Data-driven and Robust Distribution Steering

Panagiotis Tsiotras

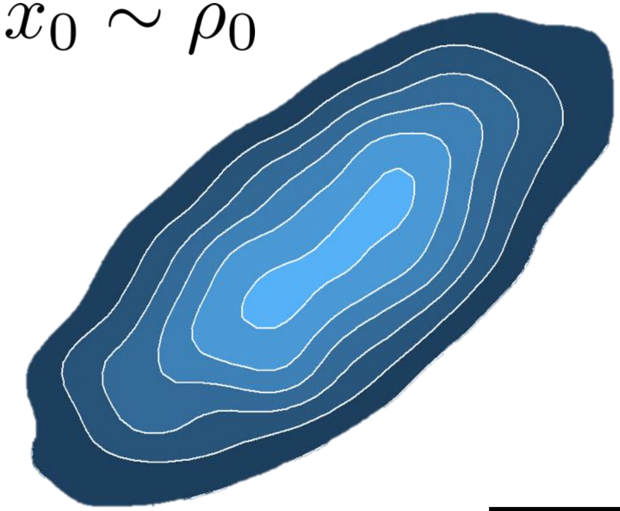
School of Aerospace Engineering
Institute for Robotics and Intelligent Machines
Georgia Institute of Technology
(joint work with J. Pilipovsky)

CDC 2025 – Stochastic Planning and Control Workshop
Rio de Janeiro, December 9, 2025

Distribution Steering

$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{\partial}{\partial x}\left(f(x,u)\rho(x,t)\right) + \tfrac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2\rho(x,t)\right)$$

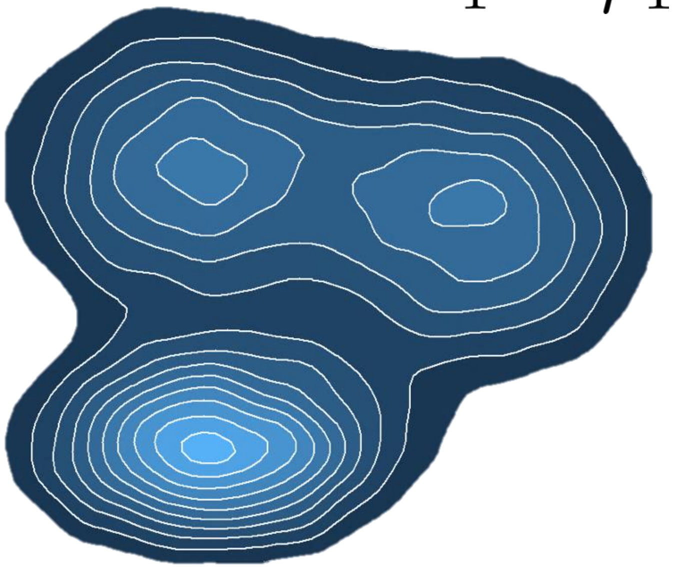
$x_0 \sim \rho_0$



$$\mathrm{d}x = f(x,u)\mathrm{d}t + \sigma \mathrm{d}w$$

$$\mathrm{d}x = f(x,u)\mathrm{d}t$$

$x_1 \sim \rho_1$

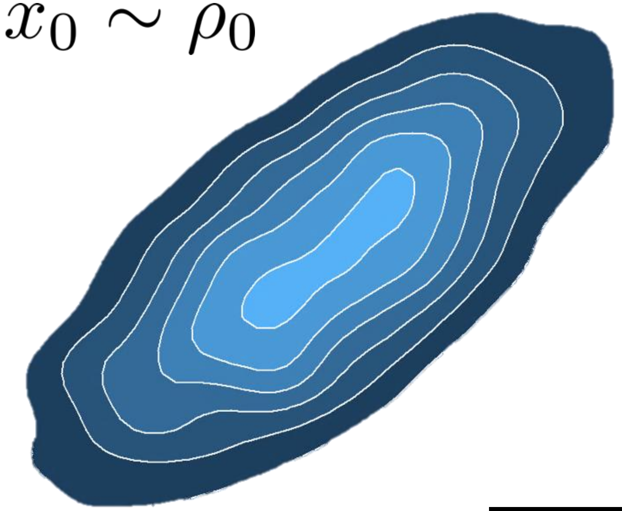


$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{\partial}{\partial x}\left(f(x,u)\rho(x,t)\right)$$

Distribution Steering

$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{\partial}{\partial x}\left(f(x,u)\rho(x,t)\right) + \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(\sigma^2\rho(x,t)\right)$$

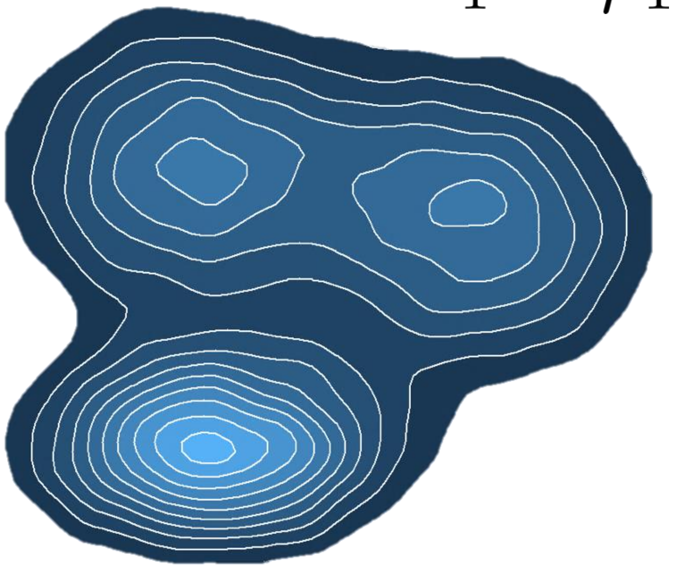
$x_0 \sim \rho_0$



Uncertainty Quantification

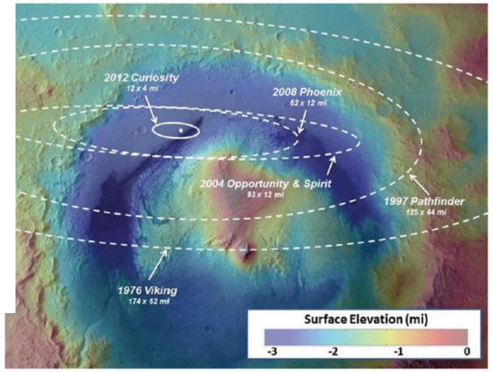
Uncertainty Mitigation

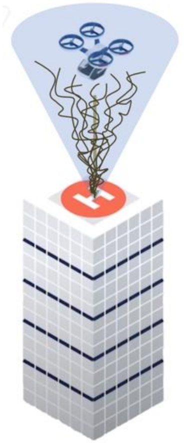
$x_1 \sim \rho_1$



$$\frac{\partial}{\partial t}\rho(x,t) = -\frac{\partial}{\partial x}\left(f(x,u)\rho(x,t)\right)$$

Many Applications





Pinpoint landing

SciTech 2019

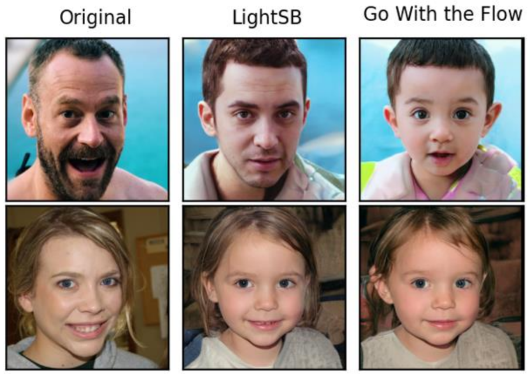
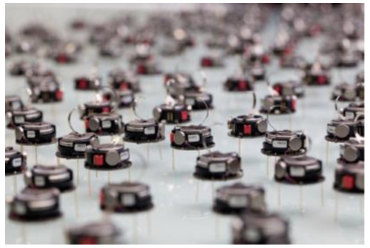


Image-to-Image translation

ChatGPT: Generate an image of a GT astronaut cat chasing a tennis ball.

Generative modeling

NeurIPS 2025

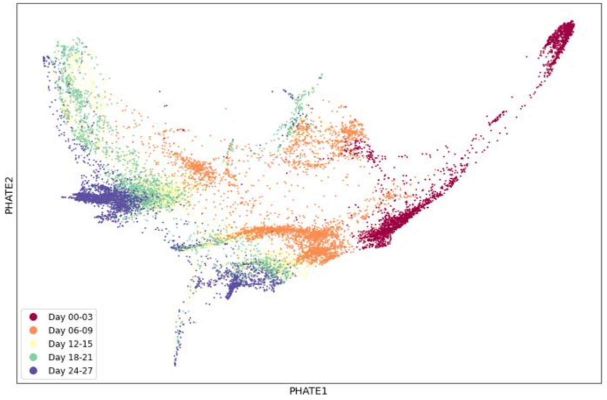


Swarm robotics

L-CSS 2025

Precision package delivery

Embryonic cell modeling



NeurIPS 2025

Problem Formulation

Consider discrete-time stochastic linear system

$$x_{k+1} = A_k x_k + B_k u_k + D_k w_k$$

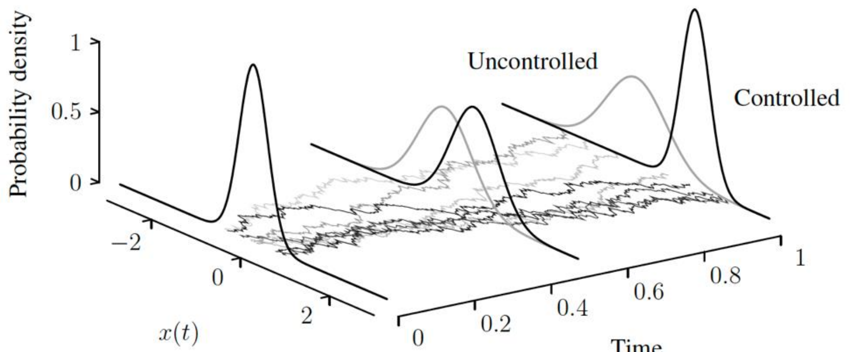
- We wish the initial and final states to be distributed according to

$$x_0 \sim \mathcal{N}(\mu_0, \Sigma_0), \qquad x_N \sim \mathcal{N}(\mu_N, \Sigma_N)$$

where $\mu_0, \Sigma_0, \mu_N, \Sigma_N$ given, while minimizing the cost function

$$J(x,u) = \mathbb{E}\left[\sum_{k=0}^{N-1} x_k^\top Q_k x_k + u_k^\top R_k u_k\right] + x_N^\top Q_N x_N$$

where $Q_k \succeq 0$ and $R_k \succ 0$ for all $k = 0, 1, \ldots, N-1$.



