Computing the Nearest Singular Univariate Polynomials with Given Root Multiplicities

Zijia Li [a,b] Lihong Zhi [b,*]

[a] *Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Linz 4040 Austria*

[b] *Key Lab of Mathematics Mechanization, AMSS, Beijing 100190 China*

Abstract

In this paper, we derive explicit expressions for the nearest singular polynomials with given root multiplicities and its distance to the given polynomial. These expressions can be computed recursively. These results extend previous results of (Zhi et al., 2004; Zhi and Wu, 1998).

Key words: Multiplicity structure, nearest singular polynomial, multiple roots.

1 Introduction

The problem of finding the nearest polynomial with given root structure has been considered by many people (Corless and Rezvani, 2007; Hitz and Kaltofen, 1998; Hitz et al., 1999; Karmarkar and Lakshman, 1996; Pope and Szanto, 2009; Rezvani and Corless, 2005; Stetter, 1999, 2004; Zeng, 2005; Zhi et al., 2004; Zhi and Wu, 1998). Substantial progress has been made by Pope and Szanto in (Pope and Szanto, 2009). They extended previous results from the univariate case to the multivariate case and presented a symbolic-numeric method for finding the closest multivariate polynomial system with given root multiplicities. Motivated by the interesting results in (Pope and Szanto, 2009), we derive explicit expressions of the nearest singular polynomials, which extend the results in (Zhi et al., 2004; Zhi and Wu, 1998) to arbitrary given multiplicity structure.

* Corresponding author.
Email addresses: zijia.li@ricam.oeaw.ac.at (Zijia Li), lzhi@mmrc.iss.ac.cn (Lihong Zhi).

Problem. Given a monic univariate polynomial $f \in \mathbb{C}[x]$ with degree m and the multiplicity structure $\mathbf{k} = (k_1, k_2, \ldots, k_s) \in \mathcal{N}^s_{\geq 1}$. Let $n = \sum_{j=1}^{s} k_j \leq m$, we want to find a polynomial $h_{\mathbf{k}} \in \mathbb{C}[x]$ and $z_1, \ldots, z_s \in \mathbb{C}$ such that

$$h_{\mathbf{k}} = \Pi_{i=1}^{s}(x - z_i)^{k_i}(x^{m-n} + \sum_{j=1}^{m-n} \phi_j x^{m-n-j}), \quad \phi_j \in \mathbb{C}, \tag{1}$$

and $\mathcal{N}_m^{(\mathbf{k})} = \parallel f - h_{\mathbf{k}} \parallel^2$ is minimal, where $\parallel f - h_{\mathbf{k}} \parallel^2$ is the square of the l^2-norm of its coefficient vector.

Prior works. In (Pope and Szanto, 2009), they generalized the explicit formula of $\mathcal{N}_m^{(\mathbf{k})}$ in (Zhi et al., 2004; Zhi and Wu, 1998) to the case $s > 1$:

$$\mathcal{N}_m^{(\mathbf{k})} = \mathbf{f}_{\mathbf{k}}^* \mathbf{M}_{\mathbf{k}}^{-1} \mathbf{f}_{\mathbf{k}}, \tag{2}$$

where they defined the column vector

$$\mathbf{f}_i = \left[f(z_i), f^{'}(z_i), \ldots, f^{(k_i-1)}(z_i) \right]^T, \tag{3}$$

and

$$\mathbf{f}_{\mathbf{k}} = \left[\mathbf{f}_1{}^T, \mathbf{f}_2{}^T, \ldots, \mathbf{f}_s{}^T \right]^T, \tag{4}$$

and $\mathbf{f}_{\mathbf{k}}^*$ denotes the conjugate transpose of $\mathbf{f}_{\mathbf{k}}$. Here and hereafter, $f^{(j)}(z_i)$ denotes the evaluation of the j-th derivative of $f(x)$ at z_i, and $\mathbf{M}_{\mathbf{k}}^{-1}$ denotes the inverse matrix of $\mathbf{M}_{\mathbf{k}}$. The matrix $\mathbf{M}_{\mathbf{k}}$ can be decomposed into

$$\mathbf{M}_{\mathbf{k}} = \mathbf{V}_{\mathbf{k}} \mathbf{V}_{\mathbf{k}}^*, \tag{5}$$

where

$$\mathbf{V}_{\mathbf{k}} = \begin{bmatrix} 1 & z_1 & \ldots & z_1^{m-1} \\ 0 & 1 & \ldots & (m-1)z_1^{m-2} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \prod_{i=1}^{k_1-1}(m-i)z_1^{m-k_1} \\ \vdots & \vdots & & \vdots \\ 1 & z_s & \ldots & z_s^{m-1} \\ 0 & 1 & \ldots & (m-1)z_s^{m-2} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & \prod_{i=1}^{k_s-1}(m-i)z_s^{m-k_s} \end{bmatrix} \in \mathbb{C}^{(k_1+\cdots+k_s)\times m}. \tag{6}$$

We define

$$\lambda_{i,j} = \sum_{t=0}^{m-1} (z_i \overline{z}_j)^t,$$

$$\mathbf{\Lambda}_{k_i,k_j} = \begin{bmatrix} \lambda_{i,j} & \frac{\partial \lambda_{i,j}}{\partial \overline{z}_j} & \cdots & \frac{\partial^{k_j-1}\lambda_{i,j}}{\partial \overline{z}_j{}^{k_j-1}} \\ \frac{\partial \lambda_{i,j}}{\partial z_i} & \frac{\partial^2 \lambda_{i,j}}{\partial z_i \partial \overline{z}_j} & \cdots & \frac{\partial^{k_j}\lambda_{i,j}}{\partial z_i \partial \overline{z}_j{}^{k_j-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^{k_i-1}\lambda_{i,j}}{\partial z_i^{k_i-1}} & \frac{\partial^{k_i}\lambda_{i,j}}{\partial z_i^{k_i-1}\partial \overline{z}_j} & \cdots & \frac{\partial^{k_i+k_j-2}\lambda_{i,j}}{\partial z_i^{k_i-1}\partial \overline{z}_j{}^{k_j-1}} \end{bmatrix} \in \mathbb{C}^{k_i \times k_j}, \tag{7}$$

where $i, j = 1, \ldots, s$, and $\overline{z}_j$ denotes the conjugate of z_j.

Note 1 *These partial derivatives denote the symbolic evaluation of* $\frac{\partial^{u+v}\lambda}{\partial x^u \partial y^v}$ *at* $x = z_i$ *and* $y = \overline{z}_j$, *where* $\lambda = \sum_{t=0}^{m-1}(xy)^t$. *We will use these notations throughout this paper.*

From (5), (6), (7), we have

$$\mathbf{M_k} = \begin{bmatrix} \mathbf{\Lambda}_{k_1,k_1} & \mathbf{\Lambda}_{k_1,k_2} & \cdots & \mathbf{\Lambda}_{k_1,k_s} \\ \mathbf{\Lambda}_{k_2,k_1} & \mathbf{\Lambda}_{k_2,k_2} & \cdots & \mathbf{\Lambda}_{k_2,k_s} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Lambda}_{k_s,k_1} & \mathbf{\Lambda}_{k_s,k_2} & \cdots & \mathbf{\Lambda}_{k_s,k_s} \end{bmatrix} \in \mathbb{C}^{(k_1+\cdots+k_s)\times(k_1+\cdots+k_s)}. \tag{8}$$

We denote the determinant of $\mathbf{M_k}$ by

$$q_{\mathbf{k}} = \det \mathbf{M_k}. \tag{9}$$

It should be noted that $q_{\mathbf{k}}$ is always different from zero, see Theorem 1 in (Zhi et al., 2004) and Definition 3 in (Pope and Szanto, 2009).

Main contribution. In previous papers (Zhi et al., 2004; Zhi and Wu, 1998), they studied the case of finding the nearest singular polynomial with one multiple root $\mathbf{k} = (k)$ and gave recursive formulas related to the determination of the nearest singular polynomials for consecutive multiplicity k. In (Pope and Szanto, 2009), they extended results in (Zhi and Wu, 1998) to find the nearest multivariate polynomial system to a given one which has roots with prescribed multiplicity structure. In univariate case, Pope and Szanto generalized the explicit formula for the gradient of the distance function to the $s > 1$ case and gave a component-wise formula for the Gauss-Newton iteration to find the optimum. We focus on extending symbolic recursive relations in (Zhi et al., 2004; Zhi and Wu, 1998) for determining the minimal distance and the nearest singular polynomial to the case when the input univariate polynomial is near to a polynomial with several multiple roots. Moreover, in (Zhi et al., 2004), they derived explicit expressions of the nearest singular polynomial for $k_1 = 2, 3, 4$. We generalize them to the case of roots with any given multiplicities $\mathbf{k} = (k_1, k_2, \ldots, k_s) \in \mathcal{N}^s_{\geq 1}$, where $s \geq 1$.

