

ERROR ESTIMATIONS FOR RANDOMIZED LOW-RANK APPROXIMATIONS

LORENZO LAZZARINO

Mathematical Institute - University of Oxford

Computational Mathematics Theme - STFC UKRI

ILAS 2026, 21th May '26



Science and
Technology
Facilities Council



Mathematical
Institute

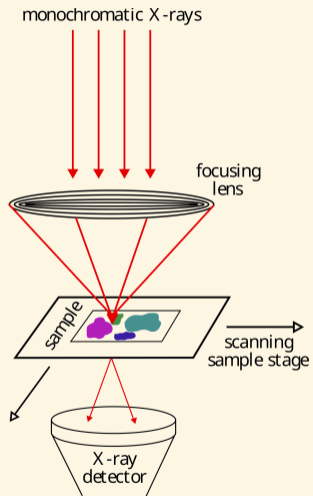
ERROR ESTIMATIONS FOR RANDOMIZED LOW-RANK APPROXIMATIONS

- 1 (ONE) MOTIVATION: SUBSAMPLING IN SPECTROMICROSCOPY
with M. Meier (RUG), B. Shustin (STFC), H. Al Daas (STFC), P. Quinn (STFC)
- 2 A-POSTERIORI ERROR ESTIMATE: GN
with K. Pearce (UT Austin), N. Pritchard (Oxford)
- 3 A-POSTERIORI ERROR ESTIMATE: CUR
with K. Pearce (UT Austin), N. Pritchard (Oxford)

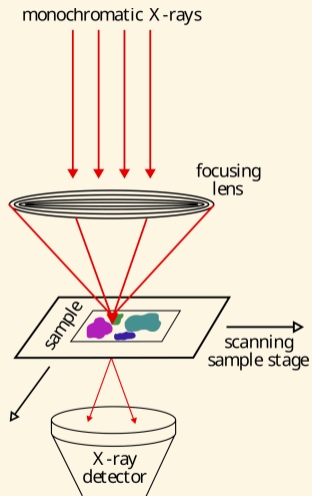
(ONE) MOTIVATION: SUBSAMPLING IN SPECTROMICROSCOPY

1

SPECTROMICROSCOPY INTRO

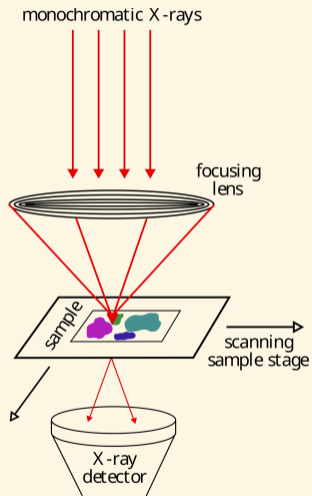


SPECTROMICROSCOPY INTRO

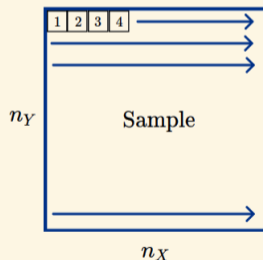


Overall Goal: Determine identity and spacial distribution of unknown materials contained in the non-homogeneous sample

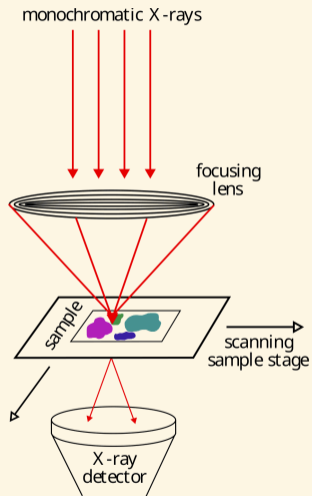
SPECTROMICROSCOPY INTRO



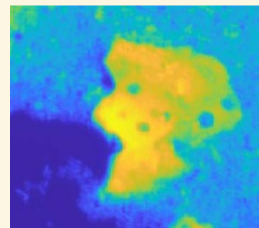
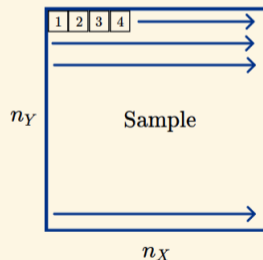
Overall Goal: Determine identity and spacial distribution of unknown materials contained in the non-homogeneous sample



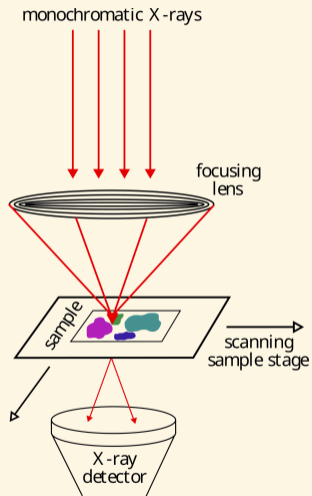
SPECTROMICROSCOPY INTRO



Overall Goal: Determine identity and spatial distribution of unknown materials contained in the non-homogeneous sample

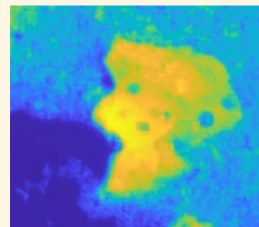
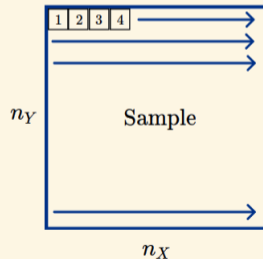


SPECTROMICROSCOPY INTRO

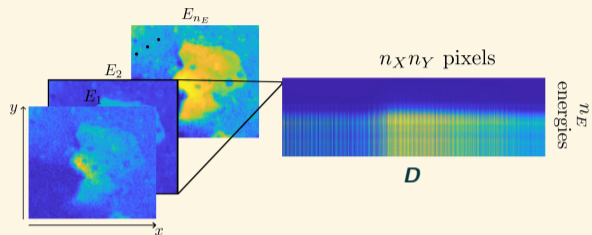


Overall Goal: Determine identity and spatial distribution of unknown materials contained in the non-homogeneous sample

