Optimality of Linear Policies in Distributionally Robust Linear Quadratic Control

Bahar Taşkesen, Dan A. Iancu, Çağıl Koçyiğit, Daniel Kuhn

Sucheth Shenoy

December 19, 2025

Outline

Outline

Why Distributionally Robust LQG?

- **Classical LQG** assumes exact knowledge of disturbance distributions.
- In practice, even small distributional mismatch can cause significant performance degradation.

- **Distributionally Robust Optimization (DRO)** replaces a single assumed distribution with an *ambiguity set* of plausible distributions.
- The controller minimizes cost against the **worst-case distribution** in this set.

- This leads to a **min–max problem**:
  - Adversary: Nature chooses a probability distribution.
  - Decision Maker: Controller chooses a control policy.

Central Questions

What is the structure of the worst-case disturbance distribution?
Are linear policies still optimal under distributional uncertainty?

Overview of Contributions

- **Finite-dimensional** formulation of the min-max problem.
- Ambiguity sets induced by common divergence measures, including **2-Wasserstein distance**, **Kullback–Leibler divergence**, and **moment-based divergences**.
- For a **finite-horizon problem** with **zero-mean Gaussian nominal noise**:
  - Optimal control policy: **affine in the observations**,
  - Worst-case distribution: **Gaussian**.
- Under a **weak condition** satisfied by the considered divergence-based ambiguity sets:
  - Optimal control policy: **linear**,
  - Worst-case distribution: **zero-mean Gaussian**.
- Adversary optimally **inflates the noise** by choosing covariance matrices that **dominate the nominal covariance in Loewner order**.
- Develop an efficient **Frank–Wolfe algorithm**, where each inner step solves a standard **LQG subproblem** via Kalman filtering and dynamic programming.
- Extend all results to the **infinite-horizon, average-cost** setting:
  - Optimal control policy: **stationary and linear**,
  - Worst-case distribution: **time-invariant, zero-mean Gaussian**.

Classical LQG: Assumptions and Separation Principle

System dynamics (finite horizon, discrete time):

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \qquad \forall t \in \{0, \dots, T-1\} \tag{1}$$

$$y_t = C_t x_t + v_t, \qquad \forall t \in \{0, \dots, T-1\} \tag{2}$$

Stochastic assumptions:

- Initial state: $x_0 \sim \mathcal{N}(\mu_0, \Sigma_0)$.
- Process noise: $w_t \sim \mathcal{N}(0, W_t)$, i.i.d.
- Measurement noise: $v_t \sim \mathcal{N}(0, V_t)$, i.i.d.
- x_0, $\{w_t\}$, and $\{v_t\}$ are mutually independent.
- Noise covariances W_t, V_t are known exactly.

Quadratic cost:

$$J = \mathbb{E}_{\mathbb{P}}\left[\sum_{t=0}^{T-1}\left(x_t^\top Q_t x_t + u_t^\top R_t u_t\right) + x_T^\top Q_T x_T\right] \tag{3}$$

Separation Principle:

- Optimal control can be decomposed into:
  - **State estimation** via a Kalman filter,
  - **State-feedback control** using the estimate.
- Estimation and control problems can be solved **independently**.

Classical LQG: Optimal Controller

Optimal control policy (finite horizon):

$$u_t = K_t \hat{x}_t, \qquad \hat{x}_t = \mathbb{E}_{\mathbb{P}}[x_t \mid y_{0:t}], \qquad \forall t \in \{0, \ldots, T-1\}. \tag{4}$$

Riccati recursion (control, backward in time):

$$\begin{aligned}
P_T &= Q_T, \\
P_t &= Q_t + A_t^\top P_{t+1} A_t - A_t^\top P_{t+1} B_t \left(R_t + B_t^\top P_{t+1} B_t \right)^{-1} B_t^\top P_{t+1} A_t, \\
K_t &= -\left(R_t + B_t^\top P_{t+1} B_t \right)^{-1} B_t^\top P_{t+1} A_t, \qquad \forall t \in \{0, \ldots, T-1\}.
\end{aligned} \tag{5}$$

Key property (certainty equivalence):

- Control gains K_t depend only on (A_t, B_t, Q_t, R_t).
- The controller acts on the state estimate $\hat{x}_t$, not the true state.

Classical LQG: Kalman Filter (Observer)

State estimate recursion:

$$\begin{aligned} \hat{x}_{0|-1} &= \mu_0, \\ \hat{x}_{t|t-1} &= A_{t-1}\hat{x}_{t-1} + B_{t-1}u_{t-1}, \qquad \forall t \in \{1, \dots, T-1\}, \\ \hat{x}_t &= \hat{x}_{t|t-1} + L_t\big(y_t - C_t\hat{x}_{t|t-1}\big), \qquad \forall t \in \{0, \dots, T-1\}. \end{aligned} \tag{6}$$

Kalman Filter (observer, forward in time):

$$\begin{aligned} \Sigma_{0|-1} &= \Sigma_0, \\ \Sigma_{t+1|t} &= A_t\Sigma_{t|t}A_t^\top + W_t, \qquad \forall t \in \{0, \dots, T-1\}, \\ \Sigma_{t|t} &= \Sigma_{t|t-1} - \Sigma_{t|t-1}C_t^\top\left(C_t\Sigma_{t|t-1}C_t^\top + V_t\right)^{-1}C_t\Sigma_{t|t-1}, \\ &\quad \forall t \in \{0, \dots, T-1\}. \end{aligned} \tag{7}$$

Kalman gain:

$$L_t = \Sigma_{t|t-1}C_t^\top\left(C_t\Sigma_{t|t-1}C_t^\top + V_t\right)^{-1}, \quad \forall t \in \{0, \dots, T-1\}. \tag{8}$$

Outline

Ambiguity Model and SMO Condition

System dynamics: Defined by Eqs. (1), (2) (discrete-time linear dynamics).

