Proceedings of Machine Learning Research vol 168:1–11, 2022 4th Annual Conference on Learning for Dynamics and Control

Sample Complexity of the Robust LQG Regulator with Coprime Factors Uncertainty

Yifei Zhang * YZHANG133@STEVENS.EDU
Stevens Institute of Technology, NJ 07030, USA

Sourav Kumar Ukil * SUKIL@STEVENS.EDU
Stevens Institute of Technology, NJ 07030, USA

Ephraim Neimand * ENEIMAND@STEVENS.EDU
Stevens Institute of Technology, NJ 07030, USA

Serban Sabău * SSABAU@STEVENS.EDU
Stevens Institute of Technology, NJ 07030, USA

Myron E. Hohil MYRON.E.HOHIL.CIV@MAIL.MIL
Armaments Center, Picatinny Arsenal, NJ 07806, USA

Editors: R. Firoozi, N. Mehr, E. Yel, R. Antonova, J. Bohg, M. Schwager, M. Kochenderfer

Abstract

This paper addresses the end-to-end sample complexity bound for learning the $\mathcal{H}_2$ optimal controller (the Linear Quadratic Gaussian (**LQG**) problem) with unknown dynamics, for potentially *unstable* Linear Time Invariant (**LTI**) systems. The robust LQG synthesis procedure is performed by considering bounded additive model uncertainty on the coprime factors of the plant. The closed-loop identification of the nominal model of the true plant is performed by constructing a Hankel-like matrix from a single time-series of noisy finite length input-output data, using the ordinary least squares algorithm from Sarkar and Rakhlin (2019). Next, an $\mathcal{H}_\infty$ bound on the estimated model error is provided and the robust controller is designed via convex optimization, much in the spirit of Boczar et al. (2018) and Zheng et al. (2020a), while allowing for bounded additive uncertainty on the coprime factors of the model. Our conclusions are consistent with previous results on learning the LQG and LQR controllers. Reference Zhang et al. (2021) is the extended version of this paper.

Keywords: Robust LQG Control, Coprime Factorization, LTI Systems, Sample Complexity.

1. Introduction

Considerable research efforts have been spent within the last few years towards approaching classical control problems with modern statistical and optimization tools from the Machine Learning framework, envisaging practical applications, see for example Dean et al. (2018), Mania et al. (2019), Dean et al. (2020), Zheng et al. (2020a). The starting point of the aforementioned research efforts has been the classical LQG control problem, which deals with partially observed linear and time-invariant dynamical systems driven by Gaussian noise and where the problem is finding the optimal output feedback law that minimizes the expected value of a quadratic cost.

In this paper an end-to-end sample-complexity bound on learning LQG controllers that stabilize the true system with high probability is established by incorporating recent advances in finite time

* The work of the authors is based upon research supported by the NSF under the CAREER Award Number # 1653756.

(non-asymptotic) system identification (Sarkar and Rakhlin (2019)). The contribution resides in the development of a tractable robust control synthesis procedure, that allows for bounded additive model uncertainty on the coprime factors of the model of the plant, thus circumventing both the need for state feedback and the restrictive assumption on the plant's open loop stability. The resulted sub-optimality gap is bounded as a function of the level of the model uncertainty. The end-to-end sample complexity bound for learning robust LQG controllers is $\mathcal{O}(\sqrt{logT/T})$, where T is the time horizon for learning. For open-loop stable systems, Zheng et al. (2020a) recently proved that the performance for LQG controllers deteriorates linearly with the model estimation error, starting from the original analysis of Dean et al. (2018) from the case of learning fully observed LQR controllers. The robust control synthesis proposed here achieves the same scaling for the sub-optimality gap as Dean et al. (2018), namely $\mathcal{O}(\gamma^2)$, where γ is the model uncertainty level.

The enclosed reference Zhang et al. (2021) provides the extended version of this paper, including the complete proofs.

1.1. The Linear Quadratic Gaussian Problem

Within the last few years, modern statistical and algorithmic methods led to new solutions for classical control problems, such as the Linear Quadratic Gaussian problem. For a discrete-time **LTI** (Linear ant Time Invariant) systems driven by Gaussian process and sensor noise:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \delta x_k, \\ y_k &= Cx_k + Du_k + \nu_k, \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state of the system, $u_k \in \mathbb{R}^m$ is the control input and $y_k \in \mathbb{R}^p$ is the measurement output with $\delta x_k \in \mathbb{R}^m$, $\nu_k \in \mathbb{R}^p$ are Gaussian noise with zero mean, covariance $\sigma_{\delta x}^2 I$ and $\sigma_\nu^2 I$ respectively, the classical LQG control problem is defined as:

$$\begin{aligned} \min_{u_0,u_1,\ldots} \quad & \lim_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\sum_{t=0}^{T}(y_t^T P_1 y_t + u_t^T P_2 u_t)\right] \\ \text{subject to} \quad & (1), \end{aligned} \tag{2}$$

where, P_1, P_2 is positive definite. Without loss of generality, it is assumed that $P_1 = I_p$, $P_2 = I_m$, $\sigma_{\delta x} = 1$, $\sigma_\nu = 1$.

In a nutshell, the problem can be stated as *learning* with high probability and in finite time the model of an unknown LTI system and subsequently designing its optimal LQG controller, while accounting for the inherent model uncertainty incurred at the *learning* stage.

