J Optim Theory Appl (2013) 159:805–821
DOI 10.1007/s10957-013-0356-y

# Using QR Decomposition to Obtain a New Instance of Mesh Adaptive Direct Search with Uniformly Distributed Polling Directions

**Benjamin Van Dyke · Thomas J. Asaki**

Received: 2 July 2012 / Accepted: 4 June 2013 / Published online: 27 June 2013


**Abstract** The purpose of this paper is to introduce a new instance of the Mesh Adaptive Direct Search (Mads) class of algorithms, which utilizes a more uniform distribution of poll directions than do other common instances, such as OrthoMads and LtMads. Our new implementation, called QrMads, bases its poll directions on an equal area partitioning of the $n$-dimensional unit sphere and the QR decomposition to obtain an orthogonal set of directions. While each instance produces directions which are dense in the limit, QrMads directions are more uniformly distributed in the unit sphere. This uniformity is the key to enhanced performance in higher dimensions and for constrained problems. The trade-off is that QrMads is no longer deterministic and at each iteration the set of polling directions is no longer orthogonal. Instead, at each iteration, the poll directions are only 'nearly orthogonal,' becoming increasingly closer to orthogonal as the mesh size decreases. Finally, we present a variety of test results on smooth, nonsmooth, unconstrained, and constrained problems and compare them to OrthoMads on the same set of problems.

**Keywords** Mesh adaptive direct search (Mads) algorithms · Derivative-free optimization

## 1 Introduction

In [1], Audet and Dennis, introduced the Mesh Adaptive Direct Search (Mads) class of algorithms for constrained, black-box optimization problems where derivative information for the objective function is either unavailable or unreliable and the set of feasible points is determined by black-box constraint functions. In most cases the objective function may not be smooth. Depending on assumptions made about the

B. Van Dyke (✉) · T.J. Asaki
Department of Mathematics, Washington State University, Pullman, WA 99164, USA
e-mail: bvandyke@math.wsu.edu

objective function, however, Mads is shown to converge to both first-order [1] and second-order [2] stationary points in the Clarke sense [3], as well as, first-order [4] stationary points in the Rockafellar sense [5].

Mads is a Search/Poll-based direct search method that was introduced as a means of extending the Gps [6] class of algorithms by allowing for an infinite number of polling directions (see Definition 2.1), in the limit, when exploring the space of variables. The first instance of Mads called LtMads [1] relied on a random lower triangular matrix to generate the set of polling directions and probabilistic arguments to obtain convergence results. This instance was later shown to have undesirably large angles between some of the poll directions at each iteration.

To eliminate these large angles and the need for probabilistic arguments, as well as make modifications for improved performance, a second instance of Mads called OrthoMads [7] was introduced. This instance is deterministic and uses a maximal orthogonal basis at each iteration for the set of polling directions. OrthoMads has been shown to behave well in practice, but for larger $n$ (for example, $n \geq 20$), it has been observed that the set of normalized directions taken over all iterations tend to cluster around the coordinate directions, so that a small set of directions is slightly favored over all other possible directions. This behavior was not apparent in [7], as that paper only looked at the distribution of the polling directions for $n = 2, 3$. In fact, we can observe in Fig. 2 that, as $n$ increases, the clustering behavior becomes more pronounced and larger proportions of the directions cluster around the coordinate directions.

The purpose of this paper is to introduce a new instance of Mads, which we call QrMads, that eliminates the clustering of directions. To accomplish this goal we use an equal area partition of the unit sphere and QR decomposition to find an orthogonal set of directions at each iteration. The orthogonal set of directions is then projected onto a hypercube centered at the origin with integer length and rounded to obtain a 'nearly orthogonal' integer set of directions. It can be observed that the set of normalized directions taken over the set of all iterations from QrMads obtains a more uniform distribution, as seen in Fig. 2.

There are trade-offs to this approach. First, QrMads is no longer deterministic. Second, at each iteration the set of polling directions is no longer orthogonal, but is instead only 'nearly orthogonal' becoming increasingly closer to orthogonal as the mesh size decreases. We believe that this trade-off is minor and that maintaining orthogonality is not a necessary feature, since in [7] it was only applied in an effort to achieve better directions than those generated by LtMads. Third, there is a higher computational cost associated with performing a QR decomposition to obtain a set of polling directions. For this reason, our target application is directed towards expensive objective functions where the QR decomposition will not significantly contribute to the overall solve time.

Sections 2 and 3 give the method used by QrMads for constructing a polling set on the current mesh and consisting of 'nearly orthogonal' polling directions. In Sect. 4, we outline the QrMads algorithm and show that it is a valid Mads instance that shares the theoretical convergence results of this class of algorithms. Finally, we give numerical results to compare QrMads with OrthoMads on a set of 60 smooth, 62 nonsmooth, and 28 constrained test problems taken from the optimization literature.

## 2 Mesh Adaptive Direct Search

The Mads class of algorithms was developed to solve black-box optimization problems of the form

$$\min_{x\in\Omega} f(x),$$

where $\Omega$ is the set of feasible points. Each iteration of Mads is characterized by two steps, an optional Search and a local Poll, performed on a mesh whose fineness must converge to zero in the limit. At each iteration, the mesh $M_k$ is defined by

$$M_k := \left\{x + \Delta_k^m Dz : x \in V_k,\ z \in \mathbb{N}^{n_D}\right\}, \tag{1}$$

where $V_k$ is the set of all evaluated points at the start of iteration $k$, $\Delta_k^m$ is the mesh size parameter at iteration $k$, and $D \in \mathbb{R}^{n\times n_D}$ is a matrix with $n_D$ directions in $\mathbb{R}^n$. The Poll step is executed if the Search step fails to find a lower objective function value than the current best solution. The Poll step evaluates mesh points adjacent to the current best solution in directions that form a positive spanning set for $\mathbb{R}^n$. The Poll step can be characterized by the following definition taken from [1]:

**Definition 2.1** At iteration $k$, the set of poll trial points for the Mads algorithm is defined to be:

$$P_k := \left\{x_k + \Delta_k^m d : d \in D_k\right\} \subset M_k,$$

where $D_k$ is a positive spanning set such that $0 \notin D_k$ and for each $d \in D_k$,

- $d$ can be written as a non-negative integer combination of the directions of $D$,
- the distance in $\ell_\infty$-norm from the poll center $x_k$ to a poll trial point is bounded above by the poll size parameter $\Delta_k^p$,
- limits (as defined by Coope and Price [8]) of the normalized sets $D_k$ are positive spanning sets.

From [1] we find that the poll and mesh size parameters produced by a Mads instance satisfy

$$\liminf_{k\to+\infty} \Delta_k^p = \liminf_{k\to+\infty} \Delta_k^m = 0. \tag{2}$$

The OrthoMads variant of Mads uses the Halton sequence [9] and the scaled Householder transformation, $H = \|q\|^2(I_n - 2vv^T)$ with $q \in \mathbb{R}^n$ and $v = q/\|q\|$, to construct an orthogonal basis for $\mathbb{R}^n$ at each iteration. It defines

$$\Delta_k^p = 2^{-\ell_k} \quad \text{and} \quad \Delta_k^m = \min\left\{1, 4^{-\ell_k}\right\}, \tag{3}$$

where $\ell_k$ is the mesh index at iteration $k$, with $\ell_0 = 0$ at the initial iteration.

