Data-driven distributionally robust LQR with multiplicative noise

Peter Coppens PETER.COPPENS@KULEUVEN.BE
Mathijs Schuurmans MATHIJS.SCHUURMANS@KULEUVEN.BE
Panagiotis Patrinos PANOS.PATRINOS@KULEUVEN.BE
STADIUS, Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium.

Editors: A. Bayen, A. Jadbabaie, G. J. Pappas, P. Parrilo, B. Recht, C. Tomlin, M. Zeilinger

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 the 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 (Dupačová, 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. Such a DR approach addresses most of the difficulties associated with learning automatically, since the ambiguity set directly models the uncertainty in the sample-based estimates against which the controllers will be robust. Similar techniques were recently studied in Schuurmans et al. (2019)

© 2020 P. Coppens, M. Schuurmans & P. Patrinos.

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.

The work of Gravell et al. (2019) also deals with learning control of linear systems with multiplicative noise, albeit from a different perspective. They employ a policy gradient algorithm, which requires the initial guess for the control gain to be stabilizing. By contrast, obtaining such a controller is the main goal of our approach.

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 Ω the sample space, $\mathcal{F}$ its associated σ-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 $\mathbb{P}_{y_0,\ldots,y_N}(A_0 \times \cdots \times A_N) := \mathbb{P}[y_0 \in A_0 \wedge \cdots \wedge y_N \in A_N] = \prod_{i=0}^N \mathbb{P}_y(A_i)$. 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. Let $\mathcal{M}$ denote the set of probability measures defined on $(\mathbb{R}^{n_w}, \mathcal{B})$, with $\mathcal{B}$ the Borel σ-algebra of $\mathbb{R}^{n_w}$.

2. Problem statement

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

$$x_{k+1} = A(w_k)x_k + B(w_k)u_k \tag{1}$$
$$\text{with} \qquad A(w) := A_0 + \sum_{i=1}^{n_w} w^{(i)} A_i, \quad B(w) := B_0 + \sum_{i=1}^{n_w} w^{(i)} B_i,$$

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 $\mathbb{P}_w$. Let $w^{(i)}$ denote the i-th element of vector w. We introduce the following shorthands: $\mathbf{A} := \left[\begin{smallmatrix} A_1^\top & \ldots & A_{n_w}^\top \end{smallmatrix}\right]^\top$, $\mathbf{B} := \left[\begin{smallmatrix} B_1^\top & \ldots & B_{n_w}^\top \end{smallmatrix}\right]^\top$, $\mathbf{A}_0 := \left[\begin{smallmatrix} A_0^\top & \mathbf{A}^\top \end{smallmatrix}\right]^\top$, $\mathbf{B}_0 := \left[\begin{smallmatrix} B_0^\top & \mathbf{B}^\top \end{smallmatrix}\right]^\top$. We also define $\Sigma_0 := \left[\begin{smallmatrix} 1 & \mu^\top \\ \mu & \Sigma + \mu\mu^\top \end{smallmatrix}\right]$, 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:

$$\begin{array}{ll} \underset{u_0,u_1,\ldots}{\textbf{minimize}} & \mathbb{E}\left[\sum_{k=0}^{\infty} x_k^\top Q x_k + u_k^\top R u_k\right] \\ \textbf{subj. to} & x_{k+1} = A(w_k)x_k + B(w_k)u_k, \quad k \in \mathbb{N} \\ & x_0 = \bar{x}, \end{array} \tag{2}$$

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 $\mathbb{E}[x_k^\top x_k] \to 0$ 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 $\mathbb{E}[x_k^\top x_k] \leq c\gamma^k \|x_0\|$ 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 are then equivalent: (i) it is* m.s.s.*, (ii) it is* e.m.s.s.*, (iii) $\exists P \succ 0$:*

$$P - \mathbf{A}_0^\top (\Sigma_0 \otimes P)\mathbf{A}_0 \succ 0. \tag{3}$$

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_∞ the solution of the following Riccati equation:*

$$P_\infty = Q + F(P_\infty) - H(P_\infty)^\top (R + G(P_\infty))^{-1} H(P), \tag{4}$$

where $F(P) := \mathbf{A}_0^\top (\Sigma_0 \otimes P)\mathbf{A}_0$, $G(P) := \mathbf{B}_0^\top (\Sigma_0 \otimes P)\mathbf{B}_0$, $H(P) := \mathbf{B}_0^\top (\Sigma_0 \otimes P)\mathbf{A}_0$.

