

On the Regret of Recursive Methods for Discrete-Time Adaptive Control with Matched Uncertainty

Aren Karapetyan, Efe C. Balta, Anastasios Tsiamis, Andrea Iannelli, and John Lygeros

Abstract—Continuous-time adaptive controllers for systems with a matched uncertainty often comprise an online parameter estimator and a corresponding parameterized controller to cancel the uncertainty. However, such methods are often impossible to implement directly, as they depend on an unobserved estimation error. We consider the equivalent discrete-time setting with a causal information structure, and propose a novel, online proximal point method-based adaptive controller, that under a sufficient excitation (SE) condition is asymptotically stable and achieves finite regret, scaling only with the time required to fulfill the SE. We show the same also for the widely-used recursive least squares with exponential forgetting controller under a stronger persistence of excitation condition.

I. INTRODUCTION

Adaptive control studies controllers that can adapt to or learn unmodelled changes in the dynamics [1], while often guaranteeing perfect asymptotic tracking and/or parameter convergence [2], [3]. The theory of online optimization was established concurrently with the developments in adaptive control, with the aim to minimize an a priori unknown, sequentially revealed cost [4]. The connections between the two have recently been highlighted [5], [6], showing that adaptive controllers are closely linked to an online cost optimization problem and are often proportional to an online estimation cost gradient. In this work, we consider such controllers that aim to minimize a cost function while controlling the system, similar to the goal-oriented control formulation in [7]. Since asymptotic stability alone does not automatically imply performance guarantees on the cost [8], [9], we also provide finite-time guarantees by characterizing the *regret* [4] of the controller: a concept borrowed from online learning to quantify the additional cost incurred by the controller over a horizon due to partial model knowledge.

We consider systems with matched uncertainty [2], where the unknown dynamics is parameterized by an unknown parameter vector and a known, state-dependant feature matrix. In this setting, the uncertainty can be exactly canceled by

the control input given knowledge of the true parameter. We aim to find causal adaptive controllers that estimate this parameter recursively while stabilizing the system and minimizing regret. Most continuous-time adaptive controllers in this setting utilize Lyapunov theory and provide only asymptotic stability guarantees, see for example [1], [2], [3]. Using the same tools, recently, a finite $\mathcal{O}(1)$ regret bound was shown in [10] for linear systems, and in [11] for nonlinear time-varying systems. In [11], the authors also extend the results to a discrete-time setting, achieving a $o(T)$, sublinear-in-time, bound with a causal formulation, and a $\mathcal{O}(1)$ with a non-causal one. In contrast to our proposed method, the Lyapunov-based adaptive controllers in [10], [11] are not realizable in discrete time, as the parameter update depends on the current, unobserved error, making it non-causal. The online gradient descent-based algorithm in [11], although causal, achieves $o(T)$ regret even in the deterministic case. The difficulty is inherent in the discrete-time formulation of the problem, which introduces a one-step delay in the error observation. In discrete time, the convenient Lyapunov function cancellation [2] no longer holds and discrete-time Lyapunov functions are known to be harder to construct [12].

The recursive learning algorithm for the adaptive controller can be set up to relate to an online or “running” cost. This setting is also studied in the time-varying optimization literature providing asymptotic tracking and convergence guarantees for online costs, e.g., [13], [14]. However, in this line of work, the problem setup considers no underlying dynamics for the state, and the regret is generally not studied.

We design causal, uncertainty-matching, recursive methods for discrete-time, nonlinear, time-varying systems. Our contributions are threefold. Firstly, we introduce a recursive proximal learning (RPL) parameter estimation algorithm based on the proximal point method [15]. Under a weak notion of persistence of excitation (PE), which we refer to as sufficient excitation (SE) as in the continuous-time adaptive control literature [16], RPL produces estimates of the unknown parameters that are contractive with respect to the true one. Secondly, we show that RPL-based adaptive control achieves finite regret scaling with the time required to achieve the SE condition. By contrast, in [11] finite regret is only achieved through access to a Lyapunov function or by a non-causal controller. Finally, we analyze the regret of the recursive least-squares with forgetting factor (RLSFF) [17]. We show that while both RPL and RLSFF achieve finite regret with similar bounds, RLSFF requires a stronger PE condition. We demonstrate the algorithms on a discrete-time model reference adaptive control (MRAC) [18] example.

This work has been supported by the Swiss National Science Foundation under NCCR Automation (grant agreement 51NF40_180545), the European Research Council under the ERC Advanced grant agreement 787845 (OCAL) and by the German Research Foundation (DFG) under Germany’s Excellence Strategy - EXC 2075 - 390740016.

A. Karapetyan, A. Tsiamis and J. Lygeros are with the Automatic Control Laboratory, Swiss Federal Institute of Technology (ETH Zürich), 8092 Zürich, Switzerland (E-mails: akarapetyan, atsiamis, jlygeros@ethz.ch).

E. C. Balta is with the Control and Automation Group, inspire AG, 8005 Zürich, Switzerland, and with the Automatic Control Laboratory efe.balta@inspire.ch.

A. Iannelli is with the Institute for Systems Theory and Automatic Control, University of Stuttgart, Stuttgart 70569, Germany (E-mail: andrea.iannelli@ist.uni-stuttgart.de).

Notation: The sets of positive real numbers, positive integers, and non-negative integers are denoted by \mathbb{R}_+ , \mathbb{N}_+ and \mathbb{N} , respectively. For a given vector x , its Euclidean norm is denoted by $\|x\|$. The spectral norm of a square matrix W is denoted by $\|W\|$, and the largest and smallest singular values and eigenvalues by $\sigma_{min}(W)$, $\sigma_{max}(W)$, and $\lambda_{min}(W)$, $\lambda_{max}(W)$ respectively. For a positive scalar $r \in \mathbb{R}_+$, we define the closed ball around the origin as $\mathcal{B}_r := \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$.

