Near-Optimal Design of Safe Output Feedback Controllers from Noisy Data

Luca Furieri, Baiwei Guo⋆, Andrea Martin⋆ and Giancarlo Ferrari-Trecate

Abstract*— As we transition towards the deployment of data-driven controllers for black-box cyberphysical systems, complying with hard safety constraints becomes a primary concern. Two key aspects should be addressed when input-output data are corrupted by noise: how much uncertainty can one tolerate without compromising safety, and to what extent is the control performance affected? By focusing on finite-horizon constrained linear-quadratic problems, we provide an answer to these questions in terms of the model mismatch incurred during a preliminary identification phase. We propose a control design procedure based on a quasiconvex relaxation of the original robust problem and we prove that, if the uncertainty is sufficiently small, the synthesized controller is safe and near-optimal, in the sense that the suboptimality gap increases *linearly* with the model mismatch level. Since the proposed method is independent of the specific identification procedure, our analysis holds in combination with state-of-the-art behavioral estimators beyond standard least-squares. The main theoretical results are validated by numerical experiments.

Index Terms— Data-driven control, Learning-Based Control, Linear Systems, Optimal Control, Robust control.

I. Introduction

Many safety-critical engineering systems that play a crucial role in our modern society are becoming too complex to be accurately modeled through white-box state-space models [1]. As a consequence, most contemporary control approaches envision unknown black-box systems for which a safe and optimal behavior must be attained by solely relying on a collection of system's output trajectories in response to different inputs.

Controllers for unknown systems can be designed according to two paradigms. *Model-based* methods follow a two-step procedure: first, data are exploited to identify the system parameters, and then a suitable controller is computed for the estimated model. On the other hand, *model-free* methods aim at directly learning an optimal control policy, without explicitly reconstructing an internal representation of the dynamical system. For a description of advantages and limitations of both approaches, we refer to [2], among recent surveys.

Authors are with the École Polytechnique Fédérale de Lausanne (EPFL), Institute of Mechanical Engineering, CH-1015 Lausanne, Switzerland. E-mails: {luca.furieri, baiwei.guo, andrea.martin, giancarlo.ferraritrecate}@epfl.ch.
⋆Baiwei Guo and Andrea Martin contributed equally to this work.

Given the intricacy of deriving rigorous suboptimality and sample-complexity bounds, most recent model-based and model-free approaches have focused on basic Linear Quadratic Regulator (LQR) and Linear Quadratic Gaussian (LQG) control problems as suitable benchmarks to establish how machine learning can be interfaced to the continuous action spaces typical of control [3]–[9]. For complex control tasks, it is more challenging to perform a thorough probabilistic analysis. Recent advances include [10], [11] for constrained and distributed LQR control with direct state measurements, respectively, and [12] for distributed output-feedback LQG.

Model-based methods may pose a difficulty when it comes to accurately identifying the state-space model of a large-scale system; this is the case, for instance, for complex networked systems such as the power grid, brain and traffic networks [1]. A promising data-driven approach that aims at bypassing a parametric state-space description of the system dynamics, while still being conceptually simple to implement for the users, hinges on the *behavioral framework* [13]. This approach has gained renewed interest with the introduction of Data-EnablEd Predictive Control (DeePC) [14]–[16], which established that constrained output reference tracking can be effectively tackled in a Model-Predictive-Control (MPC) fashion by plugging adequately generated data into a convex optimization problem. The work [17] introduces data-driven formulations for some controller design tasks, and [18] derives stability guarantees for closed-loop control.

In many scenarios, however, exact data are not available. For instance, data can be corrupted by measurement noise or even by malicious attacks intended at fatally compromising the safety [19], the quality, and the reliability of the synthesized control policies. It is therefore essential that data-driven controllers are endowed with robustness guarantees. While some approaches have been suggested in the behavioral framework, e.g. [15], [20]–[22], it remains fairly unexplored how much noise-corrupted data affect the performance and the safety of data-driven control systems. Recently, [23], [24] have derived suboptimality [24] and sample-complexity [23] bounds for LQR through direct behavioral formulations based on 1) Linear Matrix Inequalities (LMI) and 2) the System Level Synthesis (SLS) approach, respectively. A limitation is that the internal system states must be measured, which is unrealistic for several large-scale systems [1]. Furthermore, while [24] proves that for low-enough noise a high-performing and robustly stabilizing controller can be found, the corresponding suboptimality growth rate is not explicitly derived. To address these open points, [25] has formulated a Behavioral Input-

Output Parametrization (BIOP) of linear control policies which makes it possible to derive noise-dependent suboptimality analysis for output-feedback LQG, solely based on a non-parametric estimation of the system open-loop responses. The BIOP can exploit any system identification technique such as the least-square methods of [5], [26] or behavioral maximum-likelihood (ML) estimation [21], [27]. In order to deploy these systems in real-world scenarios, it is however important to include safety guarantees in the analysis. All the mentioned works [5], [23]–[25] do not include such safety requirements.

A. Contributions

We propose a method for designing safe and near-optimal output-feedback control policies for linear systems in finite-horizon. Our approach is solely based on noisy data, and we explicitly characterize the growth rate of the suboptimality as a function of the mismatch between the true and estimated system. First, we develop a new relaxed optimization problem that guarantees safety while robustly accounting for noise-corrupted data. Second, we show that the incurred level of suboptimality converges to zero approximately as a linear function of the model mismatch incurred during a preliminary identification phase. Hence, upon using a consistent system estimator, the proposed controller is near-optimal in the limit of available data growing to infinity. The corresponding analysis differs from that of [5], [25], in that a feasible solution to the proposed optimization problem must be characterized analytically while taking the safety constraints into account. In addition to dealing with constraints in an output-feedback setup — which is the main novelty with respect to [5], [25] and [10] — the effect of the *uncertain* initial condition x_0 must be explicitly tracked in the cost. Indeed, [5] assumed that $x_0 = 0$ thanks to the considered infinite-horizon setting. On a more general level, our analysis has been inspired by [10], which combined robust control tools with classical identification techniques to ensure safety of unknown systems with suboptimality guarantees when states are fully observed. As we only have access to noisy output measurements, we exploit an input-output representation of the plant and analyze four different closed-loop responses to understand how process and output measurement noises impact safety and performance. Suboptimality with respect to the best model-based *open-loop* control input has very recently been analyzed in [28] as a function of the noise-level. Instead, in the present paper we analyze the suboptimality brought about by closed-loop policies. In particular, we show a linear growth rate of the suboptimality in terms of the model mismatch level as compared to the ground-truth constrained output-feedback controller.

A preliminary version of this work has recently appeared in the 60th IEEE Conference on Decision and Control [25]. Differently from [25], this paper includes safety constraints in the analysis, thus addressing a novel and independent set of challenges and results. Furthermore, this work includes all the technical proofs. Last, new numerical experiments are developed to consider safety constraints and to explicitly include the estimation procedure of [21].

B. Paper structure

Assuming knowledge of the underlying dynamics, Section II reviews the optimal control problem of interest and its model-based solution. Section III treats the case where we only have access to noisy input and output data; we propose an optimal control problem that ensures safety against bounded model mismatches, and discuss its numerical implementation. Section IV quantifies the suboptimality incurred by our synthesis procedure as a function of the model mismatch. We present numerical experiments in Section V and conclude the paper in Section VI.

C. Notation

We use $\mathbb{R}$ and $\mathbb{N}$ to denote the sets of real numbers and non-negative integers, respectively. We use I_n to denote the identity matrix of size $n \times n$ and $0_{m \times n}$ to denote the zero matrix of size $m \times n$. We write $\mathbf{x} = \text{vec}(x_1, \ldots, x_N) \in \mathbb{R}^{Nn}$ to denote the vector obtained by stacking together the vectors $x_1, \ldots, x_N \in \mathbb{R}^n$, and $\mathbf{M} = \text{blkdiag}(M_1, \ldots, M_N)$ to denote a block-diagonal matrix with $M_1, \ldots, M_N \in \mathbb{R}^{m \times n}$ on its diagonal block entries. For $\mathbf{M} = \begin{bmatrix} M_1^\mathsf{T} & \ldots & M_N^\mathsf{T} \end{bmatrix}^\mathsf{T}$ we define the block-Toeplitz matrix

$$\text{Toep}_{m \times n}(\mathbf{M}) = \begin{bmatrix} M_1 & 0_{m \times n} & \ldots & 0_{m \times n} \\ M_2 & M_1 & \ldots & 0_{m \times n} \\ \vdots & \vdots & \ddots & \vdots \\ M_N & M_{N-1} & \ldots & M_1 \end{bmatrix}.$$

More concisely, we will write $\text{Toep}(\cdot)$ when the dimensions of the blocks are clear from the context. The Kronecker product between $M \in \mathbb{R}^{m \times n}$ and $P \in \mathbb{R}^{p \times q}$ is denoted as $M \otimes P \in \mathbb{R}^{mp \times nq}$. For a vector $v \in \mathbb{R}^n$ and a matrix $A \in \mathbb{R}^{m \times n}$ we denote as $\|v\|_p$, $\|A\|_p$, their standard p-norm and induced p-norms, respectively. For a row vector $x \in \mathbb{R}^{1 \times n}$ we define $\|x\|_1^\star = \sum_{i=1}^n |x_i|$. The Frobenius norm of a matrix $M \in \mathbb{R}^{m \times n}$ is denoted by $\|M\|_F = \sqrt{\text{Trace}(M^\mathsf{T} M)}$. For a symmetric matrix M, we write $M \succ 0$ or $M \succeq 0$ if it is positive definite or positive semidefinite, respectively. We say that $x \sim \mathcal{D}(\mu, \Sigma)$ if the random variable $x \in \mathbb{R}^n$ follows a distribution with mean $\mu \in \mathbb{R}^n$ and covariance matrix $\Sigma \in \mathbb{R}^{n \times n}, \Sigma \succeq 0$.

A finite-horizon trajectory of length T is a sequence $\omega(0), \omega(1), \ldots, \omega(T-1)$ with $\omega(t) \in \mathbb{R}^n$ for every $t = 0, 1, \ldots, T-1$, which can be compactly written as

$$\boldsymbol{\omega}_{[0,T-1]} = \begin{bmatrix} \omega^\mathsf{T}(0) & \omega^\mathsf{T}(1) & \ldots & \omega^\mathsf{T}(T-1) \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{nT}.$$

When the value of T is clear from the context, we will omit the subscript $[0, T-1]$. For a finite-horizon trajectory $\boldsymbol{\omega}_{[0,T-1]}$ we also define the Hankel matrix of depth L as

$$\mathcal{H}_L(\boldsymbol{\omega}_{[0,T-1]}) = \begin{bmatrix} \omega(0) & \omega(1) & \ldots & \omega(T-L) \\ \omega(1) & \omega(2) & \ldots & \omega(T-L+1) \\ \vdots & \vdots & \ddots & \vdots \\ \omega(L-1) & \omega(L) & \ldots & \omega(T-1) \end{bmatrix}.$$

II. PROBLEM STATEMENT: THE MODEL-BASED CASE

In this section, we review safe output-feedback controller synthesis when the system model is known. We consider a

discrete-time linear system with output observations, whose state-space representation is given by

$$x(t + 1) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + v(t), \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is the state of the system and $x(0) = x_0$ for a predefined $x_0 \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ is the control input, $y(t) \in \mathbb{R}^p$ is the observed output, and $v(t) \in \mathbb{R}^p$ denotes measurement noise $v(t) \sim \mathcal{D}(0, \Sigma_v)$, with $\Sigma_v \succ 0$. The system is controlled through a time-varying, dynamic affine control policy

$$u(t) = \sum_{k=0}^{t} K_{t,k} y(k) + g_t + w(t), \quad (2)$$

where $K_{t,k}$ and g_t are the linear and affine parts of the policy, respectively, and $w(t) \in \mathbb{R}^m$ denotes noise on the input $w(t) \sim \mathcal{D}(0, \Sigma_w)$ with $\Sigma_w \succeq 0$, which acts as process noise.[1] Furthermore, we assume that the noise is bounded with

$$\|w\|_\infty \le w_\infty, \quad \|v\|_\infty \le v_\infty,$$

where $w_\infty, v_\infty > 0$. We consider the problem of synthesizing a feedback control policy that minimizes the expected value with respect to the disturbances of a quadratic objective defined over future input-output trajectories of length $N \in \mathbb{N}$:

$$J^2 := \mathbb{E}_{w,v} \left[\sum_{t=0}^{N-1} \left(y(t)^\mathsf{T} Q_t y(t) + u(t)^\mathsf{T} R_t u(t) \right) \right], \quad (3)$$

where $Q_t \succeq 0$ and $R_t \succ 0$ for every $t = 0, \ldots, N - 1$.

The problem is made more challenging by the requirement that inputs and outputs satisfy the safety constraints

$$\begin{bmatrix} y(t) \\ u(t) \end{bmatrix} \in \Gamma_t \subseteq \mathbb{R}^{p+m}, \quad \forall t = 0, \ldots, N - 1, \quad (4)$$

where Γ_t is a nonempty polytope for every $t = 0, \ldots, N - 1$ defined as

$$\Gamma_t = \left\{ (y, u) \in (\mathbb{R}^p, \mathbb{R}^m) | F_y^t y \le b_y^t, F_u^t u \le b_u^t \right\}, \quad (5)$$

with $F_y^t \in \mathbb{R}^{s \times p}$, $F_u^t \in \mathbb{R}^{s \times m}$ and $b_y^t, b_u^t \in \mathbb{R}^s$ for every $t = 0, \ldots, N - 1$. Despite (3) being convex in the input and output trajectories and Γ_t being polytopic, we highlight that minimizing (3) subject to (1), (2) and (4) is a non-convex problem in the control policy parameters $K_{t,k}$ and g_t. We refer the interested reader to [30]–[34] for classical and recent methods to overcome the non-convexity problem. For the rest of the paper, we assume that there exists a control input (2) that complies with (4) for all possible realizations of $w(t)$ and $v(t)$.

Remark 1: In this work, we analyze a finite-horizon control problem, which represents one iteration of a receding-horizon MPC implementation. It is therefore appropriate to compare the proposed approach with a *single* iteration of open-loop prediction approaches, such as the DeePC [14], [18]. The main difference is that we perform *closed-loop predictions*, i.e., we optimize over feedback policies $\pi(\cdot)$ such that $u(t) =$

[1]The more general model $x(t + 1) = Ax(t) + Bu(t) + w(t)$ would make the cost function depend on a specific realization A, B explicitly [29, Chapter 3]. Instead, the adopted noise model ensures that the cost only depends on the covariance matrix Σ_w and the coordinate-free parameter $\mathbf{G}$, thus making our theoretical bounds meaningful in a data-driven input-output setting.

$\pi(y(t), \ldots, y(0))$, while the DeePC [14], [18] performs *open-loop predictions*, i.e., it directly optimizes over input sequences $u(0), u(1), u(N-1)$. It is well-known that closed-loop predictions are less conservative. Indeed, by setting $K_{t,k} = 0$ in (2) the closed-loop policy reduces to an open-loop one. Most notably, closed-loop policies may preserve feasibility for significantly longer prediction horizons [35]. Naturally, the price to pay is an increased computational burden due to the larger dimensionality of the problem.

