Data-driven distributionally robust LQR with multiplicative noise

Peter Coppens  
Mathijs Schuurmans  
Panagiotis Patrinos  
STADIUS, Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium.

Abstract

We present a data-driven method for solving the linear quadratic regulator problem for systems with multiplicative disturbances, the distribution of which is only known through sample estimates. We adopt a distributionally robust approach to cast the controller synthesis problem as semidefinite programs. Using results from high dimensional statistics, the proposed methodology ensures that their solution provides mean-square stabilizing controllers with high probability even for low sample sizes. As sample size increases the closed-loop cost approaches that of the optimal controller produced when the distribution is known. We demonstrate the practical applicability and performance of the method through a numerical experiment.

Keywords: data driven, distributionally robust, linear quadratic regulation, multiplicative noise, stochastic optimal control

1. Introduction

We will develop controllers for linear systems with time-varying parametric uncertainty, which may cover a wide range of system classes extensively studied in the literature. For example, we obtain Linear Parameter Varying (LPV) systems when the disturbance is observable at each time step (Wu et al., 1996; Byrnes, 1979), Linear Difference Inclusions (LDIs) when it is unknown but norm-bounded (Boyd et al., 1994) and stochastic systems with multiplicative noise when it varies stochastically (Wonham, 1967).

In many practical applications, however, the distribution of the disturbance is not known. These traditional control approaches either make a boundedness assumption on the disturbance or on its moments, which allows for a fully robust approach (El Ghaoui, 1995). Such approaches, however, disregard any statistical information that may be obtained on the distribution of the disturbances. Our aim, instead, is to design linear controllers which use sampled data to improve performance over fully robust approaches, while inheriting many of the system-theoretical guarantees of a robust control strategy. To this end, we adopt a distributionally robust (DR) approach (Dupaev, 1987; Delage and Ye, 2010) towards solving the infinite-horizon Linear Quadratic Regulator (LQR) problem, where we minimize the expected cost for the worst-case distribution in a so-called ambiguity set computed based on the available data such that it contains the true distribution with high probability. Similar techniques were recently studied in Schuurmans et al. (2019) for stochastic jump linear systems and in Dean et al. (2019) for deterministic systems, where the system matrices $A$ and $B$ are learned from data.

Our main contributions are summarized as follows. Leveraging recent results from high dimensional statistics, we provide practical high-probability confidence bounds for the ambiguity sets,
which depend only on known quantities (Section 3). We then extend the solution of the (nominal) infinite horizon LQR problem with known distribution to related DR counterparts which account for the ambiguity on the disturbance distribution. Whenever the mean of the disturbance is known, we show that the DR problem is equivalent to a semidefinite program (SDP) which has the same form as the nominal one. Next, we extend the formulation to the setting in which both the mean and the covariance are only known to lie in an ellipsoidal ambiguity set (Section 4.2), for which we can only approximate the optimal controller.

1.1. Notation

Let $\mathbb{R}$ denote the reals, $\mathbb{N}$ the naturals and $\mathbb{N}_+ = \mathbb{N} \setminus \{0\}$. For symmetric matrices $P, Q$ we write $P \succ Q$ ($P \succeq Q$) to signify that $P - Q$ is positive (semi)definite and denote by $\otimes$ the Kronecker product. We assume that all random variables are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $\Omega$ the sample space, $\mathcal{F}$ its associated $\sigma$-algebra and $\mathbb{P}$ the probability measure. Let $y : \Omega \to \mathbb{R}^n$ be a random vector defined on $(\Omega, \mathcal{F}, \mathbb{P})$. With some abuse of notation we will write $y \in \mathbb{R}^n$ to state the dimension of this random vector. Let $\mathbb{P}_y$ denote the distribution of $y$, i.e., $\mathbb{P}_y(A) = \mathbb{P}[y \in A]$, then a trajectory $\{y_i\}_{i=1}^N$ of identically and independently distributed (i.i.d.) copies of $y$ is defined by the distribution it induces. That is, for any $A_0, \ldots, A_N \in \mathcal{F}$ we define $\mathbb{P}_{y_0, \ldots, y_N}(A_0 \times \cdots \times A_N) := \mathbb{P}[y_0 \in A_0 \land \cdots \land y_N \in A_N] = \prod_{i=0}^N \mathbb{P}_y(A_i)$. This definition can be extended to infinite trajectories $\{y_i\}_{i \in \mathbb{N}}$ by Kolmogorov’s existence theorem Billingsley (1995). We will write the expectation operator as $\mathbb{E}$. We denote by $\mathbb{E}[y \mid z]$ the conditional expectation with respect to $z$.

2. Problem statement

Consider the stochastic discrete-time system with input- and state-multiplicative noise given by:

$$x_{k+1} = Ax_k + B(w_k)u_k$$  \hspace{1cm} (1)

with $A(w) := A_0 + \sum_{i=1}^{n_w} w_i A_i$, $B(w) := B_0 + \sum_{i=1}^{n_w} w_i B_i$,

where at each time $k$, $x_k \in \mathbb{R}^{n_x}$ denotes the state, $u_k \in \mathbb{R}^{n_u}$ the input and $w_k \in \mathbb{R}^{n_w}$ an i.i.d. copy of a square integrable random variable $w$ distributed according to $\mathbb{P}_w$. We introduce the following shorthands: $\mathbf{A} := [A_0^\top \ldots A_n^\top]^\top$, $\mathbf{B} := [B_0^\top \ldots B_n^\top]^\top$, $\mathbf{A}_0 := [A_0^\top \mathbf{A}^\top]^\top$, $\mathbf{B}_0 := [B_0^\top \mathbf{B}^\top]^\top$. We also define $\Sigma_0 := \begin{bmatrix} 1 & \mu^\top \\ \mu \Sigma + \mu \mu^\top \end{bmatrix}$, where $\mu := \mathbb{E}[w]$, $\Sigma := \mathbb{E}[(w-\mu)(w-\mu)^\top]$.

2.1. Nominal stochastic LQR problem and solution

The primary goal is to solve the following LQR problem:

$$\begin{align*}
\text{minimize} & \quad \mathbb{E} \left[ \sum_{k=0}^\infty x_k^\top Q x_k + u_k^\top R u_k \right] \\
\text{subj. to} & \quad x_{k+1} = Ax_k + B(w_k)u_k, \quad k \in \mathbb{N} \\
& \quad x_0 = x,
\end{align*}$$  \hspace{1cm} (2)

where we assume that $Q \succ 0$ and $R \succ 0$. The solution of (2) will yield a controller that renders the closed-loop system exponentially stable in the mean square sense, which is defined as follows.
Definition 1 ((Exponential) Mean Square Stability) We say that an autonomous system \( x_{k+1} = A(w_k)x_k \) is mean square stable (m.s.s.) iff \( \mathbb{E}[x_k^T x_k] \to 0 \) as \( k \to \infty \). It is exponentially mean square stable (e.m.s.s.) iff there exists a pair of positive constants \( \gamma \in (0, 1) \) and \( c \) such that \( \mathbb{E}[x_k^T x_k] \leq c \gamma^k \|x_0\| \) for all \( k \in \mathbb{N} \) and for each \( x_0 \in \mathbb{R}^{n_x} \).

This property can be verified using the classical Lyapunov operator (Morozan, 1983):

Theorem 2 (Lyapunov stability) For the autonomous system \( x_{k+1} = A(w_k)x_k \) the following statements about the system are then equivalent: (i) it is m.s.s., (ii) it is e.m.s.s., (iii) \( \exists P \succ 0 \): \( P - A_0^T (\Sigma_0 \otimes P) A_0 \succ 0 \).