II. PROBLEM STATEMENT

We consider nonlinear dynamical systems with matched uncertainty of the form

$$x_{k+1} = f_k(x_k) + B_k(x_k)(u_k - \alpha_k(x_k)), \quad (1)$$

where $f : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ models the known nominal dynamics and satisfies $f_k(0) = 0$ for all $k \in \mathbb{N}$, $B_k : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$ is a known input matrix, and $\alpha_k : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ is an *unknown* mapping to the input space, capturing the uncertain dynamics of the model. We assume that α_k can be exactly parameterized by possibly nonlinear, but known and bounded basis matrices, $\phi_k : \mathbb{N} \times \mathbb{R}^n \rightarrow \mathbb{R}^{p \times m}$, that map the state to a feature in $\mathbb{R}^{p \times m}$. The uncertainty is then given by $\alpha_k(x_k) = \phi_k^\top(x_k)\theta^*$, where $\theta^* \in \mathbb{R}^p$ is an unknown, constant parameter vector¹. We also assume $\|B_k(x_k)\phi_k^\top(x_k)\| \leq b$ for all $k \in \mathbb{N}$, $x \in \mathbb{R}^n$ and some $b \in \mathbb{R}_+$. The tracking problem under partial model knowledge can also be cast into this framework in the context of MRAC, as shown in Section V.

We characterize the performance of a given controller for system (1) over a time horizon T by comparing it to a benchmark control signal $u^* = [u_0^{*\top} \dots u_{T-1}^{*\top}]^\top$ that cancels the uncertainty perfectly and is defined below. We use regret to quantify the controller performance

$$\mathcal{R}_T(u) = \sum_{k=0}^{T-1} c(x_k) - c(x_k^*), \quad (2)$$

where $u = [u_0^\top \dots u_{T-1}^\top]^\top$ denotes the control signal of a given controller, x_k and x_k^* are state sequences due to the control signals u and u^* , respectively, and $c : \mathbb{R}^n \rightarrow \mathbb{R}$ is a stage cost. We restrict our analysis to locally Lipschitz continuous costs and assume for any $R \in \mathbb{R}_+$, there exists a $L_c \in \mathbb{R}_+$, such that for all $x, y \in \mathcal{B}_R$,

$$\|c(x) - c(y)\| \leq L_c \|x - y\|. \quad (3)$$

Assumption 1: (System Dynamics) The nominal dynamics $x_{k+1} = f_k(x_k)$ are uniformly, globally E - δ -ISS [19]. In other words, there exist $c_0, c_w, r_w \in \mathbb{R}_+$ and $\rho \in (0, 1)$, such that for any $x_0, y_0 \in \mathbb{R}^n$, $w_k \in \mathcal{B}_{r_w}$, and all $k \in \mathbb{N}$ the perturbed dynamics $y_{k+1} = f_k(y_k) + w_k$ satisfy

$$\|x_k - y_k\| \leq c_0 \rho^k \|x_0 - y_0\| + c_w \sum_{i=0}^{k-1} \rho^{k-i-1} \|w_i\|.$$

¹We drop the explicit state dependence on state for $B_k := B_k(x_k)$ and $\phi_k := \phi_k(x_k)$, wherever required for readability.

The E - δ -ISS property characterizes only the known and autonomous part of the dynamics. It is known that if f_k is smooth in the state and is contractive, then it is also E - δ -ISS [20], [11]. If f_k is not smooth, then under milder conditions, we show in [21] that exponential stability implies E - δ -ISS.

A. Online Estimation

We consider adaptive control laws that produce control inputs of the form $u_k = \phi_k^\top(x_k)\theta_k$ where $\theta_k \in \mathbb{R}^p$ is an online estimate of θ^* at time k . The closed-loop dynamics for a controller of this form is given by

$$x_{k+1} = f_k(x_k) + \underbrace{B_k(x_k)\phi_k^\top(x_k)}_{w_k(x_k)}(\theta_k - \theta^*), \quad (4)$$

where $w_k(x_k)$ can be thought of as a state-dependant artificial disturbance introduced due to an inexact uncertainty matching. One can also consider a benchmark counterfactual control input $u_k^* := \phi_k^\top(x_k^*)\theta^*$ and the corresponding signal u^* that perfectly cancels out the uncertainty in (1) leading to a benchmark state evolution

$$x_{k+1}^* = f_k(x_k^*). \quad (5)$$

Note that u^* is a counterfactual policy that is not realizable, as θ^* is unknown. To simplify the analysis, we take $x_0 = x_0^*$.

A standard setup in adaptive control design is to produce estimates θ_k that minimize the magnitude of $w_k(x_k)$, effectively steering the state of (4) to that of (5) thanks to the E - δ -ISS property. This is often done by setting up a least squares estimation problem, minimizing the following online estimation cost for all $k \in \mathbb{N}_+$

$$\theta_k = \arg \min_{\theta} h_{k-1}(\theta), \quad (6)$$

$$h_{k-1}(\theta) := \frac{1}{2} \sum_{i=0}^{k-1} \|B_i \phi_i^\top (\theta - \theta^*)\|^2 = \frac{1}{2} \|\Phi_k \theta - Y_k\|^2,$$

$$Y_k = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{k-1} \end{bmatrix}, \quad \Phi_k = \begin{bmatrix} B_0 \phi_0^\top \\ B_1 \phi_1^\top \\ \vdots \\ B_{k-1} \phi_{k-1}^\top \end{bmatrix},$$

$$y_i = x_{i+1} - f_i(x_i) - B_i \phi_i^\top \theta_i = B_i \phi_i^\top \theta^*, \quad \forall i \in \mathbb{N}.$$

Note that at time k , the latest available online estimation cost is h_{k-1} , as y_k is not yet available. Solving (6) directly requires computing an expensive matrix inverse at each time step and maintaining an increasing memory. In this work, we study two recursive methods that alleviate this issue, both optimizing an online cost related to (6). In particular, we introduce and analyze a recursive proximal learning method in Section III, and also study the well-established recursive least squares with forgetting factor [17] in Section IV.

