Open access • Journal Article • DOI:10.1137/080716980

OrthoMADS: A Deterministic MADS Instance with Orthogonal Directions

— Source link

Mark A. Abramson, Charles Audet, John E. Dennis, Sébastien Le Digabel

Institutions: Rice University

Published on: 01 Jul 2009 - Siam Journal on Optimization (Society for Industrial and Applied Mathematics)

OrthoMads: A deterministic Mads instance with orthogonal directions *

Mark A. Abramson † Charles Audet ‡ J.E. Dennis Jr. §
Sébastien Le Digabel ¶

28th February 2008

Abstract

The purpose of this paper is to introduce a new way of choosing directions for the Mesh Adaptive Direct Search (Mads) class of algorithms. The advantages of this new OrthoMads instantiation of Mads are that the polling directions are chosen deterministically, ensuring that the results of a given run are repeatable, and that they are orthogonal to each other, therefore the convex cones of missed directions at each iteration are minimal in size.

The convergence results for OrthoMads follow directly from those already published for Mads, and they hold deterministically, rather than with probability one, as for LtMads, the first Mads instance. The initial numerical results are quite good for both smooth and nonsmooth, and constrained and unconstrained problems considered here.

Keywords: Mesh Adaptive Direct Search algorithms (Mads), deterministic, orthogonal directions, constrained optimization, nonlinear programming.

*Work of the first author was supported by Fcar grant Nc72792 and Nserc grant 239436-05. The third author was supported by Lanl 94895-001-04 34, and both were supported by Afosr FA9550-07-1-0302, the Boeing Company, ExxonMobil Upstream Research Company.

†Air Force Institute of Technology, Department of Mathematics and Statistics, 2950 Hobson Way, Bldg 641, Wright Patterson Afb, Ohio 45433 USA, www.afit.edu/en/ENC/Faculty/MAbramson/abramson.html, Mark.Abramson@afit.edu.

‡GERAD and Département de mathématiques et de génie industriel, École Polytechnique de Montréal, C.P. 6079, Succ. Centre-ville, Montréal, Québec H3C 3A7 Canada, www.gerad.ca/Charles.Audet, Charles.Audet@gerad.ca.

§Computational and Applied Mathematics Department, Rice University - MS 134, 6100 South Main Street, Houston, Texas 77005-1892 USA, www.caam.rice.edu/∼dennis, dennis@rice.edu.

¶GERAD and Département de mathématiques et de génie industriel, Ecole Polytechnique de Montréal, C.P. 6079, Succ. Centre-ville, Montréal, Québec H3C 3A7 Canada, Sebastien.Le.Digabel@gerad.ca.

1 Introduction

This paper considers optimization problems of the form

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

where $f : \Omega \subset \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is typically evaluated through a black-box computer simulation with no available derivatives, and Ω is a set of feasible points also defined by black-box nonlinear constraint, or even Boolean, functions. Because no exploitable information on the nature of f or Ω exists, we consider direct search methods which only use functions evaluations to drive their search.

Mesh Adaptive Direct Search (MADS) is introduced in [4] as a SEARCH/POLL derivative-free class of methods with strong convergence properties. It extends the Generalized Pattern Search (GPS) method of [18]. The constraints are treated by the extreme barrier approach, which simply rejects points outside Ω by setting their objective function value to ∞. The first instance of this class of methods is called LTMADS.

LTMADS behaves well in practice, but it has drawbacks that we wish to correct in this paper. First, there is a probabilistic component to the choice of polling directions. For each new mesh size, a random direction is chosen to a current mesh point. That direction is completed somewhat randomly to a positive spanning set of directions from the current iterate to other current mesh points. The resulting algorithm is shown to have Clarke stationary point convergence with probability one. However, it has been observed [10] that this way of choosing polling directions can lead to undesirably large angles between some of the members of the LTMADS polling set at a given iteration.

The purpose of this paper is to introduce a new variant of MADS, which we call ORTHOMADS, that uses an orthogonal positive spanning set of polling directions and thus avoids large angles between polling directions. In Figure 3, we show some experiments in which the ORTHOMADS directions do seem better distributed than the LTMADS.

We show that ORTHOMADS shares the same theoretical convergence results as LTMADS, except that the convergence is not qualified by being of probability one. In the tests given here, ORTHOMADS performs generally better than LTMADS.

ORTHOMADS is detailed in Section 2, where we show a deterministic way to construct a polling set on the current mesh of orthogonal polling directions (the ORTHOMADS directions). Section 2 also gives the convergence results, based on those in [4]. Finally, we present numerical results in Section 3 and some concluding remarks in Section 4.

Notation: Throughout the text, $\| \cdot \|$ denotes the ℓ_2 norm, $e_i \in \mathbb{R}^n$ is the i^{th} coordinate vector, and $e \in \mathbb{R}^n$ is the vector whose components are all equal to 1. $B_\varepsilon(x)$ denotes the open ball of radius ε around x.

2 The ORTHOMADS algorithm

The ORTHOMADS algorithm is described in this section. We will not give details for the MADS class of algorithms and its LTMADS instantiation, since they are available in [4].

Each MADS iteration k is separated into two steps, the SEARCH and the POLL, where the objective function f and the test for feasibility are evaluated at finitely many trial points. These trial points lie on the mesh M_k defined by

$$M_k = \{x + \Delta_k^m Dz : x \in V_k, \ z \in \mathbb{N}^{n_D}\} \subset \mathbb{R}^n ,$$

where $V_k \subset \mathbb{R}^n$ is the set of all evaluated points by the start of the iteration, $\Delta_k^m \in \mathbb{R}_+$ is the mesh size parameter at iteration k, and D is a matrix in $\mathbb{R}^{n \times n_D}$ composed of n_D directions in $\mathbb{R}^n$. This paper focuses on the POLL step which is characterized by the set of trial points

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

where x_k is the POLL center at iteration k and D_k is the set of POLL directions, which have to form a positive spanning set and to be constructed so that POLL trial points lie on the mesh M_k. In GPS, a related method, the directions contained in D_k are always chosen among the columns of D. Therefore, in GPS, there is only the same finite number of possibilities for selecting the directions in every D_k.

The differences between LTMADS and ORTHOMADS lie in the way to generate the directions in D_k: With LTMADS, D_k is randomly generated and directions are not necessarily orthogonal, possibly leading to large angles between directions and large unexplored convex cones of directions at a given step. However, the union of all normalized LTMADS directions over all iterations k is dense in the unit sphere with probability one.

ORTHOMADS introduces a new way to generate the POLL directions D_k. This new method is deterministic and generates orthogonal directions, which together with their negatives form D_k, and such that the union of all normalized ORTHOMADS directions over all iterations is dense in the unit sphere. Furthermore, the components of these directions are integer, so that POLL points lie on the mesh defined with $D = [I_n \ -I_n]$, where I_n is the identity matrix in dimension n. The orthogonality of the ORTHOMADS directions offers a better distribution of the POLL trial points in the search space, and the advantage of determinism is that numerical results are now easily reproducible. Because of the random component of LTMADS, we felt that numerical experiments had to be performed on series of several runs to show the reader the variations in the results.

At each iteration of ORTHOMADS, the main steps for the construction of these directions are as follows. First, the pseudo-random Halton sequence produces one vector in $[0,1]^n$ (Subsection 2.1). Second, this vector is scaled and rounded to an appropriate length (Subsection 2.2). The resulting direction is called the *adjusted Halton* direction. Third, the Householder transformation is then applied to the adjusted Halton direction, producing n orthogonal and integer vectors, forming a basis for $\mathbb{R}^n$ (Subsection 2.3).

