# A low rank and spectrally accurate elliptic PDE solver

Heather Wilber
Center for Applied Mathematics,
Cornell University

July 9, 2018

Joint work with
Alex Townsend

# The Sylvester matrix equation

$$AX - XB = F$$

$X$ has rapidly decaying singular values $\longrightarrow$ $X$ is well-approximated by a low rank matrix

## Example: Poisson's equation on unit square

$$\begin{cases} u_{xx} + u_{yy} &= f, \\ u(x,y)\Big|_{\partial\Omega} &= 0 \end{cases} \quad \longrightarrow$$

**Finite differences:**

$$D_2 X + X D_2 = F, \quad u(\tfrac{i}{n}, \tfrac{j}{n}) \approx X_{ij}$$

D. Fortunato

**Ultraspherical-based:**

$$A\tilde{X} + \tilde{X}B = \tilde{F}, \; u(x,y) \approx \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} \tilde{X}_{ij} \phi_j(x) \phi_k(y)$$

(Beckermann & Townsend, 2017), (Antoulas, Sorenson & Zhou, 2002), (Sabino, 2008 ), (Penzl, 1999), (Benner, Truhar & Li, 2009), (Li & White, 2002), (Druskin, Knizhnerman & Simoncini, 2011), (Peaceman & Rachford, 1955), (Lu & Wachspress, 1991), (Kressner & Sirkovic, 2015), (W., Townsend, & Wright, 2017), (Townsend & Fortunato, 2017)

# Low rank approximation

$$\mathrm{rank}(Y) \leq \rho$$

$$Y = A B^*$$

$\rho$ cols

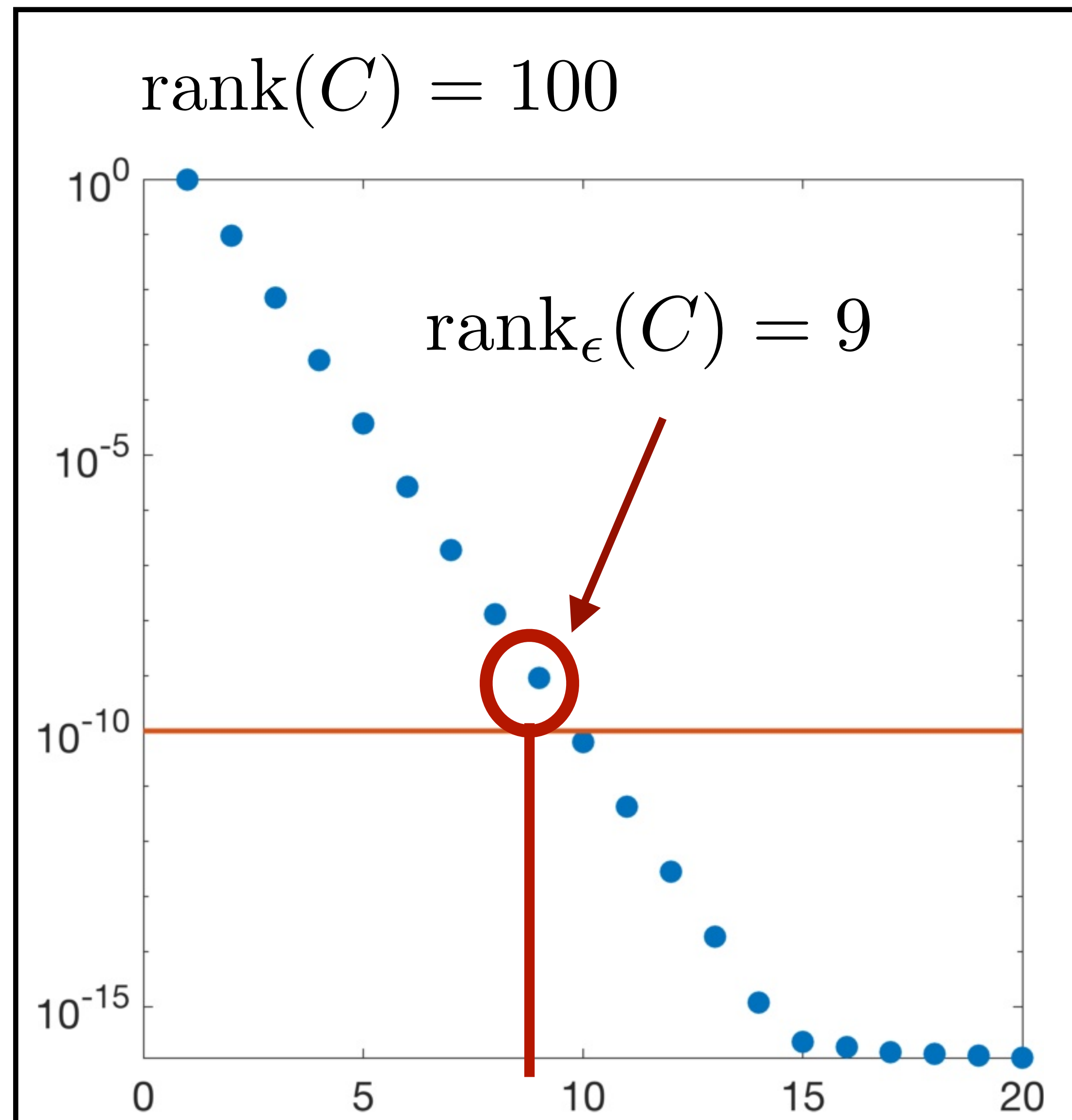
$$\mathrm{rank}_\epsilon(X) \leq \rho$$

$$X \approx A B^*$$

$\rho$ cols

$$\sigma_{\rho+1}(X) = \min\{\|X - Y\|_2,\ \mathrm{rank}(Y) = \rho\}$$

$$\mathrm{rank}_\epsilon(X) = \text{smallest } \rho \text{ where } \sigma_{\rho+1(X)} \leq \varepsilon\|X\|_2$$



# Sylvester equations with low rank F

$$AX - XB = F$$

When rank$(F)$ is very small...


