ANOVA meets control

Melanie Kircheis

joint work with
Daniel Potts, Manuel Schaller and Karl Worthmann

Research Seminar Numerics

Motivation

Unknown function (control system):

$$H : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F : \mathbb{R}^d \to \mathbb{R}^d, G : \mathbb{R}^d \to \mathbb{R}^{d \times m}$,
- $d, m \in \mathbb{N}, m \ll d$.

Motivation

Unknown function (control system):

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d, G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m},$
- $d, m \in \mathbb{N}, m \ll d.$

Given: samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$ for $i = 1, \dots, M$

Motivation

Unknown function (control system):

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d, G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d.$

Given: samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$ for $i = 1, \ldots, M$

Goal: find $\tilde{H} \approx H$

Motivation

Unknown function (control system):

$$H \colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F \colon \mathbb{R}^d \to \mathbb{R}^d, G \colon \mathbb{R}^d \to \mathbb{R}^{d \times m}$,
- $d, m \in \mathbb{N}, m \ll d$.

Given: samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$ for $i = 1, \dots, M$

Goal: find $\tilde{H} \approx H$

Problem: curse of dimensionality

Motivation

Unknown function (control system):

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d, G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d.$

Given: samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$ for $i = 1, \dots, M$

Goal: find $\tilde{H} \approx H$

Problem: curse of dimensionality

Assumption: F and G of low-dimensional structure

Motivation

Unknown function (control system):

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d, G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d.$

Given: samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$ for $i = 1, \ldots, M$

Goal: find $\tilde{H} \approx H$

Problem: curse of dimensionality

Assumption: F and G of low-dimensional structure

Example: Euler approximation of a controlled Duffing oscillator

$$\begin{aligned} H(\boldsymbol{x}, \boldsymbol{u}) &= \boldsymbol{x} + \Delta t \begin{pmatrix} x_2 \\ x_1 - 3x_1^3 u \end{pmatrix} \\ &= \begin{pmatrix} x_1 + \Delta t x_2 \\ x_2 + \Delta t x_1 \end{pmatrix} + \begin{pmatrix} 0 \\ -3\Delta t x_1^3 \end{pmatrix} u \\ &= F(\boldsymbol{x}) + G(\boldsymbol{x}) u \end{aligned}$$

i. e.,

- $d = 2,$
- $m = 1,$
- only 1D terms $\rightsquigarrow$ low-dimensional

Overview

1. ANOVA approximation for scalar-valued functions
2. ANOVA approximation for control systems
3. Numerical Example

ANOVA approximation

ANalysis Of VAriance (ANOVA) decomposition

[Caflish, Morokoff, Owen 97], [Rabitz, Alis 99], [Liu, Owen 06], [Kuo, Sloan, Wasilkowski, Wozniakowski 10], ...

Decompose a d-dimensional function f into

$$\begin{aligned} f(\boldsymbol{x}) = f_\varnothing & & \text{constant} \\ &+ f_1(x_1) + \ldots + f_d(x_d) & \text{one-dimensional terms} \\ &+ f_{1,2}(x_1, x_2) + f_{1,3}(x_1, x_3) + \ldots + f_{d-1,d}(x_{d-1}, x_d) & \text{two-dimensional terms} \\ &+ f_{1,2,3}(x_1, x_2, x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2}, x_{d-1}, x_d) & \text{three-dimensional terms} \\ &\quad \vdots & \\ &+ f_{[d]}(\boldsymbol{x}) & d\text{-dimensional term} \end{aligned}$$

$[d] := \{1, \ldots, d\}$

ANalysis Of VAriance (ANOVA) decomposition

[Caflish, Morokoff, Owen 97], [Rabitz, Alis 99], [Liu, Owen 06], [Kuo, Sloan, Wasilkowski, Wozniakowski 10], . . .

Decompose a d-dimensional function f into

$$
\begin{aligned}
f(\boldsymbol{x}) = f_\varnothing & & \text{constant}\\
&+ f_1(x_1) + \ldots + f_d(x_d) & \text{one-dimensional terms}\\
&+ f_{1,2}(x_1, x_2) + f_{1,3}(x_1, x_3) + \ldots + f_{d-1,d}(x_{d-1}, x_d) & \text{two-dimensional terms}\\
&+ f_{1,2,3}(x_1, x_2, x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2}, x_{d-1}, x_d) & \text{three-dimensional terms}\\
&\ \ \vdots & \\
&+ f_{[d]}(\boldsymbol{x}) & d\text{-dimensional term}\\
&= \sum_{\boldsymbol{v} \subseteq [d]} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}})
\end{aligned}
$$

$[d] := \{1, \ldots, d\}$

ANalysis Of VAriance (ANOVA) decomposition

[Caflish, Morokoff, Owen 97], [Rabitz, Alis 99], [Liu, Owen 06], [Kuo, Sloan, Wasilkowski, Wozniakowski 10], ...

Decompose a d-dimensional function f into

$$\begin{aligned} f(\boldsymbol{x}) = f_\varnothing & & \text{constant} \\ + f_1(x_1) + \ldots + f_d(x_d) & & \text{one-dimensional terms} \\ + f_{1,2}(x_1,x_2) + f_{1,3}(x_1,x_3) + \ldots + f_{d-1,d}(x_{d-1},x_d) & & \text{two-dimensional terms} \\ + f_{1,2,3}(x_1,x_2,x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2},x_{d-1},x_d) & & \text{three-dimensional terms} \\ \vdots \\ + f_{[d]}(\boldsymbol{x}) & & d\text{-dimensional term} \\ = \sum_{\boldsymbol{v}\subseteq[d]} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) \end{aligned}$$

$[d] := \{1,\ldots,d\}$

→ in general multiple representations possible

→ conditions for uniqueness?

Definition

Let

$$\langle f, g \rangle_{L_2(\Omega,\omega)} := \int_{\Omega} f(\boldsymbol{x})\, g(\boldsymbol{x})\, \omega(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

Definition

Let

$$\langle f, g \rangle_{L_2(\Omega,\omega)} := \int_\Omega f(\boldsymbol{x})\, g(\boldsymbol{x})\, \omega(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

$\Rightarrow$ unique decomposition satisfying:

$$\boxed{0 = \langle f_{\boldsymbol{v}}, f_{\boldsymbol{z}} \rangle_{L_2(\Omega,\omega)} \text{ for all } \boldsymbol{v} \neq \boldsymbol{z} \subseteq [d]}$$

Definition

Let

$$\langle f, g\rangle_{L_2(\Omega,\omega)} := \int_\Omega f(\boldsymbol{x})\, g(\boldsymbol{x})\, \omega(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

$\Rightarrow$ unique decomposition satisfying:

$$\boxed{0 = \langle f_{\boldsymbol{v}}, f_{\boldsymbol{z}}\rangle_{L_2(\Omega,\omega)} \text{ for all } \boldsymbol{v} \neq \boldsymbol{z} \subseteq [d]}$$

