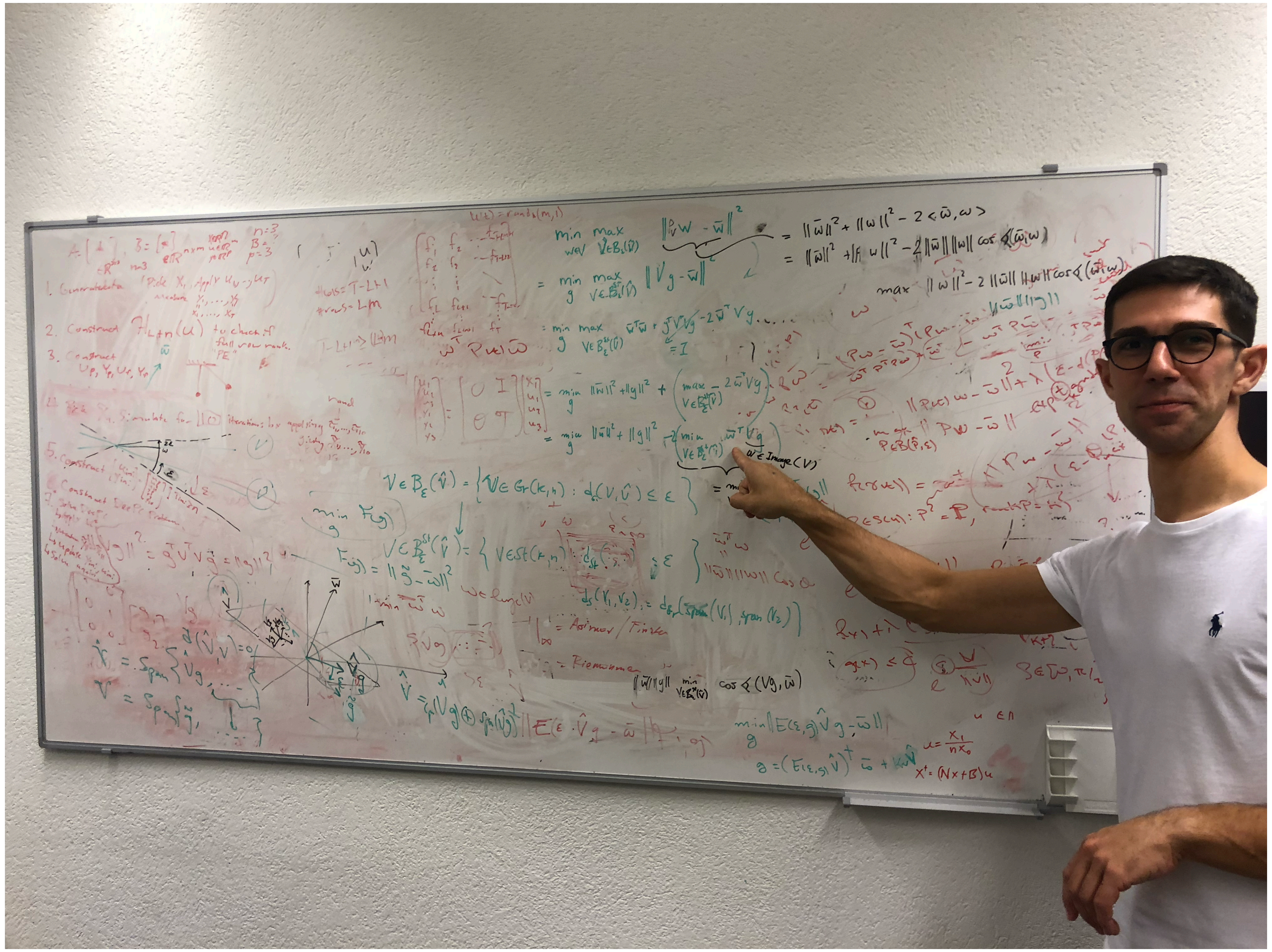
Distances between finite-horizon linear behaviors

