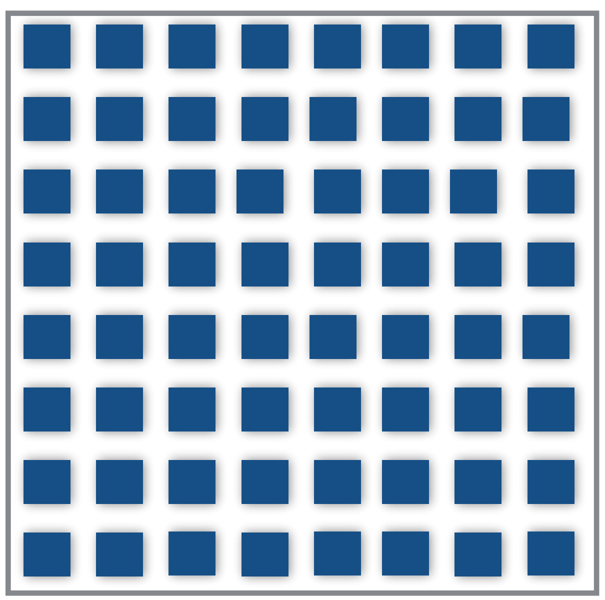
Mixture Matrix Completion:

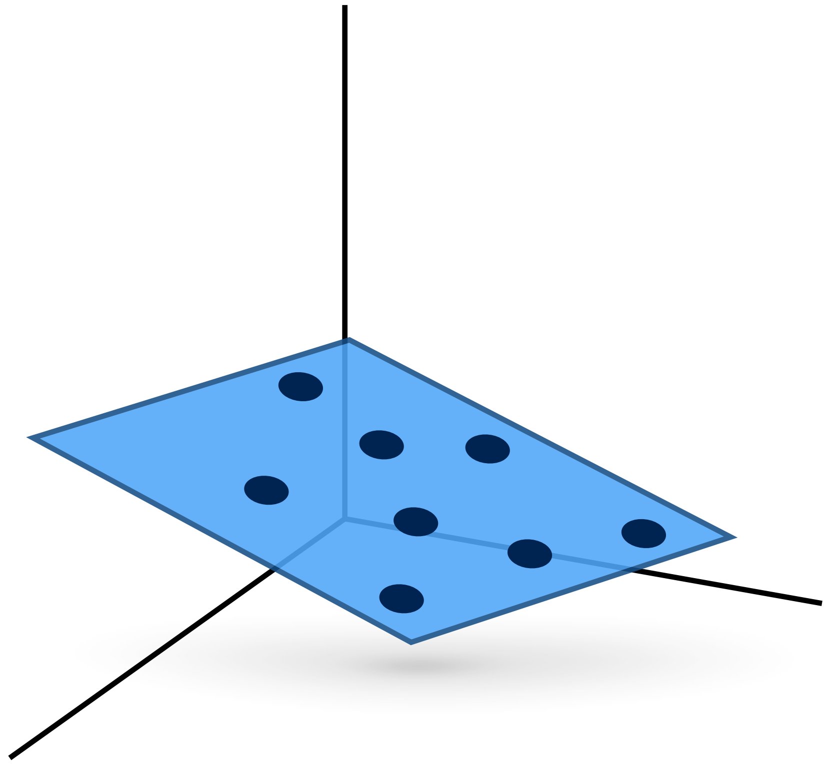
Theory, Algorithms and Open Questions

Daniel Pimentel-Alarcón

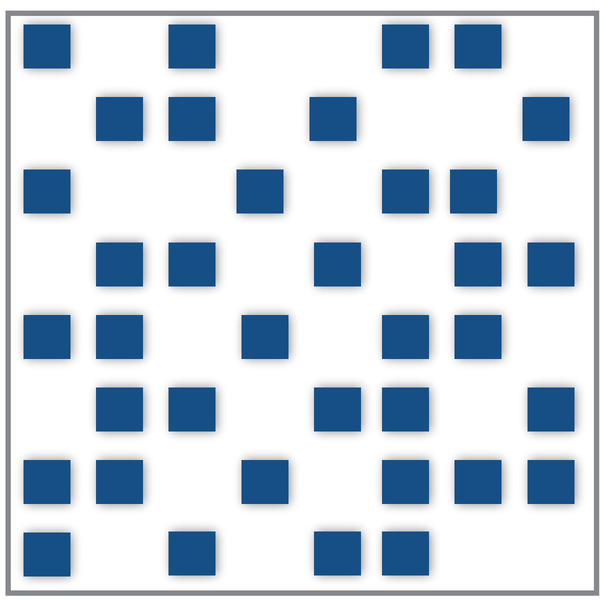
Wisconsin Institute for Discovery
UNIVERSITY *of* WISCONSIN-MADISON
Department of Electrical and Computer Engineering

