Online Control with Multiple Sensors

Matthew Faw[*] **Siva Theja Maguluri**[†]

Abstract

We consider a variant of the partially observed online control problem where there are *multiple sensors* which provide noisy linear measurements of the system. In our setting, the system state evolves according to a discrete-time linear dynamical system. The objective is to design an algorithm which minimizes regret relative to the best *single-sensor* policy (within a restricted class) in hindsight. When the system and process noise is stochastic and the costs are quadratic in the control and *unobserved state*, we design an algorithm that achieves $\mathcal{O}(\text{poly}\log(S)\sqrt{n})$ regret over time horizon n relative to our benchmark policy class if there are S sensors. Notably, this regret guarantee is with respect to the *state costs*, which are never fully observed. When the system and process noise is chosen by an oblivious adversary, and the cost functions are convex-Lipschitz (or subquadratic) functions of the control and *observations*, then a similar $\mathcal{O}(\text{poly}\log(S)\sqrt{n})$ regret is achievable. While our work is the first, to our knowledge, to study this generalization of the online control problem, we remark that a naïve application of known results in the partially observed online control literature have regret scaling $\mathcal{O}(\text{poly}(S)\sqrt{n})$, exponentially worse than our regret scaling in terms of dependence on the number of sensors. Our algorithm is based on the Follow-The-Regularized-Leader framework, and our analysis carefully exploits the ℓ_1 geometry of the policy class.

1 Introduction

Perhaps one of the most well-studied models of dynamical systems is the discrete-time linear dynamical system (LDS), where a state $\mathbf{x}_t \in \mathbb{R}^{d_{\mathbf{x}}}$ evolves linearly as a function of noise sequence $\mathbf{w}_t \in \mathbb{R}^{d_{\mathbf{x}}}$ and user-selected control inputs $\mathbf{u}_t \in \mathbb{R}^{d_{\mathbf{u}}}$ according to the (time-homogeneous) equation:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \mathbf{w}_t \quad \forall t \geq 0.$$

The control problem associated with an LDS is to select a control sequence $(\mathbf{u}_t)_{t\geq 0}$ which minimizes the total cumulative state-control costs $c_t(\mathbf{x}_t, \mathbf{u}_t)$. When the cost functions are quadratic and the noise sequence $\mathbf{w}_t = \mathbf{0}$, this problem is known as the linear quadratic regulator (LQR) problem, and the optimal control is a (time-varying) linear function $\mathbf{u}_t = K_t\mathbf{x}_t$, where K_t depends on the current a future costs. Moreover, when the state is not fully observed, but instead we are given noisy linear measurements $\mathbf{y}_{t+1} = C\mathbf{x}_t + \mathbf{v}_t$, where the initial state $\mathbf{x}_0$ is Gaussian, $\mathbf{w}_t, \mathbf{v}_t$ are zero-mean independent Gaussian random variables, the problem is known as the linear quadratic Gaussian (LQG) problem. In this case, the optimal control is $\mathbf{u}_t = K_t\widehat{\mathbf{x}}_{t|t}$, where $\widehat{\mathbf{x}}_{t|t}$ is the Kalman filter [Kal60], which maintains a linear estimate of the unknown state $\mathbf{x}_t$ as a function of the observed $\mathbf{y}_t$'s. Moreover, if $\mathbf{x}_0, (\mathbf{w}_t, \mathbf{v}_t)$ are not Gaussian but have known bounded covariance, then $\mathbf{u}_t = K_t\widehat{\mathbf{x}}_{t|t}$ is the best *linear* controller.

Since these early foundational results, significant research efforts have sought less stringent conditions under which the optimal control can be expressed in a simple (computationally feasible) manner. The

[*]Google Research. This work was completed while the author was an ARC PostDoc at Georgia Intitute of Technology. Email: mfaw3@gatech.edu

[†]H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology. Email: siva.theja@gatech.edu

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

so-called H_∞ control problem is one notable attempt, where the goal is to design a minimax-optimal controller for the worst-case (bounded) noise sequence. It turns out that, under quadratic costs, the optimal control is a (different) time-varying linear controller [BB08]. Relaxing the assumption of quadratic costs even in the stochastic noise setting to more general convex cost functions (e.g., quadratic costs with linear constraints) can lead to optimal controls which are computationally infeasible to compute [Roc87].

Over the past decade, an exciting line of works has leveraged the tools of online learning and online convex optimization to obtain strong regret guarantees for controlling linear dynamical systems when the noise sequences $\mathbf{w}_t, \mathbf{v}_t$ are adversarial, and the cost functions are arbitrary convex-Lipschitz (or sub-quadratic) functions of the observations and control. Because finding the optimal controller is computationally infeasible in general, we can instead aim to minimize regret relative to the best policy in hindsight, within some (computationally feasible) policy class Π:

$$\text{Regret}_\Pi = \sum_{t=0}^{n} c_t(\mathbf{x}_t, \mathbf{u}_t) - \min_{\pi \in \Pi} \sum_{t=0}^{n} c_t(\mathbf{x}_t^\pi, \mathbf{u}_t^\pi),$$

where $\mathbf{u}_t^\pi$ is the control executed by policy π at time t, and $\mathbf{x}_t^\pi$ is the resulting state. In applications such as controlling networked systems like robot swarms, it is common to have sensor redundancies to ensure availability. Moreover, in many IoT settings, the devices often have resource constraints (e.g., network bandwidth, power, or computation) which limit the number of sensors that the devices should use [TJP16; GCHM06; IJ18]. That is, at each time step, there are multiple available sensors providing different noisy linear measurements of the current state:

$$\mathbf{y}_t^{(s)} = C^{(s)}\mathbf{x}_t + \mathbf{v}_t^{(s)} \quad \forall s \in \mathcal{S}.$$

Given the setting described above, a natural goal is to find a controller which uses only a sparse subset of the sensors $\mathcal{S}$. Perhaps the simplest objective which captures the essence of this problem is the regret relative to the best *single-sensor policy in hindsight*. That is, we want to find some "sparse" controller which performs nearly as well as the best policy in hindsight using only one (or constant number) of the $|\mathcal{S}|$ sensors. This problem of online control with multiple sensors a natural generalization of the online control problem. While this problem, to the best of our knowledge, is new, existing standard techniques from online control imply sublinear regret bounds for this problem (see Remark H.11), in the case of cost functions depending on the *observations* $\mathbf{y}_t$. However, the standard techniques give a (suboptimal) regret scaling which depends polynomially on $|\mathcal{S}|$.

1.1 Our Results

In this work, we design an algorithm for controlling stable multi-sensor linear dynamical systems. In particular, we design an algorithm which has low regret with respect to the class Π_{sLDC} of single-sensor linear dynamical controllers (see Definition 2.3). This class of controllers is quite rich, as it captures, e.g., the optimal single-sensor controllers for both the LQG and the H_∞ problems. Moreover, it is the standard comparator class of controllers in the online control literature adapted to our multi-sensor setting (see e.g., [HS22; SSH20]). We obtain the following two results:

Sub-Gaussian noise and quadratic costs. Our first result is in the stochastic setting, where the process and measurement noise, as well as the initial state, are independent sub-Gaussian vectors with fixed, known, time-invariant covariances, and the cost functions are quadratic separately in the control and *unobserved* state (see Assumption 2 and 3).

Theorem 1.1 (Informal statement of Theorem 4.1)**.** *For the problem of controlling a stable multi-sensor LDS over a time-horizon of n in the sub-Gaussian noise and quadratic costs setting, there is an algorithm which produces a control sequence with expected regret relative to Π_{sLDC} at most $\mathcal{O}(\text{poly}\log(|\mathcal{S}|+1)\cdot\text{poly}(\dim)\cdot\sqrt{n})$, where $\dim = d_{\mathbf{x}} + d_{\mathbf{u}} + \max_{s\in\mathcal{S}} d_{y^s}$ is the system dimensions.*

We emphasize that the costs in this setting depend on the *unobserved* states $\mathbf{x}_t$. This stands in contrast to existing result for online control (see, e.g., [SSH20]), where the costs depend instead on the *observations*. We show that this stronger notion of regret can be controlled in the stochastic setting by running an online control algorithm which estimates the state costs using the state estimator given by the Kalman filter.

Adversarial noise and convex, sub-quadratic costs. Our second result is in the adversarial setting, where the process and measurement noise, as well as the initial state, are bounded vectors chosen by an

oblivious adversary. In this setting, we consider convex, Lipschitz/sub-quadratic (see Assumption 3') cost functions $c_t(\widehat{\mathbf{x}}_t, \mathbf{u}_t)$ chosen by an oblivious adversary depending on the current control $\mathbf{u}_t$ and a fixed, stable *state estimator* $\widehat{\mathbf{x}}_t$ which can depend on $\widehat{\mathbf{x}}_{t-1}$ as well as the current measurements $(\mathbf{y}_t^{(s)})_{s\in\mathcal{S}}$ (see Assumption 4'). We note that this cost function model is a generalization of the observation costs (i.e., costs depending on the current control and observations) typical of the online control literature (see, e.g., [SSH20; HS22]). Note that we consider costs depending on a fixed state estimator instead of the true state (as we did in the stochastic setting) since the noise is adversarial, and thus the true state need not be identifiable from the observations.

Theorem 1.2 (Informal statement of Theorem 4.1')**.** *For the problem of controlling a stable multi-sensor LDS over time-horizon n in the adversarial noise and convex Lipschitz/sub-quadratic costs setting, there is an algorithm which produces a control sequence with regret relative to Π_{sLDC} at most $\mathcal{O}(\operatorname{poly}\log(|\mathcal{S}|+1)\cdot\operatorname{poly}(\dim)\cdot\sqrt{n})$, where $\dim = d_{\mathbf{x}} + d_{\mathbf{u}} + \max_{s\in\mathcal{S}} d_{y^s}$ is the system dimensions.*

Comparing this result to Theorem 1.1 in the stochastic setting with quadratic costs, we observe that we obtain nearly identical regret scaling in the presence of adversarial noise and general convex Lipschitz/sub-quadratic costs. The main distinction between these two results is the dependence of the costs on the unobserved state in the stochastic setting vs the observed state estimator in the adversarial setting.

Algorithm. Our algorithm (Algorithm 1) for obtaining the guarantees from Theorems 1.1 and 1.2 builds on ideas from the online learning and online control literature, and is based on the follow-the-regularized-leader (FTRL) algorithm framework [AHR09; HK10]. As in these works, our algorithm uses ideas from convex optimization to search over a convex set of control policies. Most existing algorithms for online control make use of Online Gradient Descent (see [HS22] for a textbook overview in the online control context). However, in our setting, the dimension of the gradients scales linearly in the number of sensors, so using Online Gradient Descent would result in regret scaling polynomially, instead of logarithmically, in the number of sensors $|\mathcal{S}|$. Instead, exploiting the sparse structure of the policy class Π_{sLDC}, we use FTRL with a group-norm regularizer (which behaves similarly to a negentropy regularizer), which is key to obtaining our improved scaling in the number of sensors.

1.2 Related Works

Optimal control of linear dynamical systems is a mature field with numerous textbooks devoted to the material, e.g., [AM05; AM07; BB08; Ber05; KV15]. Much of this literature is devoted to deriving closed-form optimal controls in the case of stochastic noise sequences (often Gaussian) and quadratic cost functions. For cost functions other than quadratics, the optimal control often does not take a simple closed form [Roc87]. Much is known about relaxations of the stochastic noise assumption, however. H_∞ control aims to find the best policy for the worst-case bounded noise sequence, and many of the nice properties of optimal control under independent noise carry over to this setting [BB08].

A drawback of the H_∞ control is that its guarantees are worst-case, and the optimal control for the worst-case noise may be very different than a typical noise sequence. This shortcoming motivated the study of regret guarantees for control, where the objective is to minimize regret relative to the best policy *in hindsight*. The early works for this problem focused on the simplest setting of IID noise, quadratic costs, and full state observations (i.e., $\mathbf{y}_t = \mathbf{x}_t$) [AS11; DMMRT18; CKM19; SF20], where the optimal regret achievable over time horizon n is $\sqrt{n}$. However, a growing literature of works have extended these results to partially observed systems or more general convex cost functions [ABHKS19; HKS20; SSH20; LAHA20; CCK22; BBDH25b; BBDH25a]. These works build on the tools of online learning and online convex optimization (see, e.g., textbooks of [CL06; Haz+16; Ora19]), and many of these guarantees hold even for adversarially chosen costs and noise sequences. See [HS22] for a textbook treatment of the material.

While our setting of online multi-sensor control appears new, there is a rich literature devoted to related problems in the controls community. For instance, a large number of works have studied the design of control algorithms under various communication and quantization constraints [NFZE07; BM97; BA07]. There have also been a number of works focusing on sensor selection for state estimation (ignoring the control objective) [GCHM06; JB08; CPR17; TJP16]. Finally, a number of

works have also studied the so-called co-design problem where the goal is to minimize a (steady-state, or finite-horizon) control objective subject to a constraint on the number of sensors [IJ18; MPB15; Lim92; TCPJ20]. While their objective is morally similar to ours, their settings are solving an offline stochastic problem, while our objective is to minimize regret in the online setting in the presence of possibly adversarial noise.

2 Problem Setting

System model. We consider the following discrete-time linear dynamical system (LDS) evolving over n rounds with a set of sensors $\mathcal{S}$, each providing a noisy linear measurement of the current system state:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \mathbf{w}_t, \quad \text{and} \quad \mathbf{y}_t^{(s)} = C^{(s)}\mathbf{x}_t + \mathbf{v}_t^{(s)} \ \forall s \in \mathcal{S} \tag{LDS}$$

Here $\mathbf{x}_t \in \mathbb{R}^{d_\mathbf{x}}$ is the (unobserved) system state at time t, $\mathbf{u}_t \in \mathbb{R}^{d_\mathbf{u}}$ is the control input, and $\mathbf{y}_t^{(s)} \in \mathbb{R}^{d_{y^s}}$ is the observation of state $\mathbf{x}_t$ from a sensor s in some finite set of sensors $\mathcal{S}$. $\mathbf{w}_t \in \mathbb{R}^{d_\mathbf{x}}$ is the process noise, and $\mathbf{v}_t^{(s)} \in \mathbb{R}^{d_{y^s}}$ is the observation noise corresponding to sensor s. It will sometimes be convenient to refer to the concatenation of all observations from the sensors:

Definition 2.1 (Stacked system)**.** Given a set of sensors $\mathcal{S}$ with measurements $\mathbf{y}_t^{(s)} = C^{(s)}\mathbf{x}_t + \mathbf{v}_t^{(s)}$, the *stacked system* is: $\mathbf{y}_t = [(\mathbf{y}_t^{(1)})^\top \dots (\mathbf{y}_t^{(|\mathcal{S}|)})^\top]^\top = C\mathbf{x}_t + \mathbf{v}_t$, where C (resp., $\mathbf{v}_t$) is the concatenation of the sensor matrices $C^{(s)}$ (resp., observation noises $\mathbf{v}_t^{(s)}$). In particular, the stacked system corresponds to the concatenation of all sensor measurements.

In all of the settings we consider in this work, we will assume the following bounds on the system parameters:

Assumption 1 (System scaling)**.** The control and observation matrices are bounded in operator norm, i.e., $\|B\|_{\mathsf{op}} \leq R_B$ and $\|C^{(s)}\|_{\mathsf{op}} \leq R_{C^{(s)}}$. Moreover, the system is strongly stable, i.e., for any integer $i \geq 0$ $\|A^i\|_{\mathsf{op}} \leq R_A \rho_A^i$.

We note that the above assumption requires that the system is "internally stable," i.e., the system remains bounded when $\mathbf{u}_t = \mathbf{0}$ for all t. This assumption is made primarily to simplify the exposition of our results. There are a number of standard tricks that can apply when A is not stable (e.g., assuming access to a nominal stabilizing controller). For a more detailed discussion, see [SSH20, Appendix C].

Control objective. At each time-step t, there is an associated (convex) cost function c_t, and our goal is to design a control policy to minimize *regret* relative to a benchmark policy class Π. In the stochastic setting, we consider cost functions depending on the current control $\mathbf{u}_t$ and the *unobserved* state $\mathbf{x}_t$, and take the regret corresponding to a sequence of controls $\mathbf{u}_t$ and associated states $\mathbf{x}_t$ to be:

$$\mathrm{Regret}_\Pi = \textstyle\sum_{t=0}^n c_t(\mathbf{x}_t, \mathbf{u}_t) - \min_{\pi \in \Pi} \sum_{t=0}^n c_t(\mathbf{x}_t^\pi, \mathbf{u}_t^\pi), \tag{1}$$

where $\mathbf{u}_t^\pi$ and $\mathbf{x}_t^\pi$ are the counterfactual controls and states, had the policy $\pi \in \Pi$ been executed for all t. In the presence of adversarial noise, we consider a weaker notion of regret under *observation costs*, where costs are instead measured in terms of some fixed and known state estimator $\widehat{\mathbf{x}}_t$:

$$\widehat{\mathrm{Regret}}_\Pi = \textstyle\sum_{t=0}^n c_t(\widehat{\mathbf{x}}_t, \mathbf{u}_t) - \min_{\pi \in \Pi} \sum_{t=0}^n c_t(\widehat{\mathbf{x}}_t^\pi, \mathbf{u}_t^\pi). \tag{2}$$

In this adversarial setting, we consider linear state estimators of the form $\widehat{\mathbf{x}}_t = X\widehat{\mathbf{x}}_{t-1} + Y\mathbf{y}_t$. Intuitively, the state estimator $\widehat{\mathbf{x}}_t$ captures any stable linear combination of the (stacked) observations $\mathbf{y}_0, \dots, \mathbf{y}_t$. Taking $X = \mathbf{0}$ corresponds to costs depending only on the current observation $\mathbf{y}_t$. The norm constraint on $Y\mathbf{y}_t$ ensures that the norm of $\widehat{\mathbf{x}}_t$ does not scale polynomially in the number of sensors, and is satisfied, for example, when $Y\mathbf{y}_t = \sum_{s \in \mathcal{S}} \alpha_s \mathbf{y}_t^{(s)}$ for any mixture coefficients $(\alpha_s)_{s \in \mathcal{S}}$. One can view $\widehat{\mathbf{x}}_t$ as playing the same role as the Kalman filter plays in the stochastic noise setting.

Comparator policy class. The dynamics of the (LDS) depend on the executed control policy through the control vector $\mathbf{u}_t$. Perhaps the most well-studied class of policies are *linear dynamical controllers* (LDC). These are controllers which maintain their own internal LDS $\mathbf{h}_t^{\mathsf{ldc}}$, and execute controls based

on this hidden state $\mathbf{h}_t^{\mathsf{ldc}}$ and observations from the system being controlled. We note that the class of LDCs is rich enough to capture, e.g., the optimal LQG controller, the H_∞ optimal controller, and the controller with optimal competitive ratio [BB08; GASH23].[3]

Definition 2.2 (Linear Dynamical Controllers on a Single Sensor)**.** A *linear dynamical controller* (LDC) is parameterized by matrices $\pi_{\mathsf{ldc}}^{(s)} = (A_{\mathsf{ldc}}^{(s)}, B_{\mathsf{ldc}}^{(s)}, C_{\mathsf{ldc}}^{(s)}, D_{\mathsf{ldc}}^{(s)})$, and a sensor $s \in \mathcal{S}$. The control output $\mathbf{u}_t^{\mathsf{ldc}}$ is given by: $\mathbf{h}_{t+1}^{\mathsf{ldc}} = A_{\mathsf{ldc}}^{(s)}\mathbf{h}_t^{\mathsf{ldc}} + B_{\mathsf{ldc}}^{(s)}\mathbf{y}_t^{(s)}$ and $\mathbf{u}_t^{\mathsf{ldc}} = C_{\mathsf{ldc}}^{(s)}\mathbf{h}_t^{\mathsf{ldc}} + D_{\mathsf{ldc}}^{(s)}\mathbf{y}_t^{(s)}$, where $\mathbf{h}_t^{\mathsf{ldc}}$ is the internal state of the controller, and $\mathbf{y}_t^{(s)}$ is the sensor output of the system (LDS) observed at time t from sensor s.

Our goal is to bound the regret relative to single-sensor LDCs. Thus, it will be useful to parametrize this class of controllers. We use the following parametrization:

Definition 2.3 (Constrained single-sensor LDC class)**.** For a linear system (LDS) and parameters $R > 0$, $\rho \in (0, 1)$, the class of (R, ρ)-constrainted single-sensor LDCs is

$$\Pi_{\mathsf{sLDC}}(R, \rho) := \left\{\pi_{\mathsf{ldc}}^{(s)} = (A_{\mathsf{ldc}}^{(s)}, B_{\mathsf{ldc}}^{(s)}, C_{\mathsf{ldc}}^{(s)}, D_{\mathsf{ldc}}^{(s)}) : s \in \mathcal{S}, \|G_{\mathsf{ldc,cl}}^{(s),[i]}\|_{\mathsf{op}} \leq R\rho^i \; \forall i \geq 0\right\},$$

where $\pi_{\mathsf{ldc}}^{(s)}$ are the parameters of the single-sensor LDC from Definition 2.2. $G_{\mathsf{ldc,cl}}^{(s)}$ is the closed-loop Markov operator of $\pi_{\mathsf{ldc}}^{(s)}$ interacting in (LDS), defined by the matrices $G_{\mathsf{ldc,cl}}^{(s),[0]} = D_{\mathsf{ldc}}^{(s)}$ and $G_{\mathsf{ldc,cl}}^{(s),[i]} = C_{\mathsf{ldc,cl}}^{(s)}(A_{\mathsf{ldc,cl}}^{(s)})^{i-1}B_{\mathsf{ldc,cl}}^{(s)}$, where $(A_{\mathsf{ldc,cl}}^{(s)}, B_{\mathsf{ldc,cl}}^{(s)}, C_{\mathsf{ldc,cl}}^{(s)}, D_{\mathsf{ldc}}^{(s)})$ are the closed-loop system matrices associated with internal state $(\mathbf{x}_t, \mathbf{h}_t^{\mathsf{ldc}})$ and observation $(\mathbf{y}_t, \mathbf{u}_t)$ evolving according to (LDS) when running $\pi_{\mathsf{ldc}}^{(s)}$ (given explicitly in Definition B.14).

The parameters R and ρ provide a measure of the stability of the closed-loop dynamics of the LDC in (LDS), and our regret bounds will scale polynomially in R and $1/(1-\rho)$. Requiring that $\rho < 1$ ensures that past controls do not have significant long-term impacts on the system dynamics.

3 Online Multi-Sensor Control Algorithm

In this section, we describe Algorithm 1, our algorithm which obtains strong regret guarantees relative to $\Pi_{\mathsf{sLDC}}(R, \rho)$, the class of constrained single-sensor LDCs from Definition 2.3. As is standard in online control, our goal is to use ideas from online convex optimization to obtain such a regret bound.

At a high level, each iteration Algorithm 1 consists of three main phases, each of which we describe in detail below: (i) Executing control $\mathbf{u}_t$ depending on the current and past observations, and a current control policy π_t, (ii) Observing the cost function c_t and estimating the cost of the current (unknown) state $\mathbf{x}_t$ and control $\mathbf{u}_t$, and (iii) Selecting a new control policy π_{t+1} based on the past and current observations and costs.

Phase 1: Control execution. To specify the control execution phase, the first step is deciding what control policy class Π to use and optimize over. Since our goal is to obtain regret bounds relative to the class $\Pi_{\mathsf{sLDC}}(R, \rho)$ of LDCs, a naïve choice is $\Pi = \Pi_{\mathsf{sLDC}}(R, \rho)$. However, due to the recursive nature of LDCs in Definition 2.2, the controls $\mathbf{u}_t^{\mathsf{ldc}}$ are high-degree polynomial functions of the LDC parameters $\pi_{\mathsf{ldc}}^{(s)} = (A_{\mathsf{ldc}}^{(s)}, B_{\mathsf{ldc}}^{(s)}, C_{\mathsf{ldc}}^{(s)}, D_{\mathsf{ldc}}^{(s)})$. As a result, directly optimizing the costs c_t over $\pi_{\mathsf{ldc}}^{(s)} \in \Pi_{\mathsf{sLDC}}(R, \rho)$ is a nonconvex optimization problem. Thus, it will be convenient to search instead over a *convex relaxation* of this policy class. The relaxation we use in this work is the class of *disturbance response controllers* from [SSH20, Definition 3.2].

Nature's observations: To introduce the disturbance response controller, we must first introduce the notion of *Nature's observations*, the counterfactual observations that would have been observed, had no inputs been played (i.e., $\mathbf{u}_t = \mathbf{0}$) for all timesteps t.

Definition 3.1 (Nature's x's and y's [SSH20])**.** We let $\mathbf{x}_t^{\mathsf{nat}}$ (resp., $\mathbf{y}_t^{\mathsf{nat},(s)}$) be the (counterfactual) state (resp., observation from sensor s) of the system at time t, had we chosen $\mathbf{u}_s = 0$ for all $s \leq t$. That is, $\mathbf{x}_0^{\mathsf{nat}} = \mathbf{x}_0$, $\mathbf{x}_{t+1}^{\mathsf{nat}} = A\mathbf{x}_t^{\mathsf{nat}} + \mathbf{w}_t$, and $\mathbf{y}_t^{\mathsf{nat},(s)} = C^{(s)}\mathbf{x}_t^{\mathsf{nat}} + \mathbf{v}_t^{(s)}$.

[3] See Appendix B.5 for details on expressing linear controllers in terms of LDCs.

Algorithm 1 Online control with group-norm FTRL

Input: Time horizon $n \geq 1$, DRC parameters $R_M > 0, \rho_M \in (0,1), H \geq 1$, step-size $\eta > 0$.
Initialize $\mathbf{M}(0) \in \mathcal{M}(R_M, \rho_M, H)$ arbitrarily
for $t = 1, \ldots, n$ **do**
  Observe all $\mathbf{y}_t^{(s)}$ and compute $\mathbf{y}_t^{\mathsf{nat},(s)} = \mathbf{y}_t^{(s)} - \sum_{i=1}^{t} C^{(s)} A^{i-1} B \mathbf{u}_{t-i}$ (as in Lemma 3.2).
  Execute control $\mathbf{u}_t(\mathbf{M}(t)) = \sum_{s \in \mathcal{S}} \sum_{j=0}^{H} M_j^{(s)}(t) \mathbf{y}^{\mathsf{nat},(s)}$
  if *Stochastic noise setting* **then**
    Compute the Kalman filter estimate $\widehat{\mathbf{x}}_t := \widehat{\mathbf{x}}_{t|t}$ of $\mathbf{x}_t$,

$$\widehat{\mathbf{x}}_{0|-1} = \mu_0, \quad \widehat{\mathbf{x}}_{t|t} = (I - \tilde{L}_t C)\widehat{\mathbf{x}}_{t|t-1} + \tilde{L}_t C \quad \text{and} \quad \widehat{\mathbf{x}}_{t+1|t} = A\widehat{\mathbf{x}}_{t|t} + B\mathbf{u}_t,$$

    where $\tilde{L}_t$ is the Kalman gain matrix which depends on the system (A, C), the noise covariance W, V, and initial state covariance X_0 (defined explicitly in Definition B.5). Let $\widehat{\mathbf{x}}_t(\mathbf{M}) := \widehat{\mathbf{x}}_{t|t}(\mathbf{M})$ be the counterfactual Kalman filter as defined in Definition 3.5.
  else if *Adversarial noise setting* **then**
    Let the (fixed and given) state estimator be:

$$\widehat{\mathbf{x}}_{-1}(\mathbf{M}) = \mathbf{0} \quad \text{and} \quad \widehat{\mathbf{x}}_t(\mathbf{M}) = X\widehat{\mathbf{x}}_{t-1}(\mathbf{M}) + Y\mathbf{y}_t(\mathbf{M}),$$

    where X and Y are given estimation matrices, and $\mathbf{y}_t(\mathbf{M})$ is the counterfactual state observation from Definition 3.5.
  Observe c_t and construct counterfactual loss function $\widehat{\ell}_t(\mathbf{M}) := c_t(\widehat{\mathbf{x}}_t(\mathbf{M}), \mathbf{u}_t(\mathbf{M}))$
  Update $\mathbf{M}(t+1) = \arg\min_{\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)} \Psi(\mathbf{M}) + \sum_{i=0}^{t} \left\langle \nabla \widehat{\ell}_i(\mathbf{M}(i)), \mathbf{M} \right\rangle$, where $\Psi(\mathbf{M}) = \frac{ep}{2\eta} \|\mathbf{M}\|^2_{\rho_M^{-1/2}, q}$ is the group-norm regularizer, where $p = 2\max\{1, \ln(|\mathcal{S}|)\}$, $\frac{1}{q} = 1 - \frac{1}{p}$ (defined in Lemma 3.6).
end

Due to the linear structure of the system dynamics (LDS), the counterfactual observation $\mathbf{y}_t^{\mathsf{nat}}$ can be computed, given the observation $\mathbf{y}_t$ using the following result of [SSH20]. We give a proof for completeness in Appendix C.

Lemma 3.2 (Lemma 3.1 from [SSH20])**.** *Consider any system evolving according to* (LDS)*, with any (possibly adaptively selected) control inputs* $\mathbf{u}_t$*. Then,* $\mathbf{x}_t^{\mathsf{nat}} = \mathbf{x}_t - \sum_{i=1}^{t} A^{i-1} B \mathbf{u}_{t-i}$ *and* $\mathbf{y}_t^{\mathsf{nat},(s)} = \mathbf{y}_t^{(s)} - \sum_{i=1}^{t} C^{(s)} A^{i-1} B \mathbf{u}_{t-i}$.

Disturbance Response Control: We are now ready to define disturbance response controllers, which are a class of controllers which are linear in Nature's y's.

Definition 3.3 (Disturbance Response Controller)**.** A *disturbance response controller* (DRC) with history length $H \geq 1$ is parametrized by $|\mathcal{S}|(H+1)$ matrices $\mathbf{M} = (M_j^{(s)})_{s \in \mathcal{S}, j \in [0,H]}$. At each time t, the control is given by: $\mathbf{u}_t(\mathbf{M}) = \sum_{s \in \mathcal{S}} \sum_{i=0}^{H} M_i^{(s)} \mathbf{y}_{t-i}^{\mathsf{nat},(s)}$, where we take $\mathbf{y}_{-i}^{\mathsf{nat}} = \mathbf{0}$ for all $i \in [H]$. We take $\mathcal{M}(R_M, \rho_M, H)$ as the set of DRCs with controlled decay:

$$\mathcal{M}(R_M, \rho_M, H) := \left\{ \mathbf{M} = (M_j^{(s)})_{s \in \mathcal{S}, j \in [0,H]} : \|M_j\|_{\ell_1, \mathsf{op}} := \sum_{s \in \mathcal{S}} \|M_j^{(s)}\|_{\mathsf{op}} \leq R_M \rho_M^j \ \forall j \in [0, H] \right\}$$

Notice that the DRC control $\mathbf{u}_t(\mathbf{M})$ is *linear* in Nature's y's, and, moreover, Nature's y's do not depend on any controls (since, by definition, $\mathbf{y}^{\mathsf{nat},(s)}$ are the observations when $\mathbf{u}_{t'} = \mathbf{0}$ for all t'). This linear structure (in contrast to the highly nonlinear structure of the LDCs), together with the basic observation that $\mathcal{M}(R_M, \rho_M, H)$ is convex, allows us to use convex optimization techniques to optimize the control cost. Moreover, we note that this class of DRCs is expressive enough to capture LDCs, and thus obtaining regret guarantees against DRCs implies regret bounds against LDCs.

Claim 3.4 (Expressivity of policy class, simplified version of Claim D.1)**.** For any single-sensor LDC policy $\pi_{\mathsf{ldc}}^{(s)} \in \Pi_{\mathsf{sLDC}}(R, \rho)$ (where $R > 0$, $\rho \in (0,1)$), there is a controller $\mathbf{M} \in \mathcal{M}(R, \rho, n)$ which produces the same control sequence for all $t \in [n]$.

By [SSH20, Theorem 1], taking $H = \Omega(\log(n))$ suffices to approximate the performance of a LDC up to an additive error (in terms of regret). Because of the structure of the regularizer of our algorithm, we can set H arbitrarily large (e.g., $H = n$) without affecting our regret guarantees (just like in [KDK23]). Intuitively, this holds because of the geometric decay requirement on the control matrices $M_j^{(s)}$. We also note that, unlike the class of controllers $\Pi_{\mathsf{sLDC}}(R, \rho)$ which use *exactly* one sensor, the DRCs in $\mathcal{M}(R_M, \rho_M, H)$ can potentially use more than one measurement. However, the ℓ_1 constraint serves as a convex relaxation of the cardinality constraint, essentially allowing for using mixture distributions over single sensors.

Phase 2: Cost estimation. After executing a control, the algorithm suffers cost $c_t(\mathbf{x}_t, \mathbf{u}_t)$ in the stochastic setting, or $c_t(\widehat{\mathbf{x}}_t, \mathbf{u}_t)$ in the adversarial setting, and the cost function c_t is revealed. In order to choose a better controller, the algorithm must evaluate how well the current control $\mathbf{M}(t)$ is performing. There are two main subtleties that our algorithm must address when performing this evaluation.

The first issue, specifically in the stochastic setting (where $\mathbf{w}_t, \mathbf{v}_t$ are mean-zero), the instantaneous cost $c_t(\mathbf{x}_t, \mathbf{u}_t)$ cannot be computed exactly since $\mathbf{x}_t$ is unknown. However, we can still estimate $\mathbf{x}_t$ using the Kalman filter $\widehat{\mathbf{x}}_t$ (as specified in Algorithm 1). Our analysis establishes that using this cost estimate is a sufficient proxy for our algorithm to minimize state cost regret. For more background on the Kalman filter, refer to Appendix B.3.

The second issue, common to both stochastic and adversarial settings, is the fact that the cost functions depend on the current (estimated) state. Thus, the cost measured by the algorithm for playing a particular control policy $\mathbf{u}_t := \mathbf{u}_t(\mathbf{M}(t))$ depends not only on $\mathbf{M}(t)$, but also *all previous controls*. By contrast, our goal is to minimize regret relative to a *fixed* DRC $\mathbf{M}^*$ which plays $\mathbf{u}_t(\mathbf{M}^*)$ for all t. As a consequence, it is more natural to compute the cost a *counterfactual* states that the system would have arrived in by playing the current policy at all previous rounds:

Definition 3.5 (Counterfactuals)**.** Given any $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$, denote $\mathbf{x}_t(\mathbf{M})$ (resp., $\mathbf{y}_t^{(s)}(\mathbf{M})$) as the resulting state (resp., observation) from executing controls $\mathbf{u}_j(\mathbf{M})$ for all times $j \le t$. Similarly, the counterfactual Kalman filter $\widehat{\mathbf{x}}_{t|t}(\mathbf{M})$ is the Kalman filter on the observations $\mathbf{y}_t^{(s)}(\mathbf{M})$.

Given the cost function c_t and linear state estimator $\widehat{\mathbf{x}}_t$, our algorithm then constructs the counterfactual loss $\widehat{\ell}_t(\mathbf{M}) = c_t(\widehat{\mathbf{x}}_t(\mathbf{M}), \mathbf{u}_t(\mathbf{M}))$ associated with a DRC controller $\mathbf{M}$. Note that, since the state estimator (both in the stochastic and adversarial case) is linear in the observations and controls, $\widehat{\mathbf{x}}_t(\mathbf{M})$, and hence also $\widehat{\ell}_t(\mathbf{M})$, can be computed (see Definition C.2 for details).

Phase 3: Control policy update. Finally, we use ideas from online convex optimization to select the next controller $\mathbf{M}(t+1)$. One common approach in the online control literature is to perform this update using Online Projected Gradient Descent on the counterfactual losses $\widehat{\ell}_t(\mathbf{M})$. However, since the gradient of this loss function has dimensions linear in the number of sensors $|\mathcal{S}|$, the regret of this algorithm will scale polynomially, instead of logarithmically, in $|\mathcal{S}|$. Thus, instead, our algorithm utilizes the follow-the-regularized-leader (FTRL) framework with a group-norm regularizer. For more discussion on these regularizers, refer to Appendix E.1, or see, e.g., [KST+09] and [Ora19, Section 7.7]. For our purposes, we will only need the following:

Lemma 3.6. *For $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$, $\rho > 0$ and $q \in (1, 2]$, the (ρ, q)-group norm of $\mathbf{M}$ is $\|\mathbf{M}\|_{\rho,q} = \|(\|M^{(1)}\|_\rho, \ldots, \|M^{(|\mathcal{S}|)}\|_\rho)\|_q$ where $\|M^{(s)}\|_\rho^2 = \sum_{j=0}^{H} \rho^{2j} \|M_j^{(s)}\|_F^2$. Moreover, for $p = 2\max\{1, \ln(|\mathcal{S}|)\}$ and $\frac{1}{q} = 1 - \frac{1}{p}$, then $G_\rho(\mathbf{M}) = \frac{ep}{2}\|\mathbf{M}\|_{\rho,q}^2$ is 1-strongly convex w.r.t. $\|\cdot\|_{\rho,1}$, and the dual norm of $\|\cdot\|_{\rho,1}$ is $\|\cdot\|_{\rho^{-1},\infty}$.*

The above establishes that the regularizer $\Psi(\mathbf{M}) = {}^1\!/\!_\eta G_{1/\sqrt{\rho_M}}(\mathbf{M})$ resembles a negentropy regularizer. The inner norm $\|M^{(s)}\|_\rho$ serves the same purpose as in [KDK23], eliminating the dependence on H in the regret by exploiting the geometric decay of $M_j^{(s)}$ in j, by definition of $\mathcal{M}(R_M, \rho_M, H)$.

4 Main Results and Proof Sketch

We now describe our main result, our regret guarantee for Algorithm 1. As our results are of a slightly different form in the stochastic and adversarial settings, we detail the settings and our results in each of these settings below. Then, we provide a proof sketch of our main ideas for the stochastic setting.

Regime 1: IID noise and quadratic costs. To derive our strongest results on regret in terms of state costs $c_t(\mathbf{x}_t, \mathbf{u}_t)$, we will require the following assumptions on our system. First, we need that the noise sequence satisfies the following stochastic assumption:

Assumption 2 (Stochastic noise bounds)**.** The initial state $\mathbf{x}_0$ has (finite) mean $\mu_0 = \mathbb{E}[\mathbf{x}_0]$, variance $X_0 \succeq \mathbf{0}$, and is sub-Gaussian.[4] Additionally, the process noise $\mathbf{w}_t$ and observation noise $\mathbf{v}_t$ are mean-zero, IID sub-Gaussian, and have finite covariance $W \succeq \mathbf{0}$ and $V = \mathrm{diag}(V_1, \ldots, V_{|\mathcal{S}|}) \succ \mathbf{0}$, respectively.[5] Moreover, $(\mathbf{x}_0, (\mathbf{w}_t, \mathbf{v}_t)_{t \geq 0})$ are jointly independent.

