



A lattice matrix method for hyperspectral image unmixing

Gerhard X. Ritter^a, Gonzalo Urcid^{b,*}

^a CISE Department, University of Florida, Gainesville, FL 32611, USA

^b Optics Department, INAOE, Tonantzintla, Pue 72000, Mexico

ARTICLE INFO

Article history:

Received 13 March 2009

Received in revised form 3 December 2009

Accepted 14 March 2010

Available online 27 March 2010

Keywords:

Neural networks: Associative memories

Lattice associative memories

Lattice algebra: affine independence

Lattice independence

Lattice matrices

Hyperspectral image analysis: endmember search

Spectral unmixing

Abundance maps

ABSTRACT

In this manuscript we propose a method for the autonomous determination of endmembers in hyperspectral imagery based on recent theoretical advancements on lattice auto-associative memories. Given a hyperspectral image, the lattice algebra approach finds in a single-pass all possible candidate endmembers from which various affinely independent sets of final endmembers may be derived. In contrast to other endmember detection methods, the endmembers found using two dual canonical lattice matrices are geometrically linked to the data set spectra. The mathematical foundation of the proposed method is first described in some detail followed by application examples that illustrate the key steps of the proposed lattice based method.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

The high spectral resolution produced by current hyperspectral imaging devices facilitates identification of fundamental materials that make up a remotely sensed scene and thus supports discrimination between them. A pixel of a hyperspectral image physically represents a surface region on the ground comprising several square meters. Thus, a hyperspectral image pixel can have all or parts of many different natural or man-made objects in it. The collection of measured reflectances associated with the pixel is called the *spectrum* of the pixel. It is, therefore, useful to know the percentage of different fundamental object parts that are most represented in the spectrum of a given pixel. The most widely used spectral mixing model is the linear mixing model, which assumes that the observed reflectance spectrum of a given pixel is a linear combination of a small number of unique constituent signatures known as *endmembers* [1]. In various applications, hyperspectral image segmentation and analysis takes the form of a pattern recognition problem as the segmentation problem reduces to matching the spectra of the hyperspectral image to predetermined signatures stored in a spectral library. In many cases, however, end-members cannot be determined in advance and must be selected from the image directly by identifying the pixel spectra that are most likely to represent the fundamental materials. This comprises the *autonomous endmember detection problem*. Unfortunately, the spatial resolution of a sensor makes it often unlikely that any pixel is composed of a single endmember. Thus, the determination of endmembers becomes a search for image pixels with the least contamination from other end-members. These are also referred to as *pure pixels*. The pure pixels exhibit *maximal* or *minimal* reflectance in certain spectral bands and correspond to vertices of a high-dimensional simplex that, hopefully, encloses most if not all pixel spectra.

* Corresponding author. Tel.: +52 222 266 3100; fax: +52 222 247 2940.

E-mail addresses: ritter@cise.ufl.edu (G.X. Ritter), gurcid@inaoep.mx (G. Urcid).

In this paper we assume the *constrained linear mixing model* based on the fact that points on a simplex can be represented as a linear sum of the vertices that determine the simplex. The mathematical equations of the model and its constraints are given by

$$\begin{aligned} \mathbf{x} = S\mathbf{a} + \mathbf{r} &= \sum_{i=1}^m a_i \mathbf{s}^i + \mathbf{r}; \\ \sum_{i=1}^m a_i &= 1 \text{ and } a_i \geq 0 \quad \forall i, \end{aligned} \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the measured spectrum of an image pixel, $S = (\mathbf{s}^1, \dots, \mathbf{s}^m)$ is an $n \times m$ matrix whose columns are the m end-member spectra assumed to be affinely independent, the entries of $\mathbf{a} = (a_1, \dots, a_m)^t$ are the corresponding abundances or fractions of the endmember spectra present in \mathbf{x} , and \mathbf{r} represents a noise vector.

Endmembers may be obtained from spectral libraries for certain specific materials, or automatically from the image by a variety of techniques [5,7,28,37,39,40]. Autonomous endmember detection has received wide attention since signatures of various objects that may be present in an image are unknown before hand. Boardman [5] uses the framework of the geometry of convex sets to identify the $m+1$ endmembers as the vertices of the smallest simplex that bounds the measured data. However, the simplex vertices need not be image pixels and, hence, need not coincide with actual image data. Winter's N-FINDR method [39,40] is based on inflating a simplex within the data set to determine the largest simplex inscribed within the data. This algorithm is computationally intensive since individual pixels need to be examined and simplex volume recalculated for each image pixel. Recent methods proposed by Nascimento-Bioucas [28] and Chang et al. [7] offer similar or faster performance with respect to the N-FINDR method and are based on convex optimization techniques [6]. The autonomous endmember determination method proposed in this paper is also fast and carries little computational overhead. The method is derived from examining a lattice based auto-associative memory that stores the hyperspectral image cube. Graña et al., were the first to propose the use of lattice based auto-associative memories for autonomous endmember determination [15] as well as an evolutionary based strategy for endmember discrimination [16]. In the first approach, related to the present work, they employed the notion of morphological independence which does not necessarily lead to finding an affinely independent set of vectors that in some sense provides a maximal simplex within the data set. Furthermore, Graña's algorithm requires the user to choose a starting pixel and different starting pixels can produce different results. Improvements of algorithms using the preceding approach are based on recent discoveries of algebraic properties inherent in lattice based auto-associative memories [34] but the endmembers obtained need not be related with the hyperspectral image. The WM method described here differs from those described by Graña [17,18] and Myers [26] as the endmembers we obtain have a geometrical relationship to the pixels of the hyperspectral image under consideration. Our method will always provide the same sets of candidate endmembers based on theoretical facts given in this paper. To validate our proposed method, a brief comparison is made against two new approaches based on convex optimization, namely vertex component analysis [28] and the minimal volume enclosing simplex [7].

The paper is organized as follows. Section 2 introduces the reader to background concepts on binary lattice operations with numbers, vectors and matrices. Lattice associative memories (LAMs) and their fundamental properties are discussed in Section 3. The following three sections are devoted to the mathematical foundation that guarantees the correctness of the proposed method. Specifically, Section 4 establishes the geometric description of the fixed point set of lattice auto-associative memories. Section 5 establishes the relationships between the hyperspectral data cube, the corresponding LAMs and their fixed point set. In Section 6, we present the theoretical results needed to prove the affine independence of the sets (or proper subsets) of scaled column vectors derived from LAMs. Endmember determination using the WM method and constrained linear unmixing of hyperspectral images is presented in Section 7. Finally, a brief discussion about our proposed method and some conclusions are provided in Section 8.

2. Lattice theory fundamentals

The computational concepts for the associative neural networks used in this manuscript are governed by the *bounded lattice ordered group* $(\mathbb{R}_{\pm\infty}, \vee, \wedge, +, +')$, where \mathbb{R} denotes the set of real numbers, $\mathbb{R}_{\pm\infty} = \mathbb{R} \cup \{-\infty, \infty\}$, \vee and \wedge denote the binary operation of maximum and minimum, respectively, $+$ denotes addition, and $+'$ denotes the dual operation of $+$ defined by $a+b' = a+b$ for any $a \in \mathbb{R}$. If $a \in \mathbb{R}_{\pm\infty}$, then its *additive conjugate* is given by $a^* = -a$.

Unless stated otherwise, a vector $\mathbf{x} \in \mathbb{R}_{\pm\infty}^n$ is always viewed as a column vector, i.e., $\mathbf{x} = (x_1, \dots, x_n)^t$, where $x_i \in \mathbb{R}_{\pm\infty}$ for $i = 1, \dots, n$ and t denotes the transpose. Scalar addition of a vector $\mathbf{x} \in \mathbb{R}_{\pm\infty}^n$ is defined componentwise. That is, if $a \in \mathbb{R}_{\pm\infty}$, then $a + \mathbf{x} = (a + x_1, \dots, a + x_n)^t$. The *conjugate* of $\mathbf{x} \in \mathbb{R}_{\pm\infty}^n$ is defined as $\mathbf{x}^* = -\mathbf{x}^t$. Given two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}_{\pm\infty}^n$, then the *maximum* and *minimum* of \mathbf{x} and \mathbf{y} , denoted by $\mathbf{x} \vee \mathbf{y}$ and $\mathbf{x} \wedge \mathbf{y}$, respectively, are defined componentwise as $(\mathbf{x} \vee \mathbf{y})_i = x_i \vee y_i$ and $(\mathbf{x} \wedge \mathbf{y})_i = x_i \wedge y_i$ for $i = 1, \dots, n$. We note that the following *duality* De Morgan's type identities hold: $\mathbf{x} \vee \mathbf{y} = (\mathbf{x}^* \wedge \mathbf{y}^*)^*$ and $\mathbf{x} \wedge \mathbf{y} = (\mathbf{x}^* \vee \mathbf{y}^*)^*$. The inequalities $\mathbf{x} \leqslant \mathbf{y}$ and $\mathbf{x} < \mathbf{y}$ mean that $x_i \leqslant y_i$ and $x_i < y_i$, respectively, where $i = 1, \dots, n$. Thus, if $\mathbf{u} = \mathbf{x} \vee \mathbf{y}$ and $\mathbf{v} = \mathbf{x} \wedge \mathbf{y}$, then $\mathbf{v} \leqslant \mathbf{u}$.

As our application domain concerns only real valued vectors, we restrict our discussion to sets of vectors $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subset \mathbb{R}_{\pm\infty}^n$ for which $\mathbf{x}^\xi \in \mathbb{R}^n$ and $\xi \in K$ where $K = \{1, \dots, k\}$ is a finite set of positive integers. With this

restriction the operation of scalar addition is self-dual as $a + \mathbf{x}^\xi = a' + \mathbf{x}^\xi$ for all $\xi \in K$ and any $a \in \mathbb{R}_{\pm\infty}$. Henceforth, we suppose that $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subset \mathbb{R}^n$. A *linear minimax combination* of vectors from the set X is any vector $\mathbf{x} \in \mathbb{R}_{\pm\infty}^n$ of the form

$$\mathbf{x} = \mathfrak{S}(X) = \bigvee_{j \in J} \bigwedge_{\xi=1}^k (a_{\xi j} + \mathbf{x}^\xi), \quad (2)$$

where J is a finite set of indices and $a_{\xi j} \in \mathbb{R}_{\pm\infty} \forall j \in J$ and $\forall \xi \in K$. The expression $\mathfrak{S}(X)$ given by (2) is called a *linear minimax sum*. In particular any finite expression involving the symbols \vee, \wedge and vectors of form $a + \mathbf{x}^\xi$, where $a \in \mathbb{R}_{\pm\infty}$ and $\mathbf{x}^\xi \in X$ is a linear minimax sum with vectors from X .

A vector $\mathbf{x} \in \mathbb{R}^n$ is *lattice dependent* on X if and only if $\mathbf{x} = \mathfrak{S}(X)$ for some linear minimax sum of vectors from X . The vector \mathbf{x} is said to be *lattice independent* of X if and only if it is not lattice dependent on X . The set X is said to be *lattice independent* if and only if $\forall \lambda \in K, \mathbf{x}^\lambda$ is lattice independent of $X \setminus \{\mathbf{x}^\lambda\} = \{\mathbf{x}^\xi \in X : \xi \neq \lambda\}$ denoted for simplicity by X^λ . In other words, X is lattice independent if and only if $\mathbf{x}^\lambda \notin \mathfrak{S}(X^\lambda)$ for any set of constants $\{a_{\xi j}\} \subset \mathbb{R}_{\pm\infty}$ and any $\lambda \in K$.

Given two $m \times n$ matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ with entries from $\mathbb{R}_{\pm\infty}$, then the *pointwise maximum*, $\mathbf{A} \vee \mathbf{B}$, of \mathbf{A} and \mathbf{B} , is the $m \times n$ matrix \mathbf{C} defined by $\mathbf{A} \vee \mathbf{B} = \mathbf{C}$, where $c_{ij} = a_{ij} \vee b_{ij}$. Similarly, the *pointwise minimum* of two matrices of the same size is defined as $\mathbf{A} \wedge \mathbf{B} = \mathbf{C}$, where $c_{ij} = a_{ij} \wedge b_{ij}$. If \mathbf{A} is $m \times p$ and \mathbf{B} is $p \times n$, then the *max product* of \mathbf{A} and \mathbf{B} is the matrix $\mathbf{C} = \mathbf{A} \boxtimes \mathbf{B}$, where $c_{ij} = \bigvee_{k=1}^p (a_{ik} + b_{kj})$. The *min product* of \mathbf{A} and \mathbf{B} is the matrix $\mathbf{C} = \mathbf{A} \boxdot \mathbf{B}$ defined by $c_{ij} = \bigwedge_{k=1}^p (a_{ik} + b_{kj})$. The two matrix products are collectively referred to as *minimax products*. A vector $\mathbf{x} \in \mathbb{R}_{\pm\infty}^n$ is called a *max fixed point* of \mathbf{A} if $\mathbf{A} \boxtimes \mathbf{x} = \mathbf{x}$ and a *min fixed point* of \mathbf{A} if $\mathbf{A} \boxdot \mathbf{x} = \mathbf{x}$. The mathematical theory of lattices is exposed in [4], an overview of minimax and lattice algebra can be found in [11,31], and recent theoretical material appears in [34].