- Assume that $\Sigma_0 \succeq 0$ and $\Sigma_N \succ 0$,

Convex Programing Approach

- Let the control

$$u_k = K_k(x_k - \mu_k) + v_k,$$

Dynamics of the state mean and covariance

$$\mu_{k+1} = A_k \mu_k + B_k v_k,$$
$$\Sigma_{k+1} = (A_k + B_k K_k)\Sigma_k (A_k + B_k K_k)^\mathsf{T} + D_k D_k^\mathsf{T}.$$

- Use the standard transformation $U_k \triangleq K_k \Sigma_k$

Y_k

$$\min_{\Sigma_k, U_k, \mu_k, v_k} \quad J = \sum_{k=0}^{N-1} \mathrm{tr}(Q_k \Sigma_k) + \mathrm{tr}\left(R_k U_k \Sigma_k^{-1} {U_k}^\mathsf{T}\right) + {\mu_k}^\mathsf{T} Q_k \mu_k + {v_k}^\mathsf{T} R_k v_k$$

$$\mu_{k+1} = A_k \mu_k + B_k v_k$$

Y_k

$$A_k \Sigma_k A_k^\mathsf{T} - B_k U_k A_k^\mathsf{T} - A_k {U_k}^\mathsf{T} B_k^\mathsf{T} + B_k U_k \Sigma_k^{-1} {U_k}^\mathsf{T} B_k^\mathsf{T} + D_k D_k^\mathsf{T} - \Sigma_{k+1} = 0$$

Lossless Convexification

Theorem (Rapakoulias & PT, 2023)

The following convex relaxation is lossless

$$\min_{\Sigma_k, U_k, Y_k} \quad J_\Sigma = \sum_{k=0}^{N-1} \mathtt{tr}\big(Q_k \Sigma_k\big) + \mathtt{tr}\big(R_k Y_k\big)$$

$$U_k \Sigma_k^{-1} {U_k}^\intercal - Y_k \preceq 0 \quad \Longleftrightarrow \quad \begin{bmatrix} \Sigma_k & {U_k}^\intercal \\ U_k & Y_k \end{bmatrix} \succeq 0$$

$$A_k \Sigma_k A_k^\intercal - B_k U_k A_k^\intercal - A_k {U_k}^\intercal B_k^\intercal + B_k Y_k B_k^\intercal + D_k D_k^\intercal - \Sigma_{k+1} = 0$$

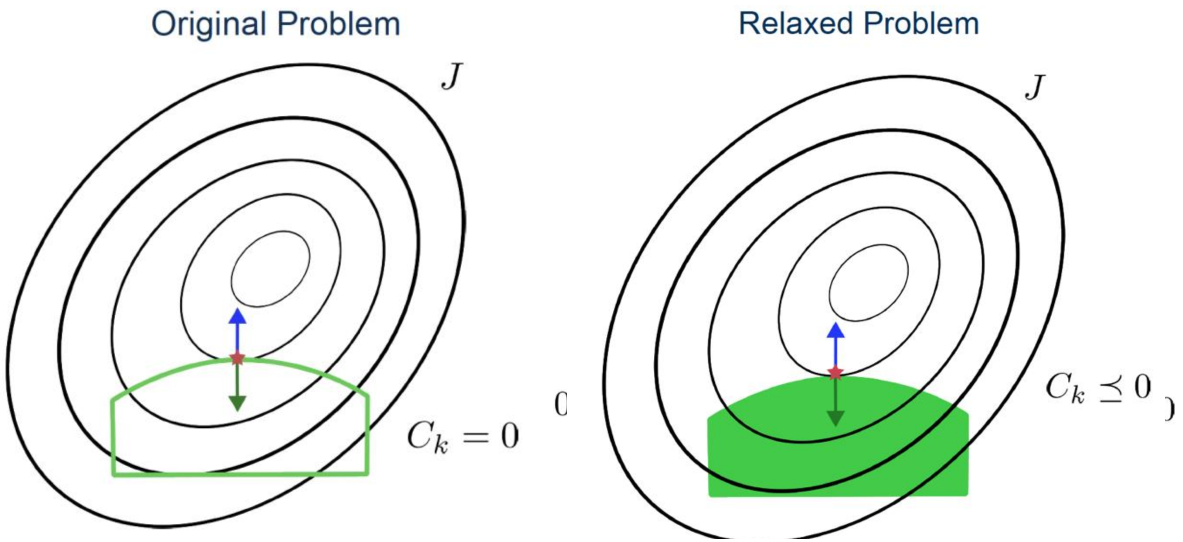
F. Liu, G. Rapakoulias, and P. Tsiotras, "Optimal Covariance Steering for Discrete-Time Linear Stochastic Systems," *IEEE Transactions on Automatic Control*, Vol. 70, No. 4, pp. 1558-2523, 2025

Chance Constraints

$$\mathtt{Pr}(\alpha_x^{\mathsf{T}} x_k \le \beta_x) \ge 1 - \epsilon_x$$
$$\mathtt{Pr}(\alpha_u^{\mathsf{T}} u_k \le \beta_u) \ge 1 - \epsilon_u$$

$$\Phi^{-1}(1-\epsilon_x)\sqrt{\alpha_x^{\mathsf{T}} \Sigma_k \alpha_x} + \alpha_x^{\mathsf{T}} \mu_k - \beta_x \le 0$$
$$\Phi^{-1}(1-\epsilon_u)\sqrt{\alpha_u^{\mathsf{T}} U_k \Sigma_k^{-1} U_k^{\mathsf{T}} \alpha_u} + \alpha_u^{\mathsf{T}} v_k - \beta_u \le 0$$

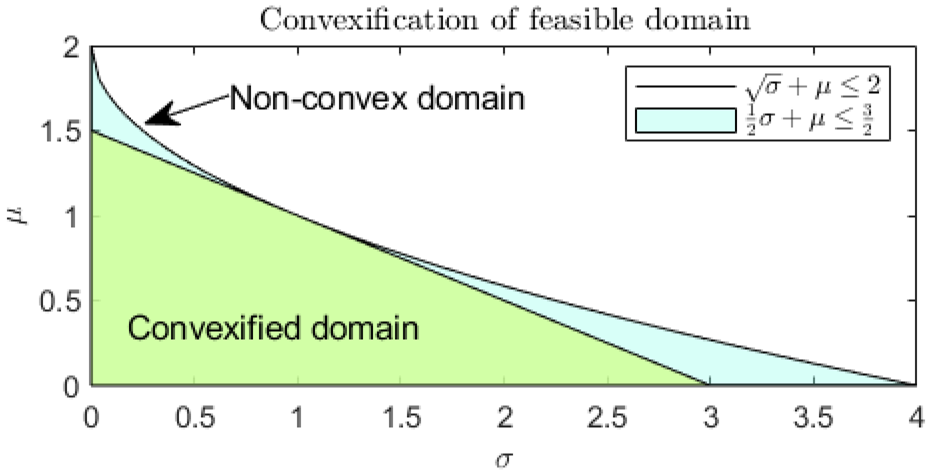
$$\ell^{\mathsf{T}} \Sigma_k \ell + \alpha_x^{\mathsf{T}} \mu_k - \beta_x \le 0$$
$$e^{\mathsf{T}} Y_k e + \alpha_u^{\mathsf{T}} v_k - \beta_u \le 0$$



Lossless CC relaxation

The optimal solution satisfies

$$C_k := U_k \Sigma_k^{-1} U_k^{\mathsf{T}} - Y_k = 0$$



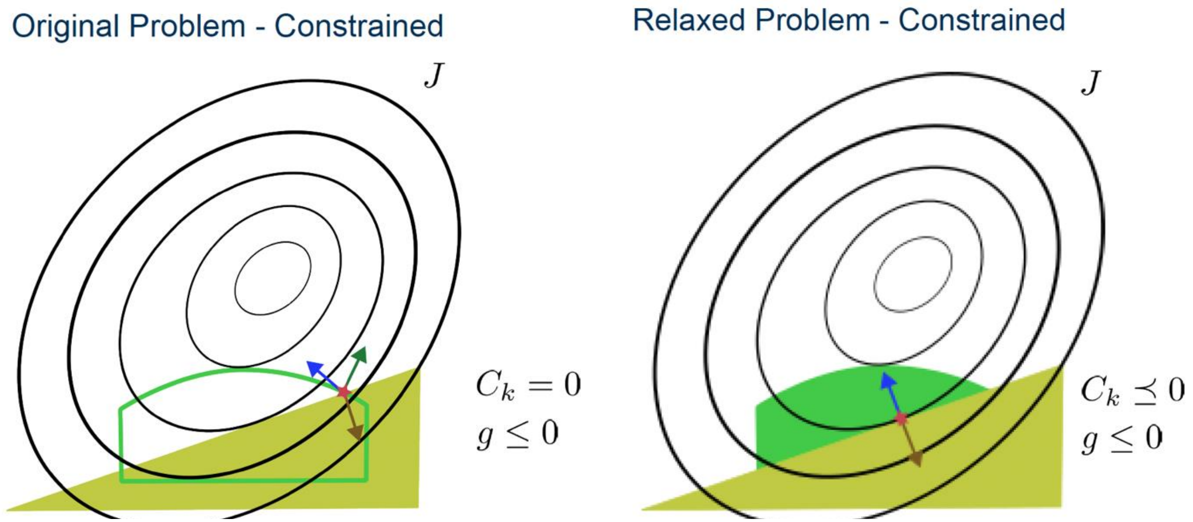
G. Rapakoulias, and P. Tsiotras, "Discrete-time Optimal Covariance Steering via Semidefinite Programming," *62nd IEEE Conference on Decision and Control*, Marina Bay Sands, Singapore, Dec. 13-15, 2023, pp. 1802-1807