Jeremy Coulson

CDC 2025

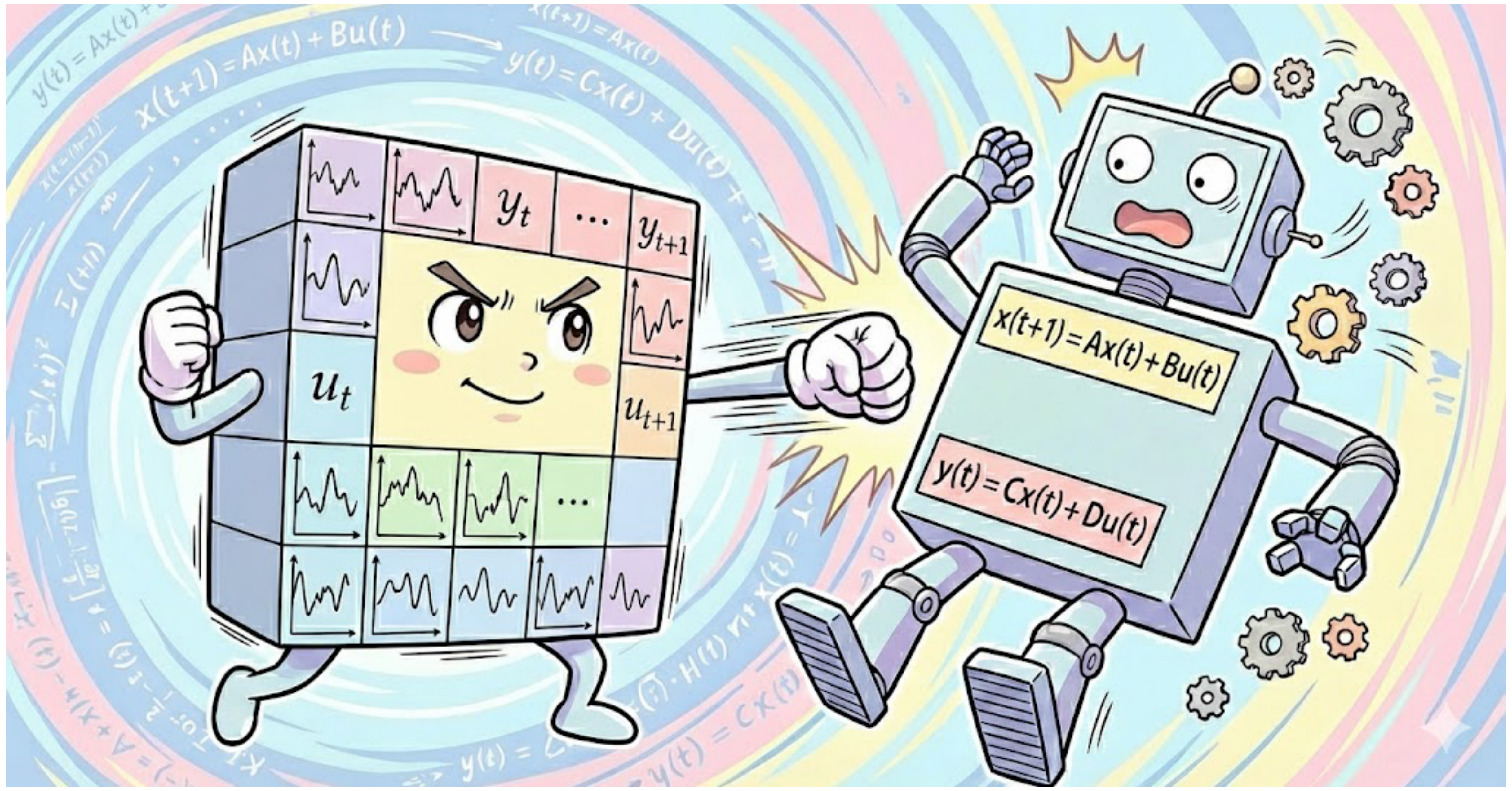
Acknowledgments

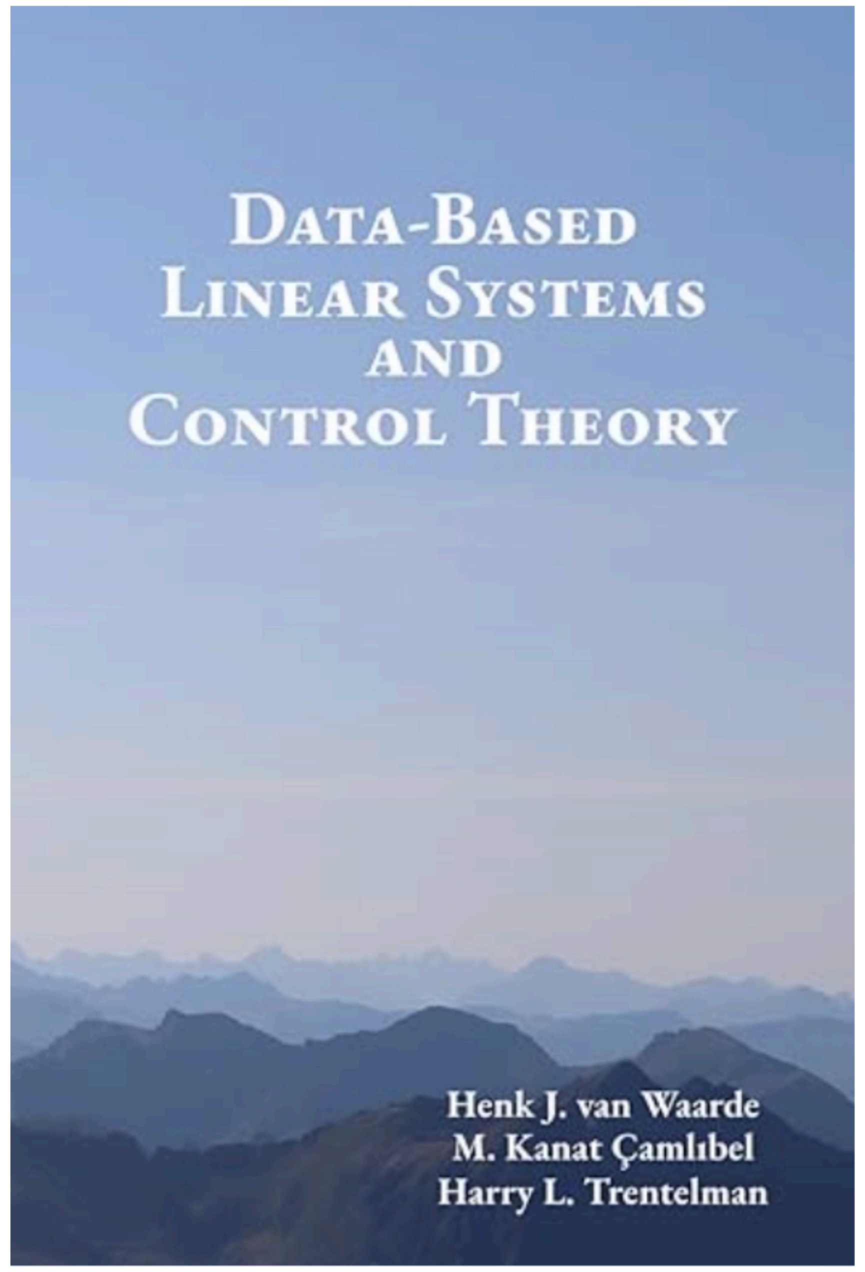


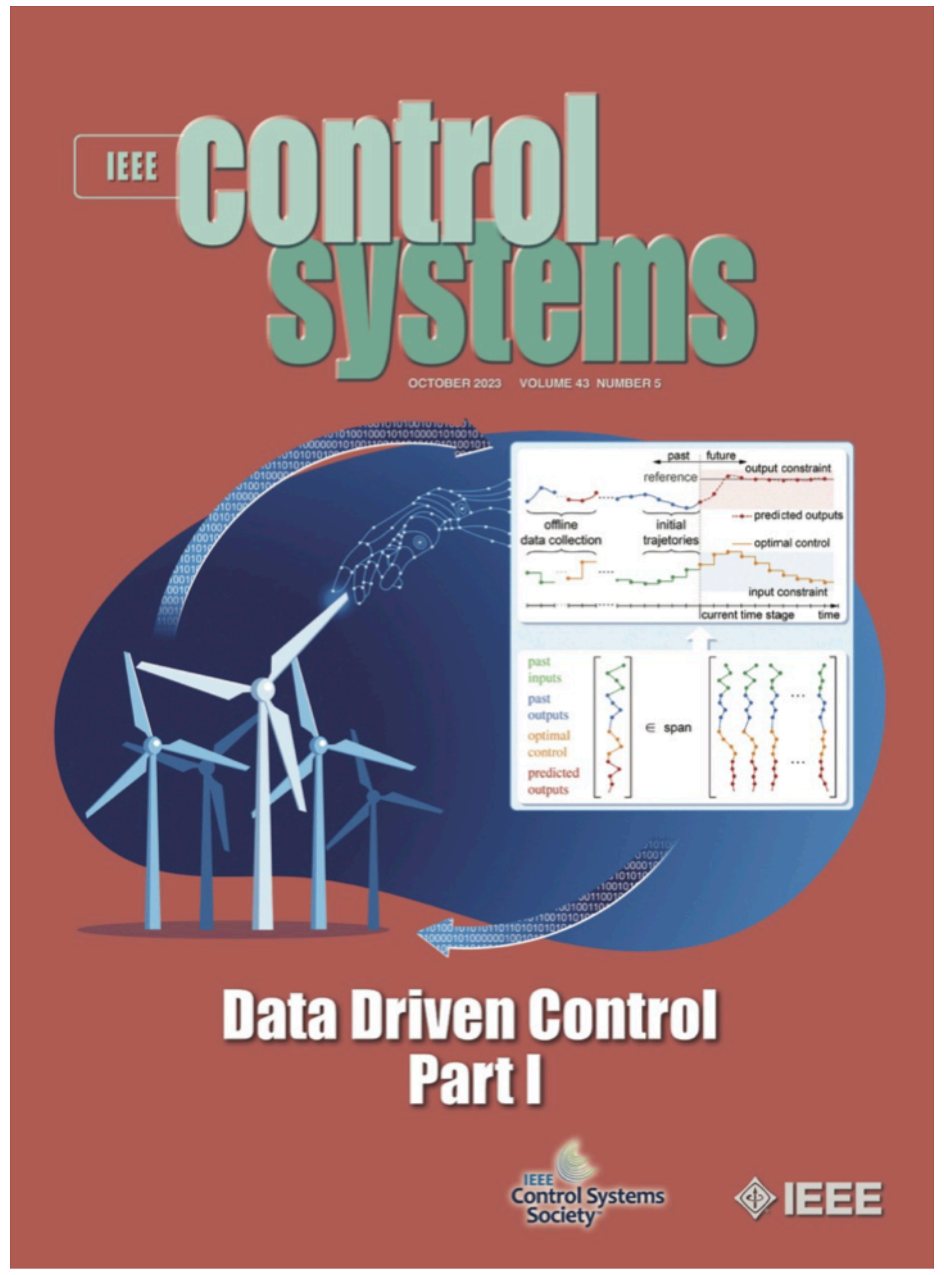
Alberto Padoan

UBC THE UNIVERSITY OF BRITISH COLUMBIA

Data matrices as models



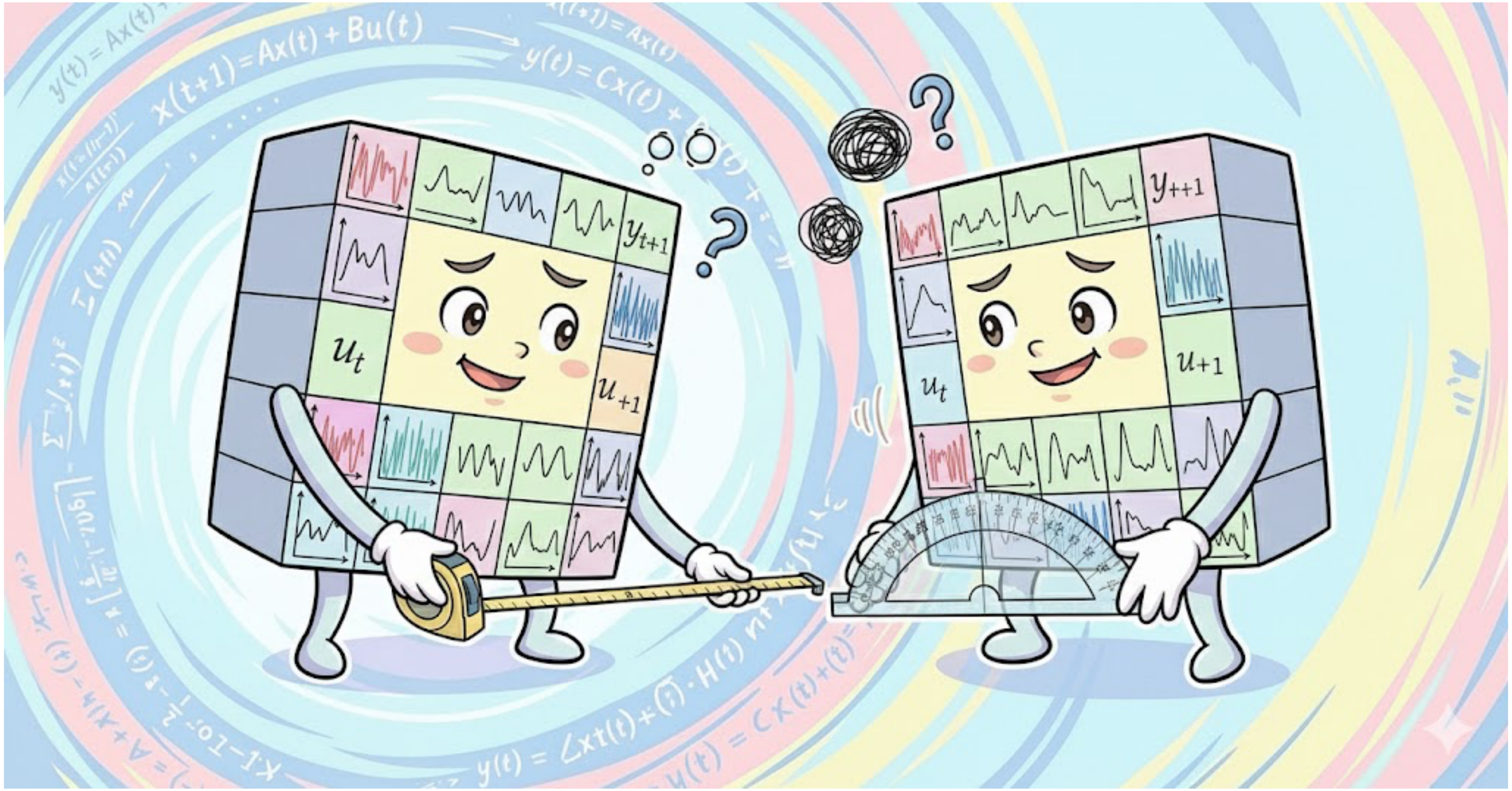




Why are data-driven approaches replacing traditional models?

