Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Nelder--Mead Method in High Dimensions: An Exact Moment-of-Inertia Identity and Dimension-Free almost sure Convergence

Martin Gugat

martin.gugat@fau.de

Friedrich-Alexander-Universitat Erlangen-Nurnberg https://orcid.org/0000-0002-5281-110X

Research Article

Keywords: Nelder–Mead simplex method, almost sure convergence, high dimensions, dimension-dependent coefficients, Sufficient decrease and oriented restart

Posted Date: November 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11605804/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

The Nelder–Mead Method in High Dimensions: An Exact Moment-of-Inertia Identity and Dimension-Free almost sure Convergence

Martin Gugat

Friedrich-Alexander-Universität Erlangen-Nürnberg, Chair for Dynamics, Control, Machine Learning and Numerics (Alexander von Humboldt Professorship), Cauerstr. 11, 91058 Erlangen, Germany, martin.gugat@fau.de

Abstract

Galántai (2023) proved that the vertices of the Nelder–Mead simplices in $\mathbb{R}^n$ converge almost surely to a common limit in a random model, but his criterion requires all $|\mathrm{T}|^{\ell}$ products of transformation matrices, $|\mathrm{T}| = 3n + 3 + (n + 1)!$, and was verified only for $n \leq 6$. We derive an exact identity for the moment of inertia Φ_k of the simplex. This yields a deterministic convergence criterion. For probabilistic models of the iteration, we show that Galántai's theorem holds in every dimension with blocks of length one, but for $n \geq 5$ its uniform model consists almost entirely of shrinks, which never occur for strictly convex functions. If the replaced vertex has uniformly distributed rank, the vertices converge almost surely to a common limit in every dimension whenever contractions are more than four times as frequent as expansions (standard coefficients), with expected rate $1 - c/n$ per iteration. For Galántai's insertion model, almost sure convergence follows from a spectral-radius condition, which we prove in every dimension with the moment of inertia as a Lyapunov function. The deterministic criterion and the result for uniformly distributed ranks extend to the original, unordered method. We combine the reflection and contraction coefficients of Puhan and Bűrmen (2026), an expansion coefficient adapted to the share of Φ_k carried by the replaced vertex, and Kelley's sufficient decrease safeguard with its constant divided by $(n + 1)^2$. We give a sufficient condition for the convergence of the resulting algorithm to a critical point. In tests up to $n = 300$, it required fewer function evaluations than the method of Gao and Han in all cases except a test example with a neural network that is motivated by machine learning, and fewer than the coefficients of Puhan and Bűrmen in eleven of the fifteen other cases.

Keywords: Nelder–Mead simplex method; almost sure convergence; high dimensions; dimension-dependent coefficients; Sufficient decrease and oriented restart
MSC: 65K10; 90C56; 60F15

1 Introduction

The Nelder–Mead (NM) simplex method [14] is one of the most widely used derivative-free methods for the minimization of a continuous function $f : \mathbb{R}^n \to \mathbb{R}$; see [12] for a survey. Its convergence theory is notoriously incomplete. McKinnon [13] constructed strictly convex functions with continuous derivatives on which the simplices converge to a nonstationary point. For strictly convex functions of two variables with bounded level sets, Lagarias, Reeds, Wright and Wright [11] proved that the function values at all vertices converge to the same value and that the diameters of the simplices tend to zero, and Lagarias, Poonen and Wright [10] proved convergence to the minimizer for a restricted version of the method. Galántai [6] showed that, for the convergence of the function values, convexity and boundedness from below suffice in two

dimensions, both for the version of [11] and for the original method of Nelder and Mead. Kelley [9] proposed a sufficient decrease condition for the average function value at the vertices. He proved that, for functions with Lipschitz continuous gradient that are bounded below, every accumulation point of the simplices is a critical point if this condition holds for all but finitely many iterations and if the product of the largest edge length at the best vertex and the condition number of the simplex directions tends to zero. If the condition fails, he replaces the simplex by a smaller one with orthogonal edges, oriented by the signs of a simplex gradient (oriented restart). Zhao, Uroševi ć, Mladenović and Hansen [19] replaced the shrink step by a step that moves the vertices toward the best vertex one at a time until an improvement is obtained, and they restarted this modified method from a simplex of the initial size around the best point found until no further improvement occurred. In their experiments with up to 100 variables, the restarted method reached global or local minima much more often than the original method. In Section 7, we discuss which of our convergence results apply to their method. Han and Neumann [8] showed for $f(x) = x^T x$, with the minimizer as one of the initial vertices, that the oriented lengths of the simplices converge linearly with a rate constant that rapidly approaches 1 as the dimension n increases. Gao and Han [7] proved that, for uniformly convex functions with modulus ρ, every expansion and contraction step decreases the sum of the function values at the vertices by at least $\frac{n-1}{2n^2}\rho(\frac{1}{2}\text{diam})$. They observed that the standard method uses an overwhelming fraction of reflection steps in high dimensions, and they proposed dimension-dependent coefficients that substantially improve its performance for large n. Puhan and Bűrmen [16] derived dimension-dependent coefficients from a model of a single step on strictly convex quadratics, in which an expected improvement of f is maximized under an assumption on the distribution of the centroid. Their coefficients follow the pattern of the heuristic schemes of [7] and others: the reflection and expansion coefficients decrease, and the inside contraction and shrink coefficients increase with n. In their experiments with quadratics in up to 100 dimensions, they outperformed the standard coefficients.

In a series of papers [2, 3, 4, 5, 6], Galántai studied the question whether the vertices of the simplices $S^{(k)}$ that are generated by the NM iteration converge to a common limit point. Writing the iteration as $S^{(k+1)} = S^{(k)} T_k P_k$ with transformation matrices T_k and permutation matrices P_k, this becomes a question about infinite products of matrices. In [5], two results were obtained. Theorem 1 of [5] proves convergence under a density condition on blocks of ℓ consecutive operations. Theorem 2 of [5] proves convergence with probability one in a random model in which these blocks are independent and identically distributed with $\mu = \mathrm{E} \| Y \|_2 < 1$, where Y is the corresponding product of $n \times n$ matrices. The verification of the hypotheses requires the norms of all $|\mathrm{T}|^\ell$ products of length ℓ, where $|\mathrm{T}| = 3n + 3 + (n+1)!$ is the number of possible operations, and was therefore carried out for $n \leq 6$. In [6], the ordered version of [11] is compared with the original, unordered method of Nelder and Mead. Theorem 1 of [5] is extended to both versions [6, Theorem 5], and for $n \leq 4$ and $\ell \leq 7$ the fractions of products of length ℓ whose matrix C (Section 2) has spectral norm at most 0.99 are computed, with and without shrink steps. Since these fractions are larger for the ordered version, Galántai concludes that the ordering of [11] improves the convergence behaviour significantly, and he conjectures that the fractions come close to one for large ℓ [6, Section 7].

In this paper we remove the dimensional restriction. Our starting point is the observation that in centered coordinates, i.e. after subtracting the centroid of the vertices, every non-shrink operation acts as a symmetric rank-one modification of the identity followed by an orthogonal permutation, and every shrink acts as a multiple of an orthogonal permutation (Proposition 3.4). In this form of the iteration, in a permutation-invariant norm, all operations have explicit norms:

- 1 for reflections and contractions,
- χ for expansions and
- σ for shrinks,

independently of n, of the reordering and of the version of the method (Corollary 3.6, Proposition 7.1). For a simplex $S = [x_1, \ldots, x_{n+1}] \in \mathbb{R}^{n\times(n+1)}$, define the centroid of all vertices

$$\bar{x} = \frac{1}{n+1} \sum_{i=1}^{n+1} x_i,$$

the centered matrix and the moment of inertia

$$D(S) = [x_1 - \bar{x}, \ldots, x_{n+1} - \bar{x}], \qquad \Phi(S) = \|D(S)\|_F^2 = \sum_{i=1}^{n+1} \|x_i - \bar{x}\|^2.$$

Then the moment of inertia Φ of the simplex satisfies the exact identity (2) from Theorem 4.1, which allows Φ_k to be updated during the iteration. This yields a dimension-free deterministic convergence criterion that can be monitored at cost $O(n)$ per iteration (see Theorem 4.5).

For random models, which are idealizations of the deterministic algorithm and ignore the dependence of the operations on f and on the geometry of the simplex, this implies that Theorem 2 of [5] on convergence with probability one holds in every dimension $n \geq 2$ with blocks of length one (Corollary 5.3). For $n \geq 5$, however, the uniform model of [5] consists almost entirely of shrink steps (Remark 5.4), and shrinks never occur for strictly convex functions [11, 6].

With a second probabilistic model, in which a uniform distribution is assumed for the choice of the exchanged vertices in addition to the distribution of the operations, we obtain almost sure convergence of the vertices also without shrinks, for the standard coefficients under a condition on the frequencies of the operations that does not depend on n (Theorem 6.2, Corollary 6.3). This result is applicable to the original method of Nelder and Mead [14], see Section 7, Corollary 7.2.

A third model, the insertion model of [5] without shrinks, assumes that the pairs $T_k P_k$ are chosen independently and uniformly from the non-shrink operations. Then the criterion for almost sure convergence is a spectral-radius condition (Theorem 6.6), which we prove in every dimension, using the moment of inertia as a Lyapunov function (Theorem 6.7); Table 1 gives the spectral radius for a selection of values $n \leq 500$.

Kelley's sufficient decrease condition [9] is a suitable tool to detect stagnation and, under additional conditions, to guarantee that the accumulation points of the simplices are stationary. However, in its original form it becomes overly strict in high dimensions, since one iteration changes a single vertex (Lemma 8.2). We introduce dimension-dependent forms. Dividing the constant in Kelley's sufficient decrease condition by $(n + 1)^2$ removed the false alarms in our tests. Combined with the reflection and contraction coefficients of [16] and an expansion coefficient adapted to the share of Φ_k carried by the replaced vertex, this defines our Algorithm 1. It converges to a critical point under conditions that can be checked along a run (Theorem 8.3). In numerical tests, Algorithm 1 reached the minimizer on McKinnon's two smooth stagnation examples and detected the stagnation on the nonsmooth one. It required fewer function evaluations than the method of [7] in all fifteen test cases of Section 8.4 and than the coefficients of [16] alone in eleven, and it came close to the target on an ill-conditioned problem with $n = 300$, on which both other methods remained far from it (Section 8.4). With the contraction coefficient of [7] instead, it was more efficient on this problem for $40 \leq n \leq 150$, but less efficient in the other cases of Section 8.4.

As in [5], we note that the convergence of the vertices to a common limit does not imply that the limit is a stationary point [13, 4]; this is the reason for the safeguard in Algorithm 1.

2 The Nelder–Mead method in matrix form

We use the version of [11], as in [5]. The simplex of iteration k is $S^{(k)} = [x_1^{(k)}, \ldots, x_{n+1}^{(k)}] \in \mathbb{R}^{n\times(n+1)}$, and the vertices are ordered so that $f(x_1^{(k)}) \leq \cdots \leq f(x_{n+1}^{(k)})$. We omit the superscript when no confusion can arise and write

$$x_c = \frac{1}{n} \sum_{i=1}^{n} x_i, \qquad X(\lambda) = (1 + \lambda)x_c - \lambda x_{n+1} \quad (\lambda \in \mathbb{R}) \tag{1}$$

for the centroid of the n best vertices and the points on the line through x_{n+1} and x_c. We allow general coefficients: expansion $\chi > 1$, contraction $\gamma \in (0, 1)$ and shrink $\sigma \in (0, 1)$; the standard choice is $\chi = 2$, $\gamma = \sigma = \frac{1}{2}$ One iteration consists of the ordering and exactly one of the following operations:

- *reflection* (type r): $X(1)$ replaces x_{n+1}; this includes the case in which $X(1)$ is accepted in the expansion branch;
- *expansion* (type e): $X(\chi)$ replaces x_{n+1};
- *outside contraction* (type oc): $X(\gamma)$ replaces x_{n+1};
- *inside contraction* (type ic): $X(-\gamma)$ replaces x_{n+1};
- *shrink* (type s): x_i is replaced by $x_1 + \sigma(x_i - x_1)$ for $i = 2, \ldots, n + 1$.

The new vertex is inserted into the ordering according to the rules of [11] (rule (2) in [5]).

A non-shrink step of type τ thus has the coefficient α_τ, where

$$\alpha_{\mathrm{r}} = 1, \qquad \alpha_{\mathrm{e}} = \chi, \qquad \alpha_{\mathrm{oc}} = \gamma, \qquad \alpha_{\mathrm{ic}} = -\gamma.$$

One NM step is defined as follows: Let the vertices be sorted, let $f_i = f(x_i)$, and let α_{r}, α_{e}, α_{oc} and α_{ic} be the coefficients of the four non-shrink step types. It computes $f_{\mathrm{r}} = f(X(\alpha_{\mathrm{r}}))$ and determines an accepted point v:

- if $f_1 \le f_{\mathrm{r}} < f_n$, then $v = X(\alpha_{\mathrm{r}})$ (reflection);
- if $f_{\mathrm{r}} < f_1$, then $f_{\mathrm{e}} = f(X(\alpha_{\mathrm{e}}))$ is computed, and $v = X(\alpha_{\mathrm{e}})$ if $f_{\mathrm{e}} < f_{\mathrm{r}}$ (expansion), otherwise $v = X(\alpha_{\mathrm{r}})$ (reflection);
- if $f_n \le f_{\mathrm{r}} < f_{n+1}$, then $f_{\mathrm{oc}} = f(X(\alpha_{\mathrm{oc}}))$ is computed, and $v = X(\alpha_{\mathrm{oc}})$ if $f_{\mathrm{oc}} \le f_{\mathrm{r}}$ (outside contraction);
- if $f_{\mathrm{r}} \ge f_{n+1}$, then $f_{\mathrm{ic}} = f(X(\alpha_{\mathrm{ic}}))$ is computed, and $v = X(\alpha_{\mathrm{ic}})$ if $f_{\mathrm{ic}} < f_{n+1}$ (inside contraction).

