Identifiability of the Simplex Volume Minimization Criterion for Blind Hyperspectral Unmixing: The No Pure-Pixel Case

[†]Chia-Hsiang Lin, [‡]Wing-Kin Ma, [†]Wei-Chiang Li, [†]Chong-Yung Chi, and [†]ArulMurugan Ambikapathi

[†]Institute of Communications Engineering, National Tsing Hua University, Taiwan, R.O.C.
Emails: chiahsiang.steven.lin@gmail.com, weichiangli@gmail.com, cychi@ee.nthu.edu.tw, aareul@ieee.org

[‡]Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong
Email: wkma@ieee.org

May 30, 2014

Abstract

In blind hyperspectral unmixing (HU), the pure-pixel assumption is well-known to be powerful in enabling simple and effective blind HU solutions. However, the pure-pixel assumption is not always satisfied in an exact sense, especially for scenarios where pixels are all intimately mixed. In the no pure-pixel case, a good blind HU approach to consider is the minimum volume enclosing simplex (MVES). Empirical experience has suggested that MVES algorithms can perform well without pure pixels, although it was not totally clear why this is true from a theoretical viewpoint. This paper aims to address the latter issue. We develop an analysis framework wherein the perfect identifiability of MVES is studied under the noiseless case. We prove that MVES is indeed robust against lack of pure pixels, as long as the pixels do not get too heavily mixed and too asymmetrically spread. Also, our analysis reveals a surprising and counter-intuitive result, namely, that MVES becomes more robust against lack of pure pixels as the number of endmembers increases. The theoretical results are verified by numerical simulations.

1 Introduction

Signal, image and data processing for hyperspectral imaging has recently received enormous attention in remote sensing [1, 2], having numerous applications such as environmental monitoring, land mapping and classification, and object detection. Such exciting developments are made possible by exploiting the unique features of hyperspectral images, most notably, their high spectral resolutions. In this scope, blind hyperspectral unmixing (HU) is one of the topics that has aroused much interest not only from remote sensing [3], but also from other communities recently [4–7]. Simply speaking, the problem of blind HU is to solve a problem reminiscent of blind source separation in signal processing, and the desired outcome is to unambiguously separate the endmember

spectral signatures and their corresponding abundance maps from the observed hyperspectal scene, with no or little prior information of the mixing system. Being given little information to solve the problem, blind HU is a challenging—but also fundamentally intriguing—problem with many possibilities. Readers are referred to some recent articles for overview of blind HU [3, 4], and here we shall not review the numerous possible ways to perform blind HU. The focus, as well as the contribution, of this paper lie in addressing a fundamental question arising from one important blind HU approach, namely, the minimum volume enclosing simplex (MVES) approach.

Also called simplex volume minimization or minimum volume simplex analysis (MVSA) [8], the MVES approach adopts a criterion that exploits the convex geometry structures of the observed hyperspectral data to blindly identify the endmember spectral signatures. In the HU context the MVES concepts were first advocated by Craig back in the 1990's [9], although as an aside it is interesting to note an earlier work in mathematical geology [10] which also described the MVES intuitions; see the historical note in [4]. In particular, Craig's work proposes the use of simplex volume as a metric for blind HU, which is later used in some other blind HU approaches such as simplex volume maximization [11–13] and non-negative matrix factorization [14]. The MVES criterion is to minimize the volume of a simplex, subject to constraints that the simplex encloses all hyperspectral data points. This amounts to a nonconvex optimization problem, and unlike the simplex volume maximization approach we do not seem to have a simple (closed-form) scheme for tackling the MVES problem. However, recent advances in optimization have enabled us to handle MVES implementations efficiently. The works in [8] and [6] independently developed practical MVES optimization algorithms based on iterative linear approximation and alternating linear programming, respectively. The GPU-implementation of the former is also considered very recently [15]. In addition, some recent MVES algorithm designs deal with noise and outlier sensitivity issues by robust formulations, such as the soft constraint formulation in SISAL [16] and the chance-constrained formulation in [17]; the pixel elimination method in [18] should also be noted. We should further mention that MVES also finds application in analytical chemistry [19], and that fundamentally MVES has a strong link to stochastic maximum-likelihood estimation [20].

What makes MVES special is that it seems to perform well even in the absence of pure pixels, i.e., pixels that are solely contributed by a single endmember. To be more accurate, extensive simulations found that MVES may obtain the ground truth quite accurately in the noiseless case and without the pure-pixel assumption; see, e.g., [6, 20, 21]. At this point we should mention that while the pure-pixel assumption is elegant and has been exploited by some other approaches, such as simplex volume maximization (also [7] for a more recent work on near-separable non-negative matrix factorization), to arrive at remarkably simple blind HU algorithms, it is also an arguably restrictive assumption in general. In the HU context it has suspected that MVES should be resistant to lack of pure pixels, but it is not known to what extent MVES can guarantee perfect identifiability under no pure pixels. Hence, we depart from existing MVES works, wherein improved algorithm designs are usually the theme, and ask the following questions: can the identifiability of the MVES criterion in the no pure-pixel case be *theoretically* pinned down? If yes, how bad (in terms of how heavy the data are mixed) can MVES withstand and where is the limit?

The contribution of this paper is theoretical. We aim to address the aforementioned questions through analysis. Previously, identifiability analysis for MVES was done only for the pure-pixel case in [6], and for the three endmember case in the preliminary version of this paper [22]. This paper considers the no pure-pixel case for any number of endmembers. We prove that MVES can indeed guarantee exact recovery of the endmembers. The key condition for attaining such exact

identifiability is that some measures concerning the pixels' purity and geometry (to be defined in Section 3.1) have to be above a certain limit. The condition mentioned above is equivalent to the pure-pixel assumption for the case of two endmembers, and is much milder than the pure-pixel assumption for the case of three endmembers or more. Our analysis also reveals a surprising and rather counter-intuitive result—that MVES becomes more robust against lack of pure pixels as the number of endmembers increases. Numerical experiments will be conducted to verify the above claims.

This paper is organized as follows. The problem statement is described in Section 2. The MVES identifiability analysis results and the associated proofs are given in Sections 3 and 4, respectively. Numerical results are provided in Section 5 to verify our theoretical claims, and we conclude the paper in Section 6.

Notations: $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ denote the sets of all real-valued n-dimensional vectors and m-by-n matrices, respectively (resp.); $\|\cdot\|$ denotes the Euclidean norm of a vector; $\boldsymbol{x}^T$ denotes the transpose of $\boldsymbol{x}$ and the same applies to matrices; given a set $\mathcal{A}\subseteq\mathbb{R}^n$, we denote $\text{aff}\mathcal{A}$ and $\text{conv}\mathcal{A}$ as the affine hull and convex hull of $\mathcal{A}$, resp. (see [23]), $\text{int}\mathcal{A}$ and $\text{bd}\mathcal{A}$ as the interior and boundary of $\mathcal{A}$, resp., and $\text{vol}\mathcal{A}$ as the volume of $\mathcal{A}$; the dimension of a set $\mathcal{A}\subseteq\mathbb{R}^n$ is defined as the affine dimension of $\text{aff}\mathcal{A}$; $\boldsymbol{x}\geq\boldsymbol{0}$ means that $\boldsymbol{x}$ is elementwise non-negative; $\boldsymbol{I}$ and $\boldsymbol{1}$ denote an identity matrix and all-one vector of appropriate dimension, resp.; $\boldsymbol{e}_i$ denotes a unit vector whose ith element is $[\boldsymbol{e}_i]_i=1$ and jth element is $[\boldsymbol{e}_i]_j=0$ for all $j\neq i$.

2 Problem Statement

In this section we review the background of the MVES identifiability analysis challenge.

2.1 Preliminaries

Before describing the problem, some basic facts about simplex should be mentioned. A convex hull

$$\text{conv}\{\boldsymbol{b}_1,\ldots,\boldsymbol{b}_N\}=\left\{\boldsymbol{x}=\sum_{i=1}^{N}\theta_i\boldsymbol{b}_i\;\middle|\;\boldsymbol{\theta}\geq\boldsymbol{0},\boldsymbol{1}^T\boldsymbol{\theta}=1\right\},$$

where $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_N\in\mathbb{R}^M$, $M\geq N-1$, is called an $(N-1)$-dimensional simplex if $\boldsymbol{b}_1,\ldots,\boldsymbol{b}_N$ are affinely independent. The volume of a simplex can be determined by [24]

$$\text{vol}(\text{conv}\{\boldsymbol{b}_1,\ldots,\boldsymbol{b}_N\})=\frac{1}{(N-1)!}\sqrt{\det(\bar{\boldsymbol{B}}^T\bar{\boldsymbol{B}})},\tag{1}$$

where $\bar{\boldsymbol{B}}=[\;\boldsymbol{b}_1-\boldsymbol{b}_N,\boldsymbol{b}_2-\boldsymbol{b}_N,\ldots,\boldsymbol{b}_{N-1}-\boldsymbol{b}_N\;]\in\mathbb{R}^{M\times(N-1)}$. A simplex is called regular if the distances between any two vertices are the same.

2.2 Blind HU Problem Setup

We consider a standard blind HU problem formulation (readers are referred to the literature, e.g., [3, 4], for coverage of the underlying modeling aspects). Concisely, we have an observed hyperspectral scene, whose pixels can be modeled as linear mixtures of endmember spectral signatures

$$\boldsymbol{x}_n=\boldsymbol{A}\boldsymbol{s}_n,\quad n=1,\ldots,L,\tag{2}$$

where $\boldsymbol{x}_n \in \mathbb{R}^M$ denotes the nth pixel vector of the observed hyperspectral image, with M being the number of spectral bands; $\boldsymbol{A} = [\ \boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\ ] \in \mathbb{R}^{M\times N}$ is the endmember signature matrix, with N being the number of endmembers; $\boldsymbol{s}_n \in \mathbb{R}^M$ is the abundance vector of the nth pixel; L is the number of pixels. The problem is to identify the unknown $\boldsymbol{A}$ from the observations $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$, thereby allowing us to unmix the abundances (also unknown) blindly. To facilitate the subsequent problem description, the noiseless case is assumed. The following assumptions are standard in the blind HU context and will be assumed throughout the paper: (i) every abundance vector satisfies $\boldsymbol{s}_n \geq 0$ and $\mathbf{1}^T\boldsymbol{s}_n = 1$ (i.e., the abundance non-negativity and sum-to-one constraints); (ii) $\boldsymbol{A}$ has full column rank; (iii) $[\ \boldsymbol{s}_1, \ldots \boldsymbol{s}_L\ ]$ has full row rank; (iv) N is known.