For high dimensions, the directions produced by OrthoMads tend to cluster around the coordinate directions. The primary reason for this seems to be that the use of the

Householder transformation,

$$H := \|q\|^2\left(I_n - 2vv^T\right) = \|q\|^2 \begin{bmatrix} 1-2v_1^2 & -2v_1v_2 & \dots & -2v_1v_n \\ -2v_1v_2 & 1-2v_2^2 & \dots & -2v_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ -2v_1v_n & -2v_2v_n & \dots & 1-2v_n^2 \end{bmatrix},$$

will cause a higher concentration of directions to occur near the axes. This is the case because if any of the coordinates of the vector $v$ are small, then the corresponding column of the resulting matrix $H$ will tend to have small coordinates off the diagonal and be nearly one on the diagonal. This behavior becomes more pronounced in high dimensions (and occurs in all the columns) since the coordinates of most unit vectors will naturally be small.

## 3 Equal Area Partition of the Unit Sphere

In [10] and [11], Leopardi introduced the recursive zonal equal area (EQ) sphere partitioning algorithm for partitioning the unit sphere $\mathbb{S}^{n-1} \subset \mathbb{R}^n$ into regions of equal area and small diameter. An equal area partition $P$ is defined to be a nonempty finite set of closed Lebesgue measurable subsets $R$ of $\mathbb{S}^{n-1}$, called regions, such that $\bigcup_{R\in P} R = \mathbb{S}^{n-1}$ and any two regions may only share points on their boundaries. The regions have equal area in the sense of the Lebesgue measure $\mathcal{L}^{n-1}$, with $\mathcal{L}^{n-1}(R) = \mathcal{L}^{n-1}(\mathbb{S}^{n-1})/N$, where $N$ is the number of partitions and the boundary of each region has measure zero. The diameter of a region is defined by $\operatorname{diam} R := \sup\{\|x-y\| : x, y \in R\}$.

In [12], Leopardi proved two useful results about this algorithm for $n \geq 2$ and $N \geq 1$ partitions. First, he showed that the EQ partition algorithm produces an equal area partition of $\mathbb{S}^{n-1}$. Second, he showed that the set of EQ partitions has diameter bound $K \in \mathbb{R}_+$ in the following sense:

$$\text{for any partition } P \text{ and for each } R \in P, \operatorname{diam} R \leq K N^{-1/(n-1)}. \tag{4}$$

As a consequence of the EQ partition algorithm, each region is defined to be a Cartesian product of spherical caps (the set of points of $\mathbb{S}^d$, $1 \leq d \leq n-1$, whose spherical distance to a given point $x$ is at most $\theta$, for some value of $\theta$) and intervals in spherical coordinates. The center of each region can then be defined to be the point that corresponds to the center of each spherical cap and/or interval. This leads to the following useful proposition which will be used to find a dense set of directions for the QrMads instance:

**Proposition 3.1** *Let $\{P_j\}_{j=1}^{\infty}$ be a sequence of EQ partitions such that the number of partitions $N_j \to \infty$. Next, build a sequence of points $\{c_i\}_{i=1}^{\infty}$ from the region centers by exhausting the centers from all the regions of a single partition, in any order, before taking centers from the next partition in the sequence. Then the resulting sequence, $\{c_i\}_{i=1}^{\infty}$, will be dense in $\mathbb{S}^{n-1}$.*

*Proof* Let $w \in \mathbb{S}^{n-1}$ and $\varepsilon > 0$ be arbitrary. Then for each partition $P_j$ there exists a region $R_j \in P_j$ with center $c_j$ such that $w \in R_j$. If $w$ is on the boundary of two or more such regions, then we may choose any one. We can then apply (4) to get

$$\|c_j - w\| \le \operatorname{diam} R_j \le K N_j^{-1/(n-1)} < \varepsilon,$$

for large enough $j$. □

## 4 Constructing a Basis

At each iteration of the Mads algorithm a positive spanning set $D_k$ is required. In the previous section we discussed a method for finding a dense set of directions, but not a dense set of positive spanning sets. In this section, we discuss a method for taking a single direction and forming an orthogonal basis that includes the original direction. We can then take the negatives of all the directions in the basis to form a positive spanning set. Since one of the directions from the dense set is included in each positive spanning set, we will then find that $\bigcup_{k=1}^{\infty} D_k$ is dense in the unit sphere.

To generate an orthogonal basis from a given partition center $c_k$, we will employ the QR decomposition of an $n \times m$ matrix $(n < m)$ $A_k$ of full rank. This matrix $A_k$ can be created by augmenting $c_k$ with any matrix of full rank (e.g. the identity or a rotation matrix). The QR decomposition then generates an orthogonal matrix $Q_k$ and an upper trapezoidal matrix $R_k$ such that $Q_k R_k = A_k$. It is important to note that the first column of $Q_k$ will be a scaler multiple of $c_k$, since $R_k$ is upper trapezoidal, and its first column will therefore be equal to $[a, 0, \dots, 0]^T$, for some constant $a \in \mathbb{R}$. Thus the resulting matrix $Q_k$ will give an orthogonal basis with the first direction corresponding to $c_k$.

To obtain a 'nearly orthogonal' integer basis, we will take each column vector $q_i$ of $Q_k$, project it onto the hypercube centered at the origin with integer side length $2 \cdot 2^{|\ell_k| + 2\ell_n}$, and then round each component to the nearest integer as follows:

$$d_i = \operatorname{round}\left(2^{|\ell_k| + 2\ell_n} \frac{q_i}{\|q_i\|_\infty}\right), \tag{5}$$

where round$(\cdot)$ refers to the rounding operation applied to each component of the vector and $\ell_k$ is as in (3). We then generate the new basis $H_k$, with column vectors $d_i$, and define $D_k := [\, H_k \; -H_k \,]$ to obtain a positive spanning set.

