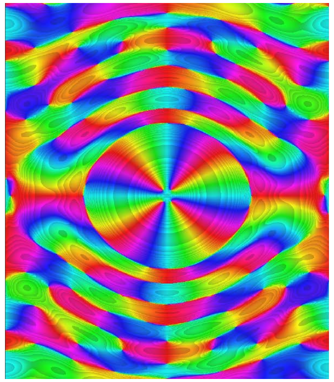
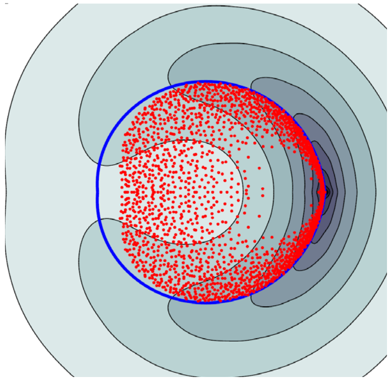
# Rigorous and data driven Koopmanism:

## Infinite-dimensional spectral computations for nonlinear systems

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

University of Cambridge + École Normale Supérieure





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

**Applications:** C., Ayton, Szőke, "*Residual Dynamic Mode Decomposition: Robust and verified Koopmanism*"

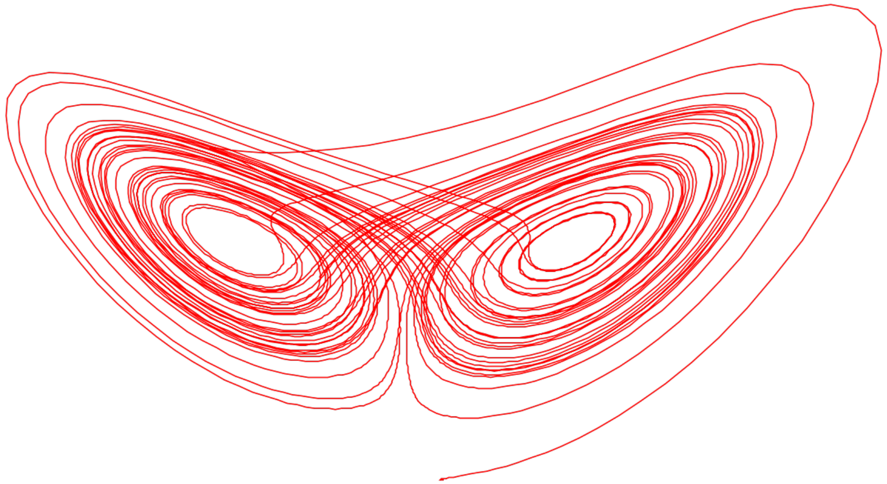
http://www.damtp.cam.ac.uk/user/mjc249/home.html: slides, papers, and code

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



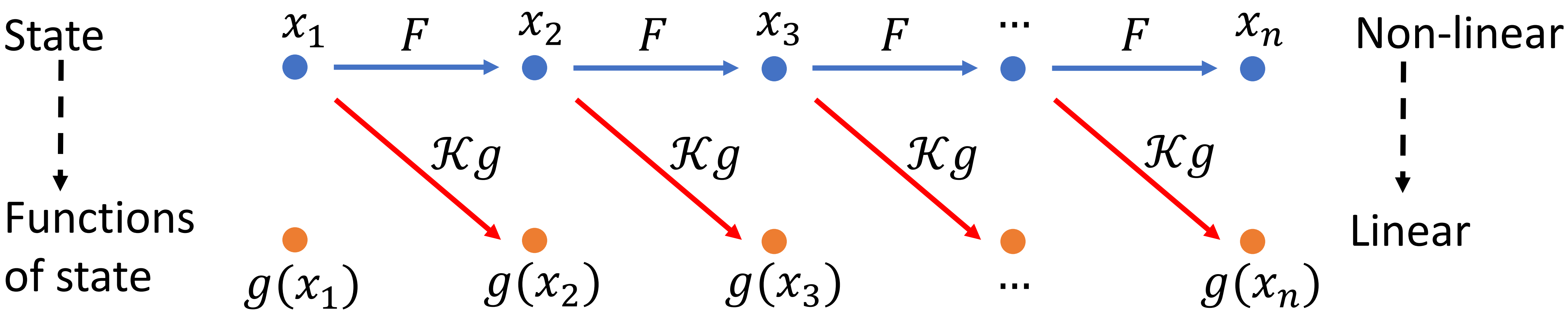
**Can we develop verified methods?**

# Operator viewpoint

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

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

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




- Work in $L^2(\Omega, \omega)$ for positive measure $\omega$, 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.

# Koopman mode decomposition

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

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

$$g(x) = \sum_{\text{eigenvalues } \lambda_j} c_{\lambda_j} \underbrace{\varphi_{\lambda_j}}_{\text{eigenfunction of } \mathcal{K}}(x) + \int_{-\pi}^{\pi} \underbrace{\phi_{\theta,g}}_{\text{generalised eigenfunction of } \mathcal{K}}(x)\, \mathrm{d}\theta$$

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

**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

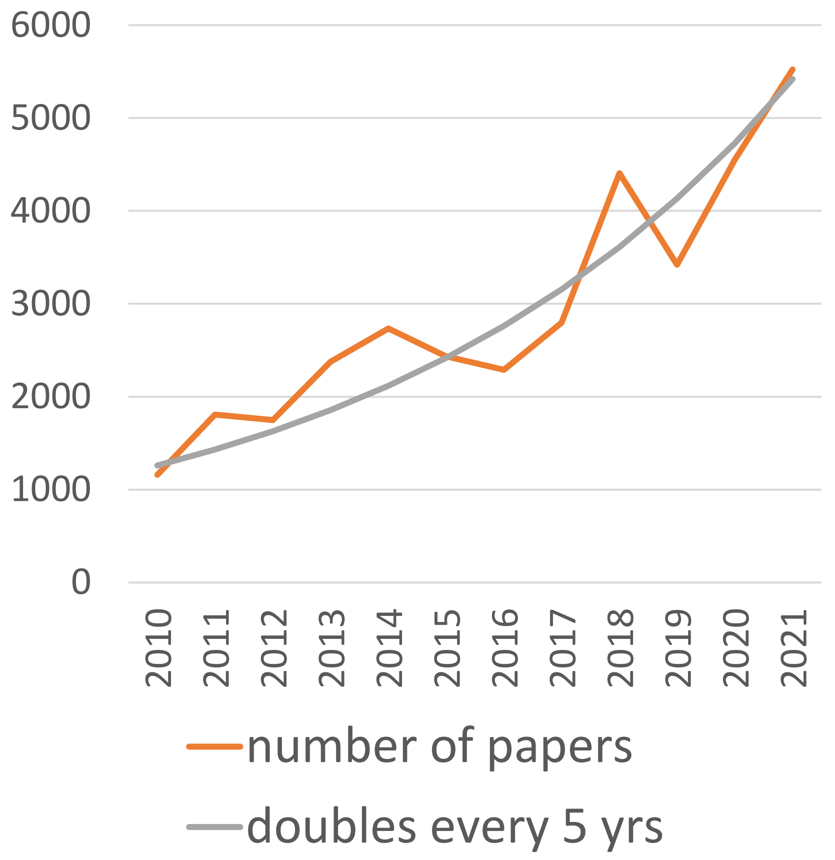
≈35,000 papers over last decade!

***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 …

**New Papers on "Koopman Operators"**




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

# Can we compute spectral properties in inf. dim.?

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

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

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

*"Most operators that arise in practice are not **diagonalized**, and it is often very hard to locate even a single point in the spectrum. Thus, one has to settle for numerical approximations. Unfortunately, there is a dearth of literature on this basic problem and **no proven <u>general</u> techniques**."* W. Arveson, Berkeley (1994)

# Four key challenges

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

1) **"Too much":** Approximate spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$ - "spectral pollution"
2) **"Too little":** Miss parts of $\mathrm{Spec}(\mathcal{K})$
3) **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.