1.2. Contributions

Recently, LQG control has been studied in a model based Reinforcement Learning framework (Zheng et al. (2020a)) and the sub-optimality performance degradation of the robust LQG controller was proved to scale as a function of the modeling error. However, the results in Zheng et al. (2020a) are valid only for open-loop stable systems, thus excluding many situations of practical interest. This paper shows how to remove the stability assumption on the unknown system, while at the same time streamlining the equivalent optimization problem, by reducing the size of the subsequent linear constraints. The proposed algorithm is consistent with previous results, while allowing for a much stronger description of the modeling error as bounded additive uncertainty on the coprime factors of the model of the plant (without restriction on the McMillan degree of the true plant or on its number of unstable poles). As expected, the presence of additive, norm-bounded factors on the

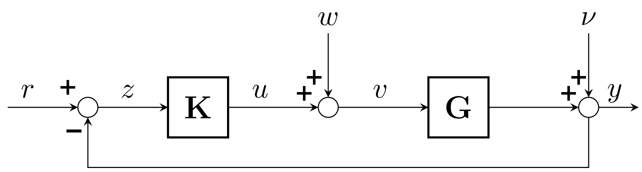

Figure 1: Standard unity feedback loop of the plant $\mathbf{G}$ with the controller $\mathbf{K}$

coprimes of the plant renders the cost functional non-convex, therefore the derivation of an upper-bound on the cost functional is needed. This is subsequently exploited to derive a quasi-convex approximation of the robust LQG problem. An inner approximation of the quasi convex problems via FIR truncation is employed. Previous results (Mania et al. (2019), Zheng et al. (2020a)) show that indeed the *certainty equivalent controller* may achieve superior sub-optimality scaling than our result, but only for the fully observed LQR settings (Mania et al. (2019)), in the setup of a stricter requirement on admissible uncertainty. Given the lack of prior gain margin for the optimal LQG controller, which is known to be notoriously fragile, even under small model uncertainty the stabilizability of the resulted controller may be lost, thus the availability of a more general framework for modelling of uncertainty is important.

Existing non-asymptotic identification methods (Sarkar et al. (2020)) have been adapted in order to yield a comparable end-to-end sample complexity. The identification of the unstable plant is performed in closed loop, directly on its the coprime factors via the dual Youla Parameterization (Anderson (1998)). The algorithm employed for system identification doesn't require the knowledge of the model's order (Oymak and Ozay (2019)), which is the common scenario in many applications. Pursuing the identification of the plant a $\mathcal{H}_2$ bound for Hankel matrix estimation with high probability is derived, followed by a $\mathcal{H}_\infty$ bound on the uncertainty on the coprime factors, which quantifies the modeling error. The robust controller design is recast as convex optimization for estimated nominal model within a *worst case scenario* on the uncertainty. For the output feedback of potentially unstable plants, the resulted sample complexity result is matched to the same level as that obtained in recent papers (Boczar et al. (2018), Dean et al. (2020), Zheng et al. (2020a)), where the robust so-called *SLP* or *IOP* procedures (Zheng et al. (2020b)) are used for design.

The paper is organized as follows: the general setup is given in Section II. The robust controller synthesis with uncertainty on the coprime factors is included in Section III. The sub-optimality guarantees are discussed in Section IV. A brief discussion on the closed-loop system identification is provided in Section V with end-to-end sample complexity results. Conclusions and future possible directions are given in Section V. For the extended version of this manuscript, including the complete proofs and algorithms we refer to Zhang et al. (2021).

2. General Setup and Technical Preliminaries

A standard unity feedback configuration is depicted in Figure 1, where $\mathbf{G} \in \mathbb{R}(z)^{p\times m}$ is a multivariable LTI plant and $\mathbf{K} \in \mathbb{R}(z)^{m\times p}$ is an LTI controller. Here w, ν and r are the input disturbance, sensor noise and reference signal respectively while u, z and y are the controls, regulated signals and measurements vectors, respectively. If all the closed–loop maps from the exogenous signals $[r^T \;\; w^T \;\; \nu^T \;]^T$ to any point inside the feedback loop of Figure 1 are stable, then $\mathbf{K}$ is said to be an (internally) stabilizing controller of $\mathbf{G}$ or equivalently that $\mathbf{K}$ stabilizes $\mathbf{G}$.

	Nomenclature of Basic Notation
TFM	Transfer Function Matrix
DCF, LCF	Doubly Coprime Factorization, Left Coprime Factorization
$\mathbb{R}(z)^{p\times m}$	Set of $p \times m$ TFMs having all entries real–rational transfer functions in z
$\mathtt{pt}$	Superscript for the true model of the plant or for a stabilizing controller designed on the basis of the aforementioned model (*e.g.* $\mathbf{G}^{\mathtt{pt}}$, $\mathbf{K}^{\mathtt{pt}}$)
$\mathtt{md}$	Superscript for the nominal/estimated model of the plant or for a stabilizing controller designed on the basis of the aforementioned model (*e.g.* $\mathbf{G}^{\mathtt{md}}$, $\mathbf{K}^{\mathtt{md}}$)
$\mathbf{G}^{\mathtt{pt}}$, $\mathbf{K}^{\mathtt{opt}}$	The true plant and its optimal $\mathcal{H}_2$ controller
$\mathbf{G}^{\mathtt{md}}$, $\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}}$	The estimated/nominal model of the plant and the stabilizing controller designed on the basis of the aforementioned model, as a function of the Youla parameter $\mathbf{Q}$
$\mathbf{T}^{\ell\varepsilon}$	The TFM of the (closed-loop) map having ε as input and ℓ as output
$\mathbf{T}^{\ell\varepsilon}_{\mathbf{Q}}$	The TFM of the (closed-loop) map from the exogenous signal ε to the signal ℓ inside the feedback loop, as a function of the Youla parameter $\mathbf{Q}$

Proposition 2.1 *Given a TFM* $\mathbf{K} \in \mathbb{R}(z)^{m\times p}$*, a fractional representation of the form* $\mathbf{K} = \mathbf{R}^{-1}\mathbf{S}$ *with* $\mathbf{R} \in \mathbb{R}(z)^{m\times m}$*,* $\mathbf{S} \in \mathbb{R}(z)^{m\times p}$ *is called a* left factorization *of* $\mathbf{K}$*. If* $\mathbf{K} = \mathbf{Y}^{-1}\mathbf{X}$ *is a left factorization of* $\mathbf{K}$ *then any other left factorization of* $\mathbf{K}$ *such as* $\mathbf{K} = \mathbf{R}^{-1}\mathbf{S}$ *is of the form* $\mathbf{R} = \mathbf{U}\mathbf{Y}$*,* $\mathbf{S} = \mathbf{U}\mathbf{X}$*, for some invertible TFM* $\mathbf{U}$*.*

