A Perturbation Bound on the Subspace Estimator from Canonical Projections

Karan Srivastava
Daniel Pimentel-Alarcón

University of Wisconsin-Madison

ISIT 2022

Outline

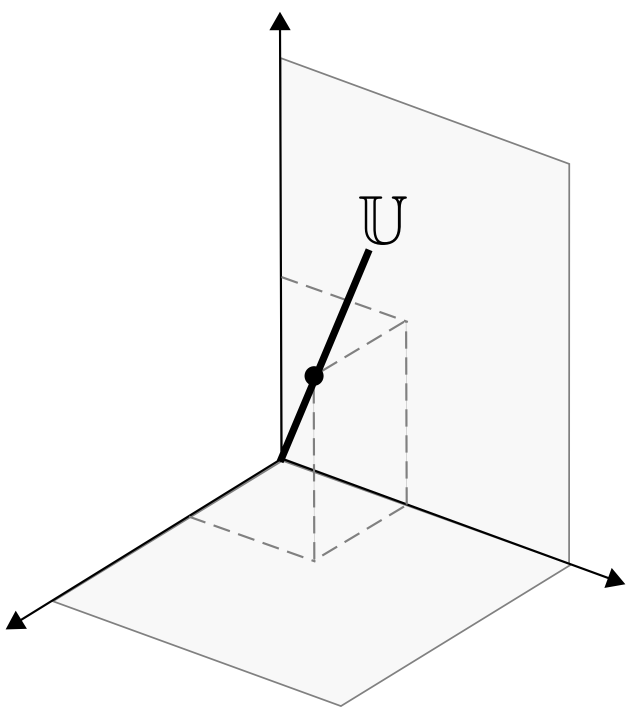
1. Problem Setup - Subspace Estimation
2. Motivation - Missing Data
3. Previous Work: Noiseless case
4. This Paper - Noisy Data and Estimation Bound
5. Applications
6. Conclusions

Outline

1. Problem Setup - Subspace Estimation
2. Motivation - Missing Data
3. Previous Work: Noiseless case
4. This Paper - Noisy Data and Estimation Bound
5. Applications
6. Conclusions

Problem Description

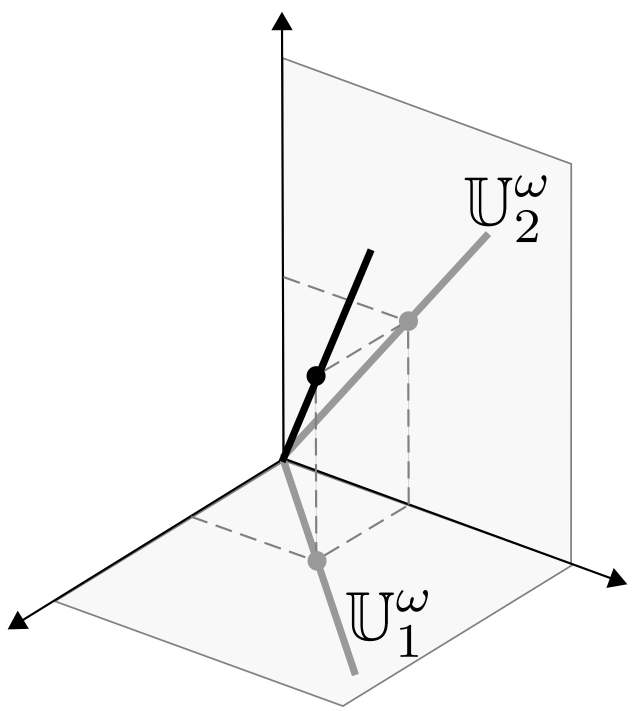
$\mathbb{U} :=$ $r-$ dimensional subspace of $\mathbb{R}^d, r < d$

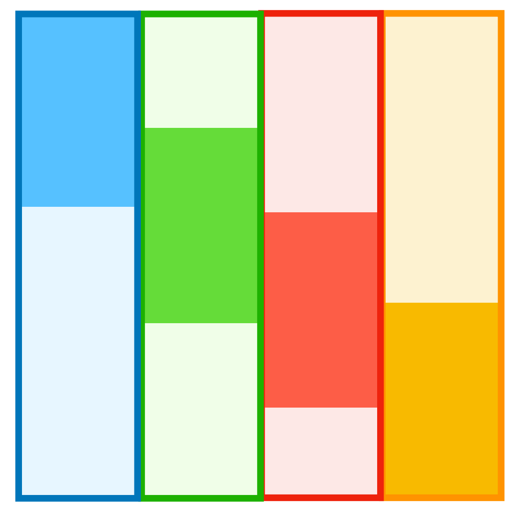


Problem Description

$\mathbb{U}$:= $r-$ dimensional subspace of $\mathbb{R}^d, r < d$

$\mathbb{U}_i^\omega$:= projections of the subspace onto canonical coordinates.



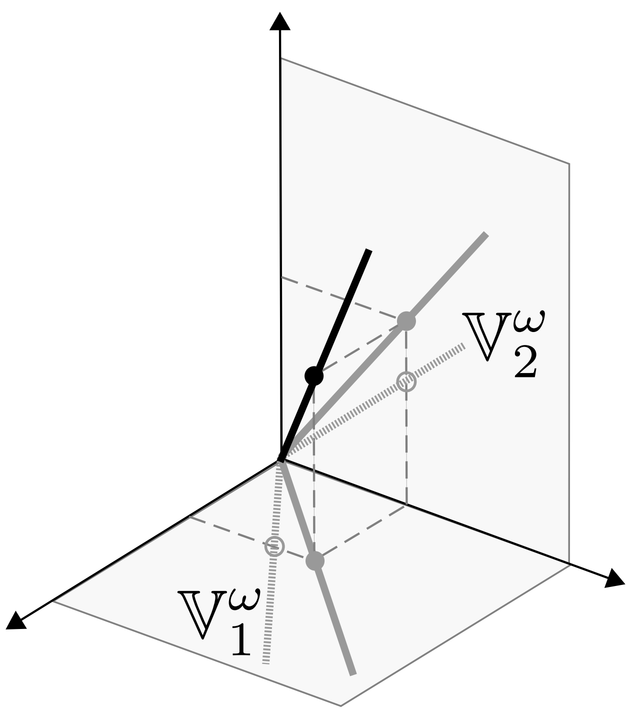


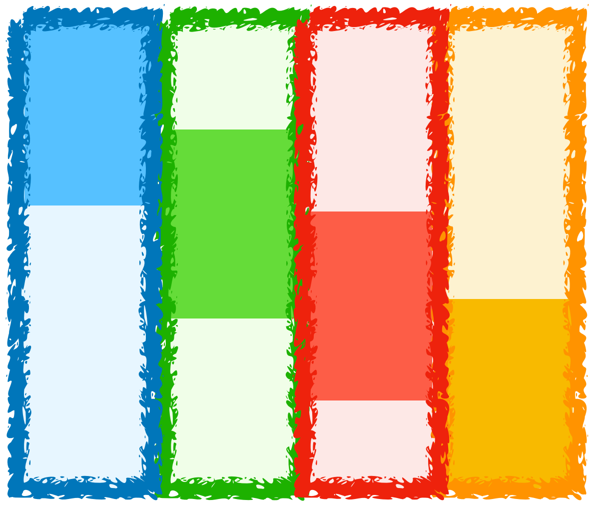
Problem Description

$\mathbb{U} :=$ $r-$ dimensional subspace of $\mathbb{R}^d, r < d$

$\mathbb{U}_i^\omega :=$ projections of the subspace onto canonical coordinates.

$\mathbb{V}_i^\omega := \mathbb{U}_i^\omega + Z_i^\omega$, where Z_i^ω is a noise matrix.



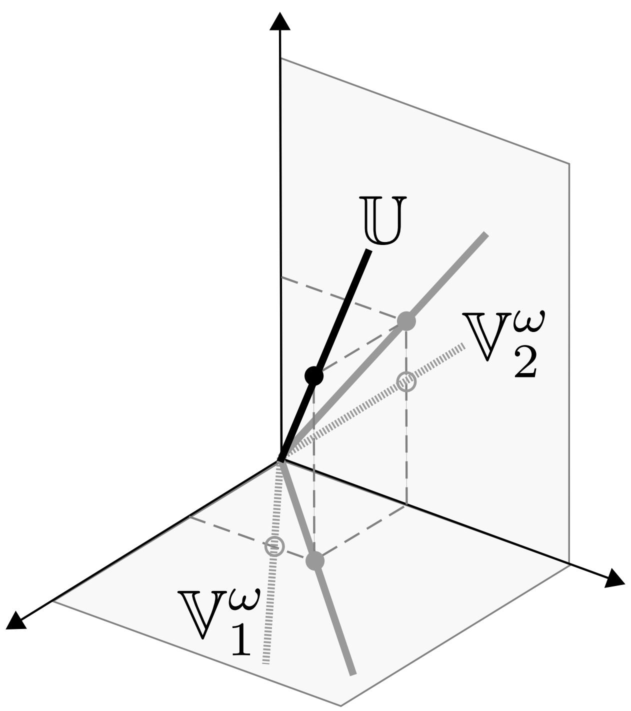


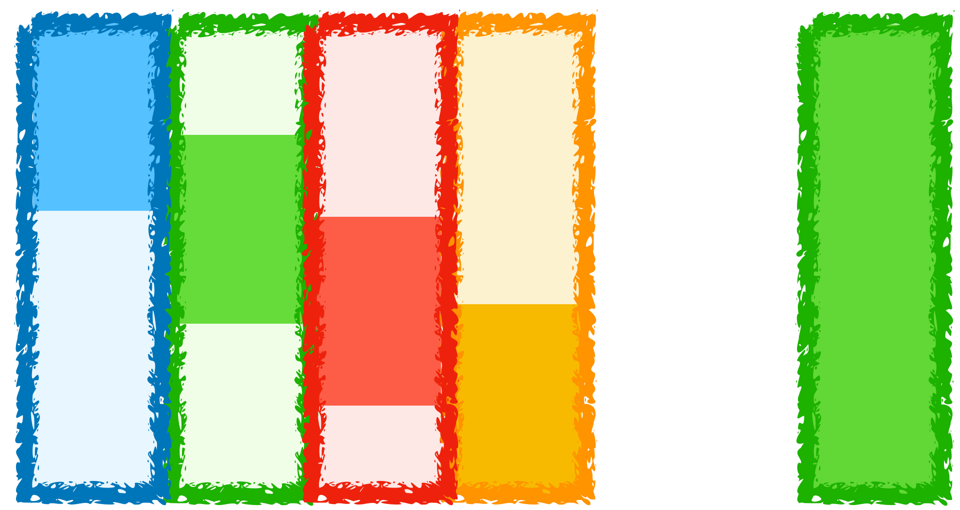
Problem Description

$\mathbb{U} :=$ $r-$ dimensional subspace of $\mathbb{R}^d, r < d$

$\mathbb{U}_i^\omega :=$ projections of the subspace onto canonical coordinates.

$\mathbb{V}_i^\omega := \mathbb{U}_i^\omega + Z_i^\omega$, where Z_i^ω is a noise matrix.





Goal: To estimate $\mathbb{U}$ from the $\{\mathbb{V}_i^\omega\}'$ s and bound the error

Outline

1. Problem Setup - Subspace Estimation
2. Motivation - Missing Data
3. Previous Work: Noiseless case
4. This Paper - Noisy Data and Estimation Bound
5. Applications
6. Conclusions

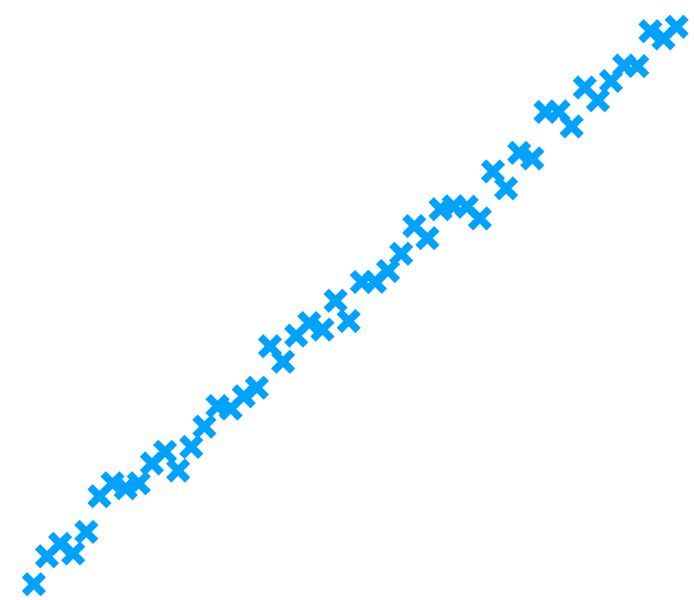
Motivation

The real world has lots of data

Motivation

Our main tool for modelling data is linear algebra

$$\begin{bmatrix} 1 & 2 & 1 & 3 & 2 & 6 \\ 1 & 2 & 1 & 3 & 2 & 6 \\ 1 & 2 & 1 & 3 & 2 & 6 \\ 2 & 4 & 2 & 6 & 4 & 12 \\ 2 & 4 & 2 & 6 & 4 & 12 \end{bmatrix}$$

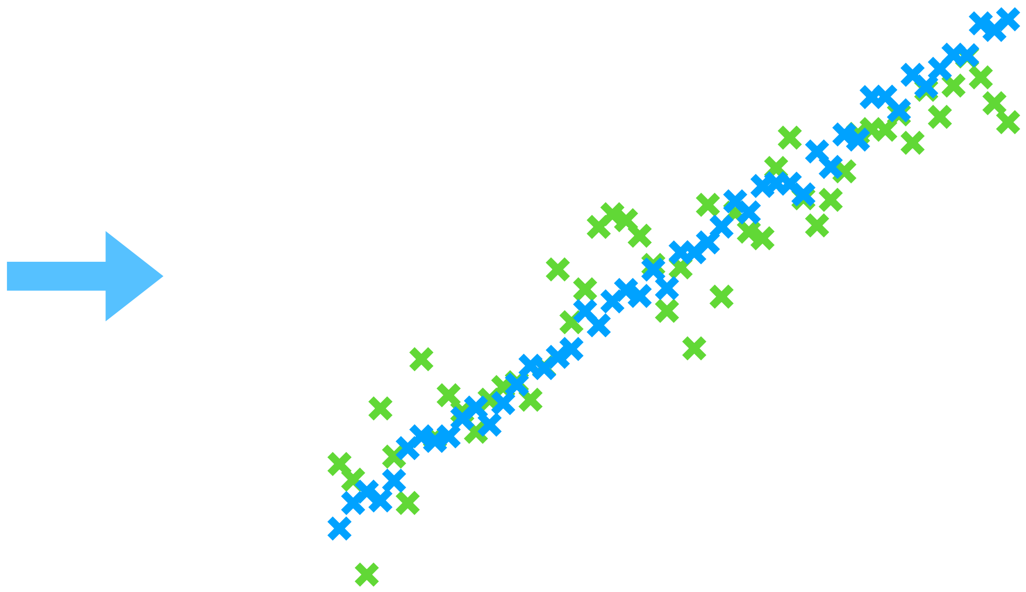


Since data is often best modelled by subspaces

Motivation

Our main tool for modelling data is linear algebra

1.1	1.9	0.8	3.1	2.2	5.8
0.9	2.2	1.2	3.2	2.1	6.2
1.2	1.8	1.1	2.9	1.8	5.9
2.1	3.9	2.2	5.8	4.1	12.1
1.9	4.1	2.1	5.9	4.2	12.2



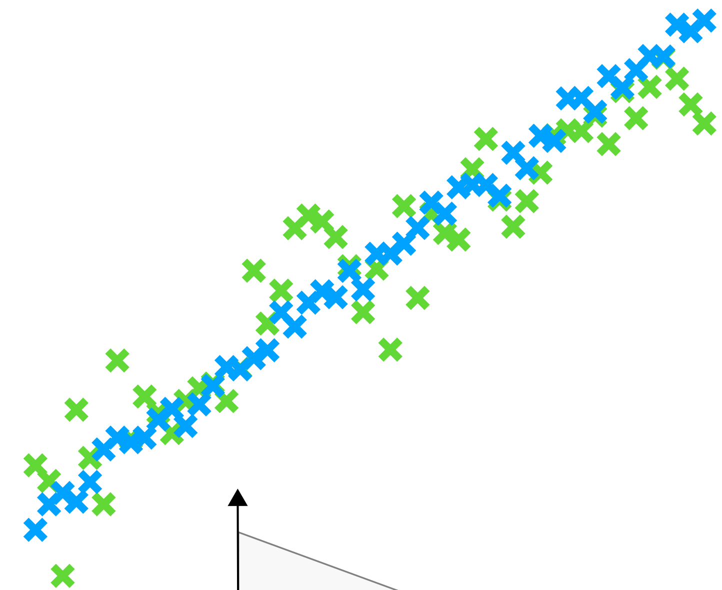
Since data is often best modelled by subspaces

But data is often noisy

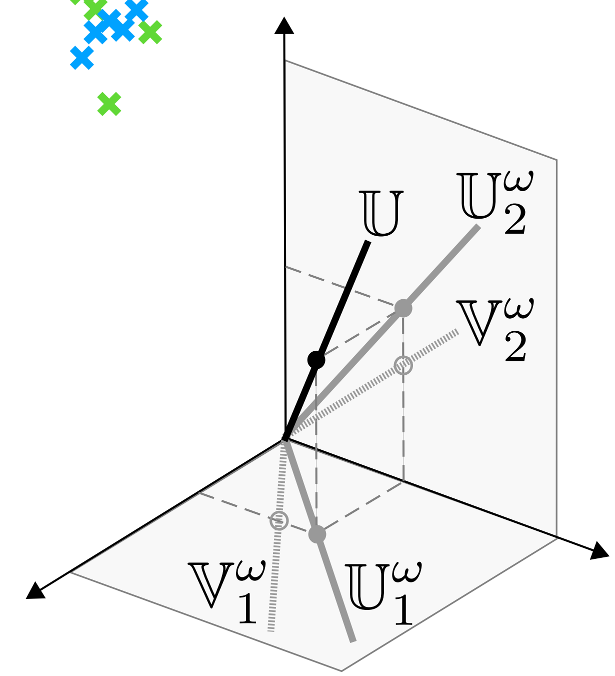
Motivation

Our main tool for modelling data is linear algebra

$$\begin{bmatrix} 1.1 & 1.9 & 0.8 & 3.1 & 2.2 & 5.8 \\ 0.9 & 2.2 & 1.2 & 3.2 & 2.1 & 6.2 \\ 1.2 & 1.8 & 1.1 & 2.9 & 1.8 & 5.9 \\ 2.1 & 3.9 & 2.2 & 5.8 & 4.1 & 12.1 \\ 1.9 & 4.1 & 2.1 & 5.9 & 4.2 & 12.2 \end{bmatrix}$$



$$\begin{bmatrix} 1.1 & 1.9 & 0.8 & \cdot & 2.2 & \cdot \\ 0.9 & 2.2 & \cdot & \cdot & 2.1 & 6.2 \\ 1.2 & \cdot & 1.1 & \cdot & 1.8 & 5.9 \\ \cdot & \cdot & 2.2 & 5.8 & \cdot & 12.1 \\ \cdot & 4.1 & 2.1 & 5.9 & 4.2 & \cdot \end{bmatrix}$$



Since data is often best modelled by subspaces

But data is often noisy and missing

Motivation

$$\begin{bmatrix} 1.1 & 1.9 & 0.8 & \cdot & 2.2 & \cdot \\ 0.9 & 2.2 & \cdot & \cdot & 2.1 & 6.2 \\ 1.2 & \cdot & 1.1 & \cdot & 1.8 & 5.9 \\ \cdot & \cdot & 2.2 & 5.8 & \cdot & 12.1 \\ \cdot & 4.1 & 2.1 & 5.9 & 4.2 & \cdot \end{bmatrix} \longrightarrow \begin{bmatrix} 1.1 & 1.9 & 0.8 & 3.1 & 2.2 & 5.8 \\ 0.9 & 2.2 & 1.2 & 3.2 & 2.1 & 6.2 \\ 1.2 & 1.8 & 1.1 & 2.9 & 1.8 & 5.9 \\ 2.1 & 3.9 & 2.2 & 5.8 & 4.1 & 12.1 \\ 1.9 & 4.1 & 2.1 & 5.9 & 4.2 & 12.2 \end{bmatrix}$$

The problem of completing matrices is
(Low Rank) Matrix Completion

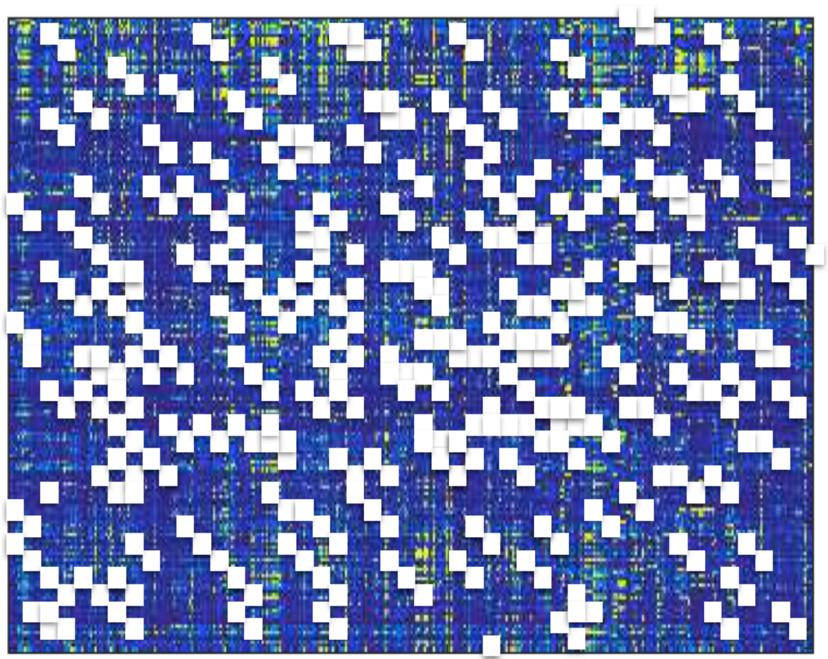
Motivation

$$\begin{bmatrix} 1.1 & 1.9 & 0.8 & \cdot & 2.2 & \cdot \\ 0.9 & 2.2 & \cdot & \cdot & 2.1 & 6.2 \\ 1.2 & \cdot & 1.1 & \cdot & 1.8 & 5.9 \\ \cdot & \cdot & 2.2 & 5.8 & \cdot & 12.1 \\ \cdot & 4.1 & 2.1 & 5.9 & 4.2 & \cdot \end{bmatrix} \longrightarrow \begin{bmatrix} 1.1 & 1.9 & 0.8 & 3.1 & 2.2 & 5.8 \\ 0.9 & 2.2 & 1.2 & 3.2 & 2.1 & 6.2 \\ 1.2 & 1.8 & 1.1 & 2.9 & 1.8 & 5.9 \\ 2.1 & 3.9 & 2.2 & 5.8 & 4.1 & 12.1 \\ 1.9 & 4.1 & 2.1 & 5.9 & 4.2 & 12.2 \end{bmatrix}$$

The problem of completing matrices is (Low Rank) Matrix Completion

Targets (proteins)

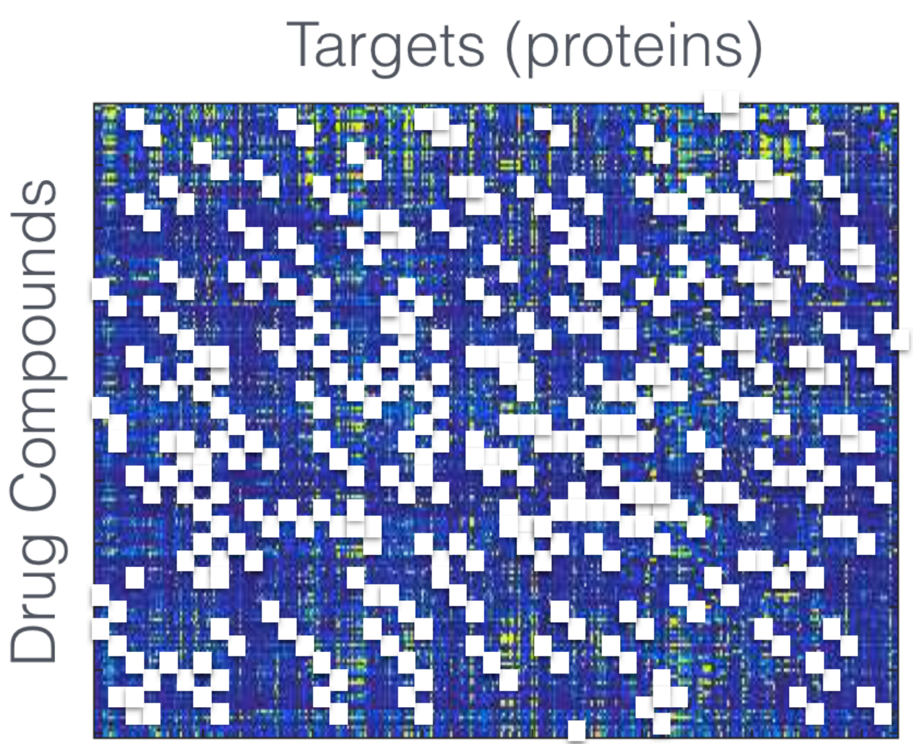
Drug Compounds

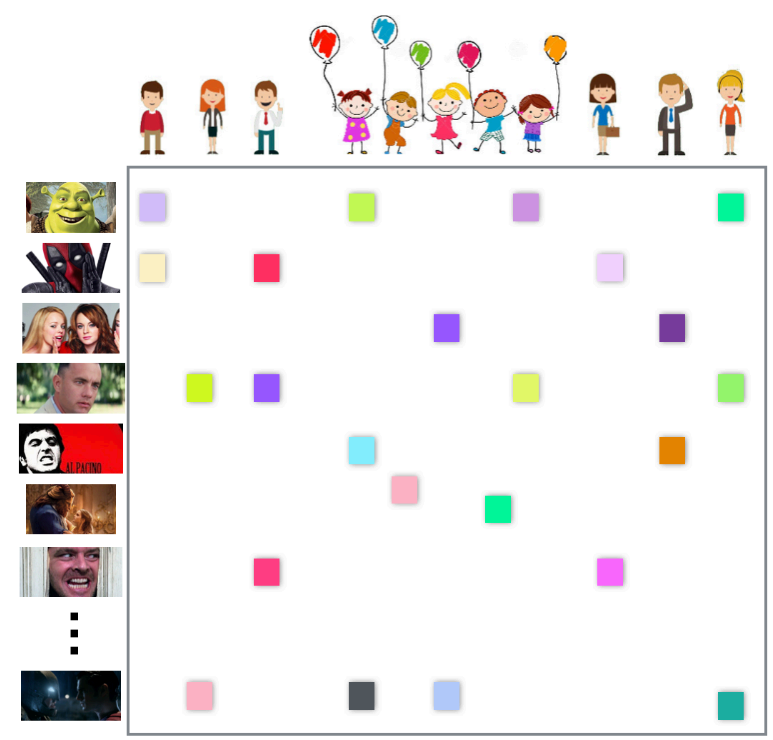


Motivation

$$\begin{bmatrix} 1.1 & 1.9 & 0.8 & \cdot & 2.2 & \cdot \\ 0.9 & 2.2 & \cdot & \cdot & 2.1 & 6.2 \\ 1.2 & \cdot & 1.1 & \cdot & 1.8 & 5.9 \\ \cdot & \cdot & 2.2 & 5.8 & \cdot & 12.1 \\ \cdot & 4.1 & 2.1 & 5.9 & 4.2 & \cdot \end{bmatrix} \rightarrow \begin{bmatrix} 1.1 & 1.9 & 0.8 & 3.1 & 2.2 & 5.8 \\ 0.9 & 2.2 & 1.2 & 3.2 & 2.1 & 6.2 \\ 1.2 & 1.8 & 1.1 & 2.9 & 1.8 & 5.9 \\ 2.1 & 3.9 & 2.2 & 5.8 & 4.1 & 12.1 \\ 1.9 & 4.1 & 2.1 & 5.9 & 4.2 & 12.2 \end{bmatrix}$$