If no point is accepted in the last two cases, the NM step calls for a shrink; otherwise, v replaces x_{n+1}. With the coefficients $\alpha_{\mathrm{r}} = 1$, $\alpha_{\mathrm{e}} = \chi$, $\alpha_{\mathrm{oc}} = \gamma$ and $\alpha_{\mathrm{ic}} = -\gamma$, and with a shrink whenever the NM step calls for one, this is the method of [11]. Algorithm 1 in Section 8 uses NM steps with other coefficients and replaces the shrink by an oriented restart.

Matrix form. Let $e \in \mathbb{R}^{n+1}$ and $e^- \in \mathbb{R}^n$ denote the vectors of ones and $e_1, \ldots, e_{n+1}$ the unit vectors of $\mathbb{R}^{n+1}$. Following [5], define for $\alpha, \sigma \in \mathbb{R}$ and $j = 1, \ldots, n + 1$

$$T(\alpha) = \begin{matrix} I_n & \frac{1+\alpha}{n} e^- \\ 0 & -\alpha \end{matrix}, \qquad T_{\mathrm{shr}}(\sigma) = \sigma I_{n+1} + (1-\sigma)\, e_1 e^T, \qquad P_j = [e_1, \ldots, e_{j-1}, e_{n+1}, e_j, \ldots, e_n].$$

Then $S\,T(\alpha) = [x_1, \ldots, x_n, X(\alpha)]$, $S\,T_{\mathrm{shr}}(\sigma)$ is the shrunk simplex, and every iteration has the form

$$S^{(k+1)} = S^{(k)}\, T_k\, P_k$$

with $T_k \in \{T(\alpha_{\mathrm{r}}), T(\alpha_{\mathrm{e}}), T(\alpha_{\mathrm{oc}}), T(\alpha_{\mathrm{ic}}), T_{\mathrm{shr}}(\sigma)\}$ and a permutation matrix P_k describing the re-ordering. All these matrices have column sums one, i.e. $\bar{e}\, T = e^T$. For the standard coefficients, the set of all possible pairs is [5, (6)]

$$\mathrm{T} = \{T(\tfrac{1}{2})P_j,\ T(-\tfrac{1}{2})P_j : j \le n + 1\} \cup \{T(1)P_j : j \le n\}$$

$$\cup \{T(2)P_1\} \cup \{T_{\text{shr}}(\tfrac{1}{2})P : P \text{ permutation}\},$$

with $|\mathrm{T}| = 3n + 3 + (n+1)!$. We write $\tilde{\mathrm{T}}$ for the subset of the $3n+3$ non-shrink elements.

With $F = \begin{smallmatrix} 1 & -e^T \\ 0 & I_n \end{smallmatrix}$, every $TP \in \mathrm{T}$ satisfies $F^{-1}TPF = \begin{smallmatrix} 1 & 0 \\ b & C \end{smallmatrix}$ with $C \in \mathbb{R}^{n\times n}$ [5, Lemma 1], [6, Lemma 15], and the convergence of the vertices to a common limit follows if the products of the matrices C tend to zero summably [5, Lemma 2].

Theorem 2 of [5] states that if, for fixed $\ell \geq 2$, the products Y of ℓ consecutive matrices C are chosen independently with probabilities p_Y and $\mu = \sum_Y p_Y \|Y\|_2 < 1$, then the vertices converge to a common limit with probability one. For the uniform distribution, $\mu < 1$ was verified in [5, (23)] for $n \leq 6$ by computing all $|\mathrm{T}|^\ell$ products.

3 Centered coordinates

Definition 3.1. Let $\Pi = I_{n+1} - \frac{1}{n+1}ee^T$ be the orthogonal projector onto

$$W = \{y \in \mathbb{R}^{n+1} : e^T y = 0\}, \quad \text{and let } W^T = \{r \in \mathbb{R}^{1\times(n+1)} : re = 0\}.$$

For a simplex S define the centered matrix and the moment of inertia

$$D(S) = S\,\Pi = [x_1 - \bar{x}, \ldots, x_{n+1} - \bar{x}], \qquad \Phi(S) = \|D(S)\|_F^2 = \sum_{i=1}^{n+1} \|x_i - \bar{x}\|^2.$$

Here $\bar{x} = \frac{1}{n+1}\sum_{i=1}^{n+1} x_i$ is the centroid of all vertices. For the iteration, we write $D_k = D(S^{(k)})$ and $\Phi_k = \Phi(S^{(k)})$.

Lemma 3.2 states properties of the moment of inertia $\Phi(S)$ and the centered simplex $D(S)$.

Lemma 3.2. *(a)* $\Phi(S) = \frac{1}{n+1}\sum_{i<j} \|x_i - x_j\|^2$.

(b) $\frac{1}{2}\operatorname{diam}(S)^2 \leq \Phi(S) \leq \frac{n}{2}\operatorname{diam}(S)^2$.

(c) $\Phi(SP) = \Phi(S)$ *for every permutation matrix* P.

(d) The rows of $D(S)$ *belong to* W^T.

Proof. Part (a) follows using Lagrange's identity componentwise. (b) For all $i \neq j$, we have $\|x_i - x_j\|^2 \leq 2\|x_i - \bar{x}\|^2 + 2\|x_j - \bar{x}\|^2 \leq 2\,\Phi(S)$. The upper bound follows from (a), since the sum has $n(n+1)/2$ terms. (c) $P\,\Pi = \Pi P$ and P is orthogonal. (d) $\Pi e = 0$. □

The next lemma contains further properties of the projection matrix Π.

Lemma 3.3. *Let* $\mathrm{A} = \{A \in \mathbb{R}^{(n+1)\times(n+1)} : e^T A = e^T\}$. *For* $A, B \in \mathrm{A}$ *we have* $AB \in \mathrm{A}$ *and* $\Pi AB\Pi = \Pi A\Pi B\Pi$. *Moreover,* $D(SA) = D(S)\,\Pi A\Pi$ *for every* S *and every* $A \in \mathrm{A}$.

Proof. $\Pi AB\Pi - \Pi A\Pi B\Pi = \Pi A(I - \Pi)B\Pi = \frac{1}{n+1}(\Pi Ae)(e^T B\Pi) = \frac{1}{n+1}(\Pi Ae)(e^T\Pi) = 0$. Similarly, $SA\Pi - S\Pi A\Pi = \frac{1}{n+1}(Se)(e^T A\Pi) = 0$. □

In the next proposition, a centered form of the transformation matrix is presented.

Proposition 3.4 (Centered form)**.** *Let* $w = e_{n+1} - \frac{1}{n+1}e$, *so that* $\|w\|^2 = \frac{n}{n+1}$, *and* $\hat{w} = w/\|w\|$. *For all* $\alpha, \sigma \in \mathbb{R}$ *and every permutation matrix* P,

$$\Pi\,T(\alpha)P\,\Pi = \Pi - (1+\alpha)\,\hat{w}\,\hat{w}^T\,P, \qquad \Pi\,T_{\text{shr}}(\sigma)P\,\Pi = \sigma\,\Pi P.$$

Consequently, $D_{k+1} = D_k\,\Pi - (1+\alpha_k)\,\hat{w}\,\hat{w}^T\,P_k$ *for non-shrink steps and* $D_{k+1} = \sigma D_k P_k$ *for shrink steps. On* W, *the matrix* $\Pi - (1+\alpha)\,\hat{w}\,\hat{w}^T$ *is symmetric with the simple eigenvalue* $-\alpha$ *(eigenvector* $\hat{w}$*) and the eigenvalue* 1 *of multiplicity* $n-1$, *and* P *is orthogonal.*

Proof. Write $T(\alpha) = I + u e_{n+1}^T$ with $u = (1+\alpha)(c - e_{n+1})$, where $c = \frac{1}{n}\sum_{i=1}^{n} e_i$. Since $c - e_{n+1} = \frac{1}{n}e - \frac{n+1}{n}e_{n+1} = -\frac{n+1}{n}w$, we have $u = -(1+\alpha)w/\|w\|^2 \in W$. Hence $\Pi T(\alpha)\Pi = \Pi + u(\Pi e_{n+1})^T = \Pi + u w^T = \Pi - (1+\alpha)\hat{w}\hat{w}^T$. For the shrink, $\Pi T_{\text{shr}}(\sigma)\Pi = \sigma\Pi + (1-\sigma)\Pi e_1 e^T \Pi = \sigma\Pi$, since $e^T\Pi = 0$. Because $P\Pi = \Pi P$, the first two identities follow, and the recursion for D_k follows from Lemma 3.3. The spectral statement is clear, and permutations preserve W and the Euclidean norm. □

Definition 3.5. For $A \in \mathrm{A}$ let $|A| = \max\{\|\Pi A \Pi\| : r \in W^T, \|r\| = 1\}$.

By Lemma 3.3, $|AB| \le |A|\,|B|$, and $\|D(SA)\| \le \|D(S)\|\,|A|$ both for the spectral and the Frobenius norm. The norm $|\cdot|$ is induced by the Euclidean norm on W^T and is therefore invariant under permutations of the vertices.

Now we obtain explicit values of the permutation-invariant operator norms for Nelder–Mead non-shrink and shrink operations.

Corollary 3.6 (Explicit norms)**.** *For every permutation matrix P, we have*

$$|T(\alpha)P| = \max\{1, |\alpha|\}, \qquad |T_{\text{shr}}(\sigma)P| = \sigma.$$

For the standard coefficients, reflections and both contractions have norm 1*, expansions norm* 2 *and shrinks norm* $\frac{1}{2}$ *independently of n and of the reordering.*

The norm $|\cdot|$ is equivalent to the spectral norm of the matrices C used in [5], with a constant that grows like $\sqrt{n+1}$.

Proposition 3.7 (Relation to [5])**.** *Let* $M = \begin{bmatrix} -e^{-T} \\ I_n \end{bmatrix} \in \mathbb{R}^{(n+1)\times n}$ *and* $R = [0 \mid I_n]\,\Pi \in \mathbb{R}^{n\times(n+1)}$*. For every product $A = T_1 P_1 \cdots T_k P_k$ of elements of* T*, the corresponding product of the matrices C of [5, Lemma 1] is $C_1 \cdots C_k = R\,\Pi A \Pi\, M$. Moreover,*

$$\|C_1 \cdots C_k\|_2 \le \sqrt{n+1}\,|A| \qquad \text{and} \qquad |A| \le \sqrt{n+1}\,\|C_1 \cdots C_k\|_2.$$

Proof. Since $F = [e_1 \mid M]$, the relation $AF = F \begin{bmatrix} 1 & 0 \\ b & C \end{bmatrix}$ gives $AM = MC$, where $C = C_1 \cdots C_k$ by the block-triangular structure. Since $e^T M = 0$, we have $\Pi M = M$ and $RM = [0 \mid I_n]M = I_n$. Hence $C = RAM = R\,\Pi A \Pi\, M$. The map $\psi : W^T \to \mathbb{R}^{1\times n}$, $r \mapsto r M$ is injective with inverse $z \mapsto zR$, and $zC = \psi(\psi^{-1}(z)\,\Pi A \Pi)$. The bounds follow from $\|M\|_2 = \sqrt{n+1}$ and $\|R\|_2 \le 1$. □

The proof uses only $e^T A = e^T$. For every $A \in \mathrm{A}$ we therefore write $C(A) = R\,\Pi A \Pi\, M$; then $F^{-1}AF$ has the block form above with $C = C(A)$, and both bounds of Proposition 3.7 hold. This applies in particular to the matrices of the original method in Section 7.

4 An exact identity for the moment of inertia

For a simplex S with ordered vertices define $d = \|x_{n+1} - x_c\|$ and the *share of the worst vertex*

$$s(S) = \frac{n}{n+1}\,\frac{d^2}{\Phi(S)}.$$

By framing the moment of inertia physically using the parallel axis theorem, the nonlinear update of moving a vertex becomes completely decoupled from the rest of the simplex, and we obtain Theorem 4.1.

Theorem 4.1 (Moment-of-inertia identity)**.** *Let S be nondegenerate. Then $s(S) \in (0, 1]$, and for every $\lambda \in \mathbb{R}$ and $\sigma > 0$, we have*

$$\Phi\,[x_1, \ldots, x_n, X(\lambda)] = \Phi(S)\,\big(1 + (\lambda^2 - 1)\,s(S)\big), \qquad \Phi\,S\,T_{\text{shr}}(\sigma) = \sigma^2\,\Phi(S). \tag{2}$$

Proof. By the Huygens–Steiner theorem, for all $m \in \{1, 2, 3, \dots\}$ we have $\sum_{i=1}^{m} \|y_i - z\|^2 = \sum_{i=1}^{m} \|y_i - \bar{y}\|^2 + m\|\bar{y} - z\|^2$ for points $y_1, \dots, y_m$ with centroid $\bar{y}$ and every z.

Let x_c be as in (1). With $\Psi = \sum_{i=1}^{n} \|x_i - x_c\|^2$ and $\bar{x} = \frac{n x_c + x_{n+1}}{n+1}$ we obtain

$$\Phi(S) = \Psi + n\|x_c - \bar{x}\|^2 + \|x_{n+1} - \bar{x}\|^2 = \Psi + \frac{n}{(n+1)^2} d^2 + \frac{n^2}{(n+1)^2} d^2 = \Psi + \frac{n}{n+1} d^2.$$

Hence $s(S) \le 1$, and $s(S) > 0$ because $d > 0$ for a nondegenerate simplex. Replacing x_{n+1} by $X(\lambda)$ changes neither Ψ nor x_c, and $\|X(\lambda) - x_c\| = |\lambda|\, d$. Therefore Φ becomes $\Psi + \lambda^2 \frac{n}{n+1} d^2 = \Phi(S) + (\lambda^2 - 1)\frac{n}{n+1} d^2$. A shrink multiplies all differences $x_i - x_j$ by σ, and the claim follows from Lemma 3.2(a). □

Remark 4.2. The identity also follows from Proposition 3.4. With $G = \Pi - (1+\alpha)\hat{w}\hat{w}^T$ we have $G^2 = \Pi + (\alpha^2 - 1)\hat{w}\hat{w}^T$ and $\|D\hat{w}\|^2 = \frac{n+1}{n}\|x_{n+1} - \bar{x}\|^2 = \frac{n}{n+1} d^2$. Hence $\|DG\|_F^2 = \Phi + (\alpha^2 - 1)\frac{n}{n+1} d^2$.