We add the positive constant $2\ell_n$ to the exponent in (5) in order to ensure that $H_k$ will be nonsingular. We do this because for small values of $2^{|\ell_k|}$, it is possible that the set of vectors $\{d_i = \operatorname{round}(2^{|\ell_k|} \frac{q_i}{\|q_i\|_\infty}) : 1 \le i \le n\}$ will not form a basis. To find the constant $\ell_n$, we use the fact that the distance from a matrix $A$ to the nearest singular matrix $B$ (in Frobenius norm) is bounded below by the smallest singular value of $A$, that is, if $B$ is singular then $\|B - A\| \ge \sigma_n(A)$. We let $A$ be the matrix with column vectors $2^{|\ell_k|+2\ell_n} q_i/\|q_i\|_\infty$,

$$A = 2^{|\ell_k|+2\ell_n} \begin{bmatrix} \frac{q_{1_1}}{\|q_1\|_\infty} & \frac{q_{2_1}}{\|q_2\|_\infty} & \cdots & \frac{q_{n_1}}{\|q_n\|_\infty} \\ \frac{q_{1_2}}{\|q_1\|_\infty} & \frac{q_{2_2}}{\|q_2\|_\infty} & \cdots & \frac{q_{n_2}}{\|q_n\|_\infty} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{q_{1_n}}{\|q_1\|_\infty} & \frac{q_{2_n}}{\|q_2\|_\infty} & \cdots & \frac{q_{n_n}}{\|q_n\|_\infty} \end{bmatrix}$$

$$= \begin{bmatrix} q_{1_1} & q_{2_1} & \cdots & q_{n_1} \\ q_{1_2} & q_{2_2} & \cdots & q_{n_2} \\ \vdots & \vdots & \ddots & \vdots \\ q_{1_n} & q_{2_n} & \cdots & q_{n_n} \end{bmatrix} \begin{bmatrix} \frac{2^{|\ell_k|+2\ell_n}}{\|q_1\|_\infty} & 0 & \cdots & 0 \\ 0 & \frac{2^{|\ell_k|+2\ell_n}}{\|q_2\|_\infty} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{2^{|\ell_k|+2\ell_n}}{\|q_n\|_\infty} \end{bmatrix}.$$

Then $A = Q_k T$, where $T$ is the diagonal matrix above. To find the singular values of $A$, we need to find the eigenvalues of $A^T A$:

$$A^T A = (Q_k T)^T (Q_k T) = T^T Q_k^T Q_k T = T^T I_n T = T^T T$$

$$= \begin{bmatrix} (\frac{2^{|\ell_k|+2\ell_n}}{\|q_1\|_\infty})^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & (\frac{2^{|\ell_k|+2\ell_n}}{\|q_n\|_\infty})^2 \end{bmatrix}.$$

This has eigenvalues $(2^{|\ell_k|+2\ell_n}/\|q_i\|_\infty)^2$, $1 \le i \le n$. Thus the singular values of $A$ are given by $2^{|\ell_k|+2\ell_n}/\|q_i\|_\infty$, $1 \le i \le n$. Next, we note that $2^{|\ell_k|+2\ell_n}/\|q_i\|_\infty \ge 2^{|\ell_k|+2\ell_n}$ and so for $A + E$ to be nonsingular, we need

$$\|A + E - A\| = \|E\| < 2^{|\ell_k|+2\ell_n},$$

where $E$ represents the perturbation matrix resulting from rounding. As a consequence of the projection and the rounding, the matrix $E$ satisfies $e_{ij} \le 1/2$ for all but $n$ choices of $i$ and $j$, and $e_{ij} = 0$ for the remaining $n$ choices. Then $\|E\| \le \frac{\sqrt{n^2-n}}{2}$ and so we require

$$\frac{\sqrt{n^2 - n}}{2} < 2^{|\ell_k|+2\ell_n},$$

for all values of $\ell_k$. Then letting $\ell_k = 0$, we obtain the following requirement for the constant $\ell_n$ determined only by the dimension $n$:

$$\sqrt{n^2 - n} < 2^{2\ell_n+1}. \tag{6}$$

The following proposition will be useful for showing our new version of Mads is a valid instance and for establishing convergence results.

**Proposition 4.1** *If* $w \in \mathbb{S}^{n-1}$ *and* $\varepsilon > 0$ *are arbitrary*, $d = \mathrm{round}(\alpha w/\|w\|_\infty)$, *and* $\alpha > 0$ *is large*; *then* $\| d/\|d\| - w \| < \varepsilon/2$.

*Proof* In Proposition 3.4 of [7] it is shown that there exists $\beta > 0$ such that

$$\left\| \frac{q}{\|q\|} - w \right\| < \frac{\varepsilon}{2},$$

when $q = \text{round}(\alpha w)$ and $\alpha > \beta$. Then for any given $w \in \mathbb{S}^{n-1}$ with $\frac{\alpha}{\|w\|_\infty} > \beta$, we have

$$\left\| \frac{d}{\|d\|} - w \right\| < \frac{\varepsilon}{2}.$$

Now, since $\|w\|_\infty \leq 1$ for all $w \in \mathbb{S}^{n-1}$, we can observe that $\frac{\alpha}{\|w\|_\infty} \geq \alpha$. Thus for $\alpha > \beta$, the result follows for all $w \in \mathbb{S}^{n-1}$. □

Although QrMads uses a 'nearly orthogonal' basis in place of an orthogonal set of directions at each iteration, we believe that it has two potential advantages over OrthoMads. First, the distribution of directions taken over all iterations is more uniform when QrMads is used in place of OrthoMads (see Fig. 2). Second, the method described above forces the distance in $\ell_\infty$-norm from the poll center to a poll trial point to equal $2^{|\ell_k|+2\ell_n}$ instead of only being bounded above by such value, ensuring that the step size taken will not be unintentionally small.

## 5 The QrMads Instance of Mads

The QrMads algorithm differs from the LtMads and OrthoMads algorithms only in the construction of the poll directions $D_k$ and the poll set $P_k$. For all three instances, the mesh $M_k$ is defined by the set of directions $D := [\, I_n \; -I_n \,]$. For QrMads, the choices of the mesh and poll size parameters are

$$\Delta_k^p = 2^{-\ell_k} \quad \text{and} \quad \Delta_k^m = \min\{4^{-\ell_k-\ell_n}, 4^{-\ell_n}\}, \tag{7}$$

with $\ell_n$ defined as in (6), so that $\Delta_k^p/\Delta_k^m = 2^{|\ell_k|+2\ell_n}$ for all $k$, as in (5). The update rules for the mesh index $\ell_k$ are given in Fig. 1, which describes the QrMads algorithm.

This algorithm is essentially the same as the algorithm from [7], with $t_k$ and $c_{t_k}$ in place of the Halton directions $u_{t_k}$ and $q_{t_k,\ell_k}$, respectively. The integer $t_k$ and the update rules are chosen so that there will be a subsequence of unsuccessful iterations with $\Delta_k^m \to 0$ and such that the directions used in the subsequence will correspond to the entire sequence of partition centers as in [7].

The poll directions $D_k$ are determined using a sequence of partition centers as described in Sect. 2, the mesh index $\ell_k$, a sequence of $n \times n$ matrices of full rank, and the index $t_k$. At each iteration, a center $c_{t_k}$ is taken from the sequence of partition centers, this is then augmented with a random orthogonal matrix to form an $n \times (n+1)$ matrix, and then QR decomposition is used to find an orthogonal basis matrix with the direction from the first column corresponding to $c_{t_k}$, as described in Sect. 3. The columns of the resulting matrix $Q_k$ are then projected onto the hypercube centered at zero with side length $2 \cdot 2^{|\ell_k|+2\ell_n}$ and rounded. The resulting matrix $H_k$ will then be 'nearly orthogonal' with the length of each column in $\ell_\infty$-norm equal to $\Delta_k^p/\Delta_k^m$.

