Convex Geometry for Blind Source Separation

Wing-Kin (Ken) Ma

Department of Electronic Engineering,
The Chinese University of Hong Kong

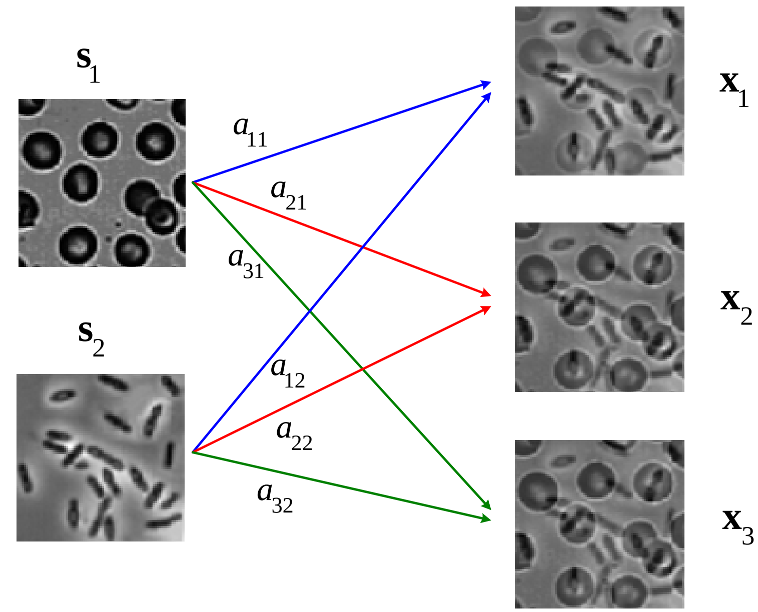
November 23, 2012

Convex Geometry for Non-Negative Blind Source Separation

Blind source separation (BSS): Problem statement

Signal model: a real-valued, N-input, M-output linear mixing model:

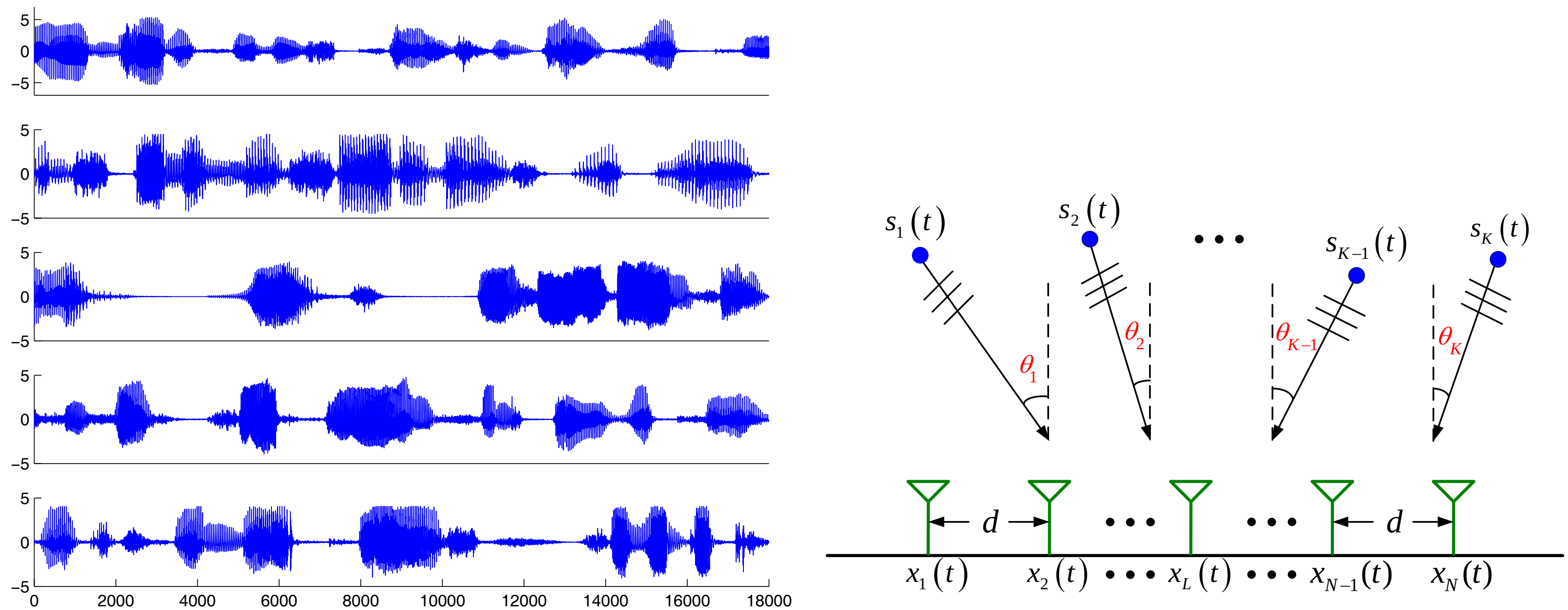
$$x_i[n] = \sum_{j=1}^{N} a_{ij} s_j[n], \quad n = 1, \ldots, L,$$



where
$x_i[n]$ is the ith observed signal, $i = 1, \ldots, M$;
$s_j[n]$ is the jth source signal, $j = 1, \ldots, N$.

Problem: recover the source signals the observed signals, without information of the mixing matrix $\mathbf{A} = \{a_{ij}\}$.

Blind speech and audio separation: A classical BSS application

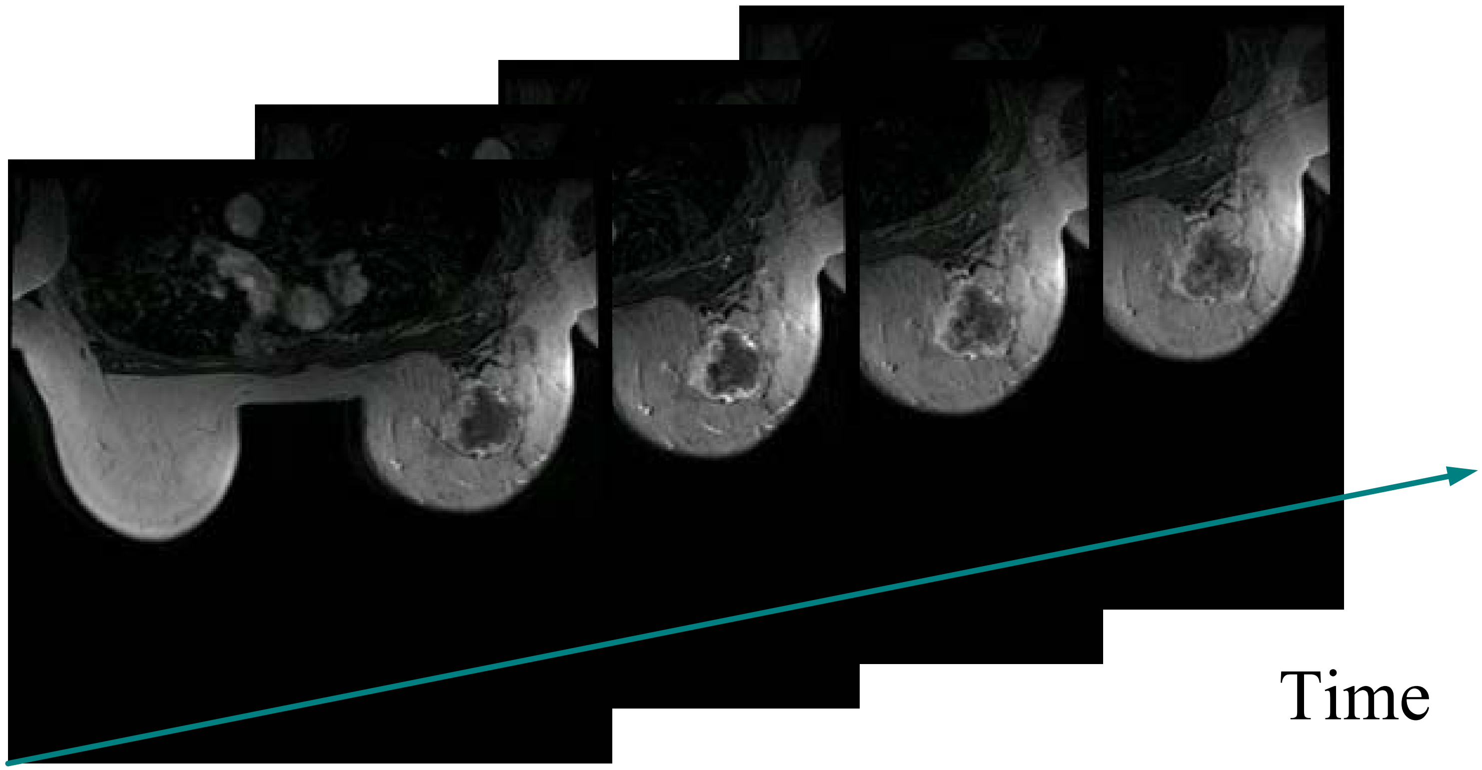


Speech and audio separation in a microphone array.

- The problem is to separate multiple speakers' voices using an array of microphones.

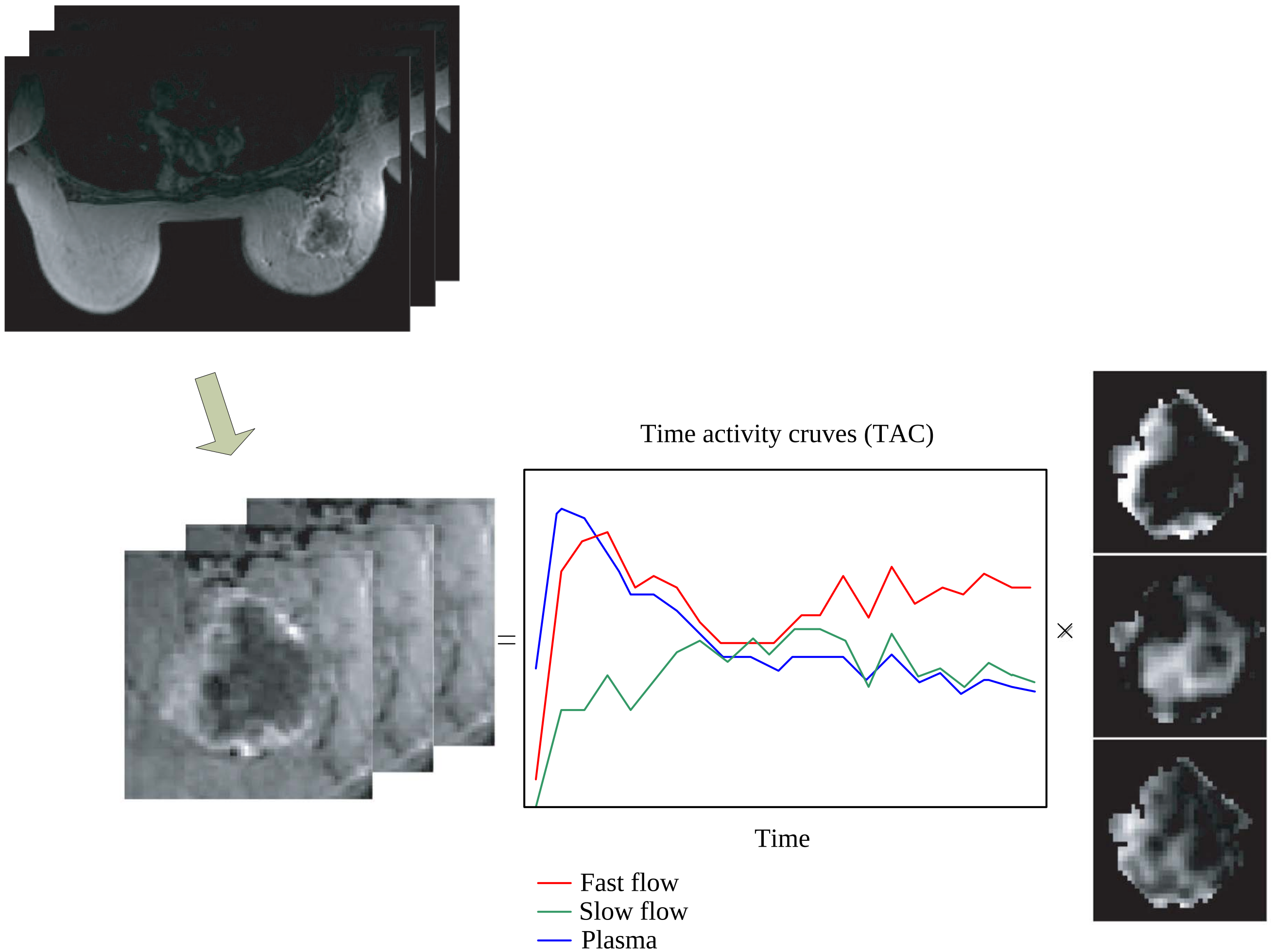
- The challenge is that the location and propagation characteristics of each speaker are not known. This results in a blind problem.

Applications in biomedical imaging



Dynamic contrast-enhanced MRI. Courtesy to [Wang *et al.* 2003].

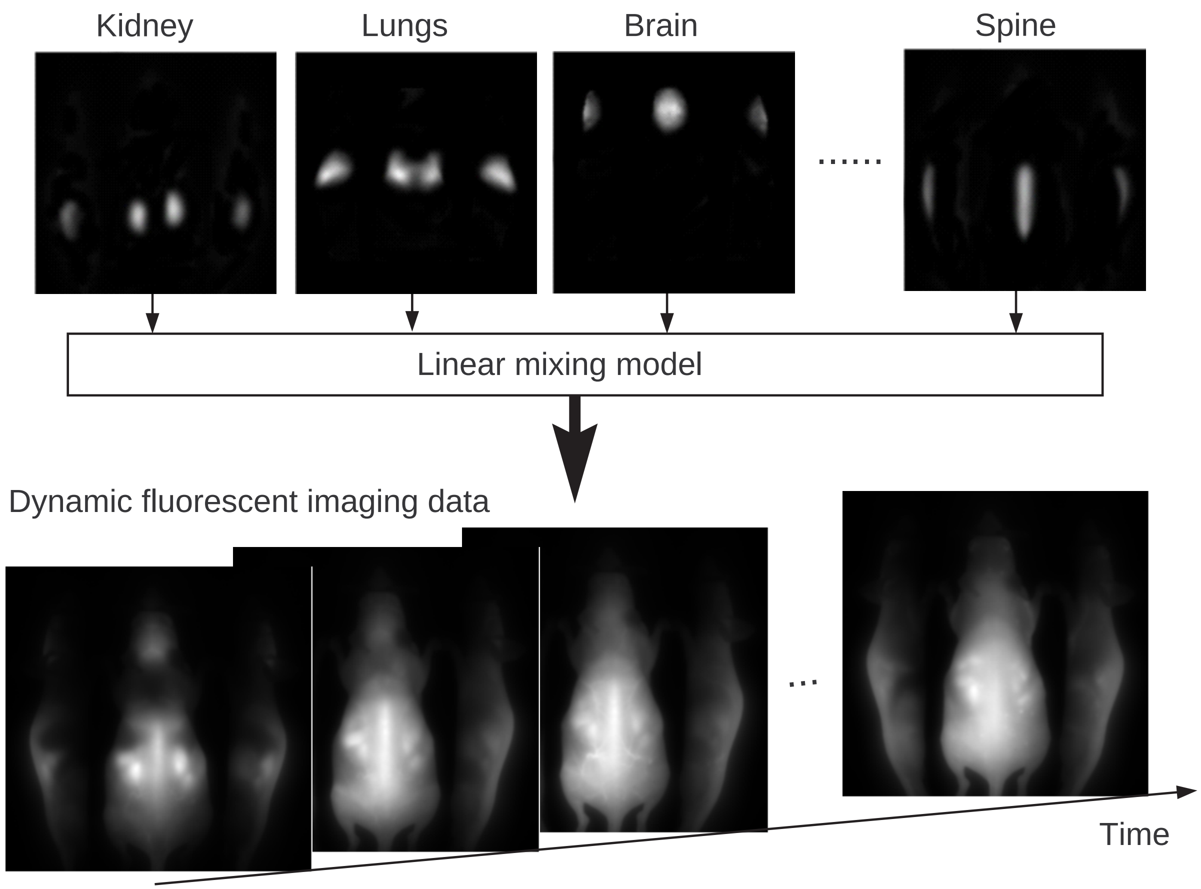
- Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) uses various molecular weight contrast agents to assess tumor vascular perfusion and permeability.



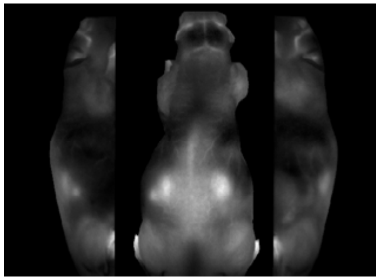
Courtesy to [Wang *et al.* 2003].

- DCE-MRI images are often linear mixtures of more than one distinct vasculature sources, since many malignant tumors show heterogeneous areas of permeability.

