



# Convergence of the Nelder-Mead method

Aurél Galántai<sup>1</sup>

Received: 12 December 2020 / Accepted: 18 October 2021 / Published online: 5 November 2021  
© The Author(s) 2021

## Abstract

We develop a matrix form of the Nelder-Mead simplex method and show that its convergence is related to the convergence of infinite matrix products. We then characterize the spectra of the involved matrices necessary for the study of convergence. Using these results, we discuss several examples of possible convergence or failure modes. Then, we prove a general convergence theorem for the simplex sequences generated by the method. The key assumption of the convergence theorem is proved in low-dimensional spaces up to 8 dimensions.

**Keywords** Nelder-Mead simplex method · Derivative-free minimization · Convergence

**Mathematics Subject Classification (2010)** MSC 65K10 · MSC 90C56

## 1 Introduction

We study the convergence of the Nelder-Mead simplex method [35] for the solution of the unconstrained minimization problem

$$f(x) \rightarrow \min \quad (f : \mathbb{R}^n \rightarrow \mathbb{R}),$$

where  $f$  is continuous. The Nelder-Mead method is widely used in derivative-free optimization and various application areas [2, 6, 21, 28, 42].

There are several forms and variants of the Nelder-Mead method. We use the version of Lagarias, Reeds, Wright and Wright [23]. The vertices of the initial simplex  $S$  are denoted by  $x_1, x_2, \dots, x_{n+1} \in \mathbb{R}^n$ . It is assumed that vertices  $x_1, \dots, x_{n+1}$  are ordered such that

$$f(x_1) \leq f(x_2) \leq \dots \leq f(x_{n+1}) \quad (1)$$

---

✉ Aurél Galántai  
galantai.aurel@uni-obuda.hu

<sup>1</sup> Óbuda University, Bécsi u. 96/b, 1034, Budapest, Hungary

and this condition is maintained during the iterations of the Nelder-Mead algorithm. Define the center  $x_c = \frac{1}{n} \sum_{i=1}^n x_i$  and  $x(\lambda) = (1 + \lambda)x_c - \lambda x_{n+1}$ . The related evaluation points are

$$x_r = x(1), \quad x_e = x(2), \quad x_{oc} = x\left(\frac{1}{2}\right), \quad x_{ic} = x\left(-\frac{1}{2}\right).$$

Then one (major) iteration of the method consists of the following operations or inner steps:

Operation	Nelder-Mead simplex method
0. Ordering	Order the vertices of $S$ such that $f(x_1) \leq \dots \leq f(x_{n+1})$ .
1. Reflect	If $f(x_1) \leq f(x_r) < f(x_n)$ , then replace $x_{n+1}$ by $x_r$ and goto 0.
2. Expand	If $f(x_r) < f(x_1)$ and $f(x_e) < f(x_r)$ , then replace $x_{n+1}$ by $x_e$ and goto 0. If $f(x_e) \geq f(x_r)$ , then replace $x_{n+1}$ by $x_r$ and goto 0.
3. Contract outside	If $f(x_n) \leq f(x_r) < f(x_{n+1})$ and $f(x_{oc}) \leq f(x_r)$ , then replace $x_{n+1}$ by $x_{oc}$ and goto 0.
4. Contract inside	If $f(x_r) \geq f(x_{n+1})$ and $f(x_{ic}) < f(x_{n+1})$ , then replace $x_{n+1}$ by $x_{ic}$ and goto 0.
5. Shrink	$x_i \leftarrow (x_i + x_1)/2, f(x_i)$ (for all $i$ ) and goto 0.

It is assumed that the above operations (or inner steps) are executed in the given order. Since the related logical conditions are mutually disjoint, any order of steps 1–5 results in the same output.

There are two rules that apply to reindexing after each iteration. If a nonshrink step occurs, then  $x_{n+1}$  is replaced by a new point  $v \in \{x_r, x_e, x_{oc}, x_{ic}\}$ . The following cases are possible:

$$f(v) < f(x_1), \quad f(x_1) \leq f(v) \leq f(x_n), \quad f(v) < f(x_{n+1}).$$

If

$$j = \begin{cases} 1, & \text{if } f(v) < f(x_1) \\ \max_{2 \leq \ell \leq n+1} \{f(x_{\ell-1}) \leq f(v) \leq f(x_\ell)\}, & \text{otherwise} \end{cases}.$$

then the new simplex vertices are

$$x_i^{new} = x_i \quad (1 \leq i \leq j - 1), \quad x_j^{new} = v, \quad x_i^{new} = x_{i-1} \quad (i = j + 1, \dots, n + 1). \quad (2)$$

This rule inserts  $v$  into the ordering with the highest possible index. If one of operations 1–4 is executed and the insertion rule (2) is used, the ordering operation can be skipped in the next iteration. If shrinking occurs, then

$$x'_1 = x_1, \quad x'_i = (x_i + x_1)/2 \quad (i = 2, \dots, n + 1)$$

and the ordering operation is necessary in the next iteration. By convention, if  $f(x'_1) \leq f(x'_i)$  ( $i = 2, \dots, n$ ), then  $x_1^{new} = x_1$ .

We adopt the following notations. The simplex of iteration  $k$  is denoted by  $S^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_{n+1}^{(k)}] \in \mathbb{R}^{n \times (n+1)}$  with vertices that satisfy the condition

$$f(x_1^{(k)}) \leq f(x_2^{(k)}) \leq \dots \leq f(x_{n+1}^{(k)}) \quad (k \geq 0). \tag{3}$$

The initial simplex is  $S^{(0)}$ . The center, reflection, expansion and contraction points of simplex  $S^{(k)}$  are denoted by  $x_c^{(k)}, x_r^{(k)}, x_e^{(k)}, x_{oc}^{(k)}$  and  $x_{ic}^{(k)}$ , respectively. The function values at the vertices  $x_j^{(k)}$  and the points  $x_c^{(k)}, x_r^{(k)}, x_e^{(k)}, x_{oc}^{(k)}$  and  $x_{ic}^{(k)}$  are denoted by  $f(x_j^{(k)}) = f_j^{(k)}$  ( $j = 1, \dots, n + 1$ ),  $f_c^{(k)}, f_r^{(k)} = f(x_r^{(k)})$ ,  $f_e^{(k)} = f(x_e^{(k)})$ ,  $f_{oc}^{(k)} = f(x_{oc}^{(k)})$  and  $f_{ic}^{(k)} = f(x_{ic}^{(k)})$ , respectively.

The insertion rule (2) guarantees that

$$f_i^{(k+1)} \leq f_i^{(k)} \quad (i = 1, 2, \dots, n + 1) \tag{4}$$

holds for any of operations 1–4. However this is not true in the case of shrinking. If function  $f$  is bounded from below on  $\mathbb{R}^n$  and only a finite number of shrink iterations occur, then each sequence  $\{f_i^{(k)}\}$  converges to some  $f_i^\infty$  for  $i = 1, \dots, n + 1$  (see Lemma 3.3 of [23]).

The original Nelder-Mead paper [35] was published in 1965 and since then it has been cited over 31000 times (see Google Scholar). However only a few results are known on the convergence.

In 1998 McKinnon [29] constructed a strictly convex function  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  with continuous derivatives for which the Nelder-Mead simplex algorithm converges to a nonstationary point.

Also in 1998 Lagarias, Reeds, Wright and Wright [23] proved convergence for one and two variable strictly convex functions. For  $n = 2$ , they summarized the main results as follows (see p. 114 of [23]):

- *The function values at all simplex vertices in the standard Nelder-Mead algorithm converge to the same value.*
- *The simplices in the standard Nelder-Mead algorithm have diameters converging to zero.*

In 1999 Kelley [19, 20] gave a sufficient decrease condition for the average of the object function values (evaluated at the vertices) and proved that if this condition is satisfied during the process, then any accumulation point of the simplices is a critical point of  $f$ .

In 2006 Han and Neumann [12] investigated the convergence and the effect of dimensionality on the Nelder–Mead method when it is applied to  $f(x) = x^T x$  ( $x \in \mathbb{R}^n$ ). They also showed that the Nelder-Mead method deteriorates as  $n$  increases.

In 2012 Lagarias, Poonen, Wright [22] significantly improved the results of the earlier paper [23] for the restricted Nelder-Mead method, where expansion steps are not allowed. Let  $\mathcal{F}$  be the class of twice-continuously differentiable functions  $\mathbb{R}^2 \rightarrow \mathbb{R}$  with bounded level sets and everywhere positive definite Hessian. They proved that for any  $f \in \mathcal{F}$  and any nondegenerate initial simplex  $S^{(0)}$ , the restricted Nelder-Mead algorithm converges to the unique minimizer of  $f$ .

Wright [43, 44] raised several questions concerning the Nelder-Mead method such as the following:

- (a) Do the function values at all vertices necessarily converge to the same value?
- (b) Do all vertices of the simplices converge to the same point?
- (c) Why is it sometimes so effective (compared to other direct search methods) in obtaining a rapid improvement in  $f$ ?
- (d) One failure mode is known (McKinnon [29]) — but are there other failure modes?
- (e) Why, despite its apparent simplicity, should the Nelder-Mead method be difficult to analyze mathematically?

Although questions (a) and (b) were positively answered for one- and two-dimensional strictly convex functions by Lagarias et al. [22, 23] no general answer is known as yet.

Our purpose is to analyze and prove the convergence of the simplex sequence generated by the method. The matrix formulation of the Nelder-Mead simplex algorithm, which is introduced in Section 2 represents the  $k$ th simplex as a product of transformation matrices and so the convergence of the simplex sequence is related to the convergence of infinite matrix products. Hence the spectra of occurring matrices which is necessary for the convergence of the simplex sequence is investigated in Section 3. Section 4 discusses several examples of possible convergence behavior or failure, and specifies the type of convergence we prove later. The main convergence theorem is proved in Section 5 under Assumption (A) of Section 5. The assumption is proved for  $1 \leq n \leq 8$  in Section 6. Related numerical data is given in the Appendix.

The results and examples may answer some of the questions raised by Wright [43, 44]. Actually Examples 4 and 5 answer questions (a) and (b) negatively. Examples of Section 4 answer question (d) positively. The main convergence theorem is related to questions (b) and (a). The connection of the Nelder-Mead method with infinite products of matrices may shed some light on question (e).

This paper is an improvement of paper [10] where the main convergence result was proved under Assumption (A) and a second assumption on the spectra of transformation matrices. The two assumptions were numerically checked only for  $n = 1, 2, 3$ . Here we prove the convergence without the second assumption which follows from Theorem 3 of Section 3. The removal of Assumption (A) is unlikely for its connection to an undecidable problem. This will be discussed at the end of Section 5.

## 2 A matrix form of the Nelder-Mead method and consequences

Assume that simplex  $S^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_{n+1}^{(k)}]$  is such that condition (3) holds. If the incoming vertex  $v$  is of the form

$$v = \frac{1 + \alpha}{n} \sum_{i=1}^n x_i^{(k)} - \alpha x_{n+1}^{(k)}$$

for some  $\alpha \in \left\{1, 2, \frac{1}{2}, -\frac{1}{2}\right\}$ , we can define the transformation matrix

$$T(\alpha) = \begin{bmatrix} I_n & \frac{1+\alpha}{n} e \\ 0 & -\alpha \end{bmatrix} \quad (e = [1, 1, \dots, 1]^T).$$

Since  $S^{(k)}T(\alpha) = \left[x_1^{(k)}, \dots, x_n^{(k)}, x(\alpha)\right]$ , we have to reorder the matrix columns according to the insertion rule (2). Define the permutation matrix

$$P_j = [e_1, \dots, e_{j-1}, e_{n+1}, e_j, \dots, e_n] \in \mathbb{R}^{(n+1) \times (n+1)} \quad (j = 1, \dots, n + 1).$$

Then  $S^{(k)}T(\alpha)P_j$  is the new simplex  $S^{(k+1)}$ . The following cases are possible

Operation	New simplex
1. Reflection	$S^{(k+1)} = S^{(k)}T(1)P_j \quad (j = 2, \dots, n)$
2a) Expansion( $v = x_e^{(k)}$ )	$S^{(k+1)} = S^{(k)}T(2)P_1$
2b) Expansion( $v = x_r^{(k)}$ )	$S^{(k+1)} = S^{(k)}T(1)P_1$
3. Outside contraction	$S^{(k+1)} = S^{(k)}T\left(\frac{1}{2}\right)P_j \quad (j = 1, \dots, n + 1)$
4. Inside contraction	$S^{(k+1)} = S^{(k)}T\left(-\frac{1}{2}\right)P_j \quad (j = 1, \dots, n + 1)$

Denote by  $\mathcal{P}_{n+1}$  the set of all possible permutation matrices of order  $n + 1$ . In the case of shrinking the new simplex is

$$S^{(k+1)} = S^{(k)}T_{shr}P \quad (P \in \mathcal{P}_{n+1}),$$

where

$$T_{shr} = \frac{1}{2}I_{n+1} + \frac{1}{2}e_1e^T$$

and the permutation matrix  $P \in \mathcal{P}_{n+1}$  is defined by the ordering condition (3).

Hence for  $k \geq 1$ , the  $k$ th simplex of the Nelder-Mead method is

$$S^{(k)} = S^{(k-1)}T_kP^{(k)} = S^{(0)}B_k, \tag{5}$$

where

$$B_k = \prod_{i=1}^k T_iP^{(i)} \quad (T_iP^{(i)} \in \mathcal{T}) \tag{6}$$

and

$$\begin{aligned} \mathcal{T} = & \left\{ T(\alpha)P_j : \alpha \in \left\{ -\frac{1}{2}, \frac{1}{2} \right\}, j = 1, \dots, n + 1 \right\} \\ & \cup \{T_{shr}P : P \in \mathcal{P}_{n+1}\} \cup \{T(1)P_j : j = 1, \dots, n\} \cup \{T(2)P_1\} \end{aligned} \tag{7}$$

Note that  $\mathcal{T}$  contains  $3n + 3 + (n + 1)!$  matrices, all transformation matrices  $TP \in \mathcal{T}$  are nonsingular and their column sums are 1.

Such matrices have the following properties.