II *The controller K_∞ stabilizes (1) in the mean square sense.*

III *The solution of the Riccati equation is found by solving the following SDP:*

$$\begin{array}{ll} \underset{P}{\textbf{minimize}} & -\mathbf{Tr}\, P \\ \textbf{subj. to} & \begin{bmatrix} Q - P + F(P) & H(P)^\top \\ H(P) & R + G(P) \end{bmatrix} \succeq 0, \\ & P \succeq 0. \end{array} \tag{5}$$

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 for the parametric form of the ambiguity set (7) used in the data-driven LQR problem, which we state in the next section.

2.2. Data-driven stochastic LQR problem

Consider now the case where $\mathbb{P}_w$ 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, we add a hat to imply that a random variable depends on these samples. 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. For instance, the *empirical approach*, where (2) is solved using $\hat{\mu} := \frac{1}{M}\sum_{i=0}^{M-1}\hat{w}_i$ and $\hat{\Sigma} := \frac{1}{M}\sum_{i=0}^{M-1}(\hat{w}_i - \hat{\mu})(\hat{w}_i - \hat{\mu})^\top$, does not guarantee stability. We will illustrate this with an example, which motivates the methodology presented in this paper.

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, but 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 $qx_k^2 + ru_k^2 = x_k^2 + 10^4 u_k^2$. 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:

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

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:

$$\hat{K}^* = -\frac{0.75\frac{\hat{p}^*}{r}}{1 + \frac{\hat{p}^*}{r}} \approx -\frac{\hat{\sigma}^2 - 0.4375}{0.75}. \tag{6}$$

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

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

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

$$\mathbb{P}\left[\hat{\sigma}^2 < 0.4697\right] = \mathbb{P}\left[\sum_{i=1}^{M-1}\hat{\xi}_i^2 < \frac{0.4697M}{\sigma^2}\right],$$

with $\hat{\xi}_i = \hat{w}_i/\sigma \sim \mathcal{N}(0,1)$ *and which corresponds to the cumulative distribution function of a* χ^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 an arbitrarily-chosen confidence level $\beta \in [0,1]$ and a corresponding ambiguity set $\hat{\mathcal{A}} : \Omega \rightrightarrows \mathcal{M}$, which represents the uncertainty of estimators $\hat{\mu}$ and $\hat{\Sigma}$. The size of $\mathcal{A}$ is then determined such that $\mathbb{P}(\mathbb{P}_w \in \hat{\mathcal{A}}) \geq 1-\beta$. In particular, we parametrize $\mathcal{A}$ as first suggested by Delage and Ye (2010):

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

The values of $r_\mu(\beta)$ and $r_\Sigma(\beta)$ such that $\mathbb{P}(\mathbb{P}_w \in \hat{\mathcal{A}}) \geq 1-\beta$ are derived in Section 3. In Section 4, the following DR counterpart of (2) is solved

$$\begin{array}{ll} \underset{u_0,u_1,\ldots}{\textbf{minimize}} & \underset{\mathbb{P}_v \in \hat{\mathcal{A}}}{\max}\, \mathbb{E}\left[\sum_{k=1}^{\infty} x_k^\top Q x_k + u_k^\top R u_k \right] \\ \textbf{subj. to} & x_{k+1} = A(v_k)x_k + B(v_k)u_k, \quad k \in \mathbb{N} \\ & x_0 = \bar{x}, \end{array} \tag{8}$$

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 the matrices* $\{\hat{A}_i\}_{i=0}^{n_w}$*, the set* $\{w_k\}_{k\in\mathbb{N}}$ *consisting of* i.i.d. *copies of a square integrable random vector* w *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* $\mathbb{P}(\mathbb{P}_w \in \hat{\mathcal{A}}) \geq 1-\beta$*, with* $\mathbb{P}_w$ *the true distribution of* w*. Then if there exists a* $P \succ 0$ *such that:*

$$P - \underset{\mathbb{P}_v \in \hat{\mathcal{A}}}{\max}\, \mathbb{E}\left[A(v)^\top P A(v)\right] \succ 0, \tag{9}$$

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 $r_\Sigma(\beta)$ and $r_\mu(\beta)$ 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 σ^2 if $\mathbb{E}[y] = 0$ and its moment generating function satisfies*

$$\mathbb{E}[\mathbf{exp}(\lambda y)] \leq \mathbf{exp}\left(\frac{\sigma^2\lambda^2}{2}\right) \quad \forall \lambda \in \mathbb{R}. \tag{10}$$

We denote this by $y \sim \mathbf{subG}(\sigma^2)$. We say that a random vector $\xi \in \mathbb{R}^{n_w}$ is sub-Gaussian, or $\xi \sim \mathbf{subG}_{n_w}(\sigma^2)$, if $z^\top \xi \sim \mathbf{subG}(\sigma^2), \forall z \in \mathbb{R}^{n_w}$ with $\|z\|_2 = 1$.

