Best of Both Worlds Guarantees for Smoothed Online Quadratic Optimization

Neelkamal Bhuyan [1] **Debankur Mukherjee** [1] **Adam Wierman** [2]

Abstract

We study the smoothed online quadratic optimization (SOQO) problem where, at each round t, a player plays an action x_t in response to a quadratic hitting cost and an additional squared ℓ_2-norm cost for switching actions. This problem class has strong connections to a wide range of application domains including smart grid management, adaptive control, and data center management, where switching-efficient algorithms are highly sought after. We study the SOQO problem in both adversarial and stochastic settings, and in this process, perform the first stochastic analysis of this class of problems. We provide the online optimal algorithm when the minimizers of the hitting cost function evolve as a general stochastic process, which, for the case of martingale process, takes the form of a *distribution-agnostic dynamic interpolation algorithm* that we call Lazy Adaptive Interpolation (LAI). Next, we present the stochastic-adversarial trade-off by proving an $\Omega(T)$ expected regret for the adversarial optimal algorithm in the literature (ROBD) with respect to LAI and, a sub-optimal competitive ratio for LAI in the adversarial setting. Finally, we present a best-of-both-worlds algorithm that obtains a robust adversarial performance while simultaneously achieving a near-optimal stochastic performance.

1 Introduction

We study a class of smoothed online quadratic optimization (SOQO) problems in which a player has to make an online decision $x_t \in \mathbb{R}^d$, in response to a quadratic hitting cost $f_t(x) = \frac{1}{2}(x - v_t)^T A(x - v_t)$ and ℓ_2-norm switching cost, i.e., $c(x_t, x_{t-1}) = \frac{1}{2}\|x_t - x_{t-1}\|_2^2$. We consider the problem over a finite horizon T, where the sequence of minimizers $\{v_t\}_{t=1}^T$ can be *adversarial* or *stochastic* and is revealed in an online fashion to the player.

In the last decade, the online optimization problem with quadratic costs has received significant attention due to its applications in several domains, including smart grid cost management (Kim & Giannakis, 2017; Wang et al., 2014; Narayanaswamy et al., 2012; Badiei et al., 2015; Mei et al., 2011; Wang & Shahidehpour, 1993; dos Santos Coelho & Lee, 2008), adaptive control (Shin et al., 2023; Li et al., 2022; 2021; Goel & Wierman, 2019; Abbasi-Yadkori & Szepesvári, 2011; Tang et al., 2021; Lale et al., 2020; Cohen et al., 2019; Dean et al., 2018; Goel et al., 2017; Li et al., 2018), data center management (Lin et al., 2011; 2012; Lu et al., 2013b), electrical vehicle charging (Kim et al., 2015; Gan et al., 2013), video transmission (Joseph & de Veciana, 2012), power systems (Narayanaswamy et al., 2012; Lu et al., 2013a; Kim & Giannakis, 2017), and chip thermal management (Zanini et al., 2009; 2010).

However, since its introduction, the smoothed online optimization problem has been studied primarily from an adversarial point of view (Andrew et al., 2013; Bansal et al., 2015; Chen et al., 2015; Antoniadis et al., 2016; Chen et al., 2016; 2018; Antonios & Schewior, 2018; Goel et al., 2019; Zhang et al., 2021; 2022; Rutten et al., 2023; Christianson et al., 2022; 2023). Attempts to incorporate noise in the related domain of LQR control consider very specific distributions and dynamics of the system (Chen et al., 2015; 2016), most of which cannot be translated to SOQO. Consequently, the existing algorithms in the online optimization community are overly pessimistic for situations that may have more structure and potentially a stochastic behavior.

In related domains, such as scheduling, stochastic modeling has been used to capture known data patterns, and algorithms can then be designed to perform well for these known models (Koutsopoulos & Tassiulas, 2011; Urgaonkar et al., 2011; Jennings et al., 1996; Gandhi et al., 2010; Joseph & de Veciana, 2012). In contrast, adversarially designed algorithms must safeguard against arbitrary inputs and thus cannot exploit the inherent structures within the environment. This trade-off between average-case and worst-case performance is long-standing in the algorithms literature (Sriskan-

[*]Equal contribution [1]H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA, USA [2]Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA, USA. Correspondence to: Neelkamal Bhuyan <nbhuyan3@gatech.edu>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

darajah & Sethi, 1989; Matsuo, 1990; Kuhn et al., 2003; Robey et al., 2022).

