

Spherical Mesh Adaptive Direct Search for Separating Quasi-uncorrelated Sources by Range-based Independent Component Analysis*

S. Easter Selvan[†]Pierre B. Borckmans[†]A. Chattopadhyay[‡]P.-A. Absil[†]

March 6, 2013

Preprint accepted for publication in Neural Computation

Abstract

It is seemingly paradoxical to the classical definition of the independent component analysis (ICA), that in reality the true sources are often not strictly uncorrelated. With this in mind, this paper concerns a framework to extract quasi-uncorrelated sources with finite supports by optimizing a range-based contrast function under unit-norm constraints (to handle the inherent scaling indeterminacy of ICA) but without orthogonality constraints. Albeit the appealing contrast properties of the range-based function, e.g., the absence of mixing local optima, the function is not differentiable everywhere. Unfortunately, there is a dearth of literature on derivative-free optimizers that effectively handle such a nonsmooth yet promising contrast function. This is the compelling reason for the design of a nonsmooth optimization algorithm on a manifold of matrices having unit-norm columns with the following objectives: (i) to ascertain convergence to a Clarke stationary point of the contrast function; (ii) to adhere to the necessary unit-norm constraints more naturally. The proposed nonsmooth optimization algorithm crucially relies on the design and analysis of an extension of the Mesh Adaptive Direct Search (MADS) method to handle locally Lipschitz objective functions defined on the sphere. The applicability of the algorithm in the ICA domain is demonstrated with simulations involving natural, face, aerial and texture images.

*This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme initiated by the Belgian Science Policy Office. This work was financially supported by the Belgian FRFC (Fonds de la Recherche Fondamentale Collective).

[†]S. E. Selvan, P. B. Borckmans, and P.-A. Absil are with the Department of Mathematical Engineering, ICTEAM Institute, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium (easter.suviseshamuthu@uclouvain.be; pierre.borckmans@uclouvain.be; <http://sites.uclouvain.be/absil/>).

[‡]A. Chattopadhyay is with the School of Computing, University of Leeds, Leeds, UK (a.chattopadhyay@leeds.ac.uk).

Keywords: Clarke stationary point, matrix manifold, nonsmooth optimizer, range-based contrast, source separation

1 Introduction

Independent component analysis (ICA) is a blind source separation technique that attempts to linearly recombine given signals into maximally statistically independent components. The task is usually formulated as minimizing¹ an objective function, termed *contrast function*, that quantifies the “level of dependence” of a collection of random variables. In practical applications, one is often given an $n \times T$ matrix \mathbf{M} containing T samples of n real-valued signals, and the ICA task is then to find an unmixing matrix \mathbf{X} such that the matrix $\mathbf{X}^T \mathbf{M}$ looks as much as possible, in the sense of a finite-sample contrast function, like T observations of n statistically independent random variables. ICA is a very active research topic—see, e.g., the surveys (Hyvärinen, 1999b, 2011)—with many real-world applications, such as biomedical signal processing, remote sensing, seismic signal analysis, denoising in electric and magnetic circuits, and medical image analysis.

A range-based contrast function for the simultaneous extraction of bounded sources was first introduced in (Pham, 2000) and further investigated in (Lee et al., 2006a,b; Vrins, 2007). A finite-sample estimate of this contrast, with low sensitivity to noise and outliers, can be obtained (see Section 6 for details) using the finite-sample estimator of the range using order statistics, as proposed in (Vrins et al., 2007, §V). We adopt this contrast in this work due to the following reasons. (i) It is endowed with the discriminatory property, i.e., it is devoid of mixing local optima (Pham and Vrins, 2006). This means that any local minimum of the contrast function corresponds to an unmixing matrix estimate. (ii) It is suitable for estimating the sources involving signals/images which are in general bounded. (iii) It has been proved to be suitable for large-scale, ill-conditioned, and noisy mixtures (Lee et al., 2006a). (iv) The finite-sample range-based contrast function is less expensive to evaluate than classical contrast functions such as the Shannon-entropy-based mutual information (MI).

However, minimizing this range-based contrast function presents two difficulties. First, in connection with the scaling indeterminacy of the ICA task, it is common to impose a *unit-norm constraint* on the columns of the candidate unmixing matrix \mathbf{X} ; see, e.g., (Vrins et al., 2007, §II) and (Absil and Gallivan, 2006; Lee et al., 2006b). Expected benefits of the unit-norm constraint are to simplify the expression of the contrast function and to avoid continuums of minimizers that may complicate convergence analyses, reduce the convergence speed, or even destroy local convergence properties. The difficulty, though, is that constrained optimization is in general more challenging than unconstrained optimization. Nevertheless, the unit-norm constraint yields a feasible set that is merely a Cartesian product of (hyper)spheres. This feasible set has a natural structure of a Riemannian manifold, and the optimization problem can thus be tackled by Rie-

¹As in (Pham and Vrins, 2006) and in keeping with the recent optimization literature, we adopt the convention that the contrast function is to be minimized. The literature that adopts the opposite convention is encompassed by taking the opposite of the contrast function considered.

mannian geometric techniques, exploited, e.g., in (Shen et al., 2009; Absil and Gallivan, 2006; Shen and Hüper, 2009; Shen and Kleinstenber, 2010; Selvan et al., 2012a). The present paper connects with the Riemannian approach in the sense that the proposed algorithm can be viewed as evolving on the Riemannian feasible set. The second difficulty is that the range-based contrast function is *nonsmooth*, as it involves order statistics. Consequently, smooth Riemannian optimization techniques, as described in (Absil et al., 2008), are not suitable for this task.

Optimization algorithms for minimizing range-based contrast functions were proposed in (Vrins, 2007) and in related papers, both in the orthogonal and nonorthogonal settings. In the orthogonal setting, orthonormality of the unmixing matrix \mathbf{X} is enforced after prewhitening (Hyvärinen et al., 2001). This produces unmixed signals that are exactly uncorrelated. A method for orthogonal range-based ICA is proposed in (Vrins et al., 2007, §VI), where the independent components (ICs) are estimated sequentially (deflation approach) by a direct-search-type method in which the search frame is built by means of Jacobi-like rotations. A similar method, referred to as *support-width ICA* (SWICA), is proposed in (Vrins, 2007, §4.3). The nonorthogonal setting, on the other hand, is motivated by the fact that the sources in real-world applications may not be exactly uncorrelated (and thus not exactly independent). It is even unreasonable to expect that the *finite-sample* estimation of the correlation between the sources be exactly zero. While prewhitening remains a worthy preprocessing for the mixtures \mathbf{M} , relaxing the orthogonality constraint offers leeway to find an unmixing matrix \mathbf{X} where the contrast function takes a lower value, with the hope that such an \mathbf{X} will also yield a better recovery of the sources. This explains why nonorthogonal ICA methods have gained popularity in the recent years (Douglas et al., 2000; Yeredor, 2002; Pham, 2001; van der Veen, 2001; Boscolo et al., 2004). In this direction, the nonorthogonal support-width ICA (NOSWICA) reported in (Lee et al., 2006b) relaxes the orthogonality constraint during the estimation process. To get around the difficulty of repeatedly estimating the same source in the deflation method, the first source is estimated by minimizing the support-width, and the rest of the sources are extracted by minimizing the penalized support. Since there is an inherent drawback of accumulating the IC estimation error in the deflation method, an algorithm for simultaneous range-based ICA was proposed in (Selvan et al., 2012b), in the form of a population-based nonsmooth optimizer, namely, a cross-entropy method with von Mises-Fisher (vMF) distribution. The attractive feature here is the ability to intrinsically maintain the unit-norm constraint of randomly drawn candidate solutions from the vMF distribution.

The aforementioned optimization methods for range-based ICA are not supported by convergence analyses. In fact, as acknowledged in (Vrins, 2007, Remark 22), the first goal of Vrins’s influential thesis was not to develop optimization schemes but rather to analyze the theoretical behavior of entropic contrast. Moreover, the cross-entropy method of (Selvan et al., 2012b) lacks convergence guarantees and suffers from the computational burden of such heuristics.

In this paper, we contribute to improving the optimization aspect of range-based ICA algorithms by proposing and analyzing a new method for nonsmooth optimization with

unit-norm constraint. The new method, when applied to the minimization of the range-based contrast function on the Cartesian product of spheres, results in a nonorthogonal ICA algorithm that is shown to strongly outperform other ICA algorithms in an image separation task.

The new optimization method, called SMADS, is an extension to the sphere of the mesh adaptive direct search (MADS). The MADS algorithm is an iterative derivative-free constrained optimizer proposed by Audet *et al.* (Audet and Dennis, 2006) to handle nonsmooth objective functions, where constraints are imposed using the extreme barrier approach—the objective function values of infeasible points are set to $+\infty$. It is an improvement over the generalized pattern search (GPS) class of methods (Torczon, 1997) that use only a finite number of *poll directions* for exploring the space of variables locally. The key advantage of the MADS algorithm over other frame-based methods, e.g., (Coope and Price, 2000), is that it produces an asymptotically dense set of polling directions, meaning that convergence to a local optimum is guaranteed under mild assumptions. In other words, the MADS algorithm produces a Clarke stationary point. Aside from this merit, it is also easy to implement.

The proposed SMADS algorithm takes the form of a general-purpose derivative-free feasible method for unconstrained optimization on the sphere. By “feasible”, we mean that all the iterates produced by SMADS belong to the sphere. By “unconstrained”, we mean that, for simplicity of the exposition, the feasible set is assumed to be the whole sphere; however, since the workings of SMADS are only slightly more complex than those of MADS, it would not be difficult to handle further constraints akin to those considered in (Audet and Dennis, 2006). (Observe that the original MADS is not suited for the sphere constraint, nor for manifold constraints in general. In particular, the hypertangent cone is empty everywhere and convergence results such as (Audet and Dennis, 2006, Th. 3.12) become vacuous.) The SMADS algorithm is built so as to retain all the convergence properties of the original MADS while remaining computationally inexpensive and easy to implement. The additional computational cost with respect to the original MADS is due to occasional scaling and rounding operations that occur during the face transition mechanism (described in Section 2); it remains comparatively negligible unless the objective function is extremely cheap to evaluate.

This paper is organized as follows. Section 2 introduces notation that enables a concise description and analysis of the proposed algorithm. After recalling the fundamentals of the original MADS algorithm, Section 3 defines and discusses SMADS, the proposed mesh adaptive direct search algorithm on the sphere. An in-depth convergence analysis of SMADS is then proposed in Section 4, where the common principles and differences with the original MADS algorithm are highlighted. In Section 5, a practical instance of the SMADS algorithm is provided, extending the LTMADS algorithm to the sphere setting. The adaptation of the SMADS algorithm for the purpose of minimizing the range-based ICA contrast function is then detailed in Section 6. Numerical experiments are conducted in Section 7, where the task of unmixing the image data is considered. Finally, conclusions and perspectives are given in Section 8.

2 Notation and Definitions

This section introduces notation, depicted in Figure 1, that enables a more concise description and analysis of the SMADS algorithm (Algorithm 1). This section can be skipped and referred to when needed.

Let $\mathbb{R}_0^n := \mathbb{R}^n \setminus \{\mathbf{0}\}$ denote the set of n -tuples of real numbers with the zero n -tuple excerpted, and let \mathbb{R}_+ denote the set of strictly positive real numbers. The unit (hyper)sphere in \mathbb{R}^n is denoted by

$$\mathbb{S}^{n-1} := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1\}, \quad (1)$$

where $\|\cdot\|_2$ denotes the Euclidean norm, and the axis-aligned circumscribed (hyper)cube is denoted by

$$\mathcal{C}^{n-1} := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty = 1\}, \quad (2)$$

where $\|\cdot\|_\infty$ denotes the maximum norm. Let $s > 0$ with $\frac{1}{s}$ integer. The *virtual mesh* of size s on the hypercube \mathcal{C}^{n-1} is the set

$$\mathcal{M}_s^{n-1} := \{\mathbf{x} \in \mathcal{C}^{n-1} : \frac{\mathbf{x}}{s} \in \mathbb{Z}^n\}. \quad (3)$$

We define the scaling map onto the sphere,

$$S : \mathbb{R}_0^n \rightarrow \mathbb{S}^{n-1} : \mathbf{x} \mapsto \frac{\mathbf{x}}{\|\mathbf{x}\|_2}, \quad (4)$$

the scaling map onto the hypercube,

$$C : \mathbb{R}_0^n \rightarrow \mathcal{C}^{n-1} : \mathbf{x} \mapsto \frac{\mathbf{x}}{\|\mathbf{x}\|_\infty}, \quad (5)$$

and the latter followed by rounding onto the virtual mesh of size s ,

$$M_s : \mathbb{R}_0^n \rightarrow \mathcal{M}_s^{n-1} : \mathbf{x} \mapsto s \lfloor \frac{1}{s} C(\mathbf{x}) \rfloor, \quad (6)$$

where $\lfloor \cdot \rfloor$ represents rounding off the argument to the nearest integer value.

Let $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ denote the canonical basis of \mathbb{R}^n . For $i = -1, \dots, -n$, define $\mathbf{e}_i := -\mathbf{e}_{|i|}$. For $i = \pm 1, \dots, \pm n$ and $\mathbf{x} \in \mathbb{R}^n$, define $x_i := \mathbf{e}_i^\top \mathbf{x}$. For $i = \pm 1, \dots, \pm n$, let $F_{\mathbf{e}_i}$ denote the face of \mathcal{C}^{n-1} that contains \mathbf{e}_i , that is,

$$F_{\mathbf{e}_i} := \{\mathbf{x} \in \mathcal{C}^{n-1} : x_i = 1\}.$$

The affine hull of a face $F_{\mathbf{e}_i}$ is $\text{aff}(F_{\mathbf{e}_i}) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{e}_i^\top \mathbf{x} = 1\} = \{\mathbf{x} \in \mathbb{R}^n : x_i = 1\}$. This is the hyperplane that contains $F_{\mathbf{e}_i}$. The tangent space of $\text{aff}(F_{\mathbf{e}_i})$ is $\text{T aff}(F_{\mathbf{e}_i}) = \{\mathbf{v} \in \mathbb{R}^n : \mathbf{e}_i^\top \mathbf{v} = 0\}$. It is the hyperplane spanned by $\{\mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n\}$. For $i = \pm 1, \dots, \pm n$ and $\mathbf{x} \in \mathbb{R}^n$ with $x_i \neq 0$, define the scaling of \mathbf{x} onto $\text{aff}(F_{\mathbf{e}_i})$ by

$$(\mathbf{x})_{\uparrow i} := \frac{\mathbf{x}}{x_i}.$$

Observe that $(\mathbf{x})_{\uparrow i}$ is the intersection between $\text{aff}(F_{\mathbf{e}_i})$ and the straight line through $\mathbf{0}$ and \mathbf{x} . Note also that for any point \mathbf{x} on the affine hull $\text{aff}(F_{\mathbf{e}_i})$ of a given face $F_{\mathbf{e}_i}$, the scaling $C(\mathbf{x})$ leaves the point \mathbf{x} unmodified if $\mathbf{x} \in F_{\mathbf{e}_i}$ and brings \mathbf{x} to the nearest face $F_{\mathbf{e}_j}$ otherwise. The scaling C —and hence also the scaling M_s —thus implicitly offers a convenient face transition mechanism that will be used to produce iterates belonging to \mathcal{C}^{n-1} at all times.

