Data-Driven Methods in Control: Error Bounds and Guaranteed Stability

Manuel Schaller

Workshop and Summer School on Applied Analysis 2025

26.09.2025

Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation

Koopman-based methods

Koopman-based predictions

For $\Omega \subset \mathbb{R}^n$ and $F : \Omega \to \Omega$, consider the dynamical system

$$x_{i+1} = F(x_i), \qquad x_0 \in \Omega.$$

Rowley, Mezić, Bagheri, Schlatter, Henningson, J. Fluid Mech. 2009
Mezić JNLS 2005, Annu. Rev. Fluid Mech 2013
Brunton, Budišić, Kaiser, Kutz, SIAM Rev. 2022
Lusch, Kutz, Brunton, Nature comm. 2018

Koopman-based predictions

For $\Omega \subset \mathbb{R}^n$ and $F : \Omega \to \Omega$, consider the dynamical system

$$x_{i+1} = F(x_i), \qquad x_0 \in \Omega.$$

Goal: Predict the dynamics of an **observable** along the flow, i.e., $\varphi(x_i)$, $i \in \mathbb{N}$

- Obvious way: Predict x_i and evaluate $\varphi(x_i)$

Rowley, Mezić, Bagheri, Schlatter, Henningson, J. Fluid Mech. 2009
Mezić JNLS 2005, Annu. Rev. Fluid Mech 2013
Brunton, Budišić, Kaiser, Kutz, SIAM Rev. 2022
Lusch, Kutz, Brunton, Nature comm. 2018

Koopman-based predictions

For $\Omega \subset \mathbb{R}^n$ and $F : \Omega \to \Omega$, consider the dynamical system

$$x_{i+1} = F(x_i), \qquad x_0 \in \Omega.$$

Goal: Predict the dynamics of an **observable** along the flow, i.e., $\varphi(x_i)$, $i \in \mathbb{N}$

- Obvious way: Predict x_i and evaluate $\varphi(x_i)$
- Alternatively: Koopman operator $\mathcal{K}$ acting on observables $\varphi : \Omega \to \mathbb{R}$:

$$(\mathcal{K}\varphi)(x_0) = \varphi(x^+(x_0)) = \varphi(F(x_0))$$

Rowley, Mezić, Bagheri, Schlatter, Henningson, J. Fluid Mech. 2009
Mezić JNLS 2005, Annu. Rev. Fluid Mech 2013
Brunton, Budišić, Kaiser, Kutz, SIAM Rev. 2022
Lusch, Kutz, Brunton, Nature comm. 2018

Koopman-based predictions

For $\Omega \subset \mathbb{R}^n$ and $F : \Omega \to \Omega$, consider the dynamical system

$$x_{i+1} = F(x_i), \qquad x_0 \in \Omega.$$

Goal: Predict the dynamics of an **observable** along the flow, i.e., $\varphi(x_i)$, $i \in \mathbb{N}$

- Obvious way: Predict x_i and evaluate $\varphi(x_i)$
- Alternatively: Koopman operator $\mathcal{K}$ acting on observables $\varphi : \Omega \to \mathbb{R}$:

$$(\mathcal{K}\varphi)(x_0) = \varphi(x^+(x_0)) = \varphi(F(x_0))$$

Central properties: $\mathcal{K}$ linear and $(\mathcal{K}^i\varphi)(x_0) = \varphi(x_i(x_0))$.

Rowley, Mezić, Bagheri, Schlatter, Henningson, J. Fluid Mech. 2009
Mezić JNLS 2005, Annu. Rev. Fluid Mech 2013
Brunton, Budišić, Kaiser, Kutz, SIAM Rev. 2022
Lusch, Kutz, Brunton, Nature comm. 2018

Koopman-based predictions

For $\Omega \subset \mathbb{R}^n$ and $F : \Omega \to \Omega$, consider the dynamical system

$$x_{i+1} = F(x_i), \qquad x_0 \in \Omega.$$

Goal: Predict the dynamics of an **observable** along the flow, i.e., $\varphi(x_i)$, $i \in \mathbb{N}$

- Obvious way: Predict x_i and evaluate $\varphi(x_i)$
- Alternatively: Koopman operator $\mathcal{K}$ acting on observables $\varphi : \Omega \to \mathbb{R}$:

$$(\mathcal{K}\varphi)(x_0) = \varphi(x^+(x_0)) = \varphi(F(x_0))$$

Central properties: $\mathcal{K}$ linear and $(\mathcal{K}^i\varphi)(x_0) = \varphi(x_i(x_0))$.

Predict observations from previous ones: $\varphi(x^+) = \mathcal{K}\varphi(x)$

Rowley, Mezić, Bagheri, Schlatter, Henningson, J. Fluid Mech. 2009
Mezić JNLS 2005, Annu. Rev. Fluid Mech 2013
Brunton, Budišić, Kaiser, Kutz, SIAM Rev. 2022
Lusch, Kutz, Brunton, Nature comm. 2018

Illustration of the Koopman scheme

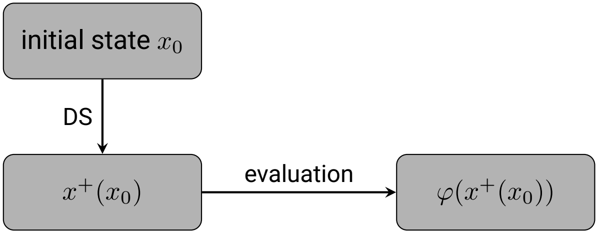


Illustration of the Koopman scheme

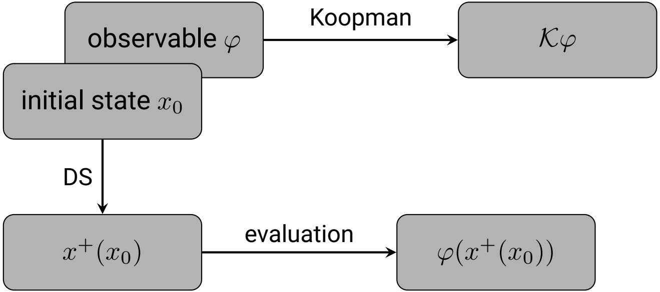


Illustration of the Koopman scheme

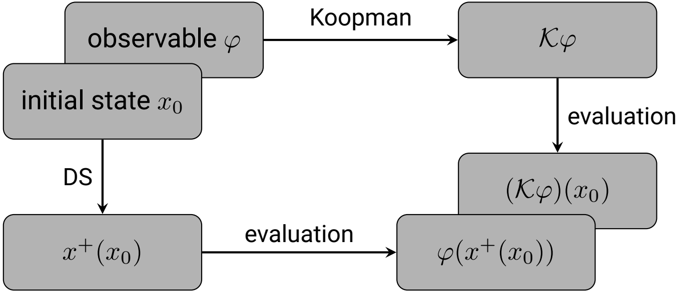
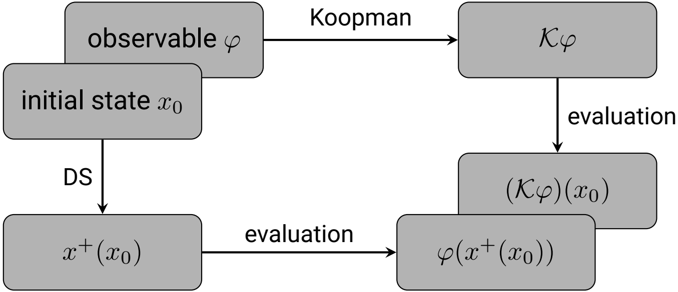
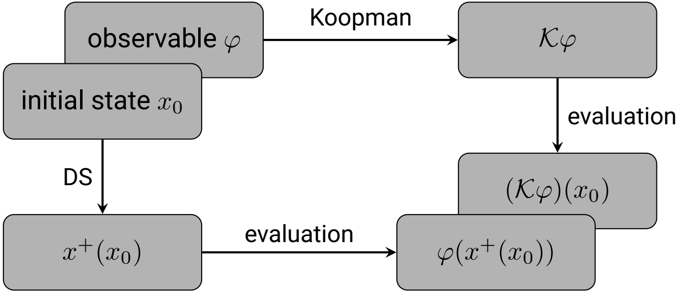


Illustration of the Koopman scheme



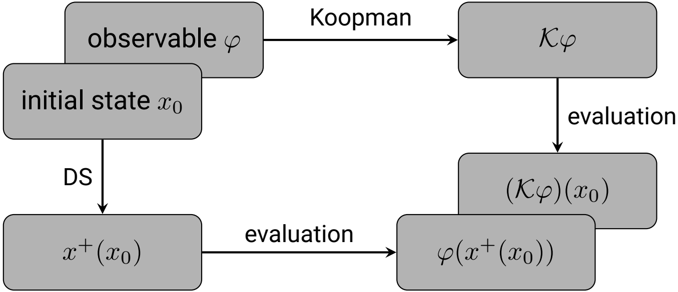
Old idea: Koopman 1931

Illustration of the Koopman scheme



Old idea: Koopman 1931
Today: Powerful tools to approximate $\mathcal{K}$ based on data: Extended Dynamic Mode Decomposition

Illustration of the Koopman scheme



Old idea: Koopman 1931
Today: Powerful tools to approximate $\mathcal{K}$ based on data: Extended Dynamic Mode Decomposition

Successfully applied in fluid dynamics, epidemiology, neuroscience, video processing and molecular dynamics [Schmid '10], [Tu et al. '14], [Brunton, Kutz et al. '16] [Noé et al. '15]

Wake behind a cylinder

$$\dot{v} + (v \cdot \nabla)v = -\nabla p + \frac{1}{\mathrm{Re}}\nabla^2 v$$
$$\operatorname{div} v = 0$$

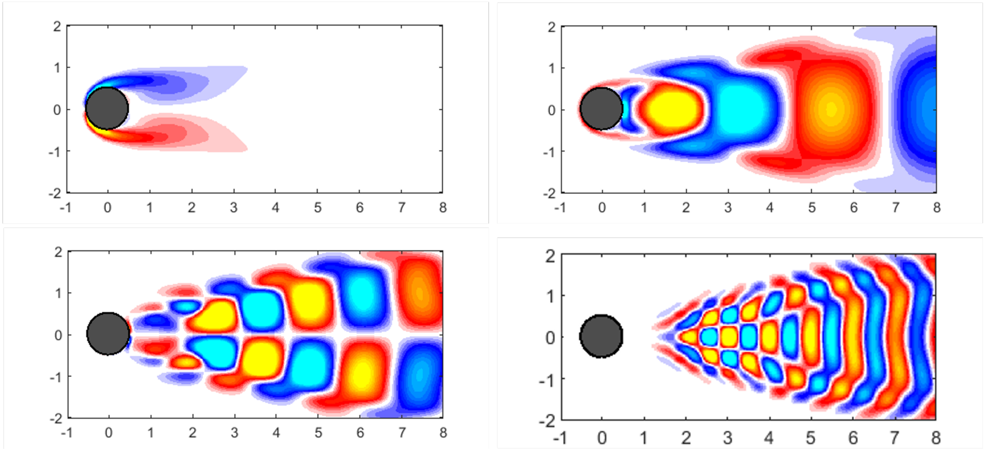
Click

Dynamic mode decomposition: data-driven modeling of complex systems, Kutz, Brunton, Brunton, Proctor, SIAM 2016
Thanks to Emilia Krendelsberger and Karl Worthmann (TU Ilmenau)