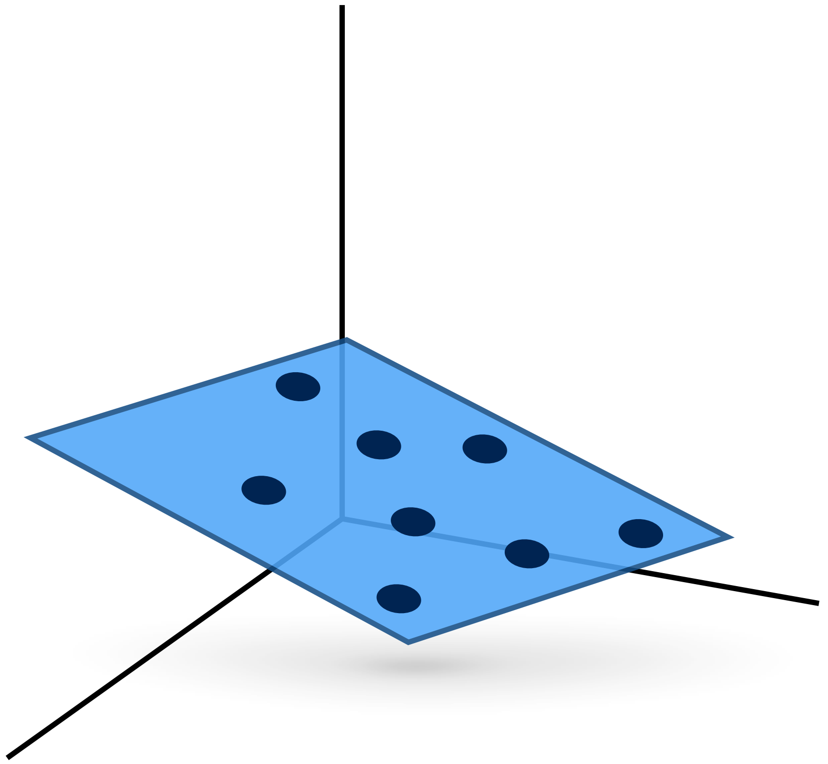
SIAM - Optimization, 2017



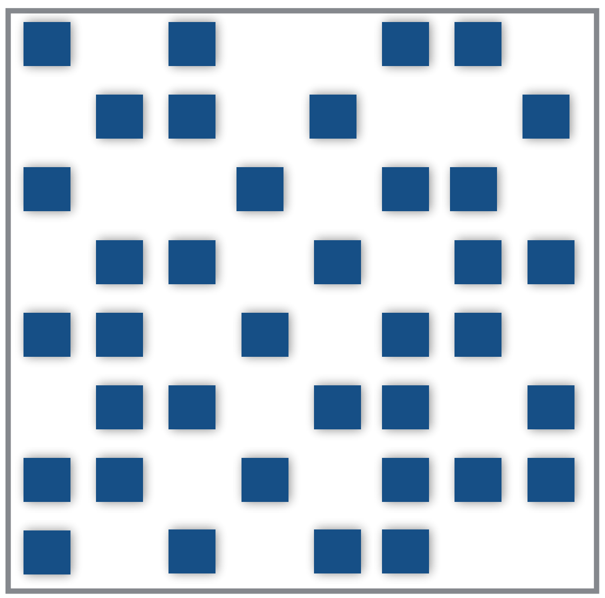


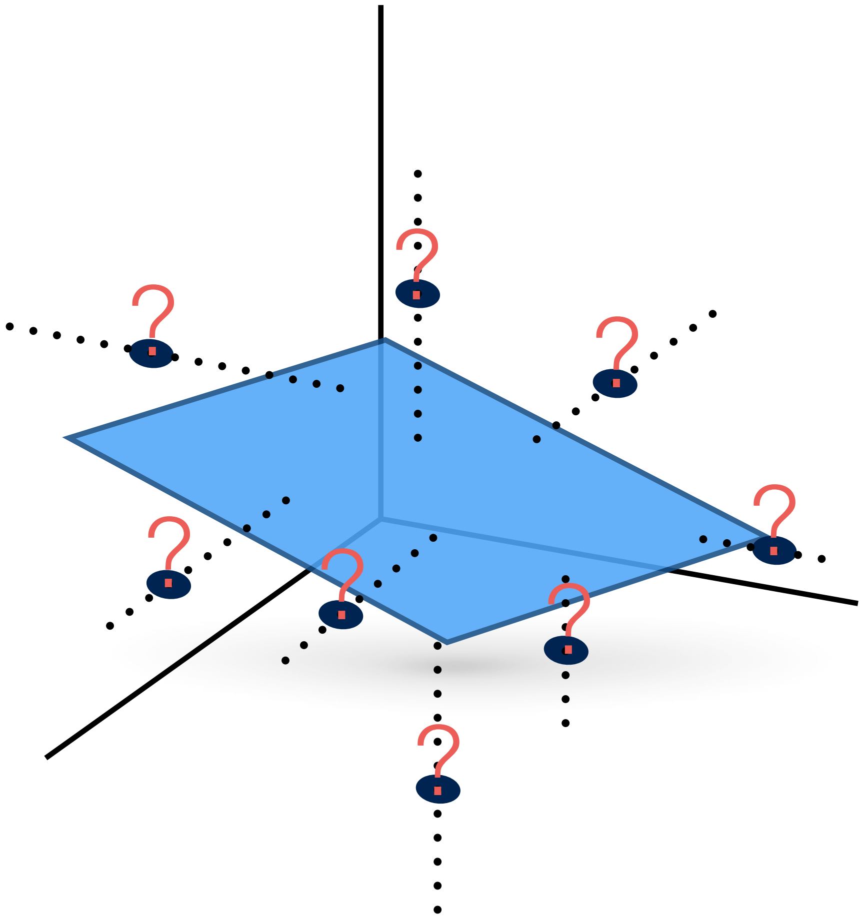
Low-Rank Matrix Completion



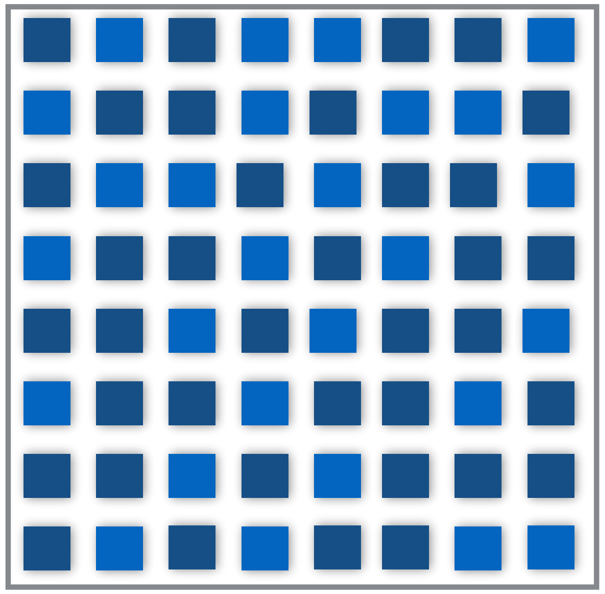


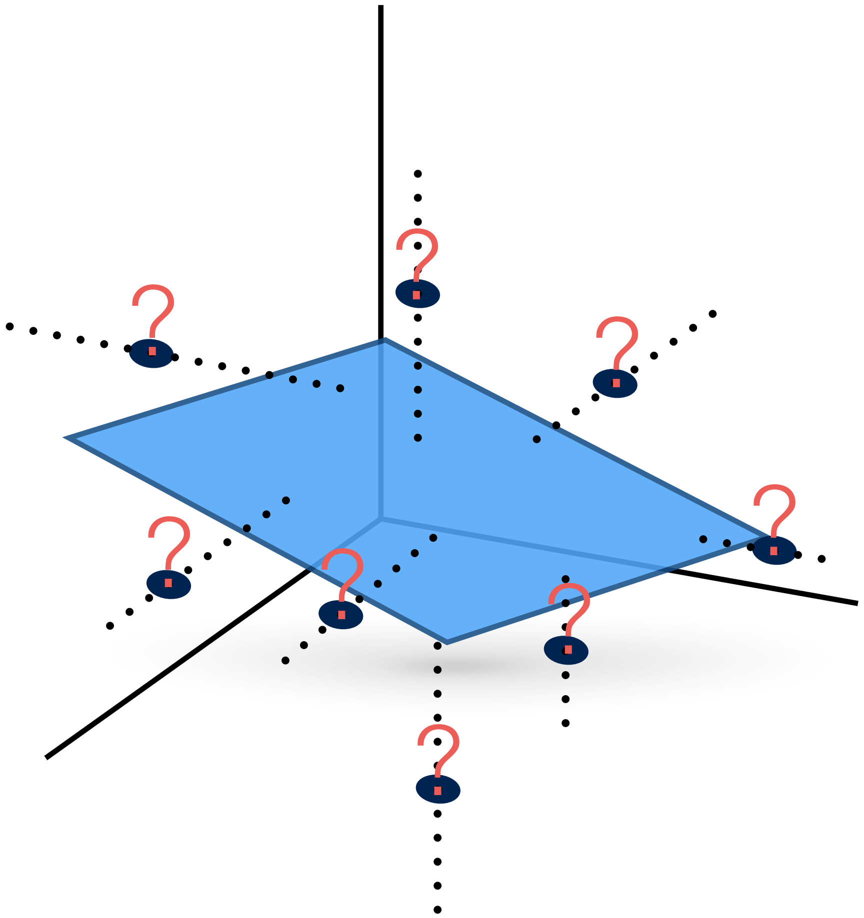
Low-Rank Matrix Completion



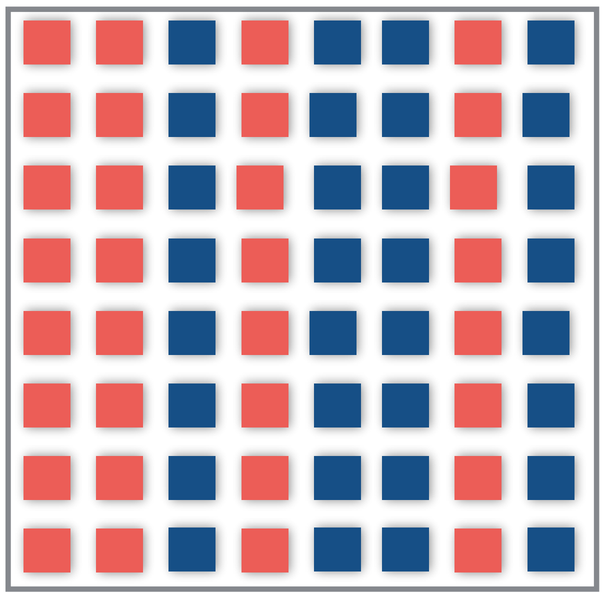


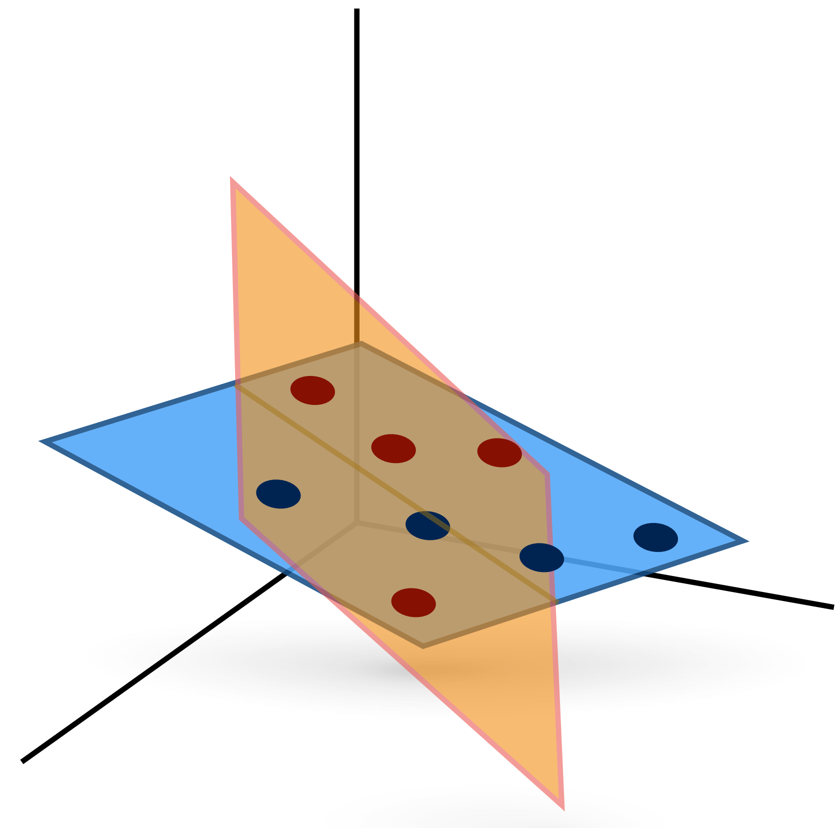
Low-Rank Matrix Completion



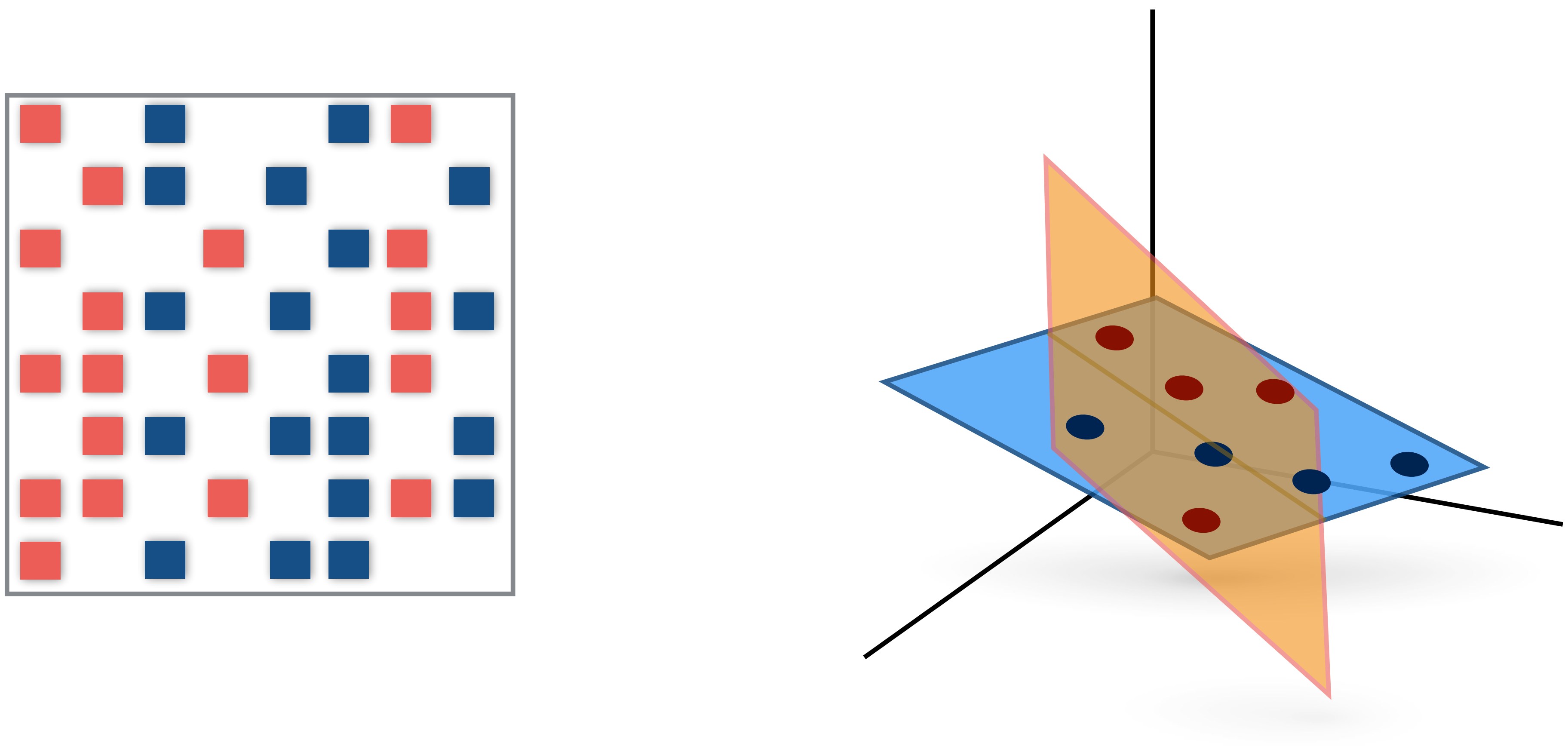


Low-Rank Matrix Completion

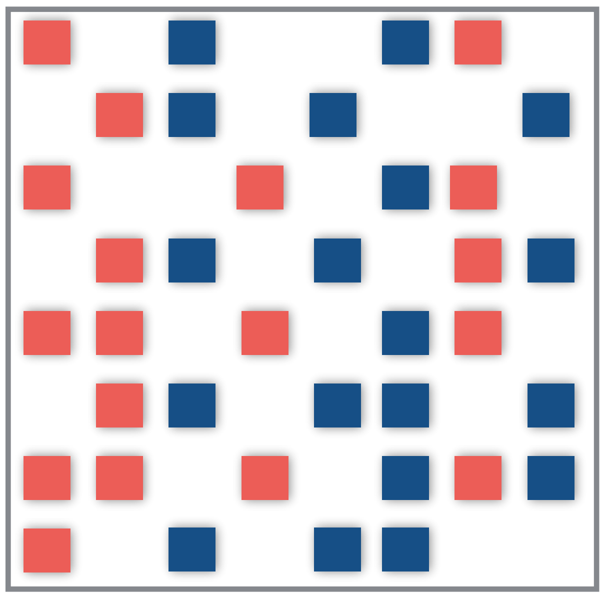
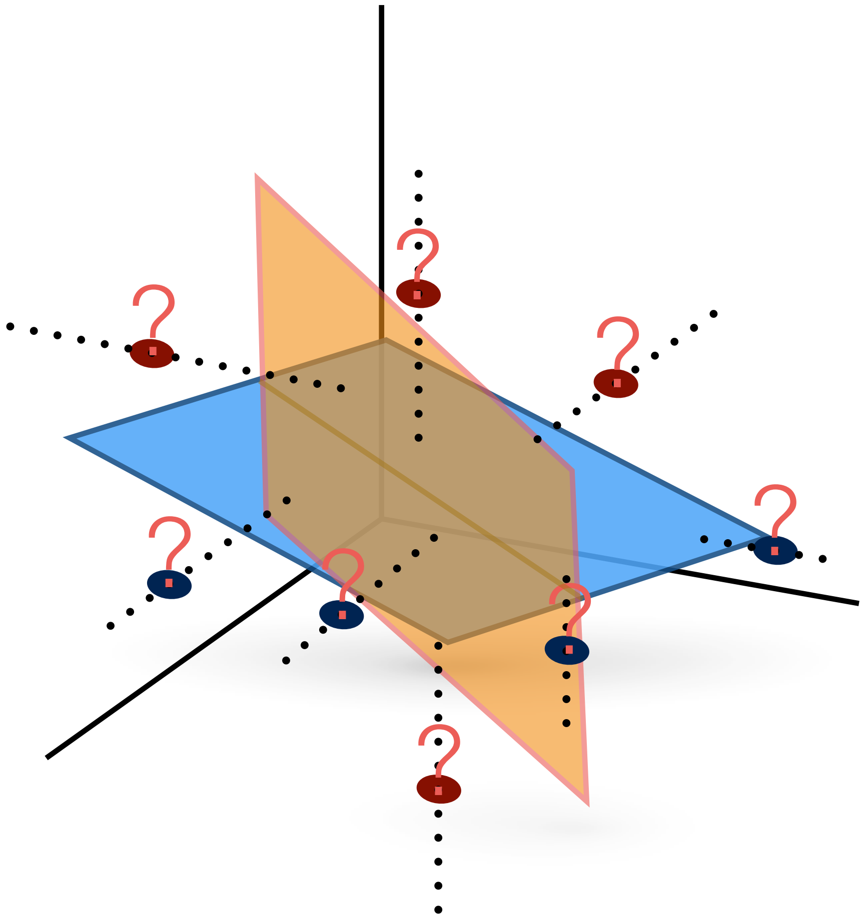




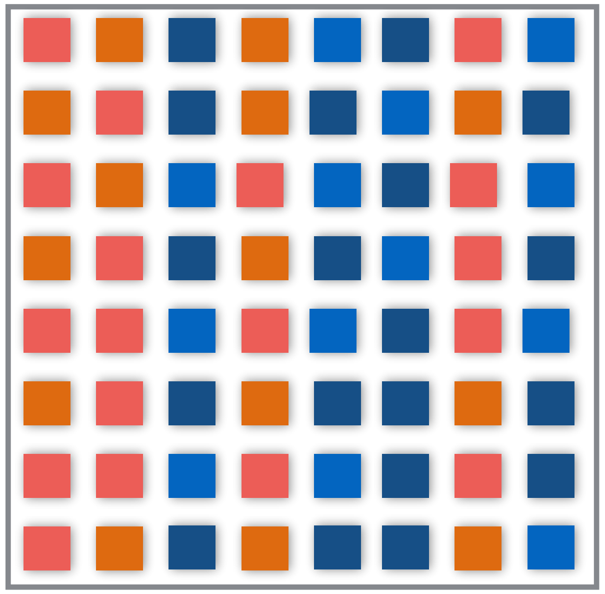
Mixture Matrix Completion

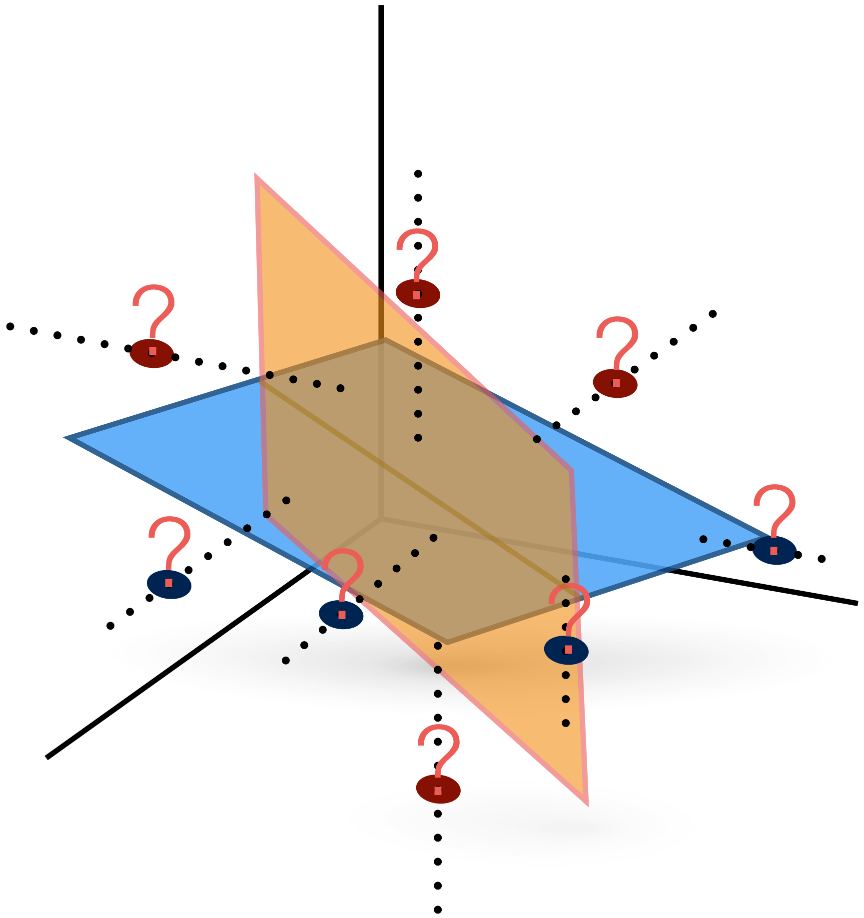


Mixture Matrix Completion

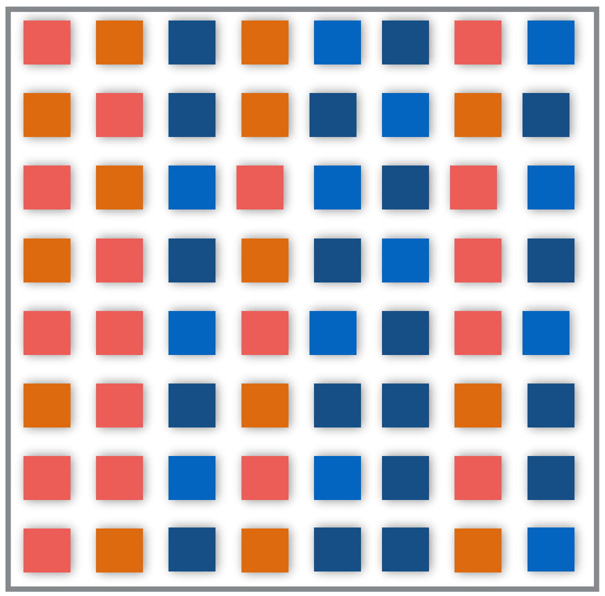
 

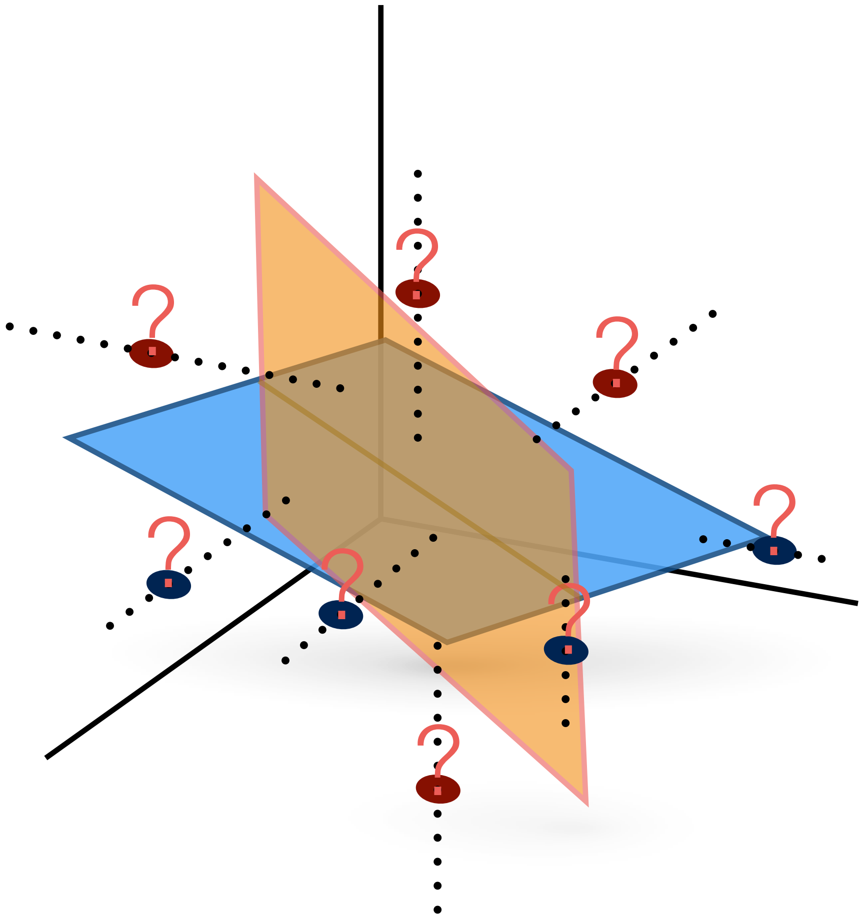
Mixture Matrix Completion





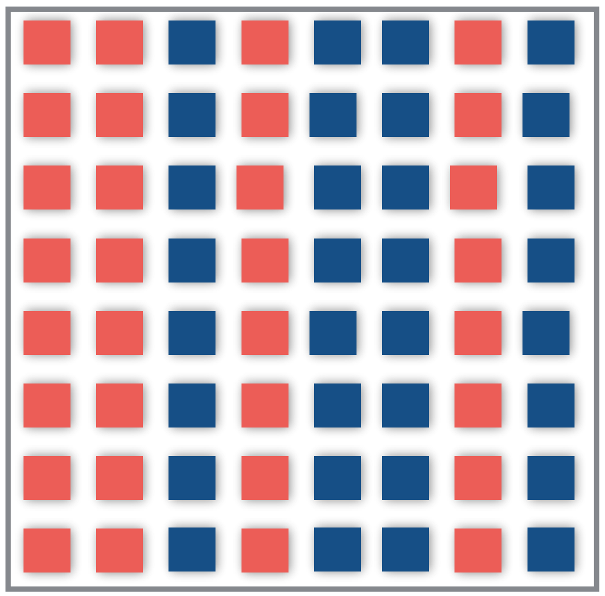
Mixture Matrix Completion

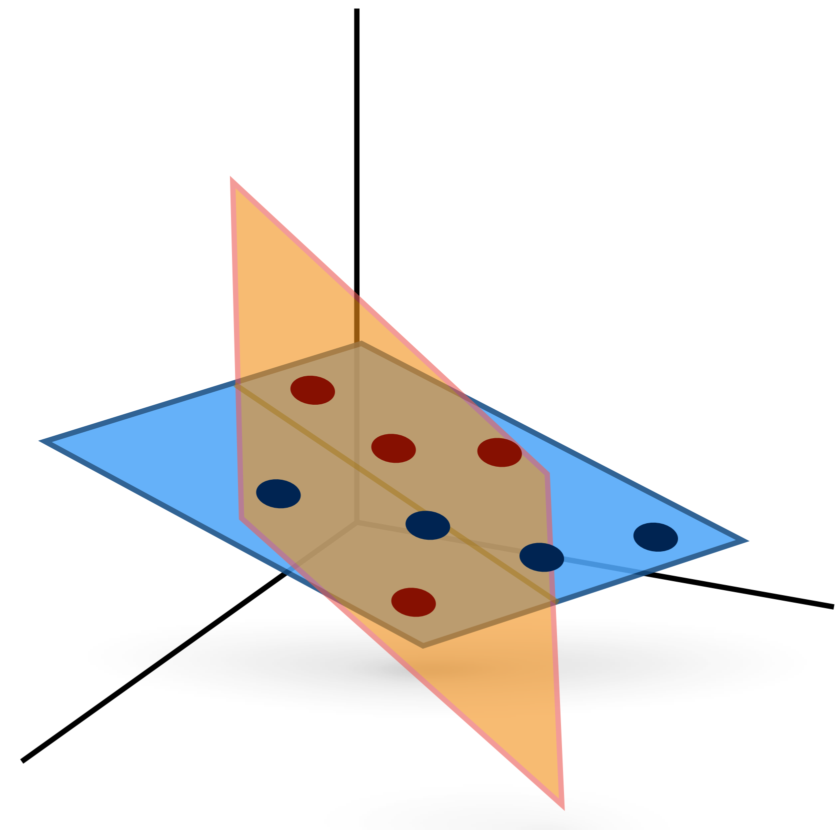




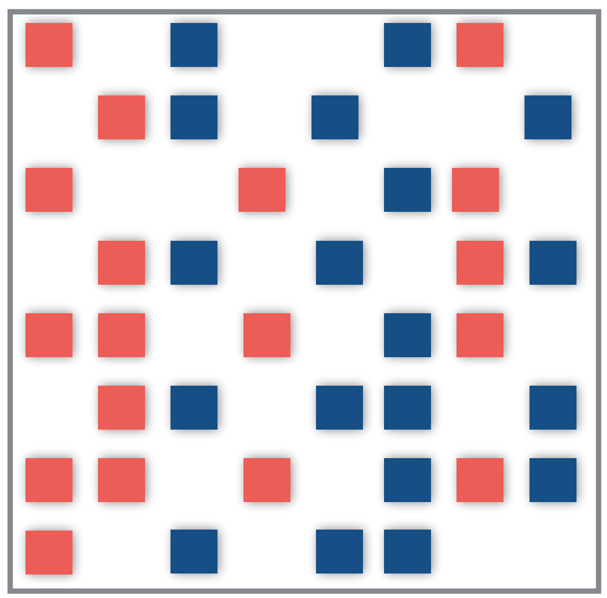
Mixture Matrix Completion

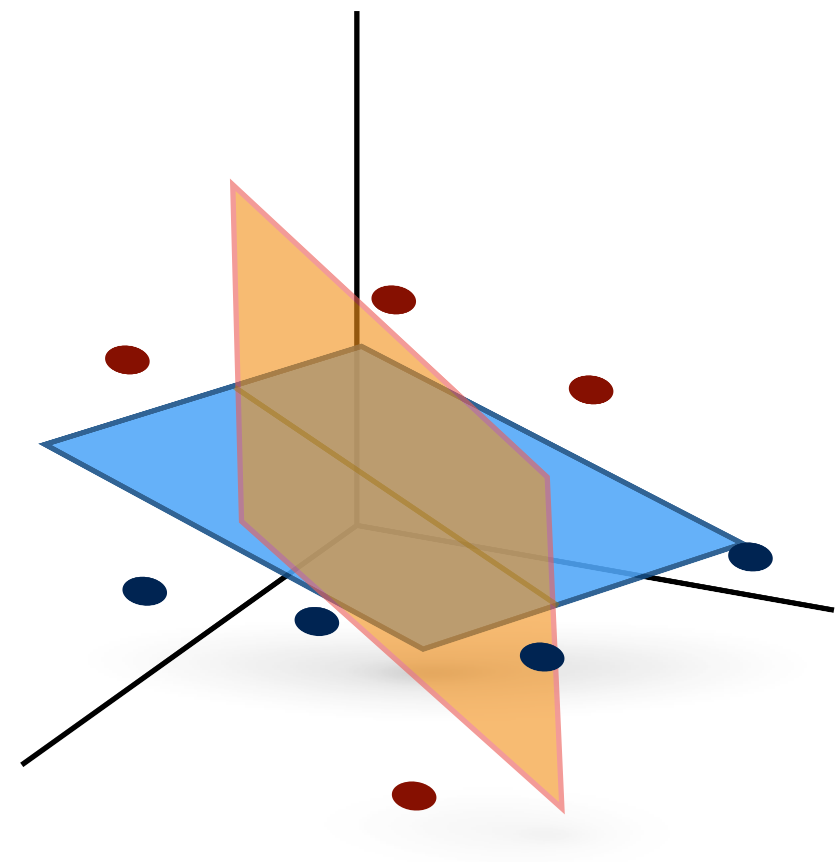
a.k.a. High-Rank Matrix Completion or
Subspace Clustering with Missing Data.



