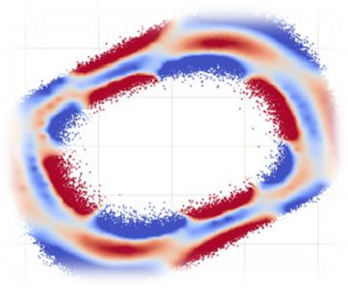
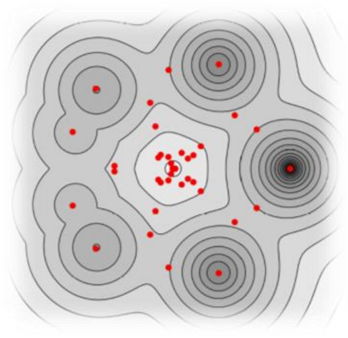



Verifying data-driven computations of Koopman spectra

Matthew Colbrook
University of Cambridge
19/09/2023

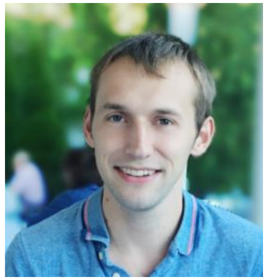
Joint work with:


Lorna Ayton


Qin Li


Ryan Raut


Matt Szőke


Alex Townsend

- C., Townsend, *“Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems”* Commun. Pure Appl. Math., 2023.
- C., Ayton, Szőke, “*Residual Dynamic Mode Decomposition,”* J. Fluid Mech.**,** 2023.
- C., Li, Raut, Townsend, *“Beyond expectations: Residual Dynamic Mode Decomposition and Variance for Stochastic Dynamical Systems,” arxiv.*

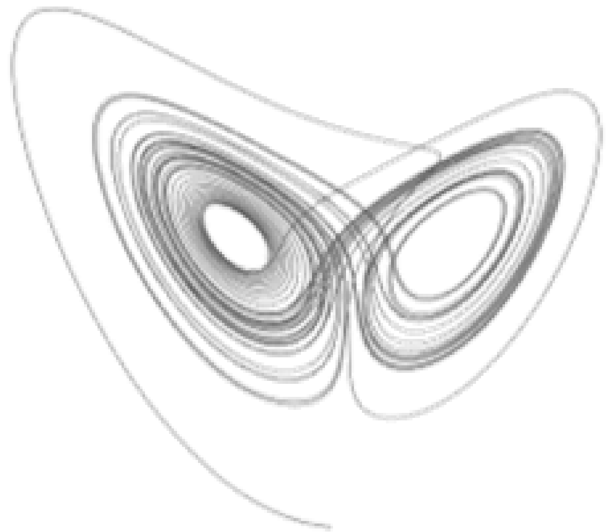
Data-driven (deterministic) dynamical systems

State $x \in \Omega \subseteq \mathbb{R}^d$.

Unknown function $F: \Omega \to \Omega$ governs dynamics: $x_{n+1} = F(x_n)$

Goal: Verified learning from data $\{x^{(m)}, y^{(m)} = F(x^{(m)})\}_{m=1}^{M}$.

Applications: chemistry, climatology, control, electronics, epidemiology, finance, fluids, molecular dynamics, neuroscience, plasmas, robotics, video processing, etc.



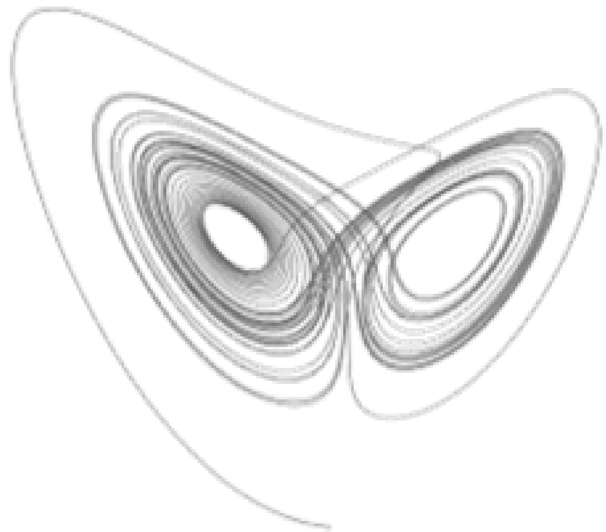
Data-driven (deterministic) dynamical systems

State $x \in \Omega \subseteq \mathbb{R}^d$.

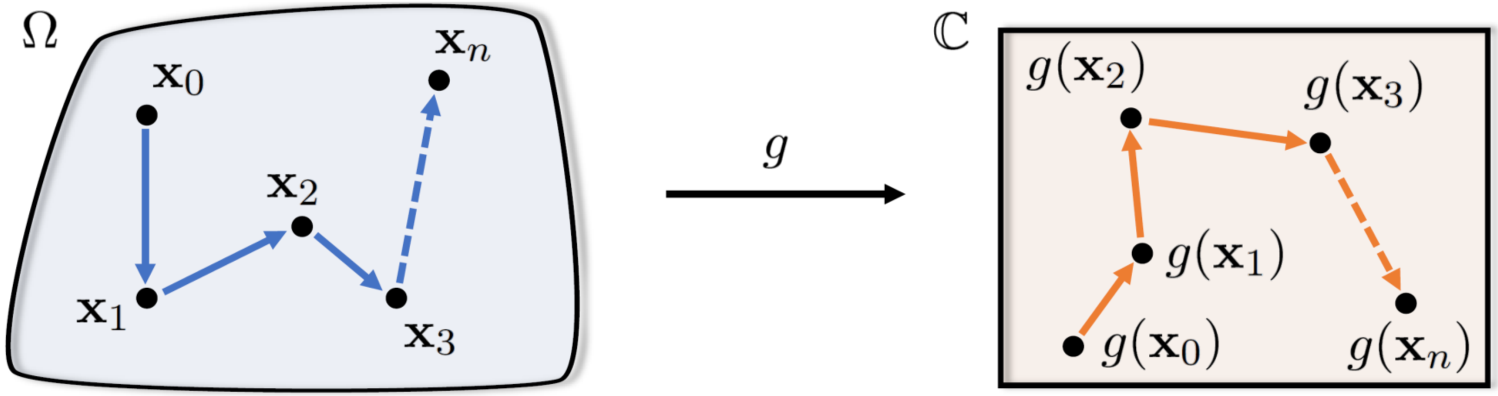
Unknown function $F: \Omega \to \Omega$ governs dynamics: $x_{n+1} = F(x_n)$

Goal: Verified learning from data $\{x^{(m)}, y^{(m)} = F(x^{(m)})\}_{m=1}^{M}$.

Applications: chemistry, climatology, control, electronics, epidemiology, finance, fluids, molecular dynamics, neuroscience, plasmas, robotics, video processing, etc.



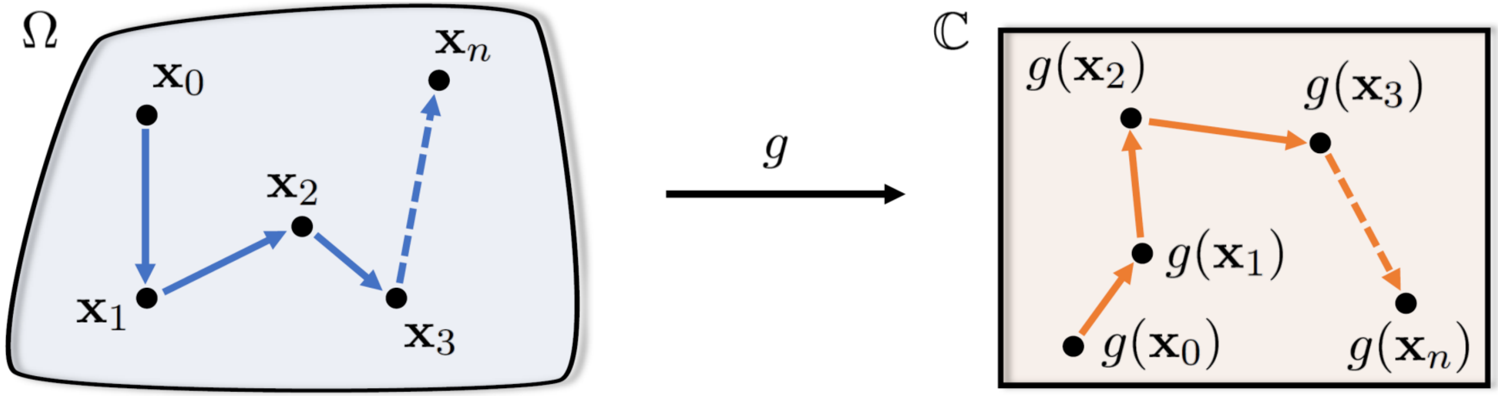
Koopman Operator $\mathcal{K}$: A global linearization



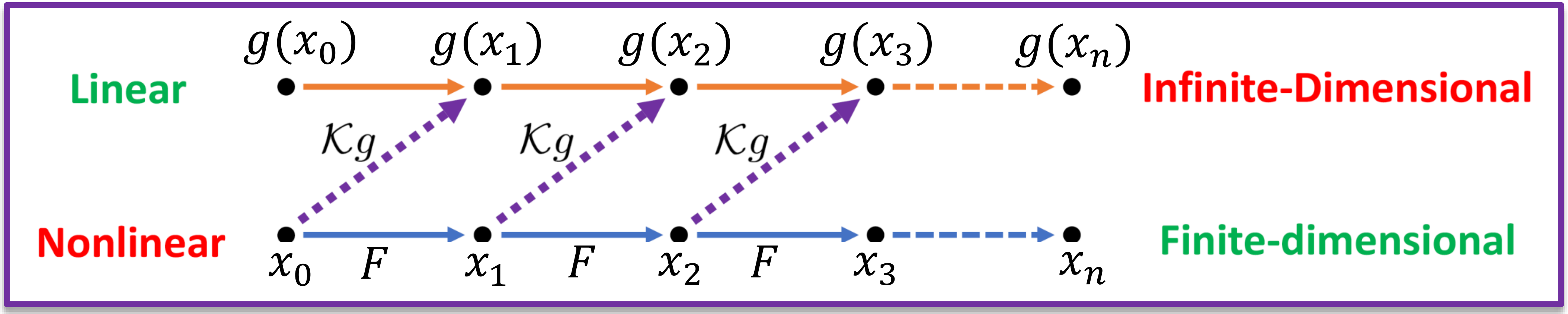
- $\mathcal{K}$ acts on functions $g: \Omega \to \mathbb{C}$, $[\mathcal{K}g](x) = g(F(x))$.
- Function space: $L^2(\Omega, \omega)$, positive measure ω, inner product $\langle \cdot, \cdot \rangle$.

- Koopman, "*Hamiltonian systems and transformation in Hilbert space,*" **Proc. Natl. Acad. Sci. USA**, 1931.
- Koopman, v. Neumann, "*Dynamical systems of continuous spectra,*" **Proc. Natl. Acad. Sci. USA**, 1932.

Koopman Operator $\mathcal{K}$: A global linearization



- $\mathcal{K}$ acts on functions $g: \Omega \to \mathbb{C}$, $[\mathcal{K}g](x) = g(F(x))$.
- Function space: $L^2(\Omega, \omega)$, positive measure ω, inner product $\langle \cdot, \cdot \rangle$.



- Koopman, *"Hamiltonian systems and transformation in Hilbert space,"* **Proc. Natl. Acad. Sci. USA**, 1931.
- Koopman, v. Neumann, *"Dynamical systems of continuous spectra,"* **Proc. Natl. Acad. Sci. USA**, 1932.

Example 1: Why is linear much easier?

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

Trivial dynamics!

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

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

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

Eigenfunction

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

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

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

Example 1: Why is linear much easier?

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

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

Trivial dynamics!

