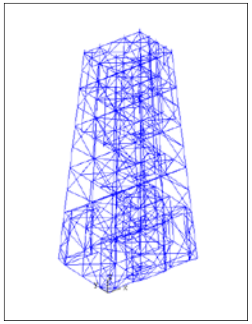
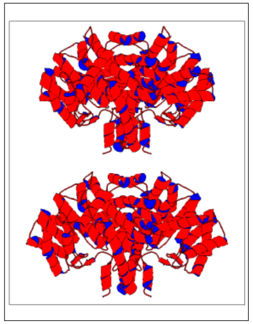
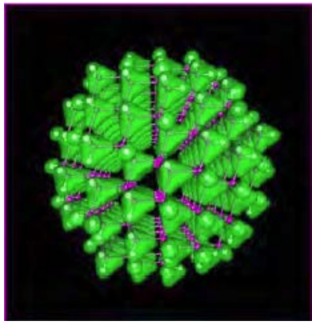
CCS-LBNL Collaborative Workshop 2018
March 5-6, 2018

Solving Nonlinear Eigenvalue Problems with Rational Approximations

Osni Marques
Lawrence Berkeley National Laboratory
oamarques@lbl.gov

Joint work with Zhaojun Bai (UC Davis)
Thanks to Roel Van Beeumen (LBNL) and Jacob Johnson (UC Davis)

Eigenvalue Problems: Linear versus Nonlinear

- Linear:
  - $Ax = \lambda x$ or $Ax = \lambda Bx$
  - Properties well-understood
  - Numerical techniques available
- Nonlinear
  - $F(\lambda)x = 0$
  - Typical solution strategy:
    1. approximation of nonlinear functions by polynomial/rational functions
       $F(\lambda) \approx P(\lambda) + \sum_{j=1}^{k} \frac{p_j(\lambda)}{q_j(\lambda)} E_j$
    2. linearization of the polynomial/rational eigenvalue problem (reformulation of the polynomial/rational eigenvalue problem as a larger generalized linear eigenvalue problem with the same eigenvalues)
    3. solution of the obtained generalized eigenvalue problem by standard methods

(Short List of) Related Work

- Ruhe, Algorithms for the Nonlinear Eigenvalue Problem, SIAM J. Numer. Anal., 10:674–689, 1973.
- Mackey, Mackey, Mehl, and Mehrmann. Vector Spaces of Linearizations for Matrix Polynomials. SIAM J. Matrix Anal. Appl., 28:971–1004, 2006.
- Su and Bai, Solving Rational Eigenvalue Problems via Linearization, SIAM. J. Matrix Anal. Appl., 32, 2011.
- Roel Van Beeumen, Rational Krylov Methods for Nonlinear Eigenvalue Problems, KU Leuven, 2015.
- Güttel and Tisseur, The Nonlinear Eigenvalue Problem, Acta Numerica, 26:1-94, 2017.

- Special Case: *polynomial eigenvalue problem*

$$\left(\lambda^k A_k + \lambda^{k-1} A_{k-1} + \cdots + \lambda A_1 + A_0\right) x = 0$$

problem of dimension *n*

$$\left(\lambda \begin{bmatrix} A_k & 0 & \cdots & 0 \\ 0 & I_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I_n \end{bmatrix} + \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I_n & 0 \end{bmatrix}\right) \begin{bmatrix} \lambda^{k-1} x \\ \vdots \\ \lambda x \\ x \end{bmatrix} = 0$$

problem of dimension *nk*

A Sample of Nonlinear Eigenvalue Problems

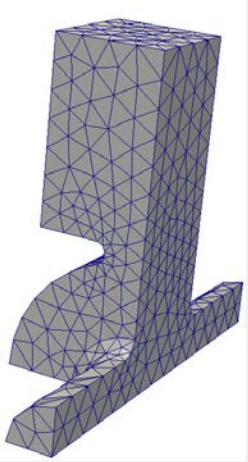
Betcke, Higham, Mehrmann, Schröder and Tisseur, NLEVP: A Collection of Nonlinear Eigenvalue Problems, ACM TOMS, 39, 2013

- $F(\lambda) = \lambda^2 M + \lambda C + K$
- $F(\lambda) = \lambda^2 M + \lambda C + (i + \mu)K$
- $F(\lambda) = A - \lambda I + s(\lambda)B$
- $F(\lambda) = (e^{\lambda} - 1)A_2 + \lambda^2 A_1 - \alpha A_0$
- $F(\lambda) = A - \lambda B + \frac{\lambda}{\lambda - \sigma} C$

$\vdots$

- $F(\lambda) = K - \lambda M + i\sqrt{\lambda - \sigma_1^2}\, W_1 + i\sqrt{\lambda - \sigma_2^2}\, W_2$
- $F(\lambda) = K - \lambda M + i\sqrt{\lambda - \sigma_1^2}\, W_1 + i\frac{\lambda}{\sqrt{\lambda - \sigma_2^2}} W_2$

Lee, Li, Ng, and Ko, Omega3P: A Parallel Finite-Element Eigenmode Analysis Code for Accelerator Cavities SLAC-PUB-13529, 2009.



$$F(\lambda) = K - \lambda M + i\sqrt{\lambda - \sigma_1^2}W_1 + i\sqrt{\lambda - \sigma_2^2}W_2$$

- Find solutions through and associated (linearized) eigenvalue problem
- Explore low-rank properties of W_k
  - ❖ $W_k \approx L_k U_k^T = Q_k R_k$
- Padé approximants for $\sqrt{\lambda - \sigma_k^2}$
  - ❖ $\sqrt{\lambda - \sigma_k^2} \approx \frac{p(\lambda)}{q(\lambda)} = a^T(C - \lambda D)^{-1} b$
  - ❖ $\frac{p(\lambda)}{q(\lambda)}$ of order $[d_p, d_q]$
  - ❖ $d_q = d_p + 1$ (easier to implement)