A. Convex design through the IOP

By leveraging tools offered by the framework of the Input-Output Parametrization[2] (IOP) [34], one can formulate a convex optimization problem that computes the optimal safe feedback control policy by searching over the input-output closed-loop responses. The state-space equations (1) provide the following relations between trajectories

$$\mathbf{x}_{[0,N-1]} = \mathbf{P}_A(:,0)x_0 + \mathbf{P}_B \mathbf{u}_{[0,N-1]}, \quad (6)$$

$$\mathbf{y}_{[0,N-1]} = \mathbf{C}\mathbf{x}_{[0,N-1]} + \mathbf{v}_{[0,N-1]}, \quad (7)$$

where $\mathbf{P}_A(:,0)$ denotes the first block-column of $\mathbf{P}_A$ and

$$\mathbf{P}_A = (I - \mathbf{ZA})^{-1}, \quad \mathbf{P}_B = (I - \mathbf{ZA})^{-1}\mathbf{ZB},$$
$$\mathbf{A} = I_N \otimes A, \quad \mathbf{B} = I_N \otimes B,$$
$$\mathbf{C} = I_N \otimes C, \quad \mathbf{Z} = \begin{bmatrix} 0_{n \times n(N-1)} & 0_{n \times n} \\ I_{n(N-1)} & 0_{n(N-1) \times n} \end{bmatrix}.$$

A few comments on the used notation are in order. First, the matrix $\mathbf{Z}$ is the block-downshift operator. Second, from now on we denote $\mathbf{G} = \mathbf{C}\mathbf{P}_B$ to highlight that $\mathbf{G}$ is a block-Toeplitz matrix containing the first N components of the impulse response of the plant $\mathbf{G}(z) = C(zI - A)^{-1}B$. Last, the matrix $\mathbf{C}\mathbf{P}_A(:,0)$ contains the entries of the observability matrix CA^i for $i = 0, \ldots, N - 1$. We denote the model-based free response of the system as $\mathbf{y}_0 = \mathbf{C}\mathbf{P}_A(:,0)x_0$. The control policy can be rewritten as:

$$\mathbf{u}_{[0,N-1]} = \mathbf{K}\mathbf{y}_{[0,N-1]} + \mathbf{g} + \mathbf{w}_{[0,N-1]}, \quad (8)$$

where $\mathbf{K}$ and $\mathbf{g}$ are defined as:

$$\mathbf{K} = \begin{bmatrix} K_{0,0} & 0_{m \times p} & \cdots & 0_{m \times p} \\ K_{1,0} & K_{1,1} & \ddots & 0_{m \times p} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N-1,0} & K_{N-1,1} & \cdots & K_{N-1,N-1} \end{bmatrix}, \mathbf{g} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{N-1} \end{bmatrix}. \quad (9)$$

The safety constraints (4)-(5) take the form

$$\max_{\|\mathbf{v}\|_\infty \le v_\infty, \|\mathbf{w}\|_\infty \le w_\infty} \mathbf{F}_y \mathbf{y} \le \mathbf{b}_y, \quad \max_{\|\mathbf{v}\|_\infty \le v_\infty, \|\mathbf{w}\|_\infty \le w_\infty} \mathbf{F}_u \mathbf{u} \le \mathbf{b}_u, \quad (10)$$

with $\mathbf{F}_y = \text{blkdiag}(F_y^0, \ldots, F_y^{N-1})$, $\mathbf{b}_y = \text{vec}(b_y^0, \ldots, b_y^{N-1})$, $\mathbf{F}_u = \text{blkdiag}(F_u^0, \ldots, F_u^{N-1})$, $\mathbf{b}_u = \text{vec}(b_u^0, \ldots, b_u^{N-1})$, and $\max(\cdot)$ to be intended row-wise. By plugging the controller (8) into (6)-(7), it is easy to derive the relationships

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{v} + \mathbf{y}_0 \\ \mathbf{w} \end{bmatrix} + \begin{bmatrix} \mathbf{Gq} \\ \mathbf{q} \end{bmatrix}, \quad (11)$$

[2]Similar to [32], [33], the IOP [34] yields a convex representation of input-output closed-loop responses. It is also numerically stable for the case of infinite-horizon stable plants and for finite-horizon control problems [36].

where

$$\mathbf{\Phi} = \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} = \begin{bmatrix} (I - \mathbf{GK})^{-1} & (I - \mathbf{GK})^{-1}\mathbf{G} \\ \mathbf{K}(I - \mathbf{GK})^{-1} & (I - \mathbf{KG})^{-1} \end{bmatrix}, \quad (12)$$

and $\mathbf{q} = (I - \mathbf{KG})^{-1}\mathbf{g} = \mathbf{\Phi}_{uu}\mathbf{g}$. The parameters $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$, where $\mathbf{\Phi}_{yy} \in \mathbb{R}^{Np \times Np}, \mathbf{\Phi}_{yu} \in \mathbb{R}^{Np \times Nm}, \mathbf{\Phi}_{uy} \in \mathbb{R}^{Nm \times Np}$ and $\mathbf{\Phi}_{uu} \in \mathbb{R}^{Nm \times Nm}$, represent the four closed-loop responses defining the relationship between disturbances and input-output signals, while $\mathbf{q} \in \mathbb{R}^{Nm}$ represents the affine part of the disturbance-feedback control policy [32], [37]. To achieve a convex reformulation of the control problem under consideration, it is not hard to extend the IOP from [34] to account for the safety constraints (10) in a convex way. The result is summarized in the next proposition, whose proof is reported in Appendix B for completeness.

Proposition 1: Consider the LTI system (1) evolving under the control policy (8) within a horizon of length $N \in \mathbb{N}$. Then:

i) For any control policy $(\mathbf{K}, \mathbf{g})$ that complies with the safety constraints, there exist four matrices $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$ and a vector $\mathbf{q}$ such that $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$, $\mathbf{g} = \mathbf{\Phi}_{uu}^{-1}\mathbf{q}$, and for all $j = 1, \ldots, sN$,

$$\begin{bmatrix} I & -\mathbf{G} \end{bmatrix} \mathbf{\Phi} = \begin{bmatrix} I & 0 \end{bmatrix}, \quad \mathbf{\Phi} \begin{bmatrix} -\mathbf{G} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad (13)$$

$$\left\| \begin{bmatrix} v_\infty(F_{y,j}\mathbf{\Phi}_{yy})^\mathsf{T} \\ w_\infty(F_{y,j}\mathbf{\Phi}_{yu})^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1^\star + F_{y,j}(\mathbf{Gq} + \mathbf{\Phi}_{yy}\mathbf{y}_0) \leq \mathbf{b}_{y,j}, \quad (14)$$

$$\left\| \begin{bmatrix} v_\infty(F_{u,j}\mathbf{\Phi}_{uy})^\mathsf{T} \\ w_\infty(F_{u,j}\mathbf{\Phi}_{uu})^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1^\star + F_{u,j}(\mathbf{q} + \mathbf{\Phi}_{uy}\mathbf{y}_0) \leq \mathbf{b}_{u,j}, \quad (15)$$

$\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu}$ with causal sparsities [3], $\quad (16)$

where $F_{y,j} \in \mathbb{R}^{1 \times Np}$, $F_{u,j} \in \mathbb{R}^{1 \times Nm}$ and $\mathbf{b}_{u,j}, \mathbf{b}_{y,j} \in \mathbb{R}$ are the j-th row of $\mathbf{F}_y$, $\mathbf{F}_u$ and $\mathbf{b}_u, \mathbf{b}_y$, respectively.

ii) For any four matrices $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$ complying with (13)-(16) and any vector $\mathbf{q} \in \mathbb{R}^{mN}$, the matrix $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ is causal as per (9) and it yields the closed-loop responses $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$. Moreover, the affine policy $(\mathbf{K}, \mathbf{g})$ with $\mathbf{g} = \mathbf{\Phi}_{uu}^{-1}\mathbf{q}$ complies with the safety constraints.

We remark that the IOP is well-suited to a data-driven output-feedback setup, as all affine control policies are directly parametrized through the impulse response parameters $\mathbf{G}$, without requiring an internal state-space representation. This is useful for two reasons. First, when dealing with unknown systems, the state-space parameters (A, B, C, x_0) can only be estimated up to an unknown change of variable, which may be problematic for defining the cost and the noise statistics [38]. Second, several large-scale systems feature a very large number of states, but a comparably small number of inputs and outputs, that is $n \gg \max(m, p)$. In such applications, it is advantageous to bypass a state-space representation and directly deal with $\mathbf{G}$, whose dimensions do not depend on n.

From now on, to simplify the expressions appearing throughout the next sections and without any loss of gen-

erality[4], we let $\mathbf{q} = \mathbf{g} = 0_{Nm \times 1}$, that is, we focus on *linear* control policies. We are ready to establish a convex formulation of the optimal control problem under study.

Proposition 2: Consider the LTI system (1). The linear control policy that achieves the minimum of the cost functional (3) is given by $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$, where $\mathbf{\Phi}_{uy}, \mathbf{\Phi}_{yy}$ are optimal solutions to the following convex optimization problem:

$$\min_{\mathbf{\Phi}} \left\| \begin{bmatrix} \mathbf{Q}^{\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_v^{\frac{1}{2}} & 0 & \mathbf{y}_0 \\ 0 & \mathbf{\Sigma}_w^{\frac{1}{2}} & 0 \end{bmatrix} \right\|_F^2 \quad (17)$$

subject to $(13) - (16)$,

where $\mathbf{Q} = \text{blkdiag}(Q_0, \ldots, Q_{N-1})$, $\mathbf{R} = \text{blkdiag}(R_0, \ldots, R_{N-1})$, $\mathbf{\Sigma}_v = I_N \otimes \Sigma_v$, $\mathbf{\Sigma}_w = I_N \otimes \Sigma_w$ and where (14)-(15) are evaluated at $\mathbf{q} = 0$.

Proof: We refer to Proposition 2 of [25] for a complete derivation of the cost function. To conclude the proof, it suffices to notice that the objective function and the safety constraints (14)-(15) are convex in $\mathbf{\Phi}$. ∎

When the system parameters (A, B, C, x_0) are known, a globally optimal solution $(\mathbf{\Phi}_{yy}^\star, \mathbf{\Phi}_{yu}^\star, \mathbf{\Phi}_{uy}^\star, \mathbf{\Phi}_{uu}^\star)$ for problem (17) can be efficiently computed with off-the-shelf solvers. The corresponding globally optimal and safe control policy is then recovered as $\mathbf{K}^\star = \mathbf{\Phi}_{uy}^\star(\mathbf{\Phi}_{yy}^\star)^{-1}$.

The rest of the paper contains our main contributions. Specifically, we address the following two questions:

Q1) How can we compute a safe control policy with performance close to that of $\mathbf{K}^\star$, solely based on libraries of *noisy* input-output trajectories?

Q2) How steeply does the suboptimality grow with respect to $\mathbf{K}^\star$ as the noise increases?

III. THE DATA-DRIVEN CASE: ROBUSTLY SAFE CONTROLLER SYNTHESIS FROM NOISY DATA

We answer question Q1) by developing a method to synthesize near-optimal safe controllers from noisy data. The main result of this section is an optimization problem based on the IOP that tightly approximates the optimal and safe control policy, despite the fact that the noise-corrupted data only yield approximate estimates of the system impulse and free response. We conclude by offering novel insights on its properties and its numerical implementation based on convex optimization.

A. From noise-corrupted data to doubly-robust optimal control

From now on, the dynamics matrices (A, B, C) and the initial state x_0 are *unknown*. Instead, only the following data are available:

D1 A noisy system trajectory $\{y^h(t), u^h(t)\}_{t=-T}^{-1}$ recorded offline during an experiment.

---

[3]Specifically, they have the block lower-triangular sparsities resulting as per the expressions (12), the sparsity of $\mathbf{K}$ in (9) and that of $\mathbf{G}$.

[4]One can redefine $\tilde{\mathbf{y}} = \begin{bmatrix} 1 & \mathbf{y}^\mathsf{T} \end{bmatrix}^\mathsf{T}$, $\overline{\mathbf{v}} = \begin{bmatrix} 1 & \mathbf{v}^\mathsf{T} \end{bmatrix}^\mathsf{T}$, $\overline{\mathbf{C}} = \begin{bmatrix} 0_{1 \times Nn} \\ \mathbf{C} \end{bmatrix}$, $\overline{\mathbf{K}} = \begin{bmatrix} \mathbf{g} & \mathbf{K} \end{bmatrix}$ and $\overline{\mathbf{\Phi}}$ as per (12) with $\overline{\mathbf{G}}$ and $\overline{\mathbf{K}}$ in place of $\mathbf{G}$ and $\mathbf{K}$, respectively. Minor modifications to (14)-(15) are needed as well.

D2 The cost matrices Q_t, R_t, the matrices Σ_v, Σ_w, the safety sets Γ_t, and the bounded sets $\mathcal{W} = \{\mathbf{w} | \; \|\mathbf{w}\|_\infty \leq w_\infty\}$ and $\mathcal{V} = \{\mathbf{v} | \; \|\mathbf{v}\|_\infty \leq v_\infty\}$ where disturbances live.

Our approach exploits the noisy data in **D1** to compute approximate system responses $\widehat{\mathbf{G}}$ and $\widehat{\mathbf{y}}_0$ in a preliminary identification step. We work under the following assumption.

Assumption 1: Let $\mathbf{\Delta} = \mathbf{G} - \widehat{\mathbf{G}}$ and $\boldsymbol{\delta}_0 = \mathbf{y}_0 - \widehat{\mathbf{y}}_0$. There exist $\epsilon_{2,G}, \epsilon_{\infty,G}, \epsilon_{2,y}, \epsilon_{\infty,y} > 0$ such that,

$$\|\mathbf{\Delta}\|_2 \leq \epsilon_{2,G}, \qquad \|\boldsymbol{\delta}_0\|_2 \leq \epsilon_{2,y},$$
$$\|\mathbf{\Delta}\|_\infty \leq \epsilon_{\infty,G}, \quad \|\boldsymbol{\delta}_0\|_\infty \leq \epsilon_{\infty,y}.$$

Note that, in practice, a meaningful bound on $\boldsymbol{\delta}_0$ is only available if A is stable or the time-horizon is sufficiently short. Let us define $\epsilon_2 = \max(\epsilon_{2,G}, \epsilon_{2,y})$ and $\epsilon_\infty = \max(\epsilon_{\infty,G}, \epsilon_{\infty,y})$. Assumption 1 can be fulfilled using different methods over the available data **D1**; for instance, one may utilize standard least-squares identification that comes with probabilistic and non-asymptotic error bounds [26], [39], or more sophisticated stochastic estimators based on behavioral theory such as maximum-likelihood predictors [21], which also come with quantifiable error bounds [40]. Our results are independent of the choice of the identification scheme. A discussion as to how recent behavioral approaches can be used for identification is reported in Appendix A. These estimators will be used in the numerical examples in Section V.

After condensing the effect of noise-corrupted data into model mismatch parameters $\mathbf{\Delta}, \boldsymbol{\delta}_0$, we formulate a *doubly-robust* control problem, that is, a problem where we enforce constraint satisfaction for 1) all possible model mismatches $(\mathbf{\Delta}, \boldsymbol{\delta}_0)$, and 2) all possible disturbances sequences $\mathbf{w} \in \mathcal{W}$ and $\mathbf{v} \in \mathcal{V}$. In particular, define $\boldsymbol{\theta} = (\mathbf{\Delta}, \boldsymbol{\delta}_0, \mathbf{w}, \mathbf{v})$ and let

$$\mathbf{y}(\mathbf{K}, \boldsymbol{\theta}) = \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0 + (\widehat{\mathbf{G}} + \mathbf{\Delta})\mathbf{u}(\mathbf{K}, \boldsymbol{\theta}) + \mathbf{v},$$
$$\mathbf{u}(\mathbf{K}, \boldsymbol{\theta}) = \mathbf{K}\mathbf{y}(\mathbf{K}, \boldsymbol{\theta}) + \mathbf{w},$$

be the closed-loop trajectories associated with a specific controller $\mathbf{K}$ and disturbance and mismatch realizations $\boldsymbol{\theta}$. Further, define the set of doubly-robust controllers as:

$$\mathcal{K} = \{\mathbf{K} \text{ in } (9)| \; (\mathbf{y}(\mathbf{K}, \boldsymbol{\theta}), \mathbf{u}(\mathbf{K}, \boldsymbol{\theta})) \in \Gamma, \; \forall \boldsymbol{\theta} \in \mathcal{E} \times \mathcal{W} \times \mathcal{V}\},$$

with $\mathcal{E} = \{(\mathbf{\Delta}, \boldsymbol{\delta}_0) | \|\mathbf{\Delta}\|_p \leq \epsilon_p, \; \|\boldsymbol{\delta}_0\|_p \leq \epsilon_p, \forall p \in \{2, \infty\}\}$, $\Gamma = \Gamma_0 \times \Gamma_1 \times \cdots \times \Gamma_{N-1}$, and assume that $\mathcal{K}$ is not empty. Then, the doubly-robust problem of interest takes the form

$$\min_{\mathbf{K} \in \mathcal{K}} \max_{(\mathbf{\Delta}, \boldsymbol{\delta}_0) \in \mathcal{E}} \sqrt{\mathbb{E}_{\mathbf{w}, \mathbf{v}} \left[\mathbf{y}(\mathbf{K}, \boldsymbol{\theta})^\mathsf{T} \mathbf{y}(\mathbf{K}, \boldsymbol{\theta}) + \mathbf{u}(\mathbf{K}, \boldsymbol{\theta})^\mathsf{T} \mathbf{u}(\mathbf{K}, \boldsymbol{\theta}) \right]},$$
$$(18)$$

where we have selected the weights $\mathbf{Q}, \mathbf{R}, \Sigma_w, \Sigma_v$ to be identity matrices with appropriate dimensions. The same assumption is used in the rest of the paper, in order to facilitate the derivations. However, we note that all our results can be easily adapted to non-identity weights. Next, we observe that the doubly-robust optimization problem admits an equivalent formulation in terms of the closed-loop response parameters.

Proposition 3: Letting $\mathbf{\Phi}_{yy} = \widehat{\mathbf{\Phi}}_{yy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}$, $\mathbf{\Phi}_{yu} = \mathbf{\Phi}_{yy}(\widehat{\mathbf{G}} + \mathbf{\Delta})$, $\mathbf{\Phi}_{uy} = \widehat{\mathbf{\Phi}}_{uy}(I - \mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}$, $\mathbf{\Phi}_{uu} = (I - \widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu}$, the optimization problem (18) is equivalent to

$$\min_{\widehat{\mathbf{\Phi}} \in \mathbf{\Pi}} \max_{(\mathbf{\Delta}, \boldsymbol{\delta_0}) \in \mathcal{E}} \left\| \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} I & 0 & \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0 \\ 0 & I & 0 \end{bmatrix} \right\|_F, \quad (19)$$

where the set of doubly-robust closed-loop responses $\mathbf{\Pi}$ is

$$\mathbf{\Pi} = \{\widehat{\mathbf{\Phi}} | \; (20) - (23), \; \forall j = 1, \ldots, sN, \; \forall (\mathbf{\Delta}, \boldsymbol{\delta}_0) \in \mathcal{E}\},$$

with

$$\begin{bmatrix} I & -\widehat{\mathbf{G}} \end{bmatrix} \widehat{\mathbf{\Phi}} = \begin{bmatrix} I & 0 \end{bmatrix}, \quad \widehat{\mathbf{\Phi}} \begin{bmatrix} -\widehat{\mathbf{G}} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad (20)$$

$$\left\| \begin{bmatrix} v_\infty (F_{y,j}\mathbf{\Phi}_{yy})^\mathsf{T} \\ w_\infty (F_{y,j}\mathbf{\Phi}_{yu})^\mathsf{T} \end{bmatrix} \right\|_1 + (F_{y,j}\mathbf{\Phi}_{yy})(\widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0) \leq \mathbf{b}_{y,j}, \quad (21)$$

$$\left\| \begin{bmatrix} v_\infty (F_{u,j}\mathbf{\Phi}_{uy})^\mathsf{T} \\ w_\infty (F_{u,j}\mathbf{\Phi}_{uu})^\mathsf{T} \end{bmatrix} \right\|_1 + (F_{u,j}\mathbf{\Phi}_{uy})(\widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0) \leq \mathbf{b}_{u,j}, \quad (22)$$

$$\widehat{\mathbf{\Phi}}_{yy}, \widehat{\mathbf{\Phi}}_{yu}, \widehat{\mathbf{\Phi}}_{uy}, \widehat{\mathbf{\Phi}}_{uu} \text{ with causal sparsities.} \quad (23)$$

The proof of Proposition 3 can be found in Appendix C. We remark that the closed-loop responses $\mathbf{\Phi}$ appearing in (19), (21) and (22) are associated with the *true* impulse response, whereas the closed-loop responses $\widehat{\mathbf{\Phi}}$ appearing in (20) and (23) are associated with the *estimated* impulse response. This is because, while we are interested in minimizing the cost and satisfying the safety constraints for the *real* system, we can only parametrize the closed-loop responses for the identified system.