Structure of the paper. The remaining part of the paper is set up as follows. In Section 2, we derive the explicit formula of $h_{\mathbf{k}}(x)$ defined in (1) for $s = 1$. In Section 3, we generalize explicit recursive formulas in (Zhi et al., 2004; Zhi and Wu, 1998) to the case $s > 1$. We illustrate two numerical examples in Section 4.

2 The Case $s = 1$

Let us consider the simplest case where $s = 1$. We take $k = k_1$, $\mathbf{k} = (k)$, $\mathbf{\Lambda}_k = \mathbf{M}_{\mathbf{k}} = \mathbf{\Lambda}_{k_1,k_1}$, $z = z_1$,

$$q_1 = \lambda_{1,1} = \sum_{t=0}^{m-1} (z\overline{z})^t, \;\; q_k = \det \mathbf{\Lambda}_k \tag{10}$$

for short. In (Zhi et al., 2004), they derived explicit expressions of the nearest singular polynomial h_k for $k = 2, 3, 4$. We generalize them to the arbitrary integer $k > 1$ case:

$$h_{k+1}(x) = \frac{\det \mathbf{H}_{k+1}}{q_k}, \tag{11}$$

where

$$\mathbf{H}_{k+1} = \begin{bmatrix} \mathbf{\Lambda}_k & \mathbf{f}_1 \\ \mathbf{q}_k(x) & f(x) \end{bmatrix},$$

$$\mathbf{f}_1 = \left[f(z), f'(z), \ldots, f^{(k-1)}(z) \right]^T,$$

$$\mathbf{q}_k(x) = \left[q_{11}(x), \frac{\partial q_{11}(x)}{\partial \overline{z}}, \ldots, \frac{\partial^{k-1} q_{11}(x)}{\partial \overline{z}^{k-1}} \right],$$

and

$$q_{11}(x) = \sum_{t=0}^{m-1} (\overline{z}x)^t. \tag{12}$$

Another expression for $\mathbf{q}_k(x)$ is

$$\mathbf{q}_k(x) = \mathbf{v}\mathbf{V}_k^*, \tag{13}$$

where

$$\mathbf{v} = \left[1, x, \ldots, x^{k-1}, \ldots, x^{m-1} \right], \tag{14}$$

and $\mathbf{V}_k^*$ is defined in Section 1 for the case $s = 1$.

Theorem 1 *For $k > 0$, suppose the minimum of the $\mathcal{N}_m^{(k+1)}$ is attained at z, then the nearest singular polynomial with a root of multiplicity $k+1$ is $h_{k+1}(x)$ defined by (11).*

Before we give the proof, we define the determinant of $\mathbf{P}_{k+1}$ by

$$p_{k+1} = \det \mathbf{P}_{k+1}, \tag{15}$$

where

$$\mathbf{P}_{k+1} = \begin{bmatrix} \mathbf{\Lambda}_k & \mathbf{f}_1 \\ \mathbf{w}_k^* & f^{(k)}(z) \end{bmatrix}, k > 0, \tag{16}$$

$$\mathbf{w}_k = \left[\frac{\partial^k q_1}{\partial \overline{z}^k}, \frac{\partial^{k+1} q_1}{\partial z \partial \overline{z}^k}, \ldots, \frac{\partial^{2k-1} q_1}{\partial z^{k-1} \partial \overline{z}^k} \right]^T, \tag{17}$$

and $p_1 = f(z)$.

Note 2 *For $k > 0$, suppose the minimum of the $\mathcal{N}_m^{(k+1)}$ is attained at z, then according to Theorem 5 in (Zhi and Wu, 1998), we have $p_{k+1} = 0$.*

Proof of Theorem 1. According to the definition of $h_{k+1}(x)$, two rows are equal in the matrix $\mathbf{H}_{k+1}$ when taken derivatives by x and evaluated at $x = z$, hence we have

$$h_{k+1}(z) = h'_{k+1}(z) = \cdots = h_{k+1}^{(k-1)}(z) = 0. \tag{18}$$

Since the minimum of $\mathcal{N}_m^{(k+1)}$ is attained at z, we derive that

$$h_{k+1}^{(k)}(z) = \frac{p_{k+1}}{q_k} = 0. \tag{19}$$

It should be noted that $p_{k+1} = 0$ follows from Note 2.

Furthermore, from the definition of $h_{k+1}(x)$ we have

$$h_{k+1}(x) = f(x) - \mathbf{q}_k(x) \mathbf{\Lambda}_k^{-1} \mathbf{f}_1.$$

As $\mathbf{\Lambda}_k$ is a Hermitian matrix (Theorem 1 in (Zhi et al., 2004)),

$$\begin{aligned} \| h_{k+1} - f \|^2 &= \| \mathbf{q}_k(x) \mathbf{\Lambda}_k^{-1} \mathbf{f}_1 \|^2 \\ &= \| \mathbf{v} \mathbf{V}_k^* \mathbf{\Lambda}_k^{-1} \mathbf{f}_1 \|^2 \\ &= \mathbf{f}_1^* \mathbf{\Lambda}_k^{-1} \mathbf{V}_k \mathbf{V}_k^* \mathbf{\Lambda}_k^{-1} \mathbf{f}_1 \\ &= \mathbf{f}_1^* \mathbf{\Lambda}_k^{-1} \mathbf{f}_1 \\ &= \mathcal{N}_m^{(k)} \\ &= \mathcal{N}_m^{(k+1)}. \end{aligned}$$

The last equality follows from $p_{k+1} = 0$ and Theorem 5 in (Zhi et al., 2004). $\square$

Let $h_1(x) = f(x)$ and $q_{11}(x) = \sum_{i=0}^{m-1} (\overline{z}x)^i$, we define

$$q_{k+1,k+1}(x) = \det \mathbf{Q}_{k+1}, \tag{20}$$

where

$$\mathbf{Q}_{k+1} = \begin{bmatrix} \mathbf{\Lambda}_k & \mathbf{w}_k \\ \mathbf{q}_k(x) & \frac{\partial^k q_{11}(x)}{\partial \overline{z}^k} \end{bmatrix}. \tag{21}$$

There is an alternative method to determine $h_k(x)$ and $q_{kk}(x)$ for $k > 1$ recursively.

Theorem 2 *For $k > 0$, the nearest singular polynomial $h_{k+1}(x)$ with a root z of multiplicity $k+1$ can be obtained recursively by the following formulas:*

$$h_{i+1}(x) = h_i(x) - \frac{h_i^{(i-1)}(z)}{q_i} q_{i,i}(x), \tag{22}$$

where

$$q_{i,i}(x) = \frac{1}{q_{i-2}} \left(q_{i-1} \frac{\partial q_{i-1,i-1}(x)}{\partial \overline{z}} - \frac{\partial q_{i-1}}{\partial \overline{z}} q_{i-1,i-1}(x) \right), \tag{23}$$

for $i = 2, \ldots, k$ and $h_2(x) = f(x) - \frac{f(z)}{q_1} q_{11}(x)$, $q_0 = 1$.

Techniques used in the proofs of Theorem 5 in (Zhi et al., 2004) and Theorem 2 in (Zhi and Wu, 1998) have been generalized to show the correctness of the recursive relations (22) and (23).

