ETH zürich

Conference on Decision and Control CDC2025

AUTOMATIC CONTROL LABORATORY

Gaussian behaviors: representations and data-driven control

Andras Sasfi

December 9, 2025

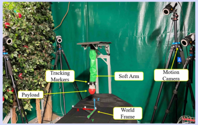
Acknowledgment

Florian Dörfler
@ ETH Zürich

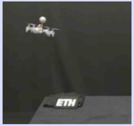
Alberto Padoan
@ UBC Vancouver

Ivan Markovsky
@ ICREA & CIMNE Barcelona

Direct data-driven control works very well across case studies



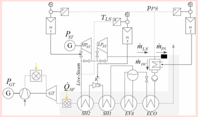
soft robotics (by H. Wang et al.)



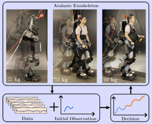
quad coptor fig-8 tracking

quadruped (by Fawcett, Afsari, Ames, & Hamed)

greenhouse automation (by Automatoes)

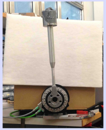


combined cycle power plant (by P Mahdavipour et. al)



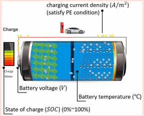
exosceleton (K. Li et. al)

robotic excavator

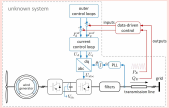


pendulum swing up

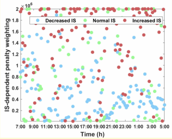
traffic coordination (by J. Wang et al.)



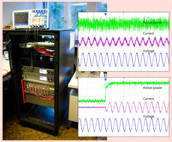
battery charging (by K. Chen et al.)



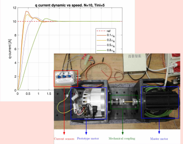
wind turbine control



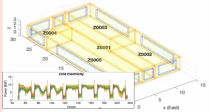
closed-loop diabetes control (X. Lu et al.)



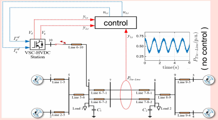
grid-connected converter



synchronous motor drive



energy hub & building automation



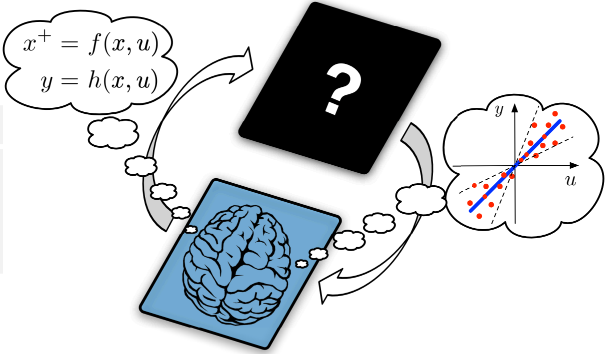
power system oscillation damping

...especially when first-principle models are hard or impossible to get

Direct data-driven control

Based on **behavioral systems theory**

"direct" data-driven control	
minimize	control cost (u, y)
subject to	trajectory (u, y) compatible with **non-parametric model** from raw data

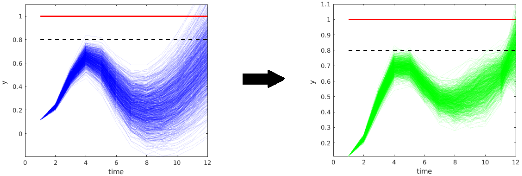


Deterministic framework $\rightarrow$ need **regularization!**

So far: **a posteriori justifications of robustness**

Bottom-up stochastic modeling framework

- simple $\rightarrow$ **tractable** control fromulations
- **robust** by design
- quantify and **shape uncertainty** of trajectories



Outline

1. Data-driven control based on behavioral systems theory

2. Bottom-up stochastic modeling framework: Gaussian behaviors

3. Predictive control with Gaussian behaviors
   - New insights on existing formulations
   - Distributionally robust control
   - Chance constraints and variance shaping

Outline

1. Data-driven control based on behavioral systems theory

2. Bottom-up stochastic modeling framework: Gaussian behaviors

3. Predictive control with Gaussian behaviors
   - New insights on existing formulations
   - Distributionally robust control
   - Chance constraints and variance shaping

Behavioral view on dynamical systems

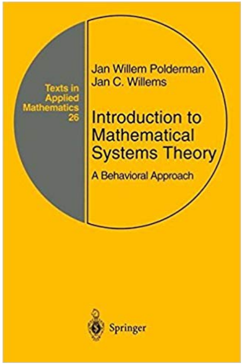
Definition: A discrete-time **dynamical system** is a 3-tuple $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathscr{B})$ where

(i) $\mathbb{Z}_{\geq 0}$ is the discrete-time axis,

(ii) $\mathbb{W}$ is the signal space, &

(iii) $\mathscr{B} \subseteq \mathbb{W}^{\mathbb{Z}_{\geq 0}}$ is the behavior.