Koopman eigenfunctions

Let $\varphi : \Omega \to \mathbb{R}$ and $\lambda \in \mathbb{C}$ satisfy

$$\mathcal{K}\varphi = \lambda\varphi.$$



Model reduction

True

$r = 3$	$r = 11$	$r = 21$

Consequences of linearity

Eigenfunctions and Koopman mode decomposition: Let $\varphi : \Omega \to \mathbb{R}$ satisfy

$$\mathcal{K}\varphi = \lambda\varphi.$$

Consequences of linearity

Eigenfunctions and Koopman mode decomposition: Let $\varphi : \Omega \to \mathbb{R}$ satisfy

$$\mathcal{K}\varphi = \lambda\varphi.$$

- Long-term predictions

$$\varphi(x_k) = (\mathcal{K}^k\varphi)(x_0) = \lambda^k\varphi(x_0).$$

Consequences of linearity

Eigenfunctions and Koopman mode decomposition: Let $\varphi : \Omega \to \mathbb{R}$ satisfy

$$\mathcal{K}\varphi = \lambda\varphi.$$

- Long-term predictions

$$\varphi(x_k) = (\mathcal{K}^k\varphi)(x_0) = \lambda^k\varphi(x_0).$$

- Invariant sets: : If $\varphi(x_0) = 0$, then

$$\varphi(F(x_0)) = (\mathcal{K}\varphi)(x_0) = \lambda\varphi(x_0) = 0 \quad \rightsquigarrow \quad S_0 = \{x : \varphi(x) = 0\} \,\text{invariant}$$

Consequences of linearity

Eigenfunctions and Koopman mode decomposition: Let $\varphi : \Omega \to \mathbb{R}$ satisfy

$$\mathcal{K}\varphi = \lambda\varphi.$$

- Long-term predictions

$$\varphi(x_k) = (\mathcal{K}^k\varphi)(x_0) = \lambda^k\varphi(x_0).$$

- Invariant sets: : If $\varphi(x_0) = 0$, then

$$\varphi(F(x_0)) = (\mathcal{K}\varphi)(x_0) = \lambda\varphi(x_0) = 0 \quad \rightsquigarrow \quad S_0 = \{x : \varphi(x) = 0\}\,\text{invariant}$$

- Conserved quantities: $\lambda = 1$, i.e.,

$$\varphi(x_k) = (\mathcal{K}^k\varphi)(x) = \varphi(x_0) = c \qquad \forall x \in S_c = \{x : \varphi(x) = c\}.$$

Consequences of linearity

Eigenfunctions and Koopman mode decomposition: Let $\varphi : \Omega \to \mathbb{R}$ satisfy

$$\mathcal{K}\varphi = \lambda\varphi.$$

- Long-term predictions

$$\varphi(x_k) = (\mathcal{K}^k\varphi)(x_0) = \lambda^k\varphi(x_0).$$

- Invariant sets: : If $\varphi(x_0) = 0$, then

$$\varphi(F(x_0)) = (\mathcal{K}\varphi)(x_0) = \lambda\varphi(x_0) = 0 \quad \rightsquigarrow \quad S_0 = \{x : \varphi(x) = 0\} \, \text{invariant}$$

- Conserved quantities: $\lambda = 1$, i.e.,

$$\varphi(x_k) = (\mathcal{K}^k\varphi)(x) = \varphi(x_0) = c \qquad \forall x \in S_c = \{x : \varphi(x) = c\}.$$

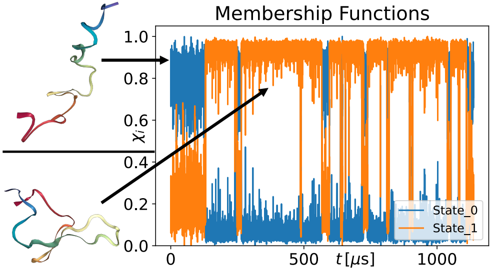
- Multiple eigenvalues: If φ_1 and φ_2 eigenfunctions with the same eigenvalue λ:

$$\frac{\varphi_1(x_k)}{\varphi_2(x_k)} = \frac{\lambda^k\varphi_1(x_0)}{\lambda^k\varphi_2(x_0)} = \frac{\varphi_1(x_0)}{\varphi_2(x_0)} = c \quad \text{invariant}.$$

More applications

Folding kinetics of a 35-amino acid protein

$$\mathrm{d}X_t = \nabla V(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$$

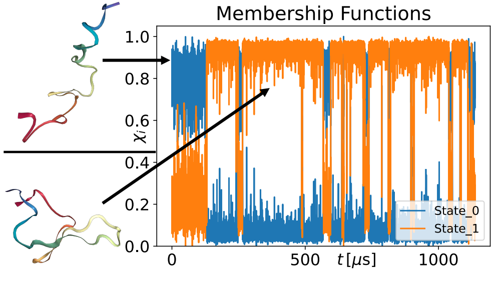


Wu, Nüske, Paul, Klus, Koltai, Noé. The Journal of Chemical Physics 2016
Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

More applications

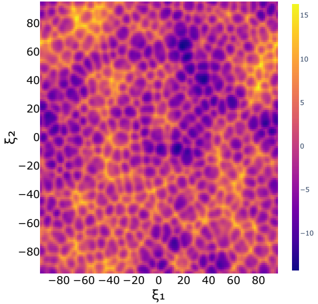
Folding kinetics of a 35-amino acid protein

$$dX_t = \nabla V(X_t)dt + \sigma(X_t)dW_t$$



Kuramoto-Sivashinsky: Chaotic flames

$$\partial_t x + \nabla^2 x + \nabla^4 x + |\nabla x|^2 = 0.$$



Wu, Nüske, Paul, Klus, Koltai, Noé. The Journal of Chemical Physics 2016
Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

Stochastic Differential Equations

Consider the SDE

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t.$$

Suitable assumptions: solution $(X_t)_{t\geq 0}$ exists and is **Markov process**.

Stochastic Differential Equations

Consider the SDE

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t.$$

Suitable assumptions: solution $(X_t)_{t\geq 0}$ exists and is **Markov process**. By

$$\rho_t(x, A) = \mathbb{P}(X_t \in A \,|\, X_0 = x)$$

we denote the associated **transition probability**.

Stochastic Differential Equations

Consider the SDE

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t.$$

Suitable assumptions: solution $(X_t)_{t\geq 0}$ exists and is **Markov process**. By

$$\rho_t(x, A) = \mathbb{P}(X_t \in A \,|\, X_0 = x)$$

we denote the associated **transition probability**. Here, one defines

$$(\mathcal{K}\varphi)(x^0) = \mathbb{E}[\varphi(X_t) \,|\, X_0 = x^0] = \int \varphi(y)\, \rho_t(x^0, dy).$$

Stochastic Differential Equations

Consider the SDE

$$dX_t = f(X_t)dt + \sigma(X_t)dW_t.$$

Suitable assumptions: solution $(X_t)_{t\geq 0}$ exists and is **Markov process**. By

$$\rho_t(x, A) = \mathbb{P}(X_t \in A \,|\, X_0 = x)$$

we denote the associated **transition probability**. Here, one defines

$$(\mathcal{K}\varphi)(x^0) = \mathbb{E}[\varphi(X_t) \,|\, X_0 = x^0] = \int \varphi(y)\, \rho_t(x^0, dy).$$

In the **deterministic case**, $\rho_t(x^0, A) = \delta_{x(t;x^0)}(A)$ such that $(\mathcal{K}\varphi)(x^0) = \varphi(x(t;x^0))$.

Extended Dynamic Mode Decomposition (EDMD)[2]

Consider **data points** $x_0, .., x_{d-1} \in \Omega$, a **dictionary** $\mathbb{V} := \operatorname{span}\{\{\psi_j\}_{j=1}^N\}$ with $\psi_j : \Omega \to \mathbb{R}$,

$$X := \left(\begin{pmatrix} \psi_1(x_0) \\ \vdots \\ \psi_N(x_0) \end{pmatrix} \Bigg| \dots \Bigg| \begin{pmatrix} \psi_1(x_{d-1}) \\ \vdots \\ \psi_N(x_{d-1}) \end{pmatrix} \right), \quad Y := \left(\begin{pmatrix} \psi_1(x^+(x_0)) \\ \vdots \\ \psi_N(x^+(x_0)) \end{pmatrix} \Bigg| \dots \Bigg| \begin{pmatrix} \psi_1(x^+(x_{d-1})) \\ \vdots \\ \psi_N(x^+(x_{d-1})) \end{pmatrix} \right).$$

[2] Williams et al., Journal of Nonlinear Science, 2015

Extended Dynamic Mode Decomposition (EDMD)[2]

Consider **data points** $x_0, .., x_{d-1} \in \Omega$, a **dictionary** $\mathbb{V} := \operatorname{span}\{\{\psi_j\}_{j=1}^N\}$ with $\psi_j : \Omega \to \mathbb{R}$,

$$X := \left(\begin{pmatrix} \psi_1(x_0) \\ \vdots \\ \psi_N(x_0) \end{pmatrix} \Bigg| \dots \Bigg| \begin{pmatrix} \psi_1(x_{d-1}) \\ \vdots \\ \psi_N(x_{d-1}) \end{pmatrix} \right), \quad Y := \left(\begin{pmatrix} \psi_1(x^+(x_0)) \\ \vdots \\ \psi_N(x^+(x_0)) \end{pmatrix} \Bigg| \dots \Bigg| \begin{pmatrix} \psi_1(x^+(x_{d-1}) \\ \vdots \\ \psi_N(x^+(x_{d-1})) \end{pmatrix} \right).$$

Data-based surrogate of the projected Koopman operator:

$$P_{\mathbb{V}} \mathcal{K}_{|\mathbb{V}} \approx \operatorname{argmin}_{K \in \mathbb{R}^{N \times N}} \|Y - KX\|_2^2$$

[2] Williams et al., Journal of Nonlinear Science, 2015

Extended Dynamic Mode Decomposition (EDMD)[2]

