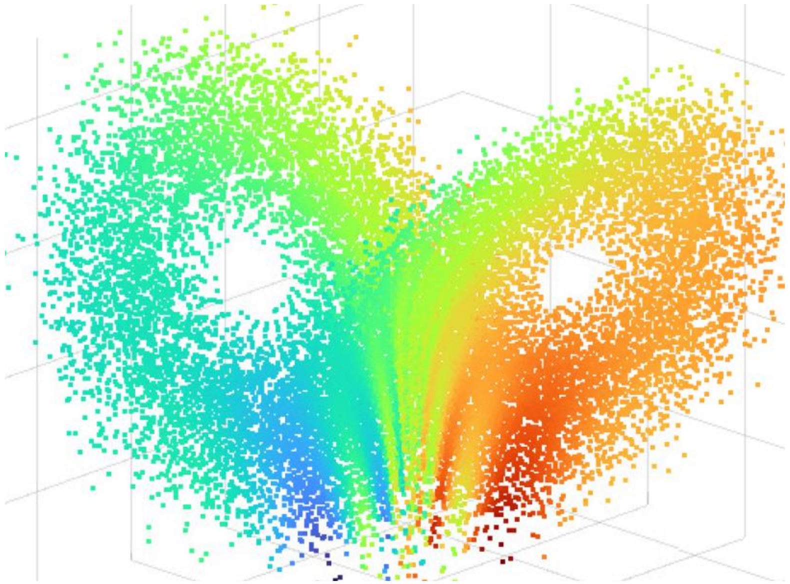
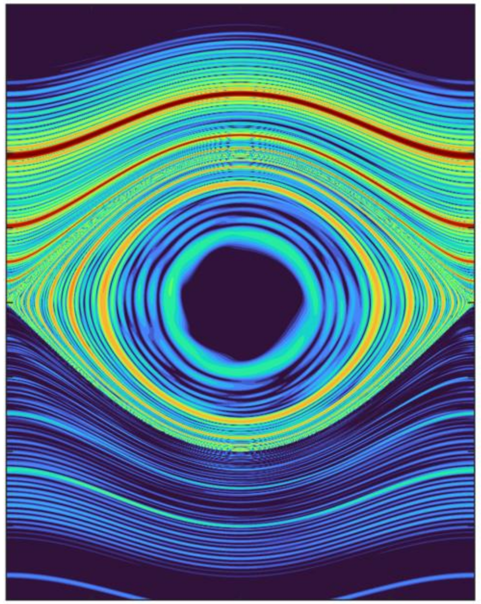
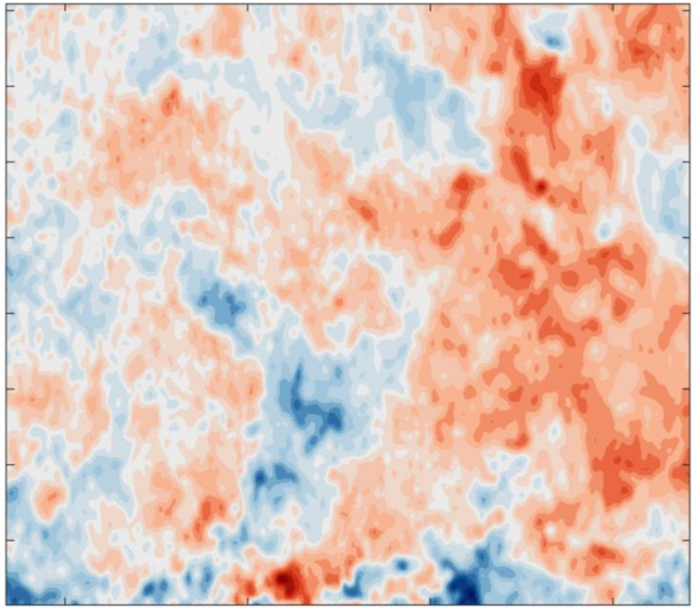
<u>m</u>easure-<u>p</u>reserving <u>E</u>xtended <u>D</u>ynamic <u>M</u>ode <u>D</u>ecomposition (**mpEDMD**):

Structure-preserving and rigorous Koopmanism

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







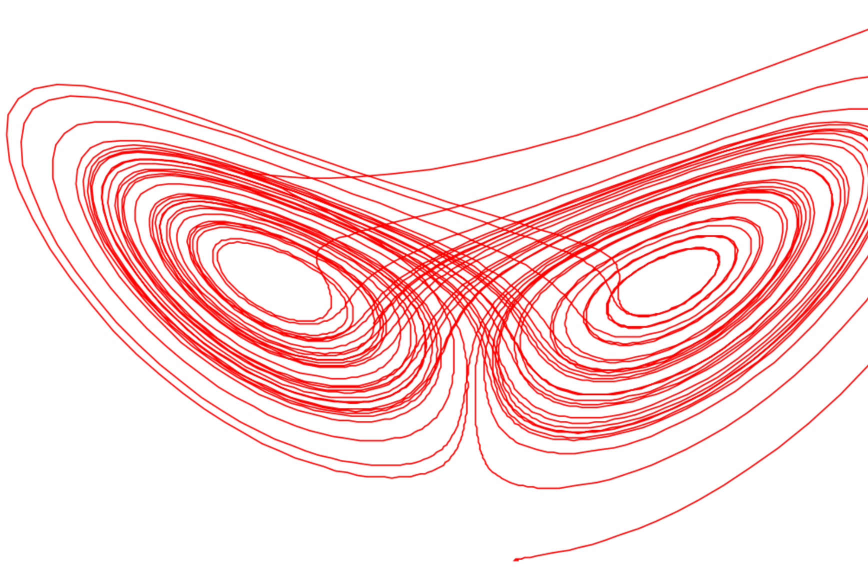
C., "*The mpEDMD Algorithm for Data-Driven Computations of Measure-Preserving Dynamical Systems,*" preprint.

Data-driven dynamical systems

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

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

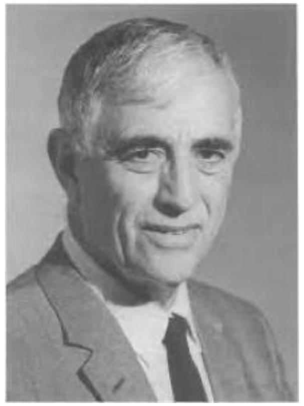
- **Goal:** Learn about system from data $\left\{x^{(m)}, y^{(m)} = F\left(x^{(m)}\right)\right\}_{m=1}^{M}$
  - **Data:** experimental measurements or numerical simulations
  - E.g., **used for** forecasting, control, design, understanding
- **Applications:** chemistry, climatology, electronics, epidemiology,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



Poincaré

Operator viewpoint

Koopman

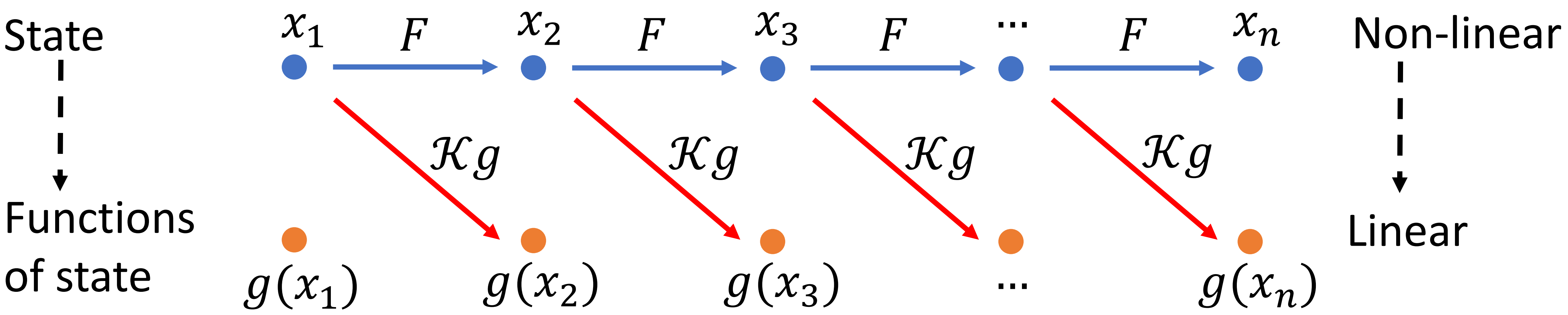


von Neumann

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

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

- $\mathcal{K}$ 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,"* **Proc. Natl. Acad. Sci. USA**, 1931.
- Koopman, v. Neumann, *"Dynamical systems of continuous spectra,"* **Proc. Natl. Acad. Sci. USA**, 1932.

Why is linear (much) easier?

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

Long-time dynamics become trivial!

- **Baby case:** $F(x) = Ax, A \in \mathbb{R}^{d\times d}$, $A = V\Lambda V^{-1}$.
- 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** F and even **chaotic** systems).

Koopman mode decomposition

generalized eigenfunction of $\mathcal{K}$

eigenfunction of $\mathcal{K}$

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

$$g(x) = \sum_{\text{eigs } \lambda_j} c_{\lambda_j} \varphi_{\lambda_j}(x) + \int_{[-\pi,\pi]_{\text{per}}} \phi_{\theta,g}(x)\,\mathrm{d}\theta$$

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

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

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

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

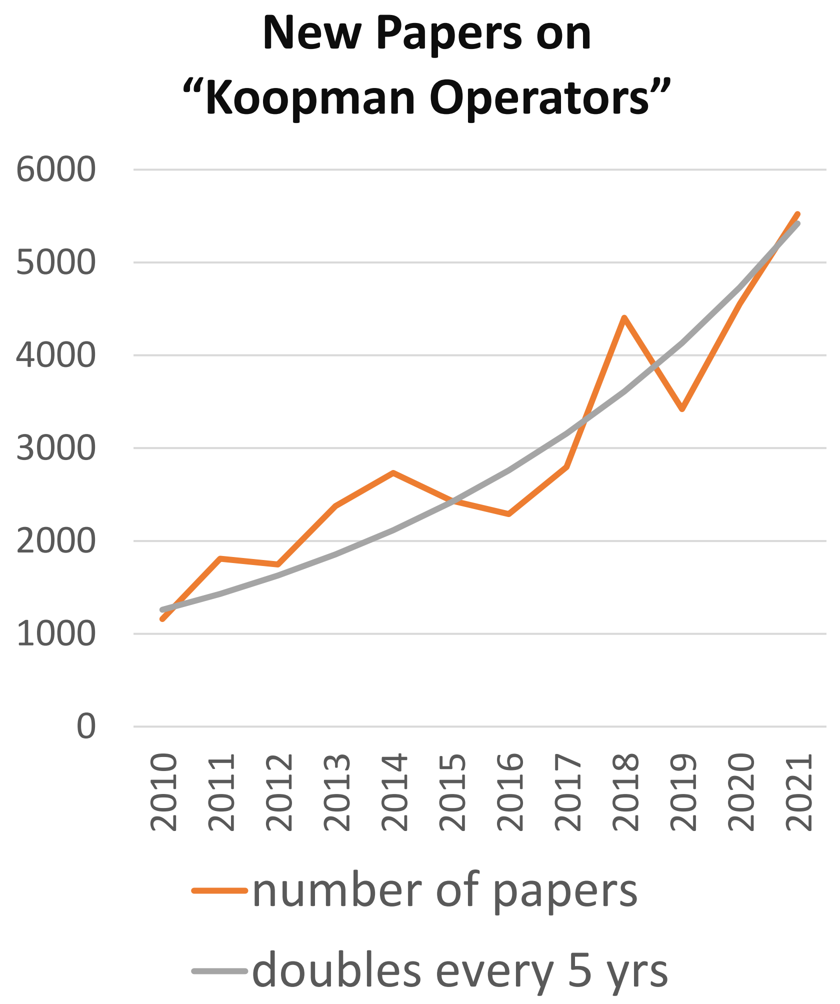
Koopmania*: A revolution in the big data era?

