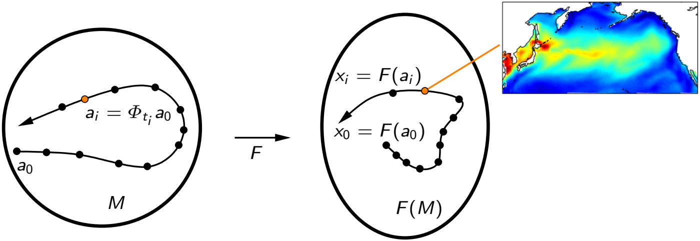
Data-Driven Spectral Decomposition and Forecasting of Ergodic Dynamical Systems

Dimitris Giannakis
Courant Institute of Mathematical Sciences, NYU

IPAM Workshop on Uncertainty Quantification for Multiscale Stochastic Systems and Applications

January 21, 2016

Setting & objectives



Ergodic dynamical system $(M, \mathcal{M}, \Phi_t, \mu)$ observed through a vector-valued function $F : M \mapsto \mathbb{R}^n$

Given time-ordered observations $\{x_0, \dots, x_{N-1}\}$ with $x_i = F(a_i)$, we seek to perform

1. **Dimension reduction** with timescale separation and invariance under changes of observation modality
2. **Nonparametric forecasting** of observables on M with deterministic or statistical initial data

Outline

1. Representation of Koopman operators in a data-driven orthonormal basis
2. Time-change techniques
3. Modes of organized tropical convection

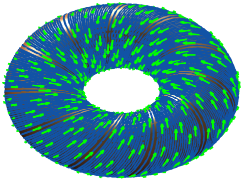
Acknowledgments
Tyrus Berry, John Harlim, Joanna Slawinska, Jane Zhao

State- and observable-centric viewpoints

- **State space viewpoint**

  In data space, we observe the manifold $F(M)$ and the vector field

  $$V|_x = \frac{dx}{dt} \quad \text{with} \quad x = F(\Phi_t a)$$



- **Operator-theoretic viewpoint** (Mezić et al. 2004, 2005, 2012, ...)
  Associated with the dynamical system is a group of unitary operators U_t on $L^2(M, \mu)$ s.t.

  $$U_t f(a) = f(\Phi_t a)$$

  The generator v of $\{U_t\}$ gives the directional derivative of functions along the dynamical flow

  $$vf(a) = \lim_{t \to 0} \frac{f(\Phi_t a) - f(a)}{t}, \quad V = DF\, v$$

Spectral characterization of ergodicity and mixing

A dynamical system $(M, \mathcal{M}, \Phi_t, \mu)$ is called

- **Ergodic** if all Φ_t-invariant sets have either zero or full measure

  Spectral characterization: 0 is a simple eigenvalue of v corresponding to a constant eigenfunction

- **Weak-mixing** if for all $A, B \in \mathcal{M}$ we have

$$\lim_{t\to\infty} \frac{1}{t} |\mu(\Phi_t(A) \cap B) - \mu(A)\mu(B)| = 0$$

  Spectral characterization: 0 is the only eigenvalue of v and this eigenvalue is simple

Systems with pure point spectra

$L^2(M, \mu)$ has an orthonormal basis consisting of eigenfunctions of v

$$v(z) = \lambda z, \quad \lambda = \mathrm{i}\omega, \quad \omega \in \mathbb{R}, \quad |z| = 1$$

The eigenvalues and eigenfunctions form a group

$$v(z\tilde{z}) = (\lambda + \tilde{\lambda})z\tilde{z}, \quad v(\tilde{z}) = \tilde{\lambda}\tilde{z}$$

- Such systems are metrically isomorphic to translations on **compact Abelian groups** equipped with the Haar measure

  The canonical phase spaces for diffeomorphisms of smooth manifolds are tori; constructions on other manifolds are available but have discontinuous eigenfunctions (Anosov & Katok 1970)

Dimension reduction for systems with pure point spectra

The group of eigenvalues for $M = \mathbb{T}^m$ is generated by m rationally independent frequencies $\Omega_i \in \mathbb{R}$ with corresponding eigenfunctions ζ_i

$$\omega_{k_1 \dots k_m} = \sum_{j=1}^{m} k_j \Omega_j, \quad z_{k_1 \dots k_m} = \prod_{j=1}^{m} \zeta_j^{k_j}, \quad k_j \in \mathbb{Z}$$

