

# Nonlinear Eigenvalue Problems: Theory and Applications

David Bindel<sup>1</sup>

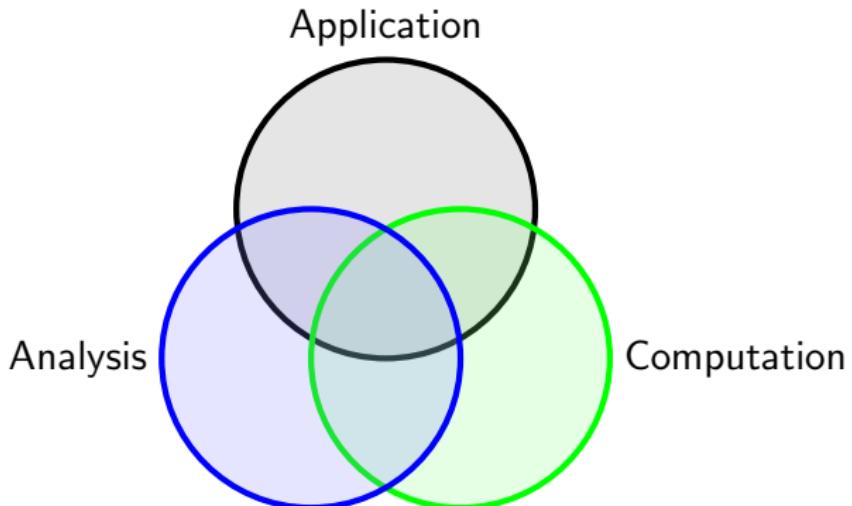
Department of Computer Science  
Cornell University

12 January 2017

---

<sup>1</sup>Joint work with Amanda Hood

# The Computational Science & Engineering Picture



- MEMS
- Smart grids
- Networks
- Systems
- Linear algebra
- Approximation theory
- Symmetry + structure
- Optimization
- HPC / cloud
- Simulators
- Solvers
- Frameworks

# Why eigenvalues?



- The polynomial connection
- The optimization connection
- The approximation connection
- The Fourier/quadrature/special function connection
- The **dynamics** connection

# Why nonlinear eigenvalues?

$$y' - Ay = 0 \xrightarrow{y=e^{\lambda t}v} (\lambda I - A)v = 0$$

$$y'' + By' + Ky = 0 \xrightarrow{y=e^{\lambda t}v} (\lambda^2 I + \lambda B + K)v = 0$$

$$y' - Ay - By(t-1) = 0 \xrightarrow{y=e^{\lambda t}v} (\lambda I - A - Be^{-\lambda})v = 0$$

$$T(d/dt)y = 0 \xrightarrow{y=e^{\lambda t}v} T(\lambda)v = 0$$

- Higher-order ODEs
- Dynamic element formulations
- Delay differential equations
- Boundary integral equation eigenproblems
- Radiation boundary conditions

## Big and little

$$y'' + By' + Ky = 0 \xrightarrow{v=y'} \begin{bmatrix} v \\ y \end{bmatrix}' + \begin{bmatrix} B & K \\ -I & 0 \end{bmatrix} \begin{bmatrix} v \\ y \end{bmatrix} = 0$$

$$(\lambda^2 I + \lambda B + K)u = 0 \xrightarrow{v=\lambda u} \left( \lambda I + \begin{bmatrix} B & K \\ -I & 0 \end{bmatrix} \right) \begin{bmatrix} v \\ u \end{bmatrix} = 0$$

**Tradeoff:** more variables = more linear.

# The big picture

$$T(\lambda)v = 0, \quad v \neq 0.$$

where

- $T : \Omega \rightarrow \mathbb{C}^{n \times n}$  analytic,  $\Omega \subset \mathbb{C}$  simply connected
- Regularity:  $\det(T) \not\equiv 0$

Nonlinear spectrum:  $\Lambda(T) = \{z \in \Omega : T(z) \text{ singular}\}$ .

What do we want?

- Qualitative information (e.g. no eigenvalues in right half plane)
- Error bounds on computed/estimated eigenvalues
- Assurances that we know all the eigenvalues in some region

# Standard solver strategies

$$T(\lambda)v = 0, \quad v \neq 0.$$

A common approach:

- ① **Approximate**  $T$  locally (linear, rational, etc)
- ② Solve **approximate** problem.
- ③ Repeat as needed.

How should we choose solver parameters? What about global behavior?  
What can we trust? What might we miss?

# Analyticity to the rescue

$$T(\lambda)v = 0, \quad v \neq 0.$$

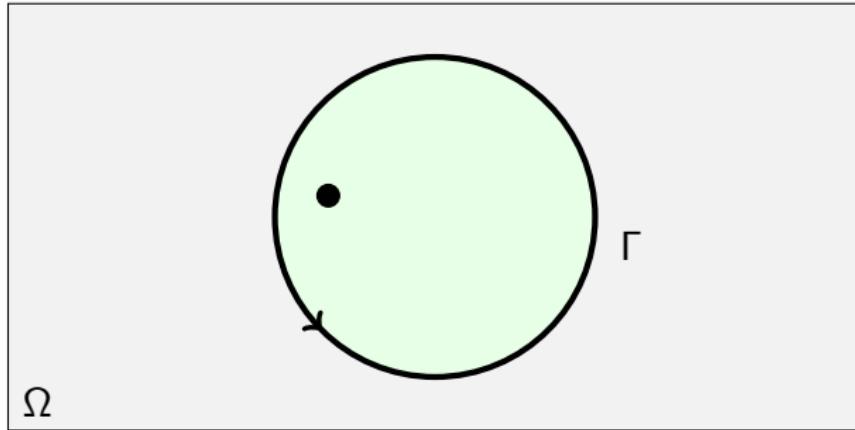
where

- $T : \Omega \rightarrow \mathbb{C}^{n \times n}$  **analytic**,  $\Omega \subset \mathbb{C}$  simply connected
- Regularity:  $\det(T) \not\equiv 0$

Nonlinear spectrum:  $\Lambda(T) = \{z \in \Omega : T(z) \text{ singular}\}$ .

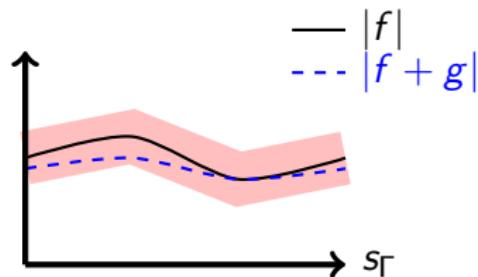
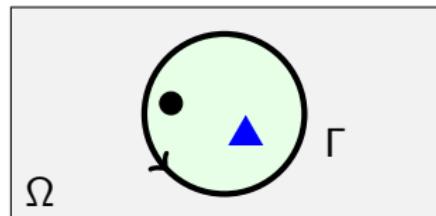
**Goal:** Use analyticity to *compare* and to *count*

# Winding and Cauchy's argument principle



$$\begin{aligned} W_{\Gamma}(f) &= \frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z)}{f(z)} dz \\ &= \# \text{ zeros} - \# \text{ poles} \end{aligned}$$

# Winding, Rouché, and Gohberg-Sigal



Analytic  $f, g : \Omega \rightarrow \mathbb{C}$

Winding #  $\frac{1}{2\pi i} \int_{\Gamma} \frac{f'(z)}{f(z)} dz$

Theorem Rouché (1862):  
 $|g| < |f|$  on  $\Gamma \implies$   
same # zeros of  $f, f + g$

$T, E : \Omega \rightarrow \mathbb{C}^{n \times n}$

$\text{tr} \left( \frac{1}{2\pi i} \int_{\Gamma} T(z)^{-1} T'(z) dz \right)$

Gohberg-Sigal (1971):  
 $\|T^{-1}E\| < 1$  on  $\Gamma \implies$   
same # eigs of  $T, T + E$

