PADERBORN UNIVERSITY

PADERBORN UNIVERSITY

KERNEL-BASED APPROXIMATION OF THE KOOPMAN GENERATOR AND SCHRÖDINGER OPERATOR

Feliks Nüske

Paderborn, 16 September, 2021

Outline

Stochastic Differential Equations

- Stochastic differential equation (SDE) on a domain $\mathbb{X}$:

$$dX_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t.$$

- Evolution of expectations: for $f \in L^\infty$, what is $f_t(x) = \mathbb{E}^x[f(X_t)]$?
- Solution is provided by *backward Kolmogorov equation*:

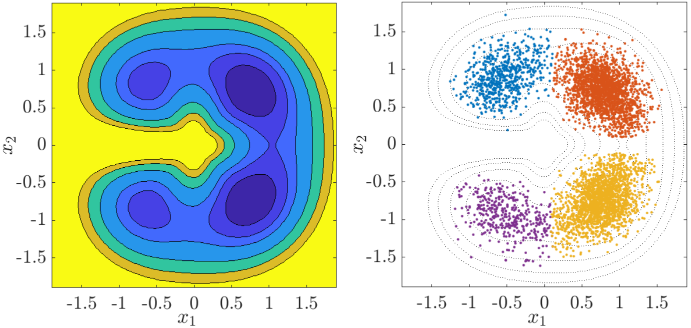
$$\frac{\partial f_t}{\partial t} = b \cdot \nabla f_t + \frac{1}{2} a : \nabla^2 f_t := \mathcal{L} f_t.$$

- If a unique invariant density ρ exists, $\mathcal{L}$ can be treated as an operator on L^2_ρ, i.e. Hilbert space methods can be used.

Analysis of the Generator

Generator $\mathcal{L}$ provides access to rich information about the system:

- Spectral analysis, identification of metastable sets based on eigenpairs.
- System identification, model reduction, control, transition rates, ...



Potential Energy and Metastable States 2d Diffusion System

Outline

Generator Approximation (gEDMD)

- Choose finite-dimensional $\mathbb{V} \subset L^2_\rho$, basis $\psi = [\psi_1, \dots, \psi_n]^T$.
- **Galerkin projection** $\mathcal{L}_\mathbb{V} = \mathcal{P}_\mathbb{V} \mathcal{L} \mathcal{P}_\mathbb{V}$, such that

$$\left\langle \phi,\, \mathcal{L}\phi' \right\rangle_\rho = \left\langle \phi,\, \mathcal{L}_\mathbb{V}\phi' \right\rangle_\rho \quad \forall \phi,\, \phi' \in \mathbb{V}.$$

- Matrix representation on $\mathbb{V}$:

$$\mathbf{L}_\mathbb{V} = \mathbf{G}^{-1}\mathbf{A}, \qquad \mathbf{G}_{ij} = \left\langle \psi_i,\, \psi_j \right\rangle_\rho = \mathbb{E}^\rho[\psi_i(\mathcal{X}_s)\psi_j(\mathcal{X}_s)],$$

$$\mathbf{A}_{ij} = \left\langle \psi_i,\, \mathcal{L}\psi_j \right\rangle_\rho = \mathbb{E}^\rho[\psi_i(\mathcal{X}_s)\mathcal{L}\psi_j(\mathcal{X}_s)]$$

Klus, **Nüske**, Peitz, et al, *Physica D* (2020)

Generator Approximation (gEDMD)

- If we have snapshots $\mathcal{X}_{s_l}$, $1 \leq l \leq m$ from long, stationary realizations of the dynamics, we can approximate:

$$\mathbf{G}_{ij} \approx \frac{1}{m}\sum_{l=1}^{m} \psi_i(\mathcal{X}_l)\psi_j(\mathcal{X}_l), \quad \mathbf{A}_{ij} \approx \frac{1}{m}\sum_{l=1}^{m} \psi_i(\mathcal{X}_l)\mathcal{L}\psi_j(\mathcal{X}_l).$$

- **Data-driven approximation** (with $\mathbf{X} = [\mathcal{X}_1, \ldots, \mathcal{X}_m]$):

$$\hat{\mathbf{L}}_{\mathbb{V}} = (\Psi(\mathbf{X})\Psi(\mathbf{X})^T)^{-1}((\Psi(\mathbf{X})\mathcal{L}\Psi(\mathbf{X})^T).$$

Klus, **Nüske**, Peitz, et al, *Physica D* (2020)

Outline

Definition RKHS

Definition

Let $\mathbb{X}$ be an open domain and $\mathbb{H}$ a space of continuous functions $f: \mathbb{X} \to \mathbb{R}$. Then $\mathbb{H}$ is called a *reproducing kernel Hilbert space* (RKHS) with inner product $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ if a function $k: \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ exists such that

1. $\mathbb{H} = \overline{\mathrm{span}\{k(x, \cdot),\ x \in \mathbb{X}\}}$,
2. $\langle f,\ k(x, \cdot) \rangle_{\mathbb{H}} = f(x)$ for all $f \in \mathbb{H}$.