B. Regret

Speed or velocity-gradient controllers are introduced in [7] to solve the continuous-time equivalent of the online estimation problem (6) in the framework of (integral) goal-oriented control. It is shown that these methods achieve

finite estimation error in the limit, under the assumption that the controller at time t can depend on the estimation error $\theta_t - \theta^*$. Recent works [10], [11] noted that if an exponential Lyapunov function exists for the nominal system, a speed-gradient descent-based uncertainty matching controller achieves finite quadratic cost on the state. In particular, the suboptimal closed-loop state trajectory x_t satisfies

$$J_{\text{cont}}(x_0, u) := \lim_{T \rightarrow \infty} \int_0^T \|x_t\|^2 dt = \mathcal{O}(1).$$

Moreover, the existence of the exponential Lyapunov function also implies that $J_{\text{cont}}(x_0, u^*)$ for the optimal state evolution is finite. Defining continuous-time regret as $\mathcal{R}_{\text{cont}}(u) := J_{\text{cont}}(x_0, u) - J_{\text{cont}}(x_0, u^*)$, an $\mathcal{O}(1)$ bound for the extra cost accumulated by the adaptive control law as compared to the benchmark follows directly. In [11] a discrete-time version of the speed-gradient algorithm is introduced with a finite cost bound, but with the same assumption that at time k the error $\theta_k - \theta^*$ is available, making the controller non-causal. Without this assumption, the authors achieve $o(T)$ sublinear regret using an online learning toolkit. In this work, we consider causal controllers that minimize (6) at time k with access only to y_{k-1} . To characterize the performance of such an adaptive, uncertainty-matching controller, we analyze the regret (2).

III. RECURSIVE PROXIMAL LEARNING

The recursive proximal learning algorithm corresponds to the ε -scaled proximal operator of (6) evaluated at the *previous* estimate

$$\begin{aligned} \theta_k &= \text{prox}_{\varepsilon h_{k-1}}(\theta_{k-1}) \\ &= \arg \min_{\theta} \left(h_{k-1}(\theta) + \frac{\varepsilon}{2} \|\theta - \theta_{k-1}\|^2 \right), \end{aligned} \quad (7)$$

for some $\theta_0 \in \mathbb{R}^p$ and $\varepsilon \in \mathbb{R}_+$. From the online learning perspective [4], θ_k can be thought of as the minimizer of the latest available cost g_{k-1} at time k , where we define

$$g_{k-1}(\theta) := h_{k-1}(\theta) + \frac{\varepsilon}{2} \|\theta - \theta_{k-1}\|^2. \quad (8)$$

While a closed-form solution to (7) exists, it requires memory increasing with time. The following equivalent set of updates to (7) provides a recursive implementation. In particular, for all $k \in \mathbb{N}$, the following is equivalent to (7)

$$\begin{aligned} P_{k+1}^{-1} &= P_k^{-1} + \phi_k B_k^\top B_k \phi_k^\top, \\ H_{k+1} &= H_k + \phi_k B_k^\top B_k \phi_k^\top \\ s_{k+1} &= s_k + \phi_k B_k^\top y_k \\ \theta_{k+1} &= \theta_k - P_{k+1} (H_{k+1} \theta_k - s_{k+1}), \end{aligned} \quad (9)$$

initializing with $H_0 = \mathbf{0}_{p \times p}$, $s_0 = \mathbf{0}_{p \times 1}$ and $P_0^{-1} = \varepsilon I_p$.

A. Contraction in Parameter Space

We define SE as follows.

Definition 1: A matrix signal $\phi_k : \mathbb{N} \rightarrow \mathbb{R}^{p \times n}$ is called sufficiently exciting if there exist $T_s \in \mathbb{N}$, and $\delta \in \mathbb{R}_+$ such that

$$0 < \delta I_p \leq \sum_{i=0}^{T_s} \phi_i \phi_i^\top.$$

We show below that under a SE assumption, the RPL update (7) enjoys a contraction property on the parameter; moreover, under an additional boundedness assumption, the lifted input vector $\Phi_k \theta_k$ is also contractive.

Assumption 2: There exists a $\beta \in \mathbb{R}_+$ such that $\lim_{T \rightarrow \infty} \sum_{i=0}^T \phi_i B_i^\top B_i \phi_i^\top \leq \beta I_p$

Lemma 1: Assume $B_k \phi_k^\top$ is sufficiently exciting, then the RPL estimate (7) satisfies

$$\begin{aligned} \|\theta_k - \theta^*\| &\leq \eta \|\theta_{k-1} - \theta^*\|, \quad \forall k \geq T_s \\ \|\theta_k - \theta^*\| &\leq \|\theta_{k-1} - \theta^*\|, \quad \forall 0 < k < T_s \end{aligned} \quad (10)$$

where $\eta = \frac{\varepsilon}{\delta + \varepsilon} \in (0, 1)$. Moreover, if Assumption 2 holds, and $\varepsilon < \frac{\delta \sqrt{\delta}}{\sqrt{\beta} - \sqrt{\delta}}$ then there exists a $\gamma \in (0, 1)$ such that

$$\begin{aligned} \|\Phi_{k+1}(\theta_k - \theta^*)\| &\leq \gamma \|\Phi_k(\theta_{k-1} - \theta^*)\|, \quad \forall k \geq T_s, \\ \|\Phi_{k+1}(\theta_k - \theta^*)\| &\leq \sqrt{\beta} \|\theta_0 - \theta^*\|, \quad \forall 0 \leq k < T_s. \end{aligned}$$