Our results additionally rely on the following special structure of the cost functions:

Assumption 3 (Quadratic costs)**.** For all time-steps t, the cost functions $c_t(\mathbf{x}, \mathbf{u}) := \mathbf{x}^\top S_t \mathbf{x} + \mathbf{u}^\top R_t \mathbf{u}$, where $S_t \succeq \mathbf{0}$ and $R_t \succ \mathbf{0}$ are symmetric (without loss of generality) and $\|S_t\|_{\mathsf{op}} \leq R_{x,\mathsf{cost}}$, $\|R_t\|_{\mathsf{op}} \leq R_{u,\mathsf{cost}}$ for all $t \geq 0$.

Finally, since the costs depend on the unobserved state $\mathbf{x}_t$, we will require the following "detectability" condition which ensures that the state $\mathbf{x}_t$ can be estimated (in particular, by the Kalman filter, see Algorithm 1). See Appendix B.2 for more background on this condition.

Assumption 4 (Detectability of Kalman Filter)**.** The Kalman filter (see Algorithm 1) is dynamically detectable, i.e., for any time steps $t' \geq t$, $\|\prod_{i=1}^{t'-t}(A - L_{t'-i}C)\|_{\mathsf{op}} \leq R_{A-LC}\rho_{A-LC}^{t'-t}$, where L_t is the Kalman gain matrix at time t.

We note that the above assumptions are standard conditions, which are sometimes referred to as LQG (linear quadratic Gaussian) setting. Given this setup, we are now ready to state our main result in this setting, a bound on the expected regret (see (1)) relative to the class of single-sensor LDCs:

Theorem 4.1. *Suppose that Algorithm 1 is run for n time steps with step-size parameter $\eta = \Theta(1/\sqrt{n})$ and DRC parameters $R_M > 0, \rho_M \in (0, 1)$, and $H = \Omega(\log(n))$. Then, if the system (LDS) satisfy Assumption 1, the initial state $\mathbf{x}_0$ and the process and observation noises $\mathbf{w}_t, \mathbf{v}_t$ satisfy Assumption 2, the system is strongly detectable, i.e., Assumption 4, and cost functions satisfy Assumption 3. Then, the controls $\mathbf{u}_t$ selected by Algorithm 1 satisfy:* $\mathbb{E}\left[\mathrm{Regret}_{\mathcal{M}(R_M, \rho_M, H)}\right] = \mathcal{O}(\mathrm{poly}\log(|S|+1) \cdot \mathrm{poly}(\dim) \cdot \sqrt{n})$, *where* $\dim = d_{\mathbf{x}} + d_{y^s,\max} + d_{\mathbf{u}}$ *is the system dimensions.*

Regime 2: Adversarial noise and convex costs. We now relax our assumptions on the noise sequence and cost functions. In this regime, we assume the following mild regularity conditions on the noise sequences.

Assumption 2' (Adversarial noise bounds)**.** The initial state $\mathbf{x}_0$ is finite $\|\mathbf{x}_0\|_2 \leq R_{x_0}$, and for all $t \geq 0$, the process and observation noise $\|\mathbf{w}_t\| \leq R_w$ and $\max_{s \in \mathcal{S}} \|\mathbf{v}_t^{(s)}\|_2 \leq R_{v,\sup}$. These quantities can be chosen by an oblivious adversary.

When the noise sequences are IID sub-Gaussian, then the above bounds R_{x_0} and R_w will generally be logarithmic in the time horizon, and $R_{v,\sup}$ will generally be logarithmic in the time horizon and number of sensors with high probability.

We now introduce our assumptions on the cost functions. As discussed earlier, in the adversarial setting, state costs are no longer a natural quantity, since the adversarial noise makes the system state nonidentifiable. As such, we consider observation costs $c_t(\widehat{\mathbf{x}}_t, \mathbf{u}_t)$, where $\widehat{\mathbf{x}}_t$ is a linear estimator satisfying the following boundedness assumption:

Assumption 4' (Bounded State Estimator)**.** The state estimator $\widehat{\mathbf{x}}_t$ defined as $\widehat{\mathbf{x}}_{-1} = \mathbf{0}$ and $\widehat{\mathbf{x}}_t = X\widehat{\mathbf{x}}_{t-1} + Y\mathbf{y}_t$ satisfies $\|X^i\|_{\mathsf{op}} \leq R_X \rho_X^i$ for all $i \geq 0$, and $\|Y\mathbf{y}_t\| \leq R_Y \max_{s \in \mathcal{S}} \|\mathbf{y}_t^{(s)}\|_2$.

Since we consider observation costs, we can now study a much broader class of cost functions than in the stochastic setting. In particular, we require the cost functions satisfy the following:

Assumption 3' (Convex, sub-quadratic costs)**.** For all t, $c_t(\widehat{\mathbf{x}}, \mathbf{u})$ is convex and satisfies, for some $L, R > 0$: $|c_t(\widehat{\mathbf{x}}, \mathbf{u}) - c_t(\widehat{\mathbf{x}}', \mathbf{u}')| \leq L \max\{R\|(\widehat{\mathbf{x}} - \widehat{\mathbf{x}}', \mathbf{u} - \mathbf{u}')\|_2, \|(\widehat{\mathbf{x}} - \widehat{\mathbf{x}}', \mathbf{u} - \mathbf{u}')\|_2^2\}$.

These cost functions capture both convex and Lipschitz functions, as well as quadratic cost functions. We are now ready to state our main result in the adversarial setting.

[4] Refer to Definition E.6 in Appendix E.2 for a formal definition

[5] Here, $\mathrm{diag}(A_1, \ldots, A_k)$ denotes the block diagonal matrix with diagonal blocks A_i, $i \in [k]$

Theorem 4.1'. Suppose that Algorithm 1 is run for n time steps with step-size parameter $\eta = \Theta(1/\sqrt{n})$ and DRC parameters $R_M > 0, \rho_M \in (0,1)$, and $H = \Omega(\log(n))$. Suppose the system (LDS) satisfies Assumption 1, the initial state $\mathbf{x}_0$ and process and observation noise $\mathbf{w}_t, \mathbf{v}_t$ satisfies Assumption 2', and the costs functions satisfy Assumption 3' and depend on a fixed and known state estimator $\widehat{\mathbf{x}}_t$ satisfying Assumption 4'. Then, the regret of this algorithm satisfies: $\widehat{\mathrm{Regret}}_{\mathcal{M}(R_M,\rho_M,H)} = \mathcal{O}(\mathrm{poly}\log(|S|+1)\cdot\mathrm{poly}(\dim)\cdot\sqrt{n})$.

4.1 Proof sketch for the stochastic setting

Here, we give an overview of the main components for proving Theorem 4.1, our regret bound in the stochastic, quadratic cost setting. Refer to Appendix H for the complete proof. We focus on bounding regret with respect to the class of DRC controllers $\mathcal{M}(R_M, \rho_M, H)$. Recall that, by Claim 3.4, bounds on $\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)}$ relative to DRCs transfer directly to bounds on $\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)}$ relative to LDCs.[6] Then, we can decompose the regret of Algorithm 1 as:

$$\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)} = \sum_{t=0}^{n} c_t(\mathbf{x}_t,\mathbf{u}_t) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=0}^{n} c_t(\mathbf{x}_t(\mathbf{M}),\mathbf{u}_t(\mathbf{M}))$$

Our goal is to make use of regret bounds for FTRL. However, there are two issues in using these bounds directly: (i) The cost $c_t(\mathbf{x}_t,\mathbf{u}_t)$ incurred by the algorithm cannot be evaluated exactly since we never observe $\mathbf{x}_t$, and (ii) The cost is *history-dependent* since $\mathbf{x}_t$ depends on the entire trajectory of prior controls. To handle these issues, we use the Kalman filter estimate $\widehat{\mathbf{x}}_{t|t}$ of the state $\mathbf{x}_t$ to decompose the cost of the algorithm in the following three terms:

$$\begin{aligned} c_t(\mathbf{x}_t,\mathbf{u}_t) &= c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)),\mathbf{u}_t(\mathbf{M}(t))) \\ &\quad + c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t(\mathbf{M}(t))) - c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)),\mathbf{u}_t(\mathbf{M}(t))) \\ &\quad + c_t(\mathbf{x}_t,\mathbf{u}_t(\mathbf{M}(t))) - c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t(\mathbf{M}(t))). \end{aligned}$$

Note that the first term above, $\widehat{\ell}_t(\mathbf{M}(t)) = c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)),\mathbf{u}_t(\mathbf{M}(t)))$, is a proxy cost term that the algorithm can evaluate. It represents the estimated counterfactual cost of playing DRC controller $\mathbf{M}(t)$ if it had been played not only at time t, but also at all previous times. The second term represents a memory error term, measuring the cost differential at the estimated true state vs the counterfactual state. The final term represents the error due to estimating the cost at $\mathbf{x}_t$ with the Kalman filter estimate. Similarly, for the benchmark term inside the min, we decompose:

$$c_t(\mathbf{x}_t(\mathbf{M}),\mathbf{u}_t(\mathbf{M})) = \widehat{\ell}_t(\mathbf{M}) + c_t(\mathbf{x}_t(\mathbf{M}),\mathbf{u}_t(\mathbf{M})) - c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}),\mathbf{u}_t(\mathbf{M}))$$

Now, the first key observation is that, under the quadratic loss assumption (Assumption 3), the cost estimation error term decomposes into the following convenient form:

$$c_t(\mathbf{x}_t(\mathbf{M}),\mathbf{u}_t(\mathbf{M})) - c_t(\widehat{\mathbf{x}}_t(\mathbf{M}),\mathbf{u}_t(\mathbf{M})) = \mathbf{e}_t(\mathbf{M})^\top S_t \mathbf{e}_t(\mathbf{M}) + 2\widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t \mathbf{e}_t(\mathbf{M}),$$

where $\mathbf{e}_t(\mathbf{M}) = \mathbf{x}_t(\mathbf{M}) - \widehat{\mathbf{x}}_{t|t}(\mathbf{M})$ is the Kalman estimation error. To reason about this cost differential, we use the well-known property of the Kalman filter (see Lemma C.1) that the error does not depend on the controls, i.e., $\mathbf{e}_t(\mathbf{M}) = \mathbf{e}_t$. As a result, only the second (mean-zero, by the Orthogonality principle Lemma B.7) term depends on the DRC $\mathbf{M}$. Thus, collecting these observations, we can decompose the regret in three parts:

$$\begin{aligned} \mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)} &\le \textstyle\sum_{t=0}^{n} \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=0}^{n} \widehat{\ell}_t(\mathbf{M}) \\ &\quad + \textstyle\sum_{t=0}^{n} c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) - c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)),\mathbf{u}_t) \\ &\quad + 2\textstyle\sum_{t=0}^{n} \widehat{\mathbf{x}}_{t|t}^\top S_t \mathbf{e}_t + 2\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)} - \textstyle\sum_{t=0}^{n} \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t \mathbf{e}_t(\mathbf{M}) \end{aligned}$$

Bounding the online learning error. The first term above, which we denote $\mathrm{R_{ol}}$, is the regret controlled by the online learning algorithm. By appealing to standard FTRL regret bounds tailored to our chosen regularizer (see Lemma H.6 for details)[7], the first regret term is bounded as:

$$\mathrm{R_{ol}} \le \Psi(\mathbf{M}^*) + \frac{\eta}{2}\sum_{t=0}^{n} \|\nabla\widehat{\ell}_t(\mathbf{M}(t))\|^2_{\sqrt{\rho_M},\infty} \le \frac{e\ln(e|\mathcal{S}|)}{\eta}\|\mathbf{M}^*\|^2_{\frac{1}{\sqrt{\rho_M}},1} + \frac{\eta}{2}\sum_{t=0}^{n} \|\nabla\widehat{\ell}_t(\mathbf{M}(t))\|^2_{\sqrt{\rho_M},\infty},$$

[6] While this Lemma is stated for $H = n$, an approximate version of the statement which introduces negligible error in the regret bound holds as long as $H = \Omega(\log(n))$. See [SSH20, Theoerm 1] for details.

[7] Another potential way to exploit the ℓ_1 geometry of our problem would be to use Hedge as a meta-algorithm, as was done in a different setting in [ZYWZ23]. While we expect this approach would also work, we chose the group-norm approach since required analyzing only a single optimization procedure.

where $\Psi(\mathbf{M}^*)$ is the regularizer from Algorithm 1 evaluated at the optimal in hindsight DRC. The first term is bounded by $\mathcal{O}(\log(|\mathcal{S}|)/\eta)$ by construction of $\mathcal{M}(R_M, \rho_M, H)$. Thus, as long as the gradients of $\widehat{\ell}_t(\mathbf{M}(t))$ are $\mathcal{O}(\text{poly}\log(|\mathcal{S}|))$-Lipschitz with respect to the $(\sqrt{\rho_M}, \infty)$-group norm, then since the step-size $\eta = \Theta(1/\sqrt{n})$, $\mathbb{E}[\mathrm{R}_{\mathrm{ol}}] = \mathcal{O}(\text{poly}\log(|\mathcal{S}|) \cdot \sqrt{n})$.

We establish the claimed Lipschitz bound (in expectation) on the gradients of $\widehat{\ell}_t(\mathbf{M}(t))$ in Lemma G.4. The fact that the regret bound depends on the ℓ_∞ (as opposed to the ℓ_2) norm is key to obtaining logarithmic (as opposed to polynomial) scaling in $|\mathcal{S}|$, and is made possible by our choice of regularizer. The main challenge in establishing this Lipschitz bound comes the dependence of $\widehat{\ell}_t(\cdot)$ on $\widehat{\mathbf{x}}_{t|t}$, which combines measurements $\mathbf{y}^{\mathsf{nat},(s)}$ from all sensors according to the update $\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} = (I - \tilde{L}_t C)\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t-1} + \tilde{L}_t \mathbf{y}^{\mathsf{nat}}_t$ (see Definition B.5). Thankfully, we can show that the Kalman filter incorporates these measurements in a way such that the norm scales (in expectation) as only $\text{poly}\log(|\mathcal{S}|)$, assuming the noise is sub-Gaussian (see Proposition B.9 and Lemma G.3 for details).

Bounding the memory error. Bounding the second, memory error term in the regret bound, denoted $\mathrm{R}_{\mathrm{mem}}$, reduces to showing that adjacent DRC control vectors are "stable": essentially, that the change $\|\mathbf{u}_t(\mathbf{M}(t)) - \mathbf{u}_t(\mathbf{M}(t-1))\|_2 = \mathcal{O}(\eta \cdot \text{poly}\log(|\mathcal{S}|))$ scales with the step-size η.[8] Obtaining a bound of this form without a $\text{poly}(|\mathcal{S}|)$ factor requires care. To see why, we can apply Hölder's inequality and the definition of $\mathbf{u}_t(\mathbf{M})$ from Definition 3.3 to obtain the following bound:

$$\|\mathbf{u}_t(\mathbf{M}(t)) - \mathbf{u}_t(\mathbf{M}(t-1))\|_2 \leq \|\mathbf{M}(t) - \mathbf{M}(t-1)\|_{\frac{1}{\sqrt{\rho_M}},1} \left(\sum_{j=1}^{H} \rho_M^j \max_{s\in\mathcal{S}} \|\mathbf{y}^{\mathsf{nat},(s)}_{t-1-j}\|_2^2\right)^{1/2}$$

Assumption 1 guarantees that norm of $\mathbf{y}^{\mathsf{nat},(s)}_t$ remains bounded, and using properties of FTRL, we know that $\|\mathbf{M}(t) - \mathbf{M}(t-1)\|_{\frac{1}{\sqrt{\rho_M}},1} \leq \eta \|\nabla \widehat{\ell}_t(\mathbf{M}(t-1))\|_{\sqrt{\rho_M},\infty}$ This is the second crucial step where we exploit the geometry of the problem: if instead we had used Online Gradient Descent, then we would not have the above contraction in the ℓ_1 norm, but in ℓ_2, which would result in a bound scaling in the ℓ_2 norm of the gradient, which would result in polynomial scaling in $|\mathcal{S}|$. Again using the Lipschitz bound on the gradients discussed above, we can conclude that $\mathbb{E}[\mathrm{R}_{\mathrm{mem}}] = \mathcal{O}(\text{poly}\log(|\mathcal{S}|) \cdot \sqrt{n})$.

Bounding the residual error. Finally, we focus on the third term, denoted $\mathrm{R}_{\mathrm{res}}$, the residual error due to estimating the cost at the unobserved state by the cost at the Kalman filter estimate of the state. While the first residual error is zero in expectation (by the Orthogonality principle, Lemma B.7), the second is the *supremum* of mean-zero quantities over all possible DRCs in $\mathcal{M}(R_M, \rho_M, H)$. To bound the supremum, we first observe that (Lemma H.7), up to dimension-dependent constants, we can bound this term as the maximum over only $|\mathcal{S}|$ terms:

$$\mathbb{E}[\mathrm{R}_{\mathrm{res}}] \lesssim \mathbb{E}\left[\max_{s\in\mathcal{S}} \|\sum_{t\in[n]} S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_t)^\top\|_F\right].$$

Notice that the sum inside the norm above is also mean-zero, by the Orthogonality principle. Thus, if we knew that $\|\sum_{t\in[n]} S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_t)^\top\|_F$ were $\mathcal{O}(\sqrt{n})$-sub-Exponential, then we could use the fact that the expected max of $|\mathcal{S}|$ such sub-Exponential random variables is $\mathcal{O}(\log(|\mathcal{S}|)\sqrt{n})$ (see Lemma E.12 for details). Now, assuming the noise $\mathbf{w}_t, \mathbf{v}_t$ are sub-Gaussian, so are $\mathbf{e}_t$ and $\mathbf{y}^{\mathsf{nat},(s)}_t$. Thus, if these terms were independent, then this goal would be simple to establish. However, these quantities do (crucially) depend on each other, more care is required. Thankfully, the dependencies between these two quantities decays geometrically due to the stability properties of the system. For this reason, we can still establish that $\mathbb{E}[\mathrm{R}_{\mathrm{res}}] = \mathcal{O}(\sqrt{n}\text{poly}\log(|\mathcal{S}|))$. Combining these three bounds yields the claimed regret bound.

Conclusion and Future Work

In this work, we study the online control problem in the generalized setting of multiple sensors. We show that, when competing against the benchmark of best single-sensor linear dynamical controller, one can obtain regret which scales poly-logarithmically with the number of sensors, improving exponentially on the best bounds obtainable from prior work. There are a number of interesting directions for future research, including characterizing the optimal dependence on the number of sensors, and studying extensions to the setting of unknown system matrices. In forthcoming work, we extend our results to a limited feedback setting where not all sensors are observed each round.

[8] See Lemmas H.3, H.4 and H.5 for precise details.

Acknowledgments and Disclosure of Funding

This work was partially supported by NSF grants EPCN-2144316 and CPS-2240982.

References

[AS11] Yasin Abbasi-Yadkori and Csaba Szepesvári. "Regret bounds for the adaptive control of linear quadratic systems". In: *Proceedings of the 24th Annual Conference on Learning Theory*. JMLR Workshop and Conference Proceedings. 2011, pp. 1–26 (page 3).

[AHR09] Jacob D Abernethy, Elad Hazan, and Alexander Rakhlin. "Competing in the dark: An efficient algorithm for bandit linear optimization". In: *Conference on Learning Theory*. 110. 2009 (page 3).

[ABHKS19] Naman Agarwal, Brian Bullins, Elad Hazan, Sham Kakade, and Karan Singh. "Online control with adversarial disturbances". In: *International Conference on Machine Learning*. PMLR. 2019, pp. 111–119 (page 3).

[AM07] Brian DO Anderson and John B Moore. *Optimal control: linear quadratic methods*. Courier Corporation, 2007 (pages 3, 22).

[AM05] Brian DO Anderson and John B Moore. *Optimal filtering*. Courier Corporation, 2005 (pages 3, 19, 20, 23).

[BA07] John Baillieul and Panos J Antsaklis. "Control and communication challenges in networked real-time systems". In: *Proceedings of the IEEE* 95.1 (2007), pp. 9–28 (page 3).

[BB08] Tamer Başar and Pierre Bernhard. *H-infinity optimal control and related minimax design problems: a dynamic game approach*. Springer Science & Business Media, 2008 (pages 2, 3, 5).

[Ber05] Dimitri Bertsekas. *Dynamic programming and optimal control*. 3rd ed. Vol. I. Athena scientific, 2005 (pages 3, 19).

[BM97] Vivek S Borkar and Sanjoy K Mitter. "LQG control with communication constraints". In: *Communications, computation, control, and signal processing: a tribute to Thomas Kailath*. Springer, 1997, pp. 365–373 (page 3).

[BBDH25a] Anand Brahmbhatt, Gon Buzaglo, Sofiia Druchyna, and Elad Hazan. "A New Approach to Controlling Linear Dynamical Systems". In: *arXiv preprint arXiv:2504.03952* (2025) (page 3).

[BBDH25b] Anand Brahmbhatt, Gon Buzaglo, Sofiia Druchyna, and Elad Hazan. "Efficient Spectral Control of Partially Observed Linear Dynamical Systems". In: *arXiv preprint arXiv:2505.20943* (2025) (page 3).

[CCK22] Asaf B Cassel, Alon Cohen, and Tomer Koren. "Efficient online linear control with stochastic convex costs and unknown dynamics". In: *Conference on Learning Theory*. PMLR. 2022, pp. 3589–3604 (page 3).

[CL06] Nicolo Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006 (page 3).

[CPR17] Luiz FO Chamon, George J Pappas, and Alejandro Ribeiro. "The mean square error in Kalman filtering sensor selection is approximately supermodular". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE. 2017, pp. 343–350 (page 3).

[CKM19] Alon Cohen, Tomer Koren, and Yishay Mansour. "Learning Linear-Quadratic Regulators Efficiently with only sqrtT Regret". In: *International Conference on Machine Learning*. PMLR. 2019, pp. 1300–1309 (page 3).

[DMMRT18] Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. "Regret bounds for robust adaptive control of the linear quadratic regulator". In: *Advances in Neural Information Processing Systems* 31 (2018) (page 3).

[GASH23] Gautam Goel, Naman Agarwal, Karan Singh, and Elad Hazan. "Best of both worlds in online control: Competitive ratio and policy regret". In: *Learning for dynamics and control conference*. PMLR. 2023, pp. 1345–1356 (page 5).

[GCHM06] Vijay Gupta, Timothy H Chung, Babak Hassibi, and Richard M Murray. "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage". In: *automatica* 42.2 (2006), pp. 251–260 (pages 2, 3).

[Haz+16] Elad Hazan et al. "Introduction to online convex optimization". In: *Foundations and Trends® in Optimization* 2.3-4 (2016), pp. 157–325 (page 3).

[HKS20] Elad Hazan, Sham Kakade, and Karan Singh. "The nonstochastic control problem". In: *Algorithmic Learning Theory*. PMLR. 2020, pp. 408–421 (page 3).

[HK10] Elad Hazan and Satyen Kale. "Extracting certainty from uncertainty: Regret bounded by variation in costs". In: *Machine learning* 80.2 (2010), pp. 165–188 (page 3).

[HS22] Elad Hazan and Karan Singh. "Introduction to online nonstochastic control". In: *arXiv preprint arXiv:2211.09619* (2022) (pages 2, 3).

[IJ18] Takuya Iwaki and Karl Henrik Johansson. "LQG control and scheduling co-design for wireless sensor and actuator networks". In: *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE. 2018, pp. 1–5 (pages 2, 4).

[JB08] Siddharth Joshi and Stephen Boyd. "Sensor selection via convex optimization". In: *IEEE Transactions on Signal Processing* 57.2 (2008), pp. 451–462 (page 3).

[KST+09] Sham Kakade, Shai Shalev-Shwartz, Ambuj Tewari, et al. "On the duality of strong convexity and strong smoothness: Learning applications and matrix regularization". In: *Unpublished Manuscript, http://ttic. uchicago. edu/shai/papers/KakadeShalevTewari09. pdf* 2.1 (2009), p. 35 (pages 7, 31–33).

[Kal60] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45 (pages 1, 19).

[KV15] Panqanamala Ramana Kumar and Pravin Varaiya. *Stochastic systems: Estimation, identification, and adaptive control*. SIAM, 2015 (pages 3, 19).

[KDK23] Raunak Kumar, Sarah Dean, and Robert Kleinberg. "Online convex optimization with unbounded memory". In: *Advances in Neural Information Processing Systems* 36 (2023), pp. 26229–26270 (page 7).

[LAHA20] Sahin Lale, Kamyar Azizzadenesheli, Babak Hassibi, and Anima Anandkumar. "Logarithmic regret bound in partially observable linear dynamical systems". In: *Advances in Neural Information Processing Systems* 33 (2020), pp. 20876–20888 (page 3).

[Lim92] KB Lim. "Method for optimal actuator and sensor placement for large flexible structures". In: *Journal of Guidance, Control, and Dynamics* 15.1 (1992), pp. 49–57 (page 4).

[MPB15] Claudia P Moreno, Harald Pfifer, and Gary J Balas. "Actuator and sensor selection for robust control of aeroservoelastic systems". In: *2015 American Control Conference (ACC)*. IEEE. 2015, pp. 1899–1904 (page 4).

[NFZE07] Girish N Nair, Fabio Fagnani, Sandro Zampieri, and Robin J Evans. "Feedback control under data rate constraints: An overview". In: *Proceedings of the IEEE* 95.1 (2007), pp. 108–137 (page 3).

[Ora19] Francesco Orabona. "A modern introduction to online learning". In: *arXiv preprint arXiv:1912.13213* (2019) (pages 3, 7, 35).

[Roc87] R Tyrell Rockafellar. "Linear-quadratic programming and optimal control". In: *SIAM Journal on Control and Optimization* 25.3 (1987), pp. 781–814 (pages 2, 3).

[Sch16] James R Schott. *Matrix analysis for statistics*. John Wiley & Sons, 2016 (page 16).

[SF20] Max Simchowitz and Dylan Foster. "Naive exploration is optimal for online lqr". In: *International Conference on Machine Learning*. PMLR. 2020, pp. 8937–8948 (page 3).

[SSH20] Max Simchowitz, Karan Singh, and Elad Hazan. "Improper learning for non-stochastic control". In: *Conference on Learning Theory*. PMLR. 2020, pp. 3320–3436 (pages 2–7, 9, 29–31, 61, 62).

[TCPJ20] Vasileios Tzoumas, Luca Carlone, George J Pappas, and Ali Jadbabaie. "LQG control and sensing co-design". In: *IEEE Transactions on Automatic Control* 66.4 (2020), pp. 1468–1483 (page 4).

[TJP16] Vasileios Tzoumas, Ali Jadbabaie, and George J Pappas. "Sensor placement for optimal Kalman filtering: Fundamental limits, submodularity, and algorithms". In: *2016 American control conference (ACC)*. IEEE. 2016, pp. 191–196 (pages 2, 3).

[vW96] Aad W van der Vaart and Jon A Wellner. "Weak convergence". In: *Weak convergence and empirical processes: with applications to statistics*. Springer, 1996, pp. 16–28 (page 34).

[Ver18] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*. Vol. 47. Cambridge university press, 2018 (pages 33–35).

[ZYWZ23] Peng Zhao, Yu-Hu Yan, Yu-Xiang Wang, and Zhi-Hua Zhou. "Non-stationary online learning with memory and non-stochastic control". In: *Journal of Machine Learning Research* 24.206 (2023), pp. 1–70 (page 9).

Overview of Appendix

A Auxiliary Results **16**

B Background in Linear Control **17**

B.1 Linear dynamical systems 17

B.2 Observability and Controllability 17

B.3 Kalman Filter 19

B.3.1 Setup and definitions 19

B.3.2 Important Properties of Kalman Filter 19

B.3.3 Useful Properties of Kalman Matrices and Error 20

B.4 Optimal LQ Control 22

B.5 Linear Dynamical Controllers 23

B.6 Deferred Proofs 23

B.6.1 Proof of Lemma B.6 23

B.6.2 Proof of Lemma B.7 24

B.6.3 Proof of Proposition B.9 25

B.6.4 Proof of Lemma B.10 26

B.6.5 Proof of Lemma B.11 27

B.6.6 Proof of Lemma B.12 28

B.6.7 Proof of Lemma B.13 29

C Nature's Observations: Deferred Proofs **29**

D Expressivity of Disturbance Response Controllers **31**

E Norm Facts **31**

E.1 Matrix Group Norms 31

E.2 Orlicz Norms, sub-Gaussian and sub-Exponential Random Variables 33

E.3 Norms on nets 35

F Convex Optimization Background **35**

G State Norm Bounds **36**

G.1 State and observation norm bounds 36

G.2 Kalman Noise Addition Bound 38

G.3 Kalman Estimate Norm Bounds 41

G.4 Estimated Loss Gradient Bounds 42

G.5 Control Matrix Norm Bounds 45

H Regret Bound: Deferred Proofs **46**

H.1 Cost Decomposition 46

H.2 Error of adjacent iterate depends on step-size 46
H.3 Error due to loss function memory 46
H.3.1 Error in state estimate function values 46
H.3.2 State and control error due to memory 48
H.4 Regret on estimated/hypothetical loss function 50
H.5 Error due to unobserved state 51
H.6 Regret decomposition 58
H.7 Proof of regret bound 59
H.8 A note on regret scaling polynomially in the number of sensors 61

I Simulations **61**

A Auxiliary Results

Lemma A.1 (Matrix Inversion Lemma, e.g., Theorem 1.9, [Sch16])**.** *Let $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times m}$, $U, V^\top \in \mathbb{R}^{n\times m}$. Suppose A and C are invertible. Then,*

$$(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$$

Lemma A.2 (Properties of Pseudoinverse, e.g., Chapter 5, [Sch16])**.** *Given a matrix $A \in \mathbb{R}^{m\times n}$, the* pseudoinverse *of A, denoted A^+, satisfies the following four conditions [Sch16, Definition 5.1, Theorem 5.1]:*

1. $AA^+A = A$
2. $A^+AA^+ = A^+$
3. $(A^+A)^\top = A^+A$
4. $(AA^+)^\top = AA^+$

If A is nonsingluar, then $A^+ = A^{-1}$. Moreover, $(\alpha A)^+ = \frac{1}{\alpha}A^+$ for any scalar $\alpha \neq 0$, $(A^+)^+ = A$, $(A^\top)^+ = (A^+)^\top$, $(A^+A)^+ = A^+A$, $(AA^+)^+ = AA^+$ [Sch16, Theorem 5.3]. If A is idempotent, then $A^+ = A$ [Sch16, Theorem 5.5]. If $A = U\Lambda V^\top$ for some U, V satisfying $U^\top U = I$ and $V^\top V = I$, then

$$(U\Lambda V^\top)^+ = (V^\top)^+\Lambda^+U^+ = V\Lambda^+U^\top.$$

[Sch16, Theorem 5.8] In particular, consider a symmetric matrix Σ with its associated projection matrix $P = \Sigma\Sigma^+ = \Sigma^+\Sigma$. Then, for any symmetric matrix B of compatible dimensions,

$$(\Sigma^+ + PBP)^+ = P(\Sigma^+ + PBP)^+ = (\Sigma^+ + PBP)^+P = P(\Sigma^+ + PBP)^+P.$$

Proof. Decompose Σ according to the compact spectral decomposition $\Sigma = U\Lambda U^T$, where $\Lambda \succ \mathbf{0}$ is the diagonal matrix of nonzero eigenvalues, and $U^\top U = I$. Then,

$$\Sigma^+ + PBP = U\Lambda^+U^\top + UU^\top BUU^\top = U(\Lambda^+ + U^\top BU)U^\top.$$

Therefore,

$$(\Sigma^+ + PBP)^+ = (U(\Lambda^+ + U^\top BU)U^\top)^+ = U(\Lambda^+ + U^\top BU)^+U^\top.$$

Since $PU = U$, the claim follows. □

Lemma A.3 (Generalized Matrix Inversion Lemma)**.** *Suppose $V \in \mathbb{R}^{n\times n}, \Sigma \in \mathbb{R}^{m\times m}$, and $C \in \mathbb{R}^{n\times m}$, where $V \succ \mathbf{0}$ and $\Sigma \succeq \mathbf{0}$. Then, we have the following equivalent forms:*

$$\begin{aligned}\Sigma C^\top(V + C\Sigma C^\top)^{-1} &= (\Sigma^+ + PC^\top V^{-1}CP)^+C^\top V^{-1}\\ &= P(\Sigma^+ + PC^\top V^{-1}CP)^+PC^\top V^{-1}\\ &= (\Sigma^+ + PC^\top V^{-1}CP)^+PC^\top V^{-1}\end{aligned}$$

where Σ^+ denotes the pseudoinverse of Σ, and $P = \Sigma\Sigma^+$ is the orthogonal projection onto colsp(Σ).

Proof. Since Σ is positive semidefinite, let us begin by decomposing $\Sigma = U\Lambda U^\top$ by its compact spectral decomposition, where $\Lambda \in \mathbb{R}^{r\times r}$ is the positive diagonal matrix $\Lambda \succ \mathbf{0}$ of $r \leq d_{\mathbf{x}}$ nonzero eigenvalues of Σ, and $U \in \mathbb{R}^{d_{\mathbf{x}}\times r}$ is the semi-orthogonal matrix $U^\top U = I$. Then, by Lemma A.1,

$$(V + C\Sigma C^\top)^{-1} = (V + CU\Lambda U^\top C^\top)^{-1} = V^{-1} - V^{-1}CU(\Lambda^{-1} + U^\top C^\top V^{-1}CU)^{-1}U^\top C^\top V^{-1}.$$

Therefore,

$$\begin{aligned}\Sigma C^\top(V + C\Sigma C^\top)^{-1} &= U\Lambda U^\top C^\top(V^{-1} - V^{-1}CU(\Lambda^{-1} + U^\top C^\top V^{-1}CU)^{-1}U^\top C^\top V^{-1})\\ &= U\Lambda(I - U^\top C^\top V^{-1}CU(\Lambda^{-1} + U^\top C^\top V^{-1}CU)^{-1})U^\top C^\top V^{-1}\\ &= U\Lambda\Lambda^{-1}(\Lambda^{-1} + U^\top C^\top V^{-1}CU)^{-1}U^\top C^\top V^{-1}\\ &= U(\Lambda^{-1} + U^\top C^\top V^{-1}CU)^{-1}U^\top C^\top V^{-1}.\end{aligned}$$

Now, since $U^\top U = I$, and using the fact that $U\Lambda^{-1}U^\top = (U\Lambda U^\top)^+ = \Sigma^+$,

$$\Sigma C^\top (V + C\Sigma C^\top)^{-1} = U(U^\top(\Sigma^+ + C^\top V^{-1}C)U)^{-1}U^\top C^\top V^{-1}.$$

Now, notice that since $U^\top U = I$, for any M such that $UMU^\top$ is invertible,

$$U(U^\top MU)^{-1}U^\top = (UU^\top MUU^\top)^+.$$

Applying this fact for $M = \Sigma^+ + C^\top V^{-1}C$, we obtain:

$$\Sigma C^\top (V + C\Sigma C^\top)^{-1} = (UU^\top(\Sigma^+ + C^\top V^{-1}C)UU^\top)^+ C^\top V^{-1}.$$

Finally, notice that, by definition of U, $UU^\top = U\Lambda U^\top U\Lambda^{-1}U^\top = \Sigma\Sigma^+ = \Sigma^+\Sigma$. Hence, denoting $P = \Sigma\Sigma^+ = \Sigma^+\Sigma$, we have that $P\Sigma^+P$. Therefore,

$$\Sigma C^\top (V + C\Sigma C^\top)^{-1} = (P(\Sigma^+ + C^\top V^{-1}C)P)^+ C^\top V^{-1} = (\Sigma^+ + PC^\top V^{-1}CP)^+ C^\top V^{-1}.$$

Finally, notice that, for any positive semidefinite A, B, since $U^\top U = I$,

$$(PAP + PBP)^+ = (U(U^\top AU + U^\top BU)U^\top)^+ = U(U^\top AU + U^\top BU)^+)U^\top$$

Therefore, since $PU = U = U^\top P$, it follows that

$$(PAP + PBP)^+ = (PAP + PBP)^+P = P(PAP + PBP)^+ = P(PAP + PBP)^+P.$$

As a consequence, we obtain

$$\Sigma C^\top (V + C\Sigma C^\top)^{-1} = P(\Sigma^+ + PC^\top V^{-1}CP)^+ PC^\top V^{-1}.$$

as claimed. □

B Background in Linear Control

B.1 Linear dynamical systems

Throughout, we will consider discrete-time linear dynamical systems of the form:

$$\begin{aligned}\mathbf{x}_{t+1} &= A\mathbf{x}_t + B\mathbf{u}_t + \mathbf{w}_t \\ \mathbf{y}_t &= C\mathbf{x}_t + \mathbf{v}_t,\end{aligned}$$

where $\mathbf{x}_t \in \mathbb{R}^{d_\mathbf{x}}$ is the (unobserved) system state at time $t \geq 0$, $\mathbf{u}_t \in \mathbb{R}^{d_\mathbf{u}}$ is the control applied to the system at time t, $\mathbf{w}_t \in \mathbb{R}^{d_\mathbf{x}}$ is the state noise, $\mathbf{y}_t \in \mathbb{R}^{d_\mathbf{y}}$ is the observed linear measurement of $\mathbf{x}_t$, and $\mathbf{v}_t \in \mathbb{R}^{d_\mathbf{y}}$ is the observation noise. $A \in \mathbb{R}^{d_\mathbf{x} \times d_\mathbf{x}}$, $B \in R^{d_\mathbf{x} \times d_\mathbf{u}}$, and $C \in \mathbb{R}^{d_\mathbf{y} \times d_\mathbf{x}}$.

B.2 Observability and Controllability

Definition B.1 (Controllability)**.** Consider the matrix

$$Q_r = [B, AB, \ldots, A^{r-1}B]$$

We say that a system (A, B) is *controllable* if $\mathsf{rank}(Q_{d_\mathbf{x}}) = d_\mathbf{x}$.

Further, we call a system (k, κ)-*strongly controllable* if

$$\mathsf{rank}(Q_k) = d_\mathbf{x} \quad \text{and} \quad \|(Q_k Q_k^\top)^{-1}\|_{\mathsf{op}} \leq \kappa.$$

Intuitively, a controllable system can be driven to any desired state. Indeed, if the system is noiseless and initialized at $\mathbf{x}_0 = \mathbf{0}$, then

$$\mathbf{x}_{d_\mathbf{x}} = A^{d_\mathbf{x}}\mathbf{x}_0 + \sum_{i=0}^{d_\mathbf{x}-1} A^{i-1}B\mathbf{u}_{d_\mathbf{x}-1-i} = Q_{d_\mathbf{x}}[\mathbf{u}_{d_\mathbf{x}-1}^\top, \mathbf{u}_{d_\mathbf{x}-2}^\top, \ldots, \mathbf{u}_0^\top]^\top.$$

Since Q has rank $d_\mathbf{x}$, there exists a sequence of controls which drive the system to any desired state $\mathbf{x}_{d_\mathbf{x}}$.

Definition B.2 (Stabilizability)**.** We say that a system (A, B) is *stable* if $\rho(A) < 1$, where the spectral radius of A, $\rho(A) = \max\{|\lambda_i|: i \in [d_{\mathbf{x}}]\}$, is the maximum modulus of the eigenvalues λ_i of A, which may be complex numbers.

Similarly, (LDS) is *stabilizable* if there exists a matrix $K \in \mathbb{R}^{d_{\mathbf{u}} \times d_{\mathbf{x}}}$ such that $\rho(A + BK) < 1$.