Wendland, *Scattered Data Approximation* (2005)

Derivative Reproducing Property

The Reproducing Property can be extended to derivatives if the kernel is smooth:

Theorem

Let $k(\cdot, \cdot) \in C^{2k}(\mathbb{X} \times \mathbb{X})$ *be a positive semi-definite function on an open set. Then all functions in* $\mathbb{H}$ *are* C^k *and we have for all* $|\alpha| \leq k$*:*

$$D^\alpha f(x) = \langle D^\alpha k(x, \cdot),\ f \rangle_{\mathbb{H}},$$

where the derivative acts on the first argument of k*.*

Wendland, *Scattered Data Approximation* (2005)

Outline

Rank-one Operators

Let $x \in \mathbb{X}$ and α a multi-index. Consider a rank-one operator on the RKHS $\mathbb{H}$:

$$\mathcal{T}_x^\alpha \phi := \langle D^\alpha k(x,\cdot),\, \phi\rangle_{\mathbb{H}}\, k(x,\cdot).$$

With derivative reproducing property, we verify that for $\phi, \phi' \in \mathbb{H}$:

$$\left\langle \mathcal{T}_x^\alpha \phi,\, \phi' \right\rangle_{\mathbb{H}} = D^\alpha \phi(x) \left\langle k(x,\cdot),\, \phi' \right\rangle_{\mathbb{H}} = D^\alpha \phi(x) \phi'(x).$$

Linear Differential Operators

Let $\mathcal{T}\phi = \sum_\alpha w_\alpha D^\alpha \phi$ be a linear differential operator and μ a measure on $\mathbb{X}$. Consider a formal operator on $\mathbb{H}$:

$$\mathcal{T}_\mathbb{H}\phi = \int_\mathbb{X} \left\langle \sum_\alpha w_\alpha(x) D^\alpha k(x,\cdot),\, \phi \right\rangle_\mathbb{H} k(x,\cdot)\,\mathrm{d}\mu(x).$$

By the same trick as on the previous slide, we find for all $\phi, \phi' \in \mathbb{H}$:

$$\begin{aligned}\left\langle \mathcal{T}_\mathbb{H}\phi,\, \phi' \right\rangle_\mathbb{H} &= \int_\mathbb{X} \mathcal{T}\phi(x) \left\langle k(x,\cdot,\, \phi' \right\rangle_\mathbb{H} \mathrm{d}\mu(x) \\ &= \int_\mathbb{X} \mathcal{T}\phi(x)\phi'(x)\,\mathrm{d}\mu(x) = \left\langle \mathcal{T}\phi,\, \phi' \right\rangle_\mu .\end{aligned}$$

RKHS Galerkin Projection

Theorem

Assume that $\mathbb{H} \subset \mathcal{D}(\mathcal{T}) \subset L^2_\mu$*, and that for all relevant* α*:*

$$\int_{\mathbb{X}} |w_\alpha(x)| \|D^\alpha k(x,\cdot)\|_{\mathbb{H}} \|k(x,\cdot)\|_{\mathbb{H}} \,\mathrm{d}\mu(x) < \infty,$$

Then, for all $\phi, \phi' \in \mathbb{H}$*,*

$$\left\langle \mathcal{T}\phi,\, \phi' \right\rangle_\mu = \left\langle \mathcal{T}_{\mathbb{H}}\phi,\, \phi' \right\rangle_{\mathbb{H}}.$$

Note: this applies in particular to the backward Kolmogorov operator $\mathcal{L}\phi = \frac{1}{2}\sum_{i,j} a_{ij}(x) D^{e_i+e_j}\phi(x) + \sum_i b_i(x) D^{e_i}\phi(x)$, and $\mu = \rho$.

Klus, **Nüske**, and Hamzi, *Entropy* (2020)

Data-driven Approximation

- If μ is a probability measure, we can use data $\{\mathcal{X}_l\}_{l=1}^m$ to approximate the integral in $\mathcal{T}_{\mathbb{H}}$.
- By further restricting the problem to linear span of $k(\mathcal{X}_l, \cdot)$, we get back to finite-dimensional problems.
- Counterparts of the standard Galerkin matrices are:

$$\mathbf{G}^k_{rs} = k(\mathcal{X}_r, \mathcal{X}_s), \qquad \mathbf{A}^k_{rs} = (\mathcal{T}k)(\mathcal{X}_r, \mathcal{X}_s).$$

Klus, **Nüske**, and Hamzi, *Entropy* (2020)

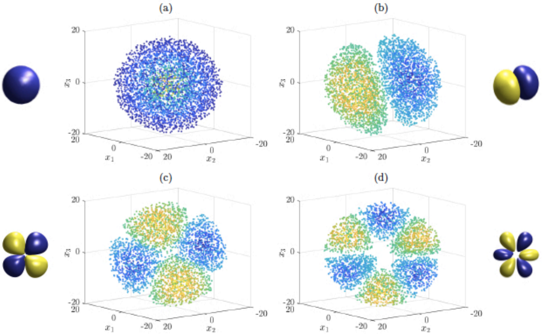
Outline

Schrödinger Operators

- The above also applies to Schrödinger operators (μ uniform):

$$\mathcal{H}\psi = -\frac{1}{2}\Delta\Psi + V\psi = -\frac{1}{2}\sum_i D^{2e_i}\psi + V\psi.$$