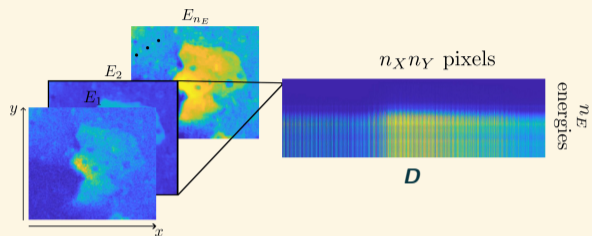
1. (Sub)sampling strategy
2. Analysis



SUBSAMPLING

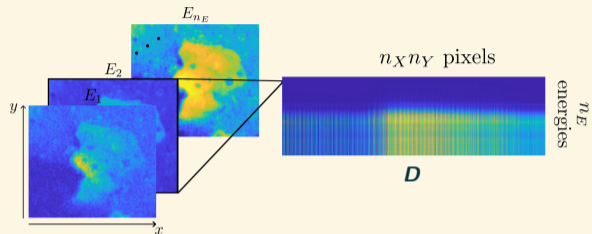


SUBSAMPLING



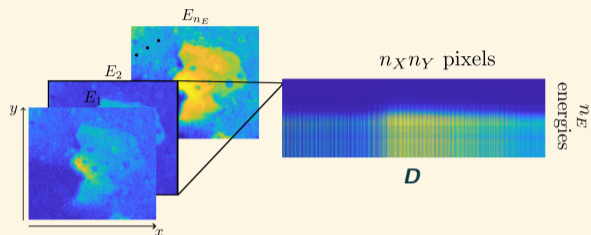
$$= \begin{matrix} S \\ \boxed{M} \end{matrix} \begin{matrix} \boxed{T} \end{matrix}$$

SUBSAMPLING



$$= \begin{matrix} S \\ \boxed{M} \end{matrix} \begin{matrix} \boxed{T} \end{matrix} + E$$

SUBSAMPLING



$$= \begin{matrix} S \\ \boxed{M} \end{matrix} \begin{matrix} \boxed{T} \end{matrix} + E$$

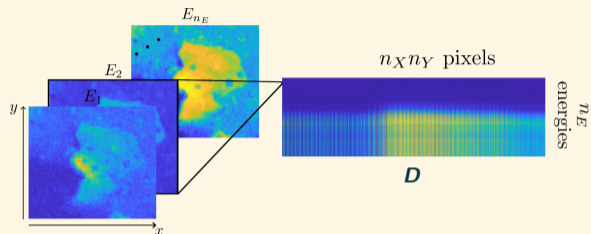
Now

- ▶ Uniform Raster sampling



- ▶ LoopedASD (Townsend et al, 2022)

SUBSAMPLING



$$= \begin{matrix} S \\ \boxed{M} \end{matrix} \begin{matrix} \boxed{T} \end{matrix} + E$$

Now

- ▶ Uniform Raster sampling



- ▶ LoopedASD (Townsend et al, 2022)

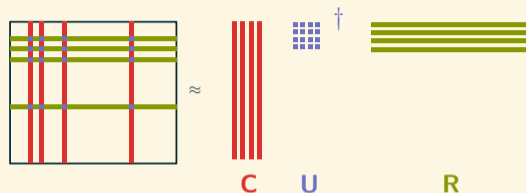
→ Importance Sampling

Leverage scores

$$\ell_i(\mathbf{D})^2 := \|\mathbf{U}_{\text{dominant}}(i, :)\|_2^2$$

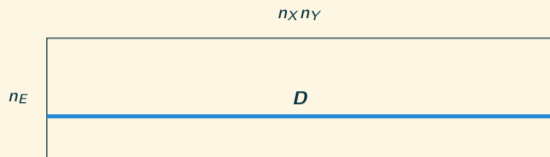
In principle, $\ell(\mathbf{D}) \approx \ell(\mathbf{M})$

CUR OVERVIEW



- ▶ Near optimal low-rank approximation
- ▶ "Motivation" of importance distributions (theoretical guarantees)
- ▶ Natural fit for experimental design
- ▶ Gives matrix completion (with interpolation of rows)

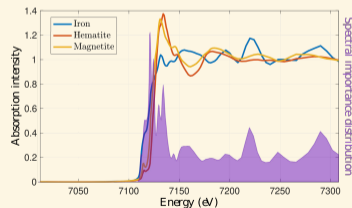
CUR IMPORTANCE SAMPLING FOR SPECTRO-MICROSCOPY (CURISS)



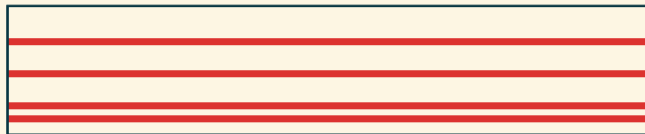
- Determine **spectral** importance distribution and measure accordingly

Data-driven approach: use a small dictionary matrix

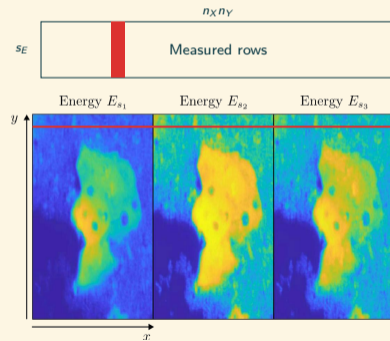
$$l_i(\mathbf{D})^2 \approx l_i(\mathbf{M}_{\text{dict}})^2$$



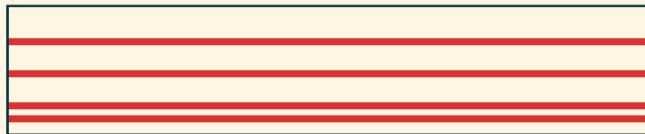
CUR IMPORTANCE SAMPLING FOR SPECTRO-MICROSCOPY (CURISS)



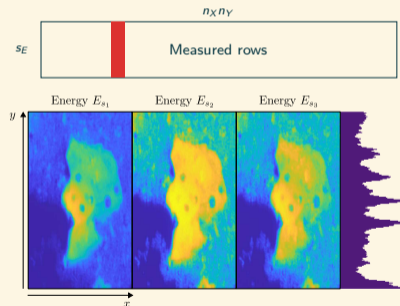
- ▶ Determine **spectral** importance distribution and measure accordingly
- ▶ Determine **spatial** importance distribution



CUR IMPORTANCE SAMPLING FOR SPECTRO-MICROSCOPY (CURISS)



- ▶ Determine **spectral** importance distribution and measure accordingly
- ▶ Determine **spatial** importance distribution



Use ARP (Cortinovic, Kressner, 2024)

CUR IMPORTANCE SAMPLING FOR SPECTRO-MICROSCOPY (CURISS)



- ▶ Determine **spectral** importance distribution and measure accordingly
- ▶ Determine **spatial** importance distribution
- ▶ Sample spatial rows and measure them at all energies

For the non-measured energies

- Set the beam energy to E .
- Measure the sampled spatial rows

CUR IMPORTANCE SAMPLING FOR SPECTRO-MICROSCOPY (CURISS)



- ▶ Determine **spectral** importance distribution and measure accordingly
- ▶ Determine **spatial** importance distribution
- ▶ Sample spatial rows and measure them at all energies

For the non-measured energies

- Set the beam energy to E .
- Measure the sampled spatial rows

CUR IMPORTANCE SAMPLING FOR SPECTRO-MICROSCOPY (CURISS)