Distributional ambiguity: The joint distribution $\mathbb{P}$ of all exogenous uncertainties $z \in \mathcal{Z}$

$$\mathcal{Z} = \{x_0, w_0, \ldots, w_{T-1}, v_0, \ldots, v_{T-1}\}$$

is unknown and assumed to belong to an ambiguity set $\mathcal{B}$:

$$\mathcal{B} = \left\{\mathbb{P} \in \mathcal{P}(\mathbb{R}^{n+T(n+p)}) : \mathbb{P}_z \in \mathcal{B}_z,\ \mathbb{E}_{\mathbb{P}}[zz'^{\top}] = 0\ \forall z \neq z' \in \mathcal{Z}\right\}. \tag{9}$$

Marginal ambiguity sets: For each component $z \in \mathcal{Z}$,

$$\mathcal{B}_z = \left\{\mathbb{P}_z \in \mathcal{P}(\mathbb{R}^{d_z}) : \mathbb{D}(\mathbb{P}_z, \hat{\mathbb{P}}_z) \leq \rho_z\right\}, \tag{10}$$

where:

- $\hat{\mathbb{P}}_z$: nominal marginal distribution, $\mathbb{P}_z$: marginal distribution of $z \in \mathcal{Z}$,
- $\mathbb{D}(\cdot,\cdot)$: divergence (e.g. Wasserstein, KL, moment-based),
- $\rho_z \geq 0$: radius controlling uncertainty size.

Second Moment Orthogonality (SMO) condition:

- $\mathbb{E}_{\mathbb{P}}[zz'^{\top}] = 0 \quad \forall z \neq z' \in \mathcal{Z}$.
- If $\mathbb{E}_{\mathbb{P}}[z] = \mathbb{E}_{\mathbb{P}}[z'] = 0$, then z and z' are uncorrelated (but not necessarily independent).

Assumptions for Tractability

Assumption 1 (Nominal distribution $\hat{\mathbb{P}}$).

- The nominal distribution $\hat{\mathbb{P}}$ is Gaussian.
- The nominal means satisfy: $\hat{\mu}_z = 0 \quad \forall z \in \mathcal{Z}$.
- The nominal measurement noise covariances satisfy: $\hat{\Sigma}_{v_t} \succ 0 \quad \forall t \in [0, T-1]$.

Assumption 2 (Ambiguity Set $\mathcal{B}$). The divergence $\mathbb{D}$ satisfies the following properties:

1. For every $(\mu_z, M_z) \in \mathcal{M}_2^{d_z}$, a Gaussian distribution minimizes the divergence from $\hat{\mathbb{P}}_z$ among all distributions with the same first and second moments:

$$\mathbb{D}\big(\mathcal{N}(\mu_z, M_z), \hat{\mathbb{P}}_z\big) = \inf_{\mathbb{P}_z \in \mathcal{P}(\mathbb{R}^{d_z})} \mathbb{D}(\mathbb{P}_z, \hat{\mathbb{P}}_z).$$

$$s.t. \quad \mathbb{E}_{\mathbb{P}_z}[z] = \mu_z, \ \mathbb{E}_{\mathbb{P}_z}[zz^\top] = M_z$$

2. The set

$$\mathcal{M}(\mu_z, M_z) = \Big\{(\mu_z, M_z) \in \mathcal{M}_2^{d_z} : \mathbb{D}\big(\mathcal{N}(\mu_z, M_z), \hat{\mathbb{P}}_z\big) \leq \rho_z\Big\}$$

is convex and compact.

Outline

Wasserstein Ambiguity Sets

The **2-Wasserstein distance** between two distributions $\mathbb{P}_z, \hat{\mathbb{P}}_z \in \mathcal{P}(\mathbb{R}^{d_z})$ is defined as

$$W(\mathbb{P}_z, \hat{\mathbb{P}}_z) = \left(\inf_{\pi \in \Pi(\mathbb{P}_z, \hat{\mathbb{P}}_z)} \int_{\mathbb{R}^{d_z} \times \mathbb{R}^{d_z}} \|z - \hat{z}\|_2^2 \, d\pi(z, \hat{z}) \right)^{1/2}, \tag{11}$$

where $\Pi(\mathbb{P}_z, \hat{\mathbb{P}}_z)$ denotes the set of all couplings of $\mathbb{P}_z$ and $\hat{\mathbb{P}}_z$.

Gelbrich distance: For distributions with first and second moments (μ_z, M_z) and $(\hat{\mu}_z, \hat{M}_z)$,

$$W(\mathbb{P}_z, \hat{\mathbb{P}}_z) \;\geq\; G\Big((\mu_z, \Sigma_z), (\hat{\mu}_z, \hat{\Sigma}_z)\Big), \tag{12}$$

where $\Sigma_z = M_z - \mu_z \mu_z^\top$, $\hat{\Sigma}_z = \hat{M}_z - \hat{\mu}_z \hat{\mu}_z^\top$, and

$$G\Big((\mu_z, \Sigma_z), (\hat{\mu}_z, \hat{\Sigma}_z)\Big) = \sqrt{\|\mu_z - \hat{\mu}_z\|^2 + \mathrm{Tr}\Big(\Sigma_z + \hat{\Sigma}_z - 2(\hat{\Sigma}_z^{1/2} \Sigma_z \hat{\Sigma}_z^{1/2})^{1/2}\Big)}. \tag{13}$$

Key implications:

- Equality in (12) holds when both distributions are Gaussian.
- Assumption 2-(i) holds: Gaussian distributions minimize the 2-Wasserstein distance among all distributions with fixed first and second moments.
- Assumption 2-(ii) holds: the induced moment set is known to be convex and compact.

Kullback–Leibler (KL) Ambiguity Sets

The **Kullback–Leibler (KL) divergence** from $\mathbb{P}_z \in \mathcal{P}(\mathbb{R}^{d_z})$ to $\hat{\mathbb{P}}_z \in \mathcal{P}(\mathbb{R}^{d_z})$ is defined as

$$\mathbb{K}(\mathbb{P}_z, \hat{\mathbb{P}}_z) = \int_{\mathbb{R}^{d_z}} \log\left(\frac{d\mathbb{P}_z}{d\hat{\mathbb{P}}_z}(z)\right) d\mathbb{P}_z(z), \tag{14}$$

if $\mathbb{P}_z$ is absolutely continuous with respect to $\hat{\mathbb{P}}_z$, and $\mathbb{K}(\mathbb{P}_z, \hat{\mathbb{P}}_z) = \infty$ otherwise.

