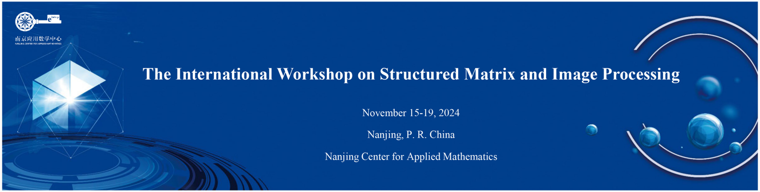
Dual Simplex Volume Maximization for Simplex-Structured Matrix Factorization

Maryam Abdolali [1] Giovanni Barbarino [2] Nicolas Gillis [2]

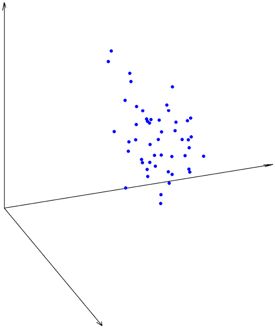


[1]K.N.Toosi University, Tehran, Iran

[2]Université de Mons, Belgium

Simplex-Structured Matrix Factorization

Simplex Identification



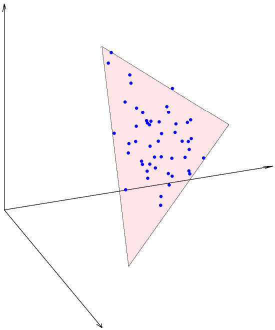
Given $X \in \mathbb{R}^{r-1\times n}$ can we find $W \in \mathbb{R}^{r-1\times r}$, $H \in \mathbb{R}^{r\times n}$ such that

$$X = WH \qquad H(:,i) \in \Delta^r = \{x \in \mathbb{R}^r_+ : x^T e = 1\} \quad \forall i$$

Since $X(:,i) = WH(:,i)$ is a *convex combination* of the columns of W

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1\times r}$$

Simplex Identification



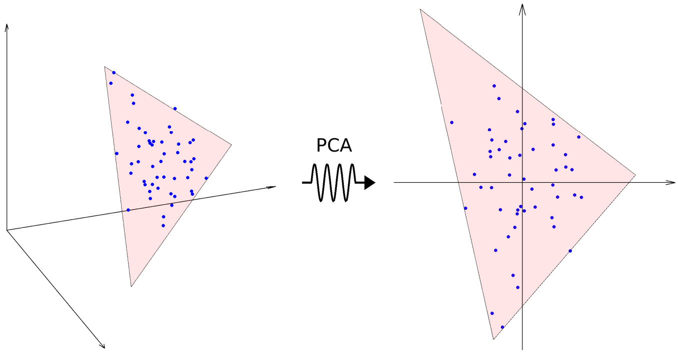
Given $X \in \mathbb{R}^{r-1 \times n}$ can we find $W \in \mathbb{R}^{r-1 \times r}$, $H \in \mathbb{R}^{r \times n}$ such that

$$X = WH \qquad H(:,i) \in \Delta^r = \{x \in \mathbb{R}^r_+ : x^T e = 1\} \quad \forall i$$

Since $X(:,i) = WH(:,i)$ is a *convex combination* of the columns of W

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Simplex Identification



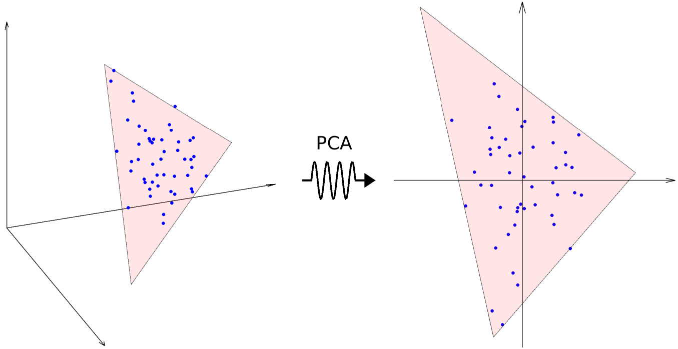
Given $X \in \mathbb{R}^{r-1 \times n}$ can we find $W \in \mathbb{R}^{r-1 \times r}$, $H \in \mathbb{R}^{r \times n}$ such that

$$X = WH \qquad H(:,i) \in \Delta^r = \{x \in \mathbb{R}^r_+ : x^T e = 1\} \quad \forall i$$

Since $X(:,i) = WH(:,i)$ is a *convex combination* of the columns of W

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Simplex Identification



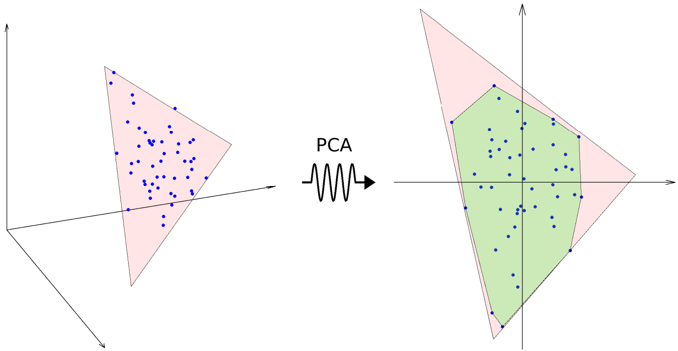
Given $X \in \mathbb{R}^{r-1 \times n}$ can we find $W \in \mathbb{R}^{r-1 \times r}$, $H \in \mathbb{R}^{r \times n}$ such that

$$X = WH \qquad H(:,i) \in \Delta^r = \{x \in \mathbb{R}^r_+ : x^T e = 1\} \quad \forall i$$

Since $X(:,i) = WH(:,i)$ is a *convex combination* of the columns of W

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Simplex Identification



Given $X \in \mathbb{R}^{r-1 \times n}$ can we find $W \in \mathbb{R}^{r-1 \times r}$, $H \in \mathbb{R}^{r \times n}$ such that

$$X = WH \qquad H(:,i) \in \Delta^r = \{x \in \mathbb{R}^r_+ : x^T e = 1\} \quad \forall i$$

Since $X(:,i) = WH(:,i)$ is a *convex combination* of the columns of W

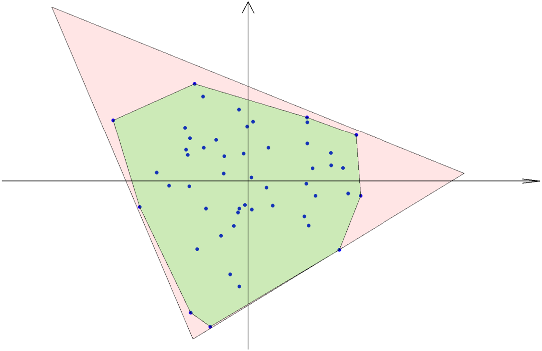
$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