- ▶ Determine **spectral** importance distribution and measure accordingly
- ▶ Determine **spatial** importance distribution
- ▶ Sample spatial rows and measure them at all energies

For the non-measured energies

- Set the beam energy to E .
- Measure the sampled spatial rows

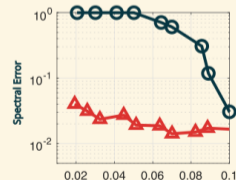
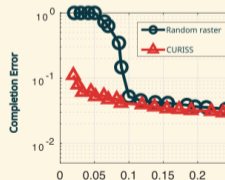
CUR IMPORTANCE SAMPLING FOR SPECTRO-MICROSCOPY (CURISS)



- ▶ Determine **spectral** importance distribution and measure accordingly
- ▶ Determine **spatial** importance distribution
- ▶ Sample spatial rows and measure them at all energies
- ▶ Complete the measured dataset using CUR



(Meier, L., Shustin, Al Daas, Quinn, 2026)



ADAPTIVE CURISS (ACURISS)

Goal: Adaptively refine the subsampling, starting from CURISS with initial ratio p_0

Refinement

Stopping Criteria

ADAPTIVE CURISS (ACURISS)

Goal: Adaptively refine the subsampling, starting from CURISS with initial ratio p_0

Refinement

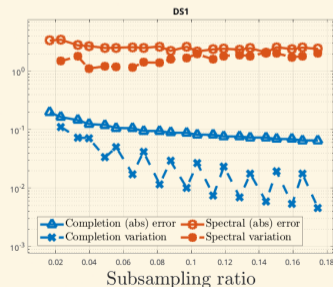
- ▶ Completion variation:

$$\|\hat{\mathbf{D}}_i - \hat{\mathbf{D}}_{i-1}\|_F \leq \eta_D$$

- ▶ Spectral variation:

$$\|\mathbf{M}_{\text{cluster}}(\hat{\mathbf{D}}_i) - \mathbf{M}_{\text{cluster}}(\hat{\mathbf{D}}_{i-1})\|_F \leq \eta_M$$

Stopping Criteria



ADAPTIVE CURISS (ACURISS)

Goal: Adaptively refine the subsampling, starting from CURISS with initial ratio p_0

Refinement

- ▶ Completion variation:

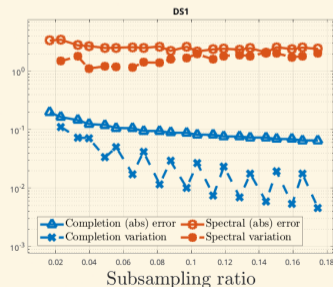
$$\|\hat{\mathbf{D}}_i - \hat{\mathbf{D}}_{i-1}\|_F \leq \eta_D$$

- ▶ Spectral variation:

$$\|\mathbf{M}_{\text{cluster}}(\hat{\mathbf{D}}_i) - \mathbf{M}_{\text{cluster}}(\hat{\mathbf{D}}_{i-1})\|_F \leq \eta_M$$

- ▶ Problem specific starting point

Stopping Criteria



A-POSTERIORI ERROR ESTIMATE: GN

2

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx AX(Y^*AX)^\dagger Y^*A =: A_{GN,X,Y}$$



(Clarkson, Woodruff, 2009)
 (Nakatsukasa, 2020)
 (Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $X \in \mathbb{R}^{n \times r}$, $Y \in \mathbb{R}^{m \times (r+\ell)}$
2. Two-side Sketch: AX and Y^*A
3. $[Q,R] = \text{qr}((Y^*A)X, 0)$
4. $A_{GN,X,Y} = ((AX)R^{-1})(Q^*(Y^*A))$

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx AX(Y^*AX)^\dagger Y^*A =: A_{GN,X,Y}$$



(Clarkson, Woodruff, 2009)
 (Nakatsukasa, 2020)
 (Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $X \in \mathbb{R}^{n \times r}$, $Y \in \mathbb{R}^{m \times (r+\ell)}$
2. Two-side Sketch: AX and Y^*A
3. $[Q,R] = \text{qr}((Y^*A)X, 0)$
4. $A_{GN,X,Y} = ((AX)R^{-1})(Q^*(Y^*A))$
 - ▶ $N_{2r+\ell} + \mathcal{O}(r^3 + (m+n)r^2)$
 - ▶ Single-pass
 - ▶ 2 multiplications by A

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx AX(Y^*AX)^\dagger Y^*A =: A_{GN,X,Y}$$



(Clarkson, Woodruff, 2009)
 (Nakatsukasa, 2020)
 (Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $X \in \mathbb{R}^{n \times r}$, $Y \in \mathbb{R}^{m \times (r+\ell)}$
2. Two-side Sketch: AX and Y^*A
3. $[Q,R] = \text{qr}((Y^*A)X, 0)$
4. $A_{GN,X,Y} = ((AX)R^{-1})(Q^*(Y^*A))$

▶ $N_{2r+\ell} + \mathcal{O}(r^3 + (m+n)r^2)$

▶ Single-pass

▶ 2 multiplications by A

Accuracy

$$\hat{r} \leq r - 2$$

$$\mathbb{E} \|A - A_{GN,X,Y}\|_F \leq \sqrt{1 + \frac{r+\ell}{\ell-1}} \sqrt{1 + \frac{r}{r-\hat{r}-1}} \|A - A_{best,\hat{r}}\|_F$$

(Tropp et al., 2017), (Nakatsukasa, 2020)

GENERALIZED NYSTRÖM APPROXIMATION

Generalized Nyström

$$A \approx AX(Y^*AX)^\dagger Y^*A =: A_{GN,X,Y}$$



(Clarkson, Woodruff, 2009)
 (Nakatsukasa, 2020)
 (Woolfe, Liberty, Rokhlin, Tygert, 2008)

1. Choose $X \in \mathbb{R}^{n \times r}$, $Y \in \mathbb{R}^{m \times (r+\ell)}$
2. Two-side Sketch: AX and Y^*A
3. $[Q,R] = \text{qr}((Y^*A)X, 0)$
4. $A_{GN,X,Y} = ((AX)R^{-1})(Q^*(Y^*A))$

▶ $N_{2r+\ell} + \mathcal{O}(r^3 + (m+n)r^2)$

▶ Single-pass

▶ 2 multiplications by A

Accuracy

$$\hat{r} \leq r - 2$$

$$\mathbb{E} \|A - A_{GN,X,Y}\|_F \leq \sqrt{1 + \frac{r+\ell}{\ell-1}} \sqrt{1 + \frac{r}{r-\hat{r}-1}} \|A - A_{best,\hat{r}}\|_F$$

(Tropp et al., 2017), (Nakatsukasa, 2020)