The shares $s_j(S)$ of all vertices, defined in Lemma 4.3 below, sum to $(n+1)/n$. This implies that the average of the shares equals exactly $1/n$, which is stated in Lemma 4.3. In our probabilistic model in Section 6, this implies that for a uniformly distributed choice of the vertices, the expected share of the moment of inertia carried by the replaced vertex is exactly the average share $1/n$, regardless of the simplex's geometric distortion.

Lemma 4.3 (Average share)**.** *For $j = 1, \dots, n+1$ let $s_j(S) = \frac{n}{n+1}\|x_j - x_c^{(-j)}\|^2 / \Phi(S)$, where $x_c^{(-j)}$ is the centroid of the vertices other than x_j. Then $\frac{1}{n+1}\sum_{j=1}^{n+1} s_j(S) = \frac{1}{n}$.*

Proof. Since $x_j - \bar{x} = \frac{n}{n+1}(x_j - x_c^{(-j)})$, we have $s_j(S) = \frac{n+1}{n}\|x_j - \bar{x}\|^2 / \Phi(S)$, and the claim follows by summation. □

Theorem 4.1 yields the following corollary.

Corollary 4.4 (Product formula)**.** *Let $s_k = s(S^{(k)})$. Then $\Phi_k = \Phi_0 \prod_{i=0}^{k-1} \varphi_i$, where $\varphi_i = 1 + (\alpha_i^2 - 1)s_i$ for non-shrink steps and $\varphi_i = \sigma^2$ for shrink steps. In particular, $\varphi_i = 1$ for reflections, $1 < \varphi_i \le \chi^2$ for expansions and $\gamma^2 \le \varphi_i < 1$ for contractions. Since x_c, d and Φ can be updated with $O(n)$ operations, the factors φ_i can be monitored along a run at cost $O(n)$ per iteration.*

The product formula yields a deterministic convergence criterion that involves only the scalar sequence $(\varphi_k)_k$ and holds in every dimension.

Theorem 4.5 (Deterministic criterion)**.** *Let $(S^{(k)})$ be generated from a nondegenerate simplex $S^{(0)}$ by non-shrink NM operations with coefficients $|\alpha_k| \le \chi$ and by shrinks with coefficient σ, and let $\kappa = \max\left\{\frac{\sqrt{1+\chi}}{\sqrt{n(n+1)}}, 1 - \sigma\right\}$. If $\sum_{k \ge 0} \sqrt{\Phi_k} < \infty$, then there is $\hat{x} \in \mathbb{R}^n$ with $x_i^{(k)} \to \hat{x}$ for all $i \in \{1, \dots, n+1\}$, and*

$$\|x_i^{(k)} - \hat{x}\| \le \sqrt{\Phi_k} + \kappa \sum_{j \ge k} \sqrt{\Phi_j}.$$

In particular, if $\limsup_{k\to\infty} \frac{1}{k} \sum_{i<k} \log \varphi_i = -2\eta < 0$, then $\|x_i^{(k)} - \hat{x}\| = O(\exp(-(\eta - \varepsilon)k))$ for every $\varepsilon \in (0, \eta)$. If f is continuous at $\hat{x}$, then $f(x_i^{(k)}) \to f(\hat{x})$.

Proof. Reordering does not change $\bar{x}$. For a non-shrink step,

$$\bar{x}^{(k+1)} - \bar{x}^{(k)} = \frac{X(\alpha_k) - x_{n+1}}{n+1} = \frac{1+\alpha_k}{n+1}(x_c - x_{n+1}),$$

and $|1 + \alpha_k| \le 1 + \chi$ together with $d^2 = \frac{n+1}{n} s_k \Phi_k \le \frac{n+1}{n}\Phi_k$ give $\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| \le \frac{\sqrt{1+\chi}}{\sqrt{n(n+1)}}\sqrt{\Phi_k}$. For a shrink, $\bar{x}^{(k+1)} - \bar{x}^{(k)} = (1-\sigma)(x_1 - \bar{x})$, whose norm is at most $(1-\sigma)\sqrt{\Phi_k}$. Hence

$\sum_k \|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| \le \kappa \sum_k \sqrt{\Phi_k} < \infty$, so $\bar{x}^{(k)}$ converges to some $\hat{x}$ with $\|\bar{x}^{(k)} - \hat{x}\| \le \kappa \sum_{j \ge k} \sqrt{\Phi_j}$. Since $\|x_i^{(k)} - \bar{x}^{(k)}\| \le \sqrt{\Phi_k}$, the estimate follows. Under the lim sup-condition, $\Phi_k \le e^{-2(\eta-\varepsilon)k}$ for sufficiently large k, which gives the rate. The continuity of f yields the last statement. □

Remark 4.6. Theorem 4.5 plays the role of [5, Theorem 1] and of its extension to both versions of the method [6, Theorem 5]. The density condition on ℓ-products, whose verification requires all $|\mathrm{T}|^\ell$ products, is replaced by a condition on the scalars $\varphi_k = 1 + (\alpha_k^2 - 1)s_k$. The identity also shows the mechanism behind the dimension dependence: a typical vertex carries the share $1/n$ (Lemma 4.3). A single contraction therefore reduces Φ only by a factor of about $1 - \frac{3}{4} \cdot \frac{1}{n}$, and a single expansion increases it by a factor of about $1 + \frac{3}{n}$ (standard coefficients). This parallels the descent factor $\frac{n-1}{2n^2}$ obtained by Gao and Han [7, Theorem 2.1] for the sum of the function values.

Remark 4.7 (Relation to [7]). For uniformly convex f and standard coefficients, Gao and Han [7, Corollary 2.2] showed that $\operatorname{diam}(S^{(k)}) \to 0$ whenever infinitely many expansion or contraction steps occur. By Lemma 3.2(b), in this case we also have $\Phi_k \to 0$. Theorem 4.5 requires the stronger condition $\sum_k \sqrt{\Phi_k} < \infty$, which in addition yields convergence of the vertices to a common point. Since reflections leave Φ unchanged (Corollary 4.4), Φ stays constant along consecutive reflections. By [7, Example 2.3], for every m there is a uniformly convex function on R^3 for which the first m iterations from a regular simplex are reflections. Hence no a priori rate for Φ_k can hold without further assumptions.

5 First-moment criteria in all dimensions

For the first result about almost sure convergence of the algorithm, as in [5], we consider a probabilistic model for the choice of the operations in the Nelder–Mead iteration. We make the following assumption.

Assumption 5.1 (Random model (R1))**.** The pairs $T_k P_k$, $k \ge 0$, are independent and identically distributed on a finite set of matrices $T(\alpha_\tau)P$ and $T_{\mathrm{shr}}(\sigma)P$, and $S^{(k+1)} = S^{(k)} T_k P_k$. The probabilities of the types are p_r, p_e, p_oc, p_ic, p_s.

Under the above distribution assumption and an additional assumption on the expected value, we obtain the following result about almost sure convergence:

Theorem 5.2. *Under Assumption 5.1, let* $\mu = \mathrm{E}\,|T_0 P_0| = 1 + (\chi - 1)p_\mathrm{e} - (1 - \sigma)p_\mathrm{s}$. *If* $\mu < 1$, *then with probability one all vertices converge to a common limit, and* $\mathrm{E}\operatorname{diam}(S^{(k)}) \le 2n\,\|D_0\|_2\,\mu^k$.

Proof. By Lemma 3.2(b), $\operatorname{diam}(S^{(k)}) \le \sqrt{2\Phi_k} = \sqrt{2}\,\|D_k\|_F \le \sqrt{2n}\,\|D_k\|_2$. By Corollary 3.6, $\|D_k\|_2 \le \|D_0\|_2 \prod_{i=0}^{k-1} \|Q_i\|_2$ with independent $X_i = \ln |T_i P_i|$ and $\mathrm{E}X_i = \ln \mu$. Hence $\mathrm{E}\,\|D_k\|_2 \le \|D_0\|_2\,\mu^k$ and $\sum_k \mathrm{E}\sqrt{\Phi_k} < \infty$. Thus $\sum_k \sqrt{\Phi_k} < \infty$ almost surely, and Theorem 4.5 applies. □

We can apply Theorem 5.2 to the setting of [5], where almost sure convergence was verified for dimensions up to 6.

Corollary 5.3 (Theorem 2 of [5] in every dimension)**.** *Let the elements of* T *(standard coefficients) be chosen independently and uniformly. Then*

$$\mu_n = \frac{3n + 4 + (n+1)!/2}{3n + 3 + (n+1)!} < 1 \qquad \text{for every } n \ge 2,$$

and the conclusion of [5, Theorem 2] holds in every dimension $n \ge 2$ *with blocks of length* $\ell = 1$.

Proof. By Corollary 3.6, the n reflections and the $2n + 2$ contractions in $\mathbb{T}^n$ have norm 1, the expansion has norm 2 and the $(n + 1)!$ shrinks have norm $\frac{1}{2}$. This gives μ_n, and $\mu_n < 1$ is equivalent to $(n + 1)! > 2$. Theorem 5.2 completes the proof. □

Remark 5.4 (Dominance of shrinks). The condition $\mu < 1$ of Theorem 5.2 is equivalent to $(1 - \sigma)p_s > (\chi - 1)p_e$, i.e. to $p_s > 2p_e$ for the standard coefficients. A first-moment criterion with blocks of length one can therefore succeed only through shrink steps. In the uniform model of [5], the probability of a shrink is $p_s = (n + 1)!/|\mathrm{T}|$, which tends to one. For $n \geq 5$, more than 97% of all operations are shrinks. This is also the regime behind the values reported in [5, (23)] for $n \geq 4$. Since no shrink steps occur when NM is applied to strictly convex functions [11, Lemma 3.5] (see [6, Lemma 8] for the original method), the relevant question in high dimensions is the shrink-free case.

6 Second-moment criteria in all dimensions

Now we modify the probabilistic model by adding a uniform distribution assumption on the indices $J_k \in \{1, \ldots, n + 1\}$ of the exchanged vertices. Note that the identity of Theorem 4.1 measures the contraction of the simplex in all directions simultaneously.

6.1 Exchangeable vertex ranking

We define the probabilistic model with uniform distribution of the choice of the exchanged vertices as follows:

Assumption 6.1 (Exchangeable model (R2))**.** Let $(\tau_k, J_k)_{k\geq 0}$ be independent and identically distributed, where $\tau_k \in \{\mathrm{r}, \mathrm{e}, \mathrm{oc}, \mathrm{ic}, \mathrm{s}\}$ has the distribution (p_τ) and J_k is uniformly distributed on $\{1, \ldots, n + 1\}$ and independent of τ_k. Starting from a nondegenerate $S^{(0)}$, $S^{(k+1)}$ is obtained from $S^{(k)}$ as follows. If $\tau_k \neq \mathrm{s}$, the vertex $x_{J_k}^{(k)}$ is replaced by $x_c + \alpha_{\tau_k}(x_c - x_{J_k}^{(k)})$, where x_c is the centroid of the other n vertices. If $\tau_k = \mathrm{s}$, a shrink with factor σ towards the vertex $x_{J_k}^{(k)}$ is performed. Let F_k denote the σ-algebra generated by $\tau_0, J_0, \ldots, \tau_{k-1}, J_{k-1}$, i.e. the information about the first k steps. Then $S^{(k)}$ is determined by F_k, whereas (τ_k, J_k) is independent of F_k.

Since Φ is invariant under reordering, (R2) describes NM operations in which the rank of the replaced vertex is uniformly distributed and carries no information about the geometry of the simplex. We obtain the following results about almost sure convergence under conditions on the iteration parameters and the corresponding probabilities:

Theorem 6.2. *Under Assumption 6.1, we have*

$$\mathrm{E}[\Phi_{k+1} \mid \mathrm{F}_k] = \vartheta_n \Phi_k, \qquad \text{where } \vartheta_n = \sigma^2 p_s + (1 - p_s) - \frac{1}{n} \sum_{\tau \neq \mathrm{s}} p_\tau (1 - \alpha_\tau^2).$$

In particular, $\mathrm{E}\Phi_k = \vartheta_n^k \Phi_0$. *If* $\vartheta_n < 1$, *then with probability one all vertices converge to a common limit* $\hat{x}$, *and* $\mathrm{E}\,\mathrm{diam}(S^{(k)}) \leq \sqrt{2\Phi_0}\,\vartheta_n^{k/2}$.

Proof. By Assumption 6.1, $S^{(k)}$, and hence Φ_k and the shares $s_j(S^{(k)})$, $j = 1, \ldots, n + 1$, of Lemma 4.3, are determined by F_k, whereas τ_k and J_k are independent of F_k and of each other, with $\mathrm{P}(\tau_k = \tau) = p_\tau$ and $\mathrm{P}(J_k = j) = \frac{1}{n+1}$.

Let $\tau_k \neq \mathrm{s}$. Since the numbering of the vertices does not change Φ (Lemma 3.2(c)), Theorem 4.1, applied to the replaced vertex $x_{J_k}^{(k)}$, gives

$$\Phi_{k+1} = \left(1 + (\alpha_{\tau_k}^2 - 1)\, s_{J_k}(S^{(k)})\right) \Phi_k.$$

Here α_{τ_k}, Φ_k and the shares $s_j(S^{(k)})$ are known if F_k and τ_k are given, and J_k is uniformly distributed and independent of F_k and τ_k. Hence the conditional expectation of $s_{J_k}(S^{(k)})$ is the average share, and Lemma 4.3 gives

$$\mathrm{E}\left[\Phi_{k+1} \mid \mathrm{F}_k, \tau_k\right] = \left(1 + (\alpha_{\tau_k}^2 - 1)\frac{1}{n+1}\right) \sum_{j=1}^{n+1} s_j(S^{(k)})\, \Phi_k = \left(1 + \frac{\alpha_{\tau_k}^2 - 1}{n}\right) \Phi_k.$$

If $\tau_k = \mathrm{s}$, then $\Phi_{k+1} = \sigma^2 \Phi_k$ by Theorem 4.1. In both cases, $\mathrm{E}[\Phi_{k+1} \mid \mathrm{F}_k, \tau_k] = g(\tau_k)\,\Phi_k$, where $g(\mathrm{s}) = \sigma^2$ and $g(\tau) = 1 + (\alpha_\tau^2 - 1)/n$ for $\tau \neq \mathrm{s}$.