A. Townsend B. Beckermann

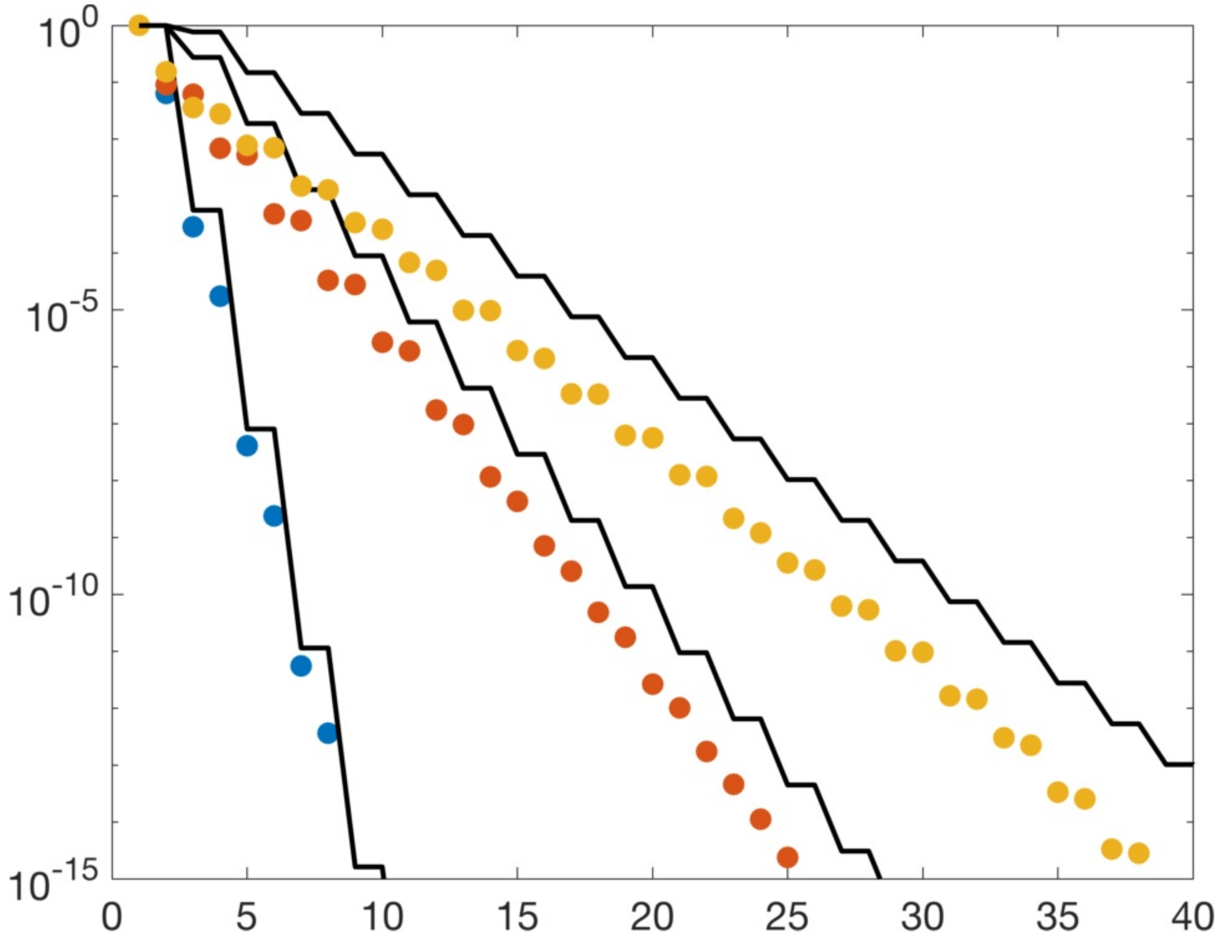
real Vandermonde
Cauchy
Pick
Löwner
real pos. def. Hankel

Large scale dynamical systems
(reduced order models)

$$AX + XA^T = BB^T$$

Poisson's equation

$$u_{xx} + u_{yy} = f, \; X_{ij} = u(\tfrac{i}{n}, \tfrac{j}{n})$$




$$\boxed{\sigma_{j\rho+1}(X) \leq C\mu^{-j}\|X\|_2, \quad \rho = \text{rank}(F)}$$

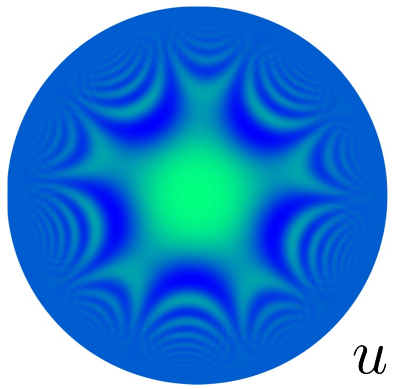
(Beckermann & Townsend, 2017), (Antoulas, Sorenson & Zhou, 2002), (Sabino, 2008 ), (Penzl, 1999), (Benner, Truhar & Li, 2009), (Li & White, 2002), (Druskin, Knizhnerman & Simoncini, 2011), (Peaceman & Rachford, 1955), (Lu & Wachspress, 1991), (Kressner & Sirkovic, 2015), (W., Townsend, & Wright, 2017), (Townsend & Fortunato, 2017)

# Sylvester equations with high rank F

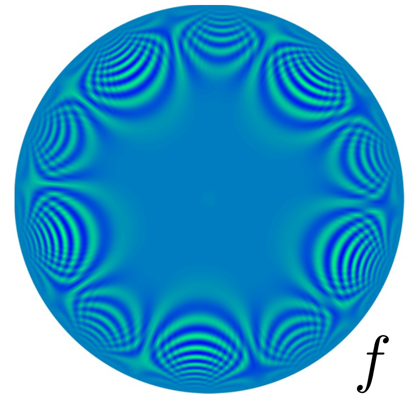
Poisson's equation with smooth right-hand sides

$$u_{xx} + u_{yy} = f$$

$$D_{xx}X + XD_{yy} = F$$

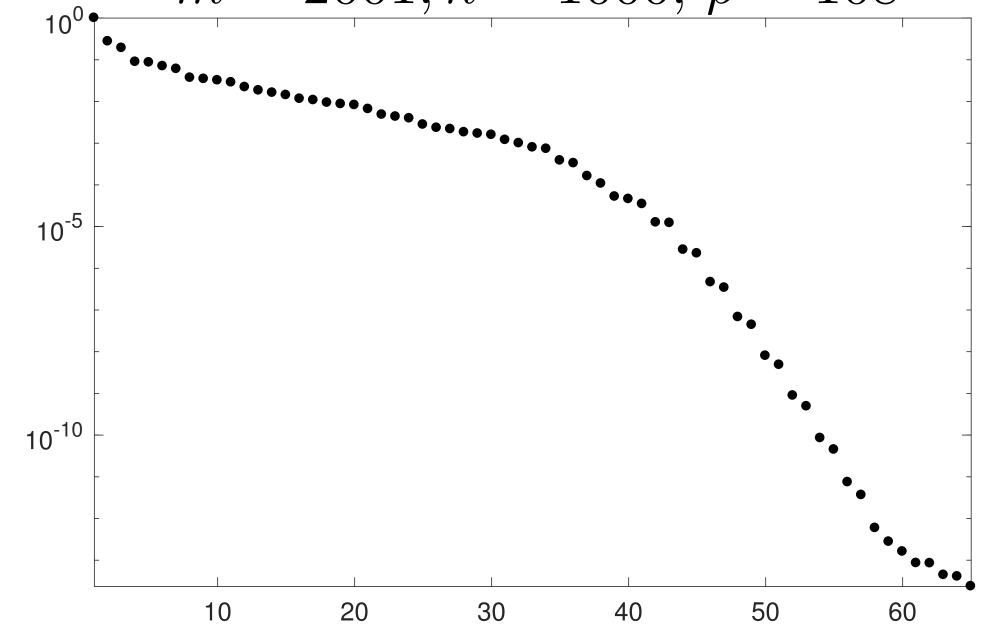


$u$



$f$

$m = 2001, n = 1000,\ \rho = 108$




2D (and d-dimensional) Vandermonde

$$\left[\ 1 \mid x \mid y \mid x^2 \mid xy \mid y^2 \mid x^3 \mid x^2y \mid xy^2 \mid y^3 \mid \cdots \mid x^{n-1} \mid \cdots \mid y^{n-1}\ \right]$$