2.3 Minimum-Volume Enclosing Simplex

This paper concentrates on the MVES approach for blind HU. MVES was inspired by the following intuition [9]: if we can find a simplex that circumscribes the data points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L$ and yields the minimum volume, then the vertices of such a simplex should be identical to, or close to, the true endmember spectral signatures $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N$ themselves. Figure 1 shows an illustration to support why the aforementioned intuition may be true. Mathematically, the MVES criterion can be formulated as an optimization problem

$$\begin{aligned} \min_{\boldsymbol{b}_1,\ldots,\boldsymbol{b}_N\in\mathbb{R}^M} \ & \mathrm{vol}(\mathrm{conv}\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_N\}) \\ \text{s.t. } & \boldsymbol{x}_n \in \mathrm{conv}\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_N\},\ n = 1, \ldots, L, \end{aligned} \tag{3}$$

wherein the solution of problem (3) is used as an estimate of $\boldsymbol{A}$. Algorithms for handling problem (3) have been previously proposed; see, e.g., [6, 8, 19, 20], and also [14, 16–18] for the noisy case. Our interest is not with MVES algorithm design. Instead, we consider the following fundamental, and very important, question: *When will the MVES criterion in* (3) *provide a solution that is unique and exactly given by the true endmember matrix* $\boldsymbol{A}$ *(up to a permutation)?*

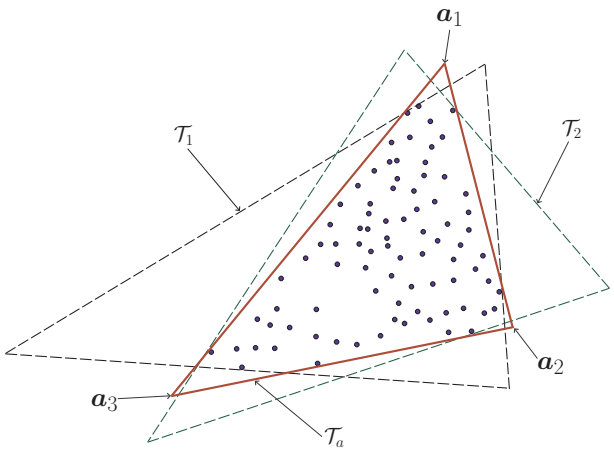


Figure 1: A geometrical illustration of MVES. The dots are the data points $\{\boldsymbol{x}_n\}$, the number of endmembers is $N = 3$, and $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_a$ are data-enclosing simplices. In particular, $\mathcal{T}_a$ is actually given by $\mathcal{T}_a = \mathrm{conv}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$. Visually, it can be seen that $\mathcal{T}_a$ has a smaller volume than $\mathcal{T}_1$ and $\mathcal{T}_2$.

It is known that MVES uniquely identifies $\boldsymbol{A}$ if the pure-pixel assumption holds [6], that is, if, for each $i \in \{1, \ldots, N\}$, there exists an abundance vector $\boldsymbol{s}_n$ such that $\boldsymbol{s}_n = \boldsymbol{e}_i$. However, empirical evidence has suggested that even when the pure-pixel assumption does not hold and the data are somehow intimately mixed, MVES may still be able to uniquely identify $\boldsymbol{A}$. At present, there are few works on providing provable MVES identifiability results in the no pure-pixel case, and this paper aims at doing so.

3 Main Results

This section describes the main results of our MVES identifiability analysis. As will be seen soon, MVES identifiability in the no pure-pixel case depends much on the level of "pixel purity" of the observed data set. To this end, we need to precisely quantify what "pixel purity" is. The first subsection will introduce two pixel purity measures. The second subsection will then present the main results, and discuss their implications in practice.

3.1 Pixel Purity Measures

A natural way to quantify pixel purity is to use the following measure

$$\rho = \sup\{\|\boldsymbol{s}_n\| \mid n = 1, \ldots, L\}. \tag{4}$$

Eq. (4) will be called the *best pixel purity level* in the sequel. A large ρ implies that there exist abundance vectors whose purity is high, while a small ρ indicates more heavily mixed data. To see it, observe that $\|\boldsymbol{s}\| \leq 1$ for any $\boldsymbol{s} \geq \boldsymbol{0}$, $\boldsymbol{1}^T\boldsymbol{s} = 1$, and equality holds if and only if $\boldsymbol{s} = \boldsymbol{e}_k$ for any k; that is, a pure pixel. Moreover, it can be shown that $\frac{1}{\sqrt{N}} \leq \|\boldsymbol{s}\|$ for any $\boldsymbol{s} \geq \boldsymbol{0}$, $\boldsymbol{1}^T\boldsymbol{s} = 1$, and equality holds if and only if $\boldsymbol{s} = \frac{1}{N}\boldsymbol{1}$; that is, a heavily mixed pixel. Without loss of generality, we may assume

$$\frac{1}{\sqrt{N}} < \rho \leq 1,$$

where we rule out $\rho = \frac{1}{\sqrt{N}}$, which implies $\boldsymbol{s}_1 = \ldots = \boldsymbol{s}_L = \frac{1}{N}\boldsymbol{1}$ and leads to a pathological case.

The previously defined pixel purity level reflects the best abundance purity among all the pixels, but says little on how the pixels are spread geometrically with respect to (w.r.t.) the various endmembers. We will also require another measure, defined as follows

$$\gamma = \sup\{r \mid \mathcal{R}(r) \subseteq \text{conv}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\}\}, \tag{5}$$

where

$$\mathcal{R}(r) = \{\boldsymbol{s} \in \mathbb{R}^N \mid \|\boldsymbol{s}\| \leq r\} \cap \text{conv}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}. \tag{6}$$

We call (5) the *uniform pixel purity level*; the reason for this will be illustrated soon. It can be shown that

$$\frac{1}{\sqrt{N}} \leq \gamma \leq \rho.$$

Also, if $\gamma = 1$, then the pure-pixel assumption is shown to hold.

To understand the differences between the pixel purity measures in (4) and (5), we first illustrate how $\mathcal{R}(r)$ looks like in Figure 2. As can be seen (and as will be shown), $\mathcal{R}(r)$ is a ball on the affine

hull aff$\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_N\}$ if $r \leq 1/\sqrt{N-1}$. Otherwise, $\mathcal{R}(r)$ takes a shape like a vertices-cropped version of the unit simplex conv$\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_N\}$. In addition, it can be shown that (4) equals

$$\rho = \inf\{r \mid \text{conv}\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_L\} \subseteq \mathcal{R}(r)\}.$$

In Figure 3, we give several examples with the abundances. From the figures, an interesting observation is that $\mathcal{R}(\rho)$ serves as a smallest $\mathcal{R}(r)$ that circumscribes the abundance convex hull conv$\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_L\}$, while $\mathcal{R}(\gamma)$ serves as a largest $\mathcal{R}(r)$ that is inscribed in conv$\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_L\}$. Moreover, we see that if the abundances are spread in a relatively symmetric manner w.r.t. all the endmembers, then ρ and γ are similar; this is the case with Figures 3(a)-3(c). However, ρ and γ can be quite different if the abundances are asymmetrically spread; this is the case with Figure 3(d) where some endmembers have pixels of high purity but some do not. Hence, the uniform pixel purity level γ quantifies a pixel purity level that applies uniformly to *all* the endmembers, not just to the best.

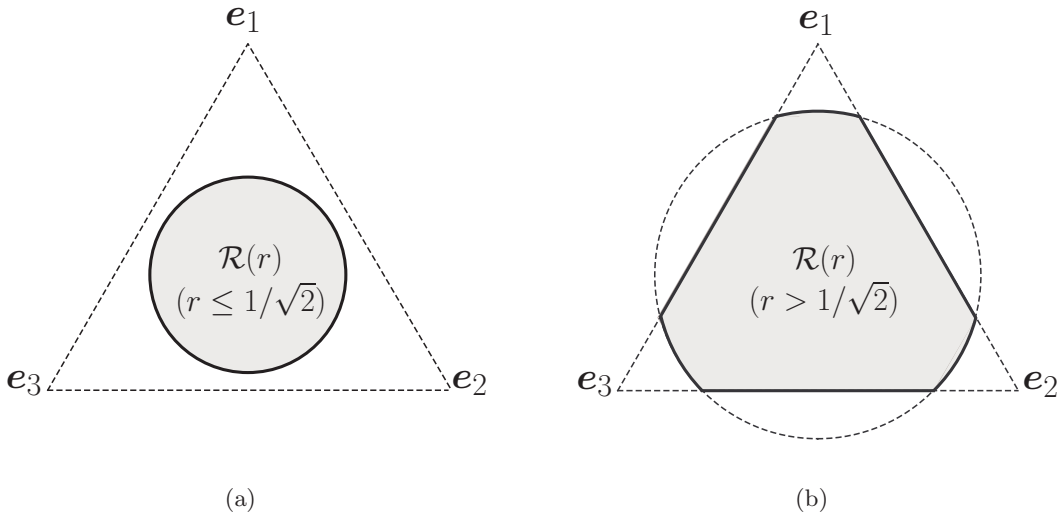


Figure 2: A geometrical illustration of $\mathcal{R}(r)$ in (6) for $N = 3$. We view $\mathcal{R}(r)$ by adjusting the viewpoint to be perpendicular to the affine hull of $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$.

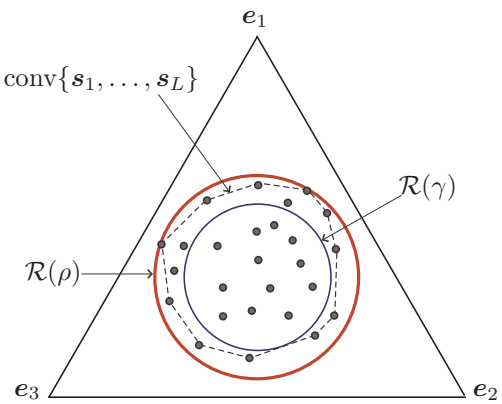
3.2 Provable MVES Identifiability and Implications

Our provable MVES identifiability results are described as follows. To facilitate our analysis, consider the following definition.

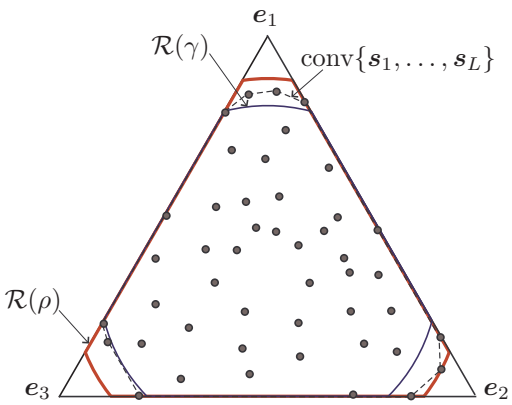
Definition 1 (minimum volume enclosing simplex) *Given an m-dimensional set $\mathcal{U} \subseteq \mathbb{R}^n$, the notation* $\mathsf{MVES}(\mathcal{U})$ *denotes the set that collects all m-dimensional minimum volume simplices that enclose $\mathcal{U}$ and lie in* aff$\mathcal{U}$.

Now, let

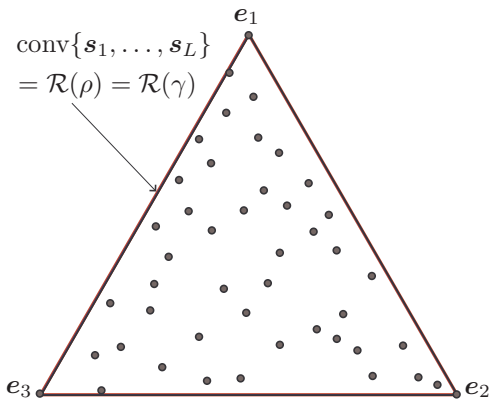
$$\mathcal{T}_e = \text{conv}\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_N\} \subseteq \mathbb{R}^N,$$
$$\mathcal{T}_a = \text{conv}\{\boldsymbol{a}_1,\ldots,\boldsymbol{a}_N\} \subseteq \mathbb{R}^M,$$



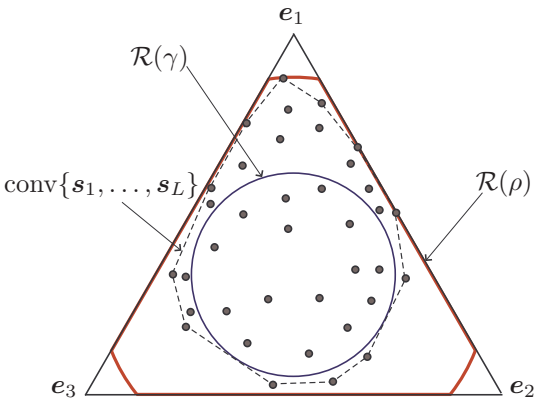
(a) $\gamma < 1/\sqrt{2}, \rho < 1/\sqrt{2}$



(b) $\gamma > 1/\sqrt{2}, \rho > 1/\sqrt{2}$



(c) $\gamma = \rho = 1$



(d) $\gamma < 1/\sqrt{2}$, $\rho > 1/\sqrt{2}$

Figure 3: Examples with the abundance distributions and the corresponding best and uniform pixel purity levels.

denote the $(N-1)$-dimensional unit simplex and the endembers' simplex, respectively. Also, for convenience, let

$$\mathcal{X}_L = \{\boldsymbol{x}_1, \dots \boldsymbol{x}_L\}, \quad \mathcal{S}_L = \{\boldsymbol{s}_1, \dots \boldsymbol{s}_L\}.$$

Under the above definition, the exact and unique identifiability problem of the MVES criterion in (3) can be posed as a problem of finding conditions under which

$$\mathsf{MVES}(\mathcal{X}_L) = \{\mathcal{T}_a\}.$$

Our first result reveals that the MVES perfect identifiability does not depend on $\boldsymbol{A}$ (as far as $\boldsymbol{A}$ has full column rank):

Proposition 1 $\mathsf{MVES}(\mathcal{X}_L) = \{\mathcal{T}_a\}$ *if and only if* $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$.