To average over τ_k, we use the tower property of conditional expectations. Since F_k is contained in the σ-algebra generated by F_k and τ_k, we have

$$\mathrm{E}\left[\Phi_{k+1} \mid \mathrm{F}_k\right] = \mathrm{E}\left[g(\tau_k)\,\Phi_k \mid \mathrm{F}_k\right].$$

As Φ_k is determined by F_k and τ_k is independent of F_k, the right-hand side equals $\Phi_k\,\mathrm{E}[g(\tau_k)] = \Phi_k \sum_\tau p_\tau\, g(\tau)$. With $\sum_{\tau \neq \mathrm{s}} p_\tau = 1 - p_\mathrm{s}$, we obtain

$$\sum_\tau p_\tau\, g(\tau) = \sigma^2 p_\mathrm{s} + \sum_{\tau \neq \mathrm{s}} p_\tau \left(1 - \frac{1-\alpha_\tau^2}{n}\right) = \sigma^2 p_\mathrm{s} + (1 - p_\mathrm{s}) - \frac{1}{n}\sum_{\tau \neq \mathrm{s}} p_\tau \left(1 - \alpha_\tau^2\right) = \vartheta_n.$$

This proves $\mathrm{E}[\Phi_{k+1} \mid \mathrm{F}_k] = \vartheta_n \Phi_k$. Taking expectations gives $\mathrm{E}\Phi_{k+1} = \vartheta_n\,\mathrm{E}\Phi_k$, and $\mathrm{E}\Phi_k = \vartheta_n^k \Phi_0$ follows by induction, since $S^{(0)}$ is deterministic.

By Jensen's inequality, $\mathrm{E}\sqrt{\Phi_k} \le \sqrt{\mathrm{E}\Phi_k} = \sqrt{\Phi_0}\,\vartheta_n^{k/2}$, hence $\mathrm{E}\sum_k \sqrt{\Phi_k} < \infty$, and Theorem 4.5 yields almost sure convergence. The bound for the diameter follows from Lemma 3.2(b). □

For the standard coefficients $\chi = 2$ and $\gamma = \sigma = \frac{1}{2}$ this yields the following result:

Corollary 6.3 (A dimension-free condition)**.** *For the standard coefficients,*

$$\vartheta_n = 1 - \tfrac{3}{4} p_\mathrm{s} - \frac{1}{n}\left(\tfrac{3}{4}(p_\mathrm{oc} + p_\mathrm{ic}) - 3p_\mathrm{e}\right),$$

and $\vartheta_n < 1$ if and only if $np_\mathrm{s} + p_\mathrm{oc} + p_\mathrm{ic} > 4p_\mathrm{e}$. Without shrinks, the condition reads

$$p_\mathrm{oc} + p_\mathrm{ic} > 4\,p_\mathrm{e},$$

independently of n: contractions must be more than four times as frequent as expansions. Then $\vartheta_n = 1 - c/n$ with $c = \frac{3}{4}(p_\mathrm{oc} + p_\mathrm{ic}) - 3p_\mathrm{e}$, so that $\mathrm{E}\Phi_k \le e^{-ck/n}\Phi_0$. Reducing $\mathrm{E}\Phi$ by a constant factor thus takes $O(n)$ iterations.

For the coefficients proposed by Gao and Han [7], the result is as follows:

Corollary 6.4 (Coefficients of Gao and Han)**.** *For the dimension-dependent coefficients $\chi = 1 + \frac{2}{n}$, $\gamma = \frac{3}{4} - \frac{1}{2n}$ and $\sigma = 1 - \frac{1}{n}$ of [7],*

$$\vartheta_n = \left(1 - \frac{1}{n}\right)^2 p_\mathrm{s} + (1 - p_\mathrm{s}) - \frac{1}{n}\,\frac{7n^2 + 12n - 4}{16n^2}(p_\mathrm{oc} + p_\mathrm{ic}) - \frac{4(n+1)}{n^2}\,p_\mathrm{e},$$

and $\vartheta_n < 1$ if and only if $16n(2n-1)\,p_\mathrm{s} + (7n^2 + 12n - 4)(p_\mathrm{oc} + p_\mathrm{ic}) > 64(n+1)\,p_\mathrm{e}$. Without shrinks, the condition reads

$$(7n^2 + 12n - 4)\,(p_\mathrm{oc} + p_\mathrm{ic}) > 64(n+1)\,p_\mathrm{e},$$

and then $\vartheta_n = 1 - c/n$ with $c = \frac{7n^2+12n-4}{16n^2}(p_\mathrm{oc} + p_\mathrm{ic}) - \frac{4(n+1)}{n^2}\,p_\mathrm{e}$, so that $\mathrm{E}\Phi_k \le e^{-ck/n}\,\Phi_0$.

Proof. Since $\chi^2 - 1 = \frac{4(n+1)}{n^2}$, $1 - \gamma^2 = \frac{7n^2+12n-4}{16n^2}$ and $1 - \sigma^2 = \frac{2n-1}{n^2}$, Theorem 6.2 gives the formula for ϑ_n and

$$16n^3\,(1 - \vartheta_n) = 16n(2n-1)\,p_\mathrm{s} + (7n^2 + 12n - 4)\,(p_\mathrm{oc} + p_\mathrm{ic}) - 64(n+1)\,p_\mathrm{e}.$$

This proves the conditions for $\vartheta_n < 1$. The bound for $\mathrm{E}\Phi_k$ follows from $\vartheta_n^k = (1 - c/n)^k \le e^{-ck/n}$. □

6.2 The insertion model of [5]

In the random model of [5], the replaced vertex is always the vertex in position $n + 1$, which is determined by the previous insertions. Its share is therefore not uniformly distributed, and we use a second-moment operator.

Assumption 6.5 (Insertion model (R3))**.** The pairs T_kP_k are independent and uniformly distributed on $\tilde{\mathrm{T}}$, and $S^{(k+1)} = S^{(k)}T_kP_k$. Let $\mathrm{F}_k = \sigma(T_iP_i : i < k)$.

Let $U \in \mathbb{R}^{(n+1)\times n}$ have orthonormal columns spanning W, so that $UU^T = \Pi$. For $TP \in \tilde{\mathrm{T}}$ let $A(TP) = U^T TPU \in \mathbb{R}^{n\times n}$, and let $Q_k = (D_kU)^T(D_kU)$. By Lemma 3.3, $D_{k+1}U = D_kU\,A(T_kP_k)$, and $\Phi_k = \operatorname{tr} Q_k$. Define the linear operator

$$\mathrm{L}_n(Q) = \frac{1}{3n+3} \sum_{TP\in\tilde{\mathrm{T}}} A(TP)^T QA(TP)$$

on the symmetric $n \times n$ matrices.

For the probabilistic model (R3) with uniformly distributed choice of the transformation matrices, we obtain the following result about almost sure convergence.

Theorem 6.6. *Under Assumption 6.5,* $\mathrm{E}[Q_{k+1} \mid \mathrm{F}_k] = \mathrm{L}_n(Q_k)$. *If*

$$\rho(\mathrm{L}_n) < 1, \tag{3}$$

then for every $r \in (\rho(\mathrm{L}_n), 1)$*, there is a constant* C_r *with* $\mathrm{E}\Phi_k \le C_r r^k \Phi_0$. *Moreover, with probability one, all vertices converge to a common limit, and* $\limsup_{k\to\infty} \frac{1}{k} \log \Phi_k \le \log \rho(\mathrm{L}_n)$.

Proof. The first statement follows from the independence of T_kP_k and F_k. Hence $\mathrm{E}Q_k = \mathrm{L}_n^k(Q_0)$. Let $\|\mathrm{L}_n^k\|$ denote the norm of L_n^k as an operator on the symmetric $n \times n$ matrices with the Frobenius norm. By Gelfand's formula, $\|\mathrm{L}_n^k\| \le C_r' r^k$ for every $r > \rho(\mathrm{L}_n)$. We use the following inequalities between the trace and the Frobenius norm:

$$\operatorname{tr} X \le \sqrt{n}\,\|X\|_F \quad \text{for all } X \in \mathbb{R}^{n\times n}, \qquad \|X\|_F \le \operatorname{tr} X \quad \text{if } X \text{ is positive semidefinite.}$$

The first is the Cauchy–Schwarz inequality for the Frobenius inner product, $\operatorname{tr} X = \operatorname{tr}(I_n^T X) \le \|I_n\|_F \, \|X\|_F$. The second follows from $\|X\|_F^2 = \sum_i \lambda_i^2 \le \sum_i \lambda_i{}^2 = (\operatorname{tr} X)^2$, where $\lambda_1, \ldots, \lambda_n \ge 0$ are the eigenvalues of X. Since $\Phi_k = \operatorname{tr} Q_k$ and Q_0 is positive semidefinite with $\operatorname{tr} Q_0 = \Phi_0$, we obtain

$$\mathrm{E}\Phi_k = \operatorname{tr} \mathrm{L}_n^k(Q_0) \le \sqrt{n}\,\| \mathrm{L}_n^k(Q_0)\|_F \le \sqrt{n}\,\| \mathrm{L}_n^k\| \, \|Q_0\|_F \le \sqrt{n}\, C_r' r^k \operatorname{tr} Q_0 = \sqrt{n}\, C_r' r^k \Phi_0,$$

which proves the bound for $\mathrm{E}\Phi_k$ with $C_r = \sqrt{n}\, C_r'$. As in the proof of Theorem 6.2, $\mathrm{E} \sum_k \sqrt{\Phi_k} < \infty$, and Theorem 4.5 applies. For $r < r' < 1$, Markov's inequality gives $\mathrm{P}(\Phi_k > r'^k) \le C_r(r/r')^k\Phi_0$. By the Borel–Cantelli lemma, $\Phi_k \le r'^k$ for all large k almost surely, which proves the last statement. □

Now we verify that (3) holds for all n.

Theorem 6.7. *Let* $n \ge 2$, $b_n = \frac{n-1}{2(n+1)}$ *and* $\tilde{w} = U^T\hat{w}$, *and let* v_n *be the smallest eigenvalue of* $M_n = \sum_{i=0}^{n-1} (\mathrm{L}_n^*)^i(\tilde{w}\tilde{w}^T)$, *where* $\mathrm{L}_n^*(X) = \frac{1}{3n+3} \sum_{TP\in\mathrm{T}} A(TP)\,X\,A(TP)^T$ *is the adjoint of* L_n. *Then*

$$\operatorname{tr} \mathrm{L}_n(Q) = \operatorname{tr} Q - b_n \tilde{w}^T Q \tilde{w}$$

for every symmetric Q, $v_n > 0$, *and* $\operatorname{tr} \mathrm{L}_n^n(Q) \le (1 - b_nv_n) \operatorname{tr} Q$ *for every* $Q \ge 0$. *In particular,* $\rho(\mathrm{L}_n)^n \le 1 - b_nv_n < 1$.

Proof. Since $UU^T = \Pi$ and $P\Pi = \Pi P$, Proposition 3.4 gives $A(T(\alpha)P) = (I_n - (1+\alpha)\tilde{w}\tilde{w}^T)\tilde{P}$ with the orthogonal matrix $\tilde{P} = U^T P U$. Hence $AA^T = I_n + (\alpha^2 - 1)\tilde{w}\tilde{w}^T$ and $\operatorname{tr}(A^T Q A) = \operatorname{tr} Q + (\alpha^2 - 1)\,\tilde{w}^T Q \tilde{w}$, which is Theorem 4.1 in these coordinates. The set $\tilde{\mathcal{T}}$ contains $n+1$ outside and $n+1$ inside contractions with $\alpha^2 - 1 = -\frac{3}{4}$, n reflections with $\alpha^2 - 1 = 0$ and one expansion with $\alpha^2 - 1 = 3$. Averaging gives the identity for $\operatorname{tr} \mathrm{L}_n(Q)$, since $\frac{3}{2}(n+1) + 3 = (3n+3) = -b_n$. Summing it along L_n yields, for $Q \succeq 0$,

$$\operatorname{tr} \mathrm{L}_n^n(Q) = \operatorname{tr} Q - b_n \sum_{i=0}^{n-1} \tilde{w}^T \mathrm{L}_n^i(Q)\,\tilde{w} = \operatorname{tr} Q - b_n \operatorname{tr}(QM_n) \le (1 - b_n v_n) \operatorname{tr} Q.$$

The map L_n preserves positive semidefiniteness: if $Q \succeq 0$, then $x^T A^T Q A x = (Ax)^T Q (Ax) \ge 0$ for every $x \in \mathbb{R}^n$, so that each term $A^T Q A$ of $\mathrm{L}_n(Q)$ is positive semidefinite, and so is their mean. By induction, every power L_n^i preserves positive semidefiniteness. (Since every $A(TP)$ is invertible, with $\det(I_n - (1+\alpha)\tilde{w}\tilde{w}^T) = -\alpha \neq 0$, the same argument shows that L_n^i also preserves positive definiteness.)

To show $v_n > 0$, let $Q \succeq 0$ with $\operatorname{tr}(QM_n) = 0$, and let $Q_i = \mathrm{L}_n^i(Q) \succeq 0$. Then $\tilde{w}^T Q_i \tilde{w} = 0$, i.e. $Q_i \tilde{w} = 0$, for $i = 0, \ldots, n-1$. If $Q_i \tilde{w} = 0$, then $(I_n - (1+\alpha)\tilde{w}\tilde{w}^T) Q_i (I_n - (1+\alpha)\tilde{w}\tilde{w}^T) = Q_i$, so that $Q_{i+1} = \frac{1}{3n+3} \sum_{TP \in \tilde{\mathrm{T}}} \tilde{P}^T Q_i \tilde{P}$, where every insertion permutation P_j, $j = 1, \ldots, n+1$, occurs. Since all terms are positive semidefinite, $Q_{i+1} v = 0$ implies $Q_i \tilde{P}_j v = 0$ for all j. Starting from $Q_m \tilde{w} = 0$ and descending m times, we find that Q_0 vanishes on $\tilde{P}_{j_m} \cdots \tilde{P}_{j_1} \tilde{w} = U^T P_{j_m} \cdots P_{j_1} \hat{w}$ for all $m \le n-1$ and all $j_1, \ldots, j_m$. Since $\hat{w}$ is a multiple of $e_{n+1} - \frac{1}{n+1} e$, $P_j e_{n+1} = e_n$ for $j \le n$, and $P_j e_m = e_{m-1}$ for $j < m$, the matrix Q_0 vanishes on $U^T(e_m - \frac{1}{n+1} e)$ for $m = 2, \ldots, n+1$. Any n of the $n+1$ vectors $e_m - \frac{1}{n+1} e$ span W, hence $Q_0 = 0$. Thus $\operatorname{tr}(QM_n) > 0$ for every $Q \succeq 0$ with $Q \neq 0$, and $v_n > 0$.