≈35,000 papers over last decade!

***BUT: Very* little on convergence guarantees or verification!**

Why is this lacking?

- Koopmanism has been (largely) distinct from NA.
- Computing spectra in infinite dimensions is notoriously hard …



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

Challenges

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

1) **"Too much":** Approximate spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$

2) **"Too little":** Miss parts of $\mathrm{Spec}(\mathcal{K})$

3) **Continuous spectra.**

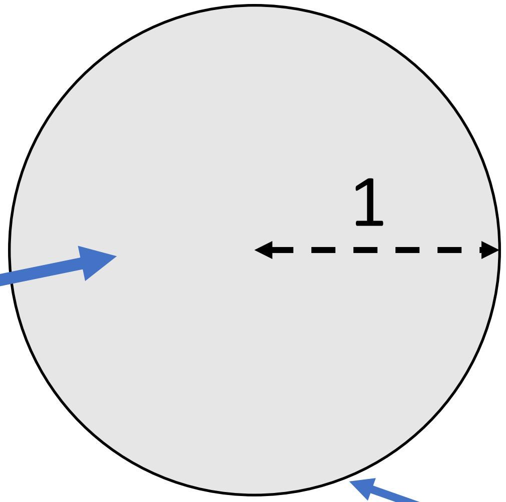
*"In practice, most operators are not presented in a representation in which they are diagonalized. Thus, one often has to settle for numerical approximations. Unfortunately, there is a dearth of literature on this basic problem and, so far as we have been able to tell, **there are no proven [general] techniques**."* W. Arveson, Berkeley (1994)

Assumption

System is measure-preserving (e.g., Hamiltonian, ergodic, post-transient etc.)

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

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



spectral measure (see later) supp. on boundary

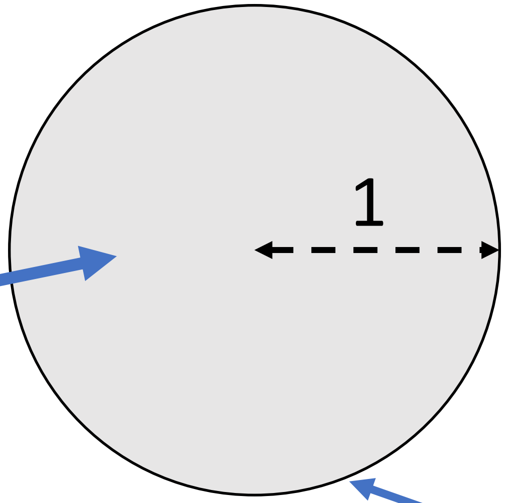
(NB: we consider unitary extensions via Wold decomposition.)

Assumption

System is measure-preserving (e.g., Hamiltonian, ergodic, post-transient etc.)

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

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

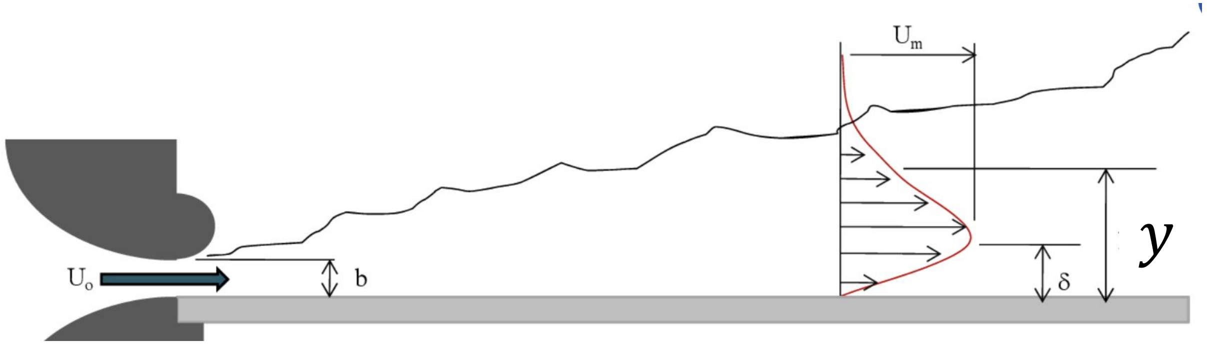


spectral measure (see later) supp. on boundary

(NB: we consider unitary extensions via Wold decomposition.)

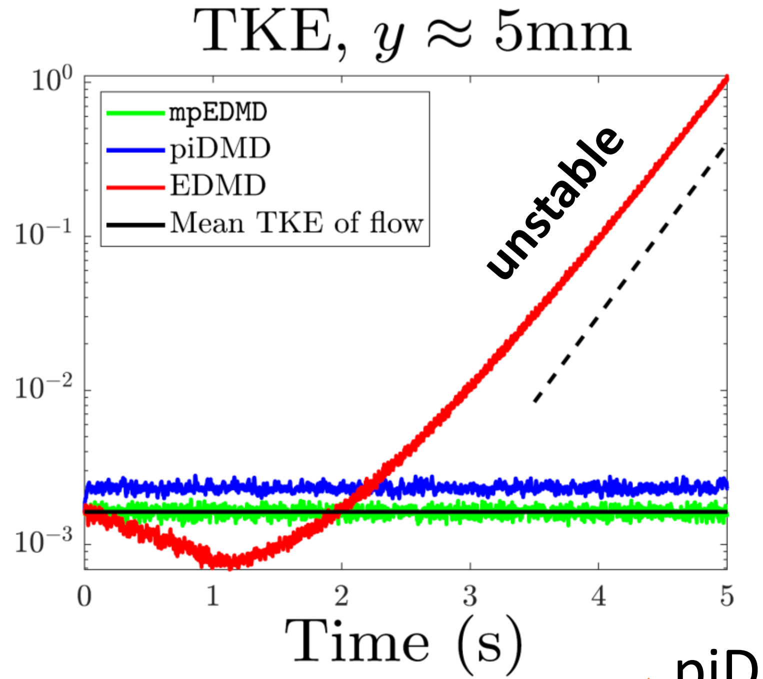
WANT: Method preserves measure (e.g., stability, long-time behavior etc.)...

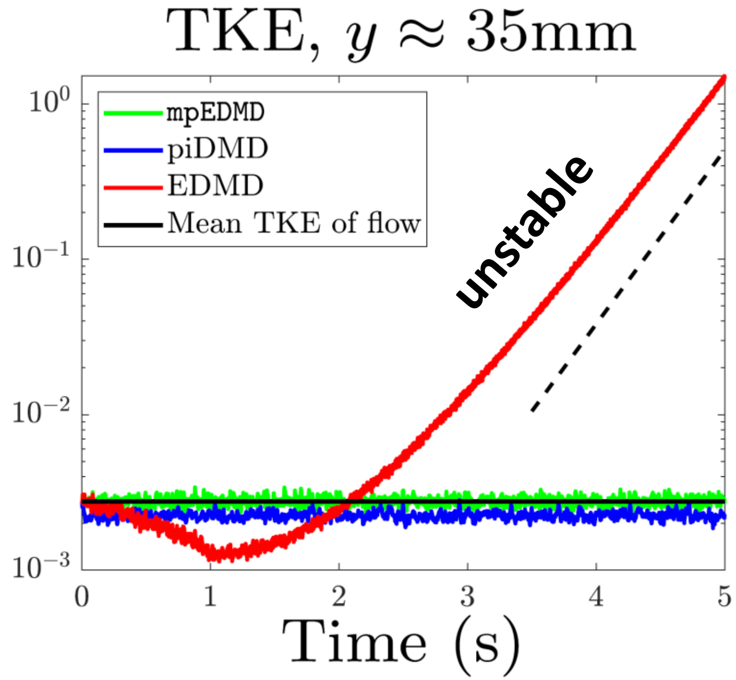
Motivating example

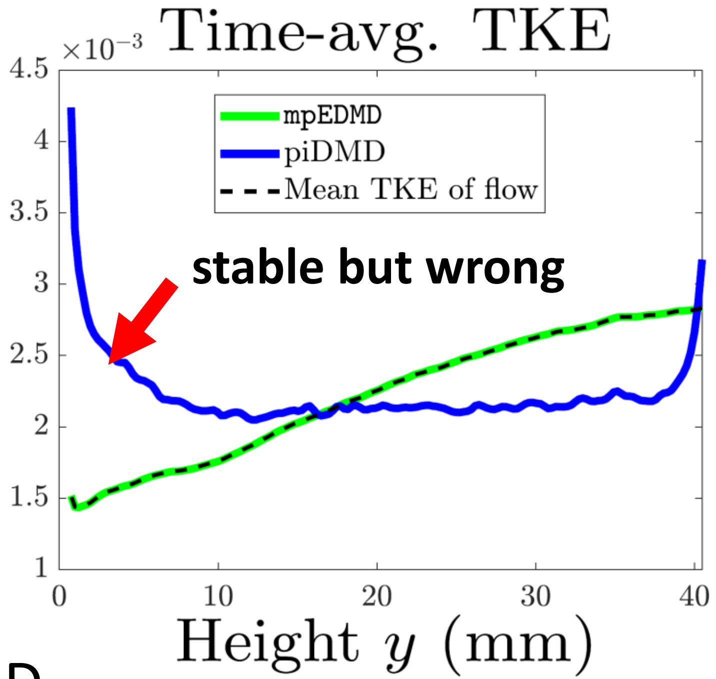


- Reynolds number $\approx 6.4 \times 10^4$
- Ambient dimension $(d) \approx 100{,}000$ (velocity at measurement points)

*Raw measurements provided by Máté Szőke (Virginia Tech)







piDMD

EDMD

- Baddoo, Herrmann, McKeon, Kutz, Brunton, *"Physics-informed dynamic mode decomposition (piDMD),"* preprint.
- Williams, Kevrekidis, Rowley *"A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition,"* **J. Nonlinear Sci.**, 2015.

The mpEDMD algorithm

Extended Dynamic Mode Decomposition (EDMD)

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

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

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

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

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

- 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.