Moment-based lower bound: For distributions with first and second moments (μ_z, M_z) and $(\hat{\mu}_z, \hat{M}_z)$,

$$\mathbb{K}(\mathbb{P}_z, \hat{\mathbb{P}}_z) \ \geq \ \mathcal{T}\Big((\mu_z, \Sigma_z), (\hat{\mu}_z, \hat{\Sigma}_z)\Big), \tag{15}$$

where $\Sigma_z = M_z - \mu_z \mu_z^\top$, $\hat{\Sigma}_z = \hat{M}_z - \hat{\mu}_z \hat{\mu}_z^\top$, and

$$\mathcal{T}\Big((\mu_z, \Sigma_z), (\hat{\mu}_z, \hat{\Sigma}_z)\Big) = \frac{1}{2}\Big((\mu_z - \hat{\mu}_z)^\top \hat{\Sigma}_z^{-1}(\mu_z - \hat{\mu}_z) + \mathrm{Tr}(\Sigma_z \hat{\Sigma}_z^{-1}) - \log\det(\Sigma_z \hat{\Sigma}_z^{-1}) - d_z\Big). \tag{16}$$

Key implications:

- Equality in (15) holds when both distributions are Gaussian.
- Assumption 2-(i) holds: Gaussian distributions minimize the KL divergence among all distributions with fixed first and second moments.
- Assumption 2-(ii) holds: the induced moment set is known to be convex and compact.

Moment Ambiguity Sets

Moment-based divergence: Consider ambiguity sets where the divergence depends only on the first two moments. For distributions $\mathbb{P}_z, \hat{\mathbb{P}}_z \in \mathcal{P}(\mathbb{R}^{d_z})$ with moment pairs (μ_z, M_z) and $(\hat{\mu}_z, \hat{M}_z)$, define

$$\mathbb{D}(\mathbb{P}_z, \hat{\mathbb{P}}_z) = \mathbb{M}\Big((\mu_z, M_z), (\hat{\mu}_z, \hat{M}_z)\Big), \tag{17}$$

where $\mathbb{M} : \mathcal{M}_2^{d_z} \times \mathcal{M}_2^{d_z} \to [0, \infty]$ is any divergence between first–second moment pairs satisfying $\mathbb{M}(m_z, m_z) = 0 \ \forall \ m_z \in \mathcal{M}_2^{d_z}$.

Moment ambiguity set: For a nominal distribution $\hat{\mathbb{P}}_z$ with moments $(\hat{\mu}_z, \hat{M}_z)$, the ambiguity set can be written as

$$\mathcal{B}_z = \left\{ \mathbb{P}_z \in \mathcal{P}(\mathbb{R}^{d_z}) : \begin{array}{l} \mathbb{E}_{\mathbb{P}_z}[z] = \mu_z, \\ \mathbb{E}_{\mathbb{P}_z}[zz^\top] = M_z, \\ \mathbb{M}\Big((\mu_z, M_z), (\hat{\mu}_z, \hat{M}_z)\Big) \leq \rho_z \end{array} \right\}. \tag{18}$$

Key implications:

- Assumption 2-(i) holds trivially: all distributions with the same first two moments attain the same divergence.
- Assumption 2-(ii) holds if the sublevel sets of $\mathbb{M}(\cdot, \hat{m}_z)$ are **convex and compact**.
- This is satisfied, for example, if $\mathbb{M}$ is **quasiconvex and coercive**.

Outline

Original DRLQ and Purified Observations

The controller solves a Distributionally Robust Linear Quadratic (DRLQ) problem:

$$p^\star = \min_{u \in \mathcal{U}_y} \max_{\mathbb{P} \in \mathcal{B}} \mathbb{E}_{\mathbb{P}} \left[u^\top R u + x^\top Q x \right] \quad \text{s.t.} \begin{cases} x = Hu + Gw, \\ y = Cx + v, \end{cases} \tag{19}$$

where x, u, y, w, v are temporally stacked vectors, R, Q, H, G, C are suitable block matrices and $\mathcal{U}_y = \left\{ u = (u_0, \ldots, u_{T-1}) : u_t = \phi_t(y_0, \ldots, y_t),\ \phi_t : \mathbb{R}^{p(t+1)} \to \mathbb{R}^m \text{ measurable} \right\}$.

Key difficulty: Cyclic dependence between control, state, and observation complicates analysis.

Purified observations: Introduce a fictitious noise-free system driven by the same control inputs:

$$\begin{aligned} x'_{t+1} &= A_t x'_t + B_t u_t, \qquad x'_0 = 0, \\ y'_t &= C_t x'_t. \end{aligned} \tag{20}$$

Define the purified observation:

$$\eta_t = y_t - y'_t, \qquad \forall t \in [0, T-1]. \tag{21}$$

Purified Observations: Properties and Reformulation

Key properties of purified observations:

- Since the control input is causal, x'_t and y'_t are computable from past observations.
- Thus, η_t is a causal function of $y_{0:t}$.
- Conversely, y_t can be reconstructed from $\eta_{0:t}$.

Equivalence of policy classes: $\mathcal{U}_y = \mathcal{U}_\eta$ and the existence of optimal affine (linear) policies in $\mathcal{U}_\eta$ is equivalent to that in $\mathcal{U}_y$.

Crucial simplification: By combining the original and fictitious systems,

$$\eta = Dw + v, \tag{22}$$

where D is a block lower-triangular matrix (due to causality).

- Purified observations depend only on exogenous uncertainties (w, v),
- They are independent of the control inputs u.