An Application to Hyperspectral Imaging

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Exists? *Yes*... but it is far from being *Unique*

This is a problem for the **Interpretability** of the solution

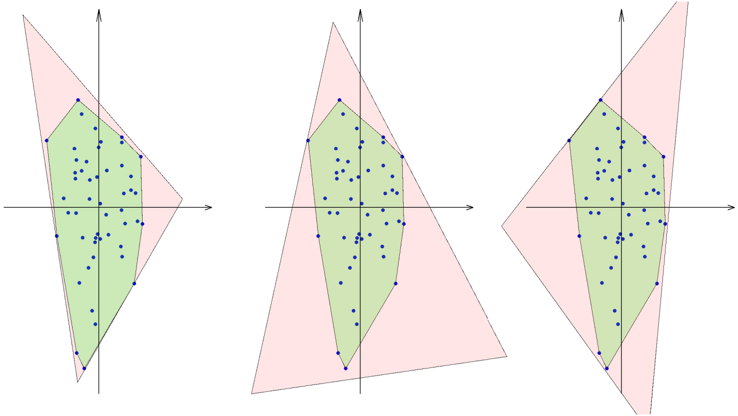


An Application to Hyperspectral Imaging

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Exists? *Yes*... but it is far from being *Unique*

This is a problem for the **Interpretability** of the solution

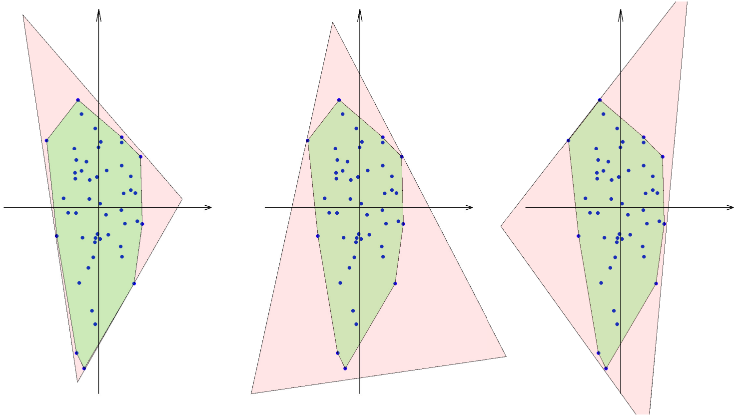


An Application to Hyperspectral Imaging

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Exists? *Yes...* but it is far from being *Unique*

This is a problem for the **Interpretability** of the solution



An Application to Hyperspectral Imaging

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Exists? *Yes*... but it is far from being *Unique*

This is a problem for the **Interpretability** of the solution

Jasper Ridge Data set

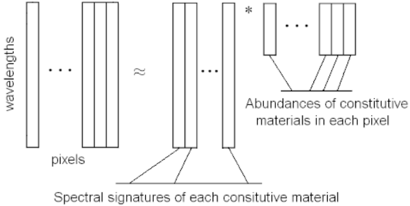
An Application to Hyperspectral Imaging

$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1 \times r}$$

Exists? *Yes*... but it is far from being *Unique*

This is a problem for the **Interpretability** of the solution

Jasper Ridge Data set



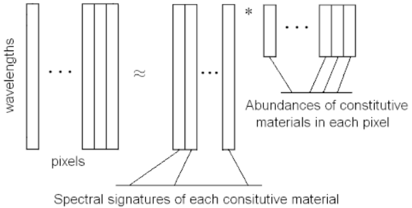
An Application to Hyperspectral Imaging

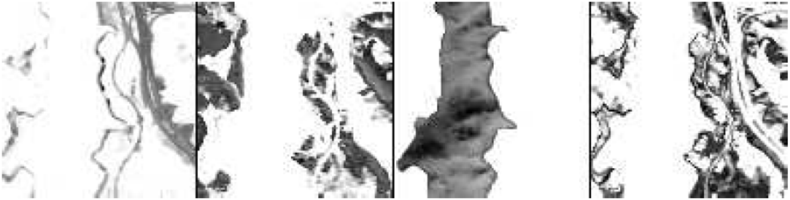
$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1\times r}$$

Exists? *Yes...* but it is far from being *Unique*

This is a problem for the **Interpretability** of the solution

Jasper Ridge Data set





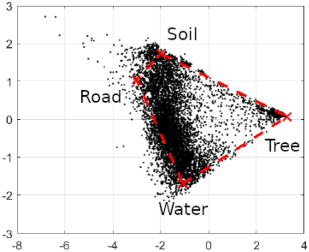
An Application to Hyperspectral Imaging

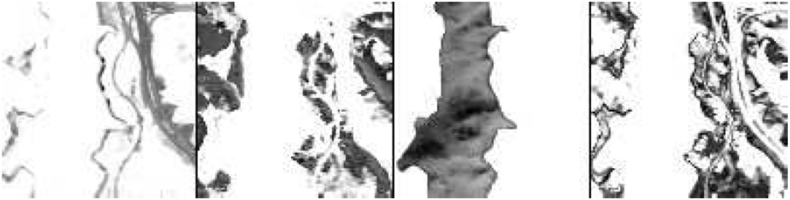
$$Conv(X) \subseteq Conv(W) \qquad W \in \mathbb{R}^{r-1\times r}$$

Exists? *Yes...* but it is far from being *Unique*

This is a problem for the **Interpretability** of the solution

Jasper Ridge Data set





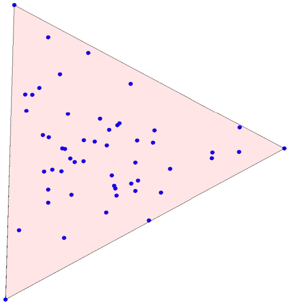
Separability

$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad$ i.e. $\qquad Conv(X) \equiv Conv(W)$