Finally, every symmetric Q is a difference $Q = Q_+ - Q_-$ of positive semidefinite matrices with $\|Q\|_* = \operatorname{tr} Q_+ + \operatorname{tr} Q_-$ for the trace norm. Since L_n preserves positive semidefiniteness, $\|\mathrm{L}_n(Q)\|_* \le \operatorname{tr} \mathrm{L}_n(Q_+) + \operatorname{tr} \mathrm{L}_n(Q_-) \le (1 - b_n v_n)\,\|Q\|_*$. Thus the operator norm of L_n induced by the trace norm is at most $1 - b_n v_n$, and Gelfand's formula, together with the submultiplicativity of this norm, gives $\rho(\mathrm{L}_n) = \lim_{m\to\infty} \|\mathrm{L}_n^{mn}\|_*^{1/(mn)} \le (1 - b_n v_n)^{1/n}$, i.e. $\rho(\mathrm{L}_n)^n \le 1 - b_n v_n$. □

The moment of inertia is thus a Lyapunov function for the insertion model: by Theorem 6.6, $\mathrm{E}[\Phi_{k+1} \mid \mathrm{F}_k] = \operatorname{tr} \mathrm{L}_n(Q_k) = (1 - b_n s_k)\Phi_k \le \Phi_k$, where s_k is the share of the worst vertex, and Φ_k decreases in expectation by at least the factor $1 - b_n v_n$ over every n consecutive steps, since the worst vertex cannot carry a zero share for n steps.

By Theorem 6.7, condition (3) holds in every dimension $n \ge 2$; hence, by Theorem 6.6, all vertices converge to a common limit with probability one in the insertion model. The bound $\rho(\mathrm{L}_n)^n \le 1 - b_n v_n$ is weak, since v_n decreases rapidly: numerically, v_n decreases from 0.53 for $n = 2$ to 0.0081 for $n = 8$. The rate is described by Table 1, which lists $\rho(\mathrm{L}_n)$ for a selection of values $n \le 500$. The decay rate per iteration behaves like $1 - c_n/n$, where $c_n = n(1 - \rho(\mathrm{L}_n))$ increases from 0.164 for $n = 2$ to 0.475 for $n = 500$.

For the uniform distribution on $\tilde{\mathrm{T}}$, the types have the frequencies $p_{\mathrm{r}} = \frac{n}{3n+3}$, $p_{\mathrm{e}} = \frac{1}{3n+3}$ and $p_{\mathrm{oc}} = p_{\mathrm{ic}} = \frac{1}{3}$ so that Corollary 6.3 gives $n(1 - \vartheta_n) = \frac{1}{2} - \frac{1}{n+1}$ for the exchangeable model with these frequencies (last column of Table 1). For every n in Table 1, the computed eigenvalue λ has a residual below 10^{-12}, and the corresponding eigenvector Y is positive definite. This identifies λ as $\rho(\mathrm{L}_n)$: since L_n maps positive semidefinite matrices to positive semidefinite matrices, $\rho(\mathrm{L}_n)$ is an eigenvalue with a positive semidefinite eigenvector Z by the Krein–Rutman theorem, and with $Z \preceq cY$ for some $c > 0$ we obtain $\rho(\mathrm{L}_n)^k Z = \mathrm{L}_n^k(Z) \preceq c\lambda^k Y$ for all k, hence $\rho(\mathrm{L}_n) \le \lambda$.

Remark 6.8 (Complexity). The operator L_n acts on a space of dimension $n(n+1)/2$. By Proposition 3.4, each $A(TP)$ is a rank-one modification of a permutation, so one application of L_n costs $O(n^3)$ operations. Since the insertion permutations P_j only shift indices, the sum over j can moreover be evaluated in closed form for each of the four coefficients, which reduces the

cost to $O(n^2)$. Hence $\rho(\mathrm{L}_n)$ can be computed, e.g. by Arnoldi iteration, in polynomial time. In contrast, [5] requires $|\mathrm{T}|^{\ell}$ matrix products. The same holds for every distribution on $\tilde{\mathrm{T}}$, with L_n replaced by the corresponding weighted sum.

Table 1: Spectral radius of the second-moment operator L_n for the insertion model (R3), compared with the exchangeable model (R2) with the same frequencies, for which $n(1-\vartheta_n) = \frac{1}{2} - \frac{1}{n+1}$.

n	$\rho(\mathrm{L}_n)$	$n\,(1-\rho(\mathrm{L}_n))$	$n\,(1-\vartheta_n)$
2	0.917811	0.164	0.167
3	0.921278	0.236	0.250
5	0.939981	0.300	0.333
10	0.964294	0.357	0.409
20	0.980234	0.395	0.452
40	0.989430	0.423	0.476
60	0.992743	0.435	0.484
100	0.995515	0.448	0.490
150	0.996953	0.457	0.493
200	0.997689	0.462	0.495
250	0.998137	0.466	0.496
300	0.998438	0.469	0.497
350	0.998655	0.471	0.497
400	0.998819	0.472	0.498
450	0.998947	0.474	0.498
500	0.999050	0.475	0.498

Note that in [5], the dimension was only considered up to $n = 6$.

7 The original method

In the original method of Nelder and Mead [14], the vertices keep their positions. Let h_k and ℓ_k be indices of a worst and of a best vertex of $S^{(k)}$, determined as in [6, Algorithm 1]. A non-shrink step replaces x_{h_k} by $x_c + \alpha_k(x_c - x_{h_k})$, where x_c is the centroid of the other vertices, and a shrink replaces every x_i by $x_{\ell_k} + \sigma(x_i - x_{\ell_k})$. Following [6, (5)–(8)], this reads $S^{(k+1)} = S^{(k)} T_{h_k}(\alpha_k)$ and $S^{(k+1)} = S^{(k)} T^{\mathrm{shr}}_{\ell_k}(\sigma)$, respectively, with

$$T_j(\alpha) = I + c_j(\alpha) - e_j\, e_j^T, \qquad T_j^{\mathrm{shr}}(\sigma) = \sigma I + (1-\sigma)\, e_j e^T,$$

where $c_j(\alpha) \in \mathbb{R}^{n+1}$ has the entry $-\alpha$ in position j and $\frac{1+\alpha}{n}$ elsewhere. For the standard coefficients, these matrices form the set

$$\mathrm{T_o} = \{T_j(\alpha) : \alpha \in \{\pm\tfrac{1}{2}, 1, 2\},\ j \le n+1\} \cup \{T_j^{\mathrm{shr}}(\tfrac{1}{2}) : j \le n+1\}$$

with $5(n+1)$ elements [6, (10)]; let $\hat{\mathrm{T}}_\mathrm{o}$ be the subset of its $4(n+1)$ non-shrink elements. Note that $T_{n+1}(\alpha) = T(\alpha)$ and $T_1^{\mathrm{shr}}(\sigma) = T_{\mathrm{shr}}(\sigma)$. If P and Q are permutation matrices with $Pe_{n+1} = e_j$ and $Qe_1 = e_j$, then

$$T_j(\alpha) = P\, T(\alpha)\, P^T, \qquad T_j^{\mathrm{shr}}(\sigma) = Q T_{\mathrm{shr}}(\sigma)\, Q^T, \tag{4}$$

cf. [6, (17), (18)]. If ties are broken consistently, both versions generate the same simplices, up to the order of the vertices [6, Section 4].

Proposition 7.1 (Original method). *Let* $\hat{w}_j = (e_j - \frac{1}{n+1}e)/\|e_j - \frac{1}{n+1}e\|$. *For all* $\alpha, \sigma \in \mathrm{R}$ *and* $j = 1, \ldots, n+1$,

$$\Pi T_j(\alpha)\Pi = \Pi - (1+\alpha)\hat{w}_j\hat{w}_j^T, \qquad \Pi T_j^{\mathrm{shr}}(\sigma)\Pi = \sigma\Pi,$$
$$|T_j(\alpha)| = \max\{1, |\alpha|\}, \qquad |T_j^{\mathrm{shr}}(\sigma)| = \sigma.$$

Replacing the vertex x_j *of a nondegenerate simplex* S *by* $c^{(-j)} + \lambda(x^{(-j)} - x_j)$ *multiplies* $\Phi(S)$ *by* $1 + (\lambda^2 - 1)s_j(S)$, *with* $x^{(-j)}$ *and* s_j *as in Lemma 4.3. Consequently, Corollary 4.4 and Theorems 4.5, 5.2 and 6.2 hold for the original method, with* s_k *the share of the replaced vertex.*

Proof. By (4), $P\Pi = \Pi P$ and Proposition 3.4, $\Pi T_j(\alpha)\Pi = P\,\Pi - (1+\alpha)\hat{w}\hat{w}^T\,P^T$, and $P\hat{w} = \hat{w}_j$. The shrink is treated in the same way. Since $r' \to rP$ maps the unit sphere of W^T onto itself, $|PAP^T| = |A|$, and Corollary 3.6 gives the norms. The statement on Φ is Theorem 4.1 for the simplex with the vertices renumbered so that x_j comes last; this does not change Φ, and the share of the last vertex becomes $s_j(S)$. The proofs of the cited results use only these properties. □

The exchangeable model (R2) is thus a model of the original method of Nelder and Mead. Its uniform case is the model behind the fractions computed for the original method in [6, Section 7].

Corollary 7.2 (Uniform model of the original method). *Let the elements of* T_o *be chosen independently and uniformly. This is the model (R2) with* $p_\tau = \frac{1}{5}$ *for all five types, and*

$$\vartheta_n = \frac{17}{20} + \frac{3}{10n}, \qquad \mathrm{E}\,|T_0| = \frac{11}{10}.$$

For $n \geq 3$, $\vartheta_n < 1$, *and with probability one all vertices converge to a common limit. For the uniform distribution on* $\tilde{\mathrm{T}}_o$, $\vartheta_n = 1 + \frac{3}{8n} > 1$.

Proof. Under the uniform distribution on T_o, the type of the step and the index j of the replaced vertex, or of the centre of the shrink, are independent and uniformly distributed, which is (R2). Theorem 6.2 gives $\vartheta_n = \frac{1}{4}\cdot\frac{1}{5} + \frac{4}{5}\cdot\frac{1}{5n}\,0 - 3 + \frac{3}{4} + \frac{3}{4}$, and Proposition 7.1 gives $\mathrm{E}\,|T_0| = \frac{1}{5}\,1 + 2 + 1 + 1 + \frac{1}{2}$. Without shrinks, $\vartheta_n = 1 - \frac{1}{4n}\,0 - 3 + \frac{3}{4} + \frac{3}{4}$. □

The method of Zhao et al. The method MNM of Zhao, Uroševi ć, Mladenović and Hansen [19] uses NM steps with the standard coefficients but replaces the shrink step: if the NM step calls for a shrink, only the worst vertex is moved toward the best one, i.e. x_{n+1} is replaced by $x_1 + \mu(x_{n+1} - x_1)$ with $\mu = 1 - \delta = \frac{1}{2}$ in their notation, and further vertices are moved in the following iterations until an improvement is obtained. The new vertex of a modified shrink does not lie on the line through x_{n+1} and x_c, but the proof of Theorem 4.1 applies to an arbitrary new vertex.

Proposition 7.3 (Modified shrink). *Let* S *be nondegenerate, let* $J \neq B$ *be two vertex indices, and let* S' *be obtained from* S *by replacing* x_J *with* $y = x_B + \mu(x_J - x_B)$, *where* $0 \leq \mu < 1$. *Then*

$$\Phi(S') = \Phi(S) + \frac{n}{n+1}\left(\|y - c^{(-J)}\|^2 - \|x_J - c^{(-J)}\|^2\right),$$

and the centroid of the vertices moves by at most $\sqrt{2}\,(1-\mu)\sqrt{\Phi(S)}/(n+1)$. *If* (J, B) *is uniformly distributed on the* $n(n+1)$ *ordered pairs of distinct indices, then*

$$\mathrm{E}\,\Phi(S') = \left(1 - \frac{2(1-\mu)(1+\mu n)}{n(n+1)}\right)\Phi(S).$$

Proof. Write $c = x_c^{(-J)}$. Replacing x_J changes neither c nor $\Psi_J = \sum_{i=J} \|x_i - c\|^2$, and $\Phi(S) = \Psi_J + \frac{n}{n+1}\|x_J - c\|^2$ as in the proof of Theorem 4.1. This gives the first assertion. The centroid moves by $(y - x_J)/(n+1) = (1-\mu)(x_B - x_J)/(n+1)$, and $\|x_B - x_J\| \le \operatorname{diam}(S) \le \sqrt{2\Phi(S)}$ by Lemma 3.2(b).

For the expectation, let $z_i = x_i - \bar{x}$, so that $\sum_i z_i = 0$, $\sum_i \|z_i\|^2 = \Phi(S)$ and $\sum_{J=B} \langle z_J, z_B\rangle = -\Phi(S)$. Since $x_J - c = \frac{n+1}{n} z_J$ and $x_B - c = z_B + \frac{1}{n} z_J$, averaging over the $n(n+1)$ pairs gives

$$\mathrm{E}\|x_J - c\|^2 = \frac{n+1}{n^2}\Phi(S), \qquad \mathrm{E}\langle x_J - c, x_B - c\rangle = 0, \qquad \mathrm{E}\|x_B - c\|^2 = \frac{n-1}{n^2}\Phi(S).$$

Since $y - c = \mu(x_J - c) + (1-\mu)(x_B - c)$, it follows that

$$\mathrm{E}\|y-c\|^2 - \mathrm{E}\|x_J - c\|^2 = \frac{\mu^2(n+1) + (1-\mu)^2(n-1) - (n+1)}{n^2}\Phi(S) = -\frac{2(1-\mu)(1+\mu n)}{n^2}\Phi(S),$$

and the first assertion yields the claim. □

A single modified shrink can increase Φ, but on average it reduces Φ like a contraction. The following remark summarizes which of our results on convergence apply to the method of [19].

