leq d_x \leq n-1 \ \wedge \ d_y = m+1), \\ d_2 = 2md_x + d_y \ (n \leq d_x \leq 2n-1 \ \wedge \ d_y = m+1) \\ d_1 = 2md_x + d_y \ (d_x = n \ \wedge \ 1 \leq d_y \leq m), \\ d_2 = 2md_x + d_y \ (d_x = n \ \wedge \ m+1 \leq d_y \leq 2m). \end{cases} \quad \triangleleft$$

Proof. The same way as in Lemma 2 in [1]. □

We note that removing only one row has possibility to satisfy $\bar{B}^{(1)}(f) > \bar{B}(f)$, since we have only Lemma 3 for the sparse Ruppert matrix. However, the above lemma guarantees that removing such two rows must decrease $\max_{i,j} \|\mathcal{R}_{i,j}\|_F$ $= \max\{\|\mathcal{R}_{n,m}\|_F, \|\mathcal{R}_{n,0}\|_F, \|\mathcal{R}_{0,m}\|_F, \|\mathcal{R}_{0,0}\|_F\}$.

By Lemma 4, we have the following simple algorithm which give us about 1.3% more accurate separation bounds, according to our experimental result.

Algorithm 1. (Removing Two Rows Sparse Version)
Input: a bivariate polynomial $f(x, y)$, *Output*: a separation bound $\bar{B}(f)$

Step 1 Construct sparse Ruppert matrix $\mathcal{R}(f)$.
Step 2 For all index pairs d_1 and d_2 in Lemma 4, compute separation bounds, and let the best separation bound be $\bar{B}^{(2)}(f)$.
Step 3 Output the separation bound $\bar{B}^{(2)}(f)$ and finish the algorithm. $\triangleleft$

- Removing Multiple Rows - For the Ruppert matrix, in the former paper, by the lemma which guarantees Lemma 3 after removing rows, the algorithms removing multiple rows were introduced. For the sparse Ruppert matrix, such a lemma does not exist since an actual sparse Ruppert matrix does not have a lots of columns and removing rows easily breaks Lemma 3. However, we can use the similar algorithms though they are not effective as before.

Algorithm 2. (Early Termination Algorithm Sparse Version)
Input: a bivariate polynomial $f(x, y)$, *Output*: a separation bound $\bar{B}(f)$

Step 1 Construct sparse Ruppert matrix $\mathcal{R}(f)$ and put $k = 1$.
Step 2 Compute contributing ratios of each rows of $\mathcal{R}(f)$.
Step 3 Construct all the index pairs d_{2k-1} and d_{2k} as in Lemma 4.
Step 4 For each index pairs constructed in Step 3, compute separation bounds with $d_1, d_2, \ldots, d_{2k}$, by ascending order of sums of contributing ratios, until an index pair for which a separation bound does not become better than that of a previous group twice, and let the best separation bound be $\bar{B}^{(2k)}(f)$.
Step 5 If $\bar{B}^{(2k-2)}(f) \leq \bar{B}^{(2k)}(f)$ then put $k = k + 1$ and goto Step 3.
Step 6 Output the separation bound $\bar{B}^{(2k-2)}(f)$ and finish the algorithm. ◁

We use Euclidean norms of corresponding row vectors of the Moore-Penrose type pseudo inverse of the transpose of $\mathcal{R}(f)$ as the contributing ratios (see [1]).

Example 2. For the polynomial in the example 1, the algorithms 1 and 2 output $\bar{B}(f) = 1.420 \times 10^{-4}$ and $\bar{B}(f) = 1.427 \times 10^{-4}$, respectively, which are slightly better than the results in the beginning example. ◁

5 Separation Bound Continuation

In this section, we consider another way to enlarge separation bounds. The key idea is that the separation bound defines a kind of ε-neighborhood of the given polynomial $f(x, y)$. From this point of view, we consider to continuate one neighborhood to others like analytic continuations.

For the given $f(x, y)$ and $0 < b \in \mathbb{R}$, let $\mathcal{A}_b(f)$ be the set of all $\tilde{f} \in \mathbb{C}[x, y]$ with $\|f - \tilde{f}\|_2 < b$ and $\deg(\tilde{f}) \leq \deg(f)$. Hence $\mathcal{A}_{B(f)}(f)$ denotes a ε-neighborhood of the given $f(x, y)$, in which all polynomials must remain absolutely irreducible.

