Decay pattern of matrices: application to matrix functions (and matrix equations)

V. Simoncini

Dipartimento di Matematica, Università di Bologna (Italy)

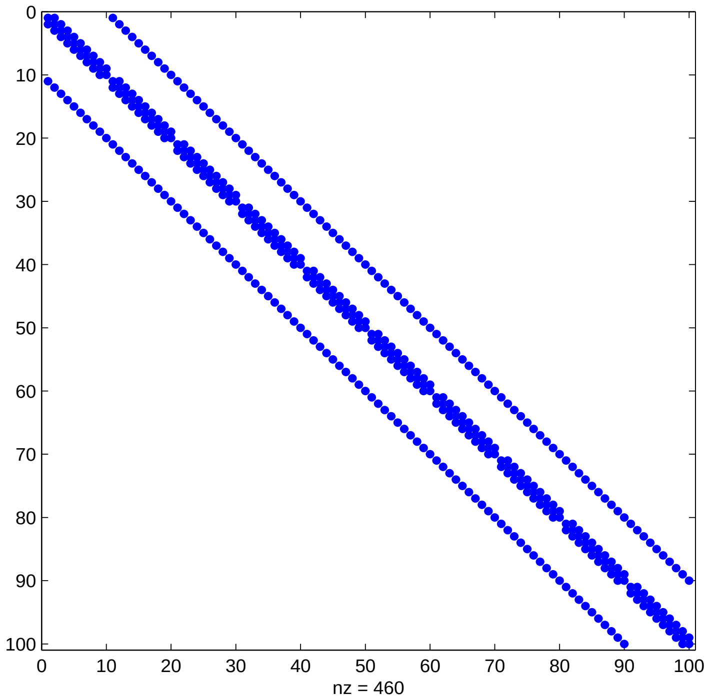
valeria.simoncini@unibo.it

Joint work with Michele Benzi, Emory University (USA)

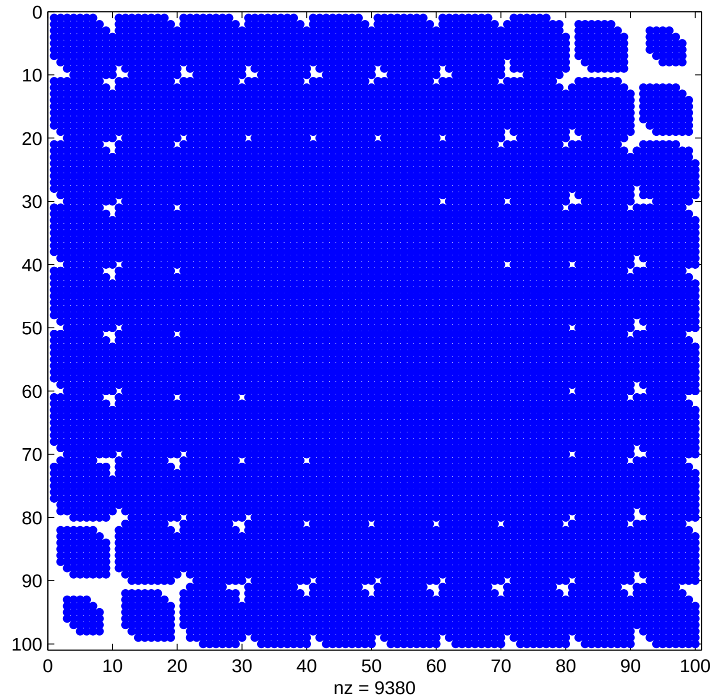
The inverse of the 2D Laplace matrix on the unit square

$$\mathcal{A} := M \otimes I_n + I_n \otimes M, \qquad M = \text{tridiag}(-1, 2, -1)$$

Sparsity pattern:



Matrix $\mathcal{A}$

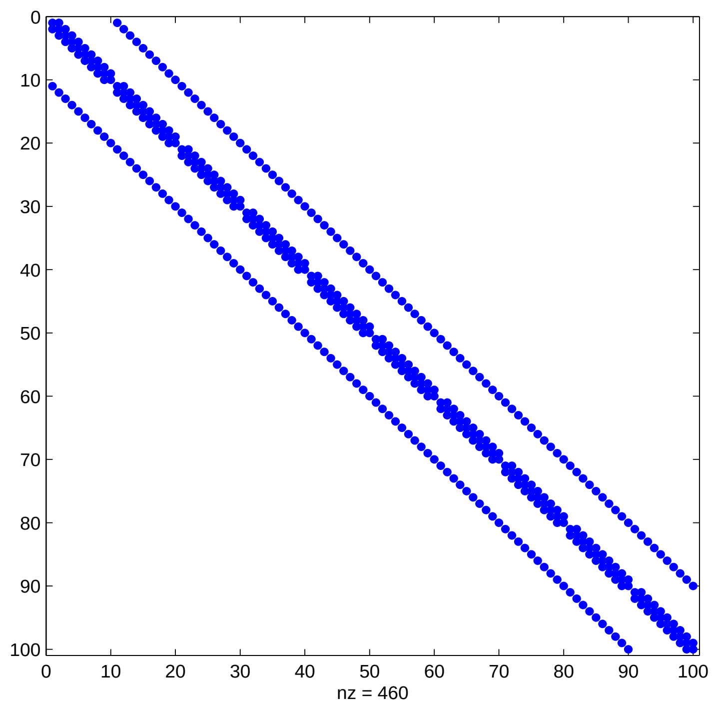


$\mathcal{A}^{-1}$

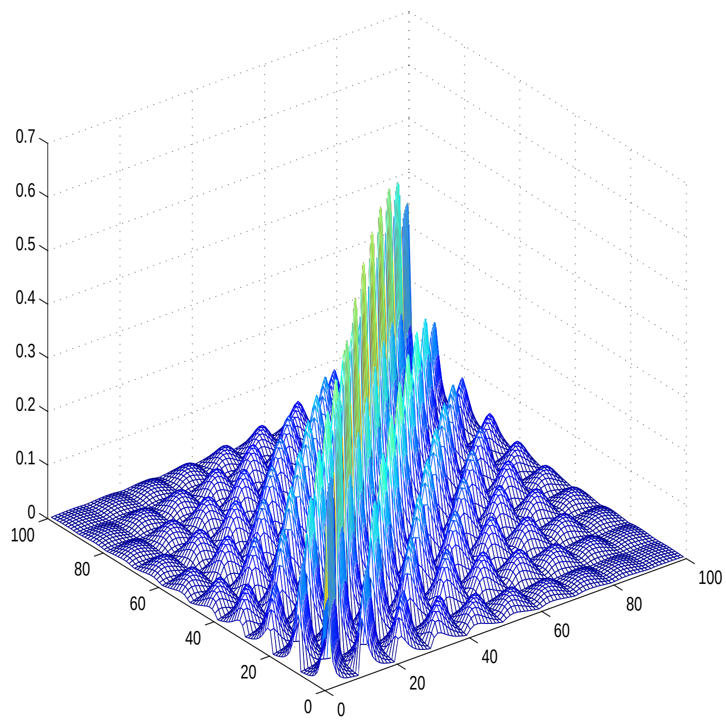
The inverse of the 2D Laplace matrix on the unit square

$$\mathcal{A} := M \otimes I_n + I_n \otimes M, \qquad M = \text{tridiag}(-1, 2, -1)$$

Sparsity pattern:



$\mathcal{A}$



$|(\mathcal{A}^{-1})_{ij}|$

The exponential decay

The classical bound (Demko, Moss & Smith):

If M spd is banded with bandwidth β, then

$$|(M^{-1})_{ij}| \leq \gamma q^{\frac{|i-j|}{\beta}}$$

where $q := \dfrac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} < 1$ (κ=cond(M)) $\qquad \gamma := \max\left\{\frac{1}{\lambda_{\min}(M)}, \frac{(1+\sqrt{\kappa})^2}{2\lambda_{\max}(M)}\right\}$

The exponential decay

The classical bound (Demko, Moss & Smith):

If M spd is banded with bandwidth β, then

$$|(M^{-1})_{ij}| \leq \gamma q^{\frac{|i-j|}{\beta}}$$

where $q := \dfrac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} < 1$ (κ=cond(M)) $\qquad \gamma := \max\left\{\frac{1}{\lambda_{\min}(M)}, \frac{(1+\sqrt{\kappa})^2}{2\lambda_{\max}(M)}\right\}$

If f analytic in region containing $\mathrm{spec}(M)$: $\qquad |f(M)_{ij}| \leq Cq^{\frac{i-j}{\beta}}$

with C, q depending on $\mathrm{spec}(M)$ and f (*Benzi & Golub, 1999*)

Many contributions: Bebendorf, Hackbusch, Benzi, Boito, Razouk, Golub, Tuma, Concus, Meurant, Mastronardi, Ng, Tyrtyshnikov, Nabben, ...