Basis representation: For orthonormal basis $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{N}_0^d}$ of $L_2(\Omega,\omega)$ we have

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x}), \qquad c_{\boldsymbol{k}}(f) := \langle f, \varphi_{\boldsymbol{k}}\rangle_{L_2(\Omega,\omega)}.$$

Definition

Let

$$\langle f, g\rangle_{L_2(\Omega,\omega)} := \int_\Omega f(\boldsymbol{x})\, g(\boldsymbol{x})\, \omega(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

$\Rightarrow$ unique decomposition satisfying:

$$\boxed{0 = \langle f_{\boldsymbol{v}}, f_{\boldsymbol{z}}\rangle_{L_2(\Omega,\omega)} \text{ for all } \boldsymbol{v} \neq \boldsymbol{z} \subseteq [d]}$$

Basis representation: For orthonormal basis $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{N}_0^d}$ of $L_2(\Omega,\omega)$ we have

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x}), \qquad c_{\boldsymbol{k}}(f) := \langle f, \varphi_{\boldsymbol{k}}\rangle_{L_2(\Omega,\omega)}.$$

[Potts, Schmischke 21]:

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d \\ \operatorname{supp} \boldsymbol{k}=\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Definition

Let

$$\langle f, g \rangle_{L_2(\Omega,\omega)} := \int_\Omega f(\boldsymbol{x})\, g(\boldsymbol{x})\, \omega(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x}$$

$\Rightarrow$ unique decomposition satisfying:

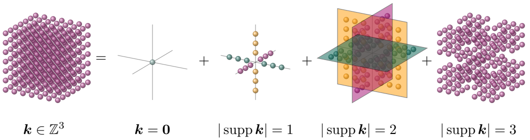
$$\boxed{0 = \langle f_{\boldsymbol{v}}, f_{\boldsymbol{z}} \rangle_{L_2(\Omega,\omega)} \text{ for all } \boldsymbol{v} \neq \boldsymbol{z} \subseteq [d]}$$

Basis representation: For orthonormal basis $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{N}_0^d}$ of $L_2(\Omega,\omega)$ we have

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{k}\in\mathbb{N}_0^d} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x}), \qquad c_{\boldsymbol{k}}(f) := \langle f, \varphi_{\boldsymbol{k}} \rangle_{L_2(\Omega,\omega)}.$$

[Potts, Schmischke 21]:

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{N}_0^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$



[image credits: Laura Weidensager]

$\Rightarrow$ decomposition in the space of basis coefficients

Numerical realization – ANOVA approximation

[Potts, Schmischke 21]

$$f = \sum_{\boldsymbol{v} \subseteq [d]} f_{\boldsymbol{v}}, \qquad f_{\boldsymbol{v}} = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Numerical realization – ANOVA approximation

[Potts, Schmischke 21]

$$f = \sum_{\boldsymbol{v} \subseteq [d]} f_{\boldsymbol{v}}, \qquad f_{\boldsymbol{v}} = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Truncation

reduce number of ANOVA terms

$$\mathcal{T}_V f = \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} f_{\boldsymbol{v}}$$

$$f_{\boldsymbol{v}} = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Numerical realization – ANOVA approximation

[Potts, Schmischke 21]

$$f = \sum_{\boldsymbol{v} \subseteq [d]} f_{\boldsymbol{v}}, \qquad f_{\boldsymbol{v}} = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Truncation

reduce number of ANOVA terms

$$\mathcal{T}_V f = \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} f_{\boldsymbol{v}}$$

$$f_{\boldsymbol{v}} = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Projection

choose finite number of basis functions $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k} \in \mathbb{Z}^d}$

$$P_{\boldsymbol{N}} f = \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} \tilde{f}_{\boldsymbol{v}}$$

$$\tilde{f}_{\boldsymbol{v}} = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{N}}^{\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Numerical realization – ANOVA approximation

[Potts, Schmischke 21]

$$f = \sum_{\boldsymbol{v} \subseteq [d]} f_{\boldsymbol{v}}, \qquad f_{\boldsymbol{v}} = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Truncation

reduce number of ANOVA terms

$$\mathcal{T}_V f = \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} f_{\boldsymbol{v}}$$

$$f_{\boldsymbol{v}} = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Projection

choose finite number of basis functions $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k} \in \mathbb{Z}^d}$

$$P_{\boldsymbol{N}} f = \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} \tilde{f}_{\boldsymbol{v}}$$

$$\tilde{f}_{\boldsymbol{v}} = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{N}}^{\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}$$

Regression

compute coefficients $c_{\boldsymbol{k}}^{\star}$ from samples

$$f^{\star} = \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} f_{\boldsymbol{v}}^{\star}$$

$$f_{\boldsymbol{v}}^{\star} = \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{N}}^{\boldsymbol{v}}} c_{\boldsymbol{k}}^{\star}\, \varphi_{\boldsymbol{k}}$$

Truncation

$$
\begin{aligned}
f(\boldsymbol{x}) = f_\varnothing & & \text{constant}\\
+ f_1(x_1) + f_2(x_2) + \ldots + f_d(x_d) & & \text{one-dimensional terms}\\
+ f_{1,2}(x_1, x_2) + f_{1,3}(x_1, x_3) + \ldots + f_{d-1,d}(x_{d-1}, x_d) & & \text{two-dimensional terms}\\
+ f_{1,2,3}(x_1, x_2, x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2}, x_{d-1}, x_d) & & \text{three-dimensional terms}\\
\vdots & & \\
+ f_{[d]}(x) & & \text{d-dimensional term}\\
= \sum_{\boldsymbol{v} \subseteq [d]} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) & &
\end{aligned}
$$

Truncation

$$
\begin{aligned}
f(\boldsymbol{x}) = f_\varnothing & & \text{constant}\\
&+ f_1(x_1) + f_2(x_2) + \ldots + f_d(x_d) & \text{one-dimensional terms}\\
&+ f_{1,2}(x_1, x_2) + f_{1,3}(x_1, x_3) + \ldots + f_{d-1,d}(x_{d-1}, x_d) & \text{two-dimensional terms}\\
&+ f_{1,2,3}(x_1, x_2, x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2}, x_{d-1}, x_d) & \text{three-dimensional terms}\\
&\quad \vdots & \\
&+ f_{[d]}(x) & \text{d-dimensional term}\\
&= \sum_{\boldsymbol{v} \subseteq [d]} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) &
\end{aligned}
$$

Problem: 2^d many terms (curse of dimensionality)

