

Discrete-Time Adaptive Control in High Dimensions: Near Dimension-Free Performance via Mirror Descent

Mohammad Boveiri¹ and Peyman Mohajerin Esfahani^{1,2}

¹Delft University of Technology, The Netherlands

²University of Toronto, Canada

ABSTRACT. Motivated by the use of modern high-capacity models in real-time control problems, this paper studies the adaptive control of high-dimensional discrete-time nonlinear systems with an unknown matrix-valued parameter. We focus on regimes where the number of unknown parameter entries is large, but the parameter matrix possesses exploitable structure, such as entrywise sparsity, group sparsity, low rank, and row-stochastic or density-matrix structure. To quantify transient performance, we consider a regret criterion relative to a nominal controller with full knowledge of the true parameter. We show that standard Euclidean update schemes, including recursive least squares and gradient descent, are ill-suited to this setting: their transient performance deteriorates as the dimension increases, even when the true parameter has low intrinsic complexity. To address this limitation, we propose a novel class of mirror-descent-type adaptive laws equipped with a non-Euclidean Polyak-type step size that exploit the geometry induced by the parameter structure. For the proposed update laws, we establish asymptotic state convergence and derive regret bounds with at most logarithmic dependence on the ambient dimension. Numerical experiments demonstrate the effectiveness of the proposed schemes.

I. INTRODUCTION

A. Problem description

In this paper, we consider the adaptive control of uncertain discrete-time nonlinear systems of the form

$$X_{t+1} = \mathcal{A}(X_t) + \mathcal{B}(\Theta\Psi(X_t, t) + U_t), \quad (1)$$

where $X_t \in \mathbb{R}^{n_1 \times n_2}$ is the system state, $U_t \in \mathbb{R}^{m \times n_2}$ is the control input, and $\Theta \in \Omega \subseteq \mathbb{R}^{m \times k}$ is an unknown parameter matrix. The known mapping $\Psi : \mathbb{R}^{n_1 \times n_2} \times \mathbb{Z}_+ \rightarrow \mathbb{R}^{k \times n_2}$ is a possibly time-varying feature map, or regressor. The known linear operators $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}^{n_1 \times n_2}$ and $\mathcal{B} : \mathbb{R}^{m \times n_2} \rightarrow \mathbb{R}^{n_1 \times n_2}$ describe the nominal dynamics and the control channel, respectively. Since the uncertain term $\Theta\Psi(X_t, t)$ enters the dynamics through the same channel as the control input, the system has a matched uncertainty structure [1]. Throughout this study, we allow the regressor to be unbounded, provided that it grows at most linearly with the state, as formalized below.

Assumption 1 (Linear growth). *The regressor Ψ satisfies*

$$\|\Psi(X, t)\| \leq c_1 + c_2\|X\|, \quad \forall (X, t) \in \mathbb{R}^{n_1 \times n_2} \times \mathbb{Z}_+, \quad (2)$$

for some constants $c_1, c_2 \in \mathbb{R}_+$, which need not be known.

The control task is to track the desired state trajectory $\{X_t^d\}_{t \in \mathbb{Z}_+}$ generated by the reference model

$$X_{t+1}^d = (\mathcal{A} - \mathcal{BK})(X_t^d) + \mathcal{B}(U_t^d), \quad (3)$$

where $U_t^d \in \mathbb{R}^{m \times n_2}$ is a bounded reference input and $\mathcal{K} : \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}^{m \times n_2}$ is a linear feedback operator chosen such that the closed-loop operator $\mathcal{A} - \mathcal{BK}$ is Schur stable. We are particularly interested in high-dimensional settings where the number of unknown parameters, $d = mk$, is large.

Date: August 3, 2026.

This work was partially supported by the ERC grant TRUST-949796 and the NSERC Grant RGPIN-2025-06544.

Our objective is to develop adaptive control schemes with performance guarantees that are dimension-free, up to logarithmic factors, by leveraging structural properties of Θ , such as sparsity, low-rank structure, or stochastic matrix structure. Although the system in (1) can be rewritten in standard vector form by stacking the columns of X_t , we retain the matrix representation in order to exploit structural properties of the uncertainty term $\Theta\Psi(X_t, t)$, such as low-rank structure, which are more naturally expressed in matrix form.

Control law. Given an estimate $\hat{\Theta}_t$ of the unknown matrix Θ at time t , we apply the certainty-equivalent control law

$$U_t(X_t, \hat{\Theta}_t) = -\mathcal{K}(X_t) - \hat{\Theta}_t\Psi(X_t, t) + U_t^d. \quad (4)$$

The term ‘certainty-equivalent’ indicates that the above controller is designed by treating the estimate $\hat{\Theta}_t$ as the true parameter Θ ; in the ideal case where $\hat{\Theta}_t = \Theta$, the control law (4) ensures the cancellation of the unknown nonlinearity and convergence of X_t to X_t^d . In the adaptive setting, the estimate $\hat{\Theta}_t$ is updated online based on the observed state evolution.

We couple the controller (4) with a family of non-Euclidean parameter update laws. These updates are driven by the instantaneous performance loss induced by the parameter-estimation error $\Theta - \hat{\Theta}_t$ in the closed-loop system. Specifically, after applying the control input at time t , we define

$$J_t(\hat{\Theta}_t) := \frac{1}{2} \left\| \mathcal{B}((\Theta - \hat{\Theta}_t)\Psi(X_t, t)) \right\|_F^2 = \frac{1}{2} \|\tilde{X}_{t+1}\|_F^2, \quad (5)$$

where $\tilde{X}_{t+1} := X_{t+1} - (\mathcal{A} - \mathcal{BK})(X_t) - \mathcal{B}(U_t^d)$. The gradient of J_t with respect to $\hat{\Theta}_t$ is

$$\nabla J_t(\hat{\Theta}_t) = -\mathcal{B}^\top(\tilde{X}_{t+1})\Psi(X_t, t)^\top, \quad (6)$$

where $\mathcal{B}^\top$ denotes the adjoint of the linear operator $\mathcal{B}$; see Section A. Observe that, although J_t is associated with the control action applied at time t , both J_t and $\nabla J_t(\hat{\Theta}_t)$ become available only at time $t + 1$, once the residual $\tilde{X}_{t+1}$ can be computed from the observed state X_{t+1} .

B. Contributions

Parameter update law. Let $\mathbb{R}^{m \times k}$ be equipped with a norm $\|\cdot\|$, and let $\|\cdot\|_*$ denote the associated dual norm. Moreover, let $f : \mathbb{R}^{m \times k} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a μ -strongly convex function on $\Omega \subset \mathbb{R}^{m \times k}$ with respect to $\|\cdot\|$, and let f^* denote its Fenchel conjugate. The estimate $\hat{\Theta}_t$ in the control law (4) is updated according to the following mirror-descent-type recursion:

$$\begin{aligned} Z_{t+1} &= Z_t - \eta_{t+1} \nabla J_t(\hat{\Theta}_t), & Z_0 &\in \partial f(\hat{\Theta}_0), \\ \hat{\Theta}_{t+1} &= \nabla f^*(Z_{t+1}), \end{aligned} \quad (7a)$$

where the step size is selected as

$$\eta_{t+1} = \frac{2\mu J_t(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2} \quad (7b)$$

for some nonnegative and bounded sequence $\{\epsilon_{t+1}\}_{t \in \mathbb{Z}_+}$.

Interestingly, (7b) may be viewed as a non-Euclidean and online generalization of the Polyak step size; see Remark 1. Within this framework, we establish the following results.

(i) Stability and state convergence. We establish boundedness and asymptotic convergence properties of the closed-loop system under the parameter update law (7). Specifically, we show that the system state X_t and parameter estimate $\hat{\Theta}_t$ remain bounded and that X_t converges asymptotically to the desired trajectory X_t^d (Theorem 1). We emphasize that the step-size choice (7b) is essential for guaranteeing stability and state convergence when the map Ψ is not necessarily bounded and may grow with the state according to Assumption 1.

(ii) **Performance.** We quantify transient performance via the cumulative loss (regret) defined by

$$\text{Reg}(T) := \sum_{t=0}^{T-1} J_t(\hat{\Theta}_t), \quad T > 0. \quad (8)$$

Observe that a controller of the form (4) with full knowledge of the true parameter Θ , namely $U_t(X_t, \Theta)$, incurs zero regret, since $J_t(\Theta) = 0$ for all t . Thus, $\text{Reg}(T)$ characterizes the cumulative performance loss over the horizon T caused by estimating Θ through (7), rather than having access to it a priori. Given this setup, we show that the regret admits the bound

$$\text{Reg}(T) \leq \sqrt{\frac{D_f(\Theta, \hat{\Theta}_0; Z_0)}{2\mu} \sum_{t=0}^{T-1} (\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2)}, \quad (9)$$

where D_f denotes the generalized Bregman divergence generated by the mirror map f (Theorem 3). Since boundedness of X_t and $\hat{\Theta}_t$ has already been established, $\|\nabla J_t(\hat{\Theta}_t)\|_*$ is bounded. Hence, (9) implies that $\text{Reg}(T)$ scales at most as $\sqrt{T}$ with the horizon.

(iii) **Dimension-free regret via geometry-aware updates.** The regret bound (9) depends on D_f and $\|\cdot\|_*$, and thus on the geometry induced by the mirror map f . For example, with the Euclidean mirror map $f(\Theta) = \frac{1}{2}\|\Theta\|_F^2$, the update law (7a) reduces to a gradient-descent-type update. In this case, the dual norm is the Frobenius norm, and, assuming $\hat{\Theta}_0 = 0$,

$$D_f(\Theta, \hat{\Theta}_0; Z_0) = \frac{1}{2}\|\Theta\|_F^2, \quad \|\nabla J_t(\hat{\Theta}_t)\|_* = \|\nabla J_t(\hat{\Theta}_t)\|_F.$$

Hence, even if Θ is s -sparse with $s \ll d$, the regret is $\mathcal{O}(\sqrt{d})$, since $\|\nabla J_t(\hat{\Theta}_t)\|_F = \mathcal{O}(\sqrt{d})$. This dimensional dependence of Euclidean-based schemes is not an artifact of conservative analysis; it is also observed empirically (see Section II).

In contrast, by selecting the mirror map according to the structure of Θ , we show that regret bounds with at most logarithmic dependence on the dimension can be achieved. We consider two classes of structured parameter sets:

- **Sparsity structures.** We consider several notions of sparsity for Θ , including (a) entrywise sparsity, where at most s entries of Θ are nonzero; (b) group sparsity, where the nonzero entries of Θ are structured into groups, such as in row-wise or column-wise sparsity; and (c) low-rank structure, where $\text{rank}(\Theta) \leq s$; see Section IV.
- **Simplex-type sets.** We also consider the case where the unknown parameter Θ belongs to a simplex-type constraint set. Specifically, we study two representative examples: the standard probability simplex and the spectraplex (i.e., the set of density matrices); see Section V.

For each case above, we identify a suitable mirror map, derive the corresponding explicit update law, and establish regret bounds with at most logarithmic dependence on the dimension.

C. Motivation and related literature

Historically, one of the primary objectives of adaptive control theory has been the development of universal controllers capable of managing broad classes of dynamical systems across diverse operating scenarios [2, 3]. This goal naturally necessitates high-dimensional modeling to capture complex system dynamics, exogenous disturbances, and potential faults. Traditionally, however, overparameterized modeling has been viewed as problematic (see, e.g., [4]) primarily due to the severe degradation in transient performance observed in existing schemes as the dimensionality increases (see Section II). This tension between expressive, high-capacity modeling and reliable transient behavior in adaptive control design has been a central obstacle to integrating modern high-capacity models into real-time control systems.

In this paper, we show that such performance degradation is not an inherent consequence of using high-dimensional models. Rather, it arises largely from the use of Euclidean parameter update laws, which

fail to preserve or exploit structural properties commonly present in high-dimensional models, such as sparsity. By incorporating such structural properties directly into the geometry of the parameter update law, we obtain transient performance guarantees that depend on the intrinsic complexity of the unknown parameter rather than on the ambient dimension. We believe that, by addressing this fundamental limitation, this paper paves the way for the use of high-capacity models in adaptive control without sacrificing transient performance.

To the best of the authors' knowledge, this work provides the first dimension-independent performance guarantees for discrete-time adaptive systems in high-dimensional settings.