Separability

$$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$

- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption



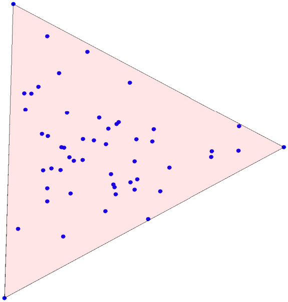
In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Separability

$$X = X(:,\mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$

- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption



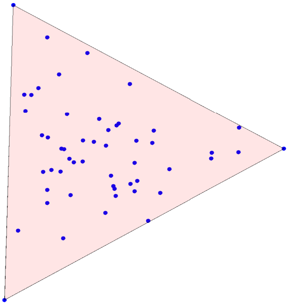
In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Separability

$$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$

- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption

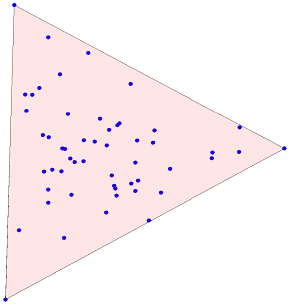


In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Separability

$$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$



- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption

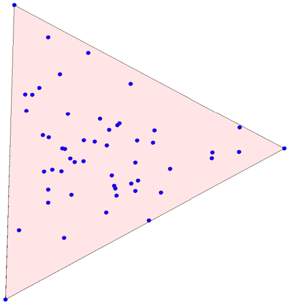
In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Separability

$$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$

- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption

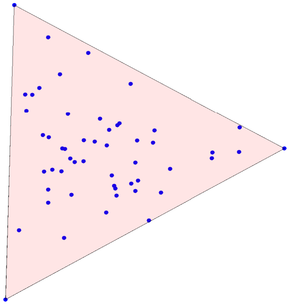


In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Separability

$$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$



- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption

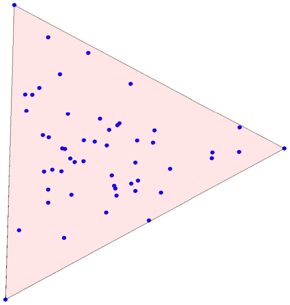
In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Separability

$$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$

- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption



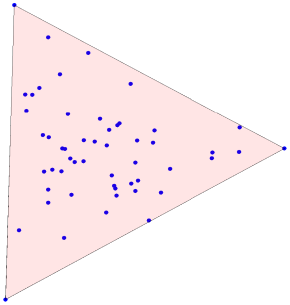
In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Separability

$$X = X(:, \mathcal{K})H \qquad |\mathcal{K}| = r \qquad \text{i.e.} \qquad Conv(X) \equiv Conv(W)$$

- ✓ Polytime algorithm
- ✓ Robust to perturbation
- ✓ Uniqueness of solution (up to permutations)
- ✓ Immediate Interpretability
- × Very strong assumption



In the Hyperspectral Imaging it means that for each material there exists a single pixel composed entirely of that material (called pure pixels)

We need **more general** assumptions with uniqueness guarantees

Sufficiently Scattered Condition

$$X = WH \quad \text{is SSC if} \quad \mathcal{C} \subset Conv(H)$$

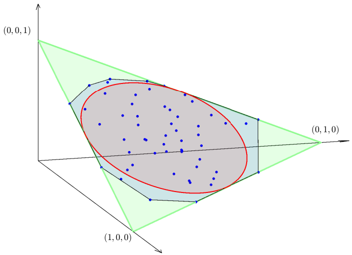
Sufficiently Scattered Condition

$$X = WH \quad \text{is SSC if} \quad \mathcal{C} \subset Conv(H)$$

Theorem (Fu, Ma, Huang, Sidiropoulos, 2015)

$X = WH$ SSC is the unique solution to

$$\min_{W \in \mathbb{R}^{r-1 \times r}} Vol(W) : Conv(X) \subseteq Conv(W)$$



× Non-convex

× Robustness to perturbation not understood

Notice: Separability $\implies$ H contains I as submatrix $\implies$ SSC

Change of Paradigm: Instead of looking for the vertices of $Conv(W)$ let us look for its *Facets*

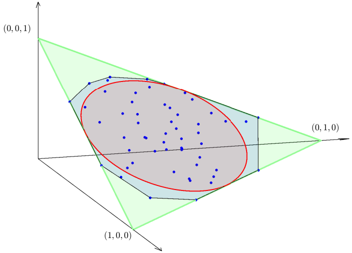
Sufficiently Scattered Condition

$$X = WH \quad \text{is SSC if} \quad \mathcal{C} \subset Conv(H)$$

Theorem (Fu, Ma, Huang, Sidiropoulos, 2015)

$X = WH$ SSC is the unique solution to

$$\min_{W \in \mathbb{R}^{r-1 \times r}} Vol(W) : Conv(X) \subseteq Conv(W)$$



- × Non-convex
- × Robustness to perturbation not understood

Notice: Separability $\Longrightarrow$ H contains I as submatrix $\Longrightarrow$ SSC

Change of Paradigm: Instead of looking for the vertices of $Conv(W)$ let us look for its *Facets*

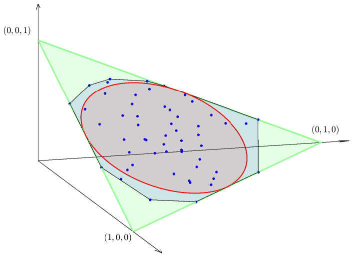
Sufficiently Scattered Condition

$$X = WH \quad \text{is SSC if} \quad \mathcal{C} \subset Conv(H)$$

Theorem (Fu, Ma, Huang, Sidiropoulos, 2015)

$X = WH$ SSC is the unique solution to

$$\min_{W \in \mathbb{R}^{r-1 \times r}} Vol(W) : Conv(X) \subseteq Conv(W)$$



- × Non-convex
- × Robustness to perturbation not understood

Notice: Separability $\implies$ H contains I as submatrix $\implies$ SSC

Change of Paradigm: Instead of looking for the vertices of $Conv(W)$ let us look for its *Facets*

Sufficiently Scattered Condition

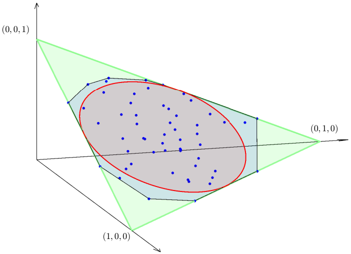
$$X = WH \quad \text{is SSC if} \quad \mathcal{C} \subset Conv(H)$$

Theorem (Fu, Ma, Huang, Sidiropoulos, 2015)

$X = WH$ SSC is the unique solution to

$$\min_{W \in \mathbb{R}^{r-1 \times r}} Vol(W) : Conv(X) \subseteq Conv(W)$$

- × Non-convex
- × Robustness to perturbation not understood



Notice: Separability $\implies$ H contains I as submatrix $\implies$ SSC

Change of Paradigm: Instead of looking for the vertices of $Conv(W)$ let us look for its *Facets*