Proof: The proximal update (7) can equivalently be represented by a single online Newton step update [22] on (8). At time k , it is equivalently represented by

$$\theta_k = \theta_{k-1} - \left(\nabla^2 g_{k-1}(\theta) \Big|_{\theta=\theta_{k-1}} \right)^{-1} \nabla g_{k-1}(\theta) \Big|_{\theta=\theta_{k-1}}.$$

Subtracting the true θ^* from both sides, denoting $\tilde{\theta}_k := \theta_k - \theta^*$, noting that $Y_k = \Phi_k \theta^*$ and recalling the definition of h_{k-1} from (6)

$$\begin{aligned} \tilde{\theta}_k &= \tilde{\theta}_{k-1} - (\Phi_k^\top \Phi_k + \varepsilon I_p)^{-1} \Phi_k^\top (\Phi_k \theta_{k-1} - Y_k) \\ &= \tilde{\theta}_{k-1} - (\Phi_k^\top \Phi_k + \varepsilon I_p)^{-1} \Phi_k^\top \Phi_k \tilde{\theta}_{k-1} \\ &= \left(I_p - (\Phi_k^\top \Phi_k + \varepsilon I_p)^{-1} \Phi_k^\top \Phi_k \right) \tilde{\theta}_{k-1} \\ &:= (I_p - M_k) \tilde{\theta}_{k-1}, \end{aligned}$$

where $M_k = (\Phi_k^\top \Phi_k + \varepsilon I_p)^{-1} \Phi_k^\top \Phi_k$. Using the orthonormality property of singular value decomposition of symmetric matrices and the sufficient excitation of $B_k \phi_k$, for all $k \geq T_s$

$$\begin{aligned} \sigma_{\min}(M_k) &\geq \frac{\delta}{\varepsilon + \delta} \in (0, 1), \\ \eta_k &:= \|I_p - M_k\| \leq \frac{\varepsilon}{\delta + \varepsilon} \in (0, 1), \end{aligned}$$

obtaining the contraction result for parameter since $\sigma_{\max}(M_k) \in (0, 1)$. For all $0 < k < T_s$, the SE condition is not fulfilled, hence $\sigma_{\min}(M_k) = 0$ and $\|\tilde{\theta}_k\| \leq \|\tilde{\theta}_{k-1}\|$.

To prove the second part, note that for a full column rank matrix A , $\|A \tilde{\theta}_{k+1}\| = \|\tilde{\theta}_{k+1}\|_{A^\top A}$, and $\lambda_{\min}(A^\top A) = \sigma_{\min}(A^\top A)$, and likewise for λ_{\max} . Then, for all $k \geq T_s$

$$\begin{aligned} \sigma_{\min} \left((\Phi_{k+1}^\top \Phi_{k+1})^{-1} \right) \|\tilde{\theta}_k\|_{\Phi_{k+1}^\top \Phi_{k+1}}^2 &\leq \|\tilde{\theta}_k\|^2 \\ &\leq \eta^2 \|\tilde{\theta}_{k-1}\|^2 \leq \eta^2 \sigma_{\max} \left((\Phi_k^\top \Phi_k)^{-1} \right) \|\tilde{\theta}_{k-1}\|_{\Phi_k^\top \Phi_k}^2. \end{aligned}$$

It follows that

$$\|\Phi_{k+1} \tilde{\theta}_k\| \leq \eta \sqrt{\frac{\sigma_{\max}(\Phi_{k+1}^\top \Phi_{k+1})}{\sigma_{\min}(\Phi_k^\top \Phi_k)}} \|\Phi_k \tilde{\theta}_{k-1}\|, \quad (11)$$

where we used the fact that $\sigma_{max}((A^\top A)^{-1}) = 1/\sigma_{min}(A^\top A)$ and $\sigma_{min}((A^\top A)^{-1}) = 1/\sigma_{max}(A^\top A)$. The proof is completed by noting that for all $k \geq T_s$

$$\gamma := \eta \sqrt{\frac{\sigma_{max}(\Phi_{k+1}^\top \Phi_{k+1})}{\sigma_{min}(\Phi_k^\top \Phi_k)}} \leq \frac{\varepsilon \sqrt{\beta}}{\varepsilon \sqrt{\delta} + \delta \sqrt{\delta}} < 1,$$

given $\varepsilon < \frac{\delta \sqrt{\delta}}{\sqrt{\beta} - \sqrt{\delta}}$. Finally, for all $k \in \mathbb{N}$, $\|\Phi_{k+1} \tilde{\theta}_k\| \leq \|\Phi_{k+1}\| \|\tilde{\theta}_k\| \leq \sqrt{\beta} \|\tilde{\theta}_0\|$. ■

B. Closed-loop Analysis

Consider the suboptimal state evolution (4) under the RPL controller $u_k^{\text{RPL}} = \phi_k^\top(x_k) \theta_k$, with θ_k given in (7). We can represent it equivalently as

$$x_{k+1} = f_k(x_k) + \underbrace{S_k (\Phi_{k+1} \theta_k - \Phi_{k+1} \theta^*)}_{w_k(x_k)}, \quad (12)$$

where $S_k := [\mathbf{0} \ \dots \ \mathbf{0} \ I_n] \in \mathbb{R}^{n \times n(k+1)}$ is a selector matrix that selects the last component of $\Phi_{k+1}(\theta_k - \theta^*)$, corresponding to $B_k \phi_k^\top(\theta_k - \theta^*)$. Similarly, the optimal state evolution under u^* is given by

$$x_{k+1}^* = f_k(x_k^*) + \underbrace{S_k (\Phi_{k+1}^* \theta^* - \Phi_{k+1}^* \theta^*)}_{\mathbf{0}}, \quad (13)$$

where $\Phi_{k+1}^* := \Phi(x_{[0, \dots, k]}^*)$. In the following theorem, we characterize the closed-loop system (12) in terms of asymptotic stability and regret. In particular, we show that the regret of the RPL controller scales with $\mathcal{O}(\|\tilde{\theta}_0\| T_s)$.