# Comparing NEPs

Suppose

$T, E : \Omega \rightarrow \mathbb{C}^{n \times n}$  analytic

$\Gamma \subset \Omega$  a simple closed contour

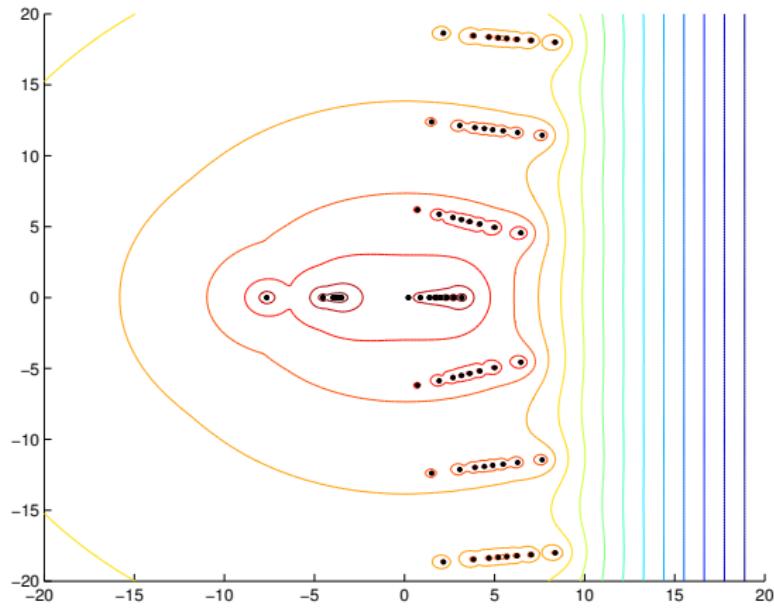
$T(z) + sE(z)$  nonsingular  $\forall s \in [0, 1], z \in \Gamma$

Then  $T$  and  $T + E$  have the same number of eigenvalues inside  $\Gamma$ .

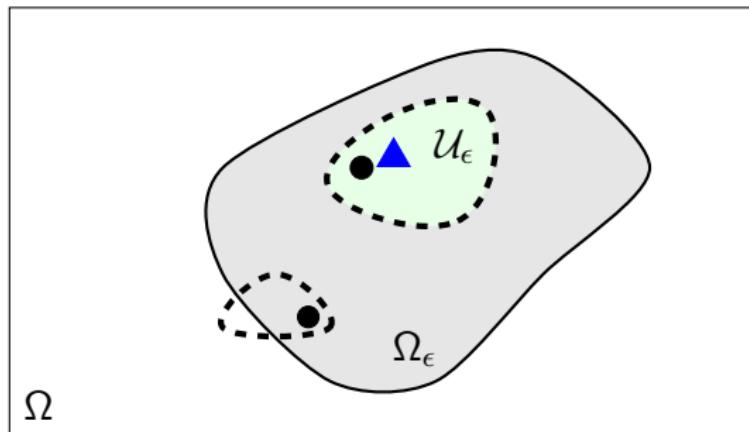
**Pf:** Constant winding number around  $\Gamma$ .

# Nonlinear pseudospectra

$$\Lambda_\epsilon(T) \equiv \{z \in \Omega : \|T(z)^{-1}\| > \epsilon^{-1}\}$$



## Pseudospectral comparison



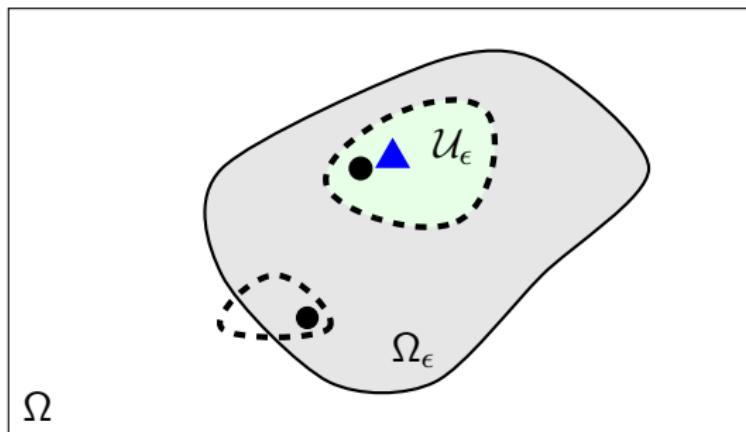
$E$  analytic,  $\|E(z)\| < \epsilon$  on  $\Omega_\epsilon$ . Then

$$\Lambda(T + E) \cap \Omega_\epsilon \subset \Lambda_\epsilon(T) \cap \Omega_\epsilon$$

Also, if  $U_\epsilon$  a component of  $\Lambda_\epsilon$  and  $\bar{U}_\epsilon \subset \Omega_\epsilon$ , then

$$|\Lambda(T + E) \cap U_\epsilon| = |\Lambda(T) \cap U_\epsilon|$$

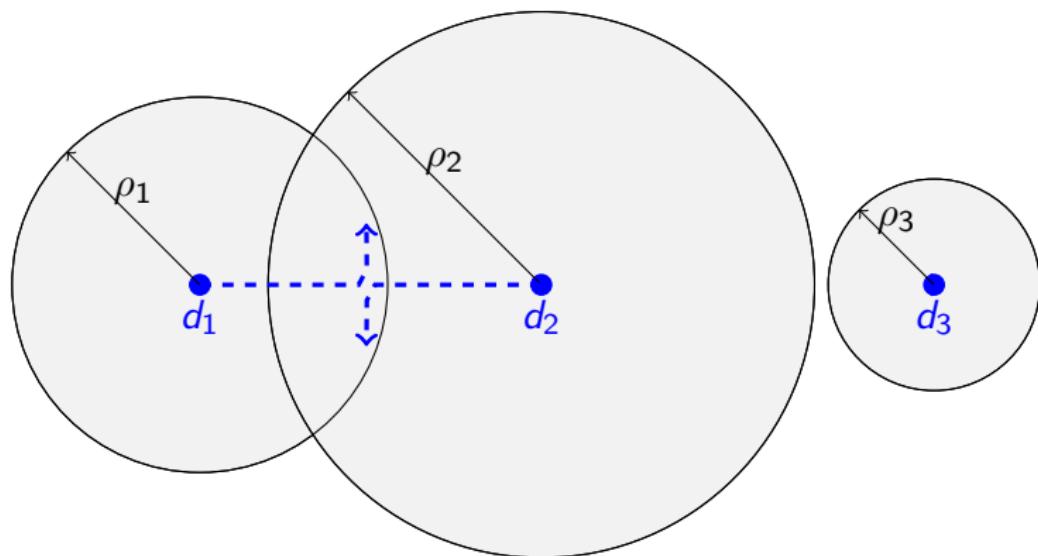
## Pseudospectral comparison



- Most useful when  $T$  is linear
- Even then, can be expensive to compute!
- What about related tools?

## The Gershgorin picture (linear case)

$$A = D + F, \quad D = \text{diag}(d_i), \quad \rho_i = \sum_j |f_{ij}|$$



## Gershgorin ( $+\epsilon$ )

Write  $A = D + F$ ,  $D = \text{diag}(d_1, \dots, d_n)$ . Gershgorin disks are:

$$G_i = \left\{ z \in \mathbb{C} : |z - d_i| \leq \sum_j |f_{ij}| \right\}.$$

Useful facts:

- Spectrum of  $A$  lies in  $\bigcup_{i=1}^m G_i$
- $\bigcup_{i \in \mathcal{I}} G_i$  disjoint from other disks  $\implies$  contains  $|\mathcal{I}|$  eigenvalues.