Finally, the basis is completed to a positive basis formed by $2n$ ORTHOMADS POLL directions D_k, by including in D_k the basis and its negatives (Subsection 2.4). Figure 1 summarizes these steps, and will be referred to throughout the section.

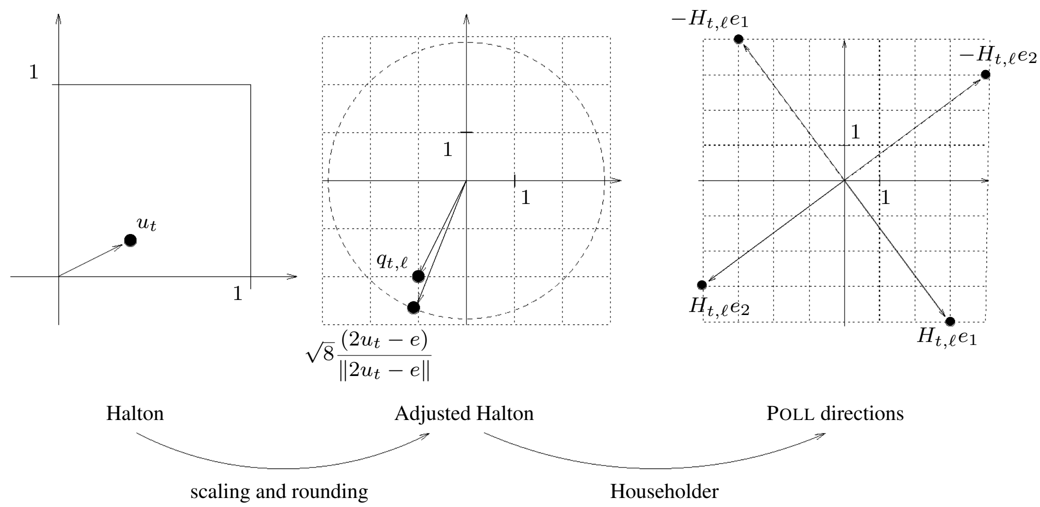

Figure 1: Example with $n = 2$ and $(t, \ell) = (6, 3)$. The Halton direction is $u_t = (3/8, 2/9)^T$, the adjusted Halton direction $q_{t,\ell} = (-1, -2)^T$ with $\alpha_{t,\ell} = 2$ and the set of POLL directions $D_k = [H_{t,\ell}\ -H_{t,\ell}]$ with $H_{t,\ell}e_1 = (3, -4)^T$ and $H_{t,\ell}e_2 = (-4, -3)^T$. Every POLL direction $d \in D_k$ satisfies $\Delta_k^m \|d\| = 5/64 < \Delta_k^p = 1/8$.

In this section we show that the ORTHOMADS directions meet all the conditions detailed in [2, 4], so that ORTHOMADS is a valid MADS instance and thus inherits all of its convergence properties.

2.1 The Halton sequence u_t

Halton [11] introduced a deterministic family of sequences that grow dense in the hypercube $[0, 1]^n$. We consider the simplest sequence of this family, whose t^{th} element is

$$u_t = (u_{t,p_1}, u_{t,p_2}, \ldots, u_{t,p_n})^T \in [0, 1]^n$$

where $p_1 = 2, p_2 = 3, p_3 = 5$ and p_j is the j^{th} prime number, and $u_{t,p}$ is the radical-inverse function in base p. More precisely,

$$u_{t,p} = \sum_{r=0}^{\infty} \frac{a_{t,r,p}}{p^{1+r}},$$

where the $a_{t,r,p} \in \mathbb{Z}_+$ are the unique coefficients of the base p expansion of t:

$$t = \sum_{r=0}^{\infty} a_{t,r,p} p^r .$$

Table 1 describes the first five elements of u_t for $n = 4$ (for example, $u_{5,3} = 1 \times 3^{-2} + 2 \times 3^{-1} = \frac{7}{9}$). Our specific sequence of u_t vectors is from this point addressed as the sequence of Halton directions.

Table 1: The sequence of Halton directions for $n = 4$ and $t = 0, 1, \ldots, 6$.

t	t in base				u_t			
	2	3	5	7	$u_{t,2}$	$u_{t,3}$	$u_{t,5}$	$u_{t,7}$
0	0	0	0	0	0	0	0	0
1	1	1	1	1	$1/2$	$1/3$	$1/5$	$1/7$
2	10	2	2	2	$1/4$	$2/3$	$2/5$	$2/7$
3	11	10	3	3	$3/4$	$1/9$	$3/5$	$3/7$
4	100	11	4	4	$1/8$	$4/9$	$4/5$	$4/7$
5	101	12	10	5	$5/8$	$7/9$	$1/25$	$5/7$
6	110	20	11	6	$3/8$	$2/9$	$6/25$	$6/7$

In order to remove the linear correlation of the last columns of u_t, it is proposed in [15] to exclude initial points of the Halton sequence. In the present work, we start the sequence at $t = n + 1$.

The following properties will be used in Subsection 2.2:

$$2u_t - e = 0 \quad \Leftrightarrow \quad n = t = 1 \tag{1}$$

$$|2u_{t,p_i} - 1| = |2u_{t,p_j} - 1| \quad \Leftrightarrow \quad t = 0 . \tag{2}$$

Property (2) follows from the fact that u_{t,p_i} and u_{t,p_j} can be written as reduced fractions with denominators that are powers of different prime numbers p_i and p_j.

The next result shows that the union of all the directions in the sequence of Halton is dense in $[0, 1]^n$, i.e. any direction $v \in [0, 1]^n$ is an accumulation point of the sequence $\{u_t\}_{t=1}^{\infty}$.

Proposition 2.1 *The Halton sequence* $\{u_t\}_{t=1}^{\infty}$ *is dense in* $[0, 1]^n$.

Proof. It suffices to show that for any vector $v \in [0, 1]^n$ and any $\varepsilon > 0$, there exists an integer t such that $\|u_t - v\| < \varepsilon$. A construction of such an integer t involves solving a system of n Diophantine equations, and existence of a solution is ensured by the Chinese Remainder Theorem [8], and by the fact that prime numbers are used in the definition of u_t. We refer the reader to [11] for a detailed proof. ∎

2.2 The adjusted Halton direction $q_{t,\ell}$

The directions in D_k used in the POLL step of MADS cannot be arbitrarily chosen, they must satisfy precise requirements. The Halton directions u_t do not satisfy these requirements and the first steps toward generating a satisfactory set D_k are to translate, scale and round u_t.

These operations depend on another integer parameter, ℓ, which is related to the mesh size parameter Δ_k^m (this relationship with Δ_k^m is unimportant at this point and will be detailed in Subsection 2.4). The parameter ℓ is used to transform the direction u_t into the *adjusted Halton direction* $q_{t,\ell} \in \mathbb{Z}^n$, a direction whose norm is close to $2^{|\ell|/2}$. Furthermore, the normalized direction $\frac{q_{t,\ell}}{\|q_{t,\ell}\|}$ will be constructed so that it is close to $\frac{2u_t - e}{\|2u_t - e\|}$. We already observed in (1) that $2u_t - e = 0$ is possible only if $n = 1$ and $t = 1$, and our algorithm never uses $t = 1$ (we begin our Halton sequence at $t = n + 1$, see Subsection 2.4).

