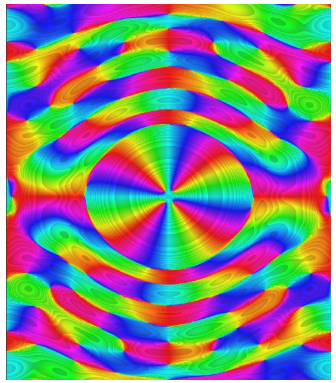
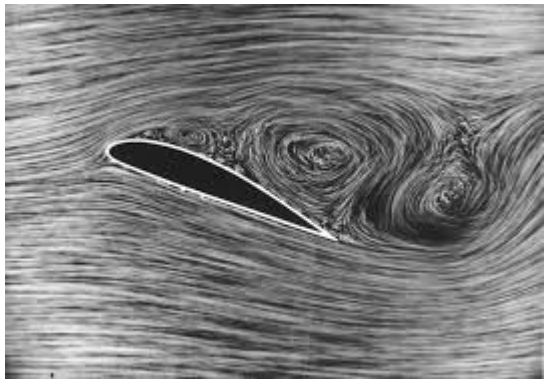
Residual DMD

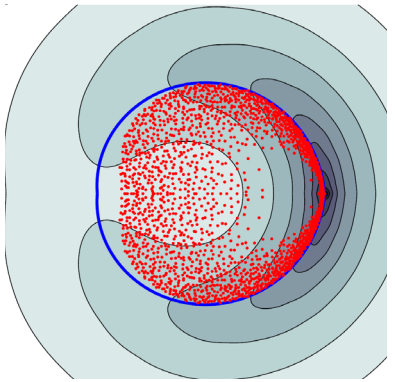
Rigorous data-driven spectral computations of Koopman operators for dynamical systems

Matthew Colbrook (m.colbrook@damtp.cam.ac.uk)

University of Cambridge + École Normale Supérieure







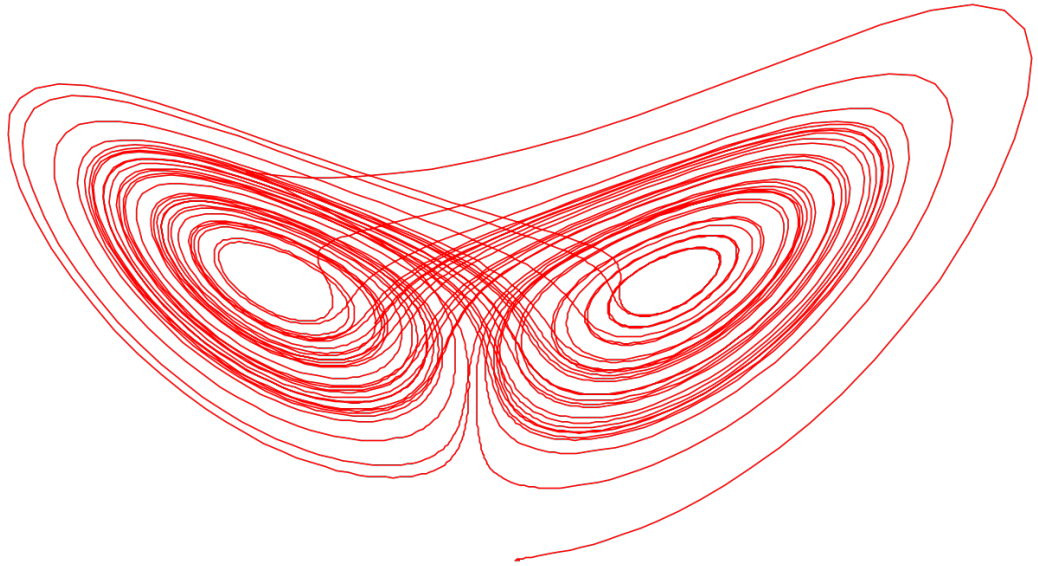
C., Townsend, "*Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems*"

Data-driven dynamical systems

- State $x \in \Omega \subseteq \mathbb{R}^d$, ***unknown*** function $F: \Omega \to \Omega$ governs dynamics

$$x_{n+1} = F(x_n)$$

- **Goal:** Learn about system from data $\left\{x^{(m)}, y^{(m)} = F(x^{(m)})\right\}_{m=1}^{M}$
  - E.g., **data from** trajectories, experimental measurements, simulations, …
  - E.g., **used for** forecasting, control, design, understanding, …
- **Applications:** chemistry, climatology, electronics, epidemiology, finance, fluids, molecular dynamics, neuroscience, plasmas, robotics, video processing, …



Data-driven dynamical systems

- **Carleman linearisation:** Carleman, *"Application de la théorie des équations intégrales linéaires aux systèmes d'équations différentielles non linéaires,"* Acta Mathematica, 1932.
- **Filtering:** Kalman, "*A new approach to linear filtering and prediction problems*," Journal of Basic Engineering, 1960.
- **Ulam's method:** Ulam, "*A Collection of Mathematical Problems*," IP, 1960.
- **Reduced order models:** Guo, Hesthaven. "*Data-driven reduced order modeling for time-dependent problems*," Computer methods in applied mechanics and engineering, 2019.
- **Sparse identification of F:** Brunton, Proctor, Kutz, "*Discovering governing equations from data by sparse identification of nonlinear dynamical systems*," Proceedings of the National Academy of Sciences, 2016.
- **Kernel analog forecasting:** Burov, Giannakis, Manohar, Stuart, "*Kernel analog forecasting: Multiscale test problems*," Multiscale Modeling & Simulation, 2021.
- **Deep learning:** Lu, Jin, Pang, Zhang, Karniadakis, "*Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators*," Nature Machine Intelligence, 2021.
- **Machine learning:** Schmidt, Lipson, "*Distilling free-form natural laws from experimental data,*" Science, 2009.

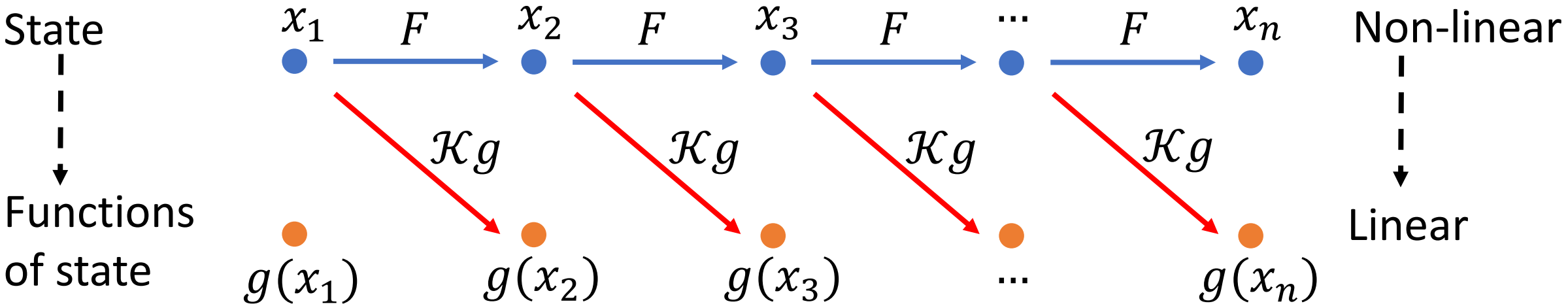
Common Q: Can we develop verified methods?

Operator viewpoint

- **Koopman operator** $\mathcal{K}$ acts on <u>functions</u> $g: \Omega \to \mathbb{C}$

$$[\mathcal{K}g](x) = g(F(x))$$

- $\mathcal{K}$ is ***linear*** but acts on an ***infinite-dimensional*** space.



- Work in $L^2(\Omega, \omega)$ for positive measure ω, with inner product $\langle \cdot, \cdot \rangle$.

- Koopman, "*Hamiltonian systems and transformation in Hilbert space,*" **Proceedings of the National Academy of Sciences**, 1931.
- Koopman, v. Neumann, "*Dynamical systems of continuous spectra,*" **Proceedings of the National Academy of Sciences**, 1932.

$x_{n+1} = F(x_n)$

Why is linear (much) easier?

Dynamics become trivial!

- Suppose $F(x) = Ax, A \in \mathbb{R}^{d\times d}, A = V\Lambda V^{-1}$.
- Set $\xi = V^{-1}x$,

$$\xi_n = V^{-1}x_n = V^{-1}A^n x_0 = \Lambda^n V^{-1}x_0 = \Lambda^n \xi_0$$

- Let $w^{\mathrm{T}}A = \lambda w$, set $\varphi(x) = w^{\mathrm{T}}x$,

$$[\mathcal{K}\varphi](x) = w^{\mathrm{T}}Ax = \lambda\varphi(x)$$

Eigenfunction

Much more general (**non-linear** and even **chaotic** F) …

Koopman mode decomposition

$[\mathcal{K}g](x) = g(F(x))$

eigenfunction of $\mathcal{K}$

generalised eigenfunction of $\mathcal{K}$

$$g(x) = \sum_{\text{eigenvalues } \lambda_j} c_{\lambda_j} \varphi_{\lambda_j}(x) + \int_{-\pi}^{\pi} \phi_{\theta,g}(x)\, \mathrm{d}\theta$$

$$g(x_n) = [\mathcal{K}^n g](x_0) = \sum_{\text{eigenvalues } \lambda_j} c_{\lambda_j} \lambda_j{}^n \varphi_{\lambda_j}(x_0) + \int_{-\pi}^{\pi} e^{in\theta} \phi_{\theta,g}(x_0)\, \mathrm{d}\theta$$

Encodes: geometric features, invariant measures, transient behaviour, long-time behaviour, coherent structures, quasiperiodicity, etc.

GOAL: Data-driven approximation of $\mathcal{K}$ and its spectral properties.

- Mezić, "*Spectral properties of dynamical systems, model reduction and decompositions,*" **Nonlinear Dynamics**, 2005.

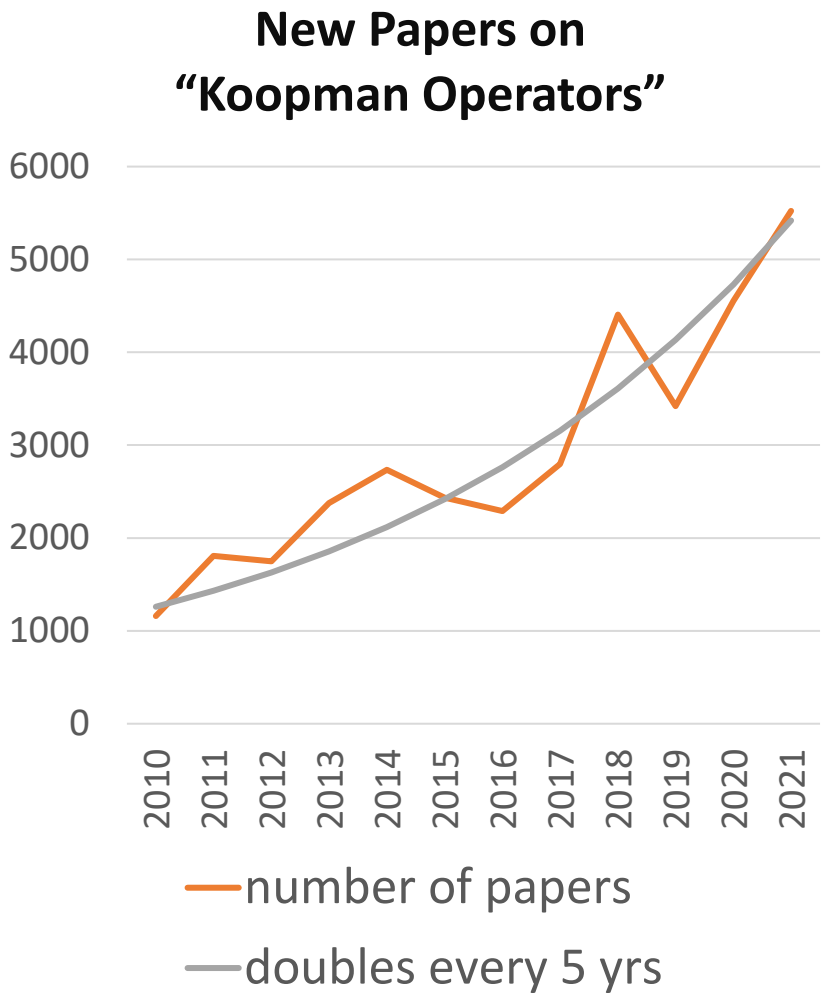
Koopmania*: a revolution in the big data era

≈35,000 papers over last decade!

***Very* little on convergence guarantees or verification.**

Why is this lacking?