The robust optimization problem (19) is non-convex in the cost and in the constraints because $\mathbf{\Phi}$ is a nonlinear function of the matrix variables $\widehat{\mathbf{\Phi}}$ and $\mathbf{\Delta}$. Therefore, it is challenging to find a feasible solution, let alone the optimal one. We note that, for the case of open-loop control policies, one may use constraint-tightening approaches such as those of [41], [42]. In this work, we propose an analysis that compares feedback control policies. Specifically, we derive suboptimality guarantees with respect to the optimal model-based linear feedback policy as a function of the model mismatch level.

B. Proposed relaxation for safe controller synthesis

Our first main result is to derive a relaxation of the intractable problem (19) that we can solve in practice. Our proposed approach is to 1) upper bound the cost function, and 2) tighten the safety constraints with more tractable expressions. In Section IV we will explicitly quantify the suboptimality incurred by these approximations. At its core, this methodology is inspired by that developed in [10] for the state-feedback case without measurement noise. However, the addition of output-feedback and measurement noise leads to new terms both in the cost and the safety constraints that are more challenging to analyze.

The following two lemmas establish the basis for our relaxation. Let $J(\mathbf{G}, \mathbf{K}) = \sqrt{\mathbb{E}_{\mathbf{w}, \mathbf{v}}[\mathbf{y}^\mathsf{T}\mathbf{y} + \mathbf{u}^\mathsf{T}\mathbf{u}]}$ denote the square root of the cost in (3). Lemma 1 provides the new expression which upper bounds $J(\mathbf{G}, \mathbf{K})$ and Lemma 2 provides a tightened form of the safety constraints. Their rather lengthy technical proof is reported in the Appendices D and E, respectively.

Lemma 1: Let $\widehat{\mathbf{\Phi}}$ denote the closed-loop responses obtained by applying $\mathbf{K}$ to $\widehat{\mathbf{G}}$. Further assume that $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \leq \gamma$, where $\gamma \in [0, \epsilon_2^{-1})$. Then, we have

$$J(\mathbf{G}, \mathbf{K}) \leq \frac{J_{UB}}{1 - \epsilon_2 \gamma} \quad (24)$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

6

where

$$J_{UB} = \left\| \begin{bmatrix} \sqrt{1+h(\epsilon_2,\gamma,\widehat{\mathbf{G}})+h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)}\widehat{\boldsymbol{\Phi}}_{yy} & \widehat{\boldsymbol{\Phi}}_{yu} & \widehat{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1+h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)}\widehat{\boldsymbol{\Phi}}_{uy} & \widehat{\boldsymbol{\Phi}}_{uu} & \widehat{\boldsymbol{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F,$$

and $h(\epsilon,\gamma,\mathbf{Y}) = \epsilon^2(2+\gamma\|\mathbf{Y}\|_2)^2 + 2\epsilon\|\mathbf{Y}\|_2(2+\gamma\|\mathbf{Y}\|_2)$.

Lemma 1 exploits the upper bound $\left\|\widehat{\boldsymbol{\Phi}}_{uy}\right\|_2 \leq \gamma$ to establish an explicit relationship between $J(\mathbf{G},\mathbf{K})$, the cost obtained by applying a controller $\mathbf{K}$ to the real system $\mathbf{G}$, and $J(\widehat{\mathbf{G}},\mathbf{K})$, the cost obtained by applying the same controller to the estimated system $\widehat{\mathbf{G}}$. To see this, notice that (24) can be equivalently rewritten as

$$J(\mathbf{G},\mathbf{K}) \leq \frac{\left(J(\widehat{\mathbf{G}},\mathbf{K})^2 + \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F^2(h(\epsilon_2,\gamma,\widehat{\mathbf{G}})+\right.}{1-\epsilon_2\gamma}$$
$$\frac{\left.+h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)) + \|\widehat{\boldsymbol{\Phi}}_{uy}\|_F^2 h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)\right)^{\frac{1}{2}}}{1-\epsilon_2\gamma}. \quad (25)$$

The expression (25) upper bounds the gap between $J(\mathbf{G},\mathbf{K})$ and $J(\widehat{\mathbf{G}},\mathbf{K})$ as a quantity that increases with ϵ_2 and with the norm of $\widehat{\mathbf{G}},\widehat{\mathbf{y}}_0,\widehat{\boldsymbol{\Phi}}$. We note that a similar result has appeared in [5, Proposition 3.2]. However, Lemma 1 additionally takes into account how an uncertain $\hat{x}(0)$ affects the cost through the free response $\widehat{\mathbf{y}}_0$. We now derive a tightened - yet more tractable - expression for the safety constraints (21)-(22).

Lemma 2: Assume $\left\|\widehat{\boldsymbol{\Phi}}_{uy}\right\|_\infty \leq \tau$, where $\tau \in [0,\epsilon_\infty^{-1})$. Then, if for all $j = 1,\dots,sN$ the closed-loop responses $\widehat{\boldsymbol{\Phi}}$ satisfy the tightened safety constraints

$$f_{1,j}(\widehat{\boldsymbol{\Phi}}) + f_{2,j}(\widehat{\boldsymbol{\Phi}}) + f_{3,j}(\widehat{\boldsymbol{\Phi}}) \leq \mathbf{b}_{y,j}, \quad (26)$$
$$f_{4,j}(\widehat{\boldsymbol{\Phi}}) + f_{5,j}(\widehat{\boldsymbol{\Phi}}) + f_{6,j}(\widehat{\boldsymbol{\Phi}}) \leq \mathbf{b}_{u,j}, \quad (27)$$

where

$$f_{1,j}(\widehat{\boldsymbol{\Phi}}) = \frac{v_\infty\left\|F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}\right\|_1^\star}{1-\epsilon_\infty\tau}, \quad f_{4,j}(\widehat{\boldsymbol{\Phi}}) = \frac{v_\infty\left\|F_{u,j}\widehat{\boldsymbol{\Phi}}_{uy}\right\|_1^\star}{1-\epsilon_\infty\tau},$$

$$f_{2,j}(\widehat{\boldsymbol{\Phi}}) = w_\infty\left\| \begin{bmatrix} \left(F_{y,j}\widehat{\boldsymbol{\Phi}}_{yu}\right)^\top \\ \epsilon_\infty\frac{1+\tau\|\widehat{\mathbf{G}}\|_\infty}{1-\epsilon_\infty\tau}\left(F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}\right)^\top \end{bmatrix} \right\|_1,$$

$$f_{5,j}(\widehat{\boldsymbol{\Phi}}) = w_\infty\left\| \begin{bmatrix} \left(F_{u,j}\widehat{\boldsymbol{\Phi}}_{uu}\right)^\top \\ \epsilon_\infty\frac{1+\tau\|\widehat{\mathbf{G}}\|_\infty}{1-\epsilon_\infty\tau}\left(F_{u,j}\widehat{\boldsymbol{\Phi}}_{uy}\right)^\top \end{bmatrix} \right\|_1,$$

$$f_{3,j}(\widehat{\boldsymbol{\Phi}}) = F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{y}}_0 + \epsilon_\infty\left\|F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}\right\|_1^\star\left(\frac{1+\tau\|\widehat{\mathbf{y}}_0\|_\infty}{1-\epsilon_\infty\tau}\right),$$

$$f_{6,j}(\widehat{\boldsymbol{\Phi}}) = F_{u,j}\widehat{\boldsymbol{\Phi}}_{uy}\widehat{\mathbf{y}}_0 + \epsilon_\infty\left\|F_{u,j}\widehat{\boldsymbol{\Phi}}_{uy}\right\|_1^\star\left(\frac{1+\tau\|\widehat{\mathbf{y}}_0\|_\infty}{1-\epsilon_\infty\tau}\right),$$

then $\widehat{\boldsymbol{\Phi}}$ satisfies the safety constraints (21)-(22) for all $(\boldsymbol{\Delta},\boldsymbol{\delta}_0) \in \mathcal{E}$.

Lemma 2 exploits the upper bound $\left\|\widehat{\boldsymbol{\Phi}}_{uy}\right\|_\infty \leq \tau$ to quantify the worst-case effect of the disturbances in increasing the values of the inputs and the outputs. In our setup, similar to [10], the feasible set shrinks in the presence of larger impulse and free response estimation error ϵ_∞. This is because (26)-(27) are more restrictive, and will eventually become infeasible for sufficiently large ϵ_∞. Instead, the effect of increasing

the value of τ is less intuitive. Indeed, as τ increases, the constraint $\left\|\widehat{\boldsymbol{\Phi}}_{uy}\right\|_\infty \leq \tau$ softens while (26)-(27) tighten. It is therefore necessary to explicitly optimize over τ. We are now ready to establish a relaxation of problem (19).

Theorem 1: Consider the following optimization problem:

$$\min_{\gamma\in[0,\epsilon_2^{-1}),\tau\in[0,\epsilon_\infty^{-1})} \frac{1}{1-\epsilon_2\gamma}\min_{\widehat{\boldsymbol{\Phi}}} \quad J_{UB} \quad (28)$$

subject to (20), (23),
$$\left\|\widehat{\boldsymbol{\Phi}}_{uy}\right\|_2 \leq \gamma, \quad \left\|\widehat{\boldsymbol{\Phi}}_{uy}\right\|_\infty \leq \tau, \quad (29)$$
$$(26)-(27), \quad \forall j = 1,\dots,sN,$$

where J_{UB} is defined in Lemma 1. Then, (28) has the following properties:

i) upon fixing any specific values for $\gamma \in [0,\epsilon_2^{-1})$ and $\tau \in [0,\epsilon_\infty^{-1})$, the optimization problem is convex in $\widehat{\boldsymbol{\Phi}}$,

ii) all of its feasible solutions yield a controller $\widehat{\mathbf{K}} = \widehat{\boldsymbol{\Phi}}_{uy}\widehat{\boldsymbol{\Phi}}_{yy}^{-1}$ complying with the safety constraints (14)-(15) for the real system,

iii) its minimal cost upper bounds that of (18).

Proof: Lemma 1 shows that the cost of (28) upper bounds $J(\mathbf{G},\mathbf{K}) = J(\mathbf{G},\widehat{\boldsymbol{\Phi}}_{uy}\widehat{\boldsymbol{\Phi}}_{yy}^{-1})$ for every feasible $\mathbf{K}$. Lemma 2 shows that (26)-(27) imply the doubly-robust constraints (21)-(22) for all $(\boldsymbol{\Delta},\boldsymbol{\delta}_0) \in \mathcal{E}$. Hence, $\widehat{\mathbf{K}} = \widehat{\boldsymbol{\Phi}}_{uy}\widehat{\boldsymbol{\Phi}}_{yy}^{-1}$ complies with safety constraints (14)-(15) for the real system. When γ and τ are fixed, it remains to optimize over $\widehat{\boldsymbol{\Phi}}$. The cost function is convex in $\widehat{\boldsymbol{\Phi}}$ and so are the constraints of the inner optimization problem. ∎

Theorem 1 shows that problem (19), which is non-convex in its matrix variables, can be approximated as the problem of solving a convex optimization problem[5] for each choice of the scalar variables γ and τ. The ϵ-dependent suboptimality introduced by such an approximation will be quantified in the next section. The global optimum of (28) is thus determined by exhaustive search over the box $(\gamma,\tau) \in [0,\epsilon_2^{-1}) \times [0,\epsilon_\infty^{-1})$, for instance through gridding, random search [43] or bisection [44]. Gridding over (τ,γ) and solving a convex optimization problem each time may significantly increase the computational burden if we are interested in determining a near-optimal solution with very low tolerance. Similarly to [5], in the next proposition we show that the inner cost function in problem (28) can be made independent of γ by introducing a parameter $\alpha \in \mathbb{R}$ that acts as an upper bound to γ. As a result, the overall cost becomes quasiconvex[6] in γ, and the globally optimal $\gamma^\star(\tau)$ for each fixed τ can be found efficiently through golden-section search [46].

Proposition 4: Fix $\alpha \in [0,\epsilon_2^{-1})$ and consider the following optimization problem

$$\min_{\gamma\in[0,\alpha],\tau\in[0,\epsilon_\infty^{-1})} \frac{1}{1-\epsilon_2\gamma}\min_{\widehat{\boldsymbol{\Phi}}} \quad J_{UB}^\alpha(\widehat{\boldsymbol{\Phi}}) \quad (30)$$

subject to (20), (23), (26), (27), (29)

---

[5]Specifically, a semidefinite program (SDP) due to the presence of quadratic $\|\cdot\|_2$ constraints.

[6]A function $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if and only if $f(\theta x_1 + (1-\theta)x_2) \leq \max(f(x_1),f(x_2))$ for every $x_1,x_2 \in \mathbb{R}^n$ and every $\theta \in [0,1]$. We refer to [45] for a comprehensive discussion.

TABLE I: Convexity properties for the proposed reformulations.

	QC in γ	QC in (τ,γ)	C for fixed (γ,τ)
(28)	X	X	✓
(30)	✓	X	✓

$$\forall j = 1, \ldots, sN,$$

where $J_{UB}^{\alpha}(\widehat{\mathbf{\Phi}})$ is defined as

$$\left\| \begin{bmatrix} \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\,\widehat{\mathbf{\Phi}}_{yy} & \widehat{\mathbf{\Phi}}_{yu} & \widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\,\widehat{\mathbf{\Phi}}_{uy} & \widehat{\mathbf{\Phi}}_{uu} & \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F.$$

Then, the statements $i)$, $ii)$ and $iii)$ of Theorem 1 hold. Furthermore

$iv)$ The cost function of problem (30) is quasiconvex in γ.

Proof: Since $\gamma \leq \alpha$, $\alpha < \epsilon_2^{-1}$, and $h(\epsilon,\gamma,\cdot)$ is a monotonically increasing function of γ, then the inequality (24) in Lemma 1 continues to hold when putting α in place of γ inside the $h(\cdot)$ functions. The constraints of (28) are unaffected. Hence, $i)$, $ii)$ and $iii)$ of Theorem 1 continue to hold. It remains to prove $iv)$. Let us fix any value for τ. First, notice that $J_{UB}^{\alpha}(\widehat{\mathbf{\Phi}})$ is a convex function of $\widehat{\mathbf{\Phi}}$ and does not depend on γ, and that the feasible set of the inner minimization in problem (30) is convex. Denote as $g(\gamma)$ the optimal value of the inner optimization problem. We are left with minimizing the functional $\frac{g(\gamma)}{1-\epsilon_2\gamma}$ over γ. We know that $g(\gamma)$ is convex in γ because it is obtained as the partial minimization of a convex functional over a convex set [47], and that $(1-\epsilon_2\gamma)$ is concave in γ. Since the ratio of a non-negative convex function and a positive concave function is quasiconvex, we conclude that the cost of problem (30) is quasiconvex in γ. ∎

In [5], the idea of using the parameter α was relying on a lemma from [48]. Here, we have derived an alternative self-contained proof that holds also for the case $x_0 \neq 0$. In summary, for a fixed $\alpha < \epsilon_2^{-1}$, for every τ gridding the interval $[0, \epsilon_\infty^{-1}]$ and for γ chosen according to golden-search, we solve the corresponding instance of the inner optimization problem in (30), which is convex in $\widehat{\mathbf{\Phi}}$. We also note that an infinite-horizon version of problem (30) can be established by adding a tail variable and adopting a finite-horizon approximation of stable transfer functions similar to [10].

Last, one may wonder whether the cost function of problem (30) is jointly quasiconvex in γ and τ, as conjectured in [10]. Here, we clarify that this may not be the case, even for the state-feedback framework of [10]. For instance, similar to the constraints (26)-(27) and those of [10], consider the function $s : \mathbb{R}^2 \to \mathbb{R}$ defined as $s(x,y) = \frac{|y|x}{(1-x)}$. Fixing $x_1 = 1, x_2 = 0, y_1 = -\frac{1}{2}, y_2 = \frac{1}{2}$, one can verify that

$$s(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) = |0.5 - \theta|\frac{\theta}{1-\theta},$$

is not quasiconvex for $\theta \in [0,1]$. Based on this reasoning, the cost of problem (30), and similarly the objective (2.3) in [10], may not be quasiconvex in τ. Hence, exhaustive search over τ remains the only solution in general. Table I summarizes the convexity properties of (28) and (30).

C. Safe exploration

In many applications, it is desirable not only that the control policy synthesized from data is safe, but also that the system operates safely during the data-collection phase. In our setup, this amounts to requiring that the available trajectories $\mathbf{D1}$ verify (4). Similar to [10], we now clarify that safe data can be collected exploiting the feasible space of the optimization problem (28) as a corollary of Theorem 1.