Truncation

$$f(\boldsymbol{x}) = f_\varnothing$$ constant

$$+ f_1(x_1) + f_2(x_2) + \ldots + f_d(x_d)$$ one-dimensional terms

$$+ f_{1,2}(x_1, x_2) + f_{1,3}(x_1, x_3) + \ldots + f_{d-1,d}(x_{d-1}, x_d)$$ two-dimensional terms

$$\cancel{+ f_{1,2,3}(x_1, x_2, x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2}, x_{d-1}, x_d)}$$ ~~three-dimensional terms~~

$$\vdots$$

$$\cancel{+ f_{[d]}(x)}$$ ~~d-dimensional term~~

$$\approx \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ |\boldsymbol{v}| \leq q}} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Problem: 2^d many terms (curse of dimensionality) $\Rightarrow$ introduce $q \in \mathbb{N}, q < d$ (superposition dimension)

Truncation

$$
\begin{aligned}
f(\boldsymbol{x}) = {} & f_\varnothing && \text{constant}\\
& + \cancel{f_1(x_1)} + f_2(x_2) + \ldots + f_d(x_d) && \text{one-dimensional terms}\\
& + f_{1,2}(x_1,x_2) + \cancel{f_{1,3}(x_1,x_3)} + \ldots + \cancel{f_{d-1,d}(x_{d-1},x_d)} && \text{two-dimensional terms}\\
& \cancel{+ f_{1,2,3}(x_1,x_2,x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2},x_{d-1},x_d)} && \cancel{\text{three-dimensional terms}}\\
& \quad \vdots \\
& \cancel{+ f_{[d]}(x)} && \cancel{\text{d-dimensional term}}\\
\approx {} & \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \mathcal{T}_V f
\end{aligned}
$$

set V according to sparsity

Projection

Recap: reduced the number of ANOVA terms

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^d \\ \operatorname{supp}\boldsymbol{k}=\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Projection

Recap: reduced the number of ANOVA terms

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^d \\ \operatorname{supp}\boldsymbol{k}=\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Problem: infinitely many coefficients needed $\Rightarrow$ impossible in practice

Projection

Recap: reduced the number of ANOVA terms

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k} \in \mathbb{Z}^d \\ \operatorname{supp} \boldsymbol{k} = \boldsymbol{v}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Problem: infinitely many coefficients needed $\Rightarrow$ impossible in practice

$\rightsquigarrow$ introduce finite index sets $\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$ and approximate by

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) \approx \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}}), \quad \boldsymbol{v} \subseteq [d]$$

Projection

Recap: reduced the number of ANOVA terms

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^d \\ \operatorname{supp}\boldsymbol{k}=\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Problem: infinitely many coefficients needed $\Rightarrow$ impossible in practice

$\rightsquigarrow$ introduce finite index sets $\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$ and approximate by

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}}), \quad \boldsymbol{v}\subseteq[d]$$

and thus

$$f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}), \qquad \mathcal{I}_{\boldsymbol{N}} := \bigcup_{|\boldsymbol{v}|\leq q} \mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$$

Projection

Recap: reduced the number of ANOVA terms

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^d \\ \operatorname{supp}\boldsymbol{k}=\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Problem: infinitely many coefficients needed $\Rightarrow$ impossible in practice

$\rightsquigarrow$ introduce finite index sets $\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$ and approximate by

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}}), \quad \boldsymbol{v}\subseteq[d]$$

and thus

$$f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}), \qquad \mathcal{I}_{\boldsymbol{N}} := \bigcup_{|\boldsymbol{v}|\le q} \mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$$

$\hat{=}$ projection using a finite dictionary $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} \subset \{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{N}_0^d}$

Projection

Recap: reduced the number of ANOVA terms

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^d \\ \operatorname{supp}\boldsymbol{k}=\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Problem: infinitely many coefficients needed $\Rightarrow$ impossible in practice

$\rightsquigarrow$ introduce finite index sets $\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$ and approximate by

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}}), \quad \boldsymbol{v}\subseteq[d]$$

and thus

$$f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}), \qquad \mathcal{I}_{\boldsymbol{N}} := \bigcup_{|\boldsymbol{v}|\leq q} \mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$$

$\hat{=}$ projection using a finite dictionary $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} \subset \{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{N}_0^d}$

Still: computation of the integrals $c_{\boldsymbol{k}}(f) = \langle f, \varphi_{\boldsymbol{k}}\rangle_{L_2(\Omega,\omega)}, \boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}$

Projection

Recap: reduced the number of ANOVA terms

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \sum_{\substack{\boldsymbol{k}\in\mathbb{Z}^d \\ \operatorname{supp}\boldsymbol{k}=\boldsymbol{v}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}})$$

Problem: infinitely many coefficients needed $\Rightarrow$ impossible in practice

$\rightsquigarrow$ introduce finite index sets $\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$ and approximate by

$$f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}_{\boldsymbol{v}}), \quad \boldsymbol{v}\subseteq[d]$$

and thus

$$f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x}), \qquad \mathcal{I}_{\boldsymbol{N}} := \bigcup_{|\boldsymbol{v}|\leq q} \mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$$

$\hat{=}$ projection using a finite dictionary $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} \subset \{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{N}_0^d}$

Still: computation of the integrals $c_{\boldsymbol{k}}(f) = \langle f,\varphi_{\boldsymbol{k}}\rangle_{L_2(\Omega,\omega)}, \boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}$ $\rightsquigarrow$ data-driven approximation

Regression

Goal: approximate coefficients $c_{\boldsymbol{k}}(f)$ in $f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x})$

- ▷ from samples of the function f
- ▷ at points $\{\boldsymbol{x}^1, \ldots, \boldsymbol{x}^M\}$ i.i.d. random according to the density ω

Regression

Goal: approximate coefficients $c_{\boldsymbol{k}}(f)$ in $f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k} \in \mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\, \varphi_{\boldsymbol{k}}(\boldsymbol{x})$

- ▷ from samples of the function f
- ▷ at points $\{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\}$ i.i.d. random according to the density ω,

i. e.,

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A} \in \mathbb{R}^{M \times |\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}}$$

Regression

Goal: approximate coefficients $c_{\boldsymbol{k}}(f)$ in $f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x})$

- ▷ from samples of the function f
- ▷ at points $\{\boldsymbol{x}^1,\dots,\boldsymbol{x}^M\}$ i.i.d. random according to the density ω,

i. e.,

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A}\in\mathbb{R}^{M\times|\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}} \qquad \rightsquigarrow \text{ minimize } \|\boldsymbol{A}\boldsymbol{c}-\boldsymbol{f}\|_2^2$$

Regression

Goal: approximate coefficients $c_{\boldsymbol{k}}(f)$ in $f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x})$

- ▷ from samples of the function f
- ▷ at points $\{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\}$ i.i.d. random according to the density ω,

i. e.,

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A}\in\mathbb{R}^{M\times|\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}} \qquad \rightsquigarrow \text{ minimize } \|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{f}\|_2^2$$