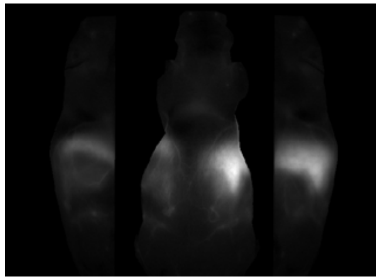
Another biomedical imaging example



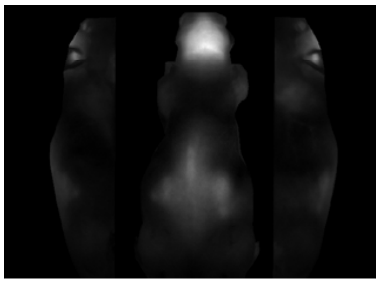
- Dynamic fluorescent imaging (DFI) exploits highly specific and bio-compatible fluorescent contrast agents to interrogate small animals for drug development and disease research.
- DFI images are linear mixtures of the anatomical maps of different organs.



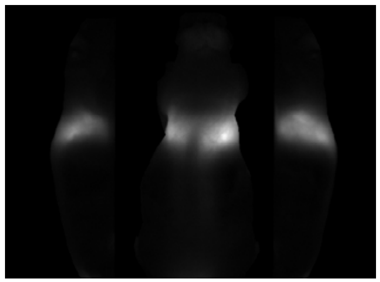
Kidney



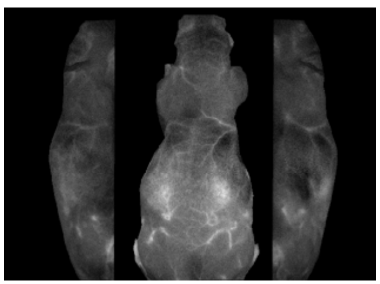
Liver



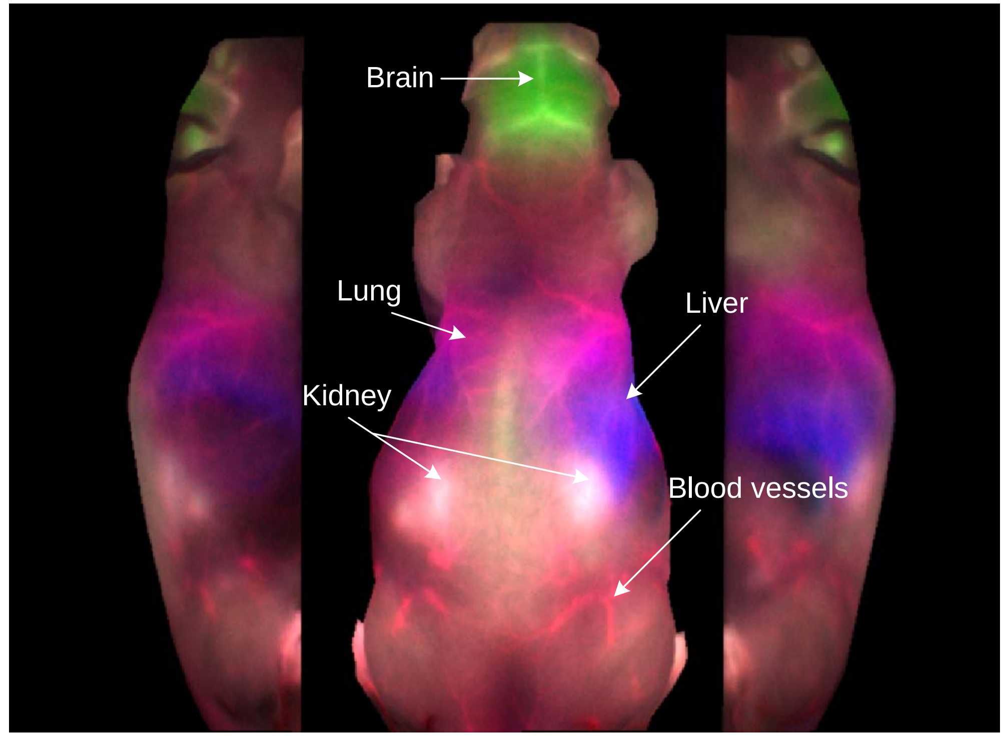
Brain



Lung



Blood vessels



Anatomical map

Separated anatomical maps, using a convex geometry-based method.

BSS techniques

- Blind source separation is *not* completely blind.
- All BSS approaches make specific assumptions on the characteristics of $\{s_i[n]\}_n$ and/or $\mathbf{A}$, and utilize them to achieve blind separation.
- The suitability of the assumptions (& the approach as a result) depends much on the applications under consideration.

Example: Independent component analysis (ICA), a well-known BSS framework, typically assumes that each $s_i[n]$ is random non-Gaussian & is mutually independent.

Mutual independence is a good assumption in speech & audio applications, but not so in hyperspectral imaging.

Non-negative blind source separation (nBSS)

- In some applications source signals are non-negative; e.g., imaging.
- nBSS approaches exploit the signal non-negativity characteristic (plus some additional assumptions).
- **Applications:**
  - biomedical imaging,
  - hyperspectral imaging,
  - analytical chemistry,
  - and most recently, speech separation [Fu-Ma'12].
- **nBSS frameworks:**
  - ICA with non-negativity incorporated; e.g., [Plumbley 2003],
  - non-negative matrix factorization (NMF) [Lee-Seung 1999],
  - **convex geometry**.

Convex Geometry

- Interestingly, different disciplines came up with similar intuitive thinking of convex geometry over different times.
  - chemometrics [Perczel *et. al*'89],
  - hyperspectral remote sensing [Craig'94],
  - nuclear magnetic resonance spectroscopy [Naanaa-Nuzillard'05],
  - SP theory and methods [Chan-Ma-Chi-Wang'08] (we got our first motivation from DCE-MRI, though).
- Our study uses convex analysis and optimization to establish rigorous signal processing frameworks for convex geometry-based nBSS.

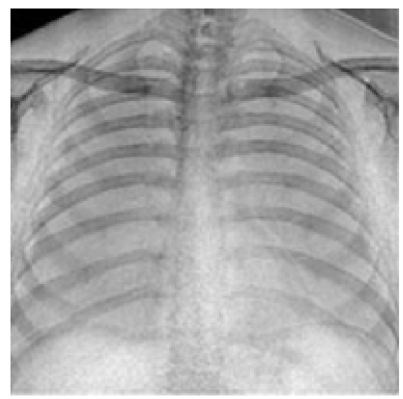
CAMNS:
Convex analysis of mixtures of non-negative sources

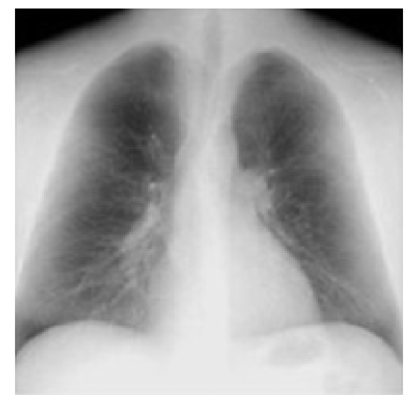
- CAMNS [Chan-Ma-Chi-Wang'08] is an nBSS approach based on convex geometry.

- Unlike ICA which is a statistical framework, convex geometry is deterministic.

- In addition to utilizing source non-negativity, CAMNS employs a special assumption called **local dominance**.

- What is local dominance? Intuitively, signals with many 'zeros' are likely to satisfy local dominance (math. def. available soon).





- Practically, we found it a good assumption for sparse or high-contrast images.

Intuition behind CAMNS

- Recall the linear mixture model

$$x_i[n] = \sum_{j=1}^{N} a_{ij} s_j[n],$$

$i = 1, \ldots, M$, $n = 1, \ldots, L$.

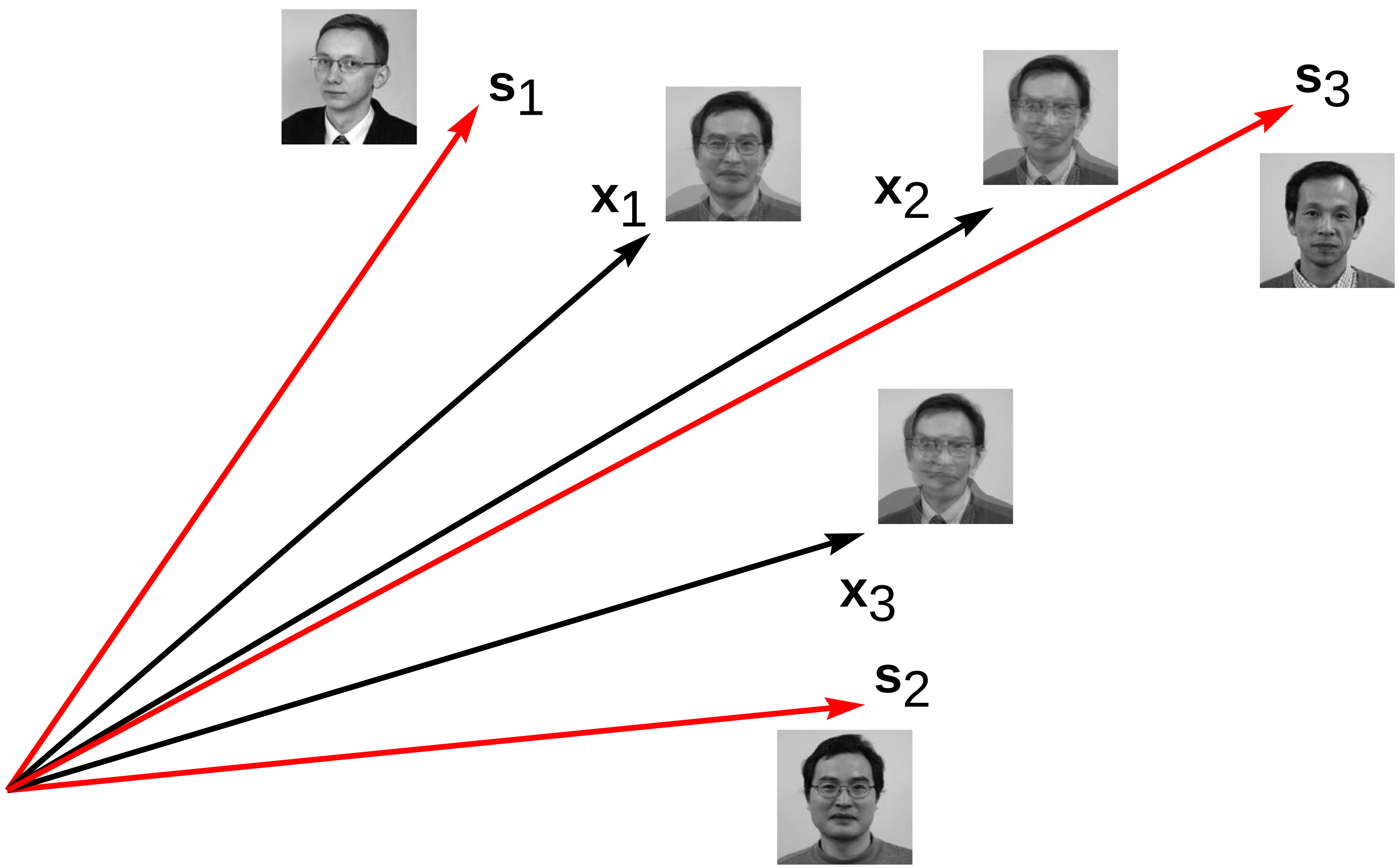
- Define

$$\boldsymbol{x}_i = \begin{bmatrix} x_i[1] \\ \vdots \\ x_i[L] \end{bmatrix}, \quad \boldsymbol{s}_i = \begin{bmatrix} s_i[1] \\ \vdots \\ s_i[L] \end{bmatrix}.$$

We can write

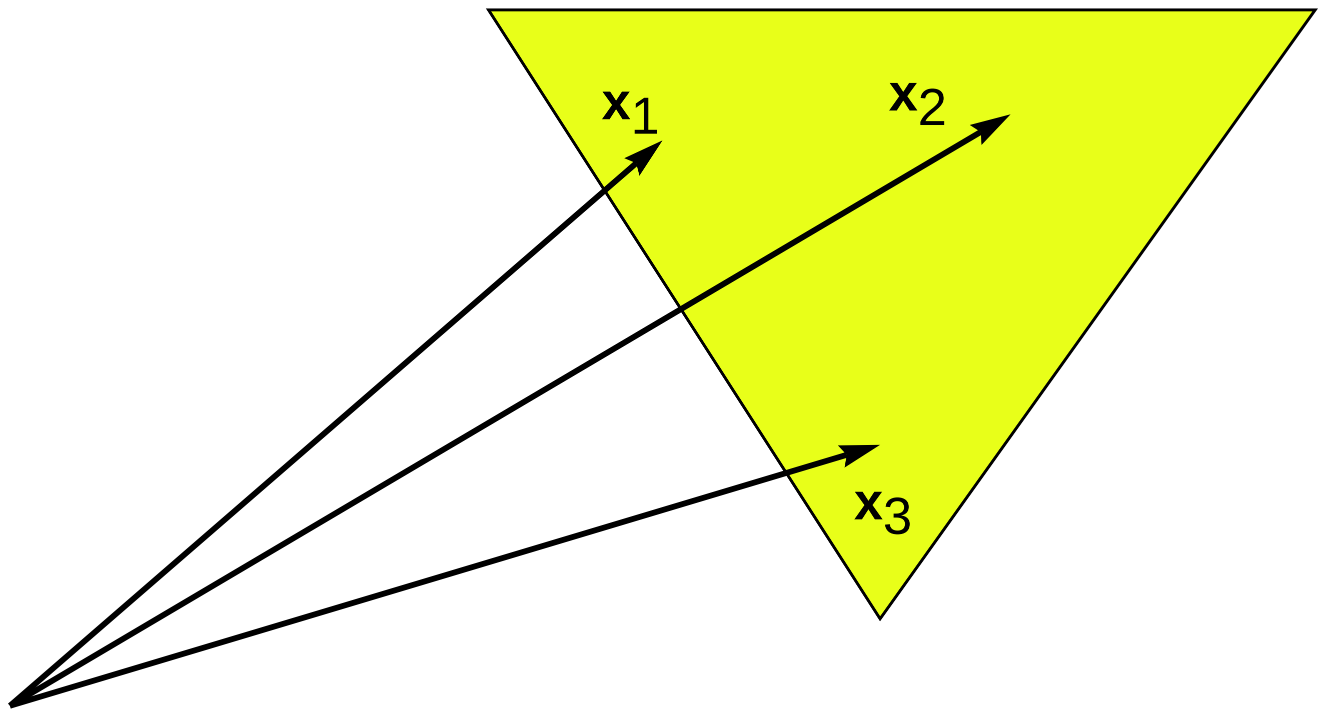
$$\boldsymbol{x}_i = \sum_{j=1}^{N} a_{ij} \boldsymbol{s}_j.$$

Intuition behind CAMNS (cont'd)



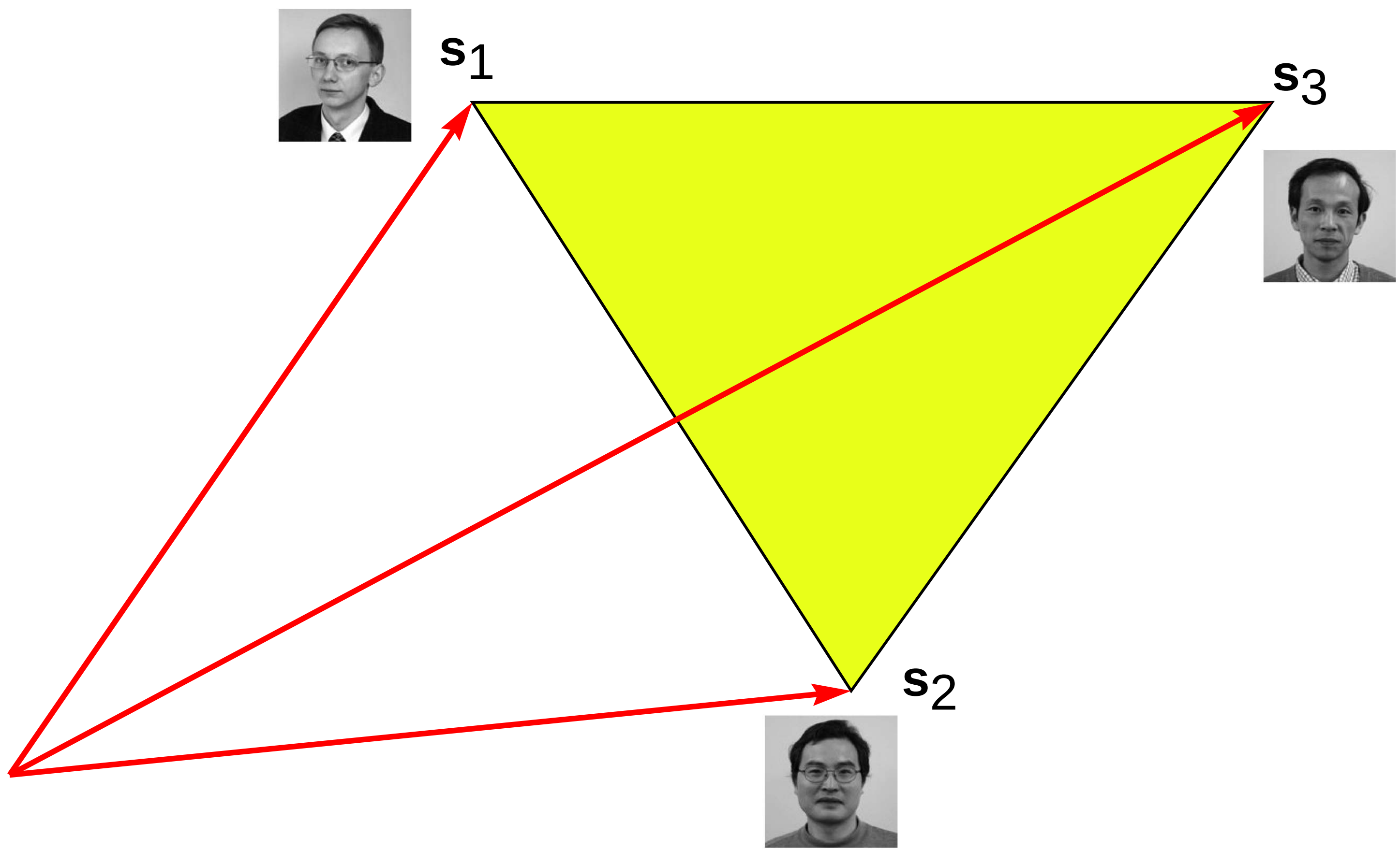
A vector space illustration of $\boldsymbol{x}_i = \sum_{j=1}^{N} a_{ij}\boldsymbol{s}_j$. How can we extract $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$ from $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}$ without knowing $\{a_{ij}\}$?

