TUHH ICS

Data-driven stochastic control with formal guarantees
– Recent results and open problems

Timm Faulwasser (Institute of Control Systems, Hamburg University of Technology)
Joint work with: ***Ruchuan Ou*** *and* ***Guanru Pan*** *(ie3, TU Dortmund)*

Workshop on Data-driven Verification and Control with Provable Guarantees

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

ECC 2024 – Stockholm, June 25, 2024

Motivation – Stochastic Optimal Control

TUHH

$$\min_{\substack{U_k \in \mathrm{L}^2,\\ k \in \mathbb{I}_{[0,N-1]}}} \mathbb{E}\left[\sum_{k=0}^{N-1} {Y_k}^\top Q Y_k + {U_k}^\top R U_k\right]$$

subject to

$$X_{k+1} = AX_k + BU_k + EW_k, \quad X_0 = X_{ini}$$
$$Y_k = CX_k + DU_k + FW_k$$

are ubiquitous in applications

- Energy systems
- Autonomous driving
- Autonomous systems
- ...

Motivation – Stochastic Optimal Control

$$\min_{\substack{U_k \in \mathrm{L}^2, \\ k \in \mathbb{I}_{[0,N-1]}}} \max_{\substack{W_k \in \mathcal{W}, \\ k \in \mathbb{I}_{[0,N-1]}}} \mathbb{E}\left[\sum_{k=0}^{N-1} {Y_k}^\top Q Y_k + {U_k}^\top R U_k\right]$$

subject to

$$X_{k+1} = AX_k + BU_k + EW_k, \quad X_0 = X_{ini}$$
$$Y_k = CX_k + DU_k + FW_k$$

are ubiquitous in applications

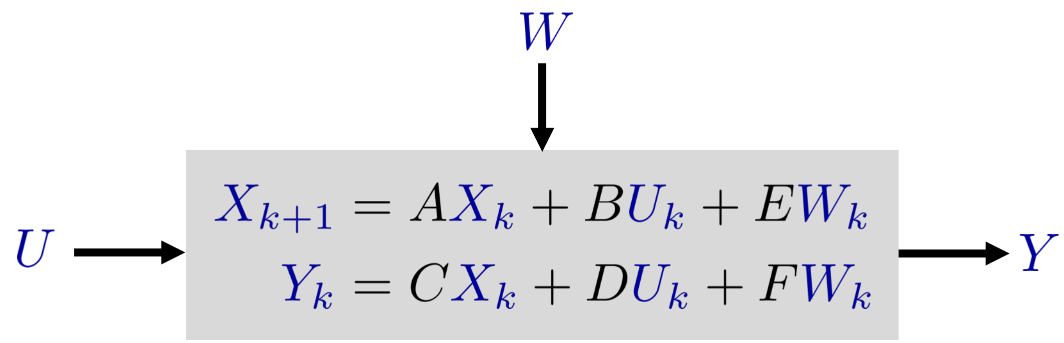
- Energy systems
- Autonomous driving
- Autonomous systems
- ...


Image: Forbes.com


Image: galooli.com

Here: data-driven distributionally robust optimal control

Problem Setting



- $i.i.d.\ W_k \sim \mu_W,\ \mu_W$ unkown and **non-Gaussian**
- System matrices are not known
- Measurements of past u_k, y_k, and w_k available
- L^2 random variables $W_k \in \mathrm{L}^2(\Omega, \mu; \mathbb{R}^{n_w}) \cdots$

- $W_k \quad \Rightarrow \quad$ stochastic disturbance
- Measurements of realizations $w_k := W_k(\omega)$ are available
- **Here:** Noise-free output and disturbance feedback
- Extension to output measurement noise is doable (but not today)

- Probability space $\mathrm{L}^2(\Omega, \mu; \mathbb{R})$
- Sample set Ω, probability measure μ
- Random variables
- Realizations of random variables

Why distributionally robust control?

Distributionally Robust Constraint Formulations

$$\min_u \ \ell(y,u) \quad \text{subject to} \quad y = f(u,w) \in \mathbb{Y}, w \in ???$$

Robustness →

Chance constraints	Distributionally robust constraints	Robust constraints
$\mathbb{P}(Y = f(U,W) \in \mathbb{Y}) \geq 1-\varepsilon$ $W \sim \mu_W$	$\mathbb{P}(Y = f(U,W) \in \mathbb{Y}) \geq 1-\varepsilon$ $W \sim \mu_W, \forall \mu_W \in \mathcal{A}$	$y = f(u,w) \in \mathbb{Y}$ $\forall w \in \mathbb{W} \subset \mathbb{R}^{n_w}$
Known distribution	Distributional ambiguity	Unknown distribution with compact support

Conservatism →

How to model distributional ambiguity?

B. P. G. Van Parys, D. Kuhn, P. J. Goulart, and M. Morari, 2016. Distributionally Robust Control of Constrained Stochastic Systems, *IEEE TAC*

P. M. Esfahani, D. Kuhn, Data-driven distributionally robust optimization using the Wasserstein metric: Performance guarantees and tractable reformulations, *Math. Program.*

Considered Distributionally Robust OCP

Min-max formulation

$$\min_{\substack{U_k \in \mathrm{L}^2, \\ k \in \mathbb{I}_{[0,N-1]}}} \max_{\substack{W_k \in \mathcal{W}, \\ k \in \mathbb{I}_{[0,N-1]}}} \mathbb{E}\left[\sum_{k=0}^{N-1} Y_k{}^\top Q Y_k + U_k{}^\top R U_k\right]$$

subject to

$$X_{k+1} = AX_k + BU_k + EW_k, \quad X_0 = X_{ini}$$
$$Y_k = CX_k + DU_k + FW_k$$
$$\mathbb{P}[Y_k \in \mathbb{Y}] \geq 1 - \varepsilon_y, \; \forall\, W_k \in \mathcal{W}$$
$$\mathbb{P}[U_k \in \mathbb{U}] \geq 1 - \varepsilon_u, \; \forall\, W_k \in \mathcal{W}$$

Questions

- How to model distributional ambiguity beyond Wasserstein sets?
- Uncertainty propagation with ambiguity through unknown dynamics?
- Tractable distributionally robust optimal control?