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

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

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

Eigenfunction

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

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

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

Example 2: Koopman mode decomposition

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

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

eigenfunction of $\mathcal{K}$

generalized eigenfunction of $\mathcal{K}$

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

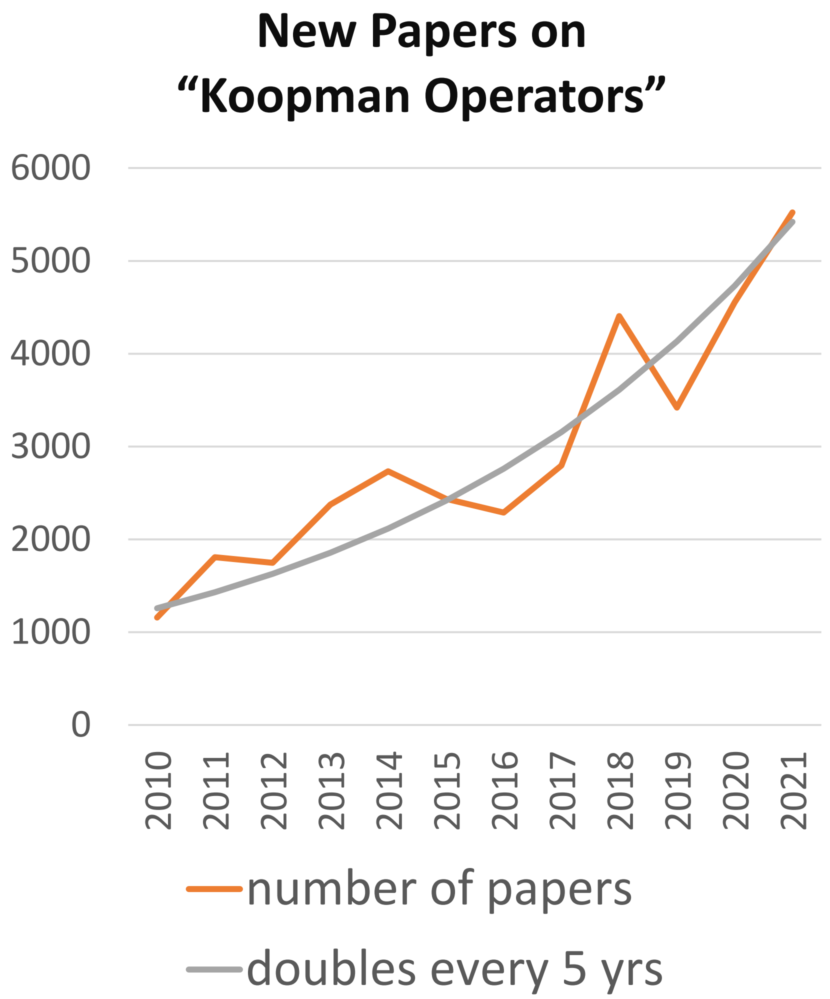
$$g(x_n) = [\mathcal{K}^n g](x_0) = \sum_{\text{eigenvalues } \lambda_j} c_{\lambda_j} {\lambda_j}^n \varphi_{\lambda_j}(x_0) + \int_{-\pi}^{\pi} e^{in\theta} \phi_{\theta,g}(x_0)\, \mathrm{d}\theta$$

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

Focus on computing spectral properties of $\mathcal{K}$ ...

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

Koopmania! ≈35,000 papers in last decade!



Few works on convergence guarantees or verification.

Why?

- Perhaps a different community with different focus?
- Dealing with infinite dimensions is notoriously hard …

Challenges of computing

$$\mathrm{Sp}(\mathcal{K}) = \{\lambda \in \mathbb{C} : \mathcal{K} - \lambda I \ \text{ is not invertible}\}$$

Truncate/discretize: $\mathcal{K} \longrightarrow \mathbb{K} \in \mathbb{C}^{N \times N}$

1) **Too much:** Spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$
2) **Too little:** Miss parts of $\mathrm{Spec}(\mathcal{K})$
3) **Continuous spectra**
4) **Verification**
5) **Instability**

• C., *"The foundations of infinite-dimensional spectral computations,"* **PhD diss.**, University of Cambridge, 2020.

Challenges of computing

$$\mathrm{Sp}(\mathcal{K}) = \{\lambda \in \mathbb{C} : \mathcal{K} - \lambda I \text{ is not invertible}\}$$

Truncate/discretize: $\mathcal{K} \longrightarrow \mathbb{K} \in \mathbb{C}^{N\times N}$

1) **Too much:** Spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$
2) **Too little:** Miss parts of $\mathrm{Spec}(\mathcal{K})$
3) **Continuous spectra**
4) **Verification**
5) **Instability**

Some inf-dim comp. spec. problems cannot be solved, regardless of computational power, time or model.

- C., *"The foundations of infinite-dimensional spectral computations,"* **PhD diss.**, University of Cambridge, 2020.

A simple example on $\ell^2(\mathbb{Z})$

$$\begin{pmatrix} \ddots & \ddots & & & & \\ & 0 & 1 & & & \\ & & 0 & 1 & & \\ & & & 0 & 1 & \\ & & & & 0 & \ddots \\ & & & & & \ddots \end{pmatrix} \quad \longrightarrow \quad \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix} \in \mathbb{C}^{N\times N}$$

Two-way infinite

- Spectrum is unit circle.
- Spectrum is stable.
- Continuous spectra.

- Spectrum is $\{0\}$.
- Spectrum is unstable.
- Discrete spectra.

Example might look silly, but lots of Koopman operators are built up from operators like these!

Dynamic Mode Decomposition (DMD)

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

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

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

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

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

Challenges: too much, too little, continuous spectra, verification, instability.

- Schmid, *"Dynamic mode decomposition of numerical and experimental data,"* **J. Fluid Mech.**, 2010.
- Rowley, Mezić, Bagheri, Schlatter, Henningson, *"Spectral analysis of nonlinear flows,"* **J. Fluid Mech.**, 2009.
- 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,"* **J. Nonlinear Sci.**, 2015.

Dynamic Mode Decomposition (DMD)

Given dictionary $\{\psi_1, \ldots, \psi_N\}$ of functions $\psi_j : \Omega \rightarrow \mathbb{C}$,

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

$$\langle \psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \psi_k(x^{(m)}) = \Bigg[\underbrace{\begin{pmatrix} \psi_1(x^{(1)}) \\ \vdots \\ \psi_1(x^{(M)}) \end{pmatrix}}_{\Psi_X} \ \underbrace{\ w_M}_{W} \ \underbrace{\begin{pmatrix} \ ) \\ \ )) \end{pmatrix}}_{\Psi_X} \Bigg]_{jk}$$

$$\langle \mathcal{K}\psi_k, \psi_j \rangle \approx \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})} = \Bigg[\begin{pmatrix} \psi_1(x^{(1)}) \\ \vdots \\ \psi_1(x^{(M)}) \end{pmatrix} \ w \ \begin{pmatrix} 1)) \\ M)) \end{pmatrix} \Bigg]_{jk}$$

$$\mathcal{K} \longrightarrow \mathbb{K} = (\Psi_X{}^* W \Psi_X)^{-1} \ldots)^{\dagger} \ldots$$

Challenges: too much, too little, c… …er

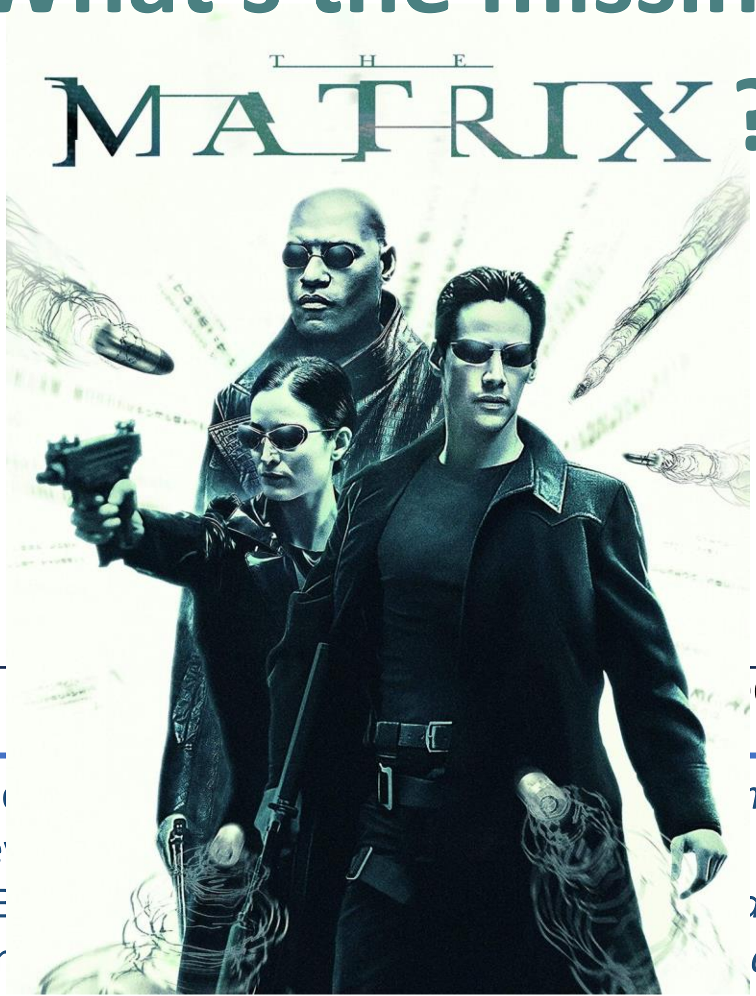
- Schmid, "*Dynamic mode decomposition of numerical and experimental data,*" **J. Fluid Mech.**, 2010.
- Rowley, Mezić, Bagheri, Schlatter, Henningson, "*Spectral analysis of nonlinear flows,*" **J. Fluid Mech.**, 2009.
- Kutz, Brunton, Brunton, Proctor, "*Dynamic mode decomposition: data-driven modeling of complex systems,*" **SIAM**, 2016.
- Williams, Kevrekidis, Rowley "*A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,*" **J. Nonlinear Sci.**, 2015.

Dynamic Mode Decomposition (DMD)

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

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

What's the missing



?

$\langle \psi$ … $x^{(1)})$ … $x^{(M)})$ … w_M … Ψ_X … W … Ψ_X … $)]_{jk}$

$\langle \mathcal{K}\psi$ … $(x^{(1)})$ ⋮ $(x^{(M)})$ … w … $1)$ … $M)$ … $]_{jk}$

$)^{-1}$ … $)^{\dagger}$

o little, c … ver

- Schmi… …nerical and experimental data," **J. Fluid Mech.**, 2010.
- Rowle… "Spectral analysis of nonlinear flows," **J. Fluid Mech.**, 2009.
- Kutz, B… …de decomposition: data-driven modeling of complex systems," **SIAM**, 2016.
- William… …ation of the Koopman operator: Extending dynamic mode decomposition," **J. Nonlinear Sci.**, 2015.

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

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

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

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

- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* **Commun. Pure Appl. Math.**, 2023.
- C., Ayton, Szőke, *"Residual Dynamic Mode Decomposition,"* **J. Fluid Mech.**, 2023.
- Code: https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition

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

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

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

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

Residuals: $g = \sum_{j=1}^{N} \mathbf{g}_j \psi_j,\ \|\mathcal{K}g - \lambda g\|^2 = \langle \mathcal{K}g - \lambda g, \mathcal{K}g - \lambda g \rangle$

- C., Townsend, "*Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems*," **Commun. Pure Appl. Math.**, 2023.
- C., Ayton, Szőke, "*Residual Dynamic Mode Decomposition*," **J. Fluid Mech.**, 2023.
- Code*:* https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition

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

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

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

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

Residuals: $g = \sum_{j=1}^{N} \mathbf{g}_j \psi_j, \ \|\mathcal{K}g - \lambda g\|^2 = \sum_{k,j=1}^{N} \mathbf{g}_k \overline{\mathbf{g}_j} \langle \mathcal{K}\psi_k - \lambda\psi_k, \mathcal{K}\psi_j - \lambda\psi_j \rangle$

- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* **Commun. Pure Appl. Math.**, 2023.
- C., Ayton, Szőke, *"Residual Dynamic Mode Decomposition,"* **J. Fluid Mech.**, 2023.
- Code: https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition

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

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

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

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

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

- C., Townsend, *"Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,"* **Commun. Pure Appl. Math.**, 2023.
- C., Ayton, Szőke, *"Residual Dynamic Mode Decomposition,"* **J. Fluid Mech.**, 2023.
- Code: https://github.com/MColbrook/Residual-Dynamic-Mode-Decomposition

ResDMD: avoiding “too much”

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

eigenvectors

eigenvalues

Algorithm 1:

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

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

eigenvectors

eigenvalues

Algorithm 1:

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

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

Algorithm 2:

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

1. Compute $G, K_1, K_2 \in \mathbb{C}^{N\times N}$.
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 $N \to \infty$)

Quadrature with trajectory data

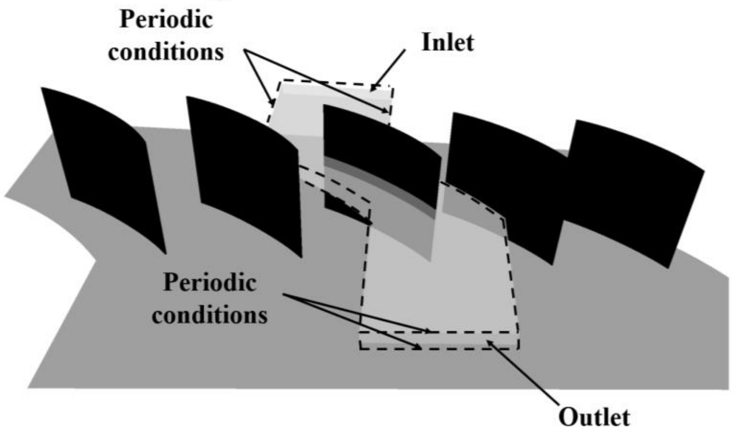
E.g., $\langle \mathcal{K}\psi_k, \psi_j \rangle = \lim_{M\to\infty} \sum_{m=1}^{M} w_m \overline{\psi_j(x^{(m)})} \underbrace{\psi_k(y^{(m)})}_{[\mathcal{K}\psi_k](x^{(m)})}$

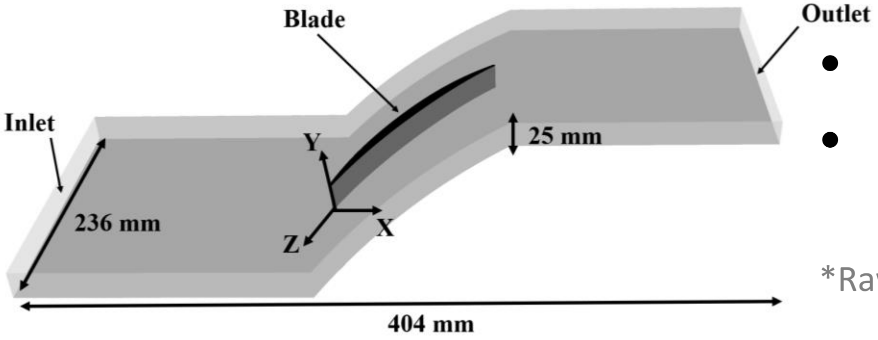
Three examples:

- **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: Verified modes

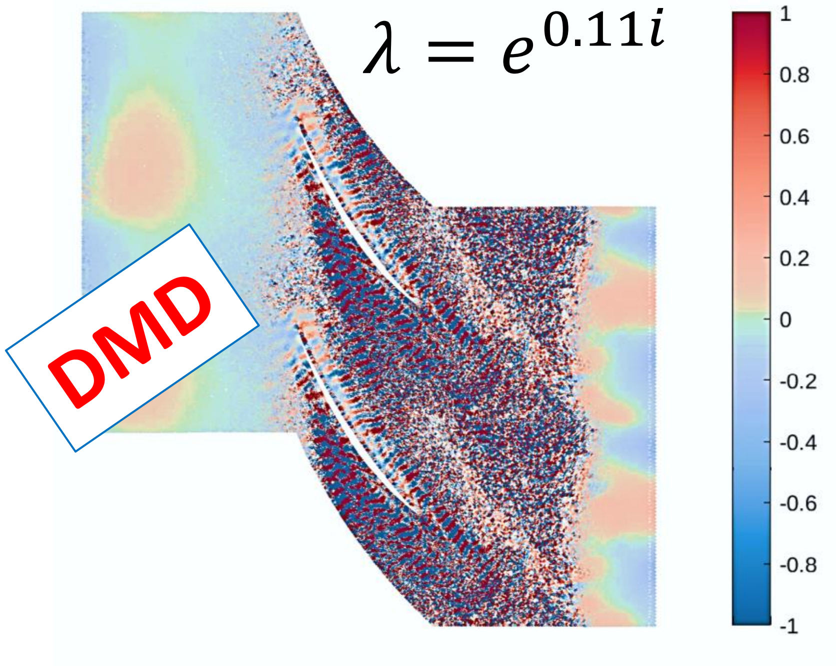




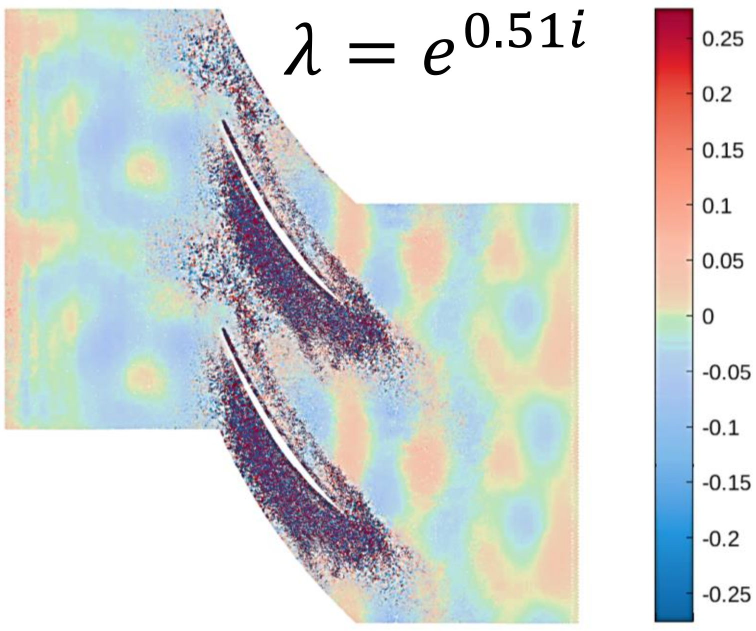
- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension $(d) \approx 300{,}000$ (number of measurement points)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

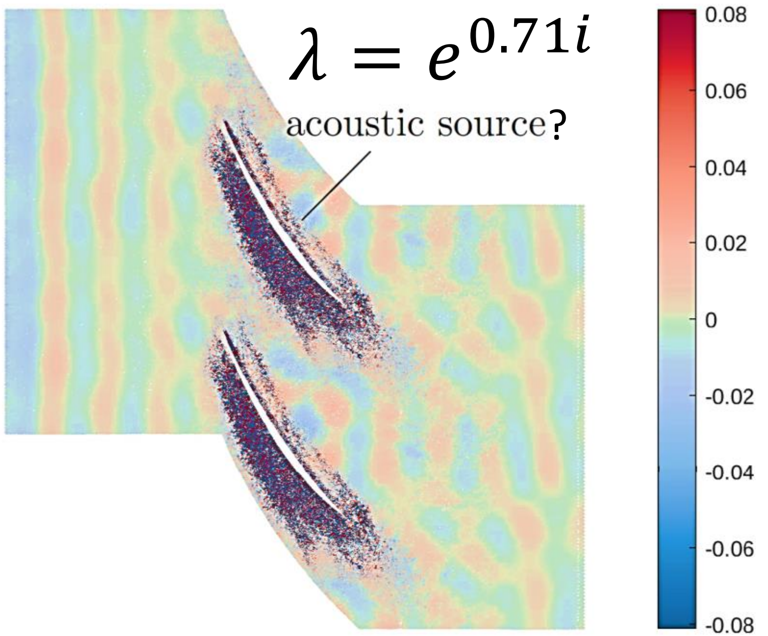
Rel. Error = ?



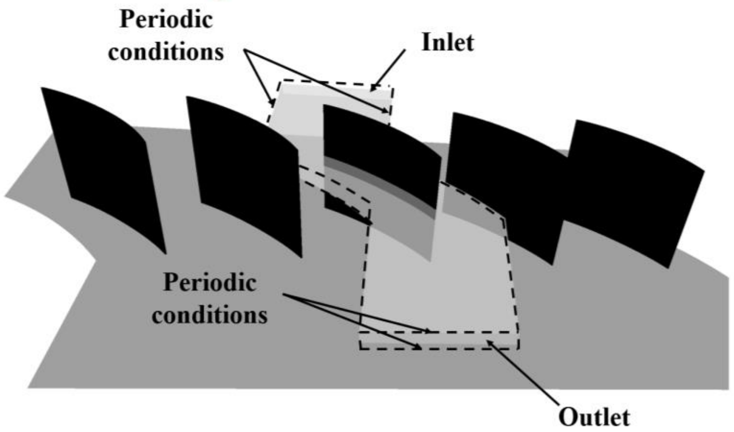
Rel. Error = ?

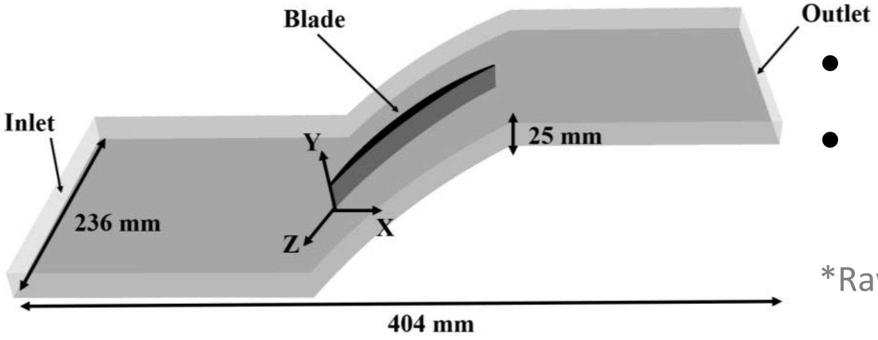


Rel. Error = ?



Example: Verified modes

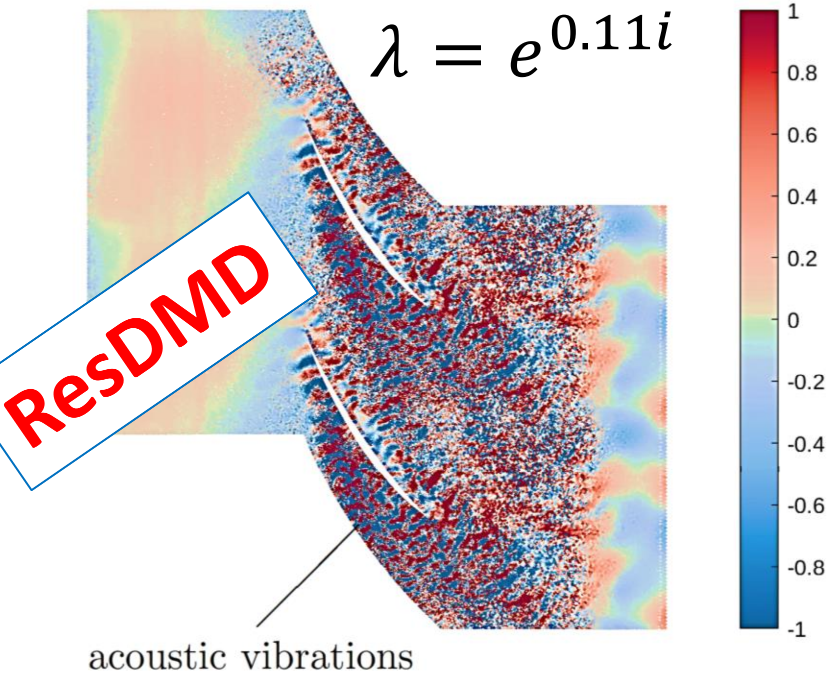




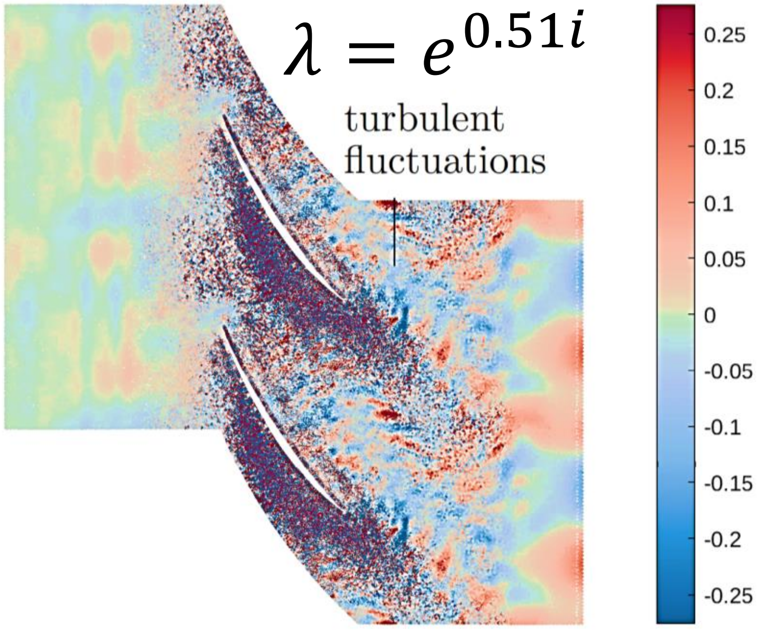
- Reynolds number $\approx 3.9 \times 10^5$
- Ambient dimension $(d) \approx 300{,}000$ (number of measurement points)

*Raw measurements provided by Stephane Moreau (Sherbrooke)