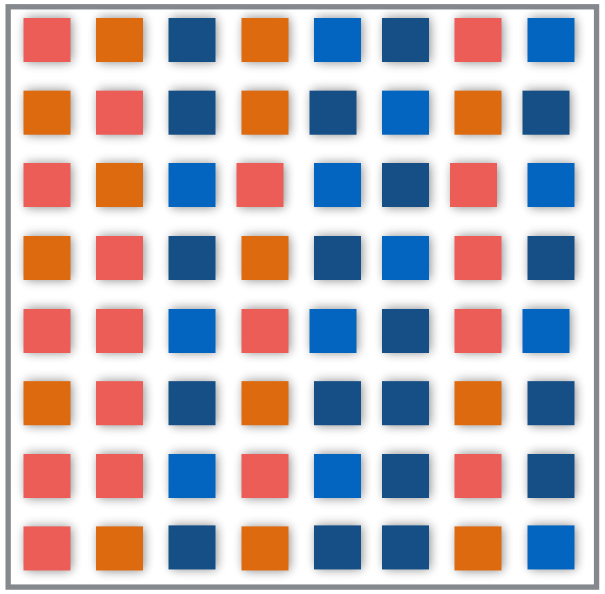
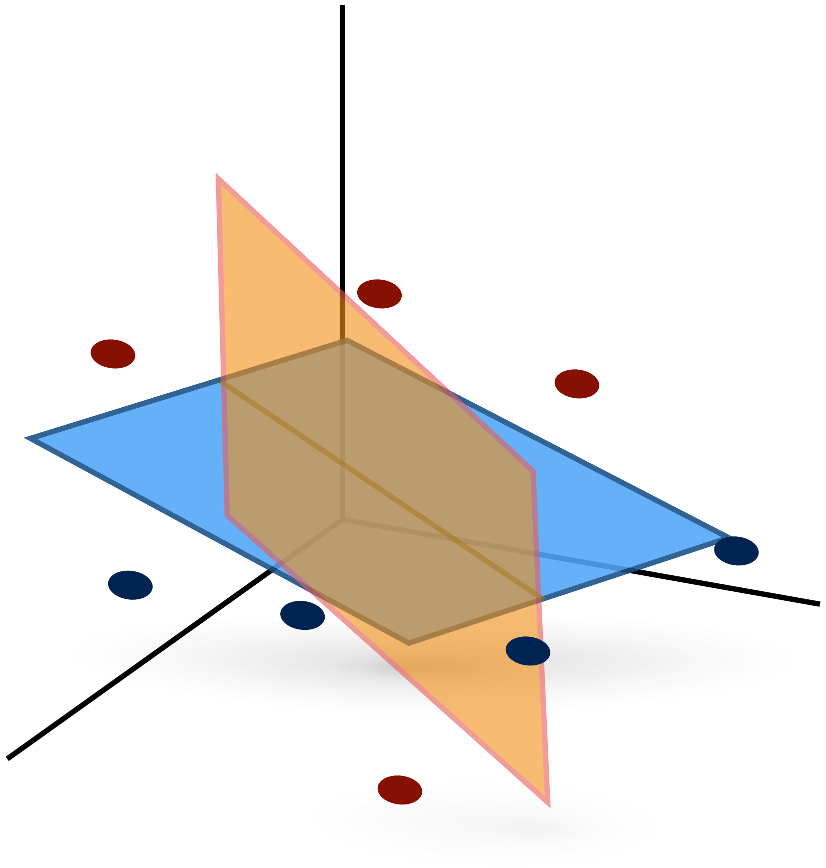


Mixture Matrix Completion

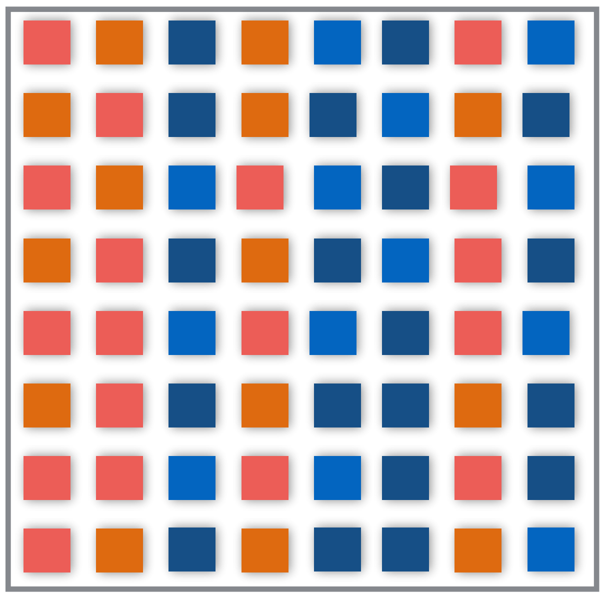


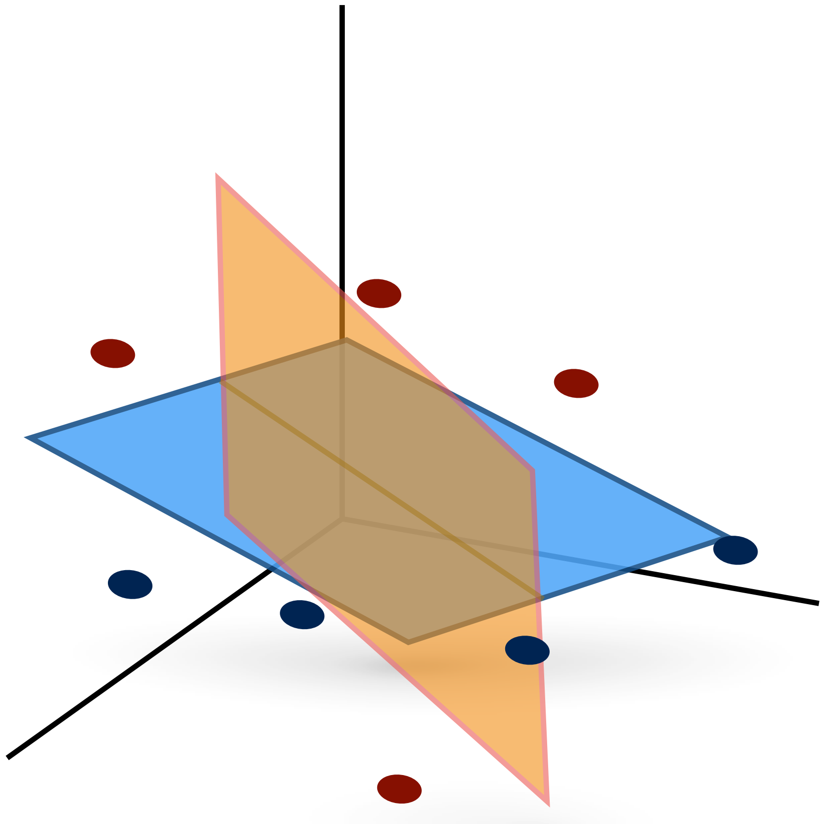


Mixture Matrix Completion

Mixture Matrix Completion

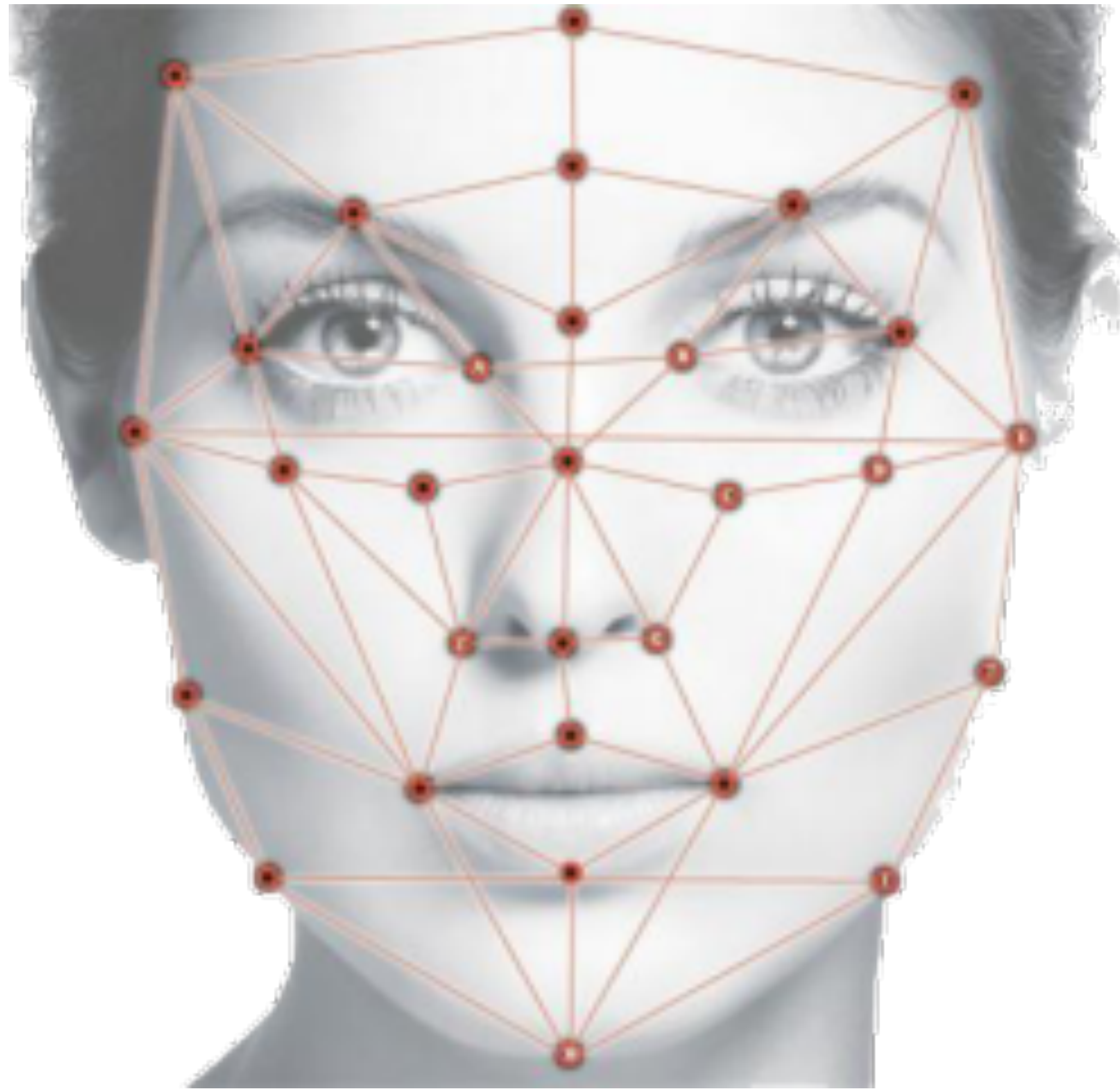




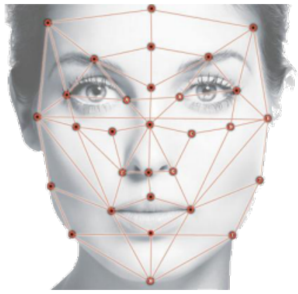
Mixture Matrix Completion

a.k.a. High-Rank Matrix Completion or
Subspace Clustering with Missing Data.

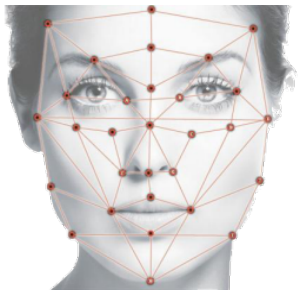

What is this good for?



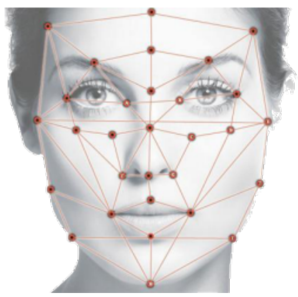
Lots of Applications

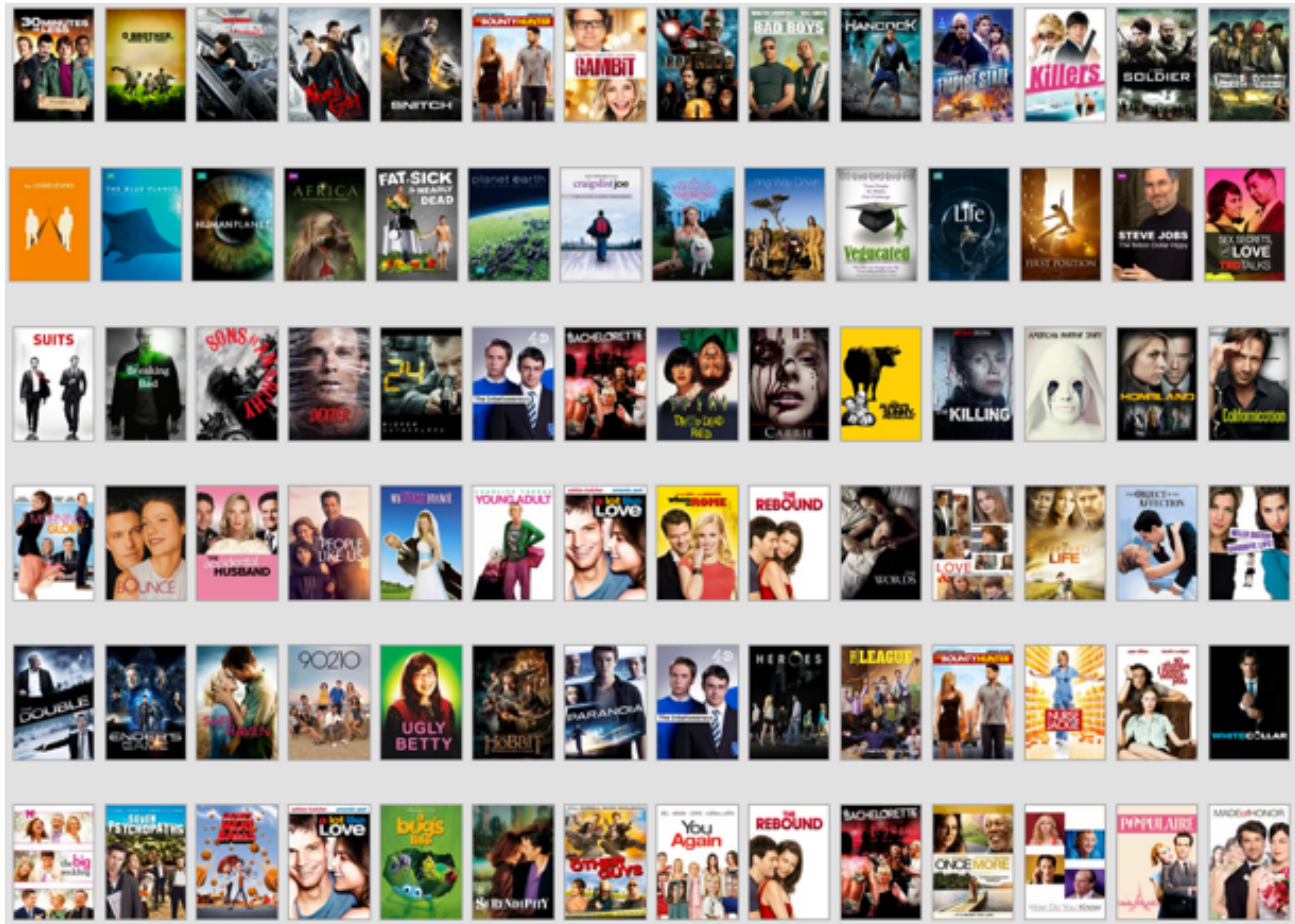


Lots of Applications

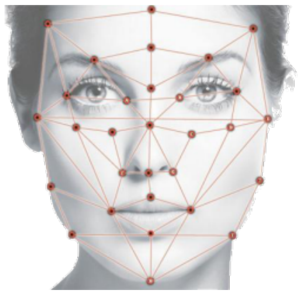


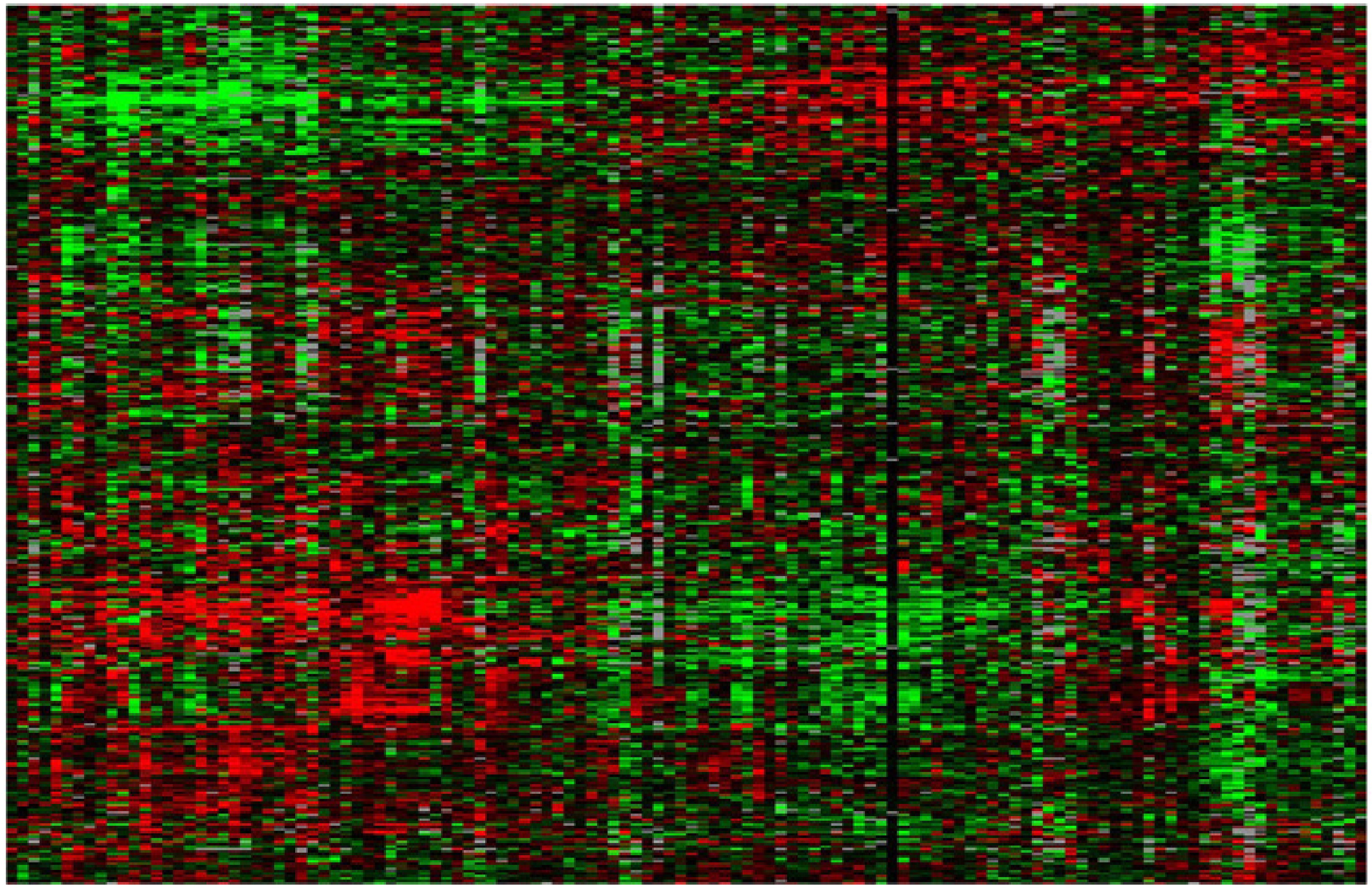
Lots of Applications

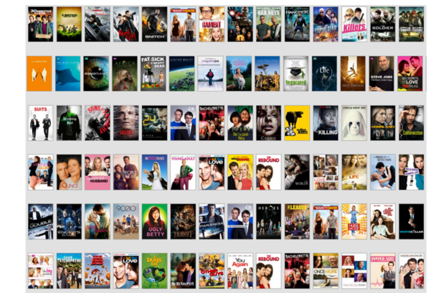




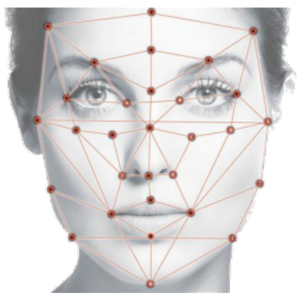
Lots of Applications



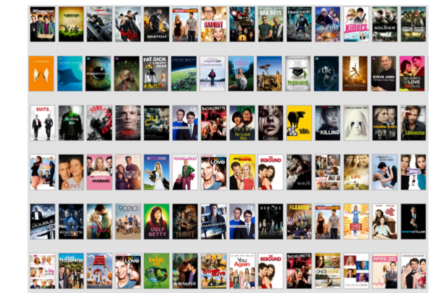
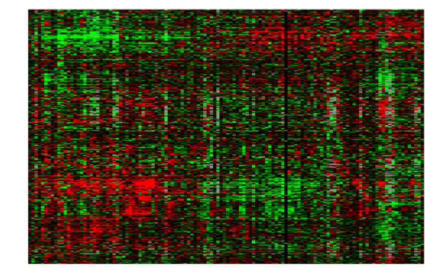