3. Lattice associative memories

Suppose $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subset \mathbb{R}^n$ and $Y = \{\mathbf{y}^1, \dots, \mathbf{y}^m\} \subset \mathbb{R}^m$ are two sets of column vectors with desired association given by the diagonal $\{(\mathbf{x}^\xi, \mathbf{y}^\xi) : \xi \in K\}$ of $X \times Y$ where $K = \{1, \dots, k\}$. The goal is to store these vector pairs in some memory device that recalls \mathbf{y}^ξ when presented with the input vector \mathbf{x}^ξ , for $\xi \in K$. With each pair of vectorial associations (X, Y) , two canonical *lattice associative memories* (LAMs) denoted by \mathbf{W}_{XY} and \mathbf{M}_{XY} can be defined. These matrices, of size $m \times n$, can be expressed in terms of minimax outer products, namely

$$\mathbf{W}_{XY} = \bigvee_{\xi=1}^k (\mathbf{y}^\xi + \mathbf{x}^{\xi*}); \quad \mathbf{M}_{XY} = \bigvee_{\xi=1}^k (\mathbf{y}^\xi - \mathbf{x}^{\xi*}), \quad (3)$$

where the *minimax outer product* is defined as $\mathbf{y} + \mathbf{x}^*$ or, componentwise, as $(\mathbf{y} + \mathbf{x}^*)_{ij} = y_i - x_j$, for $i = 1, \dots, m$ and $j = 1, \dots, n$. Therefore, the entry elements of the *min-memory* \mathbf{W}_{XY} and the *max-memory* \mathbf{M}_{XY} , respectively, are defined by

$$w_{ij} = \bigwedge_{\xi=1}^k (y_i^\xi - x_j^\xi); \quad m_{ij} = \bigvee_{\xi=1}^k (y_i^\xi - x_j^\xi). \quad (4)$$

The memories are called *auto-associative memories* whenever $Y = X$. In this case the main diagonals of the square matrices \mathbf{W}_{XX} and \mathbf{M}_{XX} consist entirely of zeros, i.e., $w_{ii} = m_{ii} = 0$ for $i = 1, \dots, n$. All entries of these matrices obey the following triangle inequalities (proven as Lemma 5.1 in [34] or Lemma 3.1 in [35]):

Theorem 3.1. *If $h, j, \ell \in \{1, \dots, n\}$, then $w_{j\ell} + w_{\ell h} \leq w_{jh}$ and $m_{h\ell} \leq m_{jh} + m_{j\ell}$.*

When speaking of fixed points of the matrices \mathbf{W}_{XX} and \mathbf{M}_{XX} , we always mean a fixed point of \mathbf{W}_{XX} with respect to the operation \boxtimes and of \mathbf{M}_{XX} with respect to the operation \boxdot . The strengths and weaknesses of these correlation type lattice matrix memories have been discussed in great detail in [30,32–34]. The following properties are pertinent to our discussion (see proofs in [34]). Recall that, $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subset \mathbb{R}^n$ and $\mathbf{x} \in \mathbb{R}^n$.

- 3.1. $\mathbf{W}_{XX} \boxtimes \mathbf{x}^\xi = \mathbf{x}^\xi = \mathbf{M}_{XX} \boxdot \mathbf{x}^\xi$ for $\xi \in K$.
- 3.2. $\mathbf{W}_{XX} \boxtimes \mathbf{x} = \mathbf{x}$ if and only if $\mathbf{M}_{XX} \boxdot \mathbf{x} = \mathbf{x}$.
- 3.3. \mathbf{x} is a fixed point of $\mathbf{W}_{XX} \iff \mathbf{x}$ is lattice dependent on X or, equivalently, \mathbf{x} is a fixed point of $\mathbf{M}_{XX} \iff \mathbf{x}$ is lattice dependent on X .

Property 3.1 establishes that \mathbf{W}_{XX} and \mathbf{M}_{XX} are *perfect recall* memories for uncorrupted input vectors and any positive integer k , no matter how large. Also, Property 3.2 says that \mathbf{W}_{XX} and \mathbf{M}_{XX} share the same set of fixed points represented by $F(X)$. Hence every max fixed point of \mathbf{W}_{XX} is also a min fixed point of \mathbf{M}_{XX} and vice versa. Finally, Property 3.3 provides an algebraic classification of the points belonging to $F(X)$.

4. The shape of $F(X)$

The algebraic classification of $F(X)$ conveys little information regarding the shape of $F(X)$ or its geometric relationships to X , \mathbf{W}_{XX} , and \mathbf{M}_{XX} . These relations are key in understanding the lattice based approach to autonomous endmember detection

and in proving the affine independence of the endmembers found by our method. The shape characterization of $F(X)$ requires some basic knowledge of vector space theory and high-dimensional Euclidean geometry. Specifically, the notions of *linear spaces*, *direction* of linear spaces, and *orientation* in higher dimensional spaces are essential. A set E in a vector space V over a field \mathbb{F} is called a *linear subspace* of V if and only if for any two distinct points $\mathbf{x}, \mathbf{y} \in E$ the set $L(\mathbf{x}, \mathbf{y}) = \{\mathbf{z} \in V : \mathbf{z} = \lambda\mathbf{x} + (1 - \lambda)\mathbf{y}, \lambda \in \mathbb{F}\} \subset E$. It follows that in vector spaces linear subspaces are equivalent to affine subspaces (but not necessarily vector subspaces) [10,19]. For spaces other than vector spaces, linear subspaces have generally no relation to affine spaces [3]. In our subsequent discussion $V = \mathbb{R}^n$, $\mathbb{F} = \mathbb{R}$, and the terms linear subspace and affine subspace can be used interchangeably. There are several linear subspaces of \mathbb{R}^n that play a vital role in describing $F(X)$. One such set of linear subspaces consists of lines defined by

$$L(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} = a + \mathbf{x}, a \in \mathbb{R}\}, \quad (5)$$

where $\mathbf{x} \in \mathbb{R}^n$ is fixed. The connection between $L(\mathbf{x})$ and $F(X)$ is given by the following observation. If $\mathbf{x} \in F(X)$, then $a + \mathbf{x} \in F(X) \forall a \in \mathbb{R}$ and, hence, $L(\mathbf{x}) \subset F(X)$ (see Theorem 6.2 in [34]). Another set of linear subspaces consists of specific types of hyperplanes. Recall that a hyperplane E in \mathbb{R}^n is defined as the set of all points $\mathbf{x} \in \mathbb{R}^n$ that satisfy an equation of the form

$$a_1x_1 + \cdots + a_nx_n = b, \quad (6)$$

where the a_i 's and b are constants and not all the a_i 's are zero. It follows from Eq. (6) that E is an $(n - 1)$ -dimensional linear subspace of \mathbb{R}^n . It will be convenient to associate a *direction* with a given linear subspace E of \mathbb{R}^n and an *orientation* with a given hyperplane. For this purpose we let $\mathbf{e}^1, \dots, \mathbf{e}^n$ denote the canonical orthonormal basis for \mathbb{R}^n ; i.e., each \mathbf{e}^i is defined by $e_i^j = 1$ if $i = j$ and $e_i^j = 0$ for $i \neq j$. In addition to the canonical basis vectors, the *directional vectors* pertinent to our discussion are the vectors \mathbf{e} and \mathbf{d}_{ij} defined by

$$\mathbf{e} = \frac{\mathbf{e}^1 + \cdots + \mathbf{e}^n}{\|\mathbf{e}^1 + \cdots + \mathbf{e}^n\|}; \quad \mathbf{d}_{ij} = \frac{\mathbf{e}^i - \mathbf{e}^j}{\|\mathbf{e}^i - \mathbf{e}^j\|}, \quad (7)$$

where $i < j$, $1 \leq i < n$, and $1 < j \leq n$. Directions and orientations in \mathbb{R}^n will be specified in terms of unit vectors emanating from the origin $\mathbf{0}$ with endpoints lying on the $(n - 1)$ -dimensional unit sphere, $S^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n x_i^2 = 1\} \subset \mathbb{R}^n$, centered at the origin. Thus, a direction \mathbf{d} is uniquely determined by a system of directional cosines that determines the coordinates of \mathbf{d} on the unit sphere. This system is defined by $\cos \theta_i = \mathbf{e}^i \cdot \mathbf{d} = d_i$ where $\sum_{i=1}^n \cos^2 \theta_i = 1$.

A directional vector \mathbf{d} is a *parallel direction*, or simply a *direction* for a k -dimensional linear subspace $E^k \subset \mathbb{R}^n$ with $0 < k < n$ if and only if for every point $\mathbf{x} \in E^k$, $\mathbf{x} + \mathbf{d} \in E^k$. If $D(E^k) = \{\mathbf{d} \in S^{n-1} : \mathbf{d}$ is a parallel direction for $E^k\}$, then $D(E^k)$ is called the *set of parallel directions affiliated with E^k* and this set forms a $(k - 1)$ -dimensional subsphere of S^{n-1} . Also, two linear subspaces E^k and E^ℓ ($k \leq \ell$) are said to be *parallel* if and only if $D(E^k) \subset D(E^\ell)$. Obviously, lines $L(\mathbf{x})$ defined by Eq. (5) are parallel one-dimensional linear subspaces with affiliated parallel directions \mathbf{e} and $-\mathbf{e}$ where $D(L(\mathbf{x})) = \{\mathbf{e}, -\mathbf{e}\}$ is a 0-dimensional subsphere of S^{n-1} .

An *oriented* hyperplane $E(\mathbf{d})$ is simply a hyperplane E with an associated directional unit vector \mathbf{d} which is perpendicular to E . Since the vector $-\mathbf{d}$ points in the opposite direction of \mathbf{d} , each hyperplane can be endowed with one of two possible directions. Given two oriented hyperplanes $E_1(\mathbf{d}_1)$ and $E_2(\mathbf{d}_2)$, then E_1 and E_2 are said to be *parallel*, denoted by $E_1 \parallel E_2$, whenever $\mathbf{d}_1 = \pm \mathbf{d}_2$ or, equivalently, whenever $\mathbf{d}_1 \cdot \mathbf{d}_2 = \pm 1$. If $\mathbf{d}_1 \cdot \mathbf{d}_2 \neq \pm 1$, then $E_1 \cap E_2$ is a linear subspace of dimension $n - 2$. In case $\mathbf{d}_1 \cdot \mathbf{d}_2 = 0$, E_1 and E_2 are said to be *perpendicular* and denoted by $E_1 \perp E_2$.

An oriented hyperplane $E(\mathbf{d})$ separates \mathbb{R}^n into two open half-spaces $H^+(\mathbf{d})$ and $H^-(\mathbf{d})$ which are bounded by $E(\mathbf{d})$. If E is given by Eq. (6), then E can also be expressed in terms of the function

$$f(\mathbf{x}) = a_1x_1 + \cdots + a_nx_n - b = 0. \quad (8)$$

We shall use the convention of identifying $H^+(\mathbf{d})$ and $H^-(\mathbf{d})$ with the half-spaces $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) > 0\}$ and $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) < 0\}$, respectively. The *closure* of $H^+(\mathbf{d})$ is the set $\overline{H^+(\mathbf{d})} = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \geq 0\}$. Similarly, $\overline{H^-(\mathbf{d})} = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq 0\}$. Therefore, $\overline{H^+(\mathbf{d})} \cap \overline{H^-(\mathbf{d})} = E(\mathbf{d})$. It is important to illustrate the geometric properties of hyperplanes of type $E(\mathbf{d}_{ij})$ as it is the half-spaces of these hyperplanes that determine the shape of $F(X)$. Specifically, there are $n(n - 1)/2$ distinct hyperplanes of type $E(\mathbf{d}_{ij})$ containing $\mathbf{y} \in \mathbb{R}^n$, where $1 \leq i < j \leq n$ (cf. [34]). We let $E_{\mathbf{y}}(\mathbf{d}_{ij}) = \{\mathbf{x} \in \mathbb{R}^n : x_i - x_j = y_i - y_j\}$ denote the unique hyperplane of type $E(\mathbf{d}_{ij})$ containing $\mathbf{y} \in \mathbb{R}^n$.

Example 4.1. Let the dimension $n = 2$ and $\mathbf{y} = (y_1, y_2)^t \in \mathbb{R}^2$, then there is only one hyperplane of type $E(\mathbf{d}_{ij})$ containing \mathbf{y} , namely the line $L(\mathbf{y}) = E_{\mathbf{y}}(\mathbf{d}_{12})$ of slope $m = 1$ given by $x_2 = x_1 + b$, where $b = y_2 - y_1$. Explicitly, $E_{\mathbf{y}}(\mathbf{d}_{12})$ has orientation $\mathbf{d}_{12} = (1/\sqrt{2}, -1/\sqrt{2})$ and direction $\mathbf{e} = (1/\sqrt{2}, 1/\sqrt{2})$. Fig. 1 illustrates the three unique hyperplanes $E_{\mathbf{y}}(\mathbf{d}_{ij})$ for $\mathbf{y} \in \mathbb{R}^3$. It is easy to verify that each of the hyperplanes $E_{\mathbf{y}}(\mathbf{d}_{12}), E_{\mathbf{y}}(\mathbf{d}_{13}), E_{\mathbf{y}}(\mathbf{d}_{23})$ contains the line $L(\mathbf{y})$ and that the equality $L(\mathbf{y}) = \bigcap_{i < j} E_{\mathbf{y}}(\mathbf{d}_{ij})$ holds in general.

Another important geometric construction is derived from the following observation. Note that for each pair (i, j) with $1 \leq i < j \leq 3$, the hyperplane $E_{\mathbf{y}}(\mathbf{d}_{ij})$ in Fig. 1 is perpendicular to the (x_i, x_j) -plane in \mathbb{R}^3 . This holds true in any dimension. The hyperplane $\mathbb{R}_i^{n-1} \subset \mathbb{R}^n$ defined by $\mathbb{R}_i^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : x_i = 0\}$ has orientation \mathbf{e}^i , and if $\ell \notin \{i, j\}$, then $\mathbf{d}_{ij} \cdot \mathbf{e}^\ell = 0$. Therefore,

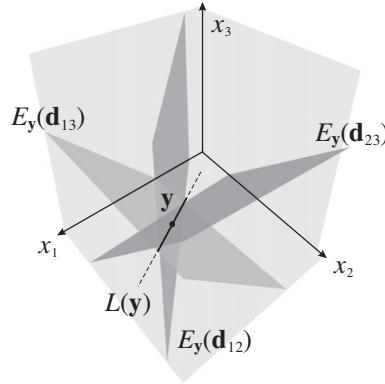


Fig. 1. The three unique hyperplanes containing $\mathbf{y} \in \mathbb{R}^3$; their intersection is the line $L(\mathbf{y})$.