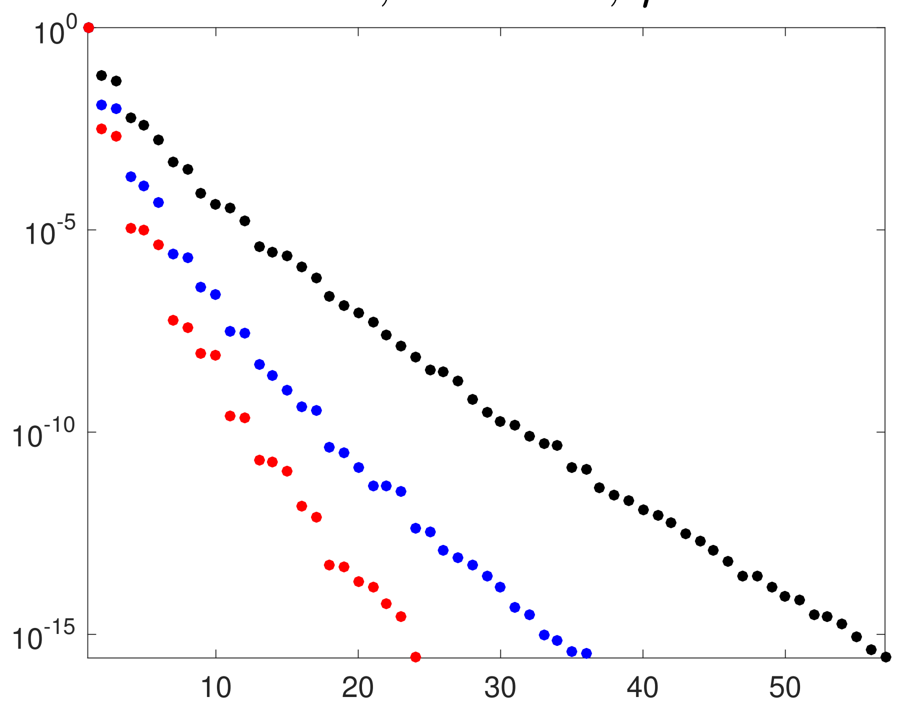
J. Sethna

K. Quinn

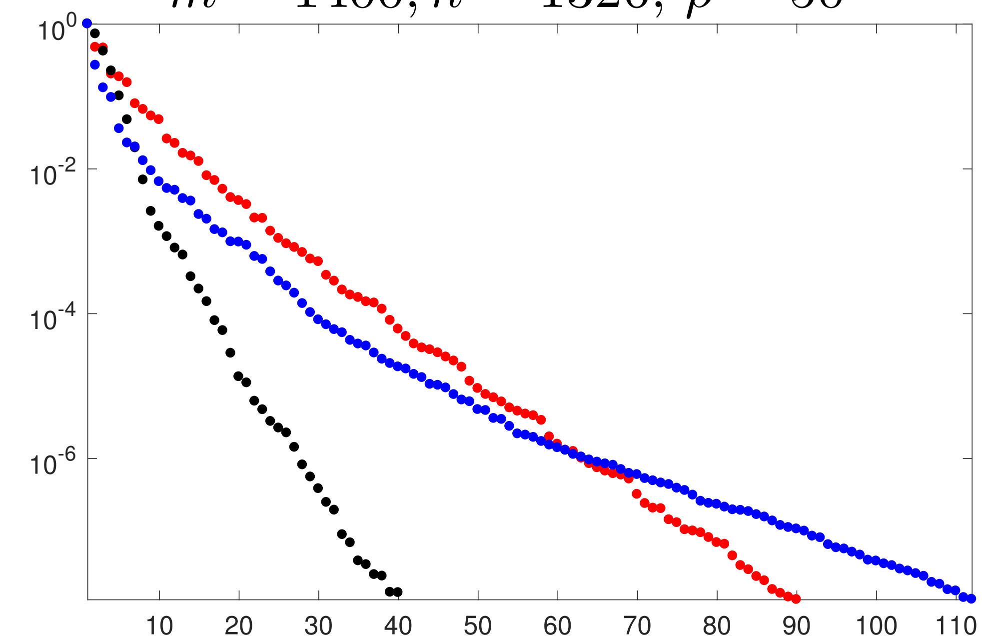
Structured matrices

$$(C_2)_{j,k} = \frac{1}{|z_j - w_k|^2}$$

$m = 100, n = 100,\ \rho = 100$




$m = 1400, n = 1326,\ \rho = 50$



# Matrices with high (A,B)-displacement rank

$$AX - XB = F$$

Standard assumption: $\operatorname{rank}(F)$ is small. $\Rightarrow$ $\boxed{\sigma_{j\rho+1}(X) \leq C\mu^{-j}\|X\|_2}$

Our assumption: the singular values of $F$ decay rapidly. $\Rightarrow$ ?

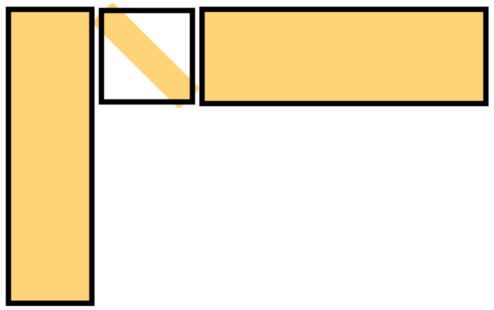
I. efficient construction of low rank approximations to $X$

II. Bounds on singular values of $X$

III. elliptic PDE solvers:
Spectral accuracy, optimal complexity + low rank approximation

# The factored ADI algorithm

$$AX - XB = F \quad \longrightarrow \quad A(ZDY^*) - (ZDY^*)B = USV^* \qquad (S \text{ of size } \rho \times \rho)$$



$$\text{shift parameters } \{\alpha_i\}_{i=1}^k, \ \{\beta_i\}_{i=1}^k$$

$$Z^{(k)} = \left[\ \hat{Z}^{(1)} \mid \hat{Z}^{(2)} \mid \cdots \mid \hat{Z}^{(k)}\ \right], \qquad \begin{cases} \hat{Z}^{(1)} = (A - \beta_1 I)^{-1} US, \\ \hat{Z}^{(i+1)} = (A - \alpha_i I)(A - \beta_{i+1} I)^{-1} Z^{(i)} \end{cases}$$

$$Y^{(k)} = \left[\ \hat{Y}^{(1)} \mid \hat{Y}^{(2)} \mid \cdots \mid \hat{Y}^{(k)}\ \right], \qquad \begin{cases} \hat{Y}^{(1)} = (B^* - \alpha_1 I)^{-1} V, \\ \hat{Y}^{(i+1)} = (B^* - \beta_i I)(B^* - \alpha_{i+1} I)^{-1} Y^{(i)} \end{cases}$$

$$D^{(k)} = \operatorname{diag}\left((\beta_1 - \alpha_1) I_\rho, \cdots, (\beta_k - \alpha_k) I_\rho\right)$$

$$X^{(k)} = Z^{(k)} D^{(k)} {Y^{(k)}}^*$$

$X^{(k)}$ is of rank at most $k\rho$

(Benner, Truhar & Li, 2009), (Li and White, 2002)