Decay bounds for Cauchy-Stieltjes (or Markov-type) functions

$$f(M) = \int_{-\infty}^{0} (M - \omega I)^{-1} \mathrm{d}\gamma(\omega), \quad \mathrm{spec}(M) \subset \mathbb{C} \setminus (-\infty, 0]$$

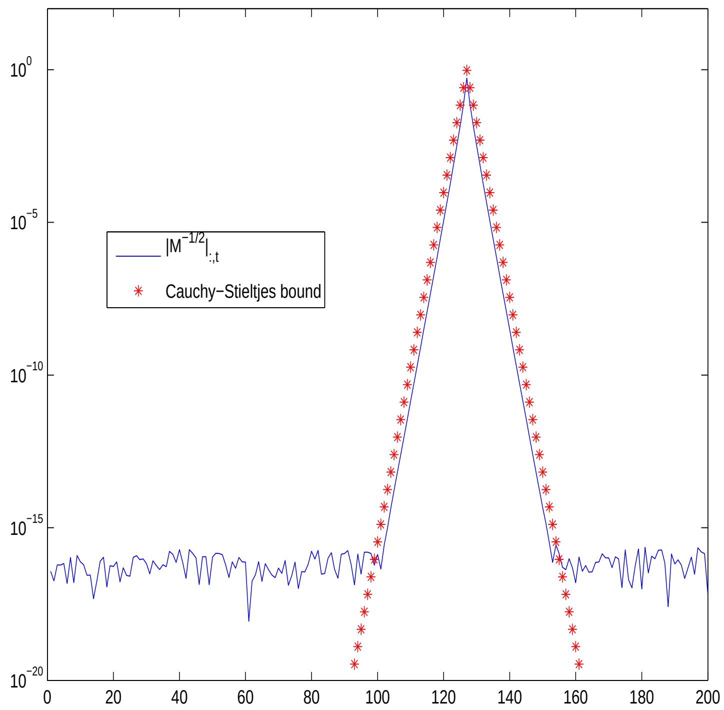
$f(x) = x^{-\frac{1}{2}}$, $f(x) = \frac{e^{-t\sqrt{x}} - 1}{x}$, $f(x) = \frac{\log(1+x)}{x}$, ...

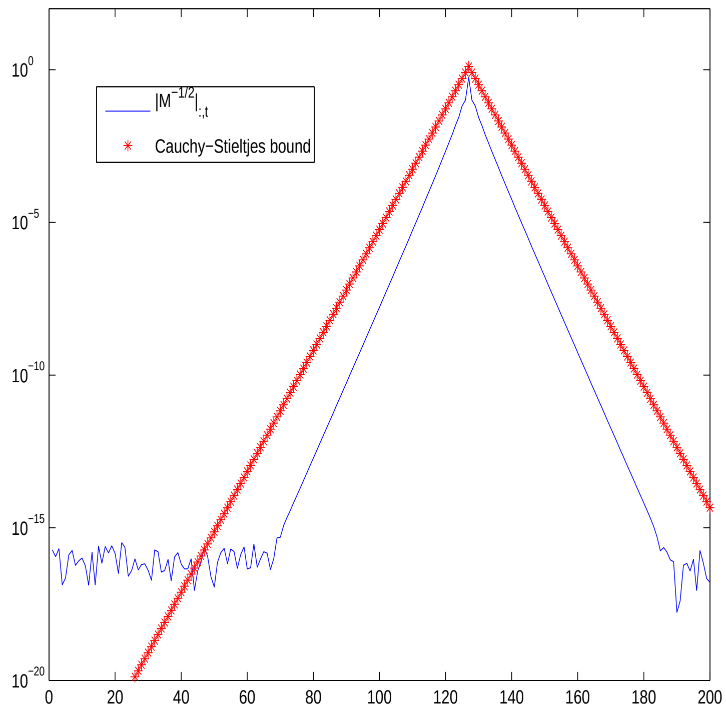
⋆ Demko etal bound useful to estimate $|f(M)|_{kt}$ for M spd and $\beta-$banded:

$$|(M^{-\frac{1}{2}})_{kt}| \leq C \left(\frac{\sqrt{\lambda_{\max}} - \sqrt{\lambda_{\min}}}{\sqrt{\lambda_{\max}} + \sqrt{\lambda_{\min}}} \right)^{\frac{|k-t|}{\beta}}$$

(C depends on spec(M))

Estimates for $|M_{:,t}^{-\frac{1}{2}}|$, $t = 127$, $n = 200$ (log-scale)