- least squares solution $\boldsymbol{c}^\star = (\boldsymbol{A}^\top \boldsymbol{A})^{-1}\boldsymbol{A}^\top \boldsymbol{f}$
  - [Kämmerer, Ullrich, Volkmer 21]: good condition number with high probability, if $|\mathcal{I}_{\boldsymbol{N}}| < \frac{M}{\log M}$
  - [Bartel, Potts, Schmischke 22]: can be computed efficiently (LSQR + fast multiplication)

Regression

Goal: approximate coefficients $c_{\boldsymbol{k}}(f)$ in $f(\boldsymbol{x}) \approx \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c_{\boldsymbol{k}}(f)\,\varphi_{\boldsymbol{k}}(\boldsymbol{x})$

- ▷ from samples of the function f
- ▷ at points $\{\boldsymbol{x}^1, \dots, \boldsymbol{x}^M\}$ i.i.d. random according to the density ω,

i. e.,

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A}\in\mathbb{R}^{M\times|\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}} \qquad \rightsquigarrow \text{ minimize } \|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{f}\|_2^2$$

- least squares solution $\boldsymbol{c}^\star = (\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top\boldsymbol{f}$
  - [Kämmerer, Ullrich, Volkmer 21]: good condition number with high probability, if $|\mathcal{I}_{\boldsymbol{N}}| < \frac{M}{\log M}$
  - [Bartel, Potts, Schmischke 22]: can be computed efficiently (LSQR + fast multiplication)
- final approximation

$$f^\star(\boldsymbol{x}) := \sum_{\boldsymbol{k}\in\mathcal{I}_{\boldsymbol{N}}} c^\star_{\boldsymbol{k}}\,\varphi_{\boldsymbol{k}}(\boldsymbol{x})$$

ANOVA approximation for control systems

Setting

Unknown function:

$$H : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F : \mathbb{R}^d \to \mathbb{R}^d$,
- $G : \mathbb{R}^d \to \mathbb{R}^{d \times m}$,
- $d, m \in \mathbb{N}$, $m \ll d$,

Setting

Unknown function:

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x},\boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d,$
- $G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m},$
- $d, m \in \mathbb{N},\, m \ll d,$

Assumption: F and G of low-dimensional structure

Setting

Unknown function:

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d$,
- $G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d$,

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i), i = 1, \dots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Setting

Unknown function:

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d$,
- $G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}$, $m \ll d$,

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i)$, $i = 1, \dots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Goal: find $\tilde{H} \approx H$

Setting

Unknown function:

$$H : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F : \mathbb{R}^d \to \mathbb{R}^d,$
- $G : \mathbb{R}^d \to \mathbb{R}^{d \times m},$
- $d, m \in \mathbb{N}, m \ll d,$

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i), i = 1, \ldots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Goal: find $\tilde{H} \approx H$

Approach: ANOVA approximation

Setting

Unknown function:

$$H : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F : \mathbb{R}^d \to \mathbb{R}^d,$
- $G : \mathbb{R}^d \to \mathbb{R}^{d \times m},$
- $d, m \in \mathbb{N}, m \ll d,$

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i), i = 1, \ldots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Goal: find $\tilde{H} \approx H$

Approach: ANOVA approximation

Problem: F and G cannot be sampled separately

Setting

Unknown function:

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d$,
- $G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d$,

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i)$, $i = 1, \ldots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Goal: find $\tilde{H} \approx H$

Approach: ANOVA approximation

Problem: F and G cannot be sampled separately
$\rightsquigarrow$ individual approximation not possible

Setting

Unknown function:

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x},\boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d$,
- $G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d$,

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i), i = 1, \ldots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Goal: find $\tilde{H} \approx H$

Approach: ANOVA approximation

Problem: F and G cannot be sampled separately

$\rightsquigarrow$ individual approximation not possible

$\rightsquigarrow$ coupled reconstruction

Setting

Unknown function:

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x}, \boldsymbol{u}) := F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d$,
- $G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d$,

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i), i = 1, \ldots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Goal: find $\tilde{H} \approx H$

Approach: ANOVA approximation

Problem: F and G cannot be sampled separately

$\rightsquigarrow$ individual approximation not possible

$\rightsquigarrow$ coupled reconstruction

Important: avoid ANOVA approximation of H in $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{u})$
for $m > 1$ introduces many additional terms – nonexistent interactions of the components of $\boldsymbol{u}$
$\rightsquigarrow$ unnecessary costs + error

Setting

Unknown function:

$$H\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,$$
$$H(\boldsymbol{x},\boldsymbol{u}) \coloneqq F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u},$$

with

- $F\colon \mathbb{R}^d \to \mathbb{R}^d$,
- $G\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$,
- $d, m \in \mathbb{N}, m \ll d$,

Assumption: F and G of low-dimensional structure

Given:

- sampling points $(\boldsymbol{x}^i, \boldsymbol{u}^i), i = 1, \ldots, M$
- samples $H(\boldsymbol{x}^i, \boldsymbol{u}^i)$

Goal: find $\tilde{H} \approx H$

Approach: ANOVA approximation

Problem: F and G cannot be sampled separately

$\rightsquigarrow$ individual approximation not possible

$\rightsquigarrow$ coupled reconstruction

Important: avoid ANOVA approximation of H in $\boldsymbol{z} = (\boldsymbol{x}, \boldsymbol{u})$
for $m > 1$ introduces many additional terms – nonexistent interactions of the components of $\boldsymbol{u}$
$\rightsquigarrow$ unnecessary costs + error

$\rightsquigarrow$ exploit linearity in $\boldsymbol{u}$ instead

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$H(\boldsymbol{x}, \mathbf{0}) = F(\boldsymbol{x}),$$

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$\begin{aligned}
H(\boldsymbol{x}, \mathbf{0}) &= F(\boldsymbol{x}), \\
H(\boldsymbol{x}, \boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\
&= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1, \dots, m,
\end{aligned}$$

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$
\begin{aligned}
H(\boldsymbol{x}, \mathbf{0}) &= F(\boldsymbol{x}), \\
H(\boldsymbol{x}, \boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\
&= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1, \dots, m,
\end{aligned}
$$

i. e.,

$$
\begin{aligned}
G_\ell(\boldsymbol{x}) &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - F(\boldsymbol{x}) \\
&= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - H(\boldsymbol{x}, \mathbf{0}), \quad \ell = 1, \dots, m
\end{aligned}
$$

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$\begin{aligned}
H(\boldsymbol{x}, \mathbf{0}) &= F(\boldsymbol{x}), \\
H(\boldsymbol{x}, \boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\
&= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1, \dots, m,
\end{aligned}$$

i. e.,

$$\begin{aligned}
G_\ell(\boldsymbol{x}) &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - F(\boldsymbol{x}) \\
&= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - H(\boldsymbol{x}, \mathbf{0}), \quad \ell = 1, \dots, m
\end{aligned}$$