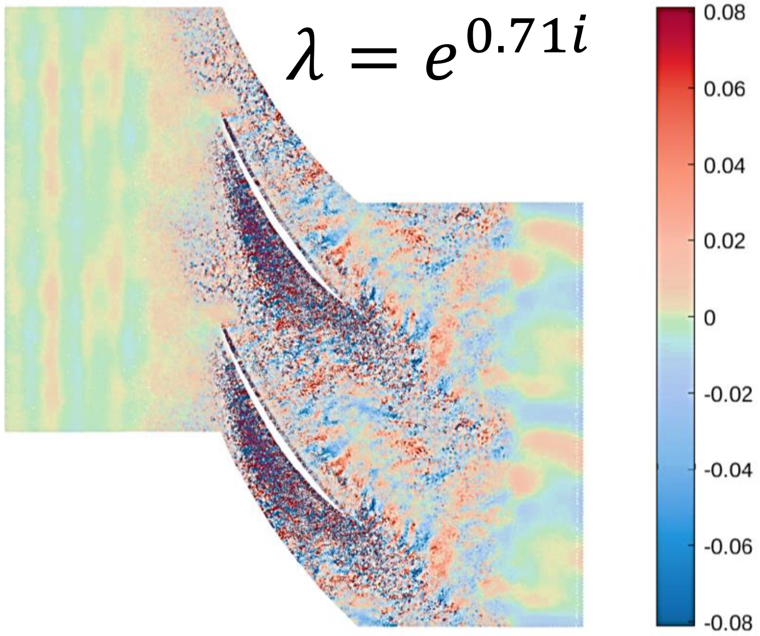
Rel. Error ≤ 0.0054



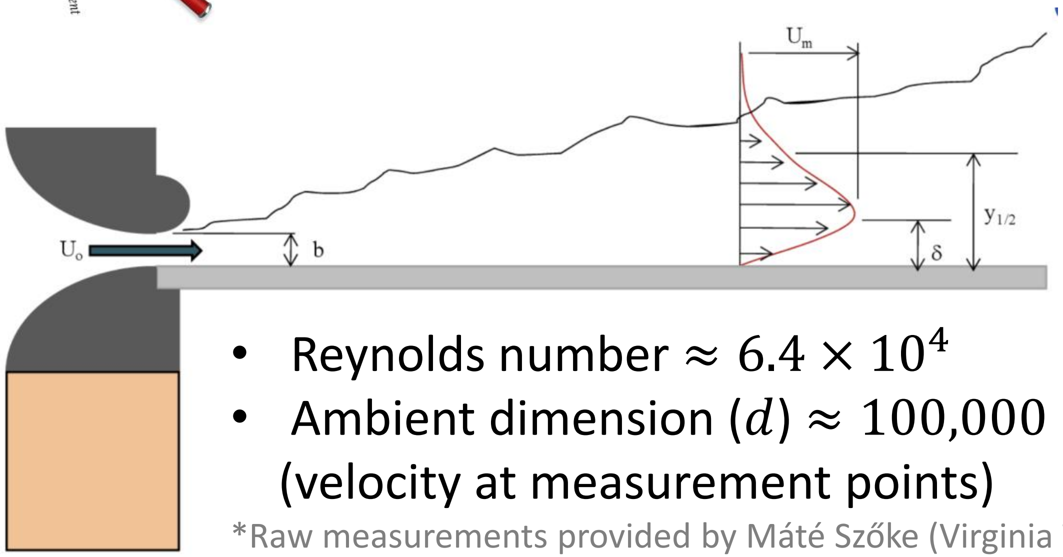
Rel. Error ≤ 0.0128



Rel. Error ≤ 0.0196

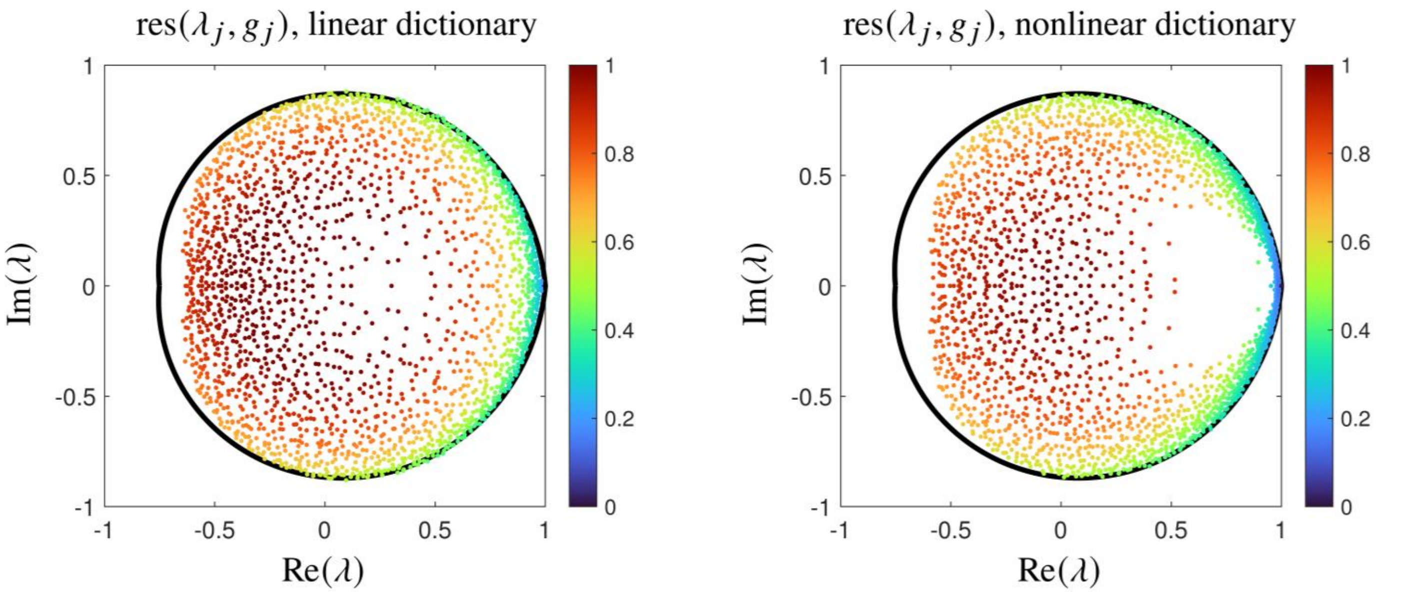


Example: Verified dictionary



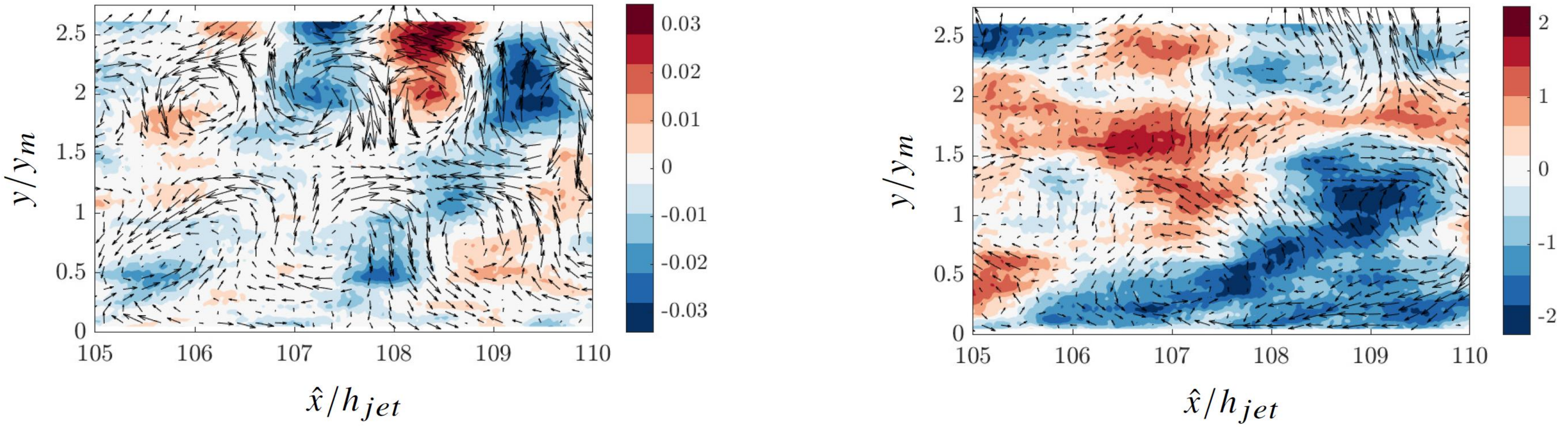
- Reynolds number $\approx 6.4 \times 10^4$
- Ambient dimension (d) $\approx 100{,}000$ (velocity at measurement points)

*Raw measurements provided by Máté Szőke (Virginia Te

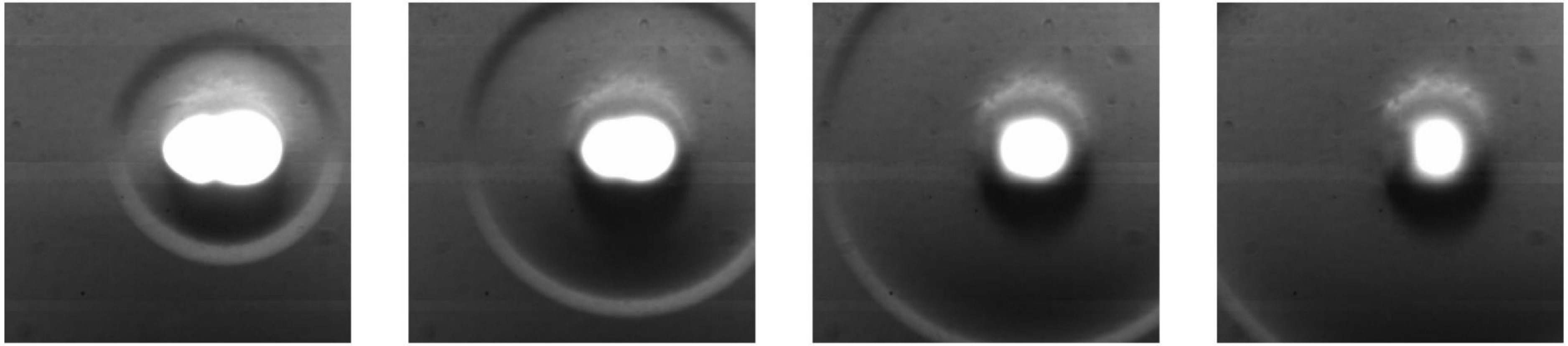


$\lambda = 0.9439 + 0.2458i$, error ≤ 0.0765

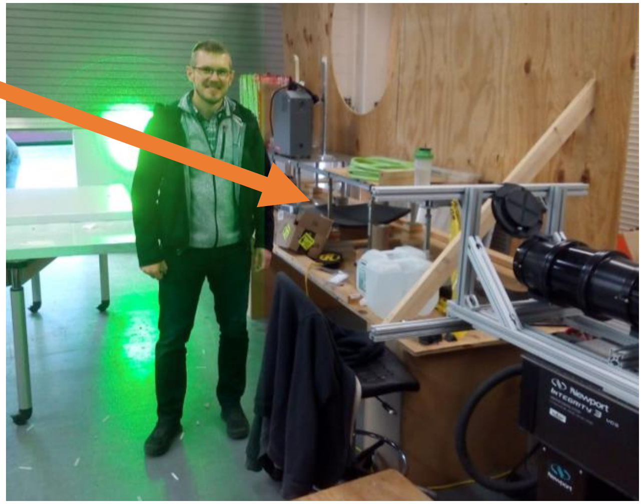
$\lambda = 0.8948 + 0.1065i$, error ≤ 0.1105



