Learning Robustly Safe Output Feedback Controllers from Noisy Data with Performance Guarantees

Luca Furieri, Andrea Martin, Baiwei Guo, Giancarlo Ferrari-Trecate*

École Polytechnique Fédérale de Lausanne

Abstract

How can we synthesize a safe and near-optimal control policy for a partially-observed system, if all we are given is one historical input/output trajectory that has been corrupted by noise? To address this challenge, we suggest a novel data-driven controller synthesis method, that exploits recent results in controller parametrizations for partially-observed systems [18] and analysis tools from robust control. We provide safety certificates for the learned control policy. Furthermore, the suboptimality of the proposed method shrinks to 0 - and linearly so - in terms of the model mismatch incurred during a preliminary system identification phase.

1 Introduction

Consider a partially-observed linear system

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

where $w(t) \in \mathcal{W}$, $v(t) \in \mathcal{V}$ are disturbances in a compact set. For a time horizon $N \in \mathbb{N}$, consider the **robust optimal control** task of designing a control policy $u(t) = \pi_t(y(t), \cdots, y(0))$ that

i) ensures that trajectories robustly remain inside a target safety set Γ, that is

$$\begin{bmatrix} y(t)^\mathsf{T} & u(t)^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \Gamma\,, \quad \forall w(t) \in \mathcal{W},\ \forall v(t) \in \mathcal{V},\ \forall t = 0, \cdots, N\,, \tag{2}$$

ii) minimizes the expected cost

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

When the linear model (1) is available, it is well-known how to synthesize $\pi_t(\cdot)$ by minimizing (3) subject to (1)-(2) (under the assumption that $\mathcal{W}$, $\mathcal{V}$ and Γ are convex for tractability - see [20, 21]).

However, in many safety-critical engineering systems [24, 2] the state-space parameters (A, B, C, x_0) are *unknown*. Instead, the policy design can only be based on a collection of recorded input and output trajectories. Such data-driven control task is particularly delicate because the historical dataset is itself corrupted by noise, thus posing a challenge for guaranteeing safety and optimality of the synthesized policies.

Recent 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 [9, 14, 26, 34, 30, 23, 31, 39]. When it comes to more complex control tasks, recent advances include [10, 13] for constrained and distributed LQR control with direct state measurements, respectively, and [17] for distributed output-feedback LQG. The approaches of [35, 5, 6, 11, 15] have showcased how constrained Model Predictive Control (MPC) tasks can be solved by plugging historical data into a convex optimization problem. In parallel, [7] introduced data-driven formulations for some controller design tasks. These works inspired several extensions, including closed-loop control with stability guarantees [3], maximum-likelihood (ML) identification for control [22, 36], and nonlinear variants [25].

Nonetheless, it remains rather unexplored how much the performance degrades, and to what extent safety is compromised, in the presence of noise in the historical data. Recently, [33, 8] have derived suboptimality [8] and sample-complexity [33] bounds for fully-observed LQR. A limitation is that the internal system

*Research supported by the Swiss National Science Foundation under the NCCR Automation (grant agreement 51NF40_80545).

states must be fully measured, which is infeasible for several large-scale systems [2]. Furthermore, while [8] proves that for low-enough noise a high-performing and robustly stabilizing controller can be found, the corresponding suboptimality is not quantified as an explicit function of the noise level. Finding such explicit dependency is important in determining whether the noise level in the historical data is tolerable or not. To address these open points, we propose a method for designing safe and near-optimal output-feedback control policies solely based on noise-corrupted data. In addition, we explicitly characterize the incurred suboptimality while guaranteeing safety margins as a function of the model mismatch level.

2 Solution Approach and Theoretical Guarantees

We introduce useful mathematical notation in Appendix A. For compactness, we denote stacked versions of signals of (1) within the horizon $N \in \mathbb{N}$ as $\mathbf{y} = \begin{bmatrix} y^\mathsf{T}(0) & \cdots & y^\mathsf{T}(N-1) \end{bmatrix}^\mathsf{T}$, and similarly for $\mathbf{u}$. We also write: 1) $\mathbf{G}$ to indicate the impulse response matrix of the unknown plant, that is, the causal block-Toeplitz matrix whose block-entries are given by the Markov parameters CA^tB; 2) $\mathbf{y}_0 = \begin{bmatrix} C^\mathsf{T} \cdots (CA^{N-1})^\mathsf{T} \end{bmatrix}^\mathsf{T} x(0)$ to denote the free system response.

The dynamics matrices (A, B, C) and the initial state x_0 are unknown. Instead, the following data are available:

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

D2 The cost matrices Q_t, R_t, the safety set Γ, and the sets of disturbances $\mathcal{W}$ and $\mathcal{V}$ against which we want to be robustly safe.

We make the following assumptions:

A1 The trajectory data (**D1**) have been used to generate estimates $(\widehat{\mathbf{G}}, \widehat{\mathbf{y}}_0)$ of $(\mathbf{G}, \mathbf{y}_0)$ and a *model mismatch* level ϵ_p such that, for both $p = 2$ and $p = \infty$, with high probability

$$\|\mathbf{\Delta}\|_p = \left\|\mathbf{G} - \widehat{\mathbf{G}}\right\|_p \leq \epsilon_p, \quad \|\boldsymbol{\delta}_0\|_p = \|\mathbf{y}_0 - \widehat{\mathbf{y}}_0\|_p \leq \epsilon_p\,.$$

A2 The disturbances $w(t) = B\hat{w}(t)$[1] and $v(t)$ are independent over time, have 0 expected value and variances given by $B\Sigma_{\hat{w}}B^\mathsf{T} \succeq 0$ and $\Sigma_v \succ 0$ respectively.

A3 The sets Γ, $\mathcal{W}$, $\mathcal{V}$ are polytopes.

While **A2** and **A3** are standard technical assumptions that enable tractability of the cost (3) and the constraints (2), assumption **A1** requires further remarks. First, **A1** involves an identification phase to estimate the underlying impulse and free responses. This paradigm is different from the DeePC approach [35, 5] and that of [33], where historical data are directly plugged into an optimization program to synthesize the optimal control policy. Identifying the impulse response matrix $\mathbf{G}$ is also different from identifying (A, B) in the full-observation setup of [9, 10]; indeed, in the output-feedback case, the state-space parameters (A, B, C, x_0) can only be deduced up to an arbitrary change of variables [27]. Instead, the impulse response matrix $\mathbf{G}$ is uniquely defined. Second, in addition to standard least-squares (LS) system identification, **A1** is compatible with state-of-the art estimation approaches based on behavioral theory, such as data-enabled Kalman filtering [1], and the recently proposed Maximum-Likelihood (ML) behavioral estimation [36, 22]; we refer to Appendix B to showcase how behavioral theory can be used to generate estimates $(\widehat{\mathbf{G}}, \widehat{\mathbf{y}}_0)$, including related numerical examples. Last, we note that the model mismatch level $\epsilon_p > 0$, for both $p = 2$ and $p = \infty$, can be obtained in practice through bootstrapping on multiple data-harvesting runs [12].

To enable tractable formulations, we search over linear output-feedback dynamic control policies in the form $u(t) = \sum_{k=0}^{t} K_{t,k}y(k)$, where $K_{t,k}$ are the decision variables, and formulate the *doubly-robust*[2] optimal control problem under consideration:

$$\begin{aligned} \min_{\mathbf{K}} \quad & \max_{\substack{\|\mathbf{\Delta}\|_p \,\leq\, \epsilon_p,\ \|\boldsymbol{\delta}_0\|_p \,\leq\, \epsilon_p \\ \forall p \,\in\, \{2,\, \infty\}}} J(\mathbf{G}, \mathbf{K}) = \sqrt{\mathbb{E}_{\mathbf{w},\mathbf{v}}\left[\mathbf{y}^\mathsf{T}\mathbf{Q}\mathbf{y} + \mathbf{u}^\mathsf{T}\mathbf{R}\mathbf{u}\right]} \\ \text{subject to} \quad & (1),\ (2),\ \mathbf{u} = \mathbf{K}\mathbf{y} + \mathbf{w}\,, \end{aligned} \tag{4}$$

where $\mathbf{K}$ stacks all the decision variables $K_{t,k}$ as a lower-block-triangular causal matrix. Any control policy that complies with the constraints of problem (4) guarantees safety of the input and output trajectories for

[1] As we are only given access to input-output trajectories, and not state trajectories, the control design can only account for process noise in the form $w(t) = B\hat{w}(t)$, where $\hat{w}(t)$ has variance $\Sigma_{\hat{w}}$ - see [16], Remark 1, for instance.

[2] robustness against 1) realizations of $w(t) \in \mathcal{W}$ and $v(t) \in \mathcal{V}$, 2) estimation error $\epsilon_p > 0$.

any realizations of $w(t)$ and $v(t)$, and for any model mismatch $\mathbf{\Delta}$ and $\boldsymbol{\delta}_0$ (within the assumed ϵ-bounds). However, the optimization program (4) is highly non-convex in the decision variables $\mathbf{K}$, $\mathbf{\Delta}$ and $\boldsymbol{\delta}_0$. Our first main result tackles this challenge by formulating a novel restriction[3] of (4), whose tractability is achieved by optimizing over *closed-loop responses* $\mathbf{\Phi} = (\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})^4$, rather than over the policy $\mathbf{K}$ directly. This is enabled by the recently developed Input-Output Parametrization (IOP) [18]; details are provided in Appendix C. From now on, we denote the closed-loop responses that minimize (3) subject to (1)-(2) as $\mathbf{\Phi}^\star$. To simplify the expressions, and without loss of generality, we select the weights $\mathbf{Q}$, $\mathbf{R}$, and the variances Σ_w, Σ_v to be identity matrices of appropriate dimensions.

Theorem 1 *Assume estimation errors $\epsilon_2, \epsilon_\infty$ and fix $0 \leq \alpha < \epsilon_2^{-1}$. Consider the problem*

$$\min_{\gamma\in[0,\alpha),\tau\in[0,\epsilon_\infty^{-1})} \frac{1}{1-\epsilon_2\gamma} \min_{\widehat{\mathbf{\Phi}}} \left\| \begin{bmatrix} \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widehat{\mathbf{\Phi}}_{yy} & \widehat{\mathbf{\Phi}}_{yu} & \widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)}\widehat{\mathbf{\Phi}}_{uy} & \widehat{\mathbf{\Phi}}_{uu} & \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0 \end{bmatrix} \right\|_F . \tag{5}$$
$$\text{subject to} \quad \mathrm{IOP}(\widehat{\mathbf{\Phi}}, \widehat{\mathbf{G}}), \ \left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \leq \gamma, \ \left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \leq \tau, \ f(\epsilon_\infty, \tau, \widehat{\mathbf{\Phi}}) \leq \mathbf{0}.$$

Then, i) all the feasible solutions of (5) *yield a controller $\widehat{\mathbf{K}} = \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}$ complying with the safety constraints* (2) *for the real underlying plant, and ii) its minimal cost upper-bounds that of* (4).