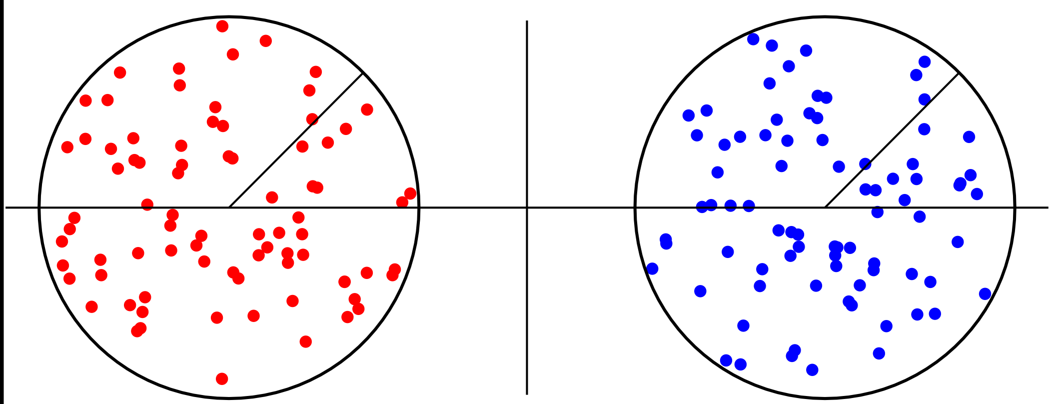
# The factored ADI algorithm

$$A({\color{red}ZDY^*})-({\color{red}ZDY^*})B = USV^* \qquad (S \text{ of size } \rho \times \rho)$$

$$\boxed{A,\ B,\ U,\ S,\ V,\ \{\alpha\}_{j=1}^{k},\ \{\beta\}_{j=1}^{k} \quad \longrightarrow \quad k \text{ steps of fADI} \rightarrow X^{(k)} = Z^{(k)}D^{(k)}Y^{(k)^*}}$$

$$X - X^{(k)} = r_k(A)Xr_k(B)^{-1},\ r_k(z) = \prod_{j=1}^{k} \frac{z-\alpha_j}{z-\beta_j}$$

"ADI-friendly" spectra



$\lambda(A)$ $\lambda(B)$

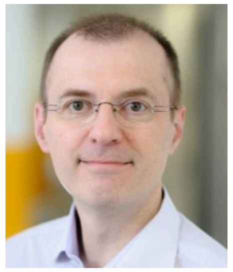
(1) Can choose $k,\ \{\alpha_j\}_{j=1}^{k},\ \{\beta_j\}_{j=1}^{k}$ so that

$$\|X - X^{(k)}\|_2 \leq C\mu^{-k}\|X\|_2$$

(Y. I. Zolotarev)

(2) $X^{(k)}$ is of rank at most $k\rho$

$$\boxed{(1) + (2) \implies \sigma_{k\rho+1}(X) \leq C\mu^{-k}\|X\|_2}$$



(G. Starke)

Key Observation:

$X^{(k)}$ is just one of many possible low rank approximations generated by fADI.

(Beckermann & Townsend, 2017), (Lu & Wachspress, 1991), (Benner, Truhar & Li, 2009), (Li and White, 2002), (Starke, 1992), (Zolotarev, 1877)

# A modification of fADI

$$A(ZDY^*) - (ZDY^*)B = USV^* \quad (S \text{ of size } \rho \times \rho)$$

$X$ can be written as a sum of rank 1 ADI terms:

$$Z^{(k)} = \left[\begin{array}{ccc|ccc|c|ccc} \mathbf{z}_{11} & \cdots & \mathbf{z}_{\rho 1} & \mathbf{z}_{12} & \cdots & \mathbf{z}_{\rho 2} & \cdots & \mathbf{z}_{1k} & \cdots & \mathbf{z}_{\rho k} \end{array}\right]$$

$$Y^{(k)} = \left[\begin{array}{ccc|ccc|c|ccc} \mathbf{y}_{11} & \cdots & \mathbf{y}_{\rho 1} & \mathbf{y}_{12} & \cdots & \mathbf{y}_{\rho 2} & \cdots & \mathbf{y}_{1k} & \cdots & \mathbf{y}_{\rho k} \end{array}\right]$$

$$D^{(k)} = \mathrm{diag}\left((\beta_1 - \alpha_1)I_\rho, \cdots, (\beta_k - \alpha_k)I_\rho\right)$$

$$\boxed{ZDY^* = \sum_{i=1}^{\rho} \sum_{j=1}^{\infty} \underbrace{d_{ij}\mathbf{z}_{ij}\mathbf{y}_{ij}^*}_{T_{ij}}}$$

Choose any $K$ terms to construct a rank $K$ approximation to $X$.

Which choice is best?

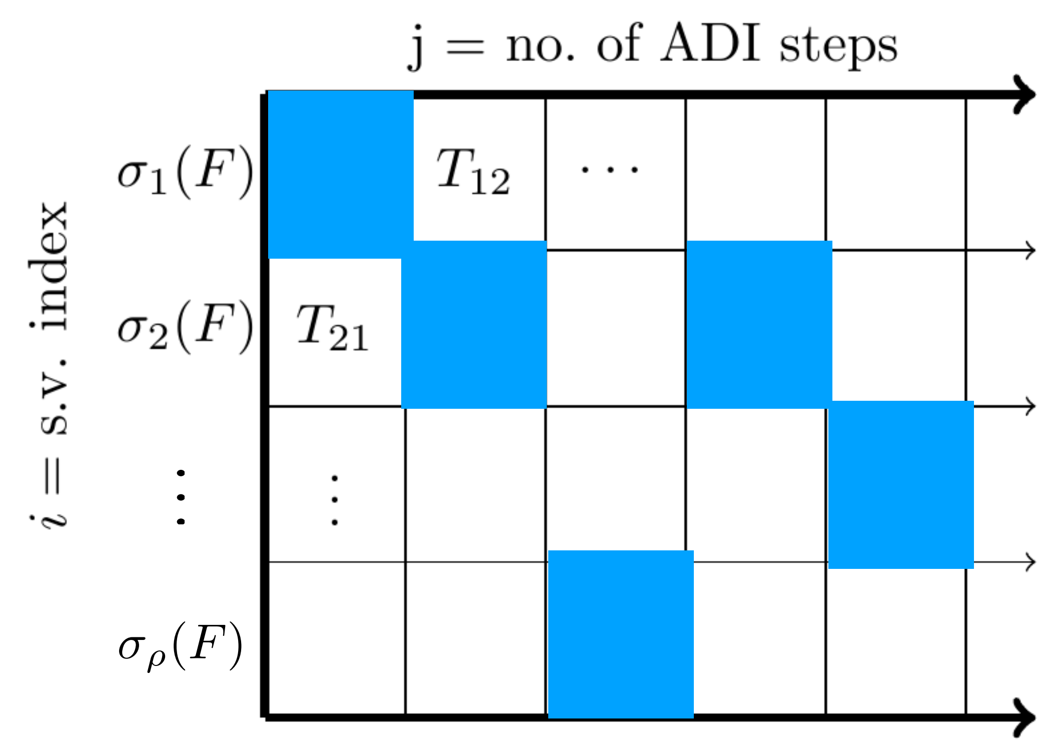
$$\|X - \tilde{X}^{(K)}\|_2 \leq \sum_{T_{ij} \notin \tilde{X}^{(K)}} \|T_{ij}\|_2$$