Proof See Appendix A.1.

The LQR problem (2) has been studied for many variations of (1) (Morozan, 1983; Costa and Kubrusly, 1997). The following proposition is then similar to many classical results in literature:

Proposition 3 Consider a system with dynamics (1) and the associated LQR problem (2). Assuming that (1) is mean square stabilizable, i.e., there exists a \( K \) and \( P \succ 0 \) such that (3) holds for the closed-loop system \( x_{k+1} = (A(w_k) + B(w_k)K)x_k \), then the following statements hold.

I The optimal solution of (2) is given by \( K_\infty = -(R + G(P_\infty))^{-1} H(P_\infty) \), with \( P_\infty \) the solution of the following Riccati equation:

\[
P_\infty = Q + F(P_\infty) - H(P_\infty)^T (R + G(P_\infty))^{-1} H(P).
\]

II The controller \( K_\infty \) stabilizes (1) in the mean square sense.
III The solution of the Riccati equation is found by solving the following SDP:

\[
\begin{align*}
\text{minimize} & \quad -\text{Tr } P \\
\text{subj. to} & \quad \begin{bmatrix} Q - P + F(P) & H(P)^T \\ H(P) & R + G(P) \end{bmatrix} \succeq 0, \\
& \quad P \succeq 0.
\end{align*}
\]

With the shorthands: \( F(P) := A_0^T (\Sigma_0 \otimes P) A_0 \), \( G(P) := B_0^T (\Sigma_0 \otimes P) B_0 \), \( H(P) := B_0^T (\Sigma_0 \otimes P) A_0 \).

Proof See Appendix A.2.

Notice that the optimal solution to the LQR problem depends solely on the first and second moment of the random disturbance. This motivates our choice of the parametric form of the ambiguity set introduced in the final problem statement which we state in the next section.
2.2. Data-driven stochastic LQR problem

Consider now the case where \( P \) is not known a priori and only a finite set of offline i.i.d. samples \( \{ \hat{w} \}_{i=0}^{M-1} \) is available. For clarity, a random variable \( y \) dependent on these samples is denoted by \( \hat{y} \).

It is apparent that under such circumstances, it is only possible to solve (2) approximately. For most applications, however, it is crucial that the approximate solution remains stabilizing, which is not trivial. After all, consider the empirical approach, where (2) is solved using

\[
\hat{\Sigma} := \frac{1}{M} \sum_{i=0}^{M-1} \hat{w}_i \hat{w}_i^T \quad \text{and} \quad \hat{\mu} := \frac{1}{M} \sum_{i=0}^{M-1} \hat{w}_i,
\]

which does not guarantee stability. We will illustrate this with an example.

**Example 1 (Motivating example)** Consider the following scalar system,

\[
x_{k+1} = (0.75 + w_k)x_k + u_k,
\]

where \( x_k \in \mathbb{R}, u_k \in \mathbb{R} \) and \( w_k \in \mathbb{R} \) are defined as before, where now \( w \) is assumed Gaussian with \( \mathbb{E}[w] = 0 \) and \( \mathbb{E}[w^2] = \sigma^2 = 0.5 \). The empirical variance is \( \hat{\sigma}^2 = \frac{1}{M} \sum_{i=0}^{M-1} \hat{w}_i^2 \), using i.i.d. samples \( \{ \hat{w}_i \}_{i=0}^{M-1} \). We will develop an optimal LQR controller with a stage cost given by \( q x_k^2 + ru_k^2 = x_k^2 + 10^4 u_k^2 \). Using the results from Proposition 3 to derive the Riccati equation, we obtain

\[
q - p + (\hat{\sigma}^2 + 0.75^2)p - (0.75p)(r + p)^{-1}(0.75p) = 0.
\]

Note that this is a quadratic equation in \( p \) with the following positive solution:

\[
\hat{p} = \frac{\frac{q}{2} - 0.4375 + \hat{\sigma}^2 + \sqrt{(\frac{q}{2} + 0.4375 - \hat{\sigma}^2)^2 + 2.25q}}{2(1 - \hat{\sigma}^2)} r \approx \frac{\hat{\sigma}^2 - 0.4375}{1 - \hat{\sigma}^2} r,
\]

where we used \( r \gg q \) and assumed that \( \hat{\sigma}^2 > 0.4375 \) and \( \hat{\sigma}^2 < 1 \). If the lower bound is not satisfied it turns out that the optimal feedback gain is given by \( \hat{K}^* \approx 0 \).

The optimal linear controller associated with this solution is given by:

\[
\hat{K}^* = -\frac{0.75 \hat{p}}{1 + \frac{\hat{p}}{r}} \approx -\frac{\hat{\sigma}^2 - 0.4375}{0.75},
\]

(6)

The closed-loop system given this controller is mean square stable iff \( (0.75 + \hat{K}^*)^2 + 0.5\hat{K}^* < 1 \), which follows from the recursive expression for the variance of the state in closed-loop:

\[
\mathbb{E} \left[ ((0.75 + \hat{K}^*) + \hat{K}^* w)x_k \right] = ((0.75 + \hat{K}^*)^2 + \hat{\sigma}^2 \hat{K}^*) \mathbb{E} \left[ x_k^2 \right].
\]

Solving this stability condition for \( \hat{K}^* \) leads to the conclusion that the system is mean square stable iff

\[
-1.4571 < \hat{K}^* < -0.0429.
\]

Filling in (6) results in a condition on \( \hat{\sigma}^2 \):

\[
0.4697 < \hat{\sigma}^2 < 1.5303.
\]

Note that \( \hat{\sigma}^2 > 1 \) implies that the system is not stabilizable (this is related to the uncertainty threshold principle of Athans et al. (1977)), so we only consider the lower bound on \( \hat{\sigma}^2 \) (which is larger than 0.4375 from our earlier assumption). In fact, we can now evaluate the probability that the empirical approach provides an unstable closed-loop controller as

\[
\mathbb{P} \left[ \hat{\sigma}^2 < 0.4697 \right] = \mathbb{P} \left[ \sum_{i=1}^{M-1} \xi_i^2 < \frac{0.4697 M}{\hat{\sigma}^2} \right],
\]
with $\hat{\xi}_i = \hat{w}_i/\sigma \sim \mathcal{N}(0, 1)$ and which corresponds to the cumulative distribution function of a $\chi^2$-distributed random variable with $M$ degrees of freedom, since we assumed that $w$ is Gaussian.

Evaluating this probability numerically, we find that with $M = 500$, there is a probability of $0.1693$ that the empirical approach provides an unstable closed-loop controller.