- *proper rational function*
- *minimal realization:*
  a and b are d_q-by-1

$$C = \begin{bmatrix} * & * & * & * & * & \cdots \\ 1 & 0 & & & & \\ & 1 & 0 & & & \\ & & 1 & 0 & & \\ & & & 1 & 0 & \\ & & & & \ddots & \ddots \end{bmatrix}$$

$$D = \begin{bmatrix} * & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & \ddots \end{bmatrix}$$

Linearization: $k = 1$ (for simplicity)

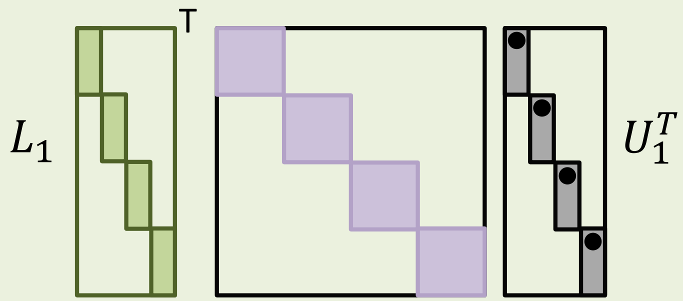
Su and Bai, 2011

- $W_k \approx L_k U_k^T$
- $\frac{p(\lambda)}{q(\lambda)} = a^T(C - \lambda D)^{-1} b$

- $Kx - \lambda Mx + iL_1(I_{r_1} \otimes a)^T (I_{r_1} \otimes C - \lambda I_{r_1} \otimes D)^{-1} (I_{r_1} \otimes b) U_1^T x = 0$
- $i(I_{r_1} \otimes C - \lambda I_{r_1} \otimes D)^{-1} (I_{r_1} \otimes b) U_1^T x = x_1$

  $-i(I_{r_1} \otimes b) U_1^T x + (I_{r_1} \otimes C) x_1 = \lambda (I_{r_1} \otimes D) x_1$
- $\begin{bmatrix} K & L_1(I_{r_1} \otimes a)^T \\ -i(I_{r_1} \otimes b)U_1^T & (I_{r_1} \otimes C) \end{bmatrix} \begin{Bmatrix} x \\ x_1 \end{Bmatrix} = \lambda \begin{bmatrix} M & \\ & (I_{r_1} \otimes D) \end{bmatrix} \begin{Bmatrix} x \\ x_1 \end{Bmatrix}$
- $\begin{bmatrix} K & L \\ U^T & \hat{C} \end{bmatrix} \hat{x} = \lambda \begin{bmatrix} M & 0 \\ 0 & \hat{D} \end{bmatrix} \hat{x}$

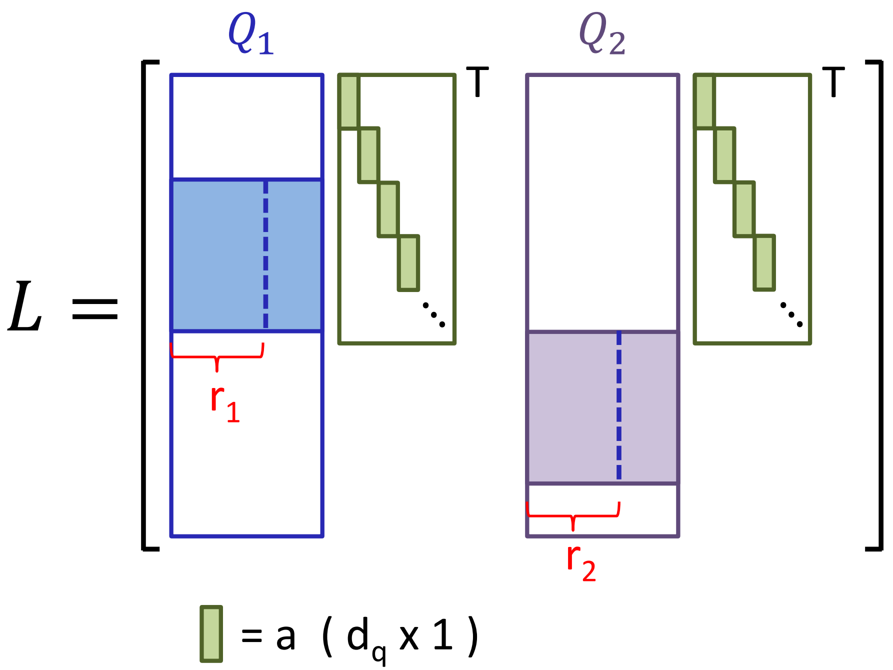
Ex. n−by−4 L_1 and U_1

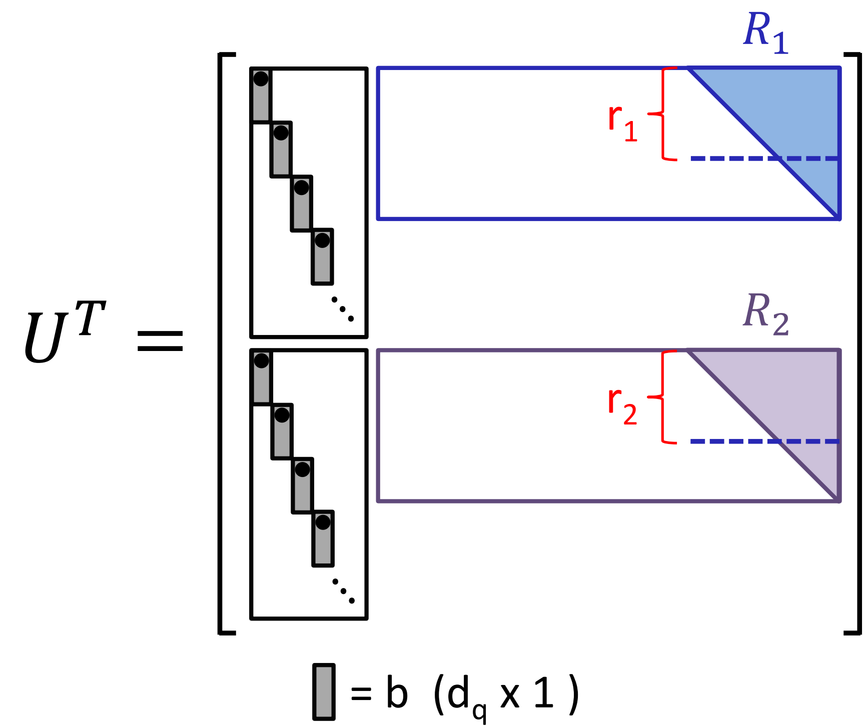


Linearization: $k = 2$

$$\begin{bmatrix} K & L \\ U^T & \hat{C} \end{bmatrix} \hat{x} = \lambda \begin{bmatrix} M & 0 \\ 0 & \widehat{D} \end{bmatrix} \hat{x}$$