The proof of Proposition 1, as well as those of the theorems to be presented, will be provided in the next section. Proposition 1 suggests that to analyze the perfect MVES identifiability w.r.t. the observed pixel vectors, it is equivalent to analyze the perfect MVES identifiability w.r.t. the abundance vectors. One may expect that perfect identifiability cannot be achieved for too heavily mixed pixels. We prove that this is indeed true.

Theorem 1 *Assume* $N \geq 3$. *If* $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$, *then the best pixel purity level must satisfy* $\rho > \frac{1}{\sqrt{N-1}}$.

To get some idea, consider the example in Figure 3(a). Since Figure 3(a) does not satisfy the condition in Theorem 1, it fails to provide exact recovery of the true endmembers. Theorem 1 is only a necessary perfect identifiability condition. We also prove a sufficient perfect identifiability condition, described as follows:

Theorem 2 *Assume* $N \geq 3$. *If the uniform pixel purity level satisfies* $\gamma > \frac{1}{\sqrt{N-1}}$, *then* $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$.

Among the four examples in Figure 3, Figure 3(b) and Figure 3(c) are cases that satisfy the condition in Theorem 2 and achieve exact and unique recovery of the true endmembers.

It is worthwhile to emphasize that the sufficient identifiability condition in Theorem 2 is much milder than the pure-pixel assumption (which is equivalent to $\gamma = 1$) for $N \geq 3$, especially when N is large. In fact, for a very large N, the pixel purity requirement $1/\sqrt{N-1}$ vanishes to zero—which implies that MVES can handle increasingly mixed cases as the number of endmembers increases. While such a conclusion is rather counter-intuitive, it for the first time provides a theoretical justification on why previous works found that MVES or related algorithms can be so robust against lack of pure pixels in practice.

Before we finish this subsection, we should mention the case of $N = 2$. While the number of endmembers in practical scenarios is often a lot more than two, it is still interesting to know the identifiability for $N = 2$.

Proposition 2 *Assume* $N = 2$. *We have* $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$ *if and only if the pure-pixel assumption holds.*

We should recall that the pure-pixel assumption corresponds to $\gamma = 1$.

4 Proof of The Main Results

This section provides the proof of the main results described in the previous section. Readers who are more interested in numerical experiments may jump to Section 5.

4.1 Proof of Proposition 1

The following lemma will be used to prove Proposition 1:

Lemma 1 *Let $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A} \in \mathbb{R}^{M\times N}$, $M \geq N$, and suppose that $\boldsymbol{A}$ has full column rank.*

(a) Let $\mathcal{T}_G \subset \mathbb{R}^N$ be an $(N-1)$-dimensional simplex, and suppose $\mathcal{T}_G \subset \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$. We have

$$\mathrm{vol}(f(\mathcal{T}_G)) = \alpha \cdot \mathrm{vol}(\mathcal{T}_G), \tag{7}$$

where $\alpha = \sqrt{\frac{\det(\bar{\boldsymbol{A}}^T\bar{\boldsymbol{A}})}{N}}$, and $\bar{\boldsymbol{A}} = [\ \boldsymbol{a}_1 - \boldsymbol{a}_N, \boldsymbol{a}_2 - \boldsymbol{a}_N, \ldots, \boldsymbol{a}_{N-1} - \boldsymbol{a}_N\ ]$. Also, it holds true that $f(\mathcal{T}_G) \subset \mathrm{aff}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\}$.

(b) Let $\mathcal{T}_H \subset \mathbb{R}^M$ be an $(N-1)$-dimensional simplex, and suppose $\mathcal{T}_H \subset \mathrm{aff}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\}$. We have

$$\mathrm{vol}(f^{-1}(\mathcal{T}_H)) = \frac{1}{\alpha} \cdot \mathrm{vol}(\mathcal{T}_H), \tag{8}$$

and $f^{-1}(\mathcal{T}_H) \subset \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$.

The proof of Lemma 1 is relegated to Appendix A. Now, suppose that $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$, but $\mathsf{MVES}(\mathcal{X}_L) \neq \{\mathcal{T}_a\}$. Let $\mathcal{T}_H$ be an MVES of $\mathcal{X}_L$. By the MVES definition (see Definition 1), we have

$$\begin{aligned} &\mathcal{X}_L \subseteq \mathcal{T}_H, \quad \mathcal{T}_H \subseteq \mathrm{aff}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}, \\ &\mathrm{vol}(\mathcal{T}_H) \leq \mathrm{vol}(\mathcal{T}_a). \end{aligned} \tag{9}$$

Recall that $[\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\ ]$ is assumed to have full row rank and satisfy $\mathbf{1}^T\boldsymbol{s}_n = 1$ for all n. From these assumptions, one can prove that $\mathrm{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\} = \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$, and $\mathrm{aff}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\} = \mathrm{aff}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\}$; see [25, Lemma 1] for example. Then, by applying Lemma 1.(b) to (9), we obtain

$$\begin{aligned} &\mathcal{S}_L \subseteq f^{-1}(\mathcal{T}_H), \quad f^{-1}(\mathcal{T}_H) \subseteq \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}, \\ &\mathrm{vol}(f^{-1}(\mathcal{T}_H)) \leq \mathrm{vol}(f^{-1}(\mathcal{T}_a)) = \mathrm{vol}(\mathcal{T}_e). \end{aligned}$$

The above equation implies that $\mathcal{T}_e$ is not the only MVES of $\mathcal{S}_L$, which is a contradiction.

On the other hand, suppose that $\mathsf{MVES}(\mathcal{X}_L) = \{\mathcal{T}_a\}$, but $\mathsf{MVES}(\mathcal{S}_L) \neq \{\mathcal{T}_e\}$. This statement can be shown to be a contradiction, by the same proof as above (particularly, the incorporation of Lemma 1.(a)). The proof of Proposition 1 is therefore complete.

4.2 Proof of Theorem 1

The proof is done by contradiction. Suppose that $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$, but $\rho \leq \frac{1}{\sqrt{N-1}}$. Recall

$$\mathcal{R}(r) = \mathcal{T}_e \cap \{\boldsymbol{s} \in \mathbb{R}^N \mid \|\boldsymbol{s}\| \leq r\}. \tag{10}$$

The proof is divided into four steps.

Step 1: We show that any $\mathcal{V} \in \mathsf{MVES}(\mathcal{R}(\rho))$ is also an MVES of $\mathcal{S}_L$. To prove it, note that

$$\mathcal{S}_L \subseteq \mathcal{R}(\rho). \tag{11}$$

Eq. (11) implies that

$$\mathrm{vol}(\mathcal{U}) \leq \mathrm{vol}(\mathcal{V}), \text{ for all } \mathcal{U} \in \mathsf{MVES}(\mathcal{S}_L),\ \mathcal{V} \in \mathsf{MVES}(\mathcal{R}(\rho)). \tag{12}$$

Also, since $\mathcal{T}_e$ encloses $\mathcal{R}(\rho)$, we have

$$\mathrm{vol}(\mathcal{V}) \leq \mathrm{vol}(\mathcal{T}_e), \text{ for all } \mathcal{V} \in \mathsf{MVES}(\mathcal{R}(\rho)). \tag{13}$$

Since we assume $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$ in the beginning, we observe from (12) and (13) that $\mathrm{vol}(\mathcal{U}) = \mathrm{vol}(\mathcal{V})$ for all $\mathcal{U} \in \mathsf{MVES}(\mathcal{S}_L)$, $\mathcal{V} \in \mathsf{MVES}(\mathcal{R}(\rho))$. The above equality, together with (11), implies that any $\mathcal{V} \in \mathsf{MVES}(\mathcal{R}(\rho))$ is an MVES of $\mathcal{S}_L$ (or satisfies $\mathcal{V} \in \mathsf{MVES}(\mathcal{S}_L)$).

Step 2: We give an alternative representation of $(N-1)$-dimensional simplices on $\mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$, which will facilitate the proof. The affine hull $\mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$ can be equivalently expressed as

$$\mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\} = \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathbb{R}^{N-1}\}, \tag{14}$$

where

$$\boldsymbol{d} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{e}_i = \frac{1}{N}\mathbf{1},$$

and $\boldsymbol{C} \in \mathbb{R}^{N\times(N-1)}$ is the first $N-1$ principal left singular vectors of $\boldsymbol{R} = [\ \boldsymbol{e}_1 - \boldsymbol{d}, \ldots, \boldsymbol{e}_N - \boldsymbol{d}\ ]$; see [6, 25]. We note that

$$\boldsymbol{R} = \boldsymbol{I} - \frac{1}{N}\mathbf{1}\mathbf{1}^T,$$

which, as a standard matrix result, its first $N-1$ principal left singular vector can be shown to be any $\boldsymbol{C}$ such that

$$\boldsymbol{U} = \left[\ \boldsymbol{C},\ \frac{1}{\sqrt{N}}\mathbf{1}\ \right] \tag{15}$$

is a unitary matrix. Or, equivalently, $\boldsymbol{C}$ is any semi-unitary matrix such that $\boldsymbol{C}^T\boldsymbol{d} = \mathbf{0}$.

Recall that an $(N-1)$-dimensional simplex $\mathcal{V} \subseteq \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$ can be written as

$$\mathcal{V} = \mathrm{conv}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N\},$$

where $\boldsymbol{v}_i \in \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$ for all i. By (14), each $\boldsymbol{v}_i \in \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$ can be represented by $\boldsymbol{v}_i = \boldsymbol{C}\boldsymbol{w}_i + \boldsymbol{d}$ for some $\boldsymbol{w}_i \in \mathbb{R}^{N-1}$. Applying this result to $\mathrm{conv}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_N\}$, we obtain the following equivalent representation of $\mathcal{V}$

$$\mathcal{V} = \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{W}\}, \tag{16}$$

where

$$\mathcal{W} = \mathrm{conv}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N\}. \tag{17}$$

Also, by the simplex volume formula (1) and the semi-unitarity of $\boldsymbol{C}$, the following relation is shown

$$\mathrm{vol}(\mathcal{V}) = \mathrm{vol}(\mathcal{W}). \tag{18}$$

Step 3: We show that there are infinitely many MVES of $\mathcal{R}(\rho)$ for $\frac{1}{\sqrt{N}} < \rho \leq \frac{1}{\sqrt{N-1}}$. Consider the following lemma.

Lemma 2 *Let*

$$\mathcal{C}(r) = \text{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\} \cap \{\boldsymbol{s} \in \mathbb{R}^N \mid \|\boldsymbol{s}\| \leq r\}. \tag{19}$$

denote a 2-norm ball on $\text{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$. *If* $\frac{1}{\sqrt{N}} < r \leq \frac{1}{\sqrt{N-1}}$, *then* $\mathcal{R}(r)$ *in* (10) *equals* $\mathcal{C}(r)$.