Pf:

$A - zI$  strictly diagonally dominant outside  $\bigcup_{i=1}^m G_i$ .

Eigenvalues of  $D - sF$ ,  $0 \leq s \leq 1$ , are continuous.

## Nonlinear Gershgorin

Write  $T(z) = D(z) + F(z)$ . Gershgorin *regions* are

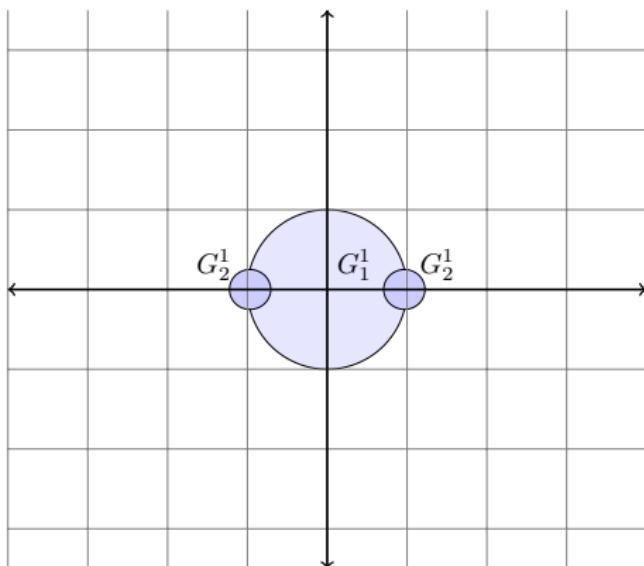
$$G_i = \left\{ z \in \mathbb{C} : |d_i(z)| \leq \sum_j |f_{ij}(z)| \right\}.$$

Useful facts:

- Spectrum of  $T$  lies in  $\bigcup_{i=1}^m G_i$
- Bdd connected component of  $\bigcup_{i=1}^m G_i$  strictly in  $\Omega$ 
  - ⇒ same number of eigs of  $D$  and  $T$  in component
  - ⇒ at least one eig per component of  $G_i$  involved

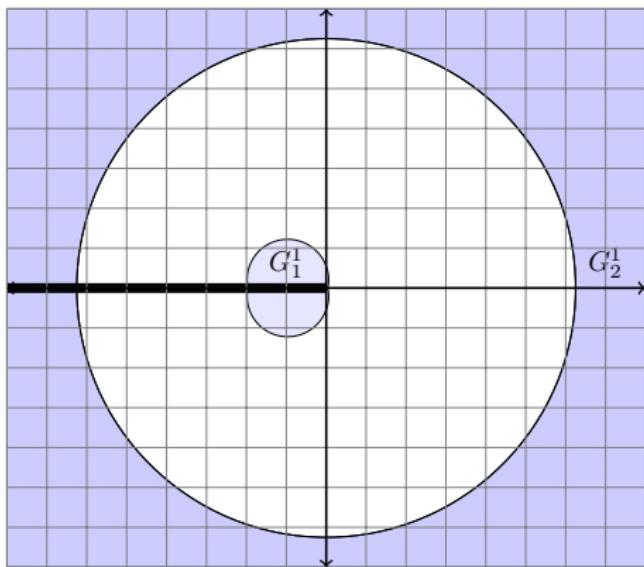
**Pf:** Strict diag dominance test + continuity of eigs

## Mini-example: Counting contributions



$$T(z) = \begin{bmatrix} z & 1 & 0 \\ 0 & z^2 - 1 & 0.5 \\ 0 & 0 & 1 \end{bmatrix}.$$

## Mini-example: Domain boundaries



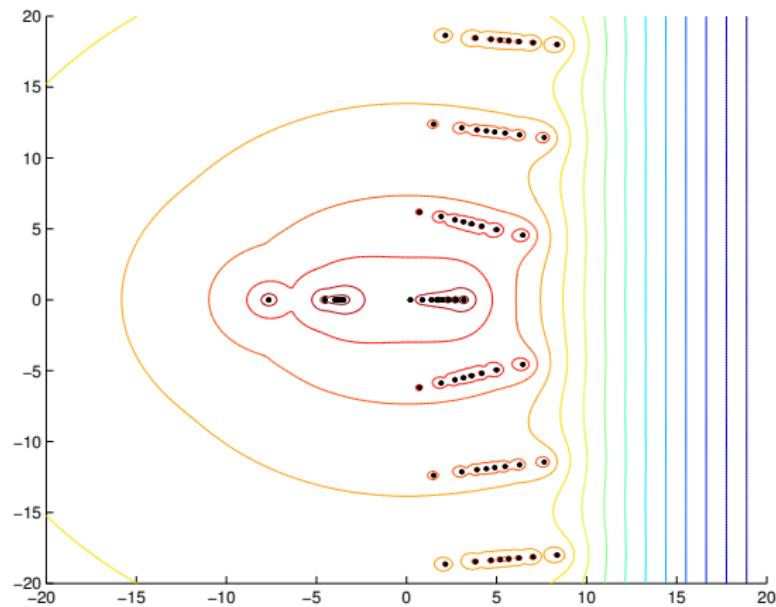
$$T(z) = \begin{bmatrix} z - 0.2\sqrt{z} + 1 & -1 \\ 0.4\sqrt{z} & 1 \end{bmatrix}$$
$$\Omega = \mathbb{C} - (-\infty, 0]$$

$$\det(D(z)) = (\sqrt{z} - 0.1 - i\sqrt{0.99})$$
$$(\sqrt{z} - 0.1 + i\sqrt{0.99})$$
$$\det(T(z)) = (\sqrt{z} + 0.1 - i\sqrt{0.99})$$
$$(\sqrt{z} + 0.1 + i\sqrt{0.99})$$

$D$  has two eigenvalues in  $\Omega$ ;  
 $T$  hides both eigenvalues behind a branch cut.

## Example I: Hadeler

$$T(z) = (e^z - 1)B + z^2 A - \alpha I, \quad A, B \in \mathbb{R}^{8 \times 8}$$



## Comparison to simplified problem

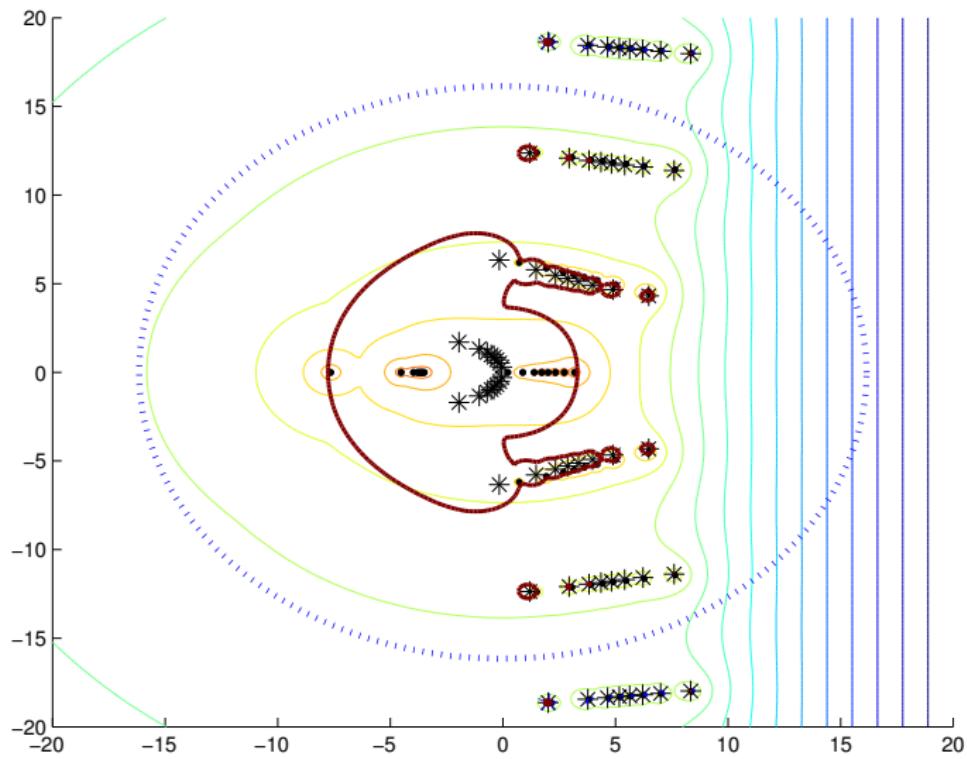
Bauer-Fike idea: apply a similarity!