- Koopman operators have so far been quite distinct from numerical analysis community.
- Dealing with infinite dim is notoriously hard …



**Wikipedia*: *"its wild surge in popularity is sometimes jokingly called 'Koopmania'"*

Can we compute spectral properties in inf. dim.?

$$\mathcal{K}\left(\sum_{l=1}^{\infty} g_l \psi_l\right) = \sum_{j=1}^{\infty}\left(\sum_{l=1}^{\infty} k_{jl} g_l\right)\psi_j\,, \qquad \mathcal{K} \text{ "=" } \begin{pmatrix} k_{11} & k_{12} & \cdots \\ k_{21} & k_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

basis expansion of $g\colon \Omega \to \mathbb{C}$

Finite-dimensional	$\Longrightarrow$	Infinite-dimensional
Eigenvalues of $B \in \mathbb{C}^{n\times n}$	$\Longrightarrow$	Spectrum, $\mathrm{Spec}(\mathcal{K})$
$\{\lambda_j \in \mathbb{C}\colon \det(B - \lambda_j I) = 0\}$	$\Longrightarrow$	$\{\lambda \in \mathbb{C}\colon \mathcal{K} - \lambda I \text{ is not invertible}\}$

*"Most operators that arise in practice are not presented in a representation in which they are **diagonalized**, and it is often very hard to locate even a single point in the spectrum. Thus, one often has to settle for numerical approximations. Unfortunately, there is a dearth of literature on this basic problem and, so far as we have been able to tell, **there are no proven general techniques**."* W. Arveson, Berkeley (1994)

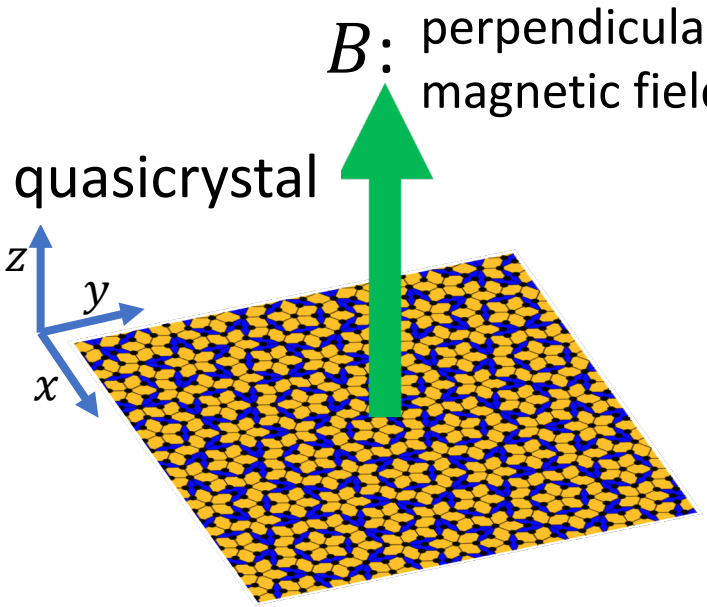
Four key challenges

Naïve: $\mathcal{K}$ "=" $\begin{pmatrix} k_{11} & k_{12} & \cdots \\ k_{21} & k_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$ → $\mathbb{K} \in \mathbb{C}^{N_K \times N_K}$ + compute e-values

1) **"Too little":** Miss parts of $\mathrm{Spec}(\mathcal{K})$
2) **"Too much":** Approximate spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$ - "spectral pollution"
3) **Lose continuous spectra.**
4) **Verification:** Which part of an approximation can we trust?

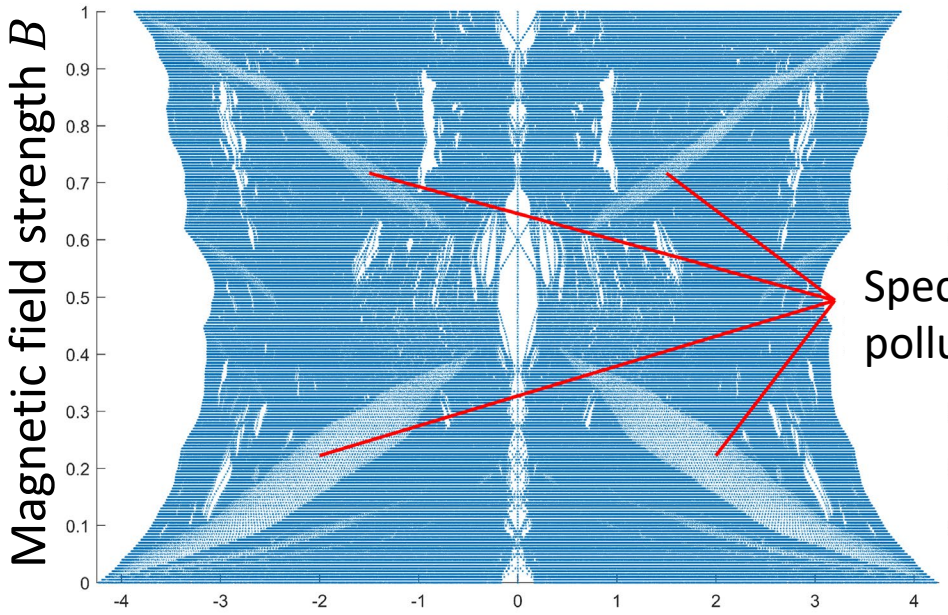
- Arveson, "*The role of C^*-algebras in infinite dimensional numerical linear algebra,*" **Contemp. Math.**, 1994.
- Davies, "*Linear operators and their spectra,*" **CUP**, 2007.
- Brunton, Kutz, "*Data-driven Science and Engineering: Machine learning, Dynamical systems, and Control*," **CUP**, 2019.

Example of "too much" (spectral pollution)



Infinite matrix: discrete Schrödinger operator

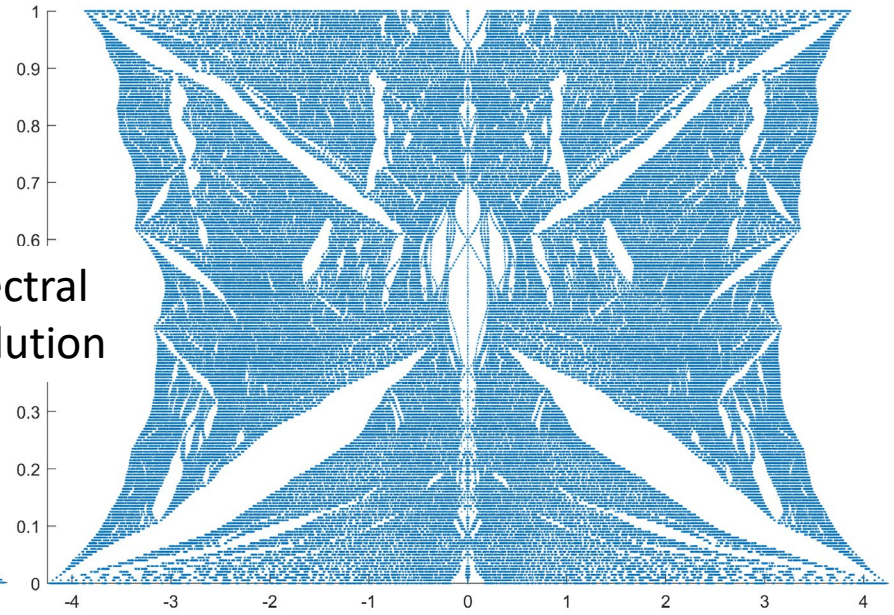
Naïve method



Spectral pollution

No error control

"Infinite-dimensional" method

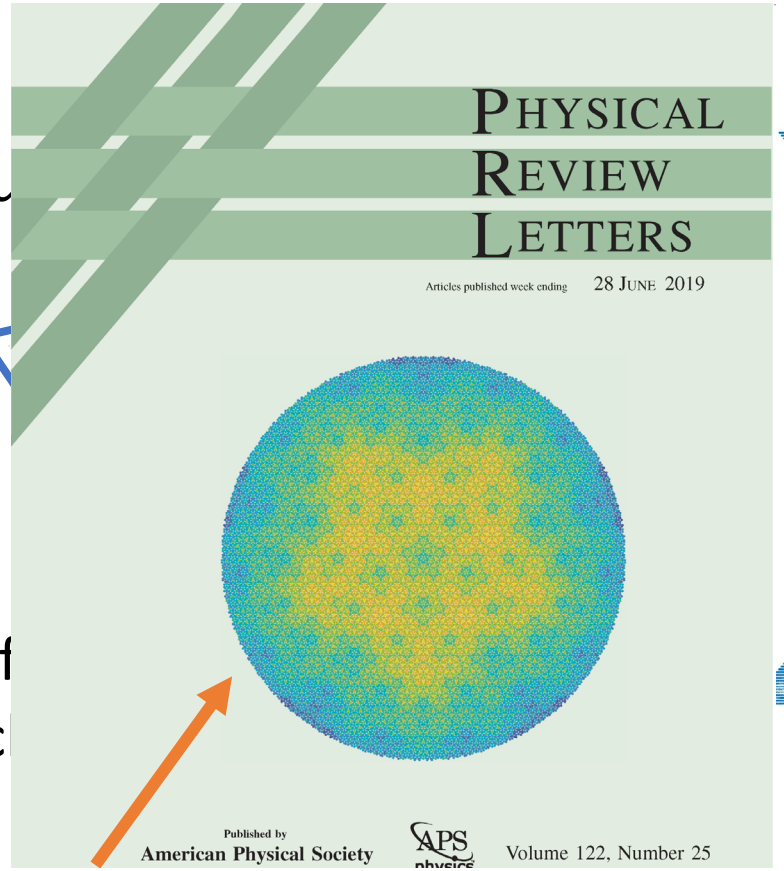


Convergent computation

Error control

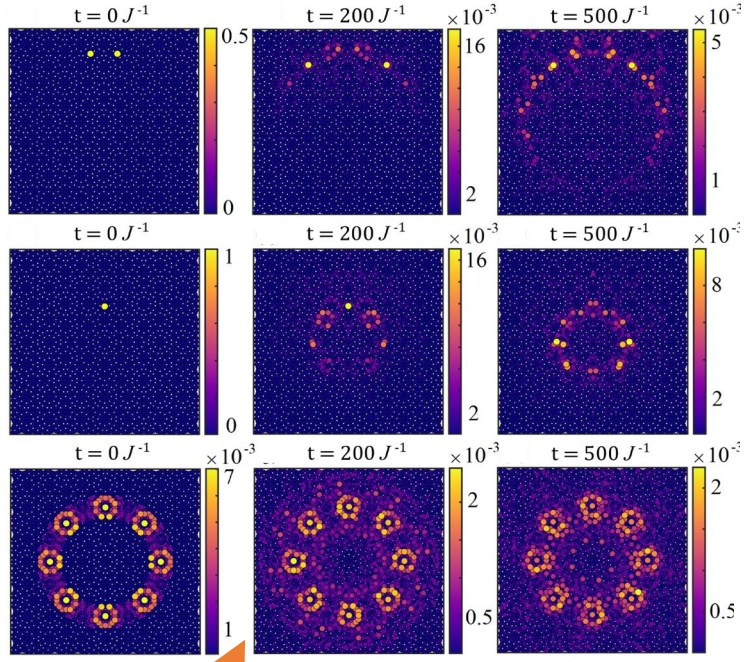
- C., Roman, Hansen, "*How to compute spectra with error control,*" **Physical Review Letters**, 2019.
- C., Horning, Townsend, "*Computing spectral measures of self-adjoint operators,*" **SIAM Review**, 2021.
- Johnstone, C., Nielsen, Öhberg, Duncan, "*Bulk Localised Transport States in Infinite and Finite Quasicrystals via Magnetic Aperiodicity.*"

Example of “too much” (spectral pollution)



E.g., ground state of quasicrystal

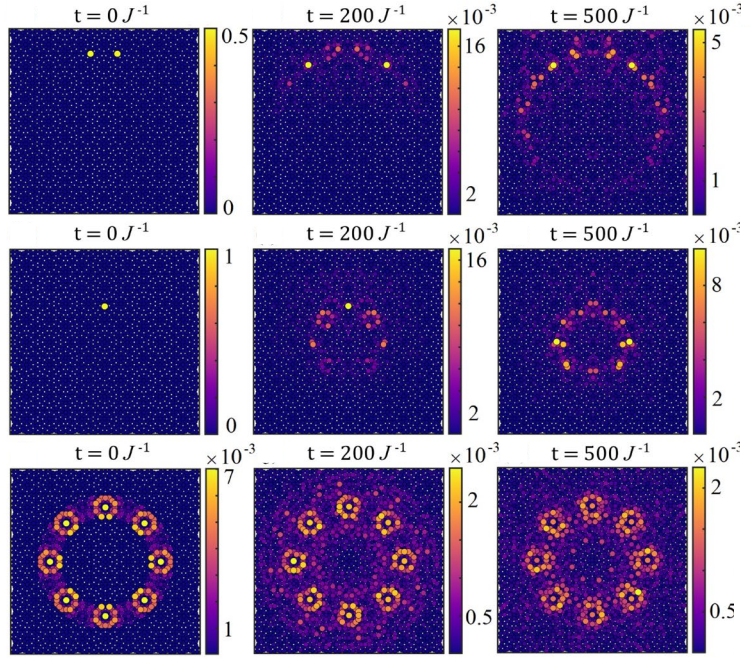
Spectral methods for continuous spectra
E.g., cts spec of graphene



E.g., new states and phenomena: bulk localised transport states

- C., Roman, Hansen, “*How to compute spectra with error control,*” **Physical Review Letters**, 2019.
- C., Horning, Townsend, “*Computing spectral measures of self-adjoint operators,*” **SIAM Review**, 2021.
- Johnstone, C., Nielsen, Öhberg, Duncan, “*Bulk Localised Transport States in Infinite and Finite Quasicrystals via Magnetic Aperiodicity.*”

Example of “too much” (spectral pollution)



Need new tools for data-driven dynamical systems …

- C., Roman, Hansen, “*How to compute spectra with error control,*” **Physical Review Letters**, 2019.
- C., Horning, Townsend, “*Computing spectral measures of self-adjoint operators,*” **SIAM Review**, 2021.
- Johnstone, C., Nielsen, Öhberg, Duncan, “*Bulk Localised Transport States in Infinite and Finite Quasicrystals via Magnetic Aperiodicity.*”

Build the matrix: Dynamic Mode Decomposition (DMD)

Given dictionary $\{\psi_1, \dots, \psi_{N_K}\}$ of functions $\psi_j : \Omega \to \mathbb{C}$

$$\{x^{(m)}, y^{(m)} = F(x^{(m)})\}_{m=1}^{M}$$

$$\langle \psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \psi_k(x^{(m)}) = \left[\underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \cdots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \cdots & \psi_{N_K}(x^{(M)}) \end{pmatrix}^*}_{\Psi_X} \underbrace{\begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_M \end{pmatrix}}_{W} \underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \cdots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \cdots & \psi_{N_K}(x^{(M)}) \end{pmatrix}}_{\Psi_X} \right]_{jk}$$

$$\langle \mathcal{K}\psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})} = \left[\underbrace{\begin{pmatrix} \psi_1(x^{(1)}) & \cdots & \psi_{N_K}(x^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(x^{(M)}) & \cdots & \psi_{N_K}(x^{(M)}) \end{pmatrix}^*}_{\Psi_X} \underbrace{\begin{pmatrix} w_1 & & \\ & \ddots & \\ & & w_M \end{pmatrix}}_{W} \underbrace{\begin{pmatrix} \psi_1(y^{(1)}) & \cdots & \psi_{N_K}(y^{(1)}) \\ \vdots & \ddots & \vdots \\ \psi_1(y^{(M)}) & \cdots & \psi_{N_K}(y^{(M)}) \end{pmatrix}}_{\Psi_Y} \right]_{jk}$$

$$\mathcal{K} \longrightarrow \mathbb{K} = (\Psi_X{}^* W \Psi_X)^{-1} \Psi_X{}^* W \Psi_Y \in \mathbb{C}^{N_K \times N_K}$$

Recall open problems: 1) "too little", 2) "too much", 3) lose continuous spectra, 4) verification.