Remark 7.4 (Convergence results for the method of [19]). By Proposition 7.3, Theorem 4.5 holds for MNM with $\kappa = \max\left\{\sqrt{\frac{1+\chi}{n(n+1)}}, \frac{\sqrt{2(1-\mu)}}{n+1}\right\}$. Theorem 5.2 gives no result, since every operation of MNM has invariant norm at least one (Corollary 3.6; in centered coordinates, a modified shrink is a rank-one modification of the identity). Theorem 6.2 applies in (R2) if a modified shrink moves a uniformly chosen vertex toward another one, with σ^2 replaced by the factor of Proposition 7.3. For the standard coefficients and $\mu = \frac{1}{2}$, $\vartheta_n < 1$ if and only if $p_{\mathrm{oc}} + p_{\mathrm{ic}} + \frac{2(n+2)}{3(n+1)} p_{\mathrm{s}} > 4p_{\mathrm{e}}$: a modified shrink acts like a further contraction.

Theorems 6.6 and 6.7 apply to MNM without shrinks, since its other steps are those of NM; in this model, all vertices converge almost surely in every dimension.

8 A modified algorithm with a sufficient decrease safeguard

The previous sections suggest two changes of the coefficients for large n. The possibility of stagnation at nonstationary points [13] calls for a safeguard in addition. In this section we combine both in Algorithm 1, prove a convergence result for it and test it numerically.

8.1 Coefficients

By Theorem 4.1, an expansion multiplies Φ by $1 + (\chi^2 - 1)s_k$. The cost of an expansion in terms of Φ can therefore be bounded by adapting the expansion coefficient to the share s_k of the replaced vertex, which is available at cost $O(n)$ (Corollary 4.4).

Proposition 8.1 (Share-capped expansion)**.** *Let $\chi > 1$ and $\omega > 0$, and let*

$$\chi_k = \min\left\{\chi, \sqrt{1 + \frac{\omega}{n s_k}}\right\}.$$

Then $1 < \chi_k \le \chi$, and an expansion with coefficient χ_k multiplies Φ_k by at most $1 + \omega/n$.

Proof. Since $s_k \in (0, 1]$ and $\omega > 0$, we have $\chi_k > 1$. By Theorem 4.1, the factor is $1+(\chi_k^2-1)s_k \le 1 + \frac{\omega}{n s_k} s_k = 1 + \frac{\omega}{n}$. □

Puhan and Bűrmen [16] write the trial points as $x_c + \rho(x_c - x_{n+1})$, $x_c + \rho\chi(x_c - x_{n+1})$, $x_c + \rho\gamma_{\mathrm{oc}}(x_c - x_{n+1})$ and $x_c + \rho\gamma_{\mathrm{ic}}(x_c - x_{n+1})$. We denote their values from [16, Theorems 1–4, Lemma 2], with the constants of [16, Section 5], by ρ^{PB}, χ^{PB}, $\gamma_{\mathrm{oc}}^{\mathrm{PB}}$, $\gamma_{\mathrm{ic}}^{\mathrm{PB}}$ and σ^{PB}, so that

$\alpha_r = \rho^{PB}$, $\alpha_e = \rho^{PB}\chi^{PB}$, $\alpha_{oc} = \rho^{PB}\gamma^{PB}_{oc}$ and $\alpha_{ic} = -\gamma^{PB}_{ic}$. In particular, $\rho^{PB} = 1 + \frac{1}{2(n+1)}$, $\chi^{PB} = 1 + \frac{2}{n} + \frac{2}{n\rho^{PB}}$ and $\sigma^{PB} = (\frac{9}{10})^{1/n}$, whereas γ^{PB}_{oc} and γ^{PB}_{ic} tend to $\frac{7}{10}$ as $n \to \infty$. For $n = 40$, for example, $\alpha_r = 1.0122$, $\alpha_e = 1.1128$, $\alpha_{oc} = 0.7392$ and $\alpha_{ic} = -0.6878$.

In contrast to the fixed expansion coefficients $\chi = 1 + \frac{2}{n}$ of [7] and $\rho^{PB}\chi^{PB} \approx 1 + \frac{9}{2n}$ of [16], the cap allows long expansions when the replaced vertex carries a small share of Φ. For $\alpha_r = 1$ and contraction coefficients $\pm\gamma$, Proposition 8.1 gives $\vartheta_n \le 1 - \frac{1}{n}(1-\gamma^2)(p_{oc} + p_{ic}) - \omega p_e$ in the exchangeable model (R2). For $\gamma = \frac{1}{2}$ and $\omega = 1$, the dimension-free condition of Corollary 6.3 thus becomes $p_{oc} + p_{ic} > \frac{4}{3}p_e$ instead of $p_{oc} + p_{ic} > 4p_e$.

The choice of the contraction coefficients is decisive. For the coefficients of [7], the contraction coefficient $\gamma_n = \frac{3}{4} - \frac{1}{2n}$ accounts for most of the improvement in our tests: for the ill-conditioned problem of Section 8.4, changing only the contraction coefficient required 6847 ($n = 40$) and 13717 ($n = 60$) function evaluations, compared with 5336 and 13199 for the full set of [7], while the standard method failed. The contraction coefficients of Puhan and Bűrmen [16] tend to $\frac{7}{10}$ instead of $\frac{3}{4}$. By Theorem 4.1, their inside contractions reduce Φ more strongly per step than those with $-\gamma_n$ for $n \ge 4$, and their outside contractions for $n \ge 42$; both $1-\alpha^2_{oc}$ and $1-\alpha^2_{ic}$ tend to 0.51 instead of $\frac{7}{16}$. Algorithm 1 therefore uses the reflection and contraction coefficients of [16], i.e. $\alpha_r = \rho^{PB}$, $\alpha_{oc} = \rho^{PB}\gamma^{PB}_{oc}$ and $\alpha_{ic} = -\gamma^{PB}_{ic}$, combined with the share-capped expansion. The shrink coefficient is not needed, since shrink steps are replaced by oriented restarts (Section 8.2). Since $\rho^{PB} > 1$, we note that for $\chi = 2$ and $\omega \ge 1$ the expansion point $X(\chi_k)$ lies beyond the reflection point, because $\chi_k^2 \ge 1 + \frac{\omega}{n} > 1 + \frac{1}{n+1} + \frac{1}{4(n+1)^2} = (\rho^{PB})^2$. With these coefficients, $\omega = 1$ improved on the uncapped expansion with $\chi_k = \chi = 2$ in all eleven test cases with $n \le 80$, and $\omega \in \{0.25, 0.5, 2\}$ in eight to ten of them. The value $\omega = 0.25$ reduced the effort on the ill-conditioned problem for $n = 40$, 60 and 80 by 12–15%, but increased it on the sphere with $n = 20$ and on the Rosenbrock function with $n = 10$. With the contraction coefficient γ_n, every $\omega \in [0.25, 2]$ improved on the uncapped expansion. We use $\omega = 1$ and $\chi = 2$. For comparison, Section 8.4 also reports the variant with $\alpha_r = 1$ and the contraction coefficient γ_n.

8.2 Sufficient decrease and oriented restart

We use the safeguard of Kelley [9]. For a sorted nondegenerate simplex S let

$$V(S) = [x_2 - x_1, \ldots, x_{n+1} - x_1], \qquad \nabla_S f = V(S)^{-T}\,\big(f(x_2) - f(x_1), \ldots, f(x_{n+1}) - f(x_1)\big)^T$$

be the matrix of simplex directions and the simplex gradient, let $\ell_+(S)$ and $\ell_-(S)$ be the largest and smallest distance $\|x_j - x_1\|$, $j \ge 2$, and let $\bar{f}(S)$ be the mean of the function values at the vertices. Kelley's sufficient decrease condition [9, (3.1)] reads

$$\bar{f}(S^{(k+1)}) - \bar{f}(S^{(k)}) < -\alpha\|\nabla_{S^{(k)}} f\|^2, \qquad \alpha = a_0 \frac{\ell_+(S^{(0)})}{\|\nabla_{S^{(0)}} f\|}, \tag{5}$$

with $a_0 = 10^{-4}$. If the condition fails, the simplex is replaced by an *oriented restart* [9, (3.5)]: a smaller simplex with orthogonal edges at the best vertex,

$$y_1 = x_1, \qquad y_{j+1} = x_1 + \tfrac{1}{2}\ell_-(S)\,\operatorname{sign}\,(\nabla_S f)_j\; e_j \quad (j = 1, \ldots, n),$$

with sign(0) := 1. Its condition number is at most $(1 + \sqrt{n})^2$ [9, Section 3.2]. Thus the shrink coefficient is not needed, since shrink steps are replaced by oriented restarts.

Lemma 8.2. *For a non-shrink step that replaces $x^{(k)}_{n+1}$ by v, we have*

$$\bar{f}(S^{(k+1)}) - \bar{f}(S^{(k)}) = \frac{f(v) - f(x^{(k)}_{n+1})}{n+1}.$$

In particular, (5) *holds if and only if $f(v) < f(x^{(k)}_{n+1}) - (n+1)\,\alpha\,\|\nabla_{S^{(k)}} f\|^2$.*

Proof. The multiset of function values changes only by replacing $f(x_{n+1}^{(k)})$ by $f(v)$. □

Lemma 8.2 is the analogue of Theorem 4.1 for the function values. One iteration changes a single vertex, so Kelley's condition requires the new vertex to improve on the worst one by an amount that grows linearly with n. In our experiments, this leads to false alarms in high dimensions (Table 3). We therefore consider the dimension-free form of the safeguard

$$f(v) < f(x_{n+1}^{(k)}) - \alpha \|\nabla_{S(k)} f\|^2, \tag{6}$$

i.e. (5) with α replaced by $\alpha/(n+1)$, and the relaxed form of the safeguard

$$f(v) < f(x_{n+1}^{(k)}) - \frac{\alpha}{n+1} \|\nabla_{S(k)} f\|^2, \tag{7}$$

i.e. (5) with α replaced by $\alpha/(n+1)^2$. In our tests, (6) still triggered restarts on the sphere from $n = 40$ on, although the standard method converges there without a safeguard; for $n = 300$, these false alarms were frequent, and one run exhausted the budget. Condition (7) triggered no restart on the sphere (Table 3). A heuristic reason is the following. For a reflection, $f(x_{n+1}) - f(X(1)) \approx 2\,\nabla f(x_c)^T (x_{n+1} - x_c)$, and in high dimensions the search direction $x_c - x_{n+1}$ tends to become nearly orthogonal to the steepest descent direction, as suggested by the experiments of Torczon [17] (see also [7, 16]). A single step can therefore improve the worst vertex only slightly even if the method does not stagnate. Algorithm 1 uses (7). Since Kelley's convergence theorem holds for every $\alpha > 0$, these modifications do not affect the theory. Following [9], shrink steps are replaced by oriented restarts, and the iteration is stopped after three consecutive restarts.

Algorithm 1 Nelder–Mead with dimension-dependent coefficients and a sufficient decrease safeguard

Require: nondegenerate simplex $S^{(0)}$; $\chi = 2$, $\omega = 1$, $a_0 = 10^{-4}$; ρ^{PB}, $\gamma_{\mathrm{oc}}^{\mathrm{PB}}$, $\gamma_{\mathrm{ic}}^{\mathrm{PB}}$ of [16]

1: sort the vertices; set $\alpha = a_0\, \ell_+(S^{(0)})/\|\nabla_{S(0)} f\|$ and $r = 0$
2: **for** $k = 0, 1, 2, \ldots$ (until a termination criterion holds) **do**
3: compute x_c, the share s_k and the simplex gradient $g_k = \nabla_{S(k)} f$
4: perform one NM step (Section 2) with the coefficients ρ^{PB}, $\chi_k = \min\{\chi, \sqrt{1 + \omega/(n s_k)}\}$, $\rho^{\mathrm{PB}}\gamma_{\mathrm{oc}}^{\mathrm{PB}}$ and $-\gamma_{\mathrm{ic}}^{\mathrm{PB}}$; let v be the accepted point (none if the NM step calls for a shrink)
5: **if** v exists and $f(v) < f(x_{n+1}^{(k)}) - \frac{\alpha}{n+1}\|g_k\|^2$ **then**
6: replace $x_{n+1}^{(k)}$ by v, sort the vertices, set $r = 0$
7: **else**
8: if v exists, replace $x_{n+1}^{(k)}$ by v and sort the vertices
9: perform an oriented restart of the current simplex and sort the vertices; set $r = r + 1$
10: **if** $r = 3$ **then stop** (stagnation detected)
11: **end if**
12: **end for**

8.3 Convergence

Combining Kelley's theorem with Theorem 4.5 yields convergence of all vertices to a single critical point, under conditions that can be checked along a run.

Theorem 8.3. *Let $f : \mathrm{R}^n \to \mathrm{R}$ be bounded below with Lipschitz continuous gradient. Suppose that Algorithm 1 does not stop and performs only finitely many oriented restarts, and that*

$$\sum_{k \geq 0} \sqrt{\Phi_k} < \infty \qquad \text{and} \qquad \sqrt{\Phi_k}\, \mathrm{cond}\, V(S^{(k)}) \to 0.$$

Then all vertices converge to a common limit $\hat{x}$, and $\nabla f(\hat{x}) = 0$.

Proof. Let k_0 be the first iteration after the last restart. For $k \geq k_0$, every iteration accepts a point v satisfying (7). By Lemma 8.2, the sequence $(S^{(k)})_{k \geq k_0}$ then satisfies (5) with the constant $\alpha/(n+1)^2 > 0$, and the mean function values decrease strictly. The simplices are sorted and remain nondegenerate, since $T(\lambda)$ is nonsingular for $\lambda \neq 0$. Hence Assumption 2.1 of [9] holds. By Lemma 3.2(b), $\ell_+(S^{(k)}) \leq \operatorname{diam}(S^{(k)}) \leq \sqrt{2\Phi_k}$, so that $\ell_+(S^{(k)}) \operatorname{cond}(V(S^{(k)})) \to 0$. By [9, Theorem 3.1], every accumulation point of the simplices is a critical point of f.