Consider **data points** $x_0, .., x_{d-1} \in \Omega$, a **dictionary** $\mathbb{V} := \operatorname{span}\{\{\psi_j\}_{j=1}^N\}$ with $\psi_j : \Omega \to \mathbb{R}$,

$$X := \left(\begin{pmatrix} \psi_1(x_0) \\ \vdots \\ \psi_N(x_0) \end{pmatrix} \middle| \dots \middle| \begin{pmatrix} \psi_1(x_{d-1}) \\ \vdots \\ \psi_N(x_{d-1}) \end{pmatrix} \right), \quad Y := \left(\begin{pmatrix} \psi_1(x^+(x_0)) \\ \vdots \\ \psi_N(x^+(x_0)) \end{pmatrix} \middle| \dots \middle| \begin{pmatrix} \psi_1(x^+(x_{d-1})) \\ \vdots \\ \psi_N(x^+(x_{d-1})) \end{pmatrix} \right).$$

Data-based surrogate of the projected Koopman operator:

$$P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}} \approx \operatorname{argmin}_{K \in \mathbb{R}^{N \times N}} \|Y - KX\|_2^2 =: K_d$$

[2] Williams et al., Journal of Nonlinear Science, 2015

Extended Dynamic Mode Decomposition (EDMD)[2]

Consider **data points** $x_0, .., x_{d-1} \in \Omega$, a **dictionary** $\mathbb{V} := \operatorname{span}\{\{\psi_j\}_{j=1}^N\}$ with $\psi_j : \Omega \to \mathbb{R}$,

$$X := \left(\begin{pmatrix} \psi_1(x_0) \\ \vdots \\ \psi_N(x_0) \end{pmatrix} \middle| \dots \middle| \begin{pmatrix} \psi_1(x_{d-1}) \\ \vdots \\ \psi_N(x_{d-1}) \end{pmatrix} \right), \quad Y := \left(\begin{pmatrix} \psi_1(x^+(x_0)) \\ \vdots \\ \psi_N(x^+(x_0)) \end{pmatrix} \middle| \dots \middle| \begin{pmatrix} \psi_1(x^+(x_{d-1}) \\ \vdots \\ \psi_N(x^+(x_{d-1})) \end{pmatrix} \right).$$

Data-based surrogate of the projected Koopman operator:

$$P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}} \approx \operatorname{argmin}_{K \in \mathbb{R}^{N \times N}} \|Y - KX\|_2^2 =: K_d$$

Approximation error can be split up into two sources:

- **Projection**: $\mathcal{K} - P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}}$ due to finite dictionary $\mathbb{V} := \operatorname{span}\{(\psi_j)_{j=1}^N\}$
- **Estimation**: $P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}} - K_d$ due to finite data points $x_0, .., x_{d-1} \in \Omega$

[2] Williams et al., Journal of Nonlinear Science, 2015

Extended Dynamic Mode Decomposition (EDMD)[2]

Consider **data points** $x_0, .., x_{d-1} \in \Omega$, a **dictionary** $\mathbb{V} := \operatorname{span}\{\{\psi_j\}_{j=1}^N\}$ with $\psi_j : \Omega \to \mathbb{R}$,

$$X := \left(\begin{pmatrix} \psi_1(x_0) \\ \vdots \\ \psi_N(x_0) \end{pmatrix} \Bigg| \dots \Bigg| \begin{pmatrix} \psi_1(x_{d-1}) \\ \vdots \\ \psi_N(x_{d-1}) \end{pmatrix} \right), \quad Y := \left(\begin{pmatrix} \psi_1(x^+(x_0)) \\ \vdots \\ \psi_N(x^+(x_0)) \end{pmatrix} \Bigg| \dots \Bigg| \begin{pmatrix} \psi_1(x^+(x_{d-1}) \\ \vdots \\ \psi_N(x^+(x_{d-1})) \end{pmatrix} \right).$$

Data-based surrogate of the projected Koopman operator:

$$P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}} \approx \operatorname{argmin}_{K \in \mathbb{R}^{N \times N}} \|Y - KX\|_2^2 =: K_d$$

Approximation error can be split up into two sources:

- **Projection**: $\mathcal{K} - P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}}$ due to finite dictionary $\mathbb{V} := \operatorname{span}\{(\psi_j)_{j=1}^N\}$
- **Estimation**: $P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}} - K_d$ due to finite data points $x_0, .., x_{d-1} \in \Omega$

Convergence in infinite-data and infinite-dictionary limit[3]

[3] Korda, Mezić, Journal of Nonlinear Science, 2018
[2] Williams et al., Journal of Nonlinear Science, 2015

Sampling strategies

Data $x_0, \ldots, x_{d-1} \in \Omega$ with successor states $y_0, \ldots, y_{d-1} \in \Omega$ from either

1) i.i.d. sampling: $X_0 = x_j$ drawn i.i.d. w.r.t. μ, $y_j = X_{\Delta t}$.

Sampling strategies

Data $x_0, \ldots, x_{d-1} \in \Omega$ with successor states $y_0, \ldots, y_{d-1} \in \Omega$ from either

1) i.i.d. sampling: $X_0 = x_j$ drawn i.i.d. w.r.t. μ, $y_j = X_{\Delta t}$.
2) sampling from a long trajectory: $x_k = X_{k\Delta t}$, $y_k = X_{(k+1)\Delta t}$.

Sampling strategies

Data $x_0, \ldots, x_{d-1} \in \Omega$ with successor states $y_0, \ldots, y_{d-1} \in \Omega$ from either

1) i.i.d. sampling: $X_0 = x_j$ drawn i.i.d. w.r.t. μ, $y_j = X_{\Delta t}$.
2) sampling from a long trajectory: $x_k = X_{k\Delta t}$, $y_k = X_{(k+1)\Delta t}$.

   Assume there is an **invariant probability measure** μ on Ω, i.e.,

$$\int \rho_t(x, A)\, d\mu(x) \equiv \mu(A)$$

Sampling strategies

Data $x_0, \dots, x_{d-1} \in \Omega$ with successor states $y_0, \dots, y_{d-1} \in \Omega$ from either

1) i.i.d. sampling: $X_0 = x_j$ drawn i.i.d. w.r.t. μ, $y_j = X_{\Delta t}$.
2) sampling from a long trajectory: $x_k = X_{k\Delta t}$, $y_k = X_{(k+1)\Delta t}$.

   Assume there is an **invariant probability measure** μ on Ω, i.e.,

$$\int \rho_t(x, A)\, d\mu(x) \equiv \mu(A)$$

   and that $(X_t)_{t\geq 0}$ is **ergodic**, i.e. for all $t > 0$

$$\rho_t(x, A) = 1\ \forall x \in A \quad \Rightarrow \quad \mu(A) \in \{0, 1\}$$

Theorem

For i.i.d. sampling w.r.t. μ, we have with $\varphi := \sum_{i=1}^{N} \psi_i^2 \in L^2_\mu$

$$\mathbb{P}\big(\|P_{\mathbb{V}}\mathcal{K}|_{\mathbb{V}} - K_d\|_F \leq \varepsilon\big) \gtrsim 1 - \frac{\|\varphi\|^2_{L^2_\mu}}{d\varepsilon^2}.$$

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024.
Nüske, Peitz, Philipp, S., Worthmann, Journal of Nonlinear Science, 2023
Kostic et al., NeurIPS, 2022/2023

Theorem

For i.i.d. sampling w.r.t. μ, we have with $\varphi := \sum_{i=1}^N \psi_i^2 \in L^2_\mu$

$$\mathbb{P}\big(\|P_{\mathbb{V}}\mathcal{K}|_{\mathbb{V}} - K_d\|_F \leq \varepsilon\big) \gtrsim 1 - \frac{\|\varphi\|^2_{L^2_\mu}}{d\varepsilon^2}.$$

If $\varphi \in L^\infty_\mu(\Omega)$,

$$\mathbb{P}\big(\|P_{\mathbb{V}}\mathcal{K}|_{\mathbb{V}} - K_d\|_F \leq \varepsilon\big) \gtrsim 1 - \exp\left(-\frac{d\varepsilon^2}{\|\varphi\|^2_{L^\infty_\mu}}\right),$$

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024.
Nüske, Peitz, Philipp, S., Worthmann, Journal of Nonlinear Science, 2023
Kostic et al., NeurIPS, 2022/2023

Theorem

For i.i.d. sampling w.r.t. μ, we have with $\varphi := \sum_{i=1}^{N} \psi_i^2 \in L_\mu^2$

$$\mathbb{P}\big(\|P_\mathbb{V}\mathcal{K}|_\mathbb{V} - K_d\|_F \leq \varepsilon\big) \gtrsim 1 - \frac{\|\varphi\|_{L_\mu^2}^2}{d\varepsilon^2}.$$

If $\varphi \in L_\mu^\infty(\Omega)$,

$$\mathbb{P}\big(\|P_\mathbb{V}\mathcal{K}|_\mathbb{V} - K_d\|_F \leq \varepsilon\big) \gtrsim 1 - \exp\left(-\frac{d\varepsilon^2}{\|\varphi\|_{L_\mu^\infty}^2}\right),$$

For ergodic sampling, assume $\lambda = 1$ is an isolated eigenvalue of $\mathcal{K}$. Then

$$\mathbb{P}\big(\|P_\mathbb{V}\mathcal{K}|_\mathbb{V} - K_d\|_F \leq \varepsilon\big) \gtrsim 1 - \frac{1}{d\varepsilon^2}.$$

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024.
Nüske, Peitz, Philipp, S., Worthmann, Journal of Nonlinear Science, 2023
Kostic et al., NeurIPS, 2022/2023

Theorem

For i.i.d. sampling w.r.t. μ, we have with $\varphi := \sum_{i=1}^N \psi_i^2 \in L^2_\mu$

$$\mathbb{P}\big(\|P_\mathbb{V}\mathcal{K}|_\mathbb{V} - K_d\|_F \le \varepsilon\big) \gtrsim 1 - \frac{\|\varphi\|^2_{L^2_\mu}}{d\varepsilon^2}.$$

If $\varphi \in L^\infty_\mu(\Omega)$,

$$\mathbb{P}\big(\|P_\mathbb{V}\mathcal{K}|_\mathbb{V} - K_d\|_F \le \varepsilon\big) \gtrsim 1 - \exp\left(-\frac{d\varepsilon^2}{\|\varphi\|^2_{L^\infty_\mu}}\right),$$

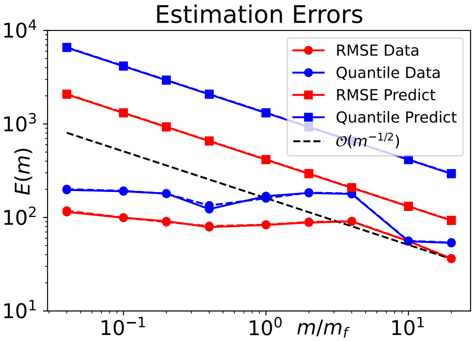
For ergodic sampling, assume $\lambda = 1$ is an isolated eigenvalue of $\mathcal{K}$. Then