Proof of Theorem 2. Firstly, for $i = 2, \ldots, k$, let

$$\mathbf{e}_i = \left[f(z), f'(z), \ldots, f^{(i-1)}(z) \right]^T,$$

then

$$\mathbf{H}_{i+1} = \begin{bmatrix} \mathbf{\Lambda}_i & \mathbf{e}_i \\ \mathbf{q}_i(x) & f(x) \end{bmatrix}.$$

We have

$$\begin{aligned}
h_{i+1}(x) - h_i(x) &= f(x) - \mathbf{q}_i(x)\mathbf{\Lambda}_i^{-1}\mathbf{e}_i - (f(x) - \mathbf{q}_{i-1}(x)\mathbf{\Lambda}_{i-1}^{-1}\mathbf{e}_{i-1}) \\
&= \mathbf{q}_{i-1}(x)\mathbf{\Lambda}_{i-1}^{-1}\mathbf{e}_{i-1} - \mathbf{q}_i(x)\mathbf{\Lambda}_i^{-1}\mathbf{e}_i \\
&= -\frac{q_{i-1}}{q_i}\left(\frac{\partial^{i-1}q_{11}(x)}{\partial \overline{z}^{i-1}} - \mathbf{q}_{i-1}(x)\mathbf{\Lambda}_{i-1}^{-1}\mathbf{w}_{i-1}\right) \\
&\quad \left(f^{(i-1)}(z) - \mathbf{w}_{i-1}^{*}\mathbf{\Lambda}_{i-1}^{-1}\mathbf{e}_{i-1}\right) \\
&= -\frac{q_{i-1}}{q_i}\frac{q_{i,i}(x)}{q_{i-1}}\frac{p_i}{q_{i-1}} \\
&= -\frac{h_i^{(i-1)}(z)}{q_i}q_{i,i}(x).
\end{aligned}$$

Secondly, by the definition of $q_{i,i}(x)$ in (20, 21), we obtain:

$$q_{2,2}(x) = q_1\frac{\partial q_{11}(x)}{\partial \overline{z}} - q_{11}(x)\frac{\partial q_1}{\partial \overline{z}}.$$

Furthermore, for $i = 3, \ldots, k$, we have

$$\begin{aligned}
q_{i,i}(x) &= q_{i-1}\left(\frac{\partial^{i-1}q_{11}(x)}{\partial \overline{z}^{i-1}} - \mathbf{q}_{i-1}(x)\mathbf{\Lambda}_{i-1}^{-1}\mathbf{w}_{i-1}\right) \\
&= q_{i-1}\left(\frac{\partial^{i-1}q_{11}(x)}{\partial \overline{z}^{i-1}} - \mathbf{q}_{i-2}(x)\mathbf{\Lambda}_{i-2}^{-1}\frac{\partial \mathbf{w}_{i-2}}{\partial \overline{z}}\right) - q_{i-2} \\
&\quad \left(\frac{\partial^{i-2}q_{11}(x)}{\partial \overline{z}^{i-2}} - \mathbf{q}_{i-2}(x)\mathbf{\Lambda}_{i-2}^{-1}\mathbf{w}_{i-2}\right)\left(\frac{\partial^{2i-3}q_1}{\partial z^{i-2}\partial \overline{z}^{i-1}} - \mathbf{w}_{i-2}^{*}\mathbf{\Lambda}_{i-2}^{-1}\frac{\partial \mathbf{w}_{i-2}}{\partial \overline{z}}\right) \\
&= q_{i-1}\left(\frac{\partial^{i-1}q_{11}(x)}{\partial \overline{z}^{i-1}} - \mathbf{q}_{i-2}(x)\mathbf{\Lambda}_{i-2}^{-1}\frac{\partial \mathbf{w}_{i-2}}{\partial \overline{z}} - \frac{q_{i-2}}{q_{i-1}}\frac{q_{i-1,i-1}(x)}{q_{i-2}}\frac{\partial q_{i-1}}{\partial \overline{z}}\frac{1}{q_{i-2}}\right) \\
&= \frac{1}{q_{i-2}}\left(q_{i-1}\frac{\partial q_{i-1,i-1}(x)}{\partial \overline{z}} - \frac{\partial q_{i-1}}{\partial \overline{z}}q_{i-1,i-1}(x)\right).
\end{aligned}$$

It should be noted that the third equality above is derived from

$$\begin{aligned}
\frac{\partial q_{i-1}}{\partial \overline{z}} &= q_{i-2}\left(\frac{\partial^{2i-3}q_1}{\partial z^{i-2}\partial \overline{z}^{i-1}} - \mathbf{w}_{i-2}^{*}\mathbf{\Lambda}_{i-2}^{-1}\frac{\partial \mathbf{w}_{i-2}}{\partial \overline{z}}\right), \\
\frac{\partial q_{i-1,i-1}(x)}{\partial \overline{z}} &= q_{i-2}\left(\frac{\partial^{i-1}q_{11}(x)}{\partial \overline{z}^{i-1}} - \mathbf{q}_{i-2}(x)\mathbf{\Lambda}_{i-2}^{-1}\frac{\partial \mathbf{w}_{i-2}}{\partial \overline{z}}\right),
\end{aligned}$$

where $\mathbf{w}_{i-2}$ comes from (17). □

Note 3 *For any given integer $k > 1$, suppose the minimum of $\mathcal{N}_m^{(k)}$ is attained at z, then the nearest singular polynomial with a root of multiplicity k can be obtained by substituting z into $h_k(x)$ computed by formulas (22) and (23). This is true by Theorem 1.*

3 The Case $s > 1$

3.1 Explicit Recursive Expression

In this section, let $\mathbf{k}$, m, n and s be given as in the introduction. We denote the determinant of $\mathbf{P}_{\mathbf{k}}$ by

$$p_{\mathbf{k}} = \det \mathbf{P}_{\mathbf{k}}, \tag{24}$$

where

$$\mathbf{P}_{\mathbf{k}} = \begin{bmatrix} \mathbf{\Lambda}_{k_1,k_1} & \mathbf{\Lambda}_{k_1,k_2} & \dots & \mathbf{\Lambda}_{k_1,k_s-1} & \mathbf{f}_1 \\ \mathbf{\Lambda}_{k_2,k_1} & \mathbf{\Lambda}_{k_2,k_2} & \dots & \mathbf{\Lambda}_{k_2,k_s-1} & \mathbf{f}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{\Lambda}_{k_s,k_1} & \mathbf{\Lambda}_{k_s,k_2} & \dots & \mathbf{\Lambda}_{k_s,k_s-1} & \mathbf{f}_s \end{bmatrix} \in \mathbb{C}^{(k_1+\cdots+k_s)\times(k_1+\cdots+k_s)}. \tag{25}$$

We derive an alternative explicit expression for $\mathcal{N}_m^{(\mathbf{k})}$ in terms of $q_{\mathbf{k}}$ and $p_{\mathbf{k}}$. Some recursive relations to generate $q_{\mathbf{k}}$ and $p_{\mathbf{k}}$ are also provided. These expressions extend results in (Zhi et al., 2004; Zhi and Wu, 1998).

We search for the nearest polynomials with the roots of multiplicity structure $\mathbf{r}_i$:

$$\mathbf{r}_i = \begin{cases} (i) & 1 \le i \le k_1, \\ (k_1, i - k_1) & k_1 < i \le k_1 + k_2, \\ \vdots & \vdots \\ (k_1, k_2, k_3, \dots, i - \sum_{j=1}^{s-1} k_j) & \sum_{j=1}^{s-1} k_j < i \le n. \end{cases} \tag{26}$$

Theorem 3 *Let $\mathbf{k} = (k_1, k_2, \dots, k_s)$, $n = \sum_{j=1}^{s} k_j$ and $1 \le i \le n$, then the distance to the nearest singular polynomial with given root multiplicities is*

$$\mathcal{N}_m^{(\mathbf{k})} = \frac{p_{\mathbf{r}_1}\overline{p_{\mathbf{r}_1}}}{q_{\mathbf{r}_1}} + \frac{p_{\mathbf{r}_2}\overline{p_{\mathbf{r}_2}}}{q_{\mathbf{r}_1}q_{\mathbf{r}_2}} + \cdots + \frac{p_{\mathbf{r}_n}\overline{p_{\mathbf{r}_n}}}{q_{\mathbf{r}_{n-1}}q_{\mathbf{r}_n}}. \tag{27}$$

Proof. Similar to the proof of Theorem 5 in (Zhi et al., 2004), we divide $\mathbf{M}_{\mathbf{r}_{i+1}}$ into

$$\mathbf{M}_{\mathbf{r}_{i+1}} = \begin{bmatrix} \mathbf{M}_{\mathbf{r}_i} & \mathbf{w}_{\mathbf{r}_i} \\ \mathbf{w}_{\mathbf{r}_i}^* & \alpha \end{bmatrix}, \tag{28}$$

where the last column of $\mathbf{M}_{\mathbf{r}_{i+1}}$ is divided into $\mathbf{w}_{\mathbf{r}_i}$ and α. Since the matrix