Uncertainty Propagation via Moments?

TUHH

W

$U \rightarrow$
$$
\begin{aligned}
X_{k+1} &= AX_k + BU_k + EW_k \\
Y_k &= CX_k + DU_k + FW_k
\end{aligned}
$$
$\rightarrow Y$

$$i.i.d.\ W_k \sim \mu_W$$

Propagation via mean and covariance

$$
\begin{aligned}
\mathbb{E}[X_{k+1}] =& A\mathbb{E}[X_k] + B\mathbb{E}[U_k] + E\mathbb{E}[W_k] \\
\Sigma[X_{k+1}] =& A\Sigma[X_k]A^\top + A\Sigma[X_k, U_k]B^\top \\
& + A\Sigma[U_k, X_k]B^\top + E\Sigma[W_k]E^\top
\end{aligned}
$$

- Moments are nonlinear functions of random variables
- Dynamics differ for different moments
- Data-driven analysis is challenging

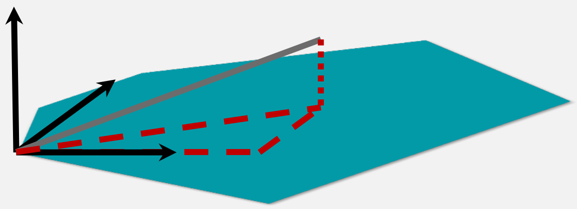
Alternatives?

Polynomial Chaos Expansion (PCE)

TUHH

- Hilbert space $\mathrm{L}^2(\Omega, \mu; \mathbb{R})$, $X : \Omega \to \mathbb{R}, \omega \mapsto X(\omega)$
- n_ξ-variate polynomials $\{\phi^j\}_{j=0}^\infty$ s.t. $\mathrm{span}\{\phi^j\}_{j=0}^\infty = \mathrm{L}^2(\Omega, \mu; \mathbb{R})$
- Orthogonality relation for all $i, j \in \mathbb{N}_0$

$$\langle \phi^i, \phi^j \rangle \doteq \int_\Omega \phi^i(\omega)\phi^j(\omega)\mathrm{d}\mu(\omega) = \mathbb{E}[\phi^i \phi^j] = \delta^{ij}\|\phi^j\|^2$$



$$W \in \mathrm{L}^2(\Omega, \mu; \mathbb{R}^{n_w})$$

PCE → $W = \sum_{j=0}^{\infty} \mathrm{w}^j \phi^j(\boldsymbol{\xi})$ PCE coefficient $\mathrm{w}^j = \dfrac{\mathbb{E}[W\phi^j]}{\mathbb{E}[\phi^j \phi^j]}$

Truncated PCE → $W \approx \sum_{j=0}^{L-1} \mathrm{w}^j \phi^j(\boldsymbol{\xi})$

Exact PCE → $W = \sum_{j \in \mathcal{J}} \mathrm{w}^j \phi^j(\boldsymbol{\xi})$ Finitely many basis functions

N. Wiener, 1938. “The homogeneous chaos,” American Journal of Mathematics.

Uncertainty propagation with exact PCEs?

T. Mühlpfordt, R. Findeisen, V. Hagenmeyer and T. Faulwasser, 2018. “Comments on Truncation Errors for Polynomial Chaos Expansions,” IEEE L-CSS.

Uncertainty Propagation via PCE

TUHH

Stochastic LTI system

$$X_{k+1} = AX_k + BU_k + EW_k$$
$$Y_k = CX_k + DU_k + FW_k$$

Exact PCEs

$U_k = \sum_{j=0}^{L-1} \mathsf{u}_k^j \phi^j(\boldsymbol{\xi}),\ Y_k = \cdots$

Orthogonal projection

LTI dynamics of PCE coefficients

$$\mathsf{x}_{k+1}^j = A\mathsf{x}_k^j + B\mathsf{u}_k^j + E\mathsf{w}_k^j$$
$$\mathsf{y}_k^j = C\mathsf{x}_k^j + D\mathsf{u}_k^j + F\mathsf{w}_k^j$$
$$j = 0, 1, \cdots, L-1$$

Uncertainty sampling

$u_k = U_k(\omega),\ y_k = Y_k(\omega),\ \cdots$

LTI dynamics of sampled trajectories

$$x_{k+1} = Ax_k + Bu_k + Ew_k$$
$$y_k = Cx_k + Du_k + Fw_k$$

Random variables, *realizations*, and *PCE coefficients* are subject to the **same** system dynamics

Data-driven uncertainty propagation?

G. Pan, R. Ou, and T. Faulwasser, 2023. On a stochastic fundamental lemma and its use for data-driven optimal control. *IEEE TAC*

T. Faulwasser et al. 2023. Behavioral theory for stochastic systems? A data-driven journey from Willems to Wiener and back again. *Annual Reviews in Control*

TUHH

Definition (Persistency of excitation):

The input data $\mathbf{u}_T^{\mathsf{d}} = \begin{bmatrix} u_1^\top & \cdots & u_T^\top \end{bmatrix}^\top$ is persistently of exciting of order L if the Hankel matrix

$$\mathscr{H}_L(\mathbf{u}_T^{\mathsf{d}}) = \begin{bmatrix} u_1 & \cdots & u_{T-L+1} \\ \vdots & \ddots & \vdots \\ u_L & \cdots & u_T \end{bmatrix}$$

is of full row rank.