Mirror descent. Mirror-descent algorithms, introduced in [5], have been widely used in optimization and machine learning problems, including sparse signal recovery [6], ℓ_1 - and nuclear-norm minimization [7], online optimization [8, 9], and saddle-point problems [10, 11]. From a purely algorithmic perspective, several aspects of our proposed mirror update law, such as its Polyak-type step size (7b) and the allowance for a nondifferentiable mirror map, are nonstandard and specifically tailored to the adaptive control framework. These features address challenges that are unique to the adaptive control setting, such as establishing stability of the interconnected plant-controller system, boundedness of the closed-loop signals, and asymptotic convergence of the tracking error.

Regret as a measure of transient performance. Regret is a performance measure that evaluates an algorithm's performance relative to a benchmark strategy. It has become a standard criterion for assessing the finite-time performance of algorithms in online decision-making problems, including online convex optimization [12], multi-armed bandits [13, 14], and reinforcement learning and control [15]. In adaptive control, regret quantifies the performance loss incurred by a learning controller relative to a benchmark controller with full knowledge of the system dynamics. Existing regret analyses for adaptive control have primarily emphasized the dependence of the bound on the time horizon T , rather than on the ambient dimension of the unknown parameters. For instance, in linear quadratic regulator problems, sublinear regret bounds, often of order $\sqrt{T}$, have been established under various assumptions on data generation, stability, excitation, and stochastic noise processes [16–20]. The regret definition adopted in this work is similar in spirit to that of [21], measuring the cumulative performance gap between our adaptive controller and a certainty-equivalence benchmark with access to the true system parameter.

Paper structure. The remainder of this paper is organized as follows. Section II presents a scalar motivating example illustrating how the transient performance of standard Euclidean-based parameter update laws deteriorates as the ambient dimension grows. Section III develops the proposed mirror descent adaptive control framework, proves the boundedness of the closed-loop signals, establishes state convergence, and derives a general regret bound for the proposed scheme. Section IV specializes the framework to parameter classes with specific sparsity structures, including entrywise sparsity, group sparsity, and low-rank structure, and derives the corresponding regret bounds. Section V extends the approach to simplex-type constraint sets, such as the standard simplex and the spectraplex. Section VI extends the proposed update law to stochastic systems subject to process noise. Section VII presents numerical simulations that validate the theoretical results. Finally, Section VIII concludes the paper and outlines potential directions for future research.

Notation. The symbols $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{Z}_+$ denote the sets of real numbers, nonnegative real numbers, and nonnegative integers, respectively. For a matrix $Q \in \mathbb{R}^{m \times k}$, $\|Q\|_p$ denotes the entrywise ℓ_p -norm, i.e., $\|Q\|_p = \left(\sum_{j,i} |Q_{ij}|^p \right)^{1/p}$ for $p \in [1, \infty)$, while $\|Q\|_\infty = \max_{i,j} |Q_{ij}|$. Moreover, $\|Q\|_{S_p}$ denotes the Schatten p -norm, and $\|Q\|_F$ denotes the Frobenius norm. For a linear operator $\mathcal{A}$, $\|\mathcal{A}\|_{\text{op}}$ denotes its operator norm. $\mathbb{S}_+^n$ denotes the cone of $n \times n$ symmetric positive semidefinite matrices. The Hadamard product is denoted by $\odot$. For a function f , ∂f denotes its subdifferential; if f is differentiable, then $\partial f = \{\nabla f\}$. The cardinality of a set $\mathcal{S}$ is indicated by $|\mathcal{S}|$.

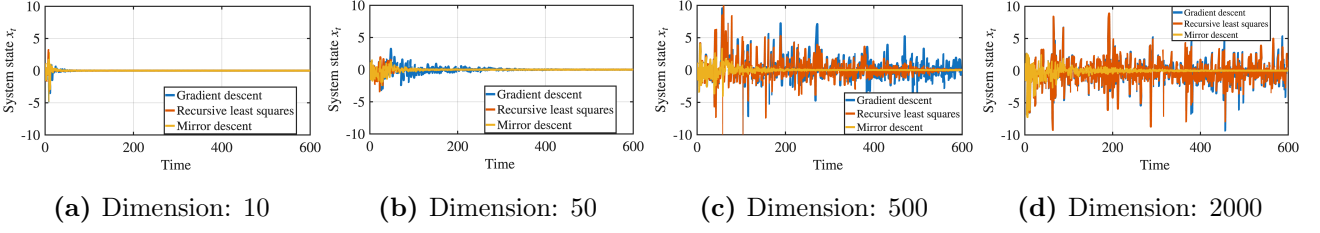


FIGURE 1. System state trajectories under control law (11) for varying dimensions k . Comparison between gradient descent (13), recursive least squares (14), and sparse mirror descent (29).

II. MOTIVATING EXAMPLE: EUCLIDEAN-BASED SCHEMES IN HIGH DIMENSIONS

In this section, we present a motivating example to illustrate the behavior of commonly used Euclidean-based adaptive update laws, namely, gradient-descent-type and recursive least squares methods, in high-dimensional settings, and to highlight the need for non-Euclidean adaptation. To this end, consider the adaptive stabilization problem for the scalar system

$$x_{t+1} = \theta^\top \psi(x_t, t) + u_t, \quad (10)$$

where $\theta \in \mathbb{R}^k$ is an unknown parameter vector. We assume $\psi(x_t, t)$ is a stochastic regressor vector whose components $\psi_i(x_t, t)$ are independent and identically distributed (i.i.d.) processes defined by:

$$\psi_i(x_t, t) = \begin{cases} 0.5x_t + 1, & \text{with probability 0.5,} \\ -0.5x_t - 1, & \text{with probability 0.5.} \end{cases}$$

We further assume that θ is sparse, with only the first three components being non-zero: $\theta_1 = \theta_2 = \theta_3 = 1$, and $\theta_i = 0$ for all $i \geq 4$. Consequently, increasing the dimension does not alter the underlying system dynamics, as it merely appends zero-valued parameters. Following the control structure in (4), we employ the certainty-equivalence control law:

$$u(x_t, \hat{\theta}_t) = -\hat{\theta}_t^\top \psi(x_t, t). \quad (11)$$

The instantaneous cost $J_t(\hat{\theta}_t)$ defined in (5) satisfies

$$J_t(\hat{\theta}_t) = \frac{1}{2} \left((\theta - \hat{\theta}_t)^\top \psi(x_t, t) \right)^2 = \frac{1}{2} x_{t+1}^2, \quad (12)$$

with the gradient given by $\nabla J_t(\hat{\theta}_t) = -x_{t+1} \psi(x_t, t)$.

We consider two standard Euclidean update laws for the parameter estimate $\hat{\theta}_t$: (normalized) gradient descent and recursive least squares. Under the Assumption 1, both update laws guarantee boundedness of the closed-loop signals and asymptotic convergence of x_t to the origin; see, e.g., [22–24].

(i) (Normalized) gradient descent. At time step $t + 1$, the estimate $\hat{\theta}_t$ is updated according to

$$\hat{\theta}_{t+1} = \hat{\theta}_t + \frac{\psi(x_t, t)x_{t+1}}{1 + \|\psi(x_t, t)\|_2^2}. \quad (13)$$

(ii) Recursive least squares. At time step $t + 1$, the parameter estimate $\hat{\theta}_t$ and the covariance matrix $P_t \in \mathbb{S}^k$ are updated according to

$$\begin{aligned} \hat{\theta}_{t+1} &= \hat{\theta}_t + \frac{P_t \psi(x_t, t) x_{t+1}}{1 + \psi(x_t, t)^\top P_t \psi(x_t, t)}, \\ P_{t+1} &= P_t - \frac{P_t \psi(x_t, t) \psi(x_t, t)^\top P_t}{1 + \psi(x_t, t)^\top P_t \psi(x_t, t)}. \end{aligned} \quad (14)$$

It is worth noting that a key drawback of recursive least squares, relative to normalized gradient descent, is that it requires maintaining and updating a $k \times k$ matrix at each iteration, which may be impractical in high-dimensional problems.

Figure 1 illustrates the state trajectories for the closed-loop system using the aforementioned update laws alongside the proposed mirror descent algorithm (7). The mirror descent scheme utilizes the mirror map (28). All algorithms are initialized with $\hat{\theta}_0 = \mathbf{0}$ and $x_0 = 1$, and simulations are conducted for $k \in \{10, 50, 500, 2000\}$. As observed, although the underlying system dynamics remain unchanged with increasing dimension, the performance of normalized gradient descent and recursive least squares degrades significantly. This behavior is consistent with the regret bound in (9). Intuitively, by propagating estimation errors across all dimensions, Euclidean-based methods fail to isolate the low-dimensional signal. In contrast, mirror descent maintains robust, nearly dimension-invariant performance by exploiting the underlying sparse geometry of the parameter space.

III. MIRROR UPDATES FOR DISCRETE-TIME ADAPTIVE CONTROL

This section develops a mirror descent framework for parameter adaptation in adaptive control of systems of the form (1). We establish the closed-loop stability and asymptotic convergence properties of the resulting adaptive system. Furthermore, we derive a regret bound to characterize the transient performance of the proposed scheme.

A. Preliminaries

This subsection reviews preliminary definitions and results from convex analysis that will be used throughout the subsequent development. For a comprehensive treatment, we refer the reader to [25–28].

We consider the vector space of matrices $\mathbb{R}^{m \times k}$ equipped with the Frobenius inner product

$$\langle X, Y \rangle := \text{Tr}(X^\top Y), \quad X, Y \in \mathbb{R}^{m \times k}.$$

Let $\|\cdot\|$ be a norm on $\mathbb{R}^{m \times k}$. Its dual norm, denoted by $\|\cdot\|_*$, is defined for any $Z \in \mathbb{R}^{m \times k}$ by

$$\|Z\|_* := \sup_{X \in \mathbb{R}^{m \times k}} \{\langle Z, X \rangle : \|X\| \leq 1\}.$$

Let $f : \mathbb{R}^{m \times k} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous, convex function, with effective domain

$$\text{dom } f := \{X \in \mathbb{R}^{m \times k} : f(X) < +\infty\}.$$

The *Fenchel conjugate* of f is defined by

$$f^*(Z) := \sup_{X \in \text{dom } f} \{\langle Z, X \rangle - f(X)\}. \quad (15)$$

For $Q \in \text{dom } f$ such that $\partial f(Q) \neq \emptyset$, and for a fixed subgradient $Z \in \partial f(Q)$, we define the generalized *Bregman divergence* associated with f by

$$D_f(P, Q; Z) := f(P) + f^*(Z) - \langle P, Z \rangle, \quad P \in \text{dom } f. \quad (16)$$

Since $Z \in \partial f(Q)$, the Fenchel–Young equality implies that

$$f(Q) + f^*(Z) = \langle Q, Z \rangle.$$

Therefore, $D_f(P, Q; Z) = f(P) - f(Q) - \langle P - Q, Z \rangle$. By the Fenchel–Young inequality, $D_f(P, Q; Z) \geq 0$. Geometrically, $D_f(P, Q; Z)$ measures the gap between $f(P)$ and the supporting affine function of f at Q induced by subgradient Z . A convex function f is $\mathcal{L}$ -smooth with respect to $\|\cdot\|$ if it is differentiable and

$$f(P) \leq f(Q) + \langle \nabla f(Q), P - Q \rangle + \frac{\mathcal{L}}{2} \|P - Q\|^2, \quad \forall P, Q \in \text{dom } f.$$

A proper convex function f is μ -strongly convex with respect to $\|\cdot\|$ if

$$f(P) \geq f(Q) + \langle Z, P - Q \rangle + \frac{\mu}{2} \|P - Q\|^2, \quad \forall P, Q \in \text{dom } f, \quad \forall Z \in \partial f(Q).$$

Throughout the paper, we assume that the mirror map f in the update law (7) is μ -strongly convex with respect to $\|\cdot\|$, but not necessarily smooth. Under this assumption, its Fenchel conjugate f^* is differentiable and $(1/\mu)$ -smooth with respect to the dual norm $\|\cdot\|_*$. Moreover, the inverse relation

$$\nabla f^* = (\partial f)^{-1}$$

holds, where ∂f denotes the subdifferential of f . Lastly, for the linear operator $\mathcal{B} : \mathbb{R}^{m \times n_2} \rightarrow \mathbb{R}^{n_1 \times n_2}$, its adjoint $\mathcal{B}^\top : \mathbb{R}^{n_1 \times n_2} \rightarrow \mathbb{R}^{m \times n_2}$ is defined through the identity

$$\langle \mathcal{B}(U), V \rangle = \langle U, \mathcal{B}^\top(V) \rangle, \quad (17)$$

for all $U \in \mathbb{R}^{m \times n_2}$ and $V \in \mathbb{R}^{n_1 \times n_2}$.