The result is based upon characterizing the relationship between the real closed-loop responses $\mathbf{\Phi}$, i.e. the ones associated with the *true* impulse response matrix $\mathbf{G}$, and the estimated closed-loop responses $\widehat{\mathbf{\Phi}}$, i.e. the ones associated with the *estimated* matrix $\widehat{\mathbf{G}}$. We provide full expressions for IOP$(\cdot)$, $h(\cdot)$ and $f(\cdot)$ in equations (19)-(23), (34) and (37)-(38) of the appendices, respectively. Here, we only highlight that: 1) the function IOP$(\widehat{\mathbf{\Phi}}, \widehat{\mathbf{G}})$ defines all the achievable closed-loop responses associated with the estimated plant $\widehat{\mathbf{G}}$; 2) the scalar function $h(\cdot) > 0$ comes into play in upper-bounding the cost $J(\mathbf{G}, \mathbf{K})$; 3) the function $f(\cdot)$ defines tightened safety constraints. We note that all the constraints can be implemented explicitly using only available information. The key property is that, upon fixing the scalars τ and γ, the remaining optimization program is convex in $\widehat{\mathbf{\Phi}}$. Hence, a near optimal solution to (5) is obtained through gridding on (γ, τ) and solving one convex program for each choice of (γ, τ).

A fundamental question that follows naturally is: how much suboptimal is problem (5) with respect to problem (4)? To answer this question, let $J^\star$ and $(\mathbf{K}^\star, \mathbf{\Phi}^\star)$ denote the min and arg min of (3) subject to (1)-(2), that is, the ground-truth optimal cost and control policy. Our second result is to show that, remarkably, the control policy $\widehat{\mathbf{K}}^\star$ corresponding to the arg min of (5) yields a **near-optimal** cost $\hat{J}$, in the sense that its suboptimality gap $(\hat{J}^2 - J^{\star 2})/J^{\star 2}$ vanishes to 0 as a *linear* function of the estimation error ϵ. The proof relies on considering a specific suboptimal solution of (5), which is characterized as the arg min of a *ground-truth doubly-robust* (GTDR) optimization program; complete details on the proof are available in Appendix E.

Theorem 2 *Let $\eta = \epsilon_2 \left\|\mathbf{\Phi}^\star_{uy}\right\|_2$ and $\zeta = \epsilon_\infty \left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty$. Assume that the model mismatch errors are small enough to guarantee $\eta < \frac{1}{5}$ and $\zeta < \frac{1}{2}$, and assume $\alpha \in [\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)}, \epsilon_2^{-1})$.*[5] *Moreover, assume that ϵ_∞ is small enough for the GTDR optimization program to be feasible, and let J^{GTDR} be its minimal cost. Then, when applying the controller $\widehat{\mathbf{K}}^\star$ optimizing* (5) ***to the ground-truth plant*** $\mathbf{G}$*, the relative error with respect to the ground-truth optimal cost is upper bounded as*

$$(\hat{J}^2 - J^{\star 2})/J^{\star 2} \leq \mathcal{O}\left(\epsilon_2 \left\|\mathbf{\Phi}^\star_{uy}\right\|_2 (\|\mathbf{G}\|_2^2 + \|\mathbf{y}_0\|_2^2)\right) + 4S(\epsilon_\infty), \tag{6}$$

where $S(\epsilon_\infty) = (J^{GTDR^2} - J^{\star 2})/J^{\star 2}$ is the suboptimality incurred by the GTDR optimization program and is such that $S(0) = 0$.

The suboptimality gap (6) has two main parts; the first addend scales linearly with ϵ_2, and the second addend $S(\epsilon_\infty)$, which is linked to the suboptimality of the tightened GTDR program, disappears as $\epsilon_\infty \to 0$. Furthermore, numerical simulations show that $S(\epsilon_\infty)$ sharply transitions from 0 to ∞ as ϵ_∞ increases; see Appendix E, Figure 3 . Hence, in practice, $S(\epsilon_\infty)$ can be interpreted as a boolean value indicating whether ϵ_∞ is small enough for our bound to hold.

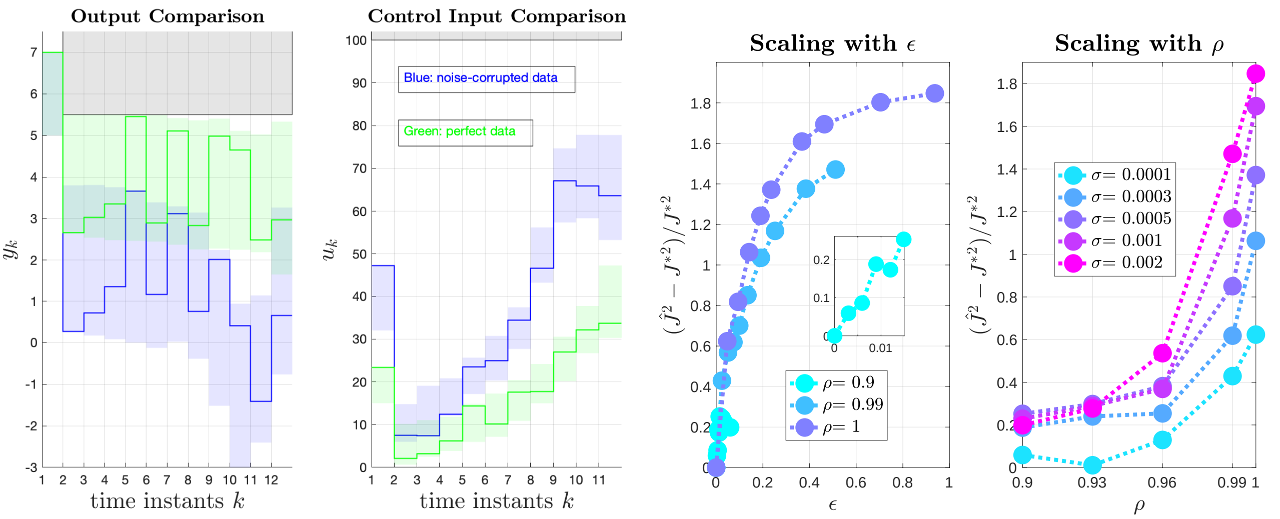
[3]that is, its feasible space is a subset of that of (4).

[4]i.e., the maps relating disturbances to input/output signals upon applying $\mathbf{u} = \mathbf{K}\mathbf{y}$, e.g. $\mathbf{y} = \mathbf{\Phi}_{yu}\mathbf{w}$, where $\mathbf{\Phi}_{yu} = (I - \mathbf{G}\mathbf{K})^{-1}\mathbf{G}$.

[5]The value $\eta = \epsilon_2 \left\|\mathbf{\Phi}^\star_{uy}\right\|_2$ is unknown in practice because so is $\mathbf{\Phi}^\star_{uy}$. One then tunes according to $\alpha < \epsilon_2^{-1}$.

3 Numerical Examples

To validate our approach[6], we consider synthesis based on noisy data harvested from an unknown underlying LTI system. The system has a bidimensional dynamic matrix A with spectral radius $0 < \rho \leq 1$. The goal is to ensure that the output $y(1) = 6$ drops below the safety value 5.5, robustly against: 1) all future disturbances with absolute value bounded by 1, and 2) model mismatch of size $\epsilon = 0.01$. Results for the ideal case



(a) Unsafe regions are colored in grey.

(b) Suboptimality gaps.

$\epsilon = 0$ and the noisy-data case are both shown in Figure (a). As expected, noisy data lead to safer, but more conservative trajectories. Figure (b) validates the suboptimality bound (6). On the left, the gap is shown for increasing levels of noise variance $\sigma > 0$ corrupting historical data samples, leading to increased ϵ levels.[7] For ϵ small enough, the growth is linear as predicted by Theorem 2. On the right, the gap is shown for increasing levels of the spectral radius ρ, showing a growth faster than linear (as indicated by the term $||\mathbf{G}||_2^2$ in (6)).

4 Conclusions

We have shown that, solely relying on *noisy* historical input-output trajectories, one can solve complex optimal control tasks with hard safety requirements. Furthermore, the suboptimality decays linearly for small model mismatch ϵ. Future work will focus on further reducing the conservatism of the method, and on analyzing recursive feasibility for receding-horizon implementations.

References

[1] Daniele Alpago, Florian Dörfler, and John Lygeros. An extended Kalman filter for data-enabled predictive control. *IEEE Control Systems Letters*, 4(4):994–999, 2020.

[2] Giacomo Baggio, Danielle S Bassett, and Fabio Pasqualetti. Data-driven control of complex networks. *Nature communications*, 12(1):1–13, 2021.

[3] Julian Berberich, Johannes Köhler, Matthias A. Müller, and Frank Allgöwer. Data-driven model predictive control with stability and robustness guarantees. *IEEE Transactions on Automatic Control*, 66(4):1702–1717, 2021.

[4] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[5] Jeremy Coulson, John Lygeros, and Florian Dörfler. Data-enabled predictive control: In the shallows of the DeePC. In *2019 18th European Control Conference (ECC)*, pages 307–312. IEEE, 2019.

[6] Jeremy Coulson, John Lygeros, and Florian Dörfler. Distributionally robust chance constrained data-enabled predictive control. *arXiv preprint arXiv:2006.01702*, 2020.

[7] Claudio De Persis and Pietro Tesi. Formulas for data-driven control: Stabilization, optimality, and robustness. *IEEE Transactions on Automatic Control*, 65(3):909–924, 2020.

[6] The code to reproduce the numerical examples is available here.

[7] The impulse response identification is performed using ML behavioral estimation [36].

[8] Claudio De Persis and Pietro Tesi. Low-complexity learning of linear quadratic regulators from noisy data. *Automatica*, 128:109548, 2021.

[9] Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. On the sample complexity of the Linear Quadratic Regulator. *Foundations of Computational Mathematics*, pages 1–47, 2019.

[10] Sarah Dean, Stephen Tu, Nikolai Matni, and Benjamin Recht. Safely learning to control the constrained Linear Quadratic Regulator. In *2019 American Control Conference (ACC)*, pages 5582–5588. IEEE, 2019.

[11] Florian Dörfler, Jeremy Coulson, and Ivan Markovsky. Bridging direct & indirect data-driven control formulations via regularizations and relaxations. *arXiv preprint arXiv:2101.01273*, 2021.