- Schmid, "*Dynamic mode decomposition of numerical and experimental data,*" **Journal of fluid mechanics**, 2010.
- Kutz, Brunton, Brunton, Proctor, "*Dynamic mode decomposition: data-driven modeling of complex systems,*" **SIAM**, 2016.
- Williams, Kevrekidis, Rowley "*A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,*" **Journal of Nonlinear Science**, 2015.

Residual DMD (ResDMD): Approx. $\mathcal{K}$ and $\mathcal{K}^*\mathcal{K}$

$$\langle \psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})}\, \psi_k(x^{(m)}) = \left[\underbrace{\Psi_X{}^* W \Psi_X}_{G}\right]_{jk}$$

$$\langle \mathcal{K}\psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})}\, \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})} = \left[\underbrace{\Psi_X{}^* W \Psi_Y}_{K_1}\right]_{jk}$$

$$\langle \mathcal{K}\psi_k, \mathcal{K}\psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(y^{(m)})}\, \psi_k(y^{(m)}) = \left[\underbrace{\Psi_Y{}^* W \Psi_Y}_{K_2}\right]_{jk}$$

Residuals: $g = \sum_{j=1}^{N_K} \mathbf{g}_j \psi_j,\ \|\mathcal{K}g - \lambda g\|^2 \approx \mathbf{g}^*[K_2 - \lambda K_1{}^* - \bar{\lambda} K_1 + |\lambda|^2 G]\mathbf{g}$

- C., Townsend, "*Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems*," **Communications on Pure and Applied Mathematics**, under review.
- Code: https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition
- C., "*Rigorous and data-driven Koopmanism*," **XXI Householder Symposium** (invited plenary).

Convergence of quadrature

$$\text{E.g.,} \quad \langle \mathcal{K}\psi_k, \psi_j \rangle = \lim_{M\to\infty} \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})}$$

Three examples:

- **High-order quadrature:** $\{x^{(m)}, w_m\}_{m=1}^{M}$ M-point quadrature rule.
  Rapid convergence. Requires free choice of $\{x^{(m)}\}_{m=1}^{M}$ and small d.
- **Random sampling:** $\{x^{(m)}\}_{m=1}^{M}$ selected at random.
  Large d. Slow Monte Carlo $O(M^{-1/2})$ rate of convergence.
- **Ergodic sampling:** $x^{(m+1)} = F(x^{(m)})$.
  Single trajectory, large d. Requires ergodicity, convergence can be slow.

Most common

ResDMD: avoiding spectral pollution

$$\mathrm{res}(\lambda, \mathbf{g})^2 = \frac{\mathbf{g}^*\left[K_2 - \lambda K_1{}^* - \bar{\lambda} K_1 + |\lambda|^2 G\right]\mathbf{g}}{\mathbf{g}^* G \mathbf{g}}$$

Algorithm 1:

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$ and eigendecomposition $K_1 V = G V \Lambda$.
2. For each eigenpair $(\lambda, \mathbf{v})$, compute $\mathrm{res}(\lambda, \mathbf{v})$.
3. **Output:** subset of e-vectors $V_{(\varepsilon)}$ & e-vals $\Lambda_{(\varepsilon)}$ with $\mathrm{res}(\lambda, \mathbf{v}) \leq \varepsilon$ ($\varepsilon =$ input tol).

Theorem (no spectral pollution): Suppose quad. rule converges. Then

$$\limsup_{M \to \infty} \max_{\lambda \in \Lambda^{(\varepsilon)}} \|(\mathcal{K} - \lambda)^{-1}\|^{-1} \leq \varepsilon$$

ResDMD: avoiding spectral pollution

$$\mathrm{res}(\lambda, \mathbf{g})^2 = \frac{\mathbf{g}^*\left[K_2 - \lambda K_1{}^* - \bar{\lambda} K_1 + |\lambda|^2 G\right]\mathbf{g}}{\mathbf{g}^* G \mathbf{g}}$$

Algorithm 1:

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$ and eigendecomposition $K_1 V = G V \Lambda$.
2. For each eigenpair $(\lambda, \mathbf{v})$, compute $\mathrm{res}(\lambda, \mathbf{v})$.
3. **Output:** subset of e-vectors $V_{(\varepsilon)}$ & e-vals $\Lambda_{(\varepsilon)}$ with $\mathrm{res}(\lambda, \mathbf{v}) \leq \varepsilon$ ($\varepsilon =$ input tol).

Theorem (no spectral pollution): Suppose quad. rule converges. Then

$$\limsup_{M \to \infty} \max_{\lambda \in \Lambda^{(\varepsilon)}} \|(\mathcal{K} - \lambda)^{-1}\|^{-1} \leq \varepsilon$$

BUT: Typically, does not capture all of spectrum! ("too little")

ResDMD: computing pseudospectra and spectra

$$\mathrm{Spec}_\varepsilon(\mathcal{K}) = \bigcup_{\|\mathcal{B}\|\le\varepsilon} \mathrm{Spec}(\mathcal{K}+\mathcal{B}), \qquad \lim_{\varepsilon\downarrow 0} \mathrm{Spec}_\varepsilon(\mathcal{K}) = \mathrm{Spec}(\mathcal{K})$$

Algorithm 2:

First convergent method for general $\mathcal{K}$

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$.
2. For z_k in comp. grid, compute $\tau_k = \min_{g=\sum_{j=1}^{N} \mathbf{g}_j \psi_j} \mathrm{res}(z_k, g)$, corresponding g_k (gen. SVD).
3. **Output:** $\{z_k : \tau_k < \varepsilon\}$ (approx. of $\mathrm{Spec}_\varepsilon(\mathcal{K})$), $\{g_k : \tau_k < \varepsilon\}$ (ε-pseudo-eigenfunctions).

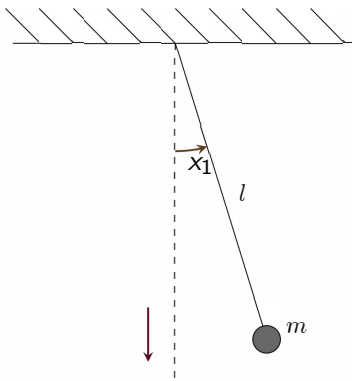
Theorem (full convergence): Suppose the quadrature rule converges.

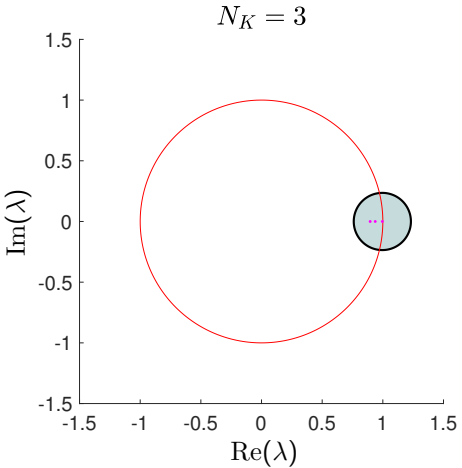
- **Error control:** $\{z_k : \tau_k < \varepsilon\} \subseteq \mathrm{Spec}_\varepsilon(\mathcal{K})$ (as $M \to \infty$)
- **Convergence:** Converges locally uniformly to $\mathrm{Spec}_\varepsilon(\mathcal{K})$ (as $N_K \to \infty$)

NB: Local optimisation strategy shrinks ε to compute $\mathrm{Spec}(\mathcal{K})$

Example: non-linear pendulum

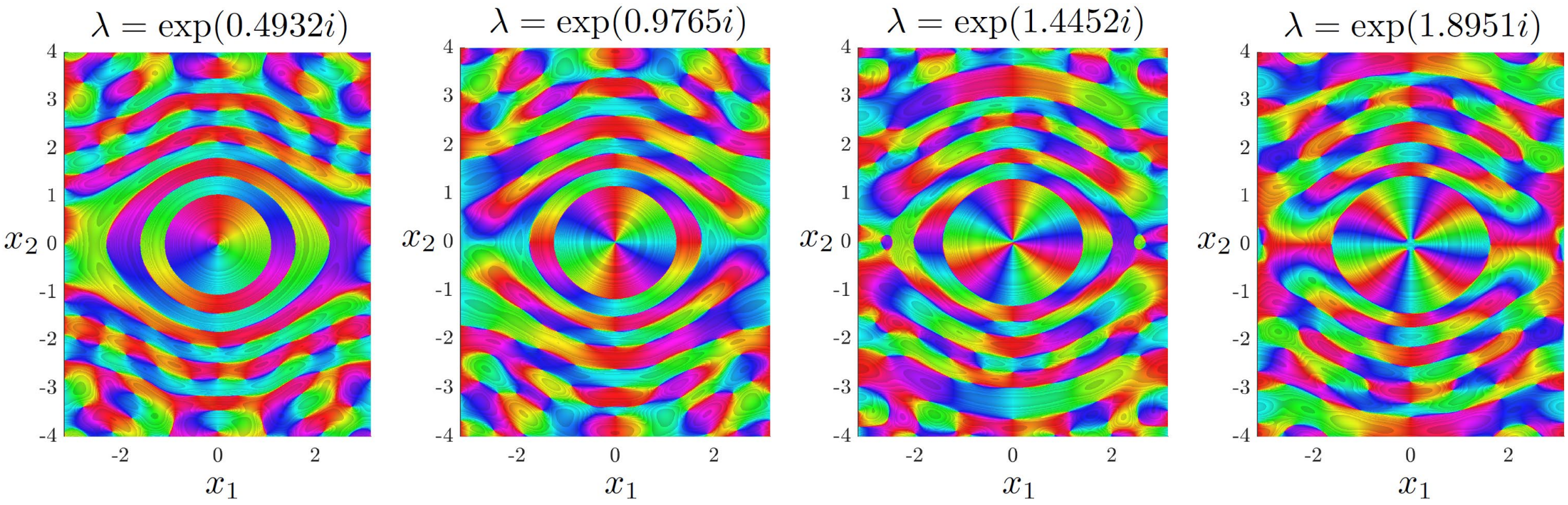
$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = -\sin(x_1), \qquad \Omega = [-\pi, \pi] \times \mathbb{R}$$





Computed pseudospectra ($\varepsilon = 0.25$). Eigenvalues of $\mathbb{K}$ shown as dots (spectral pollution).

Example: non-linear pendulum



Colour represents complex argument, constant modulus shown as shadowed steps.
All residuals smaller than $\varepsilon = 0.05$ (made smaller by increasing N_K).

Setup for continuous spectra

Suppose system is measure preserving (e.g., Hamiltonian, ergodic, …)

$$\Leftrightarrow \mathcal{K}^*\mathcal{K} = I \text{ (isometry)}$$

$$\Longrightarrow \mathrm{Spec}(\mathcal{K}) \subseteq \{z : |z| \leq 1\}$$

(For those interested: we consider canonical unitary extensions.)

Spectral measures → diagonalisation

- **Fin.-dim.:** $B \in \mathbb{C}^{n\times n}$, $B^*B = BB^*$, o.n. basis of e-vectors $\{v_j\}_{j=1}^{n}$

$$v = \left[\sum_{j=1}^{n} v_j {v_j}^*\right] v, \qquad Bv = \left[\sum_{j=1}^{n} \lambda_j v_j {v_j}^*\right] v, \qquad \forall v \in \mathbb{C}^n$$

- **Inf.-dim.:** Operator $\mathcal{L}\colon \mathcal{D}(\mathcal{L}) \to \mathcal{H}$. Typically, no basis of e-vectors!
  Spectral theorem: (projection-valued) spectral measure E

$$g = \left[\int_{\mathrm{Spec}(\mathcal{L})} 1 \,\mathrm{d}E(\lambda)\right] g, \qquad \mathcal{L}g = \left[\int_{\mathrm{Spec}(\mathcal{L})} \lambda \,\mathrm{d}E(\lambda)\right] g, \qquad \forall g \in \mathcal{H}$$

- **Spectral measures:** $\nu_g(U) = \langle E(U)g, g\rangle$ ($\|g\| = 1$) prob. measure.