Given a plant $\mathbf{G} \in \mathbb{R}(z)^{p\times m}$, a left coprime factorization of $\mathbf{G}$ is defined by $\mathbf{G} = \widetilde{\mathbf{M}}^{-1}\widetilde{\mathbf{N}}$, with $\widetilde{\mathbf{N}} \in \mathbb{R}(z)^{p\times m}$, $\widetilde{\mathbf{M}} \in \mathbb{R}(z)^{p\times p}$ both stable and satisfying $\widetilde{\mathbf{M}}\widetilde{\mathbf{Y}} + \widetilde{\mathbf{N}}\widetilde{\mathbf{X}} = I_p$, for certain stable TFMs $\widetilde{\mathbf{X}} \in \mathbb{R}(z)^{m\times p}$, $\widetilde{\mathbf{Y}} \in \mathbb{R}(z)^{p\times p}$. Analogously, a right coprime factorization of $\mathbf{G}$ is defined by $\mathbf{G} = \mathbf{N}\mathbf{M}^{-1}$ with both factors $\mathbf{N} \in \mathbb{R}(z)^{p\times m}$, $\mathbf{M} \in \mathbb{R}(z)^{m\times m}$ being stable and for which there exist $\mathbf{X} \in \mathbb{R}(z)^{m\times p}$, $\mathbf{Y} \in \mathbb{R}(z)^{m\times m}$ also stable, satisfying $\mathbf{Y}\mathbf{M} + \mathbf{X}\mathbf{N} = I_m$ (Vidyasagar, 1985, Ch. 4, Corollary 17), with I_m being the identity matrix.

Definition 2.2 *(Vidyasagar, 1985, Ch.4, Remark pp. 79) A collection of eight stable TFMs* $\big(\mathbf{M}, \mathbf{N}, \widetilde{\mathbf{M}}, \widetilde{\mathbf{N}}, \mathbf{X}, \mathbf{Y}, \widetilde{\mathbf{X}}, \widetilde{\mathbf{Y}}\big)$ *is called a* doubly coprime factorization *of* $\mathbf{G}$ *if* $\widetilde{\mathbf{M}}$ *and* $\mathbf{M}$ *are invertible, yield the factorizations* $\mathbf{G} = \widetilde{\mathbf{M}}^{-1}\widetilde{\mathbf{N}} = \mathbf{N}\mathbf{M}^{-1}$*, and satisfy the following equality (Bézout's identity):*

$$\begin{bmatrix} \widetilde{\mathbf{M}} & \widetilde{\mathbf{N}} \\ -\mathbf{X} & \mathbf{Y} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{Y}} & -\mathbf{N} \\ \widetilde{\mathbf{X}} & \mathbf{M} \end{bmatrix} = I_{p+m}, \tag{3}$$

Theorem 2.3 **(Youla-Kučera)** *(Vidyasagar, 1985, Ch.5, Theorem 1) Let* $\big(\mathbf{M}, \mathbf{N}, \widetilde{\mathbf{M}}, \widetilde{\mathbf{N}}, \mathbf{X}, \mathbf{Y}, \widetilde{\mathbf{X}}, \widetilde{\mathbf{Y}}\big)$ *be a doubly coprime factorization of* $\mathbf{G}$*. Any controller* $\mathbf{K}_{\mathbf{Q}}$ *stabilizing the plant* $\mathbf{G}$*, in the feedback interconnection of Figure 1, can be written as*

$$\mathbf{K}_{\mathbf{Q}} = \mathbf{Y}_{\mathbf{Q}}^{-1}\mathbf{X}_{\mathbf{Q}} = \widetilde{\mathbf{X}}_{\mathbf{Q}}\widetilde{\mathbf{Y}}_{\mathbf{Q}}^{-1}, \tag{4}$$

where $\mathbf{X}_{\mathbf{Q}}$*,* $\widetilde{\mathbf{X}}_{\mathbf{Q}}$*,* $\mathbf{Y}_{\mathbf{Q}}$ *and* $\widetilde{\mathbf{Y}}_{\mathbf{Q}}$ *are defined as:*

$$\mathbf{X}_{\mathbf{Q}} \overset{def}{=} \mathbf{X} + \mathbf{Q}\widetilde{\mathbf{M}}, \quad \widetilde{\mathbf{X}}_{\mathbf{Q}} \overset{def}{=} \widetilde{\mathbf{X}} + \mathbf{M}\mathbf{Q}, \quad \mathbf{Y}_{\mathbf{Q}} \overset{def}{=} \mathbf{Y} - \mathbf{Q}\widetilde{\mathbf{N}}, \quad \widetilde{\mathbf{Y}}_{\mathbf{Q}} \overset{def}{=} \widetilde{\mathbf{Y}} - \mathbf{N}\mathbf{Q} \tag{5}$$

for some stable $\mathbf{Q}$ *in* $\mathbb{R}(z)^{m\times p}$*. It also holds that* $\mathbf{K}_{\mathbf{Q}}$ *from (4) stabilizes* $\mathbf{G}$*, for any stable* $\mathbf{Q}$ *in* $\mathbb{R}(z)^{m\times p}$*.*

Proposition 2.4 *Starting from any doubly coprime factorization (3), the following identity*

$$\begin{bmatrix} \mathbf{U}_1\widetilde{\mathbf{M}} & \mathbf{U}_1\widetilde{\mathbf{N}} \\ -\mathbf{U}_2\mathbf{X}_\mathbf{Q} & \mathbf{U}_2\mathbf{Y}_\mathbf{Q} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{Y}}_\mathbf{Q}\mathbf{U}_1^{-1} & -\mathbf{N}\mathbf{U}_2^{-1} \\ \widetilde{\mathbf{X}}_\mathbf{Q}\mathbf{U}_1^{-1} & \mathbf{M}\mathbf{U}_2^{-1} \end{bmatrix} = I_{p+m}. \tag{6}$$

provides the class of all *doubly coprime factorizations of* $\mathbf{G}$*, where* $\mathbf{Q}$ *is stable in* $\mathbb{R}(z)^{m\times p}$ *and* $\mathbf{U}_1 \in \mathbb{R}(z)^{p\times p}$, $\mathbf{U}_2 \in \mathbb{R}(z)^{m\times m}$ *are both unimodular (*i.e. *stable with stable inverses).*