[12] Bradley Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.

[13] Salar Fattahi, Nikolai Matni, and Somayeh Sojoudi. Efficient learning of distributed linear-quadratic control policies. *SIAM Journal on Control and Optimization*, 58(5):2927–2951, 2020.

[14] Maryam Fazel, Rong Ge, Sham Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator. In *International Conference on Machine Learning*, pages 1467–1476. PMLR, 2018.

[15] Felix Fiedler and Sergio Lucia. On the relationship between data-enabled predictive control and subspace predictive control. *arXiv preprint arXiv:2011.13868*, 2020.

[16] Luca Furieri, Baiwei Guo, Andrea Martin, and Giancarlo Ferrari-Trecate. A behavioral input-output parametrization of control policies with suboptimality guarantees. *arXiv preprint arXiv:2102.13338*, 2021.

[17] Luca Furieri, Yang Zheng, and Maryam Kamgarpour. Learning the globally optimal distributed LQ regulator. In *Learning for Dynamics and Control*, pages 287–297. PMLR, 2020.

[18] Luca Furieri, Yang Zheng, Antonis Papachristodoulou, and Maryam Kamgarpour. An Input-Output Parametrization of stabilizing controllers: amidst Youla and System Level Synthesis. *IEEE Control Systems Letters*, 3(4):1014–1019, 2019.

[19] Jean H. Gallier. *CIS 515: Vector Norms and Matrix Norms.* University of Pennsylvania, CIS 515 lecture notes, 2020.

[20] Paul J Goulart, Eric C Kerrigan, and Jan M Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42(4):523–533, 2006.

[21] Paul James Goulart. *Affine feedback policies for robust control with constraints.* PhD thesis, University of Cambridge, 2007.

[22] Andrea Iannelli, Mingzhou Yin, and Roy S Smith. Experiment design for impulse response identification with signal matrix models. *arXiv preprint arXiv:2012.08126*, 2020.

[23] Sahin Lale, Kamyar Azizzadenesheli, Babak Hassibi, and Anima Anandkumar. Logarithmic regret bound in partially observable linear dynamical systems. *arXiv preprint arXiv:2003.11227*, 2020.

[24] Francoise Lamnabhi-Lagarrigue, Anuradha Annaswamy, Sebastian Engell, Alf Isaksson, Pramod Khargonekar, Richard M Murray, Henk Nijmeijer, Tariq Samad, Dawn Tilbury, and Paul Van den Hof. Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. *Annual Reviews in Control*, 43:1–64, 2017.

[25] Yingzhao Lian and Colin N Jones. Nonlinear data-enabled prediction and control. *arXiv preprint arXiv:2101.03187*, 2021.

[26] Dhruv Malik, Ashwin Pananjady, Kush Bhatia, Koulik Khamaru, Peter Bartlett, and Martin Wainwright. Derivative-free methods for policy optimization: Guarantees for linear quadratic systems. In *International Conference on Artificial Intelligence and Statistics*, pages 2916–2925. PMLR, 2019.

[27] Horia Mania, Stephen Tu, and Benjamin Recht. Certainty equivalent control of LQR is efficient. *arXiv preprint arXiv:1902.07826*, 2019.

[28] Ivan Markovsky and Paolo Rapisarda. Data-driven simulation and control. *International Journal of Control*, 81(12):1946–1959, 2008.

[29] Samet Oymak and Necmiye Ozay. Non-asymptotic identification of LTI systems from a single trajectory. In *2019 American Control Conference (ACC)*, pages 5655–5661. IEEE, 2019.

[30] Max Simchowitz, Karan Singh, and Elad Hazan. Improper learning for non-stochastic control. In *Conference on Learning Theory*, pages 3320–3436. PMLR, 2020.

[31] Anastasios Tsiamis, Nikolai Matni, and George Pappas. Sample complexity of kalman filtering for unknown systems. In *Learning for Dynamics and Control*, pages 435–444. PMLR, 2020.

[32] Jan C Willems, Paolo Rapisarda, Ivan Markovsky, and Bart LM De Moor. A note on persistency of excitation. *Systems & Control Letters*, 54(4):325–329, 2005.

[33] Anton Xue and Nikolai Matni. Data-driven system level synthesis. *arXiv preprint arXiv:2011.10674*, 2020.

[34] Y. Zheng and L. Furieri, Maryam Kamgarpour, and Na Li. Sample complexity of linear quadratic gaussian (LQG) control for output feedback systems. *Learning for Dynamics and Control, PMLR*, 2021.

[35] Hua Yang and Shaoyuan Li. A data-driven predictive controller design based on reduced hankel matrix. In *2015 10th Asian Control Conference (ASCC)*, pages 1–7. IEEE, 2015.

[36] Mingzhou Yin, Andrea Iannelli, and Roy S Smith. Maximum likelihood estimation in data-driven modeling and control. *arXiv preprint arXiv:2011.00925*, 2020.

[37] Dante Youla, Hamid Jabr, and Jr Bongiorno. Modern Wiener-Hopf design of optimal controllers–Part II: The multivariable case. *IEEE Trans. on Aut. Contr.*, 21(3):319–338, 1976.

[38] Mario Zanon and Sébastien Gros. Safe reinforcement learning using robust MPC. *IEEE Transactions on Automatic Control*, 2020.

[39] Kaiqing Zhang, Bin Hu, and Tamer 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*, pages 179–190. PMLR, 2020.

[40] Kemin Zhou, John Comstock Doyle, and Keith Glover. *Robust and optimal control*, volume 40. Prentice hall New Jersey, 1996.

A Notation

In this appendix, we detail the mathematical notation used in the abstract for completeness.

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 $n \times n$, $0_{m\times n}$ to denote the zero matrix of size $m \times n$, and 1_n to denote the vector of all ones of length n. We write $M = \text{blkdiag}(M_1, \dots, M_N)$ to denote a block-diagonal matrix with $M_1, \dots, M_N \in \mathbb{R}^{m\times n}$ on its diagonal block entries, and for $\mathbf{M} = \begin{bmatrix} M_1^\mathsf{T} & \dots & M_N^\mathsf{T} \end{bmatrix}^\mathsf{T}$ we define the block-Toeplitz matrix

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

More concisely, we 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 [19]. For a symmetric matrix M, we write $M \succ 0$ or $M \succeq 0$ if and only 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 $\omega(0), \omega(1), \dots, \omega(T-1)$ with $\omega(t) \in \mathbb{R}^n$ for every $t = 0, 1, \dots, T-1$, which can be compactly written as

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

When the value of T is clear from the context, we 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

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

Throughout the abstract, for a row-vector $x \in \mathbb{R}^{1\times n}$, $\|x\|_1 = \sum_{i=1}^n |x_i|$ denotes its 1-norm; this is a slight abuse of notation that allows streamlining the appearance of a few long derivations.

B Behavioral theory for impulse-response identification

First, we briefly recall *Willems' Fundamental Lemma* [32] and useful results from behavioral linear system theory. Second, we showcase how noisy historical data can be used in practice to identify the impulse response matrix $\mathbf{G}$ and free response $\mathbf{y}_0$, incurring small model mismatch.

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 1 (Theorem 3.7, [32]) *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*

$$\begin{bmatrix} \mathcal{H}_L(\mathbf{y}^h_{[0,T-1]}) \\ \mathcal{H}_L(\mathbf{u}^h_{[0,T-1]}) \end{bmatrix} g = \begin{bmatrix} \mathbf{y}^\star_{[0,L-1]} \\ \mathbf{u}^\star_{[0,L-1]} \end{bmatrix}. \tag{7}$$

Willems' Lemma above assumes that historical data are not corrupted by noise. In such ideal setup, one can exactly identify the impulse response matrix $\mathbf{G}$ and the free response $\mathbf{y}_0$ directly from historical data $\{\mathbf{y}^h_{[-T,-1]}, \mathbf{u}^h_{[-T,-1]}\}$ as follows [28].

Lemma 2 (Noiseless data-driven identification) *Assume that* (A, B) *is controllable and* (A, C) *is observable. Further, assume that* $\mathbf{u}^h_{[-T,-1]}$ *is PE of order* $n + T_{ini} + N$*, where* $T_{ini} \geq l$ *and* l *is the smallest integer such that the matrix*

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

has full row-rank. Let (G, g) *be any solutions to the linear system of equations*

$$\begin{bmatrix} U_p \\ Y_p \\ U_f \end{bmatrix} \begin{bmatrix} G & g \end{bmatrix} = \begin{bmatrix} 0_{mT_{ini}\times m} & \mathbf{u}^r_{[0,T_{ini}-1]} \\ 0_{pT_{ini}\times m} & \mathbf{y}^r_{[0,T_{ini}-1]} \\ \begin{bmatrix} I_m & 0_{m\times m(N-1)} \end{bmatrix}^\top & 0_{mN\times 1} \end{bmatrix}, \tag{8}$$

where $\begin{bmatrix} U_p \\ U_f \end{bmatrix} = \mathcal{H}_{T_{ini}+N}(\mathbf{u}^h_{[0,T-1]})$ *and* $\begin{bmatrix} Y_p \\ Y_f \end{bmatrix} = \mathcal{H}_{T_{ini}+N}(\mathbf{y}^h_{[0,T-1]})$*. Then we have*

$$\mathbf{G} = \mathrm{Toep}(Y_f G), \quad \mathbf{y}_0 = Y_f g.$$

Proof: Let G be any solution (8). By rearranging the terms, each column of G can be seen as a solution to (7) associated with a zero initial condition and input $\mathbf{u}_{[0,N-1]} = \begin{bmatrix} e_i^\top & 0_{1\times m(N-1)} \end{bmatrix}^\top$, where $e_i \in \mathbb{R}^m$ is the i-th element of the standard orthogonal basis of $\mathbb{R}^m$. Since the hypotheses of Lemma 1 are satisfied for $L = T_{ini} + N$, similar to Proposition 11 of [28] we deduce that $Y_f G$ is the first block-column of the system impulse response matrix, independent of which solution to (8) is chosen. Finally, note that $Y_f g$ corresponds to the trajectory starting at x_0 (as implicitly defined by the recent trajectory $\mathbf{y}_{[-T_{ini},-1]}$ and $\mathbf{u}_{[-T_{ini},-1]}$) when applying a zero input [28]. Therefore, it corresponds to the true free response starting from x_0.