**Lemma 1** (i) If  $A \in \mathbb{R}^{n \times n}$  is a matrix whose column sums are 1, then  $A$  has an eigenvalue  $\lambda = 1$  and a corresponding left eigenvector  $x = e^T$ . (ii) If  $A, B \in \mathbb{R}^{n \times n}$

are two matrices whose column sums are 1, then  $C = AB$  also has this property. (iii) If  $A \in \mathbb{R}^{n \times n}$  is a matrix whose column sums are 1, then  $\|A\| \geq 1$  in any induced matrix norm.

*Proof* (i) By definition  $e^T A = [\sum_{i=1}^n a_{i1}, \dots, \sum_{i=1}^n a_{in}] = 1 \cdot e^T$ . (ii)  $e^T A = A$  and  $e^T B = e^T$  imply  $e^T AB = e^T B = e^T$ . (iii)  $\rho(A) \geq 1$  and  $\|A\| \geq \rho(A)$  imply  $\|A\| \geq 1$ . □

A matrix  $A$  is called left stochastic, if  $a_{ij} \geq 0$  for all  $i, j$  and the column sums are 1. A matrix is called stochastic, if  $a_{ij} \geq 0$  for all  $i, j$  and both the column sums and the row sums are 1. All matrices  $T_{shr} P$  ( $P \in \mathcal{P}_{n+1}$ ) and  $T(\alpha)$  ( $-1 \leq \alpha \leq 0$ ) are left stochastic matrices.

A simplex  $S = [x_1, \dots, x_{n+1}]$  is said to be nondegenerate if the matrix

$$M = [x_1 - x_{n+1}, x_2 - x_{n+1}, \dots, x_n - x_{n+1}]$$

is nonsingular. Then  $S$  must be affinely independent, which is equivalent to (see, e.g., [15, 31]) that vectors  $\begin{bmatrix} 1 \\ x_1 \end{bmatrix}, \dots, \begin{bmatrix} 1 \\ x_{n+1} \end{bmatrix}$  are linearly independent. Hence  $\text{rank}\left(\begin{bmatrix} e^T \\ S \end{bmatrix}\right) = n + 1$ .<sup>1</sup> We always assume that the initial simplex  $S^{(0)}$  is nondegenerate. Since  $e^T B_k = e^T$  and  $\begin{bmatrix} e^T \\ S^{(k)} \end{bmatrix} = \begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix} B_k$  is nonsingular,  $S^{(k)}$  is also nondegenerate.

Although  $S^{(k)} \in \mathbb{R}^{n \times (n+1)}$  we can relate the convergence of  $\{S^{(k)}\}$  and  $\{B_k\}$  as follows.

**Lemma 2** *If  $\{B^{(k)}\}$  is bounded, then  $\{S^{(k)}\}$  converge to some  $S^\infty$  if and only if  $\{B_k\}$  converge to some  $B$ .*

*Proof* If  $S^{(k)} \rightarrow S^\infty$  (whatever  $S^\infty$  is) and  $B_k \rightarrow B$ , then  $S^{(0)} B_k \rightarrow S^{(0)} B = S^\infty$ . Assume that  $S^{(k)} \rightarrow S^\infty$  and  $\{B_k\}$  is not convergent. Since  $\{B_k\}$  is bounded it must have at least one accumulation point, say  $B^*$  and a subsequence  $\{B_{i_j}\} \subset \{B_k\}$  such that  $B_{i_j} \rightarrow B^*$  and  $S^{(i_j)} \rightarrow S^{(0)} B^* = S^\infty$ . Assume that there exists a second accumulation point  $B^{**} \neq B^*$  and a subsequence  $\{B_{k_j}\} \subset \{B_k\}$  such that  $B_{k_j} \rightarrow B^{**}$ . It follows that

$$\begin{bmatrix} e^T \\ S^{(i_j)} \end{bmatrix} \rightarrow \begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix} B^* = \begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix} B^{**} \leftarrow \begin{bmatrix} e^T \\ S^{(k_j)} \end{bmatrix}.$$

Since  $\begin{bmatrix} e^T \\ S^{(0)} \end{bmatrix}$  is nonsingular, we obtain that  $B^* = B^{**}$ , which is contradiction. It follows that  $\{B_k\}$  converges. □

---

<sup>1</sup>It is assumed through the whole paper that the sizes of  $e$  and the unit vectors  $e_i$  are compatible with the operation and/or partition.

If  $\{B_k\}$  is not bounded, then as Examples 1 and 2 of Section 4 show that we can have a convergence of the function values  $f_i^{(k)}$  to limit values that are not related to any extrema of the function  $f$ .

Hence we study the convergence of  $\{S^{(k)}\}$  through the convergence of  $\{B_k\}$  or the convergence of the right infinite product  $B = \prod_{i=1}^{\infty} T_i P^{(i)}$  ( $T_i P^{(i)} \in \mathcal{T}$ ).

We use the following results and definitions from the theory of infinite matrix products (see, e.g., Hartfiel [13]). A right infinite product is an expression  $A_1 A_2 \cdots A_k A_{k+1} \cdots$ . A set  $\Sigma$  of  $n \times n$  matrices has the right convergence property (RCP), if all possible right infinite products  $\prod_{i=1}^{\infty} A_i$  ( $A_i \in \Sigma$ ) converge.

It is easy to show (see, e.g., Hartfiel [13] p. 103) that if  $\Sigma$  is an RCP set,  $A_1, \dots, A_k \in \Sigma$  and  $\lambda$  is an eigenvalue of  $A_1 A_2 \cdots A_k$ , then  $|\lambda| < 1$  or  $\lambda = 1$  and this eigenvalue is simple. Hence each matrix of  $\Sigma$  must satisfy this condition.

If  $\Sigma$  is an RCP set, then there is a vector norm  $\|\cdot\|$  such that  $\|A\| \leq 1$  for all  $A \in \Sigma$  (see, e.g., [13]).

In any induced matrix norm  $\|T_i P^{(i)}\| \geq 1$  holds for every  $T_i P^{(i)} \in \mathcal{T}$  and  $\|T(2)\| > 1$ . Hence  $\mathcal{T}$  as a whole is not an RCP set. However the eigenvalue-eigenvector structure of the transformation matrices makes it possible to identify a subset of  $\mathcal{T}$ , which might have the RCP property. The next section investigates the structure of the matrices  $T_i P^{(i)} \in \mathcal{T}$ . Using these results we show several examples of possible convergence behavior or failure in Section 4. We also specify here the type of convergence we study in the rest of the paper. In Section 5 we identify a subset of  $\mathcal{T}$ , which might have the RCP property. Using this subset we give a sufficient condition under which the simplex sequences  $\{S^{(k)}\}$  converge in the specified sense.

There are many problems that may complicate the analysis of the Nelder-Mead method. Here we mention the following. The Nelder-Mead algorithm is a nonstationary iteration (see, e.g., Young [45]) of the form

$$S^{(k)} = S^{(k-1)} T_k P^{(k)} \quad (T_k P^{(k)} \in \mathcal{T}, k = 0, 1, \dots),$$

where the matrices  $T_k P^{(k)}$  are not contractive ( $\|T_k P^{(k)}\| \geq 1$ ). Operations 1–4 and the insertion rule (2) guarantee only improvement in the worst vertex ( $f_{n+1}^{(k+1)} \leq f_{n+1}^{(k)}$ ). In the case of shrinking there is no guaranteed improvement at all. It is also notable that the selection of  $T_k P^{(k)}$  at iteration  $k$  depends only on the positions of  $f_r^{(k)}$ ,  $f_e^{(k)}$ ,  $f_{oc}^{(k)}$  and  $f_{ic}^{(k)}$  relative to the ordering (3) of  $f_i^{(k)}$ 's. Example 3 of Section 4 shows that given an initial simplex  $S^{(0)}$ , the Nelder-Mead algorithm may generate the same sequence  $\{S^{(k)}\}$  for different functions such that  $\lim_{k \rightarrow \infty} S^{(k)}$  has different meanings for the different functions.

Finally, we note that matrix  $T(\alpha)$  appears in Lagarias et al. [23] (p. 149) but it is not exploited subsequently.

### 3 The spectra of the transformation matrices

In this section we study the spectra of the transformation matrices  $T(\alpha) P_j$  and  $T_{shr} P$  so that we could find a subset of  $\mathcal{T}$ , which has the RCP property. We then study the asymptotic behavior of the eigenvalues for  $n \rightarrow \infty$  which is important for

the dimensionality effects. Han, Neumann and Xu [11] and Han and Neumann [12] obtained similar results. The similarities and connections will be discussed at the end of the section.

We use the following result from Rahman and Schmeisser [38], which is one of the several Schur-Cohn type theorems or criteria (for others, see Marden [30], Henrici [14], Sheil-Small [39], Barnett [3]).

**Definition 1** Given a polynomial  $p(z) = a_0 + a_1z + \dots + a_nz^n$  of degree  $n$ , we associate with it the polynomial

$$p^*(z) := z^n \overline{p(1/\bar{z})} = \bar{a}_n + \bar{a}_{n-1}z + \dots + \bar{a}_0z^n.$$

If  $p(z) \equiv \sigma p^*(z)$ , where  $|\sigma| = 1$ , then  $p$  is said to be self-inversive.

**Theorem 1** (Cohn’s rules [38], Thm. 11.5.3) *Let  $p(z) = \sum_{i=0}^n a_i z^i$  be a polynomial of degree  $n$ . Denote by  $r$  and  $s$  the number of zeros of  $p$  inside the unit circle and on it, respectively. For polynomial  $p_j$  the corresponding numbers are respectively denoted by  $r_j$  and  $s_j$ . Then the following rules hold.*

- Rule 1 *If  $|a_0| > |a_n|$ , then  $p_1(z) := \bar{a}_0 p(z) - a_n p^*(z)$  is not identically zero, and we have  $\deg p_1 < \deg p$ ,  $r_1 = r$ , and  $s_1 = s$ .*
- Rule 2 *If  $|a_0| < |a_n|$ , then  $p_1(z) := [\bar{a}_n p(z) - a_0 p^*(z)]/z$  is of degree  $n - 1$ ,  $r_1 = r - 1$ , and  $s_1 = s$ .*
- Rule 3 *If there is an index  $k \leq n/2$  such that*

$$a_0 = \sigma \bar{a}_n, a_1 = \sigma \bar{a}_{n-1}, \dots, a_{k-1} = \sigma \bar{a}_{n-k+1}, \quad a_k \neq \sigma \bar{a}_{n-k} \quad (|\sigma| = 1),$$

then define  $b := (a_{n-k} - \sigma \bar{a}_k) / a_n$ ,

$$g(z) := \left( z^k + \frac{2b}{|b|} \right) p(z), \quad g_1(z) := \bar{g}(0) g(z) - \bar{g}^*(0) g^*(z),$$

and

$$p_1(z) := \frac{1}{z} [g_1^*(0) g_1(z) - g_1(0) g_1^*(z)],$$

which yields that  $p_1$  is of degree  $n - 1$ ,  $r_1 = r - 1$ , and  $s_1 = s$ .

- Rule 4 *If  $p$  is self-inversive, then  $p_1(z) := np(z) - zp'(z)$  is not identically zero, and we have  $\deg p_1 < \deg p$  and  $r_1 = r$ ; furthermore,  $s = n - 2r$ .*

We also use the classical Eneström-Kakeya theorem, which can be found in several sources (see, e.g., Rahman, Schmeisser [38], Corollary 8.3.5, Sheil-Small [39], Marden [30]).

**Theorem 2** (Eneström-Kakeya [8, 18]) *A polynomial  $p(z) = \sum_{i=0}^n a_i z^i$  with positive coefficients has all its zeros in the annulus*

$$R_1 \leq |z| \leq R_2,$$

where  $R_1 = \min_{0 \leq i \leq n-1} a_i/a_{i+1}$  and  $R_2 = \max_{0 \leq i \leq n-1} a_i/a_{i+1}$ .



Matrices  $T(\alpha)$  and  $T_{shr}$  have a simple eigenvalue-eigenvector structure. When they are multiplied by a permutation matrix other than  $I_{n+1}$ , this may change. For example, if  $n = 2$ ,  $T(1)$  is an involution, while  $T(1)P_2$  (reflection) is a 6-involutory matrix (for such matrices, see Trench [41]).

**Theorem 3** *The characteristic polynomial of matrix  $T(\alpha)P_j$  ( $1 \leq j \leq n + 1$ ) is*

$$p_{n+1}(\lambda) = (1 - \lambda)^{j-1} p_{n+2-j}(\lambda), \tag{8}$$

where

$$p_{n+2-j}(\lambda) = \lambda^{n+2-j} - c \sum_{i=1}^{n+1-j} \lambda^i + \alpha \quad \left( c = \frac{1 + \alpha}{n} \right). \tag{9}$$

- (i) *If  $j = 1$ , then  $p_{n+1}(\lambda)$  has at least one eigenvalue  $\lambda = 1$ .*
- (ii) *If  $\alpha = 1$ ,  $p_{n+1}(\lambda)$  has at least two eigenvalues  $\lambda = 1$ .*
- (iii) *If  $\alpha > -1$  and  $j \geq 2$ , there are exactly  $j - 1$  eigenvalues  $\lambda = 1$ .*
- (iv) *For  $|\alpha| < 1$  and  $j = 1$ ,  $p_{n+1}(\lambda)$  has exactly one zero  $\lambda_1 = 1$ , while the remaining  $n$  roots are in the open unit disk.*
- (v) *If  $|\alpha| < 1$  and  $2 \leq j \leq n + 1$ , then all zeros of  $p_{n+2-j}(\lambda)$  are in the open unit disk.*
- (vi) *If  $\alpha = 1$  and  $1 \leq j \leq n + 1$ , then all zeros of  $p_{n+2-j}(\lambda)$  are on the unit circle.*
- (vii) *If  $\alpha > 1$  and  $j = 1$ , then  $p_{n+1}(\lambda)$  has a second zero in the interval  $(1, 1 + \frac{1+\alpha}{n})$  and all its zeros are in the annulus  $1 \leq |\lambda| \leq 1 + \alpha$ .*