For $i = \pm 1, \dots, \pm n$ and $\mathbf{d} \in \mathbb{R}^{n-1}$, define the embedding $(\mathbf{d})_{\hookrightarrow i}$ to be the vector of \mathbb{R}^n obtained from \mathbf{d} by inserting a zero in the i th position. In other words, $(\mathbf{d})_{\hookrightarrow i} \in \mathbb{R}^n$ with

$$\mathbf{e}_j^\top (\mathbf{d})_{\hookrightarrow i} = \begin{cases} d_j & \text{for } j = 1, \dots, i-1 \\ 0 & \text{for } j = i \\ d_{j-1} & \text{for } j = i+1, \dots, n. \end{cases}$$

Given a sequence $\{\alpha_k\}_{k \geq 0} \in \mathbb{R}$, we let the smallest α encountered up to element k be denoted by

$$\underline{\alpha}_k := \min_{k' \leq k} \alpha_{k'}. \quad (7)$$

3 SMADS Algorithm on the Sphere

The purpose of this section is to define and discuss SMADS, the proposed mesh adaptive direct search algorithm on the sphere.

We first recall the general ideas behind the original MADS algorithm. Initially presented in (Audet and Dennis, 2006), MADS is an iterative derivative-free constrained optimization algorithm generating a finite number of trial points at each iteration k , using a set of directions around the current iterate \mathbf{x}_k . The current objective value $f(\mathbf{x}_k)$ is compared to the objective function at the feasible trial points and is updated when improvement is achieved (without any sufficient decrease requirement). The trial points must belong to a specific *mesh*, constructed using a set of directions in \mathbb{R}^n scaled by the *mesh size* parameter. Note that in order to ensure the MADS convergence properties, these directions must respect some conditions that are detailed in (Audet and Dennis, 2006), and will be discussed later in the convergence analysis of SMADS (Section 4).

Each iteration of the MADS algorithm is composed of two steps, namely, the *search step* and the *poll step*. The search step is optional and allows the exploration of the search space. Any heuristic can be used to sample candidates, as long as they lie on the mesh. This step is not involved in the convergence analysis and is typically used to enhance global exploration of the search space. The poll step on the other hand is mandatory, and is used whenever the search step is not used or fails to produce an improved iterate. The poll step is the one that leads to the convergence results of the MADS algorithm. To do so, it only produces candidate points that lie on a subset of the mesh, called the *poll frame*, to which corresponds a parameter called the *poll size*. The mesh size parameter is then updated at the end of each iteration and the poll size is updated accordingly.

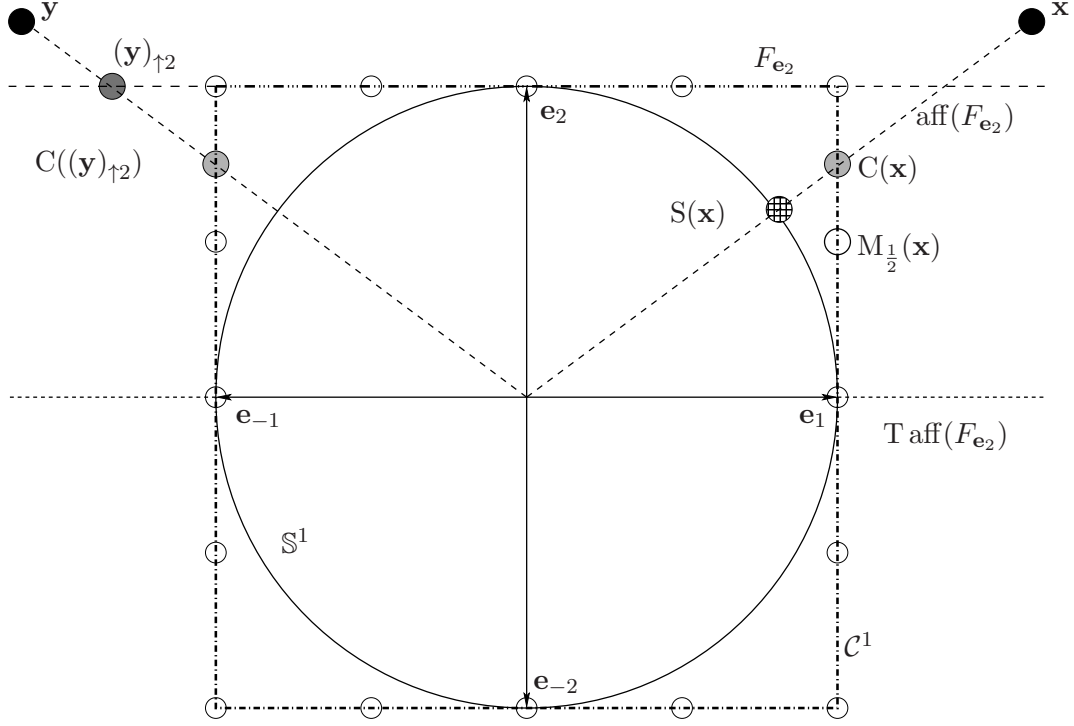


Figure 1: A pictorial representation of the introduced notation for a 2-dimensional setting, which illustrates the unit sphere \mathbb{S}^1 , the hypercube \mathcal{C}^1 (alternately long- and short-dashed lines), the virtual mesh $\mathcal{M}_{\frac{1}{2}}^1$ of size $\frac{1}{2}$ (small hollow dots), one of the four faces ($F_{\mathbf{e}_2}$ in alternately long- and triple short-dashed line), its affine hull $\text{aff}(F_{\mathbf{e}_2})$ (medium-dashed lines) and the corresponding tangent space $\text{T aff}(F_{\mathbf{e}_2})$ (short-dashed lines). A point $\mathbf{x} \in \mathbb{R}_0^2$ is depicted alongside its scaling $S(\mathbf{x})$ onto the sphere (checked dot) and its scaling $C(\mathbf{x})$ onto the hypercube (dull gray dot on the right). The scaling-and-rounding map $M_{\frac{1}{2}}^1$ of \mathbf{x} to the virtual mesh is also shown (hollow dot). Another point $\mathbf{y} \in \mathbb{R}_0^2$ is depicted together with its scaling $(\mathbf{y})_{\uparrow 2}$ (dark gray dot) onto the affine hull of the face $F_{\mathbf{e}_2}$. A face switching mechanism is implicitly present in the mapping C as typified by the point $C((\mathbf{y})_{\uparrow 2})$ (dull gray dot on the left).

Algorithm 1 : Spherical MADS (SMADS)

- 1: Refer to Section 2 for notation
 - 2: **Input**: real-valued function f on the sphere \mathbb{S}^{n-1} , extended to \mathbb{R}_0^n according to (8); convergence threshold $\varepsilon \geq 0$ (where $\varepsilon = 0$ generates an infinite sequence); integer $\tau > 1$;
 - 3: **Initialization**: $k := 0$; initial iterate $\mathbf{x}_0 := \pm \mathbf{e}_j$ for some $j \in \{1, \dots, n\}$; mesh size parameter $\alpha_0^m := \tau^{\beta_0}$ for some integer $\beta_0 \leq 0$; poll size parameter α_0^p as in (11);
 - 4: **while** $\alpha_k^p > \varepsilon$ **do**
 - 5: SEARCH STEP (*optional*)
 - 6: Generate $N \geq 0$ points $\mathbf{x}_k^1, \dots, \mathbf{x}_k^N$ on the mesh $\mathcal{M}_{\alpha_k^m}^{n-1}$ (see (3) and (7));
 - 7: **if** $N > 0$ **and** there is $i^* \in \{1, \dots, N\}$ such that $f(\mathbf{x}_k^{i^*}) < f(\mathbf{x}_k)$ **then**
 - 8: $\mathbf{x}_{k+1} := \mathbf{x}_k^{i^*}$ and declare iteration successful;
 - 9: **else**
 - 10: POLL STEP
 - 11: Construct poll frame P_k according to Definition 1; (A specific way of constructing P_k will be given in Section 5.1.)
 - 12: **if** there is $\mathbf{p}^* \in P_k$ such that $f(\mathbf{p}^*) < f(\mathbf{x}_k)$ **then**
 - 13: $\mathbf{x}_{k+1} := \mathbf{p}^*$ and declare iteration successful;
 - 14: **else**
 - 15: $\mathbf{x}_{k+1} := \mathbf{x}_k$ and declare iteration unsuccessful;
 - 16: **end if**
 - 17: **end if**
 - 18: $k := k + 1$;
 - 19: Update the parameters α_k^m as in (10) and α_k^p as in (11);
 - 20: **end while**
 - 21: **return** $\hat{\mathbf{x}} := \mathbf{S}(\mathbf{x}_k) := \mathbf{x}_k / \|\mathbf{x}_k\|_2$;
-

We now present the SMADS algorithm, our adaptation of MADS to the sphere setting. The algorithm statement is given in Algorithm 1. Several comments and precisions follow.

The problem addressed by SMADS is to minimize a given real-valued function f defined on the (hyper)sphere \mathbb{S}^{n-1} . We use the same notation f to denote its zero-degree homogeneous extension to \mathbb{R}_0^n , i.e.,

$$f(\mathbf{x}) := f(S(\mathbf{x})), \quad (8)$$

where S denotes the scaling to the sphere given by (4).

All the iterates generated by SMADS belong to the circumscribed hypercube \mathcal{C}^n . SMADS can however be viewed as an iteration on the sphere \mathbb{S}^{n-1} because the scaling map S restricted to \mathcal{C}^n is a one-to-one map that leaves f invariant. The situation is depicted in Figure 2. By working on the hypercube instead of the sphere, we benefit from the fact that the search space is locally a linear manifold almost everywhere, but we need a face-transition mechanism (provided by the scaling map M_s (6)). As we will show below, the face-transition mechanism present in SMADS introduces a very mild computational overhead and retains all the convergence properties of the original MADS.

The search step in Algorithm 1 is optional since $N = 0$ is allowed. In view of (3) and (7), the finest virtual mesh (3) at any iteration k is given by $\mathcal{M}_{\underline{\alpha}_k^m}^{n-1} = \mathcal{C}^{n-1} \cap \underline{\alpha}_k^m \mathbb{Z}^n$. As we will see, $1/\underline{\alpha}_k^m$ is an integer at every iteration k , and it follows that $\mathcal{M}_{\underline{\alpha}_k^m}^{n-1}$ consists of the nodes of a fitted grid on the circumscribed hypercube \mathcal{C}^{n-1} . This is illustrated in Figure 1 (for $n = 2$) and Figure 3 (for $n = 3$) for one face.

The poll step is performed whenever the search step does not produce a new iterate. Since the mesh $\mathcal{M}_{\underline{\alpha}_k^m}^{n-1}$ is finite and the search step does not update $\underline{\alpha}_k^m$, it follows that the poll step is performed infinitely often (when the convergence threshold ε is set to 0).

The poll frame P_k is constructed around the current iterate \mathbf{x}_k , according to Definition 1.

DEFINITION 1 (poll frame). Given a positive spanning set $\Xi_k \subset \mathbb{R}^{n-1}$ (i.e., every element of \mathbb{R}^{n-1} can be written as a nonnegative linear combination of the elements of Ξ_k), the *poll frame* of SMADS at the current iterate $\mathbf{x}_k \in F_{\mathbf{e}_i}$ is the set

$$P_k := M_{\underline{\alpha}_k^m} \left(\{ \mathbf{x}_k + \underline{\alpha}_k^m (\boldsymbol{\xi})_{\hookrightarrow i(k)} : \boldsymbol{\xi} \in \Xi_k \} \right), \quad (9)$$

where M (6) is the mapping to the virtual mesh and $i(k)$ is such that $\mathbf{x}_k \in F_{\mathbf{e}_{i(k)}}$. (When \mathbf{x}_k belongs to more than one face, the ambiguity can be removed by setting $i(k)$ to be the smallest i such that $\mathbf{x}_k \in F_{\mathbf{e}_i}$.) Furthermore, each *polling direction* $\boldsymbol{\xi} \in \Xi_k$ must satisfy the conditions below:

- $\boldsymbol{\xi} \in \mathbb{Z}_0^{n-1}$;
- $\underline{\alpha}_k^m \|\boldsymbol{\xi}\| \leq \underline{\alpha}_k^p$.

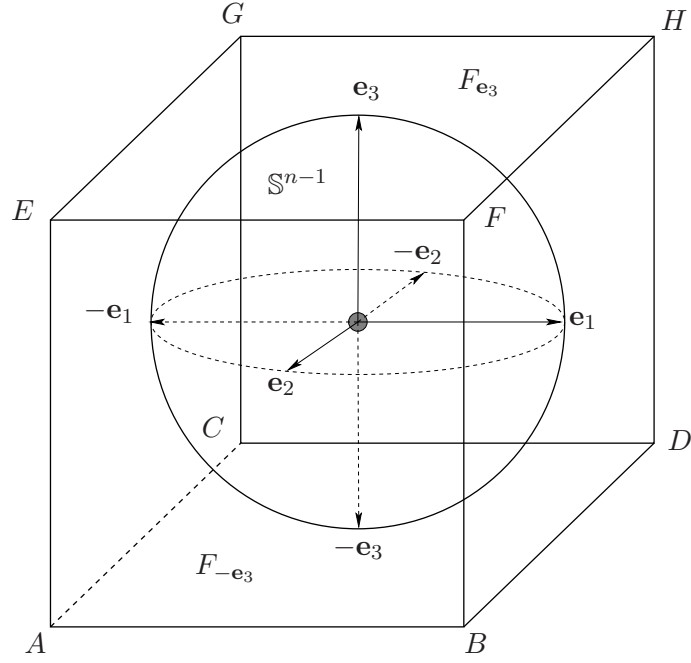


Figure 2: A pictorial representation of the algorithm setting, where the unit sphere \mathbb{S}^{n-1} is enclosed within a hypercube \mathcal{C}^{n-1} . All the iterates lie on the (bounded) faces $F_{\mathbf{e}_i}$, $i = \pm 1, \dots, \pm n$, of \mathcal{C}^{n-1} , e.g., the faces labeled as $ABCD$ and $EFGH$ are $F_{-\mathbf{e}_3}$ and $F_{\mathbf{e}_3}$, respectively; \mathbf{e}_i are the endpoints of the canonical basis vectors (dashed and continuous arrows) pointing outward from the origin of \mathbb{S}^{n-1} .