Theorem 2.5 (Dual Youla-Kuçera)*(Hof and Schrama (1992)) Let* $\left(\mathbf{M}, \mathbf{N}, \widetilde{\mathbf{M}}, \widetilde{\mathbf{N}}, \mathbf{X}, \mathbf{Y}, \widetilde{\mathbf{X}}, \widetilde{\mathbf{Y}}\right)$ *be a doubly coprime factorization of* $\mathbf{G}$*. Any plant* $\mathbf{G}_R$ *stabilized by a fixed controller* $\mathbf{K}$*, can be written as*

$$\mathbf{G}_\mathbf{R} = \widetilde{\mathbf{M}}_\mathbf{R}^{-1}\widetilde{\mathbf{N}}_\mathbf{R} = \mathbf{N}_\mathbf{R}\mathbf{M}_\mathbf{R}{}^{-1}, \tag{7}$$

where $\mathbf{M}_\mathbf{R}$, $\widetilde{\mathbf{M}}_\mathbf{R}$, $\mathbf{N}_\mathbf{R}$ *and* $\widetilde{\mathbf{N}}_\mathbf{R}$ *are defined as:*

$$\mathbf{M}_\mathbf{R} \stackrel{def}{=} \mathbf{M} - \widetilde{\mathbf{X}}\mathbf{R}, \quad \widetilde{\mathbf{M}}_\mathbf{R} \stackrel{def}{=} \widetilde{\mathbf{M}} - \mathbf{R}\mathbf{X}, \mathbf{N}_\mathbf{R} \stackrel{def}{=} \mathbf{N} + \widetilde{\mathbf{Y}}\mathbf{R}, \quad \widetilde{\mathbf{N}}_\mathbf{R} \stackrel{def}{=} \widetilde{\mathbf{N}} + \mathbf{R}\mathbf{Y} \tag{8}$$

for some stable $\mathbf{R}$ *in* $\mathbb{R}(z)^{p\times m}$.

3. The LQG Robust Controller Synthesis

Given a DCF of the nominal model of the plant $\mathbf{G}^{\mathtt{md}}$, we can write the Bezout identity that incorporates the corresponding Youla parameterization of all stabilizing controller as:

$$\begin{bmatrix} \widetilde{\mathbf{M}}^{\mathtt{md}} & \widetilde{\mathbf{N}}^{\mathtt{md}} \\ -\mathbf{X}_\mathbf{Q}^{\mathtt{md}} & \mathbf{Y}_\mathbf{Q}^{\mathtt{md}} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{Y}}_\mathbf{Q}^{\mathtt{md}} & -\mathbf{N}^{\mathtt{md}} \\ \widetilde{\mathbf{X}}_\mathbf{Q}^{\mathtt{md}} & \mathbf{M}^{\mathtt{md}} \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_m \end{bmatrix}, \tag{9}$$

where $\mathbf{Q}$ denotes as usually the Youla parameter.

Definition 3.1 (Model Uncertainty Set) *The* γ*-radius* model uncertainty *set (for the nominal plant* $\mathbf{G}^{\mathtt{md}}$ *with* $\Delta_{\widetilde{\mathbf{M}}}$, $\Delta_{\widetilde{\mathbf{N}}}$ *both stable) is defined as:*

$$\mathcal{G}_\gamma \stackrel{\text{def}}{=} \{\mathbf{G} = \widetilde{\mathbf{M}}^{-1}\widetilde{\mathbf{N}} \,|\, \widetilde{\mathbf{M}} = (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}}), \widetilde{\mathbf{N}} = (\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}}); \quad \left\| \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \right\|_\infty < \gamma\} \tag{10}$$

Definition 3.2 (γ-Robustly Stabilizable) *A stabilizing controller* $\mathbf{K}^{\mathtt{md}}$ *of the nominal plant is said to be* γ*-robustly stabilizable iff* $\mathbf{K}^{\mathtt{md}}$ *stabilizes not only* $\mathbf{G}^{\mathtt{md}}$ *but also all plants* $\mathbf{G} \in \mathcal{G}_\gamma$.

Assumption 1 *It is assumed that the true plant, denoted by* $\mathbf{G}^{\mathtt{pt}}$*, belongs to the model uncertainty set introduced in Definition 3.1, i.e. that there exist stable* $\Delta_{\widetilde{\mathbf{M}}}$, $\Delta_{\widetilde{\mathbf{N}}}$ *with* $\left\| \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \right\|_\infty < \gamma$ *for which* $\mathbf{G}^{\mathtt{pt}} = (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}})^{-1}(\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}})$.

In the presence of additive uncertainty on the coprime factors the Bezout identity in (9) no longer holds, however, the following holds for *certain stable* $\Delta_\mathbf{M}$, $\Delta_\mathbf{N}$ *factors*:

$$\begin{bmatrix} (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}}) & (\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}}) \\ -\mathbf{X}_\mathbf{Q}^{\mathtt{md}} & \mathbf{Y}_\mathbf{Q}^{\mathtt{md}} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{Y}}_\mathbf{Q}^{\mathtt{md}} & -(\mathbf{N}^{\mathtt{md}} + \Delta_\mathbf{N}) \\ \widetilde{\mathbf{X}}_\mathbf{Q}^{\mathtt{md}} & (\mathbf{M}^{\mathtt{md}} + \Delta_\mathbf{M}) \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_{11} & O \\ O & \mathbf{\Phi}_{22} \end{bmatrix}. \tag{11}$$

The block diagonal structure of the right hand side term in (11) is due to the fact that $\mathbf{G}^{\mathtt{pt}} = (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}})^{-1}(\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}}) = (\mathbf{N}^{\mathtt{md}} + \Delta_\mathbf{N})(\mathbf{M}^{\mathtt{md}} + \Delta_\mathbf{M})^{-1}$ for the aforementioned *certain stable* $\Delta_\mathbf{M}$, $\Delta_\mathbf{N}$ *factors*.