$$\hat{C} = \begin{bmatrix} \hat{C}_1 & \\ & \hat{C}_2 \end{bmatrix} \qquad \widehat{D} = \begin{bmatrix} \widehat{D}_1 & \\ & \widehat{D}_2 \end{bmatrix}$$





Solving (the linear problem) $A\hat{x} = \hat{\lambda} B \hat{x}$

- Shift-and-invert: ARPACK + SuperLU

  $t = (A - \sigma B)^{-1} z$

- Strategy 1:

  Factor $A - \sigma B = \hat{L}\hat{U}$ (e.g.)

- Strategy 2:

$$\begin{bmatrix} K - \sigma M & L \\ R^T & \hat{C} - \sigma \hat{D} \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = \begin{Bmatrix} b_1 \\ b_2 \end{Bmatrix}$$

$$u = [b_2 - R^T (K - \sigma M)^{-1} b_1]$$

$$x_2 = \left[\left(\hat{C} - \sigma \hat{D}\right) - R^T (K - \sigma M)^{-1} L\right]^{-1} u$$

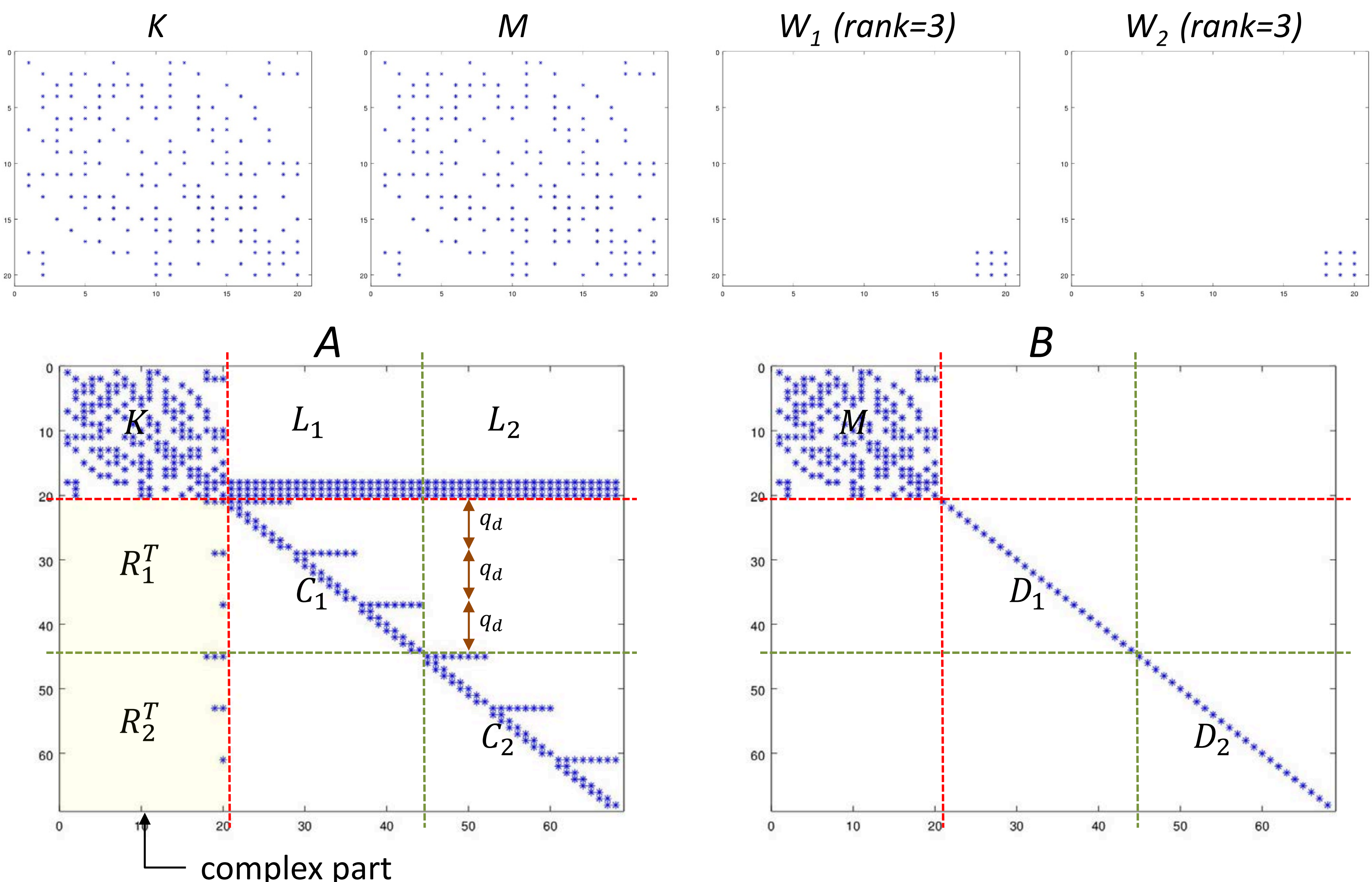
$$x_1 = (K - \sigma M)^{-1} (b_1 - L x_2)$$

$\hat{A} = \hat{C} - \sigma \hat{D}$
$U = R^T$
$V^T = (K - \sigma M)^{-1} L$
$\hat{A}^{-1}$ is block diagonal

- Sherman-Morrison-Woodbury:

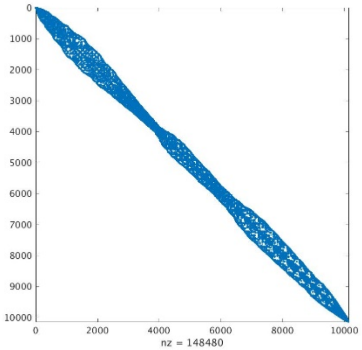
$$\left(\hat{A} - UV^T\right)^{-1} = \hat{A}^{-1} - \hat{A}^{-1} U \left(I + V^T \hat{A}^{-1} U\right)^{-1} V^T \hat{A}^{-1}$$

Synthetic Problem: $n = 20,\ d_p = 7$

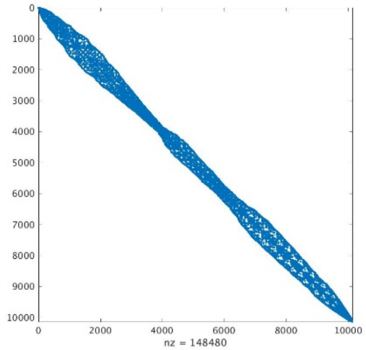


slacl problem: $n = 10142,\ d_p = 7$

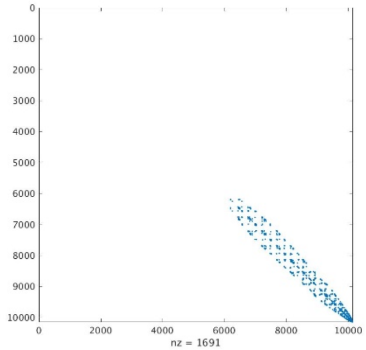
K



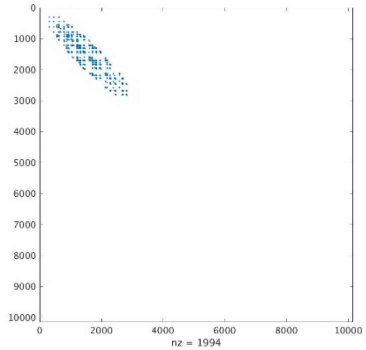
M



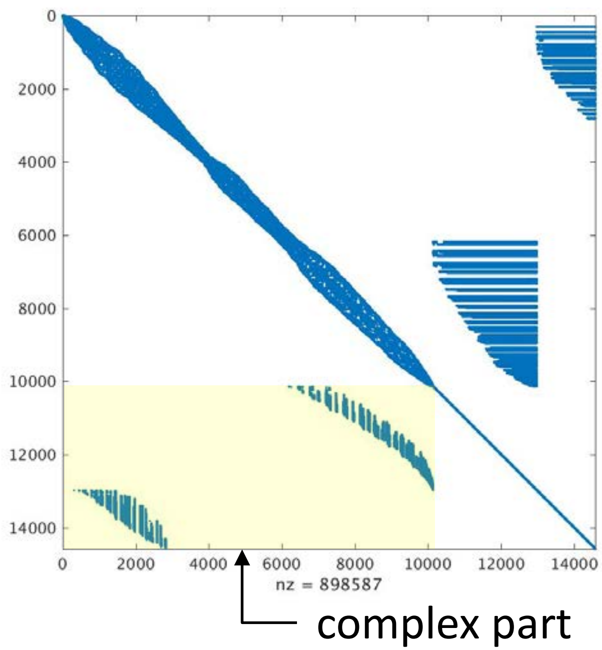
W_1 (rank=355)



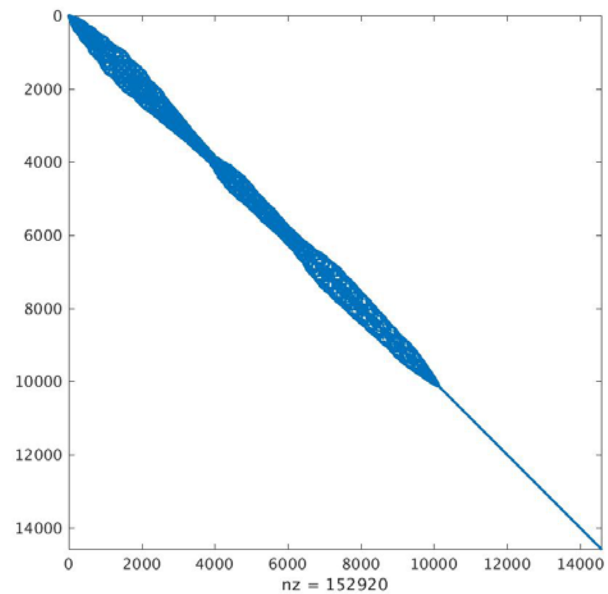
W_2 (rank=200)