From this example, it is clear that underestimation of the variance of the disturbance is directly related to the probability of failure of the controller. In order to take this into account, we introduce ambiguity sets which represent the uncertainty of estimators $\hat{\mu}$ and $\hat{\Sigma}$. In particular, we use sets of the following form, as first suggested by Delage and Ye (2010):

$$\hat{A} := \left\{ \mathbb{P}_v \in \mathcal{M} \left| \begin{array}{c} (\mathbb{E}[v] - \hat{\mu})^\top \hat{\Sigma}^{-1} (\mathbb{E}[v] - \hat{\mu}) \leq r_\mu^2 \\ \mathbb{E}[(v - \mu)(v - \mu)^\top] \leq r_\Sigma \hat{\Sigma} \end{array} \right\} \right., \quad (7)$$

where $\mathcal{M}$ is the set of probability measures defined on $(\mathbb{R}^n, \mathcal{B})$, with $\mathcal{B}$ the Borel $\sigma$-algebra of $\mathbb{R}^n$. Note that since $\hat{A}$ depends on $\hat{\Sigma}$ and $\hat{\mu}$, it is a random variable. The values of $r_\mu$ and $r_\Sigma$ such that $\mathbb{P}(\mathbb{P}_v \in \hat{A}) \geq 1 - \beta$ are derived in Section 3. In Section 4, the following DR counterpart of (2) is solved

$$\begin{array}{c} \text{minimize} \quad \max_{\mathbb{P}_v \in \hat{A}} \mathbb{E} \left[ \sum_{k=1}^{\infty} x_k^\top Q x_k + u_k^\top R u_k \right] \\ \text{subj. to} \quad x_{k+1} = A(v_k)x_k + B(v_k)u_k, \quad k \in \mathbb{N} \\ x_0 = \bar{x} \end{array} \quad (8)$$

where $\{v_k\}_{k \in \mathbb{N}}$ is a trajectory of i.i.d. copies of $v$. In doing so, we can finally establish m.s.s. of the data-driven controller with high probability, by virtue of the following generalization of Theorem 2 to the DR case.

**Theorem 4 (DR Lyapunov stability)** Consider random matrices $\{\hat{A}_i\}_{i=0}^M$ and the autonomous system $x_{k+1} = \hat{A}(w_k)x_k$, with $\hat{A}(w) = \sum_{i=1}^{M} \hat{A}_i w_i$. Say that we have an ambiguity set with $\mathbb{P}(\mathbb{P}_w \in \hat{A}) \geq 1 - \beta$, with $\mathbb{P}_w$ the true distribution of $w$. Then if there exists a $P > 0$ such that:

$$P - \max_{\mathbb{P}_v \in \hat{A}} \mathbb{E}[A(v)^\top PA(v)] > 0, \quad (9)$$

the autonomous system is e.m.s.s. with probability at least $1 - \beta$.

**Proof** See Appendix B.1.

3. Data-driven ambiguity set estimation

We now turn to the problem of estimating the parameters $r_\Sigma$ and $r_\mu$ involved in the definition of the ambiguity set (7), given that we have $M$ i.i.d. draws from the true distribution. These parameters will be estimated under the following assumption on the disturbances.

**Definition 5 (Sub-Gaussianity)** A random variable $y$ is sub-Gaussian with variance proxy $\sigma^2$ if $\mathbb{E}[y] = 0$ and its moment generating function satisfies

$$\mathbb{E}[\exp(\lambda y)] \leq \exp \left( \frac{\sigma^2 \lambda^2}{2} \right) \quad \forall \lambda \in \mathbb{R}. \quad (10)$$
We denote this by $y \sim \text{subG}(\sigma^2)$. We say that a random vector $\xi \in \mathbb{R}^{n_w}$ is sub-Gaussian, or $\xi \sim \text{subG}_{n_w}(\sigma^2)$, if $z^\top \xi \sim \text{subG}(\sigma^2)$, $\forall z \in \mathbb{R}^{n_w}$ with $\|z\|_2 = 1$.

**Assumption 1** We assume that (i) $w$ is square integrable; (ii) $w_k$ and $w_l$ are independent for all $k \neq l$; (iii) $\Sigma > 0$; and (iv) $\Sigma^{-1/2}(w_k - \mu) \sim \text{subG}_{n_w}(\sigma^2)$ for some $\sigma \geq 1$.

Note that in the specific case of Gaussian disturbances, Assumption 1(iv) holds with $\sigma^2 = 1$ and so no further prior knowledge on the distribution is required. Moreover, in this case the bound on the covariance obtained in Theorem 6 can be slightly improved (Wainwright, 2019). In the case of bounded disturbances, $\sigma^2$ can be estimated in a data-driven fashion (Delage and Ye, 2010).

For the moment, we restrict our attention to obtaining concentration inequalities for moment estimators of random vectors with zero mean and unit variance — hereafter referred to as isotropic random vectors. We will then convert these results into ambiguity sets of the form (7) using arguments from Delage and Ye (2010); So (2011). We begin by specializing the isotropic covariance bound, derived with constants in (Hsu et al., 2012a, Lemma A.1) based on a result by Litvak et al. (2005) and the isotropic mean bound by (Hsu et al., 2012b, Theorem 2.1).

**Theorem 6 (Isotropic covariance bound)** Let $\xi \sim \text{subG}_{n_w}(\sigma^2)$ be a random vector, with $\mathbb{E}[\xi] = 0$, $\mathbb{E}[\xi^\top] = I_{n_w}$. Let $\{\xi_i\}_{i=0}^{M-1}$ be $M$ independent copies of $\xi$ and $\hat{I} := \frac{1}{M} \sum_{i=0}^{M-1} \hat{\xi}_i \hat{\xi}_i^\top$. Then

$$\mathbb{P}[\| \hat{I} - I_{n_w} \|_2 \leq t_\Sigma(\beta)] \geq 1 - \beta,$$

with

$$t_\Sigma(\beta) := \frac{\sigma^2}{1 - 2\epsilon} \left( \sqrt{\frac{32q(\beta, \epsilon, n_w)}{M}} + \frac{2q(\beta, \epsilon, n_w)}{M} \right),$$

where $\epsilon \in (0, 1/2)$ is chosen freely and $q(\beta, \epsilon, n_w) := n_w \log (1 + 1/\epsilon) + \log (2/\beta)$.

**Theorem 7 (Isotropic mean bound)** Let $\{\hat{\xi}_i\}_{i=0}^{M-1}$ be as defined in Theorem 6, and $\hat{\zeta} := \frac{1}{M} \sum_{i=0}^{M-1} \hat{\xi}_i$. Then

$$\mathbb{P}[\| \hat{\zeta} \|_2 \leq t_\mu(\beta)] \geq 1 - \beta,$$

where

$$t_\mu(\beta) := \frac{\sigma^2}{M} p(\beta, n_w),$$

with $p(\beta, n_w) := \left( n_w + 2\sqrt{n_w \log (1/\beta)} + 2 \log (1/\beta) \right)$.

By combining the bounds in Theorems 6 and 7, we readily obtain the following result.

**Theorem 8 (Ambiguity set)** Let $w \in \mathbb{R}^{n_w}$ be a sub-Gaussian random vector, with $\mathbb{E}[w] = \mu$, $\mathbb{E}[(w - \mu)(w - \mu)^\top] = \Sigma$ and $\xi \sim \Sigma^{-1/2}(w - \mu) \sim \text{subG}_{n_w}(\sigma^2)$. Let $\{\hat{w}_i\}_{i=0}^{M-1}$ be independent copies of $w$. Let $\hat{\mu} := \frac{1}{M} \sum_{i=0}^{M-1} \hat{w}_i$ and $\hat{\Sigma} := \frac{1}{M} \sum_{i=0}^{M-1} (\hat{w}_i - \hat{\mu})(\hat{w}_i - \hat{\mu})^\top$ denote the empirical estimators for the mean and the covariance matrix, respectively. Let $\epsilon, p(\beta, n_w), q(\beta, \epsilon, n_w), t_\Sigma(\beta/2)$ and $t_\mu(\beta/2)$ be as defined in Theorems 6 and 7. Provided that

$$M > \left( \frac{\sigma^2 \sqrt{32q(\beta/2, \epsilon, n_w)} + \sqrt{32\sigma^4 q(\beta/2, \epsilon, n_w) + 8\sigma^2 (1 - 2\epsilon) q(\beta/2, \epsilon, n_w) + 4\epsilon^2 (1 - 2\epsilon)^2 p(\beta/2, n_w)}}{2(1 - 2\epsilon)} \right)^2,$$

(12)
then with probability at least $1 - \beta$,

$$(\hat{\mu} - \mu)^\top \Sigma^{-1}(\hat{\mu} - \mu) \leq r_\mu,$$

with $r_S := \frac{1}{1 - t_\mu(\beta/2) - t_\Sigma(\beta/2)}$ and $r_\mu := \frac{t_\mu(\beta/2)}{1 - t_\mu(\beta/2) - t_\Sigma(\beta/2)}$.