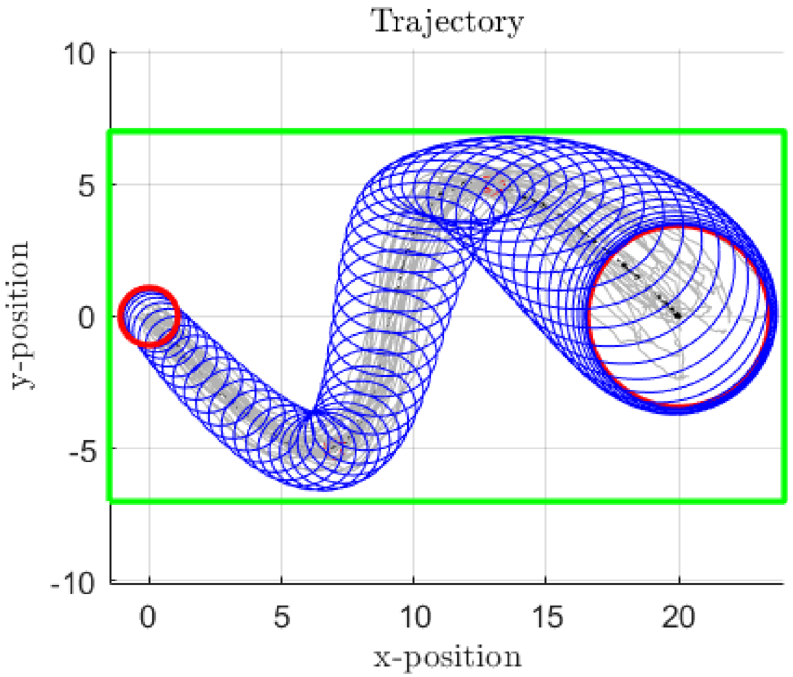
Chance Constraints

$$\texttt{Pr}(\alpha_x^\intercal x_k \leq \beta_x) \geq 1 - \epsilon_x$$
$$\texttt{Pr}(\alpha_u^\intercal u_k \leq \beta_u) \geq 1 - \epsilon_u$$

$$\Phi^{-1}(1-\epsilon_x)\sqrt{\alpha_x^\intercal \Sigma_k \alpha_x} + \alpha_x^\intercal \mu_k - \beta_x \leq 0$$
$$\Phi^{-1}(1-\epsilon_u)\sqrt{\alpha_u^\intercal Y_k \alpha_u} + \alpha_u^\intercal v_k - \beta_u \leq 0$$

$$U_k \Sigma_k^{-1} U_k^\intercal - Y_k \preceq 0$$





Lossless CC relaxation

The optimal solution satisfies

$$C_k := U_k \Sigma_k^{-1} U_k^\intercal - Y_k = 0$$

G. Rapakoulias, and P. Tsiotras, "Discrete-time Optimal Covariance Steering via Semidefinite Programming," *62nd IEEE Conference on Decision and Control*, Marina Bay Sands, Singapore, Dec. 13-15, 2023, pp. 1802-1807

Data-Driven CS

Consider the discrete-time *deterministic* system

$$x_{k+1} = Ax_k + Bu_k, \qquad A, B\text{: unknown}$$

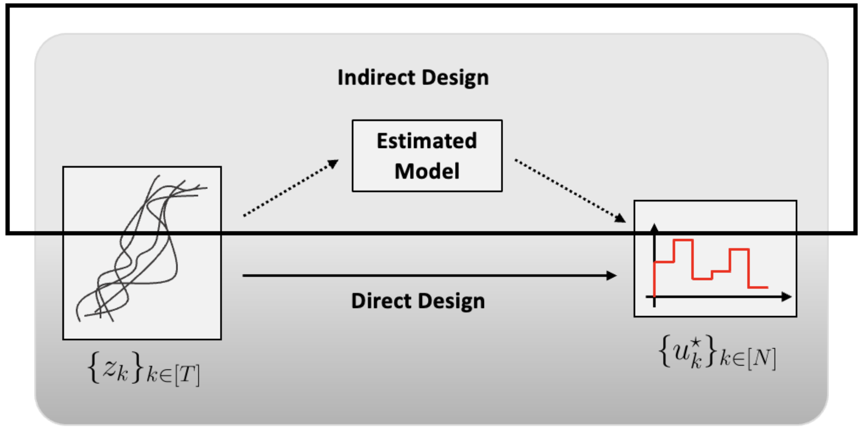
$$x_0 \sim \mathcal{N}(\mu_i, \Sigma_i), \quad x_N = x_f \sim \mathcal{N}(\mu_f, \Sigma_f)$$

$$J(u_0, \ldots, u_{N-1}) := \mathbb{E}\left[\sum_{k=0}^{N-1} x_k^\intercal Q_k x_k + u_k^\intercal R_k u_k\right]$$

Two potential approaches

- Indirect approach
- Direct approach

Willems' Fundamental Lemma, characterizes all trajectories of an LTI system through the range space of an input/output data matrix



$$U_{0,T} := [u_0^{(d)} \;\; u_1^{(d)} \;\; \ldots \;\; u_{T-1}^{(d)}],$$
$$X_{0,T} := [x_0^{(d)} \;\; x_1^{(d)} \;\; \ldots \;\; x_{T-1}^{(d)}],$$
$$X_{1,T} := [x_1^{(d)} \;\; x_2^{(d)} \;\; \ldots \;\; x_T^{(d)}].$$

$$X_{1,T} = AX_{0,T} + BU_{0,T}$$

Pilipovsky, J., and Tsiotras, P., "Data-Driven Covariance Steering Control Design," *62nd IEEE Conference on Decision and Control*, Marina Bay Sands, Singapore, Dec. 13-15, 2023, pp. 2610-2615.

Data-Driven CS

Consider the discrete-time *deterministic* system

$$x_{k+1} = Ax_k + Bu_k, \qquad A, B: \text{ unknown}$$

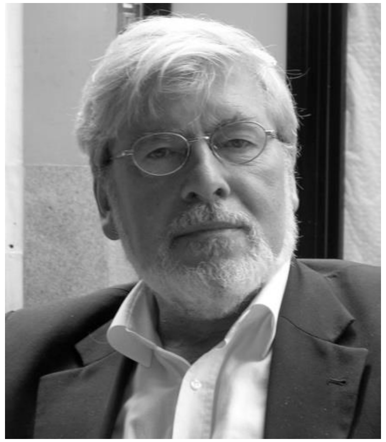
$$x_0 \sim \mathcal{N}(\mu_i, \Sigma_i), \quad x_N = x_f \sim \mathcal{N}(\mu_f, \Sigma_f)$$

$$J(u_0, \ldots, u_{N-1}) := \mathbb{E}\left[\sum_{k=0}^{N-1} x_k^\intercal Q_k x_k + u_k^\intercal R_k u_k\right]$$

Two potential approaches

- Indirect approach
- Direct approach

Willems' Fundamental Lemma, characterizes all trajectories of an LTI system through the range space of an input/output data matrix



Jan Willems (1939-2013)

$$U_{0,T} := [u_0^{(d)} \; u_1^{(d)} \; \ldots \; u_{T-1}^{(d)}],$$

$$X_{0,T} := [x_0^{(d)} \; x_1^{(d)} \; \ldots \; x_{T-1}^{(d)}],$$

$$X_{1,T} := [x_1^{(d)} \; x_2^{(d)} \; \ldots \; x_T^{(d)}].$$

$$X_{1,T} = [B \;\; A] \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix}$$

Pilipovsky, J., and Tsiotras, P., "Data-Driven Covariance Steering Control Design," *62nd IEEE Conference on Decision and Control*, Marina Bay Sands, Singapore, Dec. 13-15, 2023, pp. 2610-2615.

Assume control law

$$u_k = K_k(x_k - \mu_k) + v_k$$

Assume persistency of excitation

$$\text{rank} \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} = n + m$$

$$\begin{bmatrix} K_k \\ I_n \end{bmatrix} = \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} G_k$$

$$X_{1,T} = [B \;\; A] \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix}$$

Mean Steering

$$\min_{\mu_k, v_k} J_\mu = \sum_{k=0}^{N-1} (\mu_k^\intercal Q_k \mu_k + v_k^\intercal R_k v_k)$$

$$\mu_{k+1} = X_{1,T} \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix}^\dagger \begin{bmatrix} v_k \\ \mu_k \end{bmatrix} = \hat{A}_k \mu_k + \hat{B}_k v_k$$

$$\mu_0 = \mu_i, \quad \mu_N = \mu_f = 0$$

Alternative POV

$$[\hat{A} \quad \hat{B}] = \text{argmin}_{B,A} \left\| X_{1,T} - [B \quad A] \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} \right\|_{\mathrm{F}}$$

Covariance Steering

Assume control law

$$u_k = K_k(x_k - \mu_k) + v_k$$

Assume persistency of excitation

$$\operatorname{rank}\begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} = n + m$$

$$\begin{bmatrix} K_k \\ I_n \end{bmatrix} = \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} G_k$$

$$X_{1,T} = [B \;\; A] \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix}$$

$$\min_{\Sigma_k, K_k} \; J_\Sigma = \sum_{k=0}^{N-1} \left(\operatorname{tr}(Q_k \Sigma_k) + \operatorname{tr}(R_k K_k \Sigma_k K_k^\intercal)\right)$$

$$\Sigma_{k+1} = (A + BK_k)\Sigma_k(A + BK_k)^\intercal$$

$$\Sigma_0 = \Sigma_i \quad \Sigma_N = \Sigma_f$$

$$\Sigma_{k+1} = [B \;\; A] \begin{bmatrix} K_k \\ I_n \end{bmatrix} \Sigma_k \begin{bmatrix} K_k \\ I_n \end{bmatrix}^\intercal [B \;\; A]^\intercal$$

$$= X_{1,T} G_k \Sigma_k G_k^\intercal X_{1,T}^\intercal$$

$$S_k = G_k \Sigma_k$$

$$= X_{1,T} S_k \Sigma_k^{-1} S_k^\intercal X_{1,T}^\intercal$$

$$Y_k \succeq S_k \Sigma_k^{-1} S_k^\intercal$$

Assume control law

$$u_k = K_k(x_k - \mu_k) + v_k$$

Assume persistency of excitation

$$\operatorname{rank} \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} = n + m$$

$$\begin{bmatrix} K_k \\ I_n \end{bmatrix} = \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} G_k$$

$$X_{1,T} = [B \;\; A] \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix}$$

Convex Optimization Problem!

Assume control law

$$u_k = K_k(x_k - \mu_k) + v_k$$

Assume persistency of excitation

$$\text{rank}\begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} = n + m$$

$$\Downarrow$$

$$\boxed{\begin{bmatrix} K_k \\ I_n \end{bmatrix} = \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} G_k}$$

$$X_{1,T} = [B \;\; A] \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix}$$

Covariance Steering

$$\min_{\Sigma_k, S_k, Y_k} \sum_{k=0}^{N-1} \left(\text{tr}(Q_k \Sigma_k) + \text{tr}(R_k U_{0,T} Y_k U_{0,T}^\intercal)\right)$$

$$\Sigma_0 = \Sigma_i, \quad \Sigma_N = \Sigma_f$$

$$X_{1,T} Y_k X_{1,T}^\intercal - \Sigma_{k+1} = 0$$

$$\Sigma_k - X_{0,T} S_k = 0$$

$$\begin{bmatrix} \Sigma_k & S_k^\intercal \\ S_k & Y_k \end{bmatrix} \succeq 0$$

Noisy Data

$$x_{k+1} = Ax_k + Bu_k + \xi_k, \quad \xi_k \sim \mathcal{N}(0, \Sigma_\xi)$$

$$X_{1,T} = AX_{0,T} + BU_{0,T} + \Xi_{0,T}, \qquad \Xi_{0,T} = [\xi_0, \ldots, \xi_{T-1}]$$

Mean Propagation

$$\mu_{k+1} = (X_{1,T} - \Xi_{0,T}) \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix}^{\dagger} \begin{bmatrix} v_k \\ \mu_k \end{bmatrix} \implies \min_{B,A} \left\| X_{1,T} - \Xi_{0,T} - \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} \right\|_{\mathrm{F}}$$