More in details, assume that rough estimates $\widehat{\mathbf{G}}_r, \widehat{\mathbf{y}}_{0,r}$ are given with possibly large errors $\epsilon_{\infty,r}$ and $\epsilon_{2,r}$. Note that it is inherently impossible to guarantee safe exploration unless some prior information is available. Consider now an exploration signal $\eta(t)$ such that $\|\eta(t)\|_\infty \leq \eta_\infty$ for every $t \in \mathbb{Z}$ and define $\widetilde{w}_\infty = w_\infty + \eta_\infty$. Let $\widehat{\mathbf{K}}_r = \widehat{\mathbf{\Phi}}_{uy,r}\widehat{\mathbf{\Phi}}_{yy,r}^{-1}$ be any feasible solution to the instance of the optimization problem (28) where we use $\widetilde{w}_\infty$ in place of w_∞. Then, by Theorem 1, the control policy

$$\mathbf{u} = \mathbf{K}_r\mathbf{y} + \boldsymbol{\eta} + \mathbf{w},$$

can be applied to the real system during the exploration phase to generate safe trajectories.

IV. SUBOPTIMALITY ANALYSIS

In this section, we tackle question Q2) in Section II about performance degradation as a function of the level of model-mismatch due to noisy data. We denote as $\mathbf{K}^\star, \mathbf{\Phi}^\star$ the optimal controller for the real constrained problem (17) and corresponding closed-loop responses. Similarly, we denote as $\widehat{\mathbf{K}}^\star, \widehat{\mathbf{\Phi}}^\star$ the optimal controller for the optimization problem (30) and corresponding closed-loop responses. Further, we let $J^\star = J(\mathbf{G}, \mathbf{K}^\star)$ and $\hat{J} = J(\mathbf{G}, \widehat{\mathbf{K}}^\star)$. We aim to characterize the relative suboptimality gap $\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}}$, and specifically we will show that

$$\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}} \leq \mathcal{O}(\epsilon_2) + \tilde{S}(\epsilon_\infty, \epsilon_2),$$

where $\tilde{S}(\epsilon_\infty, \epsilon_2) = S(\epsilon_\infty)(1+\mathcal{O}(\epsilon_2))$. Here, $S(\epsilon_\infty)$ quantifies the suboptimality incurred by tightening the constraints and is such that $S(0) = 0$. We prove that if ϵ_2 and ϵ_∞ are small enough and the optimal controller $\mathbf{K}^\star$ does not activate the safety constraints, then $S(\epsilon_\infty) = 0$ and the suboptimality shrinks to 0 linearly fast as ϵ_2 converges to 0. Otherwise, the gap may decrease according to $S(\epsilon_\infty)$, for which we provide a numerical plot in Section V. In other words, for small estimation errors ϵ_2 and ϵ_∞, applying controller $\widehat{\mathbf{K}}^\star$ (which is solely computed from noisy data) to the *real* plant achieves almost optimal closed-loop performance while guaranteeing compliance with safety constraints. Surprisingly, despite the additional complexity of output-feedback and output noise, our bound matches the scaling with respect to $\epsilon = \max(\epsilon_2, \epsilon_\infty)$ that has been derived in [10] for the state-feedback case without measurement noise.

To prove the above statements, we first characterize a feasible solution to problem (30), which we later exploit to establish our suboptimality bound. The proof of Lemma 3 and Theorem 2 is reported in the Appendices F and G, respectively.

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

8

Lemma 3 (Feasible solution): Let $\eta = \epsilon_2 \left\| \mathbf{\Phi}_{uy}^\star \right\|_2$ and $\zeta = \epsilon_\infty \left\| \mathbf{\Phi}_{uy}^\star \right\|_\infty$. Assume that the estimation errors are small enough to guarantee $\eta < \frac{1}{5}$ and $\zeta < \frac{1}{2}$, and select $\alpha \in [\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, \epsilon_2^{-1}]$. Consider the following optimization problem and its optimal solutions $\mathbf{\Phi}^c$:

$$\mathbf{\Phi}^c \in \arg\min_{\mathbf{\Phi}} \quad \left\| \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} & \mathbf{\Phi}_{yy}\mathbf{y}_0 \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} & \mathbf{\Phi}_{uy}\mathbf{y}_0 \end{bmatrix} \right\|_F \quad (31)$$

$$\text{subject to } \begin{bmatrix} I & -\mathbf{G} \end{bmatrix} \mathbf{\Phi} = \begin{bmatrix} I & 0 \end{bmatrix}, \quad \mathbf{\Phi} \begin{bmatrix} -\mathbf{G} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

$$\left\| \mathbf{\Phi}_{uy} \right\|_2 \leq \left\| \mathbf{\Phi}_{uy}^\star \right\|_2, \left\| \mathbf{\Phi}_{uy} \right\|_\infty \leq \left\| \mathbf{\Phi}_{uy}^\star \right\|_\infty,$$

$$\phi_{1,j}(\mathbf{\Phi}) + \phi_{2,j}(\mathbf{\Phi}) + \phi_{3,j}(\mathbf{\Phi}) \leq \mathbf{b}_{y,j}, \quad (32)$$

$$\phi_{4,j}(\mathbf{\Phi}) + \phi_{5,j}(\mathbf{\Phi}) + \phi_{6,j}(\mathbf{\Phi}) \leq \mathbf{b}_{u,j}, \quad (33)$$

$$\forall j = 1, \ldots, sN,$$

$$\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu} \text{ with causal sparsities}.$$

where

$$\phi_{1,j}(\mathbf{\Phi}) = \frac{v_\infty \|F_{y,j}\mathbf{\Phi}_{yy}\|_1^\star}{1-2\zeta}, \quad \phi_{4,j}(\mathbf{\Phi}) = \frac{v_\infty \|F_{u,j}\mathbf{\Phi}_{uy}\|_1^\star}{1-2\zeta},$$

$$\phi_{2,j}(\mathbf{\Phi}) = w_\infty \left\| \begin{bmatrix} (F_{y,j}\mathbf{\Phi}_{yu})^\mathsf{T} \\ 2\frac{\epsilon_\infty + \zeta\|\widehat{\mathbf{G}}\|_\infty}{1-2\zeta} (F_{y,j}\mathbf{\Phi}_{yy})^\mathsf{T} \end{bmatrix} \right\|_1,$$

$$\phi_{5,j}(\mathbf{\Phi}) = w_\infty \left\| \begin{bmatrix} (F_{u,j}\mathbf{\Phi}_{uu})^\mathsf{T} \\ 2\frac{\epsilon_\infty + \zeta\|\widehat{\mathbf{G}}\|_\infty}{1-2\zeta} (F_{u,j}\mathbf{\Phi}_{uy})^\mathsf{T} \end{bmatrix} \right\|_1,$$

$$\phi_{3,j}(\mathbf{\Phi}) = F_{y,j}\mathbf{\Phi}_{yy}\widehat{\mathbf{y}}_0 + 2\frac{\epsilon_\infty + \zeta\|\widehat{\mathbf{y}}_0\|_\infty}{1-2\zeta} \|F_{y,j}\mathbf{\Phi}_{yy}\|_1^\star,$$

$$\phi_{6,j}(\mathbf{\Phi}) = F_{u,j}\mathbf{\Phi}_{uy}\widehat{\mathbf{y}}_0 + 2\frac{\epsilon_\infty + \zeta\|\widehat{\mathbf{y}}_0\|_\infty}{1-2\zeta} \|F_{u,j}\mathbf{\Phi}_{uy}\|_1^\star.$$

Then, the following expressions

$$\widetilde{\mathbf{\Phi}}_{yy} = \mathbf{\Phi}_{yy}^c(I+\mathbf{\Delta}\mathbf{\Phi}_{uy}^c)^{-1}, \quad \widetilde{\mathbf{\Phi}}_{yu} = \mathbf{\Phi}_{yy}^c(I+\mathbf{\Delta}\mathbf{\Phi}_{uy}^c)^{-1}(\mathbf{G}-\mathbf{\Delta}),$$

$$\widetilde{\mathbf{\Phi}}_{uy} = \mathbf{\Phi}_{uy}^c(I+\mathbf{\Delta}\mathbf{\Phi}_{uy}^c)^{-1}, \quad \widetilde{\mathbf{\Phi}}_{uu} = (I+\mathbf{\Phi}_{uy}^c\mathbf{\Delta})^{-1}\mathbf{\Phi}_{uu}^c,$$

$$\widetilde{\gamma} = \frac{\sqrt{2}\eta}{\epsilon_2(1-\eta)}, \quad \widetilde{\tau} = \frac{\zeta}{\epsilon_\infty(1-\zeta)}, \quad (34)$$

provide a feasible solution to problem (30).

The main idea behind Lemma 3 is to construct a feasible solution to problem (30) from the set of closed-loop responses generated applying a cautious ground-truth optimal controller $\mathbf{K}^c = \mathbf{\Phi}_{uy}^c(\mathbf{\Phi}_{yy}^c)^{-1}$ on the estimated system $\widehat{\mathbf{G}}$. In the absence of safety constraints, such a feasible solution could directly be established from the ground-truth optimal policy $\mathbf{K}^\star$ similar to [5]. In the constrained case, however, one cannot expect the optimal solution $\mathbf{K}^\star$ to be feasible for $\widehat{\mathbf{G}}$ in (30) since (26)-(27) are more stringent than (14)-(15). Hence, in (31) we first compute $\mathbf{K}^c = \mathbf{\Phi}_{uy}^c(\mathbf{\Phi}_{yy}^c)^{-1}$ as the optimal linear policy for the real system $\mathbf{G}$ under safety constraints that are more stringent than those of (30), and subsequently define $\widetilde{\mathbf{\Phi}}$ as the closed-loop responses generated applying $\mathbf{K}^c$ to $\widehat{\mathbf{G}}$. In this way, $\widetilde{\mathbf{\Phi}}$ is guaranteed to be feasible for (30), provided that the model mismatch is sufficiently small.

Clearly, the optimal solution $\mathbf{K}^c = \mathbf{\Phi}_{uy}^c(\mathbf{\Phi}_{yy}^c)^{-1}$ to (31) will yield a suboptimal cost $J(\mathbf{G}, \mathbf{K}^c) \geq J(\mathbf{G}, \mathbf{K}^\star)$. We

denote the corresponding suboptimality gap as

$$S(\epsilon_\infty) = \frac{J(\mathbf{G}, \mathbf{K}^c)^2 - J(\mathbf{G}, \mathbf{K}^\star)^2}{J(\mathbf{G}, \mathbf{K}^\star)^2}. \quad (35)$$

Note that, if the estimation error ϵ_∞ is too large, the optimization problem (31) may become infeasible. This is expected as the uncertainty level might be incompatible with the required safety. On the other hand, if the optimal solution to the non-noisy problem (17) does not activate the safety constraints, then the constraints of (31) remain inactive for small enough ϵ_∞. In such case we have that $S(\epsilon_\infty) = 0$.

We are now ready to state the main suboptimality result.

Theorem 2: Let $\eta = \epsilon_2 \left\| \mathbf{\Phi}_{uy}^\star \right\|_2$ and $\zeta = \epsilon_\infty \left\| \mathbf{\Phi}_{uy}^\star \right\|_\infty$. Assume that the estimation errors are small enough to guarantee $\eta < \frac{1}{5}$ and $\zeta < \frac{1}{2}$, and select $\alpha \in \left[\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, 5 \left\| \mathbf{\Phi}_{uy}^\star \right\|_2 \right]$. Moreover, assume that ϵ_∞ is small enough for the optimization problem (31) to be feasible. Then, when applying the controller $\widehat{\mathbf{K}}^\star$ optimizing (28) to the true plant $\mathbf{G}$, the relative error with respect to the true optimal cost is upper bounded as

$$\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}} \leq 20\eta + 4(M^c + V^c) + 4S(\epsilon_\infty)(1 + M^c + V^c)$$

$$= \mathcal{O}\left(\epsilon_2 \left(1 + \left\| \mathbf{\Phi}_{uy}^\star \right\|_2 \right) \left(1 + \|\mathbf{G}\|_2 + \|\mathbf{y}_0\|_2 \right)^2 \right) +$$

$$+ 4S(\epsilon_\infty)(1 + M^c + V^c), \quad (36)$$

where

$$M^c = h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\| \mathbf{\Phi}_{uy}^c \right\|_2, \mathbf{G})$$

$$+ h(\epsilon_2, \left\| \mathbf{\Phi}_{uy}^c \right\|_2, \mathbf{y}_0),$$

$$V^c = h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\| \mathbf{\Phi}_{uy}^c \right\|_2, \mathbf{y}_0).$$

We have expressed the suboptimality gap in the form (36) to highlight the presence of two main parts; the first addend scales as $\mathcal{O}\left(\epsilon_2 \left(1 + \left\| \mathbf{\Phi}_{uy}^\star \right\|_2 \right) \left(1 + \|\mathbf{G}\|_2 + \|\mathbf{y}_0\|_2 \right)^2 \right)$ and the second addend $S(\epsilon_\infty)(1 + M^c + V^c)$ is linked to the suboptimality of the tightened optimization program (31). The most important observation is that the suboptimality decreases at most *linearly* with ϵ_2 when $\max(\epsilon_2, \epsilon_\infty)$ is small enough. A linear suboptimality rate in the output-feedback case has first been observed for the unconstrained setup of [5]. Recovering a similar suboptimality rate for the general case with constraints is one of the main novelties of our work. Indeed, despite recovering an upper bound that scales similarly to [5], the corresponding analysis in Appendix G is significantly complicated by the fact that the feasible solution used in [5] cannot be exploited anymore. Hence, one might expect that the suboptimality rate will worsen with respect to the unconstrained case of [5]. Theorem 2 shows, however, that the bound does not deteriorate for small-enough model mismatch levels. Turning our attention to the term $S(\epsilon_\infty)$, we observe through examples (cfr. Figure 1b) that $S(\epsilon_\infty)$ sharply transitions from 0 to ∞ as ϵ_∞ increases. In practice, this example suggests that $S(\epsilon_\infty)$ might be interpreted as an indicator function; if $S(\epsilon_\infty) \approx 0$, then ϵ_∞ is small enough for the linear suboptimality rate to hold.

Our suboptimality bound (36) indicates features of the underlying unknown system that make it easier to be safely

controlled based on noisy data. Notably, the suboptimality grows quadratically with the norm of the true impulse and free responses. This fact implies that an unknown unstable system will be more difficult to control for a long horizon. Last, we note that, surprisingly, our rate in terms of ϵ_2 matches that of [10] which was valid under the assumption of exact state measurements. In other words, our analysis shows that near-optimality can be ensured in complex data-driven control scenarios that combine hard safety requirements with noisy output measurements.

V. NUMERICAL EXPERIMENTS

In this section, we demonstrate numerically the effectiveness of the proposed framework in safely controlling unknown systems. In the experiments, we consider the single-input single-output unknown LTI system characterized by the matrices

$$A = \rho \begin{bmatrix} 1 & 0.25 \\ 0 & 1 \end{bmatrix}, \ B = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}, \ C = \begin{bmatrix} 1 & -1 \end{bmatrix}, \qquad (37)$$

where $\rho > 0$ corresponds to the spectral radius of A. When $\rho < 1$, (37) is asymptotically stable, that is, its output converges to the origin at an exponential rate when the input is equal to 0. When $\rho = 1$, (37) is a marginally stable double-integrator system.

In all the following tests, the cost function is given by (3) for appropriate choices of the weights. The expectation in (3) is taken over future input/output disturbances with covariance matrices $\Sigma_w = I_m$ and $\Sigma_v = I_p$. We consider bounded disturbances between -1 and 1, that is, $w_\infty = v_\infty = 1$. Hence, each scalar disturbance is randomly chosen from $\{-1, 1\}$ with probability $\frac{1}{2}$. For solving optimization problems we use MOSEK [49], called through MATLAB via YALMIP [50][7].

A. Example: safe controller synthesis from noisy data

In our first test, we synthesize a safe output-feedback controller for system (37) with $\rho = 1$ from noisy data. We assume that $x_0 = x(1) = \begin{bmatrix} 6 & 0 \end{bmatrix}^\mathsf{T}$, where we set the initial time at $t = 1$ rather than $t = 0$ for compliance with MATLAB's indexing of vector entries.

The safety constraints are: $y(1) \in \mathbb{R}$ and

$$-5.5 \le y(t) \le 5.5, \quad \forall t = 2, \dots, 12,$$
$$-100 \le u(t) \le 100, \quad \forall t = 1, \dots, 11,$$

for all realizations of noise $\|\mathbf{w}\|_\infty, \|\mathbf{v}\|_\infty \le 1$, while minimizing the cost (3) with the weights $\mathbf{Q}$ and $\mathbf{R}$ in (17) set to the identity.

We first synthesize the optimal controller assuming that the available data are not affected by noise. To this end, we cast and solve the convex optimization problem (40). We verify that the optimal controller $\mathbf{K}^\star$ yields a cost $J(\mathbf{G}, \mathbf{K}^\star) = 69.88$. The green tubes in Figure 1a show the regions containing 50 realizations of the optimal closed-loop input and output trajectories. Due to the high level of noise, we can observe

a significant variability in the trajectory values for different noise realizations. Nonetheless, all trajectories are safe.