$\mathbf{M}_{\mathbf{r}_i}$ is an invertible Hermitian matrix, the inverse of $\mathbf{M}_{\mathbf{r}_{i+1}}$ can be written as

$$\mathbf{M}_{\mathbf{r}_{i+1}}^{-1} = \begin{bmatrix} \mathbf{M}_{\mathbf{r}_i}^{-1} & -\mathbf{M}_{\mathbf{r}_i}^{-1}\mathbf{w}_{\mathbf{r}_i}\beta^{-1} \\ 0 & \beta^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ -\mathbf{w}_{\mathbf{r}_i}^{*}\mathbf{M}_{\mathbf{r}_i}^{-1} & 1 \end{bmatrix}, \tag{29}$$

where

$$\beta = \alpha - \mathbf{w}_{\mathbf{r}_i}^{*}\mathbf{M}_{\mathbf{r}_i}\mathbf{w}_{\mathbf{r}_i} = \frac{q_{\mathbf{r}_{i+1}}}{q_{\mathbf{r}_i}}. \tag{30}$$

If we divide $\mathbf{f}_{\mathbf{r}_{i+1}}$ into $\mathbf{f}_{\mathbf{r}_i}$ and γ, then we have

$$\begin{aligned} \mathcal{N}_m^{(\mathbf{r}_{i+1})} &= \mathbf{f}_{\mathbf{r}_{i+1}}^{*}\mathbf{M}_{\mathbf{r}_{i+1}}^{-1}\mathbf{f}_{\mathbf{r}_{i+1}} \\ &= \left[\mathbf{f}_{\mathbf{r}_i}^{*}, \gamma^{*}\right] \begin{bmatrix} \mathbf{M}_{\mathbf{r}_i}^{-1} & -\mathbf{M}_{\mathbf{r}_i}^{-1}\mathbf{w}_{\mathbf{r}_i}\beta^{-1} \\ 0 & \beta^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ -\mathbf{w}_{\mathbf{r}_i}^{*}\mathbf{M}_{\mathbf{r}_i}^{-1} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{f}_{\mathbf{r}_i} \\ \gamma \end{bmatrix} \\ &= \mathbf{f}_{\mathbf{r}_i}^{*}\mathbf{M}_{\mathbf{r}_i}^{-1}\mathbf{f}_{\mathbf{r}_i} + \beta^{-1}\left(\gamma^{*} - \mathbf{f}_{\mathbf{r}_i}^{*}\mathbf{M}_{\mathbf{r}_i}^{-1}\mathbf{w}_{\mathbf{r}_i}\right)\left(\gamma - \mathbf{w}_{\mathbf{r}_i}^{*}\mathbf{M}_{\mathbf{r}_i}^{-1}\mathbf{f}_{\mathbf{r}_i}\right) \\ &= \mathcal{N}_m^{(\mathbf{r}_i)} + \frac{q_{\mathbf{r}_i}}{q_{\mathbf{r}_{i+1}}}\frac{\overline{p_{\mathbf{r}_{i+1}}}}{q_{\mathbf{r}_i}}\frac{p_{\mathbf{r}_{i+1}}}{q_{\mathbf{r}_i}} \\ &= \mathcal{N}_m^{(\mathbf{r}_i)} + \frac{p_{\mathbf{r}_{i+1}}\overline{p_{\mathbf{r}_{i+1}}}}{q_{\mathbf{r}_i}q_{\mathbf{r}_{i+1}}}. \end{aligned}$$

□

There are also recursive relationships between $p_{\mathbf{r}_i}$ and $q_{\mathbf{r}_i}$ for $i = 1, \ldots, n$. Similar to the definition of $p_{\mathbf{k}}$ in (24), for an integer l, we denote the determinant

$$p_{\mathbf{k},l} = \det \mathbf{P}_{\mathbf{k},l}, \tag{31}$$

where

$$\mathbf{P}_{\mathbf{k},l} = \begin{bmatrix} \mathbf{\Lambda}_{k_1,k_1} & \mathbf{\Lambda}_{k_1,k_2} & \ldots & \mathbf{\Lambda}_{k_1,k_s-1} & \mathbf{g}_{1,l} \\ \mathbf{\Lambda}_{k_2,k_1} & \mathbf{\Lambda}_{k_2,k_2} & \ldots & \mathbf{\Lambda}_{k_2,k_s-1} & \mathbf{g}_{2,l} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{\Lambda}_{k_s,k_1} & \mathbf{\Lambda}_{k_s,k_2} & \ldots & \mathbf{\Lambda}_{k_s,k_s-1} & \mathbf{g}_{s,l} \end{bmatrix}, \tag{32}$$

and the evaluation vectors

$$\mathbf{g}_{j,l} = \left[\lambda_{j,l}, \frac{\partial \lambda_{j,l}}{\partial z_j}, \ldots, \frac{\partial^{k_j-1}\lambda_{j,l}}{\partial z_j^{k_j-1}}\right]^T, \quad j, l = 1, \ldots, s.$$

Furthermore, for $\theta = \sum_{j=1}^{t} k_j$, we have

$$\mathbf{r}_\theta = (k_1, \ldots, k_t).$$

Define the evaluation column vectors

$$\mathbf{f}_{\mathbf{r}_\theta} = \left[\mathbf{f}_1{}^T, \mathbf{f}_2{}^T, \ldots, \mathbf{f}_t{}^T\right]^T,$$

$$\mathbf{g}_{\mathbf{r}_\theta} = \left[\mathbf{g}_{1,t+1}{}^T, \mathbf{g}_{2,t+1}{}^T, \ldots, \mathbf{g}_{t,t+1}{}^T\right]^T,$$
$$\mathbf{g}_{\mathbf{r}_\theta,l} = \left[\mathbf{g}_{1,l}{}^T, \mathbf{g}_{2,l}{}^T, \ldots, \mathbf{g}_{t,l}{}^T\right]^T.$$

Combining these notations with Theorem 3, we derive the following explicit expressions

$$\mathbf{g}^*_{\mathbf{r}_\theta}\mathbf{M}^{-1}_{\mathbf{r}_\theta}\mathbf{g}_{\mathbf{r}_\theta} = \frac{p_{\mathbf{r}_1,t+1}\overline{p_{\mathbf{r}_1,t+1}}}{q_{\mathbf{r}_1}} + \frac{p_{\mathbf{r}_2,t+1}\overline{p_{\mathbf{r}_2,t+1}}}{q_{\mathbf{r}_1}q_{\mathbf{r}_2}} + \cdots + \frac{p_{\mathbf{r}_\theta,t+1}\overline{p_{\mathbf{r}_\theta,t+1}}}{q_{\mathbf{r}_{\theta-1}}q_{\mathbf{r}_\theta}}, \tag{33}$$

$$\mathbf{g}^*_{\mathbf{r}_\theta}\mathbf{M}^{-1}_{\mathbf{r}_\theta}\mathbf{f}_{\mathbf{r}_\theta} = \frac{p_{\mathbf{r}_1}\overline{p_{\mathbf{r}_1,t+1}}}{q_{\mathbf{r}_1}} + \frac{p_{\mathbf{r}_2}\overline{p_{\mathbf{r}_2,t+1}}}{q_{\mathbf{r}_1}q_{\mathbf{r}_2}} + \cdots + \frac{p_{\mathbf{r}_\theta}\overline{p_{\mathbf{r}_\theta,t+1}}}{q_{\mathbf{r}_{\theta-1}}q_{\mathbf{r}_\theta}}, \tag{34}$$

$$\mathbf{g}^*_{\mathbf{r}_\theta}\mathbf{M}^{-1}_{\mathbf{r}_\theta}\mathbf{g}_{\mathbf{r}_\theta,l} = \frac{p_{\mathbf{r}_1,l}\overline{p_{\mathbf{r}_1,t+1}}}{q_{\mathbf{r}_1}} + \frac{p_{\mathbf{r}_2,l}\overline{p_{\mathbf{r}_2,t+1}}}{q_{\mathbf{r}_1}q_{\mathbf{r}_2}} + \cdots + \frac{p_{\mathbf{r}_\theta,l}\overline{p_{\mathbf{r}_\theta,t+1}}}{q_{\mathbf{r}_{\theta-1}}q_{\mathbf{r}_\theta}}, \tag{35}$$

where t and l are from 2 to s.

Note 4 *If $t+1=l$, equations (33) and (35) are same. Otherwise, as the lengths of vectors $\mathbf{g}_{\mathbf{r}_\theta}$ and $\mathbf{g}_{\mathbf{r}_\theta,l}$ are not equal, we can not exchange them with each other. In our algorithm, we only need to consider the case $t+1 \le l$.*

The following theorems give alternative methods to determine these $q_{\mathbf{r}_i}$, $p_{\mathbf{r}_i}$ and $p_{\mathbf{r}_i,l}$ recursively.