Covariance Propagation

$$\Sigma_{k+1} = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} K_k \\ I_n \end{bmatrix} \Sigma_k \begin{bmatrix} K_k \\ I_n \end{bmatrix}^{\intercal} \begin{bmatrix} B & A \end{bmatrix}^{\intercal} + \Sigma_\xi$$

$$= (X_{1,T} - \Xi_{0,T}) G_k \Sigma_k G_k^{\intercal} (X_{1,T} - \Xi_{0,T})^{\intercal} + \Sigma_\xi$$

ML Noise Estimation

Let

$$\mathcal{S} = \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} \implies X_{1,T} = \begin{bmatrix} B & A \end{bmatrix} \mathcal{S} + \Xi_{0,T}$$

Since $\mathcal{S}\mathcal{S}^\dagger\mathcal{S} = \mathcal{S}$ this yields the constraint

$$(X_{1,T} - \Xi_{0,T})(I_T - \mathcal{S}^\dagger\mathcal{S}) = 0$$

MLE Problem

$$\max_{\Xi_{0,T},\, \Sigma_\xi} \quad \mathcal{J}_{\mathrm{ML}} = \sum_{k=0}^{T-1} \log \rho_\xi(\xi_k)$$

$$(X_{1,T} - \Xi_{0,T})(I_T - \mathcal{S}^\dagger\mathcal{S}) = 0$$

where,

$$\rho_\xi(\xi) = \frac{1}{(2\pi)^{n/2}} (\det \Sigma_\xi)^{-1/2} \exp\left(-\frac{1}{2}\xi^\intercal \Sigma_\xi^{-1} \xi\right)$$

ML Noise Estimation

Let

$$\mathcal{S} = \begin{bmatrix} U_{0,T} \\ X_{0,T} \end{bmatrix} \implies X_{1,T} = \begin{bmatrix} B & A \end{bmatrix} \mathcal{S} + \Xi_{0,T}$$

Since $\mathcal{S}\mathcal{S}^{\dagger}\mathcal{S} = \mathcal{S}$ this yields the constraint

$$(X_{1,T} - \Xi_{0,T})(I_T - \mathcal{S}^{\dagger}\mathcal{S}) = 0$$

MLE Solution

The solution to the MLE problem for the most probable noise realization $\Xi_{0,T}$ and disturbance covariance matrix Σ_{ξ} is given by

$$\hat{\Xi}_{0,T} = X_{1,T}(I_T - \mathcal{S}^{\dagger}\mathcal{S})$$

$$\hat{\Sigma}_{\xi} = \frac{1}{T} X_{1,T}(I_T - \mathcal{S}^{\dagger}\mathcal{S}) X_{1,T}^{\intercal}$$

Estimation Error

Estimation Error

Assume $\Sigma_\xi \succ 0$ is known. For the MLE solution, we have the bound

$$\texttt{Pr}\{\Delta\Xi_{0,T} \in \boldsymbol{\Delta}\} \geq 1 - \delta$$

where,

$$\boldsymbol{\Delta} = \left\{ \|\Delta\Xi_{0,T}\| \leq \|\Sigma_\xi^{1/2}\| Q_{\chi^2_{n(n+m)}}^{1/2} (1-\delta) \right\}$$

Bound is *independent* of horizon T if data is persistently exciting!

Pilipovsky, J., and Tsiotras, P., "DUST: A Framework for Data-Driven Density Steering," 2025, arxiv.org/abs/2408.02777

Robust DD-DS

Robust DD Mean Steering


2020 59th IEEE Conference on Decision and Control (CDC)
Jeju Island, Republic of Korea, December 14-18, 2020

Robust Closed-loop Model Predictive Control via System Level Synthesis

Shaoru Chen*, Han Wang*, Manfred Morari, Victor M. Preciado, Nikolai Matni

Abstract—**In this paper, we consider the robust closed-loop model predictive control (MPC) of a linear time-variant (LTV) system with norm bounded disturbances and LTV model uncertainty, wherein a series of constrained optimal control problems (OCPs) are solved. Guaranteeing robust feasibility of these OCPs is challenging due to disturbances perturbing the predicted states, and model uncertainty, both of which can render the closed-loop system unstable. As such, a trade-off between the numerical tractability and conservativeness of the solutions is often required. We use the System Level Synthesis (SLS) framework to reformulate these constrained OCPs over closed-loop system responses, and show that this allows us to transparently account for norm bounded additive disturbances and LTV model uncertainty by computing robust state feedback policies. We further show that by exploiting the underlying linear fractional structure of the resulting robust OCPs, we can significantly reduce the conservativeness of existing SLS-based and tube-MPC-base... ...trol m...ds while also improving computationa... ...ciency. We c... ...de with numerical examples demonst... the effective... of our methods.**

I. Introduction

Model predictive control (MP... has a... ...ed r... ...k-able success in solving multivar... ...ained problems across a wide range of application areas, such as process control [1], power networks [2], and robot locomotion [3]. In MPC, a control action is computed by solving a finite-horizon constrained optimal control problem (OCP) at each sampling time, and then applying the first control action. The stability and performance of MPC depends on the accuracy of the model being used, and indeed robustness to both additive disturbances and model uncertainty must be considered. Although MPC using a nominal model (i.e., one ignoring uncertainty) offers some level of robustness [4], it has been shown that the closed-loop system achieved by nominal MPC can be destabilized by an arbitrarily small disturbance [5]. As a result, robust MPC, which explicitly deals with uncertainty, has received much attention [6].

When only additive disturbances are present, open-loop robust MPC, which optimizes over a sequence of control actions $\mathbf{u} = \{u_0, \cdots, u_{N-1}\}$ subject to suitably robust constraints, can be applied but tends to be overly conservative or even infeasible [7], as a single sequence of inputs $\mathbf{u}$ is chosen for all possible disturbance realizations. On the other hand, closed-loop MPC, which optimizes over the control policies $\pi = \{\pi_0(\cdot), \cdots, \pi_{N-1}(\cdot)\}$, can reduce the conservativeness of the solutions. However, the policy space is infinite-dimensional and renders the online OCPs intractable. The problem can be rendered tractable by restricting the policies π to lie in a function class that admits a finite-dimensional parameterization. For example, policies of the form $\pi_i(x) = Kx + v_i$ are considered in [8–10], where K is a pre-stabilizing feedback gain K that is fixed beforehand, thus reducing the decision variables to the vectors $\{v_1, \cdots, v_{...}\}$. To reduce conservativeness, an affi... ...k cont... ...aw $\pi(x) = K_i x + v_i$ can be appli... ...ith de...sion v... ...bles K_i and v_i; however, the resul... OCP is non-co... . In [11], it was observed that by re-... ...izing the ...Ps to be over disturbance based feedback ... of the ...n $\pi_i(w) = \sum_{j=0}^{i-1} M_{ij} w_j + v_i$, the resulting ... are ...vex. In [12, 13], the authors prop... an al... ...tive (t... MPC) invariant set based ap-...ba... ...ounds th... ...stem trajectories within a tube that robustly satisfy constraints.

The more challenging problem considering model uncertainty is tackled in [8, 13, 14]. When polytopic or structured feedback model uncertainty occurs, a linear matrix inequality (LMI) based robust MPC method is proposed in [14]. When both model uncertainty and additive disturbances are present, the method proposed in [13] designs tubes containing all possible trajectories under polytopic uncertainty assumptions. Alternative approaches based on dynamic programming (DP) [15] are shown to obtain tight solutions, but the computation quickly becomes intractable. Adaptive robust MPC, which considers estimation of the parametric uncertainty while implementing robust control, is proposed in [16–18].

As described above, there is a rich body of work addressing the robust MPC problem, and it remains an active area of research for which no definitive solution exists. Due to the inherent intractability of the general robust MPC problem subject to both additive disturbance and model uncertainty, all of the aforementioned methods trade off conservativeness for computational tractability in different ways. The recently developed *System Level Synthesis* (SLS) parameterization [19] provides an alternative approach to tackling the robust MPC problem and exploring this tradeoff

Han Wang is with the Department of Applied Mathematics and Computational Science, University of Pennsylvania, Philadelphia, PA, 19104, USA (e-mail: wanghan2@sas.upenn.edu).

Shaoru Chen, Manfred Morari, Victor M. Preciado and Nikolai Matni are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, 19104, USA (e-mail: {srchen, morari, preciado, nmatni}@seas.upenn.edu).

* The first two authors contributed equally to this paper.

 2152

Robust DD Covariance Steering

$$\min_{\Sigma_k, S_k, Y_k} \sum_{k=0}^{N-1} \left(\mathrm{tr}(Q_k \Sigma_k) + \mathrm{tr}(R_k U_{0,T} Y_k U_{0,T}^\intercal) \right)$$

$$\Sigma_0 = \Sigma_i, \quad \Sigma_N = \Sigma_f$$

$$(X_{1,T} - \Xi_{0,T}) S_k \Sigma_k^{-1} S_k^\intercal (X_{1,T} - \Xi_{0,T})^\intercal + \Sigma_\xi - \Sigma_{k+1} \preceq 0$$

$$\Sigma_k - X_{0,T} S_k = 0$$

$$S_k \Sigma_k^{-1} S_k^\intercal - Y_k \preceq 0$$

Robust DD-DS

Robust DD Mean Steering


2020 59th IEEE Conference on Decision and Control (CDC)
Jeju Island, Republic of Korea, December 14-18, 2020

Robust Closed-loop Model Predictive Control via System Level Synthesis

Shaoru Chen*, Han Wang*, Manfred Morari, Victor M. Preciado, Nikolai Matni

Abstract—**In this paper, we consider the robust closed-loop model predictive control (MPC) of a linear time-variant (LTV) system with norm bounded disturbances and LTV model uncertainty, wherein a series of constrained optimal control problems (OCPs) are solved. Guaranteeing robust feasibility of these OCPs is challenging due to disturbances perturbing the predicted states, and model uncertainty, both of which can render the closed-loop system unstable. As such, a trade-off between the numerical tractability and conservativeness of the solutions is often required. We use the System Level Synthesis (SLS) framework to reformulate these constrained OCPs over closed-loop system responses, and show that this allows us to transparently account for norm bounded additive disturbances and LTV model uncertainty by computing robust state feedback policies. We further show that by exploiting the underlying linear fractional structure of the resulting robust OCPs, we can significantly reduce the conservativeness of existing SLS-based and tube-MPC-base[illegible]trol m[illegible]ds while also improving computationa[illegible]ciency. We c[illegible]de with numerical examples d[illegible] the effective[illegible] of our methods.**

I. INTRODUCTION

Model predictive control (MP[illegible]) has a[illegible]ed r[illegible]k-able success in solving multivar[illegible]ained problems across a wide range of application areas, such as process control [1], power networks [2], and robot locomotion [3]. In MPC, a control action is computed by solving a finite-horizon constrained optimal control problem (OCP) at each sampling time, and then applying the first control action. The stability and performance of MPC depends on the accuracy of the model being used, and indeed robustness to both additive disturbances and model uncertainty must be considered. Although MPC using a nominal model (i.e., one ignoring uncertainty) offers some level of robustness [4], it has been shown that the closed-loop system achieved by nominal MPC can be destabilized by an arbitrarily small disturbance [5]. As a result, robust MPC, which explicitly deals with uncertainty, has received much attention [6].