The problem of completing matrices is (Low Rank) Matrix Completion





Motivation

Subspace Reconstruction sheds light on these problems

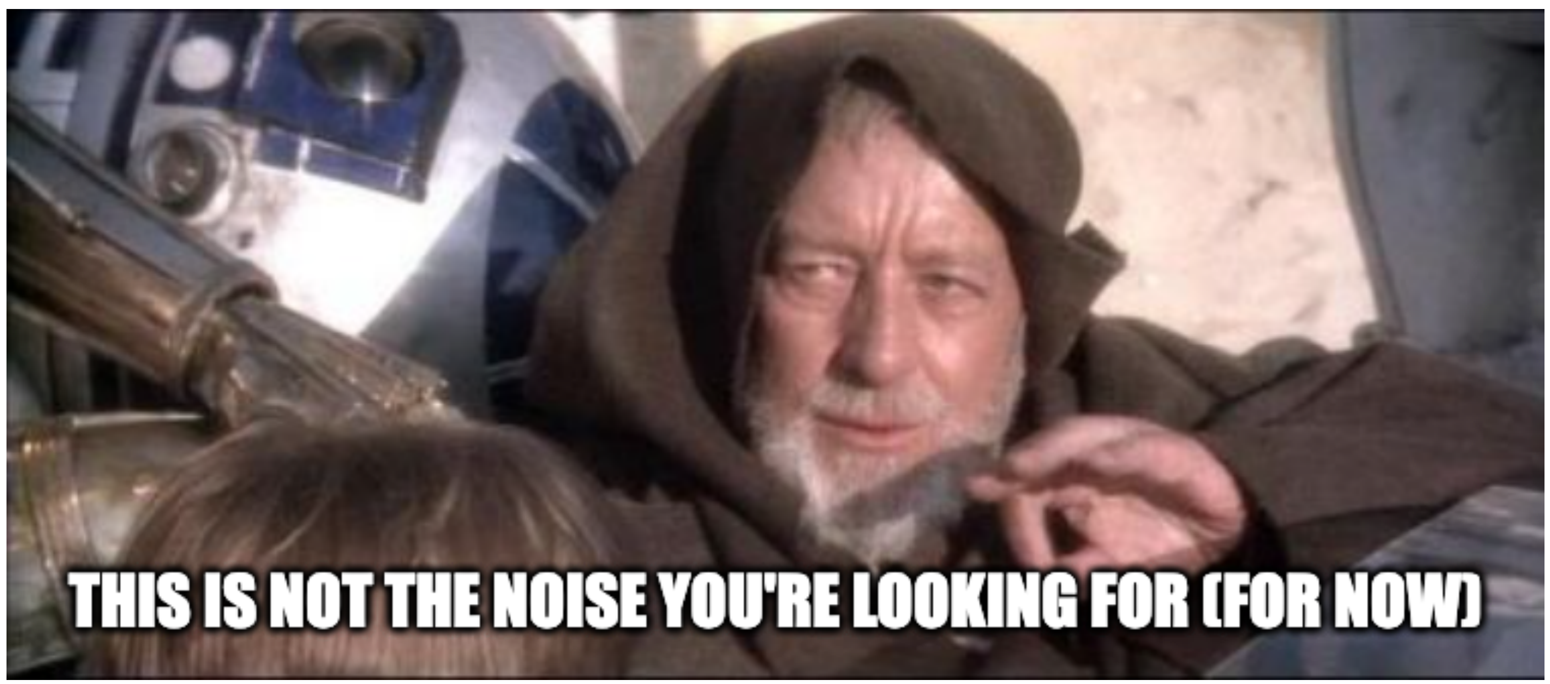
More to come later!

Outline

1. Problem Setup - Subspace Estimation
2. Motivation - Missing Data
3. Previous Work: Noiseless case
4. This Paper - Noisy Data and Estimation Bound
5. Applications
6. Conclusions

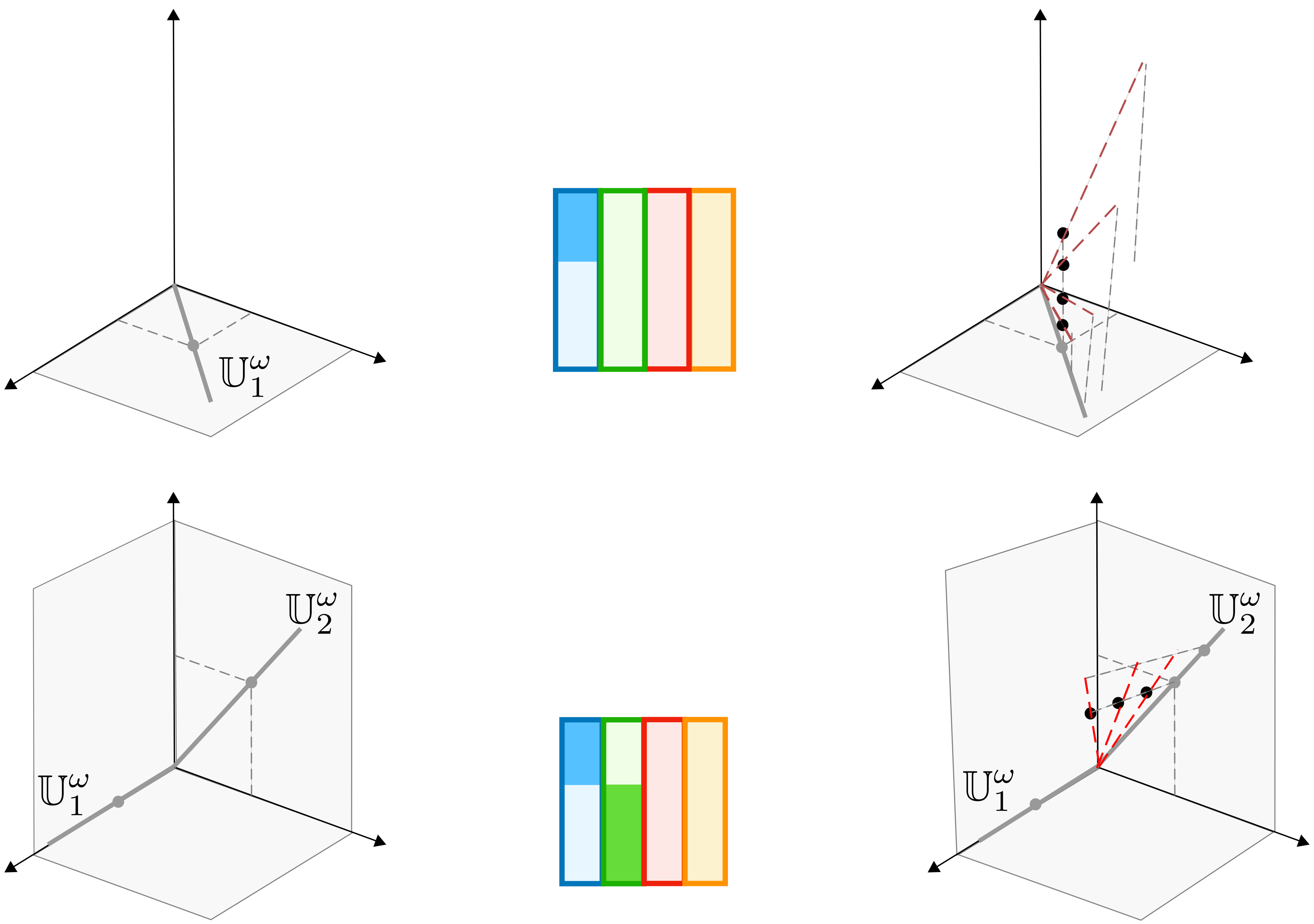
Previous Work - Noiseless Case

Let's forget about noisy data for a second.



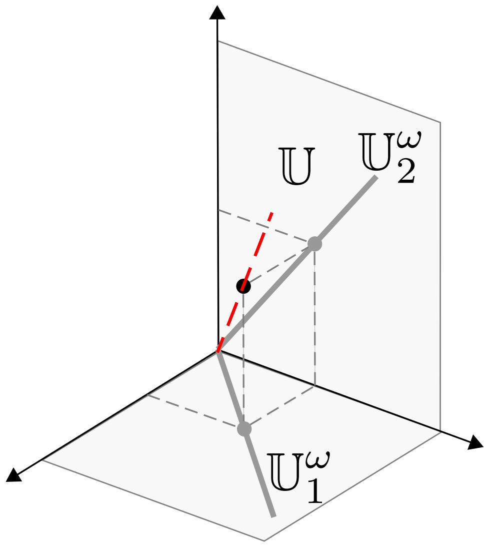
Previous Work - Noiseless Case

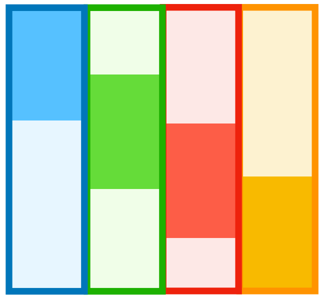
Many projections don't have unique subspace estimations

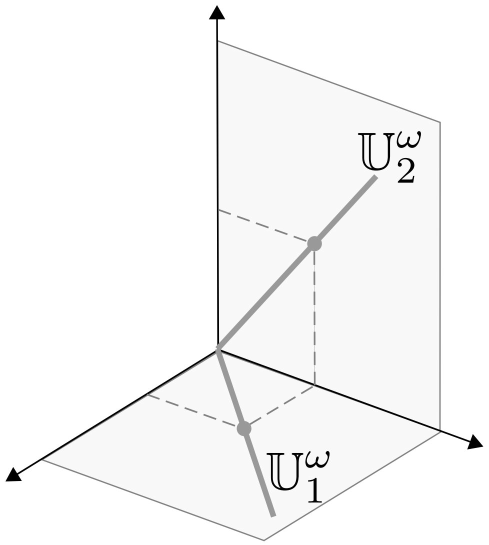


Previous Work - Noiseless Case

Many projections don't have unique subspace estimations, but many do.

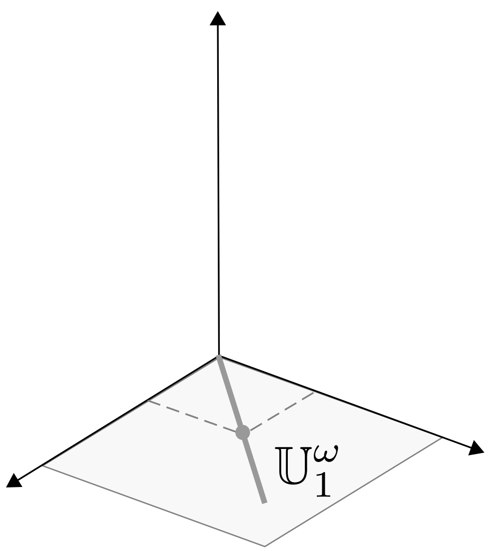


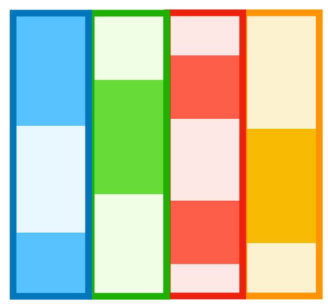


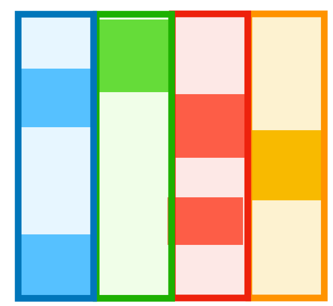


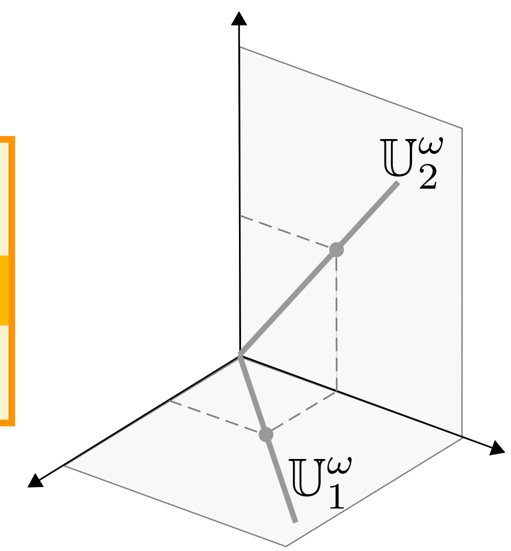
Previous Work - Noiseless Case

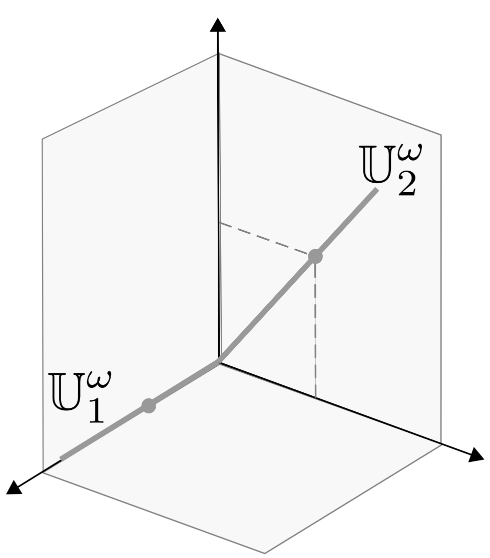
Many projections don't have unique subspace estimations, but many do.

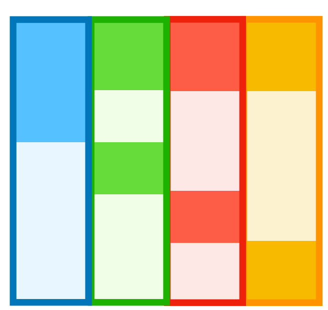












Question: Which projections do we need to observe for a unique reconstruction?

Previous Work - Noiseless Case

Theorem 1 - (Pimentel-Alarcón, Nowak, Boston, ISIT '15)

$\mathbb{U}$ can be recovered from $\{\mathbb{U}_i^\omega\}'$s if and only if $\{\mathbb{U}_i^\omega\}'$s are observed in the right places*. Meaning:

Each subset of n projections has at least n+r known entries

Previous Work - Noiseless Case

Theorem 1 - (Pimentel-Alarcón, Nowak, Boston, ISIT '15)

$\mathbb{U}$ can be recovered from $\{\mathbb{U}_i^{\omega}\}'s$ if and only if $\{\mathbb{U}_i^{\omega}\}'s$ are observed in the right places*. Meaning:

Each subset of n projections has at least n+r known entries

---

*There are some more technical conditions

- There are N = d - r observations (WLOG)
- Each observation has r+1 entries (WLOG)

But this is the essence of the deterministic result

Previous Work - Noiseless Case

Theorem 1 - (Pimentel-Alarcón, Nowak, Boston, ISIT '15)

$\mathbb{U}$ can be recovered from $\{\mathbb{U}_i^{\omega}\}'$s if and only if $\{\mathbb{U}_i^{\omega}\}'$s are observed in the right places*. Meaning:

Each subset of n projections has at least n+r known entries

More on the construction soon!

Outline

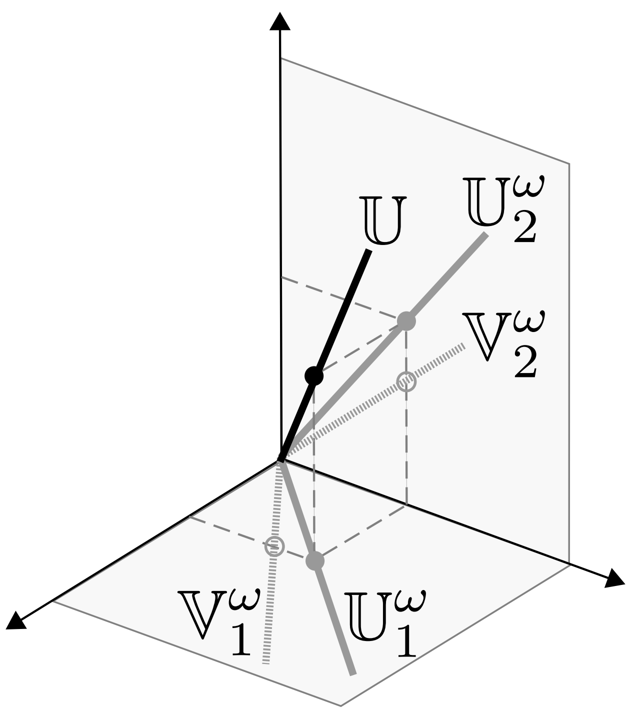
1. Problem Setup - Subspace Estimation
2. Motivation - Missing Data
3. Previous Work: Noiseless case
4. This Paper - Noisy Data and Estimation Bound
5. Applications
6. Conclusions

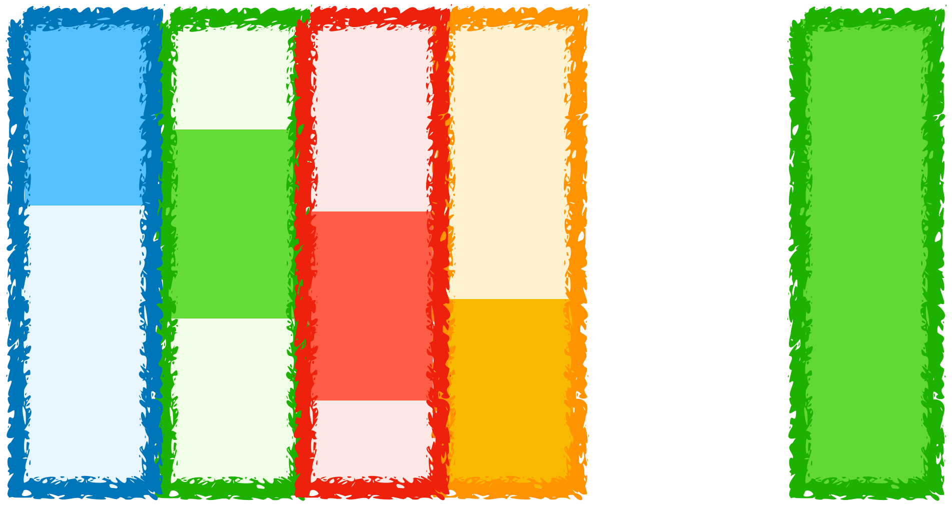
Noisy Data and Estimation Bound

$\mathbb{U}$:= $r-$ dimensional subspace of $\mathbb{R}^d, r < d$

$\mathbb{U}_i^\omega$:= projections of the subspace onto canonical coordinates.

$\mathbb{V}_i^\omega := \mathbb{U}_i^\omega + Z_i^\omega$, where Z_i^ω is a noise matrix.





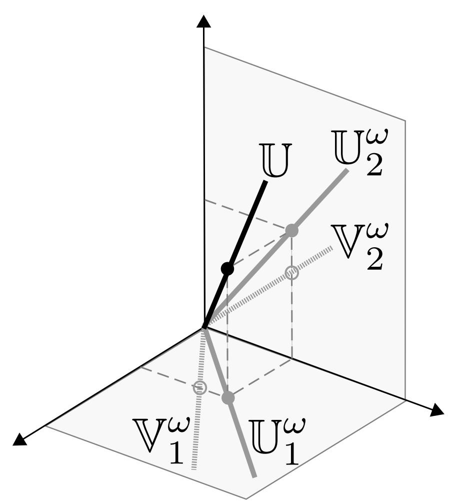
Goal: To estimate $\mathbb{U}$ from the $\{\mathbb{V}_i^\omega\}'$s and bound the error

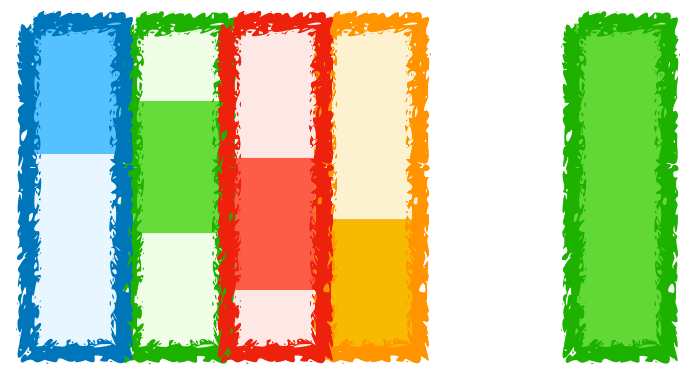
Noisy Data and Estimation Bound

$\mathbb{U} :=$ $r-$ dimensional subspace of $\mathbb{R}^d, r < d$

$\mathbb{U}_i^\omega :=$ projections of the subspace onto canonical coordinates.

$\mathbb{V}_i^\omega := \mathbb{U}_i^\omega + Z_i^\omega$, where Z_i^ω is a noise matrix.





Goal: To estimate $\mathbb{U}$ from the $\{\mathbb{V}_i^\omega\}'$s and bound the error

---

Technical assumptions:

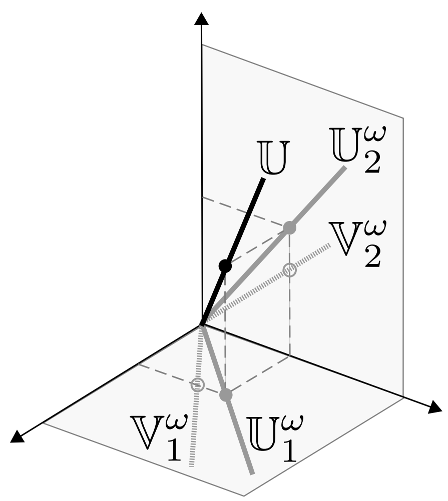
➛ The sampling satisfies the conditions in Theorem 1 (P.-A., et.al)

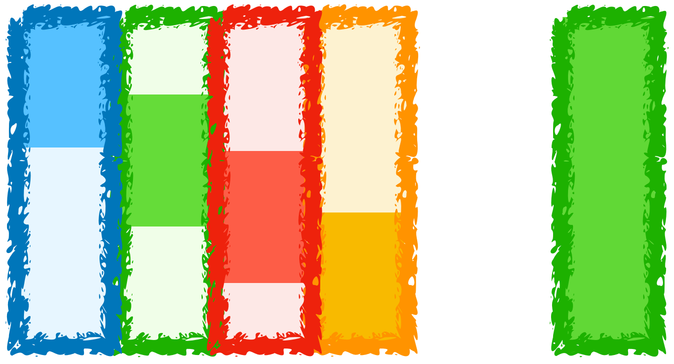
Noisy Data and Estimation Bound

$\mathbb{U} :=$ $r-$ dimensional subspace of $\mathbb{R}^d, r < d$

$\mathbb{U}_i^\omega :=$ projections of the subspace onto canonical coordinates.

$\mathbb{V}_i^\omega := \mathbb{U}_i^\omega + Z_i^\omega$, where Z_i^ω is a noise matrix.





Goal: To estimate $\mathbb{U}$ from the $\{\mathbb{V}_i^\omega\}'$s and bound the error

Technical assumptions:

- The sampling satisfies the conditions in Theorem 1 (P.-A., et.al)
- $|Z_i^\omega|_2 < \epsilon$, so ϵ is the noise
- $\min_i(\sigma(\mathbb{V}_i^\omega)) =: \delta$, and δ is the signal

Main Result

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

This refers to a set of measure zero subspaces that we will not identify.

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

This is the distance between $\mathbb{U}$ and $\hat{\mathbb{U}}$ on the grassmannian manifold - a standard metric to measure how different two subspaces are.

$$d_G(A, B) = \frac{1}{\sqrt{2}}\|P_A - P_B\|_F$$

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

This is the (multiplicative inverse of the) *signal-to-noise ratio* where

- $|Z_i^\omega|_2 < \epsilon$, for every i, bounds the *noise*
- $\min_i(\sigma(\mathbb{V}_i^\omega)) =: \delta$ requires that the *signal* power is at least δ

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

$$\epsilon = 0 \implies d_G(\mathbb{U}, \hat{\mathbb{U}}) = 0 \implies \mathbb{U} = \hat{\mathbb{U}}$$

Taking $\epsilon = 0$, this bound recovers Theorem 1 (P.-A., et. al).

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

This number is the number of degrees of freedom of an $r-$dimensional subspace of $\mathbb{R}^d$.

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

B is a matrix that encodes the sampling and the orientation of the subspace (more later). $\sigma(B)$ is its smallest singular value.