B. Closed-loop stability and finite-time performance

Before presenting the main stability results, we detail the algorithmic steps of the parameter update law (7a). Given an initial estimate $\hat{\Theta}_0$, we select a compatible dual initialization $Z_0 \in \partial f(\hat{\Theta}_0)$. The adaptation then evolves in the dual space according to the gradient-descent recursion $Z_{t+1} = Z_t - \eta_{t+1} \nabla J_t(\hat{\Theta}_t)$. At each time step, the primal parameter estimate is recovered via the mapping $\hat{\Theta}_{t+1} = \nabla f^*(Z_{t+1})$. As discussed in the previous subsection, since f is μ -strongly convex, its Fenchel conjugate f^* is differentiable and satisfies $\nabla f^* = (\partial f)^{-1}$. Consequently, the primal-dual consistency relation $Z_{t+1} \in \partial f(\hat{\Theta}_{t+1})$ is preserved by construction at each time step. We emphasize that the use of nondifferentiable mirror maps provides a convenient mechanism for incorporating constraints on Θ . Specifically, the indicator function of the feasible set can be included directly in the definition of f . The resulting extended-valued mirror map enforces feasibility through its effective domain while preserving the primal-dual update structure described above; see Section V.

The following theorem establishes the stability of the closed-loop system and the asymptotic convergence of the tracking error to zero.

Theorem 1 (Stability and convergence). *Consider the system (1) under the control law (4) and the parameter update law (7). Suppose that Assumption 1 holds. Then, the sequences $\{X_t\}_{t \geq 0}$ and $\{\hat{\Theta}_t\}_{t \geq 0}$ are uniformly bounded. Moreover,*

$$\lim_{t \rightarrow \infty} \|X_t - X_t^d\|_F = 0. \quad (18)$$

Proof. Consider the nonnegative function $V_t := D_f(\Theta, \hat{\Theta}_t; Z_t)$. The increment of V_t is given by

$$\Delta V_{t+1} := V_{t+1} - V_t = f^*(Z_{t+1}) - f^*(Z_t) - \langle \Theta, Z_{t+1} - Z_t \rangle.$$

Since f^* is $(1/\mu)$ -smooth with respect to $\|\cdot\|_*$, we have

$$f^*(Z_{t+1}) \leq f^*(Z_t) + \langle \nabla f^*(Z_t), Z_{t+1} - Z_t \rangle + \frac{1}{2\mu} \|Z_{t+1} - Z_t\|_*^2.$$

Substituting this bound into the expression for ΔV_{t+1} yields

$$\Delta V_{t+1} \leq \langle \nabla f^*(Z_t) - \Theta, Z_{t+1} - Z_t \rangle + \frac{1}{2\mu} \|Z_{t+1} - Z_t\|_*^2.$$

Substituting the update law (7a) into the inequality above gives

$$\Delta V_{t+1} \leq \eta_{t+1} \langle \nabla J_t(\hat{\Theta}_t), \Theta - \hat{\Theta}_t \rangle + \frac{\eta_{t+1}^2}{2\mu} \|\nabla J_t(\hat{\Theta}_t)\|_*^2 \quad (19)$$

By exploiting the adjoint identity (17), the inner product term can be evaluated as follows:

$$\begin{aligned} \langle \nabla J_t(\hat{\Theta}_t), \Theta - \hat{\Theta}_t \rangle &= -\langle \mathcal{B}^\top(\tilde{X}_{t+1})\Psi(X_t, t)^\top, \Theta - \hat{\Theta}_t \rangle = -\langle \mathcal{B}^\top(\tilde{X}_{t+1}), (\Theta - \hat{\Theta}_t)\Psi(X_t, t) \rangle \\ &= -\langle \tilde{X}_{t+1}, \mathcal{B}((\Theta - \hat{\Theta}_t)\Psi(X_t, t)) \rangle = -\langle \tilde{X}_{t+1}, \tilde{X}_{t+1} \rangle = -\|\tilde{X}_{t+1}\|_F^2. \end{aligned} \quad (20)$$

Substituting this result back into (19) yields

$$\Delta V_{t+1} \leq -\eta_{t+1} \|\tilde{X}_{t+1}\|_F^2 + \frac{\eta_{t+1}^2}{2\mu} \|\nabla J_t(\hat{\Theta}_t)\|_*^2. \quad (21)$$

Substituting the step size (7b) into the inequality above yields

$$\Delta V_{t+1} \leq -\frac{\mu}{2} \frac{\|\tilde{X}_{t+1}\|_F^4}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2} \leq 0. \quad (22)$$

Equation (22) implies that the nonnegative sequence $\{V_t\}_{t \geq 0}$ is monotonically nonincreasing and converges to a finite asymptotic value. Moreover, since f is μ -strongly convex and $Z_t \in \partial f(\hat{\Theta}_t)$, we have

$$\frac{\mu}{2} \|\Theta - \hat{\Theta}_t\|^2 \leq V_t \leq V_0.$$

Hence, the estimate $\hat{\Theta}_t$ remains bounded. Since V_t converges, we also have $\lim_{t \rightarrow \infty} \Delta V_{t+1} = 0$. Thus, (22) implies

$$\lim_{t \rightarrow \infty} \frac{\|\tilde{X}_{t+1}\|_F^4}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2} = 0. \quad (23)$$

The above limit does not, by itself, imply the boundedness or asymptotic convergence of $\tilde{X}_t$, since the denominator may grow faster than the numerator. We next rule out this possibility under the linear-growth condition (2). To this end, we first establish an upper bound on the growth rate of the denominator in terms of the residual norm $\|\tilde{X}_{t+1}\|_F$.

Lemma 2. *Suppose that Ψ satisfies Assumption 1. Then there exist constants $c'_1, c'_2 > 0$ such that*

$$\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2 \leq c'_1 + c'_2 \max_{0 \leq i \leq t} \|\tilde{X}_{i+1}\|_F^4, \quad \forall t \in \mathbb{Z}_+, \quad (24)$$

The proof of Lemma 2 is provided in Appendix A. Combining (24) with (23), we obtain

$$0 \leq \frac{\|\tilde{X}_{t+1}\|_F^4}{c'_1 + c'_2 \max_{0 \leq i \leq t} \|\tilde{X}_{i+1}\|_F^4} \leq \frac{\|\tilde{X}_{t+1}\|_F^4}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2}.$$

Hence, we have

$$\lim_{t \rightarrow \infty} \frac{\|\tilde{X}_{t+1}\|_F^4}{c'_1 + c'_2 \max_{0 \leq i \leq t} \|\tilde{X}_{i+1}\|_F^4} = 0. \quad (25)$$

Note that if $\{\tilde{X}_{t+1}\}_{t \in \mathbb{Z}_+}$ is bounded, then (25) implies that $\|\tilde{X}_{t+1}\|_F \rightarrow 0$, since its denominator is uniformly bounded. We now establish the boundedness of $\{\tilde{X}_{t+1}\}_{t \in \mathbb{Z}_+}$ using a contradiction argument analogous to the key technical lemma in discrete-time adaptive control [22, Lemma 6.2.1]. Suppose, for contradiction, that the sequence $\{\|\tilde{X}_{t+1}\|_F\}_{t \in \mathbb{Z}_+}$ is unbounded. Then there exists a strictly increasing sequence of indices $\{t_n\}_{n \geq 1}$ such that

$$\|\tilde{X}_{t_n+1}\|_F = \max_{0 \leq i \leq t_n} \|\tilde{X}_{i+1}\|_F \quad \text{and} \quad \|\tilde{X}_{t_n+1}\|_F \rightarrow \infty$$

as $n \rightarrow \infty$. Therefore, along this subsequence

$$\lim_{n \rightarrow \infty} \frac{\|\tilde{X}_{t_n+1}\|_F^4}{c'_1 + c'_2 \max_{0 \leq i \leq t_n} \|\tilde{X}_{i+1}\|_F^4} = \lim_{n \rightarrow \infty} \frac{\|\tilde{X}_{t_n+1}\|_F^4}{c'_1 + c'_2 \|\tilde{X}_{t_n+1}\|_F^4} = \lim_{n \rightarrow \infty} \frac{1}{c'_1 / \|\tilde{X}_{t_n+1}\|_F^4 + c'_2} = \frac{1}{c'_2} > 0.$$

This contradicts (25). Hence, $\{\tilde{X}_{t+1}\}_{t \in \mathbb{Z}_+}$ is bounded. Consequently, the denominator in (25) is uniformly bounded, and therefore $\tilde{X}_{t+1} \rightarrow 0$ as $t \rightarrow \infty$.

Finally, the tracking error dynamics can be expressed as

$$X_{t+1} - X_{t+1}^d = (\mathcal{A} - \mathcal{B}\mathcal{K})(X_t - X_t^d) + \tilde{X}_{t+1}.$$

Since $\tilde{X}_{t+1} \rightarrow 0$ as $t \rightarrow \infty$ and the closed-loop operator $(\mathcal{A} - \mathcal{B}\mathcal{K})$ is Schur stable, it follows from standard stability theory for linear systems under vanishing inputs (see e.g. [29]) that $X_t \rightarrow X_t^d$ as $t \rightarrow \infty$. $\square$

Remark 1 (Non-Euclidean Polyak step size). Since $J_t(\Theta) = 0$, the step size in (7b) can be written as

$$\eta_{t+1} = \frac{2\mu (J_t(\hat{\Theta}_t) - J_t(\Theta))}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2}.$$

In the special case where the regressor is constant, i.e., $\Psi = C \in \mathbb{R}^{k \times n_2}$, the loss J_t is time invariant. Moreover, if the Euclidean geometry is used and $\epsilon_{t+1} = 0$, then the step size reduces to the classical Polyak step size [30, 31]. Therefore, (7b) can be viewed as a non-Euclidean, online generalization of the Polyak step size. The regularization term ϵ_{t+1} provides an additional degree of freedom to avoid potential numerical issues, such as division by zero in the presence of noise. A typical choice is $\epsilon_{t+1} = \alpha(t+1)^{-\beta}$ with $\alpha, \beta \in \mathbb{R}_+$.

In the following theorem, we establish a regret bound for the proposed framework to characterize its finite-time (non-asymptotic) performance.

Theorem 3 (Regret bound). *Consider the system (1) under the control law (4) with the parameter update law (7). Then, for any horizon $T \geq 1$, the regret satisfies*

$$\text{Reg}(T) \leq \sqrt{\frac{D_f(\Theta, \hat{\Theta}_0; Z_0)}{2\mu} \sum_{t=0}^{T-1} (\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2)}, \quad (26)$$

Proof. Following the proof of Theorem 1, define $V_t := D_f(\Theta, \hat{\Theta}_t; Z_t)$. Summing both sides of (22) over $t = 0, \dots, T-1$ yields

$$\sum_{t=0}^{T-1} \frac{J_t^2(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2} = \frac{1}{4} \sum_{t=0}^{T-1} \frac{\|\tilde{X}_{t+1}\|_F^4}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2} \leq \frac{1}{2\mu} \sum_{t=0}^{T-1} (V_t - V_{t+1}) = \frac{1}{2\mu} (V_0 - V_T) \leq \frac{V_0}{2\mu}. \quad (27)$$

By applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \text{Reg}(T) &= \sum_{t=0}^{T-1} J_t(\hat{\Theta}_t) = \sum_{t=0}^{T-1} \frac{J_t(\hat{\Theta}_t)}{\sqrt{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2}} \sqrt{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2} \\ &\leq \sqrt{\sum_{t=0}^{T-1} \frac{J_t^2(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2}} \sqrt{\sum_{t=0}^{T-1} (\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2)}. \end{aligned}$$

Substituting the upper bound from (27) into the above inequality results in

$$\text{Reg}(T) \leq \sqrt{\frac{V_0}{2\mu} \left(\sum_{t=0}^{T-1} (\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2) \right)},$$

which completes the proof. $\square$

Remark 2 (Interpretation of the regret bound). As discussed in the introduction, the choice of the mirror map f in the update law (7a) affects the regret bound (26) through the Bregman divergence D_f , the dual norm $\|\cdot\|_*$, and the strong convexity parameter μ . Since f may be normalized, without loss of generality, to be 1-strongly convex, the quantities of primary interest are the Bregman divergence D_f and the dual norm $\|\cdot\|_*$ associated with the norm relative to which f is strongly convex. The main idea underlying the dimension-independent regret bounds developed in the next two sections is to choose f and the underlying norm so that the corresponding dual norm does not introduce explicit dimension-dependent factors. Examples include the ℓ_∞ norm, the spectral norm, and their appropriate analogues. The dimension dependence of the Bregman term $D_f(\Theta, \hat{\Theta}_0; Z_0)$ can then be controlled through a suitable initialization, such as zero initialization when exploiting sparse structure.