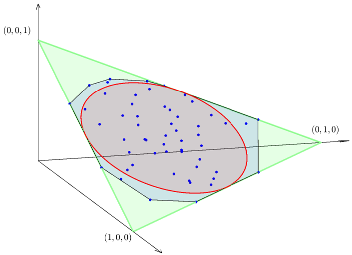
Sufficiently Scattered Condition

$$X = WH \quad \text{is SSC if} \quad \mathcal{C} \subset Conv(H)$$

Theorem (Fu, Ma, Huang, Sidiropoulos, 2015)

$X = WH$ SSC is the unique solution to

$$\min_{W \in \mathbb{R}^{r-1 \times r}} Vol(W) : Conv(X) \subseteq Conv(W)$$



- × Non-convex
- × Robustness to perturbation not understood

Notice: Separability $\implies$ H contains I as submatrix $\implies$ SSC

Change of Paradigm: Instead of looking for the vertices of $Conv(W)$ let us look for its *Facets*

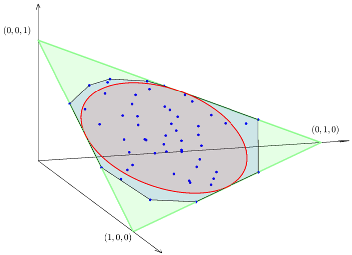
Sufficiently Scattered Condition

$$X = WH \quad \text{is SSC if} \quad \mathcal{C} \subset Conv(H)$$

Theorem (Fu, Ma, Huang, Sidiropoulos, 2015)

$X = WH$ SSC is the unique solution to

$$\min_{W \in \mathbb{R}^{r-1 \times r}} Vol(W) : Conv(X) \subseteq Conv(W)$$



× Non-convex

× Robustness to perturbation not understood

Notice: Separability $\implies$ H contains I as submatrix $\implies$ SSC

Change of Paradigm: Instead of looking for the vertices of $Conv(W)$ let us look for its *Facets*

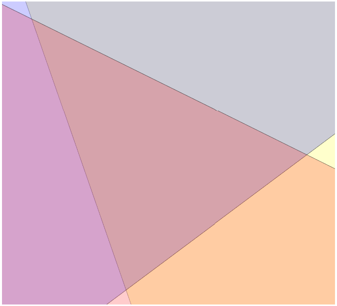
Facet Identification

Facet Based Algorithms

$$Conv(W) = \cap_{i=1}^{r} \mathcal{S}_i \quad \text{where} \quad \mathcal{S}_i := \{x : \theta_i^T x \leq 1\}$$

$$Conv(X) \subseteq Conv(W) \qquad \Longleftrightarrow \qquad \Theta = \begin{pmatrix} \theta_1 & \dots & \theta_r \end{pmatrix} \qquad \Theta^T X \leq 1$$

MVIE *Maximum Volume Inscribed Ellipsoid*
Enumerates the facets of $Conv(X)$, very expensive
(Lin, Wu, Ma, Chi, Wang, 2018)



GFPI *Greedy Facet-based Polytope Identification*
Mixed integer programming, also expensive
(Abdolali, Gillis, 2021)

In order to deal with facets GFPI works in the **Polar Space**

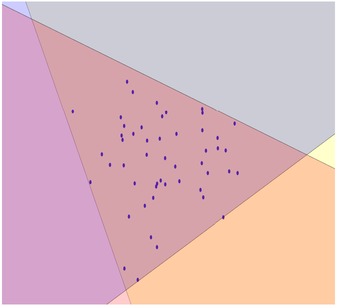
Facet Based Algorithms

$$Conv(W) = \cap_{i=1}^{r} \mathcal{S}_i \quad \text{where} \quad \mathcal{S}_i := \{x : \theta_i^T x \leq 1\}$$

$$Conv(X) \subseteq Conv(W) \qquad \Longleftrightarrow \qquad \Theta = \begin{pmatrix} \theta_1 & \dots & \theta_r \end{pmatrix} \qquad \Theta^T X \leq 1$$

MVIE *Maximum Volume Inscribed Ellipsoid*
Enumerates the facets of $Conv(X)$, very expensive
(Lin, Wu, Ma, Chi, Wang, 2018)

GFPI *Greedy Facet-based Polytope Identification*
Mixed integer programming, also expensive
(Abdolali, Gillis, 2021)



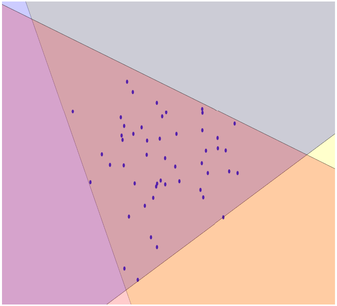
In order to deal with facets GFPI works in the **Polar Space**

Facet Based Algorithms

$$Conv(W) = \cap_{i=1}^{r} \mathcal{S}_i \quad \text{where} \quad \mathcal{S}_i := \{x : \theta_i^T x \leq 1\}$$

$$Conv(X) \subseteq Conv(W) \quad \Longleftrightarrow \quad \Theta = \begin{pmatrix} \theta_1 & \dots & \theta_r \end{pmatrix} \quad \Theta^T X \leq 1$$

MVIE *Maximum Volume Inscribed Ellipsoid*
Enumerates the facets of $Conv(X)$, very expensive
(Lin, Wu, Ma, Chi, Wang, 2018)



GFPI *Greedy Facet-based Polytope Identification*
Mixed integer programming, also expensive
(Abdolali, Gillis, 2021)

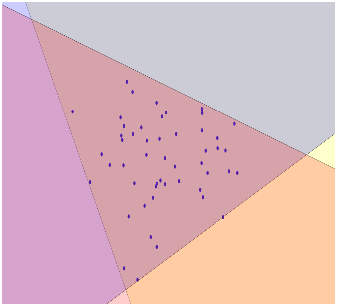
In order to deal with facets GFPI works in the **Polar Space**

Facet Based Algorithms

$$Conv(W) = \cap_{i=1}^{r} \mathcal{S}_i \quad \text{where} \quad \mathcal{S}_i := \{x : \theta_i^T x \leq 1\}$$

$$Conv(X) \subseteq Conv(W) \qquad \Longleftrightarrow \qquad \Theta = \begin{pmatrix} \theta_1 & \dots & \theta_r \end{pmatrix} \qquad \Theta^T X \leq 1$$

MVIE *Maximum Volume Inscribed Ellipsoid*
Enumerates the facets of $Conv(X)$, very expensive
(Lin, Wu, Ma, Chi, Wang, 2018)



GFPI *Greedy Facet-based Polytope Identification*
Mixed integer programming, also expensive
(Abdolali, Gillis, 2021)