Equivalent reformulation:

$$p^\star = \min_{u \in \mathcal{U}_\eta} \max_{\mathbb{P} \in \mathcal{B}} \mathbb{E}_\mathbb{P}\left[u^\top R u + x^\top Q x\right] \quad \text{s.t.} \quad x = Hu + Gw. \tag{23}$$

Upper Bound for the Primal

Outer approximation of the ambiguity set: Define an enlarged ambiguity set

$$\bar{\mathcal{B}} = \left\{ \mathbb{P} \in \mathcal{P}(\mathbb{R}^{n+T(n+p)}) : \begin{array}{l} \mathbb{P}_z \in \bar{\mathcal{B}_z} \ \forall z \in \mathcal{Z}, \\ \mathbb{E}_{\mathbb{P}}[zz'^{\top}] = 0 \ \forall z \neq z' \in \mathcal{Z} \end{array} \right\}, \tag{24}$$

where, for each $z \in \mathcal{Z}$,

$$\bar{\mathcal{B}_z} = \left\{ \mathbb{P}_z \in \mathcal{P}(\mathbb{R}^{d_z}) : \exists (\mu_z, M_z) \in \mathcal{M}_2^{d_z} \text{ s.t. } \mathbb{D}\Big(\mathcal{N}(\mu_z, M_z), \hat{\mathbb{P}}_z \Big) \leq \rho_z \right\}. \tag{25}$$

By Assumption 2-(i), $\mathbb{D}\Big(\mathcal{N}(\mu_z, M_z), \hat{\mathbb{P}}_z\Big) \leq \mathbb{D}(\mathbb{P}_z, \hat{\mathbb{P}}_z)$, hence $\mathcal{B} \subseteq \bar{\mathcal{B}}$.

Restriction to affine causal policies: Consider affine policies in purified observations:

$$u = q + U\eta = q + U(Dw + v), \tag{26}$$

where $q = (q_0, \ldots, q_{T-1})$ and U is block lower triangular:

$$U = \begin{bmatrix} U_{0,0} & & & \\ U_{1,0} & U_{1,1} & & \\ \vdots & \vdots & \ddots & \\ U_{T-1,0} & \cdots & \cdots & U_{T-1,T-1} \end{bmatrix}.$$

Block lower triangularity of U enforces causality, hence $u \in \mathcal{U}_\eta$.

Upper Bound for the Primal

Upper bound problem: Restricting the controller while enlarging nature's feasible set yields:

$$\bar{p}^\star = \begin{cases} \min\limits_{U,\,q,\,x,\,u} & \max\limits_{\mathbb{P}\in\bar{\mathcal{B}}} \quad \mathbb{E}_\mathbb{P}\left[u^\top Ru + x^\top Qx\right] \\ \quad \text{s.t.} & \qquad U \in \mathcal{U}, \quad u = q + U(Dw + v), \quad x = Hu + Gw. \end{cases} \tag{27}$$

Upper bound property: $\bar{p}^\star \;\geq\; p^\star$.

- Controller is restricted $\Rightarrow$ objective can only increase.
- Nature is relaxed $\Rightarrow$ objective can only increase.

Key observation: For fixed U, q,

- u and x are affine in (w, v),
- $\mathbb{E}_\mathbb{P}[u^\top Ru + x^\top Qx]$ depends only on first and second moments of (w, v).

Proposition 1. Problem (27) has the same optimal value as the optimization problem:

$$\begin{aligned} \bar{p}^\star = \min_{\substack{q\in\mathbb{R}^{pT} \\ U\in\mathcal{U}}} \max_{\substack{(\mu_w, M_w)\in\mathcal{M}(\mu_w, M_w) \\ (\mu_v, M_v)\in\mathcal{M}(\mu_v, M_v)}} & \operatorname{Tr}\Big(\left((UD)^\top RUD + (G + HUD)^\top Q(G + HUD)\right)M_w + U^\top \bar{R}UM_v\Big) \\ & + 2q^\top(\bar{R}UD + G^\top QH)\mu_w + 2q^\top \bar{R}U\mu_v + q^\top \bar{R}q, \end{aligned} \tag{28}$$

where $\bar{R} = R + H^\top QH$.

Lower Bound for the Dual

Restriction of nature's feasible set: Let $\mathcal{B}_{\mathcal{N}}$ denote the family of all *Gaussian* distributions contained in $\mathcal{B}$. The resulting lower-bounding problem is

$$\underline{d}^\star = \begin{cases} \max\limits_{\mathbb{P}\in\mathcal{B}_{\mathcal{N}}} & \min\limits_{x,\,u} \quad \mathbb{E}_{\mathbb{P}}\left[u^\top Ru + x^\top Qx\right] \\ & \text{s.t.} \quad u \in \mathcal{U}_\eta, \quad x = Hu + Gw. \end{cases} \tag{29}$$

Lower bound property: $\underline{d}^\star \ \leq \ d^\star$
Nature's feasible set is restricted $\Rightarrow$ objective can only decrease.

Key observation: For any fixed Gaussian distribution $\mathbb{P} \in \mathcal{B}_{\mathcal{N}}$, the inner minimization problem is solved by an **affine control policy**.

Proposition 2. Problem (29) has the same optimal value as the optimization problem:

$$\begin{aligned} \underline{d}^\star = \max_{\substack{(\mu_w, M_w)\in\mathcal{M}(\mu_w, M_w)\\(\mu_v, M_v)\in\mathcal{M}(\mu_v, M_v)}} \ \min_{\substack{q\in\mathbb{R}^{pT}\\U\in\mathcal{U}}} \ & \mathrm{Tr}\Big(\big((UD)^\top RUD + (G + HUD)^\top Q(G + HUD)\big)M_w + U^\top \bar{R}UM_v\Big) \\ & + 2q^\top(\bar{R}UD + G^\top QH)\mu_w + 2q^\top \bar{R}U\mu_v + q^\top \bar{R}q, \end{aligned} \tag{30}$$

where $\bar{R} = R + H^\top QH$.

Strong Duality and Nash Equilibrium

Theorem 3.1 (Strong duality). The optimal value of the original DRLQ problem equals the optimal values of its upper-bounding, lower-bounding, and dual formulations:

$$p^\star = \bar{p}^\star = \underline{d}^\star = d^\star,$$

and all optimal values are attained.

Proof sketch.