Since $\chi_k \leq \chi$, $\rho^{\mathrm{PB}} < \chi$ and $|\rho^{\mathrm{PB}}\gamma_{\mathrm{oc}}^{\mathrm{PB}}, \gamma_{\mathrm{ic}}^{\mathrm{PB}}| < \chi$, all step coefficients satisfy $|\alpha_k| \leq \chi$, and Theorem 4.5 applies to $(S^{(k)})_{k \geq k_0}$. Hence the vertices converge to a common limit $\hat{x}$, which is an accumulation point of the simplices and therefore critical. □

Remark 8.4. All hypotheses of Theorem 8.3 can be monitored along a run: the restarts are counted, Φ_k is available at cost $O(n)$ by Corollary 4.4, and $\operatorname{cond}(V(S^{(k)}))$ can be estimated periodically. The excluded case of infinitely many restarts corresponds to repeated detection of stagnation; there the algorithm stops after three consecutive restarts. The simplex gradient requires one linear solve with $V(S^{(k)})$ per iteration, which costs $O(n^3)$ operations. Since $\nabla_S f$ is the gradient of the affine interpolant of f at the vertices, it does not depend on the order of the vertices. It can therefore be obtained from the inverse of the matrix with rows $(x_j^T, 1)$, $j = 1, \ldots, n+1$, which changes by a rank-one matrix in each NM step; with the Sherman–Morrison formula, the cost is $O(n^2)$ per iteration.

8.4 Numerical results

As a test of the implementation, we applied the method of [11] with Kelley's original condition (5) to McKinnon's examples, with the initial simplex and termination criterion of [9, Section 4]. For $(\tau, \vartheta, \phi) = (3, 6, 400)$ and $(2, 6, 60)$, the single oriented restart occurs at iterations 21 and 19, exactly as reported in [9]. For the nonsmooth case $(1, 15, 10)$, the three restarts occur at iterations 40–42 instead of 44–46; this offset is presumably due to implementation details. Table 2 shows the results of Algorithm 1 and of its variant with $\alpha_r = 1$ and γ_n, which for $n = 2$ uses the standard contraction coefficient $\gamma_2 = \frac{1}{2}$. McKinnon's construction is tailored to the standard coefficients. With the coefficients of [16] for $n = 2$ ($\rho^{\mathrm{PB}} = \frac{7}{6}$, $\chi^{\mathrm{PB}} = \frac{20}{7}$, $\gamma_{\mathrm{oc}}^{\mathrm{PB}} \approx 0.921$, $\gamma_{\mathrm{ic}}^{\mathrm{PB}} \approx 0.570$, $\sigma^{\mathrm{PB}} \approx 0.949$) and without a safeguard, the method reached the minimizer for $(3, 6, 400)$ and $(2, 6, 60)$, but it stagnated at $(0, 0)$ for $(1, 15, 10)$. Algorithm 1 reaches the minimizer in the first two examples without any restart, and in the nonsmooth example its safeguard detects the stagnation. Dimension-dependent coefficients alone therefore do not make a safeguard superfluous.

Table 2: McKinnon's examples, $f(x, y) = \vartheta\phi|x|^\tau + y + y^2$ for $x \leq 0$ and $\vartheta x^\tau + y + y^2$ for $x > 0$, with the minimizer $(0, -\frac{1}{2})$ and $f^* = -\frac{1}{4}$. The standard method converges to the nonstationary point $(0, 0)$. Restarts: iterations of the oriented restarts (–: none); stag.: stagnation detected.

(τ, ϑ, ϕ)	standard NM	Algorithm 1		Alg. 1 with $\alpha_r = 1$, γ_n	
	final f	restarts	final f	restarts	final f
(3, 6, 400)	0	–	−0.25	25	−0.25
(2, 6, 60)	0	–	−0.25	21	−0.25
(1, 15, 10)	0	63, 64, 65	$-6.7 \cdot 10^{-5}$ (stag.)	53, 54, 55	0 (stag.)

Table 3 compares Algorithm 1 with the standard method, with the standard method combined with each of the three forms (5)–(7) of the safeguard, with the method of [7], with the coefficients of [16] in the method of [11] without safeguard, and with the variant of Algorithm 1 with $\alpha_r = 1$ and the contraction coefficient γ_n of [7]. The test functions are the sphere, the

ill-conditioned quadratic and the extended Rosenbrock function:

$$\text{sphere: } f(x) = \sum_{i=1}^{n} x_i^2,$$

$$\text{ill-conditioned: } f(x) = \sum_{i=1}^{n} i x_i^2,$$

$$\text{Rosenbrock: } f(x) = \sum_{i=1}^{n-1} 100\ (x_{i+1} - x_i^2)^2 + (1 - x_i)^2\ .$$

All three have the minimal value $f^* = 0$, attained at $x^* = 0$ for the sphere and the ill-conditioned problem and at $x^* = (1, \ldots, 1)^T$ for the Rosenbrock function. The Hessian of the ill-conditioned problem has the condition number n. For $n = 10$, 20 and 30, the Rosenbrock function also has a local minimizer near $(-1, 1, \ldots, 1)^T$, with $f \approx 3.9866$. The initial simplex is $x_0, x_0 + e_1, \ldots, x_0 + e_n$, where the three runs use three random starting points x_0 drawn from the standard normal distribution. The target is $f \leq 10^{-8}$, with a budget of $3 \cdot 10^5$ function evaluations. For $n = 300$ we report only the sphere in the table. For the ill-conditioned problem with $n = 300$, none of the methods with dimension-dependent coefficients reached the target in at least two of the three runs. The variant of Algorithm 1 with γ_n and the method of [7] ended with final values between 0.14 and 35, the method of [16] with values between 0.17 and 4.9, whereas Algorithm 1 came close to the target (see below). We did not consider the Rosenbrock function in this dimension. The runs with $n \geq 100$ use the $O(n^2)$ implementation of the simplex gradient described in Remark 8.4. It reproduces the evaluation counts and restarts of all runs of Algorithm 1 with $n \leq 80$ exactly, and of four runs with $n = 100$ and $n = 150$ that we checked, among them two runs of the standard method with condition (5), with 11 and 13 restarts.

Table 3: Median number of function evaluations to reach $f \leq 10^{-8}$ over three runs; in brackets, the median number of oriented restarts. "fail": the target was not reached in at least two runs. [a]Budget exhausted in two runs ($n \leq 80$ and $n = 150$) or in all runs ($n = 100$), with final values between $1 \cdot 10^{-8}$ and $7 \cdot 10^{-7}$ for $n \leq 100$ and of $3.2 \cdot 10^{-6}$ and $2.3 \cdot 10^{-4}$ for $n = 150$. [b]All runs were stopped by the safeguard far from the minimizer, with final values between 0.18 and 1290 for the ill-conditioned problem and between 0.033 and 184 for the sphere. [c]One run exhausted the budget, with final value $3.4 \cdot 10^{-6}$. [d]One run exhausted the budget, with final value $1.01 \cdot 10^{-8}$. [e]No run reached the target: two runs exhausted the budget, and in one run all vertices attained the same function value; final values between 17.3 and 27.8. Alg. 1, γ_n: Algorithm 1 with $\alpha_r = 1$ and the contraction coefficient γ_n of [7].

f	n	standard	+ (5)	+ (6)	+ (7)	[7]	[16]	Alg. 1, γ_n	Alg. 1
sphere	20	1624	1646 [1]	1624 [0]	1624 [0]	1690	1489	1599 [0]	1342 [0]
sphere	40	10090	4026 [3]	10090 [0]	10090 [0]	4303	3773	3796 [0]	3279 [0]
sphere	60	16903	13294 [9]	16903 [0]	16903 [0]	8501	6757	6141 [0]	5689 [0]
sphere	80	29932	31592 [10]	29202 [1]	29932 [0]	13050	10676	9408 [0]	8979 [0]
sphere	100	42924	29785 [14]	42924 [1]	42924 [0]	17174	15751	12114 [0]	11033 [0]
sphere	300	208669	fail[b] [19]	218316[c] [9]	208669 [0]	152359	125884	156007 [0]	128630 [0]
ill-cond.	20	3557	3353 [1]	3557 [0]	3557 [0]	1912	1683	1778 [0]	1652 [0]
ill-cond.	40	fail[a]	16310 [14]	21871 [1]	61816 [1]	5336	4947	4678 [0]	5011 [0]
ill-cond.	60	fail[a]	fail[b] [33]	35851 [1]	182046 [1]	13199	11753	8690 [0]	10252 [0]
ill-cond.	80	fail[a]	fail[b] [17]	82558 [1]	216815 [1]	24868	21353	15393 [0]	18749 [0]
ill-cond.	100	fail[a]	fail[b] [16]	79704 [1]	289035[d] [1]	46998	35178	22242 [0]	26289 [0]
ill-cond.	150	fail[a]	fail[b] [16]	266227 [4]	fail[a] [0]	118942	91321	57555 [0]	67087 [0]
Rosenbrock	10	2921	2921 [0]	2921 [0]	2921 [0]	2698	2152	2406 [0]	2147 [0]
Rosenbrock	20	106043	22505 [8]	36157 [4]	76591 [5]	25615	12355	20071 [0]	13185 [0]
Rosenbrock	30	fail[e]	62257 [13]	159727 [9]	245690 [9]	107919	39641	91248 [0]	41148 [0]

Table 3 leads to five observations.

Kelley's original condition. Combined with the standard method, condition (5) often helps, e.g. for the sphere with $n = 40$ and $n = 100$ and the Rosenbrock function with $n = 20$ and $n = 30$, because the restarts recondition the simplex. For the ill-conditioned problem with $n \geq 60$ and for the sphere with $n = 300$, however, it stops all runs far from the minimizer. This is the effect of the factor $n + 1$ in Lemma 8.2.

Conditions (6) *and* (7)*.* Condition (6) rescues the standard method on the ill-conditioned problem with a median of one restart for $n = 40$, 60, 80 and 100 and of four restarts for $n = 150$, and it reduces the effort for the Rosenbrock function with $n = 20$ from 106043 to 36157 evaluations and rescues the standard method for $n = 30$. On the sphere, where the standard method converges without a safeguard, it triggered one restart in one of the three runs with $n = 40$ and $n = 60$ and in two of the three runs with $n = 80$ and $n = 100$. Each of these runs needed fewer evaluations than without a safeguard, so that these false alarms did no harm. With $n = 300$, however, it triggered 7 to 11 restarts per run, and one run exhausted the budget. Condition (7) triggered no restart on the sphere for any n, including $n = 300$. It still rescues the stagnating runs of the standard method, but later; for the ill-conditioned problem with $n = 100$, one run exhausted the budget with $f = 1.01 \cdot 10^{-8}$, and for $n = 150$ it triggered no restart, so that, as without a safeguard, two runs exhausted the budget. For the ill-conditioned problem, the median effort is 61816, 182046, 216815 and 289035 evaluations for $n = 40$, 60, 80 and 100, compared with 21871, 35851, 82558 and 79704 with (6), which needed 266227 for $n = 150$; for the Rosenbrock function with $n = 20$ and 30, it is 76591 and 245690 instead of 36157 and 159727. Combined with the standard method, (7) detects the stagnation on McKinnon's examples a few iterations later than (6): the restarts occur at iterations 25, 21 and 53–55 instead of 23, 20 and 46–48. The two forms thus represent a trade-off between avoiding false alarms and detecting genuine stagnation early.

Algorithm 1. Up to $n = 150$, no restart was triggered on these problems, so the safeguard cost no function evaluations. Since (7) holds whenever (6) holds, the runs with $n \leq 100$ are the same for both forms; for the ill-conditioned problem with $n = 150$, (6) triggered two restarts in one run. Algorithm 1 needed fewer evaluations than the method of [7] in all fifteen cases, between 6% (ill-conditioned, $n = 40$) and 44% (ill-conditioned, $n = 100$) fewer on the quadratics and 20%, 49% and 62% fewer on the Rosenbrock function with $n = 10$, 20 and 30. For the sphere with $n = 300$, it needed a median of 128630 evaluations without any restart, and all runs reached the target. With (6) instead of (7), it needed 146761 evaluations, its safeguard triggered 1 to 11 restarts per run, and one run exhausted the budget with $f = 1.7 \cdot 10^{-3}$. For the ill-conditioned problem with $n = 300$, Algorithm 1 reached the target in one of three runs, after 276731 evaluations, and ended with $f = 1.9 \cdot 10^{-8}$ and $3.8 \cdot 10^{-8}$ in the other two, again without any restart. On McKinnon's examples, the safeguard is triggered only in the nonsmooth case, where it detects the stagnation (Table 2).

Coefficients of [16]. Without any safeguard, the coefficients of Puhan and Bűrmen needed fewer evaluations than those of [7] in all fifteen cases: 7–25% fewer on the quadratics and 20%, 52% and 63% fewer on the Rosenbrock function with $n = 10$, 20 and 30. With the safeguard (7), they triggered no restart for $n \leq 80$. Since neither method triggered a restart, Algorithm 1 differs from them in these runs only in the share-capped expansion; the same holds on the quadratics with $n = 100$ and $n = 150$, since Algorithm 1 triggered no restart and no shrink steps occur for strictly convex functions. The share-capped expansion reduced the effort in eleven of these fourteen cases, by up to 30%, and increased it for the ill-conditioned problem with $n = 40$ (1%) and the Rosenbrock function with $n = 20$ and 30 (7% and 4%). For the sphere with $n = 300$, Algorithm 1 needed 2% more evaluations than the method of [16].

Contraction coefficients. The variant of Algorithm 1 with $\alpha_r = 1$ and γ_n needed 7%, 15%, 18%, 15% and 14% fewer evaluations than Algorithm 1 on the ill-conditioned problem with $n = 40$, 60, 80, 100 and 150. It needed 5–21% more on the other quadratics and 12%, 52% and 122% more on the Rosenbrock function with $n = 10$, 20 and 30. For the ill-conditioned problem

with $n = 300$, it remained far from the target, whereas Algorithm 1 reached it or came close. By Theorem 4.1, the inside contractions of [16] reduce Φ more strongly per step than those with γ_n (Section 8.1); why this is a disadvantage on the ill-conditioned problem for $40 \le n \le 150$ remains open.