In order to deal with facets GFPI works in the **Polar Space**

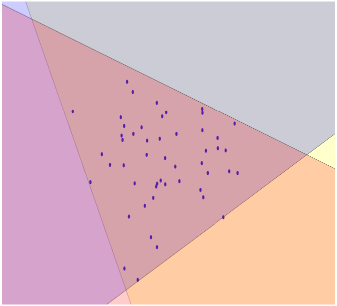
Facet Based Algorithms

$$Conv(W) = \cap_{i=1}^{r} \mathcal{S}_i \quad \text{where} \quad \mathcal{S}_i := \{x : \theta_i^T x \leq 1\}$$

$$Conv(X) \subseteq Conv(W) \qquad \Longleftrightarrow \qquad \Theta = \begin{pmatrix} \theta_1 & \dots & \theta_r \end{pmatrix} \qquad \Theta^T X \leq 1$$

MVIE *Maximum Volume Inscribed Ellipsoid*
Enumerates the facets of $Conv(X)$, very expensive
(Lin, Wu, Ma, Chi, Wang, 2018)

GFPI *Greedy Facet-based Polytope Identification*
Mixed integer programming, also expensive
(Abdolali, Gillis, 2021)



In order to deal with facets GFPI works in the **Polar Space**

Polarity

$$\mathcal{S} \subseteq \mathbb{R}^{r-1} \qquad \mathcal{S}^* := \{\theta : \theta^T x \leq 1 \ \forall x \in \mathcal{S}\}$$

- Swaps points and hyperplanes

$$\{x : \theta^T x = 1\} \rightsquigarrow \theta$$

- Sends simplexes into simplexes
- It is an involution for convex sets
- **Reverses Containments**

$$Conv(X) \subseteq Conv(W) \iff Conv(W)^* \subseteq Conv(X)^*$$

$$\iff \Theta^T X \leq 1 \quad \text{where} \quad Conv(W)^* = Conv(\Theta)$$

We can thus seek the simplex Θ with **maximum volume** inside $Conv(X)^*$ as in

$$\max_{\theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T X \leq 1 \qquad (MaxVol)$$

Polarity

$$\mathcal{S} \subseteq \mathbb{R}^{r-1} \qquad \mathcal{S}^* := \{\theta : \theta^T x \leq 1 \ \forall x \in \mathcal{S}\}$$

- Swaps points and hyperplanes

$$\{x : \theta^T x = 1\} \rightsquigarrow \theta$$

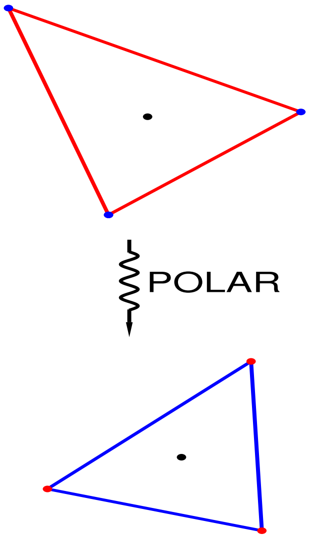
- Sends simplexes into simplexes
- It is an involution for convex sets
- **Reverses Containments**

$$Conv(X) \subseteq Conv(W) \iff Conv(W)^* \subseteq Conv(X)^*$$

$$\iff \Theta^T X \leq 1 \quad \text{where} \quad Conv(W)^* = Conv(\Theta)$$

We can thus seek the simplex Θ with **maximum volume** inside $Conv(X)^*$ as in

$$\max_{\theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T X \leq 1 \qquad (MaxVol)$$



Polarity

$$\mathcal{S} \subseteq \mathbb{R}^{r-1} \qquad \mathcal{S}^* := \{\theta : \theta^T x \leq 1 \ \forall x \in \mathcal{S}\}$$

- Swaps points and hyperplanes

$$\{x : \theta^T x = 1\} \rightsquigarrow \theta$$

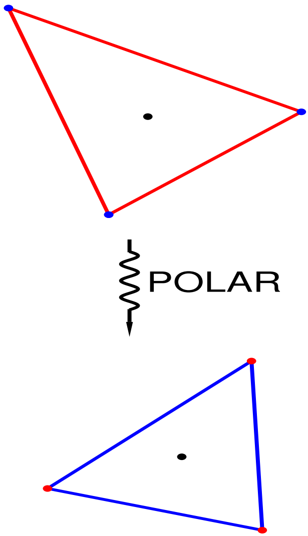
- Sends simplexes into simplexes
- It is an involution for convex sets
- **Reverses Containments**

$$Conv(X) \subseteq Conv(W) \iff Conv(W)^* \subseteq Conv(X)^*$$

$$\iff \Theta^T X \leq 1 \quad \text{where} \quad Conv(W)^* = Conv(\Theta)$$

We can thus seek the simplex Θ with **maximum volume** inside $Conv(X)^*$ as in

$$\max_{\theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T X \leq 1 \qquad (MaxVol)$$



Polarity

$$\mathcal{S} \subseteq \mathbb{R}^{r-1} \qquad \mathcal{S}^* := \{\theta : \theta^T x \leq 1 \ \forall x \in \mathcal{S}\}$$

- Swaps points and hyperplanes

$$\{x : \theta^T x = 1\} \rightsquigarrow \theta$$

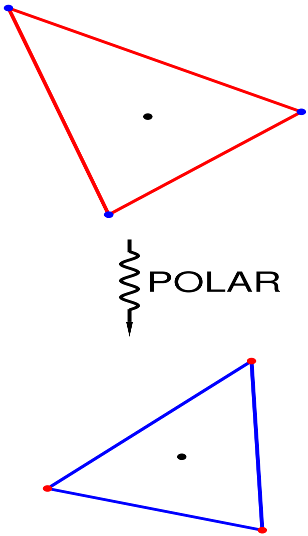
- Sends simplexes into simplexes
- It is an involution for convex sets
- **Reverses Containments**

$$Conv(X) \subseteq Conv(W) \iff Conv(W)^* \subseteq Conv(X)^*$$

$$\iff \Theta^T X \leq 1 \quad \text{where} \quad Conv(W)^* = Conv(\Theta)$$

We can thus seek the simplex Θ with **maximum volume** inside $Conv(X)^*$ as in

$$\max_{\theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T X \leq 1 \qquad (MaxVol)$$



Polarity

$$\mathcal{S} \subseteq \mathbb{R}^{r-1} \qquad \mathcal{S}^* := \{\theta : \theta^T x \le 1 \ \forall x \in \mathcal{S}\}$$