**Fig. 1** The QrMads algorithm

**[0] Initialization**

$x_0 \in \Omega$, $\ell_0 \leftarrow 0$, $k \leftarrow 0$, $t_0 \leftarrow 0$

**[1] Iteration** $k$

Search (optional)

evaluate $f$ on a finite subset $S_k \subset M_k$

Poll (optional if Search is successful)

**if** the poll size is the smallest one so far

$t_k \leftarrow \ell_k + t_0$

**else**

$t_k \leftarrow 1 + \max\{t_j : j = 0, 1, \dots, k\}$

find $c_{t_k}$, compute $H_k$ and $D_k = [H_k \; -H_k]$

evaluate $f$ on $P_k \subset M_k$

**[2] Updates**

**if** the iteration is successful (there exists $x_s \in S_k$ or $x_p \in P_k$ such that $f(x_s) < f(x_k)$ or $f(x_p) < f(x_k)$)

$x_{k+1} \leftarrow x_s$ or $x_p$

$\ell_{k+1} \leftarrow \ell_k - 1$

**else** (the iteration fails)

$x_{k+1} \leftarrow x_k$

$\ell_{k+1} \leftarrow \ell_k + 1$

$k \leftarrow k + 1$

**goto [1]** if stopping criteria are not met

This then forces the trial point $x_k + \Delta_k^m D_k e_i$ to be exactly $\Delta_k^p$ from the poll center in $\ell_\infty$-norm. Finally, $H_k$ is completed to a maximal positive basis composed of $2n$ directions, $D_k = [\, H_k \; -H_k \,]$.

The following lemma will be useful to show that QrMads has the same convergence properties as in [1].

**Lemma 5.1** *The set of normalized directions* $\{d_{i,\ell_k}/\|d_{i,\ell_k}\|\}_{i=1}^{\infty}$, *with* $d_{i,\ell_k} = \mathrm{round}(2^{|\ell_k|+2\ell_n} c_i/\|c_i\|_\infty)$ *as in* (5), *is dense on* $\mathbb{S}^{n-1}$.

*Proof* Let $w \in \mathbb{S}^{n-1}$ and $\varepsilon > 0$ be arbitrary. From Proposition 3.1, there exist a subsequence $\{c_{i_j}\}_{j=1}^{\infty}$ of region centers of the sequence of partitions and a constant $N > 0$ such that $\|w - c_{i_j}\| < \varepsilon/2$, for all $i_j > N$. From Proposition 4.1, for large values of $|\ell_k|$ we have $\|c_{i_j} - \frac{d_{i_j,\ell_k}}{\|d_{i_j,\ell_k}\|}\| < \varepsilon/2$. Then by applying the triangle inequality we obtain

$$\left\| w - \frac{d_{i_j,\ell_k}}{\|d_{i_j,\ell_k}\|} \right\| \le \|w - c_{i_j}\| + \left\| c_{i_j} - \frac{d_{i_j,\ell_k}}{\|d_{i_j,\ell_k}\|} \right\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

for some $i_j > N$ and for some large $|\ell_k|$. □

We are now ready for the main result:

**Theorem 5.1** *QrMads is a valid Mads instance with the convergence properties from* [1].

*Proof* We need to show that the poll directions for QrMads satisfy the following five properties from [1, 13]:

- *Each $D_k$ is a positive spanning set*: This follows from the use of the constant $\ell_n$ to ensure that the round procedure on $Q_k$ results in a basis.
- *Any direction $D_k e_i$ $(1 \le i \le 2n)$ can be written as a non-negative integer combination of the directions of $D$*: This is true by construction since $D = [\, I_n \ -I_n \,]$ and $D_k e_i$ is an integer vector on the hypercube centered at the origin with side length $2 \cdot 2^{|\ell_k|+2\ell_n}$.
- *The distance from the poll center $x_k$ to a poll trial point in $\ell_\infty$-norm is bounded above by $\Delta_k^p$*: This is true by construction since we ensured $\|D_k e_i\|_\infty = 2^{|\ell_k|+2\ell_n}$ and so $\|\Delta_k^m D_k e_i\|_\infty = \Delta_k^m 2^{|\ell_k|+2\ell_n} = \Delta_k^p$.
- *The set of normalized directions used over all failed iterations is dense on the unit sphere*: This follows from the strategy chosen for updating $D_k$ so that the entire sequence of partition centers corresponds to the sequence of failed iterates, by Lemma 5.1, and (2), which ensures the existence of large $|\ell_k|$.
- *Limits (as defined by [8]) of convergent subsequences of the normalized sets $\overline{D}_k = \{\frac{d}{\|d\|} : d \in D_k\}$ are positive spanning sets*: For this property we need to show $\det(\overline{H}_k) > \tau$ for all $k$, for some constant $\tau > 0$. The result will then follow as in [13].

First, note that the set $\mathcal{O} = \{A \in \mathbb{R}^{n\times n} : A \text{ is an orthogonal matrix}\}$ is a bounded set since $\mathcal{O} \subset \{A \in \mathbb{R}^{n\times n} : \|A\| = \sqrt{n}\} \subset B(0, \sqrt{n} + \varepsilon)$, where $\|\cdot\|$ denotes the Frobenius norm on $\mathbb{R}^{n\times n}$ and $B(0, \sqrt{n} + \varepsilon)$ is the ball of radius $\sqrt{n} + \varepsilon$ centered around the matrix of zeros. Furthermore, the set $V_\delta = \{B : \|A - B\| \le n\delta \text{ for any } A \in \mathcal{O}\}$ is bounded since any such $B$ would satisfy $\|B\| \le \sqrt{n} + n\delta$ and so $V_\delta \subset B(0, \sqrt{n} + n\delta + \varepsilon)$. Now, since the determinant function is $C^1$ on $\mathbb{R}^{n\times n}$, it is Lipschitz on the bounded set $V_\delta$. That is, for any $B, C \in V_\delta$ there exists a constant $K > 0$ such that

$$\bigl|\det(C) - \det(B)\bigr| \le K\|C - B\|.$$

This then yields

$$\bigl|\det(B)\bigr| \ge \bigl|\det(C)\bigr| - K\|C - B\|.$$

So, in particular, if we take $C \in \mathcal{O}$ (since $\mathcal{O} \subset V_\delta$) and $B$ such that $\|C - B\| \le n\delta$ we get

$$\bigl|\det(B)\bigr| \ge \bigl|\det(C)\bigr| - K\|C - B\| \ge 1 - Kn\delta.$$

Now, note that for any $k$, $Q_k \in \mathcal{O}$ and $\overline{H}_k \in V_{\delta_k}$ for some $\delta_k > 0$, since $\overline{H}_k = Q_k + E_k$ for some perturbation matrix $E_k$ resulting from rounding. Next, we choose $\delta$ small enough so that $1 - Kn\delta > 0$. Then by Proposition 4.1, we can choose $|\ell_k|$ large enough, say $|\ell_k| > \ell_\delta$, so that $\|\overline{H}_k - Q_k\| \le n\delta$ and conclude

$$\bigl|\det(\overline{H}_k)\bigr| \ge 1 - Kn\delta, \quad \text{for all } k \text{ such that } |\ell_k| > \ell_\delta.$$