In practice, however, exact historical and recent data are not available. We assume that historical trajectories are affected by additive noise $w^h(t), v^h(t)$[8] at all time instants, with zero expected values and

[8] where "w" and "v" denote input and output noise, respectively.

variances $\mathbf{\Sigma}_w^h \succeq 0, \mathbf{\Sigma}_v^h \succeq 0$ respectively. Due to additive noise, the matrix on the left-hand-side of (8) becomes full row-rank almost surely, and the least-squares solutions of (8) do not recover the free and impulse responses. This issue is well-known in the behavioral theory literature, and several mitigation strategies have recently been proposed [5, 6, 7, 36, 1, 22]. Specifically, these methods have the goal of yielding accurate and coherent estimators based on behavioral theory. For instance, letting the symbol " $\widehat{\ }$ " denote noisy historical trajectories the standard LS estimator is given by

$$G_{LS} = \begin{bmatrix} \hat{U}_p \\ \hat{Y}_p \\ \hat{U}_f \end{bmatrix}^+ \begin{bmatrix} 0_{mT_{ini} \times m} \\ 0_{pT_{ini} \times m} \\ \begin{bmatrix} I_m & 0_{m \times m(N-1)} \end{bmatrix}^\mathsf{T} \end{bmatrix}, \quad g_{LS} = \begin{bmatrix} \hat{U}_p \\ \hat{Y}_p \\ \hat{U}_f \end{bmatrix}^+ \begin{bmatrix} \widehat{\mathbf{u}}^h_{[-T_{ini},-1]} \\ \widehat{\mathbf{y}}^h_{[-T_{ini},-1]} \\ 0_{mN \times 1} \end{bmatrix}. \tag{9}$$

Another approach is to minimize a scalar functional $f(\cdot)$ that penalizes the residuals $\Xi_y = (Y_p - \hat{Y}_p)G$ and $\xi_y = (Y_p - \hat{Y}_p)g$ [6]. A choice that reflects the maximum-likelihood (ML) interpretation of total least squares is proposed in [36] and consists in solving the optimization problems

$$\begin{aligned} G_{ML} = \arg\min_{G} \quad & -log\left[p\left(\begin{bmatrix} \Xi_y \\ Y_f G \end{bmatrix} \mid G, Y_f\right)\right] \qquad (10) \\ \text{subject to} \quad & \begin{bmatrix} \hat{U}_p \\ \hat{U}_f \end{bmatrix} G = \begin{bmatrix} 0_{mT_{ini} \times m} \\ \begin{bmatrix} I_m & 0_{m \times m(N-1)} \end{bmatrix}^\mathsf{T} \end{bmatrix}, \\ g_{ML} = \arg\min_{g} \quad & -log\left[p\left(\begin{bmatrix} \xi_y \\ Y_f g \end{bmatrix} \mid g, Y_f\right)\right] \qquad (11) \\ \text{subject to} \quad & \begin{bmatrix} \hat{U}_p \\ \hat{U}_f \end{bmatrix} g = \begin{bmatrix} \widehat{\mathbf{u}}_{[-T_{ini},-1]} \\ 0_{mN \times 1} \end{bmatrix}, \end{aligned}$$

where $p(a|b)$ indicates the probability of event a conditioned to b.

B.1 Model mismatch beyond least-squares: numerical example

We showcase that ML-based behavioral estimation [22] may lead to significantly lower model mismatch levels than LS estimation given the same amount of noisy data. For the system $A = \begin{bmatrix} 1 & 0.25 \\ 0 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 0.1 \end{bmatrix}$, $C = \begin{bmatrix} 1 & -1 \end{bmatrix}$, we gather noisy historical trajectories with $T = 200$ and $T_{ini} = 30$ respectively. We assume that the data are corrupted by Gaussian noise with zero mean and variances $\mathbf{\Sigma}_w^h = \mathbf{\Sigma}_w^r = \sigma I_m$ and $\mathbf{\Sigma}_v^h = \mathbf{\Sigma}_v^r = \sigma I_p$, where $\sigma \geq 0$. For each experiment, we fix the variance $\sigma \geq 0$ and select a random exploration control input $\mathbf{u}$. We collect 1000 different historical 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 (9) 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 (10)-(11) 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 ϵ_2, ϵ_∞.[9] Last, we record the 90-th percentile of these values, both for ML and LS estimation.

In Figure (3) we compare the values of ϵ_2 and ϵ_∞ incurred by both estimation techniques. We observe that ML may yield significantly smaller estimation errors than LS identification. While a full sample-complexity analysis is still unavailable beyond least-squares [9, 29, 33], these examples showcase an advantage in using more sophisticated estimation techniques for safe data-driven control.

C (B)IOP

In this appendix, we recall the main concepts behind the IOP [18] which, akin to Youla-based and disturbance-feedback controllers [37, 40], expresses any linear output-feedback control policy in terms of the corresponding closed-loop responses and the affine subspace they belong to.

With the notation introduced in Section 2, we compactly write the relations between finite-horizon input-state-output trajectories of system (1) as

$$\mathbf{x} = \mathbf{P}_A(:,0)x_0 + \mathbf{P}_B\mathbf{u}, \tag{12}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{v}, \tag{13}$$

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

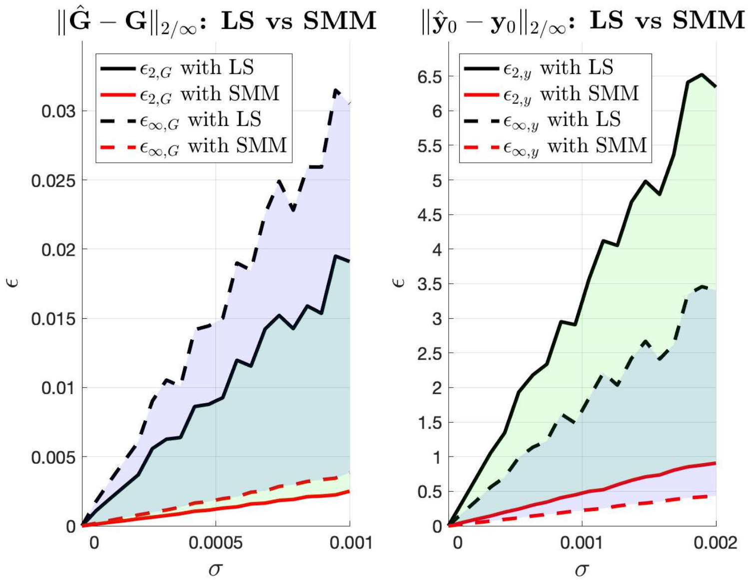


Figure 2: Estimation error as a function of the corrupting noise. ML estimation through the ML estimation yields significantly smaller errors than LS. The green and blue regions indicate the gap for the 2-norm and ∞-norm, respectively. SMM denotes the Signal Matrix Model approximation of ML estimation proposed in [36]; the MATLAB code for implementing the SMM has kindly been provided by the authors of [36].

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

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

Clearly, we have that $\mathbf{G} = \mathbf{CP}_B$ and $\mathbf{y}_0 = \mathbf{CP}_A(:,0)x_0$. Assume now that the input and measurement disturbances satisfy the box constraints $\|w_t\|_\infty \leq w_\infty$ $\|v_t\|_\infty \leq v_\infty$, for given positive constants w_∞ and v_∞. Further, let

$$\Gamma = \{(y,u) \in (\mathbb{R}^p, \mathbb{R}^m)|\ F_y y \leq b_y, F_u u \leq b_u\}\ , \tag{14}$$

with $F_y \in \mathbb{R}^{s\times p}$, $F_u \in \mathbb{R}^{s\times m}$ and $b_y, b_u \in \mathbb{R}^s$, represent the nonempty polytope that defines the safe set for input and output signals. Compliance with the safety requirements

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

over the whole prediction horizon is ensured if and only if

$$\max_{\|\mathbf{v}\|_\infty \leq v_\infty,\ \|\mathbf{w}\|_\infty \leq w_\infty} \mathbf{F}_y\mathbf{y} \leq \mathbf{b}_y\,, \quad \max_{\|\mathbf{v}\|_\infty \leq v_\infty, \|\mathbf{w}\|_\infty \leq w_\infty} \mathbf{F}_u\mathbf{u} \leq \mathbf{b}_u\,, \tag{15}$$

where $\max(\cdot)$ is intended row-wise, namely the maximizers $\mathbf{v}^\star$ and $\mathbf{w}^\star$ are potentially different for each row of $\mathbf{F}_y\mathbf{y}$ and $\mathbf{F}_u\mathbf{u}$. In (15), we define $\mathbf{F}_y = I_N \otimes F_y$, $\mathbf{F}_u = I_N \otimes F_u$, $\mathbf{b}_y = 1_N \otimes b_y$, $\mathbf{b}_u = 1_N \otimes b_u$.

Closing the loop with the causal linear control policy

$$\mathbf{u} = \mathbf{Ky} + \mathbf{w}\,, \tag{16}$$

from equations (12) and (13) we obtain the following relationships between disturbances and input-output signals

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{v}+\mathbf{y}_0 \\ \mathbf{w} \end{bmatrix}\,, \tag{17}$$

where

$$\begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} = \begin{bmatrix} (I-\mathbf{GK})^{-1} & (I-\mathbf{GK})^{-1}\mathbf{G} \\ \mathbf{K}(I-\mathbf{GK})^{-1} & (I-\mathbf{KG})^{-1} \end{bmatrix}\,. \tag{18}$$

The following proposition shows that one can compute the optimal safe feedback control policy by searching over achievable input-output closed-loop responses $\mathbf{\Phi}$, as characterized by an affine subspace defined over the plant $\mathbf{G}$.