Noisy Data and Estimation Bound

Computability: The bound is not just theoretical but very computable (i.e. it is based on the observed noisy data)

Theorem (S., P.-A., ISIT 2022)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

Signal-to-noise ratio

Constructed from kernel of noisy projections (Observed data)

Degrees of freedom of r - dimensional subspaces

Sketch of Construction and Proof

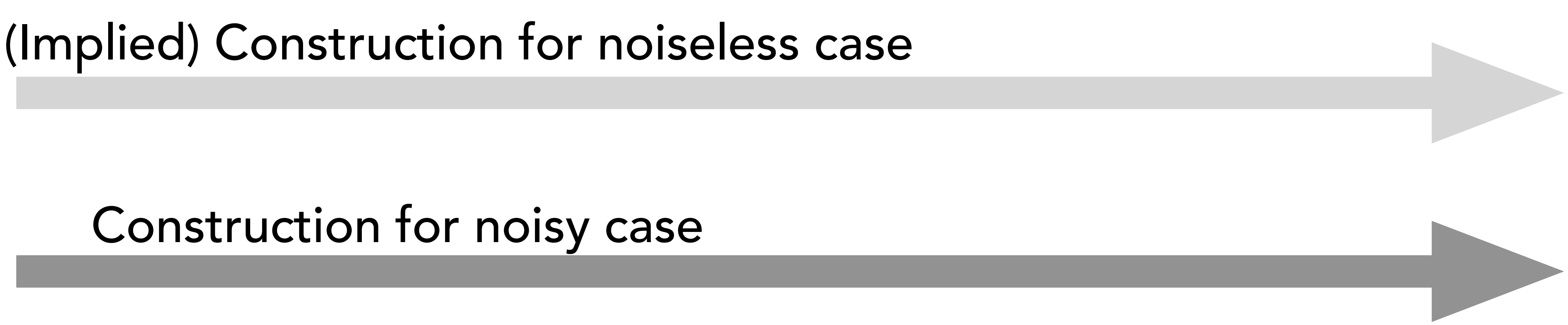
Sketch of Construction and Proof

Gameplan

Construction for noisy case

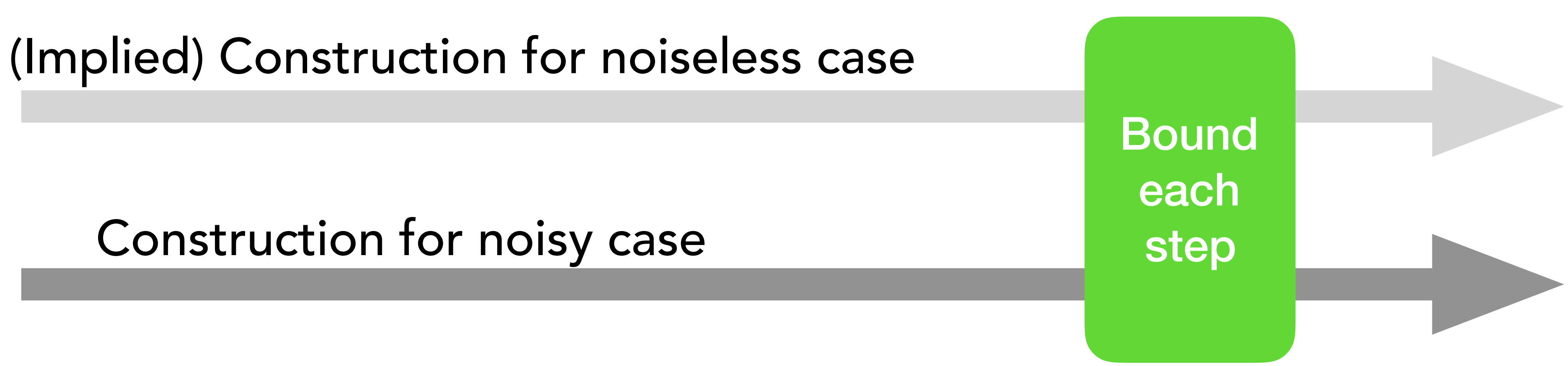
Sketch of Construction and Proof

Gameplan



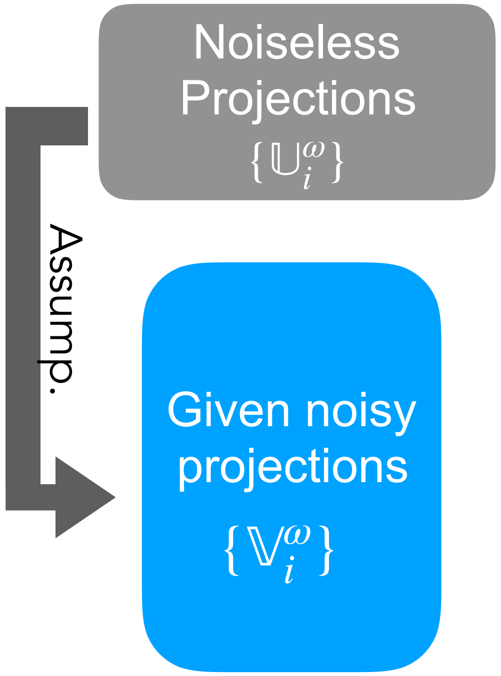
Sketch of Construction and Proof

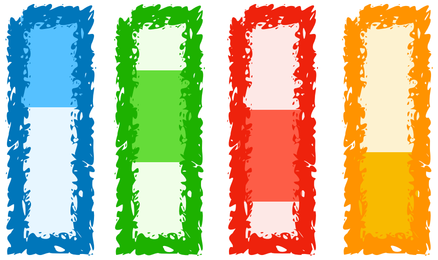
Gameplan



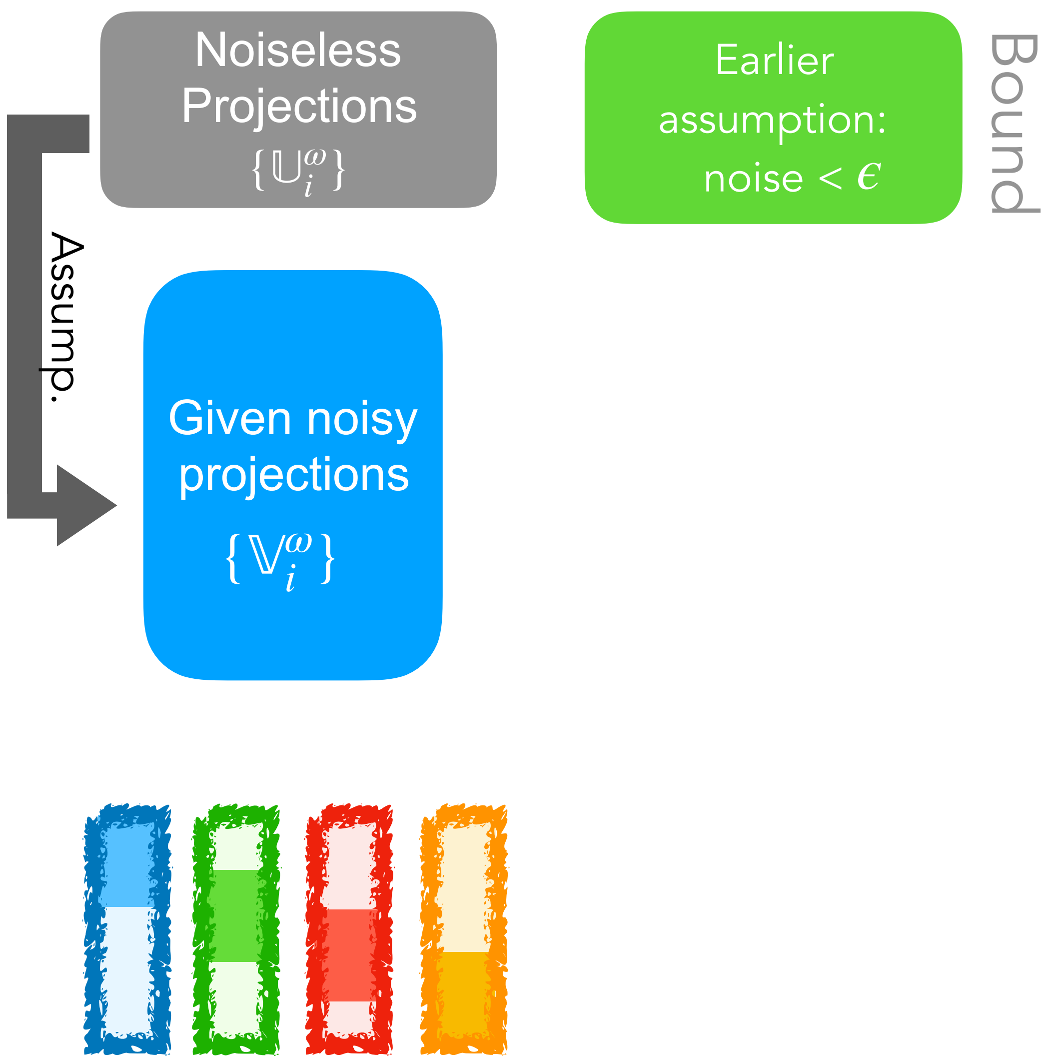
ROLL UP YOUR SLEEVES!

Noisy Data and Estimation Bound

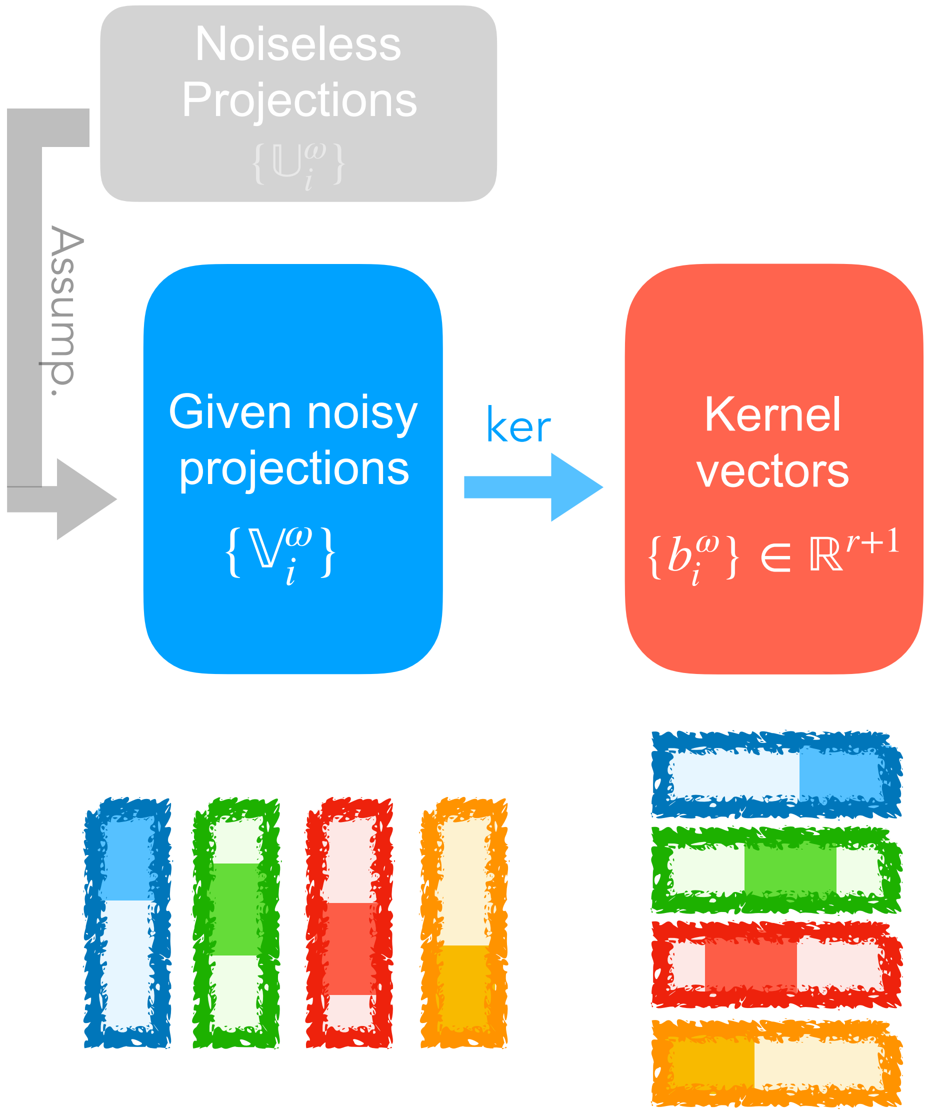




Noisy Data and Estimation Bound

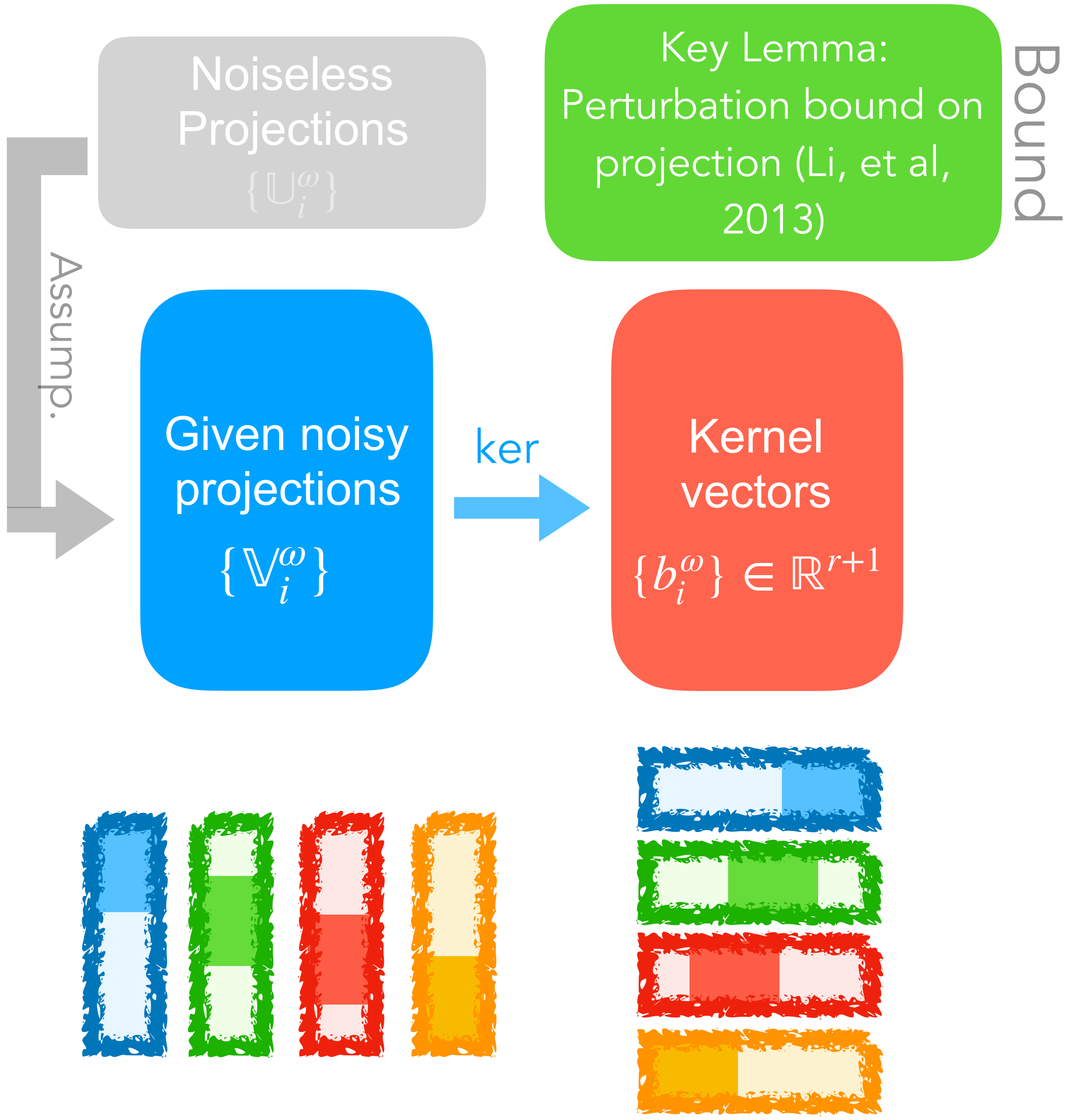


Noisy Data and Estimation Bound



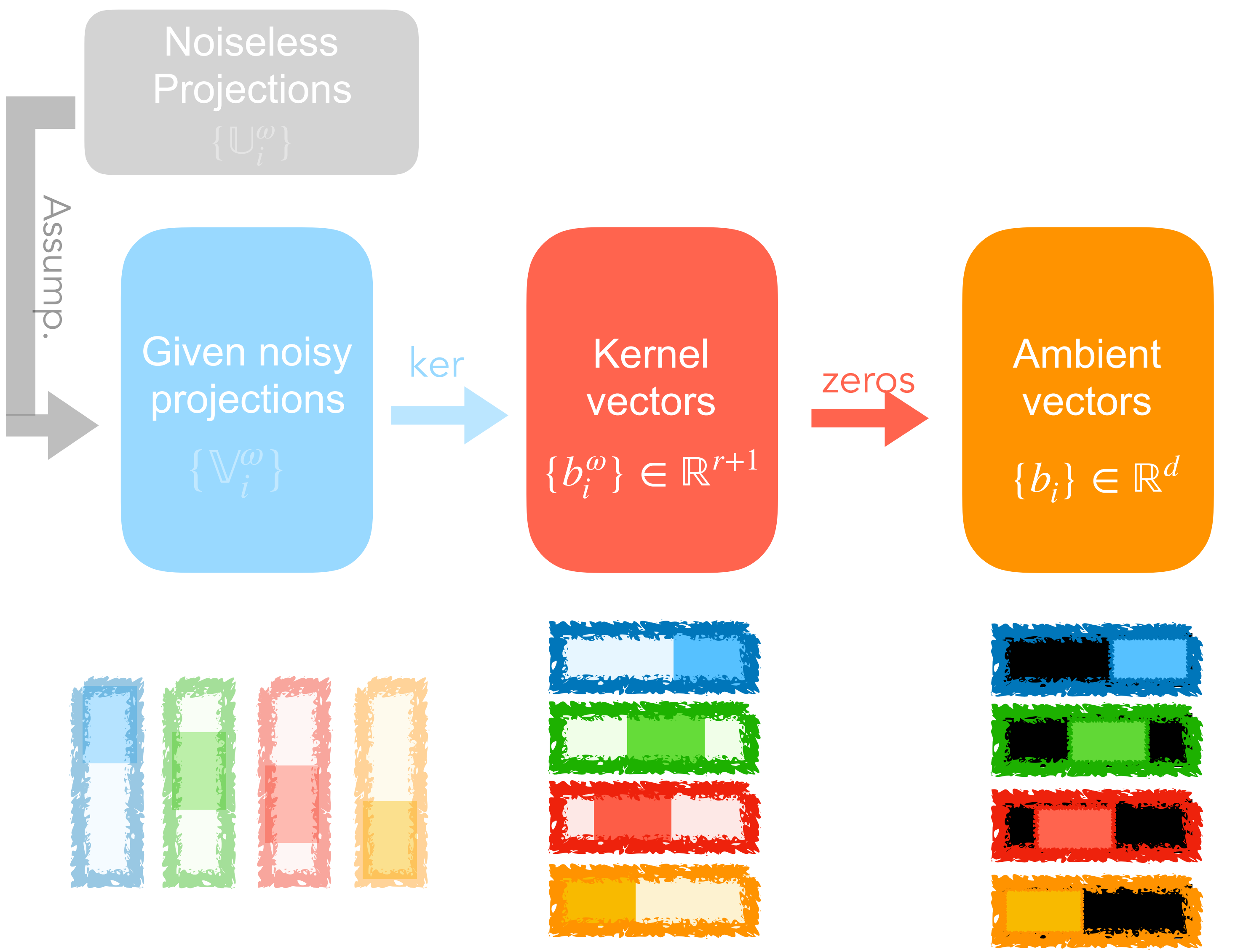
Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Noisy Data and Estimation Bound



Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

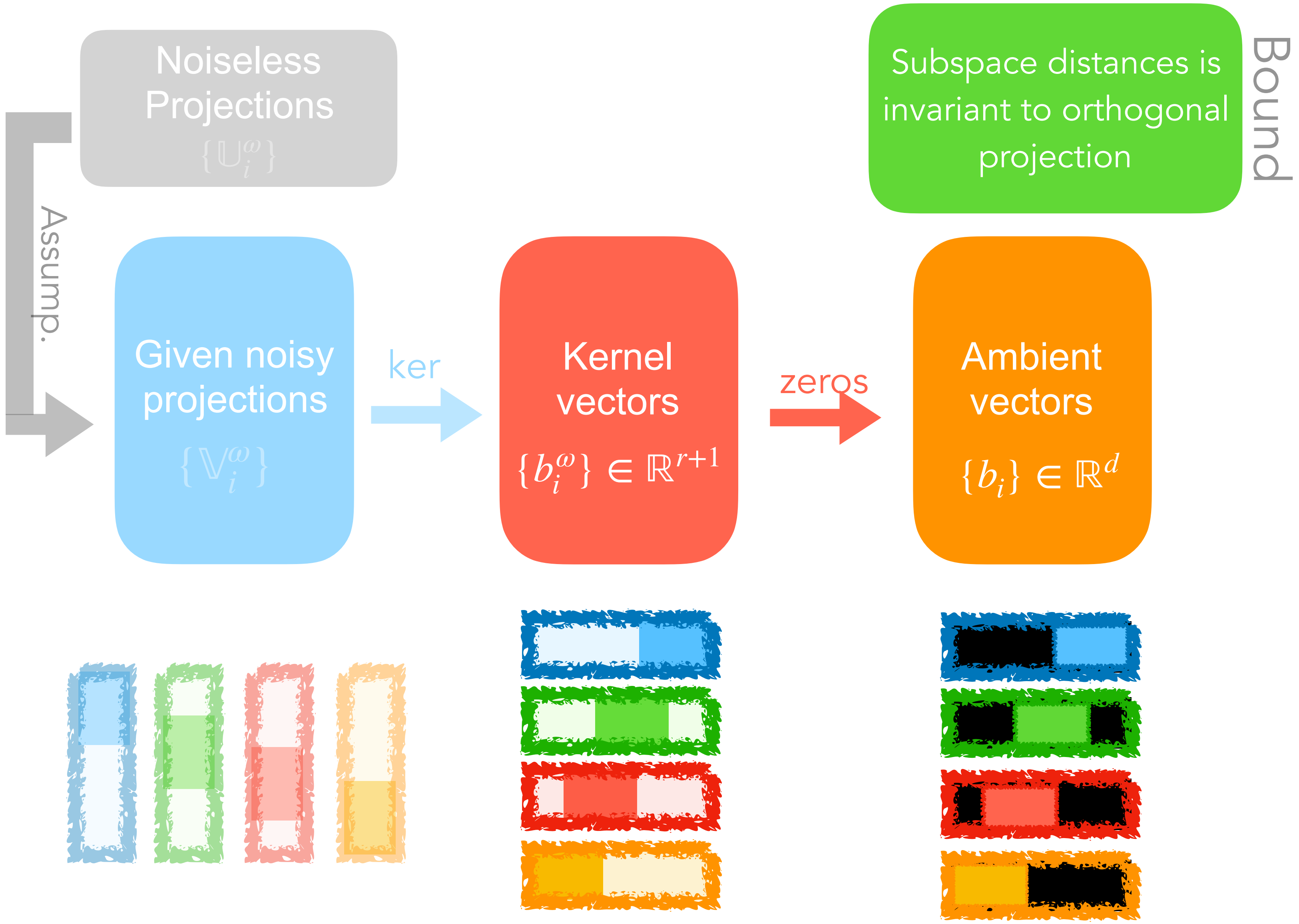
Noisy Data and Estimation Bound



Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Step 2: Construct vectors $b_i \in \mathbb{R}^d$ by padding b_i^ω with zeros where the sampling is zero

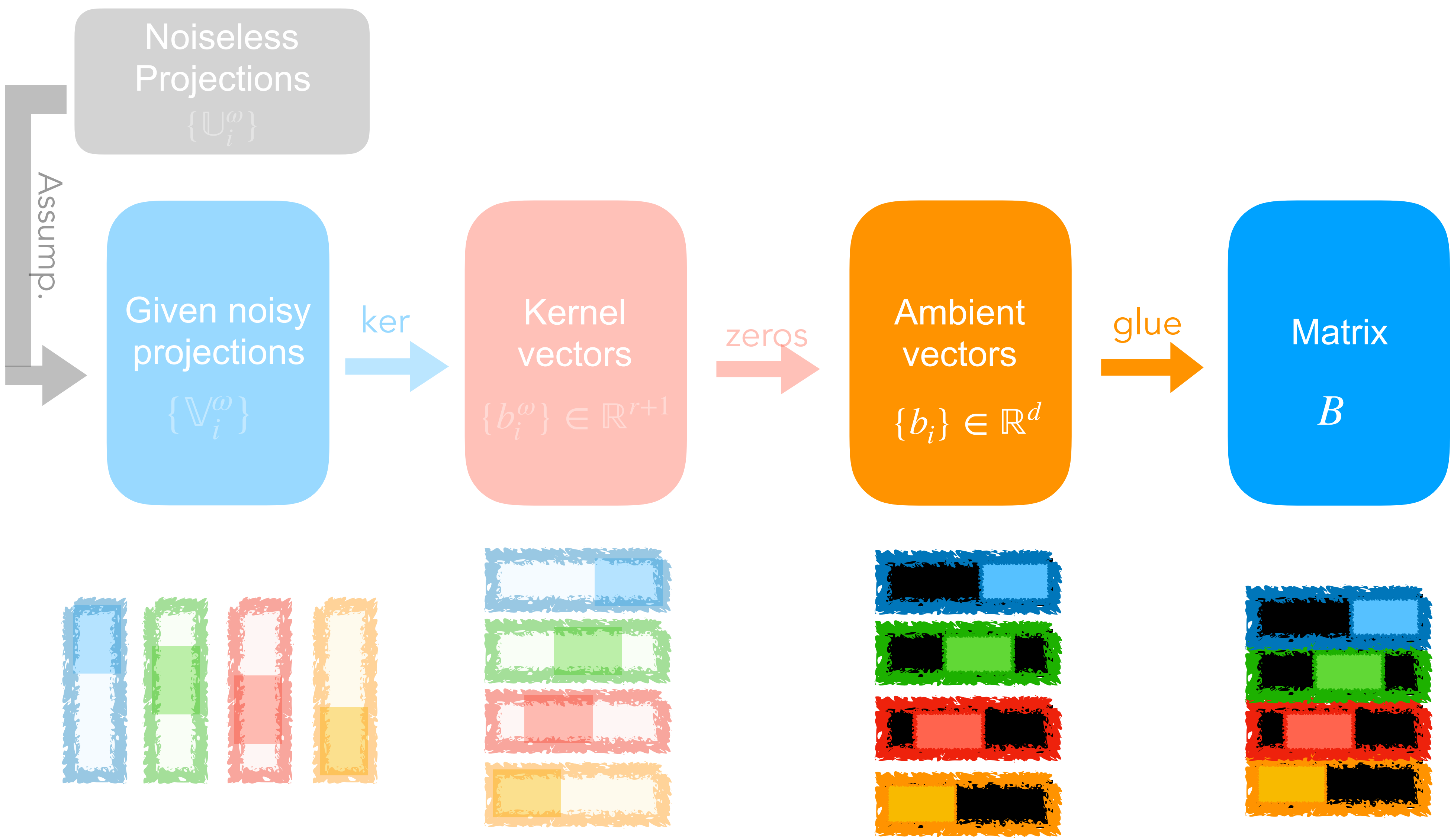
Noisy Data and Estimation Bound



Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Step 2: Construct vectors $b_i \in \mathbb{R}^d$ by padding b_i^ω with zeros according to the sampling