*Proof* For  $1 \leq j \leq n + 1$ ,

$$T(\alpha)P_j = \begin{bmatrix} I_{j-1} & cee_1^T \\ 0 & A_{n+2-j} \end{bmatrix}, \tag{10}$$

where

$$A_{n+2-j} = \begin{bmatrix} c & 1 & 0 & \cdots & 0 \\ c & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ c & 0 & \cdots & 0 & 1 \\ -\alpha & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Since  $A_{n+2-j}$  is a companion matrix (for this form, see, e.g., [16]), its characteristic polynomial is

$$p_{n+2-j}(\lambda) = \lambda^{n+2-j} - c \sum_{i=1}^{n+1-j} \lambda^i + \alpha \quad (1 \leq j \leq n + 1)$$

and the characteristic polynomial of  $T(\alpha)P_j$  is

$$\det(T(\alpha)P_j - \lambda I_{n+1}) = (1 - \lambda)^{j-1} p_{n+2-j}(\lambda).$$

Note that  $p_{n+2-j}(1) = \frac{j-1}{n}(\alpha + 1)$ . If  $j = 1$ ,  $p_{n+1}(1) = 0$ , that is  $\lambda_1 = 1$ , which proves (i). Since  $p'_{n+1}(\lambda) = \frac{1-\alpha}{2}(n + 1)$ , there is a second zero  $\lambda_2 = 1$  if  $\alpha = 1$

(Claim (ii)). If  $\alpha > -1$  and  $j \geq 2$ ,  $p_{n+2-j}(1) = (j - 1)c > 0$  proving (iii). For proving (iv) and (v) we shall consider real polynomials of the form

$$p(\lambda) = \alpha - c \sum_{i=1}^{m-1} \lambda^i + \lambda^m \quad \left( c = \frac{1 + \alpha}{n} \right),$$

where  $2 \leq m \leq n + 1$ . Since  $p^*(\lambda) = 1 - c \sum_{i=1}^{m-1} \lambda^i + \alpha\lambda^m$  and  $|\alpha| < 1$ , we can apply Rule 2 of Theorem 1:

$$p_1(\lambda) = [1p(\lambda) - \alpha p^*(\lambda)]/z = (1 - \alpha^2)\lambda^{m-1} + c(\alpha - 1) \sum_{i=0}^{m-2} \lambda^i.$$

Dividing  $p_1(\lambda)$  by  $c(1 - \alpha) > 0$ , one obtains

$$p_1(\lambda) = n\lambda^{m-1} - \sum_{i=0}^{m-2} \lambda^i = n\lambda^{m-1} - \sum_{i=1}^{m-2} \lambda^i - 1.$$

For  $|\lambda| \geq 1$ ,

$$\begin{aligned} |p_1(\lambda)| &= |\lambda|^{m-1} \left| n - \sum_{i=1}^{m-1} \frac{1}{\lambda^i} \right| \geq |\lambda|^{m-1} \left( n - \sum_{i=1}^{m-1} \frac{1}{|\lambda|^i} \right) \\ &\geq |\lambda|^{m-1} (n - m + 1). \end{aligned}$$

If  $m \leq n$  ( $j \geq 2$ ), then  $p_1(\lambda)$  has no root outside the open unit disk,  $r_1 = m - 1$ ,  $s_1 = 0$ . Hence  $p(\lambda) = 0$  has  $m$  zeros inside the open unit disk. If  $m = n + 1$ , then  $p_1(\lambda) = n\lambda^n - \sum_{i=1}^{n-1} \lambda^i - 1$ . For  $|\lambda| > 1$ ,  $|p_1(\lambda)| > 0$ , and  $p_1(1) = 0$ . We must prove that there is no other root on the unit circle, that is  $s_1 = 1$ . If  $n = 2$ , then  $p_1(\lambda) = 2\lambda^2 - \lambda - 1 = 0$  with the solution  $\lambda_1 = 1$  and  $\lambda_2 = -\frac{1}{2}$ . Hence  $r_1 = 1$  and  $s_1 = 1$ . It follows that  $r = 2$  and  $s = s_1 = 1$ . It is easy to check that

$$p_1(\lambda) = n\lambda^n - \sum_{i=1}^{n-1} \lambda^i - 1 = (\lambda - 1) \left( \sum_{i=0}^{n-1} (i + 1)\lambda^i \right) = (\lambda - 1)q(\lambda).$$

Polynomial  $q(\lambda)$  has positive real coefficients  $a_i = i + 1$  ( $i = 0, 1, \dots, n - 1$ ),

$$R_1 = \frac{1}{2} \leq \frac{a_i}{a_{i+1}} = \frac{i + 1}{i + 2} \leq \frac{n - 1}{n} = R_2 < 1,$$

and the Eneström-Kakeya theorem implies that  $q(\lambda)$  has all its zeros in the annulus  $\frac{1}{2} \leq |\lambda| \leq \frac{n-1}{n} < 1$ . Hence  $r_1 = n - 1$ ,  $s_1 = 1$ , and so  $r = n$  and  $s = 1$ .

Assume now that  $\alpha = 1$  and  $1 \leq j \leq n$ . Using the notation  $\ell = n + 2 - j$  we have

$$p(\lambda) = 1 - \frac{2}{n} \sum_{i=1}^{\ell-1} \lambda^i + \lambda^\ell \quad \left( c = \frac{2}{n} \right),$$

which is self-inversive (see Definition 1). Thus we apply Rule 4 of Cohn (see Theorem 1):

$$p_1(\lambda) = \ell p(\lambda) - \lambda p'(\lambda) = \ell - \frac{2}{n} \sum_{i=1}^{\ell-1} (\ell - i)\lambda^\ell.$$

If  $|\lambda| < 1$ , then

$$|p_1(\lambda)| \geq \ell - \frac{2}{n} \sum_{i=1}^{\ell-1} (\ell - i) |\lambda|^i > \ell - \frac{2}{n} \sum_{i=1}^{\ell-1} (\ell - i) = \frac{1}{n} \ell (n - \ell + 1).$$

Hence  $p_1(\lambda)$  has no zero inside the unit disk,  $r_1 = 0 = r$  and the zeros of  $p(\lambda)$  on the unit circle are  $s = \ell$ . Claim (vi) also follows from a result of Lakatos and Losonczi [24].

Since for  $j = 1$  and any  $\alpha$ ,  $p_{n+1}(\lambda) = (\lambda - 1) \sum_{i=0}^n (-\alpha + ic) \lambda^i$ , we investigate  $q(\lambda) = \sum_{i=0}^n (-\alpha + ic) \lambda^i$  with  $q(0) = -\alpha$ . By definition  $q^*(\lambda) = \sum_{i=0}^n (-\alpha + (n - i)c) \lambda^i$  and  $q^*(0) = -\alpha + nc = 1$ . Since  $\alpha > 1$ , we can apply Rule 1 of Theorem 1:

$$q_1(\lambda) = \sum_{i=0}^{n-1} (\alpha^2 - 1) \left(1 - \frac{i}{n}\right) \lambda^i.$$

The coefficients of  $q_1(\lambda)$  are positive and

$$1 \leq \frac{a_i}{a_{i+1}} = \frac{n - i}{n - i - 1} \leq 2.$$

The Eneström-Kakeya theorem implies that the zeros of  $q_1(\lambda)$  are in the annulus  $1 \leq |\lambda| \leq 2$ . Hence  $q_1(\lambda)$  has no zero inside the unit disk,  $r_1 = 0 = r, q$  and  $q_1$  have the same number of zeros inside the unit disk. Cauchy’s classical estimate yields the annulus  $1 \leq |\lambda| \leq 1 + \alpha$  for the zeros of  $p_{n+1}(\lambda)$ , if  $\alpha > \frac{1+\alpha}{n} = c$ , which is satisfied for  $n \geq 2$ . Note that  $q(1) = \frac{1-\alpha}{2} (n + 1) < 0$  for  $\alpha > 1$ . Assume that  $\lambda \neq 1$ . Then

$$q(\lambda) = \frac{\lambda \left(c + \alpha - \frac{\alpha}{\lambda}\right) + \lambda^{n+1} (\lambda - 1 - c)}{(\lambda - 1)^2}$$

and

$$q(1 + c) = \frac{c + \alpha + 1}{c} > 0.$$

Hence there must be a real zero in the interval  $(1, 1 + c)$ . □

For the eigenvalues of  $T_{shr} P$  ( $P \in \mathcal{P}_{n+1}$ ), we need the following result.

**Theorem 4** (Langville and Meyer [26, 27]). *If the spectrum of the stochastic matrix  $P$  is  $\{1, \lambda_2, \dots, \lambda_n\}$ , then the spectrum of*

$$W = \alpha P + (1 - \alpha) e v^T$$

*is  $\{1, \alpha \lambda_2, \alpha \lambda_3, \dots, \alpha \lambda_n\}$ , where  $v^T$  is a probability vector.<sup>2</sup>*

Since the eigenvalues of  $W$  and  $W^T$  coincide, the same result holds for the transposed matrix

$$W^T = \alpha P^T + (1 - \alpha) v e^T$$

as well.

---

<sup>2</sup>Vector  $v$  is such that  $v_i \geq 0$  and  $e^T v = 1$ .

**Corollary 1** *The spectrum of*

$$T_{shr}P = \frac{1}{2}P + \frac{1}{2}e_1e^T \quad (P \in \mathcal{P}_{n+1})$$

is  $\left\{1, \frac{1}{2}\lambda_2, \frac{1}{2}\lambda_3, \dots, \frac{1}{2}\lambda_{n+1}\right\}$ . Since the eigenvalues of a permutation matrix  $P$  are on the unit circle  $|\lambda| = 1$ , we have  $|\lambda_i(T_{shr}P)| = \frac{1}{2}$  for  $i = 2, \dots, n + 1$ .

*Remark 1* The eigenvalues  $\lambda = 1$  of  $T(\alpha)P_j$  ( $|\alpha| < 1$ ) and  $T_{shr}P$  ( $P \in \mathcal{P}_{n+1}$ ) are simple and so  $\left\{[T(\alpha)P_j]^k\right\}$  and  $\left\{[T_{shr}P]^k\right\}$  are convergent. For  $\alpha > 1$ ,  $\rho(T(\alpha)P_1) > 1$ , and so  $\left\{[T(\alpha)P_1]^k\right\}$  is unbounded. For  $\alpha = 1$ ,  $T(1)P_1$  has at least a double eigenvalue  $\lambda = 1$ , and since it is a companion matrix, its Jordan form has at least a  $2 \times 2$  block belonging to  $\lambda = 1$ . Hence  $\left\{[T(1)P_1]^k\right\}$  is also unbounded (see, e.g., [16, 32, 36]).

For the asymptotic behavior of the eigenvalues of  $T(\alpha)P_j$ , note that for  $j = n + 1 - \ell$  with fixed  $\ell$  and  $n \rightarrow \infty$ , the matrix  $T(\alpha)P_j$  has  $n - \ell$  eigenvalues  $\lambda = 1$ , while the remaining  $\ell + 1$  eigenvalues (the zeros of  $p_{\ell+1}(\lambda) = \lambda^{\ell+1} - c \sum_{i=1}^{\ell} \lambda^i + \alpha$ ) are converging to the zeros of the polynomial  $p(\lambda) = \lambda^{\ell+1} + \alpha$ .

For a more precise result, we use a less known result of E. Landau [25] (see also [1, 30] or [33]).

**Theorem 5** (E. Landau [25]). *Consider the polynomial  $p(z) = a_0 + a_1z + \dots + a_nz^n$  ( $a_0a_n \neq 0$ ). If  $z$  is a zero of  $p(z)$  and  $t$  is any positive real number, then*

$$|z| \geq g(t) := \frac{|a_0|t}{|a_0| + \max_{1 \leq i \leq n} |a_i|t^i}. \tag{11}$$

Landau’ theorem yields the following characterizations of the eigenvalues of  $T(\alpha)P_j$  when  $n \rightarrow \infty$ .

**Lemma 3** *Assume that  $|\alpha| < 1$  and consider the eigenvalues of  $T(\alpha)P_j$  for fixed  $j$  and  $n \rightarrow \infty$ . Then  $T(\alpha)P_j$  has either one eigenvalue  $\lambda = 1$  ( $j = 1$ ) or  $j - 1$  eigenvalues  $\lambda = 1$  ( $j \geq 2$ ), while the absolute values of the remaining  $n + 2 - j$  eigenvalues (the zeros of  $p_{n+2-j}(\lambda)$ ) are converging to 1.*

*Proof* In our case  $p(\lambda) = \lambda^\ell - c \sum_{i=1}^{\ell-1} \lambda^i + \alpha$  and

$$\max_{1 \leq i \leq n} |a_i|t^i = \max_{1 \leq i \leq \ell-1} \left\{ct^i, t^\ell\right\}.$$

Assume that  $n \geq 2$ . Then  $0 < c < 1$ . For  $0 < t < 1$ ,  $ct > ct^i$  ( $i \geq 2$ ) and  $ct = t^\ell$  for  $t = \ell^{-1}\sqrt[\ell]{c} < 1$ . If  $t_1 = \ell^{-1}\sqrt[\ell]{c}$  and  $\ell = n + 1 - j$  with  $j$  fixed, then for the quantity  $g(t)$  given by Theorem 5,

$$g(t_1) = \frac{|\alpha|^{n+1-j} \sqrt{\frac{1+\alpha}{n}}}{|\alpha| + \frac{1+\alpha}{n} \sqrt{\frac{1+\alpha}{n}}} \rightarrow 1 \quad (n \rightarrow \infty)$$

holds. □

**Lemma 4** *Assume that  $\alpha > 1$  and consider the eigenvalues of  $T(\alpha)P_1$  for  $n \rightarrow \infty$ . Then the absolute values of the eigenvalues of  $T(\alpha)P_1$  (the zeros of  $p_{n+1}(\lambda)$ ) are converging to 1.*