Lemma 3.3 *A stabilizing controller of the nominal plant* $\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}} = (\mathbf{Y}^{\mathtt{md}}_{\mathbf{Q}})^{-1}\mathbf{X}^{\mathtt{md}}_{\mathbf{Q}} = \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}}(\widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}})^{-1}$ *is* γ*-robustly stabilizing iff for any stable model perturbations* $\Delta_{\widetilde{\mathbf{M}}}, \Delta_{\widetilde{\mathbf{N}}}$ *with* $\left\| \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \right\|_\infty < \gamma$ *the TFM*

$$\mathbf{\Phi}_{11} = I_p + \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \tag{12}$$

from (11) *is unimodular (*i.e. *it is square, stable and has a stable inverse).*

Since $\mathbf{\Phi}_{11}$ in (12) clearly depends on the Youla parameter (via the right coprime factors of the controller), the condition for the γ-robust stabilizability of the controller can be recast in the following particular form, which will be instrumental in the sequel:

Theorem 3.4 *The Youla parameterization yields a* γ*-robustly stabilizing controller iff its corresponding Youla parameter satisfies* $\left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty \leq \dfrac{1}{\gamma}$.

Proposition 3.5 *The square root of the LQG cost function from* (2) *can be assimilated to:*

$$\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}}) \overset{def}{=} \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \mathbf{\Phi}_{11}^{-1} \begin{bmatrix} (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}}) & (\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}}) \end{bmatrix} \right\|_{\mathcal{H}_2} \tag{13}$$

whereas, $\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}})$ *denotes the following closed loop responses:*

$$\begin{bmatrix} \mathbf{T}^{y\nu}_{\mathbf{Q}} & \mathbf{T}^{yw}_{\mathbf{Q}} \\ \mathbf{T}^{u\nu}_{\mathbf{Q}} & \mathbf{T}^{uw}_{\mathbf{Q}} \end{bmatrix} = \begin{bmatrix} (I_p + \mathbf{G}^{\mathtt{pt}}\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}})^{-1} & (I_p + \mathbf{G}^{\mathtt{pt}}\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}})^{-1}\mathbf{G}^{\mathtt{pt}} \\ \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}}(I_p + \mathbf{G}^{\mathtt{pt}}\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}})^{-1} & \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}}(I_p + \mathbf{G}^{\mathtt{pt}}\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}})^{-1}\mathbf{G}^{\mathtt{pt}} \end{bmatrix}. \tag{14}$$

In this setup, the robust LQG control problem reads:

$$\begin{aligned} \min_{\mathbf{Q}\text{ stable}} \quad & \max_{\left\| \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \right\|_\infty < \gamma} \mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}}) \\ \text{s.t.} \quad & \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty \leq \frac{1}{\gamma}. \end{aligned} \tag{15}$$

whose solution, obtained for the optimal Youla parameter $\mathbf{Q}_*$ in (15) will be denoted by $\widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*}, \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*}$ such that the optimal, robust controller reads $\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}_*} = \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*}(\widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*})^{-1}$. Note that the non-convexity of the robust LQG problem is caused by the additive uncertainty on the coprime factors. In order to circumvent this, an upper bound on the $\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}})$ cost functional will be derived, much in the spirit of Dean et al. (2018) and Zheng et al. (2020a). This bound will further be exploited to derive a quasi-convex approximation for the robust LQG control problem in the next subsection.

3.1. Quasi-convex formulation

Proposition 3.6 *Given any γ-robustly stabilizing controller satisfying $\left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty < \frac{1}{\gamma}$ then for any additive model perturbations $\left\| \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \right\|_\infty < \gamma$, the cost functional of the robust LQG problem from* (15) *admits the upper bound:*

$$\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}}) \leq \frac{1}{1 - \gamma \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty} \left[h\left(\gamma, \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty \right) \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_{\mathcal{H}_2} \right] \tag{16}$$

where $h\left(\gamma, \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty \right) \overset{def}{=} \left(1 + \gamma \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty \right) \left(\left\| \begin{bmatrix} \widetilde{\mathbf{M}}^{\mathtt{md}} & \widetilde{\mathbf{N}}^{\mathtt{md}} \end{bmatrix} \right\|_\infty + \gamma \right).$

We state next one of the main results, providing an aptly designed approximation of the robust LQG control problem by means of the LQG cost upper bound from (16).

Theorem 3.7 *For the true plant, $\mathbf{G}^{\mathtt{pt}} \in \mathcal{G}_\gamma$ and $(\forall)\alpha > 0$, the robust LQG control problem in* (15) *admits the following upper bound:*

$$\begin{aligned} \min_{\delta \in [0, 1/\gamma)} \frac{1}{1 - \gamma\delta} \min_{\mathbf{Q} \textit{ stable}} & \left(h(\gamma, \alpha) \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_{\mathcal{H}_2} \right) \\ \textit{s.t.} & \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty \leq \min\{\delta, \alpha\}, \end{aligned} \tag{17}$$

where as before, $h\left(\gamma, \alpha\right) = (1 + \gamma\alpha) \left(\left\| \begin{bmatrix} \widetilde{\mathbf{M}}^{\mathtt{md}} & \widetilde{\mathbf{N}}^{\mathtt{md}} \end{bmatrix} \right\|_\infty + \gamma \right).$

Remark 1 *For each fixed δ, the inner optimization is convex but the dimension of $\mathbf{Q}(z)$ remains infinite. For numerical computation, a (FIR) truncation is considered on $\mathbf{Q}(\mathbf{z})$ which will convert the inner optimization problem into a Semi-Definite Program (Zhang et al., 2021, Appendix E).*