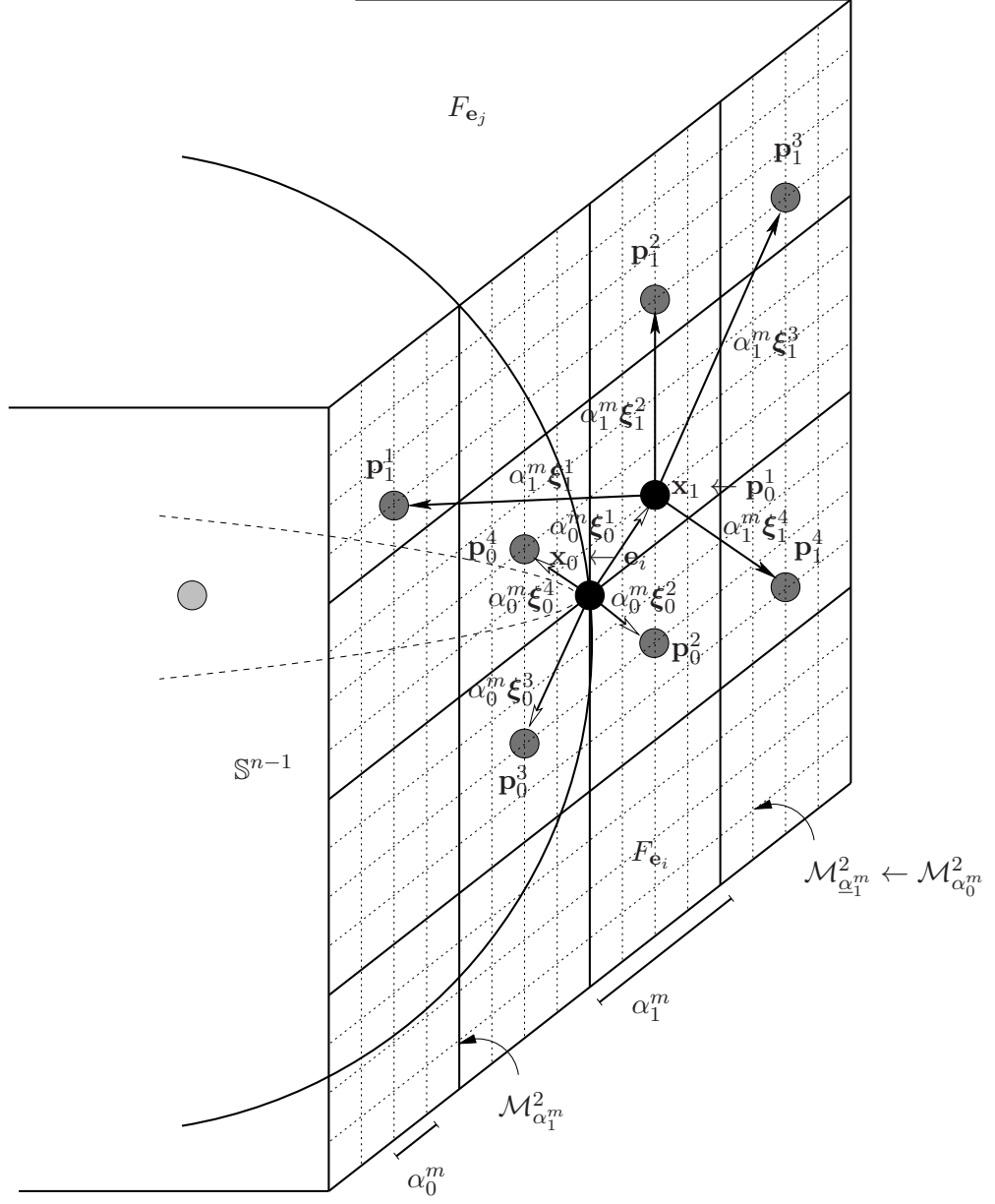


Figure 3: The initial iterate \mathbf{x}_0 has been selected as \mathbf{e}_i . As per (3), the mesh is conceptually built around \mathbf{x}_0 , which is an integer lattice (dashed grid) with the mesh size parameter α_0^m . For the poll directions $\Xi_0 = \{\xi_0^1, \dots, \xi_0^4\}$ (arrows with a hollow head for $k = 0$), the algorithm generates trial points $\mathbf{p}_0^1, \dots, \mathbf{p}_0^4$ (dark gray dots) around \mathbf{x}_0 , which constitute the poll frame denoted by a set P_0 . Note that at $k = 0$, the mesh points located at the intersections of the dashed horizontal and vertical lines form a set $\mathcal{M}_{\alpha_0^m}^2$, such that $P_0 \subset \mathcal{M}_{\alpha_0^m}^2$ as specified in Definition 1. Since $f(\mathbf{p}_0^1) < f(\mathbf{x}_0)$, \mathbf{p}_0^1 is used to update the current iterate \mathbf{x}_0 as \mathbf{x}_1 . Recall that the mesh $\mathcal{M}_{\alpha_1^m}^2$ with the mesh size parameter α_1^m given by (7) is the same as $\mathcal{M}_{\alpha_0^m}^2$, owing to the fact that $\alpha_1^m = \tau^1 \alpha_0^m$ where $\tau > 1$ is a fixed integer. At every iteration k , four trial points are generated as the set of directions Ξ_k (arrows with a solid head for $k = 1$) considered here is the maximal positive basis with $\#\Xi_k = 2(n - 1) = 4$.

Observe also that $\|(\xi)_{\hookrightarrow i(k)}\| = \|\xi\|$. In the original MADS algorithm presented in (Audet and Dennis, 2006), there is more freedom on the nature of the vectors ξ ; for simplicity, we adopt the choice made in the LTMADS method, see (Audet and Dennis, 2006, §4).

(The reader familiar with the MADS algorithm will notice that the last bullet of (Audet and Dennis, 2006, Def. 2.2) has no counterpart in Definition 1. Translated in the SMADS context, the condition requires that limits of the normalized sets $\{\frac{\xi}{\|\xi\|} : \xi \in \Xi_k\}$ are positive spanning sets. This condition is useful when f is smooth and allows to conclude, as in (Coope and Price, 2000, Th. 4.2), that each cluster point is a stationary point of f . However, as discussed in (Audet and Dennis, 2006), when f is nonsmooth, other assumptions are needed to conclude that cluster points are (Clarke) stationary points. The denseness assumption that we will encounter in Theorem 5 makes the positive spanning limit set condition superfluous.)

The mesh size parameter α_k^m is updated at the end of each iteration according to the following rule: given two integers $\beta^- \leq -1$ and $\beta^+ \geq 0$, choose

$$\alpha_k^m = \tau^{\beta_k} \alpha_{k-1}^m \quad (10)$$

for some

$$\beta_k = \begin{cases} \{0, 1, \dots, \beta^+\} & \text{if iteration is successful} \\ \{\beta^-, \beta^- + 1, \dots, -1\} & \text{otherwise,} \end{cases}$$

such that $\alpha_{k+1}^m \leq 1$. (Recall that $\tau > 1$ is a fixed integer.) Hence, for all k , $\alpha_k^m = \tau^{\sum_{i=0}^k \beta_i}$ and $\underline{\alpha}_k^m = \tau^{\min_{0 \leq k' \leq k} \sum_{i=0}^{k'} \beta_i}$, where $\min_{0 \leq k' \leq k} \sum_{i=0}^{k'} \beta_i \leq \sum_{i=0}^k \beta_i \leq 0$ are integers. In particular, as announced, $1/\underline{\alpha}_k^m$ is an integer for all k .

The poll size parameter α_k^p is then determined such that $\alpha_k^m \leq \alpha_k^p \leq 1$ and moreover, for every infinite index set K , $\lim_{k \in K} \alpha_k^m = 0$ if and only if $\lim_{k \in K} \alpha_k^p = 0$. To fix ideas, we set

$$\alpha_k^p = \sqrt{\alpha_k^m} \quad (11)$$

which satisfies both conditions. (We impose that $\alpha_k^p \leq 1$ in order to be able to apply the last point of Lemma 6 in Theorem 7.)

As we will see in Theorem 2, $\lim_{k \rightarrow \infty} \alpha_k^p = 0$, which justifies the stopping criterion.

A crucial aspect of SMADS is that all the iterates and all the poll points belong to the virtual mesh defined in (3), as shown next.

PROPOSITION 1. In SMADS (Algorithm 1), for all k , and recalling definitions (3) and (7), it holds that

$$\mathbf{x}_k \in \mathcal{M}_{\underline{\alpha}_k^m}^{n-1} \quad \text{and} \quad P_k \subset \mathcal{M}_{\underline{\alpha}_k^m}^{n-1}.$$

Proof. The claim is direct for P_k in view of the presence of the scaling-and-rounding map $M_{\underline{\alpha}_k^m}$ in (9). Regarding \mathbf{x}_k , let k' be the smallest integer such that $\mathbf{x}_{k'} = \mathbf{x}_k$. If $k' = 0$, then \mathbf{x}_k is a face center (i.e. $\mathbf{x}_k = \mathbf{e}_i$, $-n \leq i \leq n$) and the claim follows, so we henceforth assume that $k' > 0$. Then $\mathbf{x}_{k'}$ is a new point generated by Algorithm 1 at step $k' - 1$ and, in view of the workings of the algorithm, we have that $\mathbf{x}_{k'} \in \mathcal{M}_{\underline{\alpha}_{k'-1}^m}^{n-1}$. In

view of the mesh size update rule (10), we have that $\underline{\alpha}_{k'-1}^m / \underline{\alpha}_k^m$ is an integer and hence that $\mathcal{M}_{\underline{\alpha}_{k'-1}^m}^{n-1} \subseteq \mathcal{M}_{\underline{\alpha}_k^m}^{n-1}$. The claim follows. \square

An important consequence is that $\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i(k)}$ involved in the definition (9) of P_k is such that $\frac{1}{\underline{\alpha}_k^m}(\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i(k)})$ is a vector of integers for all k . Hence, in the frequently encountered situation where $\alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i(k)}$ is sufficiently small for $\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i(k)}$ to remain on the hypercube \mathcal{C}^{n-1} , we have that the scaling-and-rounding operation $M_{\underline{\alpha}_k^m}$ in P_k (9) has no effect and can thus be omitted. When it cannot be omitted, the scaling-and-rounding operation remains computationally cheap. This justifies the claim made in the introduction that the computational overhead of SMADS with respect to MADS is minor.

4 Convergence Analysis of SMADS

In adapting to SMADS the convergence analysis of MADS (Audet and Dennis, 2006, §3), the main difficulty is to take face transition into account. In other words, we need to address the fact that, whenever $\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i(k)}$ in the definition (9) of P_k is outside the hypercube \mathcal{C}^{n-1} , the action of $M_{\underline{\alpha}_k^m}$ scales it back to \mathcal{C}^{n-1} and rounds it to the nearest point of the virtual mesh $\mathcal{M}_{\underline{\alpha}_k^m}^{n-1}$.

We first show that the convergence analysis of MADS still holds if the notion of direction is replaced by the notion of *adapted direction*, defined next. We conclude by showing that if the set of refining directions is dense, then so is the set of adapted refining directions, under a reasonable condition on the size of the poll frame.

The scaling-and-rounding operation that occurs during face transition prompts us to define *adapted* search directions as follows. The concept is illustrated in Figure 4.

DEFINITION 2 (adapted search direction). Let $\mathbf{x}_k \in F_{\mathbf{e}_i}$ be the current iterate and let $\boldsymbol{\xi} \in \mathbb{R}^{n-1}$ be a search direction, where i is chosen as in Definition 1 if there is an ambiguity. The *adapted search direction* $\hat{\boldsymbol{\xi}}$ is the direction of the variation of \mathbf{x}_k obtained by going to $\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i}$, scaling to the hypercube, rounding onto the virtual mesh, and scaling back to the affine hull of face $F_{\mathbf{e}_i}$; more precisely,

$$\hat{\boldsymbol{\xi}} := \frac{1}{\alpha_k^m} \left((M_{\underline{\alpha}_k^m}(\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i}))_{\uparrow i} - \mathbf{x}_k \right). \quad (12)$$

Observe that if $\boldsymbol{\xi}$ is an integer vector and $\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i}$ remains in $F_{\mathbf{e}_i}$, then $\hat{\boldsymbol{\xi}} = \boldsymbol{\xi}$. Observe also that

$$P_k = C \left(\{ \mathbf{x}_k + \alpha_k^m \hat{\boldsymbol{\xi}} : \boldsymbol{\xi} \in \Xi_k \} \right);$$

hence $\{ \mathbf{x}_k + \alpha_k^m \hat{\boldsymbol{\xi}} : \boldsymbol{\xi} \in \Xi_k \}$ is the frame in $\text{aff}(F_{\mathbf{e}_i})$ that projects to P_k by scaling onto \mathcal{C}^{n-1} .