$$T(z) = (e^z - 1)B + z^2A - \alpha I$$

$$\begin{aligned}\tilde{T}(z) &= U^T T(z) U \\ &= (e^z - 1)D_B + z^2I - \alpha E \\ &= D(z) - \alpha E\end{aligned}$$

$$G_i = \{z : |\beta_i(e^z - 1) + z^2| < \rho_i\}.$$

# Gershgorin regions



## A different comparison

Approximate  $e^z - 1$  by a Chebyshev interpolant:

$$T(z) = (e^z - 1)B + z^2A - \alpha I$$

$$\tilde{T}(z) = q(z)B + z^2A - \alpha$$

$$T(z) = \tilde{T}(z) + r(z)B$$

Linearize  $\tilde{T}$  and transform both:

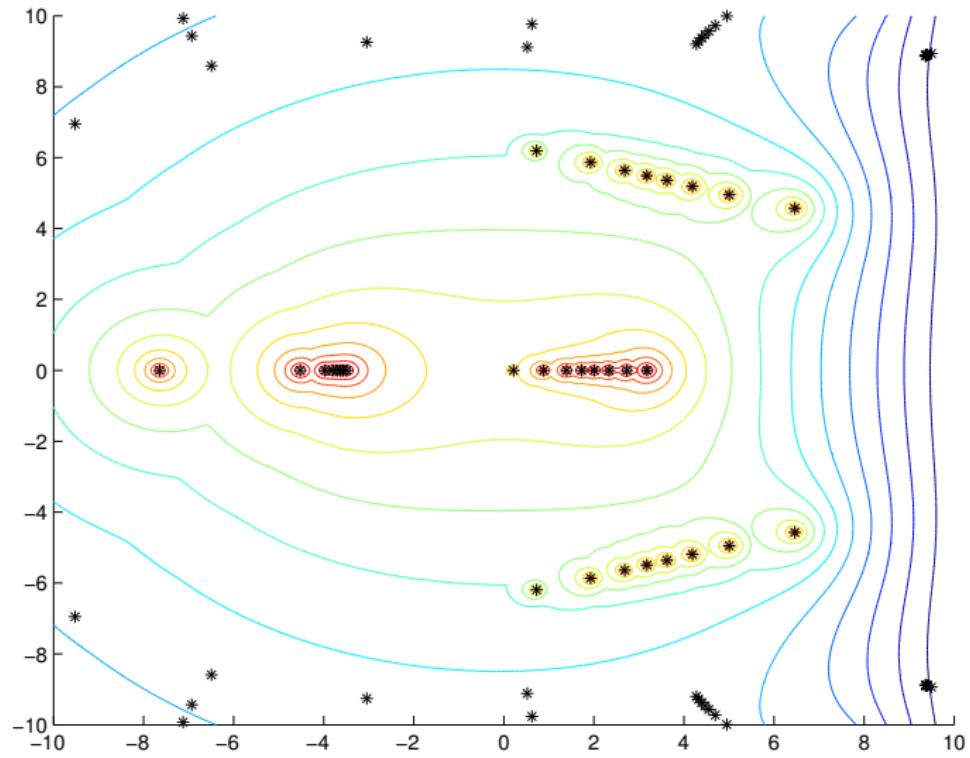
$$\tilde{T}(z) \mapsto D_C - zI$$

$$T(z) \mapsto D_C - zI + r(z)E$$

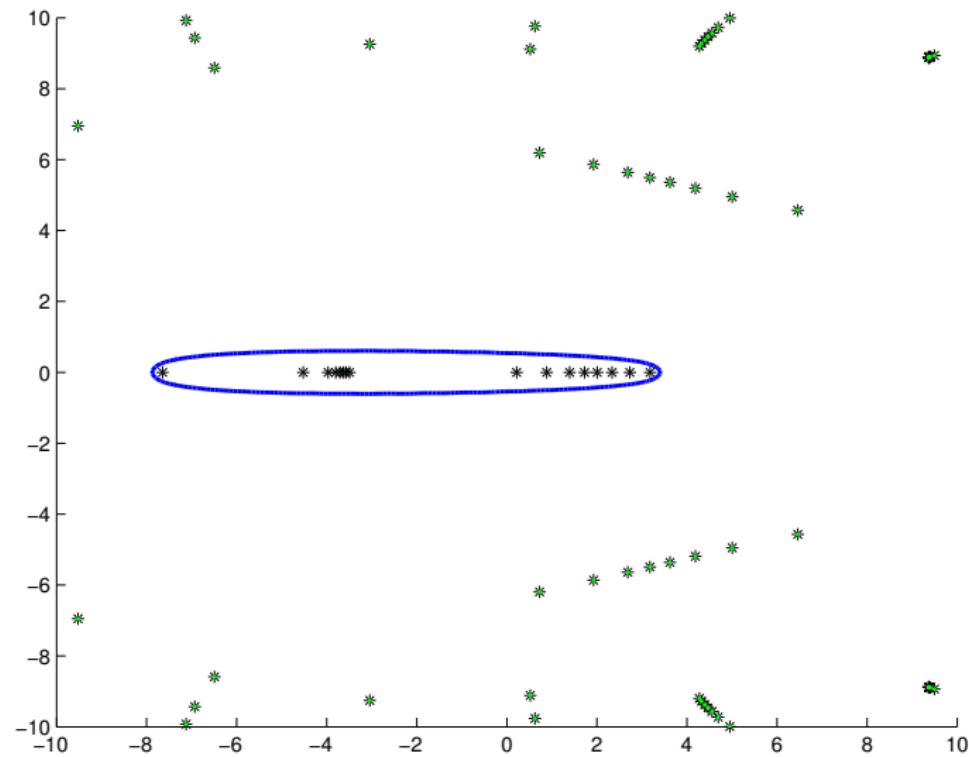
Restrict to  $\Omega_\epsilon = \{z : |r(z)| < \epsilon\}$ :

$$G_i \subset \hat{G}_i = \{z : |z - \mu_i| < \rho_i \epsilon\}, \quad \rho_i = \sum_j |e_{ij}|$$