- **Weak duality:** By construction of the bounds, $\underline{d}^\star \ \leq \ d^\star \ \leq \ p^\star \ \leq \ \bar{p}^\star$.
- **Key observation:** Problems (28) and (30) have the *same objective function*, with the order of minimization and maximization reversed.
- **Minimax structure:**
  - Minimization variables (U, q): convex, closed feasible set $\mathcal{U} \times \mathbb{R}^{pT}$, objective is convex quadratic and coercive ($R \succ 0$).
  - Maximization variables $(\mu_w, M_w), (\mu_v, M_v)$: feasible sets are convex and compact (Assumption 2-(ii)), objective is linear.
- **Sion's minimax theorem** applies $\Rightarrow$ $\underline{d}^\star = \bar{p}^\star$.

Conclusion: Strong duality holds and a **Nash equilibrium** exists.

Outline

Optimality of Affine Policies and Gaussian Distributions

Corollary 1 (Decision maker's Nash strategy is affine).
The primal DRLQ problem admits an optimal **affine** policy of the form

$$u^\star = q^\star + U^\star \eta, \qquad U^\star \in \mathcal{U},\ q^\star \in \mathbb{R}^{mT}.$$

Corollary 2 (Nature's Nash strategy is a Gaussian distribution).
The dual DRLQ problem admits an optimal solution that is a **Gaussian distribution**,

$$\mathbb{P}^\star \in \mathcal{B}_{\mathcal{N}}.$$

- The worst-case distribution $\mathbb{P}^\star$ is fully characterized by the optimal moment pairs $(\mu_w^\star, M_w^\star)$ and $(\mu_v^\star, M_v^\star)$, obtained by solving the finite-dimensional problem (30).
- This result is **non-trivial**: the ambiguity set contains many non-Gaussian distributions, yet the worst-case one is Gaussian.

Key insights:

- **Affine policies are optimal**: even under distributional ambiguity, no more complex feedback policy can outperform an affine one.
- **Gaussian noise is conservative**: assuming Gaussian disturbances captures the worst-case cost within the ambiguity set when the nominal distribution is Gaussian.

Ambiguity Sets with Zero-Mean Distributions

Assumption 3. For any $z \in \mathcal{Z}$, the moment set

$$\mathcal{M}(\mu_z, M_z) = \left\{(\mu_z, M_z) \in \mathcal{M}_2^{d_z} : \mathbb{D}\Big(\mathcal{N}(\mu_z, M_z), \hat{\mathbb{P}}_z\Big) \leq \rho_z\right\}$$

satisfies:

$$(\mu_z, M_z) \in \mathcal{M}(\mu_z, M_z) \;\Rightarrow\; (0, M_z) \in \mathcal{M}(\mu_z, M_z).$$

- If a Gaussian distribution with mean μ_z is feasible, then the *centered* Gaussian with the same second moment is also feasible.
- Nature can transfer deterministic bias (mean) into additional variance.
- Covariance changes from $M_z - \mu_z \mu_z^\top$ to M_z,
- Since $M_z \succeq M_z - \mu_z \mu_z^\top$, uncertainty is *inflated* in the Loewner order.

Proposition 3. If $\hat{\mu}_z = 0$, then:

- 2-Wasserstein and KL divergence ambiguity sets satisfy Assumption 3,
- Moment-based ambiguity sets satisfy Assumption 3 if

$$\mathbb{M}\Big((0, M_z), (0, \hat{M}_z)\Big) \leq \mathbb{M}\Big((\mu_z, M_z), (0, \hat{M}_z)\Big), \quad \forall (\mu_z, M_z) \in \mathcal{M}_2^{d_z}. \tag{31}$$

Why this is mild:

- Zero-mean nominal noise is standard in LQG models,
- For moment-based divergences, penalizing mean deviations is a sensible requirement.

Optimality of Linear Policies and Zero-Mean Distributions

Theorem 3.2 (Worst-case mean and linear controls).
Under Assumptions 1–3, the dual DRLQ problem admits an optimal solution

$$\mathbb{P}^\star \in \mathcal{B}_\mathcal{N} \quad \text{with} \quad \mathbb{E}_{\mathbb{P}^\star}[z] = 0 \ \ \forall z \in \mathcal{Z}.$$

Moreover, under such $\mathbb{P}^\star$, the decision maker admits an optimal **linear** control policy

$$u^\star = U^\star \eta, \qquad U^\star \in \mathcal{U}.$$

Proof sketch (intuition).

- From strong duality, nature chooses a Gaussian distribution and the controller restricts to affine policies $u = q + U\eta$.
- For fixed U, the objective is a convex quadratic function of the mean μ_z.
- Under Assumption 3, nature can replace a nonzero mean μ_z by zero mean and increased covariance $+\mu_z\mu_z^\top$, which *increases cost*.
- Hence, the worst-case mean is $\mu_z^\star = 0$.
- With zero-mean disturbances, the optimal intercept is $q^\star = 0$,
- The optimal policy reduces from affine to linear.

Key insights:

- **Linear policies are optimal** in DRLQ, extending classical LQG results.
- **Zero-mean Gaussian noise is worst-case conservative**.
- Predictable offsets benefit neither player in equilibrium.

Worst-Case Covariance: Structural Assumption

Under Assumptions 1–3, the worst-case distributions are zero-mean Gaussian. Thus, each worst-case distribution is fully characterized by its covariance $\Sigma_z \succeq 0$. We consider

$$\mathcal{M}_{\Sigma_z} \;=\; \mathcal{M}(\mu_z = 0, M_z = \Sigma_z).$$

Assumption 4. For each $z \in \mathcal{Z}$, there exists a function $g : \mathbb{S}_+^{d_z} \to \mathbb{R}$ such that

$$\mathcal{M}_{\Sigma_z} = \{\Sigma_z \succeq 0 : g(\Sigma_z) \leq 0\}\,,$$

where g satisfies:

- g is **convex** on $\mathbb{S}_+^{d_z}$,
- g is **differentiable** on $\mathbb{S}_{++}^{d_z}$,

(i) $g(\hat{\Sigma}_z) < 0$ (nominal covariance is interior),

(ii) $\hat{\Sigma}_z \in \arg\min_{\Sigma_z \succeq 0} g(\Sigma_z)$,

(iii) $\nabla g(\Sigma_1) \succeq \nabla g(\Sigma_2) \Rightarrow \Sigma_1 \succeq \Sigma_2$.