Remark 2 *(Feasibility) The optimization problem proposed in Theorem-3.7 is a quasi-convex relaxation of the original robust LQG problem* (15)*, due to the added constraint: $\left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}} \end{bmatrix} \right\|_\infty < \alpha$, where α is a given positive constant. Since the relaxation bound is proportional to α, it should be chosen as small as possible, however, α can't be made arbitrarily small and therefore the feasibility of the quasi-convex program from Theorem-3.7 cannot be guaranteed. This is caused by the fact that α must be assimilated to the norm of a LCF of a stabilizing controller for* the true plant*, whose H_∞ attenuation is simultaneously greater or equal to the one of the optimal H_∞ controller* for the nominal model *(see Theorem 3.4). This reflects the fact that the feasibility of the closed loop "learning" problem depends inherently on the performance (with respect to the model uncertainty of the plant) of the initially chosen stabilizing controller (the one with which the closed loop "learning" is being performed).*

4. Analysis of End-to-End Performance

If we denote by $\mathbf{K}^{\mathtt{opt}}$ the optimal $\mathcal{H}_2$ controller for the true plant, then by Assumption 1 there exist stable additive factors such that $\mathbf{G}^{\mathtt{pt}} = (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}})^{-1}(\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}}) = (\mathbf{N}^{\mathtt{md}} + \Delta_{\mathbf{N}})(\mathbf{M}^{\mathtt{md}} + \Delta_{\mathbf{M}})^{-1}$ and furthermore, there always exists a Bezout identity of the true plant that features the optimal controller $\mathbf{K}^{\mathtt{opt}} = (\mathbf{Y}^{\mathtt{opt}})^{-1}\mathbf{X}^{\mathtt{opt}} = \widetilde{\mathbf{X}}^{\mathtt{opt}}(\widetilde{\mathbf{Y}}^{\mathtt{opt}})^{-1}$ as its "central controller", thus reading:

$$\begin{bmatrix} (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}}) & (\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}}) \\ -\mathbf{X}^{\mathtt{opt}} & \mathbf{Y}^{\mathtt{opt}} \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{opt}} & -(\mathbf{N}^{\mathtt{pt}} + \Delta_{\mathbf{N}}) \\ \widetilde{\mathbf{X}}^{\mathtt{opt}} & (\mathbf{M}^{\mathtt{pt}} + \Delta_{\mathbf{M}}) \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_m \end{bmatrix} \tag{18}$$

Consequently, the square root of the LQG cost functional for optimal controller is given by:

$$\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{opt}}) \overset{def}{=} \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{opt}} \\ \widetilde{\mathbf{X}}^{\mathtt{opt}} \end{bmatrix} \begin{bmatrix} (\widetilde{\mathbf{M}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{M}}}) & (\widetilde{\mathbf{N}}^{\mathtt{md}} + \Delta_{\widetilde{\mathbf{N}}}) \end{bmatrix} \right\|_{\mathcal{H}_2} \tag{19}$$

Next, the main result on the sub-optimality guarantee for the performance of the robust controller with model uncertainty of radius γ is stated:

Theorem 4.1 *Let $\mathbf{K}^{\mathtt{opt}}$ be the optimal LQG controller and $\mathbf{G}^{\mathtt{pt}}$ be the model of the true plant, with modeling error uncertainty satisfying $\left\| \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \right\|_{\infty} < \gamma$. Furthermore, let $\mathbf{Q}_*$ and δ_* denote the solution to (17). Then, when applying the resulting controller $\mathbf{K}^{\mathtt{md}}_{\mathbf{Q}_*}$ in feedback interconnection with the true plant $\mathbf{G}^{\mathtt{pt}}$, the relative error in the LQG cost is upper bounded by:*

$$\frac{\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}_*})^2 - \mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{opt}})^2}{\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{opt}})^2} \leq \frac{1}{\left(1 - \gamma \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty}\right)^2} \times g\left(\gamma, \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty}\right)^2 - 1, \tag{20}$$

where $g\left(\gamma, \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty}\right) \overset{def}{=} \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty} \left(1 + \gamma \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty}\right) \left(\left\| \begin{bmatrix} \widetilde{\mathbf{M}}^{\mathtt{md}} & \widetilde{\mathbf{N}}^{\mathtt{md}} \end{bmatrix} \right\|_{\infty} + \gamma \right).$

Remark 3 *(Optimality vs Robustness) If $\gamma \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty} = \eta$, it's easy to observe that $\eta \in (0, 1)$. Then it's immediate to see that the upper bound of the relative error in the LQG cost increases as a function of η. The price of obtaining a faster rate is that the controller becomes less robust to model uncertainty as pointed out in Mania et al. (2019), Zheng et al. (2020a). It holds for this case too as shown in Theorem 4.1. In practice, using a relatively large value for η forces a trade-off of optimality for robustness in the controller design procedure. In general, optimality stands i.e. better controller performance is guaranteed as η goes closer to 0 and better robustness performance is guaranteed as η goes closer to 1 with the upper bound (20) of relative error in LQG cost might be large. This is shown with an example below.*
Let's set $\eta = \frac{1}{5}$. Then by Theorem 4.1 relative error in the LQG cost is

$$\frac{\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}_*})^2 - \mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{opt}})^2}{\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{opt}})^2} \leq 2 \times g\left(\gamma, \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty}\right)^2 - 1. \tag{21}$$

Hence, the the relative error in the LQG cost grows as $\mathcal{O}(\gamma^2)$ as long as $\gamma \left\| \begin{bmatrix} \widetilde{\mathbf{Y}}^{\mathtt{md}}_{\mathbf{Q}_*} \\ \widetilde{\mathbf{X}}^{\mathtt{md}}_{\mathbf{Q}_*} \end{bmatrix} \right\|_{\infty} < \frac{1}{5}$.