Example

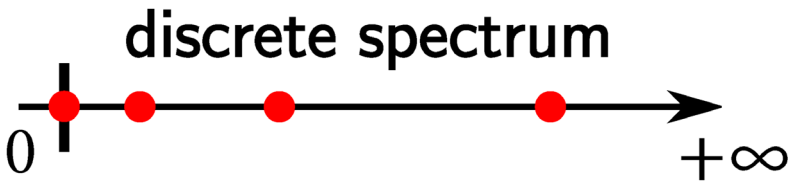
$$\mathcal{L} = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$$

Fourier transform

Spectral measures

On $[-\pi, \pi]_{\mathrm{per}}$

$$\mathrm{Spec}(\mathcal{L}) = \{n^2 : n \in \mathbb{Z}\}$$



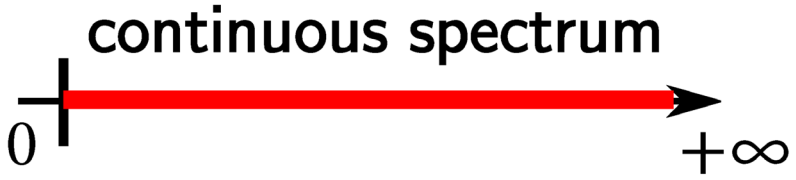
$$\hat{g}_k = \frac{1}{2\pi}\int_{[-\pi,\pi]_{\mathrm{per}}} g(x)e^{-ikx}\,\mathrm{d}x$$

$$[E([a,b])g](x) = \sum_{a \le k^2 \le b} \hat{g}_k e^{ikx}$$

$$\nu_g([a,b]) = \sum_{a \le k^2 \le b} |\hat{g}_k|^2$$

On $\mathbb{R}$

$$\mathrm{Spec}(\mathcal{L}) = [0, \infty)$$



$$\hat{g}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(x)e^{-ikx}\,\mathrm{d}x$$

$$[E([a,b])g](x) = \int_{a \le k^2 \le b} \hat{g}(k)e^{-ikx}\,\mathrm{d}x$$

$$\nu_g([a,b]) = \int_{a \le k^2 \le b} |\hat{g}(k)|^2 e^{-ikx}\,\mathrm{d}x$$

Koopman mode decomposition (again!)

ν_g probability measures on $[-\pi, \pi]_{\text{per}}$

Leb. decomp: $$\mathrm{d}\nu_g(y) = \underbrace{\sum_{\text{eigenvalues } \lambda_j = \exp(i\theta_j)} \left\langle P_{\lambda_j} g, g \right\rangle \delta(y - \theta_j)}_{\text{discrete}} + \underbrace{\rho_g(y)\mathrm{d}y + \mathrm{d}\nu_g^{\mathrm{sc}}(y)}_{\text{continuous}}$$

$$g(x) = \sum_{\text{eigenvalues } \lambda_j} c_{\lambda_j} \varphi_{\lambda_j}(x) + \int_{-\pi}^{\pi} \phi_{\theta,g}(x)\, \mathrm{d}\theta$$

eigenfunction of $\mathcal{K}$

generalised eigenfunction of $\mathcal{K}$

$$g(x_n) = [\mathcal{K}^n g](x_0) = \sum_{\text{eigenvalues } \lambda_j} c_{\lambda_j} \lambda_j{}^n \varphi_{\lambda_j}(x_0) + \int_{-\pi}^{\pi} e^{in\theta} \phi_{\theta,g}(x_0)\, \mathrm{d}\theta$$

Computing ν_g diagonalises non-linear dynamical system!

Plemelj-type formula

$$\mathcal{C}_g(z) = \int_{-\pi}^{\pi} \frac{e^{i\theta}\,\mathrm{d}\nu_g(\theta)}{e^{i\theta} - z} = \begin{cases} \langle (\mathcal{K} - zI)^{-1} g, \mathcal{K}^* g \rangle, & \text{if } |z| > 1 \\ -z^{-1} \langle g, (\mathcal{K} - \bar{z}^{-1} I)^{-1} g \rangle, & \text{if } 0 < |z| < 1 \end{cases}$$

ResDMD computes with error control

$$P_\varepsilon(\theta_0) = \frac{1}{2\pi} \frac{(1+\varepsilon)^2 - 1}{1 + (1+\varepsilon)^2 - 2(1+\varepsilon)\cos(\theta_0)}$$

Poisson kernel for unit disk

$\varepsilon =$ "smoothing parameter"

$$[P_\varepsilon * \nu_g](\theta_0) = \int_{-\pi}^{\pi} P_\varepsilon(\theta_0 - \theta)\,\mathrm{d}\nu_g(\theta) = \mathcal{C}_g\left(e^{i\theta_0}(1+\varepsilon)^{-1}\right) - \mathcal{C}_g\left(e^{i\theta_0}(1+\varepsilon)\right)$$

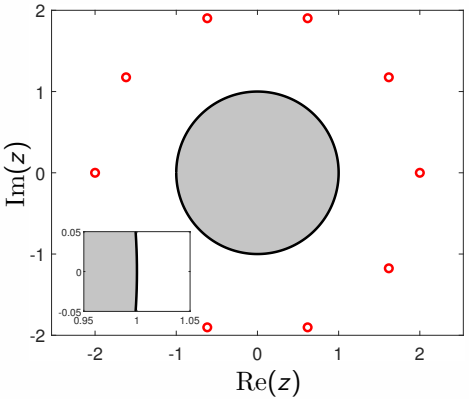
Example

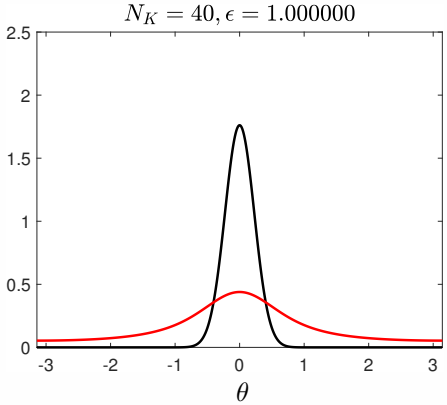
$$\mathcal{K} = \begin{pmatrix} \overline{\alpha_0} & \overline{\alpha_1}\rho_0 & \rho_0\rho_1 & & & \\ \rho_0 & -\overline{\alpha_1}\alpha_0 & -\alpha_0\rho_1 & & & \\ & \overline{\alpha_2}\rho_1 & -\overline{\alpha_2}\alpha_1 & \overline{\alpha_3}\rho_2 & \rho_3\rho_2 & \\ & \rho_2\rho_1 & -\alpha_1\rho_2 & -\overline{\alpha_3}\alpha_2 & -\rho_3\alpha_2 & \ddots \\ & & & \overline{\alpha_4}\rho_3 & -\overline{\alpha_4}\alpha_3 & \ddots \\ & & & \ddots & \ddots & \ddots \end{pmatrix}$$

$$\alpha_j = (-1)^j 0.95^{(j+1)/2}, \qquad \rho_j = \sqrt{1 - |\alpha_j|^2}$$

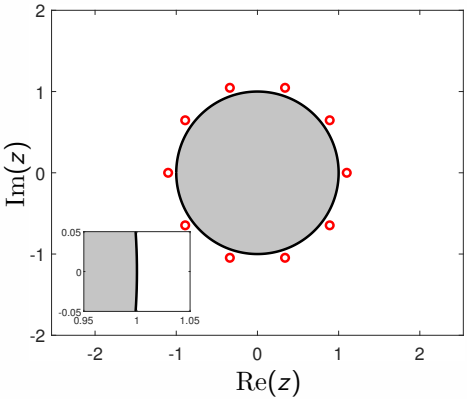
Generalised shift, typical building block of many dynamical systems.

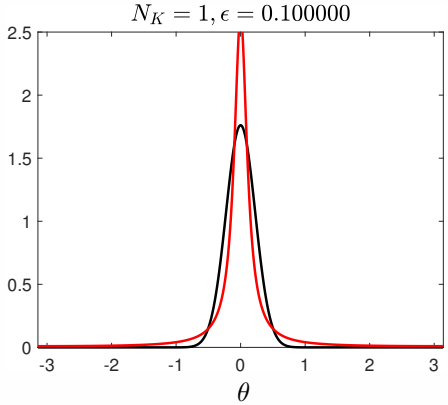
Fix N_K, vary ε: unstable!



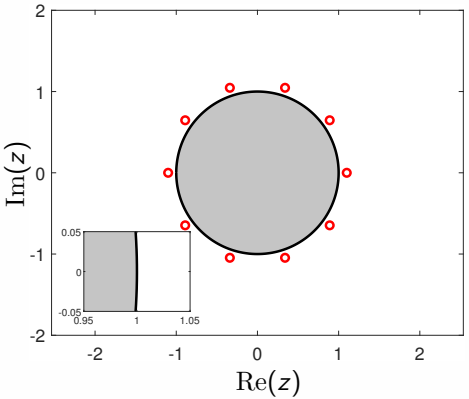


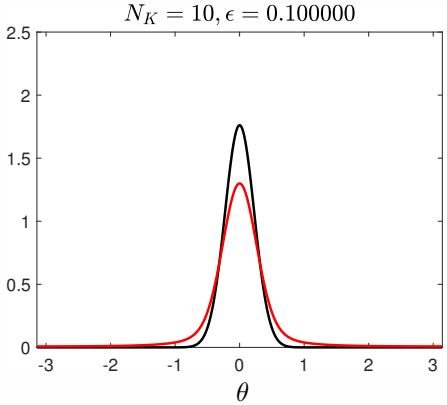
Fix ε, vary N_K: too smooth!