$E_y(\mathbf{d}_{ij}) \perp \mathbb{R}_\ell^{n-1}$. The case where $\ell \in \{i, j\}$ is of special interest. In this case, we also consider the sets $S_y(\mathbf{d}_{ij}) = E_y(\mathbf{d}_{ij}) \cap \mathbb{R}_\ell^{n-1}$ given by,

$$S_y(\mathbf{d}_{ij}) = \{\mathbf{x} \in \mathbb{R}_\ell^{n-1} : x_j = +b\} \iff \ell = i, \quad (9)$$

$$S_y(\mathbf{d}_{i\ell}) = \{\mathbf{x} \in \mathbb{R}_\ell^{n-1} : x_i = -b\} \iff \ell = j, \quad (10)$$

where $b = y_j - y_i$ is constant.

Theorem 4.1. (1) If $\ell \notin \{i, j\}$, then $E_y(\mathbf{d}_{ij}) \perp \mathbb{R}_\ell^{n-1}$, (2) $S_y(\mathbf{d}_{ij}) \parallel \mathbb{R}_{ij}^{n-2}$ in \mathbb{R}_ℓ^{n-1} and both are oriented by \mathbf{e}^j , (3) $S_y(\mathbf{d}_{i\ell}) \parallel \mathbb{R}_{i\ell}^{n-2}$ in \mathbb{R}_ℓ^{n-1} and both are oriented by \mathbf{e}^i , and (4) $S_y(\mathbf{d}_{i\ell}) \perp S_y(\mathbf{d}_{ij})$ in \mathbb{R}_ℓ^{n-1} .

The shape of $F(X)$ for $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subset \mathbb{R}^n$ can be derived from hyperboxes situated in the linear subspaces \mathbb{R}_ℓ^{n-1} , where $\ell = 1, \dots, n$. For each $\xi = 1, \dots, k$ and $\ell \in \{1, \dots, n\}$ set $\mathbf{x}^\xi(\ell) = \mathbf{x}^\xi - \mathbf{x}_\ell^\xi$, $\mathbf{v}^\ell = \bigwedge_{\xi=1}^k \mathbf{x}^\xi(\ell)$, and $\mathbf{u}^\ell = \bigvee_{\xi=1}^k \mathbf{x}^\xi(\ell)$. The ℓ -hyperbox determined by \mathbf{v}^ℓ and \mathbf{u}^ℓ is the convex set $\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell) = \{\mathbf{x} \in \mathbb{R}_\ell^{n-1} : v_i^\ell \leq x_i \leq u_i^\ell, i = 1, \dots, n\}$. The points \mathbf{v}^ℓ and \mathbf{u}^ℓ are called the *minimal* and *maximal corners* of the ℓ -hyperbox $\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$ and are directly related, respectively, to the columns of matrices \mathbf{W}_{XX} and \mathbf{M}_{XX} (see [Theorem 5.1](#)). For any value of ℓ we have the following fundamental relationships between these corners:

Theorem 4.2. For $i, j \in \{1, \dots, n\}$, $u_i^j = -v_i^j$ and $v_j^i = -u_j^i$.

Corollary 4.3. $E_{\mathbf{u}^\ell}(\mathbf{d}_{ij}) = E_{\mathbf{v}^\ell}(\mathbf{d}_{ij})$; $E_{\mathbf{v}^\ell}(\mathbf{d}_{ij}) = E_{\mathbf{u}^\ell}(\mathbf{d}_{ij})$.

For $i < j$, we have $S_y(\mathbf{d}_{ij}) \parallel S_{\mathbf{u}^\ell}(\mathbf{d}_{ij})$ in \mathbb{R}_i^{n-1} and since they are hyperplanes in \mathbb{R}_i^{n-1} each separates \mathbb{R}_i^{n-1} into two open half-spaces. For any $\mathbf{y} \in \mathbb{R}_i^{n-1}$ we let $F_y^+(\mathbf{d}_{ij})$ and $F_y^-(\mathbf{d}_{ij})$ denote the open half-spaces of \mathbb{R}_i^{n-1} associated with the hyperplane $S_y(\mathbf{d}_{ij})$. The next theorem shows how the ℓ -hyperboxes $\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$ are expressed as finite intersections of open half-spaces of $S_y(\mathbf{d}_{ij})$ type hyperplanes.

Theorem 4.4. Given any $\ell \in \{1, \dots, n\}$, then

$$\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell) = \bigcap_{i < \ell} [\bar{F}_{\mathbf{u}^\ell}^-(\mathbf{d}_{i\ell}) \cap \bar{F}_{\mathbf{v}^\ell}^+(\mathbf{d}_{i\ell})] \cap \bigcap_{\ell < j} [\bar{F}_{\mathbf{u}^\ell}^+(\mathbf{d}_{ij}) \cap \bar{F}_{\mathbf{v}^\ell}^-(\mathbf{d}_{ij})]. \quad (11)$$

Note the reversal of the positive and negative half-planes in the argument of the second big intersection operator versus those in the argument of the first big intersection operator. If $\ell = 1$ in Eq. (11), then the case $i < \ell$ does not occur and the intersection $\bigcap_{i < \ell} = \bigcap_{j=2}^n$. Similarly, if $\ell = n$, then the case $\ell < j$ does not appear in the equation and the intersection $\bigcap_{\ell < j} = \bigcap_{i=1}^{n-1}$. These two special cases also govern subsequent equations of this type. However, if $1 < \ell < n$ then $\bigcap_{i < \ell} = \bigcap_{i=1}^{\ell-1}$ and $\bigcap_{\ell < j} = \bigcap_{i=\ell+1}^n$. We also note that whenever we have $\mathbf{v}^\ell \leq \mathbf{u}^\ell$ and strict inequality ($<$) does not hold, then there exists at least one index i such that $v_i^\ell = u_i^\ell$. Thus, if $i < \ell$, then $S_{\mathbf{v}^\ell}(\mathbf{d}_{i\ell}) = S_{\mathbf{u}^\ell}(\mathbf{d}_{i\ell})$, and if $\ell < i$, then $S_{\mathbf{v}^\ell}(\mathbf{d}_{i\ell}) = S_{\mathbf{u}^\ell}(\mathbf{d}_{i\ell})$. It follows that for each index i for which $v_i^\ell = u_i^\ell$, the dimension of $\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$ is reduced by one. If $v_i^\ell < u_i^\ell$ for all $i \neq \ell$, then of course $\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$ is $(n-1)$ -dimensional.

If $\bar{H}_{\mathbf{u}^\ell}^-(\mathbf{d}_{i\ell}), \bar{H}_{\mathbf{v}^\ell}^+(\mathbf{d}_{i\ell}), \bar{H}_{\mathbf{u}^\ell}^+(\mathbf{d}_{ij})$, and $\bar{H}_{\mathbf{v}^\ell}^-(\mathbf{d}_{ij})$ are closed half-spaces in \mathbb{R}^n determined by the hyperplanes $E_{\mathbf{u}^\ell}(\mathbf{d}_{i\ell}), E_{\mathbf{v}^\ell}(\mathbf{d}_{i\ell}), E_{\mathbf{u}^\ell}(\mathbf{d}_{ij})$, and $E_{\mathbf{v}^\ell}(\mathbf{d}_{ij})$, respectively, then Eq. (11) can also be written as

$$\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell) = \left(\bigcap_{i < \ell} [\bar{H}_{\mathbf{u}^\ell}^-(\mathbf{d}_{i\ell}) \cap \bar{H}_{\mathbf{v}^\ell}^+(\mathbf{d}_{i\ell})] \cap \bigcap_{\ell < j} [\bar{H}_{\mathbf{u}^\ell}^+(\mathbf{d}_{ij}) \cap \bar{H}_{\mathbf{v}^\ell}^-(\mathbf{d}_{ij})] \right) \cap \mathbb{R}_\ell^{n-1}. \quad (12)$$

The equation follows from the observation that $\bar{F}_{\mathbf{u}^\ell}^+(\mathbf{d}_{i\ell}) = \bar{H}_{\mathbf{u}^\ell}^-(\mathbf{d}_{i\ell}) \cap \mathbb{R}_\ell^{n-1}$, $\bar{F}_{\mathbf{v}^\ell}^-(\mathbf{d}_{i\ell}) = \bar{H}_{\mathbf{v}^\ell}^-(\mathbf{d}_{i\ell}) \cap \mathbb{R}_\ell^{n-1}$, $\bar{F}_{\mathbf{u}^\ell}^+(\mathbf{d}_{ij}) = \bar{H}_{\mathbf{u}^\ell}^-(\mathbf{d}_{ij}) \cap \mathbb{R}_\ell^{n-1}$, and $\bar{F}_{\mathbf{v}^\ell}^-(\mathbf{d}_{ij}) = \bar{H}_{\mathbf{v}^\ell}^-(\mathbf{d}_{ij}) \cap \mathbb{R}_\ell^{n-1}$. By setting $F_\ell(X) = F(X) \cap \mathbb{R}_\ell^{n-1}$, we obtain the following relationship between $F_\ell(X)$ and $\mathfrak{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$.

Theorem 4.5. $F_\ell(X) \subset \mathcal{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$.

This theorem simply states that $F_\ell(X)$ is bounded by the ℓ -hyperbox. It follows from the definition of $F_\ell(X)$ that $F(X) = \bigcup_{a \in \mathbb{R}} [a + F_\ell(X)]$ where $a + F_\ell(X) = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{y} = a + \mathbf{x}, \mathbf{x} \in F_\ell(X)\}$. Similar considerations show that $E_{\mathbf{y}}(\mathbf{d}_{ij}) = \bigcup_{a \in \mathbb{R}} [a + S_{\mathbf{y}}(\mathbf{d}_{ij})]$. Thus, by Theorem 4.5 and Eq. (12) we have that

$$F(X) \subset \bigcap_{i < \ell} [\bar{H}_{\mathbf{u}^\ell}^-(\mathbf{d}_{i\ell}) \cap \bar{H}_{\mathbf{v}^\ell}^+(\mathbf{d}_{i\ell})] \cap \bigcap_{\ell < j} [\bar{H}_{\mathbf{u}^\ell}^+(\mathbf{d}_{ij}) \cap \bar{H}_{\mathbf{v}^\ell}^-(\mathbf{d}_{ij})]. \quad (13)$$

In order to replace the subset relationship by an equality it is necessary to let ℓ range over the set $\{1, \dots, n\}$.

Theorem 4.6

$$F(X) = \bigcap_{i=1}^{n-1} \bigcap_{j=i+1}^n [\bar{H}_{\mathbf{u}^i}^+(\mathbf{d}_{ij}) \cap \bar{H}_{\mathbf{v}^i}^-(\mathbf{d}_{ij})]. \quad (14)$$

According to this theorem, $F(X)$ is the intersection of particular types of half-spaces determined by the hyperplanes $E_{\mathbf{u}^i}(\mathbf{d}_{ij})$ and $E_{\mathbf{v}^i}(\mathbf{d}_{ij})$, where $i < j$ with $i = 1, \dots, n-1$ and $j = 2, \dots, n$. Since half-spaces are convex and the intersection of convex sets is again convex, $F(X)$ is a convex set. Furthermore, it follows that the boundary of $F(X)$, denoted by $\partial F(X)$, is given by

$$\partial F(X) = F(X) \cap \bigcup_{i=1}^{n-1} \bigcup_{j=i+1}^n [E_{\mathbf{u}^i}(\mathbf{d}_{ij}) \cup E_{\mathbf{v}^i}(\mathbf{d}_{ij})]. \quad (15)$$

If X is a convex subset of \mathbb{R}^n , then a hyperplane $E(\mathbf{d})$ is said to cut X if and only if $X \cap H^+(\mathbf{d}) \neq \emptyset \neq X \cap H^-(\mathbf{d})$. If $E(\mathbf{d})$ intersects the closure of X but does not cut X , then $E(\mathbf{d})$ is said to be a support hyperplane of X . Thus, $F(X)$ is characterized by pairwise parallel support hyperplanes of type $E(\mathbf{d}_{ij})$ and the maximum number of these boundary hyperplanes is $n(n-1)$. It is important to note that although $F(X)$ has non-empty boundary, it is not a bounded set. For every point $\mathbf{x} \in F(X)$, the unbounded line $L(\mathbf{x})$ is a subset of $F(X)$. Except for Theorem 4.5 that was proven in [34], the proofs of theorems and corollaries given in this section, including Eq. (15), appear in [36].