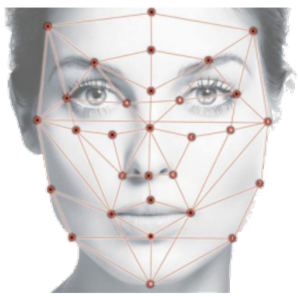
Lots of Applications

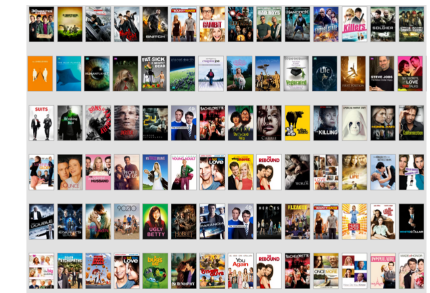


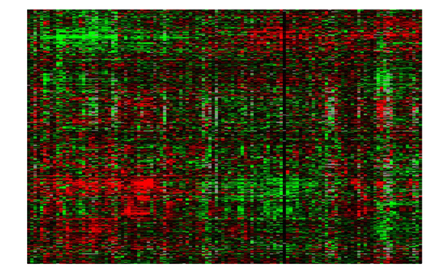


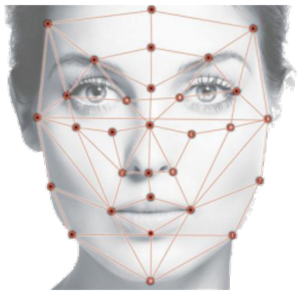
Lots of Applications

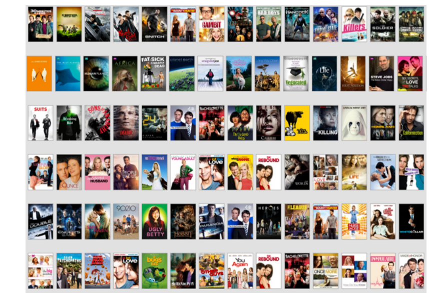


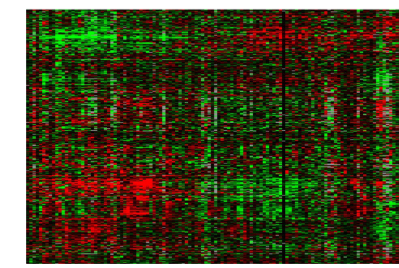




Lots of Applications

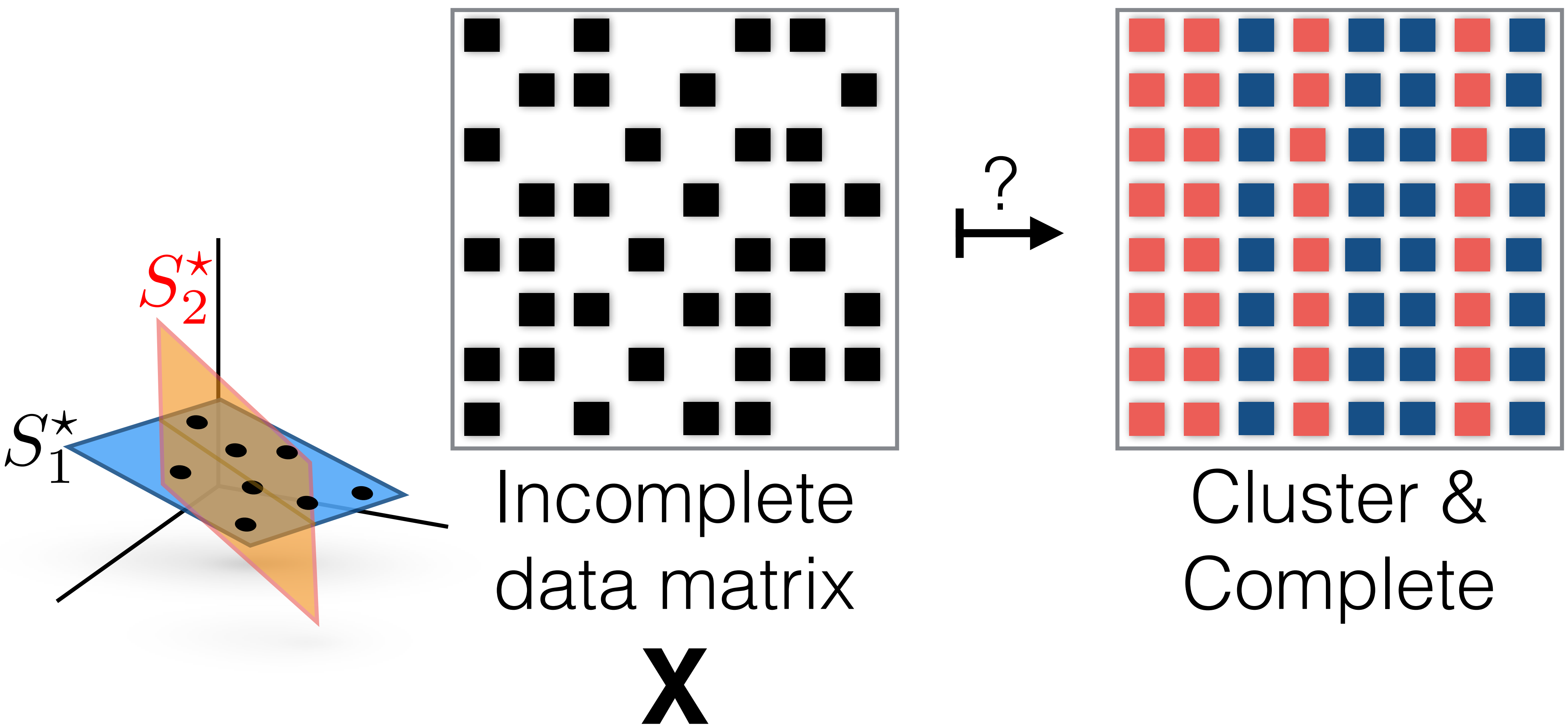




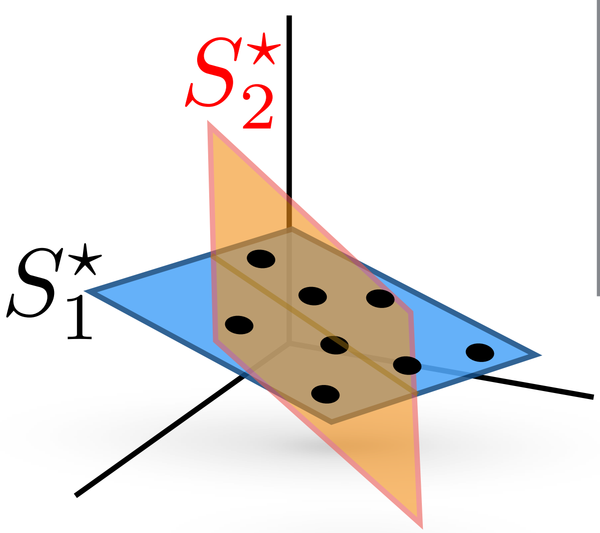


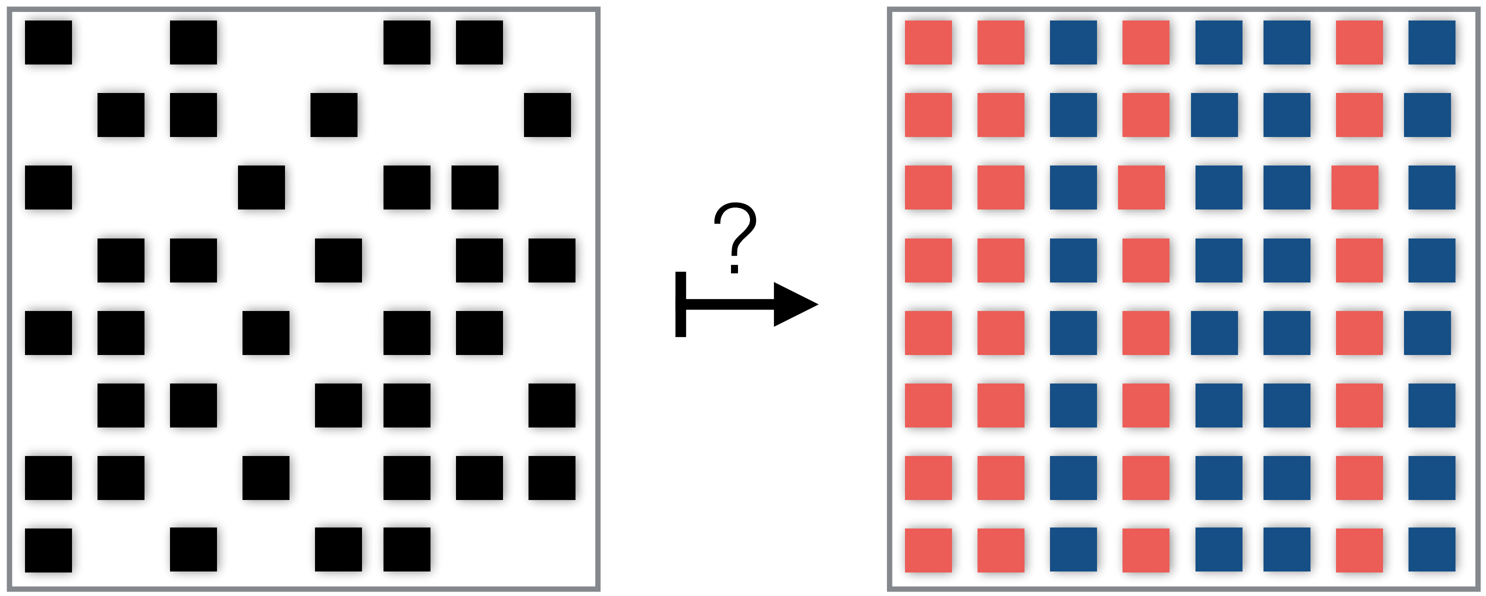
Lots of Applications

In general



In general





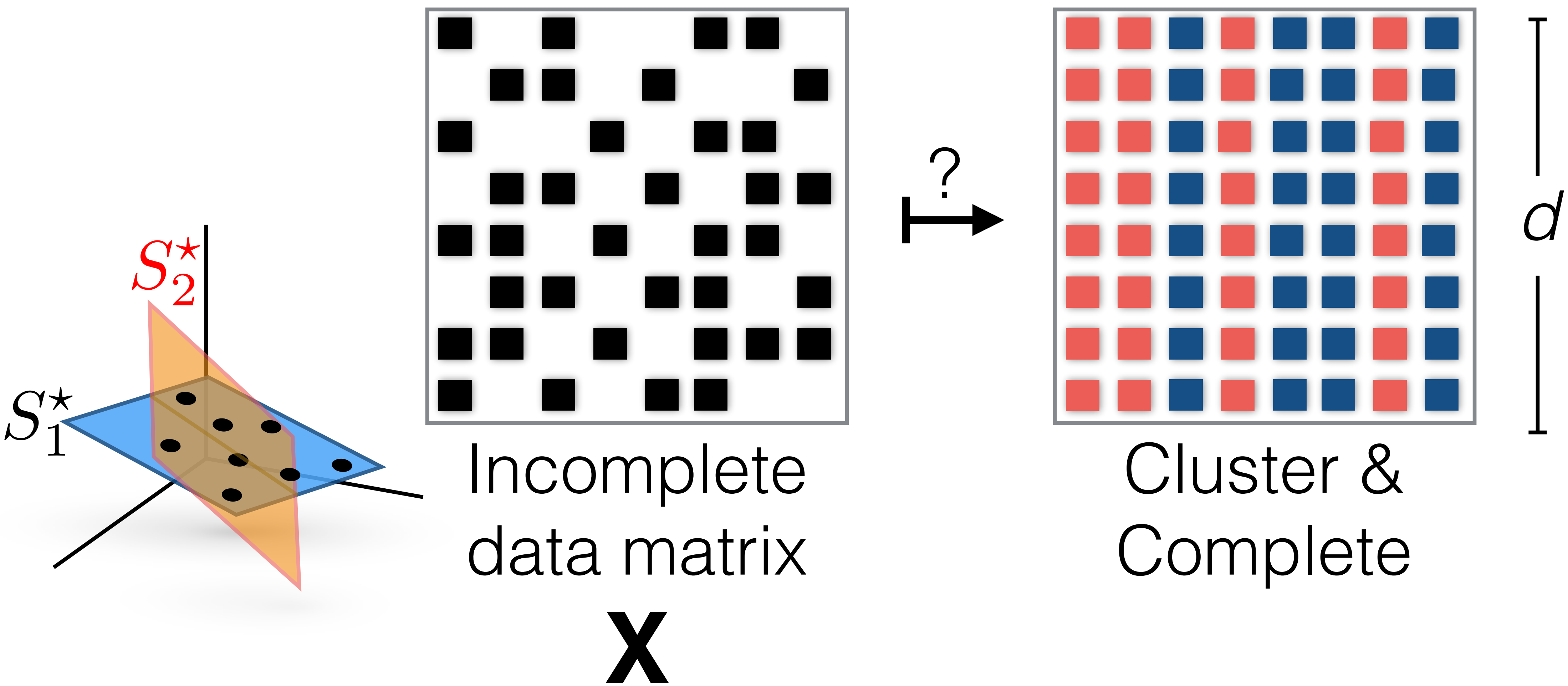
Incomplete data matrix

X

Cluster & Complete

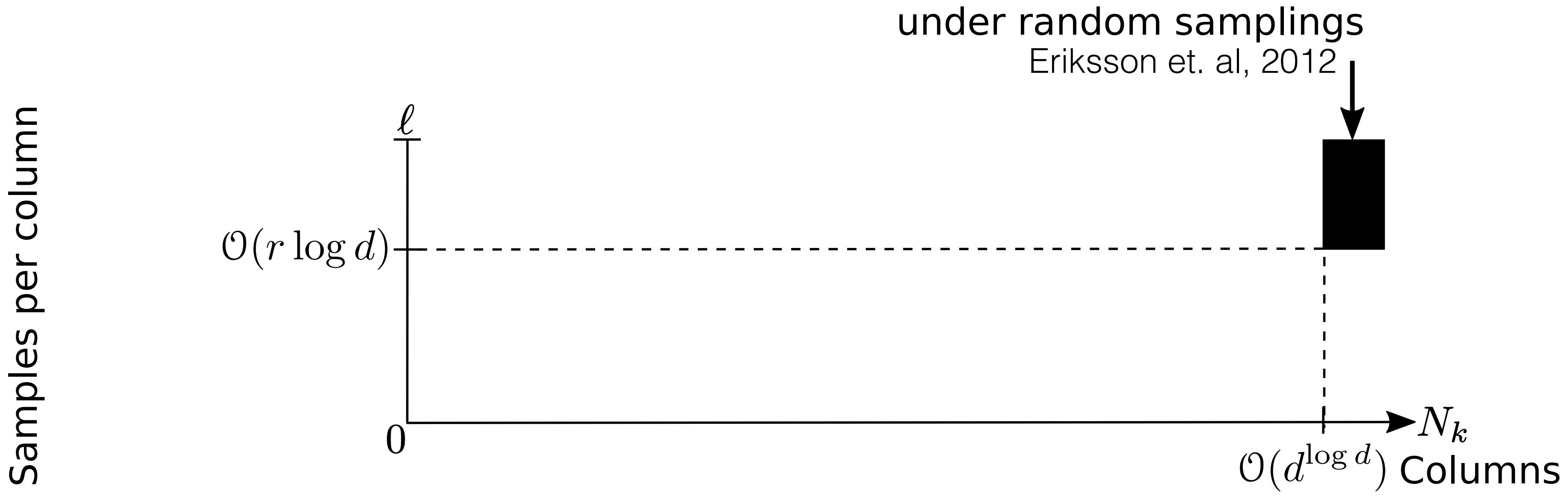
Can this be done? When? How?

In general

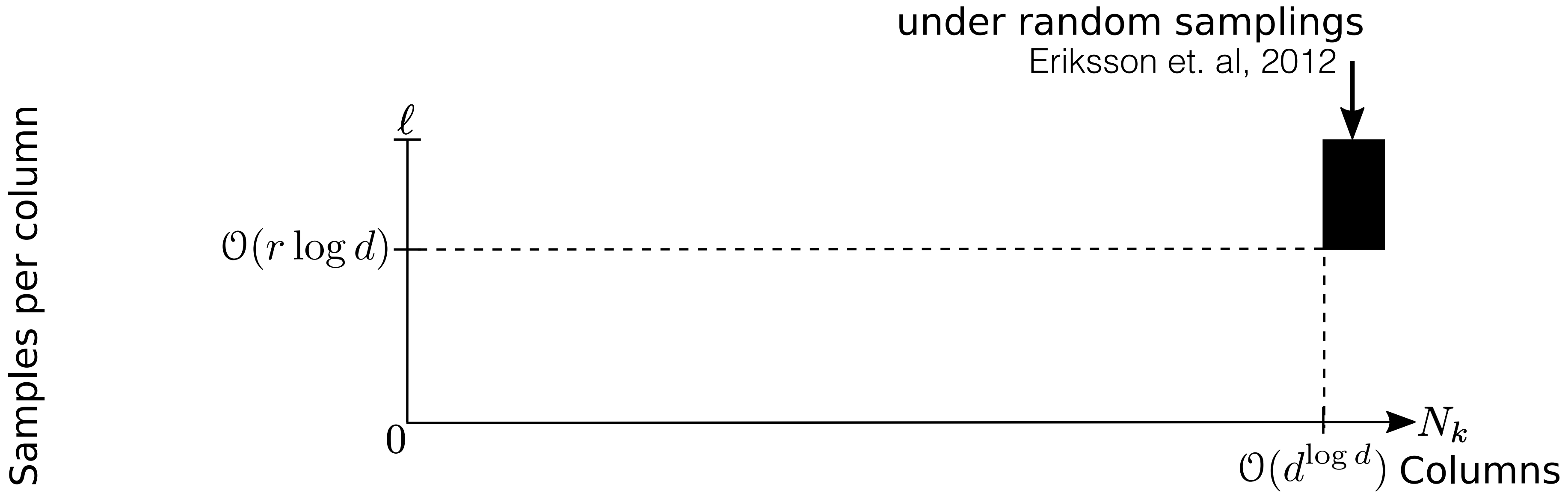


Can this be done? When? How?