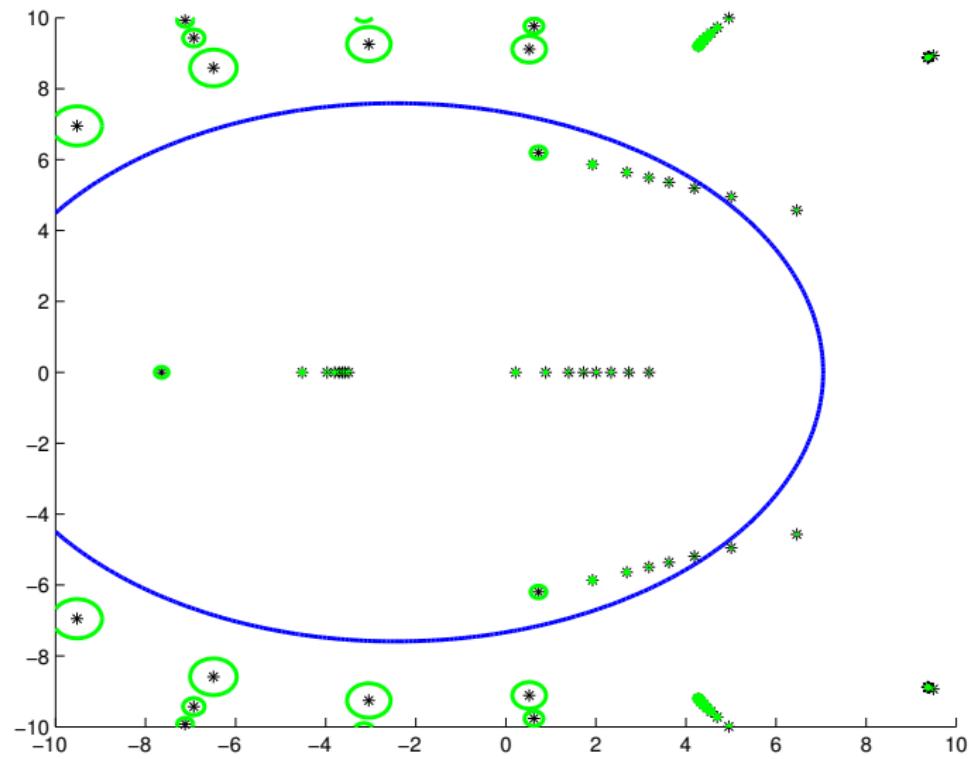
# Spectrum of $\tilde{T}$



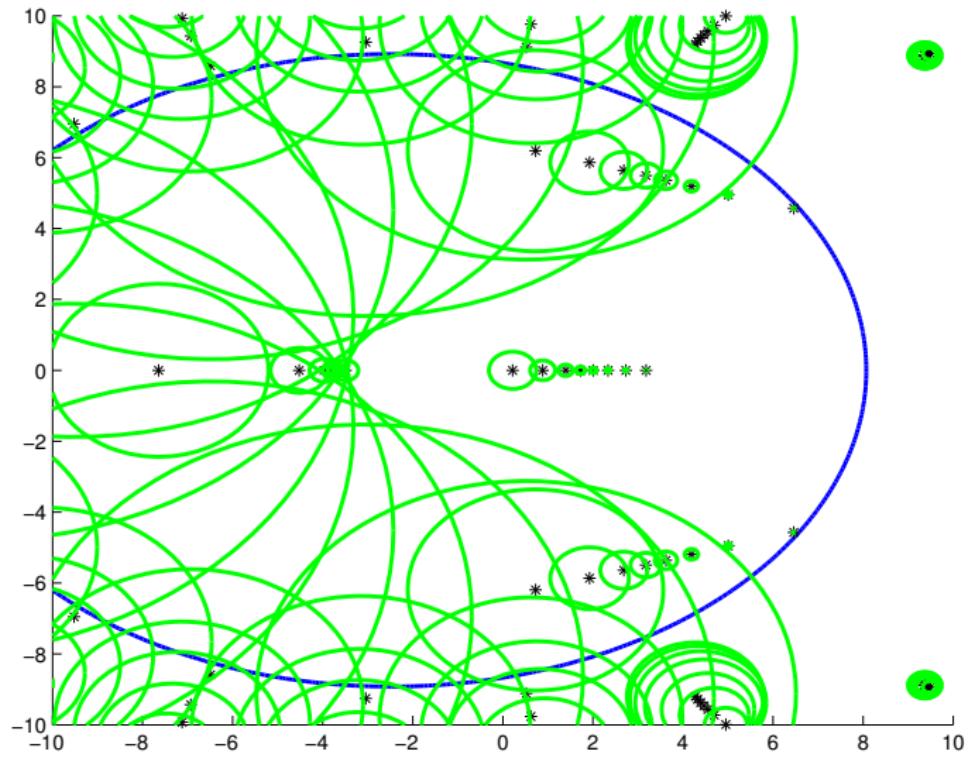
# $\hat{G}_i$ for $\epsilon < 10^{-10}$



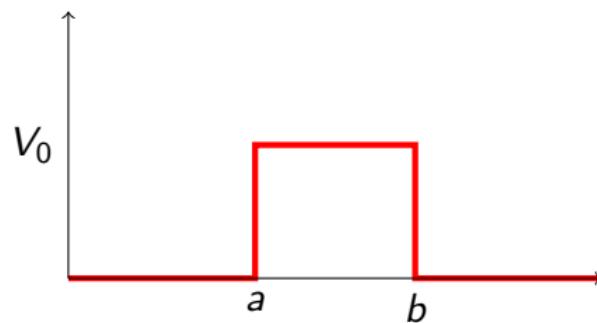
$\hat{G}_i$  for  $\epsilon = 0.1$



# $\hat{G}_i$ for $\epsilon = 1.6$



## Example II: Resonance problem

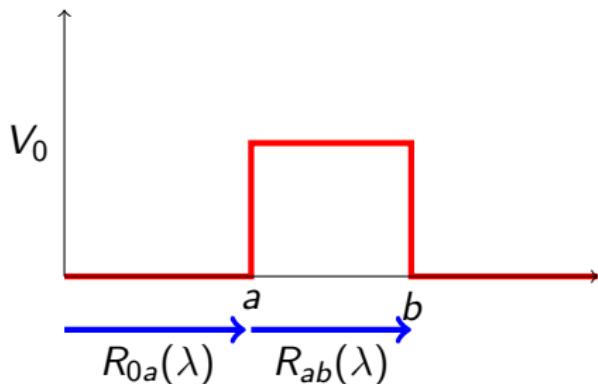


$$\psi(0) = 0$$

$$\left( -\frac{d^2}{dx^2} + V - \lambda \right) \psi = 0 \quad \text{on } (0, b),$$

$$\psi'(b) = i\sqrt{\lambda}\psi(b),$$

## Reduction via shooting



$$\psi(0) = 0,$$

$$R_{0a}(\lambda) \begin{bmatrix} \psi(0) \\ \psi'(0) \end{bmatrix} = \begin{bmatrix} \psi(a) \\ \psi'(a) \end{bmatrix},$$

$$R_{ab}(\lambda) \begin{bmatrix} \psi(b) \\ \psi'(b) \end{bmatrix} = \begin{bmatrix} \psi(b) \\ \psi'(b) \end{bmatrix},$$

$$\psi'(b) = i\sqrt{\lambda}\psi(b)$$

## Reduction via shooting

First-order form:

$$\frac{du}{dx} = \begin{bmatrix} 0 & 1 \\ V - \lambda & 0 \end{bmatrix} u, \text{ where } u(x) \equiv \begin{bmatrix} \psi(x) \\ \psi'(x) \end{bmatrix}.$$

On region  $(c, d)$  where  $V$  is constant:

$$u(d) = R_{cd}(\lambda)u(c), \quad R_{cd}(\lambda) = \exp\left((d - c) \begin{bmatrix} 0 & 1 \\ V - \lambda & 0 \end{bmatrix}\right)$$

Reduce resonance problem to 6D NEP:

$$T(\lambda)u_{\text{all}} \equiv \begin{bmatrix} R_{0a}(\lambda) & -I & 0 & -I \\ 0 & R_{ab}(\lambda) & 0 & 0 \\ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} & 0 & \begin{bmatrix} 0 & 0 \\ -i\sqrt{\lambda} & 1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} u(0) \\ u(a) \\ u(b) \end{bmatrix} = 0.$$