Dimension reduction map. $\pi : M \mapsto \mathbb{C}^m$ with

$$\pi(a) = (\pi_1(a), \dots, \pi_m(a)) = (\zeta_1(a), \dots \zeta_m(a))$$

- The π_i are **independent of observation modality**
- v is **projectible** under π_i, and the system evolves as a simple harmonic oscillator in the image space

$$\frac{d\zeta_i(\Phi_t a)}{dt} = v(\zeta_i)(\Phi_t a) = \mathrm{i}\Omega_i \zeta_i(\Phi_t a)$$

Vector field decomposition

Define the vector fields $v_i : L^2(M,\mu) \mapsto L^2(M,\mu)$ through their action on the eigenfunctions:

$$v_i(\zeta_1^{k_1} \cdots \zeta_i^{k_i} \cdots \zeta_m^{k_m}) = \mathrm{i} k_i \Omega_i \zeta_1^{k_1} \cdots \zeta_i^{k_i} \cdots \zeta_m^{k_m}$$

The v_i are linearly independent, nowhere vanishing, **mutually commuting** vector fields

$$v = \sum_{i=1}^{m} v_i, \quad [v_i, v_j] = 0$$

- Due to their vanishing commutator, the v_i can be thought of as **dynamically independent components**
- These vector fields can be realized in data space through the **pushforward map** $DF : TM \mapsto T\mathbb{R}^n$

$$V_i = DF(v_i) = v_i(F) = \sum_k A_k v(z_k), \quad A_k = \langle z_k, F \rangle$$

Data-driven basis

- Start from a **variable-bandwidth kernel** (Berry & Harlim 2014), $K_\epsilon : M \times M \mapsto \mathbb{R}_+$:

$$K_\epsilon(a_i, a_j) = \exp\left(- \frac{\|x_i - x_j\|^2}{\epsilon \hat{\sigma}_\epsilon(x_i)^{-1/m} \hat{\sigma}_\epsilon(x_j)^{-1/m}} \right),$$

$$m = \dim M, \quad x_i = F(a_i), \quad \hat{\sigma}_\epsilon(x_i) = \frac{1}{N(\pi\epsilon)^{m/2}} \sum_{j=0}^{N-1} e^{-\|x_i - x_j\|^2/\epsilon}$$

- Apply the **diffusion maps normalization** (Coifman & Lafon 2006, Berry & Sauer 2015):

$$\hat{q}_\epsilon(a_i) = \frac{1}{N} \sum_{j=0}^{N-1} K_\epsilon(a_i, a_j), \quad \hat{K}'_\epsilon(a_i, a_j) = \frac{K_\epsilon(a_i, a_j)}{\hat{q}_\epsilon(a_j)}$$

$$\hat{d}_\epsilon(a_i) = \frac{1}{N} \hat{K}'_\epsilon(a_i, a_j), \quad \hat{p}_\epsilon(a_i, a_j) = \frac{K'_\epsilon(a_i, a_j)}{\hat{d}_\epsilon(a_i)}$$

Data-driven basis

- $\hat{p}_\epsilon$ induces an **averaging operator** on $L^2(M, \hat{\mu})$ for the sampling measure $\hat{\mu} = N^{-1} \sum_{i=0}^{N-1} \delta_{a_i}$:

$$\hat{P}_\epsilon f(b) = \int_M \hat{p}_\epsilon(b, a) f(a) \, d\hat{\mu}(a) = \frac{1}{N} \sum_{j=0}^{N-1} \hat{p}_\epsilon(b, a_j) f(a_j)$$

- By ergodicity, as $N \to \infty$, $\hat{P}_\epsilon f(b)$ converges μ-a.s. to $P_\epsilon f(b)$, where

$$P_\epsilon f(b) = \int_M p_\epsilon(y, x) f(x) d\mu(x)$$

is an averaging operator on $L^2(M, \mu)$.