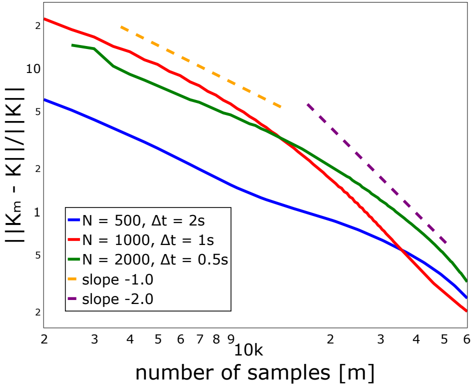
$$\mathbb{P}\big(\|P_\mathbb{V}\mathcal{K}|_\mathbb{V} - K_d\|_F \le \varepsilon\big) \gtrsim 1 - \frac{1}{d\varepsilon^2}.$$

$\lambda = 1$ is never isolated for deterministic systems: Condition related to the spectral measure

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024.
Nüske, Peitz, Philipp, S., Worthmann, Journal of Nonlinear Science, 2023
Kostic et al., NeurIPS, 2022/2023

Back to examples





Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

Intuition: Monte-Carlo Integration

We may write the projected Koopman operator via[1]

$$P_{\mathbb{V}}\mathcal{K}|_{\mathbb{V}} = C^{-1}A \qquad \text{with} \qquad C_{i,j} = \langle \psi_i, \psi_j \rangle_{L^2(\Omega)}, \quad A_{i,j} = \langle \psi_i, \mathcal{K}\psi_j \rangle_{L^2(\Omega)}$$

[1] Klus, Nüske, et al., Physica D, 2018

Intuition: Monte-Carlo Integration

We may write the projected Koopman operator via[1]

$$P_{\mathbb{V}}\mathcal{K}|_{\mathbb{V}} = C^{-1}A \qquad \text{with} \qquad C_{i,j} = \langle \psi_i, \psi_j \rangle_{L^2(\Omega)}, \quad A_{i,j} = \langle \psi_i, \mathcal{K}\psi_j \rangle_{L^2(\Omega)}$$

Data-driven approximation:

$$K_d = \operatorname{argmin}_{K \in \mathbb{R}^{N\times N}} \|Y - KX\|_2^2$$

[1] Klus, Nüske, et al., Physica D, 2018

Intuition: Monte-Carlo Integration

We may write the projected Koopman operator via[1]

$$P_{\mathbb{V}}\mathcal{K}|_{\mathbb{V}} = C^{-1}A \qquad \text{with} \qquad C_{i,j} = \langle \psi_i, \psi_j \rangle_{L^2(\Omega)}, \quad A_{i,j} = \langle \psi_i, \mathcal{K}\psi_j \rangle_{L^2(\Omega)}$$

Data-driven approximation:

$$K_d = (XX^\top)^{-1}XY^\top$$

[1] Klus, Nüske, et al., Physica D, 2018

Intuition: Monte-Carlo Integration

We may write the projected Koopman operator via[1]

$$P_{\mathbb{V}}\mathcal{K}|_{\mathbb{V}} = C^{-1}A \qquad \text{with} \qquad C_{i,j} = \langle \psi_i, \psi_j \rangle_{L^2(\Omega)}, \quad A_{i,j} = \langle \psi_i, \mathcal{K}\psi_j \rangle_{L^2(\Omega)}$$

Data-driven approximation:

$$K_d = (XX^\top)^{-1}XY^\top = C_d^{-1}A_d$$

with

$$(C_d)_{i,j} = \tfrac{1}{d}\sum_{k=0}^{d-1}\psi_i(x_k)\psi_j(x_k), \quad (A_d)_{i,j} = \tfrac{1}{d}\sum_{k=0}^{d-1}\psi_i(x_k)\cdot(\mathcal{K}\psi_j)(x_k)$$

as

$$X := \left(\begin{pmatrix} \psi_1(x_0) \\ \vdots \\ \psi_N(x_0) \end{pmatrix} \middle| \dots \middle| \begin{pmatrix} \psi_1(x_{d-1}) \\ \vdots \\ \psi_N(x_{d-1}) \end{pmatrix} \right), \quad Y := \left(\begin{pmatrix} \psi_1(x^+(x_0)) \\ \vdots \\ \psi_N(x^+(x_0)) \end{pmatrix} \middle| \dots \middle| \begin{pmatrix} \psi_1(x^+(x_{d-1})) \\ \vdots \\ \psi_N(x^+(x_{d-1})) \end{pmatrix} \right).$$

[1] Klus, Nüske, et al., Physica D, 2018

Intuition continued

It's all about $\frac{1}{d}\sum_{k=1}^{d-1}\psi_i(x_k)\psi_j(x_k) - \int \psi_i\psi_j$

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

Intuition continued

It's all about $\frac{1}{d}\sum_{k=1}^{d-1}\psi_i(x_k)\psi_j(x_k) - \int \psi_i\psi_j$

Step 1: Estimate variance:

1. i.i.d. sampling $\sigma^2_{d,i,j} = \frac{1}{d}\sigma^2_{i,j}$
2. ergodic sampling $\sigma^2_{d,i,j} = \frac{1}{d}\Big(\underbrace{\sigma^2_{\infty,i,j}}_{\text{asymptotic variance}} - \underbrace{R^d_{i,j}}_{\text{remainder term}} \Big)$

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

Intuition continued

It's all about $\frac{1}{d}\sum_{k=1}^{d-1} \psi_i(x_k)\psi_j(x_k) - \int \psi_i \psi_j$

Step 1: Estimate variance:

1. i.i.d. sampling $\sigma^2_{d,i,j} = \frac{1}{d}\sigma^2_{i,j}$
2. ergodic sampling $\sigma^2_{d,i,j} = \frac{1}{d}\Big(\underbrace{\sigma^2_{\infty,i,j}}_{\text{asymptotic variance}} - \underbrace{R^d_{i,j}}_{\text{remainder term}} \Big)$

Step 2: Apply concentration inequality (Chebychev, Hoeffding...)

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

Intuition continued

It's all about $\frac{1}{d}\sum_{k=1}^{d-1}\psi_i(x_k)\psi_j(x_k) - \int \psi_i\psi_j$

Step 1: Estimate variance:

1. i.i.d. sampling $\sigma^2_{d,i,j} = \frac{1}{d}\sigma^2_{i,j}$
2. ergodic sampling $\sigma^2_{d,i,j} = \frac{1}{d}\Big(\underbrace{\sigma^2_{\infty,i,j}}_{\text{asymptotic variance}} - \underbrace{R^d_{i,j}}_{\text{remainder term}}\Big)$

Step 2: Apply concentration inequality (Chebychev, Hoeffding...)

Decay of $\frac{1}{d}\cdot$remainder term: Ergodic theory and mixing conditions.

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

Intuition continued

It's all about $\frac{1}{d}\sum_{k=1}^{d-1}\psi_i(x_k)\psi_j(x_k) - \int \psi_i\psi_j$

Step 1: Estimate variance:

1. i.i.d. sampling $\sigma^2_{d,i,j} = \frac{1}{d}\sigma^2_{i,j}$
2. ergodic sampling $\sigma^2_{d,i,j} = \frac{1}{d}\Big(\underbrace{\sigma^2_{\infty,i,j}}_{\text{asymptotic variance}} - \underbrace{R^d_{i,j}}_{\text{remainder term}}\Big)$

Step 2: Apply concentration inequality (Chebychev, Hoeffding...)

Decay of $\frac{1}{d}$·remainder term: Ergodic theory and mixing conditions.

Approximation error:

- **Projection**: $\mathcal{K} - P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}}$ due to finite dictionary $\mathbb{V} := \operatorname{span}\{(\psi_j)_{j=1}^N\}$
- **Estimation**: $P_{\mathbb{V}}\mathcal{K}_{|\mathbb{V}} - K_d$ due to finite data points $x_0, .., x_{d-1} \in \Omega$

Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

Kernel Extended Dynamic Mode Decomposition

A **Reproducing Kernel Hilbert Space** $\mathbb{H}$ is a

- Hilbert space of functions $f : \Omega \to \mathbb{R}$
- with s.p.d. kernel $k : \Omega \times \Omega \to \mathbb{R}$ with $k(x, \cdot) \in \mathbb{H}$ for all $x \in \Omega$ and

$$\forall \varphi \in \mathbb{H} : \quad \varphi(x) = \langle \varphi, k(x, \cdot) \rangle$$ **reproducing property**

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Kernel Extended Dynamic Mode Decomposition

A **Reproducing Kernel Hilbert Space** $\mathbb{H}$ is a

- Hilbert space of functions $f : \Omega \to \mathbb{R}$
- with s.p.d. kernel $k : \Omega \times \Omega \to \mathbb{R}$ with $k(x, \cdot) \in \mathbb{H}$ for all $x \in \Omega$ and

$$\forall \varphi \in \mathbb{H} : \quad \varphi(x) = \langle \varphi, k(x, \cdot) \rangle$$ **reproducing property**

Dictionary: For data points $\mathcal{X} = \{x_1 \ldots, x_d\} \subset \Omega$ set $V_{\mathcal{X}} := \mathrm{span}\{k(x_1, \cdot), k(x_2, \cdot), \ldots, k(x_d, \cdot)\} \subset \mathbb{H}$

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Kernel Extended Dynamic Mode Decomposition

A **Reproducing Kernel Hilbert Space** $\mathbb{H}$ is a

- Hilbert space of functions $f : \Omega \to \mathbb{R}$
- with s.p.d. kernel $k : \Omega \times \Omega \to \mathbb{R}$ with $k(x, \cdot) \in \mathbb{H}$ for all $x \in \Omega$ and

$$\forall \varphi \in \mathbb{H} : \quad \varphi(x) = \langle \varphi, k(x, \cdot) \rangle \qquad \textbf{reproducing property}$$

Dictionary: For data points $\mathcal{X} = \{x_1 \ldots, x_d\} \subset \Omega$ set $V_{\mathcal{X}} := \operatorname{span}\{k(x_1, \cdot), k(x_2, \cdot), \ldots, k(x_d, \cdot)\} \subset \mathbb{H}$

$$X = \begin{bmatrix} k(x_0, x_0) & \ldots & k(x_0, x_{d-1}) \\ \vdots & \ddots & \vdots \\ k(x_{d-1}, x_0) & \ldots & k(x_{d-1}, x_{d-1}) \end{bmatrix}, \quad Y = \begin{bmatrix} k(x_0, {x_0}^+) & \ldots & k(x_0, x_{d-1}^+) \\ \vdots & \ddots & \vdots \\ k(x_{d-1}, x_0^+) & \ldots & k(x_{d-1}, x_{d-1}^+) \end{bmatrix}.$$

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Kernel Extended Dynamic Mode Decomposition

A **Reproducing Kernel Hilbert Space** $\mathbb{H}$ is a

- Hilbert space of functions $f : \Omega \to \mathbb{R}$
- with s.p.d. kernel $k : \Omega \times \Omega \to \mathbb{R}$ with $k(x, \cdot) \in \mathbb{H}$ for all $x \in \Omega$ and

$$\forall \varphi \in \mathbb{H} : \quad \varphi(x) = \langle \varphi, k(x, \cdot) \rangle$$ **reproducing property**