$\Rightarrow$ may produce samples of F and G artificially

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$H(\boldsymbol{x}, \mathbf{0}) = F(\boldsymbol{x}),$$

$$\begin{aligned} H(\boldsymbol{x}, \boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\ &= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1, \dots, m, \end{aligned}$$

i. e.,

$$\begin{aligned} G_\ell(\boldsymbol{x}) &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - F(\boldsymbol{x}) \\ &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - H(\boldsymbol{x}, \mathbf{0}), \quad \ell = 1, \dots, m \end{aligned}$$

$\Rightarrow$ may produce samples of F and G artificially

Remark: ANOVA only for scalar-valued functions

$\rightsquigarrow$ proceed rowwise:

$$H_j(\boldsymbol{x}, \boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell, \quad j = 1, \dots, d$$

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$
\begin{aligned}
H(\boldsymbol{x}, \mathbf{0}) &= F(\boldsymbol{x}), \\
H(\boldsymbol{x}, \boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\
&= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1, \dots, m,
\end{aligned}
$$

i. e.,

$$
\begin{aligned}
G_\ell(\boldsymbol{x}) &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - F(\boldsymbol{x}) \\
&= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - H(\boldsymbol{x}, \mathbf{0}), \quad \ell = 1, \dots, m
\end{aligned}
$$

$\Rightarrow$ may produce samples of F and G artificially

Remark: ANOVA only for scalar-valued functions

$\rightsquigarrow$ proceed rowwise:

$$
H_j(\boldsymbol{x}, \boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^m G_{j\ell}(\boldsymbol{x})\, u_\ell, \quad j = 1, \dots, d
$$

- samples $\left(H_j(\boldsymbol{x}^i, \mathbf{0})\right)_{i=1}^M$ for $\tilde{F}_j$

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$H(\boldsymbol{x}, \mathbf{0}) = F(\boldsymbol{x}),$$

$$\begin{aligned} H(\boldsymbol{x}, \boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\ &= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1, \dots, m, \end{aligned}$$

i. e.,

$$\begin{aligned} G_\ell(\boldsymbol{x}) &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - F(\boldsymbol{x}) \\ &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - H(\boldsymbol{x}, \mathbf{0}), \quad \ell = 1, \dots, m \end{aligned}$$

$\Rightarrow$ may produce samples of F and G artificially

Remark: ANOVA only for scalar-valued functions

$\rightsquigarrow$ proceed rowwise:

$$H_j(\boldsymbol{x}, \boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^m G_{j\ell}(\boldsymbol{x})\, u_\ell, \quad j = 1, \dots, d$$

- samples $\left(H_j(\boldsymbol{x}^i, \mathbf{0})\right)_{i=1}^M$ for $\tilde{F}_j$
- samples $\left(H_j(\boldsymbol{x}^i, \boldsymbol{e}_\ell) - H_j(\boldsymbol{x}^i, \mathbf{0})\right)_{i=1}^M$ for $\tilde{G}_{j\ell}$

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$H(\boldsymbol{x}, \mathbf{0}) = F(\boldsymbol{x}),$$

$$\begin{aligned} H(\boldsymbol{x}, \boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\ &= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1, \dots, m, \end{aligned}$$

i. e.,

$$\begin{aligned} G_\ell(\boldsymbol{x}) &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - F(\boldsymbol{x}) \\ &= H(\boldsymbol{x}, \boldsymbol{e}_\ell) - H(\boldsymbol{x}, \mathbf{0}), \quad \ell = 1, \dots, m \end{aligned}$$

$\Rightarrow$ may produce samples of F and G artificially

Remark: ANOVA only for scalar-valued functions

$\rightsquigarrow$ proceed rowwise:

$$H_j(\boldsymbol{x}, \boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^m G_{j\ell}(\boldsymbol{x})\, u_\ell, \quad j = 1, \dots, d$$

- samples $\left(H_j(\boldsymbol{x}^i, \mathbf{0})\right)_{i=1}^M$ for $\tilde{F}_j$
- samples $\left(H_j(\boldsymbol{x}^i, \boldsymbol{e}_\ell) - H_j(\boldsymbol{x}^i, \mathbf{0})\right)_{i=1}^M$ for $\tilde{G}_{j\ell}$

✓ preserves structure (linearity in $\boldsymbol{u}$)

Unit vectors as control

Simplest approach: $\mathbb{U} = \{\mathbf{0}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ with $\{\boldsymbol{e}_\ell\}_{\ell=1}^m$ unit vectors of $\mathbb{R}^m$

evaluation of $H(\boldsymbol{x},\boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$ yields

$$H(\boldsymbol{x},\mathbf{0}) = F(\boldsymbol{x}),$$

$$\begin{aligned} H(\boldsymbol{x},\boldsymbol{e}_\ell) &= \begin{pmatrix} F_1(\boldsymbol{x}) \\ \vdots \\ F_d(\boldsymbol{x}) \end{pmatrix} + \begin{pmatrix} G_{11}(\boldsymbol{x}) & \dots & G_{1m}(\boldsymbol{x}) \\ \vdots & & \vdots \\ G_{d1}(\boldsymbol{x}) & \dots & G_{dm}(\boldsymbol{x}) \end{pmatrix} \boldsymbol{e}_\ell \\ &= F(\boldsymbol{x}) + G_\ell(\boldsymbol{x}), \quad \ell = 1,\dots,m, \end{aligned}$$

i. e.,

$$\begin{aligned} G_\ell(\boldsymbol{x}) &= H(\boldsymbol{x},\boldsymbol{e}_\ell) - F(\boldsymbol{x}) \\ &= H(\boldsymbol{x},\boldsymbol{e}_\ell) - H(\boldsymbol{x},\mathbf{0}), \quad \ell = 1,\dots,m \end{aligned}$$

$\Rightarrow$ may produce samples of F and G artificially

Remark: ANOVA only for scalar-valued functions

$\rightsquigarrow$ proceed rowwise:

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^m G_{j\ell}(\boldsymbol{x})\, u_\ell, \quad j = 1,\dots,d$$

- samples $\left(H_j(\boldsymbol{x}^i,\mathbf{0})\right)_{i=1}^M$ for $\tilde{F}_j$
- samples $\left(H_j(\boldsymbol{x}^i,\boldsymbol{e}_\ell) - H_j(\boldsymbol{x}^i,\mathbf{0})\right)_{i=1}^M$ for $\tilde{G}_{j\ell}$

✓ preserves structure (linearity in $\boldsymbol{u}$)

✗ $M(m+1)$ function evaluations necessary
- at each unit vector all x^i
- non-flexible