Theorem 4 *Let $\mathbf{k} = (k_1, k_2, \ldots, k_s)$, $n = \sum_{j=1}^s k_j$ and $i = 1, \ldots, n$, we have*

$$q_{\mathbf{r}_i} = q_i, \quad i = 1, \ldots, k_1. \tag{36}$$

If $i = \sum_{j=1}^t k_j + 1$ for some $t = 1, \ldots, s-1$, we have

$$\mathbf{r}_i = (k_1, k_2, \ldots, k_t, 1),$$

and

$$q_{\mathbf{r}_i} = q_{\mathbf{r}_{i-1}}\left(\lambda_{t+1,t+1} - \mathbf{g}^*_{\mathbf{r}_{i-1}}\mathbf{M}^{-1}_{\mathbf{r}_{i-1}}\mathbf{g}_{\mathbf{r}_{i-1}}\right). \tag{37}$$

Otherwise, there exist two integers d, t with $1 < d \le k_{t+1}$ and $1 \le t \le s-1$ such that

$$\mathbf{r}_{i-d} = (k_1, k_2, \ldots, k_t),$$

and

$$q_{\mathbf{r}_{i-1}}\frac{\partial^2 q_{\mathbf{r}_{i-1}}}{\partial z_{t+1}\partial \overline{z}_{t+1}} - \frac{\partial q_{\mathbf{r}_{i-1}}}{\partial z_{t+1}}\frac{\partial q_{\mathbf{r}_{i-1}}}{\partial \overline{z}_{t+1}} = q_{\mathbf{r}_{i-2}}q_{\mathbf{r}_i}. \tag{38}$$

Proof. Firstly, if $s=1$ we obtain the recursive formula by Theorem 2 in (Zhi and Wu, 1998)

$$q_{\mathbf{r}_i} = q_i, \quad i = 1, \ldots, k_1.$$

If

$$\mathbf{r}_i = (k_1, k_2, \ldots, k_t, 1),$$

then

$$q_{\mathbf{r}_i} = \det \begin{vmatrix} \mathbf{\Lambda}_{k_1,k_1} & \mathbf{\Lambda}_{k_1,k_2} & \dots & \mathbf{\Lambda}_{k_1,k_t} & \mathbf{g}_{1,t+1} \\ \mathbf{\Lambda}_{k_2,k_1} & \mathbf{\Lambda}_{k_2,k_2} & \dots & \mathbf{\Lambda}_{k_2,k_t} & \mathbf{g}_{2,t+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{\Lambda}_{k_t,k_1} & \mathbf{\Lambda}_{k_t,k_2} & \dots & \mathbf{\Lambda}_{k_t,k_t} & \mathbf{g}_{t,t+1} \\ \mathbf{g}^*_{1,t+1} & \mathbf{g}^*_{2,t+1} & \dots & \mathbf{g}^*_{t,t+1} & \lambda_{t+1,t+1} \end{vmatrix}.$$

So we have

$$q_{\mathbf{r}_i} = q_{\mathbf{r}_{i-1}} \left(\lambda_{t+1,t+1} - \mathbf{g}^*_{\mathbf{r}_{i-1}} \mathbf{M}^{-1}_{\mathbf{r}_{i-1}} \mathbf{g}_{\mathbf{r}_{i-1}} \right).$$

If

$$\mathbf{r}_{i-d} = (k_1, k_2, \dots, k_t),$$

applying the Gaussian elimination

$$\begin{bmatrix} \mathbf{I} & 0 & 0 \\ -u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} & 1 & 0 \\ -u_2^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_1 & u_2 \\ u_1^* & \alpha & \xi \\ u_2^* & \beta & \eta \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & \begin{bmatrix} u_1 & u_2 \end{bmatrix} \\ 0 & \begin{bmatrix} \alpha & \xi \\ \beta & \eta \end{bmatrix} - \begin{bmatrix} u_1^* \\ u_2^* \end{bmatrix} \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} \begin{bmatrix} u_1 & u_2 \end{bmatrix} \end{bmatrix},$$

we have the following equalities

$$\begin{aligned} q_{\mathbf{r}_i} &= \det \mathbf{M}_{\mathbf{r}_i} \\ &= \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_1 & u_2 \\ u_1^* & \alpha & \xi \\ u_2^* & \beta & \eta \end{vmatrix} \\ &= \det \mathbf{M}_{\mathbf{r}_{i-2}} \det \left(\begin{bmatrix} \alpha & \xi \\ \beta & \eta \end{bmatrix} - \begin{bmatrix} u_1^* \\ u_2^* \end{bmatrix} \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} \begin{bmatrix} u_1 & u_2 \end{bmatrix} \right) \\ &= \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\alpha - u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_1 \right) \left(\eta - u_2^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_2 \right) \\ &- \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\beta - u_2^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_1 \right) \left(\xi - u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_2 \right), \end{aligned}$$

and

$$q_{\mathbf{r}_{i-1}} = \det \mathbf{M}_{\mathbf{r}_{i-1}} = \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_1 \\ u_1^* & \alpha \end{vmatrix} = \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\alpha - u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_1 \right),$$

$$q_{\mathbf{r}_{i-2}} = \det \mathbf{M}_{\mathbf{r}_{i-2}}.$$

In the expression of $\det \mathbf{M}_{\mathbf{r}_{i-1}}$, the partial derivative of the lth row of $\mathbf{M}_{\mathbf{r}_{i-1}}$ with respect to z_{t+1} is zero for $1 \le l \le \sum_{j=1}^t k_j$, and the partial derivative of the lth row of $\mathbf{M}_{\mathbf{r}_{i-1}}$ with respect to z_{t+1} is the $(l+1)$th row of $\mathbf{M}_{\mathbf{r}_{i-1}}$ for $\sum_{j=1}^t k_j \le l < i-1$, but the partial derivative of the last row of $\mathbf{M}_{\mathbf{r}_{i-1}}$ is the last row of $\mathbf{M}_{\mathbf{r}_i}$ upon deletion of its last element; same facts exist for the derivatives of the columns with respect to $\overline{z}_{t+1}$. Hence, we have

$$\frac{\partial q_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = \frac{\partial \det \mathbf{M}_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_1 \\ u_2^* & \beta \end{vmatrix} = \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\beta - u_2^* \mathbf{M}_{\mathbf{r}_{i-2}}^{-1} u_1 \right),$$

$$\frac{\partial q_{\mathbf{r}_{i-1}}}{\partial \overline{z}_{t+1}} = \frac{\partial \det \mathbf{M}_{\mathbf{r}_{i-1}}}{\partial \overline{z}_{t+1}} = \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_2 \\ u_1^* & \xi \end{vmatrix} = \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\xi - u_1^* \mathbf{M}_{\mathbf{r}_{i-2}}^{-1} u_2 \right),$$

$$\frac{\partial^2 q_{\mathbf{r}_{i-1}}}{\partial z_{t+1} \partial \overline{z}_{t+1}} = \frac{\partial^2 \det \mathbf{M}_{\mathbf{r}_{i-1}}}{\partial z_{t+1} \partial \overline{z}_{t+1}} = \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_2 \\ u_2^* & \eta \end{vmatrix} = \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\eta - u_2^* \mathbf{M}_{\mathbf{r}_{i-2}}^{-1} u_2 \right).$$

Then we obtain the equality (38). □

Theorem 5 *Let* $\mathbf{k} = (k_1, k_2, \ldots, k_s)$, $n = \sum_{j=1}^{s} k_j$ *and* $i = 1, \ldots, n$, *we have*

$$p_{\mathbf{r}_i} = p_i, \quad i = 1, \ldots, k_1. \tag{39}$$

If $i = \sum_{j=1}^{t} k_j + 1$ *for some* $t = 1, \ldots, s-1$, *we have*

$$\mathbf{r}_i = (k_1, k_2, \ldots, k_t, 1),$$

and

$$p_{\mathbf{r}_i} = q_{\mathbf{r}_{i-1}} \left(f(z_{t+1}) - \mathbf{g}_{\mathbf{r}_{i-1}}^* \mathbf{M}_{\mathbf{r}_{i-1}}^{-1} \mathbf{f}_{\mathbf{r}_{i-1}} \right). \tag{40}$$

Otherwise, there exist two integers d, t *with* $1 < d \leq k_{t+1}$ *and* $1 \leq t \leq s-1$ *such that*

$$\mathbf{r}_{i-d} = (k_1, k_2, \ldots, k_t),$$

and

$$q_{\mathbf{r}_{i-1}} \frac{\partial p_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} - p_{\mathbf{r}_{i-1}} \frac{\partial q_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = q_{\mathbf{r}_{i-2}} p_{\mathbf{r}_i}. \tag{41}$$