Dictionary: For data points $\mathcal{X} = \{x_1 \ldots, x_d\} \subset \Omega$ set $V_{\mathcal{X}} := \operatorname{span}\{k(x_1, \cdot), k(x_2, \cdot), \ldots, k(x_d, \cdot)\} \subset \mathbb{H}$

$$X = \begin{bmatrix} k(x_0, x_0) & \ldots & k(x_0, x_{d-1}) \\ \vdots & \ddots & \vdots \\ k(x_{d-1}, x_0) & \ldots & k(x_{d-1}, x_{d-1}) \end{bmatrix}, \quad Y = \begin{bmatrix} k(x_0, {x_0}^+) & \ldots & k(x_0, x_{d-1}^+) \\ \vdots & \ddots & \vdots \\ k(x_{d-1}, x_0^+) & \ldots & k(x_{d-1}, x_{d-1}^+) \end{bmatrix}.$$

kEDMD approximant

$$K_d = (XX^\top)^{-1} X^\top Y = X^{-1} X^{-1} XY = X^{-1} Y$$

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Representation of the Koopman operator

Compatibility assumptions such that $\mathbb{H} \overset{d}{\hookrightarrow} L^2_\mu(\Omega)$

Representation of the Koopman operator

Compatibility assumptions such that $\mathbb{H} \overset{d}{\hookrightarrow} L^2_\mu(\Omega)$

Central concept: Mercer integral operator $\mathcal{E} : L^2_\mu(\Omega) \to \mathbb{H}$

$$\mathcal{E}\psi = \int k(x, \cdot)\psi(x)\, d\mu(x).$$

Representation of the Koopman operator

Compatibility assumptions such that $\mathbb{H} \overset{d}{\hookrightarrow} L^2_\mu(\Omega)$

Central concept: Mercer integral operator $\mathcal{E} : L^2_\mu(\Omega) \to \mathbb{H}$

$$\mathcal{E}\psi = \int k(x,\cdot)\psi(x)\,d\mu(x).$$

with

$$\langle \mathcal{E}\psi, \eta\rangle = \int \psi(x)\langle k(x,\cdot),\eta\rangle\,d\mu(x) = \int \psi(x)\eta(x)\,d\mu(x) = \langle \psi,\eta\rangle_\mu \qquad \forall \psi \in L^2_\mu(\Omega), \eta \in \mathbb{H}$$

Representation of the Koopman operator

Compatibility assumptions such that $\mathbb{H} \overset{d}{\hookrightarrow} L^2_\mu(\Omega)$

Central concept: Mercer integral operator $\mathcal{E} : L^2_\mu(\Omega) \to \mathbb{H}$

$$\mathcal{E}\psi = \int k(x, \cdot)\psi(x)\, d\mu(x).$$

with

$$\langle \mathcal{E}\psi, \eta \rangle = \int \psi(x)\langle k(x,\cdot), \eta\rangle\, d\mu(x) = \int \psi(x)\eta(x)\, d\mu(x) = \langle \psi, \eta\rangle_\mu \qquad \forall \psi \in L^2_\mu(\Omega), \eta \in \mathbb{H}$$

Adjoint $\mathcal{E}^* : \mathbb{H} \to L^2_\mu(\Omega)$ is the compact inclusion operator from $\mathbb{H}$ into $L^2_\mu(\Omega)$

$$\mathcal{E}^*\eta = \eta, \qquad \eta \in \mathbb{H}.$$

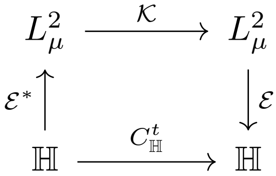
Koopman through the lens of the Mercer operator

Set

$$C_{\mathbb{H}}^t = \mathcal{E}\mathcal{K}\mathcal{E}^*$$

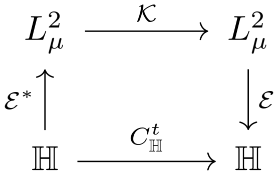
such that

$$\langle \eta, C_{\mathbb{H}}^t \psi \rangle = \langle \eta, \mathcal{K}\psi \rangle_\mu$$



Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Koopman through the lens of the Mercer operator



Set

$$C^t_{\mathbb{H}} = \mathcal{E}\mathcal{K}\mathcal{E}^*$$

such that

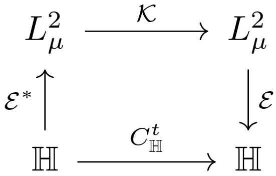
$$\langle \eta, C^t_{\mathbb{H}} \psi \rangle = \langle \eta, \mathcal{K}\psi \rangle_\mu$$

Setting $C_{\mathbb{H}} = \mathcal{E}\mathcal{E}^*$

$$(\mathcal{E}^*)^{-1}\mathcal{K}\mathcal{E}^* = (\mathcal{E}\mathcal{E}^*)^{-1}\mathcal{E}\mathcal{E}^*(\mathcal{E}^*)^{-1}\mathcal{K}\mathcal{E}^* = C_{\mathbb{H}}^{-1}\mathcal{E}\mathcal{K}\mathcal{E}^* = C_{\mathbb{H}}^{-1}C^t_{\mathbb{H}}$$

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Koopman through the lens of the Mercer operator



Set

$$C^t_{\mathbb{H}} = \mathcal{E}\mathcal{K}\mathcal{E}^*$$

such that

$$\langle \eta, C^t_{\mathbb{H}}\psi \rangle = \langle \eta, \mathcal{K}\psi \rangle_\mu$$

Setting $C_{\mathbb{H}} = \mathcal{E}\mathcal{E}^*$

$$(\mathcal{E}^*)^{-1}\mathcal{K}\mathcal{E}^* = (\mathcal{E}\mathcal{E}^*)^{-1}\mathcal{E}\mathcal{E}^*(\mathcal{E}^*)^{-1}\mathcal{K}\mathcal{E}^* = C_{\mathbb{H}}^{-1}\mathcal{E}\mathcal{K}\mathcal{E}^* = C_{\mathbb{H}}^{-1}C^t_{\mathbb{H}}$$

Strategy: Approximate $C_{\mathbb{H}}^{-1}$, $C^t_{\mathbb{H}}$ and transfer to Koopman operator

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

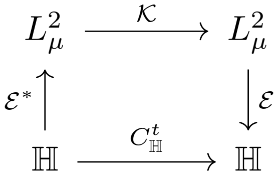
Koopman through the lens of the Mercer operator

Set

$$C_{\mathbb{H}}^t = \mathcal{E}\mathcal{K}\mathcal{E}^*$$

such that

$$\langle \eta, C_{\mathbb{H}}^t \psi \rangle = \langle \eta, \mathcal{K}\psi \rangle_\mu$$



Setting $C_{\mathbb{H}} = \mathcal{E}\mathcal{E}^*$

$$(\mathcal{E}^*)^{-1}\mathcal{K}\mathcal{E}^* = (\mathcal{E}\mathcal{E}^*)^{-1}\mathcal{E}\mathcal{E}^*(\mathcal{E}^*)^{-1}\mathcal{K}\mathcal{E}^* = C_{\mathbb{H}}^{-1}\mathcal{E}\mathcal{K}\mathcal{E}^* = C_{\mathbb{H}}^{-1}C_{\mathbb{H}}^t$$

Strategy: Approximate $C_{\mathbb{H}}^{-1}$, $C_{\mathbb{H}}^t$ and transfer to Koopman operator

Caution: $\mathcal{E}^{-1}, \mathcal{E}^{-*}$ are unbounded

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Empirical estimator for $C^t_{\mathbb{H}}$

We compute

$$C^t_{\mathbb{H}}\psi = \mathcal{E}\mathcal{K}\mathcal{E}^*\psi = \int (\mathcal{K}\psi)(x)k(x,\cdot)\,d\mu(x) = \int\int \underbrace{\psi(y)k(x,\cdot)}_{=:C_{xy}\psi}\rho_t(x,dy)d\mu(x).$$

Empirical estimator for $C^t_{\mathbb{H}}$

We compute

$$C^t_{\mathbb{H}}\psi = \mathcal{E}\mathcal{K}\mathcal{E}^*\psi = \int (\mathcal{K}\psi)(x)k(x,\cdot)\,d\mu(x) = \int\int \underbrace{\psi(y)k(x,\cdot)}_{=:C_{xy}\psi}\rho_t(x,dy)d\mu(x).$$

Empirical estimator of time-lagged cross-covariance:

$$\hat{C}^{d,t}_{\mathbb{H}} := \frac{1}{d}\sum_{k=0}^{d-1} C_{x_k,y_k} \qquad \text{with matrix rep.} \qquad \frac{1}{d}(Y)_{ij} = k(x_i,y_j)$$

Empirical estimator for $C_{\mathbb{H}}^t$

We compute

$$C_{\mathbb{H}}^t \psi = \mathcal{E}\mathcal{K}\mathcal{E}^*\psi = \int (\mathcal{K}\psi)(x)k(x,\cdot)\,d\mu(x) = \int\int \underbrace{\psi(y)k(x,\cdot)}_{=:C_{xy}\psi} \rho_t(x,dy)d\mu(x).$$

Empirical estimator of time-lagged cross-covariance:

$$\hat{C}_{\mathbb{H}}^{d,t} := \frac{1}{d}\sum_{k=0}^{d-1} C_{x_k,y_k} \qquad \text{with matrix rep.} \qquad \frac{1}{d}(Y)_{ij} = k(x_i, y_j)$$

Empirical estimator of kernel covariance:

$$C_{\mathbb{H}} = \mathcal{E}^*\mathcal{E} = \int C_{xx}\,\mathrm{d}\mu(x) \approx \frac{1}{d}\sum_{i=0}^{d-1} C_{x_k,x_k} =: \hat{C}_{\mathbb{H}}^d \qquad \text{with matrix rep.} \qquad \frac{1}{d}(X)_{ij} = k(x_i, x_j)$$

Finite-data error bound

Theorem

For all $\varepsilon > 0$*, there is* $d_0 \in \mathbb{N}$ *such that for all* $d \geq d_0$

$$\mathbb{P}(\|C^t_{\mathbb{H}} - \hat{C}^{d,t}_{\mathbb{H}}\|_{HS} > \varepsilon) \leq \frac{\mathbb{E}_0(t) + R(d)}{d\varepsilon^2},$$

Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024
Mollenhauer, Klus, Schütte, Koltai, Journal of Machine Learning Research, 2022.