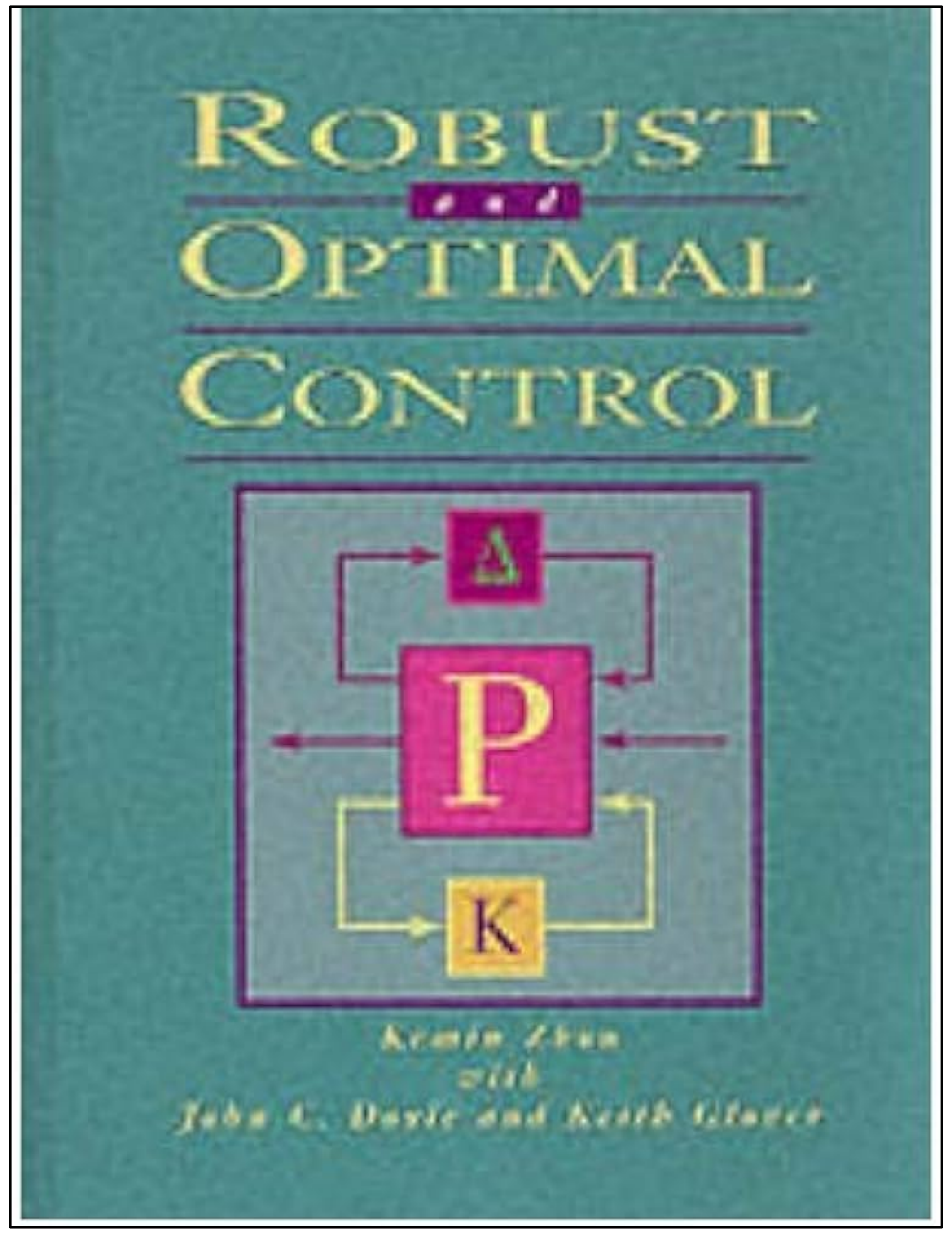
- **Complexity:** traditional models can be too complex to find or be useful (e.g., perception, human-in-the-loop)
- **Customization:** applications require custom models that are simple to build and require little expertise
- **Performance:** data-driven models often outperform traditional models when used for control

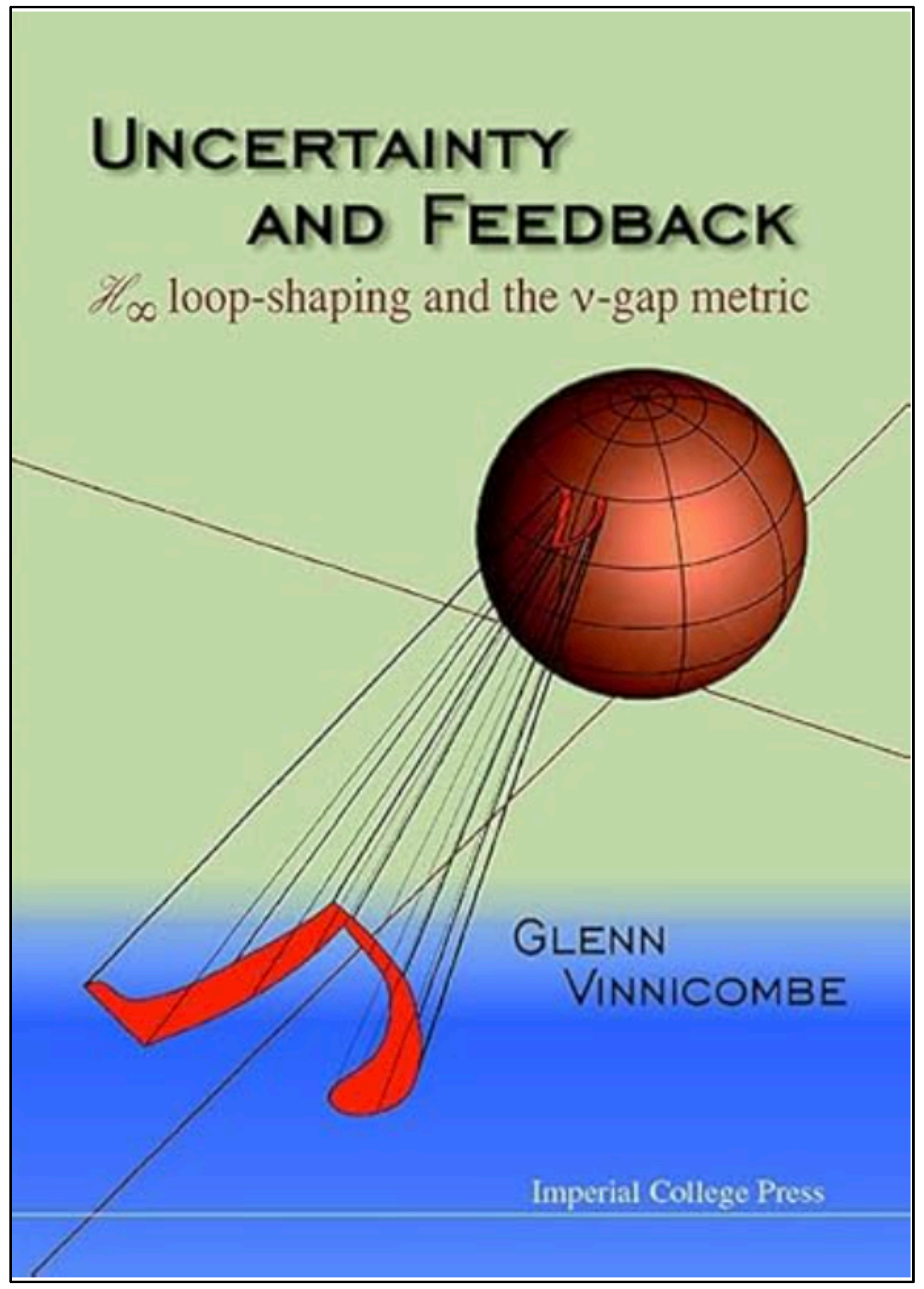
Today



How should we measure differences between data matrix representations of systems?

Long history of metrics





UNCERTAINTY IN UNSTABLE SYSTEMS: THE GAP METRIC

G. Zames and A. El-Sakkary

Distances and Riemannian metrics for multivariate spectral densities

Xianhua Jiang, Lipeng Ning, Tryphon T. Georgiou

Subspace angles between ARMA models

Katrien De Cock * , Bart De Moor

DISTANCE BETWEEN BEHAVIORS AND RATIONAL REPRESENTATIONS*

H. L. TRENTELMAN† AND S. V. GOTTIMUKKALA†

Behavioral uncertainty quantification for data-driven control

Alberto Padoan, Jeremy Coulson, Henk J. van Waarde, John Lygeros, and Florian Dörfler

Why are metrics important?

- **Robust control:** compare feedback (closed-loop) systems
- **Uncertainty quantification:** (non) parametric uncertainty, error propagation
- **Optimization:** system identification, system approximation, estimation

Thought experiment

Given two data matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 100 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Thought experiment

Given two data matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 100 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

How should I compare how “similar” these data matrices are?

Thought experiment

Given two data matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} 0 & 1 \\ 100 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

How should I compare how “similar” these data matrices are?

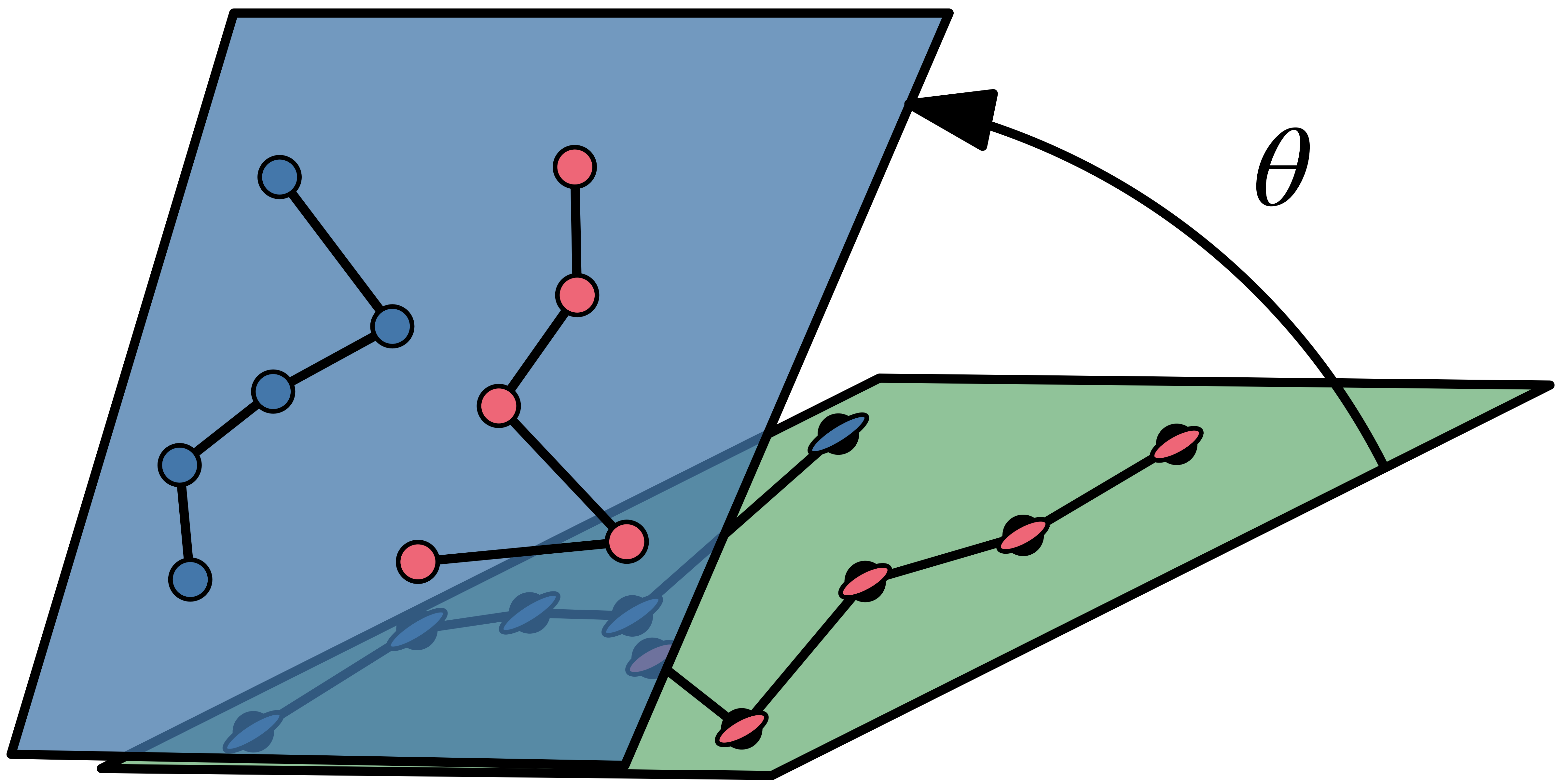
Will one perform “better” for control?