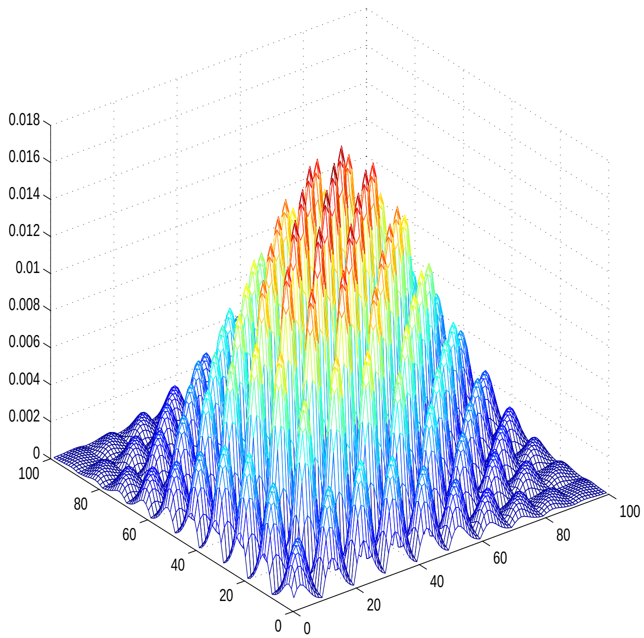


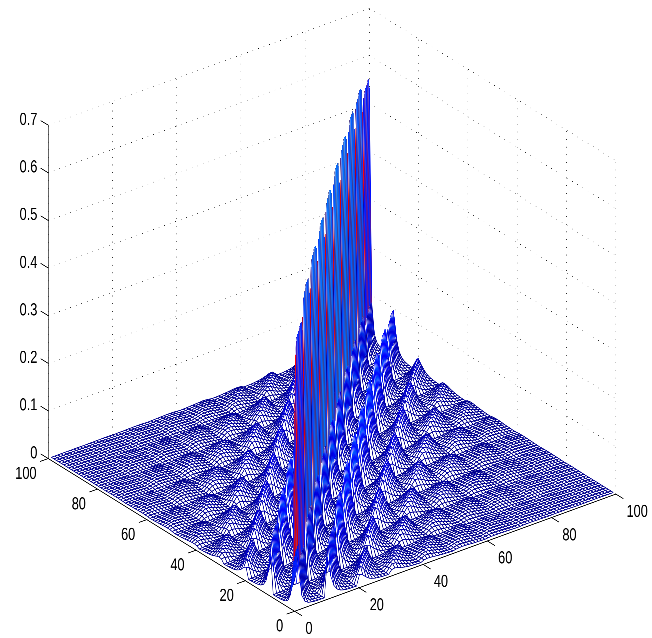


$M = \text{tridiag}(-1, 4, -1)$ $\quad$ $M = \text{pentadiag}(-0.5, -1, 4, -1, -0.5)$

Typical decay plot for $f(\mathcal{A})$

$\mathcal{A}$: Laplace operator as before





$$f(\mathcal{A}) = \exp(-5\mathcal{A}) \qquad\qquad f(\mathcal{A}) = \mathcal{A}^{-1/2}$$

Much richer structure

In general, $\mathcal{A} = M_1 \oplus M_2 := M_1 \otimes I + I \otimes M_2$, M_1, M_2 banded spd

A pause to fix the index notation "on the grid"

$$
\begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{t_1 t_2} \\ \vdots \\ x_{nm} \end{bmatrix} \leftarrow t \qquad \Leftrightarrow \qquad
\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & & \vdots \\ & & x_{t_1 t_2} & x_{2m} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}
$$

$$
t \qquad = \qquad (t_1, t_2)
$$

Decay bounds for the exponential function

Let M be spsd, β-banded; $\mathrm{spec}(M) \subset [0, 4\rho]$,

i) For $\rho\tau \geq 1$ and $\sqrt{4\rho\tau} \leq \lceil \frac{|k_j - t_j|}{\beta} \rceil \leq 2\rho\tau$,

$$|\exp(-\tau M)_{kt}| \leq 10 \exp\left(-\frac{1}{5\rho\tau} \lceil \frac{|k-t|}{\beta} \rceil \right)^2 ;$$

ii) For $\lceil \frac{|k_j - t_j|}{\beta} \rceil \geq 2\rho\tau$,

$$|\exp(-\tau M)_{kt}| \leq 10 \frac{\exp(-\rho\tau)}{\rho\tau} \left(\frac{\mathrm{e}\rho\tau}{\lceil \frac{|k-t|}{\beta} \rceil} \right)^{\lceil \frac{|k-t|}{\beta} \rceil}$$

Decay bounds for the exponential function

Let M be spsd, β-banded; $\mathrm{spec}(M) \subset [0, 4\rho]$,

i) For $\rho\tau \geq 1$ and $\sqrt{4\rho\tau} \leq \lceil \frac{|k_j - t_j|}{\beta} \rceil \leq 2\rho\tau$,

$$|\exp(-\tau M)_{kt}| \leq 10 \exp\left(-\frac{1}{5\rho\tau}\lceil\frac{|k-t|}{\beta}\rceil\right)^2;$$

ii) For $\lceil \frac{|k_j - t_j|}{\beta} \rceil \geq 2\rho\tau$,

$$|\exp(-\tau M)_{kt}| \leq 10 \frac{\exp(-\rho\tau)}{\rho\tau}\left(\frac{e\rho\tau}{\lceil\frac{|k-t|}{\beta}\rceil}\right)^{\lceil\frac{|k-t|}{\beta}\rceil}$$