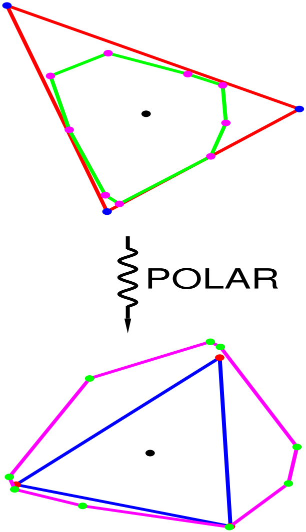
- Swaps points and hyperplanes

$$\{x : \theta^T x = 1\} \rightsquigarrow \theta$$

- Sends simplexes into simplexes
- It is an involution for convex sets
- **Reverses Containments**

$$Conv(X) \subseteq Conv(W) \iff Conv(W)^* \subseteq Conv(X)^*$$

$$\iff \Theta^T X \le 1 \quad \text{where} \quad Conv(W)^* = Conv(\Theta)$$



We can thus seek the simplex Θ with **maximum volume** inside $Conv(X)^*$ as in

$$\max_{\theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T X \le 1 \qquad (MaxVol)$$

Polarity

$$\mathcal{S} \subseteq \mathbb{R}^{r-1} \qquad \mathcal{S}^* := \{\theta : \theta^T x \leq 1 \ \forall x \in \mathcal{S}\}$$

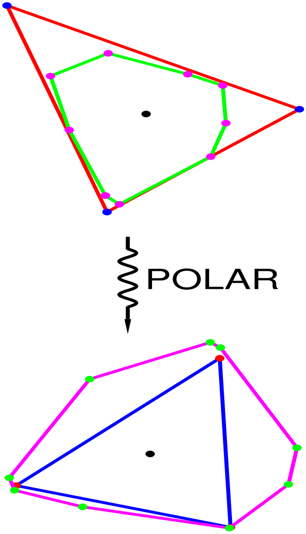
- Swaps points and hyperplanes

$$\{x : \theta^T x = 1\} \rightsquigarrow \theta$$

- Sends simplexes into simplexes
- It is an involution for convex sets
- **Reverses Containments**

$$Conv(X) \subseteq Conv(W) \iff Conv(W)^* \subseteq Conv(X)^*$$

$$\iff \Theta^T X \leq 1 \quad \text{where} \quad Conv(W)^* = Conv(\Theta)$$



We can thus seek the simplex Θ with **maximum volume** inside $Conv(X)^*$ as in

$$\max_{\theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T X \leq 1 \qquad (MaxVol)$$

Identifiability and η-Expansion

Theorem (M.A., G.B., N.G., 2023)

Let $X = WH \in \mathbb{R}^{r-\mathbf{1}\times n}$ SSC and for any $v \in \mathbb{R}^{r-\mathbf{1}}$ define

$$\mathcal{V}(v) := \max_{\Theta \in \mathbb{R}^{r-\mathbf{1}\times r}} Vol(\Theta) \quad : \quad \Theta^T(X - ve^T) \leq 1$$

Then $\mathcal{V}(v)$ is convex in v with unique minimum for $v = We/r$ and Θ polar of W

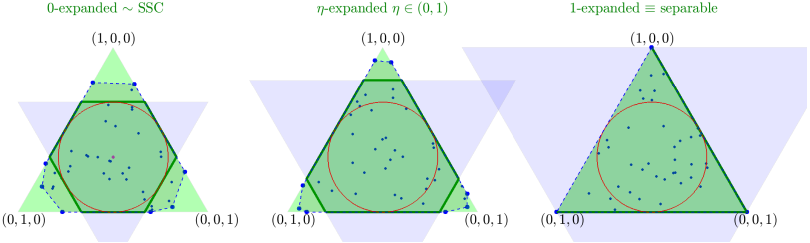
Identifiability and η-Expansion

Theorem (M.A., G.B., N.G., 2023)

Let $X = WH \in \mathbb{R}^{r-\mathbf{1} \times n}$ SSC and for any $v \in \mathbb{R}^{r-\mathbf{1}}$ define

$$\mathcal{V}(v) := \max_{\Theta \in \mathbb{R}^{r-\mathbf{1} \times r}} Vol(\Theta) \quad : \quad \Theta^T (X - ve^T) \leq 1$$

Then $\mathcal{V}(v)$ is convex in v with unique minimum for $v = We/r$ and Θ polar of W



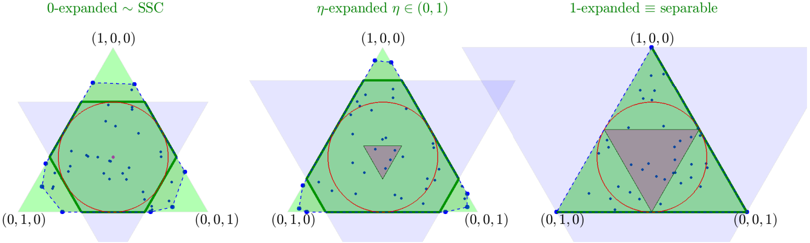
Identifiability and η-Expansion

Theorem (M.A., G.B., N.G., 2023)

Let $X = WH \in \mathbb{R}^{r-1 \times n}$ SSC and for any $v \in \mathbb{R}^{r-1}$ define

$$\mathcal{V}(v) := \max_{\Theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T (X - ve^T) \leq 1$$

Then $\mathcal{V}(v)$ is convex in v with unique minimum for $v = We/r$ and Θ polar of W



Theorem (M.A., G.B., N.G., 2023)

Let $X = WH \in \mathbb{R}^{r-1 \times n}$ be η-expanded and suppose $v = Wh$, $h \in \blacktriangledown$. Then

$$\max_{\Theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T (X - ve^T) \leq 1$$

is solved uniquely by Θ polar of W

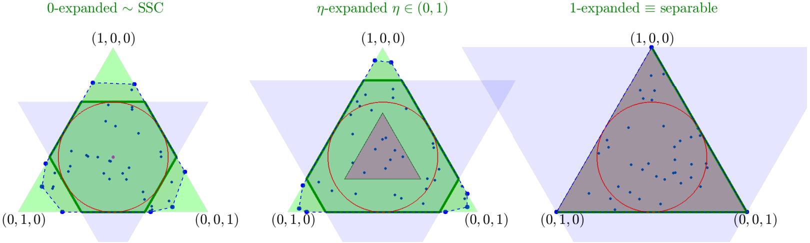
Identifiability and η-Expansion

Theorem (M.A., G.B., N.G., 2023)