# Build the matrix: Dynamic Mode Decomposition (DMD)

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

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

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

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

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

**Recall open problems: 1) "too much", 2) "too little", 3) 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 $\mathcal{K}^*\mathcal{K}$

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

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

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

**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

# ResDMD: avoiding "too much"

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

Algorithm 1:

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N_K \times N_K}$ and eigendecomposition $K_1 V = GV\Lambda$.
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 "too much"

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

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: avoiding "too little"

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

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 $\tau_k = \min_{g=\sum_{j=1}^{N_K} \mathbf{g}_j \psi_j} \mathrm{res}(z_k, g)$, corresponding $g_k$ (gen. SVD).
3. **Output:** $\{z_k : \tau_k < \varepsilon\}$ (approx. of $\mathrm{Spec}_\varepsilon(\mathcal{K})$), $\{g_k : \tau_k < \varepsilon\}$ ($\varepsilon$-pseudo-eigenfunctions).

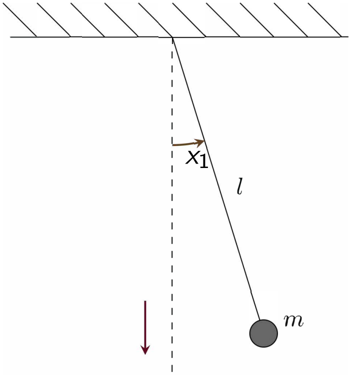
**Theorem (full convergence):** Suppose the quadrature rule converges.

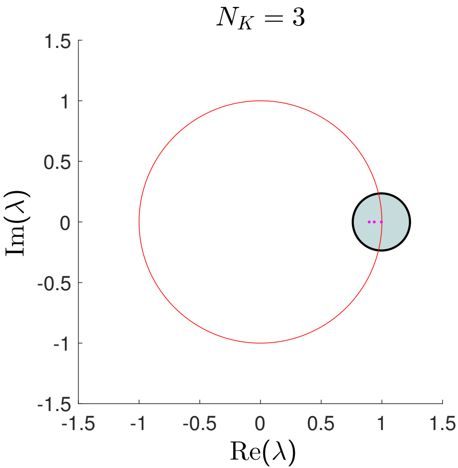
- **Error control:** $\{z_k : \tau_k < \varepsilon\} \subseteq \mathrm{Spec}_\varepsilon(\mathcal{K})$ (as $M \to \infty$)
- **Convergence:** Converges locally uniformly to $\mathrm{Spec}_\varepsilon(\mathcal{K})$ (as $N_K \to \infty$)

**NB:** Local optimisation strategy shrinks $\varepsilon$ 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).

# 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 $\mathcal{L}: \mathcal{D}(\mathcal{L}) \to \mathcal{H}$. 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.

# Koopman mode decomposition (again!)

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

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

eigenfunction of $\mathcal{K}$

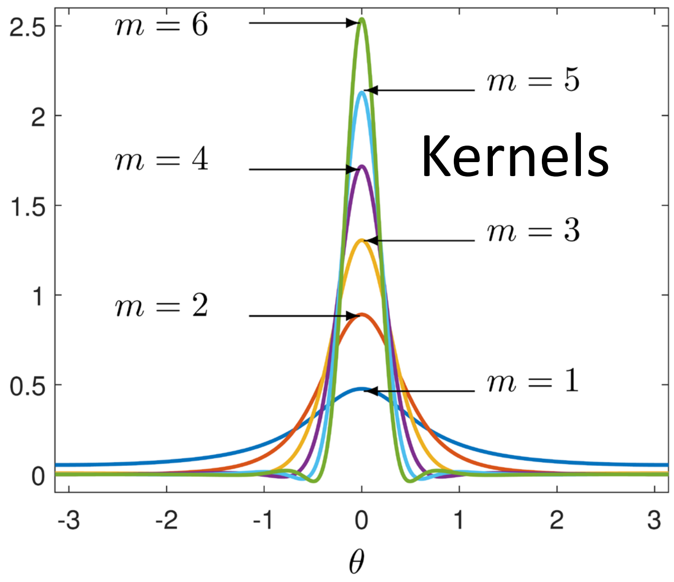
generalised eigenfunction of $\mathcal{K}$

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

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

**Computing $\nu_g$ diagonalises non-linear dynamical system!**

# $m$th order Plemelj formula



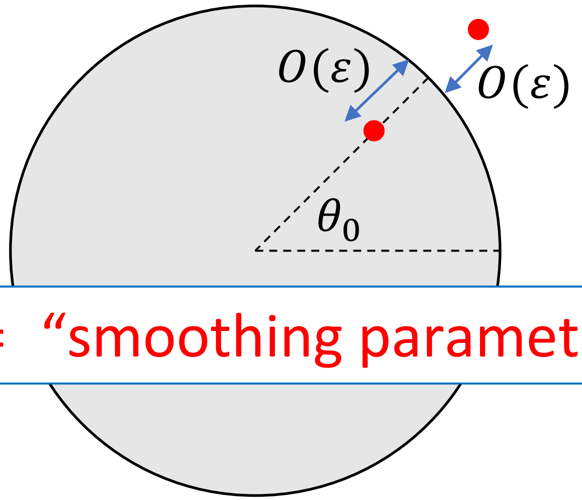

$$\mathcal{C}_g(z) = \int_{-\pi}^{\pi} \frac{e^{i\theta}\,\mathrm{d}\nu_g(\theta)}{e^{i\theta}-z} = \begin{cases} \langle (\mathcal{K}-zI)^{-1}g, \mathcal{K}^*g\rangle, \text{if } |z|>1 \\ -z^{-1}\langle g, (\mathcal{K}-\bar{z}^{-1}I)^{-1}g\,\rangle, \text{if } 0<|z|<1 \end{cases}$$

ResDMD computes with error control

$m$th order rational kernels

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

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




$\varepsilon =$ “smoothing parameter”

$O(PN_K)$ cost for evaluation at $P$ values of $\theta$

# Convergence

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

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

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

- Also recover discrete spectrum.

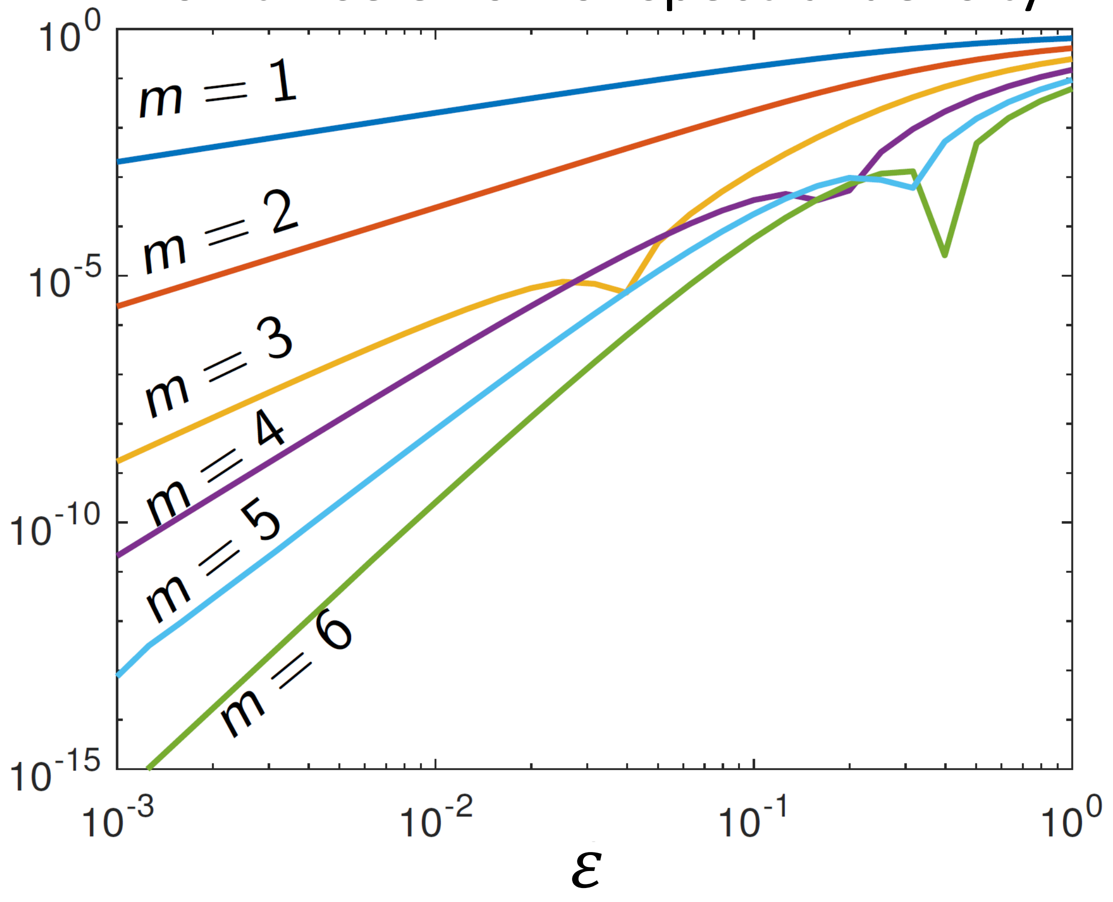
# Example

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

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

Generalised shift, typical building block of many dynamical systems.