We then discuss the case where the available data are affected by noise. In order for the tightened constraints of (30) to be feasible, we consider noisy estimates $(\widehat{\mathbf{G}}, \widehat{\mathbf{y}}_0)$ with $\epsilon = 0.01$ and compute a near-optimal solution to the proposed optimization problem (28). As discussed in Section III-B, this can be achieved by 1) extensive or random search over γ and τ, or 2) extensive search over τ and golden-section search over γ. Even if the first solution comes without strong theoretical guarantees, extensive search over γ and τ may be simpler to implement as it avoids the delicate task of tuning the parameter α. Specifically, for this example we have searched over 100 randomly extracted values of γ and τ in the interval $[0, \epsilon^{-1})$. A potential improvement to this heuristic could be to use a bisection algorithm, as proposed in [44] for example.
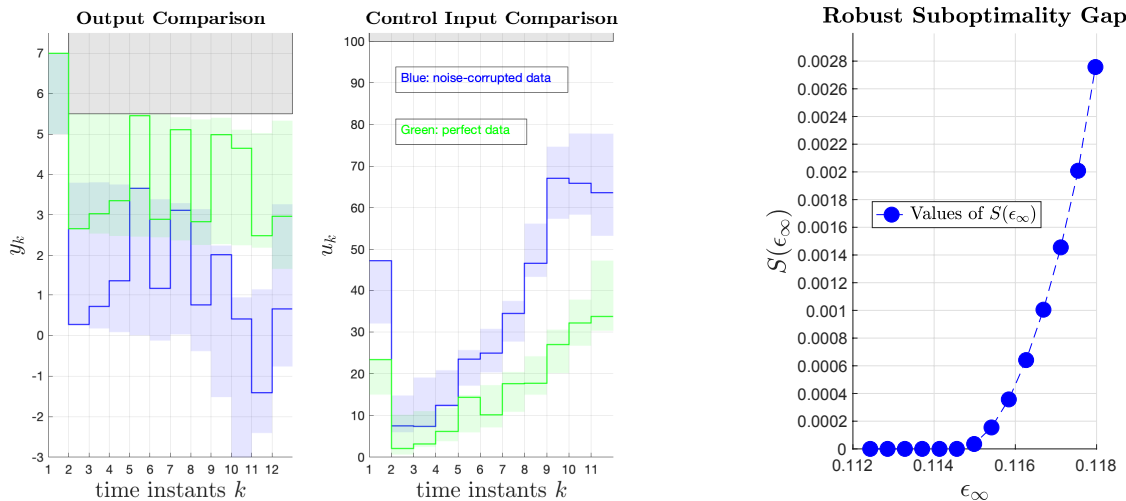
Proceeding as above, we synthesize a robustly safe controller $\widehat{\mathbf{K}}^\star$ yielding a cost of $J(\mathbf{G}, \widehat{\mathbf{K}}^\star) = 140.54$. The corresponding suboptimality gap is $\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}} = 3.049$. In Figure 1a, the trajectories and variability levels resulting from $\widehat{\mathbf{K}}^\star$ for 50 noise realizations are plotted in blue. We observe that, since $\widehat{\mathbf{K}}^\star$ is synthesized using noise-corrupted data, it leads to safer, but more conservative trajectories. Indeed, due to uncertainty, higher control effort is spent to keep the output further from the constraints.

It is informative to inspect the robust suboptimality gap $S(\epsilon_\infty)$ incurred by the tightened optimization problem (31) that we have used in the analysis to characterize a feasible solution to (30). In Figure 1b, we plot $S(\epsilon_\infty)$ assuming $x_0 = x(1) = \begin{bmatrix} 1 & 0 \end{bmatrix}^\mathsf{T}$ and requiring $-3 \le y(t) \le 3$ for $t = 1, \dots, 7$. The example exhibits a fast transition from infeasibility for $\epsilon_\infty > 0.118$ to near-optimality for $\epsilon_\infty < 0.115$. This fact leads to the following observation: high-performing safe controllers can be synthesized by solving (30) even when the optimization problem (31) is infeasible, i.e. $S(\epsilon_\infty) = \infty$. In such cases the suboptimality bound (36) is not applicable, but a robustly safe controller has been synthesized nonetheless. This phenomenon is consistent with the numerical examples of [10] for the state-feedback case.

B. Example: suboptimality scaling beyond least-squares estimation

The bound (36) in Theorem 2 states that a low estimation error level ϵ is crucial in ensuring safety and near-optimality when controlling unknown systems based on noisy data. One advantage of the proposed formulation is that it is directly compatible with behavioral estimation approaches beyond LS identification for the reconstruction of the impulse and free responses, such as data-enabled Kalman filtering [20] and SMM [21], [27]. In our last test, we drop the constraints for both the input and the outputs thus putting our focus on 1) validating the linear scaling of the suboptimality gap (36), and 2) showcasing that, for instance, SMM-based estimation [27] may lead to significantly lower error levels given the same amount of data. We consider the system (37) with different values of $\rho \in [0.9, 0.93, 0.96, 0.99, 1]$ and $x_0 = x(1) = \begin{bmatrix} 6 & 0 \end{bmatrix}^\mathsf{T}$, over a time-horizon of length $N = 11$. The cost

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

10

(a) Closed-loop trajectories. The grey region indicates unsafe input and output values. The green and blue regions contain trajectories for 50 noise realizations obtained through $\mathbf{K}^\star$ and $\widehat{\mathbf{K}}^\star$, respectively. Green and blue lines represent a specific trajectory in both settings.

(b) Robust suboptimality gap $S(\epsilon_\infty)$. This quantity can be interpreted as an indicator as to whether the guarantee (36) holds for a given ϵ_∞.

Fig. 1: Safe controller synthesis for system (37).
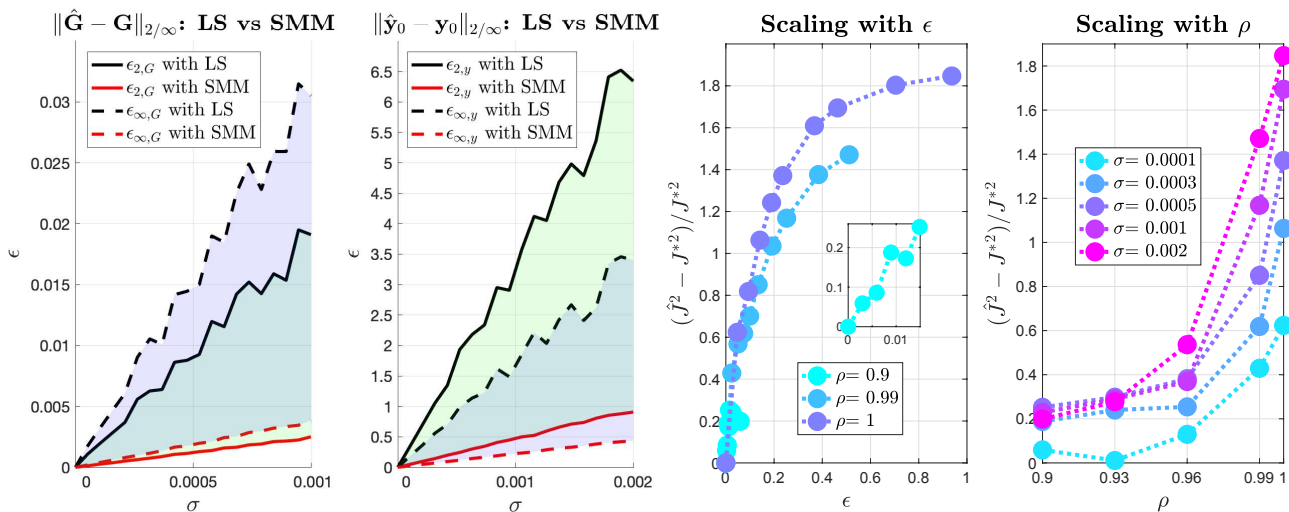
(a) Estimation error in function of the corrupting noise. ML estimation through the SMM yields significantly smaller errors than LS. The green and blue regions indicate the gap for the 2-norm and ∞-norm, respectively.

(b) Suboptimality gap as a function of ϵ_2 (obtained through SMM estimation) for increasing values of the spectral radius ρ of matrix A (on the left). Suboptimality gap as a function of ρ for increasing values of σ (on the right).

Fig. 2: Examples for estimation and suboptimality scaling.

function weights in (17) are selected as $Q(t) = I_p$ for every $t = 1, \ldots, 11$, $Q(12) = 20I_p$ and $R(t) = 0.05$ for every $t = 1, \ldots, 11$.

1) Behavioral estimation: LS vs SMM : For a fixed value of ρ, we gather system trajectories of length 200 time-steps which are corrupted by input and output Gaussian noise with covariance matrices equal to σI. For each experiment, we fix the variance $\sigma \geq 0$ and select a random exploration control input $\mathbf{u}$. We collect 1000 different trajectories for different realizations of the corrupting noise. For each realization of the trajectories, we compute 1) the LS solution (G_{LS}, g_{LS})

using (41) and the corresponding impulse and free responses $\widetilde{\mathbf{G}}_{LS}, \widetilde{\mathbf{y}}_{0,LS}$, and 2) the ML solution (G_{ML}, g_{ML}) using (42)-(43) and the corresponding impulse and free responses $\widetilde{\mathbf{G}}_{ML}, \widetilde{\mathbf{y}}_{0,ML}$. For each estimation, we determine the incurred error levels $\epsilon_{2,G}, \epsilon_{\infty,G}, \epsilon_{2,y}$ and $\epsilon_{\infty,y}$[8]. Last, we record the 90-th percentile of these values, both for SMM and LS estimation.

In Figure 2a we compare the values of ϵ_2 and ϵ_∞ incurred by both estimation techniques. We observe that SMM may yield significantly smaller estimation errors than LS identi-

---

[8]Since the real system is unavailable, in practice this can be done using a bootstrap procedure.

fication. While a full sample-complexity analysis is still unavailable beyond least-squares [3], [23], [26], these examples showcase an advantage in using more sophisticated estimation techniques for safe data-driven control.

2) Suboptimality scaling: Having exploited ML estimation to construct approximate impulse and free responses and the corresponding error-levels, we are ready to solve the optimization problem (30). Since constraints are not present in this example, (30) can be simplified to the quasiconvex formulation we have proposed in [25], where the optimization variable τ is not present. The parameter α is tuned empirically in the interval $\alpha \in [\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, \epsilon_2^{-1})$.[9]

Figure 2b shows the suboptimality gap one incurs by applying the controller $\widehat{\mathbf{K}}^\star$ obtained through the proposed approach. On the left, we consider increasing levels of the estimation error level ϵ_2 for each choice of the spectral radius $\rho = 0.9, 0.99, 1$. On the right, we conversely consider increasing levels of the spectral radius for each choice of the estimation error level ϵ_2. In both cases, we plot the suboptimality gap $\frac{\hat{J}^2 - J^{\star 2}}{J^{\star 2}}$. It can be observed that, consistently with Theorem 2, 1) the gap linearly converges to 0 as ϵ_2 converges to 0, and 2) the gap may grow faster than linearly with the spectral radius ρ as a larger ρ generally leads to larger $\|\mathbf{G}\|_2$. We also observe that larger ρ may lead to higher model mismatch values ϵ. Finally, we remark that, in finite-horizon, our formulations are valid for unstable systems with $\rho > 1$. However, it is inherently challenging to collect trajectories of an unstable system, as the values to be plugged into the corresponding optimization problems will become too large to be handled by numerical solvers. For unstable systems in a data-driven scenario, it is common to assume knowledge of a pre-stabilizing controller [5], [6].

VI. Conclusions

In this paper, we have analyzed how much the model-mismatch due to noisy data can impact the safety and performance of output-feedback control systems with constraints. By deriving a suitable problem relaxation, we have proven that, despite the presence of constraints, the suboptimality of our proposed problem relaxation increases at most linearly for small model mismatches incurred during system identification.

While the proposed approach can synthesize safe and near-optimal output-feedback controllers from noisy data, our relaxed problem might be infeasible for fairly small error levels. Feasibility issues may be significantly mitigated by using soft and chance constraints, or ellipsoidal model mismatch sets such as those of [40], [51]. Future work also includes analyzing the suboptimality in a receding-horizon setup using closed-loop predictions, as well as investigating the advantages of directly optimizing based on the data trajectories rather than performing an identification step.

Acknowledgments

---

[9]The value $\eta = \epsilon_2 \|\mathbf{\Phi}_{uy}^\star\|_2$ is unknown in practice because $\mathbf{\Phi}_{uy}^\star$ is unavailable. One can then tune α according to $\alpha < \epsilon_2^{-1}$.

References

[1] G. Baggio, D. S. Bassett, and F. Pasqualetti, "Data-driven control of complex networks," *Nature communications*, vol. 12, no. 1, pp. 1–13, 2021.

[2] B. Recht, "A tour of reinforcement learning: The view from continuous control," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, pp. 253–279, 2019.

[3] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the Linear Quadratic Regulator," *Foundations of Computational Mathematics*, pp. 1–47, 2019.

[4] M. Fazel, R. Ge, S. Kakade, and M. Mesbahi, "Global convergence of policy gradient methods for the linear quadratic regulator," in *International Conference on Machine Learning*. PMLR, 2018, pp. 1467–1476.

[5] Y. Zheng, L. Furieri, M. Kamgarpour, and N. Li, "Sample complexity of linear quadratic gaussian (lqg) control for output feedback systems," in *Learning for Dynamics and Control*. PMLR, 2021, pp. 559–570.

[6] M. Simchowitz, K. Singh, and E. Hazan, "Improper learning for non-stochastic control," in *Conference on Learning Theory*. PMLR, 2020, pp. 3320–3436.

[7] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, "Logarithmic regret bound in partially observable linear dynamical systems," *arXiv preprint arXiv:2003.11227*, 2020.

[8] K. Zhang, B. Hu, and T. Basar, "Policy optimization for $\mathcal{H}_2$ linear control with $\mathcal{H}_\infty$ robustness guarantee: Implicit regularization and global convergence," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 179–190.

[9] A. Tsiamis, N. Matni, and G. Pappas, "Sample complexity of kalman filtering for unknown systems," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 435–444.

[10] S. Dean, S. Tu, N. Matni, and B. Recht, "Safely learning to control the constrained Linear Quadratic Regulator," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 5582–5588.

[11] S. Fattahi, N. Matni, and S. Sojoudi, "Efficient learning of distributed linear-quadratic control policies," *SIAM Journal on Control and Optimization*, vol. 58, no. 5, pp. 2927–2951, 2020.

[12] L. Furieri, Y. Zheng, and M. Kamgarpour, "Learning the globally optimal distributed LQ regulator," in *Learning for Dynamics and Control*. PMLR, 2020, pp. 287–297.

[13] J. C. Willems and J. W. Polderman, *Introduction to mathematical systems theory: a behavioral approach*. Springer Science & Business Media, 1997, vol. 26.

[14] J. Coulson, J. Lygeros, and F. Dörfler, "Data-enabled predictive control: In the shallows of the DeePC," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 307–312.

[15] J. Coulson, J. Lygeros, and F. Dorfler, "Distributionally robust chance constrained data-enabled predictive control," *IEEE Transactions on Automatic Control*, 2021.

[16] F. Dörfler, J. Coulson, and I. Markovsky, "Bridging direct & indirect data-driven control formulations via regularizations and relaxations," *arXiv preprint arXiv:2101.01273*, 2021.

[17] C. De Persis and P. Tesi, "Formulas for data-driven control: Stabilization, optimality, and robustness," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 909–924, 2020.

[18] J. Berberich, J. Köhler, M. A. Müller, and F. Allgöwer, "Data-driven model predictive control with stability and robustness guarantees," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1702–1717, 2021.

[19] A. Russo and A. Proutiere, "Poisoning attacks against data-driven control methods," *arXiv preprint arXiv:2103.06199*, 2021.

[20] D. Alpago, F. Dörfler, and J. Lygeros, "An extended Kalman filter for data-enabled predictive control," *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 994–999, 2020.

[21] M. Yin, A. Iannelli, and R. S. Smith, "Maximum likelihood estimation in data-driven modeling and control," *IEEE Transactions on Automatic Control*, 2021.

[22] H. J. Vanwaarde, M. K. Camlibel, and M. Mesbahi, "From noisy data to feedback controllers: non-conservative design via a matrix S-lemma," *IEEE Transactions on Automatic Control, to appear*, 2020.

[23] A. Xue and N. Matni, "Data-driven system level synthesis," in *Learning for Dynamics and Control*. PMLR, 2021, pp. 189–200.

[24] C. De Persis and P. Tesi, "Low-complexity learning of linear quadratic regulators from noisy data," *Automatica*, vol. 128, p. 109548, 2021.

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

12

[25] L. Furieri, B. Guo, A. Martin, and G. Ferrari-Trecate, "A behavioral input-output parametrization of control policies with suboptimality guarantees," in *2021 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 2539–2544.

[26] S. Oymak and N. Ozay, "Non-asymptotic identification of LTI systems from a single trajectory," in *2019 American Control Conference (ACC)*. IEEE, 2019, pp. 5655–5661.

[27] A. Iannelli, M. Yin, and R. S. Smith, "Experiment design for impulse response identification with signal matrix models," *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 625–630, 2021.

[28] J. Berberich, J. Köhler, M. A. Muller, and F. Allgower, "Linear tracking MPC for nonlinear systems part II: The data-driven case," *IEEE Transactions on Automatic Control*, 2022.

[29] P. Van Overschee and B. De Moor, *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media, 2012.

[30] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski, "Optimization over state feedback policies for robust control with constraints," *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.

[31] J. Sieber, S. Bennani, and M. N. Zeilinger, "A system level approach to tube-based model predictive control," *IEEE Control Systems Letters*, 2021.

[32] P. J. Goulart and E. C. Kerrigan, "Output feedback receding horizon control of constrained systems," *International Journal of Control*, vol. 80, no. 1, pp. 8–20, 2007.

[33] Y. Zheng, L. Furieri, A. Papachristodoulou, N. Li, and M. Kamgarpour, "On the equivalence of Youla, System-level and Input-output parameterizations," *IEEE Transactions on Automatic Control*, pp. 1–8, 2020.

[34] L. Furieri, Y. Zheng, A. Papachristodoulou, and M. Kamgarpour, "An Input-Output Parametrization of stabilizing controllers: amidst Youla and System Level Synthesis," *IEEE Control Systems Letters*, vol. 3, no. 4, pp. 1014–1019, 2019.

[35] A. Bemporad, "Reducing conservativeness in predictive control of constrained systems with disturbances," in *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No. 98CH36171)*, vol. 2. IEEE, 1998, pp. 1384–1389.