These experiments cover three test functions and three runs each. A systematic comparison on standard test sets such as that of Moré, Garbow and Hillstrom, as in [7], remains to be done.

9 A nonconvex least-squares problem

The test functions of Section 8.4 are convex quadratics and the Rosenbrock function. As a nonconvex problem, we consider the least-squares approximation of t^2 on $[0, 1]$ by a neural network with one hidden layer of m units,

$$f(x) = \sum_{j=1}^{N} \left(t_j^2 - \sum_{k=1}^{m} x_{3k-2} \tanh\left(t_j x_{3k-1} + x_{3k} \right) \right)^2, \qquad t_j = \frac{j}{N}, \tag{8}$$

with $N = 1000$ and $m = 10$, so that $n = 3m = 30$. The variables x_{3k-2}, x_{3k-1} and x_{3k} are the output weight, the input weight and the bias of the kth unit. Since $\tanh s = 2\sigma(s) - 1$ with the continuous sigmoidal function $\sigma(s) = \frac{1}{2}(1 + \tanh s)$, and since every constant is of the form $a \tanh b$, the universal approximation theorem of Cybenko [1, Theorem 2] shows that finite sums $\sum_k a_k \tanh(w_k t + b_k)$ are dense in $C([0, 1])$. Hence the infimum of f tends to zero as $m \to \infty$. The function is smooth and nonnegative, but not convex, and its minimizers are not isolated: f is invariant under permutations of the units and under the sign change $(x_{3k-2}, x_{3k-1}, x_{3k}) \mapsto (-x_{3k-2}, -x_{3k-1}, -x_{3k})$ of a unit, and it does not depend on x_{3k-1} and x_{3k} if $x_{3k-2} = 0$. For $m = 10$, the minimal value is already practically zero: from each of 20 random starting points, the Levenberg–Marquardt method with the exact Jacobian (SciPy, at most 2000 evaluations of the residual) reached $f \le 1.3 \cdot 10^{-9}$, and in two cases $f < 10^{-12}$.

We use the methods, initial simplices and budget of Section 8.4, but ten random starting points instead of three, since the results depend strongly on the starting point. Instead of a single target, Table 4 reports the median number of function evaluations to reach $f \le 10^{-2}$, 10^{-4}, 10^{-6}, 10^{-7} and 10^{-8}, the median final value after $3 \cdot 10^5$ evaluations, and the median numbers of oriented restarts and of shrink steps. Two further variants show the effect of the safeguard: the coefficients of [16] and of [7] with the safeguard (7). With a safeguard, no shrink steps occur. All runs with a safeguard use the $O(n^2)$ implementation of the simplex gradient (Remark 8.4); for one run of Algorithm 1, the exact gradient gave identical results.

Table 4 leads to four observations.

Shrink steps. Except for the coefficients of [7], all methods without safeguard performed shrink steps in almost every run, typically thousands. The standard method reached $f \le 10^{-2}$ with the fewest evaluations, but it performed between 3521 and 8851 shrink steps per run, reached $f \le 10^{-6}$ in only one run and ended with the largest median final value. The coefficients of [16] reached $f \le 10^{-4}$ with the fewest evaluations, but they performed shrink steps in nine of ten runs, reached $f \le 10^{-6}$ in only six runs, and their median final value is more than a thousand times larger than that of [7]. The coefficients of [7] performed at most three shrink steps per run. They reached $f \le 10^{-6}$, 10^{-7} and 10^{-8} in all ten runs and gave the smallest final values, smaller than those of Algorithm 1 and of [16] in nine of ten runs.

Safeguards for the standard method. All three forms of the safeguard improved the standard method. With (7), the final value was smaller in all ten runs, and the median final value is about 140 times smaller; with (5) and (6), the final value was smaller in eight and nine runs, respectively. In this dimension, Kelley's original condition did not stop any run, although it triggered between 108 and 152 restarts per run.

Algorithm 1. Algorithm 1 reached $f \le 10^{-6}$ in all ten runs and $f \le 10^{-7}$ in nine. Its final value was smaller than that of [16] in eight of ten runs, but larger than that of [7] in nine. With

Table 4: Problem (8) with $N = 1000$ and $n = 30$: median number of function evaluations over ten runs to reach $f \le 10^{-2}, \ldots, 10^{-8}$, median final value after $3 \cdot 10^5$ evaluations, and median numbers of oriented restarts and of shrink steps. In parentheses, the number of runs that reached the threshold if it is smaller than ten; "–": fewer than six runs reached the threshold, so that the median exceeds the budget.

method	10^{-2}	10^{-4}	10^{-6}	10^{-7}	10^{-8}	final f	restarts	shrinks
standard	1938	15882 (7)	– (1)	– (1)	– (0)	$8.7 \cdot 10^{-6}$	0	8101
+ (5)	2582	15401	221149 (7)	– (3)	– (0)	$4.2 \cdot 10^{-7}$	124	0
+ (6)	2394	12208	138776 (9)	– (3)	– (1)	$2.0 \cdot 10^{-7}$	88	0
+ (7)	2300	10332	133462 (9)	267186 (6)	– (1)	$6.1 \cdot 10^{-8}$	68	0
[7]	3224	10642	47450	98139	126469	$5.1 \cdot 10^{-10}$	0	0
[16]	2592	8540	180690 (6)	– (4)	– (1)	$6.0 \cdot 10^{-7}$	0	2600
Alg. 1, γ_n	3160	12286	65560 (9)	189176 (6)	289318 (6)	$8.1 \cdot 10^{-9}$	32	0
Alg. 1	2526	10102	62639	219232 (9)	– (3)	$3.9 \cdot 10^{-8}$	39	0
[16] + (7)	2450	9596	45663 (8)	– (5)	– (1)	$1.2 \cdot 10^{-7}$	30	0
[7] + (7)	3548	12263	82178 (9)	271614 (6)	– (3)	$4.4 \cdot 10^{-8}$	14	0

the contraction coefficient γ_n, the median final value is smaller, but the final value itself was smaller in only five of ten runs. The safeguard triggered between 31 and 43 restarts per run. Of the 379 restarts, 377 were caused by an insufficient decrease rather than by a shrink step, and they recurred at intervals of roughly 4600 to 7500 iterations (10% and 90% quantiles) until the budget was exhausted.

Safeguard and shrink steps. For the coefficients of [16], the safeguard (7) improved the final value in only five of ten runs, but it led to $f \le 10^{-6}$ in eight instead of six runs. For the coefficients of [7], which almost never call for a shrink step, the safeguard was harmful: it triggered between 12 and 17 restarts per run, the final value was smaller than without safeguard in only one of ten runs, and $f \le 10^{-8}$ was reached in three instead of ten runs. Together with the results for the standard method, this suggests that the shrink steps are the main obstacle on this problem and that the safeguards help mainly by avoiding them. This agrees with the observation of Zhao et al. [19] that the classical shrink step is a main cause of failures of the NM method; they move the vertices toward the best vertex one at a time instead. Why the coefficients of [7] almost never call for a shrink step remains open.

10 Conclusions

In centered coordinates the Nelder–Mead operations have a simple structure, and the moment of inertia of the simplex obeys the exact identity $\Phi_{k+1} = \Phi_k \, 1 + (\alpha_k^2 - 1)s_k$. This yields a deterministic convergence criterion that can be monitored at cost $O(n)$ per iteration, and it removes the dimensional restriction of [5]. Theorem 2 of [5] holds in every dimension, but in the underlying uniform model it is a statement about shrink steps, which do not occur for strictly convex functions. If the rank of the replaced vertex is uniformly distributed, second-moment criteria based on the identity give almost sure convergence in every dimension also without shrinks, under dimension-free conditions and with the rate $1 - c/n$ per iteration. For the insertion model of [5], almost sure convergence follows from a spectral-radius condition, which holds in every dimension; the moment of inertia serves as a Lyapunov function in the proof. Since the identity does not depend on the labelling of the vertices, the deterministic criterion and the exchangeable model apply equally to the original method of Nelder and Mead; in its uniform random model, the vertices converge almost surely for $n \ge 3$. Because one iteration changes a single vertex, Kelley's sufficient decrease condition is overly strict for large n (Lemma 8.2); dividing its constant by $(n + 1)^2$ removed the false alarms in our tests. Algorithm 1 combines

this form of the condition with the reflection and contraction coefficients of Puhan and Bűrmen and a share-capped expansion. On McKinnon's examples, it reached the minimizer in the two smooth cases and detected the stagnation in the nonsmooth one. It required fewer function evaluations than the method of Gao and Han in all fifteen tests of Section 8.4 and than the coefficients of Puhan and Bűrmen alone in eleven; the difference to the latter is due to the share-capped expansion. On the ill-conditioned problem, the contraction coefficient of Gao and Han was more efficient for $40 \le n \le 150$, whereas for $n = 300$ only Algorithm 1 came close to the target. On the nonconvex least-squares problem of Section 9, the method of Gao and Han reached the smallest function values. The other methods without safeguard performed shrink steps in almost every run, typically thousands, and the safeguards, under which no shrink steps occur, improved the final values of the standard method in all ten runs, but worsened those of the method of Gao and Han in nine; the restarts of Algorithm 1 recurred at regular intervals.

Several questions remain open:

1. an explanation of the results of Table 3 for the variant of Algorithm 1 with $\alpha_r = 1$ and the contraction coefficient γ_n of [7]: it needed 7–18% fewer evaluations than Algorithm 1 on the ill-conditioned problem for $40 \le n \le 150$, but 5–21% more on the other quadratics and 12–122% more on the Rosenbrock function, and for $n = 300$ it remained far from the target on the ill-conditioned problem (Section 8.4); a choice of coefficients that combines the advantages of both; and an explanation of why the coefficients of [7] almost never call for a shrink step on the problem of Section 9, whereas the other coefficients call for thousands;

2. a priori conditions under which Algorithm 1 performs only finitely many restarts and $\sqrt{\Phi_k}\,\mathrm{cond}(V(S^{(k)})) \to 0$, which would make Theorem 8.3 unconditional; on the problem of Section 9, the restarts recurred until the budget was exhausted. The identity controls the size of the simplex, but not its condition. Frame-based variants [15] and fortified descent [18] offer alternative routes to stationarity, as suggested by Gao and Han [7] for their coefficients;

3. a theoretical justification of the scaling of α by $(n + 1)^{-2}$ in (7), which removed the false alarms in our tests up to $n = 300$ but delays the detection of genuine stagnation.

Remark. All numerical results were obtained with Python (NumPy/SciPy); the spectral radii in Table 1 were computed by Arnoldi iteration. For the coefficients of [16], the implementation was extended to separate coefficients for the four trial points. The scripts are available on request. The manuscript was developed with the assistance of Claude in theorem formulation, proofs and simulation under the author's verification.

References

[1] G. Cybenko, Approximation by superpositions of a sigmoidal function, *Math. Control Signals Systems* 2 (1989) 303–314, https://doi.org/10.1007/BF02551274.

[2] A. Galántai, Convergence theorems for the Nelder–Mead method, *J. Comput. Appl. Mech.* 15 (2020) 115–133.

[3] A. Galántai, A convergence analysis of the Nelder–Mead simplex method, *Acta Polytech. Hung.* 18 (2021) 93–105.

[4] A. Galántai, Convergence of the Nelder–Mead method, *Numer. Algorithms* 90 (2022) 1043–1072.

[5] A. Galántai, A stochastic convergence result for the Nelder–Mead simplex method, *Mathematics* 11 (2023) 1998, https://doi.org/10.3390/math11091998.

[6] A. Galántai, The Nelder–Mead simplex algorithm is sixty years old: new convergence results and open questions, *Algorithms* 17 (2024) 523, https://doi.org/10.3390/a17110523.

[7] F. Gao, L. Han, Implementing the Nelder–Mead simplex algorithm with adaptive parameters, *Comput. Optim. Appl.* 51 (2012) 259–277, https://doi.org/10.1007/s10589-010-9329-3.

[8] L. Han, M. Neumann, Effect of dimensionality on the Nelder–Mead simplex method, *Optim. Methods Softw.* 21 (2006) 1–16.

[9] C. T. Kelley, Detection and remediation of stagnation in the Nelder–Mead algorithm using a sufficient decrease condition, *SIAM J. Optim.* 10 (1999) 43–55.

[10] J. C. Lagarias, B. Poonen, M. H. Wright, Convergence of the restricted Nelder–Mead algorithm in two dimensions, *SIAM J. Optim.* 22 (2012) 501–532.

[11] J. C. Lagarias, J. A. Reeds, M. H. Wright, P. E. Wright, Convergence properties of the Nelder–Mead simplex method in low dimensions, *SIAM J. Optim.* 9 (1998) 112–147.

[12] J. Larson, M. Menickelly, S. M. Wild, Derivative-free optimization methods, *Acta Numer.* 28 (2019) 287–404.

[13] K. I. M. McKinnon, Convergence of the Nelder–Mead simplex method to a nonstationary point, *SIAM J. Optim.* 9 (1998) 148–158.

[14] J. A. Nelder, R. Mead, A simplex method for function minimization, *Comput. J.* 7 (1965) 308–313.

[15] C. J. Price, I. D. Coope, D. Byatt, A convergent variant of the Nelder–Mead algorithm, *J. Optim. Theory Appl.* 113 (2002) 5–19.

[16] J. Puhan, Á. Bűrmen, On the optimal parameter values of the Nelder–Mead simplex algorithm, *J. Comput. Appl. Math.* 484 (2026) 117521, https://doi.org/10.1016/j.cam.2026.117521.

[17] V. Torczon, *Multi-directional Search: A Direct Search Algorithm for Parallel Machines*, Ph.D. thesis, Rice University, Houston, TX, 1989.

[18] P. Tseng, Fortified-descent simplicial search method: a general approach, *SIAM J. Optim.* 10 (2000) 269–288.

[19] Q. H. Zhao, D. Urošević, N. Mladenović, P. Hansen, A restarted and modified simplex search for unconstrained optimization, *Comput. Oper. Res.* 36 (2009) 3263–3271, https://doi.org/10.1016/j.cor.2009.03.005.