Pointwise error for spectral density




**NB:** Small $N_K$ critical in data-driven computations.

# 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.

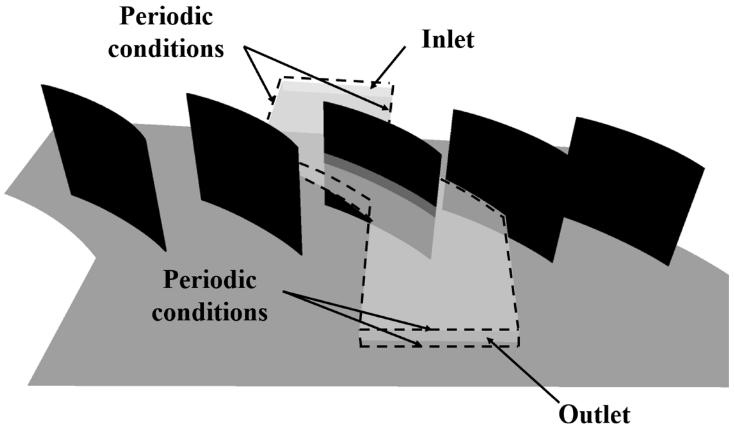
**Q: Is discretisation span$\{\psi_1, \dots, \psi_{N_K}\}$ large/rich enough?**

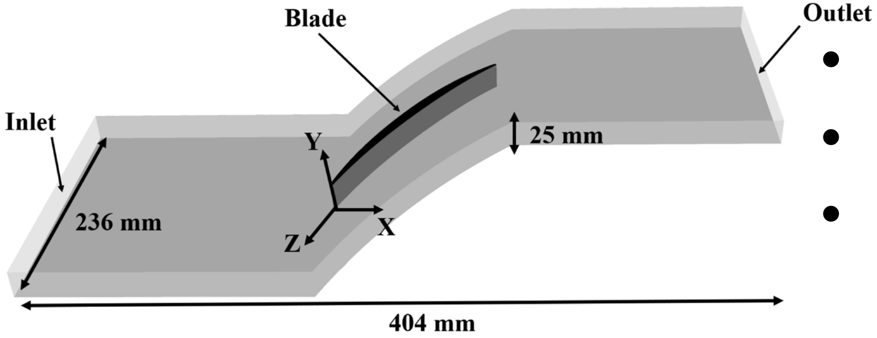
***Above algorithms:***

- 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

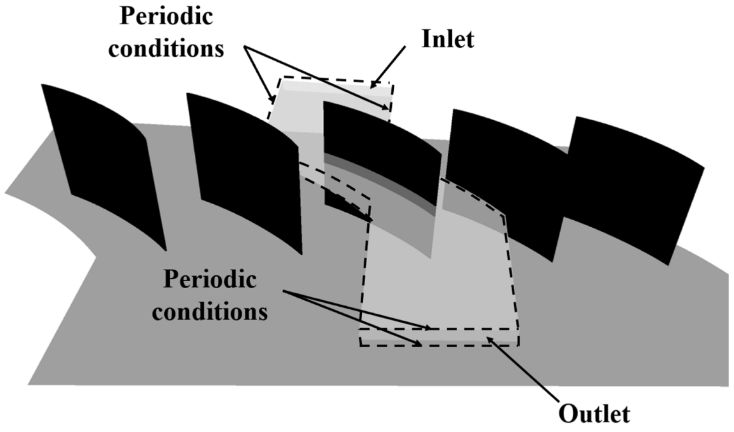


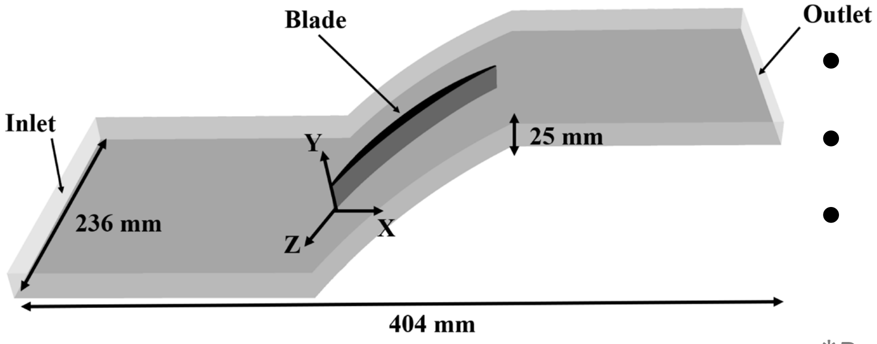




- Data collected for $2 \times 10^{-4}$s
- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension $\approx 300{,}000$ (number of measurement points*)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

# Example: pressure field of turbulent flow

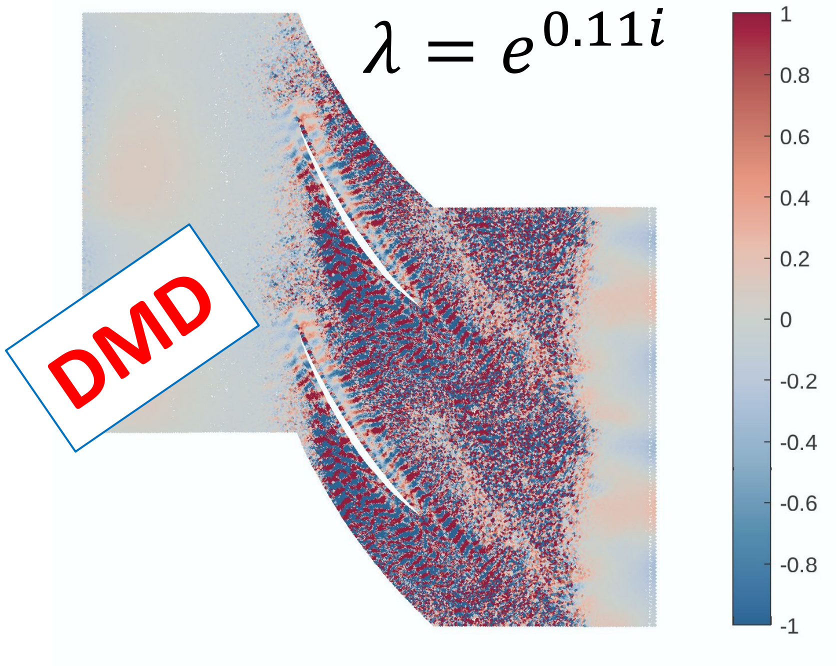


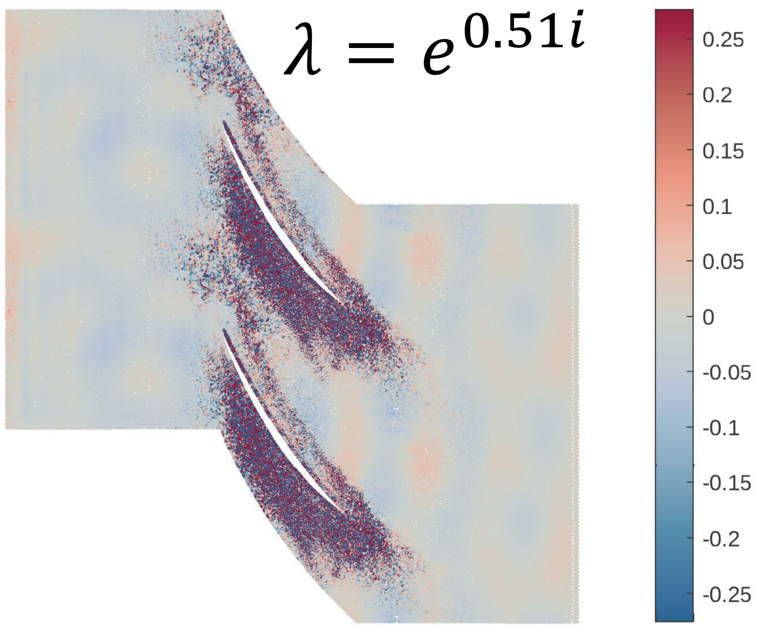

- Data collected for $2 \times 10^{-4}$s
- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension $\approx 300{,}000$ (number of measurement points*)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