Theorem 1: Let Assumption 1 hold, and $B_k \phi_k^\top(x_k)$ be sufficiently exciting under the RPL adaptive controller, then this controller

- (i) renders the closed-loop system asymptotically stable,
- (ii) achieves finite regret

$$\mathcal{R}_T(u^{\text{RPL}}) \leq c_w b L_c \|\tilde{\theta}_0\| \left(\frac{T_s}{1-\rho} + \frac{\rho^T + (1-\eta)\rho + \eta}{(1-\rho)^2(1-\eta)} \right),$$

for the η established in Lemma 1. If in addition, $B_k \phi_k$ also satisfy Assumption 2 and $\varepsilon < \frac{\delta \sqrt{\delta}}{\sqrt{\beta} - \sqrt{\delta}}$, then

$$\mathcal{R}_T(u^{\text{RPL}}) \leq c_w c_p L_c \|\tilde{\theta}_0\| \left(\frac{T_s}{1-\rho} + \frac{\rho^T + (1-\gamma)\rho + \gamma}{(1-\rho)^2(1-\gamma)} \right),$$

where $c_p := \|\Phi_{T_s}\| \leq \sqrt{\beta}$ and γ is defined in Lemma 1.

Proof: To show (i), recall the definition of E - δ -ISS, and the fact that $f_k(0) = 0$ for all $k \in \mathbb{N}$, then consider the suboptimal state evolution of the nonlinear system under the RPL adaptive controller (12) for all $k \in \mathbb{N}_+$, recalling that $\tilde{\theta}_k = \theta_k - \theta^*$

$$\begin{aligned} \|x_k\| &\leq c_0 \rho^k \|x_0\| + c_w \sum_{i=0}^{k-1} \rho^{k-i-1} \|B_i \phi_i \tilde{\theta}_i\| \\ &< c_0 \|x_0\| + \frac{c_w b \|\tilde{\theta}_0\|}{1-\rho}, \end{aligned} \quad (14)$$

where we use the submultiplicativity property of the norms, the bound on $B_k \phi_k$ and non-expansivity of $\|\tilde{\theta}_k\|$ from

Lemma 1. The asymptotic stability follows from the input-to-state stability [23] of the closed-loop state in (14): by Lemma 1, $\lim_{k \rightarrow \infty} \|\tilde{\theta}_k\| = 0$ implies that $\lim_{k \rightarrow \infty} \|x_k\| = 0$.

Next, we show (ii), by first noting that since x_k and x_k^* are bounded from (14), it follows from (3)

$$\mathcal{R}_T(u^{\text{RPL}}) \leq L_c \underbrace{\sum_{k=T_s+1}^{T-1} \|x_k - x_k^*\|}_{\text{I}} + L_c \underbrace{\sum_{k=0}^{T_s} \|x_k - x_k^*\|}_{\text{II}},$$

where II represents the cost in the initial T_s -long non-PE phase, and I the cost improvement afterwards.

Using the E - δ -ISS condition on the nominal system and noting that x_k and x_k^* , given by (12) and (13), respectively, are equal in the first time step, for all $k > T_s$

$$\begin{aligned} \|x_k - x_k^*\| &\leq c_w b \sum_{i=0}^{k-1} \rho^{k-i-1} \|\tilde{\theta}_i\| \\ &\leq c_w b \sum_{i=1}^{T_s} \rho^{k-i} \|\tilde{\theta}_{i-1}\| + c_w b \sum_{i=T_s}^{k-1} \rho^{k-i-1} \|\tilde{\theta}_i\| \\ &\leq c_w b \rho^{k-T_s} \|\tilde{\theta}_0\| \sum_{i=1}^{T_s} \rho^{T_s-i} + c_w b \sum_{i=T_s}^{k-1} \rho^{k-i-1} \|\tilde{\theta}_i\| \\ &\leq \rho c_w c_{T_s} b \|\tilde{\theta}_0\| \rho^{k-T_s} + c_w b \|\tilde{\theta}_0\| \sum_{i=T_s}^{k-1} \rho^{k-i-1} \eta^{i-T_s+1}, \end{aligned} \quad (15)$$

where, defining $c_{T_s} = \frac{1-\rho^{T_s}}{1-\rho}$, the second inequality follows from the boundedness of $B_k \phi_k$, the third from the non-expansivity of $\|\tilde{\theta}_k\|$ from Lemma 1 and rearrangement of the constants, and the last one from the contraction of $\|\tilde{\theta}_k\|$. For $0 < k \leq T_s$, we may only use the non-expansive result from Lemma 1 and hence

$$\|x_k - x_k^*\| \leq c_w b \sum_{i=1}^k \rho^{k-i} \|\tilde{\theta}_{i-1}\| \leq c_w b \|\tilde{\theta}_0\| \sum_{i=1}^k \rho^{k-i}. \quad (16)$$

