

SIMONet: a deep learning framework for learning set-valued maps. Application to Control

Francisco Periago

Polytechnic University of Cartagena. Member of the European University of Technology EUT+

<https://multisimo.com/fpe/> - <https://github.com/fperiago>
joint work with Carlos García-Cervera (UCSB, USA); Mathieu Kessler (UPCT, Spain) and Pablo Pedregal (UCLM, Spain)

Prague, April 27-30, 2026

Motivation and Main Goal



Consider the control system

$$\begin{cases} y'(t) = Ay(t) + Bu(t), & t > 0 \\ y(0) = y^0 \in X, \end{cases} \quad (1)$$

where

- $A : D(A) \subset X \rightarrow X$ is an operator,
- $y = y(t) \in D(A)$ is the state variable,
- $u = u(t) \in \mathcal{U}$ is the control variable, and
- B is an admissible control operator

Consider the control system

$$\begin{cases} y'(t) = Ay(t) + Bu(t), & t > 0 \\ y(0) = y^0 \in X, \end{cases} \quad (1)$$

where

- $A : D(A) \subset X \rightarrow X$ is an operator,
- $y = y(t) \in D(A)$ is the state variable,
- $u = u(t) \in \mathcal{U}$ is the control variable, and
- B is an admissible control operator

Examples of set- valued maps and operators arising in this context

Consider the control system

$$\begin{cases} y'(t) = Ay(t) + Bu(t), & t > 0 \\ y(0) = y^0 \in X, \end{cases} \quad (1)$$

where

- $A : D(A) \subset X \rightarrow X$ is an operator,
- $y = y(t) \in D(A)$ is the state variable,
- $u = u(t) \in \mathcal{U}$ is the control variable, and
- B is an admissible control operator

Examples of set- valued maps and operators arising in this context

- Reachability set at time T for a finite-dimensional control system:

$$\mathbb{R}^n \ni y^0 \Rightarrow \mathbb{R}^n, \quad F_T(y^0) = \mathcal{R}(T, y^0) := \left\{ y(T; y^0, u(\cdot)) \mid u(\cdot) \in \mathcal{U} \right\}$$

Consider the control system

$$\begin{cases} y'(t) = Ay(t) + Bu(t), & t > 0 \\ y(0) = y^0 \in X, \end{cases} \quad (1)$$

where

- $A : D(A) \subset X \rightarrow X$ is an operator,
- $y = y(t) \in D(A)$ is the state variable,
- $u = u(t) \in \mathcal{U}$ is the control variable, and
- B is an admissible control operator

Examples of set- valued maps and operators arising in this context

- Reachability set at time T for a finite-dimensional control system:

$$\mathbb{R}^n \ni y^0 \Rightarrow \mathbb{R}^n, \quad F_T(y^0) = \mathcal{R}(T, y^0) := \left\{ y(T; y^0, u(\cdot)) \mid u(\cdot) \in \mathcal{U} \right\}$$

- Solution map for problems where non-uniqueness of controls holds, e.g.,

$$X \ni y^0 \Rightarrow \mathcal{U}, \quad \mathcal{G}(y^0) = \left\{ u(\cdot; y^0) \in \mathcal{U} \mid y(T) = y^T \right\}$$

Consider the control system

$$\begin{cases} y'(t) = Ay(t) + Bu(t), & t > 0 \\ y(0) = y^0 \in X, \end{cases} \quad (1)$$

where

- $A : D(A) \subset X \rightarrow X$ is an operator,
- $y = y(t) \in D(A)$ is the state variable,
- $u = u(t) \in \mathcal{U}$ is the control variable, and
- B is an admissible control operator

Examples of set- valued maps and operators arising in this context

- Reachability set at time T for a finite-dimensional control system:

$$\mathbb{R}^n \ni y^0 \rightrightarrows \mathbb{R}^n, \quad F_T(y^0) = \mathcal{R}(T, y^0) := \left\{ y(T; y^0, u(\cdot)) \mid u(\cdot) \in \mathcal{U} \right\}$$

- Solution map for problems where non-uniqueness of controls holds, e.g.,

$$X \ni y^0 \rightrightarrows \mathcal{U}, \quad \mathcal{G}(y^0) = \left\{ u(\cdot; y^0) \in \mathcal{U} \mid y(T) = y^T \right\}$$

Main Goal: develop a data-driven method to
learn set-valued functions and operators

What is known ?

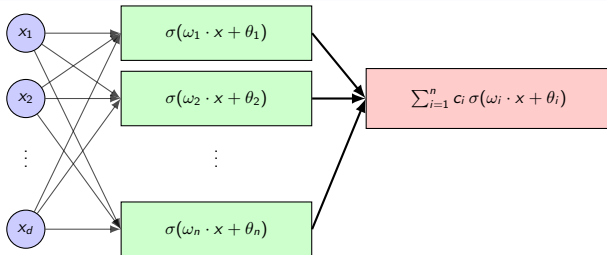
One hidden-layer networks are Universal Approximators of single-valued maps

Theorem (Cybenko (1989), Pinkus (1999), etc.)