5. Closed Loop Identification Scheme (Zhang et al., 2021, Appendix G)

Figure 2 in (Zhang et al., 2021, Subsection 4.2) depicts the closed-loop identification setup of a potentially unstable *noise contaminated plant* $\mathbf{G}^{\mathtt{md}}$ with control input u, noise ν (taken w = 0) and output measurement y (where u and ν are assumed independent and stationary), provided that some initial stabilizing controller $\mathbf{K}^{\mathtt{md}}$ is available beforehand. The key idea dating back to Anderson (1998) is to identify the *stable* dual-Youla parameter $\mathbf{R}^{\mathtt{md}}$ from Theorem 2.5 rather than $\mathbf{G}^{\mathtt{md}}$, thus recasting the problem in a standard, open-loop identification form. More specifically, direct inspection of Figure 2 from (Zhang et al., 2021, Subsection 4.2) shows that

$$e_2 = \mathbf{R}^{\mathtt{md}} e_1 + (\widetilde{\mathbf{M}}^{\mathtt{md}} - \mathbf{R}^{\mathtt{md}}\mathbf{X}^{\mathtt{md}})\nu = \mathbf{R}^{\mathtt{md}}(e_1 - \mathbf{X}^{\mathtt{md}}\nu) + \widetilde{\mathbf{M}}^{\mathtt{md}}\nu. \tag{22}$$

In (22) we have the knowledge of $\mathbf{X}^{\mathtt{md}}$; e_1 and e_2 are available from measurements with noise ν. The recent algorithm from Sarkar and Rakhlin (2019) given below is employed toward identifying the dual-Youla parameter.

5.1. Identification Algorithm (Sarkar and Rakhlin (2019))

The state space representation of $\mathbf{R}^{\mathtt{md}}$ in $e_2 = \mathbf{R}^{\mathtt{md}}u + \widetilde{\mathbf{M}}^{\mathtt{md}}\nu$ with $u = e_1 - \mathbf{X}^{\mathtt{md}}\nu$ is:

$$\begin{aligned} l_{t+1} &= A_R l_t + B_R u_t + \eta_{t+1} \\ z_t &= C_R l_t + \widetilde{\mathbf{M}}^{\mathtt{md}}\nu_t \end{aligned} \tag{23}$$

Assumption 2 *The noise process $\{\eta_t\}_{t=1}^{\infty}$ in the dynamics of $\mathbf{R}^{\mathtt{md}}$ are i.i.d., and isotropic with sub-gaussian parameter 1. The noise process $\{r_t\}_{t=1}^{\infty}$, $\{w_t\}_{t=1}^{\infty}$ and $\{\nu_t\}_{t=1}^{\infty}$ are Gaussian processes with mean $m_r(t) = m_w(t) = m_\nu(t) = 0$, and their spectral density $\phi_r(\omega)$, $\phi_w(\omega)$ and $\phi_\nu(\omega)$.*

Assumption 3 *There exist constants $\beta, R \geq 1$ s.t. $\|\mathcal{T}_{0,\infty}\|_2 < \beta$ and $\dfrac{\|\mathcal{T}\mathcal{O}_{k,d}\|_2}{\|\mathcal{T}_{0,\infty}\|_2} \leq \mathcal{R}$. Note that β exists since $\mathbf{R}^{\mathtt{md}}$ is stable.*

5.1.1. PROBABILISTIC GUARANTEES

Let's define, $T_*(\delta) = \inf\{T | d_*(T,\delta) \in \mathcal{D}(T), d_*(T,\delta) \leq 2d_*(\frac{T}{256},\delta)\}$ where, $d_*(T,\delta) = \inf\{d|\ 16\beta\mathcal{R}\alpha(d) \geq \left\|\widehat{\mathcal{H}}_{0,\widehat{d},\widehat{d}} - \widehat{\mathcal{H}}_{0,\infty,\infty}\right\|_2\}$, with $\widehat{\mathcal{H}}_{p,q,r}$ is the (p,q,r) - dimensional estimated Hankel matrix. Whenever $T \geq T_*(\delta)$ for the failure probability δ, then it follows with the probability at least $(1-\delta)$ that

$$\left\|\widehat{\mathcal{H}}_{0,\widehat{d},\widehat{d}} - \widehat{\mathcal{H}}_{0,\infty,\infty}\right\|_2 \leq 12c\beta\mathcal{R}\left(\sqrt{\frac{m\widehat{d} + p\widehat{d}^2 + \widehat{d}\log(T/\delta)}{T}}\right). \tag{24}$$

5.2. Sample Complexity

Lemma 5.1 *The norm of the identification error incurred by the proposed scheme is bounded by:*

$$\left\|\mathbf{R}^{\mathtt{md}} - \mathbf{R}^{\mathtt{pt}}\right\|_\infty \leq \left\|\widehat{\mathcal{H}}_{0,\widehat{d},\widehat{d}} - \widehat{\mathcal{H}}_{0,\infty,\infty}\right\|_2 \leq 12c\beta\mathcal{R}\left(\sqrt{\frac{m\widehat{d} + p\widehat{d}^2 + \widehat{d}\log(T/\delta)}{T}}\right)$$

Finally, the error on the model uncertainty can be directly checked using Lemma 5.1 as:

$$\left\| [\Delta_{\widetilde{\mathbf{M}}} \;\; \Delta_{\widetilde{\mathbf{N}}}] \right\|_\infty = \left\| \left[(\widetilde{\mathbf{M}}^{\mathtt{md}} + \mathbf{X}^{\mathtt{md}}\mathbf{R}^{\mathtt{md}}) \;\; (\widetilde{\mathbf{N}}^{\mathtt{md}} - \mathbf{Y}^{\mathtt{md}}\mathbf{R}^{\mathtt{md}})\right] - \left[(\widetilde{\mathbf{M}}^{\mathtt{md}} + \mathbf{X}^{\mathtt{md}}\mathbf{R}^{\mathtt{pt}}) \;\; (\widetilde{\mathbf{N}}^{\mathtt{md}} - \mathbf{Y}^{\mathtt{md}}\mathbf{R}^{\mathtt{pt}})\right] \right\|_\infty$$
$$= \left\| \begin{bmatrix} \mathbf{X}^{\mathtt{md}} & \mathbf{Y}^{\mathtt{md}} \end{bmatrix} (\mathbf{R}^{\mathtt{md}} - \mathbf{R}^{\mathtt{pt}}) \right\|_\infty \leq \left\| \begin{bmatrix} \mathbf{X}^{\mathtt{md}} & \mathbf{Y}^{\mathtt{md}} \end{bmatrix} \right\|_\infty \left\| \mathbf{R}^{\mathtt{md}} - \mathbf{R}^{\mathtt{pt}} \right\|_\infty$$

Consequently, the uncertainty level on the LCF of the model satisfies $\left\| \begin{bmatrix} \Delta_{\widetilde{\mathbf{M}}} & \Delta_{\widetilde{\mathbf{N}}} \end{bmatrix} \right\|_\infty < \gamma$.