Next, let $H_k$ have columns $d_i$ and note that by construction $\det(H_k) \ne 0$. Since $H_k$ has integer components, we know $|\det(H_k)| \ge 1$. Also, note $\|d_i\| \le \sqrt{n}2^{|\ell_k|+2\ell_n}$, $1 \le i \le n$, since each $d_i$ is on the hypercube with side length $2 \cdot 2^{|\ell_k|+2\ell_n}$. We then get the following estimate

$$\begin{aligned}\left|\det(\overline{H}_k)\right| &= \left|\det\begin{bmatrix}\frac{d_{1_1}}{\|d_1\|} & \cdots & \frac{d_{n_1}}{\|d_n\|}\\ \vdots & \ddots & \vdots\\ \frac{d_{1_n}}{\|d_1\|} & \cdots & \frac{d_{n_n}}{\|d_n\|}\end{bmatrix}\right|\\ &= \frac{|\det(H_k)|}{\prod_{i=1}^n \|d_i\|} \geq \frac{|\det(H_k)|}{n^{n/2}2^{n(|\ell_k|+2\ell_n)}} \geq n^{-n/2}2^{-n(|\ell_k|+2\ell_n)}.\end{aligned}$$

Finally, we note that $n^{-n/2}2^{-n(|\ell_k|+2\ell_n)}$ is decreasing as $|\ell_k|$ increases. Then $|\det(\overline{H}_k)| \geq n^{-n/2}2^{-n(\ell_\delta+2\ell_n)}$, if $|\ell_k| \leq \ell_\delta$, where $\ell_\delta$ is defined above. We then let

$$\tau = \min\left\{1 - Kn\delta, n^{-n/2}2^{-n(\ell_\delta+2\ell_n)}\right\}$$

and conclude $\det(\overline{H}_k) > \tau$ for all $k$. □

## 6 Numerical Tests

In this section, we compare QrMads and OrthoMads on 150 problems taken from the optimization literature. Each test was performed using our own MATLAB implementation. The test problems fall into three categories. The first category consists of 60 smooth test functions taken from [14]. The second category consists of 62 nonsmooth test functions taken from [15] and [16], where many of the problems in [16] are variable dimension generalizations of the problems from [15]. Lastly, 28 constrained test problems are also taken from [15] and [16], with the problems from [16] being constrained generalizations of some of the nonsmooth problems. Many of the constrained problems from [16] were chosen to correspond to unconstrained problems where OrthoMads performed better than QrMads.

The choices for $\Delta_k^m$ and $\Delta_k^p$ correspond to (3) and (7). For both instances the mesh index $\ell_k$ is allowed to be negative. For both instances, the initial mesh size $\Delta_0^m$ is set to one. If either the number of function evaluations reaches $1000n$ or $\Delta_k^p$ drops below $10^{-10}$, then the stopping criteria are satisfied. For both implementations, an opportunistic strategy was used for the Poll step (the Poll stops as soon as a successful point has been found) and no search step was performed. For both implementations, constraints are handled using the extreme barrier method ($f(x) = +\infty$ if $x \notin \Omega$).

For each test, a partition of the unit sphere of size $N$ was generated, with $N$ satisfying the following:

$$N = \begin{cases} 10^6 & \text{if } n \leq 6,\\ 10^n & \text{if } 6 < n \leq 15,\\ 10^{15} & \text{if } n > 15.\end{cases}$$

A random ordering of the centers of each region of the partition was then used as the sequence of directions for QrMads. This choice was made purely out of convenience as a means of testing the algorithm. There are many choices that can be made for generating a dense sequence of directions including starting with a smaller partition

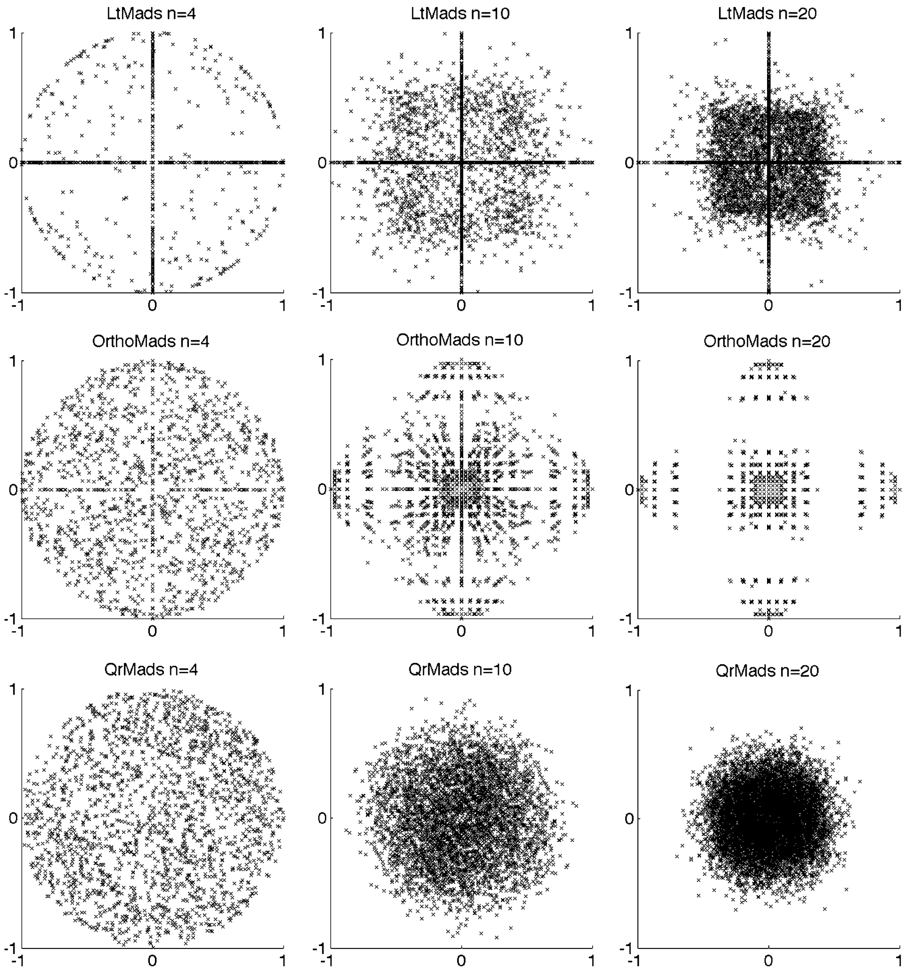


**Fig. 2** First two coordinates of the normalized poll directions of three instances of the Mads algorithm on the $n$-dimensional Rosenbrock function with $n = 4, 10, 20$

size and exhausting that partition before moving on to the centers from a larger partition. The performance of the QrMads algorithm when using different sequences of partition centers is unknown at this time and is an area for future research. Due to the random ordering of the partition centers, 30 instances of QrMads were performed on each test function and the final function values were compared to the value obtained from the OrthoMads algorithm.