Suppose that $K \subset \mathbb{R}^d$ is compact, $U \subset C(K)$ is compact, and $\sigma \in C^\infty(\mathbb{R})$ is not a polynomial. Then, for any $\varepsilon > 0$ there exist a positive integer n , real numbers $\theta_i, \omega_i \in \mathbb{R}^n$, independent of $f \in U$, and constants $c_i = c_i(f)$ depending on f , such that

$$|f(x) - \sum_{i=1}^n c_i \sigma(\omega_i \cdot x + \theta_i)| < \varepsilon$$

holds for all $x \in K$ and $f \in U$. Moreover, each $c_i(f)$ is a continuous linear functional defined on U .



DeepONets are Universal Approximators of single-valued operators

Theorem (Universal Approximation for Functionals)

... for a continuous functional $f : V \rightarrow \mathbb{R}$ we have that

$$\left| f(y) - \sum_{i=1}^n c_i \sigma \left(\sum_{j=1}^m \xi_{ij} y(x_j) + \theta_i \right) \right| < \varepsilon, \quad \text{for all } y \in V.$$

... for an operator, $\mathcal{G}(y)(t) \approx \sum_{k=1}^n c_k (\mathcal{G}(y)) \sigma(\omega_k \cdot t + \theta_k)$, where $c_k (\mathcal{G}(y))$ are continuous functionals.

DeepONets are Universal Approximators of single-valued operators

Theorem (Universal Approximation for Functionals)

... for a continuous functional $f : V \rightarrow \mathbb{R}$ we have that

$$\left| f(y) - \sum_{i=1}^n c_i \sigma \left(\sum_{j=1}^m \xi_{ij} y(x_j) + \theta_i \right) \right| < \varepsilon, \quad \text{for all } y \in V.$$

... for an operator, $\mathcal{G}(y)(t) \approx \sum_{k=1}^n c_k (\mathcal{G}(y)) \sigma(\omega_k \cdot t + \theta_k)$, where $c_k (\mathcal{G}(y))$ are continuous functionals.

Theorem (Universal Approximation for single-valued Operators)

Suppose that σ is not a polynomial, X is a Banach space, $K_1 \subset X$, $K_2 \subset \mathbb{R}^d$ are compact sets, V is a compact set in $C(K_1)$, and $\mathcal{G} : V \rightarrow C(K_2)$ is a continuous operator. Then,

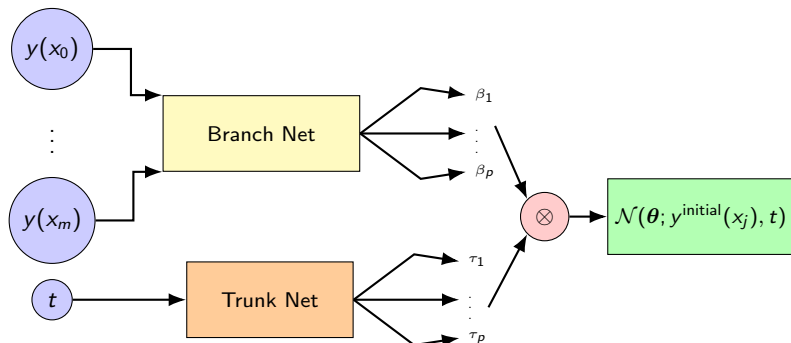
$$\left| \mathcal{G}(y)(t) - \sum_{k=1}^p \sum_{i=1}^n c_i^k \sigma \left(\sum_{j=1}^m \xi_{ij}^k y(x_j) + \theta_i^k \right) \sigma(\omega_k \cdot t + \zeta_k) \right| < \varepsilon, \quad \forall y \in V, t \in K_2$$



Chen, T., Chen, H.: **IEEE Transactions on Neural Networks** 6(4), 911-917, 1995.

DeepONet's architecture

$$\mathcal{N}(\theta; (y^{\text{initial}}(x_j); t)) := \underbrace{\sum_{k=1}^p \sum_{i=1}^n c_i^k \sigma \left(\sum_{j=1}^m \xi_{ij}^k y^{\text{initial}}(x_j) + \theta_i^k \right)}_{\beta_k} \cdot \underbrace{\sigma(w_k \cdot t + \eta_k)}_{\tau_k}$$



L. Lu, P. Jin, G. Pang, Z. Zhang, G. E. Karniadakis: **Nature Machine Intelligence** 3(3) 218–229, 2021.

What is new ?

What is new ?