Let $X = WH \in \mathbb{R}^{r-1 \times n}$ SSC and for any $v \in \mathbb{R}^{r-1}$ define

$$\mathcal{V}(v) := \max_{\Theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T(X - ve^T) \leq 1$$

Then $\mathcal{V}(v)$ is convex in v with unique minimum for $v = We/r$ and Θ polar of W



Conjecture (M.A., G.B., N.G., 2023)

Let $X = WH \in \mathbb{R}^{r-1 \times n}$ be η-expanded and suppose $v = Wh$, $h \in$ ▲. Then

$$\max_{\Theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) \quad : \quad \Theta^T(X - ve^T) \leq 1$$

is solved uniquely by Θ polar of W

Maximum Volume in Dual

Algorithm 1 Maximum Volume in the Dual (MV-Dual)

Input: Data matrix $\widetilde{X} \in \mathbb{R}^{m \times n}$ and a factorization rank r

Output: A matrix $\widetilde{W} \in \mathbb{R}^{m \times r}$ and a vector w such that $\widetilde{X} \approx w + \widetilde{W}H$ where H is column stochastic

Use PCA to reduce $\widetilde{X} = w + UX$ with $X \in \mathbb{R}^{r-1 \times n}$
Initialize $v_1 = Xe/n$, $p = 1$ and $\Theta \in \mathcal{N}(0,1)^{r-1 \times r}$
while not converged: $p = 1$ or $\frac{\|v_p - v_{p-1}\|_2}{\|v_{p-1}\|_2} > 0.01$ **do**
    Solve

$$\arg \max_{\Theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) : \Theta^T (X - v_p e^T) \leq 1$$

    via alternating optimization on the columns of Θ
    Recover W by computing the polar of $Conv\,(\Theta)$
    Let $v_{p+1} \leftarrow We/r$, and $p = p + 1$
end while
Compute $\widetilde{W} = UW$

Cost : PCA $\mathcal{O}(mnr)$ plus Maximization problem solver for a single column $\mathcal{O}(nr^2)$ times the number of iterations

Maximum Volume in Dual

Algorithm 2 Maximum Volume in the Dual (MV-Dual)

Input: Data matrix $\widetilde{X} \in \mathbb{R}^{m \times n}$ and a factorization rank r

Output: A matrix $\widetilde{W} \in \mathbb{R}^{m \times r}$ and a vector w such that $\widetilde{X} \approx w + \widetilde{W}H$ where H is column stochastic

Use PCA to reduce $\widetilde{X} = w + UX$ with $X \in \mathbb{R}^{r-1 \times n}$

Initialize $v_1 = Xe/n$, $p = 1$ and $\Theta \in \mathcal{N}(0,1)^{r-1 \times r}$

while not converged: $p = 1$ or $\frac{\|v_p - v_{p-1}\|_2}{\|v_{p-1}\|_2} > 0.01$ **do**

  Solve

$$\arg \max_{\Theta \in \mathbb{R}^{r-1 \times r}} Vol(\Theta) : \Theta^T (X - v_p e^T) \leq 1$$

  via alternating optimization on the columns of Θ

  Recover W by computing the polar of $Conv(\Theta)$

  Let $v_{p+1} \leftarrow We/r$, and $p = p + 1$

end while

Compute $\widetilde{W} = UW$

Cost : PCA $\mathcal{O}(mnr)$ plus Maximization problem solver for a single column $\mathcal{O}(nr^2)$ times the number of iterations

Algorithm 3 Maximum Volume in the Dual (MV-Dual)

Input: Data matrix $\widetilde{X} \in \mathbb{R}^{m \times n}$ and a factorization rank r

Output: A matrix $\widetilde{W} \in \mathbb{R}^{m \times r}$ and a vector w such that $\widetilde{X} \approx w + \widetilde{W}H$ where H is column stochastic

1: Use PCA to reduce $\widetilde{X} = w + UX$ with $X \in \mathbb{R}^{r-1 \times n}$

2: Initialize $v_1 = Xe/n$, $p = 1$ and $\Theta \in \mathcal{N}(0,1)^{r-1 \times r}$

3: **while** not converged: $p = 1$ or $\frac{\|v_p - v_{p-1}\|_2}{\|v_{p-1}\|_2} > 0.01$ **do**

4: Solve

$$\arg \max_{\Theta \in \mathbb{R}^{r-1 \times r}, \Delta \in \mathbb{R}^{r \times n}} Vol(\Theta)^2 - \lambda \|\Delta\|_F^2 : \Theta^T (X - v_p e^T) \leq 1 + \Delta^T$$

via alternating optimization on the columns of Θ, Δ

5: Recover W by computing the polar of $Conv(\Theta)$

6: Let $v_{p+1} \leftarrow We/r$, and $p = p + 1$

7: **end while**

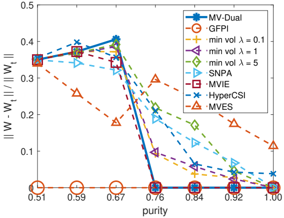
8: Compute $\widetilde{W} = UW$

Cost : PCA $\mathcal{O}(mnr)$ plus Maximization problem solver for a single column $\mathcal{O}(nr^2)$ times the number of iterations

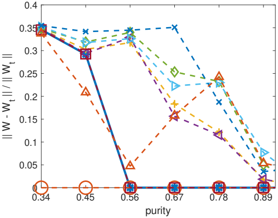
Experiments

Exact Case

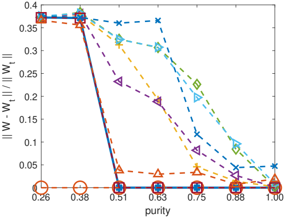
W^*, H^* ground truth $\quad ERR = \min_\pi \frac{\|W^* - W_\pi\|_F}{\|W^*\|_F} \quad$ purity $p = \max_{i,j} |H^*_{i,j}| = \eta + (1-\eta)\frac{2}{r}$