*Proof* Since for  $\alpha > 1$ , the zeros of  $p_{n+1}(\lambda) = \lambda^{n+1} - c \sum_{i=1}^n \lambda^i + \alpha$  are in the annulus  $1 \leq |z| \leq 1 + \alpha$ , the zeros of the reciprocal equation  $p_{n+1}^*(\lambda) = 1 - c \sum_{i=1}^n \lambda^i + \alpha \lambda^{n+1}$  are in the annulus  $\frac{1}{1+\alpha} \leq |z| \leq 1$ . The Landau theorem (Theorem 5) implies that for the zeros of  $p_{n+1}^*(\lambda)$ , we have the lower estimate

$$|z| \geq \frac{t}{1 + \max_{1 \leq i \leq n} \{ct^i, \alpha t^{n+1}\}}.$$

Since  $0 < c < 1$ ,  $ct > ct^i$  ( $1 \leq i \leq n$ ) for  $0 < t < 1$  and  $ct = \alpha t^{n+1}$  for  $t = \sqrt[n]{\frac{c}{\alpha}}$ . If  $t_1 = \sqrt[n]{\frac{c}{\alpha}}$ , then for the quantity  $g(t)$  given by Theorem 5,

$$g(t_1) = \frac{\sqrt[n]{\frac{1+\alpha}{n\alpha}}}{1 + \frac{1+\alpha}{n} \sqrt[n]{\frac{1+\alpha}{n\alpha}}} \rightarrow 1 \quad (n \rightarrow \infty)$$

holds. □

Han and Neumann [12] investigated the Nelder-Mead method when it generates a sequence of simplices in  $\mathbb{R}^n$  such that

$$S^{(k)} = \left[0, v^{(k+n-1)}, \dots, v^{(k+1)}, v^{(k)}\right]$$

and

$$f(0) < f\left(v^{(k+n-1)}\right) < \dots < f\left(v^{(k)}\right)$$

for  $k \geq k_0$ . They expressed the next incoming vertex  $v^{(k+n)}$  in a difference equation form

$$v^{(k+n)} = \frac{1 + \tau_k}{n} \sum_{i=1}^{n-1} v^{(k+i)} - \tau_k v^{(k)} \quad \left(\tau_k = 1, \frac{1}{2}, -\frac{1}{2}\right),$$

which has the characteristic equation  $p(\mu) = \mu^n + \frac{1+\tau_k}{n} \sum_{i=1}^{n-1} \mu^{n-i} - \tau_k = 0$ . Introducing  $M_k = \left[v^{(k+n-1)}, \dots, v^{(k+1)}, v^{(k)}\right]$  they expressed  $M_{k+1}$  in the form  $M_{k+1} = M_k A_k$ , where

$$A_k = \begin{bmatrix} \frac{1+\tau_k}{n} e & I_{n-1} \\ -\tau_k & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

The characteristic polynomial of  $A_k$  coincides with  $p(\mu)$  and  $A_k$  is close to  $T(\alpha)P_1 \in \mathbb{R}^{(n+1) \times (n+1)}$ . Han, Neumann and Xu [11] investigated the zeros of  $p(\mu)$  in the form of the two parameter polynomial

$$\widehat{p}_n(\lambda) = b - a \sum_{i=1}^{n-1} \lambda^i + \lambda^n \quad (a, b \in \mathbb{C}) \tag{12}$$

of degree  $n$  using a different version of the Schur-Cohn criterion (see Marden [30]) and a different technique. Theorems 4.1, 4.2 and 5.1 of [11] are those results that can be related to this paper.

**Theorem 6** (Han-Neumann-Xu [11], Theorem 4.1) *Suppose that  $a, b \in \mathbb{R}$  and  $a \neq 0$ .*

- (i) *Assume that  $|b| < 1$ . The polynomial (12) has one root in the interior of the unit disk and the remaining roots on the unit circle if  $\frac{b+1}{a} = -1$ .*
- (ii) *Assume that  $|b| < 1$ . The polynomial (12) has one root on the unit circle and the remaining roots in the interior of the unit disk, if  $\frac{b+1}{a} = n - 1$ .*
- (iii) *Assume that  $|b| > 1$ . The polynomial (12) has one root in the exterior of the unit disk and the remaining roots on the unit circle if  $\frac{b+1}{a} = -1$ .*
- (iv) *Assume that  $|b| > 1$ . The polynomial (12) has one root on the unit circle and the remaining roots in the exterior of the unit disk if  $\frac{b+1}{a} = n - 1$ .*

**Theorem 7** (Han-Neumann-Xu [11], Theorem 4.2) *Consider the polynomial (12). If  $b = 1$  and  $0 \leq a \leq \frac{2}{n-1}$ , then all roots of the polynomial (12) are on the unit circle.*

Denote by  $\lambda_{OC}(n)$  any root of the polynomial (12) with the coefficients  $a = \frac{3}{2n}$ ,  $b = \frac{1}{2}$  (outside contraction for  $\alpha = \frac{1}{2}$ ). Similarly, denote by  $\lambda_{IC}(n)$  any root of the polynomial (12) with the coefficients  $a = \frac{1}{2n}$ ,  $b = -\frac{1}{2}$  (inside contraction for  $\alpha = -\frac{1}{2}$ ).

**Theorem 8** (Han-Neumann-Xu [11], Theorem 5.1) *For the values  $\lambda_{OC}(n)$  and  $\lambda_{IC}(n)$ ,*

$$\lim_{n \rightarrow \infty} |\lambda_{OC}(n)| = 1$$

and

$$\lim_{n \rightarrow \infty} |\lambda_{IC}(n)| = 1.$$

In this paper we investigated a set of polynomials of the form

$$p_\ell(\lambda) = \alpha - c \sum_{i=1}^{\ell-1} \lambda^i + \lambda^\ell \quad (c = \frac{1+\alpha}{n}), \tag{13}$$

where the parameters are  $\alpha$  and  $\ell$  ( $2 \leq \ell \leq n + 1$ ). In our case  $b = \alpha$ ,  $a = \frac{1+\alpha}{n}$  implying that  $(b + 1) / a = n$  for  $2 \leq \ell \leq n + 1$  (see Theorem 3).

Theorem 4.1 of [11] assumes that either  $(b + 1) / a = -1$  or  $(b + 1) / a = n - 1$ . In the latter case we can write that  $\hat{p}_\ell(\lambda) = b - \frac{1+b}{\ell-1} \sum_{i=1}^{\ell-1} \lambda^i + \lambda^\ell$ . For  $b = \alpha$  and  $\ell = n + 1$  ( $j = 1$ ) the polynomial  $\hat{p}_\ell(\lambda)$  coincides with (13). Hence case (iv) of Theorem 3 also follows from case (ii) of Theorem 4.1 of [11]. In turn, case (iv) of Theorem 4.1 of [11] follows from case (vii) of Theorem 3. For  $\ell \neq n + 1$ , the two theorems are clearly different.

In addition, we note that case (vi) of Theorem 3 also follows from Theorem 4.2 of [11] and Lemma 3 implies Theorem 5.1 of [11] for  $\alpha = \pm \frac{1}{2}$ .

### 4 Examples of convergence behavior

Following McKinnon [29] and Han and Neumann [12] we investigate simple behavior patterns of the simplex sequences  $\{S^{(k)}\}$  generated by the Nelder-Mead method. If the simplex sequence  $\{S^{(k)}\}$  generated by the Nelder-Mead algorithm is convergent, that is  $\lim_{k \rightarrow \infty} S^{(k)} = S^\infty$  for some  $S^\infty \in \mathbb{R}^{n \times (n+1)}$ , then  $f(x_i^{(k)}) \rightarrow f(S^\infty e_i)$  ( $i = 1, 2, \dots, n + 1$ ) provided that  $f$  is continuous at the points  $S^\infty e_i$  ( $i = 1, 2, \dots, n + 1$ ). Note that if for some vector  $\hat{x}$ ,  $x_j^{(k)} \rightarrow \hat{x}$  ( $k \rightarrow \infty$ ,  $j = 1, 2, \dots, n + 1$ ), then  $S^{(k)} \rightarrow \hat{x}e^T$ , which is a rank-one matrix of special form.

We show examples where the incoming vertex  $v$  satisfies

$$f_1^{(k)} \leq \dots \leq f_{j-1}^{(k)} \leq f(v) \leq f_j^{(k)} < f_{j+1}^{(k)} \leq \dots \leq f_{n+1}^{(k)} \quad (k \geq 0)$$

with a fixed index  $j$  and the type of  $v$  (reflection, expansion, outside contraction, or inside contraction point) is the same for  $k \geq 0$ .<sup>3</sup> Hence  $S^{(k)} = S^{(0)} [T(\alpha) P_j]^k$  for  $k \geq 0$ . Using the examples we can specify the type of convergence which is studied in the rest of the paper.

Assume that  $|\alpha| < 1$  and  $S^{(0)} = [S_{j-1}^{(0)}, S_{n+2-j}^{(0)}]$  ( $S_{j-1}^{(0)} \in \mathbb{R}^{n \times (j-1)}$ ). It follows from Theorem 3 (formula (10)) that

$$\lim_{k \rightarrow \infty} S^{(k)} = \lim_{k \rightarrow \infty} S^{(0)} [T(\alpha) P_j]^k = [S_{j-1}^{(0)}, S_{j-1}^{(0)} c e e^T (I_{n+2-j} - A_{n+2-j})^{-1}]$$

If  $\text{rank}(S_{j-1}^{(0)}) \geq 2$ , then  $\lim_{k \rightarrow \infty} S^{(k)}$  cannot be of the form  $\hat{x}e^T$  for some vector  $\hat{x}$ ,  $\text{diam}(S^{(k)}) \rightarrow 0$ , and  $f_i^{(k)}$ 's do not converge to the same limit. For  $j = 1, 2$ , we can write  $T(\alpha) P_j$  in the form

$$T(\alpha) P_j = F \begin{bmatrix} 1 & 0^T \\ b_j & C_j \end{bmatrix} F^{-1} \quad \left( F = \begin{bmatrix} 1 & -e^T \\ 0 & I_n \end{bmatrix} \right).$$

For  $|\alpha| < 1$ , Theorem 3 implies that  $\rho(C) < 1$  and so

$$[T(\alpha) P_j]^k \rightarrow \begin{bmatrix} 1 - e^T (I - C_j)^{-1} b_j \\ (I - C_j)^{-1} b_j \end{bmatrix} e^T = B.$$

Hence  $\lim_{k \rightarrow \infty} S^{(k)} = S^{(0)} B = we^T$  for some vector  $w$ . For  $j = 1$ ,  $b_1 \neq 0$ , while for  $j = 2$ ,  $b_j = 0$ . In the latter case  $S^{(0)} B = x_1^{(0)} e^T$ .

Next we show five examples for different behavior patterns of  $\{S^{(k)}\}$ . Examples 1, 2, 3 and 5 are two dimensional, while Example 4 is  $n$ -dimensional. Let  $J = \{1, 2, \dots, n + 1\}$ .

In Example 1 the incoming vertex is  $x_e^{(k)}$  for  $k \geq 0$ ,  $f$  has no finite minimum,  $f_i^{(k)} \rightarrow 0$  ( $k \rightarrow \infty, i \in J$ ) and  $\text{diam}(S^{(k)}) \rightarrow \infty$ . In Example 2 the incoming vertex (as expansion point) is  $x_r^{(k)}$  for  $k \geq 0$ ,  $\text{diam}(S^{(k)})$  is constant,  $\{S^{(k)}\}$  is unbounded and three functions are given so that  $f_i^{(k)} \rightarrow -\infty$  ( $i \in J$ ) or  $f_i^{(k)} \rightarrow 0$  ( $i \in J$ ) holds.

<sup>3</sup>Here we adopt the notations defined in Section 1.

In Example 3 the incoming vertex is  $x_{ic}^{(k)}$  for  $k \geq 0$ ,  $x_i^{(k)} \rightarrow x_1^{(0)}$  ( $i = 1, 2, 3$ ) and three functions are given so that  $x_1^{(0)}$  is not a stationary point of  $f$  or it is a minimum point of  $f$ . In the  $n$ -dimensional Example 4 the incoming vertex is  $x_{ic}^{(k)}$  which replaces  $x_{n+1}^{(k)}$  for  $k \geq 0$ . Hence  $x_i^{(k)} = x_i^{(0)}$  ( $i = 1, \dots, n$ ),  $x_{n+1}^{(k)} \rightarrow x_c^{(0)}$  and  $\text{diam}(S^{(k)}) \rightarrow 0$ . For  $n \neq 3$ ,  $\lim_k x_{n+1}^{(k)} = x_c^{(0)}$  is not a stationary point of the three given functions  $f$ .

In Example 5 the incoming vertex is  $x_{ic}^{(k)}$  which replaces  $x_3^{(k)}$  for  $k \geq 0$ ,  $x_3^{(k)} \rightarrow x_c^{(0)}$ ,  $\text{diam}(S^{(k)}) \rightarrow 0$ . Two functions are given such that  $\lim_k x_3^{(k)} = x_c^{(0)}$  is not a stationary point for the first function, while it is a saddle point for the second function.

*Example 1* The expansion point  $x_e^{(k)}$  is the incoming vertex infinitely many times if

$$f_e^{(k)} < f_r^{(k)} < f_1^{(k)} \leq f_2^{(k)} \leq f_3^{(k)} \quad (k \geq 0). \tag{14}$$

In this case  $S^{(k)} = S^{(0)} B_k$  where  $B_k = [T(2) P_1]^k$  is not bounded. Select

$$S^{(0)} = \begin{bmatrix} \frac{5}{6} - \frac{7}{66}\sqrt{33} & \frac{1}{3} - \frac{1}{11}\sqrt{33} & \frac{1}{3} - \frac{1}{33}\sqrt{33} \\ \frac{7}{66}\sqrt{33} + \frac{5}{6} & \frac{1}{11}\sqrt{33} + \frac{1}{3} & \frac{1}{33}\sqrt{33} + \frac{1}{3} \end{bmatrix} = \begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix}.$$

The rows of  $S^{(0)}$  are the left eigenvectors of  $T(2) P_1$  corresponding to  $\lambda_2 = \frac{1}{4}\sqrt{33} + \frac{1}{4}$  and  $\lambda_3 = \frac{1}{4} - \frac{1}{4}\sqrt{33}$ , respectively. Hence if condition (14) holds, then

$$S^{(k)} = \begin{bmatrix} \lambda^k a & \lambda^k c & \lambda^k e \\ \mu^k b & \mu^k d & \mu^k f \end{bmatrix},$$

$$x_r^{(k)} = \begin{bmatrix} \lambda^k (a + c - e) \\ \mu^k (b + d - f) \end{bmatrix}, \quad x_e^{(k)} = \begin{bmatrix} \frac{1}{2}\lambda^k (3a + 3c - 4e) \\ \frac{1}{2}\mu^k (3b + 3d - 4f) \end{bmatrix}.$$

Let  $f(x, y) = (1 + x^2 + y^2)^{-1}$ . Since  $|a| > |c| > |e|$  and  $|b| > |d| > |f|$ , condition  $f_1^{(k)} < f_2^{(k)} < f_3^{(k)}$  clearly holds. Inequalities  $f_r^{(k)} < f_1^{(k)}$  and  $f_e^{(k)} < f_r^{(k)}$  hold if and only if

$$\lambda^{2k} (a + c - e)^2 + \mu^{2k} (b + d - f)^2 > \lambda^{2k} a^2 + \mu^{2k} b^2$$

and

$$\lambda^{2k} \frac{(3a + 3c - 4e)^2}{4} + \mu^{2k} \frac{(3b + 3d - 4f)^2}{4} > \lambda^{2k} (a + c - e)^2 + \mu^{2k} (b + d - f)^2$$

hold, respectively. The last two inequalities can be verified by direct calculation. Hence (14) holds and for  $k \rightarrow \infty$ ,  $\text{diam}(S^{(k)}) \rightarrow \infty$  and  $f_i^{(k)} \rightarrow 0$  for  $i = 1, 2, 3$ .