SIMONets are Universal Approximators of set-valued maps and operators

SIMONets are Universal Approximators of set-valued maps and operators

Notation

- Hausdorff distance

$$d_H(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A) \right\}.$$

SIMONets are Universal Approximators of set-valued maps and operators

Notation

- Hausdorff distance

$$d_H(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A) \right\}.$$

- $\mathcal{F}(Y)$ family of closed sets of a metric space Y

SIMONets are Universal Approximators of set-valued maps and operators

Notation

- Hausdorff distance

$$d_H(A, B) = \max \left\{ \sup_{a \in A} d(a, B), \sup_{b \in B} d(b, A) \right\}.$$

- $\mathcal{F}(Y)$ family of closed sets of a metric space Y

Theorem (Universal Approximation of Set-Valued Operators)

Let $K_1 \subset \mathbb{R}^n$, $K_2 \subset \mathbb{R}^m$ be compact sets, and let $V \subset C(K_1)$ be a compact family. Let

$$F : V \subset C(K_1) \rightrightarrows \mathcal{F}(C(K_2))$$

be a continuous set-valued function such that $F(u)$ is a convex compact set for all $u \in V$. Then, given $\epsilon > 0$, there exists a set-valued neural network

$$\mathcal{G} : V \rightrightarrows \mathcal{F}(C(K_2))$$

such that

$$d_H(F(u), \mathcal{G}(u)) < \epsilon, \quad \forall u \in V.$$



García-Cervera, C.J.; Kessler, M.; Pedregal, P.; Periago, F.: *Universal Approximation of Set-Valued Maps*. **Submitted**.

Theorem (Universal Approximation of Set-Valued Operators)

Let $K_1 \subset \mathbb{R}^n$, $K_2 \subset \mathbb{R}^m$ be compact sets, and let $V \subset C(K_1)$ be a compact family. Let $F : V \subset C(K_1) \rightrightarrows \mathcal{F}(C(K_2))$ be a continuous set-valued function such that $F(u)$ is a convex compact set for all $u \in V$. Then, given $\epsilon > 0$, there exists a set-valued neural network $\mathcal{G} : V \rightrightarrows \mathcal{F}(C(K_2))$ such that

$$d_H(F(u), \mathcal{G}(u)) < \epsilon, \quad \forall u \in V.$$

Theorem (Universal Approximation of Set-Valued Operators)

Let $K_1 \subset \mathbb{R}^n$, $K_2 \subset \mathbb{R}^m$ be compact sets, and let $V \subset C(K_1)$ be a compact family. Let $F : V \subset C(K_1) \rightrightarrows \mathcal{F}(C(K_2))$ be a continuous set-valued function such that $F(u)$ is a convex compact set for all $u \in V$. Then, given $\epsilon > 0$, there exists a set-valued neural network $\mathcal{G} : V \rightrightarrows \mathcal{F}(C(K_2))$ such that

$$d_H(F(u), \mathcal{G}(u)) < \epsilon, \quad \forall u \in V.$$

1) By compactness of V and $F(u)$,

$$V = \bigcup_{j=1}^N B_{\eta_j}(u_j), \quad \text{and} \quad \sup_{v \in F(u_j)} \|v\|_{C(K_2)} \leq M, \quad j = 1, 2, \dots, N.$$

Theorem (Universal Approximation of Set-Valued Operators)

Let $K_1 \subset \mathbb{R}^n$, $K_2 \subset \mathbb{R}^m$ be compact sets, and let $V \subset C(K_1)$ be a compact family. Let $F : V \subset C(K_1) \rightrightarrows \mathcal{F}(C(K_2))$ be a continuous set-valued function such that $F(u)$ is a convex compact set for all $u \in V$. Then, given $\epsilon > 0$, there exists a set-valued neural network $\mathcal{G} : V \rightrightarrows \mathcal{F}(C(K_2))$ such that

$$d_H(F(u), \mathcal{G}(u)) < \epsilon, \quad \forall u \in V.$$

- 1) By compactness of V and $F(u)$,

$$V = \bigcup_{j=1}^N B_{\eta_j}(u_j), \quad \text{and} \quad \sup_{v \in F(u_j)} \|v\|_{C(K_2)} \leq M, \quad j = 1, 2, \dots, N.$$

- 2) Construct a smooth partition of unity, and define

$$\tilde{F}(u) = \sum_{j=1}^N \xi_j(u) F(u_j), \quad \xi_j(u) = \frac{\phi_j(u)}{\sum_{k=1}^N \phi_k(u)}, \quad \phi_j(u) = \varphi\left(\frac{\|u - u_j\|}{\eta_j}\right).$$

Theorem (Universal Approximation of Set-Valued Operators)

Let $K_1 \subset \mathbb{R}^n$, $K_2 \subset \mathbb{R}^m$ be compact sets, and let $V \subset C(K_1)$ be a compact family. Let $F : V \subset C(K_1) \rightrightarrows \mathcal{F}(C(K_2))$ be a continuous set-valued function such that $F(u)$ is a convex compact set for all $u \in V$. Then, given $\epsilon > 0$, there exists a set-valued neural network $\mathcal{G} : V \rightrightarrows \mathcal{F}(C(K_2))$ such that

$$d_H(F(u), \mathcal{G}(u)) < \epsilon, \quad \forall u \in V.$$

- 1) By compactness of V and $F(u)$,

$$V = \bigcup_{j=1}^N B_{\eta_j}(u_j), \quad \text{and} \quad \sup_{v \in F(u_j)} \|v\|_{C(K_2)} \leq M, \quad j = 1, 2, \dots, N.$$