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

Another interpretation

$$\Psi(x) = [\psi_1(x) \;\; \dots \;\; \psi_N(x)], \qquad g = \sum_{j=1}^{N} \mathbf{g}_j \psi_j = \Psi \mathbf{g}, \qquad \mathcal{K} g = \Psi \mathbb{K} \mathbf{g} + R(g,\cdot)$$

$$\min_{\mathbb{K} \in \mathbb{C}^{N \times N}} \left\{ \int_\Omega \max_{\|\mathbf{g}\|_2 = 1} |R(g,x)|^2 \, d\omega(x) = \int_\Omega \|\Psi(F(x)) - \Psi(x)\mathbb{K}\|_2^2 \, d\omega(x) \right\}$$

$$\Big\downarrow \text{ quadrature}$$

$$\min_{\mathbb{K} \in \mathbb{C}^{N \times N}} \sum_{m=1}^{M} w_m \left\| \Psi\left(y^{(m)}\right) - \Psi\left(x^{(m)}\right) \mathbb{K} \right\|_2^2$$

A simple idea

$$G = \Psi_X{}^* W \Psi_X, \qquad G_{jk} \approx \langle \psi_k, \psi_j \rangle$$

Measure-preserving: $\|\Psi\mathbf{g}\| = \|\Psi\mathbb{K}\mathbf{g}\|$, $\|\Psi\mathbf{g}\|^2 \approx g^* G g$, $\|\Psi\mathbb{K}\mathbf{g}\|^2 \approx g^* \mathbb{K}^* G \mathbb{K} g$

A simple idea

$$G = \Psi_X{}^* W \Psi_X, \qquad G_{jk} \approx \langle \psi_k, \psi_j \rangle$$

Measure-preserving: $||\Psi\mathbf{g}|| = ||\Psi\mathbb{K}\mathbf{g}||$, $||\Psi\mathbf{g}||^2 \approx g^* G g$, $||\Psi\mathbb{K}\mathbf{g}||^2 \approx g^* \mathbb{K}^* G \mathbb{K} g$

Enforce: $G = \mathbb{K}^* G \mathbb{K}$

A simple idea

$$G = \Psi_X{}^* W \Psi_X, \qquad G_{jk} \approx \langle \psi_k, \psi_j \rangle$$

Measure-preserving: $\|\Psi\mathbf{g}\| = \|\Psi\mathbb{K}\mathbf{g}\|$, $\|\Psi\mathbf{g}\|^2 \approx g^* G g$, $\|\Psi\mathbb{K}\mathbf{g}\|^2 \approx g^* \mathbb{K}^* G \mathbb{K} g$

Enforce: $G = \mathbb{K}^* G \mathbb{K}$

$$\Big\downarrow \text{ quadrature}$$

$$\min_{\substack{\mathbb{K}\in\mathbb{C}^{N\times N} \\ G=\mathbb{K}^* G \mathbb{K}}} \sum_{m=1}^{M} w_m \left\| \Psi\left(y^{(m)}\right) G^{-1/2} - \Psi\left(x^{(m)}\right) \mathbb{K} G^{-1/2} \right\|_2^2$$

orthogonal Procrustes problem

A simple algorithm

Algorithm: `mpEDMD` for approximating spectral properties of $\mathcal{K}$.

Input: Snapshot data $\{\boldsymbol{x}^{(m)}, \boldsymbol{y}^{(m)} = F(\boldsymbol{x}^{(m)})\}_{m=1}^{M}$, quadrature weights $\{w_m\}_{m=1}^{M}$, and a dictionary of functions $\{\psi_j\}_{j=1}^{N}$.

1: Compute $G = \Psi_X^* W \Psi_X$ and $A = \Psi_X^* W \Psi_Y$
2: Compute an SVD of $G^{-1/2} A^* G^{-1/2} = U_1 \Sigma U_2^*$
3: Compute the eigendecomposition $U_2 U_1^* = \hat{V} \Lambda \hat{V}^*$
4: Compute $\mathbb{K} = G^{-1/2} U_2 U_1^* G^{1/2}$ and $V = G^{-1/2} \hat{V}$

Output: Koopman matrix $\mathbb{K}$, with eigenvectors V and eigenvalues Λ.

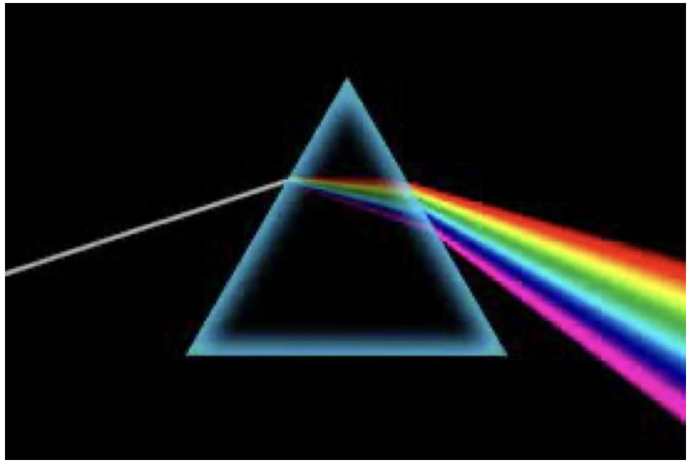
In a nutshell: *Galerkin meets polar decomposition*.

(This also allows us to prove <u>numerical stability</u>.)

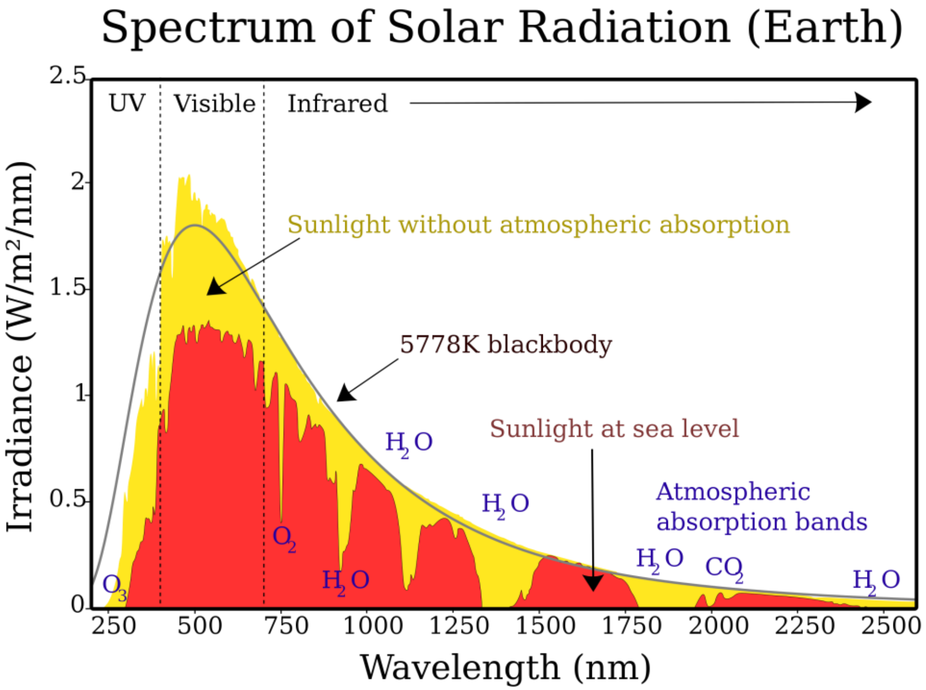
Convergence theory

Spectral measures

White light contains a continuous spectra



Often interesting to look at
the intensity of each wavelength



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 $\mathcal{E}$

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

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

Koopman mode decomposition (again!)

μ_g probability measures on $\mathbb{T}$

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

Koopman mode decomposition (again!)

μ_g probability measures on $\mathbb{T}$

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

eigenfunction of $\mathcal{K}$

generalized eigenfunction of $\mathcal{K}$

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

Koopman mode decomposition (again!)

μ_g probability measures on $\mathbb{T}$

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

eigenfunction of $\mathcal{K}$

generalized eigenfunction of $\mathcal{K}$

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

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

Koopman mode decomposition (again!)

μ_g probability measures on $\mathbb{T}$

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

eigenfunction of $\mathcal{K}$

generalized eigenfunction of $\mathcal{K}$

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

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

Computing μ_g diagonalises non-linear dynamical system!

Convergence of projection-valued measure

$$\mathrm{d}\mathcal{E}_{N,M}(\lambda) = \sum_{j=1}^{N} v_j\, v_j^* G\delta(\lambda - \lambda_j)\mathrm{d}\lambda$$

This assumption cannot be dropped in general!

Theorem: Suppose that the quadrature rule converges, $\mathcal{K}$ is unitary, $\lim_{N\to\infty} \mathrm{dist}(h, V_N) = 0$ for any $h \in L^2(\Omega,\omega)$. Then for any Lipschitz test function ξ, $g \in L^2(\Omega,\omega)$ and $\boldsymbol{g}_N \in \mathbb{C}^N$ with $\lim_{N\to\infty} \|g - \Psi \boldsymbol{g}_N\| = 0$,

$$\lim_{N\to\infty} \limsup_{M\to\infty} \left\| \int_{\mathbb{T}} \xi(\lambda)\mathrm{d}\mathcal{E}(\lambda)g - \Psi \int_{\mathbb{T}} \xi(\lambda)\mathrm{d}\mathcal{E}_{N,M}(\lambda)\boldsymbol{g}_N \right\| = 0$$

Key ingredients:

- Strong convergence of Galerkin approximation.
- Discretization by normal operators.

$\mathbb{K}$: mpEDMD matrix

λ_j: eigenvalues of $\mathbb{K}$

v_j: eigenvectors of $\mathbb{K}$

$V_N = \mathrm{span}\,\{\psi_1, \dots, \psi_N\}$

Convergence of scalar-valued measure

$$\mu_{\boldsymbol{g}}^{(N,M)}(U) = \boldsymbol{g}^* G \mathcal{E}_{N,M}(U)\boldsymbol{g} = \sum_{\lambda_j \in U} \left| v_j^* G \boldsymbol{g} \right|^2$$

$$W_1(\mu,\nu) = \sup\left\{ \int_{\mathbb{T}} \xi(\lambda) \mathrm{d}(\mu-\nu)(\lambda) : \xi \text{ Lipschitz } 1 \right\}$$

Theorem: Suppose quad. rule converges, $\lim_{N\to\infty} \mathrm{dist}(h, V_N) = 0$ for any $h \in L^2(\Omega,\omega)$. Then for $g \in L^2(\Omega,\omega)$ and $\boldsymbol{g}_N \in \mathbb{C}^N$ with $\lim_{N\to\infty} \|g - \Psi \boldsymbol{g}_N\| = 0$,

$$\lim_{N\to\infty} \limsup_{M\to\infty} W_1\left(\mu_g, \mu_{\boldsymbol{g}}^{(N,M)}\right) = 0.$$

If $\{g, \mathcal{K}g, \dots, \mathcal{K}^L g\} \subseteq V_N$ and $g = \Psi \boldsymbol{g}$, then

Matching autocorrelations!

$$\limsup_{M\to\infty} W_1\left(\mu_g, \mu_{\boldsymbol{g}}^{(N,M)}\right) \lesssim \frac{\log(L)}{L}.$$

$\mathbb{K}$: mpEDMD matrix

λ_j: eigenvalues of $\mathbb{K}$

v_j: eigenvectors of $\mathbb{K}$

$V_N = \mathrm{span}\{\psi_1, \dots, \psi_N\}$

Spectra: avoid too little!