5. X , \mathbf{W}_{XX} , \mathbf{M}_{XX} , and $F(X)$ relationships

If $X \subset \mathbb{R}^n$ is convex and $\mathbf{s} \in X$, then \mathbf{s} is an extreme point of X if and only if there do not exist two points $\mathbf{x}_1, \mathbf{x}_2 \in X$ such that $\mathbf{x}_1 \neq \mathbf{s} \neq \mathbf{x}_2$ and $\mathbf{s} \in [\mathbf{x}_1, \mathbf{x}_2]$, where $[\mathbf{x}_1, \mathbf{x}_2] \subset X$ denotes the line segment $\{\mathbf{x} \in X : \mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2, 0 \leq \lambda \leq 1\}$. It is well known that a closed bounded convex set is completely specified by its extreme points. In other words, if X is a closed bounded convex set and $S \subset X$ is the set of extreme points of X , then $S \neq \emptyset$ and $C(S) = X$, where $C(S)$ denotes the convex hull of S [14,19]. For instance, the points \mathbf{v}^ℓ and \mathbf{u}^ℓ are extreme points of the convex set $\mathcal{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$. However, they are not extreme points of the unbounded set $F(X)$. For example, the point \mathbf{u}^ℓ is a point of the line

$$L(\mathbf{u}^\ell) = \bigcap_{i < \ell} E_{\mathbf{u}^\ell}(\mathbf{d}_{i\ell}) \cap \bigcap_{\ell < j} E_{\mathbf{u}^\ell}(\mathbf{d}_{ij}), \quad (16)$$

so that $\mathbf{u}^\ell \in [\mathbf{u}^\ell - a, \mathbf{u}^\ell + b] \subset L(\mathbf{u}^\ell) \subset \partial F(X) \subset F(X)$ for any pair of real numbers $a > 0$ and $b > 0$. Similarly, the point \mathbf{v}^ℓ is a point on the line

$$L(\mathbf{v}^\ell) = \bigcap_{i < \ell} E_{\mathbf{v}^\ell}(\mathbf{d}_{i\ell}) \cap \bigcap_{\ell < j} E_{\mathbf{v}^\ell}(\mathbf{d}_{ij}). \quad (17)$$

The theoretical results embodied by the following theorems and corollaries are only stated since their proofs have been given in [36]. We shall use the notation \mathbf{W}^j and \mathbf{M}^j to denote the j th column of the matrices \mathbf{W}_{XX} and \mathbf{M}_{XX} , respectively. The relationship between the ℓ -hyperboxes $\mathcal{B}(\mathbf{v}^\ell, \mathbf{u}^\ell)$ and the lattice matrices \mathbf{W}_{XX} and \mathbf{M}_{XX} is given by the next theorem.

Theorem 5.1. For $\ell = 1, \dots, n$, $\mathbf{v}^\ell = \mathbf{W}^\ell$ and $\mathbf{u}^\ell = \mathbf{M}^\ell$.

Henceforth we let

$$\mathbf{v} = \bigwedge_{\xi=1}^k \mathbf{x}^\xi \quad \text{and} \quad \mathbf{u} = \bigvee_{\xi=1}^k \mathbf{x}^\xi, \quad (18)$$

be, respectively, the minimal and maximal corners of the hyperbox $\mathcal{B}(\mathbf{v}, \mathbf{u})$ defined by the convex set $\{\mathbf{x} \in \mathbb{R}^n : v_i \leq x_i \leq u_i, i = 1, \dots, n\}$. With each auto-associative memory \mathbf{W}_{XX} and \mathbf{M}_{XX} we build, respectively, a set of n additively scaled vectors, $W = \{\mathbf{w}^1, \dots, \mathbf{w}^n\}$ and $M = \{\mathbf{m}^1, \dots, \mathbf{m}^n\}$ defined by

$$\mathbf{w}^j = u_j + \mathbf{W}^j \quad \text{and} \quad \mathbf{m}^j = v_j + \mathbf{M}^j, \quad (19)$$

where u_j denotes the j th coordinate value of \mathbf{u} , v_j denotes the j th coordinate value of \mathbf{v} , and $j = 1, \dots, n$. As a direct consequence of Theorem 5.1 we have

Corollary 5.2. $\mathbf{w}^j = u_j + \mathbf{v}^j$ and $\mathbf{m}^j = v_j + \mathbf{u}^j$.

Also, since the diagonals of \mathbf{W}_{XX} and \mathbf{M}_{XX} consist entirely of zeros, the j th coordinate of the vectors \mathbf{w}^j and \mathbf{m}^j are given by $w_j^j = u_j$ and $m_j^j = v_j$, respectively.

Theorem 5.3. $\mathbf{u} = \bigvee_{j=1}^n \mathbf{w}^j$ and $\mathbf{v} = \bigwedge_{j=1}^n \mathbf{m}^j$.

Corollary 5.4. If $L(\mathbf{w}^\ell) = L(\mathbf{w}^j)$, then $\mathbf{w}^\ell = \mathbf{w}^j$. Similarly, if $L(\mathbf{m}^\ell) = L(\mathbf{m}^j)$, then $\mathbf{m}^\ell = \mathbf{m}^j$.

Another pertinent fact is the following:

Theorem 5.5. $\mathbf{W}_{XX} = \mathbf{W}_{WW}$; similarly, $\mathbf{M}_{XX} = \mathbf{M}_{MM}$ and $\mathbf{M}_{WW} = \mathbf{M}_{WW}$.

Since $\mathbf{v} \leq \mathbf{x}^\xi \leq \mathbf{u} \forall \xi \in \{1, \dots, k\}$, then $X \subset \mathcal{B}(\mathbf{v}, \mathbf{u})$ and, by Property 3.1, $X \subset F(X)$. In addition, $\mathcal{B}(\mathbf{v}, \mathbf{u})$ and $F(X)$ are convex sets both containing $C(X)$. Therefore, $X \subset C(X) \subset \mathcal{B}(\mathbf{v}, \mathbf{u}) \cap F(X)$. The set $\mathcal{B}(\mathbf{v}, \mathbf{u}) \cap F(X)$ is a convex polytope and the elements of $W \cup M \cup \{\mathbf{v}, \mathbf{u}\}$ are vertices of this polytope. This last observation follows from Corollary 5.2 which implies that $\mathbf{w}^j \in L(\mathbf{v}^j)$ and $\mathbf{m}^j \in L(\mathbf{u}^j)$. According to Eqs. (16) and (17), $L(\mathbf{u}^j)$ and $L(\mathbf{v}^j)$ are edges of $\partial F(X)$. Thus, \mathbf{w}^j and \mathbf{m}^j are edge points of $F(X)$. But the hyperplanes $E_{u_j} = \{\mathbf{x} \in \mathbb{R}^n : x_j = u_j\}$ and $E_{v_j} = \{\mathbf{x} \in \mathbb{R}^n : x_j = v_j\}$ both have orientation \mathbf{e}^j and, therefore, both cut $F(X)$. Since $w_j^j = u_j$ and $m_j^j = v_j$, then $\{\mathbf{w}^j\} = E_{u_j} \cap L(\mathbf{v}^j)$ and $\{\mathbf{m}^j\} = E_{v_j} \cap L(\mathbf{u}^j)$. Thus, \mathbf{w}^j and \mathbf{m}^j are vertex points of $\mathcal{B}(\mathbf{v}, \mathbf{u}) \cap F(X)$. Despite this fact, \mathbf{w}^j or \mathbf{m}^j themselves need not be elements of X .

6. Affine independence in LAMs

In this section we establish a sequence of theorems and corollaries that will aid in establishing necessary and sufficient conditions for extracting affine independent sets of vectors from W and M . The main two theorems affecting the extraction process will be Theorems 6.1 and 6.7.

Recall that a set $S = \{\mathbf{s}^0, \mathbf{s}^1, \dots, \mathbf{s}^m\} \subset \mathbb{R}^n$ is said to be *affinely independent* if and only if the set $\{\mathbf{s}^j - \mathbf{s}^0 : j = 1, \dots, m\}$ is linearly independent. Thus linear independence, by definition, implies affine independence. It follows from the definition of affine independence that any two distinct points are affinely independent, any three non-collinear points are affinely independent, any four non-coplanar points are affinely independent, and in general any m points in \mathbb{R}^n , with $m \leq n+1$, are affinely independent if and only if they are not points of a common $(m-2)$ -dimensional linear subspace of \mathbb{R}^n . The convex hulls of affinely independent points form the simplexes that are key in the linear mixing model defined by Eq. (1). In particular, if S is affinely independent, then $C(S)$ is an m -dimensional simplex, or m -simplex, denoted by σ^m or $\langle \mathbf{s}^0, \mathbf{s}^1, \dots, \mathbf{s}^m \rangle$. Thus, a 0-simplex is simply a point, a 1-simplex is a line segment determined by two affinely independent points, a 2-simplex is a triangle determined by three affinely independent points, a 3-simplex is a tetrahedron determined by four affinely independent points, and so on. Specifically,

$$\sigma^m = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_{i=0}^m a_i \mathbf{s}^i; \quad \sum_{i=0}^m a_i = 1, \quad a_i \geq 0 \right\}. \quad (20)$$

Since any non-empty subset of an affinely independent set is also affinely independent, any non-empty subset of S consisting of $j \leq m+1$ points generates a $(j-1)$ -dimensional simplex called a $(j-1)$ -dimensional face of σ^m . The $(m-1)$ -dimensional face of σ^m opposite the vertex \mathbf{s}^i is the simplex σ_i^{m-1} whose vertices are the elements of $S^i = S \setminus \{\mathbf{s}^i\}$ and the boundary of σ^m is given by $\partial \sigma^m = \bigcup_{i=0}^m \sigma_i^{m-1}$. The first theorem of this section provides six equivalent conditions that furnish a computationally simple test for the affine independence of the sets $W = \{\mathbf{w}^j : j = 1, \dots, n\}$ and $M = \{\mathbf{m}^j : j = 1, \dots, n\}$. Again, detailed proofs of theorems and corollaries stated in this section can be found in [36]. In the following discussion the symbols \mathbf{w}_i and \mathbf{m}_i will denote the i th row of W and M , respectively, that is $\mathbf{w}_i = (w_i^1, \dots, w_i^n)$. Also, the vector $\mathbf{c} = (c, \dots, c)$ will denote a constant vector, where $c \in \mathbb{R}$.

Theorem 6.1. If $j, \ell \in \{1, \dots, n\}$, then the following statements are equivalent: (1) $\mathbf{w}_j - \mathbf{w}_\ell = \mathbf{c}$, (2) $v_j^\ell = u_j^\ell$, (3) $\mathbf{w}^\ell = \mathbf{w}^j$, (4) $\mathbf{m}_j - \mathbf{m}_\ell = \mathbf{c}$, (5) $v_\ell^j = u_\ell^j$, (6) $\mathbf{m}^\ell = \mathbf{m}^j$.

Corollary 6.2. If $j < \ell$, then $E_{\mathbf{w}^j}(\mathbf{d}_{j\ell}) \| E_{\mathbf{w}^\ell}(\mathbf{d}_{j\ell})$ in \mathbb{R}^n and $E_{\mathbf{w}^j}(\mathbf{d}_{j\ell}) = E_{\mathbf{w}^\ell}(\mathbf{d}_{j\ell})$ if and only if $\mathbf{w}^j = \mathbf{w}^\ell$. Similarly, $E_{\mathbf{m}^j}(\mathbf{d}_{j\ell}) \| E_{\mathbf{m}^\ell}(\mathbf{d}_{j\ell})$ in \mathbb{R}^n and $E_{\mathbf{m}^j}(\mathbf{d}_{j\ell}) = E_{\mathbf{m}^\ell}(\mathbf{d}_{j\ell})$ if and only if $\mathbf{m}^j = \mathbf{m}^\ell$.

In various applications examples one often obtains an overlap in the elements of W and M in that $\mathbf{w}^\ell = \mathbf{m}^j$ without affecting the affine independence of either set W or M . Thus, the results of the next corollary is somewhat surprising.

Corollary 6.3. $\mathbf{w}^\ell = a + \mathbf{m}^\ell$ for $a \in \mathbb{R}^n$ and some $\ell \in \{1, \dots, n\}$ if and only if $\mathbf{w}^\ell = \mathbf{m}^j$ for $j = 1, \dots, n$.

In light of subsequent theorems, an important consequence of Theorem 6.1 is that in order to ascertain that W or M is affinely independent all one has to do is to check that no two vectors of W or M are identical.

Before stating the remaining theorems of this section we define certain concepts and sets that are critical in the proofs of these theorems. If the subset $W_\kappa = \{\mathbf{w}^{i_1}, \dots, \mathbf{w}^{i_\kappa}\} \subset W$ is affinely independent, then it generates a $(\kappa-1)$ -dimensional linear subspace $L^{\kappa-1} \subset \mathbb{R}^n$ defined by

$$L^{\kappa-1} = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_{j=1}^{\kappa} a_j \mathbf{w}^{i_j}; \quad \sum_{j=1}^{\kappa} a_j = 1, \quad a_j \in \mathbb{R} \right\}. \quad (21)$$

In an analogous fashion, we let $L^{\kappa-1}$ denote the $(\kappa-1)$ -dimensional linear subspace generated by an affinely independent subset $M_{\kappa} = \{\mathbf{m}^{i_1}, \dots, \mathbf{m}^{i_{\kappa}}\} \subset M$. Again, to simplify notation we let $J = \{1, \dots, n\}$, $J^j = J \setminus \{j\}$ for $j \in J$, $J_{\kappa} = \{i_1, \dots, i_{\kappa}\}$ will denote some arbitrary subset of J , and for $i_j \in J_{\kappa}$, we set $J_{\kappa}^j = J_{\kappa} \setminus \{i_j\}$. If $W_{\kappa} \subset W$ is affinely independent, then for $i_j \in J_{\kappa}$ we define the linear subspace $L_{i_j}^{\kappa-2}$ of $L^{\kappa-1}$ by

$$L_{i_j}^{\kappa-2} = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_{i_{\ell} \in J_{\kappa}^j} a_{i_{\ell}} \mathbf{w}^{i_{\ell}}; \quad \sum_{i_{\ell} \in J_{\kappa}^j} a_{i_{\ell}} = 1, \quad a_{i_{\ell}} \in \mathbb{R} \right\}. \quad (22)$$

Thus we may view $L^{\kappa-1}$ as being generated by the vertices of $\sigma^{\kappa-1}$ and $L_{i_j}^{\kappa-2}$ being generated by the vertices of $\sigma_{i_j}^{\kappa-2}$ so that $\sigma_{i_j}^{\kappa-2} = L_{i_j}^{\kappa-2} \cap \sigma^{\kappa-1}$. In the subsequent discussion, we shall assume without loss of generality that the elements of J_{κ} are ordered so that $i_1 < i_2 < \dots < i_{\kappa}$ and set $W^j = W \setminus \{\mathbf{w}^j\}$ and will also utilize the following notion. Given two subsets X and Y of \mathbb{R}^n , the join $X * Y$ of X and Y is defined by

$$X * Y = \{ \alpha \mathbf{x} + \beta \mathbf{y} : \mathbf{x} \in X, \mathbf{y} \in Y; \quad \alpha, \beta \geq 0, \quad \alpha + \beta = 1 \}. \quad (23)$$

The join operation is commutative and associative, and if $X = \{\mathbf{x}\}$ is a one-point set, then $\mathbf{x} * Y$ shall represent $\{\mathbf{x}\} * Y$. To illustrate these concepts and sets under discussion, we consider a couple of examples.