Data-driven basis

Uniformly on M (Coifman & Lafon 2006),

$$P_\epsilon f(a) = f(a) + \epsilon \Delta f(a) + O(\epsilon^2),$$

where Δ is the **Laplace-Beltrami operator** associated with the Riemannian metric $h = \sigma^{2/d} g$

- Effect of variable-bandwidth kernel is a **conformal transformation** such that the Riemannian measure is equal to the invariant measure
- The eigenfunctions $\{\phi_0, \phi_1, \dots\}$ of Δ are orthogonal on $L^2(M, \mu)$
- The **rescaled eigenfunctions** $\varphi_i = \phi_i / \eta_i^{1/2}$ corresponding to eigenvalue η_i are orthogonal on $H^1(M, h)$

$$\int_M \operatorname{grad}_h \varphi_i \cdot \operatorname{grad}_h \varphi_j \, d\mu = \delta_{ij}$$

- In practice, we approximate $(\eta_i, \phi_i, \varphi_i)$ by solving for

$$\hat{P}_\epsilon \hat{\phi}_i = (1 - \epsilon \hat{\eta}_i) \hat{\phi}_i, \quad \hat{\varphi}_i = \hat{\phi}_i / \hat{\eta}_i^{1/2}$$

Eigenvalue problem for the Koopman generator

$$v(z) = \lambda z$$

- In dimension $m \geq 2$ the eigenvalues form a dense set on the imaginary line
- We eliminate highly rough eigenfunctions by solving the eigenvalue problem for $L_\epsilon = v + \epsilon\Delta$.

Continuous problem. Find $z \in H^1(M, h)$ and $\lambda \in \mathbb{C}$ s.t.

$$\langle \psi, v(z) \rangle + \epsilon \langle \operatorname{grad}_h \psi, \operatorname{grad}_h z \rangle = \lambda \langle \psi, z \rangle, \quad \forall \psi \in H^1(M, h)$$

Discrete approximation. Set $\hat{H}_l = \operatorname{span}\{\hat{\varphi}_0, \ldots, \varphi_{l-1}\} \subseteq L^2(M, \hat{\mu})$. Find $\hat{z} \in \hat{H}_l$ and $\hat{\lambda} \in \mathbb{C}$ s.t.

$$\langle \psi, \hat{v}(\hat{z}) \rangle_{\hat{\mu}} + \langle \widehat{\operatorname{grad}}_h \psi, \widehat{\operatorname{grad}}_h \hat{z} \rangle_{\hat{\mu}} = \hat{\lambda} \langle \psi, \hat{z} \rangle_{\hat{\mu}}, \quad \forall \psi \in \hat{H}_l.$$

Eigenvalue problem for the Koopman generator

$$\langle \psi, \hat{v}(\hat{z}) \rangle_{\hat{\mu}} + \langle \widehat{\operatorname{grad}}_h \psi, \widehat{\operatorname{grad}}_h \hat{z} \rangle_{\hat{\mu}} = \hat{\lambda} \langle \psi, \hat{z} \rangle_{\hat{\mu}},$$

$$\hat{z} = \sum_{i=0}^{l-1} c_i \hat{\varphi}_i, \quad \psi = \sum_{i=0}^{l-1} w_i \hat{\varphi}_i$$

- $\hat{v}$ is a **finite-difference approximation** of v, e.g.,

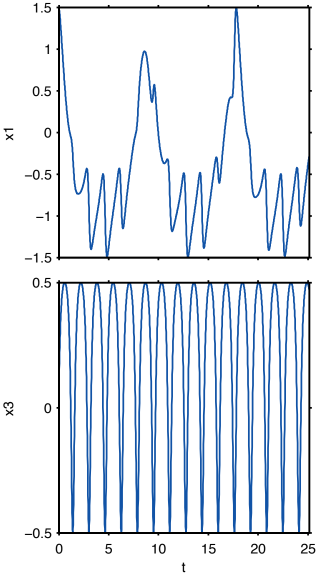
$$\begin{aligned} \langle \psi, \hat{v}(\hat{z}) \rangle_{\hat{\mu}} &= \sum_{i,j=0}^{l-1} w_i c_j \int_M \hat{\varphi}_i \hat{v}(\hat{\varphi}_j)\, d\hat{\mu} \\ &= \sum_{i,j=0}^{l-1} w_i c_j \left[\frac{1}{N} \sum_{k=1}^{N-2} \hat{\varphi}_i(a_k) \frac{\hat{\varphi}_j(a_{k+1}) - \hat{\varphi}_j(a_{k-1})}{2\,\delta t} \right] \end{aligned}$$

- By construction of the $\{\hat{\varphi}_i\}$ basis,

$$\langle \widehat{\operatorname{grad}}_h \psi, \widehat{\operatorname{grad}}_h \hat{z} \rangle_{\hat{\mu}} = \sum_{i,j=0}^{l-1} w_i c_j \delta_{ij}$$