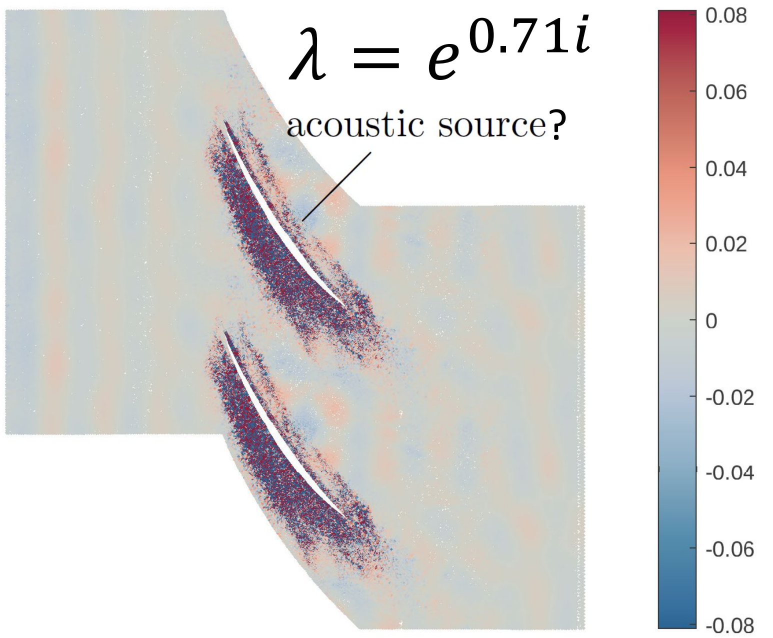
Rel. Error = ?



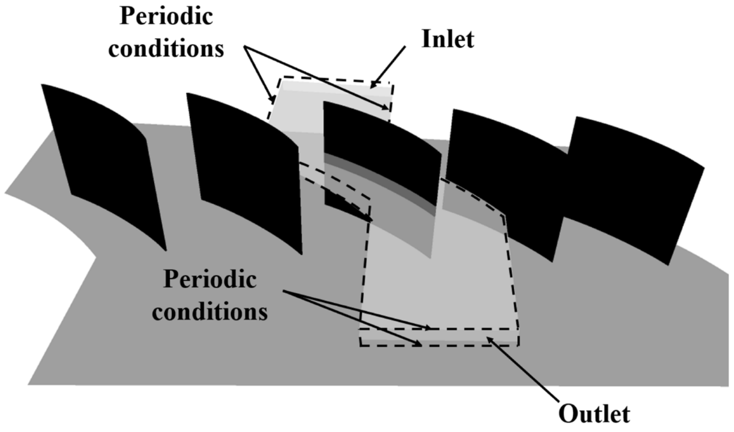

Rel. Error = ?

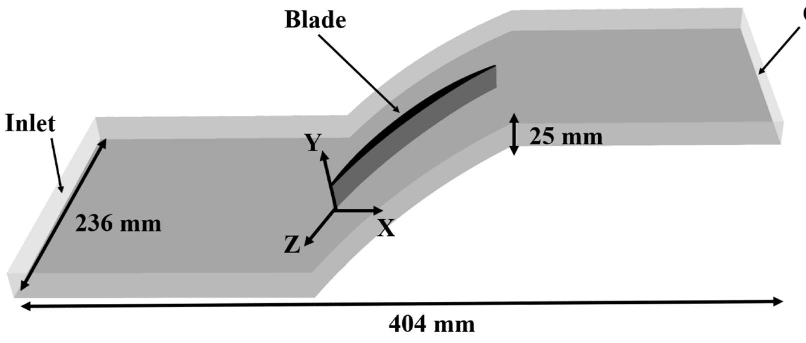



Rel. Error = ?



# Example: pressure field of turbulent flow

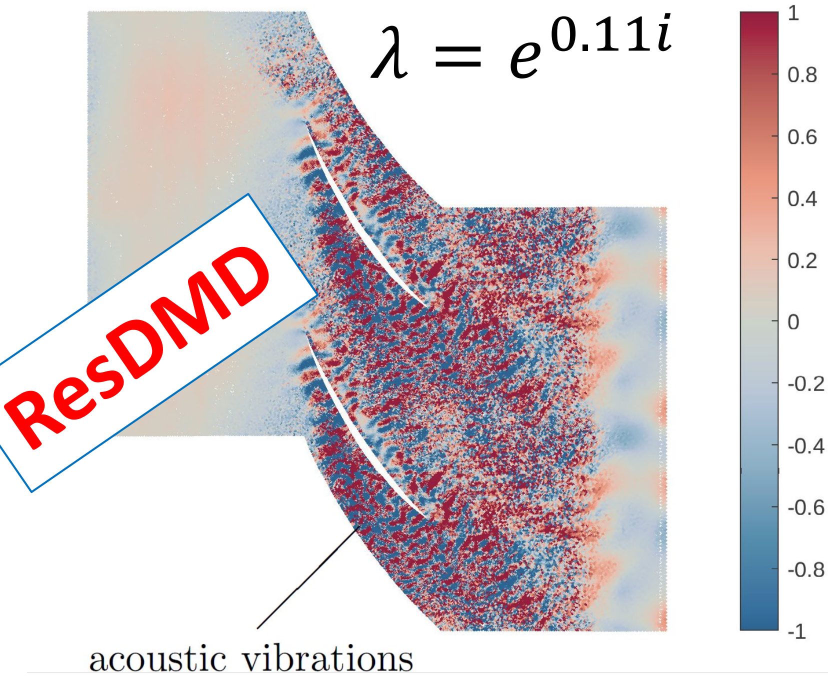


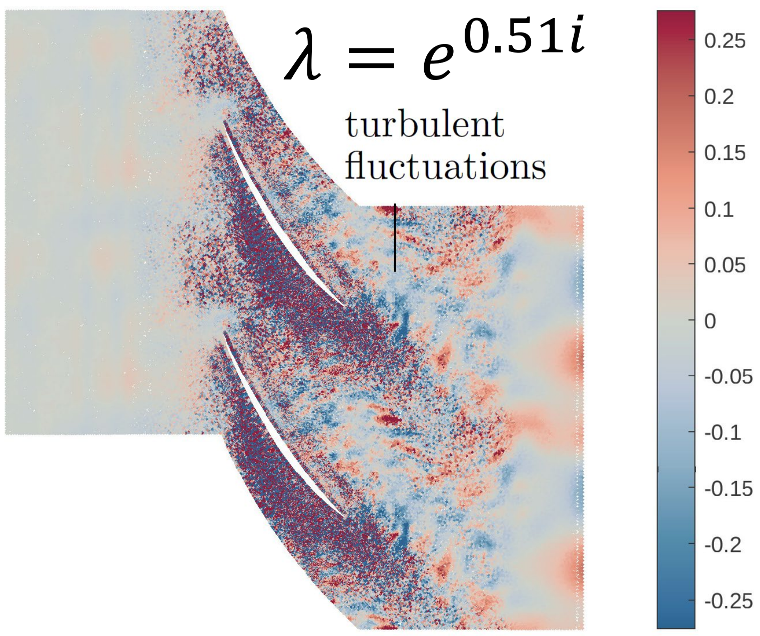

- Data collected for $2 \times 10^{-4}$s
- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension $\approx$ 300,000 (number of measurement points*)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