Figure 2 illustrates the distribution and the density of directions for three Mads instances: LtMads, OrthoMads, and QrMads. The three variants were run on the Rosenbrock function with dimensions $n = 4$, 10, and 20. In each case there are roughly $500n$ directions shown, where each direction is projected onto the plane defined by the first two coordinates. The resulting plots are typical of what can be expected regardless of the choices of the two coordinates. In each case, the QrMads directions have a relatively uniform distribution, while the LtMads directions have a large proportion with zero coordinates and the OrthoMads directions tend to cluster within specific regions. Upon further observation, it appears that many of the OrthoMads directions tend to cluster toward the coordinate directions as the dimension increases. In each case, however, the QrMads directions depict a more uniform distribution.

We also found that QrMads finds a final function value for the Rosenbrock function that is strictly better than the final value obtained by OrthoMads for 19 of the 30 runs when $n = 4$ and for all 30 runs when $n = 10$ or 20. Overall, the test results presented in this section show a similar pattern of QrMads outperforming OrthoMads as the dimension of the test functions increase.

*Note 6.1* In Fig. 2, the QrMads directions (and to some extent the LtMads directions) become more restricted towards the center as the dimension increases. This behavior is a reflection of the concentration of measure phenomenon and the fact that the observable diameter of $\mathbb{S}^{n-1}$ converges to zero as $n \to \infty$ (for more on this topic see [17]). Intuitively, one can imagine that as $n$ increases, the contribution to the norm of a vector from the first two elements gets smaller.

The results of the tests are summarized in Tables 1, 2, 3, and 4. If we let $f_O^*$ denote the final value obtained by OrthoMads, $f_{QR}^*$ denote the final value obtained by QrMads, and $f(x_0)$ denote the initial value of $f$, then we can define the three scores $S_1$, $S_2$, and $S_3$ as follows:

- $S_1 :=$ the number of final values from 30 runs of QrMads for which $f_{QR}^* < f_O^*$.
- $S_2 :=$ the number of final values from 30 runs of QrMads for which $f_{QR}^* \le f_O^* + 0.01(f(x_0) - f_O^*)$.
- $S_3 :=$ the number of final values from 30 runs of QrMads for which $f_O^* \le f_{QR}^* + 0.01(f(x_0) - f_{QR}^*)$.

In this way, we can measure the number of final values of $f$ for the 30 runs of QrMads that are strictly better than the value obtained from OrthoMads ($S_1$), as well as the number of final values of $f$ from QrMads that are within 1 % of the value obtained from OrthoMads ($S_2$) and vice versa ($S_3$).

Table 1 shows the results from the tests performed on the set of 60 unconstrained smooth functions. QrMads received a score for $S_1$ of 15 or more on 38 of these test problems, meaning that QrMads performed better than OrthoMads on more than half the 30 runs for a majority of the test functions from this set. QrMads received a score for $S_2$ of 15 or more on all of the problems from this set, indicating that QrMads finds a final value within a 1 % tolerance of OrthoMads on nearly all of the problems from this set. Finally, OrthoMads received a score for $S_3$ of 15 or more on 48 out of 60 of the test problems from this set. Of the 12 scores for $S_3$ that are below 15, there are 11 zeros and a four, indicating that QrMads showed a significant improvement over OrthoMads on 12 of the test problems from this set.

Table 2 shows similar results for the set of 62 unconstrained nonsmooth functions tested. QrMads received a score for $S_1$ of 15 or more on 30 of these test problems, meaning that QrMads performed better than OrthoMads on more than half the 30 runs for nearly half of the test functions in this set. QrMads received a score for $S_2$ of

**Table 1** Results for smooth problems

| Function | $n$ | $S_1$ | $S_2$ | $S_3$ | Function | $n$ | $S_1$ | $S_2$ | $S_3$ |
|---|---|---|---|---|---|---|---|---|---|
| Box 3D | 3 | 28 | 30 | 30 | Gaussian | 3 | 24 | 30 | 30 |
| Gulf Research and Development | 3 | 0 | 30 | 30 | Helical Valley | 3 | 0 | 30 | 30 |
| Brown Almost-Linear | 4 | 17 | 30 | 30 | Broyden Tridiagonal | 4 | 10 | 22 | 30 |
| Discrete Boundary Value | 4 | 0 | 30 | 30 | Discrete Integral Equation | 4 | 6 | 30 | 30 |
| Extended Powell Singular | 4 | 1 | 30 | 30 | Penalty 1 | 4 | 5 | 30 | 30 |
| Penalty 2 | 4 | 2 | 30 | 30 | Rosenbrock | 4 | 19 | 27 | 30 |
| Trigonometric | 4 | 30 | 30 | 18 | Variably Dimensioned | 4 | 19 | 30 | 30 |
| Wood | 4 | 0 | 30 | 30 | Biggs EXP6 | 6 | 0 | 26 | 30 |
| Watson | 8 | 9 | 30 | 30 | Brown Almost-Linear | 10 | 29 | 30 | 30 |
| Broyden Tridiagonal | 10 | 12 | 28 | 30 | Discrete Boundary Value | 10 | 30 | 30 | 0 |
| Discrete Integral Equation | 10 | 0 | 30 | 30 | Penalty 1 | 10 | 1 | 30 | 30 |
| Penalty 2 | 10 | 1 | 30 | 30 | Rosenbrock | 10 | 30 | 30 | 0 |
| Trigonometric | 10 | 18 | 26 | 30 | Variably Dimensioned | 10 | 30 | 30 | 30 |
| Extended Powell Singular | 12 | 30 | 30 | 30 | Watson | 16 | 7 | 30 | 30 |
| Brown Almost-Linear | 20 | 23 | 30 | 30 | Broyden Tridiagonal | 20 | 0 | 30 | 30 |
| Discrete Boundary Value | 20 | 30 | 30 | 0 | Discrete Integral Equation | 20 | 0 | 30 | 30 |
| Extended Powell Singular | 20 | 30 | 30 | 30 | Penalty 1 | 20 | 30 | 30 | 30 |
| Penalty 2 | 20 | 30 | 30 | 30 | Rosenbrock | 20 | 30 | 30 | 0 |
| Trigonometric | 20 | 29 | 30 | 4 | Variably Dimensioned | 20 | 30 | 30 | 30 |
| Watson | 24 | 18 | 30 | 30 | Brown Almost-Linear | 30 | 23 | 30 | 30 |
| Broyden Tridiagonal | 30 | 0 | 30 | 30 | Discrete Boundary Value | 30 | 30 | 30 | 0 |
| Discrete Integral Equation | 30 | 0 | 30 | 30 | Penalty 1 | 30 | 30 | 30 | 30 |
| Penalty 2 | 30 | 30 | 30 | 30 | Rosenbrock | 30 | 30 | 30 | 0 |
| Trigonometric | 30 | 30 | 30 | 0 | Variably Dimensioned | 30 | 30 | 30 | 30 |
| Watson | 31 | 16 | 30 | 30 | Extended Powell Singular | 32 | 14 | 30 | 30 |
| Brown Almost-Linear | 40 | 22 | 30 | 30 | Broyden Tridiagonal | 40 | 30 | 30 | 30 |
| Discrete Boundary Value | 40 | 30 | 30 | 0 | Discrete Integral Equation | 40 | 0 | 30 | 30 |
| Extended Powell Singular | 40 | 30 | 30 | 30 | Penalty 1 | 40 | 30 | 30 | 0 |
| Penalty 2 | 40 | 30 | 30 | 30 | Rosenbrock | 40 | 30 | 30 | 0 |
| Trigonometric | 40 | 30 | 30 | 0 | Variably Dimensioned | 40 | 30 | 30 | 30 |
| **median** | | | | | | | **23** | **30** | **30** |
| **mean** | | | | | | | **18.38** | **29.65** | **23.87** |

15 or more on all but ten of the test functions from this set, indicating that QrMads finds a final value within a 1 % tolerance of OrthoMads on a large majority of these problems. Finally, OrthoMads received a score for $S_3$ of 15 or more on 42 out of 62 of the test problems from this set. Of the 20 scores for $S_3$ that are below 15, there are 15 zeros, a one, and a three; indicating that QrMads showed a significant improvement on these test problems.