*Example 2* The expansion point  $x_r^{(k)}$  is the incoming vertex infinitely many times if

$$f_r^{(k)} < f_1^{(k)} \leq f_2^{(k)} \leq f_3^{(k)} \wedge f_r^{(k)} \leq f_e^{(k)} \quad (k \geq 0). \tag{15}$$

In this case  $S^{(k)} = S^{(0)} B_k$  where  $B_k = [T(1) P_1]^k$ . Since  $T(1) P_1$  has a  $2 \times 2$  Jordan block belonging to  $\lambda = 1$ ,  $B_k$  is not bounded. Define

$$S^{(0)} = \begin{bmatrix} 1 & 0 & -1 \\ 2 & -2 & 2 \end{bmatrix}.$$



If condition (15) holds, then

$$S^{(k)} = \begin{bmatrix} k + 1 & k & k - 1 \\ (-1)^k 2 & (-1)^{k+1} 2 & (-1)^k 2 \end{bmatrix}$$

and

$$x_r^{(k)} = \begin{bmatrix} k + 2 \\ (-1)^{k+1} 2 \end{bmatrix}, \quad x_e^{(k)} = \begin{bmatrix} k + \frac{7}{2} \\ (-1)^{k+1} 4 \end{bmatrix}.$$

Note that  $\{S^{(k)}\}$  is unbounded, while  $\text{diam}(S^{(k)})$  is constant. It is easy to verify that condition (15) holds for the functions  $f_1(x, y) = \frac{1}{2} - \frac{1}{2}x + \frac{1}{4}|y| + \frac{1}{4}||y| - 2|$ ,  $f_2(x, y) = \frac{1}{1+x^2} + \frac{1}{12}y^2 - \frac{1}{3}$  and  $f_3(x, y) = \frac{\sin(|y|-1.9)}{1+x^2}$ . In case of  $f_1$ ,  $f_i^{(k)} \rightarrow -\infty$  ( $k \rightarrow \infty, i = 1, 2, 3$ ). For  $f_2$  and  $f_3$ ,  $f_i^{(k)} \rightarrow 0$  ( $i = 1, 2, 3$ ). Note that  $\inf f_1 = -\infty$ ,  $(0, 0)$  is a saddle point of  $f_2$ ,  $\inf f_2 = -\frac{1}{3}$ , and  $-1 \leq f_3 \leq 1$ .

Examples 1 and 2 show that the assumption on the boundedness of  $\{B_k\}$  is justified.

McKinnon [29] proved the convergence behavior (17) of Example 3 for the convex function

$$f_M(x, y) = \begin{cases} \theta\varphi|x|^\tau + y + y^2, & \text{if } x \leq 0 \\ \theta x^\tau + y + y^2, & \text{if } x \geq 0 \end{cases} \quad (\theta = 6, \tau = 2, \varphi = 60)$$

with initial simplex

$$S^{(0)} = \begin{bmatrix} 0 & \frac{1+\sqrt{33}}{8} & 1 \\ 0 & \frac{1-\sqrt{33}}{8} & 1 \end{bmatrix} = \begin{bmatrix} 0 & a & 1 \\ 0 & b & 1 \end{bmatrix}, \tag{16}$$

where the common limit point  $[0, 0]^T$  of the simplex vertices is not a stationary point of  $f_M$ . Using simplex  $S^{(0)}$  of (16) we show other functions that generate the same simplex sequence, while the limit point  $[0, 0]^T$  is either a stationary point or a nonstationary point depending on the particular function.

*Example 3* The inside contraction point  $x_{ic}^{(k)}$  is the incoming vertex infinitely many times, if

$$f_1^{(k)} \leq f_{ic}^{(k)} < f_2^{(k)} < f_3^{(k)} \leq f_r^{(k)} \quad (k \geq 0) \tag{17}$$

If such a case occurs, then it is generally true that

$$S^{(k)} = S^{(0)} \left[ T \left( -\frac{1}{2} \right) P_2 \right]^k \rightarrow S^{(0)} e_1 e^T = [x_1^{(0)}, x_1^{(0)}, x_1^{(0)}],$$

no matter what  $x_1^{(0)}$  is. The rows of  $S^{(0)}$  are the left eigenvectors of  $T \left( -\frac{1}{2} \right) P_2$  belonging to eigenvalues  $\lambda = \frac{1+\sqrt{33}}{8}$  and  $\mu = \frac{1-\sqrt{33}}{8}$ , respectively. If (17) holds, then

$$S^{(k)} = \begin{bmatrix} 0 & \lambda^k a & \lambda^k \\ 0 & \mu^k b & \mu^k \end{bmatrix}, \quad x_r^{(k)} = \begin{bmatrix} (a - 1) \lambda^k \\ (b - 1) \mu^k \end{bmatrix}, \quad x_{ic}^{(k)} = \begin{bmatrix} \left( \frac{1}{4}a + \frac{1}{2} \right) \lambda^k \\ \left( \frac{1}{4}b + \frac{1}{2} \right) \mu^k \end{bmatrix}.$$

Let  $f(x, y) = g(x) + h(y)$ . The values of  $a, b, \lambda$  and  $\mu$  imply that  $\lambda > |\mu|$  and the arguments  $x$  and  $y$  at the points  $x_i^{(k)}, x_r^{(k)}$  and  $x_{ic}^{(k)}$  will be in the range  $x \in [-1, 1]$  and

$y \in I_y = \left[-\left(\frac{1}{8}\sqrt{33} + \frac{7}{8}\right), \left(\frac{1}{8}\sqrt{33} + \frac{7}{8}\right)\right]$ , respectively. Select  $g(x)$  and  $h(y)$  such that  $g(x) = \max(\alpha x, -\beta x)$ ,  $\alpha, \beta > 0$  and  $|h(y)| \leq \gamma |y|$  for  $y \in I_y$ . The inequality  $g(x) - \gamma |y| \leq f(x, y) \leq g(x) + \gamma |y|$  implies that for  $\alpha > \left(\frac{7}{8}\sqrt{33} + \frac{41}{8}\right) \gamma$  and  $\beta > \left(\frac{1}{2}\sqrt{33} + \frac{7}{2}\right) \alpha + \left(\frac{11}{8}\sqrt{33} + \frac{69}{8}\right) \gamma$ , condition (17) holds. Hence for functions  $f_4(x, y) = \max(7x, -53x) + \frac{1}{2} \sin(y)$  ( $\gamma = \frac{1}{2}$ ),  $f_5(x, y) = \max(13x, -100x) + 0.1y^3$  ( $\gamma = 1$ ) and  $f_6(x, y) = \max(13x, -100x) + \frac{1}{4}y^2$  ( $\gamma = 1$ ), condition (17) holds. The point  $(0, 0)$  is not a stationary point for  $f_4$  and  $f_5$ , while it is a minimum point for  $f_6$ .

Note that for all three functions, we have the same limit point, and the simplex sequence depends on the initial simplex and the relative function value distribution.

Han and Neumann [12] investigated the behavior pattern

$$0 = f_1^{(k)} < f(v) < f_2^{(k)} < \dots < f_{n+1}^{(k)} \quad (k \geq k_0). \tag{18}$$

where the incoming point  $v$  is either  $x_{oc}^{(k)}$  or  $x_{ic}^{(k)}$ . Hence

$$S^{(k+1)} = S^{(k)} T(\alpha_k) P_2 \quad \left(\alpha_k \in \left\{-\frac{1}{2}, \frac{1}{2}\right\}\right). \tag{19}$$

Under this assumption they proved the convergence  $x_i^{(k)} \rightarrow 0$  ( $k \rightarrow \infty$ ) for  $f(x) = x^T x$  ( $x \in \mathbb{R}^n$ ). For  $n = 2$ , they gave an initial simplex  $S^{(0)}$  for which condition (18) with  $\alpha_k \equiv \frac{1}{2}$  is fulfilled.

Lagarias et al. ([23], Lemma 5.1) investigated a somewhat similar but more complicated case where the best vertex  $x_1^{(k)}$  is constant for all  $k$ .

*Example 4* Assume that  $x_{ic}^{(k)}$  is the incoming vertex infinitely many times such that

$$f_1^{(k)} \leq f_2^{(k)} \leq \dots \leq f_n^{(k)} \leq f_{ic}^{(k)} < f_{n+1}^{(k)} \leq f_r^{(k)} \quad (k \geq 0) \tag{20}$$

and  $n \geq 2$ . In this case

$$S^{(k)} = S^{(0)} \left[ T\left(-\frac{1}{2}\right) \right]^k \rightarrow S^{(0)} \begin{bmatrix} I_n & \frac{1}{n}e \\ 0 & 0 \end{bmatrix} = \left[ x_1^{(0)}, \dots, x_n^{(0)}, \frac{1}{n} \sum_{i=0}^n x_i^{(0)} \right].$$

We select

$$S^{(0)} = \begin{bmatrix} I_{n-1,n} & \frac{1}{n}e \\ 0 & 1 \end{bmatrix} \quad \left( I_{n-1,n} = [\delta_{ij}]_{i,j=1}^{n-1,n} \right),$$

whose rows are the left eigenvectors of  $T\left(-\frac{1}{2}\right)$ . If (20) holds, then

$$S^{(k)} = \begin{bmatrix} I_{n-1,n} & \frac{1}{n}e \\ 0 & \frac{1}{2^k} \end{bmatrix}, \quad x_r^{(k)} = \begin{bmatrix} \frac{1}{n}e \\ -\frac{1}{2^k} \end{bmatrix}, \quad x_{ic}^{(k)} = \begin{bmatrix} \frac{1}{n}e \\ \frac{1}{2^{k+1}} \end{bmatrix}.$$

Assume that  $f(x) = \left[ \prod_{i=1}^{n-1} g(x_i) \right] h(x_n)$ , where  $g(x), h(x) > 0$  for  $x \geq 0$ ,  $h(-x) \geq h(x)$  for  $x > 0$ , and  $h(x)$  is strictly monotone increasing for  $x \geq 0$  and strictly monotone decreasing for  $x < 0$ . If  $g(1) < g(0) < g\left(\frac{1}{n}\right)$  holds,

then it follows that function values  $f_i^{(k)} = g(0)^{n-2} g(1) h(0)$  ( $i = 1, \dots, n - 2$ ),  $f_n^{(k)} = g(0)^{n-1} h(0)$ ,  $f_{ic}^{(k)} = g\left(\frac{1}{n}\right)^{n-1} h\left(\frac{1}{2^{k+1}}\right)$ ,  $f_{n+1}^{(k)} = g\left(\frac{1}{n}\right)^{n-1} h\left(\frac{1}{2^k}\right)$  and  $f_r^{(k)} = g\left(\frac{1}{n}\right)^{n-1} h\left(-\frac{1}{2^k}\right)$  satisfy inequality (20). Since  $x_{n+1}^{(k)} \rightarrow \left[\frac{1}{n}e^T, 0\right]^T$ , we have  $f_{n+1}^{(k)} \rightarrow \left[g\left(\frac{1}{n}\right)\right]^{n-1} h(0)$ . For  $g$ , we may select  $g(x) = e^{-(x-\frac{1}{3})^2}$ , which has a global maximum at  $x = \frac{1}{3}$ , while  $h$  has a minimum at  $y = 0$ . Hence  $f$  has a saddle point  $\left[\frac{1}{3}e^T, 0\right]^T$ , which is different from  $\lim_{k \rightarrow \infty} x_{n+1}^{(k)}$  if  $n \neq 3$ . However for  $n = 2$ ,  $\lim_{k \rightarrow \infty} x_3^{(k)} = x_{saddle}$ . We can select any of the functions  $f_7(x) = e^{-\sum_{i=1}^{n-1} (x_i - \frac{1}{3})^2} (1 + |x_n|)$ ,  $f_8(x) = e^{-\sum_{i=1}^{n-1} (x_i - \frac{1}{3})^2} (1 + \sin(|x_n|))$  and  $f_9(x) = e^{-\sum_{i=1}^{n-1} (x_i - \frac{1}{3})^2} \sin(1 + x_n^2)$ . Note that  $0 < f_7(x) \leq 1 + |x_n|$  and  $f_7$  has no finite minimum. Functions  $f_8$  and  $f_9$  have an infinite number of global maximum and minimum points. Also note that  $\text{diam}(S^{(k)}) \rightarrow 0$ .

For  $n = 2$ , there are plenty of similar cases.

**Lemma 5** Assume that  $S^{(0)} = [x_1^{(0)}, x_2^{(0)}, x_3^{(0)}]$  is such that  $f_1^{(0)} \leq f_2^{(0)} < f_3^{(0)}$  and define  $\varphi(t) = (1+t)x_c^{(0)} - tx_3^{(0)}$ . If, in addition,  $f$  is such that (a)  $f(\varphi(t))$  is continuous on  $[-1, 1]$ ; (b)  $f(\varphi(t)) \geq f(\varphi(-t))$  for  $t \in [0, 1]$ ; (c)  $f(\varphi(t))$  is strictly monotone decreasing on  $[-1, 0]$ ; (d)  $f(\varphi(t)) > f(\varphi(0)) = f_c^{(0)} \geq f_2^{(0)}$  ( $t \in [-1, 1], t \neq 0$ ), then

$$f_1^{(k)} \leq f_2^{(k)} \leq f_{ic}^{(k)} < f_3^{(k)} \leq f_r^{(k)} \tag{21}$$

holds for all  $k = 0, 1, 2, \dots$ ,  $x_3^{(k)} \rightarrow x_c^{(0)}$ , and  $f_3^{(k)} \rightarrow f_c^{(0)}$ .