A

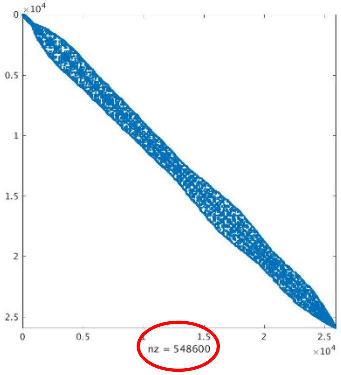


B

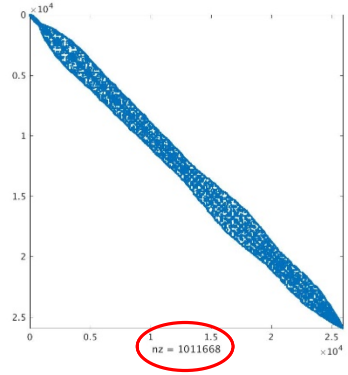


pillbox problem (coarse): $n = 25914,\ d_p = 7$

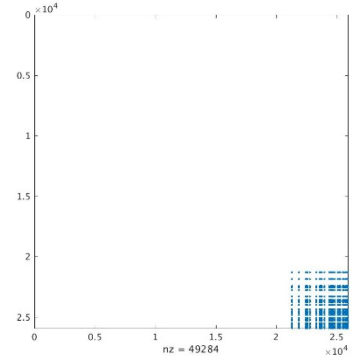
K



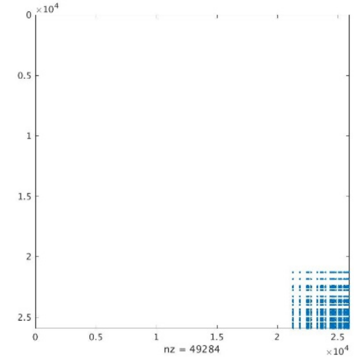
M



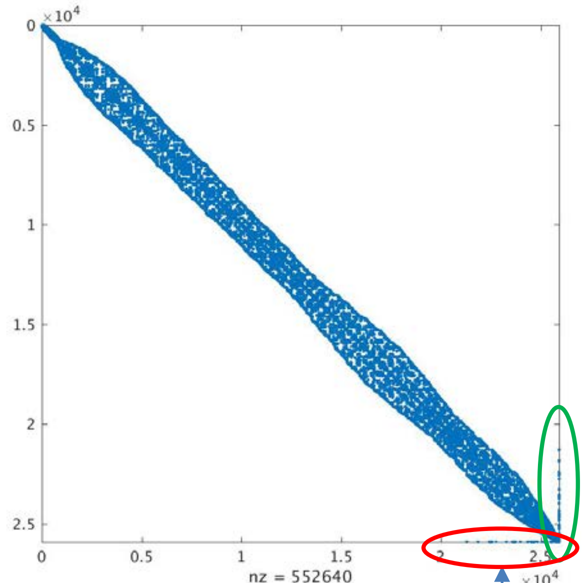
W_1 (rank=1)



W_2 (rank=1)