- 2) Construct a smooth partition of unity, and define

$$\tilde{F}(u) = \sum_{j=1}^N \xi_j(u) F(u_j), \quad \xi_j(u) = \frac{\phi_j(u)}{\sum_{k=1}^N \phi_k(u)}, \quad \phi_j(u) = \varphi\left(\frac{\|u - u_j\|}{\eta_j}\right).$$

- 3) Using the convexity and continuity of F , it is proved that

$$d_H(F(u), \tilde{F}(u)) \lesssim \epsilon, \quad \forall u \in V.$$

- 4) Since $\xi_j : V \rightarrow \mathbb{R}$ are continuous nonlinear functionals, there exist (Chen&Chen)

$$g_j(u) = \sum_{k=1}^{L_j} c_{j,k} \sigma \left(\sum_{r=1}^{m_j} \alpha_{j,kr} u(x_r) + \theta_{j,k} \right),$$

such that $\|\xi_j - g_j\|_{C(K_1)} \lesssim \epsilon$

- 4) Since $\xi_j : V \rightarrow \mathbb{R}$ are continuous nonlinear functionals, there exist (Chen&Chen)

$$g_j(u) = \sum_{k=1}^{L_j} c_{j,k} \sigma \left(\sum_{r=1}^{m_j} \alpha_{j,kr} u(x_r) + \theta_{j,k} \right),$$

such that $\|\xi_j - g_j\|_{C(K_1)} \lesssim \epsilon$

- 5) Define the set-valued neural network

$$G(u) = \sum_{j=1}^N g_j(u) F(u_j),$$

which satisfies $d_H(\tilde{F}(u), G(u)) \lesssim \epsilon$.

- 4) Since $\xi_j : V \rightarrow \mathbb{R}$ are continuous nonlinear functionals, there exist (Chen&Chen)

$$g_j(u) = \sum_{k=1}^{L_j} c_{j,k} \sigma \left(\sum_{r=1}^{m_j} \alpha_{j,kr} u(x_r) + \theta_{j,k} \right),$$

such that $\|\xi_j - g_j\|_{C(K_1)} \lesssim \epsilon$

- 5) Define the set-valued neural network

$$G(u) = \sum_{j=1}^N g_j(u) F(u_j),$$

which satisfies $d_H(\tilde{F}(u), G(u)) \lesssim \epsilon$.

- 6) By the Krein-Milman theorem,

$$F(u_j) = \overline{\text{co}}(\Gamma_j) = \overline{\text{co}}(\overline{\Gamma_j}),$$

Γ_j is the family of extremal points of $F(u_j)$

- 4) Since $\xi_j : V \rightarrow \mathbb{R}$ are continuous nonlinear functionals, there exist (Chen&Chen)

$$g_j(u) = \sum_{k=1}^{L_j} c_{j,k} \sigma \left(\sum_{r=1}^{m_j} \alpha_{j,kr} u(x_r) + \theta_{j,k} \right),$$

such that $\|\xi_j - g_j\|_{C(K_1)} \lesssim \epsilon$

- 5) Define the set-valued neural network

$$G(u) = \sum_{j=1}^N g_j(u) F(u_j),$$

which satisfies $d_H(\tilde{F}(u), G(u)) \lesssim \epsilon$.

- 6) By the Krein-Milman theorem,

$$F(u_j) = \overline{\text{co}}(\Gamma_j) = \overline{\text{co}}(\overline{\Gamma_j}),$$

Γ_j is the family of extremal points of $F(u_j)$

- 7) Since $\overline{\Gamma_j}$ is compact, we can find $d_j \in \mathbb{N}$ and a finite set

$$A_j = \{z_{j,\ell}\}_{\ell=1}^{d_j} \subset \Gamma_j \text{ such that}$$

$$d_H(F(u_j), \overline{\text{co}}(A_j)) = d_H(\overline{\text{co}}(\Gamma_j), \overline{\text{co}}(A_j)) \lesssim \epsilon.$$

- 8) We approximate each $z_{j,\ell}$ by a neural network (Cybenko-Pinkus)

$$h_{j,\ell}(\mathbf{x}) = \sum_{i=1}^{K_{j,\ell}} c_{i,j,\ell} \sigma(\omega_{i,j,\ell} \cdot \mathbf{x} + \theta_{i,j,\ell}), \quad \mathbf{x} \in K_2.$$

- 8) We approximate each $z_{j,\ell}$ by a neural network (Cybenko-Pinkus)

$$h_{j,\ell}(\mathbf{x}) = \sum_{i=1}^{K_{j,\ell}} c_{i,j,\ell} \sigma(\omega_{i,j,\ell} \cdot \mathbf{x} + \theta_{i,j,\ell}), \quad \mathbf{x} \in K_2.$$

- 9) Finally, we define what we call a **Single Input Multiple Output Net (SIMONet)** as the set-valued function

$$\mathcal{G} : V \rightrightarrows \mathcal{F}(C(K_2))$$

given by

$$\mathcal{G}(u) = \sum_{j=1}^N g_j(u) \operatorname{co} \left(\{h_{j,\ell_j}\}_{\ell_j=1}^{d_j} \right), \quad u \in V.$$

which approximates the original set-valued map F .