Adaptive: new matrix to compute residuals crucial

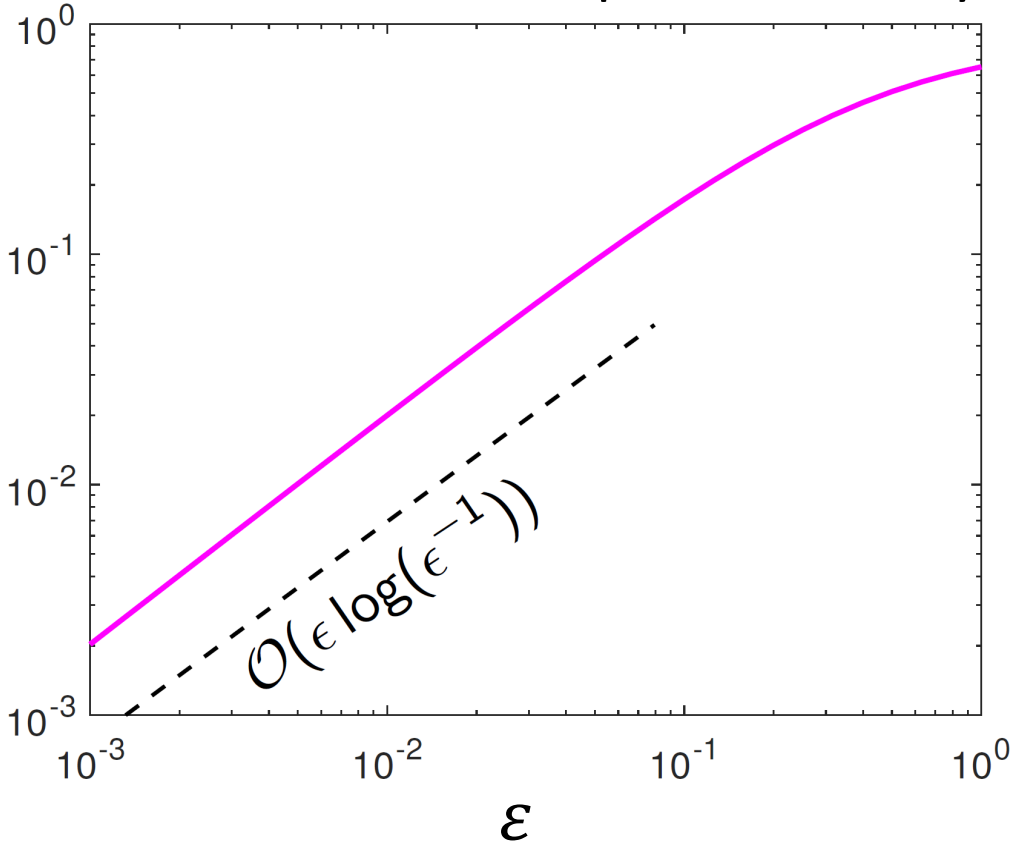




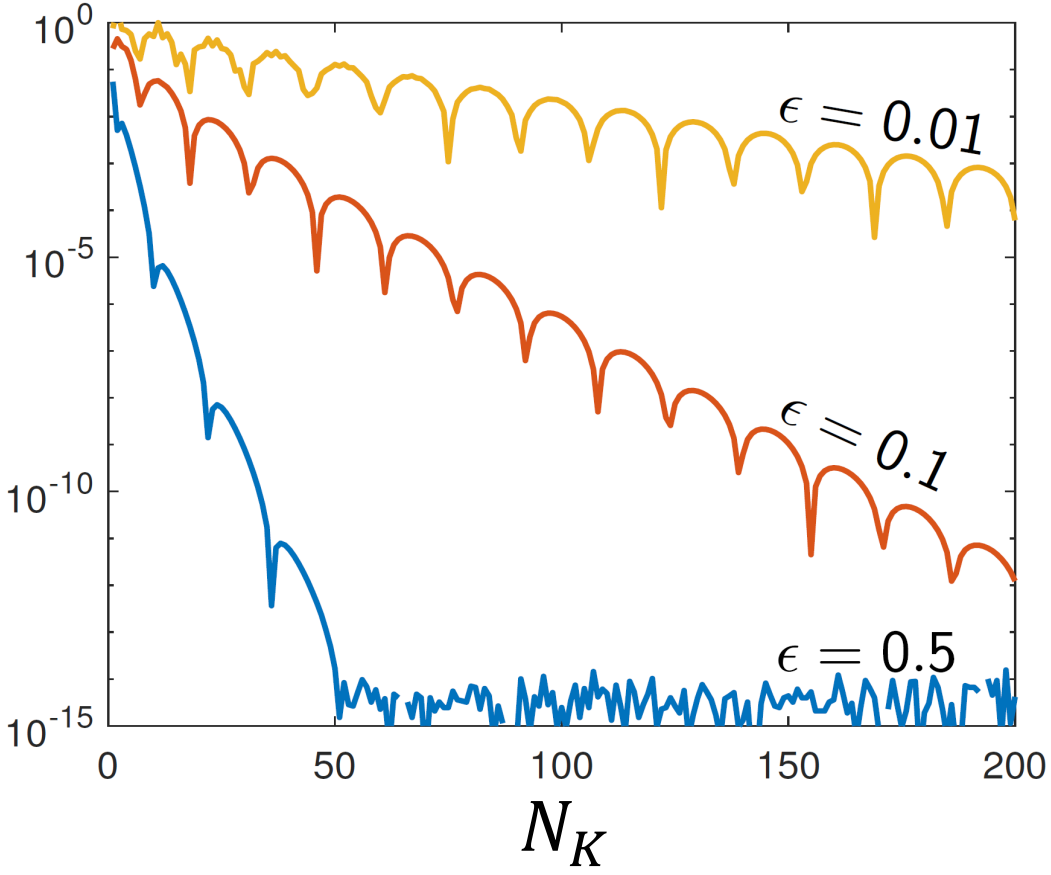
But … slow convergence

Problem: As $\varepsilon \downarrow 0$, error is $O(\varepsilon \cdot \log(1/\varepsilon))$ and $N_K(\varepsilon) \to \infty$.

Pointwise error for spectral density



Error due to discretisation



Small N_K critical in <u>data-driven</u> computations. Can we improve convergence rate?

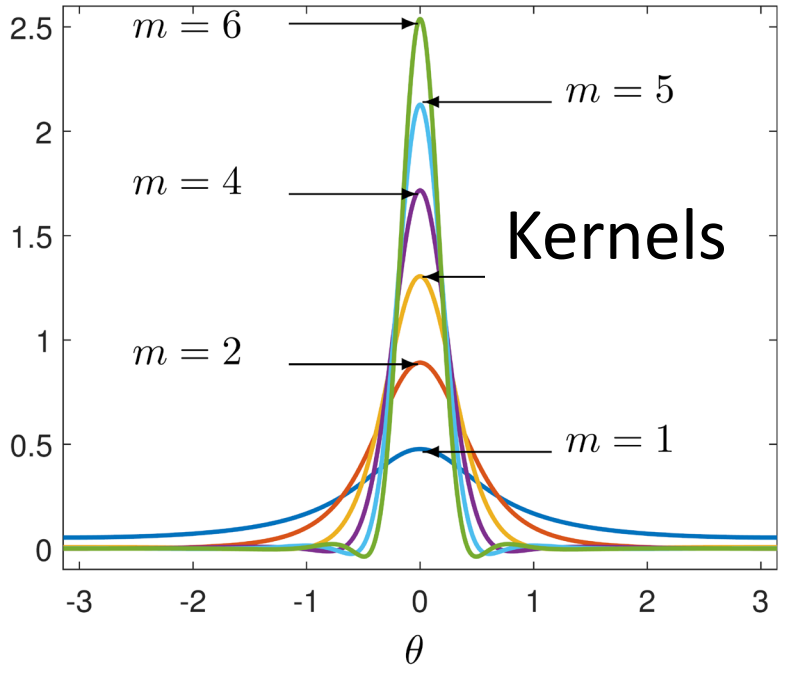
High-order rational kernels

mth order rational kernels:

$$K_\varepsilon(\theta) = \frac{e^{-i\theta}}{2\pi} \sum_{j=1}^{m} \left[\frac{c_j}{e^{-i\theta} - (1+\varepsilon\overline{z_j})^{-1}} - \frac{d_j}{e^{-i\theta} - \left(1+\varepsilon z_j\right)} \right]$$

$$\left[K_\varepsilon * \nu_g\right](\theta_0) = \sum_{j=1}^{m} \left[c_j \mathcal{C}_g\left(e^{i\theta_0}(1+\varepsilon\overline{z_j})^{-1}\right) - d_j \mathcal{C}_g\left(e^{i\theta_0}\left(1+\varepsilon z_j\right)\right) \right]$$

- Theory providing $\{c_j, d_j, z_j\}$
- Convolution computed with error control.
- $O(PN_K)$ cost for evaluation at P values of θ.



High-order rational kernels

Algorithm 3 A computational framework for evaluating an approximate spectral measure with respect to $g \in L^2(\Omega, \omega)$ at $\{\theta_k\}_{k=1}^P \subset [-\pi, \pi]_{\text{per}}$ of an isometry $\mathcal{K}$ using snapshot data.

Input: Snapshot data $\{\boldsymbol{x}^{(m)}\}_{m=1}^M, \{\boldsymbol{y}^{(m)}\}_{m=1}^M$ (such that $\boldsymbol{y}^{(m)} = F(\boldsymbol{y}^{(m)})$), quadrature weights $\{w_m\}_{m=1}^M$, a dictionary of observables $\{\psi_j\}_{j=1}^{N_K}$, $m \in \mathbb{N}$, smoothing parameter $0 < \epsilon < 1$ (accuracy goal is ϵ^m), and evaluation points $\{\theta_k\}_{k=1}^P \subset [-\pi, \pi]_{\text{per}}$.

1: Compute $\Psi_X^* W \Psi_X$, $\Psi_X^* W \Psi_Y$, and $\Psi_Y^* W \Psi_Y$.
2: Compute a generalized Schur decomposition of $\Psi_X^* W \Psi_Y$ and $\Psi_X^* W \Psi_X$, i.e., $\Psi_X^* W \Psi_Y = QSZ^*$ and $\Psi_X^* W \Psi_X = QTZ^*$, where Q, Z are unitary and S, T are upper triangular.
3: Compute $\boldsymbol{a}$ and $v_1 = TZ^*\boldsymbol{a}$, $v_2 = T^*Q^*\boldsymbol{a}$, and $v_3 = S^*Q^*\boldsymbol{a}$.
4: **for** $k = 1, \ldots, P$ **do**
5: Compute $I_j = (S - e^{i\theta_k}(1 + \epsilon z_j)T)^{-1} v_1$ for $1 \le j \le m$.
6: Compute $\nu_g^\epsilon(\theta_k) = \frac{-1}{2\pi} \sum_{j=1}^m \text{Re}\left[c_j(\epsilon) e^{-i\theta_k}(1 + \epsilon\overline{z_j})(I_j^* v_2) + d_j(v_3^* I_j)\right]$.
7: **end for**
8: Use $\Psi_Y^* W \Psi_Y$ to check whether N_K is too small (or ϵ too small). If so, increase N_K (or increase ϵ) and return to step 1.

Output: Values of the approximate spectral measure, i.e., $\{\nu_g^\epsilon(\theta_k)\}_{k=1}^P$.

Compute residuals

Rigorous convergence thm.

Convergence

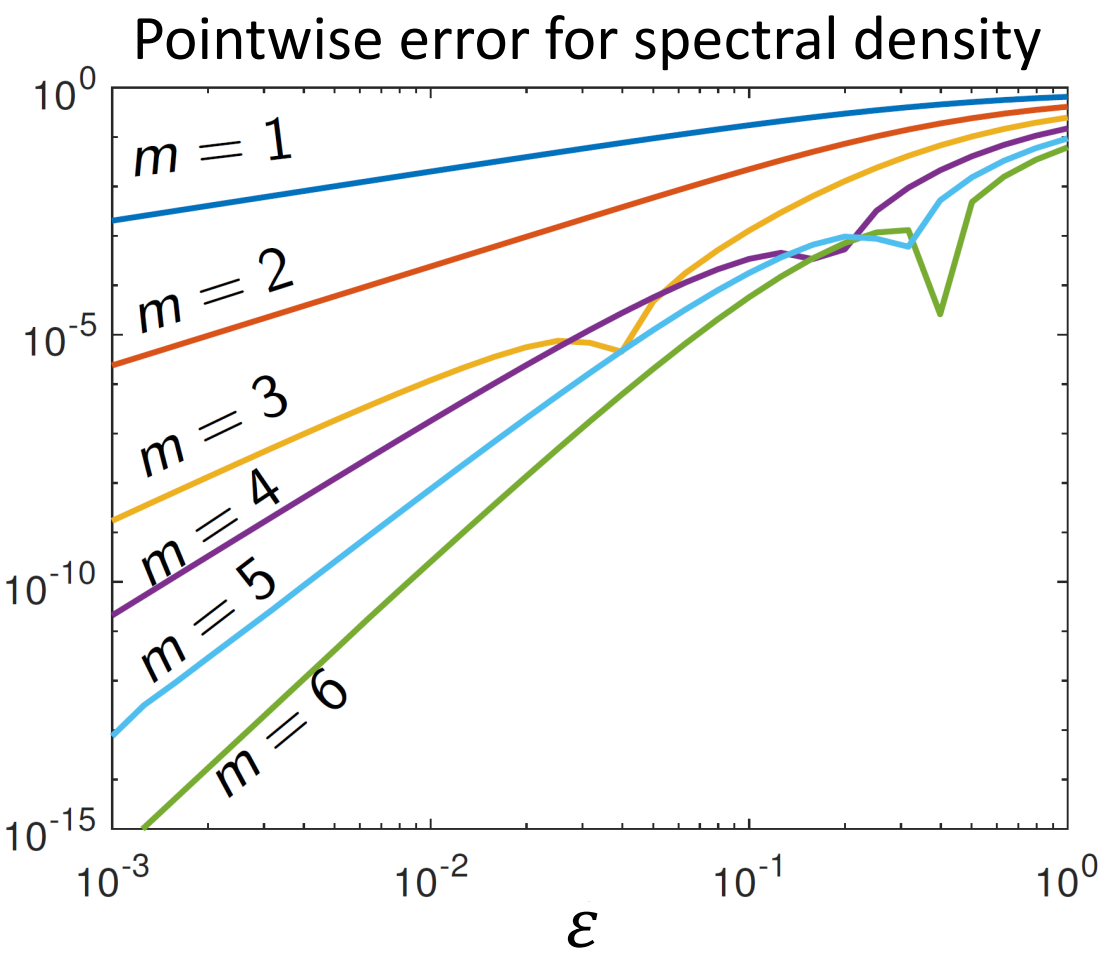
Theorem: *Automatic* selection of $N_K(\varepsilon)$ with $O(\varepsilon^m \log(1/\varepsilon))$ convergence:

- Density of continuous spectrum ρ_g.
  (pointwise and L^p)
- Integration against test functions.
  (weak convergence)

$$\int_{-\pi}^{\pi} h(\theta) \left[K_\varepsilon * \nu_g\right](\theta)\, \mathrm{d}\theta$$

$$= \int_{-\pi}^{\pi} h(\theta)\ \mathrm{d}\nu_g(\theta) + O(\varepsilon^m \log(1/\varepsilon))$$

Also recover discrete spectrum.



Large d ($\Omega \subseteq \mathbb{R}^d$): *robust* and *scalable*

Popular to learn dictionary $\{\psi_1, \dots, \psi_{N_K}\}$

E.g., DMD with truncated SVD (linear dictionary, most popular), kernel methods (this talk), neural networks, etc.

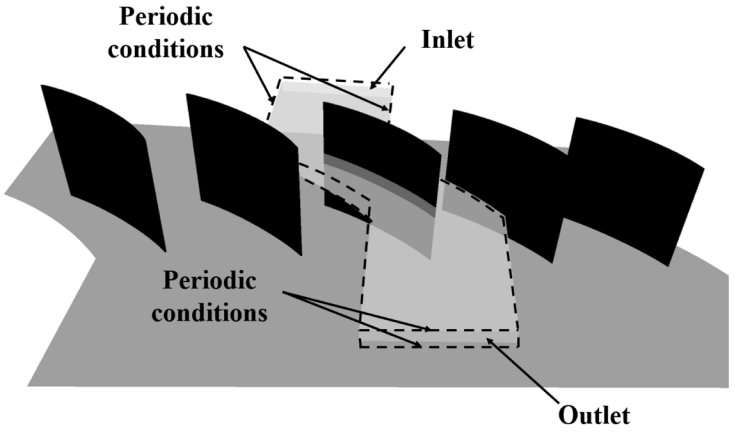
Q: Is discretisation $\mathbf{span}\{\psi_1, \dots, \psi_{N_K}\}$ large/rich enough?