IV. SPARSE STRUCTURE ON Θ

In this section, we assume a given sparsity structure on the unknown matrix $\Theta \in \Omega = \mathbb{R}^{m \times k}$. We consider three settings: entrywise sparsity, group sparsity, and low-rank structure. In each case, we first specify an appropriate mirror map function and derive the explicit form of the update law (7a). Moreover, using the general regret bound established in Theorem 3, we derive corresponding bounds for each sparsity structure.

A. Entrywise sparsity.

We consider the case where the unknown matrix $\Theta \in \mathbb{R}^{m \times k}$ is entrywise s -sparse, i.e., it has at most s nonzero entries. In this setting, we equip $\mathbb{R}^{m \times k}$ with the entrywise ℓ_1 -norm, $\|\cdot\| = \|\cdot\|_1$, whose dual norm is the entrywise ℓ_∞ -norm, $\|\cdot\|_* = \|\cdot\|_\infty$. Motivated by proximal setups for ℓ_1 -based sparse recovery [6, 7], we use the update law (7) with the mirror map

$$f_{\ell_p}(\hat{\Theta}) = \frac{d^{2-2/p}}{2(p-1)} \|\hat{\Theta}\|_p^2, \quad p = 1 + \frac{1}{\ln(d)}. \quad (28)$$

Note that, for $1 < p \leq 2$, the function $\hat{\Theta} \mapsto \frac{1}{2} \|\hat{\Theta}\|_p^2$ is $(p-1)$ -strongly convex with respect to the norm $\|\cdot\|_p$. Moreover, since $\|\hat{\Theta}\|_p \geq d^{\frac{1}{p}-1} \|\hat{\Theta}\|_1$, it follows that f_{ℓ_p} is 1-strongly convex with respect to the entrywise ℓ_1 -norm. The gradient of f_{ℓ_p} is given by

$$\nabla f_{\ell_p}(\hat{\Theta}) = \frac{d^{2-2/p}}{(p-1)} \|\hat{\Theta}\|_p^{2-p} \text{sign}(\hat{\Theta}) \odot |\hat{\Theta}|^{p-1}.$$

Here, $\odot$ denotes the Hadamard product. Let q be the dual exponent of p , defined by $q := p/(p-1) = 1 + \ln(d)$. The inverse gradient map is

$$\nabla f^*(Z) = (\nabla f_{\ell_p})^{-1}(Z) = \frac{p-1}{d^{2-2/p}} \|Z\|_q^{2-q} \text{sign}(Z) \odot |Z|^{q-1}$$

As a result, the update law (7) can be written explicitly in the primal space as

$$\begin{aligned} \hat{\Theta}_{t+1} &= \nabla f_{\ell_p}^* \left(\nabla f_{\ell_p}(\hat{\Theta}_t) - \eta_{t+1} \nabla J_t(\hat{\Theta}_t) \right), \\ \eta_{t+1} &= \frac{2 J_t(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2}. \end{aligned} \quad (29)$$

Applying the general regret bound in (26) to the mirror map (28) yields the following result.

Corollary 4 (Regret bound under entrywise sparsity). *Consider the system (1) under the control law (4) and the update law (29). Suppose that the parameter estimate is initialized as $\hat{\Theta}_0 = 0$ and that the system parameter matrix Θ has at most s nonzero entries. Then,*

$$\text{Reg}(T) \leq \frac{e}{2} s^{\frac{\ln(d)}{1+\ln(d)}} \|\Theta\|_\infty \sqrt{\ln(d) \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2 \right)}.$$

Proof. Since $\hat{\Theta}_0 = 0$, we have

$$D_f(\Theta, \hat{\Theta}_0; Z_0) = D_f(\Theta, 0; 0) = f(\Theta) = \frac{d^{2-2/p}}{2(p-1)} \|\Theta\|_p^2,$$

where we used $f(0) = 0$ and $\nabla f(0) = 0$. Moreover, since Θ has at most s nonzero entries, $\|\Theta\|_p \leq s^{1/p} \|\Theta\|_\infty$. Therefore,

$$D_f(\Theta, \hat{\Theta}_0; Z_0) \leq \frac{d^{2-2/p}}{2(p-1)} s^{2/p} \|\Theta\|_\infty^2 \leq \frac{e^2 \ln(d)}{2} s^{\frac{2 \ln(d)}{1+\ln(d)}} \|\Theta\|_\infty^2,$$

where the second inequality follows by setting $p = 1 + 1/\ln(d)$ and noting that $d^{2-2/p} \leq e^2$.

Applying the general regret bound in (26) with $\mu = 1$ and $\|\cdot\|_* = \|\cdot\|_\infty$ yields

$$\text{Reg}(T) \leq \sqrt{\frac{D_f(\Theta, \hat{\Theta}_0; Z_0)}{2\mu} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2 \right)} \leq \frac{e}{2} s^{\frac{\ln(d)}{1+\ln(d)}} \|\Theta\|_\infty \sqrt{\ln(d) \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2 \right)},$$

which proves the claim. $\square$

B. Group sparsity

We consider the case where the unknown matrix $\Theta \in \mathbb{R}^{m \times k}$ has a *group-sparse* structure. Specifically, the entries of Θ are partitioned into d_{grp} disjoint groups $\Theta_1, \Theta_2, \dots, \Theta_{d_{\text{grp}}}$, and Θ is said to be *s-group sparse* if at most s groups are nonzero, i.e., $|\{i : \Theta_i \neq 0\}| \leq s$. In this setting, we equip $\mathbb{R}^{m \times k}$ with the mixed ℓ_1/ℓ_2 norm $\|\Theta\|_{1,2} := \sum_{i=1}^{d_{\text{grp}}} \|\Theta_i\|_2$, whose dual norm is $\|Z\|_{\infty,2} := \max_{1 \leq i \leq d_{\text{grp}}} \|Z_i\|_2$. We apply the update law (7) with the block mirror map

$$f_{p,2}(\hat{\Theta}) := \frac{d_{\text{grp}}^{2-2/p}}{2(p-1)} \|\hat{\Theta}\|_{p,2}^2, \quad p = 1 + \frac{1}{\ln(d_{\text{grp}})}, \quad (30)$$

where $\|\hat{\Theta}\|_{p,2} := \left(\sum_{i=1}^{d_{\text{grp}}} \|\hat{\Theta}_i\|_2^p \right)^{1/p}$. Following the same line of reasoning as in the previous subsection, it follows that $f_{p,2}$ is 1-strongly convex with respect to the norm $\|\cdot\|_{1,2}$. The gradient of $f_{p,2}$ is given blockwise by

$$(\nabla f_{p,2}(\hat{\Theta}))_i = \frac{d_{\text{grp}}^{2-2/p}}{(p-1)} \|\hat{\Theta}\|_{p,2}^{2-p} \|\hat{\Theta}_i\|_2^{p-2} \hat{\Theta}_i, \quad i = 1, \dots, d_{\text{grp}}.$$

Let q be the dual exponent of p . The inverse gradient map is

$$(\nabla f_{p,2}^*(Z))_i = \frac{(p-1)}{d_{\text{grp}}^{2-2/p}} \|Z\|_{q,2}^{2-q} \|Z_i\|_2^{q-2} Z_i, \quad i = 1, \dots, d_{\text{grp}}.$$

Therefore, the update law (7) can be written equivalently as

$$\begin{aligned} \hat{\Theta}_{t+1} &= \nabla f_{p,2}^* \left(\nabla f_{p,2}(\hat{\Theta}_t) - \eta_{t+1} \nabla J_t(\hat{\Theta}_t) \right), \\ \eta_{t+1} &= \frac{2 J_t(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_{\infty,2}^2}. \end{aligned} \quad (31)$$

The following corollary provides a regret bound for the group sparsity structure that depends only logarithmically on d_{grp} .

Corollary 5 (Regret bound for group sparsity). *Consider the system (1) under the control law (4) and the update law (31). Suppose that the update is initialized at $\hat{\Theta}_0 = 0$ and the system parameter matrix Θ is s-group sparse. Then the regret satisfies*

$$\text{Reg}(T) \leq \frac{e}{2} s^{\frac{\ln(d_{\text{grp}})}{1+\ln(d_{\text{grp}})}} \|\Theta\|_{\infty,2} \sqrt{\ln(d_{\text{grp}}) \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_{\infty,2}^2 \right)}.$$

Proof. Since the update is initialized at $\hat{\Theta}_0 = 0$, we have

$$D_f(\Theta, \hat{\Theta}_0; Z_0) = f_{p,2}(\Theta) = \frac{d_{\text{grp}}^{2-2/p}}{2(p-1)} \|\Theta\|_{p,2}^2 \leq \frac{d_{\text{grp}}^{2-2/p}}{2(p-1)} s^{2/p} \|\Theta\|_{\infty,2}^2 \leq \frac{e^2}{2} \ln(d_{\text{grp}}) s^{\frac{2 \ln(d_{\text{grp}})}{1+\ln(d_{\text{grp}})}} \|\Theta\|_{\infty,2}^2,$$

where the first inequality follows from the s -group sparsity of Θ . As a result, employing the general regret bound in (26) with $\mu = 1$ and $\|\cdot\|_* = \|\cdot\|_{\infty,2}$ yields

$$\text{Reg}(T) \leq \sqrt{\frac{D_f(\Theta, \hat{\Theta}_0; Z_0)}{2\mu} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2 \right)} \leq \frac{e}{2} s^{\frac{\ln(d_{\text{grp}})}{1+\ln(d_{\text{grp}})}} \|\Theta\|_{\infty,2} \sqrt{\ln(d_{\text{grp}}) \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_{\infty,2}^2 \right)},$$

which completes the proof. $\square$

C. Low-rank matrix structure.

In this subsection, we consider the case where the unknown matrix $\Theta \in \mathbb{R}^{m \times k}$ is low rank. Specifically, we assume that $\text{rank}(\Theta) \leq s$. Let $d_{\text{sv}} := \min(m, k)$, which is the number of singular values of a matrix in $\mathbb{R}^{m \times k}$. Our goal is to develop an update law whose regret depends on the intrinsic rank s , rather than on the ambient singular-value dimension d_{sv} . To this end, we equip $\mathbb{R}^{m \times k}$ with the nuclear norm $\|\cdot\| = \|\cdot\|_{\text{S}_1}$, whose dual norm is the spectral norm $\|\cdot\|_* = \|\cdot\|_{\text{S}_\infty}$. Motivated by the sparse-vector construction in the previous subsection, we use the update law (7) with the Schatten- p mirror map

$$f_{\text{S}_p}(\hat{\Theta}) = \frac{d_{\text{sv}}^{2-2/p}}{2(p-1)} \|\hat{\Theta}\|_{\text{S}_p}^2, \quad p = 1 + \frac{1}{\ln(d_{\text{sv}})}. \quad (32)$$

The function f_{S_p} is 1-strongly convex with respect to the nuclear norm $\|\cdot\|_{\text{S}_1}$. Let $\hat{\Theta} = U\Sigma V^\top$ be a singular value decomposition (SVD) of $\hat{\Theta}$. The gradient of f_{S_p} is given by

$$\nabla f_{\text{S}_p}(\hat{\Theta}) = \frac{d_{\text{sv}}^{2-2/p}}{(p-1)} \|\hat{\Theta}\|_{\text{S}_p}^{2-p} U \Sigma^{p-1} V^\top.$$

Let q be the dual exponent of p . For any $Z \in \mathbb{R}^{m \times k}$ with SVD $Z = \bar{U} \bar{\Sigma} \bar{V}^\top$, the inverse gradient map is

$$\nabla f_{\text{S}_p}^*(Z) = (\nabla f_{\text{S}_p})^{-1}(Z) = \frac{(p-1)}{d_{\text{sv}}^{2-2/p}} \|Z\|_{\text{S}_q}^{2-q} \bar{U} \bar{\Sigma}^{q-1} \bar{V}^\top.$$

Consequently, the parameter update (7) can be written as

$$\begin{aligned} \hat{\Theta}_{t+1} &= \nabla f_{\text{S}_p}^* \left(\nabla f_{\text{S}_p}(\hat{\Theta}_t) - \eta_{t+1} \nabla J_t(\hat{\Theta}_t) \right), \\ \eta_{t+1} &= \frac{2 J_t(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_{\text{S}_\infty}^2}. \end{aligned} \quad (33)$$

The following corollary provides a regret bound for the low-rank structure that depends only logarithmically on d_{sv} .