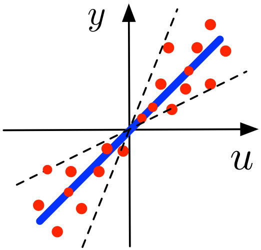
$\mathscr{B}$ is the set of all trajectories



Definition: The dynamical system $(\mathbb{Z}_{\geq 0}, \mathbb{W}, \mathscr{B})$ is

(i) **linear** if $\mathbb{W}$ is a vector space & $\mathscr{B}$ **is a subspace** of $\mathbb{W}^{\mathbb{Z}_{\geq 0}}$

(ii) & time-invariant if $w_t \in \mathscr{B} \implies w_{t+1} \in \mathscr{B}$



LTI system = shift-invariant subspace of trajectory space

$\longrightarrow$ abstract perspective suited for **data-driven control**

Fundamental Lemma

$$
\underbrace{\begin{array}{c}\text{set of all } T\text{-length trajectories } = \\ \left\{ (u, y) \in \mathbb{R}^{(m+p)T} \,:\, \exists x \in \mathbb{R}^{nT} \text{ s.t.} \right. \\ \left. x^+ = Ax + Bu \,,\; y = Cx + Du \right\}\end{array}}_{\text{parametric state-space model}}
\quad = \quad
\underbrace{\text{colspan} \begin{bmatrix} \begin{pmatrix} u^d_{1,1} \\ y^d_{1,1} \end{pmatrix} & \begin{pmatrix} u^d_{1,2} \\ y^d_{1,2} \end{pmatrix} & \begin{pmatrix} u^d_{1,3} \\ y^d_{1,3} \end{pmatrix} & \cdots \\ \begin{pmatrix} u^d_{2,1} \\ y^d_{2,1} \end{pmatrix} & \begin{pmatrix} u^d_{2,2} \\ y^d_{2,2} \end{pmatrix} & \begin{pmatrix} u^d_{2,3} \\ y^d_{2,3} \end{pmatrix} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \begin{pmatrix} u^d_{T,1} \\ y^d_{T,1} \end{pmatrix} & \begin{pmatrix} u^d_{T,2} \\ y^d_{T,2} \end{pmatrix} & \begin{pmatrix} u^d_{T,3} \\ y^d_{T,3} \end{pmatrix} & \cdots \end{bmatrix}}_{\text{raw data (every column is an experiment)}}
$$

all trajectories constructible from finitely many previous trajectories = **basis for $\mathscr{B}$**

Data-enabled Predictive Control (DeePC) $\leftrightarrow$ Model Predictive Control

$$
\begin{aligned}
\min_{g,u,y} \quad & \sum_{k=1}^{T_{\text{future}}} \|y_k - y_{\text{ref}}\|_Q^2 + \|u_k\|_R^2 \\
\text{s.t.} \quad & \mathcal{H}\begin{pmatrix} u^{\text{d}} \\ y^{\text{d}} \end{pmatrix} \cdot g = \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ \hline u \\ y \end{bmatrix} \\
& u_k \in \mathcal{U}, \quad \forall k \in \{1, \dots, T_{\text{future}}\} \\
& y_k \in \mathcal{Y}, \quad \forall k \in \{1, \dots, T_{\text{future}}\}
\end{aligned}
$$

non-parametric model for prediction and estimation

$$
\begin{aligned}
\min_{x,u,y} \quad & \sum_{k=1}^{T_{\text{future}}} \|y_k - y_{\text{ref}}\|_Q^2 + \|u_k\|_R^2 \\
\text{s.t.} \quad & x_{k+1} = Ax_k + Bu_k \\
& y_k = Cx_k + Du_k \\
& x_0 = x(t) \\
& u_k \in \mathcal{U}, \quad \forall k \in \{1, \dots, T_{\text{future}}\} \\
& y_k \in \mathcal{Y}, \quad \forall k \in \{1, \dots, T_{\text{future}}\}
\end{aligned}
$$