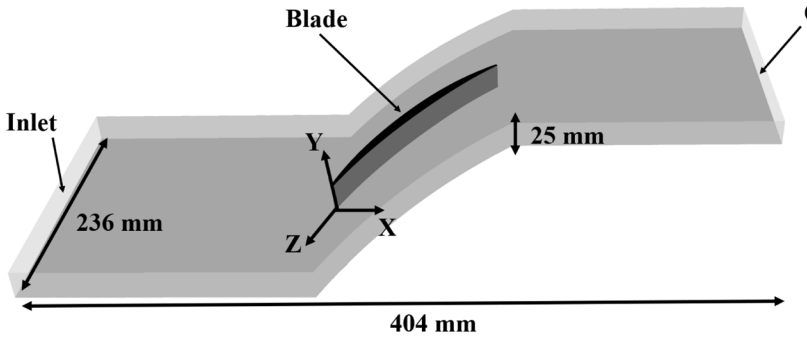
Above algorithms:

- Pseudospectra: $\{z_k : \tau_k < \varepsilon\} \subseteq \mathrm{Spec}_\varepsilon(\mathcal{K})$ **error control**
- Spectral measures: $\mathcal{C}_g(z)$ and smoothed measures **adaptive check**

$\Longrightarrow$ Rigorously ***verify*** learnt dictionary $\{\psi_1, \dots, \psi_{N_K}\}$

Example: pressure field of turbulent flow

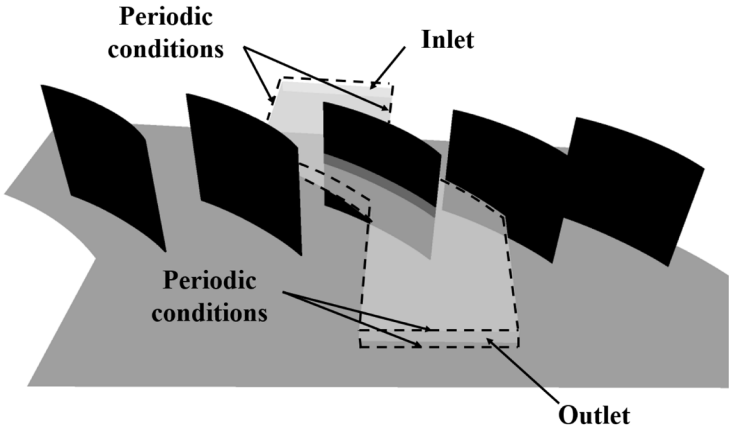


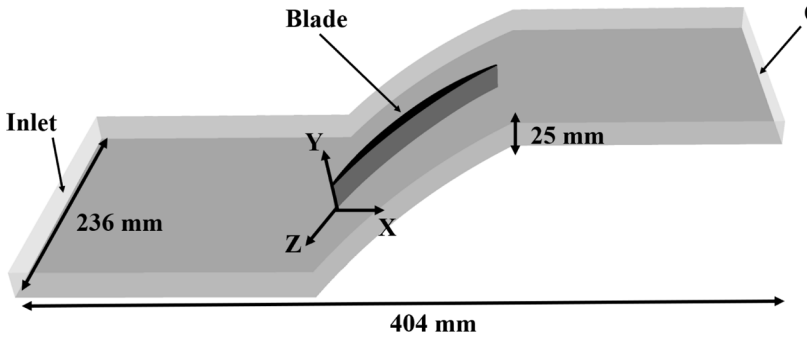


- Data collected for 2×10^{-4}s
- Reynolds number $\approx 3.9 \times 10^{5}$
- Ambient dimension $\approx 300{,}000$ (number of measurement points*)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

Example: pressure field of turbulent flow



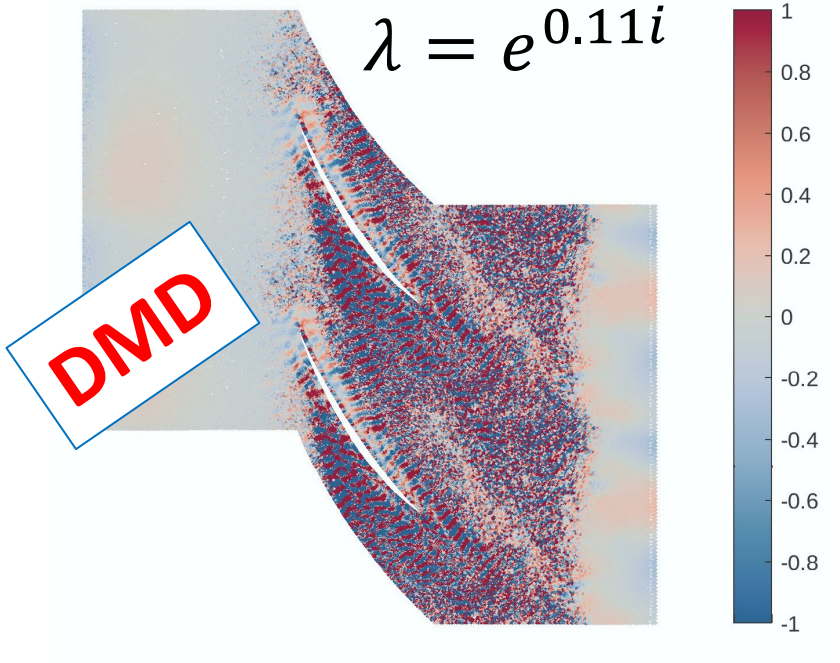


- Data collected for 2×10^{-4}s
- Reynolds number $\approx 3.9 \times 10^{5}$
- Ambient dimension $\approx 300{,}000$ (number of measurement points*)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

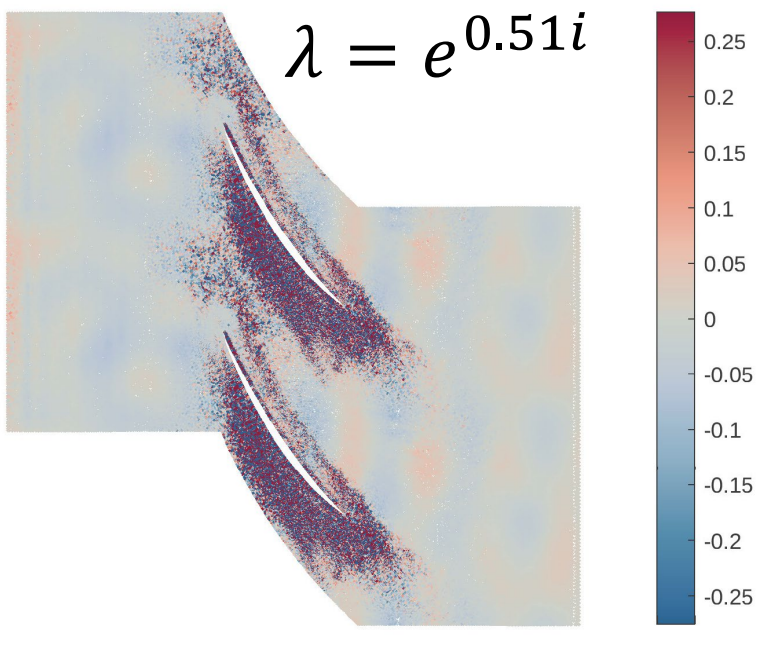
Rel. Error = ?

$\lambda = e^{0.11i}$



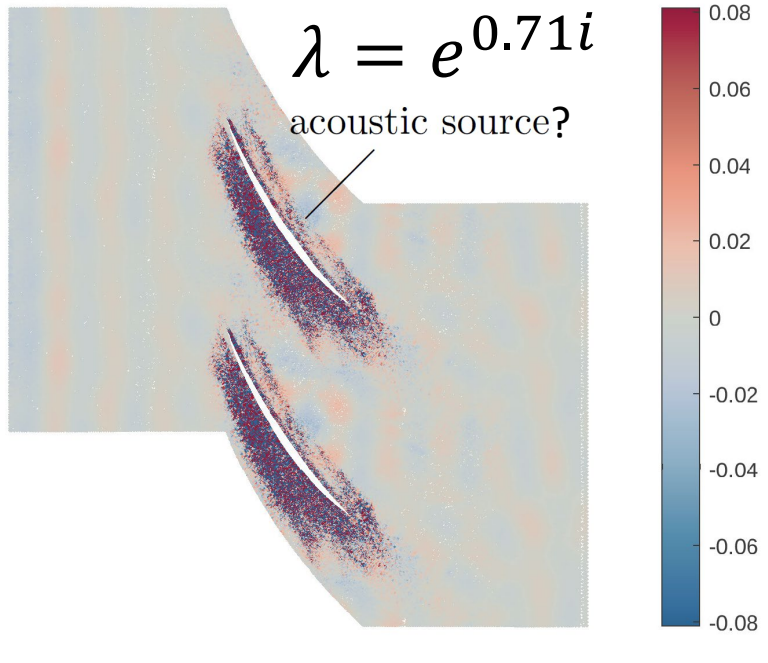
Rel. Error = ?

$\lambda = e^{0.51i}$

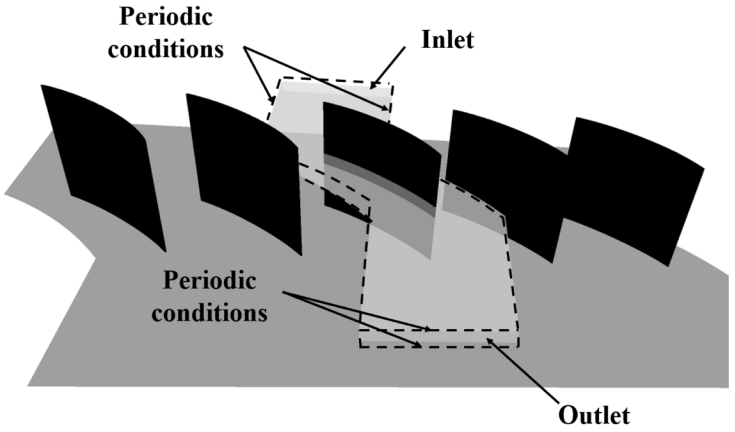


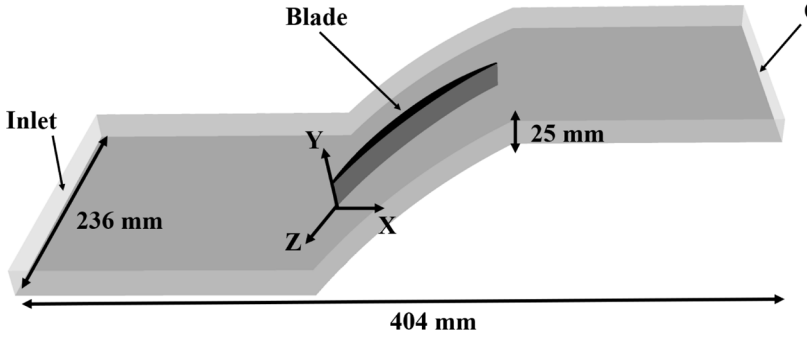
Rel. Error = ?

$\lambda = e^{0.71i}$



Example: pressure field of turbulent flow



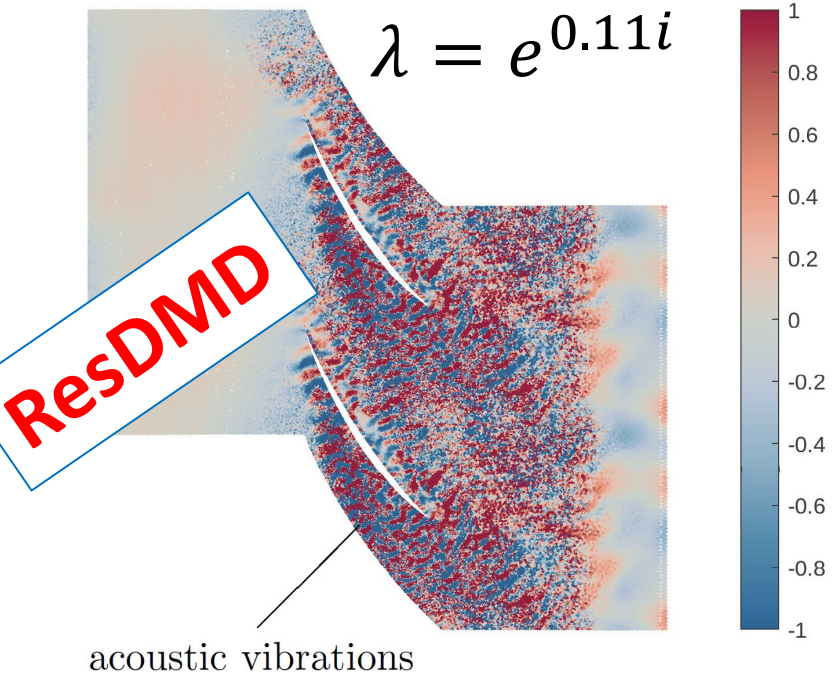


- Data collected for 2×10^{-4}s
- Reynolds number $\approx 3.9 \times 10^{5}$
- Ambient dimension $\approx 300{,}000$ (number of measurement points*)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