## Expansion via rational approximation

Consider the equation

$$\left( \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \lambda I \right) \begin{bmatrix} u \\ v \end{bmatrix} = 0$$

Partial Gaussian elimination gives the **spectral Schur complement**

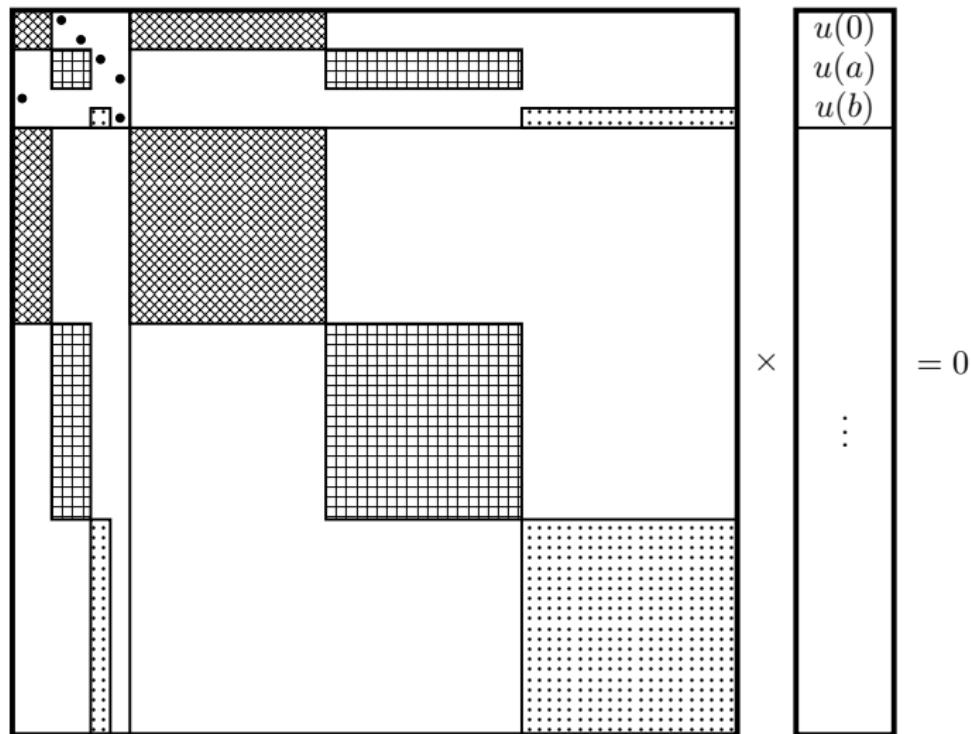
$$(A - \lambda I - B(D - \lambda I)^{-1}C) u = 0$$

**Idea:** Given

$$T(\lambda) = A - \lambda I - F(\lambda),$$

find a rational approximation  $F(\lambda) \approx B(D - \lambda I)^{-1}C$ .

# Expansion via rational approximation



## Analyzing the expanded system

- $\hat{T}(z)$  is a Schur complement in  $K - zM$ 
  - So  $\Lambda(\hat{T})$  is easy to compute.
- Or: think  $T(z)$  is a Schur complement in  $K - zM + E(z)$
- Compare  $\hat{T}(z)$  to  $T(z)$  or compare  $K - zM + E(z)$  to  $K - zM$

# Analyzing the expanded system

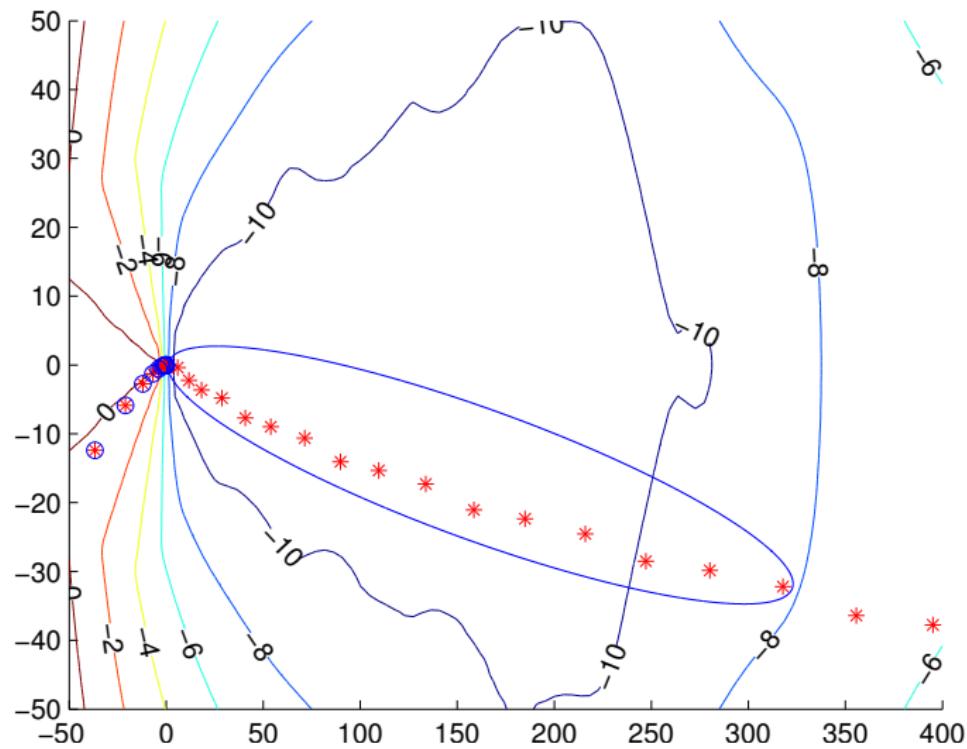
**Q:** Can we find all eigs in a region *not missing anything*?

Concrete plan ( $\epsilon = 10^{-8}$ )

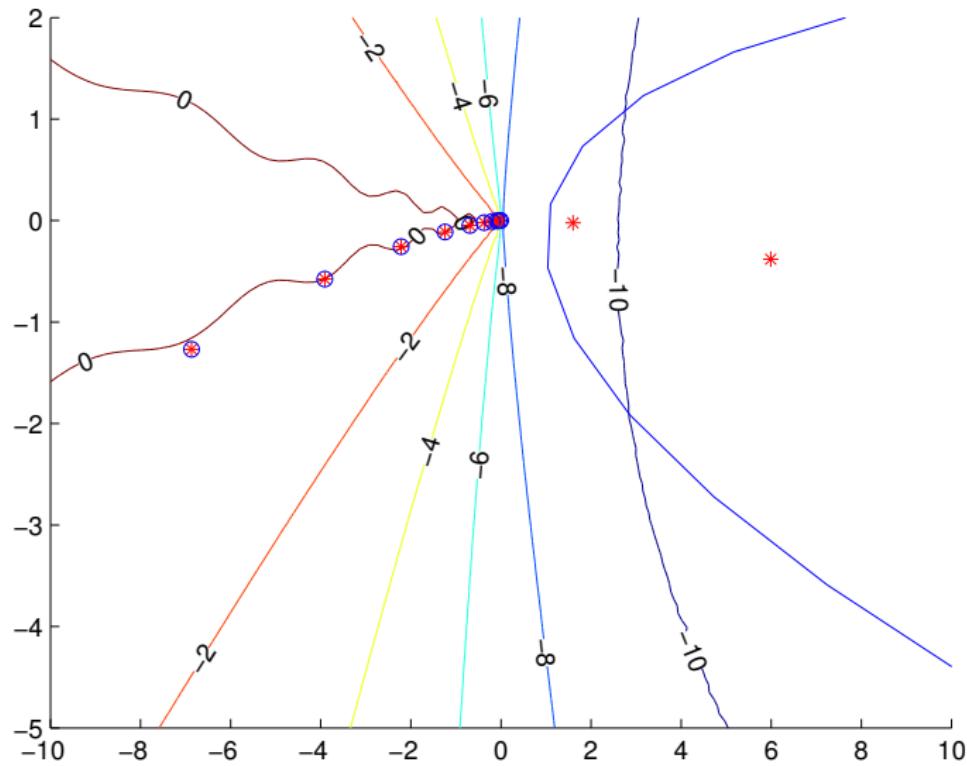
- $T$  = shooting system
- $\hat{T}$  = rational approximation
- Find region  $D$  with boundary  $\Gamma$  s.t.
  - $D \subset \Omega_\epsilon$  (i.e.  $\|T - \hat{T}\| < \epsilon$  on  $D$ )
  - $\Gamma$  does not intersect  $\Lambda_\epsilon(T)$
- $\implies$  Same eigenvalue counts for  $T, \hat{T}$
- $\implies$  Eigs of  $\hat{T}$  in components of  $\Lambda_\epsilon(T)$ 
  - Converse holds if  $D \subset \Omega_{\epsilon/2}$