Main tool: Variance representations[1]

$$\mathbb{E}\left[\|\widehat{C}_{\mathbb{H}}^{d,t} - C_{\mathbb{H}}^{t}\|_{HS}^{2}\right] = \frac{1}{d}\Bigg[\mathbb{E}_0(t) + \underbrace{2\sum_{k=1}^{d-1} \tfrac{d-k}{d} \cdot \mathbb{E}\left[\langle C_{z_k} - C_{\mathbb{H}}^{t}, C_{z_0} - C_{\mathbb{H}}^{t}\rangle_{HS}\right]}_{R(d)}\Bigg],$$

[2] Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024

[1] Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Main tool: Variance representations[1]

$$\mathbb{E}\left[\|\widehat{C}_{\mathbb{H}}^{d,t} - C_{\mathbb{H}}^{t}\|_{HS}^{2}\right] = \frac{1}{d}\Bigg[\mathbb{E}_0(t) + \underbrace{2\sum_{k=1}^{d-1}\tfrac{d-k}{d}\cdot\mathbb{E}\left[\langle C_{z_k} - C_{\mathbb{H}}^{t}, C_{z_0} - C_{\mathbb{H}}^{t}\rangle_{HS}\right]}_{R(d)}\Bigg],$$

For i.i.d. sampling, $R(d) \equiv 0$.

[2] Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024
[1] Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Main tool: Variance representations[1]

$$\mathbb{E}\left[\|\widehat{C}_{\mathbb{H}}^{d,t} - C_{\mathbb{H}}^{t}\|_{HS}^2\right] = \frac{1}{d}\Bigg[\mathbb{E}_0(t) + \underbrace{2\sum_{k=1}^{d-1} \tfrac{d-k}{d} \cdot \mathbb{E}\left[\langle C_{z_k} - C_{\mathbb{H}}^{t}, C_{z_0} - C_{\mathbb{H}}^{t}\rangle_{HS}\right]}_{R(d)}\Bigg],$$

For i.i.d. sampling, $R(d) \equiv 0$.

For ergodic sampling, if 1 is an isolated eigenvalue of $\mathcal{K}$, then with $\mathcal{K}_0 = \mathcal{K}|_{\mathbb{1}^\perp}$

$$R(d) \leq 8\mathbb{E}_0(t)\|(I - \mathcal{K}_0)^{-2}\|$$

[2]Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024
[1]Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Main tool: Variance representations[1]

$$\mathbb{E}\left[\|\widehat{C}_{\mathbb{H}}^{d,t} - C_{\mathbb{H}}^{t}\|_{HS}^2\right] = \frac{1}{d}\Bigg[\mathbb{E}_0(t) + \underbrace{2\sum_{k=1}^{d-1} \tfrac{d-k}{d} \cdot \mathbb{E}\left[\langle C_{z_k} - C_{\mathbb{H}}^{t}, C_{z_0} - C_{\mathbb{H}}^{t}\rangle_{HS}\right]}_{R(d)}\Bigg],$$

For i.i.d. sampling, $R(d) \equiv 0$.

For ergodic sampling, if 1 is an isolated eigenvalue of $\mathcal{K}$, then with $\mathcal{K}_0 = \mathcal{K}|_{\mathbb{1}^\perp}$

$$R(d) \leq 8\mathbb{E}_0(t)\|(I - \mathcal{K}_0)^{-2}\|$$

Similar variance representations for arbitrary dictionaries[2]

[2] Philipp, S., Boshoff, Peitz, Nüske, Worthmann, arXiv:2402.02494, 2024
[1] Philipp, S., Worthmann, Peitz, Nüske, Applied and Computational Harmonic Analysis, 2024

Transfering this to Koopman operator

We have $C^t_{\mathbb{H}} = \mathcal{E}\mathcal{K}\mathcal{E}^*$,

Transfering this to Koopman operator

We have $C^t_{\mathbb{H}} = \mathcal{E}\mathcal{K}\mathcal{E}^*$, but $\mathcal{E}^{-1}, \mathcal{E}^{-*}$ are unbounded.

Transfering this to Koopman operator

We have $C_{\mathbb{H}}^t = \mathcal{E}\mathcal{K}\mathcal{E}^*$, but $\mathcal{E}^{-1}, \mathcal{E}^{-*}$ are unbounded.

Remedy: Mercer basis of eigenfunctions of trace class operator $\mathcal{C}_{\mathbb{H}} = \mathcal{E}\mathcal{E}^*$ (f_j, λ_j), $\lambda_j \to 0$ with

(f_j) is ONB of $\mathbb{H}$ and $(e_j) = (\lambda_j^{-1/2} f_j)$ is ONB of L^2_μ

Transfering this to Koopman operator

We have $C^t_{\mathbb{H}} = \mathcal{E}\mathcal{K}\mathcal{E}^*$, but $\mathcal{E}^{-1}, \mathcal{E}^{-*}$ are unbounded.
Remedy: Mercer basis of eigenfunctions of trace class operator $\mathcal{C}_{\mathbb{H}} = \mathcal{E}\mathcal{E}^*$ (f_j, λ_j), $\lambda_j \to 0$ with

$$(f_j) \text{ is ONB of } \mathbb{H} \qquad \text{and} \qquad (e_j) = (\lambda_j^{-1/2} f_j) \text{ is ONB of } L^2_\mu$$

Then, for $\psi \in \mathbb{H}$,

$$\mathcal{K}\psi = \sum_{j=1}^{\infty} \langle \mathcal{K}\psi, e_j \rangle_\mu e_j = \underbrace{\sum_{j=1}^{N} \langle \mathcal{K}\psi, e_j \rangle_\mu e_j}_{=:\mathcal{K}_N \psi} + \sum_{j=N+1}^{\infty} \langle \mathcal{K}\psi, e_j \rangle_\mu e_j$$

Transfering this to Koopman operator

We have $C_{\mathbb{H}}^t = \mathcal{E}\mathcal{K}\mathcal{E}^*$, but $\mathcal{E}^{-1}, \mathcal{E}^{-*}$ are unbounded.
Remedy: Mercer basis of eigenfunctions of trace class operator $C_{\mathbb{H}} = \mathcal{E}\mathcal{E}^*$ (f_j, λ_j), $\lambda_j \to 0$ with

$$(f_j) \text{ is ONB of } \mathbb{H} \qquad \text{and} \qquad (e_j) = (\lambda_j^{-1/2} f_j) \text{ is ONB of } L^2_\mu$$

Then, for $\psi \in \mathbb{H}$,

$$\mathcal{K}\psi = \sum_{j=1}^{\infty} \langle \mathcal{K}\psi, e_j \rangle_\mu e_j = \underbrace{\sum_{j=1}^{N} \langle \mathcal{K}\psi, e_j \rangle_\mu e_j}_{=:\mathcal{K}_N\psi} + \sum_{j=N+1}^{\infty} \langle \mathcal{K}\psi, e_j \rangle_\mu e_j$$

and we approximate with $\hat{e}_j = (\hat{\lambda}_j^{-1/2} \hat{f}_j)$ are eigenfunctions of $\hat{C}_{\mathbb{H}}^d = \frac{1}{d} \sum_{k=0}^{m-1} C_{x_k, x_k}$

$$\mathcal{K}_N\psi = \sum_{j=1}^{N} \langle C_{\mathbb{H}}^t \psi, e_j \rangle_{\mathbb{H}} e_j \approx \sum_{j=1}^{N} \langle \hat{C}_{\mathbb{H}}^{t,d} \psi, \hat{e}_j \rangle_{\mathbb{H}} e_j =: \hat{\mathcal{K}}_N^d \psi$$

Transfering this to Koopman operator

We have $C^t_{\mathbb{H}} = \mathcal{E}\mathcal{K}\mathcal{E}^*$, but $\mathcal{E}^{-1}, \mathcal{E}^{-*}$ are unbounded.

Remedy: Mercer basis of eigenfunctions of trace class operator $\mathcal{C}_{\mathbb{H}} = \mathcal{E}\mathcal{E}^*$ (f_j, λ_j), $\lambda_j \to 0$ with

$$(f_j) \text{ is ONB of } \mathbb{H} \qquad \text{and} \qquad (e_j) = (\lambda_j^{-1/2} f_j) \text{ is ONB of } L^2_\mu$$

Then, for $\psi \in \mathbb{H}$,

$$\mathcal{K}\psi = \sum_{j=1}^{\infty} \langle \mathcal{K}\psi, e_j\rangle_\mu e_j = \underbrace{\sum_{j=1}^{N} \langle \mathcal{K}\psi, e_j\rangle_\mu e_j}_{=:\mathcal{K}_N\psi} + \sum_{j=N+1}^{\infty} \langle \mathcal{K}\psi, e_j\rangle_\mu e_j$$

and we approximate with $\hat{e}_j = (\hat{\lambda}_j^{-1/2} \hat{f}_j)$ are eigenfunctions of $\hat{C}^d_{\mathbb{H}} = \frac{1}{d}\sum_{k=0}^{m-1} C_{x_k, x_k}$

$$\mathcal{K}_N\psi = \sum_{j=1}^{N} \langle C^t_{\mathbb{H}}\psi, e_j\rangle_{\mathbb{H}} e_j \approx \sum_{j=1}^{N} \langle \hat{C}^{t,d}_{\mathbb{H}}\psi, \hat{e}_j\rangle_{\mathbb{H}} e_j =: \hat{\mathcal{K}}^d_N\psi$$

Back to kEDMD: $\hat{K}^d_N$ corresponds to K_d via $K_d = X^{-1}Y \approx X^\dagger_N Y$ with a rank-N truncation of X.

Main result: Estimation error

Theorem

Let $N \in \mathbb{N}$ and assume the gap condition

$$\gamma_N := \min_{j=1,\dots,N} \frac{\lambda_j - \lambda_{j+1}}{2} > 0.$$

Then, for each error bound $\varepsilon \in (0, \delta_N)$ and probabilistic tolerance $\delta \in (0,1)$ and

$$d \geq \max\{N, \frac{\mathbb{E}_0(t) + R(d)}{\varepsilon^2 \delta}\}$$

we have with probability at least $1-\delta$

$$\|\mathcal{K}_N - \hat{\mathcal{K}}_N^m\|_{\mathbb{H}\to L^2_\mu(\Omega)} \lesssim \left(\frac{1}{\sqrt{\lambda_N}} + \frac{N+1}{\gamma_N \lambda_N}\right)\varepsilon.$$

Invariance of the RKHS: Projection error

Corollary

Assume in addition that $\|\mathcal{K}_{\mathbb{H}}\| < \infty$. *Then for* $N \in \mathbb{N}$, $\varepsilon > 0$, $\delta \in (0,1)$, *there is* $m_0 \in \mathbb{N}$ *such that for all* $d \geq d_0$

$$\|\mathcal{K} - \hat{\mathcal{K}}_N^d\|_{\mathbb{H}\to L^2_\mu(\Omega)} \lesssim \left[\frac{1}{\sqrt{\lambda_N}} + \frac{N+1}{\gamma_N \lambda_N}\right]\varepsilon + \sqrt{\lambda_{N+1}}\,\|K\|_{\mathbb{H}\to\mathbb{H}}.$$