Table 3 shows the results for the set of constrained test problems. QrMads received a score for $S_1$ of 15 or more on 17 out of 28 of the test problems from this set, meaning

**Table 2** Results for nonsmooth problems

| Function | $n$ | $S_1$ | $S_2$ | $S_3$ | Function | $n$ | $S_1$ | $S_2$ | $S_3$ |
|---|---|---|---|---|---|---|---|---|---|
| Chained CB3 | 4 | 0 | 29 | 30 | Chained Crescent | 4 | 0 | 23 | 30 |
| Chained LQ | 4 | 25 | 30 | 30 | Chained Mifflin 2 | 4 | 8 | 30 | 30 |
| L1HILB | 4 | 10 | 30 | 30 | Maxq | 4 | 2 | 30 | 30 |
| MXHILB | 4 | 14 | 16 | 30 | Brown | 5 | 0 | 14 | 30 |
| Number of Active Faces | 5 | 28 | 30 | 30 | PBC1 | 5 | 30 | 30 | 3 |
| Shor | 5 | 18 | 24 | 30 | ElAttar | 6 | 30 | 30 | 0 |
| EVD61 | 6 | 10 | 15 | 25 | Wong 1 | 7 | 0 | 1 | 30 |
| Watson | 8 | 21 | 21 | 12 | Filter | 9 | 10 | 12 | 23 |
| Brown | 10 | 0 | 13 | 30 | Chained CB3 | 10 | 0 | 30 | 30 |
| Chained Crescent | 10 | 0 | 29 | 30 | Chained LQ | 10 | 1 | 30 | 30 |
| Chained Mifflin 2 | 10 | 29 | 30 | 30 | L1HILB | 10 | 30 | 30 | 0 |
| Maxq | 10 | 30 | 30 | 0 | MXHILB | 10 | 30 | 30 | 0 |
| Number of Active Faces | 10 | 30 | 30 | 0 | Polak 2 | 10 | 0 | 1 | 30 |
| Wong 2 | 10 | 0 | 9 | 30 | Osborne2 | 11 | 2 | 3 | 28 |
| Polak 3 | 11 | 2 | 30 | 30 | Steiner 2 | 12 | 0 | 1 | 30 |
| Shell Dual | 15 | 0 | 0 | 30 | Watson | 16 | 9 | 13 | 24 |
| Brown | 20 | 0 | 17 | 30 | Chained CB3 | 20 | 0 | 30 | 30 |
| Chained Crescent | 20 | 0 | 25 | 30 | Chained LQ | 20 | 4 | 30 | 30 |
| Chained Mifflin 2 | 20 | 27 | 30 | 30 | L1HILB | 20 | 30 | 30 | 10 |
| Maxq | 20 | 30 | 30 | 0 | MXHILB | 20 | 30 | 30 | 0 |
| Number of Active Faces | 20 | 30 | 30 | 0 | Wong 3 | 20 | 7 | 15 | 29 |
| Watson | 24 | 16 | 18 | 16 | Brown | 30 | 0 | 18 | 30 |
| Chained CB3 | 30 | 0 | 30 | 30 | Chained Crescent | 30 | 0 | 24 | 30 |
| Chained LQ | 30 | 27 | 30 | 30 | Chained Mifflin 2 | 30 | 30 | 30 | 30 |
| L1HILB | 30 | 30 | 30 | 1 | Maxq | 30 | 30 | 30 | 0 |
| MXHILB | 30 | 30 | 30 | 0 | Number of Active Faces | 30 | 30 | 30 | 0 |
| Watson | 31 | 18 | 20 | 14 | Brown | 40 | 0 | 20 | 30 |
| Chained CB3 | 40 | 0 | 30 | 30 | Chained Crescent | 40 | 0 | 25 | 30 |
| Chained LQ | 40 | 26 | 30 | 30 | Chained Mifflin 2 | 40 | 30 | 30 | 30 |
| L1HILB | 40 | 30 | 30 | 0 | Maxq | 40 | 30 | 30 | 0 |
| MXHILB | 40 | 30 | 30 | 0 | Number of Active Faces | 40 | 30 | 30 | 0 |
| **median** | | | | | | | **12** | **30** | **30** |
| **mean** | | | | | | | **14.74** | **24.26** | **20.35** |

that QrMads performed better than OrthoMads on a majority of these problems. For $S_2$, QrMads received a score of 15 or more on all but five of the constrained test problems, again showing that QrMads finds a final value within a 1 % tolerance of OrthoMads on a large majority of these problems. OrthoMads received a score for $S_3$ of 15 or more on 15 out of 28 of the test problems from this set. Of the 13 scores for $S_3$ that are below 15, there are ten zeros, a two, and a seven; indicating that QrMads showed significant improvement on these test problems. It is of interest to note that

**Table 3** Results for constrained problems

| Function | $n$ | $S_1$ | $S_2$ | $S_3$ |
|---|---|---|---|---|
| Brown with Broyden Tridiagonal constraint | 4 | 0 | 9 | 30 |
| Chained CB3 with MAD1 constraint | 4 | 3 | 8 | 30 |
| Chained Crescent with Broyden Tridiagonal constraint | 4 | 14 | 19 | 23 |
| Chained LQ with MAD1 constraint | 4 | 4 | 17 | 30 |
| MAD6 | 5 | 13 | 18 | 22 |
| Pentagon | 6 | 27 | 28 | 2 |
| Dembo 3 | 7 | 0 | 0 | 30 |
| Dembo 5 | 8 | 30 | 30 | 0 |
| HS114 | 9 | 0 | 0 | 30 |
| Wong 2 | 10 | 26 | 29 | 11 |
| Brown with Broyden Tridiagonal constraint | 10 | 30 | 30 | 0 |
| Chained CB3 with MAD1 constraint | 10 | 29 | 29 | 7 |
| Chained Crescent with Broyden Tridiagonal constraint | 10 | 10 | 20 | 27 |
| Chained LQ with MAD1 constraint | 10 | 0 | 23 | 30 |
| Wong 3 | 20 | 30 | 30 | 0 |
| MAD8 | 20 | 30 | 30 | 0 |
| Brown with Broyden Tridiagonal constraint | 20 | 30 | 30 | 0 |
| Chained CB3 with MAD1 constraint | 20 | 30 | 30 | 0 |
| Chained Crescent with Broyden Tridiagonal constraint | 20 | 5 | 9 | 28 |
| Chained LQ with MAD1 constraint | 20 | 13 | 30 | 29 |
| Brown with Broyden Tridiagonal constraint | 30 | 30 | 30 | 0 |
| Chained CB3 with MAD1 constraint | 30 | 30 | 30 | 0 |
| Chained Crescent with Broyden Tridiagonal constraint | 30 | 20 | 27 | 21 |
| Chained LQ with MAD1 constraint | 30 | 30 | 30 | 30 |
| Brown with Broyden Tridiagonal constraint | 40 | 30 | 30 | 0 |
| Chained CB3 with MAD1 constraint | 40 | 30 | 30 | 0 |
| Chained Crescent with Broyden Tridiagonal constraint | 40 | 18 | 29 | 24 |
| Chained LQ with MAD1 constraint | 40 | 30 | 30 | 30 |
| **median** | | **26.5** | **29** | **21.5** |
| **mean** | | **19.36** | **23.39** | **15.50** |