Can refine eigs of  $\hat{T}$  in  $D$  via Newton.

## Resonance approximation



# Resonance approximation



## Example III: Bounding dynamics

DDE model (of a laser with feedback)

$$\dot{u} = Au(t) + Bu(t-1).$$

Solution with  $u(0+) = u_0$  and  $u(t) = \phi(t)$  on  $[-1, 0]$ :

$$u(t) = \Psi(t)u_0 + \int_0^1 \Psi(t-s)B\phi(s-t) ds$$

where  $\Psi(t)$  is a fundamental matrix for “shock” solutions.

## Representing $\Psi(t)$

Associated nonlinear eigenvalue problem involves

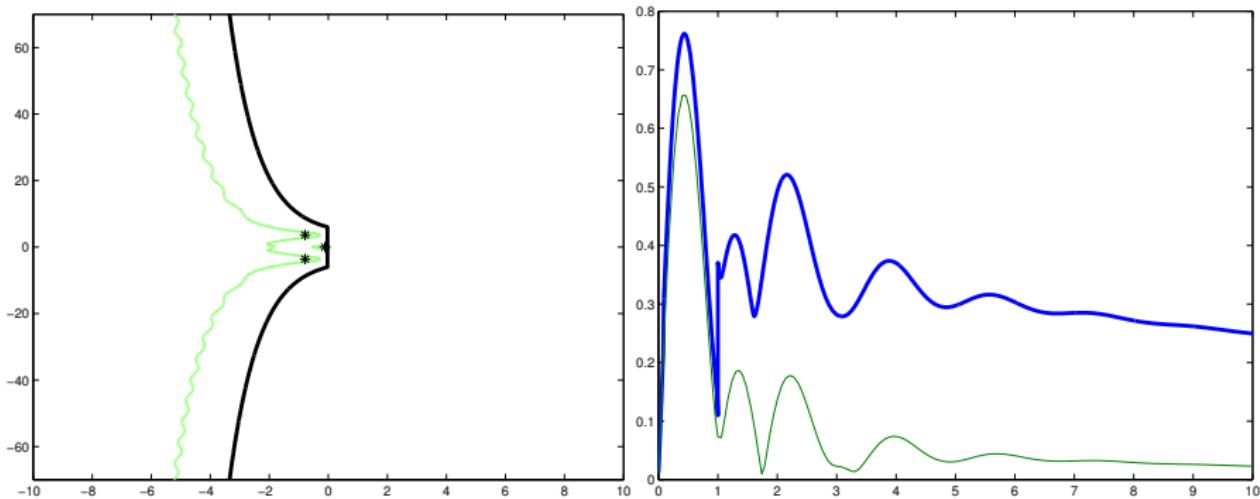
$$T(z) = zI - A - Be^{-z}$$

Assume all eigs in left half plane; via inverse Laplace transform,

$$\Psi(t) = \frac{1}{2\pi i} \int_{\Gamma} T(z)^{-1} e^{zt} dz.$$

for contour  $\Gamma$  right of the spectrum.

# Bounding pseudospectra $\implies$ bounding dynamics



Bound fundamental solution  $\Psi$  via

$$\|\Psi(t)\| = \left\| \frac{1}{2\pi i} \int_{\Gamma} T(z)^{-1} e^{zt} dz \right\| \leq \frac{1}{2\pi} \int_{\Gamma} \|T(z)^{-1}\| |e^{zt}| |dz|.$$

For more

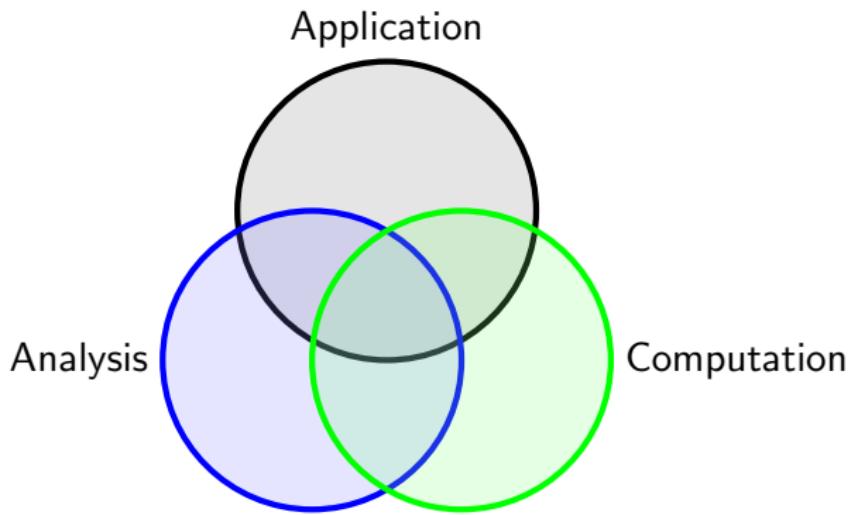
*Localization theorems for nonlinear eigenvalues.*

David Bindel and Amanda Hood, SIAM Review 57(4), Dec 2015

*Pseudospectral bounds on transient growth for higher order and constant delay differential equations.*

Amanda Hood and David Bindel, submitted to SIMAX  
arXiv:1611.05130, Nov 2016

# Trailers!

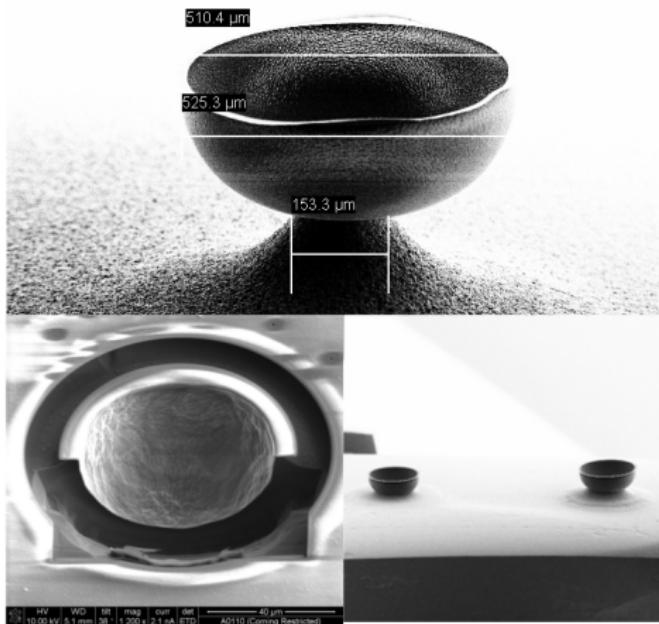
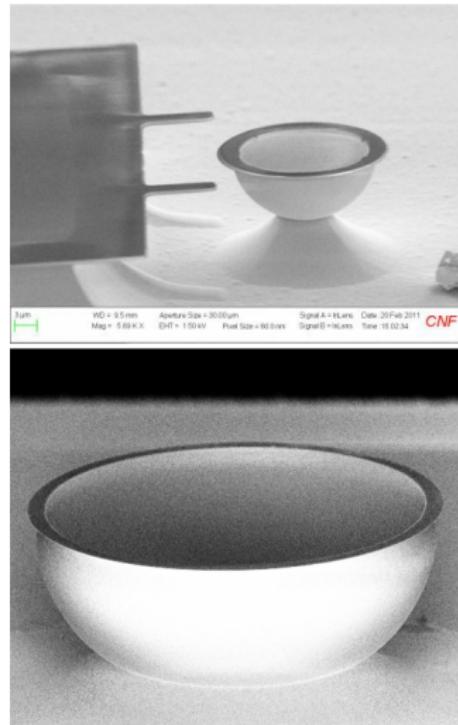