We call a controller K (κ, γ)-*stabilizing* (with $\kappa > 0$, $\gamma \in (0, 1]$) for (A, B) if

$$\|(A + BK)^t\|_{\mathsf{op}} \leq \kappa(1 - \gamma)^t \; \forall t \geq 1$$

In the case that $K = 0$, we say that A is (κ, γ)-stable.

In the case of time-varying systems (A, B), a sequence of controllers K_t is (κ, γ)-dynamically stabilizing for (A, B) if:

$$\|\prod_{t=t_1}^{t_2} A + BK_t\|_{\mathsf{op}} \leq \kappa(1 - \gamma)^{t_2 - t_1 + 1} \; \forall t_2 \geq t_1 \geq 0$$

Intuitively, a stabilizable system can be driven to the zero state (but not necessarily other states). Indeed, if we apply control $\mathbf{u}_t = K\mathbf{x}_t$ to a noiseless system at each time step, then

$$\mathbf{x}_{t+1} = (A + BK)\mathbf{x}_t = (A + BK)^{t+1}\mathbf{x}_0$$

Definition B.3 (Observability)**.** We say a pair (C, A) is *observable* if the system $(A^\top, C^\top)$ is controllable. Equivalently, we can define the matrix

$$O_r = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{bmatrix}.$$

Then we call (C, A) observable if $\mathsf{rank}(O_{d_{\mathbf{x}}}) = d_{\mathbf{x}}$.

We say (C, A) is (k, κ)-strongly observable if $(A^\top, C^\top)$ is (k, κ)-strongly controllable.

Intuitively, given $d_{\mathbf{x}}$ observations from an observable system, one can deduce the initial state. Indeed, consider a noiseless system with no inputs. Then

$$\mathbf{y}_t = CA^t\mathbf{x}_0.$$

Therefore,

$$\mathbf{y}_{0:d_{\mathbf{x}}-1} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{d_{\mathbf{x}}-1} \end{bmatrix} = O_{d_{\mathbf{x}}}\mathbf{x}_0,$$

so since $O_{d_{\mathbf{x}}}$ has rank $d_{\mathbf{x}}$, one can find $\mathbf{x}_0 = (O_{d_{\mathbf{x}}}^\top O_{d_{\mathbf{x}}})^{-1} O_{d_{\mathbf{x}}}^\top \mathbf{y}_{0:d_{\mathbf{x}}-1}$.

Definition B.4 (Detectability)**.** We say a pair (C, A) is *detectable* if $(A^\top, C^\top)$ is stabilizable. Equivalently, there exists a matrix $L \in \mathbb{R}^{d_{\mathbf{x}} \times d_{\mathbf{y}}}$ such that $\rho(A - LC) < 1$.

We say a matrix L is (κ, γ)-detectable for (C, A) if $-L^\top$ is is (κ, γ)-stabilizing for $(A^\top, C^\top)$.

Analogously, a sequence of matrices L_t is (κ, γ)-dynamically detectable for (C, A) if $-L_t^\top$ is (κ, γ)-dynamically stabilizing for $(A^\top, C^\top)$

Intuitively, in a detectable system, the current state can be estimated with vanishing error. To see this, consider the following estimate for state $\mathbf{x}_{t+1}$ (assume zero observation and state noise):

$$\hat{\mathbf{x}}_{t+1} = (A - LC)\hat{\mathbf{x}}_t + B\mathbf{u}_t + L\mathbf{y}_t.$$

Then, the estimation error satisfies:

$$e_{t+1} = \hat{\mathbf{x}}_{t+1} - \mathbf{x}_{t+1} = (A - LC)\hat{\mathbf{x}}_t + L\mathbf{y}_t - A\mathbf{x}_t = (A - LC)\hat{\mathbf{x}}_t + LC\mathbf{x}_t - A\mathbf{x}_t = (A - LC)e_t$$

Notice that this is a weaker requirement than observability – indeed, since the initial state can be determined exactly for a noiseless observable system, all future states can also be determined.

B.3 Kalman Filter

B.3.1 Setup and definitions

Here, we describe the Kalman Filter [Kal60], which provides an estimator for the unobserved system state $\mathbf{x}_t$, given the observations $\mathbf{y}_s$ and system parameters before t be and system parameters prior to before t. For a more comprehensive overview, see, e.g., the textbook treatment in [Ber05; KV15; AM05]. Consider a time-varying linear dynamical system of the form in (LDS). We assume $\mathbf{x}_0, (\mathbf{w}_t, \mathbf{v}_t)_{t\geq 0}$ are mutually independent and have finite first and second moments satisfying:

Assumption 5 (Kalman noise assumption)**.** The initial state $\mathbf{x}_0$, process noise $\mathbf{w}_t$, and observation noise $\mathbf{v}_t$ satisfy the following mean and variance conditions:

$$\begin{aligned} \mu_0 = \mathbb{E}[\mathbf{x}_0] \quad \text{and} \quad \mathbf{0} \preceq X_0 = \mathbb{E}\left[(\mathbf{x}_0 - \mu_0)(\mathbf{x}_0 - \mu_0)^\top\right] \\ \mathbb{E}[\mathbf{w}_t] = \mathbf{0} \quad \text{and} \quad \mathbf{0} \preceq W_t = \mathbb{E}\left[\mathbf{w}_t \mathbf{w}_t^\top\right] \ \forall t \geq 0 \\ \mathbb{E}[\mathbf{v}_t] = \mathbf{0} \quad \text{and} \quad \mathbf{0} \prec V_t = \mathbb{E}\left[\mathbf{v}_t \mathbf{v}_t^\top\right] \ \forall t \geq 0. \end{aligned} \tag{Noise}$$

Further, $\mathbf{x}_0, (\mathbf{w}_t, \mathbf{v}_t)_{t\geq 0}$ are uncorrelated, i.e.,

$$\begin{aligned} \mathbb{E}\left[\mathbf{x}_0 \mathbf{w}_t^\top\right] = \mathbf{0} \quad \text{and} \quad \mathbb{E}\left[\mathbf{x}_0 \mathbf{v}_t^\top\right] = \mathbf{0} = \mathbb{E}\left[\mathbf{w}_{t'} \mathbf{v}_t^\top\right] \ \forall t, t' \geq 0 \\ \text{and} \quad \mathbb{E}\left[\mathbf{w}_t \mathbf{w}_{t'}^\top\right] = \mathbf{0}, \mathbb{E}\left[\mathbf{v}_t \mathbf{v}_{t'}^\top\right] = \mathbf{0} \ \forall t' \neq t \end{aligned}$$

Definition B.5 (Kalman Filter)**.** Given a linear dynamical system (LDS) satisfying Assumption 5, the Kalman Filter is a sequence of estimates $\widehat{\mathbf{x}}_{t|t-1}$ and $\widehat{\mathbf{x}}_{t|t}$ of the unobserved state $\mathbf{x}_t$, given observations $(\mathbf{y}_s)_{s<t}$ and $(\mathbf{y}_s)_{s\leq t}$, respectively. It is defined recursively (forwards in time) according to the following linear dynamics:

$$\begin{aligned} \forall t \geq 0, \quad \widehat{\mathbf{x}}_{0|-1} = \mu_0, \quad \widehat{\mathbf{x}}_{t|t} &= \widehat{\mathbf{x}}_{t|t-1} + \tilde{L}_t(\mathbf{y}_t - C\widehat{\mathbf{x}}_{t|t-1}), \quad \widehat{\mathbf{x}}_{t+1|t} = A\widehat{\mathbf{x}}_{t|t} + B\mathbf{u}_t \\ &= (I - \tilde{L}_t C)\widehat{\mathbf{x}}_{t|t-1} + \tilde{L}_t \mathbf{y}_t, \qquad = (A - L_t C)\widehat{\mathbf{x}}_{t|t-1} + B\mathbf{u}_t + L_t \mathbf{y}_t \end{aligned}$$

The matrices $\tilde{L}_t$ and L_t are defined recursively (forwards in time) as:

$$\forall t \geq 0, \quad \tilde{L}_t = \Sigma_{t|t-1} C^\top (V_t + C\Sigma_{t|t-1} C^\top)^{-1} \quad L_t = A\tilde{L}_t,$$

where $\Sigma_{t|t-1}, \Sigma_{t|t}$ are given by:

$$\begin{aligned} \Sigma_{0|-1} &= X_0 \\ \forall t \geq 0, \quad \Sigma_{t|t} &= \Sigma_{t|t-1} - \Sigma_{t|t-1} C^\top (V_t + C\Sigma_{t|t-1} C^\top)^{-1} C\Sigma_{t|t-1} \\ \Sigma_{t+1|t} &= W_t + A^\top \Sigma_{t|t} A^\top. \end{aligned}$$

B.3.2 Important Properties of Kalman Filter

Lemma B.6 (Mean and covariance properties of Kalman filter)**.** *Suppose Assumption 5. Then the Kalman filter is unbiased, i.e.,*

$$\mathbb{E}\left[\widehat{\mathbf{x}}_{t|t-1} - \mathbf{x}_t\right] = \mathbf{0} = \mathbb{E}\left[\widehat{\mathbf{x}}_{t|t} - \mathbf{x}_t\right],$$

and the matrices $\Sigma_{t|t-1}, \Sigma_{t|t}$ satisfy:

$$\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top\right] = \Sigma_{t|t-1} \quad \textit{and} \quad \mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})^\top\right] = \Sigma_{t|t}.$$

Proof. See Appendix B.6.1 □

Another important property of the Kalman Filter is the so-called *orthogonality principle* (see, e.g., [AM05, Section 5.2])

Lemma B.7 (Orthogonality principle)**.** *Suppose that $\mathbb{E}[\mathbf{x}_0] = \mu_0$, $\mathbb{E}\left[(\mathbf{x}_0 - \mu_0)(\mathbf{x}_0 - \mu_0)^\top\right] = X_0$, $\mathbb{E}[\mathbf{v}_t] = \mathbf{0}$, $\mathbb{E}[\mathbf{w}_t] = \mathbf{0}$, $\mathbb{E}\left[\mathbf{v}_t \mathbf{v}_t^\top\right] = V_t$, $\mathbb{E}\left[\mathbf{w}_t \mathbf{w}_t^\top\right] = W_t$, and $(\mathbf{x}_0, (\mathbf{w}_t)_{t\geq 0}, (\mathbf{v}_t)_{t\geq 0})$ are jointly independent. Let a_t be any affine combination of $(\mathbf{y}_s)_{s\leq t}$, i.e., $a_t = a_{-1} + \sum_{i=0}^t M_i \mathbf{y}_i$ for fixed vector a_{-1} and matrices M_i. Let u_t be any affine function of past observations. Then,*

$$\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) a_{t-1}^\top\right] = \mathbf{0}. \quad \textit{and} \quad \mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}) a_t^\top\right] = \mathbf{0}$$

Proof. See Appendix B.6.2 □

Theorem B.8 (Best Linear Estimator, Chapter 3, Theorem 2.1, [AM05])**.** *Consider any linear system satisfying the assumptions laid out in this section. Let $\{L_t^e\}_{t\geq 0}$ be a sequence of matrices defining a linear dynamical estimator $\widehat{\mathbf{x}}_{t+1|t}^e$ for $\mathbf{x}_{t+1}$ as:*

$$\widehat{\mathbf{x}}_{t+1|t}^e = (A - L_t^e C)\widehat{\mathbf{x}}_{t|t-1}^e + L_t^e \mathbf{y}_t.$$

Define the covariance matrix of the estimation error of $\widehat{\mathbf{x}}_{t|t-1}^e$ as

$$\Sigma_{t|t-1}^e = \mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}^e)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}^e)^\top\right]$$

Then, taking $\Sigma_{t|t-1}$ as the covariance of the Kalman filter $\widehat{\mathbf{x}}_{t|t-1}$, we have

$$\Sigma_{t|t-1}^e \succeq \Sigma_{t|t-1} \quad \forall t \geq 0.$$

Moreover, the above inequality is an equality if and only if $\widehat{\mathbf{x}}_{0|-1}^e = \mu_0$ and $L_t^e = L_t$ for all $t \geq 0$.

B.3.3 Useful Properties of Kalman Matrices and Error

Proposition B.9. *Assuming Assumption 5, the Kalman filter equation $\tilde{L}_t := \Sigma_{t|t-1}C^\top(V_t + C\Sigma_{t|t-1}C^\top)^{-1}$ satisfies:*

$$\tilde{L}_t = (\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ P_t C^\top V^{-1},$$

where $P_t = \Sigma_{t|t-1}\Sigma_{t|t-1}^+$ is the orthogonal projection onto the column space of $\Sigma_{t|t-1}$ In particular, this implies that

$$(I - \tilde{L}_t C)P_t = (\Sigma_{t|t-1}^+ + C^\top V^{-1} C)^+ \Sigma_{t|t-1}^+.$$

In particular, we have that, almost surely, $P_t(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) = \mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}$, so

$$(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) = (\Sigma_{t|t-1}^+ + C^\top V^{-1} C)^+ \Sigma_{t|t-1}^+ (\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}).$$

Proof. See Appendix B.6.3 □

Lemma B.10. *We define the Kalman Filter error at a time $t \geq 0$ before and after updating with the observation $\mathbf{y}_t$ as*

$$\mathbf{e}_t^- := \mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1} \quad \textit{and} \quad \mathbf{e}_t := \mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}$$

The error almost surely satisfies the following recursion:

$$\begin{aligned}
\mathbf{e}_{t+1}^- &= A\mathbf{e}_t + \mathbf{w}_t \\
\mathbf{e}_t &= (I - \tilde{L}_t C)\mathbf{e}_t^- - \tilde{L}_t \mathbf{v}_t = (I - \tilde{L}_t C)P_t \mathbf{e}_t^- - \tilde{L}_t \mathbf{v}_t \quad \textit{where } P_t = \Sigma_{t|t-1}\Sigma_{t|t-1}^\top \\
\mathbf{e}_{t+1}^- &= (A - L_t C)\mathbf{e}_t^- + (\mathbf{w}_t - L_t \mathbf{v}_t) \\
\mathbf{e}_{t+1} &= (I - \tilde{L}_{t+1} C)P_{t+1} A \mathbf{e}_t + \widetilde{\mathbf{v}}_{t+1} \quad \textit{where } \widetilde{\mathbf{v}}_{t+1} := (I - \tilde{L}_{t+1} C)P_{t+1}\mathbf{w}_t - \tilde{L}_{t+1}\mathbf{v}_{t+1}
\end{aligned}$$

(we take $\mathbf{w}_{-1} := \mathbf{0}$). Define the following matrices:

$$\begin{aligned}
\Psi^{[t,t']} &:= \prod_{i=1}^{t'-t} (A - L_{t'-i} C) \quad \textit{and} \quad \Psi^{[t]} := \Psi^{[0,t]} \\
\Phi^{[t,t']} &:= (I - \tilde{L}_{t'} C)P_{t'} \Psi^{[t,t']} \quad \textit{and} \quad \Phi^{[t]} := \Phi^{[0,t]} \\
\widetilde{\Phi}^{[t,t']} &:= \begin{cases} \Phi^{[t+1,t']} A & \textit{if } t' > t \\ I & \textit{o.w.} \end{cases} \quad \textit{and} \quad \widetilde{\Phi}^{[t]} := \widetilde{\Phi}^{[0,t]}
\end{aligned}$$

Then, for any $t' \geq t$,

$$\begin{aligned}\mathbf{e}_{t'}^- &= [\prod_{i=1}^{t'-t}(A - L_{t'-i}C)]\mathbf{e}_t^- + \sum_{i=1}^{t'-t}[\prod_{i'=1}^{i-1}(A - L_{t'-i'}C)](\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i}) \\ &= \Psi^{[t,t']}\mathbf{e}_t^- + \sum_{i=1}^{t'-t}\Psi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i}).\end{aligned}$$

Additionally,

$$\begin{aligned}\mathbf{e}_{t'} &= \Phi^{[t,t']}\mathbf{e}_t^- - \tilde{L}_{t'}\mathbf{v}_{t'} + \sum_{i=1}^{t'-t}\Phi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i}) \\ &= \Phi^{[t,t']}\mathbf{e}_t^- - \sum_{i=1}^{t'-t+1}\widetilde{\Phi}^{[t'-i+1,t']}\tilde{L}_{t'-i+1}\mathbf{v}_{t'-i+1} + \sum_{i=1}^{t'-t}\Phi^{[t'-i+1,t']}\mathbf{w}_{t'-i}\end{aligned}$$

and

$$\begin{aligned}\mathbf{e}_{t'} &= \widetilde{\Phi}^{[t,t']}\mathbf{e}_t + \sum_{i=1}^{t'-t}\widetilde{\Phi}^{[t'-i+1,t']}((I - \tilde{L}_{t'-i+1})P_{t'-i+1}\mathbf{w}_{t'-i} - \tilde{L}_{t'-i+1}\mathbf{v}_{t'-i+1}) \\ &= \widetilde{\Phi}^{[t,t']}\mathbf{e}_t + \sum_{i=1}^{t'-t}\widetilde{\Phi}^{[t'-i+1,t']}\widetilde{\mathbf{v}}_{t'-i+1}.\end{aligned}$$

Proof. See Appendix B.6.4 □

Lemma B.11. *For any times $t' < t$, the matrix $\mathbf{e}_t(\mathbf{y}_{t'}^{\mathsf{nat}})^\top$ is mean-zero, and can be decomposed into the following sum, where each term is mean-zero:*

$$\begin{aligned}&\mathbf{e}_t(\mathbf{y}_{t'}^{\mathsf{nat}})^\top \\ &= \Phi^{[t]}\mathbf{e}_0^-\mu_0^\top(CA^{t'})^\top + \Phi^{[t]}(\mathbf{e}_0^-(\mathbf{e}_0^-)^\top - X_0)(CA^{t'})^\top \\ &\quad - \widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}(\mathbf{v}_{t'}\mathbf{v}_{t'}^\top - V) + \sum_{i=1}^{t'}\Phi^{[t'-i+1,t]}(\mathbf{w}_{t'-i}\mathbf{w}_{t'-i}^\top - W)(CA^{i-1})^\top \\ &\quad + \Phi^{[t]}\mathbf{e}_0^-(\mathbf{v}_{t'} + \sum_{i=1}^{t'}CA^{i-1}\mathbf{w}_{t'-i})^\top - \sum_{i=1}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1}(CA^{t'}\mathbf{x}_0 + \sum_{i'=1}^{t'}CA^{i'-1}\mathbf{w}_{t'-i'})^\top \\ &\quad + \sum_{i=1}^{t}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i}(CA^{t'}\mathbf{x}_0 + \mathbf{v}_{t'})^\top \\ &\quad - \sum_{\substack{i=1 \\ t-i+1\neq t'}}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1}\mathbf{v}_{t'}^\top + \sum_{i=1}^{t}\sum_{\substack{i'=1 \\ t-i\neq t'-i'}}^{t'}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i}\mathbf{w}_{t'-i'}^\top(CA^{i'-1})^\top.\end{aligned}$$

where $\Phi^{[a,b]}$ and $\widetilde{\Phi}^{[a,b]}$ are the matrices defined in Lemma B.10.

Proof. See Appendix B.6.5 □

Lemma B.12. *Suppose that the system is strongly stable, i.e., $\|A^t\|_{\mathsf{op}} \leq R_A\rho_A^t$ for all $t \geq 0$. Further, suppose that $\mathbf{x}_0$ has covariance $X_0 \succeq \mathbf{0}$ and the noise $\mathbf{v}_t, \mathbf{w}_t$ is mean-zero IID with covariance matrices $V \succ \mathbf{0}, W \succeq \mathbf{0}$ respectively. Then, for every $t \geq 0$,*

$$\mathbb{E}\left[\|\mathbf{e}_t\|_2^2\right] \leq \mathbb{E}\left[\|\mathbf{e}_t^-\|_2^2\right] \leq R_A^2(\rho_A^{2t}\mathrm{tr}(X_0) + \frac{\mathrm{tr}(W)}{1-\rho_A^2}) \leq R_A^2(\mathrm{tr}(X_0) + \frac{\mathrm{tr}(W)}{1-\rho_A^2}) := R_{e^-}.$$

Proof. See Appendix B.6.6 □

Lemma B.13 (Stability and detectability)**.** *Suppose that the system is strongly stable, i.e.,* $\|A^t\|_{\mathsf{op}} \leq R_A \rho_A^t$ *for all* $t \geq 0$*. Further, suppose that the Kalman filter is dynamically detectable, i.e.,* $\|\prod_{i=t}^{t'-1}(A - L_i C)\|_{\mathsf{op}} \leq R_{A-LC}\rho_{A-LC}^{t'-t}$ *for all* $t' \geq t$*. Then, we have that*

$$\begin{aligned}
\|\Sigma_{t|t-1}\|_{\mathsf{op}} &\leq R_{e^-} \\
\|(I - \tilde{L}_t C)\Sigma_{t|t-1}\Sigma_{t|t-1}^{+}\|_{\mathsf{op}} &\leq \max\left\{\frac{\lambda_{\max}(X_0)}{\lambda_{\min}^{+}(X_0)}, \frac{R_{e^-}}{\lambda_{\min}^{+}(W)}\right\} := R_{I-\tilde{L}C} \\
\|\Psi^{[t,t']}\|_{\mathsf{op}} &\leq R_{A-LC}\rho_{A-LC}^{t'-t} \\
\|\Phi^{[t,t']}\|_{\mathsf{op}} &\leq R_{I-\tilde{L}C}R_{A-LC}\rho_{A-LC}^{t'-t} := R_\Phi \rho_\Phi^{t'-t} \\
\|\widetilde{\Phi}^{[t,t']}\|_{\mathsf{op}} &\leq R_A R_{I-\tilde{L}C}R_{A-LC}\rho_{A-LC}^{t'-t-1} := R_{\widetilde{\Phi}}\rho_{\widetilde{\Phi}}^{t'-t},
\end{aligned}$$

where $\Psi^{[t,t']}, \Phi^{[t,t']}, \widetilde{\Phi}^{[t,t']}$ *are the quantities defined in Lemma B.10.*

Proof. See Appendix B.6.7 □

B.4 Optimal LQ Control

Here, we give an overview of some fundamental results from linear quadratic control (see, e.g., [AM07]).

Consider the same system as in the previous section (LDS), where the objective is to design a control sequence such that the following objective is minimized:

$$\sum_{t=0}^{n} \mathbf{x}_t^\top S_t \mathbf{x}_t + \mathbf{u}_t R_t \mathbf{u}_t,$$

where S_t and R_t are cost matrices (positive semi-definite WLOG) satisfying:

$$S_t \succeq \mathbf{0} \quad \text{and} \quad R_t \succ \mathbf{0}\ \forall t \geq 0.$$

Then, given the information state $I_t = (\mathbf{y}_0, \mathbf{u}_0, \ldots, \mathbf{y}_{t-1}, \mathbf{u}_{t-1}, \mathbf{y}_t)$, the value function for any time t is given as:

$$\begin{aligned}
V_t(I_t) &= \min_{\mathbf{u}_t} \mathbb{E}\left[\mathbf{x}_t^\top S_t + \mathbf{u}_t^\top R_t \mathbf{u}_t + V_{t+1}(I_t, \mathbf{u}_t, y_{t+1}) \mid I_t\right] \\
&= \widehat{\mathbf{x}}_{t|t}^\top P_t \widehat{\mathbf{x}}_{t|t} + \sum_{\ell=t}^{n} \mathrm{tr}(S_\ell \Sigma_{\ell|\ell}) + \sum_{\ell=t+1}^{n} \mathrm{tr}(P_\ell(\Sigma_{\ell|\ell-1} - \Sigma_{\ell|\ell})) \\
&= \widehat{\mathbf{x}}_{t|t}^\top P_t \widehat{\mathbf{x}}_{t|t} + \sum_{\ell=t}^{n} \mathrm{tr}(S_\ell \Sigma_{\ell|\ell}) + \sum_{\ell=t+1}^{n} \mathrm{tr}(P_\ell(A_{\ell-1}\Sigma_{\ell-1|\ell-1}A_{\ell-1}^\top + W_{\ell-1} - \Sigma_{\ell|\ell})) \\
&= \widehat{\mathbf{x}}_{t|t}^\top P_t \widehat{\mathbf{x}}_{t|t} + \mathrm{tr}(S_t \Sigma_{t|t}) + \sum_{\ell=t+1}^{n} \mathrm{tr}(P_\ell W_{\ell-1}) + \sum_{\ell=t+1}^{n} \mathrm{tr}((S_\ell - P_\ell)\Sigma_{\ell|\ell}) + \sum_{\ell=t+1}^{n} \mathrm{tr}(P_\ell(A_{\ell-1}\Sigma_{\ell-1|\ell-1}A_{\ell-1}^\top)),
\end{aligned}$$

where P_t is defined recursively (backwards in time) as:

$$\begin{aligned}
P_N &= S_N \\
P_t &= S_t + A^\top P_{t+1} A - A^\top P_{t+1} B(R_t + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A
\end{aligned}$$

The optimal cost therefore takes the form:

$$\begin{aligned}
J^* = \mathbb{E}\left[V_0(I_0)\right] = \mathbb{E}\left[\mathbf{x}_0^\top P_0 \mathbf{x}_0\right] &+ \sum_{t=1}^{n} \mathrm{tr}(P_t W_{t-1}) \\
&+ \sum_{t=0}^{n} \mathrm{tr}((S_t - P_t)\Sigma_{t|t}) + \sum_{t=1}^{n} \mathrm{tr}(P_t(A_{t-1}\Sigma_{t-1|t-1}A_{t-1}^\top)),
\end{aligned}$$

where the first two terms correspond to the LQR cost without partial observation, and the second two terms correspond to the additional cost of partial state observation.

Furthermore, the optimal control is given as:

$$\mathbf{u}_t = K_t \mathbb{E}\left[\mathbf{x}_t \mid I_t\right] \quad \text{where} \quad K_t = -(R_t + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A.$$

In the time-invariant setting when $(A, B, S_t, R_t) = (A, B, S, R)$, (A, B) is controllable, and $(A, S^{1/2})$ is observable, then $P_t \to P$, where P is the unique PSD solution to the DARE:

$$P = S + A^\top P A - A^\top P B (R + B^\top P B)^{-1} B^\top P A$$

and $K_t \to K$, where

$$K = -(R + B^\top P B)^{-1} B^\top P A.$$

B.5 Linear Dynamical Controllers

Definition B.14 (Closed-loop system)**.** Given the system (LDS) and a single-sensor LDC $(A^{(s)}_{\mathsf{ldc}}, B^{(s)}_{\mathsf{ldc}}, C^{(s)}_{\mathsf{ldc}}, D^{(s)}_{\mathsf{ldc}})$ on sensor $s \in \mathcal{S}$, the *closed-loop* dynamics of the LDS using controls specified by the LDC evolves according to

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \mathbf{h}^{\mathsf{ldc}}_{t+1} \end{bmatrix} = \overbrace{\begin{bmatrix} A + BD^{(s)}_{\mathsf{ldc}} C^{(s)} & BC^{(s)}_{\mathsf{ldc}} \\ B^{(s)}_{\mathsf{ldc}} C^{(s)} & A^{(s)}_{\mathsf{ldc}} \end{bmatrix}}^{A^{(s)}_{\mathsf{ldc,cl}}} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{h}^{\mathsf{ldc}}_t \end{bmatrix} + \overbrace{\begin{bmatrix} I & BD^{(s)}_{\mathsf{ldc}} \\ \mathbf{0} & B^{(s)}_{\mathsf{ldc}} \end{bmatrix}}^{B^{(s)}_{\mathsf{ldc,cl}}} \begin{bmatrix} \mathbf{w}_t \\ \mathbf{v}_t \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{u}_t \end{bmatrix} = \underbrace{\begin{bmatrix} C & \mathbf{0} \\ D^{(s)}_{\mathsf{ldc}} C & C^{(s)}_{\mathsf{ldc}} \end{bmatrix}}_{C^{(s)}_{\mathsf{ldc,cl}}} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{h}^{\mathsf{ldc}}_t \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{0} & I \\ \mathbf{0} & D^{(s)}_{\mathsf{ldc}} \end{bmatrix}}_{D^{(s)}_{\mathsf{ldc}}} \begin{bmatrix} \mathbf{w}_t \\ \mathbf{v}_t \end{bmatrix}.$$

We will further denote $B^{(s)}_{\mathsf{ldc,cl},\mathbf{v}}$ as the second column of $B^{(s)}_{\mathsf{ldc,cl}}$ and $C^{(s)}_{\mathsf{ldc,cl},\mathbf{u}}$ as the second row of $C^{(s)}_{\mathsf{ldc,cl}}$:

$$B^{(s)}_{\mathsf{ldc,cl},\mathbf{v}} = \begin{bmatrix} BD^{(s)}_{\mathsf{ldc}} \\ B^{(s)}_{\mathsf{ldc}} \end{bmatrix} \quad \text{and} \quad C^{(s)}_{\mathsf{ldc,cl},\mathbf{u}} = \begin{bmatrix} D^{(s)}_{\mathsf{ldc}} C^{(s)} & C^{(s)}_{\mathsf{ldc}} \end{bmatrix}$$

Remark B.15. *Any linear controller of the form $\mathbf{u}_t = K\widehat{\mathbf{x}}_{t|t}$ (in particular, the asymptotically optimal time-homogeneous Kalman filter on sensor $s \in \mathcal{S}$), can be expressed as the following LDC:*

$$A^{(s)}_{\mathsf{ldc}} = (A + BK)(I - \tilde{L}^{(s)} C^{(s)}), \qquad B^{(s)}_{\mathsf{ldc}} = (A + BK)\tilde{L}^{(s)}$$
$$C^{(s)}_{\mathsf{ldc}} = K(I - \tilde{L}^{(s)} C^{(s)}), \qquad D^{(s)}_{\mathsf{ldc}} = K\tilde{L}^{(s)},$$

where $\tilde{L}^{(s)}$ is the asymptotic Kalman gain matrix associated with sensor s. Under this parametrization, the internal state of the LDC $\mathbf{h}^{\mathsf{ldc}}_t = \widehat{\mathbf{x}}^{(s)}_{t|t-1}$ is the asymptotic Kalman filter on sensor s, given measurements before time t.

B.6 Deferred Proofs

B.6.1 Proof of Lemma B.6

Proof of Lemma B.6. This is a standard fact: see, e.g., [AM05, Section 5.4]. We prove the claims here for completeness.

We proceed by induction. In the base case of $t = 0$, notice that, by definition, $\mathbb{E}\left[\mathbf{x}_0\right] = \mu_0 = \widehat{\mathbf{x}}_{t|t-1}$, so, by definition of $\Sigma_{0|-1}$ and X_0,

$$\mathbb{E}\left[(\mathbf{x}_0 - \widehat{\mathbf{x}}_{0|-1})(\mathbf{x}_0 - \widehat{\mathbf{x}}_{0|-1})^\top\right] = \mathbb{E}\left[(\mathbf{x}_0 - \mu_0)(\mathbf{x}_0 - \mu_0)^\top\right] = X_0 = \Sigma_{0|-1}.$$

Now, suppose $\mathbb{E}\left[\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right] = \mathbf{0}$ and $\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top\right] = \Sigma_{t|t-1}$ for some $t \geq 0$. Then, by definition of $\widehat{\mathbf{x}}_{t|t}$, Definition B.5, and since $\mathbf{y}_t = C\mathbf{x}_t + \mathbf{v}_t$ and $\mathbb{E}\left[\mathbf{v}_t\right] = \mathbf{0}$,

$$\mathbb{E}\left[\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}\right] = \mathbb{E}\left[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) + \tilde{L}_t \mathbf{v}_t\right] = (I - \tilde{L}_t C)\mathbb{E}\left[\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right] = \mathbf{0},$$

where in the last equality, we apply our induction hypothesis. Additionally,

$$\begin{aligned}\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})^\top\right] &= (I - \tilde{L}_t C)\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top\right](I - \tilde{L}_t C) \\ &\quad + (I - \tilde{L}_t C)\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})\mathbf{v}_t^\top\right]\tilde{L}_t^\top \\ &\quad + \tilde{L}_t\mathbb{E}\left[\mathbf{v}_t(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top\right](I - \tilde{L}_t C)^\top \\ &\quad + \tilde{L}_t\mathbb{E}\left[\mathbf{v}_t\mathbf{v}_t^\top\right]\tilde{L}_t^\top\end{aligned}$$

Now, since $\mathbf{v}_t$ is uncorrelated with $\mathbf{x}_0, (\mathbf{w}_s)_{s\le t}, (\mathbf{v}_s)_{s<t}$, it follows from linearity of the Kalman filter and the system that $\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_t)\mathbf{v}_t^\top\right] = \mathbf{0}$. Thus, since $\mathbb{E}\left[\mathbf{v}_t\mathbf{v}_t^\top\right] = V_t$ and by the induction hypothesis,

$$\begin{aligned}\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})^\top\right] &= (I - \tilde{L}_t C)\Sigma_{t|t-1}(I - \tilde{L}_t C) + \tilde{L}_t V_t \tilde{L}_t^\top \\ &= \Sigma_{t|t-1} - \tilde{L}_t C\Sigma_{t|t-1} - \Sigma_{t|t-1}C^\top\tilde{L}_t^\top + \tilde{L}_t(V_t + C\Sigma_{t|t-1}C^\top)\tilde{L}_t^\top.\end{aligned}$$

Plugging in the definition $\tilde{L}_t = \Sigma_{t|t-1}C^\top(V_t + C\Sigma_{t|t-1}C^\top)^{-1}$, and recalling the definition of $\Sigma_{t|t}$ from Definition B.5, we conclude:

$$\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})^\top\right] = \Sigma_{t|t-1} - \Sigma_{t|t-1}C^\top(V_t + C\Sigma_{t|t-1}C^\top)^{-1}C\Sigma_{t|t-1} = \Sigma_{t|t}.$$

Further, since $\mathbb{E}[\mathbf{w}_t] = \mathbf{0}$,

$$\mathbb{E}\left[\mathbf{x}_{t+1} - \widehat{\mathbf{x}}_{t+1|t}\right] = \mathbb{E}\left[A(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}) + \mathbf{w}_t\right] = A\mathbb{E}\left[\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}\right] = \mathbf{0},$$

and since $\mathbf{w}_t$ is similarly uncorrelated with $\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}$ by the same arguments as above, and $\mathbb{E}\left[\mathbf{w}_t\mathbf{w}_t^\top\right] = W_t$,

$$\begin{aligned}\mathbb{E}\left[(\mathbf{x}_{t+1} - \widehat{\mathbf{x}}_{t+1|t})(\mathbf{x}_{t+1} - \widehat{\mathbf{x}}_{t+1|t})^\top\right] &= A\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})^\top\right]A^\top + \mathbb{E}\left[\mathbf{w}_t\mathbf{w}_t^\top\right] \\ &\quad + A\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})\mathbf{w}_t^\top\right] + \mathbb{E}\left[\mathbf{w}_t(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})^\top\right]A^\top \\ &= A\Sigma_{t|t}A^\top + W_t = \Sigma_{t+1|t}.\end{aligned}$$

Therefore, we have established the induction step, which establishes the claim. □

B.6.2 Proof of Lemma B.7

Proof of Lemma B.7. Notice that it suffices to prove the claim when $\mathbf{u}_t = \mathbf{0}$, since it implies the more general case when $\mathbf{u}_t$ is an affine function of $(\mathbf{y}_\ell)_{\ell\le t}$.

We will prove the claim inductively. In the base case of $t = 0$, notice that, since a_{-1} is fixed, and $\mathbb{E}\left[x_t - \widehat{\mathbf{x}}_{t|t-1}\right] = \mathbf{0}$,

$$\mathbb{E}\left[(\mathbf{x}_0 - \widehat{\mathbf{x}}_{0|-1})a_{-1}^\top\right] = \mathbb{E}\left[(\mathbf{x}_0 - \widehat{\mathbf{x}}_{0|-1})\right]a_{-1}^\top = \mathbf{0}.$$

Now, suppose $\mathbb{E}\left[(x_t - \widehat{\mathbf{x}}_{t|t-1})a_{t-1}^\top\right] = \mathbf{0}$ for some $t \ge 0$. We will first show that $\mathbb{E}\left[(x_t - \widehat{\mathbf{x}}_{t|t})a_t^\top\right] = \mathbf{0}$, and then will use this to show that $\mathbb{E}\left[(\mathbf{x}_{t+1} - \widehat{\mathbf{x}}_{t+1|t})a_t\right] = \mathbf{0}$.

Using the fact that $\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t} = (I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t\mathbf{v}_t$, it follows that:

$$\begin{aligned}\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})a_t^\top\right] &= \mathbb{E}\left[[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t\mathbf{v}_t](a_{t-1} + M_t\mathbf{y}_t)^\top\right] \\ &= (I - \tilde{L}_t C)\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})a_{t-1}^\top\right] - \tilde{L}_t\mathbb{E}\left[\mathbf{v}_t a_{t-1}^\top\right] \\ &\quad + \mathbb{E}\left[[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t\mathbf{v}_t]\mathbf{y}_t^\top\right]M_t^\top.\end{aligned}$$

Now, by the induction hypothesis, $\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})a_{t-1}^\top\right] = \mathbf{0}$. Further, since a_{t-1} depends only on $(\mathbf{y}_s)_{s<t}$, and $\mathbf{v}_t$ is independent of those $\mathbf{y}_s$, we have that $\mathbb{E}\left[\mathbf{v}_t a_{t-1}^\top\right] = \mathbf{0}$. Thus, focusing on the third term, we have that, by definition of $\mathbf{y}_t$,

$$\begin{aligned}\mathbb{E}\left[[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t\mathbf{v}_t]\mathbf{y}_t^\top\right] &= \mathbb{E}\left[[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t\mathbf{v}_t](C\mathbf{x}_t + \mathbf{v}_t)\right] \\ &= \mathbb{E}\left[[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t\mathbf{v}_t](C(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) + \mathbf{v}_t)^\top\right] \\ &\quad + \mathbb{E}\left[[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t\mathbf{v}_t]\widehat{\mathbf{x}}_{t|t-1}^\top\right]C^\top.\end{aligned}$$

Now, since $\widehat{\mathbf{x}}_{t|t-1}$ is an affine combination of $(\mathbf{y}_s)_{s<t}$, the induction hypothesis implies that $\mathbb{E}\left[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1} - \tilde{L}_t \mathbf{v}_t)\widehat{\mathbf{x}}_{t|t-1}^\top\right] = \mathbf{0}$. For the remaining terms, we have,

$$\begin{aligned}
&\mathbb{E}\left[[(I - \tilde{L}_t C)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) - \tilde{L}_t \mathbf{v}_t](C(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) + \mathbf{v}_t)^\top\right] \\
&= (I - \tilde{L}_t C)\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top\right] C^\top - \tilde{L}_t \mathbb{E}\left[\mathbf{v}_t \mathbf{v}_t^\top\right] \\
&\quad + (I - \tilde{L}_t C)\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})\mathbf{v}_t\right] - \tilde{L}_t \mathbb{E}\left[\mathbf{v}_t(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top\right] C^\top \\
&= (I - \tilde{L}_t C)\Sigma_{t|t-1} C^\top - \tilde{L}_t V_t \\
&= \Sigma_{t|t-1} C^\top - \tilde{L}_t (C\Sigma_{t|t-1} C^\top + V_t) \\
&= \mathbf{0}.
\end{aligned}$$

where the second equality used the fact that $\mathbf{v}_t$ is sampled independently of the past, and the third equality used the definition of $\Sigma_{t|t-1}$ and the fact that $\mathbf{v}_t$ has covariance V_t. The final equality follows from the definition of $\tilde{L}_t = \Sigma_{t|t-1} C^\top (V_t + C\Sigma_{t|t-1} C^\top)^{-1}$. Combining these facts, we conclude that

$$\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}) a_t^\top\right] = \mathbf{0}.$$

Now, using the fact that $\mathbf{x}_{t+1} - \widehat{\mathbf{x}}_{t+1|t} = A(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}) + \mathbf{w}_t$, we have that

$$\mathbb{E}\left[(\mathbf{x}_{t+1} - \widehat{\mathbf{x}}_{t+1|t}) a_t^\top\right] = A\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}) a_t^\top\right] + \mathbb{E}\left[\mathbf{w}_t a_t^\top\right] = \mathbf{0},$$

which follows by the above equality, combined with the fact that $\mathbf{w}_t$ is independent of $(\mathbf{y}_s)_{s \le t}$. Thus, we have established the claim by induction. □

B.6.3 Proof of Proposition B.9

Proof of Proposition B.9. The first equality on $\tilde{L}_t$ follows immediately from Lemma A.3, recalling the definition of $\tilde{L}_t$ from Definition B.5.