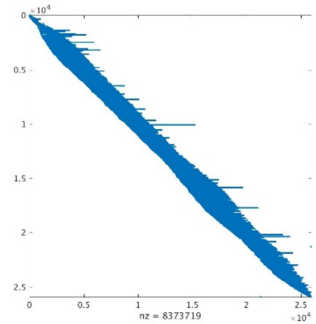


A

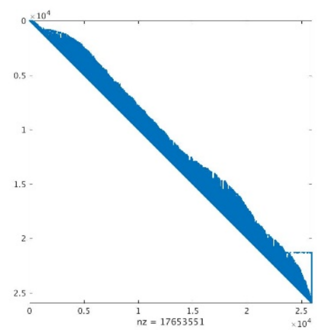


complex part

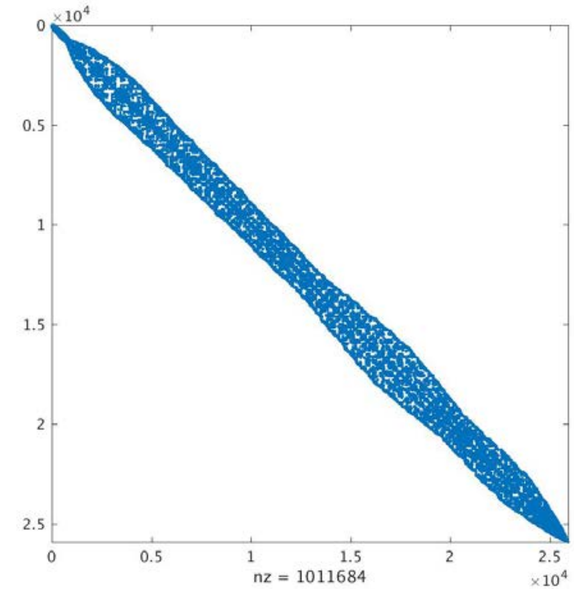
$\hat{L} =$



$\hat{U} =$



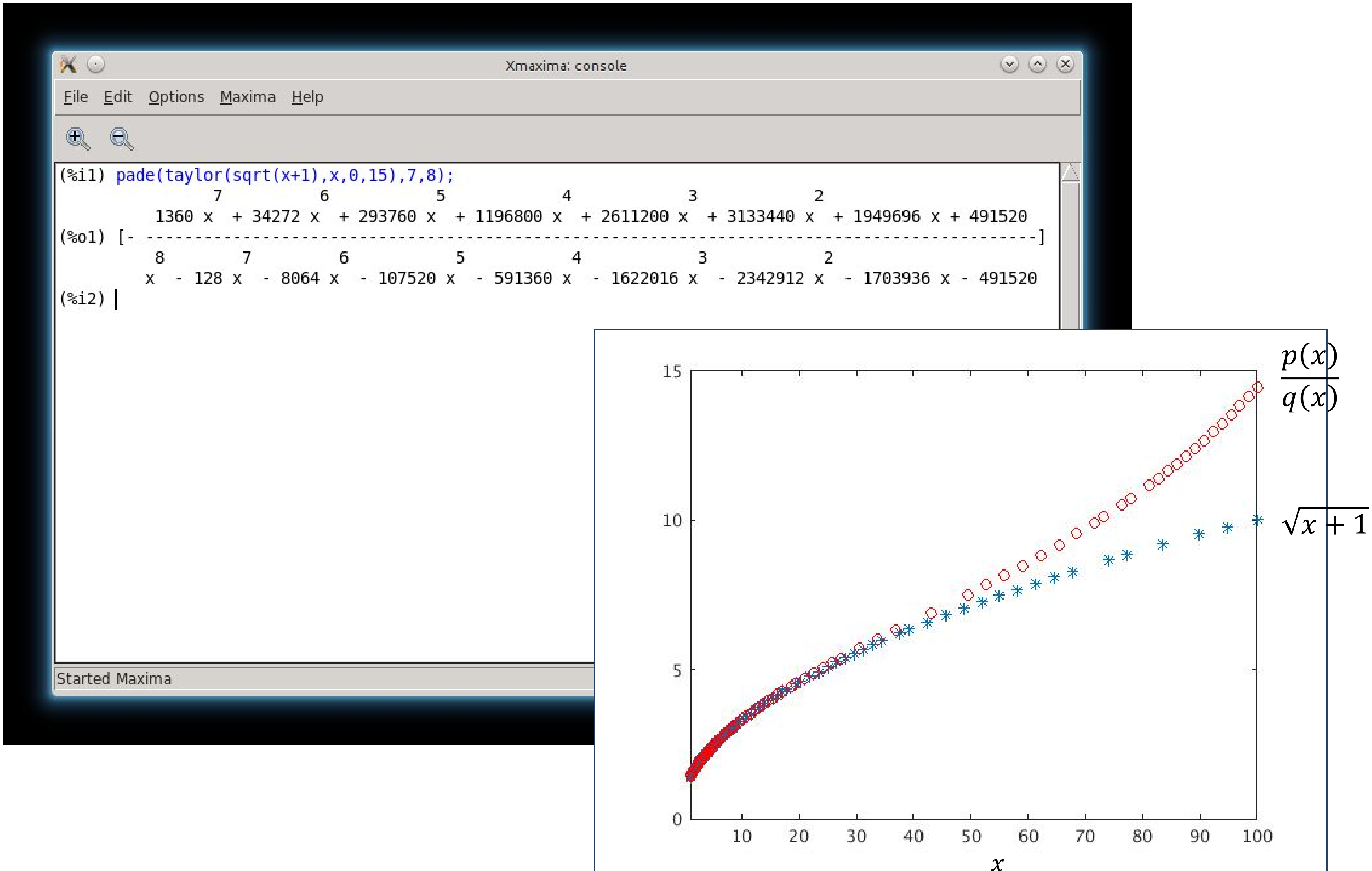
B



Padé Approximants: *change of coordinates*

- $\lambda = \mu + \lambda_0$
- $\sqrt{\lambda - \sigma_k^2} = \sqrt{\lambda_0 - \sigma_k^2}\sqrt{\left(\frac{\mu}{\lambda_0 - \sigma_k^2}\right) + 1} = c\,\sqrt{x+1} \;\approx c\,\frac{p(x)}{q(x)}$
- $(C_k - xD_k)^{-1} = \left(C_k - \frac{\mu}{\lambda_0 - \sigma_k^2} D_k\right)^{-1}$

Padé Approximants for $\sqrt{x+1}$ with Maxima

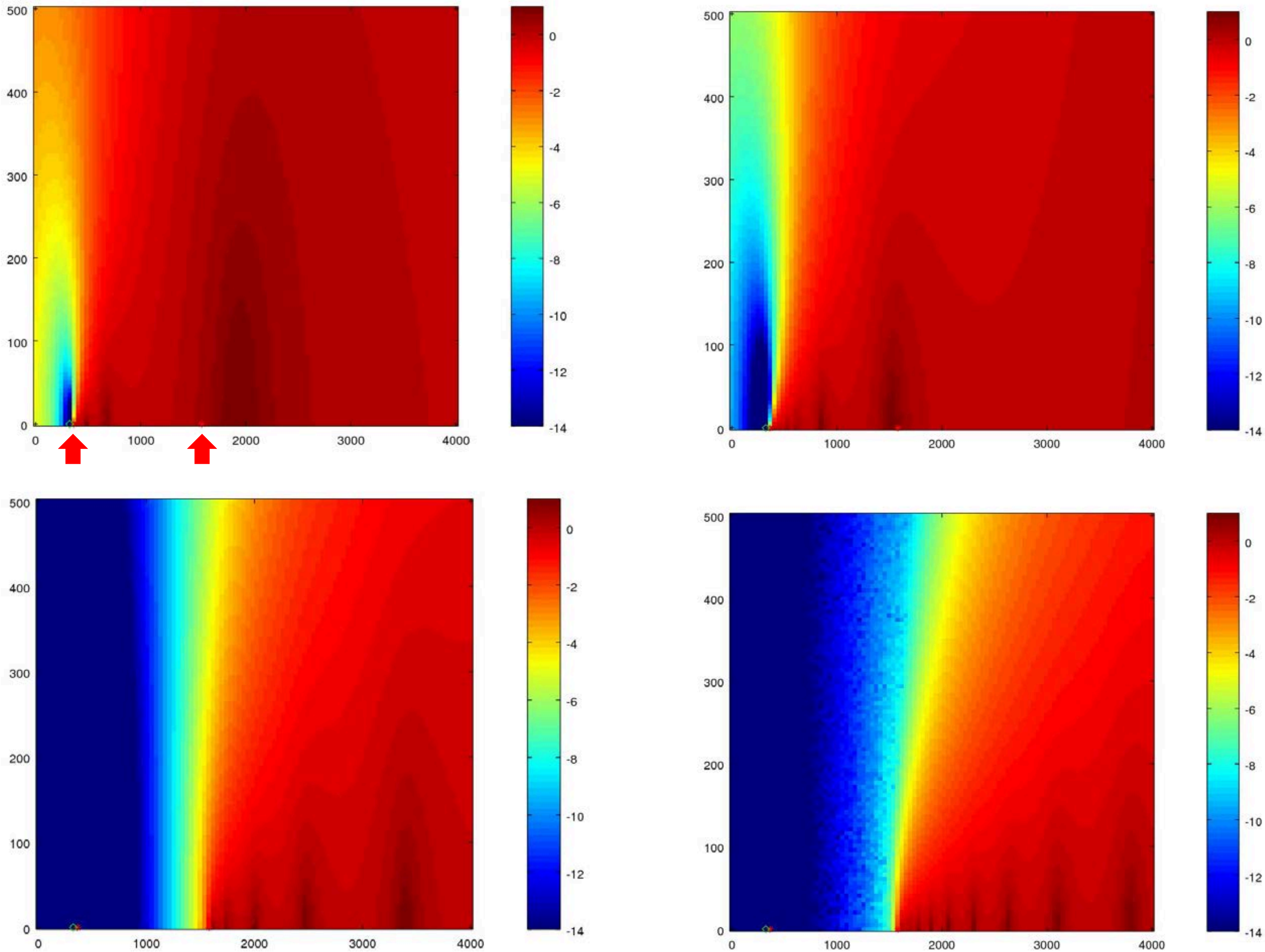


Assessing the quality of the Padé approximants

- pillbox problem: $\sigma_1 = 19.040, \sigma_2 = 39.763$
- $f_k(z) = \sqrt{z - \sigma_k^2}$
- $g(z) = c\,\frac{p(z)}{q(z)}$
- $r_k(z) = \frac{|f_k(z) - g(z)|}{|f_k(z)|}$
- pictures courtesy of Jacob Johnson