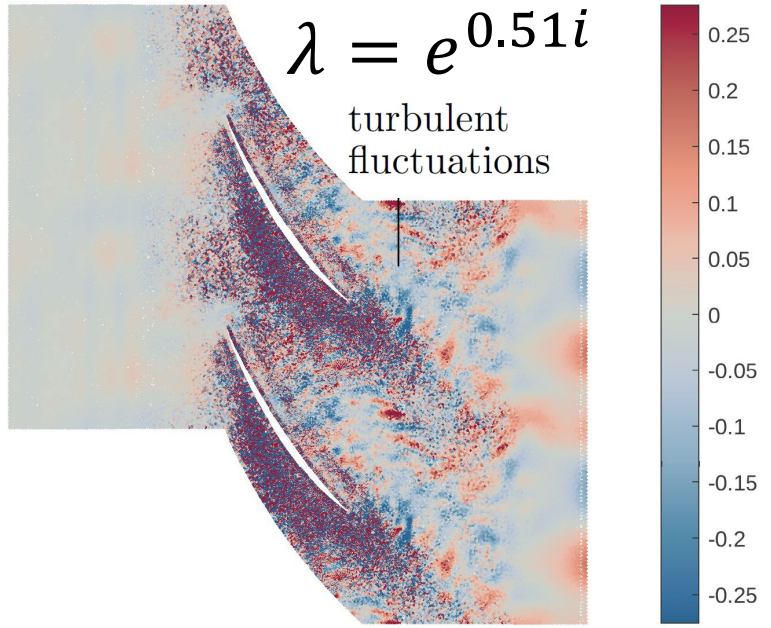
Rel. Error ≤ 0.0054

$\lambda = e^{0.11i}$



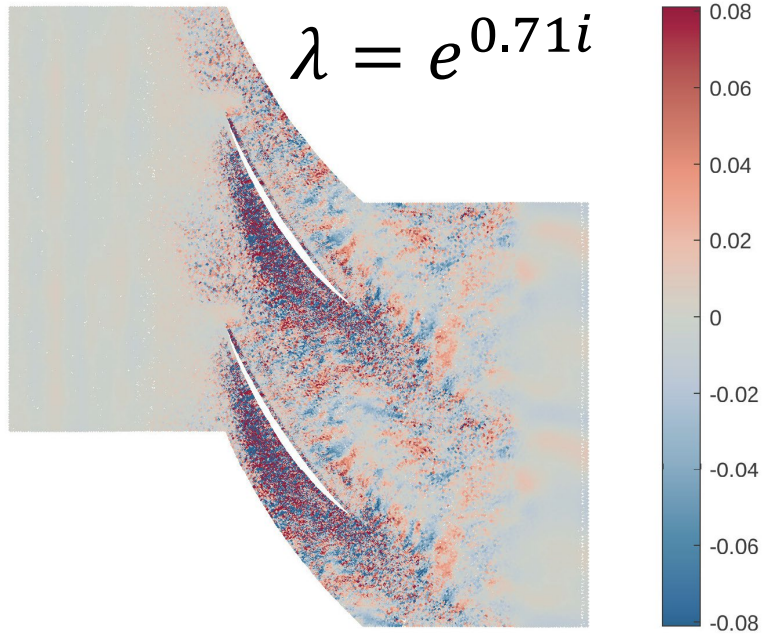
Rel. Error ≤ 0.0128

$\lambda = e^{0.51i}$

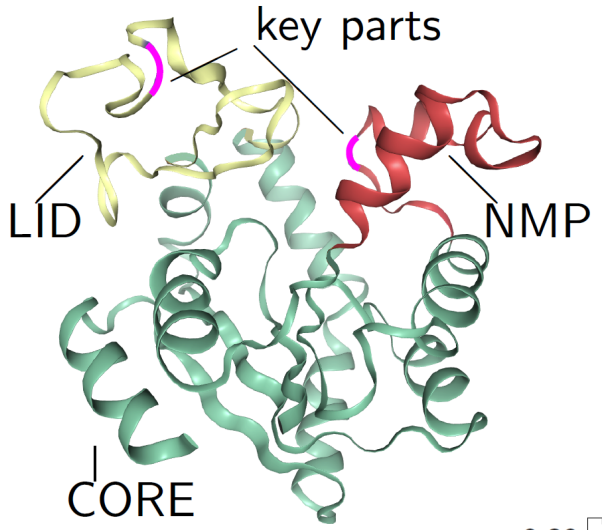


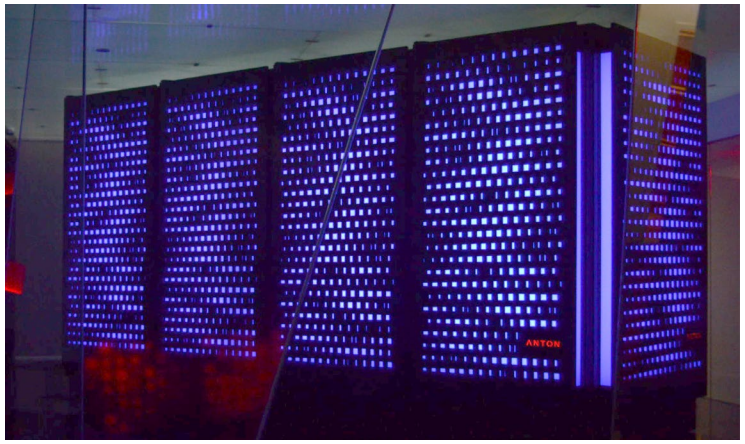
Rel. Error ≤ 0.0196

$\lambda = e^{0.71i}$



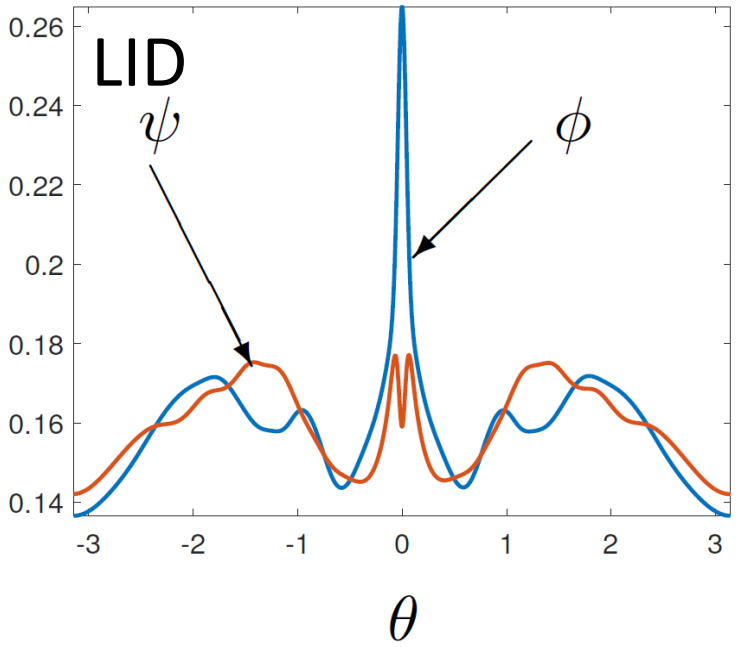
Example: molecular dynamics (Adenylate Kinase)

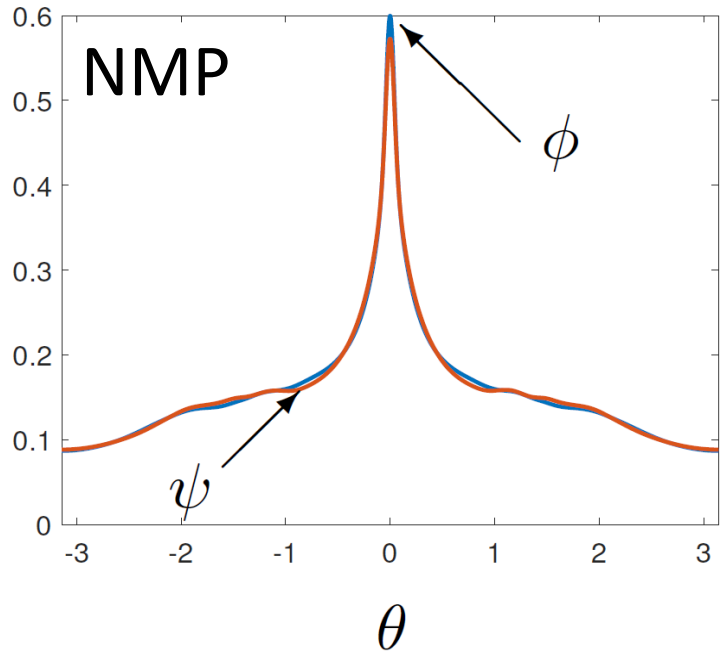




- All-atom equilibrium simulation for 1.004×10^{-6}s
- Ambient dimension $\approx 20{,}000$ (positions and momenta of atoms)
- 6th order kernel (spec res 10^{-6})

*Dataset: www.mdanalysis.org/MDAnalysisData/adk_equilibrium.html





Summary

Rigorous + practical data-driven algorithms for spectral properties of Koopman operators.

Overcame **problems: 1) too little, 2) too much, 3) continuous spectra, 4) verification.**

- Spectra, pseudospectra, and residuals of general Koopman operators (error control).
  - **Idea:** New matrix for residual ⇒ ResDMD.
- Spectral measures of measure-preserving systems with high-order convergence. Continuous spectra, discrete spectra, and weak convergence.
  - **Idea:** Convolution with rational kernels via resolvent and ResDMD.
- Dealt with high-dimensional dynamical systems.
  - **Idea:** Use ResDMD to verify learned dictionaries.

First general methods with convergence guarantees!

⟶ Opens the door to rigorous data-driven Koopmania!

Koopman mode decomposition ($\mathbb{K}V = V\Lambda$)

Standard Koopman mode decomposition (order modes by $|\Lambda|$):

$$g(x) \approx \underbrace{[\psi_1(x) \quad \cdots \quad \psi_{N_K}(x)]V}_{\text{approx Koopman e-functions}} \underbrace{\left(V\sqrt{W}\Psi_X\right)^{\dagger}\sqrt{W}[g(x^{(1)}) \quad \cdots \quad g(x^{(M)})]^{\mathrm{T}}}_{\text{Koopman modes}}$$

$$\overset{?}{\Rightarrow} g(x_n) \approx \underbrace{[\psi_1(x) \quad \cdots \quad \psi_{N_K}(x)]V}_{\text{approx Koopman e-functions}} \Lambda^n \underbrace{\left(V\sqrt{W}\Psi_X\right)^{\dagger}\sqrt{W}[g(x^{(1)}) \quad \cdots \quad g(x^{(M)})]^{\mathrm{T}}}_{\text{Koopman modes}}$$

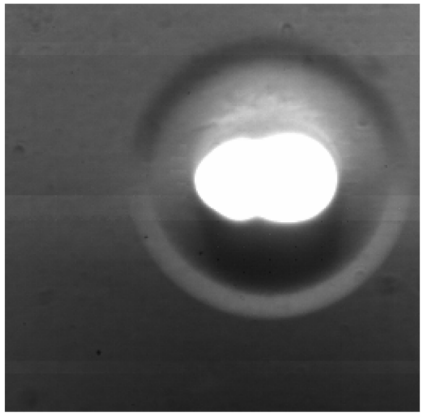
Residual Koopman mode decomposition (order modes by $\mathrm{res}(\lambda, \mathbf{v})$):

Controllable error

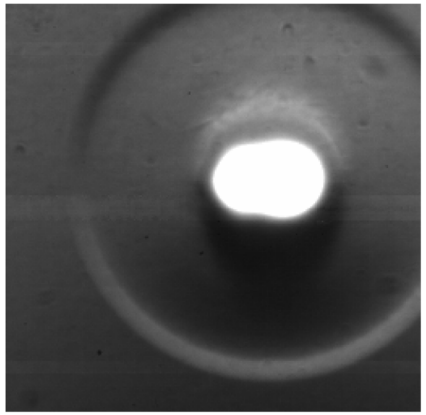
$$g(x) \approx \underbrace{[\psi_1(x) \quad \cdots \quad \psi_{N_K}(x)]V_{(\varepsilon)}}_{\text{approx Koopman e-functions}} \underbrace{\left(V_{(\varepsilon)}\sqrt{W}\Psi_X\right)^{\dagger}\sqrt{W}[g(x^{(1)}) \quad \cdots \quad g(x^{(M)})]^{\mathrm{T}}}_{\text{Koopman modes}}$$

$$g(x_n) \approx \underbrace{[\psi_1(x_0) \quad \cdots \quad \psi_{N_K}(x_0)]V_{(\varepsilon)}}_{\text{approx Koopman e-functions}} \Lambda_{(\varepsilon)}^n \underbrace{\left(V_{(\varepsilon)}\sqrt{W}\Psi_X\right)^{\dagger}\sqrt{W}[g(x^{(1)}) \quad \cdots \quad g(x^{(M)})]^{\mathrm{T}}}_{\text{Koopman modes}}$$