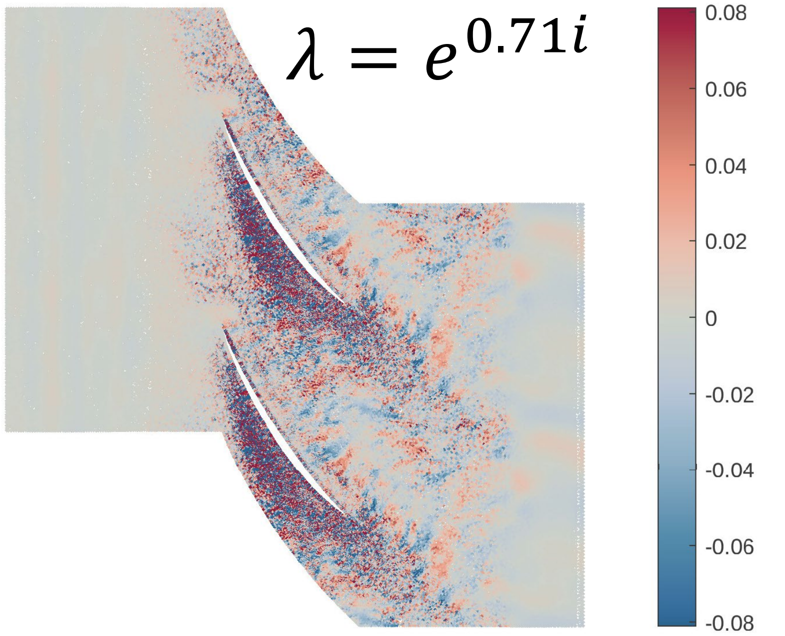
Rel. Error $\leq$ 0.0054



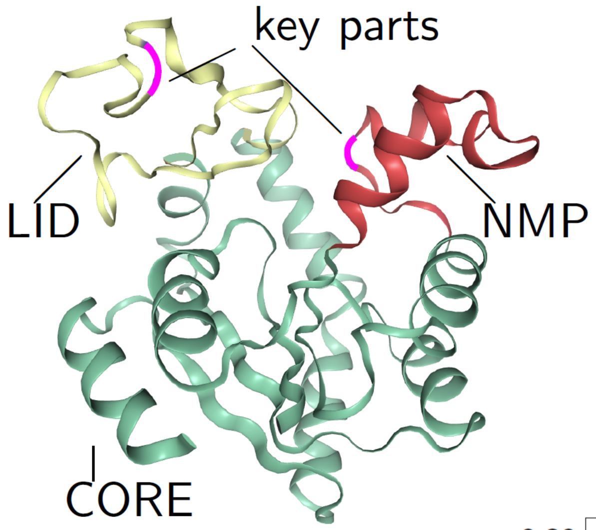

Rel. Error $\leq$ 0.0128

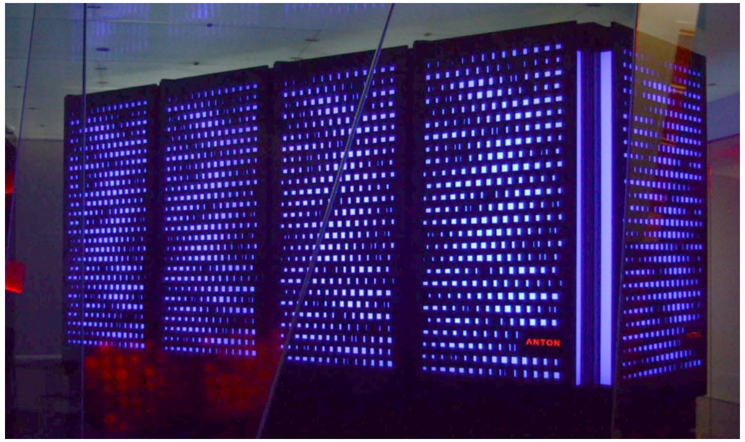



Rel. Error $\leq$ 0.0196



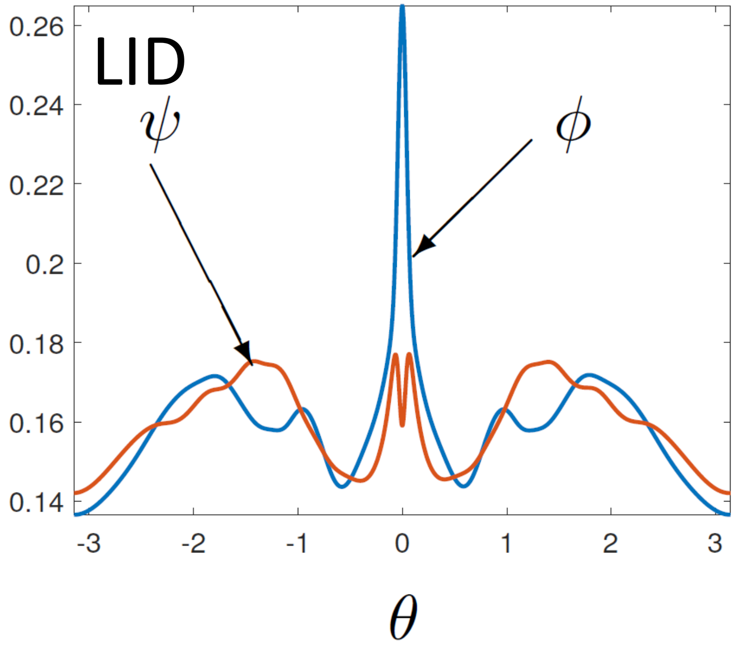
# Example: molecular dynamics (Adenylate Kinase)

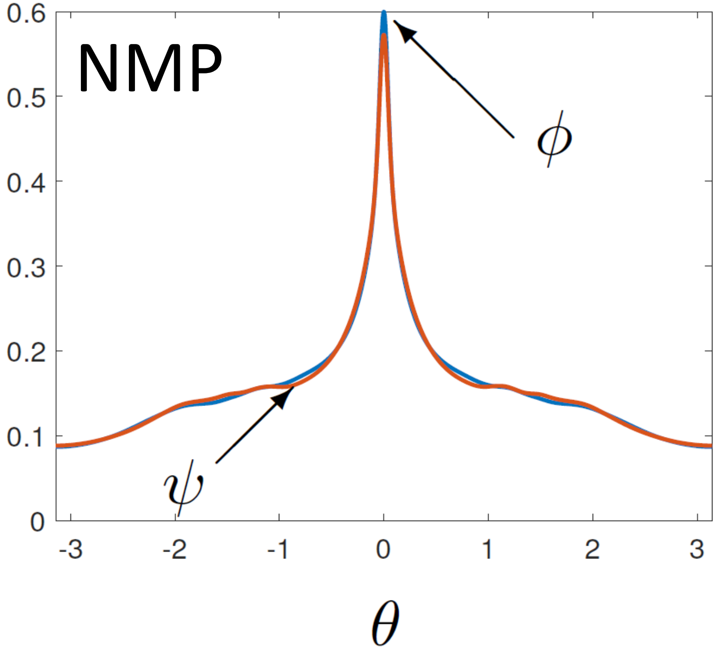


- All-atom equilibrium simulation for $1.004 \times 10^{-6}$s
- Ambient dimension $\approx 20{,}000$ (positions and momenta of atoms)
- 6th order kernel (spec res $10^{-6}$)

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



# Wider programme: a toolkit

- Infinite-dimensional NLA $\Longrightarrow$ **Compute spectral properties for the first time.**
- Solvability Complexity Index hierarchy $\Longrightarrow$ **Algorithms realise the boundaries of what's possible.**
- Builds on and extends work of **Turing**, **Smale**, and **McMullen.**
- **Extends to:** Foundations of AI, PDEs (e.g., time-dep. Schrödinger eq. on $L^2(\mathbb{R}^d)$ with error control), optimisation (e.g., guarantees), computer-assisted proofs, …