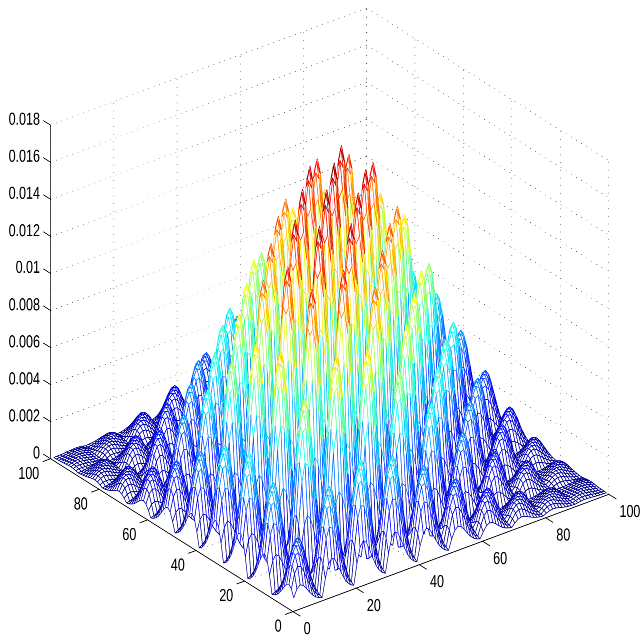
$$\text{Keynote formula :} \qquad \exp(M_1 \oplus M_2) = \exp(M_1) \otimes \exp(M_2)$$

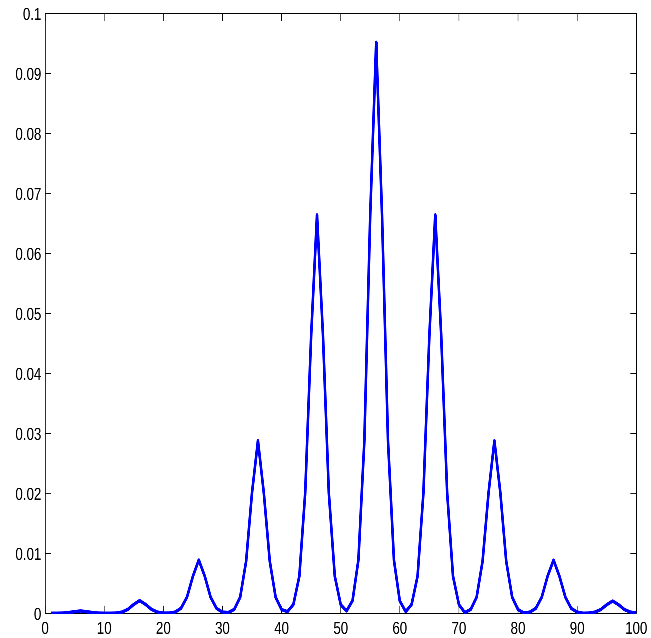
$\mathcal{A} = I \otimes M + M \otimes I$. Then

$$(\exp(-\tau\mathcal{A}))_{kt} = (\exp(-\tau M))_{k_1 t_1} (\exp(-\tau M))_{k_2 t_2}$$

for all $t = (t_1, t_2)$ and $k = (k_1, k_2)$ with $\min\{|t_1 - k_1|, |t_2 - k_2|\} \geq \sqrt{4\rho\tau}\beta$

Decay bounds for the exponential function





Left: whole pattern of $\exp(-\mathcal{A})$ Right: Row 56 of $\exp(-\mathcal{A})$

$|\exp(-\mathcal{A})_{kt}|$ with $k = 56 \Rightarrow k = (k_1, k_2) = (6, 5)$

For $t = 50 \Rightarrow t = (t_1, t_2) = (10, 4)$ so that $|k_1 - t_1| \gg 0$

For $t = 45 \Rightarrow t = (t_1, t_2) = (5, 4)$ so that $|k_1 - t_1| \not\gg 0$

Decay bounds for Laplace-Stieltjes function

$$f(M) = \int_0^\infty e^{-\tau M} \mathrm{d}\alpha(\tau)$$

e.g., $f(x) = x^{-\sigma}$ ($\sigma > 0$), $f(x) = e^{-x}$, $f(x) = e^{1/x}$, $f(x) = (1 - e^{-x})/x$, $f(x) = \log(1 + 1/x)$, ...

- For M spd and β-banded, $\widehat{M} = M - \lambda_{\min} I$

$$|f(M)_{k,t}| \le \int_0^\infty \exp(-\lambda_{\min}\tau) |(\exp(-\tau \widehat{M}))_{k,t}| \mathrm{d}\alpha(\tau)$$

Decay bounds for Laplace-Stieltjes function

$$f(M) = \int_0^\infty e^{-\tau M} \mathrm{d}\alpha(\tau)$$

e.g., $f(x) = x^{-\sigma}$ ($\sigma > 0$), $f(x) = e^{-x}$, $f(x) = e^{1/x}$, $f(x) = (1 - e^{-x})/x$, $f(x) = \log(1 + 1/x)$, ...