Example 6.1. Suppose that $\kappa = n$ and W is affinely independent. In this case $J_n = J$ and L^{n-1} is a hyperplane in \mathbb{R}^n . Since $\mathbf{w}^j = L^{n-1} \cap L(\mathbf{w}^j)$, L^{n-1} cuts the edges of $\partial F(X)$ and, hence, L^{n-1} cuts $F(X)$. We set $\mathbf{p}^j = L^{n-1} \cap L(\mathbf{m}^j)$, $P = \{\mathbf{p}^j : j = 1, \dots, n\}$, $\sigma^{n-1} = \langle \mathbf{w}^1, \dots, \mathbf{w}^n \rangle$, and let $\langle \mathbf{w}^1, \dots, \bar{\mathbf{w}}^i, \dots, \mathbf{w}^n \rangle = \sigma_i^{n-2}$ denote the $(n-2)$ -dimensional face of σ^{n-1} opposite the vertex \mathbf{w}^i . Here the symbol $\bar{\mathbf{w}}^i$ means that the vertex \mathbf{w}^i has been deleted so that the vertices of σ_i^{n-2} are the elements of the set $W^i = W \setminus \{\mathbf{w}^i\}$. Using the elements of P we form n new sets τ_i by defining $\tau_i = \mathbf{p}^i * \sigma_i^{n-2}$. Similarly, whenever M is affinely independent, then F^{n-1} will denote the hyperplane generated by M , $Q = \{\mathbf{q}^j : j = 1, \dots, n\}$, where $\mathbf{q}^j = F^{n-1} \cap L(\mathbf{w}^j)$, $s_i^{n-1} = \langle \mathbf{m}^1, \dots, \mathbf{m}^n \rangle$, $s_i^{n-2} = \langle \mathbf{m}^1, \dots, \bar{\mathbf{m}}^i, \dots, \mathbf{m}^n \rangle$, and $t_i = \mathbf{q}^i * s_i^{n-2}$. If $n = 2$, then obviously $\tau_i = \sigma_i^0 = \mathbf{p}^i = \mathbf{w}^i$ for $i, j \in \{1, 2\}$ with $i \neq j$. If $n \geq 3$ and X represents a random real data set, then generally $\mathbf{p}^i \neq \mathbf{w}^i$ and $\mathbf{q}^i \neq \mathbf{m}^i$.

Example 6.2. If $\kappa = 2$ and $n \geq 2$, then L^1 is simply the line determined by the two distinct points \mathbf{w}^{i_1} and \mathbf{w}^{i_2} . Since $\langle \mathbf{w}^{i_1}, \mathbf{w}^{i_2} \rangle \subset L^1$ and $\langle \mathbf{w}^{i_1}, \mathbf{w}^{i_2} \rangle \subset F(X)$, $\langle \mathbf{w}^{i_1}, \mathbf{w}^{i_2} \rangle \subset L^1 \cap F(X)$. Also, $F(X) \subset \bar{H}_{\mathbf{w}^{i_1}}(\mathbf{d}_{i_1 i_2}) \cap \bar{H}_{\mathbf{w}^{i_2}}(\mathbf{d}_{i_1 i_2})$, $E_{\mathbf{w}^{i_1}}(\mathbf{d}_{i_1 i_2}) \parallel E_{\mathbf{w}^{i_2}}(\mathbf{d}_{i_1 i_2})$ with $E_{\mathbf{w}^{i_1}}(\mathbf{d}_{i_1 i_2}) \cap E_{\mathbf{w}^{i_2}}(\mathbf{d}_{i_1 i_2}) = \emptyset$, and thus $L^1 \cap [E_{\mathbf{w}^{i_1}}(\mathbf{d}_{i_1 i_2}) \cup E_{\mathbf{w}^{i_2}}(\mathbf{d}_{i_1 i_2})] = \{\mathbf{w}^{i_1}, \mathbf{w}^{i_2}\}$. Therefore, $L^1 \cap F(X) \subset L^1 \cap [\bar{H}_{\mathbf{w}^{i_1}}(\mathbf{d}_{i_1 i_2}) \cap \bar{H}_{\mathbf{w}^{i_2}}(\mathbf{d}_{i_1 i_2})] = \langle \mathbf{w}^{i_1}, \mathbf{w}^{i_2} \rangle$ and, hence, $L^1 \cap F(X) = \langle \mathbf{w}^{i_1}, \mathbf{w}^{i_2} \rangle = \sigma^1$. This representation of $L^{\kappa-1} \cap F(X)$ changes somewhat when $\kappa > 2$. Considering the case $\kappa = 3$, then there are three disjoint pairs of parallel hyperplanes with each pair associated with an edge of $\sigma^2 = \langle \mathbf{w}^{i_1}, \mathbf{w}^{i_2}, \mathbf{w}^{i_3} \rangle$. It can be shown that $L^2 \cap F(X) = \sigma^2 \cup \bigcup_{j=1}^3 \tau_j$, where $\tau_j = \mathbf{p}^j * \sigma_j^1$, and also is equal to

$$L^2 \cap \bigcap_{j=1}^2 \bigcap_{\ell=j+1}^3 \left[\bar{H}_{\mathbf{w}^j}(\mathbf{d}_{i_j i_{\ell}}) \cap \bar{H}_{\mathbf{w}^{\ell}}(\mathbf{d}_{i_j i_{\ell}}) \right]. \quad (24)$$

In general there exist $\kappa(\kappa-1)/2$ pairs of disjoint parallel hyperplanes supporting the simplex $\sigma^{\kappa-1} = \langle \mathbf{w}^{i_1}, \dots, \mathbf{w}^{i_{\kappa}} \rangle$, each associated with a unique one-dimensional face of $\sigma^{\kappa-1}$. Specifically, these hyperplanes are given by $E_{\mathbf{w}^j}(\mathbf{d}_{i_j i_{j+1}}) \parallel E_{\mathbf{w}^{j+1}}(\mathbf{d}_{i_j i_{j+1}}), \dots, E_{\mathbf{w}^j}(\mathbf{d}_{i_j i_k}) \parallel E_{\mathbf{w}^{i_k}}(\mathbf{d}_{i_j i_k})$ for $1 \leq j < k$ or $E_{\mathbf{w}^j}(\mathbf{d}_{i_j i_{j-1}}) \parallel E_{\mathbf{w}^{i_{j-1}}}(\mathbf{d}_{i_j i_{j-1}}), \dots, E_{\mathbf{w}^j}(\mathbf{d}_{i_j i_j}) \parallel E_{\mathbf{w}^{i_j}}(\mathbf{d}_{i_j i_j})$ for $1 < j \leq \kappa$. Setting $\mathbf{p}^j = L(\mathbf{m}^j) \cap L^{\kappa-1}$ for $j = 1, \dots, \kappa$, $P_{\kappa} = \{\mathbf{p}^j : j = 1, \dots, \kappa\}$, $\tau_j = \mathbf{p}^j * \sigma_j^{\kappa-2}$, and assuming that W_{κ} is affinely independent, one can prove that $L^{\kappa-1} \cap F(X) = \sigma^{\kappa-1} \cup \bigcup_{j=1}^{\kappa} \tau_j$ as claimed in Corollary 6.6. Using Theorem 4.6, it turns out that $L^{\kappa-1} \cap F(X)$ is also equal to

$$L^{\kappa-1} \cap \bigcap_{j=1}^{\kappa-1} \bigcap_{\ell=j+1}^{\kappa} \left[\bar{H}_{\mathbf{w}^j}(\mathbf{d}_{i_j i_{\ell}}) \cap \bar{H}_{\mathbf{w}^{\ell}}(\mathbf{d}_{i_j i_{\ell}}) \right]. \quad (25)$$

By setting $S_{\mathbf{w}^j}(i_j, \ell) = E_{\mathbf{w}^j}(\mathbf{d}_{i_j i_{\ell}}) \cap L^{\kappa-1}$, $\bar{S}_{\mathbf{w}^j}^+(i_j, \ell) = \bar{H}_{\mathbf{w}^j}(\mathbf{d}_{i_j i_{\ell}}) \cap L^{\kappa-1}$, $\bar{S}_{\mathbf{w}^j}^-(i_j, \ell) = \bar{H}_{\mathbf{w}^j}(\mathbf{d}_{i_j i_{\ell}}) \cap L^{\kappa-1}$, etc., Eq. (25) can then be rewritten as

$$L^{\kappa-1} \cap F(X) = \bigcap_{j=1}^{\kappa-1} \bigcap_{\ell=j+1}^{\kappa} \left[\bar{S}_{\mathbf{w}^j}^+(i_j, i_{\ell}) \cap \bar{S}_{\mathbf{w}^{\ell}}^-(i_j, i_{\ell}) \right]. \quad (26)$$

Eq. (26) represents the same result in the linear subspace $L^{\kappa-1}$ of \mathbb{R}^n as does Theorem 4.6 in \mathbb{R}^n . Of course, analogous reasoning applies to affine independent subsets of $M_{\kappa} = \{\mathbf{m}^{i_1}, \dots, \mathbf{m}^{i_{\kappa}}\} \subset M$ and $F^{\kappa-1} \cap F(X)$ as well. The next theorem classifies the sets τ_j and t_j as well as the vertex points \mathbf{p}^j and \mathbf{q}^j

Theorem 6.4. If W_κ is affinely independent, then for a given integer $i_j \in J_\kappa$ either τ_{i_j} is the $(\kappa - 1)$ -simplex $\langle \mathbf{w}^{i_1}, \dots, \bar{\mathbf{w}}^{i_j}, \dots, \mathbf{w}^{i_k}, \mathbf{p}^{i_j} \rangle$ or $\tau_{i_j} = \sigma_{i_j}^{\kappa-2}$ and $\mathbf{p}^{i_j} = \mathbf{w}^{i_r}$ for some $i_r \in J_\kappa \setminus \{i_j\}$. Similarly, if M_κ is affinely independent, then t_{i_j} is either the $(\kappa - 1)$ -simplex $\langle \mathbf{m}^{i_1}, \dots, \bar{\mathbf{m}}^{i_j}, \dots, \mathbf{m}^{i_k}, \mathbf{q}^{i_j} \rangle$ or $t_{i_j} = s_{i_j}^{\kappa-2}$ and $\mathbf{q}^{i_j} = \mathbf{m}^{i_r}$ for some $i_r \in J_\kappa \setminus \{i_j\}$.

Theorem 6.5. If W_κ is affinely independent and $\mathbf{w}^\varepsilon \in W$, then $\mathbf{w}^\varepsilon \in L^{\kappa-1}$ if and only if $\mathbf{w}^\varepsilon \in W_\kappa$. Similarly, if M_κ is affinely independent and $\mathbf{m}^\varepsilon \in M$, then $\mathbf{m}^\varepsilon \in F^{\kappa-1}$ if and only if $\mathbf{m}^\varepsilon \in M_\kappa$.

A consequence of Theorem 6.5 is the geometric classification of the sets $F(X) \cap L^{\kappa-1}$ and $F(X) \cap F^{\kappa-1}$.

Corollary 6.6. $F(X) \cap L^{\kappa-1} = \sigma^{\kappa-1} \cup \bigcup_{j=1}^K \tau_{i_j}$ and $F(X) \cap F^{\kappa-1} = s^{\kappa-1} \cup \bigcup_{j=1}^K t_{i_j}$.

Theorem 6.7. Let J_κ be a non-empty subset of $J = \{1, \dots, n\}$. $W_\kappa \subset W$ is affinely independent if and only if $\mathbf{w}^j \neq \mathbf{w}^\ell$ for all distinct pairs $\{i_j, i_\ell\} \subset J_\kappa$. Similarly, $M_\kappa \subset M$ is affinely independent if and only if $\mathbf{m}^j \neq \mathbf{m}^\ell$ for all distinct pairs $\{i_j, i_\ell\} \subset J_\kappa$.

The previous theorems and corollaries provide the tools necessary for establishing the affine independence of sets derived from LAMs. Specifically, Theorem 6.1 together with Theorem 6.7 form the basis of a simple and efficient method for deriving subsets of affine independent vectors from W and M . Thus, for example, if $\mathbf{w}^j \neq \mathbf{w}^\ell$ for any pair $\{j, \ell\} \subset J$, then any non-empty subset of W can serve as a possible set of endmembers. If set X is such that $\mathbf{w}^j = \mathbf{w}^\ell$ for several pairs $\{j, \ell\}$, then an elimination algorithm of columns and rows from W results in an affinely independent set of vectors as described elsewhere [36]. So far, in our tests with hyperspectral image data, the equalities $\mathbf{w}^j = \mathbf{w}^\ell$ or $\mathbf{m}^j = \mathbf{m}^\ell$ for any $j \neq \ell$ do not occur and therefore we can select subsets of W or M as candidate sets of endmembers which are affinely independent.

In most endmember detection schemes, endmembers form a linearly independent set of m points in \mathbb{R}^n and a *dark point* is chosen to obtain a large m -simplex containing most or all of X . In order for the $m + 1$ points to be affinely independent, one must assure that the dark point is not an element of the hyperplane spanned by the $(m - 1)$ -simplex defined by the m affine independent points. Generally, the sets $W \cup \{\mathbf{v}\}$ and $M \cup \{\mathbf{u}\}$ are affine independent. We conjecture that if $\mathbf{v} < \mathbf{u}$, then $\mathbf{v} \notin L^{\kappa-1}$. Thus, we believe that the choice of \mathbf{v} for the dark point remains a good one, especially in view of the fact that it represents the low values of the actual data.