Stability

$$(AX)(Y^*AX)^\dagger_\epsilon Y^*A$$

(Nakatsukasa, 2020)

PROBLEM SETTING AND MAIN IDEA

Generalized Nyström

X, Y random matrix

$$AX(Y^TAX)^\dagger Y^T A$$

PROBLEM SETTING AND MAIN IDEA

Generalized Nyström

 X, Y random matrix

$$AX(Y^T AX)^\dagger Y^T A$$

Goal: Estimate the approximation error using only what you already have

$$\|A - AX(Y^* AX)^\dagger Y^* A\|^2$$

PROBLEM SETTING AND MAIN IDEA

Generalized Nyström

 X, Y random matrix

$$AX(Y^T AX)^\dagger Y^T A$$

Goal: Estimate the approximation error using only what you already have

$$\|A - AX(Y^* AX)^\dagger Y^* A\|^2$$

Sketch the residual

$$G \text{ random, } \|R\| \approx \|RG\|$$

- ☺ Good accuracy
- ☺ Very small size of G sufficient
- ☹ (More) multiplications by R
- ☹ G needs to be independent of R

PROBLEM SETTING AND MAIN IDEA

Generalized Nyström

 X, Y random matrix

$$AX(Y^T AX)^\dagger Y^T A$$



(Epperly, Tropp, 2024)

Goal: Estimate the approximation error using only what you already have

$$\|A - AX(Y^* AX)^\dagger Y^* A\|^2$$

Sketch the residual

$$G \text{ random, } \|R\| \approx \|RG\|$$

- ☺ Good accuracy
- ☺ Very small size of G sufficient
- ☹ (More) multiplications by R
- ☹ G needs to be independent of R

Leave-one-out

Remove one sample column and use it to sketch the error

- ☺ We use only what we have already computed
- ☹ Need fast formula

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A **SPSD** matrix, **X** random matrix

$$AX(X^TAX)^{-1}X^TA$$



(Epperly, Tropp, 2024)

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A **SPSD** matrix, **X** random matrix

$$AX(X^TAX)^{-1}X^TA$$



(Epperly, Tropp, 2024)

$$\left\| A - A \begin{bmatrix} | & & | & & | \\ x_1 & \cdots & x_j & \cdots & x_r \\ | & & | & & | \end{bmatrix} \left(\begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} | & & | & & | \\ x_1 & \cdots & x_j & \cdots & x_r \\ | & & | & & | \end{bmatrix} \right)^{-1} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \right\|^2$$

- 1.
- 2.
- 3.

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A **SPSD** matrix, **X** random matrix

$$AX(X^TAX)^{-1}X^TA$$



(Epperly, Tropp, 2024)

$$\left\| A - A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \cdots & x_j & \cdots & x_r \end{bmatrix}^{-1} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \right\|^2$$

1. Remove one column from X
- 2.
- 3.

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A **SPSD** matrix, X random matrix

$$AX(X^TAX)^{-1}X^TA$$



(Epperly, Tropp, 2024)

$$\left\| \left(A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \right)^{-1} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} x_j \end{bmatrix} \right\|^2$$

The diagram illustrates the leave-one-out error estimate. It shows the matrix $A - A X X^T A$ where the j -th column of X is shaded with diagonal lines. This matrix is inverted, and then multiplied by the j -th column of X (also shaded with diagonal lines) and the corresponding row of A (shaded with diagonal lines). The result is a vector whose squared norm is the error estimate. A green arrow points from the shaded x_j column to the shaded x_j vector.

1. Remove one column from X
2. Use it to sketch the error
- 3.

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A **SPSD** matrix, **X** random matrix

$$AX(X^TAX)^{-1}X^TA$$



(Epperly, Tropp, 2024)

$$\frac{1}{r} \sum_{j=1}^r \left\| \left(A - A \begin{bmatrix} x_1 & \cdots & \text{---} & \cdots & x_r \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} \right)^{-1} A \begin{bmatrix} x_1 & \cdots & \text{---} & \cdots & x_r \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} \right\|^2$$

The diagram illustrates the leave-one-out error estimation process. It shows a matrix A with columns $x_1, \dots, x_j, \dots, x_r$. The column x_j is highlighted with a hatched pattern. The matrix is then partitioned into blocks, with the j -th column and row removed. The inverse of the resulting matrix is computed, and the error is measured as the squared norm of the product of A and the removed column x_j .

1. Remove one column from X
2. Use it to sketch the error
3. Sum over all possible indices

LEAVE-ONE-OUT FOR SYMMETRIC APPROXIMATIONS

Nyström

A **SPSD** matrix, **X** random matrix

$$AX(X^TAX)^{-1}X^TA$$



(Epperly, Tropp, 2024)

$$\frac{1}{r} \sum_{j=1}^r \left\| \left(A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \right)^{-1} \begin{bmatrix} x_1^* \\ \vdots \\ x_j^* \\ \vdots \\ x_r^* \end{bmatrix} A \begin{bmatrix} x_j \end{bmatrix} \right\|^2$$

1. Remove one column from X
2. Use it to sketch the error
3. Sum over all possible indices

$$H = X^*AX$$

$$AX = QR$$

$$= \frac{1}{r} \left\| RH \text{diag} \left(\frac{1}{[H^{-1}]_{ii}}, i = 1, \dots, r \right) \right\|_F^2$$

= Cheap to compute formula!

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^T A$$



(L., Pearce, Pritchard, 2026)

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^T A$$



(L., Pearce, Pritchard, 2026)

$$\|A - A \begin{bmatrix} | & & | & & | \\ x_1 & \cdots & x_j & \cdots & x_r \\ | & & | & & | \end{bmatrix} \left(\begin{bmatrix} \text{---} y_1 \text{---} \\ \vdots \\ \text{---} y_\ell^* \text{---} \\ \vdots \\ \text{---} y_r^* \text{---} \end{bmatrix} A \begin{bmatrix} | & & | & & | \\ x_1 & \cdots & x_j & \cdots & x_r \\ | & & | & & | \end{bmatrix} \right)^\dagger \begin{bmatrix} \text{---} y_1^* \text{---} \\ \vdots \\ \text{---} y_\ell^* \text{---} \\ \vdots \\ \text{---} y_r^* \text{---} \end{bmatrix} A \|^2$$

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^T A$$



(L., Pearce, Pritchard, 2026)

$$\frac{1}{r^2} \sum_{j,\ell=1}^r \left\| y_\ell^* (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix}^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A) x_j \right\|^2$$

The diagram illustrates the leave-one-out error estimate. It shows the matrix A being approximated by $AX(Y^TAX)^\dagger Y^T A$. The matrix X is formed by columns $x_1, \dots, x_j, \dots, x_r$, and Y is formed by rows $y_1, \dots, y_\ell^*, \dots, y_r^*$. The error term is the squared norm of the difference between A and its approximation, evaluated at the column x_j . The j -th column of X and the ℓ -th row of Y are highlighted with red hatching, and the corresponding error term is highlighted with a red box. A red arrow points from the error term to the j -th column of X , and a green arrow points from the ℓ -th row of Y to the error term.