Intuition behind CAMNS (cont'd)



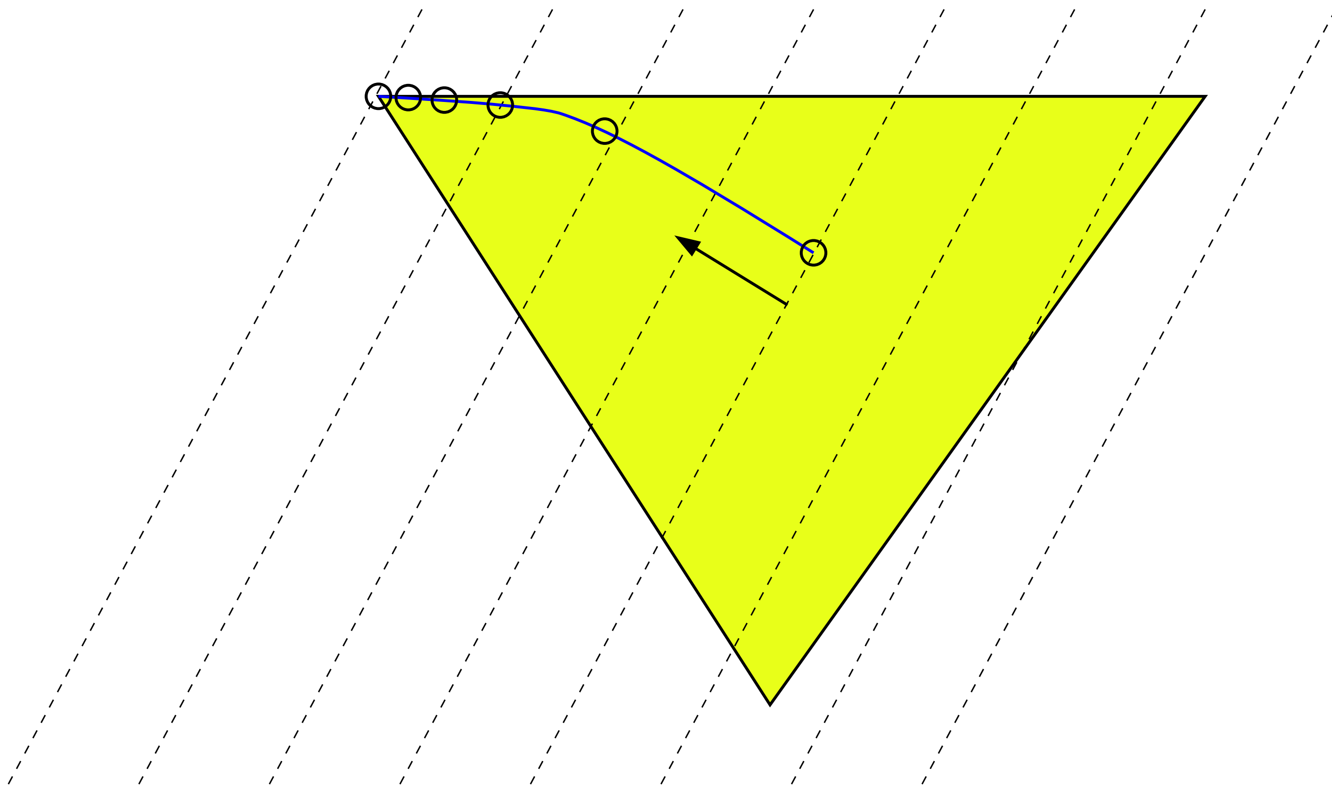
Based on some assumptions (e.g., signal non-negativity & local dominance) & by convex analysis, we use $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}$ to construct a polyhedral set.

Intuition behind CAMNS (cont'd)



We show that the 'corners' (formally speaking, extreme points) of this polyhedral set are exactly $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$.

Intuition behind CAMNS (cont'd)

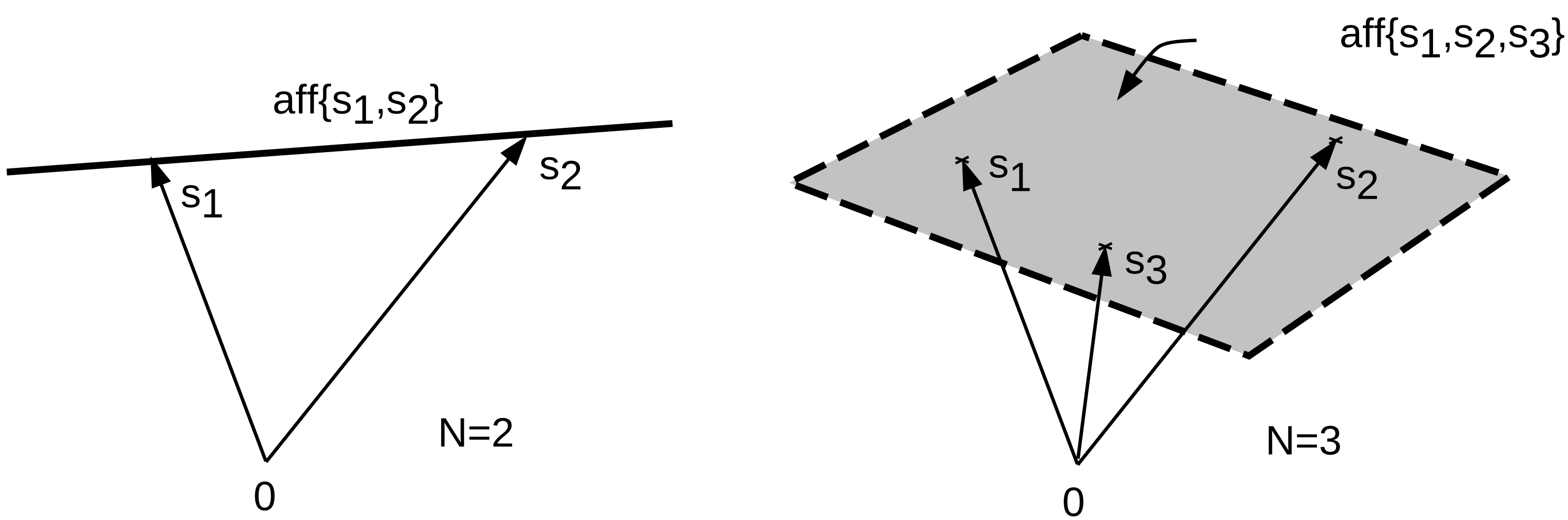


Using LP, we can locate the 'corners' of the polyhedral set effectively. As a result perfect separation can be achieved.

A quick review of some convex analysis concepts

The **affine hull** of a given set of vectors $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} \subset \mathbb{R}^L$ is defined as:

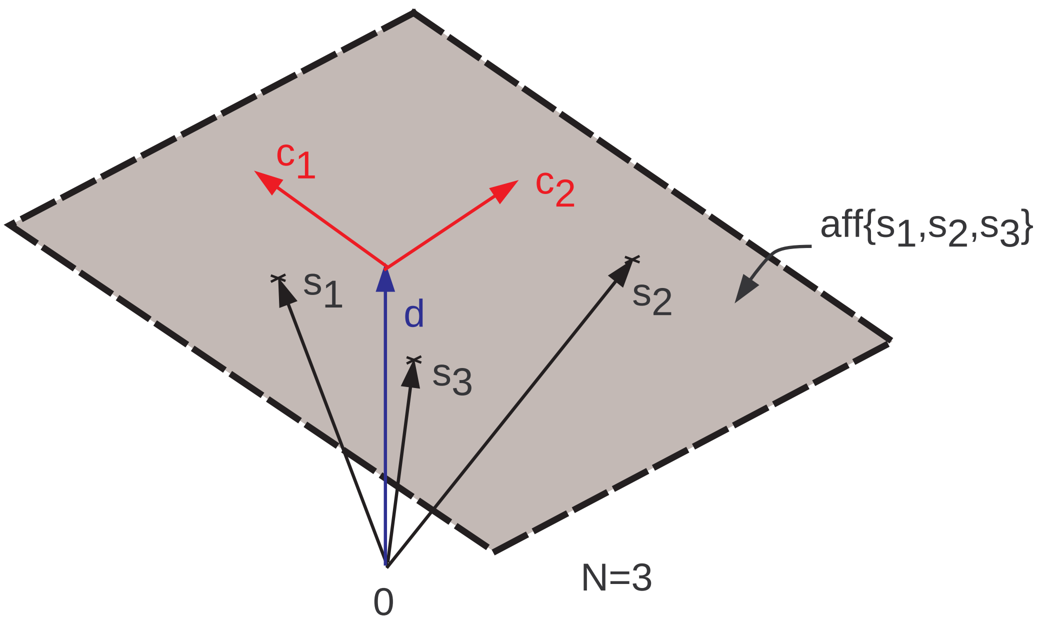
$$\mathrm{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} = \left\{ \boldsymbol{x} = \sum_{i=1}^{N} \theta_i \boldsymbol{s}_i \;\middle|\; \boldsymbol{\theta} \in \mathbb{R}^N, \sum_{i=1}^{N} \theta_i = 1 \right\}.$$



An affine hull $\text{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} = \{ \ \boldsymbol{x} = \sum_{i=1}^N \theta_i \boldsymbol{s}_i \mid \boldsymbol{\theta} \in \mathbb{R}^N, \sum_{i=1}^N \theta_i = 1 \ \}$ can always be expressed as

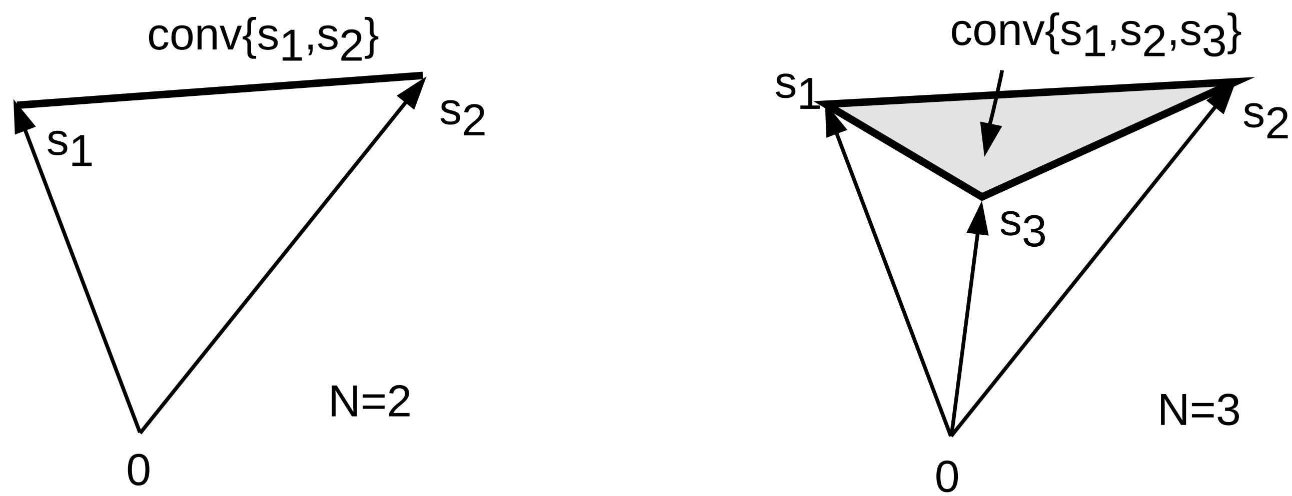
$$\text{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} = \left\{ \ \boldsymbol{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \ \middle| \ \boldsymbol{\alpha} \in \mathbb{R}^P \ \right\},$$

for some (non-unique) $\mathbf{d} \in \mathbb{R}^L$ and $\mathbf{C} \in \mathbb{R}^{L \times P}$, where $P \leq N - 1$ is the affine dimension.



The **convex hull** of a given set of vectors $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} \subset \mathbb{R}^L$ is defined as

$$\text{conv}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} = \left\{ \boldsymbol{x} = \sum_{i=1}^{N} \theta_i \boldsymbol{s}_i \;\middle|\; \boldsymbol{\theta} \in \mathbb{R}_+^N, \sum_{i=1}^{N} \theta_i = 1 \right\}$$



- A point $\boldsymbol{x} \in \text{conv}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$ is an **extreme point** of $\text{conv}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$ if $\boldsymbol{x}$ is not any nontrivial convex combination of $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$.

The assumptions in CAMNS

Recall the model $\boldsymbol{x}_i = \sum_{j=1}^{M} a_{ij} \boldsymbol{s}_j$. Our assumptions:

(A1) Source non-negativity: For each j, $\boldsymbol{s}_j \in \mathbb{R}_+^L$.

(A2) Local dominance: For each $i \in \{1, \ldots, N\}$, there exists an (unknown) index ℓ_i such that $s_i[\ell_i] > 0$ and $s_j[\ell_i] = 0$, $\forall j \neq i$.

(Reasonable assumption for sparse or high-contrast signals).

(A3) Unit row sum: For all $i = 1, \ldots, M$, $\sum_{j=1}^{N} a_{ij} = 1$.

(Already satisfied in MRI, can be relaxed).

(A4) $M \geq N$ and $\mathbf{A}$ is of full column rank. (Standard BSS assumption)

How to enforce (A3), if it does not hold

The unit row sum assumption (A3) may be relaxed.

- Suppose that $\boldsymbol{x}_i^T \mathbf{1} \neq 0$ (where $\mathbf{1}$ is an all-one vector) for all i.
- Consider a normalized version of $\boldsymbol{x}_i$:

$$\bar{\boldsymbol{x}}_i = \frac{\boldsymbol{x}_i}{\boldsymbol{x}_i^T \mathbf{1}} = \sum_{j=1}^{N} \underbrace{\left(\frac{a_{ij} \boldsymbol{s}_j^T \mathbf{1}}{\boldsymbol{x}_i^T \mathbf{1}} \right)}_{\triangleq \bar{a}_{ij}} \underbrace{\left(\frac{\boldsymbol{s}_j}{\boldsymbol{s}_j^T \mathbf{1}} \right)}_{\triangleq \bar{\boldsymbol{s}}_j}.$$

- It can be shown that $(\bar{a}_{ij})$ satisfies (A3).