Assumption 1 *We assume that (i) w is square integrable; (ii) $\{w_k\}_{k\in\mathbb{N}}$ are* i.i.d. *copies of w; (iii) $\Sigma \succ 0$; and (iv) $\Sigma^{-1/2}(w_k - \mu) \sim \mathbf{subG}_{n_w}(\sigma^2)$ for some $\sigma \geq 1$.*

Note that in the specific case of Gaussian disturbances, Assumption 1(iv) holds with $\sigma^2 = 1$, 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, σ^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 $\xi \sim \mathbf{subG}_{n_w}(\sigma^2)$ be a random vector, with $\mathbb{E}[\xi] = 0$, $\mathbb{E}[\xi\xi^\top] = I_{n_w}$. Let $\{\hat{\xi}_i\}_{i=0}^{M-1}$ be M independent copies of ξ 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

$$t_\Sigma(\beta) := \frac{\sigma^2}{1-2\epsilon}\left(\sqrt{\frac{32q(\beta,\epsilon,n_w)}{M}} + \frac{2q(\beta,\epsilon,n_w)}{M}\right), \tag{11}$$

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

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*

$$\mathbb{P}[\|\hat{\zeta}\|_2^2 \leq t_\mu(\beta)] \geq 1 - \beta.$$

where

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

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

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 $\xi = \Sigma^{-1/2}(w-\mu) \sim \mathbf{subG}_{n_w}(\sigma^2)$. 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 ϵ, $p(\beta, n_w)$, $q(\beta, \epsilon, n_w)$, $t_\Sigma(\beta/2)$ and $t_\mu(\beta/2)$ be as defined in Theorems 6 and 7. Provided that*

$$M > \left(\frac{\sigma^2\sqrt{32q(\beta/2,\epsilon,n_w)} + \sqrt{32\sigma^4 q(\beta/2,\epsilon,n_w) + 8\sigma^2(1-2\epsilon)q(\beta/2,\epsilon,n_w) + 4\sigma^2(1-2\epsilon)^2 p(\beta/2,n_w)}}{2(1-2\epsilon)} \right)^2, \tag{12}$$

then with probability at least $1-\beta$,

$$(\hat{\mu}-\mu)^\top \hat{\Sigma}^{-1}(\hat{\mu}-\mu) \le r_\mu(\beta), \quad \Sigma \preceq r_\Sigma(\beta)\hat{\Sigma},$$

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

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

$$\begin{aligned}
\mathbb{E}[\xi] &= \Sigma^{-1/2}(\mathbb{E}[w] - \mu) = 0 \\
\mathbb{E}[\xi\xi^\top] &= \Sigma^{-1/2}\mathbb{E}\left[(w-\mu)(w-\mu)^\top\right]\Sigma^{-1/2} = \Sigma^{-1/2}\Sigma\Sigma^{-1/2} = I_{n_w}.
\end{aligned}$$

Let $\hat{\zeta} = \frac{1}{M}\sum_{i=0}^{M-1}\hat{\xi}_i$ and $\hat{I} = \frac{1}{M}\sum_{i=0}^{M-1}\hat{\xi}_i\hat{\xi}_i^\top$. By Theorems 6 and 7, we then have with probability $1-\beta$ that,

$$\|\hat{I} - I_{n_w}\|_2 \le t_\Sigma(\beta/2) \tag{13a}$$

$$\text{and } \|\hat{\zeta}\|_2^2 \le t_\mu(\beta/2). \tag{13b}$$

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

$$\tilde{\Sigma} := \frac{1}{M}\sum_{i=0}^{M-1}(\hat{w}_i - \mu)(\hat{w}_i - \mu)^\top = \frac{1}{M}\sum_{i=0}^{M-1}(\Sigma^{1/2}\xi_i)(\Sigma^{1/2}\xi_i)^\top = \Sigma^{1/2}\hat{I}\Sigma^{1/2}. \tag{14}$$

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

$$\begin{aligned}
\|\hat{I} - I_{n_w}\|_2 \le t_\Sigma(\beta/2) &\Leftrightarrow (1 - t_\Sigma(\beta/2))I_{n_w} \preceq \hat{I} \preceq (1 + t_\Sigma(\beta/2))I_{n_w} \\
&\Rightarrow (1 - t_\Sigma(\beta/2))\Sigma \preceq \tilde{\Sigma} \preceq (1 + t_\Sigma(\beta/2))\Sigma,
\end{aligned} \tag{15}$$

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 $x \in \mathbb{R}^{n_w}$,

$$\begin{aligned}
x^\top(\hat{\mu}-\mu)(\hat{\mu}-\mu)^\top x = (x^\top(\hat{\mu}-\mu))^2 &= (x^\top \overbrace{\Sigma^{1/2}\Sigma^{-1/2}}^{I_{n_w}}(\hat{\mu}-\mu))^2 \\
&\le \|\Sigma^{1/2}x\|_2^2 \|\Sigma^{-1/2}(\hat{\mu}-\mu)\|_2^2 \\
&\le t_\mu(\beta/2)x^\top\Sigma x.
\end{aligned} \tag{16}$$