In order to define $q_{t,\ell}$, we first introduce the following sequence of functions:

$$q_t(\alpha) \;=\; \text{round}\left(\alpha \frac{2u_t - e}{\|2u_t - e\|}\right) \;\in\; \mathbb{Z}^n \cap \left[-\alpha - \frac{1}{2}, \alpha + \frac{1}{2}\right]^n$$

where $\alpha \in \mathbb{R}_+$ is a scaling factor, and u_t is the t^{th} Halton direction. The function $q_t(\cdot)$ is a monotone non-decreasing step function on $\mathbb{R}_+$. Let $\alpha_{t,\ell}$ be a scalar such that $\|q_t(\alpha_{t,\ell})\|$ is as close as possible to $2^{|\ell|/2}$, without exceeding it:

$$\begin{array}{rcll} \alpha_{t,\ell} & \in & \underset{\alpha \in \mathbb{R}_+}{\text{argmax}} & \|q_t(\alpha)\| \\ & & \text{s.t.} & \|q_t(\alpha)\| \leq 2^{|\ell|/2} \,. \end{array} \tag{3}$$

Problem (3) can easily be solved using a bisection method. The adjusted Halton direction $q_{t,\ell}$ is defined to be equal to $q_t(\alpha_{t,\ell})$, and the following Lemma ensures that $q_{t,\ell}$ is a nonzero integer vector:

Lemma 2.2 *If $t \neq 0$, the adjusted Halton direction satisfies $\|q_{t,\ell}\| \geq 1$.*

Proof. From (2), if $t \neq 0$ and $\alpha = \frac{\|2u_t - e\|}{2\|2u_t - e\|_\infty}$, then $\|q_t(\alpha)\| = 1 \leq 2^{|\ell|/2}$ for all ℓ. ∎

The following lemma gives a lower bound on the value of $\alpha_{t,\ell}$. It will be used later to justify that $\alpha_{t,\ell}$ grows large with ℓ.

Lemma 2.3 *The optimal solution of Problem (3) satisfies $\alpha_{t,\ell} \geq \dfrac{2^{|\ell|/2}}{\sqrt{n}} - \dfrac{1}{2}$.*

Proof. Let $\alpha_{t,\ell}$ be an optimal solution of Problem (3) and set $q_{t,\ell} = q_t(\alpha_{t,\ell})$. Then every feasible solution α to Problem (3) satisfies

$$\begin{aligned} \|q_t(\alpha)\|^2 &= \left\| \text{round}\left(\frac{\alpha(2u_t - e)}{\|2u_t - e\|} \right) \right\|^2 \\ &= \sum_{i=1}^{n} \text{round}\left(\frac{\alpha(2u_t^i - 1)}{\|2u_t - e\|} \right)^2 \\ &\leq \sum_{i=1}^{n} \left(\alpha + \frac{1}{2} \right)^2 = n \left(\alpha + \frac{1}{2} \right)^2 . \end{aligned}$$

Define $\beta = \frac{2^{|\ell|/2}}{\sqrt{n}} - \frac{1}{2}$. Then β is feasible for Problem (3), since $\|q_t(\beta)\|^2 \leq n(\beta + \frac{1}{2})^2 = 2^{|\ell|}$; therefore, $\alpha_{t,\ell} \geq \beta$. ∎

Table 2 shows elements of the sequences u_t and $q_{t,\ell}$ for $n = 4$ and eight pairs (t, ℓ) whose values are compatible with the ORTHOMADS algorithm presented in Subsection 2.4. The values of $\alpha_{t,\ell}$ and the square norm $\|q_{t,\ell}\|^2$ are also reported. One can also notice that $\alpha_{t,\ell}$ often differs from $2^{|\ell|/2}$. In the example illustrated in Figure 1, $(t, \ell) = (6, 3)$ and $q_t(\alpha) = \text{round}\left(\frac{\alpha}{\sqrt{481}}(-9, -20)^T \right)$. An optimal solution of (3) is $\alpha_{t,\ell} = 2$ and satisfies $\|q_{t,\ell}\| = \sqrt{5} < \sqrt{8} = 2^{|\ell|/2} < \|q_t(2^{|\ell|/2})\| = \|(-1, -3)^T\| = \sqrt{10}$.

Table 2: The sequence of Halton directions u_t and the adjusted Halton directions $q_{t,\ell}$ for $n = 4$ and eight pairs (t, ℓ).

(t, ℓ)	u_t				$\alpha_{t,\ell}$	$q_{t,\ell}$				$\|q_{t,\ell}\|^2$
	$u_{t,2}$	$u_{t,3}$	$u_{t,5}$	$u_{t,7}$						
$(5, 0)$	$5/8$	$7/9$	$1/25$	$5/7$	1.0	0	0	-1	0	1
$(6, 1)$	$3/8$	$2/9$	$6/25$	$6/7$	1.0	0	-1	0	1	2
$(7, 2)$	$7/8$	$5/9$	$11/25$	$1/49$	1.0	1	0	0	-1	2
$(8, 3)$	$1/16$	$8/9$	$16/25$	$8/49$	2.5	-2	1	1	-1	7
$(9, 4)$	$9/16$	$1/27$	$21/25$	$15/49$	4.0	0	-3	2	-1	14
$(10, 5)$	$5/16$	$10/27$	$2/25$	$22/49$	5.5	-2	-1	-5	-1	31
$(11, 6)$	$13/16$	$19/27$	$7/25$	$29/49$	7.7	5	4	-4	2	61
$(12, 7)$	$3/16$	$4/27$	$12/25$	$36/49$	11.0	-7	-7	0	5	123

The following proposition gives a property of the scaling and rounding operations, which transform a vector v into $q = \text{round}(\alpha v/\|u\|)$. The property states that the directions $v/\|v\|$ and $q/\|q\|$ are arbitrarily close for sufficient large values of α:

Proposition 2.4 *Let $v \neq 0$ be a vector in $\mathbb{R}^n$. For any $\varepsilon > 0$, if $\alpha > \frac{2\sqrt{n}}{\varepsilon} + \frac{\sqrt{n}}{2}$ and $q = \mathrm{round}\left(\alpha \frac{v}{\|v\|}\right) \neq 0$, then $\left\| \frac{q}{\|q\|} - \frac{v}{\|v\|} \right\| < \frac{\varepsilon}{2}$.*

Proof. Consider $\varepsilon > 0$ and $\alpha > 2\sqrt{n}/\varepsilon + \sqrt{n}/2$. The vector q may be expressed as $q = \alpha \frac{v}{\|v\|} + \delta$, where $\delta = (\delta_1, \delta_2, \ldots, \delta_n)^T$ and $|\delta_i| < 1/2$ for all $i = 1, 2, \ldots, n$. It follows that

$$\begin{aligned} \left\| \frac{q}{\|q\|} - \frac{v}{\|v\|} \right\| &= \left\| \left(\frac{\alpha}{\|q\|} - 1 \right) \frac{v}{\|v\|} + \frac{\delta}{\|q\|} \right\| \\ &\leq \left\| \left(\frac{\alpha}{\|q\|} - 1 \right) \frac{v}{\|v\|} \right\| + \left\| \frac{\delta}{\|q\|} \right\| \\ &= \frac{|\alpha - \|q\||}{\|q\|} + \frac{\|\delta\|}{\|q\|}. \end{aligned}$$

The norm of q can be bounded with $\alpha \frac{\|v\|}{\|v\|} - \|\delta\| \leq \|q\| \leq \alpha \frac{\|v\|}{\|v\|} + \|\delta\|$ and therefore $|\alpha - \|q\|| \leq \|\delta\|$. Furthermore, $\alpha > 2\sqrt{n}/\varepsilon + \sqrt{n}/2 > \sqrt{n}/2$ and $\|\delta\| < \sqrt{n}/2$ implies that α satisfies $0 < \alpha - \|\delta\|$. It follows that

$$\left\| \frac{q}{\|q\|} - \frac{v}{\|v\|} \right\| \leq \frac{2\|\delta\|}{\|q\|} \leq \frac{2\|\delta\|}{\alpha - \|\delta\|} < \frac{\sqrt{n}}{\alpha - \sqrt{n}/2} < \frac{\varepsilon}{2}.$$