- real-time measurements $(u_{\text{ini}}, y_{\text{ini}})$ for estimation — **updated online**
- trajectory matrix $\mathcal{H}\begin{pmatrix} u^{\text{d}} \\ y^{\text{d}} \end{pmatrix}$ from past experimental data — **collected offline**

Equivalent to MPC in **deterministic** LTI case ... **but does not work in case of noise!**

Regularization makes it work

$$\min_{g,u,y} \sum_{k=1}^{T_{\text{future}}} \|y_k - r_k\|_Q^2 + \|u_k\|_R^2 + \lambda_g h(g) + \lambda_\sigma \|\sigma\|$$

$$\text{subject to} \quad \mathcal{H}\begin{pmatrix} u^{\mathsf{d}} \\ y^{\mathsf{d}} \end{pmatrix} \cdot g = \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ \hline u \\ y \end{bmatrix} + \sigma$$

$$u_k \in \mathcal{U}, \quad \forall k \in \{1, \ldots, T_{\text{future}}\}$$

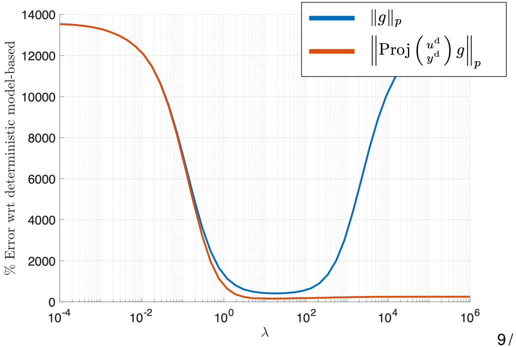
$$y_k \in \mathcal{Y}, \quad \forall k \in \{1, \ldots, T_{\text{future}}\}$$

regularization $\leftrightarrow$ **robustness**?? – hot research topic

behavioral model is **deterministic**
$\rightarrow$ **a posteriori justifications**

noisy data matrix $\mathcal{H}\begin{pmatrix} u^{\mathsf{d}} \\ y^{\mathsf{d}} \end{pmatrix}$ is usually full rank
$\rightarrow$ any $\binom{u}{y}$ is feasible

solution based on heuristics: add regularizer $h(g)$
$\rightarrow$ tuning λ_g is non-trivial
typical choices: 1-norm, 2-norm, projected 2-norm

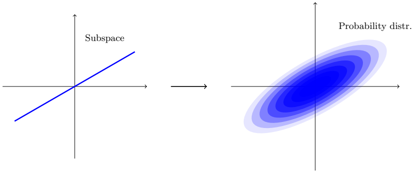


Outline

1. Data-driven control based on behavioral systems theory
2. Bottom-up stochastic modeling framework: Gaussian behaviors
3. Predictive control with Gaussian behaviors
   - New insights on existing formulations
   - Distributionally robust control
   - Chance constraints and variance shaping

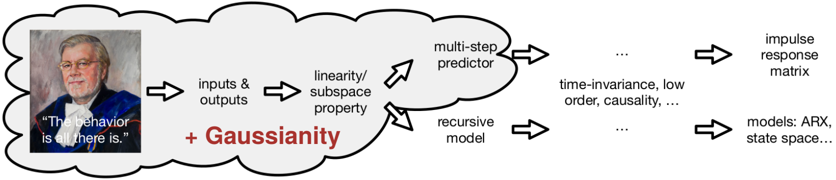
Towards a stochastic behavioral systems theory

Behavioral systems theory: trajectories are in a set $\rightarrow$ **deterministic approach**



Analogously **stochastic systems**: trajectories follow a probability distribution

How much structure do we want to impose?