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

$$\begin{aligned}
\tilde{\Sigma} &= \tfrac{1}{M}\textstyle\sum_{i=0}^{M-1}(w_i-\mu)(w_i-\mu)^\top \\
&= \tfrac{1}{M}\textstyle\sum_{i=0}^{M-1}(w_i-\hat{\mu}+\hat{\mu}-\mu)(w_i-\hat{\mu}+\hat{\mu}-\mu)^\top \\
&= \tfrac{1}{M}\textstyle\sum_{i=0}^{M-1}(w_i-\hat{\mu})(w_i-\hat{\mu})^\top + (w_i-\hat{\mu})(\hat{\mu}-\mu)^\top + (\hat{\mu}-\mu)(w_i-\hat{\mu})^\top + (\hat{\mu}-\mu)(\hat{\mu}-\mu)^\top \\
&= \hat{\Sigma} + (\hat{\mu}-\mu)(\hat{\mu}-\mu)^\top \preceq \hat{\Sigma} + t_\mu(\beta/2)\Sigma,
\end{aligned}$$

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

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

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

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 set given in (7) in two stages. Firstly, we extend the result of Proposition 3 to the case where μ is known and Σ is estimated, *i.e.,* $r_\mu(\beta) = 0$ and $r_\Sigma(\beta) > 0$. Secondly, we present the result 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*

$$\hat{\mathcal{A}}_\Sigma := \left\{ \mathbb{P}_v \in \mathcal{M} \mid \mathbb{E}\left[(v-\mu)(v-\mu)^\top\right] \preceq r_\Sigma(\beta)\hat{\Sigma},\, \mathbb{E}\left[v\right] = \mu \right\}, \tag{17}$$

where $\mathbb{P}(\mathbb{P}_w \in \hat{\mathcal{A}}_\Sigma) \geq 1-\beta$. Then applying Proposition 3 with $\Sigma = r_\Sigma(\beta)\hat{\Sigma}$ results in the optimal linear controller for (8), assuming that (1) is DR mean square stabilizable, i.e., *there exists a K such that 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 $\hat{\mathcal{A}}$ 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 $P \succ 0$ and to solve the optimization problem

$$\begin{aligned} \underset{V(x)}{\textbf{minimize}} \quad & \mathbb{E}\left[V(\tilde{x})\right] \\ \textbf{subj. to} \quad & V(x) \geq \min_u \left\{ x^\top Q x + u^\top R u + \max_{\mathbb{P}_v \in \hat{\mathcal{A}}} \mathbb{E}\left[V(A(v)x + B(v)u)\right] \right\}, \quad \forall x, \end{aligned} \tag{18}$$

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 $\bar{x}$ 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 $\hat{\mathcal{A}}$ 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 $\mathbb{E}[\tilde{x}] = 0$ and $\mathbb{E}[\tilde{x}\tilde{x}^\top] = I_{n_x}$, by solving the following SDP.*

$$\begin{aligned}
\underset{W,V,S,L}{\textbf{maximize}} \quad & \mathbf{Tr}\, W \\
\textbf{subj. to} \quad & \begin{bmatrix} S & r_\mu(\beta)H_1^\top & r_\mu(\beta)H_2^\top & \cdots & r_\mu(\beta)H_{n_w}^\top \\ r_\mu(\beta)H_1 & L & & & \\ r_\mu(\beta)H_2 & & L & & \\ \vdots & & & \ddots & \\ r_\mu(\beta)H_{n_w} & & & & L \end{bmatrix} \succeq 0, && \text{(19a)} \\
& \begin{bmatrix} W-\sqrt{2}S & (\mathbf{A}W+\mathbf{B}V)^\top & (\hat{A}W+\hat{B}V)^\top & W^\top Q^{\frac{1}{2}} & V^\top R^{\frac{1}{2}} \\ \mathbf{A}W+\mathbf{B}V & \hat{\Sigma}_{\mathrm{dr}}^{-1}\otimes W & & & \\ \hat{A}W+\hat{B}V & & W-\sqrt{2}L & & \\ Q^{\frac{1}{2}}W & & & I_{n_x} & \\ R^{\frac{1}{2}}V & & & & I_{n_u} \end{bmatrix} \succeq 0, && \text{(19b)}
\end{aligned}$$