When only additive disturbances are present, open-loop robust MPC, which optimizes over a sequence of control actions $\mathbf{u} = \{u_0, \cdots, u_{N-1}\}$ subject to suitably robust constraints, can be applied but tends to be overly conservative or even infeasible [7], as a single sequence of inputs $\mathbf{u}$ is chosen for all possible disturbance realizations. On the other hand, closed-loop MPC, which optimizes over the control policies $\pi = \{\pi_0(\cdot), \cdots, \pi_{N-1}(\cdot)\}$, can reduce the conservativeness of the solutions. However, the policy space is infinite-dimensional and renders the online OCPs intractable. The problem can be rendered tractable by restricting the policies π to lie in a function class that admits a finite-dimensional parameterization. For example, policies of the form $\pi_i(x) = Kx + v_i$ are considered in [8–10], where K is a pre-stabilizing feedback gain K that is fixed beforehand, thus reducing the decision variables to the vectors $\{v_1, \cdots, v_{[illegible]}\}$. To reduce conservativeness, an affi[illegible]k cont[illegible]aw $\pi(x) = K_i x + v_i$ can be appli[illegible]ith de[illegible]sion v[illegible]bles K_i and v_i; however, the resul[illegible]OCP is non-co[illegible]. In [11], it was observed that by re-[illegible]rizing the [illegible]Ps to be over disturbance based feedback [illegible]of the [illegible]n $\pi_i(w) = \sum_{j=0}^{i-1} M_{ij} w_j + v_i$, the resulting [illegible] are [illegible]vex. In [12, 13], the authors prop[illegible]e an al[illegible]tive (t[illegible]MPC) invariant set based ap-[illegible]a[illegible]ounds th[illegible]stem trajectories within a tube that robustly satisfy constraints.

The more challenging problem considering model uncertainty is tackled in [8, 13, 14]. When polytopic or structured feedback model uncertainty occurs, a linear matrix inequality (LMI) based robust MPC method is proposed in [14]. When both model uncertainty and additive disturbances are present, the method proposed in [13] designs tubes containing all possible trajectories under polytopic uncertainty assumptions. Alternative approaches based on dynamic programming (DP) [15] are shown to obtain tight solutions, but the computation quickly becomes intractable. Adaptive robust MPC, which considers estimation of the parametric uncertainty while implementing robust control, is proposed in [16–18].

As described above, there is a rich body of work addressing the robust MPC problem, and it remains an active area of research for which no definitive solution exists. Due to the inherent intractability of the general robust MPC problem subject to both additive disturbance and model uncertainty, all of the aforementioned methods trade off conservativeness for computational tractability in different ways. The recently developed *System Level Synthesis* (SLS) parameterization [19] provides an alternative approach to tackling the robust MPC problem and exploring this tradeoff

Han Wang is with the Department of Applied Mathematics and Computational Science, University of Pennsylvania, Philadelphia, PA, 19104, USA (e-mail: wanghan2@sas.upenn.edu).
Shaoru Chen, Manfred Morari, Victor M. Preciado and Nikolai Matni are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, 19104, USA (e-mail: {srchen, morari, preciado, nmatni}@seas.upenn.edu).
* The first two authors contributed equally to this paper.

978-1-7281-7447-1/20/$31.00 ©2020 IEEE 2152

Authorized licensed use limited to: Georgia Institute of Technology. Downloaded on October 14,2025 at 22:19:11 UTC from IEEE Xplore. Restrictions apply.

Robust DD Covariance Steering

Rewrite perturbation LMI as

$$\Delta G_k^{\Sigma} = \Theta^{\mathsf{T}}(S_k)\Delta\Xi_{0,T}^{\mathsf{T}}\Pi + \Pi^{\mathsf{T}}\Delta\Xi_{0,T}\Theta(S_k)$$

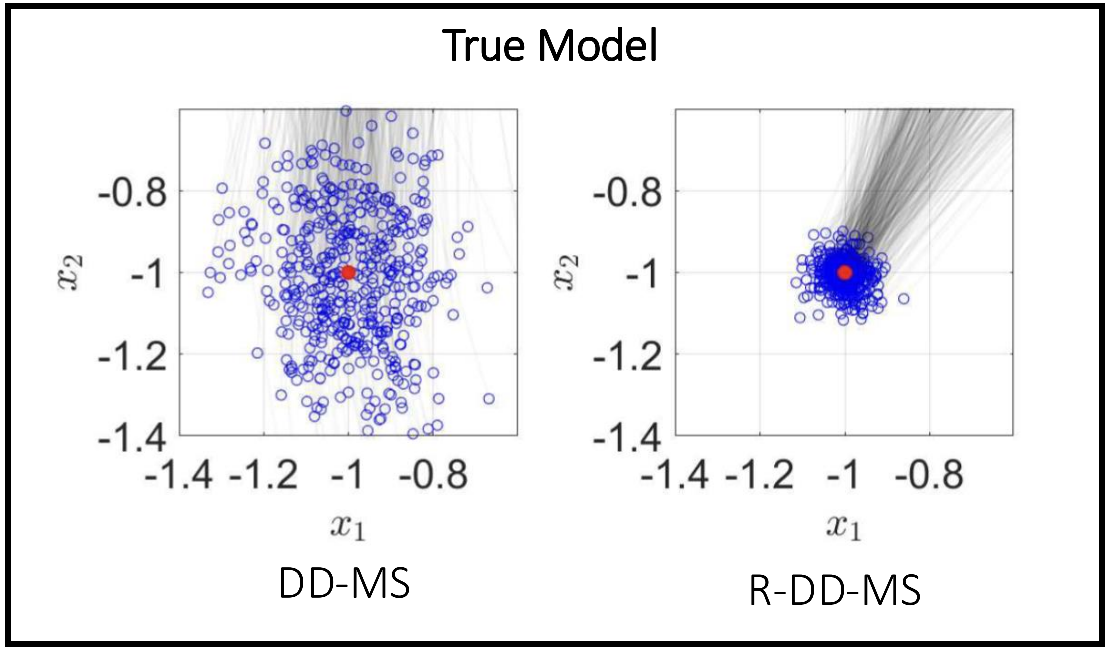
where,

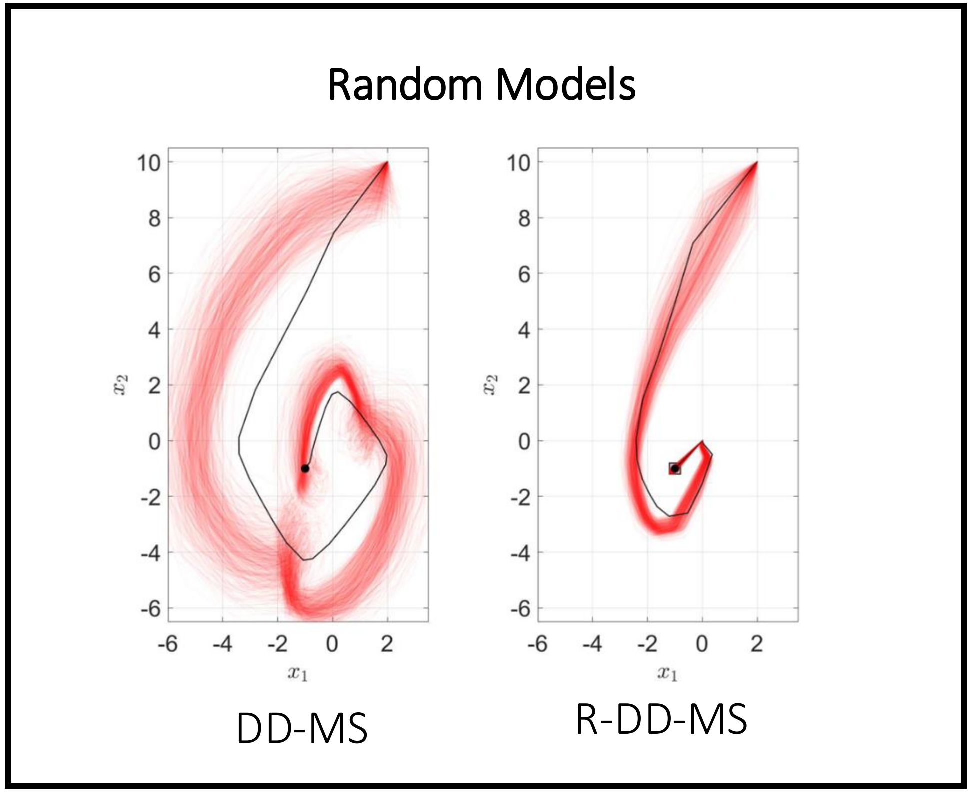
$$\Theta(S_k) = \begin{bmatrix} 0 \\ -S_k \end{bmatrix}, \qquad \Pi = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

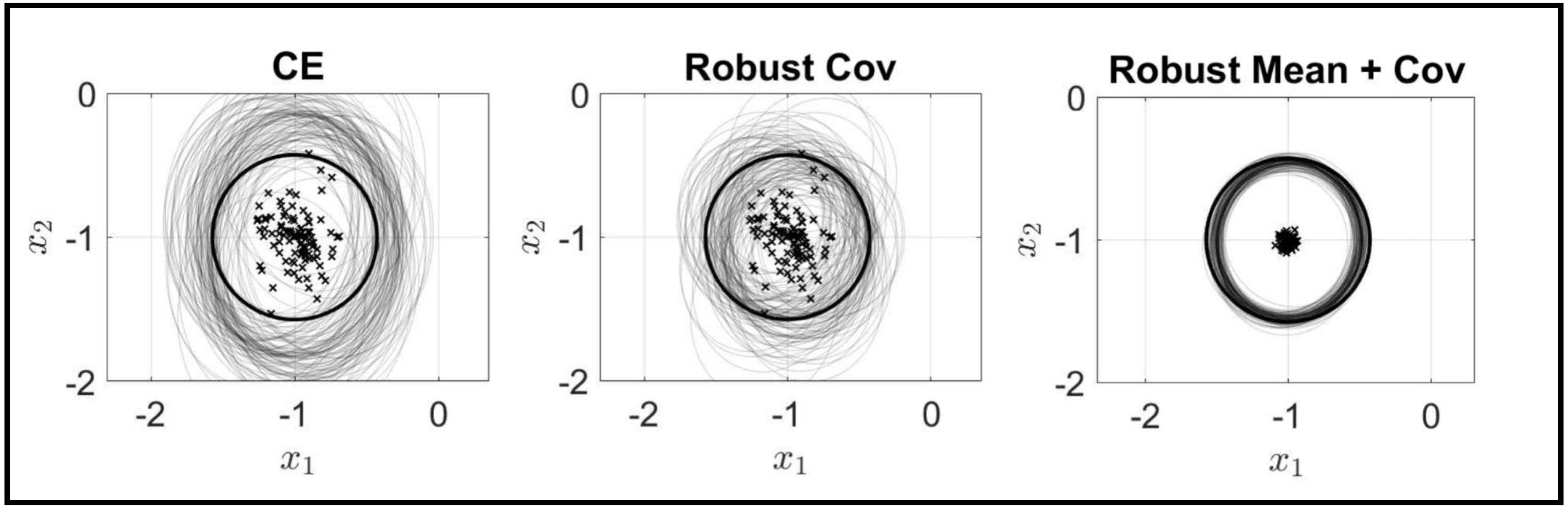
Using standard results from RO (e.g, A. Ben-Tal, L.E. Ghaoui, A. Nemirovski, 2009)

$$\begin{bmatrix} \lambda I & \rho(\delta)\,\Theta(S_k) \\ \rho(\delta)\,\Theta^{\mathsf{T}}(S_k) & \hat{G}_k^{\Sigma}(\Sigma_k, \Sigma_{k+1}, S_k) - \lambda\Pi^{\mathsf{T}}\Pi \end{bmatrix} \succeq 0$$