Theorem 5.2 *Define* $s = 144\left\| \begin{bmatrix} \mathbf{X}^{\mathtt{md}} & \mathbf{Y}^{\mathtt{md}} \end{bmatrix} \right\|_\infty^2 c^2\beta^2\mathcal{R}^2$*. Then, the robust controller will achieve the relative cost within the bound with probability* $(1-\delta)$ *provided* $T \geq \max\{T_s, T_*(\delta)\}$*. Here,* T_s *is the right most zero of* $g(T) = \gamma^2 T - s\widehat{dl}og(T/\delta) - s(m\widehat{d} + p\widehat{d}^2)$*. If* $g(T)$ *doesn't have any zero for* $T > 0$*, then define* $T_s = 0$ *and* $T_*(\delta) = \inf\{T | d_*(T,\delta) \in \mathcal{D}(T), d_*(T,\delta) \leq 2d_*(\frac{T}{256}, \delta)\}$*,* *where,* $d_*(T,\delta) = \inf\{d | 16\beta\mathcal{R}\alpha(d) \geq \left\| \widehat{\mathcal{H}}_{0,d,d} - \widehat{\mathcal{H}}_{0,\infty,\infty} \right\|_2\}$*,*
$\mathcal{D}(T) = \{d \in \mathbb{N} | d \leq \dfrac{T}{cm^2 log^3(Tm/\delta)}\}$ *and* $\alpha(h) = \sqrt{h}.(\sqrt{\dfrac{m + hp + log(T/\delta)}{T}})$*.*

Combining Theorem 5.2 with Theorem 4.1, it follows that with high probability the suboptimality gap behaves as

$$\frac{\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{md}}_{\mathbf{Q}_*})^2 - \mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{opt}})^2}{\mathcal{H}(\mathbf{G}^{\mathtt{pt}}, \mathbf{K}^{\mathtt{opt}})^2} \sim \mathcal{O}\Big(\sqrt{\frac{logT}{T}}\Big)$$

Finally, we note here that the resulted sample complexity is on par with the existing methods fromZheng et al. (2020a) and Dean et al. (2018).

6. Conclusion and Future work

In this paper, we have provided the sample complexity bounds for a robust controller synthesis procedure for LQG problems with unknown dynamics, able to cope with unstable plants. We combined finite-time, non-parametric LTI system identification with the Youla parameterization for robust stabilization under uncertainty on the coprime factors of the plant. One exciting avenue for future research is the online learning LQG control problem under the same type of model uncertainty. Another direction is to work out the sample complexity of learning the optimal state feedback (LQR) controller in tandem with the optimal state-observer (Kalman Filter (Tsiamis et al. (2020))) for a potentially unstable system. Combining these two results, should yield precisely the optimal LQG controller discussed above and reveal the *separation principle* within this framework.

References

Brian D.O. Anderson. From youla–kucera to identification, adaptive and nonlinear control. *Automatica*, 34(12):1485–1506, 1998.

Ross Boczar, Nikolai Matni, and Benjamin Recht. Finite-data performance guarantees for the output-feedback control of an unknown system. *IEEE Conference on Decision and Control (CDC)*, page 2994–2999, 2018. doi:https://doi.org/10.1109/CDC.2018.8618658.

Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. Regret bounds for robust adaptive control of the linear quadratic regulator. *Conference on Neural Information Processing Systems*, page 4188–4197, 2018.

Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. On the sample complexity of the linear quadratic regulator. *Foundations of Computational Mathematics*, 20:633–679, August 2020. doi:10.1007/s10208-019-09426-y.

Paul M. J. Van Den Hof and Ruud J. P. Schrama. Identification and control - closed-loop issues. *Automatica*, 31(12):1751 – 1770, 1992.

Horia Mania, Stephen Tu, and Benjamin Recht. Certainty equivalence is efficient for linear quadratic control. https://arxiv.org/abs/1902.07826v2, 2019.

Samet Oymak and Necmiye Ozay. Non-asymptotic identification of lti systems from a single trajectory. *American Control Conference (ACC)*, 2019. doi:10.23919/ACC.2019.8814438.

Tuhin Sarkar and Alexander Rakhlin. Near optimal finite time identification of arbitrary linear dynamical systems. *International Conference on Machine Learning*, 97:5610–5618, 2019.

Tuhin Sarkar, Alexander Rakhlin, and Munther A. Dahleh. Nonparametric finite time lti system identification. https://arxiv.org/abs/1902.01848, 2020.

Anastasios Tsiamis, Nikolai Matni, and George J. Pappas. Sample complexity of kalman filtering for unknown systems. *2nd Annual Conference on Learning for Dynamics and Control*, 120, 2020.

M. Vidyasagar. *Control System Synthesis: A Factorization Approach*. Cambridge, MA: MIT Press, Signal Processing, Optimization, and Control Series, 1985.

Yifei Zhang, Sourav Kumar Ukil, Ephraim Neimand, Serban Sabau, and Myron E Hohil. Sample complexity of the robust lqg regulator with coprime factors uncertainty. https://arxiv.org/abs/2109.14164, 2021.

Yang Zheng, Luca Furieri, Maryam Kamgarpour, and Na Li. Sample complexity of linear quadratic gaussian (lqg) control for output feedback systems. http://proceedings.mlr.press/v144/zheng21b.html, pages 1–33, 2020a.

Yang Zheng, Luca Furieri, Antonis Papachristodoulou, Na Li, and Maryam Kamgarpour. On the equivalence of youla, system-level, and input–output parameterizations. *IEEE Transactions on Automatic Control*, 68:413–420, 2020b. doi:10.1109/TAC.2020.2979785.