Definition 1. *Let $B_0(f) = B(f)$ and $B_i(f) \in \mathbb{R}$ $(i = 1, \ldots)$ be the maximum value satisfying*

$$\mathcal{A}_{B_i(f)}(f) \subseteq \bigcup_{g \in \mathcal{A}_{B_{i-1}(f)}(f)} \mathcal{A}_{B(g)}(g).$$

We call $B_i(f)$ $(i > 0)$ and $B_\infty(f)$ a continuated separation bound and the maximum continuated separation bound, of $f(x, y)$, respectively. ◁

One may think that "Does the given polynomial have an approximate factorization with tolerance $B_\infty(f)$?". The author thinks that the answer is "No" since separation bounds by the known methods are far from backward tolerances with which the given polynomials have approximate factorizations. However, this continuation helps us to enlarge separation bounds as follows.

For the problem 1, let ε be a arbitrary positive real number and $\Delta, b \in \mathbb{R}$ be

$$\Delta = B(f)/\sqrt{(n+1)(m+1)} - \varepsilon, \quad b = \min_{0 \leq i \leq n, 0 \leq j \leq m, k = -1, 1} B(f + k\Delta x^i y^j).$$

For the problem 2, let ε be a arbitrary positive real number and $\bar{\Delta}, \bar{b} \in \mathbb{R}$ be

$$\bar{\Delta} = \bar{B}(f)/\sqrt{\#\mathcal{M}} - \varepsilon, \quad \bar{b} = \min_{x^i y^j \in \mathcal{M}, k = -1, 1} \bar{B}(f + k\bar{\Delta} x^i y^j),$$

where $\mathcal{M}$ denotes the set of all the monomials $x^i y^j$ satisfying $\mathcal{P}(x^i y^j) \subseteq \mathcal{P}(f)$.

Lemma 5. *$\sqrt{b^2+\Delta^2}$ and $\sqrt{\bar{b}^2+\bar{\Delta}^2}$ are also separation bounds $B(f)$ and $\bar{B}(f)$ of $f(x,y)$, respectively, and they may be better than the original bounds.* ◁

Proof. We give the following proof only for $\sqrt{b^2+\Delta^2}$ since that for $\sqrt{\bar{b}^2+\bar{\Delta}^2}$ is proved by the same way. Let a polynomial $\tilde{f} \in \mathcal{A}_{\sqrt{b^2+\Delta^2}}(f)$ be

$$\tilde{f} = \sum_{i,j} (c_{i,j} + \tilde{c}_{i,j}) x^i y^j.$$

By the definition of $B(f)$, we have that $\tilde{f}$ is absolutely irreducible if $|\tilde{c}_{i,j}| \leq \Delta$ for all i and j. Hence, we suppose that one of variations of coefficients of $\tilde{f}$ from f is larger than Δ and such the term be $x^{i'} y^{j'}$. We rewrite $\tilde{f}$ be

$$\tilde{f} = \sum_{i,j} (c_{i,j} + \acute{c}_{i,j}) x^i y^j + k\Delta x^{i'} y^{j'}, \quad (k = -1 \text{ or } 1).$$

We have $\| f - \tilde{f} \|_2^2 = \sum_{i,j} |\acute{c}_{i,j}|^2 + 2|\acute{c}_{i',j'}|\Delta + \Delta^2 < b^2 + \Delta^2$ which means $\sum_{i,j} |\acute{c}_{i,j}|^2 < b^2$. Therefore, we have $\tilde{f} \in \mathcal{A}_b(f + k\Delta x^{i'} y^{j'})$ meaning $\tilde{f}$ remains absolutely irreducible, and the lemma is valid. □

Using the lemma, we define partial continuated separation bounds of $f(x,y)$, $B_C(f) = \max\{B(f), \sqrt{b^2+\Delta^2}\}$ and $\bar{B}_C(f) = \max\{\bar{B}(f), \sqrt{\bar{b}^2+\bar{\Delta}^2}\}$.