- C., "*On the computation of geometric features of spectra of linear operators on Hilbert spaces*," **Found. Comput. Math.**, under revisions.
- C., "*Computing spectral measures and spectral types*," **Communications in Mathematical Physics**, 2021.
- C., Horning, Townsend "*Computing spectral measures of self-adjoint operators*," **SIAM Review**, 2021.
- C., Roman, Hansen, "*How to compute spectra with error control*," **Physical Review Letters**, 2019.
- C., Hansen, "*The foundations of spectral computations via the solvability complexity index hierarchy*," **JEMS**, under revisions.
- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem*," **Proceedings of the National Academy of Sciences**, 2022.
- C., "*Computing semigroups with error control*," **SIAM Journal on Numerical Analysis**, 2022.
- Software package (MATLAB): https://github.com/SpecSolve for PDEs, integral operators, infinite matrices.
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, "*On the solvability complexity index hierarchy and towers of algorithms*," arXiv, 2020.
- Smale, "*The fundamental theorem of algebra and complexity theory*," **Bulletin of the AMS**, 1981, *36 pp.*
- McMullen, "*Families of rational maps and iterative root-finding algorithms*," **Annals of Mathematics**, 1987, *27 pp.*

# Sample of classification theorems

Increasing difficulty

**Error control**

Spectra of compact operators

Spectral stability

$$\begin{array}{ccccccccccccc}
\Pi_0 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 \\
\| & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \ddots \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 \dots \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \iddots \\
\Sigma_0 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3
\end{array}$$

Time-dependent linear PDEs, e.g.,

$$i\frac{\partial \psi}{\partial t} = \left[-\nabla^2 + V(\mathbf{x})\right]\psi$$

Spectra of $\mathcal{K}$

Continuous spectra of $\mathcal{K}$
(different regularity assumptions)

Spectral gap problem

# Summary

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

- Spectra, pseudospectra, residuals of general Koopman operators (error control).
  - **Idea:** New matrix for residual $\Rightarrow$ ResDMD.
- Spectral measures of measure-preserving systems with high-order 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!**

**$\longrightarrow$ Opens the door to rigorous data-driven Koopmania!**

# Convergence of quadrature

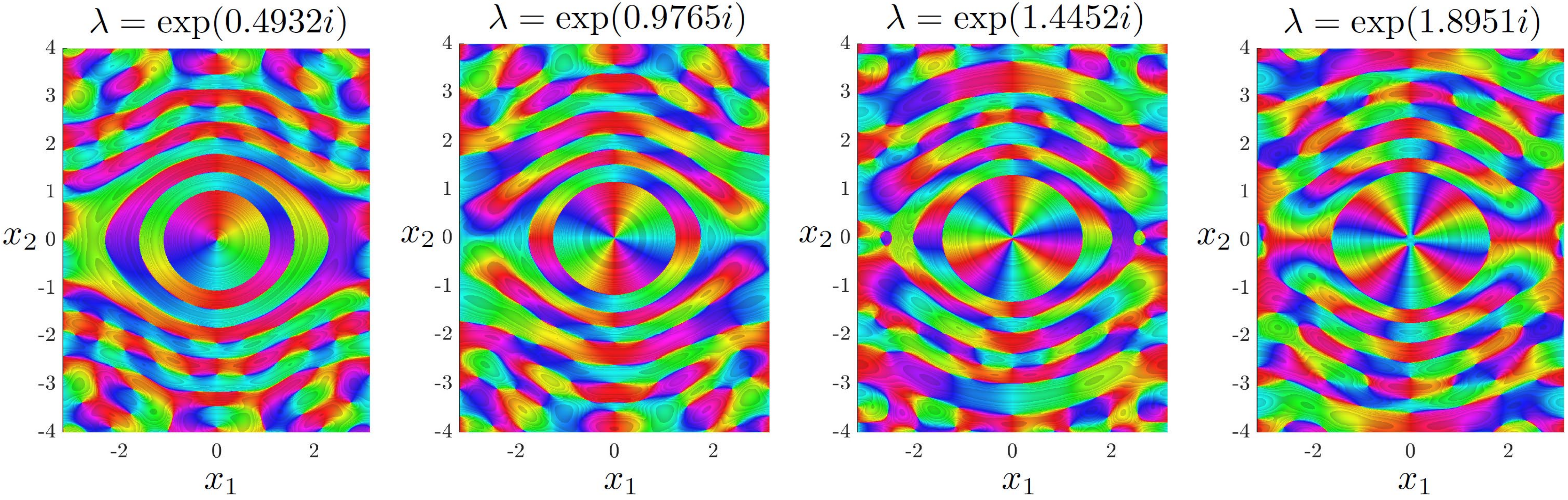
$$\text{E.g.,} \quad \langle \mathcal{K}\psi_k, \psi_j \rangle = \lim_{M\to\infty} \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})}$$

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

# 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 $N_K$).

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

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

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

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

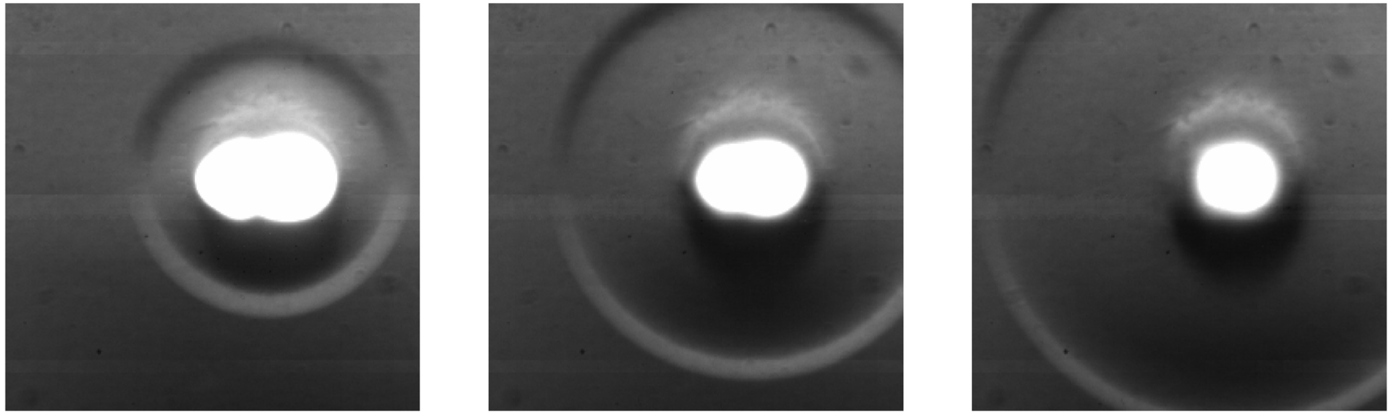
Residual Koopman mode decomposition (order modes by $\mathrm{res}(\lambda, \mathbf{v})$):