where $H_i = \sum_{j=1}^{n_w} [\hat{\Sigma}^{1/2}]_{ji}(A_j W + B_j V)$, $\hat{A} = A(\hat{\mu})$, $\hat{B} = B(\hat{\mu})$, $\hat{\Sigma}_{\mathrm{dr}} = r_\Sigma(\beta)\hat{\Sigma}$. Let $\hat{W}$ and $\hat{V}$ denote the minimizers of (19). The corresponding linear controller $u = \hat{K}x$ with $\hat{K} = \hat{V}\hat{W}^{-1}$ then achieves an upper bound of the cost (18), given by $\mathbb{E}[\tilde{x}^\top \hat{P}\tilde{x}]$, 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 $r_\mu(\beta)$ 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):*

$$\begin{aligned}
\underset{W,V,S,L,\gamma}{\textbf{minimize}} \quad & \gamma \\
\textbf{subj. to} \quad & \bar{x}^\top W^{-1}\bar{x} \leq \gamma && \text{(20a)} \\
& \text{(19a)}, \text{(19b)},
\end{aligned}$$

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 $r_\mu(\beta)$ and $r_\Sigma(\beta)$ go to zero (Balakrishnan and Vandenberghe, 2003).

Remark 12 *Invertibility of $r_\Sigma(\beta)\hat{\Sigma}$ can be guaranteed by using $r_\Sigma(\beta)\hat{\Sigma} + \lambda I$ instead of $r_\Sigma(\beta)\hat{\Sigma}$ for some small λ, which increasing the size of the ambiguity set, introducing additional conservatism.*

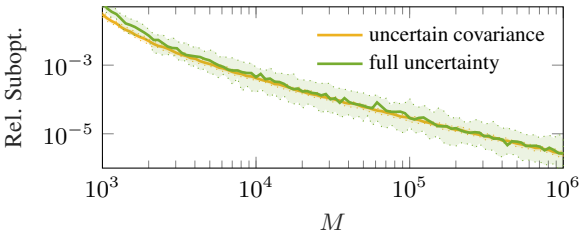


Figure 1: Relative suboptimality versus sample size for controllers. The colored area depicts the 0.3-confidence interval around the cost. The full line depicts the mean.

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 Σ and μ instead of their data-driven estimates. Consider the double integrator model with matrices:

$$A_0 = \begin{bmatrix} 1 & T_s \\ 0 & 1 - 0.4T_s \end{bmatrix}, B_0 = \begin{bmatrix} 0 \\ T_s \end{bmatrix}, A_1 = \begin{bmatrix} 0 & 0 \\ 0 & -T_s \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ T_s \end{bmatrix},$$

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 $\Sigma = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and mean $\mu = \left[\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right]$.

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

The simulation setup is as follows. We compare the nominal controller, the uncertain covariance controller (Proposition 9) and the full uncertainty controller (Theorem 10). We evaluate the expected closed-loop cost for $\bar{x} = \left[\begin{smallmatrix} 2 & 2 \end{smallmatrix}\right]^\top$ by solving the Lyapunov equation. We start with $M = 1000$ to satisfy (12). For each value of M we produce 30 realisations of both DR controllers. Figure 1 depicts confidence intervals for relative difference between closed-loop cost of the DR controllers and the nominal controller (*i.e.,* the relative suboptimality). The figure shows that both converge with a rate $\mathcal{O}(1/M)$ to the nominal cost even though Theorem 10 only solves (8) approximately.

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 operate in the setting where 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 an SDP. As a result, stability of the closed-loop system can be guaranteed with high probability.

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 Fonds Wetenschappelijk Onderzoek PhD grant 11E5520N and research projects G0A0920N, 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 Dupačová. 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.

Benjamin Gravell, Peyman Mohajerin Esfahani, and Tyler Summers. Learning robust control for LQR systems with multiplicative noise via policy gradient. *arXiv preprint arXiv:1905.13547*, 2019.

Daniel Hsu, Sham M. 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. Kakade, and Tong Zhang. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 17(52):1–6, 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, Alalin 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.*:

$$P - \mathbb{E}\left[A(w)^\top P A(w)\right] \succ 0.$$

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

$$\mathbb{E}\left[v^\top P v\right] = \mathbf{Tr}\left(P\mathbb{E}\left[\tilde{v}\tilde{v}^\top\right]\right) + \mathbb{E}\left[v\right]^\top P\mathbb{E}\left[v\right], \tag{21}$$

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

$$\begin{aligned}\mathbb{E}\left[v\right] &= A(\mu)x,\\ \mathbb{E}\left[\tilde{v}\tilde{v}^\top\right] &= \mathbb{E}\left[(\tilde{w}^\top \otimes I_{n_x})\mathbf{A}xx^\top\mathbf{A}^\top(\tilde{w}\otimes I_{n_x})\right] = \textstyle\sum_{i,j=1}^{n_w}\left(\Sigma_{ij}A_i x x^\top A_j^\top\right),\end{aligned}$$