■

2.3 Construction of an orthogonal integer basis

This subsection gives a way to transform a sequence of directions into a sequence of orthogonal bases. Given an integer nonzero vector $q \in \mathbb{Z}^n$, we apply the (symmetric) scaled Householder transformation [12] to construct an orthogonal basis for $\mathbb{R}^n$ composed of integer vectors:

$$H = \|q\|^2 (I_n - 2vv^T), \text{ where } v = \frac{q}{\|q\|}. \tag{4}$$

Proposition 2.5 *The columns of H form an integer orthogonal basis for $\mathbb{R}^n$.*

Proof. First, the columns of H are mutually orthogonal, since $v^T v = 1$ and

$$\begin{aligned} H^T H &= \|q\|^4 (I_n - 2vv^T)^T (I_n - 2vv^T) \\ &= \|q\|^4 (I_n - 2vv^T - 2vv^T + 4vv^T vv^T) = \|q\|^4 I_n. \end{aligned}$$

Second, by dividing the previous equation by $\|q\|^4$ and applying symmetry, we reveal the inverse of H as $H^{-1} = \frac{1}{\|q\|^4} H$. Since H^{-1} exists, the columns of H form a basis in $\mathbb{R}^n$. Finally, the entries of

$$H \;=\; \|q\|^2 I_n - 2\|q\|^2 \frac{q}{\|q\|} \frac{q^T}{\|q\|} \;=\; \|q\|^2 I_n - 2qq^T$$

are integer, since q and $\|q\|^2$ are integer. ∎

The next proposition shows that the Householder transformation applied to a dense set of normalized directions produces a dense set of normalized directions:

Proposition 2.6 *For $t = 1, 2, \ldots$, let $v_t = \frac{q_t}{\|q_t\|}$ and $H_t = \|q_t\|^2 (I_n - 2v_t v_t^T)$. If $\{v_t\}_{t=1}^{\infty}$ is dense on the unit sphere, then the normalized sequence composed of the i^{th} columns of H_t, $\left\{\frac{H_t e_i}{\|H_t e_i\|}\right\}_{t=1}^{\infty}$ is dense on the unit sphere.*

Proof. Let $w \in \mathbb{R}^n$ with $\|w\| = 1$ be an arbitrarily unit vector, $\varepsilon > 0$ be some small positive number, and $i \in \{1, 2, \ldots, n\}$ be the index of a column. For $n > 1$ ($n = 1$ is trivial), we need to show that there exists an index $t \in \mathbb{N}$ such that the ith column of H_t, $H_t e_i$, satisfies

$$\left\| \frac{H_t e_i}{\|H_t e_i\|} - w \right\| \;<\; \varepsilon.$$

First, observe that $\|H_t e_i\| = \sqrt{e_i^T H_t^T H_t e_i} = \|q_t\|^2$, and therefore $\frac{H_t e_i}{\|H_t e_i\|} = e_i - 2 v_t v_t^T e_i$. Now, define the vector

$$d = \begin{cases} \frac{1}{\sqrt{2(1-w_i)}}(e_i - w) & \text{if } w_i < 1, \\ e_{i+1} \text{ (the sum } i+1 \text{ is modulo } n) & \text{otherwise.} \end{cases}$$

Observe that if $w_i = 1$ then the vector d satisfies $\|d\| = 1$ and $2d_i d = 0 = e_i - w$, and if $w_i < 1$ then

$$\begin{aligned} \|d\| \;&=\; \sqrt{d^T d} \;=\; \sqrt{\frac{1}{2(1-w_i)}(e_i - w)^T (e_i - w)} \;=\; 1, \\ 2 d_i d \;&=\; \frac{1}{(1-w_i)}(e_i - w)_i (e_i - w) \;=\; e_i - w. \end{aligned}$$

By assumption, $\{v_t\}_{t=1}^{\infty}$ is dense on the unit sphere, and therefore there exists some index t such that $v_t = d + \delta$, where $\delta \in \mathbb{R}^n$ is small enough to satisfy $\|\delta_i(d+\delta) + d_i \delta\| <$

$\varepsilon/2$. The proof may be completed as follows:

$$\begin{aligned}
\left\| \frac{H_t e_i}{\|H_t e_i\|} - w \right\| &= \left\| e_i - 2 v_t v_t^T e_i - w \right\| \\
&= \left\| e_i - 2(d+\delta)(d+\delta)^T e_i - w \right\| \\
&= \| e_i - 2(d_i+\delta_i)(d+\delta) - w \| \\
&= \| e_i - 2 d_i d - w - 2(\delta_i(d+\delta) + d_i \delta) \| \\
&= \| e_i - (e_i - w) - w - 2(\delta_i(d+\delta) + d_i \delta) \| \\
&= 2 \, \| \delta_i(d+\delta) + d_i \delta \| \; < \; \varepsilon.
\end{aligned}$$

■

In Figure 1, the Householder transformation is applied to $q_{t,\ell} = (-1, -2)^T$ and produces the integer orthogonal basis $H_{t,\ell} = \begin{bmatrix} 3 & -4 \\ -4 & -3 \end{bmatrix}$.

2.4 The ORTHOMADS instance of MADS

The new ORTHOMADS instance of MADS can be now defined by combining the components introduced in Subsections 2.1–2.3. The POLL set P_k used by ORTHOMADS at iteration k is entirely determined by the values of the pair t_k and ℓ_k. The t_k^{th} element of the Halton sequence u_{t_k} is used to create the adjusted Halton direction q_{t_k,ℓ_k} whose norm is as close as possible to $2^{|\ell_k|/2}$. The Householder transformation on q_{t_k,ℓ_k} produces an orthogonal integer basis H_{t_k,ℓ_k}, and the norm of each column is close to $2^{|\ell_k|}$.

The LTMADS and ORTHOMADS algorithms are identical except for the construction of the set P_k and the POLL directions D_k. The set of directions $D = [I_n \;\; -I_n]$ defining the mesh M_k and the mesh update parameters $\tau = 4$, $w^- = -1$ and $w^+ = 1$ are the same for both algorithms. The mesh size parameter Δ_k^m and the POLL size parameter Δ_k^p are still defined with the integer ℓ_k, except that it is allowed to be negative. This extension is not specific to ORTHOMADS and can be applied in LTMADS as well: at each iteration k, the POLL and mesh size parameters are entirely defined by the value of ℓ_k:

$$\Delta_k^p = 2^{-\ell_k} \text{ and } \Delta_k^m = \begin{cases} 4^{-\ell_k} & \text{if } \ell_k > 0 \\ 1 & \text{otherwise.} \end{cases} \tag{5}$$

At iteration $k = 0$, ℓ_k is set to 1 and $\Delta_0^m = \Delta_0^p = 1$. The mesh and POLL size parameters always satisfy $\Delta_k^m \leq \Delta_k^p$ and $\Delta_k^m 2^{|\ell_k|} = \Delta_k^p$.