Controllable error

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

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

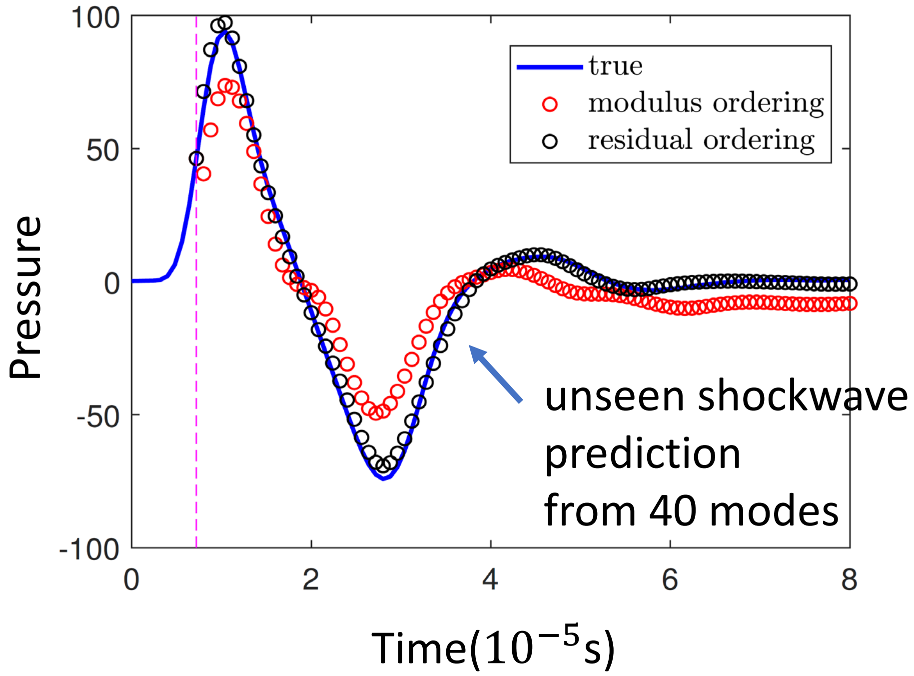
# Example: laser-induced plasma

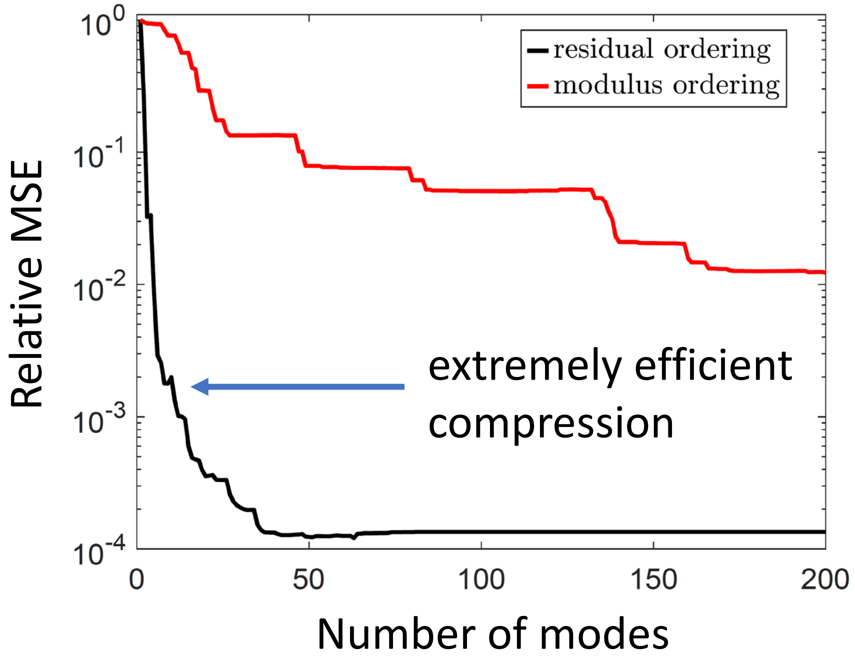


a) $t = 5\ \mu s$ b) $t = 10\ \mu s$ c) $t = 15\ \mu s$

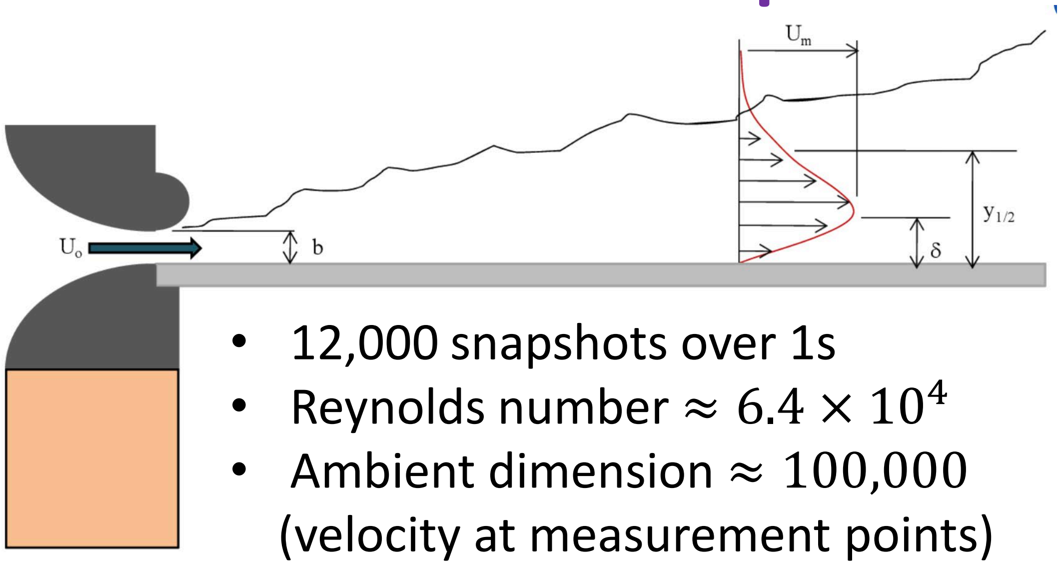
- 60 realisations ($M = 6600$)
- Ambient dimension ≈ 10 (length of initial window*)

*Raw measurements provided by Máté Szőke (Virginia Tech)



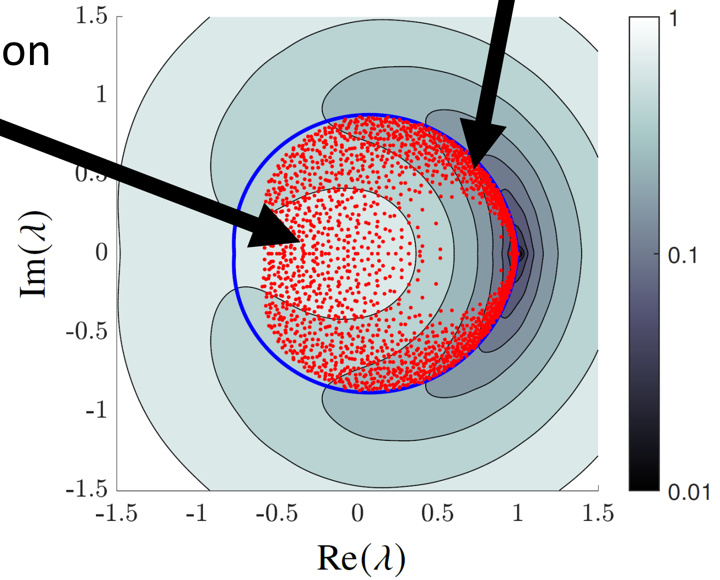



# Example: wall-jet boundary layer

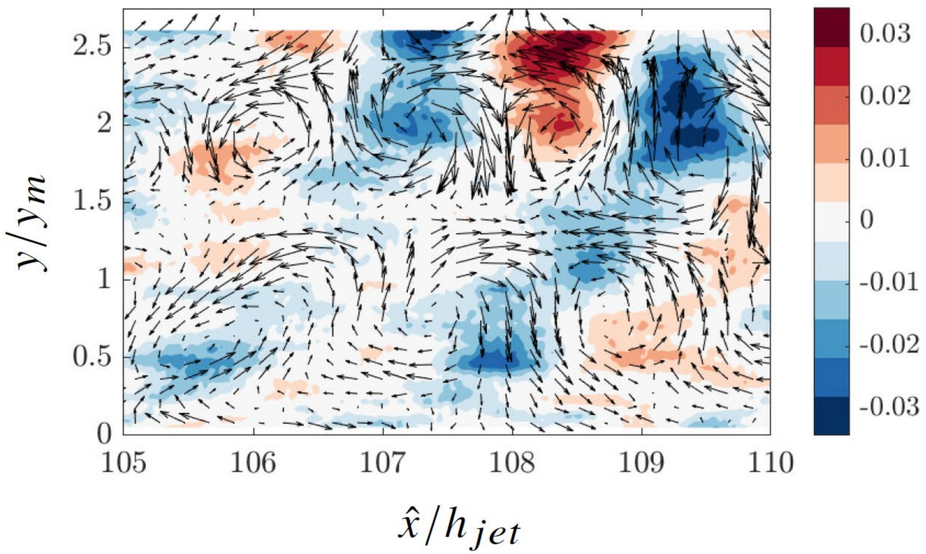



- 12,000 snapshots over 1s
- Reynolds number $\approx 6.4 \times 10^4$
- Ambient dimension $\approx 100{,}000$ (velocity at measurement points)