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

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_w}\left[x^\top A(w)^\top P A(w) x\right] &= \mathbf{Tr}\left(P\textstyle\sum_{i,j=1}^{n_w}\left(\Sigma_{ij}A_i x x^\top A_j^\top\right)\right) + x^\top A(\mu)^\top P A(\mu)x\\ &= \mathbf{Tr}\left(\textstyle\sum_{i,j=1}^{n_w}\left(\Sigma_{ij}x^\top A_j^\top P A_i x\right)\right) + x^\top A(\mu)^\top P A(\mu)x\\ &= x^\top\mathbf{A}^\top(\Sigma\otimes P)\mathbf{A}x + x^\top A(\mu)^\top P A(\mu)x\\ &= x^\top\mathbf{A}_0^\top(\Sigma_0\otimes P)\mathbf{A}_0 x.\end{aligned}$$

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):

$$(TV_k)(x) = \min_u \left\{x^\top Q x + u^\top R u + \mathbb{E}\left[V_k(A(w)x + B(w)u)\right]\right\}, \tag{22}$$

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

$$\begin{aligned} x^\top P_{k+1} x &= \min_u \left\{x^\top Q x + u^\top R u + (\mathbf{A}_0 x + \mathbf{B}_0 u)^\top (\Sigma_0 \otimes P_k)(\mathbf{A}_0 x + \mathbf{B}_0 u)\right\} \\ &= x^\top (Q + K_k^\top R K_k) x + x^\top (\mathbf{A}_0 + \mathbf{B}_0 K_k)^\top (\Sigma_0 \otimes P_k)(\mathbf{A}_0 + \mathbf{B}_0 K_k) x \qquad (23a) \\ \Leftrightarrow P_{k+1} &= Q + F(P_k) - H(P_k)^\top (R + G(P_k))^{-1} H(P_k), \qquad (23b) \end{aligned}$$

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 $u = K_k x$ 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_∞; (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_∞ 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

$$\mathbb{E}\left[x_k^\top (Q + K^\top R K) x_k\right] = \mathbf{Tr}\left((Q + K^\top R K)\mathbb{E}\left[x_k x_k^\top\right]\right) \tag{24}$$

and by definition of *e.m.s.s.*, $\mathbb{E}\left[x_k x_k^\top\right] \leq c\gamma^k \|x_0\|$, $\forall k \in \mathbb{N}$, 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 $P_\infty \succ 0$, proving (i).

Next we prove that K_∞ 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)):

$$\begin{aligned} x_k^\top P_\infty x_k &= x_k^\top (Q + K_\infty^\top R K_\infty) x_k + \mathbb{E}\left[x_{k+1}^\top P_\infty x_{k+1} \mid x_k\right] \\ x_{k+1}^\top P_\infty x_{k+1} &= x_{k+1}^\top (Q + K_\infty^\top R K_\infty) x_{k+1} + \mathbb{E}\left[x_{k+2}^\top P_\infty x_{k+2} \mid x_{k+1}\right], \end{aligned} \tag{25}$$

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 $\mathbb{E}\left[\mathbb{E}\left[x_{k+1}^\top P_\infty x_{k+1} \mid x_k\right]\right] = \mathbb{E}\left[x_{k+1}^\top P_\infty x_{k+1}\right]$ and summing the equalities for $k = 0, \ldots, N-1$ allows us to write:

$$\begin{aligned} x_0^\top P_N x_0 &\leq \mathbb{E}\textstyle\sum_{k=0}^{N-1}\left[x_k^\top (Q + K_\infty^\top R K_\infty) x_k\right] = x_0^\top P_\infty x_0 - \mathbb{E}\left[x_N^\top P_\infty x_N\right] \\ &\leq x_0^\top P_\infty x_0, \quad \forall N \in \mathbb{N}, \forall x_0 \in \mathbb{R}^{n_x}, \end{aligned} \tag{26}$$

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 $P_\infty \succ 0$. 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_∞ and K_∞ describe the optimal cost and the optimal controller respectively, proving (ii).

Notice that (23a) implies the Lyapunov condition (3) since $R \succ 0$ and $Q \succ 0$. This holds for any solution of the Riccati equation (4), proving **II**.