Key Observation:

$$\|T_{ij}\|_2 \approx C\sigma_i(F)\mu^{-(j-1)}$$

$$\begin{bmatrix} \mathbf{z}_{11} & \cdots & \mathbf{z}_{\rho 1} \end{bmatrix} = (A - \beta_1 I)^{-1} \begin{bmatrix} \sigma_1(F)\mathbf{u}_1 & \cdots & \sigma_\rho(F)\mathbf{u}_\rho \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{z}_{12} & \cdots & \mathbf{z}_{\rho 2} \end{bmatrix} = (A - \alpha_1 I)^{-1}(A - \beta_2 I) \begin{bmatrix} \mathbf{z}_{11} & \cdots & \mathbf{z}_{\rho 1} \end{bmatrix}$$

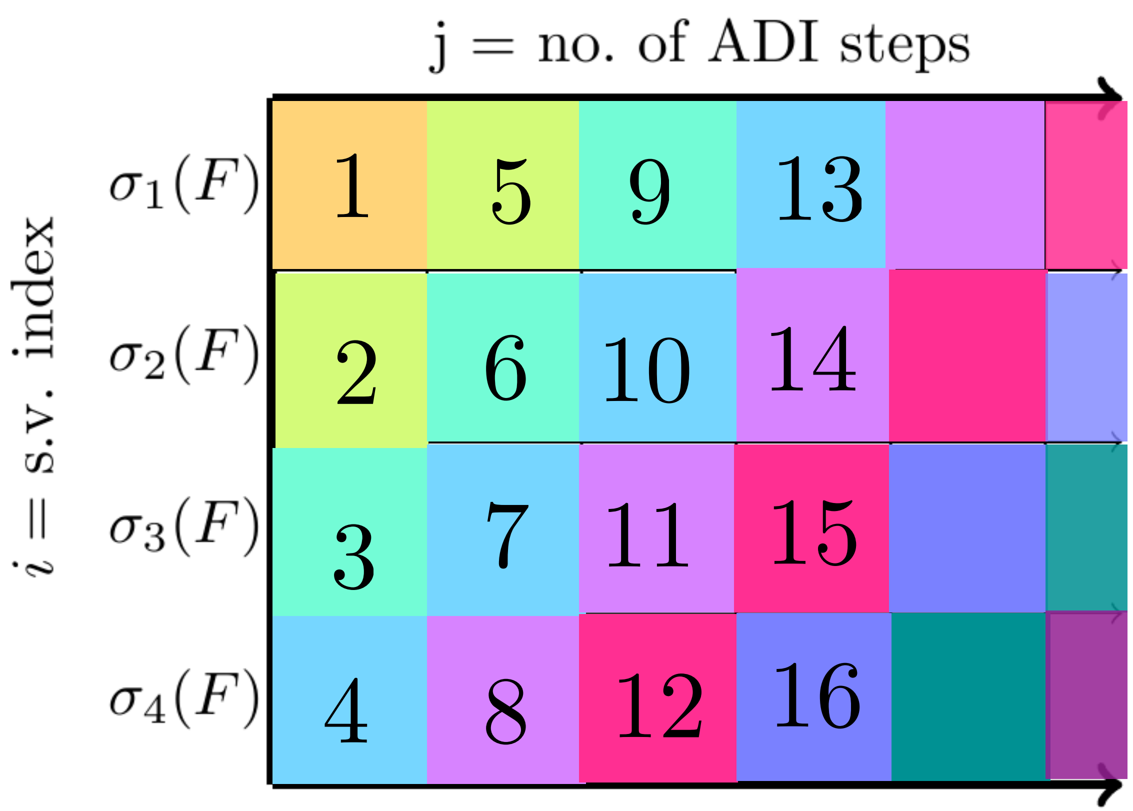


# A modification of fADI

$$A(ZDY^*) - (ZDY^*)B = USV^* \quad (S \text{ of size } \rho \times \rho)$$

Example:

Suppose that $\sigma_{i+1}(F) \leq C\mu^{-i}$, so that $\|T_{ij}\|_2 \approx C\mu^{-(i+j-2)}$




fADI approximant:

$$\|X - X^{(k)}\|_2 \leq \varepsilon \|X\|_2$$

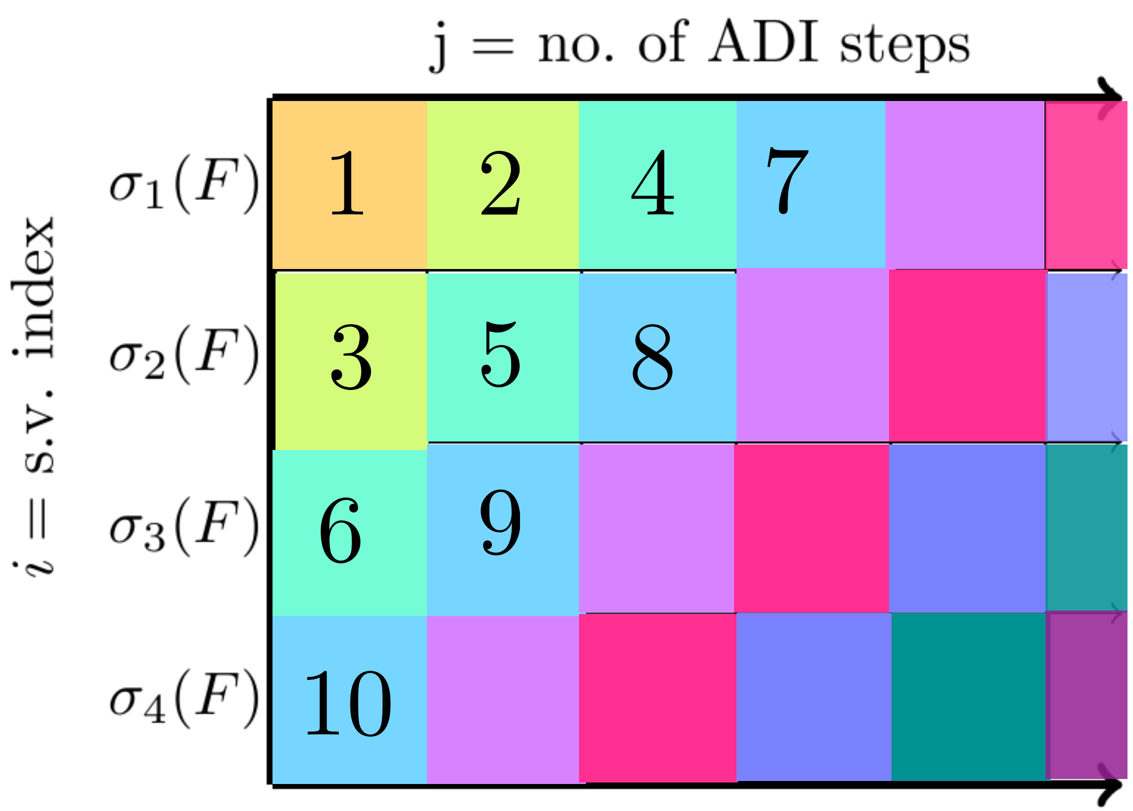
$$\text{rank}(X^{(k)}) \leq k^2$$

# A modification of fADI

$$A(ZDY^*) - (ZDY^*)B = USV^* \quad (S \text{ of size } \rho \times \rho)$$

Example:

Suppose that $\sigma_{i+1}(F) \le C\mu^{-i}$, so that $\|T_{ij}\|_2 \approx C\mu^{-(i+j-2)}$




fADI approximant:

$$\|X - X^{(k)}\|_2 \le \varepsilon \|X\|_2$$

$$\text{rank}(X^{(k)}) \le k^2$$

best choice:

$$\|X - \tilde{X}^{(k)}\|_2 \le \varepsilon \|X\|_2$$

$$\text{rank}(\tilde{X}^{(k)}) \le k(k+1)/2$$

Generalization : Explicit bounds when singular values of F decay geometrically

Automation : FI-ADI algorithm (works with any approximate SVD of F)

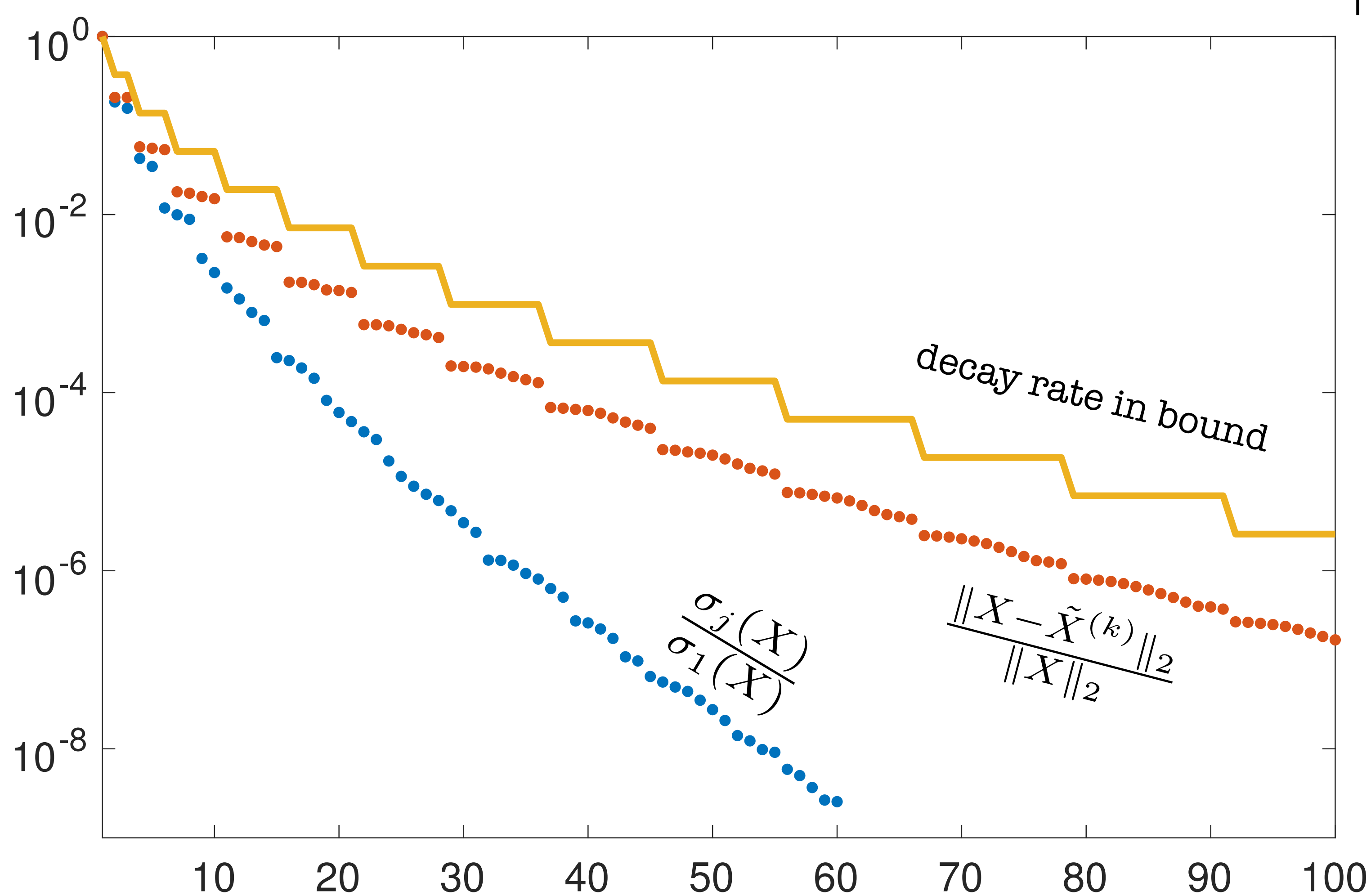
# Example: Bounding singular values

**Theorem:** (Townsend & W.) Let $\lambda(A) \subset E, \lambda(B) \subset -E$, with $E = \{z \in \mathbb{C} : |z - z_0| \leq \eta\}$, and $A, B$ normal matrices. If $AX - XB = F$ and the singular values of $F$ decay at the same rate as the ADI error, then

$$\sigma_{t+1}(X) \leq C_E \mu^{-(\sqrt{8t+1}-1)/2} \|X\|_2, \qquad t = k(k+1)/2 < n.$$

**Example (1):**

$$(C_2)_{j,k} = \frac{1}{|z_j - w_k|^2} \qquad \begin{bmatrix} \bar{z}_1 & & & \\ & \bar{z}_2 & & \\ & & \ddots & \\ & & & \bar{z}_m \end{bmatrix} C_2 - C_2 \begin{bmatrix} \bar{w}_1 & & & \\ & \bar{w}_2 & & \\ & & \ddots & \\ & & & \bar{w}_m \end{bmatrix} = \begin{bmatrix} \frac{1}{z_1 - w_1} & \frac{1}{z_1 - w_2} & \cdots \\ \frac{1}{z_2 - w_1} & \ddots & \\ \vdots & & \end{bmatrix}$$