7. Hyperspectral image unmixing

In our approach to endmember determination we consider X to be the set of pixels obtained from the hyperspectral image cube. Recall that the points \mathbf{v} and \mathbf{u} defined by Eq. (18) are, respectively, the minimal and maximal corners of the hyperbox $\mathcal{B}(\mathbf{v}, \mathbf{u})$ enclosing X , and that the scaled vectors \mathbf{w}^j and \mathbf{m}^j forming the sets W and M derived, respectively, from the min- and max-memories \mathbf{W}_{XX} and \mathbf{M}_{XX} are given by Eq. (19). The linear unmixing theory assumes that the set of endmembers are affinely independent. The vector $\mathbf{w}^j \in W$ has the feature that its j th coordinate corresponds to the *maximum measured reflectance* within the j th band of the data set X . In this sense, the elements of W can be viewed as excellent representatives of endmembers of the data cube X . However, it is important to note that for any pair $\{\mathbf{w}^j, \mathbf{m}^i\}$ the inequalities $\mathbf{w}^j \leq \mathbf{m}^i$ or $\mathbf{m}^i \leq \mathbf{w}^j$ are generally false even though $m_j^i \leq w_j^i = u_j = \bigvee_{\xi=1}^k x_\xi^i$ and $\bigwedge_{\xi=1}^k x_\xi^i = v_i = m_i^i \leq w_i^j$. In fact, when using hyperspectral data one usually obtains $w_\ell^j < m_i^i$ for several indices ℓ . For this reason various elements of M may represent important endmembers as demonstrated in this section. From the material presented in Sections 4 and 5, the elements of the set $V = W \cup M \cup \{\mathbf{v}, \mathbf{u}\}$ are vertices of the convex polytope $\mathcal{B}(\mathbf{v}, \mathbf{u}) \cap F(X)$. Thus, the image data cube represented by X is completely contained in $\mathcal{B}(\mathbf{v}, \mathbf{u}) \cap F(X)$ and lends itself to convex hull analysis using the elements of V .

The first step of the WM algorithm is to form the memories \mathbf{W}_{XX} and $\mathbf{M}_{XX} = -\mathbf{W}_{XX}^t$ and the computation of the vectors \mathbf{v} and \mathbf{u} from X . The next step consists of building the sets W and M by scaling the columns of \mathbf{W}_{XX} and \mathbf{M}_{XX} , respectively. The set V then contains all possible endmembers for analyzing X , with \mathbf{v} representing the dark point. To find affinely independent subsets of V we turn our attention to the sets W and M . In many but not necessarily all cases the sets W and M turn out to be affinely and even linearly independent. Specifically, if $\mathbf{w}^j \neq \mathbf{w}^\ell$ and $\mathbf{m}^j \neq \mathbf{m}^\ell$ for all distinct pairs $j, \ell \in \{1, \dots, n\}$, then W and M are affinely independent (see Theorem 6.7). Thus, for example, the set $W \cup \{\mathbf{v}\}$ spans an n -simplex whose vertices \mathbf{w}^j are potential endmembers each having maximal reflectance in the j th band. Sets W and M contain always a fixed but large number of endmembers. However, prior knowledge of the desired number of endmembers or use of alternative endmember reduction methods can be applied to eliminate unimportant candidate endmembers. Another possibility consists of a pre-processing stage where the hyperspectral data cube dimensionality is reduced by application of a chosen technique such as principal component analysis (PCA), minimum noise fraction transform (MNF), or adjacent band removal of highly correlated bands. Such reductions are often necessary in view of storage space, available main memory, and processing speed. For example, a hyperspectral image remotely acquired with the *airborne visible and infrared imaging spectrometer* (AVIRIS) of NASA's Jet Propulsion Laboratory, needs around 134 megabytes of storage space [38].

The aim of the following application examples, based on AVIRIS hyperspectral data, is to provide enough detail to demonstrate the effectiveness of using the dual lattice auto-associative memories, \mathbf{W}_{XX} and \mathbf{M}_{XX} , to determine sets of endmembers from which a subset of *final endmembers* is selected to accomplish hyperspectral image segmentation by means of constrained linear unmixing. Other methods for autonomous endmember determination have been described elsewhere, for example, hierarchical Bayesian models [12], minimum mutual information [17], and pattern elimination based on

minimal Chebyshev distance or angle between vector pairs [35]. At the end of this section, linear unmixing based on the vertex component analysis [28] and the minimum-volume enclosing simplex [7] approaches is described briefly, including a comparison of algorithm characteristics between them and the proposed WM method.

Example 7.1. One of the most studied cases in mineralogical and chemical composition has been the mining site of Cuprite, Nevada. The United States Geographical Survey (USGS) Laboratory has produced detailed maps of mineral and chemical compound distribution at the Cuprite site. The left part of Fig. 2, shows an approximate true color RGB image with wavelengths of 0.67 μm (red), 0.56 μm (green) and 0.48 μm (blue). The right image of Fig. 2 displays a simplified 6-false color map, built with five representative minerals chosen from the mineral groups present in the original USGS image map produced by Clark and Swayze in 1995 (see Table 1; assigned color is a visual aid to distinguish mineral distribution).

The USGS source image map shows only the geographical region of Cuprite where mineral distribution is prominent. It has a width of 534 pixels out of 614 pixels available in a scan line and is formed with the lower 358 scan lines of scene no. 3, followed by all 512 scan lines of scene no. 4, and ends with the upper 102 scan lines of scene no. 5 as registered in flight (1997). The first 80 pixels were dropped from each scan line to obtain the region studied by USGS scientists. The AVIRIS device acquires hyperspectral images in 224 channels; however, only 52 noiseless channels that fall within the short wavelength infrared band are considered for mineral detection. Specifically, channels no. 169 (1.95 μm) to 220 (2.47 μm) are used to match remote spectra against subsampled ground or laboratory spectra, since these last ones are obtained with higher spectral resolution [8]. However, only 48 AVIRIS channels were selected for spectral matching due to the existing difference in wavelength scale between ground and remote spectrometers. Table 2 gives the correspondence between the wavelength scale used by AVIRIS imaging spectrometer and the reference scale of a USGS laboratory spectrometer, showing explicitly which channels at the end of the scale were not considered. For this example, a test data cube associated with the 6-false color map is composed of “pure” pixel spectra corresponding to the minerals listed in Table 1 and illustrated in Fig. 3; each pixel spectra is a 48-dimensional vector.

Computation of the lattice auto-associative memories \mathbf{W}_{xx} and \mathbf{M}_{xx} is quite simple in this case. From Eq. (4), e.g., for $i, j = 1, \dots, n$ with $n = 48$,

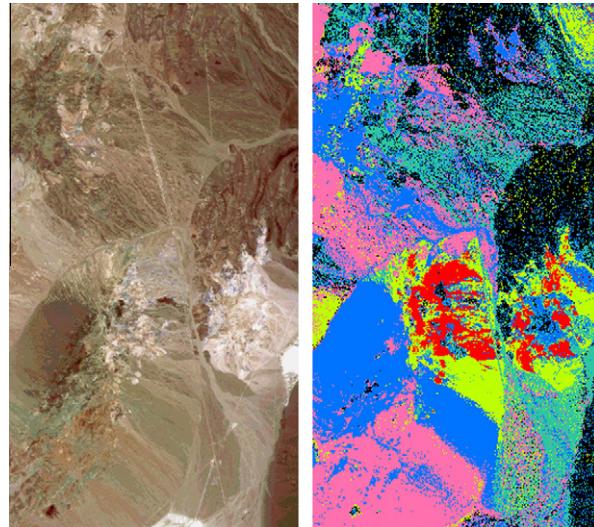


Fig. 2. Left, RGB composite image of the mining site of Cuprite, Nevada. Right, simplified 6-false color mineral distribution map. True size of each image is 534 × 972 pixels.

Table 1
Representative minerals in the cuprite site.

Description	Color	Area (%)
Alunite GDS96 (250C)	red	4.2
Calcite CO 2004	pink	22.6
Kaolinite KGa-2 (pxl)	green	10.8
Montmorillonite SCa-2.b	seagreen	16.1
Muscovite CU93-1 low-Al	blue	27.1
Unclassified	black	19.2

Table 2

AVIRIS and USGS spectrometers wavelength scales in microns (μm). AVIRIS channels not used for spectral sample matching are marked with ‘ \times ’.

Channel	AVIRIS	USGS
212	2.388	2.386
213	2.398	2.400
214	2.408	\times
215	2.418	2.418
216	2.427	\times
217	2.437	2.440
218	2.447	\times
219	2.457	\times
220	2.467	2.466

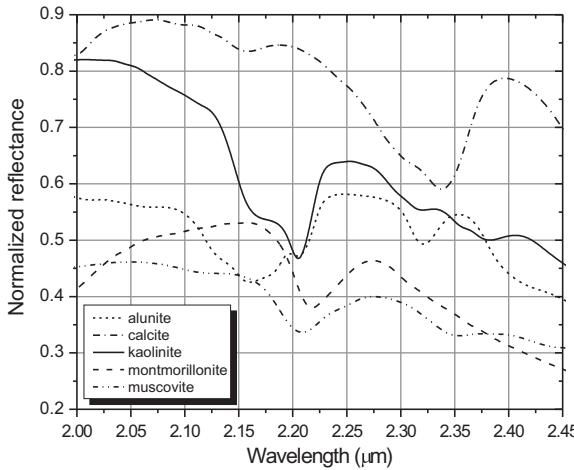


Fig. 3. Spectral curves of the minerals given in Table 1 (true endmembers).

$$w_{ij} = \bigwedge_{\xi=1}^k (x_i^\xi - x_j^\xi) = \bigwedge_{p=1}^m \bigwedge_{\xi \in \Omega_p} (x_i^\xi - x_j^\xi) = \bigwedge_{p=1}^m (x_i^{\xi_p} - x_j^{\xi_p}), \quad (27)$$

where, $k = \sum_{p=1}^m |\Omega_p| = 419,390$ (out of $519,048 = 534 \times 972$ vectors), is the number of non-zero vectors in the data cube, $m = 5$ is the number of “pure” pixels, each Ω_p is the family of indices that correspond to pixels describing the same mineral, and $\xi_p \in \{1, \dots, k\}$ is the first sequential index in the data cube equal to a different spectral vector. The first equality in Eq. (27) is a consequence of the associative property of the minimum binary operation, the second equality follows by *idempotency*. With X reduced to $\{x^{\xi_1}, \dots, x^{\xi_5}\}$, the scaled min- and max-memories are calculated in milliseconds (see Example 7.2), and the 48 columns of each memory matrix give all candidate endmembers. Since this is a test data set for which *a priori knowledge* is available about the hyperspectral image, a simple matching procedure applied to the column vectors $w^j \in W$ and $m^j \in M$, together with the v and u bounds, immediately yields a subset of final endmembers.

Let $Y = W \cup \{u\}$ and, for $p = 1, \dots, 5$ and $q = 1, \dots, 49$, let $c_{pq} = \rho(x^{\xi_p}, y^q)$, where $x^{\xi_p} \in X$ and $y^q \in Y$, be the linear correlation coefficients between true and LAM determined endmembers. In addition, for all p , let $\mu_p = \max_{1 \leq q \leq 49} \{c_{pq}\}$ be the maximum correlation coefficient of each pure pixel against all potential endmembers in Y , and compute

$$b_{pq} = \mathbf{if}(c_{pq} = \mu_p, \mathbf{if}(c_{pq} \geq \alpha, 1, 0), 0), \quad (28)$$

where α is a parameter used to threshold correlation coefficient values and **if** (condition, true, false) is the usual **if-then-else** programming construct expressed as a three argument function. A binary value, $b_{pq} = 1$, gives the row and column indices of a “very good” match between a true endmember (row index) and the potential endmember marked as final (column index). A similar spectral matching procedure is applied to the set $Z = M \cup \{v\}$.

It turns out, that *all five* “pure” pixels used to build the test data cube are correctly determined from both LAMs as final endmembers, a result that is in agreement with the theoretical background developed in Sections 5 and 6. Fig. 4 displays the set of true and final endmembers, where the curve labels give the specific element in $W \cup M \cup \{v, u\}$ whose correlation coefficient, denoted by ρ , is highest if compared to each true endmember (“pure” spectra). Final and true endmember curves are superimposed if a perfect match exists ($\rho = 1$) and, to avoid signature overlapping, the bottom two curves (muscovite) are

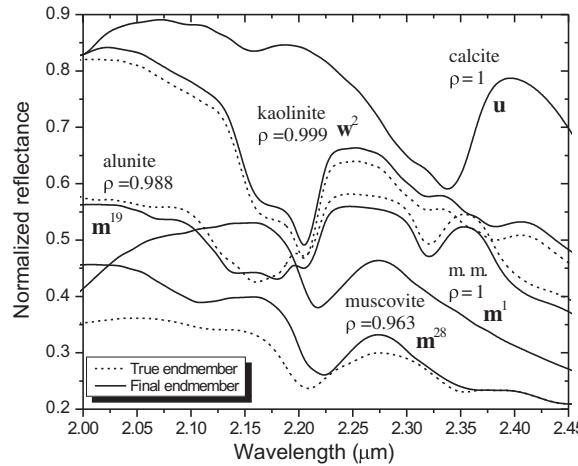


Fig. 4. The solid lines represent the final endmembers determined from the set $W \cup M \cup \{\mathbf{v}, \mathbf{u}\}$ and the dashed lines correspond to the given true endmembers; the mineral montmorillonite is abbreviated as m.m.

displaced -0.1 off their original normalized reflectances. As a final remark, the set $\{\mathbf{v}, \mathbf{m}^1, \mathbf{m}^{19}, \mathbf{m}^{28}, \mathbf{w}^2, \mathbf{u}\}$, where \mathbf{v} is taken as the dark point, is affinely independent and its convex hull forms a 5-simplex that encloses X .