General stochastic behaviors[1], polynomial chaos expansions[2] $\rightarrow$ **limited applicability**

Simple and **pragmatic** framework leading to **tractable control** methods

1) J. C. Willems, "Open stochastic systems," IEEE TAC, 2012
2) T. Faulwasser *et al.*, "Behavioral theory for stochastic systems? A data-driven journey from Willems to Wiener and back again," ARC, 2023

Gaussan behaviors

Gaussian behavior: trajectories distributed as a Gaussian Process

$$\begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{bmatrix} \sim \mathcal{N}(\mu_w, \Sigma_w)$$

Consistent with **state-space equations** with additive Gaussian noise

$$x_{k+1} = Ax_k + Bu_k + \xi_k$$
$$y_k = = Cx_k + Du_k + \eta_k$$
$$\Downarrow$$

μ_w, Σ_w **have structure!**

Deterministic LTI behaviors: special case with **singular covariance** matrix

$$\operatorname{im}(\Sigma_w) = \mathscr{B} \implies \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{bmatrix} \in \mathscr{B} \text{ with probability 1}$$

Special case of **stochastic behaviors** with the

- Borel σ-algebra
- Probability measure given by Gaussian process

Prediction

Predict y given $u_{\text{ini}}, y_{\text{ini}}$ and u by **conditioning!**

Conditioning **removes uncertainty** from $u_{\text{ini}}, y_{\text{ini}}$ and u

$$\mu_w = \begin{bmatrix} \mu_{\text{given}} \\ \mu_y \end{bmatrix}, \qquad \Sigma_w = \begin{bmatrix} \Sigma_{\text{given}} & \Sigma_{y,\text{given}}^\top \\ \Sigma_{y,\text{given}} & \Sigma_y \end{bmatrix}.$$

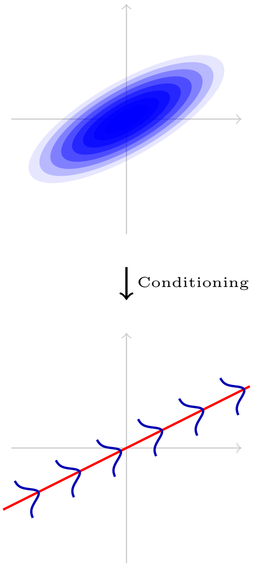
Conditional distribution is Gaussian with mean and variance

$$\mu = \mu_y + \Sigma_{y,\text{given}} \Sigma_{\text{given}}^\dagger \left(\begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \end{bmatrix} - \mu_{\text{given}} \right), \quad \textbf{(nominal prediction)}$$

$$\Sigma = \Sigma_y - \Sigma_{y,\text{given}} \Sigma_{\text{given}}^\dagger \Sigma_{y,\text{given}}^\top \quad \textbf{(uncertainty quantification)}$$

$$\begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} I \\ \Sigma_{y,\text{given}} \Sigma_{\text{given}}^\dagger \end{bmatrix}}_{\text{subspace}} \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ e \end{bmatrix}}_{\sim \mathcal{N}(\cdot, \Sigma)}$$

Partitions trajectory into deterministic and stochastic parts



Identification

Offline data matrix

$$\mathcal{H}\begin{pmatrix} u^{\mathsf{d}} \\ y^{\mathsf{d}} \end{pmatrix} = \begin{bmatrix} \mathcal{H}_{\text{ini}} \\ \mathcal{H}_u \\ \mathcal{H}_y \end{bmatrix} = \begin{bmatrix} u_{\text{ini}}^1 & u_{\text{ini}}^2 & \dots & u_{\text{ini}}^D \\ y_{\text{ini}}^1 & y_{\text{ini}}^2 & \dots & y_{\text{ini}}^D \\ u^1 & u^2 & \dots & u^D \\ y^1 & y^2 & \dots & y^D \end{bmatrix}$$

Each column is a sample from the distribution $\mathcal{N}(\mu_w, \Sigma_w)$

Estimate Σ_w by sample covariance

$$\Sigma_w \approx \frac{1}{D}\mathcal{H}\mathcal{H}^\top$$

i.i.d. zero mean samples from traj. space $\to$ ML estimate

Generalizes the non-parametric model of (deterministic) behaviors: $\operatorname{im}(\mathcal{H}\mathcal{H}^\top) = \operatorname{im}(\mathcal{H})$

Estimate predictive distribution

$$y | u_{\text{ini}}, y_{\text{ini}}, u \sim \mathcal{N}(\mu, \Sigma) \approx \mathcal{N}(\hat{\mu}, \hat{\Sigma})$$

$$\hat{\mu} = \mathcal{H}_y \begin{bmatrix} \mathcal{H}_{\text{ini}} \\ \mathcal{H}_u \end{bmatrix}^\dagger \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \\ u \end{bmatrix}$$

$$\hat{\Sigma} = \frac{1}{D}\mathcal{H}_y \left(I - \underbrace{\begin{bmatrix} \mathcal{H}_{\text{ini}} \\ \mathcal{H}_u \end{bmatrix}^\dagger \begin{bmatrix} \mathcal{H}_{\text{ini}} \\ \mathcal{H}_u \end{bmatrix}}_{\text{projection to row span}} \right) \mathcal{H}_y^\top$$

$\hat{\mu}$ is a function of u!