Example: Verified mode decomposition

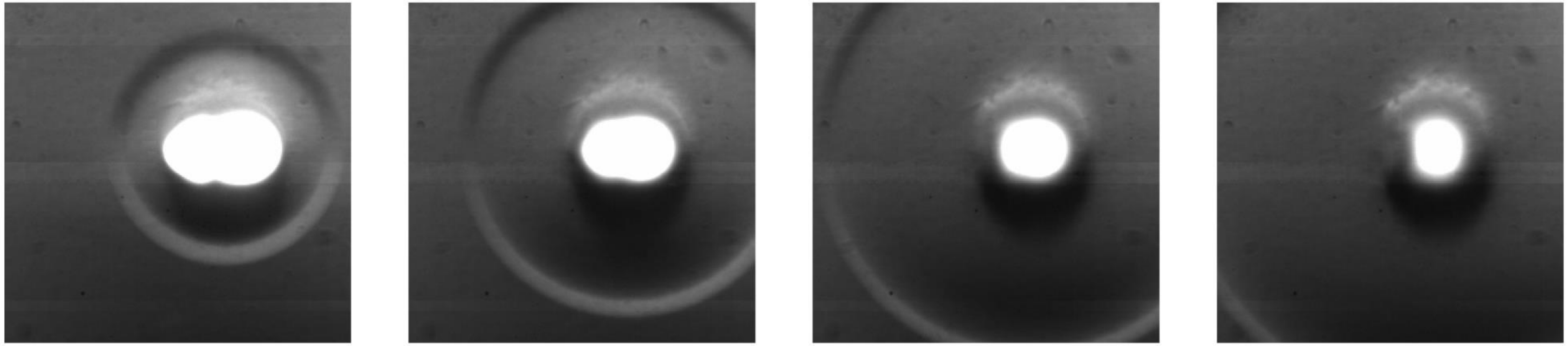


a) $t = 5\ \mu s$ b) $t = 10\ \mu s$ c) $t = 15\ \mu s$ d) $t = 20\ \mu s$

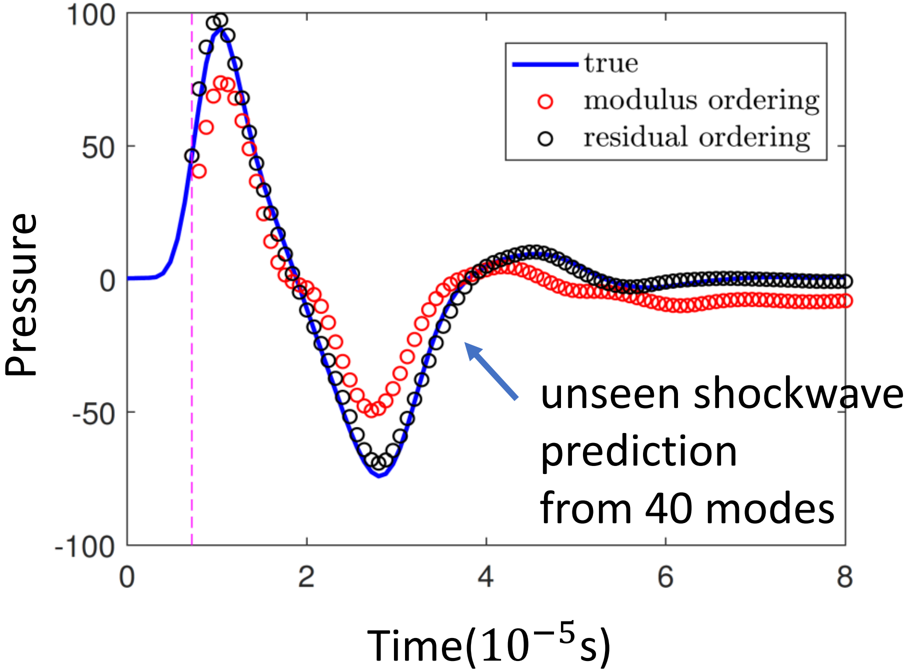


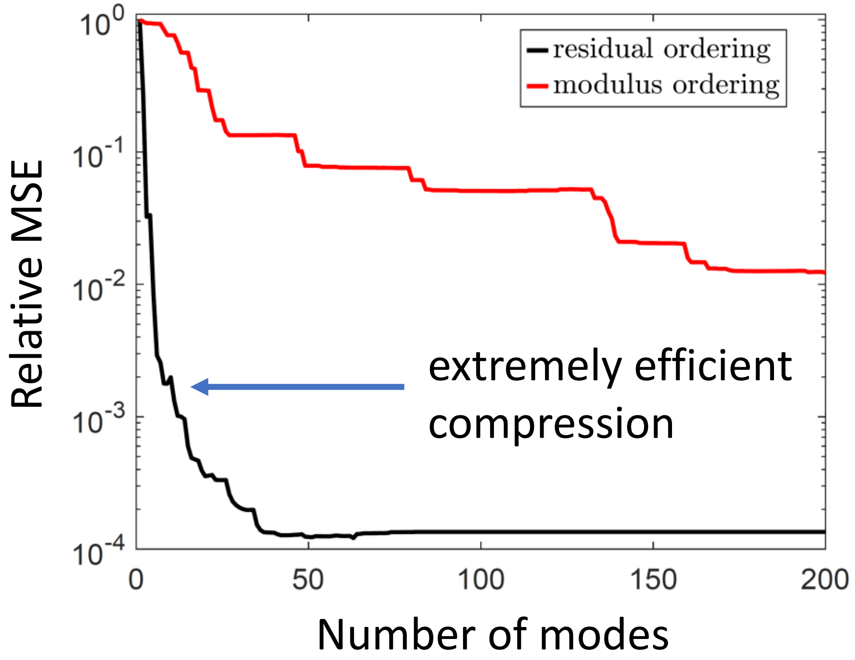
Matt Szőke showing me his laser cannon!

Example: Verified mode decomposition



a) $t = 5\ \mu s$ b) $t = 10\ \mu s$ c) $t = 15\ \mu s$ d) $t = 20\ \mu s$





Stochastic Dynamical Systems

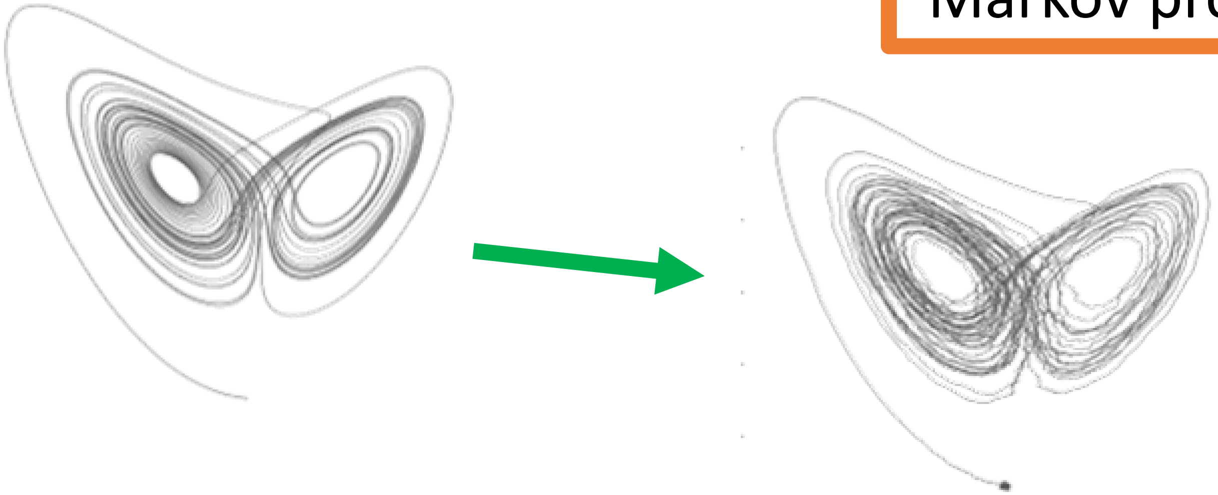
State $x \in \Omega \subseteq \mathbb{R}^d$, i.i.d. random variables $\tau_1, \tau_2, \ldots$

Unknown function F governs dynamics:

$$x_n = F(x_{n-1}, \tau_n) = F_{\tau_n}(x_{n-1})$$

Discrete-time Markov process!

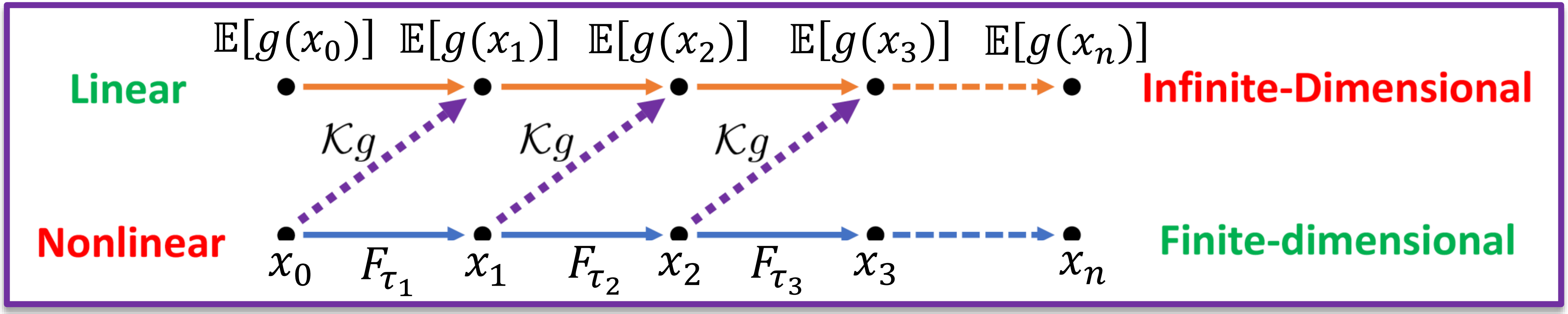
E.g., models noise or uncertainty, or truly random process.



Goal: Verified learning from data $\{x^{(m)}, y^{(m)} = F_{\tau_m}(x^{(m)})\}_{m=1}^{M}$.

Stochastic Koopman Operator

Now we have an expectation: $[\mathcal{K}g](x) = \mathbb{E}\left[g\left(F_\tau(x)\right)\right]$



Satisfies semigroup property: $\mathbb{E}[g(\underbrace{F \circ F \circ \cdots \circ F}_{n \text{ times}})] = \mathcal{K}^n g$

Time for an example!

Stochastic Van der Pol oscillator:

$$\begin{aligned} \mathrm{d}X_1 &= X_2 \mathrm{d}t, \\ \mathrm{d}X_2 &= [0.5(1 - X_1^2)X_2 - X_1]\mathrm{d}t + 0.2\mathrm{d}B_t \end{aligned}$$

Turn into discrete-time system with step 0.3.

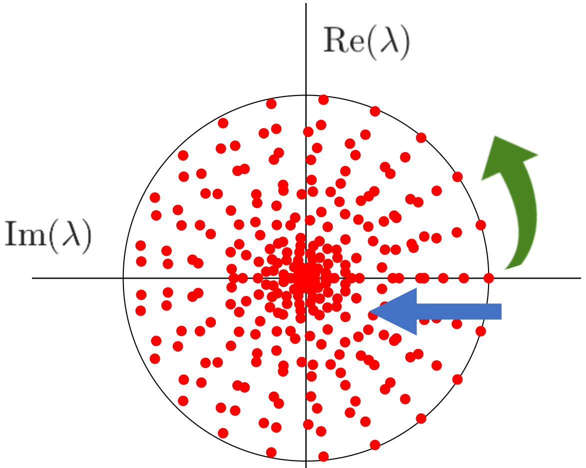
$$\begin{aligned} \mathbb{E}\left[g_\lambda\left(F_{\tau_n} \circ \cdots \circ F_{\tau_1}(x)\right)\right] \\ = [\mathcal{K}^n g_\lambda](x) \\ = \lambda^n g_\lambda(x) \end{aligned}$$

Time for an example!

Stochastic Van der Pol oscillator:

$$\begin{aligned} \mathrm{d}X_1 &= X_2 \mathrm{d}t, \\ \mathrm{d}X_2 &= [0.5(1 - X_1^2)X_2 - X_1]\mathrm{d}t + 0.2\mathrm{d}B_t \end{aligned}$$

Turn into discrete-time system with step 0.3.



Eigenvalues λ

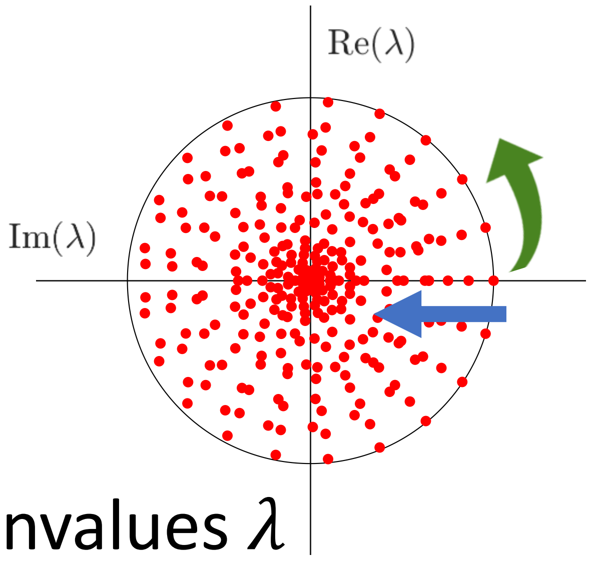
$$\begin{aligned} \mathbb{E}\left[g_\lambda\left(F_{\tau_n} \circ \cdots \circ F_{\tau_1}(x)\right)\right] \\ = [\mathcal{K}^n g_\lambda](x) \\ = \lambda^n g_\lambda(x) \end{aligned}$$

Time for an example!

Stochastic Van der Pol oscillator:

$$\mathrm{d}X_1 = X_2 \mathrm{d}t,$$
$$\mathrm{d}X_2 = [0.5(1 - X_1^2)X_2 - X_1]\mathrm{d}t + 0.2\mathrm{d}B_t$$

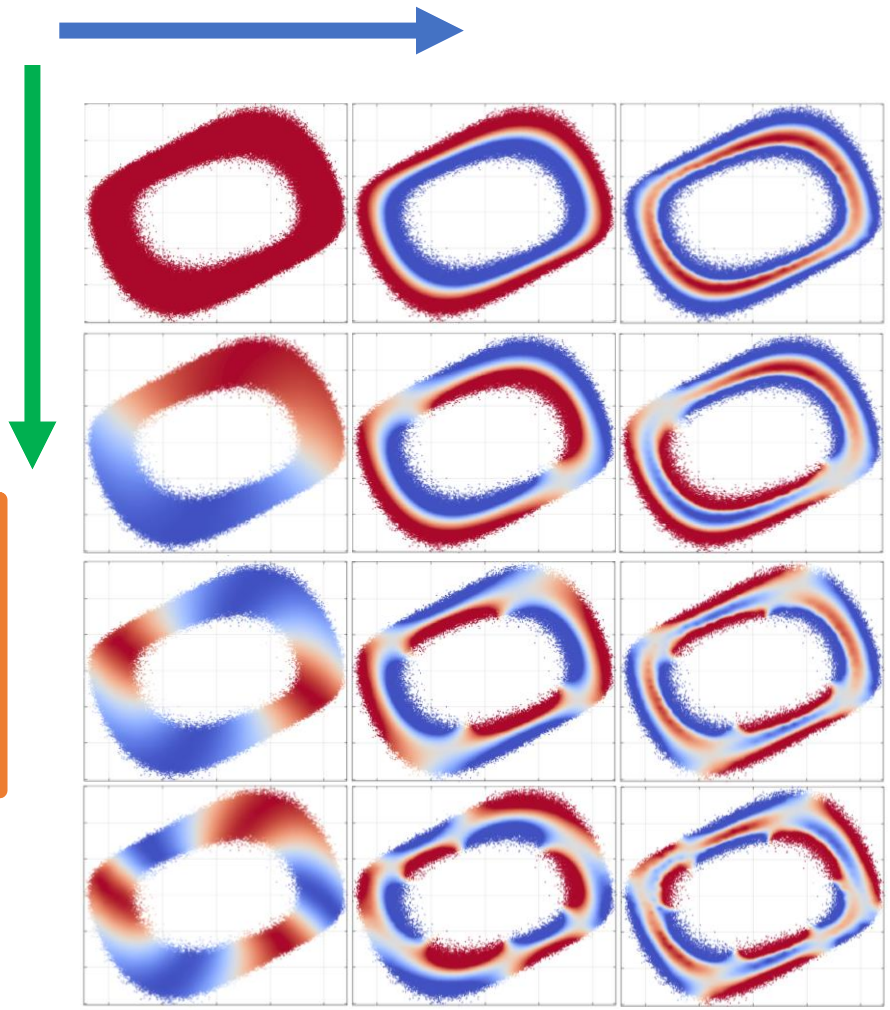
Turn into discrete-time system with step 0.3.