Interpretation:

- $g(\Sigma_z)$ acts as a convex measure of distance from the nominal covariance,
- Larger gradients correspond to larger covariances (order-reflecting in the Loewner order), ensuring monotonicity of uncertainty.

Worst-Case Covariance Inflation

Theorem 3.3 (Worst-case covariance matrix).
Under Assumptions 1–4, the dual DRLQ problem admits an optimal solution

$$\mathbb{P}_z^\star = \mathcal{N}(0, \Sigma_z^\star), \quad \forall z \in \mathcal{Z},$$

where the optimal covariance matrices satisfy

$$\Sigma_z^\star \succeq \hat{\Sigma}_z \qquad \forall z \in \mathcal{Z}.$$

Proof sketch (intuition).

- From Theorem 3.2, nature restricts to zero-mean Gaussian distributions.
- The dual problem optimizes over covariance matrices Σ_z only.
- The objective is increasing with increase in covariance in the Loewner order.
- By Assumption 4:
  - The nominal covariance $\hat{\Sigma}_z$ minimizes g,
  - Moving away from $\hat{\Sigma}_z$ increases both uncertainty and cost,
  - The worst-case covariance lies on the boundary $g(\Sigma_z) = 0$.

Key insights:

- Nature's optimal strategy is to *"inflate"* covariance,
- This generalizes the idea "larger covariance = worse case" to dynamic LQG systems,
- **Practical takeaway:** covariance inflation is a conservative and effective robustness heuristic.

Outline

Efficient Numerical Solution of DRLQ

By Theorem 3.1, the DRLQ problem is equivalent to a finite-dimensional convex–concave problem in moments.

Corollary 3. Under Assumptions 1–4, the solution to the DRLQ problem can be computed using Kalman Filtering techniques applied to classical LQG model with distribution $\mathbb{P}^\star$, where

- Nature's optimal strategy is a zero-mean Gaussian distribution $\mathbb{P}^\star$,
- The controller's optimal policy is *linear* and given by $u_t^\star = K_t \hat{x}_t$,

Why Kalman filtering applies:

- Theorem 3.3 guarantees $\Sigma_v^\star \succeq \hat{\Sigma}_v \succ 0$,
- Hence, Kalman filter recursions are well-defined under $\mathbb{P}^\star$.

The dual problem reduces to:

$$\max_{\Sigma_w \in \mathcal{M}_{\Sigma_w},\, \Sigma_v \in \mathcal{M}^+_{\Sigma_v}} f(\Sigma_w, \Sigma_v), \tag{32}$$

where $f(\Sigma_w, \Sigma_v)$ is the optimal value of the corresponding LQG problem.

Solution method:

- Use a **Frank–Wolfe algorithm** over covariance matrices,
- Each iteration solves a standard LQG subproblem via Kalman filtering and dynamic programming.

Frank–Wolfe Algorithm for Solving DRLQ

Proposition 4. Under Assumptions 1–4, the function $f(\Sigma_w, \Sigma_v)$ is **concave** and **β-smooth** on $\mathcal{M}_{\Sigma_w} \times \mathcal{M}^+_{\Sigma_v}$.

Implication:
Problem (32) can be solved using a **Frank–Wolfe (conditional gradient) algorithm**.
At iteration k, solve the linearized problem:

$$\max_{\Sigma_w \in \mathcal{M}_{\Sigma_w}, \Sigma_v \in \mathcal{M}^+_{\Sigma_v}} \langle \nabla_{\Sigma_w} f^{(k)}, \Sigma_w - \Sigma_w^{(k)} \rangle + \langle \nabla_{\Sigma_v} f^{(k)}, \Sigma_v - \Sigma_v^{(k)} \rangle \tag{33}$$

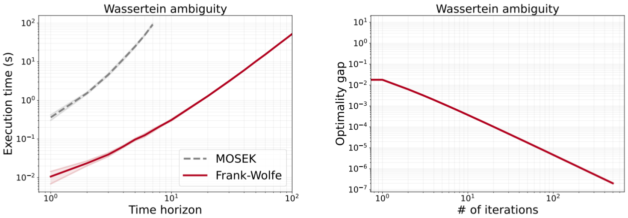
Here, $f(\Sigma_w, \Sigma_v)$ denotes the optimal value of a classical LQG problem.
Next iterates are obtained by moving towards a maximizer $(\Sigma_w^\star, \Sigma_v^\star)$

$$(\Sigma_w^{(k+1)}, \Sigma_v^{(k+1)}) \leftarrow (\Sigma_w^{(k)}, \Sigma_v^{(k)}) + \alpha \cdot (\Sigma_w^\star - \Sigma_w^{(k)}, \Sigma_v^\star - \Sigma_v^{(k)}) \tag{34}$$

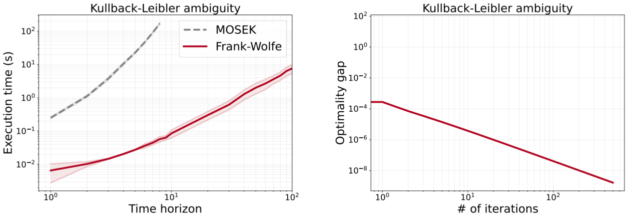
Key computational advantages:

- The direction-finding problem is **separable**:
  decomposes into $|\mathcal{Z}| = 2T + 1$ independent subproblems, that can be solved **in parallel**.
- $f(\Sigma_w, \Sigma_v)$ is the value of a classical LQG problem,
- Gradients $\nabla_{\Sigma_z} f$ are computed efficiently via **automatic differentiation** through Kalman filtering and DP.