$\to$ subspace predictor[1] from subspace SySID

1) W. Favoreel *et al.*, “SPC: Subspace predictive control,” IFAC, 1999.

Outline

1. Data-driven control based on behavioral systems theory
2. Bottom-up stochastic modeling framework: Gaussian behaviors
3. Predictive control with Gaussian behaviors
   - New insights on existing formulations
   - Distributionally robust control
   - Chance constraints and variance shaping

Predictive control with Gaussian behaviors

Est. predictor $y|u_{\text{ini}}, y_{\text{ini}}, u \sim \mathcal{N}(\hat{\mu}, \hat{\Sigma})$ is **uncertain**

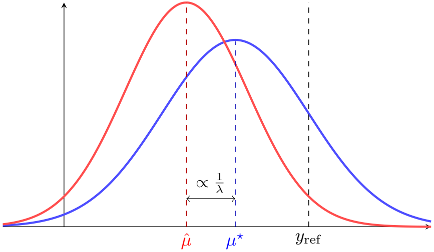
Control cost: $J(u, y) = \|y - y_{\text{ref}}\|_Q^2 + \|u\|_R^2$

Certainty-equivalence control = SPC

$$\min_{u \in \mathcal{U}} \quad \mathbb{E}_{\mathcal{N}(\hat{\mu}, \hat{\Sigma})}[J(u, y)] = J(u, \hat{\mu}) + \text{tr}(Q\hat{\Sigma})$$

Account for uncertainty in $\hat{\mu} \to$ allow for small deviation (Σ is fixed)

$$\begin{aligned} \min_{u \in \mathcal{U}} \min_{\mu} \quad & \mathbb{E}_{\mathcal{N}(\mu, \Sigma)}[J(u, y)] \\ \text{s.t.} \quad & \underbrace{D_{KL}\left(\mathcal{N}(\mu, \Sigma) \,\middle\|\, \mathcal{N}\left(\hat{\mu}, \hat{\Sigma}\right)\right)}_{\text{KL divergence}} \leq \epsilon \end{aligned}$$



Lift constraint to cost

Distributionally optimistic problem

$$\min_{u \in \mathcal{U}, \mu} \mathbb{E}_{\mathcal{N}(\mu, \Sigma)}[J(u, y)] + \lambda \cdot D_{KL}\left(\mathcal{N}(\mu, \Sigma) \,\middle\|\, \mathcal{N}\left(\hat{\mu}, \hat{\Sigma}\right)\right)$$

Theorem

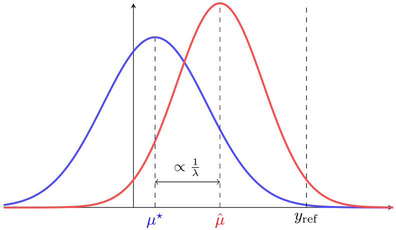
DeePC with $h(g) = \|(I - \Pi)g\|_2^2$ and without output constraints $\equiv$ DO problem with $\lambda = \lambda_g \frac{2}{D}$.

Distributionally robust control

Conservative formulation: **optimize worst-case**

Distributionally robust problem

$$\min_u \max_\mu \quad \mathbb{E}_{\mathcal{N}(\mu,\Sigma)}[J(u,y)]$$
$$\text{s.t.} \quad D_{KL}\left(\mathcal{N}(\mu,\Sigma) \,\middle\|\, \mathcal{N}\left(\hat{\mu},\hat{\Sigma}\right)\right) \leq \epsilon.$$



Theorem

For a large enough λ, the cost of DR problem is upper bounded by

$$\mathbb{E}_{\mathcal{N}(\mu^\star,\Sigma)}[J(u,y)] - \lambda\left(D_{KL}\left(\mathcal{N}(\mu^\star,\Sigma) \,\middle\|\, \mathcal{N}(\hat{\mu},\hat{\Sigma})\right) - \epsilon\right),$$

where

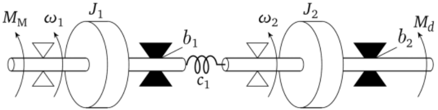
$$\mu^\star := (\lambda\hat{\Sigma}^{-1} - Q)^{-1}(\lambda\hat{\Sigma}^{-1}\hat{\mu} - Qy_{\text{ref}}).$$