**Proof** Let us define $\xi = \Sigma^{-1/2}(w - \mu) \sim \text{subG}_{nw}(\sigma^2)$ so that

$$\mathbb{E}[\xi] = \Sigma^{-1/2}(\mathbb{E}[w] - \mu) = 0,$$

$$\mathbb{E}[\xi^\top] = \Sigma^{-1/2}\mathbb{E}[(w - \mu)(w - \mu)^\top] \Sigma^{-1/2} = \Sigma^{-1/2}\Sigma \Sigma^{-1/2} = I_{nw}.$$

Let $\hat{\xi} = \frac{1}{M} \sum_{i=1}^M \hat{\xi}_i$ and $\hat{I} = \frac{1}{M} \sum_{i=1}^M \hat{\xi}_i \hat{\xi}_i^\top$. By Theorems 6 and 7, we then have with probability $1 - \beta$ that,

$$\|\hat{I} - I_{nw}\|_2 \leq t_\Sigma(\beta/2)$$

and

$$\|\hat{\xi}\|_2^2 \leq t_\mu(\beta/2).$$

Let us define the covariance estimator with respect to the true mean by

$$\hat{\Sigma} := \frac{1}{M} \sum_{i=1}^M (\hat{w}_i - \mu)(\hat{w}_i - \mu)^\top = \frac{1}{M} \sum_{i=1}^M (\Sigma^{1/2} \xi_i)(\Sigma^{1/2} \xi_i)^\top = \Sigma^{1/2} \hat{I} \Sigma^{1/2}.$$  \hfill (14)

Therefore, we can rewrite (13a) in terms of the anisotropic random variable $w$ as

$$\|\hat{I} - I_{nw}\|_2 \leq t_\Sigma(\beta/2) \iff (1 - t_\Sigma(\beta/2))I_{nw} \leq \hat{I} \leq (1 + t_\Sigma(\beta/2))I_{nw},$$

$$\Rightarrow (1 - t_\Sigma(\beta/2)) \Sigma \leq \hat{\Sigma} \leq (1 + t_\Sigma(\beta/2)) \Sigma,$$  \hfill (15)

where we have used (14) and the fact that condition (12) implies $t_\Sigma(\beta/2) < 1$. As shown in (Delage and Ye, 2010, Thm. 2), (13b) implies that for any $x \in \mathbb{R}^{nw}_{I_{nw}}$

$$x^\top (\hat{\mu} - \mu)(\hat{\mu} - \mu)^\top x = (x^\top (\hat{\mu} - \mu))^2 \leq \|\Sigma^{1/2}x\|_2^2 \|\Sigma^{-1/2}(\hat{\mu} - \mu)\|_2^2$$

$$\leq t_\mu(\beta/2) x^\top \Sigma x.$$  \hfill (16)

Using this fact, we can bound $\hat{\Sigma}$ with respect to $\hat{\Sigma}$ and $\Sigma$ as

$$\hat{\Sigma} = \frac{1}{M} \sum_{i=1}^M (w_i - \mu)(w_i - \mu)^\top$$

$$= \frac{1}{M} \sum_{i=1}^M (w_i - \hat{\mu} + \hat{\mu} - \mu)(w_i - \hat{\mu} + \hat{\mu} - \mu)^\top$$

$$= \frac{1}{M} \sum_{i=1}^M (w_i - \hat{\mu})(w_i - \hat{\mu})^\top + (w_i - \hat{\mu})(\hat{\mu} - \mu)^\top + (\hat{\mu} - \mu)(w_i - \hat{\mu})^\top + (\hat{\mu} - \mu)(\hat{\mu} - \mu)^\top$$

$$= \hat{\Sigma} + (\hat{\mu} - \mu)(\hat{\mu} - \mu)^\top \leq \hat{\Sigma} + t_\mu(\beta/2) \Sigma,$$

where we used (16) in the final step. Combining this with (15), we have that $(1 - t_\Sigma) \Sigma \leq \hat{\Sigma} + t_\mu \Sigma$, thus

$$\Sigma \leq \frac{1}{1 - t_\mu(\beta/2) - t_\Sigma(\beta/2)} \hat{\Sigma},$$

$$(1 - t_\mu - t_\Sigma)(\hat{\mu} - \mu)^\top \hat{\Sigma}^{-1}(\hat{\mu} - \mu) \leq (\hat{\mu} - \mu)^\top \Sigma^{-1}(\hat{\mu} - \mu) = \xi^\top \xi \leq t_\mu.$$

Condition 12 then follows from assuming $1 - t_\mu - t_\Sigma > 0$, which is a quadratic inequality in $\sqrt{M}$. □
4. Distributionally Robust LQR

We will tackle the solution of (2) for the ambiguity given in (7) in two stages. First we extend the result of Proposition 3 to the case where \( \mu \) is known and \( \Sigma \) is estimated, i.e., \( r_\mu = 0 \) and \( r_\Sigma > 0 \). In the second part we develop the general case where both the mean and the covariance are estimated.

4.1. Uncertain covariance

The case where the mean is known is interesting since we can still formulate an exact solution to (8). This will no longer be true for the full-uncertainty case (Section 4.2).

Proposition 9  Consider that \( v \in \mathbb{R}^n_w \) is distributed according to an element of the set

\[
\hat{A}_\Sigma := \left\{ \mathbb{P}_v \in \mathcal{M} \mid \mathbb{E} \left[ (v - \mu)(v - \mu)^\top \right] \preceq r_\Sigma \hat{\Sigma}, \mathbb{E} [v] = \mu \right\},
\]

where \( \mathbb{P}(\mathbb{P}_v \in \hat{A}_\Sigma) \geq 1 - \beta \). Then applying Proposition 3 with \( \Sigma = r_\Sigma \hat{\Sigma} \) results in the optimal linear controller for (8), assuming that (1) is DR mean square stabilizable, i.e., there exists a \( K \) such the DR Lyapunov decrease (9) holds for the closed-loop system \( x_{k+1} = (A(w_k) + B(w_k)K)x_k \).

The optimal controller is also mean square stabilizing for (1) with probability at least \( 1 - \beta \).

Proof  See Appendix B.2.

4.2. Full uncertainty

We finally consider the more general case using the full ambiguity set \( \hat{A} \) given by (7). The general min-max problem (8) for such sets is computationally intractable, which is why an upper bound on the quadratic cost is minimized instead by employing results from robust control (Boyd et al., 1994; Kothare et al., 1996). A common approach is to assume that the value function can be written in the quadratic form \( V(x) = x^\top P x \) for some \( P > 0 \) and to solve the optimization problem

\[
\begin{align*}
\text{minimize} & \quad \mathbb{E} [V(\tilde{x})] \\
\text{subj. to} & \quad V(x) \geq \min_v \left\{ x^\top Q x + u^\top R u + \max_{\mathbb{P}_v \in \hat{A}} \mathbb{E} [V(A(v)x + B(v)u)] \right\}, \quad \forall x,
\end{align*}
\]

where we introduced the random initial state \( \tilde{x} \in \mathbb{R}^{n_x} \). The optimal cost of (18) then upper bounds the true LQR cost of (8) for a given value of \( \tilde{x} \) as proven in Kothare et al. (1996) for a similar setup.