Eigenvalues λ

$$\mathbb{E}\left[g_\lambda\left(F_{\tau_n} \circ \cdots \circ F_{\tau_1}(x)\right)\right] = [\mathcal{K}^n g_\lambda](x) = \lambda^n g_\lambda(x)$$

Eigenfunctions g_λ

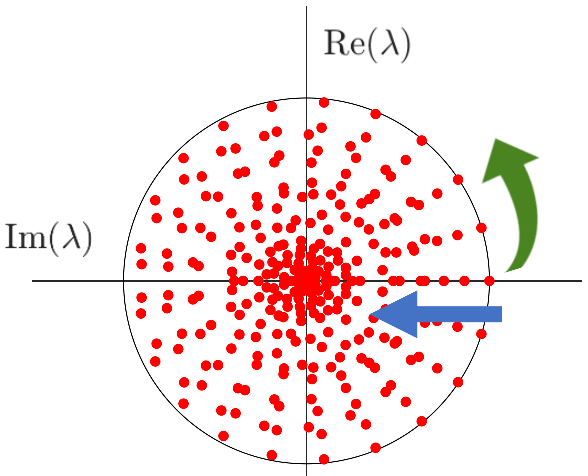


Time for an example!

Stochastic Van der Pol oscillator:

$$\mathrm{d}X_1 = X_2\mathrm{d}t,$$
$$\mathrm{d}X_2 = [0.5(1 - X_1^2)X_2 - X_1]\mathrm{d}t + 0.2\mathrm{d}B_t$$

Turn into discrete-time system with step 0.3.

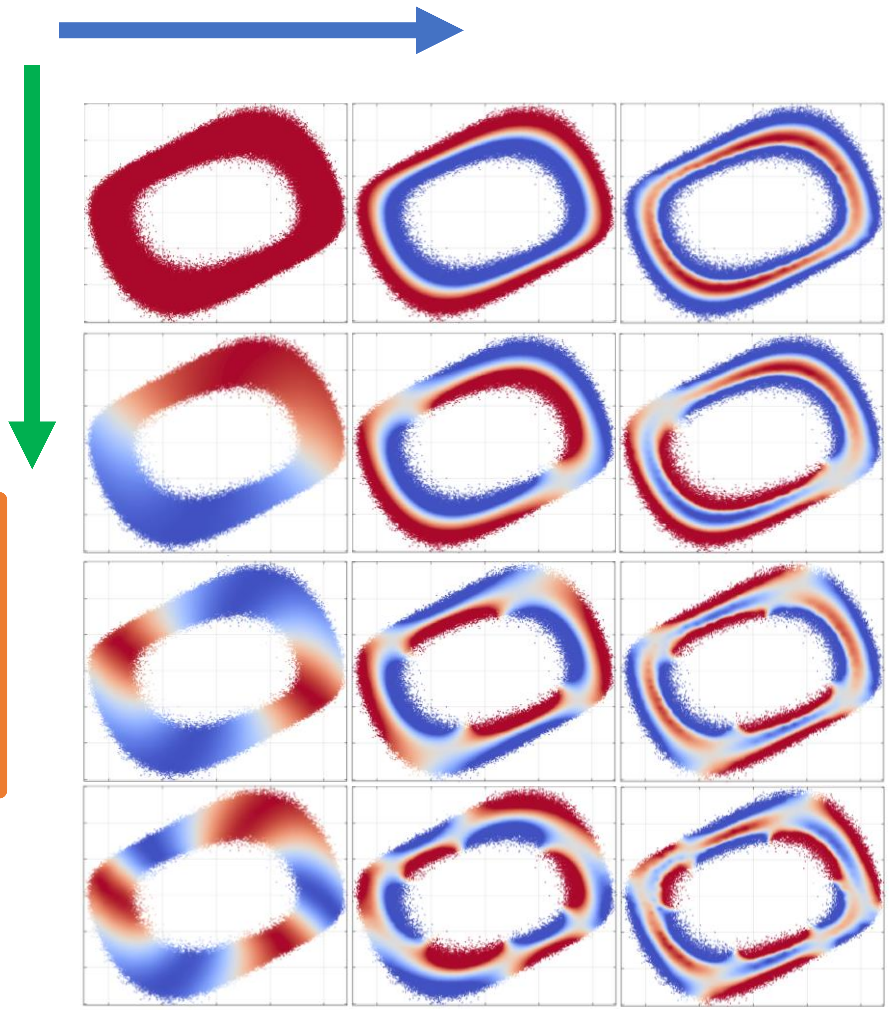


Eigenvalues λ

$$\mathbb{E}\left[g_\lambda\left(F_{\tau_n} \circ \cdots \circ F_{\tau_1}(x)\right)\right] = [\mathcal{K}^n g_\lambda](x) = \lambda^n g_\lambda(x)$$

Is this enough?

Eigenfunctions g_λ



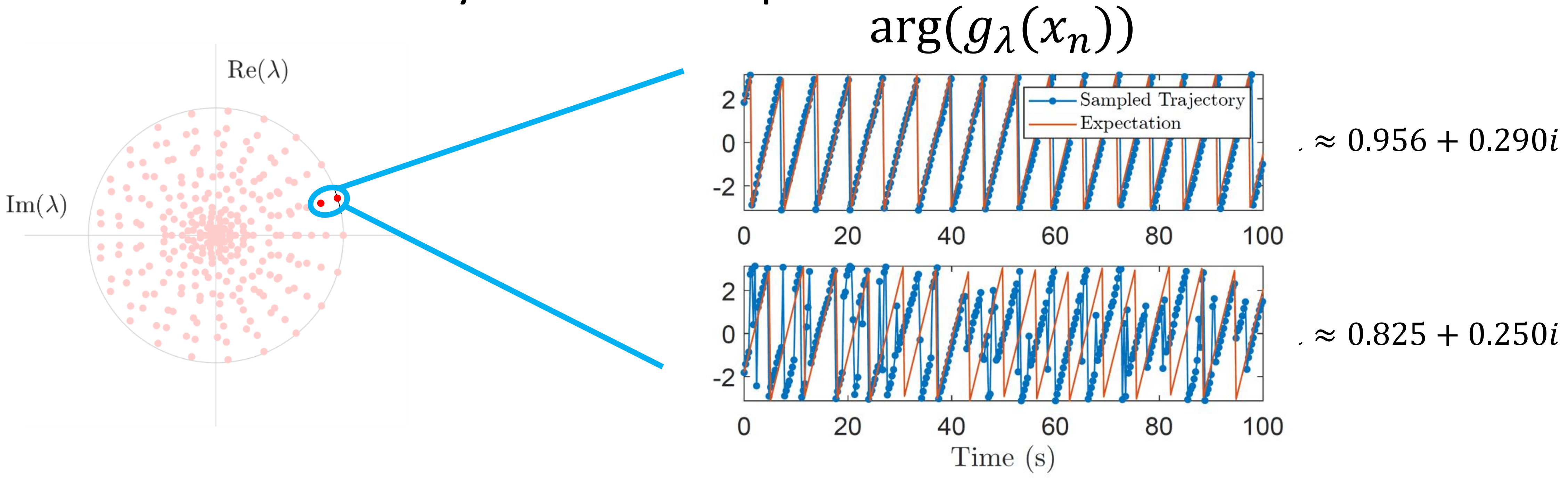
Time for an example!

Stochastic Van der Pol oscillator:

$$\mathrm{d}X_1 = X_2\mathrm{d}t,$$
$$\mathrm{d}X_2 = [0.5(1 - X_1^2)X_2 - X_1]\mathrm{d}t + 0.2\mathrm{d}B_t$$

Same phase, but clearly one is more coherent than the other!

Turn into discrete-time system with step 0.3.



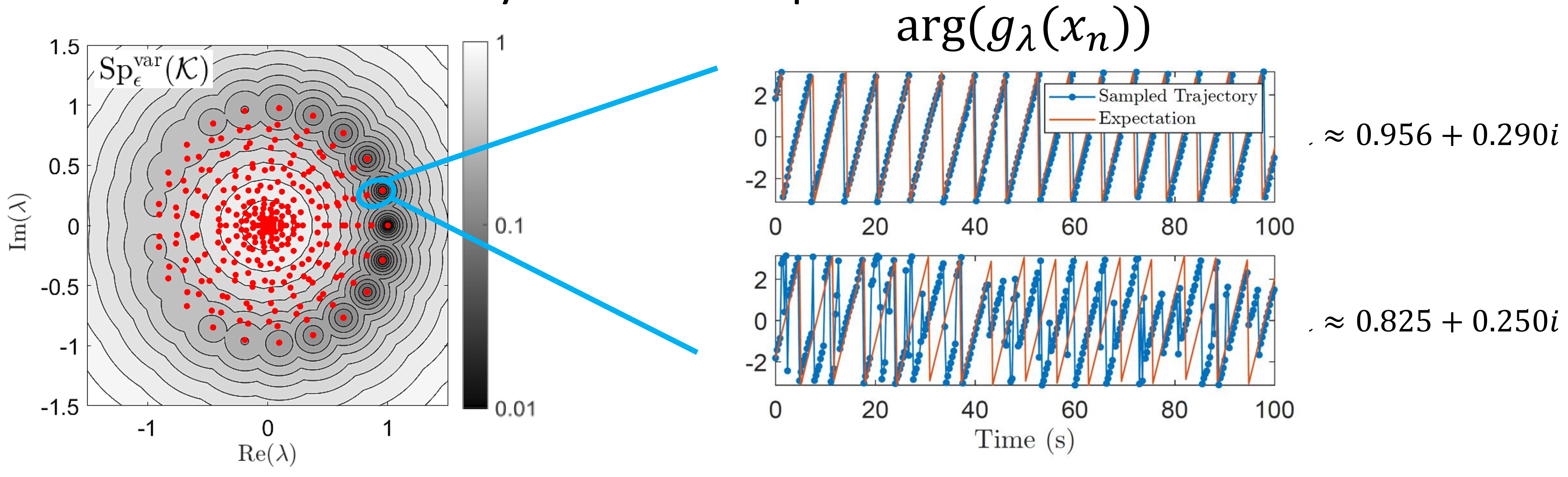
Time for an example!

Stochastic Van der Pol oscillator:

$$\mathrm{d}X_1 = X_2\mathrm{d}t,$$
$$\mathrm{d}X_2 = [0.5(1 - X_1^2)X_2 - X_1]\mathrm{d}t + 0.2\mathrm{d}B_t$$

Same phase, but clearly one is more coherent than the other!

Turn into discrete-time system with step 0.3.



Decomposing the residual

$$\lim_{M\to\infty} \mathbf{g}^*\left[K_2 - \lambda {K_1}^* - \bar{\lambda} K_1 + |\lambda|^2 G\right]\mathbf{g}$$

$$= \mathbb{E}[\|g \circ F_\tau - \lambda g\|^2]$$

$$= \|\mathcal{K}g - \lambda g\|^2 + \int_\Omega \mathrm{Var}[g(F_\tau(x))]\, \mathrm{d}\omega(x)$$

Squared residual — Integrated variance

Decomposing the residual

$$\lim_{M\to\infty} \mathbf{g}^*\left[K_2 - \lambda {K_1}^* - \bar{\lambda} K_1 + |\lambda|^2 G\right]\mathbf{g}$$

$$= \mathbb{E}[\|g \circ F_\tau - \lambda g\|^2]$$

$$= \|\mathcal{K}g - \lambda g\|^2 + \int_\Omega \mathrm{Var}[g(F_\tau(x))]\, \mathrm{d}\omega(x)$$

Squared residual — Integrated variance

Definition: For $\varepsilon > 0$ we define the variance-ε-pseudospectrum

$$\mathrm{Sp}_\varepsilon^{\mathrm{var}}(\mathcal{K}) = \{\lambda \in \mathbb{C}: \exists g \in \mathcal{D}(\mathcal{K}), \|g\| = 1, \mathbb{E}[\|g \circ F_\tau - \lambda g\|^2] < \varepsilon^2\}$$

Decomposing the residual

$$\lim_{M\to\infty} \mathbf{g}^*\left[K_2 - \lambda K_1{}^* - \bar{\lambda} K_1 + |\lambda|^2 G\right]\mathbf{g}$$
$$= \mathbb{E}[\|g \circ F_\tau - \lambda g\|^2]$$
$$= \|\mathcal{K}g - \lambda g\|^2 + \int_\Omega \mathrm{Var}[g(F_\tau(x))]\, \mathrm{d}\omega(x)$$

Squared residual — Integrated variance

Definition: For $\varepsilon > 0$ we define the variance-ε-pseudospectrum

$$\mathrm{Sp}_\varepsilon^{\mathrm{var}}(\mathcal{K}) = \{\lambda \in \mathbb{C}: \exists g \in \mathcal{D}(\mathcal{K}), \|g\| = 1, \mathbb{E}[\|g \circ F_\tau - \lambda g\|^2] < \varepsilon^2\}$$

Measure of statistical coherency.

Sampler of the results we prove

Representation of higher moments via batched Koopman operators:

$$g: \Omega^r \to \mathbb{C}, \qquad [\mathcal{K}_{(r)} g](x) = \mathbb{E}[g(F_\tau, \dots, F_\tau)]$$

Convergent algorithms for each of the terms in the residual decomposition.