In the case of SOQO, the design of algorithms for the case of stochastic hitting cost functions is unexplored, with this work being the first to address it. Since much of the literature in smoothed online optimization focuses on worst-case performance, the algorithms and analysis methods are not geared to exploit the structure which is now present in a stochastic environment. This motivates us to address the stochastic SOQO problem first in the quadratic setting, answering the following question:

> *"Is there a simple characterization of a near-optimal policy for stochastic SOQO? How do existing policies, designed primarily for the adversarial setting, compare to this (near-)optimal policy?"*

Tailoring to stochastic setting may lead to lack robustness against adversarial input. This represents the other side of the long-standing trade-off between optimizing average-case and worst-case performance, and motivates us to answer the second foundational question for SOQO:

> *"Is there an algorithm for SOQO that achieves near-optimal performance simultaneously in stochastic and adversarial settings?"*

Such *best-of-both-worlds* algorithms are akin to the seminal work by Bubeck and Slivkins (Bubeck & Slivkins, 2012), which answers the question affirmatively in the multi-arm bandit setting. Best-of-both-worlds algorithms are highly sought after in the algorithms literature; however, they are also rare, and there are key characteristics distinguishing the SOQO framework from the multi-arm bandits' framework of (Bubeck & Slivkins, 2012), e.g., the *unbounded continuous action space* and the *dynamic optimal performance benchmark* (as opposed to a focus on static regret), that make it questionable whether such a best-of-both-worlds algorithm exists for SOQO.

Contributions. This paper provides answers to both of the questions highlighted above. Specifically, the contributions of this paper are four-fold:

(a) *Characterization of a stochastic online-optimal algorithm*: We start by considering the SOQO problem in the setting when the sequence of minimizers $\{v_t\}_t$ (of the hitting costs $\{f_t\}_t$) evolves as a stochastic process. We analyze the optimization problem through the lens of dynamic programming and obtain a *closed form solution* for the online optimal algorithm (Theorem 3.2). In particular, when $\{v_t\}_t$ forms a martingale, our optimal solution translates to a *distribution-agnostic* algorithm, which we call the *Lazy Adaptive Interpolation* (LAI, Algorithm 1). We further characterize its total cost, observe it to be $\mathcal{O}(T)$ (which is optimal), and to be insensitive to the distribution beyond the variance (see Theorem 3.1). Notably, this algorithm is dynamic in nature and, as expected, has a much better performance than any static hindsight-optimal decision (see Theorem 3.4), which exhibits a $\Theta\left(T^2\right)$ cost.

(b) *Stochastic suboptimality of the adversarial optimal*: When adversarial inputs are considered, the performance is measured in terms of the competitive ratio, the worst-case ratio of the cost of an algorithm to the cost of the hindsight optimal sequence of decisions. In (Goel et al., 2019), the authors propose the ROBD algorithm and show that it has an optimal competitive ratio among all online algorithms. When the inputs are martingale, however, we show that the ROBD algorithm admits a linear regret (in the time horizon T) with respect to the stochastic optimal cost, i.e., the cost of the aforementioned LAI algorithm (see Theorem 3.5).

(c) *Adversarial analysis of stochastic optimal*: To understand the robustness of the LAI algorithm against adversarial input, we perform its competitive analysis and find it potentially suboptimal (see Theorem 3.8). En route, we also establish competitive ratio bounds for a much wider class of memoryless algorithms, which might be of independent interest. The novelty of our analysis framework lies in its applicability to algorithms with decision rules that change over time (see Theorem 3.10).

(d) *Best-of-both-worlds algorithm design*: Our final contribution focuses on the design of a best-of-both-worlds algorithm for SOQO. We present a novel hyperparametric algorithm called LAI(γ) (Algorithm 2) with $\gamma \in [0, 1]$, which delivers near-optimal performance (see Corollary 3.12) in both stochastic and adversarial environments without leveraging prior knowledge about the nature of the environment. Specifically, the hyperparameter $\gamma \in [0, 1]$ determines the algorithm's position in the trade-off between adversarial and martingale settings. For any choice of $\gamma \in [0, 1]$, the algorithm guarantees (i) a finite regret (that depends on γ) with respect to the online-optimal LAI algorithm, and (ii) a finite competitive ratio (that depends on γ) very close to that of adversarial-optimal ROBD algorithm (see Theorem 3.11).

Analytical Contributions: Analytical tools in the literature of smoothed online optimization focus solely on optimizing over the current cost function, with no regard to the future evolution of the minimizer. Further, works like (Lin et al., 2021) argue that finding the optimal action sequence and cost-to-go functions is near-impossible for general cost functions. In light of these, and to obtain insight into stochastic

smoothed online quadratic optimization, we make a number of contributions on the methodological front.