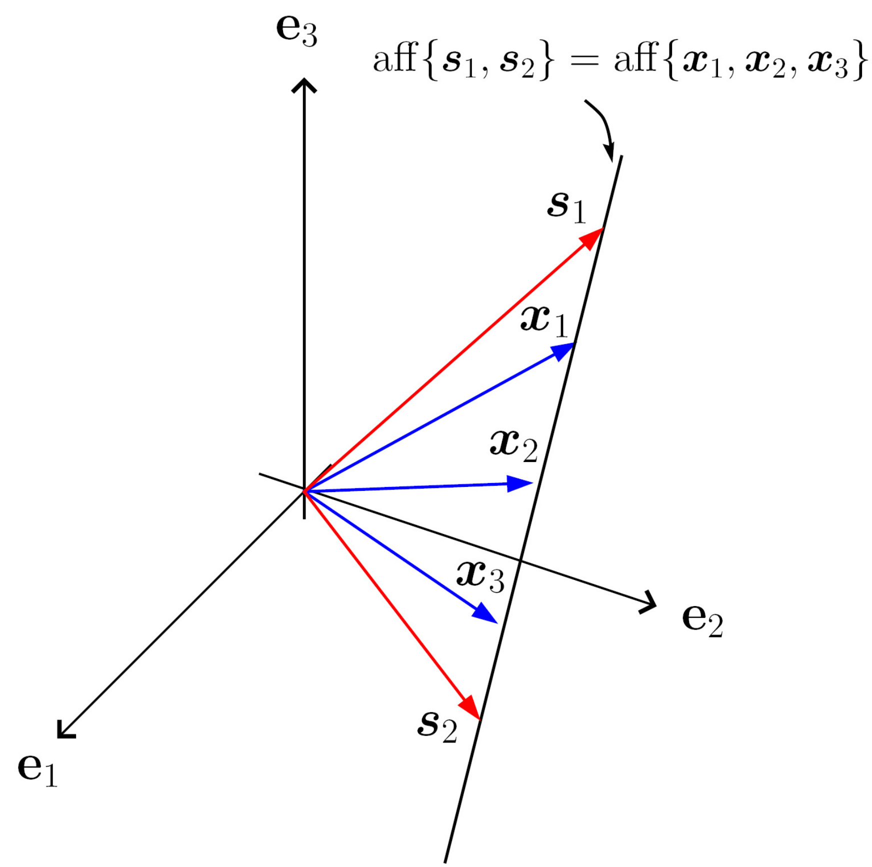
CAMNS

Since $\sum_{j=1}^{N} a_{ij} = 1$ [(A3)], we have for each observation

$$\boldsymbol{x}_i = \sum_{j=1}^{N} a_{ij} \boldsymbol{s}_j \in \text{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$$

This implies

$$\text{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} \supseteq \text{aff}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}.$$



In fact, we can show that

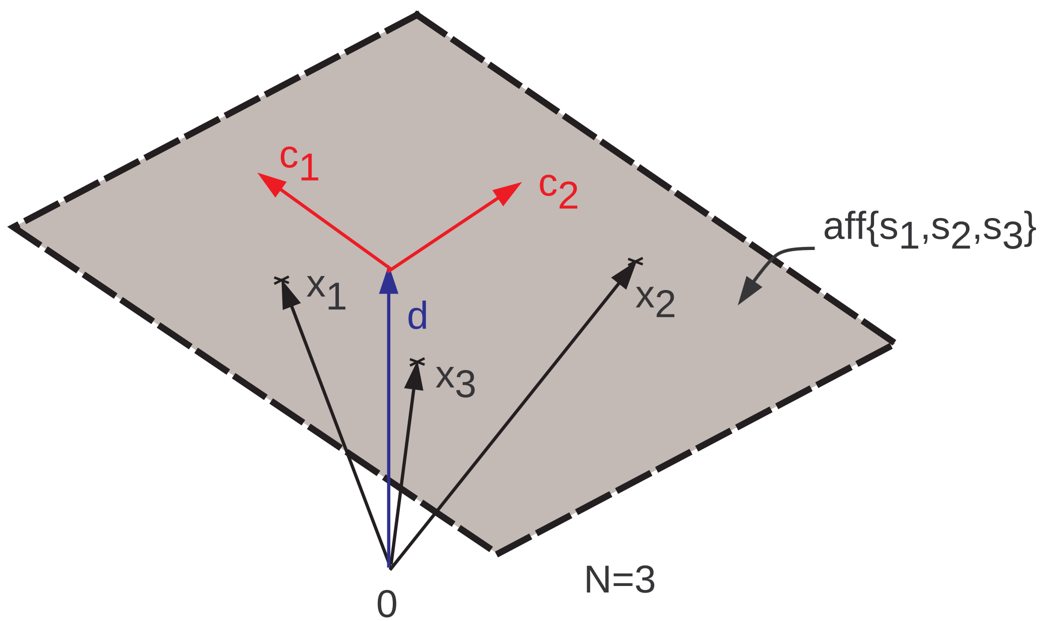
Lemma. *Under* (A3) *and* (A4), $\text{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} = \text{aff}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}$.

- Consider the representation

$$\begin{aligned}\text{aff}\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_N\} &= \text{aff}\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\} \\ &= \left\{\, \boldsymbol{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \;\middle|\; \boldsymbol{\alpha} \in \mathbb{R}^{N-1} \,\right\} \triangleq \mathcal{A}(\mathbf{C},\mathbf{d})\end{aligned}$$

for some $(\mathbf{C},\mathbf{d}) \in \mathbb{R}^{L\times(N-1)} \times \mathbb{R}^L$ with $\text{rank}(\mathbf{C}) = N-1$.

- Let us consider determining the source affine set parameters $(\mathbf{C},\mathbf{d})$ from $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_M\}$.



- Consider the representation

$$\begin{aligned}\mathrm{aff}\{\boldsymbol{s}_1,\ldots,\boldsymbol{s}_N\} &= \mathrm{aff}\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N\}\\ &= \left\{\ \boldsymbol{x}=\mathbf{C}\boldsymbol{\alpha}+\mathbf{d}\ \middle|\ \boldsymbol{\alpha}\in\mathbb{R}^{N-1}\ \right\} \triangleq \mathcal{A}(\mathbf{C},\mathbf{d})\end{aligned}$$

  for some $(\mathbf{C},\mathbf{d})\in\mathbb{R}^{L\times(N-1)}\times\mathbb{R}^L$ with $\mathrm{rank}(\mathbf{C})=N-1$.

- Let us consider determining the source affine set parameters $(\mathbf{C},\mathbf{d})$ from $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_M\}$.

- By solving an affine set fitting problem, we show that

$$\mathbf{d}=\frac{1}{M}\sum_{i=1}^{M}\boldsymbol{x}_i, \qquad \mathbf{C}=[\ \boldsymbol{q}_1(\mathbf{U}\mathbf{U}^T),\boldsymbol{q}_2(\mathbf{U}\mathbf{U}^T),\ldots,\boldsymbol{q}_{N-1}(\mathbf{U}\mathbf{U}^T)\ ]$$

  where $\mathbf{U}=[\ \boldsymbol{x}_1-\mathbf{d},\ldots,\boldsymbol{x}_M-\mathbf{d}\ ]\in\mathbb{R}^{L\times M}$, and $\boldsymbol{q}_i(\mathbf{R})$ denotes the eigenvector associated with the ith principal eigenvalue of $\mathbf{R}$.

- As a coincidence, affine set fitting is reminiscent of principal component analysis.

Be reminded that $\boldsymbol{s}_i \in \mathbb{R}_+^L$. Hence, it is true that

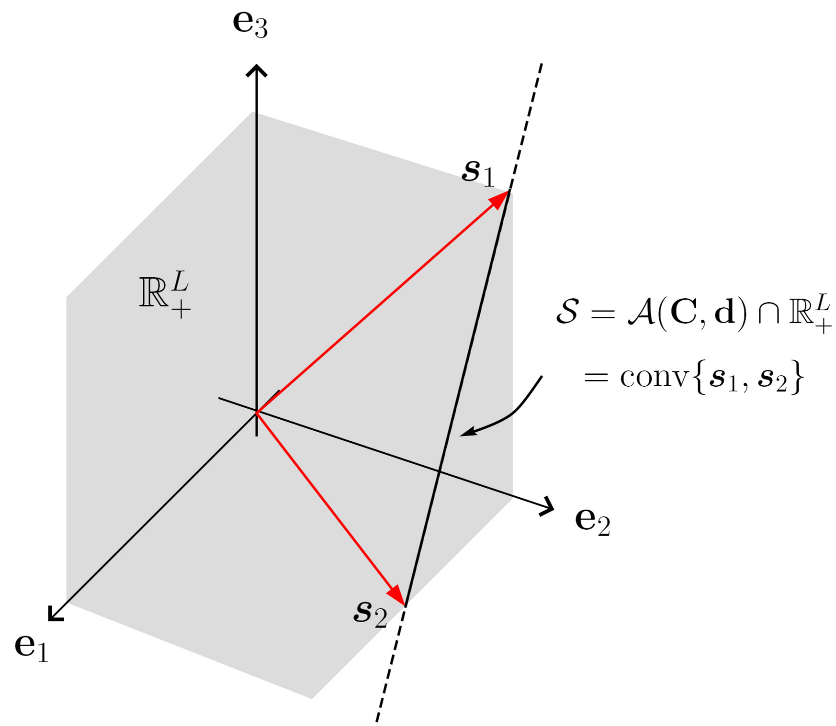
$$\boldsymbol{s}_i \in \operatorname{aff}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} \cap \mathbb{R}_+^L = \mathcal{A}(\mathbf{C}, \mathbf{d}) \cap \mathbb{R}_+^L \triangleq \mathcal{S}$$

The following lemma arises from local dominance (A2):

Lemma. *Under* (A1) *and* (A2),

$$\mathcal{S} = \operatorname{conv}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$$

Moreover, the set of all its extreme points is $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\}$.



Summarizing the above results, a new nBSS criterion is as follows:

Theorem. **(CAMNS criterion)** Under (A1)-(A4), the polyhedral set

$$\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^L \;\middle|\; \boldsymbol{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \succeq \mathbf{0},\ \boldsymbol{\alpha} \in \mathbb{R}^{N-1} \right\}$$

where $(\mathbf{C}, \mathbf{d})$ is obtained from the observation set $\{\boldsymbol{x}_1, ..., \boldsymbol{x}_M\}$ by affine set fitting, has N extreme points given by the true source vectors $\boldsymbol{s}_1, ..., \boldsymbol{s}_N$.

Practical realization of CAMNS

- CAMNS boils down to finding all the extreme points of an observation-constructed polyhedral set.

- In the optimization context this is known as vertex enumeration.

- In CAMNS, there is one important problem structure that we can take full advantage of; that is,

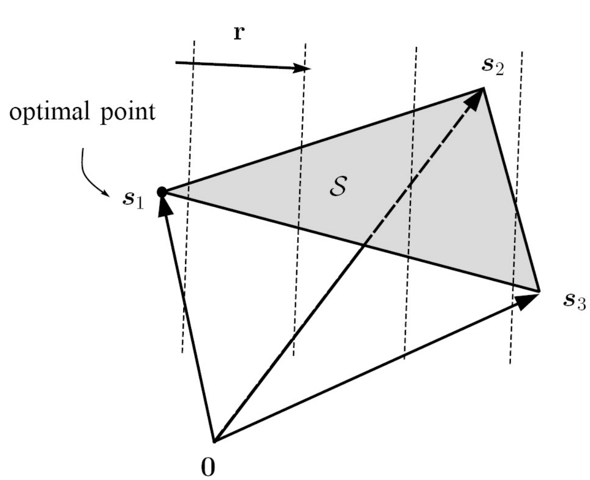
Property implied by (A2): $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$ are linear independent.

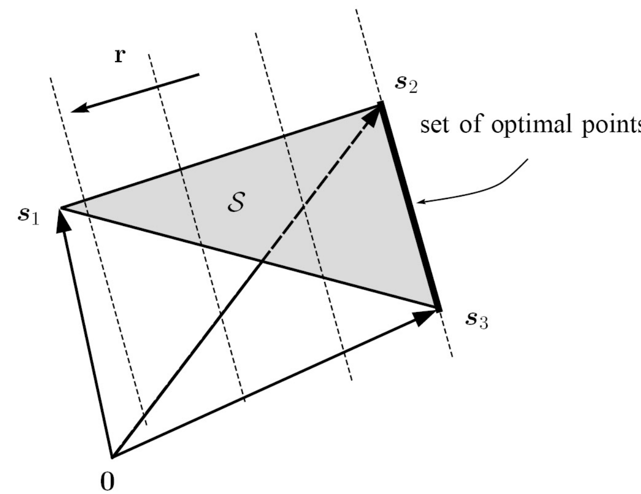
- By exploiting this property, we can locate all the extreme points by solving a sequence of LPs ($\approx 2N$ LPs at worst).

Consider the following linear program (LP)

$$p^{\star} = \min_{\boldsymbol{s}} \mathbf{r}^T \boldsymbol{s} \quad \text{s.t. } \boldsymbol{s} \in \mathcal{S} \tag{$\dagger$}$$

for an arbitrary $\mathbf{r} \in \mathbb{R}^L$. From basic LP theory, the solution of ($\dagger$) is either an extreme point of $\mathcal{S}$ (or one of the $\boldsymbol{s}_i$'s), or any point on a face of $\mathcal{S}$.

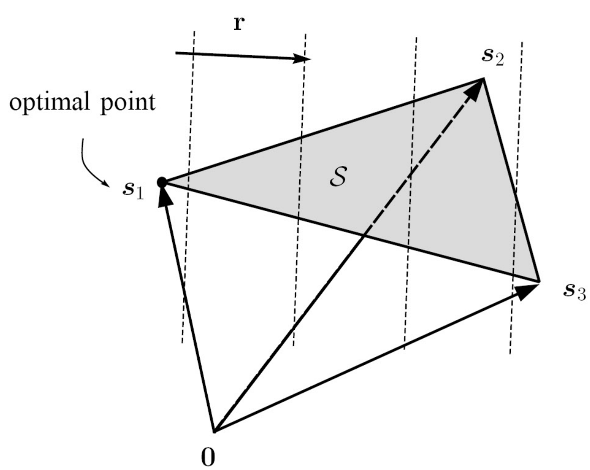


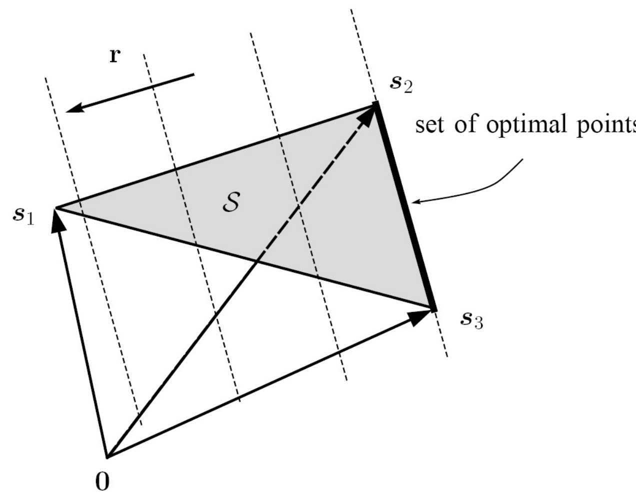


Consider the following linear program (LP)

$$p^\star = \min_{\boldsymbol{s}} \mathbf{r}^T \boldsymbol{s} \quad \text{s.t. } \boldsymbol{s} \in \mathcal{S} \tag{$\dagger$}$$

for an arbitrary $\mathbf{r} \in \mathbb{R}^L$. From basic LP theory, the solution of ($\dagger$) is either an extreme point of $\mathcal{S}$ (or one of the $\boldsymbol{s}_i$'s), or any point on a face of $\mathcal{S}$.



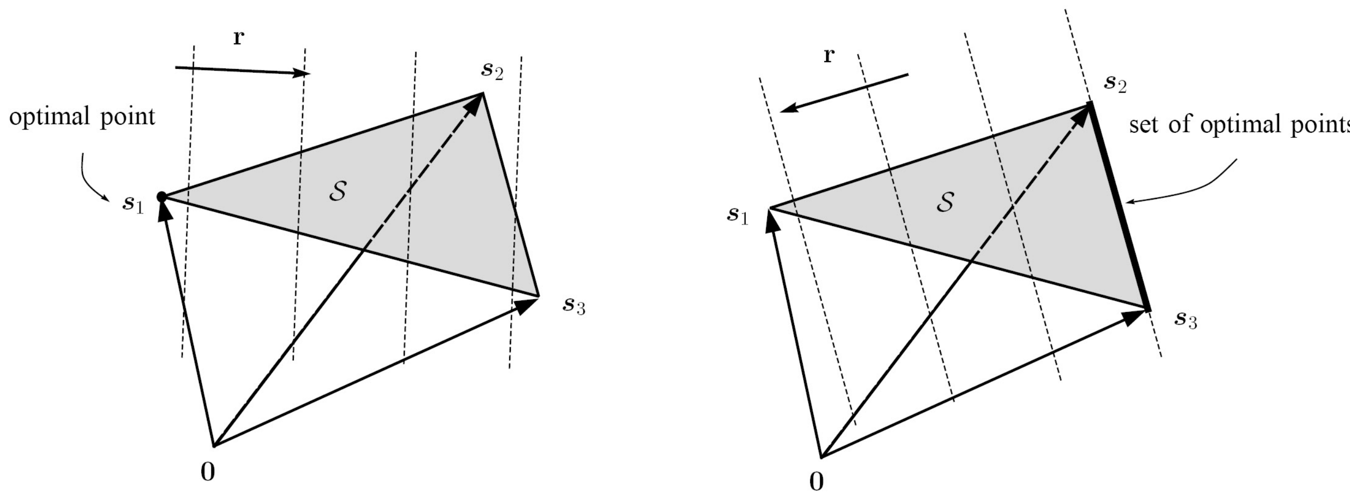


- **Question:** how to decide $\mathbf{r}$?