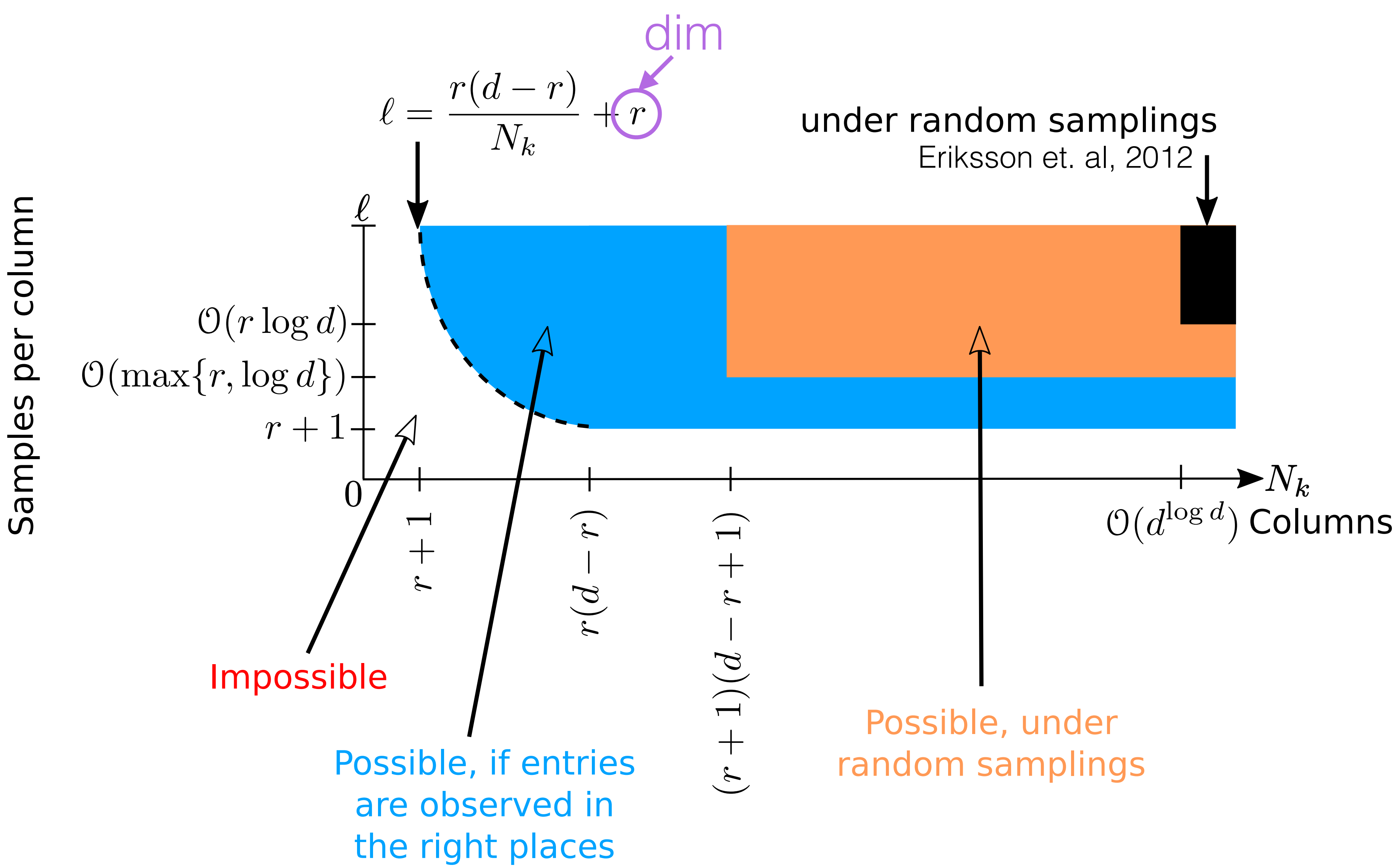
Can this be done? When? How?



Can this be done? When? How?



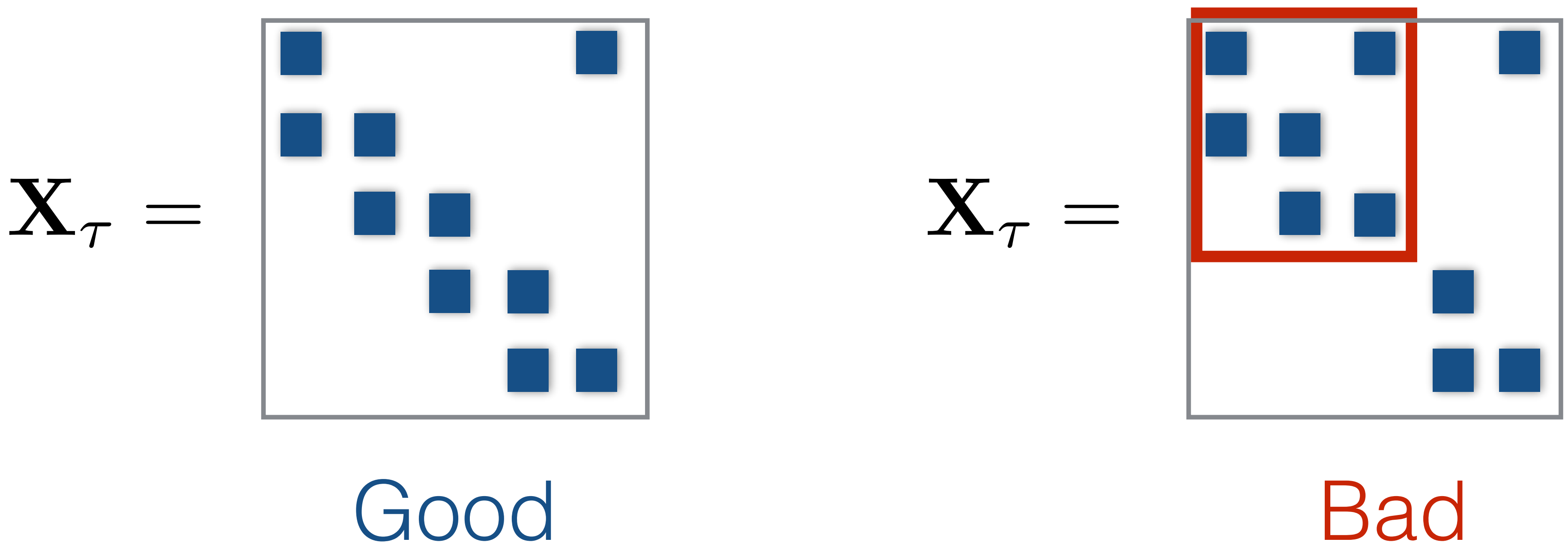
Can this be done? When? How?



Information-theoretic requirements

P.-A., Nowak, 2016

Let $\mathbf{X}_\tau$ be a matrix formed with $d - r + 1$ columns of $\mathbf{X}_k$. We say $\mathbf{X}_\tau$ is observed in the right places if every proper subset of n columns of $\mathbf{X}_\tau$ has observations on at least $n + r$ rows.



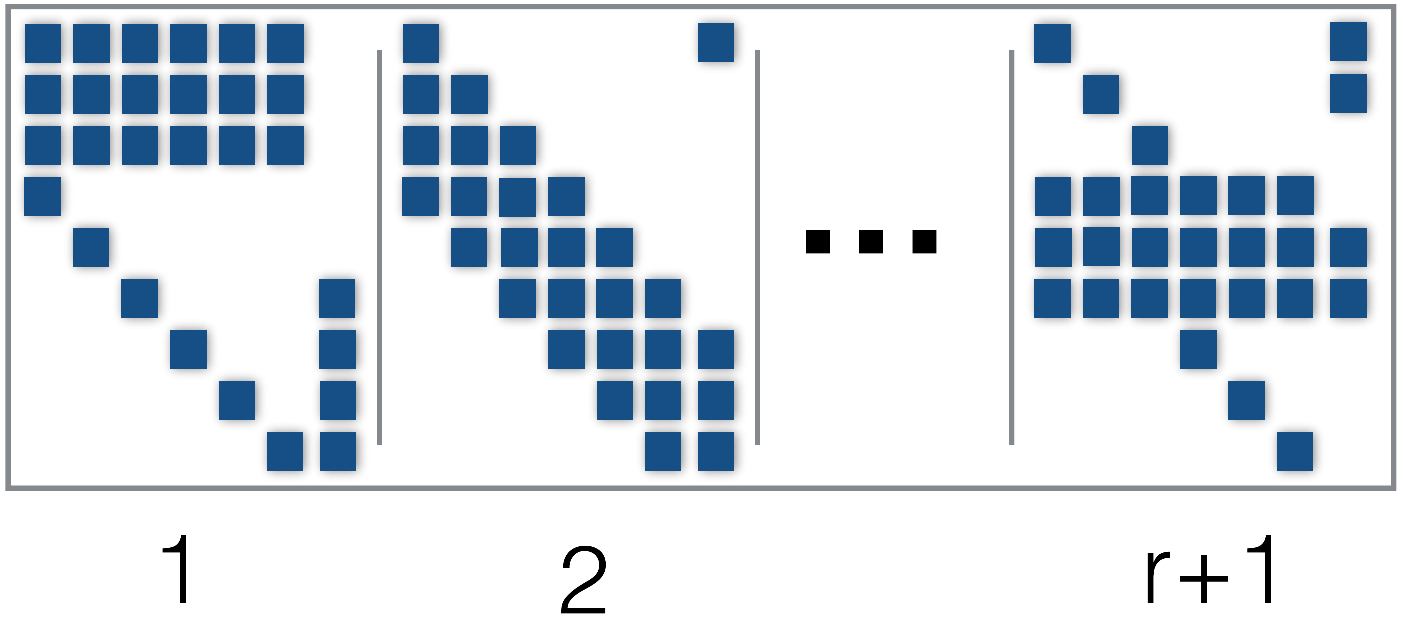
Observed in the right places

P.-A., Nowak, 2016

What do I mean?

For every k, suppose $\mathbf{X}_k$ contains $r+1$ disjoint matrices $\{\mathbf{X}_\tau\}_{\tau=1}^{r+1}$ observed in the right entries. Then $\{S_k^\star\}_{k=1}^{K}$ is the only union of subspaces that agrees with $\mathbf{X}$.

$\mathbf{X}_k =$



Information-theoretic requirements

P.-A., Nowak, 2016

THE FOLLOWING **PREVIEW** HAS BEEN APPROVED FOR

ALL AUDIENCES

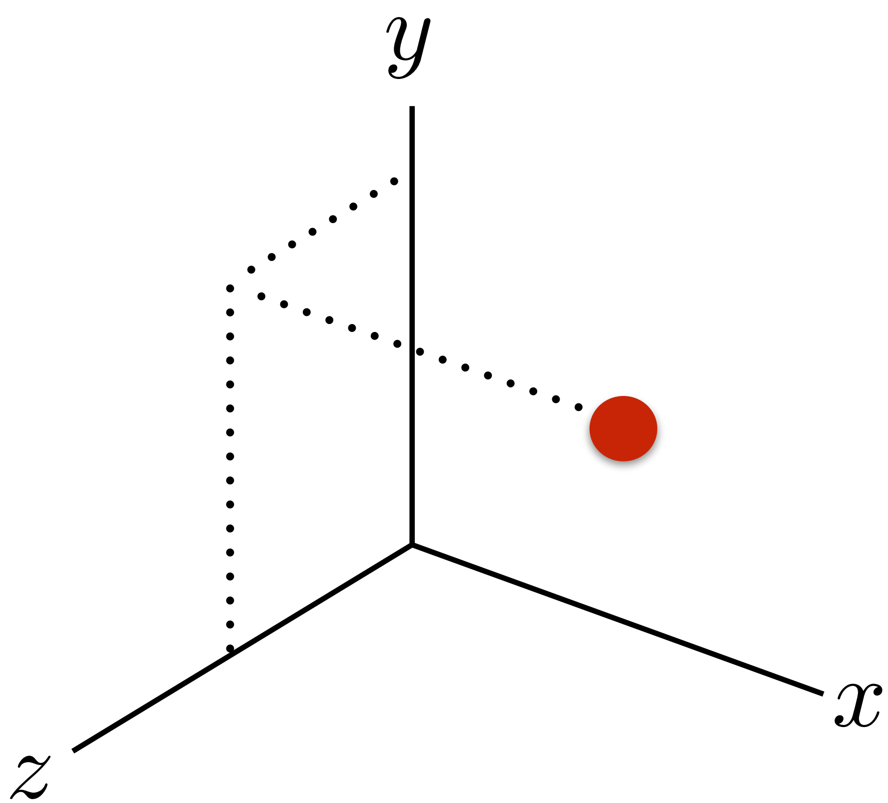
BY THE MOTION PICTURE ASSOCIATION OF AMERICA INC.

THE FILM ADVERTISED HAS BEEN RATED

R	RESTRICTED
	UNDER 17 REQUIRES ACCOMPANYING PARENT OR GUARDIAN
ALGEBRAIC GEOMETRY	

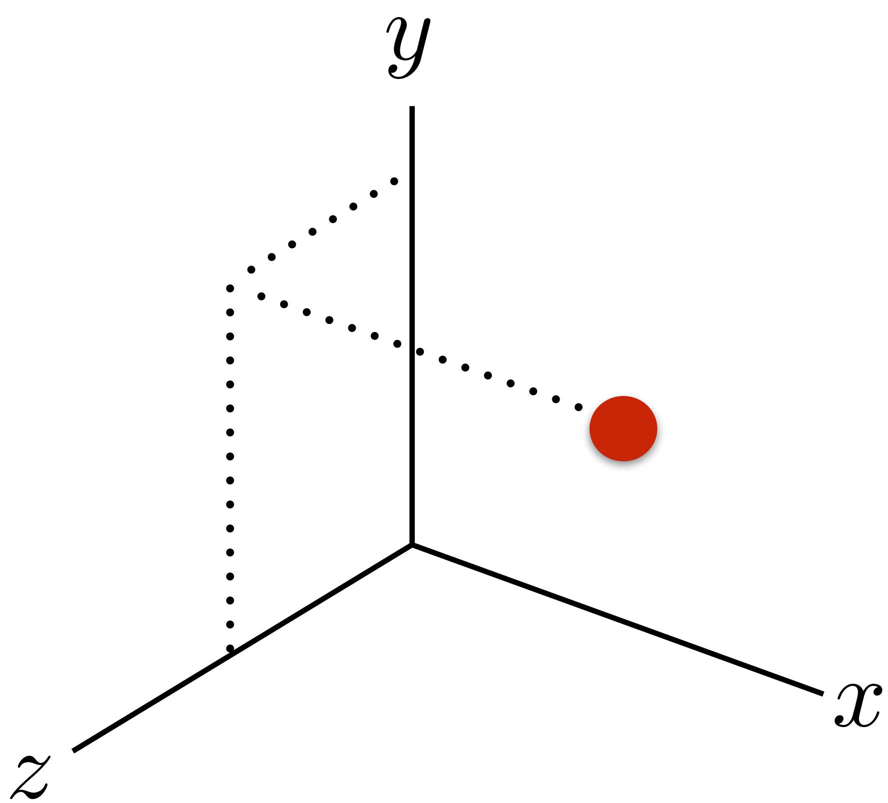
www.filmratings.com

www.mpaa.org



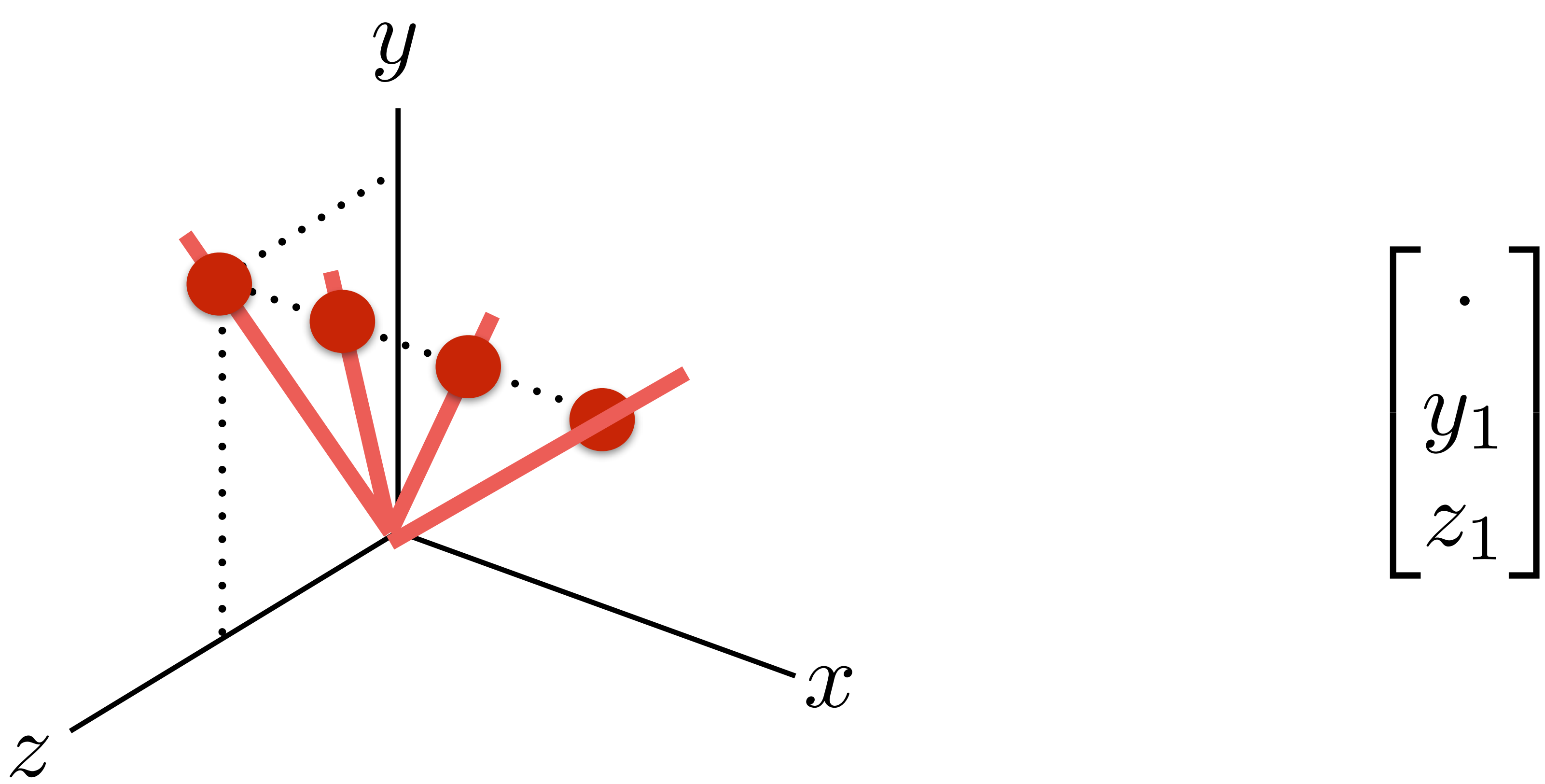
$$\begin{bmatrix} \cdot \\ y_1 \\ z_1 \end{bmatrix}$$