Example 7.2. Moffett Field is a remote sensing test site on the bay of San Francisco, a few kilometers north of the city of Mountain View, California. The site includes the Naval Air Station Moffet Field, agriculture fields, water ponds, salt banks, and man-made constructions such as several airplane hangars which today are museums. The hyperspectral data collected by AVIRIS is ideal for water variability, vegetation and urban studies [2,24,29]. Fig. 5, shows an RGB color image of scene no. 3 with wavelengths of $0.693\text{ }\mu\text{m}$ (red, channel 36), $0.557\text{ }\mu\text{m}$ (green, channel 20) and $0.458\text{ }\mu\text{m}$ (blue, channel 10). Several artificial and natural resources can be extracted from the scene using only 103 noiseless channels that fall within the visible and first half of the short wavelength infrared bands. Specifically, channels no. 4 ($0.40\text{ }\mu\text{m}$) to 106 ($1.34\text{ }\mu\text{m}$) can be used for endmember determination. Thus, the hyperspectral cube of reflectance data is formed by 314,368 pixel spectra (614×512) and each pixel is a 103-dimensional vector.

In this case, the specific format of the AVIRIS data file requires as a first step, the extraction of all pixel spectra line by line to form set $X = \{\mathbf{x}^1, \dots, \mathbf{x}^k\} \subset \mathbb{R}^n$, where $k = 314,368$ and $n = 103$. The second step performs the computation of the vector bounds \mathbf{v}, \mathbf{u} of X and the scaled sets W and M , respectively with Eqs. (29) and (30), for $i, j = 1, \dots, n$; see also Eq. (19).

$$w_{ij} = \bigvee_{\xi=1}^k x_j^\xi + \bigwedge_{\xi=1}^k (x_i^\xi - x_j^\xi), \quad (29)$$

$$m_{ij} = \bigwedge_{\xi=1}^k x_j^\xi + \bigvee_{\xi=1}^k (x_i^\xi - x_j^\xi). \quad (30)$$



Fig. 5. RGB composite “true” color image of scene no. 3 of the Moffett Field site in San Francisco, California. Real size of image is 614×512 pixels.

We point out that the scaling operation generates an “upward spike” in endmembers selected from W since $w_{ii} = u_i$, or a “downward spike” if endmembers come from M because $m_{ii} = v_i$. A simple smoothing procedure to eliminate the anomalous spikes, considers the nearest one or two spectral samples next to w_{ii} or m_{ii} [26]. It is given, for any $i \in \{1, \dots, n\}$, by

$$y_{ii} = \begin{cases} y_{2,1} & \iff i = 1, \\ \frac{1}{2}(y_{i-1,i} + y_{i+1,i}) & \iff 1 < i < n, \\ y_{n-1,n} & \iff i = n, \end{cases} \quad (31)$$

where \mathbf{y} equals \mathbf{w} or \mathbf{m} . Since $Y = W \cup \{\mathbf{u}\}$ or $Z = M \cup \{\mathbf{v}\}$ result in matrices of size 103×104 , each one gives all 104 column vectors as possible endmembers. The LAMs method always gives a number of “candidate” endmembers which is either equal or slightly less than the spectral dimensionality but in practice several contiguous columns are highly correlated. Hence, most of these potential endmembers can be discarded using appropriate techniques.

As an alternative way to automate endmember screening, we calculate a matrix whose entries are given by the linear correlation coefficients, $c_{pq} = \rho(\mathbf{y}^p, \mathbf{y}^q)$ for $p, q = 1, \dots, n$, where $\mathbf{y}^p, \mathbf{y}^q \in Y$ (resp. in Z). A threshold τ value is then applied on c_{pq} to get a subset of selected endmember pairs with low correlation coefficients. For scene no. 3 of the Moffett Field data, $\tau = 0.005$ was applied to Y (resp., $\tau = 0.0005$ for Z). The resulting set is first sorted by ascending column index and after elimination of repeated or contiguous indices, the number of potential endmembers is decreased from 104 to 10. On physical ground, it should be clear, that a low correlation value between potential endmembers, does not necessarily guarantee a clear cut criteria to obtain “good” endmembers useful to match high resolution laboratory spectra. Nevertheless, the reduction in the number of LAM column vectors is quite useful to find a smaller subset of final endmembers. Another, simpler but supervised technique for the lattice based approach forms $\lfloor \sqrt{n+1} \rfloor$ subsets, each with $\lfloor \sqrt{n+1} \rfloor$ column vectors taken from Y (resp. Z), and then, a representative from each group is selected at random as a final endmember. Thus, in this scheme, the number of possible endmembers is always diminished one order of magnitude. Both techniques provide a reasonable number of approximate true endmembers without sacrificing spectral resolution, which, from a physical point of view, is an advantage if compared to data dimensionality reduction techniques such as principal component analysis [13,23] or alternative neural network techniques [27].

Figs. 6 and 7 display the final endmember curves for scene no. 3 of the Moffett Field site determined, respectively, from the scaled versions of \mathbf{W}_{XX} and \mathbf{M}_{XX} . Normalization of reflectance data values in spectral distributions are linearly scaled from the range $[-50, 12000]$ to the unit interval $[0, 1]$. For the problem at hand, the vector bounds \mathbf{v} and \mathbf{u} as well as many other potential endmembers were rejected by the correlation coefficient technique described earlier.

For the unmixing stage, recall that Eq. (1) is an over-determined system of linear equations ($n > m$), subject to the restrictions of full additivity and non-negativity of abundance coefficients as stated in the same equation. In the present case, in matrix W (resp. M), any two column vectors are distinct and therefore its columns forms a linearly independent set of vectors. It turns out that, the set of final endmembers, $E = \{\mathbf{w}^4, \mathbf{w}^{28}, \mathbf{w}^{49}, \mathbf{w}^{78}, \mathbf{m}^1, \mathbf{m}^{24}, \mathbf{m}^{78}\} \subset W \cup M$, is a linear independent set whose pseudo-inverse matrix is unique. Although, the unconstrained solution corresponding to Eq. (1), where $n = 103 > m = |E| = 7$, has a single solution, some coefficients are negative for many pixel spectra and do not sum up to unity. If full additivity is enforced, again negative coefficients appear. Therefore, in this case, a reasonable choice consists of imposing non-negativity for the abundance proportions and simultaneously, relaxing full additivity by considering the inequality $\sum_{k=1}^m a_k < 1$. Specifically, we use the non-negative least squares (NNLS) algorithm that solves the problem of minimizing the Euclidean norm expressed by $\|\mathbf{Sa} - \mathbf{x}\|_2$, subject to the condition $\mathbf{a} \geq \mathbf{0}$. The details related to the NNLS algorithm can be found in [21,25]. Each abundance map was obtained with the NNLS numerical method as implemented in Matlab and,

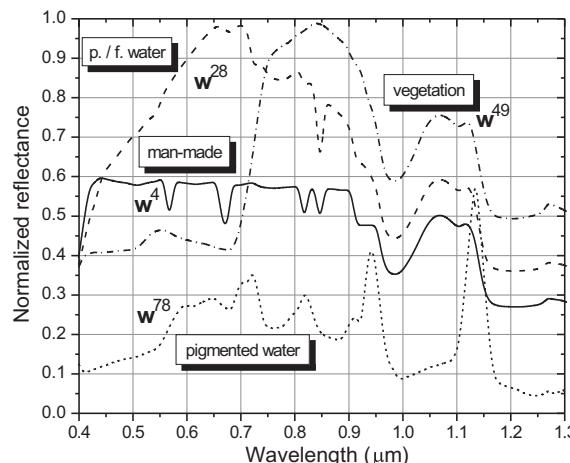


Fig. 6. Spectral curves of final endmembers determined from Y , with corresponding resource classification; p./f. = pigmented/fresh.

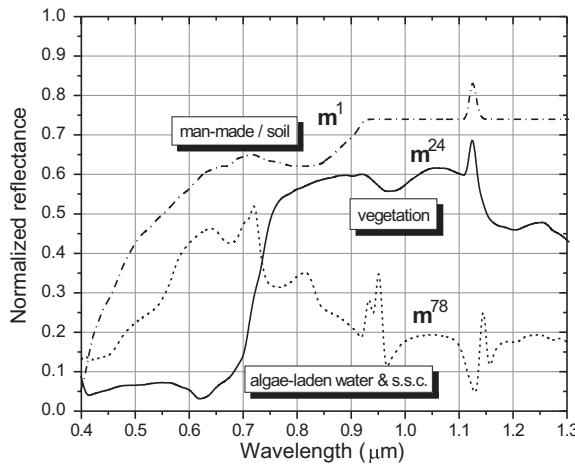


Fig. 7. Spectral curves of final endmembers determined from Z , with corresponding resource classification; s.s.c. = suspended sediment concentration.

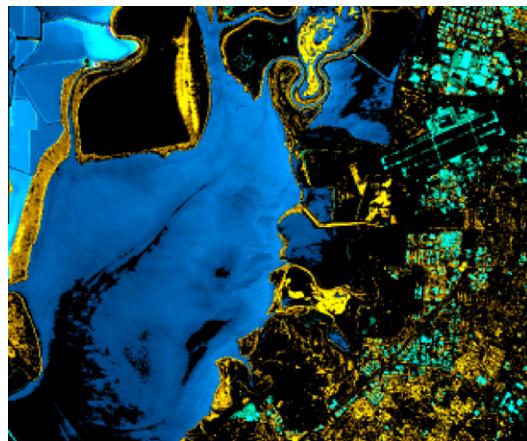


Fig. 8. Combined abundances obtained from w^4 , w^{28} , and w^{49} . Cyan colors indicate artificial or man-made resources, e.g., urban settlements. Blue colors show pigmented and fresh water distribution. The upper left bright blue regions correspond to evaporation ponds pigmented by red brine shrimp. Yellow colors indicate vegetation distribution, e.g., shore-line and in-land vegetation. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

for visual clarity, each image map has been contrast enhanced by the application of a non-linear increasing function and abundance fractions below 0.2625 were set to zero. Figs. 8 and 9 display as false RGB color images the combined abundance distribution corresponding, respectively to the final endmembers shown in Figs. 6 and 7. The abundance map obtained from w^{78} was not incorporated in Fig. 8 due to its similarity to the abundance map generated from endmember w^{28} . In this case, both spectral distributions provide information about the presence of pigmented water in wet soil or wet vegetation areas.

Note that in the present situation, we are faced with *no a priori* knowledge about the hyperspectral image content, and therefore, approximate resource classification is possible after the abundance maps are computed. However, a minimal base knowledge of hyperspectral image analysis as well as a fundamental background in spectral analysis is required for proper recognition of spectral signature characteristics in single, multiple or mixed resources. For example, in the abundance distribution shown as cyan colors in Fig. 9 (final endmember, m^1), the presence of man-made or soil is clearly differentiated. Thus, the identification labels that appear in Figs. 6 and 7 are the result of an overall spectral scrutiny, aided by a visual match between an RGB “true” color reference image and each final endmember corresponding abundance map obtained by constrained linear unmixing. Precise resource identification would require an exhaustive matching procedure between each final endmember that belongs to E against high resolution signatures available in professional spectral libraries [8,20]. Table 3 resumes the time spent by a low end personal computer (1.8 GHz processor, 512 Mb RAM, 2 Gb virtual RAM, and 80 Gb hard disk), for each computational task performed in the analysis of scene no. 3 of the Moffet Field hyperspectral image. Note that tasks 5 and 7 in the analysis process are of an interactive nature but rely, respectively, on tasks 4 and 6 which are completely automatic. Specifically, tasks 2, 3, and 4 are the fundamental steps of the lattice auto-associative

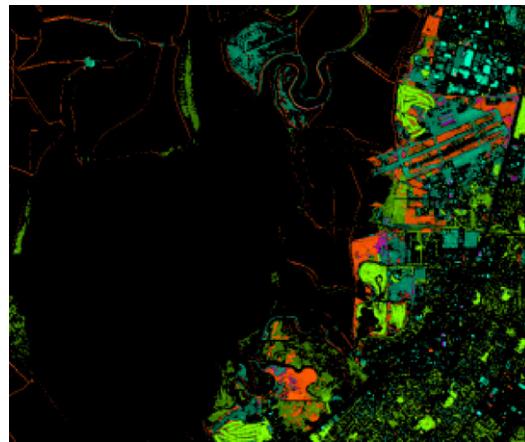


Fig. 9. Combined abundances obtained from \mathbf{m}^1 , \mathbf{m}^{24} , and \mathbf{m}^{78} . Cyan and seagreen colors reveal both man-made resources and different kinds of soil. Large irregular bright green regions with grooves correspond to golf courses. Red colors show possibly algae-laden water with some concentration of suspended sediment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Processing times in minutes for Example 7.2.

Time	Task
11.23	1 Pixel spectra extraction
5.65	2 Vector bounds & LAMs computation
0.01	3 LAMs linear independence test
2.52	4 Correlated endmembers elimination
–	5 Final endmember selection
32.00	6 NNLS abundance map generation
–	7 General resource classification

memories based technique whose computational effort is minimum in comparison to the processing times needed to perform the other tasks.

For a hyperspectral image of size $N = p \times q$ pixels acquired over n spectral bands, the computational effort required for pixel spectra extraction (task 1) is linear in N since $n \ll N$. The overall computational complexity of the LAMs based technique (tasks 2, 3, 4) is $n^2(N+3)$ which for values of n of a few hundreds is in the order of minutes (see Table 3). Endmember determination (task 5) relies on a subset of $2\lfloor\sqrt{n+1}\rfloor$ candidate extremal vectors generated after task 4 is completed and, although interactive in nature, $m \propto \lfloor\sqrt{n}\rfloor$ “final” endmembers can be selected in a lapse of minutes. On the other hand, the NNLS method needs about nm^3 arithmetical operations to find a unique set of abundance coefficients for each pixel spectra. Since m represents the number of final endmembers and $m \ll N$, it turns out that for hyperspectral image segmentation the computational complexity of the non-negative least square method (task 6) is nm^3 per pixel. Finally, general resource classification (task 7) may also be accomplished within minutes whenever a working knowledge in spectral identification or prior experience in hyperspectral image analysis is available for the problem at hand [22].