LTI system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \quad x_0 = x^0 \\ y_k &= Cx_k + Du_k \end{aligned}$$

Fundamental lemma

Consider a controllable LTI system with minimal representation order n. If $\mathbf{u}_T^{\mathsf{d}}$ is persistently exciting of order $N+n$, then $\mathbf{u}_N, \mathbf{y}_N$ is an input-output trajectory pair if and only if there exists $g \in \mathbb{R}^{T-N+1}$ such that

$$\begin{bmatrix} \mathscr{H}_N(\mathbf{u}_T^{\mathsf{d}}) \\ \mathscr{H}_N(\mathbf{y}_T^{\mathsf{d}}) \end{bmatrix} g = \begin{bmatrix} \mathbf{u}_N \\ \mathbf{y}_N \end{bmatrix}.$$

J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, 2005. A note on persistency of excitation. *Systems & Control Letters*

Data-driven Uncertainty Propagation

TUHH

Stochastic LTI system

$$X_{k+1} = AX_k + BU_k + EW_k$$
$$Y_k = CX_k + DU_k + FW_k$$

Exact PCEs

$U_k = \sum_{j=0}^{L-1} \mathsf{u}_k^j \phi^j(\boldsymbol{\xi}),\ Y_k = \cdots$

Orthogonal projection

LTI dynamics of PCE coefficients

$$\mathsf{x}_{k+1}^j = A\mathsf{x}_k^j + B\mathsf{u}_k^j + E\mathsf{w}_k^j$$
$$\mathsf{y}_k^j = C\mathsf{x}_k^j + D\mathsf{u}_k^j + F\mathsf{w}_k^j$$
$$j = 0, 1, \cdots, L-1$$

Uncertainty sampling

$u_k = U_k(\omega),\ y_k = Y_k(\omega),\ \cdots$

LTI dynamics of sampled trajectories

$$x_{k+1} = Ax_k + Bu_k + Ew_k$$
$$y_k = Cx_k + Du_k + Fw_k$$

Persistently exciting $u^{\mathrm{d}},\ w^{\mathrm{d}}$

controllable $(A, [B\ E])$

Willems' fundamental lemma

Non-parametric model

$$\begin{bmatrix} \mathscr{H}_N(u^{\mathrm{d}}) \\ \mathscr{H}_N(w^{\mathrm{d}}) \\ \mathscr{H}_N(y^{\mathrm{d}}) \end{bmatrix} g = \begin{bmatrix} u_{[0,N-1]} \\ w_{[0,N-1]} \\ y_{[0,N-1]} \end{bmatrix}$$
$$g \in \mathbb{R}^{T-N+1}$$

Realizations paths described by recorded data

J.C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, 2005. A note on persistency of excitation, *Systems & Control Letters.*

Data-driven Uncertainty Propagation

TUHH

Stochastic LTI system

$$X_{k+1} = AX_k + BU_k + EW_k$$
$$Y_k = CX_k + DU_k + FW_k$$

Exact PCEs

$U_k = \sum_{j=0}^{L-1} \mathsf{u}_k^j \phi^j(\boldsymbol{\xi}),\ Y_k = \cdots$

Orthogonal projection

LTI dynamics of PCE coefficients

$$\mathsf{x}_{k+1}^j = A\mathsf{x}_k^j + B\mathsf{u}_k^j + E\mathsf{w}_k^j$$
$$\mathsf{y}_k^j = C\mathsf{x}_k^j + D\mathsf{u}_k^j + F\mathsf{w}_k^j$$
$$j = 0, 1, \cdots, L-1$$

Uncertainty sampling

$u_k = U_k(\omega),\ y_k = Y_k(\omega),\ \cdots$

LTI dynamics of sampled trajectories

$$x_{k+1} = Ax_k + Bu_k + Ew_k$$
$$y_k = Cx_k + Du_k + Fw_k$$

Measurements $u^{\mathrm{d}},\ w^{\mathrm{d}},\ y^{\mathrm{d}}$

Persistently exciting $u^{\mathrm{d}},\ w^{\mathrm{d}}$

controllable $(A, [B\ E])$

Data-driven uncertainty propagation

$$\begin{bmatrix} \mathscr{H}_N(u^{\mathrm{d}}) \\ \mathscr{H}_N(w^{\mathrm{d}}) \\ \mathscr{H}_N(y^{\mathrm{d}}) \end{bmatrix} \mathsf{g}^j = \begin{bmatrix} \mathsf{u}_{[0,N-1]}^j \\ \mathsf{w}_{[0,N-1]}^j \\ \mathsf{y}_{[0,N-1]}^j \end{bmatrix}$$

$$\mathsf{g}^j \in \mathbb{R}^{T-N+1}$$

Stochastic fundamental lemma? How to construct basis for exact PCEs?

G. Pan, R. Ou, and T. Faulwasser, 2023. On a stochastic fundamental lemma and its use for data-driven optimal control. *IEEE TAC*

T. Faulwasser et al. 2023. Behavioral theory for stochastic systems? A data-driven journey from Willems to Wiener and back again. *Annual Reviews in Control*

A Stochastic Fundamental Lemma?