Proof of Lemma 2: Note that $\mathcal{R}(r) \subseteq \mathcal{C}(r)$. Hence, to prove Lemma 2, it suffices to show that $\mathcal{C}(r) \subseteq \mathcal{R}(r)$. By the equivalent affine hull representation in (14), we can write $\mathcal{C}(r) = \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \|\boldsymbol{s}\| \leq r\}$. By substituting $\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d}$ into $\|\boldsymbol{s}\| \leq r$, we get, for any $\boldsymbol{s} \in \mathcal{C}(r)$,

$$\|\boldsymbol{s}\|^2 \leq r^2 \Longleftrightarrow \|\boldsymbol{\theta}\|^2 + \|\boldsymbol{d}\|^2 \leq r^2 \tag{20a}$$

$$\Longleftrightarrow \|\boldsymbol{\theta}\|^2 \leq r^2 - \frac{1}{N}, \tag{20b}$$

where (20a) is obtained by using the orthogonality in (15); (20b) is by $\|\boldsymbol{d}\|^2 = \frac{1}{N}$. Hence, $\mathcal{C}(r)$ can be rewritten as

$$\mathcal{C}(r) = \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \|\boldsymbol{\theta}\|^2 \leq r^2 - 1/N\}. \tag{21}$$

Moreover, by letting $\boldsymbol{c}^i$ and $\boldsymbol{u}^i$ denote the ith rows of $\boldsymbol{C}$ and $\boldsymbol{U}$ respectively, we have

$$s_i = [\boldsymbol{c}^i]^T\boldsymbol{\theta} + d_i \tag{22a}$$

$$\geq -\|\boldsymbol{c}^i\|\|\boldsymbol{\theta}\| + \frac{1}{N} \tag{22b}$$

$$\geq -\sqrt{\frac{N-1}{N}} \cdot \sqrt{\frac{1}{(N-1)\cdot N}} + \frac{1}{N} = 0, \tag{22c}$$

where (22b) is due to the Cauchy-Schwartz inequality; (22c) is due to (20b), $r \leq \frac{1}{\sqrt{N-1}}$, and the fact that $1 = \|\boldsymbol{u}^i\|^2 = \frac{1}{N} + \|\boldsymbol{c}^i\|^2$ (see (15) and note its orthogonality). Eq. (22) suggests that any $\boldsymbol{s} \in \mathcal{C}(r)$ automatically satisfies $\boldsymbol{s} \geq \boldsymbol{0}$, and hence, $\boldsymbol{s} \in \mathcal{R}(r)$. We therefore conclude that $\mathcal{C}(r) = \mathcal{R}(r)$. ■

By Lemma 2, we can replace $\mathcal{R}(\rho)$ by $\mathcal{C}(\rho)$ and consider the MVES of the latter. Suppose that $\mathcal{V} \in \mathsf{MVES}(\mathcal{C}(\rho))$. Our argument is that a suitably rotated version of $\mathcal{V}$ is also an MVES of $\mathcal{C}(\rho)$. To be precise, use the representation in (16)-(17) to describe $\mathcal{V}$. Comparing (16)-(17) and (21), we see that $\mathcal{C}(\rho) \subseteq \mathcal{V}$ is equivalent to

$$\{\boldsymbol{\theta} \mid \|\boldsymbol{\theta}\|^2 \leq \rho^2 - 1/N\} \subseteq \mathcal{W}. \tag{23}$$

From $\mathcal{W}$, let us construct another simplex

$$\mathcal{V}' = \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{Q}\boldsymbol{\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{W}\}, \tag{24}$$

where $\boldsymbol{Q} \in \mathbb{R}^{(N-1)\times(N-1)}$ is a unitary matrix. Due to (23), $\mathcal{V}'$ can be verified to satisfy $\mathcal{C}(\rho) \subseteq \mathcal{V}'$. Also, by observing the semi-unitarity of $\boldsymbol{C}\boldsymbol{Q}$, the volume of $\mathcal{V}'$ is shown to equal

$$\text{vol}(\mathcal{V}') = \text{vol}(\mathcal{W}) = \text{vol}(\mathcal{V}).$$

In other words, $\mathcal{V}'$ is also an MVES of $\mathcal{C}(\rho)$. In fact, the argument above holds for any unitary $\boldsymbol{Q}$. Since there are infinitely many unitary $\boldsymbol{Q}$ for $N \geq 3$ (note that $\boldsymbol{Q} \in \mathbb{R}^{(N-1)\times(N-1)}$), we also have infinitely many MVESs of $\mathcal{C}(\rho)$ for $N \geq 3$.

Step 4: We combine the results in the above steps to draw conclusion. Step 1 shows that any $\mathcal{V} \in \mathsf{MVES}(\mathcal{R}(\rho))$ is also an MVES of $\mathcal{S}_L$, while Step 3 shows that $\mathcal{R}(\rho)$ has infinitely many MVESs for $\rho \leq \frac{1}{\sqrt{N-1}}$, $N \geq 3$. This contradicts the assumption that there is only one MVES of $\mathcal{S}_L$. The proof of Theorem 1 is therefore complete.

4.3 Proof of Theorem 2

To facilitate our proof, let us introduce the following fact.

Fact 1 *Let $\mathcal{C}, \mathcal{D} \subseteq \mathbb{R}^n$ be two sets of identical dimension, with $\mathcal{C} \subseteq \mathcal{D}$. If $\mathcal{D} \subseteq \mathcal{T}$ for some $\mathcal{T} \in \mathsf{MVES}(\mathcal{C})$, then $\mathcal{T} \in \mathsf{MVES}(\mathcal{D})$ and $\mathsf{MVES}(\mathcal{D}) \subseteq \mathsf{MVES}(\mathcal{C})$.*

Proof of Fact 1: Note that $\mathcal{C} \subseteq \mathcal{D}$ implies that any $\mathcal{T}' \in \mathsf{MVES}(\mathcal{D})$ is a simplex enclosing $\mathcal{C}$. Since $\mathcal{T}$ is a minimum volume simplex among all the $\mathcal{C}$-enclosing simplices, we have

$$\mathrm{vol}(\mathcal{T}) \leq \mathrm{vol}(\mathcal{T}') \text{ for all } \mathcal{T}' \in \mathsf{MVES}(\mathcal{D}). \tag{25}$$

Moreover, the condition $\mathcal{D} \subseteq \mathcal{T}$ implies that $\mathcal{T}$ is also a $\mathcal{D}$-enclosing simplex, and, as a result, equality in (25) holds. It also follows that any $\mathcal{T}' \in \mathsf{MVES}(\mathcal{D})$ is also an MVES of $\mathcal{C}$. ■

Now we proceed with the main proof.

Step 1: We show that

$$\mathcal{T}_e \in \mathsf{MVES}(\mathcal{R}(r)), \text{ for any } r \geq \frac{1}{\sqrt{N-1}}. \tag{26}$$

Note from the definition of $\mathcal{R}(r)$ in (6) that

$$\mathcal{C}\left(\tfrac{1}{\sqrt{N-1}}\right) = \mathcal{R}\left(\tfrac{1}{\sqrt{N-1}}\right) \subseteq \mathcal{R}(r) \subseteq \mathcal{T}_e, \tag{27}$$

for any $r \in [1/\sqrt{N-1}, 1]$, where the first equality is by Lemma 2. We prove that

Lemma 3 *The unit simplex $\mathcal{T}_e$ is an MVES of $\mathcal{C}(1/\sqrt{N-1})$.*

The proof of Lemma 3 is relegated to Appendix B. By applying Fact 1 and Lemma 3 to (27), we obtain $\mathcal{T}_e \in \mathsf{MVES}(\mathcal{R}(r))$ for $r \in [1/\sqrt{N-1}, 1]$.

Step 2: We prove that

$$\mathsf{MVES}(\mathcal{S}_L) \subseteq \mathsf{MVES}(\mathcal{R}(\gamma)), \text{ for } \gamma \geq \frac{1}{\sqrt{N-1}}. \tag{28}$$

By the definition of γ in (5), we have

$$\mathcal{R}(\gamma) \subseteq \mathrm{conv}\mathcal{S}_L \subseteq \mathcal{T}_e. \tag{29}$$

Also, in Step 1, it has been identified that $\mathcal{T}_e \in \mathsf{MVES}(\mathcal{R}(r))$ for $r \in [1/\sqrt{N-1}, 1]$. Hence, for $\gamma \geq 1/\sqrt{N-1}$, we can apply Fact 1 to (29) to obtain

$$\mathsf{MVES}(\mathrm{conv}\mathcal{S}_L) \subseteq \mathsf{MVES}(\mathcal{R}(\gamma)). \tag{30}$$

Next, we use a straightforward fact in convex analysis: for a convex set $\mathcal{T}$, the condition $\mathcal{C} \subset \mathcal{T}$ is the same as $\mathrm{conv}\mathcal{C} \subset \mathcal{T}$, and vice versa. In the context here, this implies that any MVES of $\mathrm{conv}\mathcal{S}_L$ also encloses $\mathcal{S}_L$, and the converse is also true. Hence, we have

$$\mathsf{MVES}(\mathrm{conv}\mathcal{S}_L) = \mathsf{MVES}(\mathcal{S}_L). \tag{31}$$

By combining (30) and (31), Eq. (28) is obtained.

Step 3: We prove that

$$\mathsf{MVES}(\mathcal{R}(\gamma)) = \{\mathcal{T}_e\}, \text{ for } \gamma > \frac{1}{\sqrt{N-1}}. \tag{32}$$

It has been shown in Step 1 that $\mathcal{T}_e \in \mathsf{MVES}(\mathcal{R}(\gamma))$. The question is whether there exists another MVES $\mathcal{T}' \in \mathsf{MVES}(\mathcal{R}(\gamma))$, with $\mathcal{T}' \neq \mathcal{T}_e$. It turns out that such a $\mathcal{T}'$ does not exist, as we prove in the following theorem.

Theorem 3 *For any $1/\sqrt{N-1} < r \leq 1$, we have $\mathsf{MVES}(\mathcal{R}(r)) = \{\mathcal{T}_e\}$; i.e., there is only one MVES of $\mathcal{R}(r)$ for $1/\sqrt{N-1} < r \leq 1$ and that MVES is always given by the unit simplex.*

Note that Theorem 3 shows a unique MVES condition for $\mathcal{R}(r)$, which is interesting in its own right and could be of independent interest in other fields. The proof of Theorem 3 is shown in the next subsection. By Theorem 3, (32) is obtained.

Step 4: We combine the results in Steps 2 and 3. Specifically, by (28) and (32), we get $\mathsf{MVES}(\mathcal{S}_L) \subseteq \{\mathcal{T}_e\}$. As $\mathcal{S}_L$ is enclosed by $\mathcal{T}_e$, we further deduce $\mathsf{MVES}(\mathcal{S}_L) = \{\mathcal{T}_e\}$. Theorem 2 is therefore proven.

4.4 Proof of Theorem 3

Let $\mathcal{T}' \in \mathsf{MVES}(\mathcal{R}(r))$ be an arbitrary MVES of $\mathcal{R}(r)$ for $1/\sqrt{N-1} < r \leq 1$. We prove Theorem 3 by showing that $\mathcal{T}' = \mathcal{T}_e$ is always true. The proof is divided into three steps.