Consider the following linear program (LP)

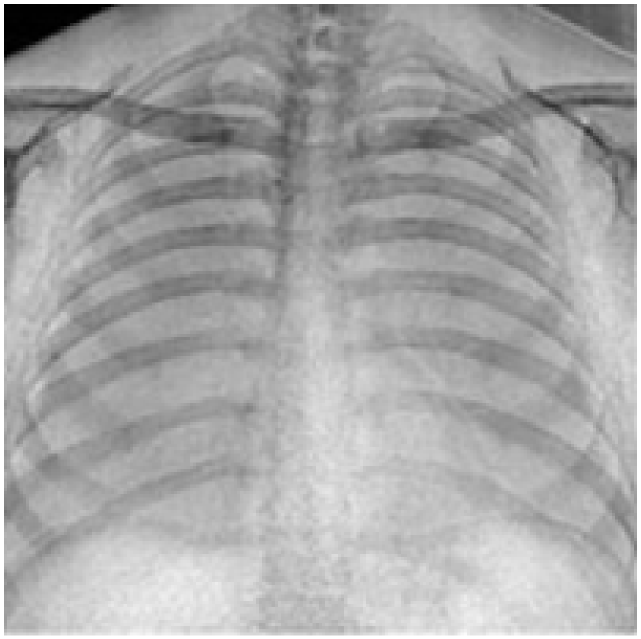
$$\begin{aligned} p^\star = \ & \min_{\boldsymbol{s}} \ \mathbf{r}^T \boldsymbol{s} \\ & \text{s.t.} \ \boldsymbol{s} \in \mathcal{S} \end{aligned} \qquad (\dagger)$$

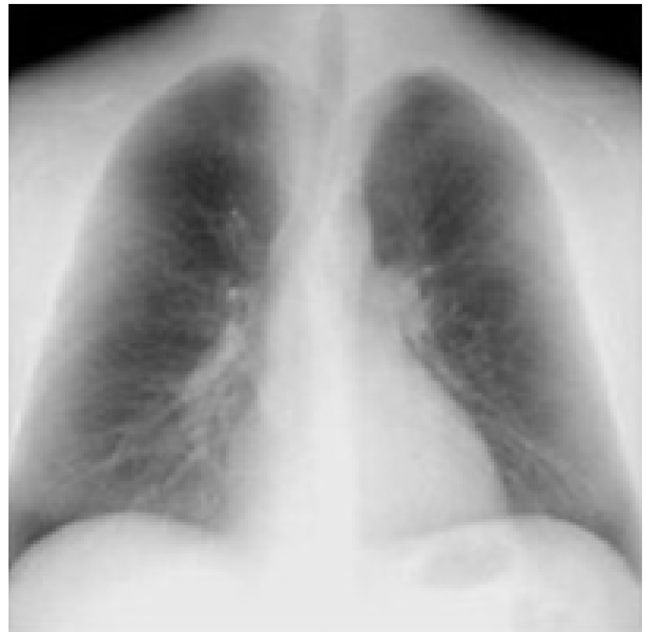
for an arbitrary $\mathbf{r} \in \mathbb{R}^L$. From basic LP theory, the solution of $(\dagger)$ is either an extreme point of $\mathcal{S}$ (or one of the $\boldsymbol{s}_i$'s), or any point on a face of $\mathcal{S}$.



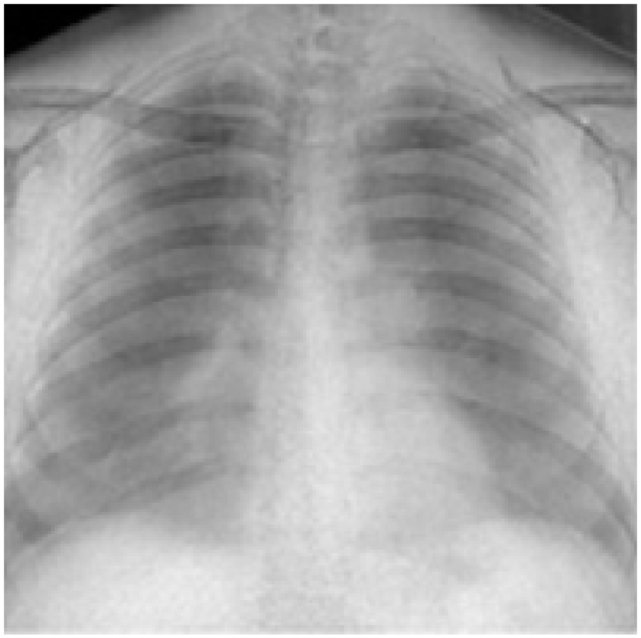
- **Question:** how to decide $\mathbf{r}$?
  - Ans: randomly! If $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_L)$, then, with probability one, the solution of $(\dagger)$ is an extreme point.
  - A systematic way to find all extreme points can also be constructed.

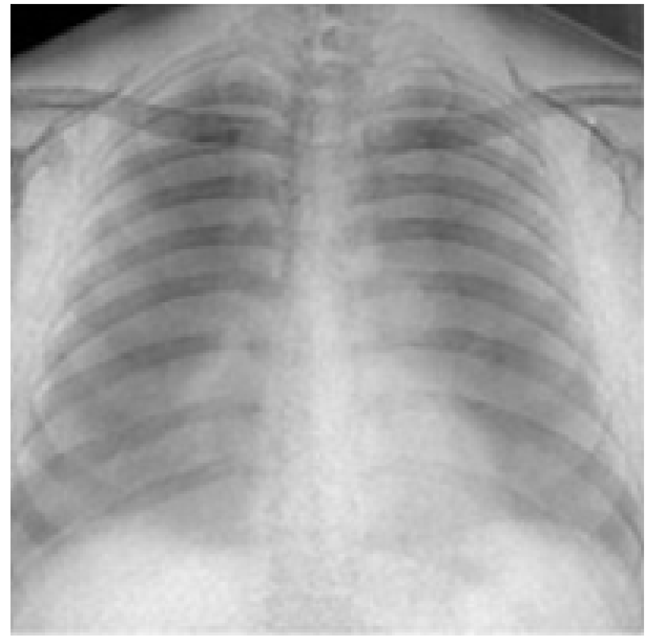
Simulation example 1: Dual energy X-Ray



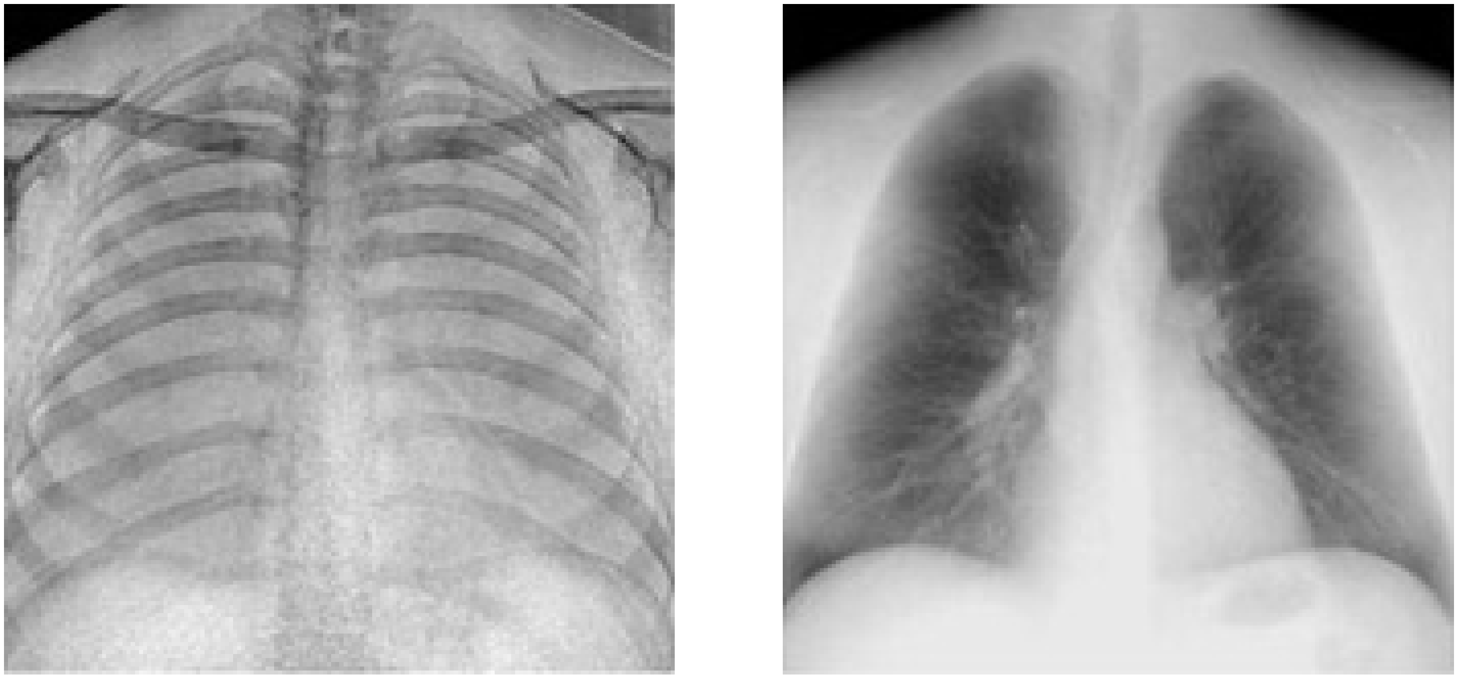


Original sources

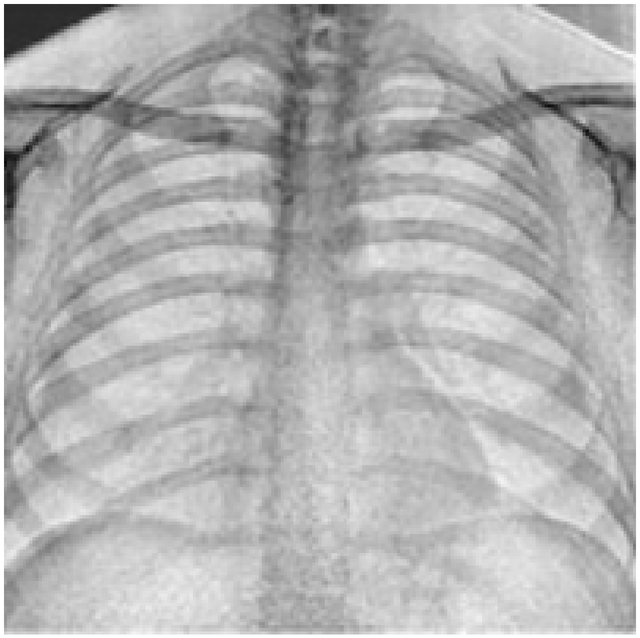


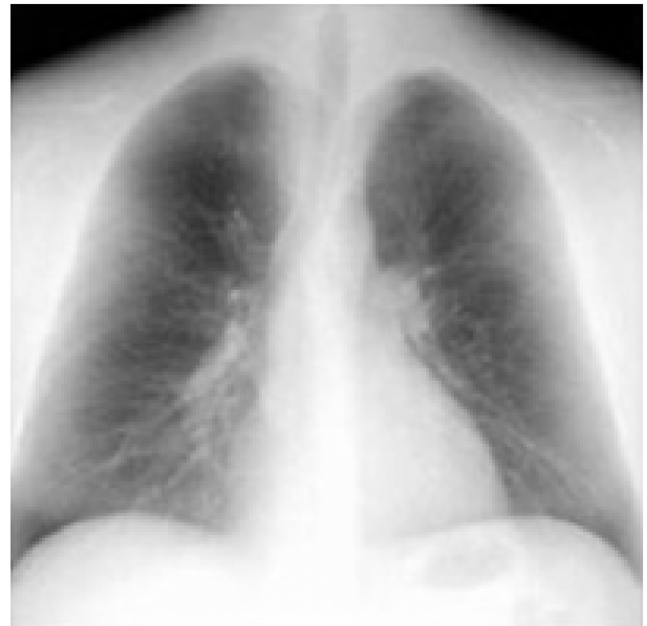


Observations



Separated sources by **CAMNS**

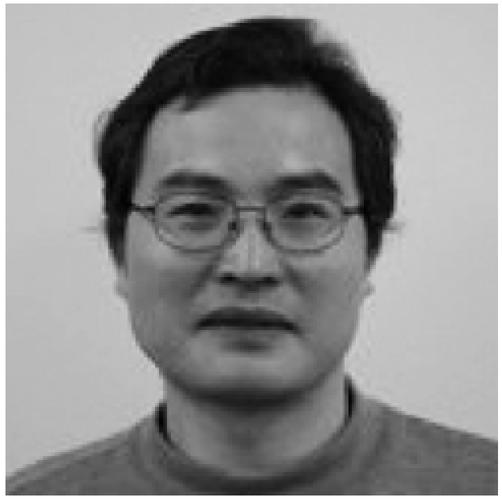




Separated sources by **nICA** (a benchmarked nBSS method)

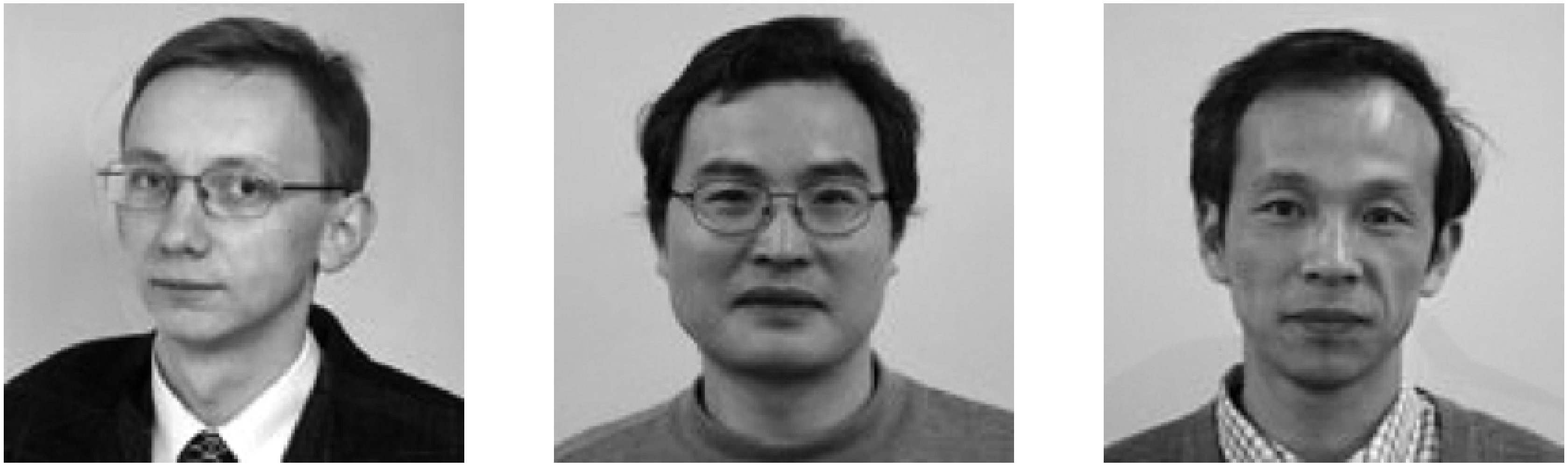
Separated sources by **NMF** (yet another benchmarked nBSS method)