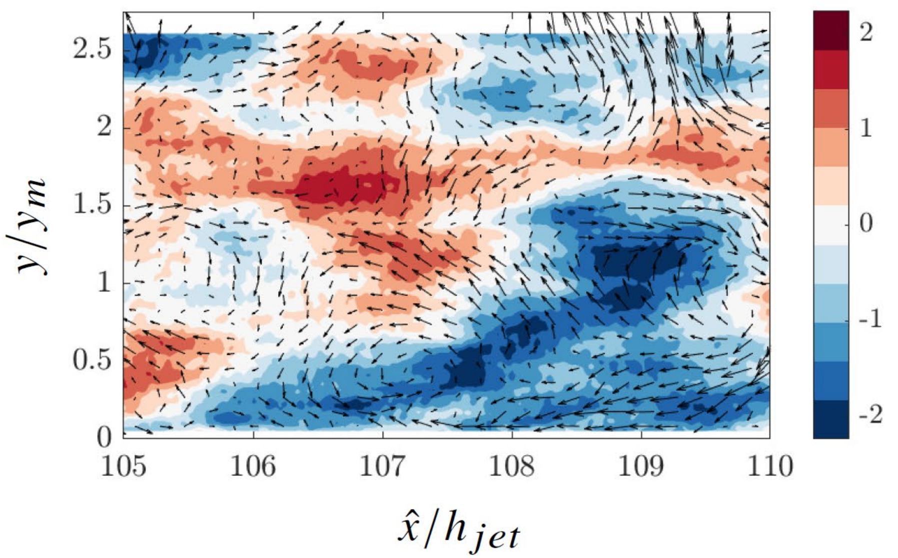
*Raw measurements provided by Máté Szőke (Virginia Tech)




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



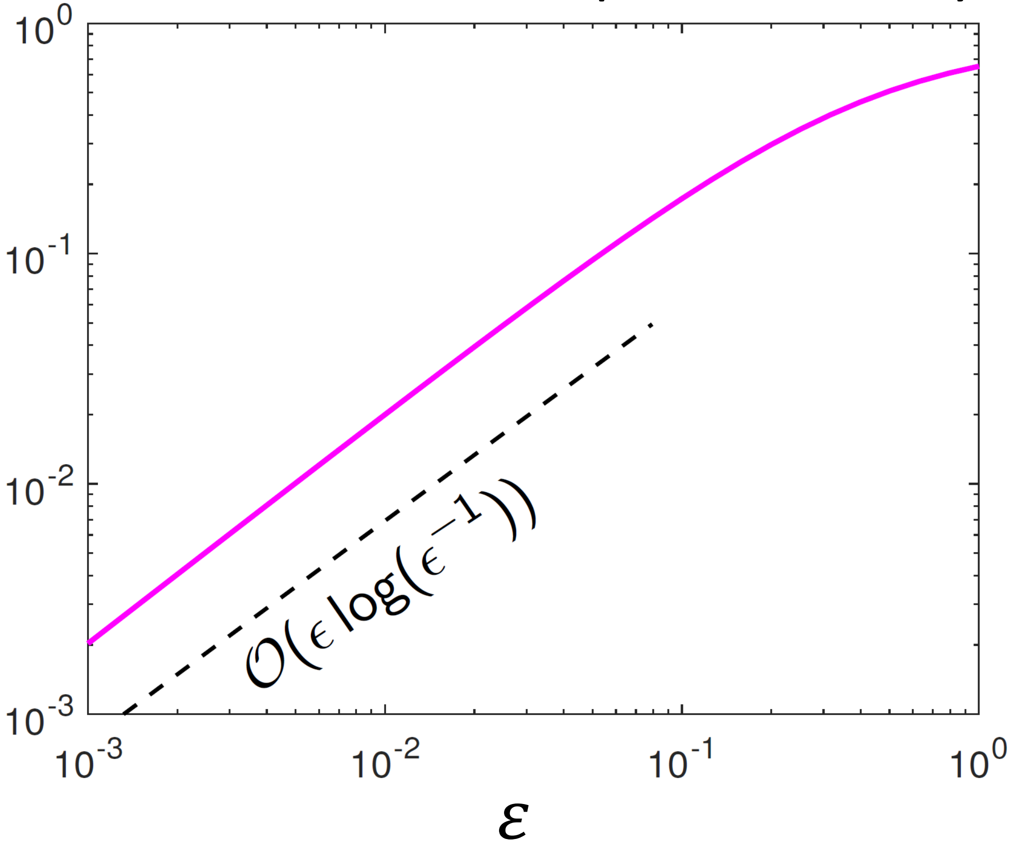

$\lambda = 0.8948 + 0.1065i$, error $\leq 0.1105$



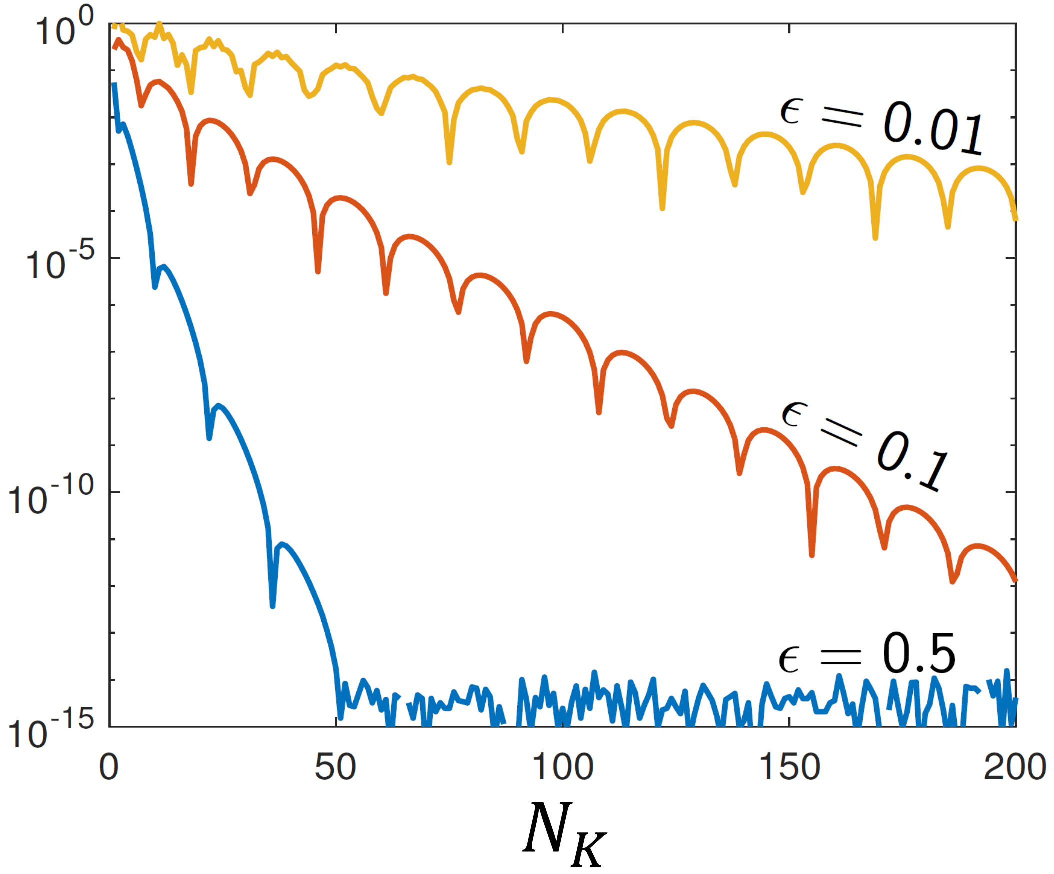
# But … slow convergence

**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 data-driven computations. Can we improve convergence rate?

# 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 $\Sigma$ from the eigendecomposition $\sqrt{W}\Psi_X\Psi_X^*\sqrt{W} = U\Sigma^2U^*$.

3: Compute the dominant $N''_K$ eigenvectors of $\widetilde{K}_{\mathrm{EDMD}} = (\Sigma^\dagger U^*)\sqrt{W}\Psi_Y\Psi_X^*\sqrt{W}(U\Sigma^\dagger)$ 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.