Corollary 6 (Regret bound for low-rank structure). *Consider the system (1) under the control law (4) and the update law (33). Suppose that the update is initialized at $\hat{\Theta}_0 = 0$ and the system parameter matrix Θ is of rank at most s . Then,*

$$\text{Reg}(T) \leq \frac{e}{2} s^{\frac{\ln(d_{\text{sv}})}{1+\ln(d_{\text{sv}})}} \|\Theta\|_{\text{S}_\infty} \sqrt{\ln(d_{\text{sv}}) \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_{\text{S}_\infty}^2 \right)}.$$

Proof. Since the update is initialized at $\hat{\Theta}_0 = 0$, we obtain

$$D_f(\Theta, \hat{\Theta}_0; Z_0) = f(\Theta) = \frac{d_{\text{sv}}^{2-2/p}}{2(p-1)} \|\Theta\|_{\text{S}_p}^2 \leq \frac{e^2}{2} \ln(d_{\text{sv}}) \|\Theta\|_{\text{S}_p}^2$$

Moreover, since Θ is of rank at most s , we find that $\|\Theta\|_{\text{S}_p} \leq s^{1/p} \|\Theta\|_{\text{S}_\infty}$, hence

$$D_f(\Theta, \hat{\Theta}_0; Z_0) \leq \frac{e^2}{2} \ln(d_{\text{sv}}) s^{2/p} \|\Theta\|_{\text{S}_\infty}^2 = \frac{e^2}{2} \ln(d_{\text{sv}}) s^{\frac{2 \ln(d_{\text{sv}})}{1+\ln(d_{\text{sv}})}} \|\Theta\|_{\text{S}_\infty}^2.$$

As a result, employing the general regret bound in (26) with $\mu = 1$ and $\|\cdot\|_* = \|\cdot\|_{\text{S}_\infty}$ yields

$$\text{Reg}(T) \leq \sqrt{\frac{D_f(\Theta, \hat{\Theta}_0; Z_0)}{2\mu} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2 \right)} \leq \frac{e}{2} \|\Theta\|_{\text{S}_\infty} s^{\frac{\ln(d_{\text{sv}})}{1+\ln(d_{\text{sv}})}} \sqrt{\ln(d_{\text{sv}}) \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_{\text{S}_\infty}^2 \right)},$$

as claimed. $\square$

V. SIMPLEX-TYPE SETS AND ENTROPIC POTENTIALS

In this section, we consider the case in which the unknown parameter set $\Omega \subset \mathbb{R}^{m \times k}$ has a simplex-type structure. Specifically, we study two representative examples: the standard probability simplex and the spectraplex, that is, the set of density matrices. For each setting, we identify a suitable mirror map and derive the explicit form of the update law (7). We then apply the general regret bound of Theorem 3 to obtain the corresponding regret guarantees and show that they are essentially dimension-free.

A. Simplex and Shannon Entropy

We consider the case where the unknown matrix Θ belongs to the entrywise probability simplex

$$\Omega_e := \left\{ \hat{\Theta} \in \mathbb{R}_+^{m \times k} : \|\hat{\Theta}\|_1 = 1 \right\}. \quad (34)$$

In this setting, we equip $\mathbb{R}^{m \times k}$ with the entrywise ℓ_1 -norm, $\|\cdot\| = \|\cdot\|_1$, whose dual norm is the entrywise ℓ_∞ -norm, $\|\cdot\|_* = \|\cdot\|_\infty$. For the update law (7), we employ the Shannon entropy plus the indicator function of Ω_e as the mirror map

$$f_{\text{Sh}}(\hat{\Theta}) = \sum_{i,j} (\hat{\Theta}_{ij} \ln(\hat{\Theta}_{ij}) - \hat{\Theta}_{ij}) + \iota_{\Omega_e}(\hat{\Theta}),$$

where ι_{Ω_e} denotes the indicator function of Ω_e , i.e.,

$$\iota_{\Omega_e}(\hat{\Theta}) = \begin{cases} 0, & \hat{\Theta} \in \Omega_e, \\ +\infty, & \text{otherwise.} \end{cases}$$

By Pinsker's inequality, f_{Sh} is 1-strongly convex on Ω_e with respect to the ℓ_1 -norm [10, 11]. Moreover, for any $\hat{\Theta} \in \text{ri}(\Omega_e) = \left\{ \hat{\Theta}_{ij} > 0 : \|\hat{\Theta}\|_1 = 1 \right\}$, its subdifferential admits the characterization

$$\partial f_{\text{Sh}}(\hat{\Theta}) = \left\{ \ln_{\text{el}}(\hat{\Theta}) + \lambda \mathbf{1} : \lambda \in \mathbb{R} \right\}, \quad (35)$$

where $\mathbf{1}$ denotes the all-ones matrix and $\ln_{\text{el}}(\cdot)$ denotes the entrywise logarithm. The conjugate of f_{Sh} is given by

$$f_{\text{Sh}}^*(Z) = \ln \left(\sum_{i,j} \exp(Z_{ij}) \right) + 1,$$

which is everywhere differentiable, with gradient

$$\nabla f_{\text{Sh}}^*(Z) = \frac{\exp_{\text{el}}(Z)}{\sum_{i,j} \exp(Z_{ij})},$$

where $\exp_{\text{el}}(\cdot)$ denotes the entrywise exponential.

Applying the update law (7), any choice of $Z_t \in \partial f_{\text{Sh}}(\hat{\Theta}_t)$ can be written as $Z_t = \ln_{\text{el}}(\hat{\Theta}_t) + \lambda_t \mathbf{1}$ for some $\lambda_t \in \mathbb{R}$. The dual update (7a) thus yields

$$Z_{t+1} = \ln_{\text{el}}(\hat{\Theta}_t) - \eta_{t+1} \nabla J_t(\hat{\Theta}_t) + \lambda_t \mathbf{1}.$$

Applying the primal recovery $\hat{\Theta}_{t+1} = \nabla f_{\text{Sh}}^*(Z_{t+1})$ yields the following parameter update law

$$\begin{aligned} \hat{\Theta}_{t+1} &= \frac{\hat{\Theta}_t \odot \exp_{\text{el}}(-\eta_{t+1} \nabla J_t(\hat{\Theta}_t))}{\|\hat{\Theta}_t \odot \exp_{\text{el}}(-\eta_{t+1} \nabla J_t(\hat{\Theta}_t))\|_1}, \\ \eta_{t+1} &= \frac{2 J_t(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2}. \end{aligned} \quad (36)$$

where $\odot$ denotes the Hadamard product. Note that the above update law is independent of the scalar λ_t . The following corollary provides the resulting regret bound.

Corollary 7 (Regret bound over the simplex). *Consider the system (1) under the control law (4) and the parameter update law (36). Suppose that the update is initialized at $\hat{\Theta}_0 = \mathbf{1}/d$ and that the system parameter matrix belongs to the simplex Ω_e defined in (34). Then, the regret satisfies*

$$\text{Reg}(T) \leq \sqrt{\frac{\ln(d)}{2} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2 \right)}. \quad (37)$$

Proof. Let $Z_0 \in \partial f_{\text{Sh}}(\hat{\Theta}_0)$. Since $\hat{\Theta}_0 = \mathbf{1}/d \in \text{ri}(\Omega_e)$, by (35) we may write $Z_0 = \ln_{\text{el}}(\hat{\Theta}_0) + \lambda_0 \mathbf{1} = -\ln(d)\mathbf{1} + \lambda_0 \mathbf{1}$ for some $\lambda_0 \in \mathbb{R}$. Hence Z_0 is a constant multiple of $\mathbf{1}$. Since $\Theta, \hat{\Theta}_0 \in \Omega_e$, we have $\langle \Theta - \hat{\Theta}_0, \mathbf{1} \rangle = 0$, and therefore $\langle \Theta - \hat{\Theta}_0, Z_0 \rangle = 0$. It follows from the definition of the generalized Bregman divergence that

$$\begin{aligned} D_{f_{\text{Sh}}}(\Theta, \hat{\Theta}_0; Z_0) &= f_{\text{Sh}}(\Theta) - f_{\text{Sh}}(\hat{\Theta}_0) - \langle \Theta - \hat{\Theta}_0, Z_0 \rangle = f_{\text{Sh}}(\Theta) - f_{\text{Sh}}(\hat{\Theta}_0) \\ &= \sum_{i,j} \Theta_{ij} \ln(\Theta_{ij}) + \ln(d) \leq \ln(d), \end{aligned} \quad (38)$$

where the last inequality follows from the convention $0 \ln 0 = 0$ and the fact that $x \ln x \leq 0, \forall x \in [0, 1]$. Consequently, applying the general regret bound in (26) with $\mu = 1$ and $\|\cdot\|_* = \|\cdot\|_\infty$, and using (38), yields

$$\text{Reg}(T) \leq \sqrt{\frac{D_f(\Theta, \hat{\Theta}_0; Z_0)}{2\mu} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_*^2 \right)} \leq \sqrt{\frac{\ln(d)}{2} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2 \right)},$$

which proves the claim. $\square$

Remark 3 (Row-stochastic simplex). The same construction applies when Θ is row-stochastic, i.e.,

$$\Omega_r := \left\{ \hat{\Theta} \in \mathbb{R}_+^{m \times k} : \sum_{j=1}^k \hat{\Theta}_{ij} = 1, i = 1, \dots, m \right\}. \quad (39)$$

In this case, we use

$$f_{\text{Sh},r}(\hat{\Theta}) = \sum_{i=1}^m \sum_{j=1}^k \left(\hat{\Theta}_{ij} \ln(\hat{\Theta}_{ij}) - \hat{\Theta}_{ij} \right) + \iota_{\Omega_r}(\hat{\Theta}).$$

For any $\Theta, \Theta' \in \Omega_r$ and $Z' \in \partial f_{\text{Sh},r}(\Theta')$, the Bregman divergence reduces to the sum of row-wise KL divergences,

$$D_{f_{\text{Sh},r}}(\Theta, \Theta'; Z') = \sum_{i=1}^m D_{\text{KL}}(\Theta_{i,:}, \Theta'_{i,:}) \geq \frac{1}{2m} \|\Theta - \Theta'\|_1^2,$$

where the inequality follows from Pinsker's and Cauchy's inequalities. Thus, $f_{\text{Sh},r}$ is $1/m$ -strongly convex on Ω_r with respect to $\|\cdot\|_1$. The update becomes the row-wise normalized multiplicative rule

$$\begin{aligned} \hat{\Theta}_{t+1} &= \mathcal{N}_r \left(\hat{\Theta}_t \odot \exp_{\text{el}} \left(-\eta_{t+1} \nabla J_t(\hat{\Theta}_t) \right) \right), \\ \eta_{t+1} &= \frac{2}{m} \frac{J_t(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2}, \end{aligned} \quad (40)$$

where $\mathcal{N}_r$ denotes row-wise normalization. If $\hat{\Theta}_0 = \mathbf{1}_{m \times k}/k$, then, for any $Z_0 \in \partial f_{\text{Sh},r}(\hat{\Theta}_0)$,

$$D_{f_{\text{Sh},r}}(\Theta, \hat{\Theta}_0; Z_0) = \sum_{i=1}^m \sum_{j=1}^k \Theta_{ij} \ln(k\Theta_{ij}) \leq m \ln(k).$$

Theorem 3, with $\mu = 1/m$ and $\|\cdot\|_* = \|\cdot\|_\infty$, gives

$$\text{Reg}(T) \leq \sqrt{\frac{m^2 \ln(k)}{2} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\nabla J_t(\hat{\Theta}_t)\|_\infty^2 \right)}.$$

B. Spectraplex and von Neumann Entropy

We consider the case where the unknown matrix Θ belongs to the spectraplex (the set of density matrices), i.e.,

$$\Omega_s := \left\{ \hat{\Theta} \in \mathbb{S}_+^k : \text{Tr}(\hat{\Theta}) = 1 \right\}. \quad (41)$$

In this subsection, the parameter matrix is square, i.e., $m = k$. In this setting, we equip $\mathbb{S}^k$ with the nuclear norm $\|\cdot\|_{\mathbb{S}_1}$, whose dual norm is the operator norm $\|\cdot\|_{\mathbb{S}_\infty}$. On Ω_s , we employ the update law (7) with the von Neumann entropy mirror map

$$f_{\text{vN}}(\hat{\Theta}) := \text{Tr}(\hat{\Theta} \ln(\hat{\Theta}) - \hat{\Theta}) + \iota_{\Omega_s}(\hat{\Theta}), \quad (42)$$

where ι_{Ω_s} denotes the indicator function of Ω_s . By the quantum Pinsker inequality, f_{vN} is 1-strongly convex on Ω_s with respect to $\|\cdot\|_{\mathbb{S}_1}$. Moreover, for any Θ in the relative interior $\text{ri}(\Omega) = \{\Theta \succ 0 : \text{Tr}(\Theta) = 1\}$, its subdifferential admits the characterization $\partial f_{\text{vN}}(\Theta) = \{\ln(\Theta) + \lambda I : \lambda \in \mathbb{R}\}$, where I denotes the identity matrix. The conjugate of f_{vN} is given by

$$f_{\text{vN}}^*(Z) = \ln(\text{Tr}(\exp(Z))) + 1,$$

which is everywhere differentiable, with gradient

$$\nabla f_{\text{vN}}^*(Z) = \frac{\exp(Z)}{\text{Tr}(\exp(Z))},$$

where $\exp(\cdot)$ denotes the matrix exponential.