Noisy Data and Estimation Bound

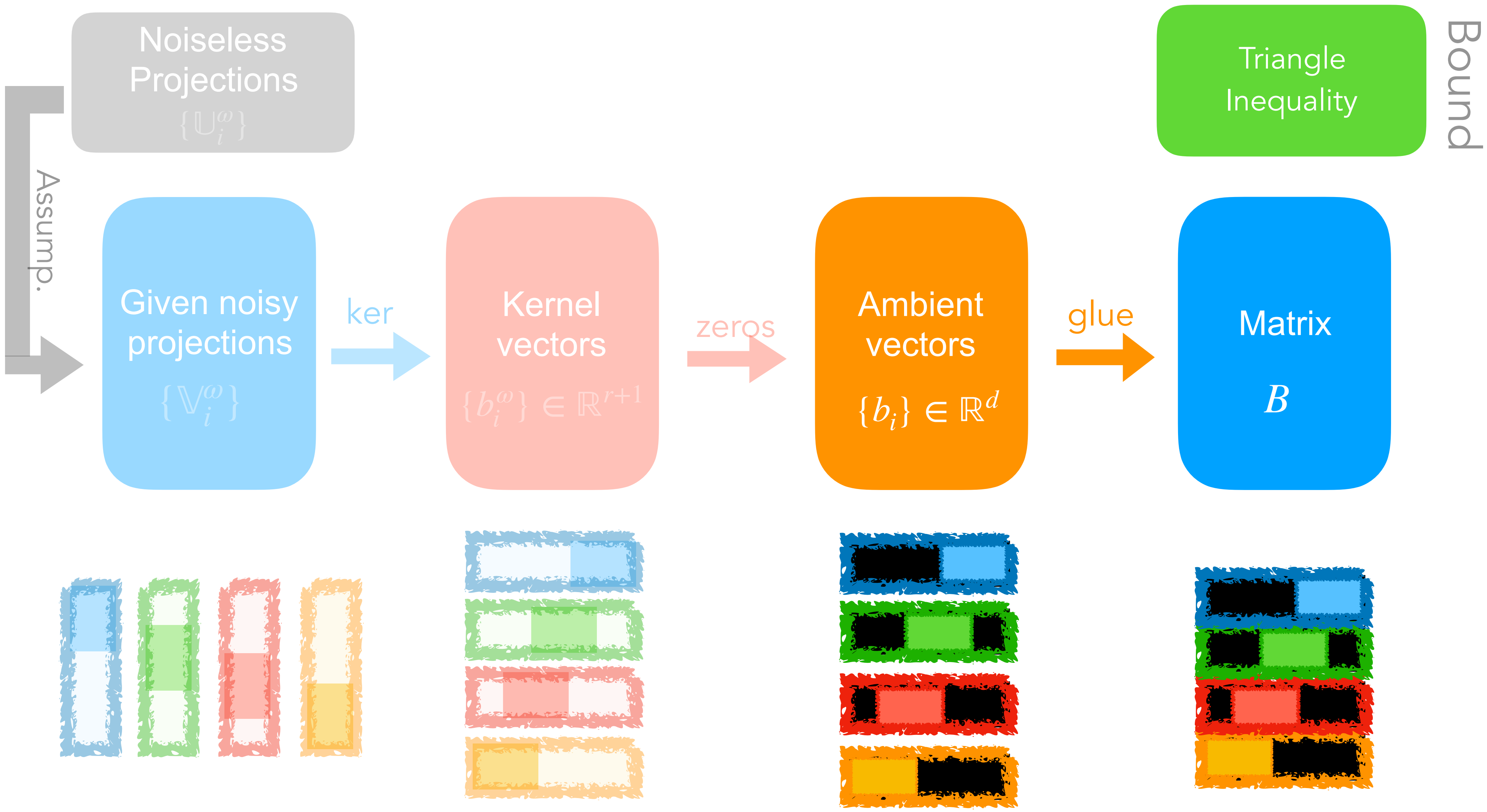


Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Step 2: Construct vectors $b_i \in \mathbb{R}^d$ by padding b_i^ω with zeros according to the sampling

Step 3: Form a matrix $B = [b_1 b_2 \ldots b_N]$

Noisy Data and Estimation Bound

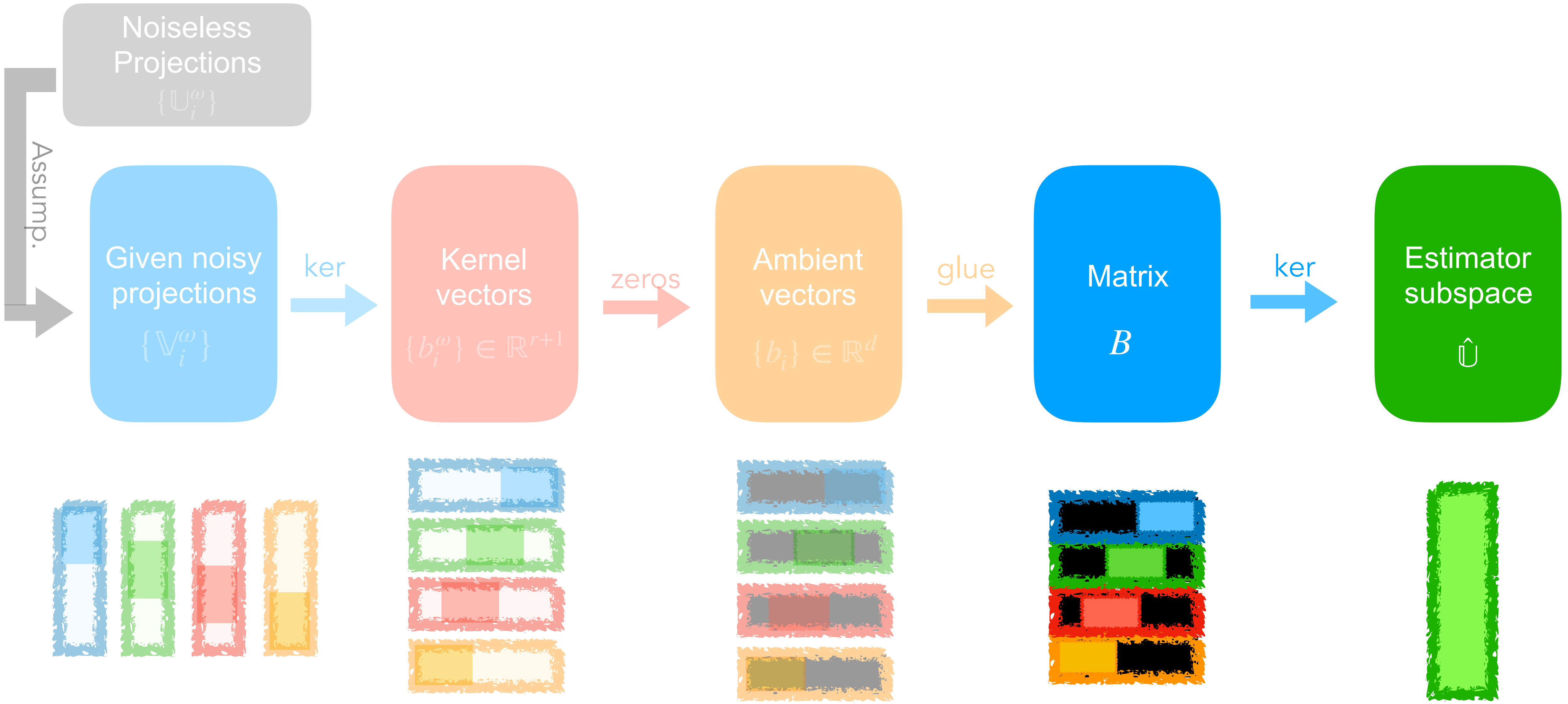


Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Step 2: Construct vectors $b_i \in \mathbb{R}^d$ by padding b_i^ω with zeros according to the sampling

Step 3: Form a matrix $B = [b_1 b_2 \ldots b_N]$

Noisy Data and Estimation Bound



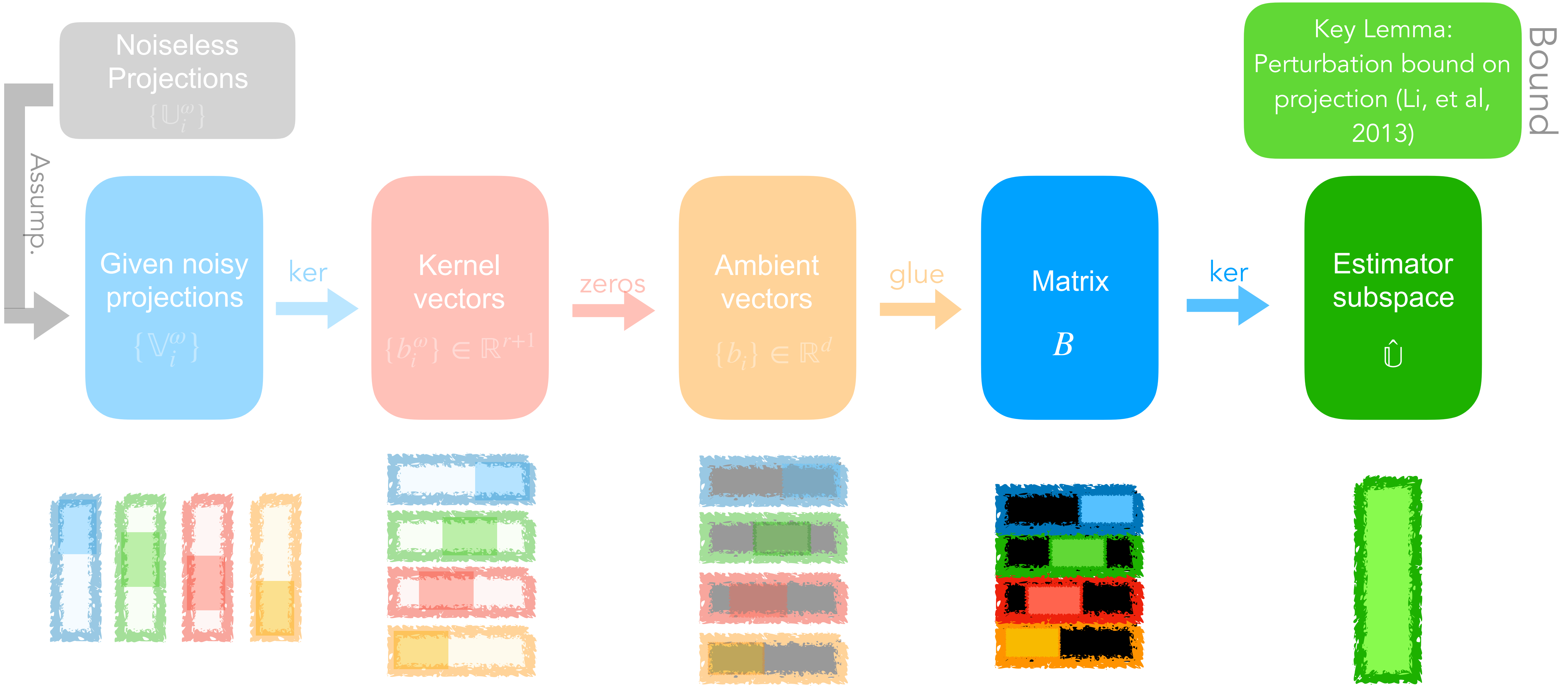
Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Step 2: Construct vectors $b_i \in \mathbb{R}^d$ by padding b_i^ω with zeros according to the sampling

Step 3: Form a matrix $B = [b_1 b_2 \ldots b_N]$

Step 4: Take $\hat{\mathbb{U}} = \ker B$

Noisy Data and Estimation Bound



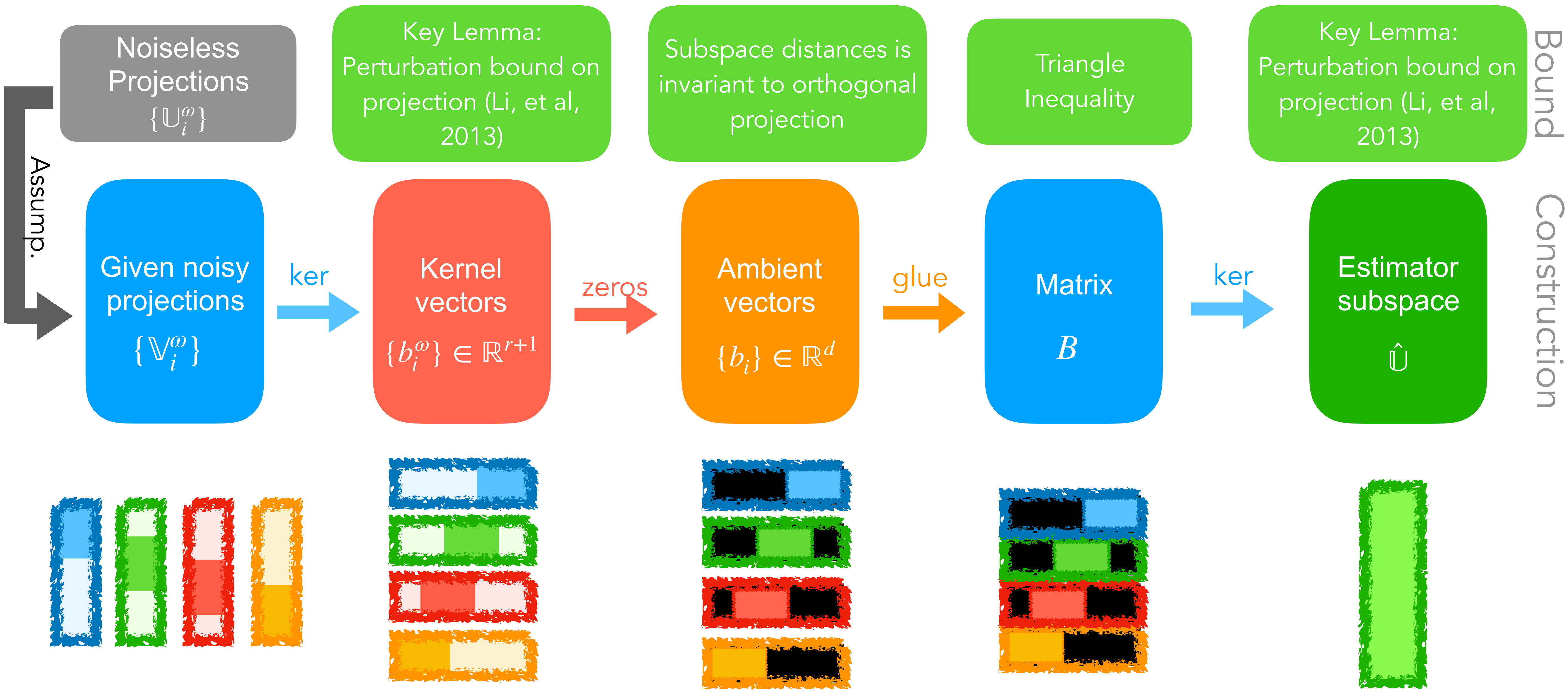
Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Step 2: Construct vectors $b_i \in \mathbb{R}^d$ by padding b_i^ω with zeros according to the sampling

Step 3: Form a matrix $B = [b_1 b_2 \ldots b_N]$

Step 4: Take $\hat{\mathbb{U}} = \ker B$

Noisy Data and Estimation Bound



Step 1: Take a normal vector $b_i^\omega \in \ker \mathbb{V}_i^\omega \subset \mathbb{R}^{d+1}$ for each projection

Step 2: Construct vectors $b_i \in \mathbb{R}^d$ by padding b_i^ω with zeros according to the sampling

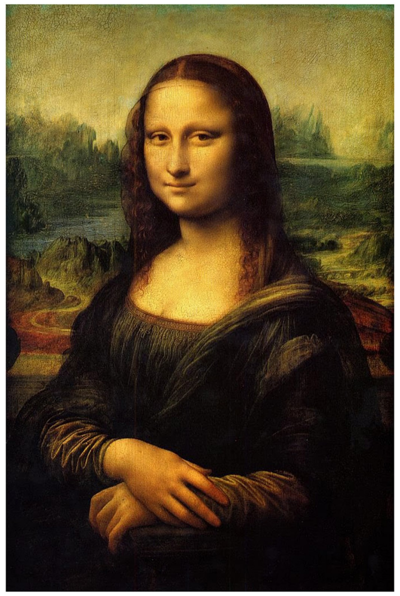
Step 3: Form a matrix $B = [b_1 b_2 \ldots b_N]$

Step 4: Take $\hat{\mathbb{U}} = \ker B$

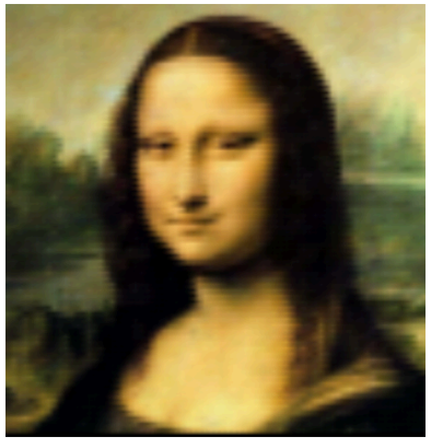
Noisy Data and Estimation Bound

Okay, but **how good is the estimator**?

True Space $\mathbb{U}$



Estimator $\hat{\mathbb{U}}$



or

?

Noisy Data and Estimation Bound

Theorem (S., P.-A., this paper)

For almost every $\mathbb{U}$,

$$d_G(\mathbb{U}, \hat{\mathbb{U}}) \leq \frac{\epsilon\sqrt{2r(d-r)}}{\delta \cdot \sigma(B)}$$

What affects this?

Noisy Data and Estimation Bound

$\sigma(B)$

Noisy Data and Estimation Bound

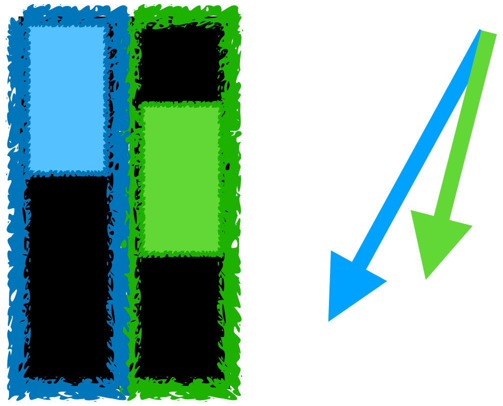
$\sigma(B)$

Residual of projecting
columns of B
onto each other

Noisy Data and Estimation Bound

$\sigma(B)$

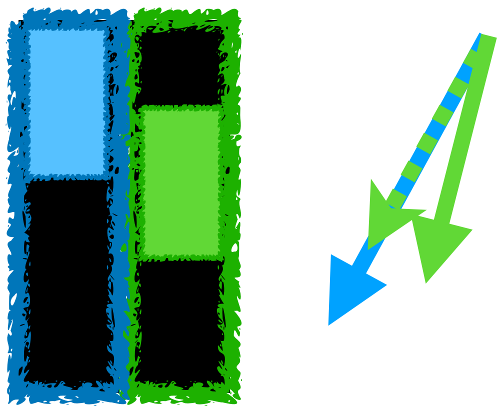
Residual of projecting
columns of B
onto each other



Noisy Data and Estimation Bound

$\sigma(B)$

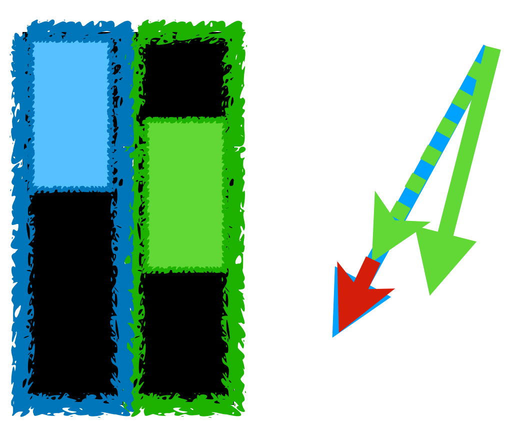
Residual of projecting
columns of B
onto each other



Noisy Data and Estimation Bound

$\sigma(B)$

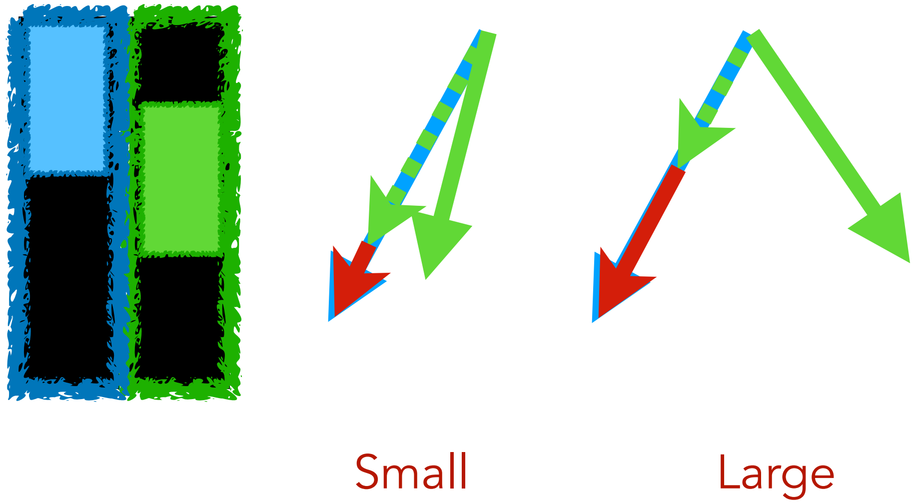
Residual of projecting columns of B onto each other



Noisy Data and Estimation Bound

$\sigma(B)$

Residual of projecting columns of B onto each other

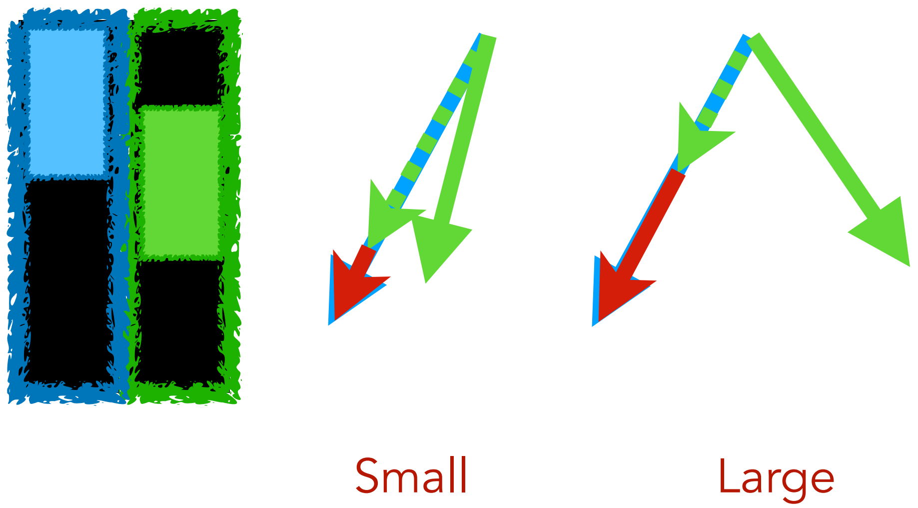


Noisy Data and Estimation Bound

$\sigma(B)$

Residual of projecting columns of B onto each other

Orientation of Projections $\mathbb{U}_i^{\omega}/\mathbb{V}_i^{\omega}$

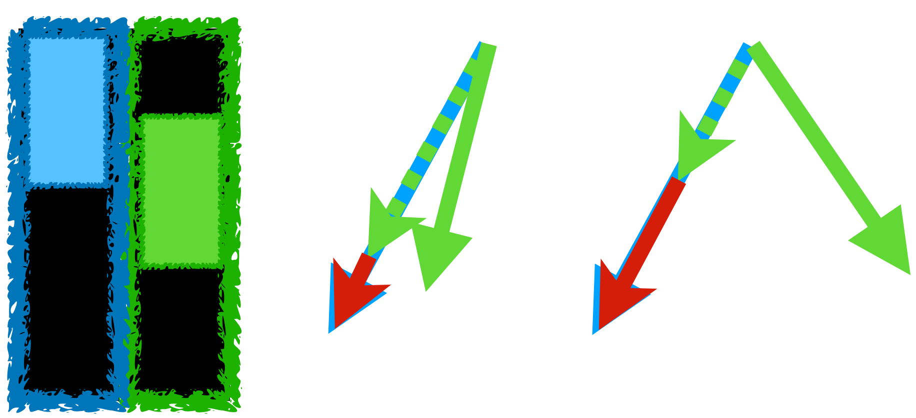


Noisy Data and Estimation Bound

$\sigma(B)$

Residual of projecting columns of B onto each other

Orientation of Projections $\mathbb{U}_i^\omega / \mathbb{V}_i^\omega$



Small

Large

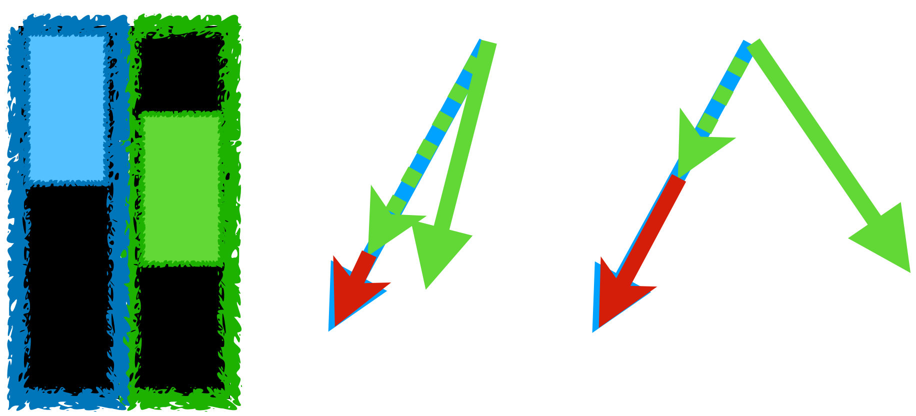
$$\mathbb{U}_i^\omega = \left[\begin{array}{c} \mathbf{I} \\ \hline \alpha_i^T \end{array} \right] \begin{array}{l} \}\, \mathrm{r} \\ \}\, 1 \end{array}$$

Noisy Data and Estimation Bound

$\sigma(B)$

Residual of projecting columns of B onto each other

Orientation of Projections $\mathbb{U}_i^{\omega} / \mathbb{V}_i^{\omega}$



Small

Large

$$\mathbb{U}_i^{\omega} = \begin{bmatrix} \mathbf{I} \\ \hline \alpha_i^T \end{bmatrix} \begin{matrix} \}\, r \\ \}\, 1 \end{matrix}$$

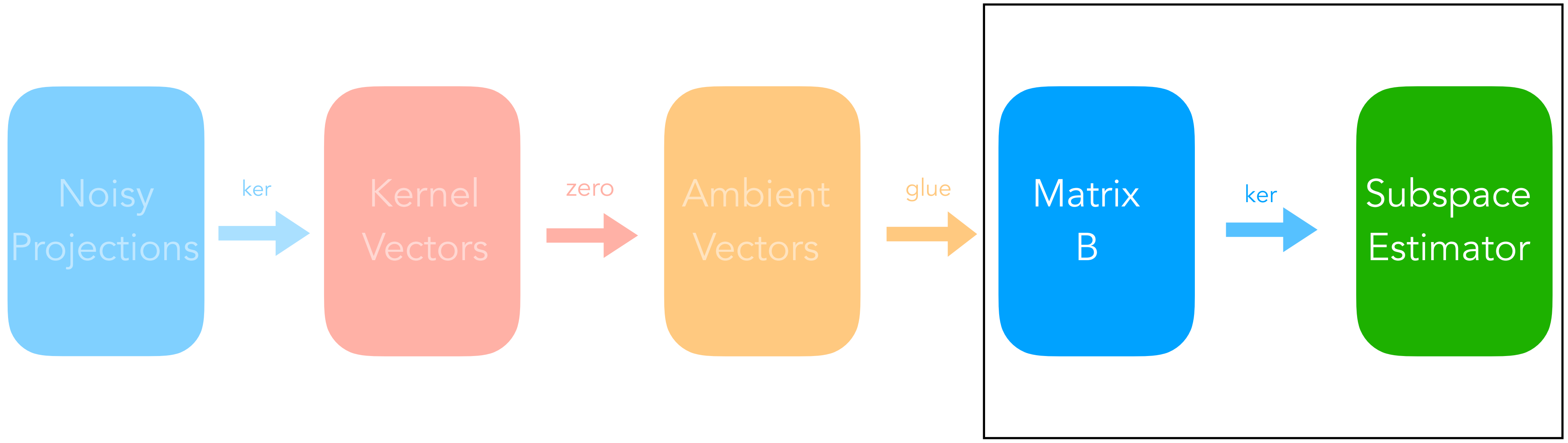
b_i is a perturbation of $[\alpha_i^T, -1]$

Noisy Data and Estimation Bound

Where does this come up?