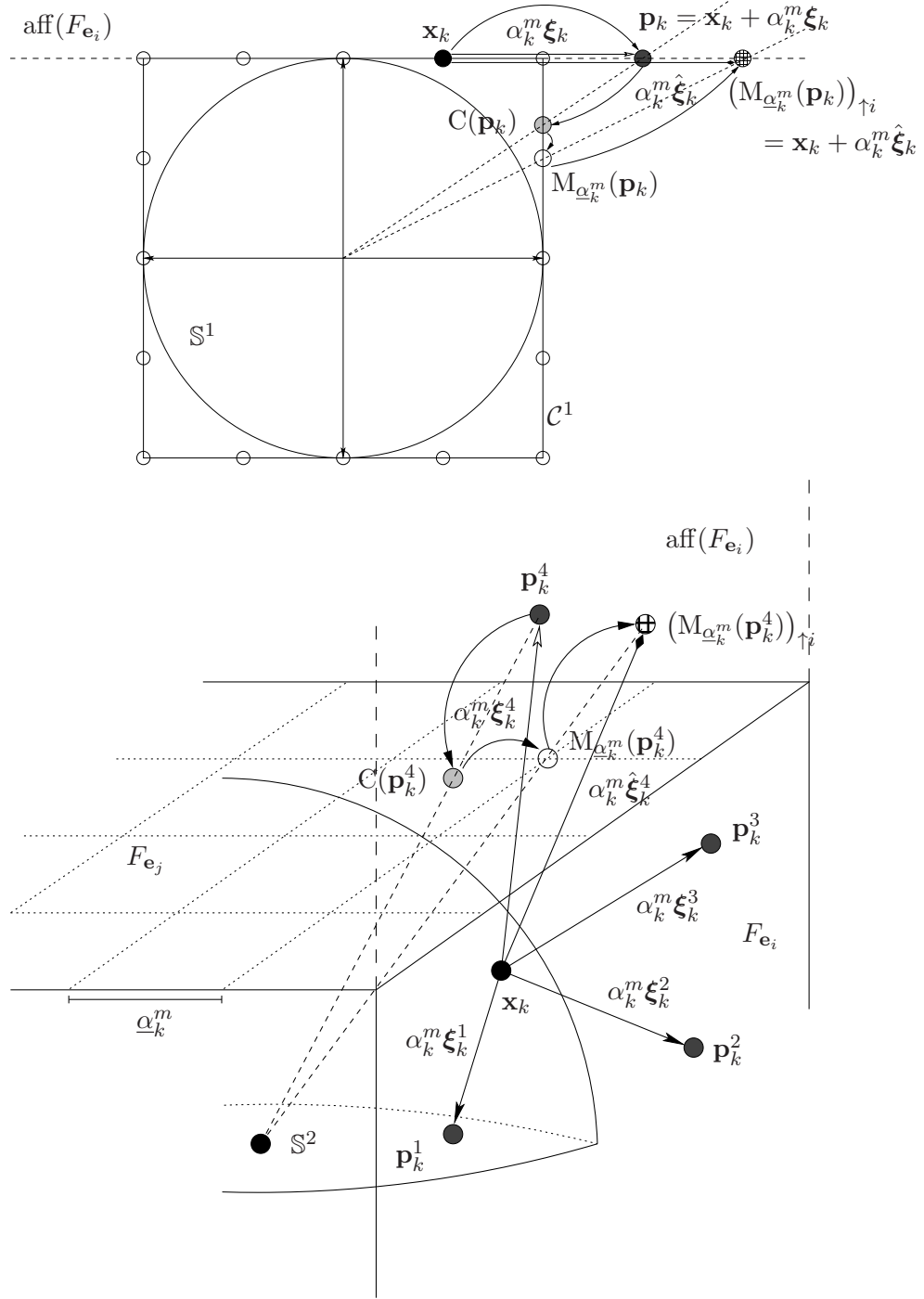


Figure 4: Illustration of the face switching mechanism and the notion of adapted direction (top: 2D; bottom: 3D). At iteration k , a poll direction (arrow with a hollow head—top: ξ_k ; bottom: ξ_k^4) produces the frame point (dark gray dot—top: \mathbf{p}_k ; bottom: \mathbf{p}_k^4) that belongs to $\text{aff}(F_{\mathbf{e}_i})$ but not to $F_{\mathbf{e}_i}$. Consequently, it must be assigned to a mesh point on the face $F_{\mathbf{e}_j}$ of \mathcal{C}^{n-1} . This is carried out in two steps: first, by scaling onto the hypercube, symbolized as C , the frame point is restricted to \mathcal{C}^{n-1} (dull gray dot); next, it is assigned to the nearest mesh point using the operation $M_{\alpha_k^m}$ (hollow dot). By scaling this point back to $\text{aff}(F_{\mathbf{e}_i})$ denoted as $(\cdot)_{\uparrow i}$ (checked dot), the face switching procedure results in an adapted search direction (arrow with a diamond head—top: $\hat{\xi}_k$; bottom: $\hat{\xi}_k^4$).

Until Section 4.1, the analysis of SMADS unfolds much as the analysis of MADS in (Audet and Dennis, 2006, §3), with directions replaced by adapted directions. We provide some details for the reader's convenience.

THEOREM 2. Under the assumption that a locally Lipschitz continuous function $f : \mathbb{S}^{n-1} \rightarrow \mathbb{R}$ and an initial iterate $\mathbf{x}_0 \in F_{\mathbf{e}_i}$ are available for the algorithm,

- (i) all the iterates $\{\mathbf{x}_k\}$ produced by the SMADS algorithm lie in a compact set;
- (ii) the update of the poll and mesh size parameters satisfy

$$\liminf_{k \rightarrow \infty} \alpha_k^p = \liminf_{k \rightarrow \infty} \alpha_k^m = 0.$$

Proof. The first claim is straightforward: all the iterates belong to the hypercube \mathcal{C}^{n-1} , which is a compact set.

The second claim can be shown by means of a contradiction argument. We know that for all $N \geq 0$, the iterates $\mathbf{x}_0, \dots, \mathbf{x}_N$ belong to the mesh $\mathcal{M}_{\underline{\alpha}_N^m}^{n-1}$. Suppose, by way of contradiction, that the second claim does not hold. Then $\underline{\alpha}_\infty^m := \lim_{N \rightarrow \infty} \underline{\alpha}_N^m > 0$, and it follows that all the iterates belong to $\mathcal{M}_{\underline{\alpha}_\infty^m}^{n-1}$. This set contains finitely many points, hence an improved mesh point can be found only finitely many times. After that, the mesh size parameter $\underline{\alpha}_N^m$ is decreased at each iterate according to rule (10), which implies that the mesh size parameter goes to zero, a contradiction. \square

DEFINITION 3 (minimal frame center). If the poll step fails to produce an improved mesh point at iteration k , then the iterate \mathbf{x}_k is said to be a *minimal frame center*.

DEFINITION 4 (refining subsequence). Let $\{\mathbf{x}_k\}$, $\{\alpha_k^p\}$ and $\{i(k)\}$ be generated by Algorithm 1. If K is an infinite index set such that $\{\mathbf{x}_k\}_{k \in K}$ are minimal frame centers, $\{\alpha_k^p\}_{k \in K}$ converges to zero, and $\{i(k)\}_{k \in K}$ is a constant, say j , then $\{\mathbf{x}_k\}_{k \in K}$ is called a *refining subsequence* (on face j).

The next result concerning the existence of at least one convergent refining subsequence can be viewed as an extension of Theorem 3.6 in (Audet and Dennis, 2002) obtained for the GPS algorithms.

THEOREM 3. Let $\{\mathbf{x}_k\}$ be the iterates produced by the SMADS algorithm optimizing a locally Lipschitz $f : \mathbb{S}^{n-1} \rightarrow \mathbb{R}$, then there exists at least one convergent refining subsequence.

Proof. The result follows immediately from Theorem 2 that guarantees the existence of a subset of indices $K' \subset K$ for which $\{\alpha_k^m\}_{k \in K'} \downarrow 0$ with K being the set of indices corresponding to the minimal frame centers. \square

In SMADS, the concept of refining direction (Audet and Dennis, 2006, Def. 3.2) is complemented by the one of *adapted* refining direction.

DEFINITION 5 (refining direction and adapted refining direction). Let $\overset{\circ}{\mathbf{x}}$ be the limit of a convergent refining subsequence. For some subset of indices $L \subseteq K$, if the limit $\lim_{k \in L} \frac{\xi_k}{\|\xi_k\|}$ exists with poll direction $\xi_k \in \Xi_k$, then this limit is said to be a *refining direction* for $\overset{\circ}{\mathbf{x}}$. For some subset of indices $L \subseteq K$, if the limit $\lim_{k \in L} \frac{\hat{\xi}_k}{\|\hat{\xi}_k\|}$ exists with poll direction $\xi_k \in \Xi_k$ (where the hat denotes the adapted direction as in (12)), then this limit is said to be an *adapted refining direction* for $\overset{\circ}{\mathbf{x}}$.

The subsequent theorem is the generalization of (Audet and Dennis, 2006, Th. 3.12) by taking into account adapted refining directions. We first recall a classical definition.

DEFINITION 6 (Clarke generalized directional derivative). The *Clarke generalized directional derivative* of f at $\mathbf{x} \in \mathbb{R}_0^n$ along a direction $\mathbf{v} \in \mathbb{R}^n$ is defined as

$$f^\circ(\mathbf{x}; \mathbf{v}) := \limsup_{\mathbf{y} \rightarrow \mathbf{x}, t \downarrow 0} \frac{f(\mathbf{y} + t\mathbf{v}) - f(\mathbf{y})}{t}.$$

THEOREM 4. Let f (as defined in (8)) be Lipschitz near a limit $\overset{\circ}{\mathbf{x}} \in F_{\mathbf{e}_j}$ of a refining subsequence on face j , and $\hat{\mathbf{v}} \in \text{T aff}(F_{\mathbf{e}_j})$ be an adapted refining direction for $\overset{\circ}{\mathbf{x}}$. Then the generalized directional derivative of f at $\overset{\circ}{\mathbf{x}}$ in the direction $\hat{\mathbf{v}}$ is nonnegative, i.e., $f^\circ(\overset{\circ}{\mathbf{x}}; \hat{\mathbf{v}}) \geq 0$.

Proof. The proof reproduces the proof of (Audet and Dennis, 2006, Th. 3.12), with \mathbb{R}^n replaced by the linear manifold $\text{aff}(F_{\mathbf{e}_j})$ and without constraints.

Let $\{\mathbf{x}_k\}_{k \in K}$ be a refining subsequence that converges to $\overset{\circ}{\mathbf{x}}$. Express $\hat{\mathbf{v}} \in \text{T aff}(F_{\mathbf{e}_j})$ as $\lim_{k \in L} \frac{\hat{\xi}_k}{\|\hat{\xi}_k\|}$ with $\xi_k \in \Xi_k$ for all $k \in L$, following Definition 5. In view of (Audet and Dennis, 2006, Prop. 3.9) (which obviously generalizes from \mathbb{R}^n to linear manifolds), we have that

$$f^\circ(\overset{\circ}{\mathbf{x}}; \hat{\mathbf{v}}) = \lim_{k \in L} f^\circ\left(\overset{\circ}{\mathbf{x}}; \frac{\hat{\xi}_k}{\|\hat{\xi}_k\|}\right). \quad (13)$$

The SMADS algorithm guarantees that $\alpha_k^m \|\hat{\xi}_k\|$ converges to zero due to the following rationale:

- (i) $\alpha_k^m \|\xi_k\|_\infty \leq \alpha_k^p$ (second bullet in the conditions to be satisfied by the poll frame in Definition 1);
- (ii) $\liminf_{k \rightarrow \infty} \alpha_k^p = 0$ (second bullet in Theorem 2).

By the definition of Clarke's generalized directional derivative, we have

$$\begin{aligned} f^\circ(\overset{\circ}{\mathbf{x}}; \hat{\mathbf{v}}) &\geq \limsup_{k \in L} \frac{f(\mathbf{x}_k + \alpha_k^m \|\hat{\xi}_k\| \frac{\hat{\xi}_k}{\|\hat{\xi}_k\|}) - f(\mathbf{x}_k)}{\alpha_k^m \|\hat{\xi}_k\|} \\ &= \limsup_{k \in L} \frac{f(\mathbf{x}_k + \alpha_k^m \hat{\xi}_k) - f(\mathbf{x}_k)}{\alpha_k^m \|\hat{\xi}_k\|} \\ &\geq 0. \end{aligned} \quad (14)$$

The first inequality stems from the fact that $f^\circ(\mathring{\mathbf{x}}; \hat{\mathbf{v}})$ given in Definition 6 uses a more general supremum limit, whereas the one on the right hand side of (14) considers a particular choice. The last inequality holds because \mathbf{x}_k is a minimal frame center. \square

Much as for MADS, we have the following convergence result to Clarke stationary points.

DEFINITION 7 (Clarke stationary point on the sphere). Let f be Lipschitz continuous near $\mathring{\mathbf{x}} \in F_{\mathbf{e}_j}$. Then $\mathring{\mathbf{x}}$ is termed a *Clarke stationary point* of f if $f^\circ(\mathring{\mathbf{x}}; \mathbf{v}) \geq 0$ for every direction $\mathbf{v} \in \mathbb{R}^n$. Since f is homogeneous of degree zero, this is equivalent to $f^\circ(\mathring{\mathbf{x}}; \mathbf{v}) \geq 0$ for every direction $\mathbf{v} \in \text{T aff}(F_{\mathbf{e}_j})$.

THEOREM 5. Let f (as defined in (8)) be Lipschitz near a limit $\mathring{\mathbf{x}} \in F_{\mathbf{e}_j}$ of a refining subsequence on face j . If the set of adapted refining directions for $\mathring{\mathbf{x}}$ is dense in $F_{\mathbf{e}_j}$, then $\mathring{\mathbf{x}}$ is a Clarke stationary point of f .

Proof. The proof follows from Theorem 4 and (Audet and Dennis, 2006, Prop. 3.9). \square

4.1 Denseness of Adapted Refining Directions

We now set out to show that, under some condition, if the set of refining directions is dense, then so is the set of adapted refining directions.

To this end, the next lemma will play a crucial role. It provides bounds on the perturbation incurred due to the scaling-and-rounding operation of the mapping M_s .

LEMMA 6. Let \mathbf{y} belong to $\text{aff}(F_{\mathbf{e}_i})$ for some $i \in \{\pm 1, \dots, \pm n\}$; in other words, $\mathbf{e}_i^\top \mathbf{y} = 1$. Then, for all mesh sizes $s > 0$ with $\frac{1}{s}$ integer, we have

$$\|M_s(\mathbf{y}) - \mathbf{C}(\mathbf{y})\|_\infty \leq \frac{s}{2}. \quad (15)$$

and

$$\|(M_s(\mathbf{y}))_{\uparrow i} - \mathbf{y}\|_\infty \leq s \frac{1}{|(M_s(\mathbf{y}))_i(\mathbf{C}(\mathbf{y}))_i|}. \quad (16)$$

If moreover $s < \frac{1}{\|\mathbf{y}\|_\infty}$, then

$$\|(M_s(\mathbf{y}))_{\uparrow i} - \mathbf{y}\|_\infty \leq 2s\|\mathbf{y}\|_\infty^2. \quad (17)$$

Finally, if moreover $\|\mathbf{y}\|_\infty \leq 2$, then

$$\|(M_s(\mathbf{y}))_{\uparrow i} - \mathbf{y}\|_\infty \leq 8s. \quad (18)$$

Proof. Claim (15) is straightforward from the definitions and from basic rounding error theory.

We show (16). To simplify the notation, let $\mathbf{a} := M_s(\mathbf{y})$ and $\mathbf{b} := \mathbf{C}(\mathbf{y})$, and observe that $(M_s(\mathbf{y}))_{\uparrow i} - \mathbf{y} = (\mathbf{a})_{\uparrow i} - (\mathbf{b})_{\uparrow i}$. Observe also that $|a_i| \leq \|\mathbf{a}\|_\infty \leq 1$ and $|b_i| \leq \|\mathbf{b}\|_\infty \leq 1$; for \mathbf{b} , this is because $\mathbf{b} = \mathbf{y}/\|\mathbf{y}\|_\infty$; for a_i , this is because the mesh fits to the hypercube. We have $(\mathbf{a})_{\uparrow i} - (\mathbf{b})_{\uparrow i} = \frac{\mathbf{a}}{a_i} - \frac{\mathbf{b}}{b_i} = \frac{b_i(\mathbf{a}-\mathbf{b}) - (a_i-b_i)\mathbf{b}}{a_i b_i}$. Thus $\|(\mathbf{a})_{\uparrow i} -$

$(\mathbf{b})_{\uparrow i} \|\infty \leq \frac{|b_i| \|\mathbf{a} - \mathbf{b}\|_\infty + |b_i - a_i| \|\mathbf{b}\|_\infty}{a_i b_i} \leq \frac{\|\mathbf{b}\|_\infty \|\mathbf{a} - \mathbf{b}\|_\infty + \|\mathbf{b} - \mathbf{a}\|_\infty \|\mathbf{b}\|_\infty}{a_i b_i} \leq \frac{2 \|\mathbf{b} - \mathbf{a}\|_\infty}{a_i b_i}$, which is the sought result. To obtain the penultimate inequality, we have used the (already shown) fact that $|b_i| \leq \|\mathbf{b}\|_\infty \leq 1$ along with $|a_i - b_i| \leq \|\mathbf{a} - \mathbf{b}\|_\infty$.