In the update step of iteration k, if no new incumbent is found, the iteration is said to be unsuccessful and $\ell_{k+1} \leftarrow \ell_k + 1$. Otherwise, the iteration is a success and $\ell_{k+1} \leftarrow \ell_k - 1$. The MADS algorithm generates POLL trial points at a distance of order Δ_k^p from the POLL center, on a mesh M_k of size Δ_k^m. At an unsuccessful iteration, Δ_k^m is reduced

[0] Initializations

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

[1] Iteration k

SEARCH (optional)

  evaluate f on $S_k \subset M_k$

POLL (optional is the SEARCH was successful)

  if the POLL size is the smallest one so far

  $\big($i.e., if $\Delta_k^p = \min\{\Delta_j^p : j = 0, 1, \ldots, k\}\big)$

    $t_k \leftarrow \ell_k + n + 1$

  else (i.e., smaller POLL sizes were considered)

    $t_k \leftarrow \max\{t_j : j = 0, 1, \ldots, k-1\}$

  compute u_{t_k}, q_{t_k,ℓ_k}, H_{t_k,ℓ_k}, and $D_k = [H_{t_k,\ell_k} \; - H_{t_k,\ell_k}]$

  evaluate f on $P_k \subset M_k$

[2] Updates

if the iteration is successful

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

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

else (iteration failure)

  $x_{k+1} \leftarrow x_k$

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

$k \leftarrow k + 1$

goto [1] if no stopping condition is met

Figure 2: The ORTHOMADS algorithm.

faster than Δ_k^p and the number of possible POLL trial points increases, allowing more flexibility in the choice of the POLL directions D_k.

Figure 2 describes our algorithm. The POLL directions D_k depend entirely on the two integers t_k and ℓ_k. These integers are chosen to ensure that there will be a sequence of unsuccessful iterations for which the mesh size parameter goes to zero, and such that the directions used in that subsequence will be the tail of the entire Halton sequence. In order to accomplish that goal, we keep track of the value of the smallest POLL size parameter visited so far. At every iteration where Δ_k^p is equal to that value, we set $t_k = \ell_k + n + 1$. A consequence of this way of fixing t_k is that the set of ordered indices

$$U := \{k_1, k_2, \ldots\} = \{k \ : \ \text{iteration } k \text{ is unsuccessful, and } \Delta_k^p \leq \Delta_j^p \ \forall j = 0, 1, \ldots k\}$$

satisfies $(t_{k_1}, \ell_{k_1}) = (n+1, 0)$, $(t_{k_2}, \ell_{k_2}) = (n+2, 1), \ldots, (t_{k_i}, \ell_{k_i}) = (n+i, i-1)$, and the set of Halton directions $\{u_{t_k}\}_{k \in U}$ is precisely $\{u_t\}_{t=n+1}^{\infty}$.

At the other iterations, those for which smaller POLL sizes were previously considered, we just keep increasing t_k so that a new Halton direction is used. Examples of

Table 3: Example of ORTHOMADS iterations for $n = 4$. Iterations $k \in \{4, 5, 8\}$ correspond to failed iterations with consecutive Halton elements $t_k = 5, 6$ and 7 satisfying $t_k = \ell_k + n + 1$.

k	Succ/Fail	$(t_k$,	$\ell_k)$	Δ_k^m	Δ_k^p	$\|D_k e_i\|$
0	S	$(\mathbf{5}$,	$\mathbf{0})$	1	1	1
1	S	$(6$,	$-1)$	1	2	2
2	F	$(7$,	$-2)$	1	4	4
3	F	$(8$,	$-1)$	1	2	2
4	F	$(\mathbf{5}$,	$\mathbf{0})$	1	1	1
5	F	$(\mathbf{6}$,	$\mathbf{1})$	$1/4$	$1/2$	2
6	S	$(\mathbf{7}$,	$\mathbf{2})$	$1/16$	$1/4$	4
7	F	$(9$,	$1)$	$1/4$	$1/2$	2
8	F	$(\mathbf{7}$,	$\mathbf{2})$	$1/16$	$1/4$	5
9	S	$(\mathbf{8}$,	$\mathbf{3})$	$1/64$	$1/8$	8

pairs (t_k, ℓ_k) can be seen in Table 3. The boldface entries are those where the POLL size parameter is the smallest one so far. In this example, the first three indices of U would be $\{4, 5, 8\}$.

As in LTMADS, the basis H_{t_k,ℓ_k} is completed to a maximal positive basis composed of $2n$ directions,

$$D_k = [H_{t_k,\ell_k} \quad -H_{t_k,\ell_k}],$$

the set of POLL directions. A minimal positive basis with $n+1$ directions is not considered in order to keep orthogonal directions. Table 4 illustrates ORTHOMADS bases H_{t_k,ℓ_k}, with possible pairs (t_k, ℓ_k).

Notice that any direction $D_k e_i$ $(1 \le i \le 2n)$ satisfies $\|D_k e_i\| = \|q_{t,\ell}\|^2 \le (2^{|\ell|/2})^2 = 2^{|\ell|}$ and $\|D_k e_i\| \le 2^{|\ell|}$. Therefore, the POLL trial point $x_k + \Delta_k^m D_k e_i$ is at an euclidean distance of at most $\Delta_k^m 2^{|\ell|} = \Delta_k^p$ from the POLL center. This distance is comparable to that used in LTMADS, where the POLL trial points are exactly at a distance Δ_k^p (using the ℓ_∞ norm) from the POLL center.

We conclude this section with the following propositions that show that ORTHOMADS has the same convergence properties as in [4] with no need for a probabilistic argument.

Proposition 2.7 *The set of normalized directions* $\left\{\frac{q_{t,\ell}}{\|q_{t,\ell}\|}\right\}_{t=1}^{\infty}$ *with* $\ell = t-n-1$ *is dense in the unit sphere.*

Proof. Let $\varepsilon > 0$ and $d \in \mathbb{R}^n$ with $\|d\| = 1$. Proposition 2.1 states that the Halton sequence $\{u_t\}_{t=1}^{\infty}$ is dense in the unit cube $[0,1]^n$. Therefore, there exists an index t such that $\frac{2^{|t-n-1|/2}}{\sqrt{n}} - \frac{1}{2} > \frac{2\sqrt{n}}{\varepsilon} + \frac{\sqrt{n}}{2}$ and $\left\|\frac{2u_t - e}{\|2u_t - e\|} - d\right\| \le \frac{\varepsilon}{2}$.

Table 4: A sequence of ORTHOMADS bases corresponding to seven consecutive failed iterations. Pairs (t_k, ℓ_k) correspond to consecutive Halton elements $t = 5, 6, \ldots, 12$ with $t_k = \ell_k + n + 1$.

(t_k, ℓ_k) $\|H_{t_k,\ell_k} e_i\|$	H_{t_k,ℓ_k}	(t_k, ℓ_k) $\|H_{t_k,\ell_k} e_i\|$	H_{t_k,ℓ_k}
$(5,0)$ 1	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$(9,4)$ 14	$\begin{bmatrix} 14 & 0 & 0 & 0 \\ 0 & -4 & 12 & -6 \\ 0 & 12 & 6 & 4 \\ 0 & -6 & 4 & 12 \end{bmatrix}$
$(6,1)$ 2	$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \end{bmatrix}$	$(10,5)$ 31	$\begin{bmatrix} 23 & -4 & -20 & -4 \\ -4 & 29 & -10 & -2 \\ -20 & -10 & -19 & -10 \\ -4 & -2 & -10 & 29 \end{bmatrix}$
$(7,2)$ 2	$\begin{bmatrix} 0 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix}$	$(11,6)$ 61	$\begin{bmatrix} 11 & -40 & 40 & -20 \\ -40 & 29 & 32 & -16 \\ 40 & 32 & 29 & 16 \\ -20 & -16 & 16 & 53 \end{bmatrix}$
$(8,3)$ 7	$\begin{bmatrix} -1 & 4 & 4 & -4 \\ 4 & 5 & -2 & 2 \\ 4 & -2 & 5 & 2 \\ -4 & 2 & 2 & 5 \end{bmatrix}$	$(12,7)$ 123	$\begin{bmatrix} 25 & -98 & 0 & 70 \\ -98 & 25 & 0 & 70 \\ 0 & 0 & 123 & 0 \\ 70 & 70 & 0 & 73 \end{bmatrix}$