A flavor of our ideas

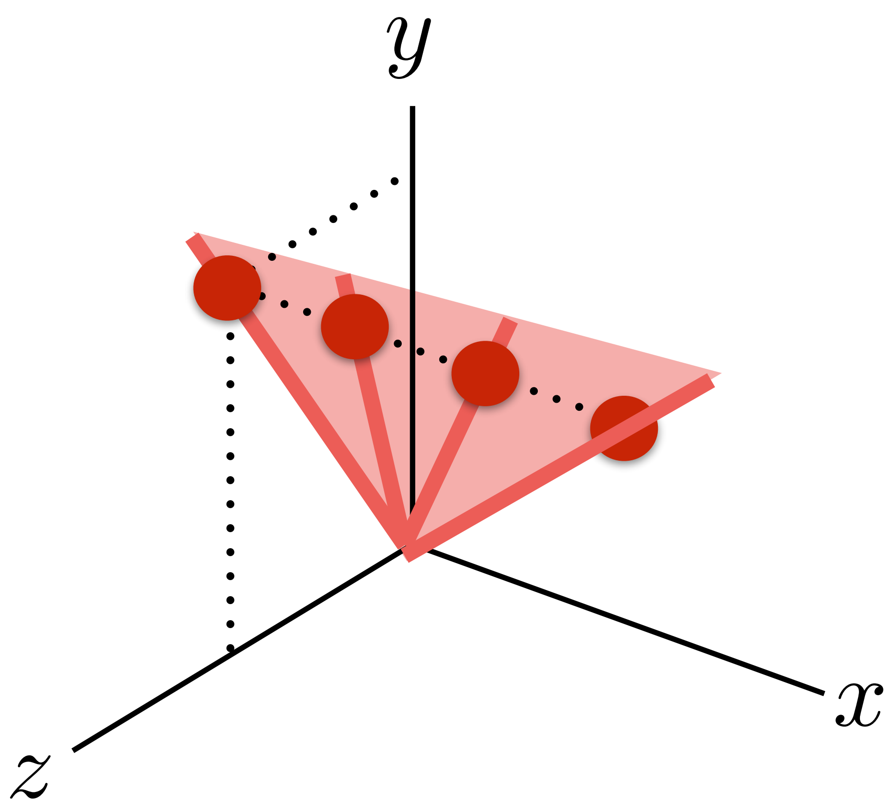


$$\begin{bmatrix} \cdot \\ y_1 \\ z_1 \end{bmatrix}$$

A flavor of our ideas



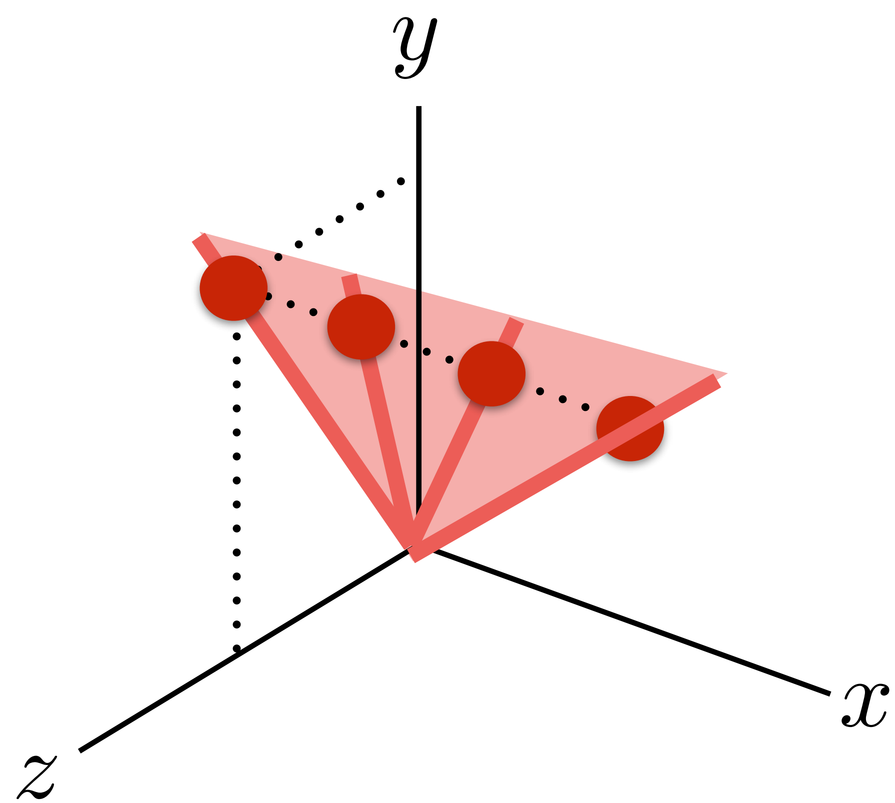
A flavor of our ideas



$$\begin{bmatrix} \cdot \\ y_1 \\ z_1 \end{bmatrix}$$

A flavor of our ideas

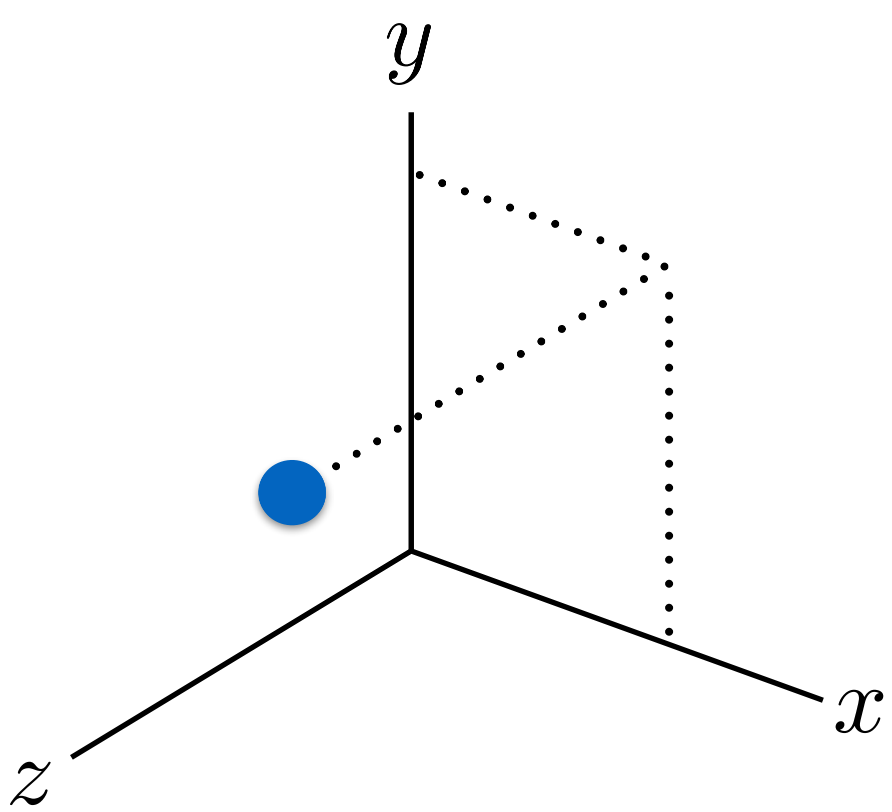
This column imposes *1* restriction on what the subspace may be



$$\begin{bmatrix} \cdot \\ y_1 \\ z_1 \end{bmatrix}$$

A flavor of our ideas

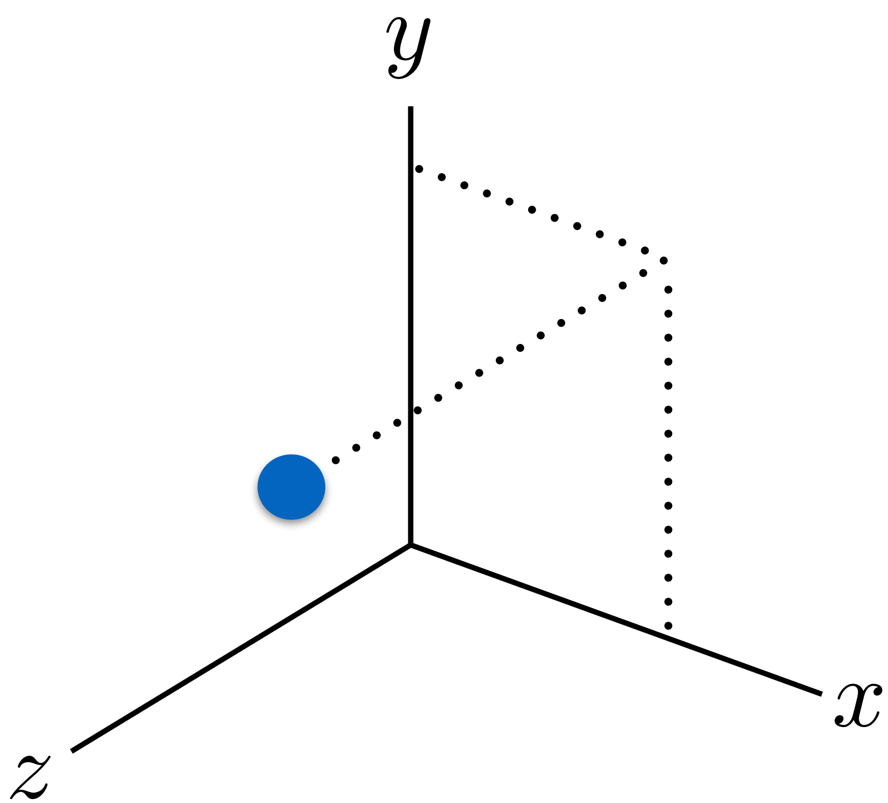
New column imposes *1* more restriction on what the subspace may be



$$\begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

A flavor of our ideas

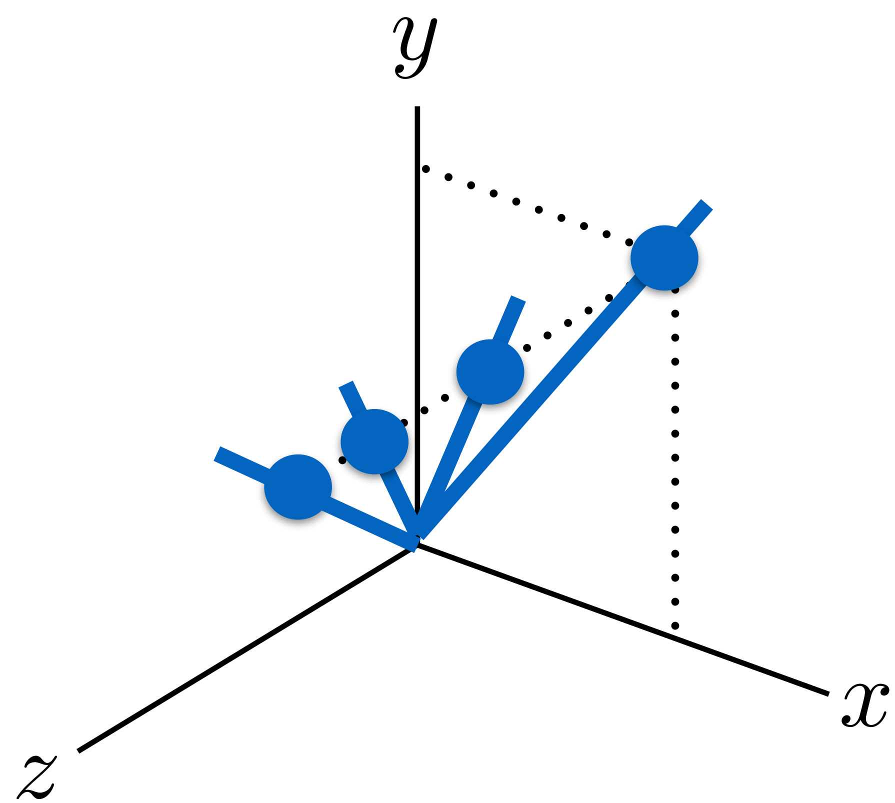
New column imposes 1 more restriction on what the subspace may be



$$\begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

A flavor of our ideas

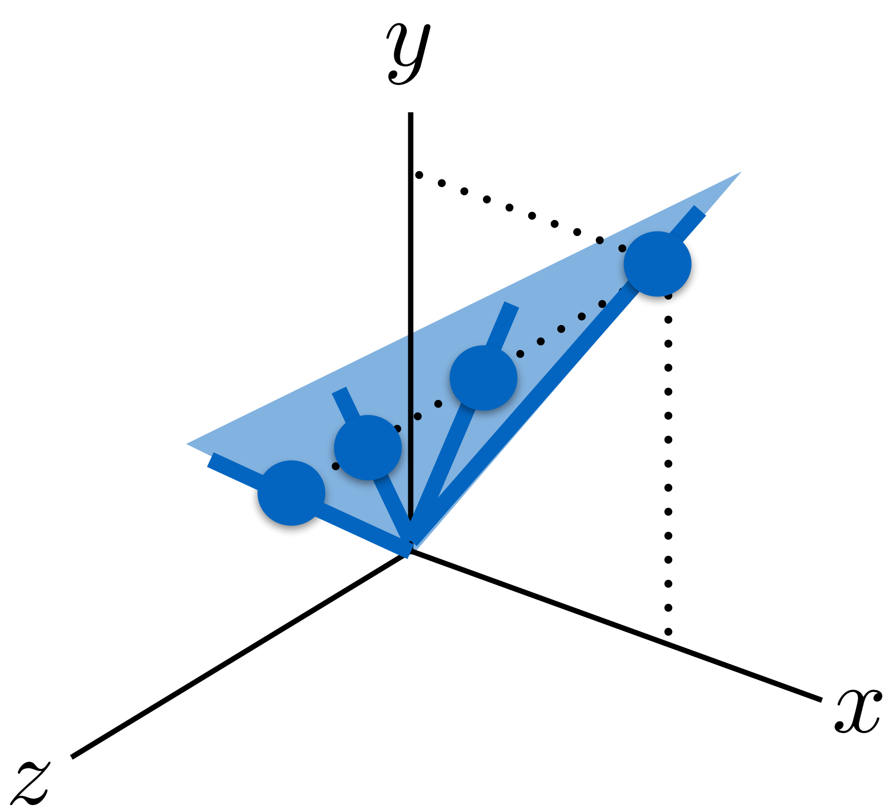
New column imposes 1 more restriction on what the subspace may be



$$\begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

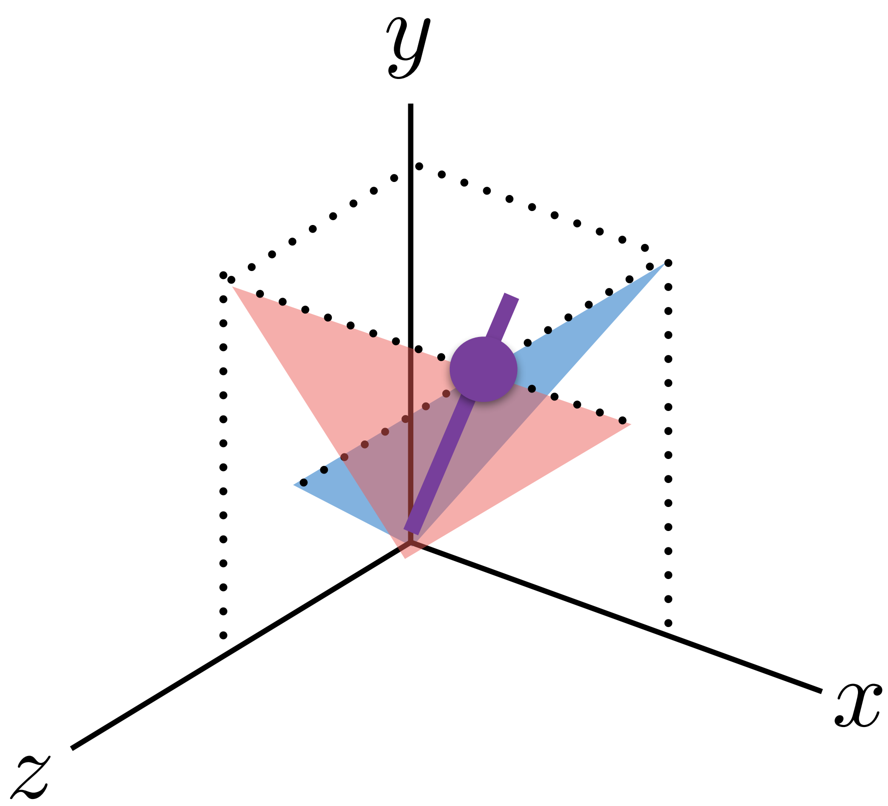
A flavor of our ideas

New column imposes *1* more restriction on what the subspace may be



$$\begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

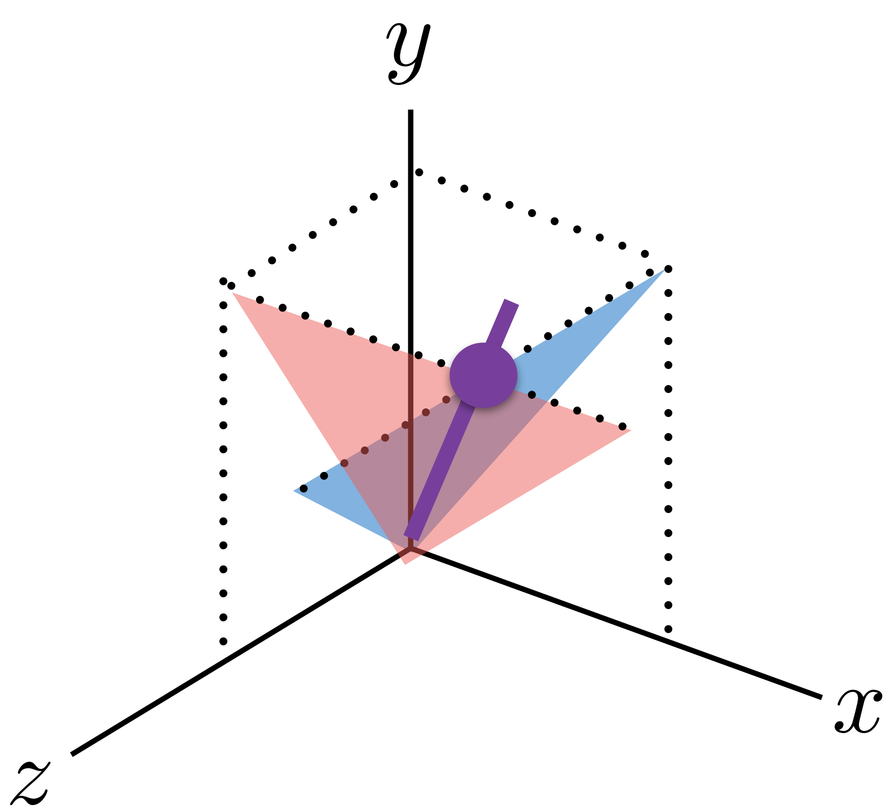
A flavor of our ideas



Each column with *r+1* entries imposes *1* restriction on what the subspace may be

$$\begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

A flavor of our ideas



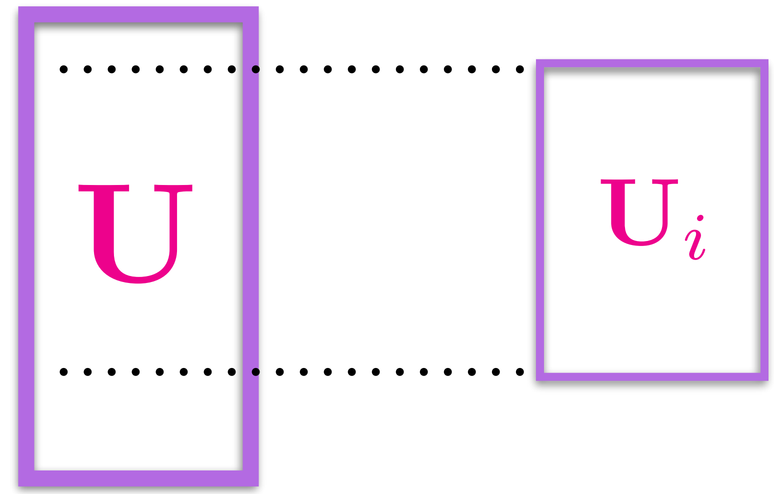
Each column with *r+1* entries imposes *1* restriction on what the subspace may be