We can then write (18) as an SDP using the following theorem.

Theorem 10  Let \( \hat{A} \) be an ambiguity set of the form (7). Then we can find an approximate solution of (18) for the system (1), assuming that the initial state is given by the random vector \( \tilde{x} \in \mathbb{R}^{n_x} \).
with $\mathbb{E} [\bar{x}] = 0$ and $\mathbb{E} [\bar{x} \bar{x}^\top] = I$, by solving the following SDP:

$$\begin{align*}
\text{maximize} & \quad \text{Tr} \ W \\
\text{subject to} & \quad \begin{bmatrix}
S & r_\mu H_1^\top & r_\mu H_2^\top & \cdots & r_\mu H_n^\top \\
r_\mu H_1 & L & & & \\
r_\mu H_2 & & L & & \\
\vdots & & & \ddots & \ddots \\
r_\mu H_n & & & & L
\end{bmatrix} \succeq 0, \quad (19a) \\
& \quad \begin{bmatrix}
W - \sqrt{\beta} S & (AW + BV)^\top (\hat{A} W + \hat{B} V) & W^\top Q^\top & V^\top R^\top \\
AW + BV & \hat{\Sigma}_{dr}^{-1} \otimes W & W - \sqrt{\beta} L & \\
\hat{A} W + BV & W - \sqrt{\beta} L & \hat{\Sigma}_{dr}^{-1} \otimes W & \\
Q^\top W & I_d & & I_n
\end{bmatrix} \succeq 0, \quad (19b)
\end{align*}$$

where $H_i = \sum_{j=1}^{n_{w}} [\hat{\Sigma}^{1/2}]_{ji} (A_j W + B_j V)$, $\hat{A} = A(\hat{\mu})$, $\hat{B} = B(\hat{\mu})$, $\hat{\Sigma}_{dr} = \hat{\Sigma}_{\Sigma}$. Let $\hat{W}$ and $\hat{V}$ denote the minimizers of (19). The corresponding linear controller $u = \hat{K} \bar{x}$ with $\hat{K} = \hat{V} \hat{W}^{-1}$ then achieves an upper bound of the cost (18), given by $\mathbb{E}[\bar{x}^\top \hat{P} \bar{x}]$, where $\hat{P} = \hat{W}^{-1}$. Moreover, $\hat{K}$ is mean-square stabilizing for (1) with probability at least $1 - \beta$.

Proof See Appendix B.3.

**Remark 11** Two approximations are made in Theorem 10. First, leveraging (Ben-Tal et al., 2000, Theorem 6.2.1), introducing an approximation error quantified by an increase of $r_\mu$ by a factor of at most $\sqrt{n_w}$. Secondly, we minimize an upper-bound of the LQR cost instead of the cost itself. We can further decrease the closed-loop cost by instead using a receding horizon controller for a given $\bar{x}$, which too can be formulated as an SDP. Then (18) is reformulated as (Kothare et al., 1996):

$$\begin{align*}
\text{minimize} & \quad \gamma \\
\text{subject to} & \quad \bar{x}^\top \hat{W}^{-1} \bar{x} \leq \gamma \quad (20a), (19a), (19b),
\end{align*}$$

where (20a) can be replaced by a linear matrix inequality using Schur’s complement (Boyd et al., 1994, Sec. 2.1). The assumption used in Theorem 10 ensures that the solution converges to the optimal one as $r_\mu$ and $r_{\Sigma}$ go to zero (Balakrishnan and Vandenberghe, 2003).

**Remark 12** Invertibility of $r_{\Sigma} \hat{\Sigma}$ can be guaranteed by using $r_{\Sigma} \hat{\Sigma} + \epsilon I$ instead of $r_{\Sigma} \hat{\Sigma}$ for a small value of $\epsilon$, which introduces additional conservatism by increasing the size of the ambiguity set.

5. Numerical Experiment

We experimentally quantify the sample complexity of our approach, i.e., how many samples are needed before the controller becomes equivalent to the nominal one based on the true $\Sigma$ and $\mu$ instead of their data-driven estimates. Consider the double integrator model with matrices:

$$A_0 = \begin{bmatrix} 1 & T_s \\ 0 & 1 - 0.4 T_s \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 \\ T_s \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 \\ 0 & -T_s \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 \\ T_s \end{bmatrix}.$$
where we chose $T_s = 0.02$. The dynamics are then given by (1) with $w_k$ an independent random sequence of Gaussian random vectors with covariance $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and mean $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

To estimate the sample complexity, we determine controllers satisfying (8) for $Q = \begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}$ and $R = 0.01$. The parameters of the ambiguity set (7) are determined for $\beta = 0.05$ and $\epsilon = 1/30$. Since $w_k$ are Gaussian random vectors we have that $\xi_k = \Sigma^{-1/2}(w_k - \mu)$ are sub-Gaussian with $\sigma^2 = 1$.

The simulation setup is as follows. We compare the nominal controller, the uncertain covariance controller produced using Proposition 9, assuming $t_\mu = 0$ and using $r_\Sigma = \frac{1}{1-t_\Sigma(\beta)}$ with $t_\Sigma(\beta)$ as defined in Theorem 6, and the full uncertainty controller produced by Theorem 10. We run 500 simulations of 100 time-steps starting from $\bar{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}^\top$. We start with $M = 500$ to satisfy (12) which requires $M > 488$. We run each closed-loop system using the same realizations of $w_k$.

Figure 1 depicts confidence intervals for the cost for all three controllers. Note that even though Theorem 10 only solves (8) approximately, it converges to the nominal performance as well.

6. Conclusion and future work

We studied the infinite horizon LQR problem for systems with multiplicative uncertainty on both the states and the inputs. We extend existing results for disturbances with known distributions to the more challenging setting in which the distributions are estimated from data. We show that using results from high-dimensional statistics, high-confidence ambiguity sets can be constructed, which allow us to formulate a DR counterpart to the stochastic optimal control problem. As a result, stability of the closed-loop system can be guaranteed with high probability. Furthermore, we show that the control synthesis problems can be stated as SDPs, and can therefore be solved efficiently.

In future work, we aim to perform an in-depth analysis of the conservatism introduced by the proposed formulations. Furthermore, we plan to study extensions towards DR Kalman filtering.

Acknowledgments

This work was supported by the Ford KU Leuven Research Alliance; the Research Foundation Flanders (FWO) PhD grant 11E5520N and research projects G086518N and G086318N; Research Council KU Leuven C1 project No. C14/18/068; Fonds de la Recherche Scientifique FNRS and the Fonds Wetenschappelijk Onderzoek Vlaanderen under EOS project no 30468160 (SeLMA).

References

Michael Athans, Richard Ku, and Stanley B. Gershwin. The uncertainty threshold principle: Some fundamental limitations of optimal decision making under dynamic uncertainty. *IEEE Transactions on Automatic Control*, 22(3):491–495, June 1977.
Venkataramanan Balakrishnan and Lieven Vandenberghe. Semidefinite programming duality and linear time-invariant systems. *IEEE Transactions on Automatic Control*, 48(1):30–41, 2003.

Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. Robustness. In *Handbook of Semidefinite Programming*, pages 139–162. Springer US, Boston, MA, 2000.