- Scheme remains well-conditioned at large spectral order l

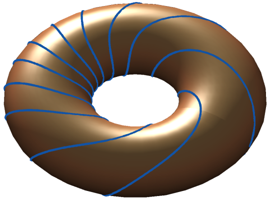
Variable-speed flow on $\mathbb{T}^2$



$$v = \sum_{\mu=1}^{2} v^{\mu} \frac{\partial}{\partial \theta^{\mu}}$$

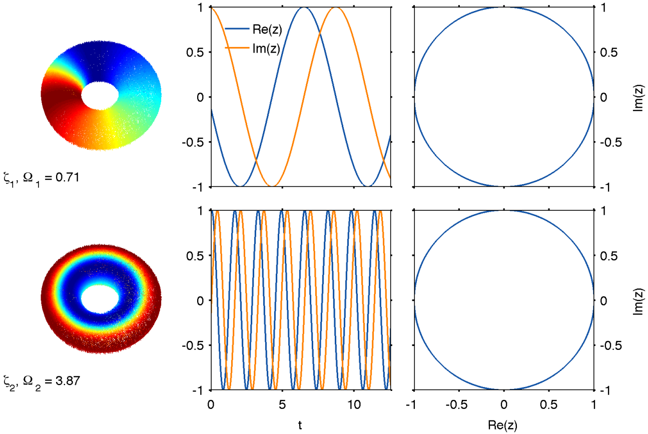
$$v^1 = 1 + \beta \cos \theta^1$$

$$v^2 = \bar{\omega}(1 - \beta \sin \theta^2)$$

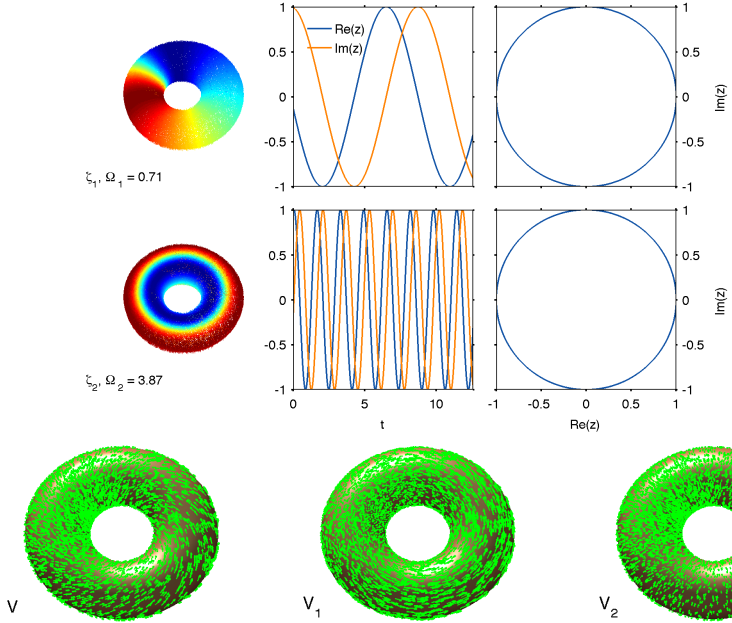


$\bar{\omega} = \sqrt{30}$, $\beta = \sqrt{1/2}$

Results for variable-speed flow on $\mathbb{T}^2$



Results for variable-speed flow on $\mathbb{T}^2$



Forecasting densities and expectation values

The adjoint U_t^* on $L^2(M,\mu)$ governs the evolution of probability densities relative to μ

$$\rho_0 = \sum_k c_k(0) z_k, \quad c_k(0) = \langle z_k, \rho_0 \rangle, \quad k = (k_1, \dots, k_m)$$

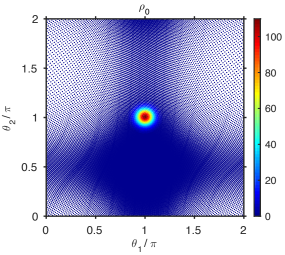
$$\rho_t = U_t^* \rho_0 = \sum_k c_k(t) z_k, \quad c_k(t) = \sum_k e^{-\mathrm{i}\omega_k t} c_k(0)$$

The time-dependent expectation value of an observable f is

$$\mathbb{E}_t f = \sum_k \hat{f}_k c_k^*(t), \quad \hat{f}_k = \langle z_k, f \rangle$$