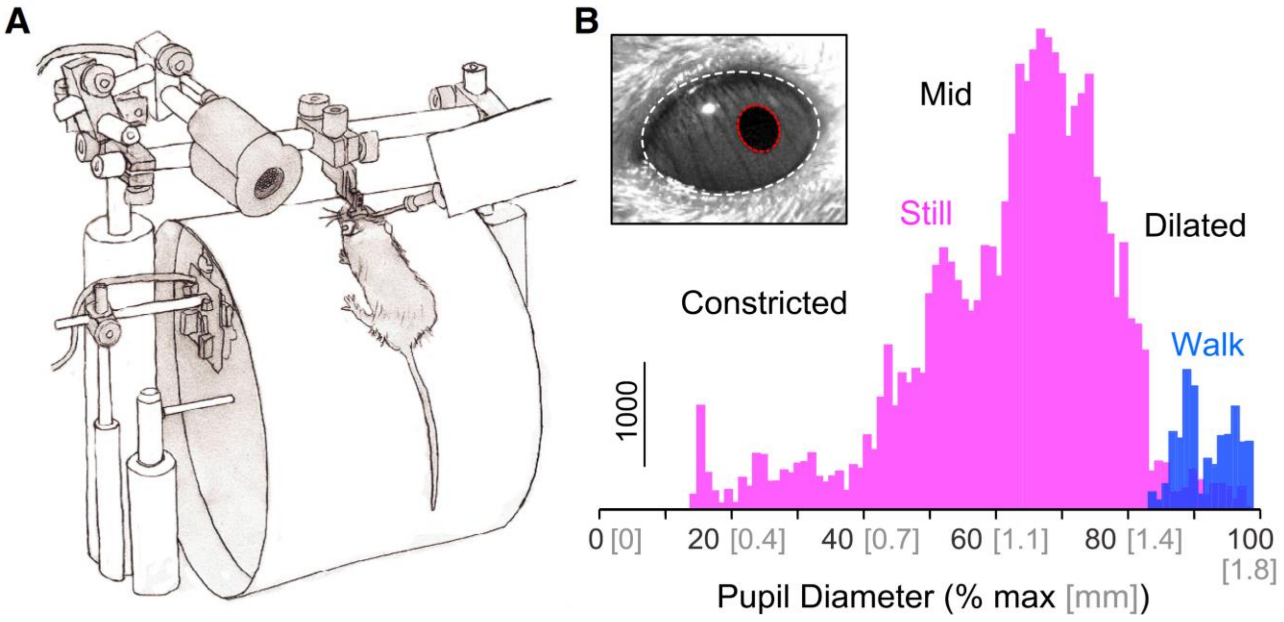
Computation of spectrum of $\mathcal{K}$ without issues such as spurious eigenvalues.

Control the error when we project to a finite-dimensional space!

Error bounds for Koopman Mode Decomposition.

Concentration bounds in terms of amount of snapshot data.

An application

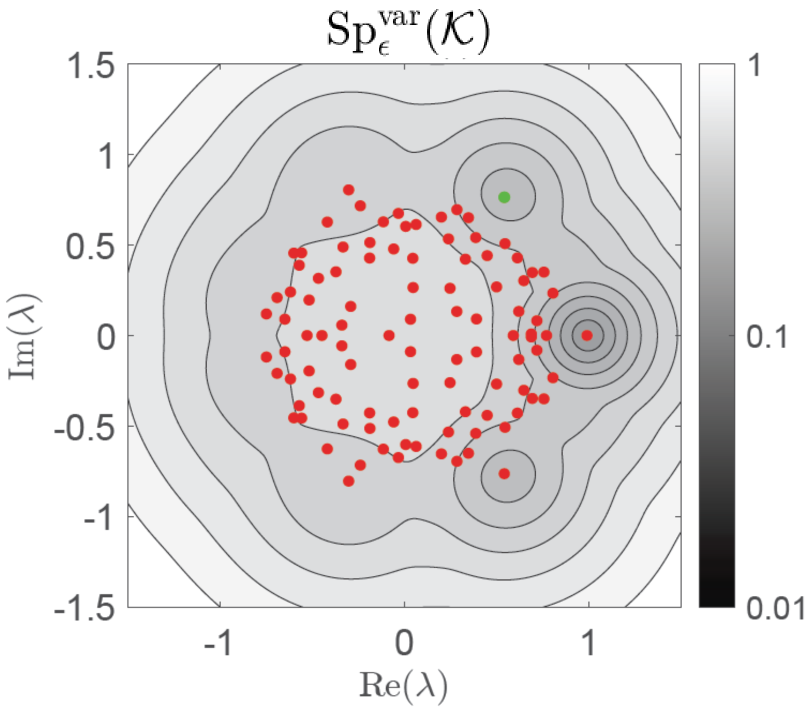


- Monitoring of large populations of neurons.
- Mice shown a drifting grating.
- Separate stochastic Koopman operators according to 15 different arousal levels (indexed by pupil diameter).

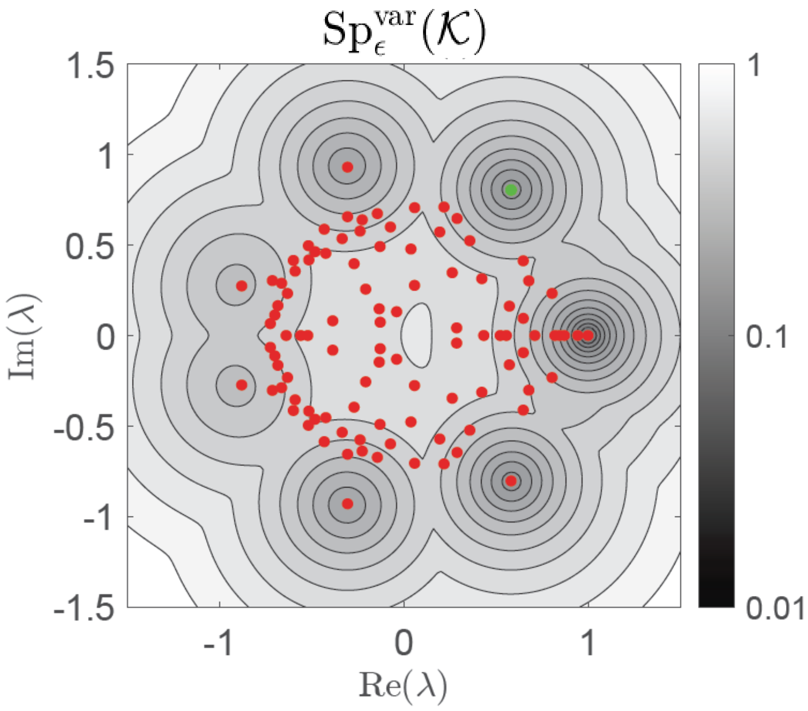
Standard DMD does not provide verification…

- Siegle, Joshua H., et al., *"Survey of spiking in the mouse visual system reveals functional hierarchy,"* **Nature**, 2021.
- McGinley, David, McCormick, "Cortical membrane potential signature of optimal states for sensory signal detection," **Neuron**, 2015.

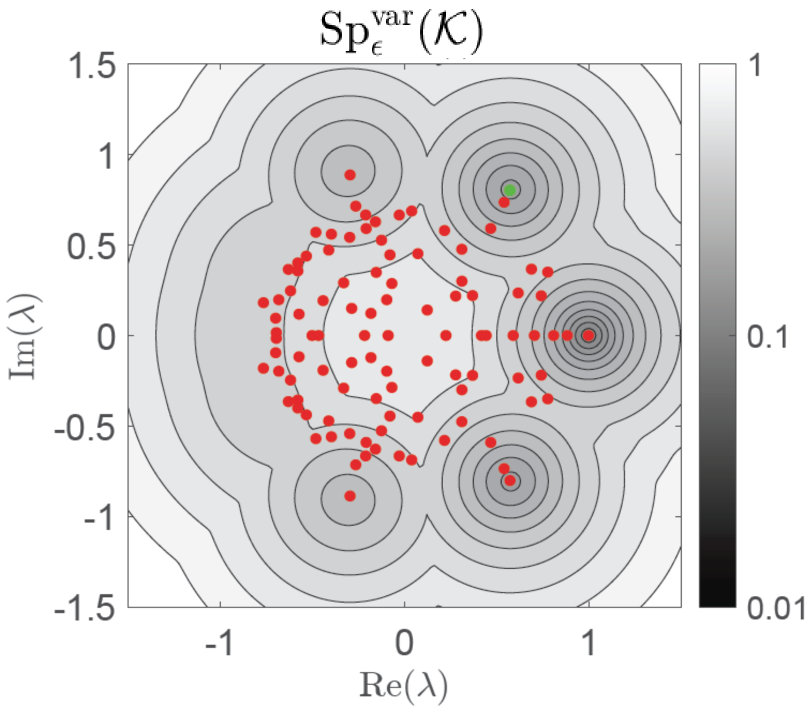
Variance pseudospectra of mouse # 11



pupil diameter 8%



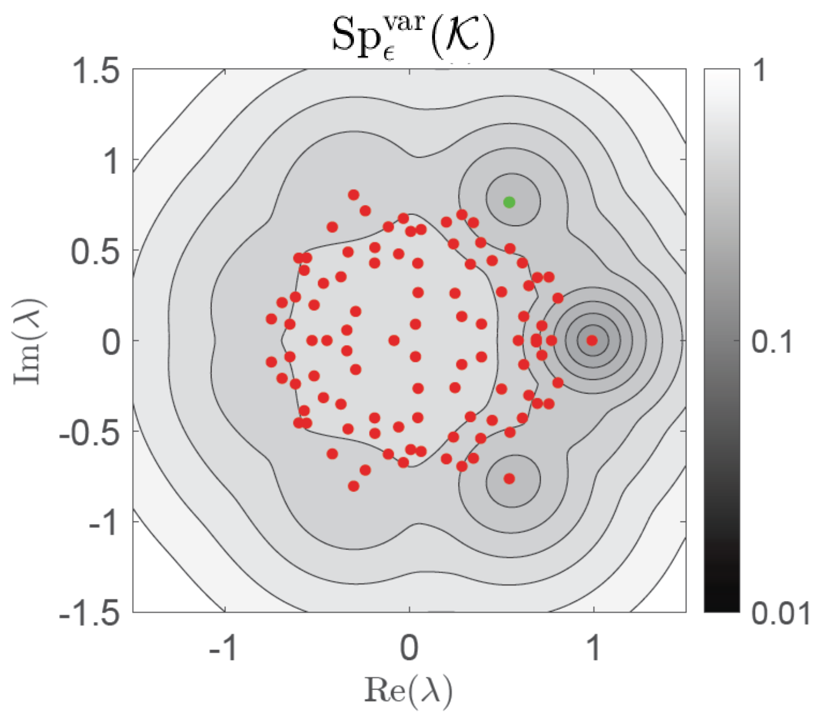
pupil diameter 28%



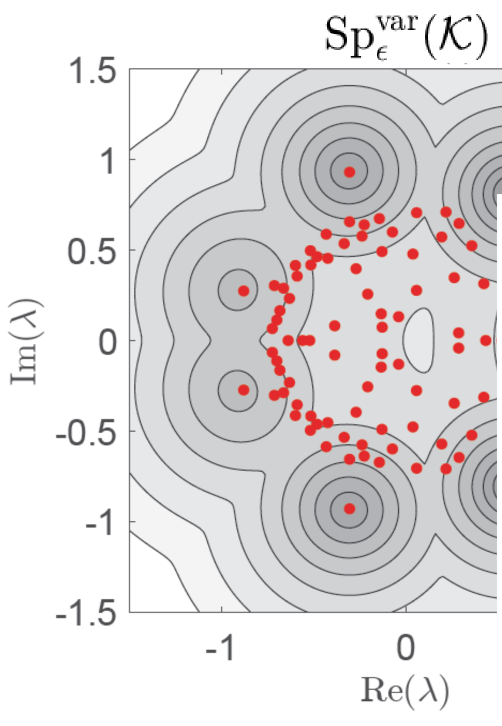
pupil diameter 43%

$$\sqrt{\|\mathcal{K}g-\lambda g\|^2+\int_\Omega \mathrm{Var}[g(F_\tau(x))]\,\mathrm{d}\omega(x)}$$