From L^2 to L^∞ bounds

With $P_{V_\mathcal{X}}$ $\mathbb{H}$-orthogonal projection onto $V_\mathcal{X} = \operatorname{span}\{k(x_i,\cdot), i = 0,\ldots,d-1\}$

$$\mathcal{K} \approx K_d := P_{V_\mathcal{X}} \mathcal{K} P_{V_\mathcal{X}} \quad \rightsquigarrow \quad \text{matrix representation} \quad K_X^{-1} K_{X,X^+}$$

with

$$(K_X)_{ij} = k(x_i, x_j), \qquad (K_{X,X^+})_{ij} = k(x_i, F(x_j)).$$

Wendland, Advances in Computational Mathematics 1995

From L^2 to L^∞ bounds

With $P_{V_\mathcal{X}}$ $\mathbb{H}$-orthogonal projection onto $V_\mathcal{X} = \operatorname{span}\{k(x_i, \cdot), i = 0, \ldots, d-1\}$

$$\mathcal{K} \approx K_d := P_{V_\mathcal{X}} \mathcal{K} P_{V_\mathcal{X}} \quad \rightsquigarrow \quad \text{matrix representation} \quad K_X^{-1} K_{X,X^+}$$

with

$$(K_X)_{ij} = k(x_i, x_j), \qquad (K_{X,X^+})_{ij} = k(x_i, F(x_j)).$$

Here: compactly supported radially symmetric Wendland kernel:

$$\mathbb{H} \cong H^{\sigma(p)}(\Omega) \qquad p \sim \text{smoothness}$$

Wendland, Advances in Computational Mathematics 1995

From L^2 to L^∞ bounds

With $P_{V_\mathcal{X}}$ $\mathbb{H}$-orthogonal projection onto $V_\mathcal{X} = \operatorname{span}\{k(x_i,\cdot), i = 0,\ldots,d-1\}$

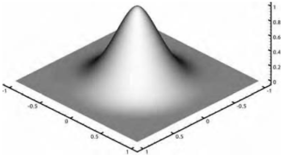
$$\mathcal{K} \approx K_d := P_{V_\mathcal{X}} \mathcal{K} P_{V_\mathcal{X}} \quad \rightsquigarrow \quad \text{matrix representation} \quad K_X^{-1} K_{X,X^+}$$

with

$$(K_X)_{ij} = k(x_i, x_j), \qquad (K_{X,X^+})_{ij} = k(x_i, F(x_j)).$$

Here: compactly supported radially symmetric Wendland kernel:

$$\mathbb{H} \cong H^{\sigma(p)}(\Omega) \qquad p \sim \text{smoothness}$$



Projection error $P_{V_\mathcal{X}} - I$ controlled by fill distance

$$h_\mathcal{X} := \sup_{x\in\Omega} \min_{1\le i\le d} \|x - x_i\|_2$$

Wendland, Advances in Computational Mathematics 1995

Error bound on $\mathcal{K} - P_{V_\mathcal{X}} \mathcal{K} P_{V_\mathcal{X}}$

Theorem[2]

If flow is in C^m, then, for $\sigma(p) \leq m$, $\mathcal{K} H^{\sigma(p)}(\Omega) \subset H^{\sigma(p)}(\Omega)$, and, for $f \in H^{\sigma(p)}(\Omega)$,

$$\|\mathcal{K}f - K_d f\|_\infty \lesssim h_\mathcal{X}^{p+1/2} \|f\|_{H^{\sigma(p)}(\Omega)}.$$

[2] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Error bound on $\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}$

Theorem[2]

If flow is in C^m, then, for $\sigma(p) \leq m$, $\mathcal{K}H^{\sigma(p)}(\Omega) \subset H^{\sigma(p)}(\Omega)$, and, for $f \in H^{\sigma(p)}(\Omega)$,

$$\|\mathcal{K}f - K_d f\|_\infty \lesssim h_\mathcal{X}^{p+1/2}\|f\|_{H^{\sigma(p)}(\Omega)}.$$

$$\left\|\mathcal{K} - K_d\right\|_{\mathbb{H}\to C_b} = \left\|\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}\right\|_{\mathbb{H}\to C_b} \lesssim (1 + \|\mathcal{K}\|_{\mathbb{H}\to\mathbb{H}})\|I - P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b}$$

[2] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Error bound on $\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}$

Theorem[2]

If flow is in C^m, then, for $\sigma(p) \leq m$, $\mathcal{K}H^{\sigma(p)}(\Omega) \subset H^{\sigma(p)}(\Omega)$, and, for $f \in H^{\sigma(p)}(\Omega)$,

$$\|\mathcal{K}f - K_d f\|_\infty \lesssim h_\mathcal{X}^{p+1/2}\|f\|_{H^{\sigma(p)}(\Omega)}.$$

$$\left\|\mathcal{K} - K_d\right\|_{\mathbb{H}\to C_b} = \left\|\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}\right\|_{\mathbb{H}\to C_b} \lesssim (1 + \|\mathcal{K}\|_{\mathbb{H}\to\mathbb{H}})\|I - P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b}$$

Theorem (Wendland 1995)

There are $C, h_0 > 0$ such that for every set $\mathcal{X} = \{x_i\}_{i=1}^d \subset \Omega$ with $h_\mathcal{X} \leq h_0$ and all $\alpha \in \mathbb{N}_0^d$, $|\alpha| \leq k$,

$$\|I - P_{V_\mathcal{X}}\|_{\mathbb{H}_{\Phi_{n,k}}\to C_b(\Omega)} \leq C h_\mathcal{X}^{k+1/2}.$$

[2] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Proof pt. 2

$$\left\|\mathcal{K} - K_d\right\|_{\mathbb{H}\to C_b} = \left\|\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}\right\|_{\mathbb{H}\to C_b} \lesssim (1 + \|\mathcal{K}\|_{\mathbb{H}\to\mathbb{H}})\|I - P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b}$$

[3] Gonzalez et al. ICLR 2025.
[4] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Proof pt. 2

$$\|\mathcal{K} - K_d\|_{\mathbb{H}\to C_b} = \|\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b} \lesssim (1 + \|\mathcal{K}\|_{\mathbb{H}\to\mathbb{H}})\|I - P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b}$$

Theorem[3]

Let $\mathbb{H}$ be a Gaussian RKHS on $\Omega = \mathbb{R}^{n_x}$. Then $\mathcal{K}\mathbb{H} \subset \mathbb{H}$ if and only if the flow is affine-linear, i.e., $x(t; x^0) = A(t)x^0 + b(t)$.

[3] Gonzalez et al. ICLR 2025.
[4] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Proof pt. 2

$$\|\mathcal{K} - K_d\|_{\mathbb{H}\to C_b} = \|\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b} \lesssim (1 + \|\mathcal{K}\|_{\mathbb{H}\to\mathbb{H}})\|I - P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b}$$

Theorem[3]

Let $\mathbb{H}$ be a Gaussian RKHS on $\Omega = \mathbb{R}^{n_x}$. Then $\mathcal{K}\mathbb{H} \subset \mathbb{H}$ if and only if the flow is affine-linear, i.e., $x(t; x^0) = A(t)x^0 + b(t)$.

Proposition[4]

The Koopman operator on RKHS is always closed.

[3] Gonzalez et al. ICLR 2025.
[4] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Proof pt. 2

$$\left\|\mathcal{K} - K_d\right\|_{\mathbb{H}\to C_b} = \left\|\mathcal{K} - P_{V_\mathcal{X}}\mathcal{K}P_{V_\mathcal{X}}\right\|_{\mathbb{H}\to C_b} \lesssim (1 + \|\mathcal{K}\|_{\mathbb{H}\to\mathbb{H}})\|I - P_{V_\mathcal{X}}\|_{\mathbb{H}\to C_b}$$

Theorem[3]

Let $\mathbb{H}$ be a Gaussian RKHS on $\Omega = \mathbb{R}^{n_x}$. Then $\mathcal{K}\mathbb{H} \subset \mathbb{H}$ if and only if the flow is affine-linear, i.e., $x(t;x^0) = A(t)x^0 + b(t)$.

Proposition[4]

The Koopman operator on RKHS is always closed. In particular, $\mathcal{K}\mathbb{H} \subset \mathbb{H}$ implies $\|\mathcal{K}\| < \infty$.

[3] Gonzalez et al. ICLR 2025.
[4] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Koopman is closed

Adjoint-like property:

$$\langle \mathcal{K}\varphi, k(x,\cdot)\rangle = \varphi(F(x)) = (\mathcal{K}\varphi)(x) = \langle \varphi, k(F(x),\cdot)\rangle$$

[5] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Koopman is closed

Adjoint-like property:

$$\langle \mathcal{K}\varphi, k(x,\cdot)\rangle = \varphi(F(x)) = (\mathcal{K}\varphi)(x) = \langle \varphi, k(F(x),\cdot)\rangle$$

Let $f, f_n \in \mathbb{H}$ and $g \in \mathbb{H}$ such that $\|f_n - f\| \to 0$ and $\|\mathcal{K}f_n - g\| \to 0$ as $n \to \infty$. To show: $g = \mathcal{K}f$.

[5] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Koopman is closed

Adjoint-like property:

$$\langle \mathcal{K}\varphi, k(x,\cdot)\rangle = \varphi(F(x)) = (\mathcal{K}\varphi)(x) = \langle \varphi, k(F(x),\cdot)\rangle$$

Let $f, f_n \in \mathbb{H}$ and $g \in \mathbb{H}$ such that $\|f_n - f\| \to 0$ and $\|\mathcal{K}f_n - g\| \to 0$ as $n \to \infty$. To show: $g = \mathcal{K}f$.

$$g(x) = \langle g, k(x,\cdot)\rangle = \lim_{n\to\infty} \langle \mathcal{K}f_n, k(x,\cdot)\rangle = \lim_{n\to\infty} \langle f_n, k(F(x),\cdot))\rangle = \langle f, k(F(x),\cdot)\rangle = \langle \mathcal{K}f, k(x,\cdot)\rangle$$

[5] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

Koopman is closed

Adjoint-like property:

$$\langle \mathcal{K}\varphi, k(x,\cdot)\rangle = \varphi(F(x)) = (\mathcal{K}\varphi)(x) = \langle \varphi, k(F(x),\cdot)\rangle$$

Let $f, f_n \in \mathbb{H}$ and $g \in \mathbb{H}$ such that $\|f_n - f\| \to 0$ and $\|\mathcal{K}f_n - g\| \to 0$ as $n \to \infty$. To show: $g = \mathcal{K}f$.

$$g(x) = \langle g, k(x,\cdot)\rangle = \lim_{n\to\infty} \langle \mathcal{K}f_n, k(x,\cdot)\rangle = \lim_{n\to\infty} \langle f_n, k(F(x),\cdot))\rangle = \langle f, k(F(x),\cdot)\rangle = \langle \mathcal{K}f, k(x,\cdot)\rangle$$

Corollary[5]

If flow F is in C^m, then for Wendland kernels with $\mathbb{H} = H^{\sigma(p)}(\Omega)$, $\sigma(p) \le m$,

$$\|\mathcal{K}\|_{\mathbb{H}\to\mathbb{H}} < \infty$$

Sketch of the proof: Chain rule, as $\mathcal{K}\varphi = \varphi \circ F$.