LEAVE-PAIR-OUT

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^T A$$



(L., Pearce, Pritchard, 2026)

$$\frac{1}{r} \sum_{j=1}^r \left| y_j^* (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix}^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A) x_j \right|^2$$

The diagram illustrates the leave-one-out error estimate for the generalized Nyström approximation. It shows the matrix A with columns $x_1, \dots, x_j, \dots, x_r$ and rows $y_1, \dots, y_j^*, \dots, y_r^*$. The matrix A is approximated by $AX(Y^TAX)^\dagger Y^T A$, where X is the matrix of columns $x_1, \dots, x_j, \dots, x_r$ and Y is the matrix of rows $y_1, \dots, y_j^*, \dots, y_r^*$. The error estimate is the squared norm of the difference between the original matrix A and the approximation, evaluated at the j -th column x_j .

LEAVE-TWINS-OUT

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM

Generalized Nyström

Low-rank approximation: A general matrix, X, Y random matrix

$$AX(Y^TAX)^\dagger Y^T A$$



(L., Pearce, Pritchard, 2026)

$$\frac{1}{r} \sum_{j=1}^r \left\| \left(A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix}^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A \right) x_j \right\|^2$$

The diagram illustrates the error estimation for a single column x_j in the generalized Nyström approximation. It shows the matrix A with columns $x_1, \dots, x_j, \dots, x_r$ and rows $y_1, \dots, y_\ell^*, \dots, y_s^*$. The matrix A is approximated by $A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix}^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix}$. The error is measured by the squared norm of the difference between the original matrix A and the approximation, applied to the column x_j .

LEAVE-**RIGHT**-OUT

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM > Cheap to compute formula

$$\text{LPO} = \frac{1}{r^2} \sum_{j,\ell=1}^r \left\| y_\ell^* (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_\ell \\ \vdots \\ y_r \end{bmatrix})^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_r^* \end{bmatrix} A x_j \right\|^2$$

$$\text{LTO} = \frac{1}{r} \sum_{j=1}^r \left\| y_j^* (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_j \\ \vdots \\ y_r \end{bmatrix})^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A x_j \right\|^2$$

$$\text{LRO} = \frac{1}{r} \sum_{j=1}^r \left\| (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix})^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_\ell^* \\ \vdots \\ y_s^* \end{bmatrix} A x_j \right\|^2$$

LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM > Cheap to compute formula

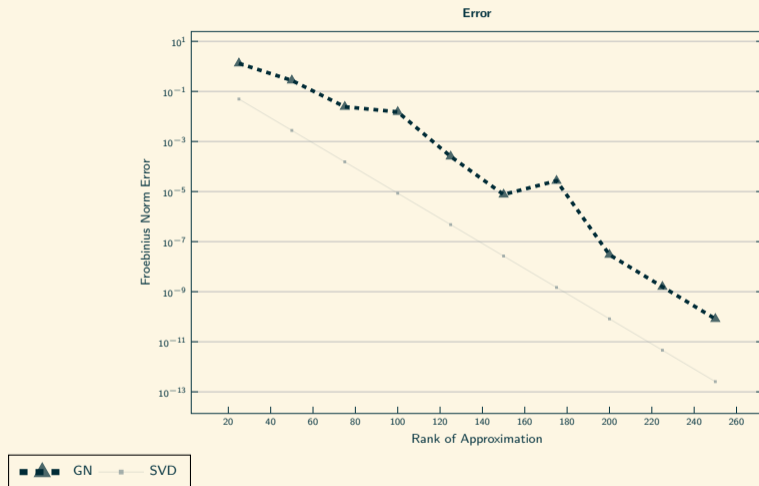
$$\text{LPO} = \frac{1}{r^2} \sum_{j,\ell=1}^r \left| y_{\ell}^* (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_{\ell} \\ \vdots \\ y_r \end{bmatrix})^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_{\ell}^* \\ \vdots \\ y_r^* \end{bmatrix} A x_j \right|^2 = \frac{1}{r^2} \sum_{j=1}^r \sum_{\ell=1}^r \left| \frac{1}{[H^{-1}]_{j,\ell}} \right|^2$$

$$H = Y^* A X \\ AX = QR$$

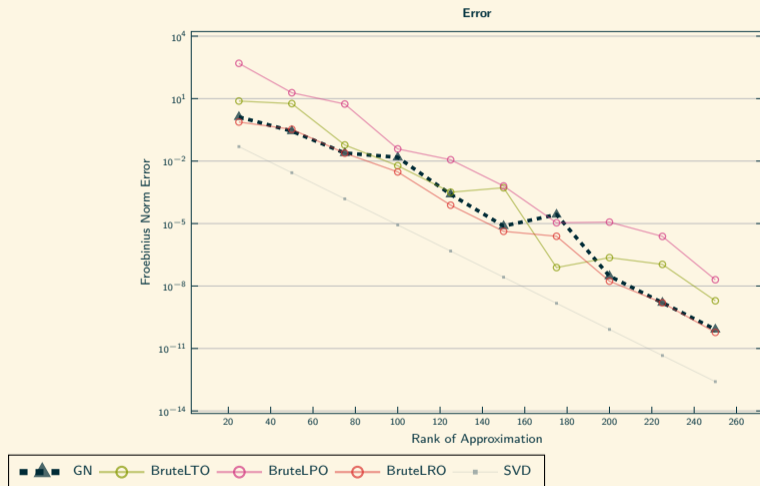
$$\text{LTO} = \frac{1}{r} \sum_{j=1}^r \left| y_j^* (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_j \\ \vdots \\ y_r \end{bmatrix})^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A x_j \right|^2 = \frac{1}{r} \sum_{j=1}^r \left| \frac{1}{[H^{-1}]_{j,j}} \right|^2$$

$$\text{LRO} = \frac{1}{r} \sum_{j=1}^r \left\| (A - A \begin{bmatrix} x_1 & \dots & x_j & \dots & x_r \end{bmatrix} \begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix})^\dagger \begin{bmatrix} y_1^* \\ \vdots \\ y_j^* \\ \vdots \\ y_r^* \end{bmatrix} A x_j \right\|^2 = \frac{1}{r} \|R(H^* H)^{-1} \text{diag}(\frac{1}{[(H^* H)^{-1}]_{i,i}}, i = 1, \dots, s)\|_F^2$$