many of the constrained problems in this set were chosen to be generalizations of nonsmooth problems on which QrMads did not perform as well as OrthoMads (based on the $S_1$ score), yet QrMads performed better than OrthoMads on most of these problems when the dimension exceeded 20.

Finally, Table 4 summarizes the results for all the test problems broken into categories of increasing dimension. The only categories where QrMads does not show a clear improvement over OrthoMads are the cases where the dimension $n$ is below 20. As the dimension of the problems increases, however, we can see that QrMads outperforms OrthoMads on a consistent basis. We believe that this reflects the decreasing quality in the distribution of the poll directions for OrthoMads as $n$ increases and the

**Table 4** Summary of results for all test problems

| | | $S_1$ | $S_2$ | $S_3$ |
|---|---|---|---|---|
| $n < 10$ (42 total tests) | median | 9.5 | 30 | 30 |
| | mean | 11.12 | 23.76 | 25.83 |
| $10 \leq n < 20$ (32 total tests) | median | 9.5 | 30 | 30 |
| | mean | 13.94 | 24.16 | 21.78 |
| $20 \leq n < 30$ (28 total tests) | median | 29.5 | 30 | 30 |
| | mean | 20.07 | 27.64 | 18.07 |
| $30 \leq n < 40$ (25 total tests) | median | 30 | 30 | 30 |
| | mean | 21.52 | 28.76 | 18.24 |
| $n = 40$ (23 total tests) | median | 30 | 30 | 30 |
| | mean | 23.74 | 29.30 | 16.70 |
| All tests (150 total tests) | median | 19.5 | 30 | 30 |
| | mean | 17.06 | 26.25 | 20.85 |

relatively uniform distribution of poll directions for the QrMads implementation for all dimensions, as illustrated in Fig. 2.

## 7 Conclusions

Just as OrthoMads was developed to generate a more uniform set of polling directions than what is produced by LtMads, this paper introduces QrMads as a means of further improving on this idea when compared to OrthoMads. Although OrthoMads produces a relatively uniform distribution of poll directions for low dimensional problems, in high dimensions this is no longer the case. It is the use of the Householder transformation for generating an orthogonal set of directions that causes a significant number of directions to cluster toward the coordinate directions, when we consider the set of directions taken over all iterations. Although the authors of [7] believed orthogonality to be important for eliminating large angles between poll directions, they overlooked a subtle mathematical implication of using Householder transformations that prevent a uniform distribution, which is masked in low dimensions. QrMads seeks to remedy this condition by using QR decomposition and an equal area partition of the unit sphere to generate a more uniformly distributed set of directions that are 'nearly orthogonal' at each iteration. The trade-off is that QrMads is not deterministic. Consequently, an area of further study is that of constructing a deterministic version of this Mads instance.

In some real-world applications, the coordinate directions often have important physical meaning, and thus may be better choices than those chosen randomly. In this case, OrthoMads may perform better. On the other hand, we believe that having a uniform distribution of poll directions is beneficial for problems with higher dimension and when polling near constraint boundaries because the more uniform set of directions should increase the chance of finding a direction close to a tangent cone generator and thus help the algorithm to avoid getting bogged down near a constraint boundary. The results from the problems tested in this paper support this conjecture since QrMads was shown to outperform OrthoMads on a large majority of the higher dimensional problems ($n \geq 20$) and a large majority of the constrained problems.

## References


1. Audet, C., Dennis, J. Jr.: Mesh adaptive direct search algorithms for constrained optimization. SIAM J. Optim. **17**(1), 188–217 (2006)
2. Abramson, M.A., Audet, C.: Convergence of mesh adaptive direct search to second-order stationary points. SIAM J. Optim. **17**(2), 606–619 (2006)
3. Clarke, F.: Optimization and Nonsmooth Analysis. SIAM, Philadelphia (1983)
4. Vicente, L., Custódio, A.: Analysis of direct searches for discontinuous functions. Math. Program. **133**(1–2), 299–325 (2012)
5. Rockafellar, R.: Generalized directional derivatives and subgradients of nonconvex functions. Can. J. Math. **32**(2), 257–280 (1980)
6. Torczon, V.: On the convergence of pattern search algorithms. SIAM J. Optim. **7**(1), 1–25 (1997)
7. Abramson, M., Audet, C., Dennis, J. Jr., Le Digabel, S.: ORTHOMADS: a deterministic MADS instance with orthogonal directions. SIAM J. Optim. **10**(2), 948–966 (2009)
8. Coope, I.D., Price, C.J.: Frame-based methods for unconstrained optimization. J. Optim. Theory Appl. **107**(2), 261–274 (2000)
9. Halton, J.H.: On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. **2**(1), 84–90 (1960)
10. Leopardi, P.: A partition of the unit sphere into regions of equal area and small diameter. Electron. Trans. Numer. Anal. **25**, 309–327 (2006)
11. Leopardi, P.: Distributing points on the sphere: partitions, separation, quadrature and energy. Ph.D. thesis, University of New South Wales (2007)
12. Leopardi, P.: Diameter bounds for equal area partitions of the unit sphere. Electron. Trans. Numer. Anal. **35**, 1–16 (2009)
13. Audet, C., Custódio, A.L., Dennis, J. Jr.: Erratum: mesh adaptive direct search algorithms for constrained optimization. SIAM J. Optim. **18**(4), 1501–1503 (2008)
14. Moré, J.J., Garbow, B.S., Hillstrom, K.E.: Testing unconstrained optimization software. ACM Trans. Math. Softw. **7**(1), 17–41 (1981)
15. Lukšan, L., Vlček, J.: Test problems for nonsmooth unconstrained and linearly constrained optimization. Tech. Rep. No. 798, Institute of Computer Science, Academy of Sciences of the Czech Republic (2000)
16. Karmitsa, N.: Test problems for large-scale nonsmooth minimization. Tech. Rep. No. B. 4/2007, University of Jyväskylä (2007)
17. Ledoux, M.: The Concentration of Measure Phenomenon. Am. Math. Soc., Providence (2001)