To see the second equality, first notice that, by definition of the projection P_t, and from the above,

$$\begin{aligned}
I - \tilde{L}_t C &= I - (\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ P_t C^\top V^{-1} C \\
&= I - P_t(\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ P_t C^\top V^{-1} C.
\end{aligned}$$

Thus, by definition of P_t, and by the properties of pseudoinverse (namely, $AA^+A = A$),

$$\begin{aligned}
(I - \tilde{L}_t C)P_t &= P_t - P_t(\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ P_t C^\top V^{-1} C P_t \\
&= P_t(I - (\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ P_t C^\top V^{-1} C P_t) \\
&= P_t(\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ (\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t) \\
&\quad - P_t(\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ P_t C^\top V^{-1} C P_t \\
&= P_t(\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ \Sigma_{t|t-1}^+ \\
&= (\Sigma_{t|t-1}^+ + P_t C^\top V^{-1} C P_t)^+ \Sigma_{t|t-1}^+,
\end{aligned}$$

which establishes the second claim.

For the final claim, it suffices to show that, with probability 1,

$$\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1} = P_t(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) = \Sigma_{t|t-1}\Sigma_{t|t-1}^\top(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}).$$

This follows by observing that, by definition of $P_t = \Sigma_{t|t-1}\Sigma_{t|t-1}^+ = \Sigma_{t|t-1}^+\Sigma_{t|t-1}$,

$$\begin{aligned}
\mathbb{E}\left[(I - P_t)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top (I - P_t)\right] &= (I - P_t)\mathbb{E}\left[(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top\right](I - P_t) \\
&= (I - P_t)\Sigma_{t|t-1}(I - P_t) \\
&= \mathbf{0}.
\end{aligned}$$

Thus, since $(I - P_t)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1})^\top (I - P_t) \succeq \mathbf{0}$, it follows that $(I - P_t)(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}) = 0$ a.s. □

B.6.4 Proof of Lemma B.10

Proof. Recall that, for the projection matrix $P_t = \Sigma_{t|t-1}\Sigma^{+}_{t|t-1}$, $P_t \mathbf{e}^{-}_t = \mathbf{e}^{-}_t$ (see Proposition B.9) and $P_{t+1}\mathbf{w}_t = \mathbf{w}_t$ (since $\Sigma_{t+1|t} = W + A\Sigma_{t|t}A^{+}$).

Given this, the one-step recursions follow immediately from the definition of the Kalman filter, Definition B.5.

The expression for $\mathbf{e}^{-}_{t'}$ follows immediately by unrolling the one-step definition recursively.

The first expression for $\mathbf{e}_{t'}$ follows from the one-step expression and the unrolled expression for $\mathbf{e}^{-}_{t'}$, since

$$\begin{aligned}
\mathbf{e}_{t'} &= (I - \tilde{L}_{t'}C)\mathbf{e}^{-}_{t'} - \tilde{L}_{t'}\mathbf{v}_{t'} \\
&= (I - \tilde{L}_{t'}C)\left(\Psi^{[t,t']}\mathbf{e}^{-}_{t'} + \sum_{i=1}^{t'-t}\Psi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i})\right) - \tilde{L}_{t'}\mathbf{v}_{t'} \\
&= \Phi^{[t,t']}\mathbf{e}^{-}_{t} - \tilde{L}_{t'}\mathbf{v}_{t'} + \sum_{i=1}^{t'-t}\Phi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i}),
\end{aligned}$$

where the last equality follows since $\Phi^{[t,t']} = (I - \tilde{L}_{t'}C)\Psi^{[t,t']}$ by definition. Additionally, recalling that $\widetilde{\Phi}^{[t',t']} = I$ and $\widetilde{\Phi}^{[t'-i+1,t']} = \widetilde{\Phi}^{[t'-i+2,t']}A$ for $i > 1$, and since $L_t = A\tilde{L}_t$, the above implies that

$$\begin{aligned}
\mathbf{e}_{t'} &= \Phi^{[t,t']}\mathbf{e}^{-}_{t} - \widetilde{\Phi}^{[t',t']}\tilde{L}_{t'}\mathbf{v}_{t'} - \sum_{i=1}^{t'-t}\Phi^{[t'-i+1,t']}A\tilde{L}_{t'-i}\mathbf{v}_{t'-i} + \sum_{i=1}^{t'-t}\Phi^{[t'-i+1,t']}\mathbf{w}_{t'-i} \\
&= \Phi^{[t,t']}\mathbf{e}^{-}_{t} - \sum_{i=1}^{t'-t+1}\widetilde{\Phi}^{[t'-i+1,t']}\tilde{L}_{t'-i+1}\mathbf{v}_{t'-i+1} + \sum_{i=1}^{t'-t}\Phi^{[t'-i+1,t']}\mathbf{w}_{t'-i}.
\end{aligned}$$

We now focus on the proof of the last statement. Let us assume without loss of generality that $t' > t$, since otherwise the statement follows immediately. Indeed, we start with

$$\mathbf{e}^{-}_{t'} = \Psi^{[t+1,t']}\mathbf{e}^{-}_{t+1} + \sum_{i=1}^{t'-t-1}\Psi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i})$$

Then, since $\mathbf{e}_{t'} = (I - \tilde{L}_{t'}C)P_{t'}\mathbf{e}^{-}_{t'} - \tilde{L}_{t'}\mathbf{v}_{t'}$ and $\mathbf{e}^{-}_{t+1} = A\mathbf{e}_t + \mathbf{w}_t$, we obtain:

$$\begin{aligned}
\mathbf{e}_{t'} &= (I - \tilde{L}_{t'}C)P_{t'}[\Psi^{[t+1,t']}\mathbf{e}^{-}_{t+1} + \sum_{i=1}^{t'-t-1}\Psi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i})] - \tilde{L}_{t'}\mathbf{v}_{t'} \\
&= (I - \tilde{L}_{t'}C)P_{t'}[\Psi^{[t+1,t']}(A\mathbf{e}_t + \mathbf{w}_t) + \sum_{i=1}^{t'-t-1}\Psi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i})] - \tilde{L}_{t'}\mathbf{v}_{t'}.
\end{aligned}$$

Now, recall that $\widetilde{\Phi}^{[t,t']} = (I - \tilde{L}_{t'}C)P_{t'}\Psi^{[t+1,t']}A$ for $t' > t$, and $\widetilde{\Phi}^{[t',t']} = I$ by definition. Thus,

$$\begin{aligned}
\mathbf{e}_{t'} &= \widetilde{\Phi}^{[t,t']}\mathbf{e}_t + (I - \tilde{L}_{t'}C)P_{t'}\Psi^{[t+1,t']}\mathbf{w}_t - \tilde{L}_{t'}\mathbf{v}_{t'} + \sum_{i=1}^{t'-t-1}(I - \tilde{L}_{t'})P_{t'}\Psi^{[t'-i+1,t']}(\mathbf{w}_{t'-i} - L_{t'-i}\mathbf{v}_{t'-i}) \\
&= \widetilde{\Phi}^{[t,t']}\mathbf{e}_t + [(I - \tilde{L}_{t'})P_{t'}\mathbf{w}_{t'-1} - \tilde{L}_{t'}\mathbf{v}_{t'}] \\
&\quad + \sum_{i=2}^{t'-t}(I - \tilde{L}_{t'}C)P_{t'}\Psi^{[t'-i+1,t']}\mathbf{w}_{t'-i} - \sum_{i=1}^{t'-t-1}(I - \tilde{L}_{t'}C)P_{t'}\Psi^{[t'-i+1,t']}L_{t'-i}\mathbf{v}_{t'-i} \\
&= \widetilde{\Phi}^{[t,t']}\mathbf{e}_t + \widetilde{\mathbf{v}}_{t'} \\
&\quad + \sum_{i=2}^{t'-t}(I - \tilde{L}_{t'}C)P_{t'}\Psi^{[t'-i+2,t']}A(I - \tilde{L}_{t'-i+1})\mathbf{w}_{t'-i} - \sum_{i=1}^{t'-t-1}(I - \tilde{L}_{t'}C)P_{t'}\Psi^{[t'-i+1,t']}A\tilde{L}_{t'-i}\mathbf{v}_{t'-i} \\
&= \widetilde{\Phi}^{[t,t']}\mathbf{e}_t + \widetilde{\mathbf{v}}_{t'} + \sum_{i=2}^{t'-t}\widetilde{\Phi}^{[t'-i+1,t']}(I - \tilde{L}_{t'-i+1})\mathbf{w}_{t'-i} - \sum_{i=1}^{t'-t-1}\widetilde{\Phi}^{[t'-i,t']}\tilde{L}_{t'-i}\mathbf{v}_{t'-i} \\
&= \widetilde{\Phi}^{[t,t']}\mathbf{e}_t + \widetilde{\mathbf{v}}_{t'} + \sum_{i=2}^{t'-t}\widetilde{\Phi}^{[t'-i+1,t']}((I - \tilde{L}_{t'-i+1})\mathbf{w}_{t'-i} + \tilde{L}_{t'-i+1}\mathbf{v}_{t'-i+1}).
\end{aligned}$$

The claim then follows from the fact that $\mathbf{w}_t = P_{t+1}\mathbf{w}_t$ (as discussed above), and by plugging in the definition of $\widetilde{\mathbf{v}}_t$. □

B.6.5 Proof of Lemma B.11

Proof. Notice that $\mathbb{E}\left[\mathbf{e}_t(\mathbf{y}_{t'}^{\mathsf{nat}})^\top\right] = \mathbf{0}$ is a consequence of the orthogonality principle, Lemma B.7. Thus, we focus on the decomposing this quantity.

By Lemma B.10, we can decompose $\mathbf{e}_t$ as

$$\mathbf{e}_t = \Phi^{[t]}\mathbf{e}_0^- - \sum_{i=1}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1} + \sum_{i=1}^{t}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i},$$

and by definition of $\mathbf{y}_{t'}^{\mathsf{nat}}$ in Definition 3.1,

$$\mathbf{y}_{t'}^{\mathsf{nat}} = CA^{t'}\mathbf{x}_0 + \mathbf{v}_{t'} + \sum_{i=1}^{t'}CA^{i-1}\mathbf{w}_{t'-i}.$$

Therefore, we have the following decomposition.

$$\begin{aligned}
&\mathbf{e}_t(\mathbf{y}_{t'}^{\mathsf{nat}})^\top \\
&= \left(\Phi^{[t]}\mathbf{e}_0^- - \sum_{i=1}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1} + \sum_{i=1}^{t}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i}\right)\left(CA^{t'}\mathbf{x}_0 + \mathbf{v}_{t'} + \sum_{i=1}^{t'}CA^{i-1}\mathbf{w}_{t'-i}\right)^\top \\
&= \Phi^{[t]}\mathbf{e}_0^-\mathbf{x}_0^\top(CA^{t'})^\top - \widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}\mathbf{v}_{t'}\mathbf{v}_{t'}^\top + \sum_{i=1}^{t'}\Phi^{[t'-i+1,t]}\mathbf{w}_{t'-i}\mathbf{w}_{t'-i}^\top(CA^{i-1})^\top \\
&\quad + \Phi^{[t]}\mathbf{e}_0^-(\mathbf{v}_{t'} + \sum_{i=1}^{t'}CA^{i-1}\mathbf{w}_{t'-i})^\top - \sum_{i=1}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1}(CA^{t'}\mathbf{x}_0 + \sum_{i'=1}^{t'}CA^{i'-1}\mathbf{w}_{t'-i'})^\top \\
&\quad + \sum_{i=1}^{t}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i}(CA^{t'}\mathbf{x}_0 + \mathbf{v}_{t'})^\top \\
&\quad - \sum_{\substack{i=1 \\ t-i+1\neq t'}}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1}\mathbf{v}_{t'}^\top + \sum_{i=1}^{t}\sum_{\substack{i'=1 \\ t-i\neq t'-i'}}^{t'}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i}\mathbf{w}_{t'-i'}^\top(CA^{i'-1})^\top.
\end{aligned}$$

Focusing on the last equality above, each individual term on the first line is not mean-zero. However, the terms on the remaining lines are all products of independent, mean-zero terms (by assumption on $\mathbf{x}_0, (\mathbf{w}_t, \mathbf{v}_t)$), and thus are themselves all mean-zero. Rewriting the terms on the first line, we have:

$$\begin{aligned}
&\Phi^{[t]}\mathbf{e}_0^-\mathbf{x}_0^\top(CA^{t'})^\top - \widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}\mathbf{v}_{t'}\mathbf{v}_{t'}^\top + \sum_{i=1}^{t'}\Phi^{[t'-i+1,t]}\mathbf{w}_{t'-i}\mathbf{w}_{t'-i}^\top(CA^{i-1})^\top \\
&= \Phi^{[t]}X_0(CA^{t'})^\top - \widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}V + \sum_{i=1}^{t'}\Phi^{[t'-i+1,t]}W(CA^{i-1})^\top \\
&\quad + \Phi^{[t]}\mathbf{e}_0^-\mu_0^\top(CA^{t'})^\top + \Phi^{[t]}(\mathbf{e}_0^-(\mathbf{e}_0^-)^\top - X_0)(CA^{t'})^\top \\
&\quad - \widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}(\mathbf{v}_{t'}\mathbf{v}_{t'}^\top - V) + \sum_{i=1}^{t'}\Phi^{[t'-i+1,t]}(\mathbf{w}_{t'-i}\mathbf{w}_{t'-i}^\top - W)(CA^{i-1})^\top.
\end{aligned}$$

Now, by the orthogonality principle, the above quantity must be mean-zero. Since all terms on the bottom two lines are mean-zero (by definition of X_0, W, V), it follows that

$$\Phi^{[t]}X_0(CA^{t'})^\top - \widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}V + \sum_{i=1}^{t'}\Phi^{[t'-i+1,t]}W(CA^{i-1})^\top = \mathbf{0}$$

(One can also check this directly, from the definitions). Thus, we have the following decomposition into mean-zero terms:

$$\begin{aligned}
&\mathbf{e}_t(\mathbf{y}_{t'}^{\mathsf{nat}})^\top \\
&= \Phi^{[t]}\mathbf{e}_0^-\mu_0^\top(CA^{t'})^\top + \Phi^{[t]}(\mathbf{e}_0^-(\mathbf{e}_0^-)^\top - X_0)(CA^{t'})^\top \\
&\quad - \widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}(\mathbf{v}_{t'}\mathbf{v}_{t'}^\top - V) + \sum_{i=1}^{t'}\Phi^{[t'-i+1,t]}(\mathbf{w}_{t'-i}\mathbf{w}_{t'-i}^\top - W)(CA^{i-1})^\top \\
&\quad + \Phi^{[t]}\mathbf{e}_0^-(\mathbf{v}_{t'} + \sum_{i=1}^{t'}CA^{i-1}\mathbf{w}_{t'-i})^\top - \sum_{i=1}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1}(CA^{t'}\mathbf{x}_0 + \sum_{i=1}^{t'}CA^{i-1}\mathbf{w}_{t'-i})^\top \\
&\quad + \sum_{i=1}^{t}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i}(CA^{t'}\mathbf{x}_0 + \mathbf{v}_{t'})^\top \\
&\quad - \sum_{\substack{i=1 \\ t-i+1\neq t'}}^{t+1}\widetilde{\Phi}^{[t-i+1,t]}\tilde{L}_{t-i+1}\mathbf{v}_{t-i+1}\mathbf{v}_{t'}^\top + \sum_{i=1}^{t}\sum_{\substack{i'=1 \\ t-i\neq t'-i'}}^{t'}\Phi^{[t-i+1,t]}\mathbf{w}_{t-i}\mathbf{w}_{t'-i'}^\top(CA^{i'-1})^\top.
\end{aligned}$$

as claimed. □

B.6.6 Proof of Lemma B.12

Proof of Lemma B.12. Recall that the covariance matrix of $\mathbf{e}_t, \mathbf{e}_{t+1}^-$ satisfies:

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1}C^\top(V + C\Sigma_{t|t-1}C^\top)^{-1}C\Sigma_{t|t-1}$$

Now, since $V \succ \mathbf{0}$ and $\Sigma_{t|t-1} \succeq \mathbf{0}$, it follows that $(V + C\Sigma_{t|t-1}C^\top)^{-1} \succ \mathbf{0}$. This, combined with the above expression for $\Sigma_{t|t}$ implies that $\Sigma_{t|t} \preceq \Sigma_{t|t-1}$. Therefore,

$$\mathbb{E}\left[\|\mathbf{e}_t\|_2^2\right] = \mathbb{E}\left[\mathrm{tr}(\mathbf{e}_t^\top\mathbf{e}_t)\right] = \mathbb{E}\left[\mathrm{tr}(\mathbf{e}_t\mathbf{e}_t^\top)\right] = \mathrm{tr}(\Sigma_{t|t}) \leq \mathrm{tr}(\Sigma_{t|t-1}) = \mathrm{tr}(\mathbb{E}\left[\mathbf{e}_t^-(\mathbf{e}_t^-)^\top\right]) = \mathbb{E}\left[\|\mathbf{e}_t^-\|_2^2\right],$$

which establishes the first inequality. For the second, notice that, by definition of $\Sigma_{t+1|t}$, and by the above,

$$\Sigma_{t+1|t} = W + A\Sigma_{t|t}A^\top \preceq W + A\Sigma_{t|t-1}A^\top.$$

In particular, this implies that

$$\Sigma_{t+1|t} \preceq A^tX_0(A^t)^\top + \sum_{i=0}^{t-1}A^iW(A^i)^\top.$$

Therefore, by our bound on $\|A\|_{\mathsf{op}}$ from the strong stability assumption,

$$\begin{aligned}\mathbb{E}\left[\|\mathbf{e}^-_{t+1}\|_2^2\right] = \mathrm{tr}(\Sigma_{t+1|t}) &\le \mathrm{tr}(A^t X_0 (A^t)^\top) + \sum_{i=0}^{t-1} \mathrm{tr}(A^i W (A^i)^\top) \\ &= \mathbb{E}\left[\|A^t \mathbf{e}^-_0\|_2^2\right] + \sum_{i=0}^{t-1} \mathbb{E}\left[\|A^i \mathbf{w}_{t-i}\|_2^2\right] \\ &\le R_A^2(\rho_A^{2t}\mathrm{tr}(X_0) + \sum_{i=0}^{t-1} \rho_A^{2i}\mathrm{tr}(W)) \\ &\le R_A^2(\rho_A^{2t}\mathrm{tr}(X_0) + \frac{\mathrm{tr}(W)}{1-\rho_A^2}).\end{aligned}$$

This gives the claimed bound. □

B.6.7 Proof of Lemma B.13

Proof of Lemma B.13. By Lemma B.12, we have $\|\Sigma_{t|t-1}\|_{\mathsf{op}} \le \mathrm{tr}(\Sigma_{t|t-1}) \le R_{e^-}$. By Proposition B.9, denoting $P_t = \Sigma_{t|t-1}$,

$$(I - \tilde{L}_t C)P_t = (\Sigma^+_{t|t-1} + C^\top V^{-1} C)^+ \Sigma^+_{t|t-1}.$$

Now, decompose $\Sigma_{t|t-1} = U\Lambda U^\top$ with its compact spectral decomposition, where $\Lambda \succ \mathbf{0}$ is the diagonal matrix of eigenvalues, and $U^\top U = I$ is a semi-orthogonal matrix. Then, we may write the above as

$$\begin{aligned}(I - \tilde{L}_t C)P_t &= (U\Lambda^{-1}U^\top + C^\top V^{-1} C)^+ U\Lambda^{-1}U^\top \\ &= U(\Lambda^{-1} + U^\top C^\top V^{-1} C U)^{-1} U^\top U \Lambda^{-1} U^\top \\ &= U(\Lambda^{-1} + U^\top C^\top V^{-1} C U)^{-1} \Lambda^{-1} U^\top \\ &= U\Lambda^{1/2}(I + \Lambda^{1/2} U^\top C^\top V^{-1} C U \Lambda^{1/2})^{-1} \Lambda^{-1/2} U^\top\end{aligned}$$

Therefore, by definition of Λ, and since $\Sigma_{t+1|t} = W + A\Sigma_{t|t}A^\top \succeq W$,

$$\begin{aligned}\|(I - \tilde{L}_t C)P_t\|_{\mathsf{op}} &\le \frac{\lambda_{\max}(\Sigma_{t|t-1})}{\lambda^+_{\min}(\Sigma_{t|t-1})} \|(I + \Lambda^{1/2} U^\top C^\top V^{-1} C U \Lambda^{1/2})^{-1}\|_{\mathsf{op}} \\ &\le \max\left\{\frac{\lambda_{\max}(X_0)}{\lambda^+_{\min}(X_0)}, \frac{R_{e^-}}{\lambda^+_{\min}(W)}\right\} := R_{I - \tilde{L}C}.\end{aligned}$$

Since $\Psi^{[t,t']} = \prod_{i=t}^{t'-1}(A - L_i C)$ by definition, the norm bound follows immediately from the assumption of dynamic detectability. Since $\Phi^{[t,t']} = (I - \tilde{L}_{t'} C)\Psi^{[t,t']}$ (defined in Lemma B.10),

$$\|\Phi^{[t,t']}\|_{\mathsf{op}} \le \|I - \tilde{L}_t C\|_{\mathsf{op}} \|\Psi^{[t,t']}\|_{\mathsf{op}} \le \frac{R_{e^-}}{\lambda_{\min}(W)} R_{A-LC} \rho_{A-LC}^{t'-t} := R_\Phi \rho_\Phi^{t'-t}$$

Finally, since $\widetilde{\Phi}^{[t,t']} = \Phi^{[t+1,t']} A$ by Lemma B.10

$$\|\widetilde{\Phi}^{[t,t']}\|_{\mathsf{op}} = \|\Phi^{[t,t']}\|_{\mathsf{op}} \|A\|_{\mathsf{op}} \le R_A \frac{R_{e^-}}{\lambda_{\min}(W)} R_{A-LC} \rho_{A-LC}^{t'-t-1} := R_{\widetilde{\Phi}} \rho_{\widetilde{\Phi}}^{t'-t}$$

□

C Nature's Observations: Deferred Proofs

Definition 3.1 (Nature's x's and y's [SSH20])**.** We let $\mathbf{x}_t^{\mathsf{nat}}$ (resp., $\mathbf{y}_t^{\mathsf{nat},(s)}$) be the (counterfactual) state (resp., observation from sensor s) of the system at time t, had we chosen $\mathbf{u}_s = 0$ for all $s \le t$. That is, $\mathbf{x}_0^{\mathsf{nat}} = \mathbf{x}_0$, $\mathbf{x}_{t+1}^{\mathsf{nat}} = A\mathbf{x}_t^{\mathsf{nat}} + \mathbf{w}_t$, and $\mathbf{y}_t^{\mathsf{nat},(s)} = C^{(s)}\mathbf{x}_t^{\mathsf{nat}} + \mathbf{v}_t^{(s)}$.

Lemma 3.2 (Lemma 3.1 from [SSH20]). *Consider any system evolving according to (LDS), with any (possibly adaptively selected) control inputs $\mathbf{u}_t$. Then, $\mathbf{x}_t^{\mathsf{nat}} = \mathbf{x}_t - \sum_{i=1}^{t} A^{i-1} B \mathbf{u}_{t-i}$ and $\mathbf{y}_t^{\mathsf{nat},(s)} = \mathbf{y}_t^{(s)} - \sum_{i=1}^{t} C^{(s)} A^{i-1} B \mathbf{u}_{t-i}$.*

Proof. We proceed via induction. The claims are trivially true for $t = 0$, since $\mathbf{x}_0^{\mathsf{nat}} = \mathbf{x}_0$ by definition, and thus also $\mathbf{y}_0^{\mathsf{nat},(s)} = \mathbf{y}_0^{(s)}$ for all s. If the claim is true for some $t \geq 0$, then by definition, and by the induction hypothesis:

$$\begin{aligned}
\mathbf{x}_{t+1}^{\mathsf{nat}} &= A\mathbf{x}_t^{\mathsf{nat}} + \mathbf{w}_t \\
&= A(\mathbf{x}_t - \sum_{i=1}^{t} A^{i-1} B \mathbf{u}_{t-i}) + \mathbf{w}_t \\
&= A\mathbf{x}_t + B\mathbf{u}_t + \mathbf{w}_t - \sum_{i=1}^{t+1} A^{i-1} B \mathbf{u}_{t+1-i} \\
&= \mathbf{x}_{t+1} - \sum_{i=1}^{t+1} A^{i-1} B \mathbf{u}_{t+1-i}.
\end{aligned}$$

This implies, by using the definition of $\mathbf{y}_t^{\mathsf{nat}}$:

$$\begin{aligned}
\mathbf{y}_{t+1}^{\mathsf{nat},(s)} &= C^{(s)} \mathbf{x}_{t+1}^{\mathsf{nat}} + \mathbf{v}_{t+1}^{(s)} \\
&= C^{(s)} (\mathbf{x}_{t+1} - \sum_{i=1}^{t+1} A^{i-1} B \mathbf{u}_{t+1-i}) + \mathbf{v}_{t+1}^{(s)} \\
&= C^{(s)} \mathbf{x}_{t+1} + \mathbf{v}_{t+1}^{(s)} - \sum_{i=1}^{t+1} C^{(s)} A^{i-1} B \mathbf{u}_{t+1-i} \\
&= \mathbf{y}_{t+1} - \sum_{i=1}^{t+1} C^{(s)} A^{i-1} B \mathbf{u}_{t+1-i},
\end{aligned}$$

as claimed. □

Lemma C.1. *The state $\mathbf{x}_t$ can be decomposed as $\mathbf{x}_t = \mathbf{x}_t^{\mathsf{nat}} + \Delta_t$, where*

$$\begin{bmatrix} \mathbf{x}_{t+1}^{\mathsf{nat}} \\ \Delta_{t+1} \end{bmatrix} = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & A \end{bmatrix} \begin{bmatrix} \mathbf{x}_t^{\mathsf{nat}} \\ \Delta_t \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ B \end{bmatrix} \mathbf{u}_t + \begin{bmatrix} \mathbf{w}_t \\ \mathbf{0} \end{bmatrix}.$$

Similarly, we can express the Kalman filters $\widehat{\mathbf{x}}_{t|t-1} = \widehat{\mathbf{x}}_{t|t-1}^{\mathsf{nat}} + \widehat{\Delta}_{t|t-1}$ and $\widehat{\mathbf{x}}_{t|t} = \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}} + \widehat{\Delta}_{t|t}$, where $\widehat{\Delta}_{t|t-1} = \Delta_t = \widehat{\Delta}_{t|t}$, and:

$$\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}} = (I - \tilde{L}_t C)\widehat{\mathbf{x}}_{t|t-1}^{\mathsf{nat}} + \tilde{L}_t \mathbf{y}_t^{\mathsf{nat}} \quad \textit{and} \quad \widehat{\mathbf{x}}_{t+1|t}^{\mathsf{nat}} = A\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}$$

In particular, we have that:

$$\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1} = \mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t-1}^{\mathsf{nat}} \quad \textit{and} \quad \mathbf{x}_t - \widehat{\mathbf{x}}_{t|t} = \mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}},$$

so these two differences do not depend on the executed controls $(\mathbf{u}_s)_{s \leq t}$.

Proof. The first claim follows immediately from the definition of the dynamics, and the fact that the dynamics are linear in $\mathbf{x}_t$, $\mathbf{u}_t$, and $\mathbf{w}_t$.

For the second claim, we begin by decomposing $\widehat{\mathbf{x}}_{t|t-1}$ as:

$$\begin{bmatrix} \widehat{\mathbf{x}}_{t+1|t}^{\mathsf{nat}} \\ \widehat{\Delta}_{t+1|t} \end{bmatrix} = \begin{bmatrix} A - L_t C & \mathbf{0} \\ \mathbf{0} & A - L_t C \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{x}}_{t|t-1}^{\mathsf{nat}} \\ \widehat{\Delta}_{t|t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ B \end{bmatrix} \mathbf{u}_t + \begin{bmatrix} L_t & \mathbf{0} \\ \mathbf{0} & L_t C \end{bmatrix} \begin{bmatrix} \mathbf{y}_t^{\mathsf{nat}} \\ \Delta_t \end{bmatrix}$$

Now, notice that, by definition, $\mathbf{x}_0^{\mathsf{nat}} = \mathbf{x}_0$ and $\widehat{\mathbf{x}}_{0|-1}^{\mathsf{nat}} = \mu_0 = \widehat{\mathbf{x}}_{0|-1}$, so $\Delta_0 = \mathbf{0} = \widehat{\Delta}_{0|-1}$. Assuming $\widehat{\Delta}_{t|t-1} = \Delta_t$ for some $t \geq 0$, the above recursion implies

$$\widehat{\Delta}_{t+1|t} = (A - L_t C)\widehat{\Delta}_{t|t-1} + B\mathbf{u}_t + L_t C \Delta_t = (A - L_t C)\Delta_t + B\mathbf{u}_t + L_t C \Delta_t = A\Delta_t + B\mathbf{u}_t = \Delta_{t+1}.$$

Therefore, $\widehat{\Delta}_{t|t-1} = \Delta_t$ for all t.

Finally, we decompose $\widehat{\mathbf{x}}_{t|t}$ and use the fact that $\widehat{\Delta}_{t|t-1} = \Delta_t$ to obtain:

$$\begin{aligned}\begin{bmatrix}\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} \\ \widehat{\Delta}_{t|t}\end{bmatrix} &= \begin{bmatrix} I - \tilde{L}_t C & \mathbf{0} \\ \mathbf{0} & I - \tilde{L}_t C \end{bmatrix} \begin{bmatrix}\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t-1} \\ \widehat{\Delta}_{t|t-1}\end{bmatrix} + \begin{bmatrix} \tilde{L}_t & \mathbf{0} \\ \mathbf{0} & \tilde{L}_t C \end{bmatrix} \begin{bmatrix}\mathbf{y}^{\mathsf{nat}}_t \\ \Delta_t\end{bmatrix} \\ &= \begin{bmatrix} I - \tilde{L}_t C & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix}\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t-1} \\ \Delta_t\end{bmatrix} + \begin{bmatrix} \tilde{L}_t \\ \mathbf{0} \end{bmatrix} \mathbf{y}^{\mathsf{nat}}_t,\end{aligned}$$

as claimed. □

Definition C.2 (Counterfactuals)**.** Given any $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$, we denote $\mathbf{x}_t(\mathbf{M})$ (resp., $\mathbf{y}^{(s)}_t(\mathbf{M})$) as the resulting state from executing controls $\mathbf{u}_j(\mathbf{M})$ for all times $j \le t$, i.e., $\mathbf{x}_0(\mathbf{M}) = \mathbf{x}_0$ and:

$$\begin{aligned}\mathbf{x}_{t+1}(\mathbf{M}) &= A\mathbf{x}_t(\mathbf{M}) + B\mathbf{u}_t(\mathbf{M}) + \mathbf{w}_t \\ \mathbf{y}^{(s)}_t(\mathbf{M}) &= C^{(s)}\mathbf{x}_t(\mathbf{M}) + \mathbf{v}_t \; \forall s \in \mathcal{S}.\end{aligned}$$

Moreover (following Lemma C.1), the counterfactual Kalman filter resulting from the observations $\mathbf{y}^{(s)}_t(\mathbf{M})$ is

$$\widehat{\mathbf{x}}_{t|t}(\mathbf{M}) := \widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} + \Delta_t(\mathbf{M}) = \widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} + \sum_{i=1}^{t} A^{i-1} B \mathbf{u}_t(\mathbf{M}).$$

D Expressivity of Disturbance Response Controllers

Claim D.1 (Expressivity of policy class)**.** For any single-sensor LDC policy $\pi^{(s)}_{\mathsf{ldc}} \in \Pi_{\mathsf{sLDC}}(R, \rho)$ (where $R > 0$, $\rho \in (0, 1)$), there is a controller $\mathbf{M} \in \mathcal{M}(R, \rho, n)$ which produces the same control sequence for all $t \in [n]$. In particular,

$$\mathbf{u}^{\mathsf{ldc}}_t = \sum_{s \in \mathcal{S}} \sum_{j=0}^{t} M^{(s)}_j \mathbf{y}^{\mathsf{nat},(s)}_{t-j},$$

where $M^{(s')}_j = \mathbf{0}$ for all $s' \neq s$, and $M^{(s)}_j = G^{(s),[j]}_{\mathsf{ldc,cl}}$ for all $j \ge 0$, where $G^{(s)}_{\mathsf{ldc,cl}}$ is the closed-loop Markov operator associated with $\pi^{(s)}_{\mathsf{ldc}}$ (see Definition 2.3).

Proof. The proof follows from Claim 5.6 in [SSH20]. □

Claim D.2 (Closure under mixtures)**.** Moreover, the policy class $\mathcal{M}(R, \rho, n)$ is closed under the following type of mixture: for any $\mathbf{M}(1), \ldots, \mathbf{M}(K) \in \mathcal{M}(R, \rho, n)$ and coefficients $(\alpha_j(k))_{j \ge 0, k \in [K]}$ satisfying $\alpha_j(k) \ge 0$ and $\sum_{k \in [K]} \alpha_j(k) \le 1$ for all $j \ge 0$, the model $\mathbf{M} := (\sum_{k \in [K]} \alpha_j(k) M^{(s)}_j(k))_{s \in \mathcal{S}, j \ge 0} \in \mathcal{M}(R, \rho, n)$.

Proof. The claim follows from the observation that, since $\mathbf{M}(k) \in \mathcal{M}(R, \rho)$ for all $k \in [K]$, and by definition of the mixture weights $(\alpha_j)_{k \in [K]}$ for each $j \ge 0$,

$$\sum_{s \in \mathcal{S}} \| \sum_{k \in [K]} \alpha_j(k) M^{(s)}_j(k) \|_{\mathsf{op}} \le \sum_{k \in [K]} \alpha_j(k) \sum_{s \in \mathcal{S}} \| M^{(s)}_j(k) \|_{\mathsf{op}} \le R\rho^j \sum_{k \in [K]} \alpha_j(k) \le R\rho^j.$$

□

E Norm Facts

E.1 Matrix Group Norms

Here, we recall some facts about matrix group norms from [KST+09], specialized to our problem setting.

Lemma E.1. *Let $M = (M_1, \dots, M_n) \in \mathbb{R}^{m\times n}$. For $r, p \geq 2$, the (r,p)-group norm of M is*

$$\|(\|M_1\|_r, \dots, \|M_n\|_r)\|_p.$$

Then the function

$$F(M) := \frac{1}{2}\|M\|_{2,q}^2 \quad \textit{where} \quad q = \begin{cases} 2 & \textit{if } n < e \\ \frac{2\ln(n)}{2\ln(n)-1} & \textit{o.w.} \end{cases}$$

is $\frac{q-1}{eq} = \frac{1}{2e\max\{1,\ln(n)\}}$-strongly convex w.r.t. $\|\cdot\|_{2,1}$. Moreover, the dual norm of $\|\cdot\|_{2,1}$ satisfies:

$$(\|\cdot\|_{2,1})_\star = \|\cdot\|_{2,\infty}.$$

Proof. By Corollary 19 in [KST+09], the Fenchel conjugate of $F(M)$, $F^\star(M) = \frac{1}{2}\|M\|_{2,p}$ is p-smooth with respect to $\|\cdot\|_{2,p}$, for $\frac{1}{p} = 1 - \frac{1}{q}$ (as long as $p \geq 2$). By definition of smoothness, this implies that

$$F^\star(M + M') \leq F^\star(M) + \langle \nabla F^\star(M), M' \rangle + \frac{p}{2}\|M'\|_{2,p}^2.$$

Thus, since

$$\|v\|_p \leq \|v\|_\infty n^{1/p} = \|v\|_\infty e^{\ln(n)/p},$$

if $p = \max\{2, 2\ln(n)\}$, we have that $F^\star$ is ep-smoooth w.r.t. $\|\cdot\|_{2,\infty}$. By Lemma 17 in [KST+09], the dual norm of $\|\cdot\|_{2,\infty}$ is $(\|\cdot\|_{2,\infty})_\star = \|\cdot\|_{2,1}$. Thus, by Theorem 6 in [KST+09], F is $\frac{1}{ep} = \frac{q-1}{eq}$-strongly convex w.r.t $\|\cdot\|_{2,1}$. Plugging in our choice of p establishes the claim. □

Corollary E.2. *For a vector $v \in \mathbb{R}^m$ and positive definite matrix $W \succ \mathbf{0} \in \mathbb{R}^{m\times m}$, define the W-reweighted 2-norm as*

$$\|v\|_W = \sqrt{v^\top W v} = \|W^{1/2}v\|_2,$$

where $W^{1/2}$ denotes the unique positive definite square root of W satisfying $W^{1/2}W^{1/2} = W$. Denote, for $M = (M_1, \dots, M_n) \in \mathbb{R}^{m\times n}$

$$\|M\|_{W,q} = \|(\|M_1\|_W, \dots, \|M_n\|_W)\|_q.$$

The, we have that the function

$$G(M) = \frac{1}{2}\|M\|_{W,q}^2 \quad \textit{where} \quad q = \begin{cases} 2 & \textit{if } n < e \\ \frac{2\ln(n)}{2\ln(n)-1} & \textit{o.w.} \end{cases}$$

is $\frac{1}{2e\max\{1,\ln(n)\}}$-strongly convex w.r.t. $\|\cdot\|_{W,1}$. Moreover, the dual norm of $\|\cdot\|_{W,1}$ is given by

$$(\|\cdot\|_{W,1})_\star = \|\cdot\|_{W^{-1},\infty}.$$

Proof. Since $F(M) = \frac{1}{2}\|M\|_{2,q}^2$ is $\beta = \frac{1}{2e\max\{1,\ln(e)\}}$-strongly convex w.r.t. $\|\cdot\|_{2,q}$, by definition, for all $\alpha \in (0,1)$, and for all M, M'

$$\frac{1}{2}\|M - M'\|_{2,q}^2 \leq \frac{\alpha}{2}\|M\|_{2,q}^2 + \frac{(1-\alpha)}{2}\|M'\|_{2,q}^2 - \frac{\beta}{2}\alpha(1-\alpha)\|M - M'\|_{2,1}^2.$$

Let $W \in \mathbb{R}^{m\times m}$ be a positive definite matrix, and let $\tilde{M} = (W^{-1/2}M_1, \dots, W^{-1/2}M_n)$. Then, recalling that $\|v\|_W = \|W^{1/2}v\|_2$, we have that, for any $p \geq 1$

$$\begin{aligned}
\|M\|_{2,p} &= \|(\|M_1\|_2, \dots, \|M_n\|_2)\|_p \\
&= \|(\|W^{1/2}W^{-1/2}M_1\|_2, \dots, \|W^{1/2}W^{-1/2}M_n\|_2)\|_p \\
&= \|(\|W^{1/2}\tilde{M}_1\|_2, \dots, \|W^{1/2}\tilde{M}_n\|_2)\|_p \\
&= \|(\|\tilde{M}_1\|_W, \dots, \|\tilde{M}_n\|_W)\|_p \\
&= \|\tilde{M}\|_{W,p}.
\end{aligned}$$

Note also that any matrix $\tilde{M} \in \mathbb{R}^{m \times m}$ can be written as above, by invertibility of $W^{1/2}$. Also, notice that $M_i - M'_i = W^{1/2}(\tilde{M}_i - \tilde{M}'_i)$ for all $i' \in [n]$. Thus, we conclude that, for any $\tilde{M}, \tilde{M}'$ and any $\alpha \in (0,1)$,

$$\frac{1}{2}\|\tilde{M} - \tilde{M}'\|_{W,q}^2 \leq \frac{\alpha}{2}\|\tilde{W}\|_{W,q}^2 + \frac{(1-\alpha)}{2}\|\tilde{M}'\|_{W,q}^2 - \frac{\beta}{2}\alpha(1-\alpha)\|\tilde{M} - \tilde{M}'\|_{W,1}^2.$$

Thus, $G(\tilde{M})$ is β-strongly convex w.r.t. $\|\cdot\|_{W,1}$, as claimed.