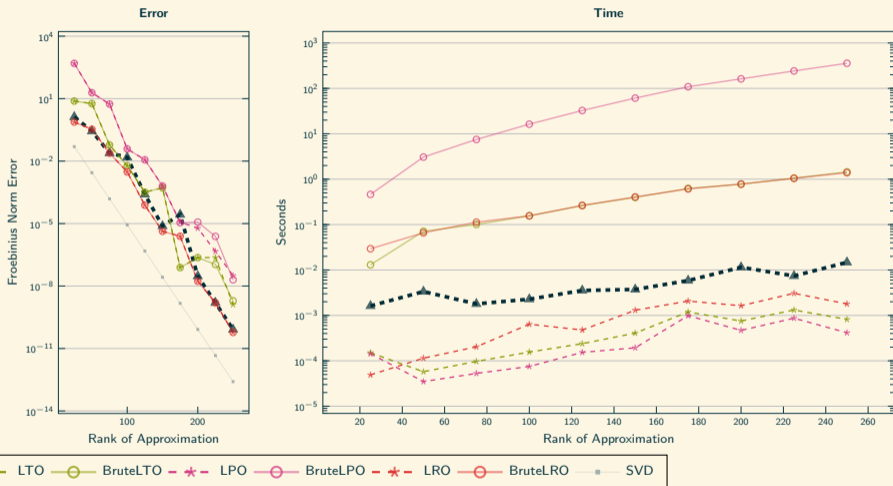
LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM > Experiments



LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM > Experiments



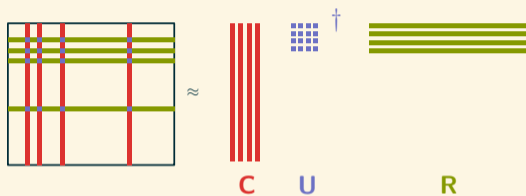
LEAVE-ONE-OUT FOR GENERALIZED NYSTRÖM > Experiments



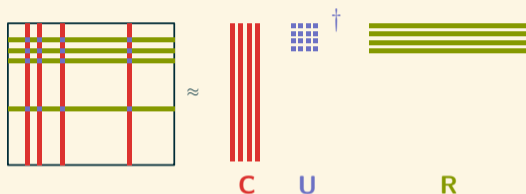
A-POSTERIORI ERROR ESTIMATE: CUR

3

MAIN CHARACTER AND PROBLEM SETTING



MAIN CHARACTER AND PROBLEM SETTING

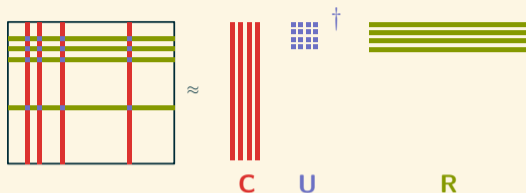


Goal

Estimate the approximation error blindly

$$\|A - CUR\|_F^2$$

MAIN CHARACTER AND PROBLEM SETTING

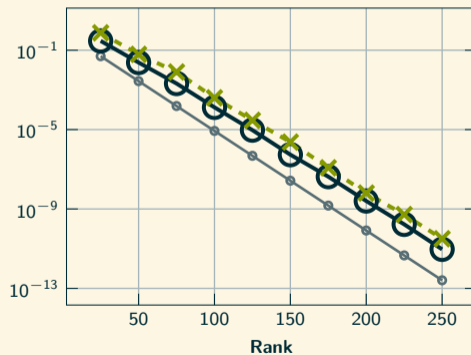


Leave-right-out for CUR

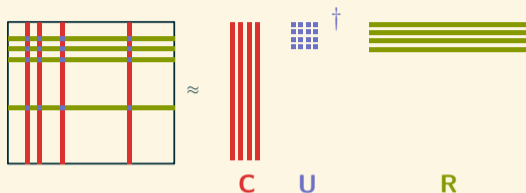
Goal

Estimate the approximation error blindly

$$\|A - CUR\|_F^2$$



MAIN CHARACTER AND PROBLEM SETTING



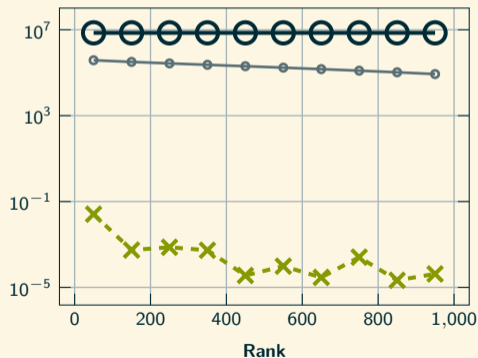
Leave-right-out for CUR

- ▶ Biased estimator
- ▶ Problem with "lack of randomness"
- ▶ Not reliable in general (we'll see it later)

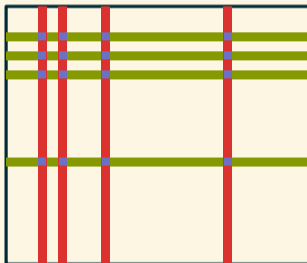
Goal

Estimate the approximation error blindly

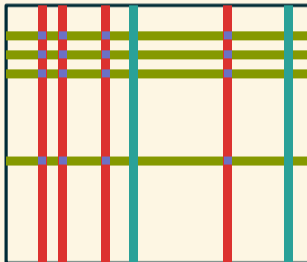
$$\|A - CUR\|_F^2$$



BLIND ERROR ESTIMATE FOR CUR

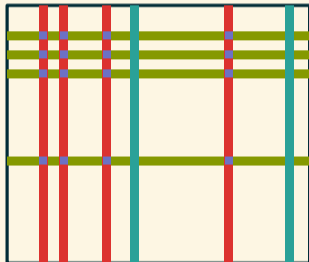
Proposed estimator

BLIND ERROR ESTIMATE FOR CUR

Proposed estimator

1. Select extra columns

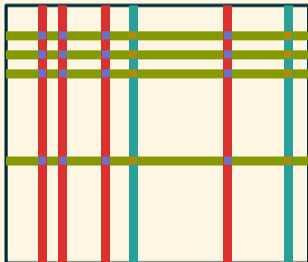
BLIND ERROR ESTIMATE FOR CUR

Proposed estimator

1. Select extra columns

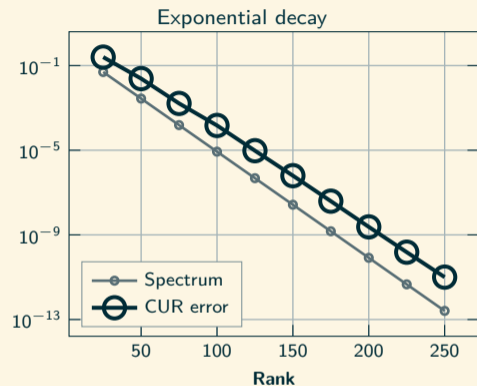
- ▶ Which extra columns?
- ▶ How many extra columns?