We show (17). We have that $(C(\mathbf{y}))_i = 1/\|\mathbf{y}\|_\infty$, and $(M_s(\mathbf{y}))_i - (C(\mathbf{y}))_i \leq s/2$. Hence $\frac{1}{|(M_s(\mathbf{y}))_i - (C(\mathbf{y}))_i|} \leq \frac{1}{|(C(\mathbf{y}))_i| - \frac{s}{2}} \frac{1}{|(C(\mathbf{y}))_i|} \leq 2\|\mathbf{y}\|_\infty^2$ if $\frac{s}{2} < \frac{1}{2} \frac{1}{\|\mathbf{y}\|_\infty}$, i.e., if $s < \frac{1}{\|\mathbf{y}\|_\infty}$. The claim follows.

Finally, (18) is direct in view of the above. \square

We can now state and prove the announced result.

THEOREM 7. Let $\hat{\mathbf{x}} \in F_{\mathbf{e}_j}$ be a limit of a refining subsequence on face j . Let \mathbf{v} be a refining direction, and let $\{\xi_k\}_{k \in L}$ be a sequence as in Definition 5 such that $\lim_{k \in L} \frac{\xi_k}{\|\xi_k\|} = \mathbf{v}$. Assume that there is a constant c_1 such that, for all $k \in L$, $c_1 \alpha_k^p \leq \alpha_k^m \|\xi_k\|$. Then $\lim_{k \in L} \frac{\hat{\xi}_k}{\|\hat{\xi}_k\|} = \mathbf{v}$, and thus \mathbf{v} is also an adapted refining direction.

Proof. From Lemma 6, one deduces that $\|\alpha_k^m \hat{\xi}_k - \alpha_k^m \xi_k\| \leq 8\alpha_k^m$. Since $c_1 \alpha_k^p \leq \alpha_k^m \|\xi_k\|$, it follows that $\frac{\|\alpha_k^m \hat{\xi}_k - \alpha_k^m \xi_k\|}{\alpha_k^m \|\xi_k\|} \leq \frac{8\alpha_k^m}{c_1 \alpha_k^p} \leq \frac{8\alpha_k^m}{c_1 \alpha_k^p} = \frac{8}{c_1} \sqrt{\alpha_k^m} \rightarrow 0$ as $k \rightarrow \infty$, and thus $\lim_{k \in L} \frac{\hat{\xi}_k}{\|\hat{\xi}_k\|} = \lim_{k \in L} \frac{\xi_k}{\|\xi_k\|}$. \square

5 A Practical SMADS

The original MADS algorithm described in (Audet and Dennis, 2006) is a general framework in the sense that its convergence properties rely on some conditions on the polling directions. In order to practically implement this algorithm, one has to provide a mechanism to produce admissible polling directions. The first proposed concrete instance of MADS is the Lower Triangular MADS (LTMADS) (Audet and Dennis, 2006, §4). We now propose a specific instance of SMADS, inspired from LTMADS. The algorithm, called LTSMADS, is described in Algorithm 2. Precisions on the algorithm statement and a convergence analysis follow.

The selection of \mathbf{x}_{k+1} in line 19 of Algorithm 2 follows the opportunistic strategy. For simplicity, we do not discuss other strategies (exhaustive, random, ...).

5.1 Generating Set of Directions

In the proposed LTSMADS, the directions in Ξ_k respecting the conditions mentioned in Definition 1 are generated as in LTMADS (Audet and Dennis, 2006, §4). The algorithm is based on a random lower triangular matrix, hence its name. A brief description of generating Ξ_k is provided below, and a more elaborate treatment is found in (Audet and Dennis, 2006, §4).

For each iteration k , the lower triangular matrix \mathbf{L} of size $(n-2) \times (n-2)$ is first constructed as a basis in \mathbb{R}^{n-2} with the diagonal elements being either 2^ℓ or -2^ℓ and the lower components being randomly selected from the set $S = \{-2^\ell + 1, -2^\ell + 2, \dots, 2^\ell - 1\}$, where $\ell = -\log_4(\alpha_k^m)$. Next, a specific column $\mathbf{b}(\ell) \in \mathbb{Z}^{n-1}$, relying only on α_k^m , is built

Algorithm 2 : LTSMADS

```
1: Refer to Section 2 for notation
2: Input: see Algorithm 1
3: Initialization: see Algorithm 1
4: while  $\alpha_k^p > \varepsilon$  do
5:   if  $k > 0$  and  $f(\mathbf{x}_k) < f(\mathbf{x}_{k-1})$  then
6:     SEARCH STEP
7:      $\mathbf{s}_k := M_{\alpha_{k-1}^m}(\mathbf{x}_{k-1} + 4\alpha_{k-1}^m(\boldsymbol{\xi})_{\hookrightarrow i});$ 
8:     if  $f(\mathbf{s}_k) < f(\mathbf{x}_k)$  then
9:        $\mathbf{x}_{k+1} := \mathbf{s}_k;$ 
10:    else
11:       $\mathbf{x}_{k+1} := \mathbf{x}_k;$ 
12:    end if
13:  else
14:    POLL STEP
15:    Construct  $P_k := M_{\alpha_k^m}(\{\mathbf{x}_k + \alpha_k^m(\boldsymbol{\xi})_{\hookrightarrow i} : \boldsymbol{\xi} \in \Xi_k\})$ , where  $\Xi_k$  is a positive basis
    generated as detailed in Section 5.1;
16:    if  $f(\mathbf{p}) \geq f(\mathbf{x}_k)$  for all  $\mathbf{p} \in P_k$  then
17:       $\mathbf{x}_{k+1} := \mathbf{x}_k;$ 
18:    else
19:      Set  $\mathbf{x}_{k+1}$  as the first element  $\mathbf{p}$  of  $P_k$  such that  $f(\mathbf{p}) < f(\mathbf{x}_k)$ ;
20:      Set  $\boldsymbol{\xi}$  as the corresponding element of  $\Xi_k$ ;
21:    end if
22:  end if
23:   $k := k + 1;$ 
24:  Update the parameters  $\alpha_k^m$  as in (10) and  $\alpha_k^p$  as in (11);
25: end while
26: return  $\hat{\mathbf{x}} := \mathbf{x}_k;$ 
```

such that $|b_i(\ell)| = 2^\ell$ for a randomly chosen index $\hat{i} \in \{1, \dots, n-1\}$, and the rest of the components of $\mathbf{b}(\ell)$ are set to be random integers in S . Then a zero vector of length $n-2$ is inserted into \mathbf{L} to be the row of index \hat{i} in a new matrix \mathbf{L}' , of which the remaining rows are permuted. The resulting $(n-1) \times (n-2)$ matrix is augmented by the column $\mathbf{b}(\ell)$ to form a basis \mathbf{B} in \mathbb{R}^{n-1} , and the columns of \mathbf{B} are permuted as well to form \mathbf{B}' with evenly distributed zeros. Finally a positive maximal basis denoted as $\Xi_k \in \mathbb{R}^{(n-1) \times 2(n-1)}$ is obtained from \mathbf{B}' by appending to it the negative of each column of \mathbf{B}' . This generation process of the poll directions is exemplified below for $n = 5$ (where $[S]$ denotes a random number in the set S and \mathbf{B}'_i denotes the i th column of the matrix \mathbf{B}'):

$$\mathbf{L} = \begin{pmatrix} \pm 2^\ell & 0 & 0 \\ [S] & \pm 2^\ell & 0 \\ [S] & [S] & \pm 2^\ell \end{pmatrix} \quad \mathbf{b}(\ell) = \begin{pmatrix} [S] \\ 2^\ell \\ [S] \\ [S] \end{pmatrix} \leftarrow \hat{i} \quad \mathbf{L}' = \begin{pmatrix} [S] & \pm 2^\ell & 0 \\ 0 & 0 & 0 \\ \pm 2^\ell & 0 & 0 \\ [S] & [S] & \pm 2^\ell \end{pmatrix} \leftarrow \hat{i}$$

$$\mathbf{B} = \begin{pmatrix} [S] & \pm 2^\ell & 0 & b_1 \\ 0 & 0 & 0 & b_2 \\ \pm 2^\ell & 0 & 0 & b_3 \\ [S] & [S] & \pm 2^\ell & b_4 \end{pmatrix} \quad \mathbf{B}' = \begin{pmatrix} 0 & b_1 & \pm 2^\ell & [S] \\ 0 & b_2 & 0 & 0 \\ 0 & b_3 & 0 & \pm 2^\ell \\ \pm 2^\ell & b_4 & [S] & [S] \end{pmatrix}$$

$$\Xi_k = (\mathbf{B}'_1 \quad \dots \quad \mathbf{B}'_n \quad -\mathbf{B}'_1 \quad \dots \quad -\mathbf{B}'_n).$$

5.2 Convergence analysis of LTSMADS

By *b-refining directions* at the limit point $\hat{\mathbf{x}}$, we mean the refining directions generated from poll directions $\mathbf{b}(\ell)$. Similar to Definition 5, we also consider the set of *adapted b-refining directions* constructed from $\hat{\mathbf{b}}(\ell)$ directions. The next result shows that the denseness hypothesis in Theorem 5 holds with probability one in LTSMADS.

THEOREM 8. Let $\{\mathbf{x}_k\}_{k \in K}$ be a convergent refining subsequence, say on face j , generated by the practical SMADS (Algorithm 2). Then the set of adapted *b-refining directions* associated to K is asymptotically dense in $F_{\mathbf{e}_j}$ with probability one.

Proof. The proof stems from the LTSMADS analysis (Audet and Dennis, 2006, Th. 4.3) and uses Theorem 7 for the adapted aspect. Let $\{\mathbf{x}_k\}_{k \in K}$ be a convergent refining subsequence on face j produced by the SMADS, and $\hat{\mathbf{x}} \in F_{\mathbf{e}_j}$ be the limit of the subsequence. Recall that there is a specific poll direction $\mathbf{b}(\ell)$ among the set of directions Ξ_k generated by the algorithm. The asymptotic denseness theorem (Audet and Dennis, 2006, Th. 4.3) shows that $\mathbf{b}(\ell)/\|\mathbf{b}(\ell)\|_\infty$ becomes arbitrarily close to a randomly picked direction $\mathbf{v} \in \text{T aff}(F_{\mathbf{e}_j})$ having $\|\mathbf{v}\|_\infty = 1$ with a probability

$$P \left[\left\| \frac{\mathbf{b}(\ell)}{\|\mathbf{b}(\ell)\|_\infty} - \mathbf{v} \right\|_\infty < \epsilon \right] \geq \frac{\left(\frac{\epsilon}{4}\right)^{n-2}}{2(n-1)}$$

for any $0 < \epsilon < 1$, provided k is sufficiently large satisfying $\sqrt{\alpha_k^m} = 2^{-\ell} \leq \frac{\epsilon}{2}$. At iterate $\mathbf{x}_k \in F_{\mathbf{e}_j}$, we have an adapted $\mathbf{b}(\ell)$ as defined above; recall that

$$\alpha_k^m \hat{\mathbf{b}}(\ell) := (\mathbf{M}_{\underline{\alpha}_k^m}(\mathbf{x}_k + \alpha_k^m \mathbf{b}(\ell)))_{\uparrow j} - \mathbf{x}_k.$$

The motivation for the final scaling to $\text{aff}(F_{\mathbf{e}_j})$ is to perform the computations in the chart $(\cdot)_{\uparrow j}$; since the topology of a manifold is defined through the charts, it follows that denseness in a chart is equivalent to denseness on the sphere. We now show that $\frac{\hat{\mathbf{b}}(\ell)}{\|\hat{\mathbf{b}}(\ell)\|_\infty}$ (which is in the direction of $\hat{\mathbf{b}}(\ell)$) is a sufficiently small perturbation of $\frac{\mathbf{b}(\ell)}{\|\mathbf{b}(\ell)\|_\infty}$ for the gist of the result to be preserved. We have $\|\frac{\hat{\mathbf{b}}(\ell)}{\|\hat{\mathbf{b}}(\ell)\|_\infty} - \frac{\mathbf{b}(\ell)}{\|\mathbf{b}(\ell)\|_\infty}\|_\infty = \frac{1}{\alpha_k^m \|\mathbf{b}(\ell)\|_\infty} \|\alpha_k^m \hat{\mathbf{b}}(\ell) - \alpha_k^m \mathbf{b}(\ell)\|_\infty = \frac{1}{\alpha_k^m \|\mathbf{b}(\ell)\|_\infty} \|(\mathbf{M}_{\underline{\alpha}_k^m}(\mathbf{x}_k + \alpha_k^m \mathbf{b}(\ell)))_{\uparrow j} - \mathbf{x}_k - \alpha_k^m \mathbf{b}(\ell)\|_\infty \leq \frac{1}{\alpha_k^m \|\mathbf{b}(\ell)\|_\infty} 8\alpha_k^m \leq 2^{3-\ell} \leq 4\epsilon$, where the antepenultimate inequality follows from (18) in Lemma 6. In conclusion, we have shown that

$$P \left[\left\| \frac{\hat{\mathbf{b}}(\ell)}{\|\hat{\mathbf{b}}(\ell)\|_\infty} - \mathbf{v} \right\|_\infty < 5\epsilon \right] \geq \frac{(\frac{\epsilon}{4})^{n-2}}{2(n-1)} \quad \text{for } \sqrt{\alpha_k^m} = 2^{-\ell} \leq \frac{\epsilon}{2}. \quad (19)$$

Since $\mathbf{b}(\ell)$ is generated independently at each iterate and since, in view of Theorem 2, $\sqrt{\alpha_k^m} \leq \frac{\epsilon}{2}$ occurs at infinitely many iterates, it follows that

$$P \left[\left\| \frac{\hat{\mathbf{b}}(\ell)}{\|\hat{\mathbf{b}}(\ell)\|_\infty} - \mathbf{v} \right\|_\infty < 5\epsilon \text{ at at least one iterate} \right] = 1.$$

This shows asymptotic denseness of the $\hat{\mathbf{b}}(\ell)$ directions. \square

6 Adapting SMADS for Minimizing Range-based ICA Contrast

Given a random vector $\mathbf{m} \in \mathbb{R}^n$, the proposed ICA scheme estimates an unmixing matrix $\mathbf{X} = [\mathbf{x}^{(1)} \ \dots \ \mathbf{x}^{(n)}]$ with column vectors $\mathbf{x}^{(j)} \in \mathbb{R}^n$, $j = 1, \dots, n$, subject to the constraint $\|\mathbf{x}^{(j)}\|_2 = 1$ such that the n components of $\mathbf{c} = \mathbf{X}^T \mathbf{m}$ are maximally independent as measured by a contrast function. As in (Pham, 2000, (4.1)), the contrast function based on a range estimation approach using order statistics is expressed as

$$f(\mathbf{X}) := \sum_{j=1}^n \log R(\mathbf{x}^{(j)T} \mathbf{m}) - \log |\det \mathbf{X}|, \quad (20)$$

where $R(Y) := \max(Y) - \min(Y)$ is the range function with Y being a random variable. An estimate of the range of Y is derived in (Vrins et al., 2007) by making use of an ordered finite sequence of observations, $y_{(l)}$, $l = 1, \dots, T$,

$$R(Y) := \frac{1}{h} \sum_{r=1}^h R_r(Y) \quad (21)$$

with $R_r(Y) := y_{(T-r+1)} - y_{(r)}$. Now we can state the expression for the sample contrast function

$$f(\mathbf{X}; \mathbf{M}) := \sum_{j=1}^n \log \left(\frac{1}{h} \sum_{r=1}^h R_r \left(\mathbf{x}^{(j)\top} \mathbf{M} \right) \right) - \log |\det \mathbf{X}|, \quad (22)$$

where \mathbf{X} is a candidate unmixing matrix and $\mathbf{M} \in \mathbb{R}^{n \times T}$ is a matrix of T observations of n variables. One possible guideline to set the default value of h , as recommended in (Vrins et al., 2007), is

$$h(T) = \max \left(1, \left\lceil \Re \left\{ \left(\frac{T-18}{6.5} \right)^{0.65} \right\} - 4.5 \right\rceil \right), \quad (23)$$

where $\bar{\rho}$ denotes the nearest integer to ρ and $\Re\{\cdot\}$ returns the real part of the argument. Furthermore, the range estimate in (21), employing the empirically determined value of h in (23), preserves the desirable discriminant contrast property inherent in the exactly evaluated range contrast function (Vrins et al., 2007, p. 816).