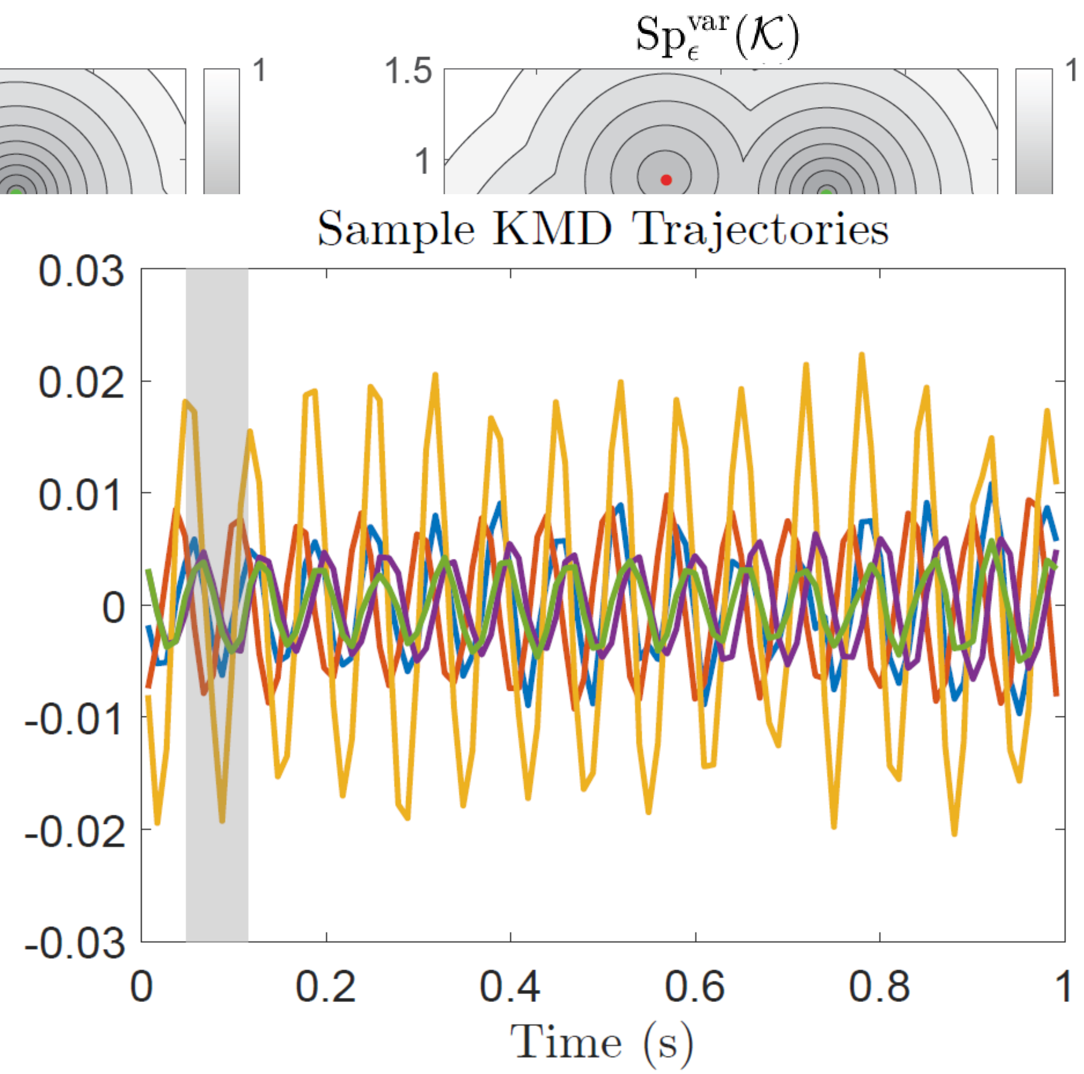
Variance pseudospectra of mouse # 11



pupil diameter 8%

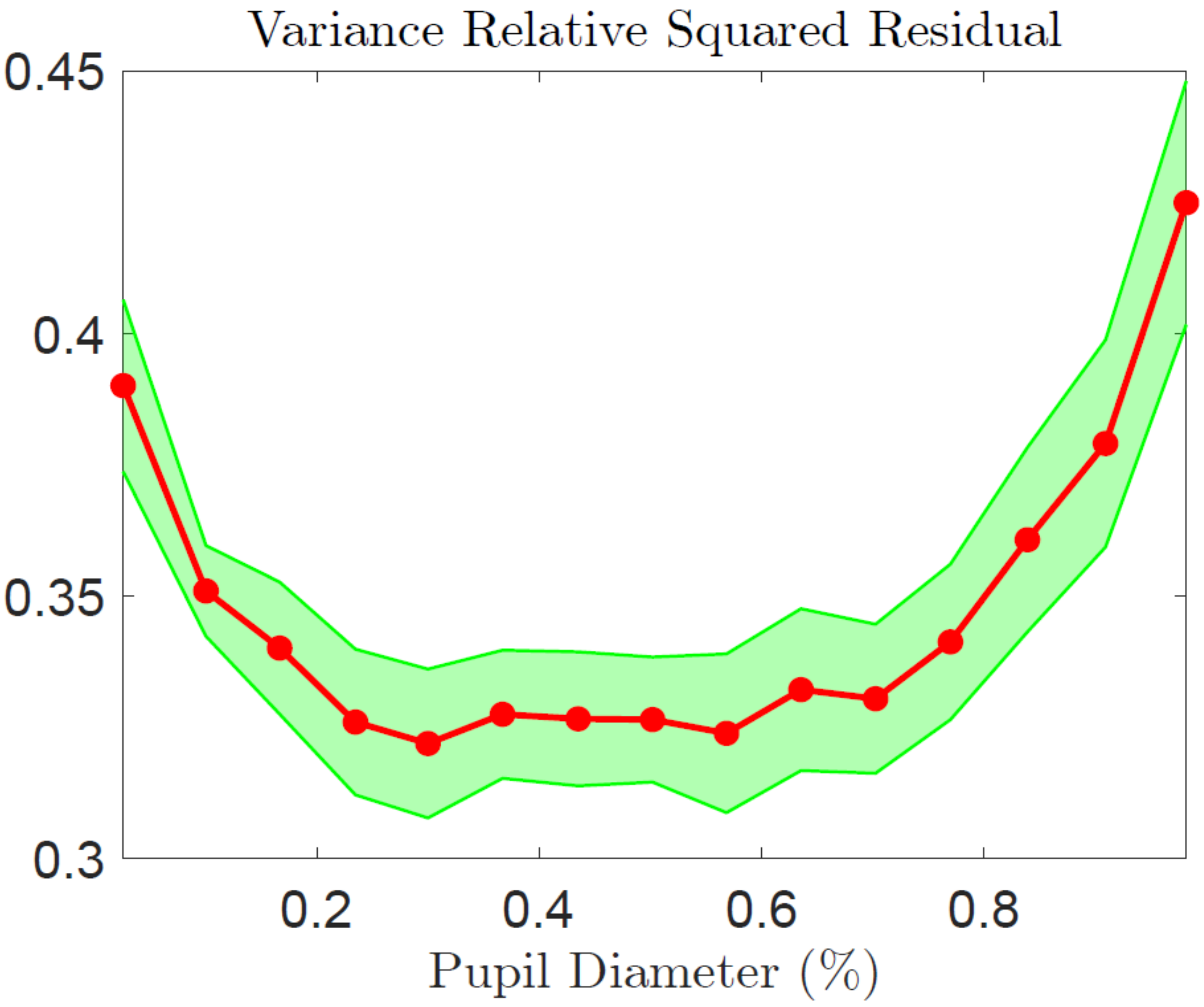


pupil diamet



$$\sqrt{\|\mathcal{K}g-\lambda g\|^2+\int_\Omega \mathrm{Var}[g(F_\tau(x))]\,\mathrm{d}\omega(x)}$$

Yerkes-Dodson law across all mice



Yerkes-Dodson law: *you reach your peak level of performance with an intermediate level of stress, or arousal. Too little or too much arousal results in poorer performance.*

Wider program: Solvability Complexity Index

- <u>Inf.-dim. computational analysis</u> $\Longrightarrow$ **Compute spectral properties rigorously, practically!**
- <u>Continuous linear algebra</u> $\Longrightarrow$ **Avoid the woes of discretization.**
- <u>Hierarchy</u> $\Longrightarrow$ **Classify difficulty of computational problems, prove algorithms are optimal.**

$$
\begin{array}{ccccccccccccccc}
\Pi_0^\alpha & & & & \Pi_1^\alpha & & & & \Pi_2^\alpha & & & & & & \\
\| & & & \subsetneq & & \supsetneq & & \subsetneq & & \supsetneq & & \subsetneq & & & \\
\Delta_0^\alpha & \subsetneq & \Delta_1^\alpha & \subsetneq & \Sigma_1^\alpha \cup \Pi_1^\alpha & \subsetneq & \Delta_2^\alpha & \subsetneq & \Sigma_2^\alpha \cup \Pi_2^\alpha & \subsetneq & \Delta_3^\alpha & \subsetneq & \cdots & & \\
\| & & & \subsetneq & & \supsetneq & & \subsetneq & & \supsetneq & & \subsetneq & & & \\
\Sigma_0^\alpha & & & & \Sigma_1^\alpha & & & & \Sigma_2^\alpha & & & & & &
\end{array}
$$

Extends to: Foundations of AI, optimization, computer-assisted proofs, and PDEs etc.

- C., *"On the computation of geometric features of spectra of linear operators on Hilbert spaces,"* **Found. Comput. Math.**, 2023.
- C., Hansen, *"The foundations of spectral computations via the solvability complexity index hierarchy,"* **J. Eur. Math. Soc.**, 2022.
- C., Antun, Hansen, "*The difficulty of computing stable and accurate neural networks,*" **Proc. Natl. Acad. Sci. USA**, 2022.
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, *"On the solvability complexity index hierarchy and towers of algorithms,"* arXiv, 2020.

Summary

- Koopman: **Nonlinearity**

Infinite dimensions

- Presented **verified data-driven methods** for Koopman operators (deterministic and stochastic)
- Methods are **cheap**, **easy-to-use**, come with **convergence guarantees**.
- For **continuous spectra,** see:
  - C., Townsend, *“Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems,”* **Commun. Pure Appl. Math.**, 2023.
  - C., "*The mpEDMD algorithm for data-driven computations of measure-preserving dynamical systems.*" **SIAM J. Numer. Anal.**, 2023.
  - C., Horning, Townsend, *“Computing spectral measures of self-adjoint operators,”* **SIAM Review**, 2021.

Brief Summaries

References

[1] Colbrook, Matthew J., and Alex Townsend. "Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems." arXiv preprint arXiv:2111.14889 (2021).

[2] Colbrook, Matthew J., Lorna J. Ayton, and Máté Szőke. "Residual dynamic mode decomposition: robust and verified Koopmanism." Journal of Fluid Mechanics 955 (2023): A21.

[3] Colbrook, Matthew J., Qin Li, Ryan Raut, and Alex Townsend. "Beyond expectations: Residual Dynamic Mode Decomposition and Variance for Stochastic Dynamical Systems", arXiv preprint arXiv:2308.10697 (2023).

[4] Colbrook, Matthew J. "The mpEDMD algorithm for data-driven computations of measure-preserving dynamical systems." SIAM Journal on Numerical Analysis 61.3 (2023): 1585-1608.

[5] Colbrook, Matthew J., and Alex Townsend. "Avoiding discretization issues for nonlinear eigenvalue problems." arXiv preprint arXiv:2305.01691 (2023).

[6] Colbrook, Matthew J. "Computing semigroups with error control." SIAM Journal on Numerical Analysis 60.1 (2022): 396-422.

[7] Colbrook, Matthew J., and Lorna J. Ayton. "A contour method for time-fractional PDEs and an application to fractional viscoelastic beam equations." Journal of Computational Physics 454 (2022): 110995.

[8] Colbrook, Matthew. The foundations of infinite-dimensional spectral computations. Diss. University of Cambridge, 2020.

[9] Ben-Artzi, J., Colbrook, M. J., Hansen, A. C., Nevanlinna, O., & Seidel, M. (2020). Computing Spectra--On the Solvability Complexity Index Hierarchy and Towers of Algorithms. arXiv preprint arXiv:1508.03280.

[10] Colbrook, Matthew J., Vegard Antun, and Anders C. 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 119.12 (2022): e2107151119.

[11] Colbrook, Matthew, Andrew Horning, and Alex Townsend. "Computing spectral measures of self-adjoint operators." SIAM review 63.3 (2021): 489-524.

[12] Colbrook, Matthew J., Bogdan Roman, and Anders C. Hansen. "How to compute spectra with error control." Physical Review Letters 122.25 (2019): 250201.

[13] Colbrook, Matthew J., and Anders C. Hansen. "The foundations of spectral computations via the solvability complexity index hierarchy." Journal of the European Mathematical Society (2022).

[14] Colbrook, Matthew J. "Computing spectral measures and spectral types." Communications in Mathematical Physics 384 (2021): 433-501.

[15] Colbrook, Matthew J., and Anders C. Hansen. "On the infinite-dimensional QR algorithm." Numerische Mathematik 143 (2019): 17-83.

[16] Colbrook, Matthew J. "On the computation of geometric features of spectra of linear operators on Hilbert spaces." Foundations of Computational Mathematics (2022): 1-82.

[17] Colbrook, Matthew. "Pseudoergodic operators and periodic boundary conditions." Mathematics of Computation 89.322 (2020): 737-766.

[18] Johnstone, Dean, et al. "Bulk localized transport states in infinite and finite quasicrystals via magnetic aperiodicity." Physical Review B 106.4 (2022): 045149.

[19] Colbrook, Matthew J., et al. "Computing spectral properties of topological insulators without artificial truncation or supercell approximation." IMA Journal of Applied Mathematics 88.1 (2023): 1-42.

[20] Colbrook, Matthew J., and Andrew Horning. "Specsolve: spectral methods for spectral measures." arXiv preprint arXiv:2201.01314 (2022).

[21] Colbrook, Matthew, Andrew Horning, and Alex Townsend. "Resolvent-based techniques for computing the discrete and continuous spectrum of differential operators." XXI Householder Symposium on Numerical Linear Algebra. 2020.

[22] Brunton, Steven L., and Matthew J. Colbrook. "Resilient Data-driven Dynamical Systems with Koopman: An Infinite-dimensional Numerical Analysis Perspective."