$$\mathrm{Spec_{ap}}(\mathcal{K}) = \left\{\lambda: \exists u_n, \|u_n\| = 1, \lim_{n\to\infty} \|(\mathcal{K}-\lambda)u_n\| = 0\right\} = \mathrm{Spec}(\mathcal{K}) \cap \mathbb{T}$$

Theorem: Suppose quad. rule converges, $\lim_{N\to\infty} \mathrm{dist}(h, V_N) = 0$ for any $h \in L^2(\Omega,\omega)$. Then

$$\lim_{N\to\infty} \limsup_{M\to\infty} \sup_{\lambda\in\mathrm{Spec_{ap}}(\mathcal{K})} \mathrm{dist}(\lambda, \mathrm{Spec}(\mathbb{K})) = 0.$$

$\mathbb{K}$: mpEDMD matrix
λ_j: eigenvalues of $\mathbb{K}$
v_j: eigenvectors of $\mathbb{K}$
$V_N = \mathrm{span}\,\{\psi_1, \dots, \psi_N\}$

Spectra: avoid too little!

$$\mathrm{Spec}_{\mathrm{ap}}(\mathcal{K}) = \left\{\lambda: \exists u_n, \|u_n\| = 1, \lim_{n\to\infty} \|(\mathcal{K}-\lambda)u_n\| = 0\right\} = \mathrm{Spec}(\mathcal{K}) \cap \mathbb{T}$$

Theorem: Suppose quad. rule converges, $\lim_{N\to\infty} \mathrm{dist}(h, V_N) = 0$ for any $h \in L^2(\Omega, \omega)$. Then

$$\lim_{N\to\infty} \limsup_{M\to\infty} \sup_{\lambda \in \mathrm{Spec}_{\mathrm{ap}}(\mathcal{K})} \mathrm{dist}(\lambda, \mathrm{Spec}(\mathbb{K})) = 0.$$

What about spectral pollution?

$\mathbb{K}$: mpEDMD matrix
λ_j: eigenvalues of $\mathbb{K}$
v_j: eigenvectors of $\mathbb{K}$
$V_N = \mathrm{span}\,\{\psi_1, \dots, \psi_N\}$

Residuals ⇒ avoid too much!

$$G = \Psi_X{}^* W \Psi_X, \; A = \Psi_X{}^* W \Psi_Y$$

$$\begin{aligned}\|(\mathcal{K}-\lambda)\Psi \boldsymbol{g}\|^2 &= \langle (\mathcal{K}-\lambda)\Psi \boldsymbol{g}, (\mathcal{K}-\lambda)\Psi \boldsymbol{g}\rangle \\ &= \lim_{M\to\infty} \boldsymbol{g}^*\left[(1+|\lambda|^2)G - \bar{\lambda}A - \lambda A^*\right]\boldsymbol{g}\end{aligned}$$

Suitable conditions $\Longrightarrow \lim_{N\to\infty} \min_{g\in V_N} \|(\mathcal{K}-\lambda)\Psi \boldsymbol{g}\| / \|g\| = \mathrm{dist}(\lambda, \mathrm{Spec}_{\mathrm{ap}}(\mathcal{K}))$

Two methods:

- Clean up procedure for tolerance ε.
- Local minimization algorithm converges to $\mathrm{Spec}_{\mathrm{ap}}(\mathcal{K})$.

- C., T., "*Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems*," preprint.
- C., Ayton, Szőke, "*Residual Dynamic Mode Decomposition*," **J. Fluid Mech.,** under minor rev.

Challenges

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

~~**1) “Too much”:** Approximate spurious modes $\lambda \notin \mathrm{Spec}(\mathcal{K})$~~ ✓

~~**2) “Too little”:** Miss parts of $\mathrm{Spec}(\mathcal{K})$~~ ✓

~~**3) Continuous spectra.**~~ ✓

“In practice, most o[illegible]rs are not presented in a repres[illegible]n in which they are diagonalized. Thus, one ofte[illegible]pproximations. Unfortunately, there is a dearth of literat[illegible]pr[illegible]far as we have been able to tell, ***there are no pro[illegible] [general] techniques****.”* W[illegible]son, Berkeley (1994)

Executive summary

	DMD	EDMD	piDMD	mpEDMD
Aux. SVD matrices	n/a	n/a	$YX^* = V_1 S V_2^*$	$G^{-\frac{1}{2}} A^* G^{-\frac{1}{2}} = U_1 \Sigma U_2^*$
Koopman matrix	$(YX^\dagger)^\top$	$G^\dagger A$	$V_2 V_1^*$	$G^{-\frac{1}{2}} U_2 U_1^* G^{\frac{1}{2}}$
Nonlinear dictionary	✗	✓	✗	✓
Conv. spec. meas.	✗	✗	✗	✓
Conv. spectra	✗	✗	✗	✓
Conv. KMD	✗	✓	✗	✓
Measure-preserving	✗	✗	✗/✓‡	✓

$X = [x^{(1)} \dots x^{(M)}], Y = [y^{(1)} \dots y^{(M)}] \in \mathbb{C}^{d\times M}$ are matrices of the snapshots (linear dictionary), common to combine DMD with truncated SVD. $G = \Psi_X{}^* W \Psi_X$, $A = \Psi_X{}^* W \Psi_Y$.
NB: piDMD is measure-preserving only if XX^* and W are multiples of the identity.

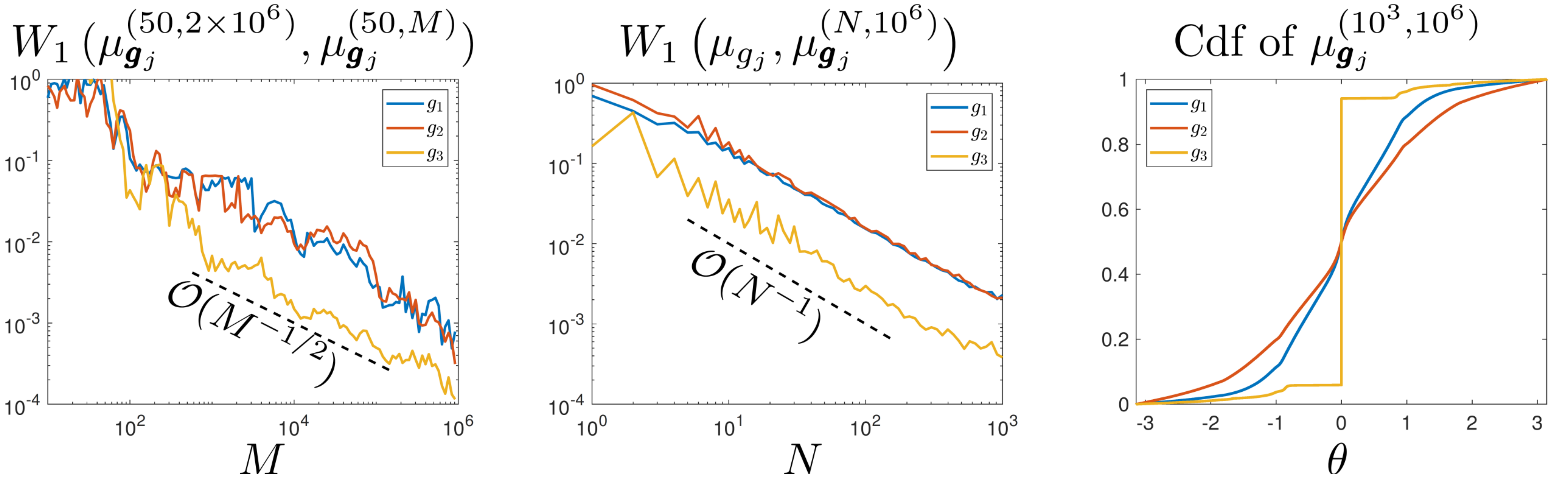
$\boxed{\textbf{mpEDMD: First convergent Galerkin method!}}$

Numerical examples

Lorenz system: scalar-valued spec meas

$$\dot{x}_1 = 10(x_2 - x_1), \qquad \dot{x}_2 = x_1(28 - x_3) - x_2, \qquad \dot{x}_3 = x_1 x_2 - \frac{8}{3}x_3, \qquad \Delta_t = 0.1$$

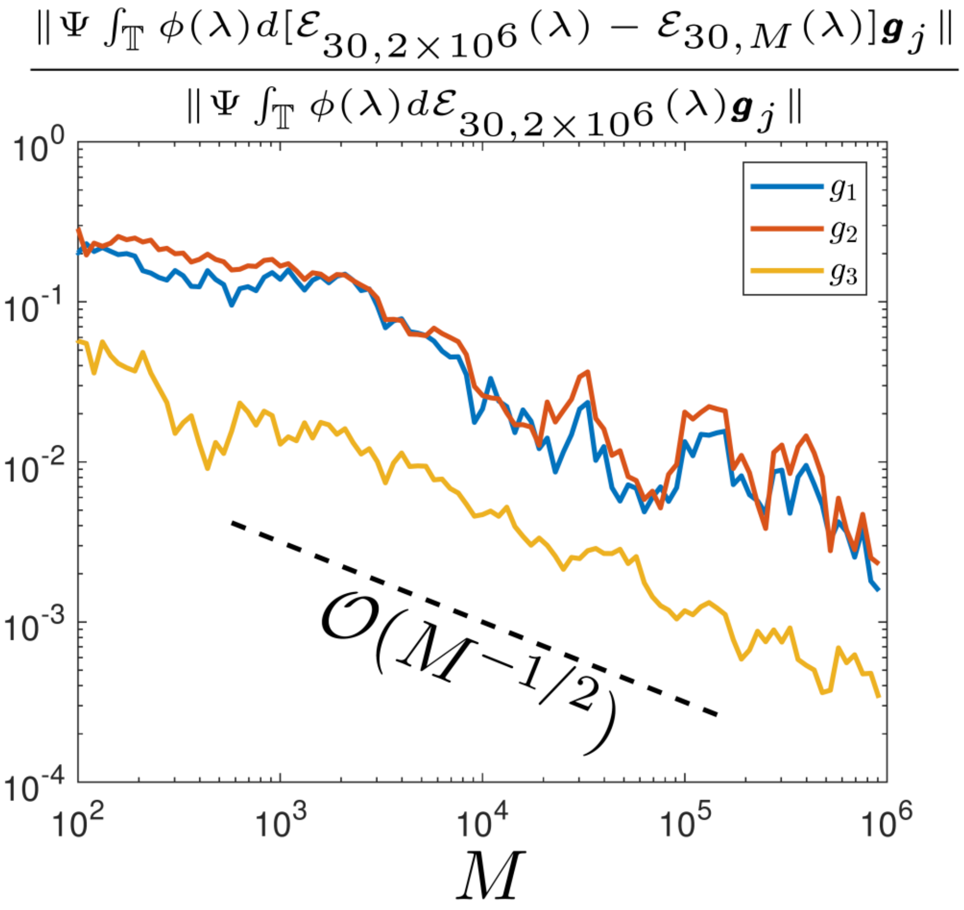
$$g_j = c_j[x]_j, \qquad V_N = \mathrm{span}\{g_j, \mathcal{K}g_j, \dots, \mathcal{K}^{N-1}g_j\}$$