Degree-*r* polynomial

$$\begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

A flavor of our ideas

Take a basis of an arbitrary subspace



This subspace agrees with $\mathbf{x}_i$ if and only if

$$\mathbf{x}_i = \mathbf{U}_i \boldsymbol{\theta}_i$$

A flavor of our ideas

- We can split this as:

$$\begin{matrix} r\left\{ \right. \\ 1\left\{ \right. \end{matrix} \left[\begin{array}{c} \boldsymbol{x}_{\boldsymbol{\Delta}_i} \\ \hline \boldsymbol{x}_{\boldsymbol{\nabla}_i} \end{array} \right] = \left[\begin{array}{c} \mathbf{U}_{\boldsymbol{\Delta}_i} \\ \hline \mathbf{U}_{\boldsymbol{\nabla}_i} \end{array} \right] \boldsymbol{\theta}_i.$$

- We can use the top block to solve for $\boldsymbol{\theta}_i$:

$$\boldsymbol{\theta}_i = \mathbf{U}_{\boldsymbol{\Delta}_i}^{-1} \boldsymbol{x}_{\boldsymbol{\Delta}_i}.$$

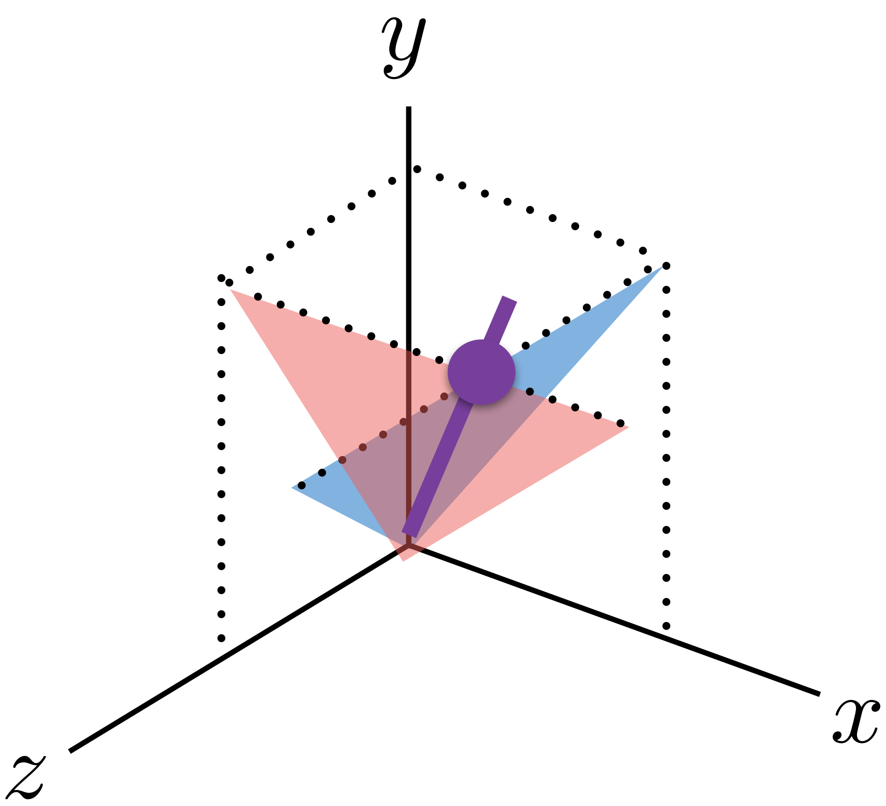
- Plug this in the last row:

$$\boldsymbol{x}_{\boldsymbol{\nabla}_i} = \mathbf{U}_{\boldsymbol{\nabla}_i} \mathbf{U}_{\boldsymbol{\Delta}_i}^{-1} \boldsymbol{x}_{\boldsymbol{\Delta}_i}.$$

- Or equivalently

$$\underbrace{\boldsymbol{x}_{\boldsymbol{\nabla}_i} - \mathbf{U}_{\boldsymbol{\nabla}_i} \mathbf{U}_{\boldsymbol{\Delta}_i}^{-1} \boldsymbol{x}_{\boldsymbol{\Delta}_i}}_{f_i(\mathbf{U}_i | x_i)} = 0.$$

A flavor of our ideas



$$\mathbf{X} = \begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

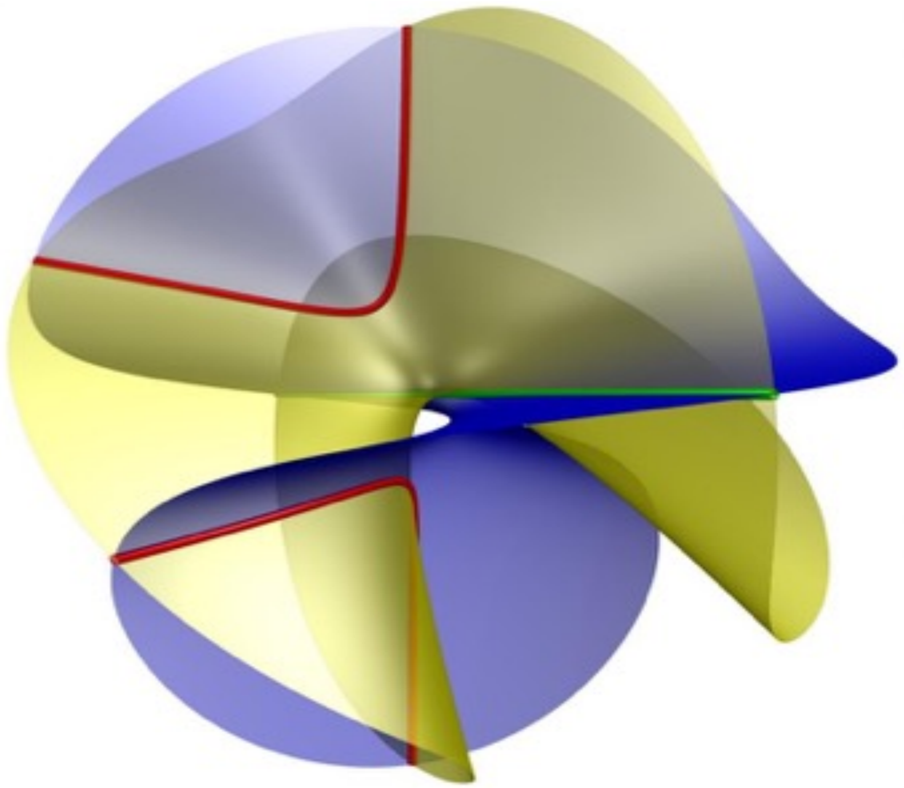
A subspace S agrees with $\mathbf{X}$

$$\Updownarrow$$

$$f_1(\mathbf{U}_1|x_1) = 0$$

$$f_2(\mathbf{U}_2|x_2) = 0$$

A flavor of our ideas



$$\mathbf{X} = \begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

A subspace S agrees with $\mathbf{X}$

$$\Updownarrow$$

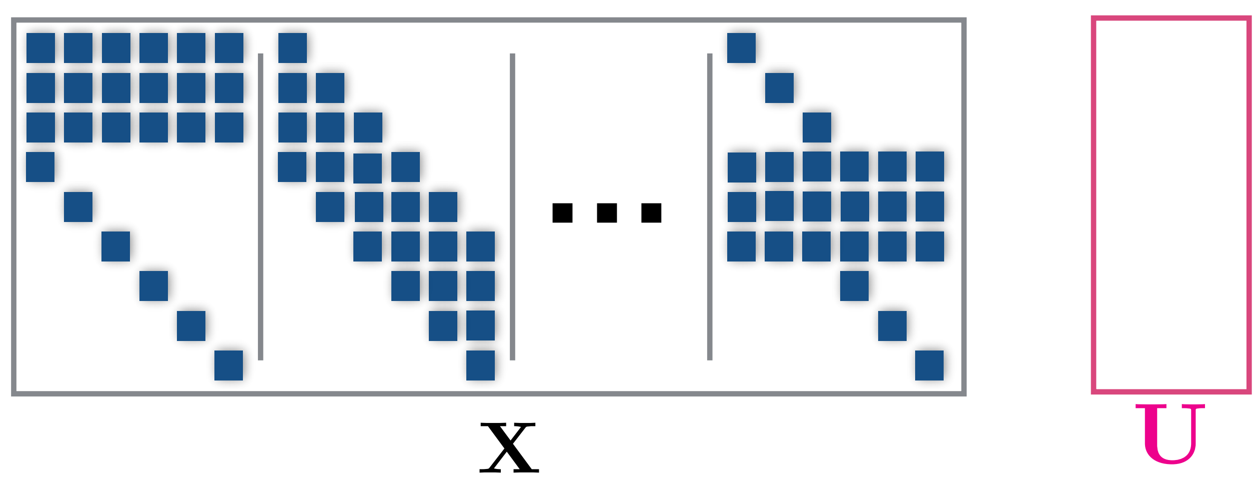
$$f_1(\mathbf{U}_1 | x_1) = 0$$
$$f_2(\mathbf{U}_2 | x_2) = 0$$

A flavor of our ideas

- Each of column produces one polynomial

$$f_1(\mathbf{U}_1|x_1),\ f_2(\mathbf{U}_2|x_2),\ \ldots,\ f_N(\mathbf{U}_N|x_N)$$

- The observed rows indicate the variables involved
- If data is observed in *the right entries*, all variables will be pined down

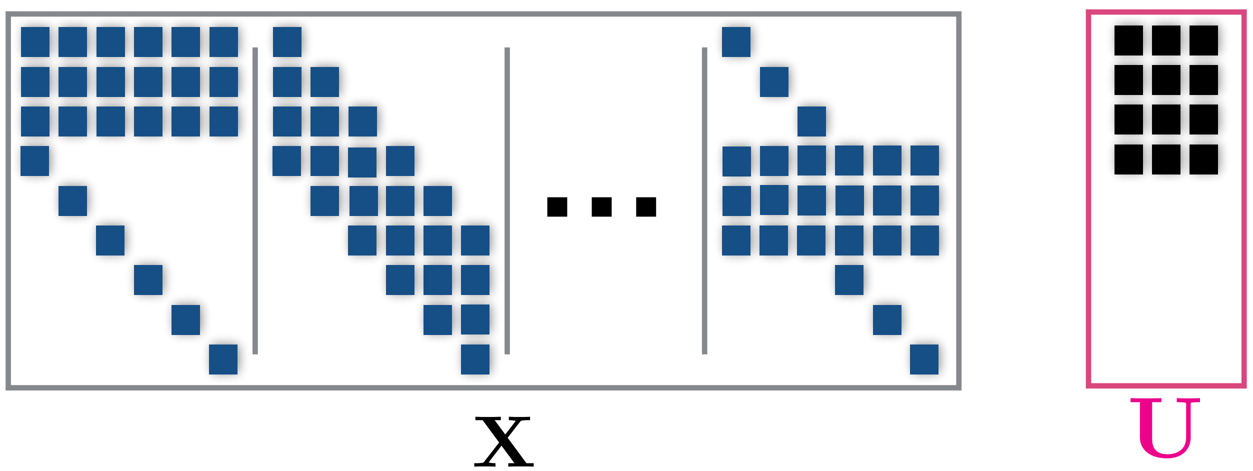


A flavor of our ideas

- Each of column produces one polynomial

$$f_1(\mathbf{U}_1|x_1),\ f_2(\mathbf{U}_2|x_2),\ \ldots,\ f_N(\mathbf{U}_N|x_N)$$

- The observed rows indicate the variables involved
- If data is observed in *the right entries*, all variables will be pined down

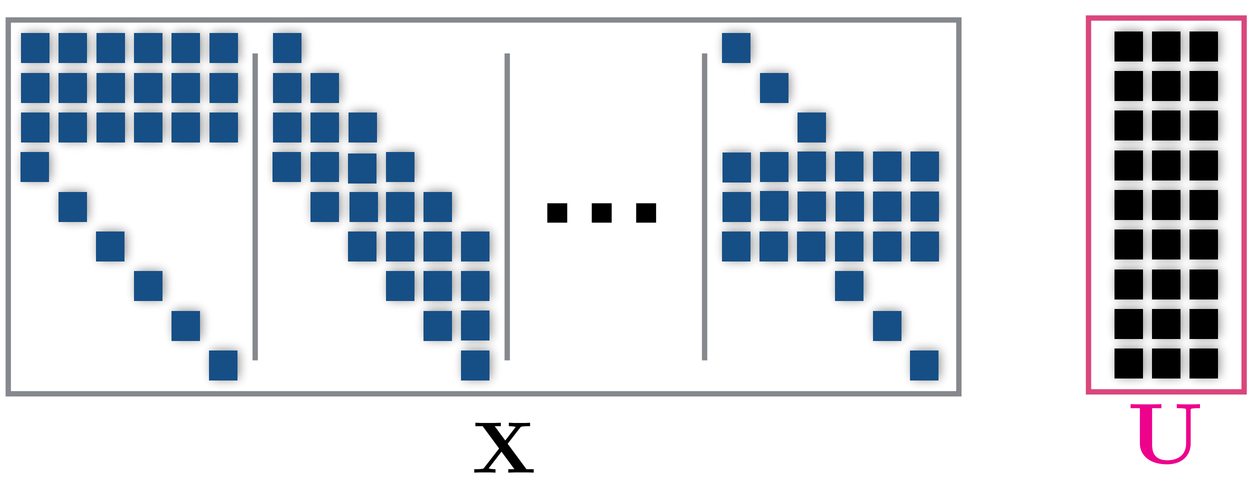


A flavor of our ideas

- Each of column produces one polynomial

$$f_1(\mathbf{U}_1|x_1),\ f_2(\mathbf{U}_2|x_2),\ \ldots,\ f_N(\mathbf{U}_N|x_N)$$

- The observed rows indicate the variables involved
- If data is observed in *the right entries*, all variables will be pined down

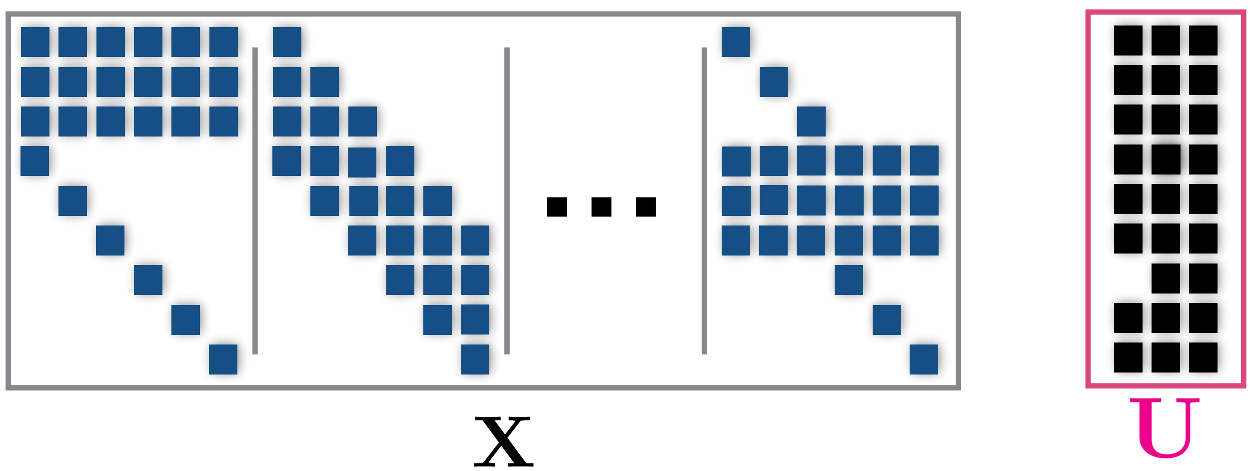


A flavor of our ideas

- Each of column produces one polynomial

$$f_1(\mathbf{U}_1|x_1),\ f_2(\mathbf{U}_2|x_2),\ \ldots,\ f_N(\mathbf{U}_N|x_N)$$

- The observed rows indicate the variables involved
- If data is observed in *the right entries*, all variables will be pined down

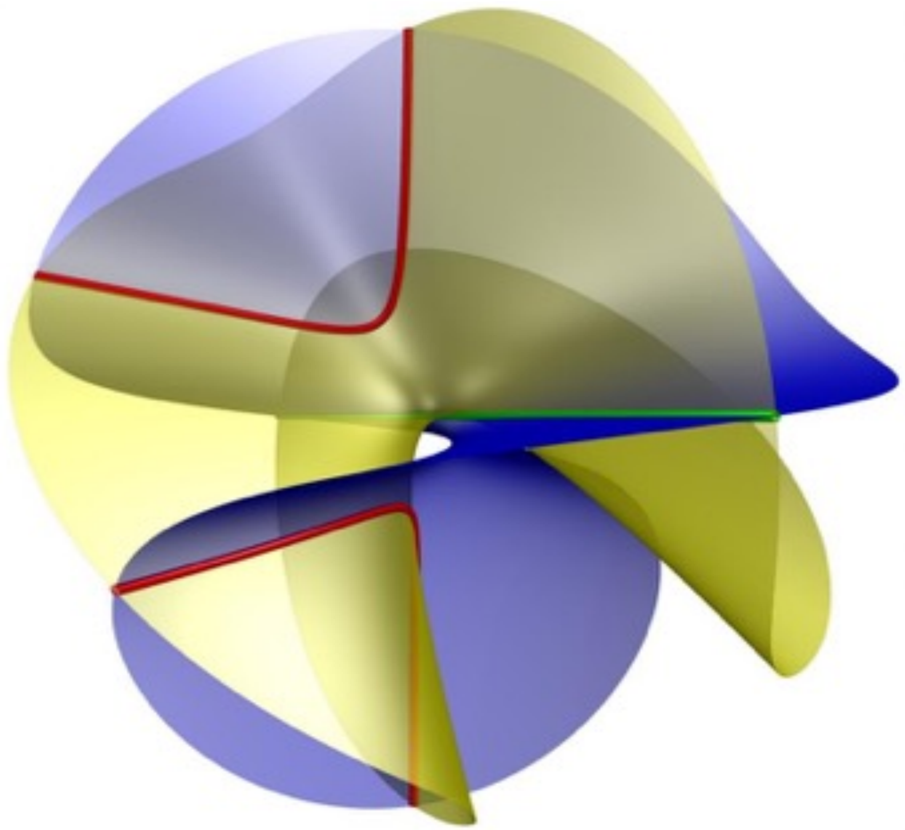


A flavor of our ideas

- If data is observed in *the right entries*
    - Polynomials are algebraically independent

- After this, use cool Algebraic Geometry tricks:
    - Polynomials are a regular sequence
    - Polynomials define a zero-dimensional variety
    - At most finitely many solutions
    - Unique solution (with a bit more work)

A flavor of our ideas



Each column with *r+1* entries imposes *1* restriction on what the subspace may be

Degree-*r* polynomial

$$\begin{bmatrix} \cdot & x_2 \\ y_1 & y_2 \\ z_1 & \cdot \end{bmatrix}$$

Entries observed in *the right places.*

↓

Polynomials are independent.