Example

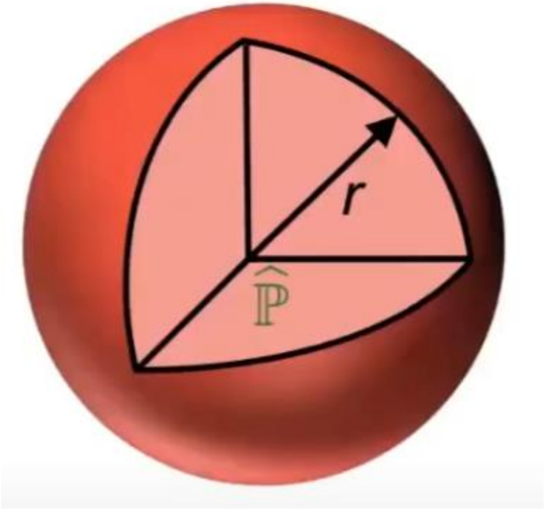

True Model
-0.8
-1
-1.2
-1.4
-1.4 -1.2 -1 -0.8
x_1
x_2
DD-MS
R-DD-MS


Random Models
10
8
6
4
2
0
-2
-4
-6
-6 -4 -2 0 2
x_1
x_2
DD-MS
R-DD-MS


CE
Robust Cov
Robust Mean + Cov
0
-1
-2
-2 -1 0
x_1
x_2

Distributionally Robust CS

Nominal Distribution



$$\mathbb{P}_w = \mathcal{N}(0, I)$$

Wasserstein Ambiguity Set

$$\mathbb{B}^c_{\epsilon,p}(\mathbb{P}) = \{\mathbb{Q} \in \mathcal{S} : \mathbb{W}^c_p(\mathbb{Q}, \mathbb{P}) \leq \varepsilon\}$$

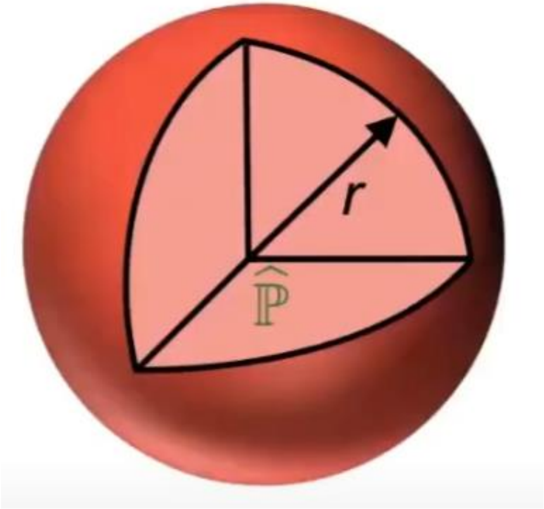
$$\mathbb{W}^c_p(\mathbb{P}, \mathbb{P}') \triangleq \left(\inf_{\pi \in \Pi(\mathbb{P}, \mathbb{P}')} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(\xi, \xi')^p \, \pi(\mathrm{d}\xi, \mathrm{d}\xi') \right)^{\frac{1}{p}}$$

Joint distributions with marginals $\mathbb{P}, \mathbb{P}'$

Transportation Cost

J. Pilipovsky, and P. Tsiotras, "Distributionally Robust Density Control with Wasserstein Ambiguity Sets", *63th IEEE Conference on Decision and Control*, Milan, Italy, Dec. 16–19, 2024, pp. 1081-1086.

Distributionally Robust CS



$$\mathbb{P}_w = \mathcal{N}(0, I)$$

Wasserstein Ambiguity Set

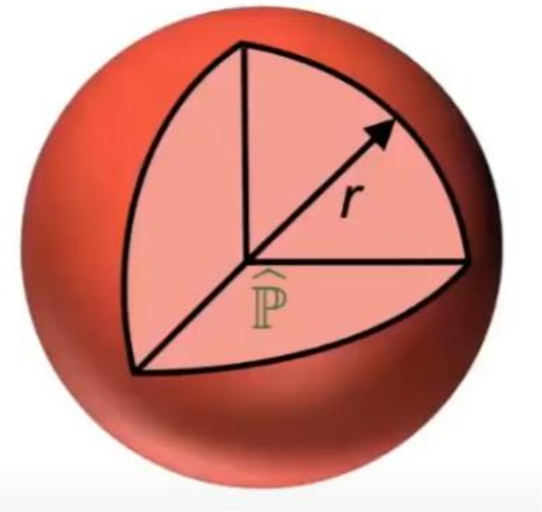
$$\mathbb{B}^c_{\epsilon,p}(\mathbb{P}) = \{\mathbb{Q} \in \mathcal{S} : \mathbb{W}^c_p(\mathbb{Q}, \mathbb{P}) \leq \varepsilon\}$$

$$\mathbb{W}^c_p(\mathbb{P}, \mathbb{P}') \triangleq \left(\inf_{\pi \in \Pi(\mathbb{P}, \mathbb{P}')} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(\xi, \xi')^p \, \pi(\mathrm{d}\xi, \mathrm{d}\xi') \right)^{\frac{1}{p}}$$

GOAL: Ensure robust performance of the system under a range of uncertainties that can potentially affect the system

J. Pilipovsky, and P. Tsiotras, "Distributionally Robust Density Control with Wasserstein Ambiguity Sets", *63th IEEE Conference on Decision and Control*, Milan, Italy, Dec. 16–19, 2024, pp. 1081-1086.

Distributionally Robust CS



$$\mathbb{P}_w = \mathcal{N}(0, I)$$

Wasserstein Ambiguity Set

$$\mathbb{B}^c_{\epsilon,p}(\mathbb{P}) = \{\mathbb{Q} \in \mathcal{S} : \mathbb{W}^c_p(\mathbb{Q}, \mathbb{P}) \leq \varepsilon\}$$

$$\mathbb{W}^c_p(\mathbb{P}, \mathbb{P}') \triangleq \left(\inf_{\pi \in \Pi(\mathbb{P}, \mathbb{P}')} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(\xi, \xi')^p \, \pi(\mathrm{d}\xi, \mathrm{d}\xi') \right)^{\frac{1}{p}}$$

APPROACH: Steer the state distribution of a dynamical system subject to partially known uncertainties

J. Pilipovsky, and P. Tsiotras, "Distributionally Robust Density Control with Wasserstein Ambiguity Sets", *63th IEEE Conference on Decision and Control*, Milan, Italy, Dec. 16–19, 2024, pp. 1081-1086.

Problem Statement

GOAL: Steer the ambiguity set of the state to a prescribed, terminal ambiguity set $\mathbb{S}_f \triangleq \mathbb{B}_\delta^{\|\cdot\|}(\mathbb{P}_f)$ while minimizing the distributionally-robust cost

$$\mathcal{W} = \mathbb{B}_\varepsilon^{\|\cdot\|}(\hat{\mathbb{P}}_w)$$

$$\mathcal{J} = \beta \sum_{k=0}^{N-1} \|\bar{u}_k\| + \max_{\mathbb{P}\in\mathcal{W}} \mathbb{E}_{\mathbb{P}} \left[\sum_{k=0}^{N-1} \tilde{\boldsymbol{x}}_k^\intercal Q_k \tilde{\boldsymbol{x}}_k + \tilde{\boldsymbol{u}}_k^\intercal R_k \tilde{\boldsymbol{u}}_k \right]$$

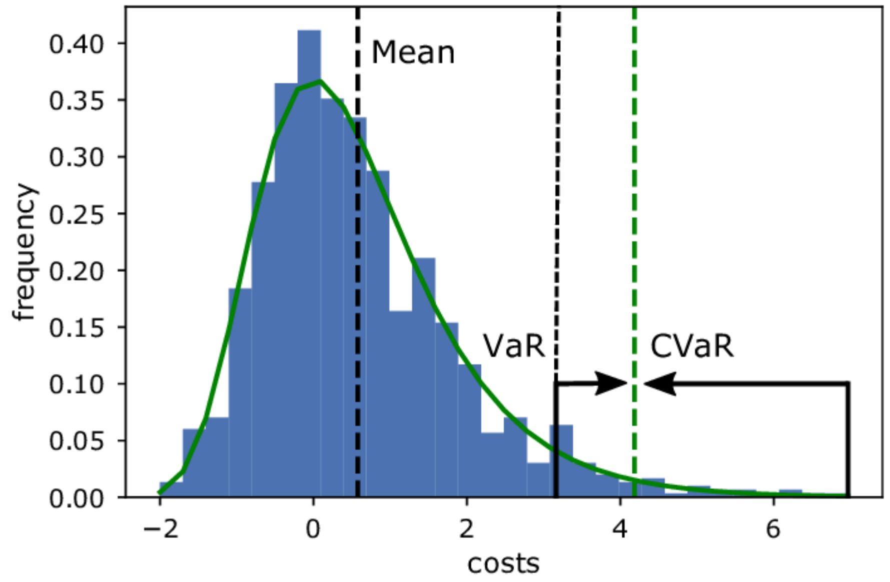
Enforce state constraints along the planning horizon

$$\mathcal{X} \triangleq \{x : \max_{j\in[J]} \alpha_j^\intercal x + \beta_j \leq 0\}$$

Probability of violating the constraints is less than γ

Why CVaR?