- For M spd and β-banded, $\widehat{M} = M - \lambda_{\min} I$

$$|f(M)_{k,t}| \leq \int_0^\infty \exp(-\lambda_{\min}\tau) |(\exp(-\tau \widehat{M}))_{k,t}| \mathrm{d}\alpha(\tau)$$

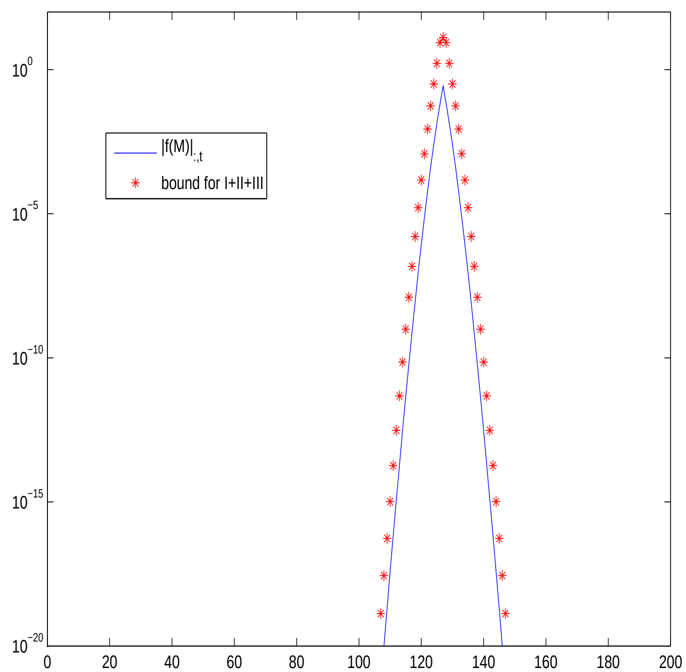
- For $\mathcal{A} = M \otimes I + I \otimes M$

$$(f(\mathcal{A}))_{kt} = \int_0^\infty (\exp(-\tau M))_{k_1 t_1} (\exp(-\tau M))_{t_2 k_2} \mathrm{d}\alpha(\tau)$$

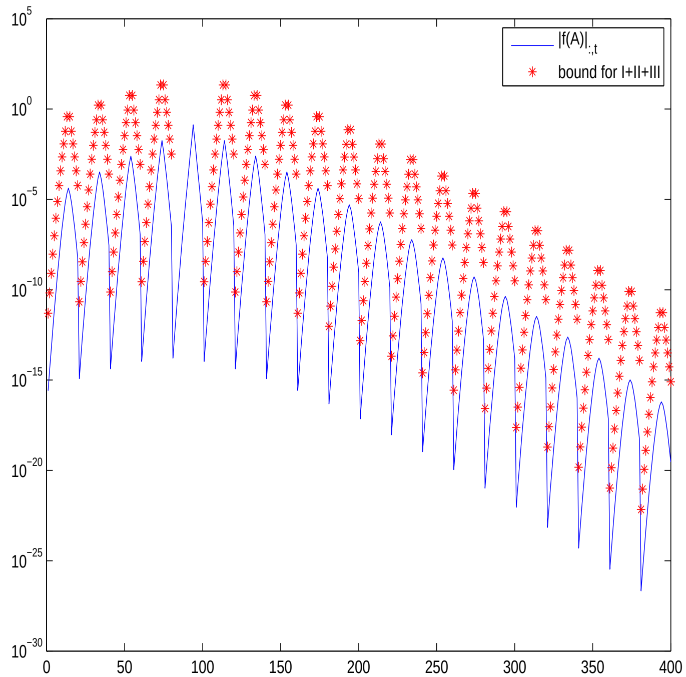
then, more precise bounds for specific choices of $\mathrm{d}\alpha(\tau)$

An example: $f(x) = \frac{1-e^{-x}}{x}$

$M = \mathrm{tridiag}(-1, 4, -1)$, $n = 200$



Typical row of $f(M)$



Typical row of $f(\mathcal{A})$

Cauchy-Stieltjes functions of Kronecker sum: $f(\mathcal{A}) = \int_\Gamma (\mathcal{A} - \omega I)^{-1} \mathrm{d}\gamma(\omega)$

$$e_k^T f(\mathcal{A}) e_t = \int_\Gamma e_k^T (\mathcal{A} - \omega I)^{-1} e_t \mathrm{d}\gamma(\omega),$$

where we can write $\mathcal{A} - \omega I = M \otimes I + I \otimes (M - \omega I)$