An example of application in Control Theory



An example of application in Control Theory



Reachability set of a linear infinite-dimensional control system

Reachability set of a linear infinite-dimensional control system

Model and physical interpretation

$$\begin{cases} \partial_t y(x, t) = \partial_{xx} y(x, t) + b(x)u(t), & x \in (0, 1), \quad t \in (0, T), \\ y(0, t) = y(1, t) = 0, & t \in (0, T), \\ y(x, 0) = 0, & x \in (0, 1), \end{cases} \quad (2)$$

where $b \in H := L^2(0, 1)$ is a spatial profile and the control satisfies

$$\|u\|_{L^2(0, T)} \leq 1.$$

Reachability set of a linear infinite-dimensional control system

Model and physical interpretation

$$\begin{cases} \partial_t y(x, t) = \partial_{xx} y(x, t) + b(x)u(t), & x \in (0, 1), \quad t \in (0, T), \\ y(0, t) = y(1, t) = 0, & t \in (0, T), \\ y(x, 0) = 0, & x \in (0, 1), \end{cases} \quad (2)$$

where $b \in H := L^2(0, 1)$ is a spatial profile and the control satisfies

$$\|u\|_{L^2(0, T)} \leq 1.$$

Goal

To develop a data-driven (one-shot) method to learn the set-valued operator

$$\mathcal{F} : b \in H \Rightarrow \mathcal{R}_T(b) = \{y(T; u) : \|u\|_{L^2(0, T)} \leq 1\} \subset H$$

$\mathcal{F}$ satisfies all hypotheses of our set-valued UAT



Expanding in the Laplacian eigenpairs $e_k^{(L)}(x) = \sqrt{2} \sin(k\pi x)$, $\lambda_k^{(L)} = (k\pi)^2$

$$y(T; b) = \int_0^T e^{A(T-s)} b u(s) ds, \quad e^{At} f = \sum_{k=1}^{\infty} e^{-\lambda_k t} f_k e_k, \quad t > 0.$$

Expanding in the Laplacian eigenpairs $e_k^{(L)}(x) = \sqrt{2} \sin(k\pi x)$, $\lambda_k^{(L)} = (k\pi)^2$

$$y(T; b) = \int_0^T e^{A(T-s)} b u(s) ds, \quad e^{At} f = \sum_{k=1}^{\infty} e^{-\lambda_k t} f_k e_k, \quad t > 0.$$

Define the linear and bounded operator

$$K_b : L^2(0, T) \rightarrow H, \quad K_b u := \int_0^T e^{A(T-s)} b u(s) ds.$$

The reachable set at time T is

$$\mathcal{R}_T(b) = \{K_b u : \|u\|_{L^2(0, T)} \leq 1\}.$$

Expanding in the Laplacian eigenpairs $e_k^{(L)}(x) = \sqrt{2} \sin(k\pi x)$, $\lambda_k^{(L)} = (k\pi)^2$

$$y(T; b) = \int_0^T e^{A(T-s)} b u(s) ds, \quad e^{At} f = \sum_{k=1}^{\infty} e^{-\lambda_k t} f_k e_k, \quad t > 0.$$

Define the linear and bounded operator

$$K_b : L^2(0, T) \rightarrow H, \quad K_b u := \int_0^T e^{A(T-s)} b u(s) ds.$$

The reachable set at time T is

$$\mathcal{R}_T(b) = \{K_b u : \|u\|_{L^2(0, T)} \leq 1\}.$$

K_b is a Hilbert-Schmidt operator. As a consequence, $\overline{\mathcal{R}_T(b)}$ is an ellipsoid¹

¹A subset $E \subset H$ is called an *ellipsoid* if there exist a point $a \in H$ (the center), an orthonormal family $\{f_n\}_{n \geq 1} \subset H$, and a sequence of nonnegative real numbers $\{\sigma_n\}_{n \geq 1}$ (the semiaxes) such that

$$E = \left\{ a + \sum_{n=1}^{\infty} \sigma_n \alpha_n f_n : \sum_{n=1}^{\infty} \alpha_n^2 \leq 1 \right\}.$$

Dataset

Dataset

We sample b from the (truncated) Karhunen–Loève expansion expansion associated to the squared-exponential covariance kernel

$$k(x, x') = \sigma^2 \exp\left(-\frac{(x - x')^2}{2\ell^2}\right), \quad x, x' \in (0, 1).$$

That is, if $\{\lambda_j^{(KL)}, e_j^{(KL)}\}_{j \geq 1}$ denote the eigenpairs of the associated integral operator

$$(Kf)(x) = \int_0^1 k(x, x') f(x') dx',$$

then

$$b^{(m, KL)}(x) = \sum_{j=1}^m \sqrt{\lambda_j^{(KL)}} \xi_j e_j^{(KL)}(x)$$

Dataset

We sample b from the (truncated) Karhunen–Loève expansion expansion associated to the squared-exponential covariance kernel

$$k(x, x') = \sigma^2 \exp\left(-\frac{(x - x')^2}{2\ell^2}\right), \quad x, x' \in (0, 1).$$