[5] Köhne, Philipp, S., Schiela, Worthmann, SIADS 2025

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t))u_i(t)$$

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t)) u_i(t)$$

Aim: Find $u = \mu(x)$ such that above system $\dot{x} = f(x, \mu(x))$ is stable towards x^*.

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t))u_i(t)$$

Aim: Find $u = \mu(x)$ such that above system $\dot{x} = f(x, \mu(x))$ is stable towards x^*.

Assuming $u \in L^\infty$, for fixed $\Delta t > 0$, we may define $x^+ = F(x, u) := \varphi(\Delta t; x, u)$.

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t))u_i(t)$$

Aim: Find $u = \mu(x)$ such that above system $\dot{x} = f(x, \mu(x))$ is stable towards x^*.

Assuming $u \in L^\infty$, for fixed $\Delta t > 0$, we may define $x^+ = F(x, u) := \varphi(\Delta t; x, u)$.

Linear surrogate model (eDMDc) $\psi^+ = \psi(x^+) = \mathcal{K}\psi + \mathcal{B}u$ [Proctor et al. '16] [Korda, Mezić '18]

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t))u_i(t)$$

Aim: Find $u = \mu(x)$ such that above system $\dot{x} = f(x, \mu(x))$ is stable towards x^*.

Assuming $u \in L^\infty$, for fixed $\Delta t > 0$, we may define $x^+ = F(x, u) := \varphi(\Delta t; x, u)$.

Linear surrogate model (eDMDc) $\psi^+ = \psi(x^+) = \mathcal{K}\psi + \mathcal{B}u$ [Proctor et al. '16] [Korda, Mezić '18]
But: Even if $\dot{x} = f(x, u) = Ax + Bu$

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t))u_i(t)$$

Aim: Find $u = \mu(x)$ such that above system $\dot{x} = f(x, \mu(x))$ is stable towards x^*.

Assuming $u \in L^\infty$, for fixed $\Delta t > 0$, we may define $x^+ = F(x,u) := \varphi(\Delta t; x, u)$.

Linear surrogate model (eDMDc) $\psi^+ = \psi(x^+) = \mathcal{K}\psi + \mathcal{B}u$ [Proctor et al. '16] [Korda, Mezić '18]
But: Even if $\dot{x} = f(x,u) = Ax + Bu \rightsquigarrow x(t) = e^{tA}x^0 + \int_0^t e^{(t-s)A}Bu(s)\,\mathrm{d}s$:

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t))u_i(t)$$

Aim: Find $u = \mu(x)$ such that above system $\dot{x} = f(x, \mu(x))$ is stable towards x^*.

Assuming $u \in L^\infty$, for fixed $\Delta t > 0$, we may define $x^+ = F(x,u) := \varphi(\Delta t; x, u)$.

Linear surrogate model (eDMDc) $\psi^+ = \psi(x^+) = \mathcal{K}\psi + \mathcal{B}u$ [Proctor et al. '16] [Korda, Mezić '18]
But: Even if $\dot{x} = f(x,u) = Ax + Bu \rightsquigarrow x(t) = e^{tA}x^0 + \int_0^t e^{(t-s)A}Bu(s)\,\mathrm{d}s$:

$$\psi^+ = \psi(x(\Delta t, u, x^0)) = \psi\left(e^{\Delta t A}x^0 + \int_0^{\Delta t} e^{(\Delta t - s)A}Bu(s)\right).$$

EDMD for control systems

Consider the control-affine system

$$\dot{x}(t) = f(x(t), u(t)) = g_0(x(t)) + \sum_{i=1}^{n_c} g_i(x(t)) u_i(t)$$

Aim: Find $u = \mu(x)$ such that above system $\dot{x} = f(x, \mu(x))$ is stable towards x^*.

Assuming $u \in L^\infty$, for fixed $\Delta t > 0$, we may define $x^+ = F(x, u) := \varphi(\Delta t; x, u)$.

Linear surrogate model (eDMDc) $\psi^+ = \psi(x^+) = \mathcal{K}\psi + \mathcal{B}u$ [Proctor et al. '16] [Korda, Mezić '18]
But: Even if $\dot{x} = f(x, u) = Ax + Bu \rightsquigarrow x(t) = e^{tA}x^0 + \int_0^t e^{(t-s)A} Bu(s)\,\mathrm{d}s$:

$$\psi^+ = \psi(x(\Delta t, u, x^0)) = \psi\left(e^{\Delta t A}x^0 + \int_0^{\Delta t} e^{(\Delta t - s)A} Bu(s)\right).$$

No linearity to be expected for nonlinear liftings.

Bilinear surrogates

Bilinear surrogate model [Williams et al. '16, Surana '16, Peitz et al. '20]

Let $u \in \mathbb{R}^{n_u}$ and consider the Koopman operator

$$(\mathcal{K}_u^t \varphi)(x^0) = \varphi(x(t; x^0, u))$$

Philipp, S., Worthmann, Peitz, Nüske, Journal of Nonlinear Science 2023, 2025

Bilinear surrogates

Bilinear surrogate model [Williams et al. ’16, Surana ’16, Peitz et al. ’20]

Let $u \in \mathbb{R}^{n_u}$ and consider the Koopman operator

$$(\mathcal{K}_u^t \varphi)(x^0) = \varphi(x(t; x^0, u))$$

Strongly continuous semigroup (in L^2 or C):

$$(\mathcal{K}^{t+s} \varphi)(x^0) = \varphi(x(t+s; x^0)) = \varphi(x(t; x(s; x^0))) = (\mathcal{K}^t \varphi)(x(s; x^0)) = (\mathcal{K}^t \mathcal{K}^s \varphi)(x^0)$$

Philipp, S., Worthmann, Peitz, Nüske, Journal of Nonlinear Science 2023, 2025

Bilinear surrogates

Bilinear surrogate model [Williams et al. ’16, Surana ’16, Peitz et al. ’20]

Let $u \in \mathbb{R}^{n_u}$ and consider the Koopman operator

$$(\mathcal{K}_u^t \varphi)(x^0) = \varphi(x(t;x^0,u))$$

Strongly continuous semigroup (in L^2 or C):

$$(\mathcal{K}^{t+s}\varphi)(x^0) = \varphi(x(t+s;x^0)) = \varphi(x(t;x(s;x^0))) = (\mathcal{K}^t\varphi)(x(s;x^0)) = (\mathcal{K}^t\mathcal{K}^s\varphi)(x^0)$$

Then, the densely defined generator

$$\mathcal{L}_u\varphi := \lim_{t\to 0}\frac{\mathcal{K}_u^t\varphi - \varphi}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(x(t;\cdot,u))_{|t=0} = \nabla\varphi\cdot\left(g_0 + \sum g_i u_i\right)$$

Philipp, S., Worthmann, Peitz, Nüske, Journal of Nonlinear Science 2023, 2025

Bilinear surrogates

Bilinear surrogate model [Williams et al. '16, Surana '16, Peitz et al. '20]

Let $u \in \mathbb{R}^{n_u}$ and consider the Koopman operator

$$(\mathcal{K}_u^t \varphi)(x^0) = \varphi(x(t; x^0, u))$$

Strongly continuous semigroup (in L^2 or C):

$$(\mathcal{K}^{t+s}\varphi)(x^0) = \varphi(x(t+s; x^0)) = \varphi(x(t; x(s; x^0))) = (\mathcal{K}^t\varphi)(x(s; x^0)) = (\mathcal{K}^t\mathcal{K}^s\varphi)(x^0)$$

Then, the densely defined generator

$$\mathcal{L}_u\varphi := \lim_{t\to 0} \frac{\mathcal{K}_u^t\varphi - \varphi}{t} = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(x(t; \cdot, u))_{|t=0} = \nabla\varphi \cdot \left(g_0 + \sum g_i u_i\right)$$

hence $\mathcal{L}_u = \mathcal{L}_0 + \sum_{i=1}^{n_c} u_i(\mathcal{L}_{e_i} - \mathcal{L}_0)$, such that

$$\dot{\varphi} = \mathcal{L}_u\varphi = \mathcal{L}_0 + \sum_{i=1}^{n_c} u_i(\mathcal{L}_{e_i} - \mathcal{L}_0)\varphi$$

Philipp, S., Worthmann, Peitz, Nüske, Journal of Nonlinear Science 2023, 2025

EDMD-based exponentially stabilizing controller

Approximately bilinear system

$$\varphi^+ = \mathcal{K}\varphi + u^\top \mathcal{B}\varphi + \mathcal{O}(\Delta t^2)$$

Strässer, S., Worthmann, Berberich, Allgöwer, IEEE TAC 2025
Strässer, Worthmann, Mézic, Berberich, S. Allgöwer, submitted 2025

EDMD-based exponentially stabilizing controller

Approximately bilinear system

$$\varphi^+ = \mathcal{K}\varphi + u^\top \mathcal{B}\varphi + \mathcal{O}(\Delta t^2)$$

$\hookrightarrow$ LMI-based tools from robust control to design state-feedback controller

$$\mu(x) = (I - L_w(\Lambda^{-1} \otimes \hat{\Phi}(x)))^{-1} L P^{-1} \hat{\Phi}(x)$$

ensuring exponential stability (with probability $1-\delta$) for all initial conditions in the safe operating region

$$\hat{x} \in \{x \in \mathbb{R}^n \mid \hat{\Phi}(x)^\top P^{-1} \hat{\Phi}(x) \le 1\},$$

where $P, L, L_w, \Lambda, \ldots$ solve two Linear Matrix Inequalities.

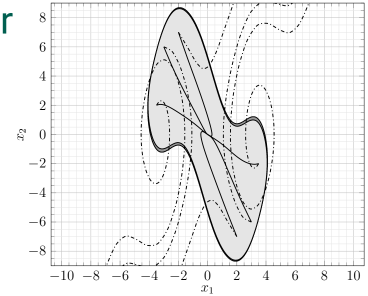
Strässer, S., Worthmann, Berberich, Allgöwer, IEEE TAC 2025
Strässer, Worthmann, Mézic, Berberich, S. Allgöwer, submitted 2025

EDMD-based exponentially stabilizing controller

Nonlinear inverted pendulum

$$\dot{x}_1(t) = x_2(t),$$
$$\dot{x}_2(t) = \frac{g}{l}\sin(x_1(t)) - \frac{b}{ml^2}x_2(t) + \frac{1}{ml^2}u(t)$$

with mass m, length l, rotational friction coefficient b, and gravitational constant g.



Nonlinear system

$$\dot{x}_1(t) = \rho x_1(t),$$
$$\dot{x}_2(t) = \lambda(x_2(t) - x_1(t)^2) + u(t)$$

with $\rho, \lambda \in \mathbb{R}$