Lemma 2.3 ensures that that $\alpha_{t,\ell} \geq \frac{2^{|\ell|/2}}{\sqrt{n}} - \frac{1}{2} > \frac{2\sqrt{n}}{\varepsilon} + \frac{\sqrt{n}}{2}$. Combining this last inequality with Proposition 2.4 gives

$$\begin{aligned}\left\| \frac{q_{t,\ell}}{\|q_{t,\ell}\|} - d \right\| &\leq \left\| \frac{q_{t,\ell}}{\|q_{t,\ell}\|} - \frac{2u_t - e}{\|2u_t - e\|} \right\| + \left\| \frac{2u_t - e}{\|2u_t - e\|} - d \right\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \,.\end{aligned}$$

■

This allows us to state our main result:

Theorem 2.8 ORTHOMADS *is a valid* MADS *instance.*

Proof. In order to show that ORTHOMADS is a valid MADS instance we need to show that the POLL directions satisfy the following four properties [2, 4]:

- *Any direction $D_k e_i$ $(1 \leq i \leq 2n)$ can be written as a non-negative integer combination of the directions of D*: This is the case by construction.

- *The distance from the* POLL *center x_k to a* POLL *trial point (in ℓ_∞ norm) has to be bounded above by Δ_k^p*: This is also the case by construction because we ensured that $\|D_k e_i\| \leq 2^{|\ell_k|}$ for all i in $\{1, 2, ..., 2n\}$ and $\|\Delta_k^m D_k e_i\|_\infty \leq \|\Delta_k^m D_k e_i\| \leq \Delta_k^m 2^{|\ell_k|} = \Delta_k^p$.

- *Limits (as defined in [7]) of convergent subsequences of the normalized sets $\overline{D_k} = \{d/\|d\| : d \in D_k\}$ are positive spanning sets.* This can be shown the same way as in [2] where the proof for LTMADS is detailed, since, for ORTHOMADS and with $\overline{H_{t_k,\ell_k}} = \{d/\|d\| : d \in H_{t_k,\ell_k}\}$, $\det\left(\overline{H_{t_k,\ell_k}}\right) = -1$.

- *The set of normalized directions used over all failed iterations is dense in the unit sphere*: The strategy chosen for the values of t_k and ℓ_k ensures that there exists a sequence of failed iterations corresponding to consecutive values of t_k. These iterations $k \in U$ can be chosen to correspond to large values of ℓ_k because, from [4], $\lim\limits_{\substack{k \in U \\ k \to \infty}} \Delta_k^m = 0$, and $\Delta_k^m = 4^{-\ell_k}$ for $\ell_k \geq 0$. For $k \in U$, the sets of directions $\{D_k\}_{k \in U}$ are constructed from consecutive directions q_{t_k,ℓ_k}, which are dense in the unit sphere after normalization (Proposition 2.7). Then, from Proposition 2.6 and since $D_k = [H_{t_k,\ell_k} \;\; -H_{t_k,\ell_k}]$, the set of normalized directions $\left\{\frac{D_k e_i}{\|D_k e_i\|}\right\}_{k \in U}$ is also dense in the unit sphere for all $i = 1, 2, \ldots, 2n$.

■

3 Numerical Tests

In this section, OrthoMads is compared to its predecessor LtMads [4] and to the Gps method [18], on 45 problems from the literature. In the Mads algorithms, the theory supports handling constraints by the extreme barrier approach: Points outside Ω are simply ignored and f is not evaluated. For Gps, the extreme barrier approach is supported by the theory only for a finite number of linear constraints [13]. Still, for comparison, we apply two different approaches: the extreme barrier (Gps-EB), and the filter method described in [3] (Gps-Filter), which has stronger theoretical support.

Because of its random behavior, 30 instances of LtMads are performed for each problem. Gps and OrthoMads are scored by comparing them against the 30 LtMads instances. A score of s for Gps or OrthoMads means that this instance gave a value of f at least as good as s of the 30 LtMads instances, with a relative precision of 1%. The worst score is 0 and the perfect score corresponds to 30. We consider that a bad instance has a score less than 10, an acceptable instance is between 10 and 19, and a good instance has a score greater than or equal to 20.

The integer ℓ_k (see (5)), defining the mesh and poll size parameters Δ_k^m and Δ_k^p at iteration k, is allowed to be negative for both LtMads and OrthoMads. Maximal positive bases ($2n$ directions) are used in the three methods, as is the opportunistic strategy (the poll is interrupted at the first success), and the optimist strategy: after a successful point has been found, a search point is generated further along the same direction. No other search is performed. The stopping criteria is satisfied when the poll size parameter Δ_k^p drops below 1E-12 or when the number of function evaluations reaches $1000n$.

The methods are tested on 45 problems divided into 4 groups: our choice of smooth and nonsmooth unconstrained problems is the same as in [10] and [9], respectively, with 21 smooth problems from the CUTEr test set [16] and 13 nonsmooth problems from [14], which is a compilation of nonsmooth problems from the literature. We also tested on 9 constrained problems from [5, 6, 14], and in addition, we added two problems from [1] that correspond to real applications.

All results and problem descriptions are summarized in Tables 5–9, where $f(x^*)$ corresponds to the best known minimal value of f, $value$ to the final value of f for each method, and $evals$ to the number of function evaluations that each method performed. Tables 5 and 6 show results on the 21 unconstrained smooth problems from CUTEr. OrthoMads has a perfect score on 17 of these problems. Table 7 displays results on the 13 unconstrained nonsmooth problems, where OrthoMads achieves good scores on 7 problems. Table 8 shows results for the 9 constrained problems. The same number of problems (4) is considered good and bad for OrthoMads. Finally, Table 9 presents results for the two real applications, and OrthoMads has perfect scores on both of them.

Table 10 summarizes the results. The first observation is that both Mads instances

Table 5: CUTEr unconstrained smooth problems (1 of 2). A score of s for a method indicates that the final f value is at least as good as s of the 30 LTMADS runs (with a relative error of 1%).

Problem n $f(x^*)$	LTMADS×30 worst *evals* *value*	 best *evals* *value*	GPS *evals* *value*	 *score*	ORTHOMADS *evals* *value*	 *score*
ARWHEAD 10 0.00	6128 0.00	650 0.00	1039 0.00	30	660 0.00	30
ARWHEAD 20 0.00	20000 0.00	1285 0.00	2079 0.00	30	1320 0.00	30
BDQRTIC 10 11.9	6320 18.3	4497 18.3	3510 18.3	30	5763 18.3	30
BDQRTIC 20 35.4	20000 58.3	17884 58.3	17074 58.3	30	20000 58.3	30
BIGGS6 6 0.00	831 2.06	570 2.06	764 2.06	30	713 2.06	30
BROWNAL10 10 0.00	10000 0.07	10000 0.00	10000 0.06	1	10000 0.04	3
BROWNAL20 20 0.00	20000 0.34	20000 0.00	20000 0.16	3	20000 0.01	20
PENALTY1 10 7.09E-5	10000 7.09E-5	10000 7.09E-5	10000 8.82E-5	0	10000 7.09E-5	30
PENALTY1 20 1.58E-4	20000 1.58E-4	20000 1.58E-4	20000 1.88E-4	0	20000 1.58E-4	30
PENALTY2 10 0.294E-3	10000 1.280E-3	10000 1.241E-3	10000 1.243E-3	30	10000 1.250E-3	30
PENALTY2 20 0.829E-2	20000 1.080E-2	20000 1.078E-2	20000 1.152E-2	0	20000 1.079E-2	30

. . . continued on Table 6

Table 6: CUTEr unconstrained smooth problems (2 of 2). A score of s for a method indicates that the final f value is at least as good as s of the 30 LTMADS runs (with a relative error of 1%).