Simulation example 2: Human faces


Original sources

Observations



Separated sources by **CAMNS**

Separated sources by **nICA**

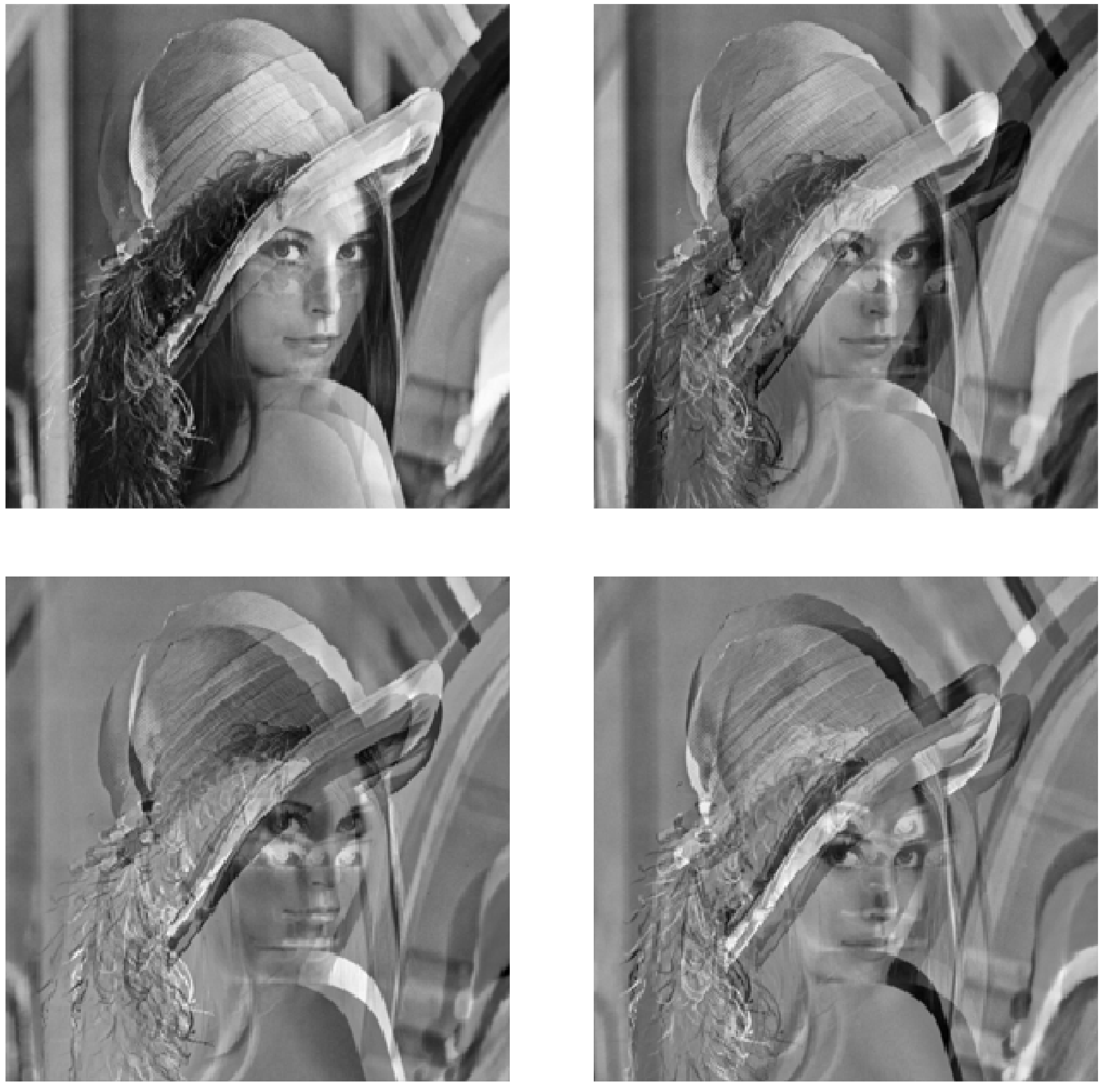
Separated sources by **NMF**

Simulation example 3: Ghosting

Original sources

Observations

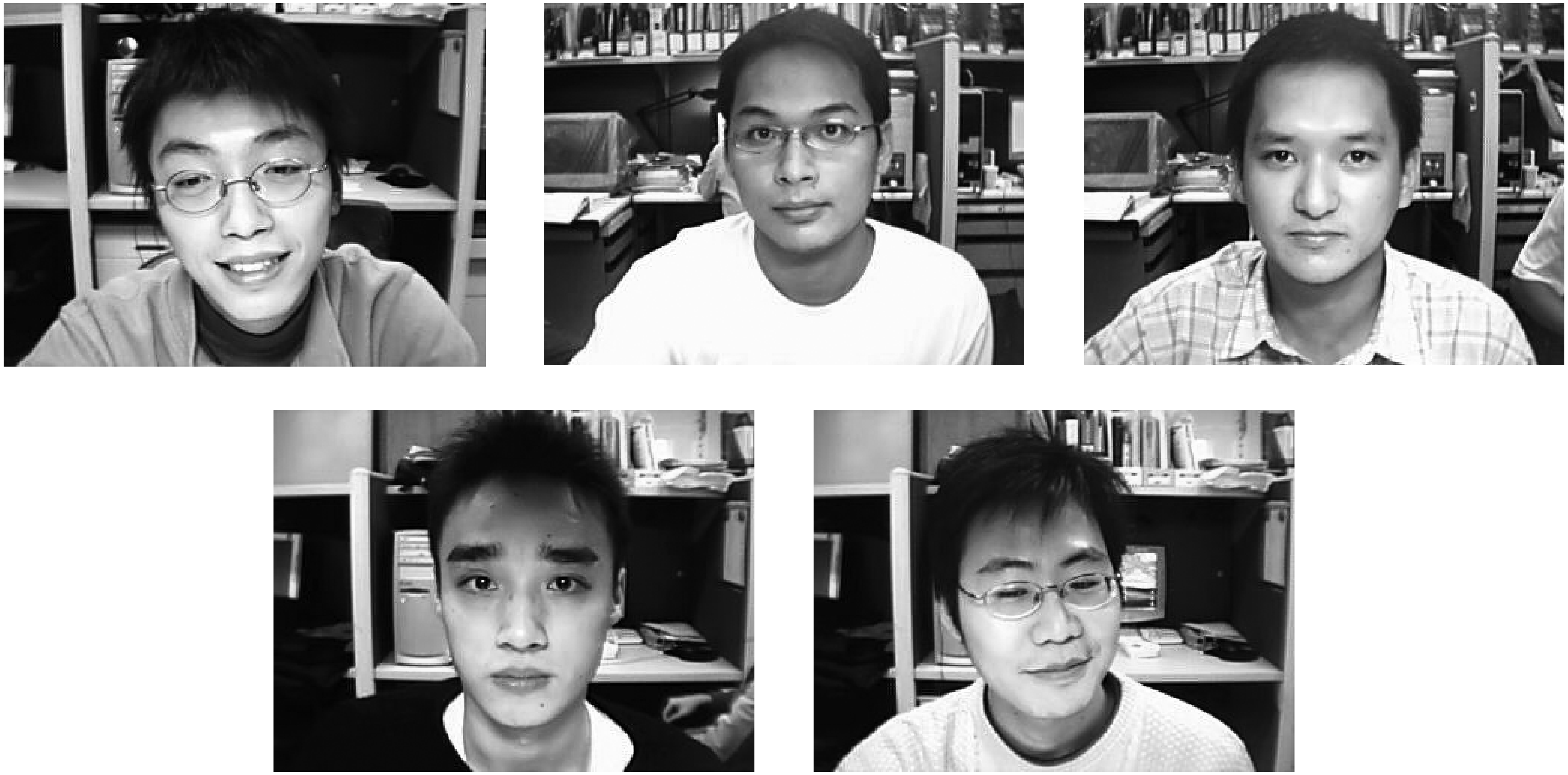
Separated sources by **CAMNS**



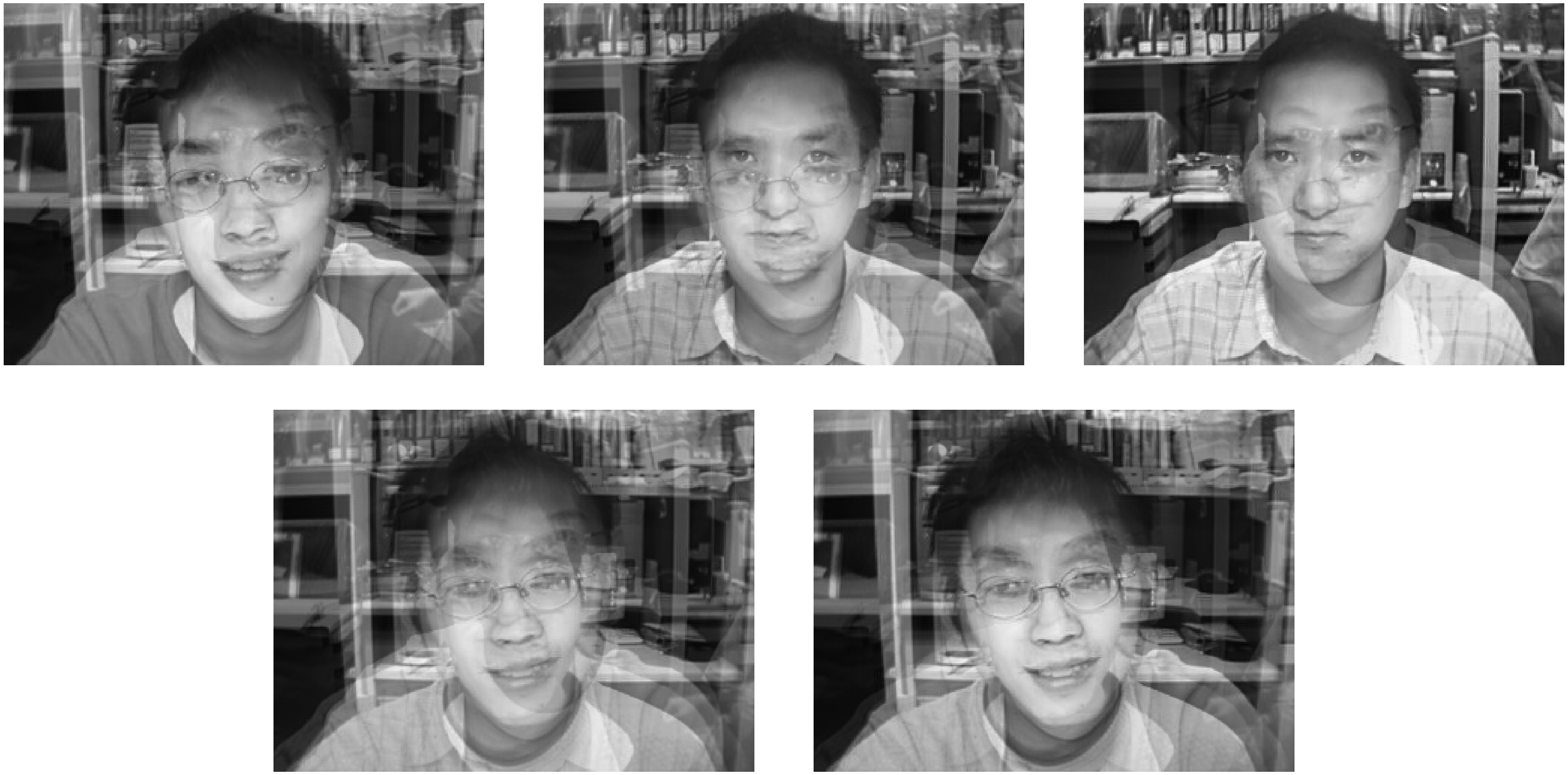
Separated sources by **nICA**

Separated sources by **NMF**

Simulation example 4: Five of my students



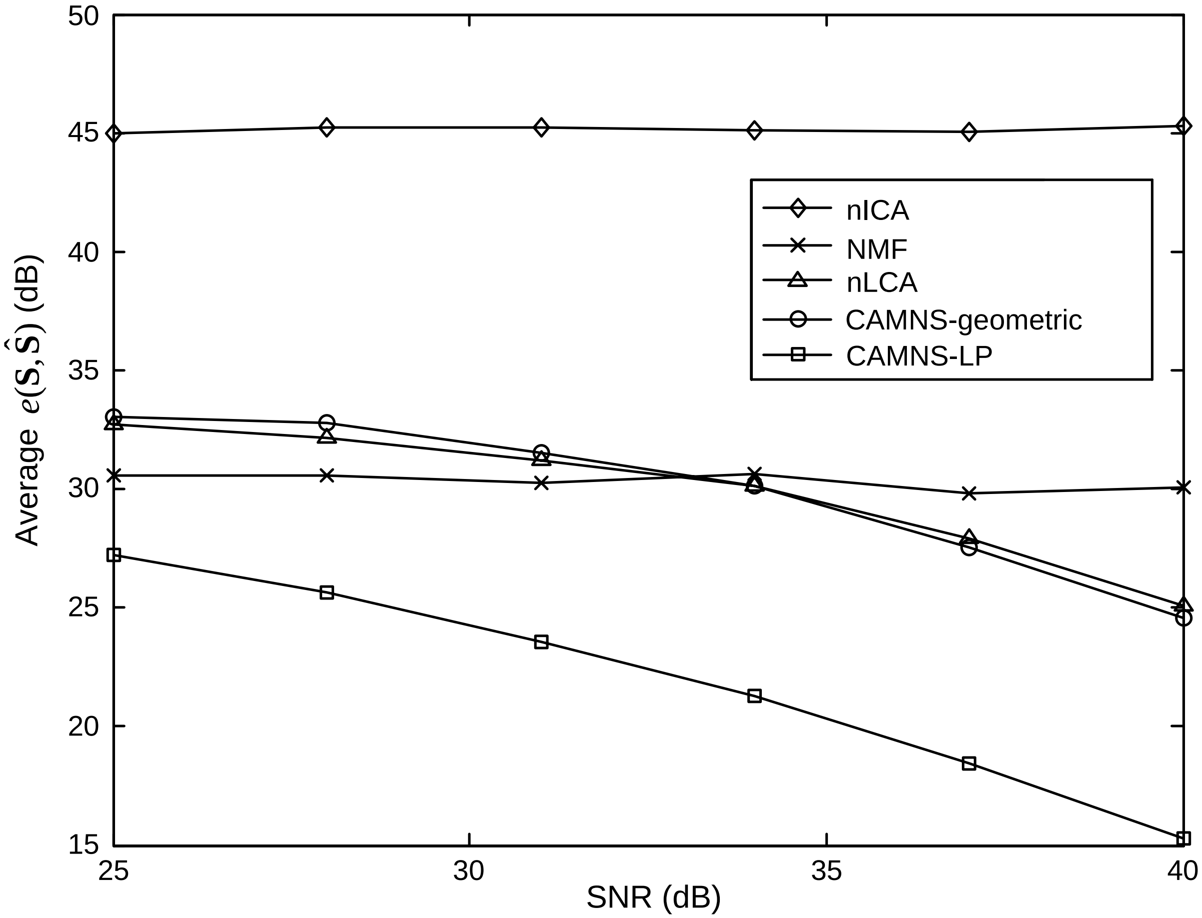
Original sources



Observations

Separated sources by **CAMNS**

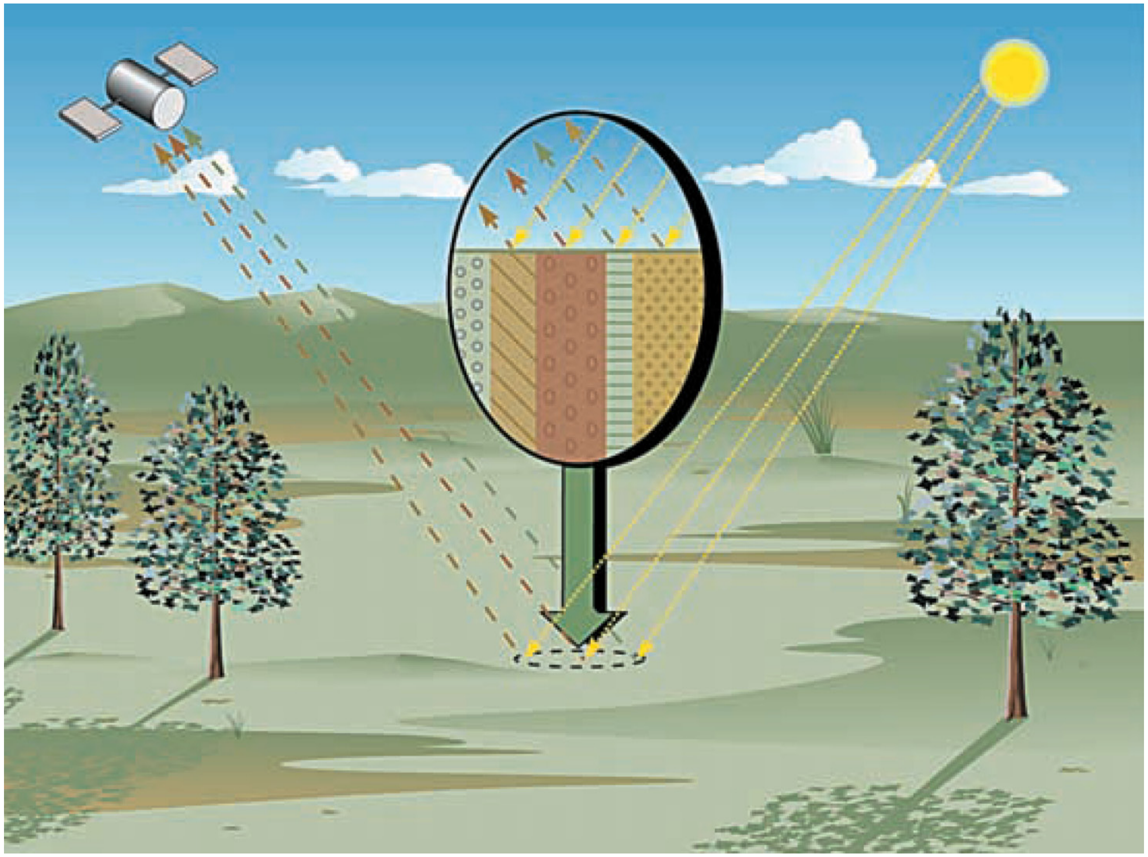
Simulation example 5: Monte Carlo performance for $N = 3$



Average sum squared errors of the sources with respect to SNRs.