That is, if $\{\lambda_j^{(KL)}, e_j^{(KL)}\}_{j \geq 1}$ denote the eigenpairs of the associated integral operator

$$(Kf)(x) = \int_0^1 k(x, x') f(x') dx',$$

then

$$b^{(m, KL)}(x) = \sum_{j=1}^m \sqrt{\lambda_j^{(KL)}} \xi_j e_j^{(KL)}(x)$$

Projecting onto the Laplacian basis,

$$\mathbf{b}^{(L)} = M \operatorname{diag}(\sqrt{\lambda_1^{(KL)}}, \dots, \sqrt{\lambda_m^{(KL)}}) \boldsymbol{\xi} \in \mathbb{R}^N, \quad M_{jk} := \langle e_j^{(KL)}, e_k^{(L)} \rangle_H, \quad \boldsymbol{\xi} = (\xi_1, \dots, \xi_m)^\top$$

This is a **feature vector** for learning.

Dataset

Dataset

Discrete control operator in spectral coordinates: $K_{\mathbf{b}} : \mathbb{R}^{N_t} \rightarrow \mathbb{R}^N$ is

$$K_{\mathbf{b}} = \begin{bmatrix} b_1^{(L)} e^{-\lambda_1^{(L)}(T-t_1)} & \dots & b_1^{(L)} e^{-\lambda_1^{(L)}(T-t_{N_t})} \\ \vdots & \ddots & \vdots \\ b_N^{(L)} e^{-\lambda_N^{(L)}(T-t_1)} & \dots & b_N^{(L)} e^{-\lambda_N^{(L)}(T-t_{N_t})} \end{bmatrix} \Delta t.$$

For a discretized control $\mathbf{u} = (u(t_1), \dots, u(t_{N_t}))^\top$, $\mathbf{y}_T \approx K_{\mathbf{b}} \mathbf{u}$.

Dataset

Discrete control operator in spectral coordinates: $K_{\mathbf{b}} : \mathbb{R}^{N_t} \rightarrow \mathbb{R}^N$ is

$$K_{\mathbf{b}} = \begin{bmatrix} b_1^{(L)} e^{-\lambda_1^{(L)}(T-t_1)} & \dots & b_1^{(L)} e^{-\lambda_1^{(L)}(T-t_{N_t})} \\ \vdots & \ddots & \vdots \\ b_N^{(L)} e^{-\lambda_N^{(L)}(T-t_1)} & \dots & b_N^{(L)} e^{-\lambda_N^{(L)}(T-t_{N_t})} \end{bmatrix} \Delta t.$$

For a discretized control $\mathbf{u} = (u(t_1), \dots, u(t_{N_t}))^\top$, $\mathbf{y}_T \approx K_{\mathbf{b}} \mathbf{u}$.

Truncated SVD and parametric ellipsoid: $K_{\mathbf{b}} \approx U_r(\mathbf{b}) \Sigma_r(\mathbf{b}) V_r(\mathbf{b})^\top$

$$U_r(\mathbf{b}) = [\psi_1(\mathbf{b}), \dots, \psi_r(\mathbf{b})] \in \mathbb{R}^{N \times r},$$

$$\Sigma_r(\mathbf{b}) = \text{diag}(\sigma_1(\mathbf{b}), \dots, \sigma_r(\mathbf{b})),$$

$$V_r(\mathbf{b})^\top = [\phi_1(\mathbf{b}), \dots, \phi_r(\mathbf{b})] \in \mathbb{R}^{r \times N_t}.$$

Dataset

Discrete control operator in spectral coordinates: $K_{\mathbf{b}} : \mathbb{R}^{N_t} \rightarrow \mathbb{R}^N$ is

$$K_{\mathbf{b}} = \begin{bmatrix} b_1^{(L)} e^{-\lambda_1^{(L)}(T-t_1)} & \dots & b_1^{(L)} e^{-\lambda_1^{(L)}(T-t_{N_t})} \\ \vdots & \ddots & \vdots \\ b_N^{(L)} e^{-\lambda_N^{(L)}(T-t_1)} & \dots & b_N^{(L)} e^{-\lambda_N^{(L)}(T-t_{N_t})} \end{bmatrix} \Delta t.$$

For a discretized control $\mathbf{u} = (u(t_1), \dots, u(t_{N_t}))^\top$, $\mathbf{y}_T \approx K_{\mathbf{b}} \mathbf{u}$.

Truncated SVD and parametric ellipsoid: $K_{\mathbf{b}} \approx U_r(\mathbf{b}) \Sigma_r(\mathbf{b}) V_r(\mathbf{b})^\top$

$$U_r(\mathbf{b}) = [\psi_1(\mathbf{b}), \dots, \psi_r(\mathbf{b})] \in \mathbb{R}^{N \times r},$$

$$\Sigma_r(\mathbf{b}) = \text{diag}(\sigma_1(\mathbf{b}), \dots, \sigma_r(\mathbf{b})),$$

$$V_r(\mathbf{b})^\top = [\phi_1(\mathbf{b}), \dots, \phi_r(\mathbf{b})] \in \mathbb{R}^{r \times N_t}.$$