Step 1: We show that

$$\mathcal{T}' \in \mathsf{MVES}(\mathcal{R}(1/\sqrt{N-1})).$$

To prove this, note that $\mathcal{R}(1/\sqrt{N-1}) \subseteq \mathcal{R}(r)$ for all $1/\sqrt{N-1} \leq r \leq 1$. Also, it has been shown in (26) that $\mathcal{T}_e \in \mathsf{MVES}(\mathcal{R}(r))$ for all $1/\sqrt{N-1} \leq r \leq 1$. Applying Fact 1 to the above two results yields

$$\mathsf{MVES}(\mathcal{R}(r)) \subseteq \mathsf{MVES}(\mathcal{R}(1/\sqrt{N-1})),$$

for all $1/\sqrt{N-1} \leq r \leq 1$. Since $\mathcal{T}' \in \mathsf{MVES}(\mathcal{R}(r))$ for $1/\sqrt{N-1} < r \leq 1$, it follows that $\mathcal{T}' \in \mathsf{MVES}(\mathcal{R}(1/\sqrt{N-1}))$ is also true.

Step 2: To proceed further, we apply the equivalent representation in (16)-(17) to rewrite $\mathcal{T}_e$ as

$$\mathcal{T}_e = \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{W}_e\} \tag{33}$$

for some $(N-1)$-dimensional simplex $\mathcal{W}_e \subseteq \mathbb{R}^{N-1}$. Similarly, we can characterize $\mathcal{T}'$ by

$$\mathcal{T}' = \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{W}'\} \tag{34}$$

for some $(N-1)$-dimensional simplex $\mathcal{W}' \subseteq \mathbb{R}^{N-1}$. Also, by noting $\mathcal{R}(r) = \mathcal{T}_e \cap \mathcal{C}(r)$, the expression of $\mathcal{C}(r)$ in (21), and $\mathcal{R}(r) = \mathcal{C}(r)$ for $r = 1/\sqrt{N-1}$ (see Lemma 2), $\mathcal{R}(r)$ can be expressed as

$$\mathcal{R}(r) = \begin{cases} \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{B}(\sqrt{r^2 - 1/N})\}, & r = \frac{1}{\sqrt{N-1}} \\ \{\boldsymbol{s} = \boldsymbol{C}\boldsymbol{\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{W}_e \cap \mathcal{B}(\sqrt{r^2 - 1/N})\}, & r > \frac{1}{\sqrt{N-1}} \end{cases} \tag{35}$$

where

$$\mathcal{B}(r) = \{\boldsymbol{\theta} \in \mathbb{R}^{N-1} \mid \|\boldsymbol{\theta}\| \leq r\}. \tag{36}$$

Now, by comparing (34)-(35), the following result can be proven:

$$\mathcal{T}' \in \mathsf{MVES}(\mathcal{R}(r)) \Longleftrightarrow \mathcal{W}' \in \begin{cases} \mathsf{MVES}\left(\mathcal{B}(\sqrt{r^2-1/N})\right), & r = \frac{1}{\sqrt{N-1}} \\ \mathsf{MVES}\left(\mathcal{W}_e \cap \mathcal{B}(\sqrt{r^2-1/N})\right), & r > \frac{1}{\sqrt{N-1}} \end{cases} \tag{37}$$

The proof of (37) is analogous to that of Proposition 1, and will not be repeated here.

Step 3: From the equivalent representation (37), we further deduce the following results: i) $\mathcal{W}_e, \mathcal{W}' \in \mathsf{MVES}(\mathcal{B}(\sqrt{r^2-1/N}))$ for $r = 1/\sqrt{N-1}$, which is due to Step 1 and (26); ii) $\mathcal{W}_e \cap \mathcal{B}(\sqrt{r^2-1/N}) \subseteq \mathcal{W}'$ for all $r > 1/\sqrt{N-1}$, which is due to the underlying assumption that $\mathcal{T}' \in \mathsf{MVES}(\mathcal{R}(r))$ for $1/\sqrt{N-1} < r \leq 1$. Consider the following lemma:

Lemma 4 *Suppose that $\mathcal{W}, \mathcal{W}' \in \mathsf{MVES}(\mathcal{B}(r))$, where $\mathcal{B}(r)$ is defined in* (36)*. Also, suppose that $\mathcal{R} = \mathcal{W} \cap \mathcal{B}(\bar{r}) \subseteq \mathcal{W}'$ for some $\bar{r} > r > 0$. Then we have $\mathcal{W} = \mathcal{W}'$.*

The proof of Lemma 4 is relegated to Appendix C. By Lemma 4, we obtain $\mathcal{W}_e = \mathcal{W}'$, and consequently, $\mathcal{T}_e = \mathcal{T}'$.

4.5 Proof of Proposition 2

First, we consider an alternative representation of $\{\boldsymbol{s}_n\}$ for $N = 2$. It can be shown from (14) that

$$\mathrm{aff}\{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \{\boldsymbol{s} = \boldsymbol{c}\theta + \boldsymbol{d} \mid \theta \in \mathbb{R}\}, \tag{38}$$

where $\boldsymbol{d} = \frac{1}{2}(\boldsymbol{e}_1 + \boldsymbol{e}_2) = \frac{1}{2}[\ 1\ 1\ ]^T$, and $\boldsymbol{c} = \frac{1}{\sqrt{2}}(\boldsymbol{e}_1 - \boldsymbol{e}_2) = \frac{1}{\sqrt{2}}[\ 1\ \ -1\ ]^T$. Using the facts that $\boldsymbol{s}_n \in \mathrm{aff}\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ and $\boldsymbol{s}_n \geq \boldsymbol{0}$, we obtain the alternative expression

$$\boldsymbol{s}_n = \boldsymbol{c}\theta_n + \boldsymbol{d}, \tag{39}$$

for all n, where each θ_n must satisfy

$$-\frac{1}{\sqrt{2}} \leq \theta_n \leq \frac{1}{\sqrt{2}}. \tag{40}$$

Also, $\theta_n = 1/\sqrt{2}$ and $\theta_n = -1/\sqrt{2}$ can be shown to be equivalent to $\boldsymbol{s}_n = \boldsymbol{e}_1$ and $\boldsymbol{s}_n = \boldsymbol{e}_2$, resp.

Second, we consider the MVES. Let $\mathrm{conv}\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ be an arbitrary MVES of $\mathcal{S}_L$. Since $\boldsymbol{b}_1, \boldsymbol{b}_2 \in \mathrm{aff}\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ must hold, we get

$$\boldsymbol{b}_1 = \boldsymbol{c}\alpha_1 + \boldsymbol{d}, \quad \boldsymbol{b}_2 = \boldsymbol{c}\alpha_2 + \boldsymbol{d} \tag{41}$$

for some α_1, α_2. Let us assume $\alpha_2 \geq \alpha_1$ without loss of generality. Since $\mathrm{conv}\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ encloses all $\boldsymbol{s}_n$'s, we see from (39)-(41) that

$$\alpha_1 \leq \theta_n \leq \alpha_2, \quad \text{for all } n. \tag{42}$$

Also, the volume of $\mathrm{conv}\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ is

$$\mathrm{vol}(\mathrm{conv}\{\boldsymbol{b}_1, \boldsymbol{b}_2\}) = \alpha_2 - \alpha_1, \tag{43}$$

which is obtained by applying (41) to the volume formula in (1). Since $\mathrm{conv}\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ has the minimum volume among all enclosing simplices, both $-\alpha_1$ and α_2 must be minimized subject to the constraints in (42); the solution is uniquely given by

$$\alpha_1 = \min_{n=1,\ldots,L} \theta_n, \quad \alpha_2 = \max_{n=1,\ldots,L} \theta_n. \tag{44}$$

The above equations also suggest that $\mathcal{S}_L$ has a unique MVES.

Third, we compare the volumes of $\text{conv}\{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ and $\mathcal{T}_e$. It can be shown that $\text{vol}(\mathcal{T}_e) = \sqrt{2}$ (for $N = 2$). On the other hand, from (43), (44) and (40) we get $\text{vol}(\text{conv}\{\boldsymbol{b}_1, \boldsymbol{b}_2\}) \leq \sqrt{2}$, and equality holds if and only if $\alpha_1 = -1/\sqrt{2}$ and $\alpha_2 = 1/\sqrt{2}$—which suggests the pure-pixel assumption. We therefore conclude that $\mathcal{T}_e$ is the only MVES of $\mathcal{S}_L$ if and only if the pure-pixel assumption holds.

5 Numerical Experiments

In this section, we provide numerical simulation results that aim to verify the theoretical MVES identifiability results proven in the previous section. The signals are generated by the following way. The observed data set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_L\}$ follows the basic model in (2). The endmember signature vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N$ are selected from the U.S. geological survey (USGS) library [26], and the number of spectral bands is $M = 224$. The generation of the abundance vectors is similar to that in [6]. Specifically, we generate a large pool of random vectors following a Dirichlet distribution with parameter $\boldsymbol{\mu} = \frac{1}{N}\mathbf{1}$, and then select a number of L such random vectors as the abundance set $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\}$. During the selection, we do not choose vectors whose 2-norm exceeds a given parameter r; the reason of doing so is to allow us to control the pixel purity level of $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_L\}$ at or below r in the simulations. Note that if the number of pixels L is large, then one should expect that r be close to the best pixel purity level ρ and uniform pixel purity level γ. In the simulations, we set $L = 1,000$.

The simulation settings are as follows. MVES is implemented by the alternating linear programming method in [6]. We measure its identification performance by using the root-mean-square (RMS) angle error

$$\phi = \min_{\boldsymbol{\pi} \in \Pi_N} \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\arccos\left(\frac{\boldsymbol{a}_i^T \hat{\boldsymbol{a}}_{\pi_i}}{\|\boldsymbol{a}_i\| \cdot \|\hat{\boldsymbol{a}}_{\pi_i}\|}\right)\right]^2},$$

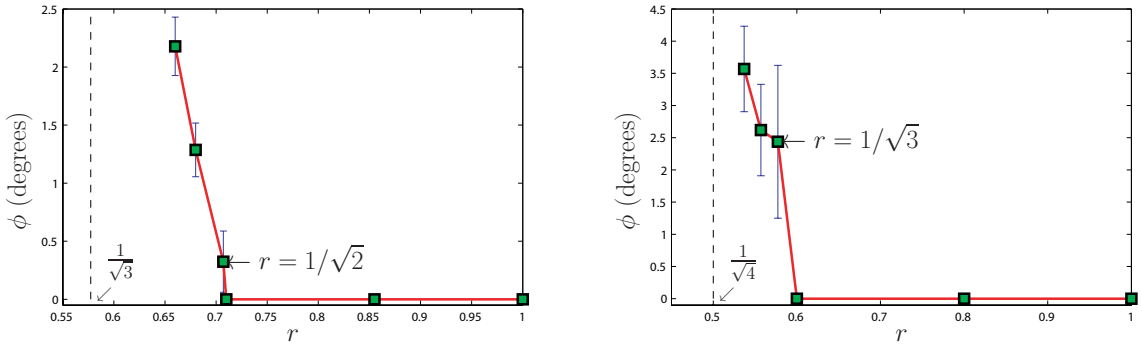
where $\{\hat{\boldsymbol{a}}_1, \ldots, \hat{\boldsymbol{a}}_N\}$ denotes the MVES endmember estimate, and Π_N denotes the set of all permutations of $\{1, \ldots, N\}$. A number of 50 randomly generated realizations were run to evaluate the means and standard deviations of ϕ.

The obtained RMS angle error results are shown in Figure 4. We see that zero RMS angle error, or equivalently, perfect identifiability, is attained when $r > 1/\sqrt{N-1}$ — which is a good match with the sufficient MVES identifiability result in Theorem 2. Also, we observe non-zero errors for $r \leq 1/\sqrt{N-1}$, which verifies the necessary MVES identifiability result in Theorem 1.

Before closing this experiment section, we should mention that previous papers, such as [6, 15, 17–21], have together provided a nice and rather complete coverage on MVES's performance under both synthetic and real-data experiments. Hence, readers are referred to such papers for more experimental results. The results reported therein also indicate that MVES-based algorithms are robust against lack of pure pixels. The numerical (and also theoretical) results above further show the limit of robustness—$1/\sqrt{N-1}$ with the uniform pixel purity level.