Proposition 1 *Consider the LTI system* (1) *evolving under the control policy* (16) *within a horizon of length* $N \in \mathbb{N}$*. Then:*

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

$$\begin{bmatrix} I & -\mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix}, \tag{19}$$

$$\begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} -\mathbf{G} \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \tag{20}$$

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

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

$$\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu} \text{ with causal sparsities }^{10}, \tag{23}$$

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 $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$ *lying in the affine subspace* (19)-(23)*, the matrix* $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ *is causal and it yields the closed-loop responses* $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$*. Moreover, the linear control policy* $\mathbf{K}$ *complies with the safety constraints.*

Proof of Proposition 1 For the first statement, notice that the controller $\mathbf{K}$ achieves the closed-loop responses (18). Now select $(\mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu})$ as

$$\begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} = \begin{bmatrix} (I-\mathbf{GK})^{-1} & (I-\mathbf{GK})^{-1}\mathbf{G} \\ \mathbf{K}(I-\mathbf{GK})^{-1} & (I-\mathbf{KG})^{-1} \end{bmatrix}, \tag{24}$$

Clearly, $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$. By plugging the corresponding expressions, we verify that (19), (20) and (23) are satisfied. It remains to prove that (21)-(22) are satisfied. In (15), substitute $\mathbf{y}$ and $\mathbf{u}$ with their closed-loop expressions (17). It follows that the addends separately depend on $\mathbf{w}$ or $\mathbf{v}$. Hence, (15) can be rewritten as

$$\max_{\|\mathbf{v}\|_\infty \leq v_\infty} (\mathbf{F}_y\mathbf{\Phi}_{yy})\,\mathbf{v} + \max_{\|\mathbf{w}\|_\infty \leq w_\infty} (\mathbf{F}_y\mathbf{\Phi}_{yu})\,\mathbf{w} + (\mathbf{F}_y\mathbf{\Phi}_{yy})\,\mathbf{CP}_A(:,0)x_0 \leq \mathbf{b}_y\,, \tag{25}$$

$$\max_{\|\mathbf{v}\|_\infty \leq v_\infty} (\mathbf{F}_u\mathbf{\Phi}_{uy})\,\mathbf{v} + \max_{\|\mathbf{w}\|_\infty \leq w_\infty} (\mathbf{F}_u\mathbf{\Phi}_{uu})\,\mathbf{w} + (\mathbf{F}_y\mathbf{\Phi}_{uy})\,\mathbf{CP}_A(:,0)x_0 \leq \mathbf{b}_u\,, \tag{26}$$

where the $\max(\cdot)$ is to be intended row-wise, namely the maximizers $\mathbf{v}^\star$ and $\mathbf{w}^\star$ are potentially different for each row. The expressions (25)-(26) are already convex in $\mathbf{\Phi}$. 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 [4], that is $k\,\|x\|_1 = \max_{\|w\|_\infty \leq k} x^\mathsf{T} w$. The result follows immediately by inspecting (25)-(26) and letting x^T be equal to either $F_{y,j}\mathbf{\Phi}_{yy}$, $F_{y,j}\mathbf{\Phi}_{uy}$, $F_{y,j}\mathbf{\Phi}_{yu}$ or $F_{y,j}\mathbf{\Phi}_{uu}$, 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 $\mathbf{\Phi}_{uy}$ and $\mathbf{\Phi}_{yy}$ are block lower-triangular. By selecting the controller $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ one has

$$\begin{aligned}(I - \mathbf{G}\mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1})^{-1} &= (I - \mathbf{G}\mathbf{\Phi}_{uy}(I + \mathbf{G}\mathbf{\Phi}_{uy})^{-1})^{-1} \\ &= ((I + \mathbf{G}\mathbf{\Phi}_{uy} - \mathbf{G}\mathbf{\Phi}_{uy})(I + \mathbf{G}\mathbf{\Phi}_{uy})^{-1})^{-1} = I + \mathbf{G}\mathbf{\Phi}_{uy} = \mathbf{\Phi}_{yy}\,,\end{aligned}$$

which shows that $\mathbf{\Phi}_{yy}$ 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 (18). Similar computations hold for the remaining closed-loop responses. For the safety constraints, select any $\mathbf{\Phi}$ complying with (21)-(22). It is easy to verify by direct computation that, for any $\mathbf{w}$ and $\mathbf{v}$, the same input and output trajectories defined at (17) are obtained by letting $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1}$ in (12), (13), (16). Hence, the safety constraints are satisfied for any disturbance realization.■

Next, linearity of the expectation operator and the identity $\mathbb{E}_x(x^\mathsf{T} M x) = \mathrm{Tr}(M\Sigma_x) + \mu_x^\mathsf{T} M \mu_x$, that holds true for the expected value of a quadratic function, allow to rewrite the objective function $\mathbb{E}_{\mathbf{v},\mathbf{w}}[\mathbf{y}^\mathsf{T}\mathbf{L}\mathbf{y} + \mathbf{u}^\mathsf{T}\mathbf{R}\mathbf{u}]$ in terms of the closed loop responses. Considering the LTI system (1), the linear control policy that achieves

[10] Specifically, they have the block lower-triangular sparsities resulting as per the expressions (18), the sparsity of $\mathbf{K}$ and that of $\mathbf{G}$.

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 program:

$$\begin{aligned} \min_{\mathbf{\Phi}} \; & \left\| \begin{bmatrix} \mathbf{Q}^{\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_v^{\frac{1}{2}} & 0 & \mathbf{y}_0 \\ 0 & \mathbf{\Sigma}_w^{\frac{1}{2}} & 0 \end{bmatrix} \right\|_F^2 \\ & \text{subject to } \; (19) - (23)\,, \end{aligned} \tag{27}$$

where $\mathbf{Q} = \text{blkdiag}(Q_0, \ldots, Q_{N-1})$, $\mathbf{R} = \text{blkdiag}(R_0, \ldots, R_{N-1})$, $\mathbf{\Sigma}_v = I_N \otimes \Sigma_v$, $\mathbf{\Sigma}_w = I_N \otimes \Sigma_w$.

In problem (27), the achievability constraints (19)-(20), the safety constraints (21)-(22) and the cost function are all expressed in terms of the impulse response matrix $\mathbf{G}$ and of the free response $\mathbf{y}_0$. Hence, an explicit state-space description (A, B, C, x_0) is not needed for optimal controller synthesis. In light of this observation, Lemma 2 can be used to embed previously-recorded input-output trajectories into an equivalent data-driven optimization problem. In particular, it naturally follows from equation (8) that the behavioral counterpart of problem (27) is given by

$$\begin{aligned} \min_{\mathbf{\Phi}} \; & \left\| \begin{bmatrix} \mathbf{Q}^{\frac{1}{2}} & 0 \\ 0 & \mathbf{R}^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_v^{\frac{1}{2}} & 0 & Y_f g \\ 0 & \mathbf{\Sigma}_w^{\frac{1}{2}} & 0 \end{bmatrix} \right\|_F^2 \\ \text{subject to} \; & \begin{bmatrix} I & -\text{Toep}(Y_f G) \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} = \begin{bmatrix} I & 0 \end{bmatrix}\,, \\ & \begin{bmatrix} \mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} \\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} \end{bmatrix} \begin{bmatrix} -\text{Toep}(Y_f G) \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ I \end{bmatrix}\,, \\ & \left\| \begin{bmatrix} v_\infty \left(F_{y,j}\mathbf{\Phi}_{yy}\right)^\mathsf{T} \\ w_\infty \left(F_{y,j}\mathbf{\Phi}_{yu}\right)^\mathsf{T} \end{bmatrix} \right\|_1 + \left(F_{y,j}\mathbf{\Phi}_{yy}\right) Y_f g \leq \mathbf{b}_{y,j}\,, \\ & \left\| \begin{bmatrix} v_\infty \left(F_{u,j}\mathbf{\Phi}_{uy}\right)^\mathsf{T} \\ w_\infty \left(F_{u,j}\mathbf{\Phi}_{uu}\right)^\mathsf{T} \end{bmatrix} \right\|_1 + \left(F_{u,j}\mathbf{\Phi}_{uy}\right) Y_f g \leq \mathbf{b}_{u,j}\,, \\ & \forall j = 1, \ldots, sN\,, \\ & \mathbf{\Phi}_{yy}, \mathbf{\Phi}_{yu}, \mathbf{\Phi}_{uy}, \mathbf{\Phi}_{uu} \text{ with causal sparsities.} \end{aligned} \tag{28}$$

Note that the optimization problem (28) requires no direct state observation and solely relies on input-output data Hankel matrices. We denote the resulting parametrization of output-feedback linear policies as Behavioral IOP (BIOP). Future work will focus on direct formulations that allow to completely bypass the non-parametric identification step of the impulse response matrix $\mathbf{G}$.

D Proof of Theorem 1

This appendix presents the proposed reformulation of the intractable problem (4) into a more convenient form. Specifically, we prove two technical lemmas that provide, respectively, a tight upper bound of the cost function, and a tightened - yet more favorable - expression for the safety constraints. Remarkably, the suboptimality incurred by these approximations scales linearly with the model mismatch whenever such discrepancy is sufficiently small, as per Theorem 2.

We start by analytically characterizing the relationship between the real closed-loop responses $\mathbf{\Phi}$, i.e. the ones associated with the *true* impulse response matrix $\mathbf{G}$, and the estimated closed-loop responses $\widehat{\mathbf{\Phi}}$, i.e. the ones associated with the *estimated* impulse matrix $\widehat{\mathbf{G}}$. In particular, recalling that $\mathbf{K} = \mathbf{\Phi}_{uy}\mathbf{\Phi}_{yy}^{-1} = \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}$, $\mathbf{G} = \widehat{\mathbf{G}} + \mathbf{\Delta}$, $\mathbf{y}_0 = \widehat{\mathbf{y}}_0 + \boldsymbol{\delta}_0$, it follows from the achievability constraints over $\widehat{\mathbf{\Phi}}$ and from the Woodbury matrix identity that

$$\begin{aligned}
\mathbf{\Phi}_{yy} &= (I-\mathbf{G}\mathbf{K})^{-1} = \left(I-(\widehat{\mathbf{G}}+\mathbf{\Delta})\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}\right)^{-1} = \left(I-\widehat{\mathbf{G}}\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}\right)^{-1} \\
&= \left((\widehat{\mathbf{\Phi}}_{yy}-\widehat{\mathbf{G}}\widehat{\mathbf{\Phi}}_{uy}-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})\widehat{\mathbf{\Phi}}_{yy}^{-1}\right)^{-1} = \widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}, \qquad (29)\\
\mathbf{\Phi}_{yu} &= (I-\mathbf{G}\mathbf{K})^{-1}\mathbf{G} = \mathbf{\Phi}_{yy}\mathbf{G} = \widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta}), \qquad (30)\\
\mathbf{\Phi}_{uy} &= \mathbf{K}(I-\mathbf{G}\mathbf{K})^{-1} = \mathbf{K}\mathbf{\Phi}_{yy} = \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1}\widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1} = \widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}, \qquad (31)\\
\mathbf{\Phi}_{uu} &= (I-\mathbf{K}\mathbf{G})^{-1} = \mathbf{K}(I-\mathbf{G}\mathbf{K})^{-1}\mathbf{G}+I = \mathbf{\Phi}_{uy}\mathbf{G}+I \\
&= \widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})+I = \left(I-\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta}\right)^{-1}\widehat{\mathbf{\Phi}}_{uy}(\widehat{\mathbf{G}}+\mathbf{\Delta})+I \\
&= (I-\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}(\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{G}}+\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta}+I-\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta}) \\
&= (I-\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}(\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{G}}+I) = (I-\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu}. \qquad (32)
\end{aligned}$$