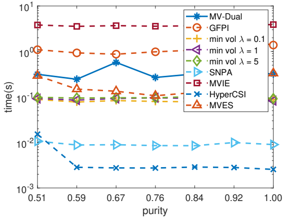
ERR for $r = 3,\ n = 30r$



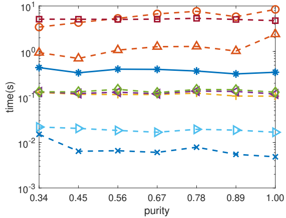
ERR for $r = 4,\ n = 30r$



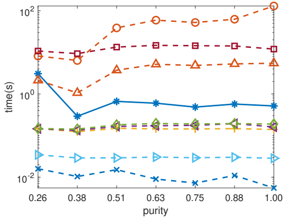
ERR for $r = 5,\ n = 30r$



Time for $r = 3,\ n = 30r$



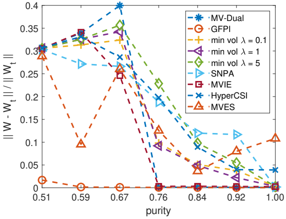
Time for $r = 4,\ n = 30r$



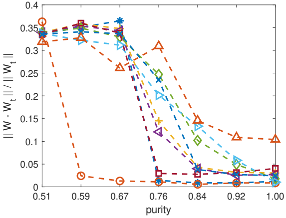
Time for $r = 5,\ n = 30r$

Noisy Case

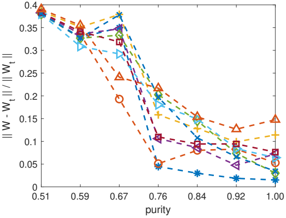
W^*, H^* ground truth $\quad ERR = \min_\pi \frac{\|W^* - W_\pi\|_F}{\|W^*\|_F} \quad$ purity $p = \max_{i,j} |H^*_{i,j}| = \eta + (1-\eta)\frac{2}{r}$



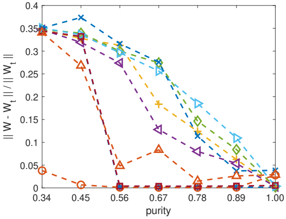
ERR for $r = 3$, SNR = 60



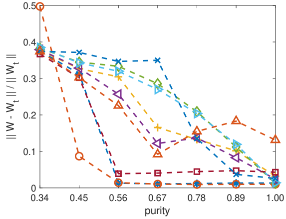
ERR for $r = 3$, SNR = 40



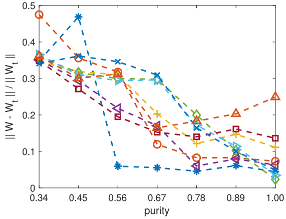
ERR for $r = 3$, SNR = 30



ERR for $r = 4$, SNR = 60



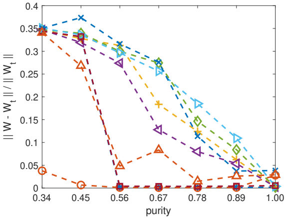
ERR for $r = 4$, SNR = 40



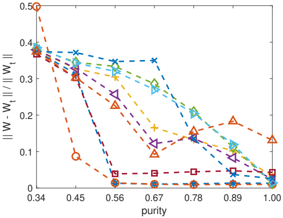
ERR for $r = 4$, SNR = 30

Noisy Case

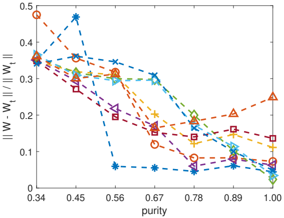
W^*, H^* ground truth $\quad ERR = \min_\pi \frac{\|W^* - W_\pi\|_F}{\|W^*\|_F} \quad$ purity $p = \max_{i,j} |H^*_{i,j}| = \eta + (1-\eta)\frac{2}{r}$



ERR for $r = 4$, SNR $= 60$



ERR for $r = 4$, SNR $= 40$

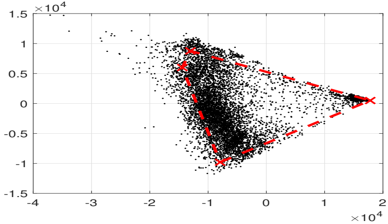


ERR for $r = 4$, SNR $= 30$

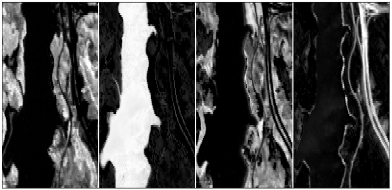
SNR	MVDual	GFPI	min vol $\lambda = 0.1$	min vol $\lambda = 1$	min vol $\lambda = 5$	SNPA	MVIE	HyperCSI	MVES
30	0.56±0.11	7.76±3.51	0.12±0.01	0.13±0.01	0.14±0.02	0.01±0.001	5.28±0.23	0.01±0.004	0.30±0.04
40	0.45±0.06	4.18±1.12	0.10±0.01	0.11±0.01	0.13±0.01	0.01±0.00	4.96±0.12	0.005±0.004	0.30±0.05
60	0.42±0.06	1.47±0.45	0.07±0.01	0.08±0.01	0.09±0.01	0.01±0.00	3.78±0.12	0.001±0.00	0.26±0.07

Unmixing Hyperspectral Imaging

$$\mathrm{MRSA}(x,y) = \frac{100}{\pi}\cos^{-1}\left(\frac{(x-\bar{x}e)^\top (y-\bar{y}e)}{\|x-\bar{x}e\|_2 \|y-\bar{y}e\|_2}\right) \qquad ERR = \min_\pi \mathrm{MRSA}(W_k^*, W_{\pi(k)})$$



Projection of data points
and the symplex computed by MV-Dual



Abundance maps estimated by MV-Dual
From left to right: road, tree, soil, water

	SNPA	Min-Vol	HyperCSI	GFPI	MV-Dual
MRSA	22.27	6.03	17.04	4.82	3.74
Time (s)	0.60	1.45	0.88	100*	43.51

Comparing the performances of MV-Dual with the state-of-the-art SSMF algorithms on Jasper-Ridge data set. Numbers marked with * indicate that the corresponding algorithms did not converge within 100 seconds.

Thank You!

- Abdolali M., Barbarino G., and Gillis N. **Dual simplex volume maximization for simplex-structured matrix factorization.** *Arxiv*, 2024.
- Abdolali M. and Gillis N. **Simplex-structured matrix factorization: Sparsity-based identifiability and provably correct algorithms.** *SIAM Journal on Mathematics of data Science*, 3(2):593–623, 2021.
- Fu X., Ma W.K., Huang K., and Sidiropoulos N.D. **Blind separation of quasi-stationry sources: exploiting convex geometry in covariance domain.** *IEEE Transactions on Signal Processing*, 63(9):2306–2320, 2015.
- Zu F. **Hyperspectral unmixing: ground truth labeling, datasets, benchmark performances and survey.** *Arxiv*, 2017.
- Lin C.H., Wu R., Ma W.K., Chi C.Y., and Wang Y. **Maximum volume inscribed ellipsoid: A new simplex-structured matrix factorization framework via facet enumeration and convex optimization.** *SIAM Journal on Imaging Sciences*, 11(2):1651–1679, 2018.