Convex Geometry for Hyperspectral Unmixing in Remote Sensing

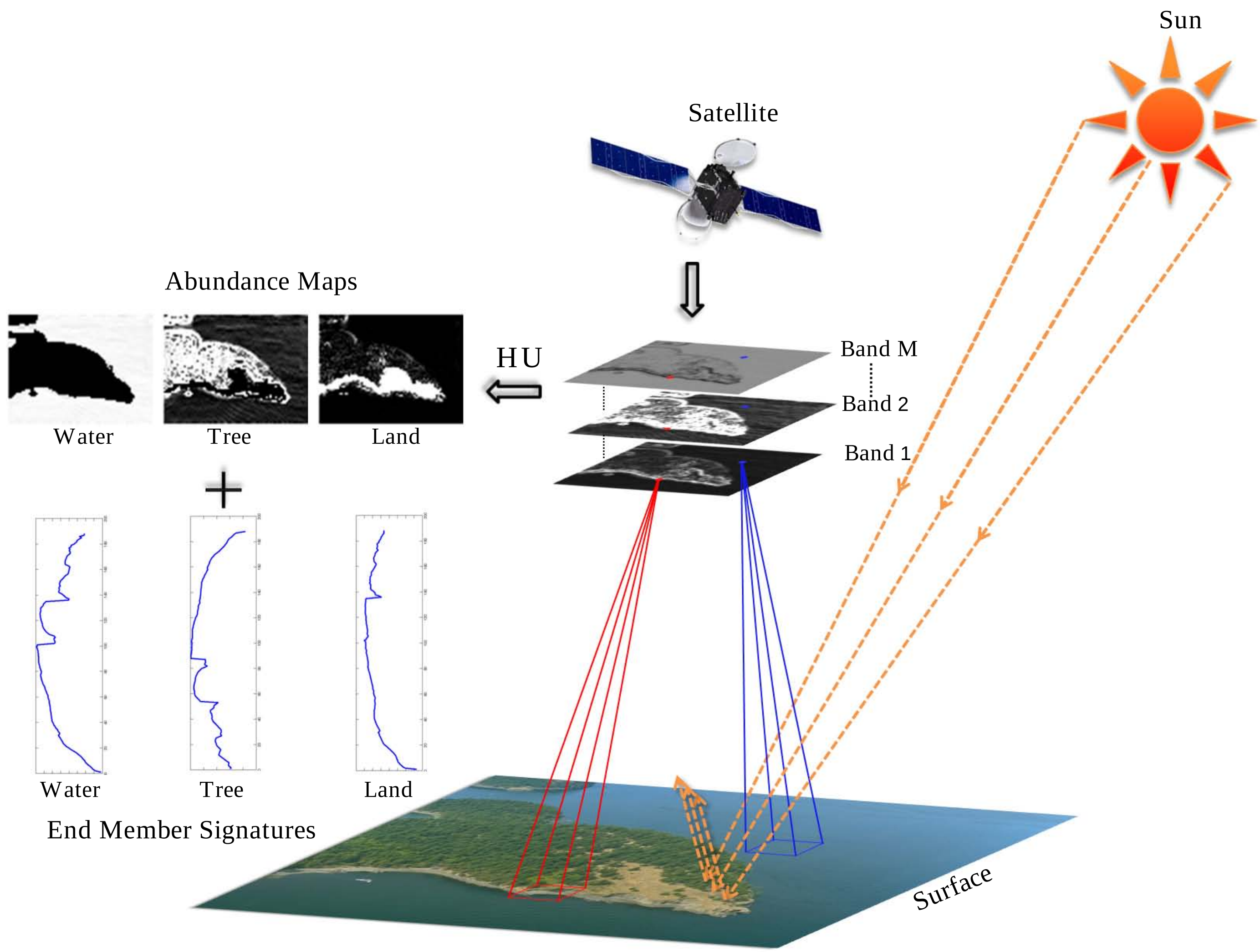
Hyperspectral Imaging: A Key Area in Geoscience and Remote Sensing



Courtesy to [Keshava *et al.* '02].

- Hyperspectral sensors record EM scattering patterns of distinct materials over > 200 spectral bands, from visible to near-infrared wavelength at a resolution of 10nm.

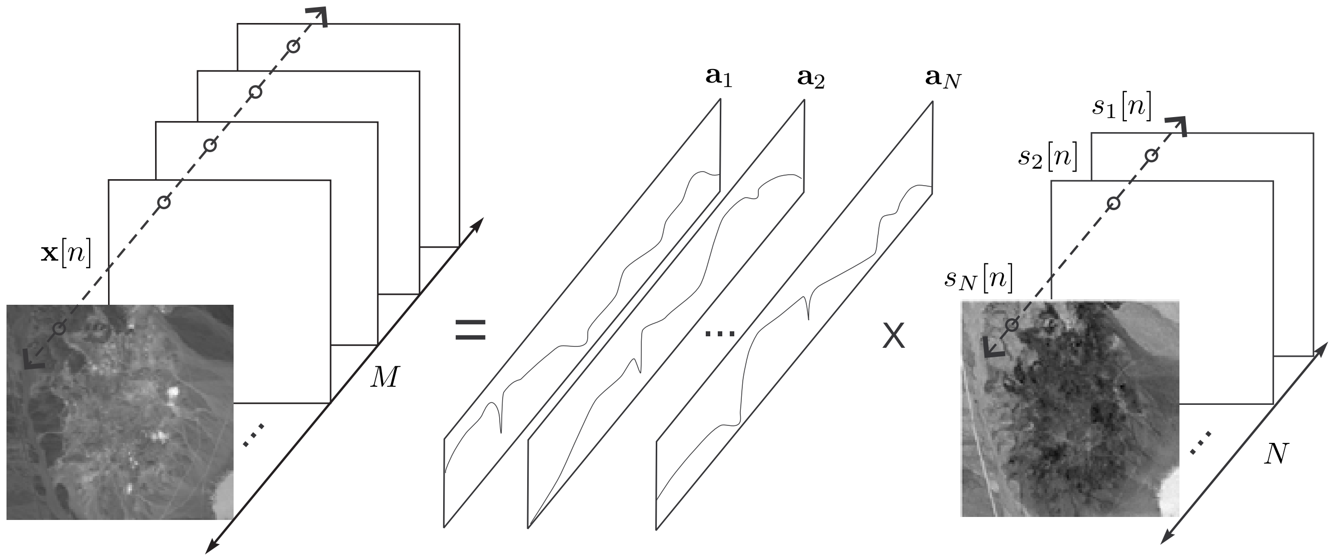
- The high spectral degrees of freedom enable us to identify the unknown materials, as revealed by their spectral signatures, and their compositions in the scene — this is fundamentally connected to source separation.



Applications of Hyperspectral Imaging

- Hyperspectral imaging has found numerous applications in remote sensing, such as
  - mineral identification,
  - agriculture,
  - environment monitoring,
  - terrain classification, land-cover mapping,
  - object detection, change detection, ...
- It is also a crucial technique for planetary exploration (like Mars) and astrophysics.
- It also has non-remote sensing applications, such as food inspection, forensics, medical imaging and chemometrics.
- A key problem in hyperspectral imaging is blind hyperspectral unmixing (also called unsupervised hyperspectral unmixing), which is essentially nBSS.

Hyperspectral Linear Mixing Model

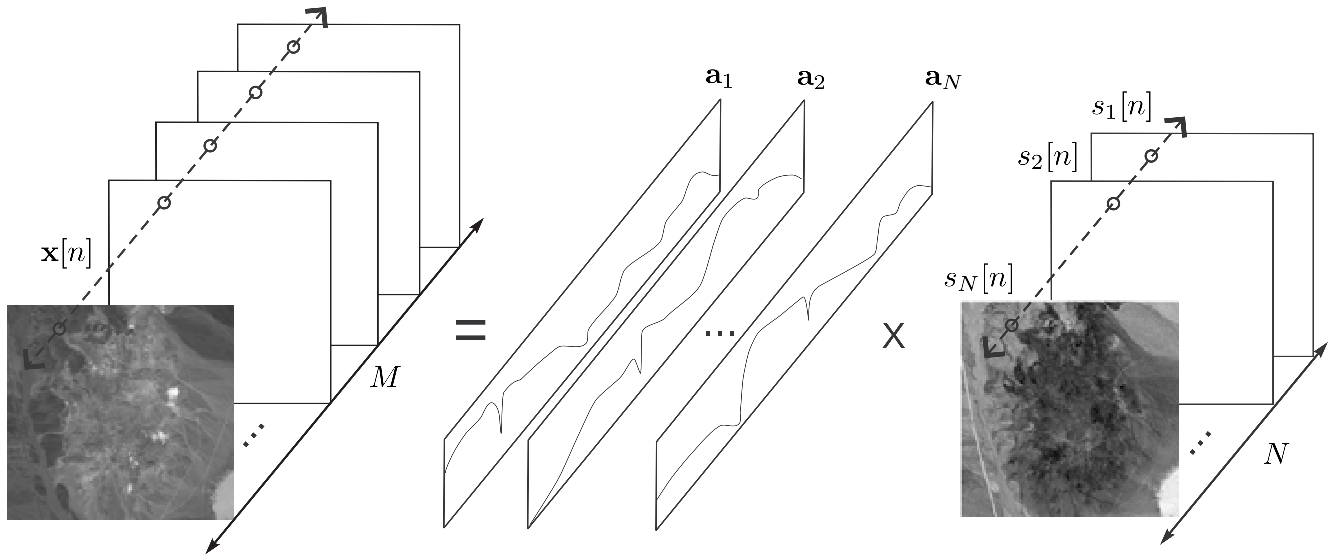


- Signal model:

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] = \sum_{i=1}^{N} s_i[n]\mathbf{a}_i, \qquad n = 1, \ldots, L$$

$\mathbf{x}[n] = [x_1[n], \ldots, x_M[n]]^T \in \mathbb{R}^M$	observed pixel vector;
$\mathbf{A} = [\ \mathbf{a}_1, \ldots, \mathbf{a}_N\ ] \in \mathbb{R}^{M \times N}$	endmember signature matrix;
$\mathbf{s}[n] = [\ s_1[n], \ldots, s_N[n]\ ]^T \in \mathbb{R}^N$	abundance vector;
$M,\ N,\ \& L$	$\#$ of spectral bands, endmembers, $\&$ pixels.

Hyperspectral Linear Mixing Model



- Signal model:

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] = \sum_{i=1}^{N} s_i[n]\mathbf{a}_i, \qquad n = 1, \ldots, L$$

- Assumptions:
  - $s_i[n] \geq 0$ for all i, n (non-negativity), $\sum_{i=1}^{N} s_i[n] = 1$ for all n (sum-to-one).
  - $\mathrm{rank}(\mathbf{A}) = N$.

Convex Geometry Observation

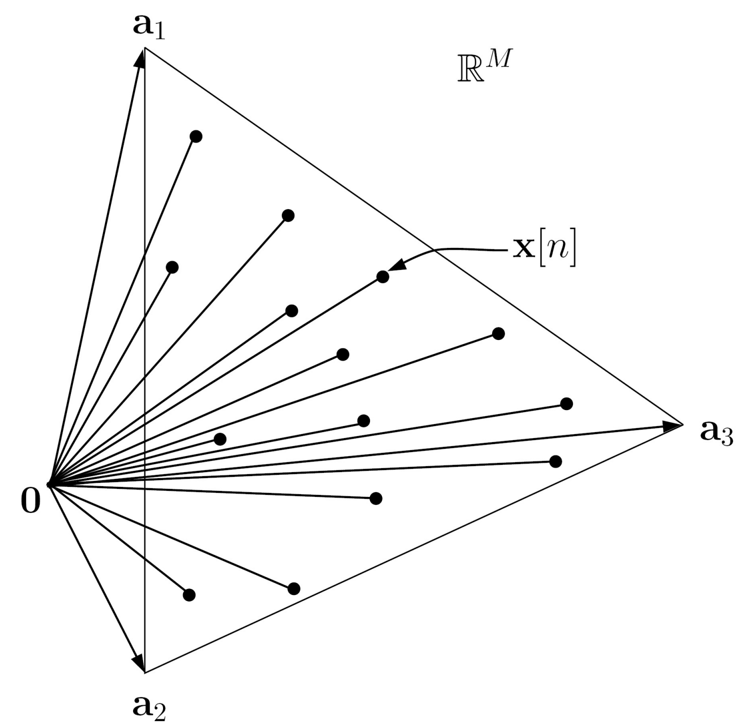
- Recall the signal model

$$\mathbf{x}[n] = \sum_{i=1}^{N} s_i[n]\mathbf{a}_i,$$

and that $s_i[n] \geq 0$, $\sum_{i=1}^{N} s_i[n] = 1$.

- Apparently, we have

$$\mathbf{x}[n] \in \mathrm{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}.$$



- **Observation:** each hyperspectral pixel $\mathbf{x}[n]$ lies in the convex hull of the ground-truth endmembers $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$.

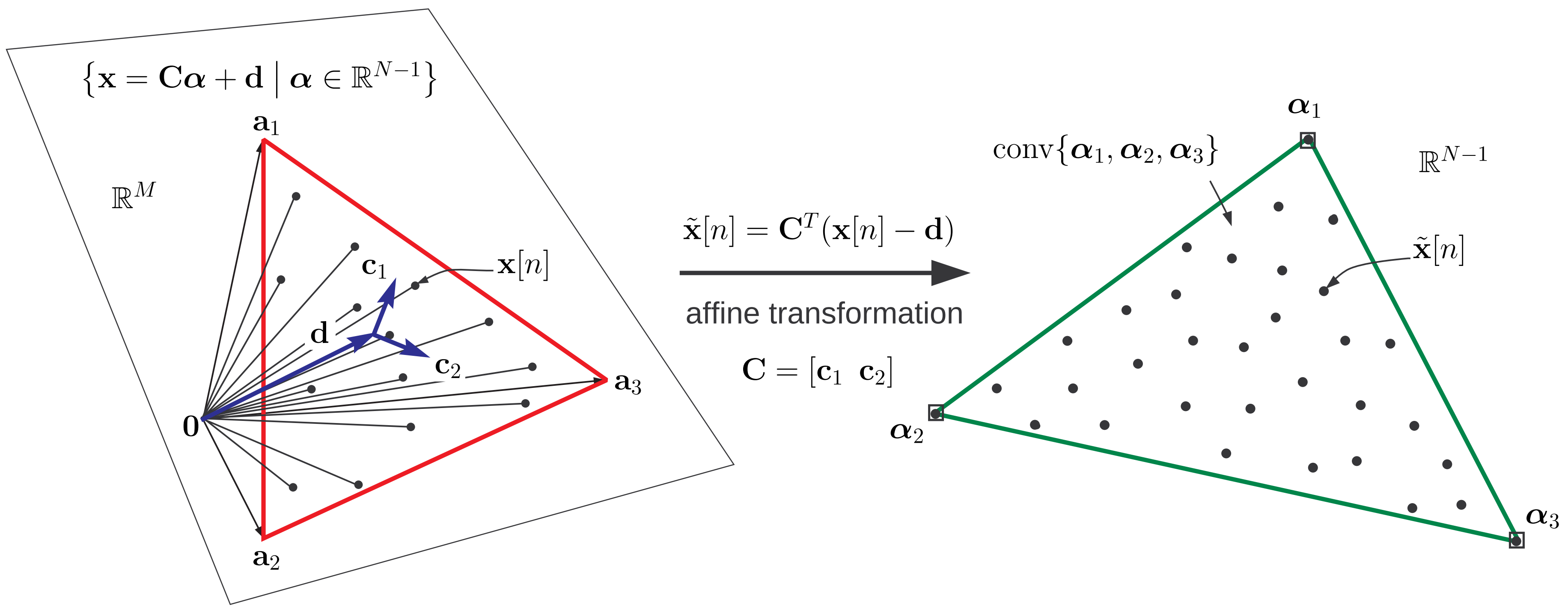
- Intuitively, if we can find the 'corners', then we are done!

Simplex Geometry Preserved Via Affine Transformation

- By affine set fitting, we can perform dimension reduction. The dimension-reduced pixels $\tilde{\mathbf{x}}[n]$ adhere to a similar model

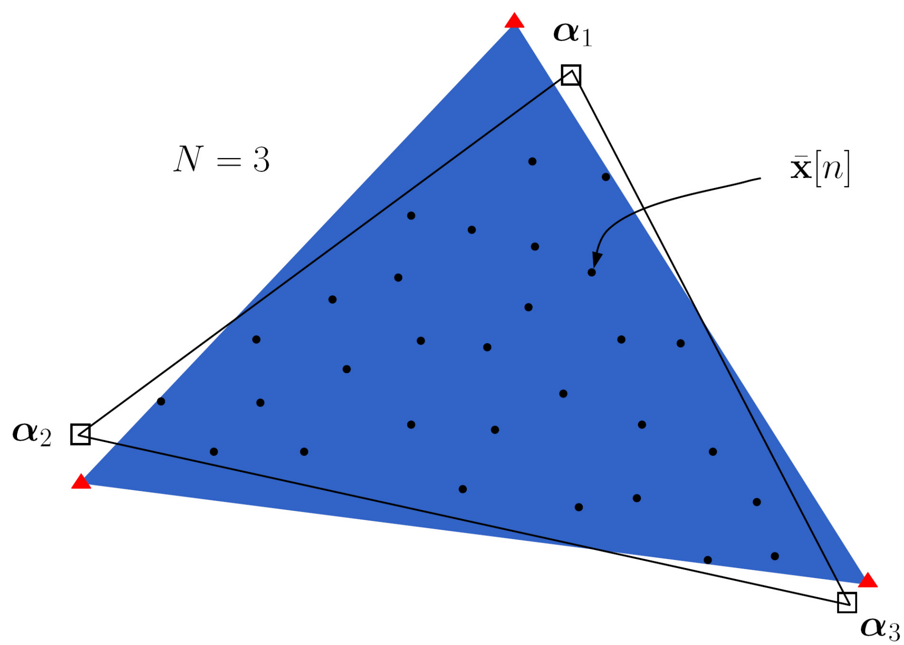
$$\tilde{\mathbf{x}}[n] = \sum_{i=1}^{N} s_i[n]\boldsymbol{\alpha}_i$$

where $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$ are dimension-reduced endmembers.