Proof. Firstly, for $s = 1$ case, we can obtain the recursive formula by Theorem 2 and Theorem 4 in (Zhi and Wu, 1998)

$$p_{\mathbf{r}_i} = p_i, \quad i = 1, \ldots, k_1.$$

If

$$\mathbf{r}_i = (k_1, k_2, \ldots, k_t, 1),$$

then

$$p_{\mathbf{r}_i} = \det \begin{vmatrix} \mathbf{\Lambda}_{k_1,k_1} & \mathbf{\Lambda}_{k_1,k_2} & \ldots & \mathbf{\Lambda}_{k_1,k_t} & \mathbf{f}_1 \\ \mathbf{\Lambda}_{k_2,k_1} & \mathbf{\Lambda}_{k_2,k_2} & \ldots & \mathbf{\Lambda}_{k_2,k_t} & \mathbf{f}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{\Lambda}_{k_t,k_1} & \mathbf{\Lambda}_{k_t,k_2} & \ldots & \mathbf{\Lambda}_{k_t,k_t} & \mathbf{f}_t \\ \mathbf{g}_{1,t+1}^* & \mathbf{g}_{2,t+1}^* & \ldots & \mathbf{g}_{t,t+1}^* & f(z_{t+1}) \end{vmatrix}.$$

So we have

$$p_{\mathbf{r}_i} = q_{\mathbf{r}_{i-1}} \left(f(z_{t+1}) - \mathbf{g}^*_{\mathbf{r}_{i-1}} \mathbf{M}^{-1}_{\mathbf{r}_{i-1}} \mathbf{f}_{\mathbf{r}_{i-1}} \right).$$

If

$$\mathbf{r}_{i-d} = (k_1, k_2, \ldots, k_t),$$

applying the Gaussian elimination

$$\begin{bmatrix} \mathbf{I} & 0 \; 0 \\ -u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} & 1 \; 0 \\ -\frac{\partial u_1^*}{\partial z_{t+1}} \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} & 0 \; 1 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_1 & u_2 \\ u_1^* & \alpha & \beta \\ \frac{\partial u_1^*}{\partial z_{t+1}} & \frac{\partial \alpha}{\partial z_{t+1}} & \frac{\partial \beta}{\partial z_{t+1}} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & \begin{bmatrix} u_1 \; u_2 \end{bmatrix} \\ 0 & \begin{bmatrix} \alpha & \beta \\ \frac{\partial \alpha}{\partial z_{t+1}} & \frac{\partial \beta}{\partial z_{t+1}} \end{bmatrix} - \begin{bmatrix} u_1^* \\ \frac{\partial u_1^*}{\partial z_{t+1}} \end{bmatrix} \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} \begin{bmatrix} u_1 \; u_2 \end{bmatrix} \end{bmatrix},$$

we have

$$\begin{aligned}
p_{\mathbf{r}_i} &= \det \mathbf{P}_{\mathbf{r}_i} \\
&= \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_1 & u_2 \\ u_1^* & \alpha & \beta \\ \frac{\partial u_1^*}{\partial z_{t+1}} & \frac{\partial \alpha}{\partial z_{t+1}} & \frac{\partial \beta}{\partial z_{t+1}} \end{vmatrix} \\
&= \det \mathbf{M}_{\mathbf{r}_{i-2}} \det \left(\begin{bmatrix} \alpha & \beta \\ \frac{\partial \alpha}{\partial z_{t+1}} & \frac{\partial \beta}{\partial z_{t+1}} \end{bmatrix} - \begin{bmatrix} u_1^* \\ \frac{\partial u_1^*}{\partial z_{t+1}} \end{bmatrix} \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} \begin{bmatrix} u_1 \; u_2 \end{bmatrix} \right) \\
&= \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\alpha - u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_1 \right) \left(\frac{\partial \beta}{\partial z_{t+1}} - \frac{\partial u_1^*}{\partial z_{t+1}} \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_2 \right) \\
&- \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\frac{\partial \alpha}{\partial z_{t+1}} - \frac{\partial u_1^*}{\partial z_{t+1}} \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_1 \right) \left(\beta - u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_2 \right),
\end{aligned}$$

and

$$p_{\mathbf{r}_{i-1}} = \det \mathbf{P}_{\mathbf{r}_{i-1}} = \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_2 \\ u_1^* & \beta \end{vmatrix} = \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\beta - u_1^* \mathbf{M}^{-1}_{\mathbf{r}_{i-2}} u_2 \right),$$

$$q_{\mathbf{r}_{i-2}} = \det \mathbf{M}_{\mathbf{r}_{i-2}}.$$

In $\det \mathbf{P}_{\mathbf{r}_{i-1}}$, the partial derivative of the lth row of $\mathbf{P}_{\mathbf{r}_{i-1}}$ with respect to z_{t+1} is zero for $1 \le l \le \sum_{j=1}^t k_j$, and the partial derivative of the lth row of $\mathbf{P}_{\mathbf{r}_{i-1}}$ with respect to z_{t+1} is the $(l+1)$th row of $\mathbf{P}_{\mathbf{r}_{i-1}}$ for $\sum_{j=1}^t k_j \le l < i-1$, but the partial derivative of the last row of $\mathbf{P}_{\mathbf{r}_{i-1}}$ is the last row of $\mathbf{P}_{\mathbf{r}_i}$ upon deletion of the last second element; similar facts exist for $\det \mathbf{M}_{\mathbf{r}_{i-1}}$. Hence,

$$\frac{\partial p_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = \frac{\partial \det \mathbf{P}_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_2 \\ \frac{\partial u_1^*}{\partial z_{t+1}} & \frac{\partial \beta}{\partial z_{t+1}} \end{vmatrix} = \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\frac{\partial \beta}{\partial z_{t+1}} - \frac{\partial u_1^*}{\partial z_{t+1}} \mathbf{M}_{\mathbf{r}_{i-2}}^{-1} u_2 \right),$$

$$\frac{\partial q_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = \frac{\partial \det \mathbf{M}_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = \det \begin{vmatrix} \mathbf{M}_{\mathbf{r}_{i-2}} & u_1 \\ \frac{\partial u_1^*}{\partial z_{t+1}} & \frac{\partial \alpha}{\partial z_{t+1}} \end{vmatrix} = \det \mathbf{M}_{\mathbf{r}_{i-2}} \left(\frac{\partial \alpha}{\partial z_{t+1}} - \frac{\partial u_1^*}{\partial z_{t+1}} \mathbf{M}_{\mathbf{r}_{i-2}}^{-1} u_1 \right).$$

Then we obtain the equality (41). □

Theorem 6 *Let $\mathbf{k} = (k_1, k_2, \ldots, k_s)$, $n = \sum_{j=1}^{s} k_j$ and $i = 1, \ldots, n$. If $i = 1, \ldots, k_1$, we can obtain all $p_{\mathbf{r}_i,l}$ by replacing $\mathbf{f}_1$ in p_i with $\mathbf{g}_{1,l}$. If $i = \sum_{j=1}^{t} k_j + 1$ for some $t = 1, \ldots, s-1$, we have*

$$\mathbf{r}_i = (k_1, k_2, \ldots, k_t, 1),$$

and

$$p_{\mathbf{r}_i,l} = q_{\mathbf{r}_{i-1}} \left(\lambda_{t+1,l} - \mathbf{g}_{\mathbf{r}_{i-1}}^* \mathbf{M}_{\mathbf{r}_{i-1}}^{-1} \mathbf{g}_{\mathbf{r}_{i-1},l} \right). \tag{42}$$

Otherwise, there exist two integers d, t with $1 < d \leq k_{t+1}$ and $1 \leq t \leq s-1$ such that

$$\mathbf{r}_{i-d} = (k_1, k_2, \ldots, k_t),$$

and

$$q_{\mathbf{r}_{i-1}} \frac{\partial p_{\mathbf{r}_{i-1},l}}{\partial z_{t+1}} - p_{\mathbf{r}_{i-1},l} \frac{\partial q_{\mathbf{r}_{i-1}}}{\partial z_{t+1}} = q_{\mathbf{r}_{i-2}} p_{\mathbf{r}_i,l}. \tag{43}$$

The proof of Theorem 6 is similar to the proof of Theorem 5, since we only need to replace all $\mathbf{f}_j$ in $\mathbf{P}_{\mathbf{r}_i}$ by $\mathbf{g}_{j,l}$, where $j = 1, \ldots, s$.

