Safe Output Feedback Improvement with Baselines

Yibo Wang, Ruoqi Zhang, Per Mattsson

***Abstract*—It is important to account for the presence of uncertainties when designing data-driven controllers. Min-max optimization is an often used approach for addressing this challenge in the way of minimizing a given design criterion over the set of worst-case uncertainties. Although this algorithm guarantees the safety of the system, the resulting control laws can be overly conservative. In order to mitigate this issue, a baseline regret cost is incorporated into the objective function so that the controller performance can be improved within a safe region. The design consists of two steps. First, an uncertainty set is constructed via system identification based on finite impulse response (FIR) models. Then, a control design criterion based on model reference control is used. The numerical examples show that the inclusion of baseline regret indeed improves the performance of the resulting controllers, and that the use of regularization techniques in the FIR-modeling can be beneficial.**

I. Introduction

System uncertainties is a significant concern, especially in a population of dynamical systems within the context of data-driven control. In the context of mass production, for instance, the same controller is typically used for all produced units. Consequently, it is crucial to design a controller that is sufficiently robust to handle a wide range of scenarios, such as varying environmental conditions and small perturbations. However, such a robust controller may be overly conservative. To address this issue, while maintaining the safety, [1] proposed an approach that starts from the a robust baseline controller and then improves this controller in a data-driven manner.

A. Problem Formulation

The problem is based on a general stable discrete-time linear systems,

$$y_t = G_\circ(q)u_t + v_t, \tag{1}$$

where u_t is the input, y_t is the output and v_t is white noise with zero mean and variance σ^2. $G_\circ(q)$ is an unknown transfer operator and q^{-1} is the backward shift operator, $q^{-1}u_t = u_{t-1}$. The aim is to design a linear output feedback controller, $u_t = C(q)\left(r_t - y_t\right)$, for tracking a given reference signal r_t, such that some performance criterion J is minimized.

B. Baseline Regret Optimization

Before introducing baseline regret optimization, one commonly used criteria for finding optimal controllers, known as min-max, is given by,

$$C \in \arg\min_{C\in\mathcal{C}}\max_{G\in\mathcal{G}} J(C,G), \tag{2}$$

This work was supported by the *Swedish Research Council (VR)* under contract 2023-04546.

Y. Wang, R. Zhang, P. Mattsson are with Department of Information Technology, Uppsala University, SE-75105 Uppsala, SWEDEN (e-mail: yibo.wang@it.uu.se, ruoqi.zhang@it.uu.se, per.mattsson@it.uu.se)

where $\mathcal{C}$ is the set of controllers considered, and $\mathcal{G}$ describes possible variations in the true system.

It is widely used due to its robust stabilization properties for uncertain systems [2], [3]. However, the resulting control policy can be very conservative, e.g., takes longer time to reach the final state, particularly when the uncertainty set $\mathcal{G}$ is large. Inspired by the developments in the reinforcement learning literature, where safe policy improvement is delineated by minimizing the negative regret with respect to the baseline policy over certain processes [4], the following baseline regret criterion equation is proposed,

$$C \in \arg\min_{C\in\mathcal{C}}\max_{G\in\mathcal{G}} J(C,G) - J\left(C_b, G\right). \tag{3}$$

This strategy ensures that the resulting controller is an improvement over the baseline C_b. The corresponding theoretical proof can be found in Appendix A.

It is then critical to define a set of uncertainties $\mathcal{G}$ and a corresponding set of feasible controllers $\mathcal{C}$. The least squares identification approach is first employed to estimate finite impulse response (FIR) models that characterize the uncertainty set, followed by the construction of a well-defined controller set. To formulate a convex optimization problem with respect to the controller C, a model reference criterion is utilized, where the objective is to find a controller C such that the closed-loop system resembles a reference model $W(q)$ as close as possible.

Furthermore, the scenario approach [5] is introduced in [1] to tackle the issue of the uncountable uncertainty set $\mathcal{G}$, allowing the problem to be simplified by drawing only a finite number of samples from the set $\mathcal{G}$.

C. Numerical Example

To evaluate the performance of the proposed strategy in comparison to the min-max benchmark, a numerical example with the use of a proportional controller is studied. The reference model is chosen so that a proportional controller with gain 0.5 is optimal for the true system.

The simulation results indicate that the traditional min-max optimization yields a relatively conservative controller under high uncertainty, with a gain of approximately 0.13, as shown in the left plot of Figure 1. In contrast, when the baseline regret cost is incorporated, the final controller increases to 0.25, which is closer to the desired reference value, as illustrated in the right plot of Figure 1. Moreover, as the size of training data increases, the uncertainty set accordingly shrinks, and the gap between 0.5 and optimized controllers become smaller regardless of the selected baseline controller. Nevertheless, the results obtained with the baseline regret criterion remain more aggressive than those using original min-max formulation within a safe region. The corresponding figure is not included here in order to conserve space.

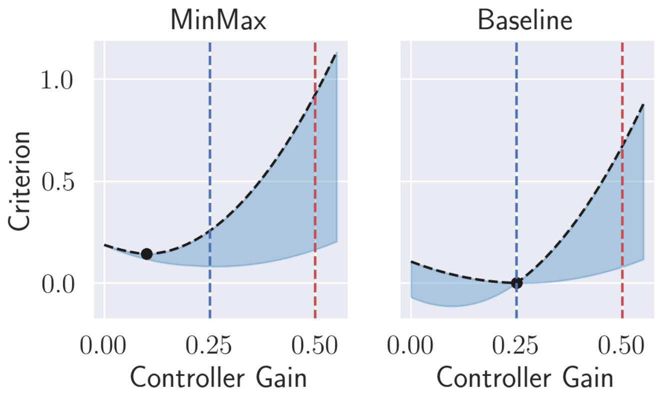

Fig. 1. Controller gain under different performance criteria. The left plot illustrates the result using the min-max criterion in (2), while the right plot shows the result using the proposed baseline regret criterion in (3). The *black dashed line* indicates the worst-case cost under model uncertainty (represented by shaded area), and the *dot* marks the solution. The *blue dashed line* denotes the baseline gain, and the *red dashed line* represents the optimal gain for the unknown system G_o.