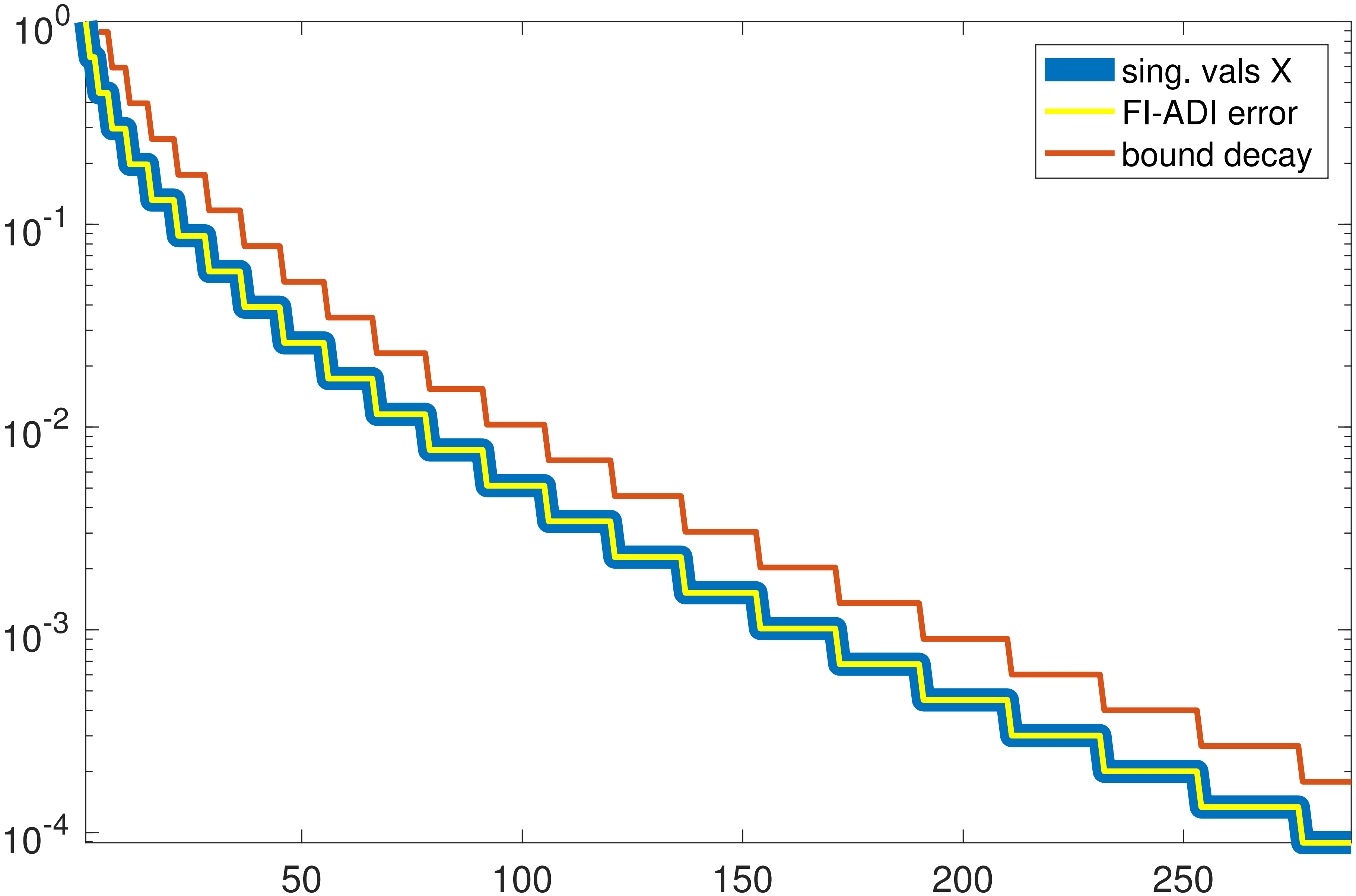


# Example: Bounding singular values

<u>Theorem:</u> (Townsend & W.) Let $\lambda(A) \subset E, \lambda(B) \subset -E$, with $E = \{z \in \mathbb{C} : |z - z_0| \leq \eta\}$, and $A, B$ normal matrices. If $AX - XB = F$ and the singular values of $F$ decay at the same rate as the ADI error, then

$$\sigma_{t+1}(X) \leq C_E \mu^{-(\sqrt{8t+1}-1)/2} \|X\|_2, \qquad t = k(k+1)/2 < n.$$

<u>Example (2):</u> FI-ADI error $\|X - \tilde{X}^{(k)}\|_2$ is a sharp bound.



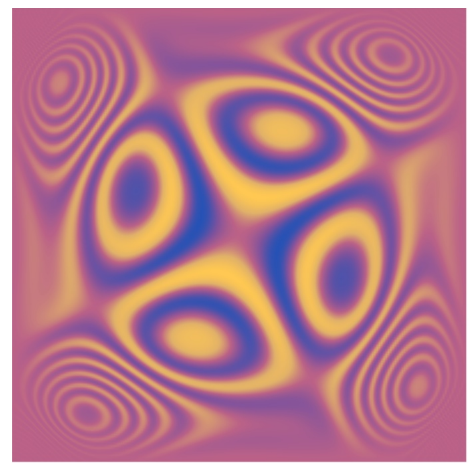
# Application: low rank Poisson solver

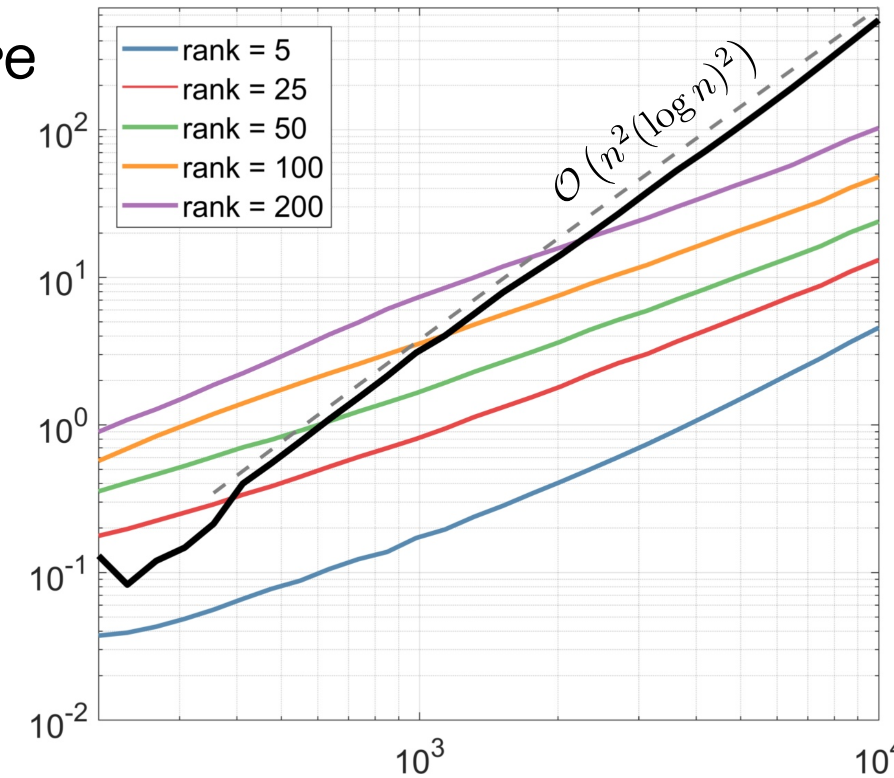
ADI-friendly spectral discretizations → Optimal complexity solvers