Using the Cauchy Product inequality defined for two finite series $\{a_i\}_{i=1}^T$ and $\{b_i\}_{i=1}^T$

$$\sum_{i=0}^T \left| \sum_{j=0}^i a_j b_{i-j} \right| \leq \left(\sum_{i=0}^T |a_i| \right) \left(\sum_{j=0}^T |b_j| \right),$$

and summing (16) and (15) over $T_s - 1$ and $T - T_s - 2$, respectively, we get

$$\begin{aligned} \text{II} &\leq c_w b L_c \|\tilde{\theta}_0\| \sum_{k=0}^{T_s-1} \sum_{i=0}^k \rho^{k-i} \leq c_w c_{T_s} T_s b L_c \|\tilde{\theta}_0\| \\ \text{I} &\leq c_w b L_c \|\tilde{\theta}_0\| \left(\rho c_{T_s} \sum_{k=1}^{T-T_s-1} \rho^k + \sum_{k=0}^{T-T_s-2} \sum_{i=0}^k \rho^{k-i} \eta^{i+1} \right) \\ &\leq \rho c_w b L_c \|\tilde{\theta}_0\| \frac{\rho^T + \rho}{(1-\rho)^2} \\ &\quad + c_w b \eta L_c \|\tilde{\theta}_0\| \frac{(1-\rho^{T-T_s-1})(1-\eta^{T-T_s-1})}{(1-\rho)(1-\eta)}. \end{aligned}$$

The first regret bound then follows. When Assumption 2 is satisfied and $\varepsilon < \frac{\delta \sqrt{\delta}}{\sqrt{\beta} - \sqrt{\delta}}$, one can make use of the lifted

input contraction in Lemma 1 to show the second regret bound. The full proof can be found in [24]. ■

The theorem states that given the assumptions on the nominal system, both (12) and (13) state trajectories converge to the origin. It is important to note that no further restrictions are put on the stage costs c , other than local Lipschitz continuity. Hence, the finite regret result of the theorem does not imply the cost $\sum_{k=0}^{T-1} c(x_k)$ is minimized, but the relative cost performance is bounded. Moreover, it shows that this performance, in terms of regret scales with the initial exploratory time period T_s , and cannot be avoided, as during these transients the parameter updates are only nonexpansive by Lemma 1. If the additional assumptions on ε and $B_k \phi_k$ are satisfied, then the constant term is updated from a uniform bound b to $c_p := \|\Phi_{T_s}\|$, which relates to the regression matrix only in the initial period.

The following lemma states that Lipschitz continuity of the basis matrices and $\phi(0) = \mathbf{0}_{p \times m}$ is a sufficient condition to fulfill Assumption 2. The proof can be found in [24].

Lemma 2: Let $B_k \phi_k$ be sufficiently exciting, $\phi_k(0) = \mathbf{0}_{p \times m}$, and ϕ_k be uniformly L -Lipschitz continuous, that is, there exists a $L \in \mathbb{R}_+$ such that

$$\|\phi_k(x) - \phi_k(y)\| \leq L\|x - y\|. \quad \forall x, y \in \mathbb{R}^n, \forall k \in \mathbb{N},$$

then Assumption 2 holds.

IV. RLS WITH EXPONENTIAL FORGETTING

The RLSFF algorithm minimizes the online cost

$$g_{k-1}^f(\theta) := \frac{1}{2} \sum_{i=0}^{k-1} \lambda^{2(k-1-i)} \|B_i \phi_i^\top (\theta - \theta^*)\|^2 + \frac{\lambda^{2(k-1)} \varepsilon}{2} \|\theta_0 - \theta^*\|^2, \quad (17)$$

for a fixed exponential forgetting factor $\lambda^2 \in (0, 1)$, where the first term can be thought of as the discounted version of (6) and the second, as a discounted regularizer on the *initial* parameter error. The update can be recursively implemented as follows [25] for all $k \in \mathbb{N}$

$$P_{k+1}^{-1} = \lambda^2 P_k^{-1} + \phi_k B_k^\top B_k \phi_k^\top, \quad (18)$$

$$\theta_{k+1} = \theta_k - P_{k+1} \phi_k B_k^\top (B_k \phi_k^\top \theta_k - y_k), \quad (19)$$

where $\theta_0 \in \mathbb{R}^p$ is an initial guess and $P_0^{-1} = \varepsilon I_p$. The convergence of the above update requires a persistence of excitation condition in the following sense [17].

Definition 2: A matrix signal $\phi_k : \mathbb{N} \rightarrow \mathbb{R}^{p \times n}$ is called persistently exciting if it is bounded and there exist $\delta \in \mathbb{R}_+$, and $T_s \in \mathbb{N}$ such that for all $k_0 \in \mathbb{N}$

$$\delta I_p \leq \sum_{i=k_0}^{k_0+T_s} \phi_i \phi_i^\top.$$

This PE condition, commonly found in the adaptive control literature [1], [17], is a stronger version of SE in Definition 1. The latter requires the positive definiteness of the regressor matrix only in some finite interval of time, rather than, persistently, in all T_s -long time intervals. The following Lemma from [25] shows convergence of θ_k given PE holds.

Lemma 3: Assume $B_k \phi_k^\top$ is persistently exciting, then the RLSFF update (19) satisfies

$$\begin{aligned} \|\theta_k - \theta^*\| &\leq c_r \lambda^{k-T_s} \|\theta_0 - \theta^*\|, & \forall k \geq T_s \\ \|\theta_k - \theta^*\| &\leq \|\theta_0 - \theta^*\|, & \forall 0 < k < T_s \end{aligned}$$

where $c_r^2 = \varepsilon (\lambda^{2T_s} - \lambda^{-2}) / (\delta (1 - \lambda^{-2}))$.