Thus, we get the r -dimensional ellipsoid

$$\mathcal{E}_r(\mathbf{b}) := \left\{ \mathbf{y} = \sum_{n=1}^r \sigma_n(\mathbf{b}) \alpha_n \psi_n(\mathbf{b}) : \sum_{n=1}^r \alpha_n^2 \leq 1 \right\} \subset \mathbb{R}^N.$$

$$\boldsymbol{\sigma}(\mathbf{b}) = (\sigma_1(\mathbf{b}), \dots, \sigma_r(\mathbf{b})) \in \mathbb{R}^r, \quad \boldsymbol{\Psi}(\mathbf{b}) = (\psi_1(\mathbf{b}), \dots, \psi_r(\mathbf{b})) \in \mathbb{R}^{N \times r}.$$

These serve as the **labels** for the learning process.

$$\mathcal{F}_{N,r} : \mathbb{R}^N \longrightarrow \mathbb{R}^r \times (\mathbb{R}^N)^r, \quad \mathbf{b}^{(L)} \longmapsto (\boldsymbol{\sigma}(\mathbf{b}), \boldsymbol{\Psi}(\mathbf{b})),$$

where

$$\boldsymbol{\sigma}(\mathbf{b}) = (\sigma_1(\mathbf{b}), \dots, \sigma_r(\mathbf{b})) \in \mathbb{R}^r, \quad \boldsymbol{\Psi}(\mathbf{b}) = (\psi_1(\mathbf{b}), \dots, \psi_r(\mathbf{b})) \in \mathbb{R}^{N \times r}.$$

$$\mathcal{F}_{N,r} : \mathbb{R}^N \longrightarrow \mathbb{R}^r \times (\mathbb{R}^N)^r, \quad \mathbf{b}^{(L)} \longmapsto (\boldsymbol{\sigma}(\mathbf{b}), \boldsymbol{\Psi}(\mathbf{b})),$$

where

$$\boldsymbol{\sigma}(\mathbf{b}) = (\sigma_1(\mathbf{b}), \dots, \sigma_r(\mathbf{b})) \in \mathbb{R}^r, \quad \boldsymbol{\Psi}(\mathbf{b}) = (\psi_1(\mathbf{b}), \dots, \psi_r(\mathbf{b})) \in \mathbb{R}^{N \times r}.$$

Configuration of the numerical experiment

Parameters: $\ell = 0.3$, $\sigma = 1$ and $m = 3$ for KL. $N = 5$ for truncation of the Dirichlet Laplacian spectral basis, $T = 0.25$ and $\Delta t = 0.01$. The SVD is truncated at $r = 3$.

$$\mathcal{F}_{N,r} : \mathbb{R}^N \longrightarrow \mathbb{R}^r \times (\mathbb{R}^N)^r, \quad \mathbf{b}^{(L)} \longmapsto (\boldsymbol{\sigma}(\mathbf{b}), \boldsymbol{\Psi}(\mathbf{b})),$$

where

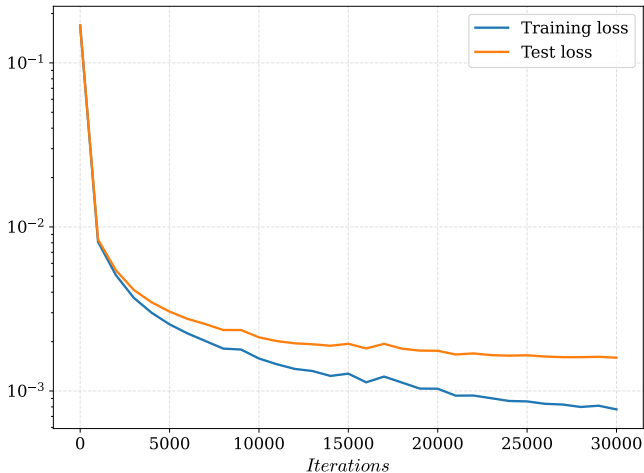
$$\boldsymbol{\sigma}(\mathbf{b}) = (\sigma_1(\mathbf{b}), \dots, \sigma_r(\mathbf{b})) \in \mathbb{R}^r, \quad \boldsymbol{\Psi}(\mathbf{b}) = (\psi_1(\mathbf{b}), \dots, \psi_r(\mathbf{b})) \in \mathbb{R}^{N \times r}.$$

Configuration of the numerical experiment

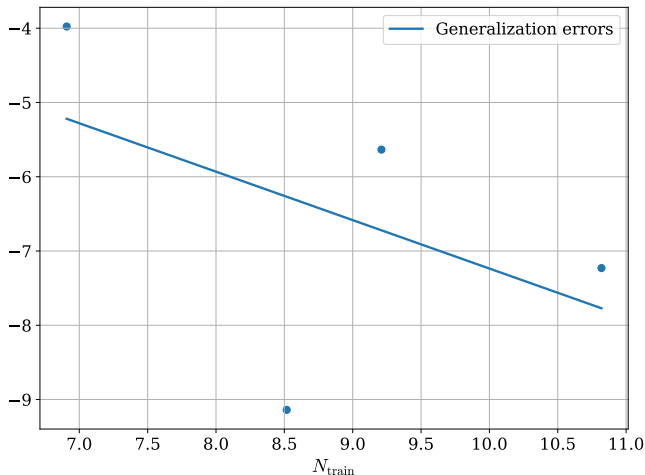
Parameters: $\ell = 0.3$, $\sigma = 1$ and $m = 3$ for KL. $N = 5$ for truncation of the Dirichlet Laplacian spectral basis, $T = 0.25$ and $\Delta t = 0.01$. The SVD is truncated at $r = 3$.