TUHH

Hankel matrix description for PCE coefficients

$$\begin{bmatrix} \mathscr{H}_N(\mathbf{u}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{y}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{w}^{\mathrm{d}}) \end{bmatrix} \mathsf{g}^j = \begin{bmatrix} \mathsf{u}^j_{[0,N-1]} \\ \mathsf{y}^j_{[0,N-1]} \\ \mathsf{w}^j_{[0,N-1]} \end{bmatrix}, \quad j = 0, \dots, L-1$$

Multiply with scalar PCE basis ϕ^j

$$\begin{bmatrix} \mathscr{H}_N(\mathbf{u}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{y}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{w}^{\mathrm{d}}) \end{bmatrix} \mathsf{g}^j \phi^j = \begin{bmatrix} \mathsf{u}^j_{[0,N-1]} \\ \mathsf{y}^j_{[0,N-1]} \\ \mathsf{w}^j_{[0,N-1]} \end{bmatrix} \phi^j, \quad j = 0, \dots, L-1$$

Sum over $j = 0, \dots, L-1$

$$\sum_{j=0}^{L-1} \begin{bmatrix} \mathscr{H}_N(\mathbf{u}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{y}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{w}^{\mathrm{d}}) \end{bmatrix} \mathsf{g}^j \phi^j = \sum_{j=0}^{L-1} \begin{bmatrix} \mathsf{u}^j_{[0,N-1]} \\ \mathsf{y}^j_{[0,N-1]} \\ \mathsf{w}^j_{[0,N-1]} \end{bmatrix} \phi^j$$

$$\begin{bmatrix} \mathscr{H}_N(\mathbf{u}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{y}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{w}^{\mathrm{d}}) \end{bmatrix} G = \begin{bmatrix} \mathscr{H}_N(\mathbf{u}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{y}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{w}^{\mathrm{d}}) \end{bmatrix} \sum_{j=0}^{L-1} \mathsf{g}^j \phi^j = \sum_{j=0}^{L-1} \begin{bmatrix} \mathsf{u}^j_{[0,N-1]} \\ \mathsf{y}^j_{[0,N-1]} \\ \mathsf{w}^j_{[0,N-1]} \end{bmatrix} \phi^j = \begin{bmatrix} \mathbf{U}_{[0,N-1]} \\ \mathbf{Y}_{[0,N-1]} \\ \mathbf{W}_{[0,N-1]} \end{bmatrix}$$

A Stochastic Fundamental Lemma

Stochastic LTI system with $W_k \in \mathrm{L}^2(\Omega, \mu, \mathbb{R}^m), U_k \in \mathrm{L}^2(\Omega, \mu, \mathbb{R}^p)$

$$\begin{aligned} X_{k+1} &= AX_k + BU_k + EW_k, \qquad X_0 = X^0 \in \mathrm{L}^2(\Omega, \mu, \mathbb{R}^n) \qquad (1^*) \\ Y_k &= CX_k + DU_k + FW_k \end{aligned}$$

Stochastic fundamental lemma (input-output setting)

Consider (1*) with $(A, [B, E])$ controllable and with exact PCEs for W, X, U. Let the realizations $\mathbf{u}^{\mathrm{d}}, \mathbf{w}^{\mathrm{d}}$ be persistently exciting of order $N + n$.

The tuple $(\mathbf{U}_{[0,N-1]}, \mathbf{Y}_{[0,N-1]}, \mathbf{W}_{[0,N-1]})$ is an input-output-disturbance trajectory of (1*) if and only if there exists $G \in \mathrm{L}^2(\Omega, \mu; \mathbb{R}^{T-N+1})$ such that

$$\begin{bmatrix} \mathscr{H}_N(\mathbf{u}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{y}^{\mathrm{d}}) \\ \mathscr{H}_N(\mathbf{w}^{\mathrm{d}}) \end{bmatrix} G = \begin{bmatrix} \mathbf{U}_{[0,N-1]} \\ \mathbf{Y}_{[0,N-1]} \\ \mathbf{W}_{[0,N-1]} \end{bmatrix}.$$

Hankel matrices of measured realizations

Input-output trajectories of (1*)

. Pan, R. Ou, and T. Faulwasser, 2023. On a stochastic fundamental lemma and its use for data-driven optimal control. *IEEE TAC*

Data-driven Uncertainty Propagation

TUHH

Stochastic LTI system

$$X_{k+1} = AX_k + BU_k + EW_k$$
$$Y_k = CX_k + DU_k + FW_k$$

Exact PCEs

$U_k = \sum_{j=0}^{L-1} \mathsf{u}_k^j \phi^j(\boldsymbol{\xi}),\ Y_k = \cdots$

Orthogonal projection

LTI dynamics of PCE coefficients

$$\mathsf{x}_{k+1}^j = A\mathsf{x}_k^j + B\mathsf{u}_k^j + E\mathsf{w}_k^j$$
$$\mathsf{y}_k^j = C\mathsf{x}_k^j + D\mathsf{u}_k^j + F\mathsf{w}_k^j$$
$$j = 0, 1, \cdots, L-1$$

Uncertainty sampling

$u_k = U_k(\omega),\ y_k = Y_k(\omega),\ \cdots$

Stochastic fundamental lemma

LTI dynamics of sampled trajectories

$$x_{k+1} = Ax_k + Bu_k + Ew_k$$
$$y_k = Cx_k + Du_k + Fw_k$$

Measurements $u^{\mathrm{d}},\ w^{\mathrm{d}},\ y^{\mathrm{d}}$

Persistently exciting $u^{\mathrm{d}},\ w^{\mathrm{d}}$
controllable $(A, [B\ E])$

Data-driven uncertainty propagation

$$\begin{bmatrix} \mathscr{H}_N(u^{\mathrm{d}}) \\ \mathscr{H}_N(w^{\mathrm{d}}) \\ \mathscr{H}_N(y^{\mathrm{d}}) \end{bmatrix} G = \begin{bmatrix} U_{[0,N-1]} \\ W_{[0,N-1]} \\ Y_{[0,N-1]} \end{bmatrix}$$
$$G \in \mathrm{L}^2(\Omega, \mu; \mathbb{R}^{T-N+1})$$

How to consider distributional ambiguity?