3.2 Explicit Expression of the Nearest Singular Polynomial

Let $\mathbf{k}$, n, i and $\mathbf{r}_i$ be given as in the previous subsection. We introduce auxiliary polynomials $q_{\mathbf{k},\mathbf{k}}(x)$ and $h_{\mathbf{k}}(x)$ to obtain the generalized explicit expression of the nearest singular polynomial.

We denote the auxiliary polynomial

$$q_{\mathbf{k},\mathbf{k}}(x) = \det \mathbf{Q}_{\mathbf{k}}, \tag{44}$$

where

$$\mathbf{Q_k} = \begin{cases} \begin{bmatrix} \mathbf{M}_{\mathbf{r}_{n-1}} & \mathbf{g}_{\mathbf{r}_{n-1}} \\ \mathbf{q}_{\mathbf{r}_{n-1}}(x) & q_{1,s}(x) \end{bmatrix}, & k_s = 1, \\ \\ \begin{bmatrix} \mathbf{M}_{\mathbf{r}_{n-1}} & \frac{\partial \mathbf{M}_{\mathbf{r}_{n-1}}(\cdot, n-1)}{\partial \overline{z}_s} \\ \mathbf{q}_{\mathbf{r}_{n-1}}(x) & \frac{\partial^{k_s-1} q_{1,s}(x)}{\partial \overline{z}_s{}^{k_s-1}} \end{bmatrix}, & k_s > 1, \end{cases} \tag{45}$$

$$\mathbf{q}_{\mathbf{r}_{n-1}}(x) = \mathbf{v}\mathbf{V}^*_{\mathbf{r}_{n-1}}, q_{1,s}(x) = \sum_{i=0}^{m-1} (\overline{z}_s x)^i, \tag{46}$$

and $\mathbf{M}_{\mathbf{r}_{n-1}}(\cdot, n-1)$ denotes the last column of $\mathbf{M}_{\mathbf{r}_{n-1}}$. We define the polynomial

$$h_{\mathbf{k}}(x) = \frac{\det \mathbf{H_k}}{q_{\mathbf{r}_{n-1}}}, \tag{47}$$

where

$$\mathbf{H_k} = \begin{bmatrix} \mathbf{M}_{\mathbf{r}_{n-1}} & \mathbf{f}_{\mathbf{r}_{n-1}} \\ \mathbf{q}_{\mathbf{r}_{n-1}}(x) & f(x) \end{bmatrix}.$$

Note 5 *Suppose the minimum of $\mathcal{N}_m^{(\mathbf{k})}$ is attained at $z_1, \ldots, z_s$, according to Proposition 13 and Remark 14 in (Pope and Szanto, 2009), we have $p_{\mathbf{k}} = 0$.*

Theorem 7 *Let $\mathbf{k} = (k_1, \ldots, k_s)$, $n = \sum_{j=1}^{s} k_j$ and $\mathbf{r}_i$ be defined in (26) for $1 \leq i \leq n$. Suppose the minimum of $\mathcal{N}_m^{(\mathbf{k})}$ is attained at $z_1, \ldots, z_s$, then the nearest singular polynomial with roots of multiple structure $\mathbf{k}$ is $h_{\mathbf{k}}(x)$.*

Proof. If $k_s = 1$, similar to Theorem 1, we obtained

$$h_{\mathbf{k}}(z_j) = h'_{\mathbf{k}}(z_j) = \cdots = h_{\mathbf{k}}^{(k_j-1)}(z_j) = 0,$$

where $j = 1, \ldots, s-1$. Furthermore, as the minimum of $\mathcal{N}_m^{(\mathbf{k})}$ attained at $z_1, \ldots, z_s$, we have

$$h_{\mathbf{k}}(z_s) = \frac{p_{\mathbf{k}}}{q_{\mathbf{r}_{n-1}}} = 0$$

according to Note 5. If $k_s > 1$, similarly, we have

$$h_{\mathbf{k}}(z_j) = h'_{\mathbf{k}}(z_j) = \cdots = h_{\mathbf{k}}^{(k_j-1)}(z_j) = 0,$$

for $j = 1, \ldots, s-1$ and

$$h_{\mathbf{k}}(z_s) = h'_{\mathbf{k}}(z_s) = \cdots = h_{\mathbf{k}}^{(k_s-2)}(z_s) = 0.$$

Furthermore,

$$h_{\mathbf{k}}^{(k_s-1)}(z_s) = \frac{p_{\mathbf{k}}}{q_{\mathbf{r}_{n-1}}} = 0.$$

It follows that every z_j is a root of $h_{\mathbf{k}}(x)$ with multiplicity k_j. In the end, we have

$$\begin{aligned}
\parallel h_{\mathbf{k}} - f \parallel^2 &= \parallel \mathbf{q}_{\mathbf{r}_{n-1}}(x)\mathbf{M}_{\mathbf{r}_{n-1}}^{-1}\mathbf{f}_{\mathbf{r}_{n-1}} \parallel^2 \\
&= \parallel \mathbf{v}\mathbf{V}_{\mathbf{r}_{n-1}}^{*}\mathbf{M}_{\mathbf{r}_{n-1}}^{-1}\mathbf{f}_{\mathbf{r}_{n-1}} \parallel^2 \\
&= \mathbf{f}_{\mathbf{r}_{n-1}}^{*}\mathbf{M}_{\mathbf{r}_{n-1}}^{-1}\mathbf{V}_{\mathbf{r}_{n-1}}\mathbf{V}_{\mathbf{r}_{n-1}}^{*}\mathbf{M}_{\mathbf{r}_{n-1}}^{-1}\mathbf{f}_{\mathbf{r}_{n-1}} \\
&= \mathbf{f}_{\mathbf{r}_{n-1}}^{*}\mathbf{M}_{\mathbf{r}_{n-1}}^{-1}\mathbf{f}_{\mathbf{r}_{n-1}} \\
&= \mathcal{N}_m^{(\mathbf{r}_{n-1})} \\
&= \mathcal{N}_m^{(\mathbf{k})}.
\end{aligned}$$

The last equality is derived from Note 5 and Theorem 3. □

Similarly, we obtain a recursive method to determine $h_{\mathbf{k}}(x)$ and $q_{\mathbf{k},\mathbf{k}}(x)$ for $\mathbf{k} = (k_1, \ldots, k_s)$.

Theorem 8 *Let $\mathbf{k}$, n and $\mathbf{r}_i$ be defined before. For $h_{\mathbf{r}_1}(x) = f(x)$ and $q_{\mathbf{r}_0} = 1$, $h_{\mathbf{k}}(x)$ can be obtained recursively for i from 2 to n by the following recursive formula:*

$$h_{\mathbf{r}_i}(x) = h_{\mathbf{r}_{i-1}}(x) - \frac{p_{\mathbf{r}_{i-1}}}{q_{\mathbf{r}_{i-1}}q_{\mathbf{r}_{i-2}}} q_{\mathbf{r}_{i-1},\mathbf{r}_{i-1}}(x), \tag{48}$$

where

$$q_{\mathbf{r}_i,\mathbf{r}_i}(x) = q_{i,i}(x), \quad i = 2, \ldots, k_1. \tag{49}$$

If $i = \sum_{j=1}^{t} k_j + 1$ for some $t = 1, \ldots, s-1$, we have

$$\mathbf{r}_i = (k_1, k_2, \ldots, k_t, 1),$$

and then

$$q_{\mathbf{r}_i,\mathbf{r}_i}(x) = \mathbf{subs}(\overline{z}_s = x, q_{\mathbf{r}_i}). \tag{50}$$

Otherwise, there exist two integers d, t with $1 < d \le k_{t+1}$ and $1 \le t \le s-1$ such that

$$\mathbf{r}_{i-d} = (k_1, k_2, \ldots, k_t),$$

and

$$q_{\mathbf{r}_i,\mathbf{r}_i}(x) = \frac{1}{q_{\mathbf{r}_{i-2}}}\left(q_{\mathbf{r}_{i-1}}\frac{\partial q_{\mathbf{r}_{i-1},\mathbf{r}_{i-1}}(x)}{\partial \overline{z}_{t+1}} - \frac{\partial q_{\mathbf{r}_{i-1}}}{\partial \overline{z}_{t+1}} q_{\mathbf{r}_{i-1},\mathbf{r}_{i-1}}(x)\right). \tag{51}$$