In particular, at the core of our analysis, we characterize the online optimal algorithm in the stochastic setting by leveraging the dynamic programming (DP) machinery. Although the DP framework offers a powerful set of tools, a fundamental difficulty lies in understanding the structure of the value function. This is compounded by the fact that our setting involves *finite horizon and unbounded continuous state space and action space*. The horizon being finite makes the value function's structure *time-variant*, decoding which was one of the major hurdles. The infinite state/action space ruled out tabular methods and demanded a *closed form* for the time-variant value function. We overcame these by inductively identifying structure in the value function (see Proposition B.2 and Theorem C.1), enabling us to express the optimal algorithm in a simple closed form. Beyond our stochastic analysis, in the adversarial setting, the non-stationary nature of our algorithms necessitated an extension of the existing adversarial analysis methods to encompass a broader range of algorithms, allowing tighter analysis than that provided by existing methods. Lastly, the design of LAI(γ) required careful weaving of properties of both stochastic optimal and adversarial optimal, leading to a smooth transition between the two worlds.

2 Model and Preliminaries

We study smoothed online quadratic optimization (SOQO) and consider an action space $\mathbb{R}^d$, $d \geq 1$, over a finite time horizon of T rounds. The hitting cost for round t is given by $f_t(x) = \frac{1}{2}(x - v_t)^T A(x - v_t)$, where A represents a known, positive definite $d \times d$ matrix, denoted as $A \succ 0$, and v_t's are revealed in an online fashion to the decision maker. Additionally, there is a switching cost of $\frac{1}{2}\|x_t - x_{t-1}\|_2^2$ as the player transitions between actions. Therefore, the player aims to minimize $\sum_{t=1}^{T} \left(f_t(x_t) + \frac{1}{2}\|x_t - x_{t-1}\|_2^2\right)$ in an online fashion. Our work explores SOQO in both adversarial and stochastic environments.

2.1 Adversarial SOQO

Prior work on smoothed online optimization has considered an adversarial setting in which algorithms seek to ensure a performance guarantee against any sequence of minimizers $\{v_t\}_{t=1}^T$. Performance in this context is assessed most commonly using the **competitive ratio**, defined for an online algorithm ALG as *the worst-case ratio of the total cost of* ALG *and that of the offline optimal sequence of decisions*. For any finite time horizon $\{1, \ldots, T\}$, we denote the cost of ALG and the hindsight optimal solution by $\text{Cost}_{\text{ALG}}[1, T]$ and $\text{Cost}_{\text{OPT}}[1, T]$, respectively. Within the class of hitting and switching costs considered in this work, recent work in (Goel et al., 2019) has identified an algorithm, ROBD, which maintains the optimal competitive ratio possible by any online algorithm. ROBD is defined as follows.

Definition 2.1. *The action of* ROBD *in round t is $x_t = \text{argmin}_x f_t(x) + \frac{\mu_1}{2}\|x - x_{t-1}\|_2^2 + \frac{\mu_2}{2}\|x - v_t\|_2^2$, where $\mu_1 \in [0, 1]$ and $\mu_2 > 0$ are hyperparameters. Optimizing over μ_1 and μ_2, the* ROBD *algorithm in our setting is $x_t = C_{\text{ROBD}} x_{t-1} + (I - C_{\text{ROBD}}) v_t$ where $C_{\text{ROBD}} = \left(A + I + \mu_2^*\left(\lambda_{\min}^A\right) I\right)^{-1}$, with $\mu_2^*\left(\lambda_{\min}^A\right) = \frac{\lambda_{\min}^A}{2}\left(\sqrt{1 + \frac{4}{\lambda_{\min}^A}} - 1\right)$ and $\lambda_{\min}^A$ denoting the smallest eigenvalue of A.*

The competitive ratio of ROBD, as defined above, is given by $1 + \frac{1}{2}\left(\sqrt{1 + \frac{4}{\lambda_{\min}^A}} - 1\right)$ (Goel et al., 2019). A notable property of this competitive ratio used in this paper is its $1/\sqrt{\lambda_{\min}^A}$ growth as $\lambda_{\min}^A \to 0$, which has been shown to be optimal in (Goel et al., 2019).

Remark 2.2. *The limiting behavior of the competitive ratio under $\lambda_{\min}^A \to 0$ (or $\lambda_i^A \to 0 \; \forall \, i$) is the focus of smoothed online optimization literature (Bansal et al., 2015; Antoniadis et al., 2016; Antonios & Schewior, 2018; Chen et al., 2018; Goel et al., 2019; Goel & Wierman, 2019; Zhang et al., 2021; Christianson et al., 2022; 2023; Rutten et al., 2023). Here, hitting costs are very low compared to switching costs which renders follow-the-minimizer algorithm as sub-optimal, demanding search for non-trivial algorithms.*