[36] Y. Zheng, L. Furieri, M. Kamgarpour, and N. Li, "System-level, input–output and new parameterizations of stabilizing controllers, and their numerical computation," *Automatica*, vol. 140, p. 110211, 2022.

[37] L. Furieri and M. Kamgarpour, "Unified approach to convex robust distributed control given arbitrary information structures," *IEEE Transactions on Automatic Control*, 2019.

[38] H. Mania, S. Tu, and B. Recht, "Certainty equivalence is efficient for linear quadratic control," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[39] Y. Zheng and N. Li, "Non-asymptotic identification of linear dynamical systems using multiple trajectories," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1693–1698, 2020.

[40] M. Yin, A. Iannelli, and R. S. Smith, "Data-driven prediction with stochastic data: Confidence regions and minimum mean-squared error estimates," *arXiv preprint arXiv:2111.04789*, 2021.

[41] M. Tanaskovic, L. Fagiano, R. Smith, and M. Morari, "Adaptive receding horizon control for constrained mimo systems," *Automatica*, vol. 50, no. 12, pp. 3019–3029, 2014.

[42] E. Terzi, L. Fagiano, M. Farina, and R. Scattolini, "Learning-based predictive control for linear systems: A unitary approach," *Automatica*, vol. 108, p. 108473, 2019.

[43] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *Journal of machine learning research*, vol. 13, no. 2, 2012.

[44] S. Chen, H. Wang, M. Morari, V. M. Preciado, and N. Matni, "Robust closed-loop model predictive control via system level synthesis," in *2020 59th IEEE Conference on Decision and Control (CDC)*. IEEE, 2020, pp. 2152–2159.

[45] A. Agrawal and S. Boyd, "Disciplined quasiconvex programming," *Optimization Letters*, pp. 1–15, 2020.

[46] J. Kiefer, "Sequential minimax search for a maximum," *Proceedings of the American mathematical society*, vol. 4, no. 3, pp. 502–506, 1953.

[47] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[48] N. Matni, Y.-S. Wang, and J. Anderson, "Scalable system level synthesis for virtually localizable systems," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 3473–3480.

[49] MOSEK Aps, "The MOSEK optimization toolbox for MATLAB manual. Version 8.1." 2017.

[50] J. Löfberg, "YALMIP : A Toolbox for Modeling and Optimization in MATLAB," in *In Proc. of the CACSD Conf.*, Taipei, Taiwan, 2004.

[51] Y. Abbasi-Yadkori and C. Szepesvári, "Regret bounds for the adaptive control of linear quadratic systems," in *Proceedings of the 24th Annual Conference on Learning Theory*, 2011, pp. 1–26.

[52] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.

[53] I. Markovsky and P. Rapisarda, "Data-driven simulation and control," *International Journal of Control*, vol. 81, no. 12, pp. 1946–1959, 2008.

APPENDIX

A. Willems' lemma and behavioral theory for synthesizing safe controllers

We recall the definition of persistency of excitation and the result known as the *Fundamental Lemma* for LTI systems [52].

Definition 1: We say that $\mathbf{u}^h_{[0,T-1]}$ is *persistently exciting* (PE) of order L if the Hankel matrix $\mathcal{H}_L(\mathbf{u}^h_{[0,T-1]})$ is full row-rank.

A necessary condition for the matrix $\mathcal{H}_L(\mathbf{u}^h_{[0,T-1]})$ to be full row-rank is that it has at least as many columns as rows. It follows that the input trajectory $\mathbf{u}^h_{[0,T-1]}$ must be long enough to satisfy $T \geq (m+1)L - 1$.

Lemma 4 (Theorem 3.7, [52]): Consider system (1). Assume that (A, B) is controllable and that there is no noise. Let $\{\mathbf{y}^h_{[0,T-1]}, \mathbf{u}^h_{[0,T-1]}\}$ be a system trajectory of length T that has been recorded during a past experiment. Then, if $\mathbf{u}^h_{[0,T-1]}$ is PE of order $n + L$, the signals $\mathbf{y}^\star_{[0,L-1]} \in \mathbb{R}^{pL}$ and $\mathbf{u}^\star_{[0,L-1]} \in \mathbb{R}^{mL}$ are trajectories of (1) if and only if there exists $g \in \mathbb{R}^{T-L+1}$ such that

$$\begin{bmatrix} \mathcal{H}_L(\mathbf{y}^h_{[0,T-1]}) \\ \mathcal{H}_L(\mathbf{u}^h_{[0,T-1]}) \end{bmatrix} g = \begin{bmatrix} \mathbf{y}^\star_{[0,L-1]} \\ \mathbf{u}^\star_{[0,L-1]} \end{bmatrix} . \tag{38}$$

We proceed by showing how Lemma 4 allows one to derive a data-driven formulation of (17) when the data are not noisy. We work under the following assumptions that are standard in the behavioral framework.

Assumption 2: The data-generating LTI system (1) is such that (A, B) is controllable and (A, C) is observable.

Assumption 3: The historical input trajectory $\mathbf{u}^h_{[0,\tilde{T}-1]}$ is PE of order $n + T_{ini} + N$, where $T_{ini} \geq l$ and l is the smallest integer such that the matrix

$$\begin{bmatrix} C^\mathsf{T} & (CA)^\mathsf{T} & \dots & (CA^{l-1})^\mathsf{T} \end{bmatrix}^\mathsf{T} ,$$

has full row-rank. Note that if Assumption 2 holds, then $l \leq n$.

Further, we give the following definition.

Definition 2: The available data in **D1** are further split as follows:

i) a *recent* system trajectory of length T_{ini}: $\left\{ \mathbf{y}^r_{[0,T_{ini}-1]}, \mathbf{u}^r_{[0,T_{ini}-1]} \right\}$, with $\mathbf{y}^r_{[0,T_{ini}-1]} = \mathbf{y}_{[-T_{ini},-1]}$ and $\mathbf{u}^r_{[0,T_{ini}-1]} = \mathbf{u}_{[-T_{ini},-1]}$,

ii) a *historical* system trajectory of length $\tilde{T}$: $\left\{ \mathbf{y}^h_{[0,\tilde{T}-1]}, \mathbf{u}^h_{[0,\tilde{T}-1]} \right\}$, with $\mathbf{y}^h_{[0,\tilde{T}-1]} = \mathbf{y}_{[-T_h,-T_h+\tilde{T}-1]}$ and $\mathbf{u}^h_{[0,\tilde{T}-1]} = \mathbf{u}_{[-T_h,-T_h+\tilde{T}-1]}$ for $T_h \in \mathbb{N}$ such that $T_h \geq \tilde{T}$ and $\tilde{T} \leq T$.

The *historical* data are to be used in substitution of the system model, while the *recent* data reflect the system initial state $x_0 \in \mathbb{R}^n$ [53]. By exploiting (38), one can derive a constrained version of the BIOP derived in [25] as follows:

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

13

Proposition 5 (Safe Behavioral IOP): Consider the LTI system (1), whose parameters (A, B, C, x_0) are *unknown*, and let Assumptions 2-3 hold. Further assume that the historical and recent trajectories are *not* affected by noise. Let (G, g) be any solutions to the linear system of equations

$$\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix} \begin{bmatrix} G & g \end{bmatrix} = \begin{bmatrix} 0_{mT_{ini} \times m} & \mathbf{u}^r_{[0,T_{ini}-1]} \\ 0_{pT_{ini} \times m} & \mathbf{y}^r_{[0,T_{ini}-1]} \\ \begin{bmatrix} I_m & 0_{m \times m(N-1)} \end{bmatrix}^\mathsf{T} & 0_{mN \times 1} \end{bmatrix}, \quad (39)$$

where $\begin{bmatrix} U_p \\ U_f \end{bmatrix} = \mathcal{H}_{T_{ini}+N}(\mathbf{u}^h_{[0,\tilde{T}-1]})$ and $\begin{bmatrix} Y_p \\ Y_f \end{bmatrix} = \mathcal{H}_{T_{ini}+N}(\mathbf{y}^h_{[0,\tilde{T}-1]})$. Then, the optimization problem (17) is equivalent to

$$\min_{\boldsymbol{\Phi}} \left\| \begin{bmatrix} \mathbf{Q}^{\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}_{yy} & \boldsymbol{\Phi}_{yu} \\ \boldsymbol{\Phi}_{uy} & \boldsymbol{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_v^{\frac{1}{2}} & 0 & Y_f g \\ 0 & \boldsymbol{\Sigma}_w^{\frac{1}{2}} & 0 \end{bmatrix} \right\|_F^2 \quad (40)$$

$$\text{subject to } \begin{bmatrix} I & -\text{Toep}(Y_f G) \end{bmatrix} \boldsymbol{\Phi} = \begin{bmatrix} I & 0 \end{bmatrix},$$

$$\boldsymbol{\Phi} \begin{bmatrix} -\text{Toep}(Y_f G) \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

$$\left\| \begin{bmatrix} v_\infty (F_{y,j} \boldsymbol{\Phi}_{yy})^\mathsf{T} \\ w_\infty (F_{y,j} \boldsymbol{\Phi}_{yu})^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1 + (F_{y,j} \boldsymbol{\Phi}_{yy}) Y_f g \le \mathbf{b}_{y,j},$$

$$\left\| \begin{bmatrix} v_\infty (F_{u,j} \boldsymbol{\Phi}_{uy})^\mathsf{T} \\ w_\infty (F_{u,j} \boldsymbol{\Phi}_{uu})^\mathsf{T} \end{bmatrix}^\mathsf{T} \right\|_1 + (F_{u,j} \boldsymbol{\Phi}_{uy}) Y_f g \le \mathbf{b}_{u,j},$$

$$\forall j = 1, \dots, sN,$$

$$\boldsymbol{\Phi}_{yy}, \boldsymbol{\Phi}_{yu}, \boldsymbol{\Phi}_{uy}, \boldsymbol{\Phi}_{uu} \text{ with causal sparsities.}$$

The proof of Proposition 5 is analogous to that of Theorem 1 in [25], with the addition of the safety constraints as per Proposition 1. Since the historical and recent data are not noisy, $Y_f G$ and $Y_f g$ yield the *true* impulse response matrix $\mathbf{G}$ and free response $\mathbf{y}_0$ and the optimal solution of (40) recovers the optimal safe controller $\mathbf{K}^\star$ for the real system.

In practice, exact historical and recent data are not available. As per the noise model in the dynamics (1)-(2), one may assume that historical and recent trajectories are affected by additive noise $w^h(t), w^r(t), v^h(t), v^r(t)$[10] at all time instants, with zero expected values and variances $\boldsymbol{\Sigma}_w^h, \boldsymbol{\Sigma}_w^r, \boldsymbol{\Sigma}_v^h, \boldsymbol{\Sigma}_v^r$ respectively. Hence, the matrix on the left-hand-side of (39) becomes full row-rank almost surely, and any solution to (39) leads to potentially different estimates of the system free and impulse responses, which do not necessarily match the exact ones. This issue is well-known in the behavioral theory literature, and several mitigation strategies have recently been proposed [14], [15], [17], [20], [21], [27]. For instance, a behavioral LS estimator akin to the impulse-response identification of [5], [26] is given by

$$\begin{bmatrix} G_{LS} & g_{LS} \end{bmatrix} = \begin{bmatrix} \hat{U}_p \\ \hat{Y}_p \\ \hat{U}_f \end{bmatrix}^+ \begin{bmatrix} 0_{mT_{ini} \times m} & \mathbf{u}^r_{[0,T_{ini}-1]} \\ 0_{pT_{ini} \times m} & \mathbf{y}^r_{[0,T_{ini}-1]} \\ \begin{bmatrix} I_m & 0_{m \times m(N-1)} \end{bmatrix}^\mathsf{T} & 0_{mN \times 1} \end{bmatrix}, \quad (41)$$

while the ML estimator [21] is computed as

$$G_{ML} = \arg\min_G \quad -log \left[p \left(\begin{bmatrix} \Xi_y \\ Y_f G \end{bmatrix} \mid G, Y_f \right) \right] \quad (42)$$

$$\text{subject to } \begin{bmatrix} \hat{U}_p \\ \hat{U}_f \end{bmatrix} G = \begin{bmatrix} 0_{mT_{ini} \times m} \\ \begin{bmatrix} I_m & 0_{m \times m(N-1)} \end{bmatrix}^\mathsf{T} \end{bmatrix},$$

$$g_{ML} = \arg\min_g \quad -log \left[p \left(\begin{bmatrix} \xi_y \\ Y_f g \end{bmatrix} \mid g, Y_f \right) \right] \quad (43)$$

$$\text{subject to } \begin{bmatrix} \hat{U}_p \\ \hat{U}_f \end{bmatrix} g = \begin{bmatrix} \mathbf{u}^r_{[0,T_{ini}-1]} \\ 0_{mN \times 1} \end{bmatrix},$$

where the residuals $\Xi_y = (Y_p - \hat{Y}_p)G$ and $\xi_y = (Y_p - \hat{Y}_p)g$ denote the fitting deviation from the most recent output measurements, and $p(a|b)$ indicates the probability of event a conditioned to b.

B. Proof of Proposition 1

For the first statement, notice that the controller $\mathbf{K}$ achieves the closed-loop responses (12). Now select $(\boldsymbol{\Phi}_{yy}, \boldsymbol{\Phi}_{yu}, \boldsymbol{\Phi}_{uy}, \boldsymbol{\Phi}_{uu})$ as

$$\begin{bmatrix} \boldsymbol{\Phi}_{yy} & \boldsymbol{\Phi}_{yu} \\ \boldsymbol{\Phi}_{uy} & \boldsymbol{\Phi}_{uu} \end{bmatrix} = \begin{bmatrix} (I - \mathbf{GK})^{-1} & (I - \mathbf{GK})^{-1}\mathbf{G} \\ \mathbf{K}(I - \mathbf{GK})^{-1} & (I - \mathbf{KG})^{-1} \end{bmatrix}, \quad (44)$$

and $\mathbf{q} = \boldsymbol{\Phi}_{uu}\mathbf{g}$. Clearly, $\mathbf{K} = \boldsymbol{\Phi}_{uy}\boldsymbol{\Phi}_{yy}^{-1}$ and $\mathbf{g} = \boldsymbol{\Phi}_{uu}^{-1}\mathbf{q}$, and by plugging the corresponding expressions, we verify that (13) and (16) are satisfied. It remains to prove that (14)-(15) are satisfied. In (10), substitute $\mathbf{y}$ and $\mathbf{u}$ with their closed-loop expressions (11). It follows that the addends separately depend on $\mathbf{w}$ or $\mathbf{v}$. Hence, (10) can be rewritten as

$$\max_{\|\mathbf{v}\|_\infty \le v_\infty} (\mathbf{F}_y \boldsymbol{\Phi}_{yy}) \mathbf{v} + \max_{\|\mathbf{w}\|_\infty \le w_\infty} (\mathbf{F}_y \boldsymbol{\Phi}_{yu}) \mathbf{w} +$$
$$+ \mathbf{F}_y \mathbf{G} \mathbf{q} + (\mathbf{F}_y \boldsymbol{\Phi}_{yy}) \mathbf{CP}_A(:,0)x_0 \le \mathbf{b}_y, \quad (45)$$

$$\max_{\|\mathbf{v}\|_\infty \le v_\infty} (\mathbf{F}_u \boldsymbol{\Phi}_{uy}) \mathbf{v} + \max_{\|\mathbf{w}\|_\infty \le w_\infty} (\mathbf{F}_u \boldsymbol{\Phi}_{uu}) \mathbf{w} +$$
$$+ \mathbf{F}_u \mathbf{q} + (\mathbf{F}_y \boldsymbol{\Phi}_{uy}) \mathbf{CP}_A(:,0)x_0 \le \mathbf{b}_u, \quad (46)$$

where the $\max(\cdot)$ is to be intended row-wise. The expressions (45)-(46) are already convex in $\boldsymbol{\Phi}, \mathbf{q}$. To have a more explicit expression, similar to [10] we utilize the well-known property that the $\|\cdot\|_1$ and the $\|\cdot\|_\infty$ vector norms are dual of each other [47], that is $k\|x\|_1 = \max_{\|w\|_\infty \le k} x^\mathsf{T} w$. The result follows immediately by inspecting (45)-(46) and letting x^T be equal to either $F_{y,j}\boldsymbol{\Phi}_{yy}$, $F_{y,j}\boldsymbol{\Phi}_{uy}$, $F_{y,j}\boldsymbol{\Phi}_{yu}$ or $F_{y,j}\boldsymbol{\Phi}_{uu}$, and letting k be equal to either v_∞ or w_∞.

For the second statement, it is easy to notice that $\mathbf{K}$ is causal by construction because $\boldsymbol{\Phi}_{uy}$ and $\boldsymbol{\Phi}_{yy}$ are block lower-triangular. By selecting the controller $\mathbf{K} = \boldsymbol{\Phi}_{uy}\boldsymbol{\Phi}_{yy}^{-1}$ one has

$$(I - \mathbf{G}\boldsymbol{\Phi}_{uy}\boldsymbol{\Phi}_{yy}^{-1})^{-1} = (I - \mathbf{G}\boldsymbol{\Phi}_{uy}(I + \mathbf{G}\boldsymbol{\Phi}_{uy})^{-1})^{-1}$$
$$= ((I + \mathbf{G}\boldsymbol{\Phi}_{uy} - \mathbf{G}\boldsymbol{\Phi}_{uy})(I + \mathbf{G}\boldsymbol{\Phi}_{uy})^{-1})^{-1}$$
$$= I + \mathbf{G}\boldsymbol{\Phi}_{uy} = \boldsymbol{\Phi}_{yy},$$

which shows that $\boldsymbol{\Phi}_{yy}$ is the closed-loop response from $\mathbf{v}_{[0,N-1]} + \mathbf{CP}_A(:,0)x_0$ to $\mathbf{y}_{[0,N-1]}$ as per (12). Similar computations hold for the remaining closed-loop responses. For the safety constraints, select any $\boldsymbol{\Phi}$ and $\mathbf{q}$ complying with (14)-(15). It is easy to verify by direct computation that, for any $\mathbf{w}$ and $\mathbf{v}$, the same input and output trajectories defined at (11) are obtained by letting $\mathbf{K} = \boldsymbol{\Phi}_{uy}\boldsymbol{\Phi}_{yy}^{-1}$ and $\mathbf{g} = \boldsymbol{\Phi}_{uu}^{-1}\mathbf{q}$ in (6), (7), (8). Hence, the safety constraints are satisfied for any disturbance realization.