Craig's Minimum Volume Simplex Belief

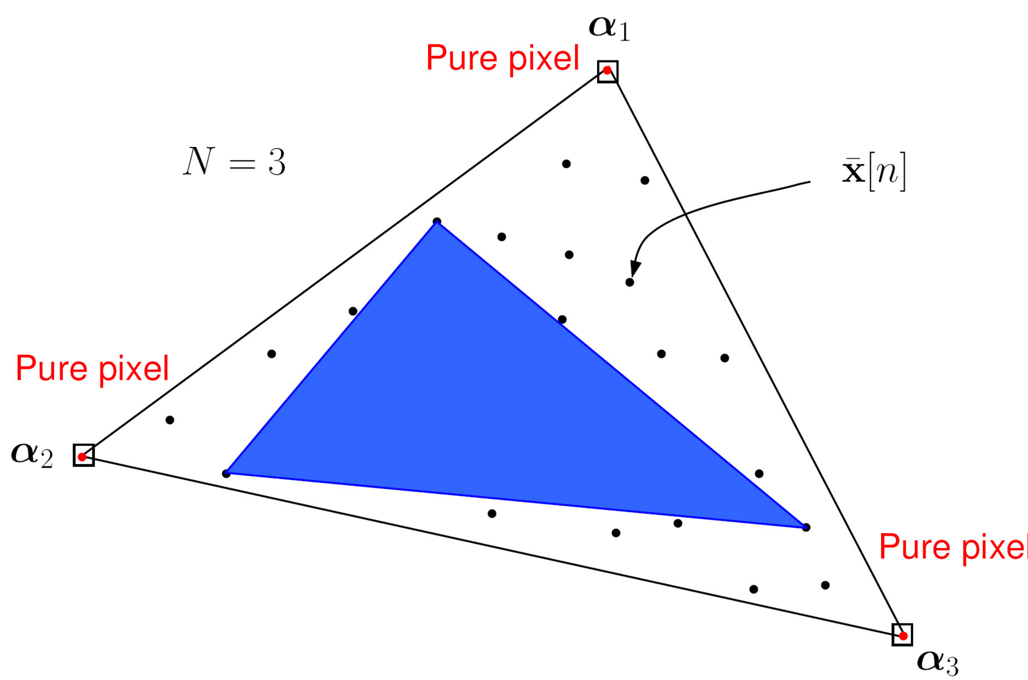
Craig's belief [Craig'94]: the true endmembers may be located by finding a data enclosing simplex whose volume is the smallest.



- Craig's belief provided significant insights to blind hyperspectral unmixing.
- But is Craig's belief fundamentally sound? Can it be efficiently implemented?

Winter's Maximum Volume Simplex Belief

Winter's belief [Winter'99]: the true endmembers may be located by finding a collection of pixel vectors whose simplex volume is the largest.

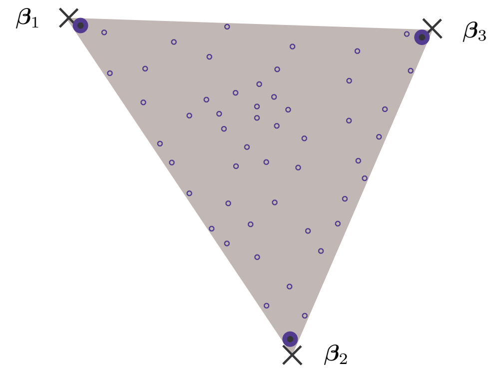


- Winter's belief led to N-FINDR, a class of widely used blind unmixing algorithms.

- Again, is Winter's belief fundamentally sound? Is N-FINDR the only way to go?

An Opt. Approach to Craig's Belief [Chan-Chi-Huang-Ma'09]

- We employ convex analysis and optimization to treat Craig's belief.
- We formulate Craig's belief as



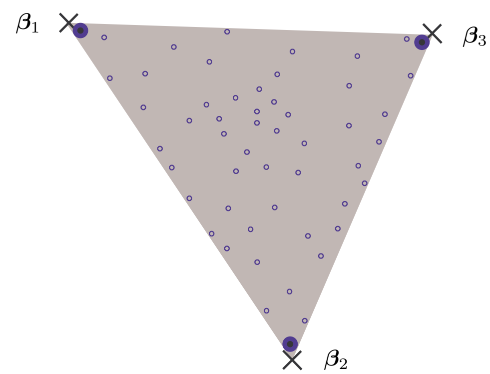
$$(\dagger) \qquad \begin{aligned} \min_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \quad & \mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t.} \quad & \tilde{\mathbf{x}}[n] \in \mathrm{conv}\{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\}, \\ & i = 1,\ldots,L, \end{aligned}$$

where $\mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) = |\det(\tilde{\boldsymbol{\beta}}_1,\ldots,\tilde{\boldsymbol{\beta}}_N)|/(N-1)!$, $\tilde{\boldsymbol{\beta}}_i = [\ \boldsymbol{\beta}_i^T, 1\ ]^T$, is the simplex volume.

An Opt. Approach to Craig's Belief [Chan-Chi-Huang-Ma'09]

- We employ convex analysis and optimization to treat Craig's belief.

- We formulate Craig's belief as



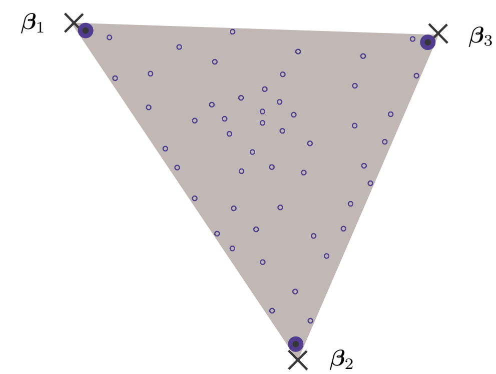
$$
(\dagger) \quad \begin{aligned} \min_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \quad & \mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t.} \quad & \tilde{\mathbf{x}}[n] \in \mathrm{conv}\{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\}, \\ & i = 1,\ldots,L, \end{aligned}
$$

where $\mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) = |\det(\tilde{\boldsymbol{\beta}}_1,\ldots,\tilde{\boldsymbol{\beta}}_N)|/(N-1)!$, $\tilde{\boldsymbol{\beta}}_i = [\ \boldsymbol{\beta}_i^T, 1\ ]^T$, is the simplex volume.

- **Question 1:** Is Craig's belief fundamentally sound?
  - **Yes.** We prove that local dominance (or pure pixels) is a sufficient condition for problem (†) to perfectly identify the true endmembers.

An Opt. Approach to Craig's Belief [Chan-Chi-Huang-Ma'09]

- We employ convex analysis and optimization to treat Craig's belief.
- We formulate Craig's belief as



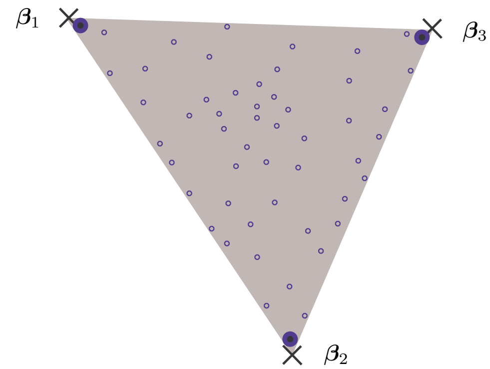
$$
(\dagger) \quad \begin{aligned} \min_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \quad & \mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t.} \quad & \tilde{\mathbf{x}}[n] \in \mathrm{conv}\{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\}, \\ & i = 1,\ldots,L, \end{aligned}
$$

where $\mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) = |\det(\tilde{\boldsymbol{\beta}}_1,\ldots,\tilde{\boldsymbol{\beta}}_N)|/(N-1)!$, $\tilde{\boldsymbol{\beta}}_i = [\ \boldsymbol{\beta}_i^T, 1\ ]^T$, is the simplex volume.

- **Question 2:** Can Craig's belief be efficiently implemented?
  - Problem ($\dagger$) is nonconvex.
  - We derive a pragmatic algorithm using alternating LP optimization.

An Opt. Approach to Craig's Belief [Chan-Chi-Huang-Ma'09]

- We employ convex analysis and optimization to treat Craig's belief.
- We formulate Craig's belief as



$$(\dagger) \quad \begin{aligned} \min_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \quad & \mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t.} \quad & \tilde{\mathbf{x}}[n] \in \mathrm{conv}\{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\}, \\ & i = 1,\ldots,L, \end{aligned}$$

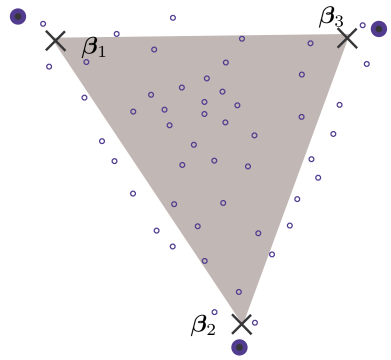
where $\mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) = |\det(\tilde{\boldsymbol{\beta}}_1,\ldots,\tilde{\boldsymbol{\beta}}_N)|/(N-1)!$, $\tilde{\boldsymbol{\beta}}_i = [\ \boldsymbol{\beta}_i^T, 1\ ]^T$, is the simplex volume.

- **Question 2:** Can Craig's belief be efficiently implemented?
  - Problem (†) is nonconvex.
  - We derive a pragmatic algorithm using alternating LP optimization.
    - ∗ A similar idea, MVSA [Li-Bioucas'08], was proposed about the same time.

An Opt. Approach to Winter's Belief [Chan-Ma-Ambikapathi-Chi'11]

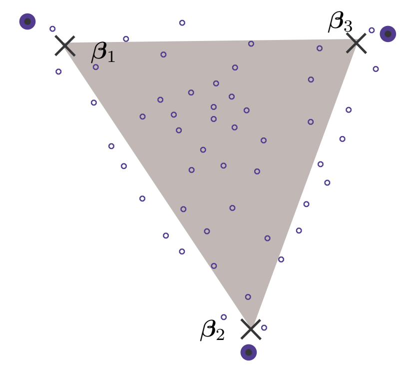
- We formulate Winter's belief as

$$
(\ddagger) \quad \begin{aligned} \max_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \ & \mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t. } & \boldsymbol{\beta}_i \in \mathrm{conv}\{\tilde{\mathbf{x}}[1],\ldots,\tilde{\mathbf{x}}[L]\}, \\ & i = 1,\ldots,N. \end{aligned}
$$



An Opt. Approach to Winter's Belief [Chan-Ma-Ambikapathi-Chi'11]

- We formulate Winter's belief as

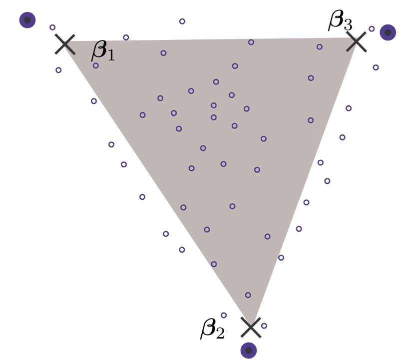


$$(\ddagger) \quad \begin{aligned} \max_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \ & \mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t.} \ & \boldsymbol{\beta}_i \in \mathrm{conv}\{\tilde{\mathbf{x}}[1],\ldots,\tilde{\mathbf{x}}[L]\}, \\ & i = 1,\ldots,N. \end{aligned}$$

- **Question 1:** Is Winter's belief fundamentally sound?
  - Yes. But we prove that local dominance is a sufficient and necessary condition for problem ($\ddagger$) to perfectly identify the true endmembers.
  - This implies that Winter is fundamentally weaker than Craig.

An Opt. Approach to Winter's Belief [Chan-Ma-Ambikapathi-Chi'11]

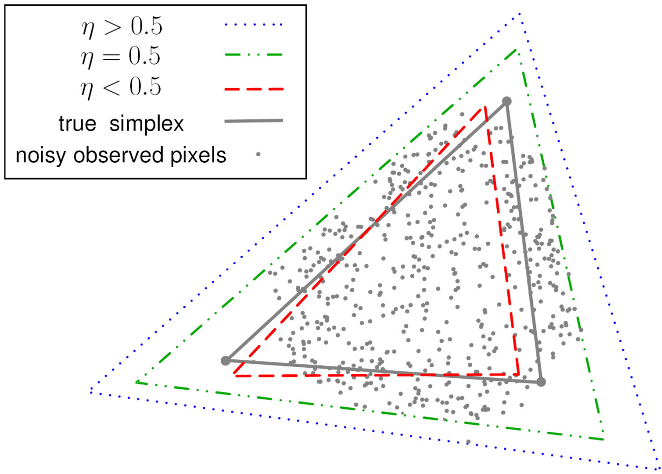
- We formulate Winter's belief as



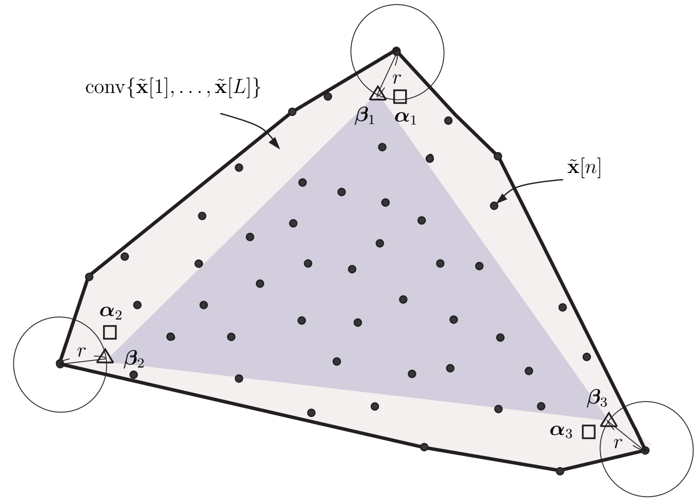
$$(\ddagger) \quad \begin{aligned} \max_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \ & \mathrm{vol}(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t.} \ & \boldsymbol{\beta}_i \in \mathrm{conv}\{\tilde{\mathbf{x}}[1],\ldots,\tilde{\mathbf{x}}[L]\}, \\ & i = 1,\ldots,N. \end{aligned}$$

- **Question 2:** Is N-FINDR the only way for Winter's belief?
  - Not really.
  - N-FINDR may be viewed as an alternating opt. algorithm for ($\ddagger$).
  - VCA [Nascimento-Bioucas'06] was previously not seen as Winter-based. We show that VCA may be interpreted as a greedy opt. algorithm for ($\ddagger$).
  - Our top-down study unifies such existing algorithms under one umbrella, and gives new theoretical insight and implication to them.

Robust Generalizations



(a) Robust Craig geometry.

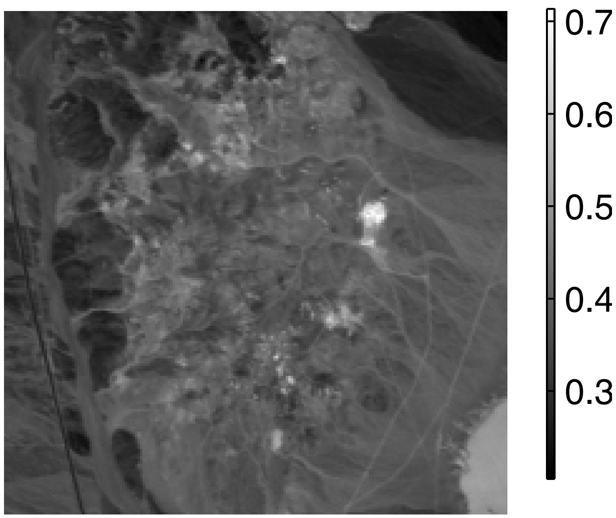


(b) Robust Winter geometry.

- Hyperspectral data may be corrupted by measurement noise.

- Our top-down optimization approach enables us to develop robust generalizations of Craig's and Winter's formulations [Chan-Ambikapathi-Ma-Chi'11], [Chan-Ma-Ambikapathi-Chi'11].

Real Data Experiment



(a) Cuprite AVIRIS data at band 10.

	RMVES	MVES	VCA
Andradite#1	**9.36**	25.61	-
Andradite#2	24.52	-	**18.49**
Alunite#1	**15.92**	21.85	17.74
Alunite#2	-	**17.72**	-
Buddingtonite	23.54	**22.98**	27.25
Chalcedony	**27.74**	38.25	31.9
Desert Varnish#1	20.99	18.64	**12.12**
Desert Varnish#2	-	**43.04**	-
Dumortierite	**20.77**	29.32	31.95 (32.01)
Goethite	**17.71**	19.05	-
Kaolinite#1	27.25	**26.50**	-
Kaolinite#2	-	**21.09**	32.49
Montmorillonite#1	22.99	-	**18.06**
Montmorillonite#2	**24.34**	26.00	-
Muscovite	39.63	44.19	**32.7**
Nontronite#1	**22.95**	28.83	24.66
Nontronite#2	-	-	**21.51**
Paragonite	-	-	**35.91**
Pyrope	-	-	**25.59**
Smectite	**22.53**	-	-
Average ϕ	**22.87**	27.11	25.73

(b) Spectral angle performance.

- Robust minimum volume simplex analysis (MVES) can perform better than our previous MVES algorithm.