Matrix approach

Aim: preserve structure + use only one set of samples of H

Matrix approach

Aim: preserve structure + use only one set of samples of H

Observation: ANOVA approximation separately for F and G would mean

$$H_j(\boldsymbol{x}, \boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell \qquad , \quad j = 1, \dots, d$$

Matrix approach

Aim: preserve structure + use only one set of samples of H

Observation: ANOVA approximation separately for F and G would mean

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell \approx \sum_{\boldsymbol{k}\in\mathcal{I}(V_1)} c^1_{\boldsymbol{k}} \varphi^1_{\boldsymbol{k}}(\boldsymbol{x}) + \qquad\qquad , \quad j = 1,\dots,d$$

Matrix approach

Aim: preserve structure + use only one set of samples of H

Observation: ANOVA approximation separately for F and G would mean

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\,u_\ell \approx \sum_{\boldsymbol{k}\in\mathcal{I}(V_1)} c_{\boldsymbol{k}}^1 \varphi_{\boldsymbol{k}}^1(\boldsymbol{x}) + \sum_{\ell=1}^{m}\left(\sum_{\boldsymbol{s}\in\mathcal{I}(V_{\ell+1})} c_{\boldsymbol{s}}^{\ell+1}\varphi_{\boldsymbol{s}}^{\ell+1}(\boldsymbol{x})\right)u_\ell, \quad j = 1,\dots,d$$

Matrix approach

Aim: preserve structure + use only one set of samples of H

Observation: ANOVA approximation separately for F and G would mean

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell \approx \sum_{\boldsymbol{k}\in\mathcal{I}(V_1)} c_{\boldsymbol{k}}^1 \varphi_{\boldsymbol{k}}^1(\boldsymbol{x}) + \sum_{\ell=1}^{m}\left(\sum_{\boldsymbol{s}\in\mathcal{I}(V_{\ell+1})} c_{\boldsymbol{s}}^{\ell+1}\varphi_{\boldsymbol{s}}^{\ell+1}(\boldsymbol{x})\right) u_\ell, \quad j=1,\dots,d$$

$\Rightarrow$ for $(\boldsymbol{x}^i,\boldsymbol{u}^i)$ i.i.d. random:

$$\boldsymbol{c}^\ell := \left(c_{\boldsymbol{k}}^\ell\right)_{\boldsymbol{k}\in\mathcal{I}(V_\ell)}, \quad \boldsymbol{A}_\ell := \left(\varphi_{\boldsymbol{k}}^\ell(\boldsymbol{x}^i)\right)_{i=1,\,\boldsymbol{k}\in\mathcal{I}(V_\ell)}^{M}, \quad \boldsymbol{U}_\ell := \operatorname{diag}(u_\ell^1,\dots,u_\ell^M)$$

Matrix approach

Aim: preserve structure + use only one set of samples of H

Observation: ANOVA approximation separately for F and G would mean

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell \approx \sum_{\boldsymbol{k}\in\mathcal{I}(V_1)} c_{\boldsymbol{k}}^1 \varphi_{\boldsymbol{k}}^1(\boldsymbol{x}) + \sum_{\ell=1}^{m} \left(\sum_{\boldsymbol{s}\in\mathcal{I}(V_{\ell+1})} c_{\boldsymbol{s}}^{\ell+1} \varphi_{\boldsymbol{s}}^{\ell+1}(\boldsymbol{x}) \right) u_\ell, \quad j = 1,\dots,d$$

$\Rightarrow$ for $(\boldsymbol{x}^i, \boldsymbol{u}^i)$ i.i.d. random:

$$\boldsymbol{c}^\ell := \left(c_{\boldsymbol{k}}^\ell\right)_{\boldsymbol{k}\in\mathcal{I}(V_\ell)}, \quad \boldsymbol{A}_\ell := \left(\varphi_{\boldsymbol{k}}^\ell(\boldsymbol{x}^i)\right)_{i=1,\,\boldsymbol{k}\in\mathcal{I}(V_\ell)}^{M}, \quad \boldsymbol{U}_\ell := \operatorname{diag}(u_\ell^1,\dots,u_\ell^M)$$

yields

$$\left(H_j(\boldsymbol{x}^i,\boldsymbol{u}^i)\right)_{i=1}^{M} \approx \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{U}_1\boldsymbol{A}_2 & \dots & \boldsymbol{U}_m\boldsymbol{A}_{m+1} \end{pmatrix} \begin{pmatrix} \boldsymbol{c}^1 \\ \vdots \\ \boldsymbol{c}^{m+1} \end{pmatrix}$$

Matrix approach

Aim: preserve structure + use only one set of samples of H

Observation: ANOVA approximation separately for F and G would mean

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell \approx \sum_{\boldsymbol{k}\in\mathcal{I}(V_1)} c_{\boldsymbol{k}}^1 \varphi_{\boldsymbol{k}}^1(\boldsymbol{x}) + \sum_{\ell=1}^{m} \left(\sum_{\boldsymbol{s}\in\mathcal{I}(V_{\ell+1})} c_{\boldsymbol{s}}^{\ell+1} \varphi_{\boldsymbol{s}}^{\ell+1}(\boldsymbol{x}) \right) u_\ell, \quad j = 1,\dots,d$$

$\Rightarrow$ for $(\boldsymbol{x}^i, \boldsymbol{u}^i)$ i.i.d. random:

$$\boldsymbol{c}^\ell := \left(c_{\boldsymbol{k}}^\ell\right)_{\boldsymbol{k}\in\mathcal{I}(V_\ell)}, \quad \boldsymbol{A}_\ell := \left(\varphi_{\boldsymbol{k}}^\ell(\boldsymbol{x}^i)\right)_{i=1,\,\boldsymbol{k}\in\mathcal{I}(V_\ell)}^{M}, \quad \boldsymbol{U}_\ell := \operatorname{diag}(u_\ell^1,\dots,u_\ell^M)$$

yields

$$\left(H_j(\boldsymbol{x}^i,\boldsymbol{u}^i)\right)_{i=1}^{M} \approx \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{U}_1\boldsymbol{A}_2 & \dots & \boldsymbol{U}_m\boldsymbol{A}_{m+1} \end{pmatrix} \begin{pmatrix} \boldsymbol{c}^1 \\ \vdots \\ \boldsymbol{c}^{m+1} \end{pmatrix}$$

$\Rightarrow$ find least squares solution

Efficient implementation

[https://github.com/NFFT/ANOVAapprox.jl]

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\,u_\ell, \quad j = 1,\dots,d$$

Recap – ANOVA approximation:

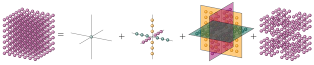
Truncation

$$\begin{aligned} f(\boldsymbol{x}) = {} & f_\varnothing \\ & + \cancel{f_1(x_1)} + f_2(x_2) + \ldots + f_d(x_d) \\ & + f_{1,2}(x_1,x_2) + \cancel{f_{1,3}(x_1,x_3)} + \ldots + \cancel{f_{d-1,d}(x_{d-1},x_d)} \\ & \cancel{+ f_{1,2,3}(x_1,x_2,x_3) + \ldots + f_{d-2,d-1,d}(x_{d-2},x_{d-1},x_d)} \\ & \vdots \\ & \cancel{+ f_{[d]}(\boldsymbol{x})} \\ \approx {} & \sum_{\substack{\boldsymbol{u}\subseteq[d] \\ \boldsymbol{u}\in V}} f_{\boldsymbol{u}}(\boldsymbol{x}_{\boldsymbol{u}}) = T_V f \end{aligned}$$

reduce number of ANOVA terms

Projection

choose finite number of basis functions $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$



$\boldsymbol{k}\in\mathbb{Z}^3$ $\quad$ $\boldsymbol{k}=\boldsymbol{0}$ $\quad$ $|\operatorname{supp}\boldsymbol{k}|=1$ $\quad$ $|\operatorname{supp}\boldsymbol{k}|=2$ $\quad$ $|\operatorname{supp}\boldsymbol{k}|=3$

Regression

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A}\in\mathbb{R}^{M\times|\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}}$$

$\rightsquigarrow$ minimize $\|\boldsymbol{A}\boldsymbol{c}-\boldsymbol{f}\|_2^2$

Efficient implementation

[https://github.com/NFFT/ANOVAapprox.jl]

$$H_j(\boldsymbol{x},\boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\,u_\ell, \quad j = 1,\dots,d$$

Recap – ANOVA approximation:

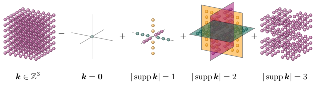
Truncation

$$\begin{aligned} f(\boldsymbol{x}) = {} & f_\varnothing \\ & + \cancel{f_1(x_1)} + f_2(x_2) + \dots + f_d(x_d) \\ & + f_{1,2}(x_1,x_2) + \cancel{f_{1,3}(x_1,x_3)} + \dots + \cancel{f_{d-1,d}(x_{d-1},x_d)} \\ & \cancel{+ f_{1,2,3}(x_1,x_2,x_3) + \dots + f_{d-2,d-1,d}(x_{d-2},x_{d-1},x_d)} \\ & \vdots \\ & \cancel{+ f_{[d]}(\boldsymbol{x})} \\ \approx {} & \sum_{\substack{\boldsymbol{v}\subseteq[d] \\ \boldsymbol{v}\in V}} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \mathrm{T}_V f \end{aligned}$$

$\rightsquigarrow$ set V without interactions between the u_ℓ

Projection

choose finite number of basis functions $\{\varphi_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\mathbb{Z}^d}$



Regression

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A}\in\mathbb{R}^{M\times|\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}}$$

$\rightsquigarrow$ minimize $\|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{f}\|_2^2$

Efficient implementation

[https://github.com/NFFT/ANOVAapprox.jl]

$$H_j(\boldsymbol{x}, \boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell, \quad j = 1, \dots, d$$

Recap – ANOVA approximation:

Truncation

$$\begin{aligned} f(\boldsymbol{x}) = {} & f_\varnothing \\ & + \cancel{f_1(x_1)} + f_2(x_2) + \dots + f_d(x_d) \\ & + f_{1,2}(x_1, x_2) + \cancel{f_{1,3}(x_1, x_3)} + \dots + \cancel{f_{d-1,d}(x_{d-1}, x_d)} \\ & \cancel{+ f_{1,2,3}(x_1, x_2, x_3) + \dots + f_{d-2,d-1,d}(x_{d-2}, x_{d-1}, x_d)} \\ & \quad \vdots \\ & \cancel{+ f_{[d]}(\boldsymbol{x})} \\ \approx {} & \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \mathcal{T}_V f \end{aligned}$$

$\rightsquigarrow$ set V without interactions between the u_ℓ

Projection

use the Chebyshev basis

$$\{\varphi_k^{\text{cheb}}\}_{k \in \mathbb{N}_0} = \{1, \sqrt{2}x, \dots\}$$

+

choose only one basis function for all u_ℓ

$\Rightarrow$ exact reconstruction of all u_ℓ

Regression

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A} \in \mathbb{R}^{M \times |\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}}$$

$\rightsquigarrow$ minimize $\|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{f}\|_2^2$

Efficient implementation

[https://github.com/NFFT/ANOVAapprox.jl]

$$H_j(\boldsymbol{x}, \boldsymbol{u}) = F_j(\boldsymbol{x}) + \sum_{\ell=1}^{m} G_{j\ell}(\boldsymbol{x})\, u_\ell, \quad j = 1, \dots, d$$

Recap – ANOVA approximation:

Truncation

$$\begin{aligned} f(\boldsymbol{x}) = & f_\varnothing \\ & + \cancel{f_1(x_1)} + f_2(x_2) + \dots + f_d(x_d) \\ & + f_{1,2}(x_1, x_2) + \cancel{f_{1,3}(x_1, x_3)} + \dots + \cancel{f_{d-1,d}(x_{d-1}, x_d)} \\ & \cancel{+ f_{1,2,3}(x_1, x_2, x_3) + \dots + f_{d-2,d-1,d}(x_{d-2}, x_{d-1}, x_d)} \\ & \quad \vdots \\ & \cancel{+ f_{[d]}(x)} \\ \approx & \sum_{\substack{\boldsymbol{v} \subseteq [d] \\ \boldsymbol{v} \in V}} f_{\boldsymbol{v}}(\boldsymbol{x}_{\boldsymbol{v}}) = \mathcal{T}_V f \end{aligned}$$

$\rightsquigarrow$ set V without interactions between the u_ℓ

Projection

use the Chebyshev basis

$\{\varphi_k^{\mathrm{cheb}}\}_{k \in \mathbb{N}_0} = \{1, \sqrt{2}x, \dots\}$

+

choose only one basis function for all u_ℓ

$\Rightarrow$ exact reconstruction of all u_ℓ

Regression

$$\underbrace{\begin{pmatrix} \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^1) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^1) \\ \vdots & & \vdots \\ \varphi_{\boldsymbol{k}_1}(\boldsymbol{x}^M) & \cdots & \varphi_{\boldsymbol{k}_N}(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{A} \in \mathbb{R}^{M \times |\mathcal{I}_{\boldsymbol{N}}|}} \underbrace{\begin{pmatrix} c_{\boldsymbol{k}_1} \\ \vdots \\ c_{\boldsymbol{k}_N} \end{pmatrix}}_{\boldsymbol{c}} \approx \underbrace{\begin{pmatrix} f(\boldsymbol{x}^1) \\ \vdots \\ f(\boldsymbol{x}^M) \end{pmatrix}}_{\boldsymbol{f}}$$