Take-home message

We require metrics tailored for data matrix system representations

Typical matrix metrics are not suitable for data matrices

Metrics should encode the geometry (complexity and misfit)



Menu

Behaviors & Subspaces

Metric on data matrices

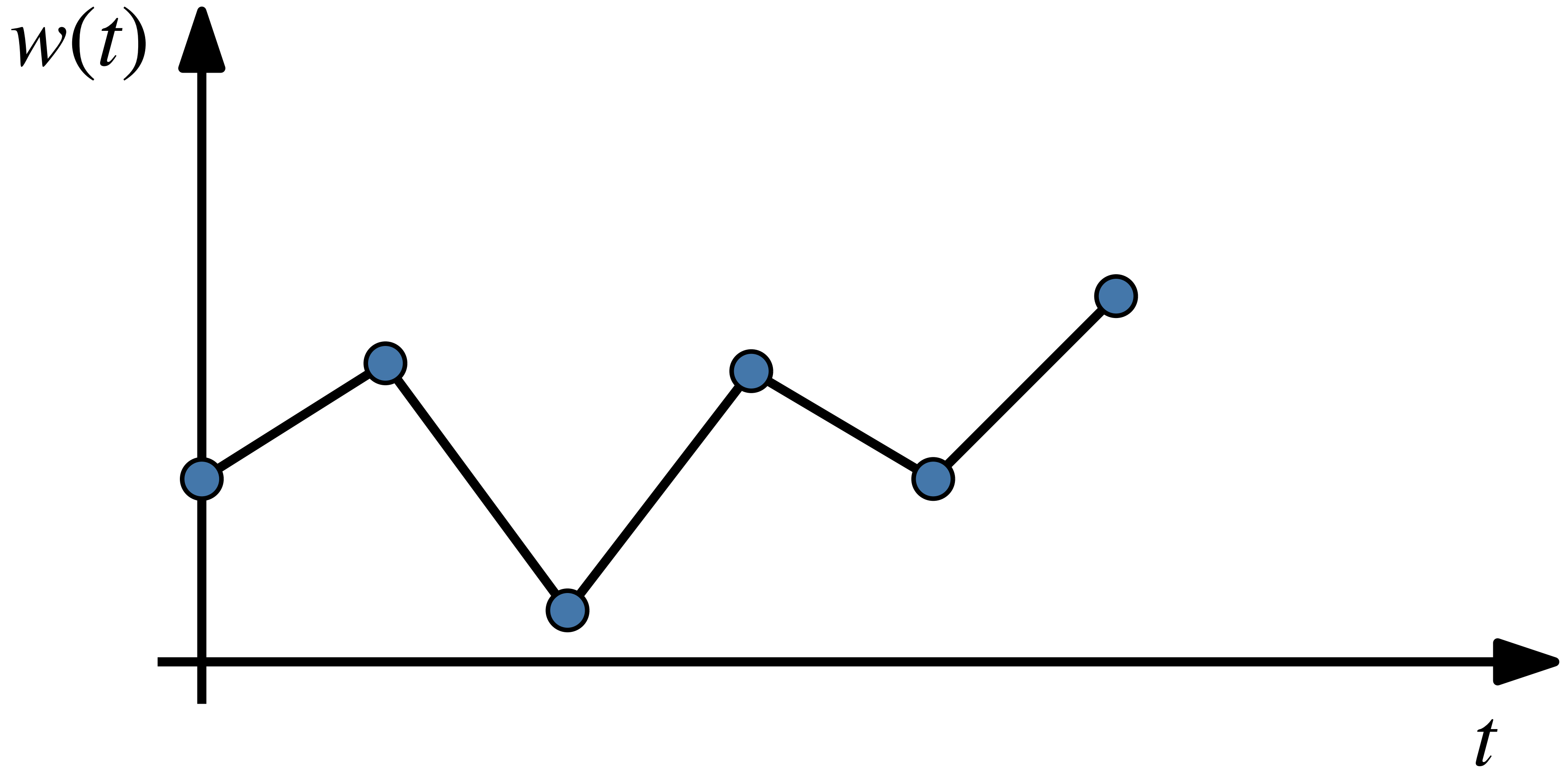
Example: anomaly detection

Behavioral systems

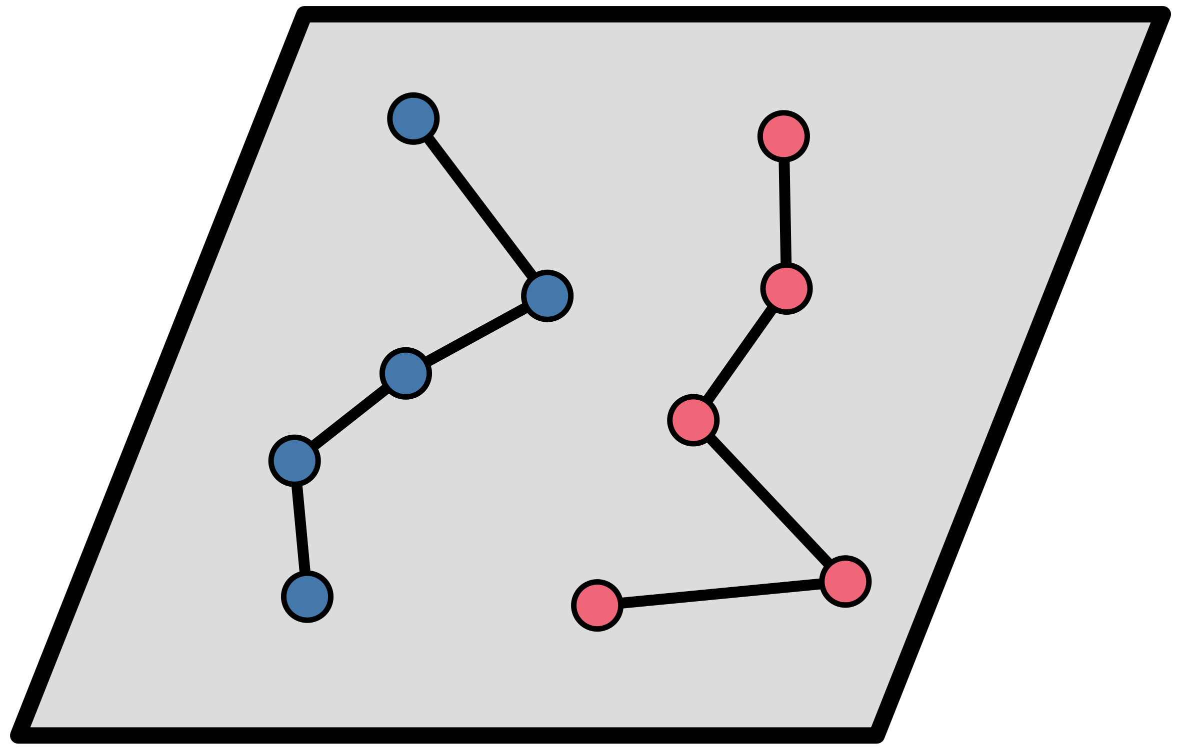
A discrete-time dynamical system is a 3-tuple $(\mathbb{N}, \mathbb{R}^q, \mathscr{B})$ with:

- time axis $\mathbb{N}$
- signal space $\mathbb{R}^q$
- behavior $\mathscr{B}$

J.C. Willems



$\in$



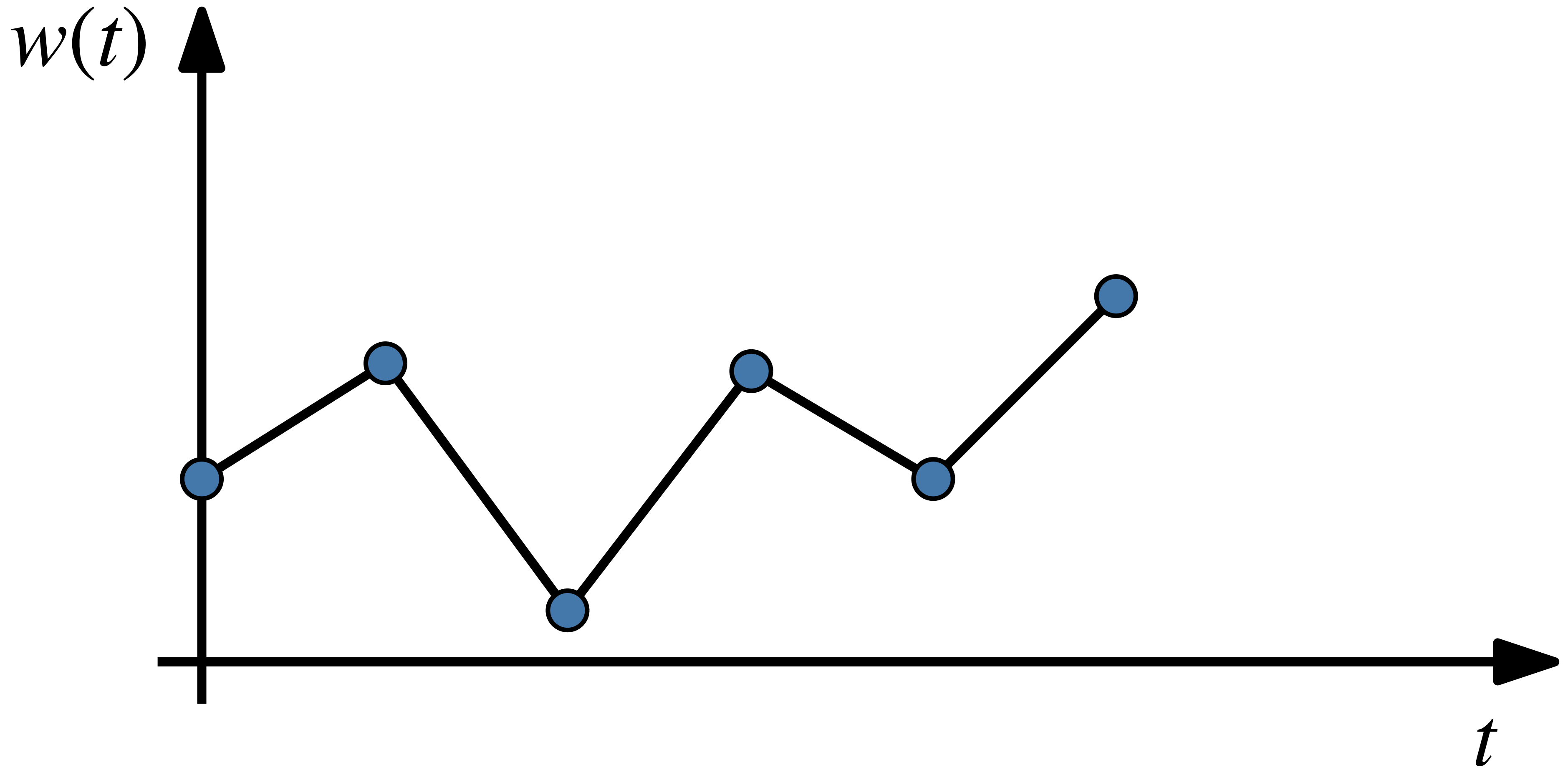
Behavioral systems

A discrete-time dynamical system is a 3-tuple $(\mathbb{N}, \mathbb{R}^q, \mathscr{B})$ with:

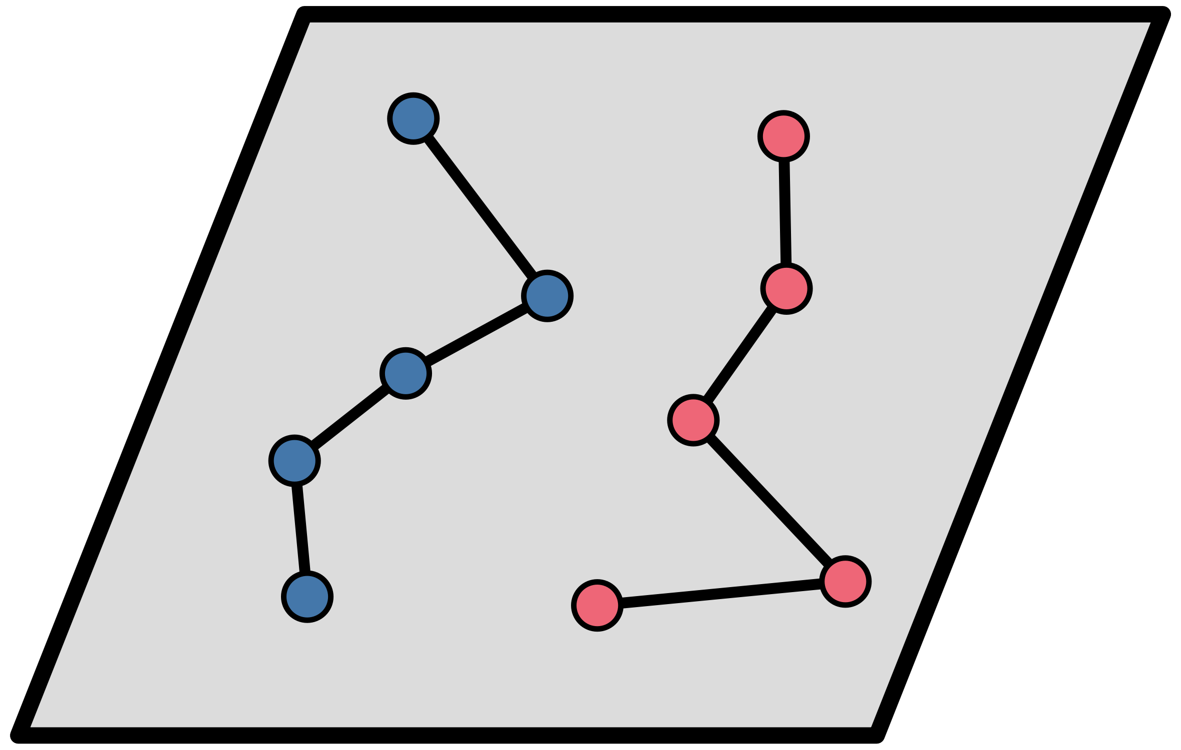
- time axis $\mathbb{N}$
- signal space $\mathbb{R}^q$
- behavior $\mathscr{B}$

Restricted behavior: $\mathscr{B}_{[0,L-1]}$ = set of L-length trajectories

J.C. Willems



$\in$



Behavioral systems

J.C. Willems

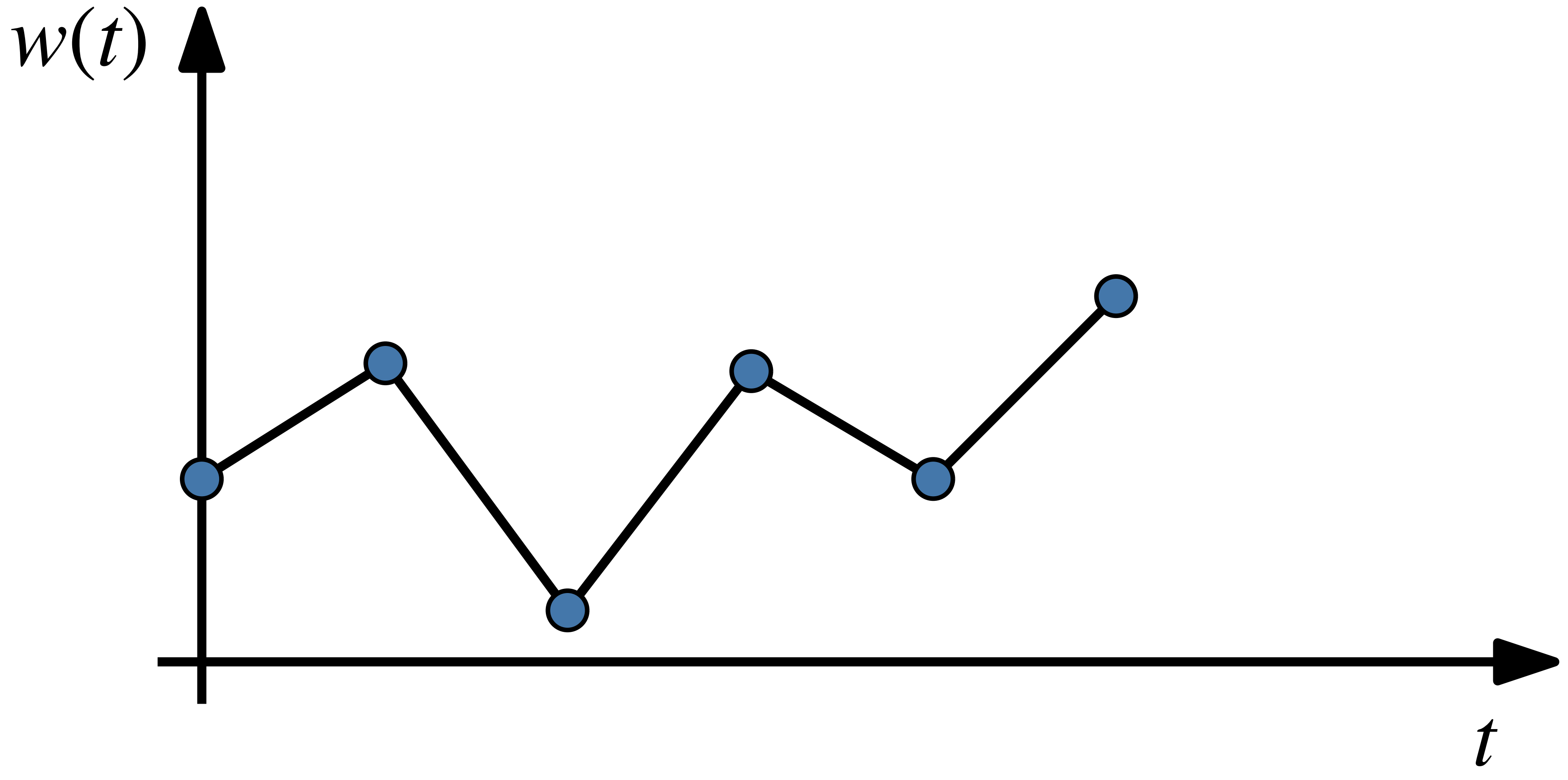
A discrete-time dynamical system is a 3-tuple $(\mathbb{N}, \mathbb{R}^q, \mathscr{B})$ with:

- time axis $\mathbb{N}$
- signal space $\mathbb{R}^q$
- behavior $\mathscr{B}$

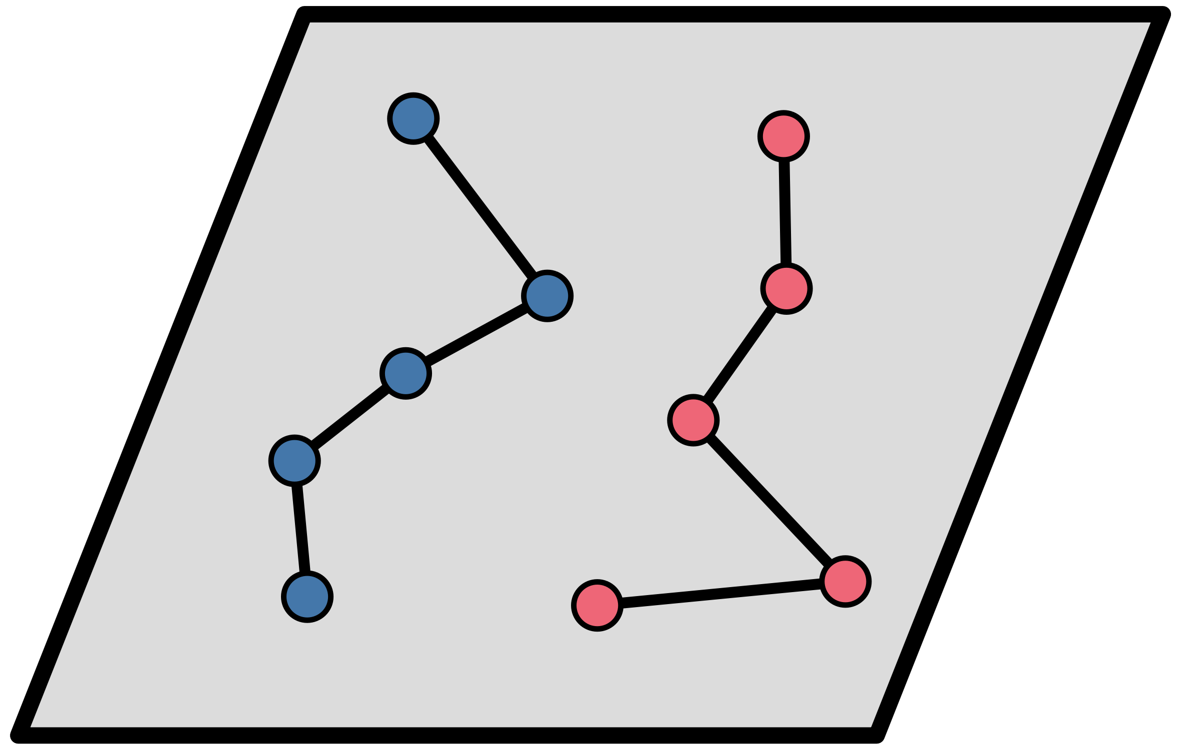
Restricted behavior: $\mathscr{B}_{[0,L-1]}$ = set of L-length trajectories

Linearity: $w_1, w_2 \in \mathscr{B} \Rightarrow \alpha w_1 + \beta w_2 \in \mathscr{B}$

Time-invariance: $w \in \mathscr{B} \Rightarrow \sigma w \in \mathscr{B},$ where $(\sigma w)(t) = w(t+1)$



$\in$



Behaviors as subspaces

Data matrix:

$$H = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-L) \\ w(1) & w(2) & \cdots & w(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ w(L-1) & w(L) & \cdots & w(T-1) \end{bmatrix}$$

Behaviors as subspaces

Data matrix:

$$H = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-L) \\ w(1) & w(2) & \cdots & w(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ w(L-1) & w(L) & \cdots & w(T-1) \end{bmatrix}$$

Fundamental lemma [Willems et al. '05]:

$$\mathscr{B}_{[0,L-1]} = \operatorname{im} H$$

if and only if

$$\operatorname{rank} H = mL + n.$$

Behaviors as subspaces

Data matrix:

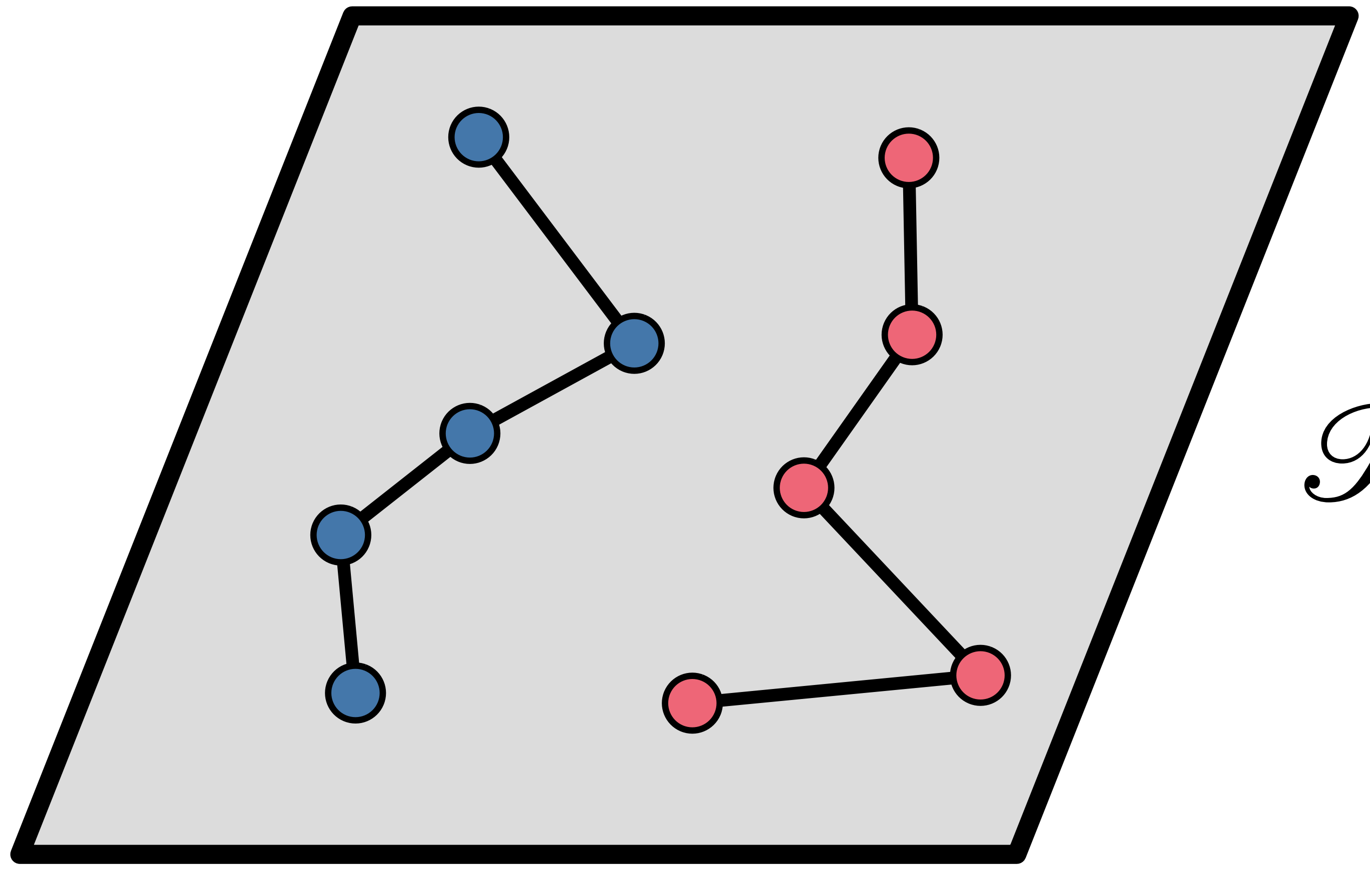
$$H = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-L) \\ w(1) & w(2) & \cdots & w(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ w(L-1) & w(L) & \cdots & w(T-1) \end{bmatrix}$$

Fundamental lemma [Willems et al. '05]:

$$\mathscr{B}_{[0,L-1]} = \operatorname{im} H$$

if and only if

$$\operatorname{rank} H = mL + n.$$



$$\mathscr{B}_{[0,L-1]} \in \underbrace{\mathrm{Gr}(mL+n, qL)}_{\text{set of subspaces}}$$

H. G. Grassmann

Behaviors as subspaces

Data matrix:

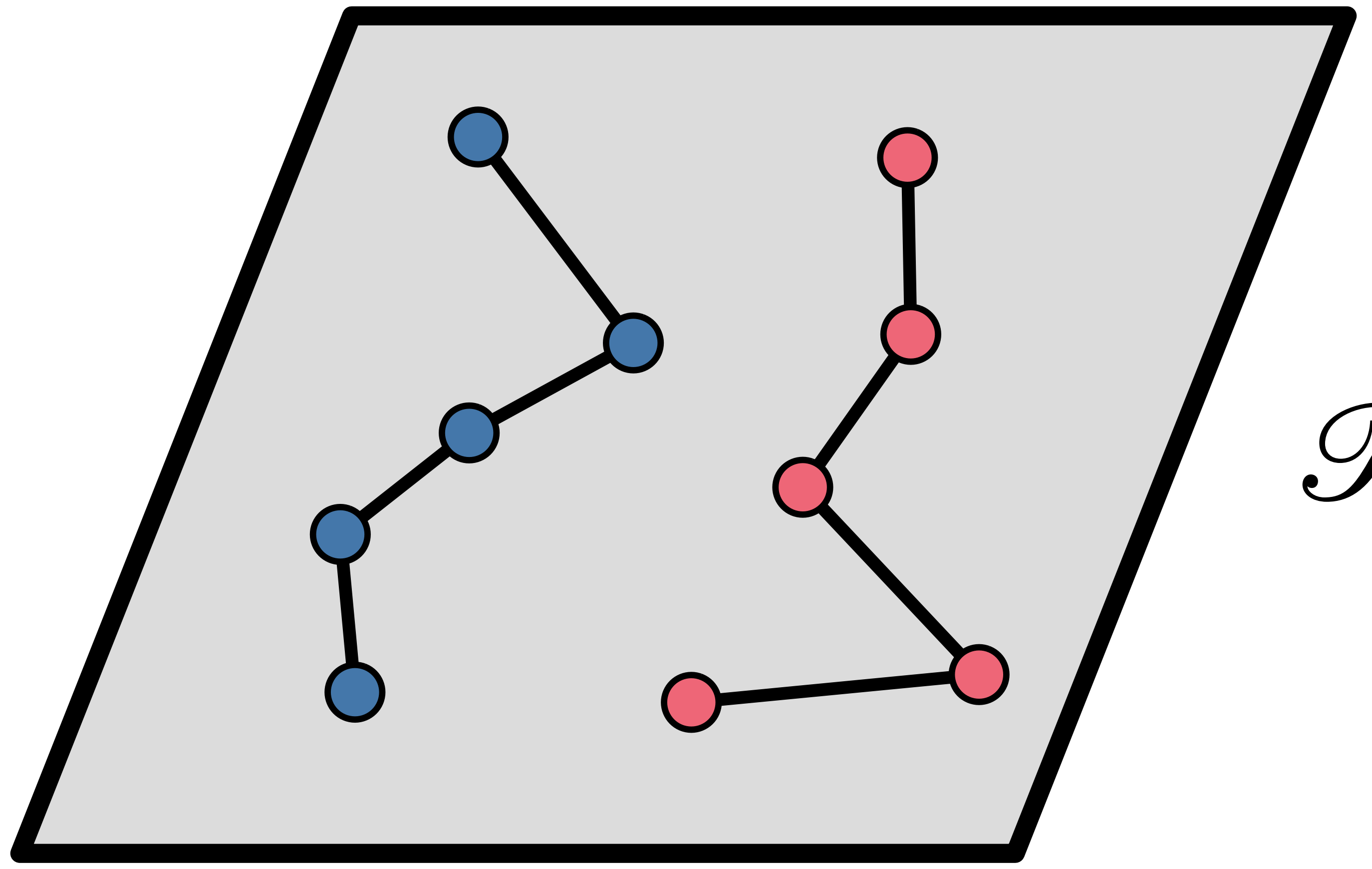
$$H = \begin{bmatrix} w(0) & w(1) & \cdots & w(T-L) \\ w(1) & w(2) & \cdots & w(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ w(L-1) & w(L) & \cdots & w(T-1) \end{bmatrix}$$