BLIND ERROR ESTIMATE FOR CUR

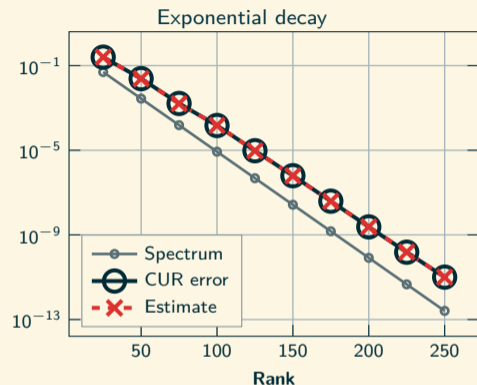
Proposed estimator

1. Select extra columns
 - ▶ Which extra columns?
 - ▶ How many extra columns?
2. Compute the error "on the extra column"

$$\| \begin{matrix} \text{teal} \\ \text{columns} \end{matrix} - \begin{matrix} \text{red} \\ \text{columns} \end{matrix} \begin{matrix} \text{blue} \\ \text{matrix} \end{matrix}^\dagger \begin{matrix} \text{orange} \\ \text{columns} \end{matrix} \|_F$$

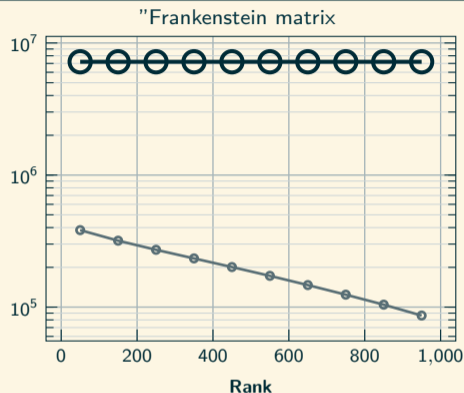
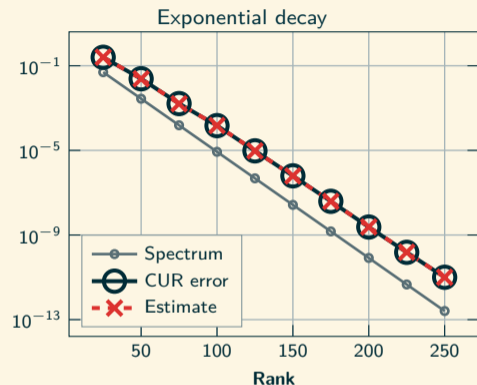
BLIND ERROR ESTIMATE FOR CUR > *Does it work?*

- ▶ Haar distributed singular vectors
- ▶ CUR via Sketch QR

BLIND ERROR ESTIMATE FOR CUR > *Does it work?*

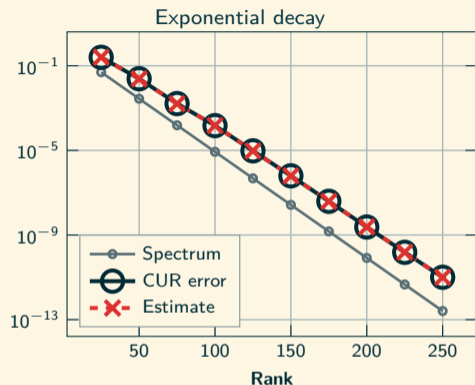
- ▶ Haar distributed singular vectors
- ▶ CUR via Sketch QR
- ▶ 5 uniform sampled extra columns
- ▶ 10 trials

BLIND ERROR ESTIMATE FOR CUR > Does it work?

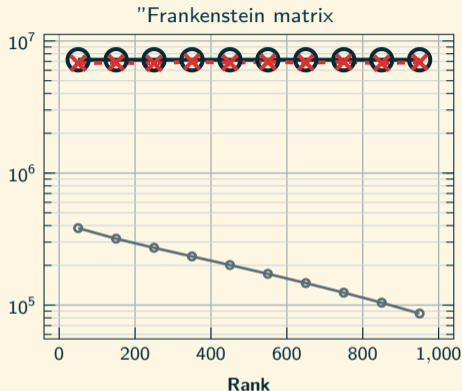


- ▶ Haar distributed singular vectors
- ▶ CUR via Sketch QR
- ▶ 5 uniform sampled extra columns
- ▶ 10 trials

- ▶ $A = \text{blockdiag}\{\text{cauchy}, \text{golub}, \text{randcorr}, \text{hilb}\}$
- ▶ CUR with first columns/rows

BLIND ERROR ESTIMATE FOR CUR > *Does it work?*

- ▶ Haar distributed singular vectors
- ▶ CUR via Sketch QR
- ▶ 5 uniform sampled extra columns
- ▶ 10 trials

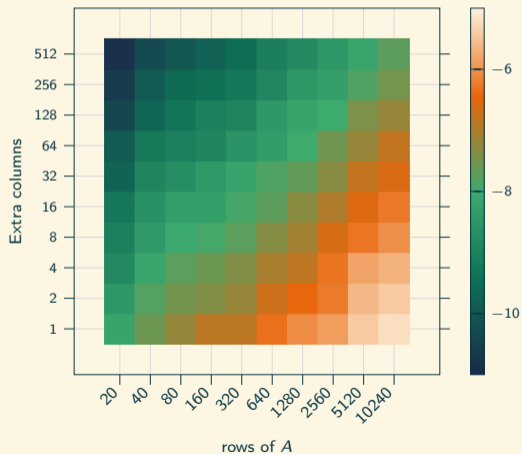


- ▶ $A = \text{blockdiag}\{\text{cauchy}, \text{golub}, \text{randcorr}, \text{hilb}\}$
- ▶ CUR with first columns/rows
- ▶ 5 uniform sampled extra columns
- ▶ 10 trials

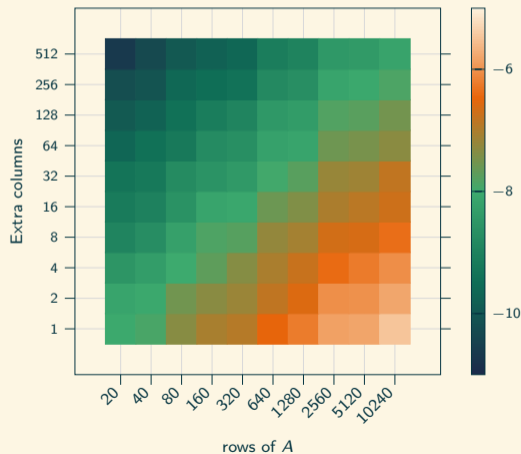
BLIND ERROR ESTIMATE FOR CUR > How many extra columns?