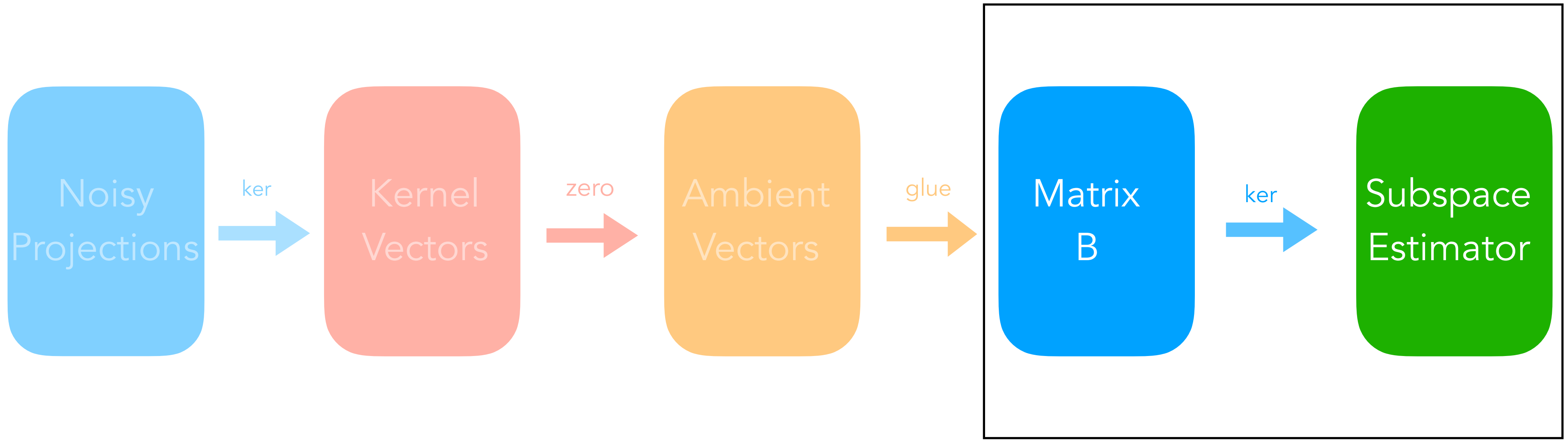
Noisy Data and Estimation Bound

Where does this come up?



Noisy Data and Estimation Bound

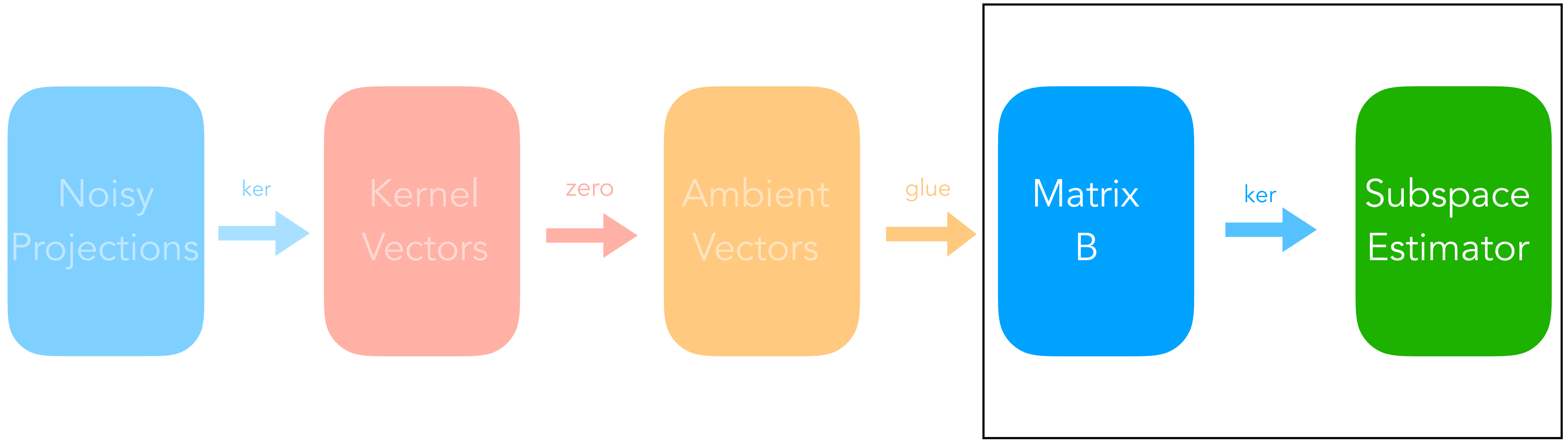
Where does this come up?



Here, we are taking an orthogonal space

Noisy Data and Estimation Bound

Where does this come up?



Here, we are taking an orthogonal space
How sensitive is this to perturbations?

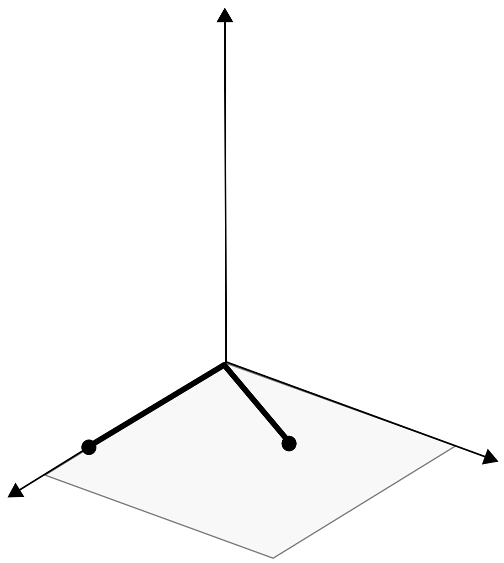
Noisy Data and Estimation Bound

How much a subspace gets perturbed depends on how singular its basis is.

Noisy Data and Estimation Bound

How much a subspace gets perturbed depends on how singular its basis is.

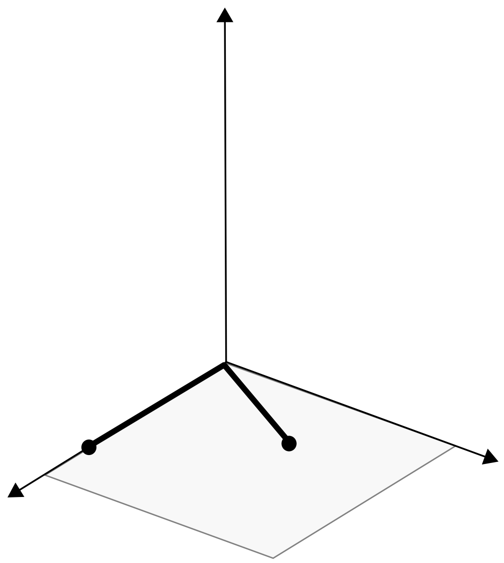
Non-singular Basis

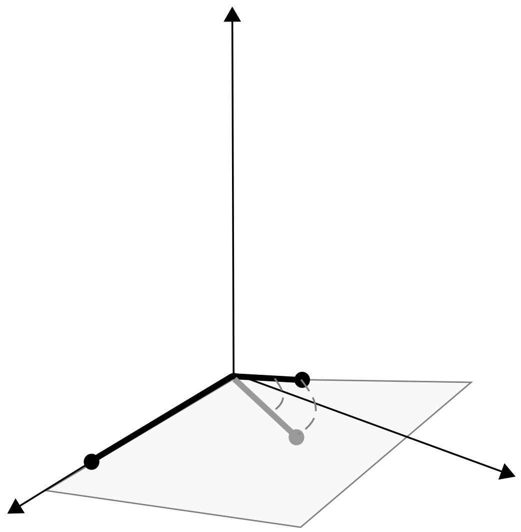


Noisy Data and Estimation Bound

How much a subspace gets perturbed depends on how singular its basis is.

Non-singular Basis



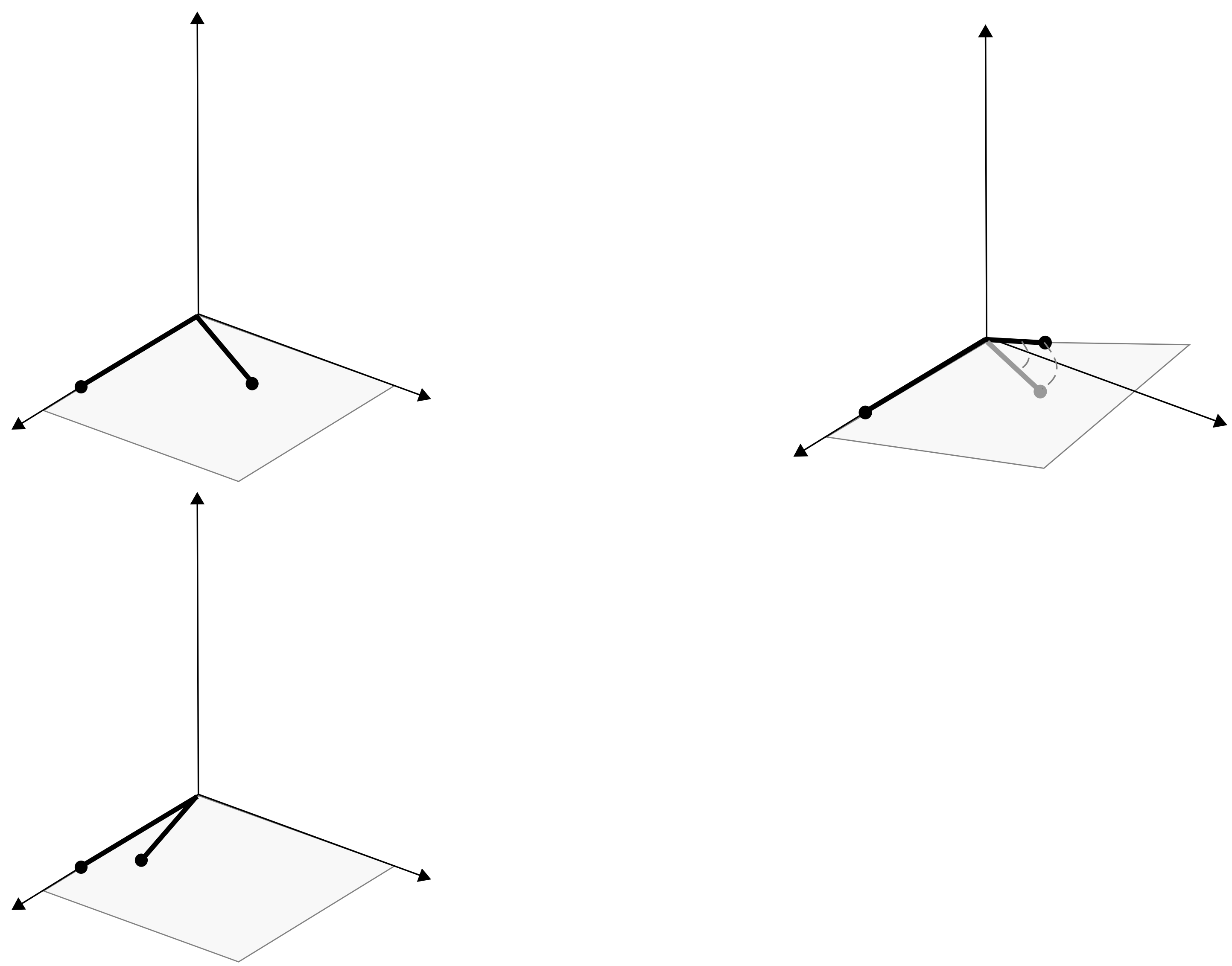


Noisy Data and Estimation Bound

How much a subspace gets perturbed depends on how singular its basis is.

Non-singular Basis

Singular Basis

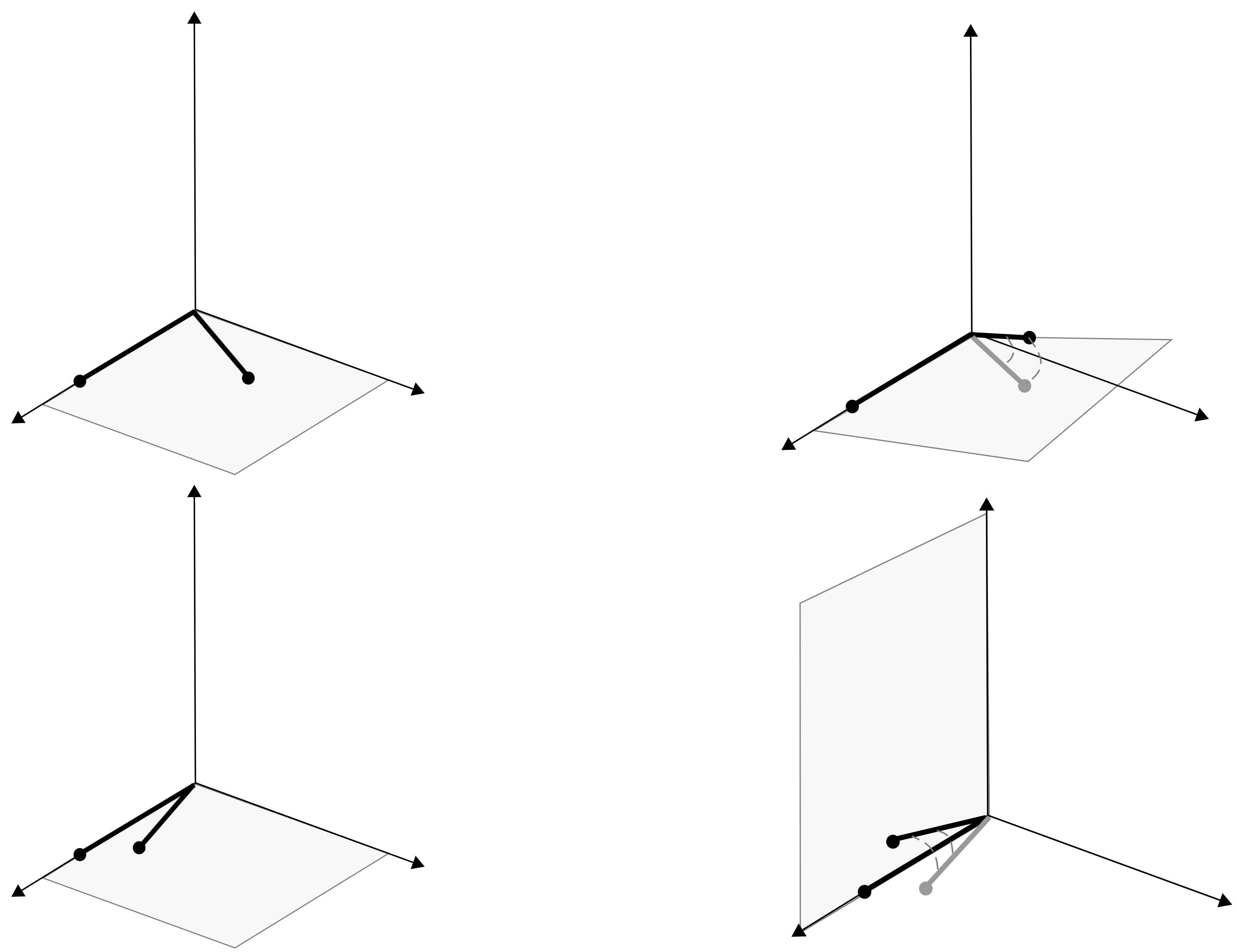


Noisy Data and Estimation Bound

How much a subspace gets perturbed depends on how singular its basis is.

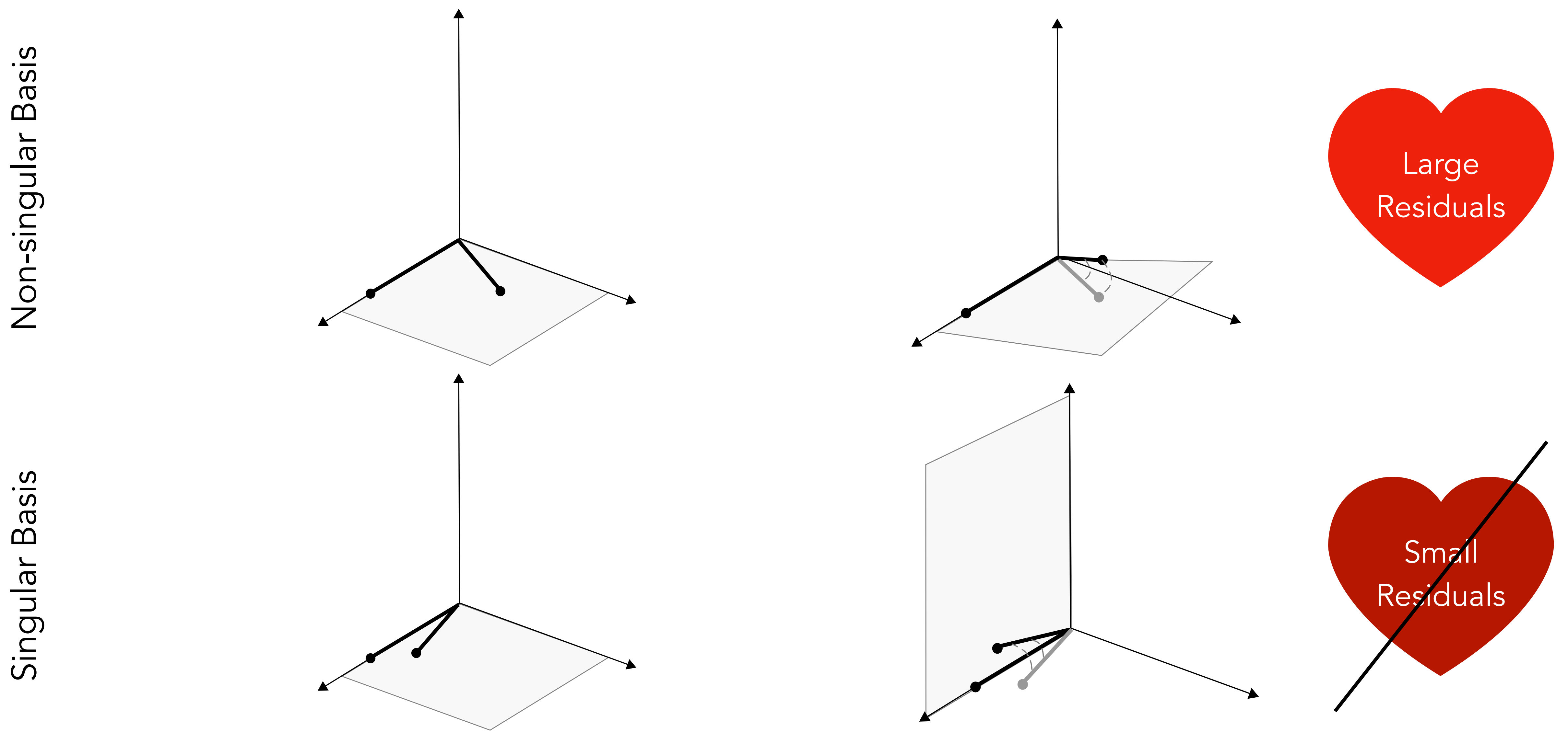
Non-singular Basis

Singular Basis



Noisy Data and Estimation Bound

How much a subspace gets perturbed depends on how singular its basis is.



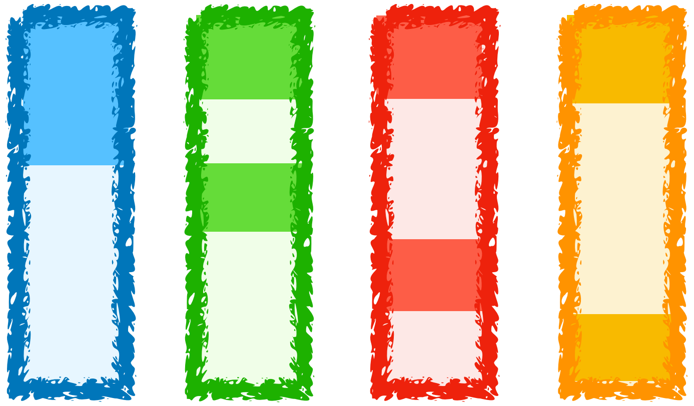
Noisy Data and Estimation Bound

Bottom Line: Different Samplings can lead to different bounds.

Noisy Data and Estimation Bound

Bottom Line: Different Samplings can lead to different bounds.

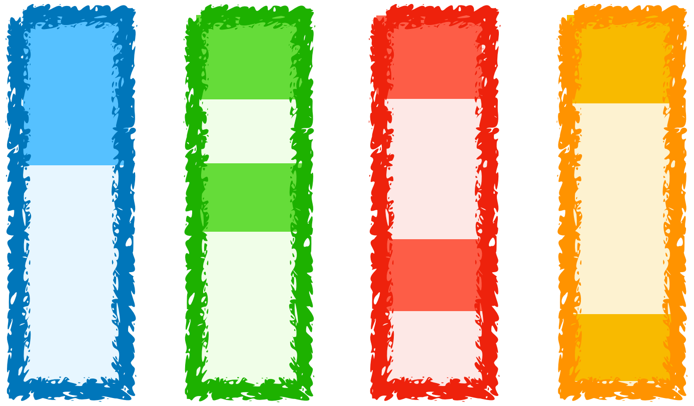
$$\Omega_1 =$$



Noisy Data and Estimation Bound

Bottom Line: Different Samplings can lead to different bounds.

$$\Omega_1 =$$

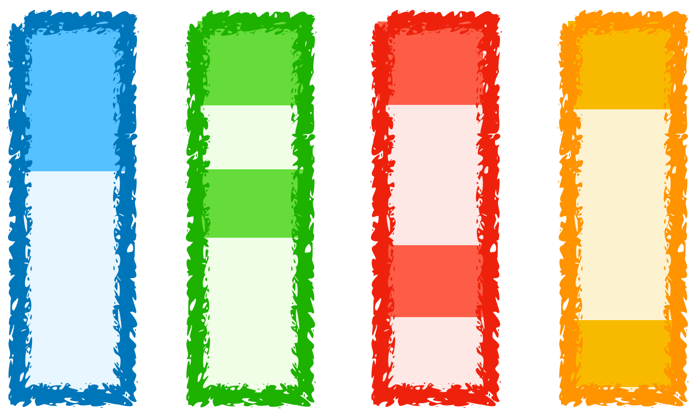


- More overlaps, but
- Each column has a unique coordinate
- Less likely to have small residuals

Noisy Data and Estimation Bound

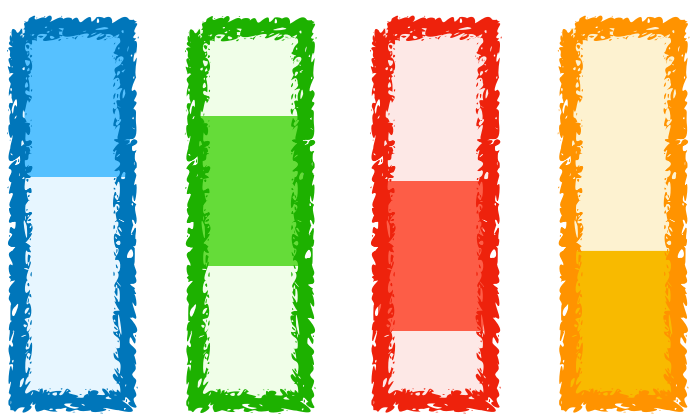
Bottom Line: Different Samplings can lead to different bounds.

$\boldsymbol{\Omega}_1 =$



- More overlaps, but
- Each column has a unique coordinate
- Less likely to have small residuals

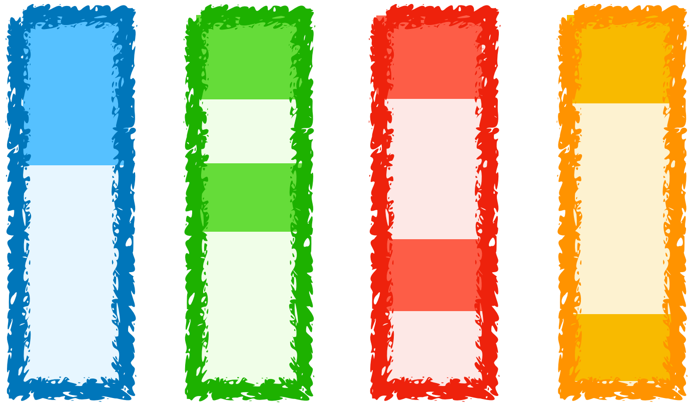
$\boldsymbol{\Omega}_2 =$



Noisy Data and Estimation Bound

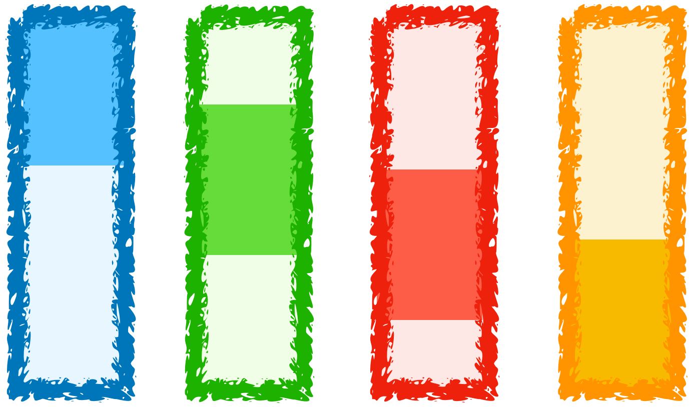
Bottom Line: Different Samplings can lead to different bounds.

$\boldsymbol{\Omega}_1 =$



- More overlaps, but
- Each column has a unique coordinate
- Less likely to have small residuals

$\boldsymbol{\Omega}_2 =$

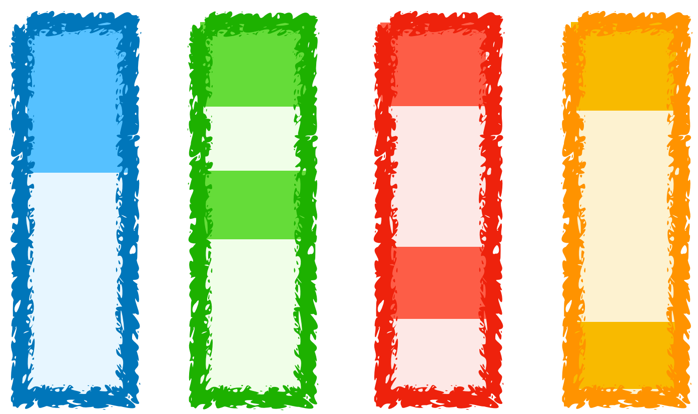


- Fewer overlaps, but
- Most columns don't have unique coordinates
- More likely to have small residuals (than sampling 1)

Noisy Data and Estimation Bound

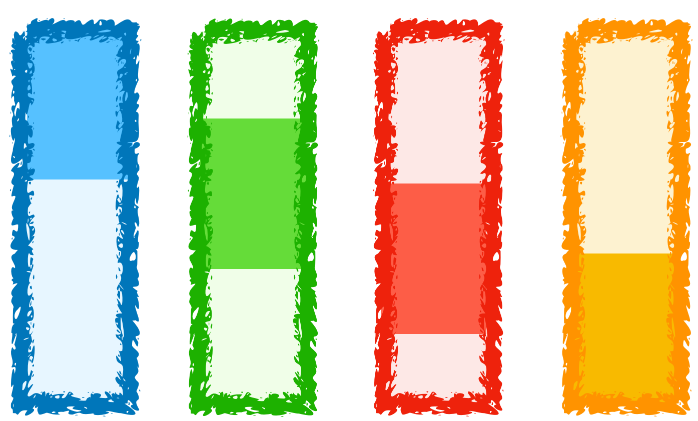
Bottom Line: Different Samplings can lead to different bounds.