G. Pan, R. Ou, and T. Faulwasser, 2023. On a stochastic fundamental lemma and its use for data-driven optimal control. *IEEE TAC*
T. Faulwasser et al. 2023. Behavioral theory for stochastic systems? A data-driven journey from Willems to Wiener and back again. *Annual Reviews in Control*

Considered Distributionally Robust OCP

Min-max formulation

$$\min_{\substack{U_k \in \mathrm{L}^2,\\ k\in\mathbb{I}_{[0,N-1]}}} \max_{\substack{W_k \in \mathcal{W},\\ k\in\mathbb{I}_{[0,N-1]}}} \mathbb{E}\left[\sum_{k=0}^{N-1} Y_k{}^\top Q Y_k + U_k{}^\top R U_k\right]$$

subject to

$$X_{k+1} = AX_k + BU_k + EW_k, \quad X_0 = X_{ini}$$

$$Y_k = CX_k + DU_k + FW_k$$

$$\mathbb{P}[Y_k \in \mathbb{Y}] \geq 1-\varepsilon_y, \ \forall\, W_k \in \mathcal{W}$$

$$\mathbb{P}[U_k \in \mathbb{U}] \geq 1-\varepsilon_u, \ \forall\, W_k \in \mathcal{W}$$

Empirial distribution $\bar{\mu}_W$ with expectation $\bar{m}$ and covariance $\bar{\Sigma}$

Distributions *close* to the empirical distribution $\bar{\mu}_W$

- Wasserstein metric $d_{\mathsf{W}}(\mu_W, \bar{\mu}_W) = \inf_{W \sim \mu_W, \bar{W} \sim \bar{\mu}_W} \left(\mathbb{E}\left[\|W - \bar{W}\|^2 \right] \right)^{\frac{1}{2}}$
- Wasserstein ambiguity set $\mathcal{W}(\bar{\mu}_W, \rho) = \{ \mu_W \mid d_{\mathsf{W}}(\mu_W, \bar{\mu}_W) \leq \rho \}$

$$\mathcal{W}(\bar{\mu}_W, \rho) \subseteq \mathcal{G}(\bar{m}, \bar{\Sigma}, \rho)$$

Distributions *close* to the empirical moments $(\bar{m}, \bar{\Sigma})$

- Gelbrich distance $d_{\mathsf{G}}((m, \Sigma), (\bar{m}, \bar{\Sigma})) = \sqrt{\|m - \bar{m}\|^2 + \mathrm{tr}\left(\Sigma + \bar{\Sigma} - 2(\bar{\Sigma}^{\frac{1}{2}} \Sigma \bar{\Sigma}^{\frac{1}{2}})^{\frac{1}{2}} \right)}$
- Gelbrich ambiguity set $\mathcal{G}(\bar{m}, \bar{\Sigma}, \rho) = \{ \mu_W \mid d_{\mathsf{G}}((\mathbb{E}[W], \Sigma[W]), (\bar{m}, \bar{\Sigma})) \leq \rho \}$

Gelbrich ambiguity for uncertainty propagation?

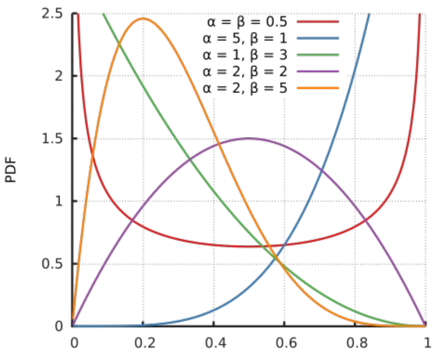
C. R. Givens and R. M. Shortt, 1984. A class of Wasserstein metrics for probability distributions. *Michigan Math. J.*

Uncertainty Quantification Perspective on PCE Bases

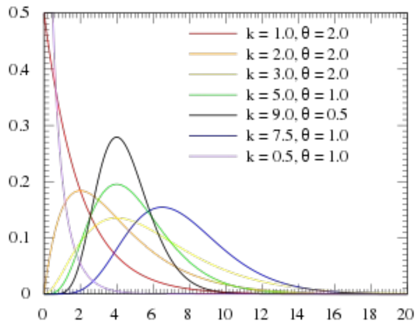
TUHH

Exact PCEs with 2 coefficients for specific distributions

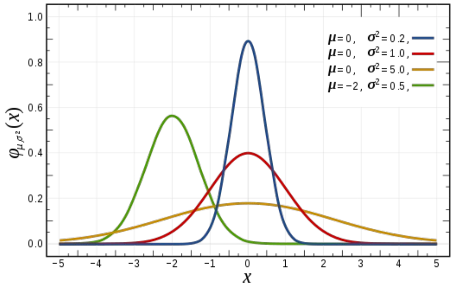
Beta



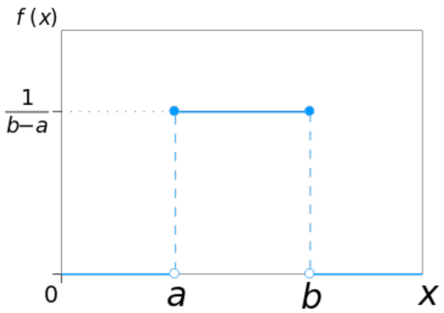
Gamma



Gaussian



Uniform



Images: wikipedia.org

$$W = \sum_{j=0}^{L-1} \mathrm{w}^j \phi^j(\boldsymbol{\xi})$$

Distribution	Support	Orthogonal basis polynomials	Argument
Gaussian	$(-\infty, \infty)$	Hermite	$\xi \sim \mathcal{N}(0,1)$
Uniform	$[a, b]$	Legendre	$\xi \sim \mathcal{U}([-1,1])$
Beta	$[a, b]$	Jacobi	$\xi \sim \mathcal{B}(\alpha, \beta, [-1,1])$
Gamma	$(0, \infty)$	Laguerre	$\xi \sim \Gamma(\alpha, \beta, (0,\infty))$

Design degrees of freedom for PCE bases are two-fold:

- **Algebraic structure** of basis functions
- **Arguments** of basis

Sullivan, Timothy John. *Introduction to uncertainty quantification*. Vol. 63. Springer, 2015.