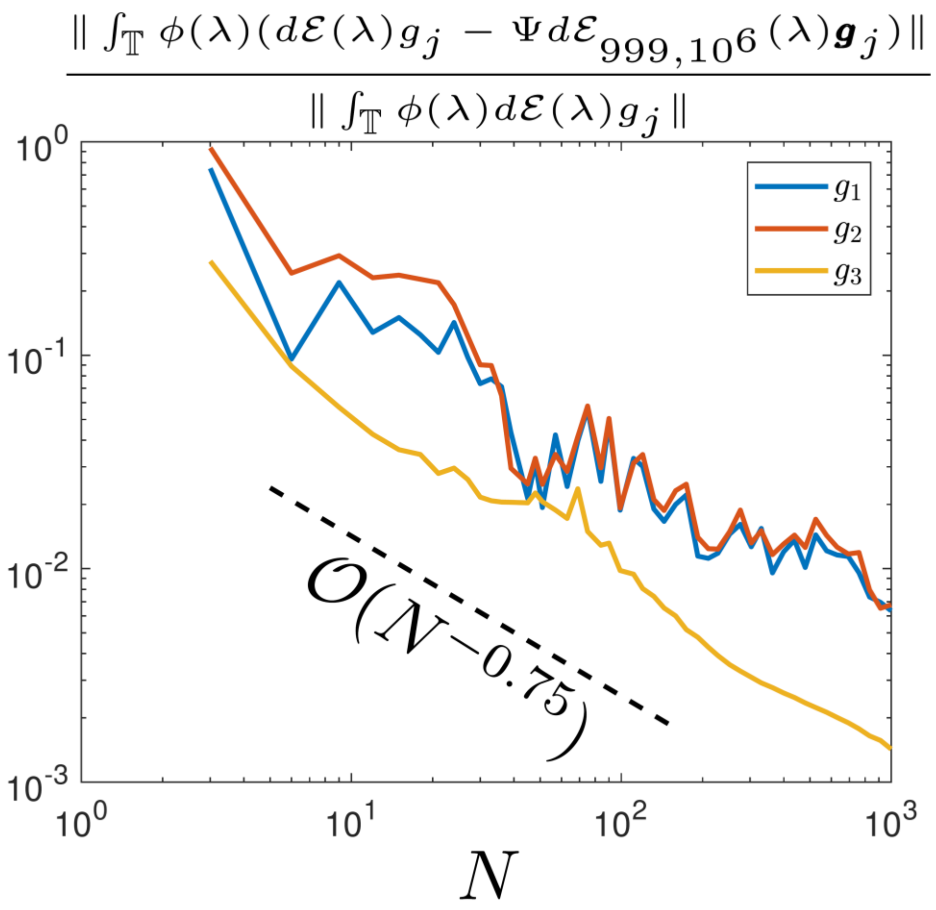


Lorenz system: projection-valued spec meas

$$V_N = \mathrm{span}\{g_1, g_2, g_3, \mathcal{K}g_1, \mathcal{K}g_2, \mathcal{K}g_3, \ldots, \mathcal{K}^{q-1}g_1, \mathcal{K}^{q-1}g_2, \mathcal{K}^{q-1}g_3\}$$

$$\phi(\theta) = \exp(\sin(\theta))$$

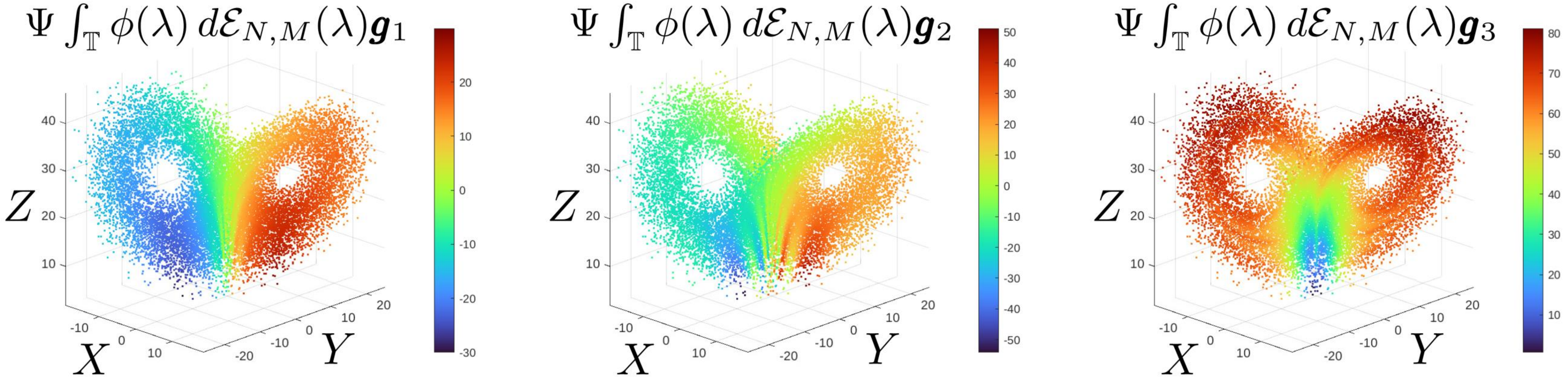




Lorenz system: projection-valued spec meas

$$V_N = \mathrm{span}\{g_1, g_2, g_3, \mathcal{K}g_1, \mathcal{K}g_2, \mathcal{K}g_3, \ldots, \mathcal{K}^{q-1}g_1, \mathcal{K}^{q-1}g_2, \mathcal{K}^{q-1}g_3\}$$

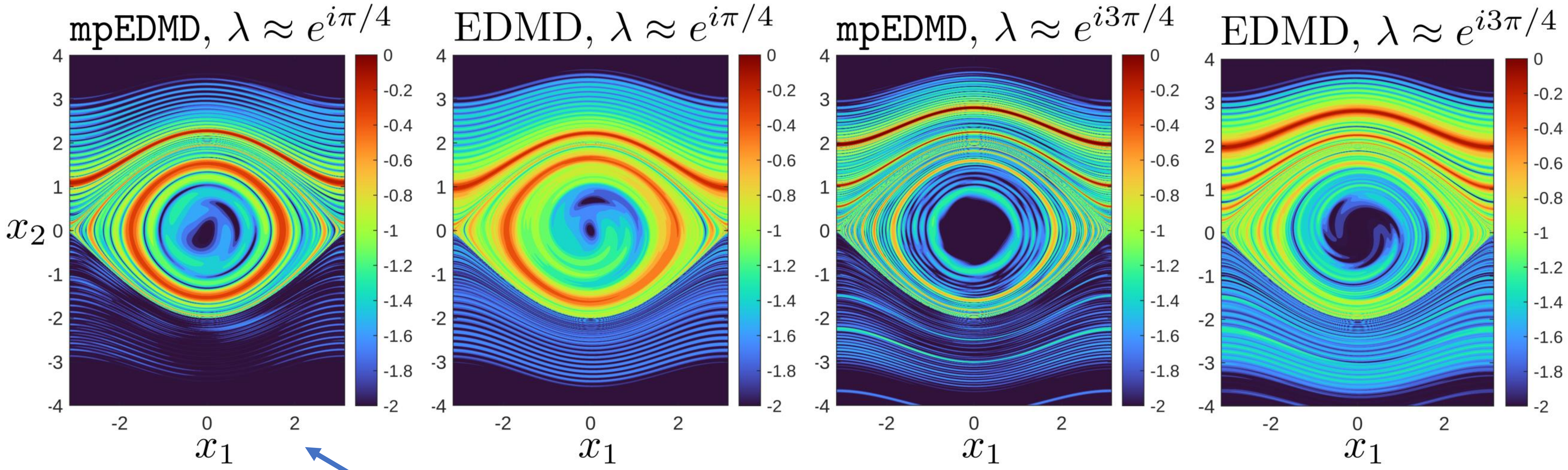
$$\phi(\theta) = \exp(\sin(\theta))$$



Nonlinear pendulum

$$\dot{x}_1 = x_2, \qquad \dot{x}_2 = -\sin(x_1), \qquad \Omega = [-\pi,\pi]_{\mathrm{per}} \times \mathbb{R}, \qquad \Delta_t = 0.5$$

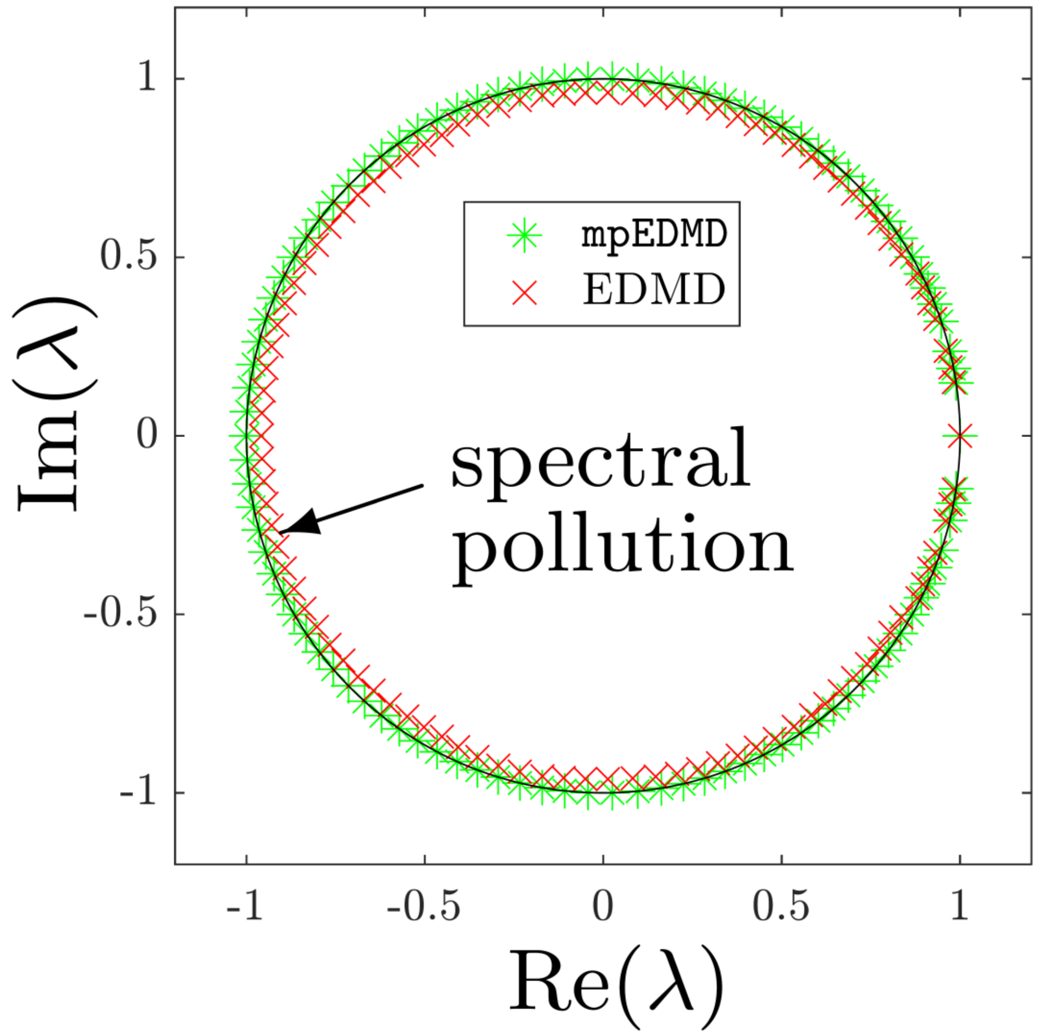
$$g(x) = \exp(i x_1)\, x_2 \exp(-x_2^2/2), \qquad V_N = \mathrm{span}\{g, \mathcal{K}g, \dots, \mathcal{K}^{99}g\}$$