Lemma 3 *Let $\widehat{\mathbf{\Phi}}$ denote the closed-loop responses obtained by applying $\mathbf{K}$ to $\widehat{\mathbf{G}}$. Further, assume that $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \le \gamma$, where $\gamma \in [0, \epsilon_2^{-1})$. Then, we have*

$$J(\mathbf{G},\mathbf{K}) \le \frac{1}{1-\epsilon_2\gamma}\left\|\begin{bmatrix}\sqrt{1+h(\epsilon_2,\gamma,\widehat{\mathbf{G}})+h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)}\,\widehat{\mathbf{\Phi}}_{yy} & \widehat{\mathbf{\Phi}}_{yu} & \widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{y}}_0 \\ \sqrt{1+h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)}\,\widehat{\mathbf{\Phi}}_{uy} & \widehat{\mathbf{\Phi}}_{uu} & \widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{y}}_0\end{bmatrix}\right\|_F, \qquad (33)$$

where

$$h(\epsilon,\gamma,\mathbf{Y}) = \epsilon^2(2+\gamma\|\mathbf{Y}\|_2)^2 + 2\epsilon\,\|\mathbf{Y}\|_2\,(2+\gamma\,\|\mathbf{Y}\|_2). \qquad (34)$$

Proof of Lemma 3 In order to simplify the notation going forward, let the weights $\mathbf{Q}$, $\mathbf{R}$ in the cost function and the noise variances $\mathbf{\Sigma}_w$, $\mathbf{\Sigma}_v$ be identity matrices of appropriate dimensions. The objective in Problem (4) is given by

$$\left\|\begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu}\\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix}\begin{bmatrix}I & 0 & \mathbf{y}_0\\ 0 & I & 0\end{bmatrix}\right\|_F^2 = \left\|\begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} & \mathbf{\Phi}_{yy}\mathbf{y}_0\\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} & \mathbf{\Phi}_{uy}\mathbf{y}_0\end{bmatrix}\right\|_F^2, \qquad (35)$$

or, equivalently, as the square-root of the sum of the square of the Frobenius norms of each of its six blocks. We proceed by upper-bounding each one of them individually. For the upper-left block, we have

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

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

$$\|\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\|_F \le \frac{\|\widehat{\mathbf{\Phi}}_{uy}\|_F}{1-\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2},\ \|(I-\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu}\|_F \le \frac{\|\widehat{\mathbf{\Phi}}_{uu}\|_F}{1-\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2}.$$

Next, we have

$$\begin{aligned}
\|\widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})\|_F &\le \|\widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{G}}\|_F+\|\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\|_F+\left\|\widehat{\mathbf{\Phi}}_{yy}\left(\sum_{k=1}^{\infty}(\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^k\right)(\widehat{\mathbf{G}}+\mathbf{\Delta})\right\|_F \\
&\le \|\widehat{\mathbf{\Phi}}_{yu}\|_F+\epsilon_2\|\widehat{\mathbf{\Phi}}_{yy}\|_F+\|\widehat{\mathbf{\Phi}}_{yy}\|_F\frac{\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2(\|\widehat{\mathbf{G}}\|_2+\epsilon_2)}{1-\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2} \\
&\le \frac{\|\widehat{\mathbf{\Phi}}_{yu}\|_F+\epsilon_2\|\widehat{\mathbf{\Phi}}_{yy}\|_F(2+\|\widehat{\mathbf{\Phi}}_{uy}\|_2\|\widehat{\mathbf{G}}\|_2)}{1-\epsilon_2\|\widehat{\mathbf{\Phi}}_{uy}\|_2},
\end{aligned}$$

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

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

Proceeding analogously, one can also prove that

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

Therefore, combining the above inequalities we finally conclude that

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

Lemma 3 exploits the upperbound $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_2 \leq \gamma$ to establish an explicit relationship between the cost $J(\mathbf{G},\mathbf{K})$, obtained by applying a controller $\mathbf{K}$ to the real system $\mathbf{G}$, and the cost $J(\widehat{\mathbf{G}},\mathbf{K})$, obtained by applying the same controller to the estimated system $\widehat{\mathbf{G}}$. To see this, notice that (33) is equivalently rewritten as

$$J(\mathbf{G},\mathbf{K}) \leq \frac{\sqrt{J(\widehat{\mathbf{G}},\mathbf{K})^2 + \|\widehat{\mathbf{\Phi}}_{yy}\|_F^2(h(\epsilon_2,\gamma,\widehat{\mathbf{G}})+h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)) + \|\widehat{\mathbf{\Phi}}_{uy}\|_F^2 h(\epsilon_2,\gamma,\widehat{\mathbf{y}}_0)}}{1-\epsilon_2\gamma}. \tag{36}$$

The expression (36) 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 $\widehat{\mathbf{G}},\widehat{\mathbf{y}}_0,\widehat{\mathbf{\Phi}}$.

Lemma 4 *Assume $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \leq \tau$, where $\tau \in [0,\epsilon_\infty^{-1})$. Then, if $\widehat{\mathbf{\Phi}}$ satisfies the tightened safety constraints*

$$f_{1,j}(\widehat{\mathbf{\Phi}}) + f_{2,j}(\widehat{\mathbf{\Phi}}) + f_{3,j}(\widehat{\mathbf{\Phi}}) \leq \mathbf{b}_{y,j}\,, \tag{37}$$
$$f_{4,j}(\widehat{\mathbf{\Phi}}) + f_{5,j}(\widehat{\mathbf{\Phi}}) + f_{6,j}(\widehat{\mathbf{\Phi}}) \leq \mathbf{b}_{u,j}\,, \tag{38}$$
$$\forall j = 1,\ldots,sN\,,$$

where

$$f_{1,j}(\widehat{\mathbf{\Phi}}) = \frac{v_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1}{1-\epsilon_\infty\tau}, \quad f_{2,j}(\widehat{\mathbf{\Phi}}) = w_\infty\left\|\begin{bmatrix}\left(F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right)^\mathsf{T}\\ \epsilon_\infty\frac{1+\tau\|\widehat{\mathbf{G}}\|_\infty}{1-\epsilon_\infty\tau}\left(F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right)^\mathsf{T}\end{bmatrix}\right\|_1,$$
$$f_{4,j}(\widehat{\mathbf{\Phi}}) = \frac{v_\infty\left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right\|_1}{1-\epsilon_\infty\tau}, \quad f_{5,j}(\widehat{\mathbf{\Phi}}) = w_\infty\left\|\begin{bmatrix}\left(F_{u,j}\widehat{\mathbf{\Phi}}_{uu}\right)^\mathsf{T}\\ \epsilon_\infty\frac{1+\tau\|\widehat{\mathbf{G}}\|_\infty}{1-\epsilon_\infty\tau}\left(F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right)^\mathsf{T}\end{bmatrix}\right\|_1,$$
$$f_{3,j}(\widehat{\mathbf{\Phi}}) = F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\widehat{\boldsymbol{y}}_0 + \epsilon_\infty\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1\left(\frac{1+\tau\|\widehat{\mathbf{y}}_0\|_\infty}{1-\epsilon_\infty\tau}\right),$$
$$f_{6,j}(\widehat{\mathbf{\Phi}}) = F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\widehat{\boldsymbol{y}}_0 + \epsilon_\infty\left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right\|_1\left(\frac{1+\tau\|\widehat{\mathbf{y}}_0\|_\infty}{1-\epsilon_\infty\tau}\right),$$

then $\widehat{\mathbf{\Phi}}$ satisfies also the safety constraints

$$\left\|\begin{bmatrix}v_\infty\left(F_{y,j}\mathbf{\Phi}_{yy}\right)^\mathsf{T}\\ w_\infty\left(F_{y,j}\mathbf{\Phi}_{yu}\right)^\mathsf{T}\end{bmatrix}^\mathsf{T}\right\|_1 + \left(F_{y,j}\mathbf{\Phi}_{yy}\right)(\widehat{\mathbf{y}}_0+\boldsymbol{\delta}_0) \leq \mathbf{b}_{y,j}\,, \tag{39}$$
$$\left\|\begin{bmatrix}v_\infty\left(F_{u,j}\mathbf{\Phi}_{uy}\right)^\mathsf{T}\\ w_\infty\left(F_{u,j}\mathbf{\Phi}_{uu}\right)^\mathsf{T}\end{bmatrix}^\mathsf{T}\right\|_1 + \left(F_{u,j}\mathbf{\Phi}_{uy}\right)(\widehat{\mathbf{y}}_0+\boldsymbol{\delta}_0) \leq \mathbf{b}_{u,j}\,, \tag{40}$$

for the real system.

Proof of Lemma 4 By using the fact that for $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ we have that $\left\|\left[x^\top \quad y^\top\right]^\top\right\|_1 = \|x\|_1 + \|y\|_1$, the left-hand-sides of (39)-(40) are each made of three addends. The proof hinges on upper-bounding each one of them. We report the full derivations for the most informative of them. We have

$$\begin{aligned}
v_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\right\|_1 &= \max_{\|\mathbf{v}\|_\infty \le v_\infty} \left(F_{y,j}\widehat{\mathbf{\Phi}}_{yy} + F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\right)\mathbf{v} \\
&\le v_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 + \max_{\|\mathbf{v}\|_\infty \le v_\infty} |F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\mathbf{v}| \\
&\le v_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 + v_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 \left\|\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}\right\|_\infty \\
&\le v_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 + v_\infty\epsilon_\infty \frac{\left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 \left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty}{1-\epsilon_\infty\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty} \\
&\le v_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 \left(1 + \frac{\epsilon_\infty\tau}{1-\epsilon_\infty\tau}\right) \\
&= \frac{v_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1}{1-\epsilon_\infty\tau} = f_{1,j}(\widehat{\mathbf{\Phi}})\,.
\end{aligned}$$

Next, recalling $\widehat{\mathbf{\Phi}}_{yu} = \widehat{\mathbf{\Phi}}_{yy}\widehat{\mathbf{G}}$, we have

$$\begin{aligned}
& w_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})\right\|_1 \\
&\le w_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1 + \max_{\|\mathbf{w}\|_\infty \le w_\infty} |F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\mathbf{w}| + \max_{\|\mathbf{w}\|_\infty \le w_\infty} |F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})\mathbf{w}| \\
&\le w_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1 + w_\infty\epsilon_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 + w_\infty\epsilon_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 \left\|\widehat{\mathbf{\Phi}}_{uy}(I-\mathbf{\Delta}\widehat{\mathbf{\Phi}}_{uy})^{-1}(\widehat{\mathbf{G}}+\mathbf{\Delta})\right\|_\infty \\
&\le w_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1 + w_\infty\epsilon_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 \left(1+\tau\frac{\left\|\widehat{\mathbf{G}}\right\|_\infty + \epsilon_\infty}{1-\epsilon_\infty\tau}\right) \\
&= w_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yu}\right\|_1 + w_\infty\epsilon_\infty \left\|F_{y,j}\widehat{\mathbf{\Phi}}_{yy}\right\|_1 \left(\frac{1+\tau\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\tau}\right) = f_{2,j}(\widehat{\mathbf{\Phi}})\,.
\end{aligned}$$