Patrick Billingsley. *Probability and measure*. Wiley series in probability and mathematical statistics. Wiley, New York, 3rd edition edition, 1995.

Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics, 1994.

Christopher I. Byrnes. On the stabilizability of linear control systems depending on parameters. In *Proceedings of the IEEE Conference on Decision and Control*, volume 1, pages 233–236. IEEE, 1979.

Oswaldo L.V. Costa and Carlos S. Kubrusly. Quadratic optimal control for discrete-time infinite-dimensional stochastic bilinear systems. *IMA Journal of Mathematical Control and Information*, 14(4):385–399, 1997.

Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. On the Sample Complexity of the Linear Quadratic Regulator. *Foundations of Computational Mathematics*, 2019.

Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612, 2010.

Jitka Dupaov. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20(1):73–88, 1987.

Laurent El Ghaoui. State-feedback control of systems with multiplicative noise via linear matrix inequalities. *Systems and Control Letters*, 24(3):223–228, 1995.

Daniel Hsu, Sham Kakade, and Tong Zhang. Tail inequalities for sums of random matrices that depend on the intrinsic dimension. *Electronic Communications in Probability*, 17(52):1–13, 2012a.

Daniel Hsu, Sham M. Kakadey, and Tong Zhang. Tail inequalities for sums of random matrices that depend on the intrinsic dimension. *Electronic Communications in Probability*, 17(14), 2012b.

Mayuresh V. Kothare, Venkataramanan Balakrishnan, and Manfred Morari. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32(10):1361–1379, 1996.

Alexander E. Litvak, Alain Pajor, M. Rudelson, and N. Tomczak-Jaegermann. Smallest singular value of random matrices and geometry of random polytopes. *Advances in Mathematics*, 195(2):491–523, 2005.

Toader Morozan. Stabilization of some stochastic discrete-time control systems. *Stochastic Analysis and Applications*, 1(1):89–116, 1983.
Mathijs Schuurmans, Pantelis Sopasakis, and Panagiotis Patrinos. Safe learning-based control of stochastic jump linear systems: a distributionally robust approach. In 2019 IEEE Conference on Decision and Control (CDC), pages 6498–6503. IEEE, 2019.

Anthony Man Cho So. Moment inequalities for sums of random matrices and their applications in optimization. Mathematical Programming, 130(1):125–151, 2011.

Martin J. Wainwright. High-Dimensional Statistics. Cambridge University Press, 2019.

W. Murray Wonham. Optimal Stationary Control of a Linear System with State-Dependent Noise. SIAM Journal on Control, 5(3):486–500, 1967.

Fen Wu, Xin Hua Yang, Andy Packard, and Greg Becker. Induced L2-norm control for LPV systems with bounded parameter variation rates. International Journal of Robust and Nonlinear Control, 6(9-10):983–998, 1996.

Appendix A. Proofs for nominal case

A.1. Proof of Theorem 2

(Morozan, 1983, Lemma 1) states that the following is a necessary and sufficient condition for m.s.s.:

\[ P - \mathbb{E} [A(w)^\top PA(w)] > 0. \]

The expected value is of the form \( \mathbb{E}[v^\top Pv] \), with \( v \) a random vector. We will leverage the following, easily verified result:

\[
\mathbb{E}[v^\top Pv] = \text{Tr} (P\mathbb{E}[\tilde{v}\tilde{v}^\top]) + \mathbb{E}[v^\top] P\mathbb{E}[v],
\]

(21)

with \( \tilde{v} = v - \mathbb{E}[v] \). Taking \( v = A(w)x \) for a given \( x \) results in:

\[
\mathbb{E}[v] = A(\mu)x,
\]

\[
\mathbb{E}[\tilde{v}\tilde{v}^\top] = \mathbb{E}[(\tilde{w}^\top \otimes I_d)Axx^\top A^\top(\tilde{w} \otimes I_d)] = \sum_{i,j=1}^{n_w} \left( \Sigma_{ij} A_i x x^\top A_j^\top \right),
\]

with \( \tilde{w} = w - \mu \). Applying (21) then gives:

\[
\mathbb{E}_{P,w} [x^\top A(w)^\top PA(w)x] = \text{Tr} \left( P \sum_{i,j=1}^{n_w} \left( \Sigma_{ij} A_i x x^\top A_j^\top \right) \right) + x^\top A(\mu)^\top PA(\mu)x
\]

\[
= \text{Tr} \left( \sum_{i,j=1}^{n_w} \left( \Sigma_{ij} x^\top A_j^\top PA_i x \right) \right) + x^\top A(\mu)^\top PA(\mu)x
\]

\[
= x^\top A^\top (\Sigma \otimes P)Ax + x^\top A(\mu)^\top PA(\mu)x
\]

\[
= x^\top A_0^\top (\Sigma_0 \otimes P)A_0x.
\]

By (Morozan, 1983, Lemma 1) it follows that m.s.s. is equivalent to e.m.s.s. for system (1).
A.2. Proof of Proposition 3

We will follow the proof of (Morozan, 1983, Theorem 1). We look at properties of the Bellman operator associated with (2):

\[
(TV_k)(x) = \min_u \{ x^\top Qx + u^\top Ru + \mathbb{E}[V_k(A(w)x + B(w)u)] \},
\]

(22)

Assuming a quadratic value function of the form \( V_k(x) = V(x) = x^\top P_k x, \forall k \in \mathbb{N} \), we can for all \( x \in \mathbb{R}^{nx} \), write the \( k + 1 \)th step of value iteration with \( P_0 = 0 \) as

\[
x^{\top}P_{k+1}x = \min_u \{ x^\top Qx + u^\top Ru + (A_0x + B_0u)^\top(\Sigma_0 \otimes P_k)(A_0x + B_0u) \}
\]

\[
= x^\top(Q + K_k^\top RK_k)x + x^\top(A_0 + B_0K_k)^\top(\Sigma_0 \otimes P_k)(A_0 + B_0K_k)x
\]

\[
\Leftrightarrow P_{k+1} = Q + F(P_k) - H(P_k)^\top(R + G(P_k))^{-1}H(P_k),
\]

(23a)

(23b)

where the first equality follows from the same arguments as those in the proof of Theorem 2. The second equality follows from the fact that the optimal value of \( u \) in (22) is given by \( u = K_kx \) with \( K_k = -(R + G(P_k))^{-1}H(P_k) \).

Statement I follows from the following statements, which we will prove separately: (i) value iteration converges to some \( P_\infty \); (ii) Given an initial state \( x_0, x_0^\top P_\infty x_0 \) is the optimal cost of (2), realized by the state feedback law \( u = K_\infty x \); and (iii) \( P_\infty \) is the unique solution to (4).

Since there exists a stabilizing controller \( u = Kx \), the optimal cost of (2) is bounded above. Indeed, we can write the closed-loop stage cost at time \( k \) equivalently as

\[
\mathbb{E}[x_k^\top (Q + K^\top RK) x_k] = \text{Tr} \left( (Q + K^\top RK) \mathbb{E}[x_k x_k^\top] \right)
\]

(24)

and by definition of e.m.s.s., \( \mathbb{E}[x_k x_k^\top] \leq c \gamma^k ||x_0||, \forall k \in \mathbb{N} \), with \( c > 0 \) and \( \gamma \in (0, 1) \), which implies that (24) is summable. Furthermore, positive definiteness of \( Q \) and \( R \) guarantees that \( \{P_k\}_{k \in \mathbb{N}^+} \) is a monotone increasing sequence (i.e., \( P_{k+1} \succeq P_k \), \( \forall k \in \mathbb{N}^+ \)) of positive definite matrices if \( P_0 = 0 \). Therefore the sequence \( \{P_k\}_{k \in \mathbb{N}} \) converges to some \( P_\infty > 0 \), proving (i).