## Ultraspherical-based spectral discretization on square

$$\Delta u = f, \quad u(x,y)\Big|_{\partial\Omega} = 0$$

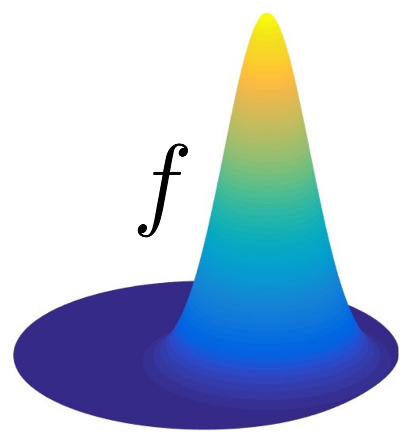
$u$

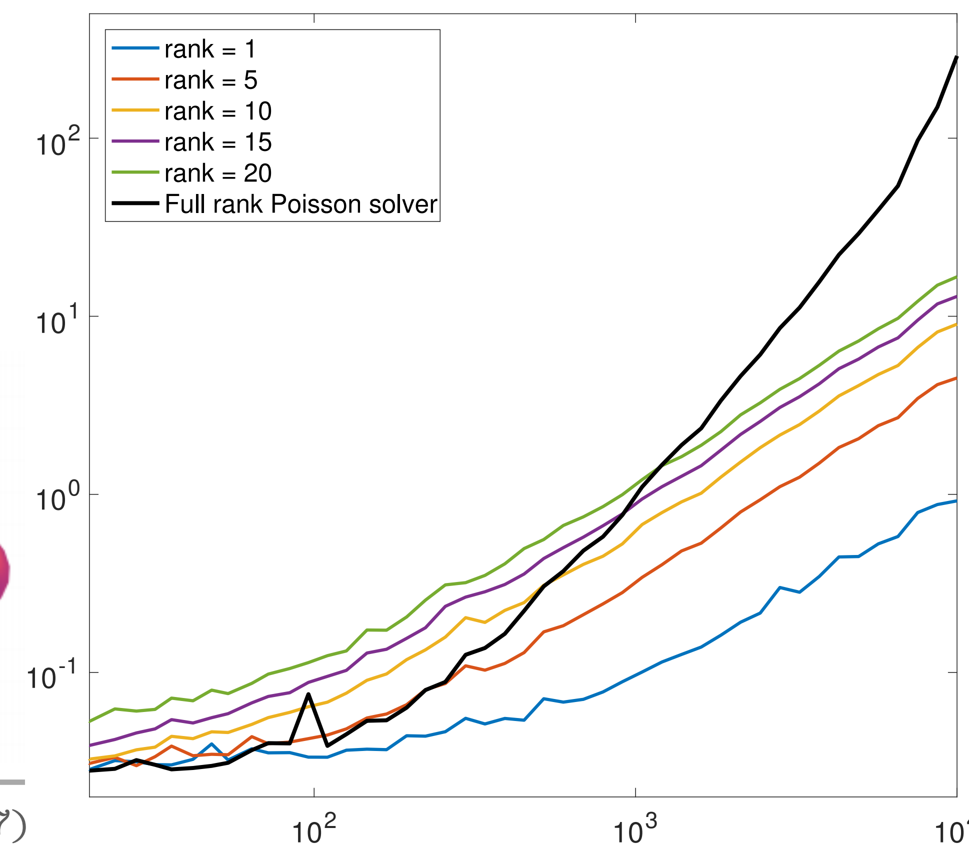




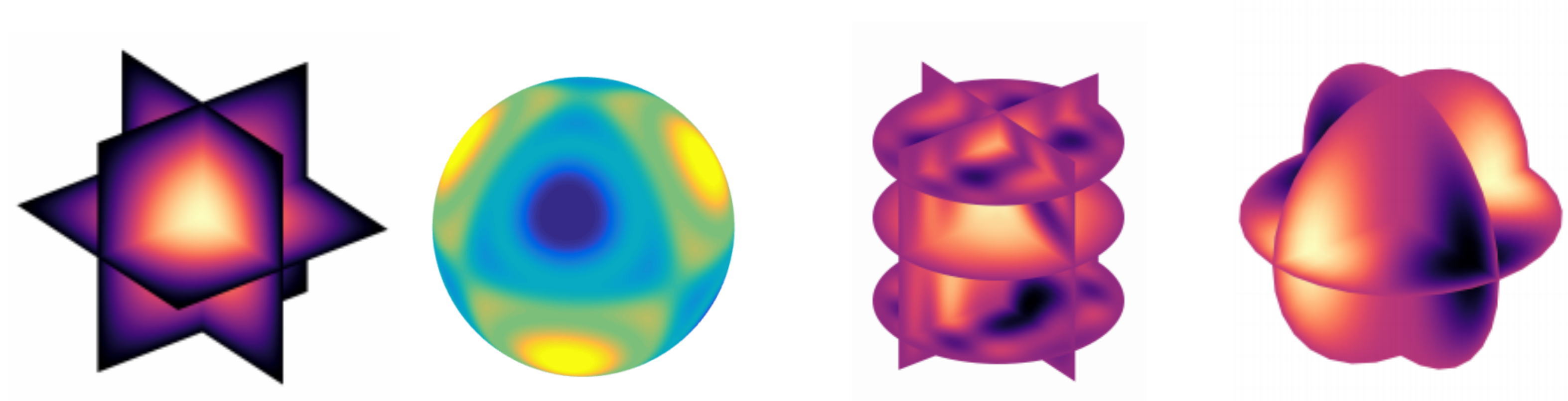

## Ultraspherical spectral discretization on disk

$$\Delta u = f, \quad u(\theta,\rho)\Big|_{\rho=1} = 0$$




works on a variety of domains:

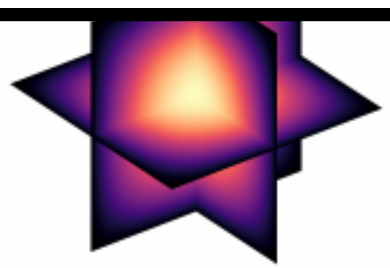
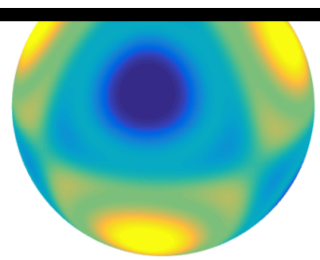
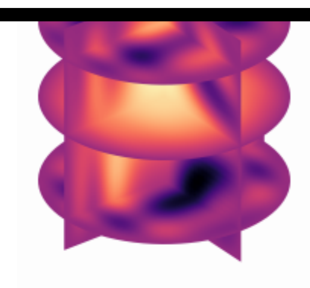
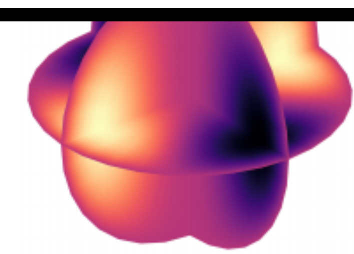
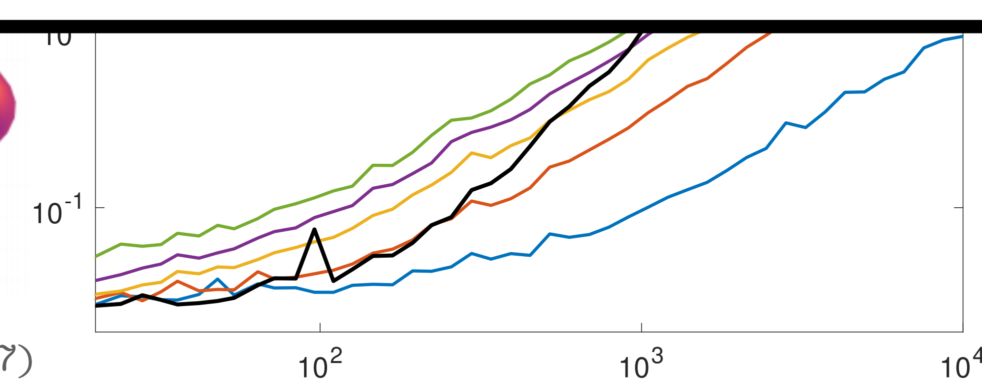


(Fortunato & Townsend, 2017), (Olver & Townsend, 2013), (W., Townsend & Wright, 2017)

# Application: low rank Poisson solver

ADI-friendly spectral discretizations + FI-ADI ⇨ Fast, low rank solvers

Code: - github.com/ajt60gaibb/freeLYAP
- www.chebfun.org


(Fortunato & Townsend, 2017), (Olver & Townsend, 2013), (W., Townsend & Wright, 2017)