A SIMONet composed of three hidden layers with 100 neurons each. ReLU is used as activation function, initialization is carried out with Glorot normal, and ADAM is the chosen optimizer.

Training performance

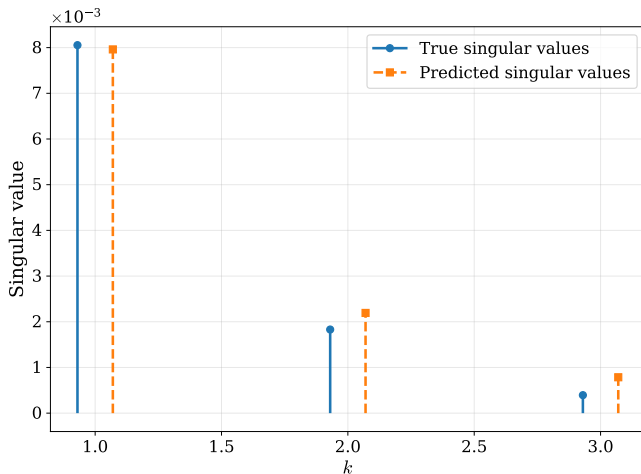


- 5×10^4 samples for training (80%) and test (20%)
- generalization gap = $\mathcal{O}(10^{-4})$



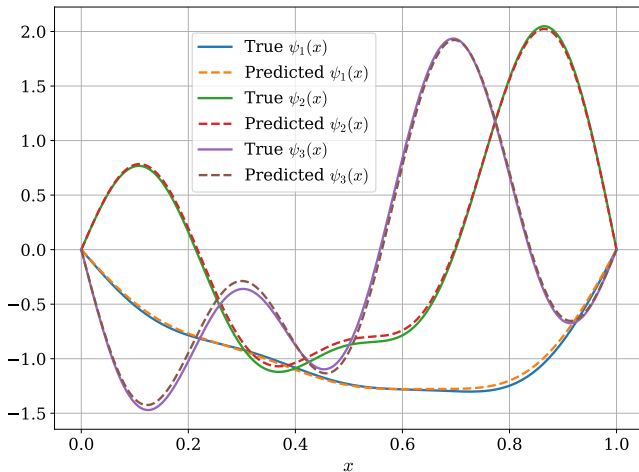
- Generalization gap and linear regression curves in log-log scale after training datasets of size $N_{\text{train}} = 10^3$, 5×10^3 , 10^4 and 5×10^4
- Convergence order = $\mathcal{O}(N_{\text{train}}^{-0.65})$

Qualitative comparison of semi-axis



$$\frac{\|\sigma_{\text{True}} - \sigma_{\text{Predicted}}\|_{\infty}}{\|\sigma_{\text{True}}\|_{\infty}} = 4.4 \times 10^{-2}.$$

Qualitative comparison of principal directions



$$\max_{1 \leq j \leq 3} \frac{\|\psi_j^{\text{True}} - \psi_j^{\text{Predicted}}\|_2}{\|\psi_j^{\text{True}}\|_2} = 2.7 \times 10^{-2}.$$

Contributions

- **Theoretical:** We have established universal approximation theorems for mappings with values in compact and convex sets

Contributions

- **Theoretical:** We have established universal approximation theorems for mappings with values in compact and convex sets
- **Application:** We have applied it to design one-shot, data-driven, approach to compute reachability sets of linear control problems

Contributions

- **Theoretical:** We have established universal approximation theorems for mappings with values in compact and convex sets
- **Application:** We have applied it to design one-shot, data-driven, approach to compute reachability sets of linear control problems

Challenges

- **Theoretical:** Extend universal approximation theorems for mappings with **non-convex** values

Contributions

- **Theoretical:** We have established universal approximation theorems for mappings with values in compact and convex sets
- **Application:** We have applied it to design one-shot, data-driven, approach to compute reachability sets of linear control problems

Challenges

- **Theoretical:** Extend universal approximation theorems for mappings with **non-convex** values
- **Application:** Design one-shot, data-driven, approaches to compute reachability sets of **non-linear** control problems

Work supported by

- Grant PID2022-141957OA-C22 funded by MCIN/AEI/10.13039/501100011033 and by ERDF A way of making Europe



- Autonomous Community of the Región of Murcia, Spain, through the programme for the development of scientific and technical research by competitive groups (21996/PI/22), included in the Regional Program for the Promotion of Scientific and Technical Research of Fundación Séneca – Agencia de Ciencia y Tecnología de la Región de Murcia..



f SéNeCa⁽⁺⁾

Agencia de Ciencia y Tecnología
Región de Murcia

Thank you for your attention !