Example 3. For the polynomial in the example 1, the algorithm using the above lemma (let it be Algorithm C) outputs $B_C(f) = 4.068 \times 10^{-5}$ and $\bar{B}_C(f) = 1.467 \times 10^{-4}$, which are slightly better though it is very time-consuming. ◁

6 Numerical Experiment and Remarks

We have generated 100 bivariate sparse polynomials of degrees 6 and 5 w.r.t. x and y, respectively, with coefficients randomly chosen in the real interval $[-1, 1]$, where each sample is irreducible and about 25% of coefficients are non-zero. With those polynomials, we have tested the new algorithm 1, 2 and C, using our preliminary implementations. We note that the results of our experiments are small so we have to take care of precisions. Basically, we have tested it using the same way in the paper [3] (bounding errors of singular values). The upper part of the table 1 shows the results. According to the results, our improvements give us more accurate separation bounds.

Moreover, we have generated 100 bivariate reducible polynomials. Each polynomial is a product of two dense polynomials of total-degrees 5 and 4, respectively, with coefficients randomly chosen in the integer interval $[-5, 5]$. Using those polynomials, we have generated 100 approximately reducible polynomials. Each polynomial is a sum of a reducible polynomial and a polynomial which has the same degree as the reducible polynomial, about 25% as many terms and coefficients randomly chosen in the real interval $[-10^{-4}, 10^{-4}]$.

With those polynomials, we have tested the new algorithms except for the algorithm C . The lower part of the table 1 shows the results. According to the results, our improvements give us more accurate separation bounds. Although we could not use the algorithm C for all the generated polynomials due to its time-complexity, it gave us better results. We note that an average of backward errors of those approximately reducible polynomials by the method [9] is 2.829×10^{-4}.

Table 1. Experimental results

		Algorithm	$B(f)/\|f\|$ or $\tilde{B}(f)/\|f\|$	Ratio to KM03's
Irr.	$B(f)$	KM03	1.412×10^{-2}	–
		Algorithm 2 in [1]	1.463×10^{-2}	1.036
		Algorithm C	1.473×10^{-2}	1.043
	$\tilde{B}(f)$	KM03 (Polytope)	1.639×10^{-2}	–
		Algorithm 1	1.661×10^{-2}	1.013
		Algorithm 2	1.680×10^{-2}	1.024
		Algorithm C	1.703×10^{-2}	1.038
Red.	$B(f)$	KM03	1.074×10^{-6}	–
		Algorithm 2 in [1]	1.083×10^{-6}	1.008
	$\tilde{B}(f)$	KM03 (Polytope)	2.145×10^{-6}	–
		Algorithm 1	2.177×10^{-6}	1.015
		Algorithm 2	2.204×10^{-6}	1.027

The methods revised by the former and this, are more time-consuming than the originals though their separation bounds are better. The reason is that we have to compute singular values after deleting unnecessary rows. Furthermore, the author wishes to thank the anonymous referees for their suggestions.

References

1. Nagasaka, K.: Towards more accurate separation bounds of empirical polynomials. SIGSAM/CCA **38** (2004) 119–129
2. Kaltofen, E.: Effective noether irreducibility forms and applications. J. Computer and System Sciences **50** (1995) 274–295
3. Nagasaka, K.: Towards certified irreducibility testing of bivariate approximate polynomials. In: Proc. ISSAC '02. (2002) 192–199
4. Sasaki, T.: Approximate multivariate polynomial factorization based on zero-sum relations. In: Proc. ISSAC 2001. (2001) 284–291
5. Nagasaka, K.: Neighborhood irreducibility testing of multivariate polynomials. In: Proc. CASC 2003. (2003) 283–292
6. Kaltofen, E., May, J.: On approximate irreducibility of polynomials in several variables. In: Proc. ISSAC '03. (2003) 161–168
7. Ruppert, W.M.: Reducibility of polynomials $f(x, y)$ modulo p. J. Number Theory **77** (1999) 62–70
8. Gao, S., Rodrigues, V.M.: Irreducibility of polynomials modulo p via newton polytopes. J. Number Theory **101** (2003) 32–47
9. Gao, S., Kaltofen, E., May, J., Yang, Z., Zhi, L.: Approximate factorization of multivariate polynomials via differential equations. In: Proc. ISSAC '04. (2004) 167–174