Next we prove that \( K_\infty \) is the optimal controller. To do so, we write the Riccati equation for states \( x_k \) and \( x_{k+1} \) at subsequent time steps as (see (23a)):

\[
x_k^\top P_\infty x_k = x_k^\top (Q + K_\infty^\top RK_\infty)x_k + \mathbb{E}[x_{k+1}^\top P_\infty x_{k+1} | x_k]
\]

\[
x_{k+1}^\top P_\infty x_{k+1} = x_{k+1}^\top (Q + K_\infty^\top RK_\infty)x_{k+1} + \mathbb{E}[x_{k+2}^\top P_\infty x_{k+2} | x_{k+1}],
\]

(25)

where \( x_{k+1} = A(w_k)x_k + B(w_k)K_\infty x_k \). Taking the expected value of both sides of the equalities in (25), noting that \( \mathbb{E}[x_{k+1}^\top P_\infty x_{k+1} | x_k] = \mathbb{E}[x_{k+1}^\top P_\infty x_{k+1}] \) and summing the equalities for \( k = 0, \ldots, N - 1 \) allows us to write:

\[
x_0^\top P_N x_0 \leq \mathbb{E} \sum_{k=0}^{N-1} [x_k^\top (Q + K_\infty^\top RK_\infty)x_k] = x_0^\top P_\infty x_0 - \mathbb{E}[x_N^\top P_\infty x_N]
\]

\[
\leq x_0^\top P_0 x_0, \quad \forall N \in \mathbb{N}, \forall x_0 \in \mathbb{R}^{nx},
\]

(26)

with \( x_N \) produced by running (1) for \( u_k = K_\infty x_k \) for \( N \) time steps, starting from some \( x_0 \). The first inequality in (26) follows from dynamic programming, since we know that for the finite horizon case \( P_N \) describes the optimal cost and the final inequality follows from \( P_\infty > 0 \). Taking the limit of \( N \to \infty \) we have that \( x_0^\top P_\infty x_0 \) is the closed-loop cost for \( u_k = K_\infty x_k \) and that \( P_\infty \) and \( K_\infty \) describe the optimal cost and the optimal controller respectively, proving (ii).
Notice that (23a) implies the Lyapunov condition (3) since $R > 0$ and $Q > 0$. This holds for any solution of the Riccati equation (4), proving II.

Note that by (23b), $P_\infty$ satisfies the Riccati equation (4). Since (26) holds for any $\hat{P}_\infty$ that satisfies (4) we can see that $P_\infty \preceq \hat{P}_\infty$ when we take $N \to \infty$, since $E[x_N \hat{P}_\infty x_N] \to 0$ by definition of $m.s.s$. We use (23a) to write:

$$x_k^T \hat{P}_\infty x_k \leq x_k^T (Q + K_k^T R K_k) x_k + E\left[ x_{k+1}^T \hat{P}_\infty x_{k+1} \mid x_k \right],$$

(27)

where $x_{k+1} = A(w_k)x_k + B(w_k)K_kx_k$. The inequalities follow from the fact that we did not use the optimal controller $K_\infty = -(R + G(\hat{P}_\infty))^{-1} H(\hat{P}_\infty)$, instead we use $K_k$, which denote the finite horizon optimal controllers. Then summing (27) similarly to what we did for (25) results in:

$$\bar{x}_0^T \hat{P}_\infty x_0 \leq \bar{x}_0^T P_N x_0 + E\left[ x_N \bar{P}_\infty x_N \right]$$

(28)

where $\bar{x}_{k+1} = A(w_k) + B(w_k)\hat{K}_\infty$. Then taking $N \to \infty$ in (28) implies $E[\bar{x}_N \bar{P}_\infty x_N] \to 0$ as before. So $\hat{P}_\infty \preceq P_\infty$ and we know $P_\infty \preceq \hat{P}_\infty$ from earlier. Therefore $\hat{P}_\infty = \bar{P}_\infty$, proving (iii).

We can use the arguments by Balakrishnan and Vandenberghe (2003) to show that the solution (4) is obtained by solving (5), proving III. This follows from the complementary slack condition.

Appendix B. Proofs for distributionally robust case

B.1. Proof of Theorem 4

The Lyapunov inequality (9) directly implies

$$P - E\left[ \hat{A}(v)^T P \hat{A}(v) \right] > 0, \quad \forall \mathbb{P}_v \in \hat{A}.$$  (29)

Since $P(\mathbb{P}_w \in \hat{A}) \geq 1 - \beta$ we have that (3) holds with probability at least $1 - \beta$, proving the required result.

B.2. Proof of Proposition 9

The Bellman operator associated with (8) is given by:

$$(TV_{k+1})(x) = \min_a \left\{ x^T Q x + u^T R u + \max_{\mathbb{P}_v \in \hat{A}_k} E[V_k(A(v)x + B(v)u)] \right\}. \quad (30)$$

We can evaluate the expectation as in the proof of Theorem 2, resulting in a similar statement to (23a). After extracting the terms that are independent of $\Sigma$ from the maximum and grouping the remaining ones together, only the following needs to be evaluated:

$$\max_{\Sigma \leq r \Sigma} \Sigma \otimes P \Sigma \otimes P\Sigma \Sigma \otimes P\Sigma$$

This is equivalent to:

$$\max_{\Sigma \leq r \Sigma} \text{Tr} \begin{bmatrix} z_1^T P z_1 & \ldots & z_1^T P z_{n_w} \\ \vdots & \ddots & \vdots \\ z_{n_w}^T P z_1 & \ldots & z_{n_w}^T P z_{n_w} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \ldots & \Sigma_{1n_w} \\ \ldots & \ddots & \ldots \\ \Sigma_{n_w1} & \ldots & \Sigma_{n_wn_w} \end{bmatrix} = \max_{\Sigma \leq r \Sigma} \text{Tr} (X \Sigma),$$
where \( z_i = A_i x + B_i u \). Since \( X = Z^T P Z \succeq 0 \), with \( Z = [z_1 \ldots z_n] \). The maximum corresponds to a support function for which one can easily verify using the optimality conditions that:

\[
\max_{\Sigma \succeq r_\Sigma \Sigma} \text{Tr} \left( X \Sigma \right) = \text{Tr} \left( r_\Sigma X \Sigma \right) = r_\Sigma (Ax + Bu)^\top (\hat{\Sigma} \otimes P)(Ax + Bu)^\top.
\]

As such the Bellman operator is still of a similar form to (23a). Therefore the remaining arguments from the proof of Proposition 3 are all applicable using \( r_\Sigma \hat{\Sigma} \) instead of \( \Sigma \). More specifically, we know that the cost is bounded above since there exists a stabilizing \( K \) for \( r_\Sigma \hat{\Sigma} \) by assumption. This means that value iteration will converge to some \( P_\infty \). We can once again telescope the Riccati equation and prove that this \( P_\infty \) and \( K_\infty \) are optimal and the unique solution to the Riccati equation. The arguments of (Balakrishnan and Vandenberghe, 2003) are still applicable as well.