*Proof* Assume that for some  $-1 \leq t < 0$ ,  $x_3 = \varphi(t)$  and  $f_1^{(0)} \leq f_2^{(0)} < f(\varphi(t))$ . Then  $x_r = \varphi(-t)$ ,  $x_{ic} = \varphi\left(\frac{t}{2}\right)$ , and (b) and (c) imply that

$$f(x_r) = f(\varphi(-t)) \geq f(\varphi(t)) = f(x_3) > f\left(\varphi\left(\frac{t}{2}\right)\right) = f(x_{ic}).$$

Condition (d) implies that  $f_1^{(0)} \leq f_2^{(0)} < f(\varphi\left(\frac{t}{2}\right))$ . □

*Example 5* Consider function  $f(x, y) = \frac{1}{4}(x + |x|) + \frac{1}{2}|x - |x|| + g(y)$ , where

$$g(y) = \begin{cases} 0.2 \sin(10\pi y - 5\pi), & \text{if } 0.5 \leq y \leq 0.7 \\ 0, & \text{otherwise} \end{cases}.$$

Select  $x_1^{(0)} = [0, 0.5]^T$ ,  $x_2^{(0)} = [0, 0.7]^T$ ,  $x_3^{(0)} = [0.5, 0.6]^T$ . Then by Lemma 5  $x_3^{(k)} \rightarrow x_{ic}^{(0)} = [0, 0.6]^T$ , which is not a stationary point. Assume now that  $f(x, y) = g(x) - h(y)$ , where  $g$  and  $h$  are continuous real functions,  $g(x) > 0$  for  $x \neq 0$ ,  $g(0) = 0$ ,  $g(x)$  is strictly monotone increasing for  $x \geq 0$ ,  $g(x)$  is strictly monotone decreasing for  $x < 0$ ,  $g(-x) \geq g(x)$  ( $x \geq 0$ ),  $h(y) \geq 0$  for  $y \neq 0$ ,  $h(0) = 0$  and

$h(-y) \geq h(y)$  for  $y \geq 0$ . Select  $x_1^{(0)} = [0, -a]^T$ ,  $x_2^{(0)} = [0, a]^T$  and  $x_3^{(0)} = [b, 0]^T$  with  $a, b > 0$ . Lemma 5 implies that  $x_3^{(i)}$  converges to the saddle point  $x_c^{(0)} = [0, 0]^T$ .

The examples show that for different functions the Nelder-Mead algorithm may generate the same simplex sequence whose limit vertices may be different from or equal to a stationary point of the function. They also show that in case of convergence the simplex vertices  $x_j^{(k)}$  either converge to the same vector  $\hat{x}$  or converge to different vectors as shown by Examples 4 and 5.

Assume that the simplex vertices  $x_j^{(k)}$  ( $j = 1, 2, \dots, n + 1$ ) converge to the same vector  $\hat{x}$  as  $k \rightarrow \infty$  and  $f$  is continuous at  $\hat{x}$ . Then

$$\lim_{k \rightarrow \infty} S^{(k)} = [\hat{x}, \dots, \hat{x}] = \hat{x}e^T \tag{22}$$

and  $f_i^{(k)} \rightarrow f(\hat{x})$  ( $i = 1, \dots, n + 1$ ). Assume that  $B_k \rightarrow B = we^T$ . Then  $S^{(k)} = S^{(0)}B_k \rightarrow S^{(0)}we^T = \hat{x}e^T$  and

$$\text{diam}(S^{(k)}) = \max_{i,j} \|S^{(k)}(e_i - e_j)\| \leq \max_{i,j} \|S^{(0)}\| \|B_k(e_i - e_j)\| \rightarrow 0. \tag{23}$$

Since  $B(e_i - e_j) = 0$ ,  $B_k(e_i - e_j) = (B_k - B)(e_i - e_j)$ , we also have the speed estimate

$$\text{diam}(S^{(k)}) \leq \sqrt{2} \|S^{(0)}\| \|B_k - B\|. \tag{24}$$

Note that properties  $f_i^{(k)} \rightarrow \hat{f}$  ( $i = 1, \dots, n + 1$ ) and  $\text{diam}(S^{(k)}) \rightarrow 0$  ( $k \rightarrow \infty$ ) were proved directly for strictly convex two-dimensional functions by Lagarias et al. [23] without relating the results to the stationary point of  $f$ . Except for Kelley [20], [19] and Lagarias et al. [22] no general result is known as yet on the convergence to a stationary point of the target function  $f$ .

Upon the basis of the preceding arguments we restrict our study to the convergence  $S^{(k)} \rightarrow \hat{x}e^T$ , which implies  $f_i^{(k)} \rightarrow f(\hat{x})$  ( $i = 1, \dots, n + 1$ ) and the speed estimate (24).

## 5 The convergence of the Nelder-Mead method

The convergence result will be proved in several steps. First using a fixed similarity we transform the matrices  $T_i P^{(i)} \in \mathcal{T}$  to a common lower block triangular form and identify a subset of  $\mathcal{T}$  that might have the RCP property. Next we prove a lemma on the convergence of the product of lower block triangular matrices. Finally, under Assumption (A) we prove the convergence in Section 5.3

### 5.1 A common similarity transformation

Since  $e^T$  is a left eigenvector of each  $T_i P^{(i)} \in \mathcal{T}$ , there exists a common similarity transformation that makes them block lower triangular (for a more general case, see Theorem 6.10 of Hartfiel [13]).

**Lemma 6** For all  $T_i P^{(i)} \in \mathcal{T}$ , matrix  $F^{-1} T_i P^{(i)} F$  has the form

$$F^{-1} T_i P^{(i)} F = \begin{bmatrix} 1 & 0 \\ b_i & C_i \end{bmatrix} \quad \left( F = \begin{bmatrix} 1 & -e^T \\ 0 & I_n \end{bmatrix} \right), \tag{25}$$

where  $b_i \in \mathbb{R}^n$  and  $C_i \in \mathbb{R}^{n \times n}$  are defined by  $T_i P^{(i)}$ .

*Proof* For  $j > 1$  we can write

$$T(\alpha) P_j = \begin{bmatrix} 1 & ce_{j-1}^T \\ 0 & W \end{bmatrix} \quad (W \in \mathbb{R}^{n \times n}),$$

and so

$$F^{-1} T(\alpha) P_j F = \begin{bmatrix} 1 & -e^T + ce_{j-1}^T + e^T W \\ 0 & W \end{bmatrix}.$$

Since  $e^T W e_{j-1} = (n - 1)c - \alpha = 1 - c$ ,  $e^T W = [1, \dots, 1 - c, 1, \dots, 1]$ , we obtain the form

$$F^{-1} T(\alpha) P_j F = \begin{bmatrix} 1 & 0 \\ 0 & W \end{bmatrix}.$$

For  $j = 1$ ,

$$T(\alpha) P_1 = \begin{bmatrix} c & e_1^T \\ z & W \end{bmatrix} \quad (W \in \mathbb{R}^{n \times n})$$

with  $z = [c, \dots, c, -\alpha]^T$ . Hence

$$F^{-1} T(\alpha) P_1 F = \begin{bmatrix} c + e^T z & -ce^T + e_1^T - e^T z e^T + e^T W \\ z & -ze^T + W \end{bmatrix}.$$

Since  $e^T W = [0, 1, \dots, 1]$ ,  $e^T z = 1 - c$ ,  $c + e^T z = 1$ ,

$$-ce^T + e_1^T - e^T z e^T + e^T W = -ce^T + e_1^T - (1 - c)e^T + e^T W = 0.$$

The final result is

$$F^{-1} T(\alpha) P_1 F = \begin{bmatrix} 1 & 0 \\ z & -ze^T + W \end{bmatrix}.$$

For  $T(\alpha) P_j$  ( $j > 1$ ),  $b = 0$ , and for  $T(\alpha) P_1$ ,  $\|b\|_2 = ((n - 1)c^2 + \alpha^2)^{\frac{1}{2}}$ .

Note that  $T_{shr} P = \frac{1}{2} P + \frac{1}{2} e_1 e^T$  and  $P = [e_{i_1}, \dots, e_{i_{n+1}}]$ . If  $i_1 = 1$ , then

$$T_{shr} P = \begin{bmatrix} 1 & \frac{1}{2} e^T \\ 0 & W_1 \end{bmatrix}$$

where  $W_1$  is an  $n \times n$  permutation matrix multiplied by  $\frac{1}{2}$ . Hence  $e^T W_1 = \frac{1}{2} e^T$  and

$$F^{-1} T_{shr} P F = \begin{bmatrix} 1 & -\frac{1}{2} e^T + e^T W_1 \\ 0 & W_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & W_1 \end{bmatrix}.$$

If  $i_1 > 1$  and  $T_{shr} P e_j = e_1$ , then

$$T_{shr} P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} e^T + \frac{1}{2} e_{j-1}^T \\ \frac{1}{2} e_{i_1-1} & W_2 \end{bmatrix},$$

where  $W_2 e_{j-1} = 0$ ,  $e^T W_2 e_i = \frac{1}{2}$  ( $i \neq j - 1$ ) and  $e^T W_2 = \frac{1}{2} (e^T - e_{j-1}^T)$ . Since  $e^T e_{i_1-1} = 1$ ,

$$\begin{aligned}
 F^{-1} T_{shr} P F &= \begin{bmatrix} \frac{1}{2} + \frac{1}{2} e^T e_{i_1-1} & \frac{1}{2} e_{j-1}^T - \frac{1}{2} e^T e_{i_1-1} e^T + e^T W_2 \\ \frac{1}{2} e_{i_1-1} & -\frac{1}{2} e_{i_1-1} e^T + W_2 \end{bmatrix} \\
 &= \begin{bmatrix} 1 & 0 \\ \frac{1}{2} e_{i_1-1} & -\frac{1}{2} e_{i_1-1} e^T + W_2 \end{bmatrix}
 \end{aligned}$$

If  $i_1 = 1$ , then the first column entries of  $F^{-1} T_{shr} P F$  are 0 except for entry  $(1, 1)$ . If  $i_1 \geq 2$ , then entry  $(i_1, 1)$  is  $\frac{1}{2}$ , while the remaining entries are 0 ( $\ell \neq 1, i_1$ ). Hence  $\|b\|_2 \leq \frac{1}{2}$ . The entries of submatrix  $C$  are only 0,  $\frac{1}{2}$  and  $-\frac{1}{2}$ . In column  $j$ , there can be at most two nonzero elements. Note that  $\rho(C_i) = \frac{1}{2}$  and  $\|C_i\|_1 \leq 1$  for  $T_i P^{(i)} = T_{shr} P$  ( $P \in \mathcal{P}_{n+1}$ ). □

Note that matrices  $C_i$  and their norms play the key role in the convergence proof. Accordingly we divide the set  $\mathcal{T}$  in two disjoint sets

$$\mathcal{W}_1 = \left\{ T \left( \frac{1}{2} \right) P_j, T \left( -\frac{1}{2} \right) P_j : j = 1, 2 \right\} \cup \{ T_{shr} P : P \in \mathcal{P}_{n+1} \} \tag{26}$$

and

$$\begin{aligned}
 \mathcal{W}_2 = \{ &T(2) P_1, T(1) P_j : j = 1, \dots, n \} \\
 &\cup \left\{ T \left( \frac{1}{2} \right) P_j, T \left( -\frac{1}{2} \right) P_j : j = 3, \dots, n + 1 \right\}. \tag{27}
 \end{aligned}$$

The matrices of  $\mathcal{W}_1$  correspond to the inside and outside contraction operations, when the incoming vertices are inserted in the first or the second position of the ordering (3) or they correspond to any shrinking operation. The matrices of  $\mathcal{W}_2$  correspond to the remaining operations of  $\mathcal{T}$ . Theorem 3 and Corollary 1 imply that

$$\rho(C_i) < 1 \quad (T_i P^{(i)} \in \mathcal{W}_1) \tag{28}$$

and

$$\rho(C_i) \geq 1 \quad (T_i P^{(i)} \in \mathcal{W}_2). \tag{29}$$

Note that for each matrix  $T_i P^{(i)} \in \mathcal{W}_1$ , an induced matrix norm  $\|\cdot\|$  exists such that  $\rho(C_i) \leq \|C_i\| < 1$ . However for any  $T_i P^{(i)} \in \mathcal{W}_2$  and any induced matrix norm  $\|\cdot\|$ , only  $1 \leq \rho(C_i) \leq \|C_i\|$  holds. Since  $\max \{ \rho(C_i) : T_i P^{(i)} \in \mathcal{W}_1 \} < 1$ , set  $\mathcal{W}_1$  might be an RCP set if we find a proper induced norm for which  $\|C_i\| \leq 1$  for all  $T_i P^{(i)} \in \mathcal{W}_1$ . In fact we make a stronger restriction in the form of Assumption (A).

### 5.2 A lemma on the convergence of lower block triangular matrices

For  $i \geq 1$ , let

$$A_i = \begin{bmatrix} 1 & 0 \\ b_i & C_i \end{bmatrix} \in \mathbb{R}^{(n+1) \times (n+1)} \quad (C_i \in \mathbb{R}^{n \times n}). \tag{30}$$

**Lemma 7** Assume that  $\left\| \prod_{j=1}^k C_j \right\| \leq c_k$ ,  $\sum_{k=1}^{\infty} c_k$  is convergent ( $< \infty$ ) and  $\|b_k\| \leq \gamma$  for all  $k$ . Then  $L_k = \prod_{j=1}^k A_j$  converges and

$$\lim_{k \rightarrow \infty} L_k = \begin{bmatrix} 1 & 0 \\ \tilde{x} & 0 \end{bmatrix} \tag{31}$$

for some  $\tilde{x}$ .