$\mathbf{\Omega}_1 =$



- More overlaps, but
- Each column has a unique coordinate
- Less likely to have small residuals

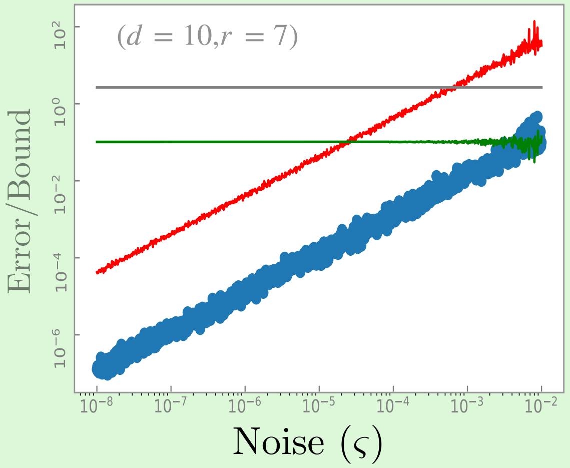
$\mathbf{\Omega}_2 =$



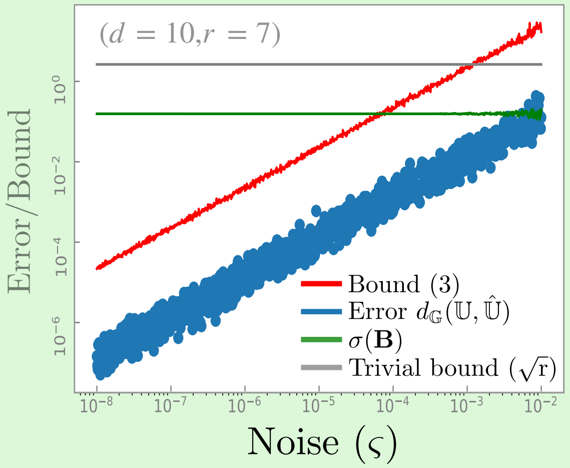
- Fewer overlaps, but
- Most columns don't have unique coordinates
- More likely to have small residuals (than sampling 1)

Noisy Data and Estimation Bound

Sampling $\mathbf{\Omega}_1$



Sampling $\mathbf{\Omega}_2$



Low dimensions

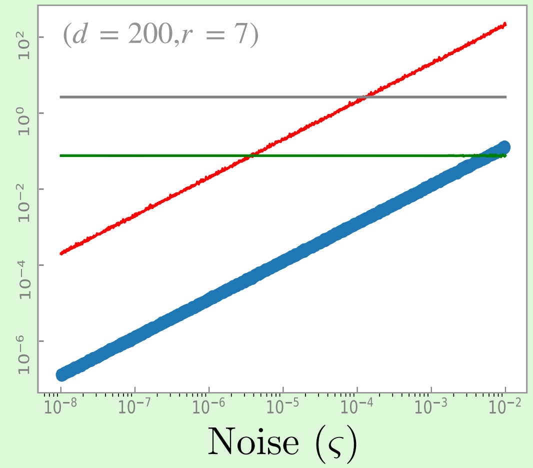
+

Varying Noise

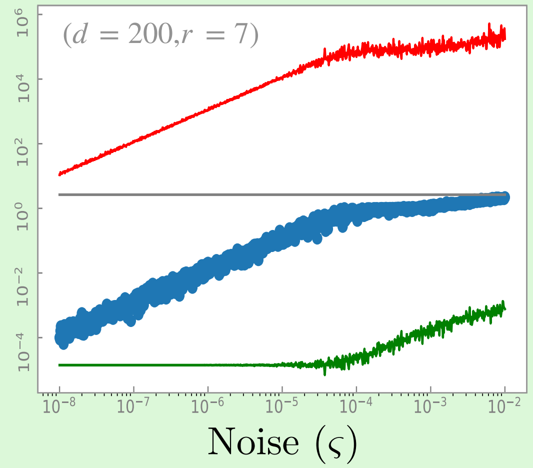
Bound tight - ✓

Noisy Data and Estimation Bound

Sampling $\boldsymbol{\Omega}_1$



Sampling $\boldsymbol{\Omega}_2$



High dimensions
+
Varying Noise

Bound tight - ✓

Depends on Sampling Pattern

Noisy Data and Estimation Bound

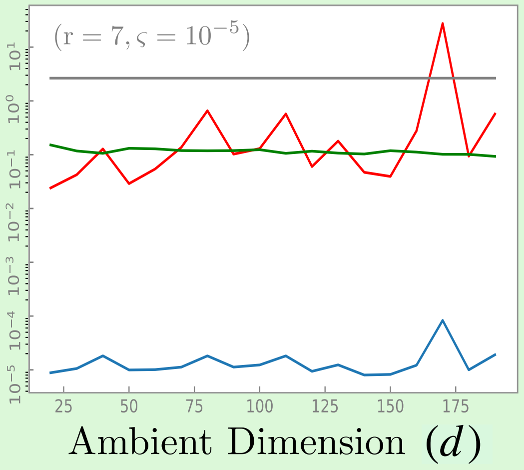
Sampling $\boldsymbol{\Omega}_1$

Error/Bound

Low Subspace Rank
+
Varying Ambient Dim

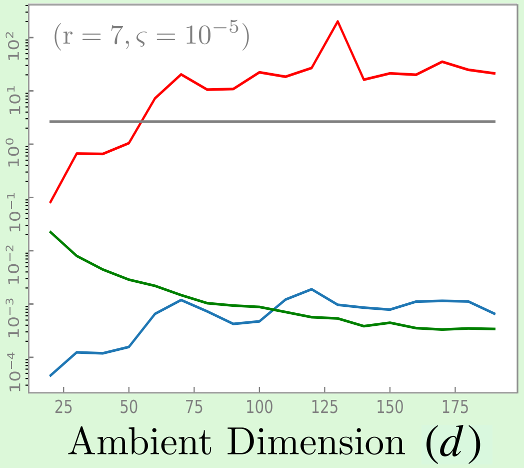
Bound tight - ✓

Depends on sampling.
Good sampling shows bound follow error



Sampling $\boldsymbol{\Omega}_2$

Error/Bound



Noisy Data and Estimation Bound

Sampling $\mathbf{\Omega}_1$

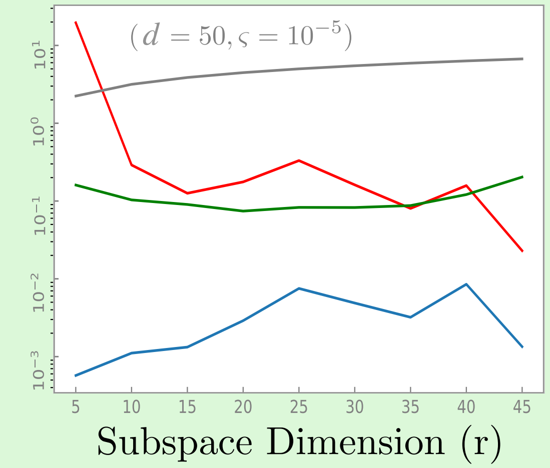
Error/Bound

High ambient dim

+

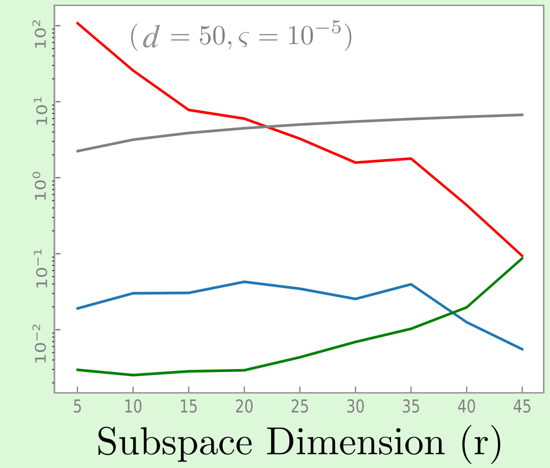
Varying Subspace Dim

Bound tight - ✔

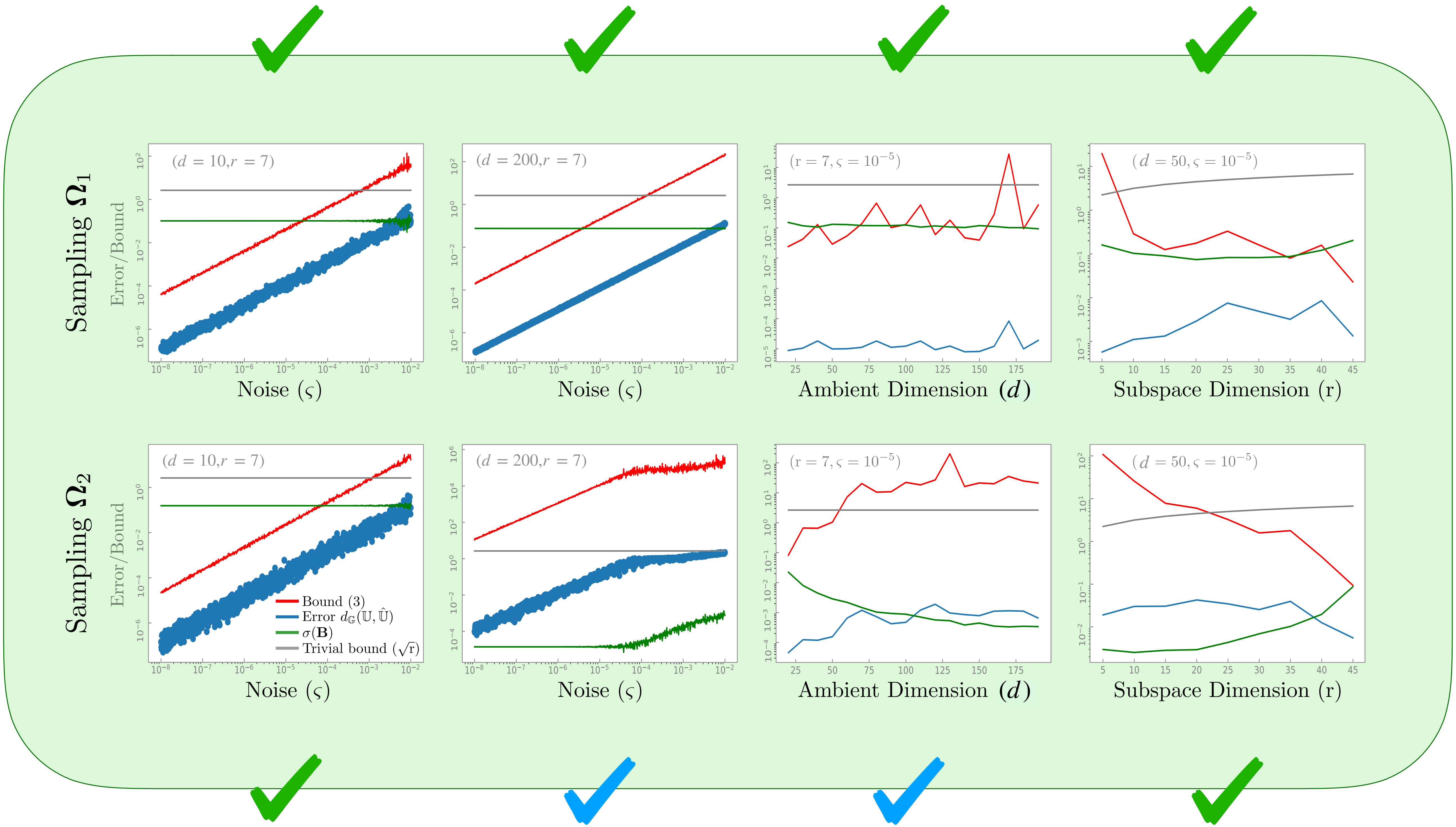


Sampling $\mathbf{\Omega}_2$

Error/Bound



Noisy Data and Estimation Bound



Outline

1. Problem Setup - Subspace Estimation
2. Motivation - Missing Data
3. Previous Work: Noiseless case
4. This Paper - Noisy Data and Estimation Bound
5. Applications
6. Conclusions

Applications - LRMC Theory

Subspace reconstruction has shed light on various LRMC!

Applications - LRMC Theory

Subspace reconstruction has shed light on various LRMC!

❶ Deterministic sampling conditions for unique completability in LRMC (Theorem 2, Lemma 8 in [2]).
❷ The information-theoretic requirements and sample complexity of HRMC (Theorems 1, 2 in [3]).
❸ The fundamental conditions for learning mixtures in MMC (Theorem 1 in [4]).
❹ Identifiability conditions to learn tensorized subspaces in LADMC (Theorem 1 in [5], Lemmas 2, 3 in [6]).
❻ Unique completability conditions for LTRTC (Lemma 4, Theorem 4 in [7], Lemma 9, Theorem 7 in [8]).
❼ Deterministic conditions for unique completability in LCRTC (Lemma 18 in [9]).

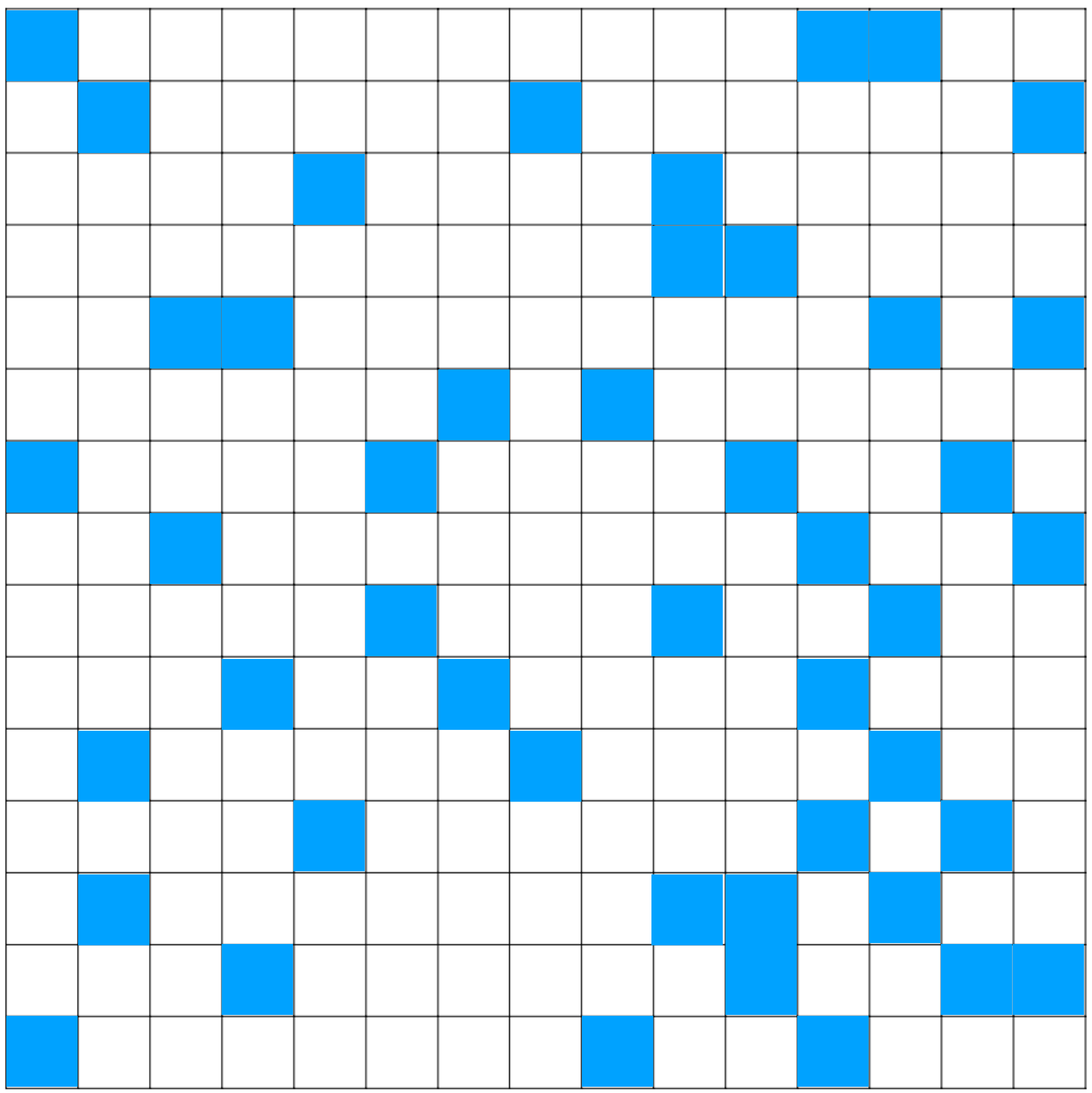
Applications - LRMC Theory

Subspace reconstruction has shed light on various LRMC!

❶ Deterministic sampling conditions for unique completability in LRMC (Theorem 2, Lemma 8 in [2]).
❷ The information-theoretic requirements and sample complexity of HRMC (Theorems 1, 2 in [3]).
❸ The fundamental conditions for learning mixtures in MMC (Theorem 1 in [4]).
❹ Identifiability conditions to learn tensorized subspaces in LADMC (Theorem 1 in [5], Lemmas 2, 3 in [6]).
❻ Unique completability conditions for LTRTC (Lemma 4, Theorem 4 in [7], Lemma 9, Theorem 7 in [8]).
❼ Deterministic conditions for unique completability in LCRTC (Lemma 18 in [9]).

Project incomplete data onto candidate subspaces; reject incompatible spaces.

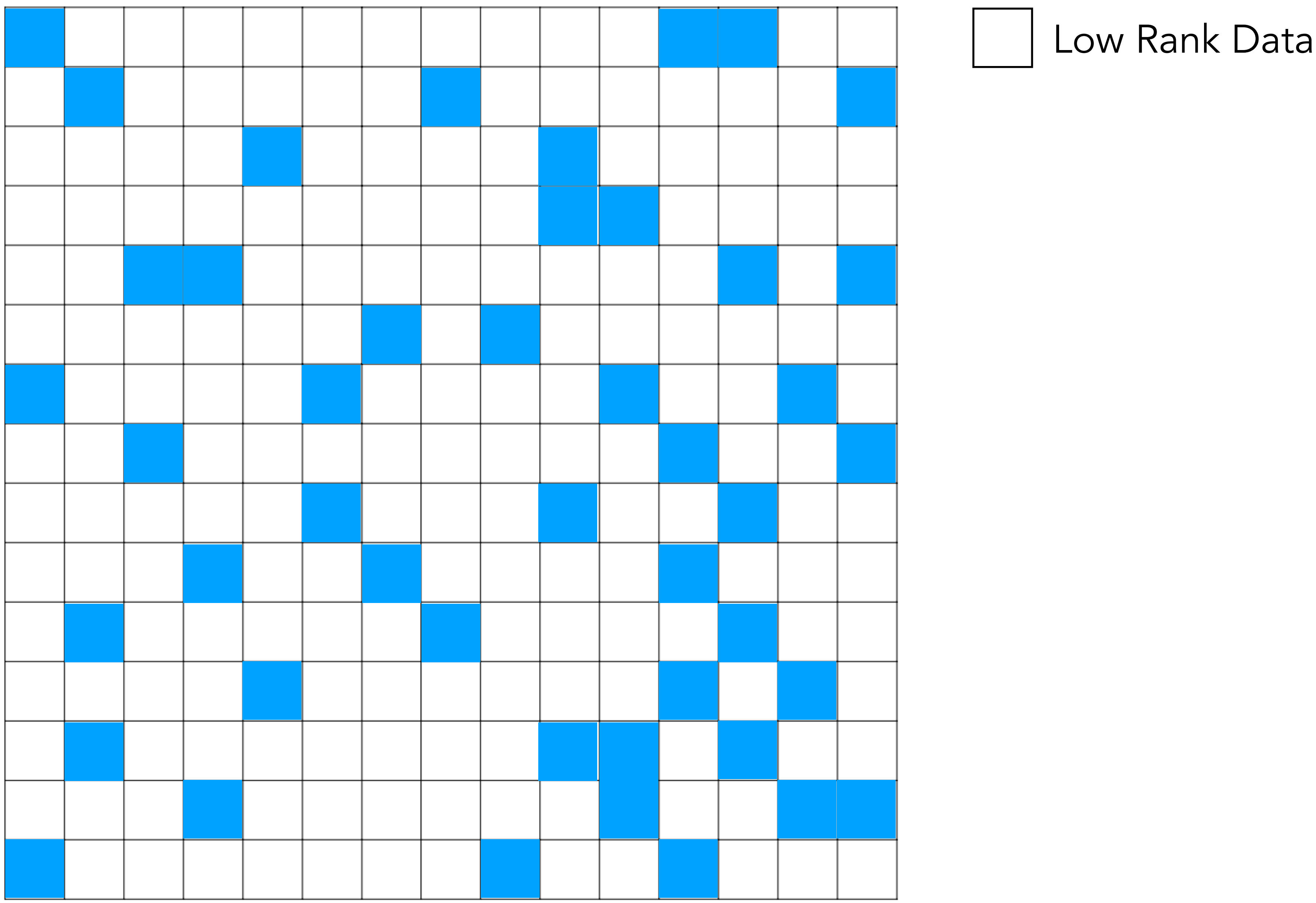
Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

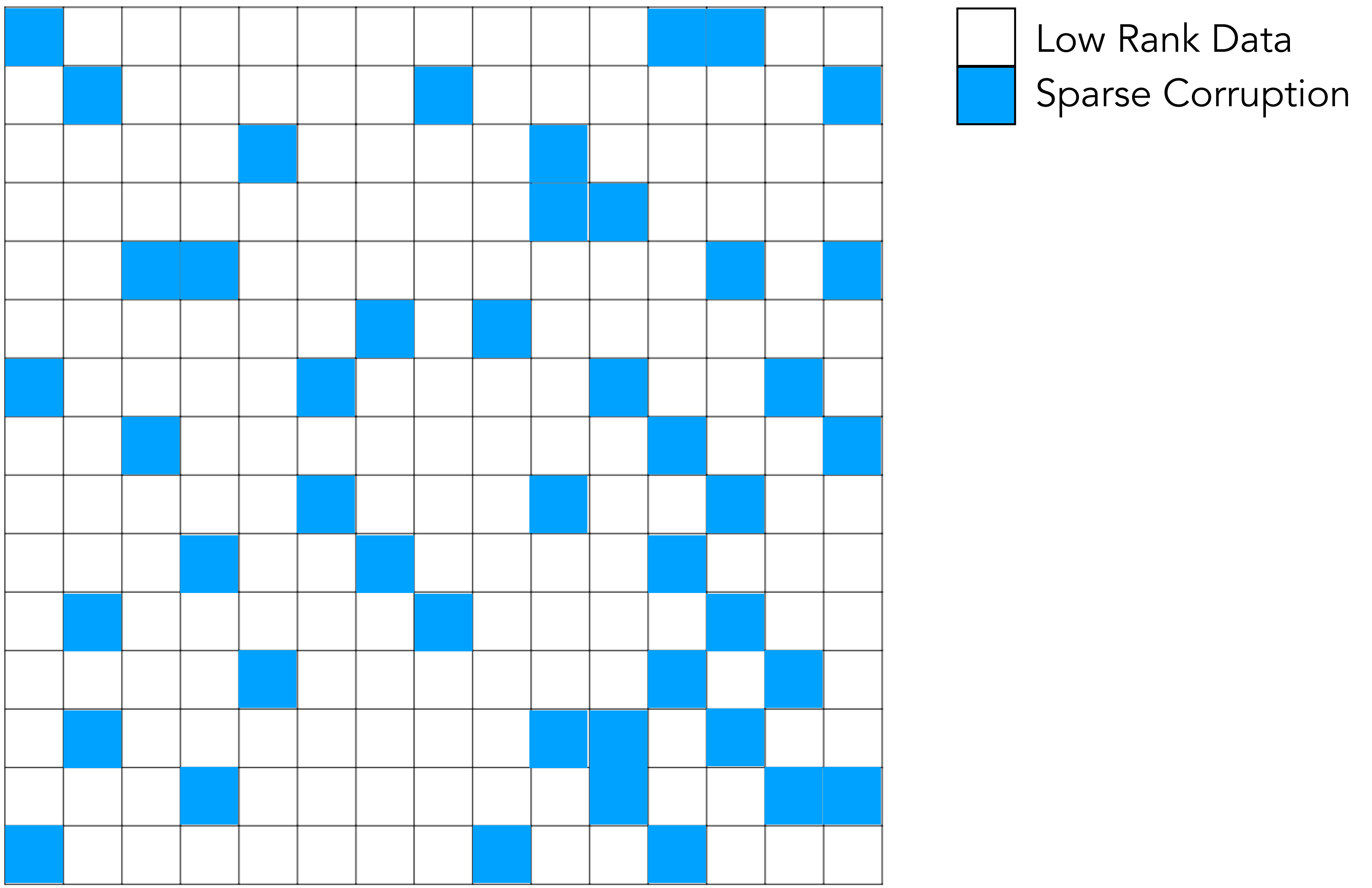
Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