$\rightsquigarrow$ minimize $\|\boldsymbol{A}\boldsymbol{c} - \boldsymbol{f}\|_2^2$

$\Rightarrow$ only need suitable preprocessing step (adjust the index sets V and $\mathcal{I}_{\boldsymbol{N}^{\boldsymbol{v}}}$ appropriately)

Numerical Example

Toy example

- $d = 8$ and $m = 1$
- consider only one component

$$H_j(\boldsymbol{x}, u) = \underbrace{(x_4 - x_1)\, x_2}_{F(\boldsymbol{x})} - \underbrace{x_3}_{G(\boldsymbol{x})} u$$

- maximum two-dimensional terms
$\rightsquigarrow q = 2$

Toy example

- $d = 8$ and $m = 1$
- consider only one component

$$H_j(\boldsymbol{x}, u) = \underbrace{(x_4 - x_1)\, x_2}_{F(\boldsymbol{x})} - \underbrace{x_3}_{G(\boldsymbol{x})} u$$

- maximum two-dimensional terms
  $\rightsquigarrow q = 2$
- Chebyshev basis for both approaches
  - ▷ approach 1 – unit vectors
  - ▷ approach 2 – matrix approach

Toy example

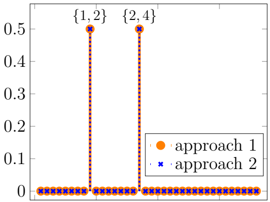
- $d = 8$ and $m = 1$
- consider only one component

$$H_j(\boldsymbol{x}, u) = \underbrace{(x_4 - x_1)\, x_2}_{F(\boldsymbol{x})} - \underbrace{x_3}_{G(\boldsymbol{x})}\, u$$

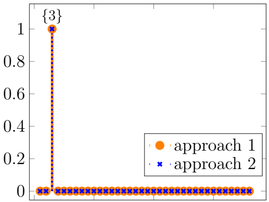
- maximum two-dimensional terms
  $\rightsquigarrow q = 2$
- Chebyshev basis for both approaches
  - approach 1 – unit vectors
  - approach 2 – matrix approach

1. relevant ANOVA terms

of F



of G



$\Rightarrow$ correctly identified

Toy example

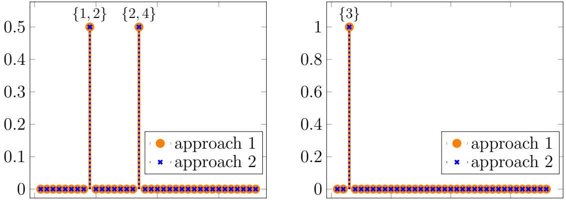
- $d = 8$ and $m = 1$
- consider only one component

$$H_j(\boldsymbol{x}, u) = \underbrace{(x_4 - x_1)\, x_2}_{F(\boldsymbol{x})} - \underbrace{x_3}_{G(\boldsymbol{x})}\, u$$

- maximum two-dimensional terms
  $\rightsquigarrow q = 2$
- Chebyshev basis for both approaches
  - approach 1 – unit vectors
  - approach 2 – matrix approach

1. relevant ANOVA terms

of F of G



$\Rightarrow$ correctly identified

2. error results $\rightsquigarrow$ same order of magnitude 10^{-10}

Summary

- investigated ANOVA in the context of control problems $H(\boldsymbol{x},\boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$

Summary

- investigated ANOVA in the context of control problems $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$
- introduced 2 methods to exploit low-dimensional structures of F and G, especially linearity in $\boldsymbol{u}$
  - ✓ easy to use due to free software package [https://github.com/NFFT/ANOVAapprox.jl]
  - ✓ efficient computation

Summary

- investigated ANOVA in the context of control problems $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$
- introduced 2 methods to exploit low-dimensional structures of F and G, especially linearity in $\boldsymbol{u}$
  - ✓ easy to use due to free software package [https://github.com/NFFT/ANOVAapprox.jl]
  - ✓ efficient computation

1. simple approach using unit vectors as control
   - ✓ allows to approximate F and G separately due to artificial samples
   - ✗ many samples needed $\rightsquigarrow$ possibly unfeasible in practice

Summary

- investigated ANOVA in the context of control problems $H(\boldsymbol{x},\boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$
- introduced 2 methods to exploit low-dimensional structures of F and G, especially linearity in $\boldsymbol{u}$
  - ✓ easy to use due to free software package [https://github.com/NFFT/ANOVAapprox.jl]
  - ✓ efficient computation

1. simple approach using unit vectors as control
   - ✓ allows to approximate F and G separately due to artificial samples
   - ✗ many samples needed $\rightsquigarrow$ possibly unfeasible in practice
2. additional approach to only need one set of samples $\rightsquigarrow$ determined suitable preprocessing
   - ✓ preserve linearity – exact reconstruction of components of $\boldsymbol{u}$
   - ✓ only one set of samples
   - ✓ no additional ANOVA terms introduced

Summary

- investigated ANOVA in the context of control problems $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$
- introduced 2 methods to exploit low-dimensional structures of F and G, especially linearity in $\boldsymbol{u}$
  - ✓ easy to use due to free software package [`https://github.com/NFFT/ANOVAapprox.jl`]
  - ✓ efficient computation

1. simple approach using unit vectors as control
   - ✓ allows to approximate F and G separately due to artificial samples
   - ✗ many samples needed $\rightsquigarrow$ possibly unfeasible in practice
2. additional approach to only need one set of samples $\rightsquigarrow$ determined suitable preprocessing
   - ✓ preserve linearity – exact reconstruction of components of $\boldsymbol{u}$
   - ✓ only one set of samples
   - ✓ no additional ANOVA terms introduced

- **K., Potts, Schaller, Worthmann:** **ANOVA approximation for control systems**. In preparation.

Summary

- investigated ANOVA in the context of control problems $H(\boldsymbol{x}, \boldsymbol{u}) = F(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$
- introduced 2 methods to exploit low-dimensional structures of F and G, especially linearity in $\boldsymbol{u}$
  - ✓ easy to use due to free software package [https://github.com/NFFT/ANOVAapprox.jl]
  - ✓ efficient computation

1. simple approach using unit vectors as control
   - ✓ allows to approximate F and G separately due to artificial samples
   - ✗ many samples needed $\rightsquigarrow$ possibly unfeasible in practice
2. additional approach to only need one set of samples $\rightsquigarrow$ determined suitable preprocessing
   - ✓ preserve linearity – exact reconstruction of components of $\boldsymbol{u}$
   - ✓ only one set of samples
   - ✓ no additional ANOVA terms introduced

- **K., Potts, Schaller, Worthmann:** **ANOVA approximation for control systems**. In preparation.

Thank you for your attention!