How to choose the basis for PCE?

- **Exact PCEs with 2 coefficients** for any L² distributions $$W = \sum_{j=0}^{L-1} \mathrm{w}^j \phi^j(\boldsymbol{\xi})$$

Scalar Gaussian

$W \in \mathrm{L}^2(\Omega, \mu; \mathbb{R}), W \sim \mathcal{N}(m, \sigma^2)$ → $W = m + \sigma \cdot \xi, \quad \xi \sim \mathcal{N}(0,1)$

Arbitrary L² random variable

$W \in \mathrm{L}^2(\Omega, \mu; \mathbb{R}), \mathbb{E}[W] = m$ → $W = m \cdot 1 + 1 \cdot (W - m), \quad \phi_0 = 1, \ \phi_1 = W - m$

Moments from PCE coefficients?

- Expectation: w^0
- Variance: $\sum_{j=1}^{L-1} (\mathrm{w}^j)^2$, Covariance: ...

Generalization: construct PCE basis by covariance decomposition

PCE Representation for Gelbrich Ambiguity

TUHH

Covariance decomposition leads to exact vector-valued PCE

$W \in \mathrm{L}^2(\Omega,\mu;\mathbb{R}^{n_w}),\ \mathbb{E}[W]=m,\ \Sigma[W]=MM^\top \Rightarrow W = m + M\cdot\xi, \quad \xi\in\mathrm{L}^2(\Omega,\mu;\mathbb{R}^{n_w}), \mathbb{E}[\xi]=0, \Sigma[\xi]=I_{n_w}$

Nonconvex Gelbrich ambiguity set

$\mu_W \in \mathcal{G}(\bar m,\bar\Sigma,\rho)$

$$\left\{\mu_W \,\middle|\, \sqrt{\|m-\bar m\|^2 + \mathrm{tr}\left(\Sigma+\bar\Sigma-2(\bar\Sigma^{\frac12}\Sigma\bar\Sigma^{\frac12})^{\frac12}\right)} \le \rho\right\}$$

Convex PCE coefficient uncertainty set

$[m\,|\,M] \in \mathbb{G}(\bar m,\bar M,\rho)$

$$\left\{[m\,|\,M] \,\middle|\, \bar\Sigma^{\frac12} M \succeq 0, \left\|[m\,|\,M]-[\bar m\,|\,\bar M]\right\|_F \le \rho\right\}$$

Germs ξ with normalized ambiguity

$\mu_\xi \in \mathcal{G}(0, I_{n_w}, 0)$

Tailored covariance decomposition for Gelbrich ambiguity

$$W = m + M\xi,\ M = \bar\Sigma^{-\frac12}(\bar\Sigma^{\frac12}\Sigma\bar\Sigma^{\frac12})^{\frac12},\ \bar M = \bar\Sigma^{\frac12}$$

Representation of Gelbrich ambiguity: coefficients m, M *and* germs $\xi \in \mathrm{L}^2(\Omega,\mu;\mathbb{R}^{n_w})$

G. Pan, R. Ou, and T. Faulwasser, 2023. On Data-Driven Stochastic Output-Feedback Predictive Control, *arXiv preprint*

Representation of Input and Output Random Variables?

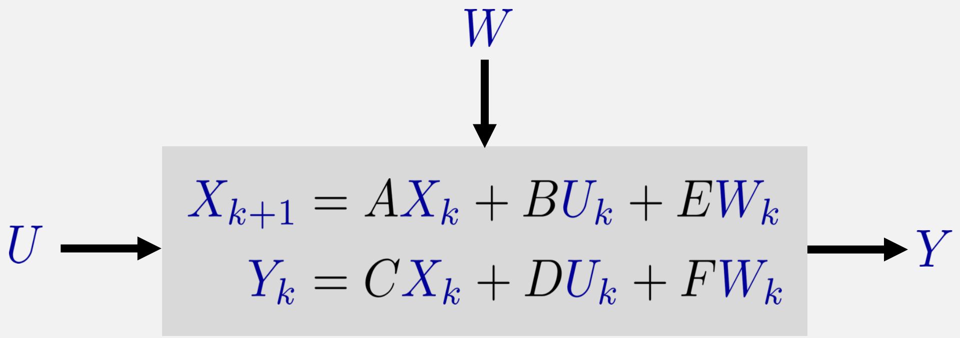
TUHH

Exact PCEs for all random variables?