Finally, the claim that $(\|\cdot\|_{W,1})_\star = \|\cdot\|_{W^{-1},\infty}$ follows from [KST+09, Lemma 17] (since $\|\cdot\|_1$ is an absolutely symmetric norm), together with the observation that $(\|\cdot\|_W)_\star = \|\cdot\|_{W^{-1}}$ since

$$(\|z\|_W)_\star = \sup_{\|v\|_W = 1} \langle z, v\rangle = \sup_{\|W^{1/2}v\|_2 = 1} \langle W^{-1/2}z, W^{1/2}v\rangle = \sup_{\|\tilde{v}\|_2 = 1} \langle W^{-1/2}z, \tilde{v}\rangle = \|W^{-1/2}z\|_2 = \|z\|_{W^{-1}},$$

since the 2-norm is self-dual. □

We apply the above bounds to our setting by a minor extension of the above notation.

Definition E.3. Let $\mathbf{M} = (M_j^{(s)})_{j\in[H],s\in\mathcal{S}}$, where $M_j^{(s)} \in \mathbb{R}^{d_{\mathbf{u}} \times d_{y^s}}$. Denote $M^{(s)} = (M_j^{(s)})_{j=0}^H$ Then, for $p \geq 1$, the (r,p)-group norm of $\mathbf{M}$ is

$$\|(\|M^{(1)}\|_r, \ldots, \|M^{(|\mathcal{S}|)}\|_r)\|_p,$$

where $\|M^{(s)}\|_F$ denotes the Frobenius norm of $M^{(s)}$. For any $\rho > 0$ and $q \geq 1$, define the ρ-reweighted $(2,q)$-group norm as

$$\|(\|M^{(1)}\|_\rho, \ldots, \|M^{(|\mathcal{S}|)}\|_\rho)\|_q, \quad \text{where} \quad \|M^{(s)}\|_\rho^2 = \sum_{j=0}^H \rho^{2j}\|M_j^{(s)}\|_F^2.$$

Lemma 3.6. *For $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$, $\rho > 0$ and $q \in (1,2]$, the (ρ,q)-group norm of $\mathbf{M}$ is $\|\mathbf{M}\|_{\rho,q} = \|(\|M^{(1)}\|_\rho, \ldots, \|M^{(|\mathcal{S}|)}\|_\rho)\|_q$ where $\|M^{(s)}\|_\rho^2 = \sum_{j=0}^H \rho^{2j}\|M_j^{(s)}\|_F^2$. Moreover, for $p = 2\max\{1, \ln(|\mathcal{S}|)\}$ and $\frac{1}{q} = 1 - \frac{1}{p}$, then $G_\rho(\mathbf{M}) = \frac{ep}{2}\|\mathbf{M}\|_{\rho,q}^2$ is 1-strongly convex w.r.t. $\|\cdot\|_{\rho,1}$, and the dual norm of $\|\cdot\|_{\rho,1}$ is $\|\cdot\|_{\rho^{-1},\infty}$.*

Proof. Notice that the claim follows directly from Corollary E.2. To see this, we can view each $M^{(s)}$ as a vector of length $H \times d_{\mathbf{u}} \times d_{y^s}$. Then, take W to be the diagonal matrix where rows corresponding to $M_j^{(s)}$ have weight ρ^j for each $j \in [0,H]$. □

E.2 Orlicz Norms, sub-Gaussian and sub-Exponential Random Variables

Here we review some basic properties of Orlicz norms. For details, see, e.g., [Ver18].

Definition E.4. For $q \geq 1$, the ψ_q-*Orlicz norm* of a mean-zero random variable X is defined as:

$$\|X\|_{\psi_q} := \inf\{\lambda > 0 : \mathbb{E}[\psi_q(|X|/\lambda)] \leq 1\} \quad \text{where} \quad \psi_q(x) = \exp(x^q) - 1$$

One may verify that this is indeed a norm when $q \geq 1$, and, in particular, satisfies the triangle inequality:

$$\|X + Y\|_{\psi_q} \leq \|X\|_{\psi_q} + \|Y\|_{\psi_q}$$

Definition E.5 (sub-Gaussian random variable)**.** We say a random variable X is sub-Gaussian if $\|X\|_{\psi_2} < \infty$.

Definition E.6 (sub-Gaussian random vector)**.** We say a random vector $X \in \mathbb{R}^d$ is sub-Gaussian if

$$\sup_{v \in \mathbb{R}^d : \|v\|_2 = 1} \|\langle X, v\rangle\|_{\psi_2} < \infty.$$

Definition E.7 (sub-Exponential random variable)**.** We say a random variable X is sub-Exponential if $\|X\|_{\psi_1} < \infty$.

Lemma E.8. *Suppose that X is sub-Exponential. Then, there is an absolute constant $c > 0$ such that*

$$\mathbb{E}[|X|^p]^{1/p} \leq cp\|X\|_{\psi_1} \ \forall p \geq 1.$$

Proof. This is the (equivalent) moment characterization of sub-Exponential random variables: see [Ver18, Lemma 2.8.1 (ii)] for a proof. □

Lemma E.9 (Proposition 2.7.1, [Ver18])**.** *Let X_i, $i \in [n]$ be independent, mean-zero sub-Gaussian random variables. Then, there exists an absolute constant $C > 0$ such that*

$$\|\sum_{i=1}^n X_i\|_{\psi_2} \leq R_{\psi_2}^2 \sum_{i=1}^n \|X_i\|_{\psi_2}.$$

Lemma E.10. *Suppose $X_1, \ldots, X_n$ are independent, mean-zero, and $\|X_i\|_{\psi_1} < \infty$ for all i. Then, there is an absolute constant $R_{\psi_1} > 0$ such that*

$$\|\sum_{i=1}^n X_i\|_{\psi_1}^2 \leq R_{\psi_1}^2 \sum_{i=1}^n \|X_i\|_{\psi_1}^2$$

Proof. This result is an immediate consequence of [Ver18, Proposition 2.8.1]; we give a proof here for completeness.

By [Ver18, Proposition 2.8.1], if X_i is mean-zero and sub-Exponential, then there exists an absolute constant $c > 0$ such that

$$\mathbb{E}\left[\exp(\lambda X_i)\right] \leq \exp(c^2 \|X_i\|_{\psi_1}^2 \lambda^2) \quad \forall |\lambda| \leq \frac{1}{c\|X_i\|_{\psi_1}}.$$

Therefore, by the fact that $(X_i)_{i\in[n]}$ are independent, we have that

$$\begin{aligned}\mathbb{E}\left[\exp(\lambda \sum_{i=1}^n X_i)\right] &= \prod_{i=1}^n \mathbb{E}\left[\exp(\lambda X_i)\right] \\ &\leq \exp(c^2 \sum_{i=1}^n \|X_i\|_{\psi_1}^2 \lambda^2) \quad \forall |\lambda| \leq \frac{1}{\sqrt{c^2 \sum_{i=1}^n \|X_i\|_{\psi_1}}} \leq \min_{i\in[n]} \frac{1}{c\|X_i\|_{\psi_1}}.\end{aligned}$$

Again by [Ver18, Proposition 2.8.1], the above inequality implies that there exists an absolute constant $R_{\psi_1} > 0$ such that

$$\mathbb{E}\left[\exp\left(\frac{|\sum_{i=1}^n X_i|}{R_{\psi_1}\sqrt{\sum_{i=1}^n \|X_i\|_{\psi_1}^2}}\right)\right] \leq 2,$$

which, by Definition E.7, implies

$$\|\sum_{i=1}^n X_i\|_{\psi_1}^2 \leq R_{\psi_1}^2 \sum_{i=1}^n \|X_i\|_{\psi_1}^2,$$

as claimed. □

Lemma E.11 (Lemma 2.8.6, [Ver18])**.** *Suppose that X and Y are sub-Gaussian, i.e., mean-zero (not necessarily independent) random variables with $\|X\|_{\psi_2}, \|Y\|_{\psi_2} < \infty$. Then*

$$\|XY\|_{\psi_1} \leq \|X\|_{\psi_2} \|Y\|_{\psi_2}.$$

In particular, XY is sub-Exponential.

Lemma E.12 (Lemma 2.2.2, [vW96])**.** *Let $q \geq 1$, and $Y_1, \ldots, Y_m$ be random variables (not necessarily independent) such that $\|Y_j\|_{\psi_q} < \infty$ for all $j \in [m]$. Then, for some constant K_q depending only on q,*

$$\|\max_{j\in[m]} Y_j\|_{\psi_q} \leq K_q \ln(m+1)^{1/q} \max_{j\in[m]} \|Y_j\|_{\psi_q}.$$

Lemma E.13. *Suppose $X \in \mathbb{R}^d$ is a sub-Gaussian random vector. Then, for any $v \in \mathbb{R}^d$,*

$$\|\langle X, v\rangle\|_{\psi_2} \leq \|v\|_2 \|X\|_{\psi_2}$$

Proof. The proof follows easily from [Ver18, Section 3.4]. We give a proof here for completeness.

Recall that, since $\|\cdot\|_{\psi_2}$ is a norm, $\|cY\|_{\psi_2} = c\|Y\|_{\psi_2}$ for any $c \geq 0$. Therefore, since $\|v/\|v\|_2\|_2 = 1$,

$$\frac{1}{\|v\|_2}\|\langle X, v\rangle\|_{\psi_2} = \|\langle X, v/\|v\|_2\rangle\|_{\psi_2} \leq \sup_{x\in\mathbb{R}^d : \|x\|_2=1} \|\langle X, x\rangle\|_{\psi_2} = \|X\|_{\psi_2},$$

which establishes the claimed inequality. □

E.3 Norms on nets

Lemma E.14 (Lemma 4.4.1 and Corollary 4.2.11 [Ver18]). *Let $M \in \mathbb{R}^{m\times n}$, and $\varepsilon \in [0,1)$. Then, there exists a set $\mathcal{N} \subset \mathbb{R}^n$ such that, for all $x \in \mathcal{N}$, $\|x\|_2 = 1$, $|\mathcal{N}| \le (1 + 2/\varepsilon)^n$, and*

$$\|M\|_{\mathsf{op}} \le \frac{1}{1-\varepsilon} \sup_{x\in\mathcal{N}} \|Mx\|_2.$$

Lemma E.15 (Lemma 4.4.2 and Corollary 4.2.11 [Ver18]). *Let $M \in \mathbb{R}^{m\times n}$, and $\varepsilon \in [0, 1/2)$. Then, there exists sets $\mathcal{N} \subset \mathbb{R}^n$, $\mathcal{M} \subset \mathbb{R}^m$ such that, for all $x \in \mathcal{N}$, $\|x\|_2 = 1$, $|\mathcal{N}| \le (1 + 2/\varepsilon)^n$, $\|y\|_2 = 1$ for all $y \in \mathcal{M}$, $|\mathcal{M}| = (1 + 2/\varepsilon)^m$, and*

$$\|M\|_{\mathsf{op}} \le \frac{1}{1-2\varepsilon} \max_{x\in\mathcal{N}, y\in\mathcal{M}} \langle y, Ax\rangle .$$

F Convex Optimization Background

Here, we summarize some useful standard results from online convex optimization. For a textbook treatment of these results, see, e.g., [Ora19].

Theorem F.1 (Regret of FTRL for OCO, Corollary 7.7, [Ora19]). *Consider the (linearlized) FTRL algorithm which, given a constraint set $\mathcal{V}$, a sequence of regularizers $\Psi_t(x) = \Psi(x)/\eta_{t-1}$, where $\Psi(x)$ is μ-strongly convex with respect to norm $\|\cdot\|$, positive step-sizes $\eta_{t+1} \le \eta_t$ for all t, and a sequence of convex, closed subdifferentiable functions ℓ_t, maintains iterates x_t according to:*

$$x_t \in \arg\min_{x\in\mathcal{V}} \Psi_t(x) + \sum_{i=1}^{t-1} \langle g_i, x\rangle ,$$

where $g_t \in \partial\ell_t(x_t)$. Then, for any comparator $u \in \mathcal{V}$, the regret is bounded as

$$\sum_{t=1}^{n} \ell_t(x_t) - \ell_t(u) \le \frac{\Psi(u) - \min_{x\in\mathcal{V}} \Psi(x)}{\eta_{n-1}} + \frac{1}{2\mu} \sum_{t=1}^{n} \eta_{t-1} \|g_t\|_*^2 \quad \forall u \in \mathcal{V}.$$

Lemma F.2. *Consider the FTRL algorithm with λ-strongly convex regularizer Ψ w.r.t. $\|\cdot\|$:*

$$x_t = \arg\min_{x\in\mathcal{V}} \Psi(x) + \sum_{i=0}^{t-1} \ell_i(x) = \arg\min_{x\in\mathcal{V}} F_t(x).$$

Suppose that ℓ_i convex and differentiable for all i. Then

$$\|x_{t+1} - x_t\| \le \frac{\|\nabla\ell_t(x_t)\|_*}{\lambda}$$

Proof. This is a well-known consequence of strongly convexity and Lipschitzness. We give a proof here for completeness.

Since we consider a fixed regularizer, denote $F_t(x) = \Psi(x) + \sum_{i=1}^{t-1} \ell_i(\mathbf{x})$ from the previous theorem. Since Ψ is λ-strongly convex and ℓ_i is convex for all i, it follows that F_t is λ-strongly convex. Therefore,

$$F_t(x_{t+1}) - F_t(x_t) \ge \langle \nabla F_t(x_t), x_{t+1} - x_t\rangle + \frac{\lambda}{2}\|x_{t+1} - x_t\|^2 \ge \frac{\lambda}{2}\|x_{t+1} - x_t\|^2,$$

where the second inequality follows since $\langle \nabla F_t(x_t), y - x_t\rangle \ge 0$ for all $y \in \mathcal{V}$ by minimality of x_t (see, e.g., Theorem 2.8 in [Ora19]). Similarly,

$$F_{t+1}(x_t) - F_{t+1}(x_{t+1}) \ge \frac{\lambda}{2}\|x_{t+1} - x_t\|^2.$$

Summing the above two expressions, notice that:

$$\begin{aligned}
\lambda\|x_{t+1} - x_t\|^2 &\le (F_t(x_{t+1}) - F_t(x_t)) + (F_{t+1}(x_t) - F_{t+1}(x_{t+1})) \\
&= (\Psi(x_{t+1}) - \Psi(x_t) + \sum_{i=1}^{t-1} \ell_i(x_{t+1}) - \ell_i(x_t)) \\
&\quad + (\Psi(x_t) - \Psi(x_{t+1}) + \sum_{i=1}^{t} \ell_i(x_t) - \ell_i(x_{t+1})) \\
&= \ell_t(x_t) - \ell_t(x_{t+1}).
\end{aligned}$$

Combining the above bound with the fact that, by definition of convexity ℓ_t and Hölder's inequality,

$$\begin{aligned}\ell_t(x_t) - \ell_t(x_{t+1}) &\leq \langle \nabla \ell_t(x_t), x_t - x_{t+1} \rangle \\ &\leq \|\nabla \ell_t(x_t)\|_* \|x_{t+1} - x_t\|,\end{aligned}$$

Thus, dividing both sides by $\|x_{t+1} - x_t\|$, the claimed bound follows. □

G State Norm Bounds

G.1 State and observation norm bounds

Here, we show moment bounds for the states, observations, and controls of the algorithm in the stochastic setting. Note that these bounds are obtained under slightly weaker assumptions than Assumption 2.

Lemma G.1 (State, observation, and control norm bounds)**.** *Suppose that the initial state* $\mathbf{x}_0$ *and the state and observation noises* $\mathbf{w}_t$, $\mathbf{v}_t^{(s)}$ *satisfy*

$$\mathbb{E}\left[\|\mathbf{e}_0^-\|^8\right] \leq R_{x_0}^8, \quad \mathbb{E}\left[\|\mathbf{w}_t\|^8\right] \leq R_w^8, \quad \mathbb{E}\left[\|\mathbf{v}_t^{(s)}\|^8\right] \leq R_v^8,$$

$$\mathbb{E}\left[\max_{s\in\mathcal{S}} \|\mathbf{v}_t^{(s)}\|^8\right] \leq R_{v,\sup}^8, \quad \mathbb{E}\left[\max_{s\in\mathcal{S}} \|V_s^{-1/2}\mathbf{v}_t^{(s)}\|^8\right] \leq R_{\bar{v}^s}^8, \quad \forall t \geq 0, s \in \mathcal{S}.$$

Then, we have that, for any $\mathbf{M}$ *such that* $\sum_{s\in\mathcal{S}} \|M_j^{(s)}\|_{\mathsf{op}} \leq R_M \rho_M^j$ *for all* $j \geq 0$*, and any time* t

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{x}_t^{\mathsf{nat}}\|_2^8\right] &\leq R_A^8(1 + \frac{1}{1-\rho_A})^7(R_{e_0^-}^8 + \|\mu_0\|_2^8 + \sum_{i=1}^t \frac{R_w^8}{1-\rho_A}) := R_{x,\mathsf{nat}}^8 \\
\mathbb{E}\left[\|\mathbf{y}_t^{\mathsf{nat},(s)}\|_2^8\right] &\leq 128(R_{C^{(s)}}^8 R_x^8 + R_v^8) := R_{y^{(s)},\mathsf{nat}}^8 \\
\mathbb{E}\left[\max_{s\in\mathcal{S}} \|\mathbf{y}_t^{\mathsf{nat},(s)}\|_2^8\right] &\leq 128(R_{C^{(s)},\max}^8 R_x^8 + R_{v,\sup}^8) := R_{y^{(s)},\mathsf{nat},\sup}^8 \\
\mathbb{E}\left[\|\mathbf{u}_t(\mathbf{M})\|_2^8\right] &\leq \frac{R_M^8 R_{y^{(s)},\mathsf{nat},\sup}^8}{(1-\rho_M)^8} := R_u^8 \\
\mathbb{E}\left[\|\mathbf{x}_t(\mathbf{M})\|_2^8\right] &\leq (1 + \frac{R_A R_B}{1-\rho_A})^7(R_{x,\mathsf{nat}}^8 + \frac{R_u^8}{1-\rho_A}) := R_x \\
\mathbb{E}\left[\|\mathbf{x}_t\|_2^8\right] &\leq R_x^8 \\
\mathbb{E}\left[\max_{s\in\mathcal{S}} \|\mathbf{y}_t^{(s)}(\mathbf{M})\|_2^8\right] &\leq 128(R_{C^{(s)},\max}^8 R_x^8 + R_{v,\sup}) := R_{y^{(s)},\max}^8 \\
\mathbb{E}\left[\max_{s\in\mathcal{S}} \|\mathbf{y}_t^{(s)}\|_2^8\right] &\leq R_{y^{(s)},\max}^8.
\end{aligned}$$

Proof. From Definition 3.1, we know that

$$\mathbf{x}_t^{\mathsf{nat}} = A\mathbf{x}_{t-1}^{\mathsf{nat}} + \mathbf{w}_{t-1} = A^t\mathbf{x}_0 + \sum_{i=1}^t A^{i-1}\mathbf{w}_{t-i},$$

therefore, using our assumption on the norm decay of A Assumption 1, and applying Hölder's inequality,

$$\begin{aligned}
\|\mathbf{x}_t^{\mathsf{nat}}\|_2 &\leq \|A\|_{\mathsf{op}}\|x_0\|_2 + \sum_{i=1}^t \|A^{i-1}\|_{\mathsf{op}}\|\mathbf{w}_{t-i}\|_2 \\
&\leq R_A(\rho_A^t(\|\mathbf{e}_0^-\|_2 + \|\mu_0\|_2) + \sum_{i=1}^t \rho_A^{i-1}\|\mathbf{w}_{t-1}\|_2) \\
&\leq R_A(2 + \sum_{i=1}^t \rho_A^{(i-1)})^{7/8}(\|\mathbf{e}_0^-\|_2^8 + \|\mu_0\|_2^8 + \sum_{i=1}^t \rho_A^{(i-1)}\|\mathbf{w}_{t-1}\|_2^8)^{1/8}.
\end{aligned}$$

Therefore,

$$\mathbb{E}\left[\|\mathbf{x}_t^{\mathsf{nat}}\|_2^8\right] \le R_A^8(1+\frac{1}{1-\rho_A})^7\mathbb{E}\left[\|\mathbf{e}_0^-\|_2^8+\|\mu_0\|_2^8+\sum_{i=1}^{t}\rho_A^{(i-1)}\|\mathbf{w}_{t-1}\|_2^8\right]$$
$$\le R_A^8(1+\frac{1}{1-\rho_A})^7(R_{e_0^-}^8+\|\mu_0\|_2^8+\sum_{i=1}^{t}\frac{R_w^8}{1-\rho_A}) := R_{x,\mathsf{nat}}^8$$

Now, recalling the definition of $\mathbf{y}_t^{\mathsf{nat},(s)} = C^{(s)}\mathbf{x}_t^{\mathsf{nat}} + \mathbf{v}_t^{(s)}$ from Definition 3.1, by the above bound, we obtain:

$$\mathbb{E}\left[\|\mathbf{y}_t^{\mathsf{nat},(s)}\|_2^8\right] \le 128(\|C^{(s)}\|_{\mathsf{op}}^8\mathbb{E}\left[\|\mathbf{x}_t^{\mathsf{nat}}\|_2^8\right]+\mathbb{E}\left[\|\mathbf{v}_t^{(s)}\|_2^8\right]) \le 128(R_{C^{(s)}}^8R_x^8+R_v^8) := R_{y^{(s)},\mathsf{nat}}^8$$

Similarly,

$$\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\mathbf{y}_t^{\mathsf{nat},(s)}\|_2^8\right] \le 128\mathbb{E}\left[\max_{s\in\mathcal{S}}(\|C^{(s)}\|_{\mathsf{op}}^8\|\mathbf{x}_t^{\mathsf{nat}}\|_2^8+\|\mathbf{v}_t^{(s)}\|_2^8)\right]$$
$$\le 128(R_{C^{(s)},\max}^8R_x^8+R_{v,\sup}^8) := R_{y^{(s)},\mathsf{nat},\sup}^8$$

Hence, for any $\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)$, by definition of $\mathbf{u}_t(\mathbf{M}) := \sum_{s\in\mathcal{S}}\sum_{j=0}^{H}M_j^{(s)}\mathbf{y}_{t-j}^{\mathsf{nat},(s)}$,

$$\|\mathbf{u}_t(\mathbf{M})\|_2 \le \sum_{j=0}^{H}\sum_{s\in\mathcal{S}}\|M_j^{(s)}\|_{\mathsf{op}}\|\mathbf{y}_{t-j}^{\mathsf{nat},(s)}\|_2 \le \sum_{j=0}^{H}(\sum_{s\in\mathcal{S}}\|M_j^{(s)}\|_{\mathsf{op}})\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-j}^{\mathsf{nat},(s)}\|_2.$$

By definition of $\mathcal{M}(R_M,\rho_M,H)$, we have deterministically that $\sum_{s\in\mathcal{S}}\|M_j^{(s)}\|_{\mathsf{op}} \le R_M\rho_M^j$ for all j. Therefore, applying the same type of arguments as before,

$$\mathbb{E}\left[\|\mathbf{u}_t(\mathbf{M})\|_2^8\right] \le \frac{R_M^8}{(1-\rho_M)^7}\sum_{j=0}^{H}\rho_M^j\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-j}^{\mathsf{nat},(s)}\|_2^8\right] \le \frac{R_M^8R_{y^{(s)},\mathsf{nat},\sup}^8}{(1-\rho_M)^8} := R_u^8.$$

Next, using Lemma C.1,

$$\mathbf{x}_t(\mathbf{M}) = \mathbf{x}_t^{\mathsf{nat}} + \Delta_t(\mathbf{M}) = \mathbf{x}_t^{\mathsf{nat}} + \sum_{i=1}^{t}A^{i-1}B\mathbf{u}_{t-i}(\mathbf{M})$$

Thus, proceeding in a similar manner as before:

$$\mathbb{E}\left[\|\mathbf{x}_t(\mathbf{M})\|_2^8\right] \le (1+\frac{R_AR_B}{1-\rho_A})^7\mathbb{E}\left[\|\mathbf{x}_t^{\mathsf{nat}}\|_2^8+\sum_{i=1}^{t}\rho_A^{i-1}\|\mathbf{u}_{t-i}(\mathbf{M})\|_2^8\right]$$
$$\le (1+\frac{R_AR_B}{1-\rho_A})^7(R_{x,\mathsf{nat}}^8+\frac{R_u^8}{1-\rho_A}) := R_x$$

Similarly, since

$$\mathbf{x}_t = \mathbf{x}_t^{\mathsf{nat}} + \Delta_t = \mathbf{x}_t^{\mathsf{nat}} + \sum_{i=1}^{t}A^{i-1}B\mathbf{u}_{t-i}(\mathbf{M}(t)),$$

where $\mathbf{M}(t)\in\mathcal{M}(R_M,\rho_M,H)$ for all t by definition of Algorithm 1, it follows that

$$\mathbb{E}\left[\|\mathbf{x}_t\|_2^8\right] \le R_x^8.$$

Since $\mathbf{y}_t^{(s)}(\mathbf{M}) = C^{(s)}\mathbf{x}_t(\mathbf{M}) + \mathbf{v}_t^{(s)}$ and $\mathbf{y}_t^{(s)} = C^{(s)}\mathbf{x}_t + \mathbf{v}_t^{(s)}$, we have that, for any $\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)$,

$$\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\mathbf{y}_t^{(s)}(\mathbf{M})\|_2^8\right] \le 128(R_{C^{(s)},\max}^8R_x^8+R_{v,\sup}) := R_{y^{(s)},\max}^8,$$

and similarly, $\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\mathbf{y}_t^{(s)}\|_2^8\right] \le R_{y^{(s)},\max}^8$. □

G.2 Kalman Noise Addition Bound

Lemma G.2 (Bounds on Kalman noise additions)**.** *Suppose that $W, X_0, V \succ \mathbf{0}$ for all t. Further, suppose that the stacked system*

$$\|\tilde{L}_t \mathbf{v}_t\|_2 \leq \sqrt{\mathsf{rank}(C)}(\max_{s\in\mathcal{S}} \|V_s^{-1/2}\mathbf{v}_t^{(s)}\|_2) \max\left\{\frac{\lambda_{\max}(X_0)}{\lambda^+_{\min}(X_0)}, \frac{R_{e^-}}{\lambda^+_{\min}(W)}\right\} \frac{\max_{s\in\mathcal{S}} \sqrt{\lambda_{\max}((C^{(s)})^\top V_s^{-1} C^{(s)})}}{\min_{s\in\mathcal{S}} \lambda^+_{\min}((C^{(s)})^\top V_s^{-1} C^{(s)})}$$
$$= R_{\tilde{L}v}(\max_{s\in\mathcal{S}} \|V_s^{-1/2}\mathbf{v}_t^{(s)}\|_2),$$

where R_{e^-} is as defined in Lemma B.12.

Proof. In the following, we will use the notation of the stacked system from Definition 2.1

$$\mathbf{y}_t = C\mathbf{x}_t + \mathbf{v}_t = \begin{bmatrix} C^{(1)}\mathbf{x}_t \\ \vdots \\ C^{(|\mathcal{S}|)}\mathbf{x}_t \end{bmatrix} + \begin{bmatrix} \mathbf{v}_t^{(1)} \\ \vdots \\ \mathbf{v}_t^{(|\mathcal{S}|)} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_t^{(1)} \\ \vdots \\ \mathbf{y}_t^{(|\mathcal{S}|)} \end{bmatrix}, \quad \text{where} \quad C = \begin{bmatrix} C^{(1)} \\ \vdots \\ C^{(|\mathcal{S}|)} \end{bmatrix}, \quad V = \begin{bmatrix} V_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & V_{|\mathcal{S}|} \end{bmatrix}.$$

Our goal will be to show that, even though $\mathbf{v}_t \in \mathbb{R}^{d_\mathbf{y}\cdot|\mathcal{S}|}$, $\|\tilde{L}_t\mathbf{v}_t\|_2$ does not scale polynomially in $|\mathcal{S}|$. Now, by Proposition B.9, and the above definition of the stacked system, and denoting $P_t = \Sigma_{t|t-1}\Sigma^+_{t|t-1}$,

$$\tilde{L}_t\mathbf{v}_t = (\Sigma^+_{t|t-1} + \sum_{s\in\mathcal{S}} P_t(C^{(s)})^\top V_s^{-1} C^{(s)} P_t)^+ \sum_{s\in\mathcal{S}} P_t(C^{(s)})^\top V_s^{-1}\mathbf{v}_t^{(s)}.$$

This expression makes clear that, at least the case when $d_\mathbf{y} = 1$, $\tilde{L}_t\mathbf{v}_t$ should not scale polynomially in $|\mathcal{S}|$, since both numerator and denominator have a summation over $s \in \mathcal{S}$. Thus, in the following, we carefully exploit this structure for $d_\mathbf{y} > 1$.

To proceed, it will be convenient to rescale the terms in the above expression as follows. First, let

$$\widehat{C}^{(s)} := V_s^{-1/2} C^{(s)} \Sigma^{1/2}_{t|t-1}, \quad \widehat{C} := \begin{bmatrix} \widehat{C}^{(1)} \\ \vdots \\ \widehat{C}^{(|\mathcal{S}|)}, \end{bmatrix} = V^{-1/2} C \Sigma^{1/2}_{t|t-1} \quad \text{and} \quad \hat{\mathbf{v}}_t^{(s)} := V_s^{-1/2}\mathbf{v}_t^{(s)}.$$

Then, under this notation, notice that

$$\begin{aligned}
\tilde{L}_t\mathbf{v}_t &= (\Sigma^+_{t|t-1} + \sum_{s\in\mathcal{S}} P_t(C^{(s)})^\top V_s^{-1} C^{(s)} P_t)^+ \sum_{s\in\mathcal{S}} P_t(C^{(s)})^\top V_s^{-1}\mathbf{v}_t^{(s)} \\
&= \Sigma^{1/2}_{t|t-1}(I + \sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \widehat{C}^{(s)})^+ \Sigma^{1/2}_{t|t-1} \sum_{s\in\mathcal{S}} (C^{(s)})^\top V_s^{-1}\mathbf{v}_t^{(s)} \\
&= \Sigma^{1/2}_{t|t-1}(I + \sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \widehat{C}^{(s)})^+ \sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \hat{\mathbf{v}}_t^{(s)} \\
&= \Sigma^{1/2}_{t|t-1}(I + \widehat{C}^\top\widehat{C})^+ \sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \hat{\mathbf{v}}_t^{(s)}.
\end{aligned}$$

Now, let us decompose $\widehat{C}$ and $\widehat{C}^{(s)}$ by their (compact) singular value decompositions:

$$\widehat{C} = G\Lambda H^\top \quad \text{and} \quad \widehat{C}^{(s)} = G_s\Lambda_s H_s^\top,$$

where, taking $r, r_s \leq \min\{d_\mathbf{y}, d_\mathbf{x}\}$ to be the rank of $\widehat{C}$ and $\widehat{C}^{(s)}$ respectively, then G is $d_\mathbf{y} \times r$, G_s is $d_\mathbf{y} \times r_s$, H is $d_\mathbf{x} \times r$, and $H_s is d_\mathbf{x} \times r_s$. Λ, Λ_s are $r \times r$ and $r_s \times r_s$ (respectively) invertible positive diagonal matrices, and G, G_s, H, H_s are semi-orthogonal matrices satisfying $H^\top H = G^\top G = I_r$ and $H_s^\top H_s = G_s^\top G_s = I_{r_s}$ for all $s \in \mathcal{S}$. Thus, we can decompose:

$$\begin{aligned}
\tilde{L}_t\mathbf{v}_t &= \Sigma^{1/2}_{t|t-1}(I + \widehat{C}^\top\widehat{C})^+ \sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \mathbf{v}_t^{(s)} \\
&= \Sigma^{1/2}_{t|t-1}(I + H\Lambda G^\top G\Lambda H^\top)^+ \sum_{s\in\mathcal{S}} H_s\Lambda_s G_s^\top \mathbf{v}_t^{(s)} \\
&= \Sigma^{1/2}_{t|t-1}(I + H\Lambda^2 H^\top)^+ \sum_{s\in\mathcal{S}} H_s\Lambda_s G_s^\top \mathbf{v}_t^{(s)}.
\end{aligned}$$

Now, notice that $\mathsf{colsp}(H_s) \subseteq \mathsf{colsp}(H)$ for all $s \in \mathcal{S}$, which follows from the facts that $\Lambda \succ \mathbf{0}$, $\Lambda_s \succ \mathbf{0}$, and

$$H\Lambda^2 H^\top = \widehat{C}^\top \widehat{C} = \sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \widehat{C}^{(s)} = \sum_{s\in\mathcal{S}} H_s \Lambda_s G_s^\top G_s \Lambda_s H_s^\top = \sum_{s\in\mathcal{S}} H_s \Lambda_s^2 H_s^\top.$$

Therefore, we have that $HH^\top H_s = H_s$ for all s, which implies:

$$\Sigma_{t|t-1}^{1/2}(I + H\Lambda^2 H^\top)^+ H_s = \Sigma_{t|t-1}^{1/2} H(I+\Lambda^2)^{-1} H^\top H_s.$$

Thus, our above expression further simplifies to:

$$\begin{aligned}\tilde{L}_t \mathbf{v}_t &= \Sigma_{t|t-1}^{1/2}(I + H\Lambda^2 H^\top)^+ \sum_{s\in\mathcal{S}} H_s \Lambda_s G_s^\top \mathbf{v}_t^{(s)} \\ &= \Sigma_{t|t-1}^{1/2} \sum_{s\in\mathcal{S}} H(I+\Lambda^2)^{-1} H^\top H_s \Lambda_s G_s^\top \mathbf{v}_t^{(s)}.\end{aligned}$$

Now, denoting $g_{s,i}, h_{s,i}$ as the ith column vector of G_s, H_s respectively ($i \in [r_s]$) and ξ_j as the jth column vector of H ($j \in [r]$), $\lambda_{s,i}, \sigma_j$ as the ith and jth diagonal entries of Λ_s, Λ ($i \in [r_s], j \in [r]$)

$$\sum_{s\in\mathcal{S}} H(I+\Lambda^2)^{-1} H^\top H_s \Lambda_s G_s^\top \hat{\mathbf{v}}_t^{(s)} = \sum_{j\in[r]} \sum_{s\in\mathcal{S}} \sum_{i\in[r_s]} \frac{\lambda_{s,i}}{1+\sigma_j^2} \left\langle g_{s,i}, \hat{\mathbf{v}}_t^{(s)} \right\rangle \langle h_{s,i}, \xi_j \rangle \xi_j,$$

so, since G_s, H_s, H are semi-orthogonal matrices, it follows that the ℓ_2 norm of the above expression can be bounded as follows:

$$\begin{aligned}&\|\sum_{s\in\mathcal{S}} H(I+\Lambda^2)^{-1} H^\top H_s \Lambda_s G_s^\top \hat{\mathbf{v}}_t^{(s)}\|_2^2 \\ &= \sum_{j\in[r]} \left(\sum_{s\in\mathcal{S}} \sum_{i\in[r_s]} \frac{\lambda_{s,i}}{1+\sigma_j^2} \langle h_{s,i}, \xi_j \rangle \left\langle g_{s,i}, \hat{\mathbf{v}}_t^{(s)} \right\rangle \right)^2 \\ &\le \sum_{j\in[r]} \frac{1}{(1+\sigma_j^2)^2} \left(\sum_{s\in\mathcal{S}} \left(\sum_{i\in[r_s]} \lambda_{s,i}^2 \langle h_{s,i}, \xi_j \rangle^2 \right)^{1/2} \left(\sum_{i\in[r_s]} \left\langle g_{s,i}, \hat{\mathbf{v}}_t^{(s)} \right\rangle^2 \right)^{1/2} \right)^2 \\ &= \sum_{j\in[r]} \frac{1}{(1+\sigma_j^2)^2} \left(\sum_{s\in\mathcal{S}} \|\Lambda_s H_s^\top \xi_j\|_2 \|G_s^\top \hat{\mathbf{v}}_t^{(s)}\|_2 \right)^2,\end{aligned}$$

where the inequality follows by application of Cauchy-Schwarz. Further, by definition of σ_j,

$$\begin{aligned}\sigma_j^2 = \xi_j^\top H\Lambda^2 H^\top \xi_j = \xi_j^\top \widehat{C}^\top \widehat{C} \xi_j &= \xi_j^\top (\sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \widehat{C}^{(s)}) \xi_j \\ &= \xi_j^\top (\sum_{s\in\mathcal{S}} H_s \Lambda_s^2 H_s^\top) \xi_j = \sum_{s\in\mathcal{S}} \xi_j^\top H_s \Lambda_s^2 H_s^\top \xi_j = \sum_{s\in\mathcal{S}} \|\Lambda_s H_s^\top \xi_j\|_2^2.\end{aligned}$$