Consider the RLSFF controller $u_k^{\text{RLS}} = \phi_k^\top(x_k) \theta_k$ with θ_k given by the recursive law (18)-(19).

Theorem 2: Let Assumption 1 hold, and $B_k \phi_k^\top(x_k)$ be persistently exciting under the RLSFF adaptive controller, then this controller

- (i) renders the closed-loop system asymptotically stable
- (ii) achieves finite regret

$$\mathcal{R}_T(u^{\text{RLS}}) \leq c_w b L_c \|\tilde{\theta}_0\| \left(\frac{T_s}{1-\rho} + \frac{c_r (\rho^T + 1)}{(1-\rho)^2 (1-\lambda)} \right).$$

Unlike RPL, the RLSFF controller requires the stronger PE condition, but does not restrict the features as in Lemma 2. A sufficient condition for Theorem 2 to hold is that of PE of $B_k \phi_k^\top$. Given that these are time-varying, PE can hold also in the limit, as the state converges to the origin. This happens, notably, in the case of tracking of time-varying references. A numerical example showcasing this is presented in the next section. The proof for Theorem 2 be found in [24].

A. Analysis of the Bounds

Both the RPL and the RLSFF adaptive controllers achieve a finite regret containing three separate terms of potential interest. A constant term, an exponentially decaying one with rate ρ , and a linear term in the PE time T_s . The latter is of most interest, which arises from the non-expansive properties of both estimators and signifies that no improvement in the worst-case can be expected until the time T_s when the PE condition is met. Comparing the regret bounds directly, when $\gamma < \lambda$ and $c_r > 1$, i.e. $\varepsilon > \delta$, then the upper bound of the RLSFF regret is strictly larger than that of the RPL controller.

V. NUMERICAL SIMULATION

In this section, we show that the MRAC problem can be cast into the considered setting and provide a numerical example to showcase the performance of the discussed adaptive controllers. Consider the uncertainty-matched system

$$x_{k+1} = A x_k + B(u_k - \psi^\top(x_k) \theta^*),$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^{p \times m}$ is a feature matrix, and $\theta^* \in \mathbb{R}^p$ is unknown. The MRAC goal is to track the reference dynamics $\bar{x}_{k+1} = A_r \bar{x}_k + B_r r_k$, where $\bar{x}_k \in \mathbb{R}^n$, is the ideal reference vector to be tracked, $r_k \in \mathbb{R}^m$, is the bounded input command, $B_r \in \mathbb{R}^{n \times m}$ and $A_r \in \mathbb{R}^{n \times n}$ is Schur stable. As in [18], we assume (A, B) to be controllable and B to be full column rank, and consider the input $u_k = -K_1 x_k + K_2 r_k + \psi^\top(x_k) \theta_k$ with matrices $K_1, K_2 \in \mathbb{R}^{m \times n}$ chosen such that $A - BK_1 = A_r$ and $BK_2 = B_r$ hold and θ_k being the parameter estimate. Denoting $e_k = x_k - \bar{x}_k$, the error dynamics are represented in the form of (1)

$$e_{k+1} = A_r e_k + B \phi_k^\top(e_k) \tilde{\theta}_k, \quad (20)$$

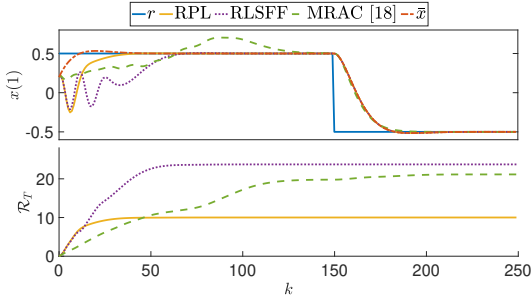


Fig. 1. Comparison of adaptive controllers for the MRAC example in terms of asymptotic tracking (on the top) and regret (on the bottom).

where we define $\phi_k^\top(e_k) := \psi^\top(e_k + \bar{x}_k)$. Since A_r is stable, the E - δ -ISS condition is satisfied. We take the cost to be $c(e) := \|e\|^2$, which is locally Lipschitz since the closed-loop states are bounded. Note that, with this formulation, the SE/PE condition can hold even when the error dynamics are asymptotically stable, given that the reference r_k is such that the feature matrices $\phi_k(e_k)$ are SE/PE.

We consider the following example system [18]

$$A = \begin{bmatrix} 1.0314 & 0.2526 \\ 0.2526 & 1.0314 \end{bmatrix}, B = \begin{bmatrix} 0.0314 \\ 0.2526 \end{bmatrix},$$

with the reference model dynamics

$$A_r = \begin{bmatrix} -0.9929 & 0.2253 \\ -0.0569 & 0.8117 \end{bmatrix}, B_r = \begin{bmatrix} 0.0314 \\ 0.2526 \end{bmatrix}.$$

The true parameter is $\theta^* = [-0.75 \quad -0.50]^\top$ and $\psi(x_k) = x_k$, with the corresponding $\phi_k(e_k) = e_k + \bar{x}_k$. Figure 1 shows the tracking of the first component of state x_k for the above example under the RPL, RLSFF controllers and the command governor-based MRAC controller proposed in [18] without regret guarantees. For all we initialize the estimate by $\theta_0 = [5.00 \quad -1.00]^\top$, take $\varepsilon = 1$ and start the simulation at $x_0 = [0.2 \quad 0.2]^\top$. For RLSFF, we take $\lambda^2 = 0.99$, as lower values led to conditioning problems. For the controller of [18] we used the same parameters as in [18, Example 1] and a tuned saturation value of 1.5 for the command governor update for best results. The superior performance of RPL is evident from Fig. 1, despite a less stringent PE condition. The top figure shows the tracking performance of the controllers, and the bottom, the regret for the quadratic costs, which in this case is $\mathcal{R}_T(u) := \sum_{k=0}^{T-1} \|e_k\|^2$, since by definition $\bar{x}_0 = x_0$ and therefore $e_k^* = 0$ for all $k \in \mathbb{N}$.