Cauchy-Stieltjes functions of Kronecker sum: $f(\mathcal{A}) = \int_\Gamma (\mathcal{A} - \omega I)^{-1} \mathrm{d}\gamma(\omega)$

$$e_k^T f(\mathcal{A}) e_t = \int_\Gamma e_k^T (\mathcal{A} - \omega I)^{-1} e_t \mathrm{d}\gamma(\omega),$$

where we can write $\mathcal{A} - \omega I = M \otimes I + I \otimes (M - \omega I)$

- For each t, $x_t := (\mathcal{A} - \omega I)^{-1} e_t$, so that $X_t = X_t(\omega) \in \mathbb{C}^{n \times n}$ solution to

$$M X_t + X_t (M - \omega I) = E_t, \qquad x_t = \mathrm{vec}(X_t), \quad e_t = \mathrm{vec}(E_t)$$

Then (e.g., Lancaster 1970)

$$X_t = -\int_0^\infty \exp(-\tau M) E_t \exp(-\tau(M - \omega I)) \mathrm{d}\tau$$

so that (with $k = (k_1, k_2), t = (t_, t_2)$)

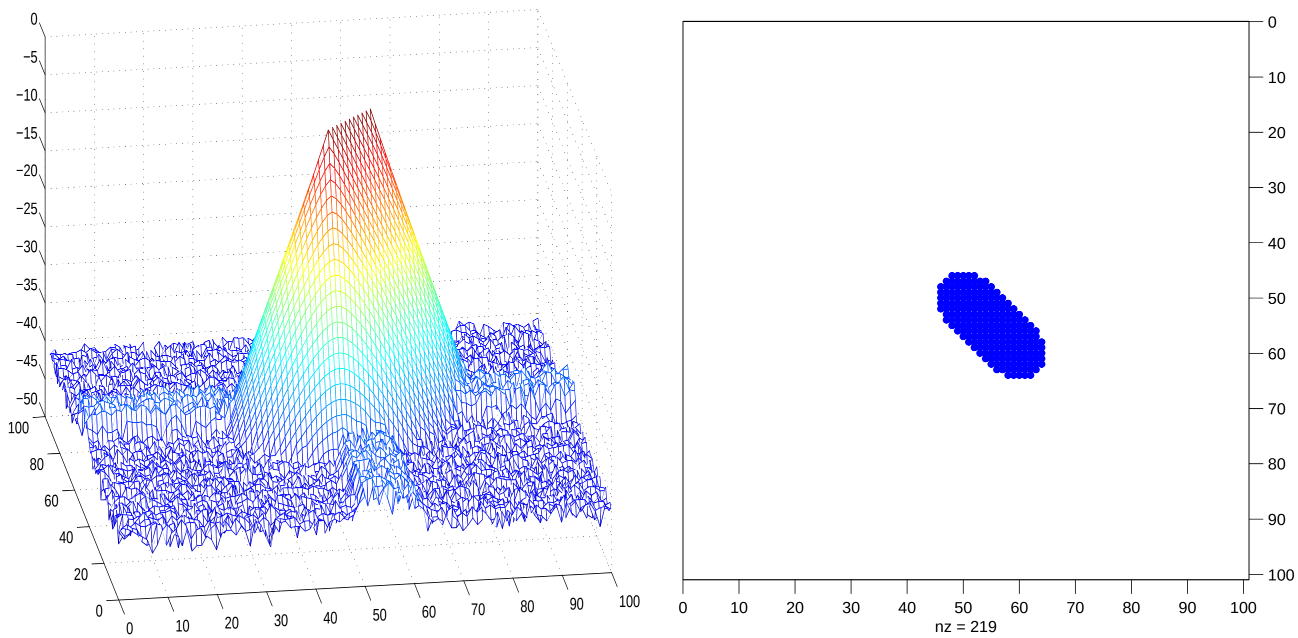
$$e_k^T (\omega I - \mathcal{A})^{-1} e_t = e_{k_1}^T X_t e_{k_2} = -\int_0^\infty |\exp(-\tau M)_{k_1,t_1}|| \exp(-\tau(M - \omega I))_{t_2,k_2}| \mathrm{d}\tau$$

then, more precise bounds for specific choices of f...