To preserve the symmetry of the parameter estimates, we use the symmetric part of the gradient, defined by

$$\text{Sym}(\nabla J_t(\hat{\Theta}_t)) := \frac{1}{2} \left(\nabla J_t(\hat{\Theta}_t) + \nabla J_t(\hat{\Theta}_t)^\top \right).$$

Indeed, for every $H \in \mathbb{S}^k$, we have

$$\begin{aligned} \langle \text{Sym}(\nabla J_t(\hat{\Theta}_t)), H \rangle &= \frac{1}{2} \langle \nabla J_t(\hat{\Theta}_t), H \rangle + \frac{1}{2} \langle \nabla J_t(\hat{\Theta}_t)^\top, H \rangle = \frac{1}{2} \langle \nabla J_t(\hat{\Theta}_t), H \rangle + \frac{1}{2} \langle \nabla J_t(\hat{\Theta}_t), H^\top \rangle \\ &= \langle \nabla J_t(\hat{\Theta}_t), H \rangle. \end{aligned}$$

Thus, when $\Theta, \hat{\Theta}_t \in \mathbb{S}^k$, replacing the gradient by its symmetric part does not change the key inner-product identity (20) used in the stability and regret analyses of Section III.

Applying the update law (7a), any choice of $Z_t \in \partial f_{\text{vN}}(\hat{\Theta}_t)$ can be written as $Z_t = \ln \hat{\Theta}_t + \lambda_t I$ for some $\lambda_t \in \mathbb{R}$. The dual update $Z_{t+1} = Z_t - \eta_{t+1} \text{Sym}(\nabla J_t(\hat{\Theta}_t))$ thus yields

$$Z_{t+1} = \ln(\hat{\Theta}_t) - \eta_{t+1} \text{Sym}(\nabla J_t(\hat{\Theta}_t)) + \lambda_t I.$$

Applying the primal recovery $\hat{\Theta}_{t+1} = \nabla f_{\text{vN}}^*(Z_{t+1})$ gives

$$\begin{aligned} \hat{\Theta}_{t+1} &= \frac{\exp(\ln(\hat{\Theta}_t) - \eta_{t+1} \text{Sym}(\nabla J_t(\hat{\Theta}_t)))}{\text{Tr}(\exp(\ln(\hat{\Theta}_t) - \eta_{t+1} \text{Sym}(\nabla J_t(\hat{\Theta}_t))))}, \\ \eta_{t+1} &= \frac{2 J_t(\hat{\Theta}_t)}{\epsilon_{t+1} + \|\text{Sym}(\nabla J_t(\hat{\Theta}_t))\|_{\mathbb{S}_\infty}^2}. \end{aligned} \quad (43)$$

The following corollary provides the resulting regret bound.

Corollary 8 (Regret bound over the spectraplex). *Consider the closed-loop system (1) under the control law (4) and the update law (43). Suppose that $\Theta \in \Omega_s$ and $\hat{\Theta}_0 = I/k$. Then*

$$\text{Reg}(T) \leq \sqrt{\frac{\ln(k)}{2} \sum_{t=0}^{T-1} \left(\epsilon_{t+1} + \|\text{Sym}(\nabla J_t(\hat{\Theta}_t))\|_{S_\infty}^2 \right)}. \quad (44)$$

Proof. Let $Z_0 \in \partial f_{\text{vN}}(\hat{\Theta}_0)$. Since $\hat{\Theta}_0 = I/k$, we have $Z_0 = \ln(\hat{\Theta}_0) + \lambda_0 I = (-\ln(k) + \lambda_0)I$ for some $\lambda_0 \in \mathbb{R}$. Hence, since $\Theta, \hat{\Theta}_0 \in \Omega_s$, we have $\langle \Theta - \hat{\Theta}_0, Z_0 \rangle = 0$. Therefore,

$$D_{f_{\text{vN}}}(\Theta, \hat{\Theta}_0; Z_0) = f_{\text{vN}}(\Theta) - f_{\text{vN}}(\hat{\Theta}_0) - \langle \Theta - \hat{\Theta}_0, Z_0 \rangle = \text{Tr}(\Theta \ln(\Theta)) + \ln(k) \leq \ln(k), \quad (45)$$

where the last inequality follows from $\text{Tr}(\Theta \ln(\Theta)) = \sum_i \alpha_i \ln(\alpha_i) \leq 0$, with $\{\alpha_i\}$ denoting the positive eigenvalues of Θ . Applying (26) with $\mu = 1$, $\|\cdot\|_* = \|\cdot\|_{S_\infty}$, replacing the $\nabla J_t(\hat{\Theta}_t)$ with $\text{Sym}(\nabla J_t(\hat{\Theta}_t))$ and using (45) yields (44). $\square$

VI. EXTENSION TO STOCHASTIC SYSTEMS

In this section, we briefly discuss the effect of process noise on the proposed framework. Consider the stochastic system

$$X_{t+1} = \mathcal{A}(X_t) + \mathcal{B}(\Theta \Psi(X_t, t) + U_t) + W_t, \quad (46)$$

where $W_t \in \mathbb{R}^{n_1 \times n_2}$ is an independent, zero-mean white-noise process, independent of the initial state and parameter estimate. Under the control law (4), the residual becomes

$$\tilde{X}_{t+1} := X_{t+1} - (\mathcal{A} - \mathcal{BK})(X_t) - \mathcal{B}(U_t^d) = \mathcal{B}((\Theta - \hat{\Theta}_t)\Psi(X_t, t)) + W_t.$$

For each realization of X_t , define the instantaneous loss

$$J_t(\hat{\Theta}; X_t) := \frac{1}{2} \left\| \mathcal{B}((\Theta - \hat{\Theta})\Psi(X_t, t)) \right\|_F^2,$$

Conditional on X_t , this is a deterministic function of $\hat{\Theta}$. The corresponding expected regret is defined as

$$\text{Reg}(T) := \sum_{t=0}^{T-1} \mathbb{E} J_t(\hat{\Theta}_t; X_t), \quad (47)$$

where the expectation is taken over the randomness induced by the process noise. Even in the noisy setting, a controller of the form (4) with full knowledge of the true parameter Θ , namely $U_t(X_t, \Theta)$, incurs zero regret because $J_t(\Theta; X_t) = 0$ for every realization of X_t .

After observing X_{t+1} , define the noisy gradient estimator

$$\hat{G}_t := -\mathcal{B}^\top(\tilde{X}_{t+1})\Psi(X_t, t)^\top.$$

Using the conditional zero-mean property of W_t , we obtain

$$\begin{aligned} \mathbb{E}[\hat{G}_t \mid X_t, \hat{\Theta}_t] &= -\mathcal{B}^\top \mathbb{E}[\tilde{X}_{t+1} \mid X_t, \hat{\Theta}_t] \Psi(X_t, t)^\top = -\mathcal{B}^\top \mathcal{B}((\Theta - \hat{\Theta}_t)\Psi(X_t, t)) \Psi(X_t, t)^\top \\ &= \nabla J_t(\hat{\Theta}_t; X_t). \end{aligned}$$

Thus, the update law (7), when driven by $\hat{G}_t$, can be viewed as a stochastic online mirror-descent scheme with an unbiased conditional gradient oracle. The corresponding Polyak-type step size takes the form

$$\eta_{t+1} = \frac{\mu \|\tilde{X}_{t+1}\|_F^2}{\epsilon_{t+1} + \|\hat{G}_t\|_*^2}. \quad (48)$$

The preceding discussion motivates the following question: To what extent can the stability and regret guarantees of Section III be recovered in the presence of process noise?

Empirical observations. In Section VII, we empirically demonstrate that the proposed method remains stable and continues to exhibit essentially dimension-independent performance in the presence of process noise, while significantly outperforming Euclidean-based adaptive schemes. In contrast, recursive

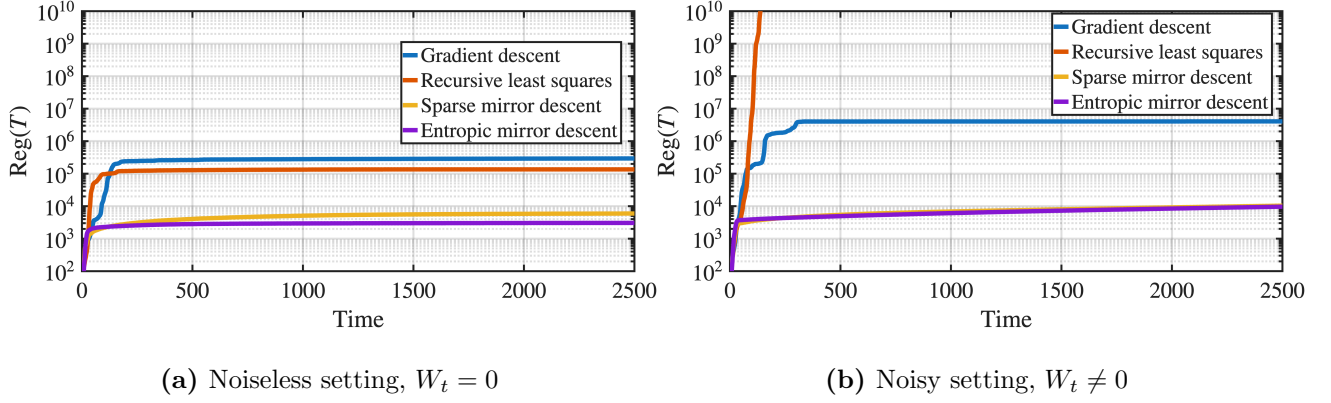


FIGURE 2. Evolution of the regret $\text{Reg}(T)$ for the system (49) under the control law (51) with four different parameter-update algorithms in the noiseless and noisy settings.

least-squares-based adaptive controllers may become ill-conditioned and exhibit unstable behavior in high-dimensional noisy settings.

Robust modifications. Standard robustification mechanisms, such as dead-zone modifications or projection-based updates, can also be incorporated to guarantee boundedness of the parameter estimates $\hat{\Theta}_t$. For instance, the constrained mirror maps used in Section V provide a natural way to enforce such boundedness through the inclusion of a set-indicator function, which keeps the iterates within a prescribed constraint set.

Theoretical challenges. A complete theoretical treatment, however, requires overcoming two key obstacles. First, the stochastic closed-loop system must be shown to satisfy appropriate stability or moment bounds without imposing uniformly bounded regressors. This is important because the deterministic analysis relies only on the linear-growth condition (2), which is more natural for dynamical systems but more delicate under process noise. Second, the Polyak-type step size (48) introduces additional technical difficulties. Indeed, in the noisy setting, the step size depends on the random residual $\tilde{X}_{t+1}$, and therefore explicitly depends on the disturbance W_t . Consequently, the step size and the gradient estimator become statistically coupled, preventing a direct application of standard unbiased stochastic approximation arguments. Therefore, extending the regret bound (26) to the stochastic setting likely requires either an alternative step-size mechanism, such as predictable or delayed step sizes, or a more refined stochastic analysis. We leave such developments for future work. Nevertheless, the unbiased-gradient structure derived above, together with the empirical evidence in Section VII, suggests that the canonical $O(\sqrt{T})$ stochastic regret rate, with logarithmic dependence on the ambient dimension for sparse and simplex-type parameter structures, should remain achievable under suitable stability and moment conditions.