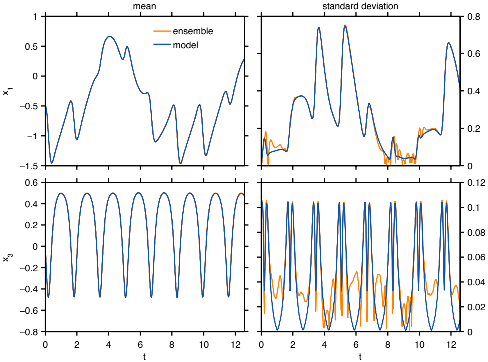
- By computing $\mathbb{E}_t f$ and $\mathbb{E}_t f^2$ the method keeps track of both the **forecast mean** and the **forecast uncertainty**
- The forecast accuracy depends on the **bandwidth** of f in the $\{z_k\}$ basis

Nonparametric forecasting of the variable-speed system on $\mathbb{T}^2$



Initial distribution has circular Gaussian density relative to μ with mean (π, π) and variance $(30^{-1}, 30^{-1})$

Forecasts with the nonparametric model are in good agreement with ensemble forecasts with the perfect model for both the mean and uncertainty

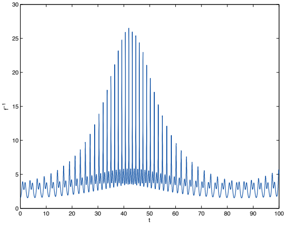


Time change and mixing

Given a flow Φ_t with generator v and a positive function f, construct the flow Ψ_t with generator $w = f^{-1} v$

Ψ_t has the same orbits as Φ_t, but evolves at different speed and has invariant measure ν with $d\nu = f\, d\mu$

- For any ergodic Φ_t there exists a time change s.t. Ψ_t is **mixing** (Kochergin 1973)



Example on $\mathbb{T}^3$ (Fayad 2002). Φ_t is an irrational flow, and

$$f(\theta^1, \theta^2, \theta^3) = 1 + \mathrm{Re} \sum_{k=1}^{\infty} \sum_{|l| \le k} \frac{e^{-k}}{k} \left(e^{\mathrm{i}k\theta^1} + e^{\mathrm{i}k\theta^2} \right) e^{\mathrm{i}lz}$$

Regularization by time change

Construct an orthonormal basis for $L^2(M, \nu)$ with $d\nu = \|v\| \, d\mu$ using the kernel

$$K_\epsilon(a_i, a_j) = \exp\left(-\frac{\|x_i - x_j\|^2 \| V|_{x_i} \|^{1/m} \| V|_{x_j} \|^{1/m}}{\epsilon \sigma_\epsilon(x_i)^{-1/m} \sigma_\epsilon(x_j)^{-1/m}} \right)$$

Solve the eigenvalue problem

$$w(z) = \mathrm{i}\omega z, \quad w = \frac{v}{\|v\|}, \quad z \in L^2(M, \nu)$$

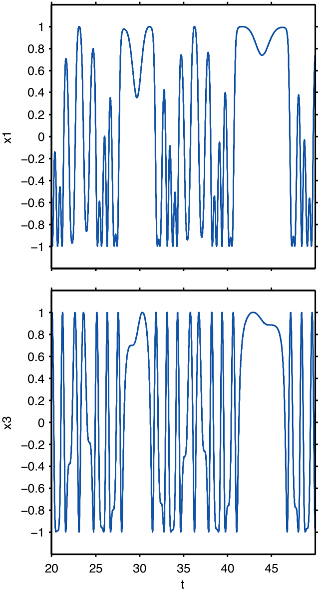
- In the reduced coordinates $\pi(a) = z(a) \in \mathbb{C}$ the system evolves as an oscillator with **variable frequency** $\omega\|v\|$

Make the vector field decomposition

$$v = \sum_{i=1}^{m} v_i, \quad v_i = \|v\| w_i, \quad w = \sum_{i=1}^{m} w_i, \quad [w_i, w_j] = 0$$

- In general, the v_i are **non-commuting** vector fields

Flow on $\mathbb{T}^2$ with fixed points (Oxtoby 1953)