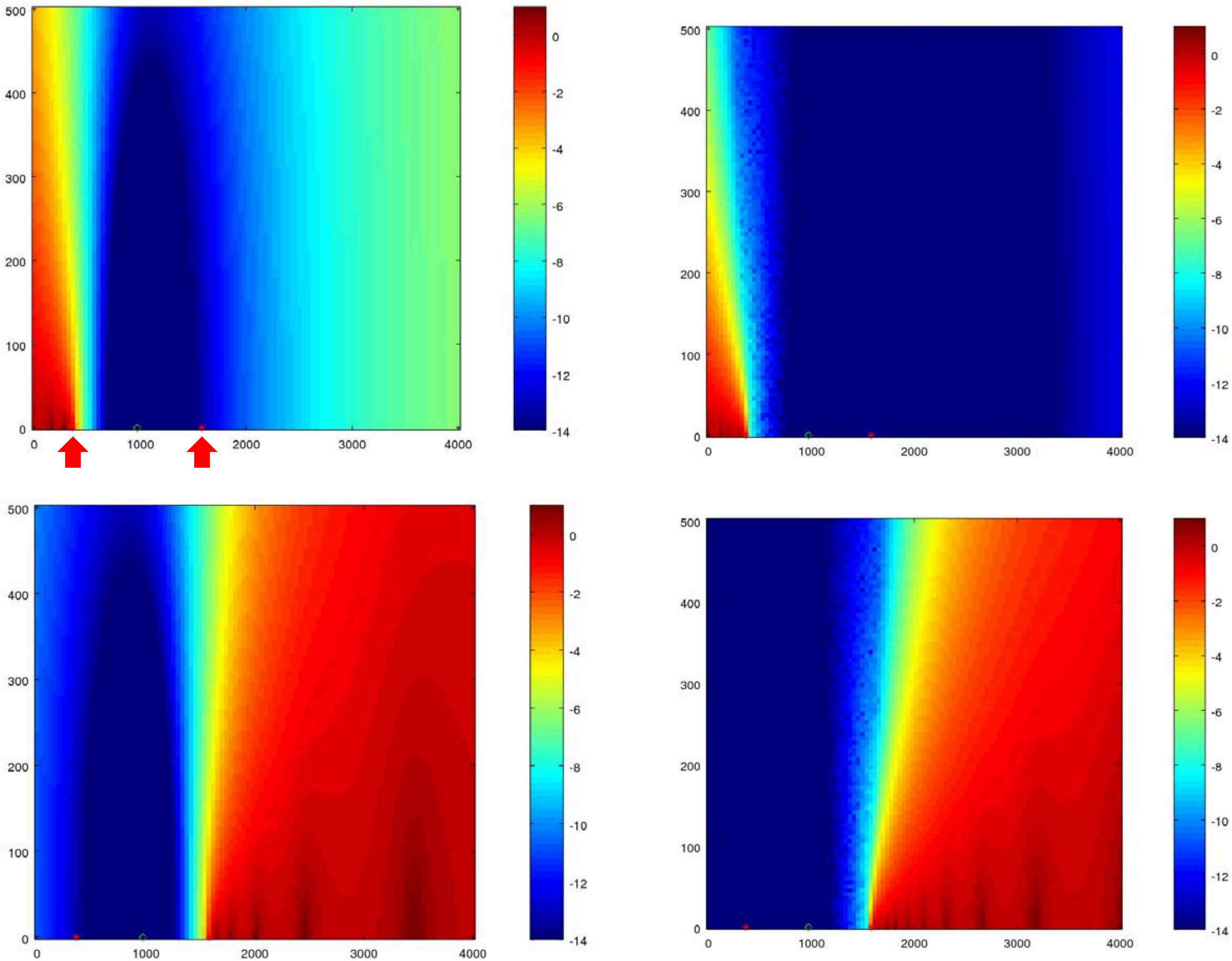
$\lambda_0 = 18.0$

$d_q = 8$ $\qquad\qquad$ $d_q = 16$



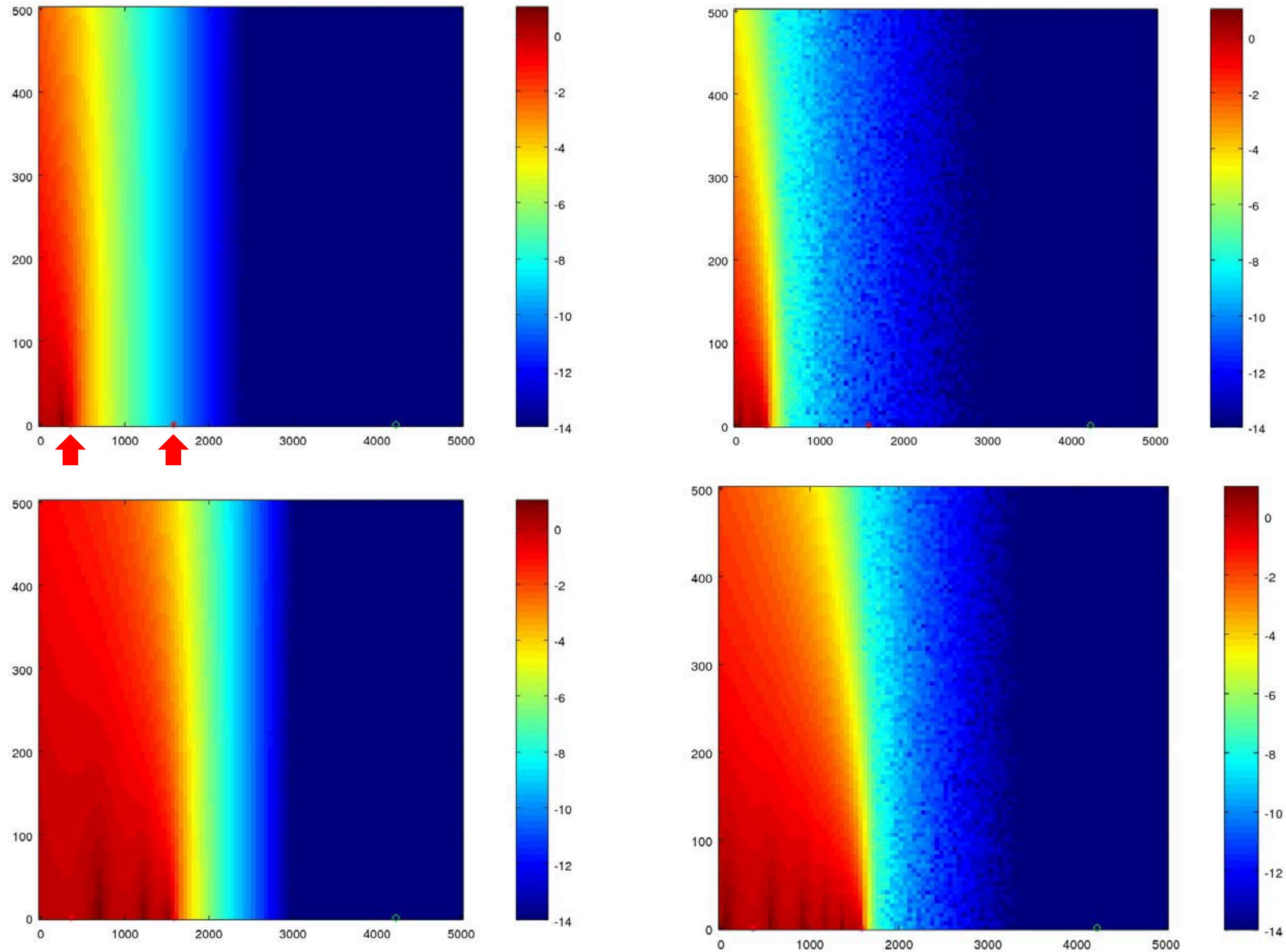
$$\lambda_0 = 31.0$$

$d_a = 8$ $\qquad$ $d_q = 16$



$\lambda_0 = 65.0$

$d_q = 8$ $d_q = 16$



$$\mathrm{Res}(\lambda, x) = \frac{\|\mathcal{N}(\lambda)x\|_2/\|x\|_2}{\|K\|_1 + |\lambda|\|M\|_1 + \sqrt{|\lambda - \sigma_1^2|}\,\|W_1\|_1 + \sqrt{|\lambda - \sigma_2^2|}\,\|W_2\|_1}$$

$\lambda_0 = 31.0, d_q = 8$

#	Qe	sqrt(eig) NEP		res
1	2.31e+00	1.451536524954908e+01	3.143497825350441e+00 i	1.71e-07
2	8.87e+00	1.716622176455362e+01	9.675886796702055e-01 i	4.23e-07
3	4.71e+01	1.859831897401196e+01	1.974846413698813e-01 i	9.12e-07
4	3.47e+02	2.434947852128341e+01	3.509420612647874e-02 i	7.75e-12
5	4.33e+10	3.992821523738079e+01	4.608556138970174e-10 i	2.08e-06
6	-1.35e+10	4.044128017712303e+01	-1.493842415218597e-09 i	2.23e-06
7	1.38e+09	4.133081006415875e+01	1.496907312533786e-08 i	1.86e-06
8	2.17e+08	4.165656716124850e+01	9.579114752698678e-08 i	5.58e-07

$\lambda_0 = 65.0, d_q = 8$

#	Qe	sqrt(eig) NEP		res
1	7.32e+03	6.434788618668856e+01	4.396032402504043e-03 i	2.62e-16
2	9.52e+01	6.452288035389111e+01	3.387203103962095e-01 i	2.55e-16
3	1.91e+01	6.629579759528129e+01	1.732550445193188e+00 i	7.94e-16
4	1.89e+02	6.674594288038902e+01	1.761574035107226e-01 i	5.94e-16