Computation Time and Optimality Gap

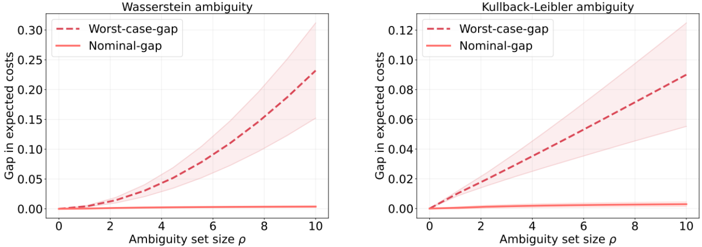


Wasserstein ambiguity



Kullback–Leibler ambiguity

Optimality Gap



$$\text{Worst-case gap} = \mathbb{E}_{\mathbb{P}^\star}[J(\hat{u})] - \mathbb{E}_{\mathbb{P}^\star}[J(u^\star)]\,, \tag{35}$$

$$\text{Nominal gap} = \mathbb{E}_{\hat{\mathbb{P}}}[J(u^\star)] - \mathbb{E}_{\hat{\mathbb{P}}}[J(\hat{u})]\,. \tag{36}$$

Outline

Problem Setup

We consider a linear time-invariant system with average-cost criterion:

$$\begin{aligned} x_{t+1} &= A_0 x_t + B_0 u_t + w_t, \\ y_t &= C_0 x_t + v_t, \qquad t \in \mathbb{N}, \end{aligned} \tag{37}$$

where all random variables are functions of x_0, $\{w_t\}_{t\geq 0}$, and $\{v_t\}_{t\geq 0}$.
The ambiguity set $\mathcal{B}_\infty$ is defined analogously to the finite-horizon case.

Control policies: $\mathcal{U}_y = \left\{ u : u_t = \phi_t(y_0, \ldots, y_t),\ \phi_t : \mathbb{R}^{p(t+1)} \to \mathbb{R}^m \text{ measurable } \forall t \in \mathbb{N} \right\}$.

Infinite-horizon DRLQ Problem

$$\min_{u \in \mathcal{U}_y} \max_{\mathbb{P} \in \mathcal{B}_\infty} J_{\mathbb{P}}(x, u), \qquad J_{\mathbb{P}}(x, u) = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}_{\mathbb{P}} \left[x_t^\top Q_0 x_t + u_t^\top R_0 u_t \right]. \tag{38}$$

Assumption 5. The matrices A_0, B_0, C_0 and Q_0 satisfy:

(i) (A_0, B_0) is **stabilizable**,
(ii) (A_0, C_0) is **detectable**,
(iii) $Q_0 \succ 0$.

Purpose of assumptions: ensure stability, finite average cost, and existence of stationary optimal policies.

Problem Setup

Assumption 6. The nominal distribution $\hat{\mathbb{P}}$ and ambiguity set $\mathcal{B}^\infty$ satisfy:

(i) $\hat{\mathbb{P}} \in \mathcal{B}^\infty_{\mathcal{N}}$ is time-invariant, zero-mean Gaussian with $\hat{\Sigma}_z \succ 0$ for all $z \in \mathcal{Z}$,

(ii) Ambiguity radii are time-invariant: $\rho_z = \rho_w$ for $z \in \{x_0, w_t\}$, $\rho_z = \rho_v$ for $z \in \{v_t\}$,

(iii) All distributions in $\mathcal{B}^\infty$ have zero mean.

Implications:

- Ambiguity sets reduce to covariance uncertainty only,
- Moment sets are convex and compact:

$$\mathcal{M}_{\Sigma_w} = \{\Sigma \succeq 0 : \mathbb{D}(\mathcal{N}(0,\Sigma), \hat{\mathbb{P}}) \leq \rho_w\}, \quad \mathcal{M}_{\Sigma_v} = \{\Sigma \succeq 0 : \mathbb{D}(\mathcal{N}(0,\Sigma), \hat{\mathbb{P}}) \leq \rho_v\}.$$

Infinite-horizon DRLQ (primal problem): Considering purified observations,

$$\begin{aligned} p^\star = \inf_{x,u} \sup_{\mathbb{P} \in \mathcal{B}^\infty} \; & J_{\mathbb{P}}(x,u), \\ \text{s.t.} \quad & u \in \mathcal{U}_\eta, \quad x = Hu + Gw. \end{aligned} \tag{39}$$

Upper Bound - Idea and Construction

Basic idea: As in the finite-horizon case, we obtain an upper bound on $p^\star$ by:

- **Inflating nature's ambiguity set** from $\mathcal{B}_\infty$ to $\bar{\mathcal{B}}_\infty$,
- **Restricting the controller** to a structured class of stationary policies.

Restriction to stationary purified-output feedback: To ensure finite long-run average cost and enable duality, we restrict attention (without loss of optimality under Assumption 5) to

$$u = U(Dw + v),$$

where $U \in \mathcal{U}_\infty$ is a block lower-triangular *Toeplitz* operator.

This guarantees:

- Stationarity of the closed-loop system,
- Convergence of state and control covariances,
- Well-defined average cost.

Upper-bounding problem.

$$\bar{p}^\star = \begin{cases} \min\limits_{U \in \mathcal{U}_\infty} \max\limits_{\mathbb{P} \in \bar{\mathcal{B}}_\infty} & J_{\mathbb{P}}(x, u) \\ \quad \text{s.t.} & u = U(Dw + v), \quad x = Hu + Gw. \end{cases} \tag{40}$$

Upper Bound - Gaussian Reduction

Proposition 5. *Under Assumptions 2, 5, and 6, the inner maximization problem in* (40) *is solved by a* **time-invariant Gaussian distribution** $\mathbb{P}^\star \in \mathcal{B}_{\mathcal{N}}^{\infty}$.

Key implication: Any $\mathbb{P} \in \mathcal{B}_{\mathcal{N}}^{\infty}$ is uniquely characterized by covariance matrices (Σ_w, Σ_v). Hence, the upper bound reduces to the finite-dimensional minimax problem

$$\bar{p}^\star = \min_{U \in \mathcal{U}_\infty} \max_{\Sigma_w \in \mathcal{M}_{\Sigma_w},\, \Sigma_v \in \mathcal{M}_{\Sigma_v}} J(U; \Sigma_w, \Sigma_v). \tag{41}$$

Further refinement: By Assumption 4 and the infinite-horizon analogue of Theorem 3.3, the worst-case covariances satisfy

$$\Sigma_w^\star \succeq \hat{\Sigma}_w, \qquad \Sigma_v^\star \succeq \hat{\Sigma}_v,$$

allowing us to restrict to $\mathcal{M}_{\Sigma_w}^{+}$ and $\mathcal{M}_{\Sigma_v}^{+}$.