Problem n $f(x^*)$	LTMADS×30 worst *evals* *value*	 best *evals* *value*	GPS *evals* *value*	 *score*	ORTHOMADS *evals* *value*	 *score*
POWELLSG 12 0.00	12000 0.00	12000 0.00	10093 0.00	30	12000 0.00	30
POWELLSG 20 0.00	20000 0.00	20000 0.00	20000 0.00	30	20000 0.00	30
SROSENBR 10 0.00	10000 6.31	10000 0.00	10000 0.00	30	10000 0.06	25
SROSENBR 20 0.00	20000 16.52	20000 0.48	20000 0.00	30	1958 0.00	30
TRIDIA 10 0.00	10000 0.00	7591 0.00	9317 0.00	30	10000 0.00	30
TRIDIA 20 0.00	20000 0.00	20000 0.00	20000 0.00	30	20000 0.00	30
VARDIM 10 0.00	1000 0.00	8163 0.00	10000 4.01	0	10000 0.00	30
VARDIM 20 0.00	20000 0.00	20000 0.00	20000 110.59	0	20000 110.84	0
WOODS 12 0.00	10951 104.91	7327 104.91	8433 104.91	30	9479 104.91	30
WOODS 20 0.00	20000 174.84	19181 174.84	20000 174.84	30	20000 174.84	30
			average scores	20.2		26.6

Table 7: Results for unconstrained nonsmooth problems from [14]. A score of s for a method indicates that the final f value is at least as good as s of the 30 LTMADS runs (with a relative error of 1%).

Problem	LTMADS×30		GPS		ORTHOMADS	
	worst	best				
n $f(x^*)$	*evals* *value*	*evals* *value*	*evals* *value*	*score*	*evals* *value*	*score*
ELATTAR	456	1795	2392	16	3984	20
6 0.560	8.021	0.563	1.714		1.504	
EVD61	490	3280	920	5	4224	22
6 0.0349	1.6001	0.0417	0.5443		0.0709	
FILTER	1293	1761	1132	7	1332	9
9 0.00619	0.00971	0.00797	0.00950		0.00935	
GOFFIN	50000	50000	24097	30	16842	30
50 0.00	1.10	0.06	0.00		0.00	
HS78	403	1026	819	13	405	13
5 -2.92	10.00	-2.88	0.00		0.00	
L1HILB	17953	50000	8738	0	50000	14
50 0.00	1.84	0.04	3.95		0.22	
MXHILB	11523	20377	9384	0	20755	3
50 0.00	0.280	0.003	0.976		0.197	
OSBORNE2	2046	5414	1660	0	4555	20
11 0.0480	0.1703	0.0549	0.2799		0.1089	
PBC1	1211	1127	677	2	1291	17
5 0.0223	0.4146	0.0343	0.3845		0.1602	
POLAK2	1449	1742	1327	30	949	30
10 54.6	54.6	54.6	54.6		54.6	
SHOR	1345	882	787	0	1087	30
5 22.6	22.9	22.6	23.5		22.8	
WONG1	1161	2109	1100	30	1823	30
7 681	699	693	697		693	
WONG2	5403	5403	1871	0	4977	0
10 24.3	31.4	24.8	47.4		32.8	
			average scores	10.2		18.3

Table 8: Results for constrained problems. A score of s for a method indicates that the final f value is at least as good as s of the 30 LTMADS runs (with a relative error of 1%).

Problem n m $f(x^*)$	LTMADS×30 worst *evals* *value*	LTMADS×30 best *evals* *value*	GPS-FILTER *evals* *value*	GPS-FILTER *score*	GPS-EB *evals* *value*	GPS-EB *score*	ORTHOMADS *evals* *value*	ORTHOMADS *score*
CRESCENT10 [5]	1279	4473	2152	0	1172	0	5497	30
10 2 -9.00	-8.26	-8.95	-6.19		-2.32		-8.97	
DISK10 [5]	1909	2626	2322	0	1143	0	2359	30
10 1 -17.3	-17.2	-17.3	-13.0		-10.0		-17.2	
B250 [6]	60000	60000	15412	0	27773	0	60000	0
60 1 7.95	15.41	7.99	1142.01		1116.03		16.92	
B500 [6]	16705	15359	11189	0	18912	0	29858	6
60 1 104	557	104	1235		1254		277	
G2 [6]	3880	6461	2056	4	2689	29	5414	22
10 2 -0.728	-0.181	-0.728	-0.221		-0.706		-0.561	
G2 [6]	10877	15722	6376	3	6551	29	20000	29
20 2 -0.804	-0.203	-0.736	-0.241		-0.721		-0.711	
HS114 [14]	1506	2135	1756	0	1756	0	1661	4
9 6 -1769	-1012	-1312	-968		-968		-1016	
MAD6 [14]	1122	1542	1378	22	1378	22	1671	7
5 7 0.102	0.113	0.102	0.103		0.103		0.108	
PENTAGON [14]	859	2525	601	0	601	0	980	19
6 15 -1.86	-1.60	-1.86	0.00		0.00		-1.81	
			average scores	3.2		8.9		16.3

Table 9: Results for real applications. A score of s for a method indicates that the final f value is at least as good as s of the 30 LTMADS runs (with a relative error of 1%). Displayed z values for problem STY are divided by 10^7.

Problem	LTMADS×30 worst	LTMADS×30 best	GPS-FILTER		GPS-EB		ORTHOMADS	
n m $f(x^*)$	$evals$ $value$	$evals$ $value$	$evals$ $value$	$score$	$evals$ $value$	$score$	$evals$ $value$	$score$
MDO [1]	1767	15	2719	0	2048	0	1212	30
10 10 -3964	-2530	-3964	-1386		-1386		-3964	
STY [1]	1590	1189	2073	22	2113	0	1214	30
8 11 -3.35	-2.88	-3.29	-3.11		-2.82		-3.27	
			average scores	11.0		0.0		30.0

outperform GPS. For 25 problems out of 45, ORTHOMADS found the same solution as the best of 30 LTMADS runs. The new method solved 32 problems out of 45 problems efficiently enough that, for these problems, the single run of ORTHOMADS was better than two thirds of the 30 LTMADS runs. For 4 problems, the two methods performed equally well, and for 9 problems, at least two thirds of the LTMADS runs gave a better solution than the one produced by ORTHOMADS.