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

14

C. Proof of Proposition 3

We first prove that $\mathbf{K} \in \mathcal{K} \implies \widehat{\boldsymbol{\Phi}} \in \boldsymbol{\Pi}$, where

$$\widehat{\boldsymbol{\Phi}} := \begin{bmatrix} (I - \widehat{\mathbf{G}}\mathbf{K})^{-1} & (I - \widehat{\mathbf{G}}\mathbf{K})^{-1}\widehat{\mathbf{G}} \\ \mathbf{K}(I - \widehat{\mathbf{G}}\mathbf{K})^{-1} & (I - \mathbf{K}\widehat{\mathbf{G}})^{-1} \end{bmatrix}. \qquad (47)$$

Let us fix $(\boldsymbol{\Delta}, \boldsymbol{\delta}_0) \in \mathcal{E}$. By substitution of $\widehat{\boldsymbol{\Phi}}$ inside the blocks of $\boldsymbol{\Phi}$ defined in the proposition statement, one has

$$\boldsymbol{\Phi} = \begin{bmatrix} (I - (\widehat{\mathbf{G}} + \boldsymbol{\Delta})\mathbf{K})^{-1} & (I - (\widehat{\mathbf{G}} + \boldsymbol{\Delta})\mathbf{K})^{-1}(\widehat{\mathbf{G}} + \boldsymbol{\Delta}) \\ \mathbf{K}(I - (\widehat{\mathbf{G}} + \boldsymbol{\Delta})\mathbf{K})^{-1} & (I - \mathbf{K}(\widehat{\mathbf{G}} + \boldsymbol{\Delta}))^{-1} \end{bmatrix}.$$

From (11)-(12) the closed-loop trajectories obtained by applying $\mathbf{K}$ to the system $\widehat{\mathbf{G}} + \boldsymbol{\Delta}$ are given by

$$\begin{bmatrix} \mathbf{y}(\mathbf{K}, \boldsymbol{\theta}) \\ \mathbf{u}(\mathbf{K}, \boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_{yy} & \boldsymbol{\Phi}_{yu} \\ \boldsymbol{\Phi}_{uy} & \boldsymbol{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{v} + \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0 \\ \mathbf{w} \end{bmatrix}.$$

Proceeding as in the proof of Proposition 1, one can show that "$(\mathbf{y}(\mathbf{K}, \boldsymbol{\theta}), \mathbf{u}(\mathbf{K}, \boldsymbol{\theta})) \in \boldsymbol{\Gamma}$" for every $\boldsymbol{\theta} \in \mathcal{E} \times \mathcal{W} \times \mathcal{V}$ is the same as "(21)-(22)" for every $(\boldsymbol{\Delta}, \boldsymbol{\delta}_0) \in \mathcal{E}$. Since (20) and (23) are verified by construction, the proof is concluded. Further, for any $(\boldsymbol{\Delta}, \boldsymbol{\delta}_0)$, the cost of (18) achieved by $\mathbf{K}$ is identical to the cost of (19) achieved by $\widehat{\boldsymbol{\Phi}}$ as proven in Proposition 2.

Next, we show $\widehat{\boldsymbol{\Phi}} \in \boldsymbol{\Pi} \implies \widehat{\mathbf{K}} \in \mathcal{K}$, where $\widehat{\mathbf{K}} := \widehat{\boldsymbol{\Phi}}_{uy}\widehat{\boldsymbol{\Phi}}_{yy}^{-1}$. Using (20), one can verify that

$$\boldsymbol{\Phi}_{yy} := \widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1} = (I - (\widehat{\mathbf{G}} + \boldsymbol{\Delta})\widehat{\mathbf{K}})^{-1},$$

and similarly, that all other equalities in (47) hold by substituting $\widehat{\boldsymbol{\Phi}}$ with $\boldsymbol{\Phi}$ and $\mathbf{K}$ with $\widehat{\mathbf{K}}$. Then

$$\begin{bmatrix} \mathbf{y}(\widehat{\mathbf{K}}, \boldsymbol{\theta}) \\ \mathbf{u}(\widehat{\mathbf{K}}, \boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Phi}_{yy} & \boldsymbol{\Phi}_{yu} \\ \boldsymbol{\Phi}_{uy} & \boldsymbol{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{v} + \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0 \\ \mathbf{w} \end{bmatrix}.$$

But "(21)-(22)" for every $(\boldsymbol{\Delta}, \boldsymbol{\delta}_0) \in \mathcal{E}$, which hold by definition, imply that $(\mathbf{y}(\widehat{\mathbf{K}}, \boldsymbol{\theta}), \mathbf{u}(\widehat{\mathbf{K}}, \boldsymbol{\theta})) \in \boldsymbol{\Gamma}$ for every $\boldsymbol{\theta} \in \mathcal{E} \times \mathcal{W} \times \mathcal{V}$ (see the proof of Proposition 1). Further, for any $(\boldsymbol{\Delta}, \boldsymbol{\delta}_0)$, the cost of (19) achieved by $\widehat{\boldsymbol{\Phi}}$ is identical to the cost of (18) achieved by $\widehat{\mathbf{K}}$ as proven in Proposition 2.

D. Proof of Lemma 1

The objective function in Proposition 3 can be written as the square-root of the sum of the square of the Frobenius norms of each of its six blocks. For the upper-left block, since $\left\| \widehat{\boldsymbol{\Phi}}_{uy} \right\|_2 \leq \gamma < \epsilon_2^{-1}$ by assumption, we have

$$\|\widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}\|_F \leq \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F \left\| \sum_{k=0}^{\infty} (\boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^k \right\|_2$$
$$\leq \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F \sum_{k=0}^{\infty} \left\| \epsilon_2 \widehat{\boldsymbol{\Phi}}_{uy} \right\|_2^k = \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F \left(1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2 \right)^{-1},$$

where the convergence of the series follows from $\boldsymbol{\Delta}$ and $\widehat{\boldsymbol{\Phi}}_{uy}$ having zero-entries diagonal blocks by construction. Similarly,

$$\|\widehat{\boldsymbol{\Phi}}_{uy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}\|_F \leq \|\widehat{\boldsymbol{\Phi}}_{uy}\|_F \left(1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2 \right)^{-1},$$
$$\|(I - \widehat{\boldsymbol{\Phi}}_{uy}\boldsymbol{\Delta})^{-1}\widehat{\boldsymbol{\Phi}}_{uu}\|_F \leq \|\widehat{\boldsymbol{\Phi}}_{uu}\|_F \left(1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2 \right)^{-1}.$$

Next, we have

$$\|\widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}} + \boldsymbol{\Delta})\|_F$$

$$\leq \|\widehat{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{G}}\|_F + \|\widehat{\boldsymbol{\Phi}}_{yy}\boldsymbol{\Delta}\|_F + \left\| \widehat{\boldsymbol{\Phi}}_{yy} \left(\sum_{k=1}^{\infty} (\boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^k \right) (\widehat{\mathbf{G}} + \boldsymbol{\Delta}) \right\|_F$$

$$\leq \|\widehat{\boldsymbol{\Phi}}_{yu}\|_F + \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F + \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F \frac{\epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2 (\|\widehat{\mathbf{G}}\|_2 + \epsilon_2)}{1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2}$$

$$\leq \frac{\|\widehat{\boldsymbol{\Phi}}_{yu}\|_F + \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F (2 + \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2 \|\widehat{\mathbf{G}}\|_2)}{1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2},$$

and therefore, by developing the squares and using that $\left\| \widehat{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{G}} \right\|_F \leq \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F \|\widehat{\mathbf{G}}\|_2$ we obtain

$$\|\widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}} + \boldsymbol{\Delta})\|_F^2$$
$$\leq \frac{\left(\|\widehat{\boldsymbol{\Phi}}_{yu}\|_F^2 + \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F^2 h(\epsilon_2, \gamma, \widehat{\mathbf{G}}) \right)}{(1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2)^2}.$$

Proceeding analogously, one can also prove that

$$\|\widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}(\widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0)\|_F^2$$
$$\leq \frac{1}{(1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2)^2} \left(\|\widehat{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{y}}_0\|_F^2 + \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F^2 h(\epsilon_2, \gamma, \widehat{\mathbf{y}}_0) \right),$$
$$\|\widehat{\boldsymbol{\Phi}}_{uy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}(\widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0)\|_F^2 \leq$$
$$\frac{1}{(1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2)^2} \left(\|\widehat{\boldsymbol{\Phi}}_{uy}\widehat{\mathbf{y}}_0\|_F^2 + \|\widehat{\boldsymbol{\Phi}}_{uy}\|_F^2 h(\epsilon_2, \gamma, \widehat{\mathbf{y}}_0) \right).$$

Therefore, combining the above inequalities we finally conclude that

$$J(\mathbf{G}, \mathbf{K}) \leq \frac{\left(\left\| \begin{bmatrix} \widehat{\boldsymbol{\Phi}}_{yy} & \widehat{\boldsymbol{\Phi}}_{yu} & \widehat{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \widehat{\boldsymbol{\Phi}}_{uy} & \widehat{\boldsymbol{\Phi}}_{uu} & \widehat{\boldsymbol{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F^2 + \|\widehat{\boldsymbol{\Phi}}_{yy}\|_F^2 (h(\epsilon_2, \gamma, \widehat{\mathbf{G}}) + \right.}{1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2}$$
$$\frac{\left. + h(\epsilon_2, \gamma, \widehat{\mathbf{y}}_0)) + \|\widehat{\boldsymbol{\Phi}}_{uy}\|_F^2 h(\epsilon_2, \gamma, \widehat{\mathbf{y}}_0) \right)^{\frac{1}{2}}}{1 - \epsilon_2 \|\widehat{\boldsymbol{\Phi}}_{uy}\|_2}.$$

E. Proof of Lemma 2

By using the fact that for $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ we have that $\left\| \begin{bmatrix} x^\top & y^\top \end{bmatrix} \right\|_1 = \left\| x^\top \right\|_1 + \left\| y^\top \right\|_1$, the left-hand-sides of (21)-(22) are each made of three addends. The proof hinges on upper bounding each one of them for a generic $(\boldsymbol{\Delta}, \boldsymbol{\delta}_0) \in \mathcal{E}$. We report the full derivations for the most informative of them. Exploiting Holder's inequality and the relation $\left\| I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy} \right\|_\infty \leq \frac{1}{1 - \epsilon_\infty \tau}$, which can be derived by proceeding as in the proof of Lemma 1, we have

$$v_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1} \right\|_1^\star$$

$$\leq v_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy} \right\|_1^\star \left\| (I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1} \right\|_\infty \leq \frac{v_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy} \right\|_1^\star}{1 - \epsilon_\infty \tau},$$

which is equal to $f_{1,j}(\widehat{\boldsymbol{\Phi}})$. Next, recalling $\widehat{\boldsymbol{\Phi}}_{yu} = \widehat{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{G}}$,

$$w_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}} + \boldsymbol{\Delta}) \right\|_1^\star$$
$$\leq w_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yu} \right\|_1^\star + \max_{\|\mathbf{w}\|_\infty \leq w_\infty} |F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}\boldsymbol{\Delta}\mathbf{w}| +$$
$$+ \max_{\|\mathbf{w}\|_\infty \leq w_\infty} |F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy}\boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}} + \boldsymbol{\Delta})\mathbf{w}|$$
$$\leq w_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yu} \right\|_1^\star + w_\infty \epsilon_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy} \right\|_1^\star +$$
$$+ w_\infty \epsilon_\infty \left\| F_{y,j}\widehat{\boldsymbol{\Phi}}_{yy} \right\|_1^\star \left\| \widehat{\boldsymbol{\Phi}}_{uy}(I - \boldsymbol{\Delta}\widehat{\boldsymbol{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}} + \boldsymbol{\Delta}) \right\|_\infty$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

15

$$\leq w_\infty \left\| F_{y,j} \widehat{\boldsymbol{\Phi}}_{yu} \right\|_1^\star + w_\infty \epsilon_\infty \left\| F_{y,j} \widehat{\boldsymbol{\Phi}}_{yy} \right\|_1^\star \left(1 + \tau \frac{\left\| \widehat{\mathbf{G}} \right\|_\infty + \epsilon_\infty}{1 - \epsilon_\infty \tau} \right)$$

$$= w_\infty \left\| F_{y,j} \widehat{\boldsymbol{\Phi}}_{yu} \right\|_1^\star + w_\infty \epsilon_\infty \left\| F_{y,j} \widehat{\boldsymbol{\Phi}}_{yy} \right\|_1^\star \left(\frac{1 + \tau \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \epsilon_\infty \tau} \right)$$

$$= f_{2,j}(\widehat{\boldsymbol{\Phi}}).$$

Lastly, remembering that $\widehat{\boldsymbol{\Phi}}_{uu} = I + \widehat{\boldsymbol{\Phi}}_{uy} \widehat{\mathbf{G}}$ and noticing that

$$(I - \widehat{\boldsymbol{\Phi}}_{uy} \boldsymbol{\Delta})^{-1} \widehat{\boldsymbol{\Phi}}_{uu} = \widehat{\boldsymbol{\Phi}}_{uu} + \widehat{\boldsymbol{\Phi}}_{uy} \boldsymbol{\Delta} (I - \widehat{\boldsymbol{\Phi}}_{uy} \boldsymbol{\Delta})^{-1} \widehat{\boldsymbol{\Phi}}_{uu},$$

we have

$$w_\infty \left\| F_{u,j} (I - \widehat{\boldsymbol{\Phi}}_{uy} \boldsymbol{\Delta})^{-1} \widehat{\boldsymbol{\Phi}}_{uu} \right\|_1^\star$$

$$\leq w_\infty \left\| F_{u,j} \widehat{\boldsymbol{\Phi}}_{uu} \right\|_1^\star$$

$$+ w_\infty \epsilon_\infty \frac{\left\| F_{u,j} \widehat{\boldsymbol{\Phi}}_{uy} \right\|_1^\star}{1 - \epsilon_\infty \left\| \widehat{\boldsymbol{\Phi}}_{uy} \right\|_\infty} \left(\left\| \widehat{\mathbf{G}} \right\|_\infty \left\| \widehat{\boldsymbol{\Phi}}_{uy} \right\|_\infty + 1 \right)$$

$$\leq w_\infty \left\| F_{u,j} \widehat{\boldsymbol{\Phi}}_{uu} \right\|_1^\star + w_\infty \epsilon_\infty \left\| F_{u,j} \widehat{\boldsymbol{\Phi}}_{uy} \right\|_1^\star \frac{1 + \tau \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \epsilon_\infty \tau}$$

$$= f_{5,j}(\widehat{\boldsymbol{\Phi}}).$$

Similar computations allows one to derive the upper bounds for the remaining terms.