A with algebraically decay singular values, Haar distributed singular vectors;
target rank = 10, exact rank = 15 (fixed or doubled)

Fixed rank



Increased rank



Theory

- ▶ Unbiased estimator
- ▶ Let J be the set of indices corresponding to selected columns in the CUR. Let I be a random subset of $\{1, \dots, n\} \setminus J$, and $p_k > 0$ be the probability of $k \in I$. Let S_I be the matrix whose columns are scaled canonical vectors indexed by I , with scaling $\omega_k > 0$ for the column corresponding to the index $k \in I$. Then,

$$\frac{\|(A - CUR)S_I\|_F^2 - \mathbb{E}[\|(A - CUR)S_I\|_F^2]}{\|(A - CUR)S_I\|_F^2} \sim SE\left(\frac{\max_i \omega_k^2}{2 \min_k p_k \omega_k^2}, 0\right)$$

- ▶ Also,

$$\frac{\mathbb{E}[\|A - CUR\|_F^2 - \|(A - CUR)S_I\|_F^2]^2}{\|A - CUR\|_F^4} \leq (\max_k (1 - p_k))^2.$$

Practice

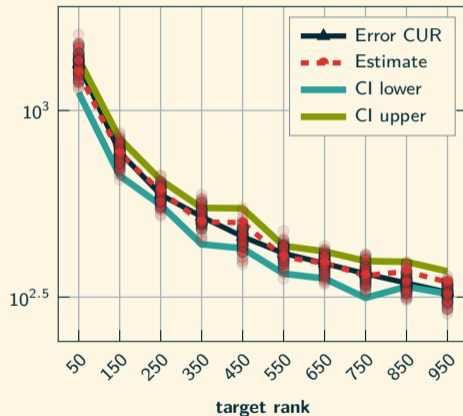
Bootstrap

Provide uncertainty sets for the estimate by:

- ▶ Repeatedly resample from observed data to create many simulated datasets
- ▶ Compute statistic on each one to build an empirical distribution of it

Experiment: Chan 10000×10000 matrix

Bootstrap: Column norm of residual at the extra columns on first trial



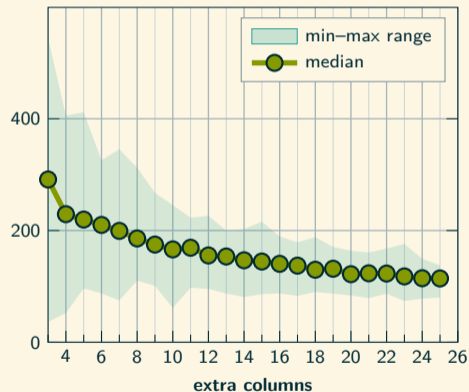
Practice**Bootstrap**

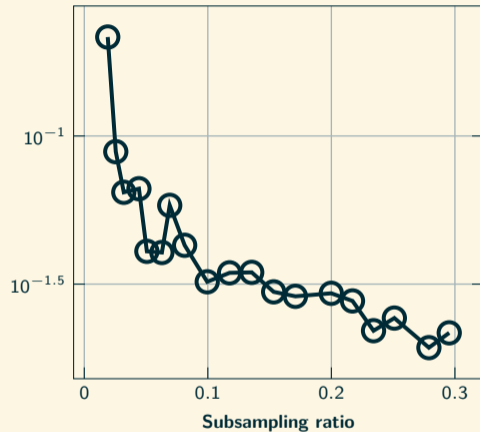
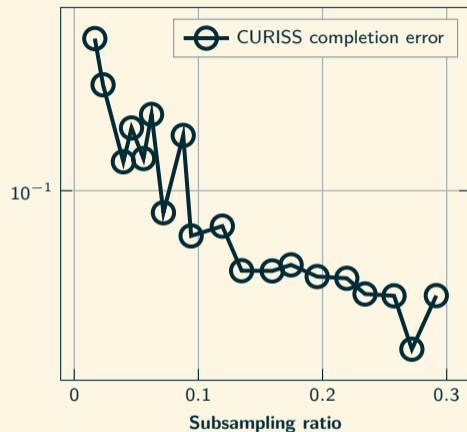
Provide uncertainty sets for the estimate by:

- ▶ Repeatedly resample from observed data to create many simulated datasets
- ▶ Compute statistic on each one to build an empirical distribution of it

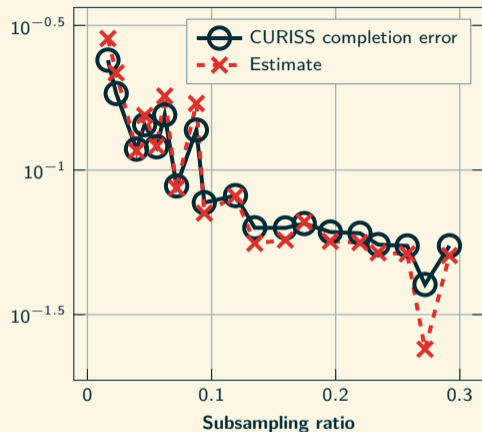
Experiment: Chan 10000×10000 matrix

Bootstrap: Column norm of residual at the extra columns on first trial

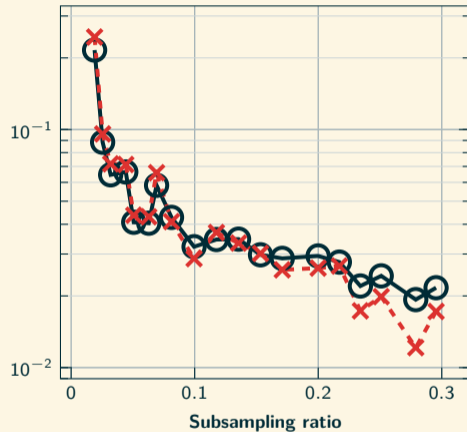


BLIND ERROR ESTIMATE FOR CUR > *Spectromicroscopy*

- ▶ Spectromicroscopy datasets
- ▶ CUR via CURISS (presented at the start)

BLIND ERROR ESTIMATE FOR CUR > *Spectromicroscopy*

- ▶ Spectromicroscopy datasets
- ▶ CUR via CURISS (presented at the start)



- ▶ 2 extra spatial rows
- ▶ 100 trials

THANK YOU!



ERROR ESTIMATIONS FOR RANDOMIZED LOW-RANK APPROXIMATIONS

LORENZO LAZZARINO

- [1] Reducing acquisition time and radiation damage: data-driven subsampling for spectro-microscopy, M. Meier, L. L., B. Shustin, H. Al Daas, P. Quinn, 2026, Arxiv
- [2] Efficient error estimators for generalized Nystrom, L. L., K. Pearce, N. Pritchard, 2026, Arxiv
- [3] Blind error estimator for CUR, L. L., K. Pearce, N. Pritchard, Hopefully soon!