Table 10: Summary for the GPS and ORTHOMADS performances. F, E and O correspond respectively to GPS-FILTER, GPS-EB, and ORTHOMADS. A bad instance has a score between 0 and 9, an acceptable (acc.) instance a score between 10 and 19, a good instance a score higher than 20 and a perfect (perf.) instance has a score of 30.

problems	average scores (on 30)			# of problems	# of bad instances			# of acc. instances			# of good instances			# of perf. instances		
	F	E	O		F	E	O	F	E	O	F	E	O	F	E	O
smooth	20.2	20.2	26.6	21	7	7	2	0	0	0	14	14	19	14	14	17
nonsmooth	10.2	10.2	18.3	13	8	8	3	2	2	3	3	3	7	3	3	4
constrained	3.2	8.9	16.3	9	8	6	4	0	0	1	1	3	4	0	0	2
real appli.	11.0	0.0	30.0	2	1	2	0	0	0	0	1	0	2	0	0	2
total or avg	13.5	14.2	22.3	45	24	23	9	2	2	4	19	20	32	17	17	25

Figure 3 illustrates the spread of the directions for both LTMADS and ORTHOMADS. Rosenbrock's function [17] with $n = 2$ and $n = 3$ was used with 2000 and 3000 evaluations, respectively. In the two-dimensional case, all the normalized directions used to generate POLL trial points are directly represented on the top two subfigures. It is clear that ORTHOMADS directions are well distributed on the unit circle. This is not the case with LTMADS because half the directions correspond to either $\pm e_1$ or $\pm e_2$. For $n = 3$, the two plots on the bottom represent the standard angles of the normalized directions in spherical coordinates. There again it can be seen that ORTHOMADS directions have a better distribution than those of LTMADS, since at least two thirds of the LTMADS directions possess some null coordinates. On the subfigure using LTMADS with $n = 3$, the horizontal bar at $\Phi = \pi/2$ corresponds to the set of directions where $z = 0$. The vertical bars at $\theta = \pm\pi/2$ correspond to directions with $x = 0$, and the one at $\theta = 0$ and $\theta = \pi$ correspond to directions with $y = 0$.

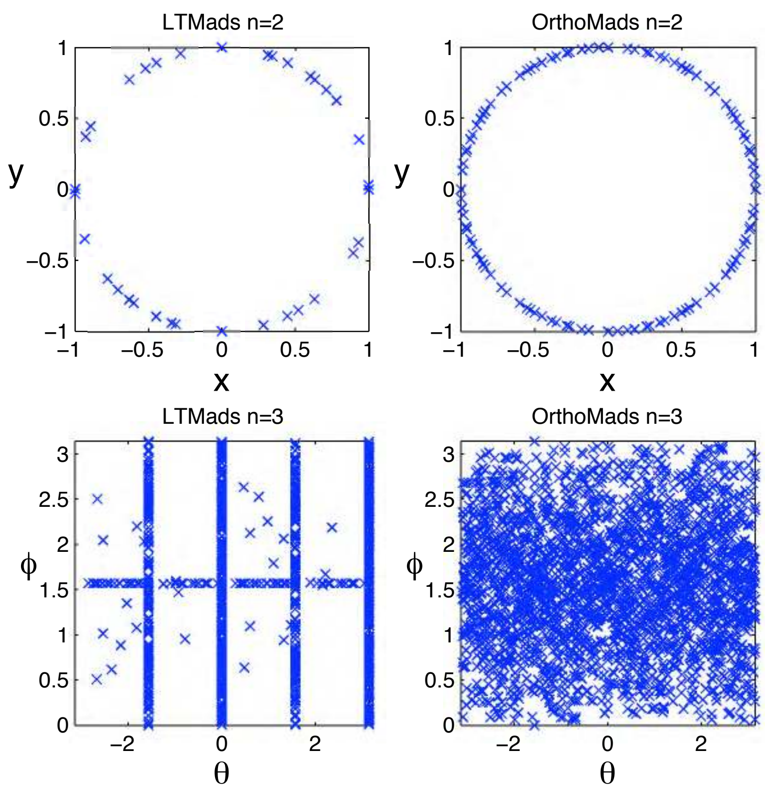

Figure 3: LTMADS and ORTHOMADS normalized POLL directions on the Rosenbrock function with $n = 2$ and $n = 3$.

4 Discussion

This paper introduced ORTHOMADS, an alternative instantiation of the MADS class of algorithms. The advantages of ORTHOMADS over the original LTMADS are that the MADS directions are chosen deterministically, and that those directions are orthogonal to each other. Moreover, ORTHOMADS inherits all of the MADS convergence properties, without probabilistic arguments, and without additional parameters.

Intensive tests on 45 problems from the literature showed that both MADS instances outperform the GPS algorithm, and that ORTHOMADS is at least as competitive as LTMADS, with a better distribution of the POLL directions.

Acknowledgements

We would like to thank Ana Custódio and Luis Vicente for helpful remarks about missed cones of directions in single iterations of LTMADS. These remarks and similar observations of our own helped motivate us to discover ORTHOMADS. We also wish to thank Andrew Booker for suggesting quasi-Monte Carlo methods which started us in directions that led to our use of Halton sequences.

References

[1] C. Audet, V. Béchard, and S. Le Digabel. Nonsmooth optimization through Mesh Adaptive Direct Search and Variable Neighborhood Search. To appear in *Journal of Global Optimization*, DOI: 10.1007/s10898-007-9234-1, 2007.

[2] C. Audet, A. L. Custódio, and J. E. Dennis, Jr. Erratum : Mesh Adaptive Direct Search Algorithms for Constrained Optimization. SIAM *Journal on Optimization*, 18(4):1501–1503, 2008.

[3] C. Audet and J. E. Dennis, Jr. A Pattern Search Filter Method for Nonlinear Programming without Derivatives. SIAM *Journal on Optimization*, 14(4):980–1010, 2004.

[4] C. Audet and J. E. Dennis, Jr. Mesh Adaptive Direct Search Algorithms for Constrained Optimization. SIAM *Journal on Optimization*, 17(1):188–217, 2006.

[5] C. Audet and J. E. Dennis, Jr. A MADS Algorithm with a Progressive Barrier for Derivative-Free Nonlinear Programming. Technical Report G-2007-37, Les Cahiers du GERAD, May 2007.

[6] C. Audet, J. E. Dennis, Jr., and S. Le Digabel. Parallel Space Decomposition of the Mesh Adaptive Direct Search Algorithm. Technical Report G-2007-81, Les Cahiers du GERAD, November 2007.

[7] I. D. Coope and C. J. Price. Frame-Based Methods for Unconstrained Optimization. *Journal of Optimization Theory and Applications*, 107(2):261–274, 2000.

[8] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms, Second Edition*, pages 873–876. MIT Press and McGraw-Hill, 2001.

[9] A. L. Custódio, J. E. Dennis, Jr., and L. N. Vicente. Using Simplex Gradients of Nonsmooth Functions in Direct Search Methods. To appear in IMA *Journal of Numerical Analysis*, 2007.

[10] A. L. Custódio and L. N. Vicente. Using Sampling and Simplex Derivatives in Pattern Search Methods. SIAM *Journal on Optimization*, 18:537–555, 2007.

[11] J. H. Halton. On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals. *Numerische Mathematik*, 2(1):84–90, 1960.

[12] A. S. Householder. Unitary Triangularization of a Nonsymmetric Matrix. *Journal of the Association for Computing Machinery (*ACM*)*, 5(4):339–342, 1958.

[13] R. M. Lewis and V. Torczon. Pattern Search Methods for Linearly Constrained Minimization. SIAM *Journal on Optimization*, 10(3):917–941, 2000.

[14] L. Lukšan and J. Vlček. Test Problems for Nonsmooth Unconstrained and Linearly Constrained Optimization. Technical Report V-798, ICS AS CR, February 2000.

[15] J. Matoušek. On the L_2-Discrepancy for Anchored Boxes. *Journal of Complexity*, 14(4):527–556, 1998.

[16] D. Orban N. I. M. Gould and Ph. L. Toint. CUTEr (and SIFDEC): a Constrained and Unconstrained Testing Environment, revisited. ACM *Transactions on Mathematical Software*, 29(4):373–394, 2003.

[17] H. H. Rosenbrock. An Automatic Method for Finding the Greatest or Least Value of a Function. *The Computer Journal*, 3(3):175–184, October 1960.

[18] V. Torczon. On the Convergence of Pattern Search Algorithms. SIAM *Journal on Optimization*, 7(1):1–25, February 1997.