- Exact PCEs $W_k = m + M\xi_k$, $\mu_{\boldsymbol{\xi}_k} \in \mathcal{G}(0, I_{Nn_w}, 0)$
- Let $\boldsymbol{\xi} = [\xi_0^\top, \cdots, \xi_{N-1}^\top]$, $\mu_{\boldsymbol{\xi}} \in \mathcal{G}(0, I_{Nn_w}, 0)$
- LTI dynamics
- Control U is affine and modelled as causal disturbance feedback

$$U_{[0,N-1]} = \bar{u}_{[0,N-1]} + K_w W_{[0,N-1]},$$

$$K_w = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ K_{1,0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \mathbf{0} & \vdots \\ K_{N-1,0} & \cdots & K_{N-1,N-2} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{Nn_u \times Nn_w}.$$



$\Rightarrow$ U_k and Y_k depend affinely on $\boldsymbol{\xi}$

$\Rightarrow$ Exact PCEs: $U_k = \sum_{j=0}^{L-1} \mathsf{u}_k^j \phi^j(\boldsymbol{\xi})$, $Y_k = \sum_{j=0}^{L-1} \mathsf{y}_k^j \phi^j(\boldsymbol{\xi})$, $\cdots$

PCE reformulation of distributionally robust chance constraints?

G.C. Calafiore and L.E. Ghaoui, 2006. On distributionally robust chance constrained linear programs, *J. Optim. Theory Appl.*

Distributionally Robust Chance Constraints?

Distributionally robust constraints in PCE coefficients

$$\mathbb{P}[a^\top U_k(\boldsymbol{\xi}) \leq 1] \geq 1-\varepsilon,\ \forall \mu_{\boldsymbol{\xi}} \in \mathcal{G}(0, I_{Nn_w}, 0) \quad \Leftrightarrow \quad a^\top \mathsf{u}_k^0 + \sqrt{\tfrac{1-\varepsilon}{\varepsilon} \textstyle\sum_{j=1}^{Nn_w} (a^\top \mathsf{u}_k^j)^2} \leq 1$$

$$\mathbb{P}[a^\top Y_k(\boldsymbol{\xi}) \leq 1] \geq 1-\varepsilon,\ \forall \mu_{\boldsymbol{\xi}} \in \mathcal{G}(0, I_{Nn_w}, 0) \quad \Leftrightarrow \quad a^\top y_k^0 + \sqrt{\tfrac{1-\varepsilon}{\varepsilon} \textstyle\sum_{j=1}^{Nn_w} (a^\top \mathsf{y}_k^j)^2} \leq 1$$

U_k and Y_k depend affinely on $\boldsymbol{\xi}$ + orthonormality of $\boldsymbol{\xi}$
$\Rightarrow$ Exact second-order cone reformulation independent of $\boldsymbol{\xi}$

Do we need $\boldsymbol{\xi}$ for distributionally robust optimal control?

G.C. Calafiore and L.E. Ghaoui, 2006. On distributionally robust chance constrained linear programs, *J. Optim. Theory Appl.*

Distributionally Robust Optimal Control Problem

TUHH

$$\min_{\bar{u},K_w} \max_{m,M} \sum_{k=0}^{N-1} \sum_{j=0}^{L-1} (\|\mathsf{y}_k^j\|_Q^2 + \|\mathsf{u}_k^j\|_R^2)$$

s.t. $\forall [m, M] \in \mathbb{G}(\bar{m}, \bar{\Sigma}, \rho),\ \forall k \in \mathbb{I}_{[0,N-1]},$

$$\mathsf{w}_{[0,N-1]}^{[0,L-1]} = [\mathbf{1}_N \otimes m, I_N \otimes M]$$

$$\mathscr{H}_{N+l}(\mathbf{u}^{\mathrm{d}}, \mathbf{w}^{\mathrm{d}}, \mathbf{y}^{\mathrm{d}})\mathsf{g}^j = (\mathsf{u}, \mathsf{w}, \mathsf{y})_{[-l,N-1]}^j$$

$$(\mathsf{u}, \mathsf{w}, \mathsf{y})_{[-l,-1]}^0 = (u, w, y)_{[-l,-1]}$$

$$\mathsf{u}_{[0,N-1]}^j = \delta^{0j} \bar{u}_{[0,N-1]} + K_w \mathsf{w}_{[0,N-1]}^j$$

$$a_{u,i}^\top \mathsf{u}_k^0 + \sqrt{\tfrac{1-\varepsilon_u}{\varepsilon_u} \textstyle\sum_{j=1}^{Nn_w} (a_{u,i}^\top \mathsf{u}_k^j)^2} \le 1, \forall i \in \mathbb{I}_{[1,N_u]}$$

$$a_{y,i}^\top \mathsf{y}_k^0 + \sqrt{\tfrac{1-\varepsilon_y}{\varepsilon_y} \textstyle\sum_{j=1}^{Nn_w} (a_{y,i}^\top \mathsf{y}_k^j)^2} \le 1, \forall i \in \mathbb{I}_{[1,N_y]}$$