VII. SIMULATIONS: ONLINE OPTIMIZATION WITH ADAPTIVE CONTROL

Consider a multi-agent system consisting of N agents, whose collective dynamics are governed by

$$X_{t+1} = \Theta \Psi(X_t, t) + U_t + W_t, \quad (49)$$

where $X_t \in \mathbb{R}^{n \times N}$ denotes the collective state at time t , $U_t \in \mathbb{R}^{n \times N}$ is the control input, and $W_t \in \mathbb{R}^{n \times N}$ is a stochastic disturbance. The matrix $\Theta \in \mathbb{R}^{n \times k}$ represents unknown system parameters, while the regressor $\Psi(X_t, t)$ accounts for the inherent internal dynamics, potential inter-agent coupling, and exogenous disturbances. Furthermore, let $\{\mathcal{F}_t : \mathbb{R}^{n \times N} \rightarrow \mathbb{R}_+\}_{t \geq 0}$ be a sequence of time-varying, $\mathcal{L}$ -smooth, and convex cost functions that characterize the performance objective at each time step.

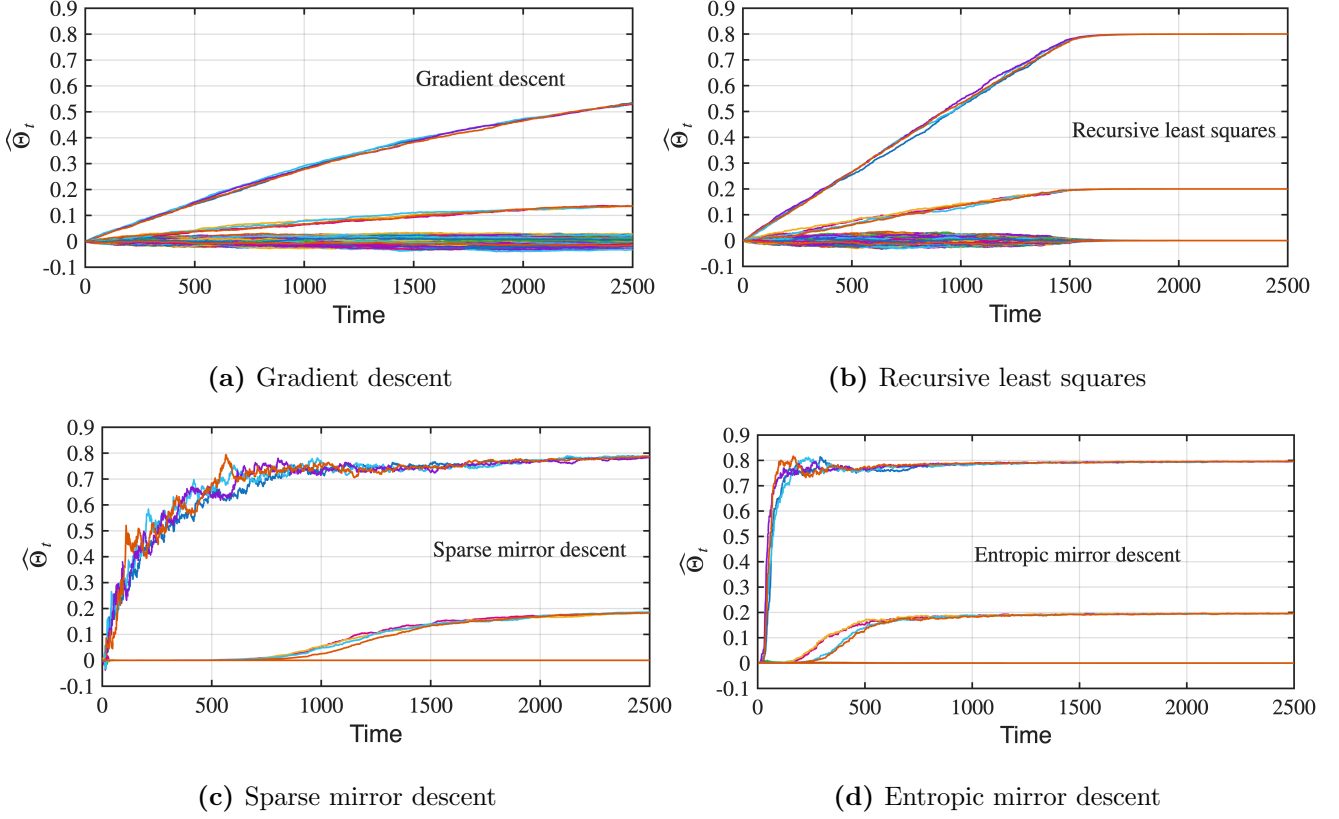


FIGURE 3. Parameter estimates in the noiseless setting, $W_t = 0$.

The control objective is to design an adaptive controller such that the state of the system (49) tracks the reference trajectory $\{X_t^d\}_{t \in \mathbb{Z}_+}$ produced by the online gradient descent iteration

$$X_{t+1}^d = X_t^d - \frac{1}{\mathcal{L}} \nabla \mathcal{F}_t(X_t^d).$$

For the numerical study, we consider the Laplacian-coupled quadratic objective

$$\mathcal{F}_t(X) = \frac{1}{2} \text{Tr}((X - C_t)^\top H(X - C_t)) + \frac{1}{2} \text{Tr}(X L_g X^\top), \quad (50)$$

where $H \in \mathbb{S}_+^n$ is a symmetric positive semidefinite matrix, possibly singular, $C_t \in \mathbb{R}^{n \times N}$ is a time-varying reference matrix, and $L_g \in \mathbb{R}^{N \times N}$ is the graph Laplacian of the communication network. The second term in (50) is the standard Laplacian regularizer

$$\text{Tr}(X L_g X^\top) = \frac{1}{2} \sum_{i,j=1}^N w_{ij} \|x_i - x_j\|_2^2,$$

which penalizes disagreement between neighboring agents. The gradient of (50) is given by

$$\nabla \mathcal{F}_t(X) = H(X - C_t) + X L_g.$$

Thus, the gradient is Lipschitz continuous with constant

$$\mathcal{L} = \lambda_{\max}(H) + \lambda_{\max}(L_g).$$

In the simulations, L_g is chosen as the Laplacian of an unweighted ring graph, namely,

$$(L_g)_{ij} = \begin{cases} 2, & i = j, \\ -1, & j = i \pm 1 \pmod{N}, \\ 0, & \text{otherwise,} \end{cases}$$

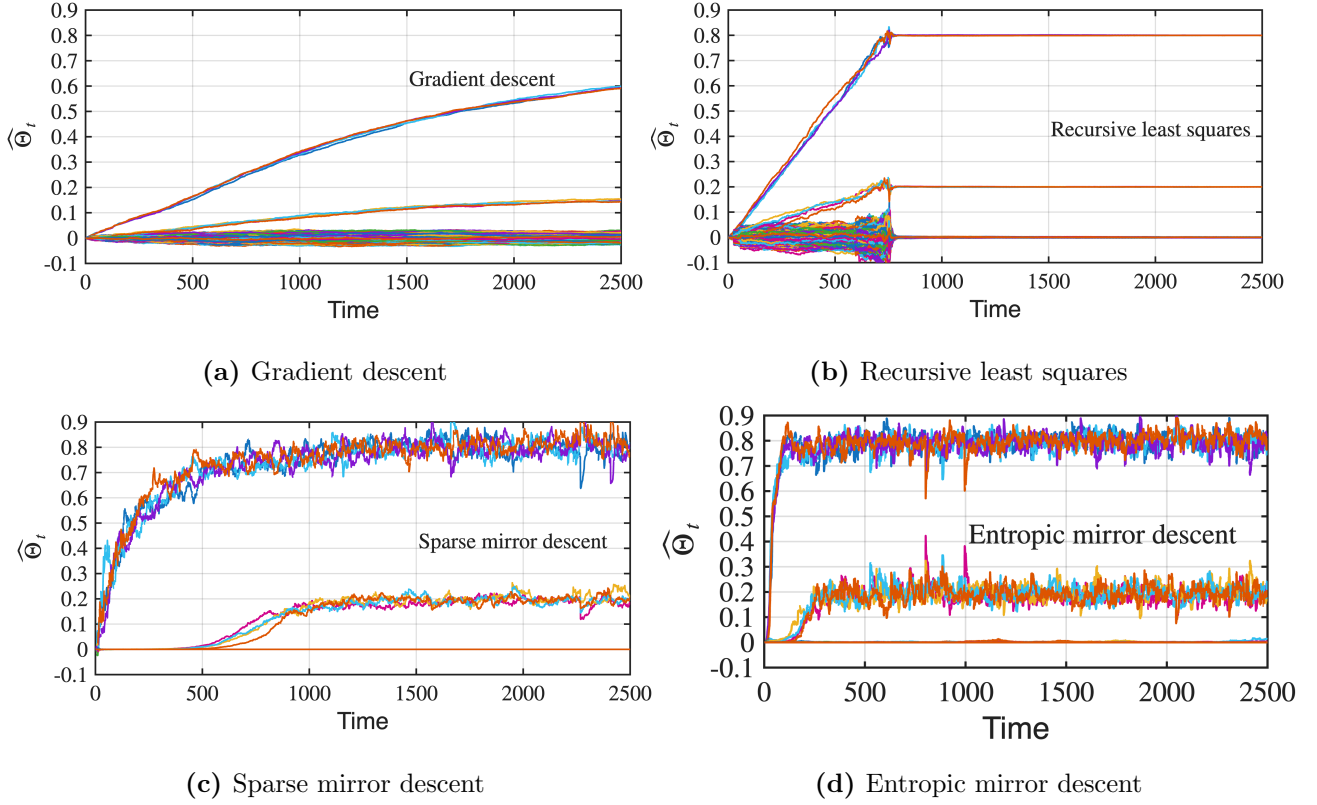


FIGURE 4. Parameter estimates in the noisy setting, $W_t \neq 0$.

so that each agent communicates only with its two immediate neighbors. Moreover, we set

$$H = \text{diag}(h_1, \dots, h_n) \in \mathbb{S}_+^n, \quad h_i = \begin{cases} 1, & i \text{ odd}, \\ 0, & i \text{ even}, \end{cases}$$

so that the tracking term penalizes only the odd-indexed state coordinates. Following the control law (4), we set

$$U_t = -\hat{\Theta}_t \Psi(X_t, t) + X_t^d - \frac{1}{\mathcal{L}} \nabla \mathcal{F}_t(X_t^d). \quad (51)$$

In the simulations, we take $n = N = 4$ and $k = 3000$. The reference signal is chosen as

$$(C_t)_{ij} = (-1)^{i+j} \sin(0.05t), \quad i = 1, \dots, n, \quad j = 1, \dots, N.$$

The feature map is set as

$$\Psi(X_t, t) = \Gamma(X_t \odot \sin(X_t) + \mathbf{1}_{n \times N}),$$

where $\Gamma \in \mathbb{R}^{k \times n}$ is a random feature matrix with independent Rademacher entries, i.e.,

$$\mathbb{P}(\Gamma_{ij} = 1) = \mathbb{P}(\Gamma_{ij} = -1) = \frac{1}{2}.$$

The true parameter matrix Θ is chosen to be sparse and row-stochastic. Specifically, for each row $i = 1, \dots, n$, we set

$$\Theta_{ij} = 0.8 \times \mathbf{1}_{\{j=i\}} + 0.2 \times \mathbf{1}_{\{j=k-i\}}, \quad j = 1, \dots, k.$$

where $\mathbf{1}_{\{\mathcal{E}\}}$ equals 1 if $\mathcal{E}$ holds and 0 otherwise. Thus, each row of Θ has two nonzero entries, whose values sum to one.

We consider two simulation scenarios: a noiseless setting with $W_t = 0$, and a noisy setting where the entries of W_t are generated independently from the uniform distribution on $[-1, 1]$. In both cases, we compare four parameter-estimation schemes: recursive least squares (RLS); gradient descent, corresponding to (7) with the mirror map $f(\hat{\Theta}) = \frac{1}{2} \|\hat{\Theta}\|_{\mathbb{F}}^2$; the row-wise entropic mirror descent update (40); and the sparse mirror descent update (29).