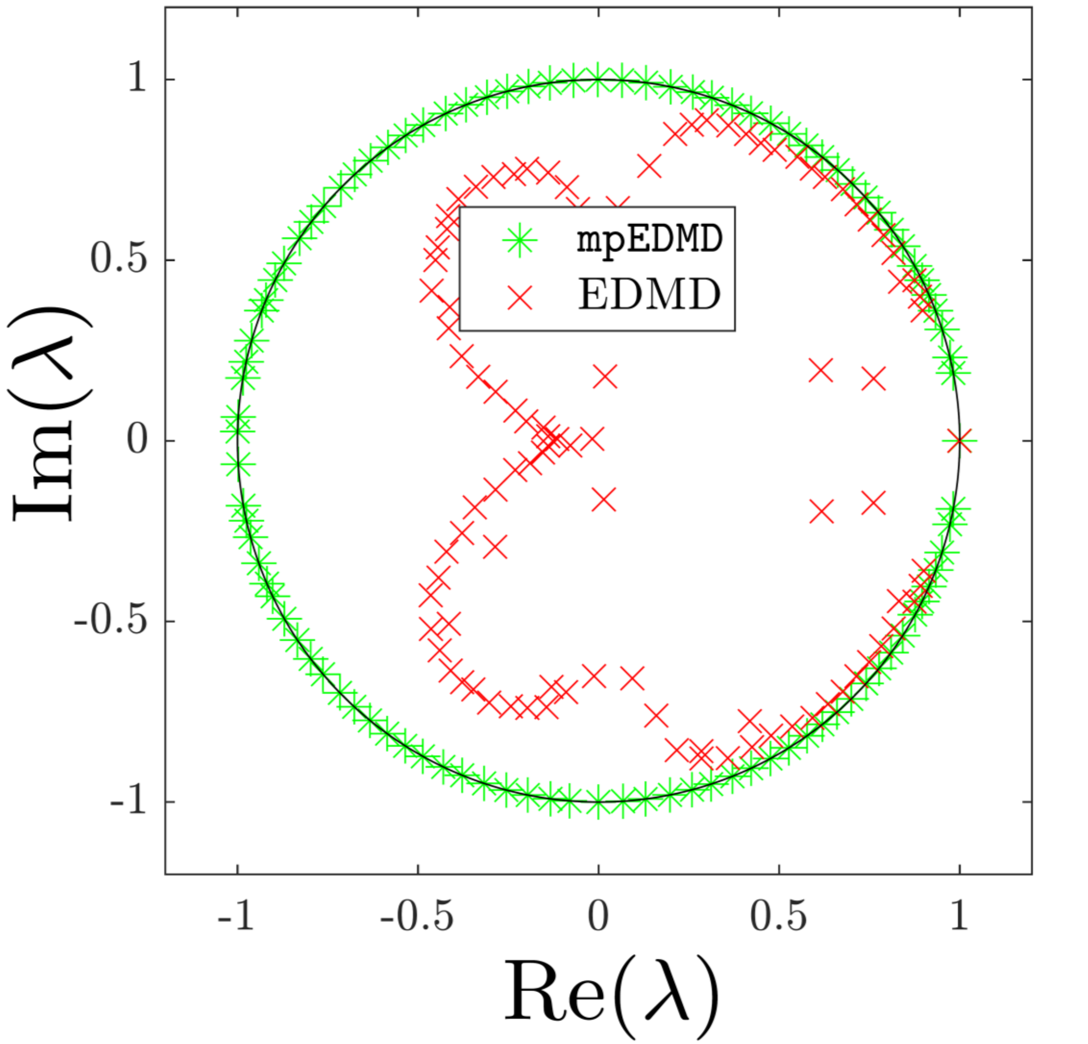
$\log_{10}(|v_j|)$

Nonlinear pendulum

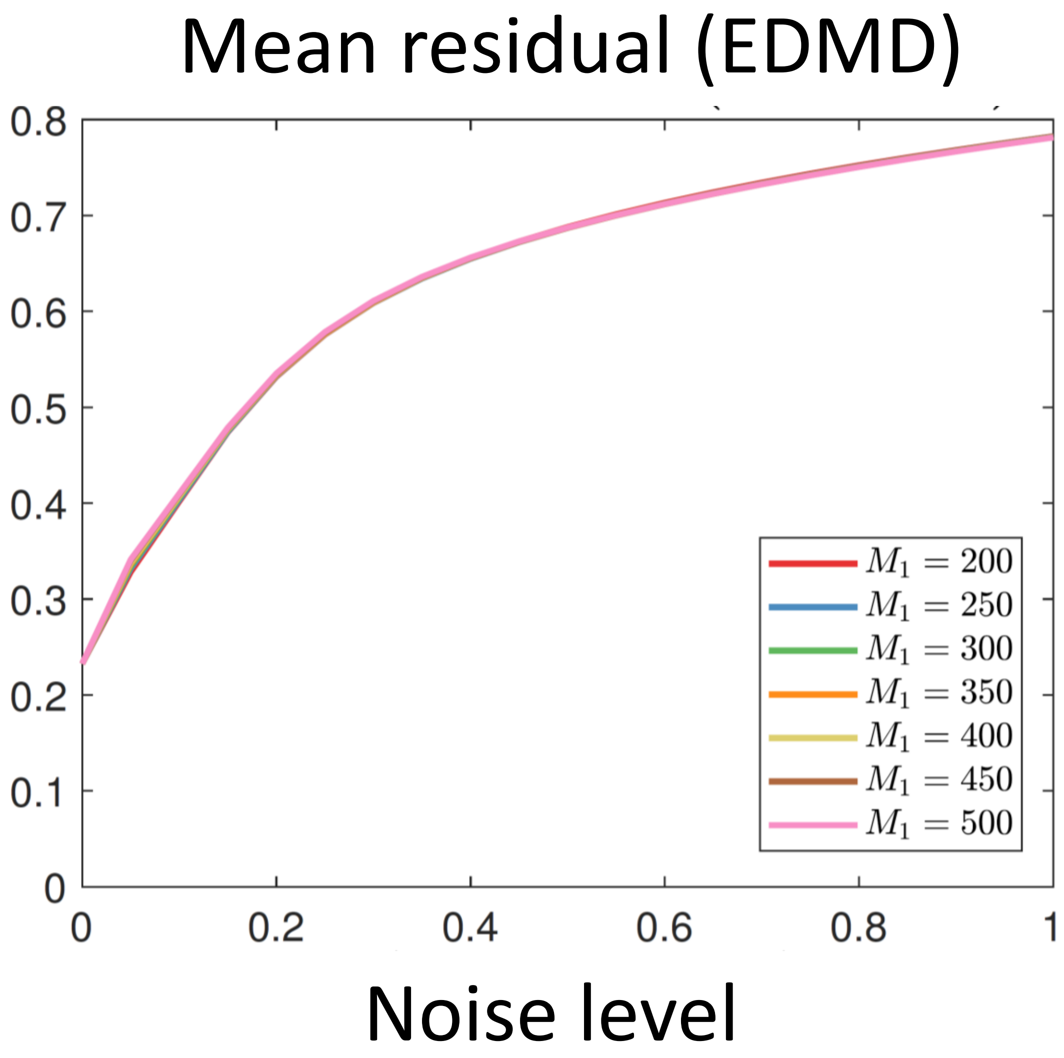
Noise free

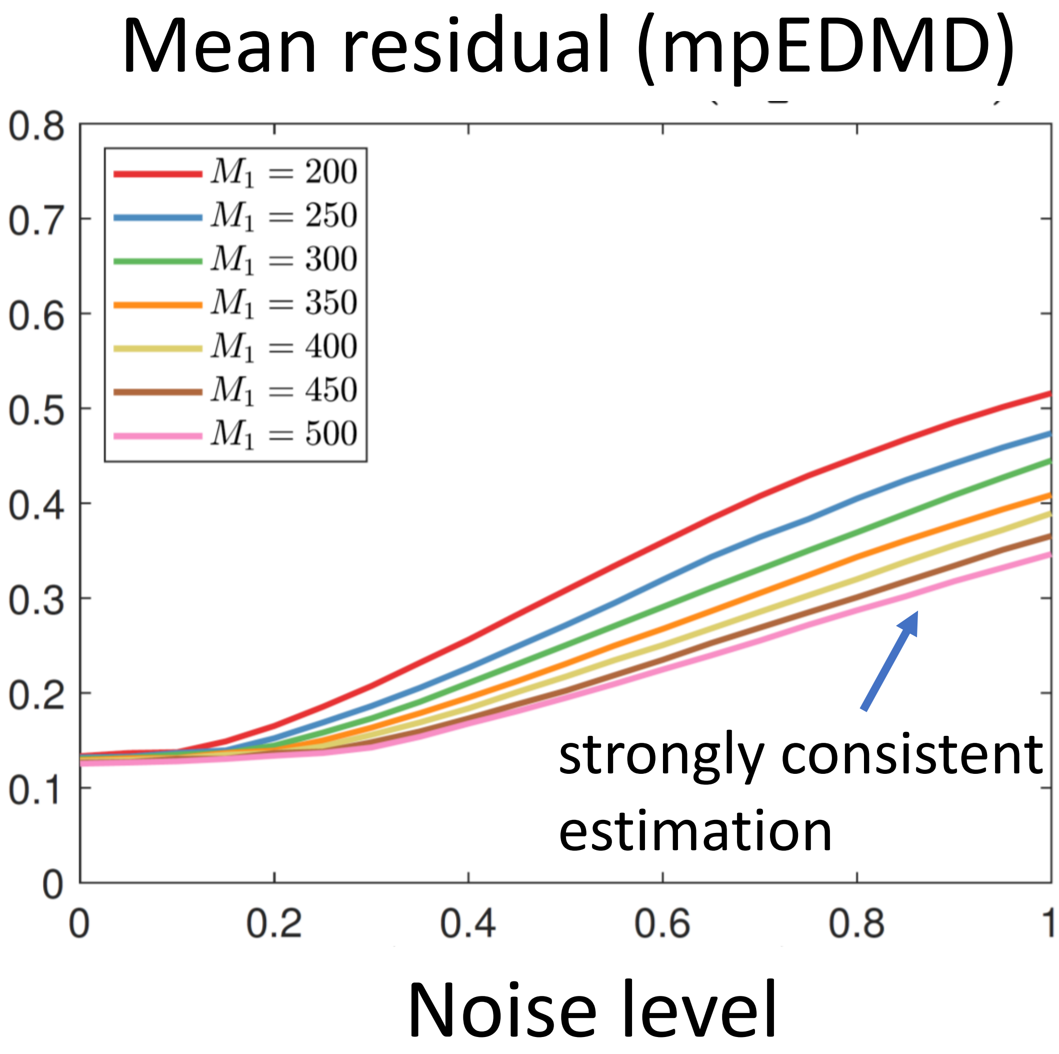


10% Gauss. noise for Ψ_X, Ψ_Y

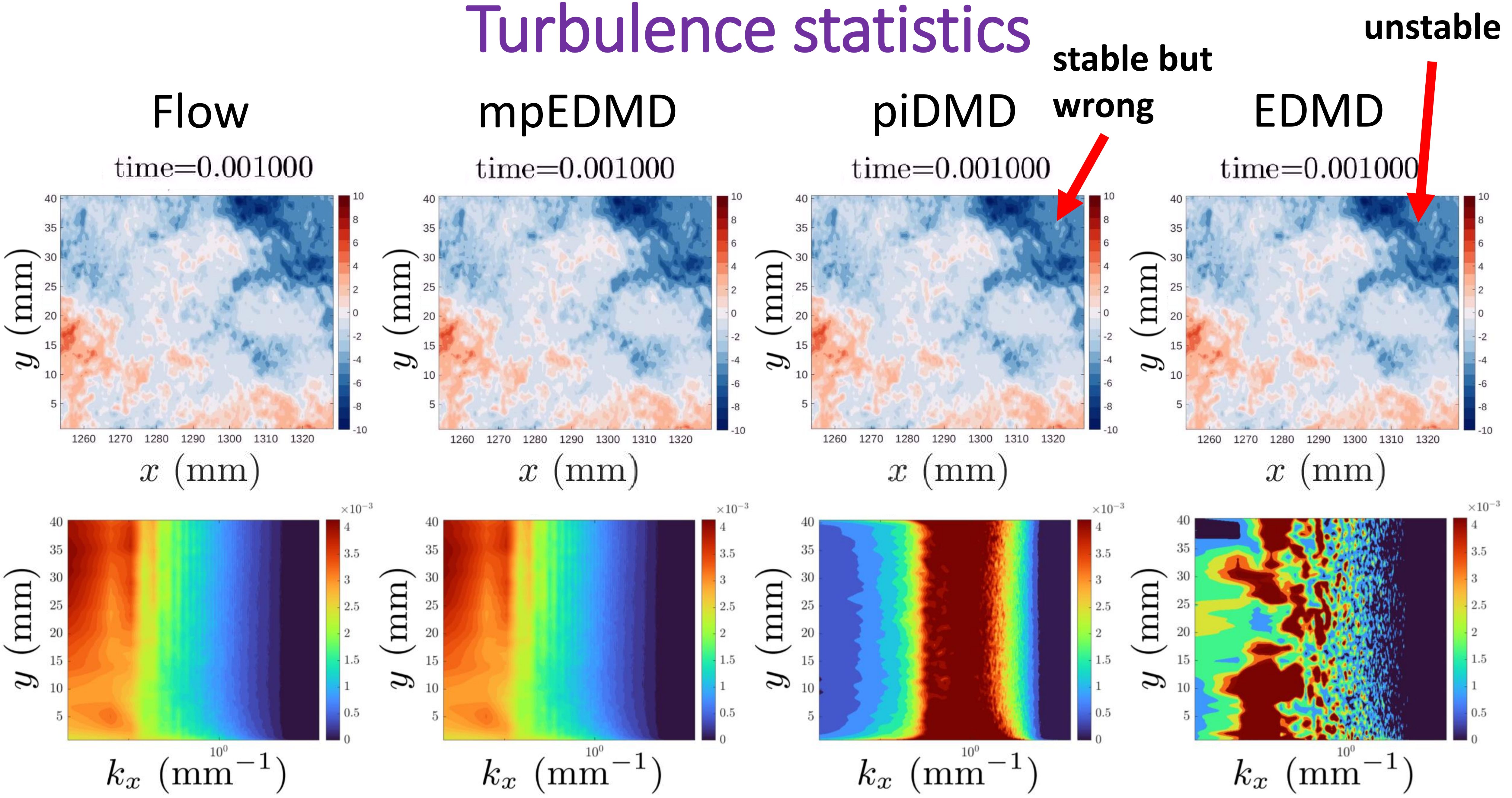


Nonlinear pendulum





Turbulence statistics



Wider programme

- Inf.-dim. computational analysis $\Longrightarrow$ **Compute spectral properties rigorously.**
- Continuous linear algebra $\Longrightarrow$ **Avoid the woes of discretization**
- Solvability Complexity Index hierarchy $\Longrightarrow$ **Classify diff. of comp. problems, prove algs are optimal.**
- **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.**, to appear.
- C., "*Computing spectral measures and spectral types,*" **Comm. Math. Phys.,** 2021.
- C., Horning, Townsend "*Computing spectral measures of self-adjoint operators,*" **SIAM Rev.**, 2021.
- C., Roman, Hansen, "*How to compute spectra with error control,*" **Phys. Rev. Lett.**, 2019.
- 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: On the barriers of deep learning and Smale's 18th problem,*" **Proc. Natl. Acad. Sci. USA**, 2022.
- C., "*Computing semigroups with error control,*" **SIAM J. Numer. Anal.**, 2022.
- 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,*" **Bull. Amer. Math. Soc.**, 1981, *36 pp.*
- McMullen, "*Families of rational maps and iterative root-finding algorithms,*" **Ann. of Math.**, 1987, *27 pp.*

Summary

Structure-preserving Koopmanism for arbitrary measure-preserving systems.

Some advantages:

- Converges for spectral measures, spectra, Koopman mode decomposition.
- Long-time stability and improved qualitative behavior.
- Increased stability to noise (e.g., measurements).
- Easy to combine with any DMD-type method!

What other areas of NA can successfully be combined with Koopmanism?

Code: https://github.com/MColbrook/Measure-preserving-Extended-Dynamic-Mode-Decomposition

Additional slides…

Solvability Complexity Index Hierarchy

Class $\Omega \ni A$, want to compute $\Xi\colon \Omega \to (\mathcal{M}, d)$ metric space

- Δ_0: Problems solved in finite time (v. rare for cts problems).
- Δ_1: Problems solved in "one limit" with full error control:

$$d(\Gamma_n(A), \Xi(A)) \leq 2^{-n}$$

- Δ_2: Problems solved in "one limit":

$$\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$$

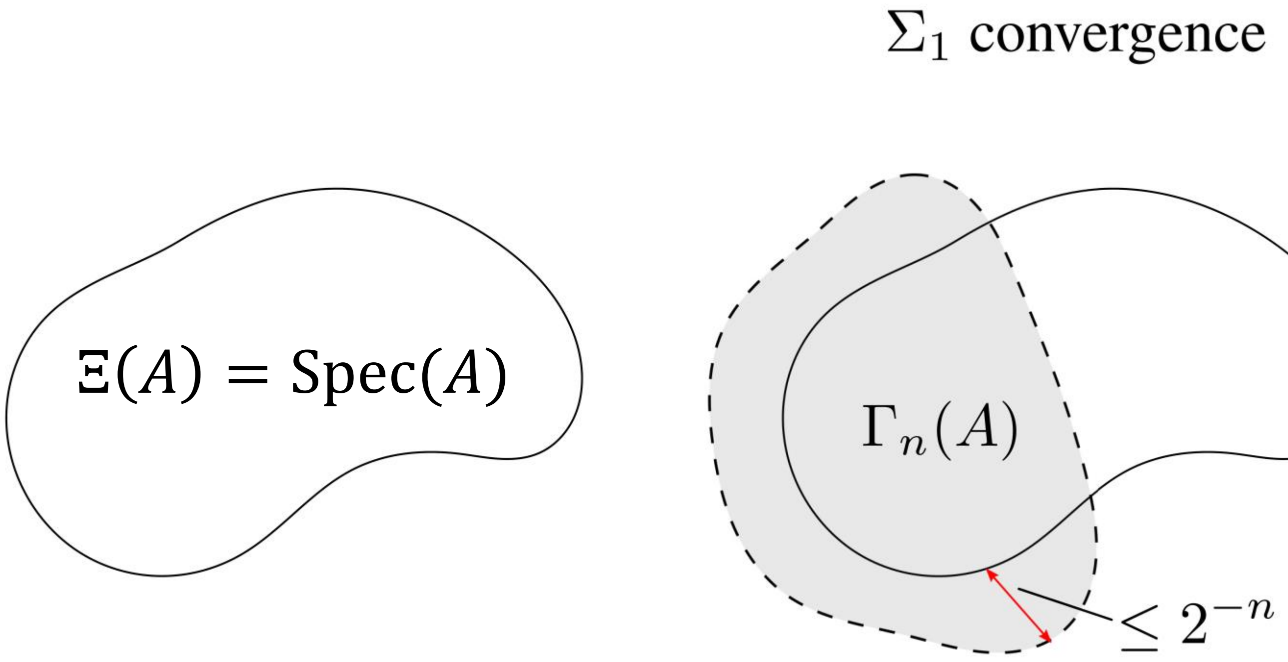
- Δ_3: Problems solved in "two successive limits":

$$\lim_{n\to\infty} \lim_{m\to\infty} \Gamma_{n,m}(A) = \Xi(A)$$

$\vdots$

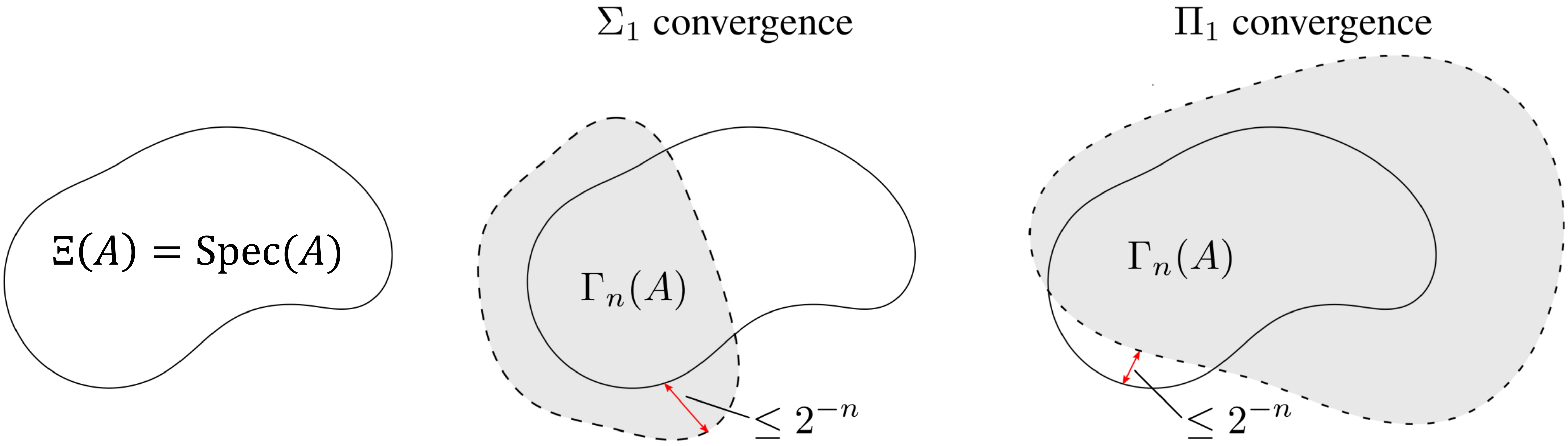
- Ben-Artzi, C., Hansen, Nevanlinna, Seidel, "*On the solvability complexity index hierarchy and towers of algorithms,*" preprint.
- Hansen, "*On the solvability complexity index, the n-pseudospectrum and approximations of spectra of operators,*" **J. Amer. Math. Soc.**, 2011.
- McMullen, "*Families of rational maps and iterative root-finding algorithms,*" **Ann. of Math.**, 1987.
- Doyle, McMullen, "*Solving the quintic by iteration,*" **Acta Math.**, 1989.
- Smale, "*The fundamental theorem of algebra and complexity theory,*" **Bull. Amer. Math. Soc.**, 1981.