F. Proof of Lemma 3

First, it is easy to verify that $\widetilde{\boldsymbol{\Phi}}$ satisfies the constraints in (30); indeed, $\widetilde{\boldsymbol{\Phi}}$ comprises the closed-loop responses when we apply $\mathbf{K}^c$ to the estimated plant $\widehat{\mathbf{G}}$. Next, we have

$$\left\| \widetilde{\boldsymbol{\Phi}}_{uy} \right\|_2 = \left\| \boldsymbol{\Phi}_{uy}^c (I + \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c)^{-1} \right\|_2$$

$$\leq \frac{\left\| \boldsymbol{\Phi}_{uy}^c \right\|_2}{1 - \epsilon_2 \left\| \boldsymbol{\Phi}_{uy}^c \right\|_2} \leq \sqrt{2} \frac{\left\| \boldsymbol{\Phi}_{uy}^c \right\|_2}{1 - \epsilon_2 \left\| \boldsymbol{\Phi}_{uy}^c \right\|_2}$$

$$\leq \sqrt{2} \frac{\left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2}{1 - \epsilon_2 \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2} = \sqrt{2} \frac{\eta}{\epsilon_2 (1 - \eta)} = \widetilde{\gamma}.$$

Since $\alpha \in [\sqrt{2} \frac{\eta}{\epsilon_2(1-\eta)}, \epsilon_2^{-1})$ and $\eta < \frac{1}{5}$, then $\tilde{\gamma} \leq \alpha < \epsilon_2^{-1}$. Hence $\widetilde{\gamma}$ is feasible. Similarly,

$$\left\| \widetilde{\boldsymbol{\Phi}}_{uy} \right\|_\infty = \left\| \boldsymbol{\Phi}_{uy}^c (I + \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c)^{-1} \right\|_\infty$$

$$\leq \frac{\left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty}{1 - \epsilon_\infty \left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty} \leq \frac{\left\| \boldsymbol{\Phi}_{uy}^\star \right\|_\infty}{1 - \epsilon_\infty \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_\infty} = \frac{\zeta}{\epsilon_\infty (1 - \zeta)} = \widetilde{\tau}.$$

Since $\zeta < \frac{1}{2}$, then $\tilde{\tau} < \epsilon_\infty^{-1}$ and hence it is a feasible value for τ. It remains to show that $\widetilde{\boldsymbol{\Phi}}$ satisfies the safety constraints (26)-(27). We know that $\boldsymbol{\Phi}^c$ is feasible for (31), and hence $\phi_{1,j}(\boldsymbol{\Phi}^c) + \phi_{2,j}(\boldsymbol{\Phi}^c) + \phi_{3,j}(\boldsymbol{\Phi}^c) \leq \mathbf{b}_{y,j}$ and $\phi_{4,j}(\boldsymbol{\Phi}^c) + \phi_{5,j}(\boldsymbol{\Phi}^c) + \phi_{6,j}(\boldsymbol{\Phi}^c) \leq \mathbf{b}_{u,j}$. We conclude the proof by showing that $f_{i,j}(\widetilde{\boldsymbol{\Phi}}) \leq \phi_{i,j}(\boldsymbol{\Phi}^c)$ for every $i = 1, \ldots, 6$. We report the full derivations for the most informative terms.

$$f_{1,j}(\widetilde{\boldsymbol{\Phi}}) = \frac{v_\infty \left\| F_{y,j} \left(\boldsymbol{\Phi}_{yy}^c - \boldsymbol{\Phi}_{yy}^c \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c \left(I + \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c \right)^{-1} \right) \right\|_1^\star}{1 - \epsilon_\infty \widetilde{\tau}}$$

$$\leq \frac{v_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star + \frac{v_\infty \epsilon_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star \left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty}{1 - \epsilon_\infty \left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty}}{1 - \epsilon_\infty \widetilde{\tau}}$$

$$\leq \frac{v_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star + \frac{v_\infty \epsilon_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_\infty}{1 - \epsilon_\infty \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_\infty}}{1 - \epsilon_\infty \widetilde{\tau}}$$

$$\leq \frac{v_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star}{1 - 2\zeta} = \phi_{1,j}(\boldsymbol{\Phi}^c).$$

Similarly, it is easy to show that $f_{4,j}(\widetilde{\boldsymbol{\Phi}}) \leq \phi_{4,j}(\boldsymbol{\Phi}^c)$. Next, recalling (34) and observing that

$$\widetilde{\boldsymbol{\Phi}}_{yu} = \boldsymbol{\Phi}_{yu}^c - \boldsymbol{\Phi}_{yy}^c \boldsymbol{\Delta} - \boldsymbol{\Phi}_{yy}^c \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c (I + \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c)^{-1} \widehat{\mathbf{G}},$$

$$\widetilde{\boldsymbol{\Phi}}_{yy} = \boldsymbol{\Phi}_{yy}^c - \boldsymbol{\Phi}_{yy}^c \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c \left(I + \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c \right)^{-1},$$

we have

$$f_{2,j}(\widetilde{\boldsymbol{\Phi}})$$

$$\leq w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c (I + \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c)^{-1} (\mathbf{G} - \boldsymbol{\Delta}) \right\|_1^\star +$$

$$+ w_\infty \epsilon_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c (I + \boldsymbol{\Delta} \boldsymbol{\Phi}_{uy}^c)^{-1} \right\|_1^\star \left(\frac{1 + \widetilde{\tau} \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \epsilon_\infty \widetilde{\tau}} \right)$$

$$\leq w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yu}^c \right\|_1^\star + w_\infty \epsilon_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star \left(1 + \frac{\left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \epsilon_\infty \left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty} \right) +$$

$$+ w_\infty \epsilon_\infty \frac{\left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star \left(\frac{1 + \widetilde{\tau} \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \epsilon_\infty \widetilde{\tau}} \right)}{1 - \epsilon_\infty \left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty} \leq w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yu}^c \right\|_1^\star +$$

$$+ w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star \left(\epsilon_\infty + \frac{\zeta \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \zeta} + \frac{\epsilon_\infty + \frac{\zeta \left\| \widehat{\mathbf{G}} \right\|_\infty}{1 - \zeta}}{\left(1 - \frac{\zeta}{1-\zeta} \right)(1 - \zeta)} \right)$$

$$= w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yu}^c \right\|_1^\star + 2 w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star \frac{(1 - \zeta) \left(\epsilon_\infty + \zeta \left\| \widehat{\mathbf{G}} \right\|_\infty \right)}{1 - 2\zeta}$$

$$\leq w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yu}^c \right\|_1^\star + 2 w_\infty \left\| F_{y,j} \boldsymbol{\Phi}_{yy}^c \right\|_1^\star \frac{\left(\epsilon_\infty + \zeta \left\| \widehat{\mathbf{G}} \right\|_\infty \right)}{1 - 2\zeta}$$

$$\leq \phi_{2,j}(\boldsymbol{\Phi}^c).$$

Similarly, $f_{3,j}(\widetilde{\boldsymbol{\Phi}}) \leq \phi_{3,j}(\boldsymbol{\Phi}^c)$ and $f_{6,j}(\widetilde{\boldsymbol{\Phi}}) \leq \phi_{6,j}(\boldsymbol{\Phi}^c)$. By only noticing that $\left\| \boldsymbol{\Phi}_{uu}^c \right\|_\infty \leq 1 + \left\| \boldsymbol{\Phi}_{uy}^c \right\|_\infty \left(\left\| \widehat{\mathbf{G}} \right\|_\infty + \epsilon_\infty \right)$ and that $(1 + \zeta)(1 - 2\zeta) \leq 1 - \zeta$ for every $\zeta > 0$, analogous computations lead to $f_{5,j}(\widetilde{\boldsymbol{\Phi}}) \leq \phi_{5,j}(\boldsymbol{\Phi}^c)$.

G. Proof of Theorem 2

By denoting as $\widehat{\boldsymbol{\Phi}}^\star$ the closed-loop responses obtained by applying $\widehat{\mathbf{K}}^\star$ to $\widehat{\mathbf{G}}$, we have by Lemma 1 and by $\gamma \leq \alpha$

$$J(\mathbf{G}, \widehat{\mathbf{K}}^\star) \leq \frac{1}{1 - \epsilon_2 \gamma^\star} \times$$

$$\times \left\| \begin{bmatrix} \sqrt{1 + h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0)} \widehat{\boldsymbol{\Phi}}_{yy}^\star & \widehat{\boldsymbol{\Phi}}_{yu}^\star & \widehat{\boldsymbol{\Phi}}_{yy}^\star \widehat{\mathbf{y}}_0 \\ \sqrt{1 + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0)} \widehat{\boldsymbol{\Phi}}_{uy}^\star & \widehat{\boldsymbol{\Phi}}_{uu}^\star & \widehat{\boldsymbol{\Phi}}_{uy}^\star \widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F,$$

where $\gamma^\star$ is optimal for (30). By Lemma 3, under the assumptions on η, ζ, α we have that $(\widetilde{\gamma}, \widetilde{\tau}, \widetilde{\boldsymbol{\Phi}})$ belongs to the feasible set of (30). Hence, by suboptimality of any feasible solution:

$$J(\mathbf{G}, \widehat{\mathbf{K}}^\star) \leq \frac{1}{1 - \epsilon_2 \widetilde{\gamma}} \times$$

This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2022.3180692

16

$$\times \left\| \begin{bmatrix} \sqrt{1 + h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0)} \widetilde{\boldsymbol{\Phi}}_{yy} & \widetilde{\boldsymbol{\Phi}}_{yu} & \widetilde{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1 + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0)} \widetilde{\boldsymbol{\Phi}}_{uy} & \widetilde{\boldsymbol{\Phi}}_{uu} & \widetilde{\boldsymbol{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F.$$

Using the definition of $\widetilde{\boldsymbol{\Phi}}$ from Lemma 3, we now relate

$$\widetilde{C} = \left\| \begin{bmatrix} \sqrt{1 + h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0)} \widetilde{\boldsymbol{\Phi}}_{yy} & \widetilde{\boldsymbol{\Phi}}_{yu} & \widetilde{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1 + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0)} \widetilde{\boldsymbol{\Phi}}_{uy} & \widetilde{\boldsymbol{\Phi}}_{uu} & \widetilde{\boldsymbol{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F,$$

to the optimal cost of problem (31). Recalling the expressions of M^c and V^c, and similarly to Lemma 3,

$$\widetilde{C} = \left(\left\| \begin{bmatrix} \widetilde{\boldsymbol{\Phi}}_{yy} & \widetilde{\boldsymbol{\Phi}}_{yu} & \widetilde{\boldsymbol{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \widetilde{\boldsymbol{\Phi}}_{uy} & \widetilde{\boldsymbol{\Phi}}_{uu} & \widetilde{\boldsymbol{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F^2 + \right.$$
$$\left. + \left(h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) \right) \left\| \widetilde{\boldsymbol{\Phi}}_{yy} \right\|_F^2 + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) \left\| \widetilde{\boldsymbol{\Phi}}_{uy} \right\|_F^2 \right)^{\frac{1}{2}}$$
$$\leq \frac{\sqrt{J(\mathbf{G}, \mathbf{K}^c)^2 + M^c \left\| \boldsymbol{\Phi}_{yy}^c \right\|_F^2 + V^c \left\| \boldsymbol{\Phi}_{uy}^c \right\|_F^2}}{1 - \epsilon_2 \left\| \boldsymbol{\Phi}_{uy}^c \right\|_2}.$$

Thus, we have established the chain of inequalities

$$J(\mathbf{G}, \widehat{\mathbf{K}}^\star)^2 \leq \frac{\widetilde{C}^2}{(1 - \epsilon_2\widetilde{\gamma})^2}$$
$$\leq \frac{\left(J(\mathbf{G}, \mathbf{K}^c)^2 + M^c \left\| \boldsymbol{\Phi}_{yy}^c \right\|_F^2 + V^c \left\| \boldsymbol{\Phi}_{uy}^c \right\|_F^2 \right)}{(1 - \epsilon_2\widetilde{\gamma})^2 (1 - \epsilon_2 \left\| \boldsymbol{\Phi}_{uy}^c \right\|_2)^2}.$$

Next, notice that, by definition, we have $J(\mathbf{G}, \mathbf{K}^c)^2 = (S(\epsilon_\infty) + 1)J(\mathbf{G}, \mathbf{K}^\star)^2$. Recalling that $\left\| \boldsymbol{\Phi}_{uy}^c \right\|_2 \leq \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2$ and $\left\| \boldsymbol{\Phi}_{yy}^c \right\|_2 \leq \left\| \boldsymbol{\Phi}_{yy}^\star \right\|_2$, observing that $\eta < \frac{1}{5}$ implies $1 - (1 + \sqrt{2})\eta \leq 2$, and further noticing that if $M, V > 0$, then $Ma^2 + Vb^2 \leq (M + V)(a^2 + b^2)$, we can establish:

$$\frac{J(\mathbf{G}, \widehat{\mathbf{K}}^\star)^2 - J(\mathbf{G}, \mathbf{K}^\star)^2}{J(\mathbf{G}, \mathbf{K}^\star)^2} \leq \left(\frac{1}{(1 - \epsilon_2 \left\| \boldsymbol{\Phi}_{uy}^c \right\|_2)^2 (1 - \epsilon_2\widetilde{\gamma})^2} \right) \times$$
$$\times \left(S(\epsilon_\infty) + 1 + \frac{M^c \left\| \boldsymbol{\Phi}_{yy}^c \right\|_F^2 + V^c \left\| \boldsymbol{\Phi}_{uy}^c \right\|_F^2}{J(\mathbf{G}, \mathbf{K}^\star)^2} \right) - 1$$
$$\leq \left(\frac{1}{(1 - \eta)^2 (1 - \sqrt{2}\frac{\eta}{1-\eta})^2} - 1 + \frac{S(\epsilon_\infty)}{(1 - \eta)^2 (1 - \sqrt{2}\frac{\eta}{1-\eta})^2} \right) +$$
$$+ \frac{M^c \left\| \boldsymbol{\Phi}_{yy}^c \right\|_F^2 + V^c \left\| \boldsymbol{\Phi}_{uy}^c \right\|_F^2}{(1 - \eta)^2 (1 - \sqrt{2}\frac{\eta}{1-\eta})^2 J(\mathbf{G}, \mathbf{K}^\star)^2}$$
$$\leq \eta \left(\frac{2(1 + \sqrt{2}) - (1 + \sqrt{2})^2\eta}{(1 - (1 + \sqrt{2})\eta)^2} \right) + \frac{S(\epsilon_\infty)}{(1 - (1 + \sqrt{2})\eta)^2} +$$
$$+ \frac{(M^c + V^c)J(\mathbf{G}, \mathbf{K}^c)^2}{(1 - (1 + \sqrt{2})\eta)^2 J(\mathbf{G}, \mathbf{K}^\star)^2}$$
$$\leq 20\eta + 4(M^c + V^c) + 4S(\epsilon_\infty)(1 + M^c + V^c).$$

Last, we prove that $20\eta + 4(M^c + V^c) = \mathcal{O}\left(\epsilon_2 \left(1 + \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2 \right) (1 + \left\| \mathbf{G} \right\|_2 + \left\| \mathbf{y}_0 \right\|_2)^2 \right)$. First, notice that $M^c + V^c \leq M^\star + V^\star$, where
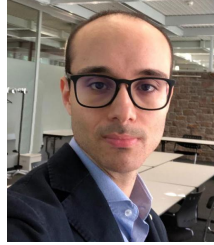
$$M^\star = h(\epsilon_2, \alpha, \widehat{\mathbf{G}}) + h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0)$$
$$+ h(\epsilon_2, \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2, \mathbf{G}) + h(\epsilon_2, \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2, \mathbf{y}_0),$$
$$V^\star = h(\epsilon_2, \alpha, \widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2, \mathbf{y}_0).$$

Using $\alpha \leq 5 \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2$, $\eta < \frac{1}{5}$, $\left\| \widehat{\mathbf{G}} \right\|_2 \leq \left\| \mathbf{G} \right\|_2 + \epsilon_2$ and $\left\| \widehat{\mathbf{y}}_0 \right\|_2 \leq \left\| \mathbf{y}_0 \right\|_2 + \epsilon_2$, we deduce that

$$M^\star \leq 2 \left[\epsilon_2^2(2 + 5 \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2 \left\| \mathbf{G} \right\|_2)^2 + 2\epsilon_2 \left\| \mathbf{G} \right\|_2 (2 + 5 \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2 \left\| \mathbf{G} \right\|_2) \right.$$

$$+ \epsilon_2^2(2 + 5 \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2 \left\| \mathbf{y}_0 \right\|_2)^2$$
$$+ 2\epsilon_2 \left\| \mathbf{y}_0 \right\|_2 (2 + 5 \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2 \left\| \mathbf{y}_0 \right\|_2) \right] + \mathcal{O}(\epsilon_2^2)$$
$$= \mathcal{O}\left(\epsilon_2 \left(1 + \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2 \right) \left(1 + \left\| \mathbf{G} \right\|_2 + \left\| \mathbf{y}_0 \right\|_2 \right)^2 \right),$$

and, similarly, $V^\star = \mathcal{O}\left(\epsilon_2 \left(1 + \left\| \boldsymbol{\Phi}_{uy}^\star \right\|_2 \right) (1 + \left\| \mathbf{y}_0 \right\|_2)^2 \right)$.

Luca Furieri (S'17, M'21) received his B.Sc. and M.Sc. degrees in Automation Engineering from the University of Bologna, Italy, in 2014 and 2016 respectively. He received the Ph.D. degree from ETH Zürich, Switzerland, in 2020. He is currently a Postdoctoral scientist with the Automatic Control Laboratory, EPFL Lausanne, Switzerland. He received the O. Hugo Schuck Best Paper Award in 2018 and was a finalist for the Best Student Paper Award at the 2019 ECC. His research interests lie in the broad areas of learning and optimal control for distributed decision making and safety critical applications.
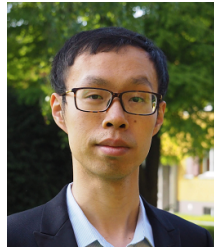
Baiwei Guo received his B.Sc. and M.Sc. degrees in Electrical Engineering respectively from Harbin Institute of Technology in 2016 and ETH Zürich in 2019. He is currently a Ph.D. candidate at the Automatic Control Laboratory at EPF Lausanne. His research interests include statistical learning with applications in control theory and combinatorial optimization.

Andrea Martin is a Ph.D. Student with the Automatic Control Laboratory at EPFL, Lausanne, Switzerland. He received his Double M.Sc. in Automation Engineering and in Automatic Control and Robotics from the University of Padua, Italy, and from the Polytechnic University of Catalonia, Spain, in 2020. He received his B.Sc. in Information Engineering from the University of Padua, Italy, in 2017. His research interests include stochastic system modelling and data-driven control.

Giancarlo Ferrari-Trecate (SM'12) received the Ph.D. degree in Electronic and Computer Engineering from the Università degli Studi di Pavia in 1999. Since September 2016 he is Professor at EPFL, Lausanne, Switzerland. In spring 1998, he was a Visiting Researcher at the Neural Computing Research Group, University of Birmingham, UK. In fall 1998, he joined as a Postdoctoral Fellow the Automatic Control Laboratory, ETH, Zurich, Switzerland. He was appointed Oberassistent at ETH, in 2000. In 2002, he joined INRIA, Rocquencourt, France, as a Research Fellow. From March to October 2005, he was researcher at the Politecnico di Milano, Italy. From 2005 to August 2016, he was Associate Professor at the Dipartimento di Ingegneria Industriale e dell'Informazione of the Universita' degli Studi di Pavia.

His research interests include scalable control, microgrids, machine learning, networked control systems and hybrid systems.

Giancarlo Ferrari Trecate is the founder of the Swiss chapter of the IEEE Control Systems Society and is currently a member of the IFAC Technical Committees on Control Design and Optimal Control. He has been serving on the editorial board of Automatica for nine years and of Nonlinear Analysis: Hybrid Systems.