# Spectral Topic Modeling

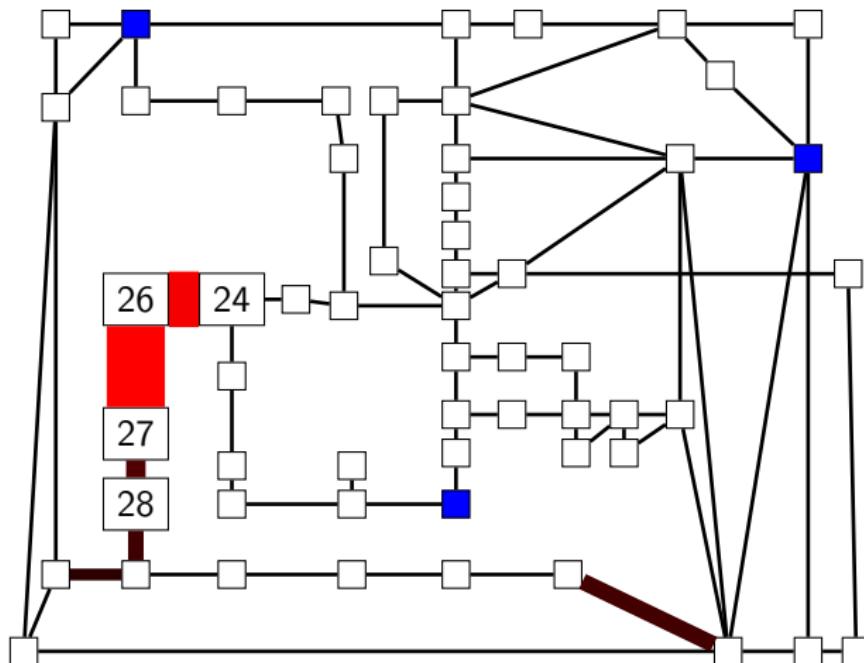
$$C \approx B \times A \times B^T$$

The diagram illustrates the matrix factorization process in spectral topic modeling. On the left, a large pink rectangle labeled  $C$  represents the original document-term matrix. To its right is the approximation symbol  $\approx$ . Following this are three smaller rectangles: a red rectangle labeled  $B$ , a blue rectangle labeled  $A$ , and another red rectangle labeled  $B^T$ . The  $B$  and  $B^T$  rectangles are positioned side-by-side, indicating they are multiplied together. The  $\times$  symbols between  $B$  and  $A$ , and between  $A$  and  $B^T$ , represent matrix multiplication.

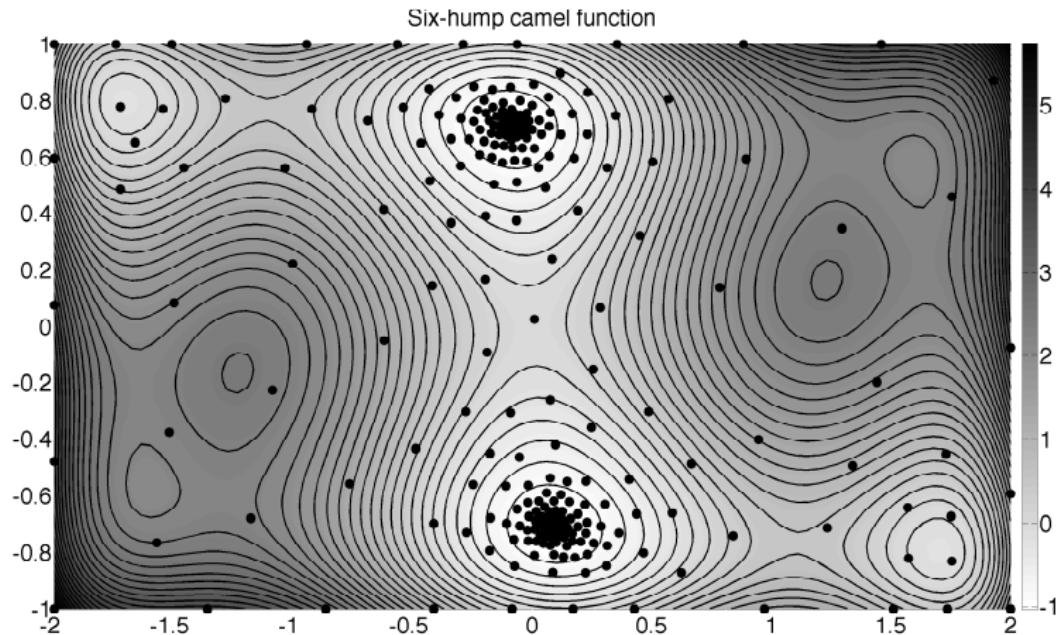
# Music of the Microspheres



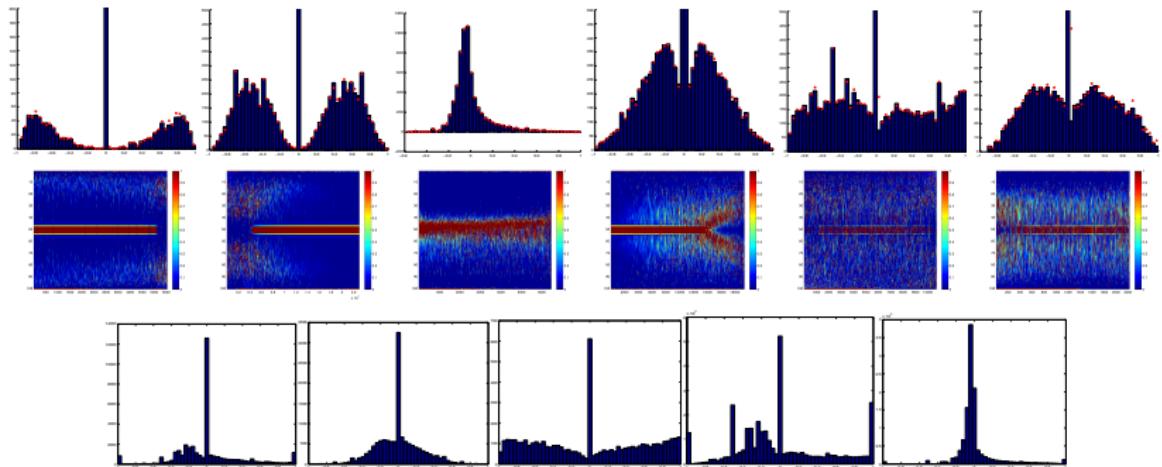
# Fast Fingerprints for Power Systems



# Response Surfaces for Global Optimization



# Graph Densities of States



Fin

<http://www.cs.cornell.edu/~bindel>