Collecting our results,

$$\begin{aligned}\|\tilde{L}_t \mathbf{v}_t\|_2^2 &\le \|\Sigma_{t|t-1}^{1/2}(I + \widehat{C}^\top \widehat{C})^+ \sum_{s\in\mathcal{S}} (\widehat{C}^{(s)})^\top \hat{\mathbf{v}}_t^{(s)}\|_2^2 \\ &\le \|\Sigma_{t|t-1}^{1/2}\|_{\mathsf{op}}^2 \sum_{j\in[r]} \frac{\left(\sum_{s\in\mathcal{S}} \|\Lambda_s H_s^\top \xi_j\|_2 \|G_s^\top \hat{\mathbf{v}}_t^{(s)}\|_2\right)^2}{(1+\sum_{s\in\mathcal{S}} \|\Lambda_s H_s^\top \xi_j\|_2^2)^2} \\ &\le \|\Sigma_{t|t-1}\|_{\mathsf{op}} \max_{s\in\mathcal{S}} \|G_s^\top \hat{\mathbf{v}}_t^{(s)}\|_2^2 \sum_{j\in[r]} \left(\frac{\sum_{s\in\mathcal{S}} \|\Lambda_s H_s^\top \xi_j\|_2}{1+\sum_{s\in\mathcal{S}} \|\Lambda_s H_s^\top \xi_j\|_2^2} \right)^2.\end{aligned}$$

The inner terms in the above summation can be bounded as follows: let $N_j = \sum_{s\in\mathcal{S}} \mathbb{1}\{\|\Lambda_s H_s^\top \xi_j\|_2 > 0\}$ denote the number of nonzero terms $\|\Lambda_s H_s^\top \xi_j\|_2$. Then,

$$\frac{\sum_{s\in\mathcal{S}}\|\Lambda_s H_s^\top \xi_j\|_2}{1+\sum_{s\in\mathcal{S}}\|\Lambda_s H_s^\top \xi_j\|_2^2} \le \frac{\max_{s\in\mathcal{S}}\|\Lambda_s H_s^\top \xi_j\|_2 N_j}{1+\min_{s\in\mathcal{S}:\|\Lambda_s H_s^\top \xi_j\|_2>0}\|\Lambda_s H_s^\top \xi_j\|_2^2 N_j} \le \frac{\max_{s\in\mathcal{S}}\lambda_{\max}(\Lambda_s) N_j}{1+\min_{s\in\mathcal{S}}\lambda_{\min}^+(\Lambda_s)^2 N_j}$$
$$\le \frac{\max_{s\in\mathcal{S}}\lambda_{\max}(\Lambda_s)}{\min_{s\in\mathcal{S}}\lambda_{\min}^+(\Lambda_s)^2}.$$

where $\lambda_{\min}^+(A)$ denotes the smallest nonzero eigenvalue of a matrix A. Thus, our bound can be further simplified to:

$$\|\tilde{L}_t \mathbf{v}_t\|_2^2 \le \|\Sigma_{t|t-1}\|_{\mathsf{op}} \max_{s\in\mathcal{S}} \|G_s^\top \hat{\mathbf{v}}_t^{(s)}\|_2^2 \sum_{j\in[r]} \left(\frac{\max_{s\in\mathcal{S}}\lambda_{\max}(\Lambda_s)}{\min_{s\in\mathcal{S}}\lambda_{\min}^+(\Lambda_s)^2}\right)^2$$
$$\le \|\Sigma_{t|t-1}\|_{\mathsf{op}} r \max_{s\in\mathcal{S}} \|G_s^\top \hat{\mathbf{v}}_t^{(s)}\|_2^2 \left(\frac{\max_{s\in\mathcal{S}}\lambda_{\max}(\Lambda_s)}{\min_{s\in\mathcal{S}}\lambda_{\min}^+(\Lambda_s)^2}\right)^2$$

Now, by definition, $G_s^\top$ is semi-orthogonal, and $\hat{\mathbf{v}}_t^{(s)} = V_s^{-1/2}\mathbf{v}_t^{(s)}$, so

$$\|G_s^\top \hat{\mathbf{v}}_t^{(s)}\|_2 \le \|G_s^\top\|_{\mathsf{op}} \|V_s^{-1/2}\hat{\mathbf{v}}_t^{(s)}\|_2 \le \|V_s^{-1/2}\|_{\mathsf{op}} \|\mathbf{v}_t^{(s)}\|_2.$$

Also, recall that $\Lambda_s \succ \mathbf{0}$ is the diagonal matrix of the (compact) SVD representation of $\widehat{C}^{(s)} = G_s \Lambda_s H_s^\top$. Recall $\widehat{C}^{(s)} = V_s^{-1/2} C^{(s)} \Sigma_{t|t-1}^{1/2}$. Therefore,

$$\lambda_{\max}(\Lambda_s) = \|\widehat{C}^{(s)}\|_{\mathsf{op}} = \|V_s^{-1/2} C^{(s)} \Sigma_{t|t-1}^{1/2}\|_{\mathsf{op}} \le \|V_s^{-1/2} C^{(s)}\|_{\mathsf{op}} \|\Sigma_{t|t-1}^{1/2}\|_{\mathsf{op}}$$
$$\le \sqrt{\lambda_{\max}((C^{(s)})^\top V_s^{-1} C^{(s)}) \lambda_{\max}(\Sigma_{t|t-1})},$$

Similarly,

$$\lambda_{\min}^+(\Lambda_s)^2 = \min\left\{ \frac{x^\top \Sigma_{t|t-1}^{1/2} (C^{(s)})^\top V_s^{-1} C^{(s)} \Sigma_{t|t-1}^{1/2} x}{x^\top x} : x^\top \Sigma_{t|t-1}^{1/2} (C^{(s)})^\top V_s^{-1} C^{(s)} \Sigma_{t|t-1}^{1/2} x > 0 \right\}$$
$$= \min\left\{ \frac{u^\top (C^{(s)})^\top V_s^{-1} C^{(s)} \Sigma_{t|t-1}^{1/2} u}{u^\top \Sigma_{t|t-1}^{-1} u} : u^\top (C^{(s)})^\top V_s^{-1} C^{(s)} u > 0 \right\}$$
$$\ge \frac{\lambda_{\min}^+((C^{(s)})^\top V_s^{-1} C^{(s)})}{\lambda_{\max}(\Sigma_{t|t-1}^+)}$$
$$\ge \lambda_{\min}^+((C^{(s)})^\top V_s^{-1} C^{(s)}) \lambda_{\min}^+(\Sigma_{t|t-1}).$$

Recall that, by definition, $\Sigma_{t+1|t} = W + A^\top \Sigma_{t|t} A^\top \succeq W$, so $\lambda_{\min}(\Sigma_{t|t-1}) \ge \min\{\lambda_{\min}(W), \lambda_{\min}(X_0)\}$. By Lemma B.12, $\lambda_{\max}(\Sigma_{t|t-1}) \le \mathrm{tr}(\Sigma_{t|t-1}) \le R_{e^-}$ for all $t \ge 0$. Finally, note that, since $V \succ \mathbf{0}$ and $\Sigma_{t|t-1} \succ \mathbf{0}$, $\mathsf{colsp}(\widehat{C}) = \mathsf{colsp}(V^{-1/2} C \Sigma_{t|t-1}^{1/2}) = \mathsf{colsp}(V^{-1/2} C)$, so

$$r = \mathsf{rank}(\widehat{C}) = \mathsf{rank}(V^{-1/2} C) = \mathsf{rank}(C).$$

Therefore,

$$\|\tilde{L}_t \mathbf{v}_t\|_2 \le \sqrt{r}(\max_{s\in\mathcal{S}} \|V_s^{-1/2}\mathbf{v}_t^{(s)}\|_2) \|\Sigma_{t|t-1}^{1/2}\|_{\mathsf{op}} \frac{\max_{s\in\mathcal{S}}\|V_s^{-1/2} C^{(s)}\|_{\mathsf{op}} \|\Sigma_{t|t-1}^{1/2}\|_{\mathsf{op}}}{\min_{s\in\mathcal{S}}\lambda_{\min}^+((C^{(s)})^\top V_s^{-1} C^{(s)}) \lambda_{\min}^+(\Sigma_{t|t-1})}$$
$$\le \sqrt{r}(\max_{s\in\mathcal{S}} \|V_s^{-1/2}\mathbf{v}_t^{(s)}\|_2) \frac{\|\Sigma_{t|t-1}\|_{\mathsf{op}}}{\lambda_{\min}^+(\Sigma_{t|t-1})} \frac{\max_{s\in\mathcal{S}}\sqrt{\lambda_{\max}((C^{(s)})^\top V_s^{-1} C^{(s)})}}{\min_{s\in\mathcal{S}}\lambda_{\min}^+((C^{(s)})^\top V_s^{-1} C^{(s)})}$$
$$\le \sqrt{\mathsf{rank}(C)}(\max_{s\in\mathcal{S}} \|V_s^{-1/2}\mathbf{v}_t^{(s)}\|_2) \max\left\{\frac{\lambda_{\max}(X_0)}{\lambda_{\min}^+(X_0)}, \frac{R_{e^-}}{\lambda_{\min}^+(W)}\right\} \frac{\max_{s\in\mathcal{S}}\sqrt{\lambda_{\max}((C^{(s)})^\top V_s^{-1} C^{(s)})}}{\min_{s\in\mathcal{S}}\lambda_{\min}^+((C^{(s)})^\top V_s^{-1} C^{(s)})}$$
$$= R_{\tilde{L}v}(\max_{s\in\mathcal{S}} \|V_s^{-1/2}\mathbf{v}_t^{(s)}\|_2),$$

as claimed. □

G.3 Kalman Estimate Norm Bounds

Lemma G.3 (Kalman estimate bound)**.** *Assuming Assumption 1 and Assumption 4', then we have the following bounds:*

$$\begin{aligned}
\|\mathbf{e}_t^-\|_2 &\le R_{A-LC}(\rho_{A-LC}^t\|\mathbf{e}_0^-\|_2+\sum_{i=1}^t \rho_{A-LC}^{i-1}(\|\mathbf{w}_{t-i}\|_2+R_A R_{\tilde{L}v}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}_{t-i}\|_2)\\
\|\mathbf{e}_t\|_2 &\le R_{I-\tilde{L}C}\|\mathbf{e}_t^-\|_2+R_{\tilde{L}v}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}_t^{(s)}\|_2\\
\|\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}\|_2 &\le \|\mathbf{e}_t\|_2+\|\mathbf{x}_t^{\mathsf{nat}}\|_2\\
\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2 &\le \|\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}\|_2+R_A R_B\sum_{i=1}^t \rho_A^{i-1}\|\mathbf{u}_t(\mathbf{M})\|_2.
\end{aligned}$$

In particular, under Assumption 2, for any $\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)$,

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{e}_t^-\|_2^8\right] &\le R_{A-LC}^8\left(1+\frac{1+(R_AR_{\tilde{L}v})^{8/7}}{1-\rho_{A-LC}}\right)^7(R_{e_0^-}^8+\frac{R_w^8+R_{\bar{v}^s}^8}{1-\rho_{A-LC}}) := R_{e^-}^8\\
\mathbb{E}\left[\|\mathbf{e}_t\|_2^8\right] &\le 128(R_{I-\tilde{L}C}^8R_{e^-}^8+R_{\tilde{L}v}^8R_{\bar{v}^s}^8) := R_e^8\\
\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}\|_2^8\right] &\le 128(R_e^8+R_{x,\mathsf{nat}}^8) := R_{x,\mathsf{kal,nat}}^8\\
\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2^8\right] &\le (1+\frac{(R_AR_B)^{8/7}}{1-\rho_A})^7(R_{x,\mathsf{kal,nat}}+\frac{R_u^8}{1-\rho_A}) := R_{x,\mathsf{kal}}^8.
\end{aligned}$$

Proof. By Lemma B.10, we can decompose $\mathbf{e}_t^-$ as:

$$\mathbf{e}_t^- = \Psi^{[t]}\mathbf{e}_0^- + \sum_{i=1}^t \Psi^{[t-i+1,t]}(\mathbf{w}_{t-i}-L_{t-i}\mathbf{v}_{t-i})$$

By Lemma B.13,

$$\begin{aligned}
\|\mathbf{e}_t^-\|_2 &\le \|\Psi^{[t]}\|_{\mathsf{op}}\|\mathbf{e}_0^-\|_2+\sum_{i=1}^t\|\Psi^{[t-i+1,t]}\|_{\mathsf{op}}\|(\mathbf{w}_{t-i}-L_{t-i}\mathbf{v}_{t-i})\|_2\\
&\le R_{A-LC}(\rho_{A-LC}^t\|\mathbf{e}_0^-\|_2+\sum_{i=1}^t\rho_{A-LC}^{i-1}(\|\mathbf{w}_{t-i}\|_2+\|L_{t-i}\mathbf{v}_{t-i}\|_2).
\end{aligned}$$

By Lemma G.2,

$$\|\mathbf{e}_t^-\|_2 \le R_{A-LC}(\rho_{A-LC}^t\|\mathbf{e}_0^-\|_2+\sum_{i=1}^t\rho_{A-LC}^{i-1}(\|\mathbf{w}_{t-i}\|_2+R_AR_{\tilde{L}v}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}_{t-i}^{(s)}\|_2).$$

Therefore, by Lemma G.1,

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{e}_t^-\|_2^8\right] &\le R_{A-LC}^8\left(1+\frac{1+(R_AR_{\tilde{L}v})^{8/7}}{1-\rho_{A-LC}}\right)^7\\
&\quad\times(\mathbb{E}\left[\|\mathbf{e}_0^-\|_2^8\right]+\sum_{i=1}^t\rho_{A-LC}^{i-1}(\mathbb{E}\left[\|\mathbf{w}_{t-i}\|_2^8\right]+\mathbb{E}\left[\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}_{t-i}^{(s)}\|_2^8\right]))\\
&\le R_{A-LC}^8\left(1+\frac{1+(R_AR_{\tilde{L}v})^{8/7}}{1-\rho_{A-LC}}\right)^7\times(R_{e_0^-}^8+\frac{R_w^8+R_{\bar{v}^s}^8}{1-\rho_{A-LC}})\\
&:= R_{e^-}^8
\end{aligned}$$

For the second, by Lemma B.10,

$$\mathbf{e}_t = (I-\tilde{L}_tC)\mathbf{e}_t^- - \tilde{L}_t\mathbf{v}_t.$$

Now, by Lemma B.13, $\|I - \tilde{L}_t C\|_{\mathsf{op}} \leq R_{I-\tilde{L}C}$ for all t. Therefore, by Lemma G.2 together with the above bound,

$$\begin{aligned}\|\mathbf{e}_t\| &\leq \|I - \tilde{L}_t C\|_{\mathsf{op}} \|\mathbf{e}_t^-\|_2 + \|\tilde{L}_t \mathbf{v}_t\|_2 \\ &\leq R_{I-\tilde{L}C} \|\mathbf{e}_t^-\|_2 + R_{\tilde{L}v} \max_{s \in \mathcal{S}} \|V_s^{-1/2} \mathbf{v}_t^{(s)}\|_2\end{aligned}$$

Therefore,

$$\mathbb{E}\left[\|\mathbf{e}_t\|_2^8\right] \leq 128(R_{I-\tilde{L}C}^8 R_{e^-}^8 + R_{\tilde{L}v}^8 R_{\bar{v}^s}^8) := R_e^8.$$

Now, recalling that $\mathbf{e}_t = \mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}$,

$$\|\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}\|_2 \leq \|\mathbf{e}_t\|_2 + \|\mathbf{x}_t^{\mathsf{nat}}\|_2,$$

so

$$\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}\|_2^8\right] \leq 128(R_e^8 + R_{x,\mathsf{nat}}^8) := R_{x,\mathsf{kal,nat}}^8.$$

By Lemma C.1,

$$\widehat{\mathbf{x}}_{t|t}(\mathbf{M}) = \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}} + \sum_{i=1}^{t} A^{i-i} B \mathbf{u}_{t-i}(\mathbf{M}).$$

Therefore,

$$\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2 \leq \|\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}\|_2 + R_A R_B \sum_{i=1}^{t} \rho_A^{i-1} \|\mathbf{u}_t(\mathbf{M})\|_2.$$

Thus,

$$\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2^8\right] \leq (1 + \frac{(R_A R_B)^{8/7}}{1-\rho_A})^7 (R_{x,\mathsf{kal,nat}} + \frac{R_u^8}{1-\rho_A}) := R_{x,\mathsf{kal}}^8.$$

□

G.4 Estimated Loss Gradient Bounds

Lemma G.4 (Gradient bound)**.** *Recall the loss function $\widehat{\ell}_t(\mathbf{M}) := c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}), \mathbf{u}_t(\mathbf{M}))$. Suppose that the cost functions c_t satisfy Assumption 3, and the noise satisfies Assumption 2. Then, for any sensor $s \in \mathcal{S}$, and any $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$,*

$$\nabla_{M_j^{(s)}} \widehat{\ell}_t(\mathbf{M}) = 2 \sum_{i=1}^{t} (A^{i-1}B)^\top S_t \widehat{\mathbf{x}}_{t|t}(\mathbf{M}) (\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)})^\top + 2 R_t \mathbf{u}_t(\mathbf{M}) (\mathbf{y}_{t-j}^{\mathsf{nat},(s)})^\top.$$

Further, for any $j \in [0, H]$,

$$\begin{aligned}\mathbb{E}\left[\max_{s \in \mathcal{S}} \|\nabla_{M_j^{(s)}} \widehat{\ell}_t(\mathbf{M})\|_F^4\right] &\leq 16(16 + \frac{1}{(1-\rho_A)^3}) \sqrt{R_{u,\mathsf{cost}}^8 R_u^8 + (R_A R_B R_{x,\mathsf{cost}})^8 R_{x,\mathsf{kal}}^8} \\ &\quad \times \left(1 + \frac{1}{1-\rho_A}\right)^{3/2} R_{y^{(s)},\mathsf{nat,sup}}^4 \\ &:= R_{\mathsf{grad,coord}}.\end{aligned}$$

and

$$\mathbb{E}\left[\|\nabla \widehat{\ell}_t(\mathbf{M})\|_{\sqrt{\rho_M},\infty}^4\right] \leq \frac{R_{\mathsf{grad,coord}}}{(1-\rho_M)^2} := R_{\mathsf{grad},\infty},$$

where $\|\cdot\|_{\rho,\infty}$ is the group norm defined in Definition E.3.

Alternatively, suppose that the costs satisfy Assumption 3', and the noise is adversarial satisfying Assumption 2'. Then, $\widehat{\ell}_t(\mathbf{M})$ is $R_{\mathsf{grad},\infty}$-Lipschitz with respect to $\|\cdot\|_{1/\sqrt{\rho_M},1}$ for all $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$, where

$$R_{\mathsf{grad},\infty} := LR\frac{R_{y^{(s)},\mathsf{nat},\sup}}{\sqrt{1-\rho_M}}\left(1+\frac{R_X R_Y R_C R_A R_B}{(1-\rho_A)(1-\rho_X)}\right) + 2L\left(\frac{R_{y^{(s)},\mathsf{nat},\sup}}{\sqrt{1-\rho_M}}\left(1+\frac{R_X R_Y R_C R_A R_B}{(1-\rho_A)(1-\rho_X)}\right)\right)^2\sqrt{\min\left\{d_{y^s,\max}, d_{\mathbf{u}}\right\}}\frac{R_M}{1-\rho_M}.$$

In particular, this implies that

$$\|g_t\|_{\sqrt{\rho_M},\infty} \le R_{\mathsf{grad},\infty} \quad \forall g_t \in \partial\widehat{\ell}_t(\mathbf{M}), \mathbf{M} \in \mathcal{M}(R_M, \rho_M, H).$$

Proof. We first consider the stochastic setting, where the cost functions satisfy Assumption 3 and the noise satisfies Assumption 2. Recalling our definition of $\widehat{\ell}_t(\mathbf{M})$, and by Lemma C.1 and definition of controls,

$$\begin{aligned}
\widehat{\ell}_t(\mathbf{M}) &= \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t \widehat{\mathbf{x}}_{t|t}(\mathbf{M}) + \mathbf{u}_t(\mathbf{M})^\top R_t \mathbf{u}_t(\mathbf{M}) \\
\widehat{\mathbf{x}}_{t|t}(\mathbf{M}) &= \widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} + \Delta_t(\mathbf{M}) = \widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} + \sum_{i=1}^{t} A^{i-1} B \mathbf{u}_{t-i}(\mathbf{M}) \\
\mathbf{u}_t(\mathbf{M}) &= \sum_{j=0}^{H}\sum_{s\in\mathcal{S}} M_j^{(s)} \mathbf{y}^{\mathsf{nat},(s)}_{t-j}
\end{aligned}$$

Thus, by application of the chain rule,

$$\nabla_{M_j^{(s)}}\widehat{\ell}_t(\mathbf{M}) = 2\sum_{i=1}^{t}(A^{i-1}B)^\top S_t \widehat{\mathbf{x}}_{t|t}(\mathbf{M})(\mathbf{y}^{\mathsf{nat},(s)}_{t-i-j})^\top + 2R_t\mathbf{u}_t(\mathbf{M})(\mathbf{y}^{\mathsf{nat},(s)}_{t-j})^\top.$$

Therefore, since $\|xy^\top\|_F^2 = \mathrm{tr}(yx^\top xy^\top) = \|x\|_2^2\|y\|_2^2$,

$$\begin{aligned}
\|\nabla_{M_j^{(s)}}\widehat{\ell}_t(\mathbf{M})\|_F &\le 2\sum_{i=1}^{t}\|(A^{i-1}B)^\top S_t\widehat{\mathbf{x}}_{t|t}(\mathbf{M})(\mathbf{y}^{\mathsf{nat},(s)}_{t-i-j})^\top\|_F + 2\|R_t\mathbf{u}_t(\mathbf{M})(\mathbf{y}^{\mathsf{nat},(s)}_{t-j})^\top\|_F \\
&\le 2R_A R_B R_{x,\mathsf{cost}}\sum_{i=1}^{t}\rho_A^{i-1}\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2\|\mathbf{y}^{\mathsf{nat},(s)}_{t-i-j}\|_2 \\
&\quad + 2R_{u,\mathsf{cost}}\|\mathbf{u}_t(\mathbf{M})\|_2\|\mathbf{y}^{\mathsf{nat},(s)}_{t-j}\|_2,
\end{aligned}$$

where, in the second inequality, we applied Assumption 1 and 3. Therefore,

$$\begin{aligned}
\|\nabla_{M_j^{(s)}}\widehat{\ell}_t(\mathbf{M})\|_F^4 &\le \left(R_{u,\mathsf{cost}}^2\|\mathbf{u}_t(\mathbf{M})\|_2^2 + (R_A R_B R_{x,\mathsf{cost}})^2\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2^2\right)^2 \\
&\quad\times\left(4\|\mathbf{y}^{\mathsf{nat},(s)}_{t-j}\|_2^2 + \frac{1}{1-\rho_A}\sum_{i=1}^{t}\rho_A^{i-1}\|\mathbf{y}^{\mathsf{nat},(s)}_{t-i-j}\|_2^2\right)^2 \\
&\le 8(16+\frac{1}{(1-\rho_A)^3})\left(R_{u,\mathsf{cost}}^4\|\mathbf{u}_t(\mathbf{M})\|_2^4 + (R_A R_B R_{x,\mathsf{cost}})^4\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2^4\right) \\
&\quad\times\left(\|\mathbf{y}^{\mathsf{nat},(s)}_{t-j}\|_2^4 + \sum_{i=1}^{t}\rho_A^{i-1}\|\mathbf{y}^{\mathsf{nat},(s)}_{t-i-j}\|_2^4\right)
\end{aligned}$$

Using the above bound, observe that

$$\begin{aligned}
&\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\nabla_{M_j^{(s)}}\widehat{\ell}_t(\mathbf{M})\|_F^4\right] \\
&\le 8(16+\frac{1}{(1-\rho_A)^3})\sqrt{\mathbb{E}\left[\left(R_{u,\mathsf{cost}}^4\|\mathbf{u}_t(\mathbf{M})\|_2^4+(R_AR_BR_{x,\mathsf{cost}})^4\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2^4\right)^2\right]} \\
&\quad\times\sqrt{\mathbb{E}\left[\max_{s\in\mathcal{S}}\left(\|\mathbf{y}_{t-j}^{\mathsf{nat},(s)}\|_2^4+\sum_{i=1}^t\rho_A^{i-1}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^4\right)^2\right]} \\
&\le 16(16+\frac{1}{(1-\rho_A)^3})\sqrt{R_{u,\mathsf{cost}}^8\mathbb{E}\left[\|\mathbf{u}_t(\mathbf{M})\|_2^8\right]+(R_AR_BR_{x,\mathsf{cost}})^8\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M})\|_2^8\right]} \\
&\quad\times\left(1+\frac{1}{1-\rho_A}\right)\sqrt{\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-j}^{\mathsf{nat},(s)}\|_2^8\right]+\sum_{i=1}^t\rho_A^{i-1}\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^8\right]}.
\end{aligned}$$

Therefore, by Lemma G.1 and Lemma G.3,

$$\begin{aligned}
\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\nabla_{M_j^{(s)}}\widehat{\ell}_t(\mathbf{M})\|_F^4\right] &\le 16(16+\frac{1}{(1-\rho_A)^3})\sqrt{R_{u,\mathsf{cost}}^8R_u^8+(R_AR_BR_{x,\mathsf{cost}})^8R_{x,\mathsf{kal}}^8} \\
&\quad\times\left(1+\frac{1}{1-\rho_A}\right)^{3/2}R_{y^{(s)},\mathsf{nat},\mathsf{sup}}^4 \\
&:= R_{\mathsf{grad},\mathsf{coord}}.
\end{aligned}$$

Then, recalling the definition of $\|\cdot\|_{\sqrt{\rho_M},\infty}$ from Definition E.3, we have that

$$\|\nabla\widehat{\ell}_t(\mathbf{M})\|_{\sqrt{\rho_M},\infty}^4=\max_{s\in\mathcal{S}}\left(\sum_{j=0}^H\rho_M^j\|\nabla_{M_j^{(s)}}\widehat{\ell}_t(\mathbf{M})\|_F^2\right)^2\le\frac{1}{1-\rho_M}\sum_{j=0}^H\rho_M^j\max_{s\in\mathcal{S}}\|\nabla_{M_j^{(s)}}\widehat{\ell}_t(\mathbf{M})\|_F^4$$

Therefore,

$$\mathbb{E}\left[\|\nabla\widehat{\ell}_t(\mathbf{M})\|_{\sqrt{\rho_M},\infty}^4\right]\le\frac{R_{\mathsf{grad},\mathsf{coord}}}{(1-\rho_M)^2}:=R_{\mathsf{grad},\infty},$$

as claimed.

Now, in the adversarial setting, we consider cost functions satisfying Assumption 3'. Thus, we know that

$$|c_t(\mathbf{x},\mathbf{u})-c_t(\mathbf{x}',\mathbf{u}')|\le L\max\left\{R\left\|\begin{bmatrix}\mathbf{x}-\mathbf{x}'\\\mathbf{u}-\mathbf{u}'\end{bmatrix}\right\|_2,\left\|\begin{bmatrix}\mathbf{x}-\mathbf{x}'\\\mathbf{u}-\mathbf{u}'\end{bmatrix}\right\|_2^2\right\}.$$

Now, by definition of $\mathbf{u}_t(\mathbf{M})$,

$$\mathbf{u}_t(\mathbf{M})-\mathbf{u}_t(\tilde{\mathbf{M}})=\sum_{s\in\mathcal{S}}\sum_{j=0}^H(M_j^{(s)}-\tilde{M}_j^{(s)})\mathbf{y}_{t-j}^{\mathsf{nat},(s)},$$

so, by Lemma G.1,

$$\begin{aligned}
\|\mathbf{u}_t(\mathbf{M})-\mathbf{u}_t(\tilde{\mathbf{M}})\|_2&\le\sum_{s\in\mathcal{S}}\sum_{j=0}^H\frac{1}{\sqrt{\rho_M^j}}\|M_j^{(s)}-\tilde{M}_j^{(s)}\|_{\mathsf{op}}\sqrt{\rho_M^j}\|\mathbf{y}_{t-j}^{\mathsf{nat},(s)}\|_2 \\
&\le\|\mathbf{M}-\tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1}\max_{s\in\mathcal{S}}\sqrt{\sum_{j=0}^H\rho_M^j\|\mathbf{y}_{t-j}^{\mathsf{nat},(s)}\|_2^2} \\
&\le\frac{R_{y^{(s)},\mathsf{nat},\mathsf{sup}}}{\sqrt{1-\rho_M}}\|\mathbf{M}-\tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1}.
\end{aligned}$$

Similarly, since the state estimator satisfies Assumption 4', and by Lemma C.1,

$$\begin{aligned}\widehat{\mathbf{x}}_{t|t}(\mathbf{M}) - \widehat{\mathbf{x}}_{t|t}(\tilde{\mathbf{M}}) &= \sum_{i=0}^{t} X^i Y(\mathbf{y}_{t-i}(\mathbf{M}) - \mathbf{y}_{t-i}(\tilde{\mathbf{M}})) \\ &= \sum_{i=0}^{t} X^i YC(\mathbf{x}_{t-i}(\mathbf{M}) - \mathbf{x}_{t-i}(\tilde{\mathbf{M}})) \\ &= \sum_{i=0}^{t} X^i YC(\Delta_{t-i}(\mathbf{M}) - \Delta_{t-i}(\tilde{\mathbf{M}})) \\ &= \sum_{i=0}^{t}\sum_{\ell=1}^{t-i} X^i YCA^{\ell-1}B(\mathbf{u}_{t-i-\ell}(\mathbf{M}) - \mathbf{u}_{t-i-\ell}(\tilde{\mathbf{M}})).\end{aligned}$$

Therefore,

$$\begin{aligned}\|\widehat{\mathbf{x}}_{t|t}(\mathbf{M}) - \widehat{\mathbf{x}}_{t|t}(\tilde{\mathbf{M}})\| &\le R_X R_Y R_C R_A R_B \sum_{i=0}^{t} \rho_X^i \sum_{\ell=1}^{t-i} \rho_A^{\ell-1} \|\mathbf{u}_{t-i-\ell}(\mathbf{M}) - \mathbf{u}_{t-i-\ell}(\tilde{\mathbf{M}})\|_2 \\ &\le \frac{R_X R_Y R_C R_A R_B R_{y^{(s)},\mathsf{nat},\sup}}{\sqrt{1-\rho_M}(1-\rho_A)(1-\rho_X)} \|\mathbf{M} - \tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1}\end{aligned}$$

Additionally, by Lemma G.5, for any $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$,

$$\|\mathbf{M}\|_{1/\sqrt{\rho_M},1} \le \sqrt{\min\{d_{y^s,\max}, d_{\mathbf{u}}\}} \frac{R_M}{1-\sqrt{\rho_M}}.$$

Combining these bounds, we conclude that, by Assumption 3', for any $\mathbf{M}, \tilde{\mathbf{M}} \in \mathcal{M}(R_M, \rho_M, H)$,

$$\begin{aligned}&|\widehat{\ell}_t(\mathbf{M}) - \widehat{\ell}_t(\tilde{\mathbf{M}})| \\ &= |c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}), \mathbf{u}_t(\mathbf{M})) - c_t(\widehat{\mathbf{x}}_{t|t}(\tilde{\mathbf{M}}), \mathbf{u}_t(\tilde{\mathbf{M}}))| \\ &\le L \max\left\{ R \left\| \begin{bmatrix} \widehat{\mathbf{x}}_{t|t}(\mathbf{M}) - \widehat{\mathbf{x}}_{t|t}(\tilde{\mathbf{M}}) \\ \mathbf{u}(\mathbf{M}) - \mathbf{u}(\tilde{\mathbf{M}}) \end{bmatrix} \right\|_2, \left\| \begin{bmatrix} \widehat{\mathbf{x}}_{t|t}(\mathbf{M}) - \widehat{\mathbf{x}}_{t|t}(\tilde{\mathbf{M}}) \\ \mathbf{u}(\mathbf{M}) - \mathbf{u}(\tilde{\mathbf{M}}) \end{bmatrix} \right\|_2^2 \right\} \\ &\le LR \frac{R_{y^{(s)},\mathsf{nat},\sup}}{\sqrt{1-\rho_M}} \left(1 + \frac{R_X R_Y R_C R_A R_B}{(1-\rho_A)(1-\rho_X)}\right) \|\mathbf{M} - \tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1} \\ &\quad + L \left(\frac{R_{y^{(s)},\mathsf{nat},\sup}}{\sqrt{1-\rho_M}} \left(1 + \frac{R_X R_Y R_C R_A R_B}{(1-\rho_A)(1-\rho_X)}\right) \right)^2 \|\mathbf{M} - \tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1}^2 \\ &\le LR \frac{R_{y^{(s)},\mathsf{nat},\sup}}{\sqrt{1-\rho_M}} \left(1 + \frac{R_X R_Y R_C R_A R_B}{(1-\rho_A)(1-\rho_X)}\right) \|\mathbf{M} - \tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1} \\ &\quad + 2L \left(\frac{R_{y^{(s)},\mathsf{nat},\sup}}{\sqrt{1-\rho_M}} \left(1 + \frac{R_X R_Y R_C R_A R_B}{(1-\rho_A)(1-\rho_X)}\right) \right)^2 \sqrt{\min\{d_{y^s,\max}, d_{\mathbf{u}}\}} \frac{R_M}{1-\rho_M} \|\mathbf{M} - \tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1} \\ &:= R_{\mathsf{grad},\infty} \|\mathbf{M} - \tilde{\mathbf{M}}\|_{1/\sqrt{\rho_M},1}.\end{aligned}$$

In particular, this establishes that $\widehat{\ell}_t(\cdot)$ is $R_{\mathsf{grad},\infty}$-Lipschitz with respect to $\|\cdot\|_{1/\sqrt{\rho_M},1}$ for $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$. Hence, we have that $\|g_t\|_{\sqrt{\rho_M},\infty} \le R_{\mathsf{grad},\infty}$ for all $g_t \in \partial\widehat{\ell}_t(\mathbf{M})$. □

G.5 Control Matrix Norm Bounds

Lemma G.5 (Control group-norm bound)**.** *For any $\mathbf{M} \in \mathcal{M}(R_M, \rho_M, H)$, the $\rho_M^{-1/2}$-reweighted $(2,1)$-group norm (see Definition E.3) of $\mathbf{M}$ is:*

$$\|\mathbf{M}\|_{\frac{1}{\sqrt{\rho_M}},1} \le \sqrt{\min\{d_{y^s,\max}, d_{\mathbf{u}}\}} \frac{R_M}{1-\sqrt{\rho_M}}.$$

Proof. By the choice of constraints of the algorithm, each iterate $\mathbf{M}(t)$ satisfies $\sum_{s\in\mathcal{S}}\|M_j^{(s)}(t)\|_{\mathsf{op}} \leq R_M\rho_M^j$. Using this observation, together with the fact that $\|M_j^{(s)}\|_F \leq \sqrt{\min\{d_{y^s,\max}, d_{\mathbf{u}}\}}\|M_j^{(s)}\|_{\mathsf{op}}$, it follows that, by Definition E.3,

$$\begin{aligned}
\|\mathbf{M}\|_{\frac{1}{\sqrt{\rho_M}},1} &= \sum_{s\in\mathcal{S}}\sqrt{\sum_{j=0}^{H}\frac{1}{\rho_M^j}\|M_j^{(s)}\|_F^2} \\
&\leq \sum_{s\in\mathcal{S}}\sum_{j=0}^{H}\frac{1}{\rho_M^{j/2}}\|M_j^{(s)}\|_F \\
&\leq \sqrt{\min\{d_{y^s,\max}, d_{\mathbf{u}}\}}\sum_{j=0}^{H}\frac{1}{\rho_M^{j/2}}\sum_{s\in\mathcal{S}}\|M_j^{(s)}\|_{\mathsf{op}} \\
&\leq \sqrt{\min\{d_{y^s,\max}, d_{\mathbf{u}}\}}\sum_{j=0}^{H}\frac{1}{\rho_M^{j/2}}R_M\rho_M^j \\
&\leq \sqrt{\min\{d_{y^s,\max}, d_{\mathbf{u}}\}}\frac{R_M}{1-\sqrt{\rho_M}},
\end{aligned}$$

as claimed. □

H Regret Bound: Deferred Proofs

H.1 Cost Decomposition

Lemma H.1 (Cost decomposition)**.** *Suppose $c_t(\mathbf{x},\mathbf{u}) = \mathbf{x}^\top S_t\mathbf{x} + \mathbf{u}^\top R_t\mathbf{u}$. Then,*

$$c_t(\mathbf{x}_t,\mathbf{u}_t) = c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) + (\mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}})^\top S_t(\mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}) + 2\widehat{\mathbf{x}}_{t|t}^\top S_t(\mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}})$$

Proof. First, notice that

$$c_t(\mathbf{x}_t,\mathbf{u}_t) = c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) + (\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t})^\top S_t(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}) + 2\widehat{\mathbf{x}}_{t|t}^\top S_t(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t}).$$

Further, by Lemma C.1,

$$\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t} = \mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}.$$

Combining these two equalities establishes the claim. □

H.2 Error of adjacent iterate depends on step-size

Lemma H.2 (Adjacent iterate bound)**.** *Suppose that Algorithm 1 is run with regularizer* (Reg)*. Then the iterates $\mathbf{M}(t)$ satisfy:*

$$\|\mathbf{M}(t+1) - \mathbf{M}(t)\|_{\frac{1}{\sqrt{\rho_M}},1} \leq \eta\|\nabla\widehat{\ell}_t(\mathbf{M}(t))\|_{\sqrt{\rho_M},\infty},$$

where $\|\cdot\|_{\rho,q}$ is the ρ-reweighted $(2,q)$-group norm from Definition E.3.