Figures 2(a) and 2(b) illustrate the evolution of the regret $\text{Reg}(T)$ in the noiseless and noisy settings, respectively. As observed, both the entropic and sparse mirror descent methods significantly outperform gradient descent and RLS. Furthermore, the results highlight varying degrees of robustness to process noise: while both entropic and sparse mirror descent maintain their performance, RLS degrades substantially. Specifically, although RLS outperforms gradient descent in the noiseless case, its regret grows rapidly in the presence of noise. This behavior suggests that, in high-dimensional regimes, first-order methods such as mirror descent can offer greater robustness than RLS. This disparity may be attributed to the fact that RLS relies on recursively updating an inverse covariance matrix, which is prone to ill-conditioning and heightened sensitivity to persistent noise in high dimensions.

Figures 3 and 4 show the corresponding parameter estimates obtained by each algorithm in the noiseless and noisy settings, respectively. The results demonstrate that the entropic and sparse mirror descent updates are better able to identify the low-dimensional sparse structure of the true parameter matrix Θ . In contrast, gradient descent and RLS tend to spread the estimation error across many coordinates, producing dense parameter estimates even though the true parameter matrix is sparse. It is also worth noting that, for entropic and sparse mirror descent, the step-size denominator depends on the ℓ_∞ -norm of the gradient, which is much smaller in high dimensions than the Frobenius norm utilized in gradient descent. Consequently, the resulting step size for gradient descent can be considerably smaller in high-dimensional regimes, causing its estimates to evolve at a much slower rate.

VIII. CONCLUSIONS AND FUTURE DIRECTIONS

This paper studied adaptive control of discrete-time nonlinear systems with high-dimensional, matrix-valued parametric uncertainty. We showed that standard Euclidean adaptive schemes, such as recursive least squares and gradient-descent-based updates, can exhibit transient performance degradation as the ambient dimension increases, even when the true parameter has low intrinsic complexity. To address this limitation, we developed mirror-descent-based adaptive laws that exploit the geometry of structured parameter classes, including sparse, low-rank, simplex-constrained, and spectraplex-constrained uncertainty sets. The proposed framework guarantees boundedness of all closed-loop signals, asymptotic tracking, and regret bounds with at most logarithmic dependence on the ambient dimension.

Several directions remain open for future research. An important next step is to extend the regret bound in (26) to stochastic systems subject to process noise. Another direction is to sharpen the dependence of the regret bound on the horizon T under stronger excitation conditions, such as persistence of excitation [32]. Finally, incorporating prediction-based online learning methods, including gradient prediction and optimistic mirror descent [33, 34], may further improve transient performance in predictable or slowly varying environments.

APPENDIX A. PROOF OF AUXILIARY LEMMAS

A. Proof of Lemma 2

Throughout the proof, c, c_1 , and c_2 denote positive constants whose values may change from line to line. Since all norms are equivalent in finite-dimensional spaces, there exists a constant $c > 0$ such that

$$\|\nabla J_t(\hat{\Theta}_t)\|_* \leq c \|\nabla J_t(\hat{\Theta}_t)\|_F.$$

Using the definition of $\nabla J_t(\hat{\Theta}_t)$ and the submultiplicativity of the Frobenius norm, we find that

$$\begin{aligned} \|\nabla J_t(\hat{\Theta}_t)\|_* &\leq c \|\nabla J_t(\hat{\Theta}_t)\|_F = c \|\mathcal{B}^\top(\tilde{X}_{t+1})\Psi(X_t, t)^\top\|_F \leq c \|\mathcal{B}^\top(\tilde{X}_{t+1})\|_F \|\Psi(X_t, t)\|_F \\ &\leq c \|\mathcal{B}^\top\|_{\text{op}} \|\tilde{X}_{t+1}\|_F \|\Psi(X_t, t)\|_F. \end{aligned} \quad (52)$$

where $\|\mathcal{B}^\top\|_{\text{op}}$ denotes the operator norm of $\mathcal{B}^\top$. Next, consider the recursion

$$X_{t+1} = (\mathcal{A} - \mathcal{B}\mathcal{K})(X_t) + \tilde{X}_{t+1} + \mathcal{B}(U_t^d).$$

Since the closed-loop operator $\mathcal{A} - \mathcal{BK}$ is Schur stable, there exist constants $c > 0$ and $\gamma \in (0, 1)$ such that $\|(\mathcal{A} - \mathcal{BK})^t\|_{\text{op}} \leq c\gamma^t$ for all $t \geq 0$. Iterating the recursion yields

$$\begin{aligned} \|X_t\|_{\text{F}} &\leq c\gamma^t \|X_0\|_{\text{F}} + c \sum_{i=0}^{t-1} \gamma^{t-1-i} (\|\tilde{X}_{i+1}\|_{\text{F}} + \|\mathcal{B}(U_i^{\text{d}})\|_{\text{F}}) \\ &\leq c\gamma^t \|X_0\|_{\text{F}} + \frac{c}{1-\gamma} \max_{0 \leq i \leq t-1} (\|\tilde{X}_{i+1}\|_{\text{F}} + \|\mathcal{B}(U_i^{\text{d}})\|_{\text{F}}) \\ &\leq c_1 + c_2 \max_{0 \leq i \leq t-1} \|\tilde{X}_{i+1}\|_{\text{F}}, \end{aligned}$$

where in the second inequality we used the fact that the desired input U_t^{d} is bounded. Furthermore, due to the linear growth condition on $\Psi(X_t, t)$, we have

$$\|\Psi(X_t, t)\|_{\text{F}} \leq c_1 + c_2 \|X_t\|_{\text{F}} \leq c_1 + c_2 \max_{0 \leq i \leq t-1} \|\tilde{X}_{i+1}\|_{\text{F}}.$$

As a result, we find that

$$\|\tilde{X}_{t+1}\|_{\text{F}} \|\Psi(X_t, t)\|_{\text{F}} \leq c_1 + c_2 \max_{0 \leq i \leq t} \|\tilde{X}_{i+1}\|_{\text{F}}^2.$$

Finally, combining this bound with (52), we obtain

$$\|\nabla J_t(\hat{\Theta}_t)\|_*^2 \leq c \|\mathcal{B}^\top\|_{\text{op}}^2 \|\tilde{X}_{t+1}\|_{\text{F}}^2 \|\Psi(X_t, t)\|_{\text{F}}^2 \leq c_1 + c_2 \max_{0 \leq i \leq t} \|\tilde{X}_{i+1}\|_{\text{F}}^4,$$

as claimed.

REFERENCES

- [1] M. Krstic, I. Kanellakopoulos, and P. Kokotovic, *Nonlinear and Adaptive Control Design*. John Wiley, 1995.
- [2] K. J. Åström, “History of adaptive control,” in *Encyclopedia of Systems and Control*, J. Baillieul and T. Samad, Eds. Cham: Springer International Publishing, 2021, pp. 902–909.
- [3] A. M. Annaswamy and A. L. Fradkov, “A historical perspective of adaptive control and learning,” *Annu. Rev. Control*, vol. 52, pp. 18–41, 2021.
- [4] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović, “Adaptive nonlinear control without overparametrization,” *Systems & Control Letters*, vol. 19, no. 3, pp. 177–185, 1992.
- [5] A. S. Nemirovski and D. B. Yudin, *Problem Complexity and Method Efficiency in Optimization* (Wiley-Interscience Series in Discrete Mathematics). Chichester, UK: John Wiley & Sons, 1983, Translated from Russian by E. R. Dawson.
- [6] A. Juditsky, J. Kwon, M. Molinaro, and C. Tauvel, “Sparse recovery by reduced variance stochastic approximation,” *Information and Inference: A Journal of the IMA*, vol. 12, no. 2, pp. 851–896, 2023.
- [7] Y. Nesterov and A. Nemirovski, “On first-order algorithms for ℓ_1 /nuclear norm minimization,” *Acta Numerica*, vol. 22, pp. 509–575, 2013.
- [8] S. Shahrampour and A. Jadbabaie, “Distributed online optimization in dynamic environments using mirror descent,” *IEEE Transactions on Automatic Control*, vol. 63, no. 3, pp. 714–725, 2018.
- [9] D. Yuan, Y. Hong, D. W. C. Ho, and S. Xu, “Distributed mirror descent for online composite optimization,” *IEEE Transactions on Automatic Control*, vol. 66, no. 2, pp. 714–729, 2021.
- [10] A. Nemirovski, “Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems,” *SIAM Journal on Optimization*, vol. 15, no. 1, pp. 229–251, 2004.
- [11] A. Nemirovski, A. Juditsky, G. Lan, et al., “Robust stochastic approximation approach to stochastic programming,” *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1574–1609, 2009.
- [12] E. Hazan, “Introduction to online convex optimization,” *Foundations and Trends in Optimization*, vol. 2, no. 3–4, pp. 157–325, 2016.
- [13] T. Lattimore and C. Szepesvári, *Bandit Algorithms*. Cambridge University Press, 2020.

- [14] S. Bubeck and N. Cesa-Bianchi, “Regret analysis of stochastic and nonstochastic multi-armed bandit problems,” *Foundations and Trends in Machine Learning*, vol. 5, no. 1, pp. 1–122, 2012.
- [15] E. Hazan and K. Singh, *Introduction to Online Control*. Cambridge University Press, 2026.
- [16] Y. Abbasi-Yadkori and C. Szepesvári, “Regret bounds for the adaptive control of linear quadratic systems,” in *Proceedings of the 24th Annual Conference on Learning Theory*, ser. Proceedings of Machine Learning Research, vol. 19, 2011, pp. 1–26.
- [17] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, “Regret bounds for robust adaptive control of the linear quadratic regulator,” in *Advances in Neural Information Processing Systems*, vol. 31, 2018, pp. 4192–4201.
- [18] M. K. S. Faradonbeh, A. Tewari, and G. Michailidis, “Optimism-based adaptive regulation of linear-quadratic systems,” *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1802–1808, 2021.
- [19] F. Zhao, F. Dörfler, A. Chiuso, and K. You, “Data-enabled policy optimization for direct adaptive learning of the LQR,” *IEEE Transactions on Automatic Control*, vol. 70, no. 11, pp. 7217–7232, 2025.
- [20] H. Mania, S. Tu, and B. Recht, “Certainty equivalence is efficient for linear quadratic control,” in *Advances in Neural Information Processing Systems*, vol. 32, 2019.
- [21] N. M. Boffi, S. Tu, and J.-J. E. Slotine, “Regret bounds for adaptive nonlinear control,” in *Proceedings of the 3rd Conference on Learning for Dynamics and Control*, ser. Proceedings of Machine Learning Research, vol. 144, PMLR, Jul. 2021, pp. 471–483.
- [22] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*. Mineola, NY: Dover Publications, 2009.
- [23] K. J. Åström and B. Wittenmark, *Adaptive control*, 2nd ed. Mineola, NY: Dover Publications, 2008.
- [24] I. D. Landau, R. Lozano, M. M’Saad, and A. Karimi, *Adaptive Control: Algorithms, Analysis and Applications*, 2nd ed. London: Springer, 2011.
- [25] R. T. Rockafellar, *Convex Analysis*. Princeton, NJ: Princeton University Press, 1970.
- [26] J. M. Borwein and A. S. Lewis, *Convex Analysis and Nonlinear Optimization: Theory and Examples*, 2nd ed. New York: Springer, 2006.
- [27] D. P. Bertsekas, A. Nedić, and A. E. Ozdaglar, *Convex Analysis and Optimization*. Belmont, MA: Athena Scientific, 2003.
- [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2004.
- [29] J. P. Hespanha, *Linear Systems Theory*, 2nd ed. Princeton, NJ: Princeton University Press, 2018.
- [30] B. T. Polyak, *Introduction to Optimization*. New York: Optimization Software, Inc., 1987.
- [31] E. Hazan and S. Kakade, “Revisiting the polyak step size,” *arXiv preprint arXiv:1905.00313*, 2019.
- [32] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. De Moor, “A note on persistency of excitation,” *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, 2005.
- [33] A. Rakhlin and K. Sridharan, “Online learning with predictable sequences,” in *Annual Conference on Learning Theory*, S. Shalev-Shwartz and I. Steinwart, Eds., vol. 30, 2013, pp. 993–1019.
- [34] P. Z. Scroccaro, A. S. Kolarijani, and P. Mohajerin Esfahani, “Adaptive composite online optimization: Predictions in static and dynamic environments,” *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 2906–2921, 2023.