Convex Gelbrich ambiguity

Data-driven uncertainty propagation via PCE

Affine disturbance feedback policy

Second-order cone reformulation

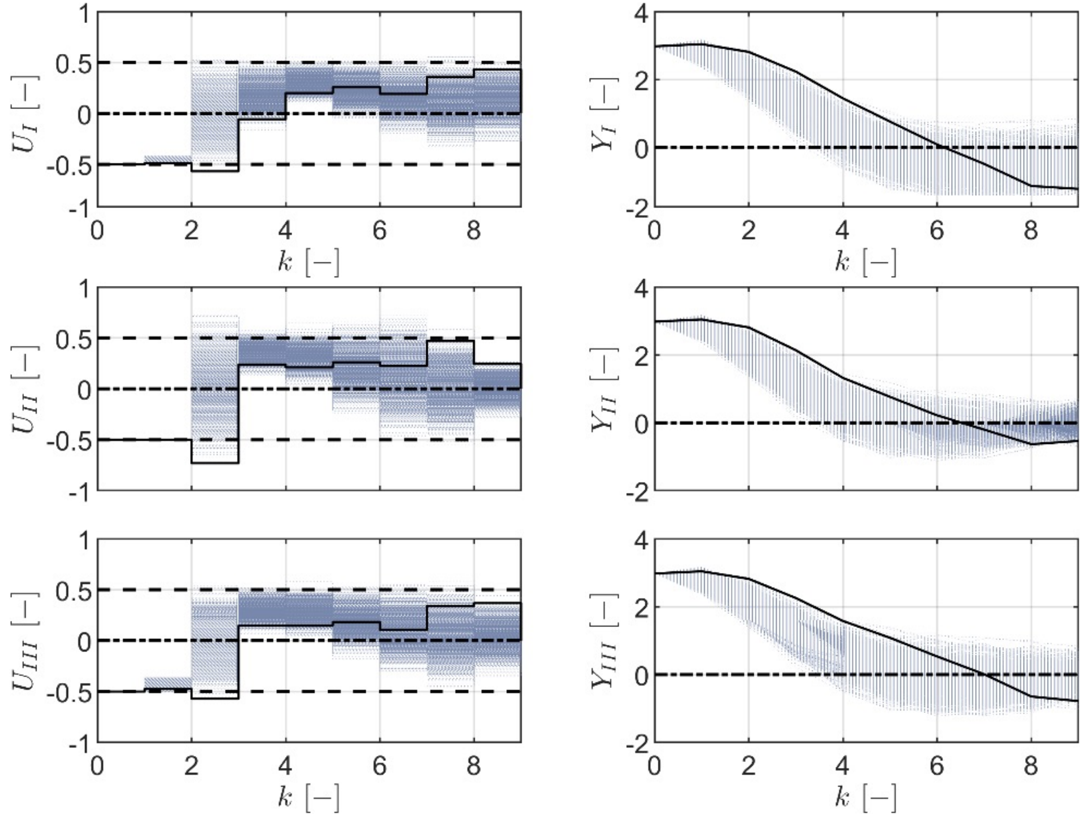
- Robust optimization with convex ambiguity set independent of ξ
- Approximation of $\mathbb{G}$ by polytope $\Rightarrow$ tractable second-order cone program
- Open-loop distributionally robust solution: $u_{[0,N-1]} = \bar{u}_{[0,N-1]} + K_w w_{[0,N-1]}$

Numerical Example

$$X_{k+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} X_k + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} U_k + W_k,$$
$$Y_k = \begin{bmatrix} 1 & 0 \end{bmatrix} X_k$$

Non Gaussian μ_W: mixture of $\mathcal{N}([\begin{smallmatrix} 0.1 \\ 0.1 \end{smallmatrix}], 0.01 I_2)$ and $\mathcal{N}([\begin{smallmatrix} -0.1 \\ -0.1 \end{smallmatrix}], 0.01 I_2)$

$\mathbb{P}[U_k \leq 0.5] \geq 80\%,\ \mathbb{P}[U_k \geq -0.5] \geq 80\%,\ Q = R = 1,\ N = 10$



	Averaged cost	Violations at $k = 3$
Ideal $\mathcal{G}(m_{\text{true}}, \Sigma_{\text{true}}, 0)$	25.04	0.3 %
Optimistic $\mathcal{G}(\bar{m}, \bar{\Sigma}, 0)$	24.47	5.8 %
Distributionally robust $\mathcal{A}(\bar{m}, \bar{\Sigma}, \rho)$, $\rho = 0.5\|[\bar{m}, \bar{\Sigma}]\|_F$	25.79	0.8 %

G. Pan and T. Faulwasser, 2023. Distributionally robust uncertainty quantification via data-driven stochastic optimal control. IEEE L-CSS

Summary

Distributional ambiguity via Gelbrich sets

- Outer approximation of Wasserstein ambiguity set in moment-based formulation
- Convex reformulation by a tailored decomposition of covariance matrix
- Decoupling into ambiguities of coefficients and germs via polynomial chaos

G. Pan and T. Faulwasser, 2023. Distributionally robust uncertainty quantification via data-driven stochastic optimal control. IEEE L-CSS

Data-driven uncertainty propagation

- Prediction of future random variables by previously recorded data
- Tractable propagation by exact PCE
- Second order cone reformulation of distributionally robust constraints

G. Pan, R. Ou, and T. Faulwasser, 2023. On a stochastic fundamental lemma and its use for data-driven optimal control. IEEE TAC

Data-driven distributionally robust optimal control

Extensions to predictive control and closed-loop analysis

- Data-driven stochastic output-feedback predictive control with closed-loop guarantees

G. Pan, R. Ou, and T. Faulwasser, 2023. On Data-Driven Stochastic Output-Feedback Predictive Control, *arXiv preprint*

Summary

Distributional ambiguity via Gelbrich sets

- Outer approximation of Wasserstein ambiguity set in moment-based formulation
- Convex reformulation by a tailored decomposition of covariance matrix
- Decoupling into ambiguities of coefficients and germs via polynomial chaos

G. Pan and T. Faulwasser, 2023. Distributionally robust uncertainty quantification via data-driven stochastic optimal control. IEEE L-CSS

Data-driven uncertainty propagation

- Prediction of future random variables by previously recorded data
- Tractable propagation by exact PCE
- Second order cone reformulation of distributionally robust constraints

G. Pan, R. Ou, and T. Faulwasser, 2023. On a stochastic fundamental lemma and its use for data-driven optimal control. IEEE TAC

Data-driven distributionally robust optimal control

Open problems

- Tools for numerical implementation, **fundamental lemma for nonlinear systems**?

O. Molodchyk & T. Faulwasser, 2023. Exploring the Links between the Fundamental Lemma and Kernel Regression. *IEEE L-CSS*

Thank you for your attention!