We still need to prove that the resulting controller stabilizes (1) for the true \( \Sigma \) with probability at least \( 1 - \beta \). For this consider the Riccati equation (4) which is equivalent to:

\[
P = \max_{\Sigma \succeq r_\Sigma \Sigma} (A_0 + B_0 \hat{K}_\infty)^\top (\Sigma_0 \otimes P) (A_0 + B_0 \hat{K}_\infty) = Q + \hat{K}_\infty^\top R \hat{K}_\infty \succeq 0,
\]

Due to Theorem 4, \( \hat{K}_\infty \) then stabilizes (1) in the mean square sense with probability at least \( 1 - \beta \).

\[\blacksquare\]

### B.3. Proof of Theorem 10

We need to prove the following statements: (i) constraints (19a) and (19b) are equivalent to the constraint in (18), (ii) the solution of (18) upper bounds the true optimal cost of (8) (iii) the cost of (19) is equivalent to that of (18) (iv) the resulting controller is mean square stabilizing for (1).

Using the quadratic parametrization of \( V(x) \) and applying the same tricks as in the proof of Proposition 9 we can rewrite the constraint of (18) as follows:

\[
P - Q - K^\top R K - (A + B K)^\top (\hat{\Sigma}_{dr} \otimes P) (A + B K) - (A(\mu) + B(\mu) K)^\top P (A(\mu) + B(\mu) K) \succeq 0, \quad \forall \mu \in \mathcal{D},
\]

where we used \( u = K x \) and define \( \mathcal{D} := \{ (\mu - \bar{\mu})^\top \hat{\Sigma}^{-1} (\mu - \bar{\mu}) \leq r^2_m \} \). Pre- and post-multiplying by \( W = P^{-1} \) and setting \( V = KW \) results in:

\[
W - W^\top Q W - V^\top R V - (A W + B V)^\top (\hat{\Sigma}_{dr} \otimes W) (A W + B V)
- (A(\mu) W + B(\mu) V)^\top W^{-1} (A(\mu) W + B(\mu) V), \quad \forall \mu \in \mathcal{D}.
\]

Applying Schur’s complement lemma (Boyd et al., 1994, Sec. 2.1) then gives the following:

\[
\begin{bmatrix}
W
A W + B V
(\hat{A} W + \hat{B} V) + \sum_{i=1}^{n_w} \mu_i F_i
\end{bmatrix}
\begin{bmatrix}
\hat{\Sigma}_{dr}^{-1} \otimes W
W
Q \frac{1}{2} W
R \frac{1}{2} V
I_d
I_u
\end{bmatrix}
\succeq 0, \quad \forall \bar{\mu} \in \bar{\mathcal{D}},
\]

where \( \bar{\mathcal{D}} := \{ \mu^\top \hat{\Sigma}^{-1} \bar{\mu} \leq r^2_m \} \) and \( F_i = A_i W + B_i V \). Defining \( \zeta = \hat{\Sigma}^{-1/2} \bar{\mu} \) allows us to write \( \sum_{i=1}^{n_w} \mu_i F_i = \sum_{i=1}^{n_w} \zeta_i H_i \). Note that, after using this equality, (31) is of the form

\[
U_0 + \sum_{i=1}^{n_w} \delta_i U_i \succeq 0, \quad \forall \delta \in \{ \delta \mid \| \delta \|_2 \leq \rho \}.
\]
Hence we can apply a slightly modified version of (Ben-Tal et al., 2000, Theorem 6.2.1) where we exploit the sparsity of $U_i$, which results in conditions (19a) and (19b). More specifically we have:

\[
\xi^\top (U_0 + \sum_{i=1}^{n_u} \delta_i U_i) \xi = \xi^\top U_0 \xi + 2 \sum_{i=1}^{n_u} \zeta_i \xi_i^\top H_i \xi_i
\]

\[
(\xi = [\xi^0_1, \xi^1_1, \xi^1_2, \xi^2_1, \xi^2_2]^\top)
\]

\[
= \xi^\top U_0 \xi + 2 \xi^\top_L \left[ \sum_{i=1}^{n_u} \zeta_i L^{-1/2} H_i S^{-1/2} \xi_S \right]
\]

\[
(\xi_S = S^{1/2} \xi_1, \xi_L = L^{1/2} \xi_2)
\]

\[
\geq \xi^\top U_0 \xi - 2\|\xi_L\|_2 \left[ \sum_{i=1}^{n_u} |\zeta_i| \|L^{-1/2} H_i S^{-1/2} \xi_S\|_2 \right]
\]

\[
\geq \xi^\top U_0 \xi - 2\|\xi_L\|_2 \sqrt{\sum_{i=1}^{n_u} r_\mu^2 \|L^{-1/2} H_i S^{-1/2} \xi_S\|_2^2} \tag{32a}
\]

\[
= \xi^\top U_0 \xi - 2\|\xi_L\|_2 \sqrt{\sum_{i=1}^{n_u} r_\mu^2 \|L^{-1/2} H_i S^{-1/2} \xi_S\|_2^2} \tag{32b}
\]

where we used $\|\zeta\| \leq r_\mu$ for (32a) and (32b) holds when:

\[
\sum_{i=1}^{n_u} S^{-1/2}(r_\mu H_i^\top) L^{-1} (H_i r_\mu) S^{-1/2} \leq I,
\]

which after pre- and post-multiplying by $S^{1/2}$ and applying Schur’s complement lemma (Boyd et al., 1994, Sec. 2.1) is shown to be equivalent to (19a). The result then holds when:

\[
\xi^\top U_0 \xi \geq 2 \sqrt{(\xi^0_1 \xi_L)(\xi^1_2 \xi_S)}, \tag{33}
\]

for which (19b) is a sufficient condition since:

\[
\sqrt{2}(\xi^0_1 \xi_L + \xi^1_2 \xi_S) \geq 2 \sqrt{(\xi^0_1 \xi_L)(\xi^1_2 \xi_S)}
\]

\[
2 \xi^0_1 \xi_L + 4(\xi^0_1 \xi_L)(\xi^1_2 \xi_S) + 2 \xi^1_2 \xi_S \geq 4(\xi^0_1 \xi_L)(\xi^1_2 \xi_S).
\]

So therefore (19b), which can be written as $\xi^\top U_0 \xi \geq \sqrt{2}(\xi^0_1 \xi_L + \xi^1_2 \xi_S)$, implies (33), proving (i).

Next we will prove that the result of (18) upper bounds the true cost of (8). We introduce $\ell(x) = x^\top (Q + K^\top RK)x$ and $A_K(w) = A(w) + B(w)K$. Then we can use the constraint in (18) to write

\[
\bar{x}^\top P \bar{x} \geq l(\bar{x}) + \max_{P_v \in A} E[\bar{x}^\top A_K(v_0)\top PA_K(v_0)\bar{x}]
\]

\[
\geq \max_{P_v \in A} E[l(\bar{x}) + l(A_K(v_0)\bar{x}) + \bar{x}^\top A_K(v_0)\top A_K(v_1)\top PA_K(v_1)A_K(v_0)\bar{x}]
\]

which can be recursively applied to show that

\[
\bar{x}^\top P \bar{x} \geq \max_{P_v \in A} E \left[ \sum_{k=0}^{\infty} x_k^\top (Q + K^\top RK)x_k \right]. \tag{34}
\]

We can then take the expectation with respect to $\bar{x}$ of both sides. Noting that $E[\bar{x}^\top P \bar{x}] = Tr(P E[\bar{x}^\top \bar{x}]) = Tr(P I)$ explains why the trace of $W^{-1}$ is maximized in (19), proving (ii). Finally consider the constraint of (18), which is a DR Lyapunov decrease condition. As such we can use Theorem 4, proving (iii).