- Convex
- Penalizes worst-case violations ("black swan")
- Implicitly satisfies VaR



Problem Statement

GOAL: Steer the ambiguity set of the state to a prescribed, terminal ambiguity set $\mathbb{S}_f \triangleq \mathbb{B}_\delta^{\|\cdot\|}(\mathbb{P}_f)$ while minimizing the distributionally-robust cost

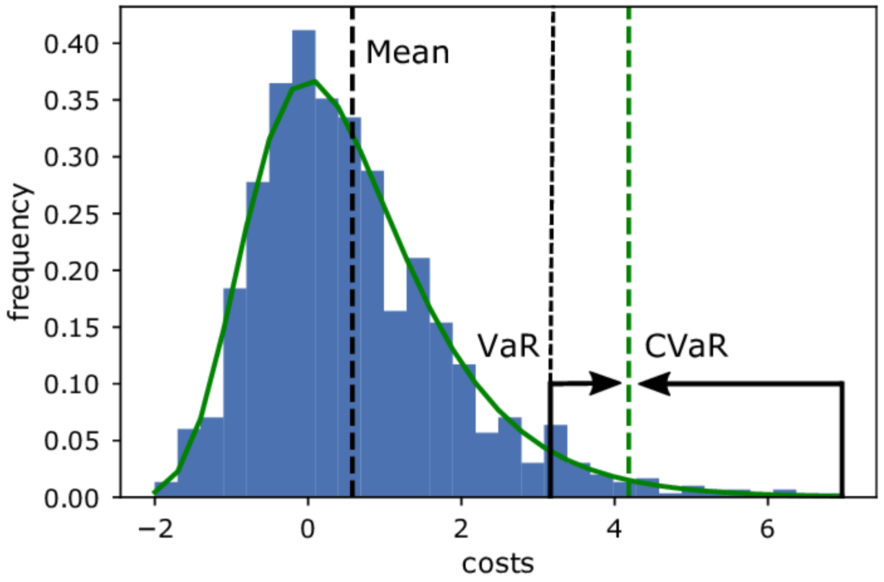
$$\mathcal{J} = \beta \sum_{k=0}^{N-1} \|\bar{u}_k\| + \max_{\mathbb{P}\in\mathcal{W}} \mathbb{E}_{\mathbb{P}} \left[\sum_{k=0}^{N-1} \tilde{\boldsymbol{x}}_k^\intercal Q_k \tilde{\boldsymbol{x}}_k + \tilde{\boldsymbol{u}}_k^\intercal R_k \tilde{\boldsymbol{u}}_k \right]$$

Enforce the distributionally-robust CVaR (DR-CVaR) constraints

$$\sup_{\mathbb{P}_k \in \mathbb{S}_k} \mathrm{CVaR}_{1-\gamma}^{\mathbb{P}_k} \left(\max_{j\in[J]} \alpha_j^\intercal x_k + \beta_j \right) \leq 0$$

Why CVaR?

- Convex
- Penalizes worst-case violations (“black swan”)
- Implicitly satisfies VaR



Problem Formulation

Consider augmented system

$$\boldsymbol{x} = \mathcal{A}x_0 + \mathcal{B}\boldsymbol{u} + \mathcal{D}\boldsymbol{w},$$

with control law

$$u_k = v_k + K_k \tilde{x}_k$$

$$\bar{\boldsymbol{x}} = \mathcal{A}x_0 + \mathcal{B}\boldsymbol{v}$$

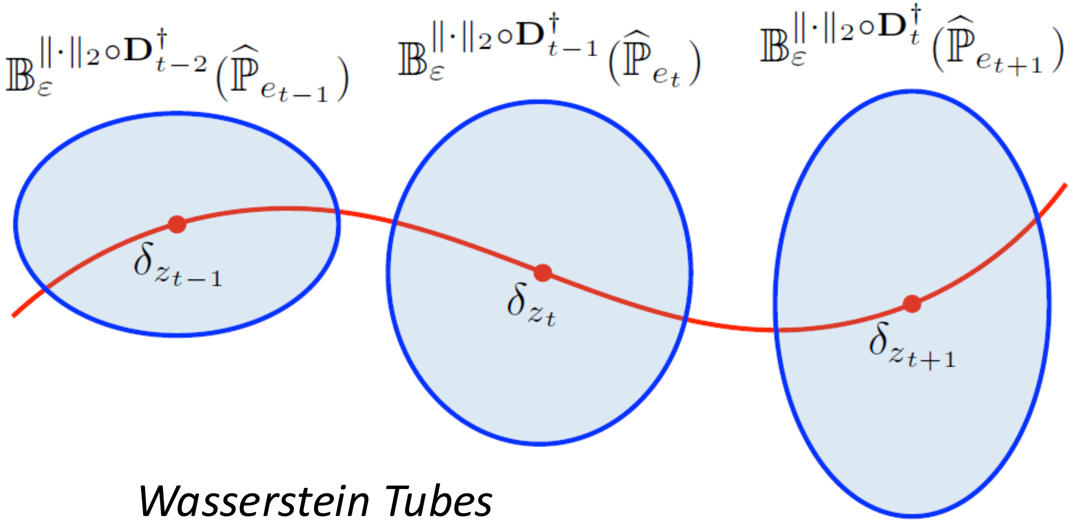
$$\tilde{\boldsymbol{x}} = (I + \mathcal{B}L)\mathcal{D}\boldsymbol{w}$$

New decision variable
$K = L(I + \mathcal{B}L)^{-1}$

Distributional state uncertainty at *k*

$$\mathbb{S}_k = \delta_{\mathcal{A}x_0 + \mathcal{B}\boldsymbol{v}} * \mathbb{B}_{\varepsilon}^{\|\cdot\| \circ \tilde{L}_k^{\dagger}}((\tilde{L}_k)_{\#}\hat{\mathbb{P}}_w)$$

$$(\delta_x * \mathbb{P})(\mathcal{B}) = \mathbb{P}(\mathcal{B} - x)$$



Wasserstein Tubes

Problem Formulation

Consider augmented system

$$\boldsymbol{x} = \mathcal{A}x_0 + \mathcal{B}\boldsymbol{u} + \mathcal{D}\boldsymbol{w},$$

with control law

$$u_k = v_k + K_k \tilde{x}_k$$

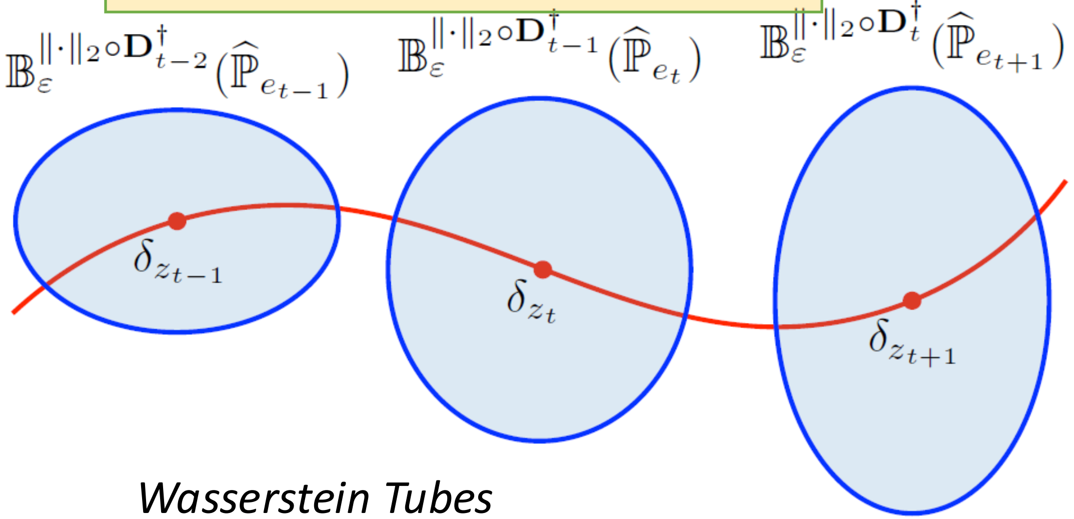
$$\bar{\boldsymbol{x}} = \mathcal{A}x_0 + \mathcal{B}\boldsymbol{v}$$

$$\tilde{\boldsymbol{x}} = (I + \mathcal{B}L)\mathcal{D}\boldsymbol{w}$$

Distributional state uncertainty at *k*

$$\mathbb{S}_k = \delta_{\mathcal{A}x_0 + \mathcal{B}\boldsymbol{v}} * \mathbb{B}_\varepsilon^{\|\cdot\| \circ \tilde{L}_k^\dagger}\left((\tilde{L}_k)_\# \hat{\mathbb{P}}_w\right)$$

$$\tilde{L}_k \triangleq E_k(I + \mathcal{B}L)\mathcal{D}$$

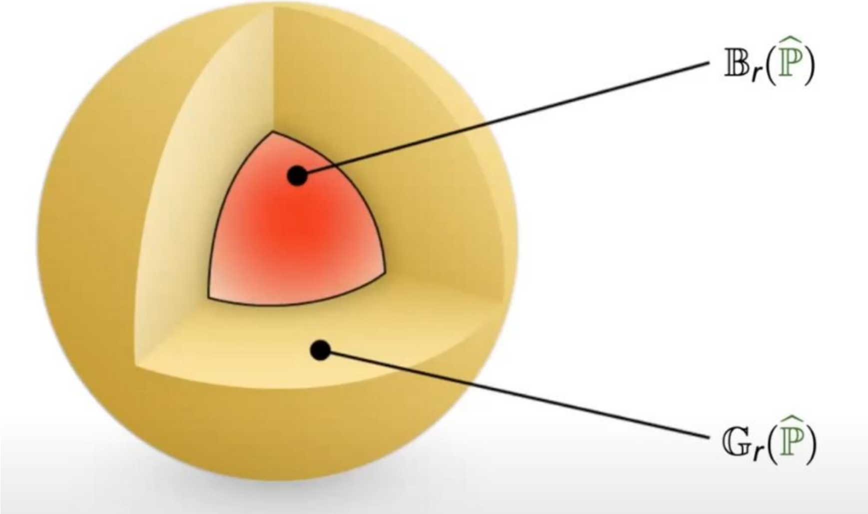


Wasserstein Tubes

Key Result: Propagation of uncertainty through stochastic LTI systems (L. Aolaritei, N. Lanzetti, H. Chen, and F. Dörfler, 2023)

$$A_\# \mathbb{B}_\varepsilon^c(\mathbb{P}) = \mathbb{B}_\varepsilon^{c \circ A^\dagger}(A_\# \mathbb{P})$$

Gelbrich Ambiguity Set



$$\mathcal{G}_\varepsilon(\mu, \Sigma) = \{\mathbb{Q} \in \mathcal{S} : (\mathbb{E}_\mathbb{Q}[\xi], \mathrm{Cov}_\mathbb{Q}[\xi]) \in \mathcal{U}_\varepsilon(\mu, \Sigma)\}$$

$$\mathcal{U}_\varepsilon(\hat{\mu}, \hat{\Sigma}) = \{(\mu, \Sigma) \in \mathbb{R}^d \times \mathbb{S}_+^d : \mathbb{G}((\mu, \Sigma), (\hat{\mu}, \hat{\Sigma})) \leq \varepsilon\}$$