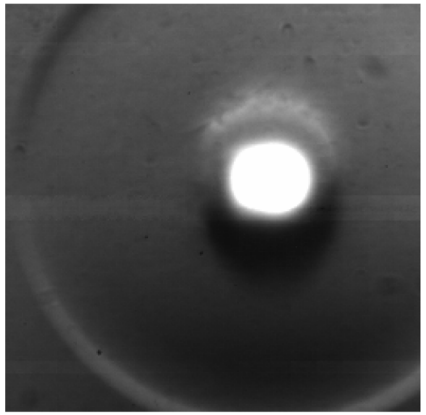
Example: laser-induced plasma



a) $t = 5\ \mu s$



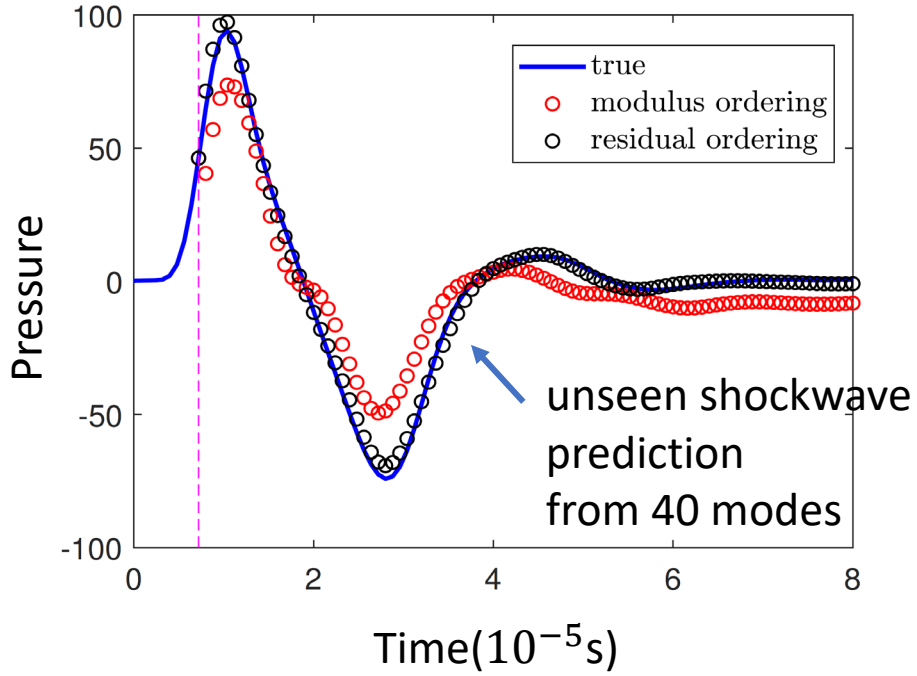
b) $t = 10\ \mu s$

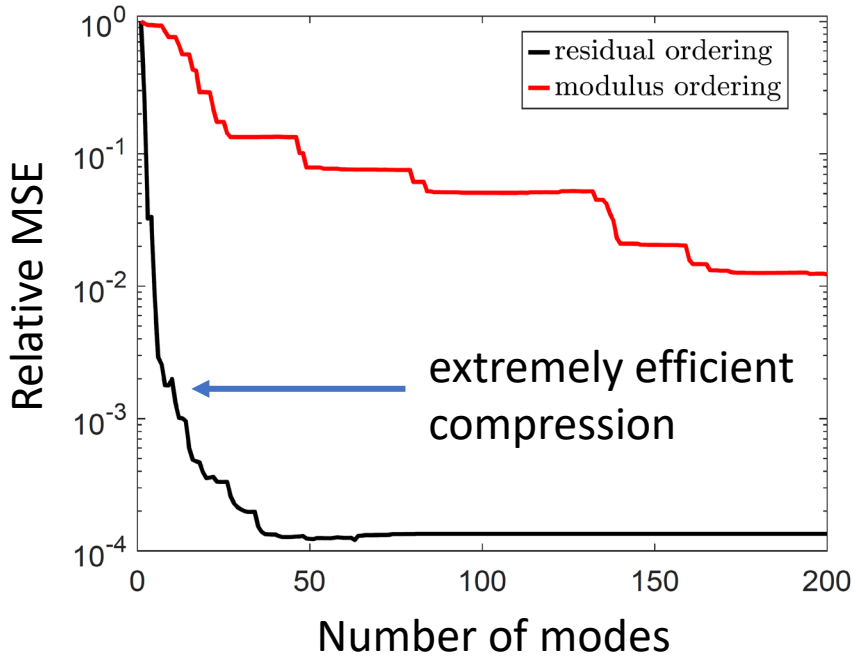


c) $t = 15\ \mu s$

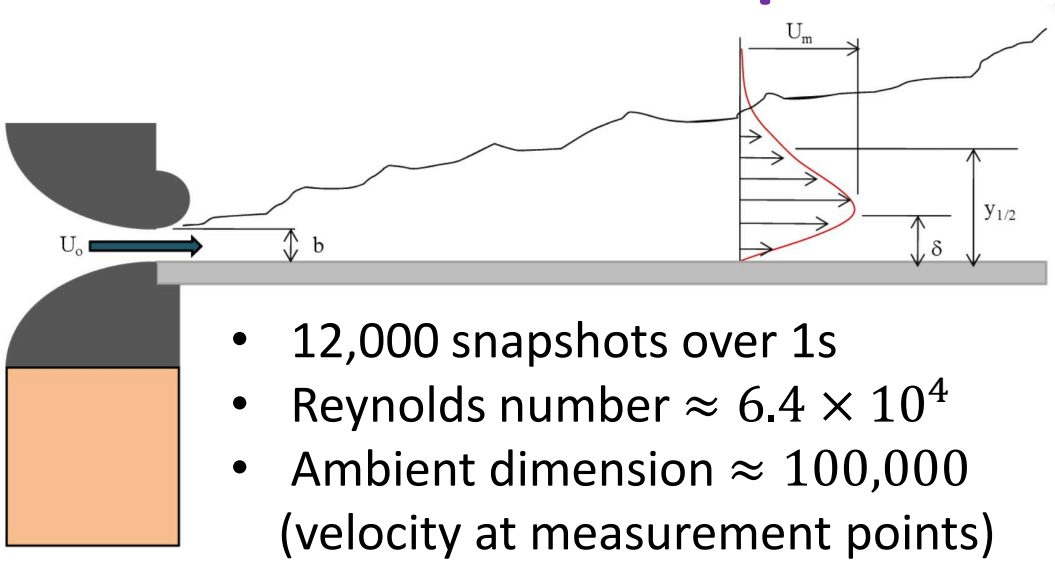
- 60 realisations ($M = 6600$)
- Ambient dimension ≈ 10 (length of initial window*)

*Raw measurements provided by Máté Szőke (Virginia Tech)



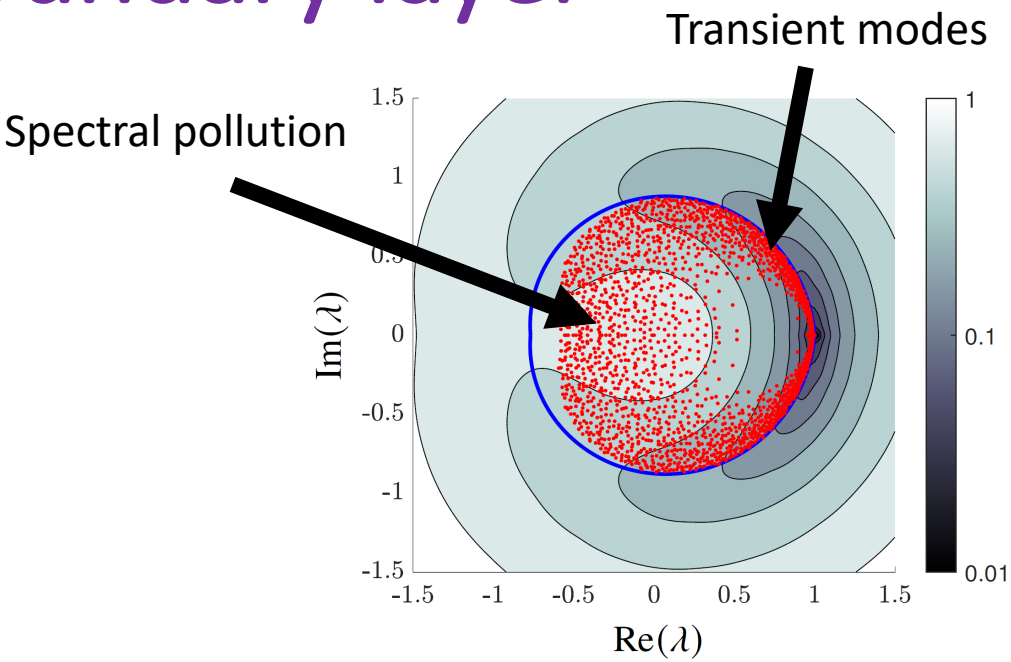


Example: wall-jet boundary layer

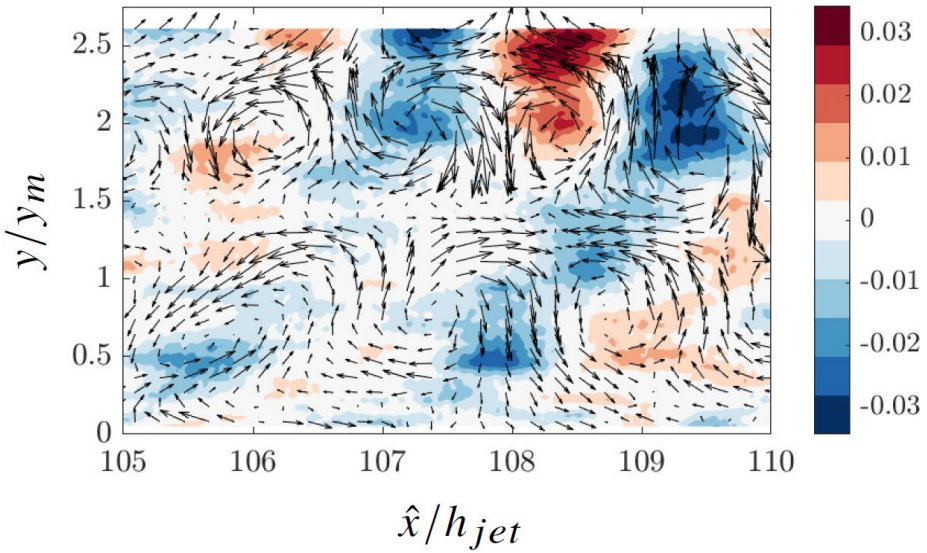


- 12,000 snapshots over 1s
- Reynolds number $\approx 6.4 \times 10^4$
- Ambient dimension $\approx 100{,}000$ (velocity at measurement points)

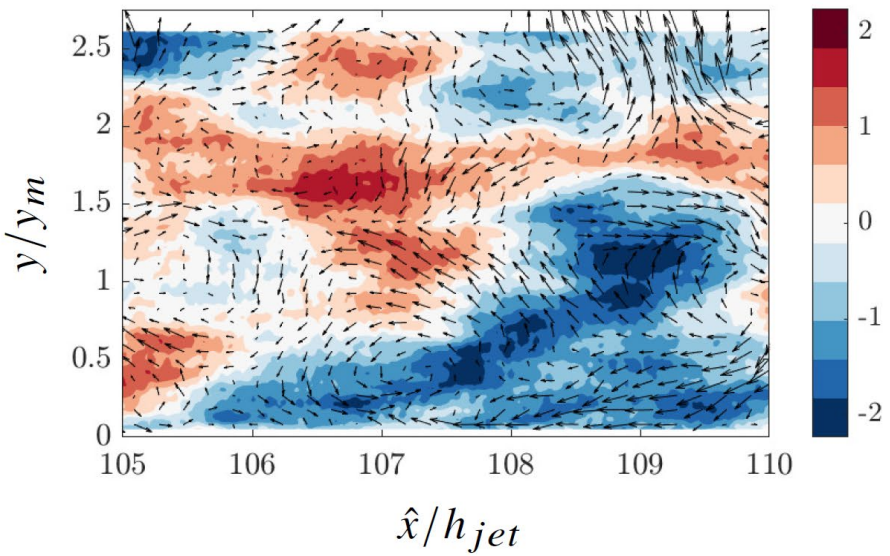
*Raw measurements provided by Máté Szőke (Virginia Tech)



$\lambda = 0.9439 + 0.2458i$, error ≤ 0.0765



$\lambda = 0.8948 + 0.1065i$, error ≤ 0.1105



Kernel method

Algorithm 4 A computational framework for kernelized versions of Algorithms 1 to 3.

Input: Snapshot data $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\}_{m=1}^{M'}$ and $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$, positive-definite kernel function $\mathcal{S} : \Omega \times \Omega \to \mathbb{R}$, and positive integer $N_K'' \leq M'$.

1: Apply kernel EDMD to $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)}\}_{m=1}^{M'}$ with kernel $\mathcal{S}$ to compute the matrices $\sqrt{W}\Psi_X\Psi_X^*\sqrt{W}$ and $\sqrt{W}\Psi_Y\Psi_X^*\sqrt{W}$ using the kernel trick.

2: Compute U and Σ from the eigendecomposition $\sqrt{W}\Psi_X\Psi_X^*\sqrt{W} = U\Sigma^2U^*$.

3: Compute the dominant N_K'' eigenvectors of $\widetilde{K}_{\text{EDMD}} = (\Sigma^\dagger U^*)\sqrt{W}\Psi_Y\Psi_X^*\sqrt{W}(U\Sigma^\dagger)$ and stack them column-by-column into $Z \in \mathbb{C}^{M' \times N_K''}$.

4: Apply a QR decomposition to orthogonalize Z to $Q = \begin{bmatrix} Q_1 & \cdots & Q_{N_K''} \end{bmatrix} \in \mathbb{C}^{M' \times N_K''}$.

5: Apply Algorithms 1 to 3 with $\{\hat{\boldsymbol{x}}^{(m)}, \hat{\boldsymbol{y}}^{(m)}\}_{m=1}^{M''}$ and the dictionary $\{\psi_j\}_{j=1}^{N_K''}$, where

$$\psi_j(\boldsymbol{x}) = \begin{bmatrix} \mathcal{S}(\boldsymbol{x}, \boldsymbol{x}^{(1)}) & \mathcal{S}(\boldsymbol{x}, \boldsymbol{x}^{(2)}) & \cdots & \mathcal{S}(\boldsymbol{x}, \boldsymbol{x}^{(M')}) \end{bmatrix} (U\Sigma^+)Q_j, \qquad 1 \leq j \leq N_K''.$$

Output: Spectral properties of Koopman operator according to Algorithms 1 to 3.