Proof. By Corollary E.2, it follows that $\Psi(\mathbf{M})$ is $1/\eta$-strongly convex with respect to $\|\cdot\|_{\rho_M^{-1/2},1}$, where the dual norm $(\|\cdot\|_{\rho_M^{-1/2},1})_* = \|\cdot\|_{\rho_M^{1/2},\infty}$. Therefore, the claim follows immediately from Lemma F.2. □

H.3 Error due to loss function memory

H.3.1 Error in state estimate function values

Lemma H.3 (Cost vs FTRL loss)**.** *Suppose that Algorithm 1 is run using the same regularizer from Lemma H.2. Then, at any time t, the expected difference between the cost $c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t)$ and the loss function $\widehat{\ell}_t(\mathbf{M}(t)) = c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)),\mathbf{u}_t(\mathbf{M}(t)))$ can be bounded as:*

1. *(Stochastic): If the cost functions satisfy Assumption 3, then*
$$\mathbb{E}\left[c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] \le \eta R_{x,\mathsf{cost}}(\eta R^2_{\widehat{\mathbf{x}},\mathsf{mem}} + 2R_x R_{\widehat{\mathbf{x}},\mathsf{mem}})$$
2. *(Adversarial): If the cost functions satisfy Assumption 3' and the state estimator satisfies Assumption 4',*
$$c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t)) \le \eta L \max\left\{R R_{\widehat{\mathbf{x}},\mathsf{mem}}, \eta R^2_{\widehat{\mathbf{x}},\mathsf{mem}}\right\},$$

Proof. In the stochastic case, by Assumption 3, $c_t(\mathbf{x}, \mathbf{u}) = \mathbf{x}^\top S_t \mathbf{x} + \mathbf{u}^\top R_t \mathbf{u}$, and since $\mathbf{u}_t = \mathbf{u}_t(\mathbf{M}(t))$, we have that:
$$\begin{aligned}
c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t)) &= \widehat{\mathbf{x}}_{t|t}^\top S_t \widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))^\top S_t \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)) \\
&= (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)))^\top S_t (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))) \\
&\quad + 2\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))^\top S_t (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))) \\
&= (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)))^\top S_t (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))) \\
&\quad + 2\mathbf{x}_t(\mathbf{M}(t))^\top S_t (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))) \\
&\quad + 2(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)) - \mathbf{x}_t(\mathbf{M}(t)))^\top S_t (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))).
\end{aligned}$$
To proceed, we begin by noticing the following: First, by Lemma C.1,
$$\begin{aligned}
\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}_H(t)) &= \Delta_t - \Delta_t(\mathbf{M}_H(t)) \\
&= \sum_{i=1}^{t} A^{i-1} B(\mathbf{u}_{t-i}(\mathbf{M}_H(t-i)) - \mathbf{u}_{t-i}(\mathbf{M}_H(t))) \\
&= \sum_{i=1}^{t} A^{i-1} B \sum_{s\in\mathcal{S}} \sum_{j=0}^{H} (M_j^{(s)}(t-i) - M_j^{(s)}(t)) \mathbf{y}^{\mathsf{nat},(s)}_{t-i-j}. \qquad (3)
\end{aligned}$$
In particular, this difference is a linear combination of $\mathbf{y}^{\mathsf{nat}}_\ell$, $\ell < t$ Therefore, by the Orthogonality principle Lemma B.7,
$$\mathbb{E}\left[(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)) - \mathbf{x}_t(\mathbf{M}))^\top S_t (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)))\right] = 0.$$
Hence, the above expression simplifies to:
$$\begin{aligned}
\mathbb{E}\left[c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] &= \mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|^2_{S_t}\right] + 2\mathbb{E}\left[\mathbf{x}_t(\mathbf{M})^\top S_t (\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)))\right] \\
&\le \|S_t\|_{\mathsf{op}} \mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2^2\right] \\
&\quad + 2\|S_t\|_{\mathsf{op}} \sqrt{\mathbb{E}\left[\|\mathbf{x}_t(\mathbf{M}(t))\|_2^2\right] \mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2^2\right]}.
\end{aligned}$$
Using Lemma G.1, $\mathbb{E}\left[\|\mathbf{x}_t(\mathbf{M}(t))\|_2^2\right] \le R_x^2$, so since $\|S_t\|_{\mathsf{op}} \le R_{x,\mathsf{cost}}$ (by Assumption 3),
$$\begin{aligned}
\mathbb{E}\left[c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] \le R_{x,\mathsf{cost}} (&\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}_H(t))\|_2^2\right] \\
&+ 2R_x \sqrt{\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}_H(t))\|_2^2\right]}) \qquad (4)
\end{aligned}$$
Finally, by Lemma H.5, we conclude
$$\mathbb{E}\left[c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] \le \eta R_{x,\mathsf{cost}}(\eta R^2_{\widehat{\mathbf{x}},\mathsf{mem}} + 2R_x R_{\widehat{\mathbf{x}},\mathsf{mem}}),$$
as claimed.

In the adversarial case, by Assumption 3' and by definition of $\widehat{\ell}_t(\mathbf{M}(t))$ and $\mathbf{u}_t$,
$$\begin{aligned}
c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t)) &= c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t(\mathbf{M}(t))) - c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)), \mathbf{u}_t(\mathbf{M}(t))) \\
&\le L \max\left\{R\|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2, \|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2^2\right\}.
\end{aligned}$$
Therefore, by Lemma H.5, we conclude
$$c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t)) \le \eta L \max\left\{R R_{\widehat{\mathbf{x}},\mathsf{mem}}, \eta R^2_{\widehat{\mathbf{x}},\mathsf{mem}}\right\},$$
as claimed. □

H.3.2 State and control error due to memory

Lemma H.4 (Control memory bound)**.** *For any times $0 \le t-i \le t$, assuming Assumption 2, the iterates of Algorithm 1 satisfy*

$$\mathbb{E}\left[\|\mathbf{u}_{t-i}(\mathbf{M}(t-i)) - \mathbf{u}_{t-i}(\mathbf{M}(t))\|_2^2\right] \le \eta^2 i^2 \frac{\sqrt{R_{\mathsf{grad},\infty} R_{y^{(s)},\mathsf{nat},\sup}}}{1-\rho_M} := \eta^2 i^2 R_{\mathbf{u},\mathsf{mem}}$$

Moreover, under Assumption 2', the above bound holds deterministically:

$$\mathbb{E}\left[\|\mathbf{u}_{t-i}(\mathbf{M}(t-i)) - \mathbf{u}_{t-i}(\mathbf{M}(t))\|_2^2\right] \le \eta^2 i^2 R_{\mathbf{u},\mathsf{mem}}^2$$

Proof. Now, we can bound the inner-most term above as follows:

$$\begin{aligned}
&\|\sum_{s\in\mathcal{S}}\sum_{j=0}^{H}(M_j^{(s)}(t-i) - M_j^{(s)}(t))\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2 \\
&\le \sum_{s\in\mathcal{S}}\sum_{j=0}^{H}\|M_j^{(s)}(t-i) - M_j^{(s)}(t)\|_{\mathsf{op}}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2 \\
&= \sum_{s\in\mathcal{S}}\sum_{j=0}^{H}\frac{1}{\sqrt{\rho_M^j}}\|M_j^{(s)}(t-i) - M_j^{(s)}(t)\|_{\mathsf{op}}\sqrt{\rho_M^j}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2 \\
&\le \sum_{s\in\mathcal{S}}\left(\sum_{j=0}^{H}\frac{1}{\rho_M^j}\|M_j^{(s)}(t-i) - M_j^{(s)}(t)\|_{\mathsf{op}}^2\right)^{1/2}\left(\sum_{j=0}^{H}\rho_M^j\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^2\right)^{1/2} \\
&\le \|\mathbf{M}(t-i) - \mathbf{M}(t)\|_{\rho_M^{-1/2},1}\left(\sum_{j=0}^{H}\rho_M^j\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^2\right)^{1/2},
\end{aligned}$$

where $\|\cdot\|_{\rho,1}$ is the ρ-reweighted $(2,1)$-group norm (see Definition E.3). Therefore, by Lemma H.2,

$$\begin{aligned}
\|\mathbf{M}(t-i) - \mathbf{M}(t)\|_{\rho_M^{-1/2},1} &= \|\sum_{\ell=0}^{i-1}\mathbf{M}(t-\ell) - \mathbf{M}(t-\ell+1)\|_{\rho_M^{-1/2},1} \\
&= \sum_{\ell=0}^{i-1}\|\mathbf{M}(t-\ell) - \mathbf{M}(t-\ell+1)\|_{\rho_M^{-1/2},1} \\
&= \sum_{\ell=0}^{i-1}\eta\|\nabla\widehat{\ell}_{t-i}(\mathbf{M}(t-i))\|_{\sqrt{\rho_M},\infty}.
\end{aligned}$$

By the above bounds, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\|\sum_{s\in\mathcal{S}}\sum_{j=0}^{H}(M_j^{(s)}(t-i) - M_j^{(s)}(t))\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^2\right] \\
&\le \eta^2\mathbb{E}\left[\left(\sum_{\ell=0}^{i-1}\|\nabla\widehat{\ell}_{t-i}(\mathbf{M}(t-i))\|_{\sqrt{\rho_M},\infty}\right)^2\left(\sum_{j=0}^{H}\rho_M^j\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^2\right)\right] \\
&\le \eta^2 i\sum_{\ell=0}^{i-1}\sum_{j=0}^{H}\rho_M^j\mathbb{E}\left[\|\nabla\widehat{\ell}_{t-i}(\mathbf{M}(t-i))\|_{\sqrt{\rho_M},\infty}^2\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^2\right] \\
&\le \eta^2 i\sum_{\ell=0}^{i-1}\sum_{j=0}^{H}\rho_M^j\sqrt{\mathbb{E}\left[\|\nabla\widehat{\ell}_{t-i}(\mathbf{M}(t-i))\|_{\sqrt{\rho_M},\infty}^4\right]\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^4\right]}.
\end{aligned}$$

Therefore, using Lemmas G.1 and G.4, we obtain:

$$\mathbb{E}\left[\|\sum_{s\in\mathcal{S}}\sum_{j=0}^{H}(M_j^{(s)}(t-i)-M_j^{(s)}(t))\mathbf{y}_{t-i-j}^{\mathsf{nat},(s)}\|_2^2\right] \le \eta^2 i^2 \frac{\sqrt{R_{\mathsf{grad},\infty}R_{y^{(s)},\mathsf{nat},\mathrm{sup}}}}{1-\rho_M},$$

as claimed.

Note that, in the adversarial case, under Assumption 2', the bounds on the gradient and $\mathbf{y}^{\mathsf{nat},(s)}$ hold deterministically. Thus, exactly the same arguments given above go through without taking expectations. □

Lemma H.5 (State estimate memory bound)**.** *Under Assumption 2, the error between the Kalman filter estimate of Algorithm 1's state at time t, and the counterfactual state estimate under the algorithm's current $\mathbf{M}(t)$ is bounded as*

$$\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t}-\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2^2\right] \le \eta^2 R_{\mathbf{u},\mathsf{mem}}\frac{R_A^2R_B^2}{1-\rho_A}\frac{1+\rho_A}{(1-\rho_A)^3} := \eta^2 R_{\widehat{\mathbf{x}},\mathsf{mem}}^2$$

Moreover, under Assumption 2', the error of any state estimate satisfying Assumption 4' is

$$\|\widehat{\mathbf{x}}_{t|t}-\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2 \le \eta\frac{R_XR_YR_CR_AR_BR_{\mathbf{u},\mathsf{mem}}(1-\rho_A\rho_X)}{(1-\rho_A)^2(1-\rho_X)^2} := \eta R_{\widehat{\mathbf{x}},\mathsf{mem}}$$

Proof. We begin with the stochastic case, where $\widehat{\mathbf{x}}_{t|t}$ denotes the Kalman filter Definition B.5. We notice the following: First, by Lemma C.1,

$$\begin{aligned}\widehat{\mathbf{x}}_{t|t}-\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)) &= \Delta_t-\Delta_t(\mathbf{M}(t))\\ &= \sum_{i=1}^{t}A^{i-1}B(\mathbf{u}_{t-i}(\mathbf{M}(t-i))-\mathbf{u}_{t-i}(\mathbf{M}(t)))\end{aligned}$$

Therefore, by Lemma H.4 and the above decomposition,

$$\begin{aligned}\mathbb{E}\left[\|\widehat{\mathbf{x}}_{t|t}-\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2^2\right] &\le \frac{R_A^2R_B^2}{1-\rho_A}\sum_{i=1}^{t}\rho_A^{i-1}\mathbb{E}\left[\|\mathbf{u}_{t-i}(\mathbf{M}(t-i))-\mathbf{u}_{t-i}(\mathbf{M})\|_2^2\right]\\ &\le \eta^2R_{\mathbf{u},\mathsf{mem}}\frac{R_A^2R_B^2}{1-\rho_A}\sum_{i=1}^{t}i^2\rho_A^{i-1}\\ &\le \eta^2R_{\mathbf{u},\mathsf{mem}}\frac{R_A^2R_B^2}{1-\rho_A}\frac{1+\rho_A}{(1-\rho_A)^3},\end{aligned}$$

as claimed.

In the adversarial case, we consider the state estimator from Assumption 4', parametrized by estimation matrices X, Y. By definition,

$$\widehat{\mathbf{x}}_{t|t}-\widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t)) = \sum_{i=0}^{t}X^iY(\mathbf{y}_{t-i}-\mathbf{y}_{t-i}(\mathbf{M}(t))).$$

By Lemma C.1,

$$\begin{aligned}\mathbf{y}_{t-i}-\mathbf{y}_{t-i}(\mathbf{M}(t)) &= C(\mathbf{x}_{t-i}-\mathbf{x}_{t-i}(\mathbf{M}(t)))\\ &= C(\Delta_{t-i}-\Delta_{t-i}(\mathbf{M}(t)))\\ &= \sum_{\ell=1}^{t-i}CA^{\ell-1}B(\mathbf{u}_{t-i-\ell}(\mathbf{M}(t-i-\ell))-\mathbf{u}_{t-i-\ell}(\mathbf{M}(t))).\end{aligned}$$

Therefore, by Lemma H.4, and by Assumption 4',

$$\begin{aligned}\|\widehat{\mathbf{x}}_{t|t} - \widehat{\mathbf{x}}_{t|t}(\mathbf{M}(t))\|_2 &\le R_X R_Y \sum_{i=0}^{t} \rho_X^i \|\mathbf{y}_{t-i} - \mathbf{y}_{t-i}(\mathbf{M}(t))\|_2 \\ &\le R_X R_Y R_C R_A R_B \sum_{i=0}^{t} \rho_X^i \sum_{\ell=1}^{t-i} \rho_A^{\ell-1} \|\mathbf{u}_{t-i-\ell}(\mathbf{M}(t-i-\ell)) - \mathbf{u}_{t-i-\ell}(\mathbf{M}(t))\|_2 \\ &\le \eta R_X R_Y R_C R_A R_B R_{\mathbf{u},\mathsf{mem}} \sum_{i=0}^{t} \rho_X^i \sum_{\ell=1}^{t-i} (i+\ell)\rho_A^{\ell-1} \\ &\le \eta \frac{R_X R_Y R_C R_A R_B R_{\mathbf{u},\mathsf{mem}}(1-\rho_A\rho_X)}{(1-\rho_A)^2(1-\rho_X)^2},\end{aligned}$$

as claimed. □

H.4 Regret on estimated/hypothetical loss function

Lemma H.6 (FTRL Regret)**.** *Let $\mathbf{M}(t)$ be the iterates of Algorithm 1 with regularizer $\Psi(\mathbf{M})$ defined as in (Reg). Suppose the cost functions c_t satisfy Assumption 3 or 3'. Let $\widehat{\ell}_t(\mathbf{M}) = c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}), \mathbf{u}_t(\mathbf{M}))$ denote the hypothetical cost that would be incurred, had the algorithm played control $\mathbf{u}_\ell(\mathbf{M})$ for all time steps $\ell \le t$, and had the resulting state estimate $\widehat{\mathbf{x}}_{t|t}(\mathbf{M})$ been the true state $\mathbf{x}_t(\mathbf{M})$. Then, for any $\mathbf{M}^* \in \mathcal{M}(R_M, \rho_M, H)$, the regret relative to the hypothetical losses $\widehat{\ell}_t(\cdot)$ is:*

$$\begin{aligned}\sum_{t=0}^{n} \widehat{\ell}_t(\mathbf{M}(t)) - \sum_{t=0}^{n} \widehat{\ell}_t(\mathbf{M}^*) &\le \Psi(\mathbf{M}^*) + \frac{\eta}{2}\sum_{t=1}^{n} \|\nabla \widehat{\ell}_t(\mathbf{M}(t))\|^2_{\sqrt{\rho_M},\infty} \\ &\le \frac{2e \min\{d_{y^s,\max}, d_{\mathbf{u}}\} R_M^2 \max\{1, \ln(|\mathcal{S}|)\}}{(1-\sqrt{\rho_M})^2 \eta} \\ &\quad + \frac{\eta}{2}\sum_{t=0}^{n} \|\nabla \widehat{\ell}_t(\mathbf{M}(t))\|^2_{\sqrt{\rho_M},\infty}\end{aligned}$$

Proof. First, notice that $\widehat{\ell}_t(\mathbf{M})$ is convex. Indeed, this follows from the fact that $\mathbf{u}_t(\mathbf{M})$, and hence also $\widehat{\mathbf{x}}_{t|t}(\mathbf{M})$ is linear in $\mathbf{M}$ by construction. Thus, since $c_t(\mathbf{x}, \mathbf{u})$ is convex, and the composition of a convex function with a linear function preserves convexity, it follows that $\widehat{\ell}_t(\mathbf{M})$ is convex. Thus, we can apply Theorem F.1, using the fact that, by Corollary E.2 and definition of $\Psi(\mathbf{M})$ in (Reg), $\Psi(\mathbf{M})$ is $1/\eta$-strongly convex with respect to $\|\cdot\|_{\rho_M^{-1/2},1}$, where $(\|\cdot\|_{\rho_M^{-1/2},1})_* = \|\cdot\|_{\sqrt{\rho_M},\infty}$. The result follows immediately.

The second bound follows from the first, noting that, by definition of the regularizer in (Reg), together with Lemma G.5, we have that, since $\mathbf{M}^* \in \mathcal{M}(R_M, \rho_M, H)$,

$$\Psi(\mathbf{M}^*) = \frac{e\max\{1, \ln(|\mathcal{S}|)\}}{\eta}\|\mathbf{M}^*\|^2_{1/\sqrt{\rho_M},1} \le \frac{e\max\{1, \ln(|\mathcal{S}|)\}}{\eta} \min\{d_{y^s,\max}, d_{\mathbf{u}}\} \frac{R_M^2}{(1-\sqrt{\rho_M})^2}.$$

□

H.5 Error due to unobserved state

Lemma H.7 (Rewriting the unobserved state error bound)**.** *Suppose that* $\mathbf{x}_0$, $(\mathbf{w}_t, \mathbf{v}_t)_{t\geq 0}$ *are independent with bounded second moments. Then,*

$$\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^{n} \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} - \mathbf{x}^{\mathsf{nat}}_t)\right]$$
$$\leq \sqrt{\min\{d_{\mathbf{u}}, d_{y^s}\}} R_M R_A R_B \sum_{i=1}^{n} \rho_A^{i-1} \sum_{j=0}^{H} \rho_M^j \mathbb{E}\left[\max_{s\in\mathcal{S}} \| \sum_{t=i+j}^{n} S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top \|_F\right].$$

Proof. First, observe that, by Lemma C.1,

$$\widehat{\mathbf{x}}_{t|t}(\mathbf{M}) = \widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} + \Delta_t(\mathbf{M}), \quad \text{where} \quad \Delta_t(\mathbf{M}) = \sum_{i=1}^{t} A^{i-1} B \mathbf{u}_{t-i}(\mathbf{M}) = \sum_{i=1}^{t} A^{i-1} B \sum_{s\in\mathcal{S}} \sum_{j=0}^{H} M_j^{(s)} \mathbf{y}^{\mathsf{nat}}_{t-i-j}.$$

Thus, by the Orthogonality principle Lemma B.7, and since $\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t}$ does not depend on $\mathbf{M}$, we have that, taking $\mathbf{e}_t = \widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} - \mathbf{x}^{\mathsf{nat}}_t$,

$$\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^{n} \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t \mathbf{e}_t\right] = \mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)(H,R)} \sum_{t=1}^{n} \Delta_t(\mathbf{M})^\top S_t \mathbf{e}_t\right]$$

Now, using the definition of $\Delta_t(\mathbf{M})$, we can rearrange the above as:

$$\begin{aligned}\sum_{t=1}^{n} \Delta_t(\mathbf{M})^\top S_t \mathbf{e}_t &= \sum_{t=1}^{n} \sum_{i=1}^{t} \sum_{j=0}^{H} \sum_{s\in\mathcal{S}} (A^{i-1} B M_j^{(s)} \mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top S_t \mathbf{e}_t \\ &= \sum_{i=1}^{n} \sum_{j=0}^{H} \sum_{s\in\mathcal{S}} \sum_{t=i}^{n} \left\langle M_j^{(s)} \mathbf{y}^{\mathsf{nat},(s)}_{t-j-i}, (A^{i-1}B)^\top S_t \mathbf{e}_t \right\rangle\end{aligned}$$

Thus, by the cyclic property of trace, and Hölder's inequality for Schatten norms, we can bound the inner expression above as:

$$\begin{aligned}\sum_{t=i}^{n} \left\langle M_j^{(s)} \mathbf{y}^{\mathsf{nat},(s)}_{t-j-i}, (A^{i-1}B)^\top S_t \mathbf{e}_t \right\rangle &= \sum_{t=i}^{n} \mathrm{tr}((M_j^{(s)} \mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top (A^{i-1}B)^\top S_t \mathbf{e}_t) \\ &= \mathrm{tr}((M_j^{(s)})^\top (A^{i-1}B)^\top S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top) \\ &= \left\langle M_j^{(s)}, (A^{i-1}B)^\top \sum_{t=i}^{n} S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top \right\rangle_F \\ &\leq \|M_j^{(s)}\|_F \|(A^{i-1}B)^\top S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top\|_F \\ &\leq \|M_j^{(s)}\|_F \|A^{i-1}B\|_{\mathsf{op}} \|\sum_{t=i}^{n} S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top\|_F \\ &\leq \sqrt{\min\{d_{\mathbf{u}}, d_{\mathbf{y}}\}} \|M_j\|_{\mathsf{op}} \|A^{i-1}B\|_{\mathsf{op}} \|\sum_{t=i}^{n} S_t \mathbf{e}_t (\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top\|_F,\end{aligned}$$

where $\langle \cdot, \cdot \rangle_F$ denotes the Frobenius inner product. Note that we used the facts that $\|X^\top Y\|_F \leq \|X^\top\|_{\mathsf{op}} \|Y\|_F = \|X\|_{\mathsf{op}} \|Y\|_F$ for any two real matrices X, Y of appropriate dimensions, and $\|X\|_F \leq \sqrt{\min\{m,n\}} \|X\|_{\mathsf{op}}$ for an $m \times n$ matrix X. Now, recall by definition of $\mathcal{M}(R_M, \rho_M, H)$, $\sum_{s\in\mathcal{S}} \|M_j^{(s)}\|_{\mathsf{op}} \leq R_M \rho_M^j$ and by Assumption 1, $\|A^{i-1}B\|_{\mathsf{op}} \leq R_A R_B \rho_A^{i-1}$. Thus, collecting the

above bounds, we obtain:

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)}\sum_{t=1}^{n}\widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t\mathbf{e}_t\right]\\
&\le\sqrt{\min\{d_{\mathbf{u}},d_{\mathbf{y}}\}}R_MR_AR_B\sum_{i=1}^{n}\rho_A^{i-1}\sum_{j=0}^{H}\rho_M^j\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\sum_{t=i}^{n}S_t\mathbf{e}_t(\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top\|_F\right]\\
&=\sqrt{\min\{d_{\mathbf{u}},d_{\mathbf{y}}\}}R_MR_AR_B\sum_{i=1}^{n}\rho_A^{i-1}\sum_{j=0}^{H}\rho_M^j\mathbb{E}\left[\max_{s\in\mathcal{S}}\|\sum_{t=i+j}^{n}S_t\mathbf{e}_t(\mathbf{y}^{\mathsf{nat},(s)}_{t-j-i})^\top\|_F\right],
\end{aligned}$$

where, in the last line, we used the fact that $\mathbf{y}^{\mathsf{nat}}_\ell = 0$ for $\ell < 0$. Thus, we obtain the claimed bound. □

Lemma H.8 (Error bound)**.** *Suppose that* $\mathbf{x}_0$, $(\mathbf{w}_t,\mathbf{v}_t)_{t\ge0}$ *are independent sub-Gaussian random vectors, and* $\mathbf{w}_t,\mathbf{v}_t$ *are each mean-zero. Then, denoting* $D_s=\sum_{t=i}^{n}S_t\mathbf{e}_t(\mathbf{y}^{\mathsf{nat},(s)}_{t-i-j})^\top$,

$$\mathbb{E}\left[\max_{s\in\mathcal{S}}\|D_s\|_F\right]\lesssim(d_{\mathbf{x}}+d_{y^s,\max})\sqrt{\min\{d_{\mathbf{x}},d_{y^s,\max}\}}\sqrt{d_{y^s,\max}}(\log(|\mathcal{S}|{+}1))^{3/2}\sqrt{n(i+j)}.$$

In particular, combined with Lemma H.7, this implies that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)}\sum_{t=1}^{n}\widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t}-\mathbf{x}^{\mathsf{nat}}_t)\right]\\
&\lesssim\sqrt{\min\{d_{\mathbf{u}},d_{y^s,\max}\}}(d_{\mathbf{x}}+d_{y^s,\max})\sqrt{\min\{d_{\mathbf{x}},d_{y^s,\max}\}}\sqrt{d_{y^s,\max}}(\log(|\mathcal{S}|{+}1))^{3/2}\sqrt{n}.
\end{aligned}$$

Proof. To begin, let us introduce some simplifying notation: Let $\Pi_s\in\mathbb{R}^{d_{y^s}\times d_{\mathbf{y}}}$ denote the projection matrix satisfying $\mathbf{y}^{\mathsf{nat},(s)}=\Pi_s\mathbf{y}^{\mathsf{nat}}$. Then

$$\bar{\mathbf{e}}_t=S_t\mathbf{e}_t\quad\text{and}\quad t'=t-i-j\quad\text{and}\quad D_s=\sum_{t=i}^{n}S_t\mathbf{e}_t(\mathbf{y}^{\mathsf{nat},(s)}_{t-i-j})^\top=\sum_{t=i}^{n}\bar{\mathbf{e}}_t(\mathbf{y}^{\mathsf{nat}}_{t'})^T\Pi_s^\top.$$

To prove the claim, we will show that $\|D_s\|_F$ is sub-Exponential (see Definition E.7). Now, recall that

$$\|D_s\|_F\le\sqrt{\min\{d_{\mathbf{x}},d_{y^s}\}}\|D_s\|_{\mathsf{op}}.$$

Now, by Lemma E.15, there exists nets $\mathcal{N}_{\mathbf{x}},\mathcal{N}_{\mathbf{y}^{(s)}}$ of $\{a\in\mathbb{R}^{d_{\mathbf{x}}}:\|a\|_2=1\}$ and $\{b\in\mathbb{R}^{d_{y^s}}:\|b\|_2=1\}$ respectively, such that $|\mathcal{N}_{d_{\mathbf{x}}}|=9^{d_{\mathbf{x}}}$, $|\mathcal{N}_{d_{y^s}}|=9^{d_{y^s}}$, such that

$$\|D_s\|_{\mathsf{op}}\le2\max_{a\in\mathcal{N}_{\mathbf{x}},b\in\mathcal{N}_{\mathbf{y}^{(s)}}}\langle a,D_sb\rangle.$$

Thus, by Lemma E.12, there exists an absolute constant $C>0$ such that

$$\|\max_{s\in\mathcal{S}}\|D_s\|_F\|_{\psi_1}\le C(d_{\mathbf{x}}+d_{y^s,\max})\sqrt{\min\{d_{\mathbf{x}},d_{y^s,\max}\}}\log(|\mathcal{S}|{+}1)\max_{s\in\mathcal{S},a\in\mathcal{N}_{d_{\mathbf{x}}},b\in\mathcal{N}_{d_{y^s}}}\|\langle a,D_sb\rangle\|_{\psi_1}.\tag{5}$$

Recalling that, by Lemma E.8, there exists an absolute constant C' such that

$$\mathbb{E}\left[\max_{s\in\mathcal{S}}\|D_s\|_F\right]\le C'\|\max_{s\in\mathcal{S}}\|D_s\|_F\|_{\psi_1}.\tag{6}$$

So, to conclude, it suffices to upper-bound $\|\langle a,D_sb\rangle\|_{\psi_1}$ for an arbitrary, fixed $s\in\mathcal{S}$ and $a\in\mathbb{R}^{d_{\mathbf{x}}},b\in\mathbb{R}^{d_{y^s}}$ with $\|a\|_2=1=\|b\|_2$.

Now, by Lemma B.11, we can decompose $\mathbf{e}_t(\mathbf{y}^{\mathsf{nat}}_{t'})^\top$ as:

$$\mathbf{e}_t(\mathbf{y}^{\mathsf{nat}}_{t'})^\top=(T_{1,t})+(T_{2,t})+(T_{3,t}),$$

where

$$
\begin{aligned}
(T_{1,t}) &= \Phi^{[t]}\mathbf{e}_0^-\mu_0^\top(CA^{t'})^\top + \Phi^{[t]}(\mathbf{e}_0^-(\mathbf{e}_0^-)^\top - X_0)(CA^{t'})^\top \\
&\quad + \Phi^{[t]}\mathbf{e}_0^-(\mathbf{v}_{t'} + \sum_{k=1}^{t'} CA^{k-1}\mathbf{w}_{t'-k})^\top - \sum_{k=1}^{t+1}\widetilde{\Phi}^{[t-k+1,t]}\tilde{L}_{t-k+1}\mathbf{v}_{t-k+1}\mathbf{x}_0^\top(CA^{t'})^\top \\
&\quad + \sum_{k=1}^{t}\Phi^{[t-k+1,t]}\mathbf{w}_{t-k}\mathbf{x}_0^\top(CA^{t'})^\top \\
(T_{2,t}) &= -\widetilde{\Phi}^{[t',t]}\tilde{L}_{t'}(\mathbf{v}_{t'}\mathbf{v}_{t'}^\top - V) + \sum_{k=1}^{t'}\Phi^{[t'-k+1,t]}(\mathbf{w}_{t'-k}\mathbf{w}_{t'-k}^\top - W)(CA^{k-1})^\top \\
&\quad - \sum_{k=1}^{t+1}\sum_{k'=1}^{t'}\widetilde{\Phi}^{[t-k+1,t]}\tilde{L}_{t-k+1}\mathbf{v}_{t-k+1}\mathbf{w}_{t'-k'}^\top(CA^{k'-1})^\top + \sum_{k=1}^{t}\Phi^{[t-k+1,t]}\mathbf{w}_{t-k}\mathbf{v}_{t'}^\top \\
(T_{3,t}) &= -\sum_{\substack{k=1 \\ t-k+1\neq t'}}^{t+1}\widetilde{\Phi}^{[t-k+1,t]}\tilde{L}_{t-k+1}\mathbf{v}_{t-k+1}\mathbf{v}_{t'}^\top + \sum_{k=1}^{t}\sum_{\substack{k'=1 \\ t-k\neq t'-k'}}^{t'}\Phi^{[t-k+1,t]}\mathbf{w}_{t-k}\mathbf{w}_{t'-k'}^\top(CA^{k'-1})^\top
\end{aligned}
$$

Note that the terms in $(T_{1,t})$ decay geometrically in t, so their sum can be bounded by a constant. The terms in $(T_{2,t})$ do not decay with t, but $\sum_{t=i+j}^{n}(T_{2,t})$ involves sums of independent random variables, so we can use Lemma E.10 to bound these terms by $\mathcal{O}(\sqrt{n-(i+j)+1})$. Finally, the terms in $(T_{3,t})$ also do not decay with t, and $\sum_{t=i+j}^{n}(T_{3,t})$ is not a sum of independent variables, due to the overlapping terms. However, we can further decompose this term into $\mathcal{O}(i+j+1)$ sums of independent random variables, which suffices for our argument.

Therefore, by the triangle inequality for ψ_1-norms, and using the fact that $\Pi_s C = C^{(s)}$ and $\Pi_s\mathbf{v} = \mathbf{v}^{(s)}$ by definition, we can decompose

$$
\|\langle a, D_s b\rangle\|_{\psi_1} = \|\sum_{t=i+1}^{n} a^\top S_t\mathbf{e}_t(\mathbf{y}_{t'}^{\mathsf{nat}})^\top\Pi_s b\|_{\psi_1} \leq (T_1') + (T_2') + (T_3'),
$$

where

$$
\begin{aligned}
(T_1') = &\sum_{t=i+j}^{n} \|a^\top \Phi^{[t]} \mathbf{e}_0^- \mu_0^\top (C^{(s)} A^{t'})^\top b\|_{\psi_1} + \sum_{t=i+j}^{n} \|a^\top \Phi^{[t]} (\mathbf{e}_0^- (\mathbf{e}_0^-)^\top - X_0)(C^{(s)} A^{t'})^\top b\|_{\psi_1} \\
&+ \sum_{t=i+j}^{n} \|a^\top \Phi^{[t]} \mathbf{e}_0^- (\mathbf{v}_{t'}^{(s)})^\top b\|_{\psi_1} + \sum_{t=i+j}^{n} \sum_{k=1}^{t'} \|a^\top \Phi^{[t]} \mathbf{e}_0^- \mathbf{w}_{t'-k}^\top (C^{(s)} A^{k-1})^\top b\|_{\psi_1} \\
&+ \sum_{t=i+j}^{n} \sum_{k=1}^{t+1} \|a^\top \widetilde{\Phi}^{[t-k+1,t]} \tilde{L}_{t-k+1} \mathbf{v}_{t-k+1} \mathbf{x}_0 (C^{(s)} A^{t'})^\top\|_{\psi_1} \\
&+ \sum_{t=i+j}^{n} \sum_{k=1}^{t} \|a^\top \Phi^{[t-k+1,t]} \mathbf{w}_{t-k} \mathbf{x}_0^\top (C^{(s)} A^{t'})^\top\|_{\psi_1} \\
(T_2') = &\| \sum_{t=i+j}^{n} a^\top \widetilde{\Phi}^{[t',t]} \tilde{L}_{t'} (\mathbf{v}_{t'} \mathbf{v}_{t'}^\top - V) \Pi_s^\top b\|_{\psi_1} \\
&+ \sum_{k=1}^{n-i-j} \| \sum_{t=k+i+j}^{n} a^\top \Phi^{[t'-k+1,t]} (\mathbf{w}_{t'-k} \mathbf{w}_{t'-k}^\top - W)(C^{(s)} A^{k-1})^\top b\|_{\psi_1} \\
&+ \sum_{k=1}^{n+1} \sum_{k'=1}^{n-i-j} \| \sum_{t=\max\{k-1,k'+i+j\}}^{n} a^\top \widetilde{\Phi}^{[t-k+1,t]} \tilde{L}_{t-k+1} \mathbf{v}_{t-k+1} \mathbf{w}_{t'-k'} (C^{(s)} A^{k'-1})^\top b\|_{\psi_1} \\
&+ \sum_{k=1}^{n} \| \sum_{t=\max\{i+j,k\}}^{n} a^\top \Phi^{[t-k+1,t]} \mathbf{w}_{t-k} (\mathbf{v}_{t'}^{(s)})^\top b\|_{\psi_1} \\
(T_3') = &\sum_{k=1}^{n+1} \| \sum_{\substack{t=\max\{i+j,k-1\} \\ t-k+1 \neq t'}}^{n} a^\top \widetilde{\Phi}^{[t-k+1,t]} \tilde{L}_{t-k+1} \mathbf{v}_{t-k+1} (\mathbf{v}_{t'}^{(s)})^\top b\|_{\psi_1} \\
&+ \sum_{k=1}^{n} \sum_{k'=1}^{n-i-j} \| \sum_{\substack{t=\max\{k,k'+i+j\} \\ t-k \neq t'-k'}} a^\top \Phi^{[t-k+1,t]} \mathbf{w}_{t-k} \mathbf{w}_{t'-k'}^\top (C^{(s)} A^{k'-1})^\top b\|_{\psi_1}.
\end{aligned}
$$

We now bound each of these terms. For (T_1'), using Lemmas E.11 and E.13, we use the fact that, for any sub-Gaussian vectors X, Y, $\|a^\top X Y b\|_{\psi_1} \leq \|a^\top X\|_{\psi_2} \|b^\top Y\|_{\psi_2} \leq \|a\|_2 \|b\|_2 \|X\|_{\psi_2} \|Y\|_{\psi_2}$. Hence, combining this with the fact that $\|Ma\|_2 \leq \|M\|_{\mathsf{op}} \|a\|_2$ and $\|a\|_2 = 1 = \|b\|_2$,

$$
\begin{aligned}
(T_1') \leq &\sum_{t=i+j}^{n} \|\Phi^{[t]}\|_{\mathsf{op}} \|C^{(s)} A^{t'}\|_{\mathsf{op}} \|\mathbf{e}_0^-\|_{\psi_2} \|\mu_0\|_2 + \sum_{t=i+j}^{n} \|\Phi^{[t]}\|_{\mathsf{op}} \|C^{(s)} A^{t'}\|_{\mathsf{op}} (\|\mathbf{e}_0^-\|_{\psi_2}^2 + \|X_0\|_{\mathsf{op}}) \\
&+ \sum_{t=i+j}^{n} \|\Phi^{[t]}\|_{\mathsf{op}} \|\mathbf{e}_0^-\|_{\psi_2} \|\mathbf{v}_{t'}^{(s)}\|_{\psi_2} + \sum_{t=i+j}^{n} \sum_{k=1}^{t'} \|\Phi^{[t]}\|_{\mathsf{op}} \|C^{(s)} A^{k-1}\|_{\mathsf{op}} \|\mathbf{e}_0^-\|_{\psi_2} \|\mathbf{w}_{t'-k}\|_{\psi_2} \\
&+ \sum_{t=i+j}^{n} \sum_{k=1}^{t+1} \|\widetilde{\Phi}^{[t-k+1,t]}\|_{\mathsf{op}} \|C^{(s)} A^{t'}\|_{\mathsf{op}} \|\tilde{L}_{t-k+1} \mathbf{v}_{t-k+1}\|_{\psi_2} (\|\mathbf{e}_0^-\|_{\psi_2} + \|\mu_0\|_2) \\
&+ \sum_{t=i+j}^{n} \sum_{k=1}^{t} \|\Phi^{[t-k+1,t]}\|_{\mathsf{op}} \|C^{(s)} A^{t'}\|_{\mathsf{op}} \|\mathbf{w}_{t-k}\|_{\psi_2} (\|\mathbf{e}_0^-\|_{\psi_2} + \|\mu_0\|_2).
\end{aligned}
$$

Now, using Lemmas E.12 and G.2, there exists an absolute constant $c > 0$ such that

$$
\begin{aligned}
\|\tilde{L}_t \mathbf{v}_t\|_{\psi_2} \leq \|\|\tilde{L}_t \mathbf{v}_t\|_2\|_{\psi_2} &\leq R_{\tilde{L}v} \|\max_{s \in \mathcal{S}} \|V_s^{-1/2} \mathbf{v}_t^{(s)}\|_2\|_{\psi_2} \\
&\leq c R_{\tilde{L}v} \sqrt{d_{y^s,\max}} \sqrt{\log(|\mathcal{S}|+1)} \max_{s \in \mathcal{S}} \|V_s^{-1/2} \mathbf{v}_t^{(s)}\|_{\psi_2},
\end{aligned}
$$

so

$$(T_1') \leq R_\Phi R_A R_{C^{(s)}} \rho_\Phi^{i+j} \frac{\|\mathbf{e}_0^-\|_{\psi_2}\|\mu_0\|_2 + \|\mathbf{e}_0^-\|_{\psi_2}^2 + \|X_0\|_{\mathsf{op}}}{1-\rho_A\rho_\Phi}$$
$$+ R_\Phi \frac{\|\mathbf{e}_0^-\|_{\psi_2}}{1-\rho_\Phi}\left(\|\mathbf{v}^{(s)}\|_{\psi_2} + R_{C^{(s)}} R_A \frac{\|\mathbf{w}\|_{\psi_2}}{1-\rho_A}\right)$$
$$+ \frac{R_{\widetilde{\Phi}} R_A R_{C^{(s)}}}{(1-\rho_A)(1-\rho_{\widetilde{\Phi}})}(\|\mathbf{e}_0^-\|_{\psi_2} + \|\mu_0\|_2)\left(cR_{\tilde{L}v}\sqrt{d_{y^s,\max}}\sqrt{\log(|\mathcal{S}|+1)}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}^{(s)}\|_{\psi_2} + \|\mathbf{w}\|_{\psi_2}\right).$$