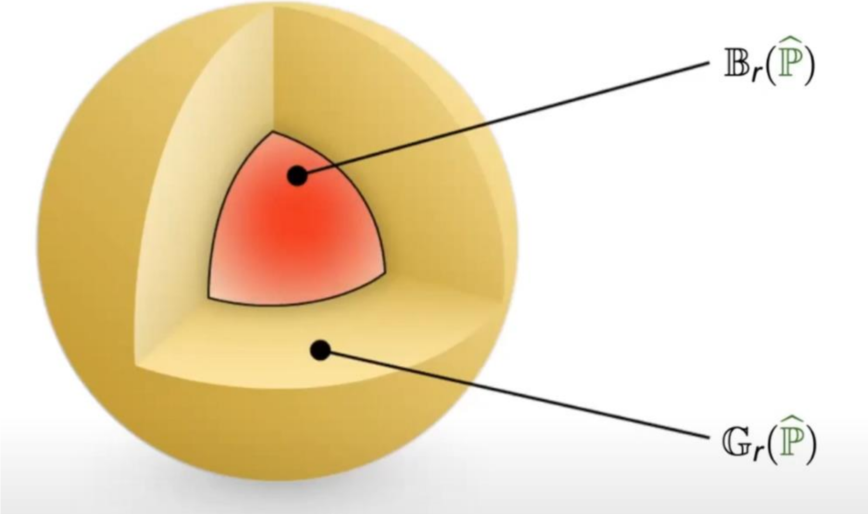
$$\mathbb{G}^2((\mu_1, \Sigma_1), (\mu_2, \Sigma_2)) \triangleq \|\mu_1 - \mu_2\|^2 + \mathrm{tr}\left[\Sigma_1 + \Sigma_2 - 2(\Sigma_2^{\frac{1}{2}} \Sigma_1 \Sigma_2^{\frac{1}{2}})^{\frac{1}{2}}\right]$$

$$\sup_{\mathbb{P}_k \in \mathcal{G}_k} \mathrm{CVaR}_{1-\gamma}^{\mathbb{P}_k}\left(\max_{j \in [J]} \alpha_j^\intercal x_k + \beta_j\right) \leq 0 \implies \sup_{\mathbb{P}_k \in \mathbb{S}_k} \mathrm{CVaR}_{1-\gamma}^{\mathbb{P}_k}\left(\max_{j \in [J]} \alpha_j^\intercal x_k + \beta_j\right) \leq 0$$

Gelbrich Ambiguity Set

$$\mathcal{G}_\varepsilon(\mu, \Sigma) = \{\mathbb{Q} \in \mathcal{S} : (\mathbb{E}_\mathbb{Q}[\xi], \mathrm{Cov}_\mathbb{Q}[\xi]) \in \mathcal{U}_\varepsilon(\mu, \Sigma)\}$$

$$\mathcal{U}_\varepsilon(\hat{\mu}, \hat{\Sigma}) = \{(\mu, \Sigma) \in \mathbb{R}^d \times \mathbb{S}^d_+ : \mathbb{G}((\mu, \Sigma), (\hat{\mu}, \hat{\Sigma})) \leq \varepsilon\}$$



$$\mathbb{G}^2((\mu_1, \Sigma_1), (\mu_2, \Sigma_2)) \triangleq \|\mu_1 - \mu_2\|^2 + \mathrm{tr}\left[\Sigma_1 + \Sigma_2 - 2(\Sigma_2^{\frac{1}{2}} \Sigma_1 \Sigma_2^{\frac{1}{2}})^{\frac{1}{2}}\right]$$

Satisfaction of Gelbrich DR-CVaR constraints implies satisfaction of Wasserstein DR-CVaR constraints

Chance Constraints

The **individual** DR-CVAR constraints

$$\sup_{\mathbb{P}_k \in \mathbb{S}_k} \mathrm{CVaR}^{\mathbb{P}_k}_{1-\gamma_{jk}}(\alpha_j^\intercal x_k + \beta_j) \leq 0,$$

are satisfied if the following **convex** constraints are satisfied

$$\beta_j + \alpha_j^\intercal \hat{\mu}_k(v) + \tau_{jk}\sqrt{\alpha_j^\intercal \hat{\Sigma}_k(\tilde{L}_k)\alpha_j} + \tilde{\varepsilon}_k(\tilde{L}_k)\|\alpha_j\| \leq 0$$

where,

$$\hat{\Sigma}_k \triangleq \tilde{L}_k \Sigma_w \tilde{L}_k^\intercal \qquad \tilde{\varepsilon}_k \triangleq \varepsilon(1+\tau_{jk}^2)^{1/2}\sigma_{\max}^2(\tilde{L}_k) \qquad \tau_{jk} \triangleq \sqrt{\frac{1-\gamma_{jk}}{\gamma_{jk}}}$$

J. Pilipovsky, and P. Tsiotras, "Distributionally Robust Density Control with Wasserstein Ambiguity Sets", *63th IEEE Conference on Decision and Control*, Milan, Italy, Dec. 16–19, 2024, pp. 1081-1086.

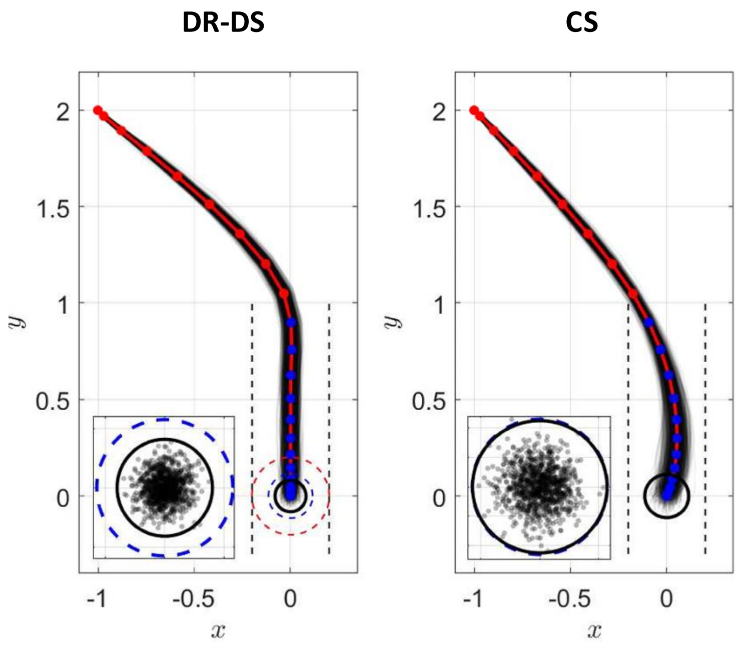
Chance Constraints

Similarly, the DR Objective and terminal ambiguity set constraint can also be formulated in terms of LMIs

Using Schur complement we can further reformulate these constraints as tractable second-order cone constraints (SOCC) and linear matrix inequalities (LMIs)

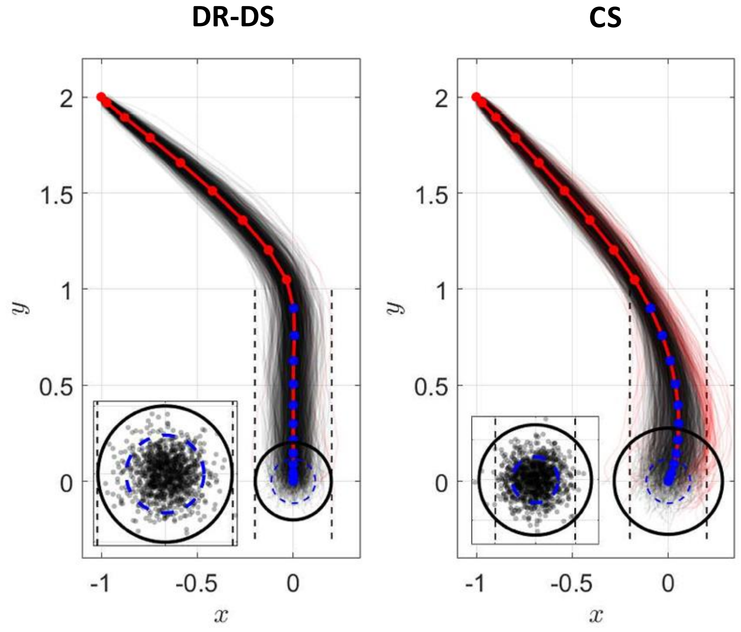
J. Pilipovsky, and P. Tsiotras, "Distributionally Robust Density Control with Wasserstein Ambiguity Sets", *63th IEEE Conference on Decision and Control*, Milan, Italy, Dec. 16–19, 2024, pp. 1081-1086.

Performance Comparison – 2D System



Nominal disturbance

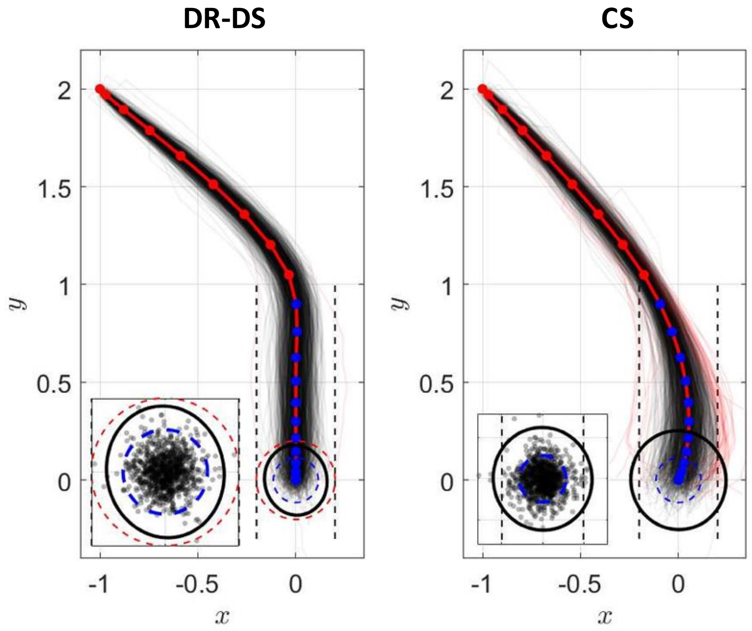
$$\mathbb{P}_w = \mathcal{N}(0, I)$$



Extreme Gaussian disturbance

$$w_k \sim \mathcal{N}(0, \eta^2 I)$$

$$\mathbb{W}(\mathbb{P}_w, \hat{\mathbb{P}}_w) = \varepsilon$$



Non-Gaussian disturbance
(3-DOF *t-distribution*)

Take-Aways

- Directly controlling **distributions of trajectories** leads to strict performance guarantees
- Uncertainty Synthesis
    - Control of system with uncertainty
    - Control of uncertainty
- Many, many, applications
    - Pinpoint landing
    - Swarms, Ensemble control
- For linear systems with Gaussian noise, theory well-developed
- **Have extended CS theory to**
    - **Unknown system matrices**
    - **Unknown noise statistics**