*Proof* It is easy to see that

$$L_k = \prod_{j=1}^k A_j = \begin{bmatrix} 1 & 0 \\ \sum_{i=1}^k \left( \prod_{j=1}^{i-1} C_j \right) b_i & \prod_{j=1}^k C_j \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ x_k & \prod_{j=1}^k C_j \end{bmatrix}. \tag{32}$$

If  $\sum_{k=1}^{\infty} c_k$  is convergent, then  $c_k \rightarrow 0$ . Hence  $\prod_{j=1}^k C_j \rightarrow 0$  as  $k \rightarrow \infty$ . Since  $s_k = \sum_{j=1}^k c_j$  is convergent, for any  $\varepsilon > 0$  there is a number  $k_0 = k_0(\varepsilon)$  such that for  $m > k \geq k_0$ ,  $|s_m - s_k| < \varepsilon$ . Thus for  $m > k \geq k_0$ , we obtain

$$\|x_m - x_k\| \leq \sum_{i=k+1}^m \left\| \prod_{j=1}^{i-1} C_j \right\| \|b_i\| \leq \gamma \sum_{i=k+1}^m c_{i-1} \leq \gamma \varepsilon.$$

Hence  $x_k \rightarrow \tilde{x}$  for some  $\tilde{x}$ . □

If  $\|C_j\| \leq q < 1$  for  $j \geq 1$ , then  $\left\| \prod_{j=1}^k C_j \right\| \leq q^k$  and the series  $\sum_{i=1}^{\infty} q^i$  is convergent.

### 5.3 The convergence theorem

Formula (25) implies that  $B_k = \prod_{i=1}^k T_i P^{(i)} = F L_k F^{-1}$ , where

$$L_k = \prod_{i=1}^k \begin{bmatrix} 1 & 0 \\ b_i & C_i \end{bmatrix}, \tag{33}$$

and  $B_k$  is convergent if and only if  $L_k$  is convergent. The convergence of the Nelder-Mead algorithm will be proved under the following key condition.

**Assumption (A):** There is an induced matrix norm  $\|A\|_{\vartheta}$  such that if  $T_i P^{(i)} \in \mathcal{W}_1$ , then  $\|C_i\|_{\vartheta} < 1$ .

If Assumption (A) holds, there exist constants  $0 < q < 1 \leq Q$  such that  $\|C_i\|_{\vartheta} \leq q < 1$  ( $T_i P^{(i)} \in \mathcal{W}_1$ ) and  $1 \leq \|C_i\|_{\vartheta} \leq Q$  ( $T_i P^{(i)} \in \mathcal{W}_2$ ). Also there is a constant  $\gamma > 0$  such that for every  $T_i P^{(i)} \in \mathcal{W}_1 \cup \mathcal{W}_2$ ,  $\|b_i\|_{\vartheta} \leq \gamma$ .

Under Assumption (A) the matrix set  $\mathcal{C} = \{C_i : T_i P^{(i)} \in \mathcal{W}_1\}$  is an RCP set and all infinite products  $\prod_{i=1}^{\infty} C^{(i)}$  ( $C^{(i)} \in \mathcal{C}$ ) converge to the zero matrix. By Lemma 7 the sequences  $B_k = \prod_{i=1}^k T_i P^{(i)}$  and  $S^{(k)} = S^{(0)} B_k$  ( $T_i P^{(i)} \in \mathcal{W}_1$ ) are also convergent. Hence if Assumption (A) holds, then  $\mathcal{W}_1$  is an RCP set. We show in Section 6 and the Appendix that Assumption (A) holds at least for  $1 \leq n \leq 8$ . Note again that matrices of  $\mathcal{W}_1$  correspond to the inside and outside contraction operations, when the incoming vertices are inserted in the first or the second position of the ordering (3) or they correspond to any shrinking operation.

**Theorem 9** Suppose that Assumption (A) is satisfied and  $S^{(0)}$  is nondegenerate. Let  $t_1(k)$  be the number of operations  $T_i P^{(i)}$  that belong to  $\mathcal{W}_1$ , and  $t_2(k)$  be the number of those operations  $T_i P^{(i)}$  that belong to  $\mathcal{W}_2$  during the first  $k$  iterations of the Nelder-Mead method. Also assume that for  $\kappa \in \mathbb{N}$ ,  $q^{1-\kappa} \leq Q \leq q^{-\kappa}$  and for  $\mu \in (0, 1)$ ,  $t_1(k) \geq \mu k + \kappa t_2(k)$  holds ( $k \geq k_0$ ). Then the Nelder-Mead algorithm converges in the sense that

$$\lim_{k \rightarrow \infty} x_j^{(k)} = \hat{x} \quad (j = 1, \dots, n + 1) \tag{34}$$

for some vector  $\hat{x}$  with a convergence speed proportional to  $O(q^{\mu k})$ . If  $f$  is continuous at  $\hat{x}$ , then

$$\lim_{k \rightarrow \infty} f(x_j^{(k)}) = f(\hat{x}) \quad (j = 1, \dots, n + 1) \tag{35}$$

holds as well.

*Proof* We first investigate the product (33). By assumption  $t_1(k)$  is the number of those  $C_i$ 's that satisfies  $\|C_i\| \leq q < 1$  ( $1 \leq i \leq k$ ) and  $t_2(k)$  is the number of those  $C_i$ 's that satisfies  $1 \leq \|C_i\| \leq Q$  ( $1 \leq i \leq k$ ). Clearly,  $0 \leq t_i(k) \leq k$  and  $t_1(k) + t_2(k) = k$ . Then

$$\left\| \prod_{j=1}^k C_j \right\|_{\vartheta} \leq q^{t_1(k)} Q^{t_2(k)} \leq q^{t_1(k) - \kappa t_2(k)} \leq q^{\mu k} = c_k$$

and  $\sum_{i=1}^{\infty} c_k$  is clearly convergent. Hence it follows from Lemma 7 that

$$\lim_{k \rightarrow \infty} L_k = \begin{bmatrix} 1 & 0 \\ \tilde{x} & 0 \end{bmatrix} = \tilde{L}$$

for some vector  $\tilde{x}$ . Since

$$\begin{aligned} \|\tilde{x} - x_k\| &= \left\| \sum_{i=k+1}^{\infty} \left( \prod_{j=1}^{i-1} C_j \right) b_i \right\|_{\vartheta} \leq \gamma \sum_{i=k}^{\infty} q^{\mu i} \leq \Gamma_1 q^{\mu k}, \\ \|L_k - \tilde{L}\|_{\vartheta} &\leq \Gamma_2 q^{\mu k} \end{aligned}$$

holds with a suitable constant  $\Gamma_2 > 0$ . Hence

$$B_k \rightarrow F \begin{bmatrix} 1 & 0 \\ \tilde{x} & 0 \end{bmatrix} F^{-1} = \begin{bmatrix} 1 - e^T \tilde{x} \\ \tilde{x} \end{bmatrix} e^T = w e^T = B \tag{36}$$

and

$$\|B_k - B\|_{\vartheta} \leq \Gamma_2 \text{cond}(F) q^{\mu k}. \tag{37}$$

□

**Corollary 2**  $\text{diam}(S^{(k)}) \rightarrow 0$  ( $k \rightarrow \infty$ ) with a speed of  $O(q^{\mu k})$ .

For higher dimension, we can expect slower convergence, since Lemma 3 implies that  $q$  must approach 1.



Except for Lagarias et al. [23] we do not know that what kind of steps can follow each others when the Nelder-Mead method applied to a function. Under Assumption (A), if the Nelder-Mead steps are taken only from  $\mathcal{W}_1$ , then the algorithm converges in the sense of Theorem 9. Thus the convergence of the Han-Neumann case (18)–(19) also follows. In general the method also takes steps from  $\mathcal{W}_2$ . If it occurs in a finite number of occasions, that is  $t_2(k) \leq k_0$ , then  $t_1(k) \geq k - k_0$  and we can set  $\mu = 1$  in the theorem. If not, we must assume that the elements from  $\mathcal{W}_1$  counterbalance the effect of those from  $\mathcal{W}_2$ . This is provided by the simple assumption  $t_1(k) \geq \mu k + \kappa t_2(k)$ .

The difficulty of Theorem 9 is to find a suitable norm  $\|\cdot\|_\vartheta$  for which Assumption (A) holds. The reason for this is the following.

If all infinite products from a matrix set  $\Sigma$  converge, that is  $\Sigma$  is an RCP set, then  $\Sigma$  is also product bounded (see, e.g., [13]). A set  $\Sigma$  of  $n \times n$  matrices is product bounded if there is a constant  $\beta > 0$  such that  $\|A_1 \cdots A_k\| \leq \beta$  for all  $k$  and all  $A_1, \dots, A_k \in \Sigma$ . A matrix set  $\Sigma$  is product bounded if and only if there exists a multiplicative matrix norm  $\|\cdot\|$  such that  $\|A\| \leq 1$  for all  $A \in \Sigma$  (see, e.g., [4, 13]).

If  $\mathcal{W}_1$  is an RCP set, then it is also product bounded. Blondel and Tsitsiklis [5] proved that the product boundedness of a finite matrix set  $\Sigma$  is algorithmically undecidable and it remains undecidable even in the special case, when  $\Sigma$  consists of only two matrices. Since product boundedness is a weaker property than the RCP, and yet it is algorithmically undecidable, it seems difficult to decide the RCP property in general. In Section 6 we present a technique that circumvents this problem for  $\mathcal{W}_1$  at least for  $n \leq 8$ .

## 6 The convergence in low dimensions

Here we show that Assumption (A) holds for  $n = 1, 2, \dots, 8$ , which implies that Theorem 9 also holds for  $n = 1, 2, \dots, 8$ . Case  $n = 1$  is simple, but for cases  $n = 2, \dots, 8$ , we have to construct induced matrix norms that satisfy (A).

### 6.1 Convergence for $n = 1$

In this case

$$\mathcal{W}_1 = \left\{ T \left( \frac{1}{2} \right) P_j, T \left( -\frac{1}{2} \right) P_j : j = 1, 2 \right\}$$

and

$$\mathcal{W}_2 = \{ T(1) P_2, T(2) P_2, T(1) P_1 \}.$$

Lemma 6 implies

$$F^{-1} B_k F = \prod_{i=1}^k \begin{bmatrix} 1 & 0 \\ b_i & c_i \end{bmatrix},$$

where  $|c_i| = \frac{1}{2} = q$  ( $T_i P^{(i)} \in \mathcal{W}_1$ ) and  $1 \leq |c_i| \leq 2 = Q$  ( $T_i P^{(i)} \in \mathcal{W}_2$ ). Here the norm  $\|\cdot\|_w = |\cdot|$  and the convergence Theorem 9 holds with  $\kappa = 1$ .

### 6.2 Convergence for $2 \leq n \leq 8$

Using the result of Stein [40], Householder [17] proved that for each matrix  $A \in \mathbb{R}^{m \times m}$  with  $\rho(A) < 1$ , there is a matrix  $R$  such that  $\|RAR^{-1}\|_2 < 1$ . A related result is given by Deutsch [7].

Here we need a matrix  $S$  such that  $\|C_i\|_w = \|S^{-1}C_iS\|_2 < 1$  holds for all matrices  $T_iP^{(i)} \in \mathcal{W}_1$ . A simultaneous diagonalization of these matrices clearly would do it. However the matrices of  $\mathcal{W}_1$  are not pairwise commuting and so they are not simultaneously diagonalizable (see, e.g., [9, 34, 37]). Hence we tried to solve the optimization problem

$$\min_S \max \left\{ \|S^{-1}C_iS\|_2 : T_iP^{(i)} \in \left\{ T\left(\frac{1}{2}\right)P_j, T\left(-\frac{1}{2}\right)P_j : j = 1, 2 \right\} \right\}. \tag{38}$$

using the standard matrix routines of Matlab R2013b and the ‘fminsearch’ (Nelder-Mead) algorithm starting from several initial points. All numerical computations were done on a PC with Intel i7-8700 CPU @ 3.20GHz and Windows 10 operating system. The numerical results are presented in this section and the Appendix are given in Matlab’s short format (scaled fixed point format with 5 digits).

Since the number of possible  $(n + 1) \times (n + 1)$  matrices  $T_iP^{(i)}$  is  $N = 3n + 3 + (n + 1)!$  we only present the following computed quantities

$$\rho_1 = \max \left\{ \rho(C_i) : T_iP^{(i)} \in \mathcal{W}_1 \right\}, \quad \nu_1 = \max \left\{ \|C_i\|_2 : T_iP^{(i)} \in \mathcal{W}_1 \right\},$$

$$\rho_2 = \max \left\{ \rho(C_i) : T_iP^{(i)} \in \mathcal{W}_2 \right\}, \quad \nu_2 = \max \left\{ \|C_i\|_2 : T_iP^{(i)} \in \mathcal{W}_2 \right\},$$

$$q = \max \left\{ \|S^{-1}C_iS\|_2 : T_iP^{(i)} \in \mathcal{W}_1 \right\}, \quad Q = \max \left\{ \|S^{-1}C_iS\|_2 : T_iP^{(i)} \in \mathcal{W}_2 \right\}.$$

The computed matrices  $S$  will be presented in the Appendix for  $2 \leq n \leq 8$ .

$n$	$T_iP^{(i)} \in \mathcal{W}_1$			$T_iP^{(i)} \in \mathcal{W}_2$			$\kappa$
	$\rho_1$	$\nu_1$	$q$	$\rho_2$	$\nu_2$	$Q$	
2	0.8431	1.2892	0.8431	1.6861	3.1787	2.6568	6
3	0.9275	1.2622	0.9275	1.5214	3.8378	2.7437	14
4	0.9587	1.2271	0.9590	1.4201	4.3635	2.7298	24
5	0.9735	1.2171	0.9735	1.3517	4.8195	2.7560	38
6	0.9815	1.3155	0.9836	1.3024	5.2301	3.0442	68
7	0.9864	1.4075	0.9885	1.2652	5.6075	3.0901	98
8	0.9896	1.4939	0.9913	1.2361	5.9592	3.1962	133

The data of the table implies that Assumption (A) holds and so Theorem 9 implies the convergence of the Nelder-Mead method for  $n = 2, \dots, 8$ . Note that  $\rho_1$  and  $\rho_2$  approach 1 as indicated by Lemmas 3 and 4.

If we exclude the expansion operations  $(T(2)P_1, T(1)P_1)$ , we have smaller  $\kappa$  values as shown in the table below

$n$	2	3	4	5	6	7	8
$q$	0.8431	0.9275	0.9590	0.9735	0.9836	0.9885	0.9913
$Q'$	1.3520	1.3918	1.3858	1.3889	1.5442	1.5545	1.6144
$\kappa'$	2	5	8	13	27	39	56

Consequently we have a faster convergence speed, although the estimated speed still slows down for increasing  $n$ .