More applications. Using sparsity in solution strategies

$$MX + XM = BB^T$$

$M = \mathrm{tridiag}(-1, 4, -1) \in \mathbb{R}^{n\times n}$, $n = 100$ and $B = [e_{50}, \ldots, e_{60}]$



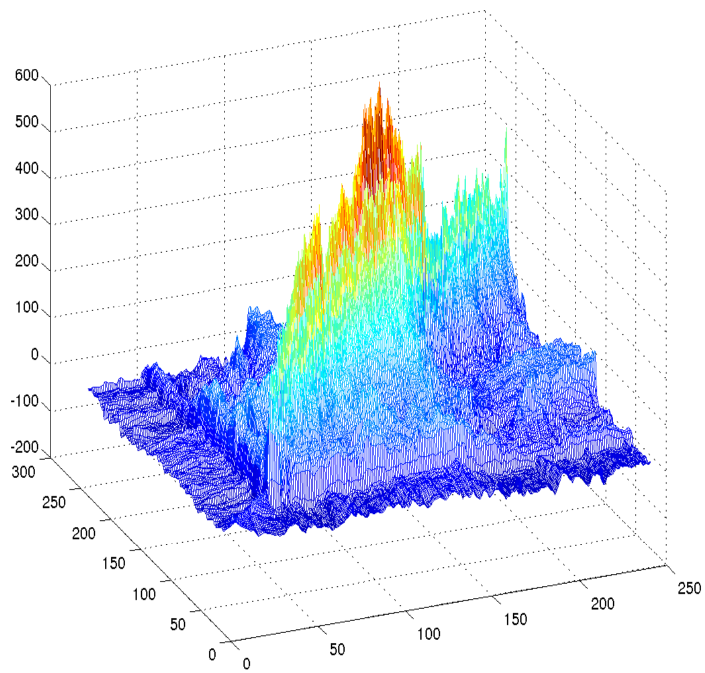
Left: pattern of X with log scale, $\mathtt{nnz}(X) = 9724$

Right: Sparsity pattern of truncated ver. of X: all entries below 10^{-5} are omitted

More applications. Images

M: image $\qquad A^{\frac{1}{2}}$ with $A = M^T M$

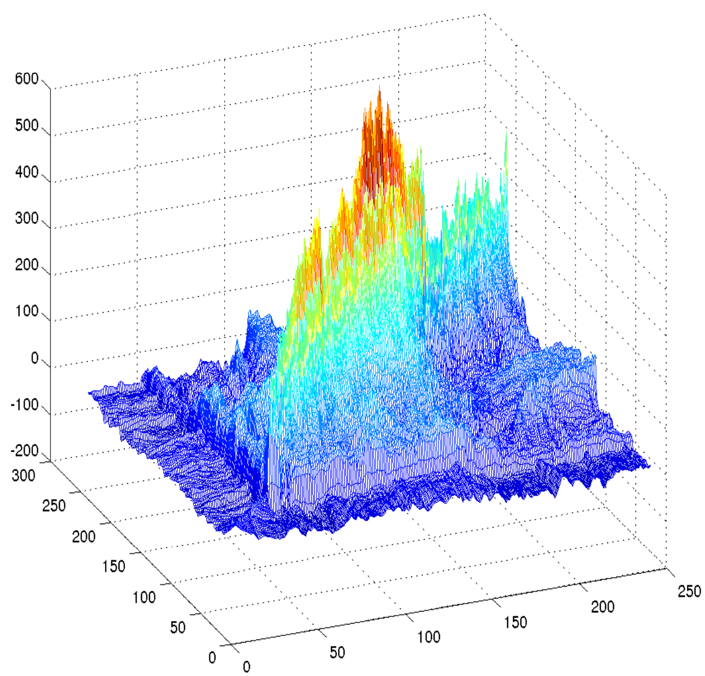
A “more than a man” structure



More applications. Images

M: image $\qquad A^{\frac{1}{2}}$ with $A = M^T M$

A "more than a man" structure



Conclusions and outlook

- Exploring/Exploiting structure is beneficial
- Generalization to d-Kronecker sum is possible, e.g.,

$$\mathcal{A} = M_1 \otimes I \otimes I + I \otimes M_2 \otimes I + I \otimes I \otimes M_3$$

- Possibility of using quasi-sparsity (decay) information in applications ?
  (already done for $f(x) = x^{-1}$)

Conclusions and outlook

- Exploring/Exploiting structure is beneficial
- Generalization to d-Kronecker sum is possible, e.g.,

$$\mathcal{A} = M_1 \otimes I \otimes I + I \otimes M_2 \otimes I + I \otimes I \otimes M_3$$

- Possibility of using quasi-sparsity information in applications ?
(already done for $f(x) = x^{-1}$)

REFERENCES

1. V. Simoncini
The Lyapunov matrix equation. Matrix analysis from a computational perspective
pp. 1-14, Dip. Matematica, UniBo, Jan. 2015. arXiv:1501.07564

2. Michele Benzi and V. Simoncini
Decay bounds for functions of matrices with banded or Kronecker structure
pp.1-20, Dip. Matematica, UniBo, Jan. 2015. arXiv:1501.07376

3. Michele Benzi and V. Simoncini
Approximation of functions of large matrices with Kronecker structure
pp.1-21, Dip. Matematica, UniBo, March 2015. arXiv:1503.02615