We work under the assumption that the data matrix \mathbf{M} is such that the term $\sum_{j=1}^n \log \left(\frac{1}{h} \sum_{r=1}^h R_r \left(\mathbf{x}^{(j)\top} \mathbf{M} \right) \right)$ in (22) is $> -\infty$ under the constraint $\|\mathbf{x}^{(j)}\|_2 = 1$. Note that $\sum_{r=1}^h R_r \left(\mathbf{x}^{(j)\top} \mathbf{M} \right) = 0$ if and only if $\mathbf{x}^{(j)\top} \mathbf{M} = \alpha \mathbf{1}^\top$, where α is a constant. Therefore, we have that $\sum_{j=1}^n \log \left(\frac{1}{h} \sum_{r=1}^h R_r \left(\mathbf{x}^{(j)\top} \mathbf{M} \right) \right) > -\infty$ as long as $\mathbf{1} \notin \text{Im}(\mathbf{M}^\top)$. We remark that the assumption $\mathbf{1} \notin \text{Im}(\mathbf{M}^\top)$ does not apply to ICA problems where the original sources are believed to be constant signals/images, i.e., their distributions are degenerate. In practice, the situation $\mathbf{1} \in \text{Im}(\mathbf{M}^\top)$ will not usually arise, because the condition $\mathbf{1} \in \text{Im}(\mathbf{M}^\top)$ means that all the T observations in \mathbb{R}^n belong to a same hyper-plane. As soon as $T \geq n$, this situation occurs with zero probability under reasonable assumptions on the observation noise.

The following lemma guarantees that the Lipschitz assumption frequently made in the above convergence analysis of (LT)SMADS holds for the contrast function (22).

LEMMA 9. The range-based ICA contrast function f (22) is Lipschitz near every \mathbf{X} such that $-\infty < f(\mathbf{X}) < \infty$.

Proof. The range-based contrast function in (22) involves the sum of max functions. Since the max function is Lipschitz and the log function is differentiable on $(0, \infty)$, it follows that the term $\sum_{j=1}^n \log \left(\frac{1}{h} \sum_{r=1}^h R_r \left(\mathbf{x}^{(j)\top} \mathbf{M} \right) \right)$ is Lipschitz near \mathbf{x} if and only if $\sum_{r=1}^h R_r \left(\mathbf{x}^{(j)\top} \mathbf{M} \right) \neq 0$, $j = 1, \dots, n$. Besides, the barrier term $-\log |\det \mathbf{X}|$ is Lipschitz near every \mathbf{X} where it is finite. \square

Our objective function is thus the range-based ICA contrast in (22) defined on the set of $n \times n$ matrices whose columns have unit Euclidean norm, denoted here by

$$\mathcal{OB}(n) = \left\{ \mathbf{X} \in \mathbb{R}^{n \times n} : \text{ddiag}(\mathbf{X}^\top \mathbf{X}) = \mathbf{I}_n \right\}, \quad (24)$$

where $\text{ddiag}(\cdot)$ represents the diagonal matrix whose diagonal elements are those of the matrix in the argument, and \mathbf{I}_n is the $n \times n$ identity matrix. Since $\mathcal{OB}(n)$ is merely the Cartesian product of n copies of \mathbb{S}^{n-1} , the generalization of SMADS (Algorithm 1) to $\mathcal{OB}(n)$ is straightforward.

We point out that our definition of \mathcal{OB} in (24) does not exactly coincide with its definition given in (Trendafilov and Lippert, 2002), as it does not impose that its elements have full rank. The distinction is irrelevant for the objective function (22) since the $-\log|\det \mathbf{X}|$ term guarantees that the iterates of any descent iteration stay away from the set of rank-deficient matrices.

What follows is a theorem that justifies the choice of the SMADS algorithm to seek for a local minimum of the contrast function defined in (22).

THEOREM 10. For the range-based ICA contrast function f in (22), the sequence generated by the SMADS algorithm (Algorithm 2) admits at least one convergent refining subsequence, and the limit point of any such subsequence is a Clarke stationary point of f .

Proof. The result follows from compactness of \mathcal{OB} and from Theorem 5. \square

REMARK 1. In our numerical experiments, we have always observed that the *whole* SMADS sequence converges, and it then follows from Theorem 10 that it converges to a Clarke stationary point. Moreover, since SMADS is a descent method, convergence to points that are not local minimizers is unstable under perturbations, and is not expected to occur in practice.

7 Experimental Results

7.1 Simulation with Various Image Datasets

The range-based contrast function assumes that the source support measure is finite, in other words, that the underlying sources are bounded. This assumption holds good in the case of images. Moreover, the image sources are in general not uncorrelated and their density distributions are usually multimodal, which exacerbate the challenges faced by an ICA technique. Furthermore the earlier works using the range-based contrast either recommend the use of image data (Vrins et al., 2007) or validate the algorithm with face images (Vrins, 2007). Therefore we have considered the following image categories for assessing the performance of our approach: (i) two sets of 12 natural images including sceneries, human portraits, man-made structures, animals and birds from the Berkeley segmentation dataset and benchmark (Martin et al., 2001); (ii) nine images under each category of aerial and texture images acquired from the Signal and Image Processing Institute, University of Southern California at “<http://sipi.usc.edu/database>”; (iii) nine face images from the dataset of Informatics and Mathematical Modelling, Technical University of Denmark (Stegmann et al., 2003). All the test images were resized to have 200×200 pixels each and converted into gray-scale images. From each dataset comprising natural, aerial, texture and face images, 25 random combinations of six images were

generated to assess the range-based ICA on \mathcal{OB} in terms of the source separation quality. During each trial, the multidimensional data of size 40000×6 consisting of column-wise concatenated pixel gray-levels was mixed by a mixing matrix built from 36 coefficients drawn from a uniform distribution on the interval $(0, 1)$. It is remarked that the mixing procedure is advocated in the works of (Karhunen et al., 1997, p. 497) concerning the source separation of multidimensional image data. The resulting mixture was whitened, because pre-whitening is deemed as the standard preprocessing procedure (Li and Adalı, 2010) and it is believed to improve the convergence, although strictly speaking, ICA algorithms on \mathcal{OB} obviate the need for pre-whitening. The mixture/pre-whitened data thus obtained was supplied as the input for the below-mentioned ICA algorithms for a relative assessment of the source separation performance:

- fastICA (Hyvärinen, 1999a), JADE (Cardoso and Souloumiac, 1993), and info-max (Makeig et al., 1996), which are widely reported in the ICA literature;
- SWICA (Vrins et al., 2007), where the range-based function in (22) was first introduced;
- NOSWICA (Lee et al., 2006b) meant for the separation of highly correlated sources;
- ICA by entropy bound minimization (ICA-EBM) (Li and Adalı, 2010) that uses an accurate entropy estimator and adopts a line-search optimization procedure;
- ICA by optimizing quadratic measures of independence (QICA) with an exhaustive search strategy (Seth et al., 2011) in the quest of avoiding local optima;
- our proposed LTSMADS (Algorithm 2) on the \mathcal{OB} manifold (OBMADS) minimizing the range-based function in (22).

It is pointed out that the image mixture is generated in an artificial manner to enable us to compare the unmixing matrix/source estimate with the true unmixing matrix/source. Further research is underway, where the algorithm can be potentially employed to separate signal mixtures arising from real scenarios, e.g., unmixing electroencephalogram recordings from scalp electrodes prior to generating brain maps to help diagnose focal epilepsy seizures. In such real scenarios, since the true unmixing matrix/source are unknown, one must resort to indirect quality assessment methods in order to compare the various ICA algorithms; for example, in electroencephalogram experiments that we will report elsewhere, an expert neurologist evaluates the accuracy with which the brain maps generated from the ICA output matches the patient's pathological condition.

The parameter values for OBMADS were set as prescribed in (Audet and Dennis, 2006): $\tau = 4$, $\beta^+ = 1$, $\beta^- = -1$, $\varepsilon = 10^{-10}$, $\alpha_0^m = \alpha_0^p = 1$, and total function evaluations $f_{\text{emax}} = 10^5$. The ICA unmixing matrices yielded by the methods included in the empirical study were used to reconstruct the sources from the mixtures. Since the ambiguity of sign and the indeterminacy of the order of the ICs hold for the ICA model, the source estimates are necessarily rescaled and reordered prior to computing

Table 1: Mean and standard deviation of RMSE values from the investigated ICA schemes with a face, aerial, texture and natural image dataset. This empirical study involves 25 trial runs with the mixture of six randomly selected 200×200 images from each dataset. The face, aerial and texture dataset comprises nine images each, and two natural image datasets contain 12 images each. The values in bold face represent the minimum obtained among the experimented schemes.

ICA algorithm	RMSE mean and std. dev.				
	natural I	natural II	face	aerial	texture
fastICA	0.169 ± 0.054	0.254 ± 0.050	0.439 ± 0.025	0.109 ± 0.026	0.018 ± 0.002
JADE	0.143 ± 0.037	0.230 ± 0.055	0.435 ± 0.031	0.092 ± 0.024	0.017 ± 0.002
infomax	0.158 ± 0.050	0.257 ± 0.052	0.423 ± 0.033	0.108 ± 0.031	0.018 ± 0.002
SWICA	0.138 ± 0.051	0.196 ± 0.045	0.335 ± 0.036	0.095 ± 0.026	0.099 ± 0.042
NOSWICA	0.297 ± 0.131	0.321 ± 0.129	0.356 ± 0.101	0.211 ± 0.127	0.237 ± 0.108
ICA-EBM	0.100 ± 0.024	0.113 ± 0.066	0.410 ± 0.045	0.058 ± 0.016	0.017 ± 0.002
QICA	0.182 ± 0.052	0.192 ± 0.053	0.398 ± 0.043	0.115 ± 0.038	0.091 ± 0.035
OBMADS	0.034 ± 0.008	0.040 ± 0.036	0.086 ± 0.011	0.042 ± 0.013	0.008 ± 0.001

the root-mean-square error (RMSE) values with respect to the original sources. Given the original vectorized images $\mathbf{a}^{(j)} = [a_1^{(j)}, \dots, a_T^{(j)}]$, $j = 1, \dots, n$, and the estimated ones after reordering $\mathbf{c}^{(j)} = \mathbf{x}^{(j)\top} \mathbf{M} = [c_1^{(j)}, \dots, c_T^{(j)}]$, $j = 1, \dots, n$, the RMSE is computed as

$$\text{RMSE} = \sqrt{\frac{\sum_{j=1}^n \sum_{l=1}^T \left(a_l^{(j)} - c_l^{(j)}\right)^2}{\sum_{j=1}^n \sum_{l=1}^T \left(a_l^{(j)}\right)^2}}. \quad (25)$$

The RMSE measure in (25) is always nonnegative, scale invariant, and it circumvents the permutation ambiguity due to the reordering of the source estimates.

The mean and standard deviation of the RMSE values in 25 trial runs of the experimented schemes with a face, aerial, texture dataset containing nine images each, and two datasets comprising 12 natural images each are reported in Table 1. Though the fastICA that maximizes the negentropy of the estimated sources with an approximate Newton method offers computational advantage, the mean RMSE is high. The JADE algorithm aims at minimizing the sum of the squared cross-cumulants of the source estimates; the tensorial algorithm faces limitations in higher dimensions as evidenced by the mean RMSE. The infomax employs a stochastic gradient ascent rule to maximize the entropy of the ICs that are nonlinearly transformed; as expected, the surrogate contrast resulted in an inferior outcome.