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

$$x_k^\top \tilde{P}_\infty x_k \leq x_k^\top (Q + K_k^\top R K_k) x_k + \mathbb{E}\left[x_{k+1}^\top \tilde{P}_\infty x_{k+1} \mid x_k \right], \tag{27}$$

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

$$x_0^\top \tilde{P}_\infty x_0 \leq x_0^\top P_N x_0 + \mathbb{E}\left[\tilde{x}_N \tilde{P}_\infty x_N \right] \tag{28}$$

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

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

Since $\mathbb{P}(\mathbb{P}_w \in \hat{\mathcal{A}}) \geq 1 - \beta$ 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:

$$(TV_{k+1})(x) = \min_u \left\{ x^\top Q x + u^\top R u + \max_{\mathbb{P}_v \in \hat{\mathcal{A}}_\Sigma} \mathbb{E}\left[V_k(A(v)x + B(v)u) \right] \right\}. \tag{30}$$

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 Σ from the maximum and grouping the remaining ones together, only the following needs to be evaluated:

$$\max_{\Sigma \preceq r_\Sigma(\beta)\hat{\Sigma}} (\mathbf{A}x + \mathbf{B}u)^\top (\Sigma \otimes P)(\mathbf{A}x + \mathbf{B}u).$$

This is equivalent to:

$$\max_{\Sigma \preceq r_\Sigma(\beta)\hat{\Sigma}} \mathbf{Tr}\left(\begin{bmatrix} z_1^\top P z_1 & \dots & z_1^\top P z_{n_w} \\ \vdots & \ddots & \vdots \\ z_{n_w}^\top P z_1 & \dots & z_{n_w}^\top P z_{n_w} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \dots & \Sigma_{1n_w} \\ \vdots & \ddots & \vdots \\ \Sigma_{n_w 1} & \dots & \Sigma_{n_w n_w} \end{bmatrix}\right) = \max_{\Sigma \preceq r_\Sigma(\beta)\hat{\Sigma}} \mathbf{Tr}\left(X\Sigma\right),$$

where $z_i = A_i x + B_i u$. Since $X = Z^\top P Z \succ 0$, with $Z = \begin{bmatrix} z_1 & \dots & z_{n_w} \end{bmatrix}$. The maximum corresponds to a support function for which one can easily verify using the optimality conditions that:

$$\max_{\Sigma \preceq r_\Sigma(\beta)\hat{\Sigma}} \mathbf{Tr}\left(X\Sigma\right) = \mathbf{Tr}\left(r_\Sigma(\beta) X \hat{\Sigma}\right) = r_\Sigma(\beta)(\mathbf{A}x + \mathbf{B}u)^\top (\hat{\Sigma} \otimes P)(\mathbf{A}x + \mathbf{B}u)^\top.$$

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 $r_\Sigma(\beta)\hat{\Sigma}$ instead of Σ. More specifically, we know that the cost is bounded above since there exists a stabilizing K for $r_\Sigma(\beta)\hat{\Sigma}$ by assumption. This means that value iteration will converge to some P_∞. We can once again telescope the Riccati equation and prove that this P_∞ and K_∞ 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 Σ with probability at least $1-\beta$. For this consider the Riccati equation (4) which is equivalent to:

$$P - \max_{\Sigma \preceq r_\Sigma(\beta)\hat{\Sigma}} (\mathbf{A}_0 + \mathbf{B}_0 \hat{K}_\infty)^\top (\Sigma_0 \otimes P)(\mathbf{A}_0 + \mathbf{B}_0 \hat{K}_\infty) = Q + K_\infty^\top R K_\infty \succ 0,$$

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

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:

$$\begin{aligned} P - Q - K^\top R K - (\mathbf{A} + \mathbf{B}K)^\top (\hat{\Sigma}_{\text{dr}} \otimes P)(\mathbf{A} + \mathbf{B}K) \\ - (A(\mu) + B(\mu)K)^\top P (A(\mu + B(\mu)K) \succeq 0, \quad \forall \mu \in \mathcal{D}, \end{aligned}$$

where we used $u = Kx$ and define $\mathcal{D} := \left\{(\mu - \hat{\mu})^\top \hat{\Sigma}^{-1} (\mu - \hat{\mu}) \leq r_\mu(\beta)^2\right\}$. Pre- and post-multiplying by $W = P^{-1}$ and setting $V = KW$ results in:

$$\begin{aligned} W - W^\top Q W - V^\top R V - (\mathbf{A}W + \mathbf{B}V)^\top (\hat{\Sigma}_{\text{dr}} \otimes W)(\mathbf{A}W + \mathbf{B}V) \\ - (A(\mu)W + B(\mu)V)^\top W^{-1} (A(\mu)W + B(\mu)V), \quad \forall \mu \in \mathcal{D}. \end{aligned}$$