VI. CONCLUSION

Two recursive learning algorithms are considered for the nonlinear adaptive control problem with matched uncertainty in the context of regret minimization. We show that both the recursive least squares with forgetting factor and the novel recursive proximal learning algorithm are asymptotically stable and achieve finite regret scaling with the time required to achieve persistence of excitation. Possible extensions include the consideration of inexact basis matrices, bounded, non-stochastic noise, and goal-oriented controllers that instead of the estimation cost optimize over the objective cost directly.

REFERENCES

- [1] K. J. Åström and B. Wittenmark, *Adaptive control*. Addison Wesley Publishing Company, 1989.
- [2] A. Astolfi, D. Karagiannis, and R. Ortega, *Nonlinear and adaptive control with applications*. Springer, 2008, vol. 187.
- [3] K. S. Narendra and A. M. Annaswamy, *Stable adaptive systems*. Courier Corporation, 2012.
- [4] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*. Cambridge university press, 2006.
- [5] M. Raginsky, A. Rakhlin, and S. Yüksel, “Online convex programming and regularization in adaptive control,” in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 1957–1962.
- [6] A. M. Annaswamy and A. L. Fradkov, “A historical perspective of adaptive control and learning,” *Annual Reviews in Control*, vol. 52, pp. 18–41, 2021.
- [7] A. Fradkov, I. Miroshnik, and V. Nikiforov, *Nonlinear and Adaptive Control of Complex Systems*. Springer Science & Business Media, 1999, vol. 491.
- [8] M. Nonhoff and M. A. Müller, “On the relation between dynamic regret and closed-loop stability,” *Systems & Control Letters*, vol. 177, p. 105532, 2023.
- [9] A. Karapetyan, A. Tsiamis, E. C. Balta, A. Iannelli, and J. Lygeros, “Implications of regret on stability of linear dynamical systems,” *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 2583–2588, 2023.
- [10] J. E. Gaudio, T. E. Gibson, A. M. Annaswamy, M. A. Bolender, and E. Lavretsky, “Connections between adaptive control and optimization in machine learning,” in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 4563–4568.
- [11] N. M. Boffi, S. Tu, and J.-J. E. Slotine, “Regret bounds for adaptive nonlinear control,” in *Learning for Dynamics and Control*. PMLR, 2021, pp. 471–483.
- [12] S. Akhtar, R. Venugopal, and D. S. Bernstein, “Logarithmic lyapunov functions for direct adaptive stabilization with normalized adaptive laws,” *International Journal of Control*, vol. 77, no. 7, pp. 630–638, 2004.
- [13] A. Simonetto, A. Mokhtari, A. Koppel, G. Leus, and A. Ribeiro, “A class of prediction-correction methods for time-varying convex optimization,” *IEEE Transactions on Signal Processing*, vol. 64, no. 17, pp. 4576–4591, 2016.
- [14] E. Dall’Anese, A. Simonetto, and A. Bernstein, “On the convergence of the inexact running Krasnosel’skii–mann method,” *IEEE Control Systems Letters*, vol. 3, no. 3, pp. 613–618, 2019.
- [15] N. Parikh, S. Boyd *et al.*, “Proximal algorithms,” *Foundations and trends® in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.
- [16] Y. Pan, T. Shi, and R. Ortega, “Comparative analysis of parameter convergence for several least-squares estimation schemes,” *IEEE Transactions on Automatic Control*, 2023.
- [17] R. M. Johnstone, C. R. Johnson Jr, R. R. Bitmead, and B. D. Anderson, “Exponential convergence of recursive least squares with exponential forgetting factor,” *Systems & Control Letters*, vol. 2, no. 2, pp. 77–82, 1982.
- [18] K. M. Dogan, T. Yucelen, W. M. Haddad, and J. A. Muse, “Improving transient performance of discrete-time model reference adaptive control architectures,” *International Journal of Adaptive Control and Signal Processing*, vol. 34, no. 7, pp. 901–918, 2020.
- [19] D. Angeli, “A Lyapunov approach to incremental stability properties,” *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 410–421, 2002.
- [20] D. N. Tran, B. S. Rüffer, and C. M. Kellett, “Convergence properties for discrete-time nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 64, no. 8, pp. 3415–3422, 2018.
- [21] A. Karapetyan, E. C. Balta, A. Iannelli, and J. Lygeros, “Closed-loop finite-time analysis of suboptimal online control,” *arXiv preprint arXiv:2312.05607*, 2023.
- [22] E. Hazan, A. Kalai, S. Kale, and A. Agarwal, “Logarithmic regret algorithms for online convex optimization,” in *International Conference on Computational Learning Theory*. Springer, 2006, pp. 499–513.
- [23] Z.-P. Jiang and Y. Wang, “Input-to-state stability for discrete-time nonlinear systems,” *Automatica*, vol. 37, no. 6, pp. 857–869, 2001.
- [24] A. Karapetyan, E. C. Balta, A. Tsiamis, A. Iannelli, and J. Lygeros, “On the regret of recursive methods for discrete-time adaptive control with matched uncertainty,” *arXiv preprint arXiv:2404.02023*, 2024.
- [25] S. Brüggemann and R. R. Bitmead, “Exponential convergence of recursive least squares with forgetting factor for multiple-output systems,” *Automatica*, vol. 124, p. 109389, 2021.