Fundamental lemma [Willems et al. '05]:

$$\mathscr{B}_{[0,L-1]} = \operatorname{im} H$$

if and only if

$$\operatorname{rank} H = mL + n.$$



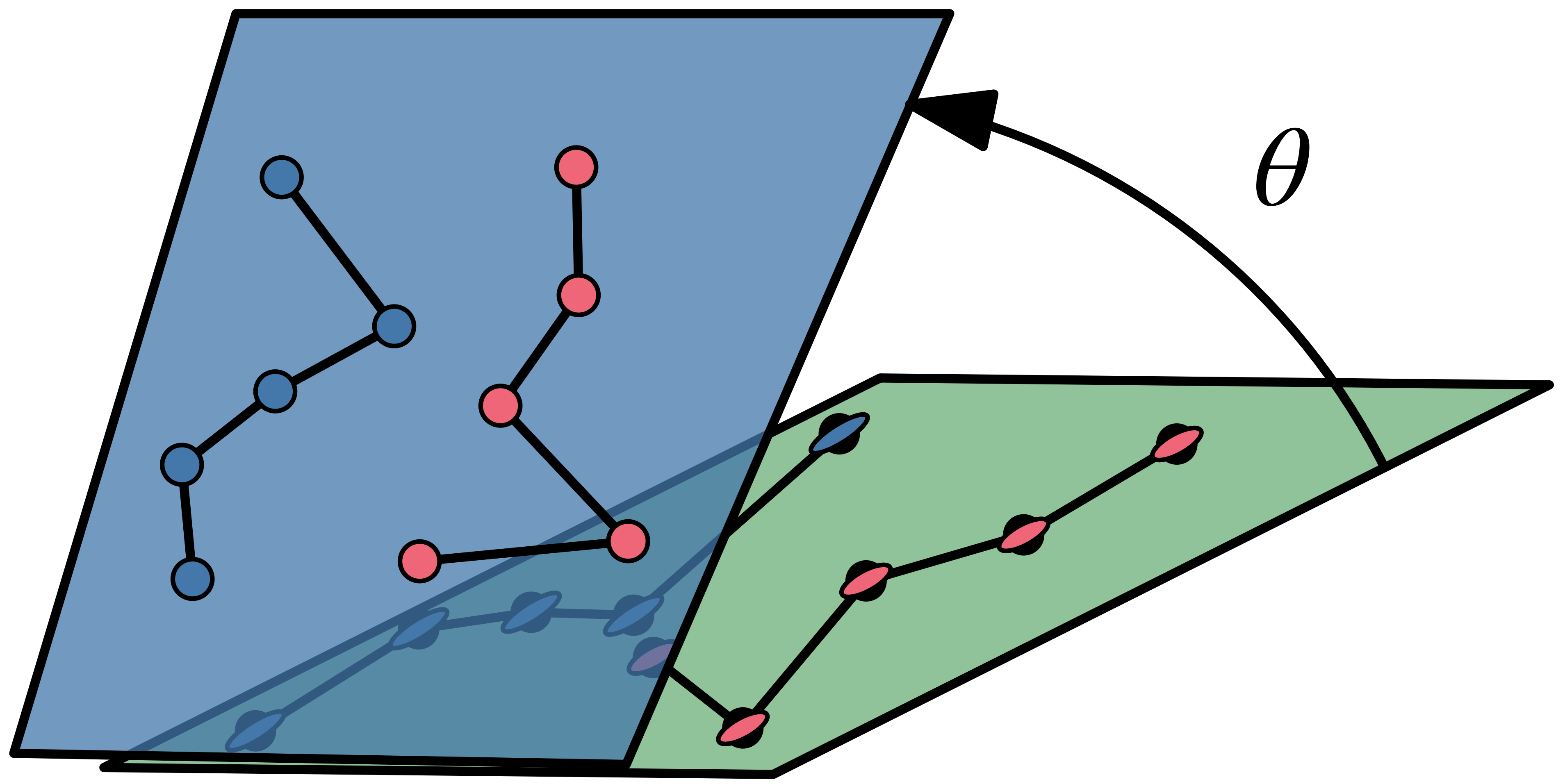
$$\mathscr{B}_{[0,L-1]} \in \underbrace{\operatorname{Gr}(mL+n, qL)}_{\text{set of subspaces}}$$

H. G. Grassmann

We can do systems theory, control, optimization with data & geometry!

Metrics on data matrices

$$d(\mathscr{B}_{[0,L-1]}, \mathscr{B}'_{[0,L-1]})^2 = \underbrace{|\dim \mathscr{B}_{[0,L-1]} - \dim \mathscr{B}'_{[0,L-1]}|}_{\text{complexity}} + \underbrace{\sum_{i=1}^{\min\{\dim \mathscr{B}_{[0,L-1]}, \mathscr{B}'_{[0,L-1]}\}} \sin^2(\theta_i)}_{\text{misfit}}$$

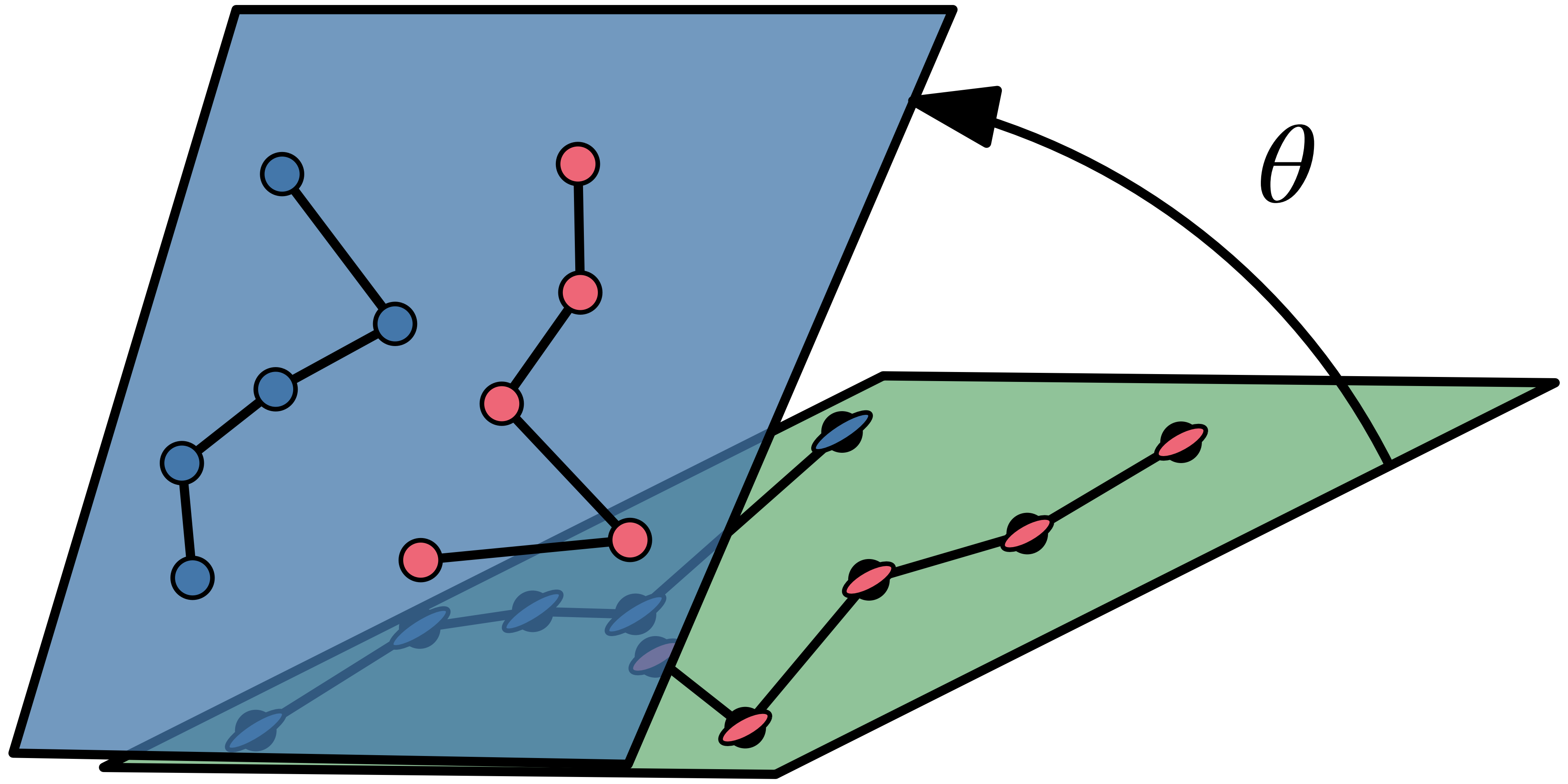


[Ye, Lim, '16]