5	1.42e+01	6.732660518695030e+01	2.373494447683829e+00 i	1.10e-15
6	8.12e+03	6.793085224226957e+01	4.182508463304461e-03 i	5.78e-16
7	7.43e+02	6.893837425032825e+01	4.641064437184546e-02 i	8.73e-16
8	9.00e+02	7.001107674971578e+01	3.888757621137510e-02 i	5.85e-16

- Similar to case $k = 2$ (slide 7)
- $\begin{bmatrix} K & L \\ U^T & \hat{C} \end{bmatrix} \hat{x} = \lambda \begin{bmatrix} M & 0 \\ 0 & \widehat{D} \end{bmatrix} \hat{x}$
- $\hat{C} = \begin{bmatrix} \hat{C}_1 & \\ & \hat{C}_2 \end{bmatrix}$ and $\widehat{D} = \begin{bmatrix} \widehat{D}_1 & \\ & \widehat{D}_2 \end{bmatrix}$
- $\hat{C}_i$ and $\widehat{D}_i$ correspond to different Pade approximants
- $\hat{C}_i$ and $\widehat{D}_i$ themselves can correspond to a sum of various terms
- $\frac{\lambda}{\sqrt{\lambda - \sigma_2^2}} \xRightarrow{\lambda = \mu + \lambda_0} \frac{1}{c}\left(c_1 \frac{x}{\sqrt{x+1}} + c_2 \frac{1}{\sqrt{x+1}}\right)$

Conclusions

- Padé approximants is a good alternative for the linearization of a class of problems $F(\lambda)x = 0$
- Approximations for solutions near branch points need to be better understood
- Future work:
  - tests with polynomials of higher degrees
  - $F(\lambda) = K - \lambda M + i\sqrt{\lambda - \sigma_1^2} W_1 + i \frac{\lambda}{\sqrt{\lambda - \sigma_2^2}} W_2$
  - integration of the solver into a parallel simulation code and further experiments