II. Benefit of Regularization

The approach of baseline regret has demonstrated strong potential to improve the controller performance while maintaining safety, as discussed in the previous section. Nevertheless, there remains room for further enhancement in terms of robustness. One key limitation is that, even with the use of the scenario approach, the sampled distribution does not sufficiently capture the ground-truth system. This is possibly caused by the fact that the estimated covariance is not sufficiently large to capture the true system. The rest of this section studies a potential solution to this challenge.

A. Bayesian Regularization

To address the problem that the sampled distribution fails to include the criterion curve of the true system, we smooth the result of system identification step by embedding a regularization term. Several methods have been proposed for this purpose [6], [7]. In particular, [6] suggested a Bayesian kernel-based approach for linear controller design, which we adopt and integrate into the proposed control framework.

By treating the parameters as stochastic and introducing a prior for regularization, a Bayesian approach is applied to enhance the smoothness of impulse responses in FIR models. The problem now becomes,

$$\hat{g} = \arg\min_{g} \|Y - \Phi g\|^2 + \frac{\sigma^2}{\lambda} g^T K_{ij}^{-1} g,$$

where g is the true impulse response, $\hat{g}$ is its estimate, Φ and Y contain the input and output data, respectively, λ is a positive scalar, K is a regularization matrix defined by the class of the stable spline kernels. Additionally, the prior covariance is given by $\Sigma = \lambda K$, and the first-order stable spline kernel,

$$K_{ij} := \alpha^{\max(i,j)}, \quad 0 \leq \alpha < 1, \tag{4}$$

is used in this study. Note that both λ and α can be tuned based on specific requirement. The former one determines the weight of the penalty term, while the latter hyperparameter controls the rate at which the impulse response decays to zero. The resulting estimated impulse response, along with the corresponding posterior covariance, are therefore given by,

$$\begin{aligned} \hat{g} &= \left(\Phi^T\Phi + \frac{\sigma_v^2}{\lambda} K^{-1}\right)^{-1} \Phi^T Y, \\ \hat{\Sigma} &= \sigma^2 \left(\Phi^T\Phi + \frac{{\sigma_v^2}^2}{\lambda} K^{-1}\right)^{-1}. \end{aligned} \tag{5}$$

Based on the estimated plant models and uncertainties obtained via (5), the subsequent steps involve sampling a finite set of impulse responses and optimizing the controller using the proposed baseline-regret criterion.

In terms of the final optimal controller, the solutions computed with the regularization technique outperform the result of the original formulation. This conclusion is drawn by noticing that the value of the final optimal controller is closer the desired control gain used in the reference model. This improvement is particularly evident when the high uncertainty exists. Moreover, the proposed baseline regret cost introduces flexibility in the final controller design by allowing adjustment of the baseline controller value. This ensures that the resulting controller not only improves performance but can also be tailored to specific application requirements. To be more specific, when a more aggressive baseline gain is selected, the resulting gain exceeds the reference gain within a reasonable region.

Appendix A
Proof of Performance Guarantee for Baseline Regret Criterion

Proposition 1: If $G_\circ \in \mathcal{G}$, and the solution to (3) is $\hat{C}^*$, then $J\left(\hat{C}^*, G_\circ\right) \leq J\left(C_b, G_\circ\right)$.

Proof: If the true system $G_\circ \in \mathcal{G}$, it can be guaranteed that $J\left(\hat{C}^*, G_\circ\right) - J\left(C_b, G_\circ\right) \leq \max_{G\in\mathcal{G}} J\left(\hat{C}^*, G\right) - J\left(C_b, G\right) \leq J\left(C_b, G\right) - J\left(C_b, G\right) = 0$

References

[1] R. Zhang, P. Mattsson, and D. Zachariah, "Safe output feedback improvement with baselines," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, pp. 1899–1904, 2024.

[2] D. M. Raimondo, D. Limon, M. Lazar, L. Magni, and E. F. ndez Camacho, "Min-max model predictive control of nonlinear systems: A unifying overview on stability," *European Journal of Control*, vol. 15, no. 1, pp. 5–21, 2009.

[3] Y. Xie, J. Berberich, and F. Allgöwer, "Data-driven min-max mpc for linear systems," in *2024 American Control Conference (ACC)*, pp. 3184–3189, 2024.

[4] M. Petrik, Y. Chow, and M. Ghavamzadeh, "Safe policy improvement by minimizing robust baseline regret," 2016.

[5] G. C. Calafiore and M. C. Campi, "The scenario approach to robust control design," *IEEE Transactions on Automatic Control*, vol. 51, no. 5, pp. 742–753, 2006.

[6] A. Scampicchio, A. Chiuso, S. Formentin, and G. Pillonetto, "Bayesian kernel-based linear control design," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 822–827, 2019.

[7] G. Pillonetto, T. Chen, A. Chiuso, G. Nicolao, and L. Ljung, *Regularized System Identification: Learning Dynamic Models from Data*. 01 2022.