The system

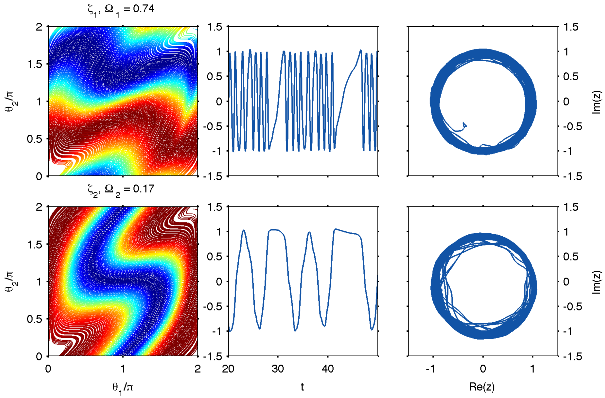
$$v = \sum_{\mu=1}^{2} v^\mu \frac{\partial}{\partial \theta^\mu}$$

$$v^1 = \alpha(1 - \cos(\theta^1 - \theta^2)) + (1-\alpha)(1 - cos\theta^2)$$

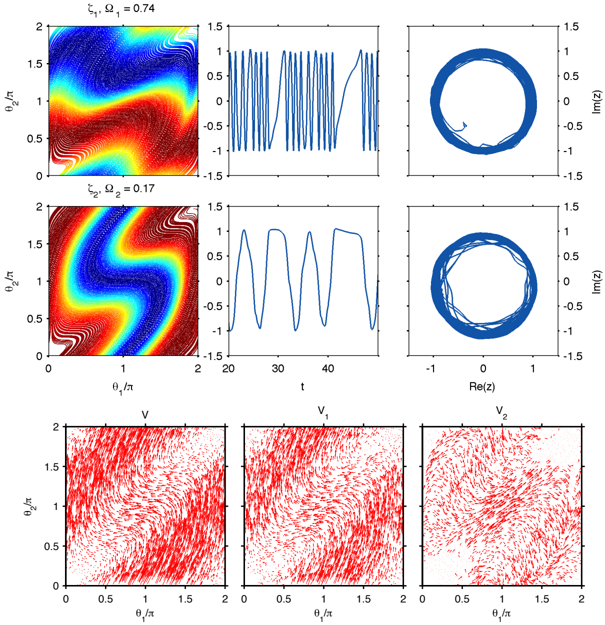
$$v^2 = \alpha(1 - \cos(\theta^1 - \theta^2))$$

has a fixed point at $\theta^1 = \theta^2 = 0$, and preserves the Haar measure

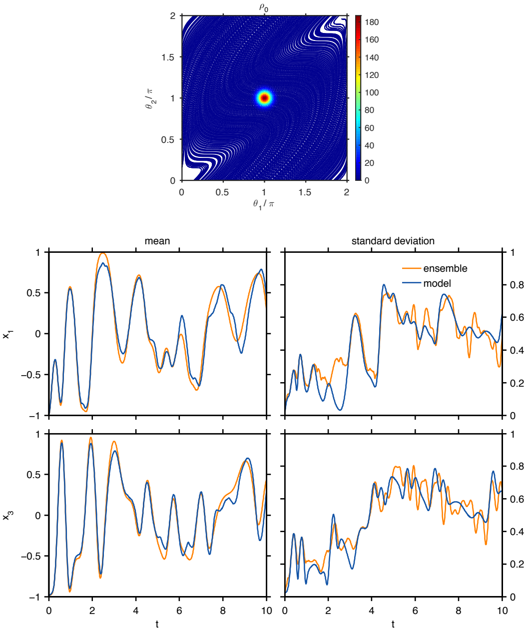
Vector field decomposition for the fixed-point system



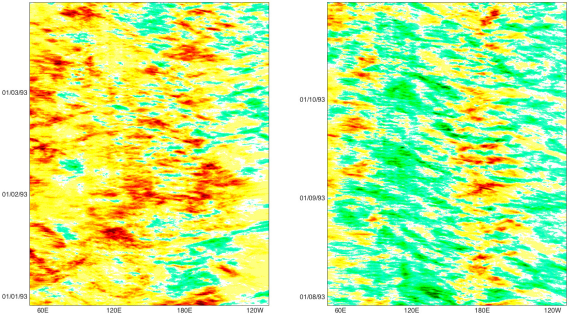
Vector field decomposition for the fixed-point system



Nonparametric forecasting



Analysis of organized tropical convection



- **Brightness temperature**, T_b, is the blackbody emission temperature received from Earth by satellites
- Deep convective clouds become cold, and present as a negative T_b anomaly against the Earth's surface
- We analyze T_b data from the CLAUS archive averaged over the latitudes 15°S–15°N
- Sampling is 8 $\times$ daily at 0.5° resolution for the period 1983–2006