### 7 Summary

We analyzed the Nelder-Mead algorithm in the iterative form

$$S^{(k)} = S^{(k-1)}T_kP^{(k)} = S^{(0)} \prod_{i=1}^k T_iP^{(i)},$$

where  $T_iP^{(i)} \in \mathcal{T}$  is the matrix of the executed inner step at iteration  $i$ . Since the convergence of the sequence  $\{S^{(k)}\}$  clearly depends on the convergence of the infinite matrix products  $\prod_{i=1}^\infty T_iP^{(i)}$ , we used techniques from the theory of infinite matrix products [13]. First, we investigated the spectra of the matrices  $T_iP^{(i)}$ , then using a simultaneous similarity reduction on  $\mathcal{T}$  to block lower triangular matrices we proved a convergence result (Theorem 9) for the simplex sequence  $\{S^{(k)}\}$  to rank-one matrices of the form  $\widehat{x}e^T$  for some vector  $\widehat{x}$ . This implies the convergence  $f_i^{(k)} \rightarrow f(\widehat{x})$ . The examples of Section 4 support the study of this type of convergence. The main idea of the convergence theorem is to identify a subset  $\mathcal{W}_1$  of operations  $T$  that has the RCP property. This property follows from Assumption (A) which is proved for  $1 \leq n \leq 8$  using numerical optimization. It is not yet known if this bound can be increased in the same way or not. Theorem 9 has a deficiency that  $\widehat{x}$  is not related to any stationary point of  $f$  (for a similar result, see also Lagarias et al. [23]). The results of Kelley [20], Lagarias et al. [22] and also the examples of Section 4 indicate that new techniques are to be developed for such results.

### Appendix

Here we present the similarity transformation matrices  $S$  (the approximate solutions of optimization problem (38)) for  $n = 2, \dots, 8$ . The results are given in Matlab's short format (scaled fixed point format with 5 digits).

- $n = 2$

$$\begin{bmatrix} 1.9750 & -0.8063 \\ -0.0000 & 1.6126 \end{bmatrix}$$

2.  $n = 3$ 

$$\begin{bmatrix} -0.6128 & 1.5599 & 0.3998 \\ 1.4872 & -0.0621 & 0.4384 \\ -0.6508 & -0.8054 & 0.7748 \end{bmatrix}$$

3.  $n = 4$ 

$$\begin{bmatrix} 0.5931 & 0.6016 & 0.8564 & 0.0438 \\ 0.2918 & -0.7842 & 0.0464 & -0.7280 \\ 0.3339 & 0.4243 & -0.8902 & -0.0357 \\ 0.0252 & -0.4527 & -0.0132 & 0.8120 \end{bmatrix}$$

4.  $n = 5$ 

$$\begin{bmatrix} -2.9412 & 0.6476 & -1.0913 & -1.3439 & -0.0418 \\ -0.1282 & 0.0171 & 0.1649 & 2.9236 & -1.5885 \\ 0.3674 & 1.3641 & 2.3126 & -0.9888 & -0.4276 \\ 0.0280 & -2.6337 & 0.7949 & -0.4218 & -0.0538 \\ 0.0278 & 0.3617 & -0.0620 & 0.8253 & 2.4192 \end{bmatrix}$$

5.  $n = 6$ 

$$\begin{bmatrix} 0.3119 & 1.0592 & -0.1946 & -0.5866 & -0.3932 & 0.0921 \\ 0.4206 & 0.3139 & 0.5872 & 0.6986 & 0.3377 & 0.3883 \\ 0.2407 & -0.3890 & 0.6421 & -0.6398 & 0.1694 & -0.4387 \\ 0.2614 & -0.5562 & -0.0797 & 0.3364 & -0.8387 & -0.0254 \\ 0.2237 & -0.4478 & -0.6062 & -0.2513 & 0.4292 & 0.4870 \\ -0.0146 & 0.0967 & -0.4087 & 0.3929 & 0.3054 & -0.7185 \end{bmatrix}$$

6.  $n = 7$ 

$$\begin{bmatrix} -0.3626 & -0.8852 & 0.2800 & -0.5387 & 0.2400 & -0.4203 & -0.0043 \\ -0.0908 & -0.5032 & -0.4429 & 0.8128 & -0.1211 & 0.1963 & 0.4403 \\ -0.3671 & 0.2001 & -0.7056 & -0.5943 & -0.2162 & 0.4116 & -0.1651 \\ -0.3000 & 0.4885 & -0.1812 & 0.3925 & 0.6319 & -0.4248 & -0.3131 \\ -0.2462 & 0.4991 & 0.3233 & -0.0929 & -0.4733 & -0.3885 & 0.5635 \\ -0.0391 & 0.1224 & 0.5386 & 0.0662 & 0.3831 & 0.7052 & 0.1888 \\ 0.0088 & -0.1091 & 0.2667 & 0.1922 & -0.5099 & 0.0707 & -0.6996 \end{bmatrix}$$

7.  $n = 8$ 

$$\begin{bmatrix} -0.1239 & -0.2497 & -0.9519 & -0.0070 & 0.5374 & 0.3386 & 0.1665 & -0.4338 \\ -0.2209 & 0.3785 & -0.5796 & 0.4922 & -0.1801 & -0.4578 & 0.1531 & 0.4066 \\ -0.2539 & 0.6514 & -0.0050 & -0.1626 & -0.4501 & 0.4870 & -0.3204 & -0.2528 \\ -0.2075 & 0.4321 & 0.4152 & -0.4207 & 0.3413 & -0.4255 & 0.4287 & -0.2011 \\ -0.2447 & -0.0032 & 0.4940 & 0.2731 & 0.5119 & 0.4115 & -0.1839 & 0.4731 \\ -0.2273 & -0.4122 & 0.4021 & 0.4897 & -0.1784 & -0.2661 & -0.1448 & -0.5271 \\ -0.1404 & -0.4743 & 0.1264 & -0.2262 & -0.4800 & 0.2640 & 0.5346 & 0.2787 \\ -0.0623 & -0.3054 & -0.1097 & -0.5037 & -0.0189 & -0.3532 & -0.5928 & 0.2230 \end{bmatrix}$$

**Acknowledgements** The author is indebted to the unknown referee whose remarks and suggestions improved the paper.

**Funding** Open access funding provided by Öbuda University.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Alzer, H.: Refinement of an Inequality of E. Landau. *Acta Math. Univ. Comenianae.* **62**(1), 13–16 (1993)
2. Audet, C., Hare, W.: *Derivative-free and Blackbox Optimization*. Springer. <https://doi.org/10.1007/978-3-319-68913-5> (2017)
3. Barnett, S.: *Polynomials and Linear Control Systems*. Marcel Dekker Inc. (1983)
4. Beyn, W.-J., Elsner, L.: Infinite products and paracontracting matrices. *Electron. J. Linear Al.* **2**, 1–8 (1997)
5. Blondel, V.D., Tsitsiklis, J.N.: The boundedness of all products of a pair of matrices is undecidable. *Syst. Control Lett.* **41**, 135–140 (2000)
6. Conn, A.R., Scheinberg, K., Vicente, L.N.: *Introduction to Derivative-Free optimizations*. SIAM. <https://doi.org/10.1137/1.9780898718768> (2009)
7. Deutsch, E.: On matrix norms and logarithmic norms. *Numer. Math.* **24**, 49–51 (1975)
8. Eneström, G.: Remarque sur un theoreme relatif aux racines de l'equation  $a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$  ou tous les coefficients a sont réels at positifs. *Tohoku. Math. J.* **18**, 34–36 (1920)
9. Friedland, S.: *Matrices: algebra, analysis, an applications*. World Scientific Publishing (2015)
10. Galántai, A.: Convergence theorems for the Nelder-Mead method. *J. Comput. Appl. Mechanics.* **15**(2), 115–133 (2020). <https://doi.org/10.32973/jcam.2020.008>
11. Han, L., Neumann, M., Xu, J.: On the roots of certain polynomials arising from the analysis of the Nelder-Mead simplex method. *Linear Algebra Appl.* **363**, 109–124 (2003). [https://doi.org/10.1016/S0024-3795\(02\)00485-8](https://doi.org/10.1016/S0024-3795(02)00485-8)
12. Han, L., Neumann, M.: Effect of dimensionality on the Nelder-Mead simplex method. *Optimization Methods and Software* **21**(1), 1–16 (2006). <https://doi.org/10.1080/10556780512331318290>
13. Hartfiel, D.J.: *Nonhomogeneous matrix products*. World Scientific. <https://doi.org/10.1142/4707> (2002)
14. Henrici, P.: *Applied and computational complex analysis*. Volume I John Wiley and Sons (1974)
15. Hocking, J.G., Young, G.S.: *Topology*. Addison–Wesley (1961)
16. Horn, R.A., Johnson, C.R.: *Matrix analysis*. Cambridge University Press. <https://doi.org/10.1017/9781139020411> (2013)
17. Householder, A.S.: *The theory of matrices in numerical analysis*. Blaisdell (1964)
18. Kakeya, S.: On the limits of the roots of an algebraic equation with positive coefficients. *Tohoku Math. J.* **2**, 140–142 (1912)
19. Kelley, C.T.: *Iterative methods for optimization*. SIAM. <https://doi.org/10.1137/1.9781611970920> (1999)
20. Kelley, C.T.: Detection and remediation of stagnation in the Nelder-Mead algorithm using an sufficient decrease condition. *SIAM J. Optimiz.* **10**, 33–45 (1999). <https://doi.org/10.1137/S1052623497315203>
21. Kolda, T.G., Lewis, R.M., Torczon, V.: Optimization by direct search: new perspectives on some classical and modern methods. *SIAM Rev.* **45**(3), 385–482 (2003). <https://doi.org/10.1137/S003614450242889>
22. Lagarias, J.C., Poonen, B., Wright, M.H.: Convergence of the restricted Nelder-Mead algorithm in two dimensions. *SIAM J. Optimiz.* **22**(2), 501–532 (2012). <https://doi.org/10.1137/110830150>
23. Lagarias, J.C., Reeds, J.A., Wright, M.H., Wright, P.E.: Convergence properties of the Nelder-Mead simplex method in low dimensions. *SIAM J. Optimiz.* **9**(1), 112–147 (1998). <https://doi.org/10.1137/S1052623496303470>

24. Lakatos, P., Losonczy, L.: Self-inversive polynomials whose zeros are on the unit circle. *Publ. Math. Debrecen*. **65**(3–4), 409–420 (2004)
25. Landau, E.: Über eine Aufgabe aus der Funktionentheorie. *Tohoku Math. J.* **5**, 97–116 (1914)
26. Langville, A.N., Moxley, C.D.: Deeper inside PageRank. *Internet Mathematics Journal* **1**(3), 335–80 (2005). <https://doi.org/10.1080/15427951.2004.10129091>
27. Langville, A.N., Moxley, C.D.: *Google's PageRank and beyond: The Science of Search Engine Rankings*. Princeton University Press, Princeton (2006)
28. Lewis, R.M., Torczon, V., Trosset, M.W.: Direct search methods: then and now. *J. Comput. Appl. Math.* **124**, 191–207 (2000). [https://doi.org/10.1016/S0377-0427\(00\)00423-4](https://doi.org/10.1016/S0377-0427(00)00423-4)
29. McKinnon, K.I.M.: Convergence of the Nelder-Mead simplex method to a nonstationary point. *SIAM J. Optimiz.* **9**(1), 148–158 (1998). <https://doi.org/10.1137/S1052623496303482>
30. Marden, M.: *Geometry of Polynomials*. American Mathematical Society, Providence, Rhode Island (1966)
31. Matousek, J.: *Using the Borsuk-Ulam theorem*. Springer. <https://doi.org/10.1007/978-3-540-76649-0> (2008)
32. Meyer, C.D.: *Matrix analysis and applied linear algebra*. SIAM (2000)
33. Milovanovic, G.V., Rassias, T.M.: Inequalities for polynomial zeros. In: Rassias, T.M. (ed.) *Survey on Classical Inequalities*, pp. 165–202. Kluwer Academic Publishers (2000). <https://doi.org/10.1007/978-94-011-4339-4>
34. Mirsky, L.: *An Introduction to Linear Algebra*. Oxford University Press, Oxford (1955)
35. Nelder, J.A., Mead, R.: A simplex method for function minimization. *Comput. J* **7**, 308–313 (1965). <https://doi.org/10.1093/comjnl/7.4.308>
36. Oldenburger, R.: Infinite powers of matrices and characteristic roots. *Duke Math. J* **6**, 357–361 (1940). <https://doi.org/10.1215/S0012-7094-40-00627-5>
37. O'Meara, K.C., Clark, J., Vinsonhaler, C.L.: *Advanced topics in linear algebra : weaving matrix problems through the weyr form*. Oxford University Press, Oxford (2011)
38. Rahman, Q.I., Schmeisser, G.: *Analytic theory of polynomials*. Clarendon Press (2002)
39. Sheil-Small, T.: *Complex polynomials*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511543074> (2002)
40. Stein, P.: Some general theorems on iterants. *J. Res. Nat. Bur. Stand.* **48**(1), 82–83 (1952)
41. Trench, W.F.: Characterization and properties of matrices with  $k$ -involutory symmetries. *Linear Algebra Appl.* **429**, 2278–2290 (2008). <https://doi.org/10.1016/j.laa.2008.07.002>
42. Walters, F.H., Morgan, S.L., Parker, L.R., Deming, S.N.: *Sequential simplex optimization*. CRC Press LLC (1991)
43. Wright, M.H.: Direct search methods: once scorned, now respectable. In: Griffiths, D.F., Watson, G.A. (eds.) *Numerical Analysis 1995 (Proceedings of the 1995 Dundee Biennial Conference in Numerical Analysis)*, pp. 191–208. Addison-Wesley Longman, Harlow, United Kingdom (1996)
44. Wright, M.H.: Nelder, Mead, and the other simplex method. *Documenta Mathematica. Extra Volume: Optimization Stories*, pp. 271–276 (2012)
45. Young, D.M.: *Iterative solution of large linear systems*. Academic Press Inc (1971)

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.