The SWICA, NOSWICA and OBMADS minimize the same contrast function in (22). While the SWICA insists on the orthogonality constraint with the Givens rotation ma-

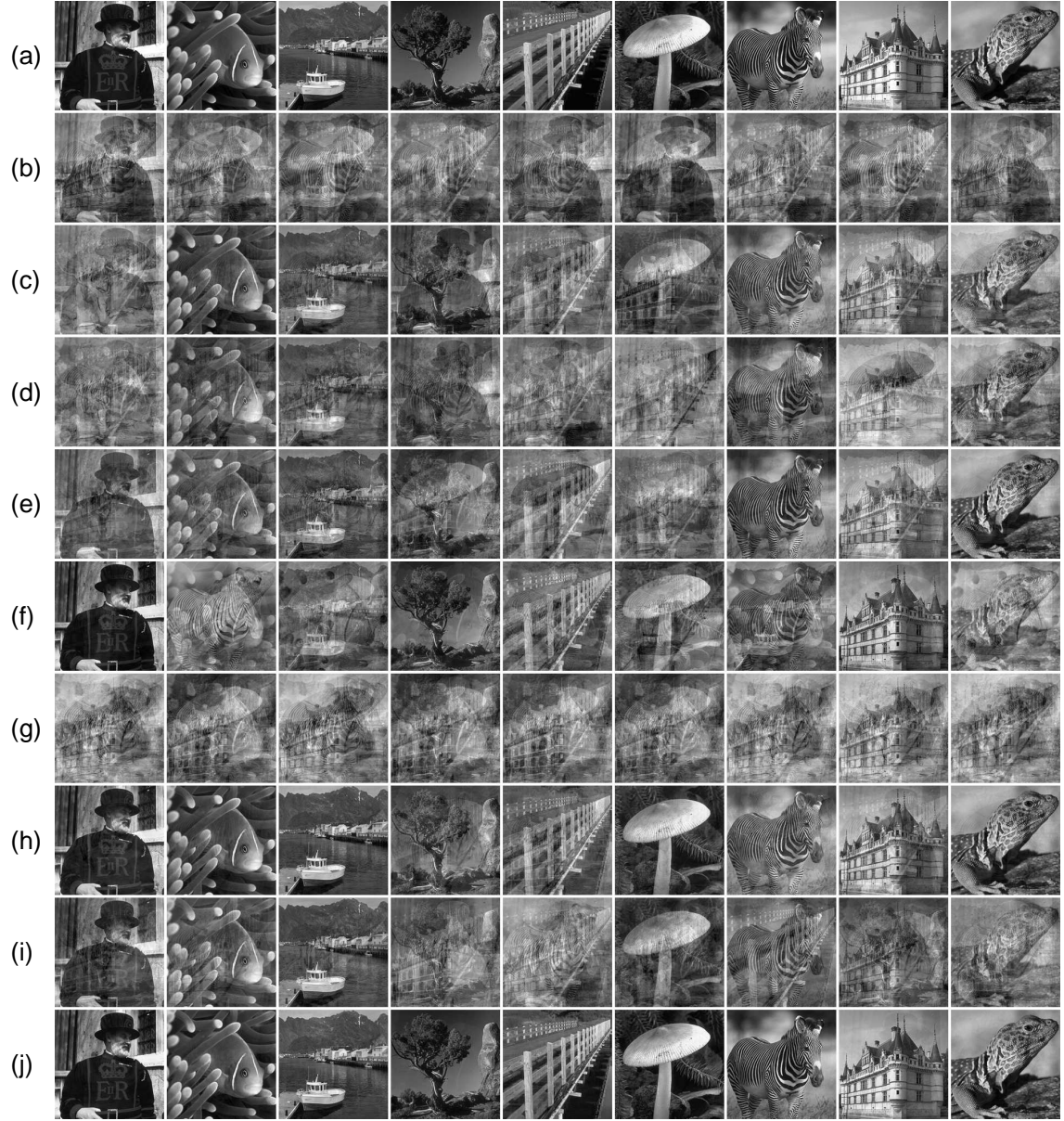


Figure 5: Simulation results with nine natural images of size 200×200 for subjective assessment. (a) Original source images. (b) Mixed images. (c)–(j) Estimated and reordered sources from (c) fastICA, (d) JADE, (e) infomax, (f) SWICA, (g) NOSWICA, (h) ICA-EBM, (i) QICA and (j) OBMADS with the RMSE values of 0.265, 0.353, 0.267, 0.248, 0.396, 0.127, 0.268 and 0.039, respectively.

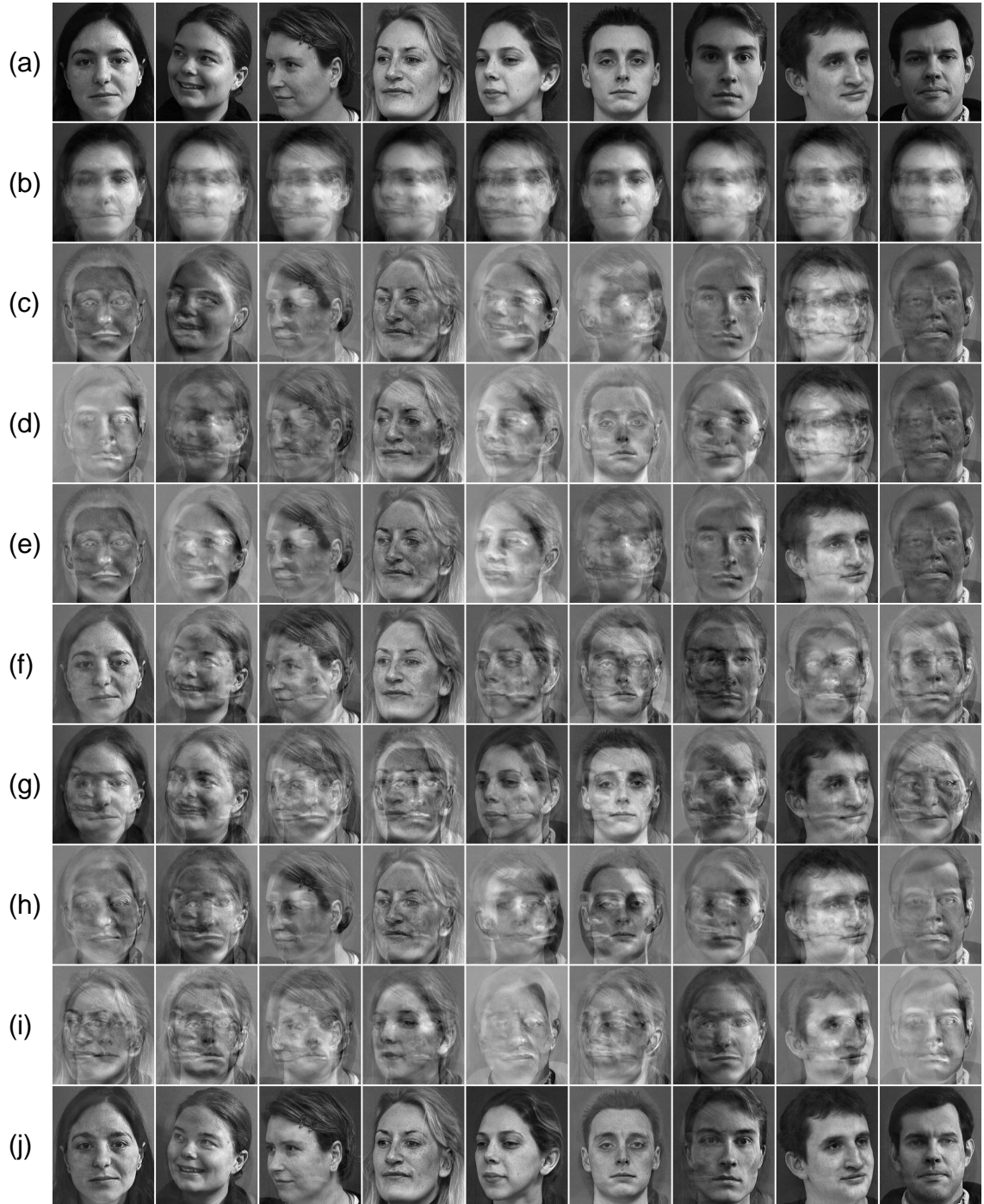


Figure 6: Unmixing results using nine face images of size 350×300 for visual scrutiny. (a) Original source images. (b) Mixed images. (c)–(j) Estimated and reordered sources from (c) fastICA, (d) JADE, (e) infomax, (f) SWICA, (g) NOSWICA, (h) ICA-EBM, (i) QICA and (j) OBMADS with the RMSE values of 0.458, 0.466, 0.468, 0.375, 0.365, 0.451, 0.467 and 0.114, respectively.

trices, the NOSWICA intentionally relaxes this constraint in estimating the ICs which hopefully represent the correlated sources. The support-width ICA methods have been claimed to enjoy the following advantages (Vrins et al., 2007): (i) an improvement in the separation performance; (ii) applicability in situations where sources may be correlated; (iii) ability to recover the sources with strongly bimodal densities; (iv) robustness to the dimensionality of the source space. The emphasis in (Vrins et al., 2007) is a systematic study of the range-based contrast function, and the nonsmooth optimization recommended in the SWICA is admittedly simple. The NOSWICA motivates further research in the direction of allowing more degrees of freedom in the ICA estimation by taking into account the correlatedness assumption among the sources constructed from real-world data. It has been remarked in (Vrins, 2007, p. 262) that the robustness of NOSWICA is disappointing as the relaxation of orthogonality constraint comes with a price of a larger space of solutions, which may account for the variability in results. Indeed the aforesaid reasons underpin the development of a robust nonsmooth optimization algorithm on the \mathcal{OB} manifold that can efficiently handle the range-based function as corroborated by the mean RMSE values listed in Table 1.

Since we intend to compare our approach with some recent ICA algorithms, the ICA-EBM and QICA were incorporated into the empirical study. The QICA offers a unified framework for a number of independence measures, and it generalizes the concept of information theoretic learning (ITL); the similarity between two random variables is evaluated by the inner product of their probability density functions. Since the QICA stresses the accurate estimation of ICs rather than speeding up the estimation, we find it appropriate for the comparative study. The ICA-EBM relies on accurately estimating the entropy of a random variable with the help of the maximum entropy bound, given the observations. The method is claimed to be effective in separating the sources following a wide range of distributions. The fact that it is more common to encounter not-so-trivial distributions of the image pixel gray-levels justifies the inclusion of the ICA-EBM in simulations. In spite of increased computational load, the QICA gave rise to a large mean RMSE. On the other hand, the source estimates of acceptable quality were obtained in the ICA-EBM.

Noteworthy, the OBMADS is capable of estimating the ICs more accurately, and the reduction in the mean RMSE of the source estimates can be ascribed to the efficiency of the Riemannian nonsmooth optimizer which ensures convergence to a Clarke stationary point. Furthermore, the Wilcoxon’s signed-rank test has been carried out to affirm that the RMSE reduction attained by OBMADS in unmixing the sources from all the experimented image datasets is statistically significant (at the 0.05 significance level) compared to the competing approaches. Except in three² out of 35 instances, the p -value obtained is 1.2290×10^{-05} , which implies that the RMSE from OBMADS is less than that of the method under comparison in all the 25 trial runs. It is remarked that in

² p -values corresponding to the ICA-EBM versus OBMADS in the aerial and natural II dataset are 1.3898×10^{-05} and 3.6243×10^{-05} , respectively, as OBMADS outclassed the ICA-EBM in 24 trial runs in the former and in 23 in the latter case. Similarly, in the natural II dataset, the SWICA versus OBMADS resulted in a p -value of 1.3898×10^{-05} .

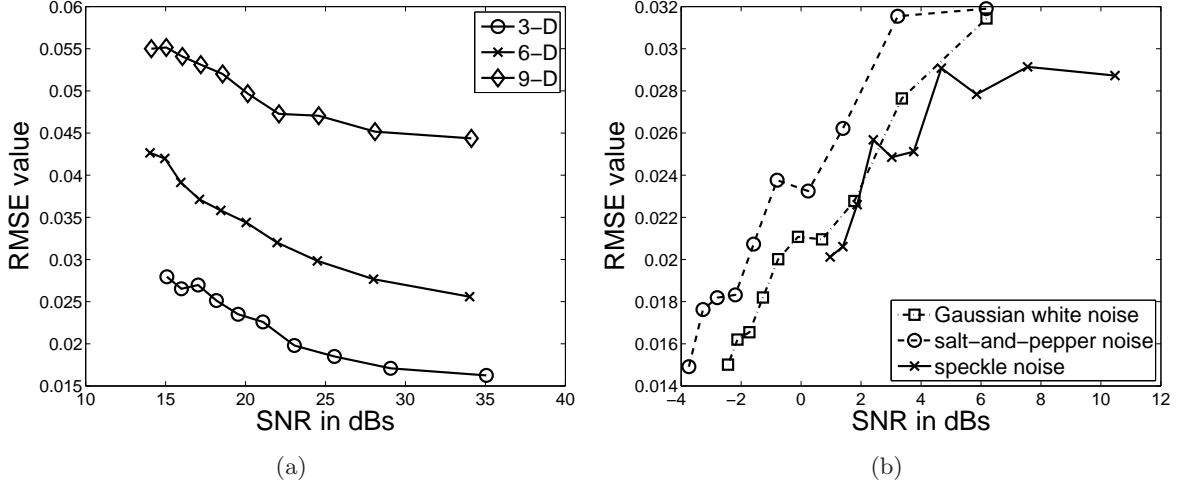


Figure 7: Effect of noise perturbation on the unmixing performance of OBICA measured by the RMSE value, when the image sources are corrupted using (a) unbounded AWGN; (b) bounded noise models (`imnoise` command in Matlab) prescribed for images. The signal-to-noise-ratio (SNR) is specified in decibels (dBs).

the Wilcoxon’s signed-rank test, when a random sample of size z is element-wise strictly less (or more) than the other sample, then the statistic assumes a fixed value given by $z(z+1)/2$ regardless of the actual values of the samples. As a consequence, the p -value will remain the same for the comparisons wherein the investigated method outperforms the competitors in a fixed number of trial runs. The Wilcoxon’s signed-rank test is more meaningful here, since it belongs to the category of nonparametric hypothesis testing methods, where the assumption that two compared populations are normally distributed does not necessarily hold.

Finally, to bear evidence subjectively for the appealing performance of the proposed optimizer coupled with a range-based contrast function, two instances of the natural and face image simulation results from all the investigated ICA algorithms are portrayed in Figures 5 and 6 along with the respective RMSE values.

7.2 Discussion on OBMADS Behavior

7.2.1 Noise Perturbation

To investigate the robustness of OBMADS under noise perturbations, the vector of concatenated pixel gray-levels from a noisy image is first considered as an unbounded source, i.e., the data-points are not necessarily confined to the interval $[0, 255]$. A linear combination of three, six or nine natural images contaminated with the additive white Gaussian noise (AWGN), having zero mean and the standard deviation incrementally selected in the interval $[1, 10]$, was supplied as the input for the OBMADS algorithm. As can be envisaged, the unmixing performance measured by the RMSE degrades as

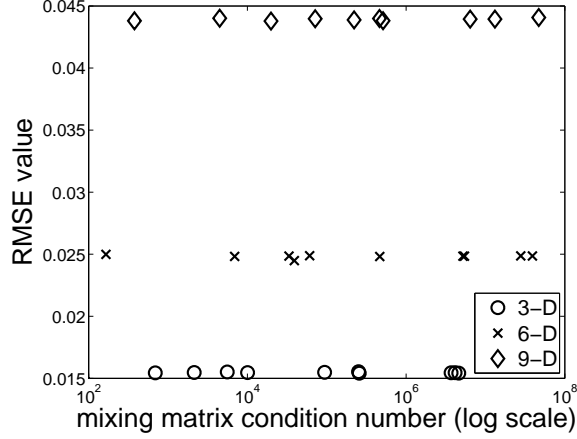


Figure 8: The plot of the RMSE between the true and estimated sources versus the condition number of the mixing matrices suggests that the performance of OBMADS is not influenced by an ill-conditioned mixing matrix.

shown in Figure 7(a). Nevertheless, in the aforementioned setting, the boundedness assumption to be valid for the sources in a range-based ICA model is clearly violated. Therefore an interesting observation is included in Figure 7(b) for a set of six natural images, while various types of noise—AWGN, salt-and-pepper noise and multiplicative speckle noise—commonly used by the image processing community were introduced in an artificial manner³. When the noisy pixel values are bounded, the situation turns out to be in favor of our approach, since a consequence of noise addition is an increase in the number of pixels with extreme gray-levels. Apparently the pixels whose gray-levels are either zero or 255 facilitate any range-based ICA algorithm to estimate the range of the unmixed sources more accurately.