Minimizing the upper bound leads to the **convex problem**

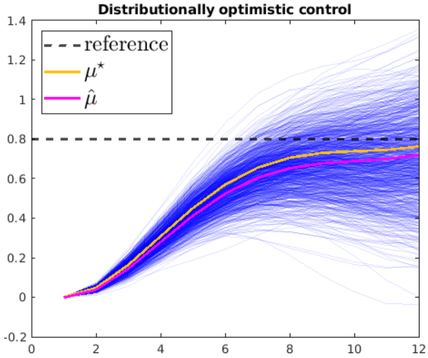
$$\min_u \quad \|\lambda\hat{\Sigma}^{-1}\hat{\mu} - Qy_{\text{ref}}\|^2_{(\lambda\hat{\Sigma}^{-1}-Q)^{-1}}$$
$$- \lambda\|\hat{\mu}\|^2_{\hat{\Sigma}^{-1}} + \|u\|^2_R$$

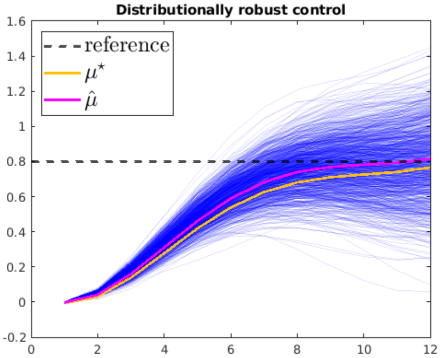
Case study

- Rotating discs with springs + dampers
- Input is a single torque
- Process and measurement noise
- Single trajectory of length 1000 $\rightarrow\ \hat{\mu},\ \hat{\Sigma}$



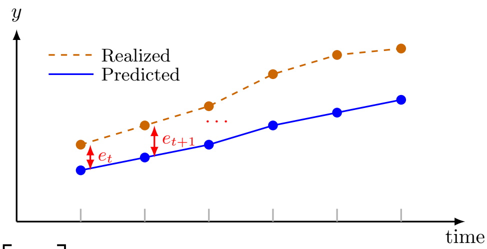
S. Kerz *et al.*, “Data-driven tube-based stochastic predictive control", IEEE OJCS, 2023





More interesting: chance constraints and optimization over Σ

Leveraging the predictor's variance



$$\begin{bmatrix} e_t \\ e_{t+1} \\ \vdots \end{bmatrix} \sim \mathcal{N}(0, \Sigma) \rightarrow e_t, e_{t+1}, \ldots$$ **correlated!**

Shape the trajectory's variance through **feedback on error**

$$u = u_{\text{ff}} + K_{\text{causal}} e$$

$$\rightarrow y = \underbrace{M_{\text{ini}} \begin{bmatrix} u_{\text{ini}} \\ y_{\text{ini}} \end{bmatrix} + M_u u_{\text{ff}}}_{\hat{\mu}} + (M_u K_{\text{causal}} + I) e$$

u and y are stochastic $\rightarrow$ add **chance constraints**

$\mathrm{Pr}(y_k \in \mathcal{Y}),\ \mathrm{Pr}(u_k \in \mathcal{U})$ with at least 99% probability