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

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

we have

$$\begin{aligned}
& w_\infty \left\|F_{u,j}(I-\widehat{\mathbf{\Phi}}_{uy}\mathbf{\Delta})^{-1}\widehat{\mathbf{\Phi}}_{uu}\right\|_1 \\
&\le w_\infty \left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uu}\right\|_1 + w_\infty\left\|\widehat{\mathbf{G}}\right\|_\infty \epsilon_\infty \frac{\left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right\|_1 \left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty}{1-\epsilon_\infty\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty} + w_\infty\epsilon_\infty \frac{\left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right\|_1}{1-\epsilon_\infty\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty} \\
&\le w_\infty\left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uu}\right\|_1 + w_\infty\epsilon_\infty \left\|F_{u,j}\widehat{\mathbf{\Phi}}_{uy}\right\|_1 \frac{1+\tau\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\tau} = f_{5,j}(\widehat{\mathbf{\Phi}})\,.
\end{aligned}$$

Similar computations allows one to derive the upperbounds for the remaining terms. ■

Lemma 4 exploits the upperbound $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \le \tau$ to take into account the worst-case increase in the values of the input and output signals due to uncertainty. One can interpret (37)-(38) as a kind of constraint tightening approach related to the one that is used in the robust learning-based MPC literature [38]. 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 (37)-(38) 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 $\left\|\widehat{\mathbf{\Phi}}_{uy}\right\|_\infty \le \tau$ softens while (37)-(38) tightens. It is therefore necessary to explicitly optimize over τ. We are ready to finalize the proof of Theorem 1

Proof of Theorem 1 Lemma 3 shows that the the cost of (5) upper-bounds $J(\mathbf{G},\mathbf{K}) = J(\mathbf{G},\widehat{\mathbf{\Phi}}_{uy}\widehat{\mathbf{\Phi}}_{yy}^{-1})$ for every feasible $\mathbf{K}$. Lemma 4 shows that (37)-(38) imply (39)-(40), which are equivalent to the safety constraints (21)-(22) for the real system.

Note that the optimization problem in (5) can be written as

$$\min_{\gamma\in[0,\epsilon_2^{-1}),\tau\in[0,\epsilon_\infty^{-1})} J_{robust}(\gamma,\tau)\,, \tag{41}$$

where the function $J_{robust}(\gamma,\tau)$ is the optimal objective of the inner minimization in (5). This inner optimization program is convex in $\widehat{\mathbf{\Phi}}$. ■

E Proof of Theorem 2

It is first necessary to characterize a feasible solution to problem (5), which we later exploit to establish our suboptimality bound. Such feasible solution is constructed as the optimal solution to a *ground-truth doubly-robust* (GTDR) optimization program fully stated in Lemma 5 below.

Lemma 5 (Feasible solution to problem (5)**)** *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 assume* $\alpha\in[\sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)},\epsilon_2^{-1})$*. Consider the following optimization problem and its optimal solution* $\mathbf{\Phi}^c$*:*

$$\mathbf{\Phi}^c\in\arg\min_{\mathbf{\Phi}} \left\|\begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu} & \mathbf{\Phi}_{yy}\mathbf{y}_0\\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu} & \mathbf{\Phi}_{uy}\mathbf{y}_0\end{bmatrix}\right\|_F \tag{42}$$

$$\begin{aligned}
\text{subject to}\ & \begin{bmatrix}I & -\mathbf{G}\end{bmatrix}\begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu}\\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix} = \begin{bmatrix}I & 0\end{bmatrix},\\
& \begin{bmatrix}\mathbf{\Phi}_{yy} & \mathbf{\Phi}_{yu}\\ \mathbf{\Phi}_{uy} & \mathbf{\Phi}_{uu}\end{bmatrix}\begin{bmatrix}-\mathbf{G}\\ I\end{bmatrix} = \begin{bmatrix}0\\ I\end{bmatrix},\\
& \left\|\mathbf{\Phi}_{uy}\right\|_2 \le \left\|\mathbf{\Phi}_{uy}^\star\right\|_2,\ \left\|\mathbf{\Phi}_{uy}\right\|_\infty \le \left\|\mathbf{\Phi}_{uy}^\star\right\|_\infty\,,\\
& \phi_{1,j}(\mathbf{\Phi})+\phi_{2,j}(\mathbf{\Phi})+\phi_{3,j}(\mathbf{\Phi}) \le \mathbf{b}_{y,j}\,, && (43)\\
& \phi_{4,j}(\mathbf{\Phi})+\phi_{5,j}(\mathbf{\Phi})+\phi_{6,j}(\mathbf{\Phi}) \le \mathbf{b}_{u,j}\,, && (44)\\
& \forall j=1,\dots,sN\,,\\
& \mathbf{\Phi}_{yy},\mathbf{\Phi}_{yu},\mathbf{\Phi}_{uy},\mathbf{\Phi}_{uu}\ \text{with causal sparsities}\,.
\end{aligned}$$

where

$$\phi_{1,j}(\mathbf{\Phi}) = \frac{v_\infty\left\|F_{y,j}\mathbf{\Phi}_{yy}\right\|_1}{1-2\zeta}\,,\quad \phi_{4,j}(\mathbf{\Phi}) = \frac{v_\infty\left\|F_{u,j}\mathbf{\Phi}_{uy}\right\|_1}{1-2\zeta}\,,\quad \phi_{2,j}(\mathbf{\Phi}) = w_\infty\left\|\begin{bmatrix}(F_{y,j}\mathbf{\Phi}_{yu})^\mathsf{T}\\ 2\frac{\epsilon_\infty+\zeta\|\widehat{\mathbf{G}}\|_\infty}{1-2\zeta}(F_{y,j}\mathbf{\Phi}_{yy})^\mathsf{T}\end{bmatrix}^\mathsf{T}\right\|_1\,,$$

$$\phi_{5,j}(\mathbf{\Phi}) = w_\infty\left\|\begin{bmatrix}(F_{u,j}\mathbf{\Phi}_{uu})^\mathsf{T}\\ 2\frac{\epsilon_\infty+\zeta\|\widehat{\mathbf{G}}\|_\infty}{1-2\zeta}(F_{u,j}\mathbf{\Phi}_{uy})^\mathsf{T}\end{bmatrix}^\mathsf{T}\right\|_1\,,\quad \phi_{3,j}(\mathbf{\Phi}) = F_{y,j}\mathbf{\Phi}_{yy}\widehat{\mathbf{y}}_0 + 2\frac{\epsilon_\infty+\zeta\left\|\widehat{\mathbf{y}}_0\right\|_\infty}{1-2\zeta}\left\|F_{y,j}\mathbf{\Phi}_{yy}\right\|_1\,,$$

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

Then, the following expressions

$$\begin{aligned}
&\widetilde{\mathbf{\Phi}}_{yy} = \mathbf{\Phi}_{yy}^c(I+\mathbf{\Delta}\mathbf{\Phi}_{uy}^c)^{-1},\ \widetilde{\mathbf{\Phi}}_{yu} = \mathbf{\Phi}_{yy}^c(I+\mathbf{\Delta}\mathbf{\Phi}_{uy}^c)^{-1}(\mathbf{G}-\mathbf{\Delta}),\\
&\widetilde{\mathbf{\Phi}}_{uy} = \mathbf{\Phi}_{uy}^c(I+\mathbf{\Delta}\mathbf{\Phi}_{uy}^c)^{-1},\ \widetilde{\mathbf{\Phi}}_{uu} = (I+\mathbf{\Phi}_{uy}^c\mathbf{\Delta})^{-1}\mathbf{\Phi}_{uu}^c,\\
&\widetilde{\gamma} = \frac{\sqrt{2}\eta}{\epsilon_2(1-\eta)},\qquad \widetilde{\tau} = \frac{\zeta}{\epsilon_\infty(1-\zeta)}\,,
\end{aligned} \tag{45}$$

provide a feasible solution to problem (5).

The achievability constraints of the GTDR program are based upon the ground-truth impulse response $\mathbf{G}$, and its safety constraints may be more stringent than those of problem (5), thus allowing to guarantee that $\mathbf{\Phi}^c$ is feasible for problem (5). Letting $\mathbf{K}^c$ denote the controller being optimal for the GTDR problem, we define the corresponding suboptimality gap with respect to ground-truth as

$$S(\epsilon_\infty) = \frac{J(\mathbf{G},\mathbf{K}^c)^2 - J(\mathbf{G},\mathbf{K}^\star)^2}{J(\mathbf{G},\mathbf{K}^\star)^2}\,, \tag{46}$$

where we have equivalently denoted $J^{\mathrm{GTDR}} = J(\mathbf{G}, \mathbf{K}^c)$ in the main body of this extended abstract to simplify the notation. We further note that if the estimation error ϵ_∞ is too large, the GTDR program (42) 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 ground-truth problem (3) subject to (1)-(2) does not activate the safety constraints (2), then the constraints of (42) remain inactive for small enough ϵ_∞. In such case we have that $S(\epsilon_\infty) = 0$.

Exploiting the feasible solution characterized by Lemma 5, we are ready to prove Theorem 2.

Proof of Theorem 2 By denoting as $\widehat{\mathbf{\Phi}}^\star$ the closed-loop responses obtained by applying $\widehat{\mathbf{K}}^\star$ to $\widehat{\mathbf{G}}$, we have by Lemma 3

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

where $\gamma^\star$ is optimal for (5). By Lemma 5, under the assumptions on η and ζ we have that $(\widetilde{\gamma}, \widetilde{\tau}, \widetilde{\mathbf{\Phi}})$ defined in (45) belongs to the feasible set of (5). Hence, by suboptimality of any feasible solution, one has

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

Using the definition of $\widetilde{\mathbf{\Phi}}$ from Lemma 5, we now relate the term

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

to $J^{\mathrm{GTDR}} = J(\mathbf{G}, \mathbf{K}^c)$. Define the quantities

$$\begin{aligned} M^c &= h(\epsilon_2,\alpha,\widehat{\mathbf{G}}) + h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^c_{uy}\right\|_2, \mathbf{G}) + h(\epsilon_2, \left\|\mathbf{\Phi}^c_{uy}\right\|_2, \mathbf{y}_0), \\ M^\star &= h(\epsilon_2,\alpha,\widehat{\mathbf{G}}) + h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^\star_{uy}\right\|_2, \mathbf{G}) + h(\epsilon_2, \left\|\mathbf{\Phi}^\star_{uy}\right\|_2, \mathbf{y}_0), \\ V^c &= h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^c_{uy}\right\|_2, \mathbf{y}_0), \\ V^\star &= h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0) + h(\epsilon_2, \left\|\mathbf{\Phi}^\star_{uy}\right\|_2, \mathbf{y}_0). \end{aligned}$$

By using analogous expressions as those derived in Lemma 5, we obtain

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

Thus, we have established the chain of inequalities

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

Next, notice that, by definition of the GTDR suboptimality gap (46), we have

$$J(\mathbf{G},\mathbf{K}^c)^2 = (S(\epsilon_\infty)+1)J(\mathbf{G},\mathbf{K}^\star)^2.$$

Recalling that $\eta < \frac{1}{5}$, $\left\|\mathbf{\Phi}^c_{uy}\right\|_2 \leq \left\|\mathbf{\Phi}^\star_{uy}\right\|_2$ and $\left\|\mathbf{\Phi}^c_{yy}\right\|_2 \leq \left\|\mathbf{\Phi}^\star_{yy}\right\|_2$, and noticing that if $M, V > 0$, then

$Ma^2 + Vb^2 \leq (M+V)(a^2+b^2)$, we can establish the following inequalities.