↓

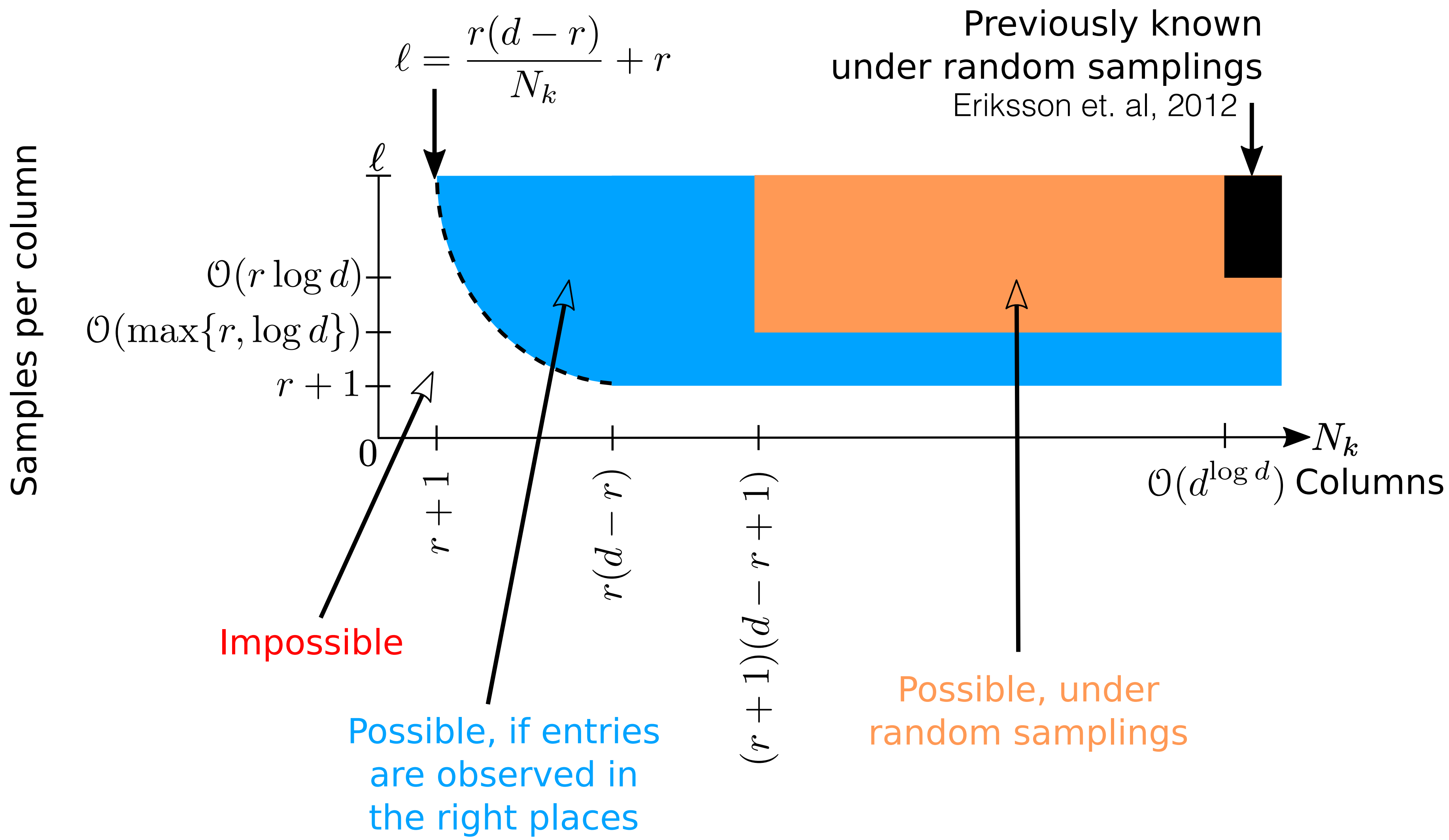
Completion is solution to polynomials.

A flavor of our ideas

Summary


WOW, AMAZING
PLEASE TELL
ME MORE

Can this be done? When? How?



Information-theoretic requirements

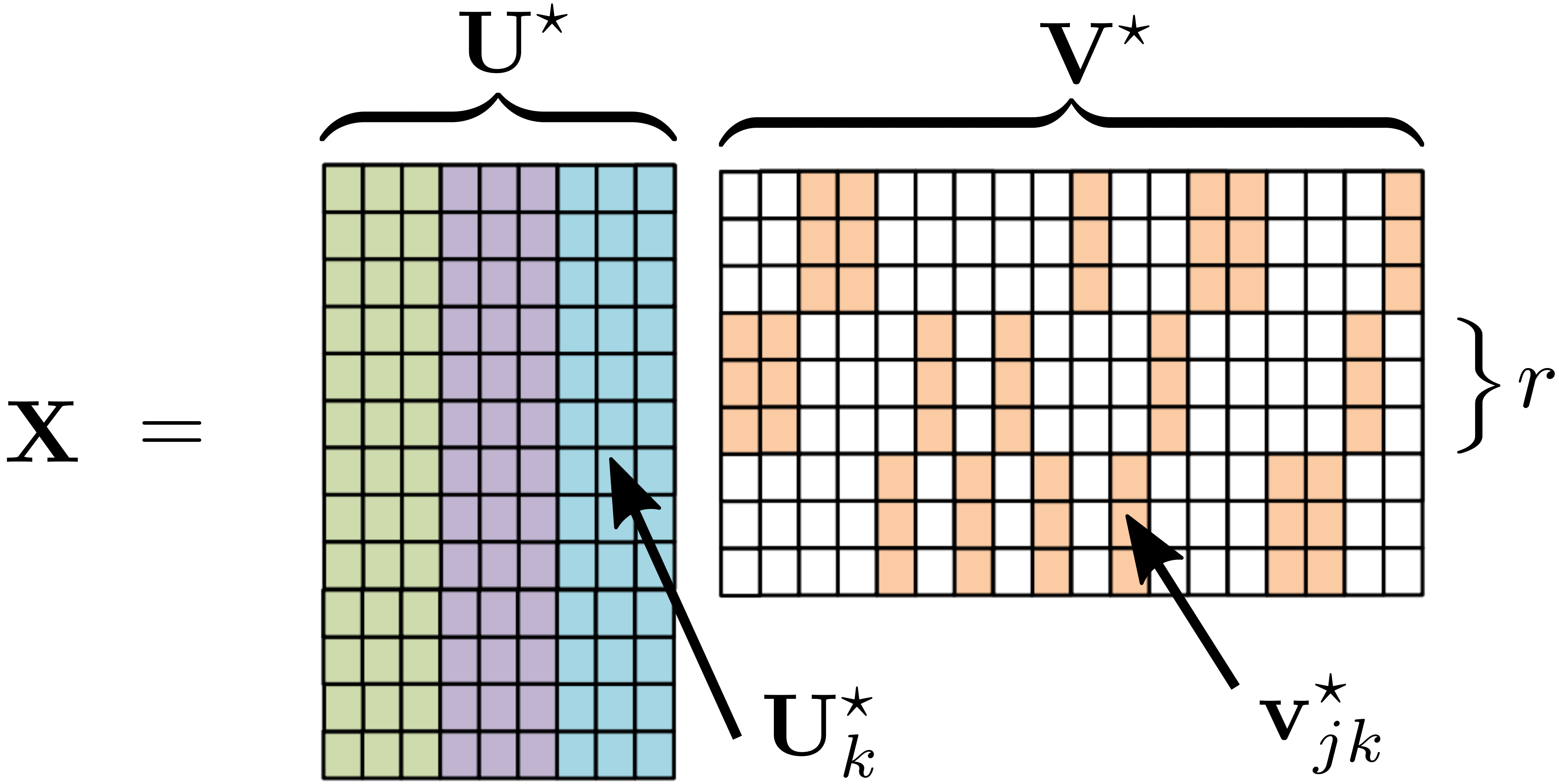
P.-A., Nowak, 2016

Can this be done? When? How?

Samples \ Computation	Good	Bad
Good	• EM [P.-A. Balzano, Nowak, 2014] • GSSC [P.-A. et. al, 2016] • MSC [P.-A. al, 2016] • SSC-EWZF [Wang et. al, 2016] • K-GROUSE [Balzano et. al, 2016]	Polynomials [P.-A. et. al, 2016]
Bad	HRMC [Eriksson et. al, 2012]	Who Cares

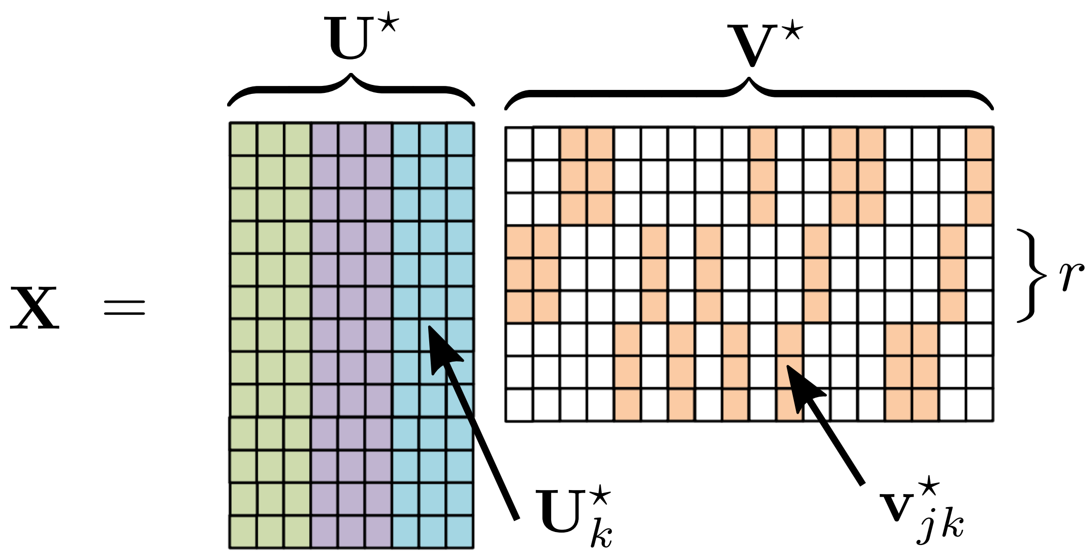
Can this be done? When? How?

Samples \ Computation	Good	Bad
Good	• EM [P.-A., Balzano, Nowak, 2014] • **GSSC [P.-A. et. al, 2016]** • MSC [P.-A. et. al, 2016] • SSC-EWZF [Wang et. al, 2016] • K-GROUSE [Balzano et. al, 2016]	Polynomials [P.-A. et. al, 2016]
Bad	HRMC [Eriksson et. al, 2012]	Who Cares



State-of-the-art Algorithms

P.-A. et. al, 2016



Algorithm 1: Group-Sparse Subspace Clustering

Input: $\mathbf{X}_\mathbf{\Omega}, K, r, \lambda$.

Initialize $\widehat{\mathbf{U}} \in \mathbb{R}^{d \times Kr}$ (e.g., using SSC-EWZF).

repeat

$$\widehat{\mathbf{V}} = \arg\min_{\mathbf{V}} \|\mathbf{\Omega}(\mathbf{X} - \widehat{\mathbf{U}}\mathbf{V})\|_F^2 + \lambda \sum_{j,k=1}^{N,K} \|\mathbf{v}_{jk}\|_2.$$

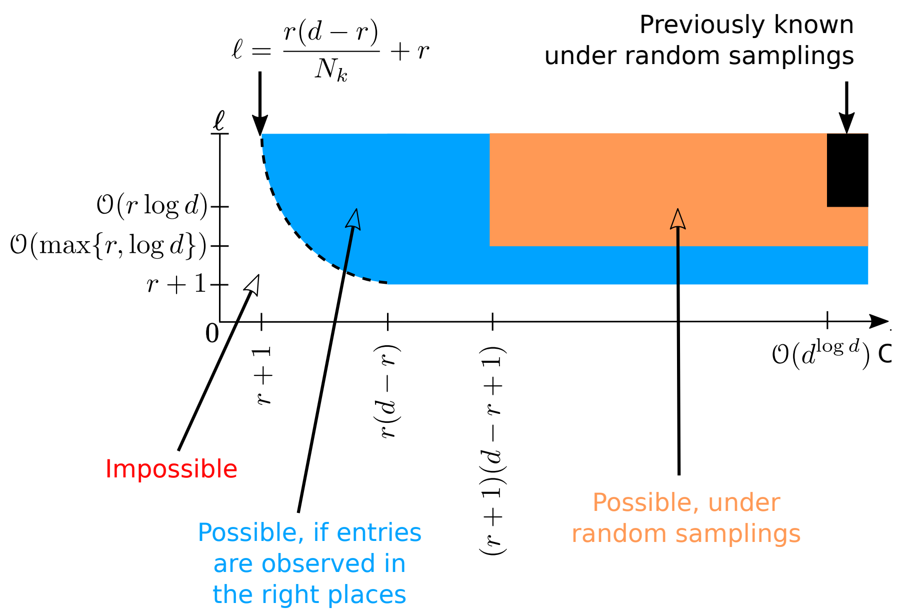
$$\widehat{\mathbf{U}} = \arg\min_{\mathbf{U} \,:\, \|\mathbf{U}\|_F \leq 1} \|\mathbf{\Omega}(\mathbf{X} - \mathbf{U}\widehat{\mathbf{V}})\|_F.$$

until convergence;

Output: $\widehat{\mathbf{U}}, \widehat{\mathbf{V}}$.

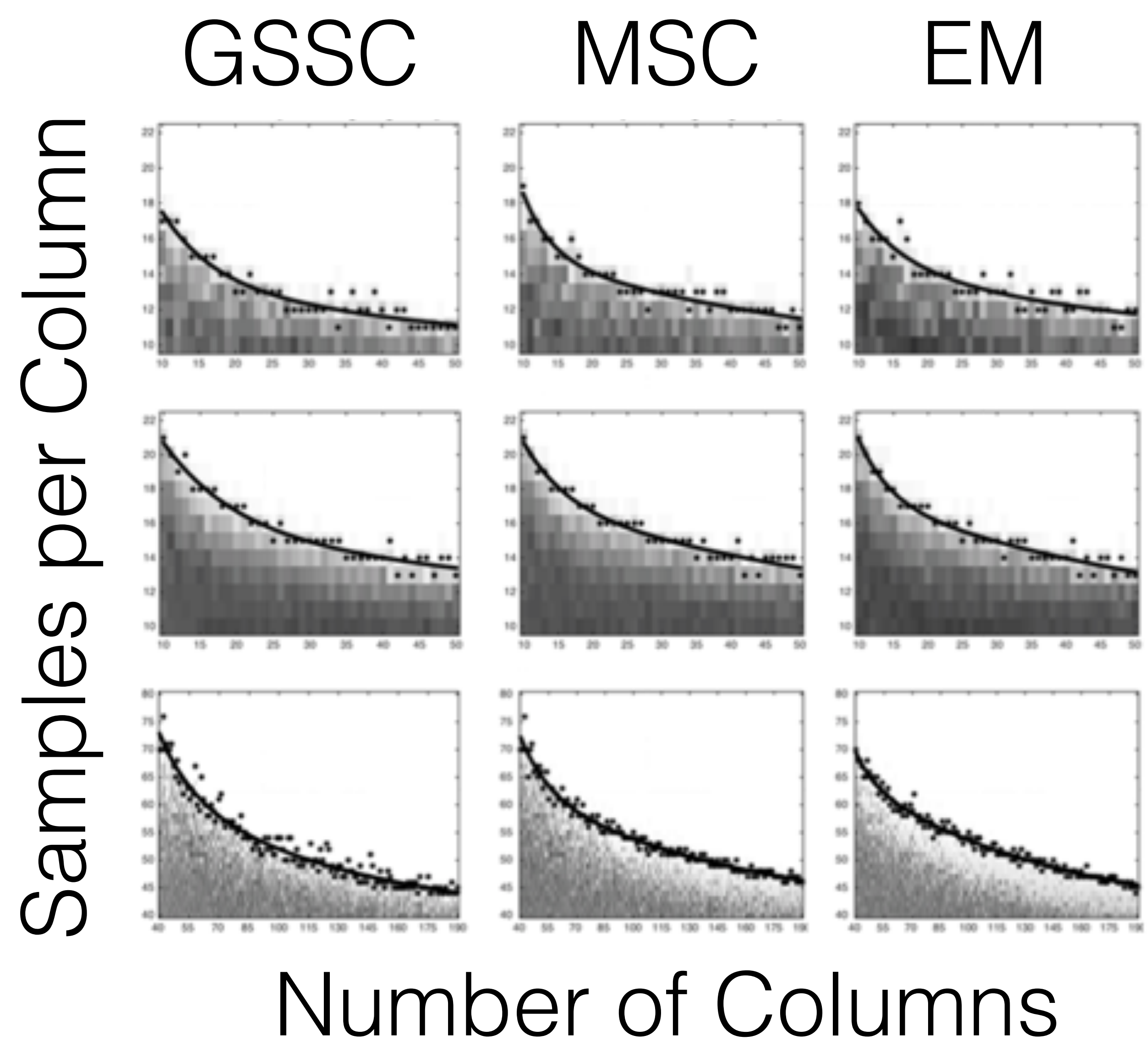
State-of-the-art Algorithms

P.-A. et. al, 2016



P.-A., Nowak, 2016

P.-A. et. al, 2016



Theory matches Practice

Can this be done? When? How?

Samples \ Computation	Good	Bad
Good	• EM [14] P.-A. et. al,, 2014 • GSSC [7] P.-A. et. al, 2016 • MSC [7] P.-A. et. al, 2016 • SSC-EWZF Wang et. al, 2016 • K-GROUSE Balzano et. al, 2016	Polynomials [10] P.-A. et. al, 2016
Bad	HRMC Eriksson et. al, 2012	Who Cares

Can this be done? When? How?

Samples \ Computation	**Good**	**Bad**
Good	• EM [14] P.-A. et. al,, 2014 • GSSC [7] P.-A. et. al, 2016 • MSC [7] P.-A. et. al, 2016 • SSC-EWZF Wang et. al, 2016 • K-GROUSE Balzano et. al, 2016 **Provable?**	Polynomials [10] P.-A. et. al, 2016
Bad	HRMC Eriksson et. al, 2012	Who Cares


Rob Nowak


Nigel Boston


Laura Balzano

Joint work with:


Becca Willett


Roummel Marcia


Steve Wright

Thank you!