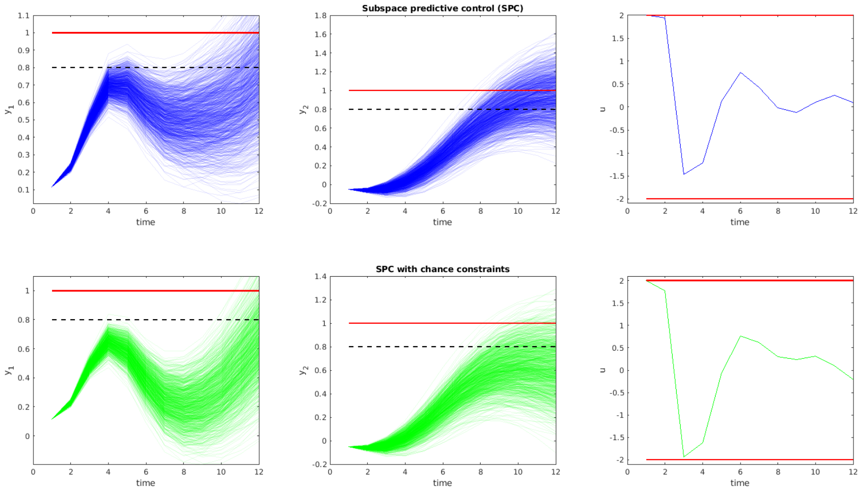
$$\min_{u_{\text{ff}}, K_{\text{causal}}} \quad \mathbb{E}[J(u,y)] = J(u_{\text{ff}}, \hat{\mu}) + \mathrm{tr}\Big(R \cdot \overbrace{K_{\text{causal}} \Sigma K_{\text{causal}}^\top}^{\Sigma_u}\Big) + \mathrm{tr}\Big(Q \cdot \underbrace{(M_u K_{\text{causal}} + I)\Sigma(M_u K_{\text{causal}} + I)^\top}_{\Sigma_y}\Big)$$

$$\text{s.t.} \quad \left.\begin{aligned} &\mathrm{Pr}(a_i^\top u \leq b_i) \geq \delta_i, \quad i = 1, 2, \ldots \\ &\mathrm{Pr}(c_i^\top y \leq d_i) \geq \gamma_i, \quad i = 1, 2, \ldots \end{aligned}\right\} \text{chance constraints}$$

For polytopic $\mathcal{Y}$ and $\mathcal{U}$ $\rightarrow$ convex problem

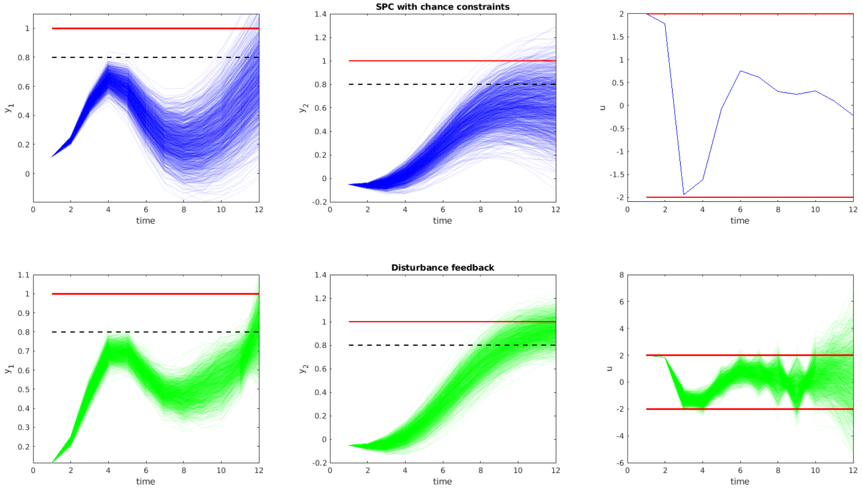
Case study – chance constraints

More cautious controller using the predictor's uncertainty estimate

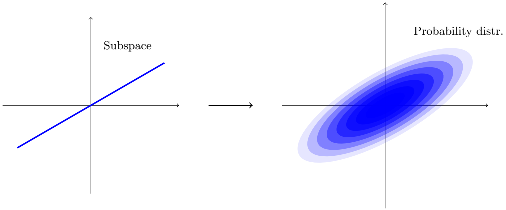


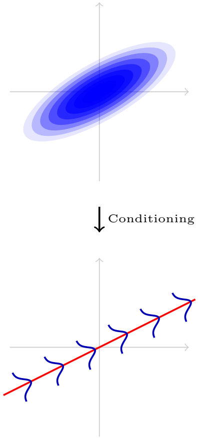
Case study – disturbance feedback

Feedback reduces the variance of open-loop output trajectories



Summary





- Proposed Gaussian behaviors for modeling stochastic systems
- Model deterministic and stochastic parts of systems
- Stochastic interpretation of existing methods (DeePC and SPC)
- Proposed novel distributionally robust formulation
- Leveraging the uncertainty quantification: chance constraints and disturbance feedback

Outlook

- Extend framework beyond linear systems and Gaussian noise
- Online adaptation of Gaussian behaviors