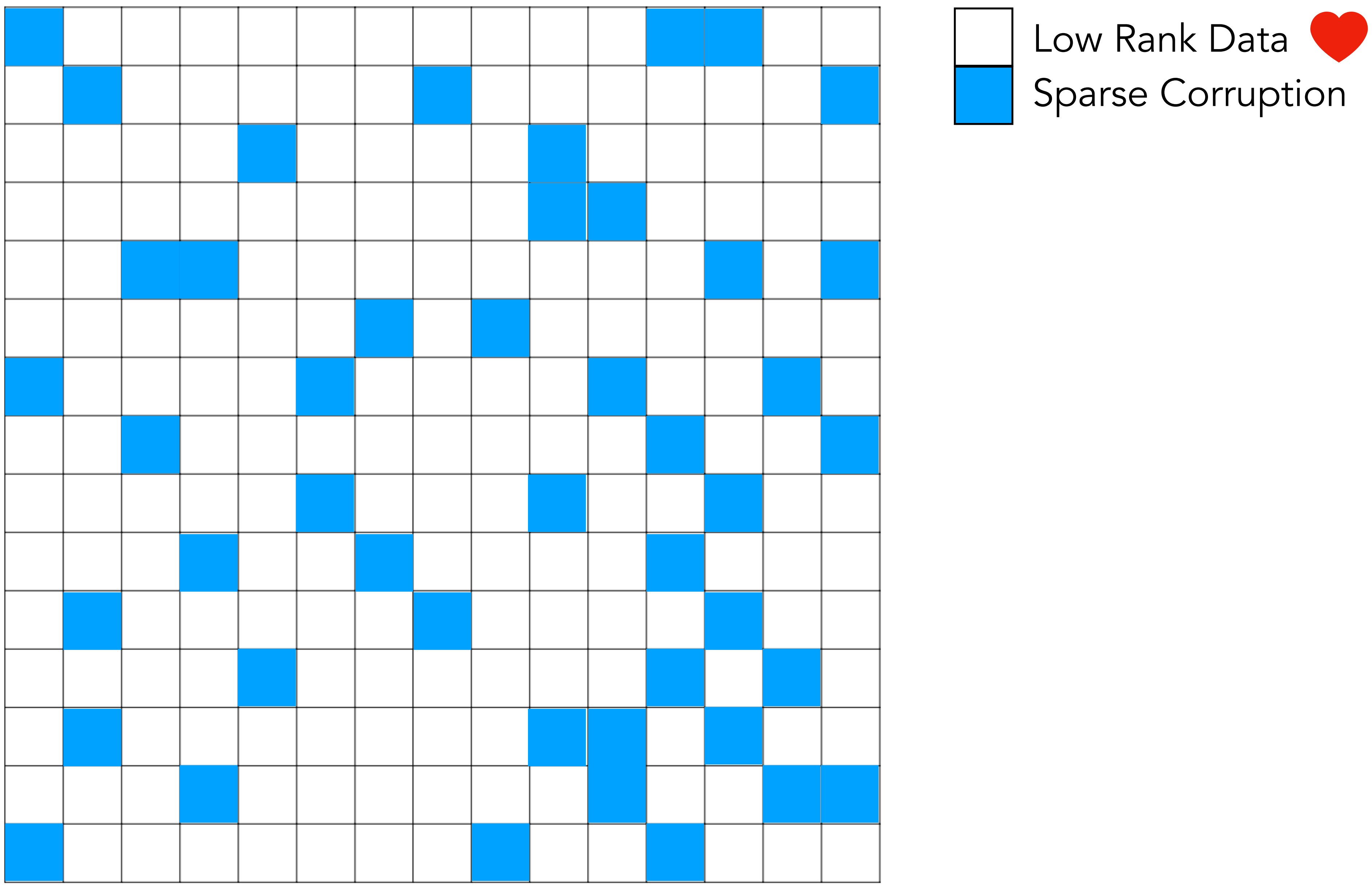
Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

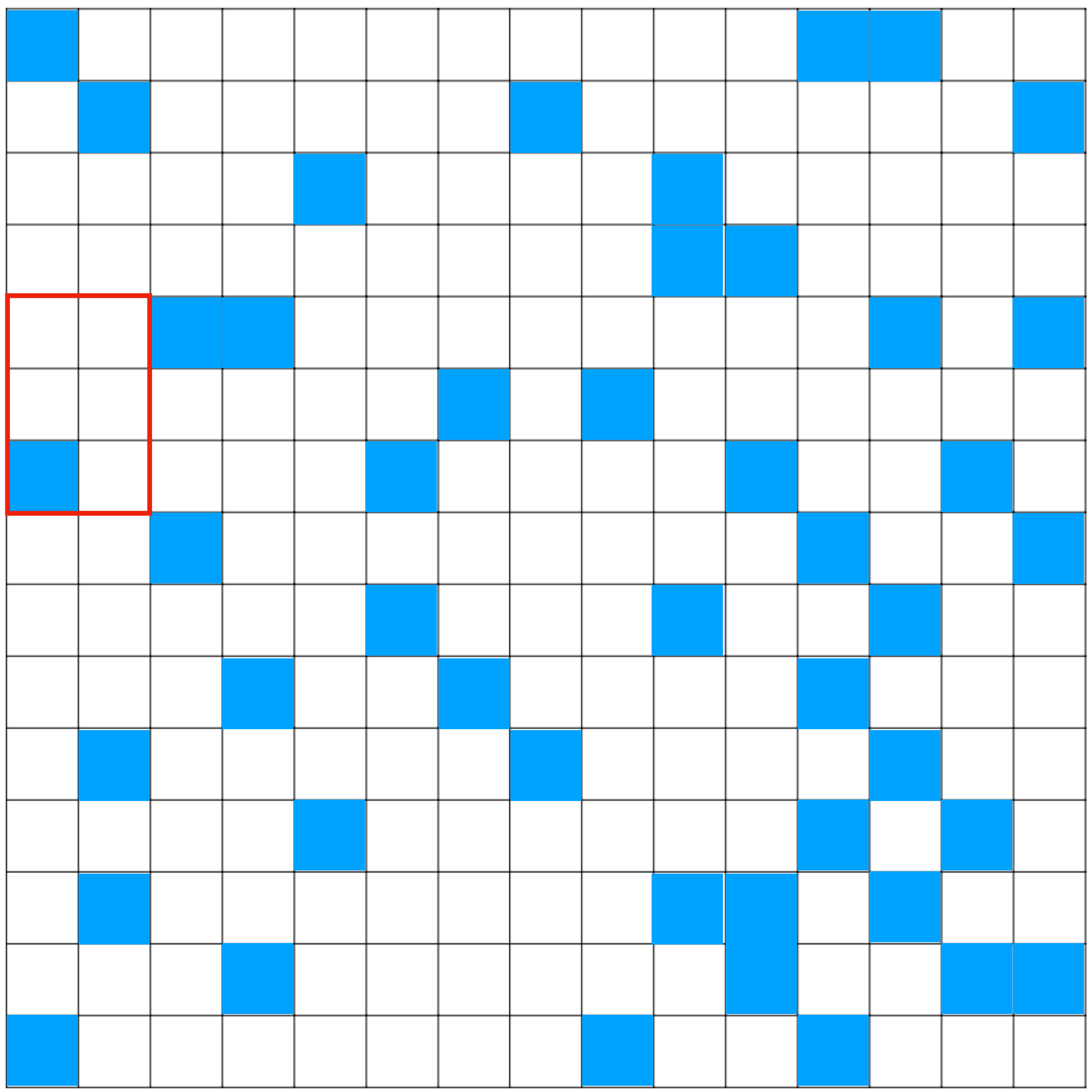
Applications - RPCA Algorithms

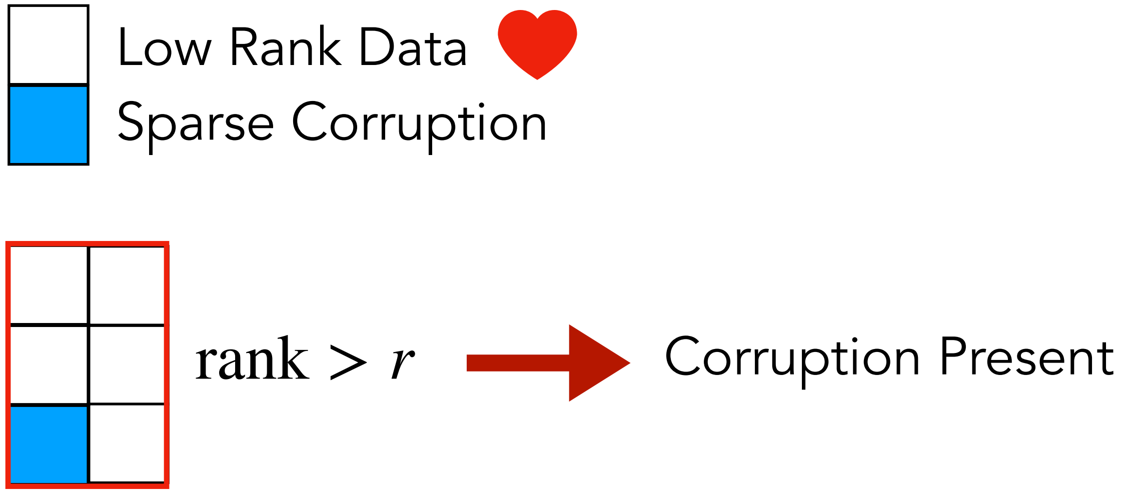


Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Applications - RPCA Algorithms

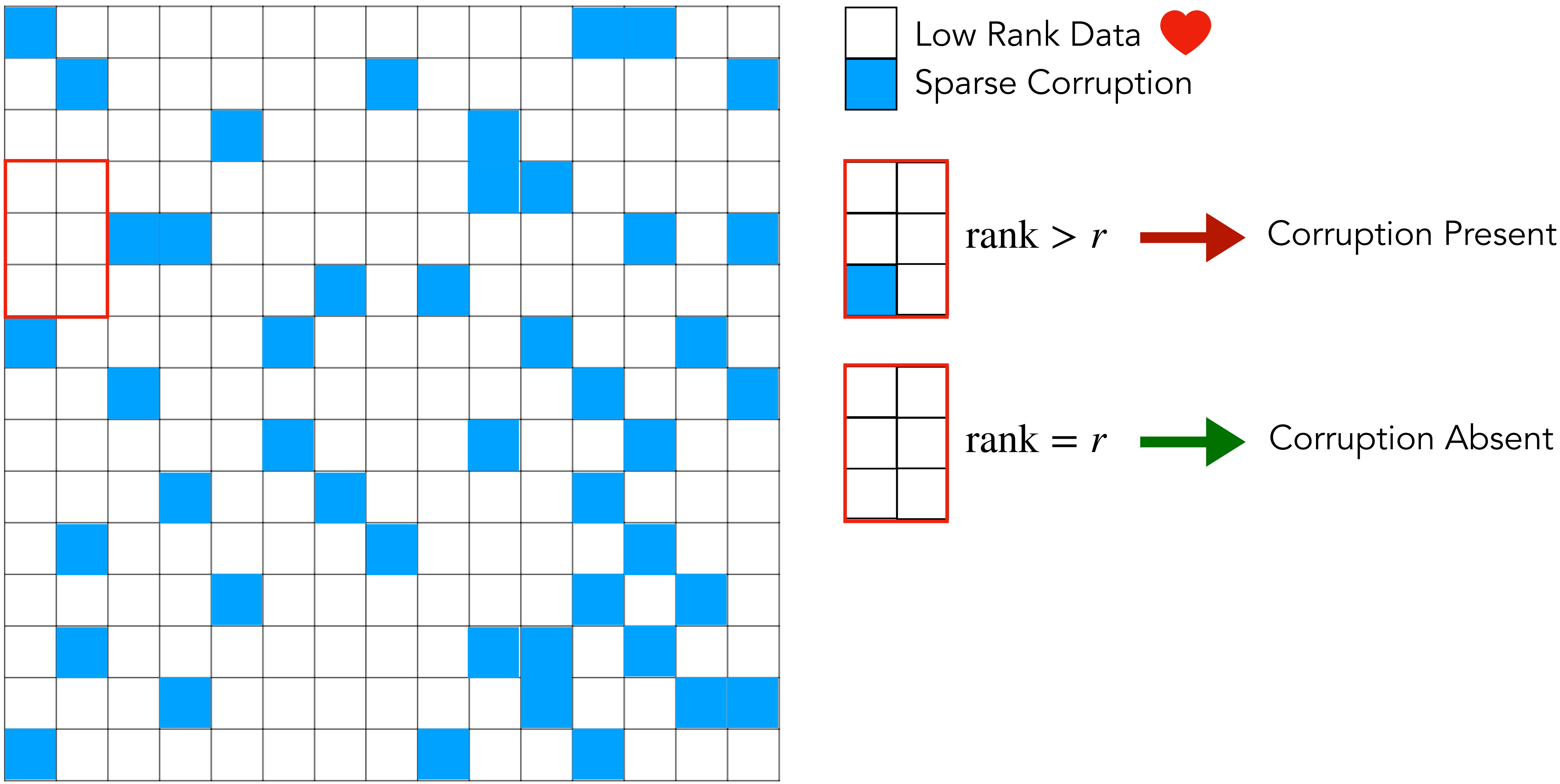




Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

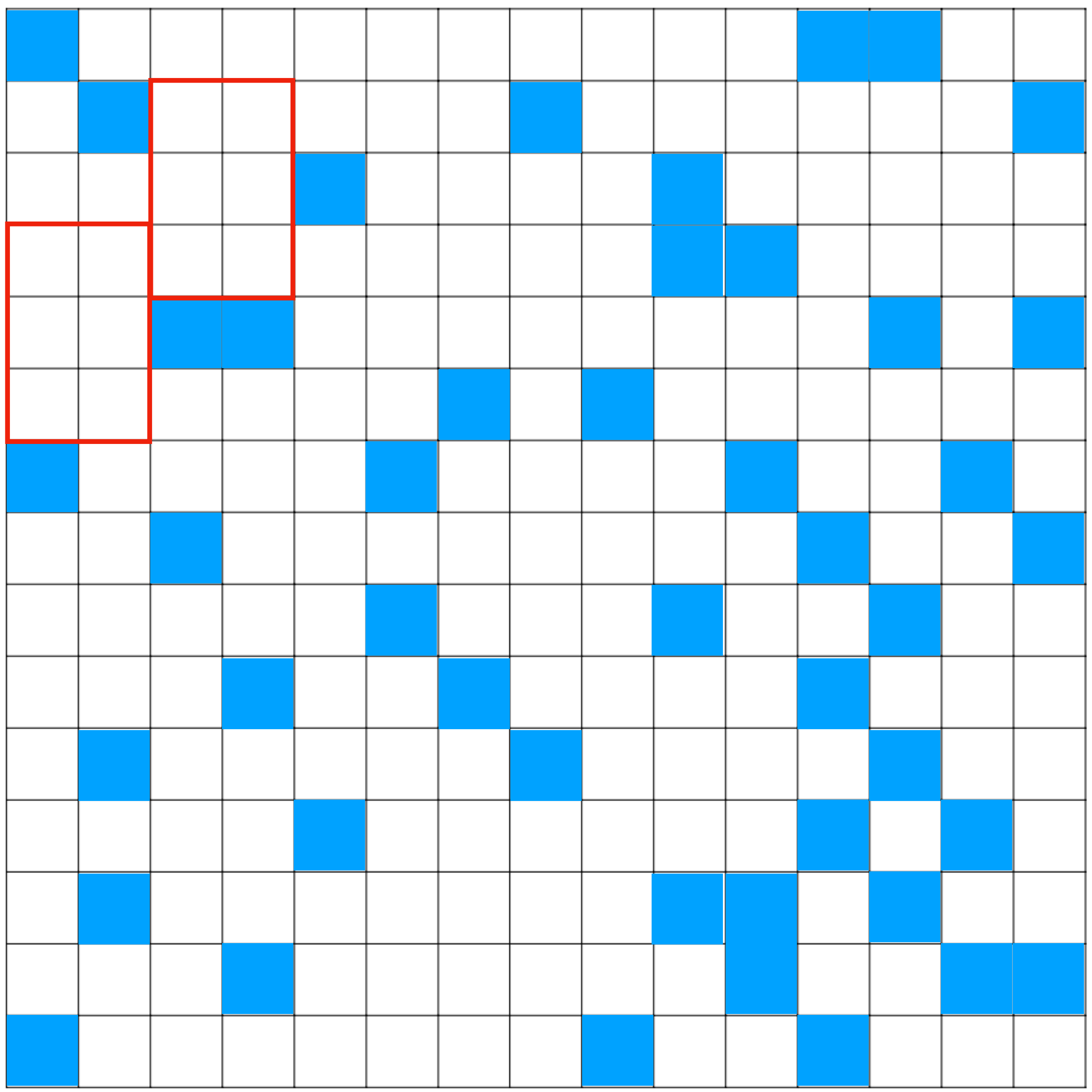
Applications - RPCA Algorithms

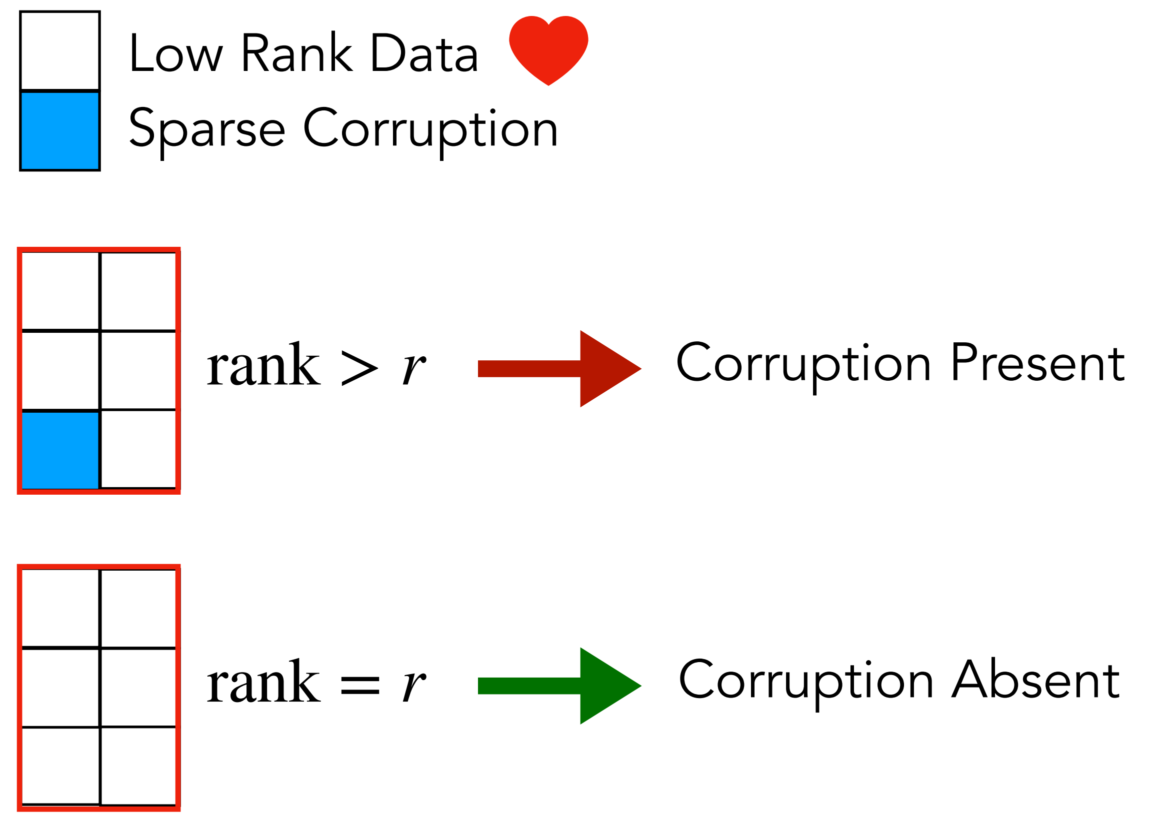


Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Applications - RPCA Algorithms

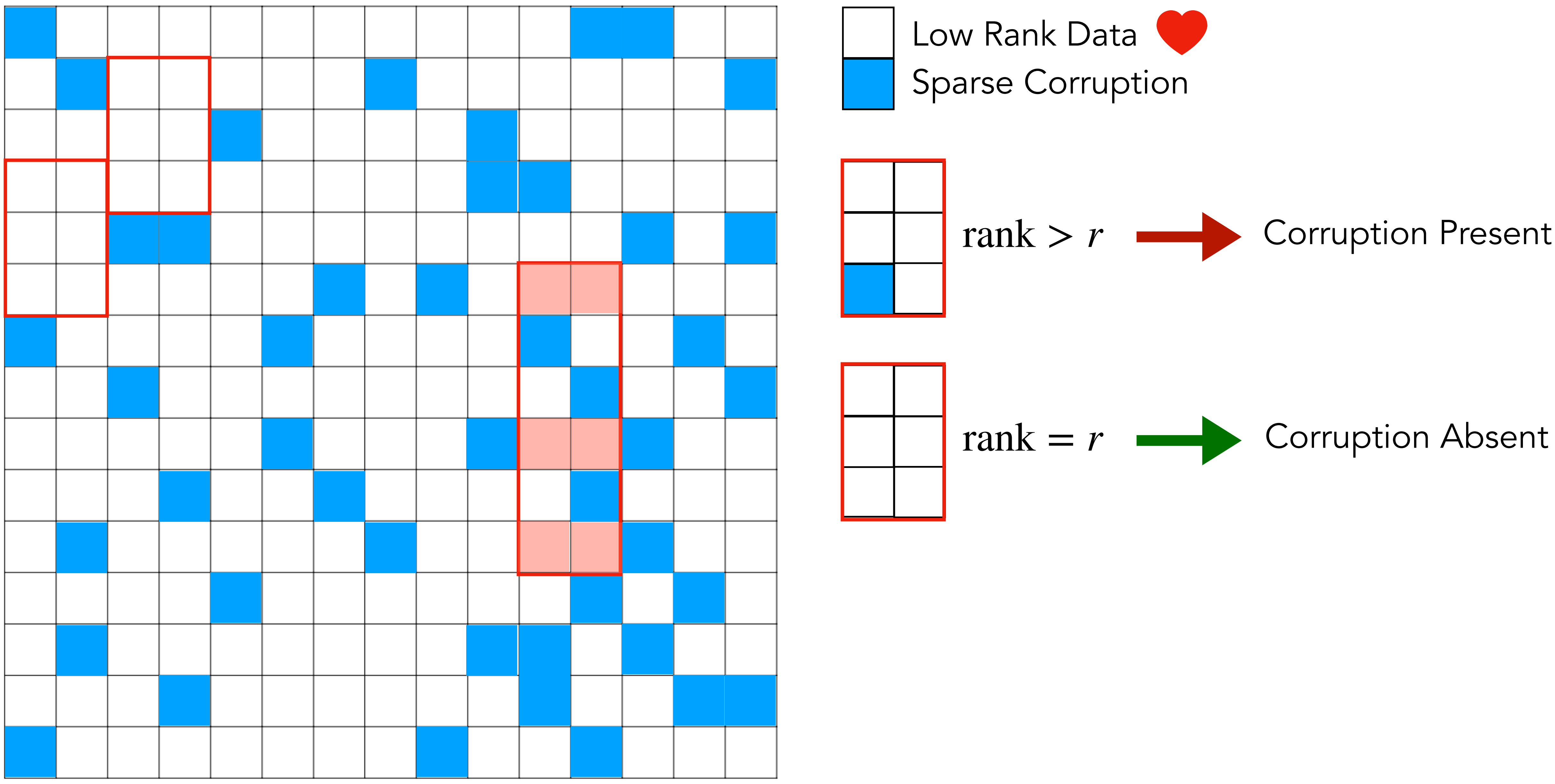




Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

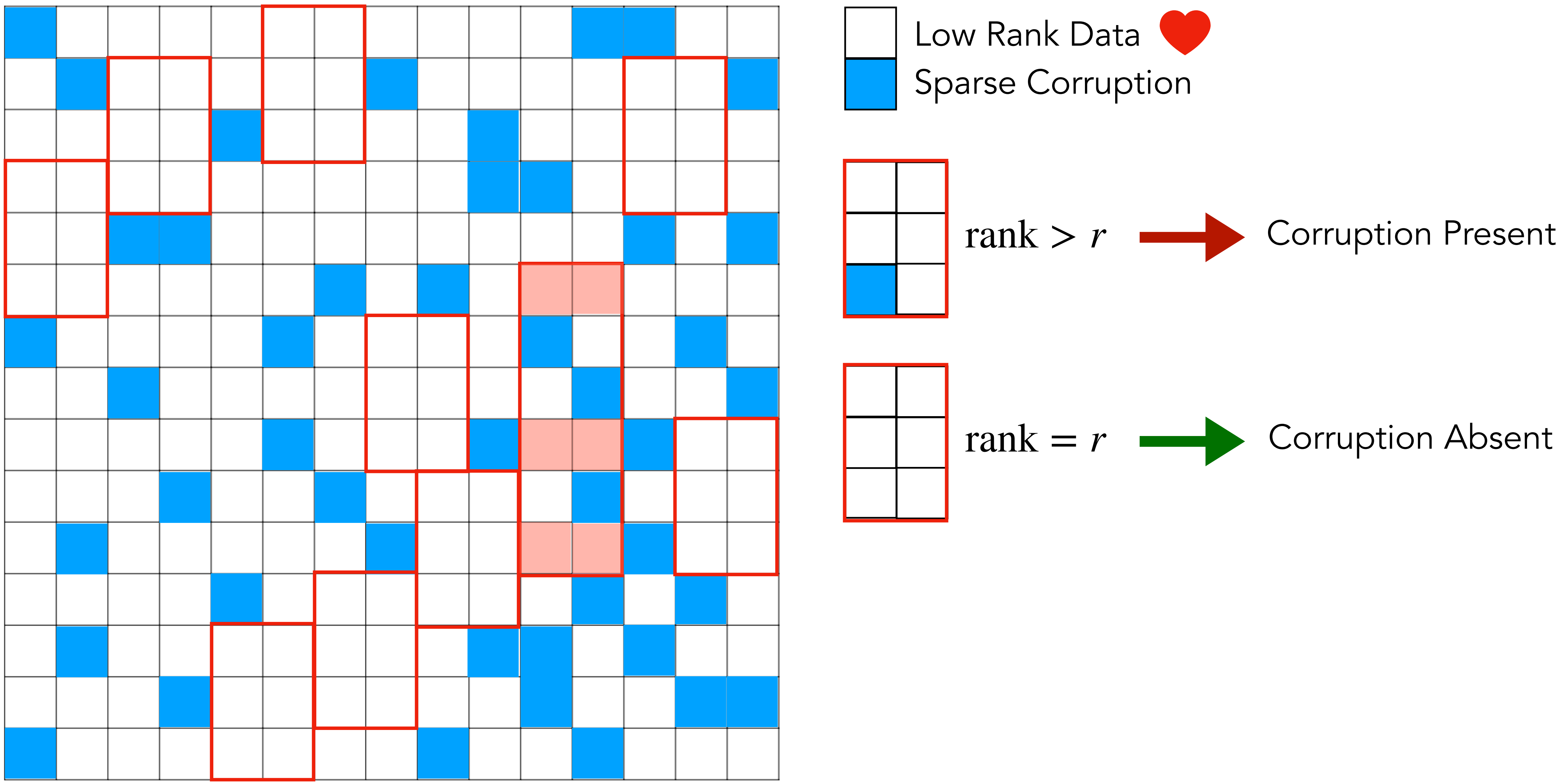
Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