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

Last, we prove that $20\eta + 4(M^c+V^c) = \mathcal{O}\left(\epsilon_2\left\|\mathbf{\Phi}^\star_{uy}\right\|_2(\|\mathbf{G}\|_2^2+\|\mathbf{y}_0\|_2^2)\right)$. First, notice that $M^c+V^c \leq M^\star+V^\star$. By considering the expressions of $M^\star$ and $V^\star$, using $\alpha \leq 5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2$, $\eta < \frac{1}{5}$, $\left\|\widehat{\mathbf{G}}\right\|_2 \leq \|\mathbf{G}\|_2+\epsilon_2$ and $\|\widehat{\mathbf{y}}_0\|_2 \leq \|\mathbf{y}_0\|_2+\epsilon_2$, we deduce that

$$
\begin{aligned}
M^\star &= h(\epsilon_2,\alpha,\widehat{\mathbf{G}})+h(\epsilon_2,\alpha,\widehat{\mathbf{y}}_0)+h(\epsilon_2,\left\|\mathbf{\Phi}^\star_{uy}\right\|_2,\mathbf{G})+h(\epsilon_2,\left\|\mathbf{\Phi}^\star_{uy}\right\|_2,\mathbf{y}_0) \\
&\leq 2\Big[\epsilon_2^2(2+5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2\|\mathbf{G}\|_2)^2+2\epsilon_2\|\mathbf{G}\|_2(2+5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2\|\mathbf{G}\|_2)+\epsilon_2^2(2+5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2\|\mathbf{y}_0\|_2)^2+ \\
&\quad +2\epsilon_2\|\mathbf{y}_0\|_2(2+5\left\|\mathbf{\Phi}^\star_{uy}\right\|_2\|\mathbf{y}_0\|_2)\Big]+\mathcal{O}(\epsilon_2^2\left\|\mathbf{\Phi}^\star_{uy}\right\|_2(\|\mathbf{G}\|_2^2+\|\mathbf{y}_0\|_2^2)) \\
&= \mathcal{O}\left(\epsilon_2\left\|\mathbf{\Phi}^\star_{uy}\right\|_2(\|\mathbf{G}\|_2^2+\|\mathbf{y}_0\|_2^2)\right)\,,
\end{aligned}
$$

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

Proof of Lemma 5 It is easy to verify that $\widetilde{\mathbf{\Phi}}$ satisfies the constraints in (5); indeed, $\widetilde{\mathbf{\Phi}}$ comprises the closed-loop responses when we apply $\mathbf{K}^c$ to the estimated plant $\widehat{\mathbf{G}}$. Next, we have

$$
\begin{aligned}
\left\|\widetilde{\mathbf{\Phi}}_{uy}\right\|_2 &= \left\|\mathbf{\Phi}^c_{uy}(I+\mathbf{\Delta\Phi}^c_{uy})^{-1}\right\|_2 \leq \frac{\left\|\mathbf{\Phi}^c_{uy}\right\|_2}{1-\epsilon_2\left\|\mathbf{\Phi}^c_{uy}\right\|_2} \leq \sqrt{2}\frac{\left\|\mathbf{\Phi}^c_{uy}\right\|_2}{1-\epsilon_2\left\|\mathbf{\Phi}^c_{uy}\right\|_2} \\
&\leq \sqrt{2}\frac{\left\|\mathbf{\Phi}^\star_{uy}\right\|_2}{1-\epsilon_2\left\|\mathbf{\Phi}^\star_{uy}\right\|_2} = \sqrt{2}\frac{\eta}{\epsilon_2(1-\eta)} = \widetilde{\gamma}\,.
\end{aligned}
$$

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 a feasible value for γ in problem (5). Similarly,

$$
\left\|\widetilde{\mathbf{\Phi}}_{uy}\right\|_\infty = \left\|\mathbf{\Phi}^c_{uy}(I+\mathbf{\Delta\Phi}^c_{uy})^{-1}\right\|_\infty \leq \frac{\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}{1-\epsilon_\infty\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty} \leq \frac{\left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty}{1-\epsilon_\infty\left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty} = \frac{\zeta}{\epsilon_\infty(1-\zeta)} = \widetilde{\tau}\,.
$$

Since $\zeta < \frac{1}{2}$, then $\tilde{\tau} < \epsilon_\infty^{-1}$. Hence, it is a feasible value for τ in problem (5). It remains to show that $\widetilde{\mathbf{\Phi}}$ satisfies the safety constraints (37)-(38). We know that $\mathbf{\Phi}^c$ is feasible for (42), and hence $\phi_{1,j}(\mathbf{\Phi}^c)+\phi_{2,j}(\mathbf{\Phi}^c)+\phi_{3,j}(\mathbf{\Phi}^c) \leq \mathbf{b}_{y,j}$ and $\phi_{4,j}(\mathbf{\Phi}^c)+\phi_{5,j}(\mathbf{\Phi}^c)+\phi_{6,j}(\mathbf{\Phi}^c) \leq \mathbf{b}_{u,j}$. We conclude the proof by showing that $f_{i,j}(\widetilde{\mathbf{\Phi}}) \leq \phi_{i,j}(\mathbf{\Phi}^c)$ for every $i = 1,\dots,6$. We report the full derivations for the most informative terms.

$$
\begin{aligned}
f_{1,j}(\widetilde{\mathbf{\Phi}}) &= \frac{v_\infty\left\|F_{y,j}\left(\mathbf{\Phi}^c_{yy}-\mathbf{\Phi}^c_{yy}\mathbf{\Delta\Phi}^c_{uy}\left(I+\mathbf{\Delta\Phi}^c_{uy}\right)^{-1}\right)\right\|_1}{1-\epsilon_\infty\widetilde{\tau}} \\
&\leq \frac{v_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1+\frac{v_\infty\epsilon_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}{1-\epsilon_\infty\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}}{1-\epsilon_\infty\widetilde{\tau}} \\
&\leq \frac{v_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1+\frac{v_\infty\epsilon_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1\left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty}{1-\epsilon_\infty\left\|\mathbf{\Phi}^\star_{uy}\right\|_\infty}}{1-\epsilon_\infty\widetilde{\tau}} \\
&\leq \frac{v_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1}{1-2\zeta} = \phi_{1,j}(\mathbf{\Phi}^c)\,.
\end{aligned}
$$

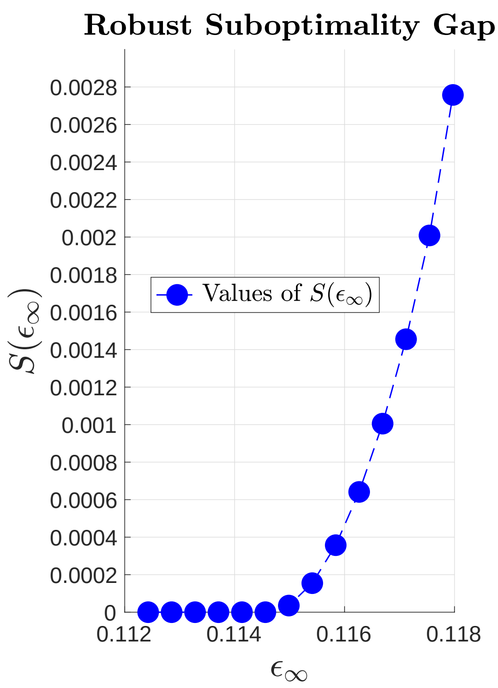


Figure 3: Robust suboptimality gap $S(\epsilon_\infty)$. This quantity can be interpreted as an indicator as to whether the guarantee (6) holds for a given ϵ_∞.

Similarly, it is easy to show that $f_{4,j}(\widetilde{\mathbf{\Phi}}) \leq \phi_{4,j}(\mathbf{\Phi}^c)$. Next, by remembering that

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

we have

$$\begin{aligned}
&f_{2,j}(\widetilde{\mathbf{\Phi}})\\
&\leq w_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}(\mathbf{G}-\mathbf{\Delta})\right\|_1 + w_\infty\epsilon_\infty \left\|F_{y,j}\mathbf{\Phi}^c_{yy}(I+\mathbf{\Delta}\mathbf{\Phi}^c_{uy})^{-1}\right\|_1 \left(\frac{1+\widetilde{\tau}\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\widetilde{\tau}}\right)\\
&\leq w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1 + w_\infty\epsilon_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1\left(1+\frac{\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}\right) + w_\infty\epsilon_\infty\frac{\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1\left(\frac{1+\widetilde{\tau}\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\epsilon_\infty\widetilde{\tau}}\right)}{1-\epsilon_\infty\left\|\mathbf{\Phi}^c_{uy}\right\|_\infty}\\
&\leq w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1 + w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1\left(\epsilon_\infty + \frac{\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\zeta} + \frac{\epsilon_\infty + \frac{\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty}{1-\zeta}}{\left(1-\frac{\zeta}{1-\zeta}\right)(1-\zeta)}\right)\\
&= w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1 + 2w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1\frac{(1-\zeta)\left(\epsilon_\infty+\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty\right)}{1-2\zeta}\\
&\leq w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yu}\right\|_1 + 2w_\infty\left\|F_{y,j}\mathbf{\Phi}^c_{yy}\right\|_1\frac{\left(\epsilon_\infty+\zeta\left\|\widehat{\mathbf{G}}\right\|_\infty\right)}{1-2\zeta}\\
&\leq \phi_{2,j}(\mathbf{\Phi}^c)\,.
\end{aligned}$$

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