7.2.2 Ill-conditioned Mixing Matrix

To examine whether the performance of OBMADS relies on the condition number of the mixing matrix, mixtures of three, six or nine natural images generated using mixing matrices with a wide-range of condition numbers were allowed to be source-separated. The RMSEs between the true and estimated sources that correspond to different condition numbers of the mixing matrices were recorded. The plot in Figure 8 enables us to conclude that the separation performance of OBMADS is not adversely affected by the ill-conditioned mixing matrix.

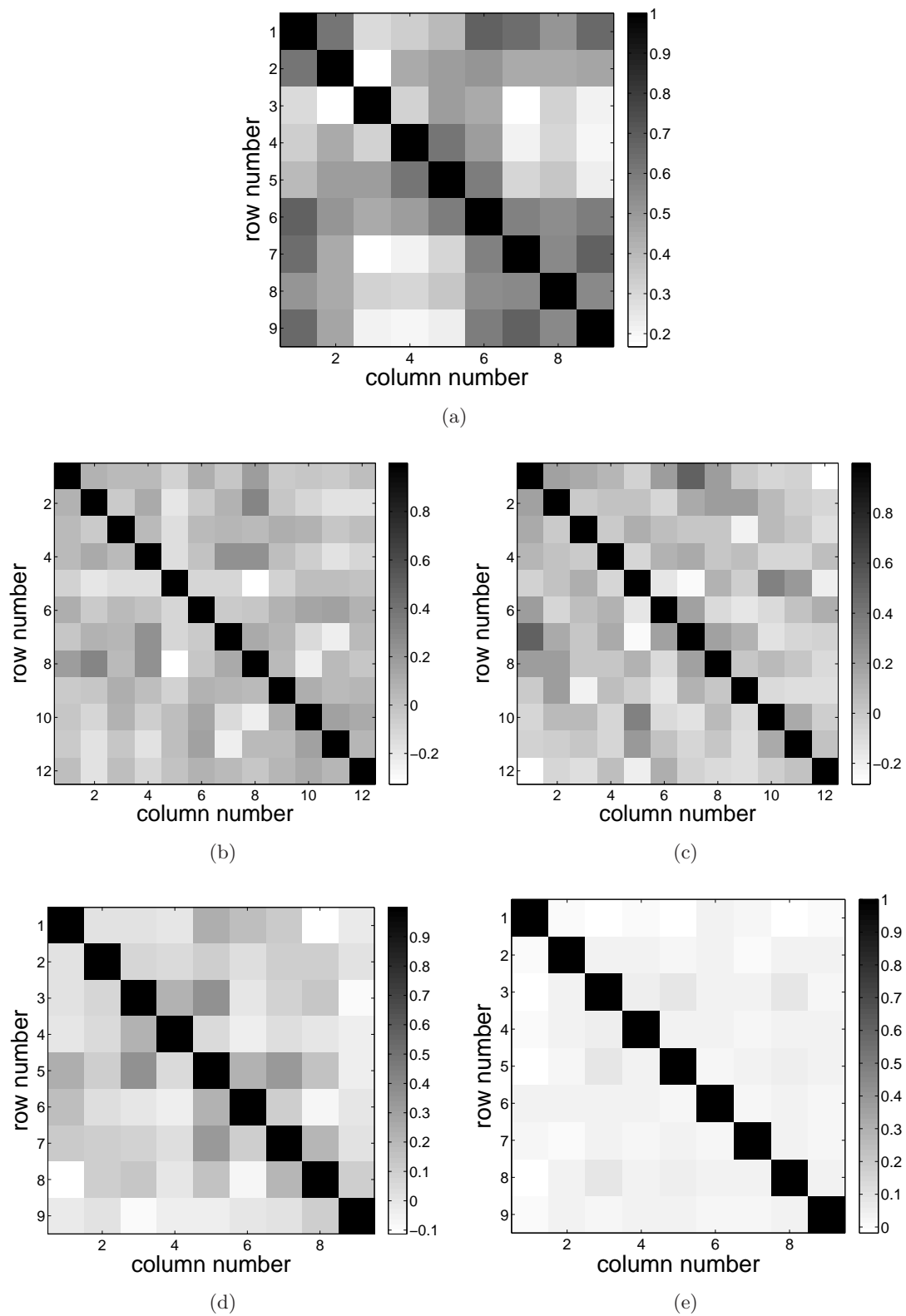


Figure 9: Correlation matrices generated with the source images are displayed for the following datasets: (a) face; (b) natural I; (c) natural II; (d) aerial; (e) texture.

Table 2: Mean and standard deviation of the computational time (in seconds) taken by various ICA methods in ten trial executions. In each run, all the algorithms were supplied with the same mixture generated from 6-, 7-, 8- or 9-D natural image data of size 50×50 . This table should be considered together with Figure 5, which shows that OBMADS noticeably dominates the other methods in terms of unmixing quality.

ICA algorithm	mean and std. dev. of computational time (seconds)			
	6-D	7-D	8-D	9-D
fastICA	0.283 ± 0.522	0.322 ± 0.507	0.297 ± 0.579	0.296 ± 0.673
JADE	0.012 ± 0.006	0.016 ± 0.006	0.019 ± 0.007	0.028 ± 0.007
infomax	2.118 ± 0.310	2.309 ± 0.266	2.727 ± 0.611	3.659 ± 1.356
SWICA	1.032 ± 0.009	1.534 ± 0.014	2.147 ± 0.015	3.481 ± 0.034
NOSWICA	0.920 ± 0.175	1.304 ± 0.243	1.772 ± 0.253	2.579 ± 0.345
ICA-EBM	0.308 ± 0.053	0.361 ± 0.083	0.553 ± 0.117	0.600 ± 0.119
QICA	32.373 ± 0.287	45.365 ± 0.370	59.963 ± 0.901	77.398 ± 1.021
OBMADS	19.072 ± 1.161	31.391 ± 0.288	39.097 ± 0.092	49.711 ± 0.383
CE-vMF	32.307 ± 5.943	62.567 ± 9.572	105.657 ± 15.501	196.412 ± 29.310

7.2.3 Source Correlatedness

As the proposed optimization relaxes the orthogonality constraint and remains on \mathcal{OB} , it is ideally suited for unmixing the sources which are not uncorrelated. Importantly, it is not uncommon to come across sources in practice which do not comply with the strict uncorrelatedness assumption imposed by several ICA algorithms. For instance, the face or landscape images share a common underlying pattern and tend to be highly correlated (Lee et al., 2006b). In such circumstances, as emphatically stated in (Selvan et al., 2012a), the whitening process does not restrict the search for the unmixing matrix to the space of orthogonal matrices. To visualize the degree of correlation amongst the images in each dataset (face, natural, aerial and texture), a pictorial representation of the correlation matrices is provided in Figure 9. A closer scrutiny would reveal that the face (Figure 9(a)) or natural images (Figures 9(b) and 9(c)) are more correlated than the aerial (Figure 9(d)) or texture images (Figure 9(e)). The efficacy of the proposed \mathcal{OB} optimization algorithm can be endorsed by the illustrations in Figures 5 and 6, and the relatively small RMSE values yielded by OBMADS in comparison with the state-of-the-art approaches for face and natural image sources as consolidated in Table 1.

7.2.4 Computational Cost

The software system was implemented in Matlab R2011b on a MacBook Pro (Intel Core i7 2.2 GHz CPU, 8 GB 1333 MHz DDR3) using Mac OS X Lion 10.7.4.

In Figure 5, the proposed OBMADS is the only tested method that arguably produces an unmixing of sufficient quality. Its computational time was of the order of one minute, which makes it adequate for offline source separation tasks. For information, timing comparisons are furnished in Table 2. The table also features the CE-vMF method of (Selvan et al., 2012b), which produces results of similar quality as OBMADS but is seen to be slower.

8 Conclusion

We have proposed and analyzed an extension to the sphere of the MADS algorithm of (Audet and Dennis, 2006). A specific instance, called LTSMADS, of the resulting derivative-free algorithmic framework has been proposed and extended to an optimization algorithm, called OBMADS, evolving on the set of matrices with unit-norm column constraint. OBMADS has then been applied to a range-based contrast function for ICA due to Pham, Vrms and collaborators. An analysis has been carried out to show that OBMADS converges to Clarke stationary points of the contrast function. Since the contrast function has the discriminatory property, it follows that bounded and independent sources are recovered in practice from their linear mixtures by OBMADS in the infinite-sample limit. The efficacy of OBMADS when the unmixing matrix is estimated from a mixture of quasi-uncorrelated sources has been demonstrated using simulations with natural, face, aerial and texture images. Though the paper focuses on an ICA application, the proposed optimization algorithms are applicable to several potential problems in science where one has to optimize a locally Lipschitz continuous function subject to unit-norm constraints.

References

- P.-A. Absil and K. A. Gallivan. Joint diagonalization on the oblique manifold for independent component analysis. In *Proceedings of the 31st IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 945–948. IEEE, 2006.
- P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008. ISBN 978-0-691-13298-3.
- C. Audet and J. E. Dennis, Jr. Analysis of generalized pattern searches. *SIAM Journal on Optimization*, 13(3):889–903, 2002.

³The Matlab command `imnoise` adds a specific type of noise to an image such that the noisy image pixel values are bounded in the interval $[0, 255]$.

- C. Audet and J. E. Dennis, Jr. Mesh adaptive direct search algorithms for constrained optimization. *SIAM Journal on Optimization*, 17(1):188–217, 2006.
- R. Boscolo, H. Pan, and V. P. Roychowdhury. Independent component analysis based on nonparametric density estimation. *IEEE Transactions on Neural Networks*, 15(1):55–65, January 2004.
- J.-F. Cardoso and A. Souloumiac. Blind beamforming for non Gaussian signals. *IEE Proceedings-F Radar and Signal Processing*, 140(6):362–370, December 1993.
- I. D. Coope and C. J. Price. Frame based methods for unconstrained optimization. *Journal of Optimization Theory and Applications*, 107(2):261–274, November 2000.
- S. C. Douglas, S.-i. Amari, and S.-Y. Kung. On gradient adaptation with unit-norm constraints. *IEEE Transactions on Signal Processing*, 48(6):1843–1847, June 2000.
- A. Hyvärinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634, May 1999a.
- A. Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2(4):94–128, 1999b.
- A. Hyvärinen. Independent component analysis: Recent advances. *Philosophical Transactions of the Royal Society A*, 1:1–25, 2011.
- A. Hyvärinen, J. Karhunen, and E. Oja. *Independent component analysis*. Wiley-Interscience, 2001.
- J. Karhunen, E. Oja, L. Wang, R. Vigário, and J. Joutsensalo. A class of neural networks for independent component analysis. *IEEE Transactions on Neural Networks*, 8(3):486–504, May 1997.
- J. A. Lee, F. Vrins, and M. Verleysen. A least absolute bound approach to ICA-application to the MLSP 2006 competition. In *Proceedings of the 16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing*, pages 41–46, Maynooth, Ireland, September 2006a. IEEE.
- J. A. Lee, F. Vrins, and M. Verleysen. Non-orthogonal support-width ICA. In *Proceedings of the 14th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, pages 351–358, Bruges, Belgium, April 2006b.
- X.-L. Li and T. Adalı. Independent component analysis by entropy bound minimization. *IEEE Transactions on Signal Processing*, 58(10):5151–5164, October 2010.
- S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski. Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems*, 8:145–151, 1996.

- D. R. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th IEEE International Conference on Computer Vision*, volume 2, pages 416–425, Vancouver, Canada, July 2001. IEEE.
- D.-T. Pham. Blind separation of instantaneous mixture of sources based on order statistics. *IEEE Transactions on Signal Processing*, 48(2):363–375, February 2000.
- D.-T. Pham. Joint approximate diagonalization of positive definite Hermitian matrices. *SIAM Journal on Matrix Analysis and Applications*, 22(4):1136–1152, March 2001.
- D.-T. Pham and F. Vrins. Discriminacy of the minimum range approach to the simultaneous blind separation of bounded sources. In M. Verleysen, editor, *Advances in Computational Intelligence and Learning*, pages 377–382, April 2006. 14th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.
- S. E. Selvan, U. Amato, C. Qi, K. A. Gallivan, M. F. Carfora, M. Larobina, and B. Alfano. Descent algorithms on oblique manifold for source-adaptive ICA contrast. *IEEE Transactions on Neural Networks and Learning Systems*, 23(12):1930–1947, December 2012a.
- S. E. Selvan, A. Chattopadhyay, U. Amato, and P.-A. Absil. Range-based non-orthogonal ICA using cross-entropy method. In *Proceedings of the 20th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, pages 519–524, Bruges, Belgium, April 2012b.
- S. Seth, M. Rao, I. Park, and J. C. Príncipe. A unified framework for quadratic measures of independence. *IEEE Transactions on Signal Processing*, 59(8):3624–3635, August 2011.
- H. Shen and K. Hüper. Block Jacobi-type methods for non-orthogonal joint diagonalisation. In *Proceedings of the 34th IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3285–3288, Taipei, Taiwan, April 2009. IEEE.
- H. Shen and M. Kleinstenber. Complex blind source separation via simultaneous strong uncorrelating transform. *Latent Variable Analysis and Signal Separation*, 6365:287–294, 2010.
- H. Shen, K. Diepold, and K. Hüper. Geometric algorithms for the non-whitened one-unit linear independent component analysis problem. In *Proceedings of the IEEE/SP 15th Workshop on Statistical Signal Processing*, pages 381–384, Cardiff, UK, August–September 2009. IEEE.
- M. B. Stegmann, B. K. Ersbøll, and R. Larsen. FAME – A flexible appearance modelling environment. *IEEE Transactions on Medical Imaging*, 22(10):1319–1331, October 2003.

- V. Torczon. On the convergence of pattern search algorithms. *SIAM Journal on Optimization*, 7(1):1–25, February 1997.
- N. T. Trendafilov and R. A. Lippert. The multimode Procrustes problem. *Linear Algebra and Its Applications*, 349(1):245–264, July 2002.
- A.-J. van der Veen. Joint diagonalization via subspace fitting techniques. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 2773–2776, Salt Lake City, UT, May 2001. IEEE.
- F. Vrins. *Contrast properties of entropic criteria for blind source separation: A unifying framework based on information-theoretic inequalities*. PhD thesis, Faculté des Sciences Appliquées, Université catholique de Louvain, Louvain-la-Neuve, Belgium, March 2007.
- F. Vrins, J. A. Lee, and M. Verleysen. A minimum-range approach to blind extraction of bounded sources. *IEEE Transactions on Neural Networks*, 18(3):809–822, May 2007.
- A. Yeredor. Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. *IEEE Transactions on Signal Processing*, 50(7):1545–1553, July 2002.