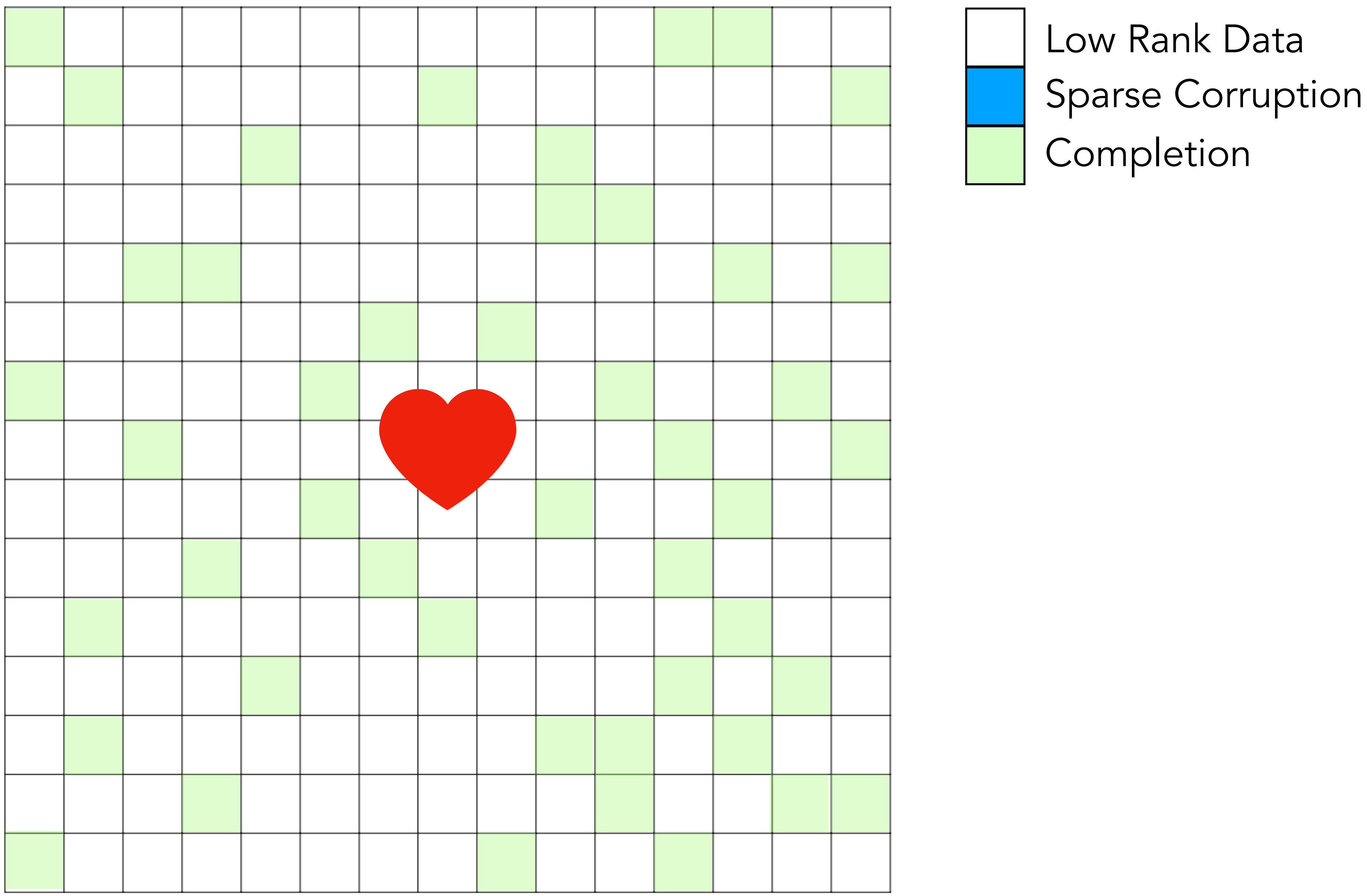
Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

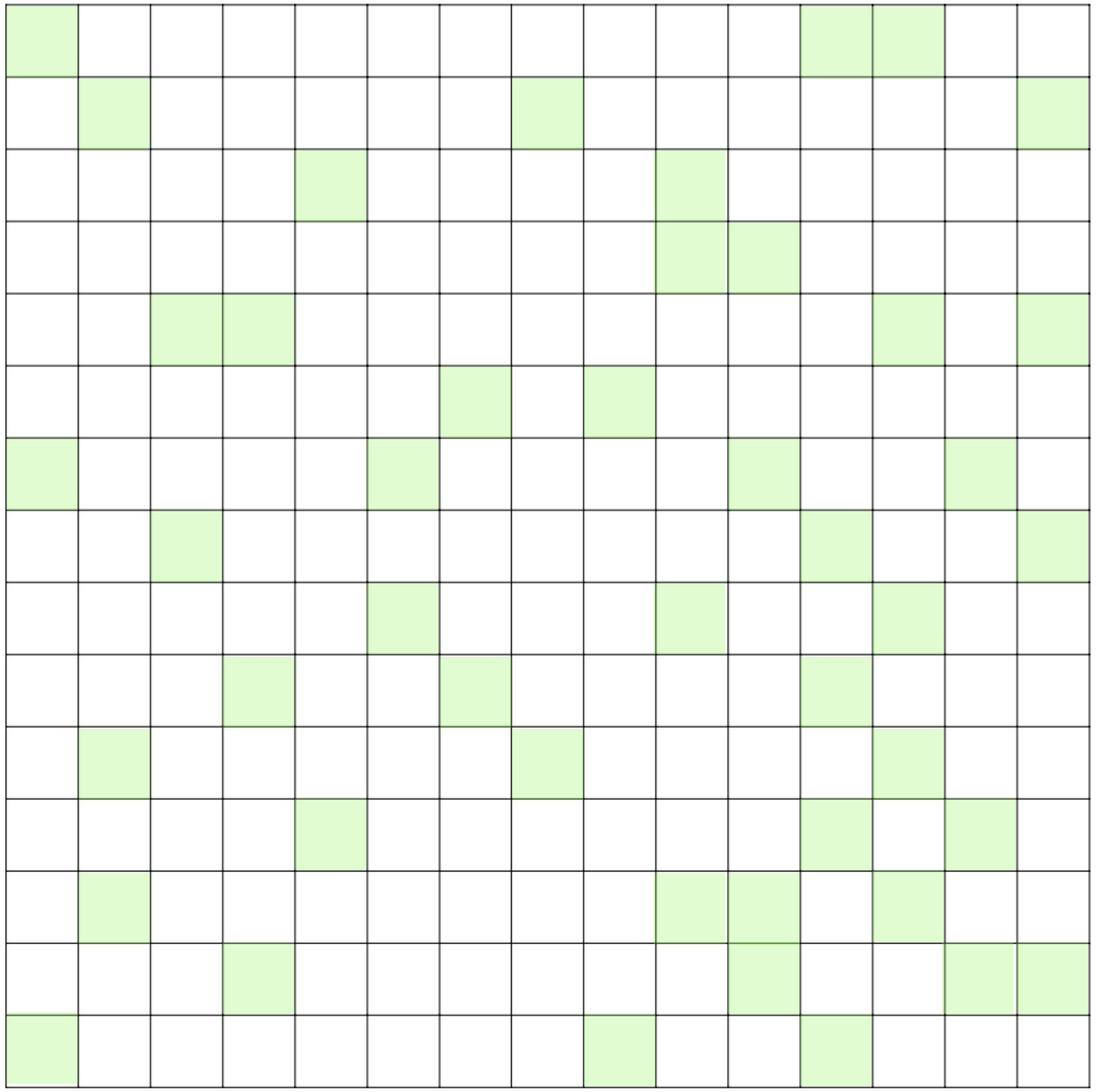
Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Applications - RPCA Algorithms



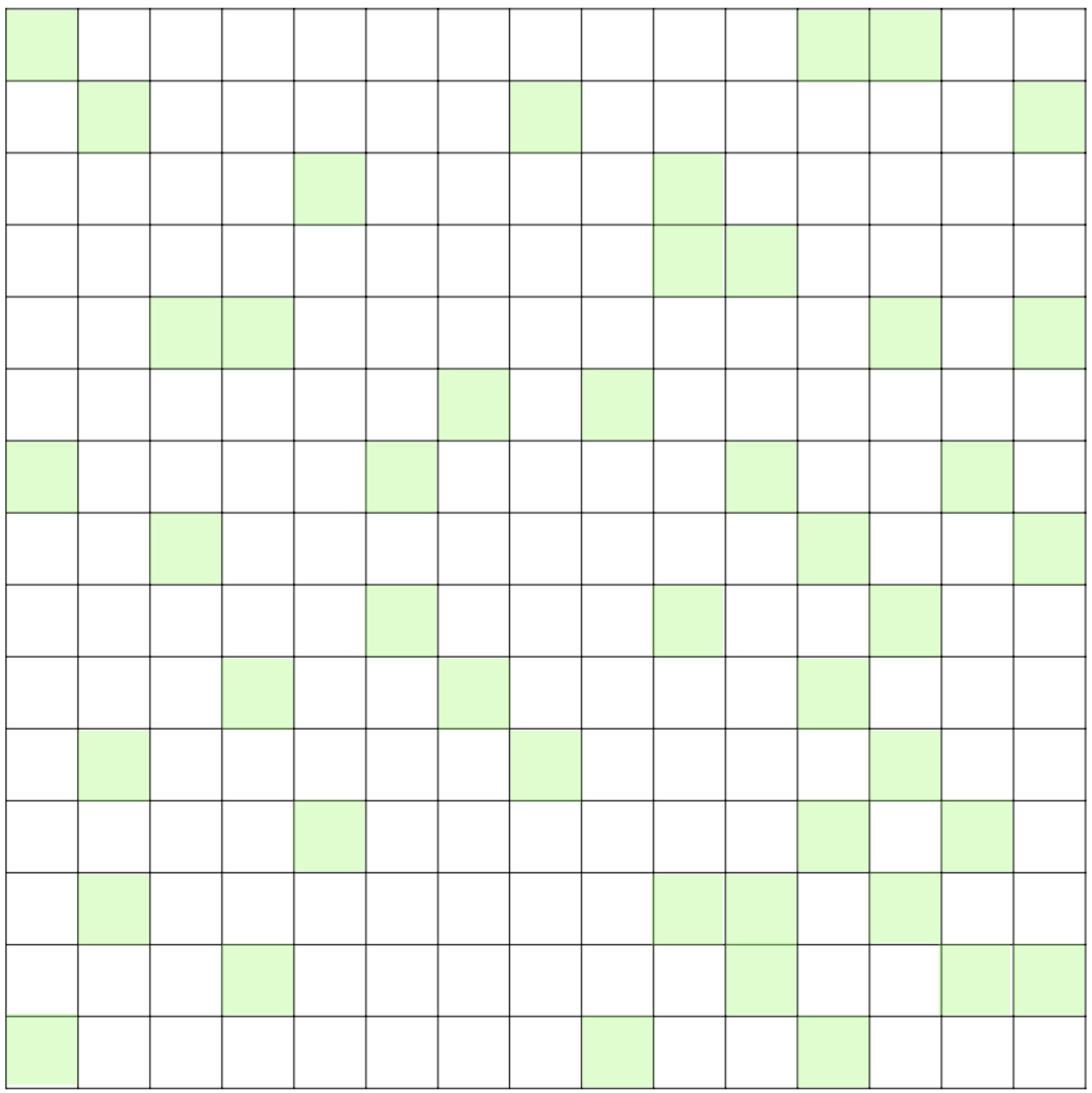
Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Standard RPCA Methods:

- Use Optimization approaches
- Require coherence / incoherence in data
- Work with high probability

Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Standard RPCA Methods:

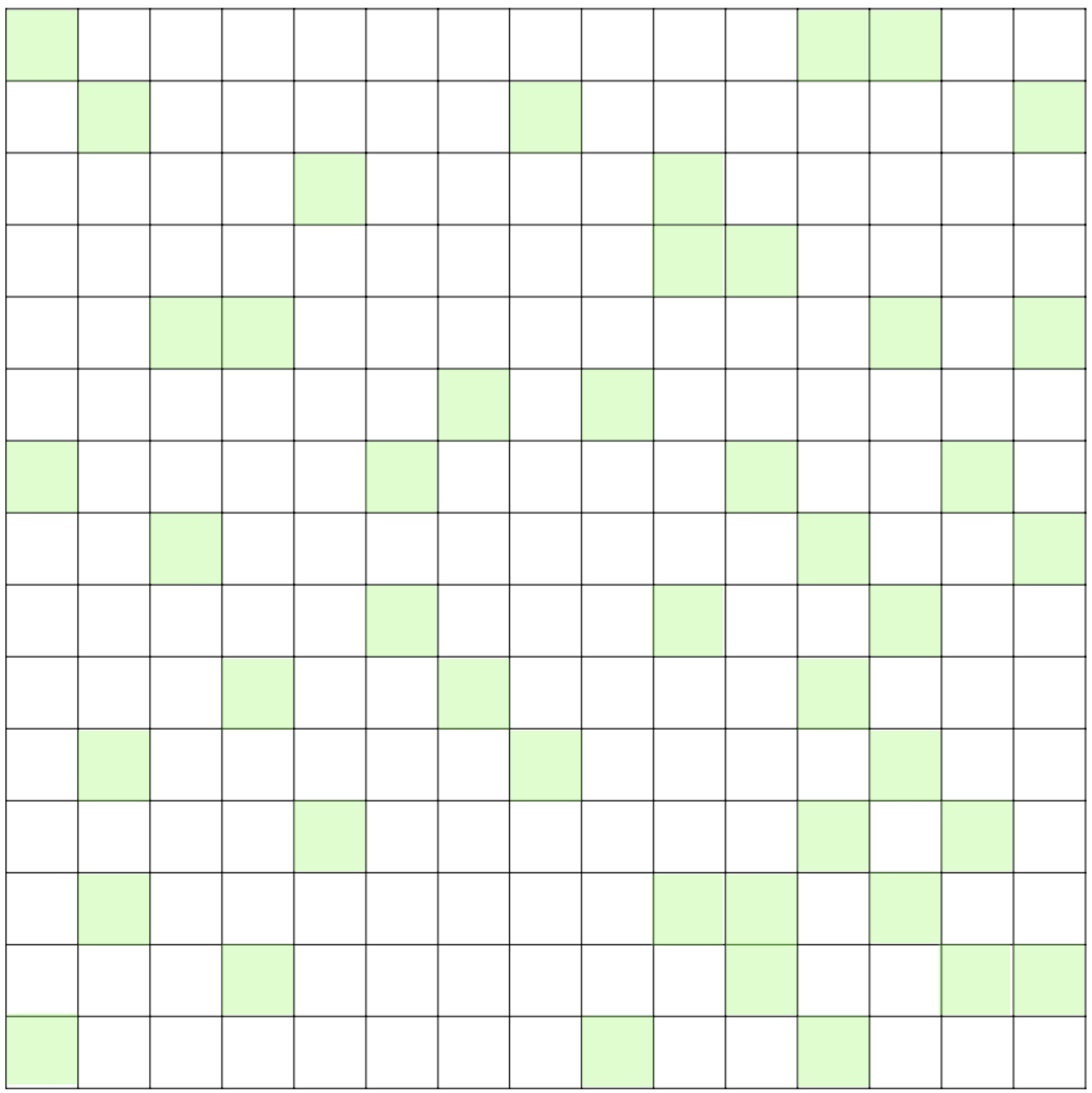
- Use Optimization approaches
- Require coherence / incoherence in data
- Work with high probability

Subspace Reconstruction Based Algorithm:

- Algebraic / Geometric approach
- Works with probability 1
- No assumptions on the data
- Deterministic

:)

Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Standard RPCA Methods:

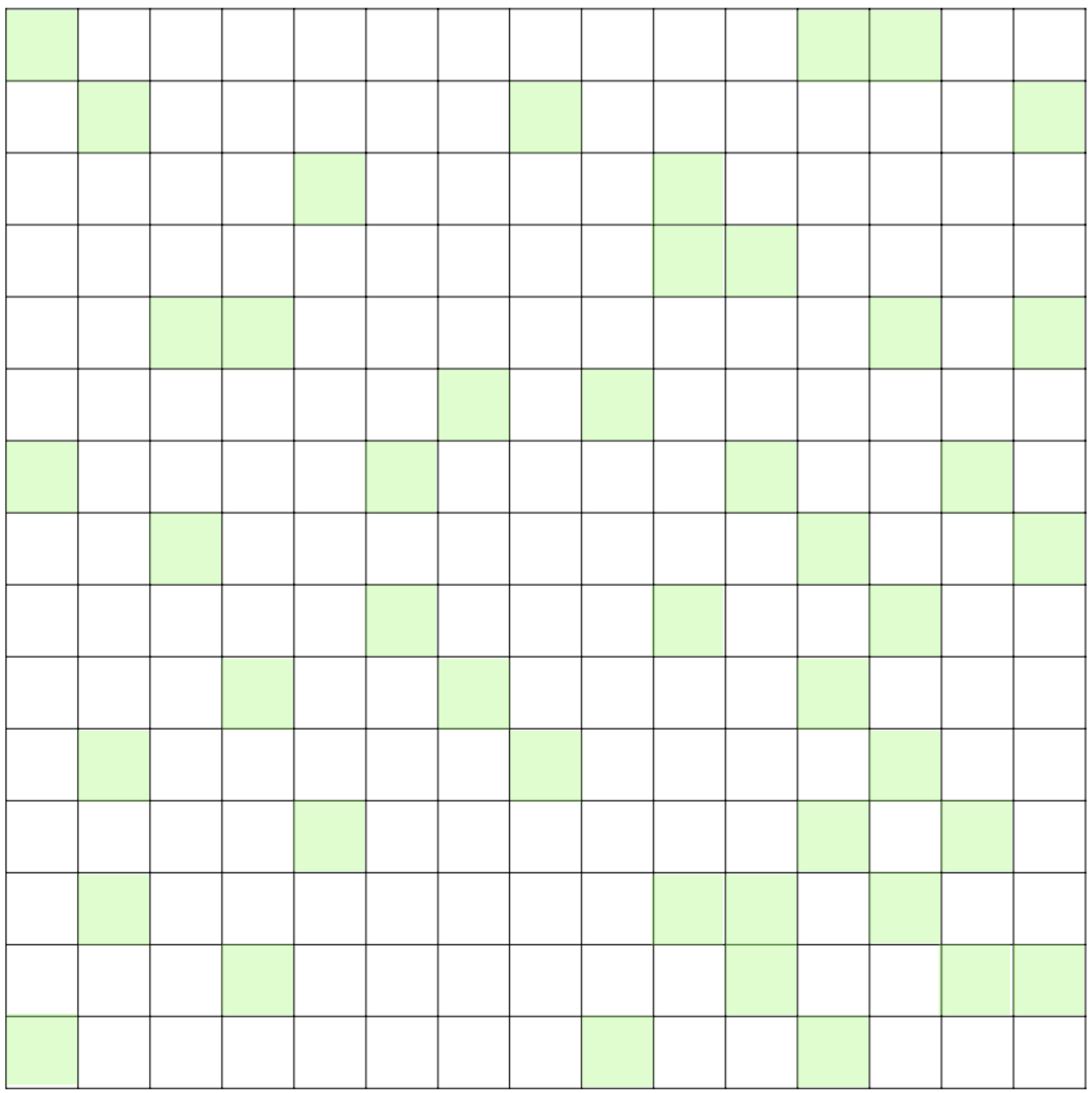
- Use Optimization approaches
- Require coherence / incoherence in data
- Work with high probability
- HAVE NOISY BOUNDS

Subspace Reconstruction Based Algorithm:

- Algebraic / Geometric approach
- Works with probability 1
- No assumptions on the data
- Deterministic

:)

Applications - RPCA Algorithms



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Standard RPCA Methods:

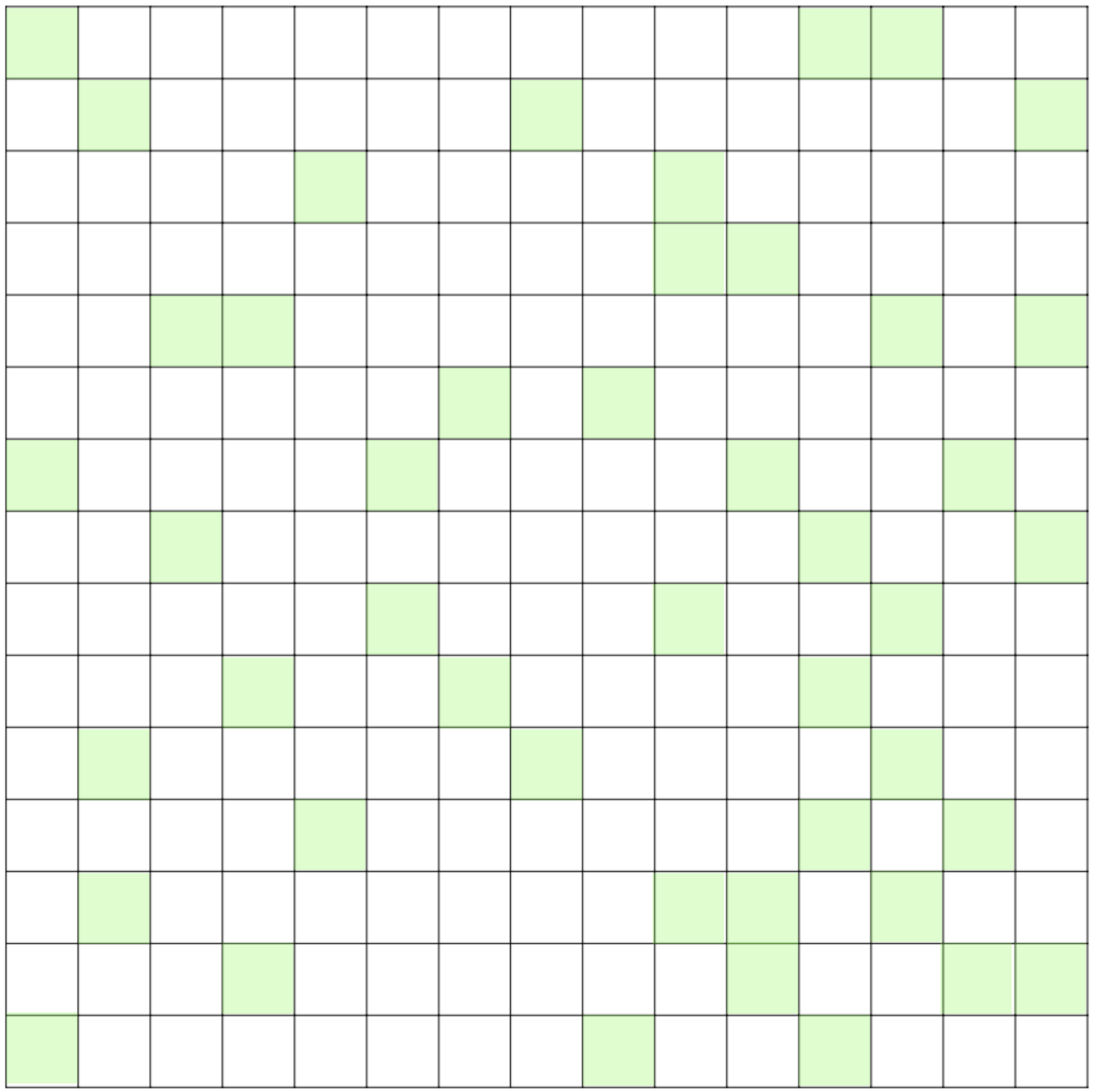
- Use Optimization approaches
- Require coherence / incoherence in data
- Work with high probability
- HAVE NOISY BOUNDS

Subspace Reconstruction Based Algorithm:

- Algebraic / Geometric approach
- Works with probability 1
- No assumptions on the data
- Deterministic
- NO NOISY BOUND

~~:)~~ :(

Applications - RPCA



Subspace reconstruction based RPCA Algorithm

D. Pimentel-Alarcón and R. Nowak, "Random consensus robust pca," *Electronic Journal of Statistics*, vol. 11, no. 2, pp. 5232–5253, 2017.

Standard RPCA Methods:

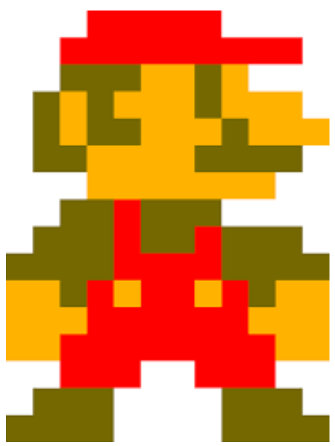
- Use Optimization approaches
- Require coherence / incoherence in data
- Work with high probability
- HAVE NOISY BOUNDS

Subspace Reconstruction Based Algorithm:

- Algebraic / Geometric approach
- Works with probability 1
- No assumptions on the data
- Deterministic
- NO NOISY BOUND…… till now!!

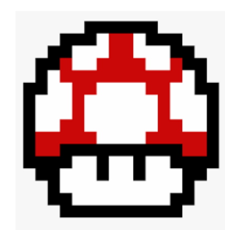
:)

Applications - Our Paper



Existing Subspace-Reconstruction Based LRMC Theory, RPCA Algorithms, etc (Noiseless)

+



Understanding of Noisy Subspace Reconstruction (Noisy Bound)

=

Generalization of existing results to Noisy Cases

(Or at least a step in this direction)

Outline

1. Problem Setup - Subspace Estimation
2. Motivation - Missing Data
3. Previous Work: Noiseless case
4. This Paper - Noisy Data and Estimation Bound
5. Applications
6. Conclusions

Conclusions

In our work, we have

- Generalized subspace estimation method to deal with noisy data

Conclusions

In our work, we have

- Generalized subspace estimation method to deal with noisy data
- Bounded the error in approximating the optimal subspace estimator using the deterministic conditions for subspace identifiability

Conclusions

In our work, we have

- Generalized subspace estimation method to deal with noisy data
- Bounded the error in approximating the optimal subspace estimator using the deterministic conditions for subspace identifiability
- Experimentally verified that sampling patterns affect the construction and bound

References

[1] D. L. Pimentel-Alarcón, N. Boston, and R. D. Nowak, "Deterministic conditions for subspace identifiability from incomplete sampling," in *Information Theory (ISIT), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 2191–2195.

[2] ——, "A characterization of deterministic sampling patterns for low-rank matrix completion," *IEEE Journal of Selected Topics in Signal Processing*, vol. 10, no. 4, pp. 623–636, 2016.

[3] D. Pimentel-Alarcon and R. Nowak, "The information-theoretic requirements of subspace clustering with missing data," in *International Conference on Machine Learning*, 2016, pp. 802–810.

[4] D. Pimentel-Alarcon, "Mixture matrix completion," *Advances in Neural Information Processing Systems*, vol. 31, pp. 2193–2203, 2018.

[5] D. Pimentel-Alarcón, G. Ongie, L. Balzano, R. Willett, and R. Nowak, "Low algebraic dimension matrix completion," in *Communication, Control, and Computing (Allerton), 2017 55th Annual Allerton Conference on*. IEEE, 2017, pp. 790–797.

[6] G. Ongie, D. Pimentel-Alarcón, L. Balzano, R. Willett, and R. D. Nowak, "Tensor methods for nonlinear matrix completion," *SIAM Journal on Mathematics of Data Science*, vol. 3, no. 1, pp. 253–279, 2021.

[7] M. Ashraphijuo, V. Aggarwal, and X. Wang, "A characterization of sampling patterns for low-tucker-rank tensor completion problem," in *Information Theory (ISIT), 2017 IEEE International Symposium on*. IEEE, 2017, pp. 531–535.

[8] ——, "Deterministic and probabilistic conditions for finite completability of low rank tensor," *arXiv preprint*, vol. 1612, 2016.

[9] M. Ashraphijuo and X. Wang, "Fundamental conditions for low-cp-rank tensor completion," *Journal of Machine Learning Research*, vol. 18, no. 63, pp. 1–29, 2017.

Karan Srivastava

ksrivastava4@wisc.edu

All code for experiments can be found at github.io/ksrivastava1