- Hydrogen atom ($\mathbb{X} = \mathbb{R}^3$, $V = -\frac{1}{\|x\|}$, $m = 5000$, Gaussian kernel):



Symmetries of Quantum Systems

- Quantum systems often require (anti-)symmetry of the wavefunction ψ. If $\mathbb{X} = \mathbb{R}^{dN}$ (e.g. N particles in d-dim. space), and S_N is the permutation group, we need for all permutations $\pi \in S_N$:

$$\psi(x_1, \ldots, x_N) = \psi(\pi(x_1, \ldots, x_N))$$
$$\text{or } \psi(x_1, \ldots, x_N) = \operatorname{sgn}(\pi)\psi(\pi(x_1, \ldots, x_N)).$$

- Can these symmetries be built into data-driven approximations?

Anti-symmetric Kernels

Lemma

Let $k : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ be a kernel, and $\mathbb{X} \subset \mathbb{R}^d$. We define an antisymmetric function $k_a : \mathbb{X} \times \mathbb{X} \to \mathbb{R}$ by

$$k_a(x, x') = \frac{1}{d!^2} \sum_{\pi \in S_d} \sum_{\pi' \in S_d} \operatorname{sgn}(\pi) \operatorname{sgn}(\pi') k\big(\pi(x), \pi'(x')\big).$$

Then k_a is a symmetric positive semi-definite kernel and generates an RKHS of anti-symmetric functions.

Klus, Gelß, **Nüske**, and Noé, *Machine Learning: Science and Technology* (2021)

Reducing the Computational Effort

Lemma

A kernel k is called permutation-invariant if $k(x, x') = k(\pi(x), \pi(x'))$ for all $\pi \in S_d$. If this condition holds, we have:

$$k_a(x, x') = \frac{1}{d!} \sum_{\pi \in S_d} \operatorname{sgn}(\pi) k\big(\pi(x), x'\big) = \frac{1}{d!} \sum_{\pi \in S_d} \operatorname{sgn}(\pi) k\big(x, \pi(x')\big).$$

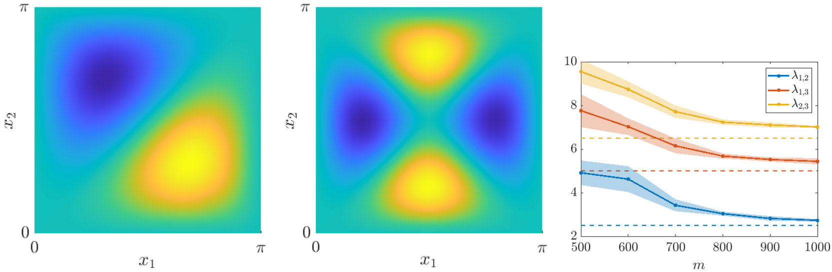
For the Gaussian kernel with bandwidth σ, we obtain:

$$k_a(x, x') = \frac{1}{d!} \begin{vmatrix} e^{-\frac{(x_1 - x_1')^2}{2\sigma^2}} & \dots & e^{-\frac{(x_1 - x_d')^2}{2\sigma^2}} \\ \vdots & \ddots & \vdots \\ e^{-\frac{(x_d - x_1')^2}{2\sigma^2}} & \dots & e^{-\frac{(x_d - x_d')^2}{2\sigma^2}} \end{vmatrix}.$$

Klus, Gelß, **Nüske**, and Noé, *Machine Learning: Science and Technology* (2021)

Particle in a Box

- Particle in a box system:
  $\mathbb{X} = \mathbb{R}^2, \quad V(x) = 0 \; x \in [0, \pi]^2; \quad V(x) = \infty$ otherwise.
- Analytical eigenfunctions $\psi_{l_1,l_2}(x_1, x_2) = \frac{2}{\pi} \sin(l_1 x_1) \sin(l_2 x_2)$ are either symmetric or anti-symmetric.
- Kernel-discretization with anti-symmetrized Gaussian kernel picks up anti-symmetric eigenpairs:



Klus, Gelß, **Nüske**, and Noé, *Machine Learning: Science and Technology* (2021)

Thanks

Thank you for your attention!

Joint work with: Stefan Klus (U Surrey), Sebastian Peitz (UPB), Boumediene Hamzi (Imperial), Patrick Gelß (FU Berlin), Frank Noé (FU Berlin)

Main Papers:

Klus, **Nüske**, Peitz, et al, *Physica D: Nonlinear Phenomena*, 406, 132416 (2020)

Klus, **Nüske**, and Hamzi, *Entropy*, 22 (7), 722 (2020)

Klus, Gelß, **Nüske**, and Noé, *Machine Learning: Science and Technology*, 2 (4), 045016 (2021)