Proof sketch (intuition).

- For fixed stationary U, x and u are stationary linear processes,
- The average cost depends only on second moments,
- Gaussian distributions are worst-case for fixed moments (Assumption 2),
- Nature increases cost by inflating covariance in Loewner order.

Lower Bound for the Dual

Restricted ambiguity set: $\underline{\mathcal{B}}_{\mathcal{N}}^{\infty} = \left\{ \mathbb{P} \in \mathcal{B}_{\mathcal{N}}^{\infty} : \mathbb{E}_{\mathbb{P}}[zz^\top] \succeq \mathbb{E}_{\hat{\mathbb{P}}}[zz^\top], \ \forall z \in \mathcal{Z} \right\}.$

Lower-bounding problem:

$$\underline{d}^\star = \begin{cases} \sup\limits_{\mathbb{P} \in \underline{\mathcal{B}}_{\mathcal{N}}^{\infty}} \inf\limits_{x,u} & J_{\mathbb{P}}(x, u) \\ \text{s.t.} & u \in \mathcal{U}_\eta, \quad x = Hu + Gw. \end{cases} \tag{42}$$

By construction, $\underline{d}^\star \leq d^\star$.

Proposition 6. *Under Assumptions 2, 5, and 6, for any* $\mathbb{P} \in \mathcal{B}_{\infty,\mathcal{N}}^{+}$, *the inner minimization is solved by a* **stationary linear policy** $u = U\eta$ *with* $U \in \mathcal{U}_\infty$.

Resulting reformulation:

$$\underline{d}^\star = \max_{\Sigma_w \in \mathcal{M}_{\Sigma_w}^{+},\, \Sigma_v \in \mathcal{M}_{\Sigma_v}^{+}} \ \min_{U \in \mathcal{U}_\infty} J(U; \Sigma_w, \Sigma_v). \tag{43}$$

Key takeaway: Both players can be restricted to **stationary, linear strategies**, yielding a finite-dimensional **convex–concave saddle-point problem**.

Strong Duality and Nash Equilibrium

Theorem 6.1 (Nash strategies and strong duality). Under Assumptions 2, 4, and 5–6:

(i) **Strong duality holds:** $p^\star = \bar{p}^\star = \underline{d}^\star = d^\star$, and all optimal values are attained.

(ii) **Optimal controller:** the primal DRLQ problem admits an optimal *stationary linear policy* $u_t^\star = U^\star \eta_t, \qquad U^\star \in \mathcal{U}_\infty.$

(iii) **Optimal adversary:** the dual problem admits an optimal *time-invariant Gaussian distribution* $\mathbb{P}^\star \in \mathcal{B}_{\mathcal{N}}^\infty$.

(iv) **Covariance inflation:** $\Sigma_w^\star \succeq \hat{\Sigma}_w, \qquad \Sigma_v^\star \succeq \hat{\Sigma}_v.$

Proof sketch.

- By construction of bounds: $\underline{d}^\star \le d^\star \le p^\star \le \bar{p}^\star$.
- Propositions 5 and 6 reduce the primal and dual bounds to the *same saddle function* $J(U; \Sigma_w, \Sigma_v)$, differing only in the order of min/max.
- Feasible sets:
  - U_∞: convex subset of a linear space (Toeplitz operators),
  - $\mathcal{M}_{\Sigma_w}^+ \times \mathcal{M}_{\Sigma_v}^+$: convex and compact.
- J is convex in U and linear (thus concave and upper semicontinuous) in (Σ_w, Σ_v).
- **Fan–Sion minimax theorem** applies $\Rightarrow$ minimax equality.

Key insights:

- Infinite-horizon DRLQ admits a **stationary Nash equilibrium**.
- **Linear control and Gaussian noise remain optimal** under distributional ambiguity.
- Worst-case robustness is achieved via **time-invariant covariance inflation**.

Outline

Conclusions

- The paper formulates a **distributionally robust LQG (DRLQ)** problem, replacing the fixed disturbance model with a **divergence-based ambiguity set** around a nominal distribution.
- Under:
  - zero-mean Gaussian nominal noise,
  - orthogonality (uncorrelatedness) of second moments,
  - mild structural assumptions on the divergence,

  the paper proves that:
  - **linear output-feedback policies** and
  - a **Gaussian worst-case distribution**

  form a **Nash equilibrium** of the minimax game.
- It is shown that the adversary optimally chooses **zero-mean noise** and **inflates the nominal covariance matrix**, while the controller's optimal policy becomes **linear**.
- These results:
  - generalize and rationalize several classical and robust LQG results,
  - provide an intuitive robustness heuristic based on **covariance inflation**,
  - enable an efficient **Frank–Wolfe algorithm** whose iterations correspond to standard LQG subproblems.
- All results extend to the **infinite-horizon, average-cost** setting, yielding **stationary linear policies** and a **time-invariant Gaussian worst-case model**.

Limitations and Future Directions

- **Performance guarantees under general noise:**
  - If a true disturbance distribution Q belongs to an ambiguity set $\mathcal{B}$, the DRLQ solution is feasible for the (intractable) LQG problem under Q,
  - The optimality gap can be bounded via robust and optimistic formulations.
- **Open challenges:**
  - Solving the optimistic problem $\inf_{P \in \mathcal{B}} \inf_{u \in \mathcal{U}_y} \mathbb{E}_P[J(u)]$, which may be nonconvex,
  - Testing membership of a given distribution Q in $\mathcal{B}$,
  - Calibrating ambiguity sets without excessive conservatism.
- **Extensions to other control models:**
  - Linear–exponential–quadratic Gaussian (LEQG) control,
  - Affine dynamics with extended quadratic costs,
  - Other stochastic control formulations beyond LQG.
- **Learning and constraints:**
  - State and control constraints with robustness guarantees,
  - Joint control and ambiguity-set learning.

The End

Thank You!

Questions or Feedback?