To bound (T_2'), since each term is a sum of mean-zero, independent random variables, we may apply Lemma E.10, together with the arguments from above, and the fact that, by Proposition B.9, $\tilde{L}_{t'} = (\Sigma^+_{t'|t'-1} + P_{t'}C^\top V^{-1} C P_{t'})^+ P_{t'} C^\top V^{-1}$, to obtain:

$$(T_2') \leq R_{\psi_1}\sqrt{\sum_{t=i+j}^{n}\|\widetilde{\Phi}^{[t',t]}\|_{\mathsf{op}}^2(\|\tilde{L}_{t'}\mathbf{v}_{t'}\|_{\psi_2}\|\mathbf{v}^{(s)}\|_{\psi_2} + \|(\Sigma^+_{t'|t'-1} + P_{t'}C^\top V^{-1}C^\top P_{t'})^+ P_{t'}(C^{(s)})^\top\|_{\mathsf{op}})^2}$$
$$+ \sum_{k=1}^{n-i-j}\sqrt{\sum_{t=k+i+j}^{n}\|\Phi^{[t'-k+1,t]}\|_{\mathsf{op}}^2\|C^{(s)}A^{k-1}\|_{\mathsf{op}}^2(\|\mathbf{w}_{t'-k}\|_{\psi_2}^2 + \|W\|_{\mathsf{op}})^2}$$
$$+ \sum_{k=1}^{n+1}\sum_{k'=1}^{n-i-j}\sqrt{\sum_{t=\max\{k-1,k'+i+j\}}^{n}\|\widetilde{\Phi}^{[t-k+1,t]}\|_{\mathsf{op}}^2\|C^{(s)}A^{k'-1}\|_{\mathsf{op}}^2\|\tilde{L}_{t-k+1}\mathbf{v}_{t-k+1}\|_{\psi_2}^2\|\mathbf{w}\|_{\psi_2}^2}$$
$$+ \sum_{k=1}^{n}\sqrt{\sum_{t=\max\{i+j,k\}}^{n}\|\Phi^{[t-k+1,t]}\|_{\mathsf{op}}^2\|\mathbf{w}_{t-k}\|_{\psi_2}^2\|\mathbf{v}_{t'}^{(s)}\|_{\psi_2}^2}$$

using the fact that $\|\Sigma_{t'|t'-1}\|_{\mathsf{op}} \leq R_{e^-}$ by Lemma B.13,

$$(T_2') \leq R_{\widetilde{\Phi}}\rho_{\widetilde{\Phi}}^{i+j} R_{\psi_1}(cR_{\tilde{L}v}\sqrt{d_{y^s,\max}}\sqrt{\log(|S|+1)}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}^{(s)}\|_{\psi_2}\|\mathbf{v}^{(s)}\|_{\psi_2} + R_{C^{(s)}}R_{e^-})\sqrt{n-(i+j)-1}$$
$$+ R_\Phi R_A R_{C^{(s)}}\frac{\rho_\Phi^{i+j}}{1-\rho_A\rho_\Phi}(\|\mathbf{w}\|_{\psi_2}^2 + \|W\|_{\mathsf{op}})\sqrt{n-(i+j)}$$
$$+ \frac{R_{\widetilde{\Phi}}R_{C^{(s)}}R_A}{(1-\rho_A)(1-\rho_{\widetilde{\Phi}})}cR_{\tilde{L}v}\sqrt{d_{y^s,\max}}\sqrt{\log(|\mathcal{S}|+1)}\max_{s\in\mathcal{S}}\|\mathbf{v}^{(s)}\|_{\psi_2}\|\mathbf{w}\|_{\psi_2}\sqrt{n-(i+j)}$$
$$+ \frac{R_\Phi\|\mathbf{w}\|_{\psi_2}\|\mathbf{v}^{(s)}\|_{\psi_2}}{1-\rho_\Phi}\sqrt{n-(i+j)+1}.$$

Finally, for (T_3'), we first split the terms as follows (recall that $t' = t - i - j$):

$$
\begin{aligned}
(T_3') \le & \sum_{k=1}^{\min\{i+j+1,n+1\}} \sum_{r=0}^{i+j-k+1} \Big\| \sum_{\substack{t=\max\{i+j,k-1\} \\ t-k+1\neq t' \\ t\equiv r \mod (i+j-k+2)}}^{n} a^\top \widetilde{\Phi}^{[t-k+1,t]} \tilde{L}_{t-k+1} \mathbf{v}_{t-k+1} (\mathbf{v}_{t'}^{(s)})^\top b \Big\|_{\psi_1} \\
& + \sum_{k=i+j+2}^{n+1} \sum_{r=0}^{k-1-i-j} \Big\| \sum_{\substack{t=\max\{i+j,k-1\} \\ t-k+1\neq t' \\ t\equiv r \mod (k-i-j)}}^{n} a^\top \widetilde{\Phi}^{[t-k+1,t]} \tilde{L}_{t-k+1} \mathbf{v}_{t-k+1} (\mathbf{v}_{t'}^{(s)})^\top b \Big\|_{\psi_1} \\
& + \sum_{k=1}^{n} \sum_{k'=\max\{k-i-j,1\}}^{n-i-j} \sum_{r=0}^{i+j+k'-k} \Big\| \sum_{\substack{t=\max\{k,k'+i+j\} \\ t-k\neq t'-k' \\ t\equiv r \mod (i+j+k'-k+1)}}^{n} \Phi^{[t-k+1,t]} \mathbf{w}_{t-k} \mathbf{w}_{t'-k'}^\top (C^{(s)} A^{k'-1})^\top b \Big\|_{\psi_1} \\
& + \sum_{k=1}^{n} \sum_{k'=1}^{k-i-j} \sum_{r=0}^{k-k'-i-j} \Big\| \sum_{\substack{t=\max\{k,k'+i+j\} \\ t-k\neq t'-k' \\ t\equiv r \mod (k-k'-i-j+1)}}^{n} \Phi^{[t-k+1,t]} \mathbf{w}_{t-k} \mathbf{w}_{t'-k'}^\top (C^{(s)} A^{k'-1})^\top b \Big\|_{\psi_1}.
\end{aligned}
$$

Notice that, after splitting the summation ranges, the inner terms are all now independent (and mean-zero). Therefore, we may apply Lemma E.10, and the similar arguments as above, to obtain

$$
\begin{aligned}
(T_3') \le & R_{\widetilde{\Phi}} \|\mathbf{v}^{(s)}\|_{\psi_2} c \sqrt{d_{y^s,\max}} \sqrt{\log(|\mathcal{S}|+1)} \max_{s\in\mathcal{S}} \|V_s^{-1/2} \mathbf{v}^{(s)}\|_{\psi_2} \sqrt{n+2} \sum_{k=1}^{\min\{i+j+1,n+1\}} \rho_{\widetilde{\Phi}}^{k-1} \sqrt{i+j-k+2} \\
& + R_{\widetilde{\Phi}} \|\mathbf{v}^{(s)}\|_{\psi_2} c \sqrt{d_{y^s,\max}} \sqrt{\log(|\mathcal{S}|+1)} \max_{s\in\mathcal{S}} \|V_s^{-1/2} \mathbf{v}^{(s)}\|_{\psi_2} \sqrt{n-i-j+1} \sum_{k=i+j+2}^{n+1} \rho_{\widetilde{\Phi}}^{k-1} \sqrt{k-i-j} \\
& + R_\Phi R_{C^{(s)}} R_A \|\mathbf{w}\|_{\psi_2}^2 \sqrt{n+1} \sum_{k=1}^{n} \rho_\Phi^{k-1} \sum_{k'=\max\{k-i-j,1\}}^{n-i-j} \rho_A^{k'-1} \sqrt{i+j+k'-k+1} \\
& + R_\Phi R_{C^{(s)}} R_A \|\mathbf{w}\|_{\psi_2}^2 \sqrt{n-i-j+1} \sum_{k=1}^{n} \rho_\Phi^{k-1} \sum_{k'=1}^{k-i-j} \rho_A^{k'-1} \sqrt{k-k'-i-j+1}.
\end{aligned}
$$

Thus, simplifying the above, we obtain

$$
\begin{aligned}
(T_3') \le & R_{\widetilde{\Phi}} \|\mathbf{v}^{(s)}\|_{\psi_2} c \sqrt{d_{y^s,\max}} \sqrt{\log(|\mathcal{S}|+1)} \max_{s\in\mathcal{S}} \|V_s^{-1/2} \mathbf{v}^{(s)}\|_{\psi_2} \sqrt{n+2} \Big(\frac{\sqrt{i+j+1}}{1-\rho_{\widetilde{\Phi}}} + \frac{\rho_{\widetilde{\Phi}}^{i+j+1}(2-\rho_{\widetilde{\Phi}})}{(1-\rho_{\widetilde{\Phi}})^2} \Big) \\
& + R_\Phi R_{C^{(s)}} R_A \|\mathbf{w}\|_{\psi_2}^2 \sqrt{n+1} \frac{2-\rho_A}{(1-\rho_A)^2(1-\rho_\Phi)}.
\end{aligned}
$$

Collecting our results, we have shown, for any $s \in \mathcal{S}$ and any $\|a\|_2 = 1 = \|b\|_2$,
$\|\langle a, D_s b\rangle\|_{\psi_1}$

$$
\begin{aligned}
&\le R_\Phi R_A R_{C^{(s)}} \rho_\Phi^{i+j} \frac{\|\mathbf{e}_0^-\|_{\psi_2}\|\mu_0\|_2 + \|\mathbf{e}_0^-\|_{\psi_2}^2 + \|X_0\|_{\mathsf{op}}}{1-\rho_A\rho_\Phi} \\
&\quad + R_\Phi \frac{\|\mathbf{e}_0^-\|_{\psi_2}}{1-\rho_\Phi}\left(\|\mathbf{v}^{(s)}\|_{\psi_2} + R_{C^{(s)}} R_A \frac{\|\mathbf{w}\|_{\psi_2}}{1-\rho_A}\right) \\
&\quad + \frac{R_{\widetilde{\Phi}} R_A R_{C^{(s)}}}{(1-\rho_A)(1-\rho_{\widetilde{\Phi}})}(\|\mathbf{e}_0^-\|_{\psi_2} + \|\mu_0\|_2)\left(cR_{\tilde{L}v}\sqrt{d_{y^s,\max}}\sqrt{\log(|\mathcal{S}|+1)}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}^{(s)}\|_{\psi_2} + \|\mathbf{w}\|_{\psi_2}\right) \\
&\quad + R_{\widetilde{\Phi}}\rho_{\widetilde{\Phi}}^{i+j} R_{\psi_1}(cR_{\tilde{L}v}\sqrt{d_{y^s,\max}}\sqrt{\log(|S|+1)}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}^{(s)}\|_{\psi_2}\|\mathbf{v}^{(s)}\|_{\psi_2} + R_{C^{(s)}}R_{e^-})\sqrt{n-(i+j)-1} \\
&\quad + R_\Phi R_A R_{C^{(s)}} \frac{\rho_\Phi^{i+j}}{1-\rho_A\rho_\Phi}(\|\mathbf{w}\|_{\psi_2}^2 + \|W\|_{\mathsf{op}})\sqrt{n-(i+j)} \\
&\quad + \frac{R_{\widetilde{\Phi}} R_{C^{(s)}} R_A}{(1-\rho_A)(1-\rho_{\widetilde{\Phi}})} cR_{\tilde{L}v}\sqrt{d_{y^s,\max}}\sqrt{\log(|\mathcal{S}|+1)}\max_{s\in\mathcal{S}}\|\mathbf{v}^{(s)}\|_{\psi_2}\|\mathbf{w}\|_{\psi_2}\sqrt{n-(i+j)} \\
&\quad + \frac{R_\Phi\|\mathbf{w}\|_{\psi_2}\|\mathbf{v}^{(s)}\|_{\psi_2}}{1-\rho_\Phi}\sqrt{n-(i+j)+1} \\
&\quad + R_{\widetilde{\Phi}}\|\mathbf{v}^{(s)}\|_{\psi_2} c\sqrt{d_{y^s,\max}}\sqrt{\log(|\mathcal{S}|+1)}\max_{s\in\mathcal{S}}\|V_s^{-1/2}\mathbf{v}^{(s)}\|_{\psi_2}\sqrt{n+2}\left(\frac{\sqrt{i+j+1}}{1-\rho_{\widetilde{\Phi}}} + \frac{\rho_{\widetilde{\Phi}}^{i+j+1}(2-\rho_{\widetilde{\Phi}})}{(1-\rho_{\widetilde{\Phi}})^2}\right) \\
&\quad + R_\Phi R_{C^{(s)}} R_A \|\mathbf{w}\|_{\psi_2}^2\sqrt{n+1}\frac{2-\rho_A}{(1-\rho_A)^2(1-\rho_\Phi)}.
\end{aligned}
$$

Combined with (5) and (6), we conclude that

$$
\mathbb{E}\left[\max_{s\in\mathcal{S}}\|D_s\|_F\right] \lesssim (d_{\mathbf{x}} + d_{y^s,\max})\sqrt{\min\{d_{\mathbf{x}}, d_{y^s,\max}\}}\sqrt{d_{y^s,\max}}(\log(|\mathcal{S}|+1))^{3/2}\sqrt{n(i+j)},
$$

as claimed.

Now, to conclude with the bound on $\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)}\sum_{t=1}^n \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}})\right]$, by Lemma H.7 and the above bound, it suffices to bound

$$
\sum_{i=1}^n \rho_A^{i-1}\sum_{j=0}^H \rho_M^j\sqrt{i+j}.
$$

Now, upper bounding $\sqrt{i+j} \le i+j$, we have that, for any $i \ge 1$,

$$
\sum_{j=0}^H \rho_M^j(i+j) \le \frac{1}{1-\rho_M} + (i-1) + \rho_M\sum_{j=0}^\infty (i+j)\rho_M^j = \frac{1}{1-\rho_M}\left(\frac{1}{1-\rho_M} + (i-1)\right).
$$

Therefore,

$$
\begin{aligned}
\sum_{i=1}^n \rho_A^{i-1}\sum_{j=0}^H \rho_M^j\sqrt{i+j} &\le \frac{1}{1-\rho_M}\sum_{i=1}^n \rho_A^{i-1}\left(\frac{1}{1-\rho_M} + (i-1)\right) \\
&\le \frac{1}{(1-\rho_M)^2(1-\rho_A)} + \frac{1}{1-\rho_M}\sum_{i=1}^n (i-1)\rho_A^{i-1} \\
&\le \frac{1}{(1-\rho_M)^2(1-\rho_A)} + \frac{\rho_A}{(1-\rho_M)(1-\rho_A)^2} \\
&\le \frac{1-\rho_A\rho_M}{(1-\rho_M)^2(1-\rho_A)^2}.
\end{aligned}
$$

Collecting our results, we thus conclude that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)}\sum_{t=1}^n \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}})\right] \\
&\lesssim \sqrt{\min\{d_{\mathbf{u}}, d_{y^s,\max}\}}(d_{\mathbf{x}} + d_{y^s,\max})\sqrt{\min\{d_{\mathbf{x}}, d_{y^s,\max}\}}\sqrt{d_{y^s,\max}}(\log(|\mathcal{S}|+1))^{3/2}\sqrt{n(i+j)},
\end{aligned}
$$

as claimed. □

H.6 Regret decomposition

Lemma H.9 (Regret decomposition)**.** *Suppose that the noise satisfies Assumption 2 and the cost functions satisfy Assumption 3. Then, the expected regret can be decomposed as:*

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)}\right] &= \mathbb{E}\left[\sum_{t=1}^{n} c_t(\mathbf{x}_t,\mathbf{u}_t) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} c_t(\mathbf{x}_t(\mathbf{M}^*),u_t(\mathbf{M}^*))\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^{n} \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^{n} \widehat{\ell}_t(\mathbf{M}^*)\right] \\
&\quad + \mathbb{E}\left[\sum_{t=1}^{n} c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] \\
&\quad + 2\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^{n} \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}} - \mathbf{x}_t^{\mathsf{nat}})\right],
\end{aligned}$$

where $\widehat{\ell}_t(\mathbf{M}) := c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}),\mathbf{u}_t(\mathbf{M}))$*. Moreover, in the adversarial setting, for any fixed reference state estimate* $\widehat{\mathbf{x}}_{t|t}$ *satisfying Assumption 4',*

$$\begin{aligned}
\widehat{\mathrm{Regret}}_{\mathcal{M}(R_M,\rho_M,H)} &= \sum_{t=1}^{n} c_t(\widehat{\mathbf{x}}_{t|t},u_t) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^{n} c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}^*),u_t(\mathbf{M}^*)) \\
&= \sum_{t=1}^{n} \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^{n} \widehat{\ell}_t(\mathbf{M}^*) \\
&\quad + \sum_{t=1}^{n} c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t)).
\end{aligned}$$

Proof. By Lemma H.1,

$$\begin{aligned}
\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)} &= \sum_{t=1}^{n} c_t(\mathbf{x}_t,\mathbf{u}_t) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^{n} c_t(\mathbf{x}_t(\mathbf{M}^*),\mathbf{u}_t(\mathbf{M}^*)) \\
&= \sum_{t=1}^{n} c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) + (\mathbf{x}_t-\widehat{\mathbf{x}}_{t|t})^\top S_t(\mathbf{x}_t-\widehat{\mathbf{x}}_{t|t}) + 2\widehat{\mathbf{x}}_{t|t}^\top S_t(\mathbf{x}_t^{\mathsf{nat}}-\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}) \\
&\quad - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \Bigg(c_t(\widehat{\mathbf{x}}_t(\mathbf{M}^*),\mathbf{u}_t(\mathbf{M}^*)) + (\mathbf{x}_t(\mathbf{M}^*)-\widehat{\mathbf{x}}_{t|t}(\mathbf{M}^*))^\top S_t(\mathbf{x}_t(\mathbf{M}^*)-\widehat{\mathbf{x}}_{t|t}(\mathbf{M}^*)) \\
&\qquad + 2\widehat{\mathbf{x}}_{t|t}(\mathbf{M}^*)^\top S_t(\mathbf{x}_t^{\mathsf{nat}}-\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}) \Bigg).
\end{aligned}$$

Now, by Lemma C.1, it follows that $\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t} = \mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}} = \mathbf{x}_t(\mathbf{M}^*) - \mathbf{x}_{t|t}^{\mathsf{nat}}(\mathbf{M}^*)$. In particular, this implies that the quadratic terms above do not depend on the policy, and thus cancel. Therefore, the above simplifies to:

$$\begin{aligned}
\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)} &= \sum_{t=1}^{n} c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) + 2\widehat{\mathbf{x}}_{t|t}^\top S_t(\mathbf{x}_t^{\mathsf{nat}}-\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}) \\
&\quad - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \left(c_t(\widehat{\mathbf{x}}_t(\mathbf{M}^*),\mathbf{u}_t(\mathbf{M}^*)) + 2\widehat{\mathbf{x}}_{t|t}(\mathbf{M}^*)^\top S_t(\mathbf{x}_t^{\mathsf{nat}}-\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}}) \right).
\end{aligned}$$

Next, by the Orthogonality principle (Lemma B.7), since $\mathbf{x}_{t|t}$ is a linear combination of $(\mathbf{y}_\ell^{\mathsf{nat}})_{\ell \le t}$, it follows that $\mathbb{E}\left[\widehat{\mathbf{x}}_{t|t}^\top S_t(\mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}})\right] = 0$. Therefore, we conclude:

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)}\right] &= \mathbb{E}\left[\sum_{t=1}^n c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t)\right] \\
&\quad - \mathbb{E}\left[\min_{\mathbf{M}^* \in \mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n c_t(\widehat{\mathbf{x}}_t(\mathbf{M}^*), \mathbf{u}_t(\mathbf{M}^*)) + 2\widehat{\mathbf{x}}_{t|t}(\mathbf{M}^*) S_t(\mathbf{x}_t^{\mathsf{nat}} - \widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}})\right] \\
&\le \mathbb{E}\left[\sum_{t=1}^n c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \min_{\mathbf{M}^* \in \mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n c_t(\widehat{\mathbf{x}}_t(\mathbf{M}^*), \mathbf{u}_t(\mathbf{M}^*))\right] \\
&\quad + 2\mathbb{E}\left[\sup_{\mathbf{M} \in \mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}} - \mathbf{x}_t^{\mathsf{nat}})\right].
\end{aligned}$$

Finally, using the fact that, by definition, $\widehat{\ell}_t(\mathbf{M}^*) = c_t(\widehat{\mathbf{x}}_{t|t}(\mathbf{M}^*), \mathbf{u}_t(\mathbf{M}^*))$ and $\mathbf{u}_t = \mathbf{u}_t(\mathbf{M}(t))$, we can further decompose the above as:

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)}\right] &\le \mathbb{E}\left[\sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}^* \in \mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}^*)\right] \\
&\quad + \mathbb{E}\left[\sum_{t=1}^n c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] \\
&\quad + 2\mathbb{E}\left[\sup_{\mathbf{M} \in \mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\widehat{\mathbf{x}}_{t|t}^{\mathsf{nat}} - \mathbf{x}_t^{\mathsf{nat}})\right],
\end{aligned}$$

which gives the first claimed bound.

The equality for $\widehat{\mathrm{Regret}}_{\mathcal{M}(R_M,\rho_M,H)}$ follows immediately from the definition and the definition of $\widehat{\ell}_t(\mathbf{M})$. □

H.7 Proof of regret bound

Theorem H.10. *Suppose that, for some $\eta > 0$, Algorithm 1 is run with group-norm regularizer*

$$\Psi(\mathbf{M}) := \frac{1}{2}\frac{2e\max\{1, \ln(|\mathcal{S}|)\}}{\eta}\|\mathbf{M}\|_{\rho_M^{-1/2},q} \quad \textit{where} \quad q = \begin{cases} 2 & \textit{if } |\mathcal{S}| < e \\ \frac{2\ln(|\mathcal{S}|)}{2\ln(|\mathcal{S}|)-2} & \textit{o.w.} \end{cases} \tag{Reg}$$

Then, if the system (LDS) satisfy Assumption 1, we have the following regret guarantees:

1. *(Stochastic): Suppose that the initial state $\mathbf{x}_0$ and the process and observation noises $\mathbf{w}_t, \mathbf{v}_t$ satisfy Assumption 2. Further, suppose that the system is strongly detectable under the Kalman filter, i.e., Assumption 4. Then, assuming the cost functions satisfy Assumption 3, then the controls $\mathbf{u}_t$ selected by Algorithm 1 with $\eta = \Theta(1/\sqrt{n})$ and resulting states $\mathbf{x}_t$ satisfy:*

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Regret}_{\mathcal{M}(R_M,\rho_M,H)}\right] &= \mathbb{E}\left[\sum_{t\in[n]} c_t(\mathbf{x}_t, \mathbf{u}_t) - \min_{\mathbf{M}^* \in \mathcal{M}(R_M,\rho_M,H)} \sum_{t\in[n]} c_t(\mathbf{x}_t(\mathbf{M}^*), \mathbf{u}_t(\mathbf{M}^*))\right] \\
&= \mathcal{O}((\mathrm{poly}\log(|S|+1))(\mathrm{poly}(\mathrm{dimension})\sqrt{n}).
\end{aligned}$$

2. *(Adversarial): Suppose that the initial state $\mathbf{x}_0$ and the process and observation noises $\mathbf{w}_t, \mathbf{v}_t$ are chosen by an oblivious adversary according to Assumption 2'. Let $\widehat{\mathbf{x}}_{0|-1} = \mathbf{0}$ and*

$$\widehat{\mathbf{x}}_{t|t} = (I - \tilde{L}C)\widehat{\mathbf{x}}_{t|t-1} + \tilde{L}\mathbf{y}_t \quad \textit{and} \quad \widehat{\mathbf{x}}_{t+1|t} = A\widehat{\mathbf{x}}_{t|t}$$

be any fixed linear dynamical state estimator satisfying Assumption 4'. Then, assuming the cost functions (chosen by an oblivious adversary) satisfy Assumption 3', then the controls $\mathbf{u}_t$ *selected by Algorithm 1 with* $\eta = \Theta(1/\sqrt{n})$ *and resulting states* $\mathbf{x}_t$ *satisfy:*

$$\begin{aligned}\widehat{\text{Regret}}_{\mathcal{M}(R_M,\rho_M,H)} &= \sum_{t\in[n]} c_t(\widehat{\mathbf{x}}_t,\mathbf{u}_t) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t\in[n]} c_t(\widehat{\mathbf{x}}_t(\mathbf{M}^*),\mathbf{u}_t(\mathbf{M}^*)) \\ &= \mathcal{O}((\text{poly}\log(|S|+1))(\text{poly}(\text{dimension})\sqrt{n}).\end{aligned}$$

Proof. For the stochastic case, we begin by decomposing the regret using Lemma H.9. Indeed,

$$\begin{aligned}\mathbb{E}\left[\text{Regret}_{\mathcal{M}(R_M,\rho_M,H)}\right] \le \mathbb{E}&\left[\sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}^*)\right] \\ &+ \mathbb{E}\left[\sum_{t=1}^n c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] \\ &+ 2\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} - \mathbf{x}^{\mathsf{nat}}_t)\right],\end{aligned} \tag{7}$$

The first term in (7) can be bounded using standard regret guarantees for the FTRL algorithm. In particular, using Lemma H.6,

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}^*)\right] \le{}& \frac{2e\min\{d_{y^s,\max}, d_{\mathbf{u}}\} R_M^2 \max\{1,\ln(|\mathcal{S}|)\}}{(1-\sqrt{\rho_M})^2\eta} \\ &+ \frac{\eta}{2}\sum_{t=1}^n \mathbb{E}\left[\|\nabla\widehat{\ell}_t(\mathbf{M}(t))\|^2_{\sqrt{\rho_M},\infty}\right].\end{aligned}$$

Therefore, by Lemma G.4,

$$\mathbb{E}\left[\sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}^*)\right] \le \frac{2e\min\{d_{y^s,\max}, d_{\mathbf{u}}\} R_M^2 \max\{1,\ln(|\mathcal{S}|)\}}{(1-\sqrt{\rho_M})^2\eta} + \frac{\eta}{2} R_{\mathsf{grad},\infty} n$$

For the second term in (7), by Lemma H.3

$$\mathbb{E}\left[\sum_{t=1}^n c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t))\right] \le \eta R_{x,\mathsf{cost}}(\eta R^2_{\widehat{\mathbf{x}},\mathsf{mem}} + 2R_x R_{\widehat{\mathbf{x}},\mathsf{mem}})$$

For the third term, by Lemma H.8

$$\begin{aligned}&\mathbb{E}\left[\sup_{\mathbf{M}\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\mathbf{x}}_{t|t}(\mathbf{M})^\top S_t(\widehat{\mathbf{x}}^{\mathsf{nat}}_{t|t} - \mathbf{x}^{\mathsf{nat}}_t)\right] \\ &\lesssim \sqrt{\min\{d_{\mathbf{u}}, d_{y^s,\max}\}}(d_{\mathbf{x}} + d_{y^s,\max})\sqrt{\min\{d_{\mathbf{x}}, d_{y^s,\max}\}}\sqrt{d_{y^s,\max}}(\log(|\mathcal{S}|+1))^{3/2}\sqrt{n}.\end{aligned}$$

Therefore, as long as $\eta = \Theta(1/\sqrt{n})$, then

$$\mathbb{E}\left[\text{Regret}_{\mathcal{M}(R_M,\rho_M,H)}\right] \lesssim \log(|\mathcal{S}|+1)^{3/2}\sqrt{n}.$$

Moreover, in the adversarial case, we fix a reference state estimator $\widehat{\mathbf{x}}_{t|t}$ satisfying Assumption 4'. Then, by Lemma H.9,

$$\begin{aligned}\widehat{\text{Regret}}_{\mathcal{M}(R_M,\rho_M,H)} = \sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}(t)) &- \min_{\mathbf{M}^*\in\mathcal{M}(R_M,\rho_M,H)} \sum_{t=1}^n \widehat{\ell}_t(\mathbf{M}^*) \\ &+ \sum_{t=1}^n c_t(\widehat{\mathbf{x}}_{t|t},\mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t)).\end{aligned}$$

For the first term, we again apply Lemma H.6 and Lemma G.4 to obtain:

$$\begin{aligned}
&\sum_{t=1}^{n} \widehat{\ell}_t(\mathbf{M}(t)) - \min_{\mathbf{M}^* \in \mathcal{M}(R_M, \rho_M, H)} \sum_{t=1}^{n} \widehat{\ell}_t(\mathbf{M}^*) \\
&\leq \frac{2e \min\{d_{y^s,\max}, d_{\mathbf{u}}\} R_M^2 \max\{1, \ln(|\mathcal{S}|)\}}{(1-\sqrt{\rho_M})^2 \eta} + \frac{\eta}{2} \sum_{t=1}^{n} \|\nabla \widehat{\ell}_t(\mathbf{M}(t))\|^2_{\sqrt{\rho_M},\infty} \\
&\leq \frac{2e \min\{d_{y^s,\max}, d_{\mathbf{u}}\} R_M^2 \max\{1, \ln(|\mathcal{S}|)\}}{(1-\sqrt{\rho_M})^2 \eta} + \frac{\eta}{2} R_{\mathsf{grad},\infty} n.
\end{aligned}$$

For the remaining term, by Lemma H.3

$$\sum_{t=1}^{n} c_t(\widehat{\mathbf{x}}_{t|t}, \mathbf{u}_t) - \widehat{\ell}_t(\mathbf{M}(t)) \leq \eta L \max\left\{R R_{\widehat{\mathbf{x}},\mathsf{mem}}, \eta R^2_{\widehat{\mathbf{x}},\mathsf{mem}}\right\}.$$

Combining these two terms, and choosing $\eta = \Theta(1/\sqrt{n})$, we conclude that

$$\widehat{\mathrm{Regret}}_{\mathcal{M}(R_M, \rho_M, H)} \lesssim \log(|\mathcal{S}|+1)\sqrt{n}.$$

□

H.8 A note on regret scaling polynomially in the number of sensors

Remark H.11. *The first observation that one can make for our problem is that, given the stacked system from Definition 2.1, one can immediately apply the algorithm and analysis from, e.g., [SSH20] to establish $\mathcal{O}(\sqrt{n})$ regret (in terms of observation costs as in (2)). We make several observations: (i) After lifting, the dimension of the observations has changed from d_{y^s} to $|\mathcal{S}|d_{y^s}$. Therefore, since the regret guarantees from [SSH20] scale polynomially in the output dimension, the regret will scale polynomially in $|\mathcal{S}|$. (ii) Because the algorithm of [SSH20] is also based on the disturbance response controller, after lifting, we have also lifted the number of parameters of $\mathbf{M}$ we need to learn linearly in $|\mathcal{S}|$. (iii) Note that the regret guarantee we inherit from this lifting is actually* stronger *than what we are aiming for! Indeed, the class of all disturbance response controllers in the lifted system includes all controllers which may combine (potentially) all of the sensors. However, we only want to compete against the best sensor in hindsight. For these reasons, we can hope to improve upon this approach.*

I Simulations

In this section, we discuss an empirical evaluation of our algorithm on a toy setting to demonstrate the advantages of our algorithm relative to some natural benchmarks. We consider a two-dimensional linear system with stochastic noise, where $\mathbf{x}_0 \sim \mathcal{N}(0, I)$, each $\mathbf{w}_t \sim \mathcal{N}(0, .03I)$, and

$$\mathbf{x}_{t+1} = \begin{pmatrix} .82 & .12 \\ -.08 & .75 \end{pmatrix} \mathbf{x}_t + \begin{pmatrix} .4 & .05 \\ .02 & .35 \end{pmatrix} \mathbf{u}_t + \mathbf{w}_t.$$

We consider quadratic costs of the form:

$$c_t(\mathbf{x}_t, \mathbf{u}_t) = \mathbf{x}_t^\top S \mathbf{x}_t + \mathbf{u}_t^\top R \mathbf{u}_t, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & 1.2 \end{pmatrix}, \; R = \begin{pmatrix} .15 & 0 \\ 0 & .15 \end{pmatrix},$$

The observations from each sensor are also two-dimensional and of the form

$$\mathbf{y}_t^{(s)} = C^{(s)} \mathbf{x}_t + \mathbf{v}_t^{(s)} \; \forall s \in \mathcal{S},$$

where $\mathbf{v}_t^{(s)} \sim \mathcal{N}(\mathbf{0}, V_s)$ are all independent, with the first three sensors specified as:

$$\begin{aligned}
C^{(1)} &= \begin{pmatrix} 1 & 0 \\ 0 & .25 \end{pmatrix}, \quad V_1 = \begin{pmatrix} .04 & 0 \\ 0 & .4 \end{pmatrix} \\
C^{(2)} &= \begin{pmatrix} .25 & 0 \\ 0 & 1 \end{pmatrix}, \quad V_2 = \begin{pmatrix} .4 & 0 \\ 0 & .04 \end{pmatrix} \\
C^{(3)} &= \begin{pmatrix} .75 & .25 \\ -.2 & .85 \end{pmatrix}, \quad V_3 = \begin{pmatrix} .12 & 0 \\ 0 & .12 \end{pmatrix}
\end{aligned}$$

The remaining sensors $s \geq 4$ are of the form:

$$C^{(s)} = \begin{pmatrix} \cos(\theta_s) & \sin(\theta_s) \\ -.25\sin(\theta_s) & .25\cos(\theta_s) \end{pmatrix}, \; V_s = (1 + .08(s-3)) \begin{pmatrix} .08 & 0 \\ 0 & .5 \end{pmatrix},$$

where $\theta_s = \pi(3 - \sqrt{5})(s-3) \mod \pi$. Thus, sensors with a larger index have a larger variance. In the experiments, we vary the number of sensors $S \in \{5, 10, 50, 100\}$.

We consider the DRC controller class $\mathcal{M}(R_M, \rho_M, H)$ with $R_M = .3$, $\rho_M = .75$, and $H = 20$. We run each algorithm over a time-horizon of $n = 15,000$.

We implement our Algorithm 1 as specified in Theorem 4.1, with step-size $\eta = 2/\sqrt{n}$. To perform the (approximate) projection step onto $\mathcal{M}(R_M, \rho_M, H)$, we solve for the optimal DRC parameters for any fixed dual constraint, run an inner optimization procedure to solve for the (approximately) optimal dual variables, then returning these optimal DRC parameters at the computed approximate dual solution.

We compare our algorithm to several natural benchmarks:

1. Weighted $(2,2)$ FTRL: This is the same FTRL algorithm as our Algorithm 1, but with the regularizer set as the $(\rho, 2)$-group norm regularizer instead of the (ρ, q)-group norm regularizer. We use step-size $\eta = 1/\sqrt{Sn}$. Because of the Euclidean constraints, we can implement this procedure by first computing the unprojected update, then performing an $\ell_{1,\infty}$ ball projection (see, e.g., [SSH20, Appendix A.5])
2. Online Projected Gradient Descent (OPGD): Here, in place of our FTRL step, we instead compute the next DRC using online projected gradient descent, using step-size $\eta = 1/\sqrt{Sn}$.
3. No control: Here, we compare with the simplest baseline where $\mathbf{u}_t = \mathbf{0}$ for all t (recall that, since the system is stable, the average cost will remain finite).
4. Best hindsight feasible DRC: Here, after observing the entire trajectory, we search over all feasible DRC controllers of the form $\mathbf{u}_t(\mathbf{M})$. Since the states $\mathbf{x}_t(\mathbf{M})$ and controls $\mathbf{u}_t(\mathbf{M})$ are affine in $\mathbf{M}$, this is a convex quadratic optimization problem.

We run each of these algorithms over time horizon $n = 15,000$. Each algorithm is initialized with (trivially feasible) DRC $\mathbf{M}(0) = \mathbf{0}$. We average our results over 10 independent runs. Note that we used the same 10 seeds to evaluate each of the benchmarks, so they are encountering the same noise across the trajectory. The results of our experiment are summarized in Table 1

Table 1: Average cost of each control policy over time horizon $n = 15{,}000$. We report the cost averaged over 10 runs, as well as the standard error.

S	H	**Ours**	Weighted $(2,2)$ FTRL	OPGD	No control	Best hindsight feasible DRC
5	20	0.12963 ± 0.00047	0.13153 ± 0.00051	0.14256 ± 0.00056	0.18011 ± 0.00072	0.12518 ± 0.00047
10	20	0.12924 ± 0.00042	0.13263 ± 0.00048	0.14850 ± 0.00057	0.18011 ± 0.00072	0.12340 ± 0.00043
50	20	0.13106 ± 0.00039	0.14055 ± 0.00050	0.17225 ± 0.00067	0.18011 ± 0.00072	0.11910 ± 0.00039
100	20	0.13389 ± 0.00041	0.15072 ± 0.00056	0.18075 ± 0.00058	0.18011 ± 0.00072	0.11635 ± 0.00040

From the results displayed in Table 1, our algorithm performs better than all online baselines (OPGD, FTRL (2,2), and No Control). Moreover, the performance of our algorithm only has a mild degradation as we increase the number of sensors, while the other online control benchmarks degrade more significantly as we increase the number of sensors. Moreover, as we would expect, the performance of our algorithm is close to that of the best benchmark DRC in hindsight.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: Yes. We list our main contributions in Section 1.1. As our contributions are theoretical, our main results are our algorithm Algorithm 1 and the corresponding regret guarantees Theorem 4.1 and 4.1'.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: We discuss limitations and possible directions for future research building upon our results in Section 4.1.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

Justification: Yes. We list all assumptions made in our paper, and reference these assumptions in our main theorem statements Theorem 4.1 and 4.1'.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: The contributions of this paper are theoretical, and there are not significant experimental results to reproduce. We describe the setup of a toy experiment in Appendix I.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
     (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: The contributions of this paper are theoretical, and there are not significant experimental results to reproduce. We provide details of the toy experimental setup in Appendix I.

Guidelines:

- The answer [N/A] means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (https://neurips.cc/public/guides/CodeSubmissionPolicy) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

Answer: [Yes]

Justification: We describe our basic toy experiment setup in Appendix I.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: These are reported in Table 1.

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.

- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Since the only experiments in this paper are in a toy setting, no special setup should be needed to run the experiments in a short time-frame of a few hours.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: We have reviewed the NeurIPS Code of Ethics, and certify that we have adhered to it. This paper does not involve human subjects or other data. Moreover, since the focus of our work is theoretical, we are not aware of any societal impacts or potential harmful consequences of our work.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

    Justification: The focus of our work is theoretical. We are not aware of any immediate societal impacts of our work.

    Guidelines:

    - The answer [N/A] means that there is no societal impact of the work performed.
    - If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A]

    Justification: Our work is purely theoretical.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [N/A]

    Justification: Our paper is purely theoretical, so no code, data, or models were used in the paper. We provide citations to all relevant work which we build upon.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
    - If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
    - For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
    - If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: There are no assets to release for this paper.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: Our paper does not involve crowdsourcing or research with human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

Answer: [N/A]

Justification: LLMs were not used in any important way in obtaining the results for this paper. The authors used ChatGPT 5.6 Pro to help implement the toy experiment presented in Appendix I and for minor edits. The authors take full responsibility for the contents of this work.

Guidelines:

- The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.