Proof. We have

$$\begin{aligned}
h_{\mathbf{r}_i}(x) - h_{\mathbf{r}_{i-1}}(x) &= f(x) - \mathbf{q}_{\mathbf{r}_{i-1}}(x)\mathbf{M}_{\mathbf{r}_{i-1}}^{-1}\mathbf{f}_{\mathbf{r}_{i-1}} - (f(x) - \mathbf{q}_{\mathbf{r}_{i-2}}(x)\mathbf{M}_{\mathbf{r}_{i-2}}^{-1}\mathbf{f}_{\mathbf{r}_{i-2}}) \\
&= \mathbf{q}_{\mathbf{r}_{i-2}}(x)\mathbf{M}_{\mathbf{r}_{i-2}}^{-1}\mathbf{f}_{\mathbf{r}_{i-2}} - \mathbf{q}_{\mathbf{r}_{i-1}}(x)\mathbf{M}_{\mathbf{r}_{i-1}}^{-1}\mathbf{f}_{\mathbf{r}_{i-1}} \\
&= -\frac{q_{\mathbf{r}_{i-2}}}{q_{\mathbf{r}_{i-1}}}\frac{q_{\mathbf{r}_{i-1},\mathbf{r}_{i-1}}(x)}{q_{\mathbf{r}_{i-2}}}\frac{p_{\mathbf{r}_{i-1}}}{q_{\mathbf{r}_{i-2}}} \\
&= -\frac{p_{\mathbf{r}_{i-1}}}{q_{\mathbf{r}_{i-1}}q_{\mathbf{r}_{i-2}}} q_{\mathbf{r}_{i-1},\mathbf{r}_{i-1}}(x),
\end{aligned}$$

where $i = 2, \ldots, n$.

The proof of the recursive relation of $q_{\mathbf{r}_i,\mathbf{r}_i}(x)$ is similar to proofs of Theorem 2 and Theorem 4 for $s = 1$ case. $\square$

Note 6 *For any given multiplicity structure defined by* $\mathbf{k}$*, suppose the minimum of the* $\mathcal{N}_m^{(\mathbf{k})}$ *is attained at* $z_1, \ldots, z_s$*, then the nearest singular polynomial with the roots of multiplicity* $\mathbf{k}$ *can be obtained by substituting* $z_1, \ldots, z_s$ *into the* $h_{\mathbf{k}}(x)$ *computed from the above formula. This is true by Theorem 7.*

4 Examples

We are now ready to describe two examples of computing the nearest singular polynomials. All experiments are run with Digits=10 in Maple 13 under Windows XP.

Example 1

$$f = x^4 - 1.999\,x^3 + 0.997998\,x^2 + 0.001004\,x + 0.000398.$$

For $\mathbf{k} = (2)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^4 - 1.999100023\,x^3 + 0.9978980072\,x^2 + 0.9040371317e - 3\,x \\ &\quad + 0.2980670675e - 3; \\ z_1 &= 1.000299559; \quad \text{(double root)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.3998232663e - 7. \end{aligned}$$

For $\mathbf{k} = (3)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^4 - 1.929265099\,x^3 + 1.046163212\,x^2 + 0.9656143200e - 3\,x \\ &\quad - 0.9167052023e - 1; \\ z_1 &= 0.7237048697; \quad \text{(triple root)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.1565945795e - 1. \end{aligned}$$

For $\mathbf{k} = (2, 2)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^4 - 1.999000817\,x^3 + 0.9979971828\,x^2 + 0.1003382271e - 2\,x \\ &\quad + 0.2519456319e - 6; \\ z_1 &= -0.50194072190e - 3, \quad z_2 = 1.000002349; \quad \text{(two double roots)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.1582052317e - 6. \end{aligned}$$

Suppose the given tolerance $\epsilon = 10^{-3}$, then f has two 2-cluster of zeros. If $\epsilon = 10^{-7}$, then f only has a 2-cluster of zeros.

Example 2

$$f = x^5 - 3.000\,x^4 + 2.998997\,x^3 - 0.997991998\,x^2 - 0.1007004e-2\,x \\ + 0.402002e-3.$$

For $\mathbf{k} = (2)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^5 - 3.000079054\,x^4 + 2.998916651\,x^3 - 0.9980736638\,x^2 \\ &\quad - 0.1090008152e-2\,x + 0.3176375994e-3; \\ z_1 &= 0.9838765078; \quad \text{(double root)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.3338184094e-7. \end{aligned}$$

For $\mathbf{k} = (3)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^5 - 2.999919943\,x^4 + 2.998996997\,x^3 - 0.9980720296\,x^2 \\ &\quad - 0.1167031531e-2\,x + 0.1620105104e-3; \\ z_1 &= 1.000200211; \quad \text{(triple root)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.3524545527e-6. \end{aligned}$$

For $\mathbf{k} = (4)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^5 - 2.916630494\,x^4 + 2.984063932\,x^3 - 1.080981545\,x^2 \\ &\quad - 0.7456574864e-1\,x + 0.9061192588e-1; \\ z_1 &= 0.2978402953; \quad \text{(quadruple root)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.1525953438. \end{aligned}$$

For $\mathbf{k} = (2, 2)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^5 - 2.999999511\,x^4 + 2.998997489\,x^3 - 0.9979915087\,x^2 \\ &\quad - 0.1006214604e-2\,x - 0.2532432402e-6; \\ z_1 &= -0.502978703700000027e-3, \quad z_2 = 0.999683127399999982; \\ &\quad \text{(two double roots)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.1618668066e-6. \end{aligned}$$

For $\mathbf{k} = (3, 2)$,

$$\begin{aligned} h_{\mathbf{k}} &= x^5 - 3.000000991\,x^4 + 2.998997020\,x^3 - 0.9979909797\,x^2 \\ &\quad - .1004797476e-2\,x - 0.25253117e-6; \\ z_1 &= 1.00033517600000010, \quad z_2 = -0.502270004600000042e-3; \\ &\quad \text{(one double root and one triple root)} \\ \mathcal{N}_m^{(\mathbf{k})} &= 0.1591303726e-4. \end{aligned}$$

Suppose the given tolerance $\epsilon = 10^{-3}$, then f has a 3-cluster of zeros and a 2-cluster of zeros. If $\epsilon = 10^{-7}$, then f only has a 2-cluster of zeros.

Acknowledgement: This work is supported by the National Key Basic Research Project of China NKBRPC 2011CB302400, the Sino-French Lab for Computer Science, Automation and Applied Mathematics LIAMA through the ECCA project, and the Chinese National Natural Science Foundation under grant NSFC60821002/F02, 91118001, 10871194, 60911130369. The research was partially supported by the Austrian Science Fund (FWF): W1214-N15, project DK9. We thank referees for their careful reading and useful suggestions that help us improve the paper. The paper is also in memory of Wenda Wu (1929-2009) who proposed and studied the problem in 1997.

References

Corless, R. M., Rezvani, N., 2007. The nearest polynomial of lower degree. In: Proceedings of the 2007 international workshop on Symbolic-numeric computation. SNC '07. ACM, New York, NY, USA, pp. 199–200.

Hitz, M. A., Kaltofen, E., 1998. Efficient algorithms for computing the nearest polynomial with constrained roots. In: Proceedings of the 1998 international symposium on Symbolic and algebraic computation. ISSAC '98. ACM, New York, NY, USA, pp. 236–243.

Hitz, M. A., Kaltofen, E., Lakshman, Y. N., 1999. Efficient algorithms for computing the nearest polynomial with a real root and related problems. In: Proceedings of the 1999 international symposium on Symbolic and algebraic computation. ISSAC '99. ACM, New York, NY, USA, pp. 205–212.

Karmarkar, N., Lakshman, Y. N., 1996. Approximate polynomial greatest common divisors and nearest singular polynomials. In: Proceedings of the 1996 international symposium on Symbolic and algebraic computation. ISSAC '96. ACM, New York, NY, USA, pp. 35–39.

Pope, S. R., Szanto, A., June 2009. Nearest multivariate system with given root multiplicities. J. Symb. Comput. 44, 606–625.

Rezvani, N., Corless, R. M., September 2005. The nearest polynomial with a given zero, revisited. SIGSAM Bull. 39, 73–79.

Stetter, H. J., December 1999. The nearest polynomial with a given zero, and similar problems. SIGSAM Bull. 33, 2–4.

Stetter, H. J., 2004. Numerical Polynomial Algebra. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Zeng, Z., 2005. Computing multiple roots of inexact polynomials. Math. Comp 74, 869–903.

Zhi, L., Noda, M.-T., Kai, H., Wu, W., 2004. Hybrid method for computing the nearest singular polynomials. Japan Journal of Industrial and Applied Mathematics 21, 149–162.

Zhi, L., Wu, W., December 1998. Nearest singular polynomials. J. Symb. Comput. 26, 667–675.