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

$$\begin{bmatrix} W & \star & \star & \star & \star \\ \mathbf{A}W + \mathbf{B}V & \hat{\Sigma}_{\text{dr}}^{-1} \otimes W & & & \\ (\hat{A}W + \hat{B}V) + \sum_{i=1}^{n_w} \mu_i F_i & & W & & \\ Q^{\frac{1}{2}} W & & & I_{n_x} & \\ R^{\frac{1}{2}} V & & & & I_{n_u} \end{bmatrix} \succeq 0, \quad \forall \tilde{\mu} \in \tilde{\mathcal{D}}, \tag{31}$$

where $\tilde{\mathcal{D}} := \left\{\tilde{\mu}^\top \hat{\Sigma}^{-1}\tilde{\mu} \leq r_\mu(\beta)^2\right\}$ and $F_i = A_i W + B_i V$. Defining $\zeta = \hat{\Sigma}^{-1/2}\tilde{\mu}$ 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

$$U_0 + \textstyle\sum_{i=1}^{n_w} \delta_i U_i \succeq 0, \quad \forall \delta \in \{\delta \mid \|\delta\|_2 \leq \rho\}\,.$$

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:

$$\begin{aligned}
\xi^\top(U_0 + \textstyle\sum_{i=1}^{n_w} \delta_i U_i)\xi &= \xi^\top U_0 \xi + 2\textstyle\sum_{i=1}^{n_w} \zeta_i \xi_3^\top H_i \xi_1 \\
&\qquad \left(\xi{=}[\xi_1^\top\ \xi_2^\top\ \xi_3^\top\ \xi_4^\top\ \xi_5^\top\,]^\top\right) \\
&= \xi^\top U_0 \xi + 2\xi_L^\top \left[\textstyle\sum_{i=1}^{n_w} \zeta_i L^{-1/2} H_i S^{-1/2} \xi_S\right] \\
&\qquad (\xi_S{=}S^{1/2}\xi_1, \xi_L{=}L^{1/2}\xi_3) \\
&\geq \xi^\top U_0 \xi - 2\|\xi_L\|_2 \left[\textstyle\sum_{i=1}^{n_w} |\zeta_i| \|L^{-1/2} H_i S^{-1/2} \xi_S\|_2\right] \\
&\geq \xi^\top U_0 \xi - 2\|\xi_L\|_2 \sqrt{\textstyle\sum_{i=1}^{n_w} r_\mu(\beta)^2 \|L^{-1/2} H_i S^{-1/2} \xi_S\|_2^2} \qquad (32a) \\
&= \xi^\top U_0 \xi - 2\|\xi_L\|_2 \sqrt{r_\mu(\beta)^2 \xi_S^\top \left[\textstyle\sum_{i=1}^{n_w} S^{-1/2} H_i^\top L^{-1} H_i S^{-1/2}\right] \xi_S} \\
&\geq \xi^\top U_0 \xi - 2\|\xi_L\|_2 \|\xi_S\|_2, \qquad (32b)
\end{aligned}$$

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

$$\textstyle\sum_{i=1}^{n_w} S^{-1/2}(r_\mu(\beta) H_i^\top) L^{-1} (H_i r_\mu(\beta)) S^{-1/2} \leq I_{n_w},$$

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:

$$\xi^\top U_0 \xi \geq 2\sqrt{(\xi_L^\top \xi_L)(\xi_S^\top \xi_S)}, \qquad (33)$$

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

$$\begin{aligned}
\sqrt{2}(\xi_L^\top \xi_L + \xi_S^\top \xi_S) &\geq 2\sqrt{(\xi_L^\top \xi_L)(\xi_S^\top \xi_S)} \\
2\xi_L^\top \xi_L + 4(\xi_L^\top \xi_L)(\xi_S^\top \xi_S) + 2\xi_S^\top \xi_S &\geq 4(\xi_L^\top \xi_L)(\xi_S^\top \xi_S).
\end{aligned}$$

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

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

$$\begin{aligned}
\tilde{x}^\top P \bar{x} &\geq \ell(\tilde{x}) + \max_{\mathbb{P}_v \in \hat{\mathcal{A}}} \mathbb{E}\left[\tilde{x}^\top A_K(v_0)^\top P A_K(v_0)\tilde{x}\right], \\
&\geq \max_{\mathbb{P}_v \in \hat{\mathcal{A}}} \mathbb{E}\left[\ell(\tilde{x}) + \ell(A_K(v_0)\tilde{x}) + \tilde{x}^\top A_K(v_0)^\top A_K(v_1)^\top P A_K(v_1) A_K(v_0)\tilde{x}\right]
\end{aligned}$$

which can be recursively applied to show that

$$\tilde{x}^\top P \tilde{x} \geq \max_{\mathbb{P}_v \in \hat{\mathcal{A}}} \mathbb{E}\left[\textstyle\sum_{k=0}^{\infty} x_k^\top (Q + K^\top R K) x_k\right]. \qquad (34)$$

We can then take the expectation with respect to $\tilde{x}$ of both sides. Noting that $\mathbb{E}[\tilde{x}^\top P \tilde{x}] = \mathbf{Tr}(P\mathbb{E}[\tilde{x}\tilde{x}^\top]) = \mathbf{Tr}(P I_{n_x})$ 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). ■