Two recent approaches to linear spectral unmixing, comparable in performance to N-FINDR and based on the geometry of convex sets and convex optimization algorithms are *vertex component analysis* (VCA) [28] and the *minimum-volume enclosing simplex* (MVES) [7]. Vertex component analysis is an unsupervised technique that relies on singular value decomposition (SVD) and principal component analysis (PCA) as subprocedures, and assumes the existence of pure pixels. Specifically, VCA exploits the fact that endmembers are vertices of a simplex and that the affine transformation of a simplex is again a simplex. This algorithm iteratively projects data onto a direction orthogonal to the subspace generated by the endmembers already determined. The new endmember spectrum is the extreme of the projection and the main loop continues until all given endmembers are exhausted. Similarly, the minimum-volume enclosing simplex is an autonomous technique supported on a linear programming (LP) solver but does not require the existence of pure pixels in the hyperspectral image. However, when pure pixels exist, the MVES technique leads to unique identification of endmembers. In particular, dimension reduction is accomplished by affine set fitting, and Craig's unmixing criterion [9] is applied to formulate hyperspectral unmixing as an MVES optimization problem. Table 4 gives the main characteristics of the VCA and MVES convex geometry based algorithms as well as the lattice algebra approach based on the W & M vector sets derived respectively from the W_{XX} and M_{XX} LAMs.

Note that the proposed lattice algebra based method *does not* require to know in advance the number of endmembers as specified in row 1 of Table 4. In row 2 of the same table, the numerical procedures used by VCA are SVD for projections onto a subspace of dimension m and PCA for projections onto a subspace of dimension $m - 1$. Algorithm MVES first determines the

Table 4

Characteristics of three autonomous linear unmixing techniques. (1) input data, (2) endmember search numerical methods, (3) endmember search computational complexity, (4) and (5) abundance map generation algorithm and type.

	VCA	MVES	W & M
1	X, m	X, m	X
2	SVD, PCA	APS, LPs	LAMs, CMs
3	$2m^2N$	$\alpha m^2 N^{1.5}$	$n^2(N + 3)$
4	M-P inverse	M-P inverse	LNNS
5	unconstrained	unconstrained	constrained

affine parameters set (APS), solves by LP an initial feasibility problem with linear convex constraints, and iteratively optimizes two LP problems with non-convex objective functions. The WM algorithm computes first the scaled min- and max LAMs and their corresponding correlation coefficient matrices (CMs) for further endmember discrimination. Abundance coefficients are determined in the VCA and MVES algorithms using the Moore–Penrose (M–P) pseudo-inverse and, as explained earlier, the WM algorithm makes use of the non-negative least squares method (LNNS). Computational complexity given in row 3 is expressed as a function of the number of endmembers m , the number of pixels or observed spectra N , the number of iterations α , and the number of spectral bands n . These expressions assume that $N \gg n \approx m^2$ and consider only the computations required by the numerical endmember search methods. Although the M–P inverse operation is faster than the LNNS numerical method for the calculation of the abundance coefficients for all pixels in the input image X , we favored the LNNS algorithm since it enforces fraction positivity and consequently allows for a better rendering of the corresponding abundance maps.

We remark that since VCA and MVES require to know in advance the number m of endmembers to be found, their application to a real hyperspectral image must probe all values of m in a specified interval, e.g., from 1 to m_{\max} . Hence, if no *a priori* information is known about the number of pure pixels existing in a hyperspectral image, the computational performance for finding endmembers and determining their abundance fractions increases respectively, to $2N\sum_{m=1}^{m_{\max}} m^2$ and $\alpha N^{1.5} \sum_{m=1}^{m_{\max}} m^2$, which are proportional to $m_{\max}^3 N$ and $\alpha m_{\max}^3 N^{1.5}$. Therefore, in practical situations, overall complexity of our proposed WM method is lower by one order of magnitude than VCA though better in performance than MVES. Furthermore, the VCA and MVES Matlab codes provided by their respective authors were applied to specific subimages of Examples 7.1 and 7.2 producing similar results to those obtain by the WM technique.

8. Conclusions

We presented a new tool to find a set of endmembers in a given hyperspectral image. The mathematical foundation of the autonomous stage in the proposed method is based on several algebraic and geometric properties derived from the set of fixed points of lattice auto-associative memories. Given a set X of n -dimensional pixel spectra, we have shown that the sets W and M derived from the column vectors of the \mathbf{W}_{XX} and \mathbf{M}_{XX} lattice matrices, including the bounds $\{\mathbf{v}, \mathbf{u}\}$, provide all candidate endmembers. In fact, an important contribution of the proposed method is that each candidate endmember is related geometrically with the given hyperspectral data since the j th coordinate of \mathbf{w}^j corresponds to the maximal measured reflectance in the j th band and similarly, the j th coordinate of \mathbf{m}^j corresponds to the minimal measured reflectance in the j th band. Determination of all potential endmembers is computationally fast since only a single scan of the hyperspectral image cube X is required. However, determination of final endmembers relies on appropriate reduction techniques, such as, discrimination by non-contiguous maximal difference correlation values or random selection of a single representative endmember in contiguous equal sized blocks of candidate endmembers. We remark that the extraction of “pure” pixels is problem dependent and needs a certain degree of *a priori* knowledge for proper spectral identification of resource materials against available spectral libraries. In the final stage, abundance coefficients were computed using the NNLS numerical algorithm from which resource distribution images have been illustrated for a typical application example in hyperspectral imagery. Comparison with other recent developments such as VCA and MVES methods shows that our WM method is an alternative competitive solution for hyperspectral image unmixing.

Acknowledgments

Gonzalo Urcid thanks the National System of Research (SNI) and the National Council of Science and Technology (CONACYT) in Mexico for partial support through Grant No. 22036.

References

- [1] J.B. Adam, M.O. Smith, P.E. Johnson, Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site, *Journal of Geophysical Research* 91 (B8) (1986) 8098–8112.
- [2] J.C. Anderson, L. Gosink, M.A. Duchaineau, K.I. Joy, Feature identification and extraction in function fields, in: Proceedings Eurographics/IEEE-VGTC Symposium on Visualization, 2007, pp. 195–201.
- [3] L.M. Batten, A. Rentelspacher, *The Theory of Finite Linear Spaces*, Cambridge University Press, Cambridge, UK, 1993.
- [4] G. Birkhoff, *Lattice Theory*, third ed., American Mathematical Society, Providence, RI, 1967.

- [5] J.W. Boardman, Automated spectral unmixing of AVIRIS data using convex geometry concepts, in: Proceedings AVIRIS Workshop, vol. 4, JPL Publication, 1993, pp. 11–14.
- [6] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, New York, NY, 2008.
- [7] T-H. Chan, C-Y. Chi, Y-M. Huang, W-K. Ma, A convex analysis based minimum-volume enclosing simplex algorithm for hyperspectral unmixing, *IEEE Transactions on Signal Processing* 57 (11) (2009) 4418–4432.
- [8] R.N. Clark, G.A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, S.J. Sutley, USGS digital spectral library splib06a: US Geological Survey, Digital Data Series 231, <<http://speclab.cr.usgs.gov/spectral.lib06>>, 2007.
- [9] M.D. Craig, Minimum-volume transforms for remotely sensed data, *IEEE Transactions on Geoscience and Remote Sensing* 32 (3) (1994) 542–552.
- [10] R. Critescu, Topological Vector Spaces, Nordhoff International Publishing, Leyden, Holland, 1977.
- [11] R. Cuninghame-Green, Minimax algebra and applications, in: P. Hawkes (Ed.), Advances in Imaging and Electron Physics, vol. 90, Academic Press, San Diego, CA, 1995, pp. 1–121.
- [12] N. Dobigeon, J.Y. Tourneret, C.I. Chang, Semi-supervised linear spectral unmixing using a hierarchical Bayesian model for hyperspectral imagery, IRIT-ENSEEIH-TéSA, France, Tech. Report, 2007, pp. 1–34.
- [13] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second ed., John Wiley and Sons, New York, 2000.
- [14] H.G. Eggleston, Convexity, Cambridge University Press, Cambridge, 1963.
- [15] M. Graña, P. Sussner, G.X. Ritter, Associative morphological memories for endmember determination in spectral unmixing, in: Proceedings of FUZZ-IEEE, 2003, pp. 1285–1290.
- [16] M. Graña, C. Hernández, J. Gallego, A single individual evolutionary strategy for endmember search in hyperspectral images, *Information Sciences* 161 (3–4) (2004) 181–197.
- [17] M. Graña, J.L. Jiménez, C. Hernández, Lattice independence, auto-associative morphological memories and unsupervised segmentation of hyperspectral images, in: Proceedings 10th Joint Conference on Information Sciences, 2007, pp. 1624–1631.
- [18] M. Graña, I. Villaverde, J.O. Maldonado, C. Hernández, Two lattice computing approaches for the unsupervised segmentation of hyperspectral images, *Neurocomputing* 72 (10–12) (2009) 2111–2120.
- [19] H. Hadwiger, Vorlesungen Über Inhalt, Oberfläche und Isoperimetrie, Springer, Berlin, 1957.
- [20] S.J. Hook, ASTER spectral library: Jet Propulsion Laboratory (<<http://speclib.jpl.nasa.gov>>), California Institute of Technology, Pasadena, CA, 1999.
- [21] Kh.D. Ikramov, M. Matin-far, Computer-algebra implementation of the least squares method on the non-negative orthant, *Journal of Mathematical Sciences* 132 (2) (2006) 156–159.
- [22] J.R. Jensen, Remote Sensing of the Environment: An Earth Resource Perspective, second ed., Prentice-Hall Series in Geographic Information Systems, Upper Saddle River, NJ, 2007.
- [23] T. Korenius, J. Laurikkala, M. Juhola, On principal component analysis, cosine, and Euclidean measures in information retrieval, *Information Sciences* 177 (22) (2007) 4893–4905.
- [24] F.A. Kruse, L.L. Richardson, V.G. Ambrosia, Techniques developed for geological analysis of hyperspectral data applied to near-shore hyperspectral ocean data, in: Proceedings of the Fourth International Conference on Remote Sensing for Marine and Coastal Environments, vol. I, Ann Arbor, MI, 1997, pp. 233–246.
- [25] C.L. Lawson, R.J. Hanson, Solving Least Squares Problems, Prentice-Hall, Englewood Cliffs, NJ, 1974.
- [26] D.S. Myers, Hyperspectral Endmember Detection Using Morphological Autoassociative Memories, M.S. Thesis, University of Florida, Gainesville, FL, 2005, pp. 1–51.
- [27] G. Myklebust, E. Steen, Neural networks in visualization of multispectral medical images, *Information Sciences* 89 (3–4) (1996) 297–307.
- [28] J.M.P. Nascimento, J.M. Bioucas-Dias, Vertex component analysis: a fast algorithm to unmix hyperspectral data, *IEEE Transactions on Geoscience and Remote Sensing* 43 (4) (2005) 898–910.
- [29] L.L. Richardson, D. Buisson, C.J. Liu, V.G. Ambrosia, The detection of algal photosynthetic accessory pigments using AVIRIS spectral data, *Marine Technology Society Journal* 28 (1994) 10–21.
- [30] G.X. Ritter, P. Sussner, J.L. Díaz de León, Morphological associative memories, *IEEE Transactions on Neural Networks* 9 (2) (1998) 281–293.
- [31] G.X. Ritter, Lattice algebra and minimax algebra, in: Image Algebra, CISE Department, University of Florida, Gainesville, FL, 1999 (unpublished manuscript available via anonymous FTP).
- [32] G.X. Ritter, J.L. Díaz de León, P. Sussner, Morphological bidirectional associative memories, *Neural Networks* 12 (1999) 851–867.
- [33] G.X. Ritter, G. Urcid, L. Iancu, Reconstruction of patterns from noisy inputs using morphological associative memories, *Journal of Mathematical Imaging and Vision* 19 (2) (2003) 95–111.
- [34] G.X. Ritter, P. Gader, Fixed points of lattice transforms and lattice associative memories, in: P. Hawkes (Ed.), Advances in Imaging and Electron Physics, vol. 144, Academic Press, San Diego, CA, 2006, pp. 165–242.
- [35] G.X. Ritter, G. Urcid, M.S. Schmalz, Autonomous single-pass endmember approximation using lattice auto-associative memories, *Neurocomputing* 72 (10–12) (2009) 2101–2110.
- [36] G.X. Ritter, G. Urcid, Lattice algebra approach to endmember determination in hyperspectral imagery, in: P. Hawkes (Ed.), Advances in Imaging and Electron Physics, vol. 160, Academic Press, Burlington, MA, 2010, pp. 113–169.
- [37] D.A. Roberts, M. Gardner, R. Church, S. Ustin, G. Sheer, R.O. Green, Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models, *Remote Sensing Environment* 65 (3) (1998) 267–279.
- [38] G. Vane, R.O. Green, T.G. Chrien, H.T. Enmark, E.G. Hansen, W.M. Porter, The airborne visible/infrared imaging spectrometer (AVIRIS), *Remote Sensing Environment* 44 (1993) 127–143.
- [39] M.E. Winter, Fast autonomous spectral endmember determination in hyperspectral data, in: Proceedings 13th International Conference on Applied Remote Sensing, vol. 2, 1999, pp. 367–344.
- [40] M.E. Winter, An algorithm for fast autonomous spectral endmember determination in hyperspectral analysis, in: Imaging Spectrometry, Proceedings of SPIE 3753, 1999, pp. 266–275.