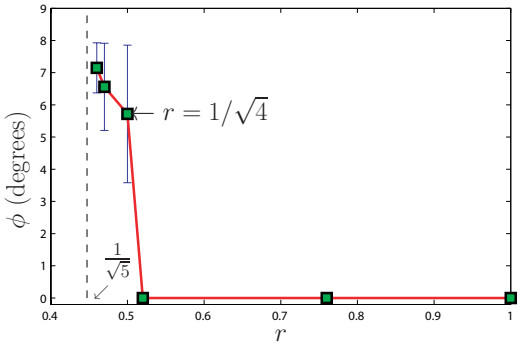
6 Conclusion and Discussion

In this paper, a theoretical analysis for the identifiablility of MVES in blind HU was performed. The results suggest that under some mild assumptions which are considerably more relaxed than

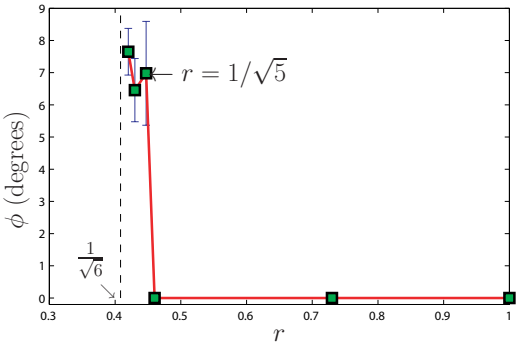


(a) $N = 3$

(b) $N = 4$



(c) $N = 5$



(d) $N = 6$

Figure 4: MVES performance with respect to the numerically control pixel purity level r.

those for the pure-pixel case, MVES exhibits robustness against lack of pure pixels. Hence, our study provides a theoretical explanation on why numerical studies usually found that MVES can recover the endmembers accurately in the no pure-pixel case. As future research, it would be interesting to investigate implications of the identifiability analysis on efficient algorithm designs, and extensions to even more challenging analysis scenarios such as noise sensitivity analysis.

Appendix

A Proof of Lemma 1

Let us first prove Lemma 1.(a). The set $\mathcal{T}_G$ can be explicitly represented by

$$\mathcal{T}_G = \mathrm{conv}\{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_N\},$$

where $\boldsymbol{g}_i \in \mathbb{R}^N$ for all i. Also, by letting $\boldsymbol{h}_i = \boldsymbol{A}\boldsymbol{g}_i$ for all i, one can easily show that

$$f(\mathcal{T}_G) = \mathrm{conv}\{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N\}.$$

Since $\mathcal{T}_G \subset \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$, we have $\boldsymbol{g}_i \in \mathrm{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$ for all i. This means that each $\boldsymbol{g}_i$ satisfies $\boldsymbol{1}^T\boldsymbol{g}_i = 1$, or equivalently, $g_{i,N} = 1 - \sum_{j=1}^{N-1} g_{i,j}$. Using the above fact, we can write

$$\boldsymbol{g}_i = \boldsymbol{C}\boldsymbol{\theta}_i + \boldsymbol{e}_N,$$

where $\boldsymbol{\theta}_i = [\boldsymbol{g}_i]_{1:(N-1)}$, and

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{I} \\ -\boldsymbol{1}^T \end{bmatrix} \in \mathbb{R}^{N\times(N-1)}.$$

Let $\bar{\boldsymbol{G}} = [\ \boldsymbol{g}_1 - \boldsymbol{g}_N, \ldots, \boldsymbol{g}_{N-1} - \boldsymbol{g}_N\ ]$. We get

$$\bar{\boldsymbol{G}} = \boldsymbol{C}\bar{\boldsymbol{\Theta}},$$

where $\bar{\boldsymbol{\Theta}} = [\ \boldsymbol{\theta}_1 - \boldsymbol{\theta}_N, \ldots, \boldsymbol{\theta}_{N-1} - \boldsymbol{\theta}_N\ ] \in \mathbb{R}^{(N-1)\times(N-1)}$. We therefore obtain

$$\begin{aligned} \det(\bar{\boldsymbol{G}}^T\bar{\boldsymbol{G}}) &= \det(\bar{\boldsymbol{\Theta}}^T\boldsymbol{C}^T\boldsymbol{C}\bar{\boldsymbol{\Theta}}) && \text{(45a)} \\ &= \det(\bar{\boldsymbol{\Theta}})\det(\boldsymbol{C}^T\boldsymbol{C})\det(\bar{\boldsymbol{\Theta}}) && \text{(45b)} \\ &= N\cdot|\det(\bar{\boldsymbol{\Theta}})|^2, && \text{(45c)} \end{aligned}$$

where (45b) is due to $\det(\boldsymbol{A}\boldsymbol{B}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$ for square $\boldsymbol{A}, \boldsymbol{B}$, and (45c) is due to the following result

$$\det(\boldsymbol{C}^T\boldsymbol{C}) = \det(\boldsymbol{I} + \boldsymbol{1}\boldsymbol{1}^T) = N$$

(note that the matrix result $\det(\boldsymbol{I} + \boldsymbol{q}\boldsymbol{q}^T) = \|\boldsymbol{q}\|^2 + 1$ has been used). Likewise, by letting $\bar{\boldsymbol{H}} = [\ \boldsymbol{h}_1 - \boldsymbol{h}_N, \ldots, \boldsymbol{h}_{N-1} - \boldsymbol{h}_N\ ]$, we have

$$\bar{\boldsymbol{H}} = \boldsymbol{A}\bar{\boldsymbol{G}} = \boldsymbol{A}\boldsymbol{C}\bar{\boldsymbol{\Theta}} = \bar{\boldsymbol{A}}\bar{\boldsymbol{\Theta}},$$

and

$$\det(\bar{\boldsymbol{H}}^T\bar{\boldsymbol{H}}) = \det(\bar{\boldsymbol{A}}^T\bar{\boldsymbol{A}})\cdot|\det(\bar{\boldsymbol{\Theta}})|^2. \quad (46)$$

Now, by (1), (45) and (46), Eq. (7) is obtained. Also, the property $f(\mathcal{T}_G) \subset \text{aff}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\}$ can be easily proven by the fact that $\boldsymbol{H} = \boldsymbol{AG}$ and $\mathbf{1}^T \boldsymbol{g}_i = 1$ for all i.

Next, we prove Lemma 1.(b). The set $\mathcal{T}_H$ can be written as

$$\mathcal{T}_H = \text{conv}\{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N\},$$

where $\boldsymbol{h}_i \in \mathbb{R}^M$ for all i. Since $\mathcal{T}_H \subset \text{aff}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\}$, we have $\boldsymbol{h}_i \in \text{aff}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N\}$ for all i. Hence, each $\boldsymbol{h}_i$ can be expressed as $\boldsymbol{h}_i = \boldsymbol{A}\boldsymbol{g}_i$, where $\boldsymbol{g}_i \in \mathbb{R}^N$, $\mathbf{1}^T \boldsymbol{g}_i = 1$. This leads to

$$\begin{aligned}
f^{-1}(\mathcal{T}_H) &= \{\ \boldsymbol{x} \mid \boldsymbol{Ax} \in \text{conv}\{\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N\}\ \} && \text{(47a)}\\
&= \{\ \boldsymbol{x} \mid \boldsymbol{Ax} = \boldsymbol{H\theta},\ \boldsymbol{\theta} \geq \mathbf{0}, \mathbf{1}^T\boldsymbol{\theta} = 1\ \} && \text{(47b)}\\
&= \{\ \boldsymbol{x} \mid \boldsymbol{Ax} = \boldsymbol{AG\theta},\ \boldsymbol{\theta} \geq \mathbf{0}, \mathbf{1}^T\boldsymbol{\theta} = 1\ \} && \text{(47c)}\\
&= \{\ \boldsymbol{x} \mid \boldsymbol{x} = \boldsymbol{G\theta},\ \boldsymbol{\theta} \geq \mathbf{0}, \mathbf{1}^T\boldsymbol{\theta} = 1\ \} && \text{(47d)}\\
&= \text{conv}\{\boldsymbol{g}_1, \ldots, \boldsymbol{g}_N\} && \text{(47e)}\\
&\subset \text{aff}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}, && \text{(47f)}
\end{aligned}$$

where (47d) is due to the full column rank condition of $\boldsymbol{A}$, and (47f) uses the structure $\mathbf{1}^T\boldsymbol{g}_i = 1$. The rest of the proof is the same as that of Lemma 1.(a).

B Proof of Lemma 3

Fix $r = 1/\sqrt{N-1}$. From (21), $\mathcal{C}(r)$ can be re-expressed as

$$\mathcal{C}(r) = \{\boldsymbol{s} = \boldsymbol{C\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{B}(\mu)\}, \tag{48}$$

where $\mu = \sqrt{r^2 - 1/N} = 1/\sqrt{(N-1)N}$, and

$$\mathcal{B}(r') = \{\boldsymbol{\theta} \in \mathbb{R}^{N-1} \mid \|\boldsymbol{\theta}\| \leq r'\} \tag{49}$$

is a ball on $\mathbb{R}^{N-1}$. Also, recall from (16)-(17) that an MVES $\mathcal{V} \in \mathsf{MVES}(\mathcal{C}(r))$ can be written as

$$\mathcal{V} = \{\boldsymbol{s} = \boldsymbol{C\theta} + \boldsymbol{d} \mid \boldsymbol{\theta} \in \mathcal{W}\}, \tag{50}$$

where $\mathcal{W} = \text{conv}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N\} \subseteq \mathbb{R}^{N-1}$; and that $\text{vol}(\mathcal{V}) = \text{vol}(\mathcal{W})$ (see (18)). From the expressions above, we can deduce the following result: $\mathcal{W}$ must be an MVES of $\mathcal{B}(\mu)$ if $\mathcal{V}$ is an MVES of $\mathcal{C}(r)$, and the converse is also true.

Next, we will use the following fact:

Fact 2 [27, Theorem 3.2] *The volume of an $(N-1)$-dimensional simplex $\mathcal{W}$ enclosing $B(r')$ in (49) satisfies*

$$\text{vol}(\mathcal{W}) \geq \frac{1}{(N-1)!} N^{\frac{N}{2}} (N-1)^{\frac{1}{2}(N-1)} (r')^{N-1} \tag{51}$$

with equality only for the regular simplex.

Using Fact 2 and the result $\text{vol}(\mathcal{V}) = \text{vol}(\mathcal{W})$, we obtain

$$\text{vol}(\mathcal{V}) = \frac{1}{(N-1)!}\sqrt{N},$$

where we should note that the right-hand side of the above equation is obtained by putting $r' = \mu = 1/\sqrt{(N-1)N}$ into (51). On the other hand, consider $\mathcal{T}_e = \mathrm{conv}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_N\}$, which encloses $\mathcal{C}(r)$ (for $r = 1/\sqrt{N-1}$). From the simplex volume formula (1), one can show that

$$\mathrm{vol}(\mathcal{T}_e) = \frac{1}{(N-1)!}\sqrt{N}.$$

Since $\mathcal{T}_e$ attains the same volume as $\mathcal{V}$, $\mathcal{T}_e$ is an MVES of $\mathcal{C}(r)$.

C Proof of Lemma 4

The following lemma will be required:

Lemma 5 *Let $\mathcal{B}(r) = \{\boldsymbol{\theta} \in \mathbb{R}^{N-1} \mid \|\boldsymbol{\theta}\| \leq r\}$, where $r > 0$. For any $\mathcal{W} \in \mathsf{MVES}(\mathcal{B}(r))$, the boundaries of $\mathcal{B}(r)$ and $\mathcal{W}$ have exactly N intersecting points. Also, by letting $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}$ be the set of those intersecting points, we have the following properties:*

(a) The points $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N$ are affinely independent.