Metrics on data matrices

$$d(\mathscr{B}_{[0,L-1]}, \mathscr{B}'_{[0,L-1]})^2 = \underbrace{|\dim \mathscr{B}_{[0,L-1]} - \dim \mathscr{B}'_{[0,L-1]}|}_{\text{complexity}} + \underbrace{\sum_{i=1}^{\min\{\dim \mathscr{B}_{[0,L-1]}, \mathscr{B}'_{[0,L-1]}\}} \sin^2(\theta_i)}_{\text{misfit}}$$

Complexity and misfit (principal angles) are readily computed through SVD



[Ye, Lim, ’16]

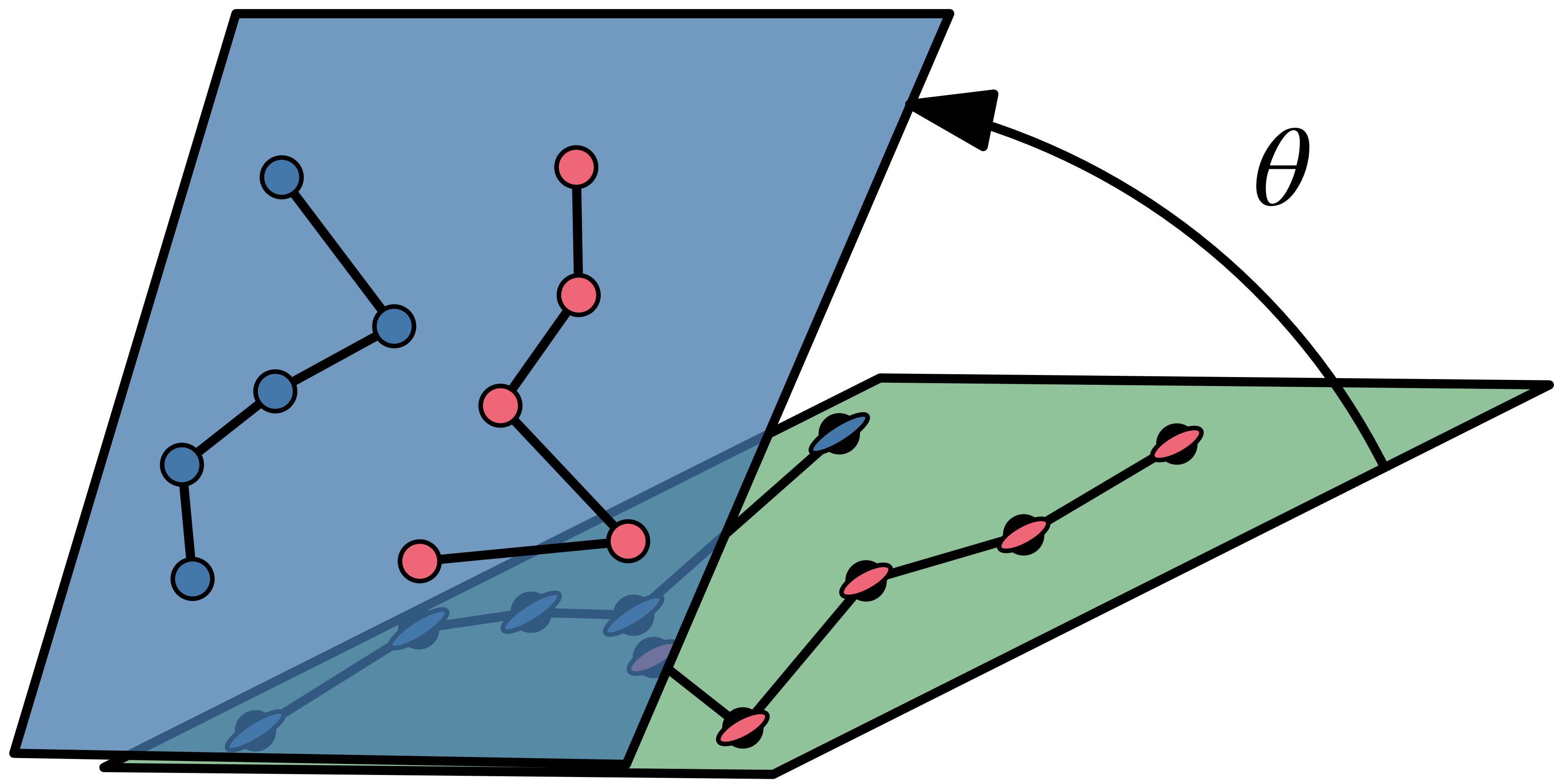
Metrics on data matrices

$$d(\mathscr{B}_{[0,L-1]}, \mathscr{B}'_{[0,L-1]})^2 = \underbrace{|\dim \mathscr{B}_{[0,L-1]} - \dim \mathscr{B}'_{[0,L-1]}|}_{\text{complexity}} + \underbrace{\sum_{i=1}^{\min\{\dim \mathscr{B}_{[0,L-1]}, \mathscr{B}'_{[0,L-1]}\}} \sin^2(\theta_i)}_{\text{misfit}}$$

Complexity and misfit (principal angles) are readily computed through SVD

Applications

- mode recognition [Padoan et al. ’22]
- fault detection [Padoan, Coulson ’25]
- modelling with accuracy complexity tradeoff [Padoan, Coulson ’25]

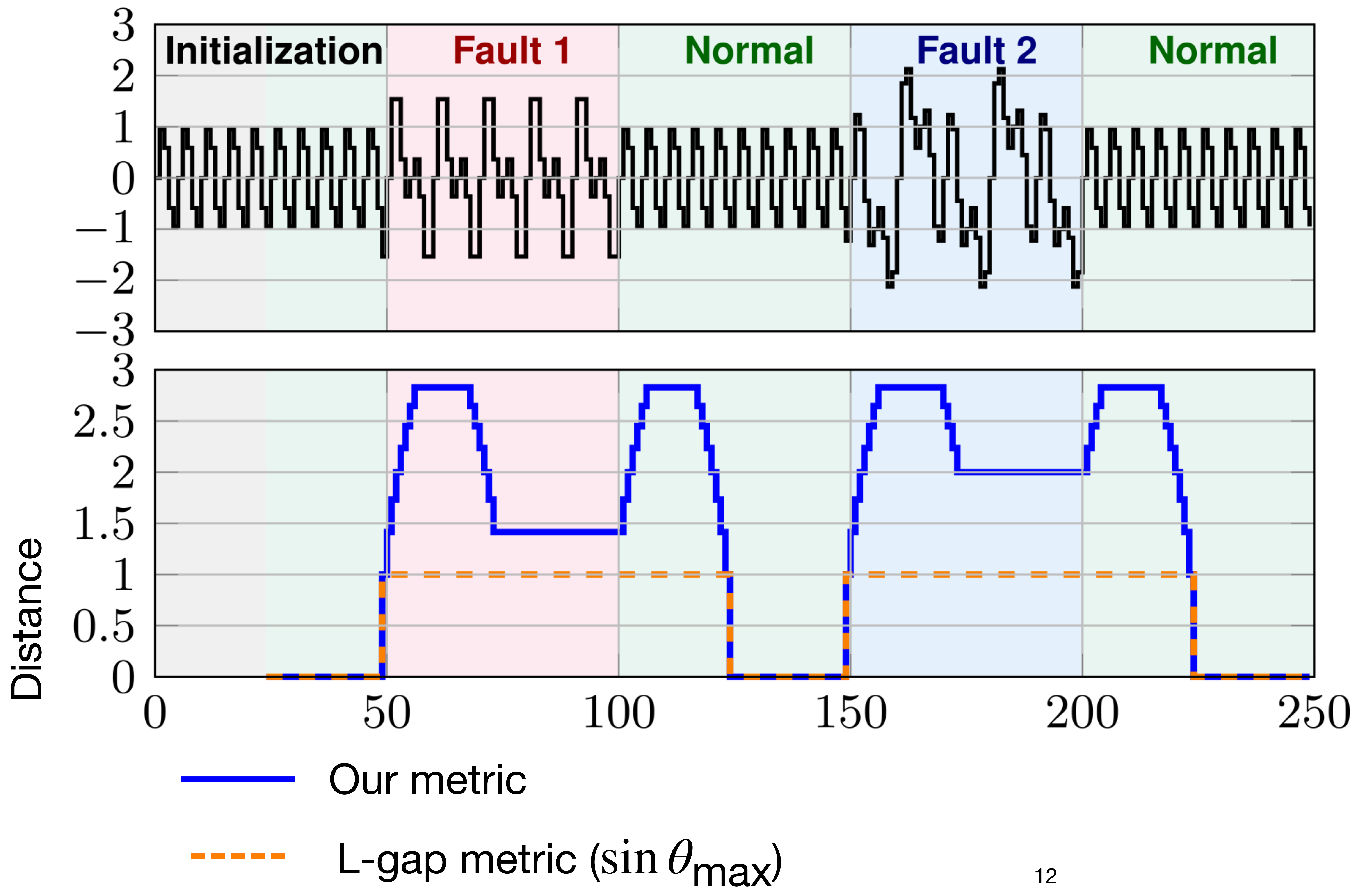


[Ye, Lim, ’16]

Numerical example: anomaly detection

Setup:
Nominal behavior: uniformly sampled sinusoid
Faulty behavior: additive harmonic perturbations

Data collection:
Behaviors defined by moving window Hankel matrices



Our metric shows finer resolution (complexity and misfit)

Summary

Metric for finite-horizon behaviors:
- Combines complexity and misfit
- Readily computed using data or models

Ongoing work
- Prediction/control error analysis
- Robust data-driven control
- Subspace tracking
- Approximate model reference control
- Model reduction
- Weighted distances

Thank you!