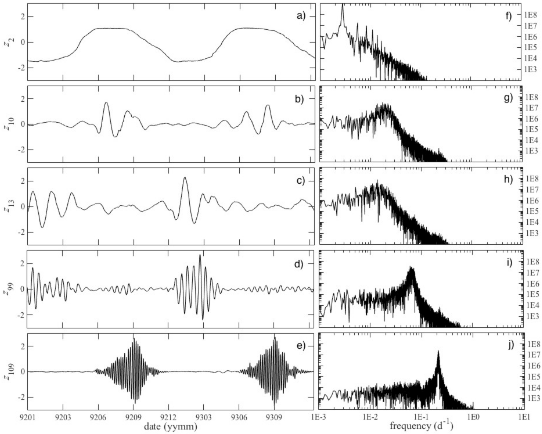
Delay-coordinate embeddings

- The raw $T_b(t)$ timeseries is highly non-Markovian
- To recover information lost in partial observations, we perform **delay-coordinate mapping** (Packard et al. 1980, Sauer et al. 1991)

$$x(t) = (\, T_b(t), T_b(t - \delta t), \dots, T_b(t - (q-1)\,\delta t))$$

- This procedure significantly improves the quality of the diffusion eigenfunction basis (G. & Majda 2012, Berry et al. 2013)
- We use $q = 512$, equivalent to 64 days (intraseasonal timescale)

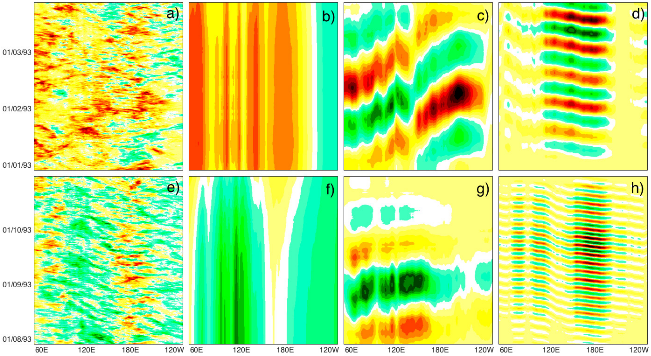
Koopman eigenfunctions



Multiple timescales are resolved including:

- (a) The annual cycle.
- (b, c) Intraseasonal oscillations.
- (d, e) Convectively coupled equatorial waves (CCEWs).

Spatiotemporal reconstructions



- (b, f) Annual cycle.
- (c) Madden-Julian oscillation.
- (d,h) Westward-propagating CCEWs.
- (g) Boreal summer intraseasonal oscillation (Indian Monsoon).

Effects of partial observations

- The computed Koopman eigenfunctions have amplitude modulations which are not consistent with the skew-symmetry of v
- Despite delay-coordinate mapping, it is unrealistic to expect that we have recovered the full attractor of the climate system
- Instead of the full generator, it is more likely that we are approximating an operator of the form

$$\tilde{v} = \Pi v \Pi,$$

  where Π is a projector to the subspace of the full L^2 space on the attractor spanned by the diffusion eigenfunctions obtained from T_b
- It is plausible that an effective description of $\tilde{v}$ is through a nonautonomous advection-diffusion process
- Consistent with stochastic oscillator models for the MJO (Chen et al. 2014)

Summary

- The spectral properties of Koopman operators have attractive properties for dimension reduction and nonparametric forecasting of dynamical systems
- These operators can be approximated from time-ordered data with no a priori knowledge of the equations of motion using kernel methods
- In systems with pure point spectra, the eigenfunctions of the Koopman group lead to a decomposition of the dynamics into a collection of independent harmonic oscillators
- Time change extends the applicability of Koopman eigendecomposition techniques to certain classes of mixing systems, which are now decomposed into coupled oscillators with time-dependent frequencies

References

- Giannakis, D. (2015). Data-driven spectral decomposition and forecasting of ergodic dynamical systems. arXiv:1507.02338
- Berry, T., D. Giannakis, J. Harlim (2015). Nonparametric forecasting of low-dimensional dynamical systems. *Phys. Rev. E*, 91, 032915
- Giannakis, D., J. Slawinska, Z. Zhao (2015). Spatiotemporal feature extraction with data-driven Koopman operators. *J. Mach. Learn. Res. Proceedings*, 44, 103–115