Error control for spectral problems



- Σ_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A)$, $\max_{z\in\Gamma_n(A)} \mathrm{dist}\big(z, \Xi(A)\big) \leq 2^{-n}$

Error control for spectral problems



- Σ_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A),\ \max_{z\in\Gamma_n(A)} \mathrm{dist}\big(z, \Xi(A)\big) \leq 2^{-n}$
- Π_1: $\exists$ alg. $\{\Gamma_n\}$ s.t. $\lim_{n\to\infty} \Gamma_n(A) = \Xi(A),\ \max_{z\in\Xi(A)} \mathrm{dist}(z, \Gamma_n(A)) \leq 2^{-n}$

Such problems can be used in a proof!

Small sample of classification theorems

Increasing difficulty

Error control

$$
\begin{tikzcd}[column sep=small]
\Pi_0 & & & \Pi_1 \arrow[dr, phantom, "\subsetneq" sloped] & & & \Pi_2 \arrow[dr, phantom, "\subsetneq" sloped] & & & \Pi_3 & \cdots \\
\Delta_0 \arrow[u, phantom, "\|"] \arrow[d, phantom, "\|"] & \subsetneq & \Delta_1 \arrow[ur, phantom, "\subsetneq" sloped] \arrow[dr, phantom, "\subsetneq" sloped] & \subsetneq \Sigma_1 \cup \Pi_1 \subsetneq & \Delta_2 \arrow[ur, phantom, "\subsetneq" sloped] \arrow[dr, phantom, "\subsetneq" sloped] & & \subsetneq \Sigma_2 \cup \Pi_2 \subsetneq & \Delta_3 \arrow[ur, phantom, "\subsetneq" sloped] \arrow[dr, phantom, "\subsetneq" sloped] & & \subsetneq \Sigma_3 \cup \Pi_3 \ldots & \\
\Sigma_0 & & & \Sigma_1 \arrow[ur, phantom, "\subsetneq" sloped] & & & \Sigma_2 \arrow[ur, phantom, "\subsetneq" sloped] & & & \Sigma_3 & \cdots
\end{tikzcd}
$$

Small sample of classification theorems

Increasing difficulty

Error control

$$\begin{tikzcd}
\Pi_0 \arrow[d, equal] & & \Pi_1 \arrow[dr, phantom, sloped, "\subsetneq"] & & \Pi_2 \arrow[dr, phantom, sloped, "\subsetneq"] & & \Pi_3 & \ddots \\
\Delta_0 \arrow[d, equal] \arrow[r, phantom, "\subsetneq"] & \Delta_1 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_1 \cup \Pi_1 \arrow[r, phantom, "\subsetneq"] & \Delta_2 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_2 \cup \Pi_2 \arrow[r, phantom, "\subsetneq"] & \Delta_3 \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] \arrow[r, phantom, "\subsetneq"] & \Sigma_3 \cup \Pi_3 & \ldots \\
\Sigma_0 & & \Sigma_1 \arrow[ur, phantom, sloped, "\subsetneq"] & & \Sigma_2 \arrow[ur, phantom, sloped, "\subsetneq"] & & \Sigma_3 & \iddots
\end{tikzcd}$$

Spectra of $\mathcal{K}$

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

Small sample of classification theorems

Increasing difficulty

Error control

Spectra of compact operators

$$
\begin{tikzcd}[column sep=small]
\Pi_0 \arrow[d, phantom, "\|"] & & & \Pi_1 \arrow[dr, phantom, sloped, "\subsetneq"] & & & \Pi_2 \arrow[dr, phantom, sloped, "\subsetneq"] & & & \Pi_3 \arrow[dr, phantom, sloped, "\cdots"] & \\
\Delta_0 \arrow[d, phantom, "\|"] \arrow[r, phantom, "\subsetneq"] & \Delta_1 \arrow[r, phantom, "\subsetneq"] \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] & \Sigma_1 \cup \Pi_1 \arrow[r, phantom, "\subsetneq"] & & \Delta_2 \arrow[r, phantom, "\subsetneq"] \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] & \Sigma_2 \cup \Pi_2 \arrow[r, phantom, "\subsetneq"] & & \Delta_3 \arrow[r, phantom, "\subsetneq"] \arrow[ur, phantom, sloped, "\subsetneq"] \arrow[dr, phantom, sloped, "\subsetneq"] & \Sigma_3 \cup \Pi_3 & \ldots & \\
\Sigma_0 & & & \Sigma_1 \arrow[ur, phantom, sloped, "\subsetneq"] & & & \Sigma_2 \arrow[ur, phantom, sloped, "\subsetneq"] & & & \Sigma_3 \arrow[ur, phantom, sloped, "\cdots"] &
\end{tikzcd}
$$

Spectra of $\mathcal{K}$

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

Small sample of classification theorems

Increasing difficulty

Error control

Spectra of compact operators

Spectra of Schrödinger*
(different potential classes)

$$
\begin{array}{ccccccccccccccc}
\Pi_0 & & & & \Pi_1 & & & & \Pi_2 & & & & \Pi_3 & & \\
\| & & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \ddots & \\
\Delta_0 & \subsetneq & \Delta_1 & \subsetneq & \Sigma_1 \cup \Pi_1 & \subsetneq & \Delta_2 & \subsetneq & \Sigma_2 \cup \Pi_2 & \subsetneq & \Delta_3 & \subsetneq & \Sigma_3 \cup \Pi_3 & \ldots & \\
\| & & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \rotatebox{-45}{$\subsetneq$} & & \rotatebox{45}{$\subsetneq$} & & \cdot^{\cdot^{\cdot}} & \\
\Sigma_0 & & & & \Sigma_1 & & & & \Sigma_2 & & & & \Sigma_3 & &
\end{array}
$$

Spectra of $\mathcal{K}$

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

**Open problem of Schwinger*: “The special canonical group,” “Unitary operator bases,” PNAS, 1960.

Small sample of classification theorems

Increasing difficulty

Error control

Spectra of compact operators

Spectra of Schrödinger* (different potential classes)

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

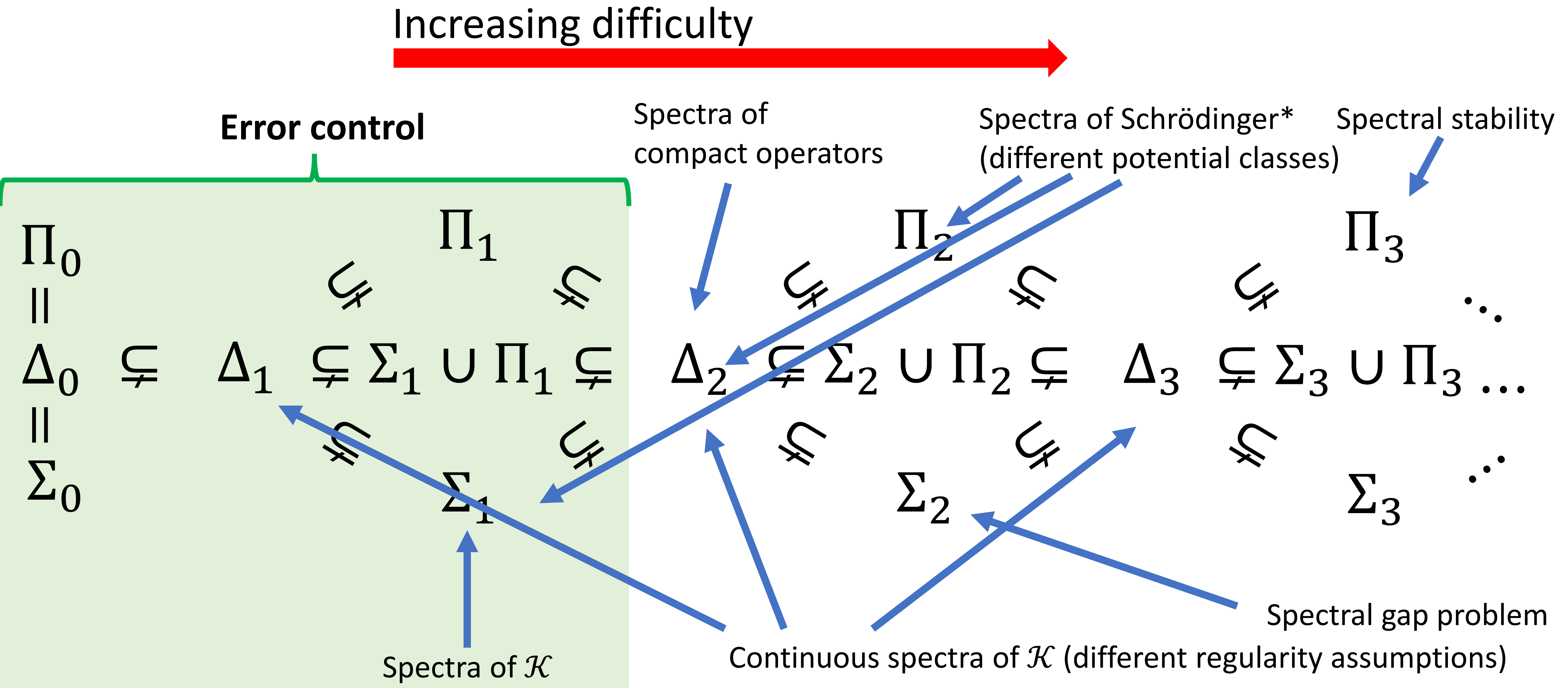
Spectra of $\mathcal{K}$

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

Spectral gap problem

**Open problem of Schwinger*: "The special canonical group," "Unitary operator bases," PNAS, 1960.

Small sample of classification theorems



**Open problem of Schwinger*: "The special canonical group," "Unitary operator bases," PNAS, 1960.