(b) The simplex $\mathcal{W}$ can be constructed from $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N$ via

$$\mathcal{W} = \bigcap_{i=1}^{N} \left\{\boldsymbol{\theta} \in \mathbb{R}^{N-1} \mid r^2 \geq \boldsymbol{t}_i^T \boldsymbol{\theta}\right\}.$$

The proof of Lemma 5 is given in Appendix D. Let

$$\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W},$$
$$\{\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N\} = \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}',$$

which, by Lemma 5, always exist. Since $\mathcal{B}(r) \subset \mathcal{W}$ and $\mathcal{B}(r) \subset \mathcal{W}'$, the above two equations can be equivalently expressed as

$$\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \mathrm{bd}\mathcal{B}(r) \setminus \mathrm{int}\mathcal{W}, \tag{52}$$
$$\{\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N\} = \mathrm{bd}\mathcal{B}(r) \setminus \mathrm{int}\mathcal{W}'. \tag{53}$$

Also, by Lemma 5.(b), we have $\mathcal{W} = \mathcal{W}'$ if $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \{\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N\}$. In the following steps we focus on proving $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \{\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N\}$.

Step 1: We first prove

$$\mathrm{bd}\left(\mathcal{W} \cap \mathcal{B}(\bar{r})\right) \subseteq \mathrm{bd}\mathcal{W} \cup \mathrm{bd}\mathcal{B}(\bar{r}) \tag{54}$$

by contradiction. Suppose that (54) does not hold, namely, there exists an $\boldsymbol{x} \in \mathbb{R}^{N-1}$ satisfying

$$\boldsymbol{x} \in \mathrm{bd}\left(\mathcal{W} \cap \mathcal{B}(\bar{r})\right), \text{ but} \tag{55}$$
$$\boldsymbol{x} \notin \mathrm{bd}\mathcal{W} \cup \mathrm{bd}\mathcal{B}(\bar{r}). \tag{56}$$

Now, since $\mathcal{W} \cap \mathcal{B}(\bar{r})$ is a closed set, (55) implies

$$\boldsymbol{x} \in \mathcal{W} \cap \mathcal{B}(\bar{r}). \tag{57}$$

Equations (56) and (57) imply that $\boldsymbol{x} \in \mathrm{int}\mathcal{W}$ and that $\boldsymbol{x} \in \mathrm{int}\mathcal{B}(\bar{r})$. Thus, we have $\boldsymbol{x} \in \mathrm{int}(\mathcal{W} \cap \mathcal{B}(\bar{r}))$ which contradicts (55). Hence, (54) must hold.

Step 2: We show that $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{R}$. Let us first consider proving $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} \subseteq \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{R}$. We observe from $\mathcal{B}(r) \subseteq \mathcal{B}(\bar{r})$ and $\mathcal{B}(r) \subseteq \mathcal{W}$ that

$$\mathcal{B}(r) \subseteq \mathcal{B}(\bar{r}) \cap \mathcal{W} = \mathcal{R}. \tag{58}$$

Subsequently, the following inequality chain can be derived:

$$\begin{aligned}
\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} =& \mathrm{bd}\mathcal{B}(r) \setminus \mathrm{int}\mathcal{W} && \text{(59a)}\\
\subseteq& \mathrm{bd}\mathcal{B}(r) \setminus (\mathrm{int}\mathcal{W} \cap \mathrm{int}\mathcal{B}(\bar{r})) && \text{(59b)}\\
=& \mathrm{bd}\mathcal{B}(r) \setminus \mathrm{int}\mathcal{R} && \text{(59c)}\\
=& \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{R}, && \text{(59d)}
\end{aligned}$$

where (59a) is by (52); (59c) is by $\mathrm{int}(\mathcal{W} \cap \mathcal{B}(\bar{r})) = \mathrm{int}\mathcal{W} \cap \mathrm{int}\mathcal{B}(\bar{r})$; (59d) is by (58).

Moreover, we have $\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{R} \subseteq \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$, obtained from the following chain:

$$\begin{aligned}
\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{R} =& \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}(\mathcal{W} \cap \mathcal{B}(\bar{r})) && \text{(60a)}\\
\subseteq& \mathrm{bd}\mathcal{B}(r) \cap (\mathrm{bd}\mathcal{W} \cup \mathrm{bd}\mathcal{B}(\bar{r})) && \text{(60b)}\\
=& (\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}) \cup (\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{B}(\bar{r})) && \text{(60c)}\\
=& (\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}) \cup \emptyset && \text{(60d)}\\
=& \mathrm{bd}\mathcal{B}(r) \setminus \mathrm{int}\mathcal{W} && \text{(60e)}\\
=& \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}, && \text{(60f)}
\end{aligned}$$

where (60b) is by (54); (60d) is by $\bar{r} > r$; (60e) is by $\mathrm{bd}\mathcal{B}(r) \subseteq \mathcal{B}(r) \subseteq \mathcal{W}$; (60f) is by (52).

Step 3: We prove $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \{\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N\}$. In Step 2, it is shown that

$$\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\} = \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{R}. \tag{61}$$

By the fact that $\boldsymbol{t}'_i \in \mathcal{B}(r)$ and by (58), we have

$$\boldsymbol{t}'_i \in \mathcal{R}. \tag{62}$$

Moreover, from the assumption that $\mathcal{R} \subseteq \mathcal{W}'$, we have $\mathrm{bd}\mathcal{W}' \cap \mathrm{int}\mathcal{R} = \emptyset$. But from (53), we note that $\boldsymbol{t}'_i \in \mathrm{bd}\mathcal{W}'$. Thus we can conclude $\boldsymbol{t}'_i \notin \mathrm{int}(\mathcal{R})$, which together with (62) yields

$$\boldsymbol{t}'_i \in \mathrm{bd}\mathcal{R}. \tag{63}$$

Combining $\boldsymbol{t}'_i \in \mathrm{bd}\mathcal{B}(r)$ (cf. (53)) with (61) and (63), we obtain $\boldsymbol{t}'_i \in \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$. Since Property (a) in Lemma 5 restricts $\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N$ to be affinely independent, the only possible choice of $\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N$ is $\{\boldsymbol{t}'_1, \ldots, \boldsymbol{t}'_N\} = \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$. Lemma 4 is therefore proven.

D Proof of Lemma 5

The proof of Lemma 5 requires several convex analysis results. To start with, consider the following results:

Fact 3 *Let $\mathcal{W} = \mathrm{conv}\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N\} \subset \mathbb{R}^{N-1}$ denote an $(N-1)$-dimensional simplex. Also, let*

$$\mathcal{P}(\boldsymbol{g}, \boldsymbol{H}) = \{\boldsymbol{\theta} \in \mathbb{R}^{N-1} \mid \boldsymbol{H}^T\boldsymbol{\theta} + \boldsymbol{g} \geq \boldsymbol{0}, \ -(\boldsymbol{H}\boldsymbol{1})^T\boldsymbol{\theta} + (1 - \boldsymbol{1}^T\boldsymbol{g}) \geq 0\} \tag{64}$$

denote a polyhedron, where $(\boldsymbol{g}, \boldsymbol{H}) \in \mathbb{R}^{N-1} \times \mathbb{R}^{(N-1)\times(N-1)}$ is given.

(a) Any $\mathcal{W}$ can be equivalently represented by $\mathcal{P}(\boldsymbol{g}, \boldsymbol{H})$ via setting

$$\boldsymbol{H} = \bar{\boldsymbol{W}}^{-T}, \quad \boldsymbol{g} = -\bar{\boldsymbol{W}}^{-T}\boldsymbol{w}_N, \tag{65}$$

where $\bar{\boldsymbol{W}} = [\ \boldsymbol{w}_1 - \boldsymbol{w}_N, \ldots, \boldsymbol{w}_{N-1} - \boldsymbol{w}_N\ ]$.

(b) Suppose that $\boldsymbol{H}$ has full rank. Under the above restriction, the set $\mathcal{P}(\boldsymbol{g}, \boldsymbol{H})$ for any $(\boldsymbol{g}, \boldsymbol{H})$ can be equivalently represented by $\mathcal{W}$, whose vertices $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$ can be determined by solving the inverse of (65)*. Also, the corresponding volume is*

$$\mathrm{vol}(\mathcal{P}(\boldsymbol{g}, \boldsymbol{H})) = \frac{1}{(N-1)!}|\det(\boldsymbol{H})|^{-1}. \tag{66}$$

The proof of Fact 3 has been shown in the literature [6, 23]. Also, (66) is determined by the simplex volume formula (1) and the relation in (65). From Fact 3, we derive several convex analysis properties for proving Lemma 5.

Fact 4 *Let $\mathcal{W}$ be an $(N-1)$-dimensional simplex on $\mathbb{R}^{N-1}$, and consider the polyhedral representation of $\mathcal{W}$ in* (64)-(65)*. Also, recall the definition $\mathcal{B}(r) = \{\boldsymbol{\theta} \in \mathbb{R}^{N-1} \mid \|\boldsymbol{\theta}\| \leq r\}$.*

(a) If $\mathcal{B}(r) \subseteq \mathcal{W}$, then the following equations hold

$$-r\|\boldsymbol{h}_i\| + g_i \geq 0, \quad i = 1, \ldots, N-1, \tag{67a}$$

$$-r\|\boldsymbol{H}\boldsymbol{1}\| + (1 - \boldsymbol{1}^T\boldsymbol{g}) \geq 0, \tag{67b}$$

where $\boldsymbol{h}_i$ and g_i denote the ith column of $\boldsymbol{H}$ and ith element of $\boldsymbol{g}$, resp. Conversely, if (67) *holds, then $\mathcal{B}(r) \subseteq \mathcal{W}$.*

(b) Suppose $\mathcal{B}(r) \subseteq \mathcal{W}$. The boundaries of $\mathcal{B}(r)$ and $\mathcal{W}$ have at most N intersecting points. Specifically, we have $\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W} \subseteq \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$ where

$$\boldsymbol{t}_i = -\frac{r}{\|\boldsymbol{h}_i\|}\boldsymbol{h}_i, \quad i = 1, \ldots, N-1, \tag{68a}$$

$$\boldsymbol{t}_N = \frac{r}{\|\boldsymbol{H}\boldsymbol{1}\|}\boldsymbol{H}\boldsymbol{1}. \tag{68b}$$

Also, if $\boldsymbol{t}_i \in \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}$, then

$$\begin{cases} -r\|\boldsymbol{h}_i\| + g_i = 0, & i \in \{1, \ldots, N-1\}, \\ -r\|\boldsymbol{H}\boldsymbol{1}\| + (1 - \boldsymbol{1}^T\boldsymbol{g}) = 0, & i = N; \end{cases} \tag{69}$$

otherwise

$$\begin{cases} -r\|\boldsymbol{h}_i\| + g_i > 0, & i \in \{1, \ldots, N-1\}, \\ -r\|\boldsymbol{H}\boldsymbol{1}\| + (1 - \boldsymbol{1}^T\boldsymbol{g}) > 0, & i = N. \end{cases} \tag{70}$$

Proof of Fact 4: The proof of Fact 4.(a) basically follows the development in [23, pp.148-149], and is omitted here for conciseness. To prove Fact 4.(b), observe that a point $\tilde{\boldsymbol{\theta}} \in \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}$ satisfies i) $\|\tilde{\boldsymbol{\theta}}\| = r$; and ii) either

$$\boldsymbol{h}_i^T \tilde{\boldsymbol{\theta}} + g_i = 0, \tag{71}$$

for some $i \in \{1, \ldots, N-1\}$, or

$$-(\boldsymbol{H}\mathbf{1})^T \tilde{\boldsymbol{\theta}} + (1 - \mathbf{1}^T \boldsymbol{g}) = 0. \tag{72}$$

Suppose that $\tilde{\boldsymbol{\theta}}$ satisfies (71). Recall that the assumption $\mathcal{B}(r) \subseteq \mathcal{W}$ implies

$$\boldsymbol{h}_i^T \boldsymbol{\theta} + g_i \geq 0, \quad \text{for all } \|\boldsymbol{\theta}\| \leq r, \tag{73}$$

and that the left-hand side of (73) attains its minimum if and only if $\boldsymbol{\theta} = -(r/\|\boldsymbol{h}_i\|)\boldsymbol{h}_i = \boldsymbol{t}_i$. Thus, if (71) is to be satisfied, then $\tilde{\boldsymbol{\theta}}$ must equal $\boldsymbol{t}_i$, and subsequently (71) becomes

$$-r\|\boldsymbol{h}_i\| + g_i = 0. \tag{74}$$

Likewise, it is shown that if $\tilde{\boldsymbol{\theta}}$ satisfies (72), then $\tilde{\boldsymbol{\theta}} = (r/\|\boldsymbol{H}\mathbf{1}\|)\boldsymbol{H}\mathbf{1} = \boldsymbol{t}_N$ is the only choice and (72) becomes

$$-r\|\boldsymbol{H}\mathbf{1}\| + (1 - \mathbf{1}^T \boldsymbol{g}) = 0. \tag{75}$$

We therefore complete the proof that $\tilde{\boldsymbol{\theta}} \in \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}$ implies $\tilde{\boldsymbol{\theta}} \in \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$.

We should also mention (69)-(70). From the proof above, it is clear that $\boldsymbol{t}_i \in \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}$ holds if and only if (74) holds for $i = 1, \ldots, N-1$, and (75) holds for $i = N$, respectively. By considering (67) as well, we obtain the conditions in (69)-(70). ■

We are now ready to prove Lemma 5. Recall that $\mathcal{W} \in \mathsf{MVES}(\mathcal{B}(r))$ is assumed. By Fact 3.(a), we can write $\mathcal{W} = \mathcal{P}(\boldsymbol{g}, \boldsymbol{H})$ for some $(\boldsymbol{g}, \boldsymbol{H})$, with $\boldsymbol{H}$ being of full rank. Then, by Fact 4.(b), we obtain $\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W} \subseteq \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$. We consider two cases.

Case 1: Suppose that $\boldsymbol{t}_i \notin \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}$ for some $i \in \{1, \ldots, N-1\}$. For simplicity but without loss of generality, assume $i = 1$. By Fact 4.(a)-(b), we have

$$-r\|\boldsymbol{h}_1\| + g_1 > 0, \tag{76a}$$

$$-r\|\boldsymbol{h}_i\| + g_i \geq 0, \quad i = 2, \ldots, N-1, \tag{76b}$$

$$-r\|\boldsymbol{H}\mathbf{1}\| + (1 - \mathbf{1}^T \boldsymbol{g}) \geq 0. \tag{76c}$$

Let us construct another polyhedron, denoted by $\mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})$, where the 2-tuple $(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}}) \in \mathbb{R}^{N-1} \times \mathbb{R}^{(N-1)\times(N-1)}$ is chosen as

$$\tilde{g}_1 = g_1 - N\epsilon, \tag{77a}$$

$$\tilde{g}_i = g_i + \epsilon, \quad i = 2, \ldots, N-1, \tag{77b}$$

$$\tilde{\boldsymbol{H}} = \left(\frac{r+\delta}{r}\right)\boldsymbol{H}, \tag{77c}$$

where

$$\epsilon = \frac{-r\|\boldsymbol{h}_1\| + g_1}{2N} > 0, \tag{78}$$

$$\delta = \frac{\epsilon}{\max\{\|\boldsymbol{h}_1\|, \ldots, \|\boldsymbol{h}_{N-1}\|, \|\boldsymbol{H}\mathbf{1}\|\}} > 0. \tag{79}$$

The polyhedron $\mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})$ is also an $(N-1)$-dimensional simplex; this is shown by Fact 3.(b) and the fact that the rank of $\tilde{\boldsymbol{H}}$ is the same as that of $\boldsymbol{H}$ (which is full). Now, we claim that $\mathcal{B}(r) \subseteq \mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})$ and $\mathrm{vol}(\mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})) < \mathrm{vol}(\mathcal{P}(\boldsymbol{g}, \boldsymbol{H})) = \mathrm{vol}(\mathcal{W})$. For the first claim, one can verify from (76)-(77) that

$$\begin{aligned} -r\|\tilde{\boldsymbol{h}}_1\| + \tilde{g}_1 &\geq (N-1)\epsilon \geq 0, \\ -r\|\tilde{\boldsymbol{h}}_i\| + \tilde{g}_i &\geq 0, \quad i = 2, \ldots, N-1, \\ -r\|\tilde{\boldsymbol{H}}\boldsymbol{1}\| + (1 - \boldsymbol{1}^T\tilde{\boldsymbol{g}}) &\geq \epsilon \geq 0, \end{aligned}$$

where $\tilde{\boldsymbol{h}}_i$ and $\tilde{g}_i$ denote the ith column of $\tilde{\boldsymbol{H}}$ and ith element of $\tilde{\boldsymbol{g}}$, resp. The above equations, together with Fact 4.(a), implies that $\mathcal{B}(r) \subseteq \mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})$. The second claim follows from (66) in Fact 3.(b) and (77c):

$$\begin{aligned} \mathrm{vol}(\mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})) &= \frac{1}{(N-1)!}\left(\frac{r}{r+\delta}\right)^{N-1} |\det(\boldsymbol{H})|^{-1} \\ &< \frac{1}{(N-1)!}|\det(\boldsymbol{H})|^{-1} = \mathrm{vol}(\mathcal{W}), \end{aligned} \tag{80}$$

for $N \geq 2$ (note that $N = 1$ is meaningless). The above two claims contradicts the assumption that $\mathcal{W}$ is an MVES of $\mathcal{B}(r)$.

Case 2: Suppose that $\boldsymbol{t}_N \notin \mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W}$. The proof is similar to that of Case 1. Very concisely, this case has $-r\|\boldsymbol{H}\boldsymbol{1}\| + (1 - \boldsymbol{1}^T\boldsymbol{g}) > 0$ and $-r\|\boldsymbol{h}_i\| + g_i \geq 0$ for all $i \in \{1, \ldots, N-1\}$. By constructing a polyhedron $\mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})$ where

$$\begin{aligned} \tilde{\boldsymbol{g}} &= \boldsymbol{g} + \epsilon\boldsymbol{1}, \quad \tilde{\boldsymbol{H}} = \left(\frac{r+\delta}{r}\right)\boldsymbol{H}, \\ \epsilon &= \frac{-r\|\boldsymbol{H}\boldsymbol{1}\| + (1 - \boldsymbol{1}^T\boldsymbol{g})}{2N}, \end{aligned}$$

and δ is the same as (79), we show that $\mathcal{B}(r) \subseteq \mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})$ and $\mathrm{vol}(\mathcal{P}(\tilde{\boldsymbol{g}}, \tilde{\boldsymbol{H}})) < \mathrm{vol}(\mathcal{W})$. The above two claims contradict the MVES assumption with $\mathcal{W}$.

The above two cases imply that $\mathrm{bd}\mathcal{B}(r) \cap \mathrm{bd}\mathcal{W} = \{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N\}$, the desired result. In addition to this, Property (a) in Lemma 5 is obvious since the expression of $\boldsymbol{t}_i$'s in (68), as well as (65), already suggest the affine independence of $\boldsymbol{t}_1, \ldots, \boldsymbol{t}_N$. As for Property (b) in Lemma 5, note that (69) are all satisfied. It can be verified that by substituting (68) and (69) into (64), $\mathcal{W}$ can be rewritten as $\mathcal{W} = \cap_{i=1}^{N} \left\{\boldsymbol{\theta} \in \mathbb{R}^{N-1} \mid r^2 \geq \boldsymbol{t}_i^T\boldsymbol{\theta}\right\}$.

References

[1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[2] W.-K. Ma, J. M. Bioucas-Dias, J. Chanussot, and P. Gader, Eds., *Special Issue on Signal and Image Processing in Hyperspectral Remote Sensing, IEEE Signal Process. Mag.*, vol. 31, no. 1, Jan. 2014.

[3] J. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 5, no. 2, pp. 354–379, 2012.

[4] W.-K. Ma, J. M. Bioucas-Dias, T.-H. Chan, N. Gillis, P. Gader, A. J. Plaza, A. Ambikapathi, and C.-Y. Chi, "A signal processing perspective on hyperspectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 67–81, 2014.

[5] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.

[6] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, 2009.

[7] N. Gillis and S. A. Vavasis, "Fast and robust recursive algorithms for separable nonnegative matrix factorization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 4, pp. 698–714, 2014.

[8] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *Proc. IEEE IGARSS*, Aug. 2008.

[9] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 542–552, May 1994.

[10] W. E. Full, R. Ehrlich, and J. E. Klovan, "EXTENDED QMODEL—objective definition of external endmembers in the analysis of mixtures," *Mathematical Geology*, vol. 13, no. 4, pp. 331–344, 1981.

[11] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proc. SPIE Conf. Imaging Spectrometry*, Pasadena, CA, Oct. 1999, pp. 266–275.

[12] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "End-member extraction for hyperspectral image analysis," *Applied Optics*, vol. 47, no. 28, pp. F77–F84, 2008.

[13] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4177–4193, 2011.

[14] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, 2007.

[15] A. Agathos, J. Li, D. Petcu, and A. Plaza, "Multi-GPU implementation of the minimum volume simplex analysis algorithm for hyperspectral unmixing," to appear in *IEEE J. Sel. Topics Appl. Earth Observ.*, 2014.

[16] J. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. IEEE WHISPERS*, Aug. 2009.

[17] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "Chance-constrained robust minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4194–4209, 2011.

[18] E. M. Hendrix, I. García, J. Plaza, G. Martin, and A. Plaza, "A new minimum-volume enclosing algorithm for endmember identification and abundance estimation in hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2744–2757, 2012.

[19] M. B. Lopes, J. C. Wolff, J. Bioucas-Dias, and M. Figueiredo, "NIR hyperspectral unmixing based on a minimum volume criterion for fast and accurate chemical characterisation of counterfeit tablets," *Analytical Chemistry*, vol. 82, no. 4, pp. 1462–1469, 2010.

[20] J. Nascimento and J. Bioucas-Dias, "Hyperspectral unmixing based on mixtures of Dirichlet components," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 863–878, 2012.

[21] J. Plaza, E. M. Hendrix, I. García, G. Martín, and A. Plaza, "On endmember identification in hyperspectral images without pure pixels: A comparison of algorithms," *Journal of Mathematical Imaging and Vision*, vol. 42, no. 2-3, pp. 163–175, 2012.

[22] C.-H. Lin, A. Ambikapathi, W.-C. Li, and C.-Y. Chi, "On the endmember identifiability of Craig's criterion for hyperspectral unmixing: A statistical analysis for three-source case," in *Proc. IEEE ICASSP*, May 2013, pp. 2139–2143.

[23] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[24] P. Gritzmann, V. Klee, and D. Larman, "Largest j-simplices in n-polytopes," *Discrete and Computational Geometry*, vol. 13, no. 1, pp. 477–515, 1995.

[25] T.-H. Chan, W.-K. Ma, C.-Y. Chi, and Y. Wang, "A convex analysis framework for blind separation of non-negative sources," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5120–5134, 2008.

[26] R. Clark, G. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. Sutley, "USGS digital spectral library splib06a: U.S. Geological Survey, Digital Data Series 231," http://speclab.cr.usgs.gov/spectral.lib06, 2007.

[27] L. Gerber, "The orthocentric simplex as an extreme simplex," *Pacific Journal of Mathmatics*, vol. 56, no. 1, pp. 97–111, Nov. 1975.