2.2 Stochastic SOQO

We introduce and, for the first time, study a stochastic version of SOQO. The stochasticity in the problem arises from the sequence of minimizers $\{v_t\}_{t=1}^T$ being a stochastic process, with $\mathcal{F}_t = \sigma\left(\{v_s\}_{s=1}^t\right)$, $t = 1, \ldots, T$, being the natural filtration generated by $\{v_s\}_{s=1}^t$. In this work, we employ the dynamic programming (DP) machinery in the stochastic context, introduced by Bellman (Bellman, 1966), as a method to find the optimal *online* action sequence with respect to the expected total cost over the horizon. Here, the optimal online action at round t, x_t^{DP}, is given as the minimizer of

$$\min_{x \in \mathbb{R}^d} \left\{ f_t(x) + c\left(x, x_{t-1}^{DP}\right) + \mathbb{E}\left[V_{t+1}(x, v_{t+1}) | \mathcal{F}_t\right] \right\}, \quad (2.1)$$

where $V_t\left(x_{t-1}^{DP}, v_t\right)$, called the *value function*, is the minimum online expected current and future cost (or cost-to-go) at round t. The algorithm derived from solving the aforementioned problem represents the online optimal. It is worthwhile to note that the optimal online algorithm is usually characterized implicitly using the above-mentioned DP approach. However, in most cases, it is non-trivial and sometimes impossible to formulate a simple and explicit algorithm that is online-optimal (Lin et al., 2021). We achieve this is Section 3.1.

We put special focus on the case where the sequence of minimizers $\{v_t\}_{t=1}^T$ forms a martingale in $\mathbb{R}^d$, i.e,

$$\mathbb{E}[v_s|\mathcal{F}_t] = v_t \; \forall \; s \geq t. \tag{2.2}$$

In this case, we will design an algorithm that can simultaneously achieve near-optimal performance both in the stochastic (in terms of regret with respect to the online optimal algorithm) and in the adversarial case – thus, furnishing a best-of-both worlds performance. The increments to the minimizers, that is, $v_t - v_{t-1}$, are considered to have a finite covariance, denoted as Σ_t. With a slight abuse of notation, we define the 'variance' of $v_t - v_{t-1}$ as $\sigma_t^2 := \mathbb{E}\|v_t - v_{t-1}\|_2^2 = tr(\Sigma_t)$.

For the stochastic setting, we use the **dynamic online optimal algorithm** as the performance benchmark instead of the *static* hindsight optimal commonly used in many online optimization works. This is because the latter exhibits a poor performance in the current setting, with a total cost that is an order of magnitude higher than that of the online optimal (Theorem 3.4). Such a benchmark is not uncommon in the literature, as evidenced by (Vera et al., 2021). We thus, consider the **(dynamic) regret as the performance metric** in the stochastic case: For any online algorithm ALG define

$$\text{Regret}_{\text{ALG}}[1,T] := \mathbb{E}[\text{Cost}_{\text{ALG}}[1,T]] - \mathbb{E}[\text{Cost}^*[1,T]], \tag{2.3}$$

where $\mathbb{E}[\text{Cost}^*[1,T]]$ is the total cost of the online optimal algorithm in the stochastic setting.

Remark 2.3. *The stochastic and the adversarial settings have different metrics due to the different behaviors of the environments. Although an additive gap is preferred over competitive ratio, the worst-case additive gap to OPT in the adversarial setting is unbounded for finite T, unlike the stochastic environment. This concept of different metrics for separate assumptions already exists in the literature (Zhang et al., 2021; Chen et al., 2018; Goel et al., 2019).*

3 Algorithms and Main Results

We now present our main contributions, which include a characterization of an online optimal stochastic SOQO algorithm and an analysis of its performance in both stochastic and adversarial settings, an analysis of an optimal adversarial SOQO algorithm (ROBD) in the martingale setting, and the design and analysis of a best-of-both-worlds algorithm that is near-optimal in both martingale and adversarial settings.

3.1 Optimal Stochastic SOQO

The optimal online algorithm in the stochastic setting (2.2) can be characterized using the dynamic programming (DP) machinery (2.1) discussed in the preliminaries. Our primary contribution lies in simplifying the formulation of the optimal action. The simple form we obtain is a departure from the complexity typically associated with the underlying DP methodology.

Algorithm 1 Lazy Adaptive Interpolation

Input: $A \succ 0, b \in \mathbb{R}^d, T$
Initialize: $\{C_t\}_{t=1}^T$ such that $C_t^{-1} = 2I + A - C_{t+1} \; \forall \; t \in \{1, \ldots, T-1\}$ and $C_T = (I+A)^{-1}$
  for $t = 1, 2, \ldots, T$ **do**
    $x_t^{\text{LAI}} \leftarrow C_t x_{t-1}^{\text{LAI}} + (I - C_t) v_t$
  end for

The *online optimal* action, as described by Algorithm 1, is an interpolation between v_t and the previous action. The crux of the algorithm lies in the recursion followed by the matrix sequence $\{C_t\}_t$, that is, $C_t^{-1} = 2I + A - C_{t+1}$, obtained by solving the DP problem in the SOQO context. The significance of this recursion lies in its connection to (near-)optimal stochastic performance for the SOQO setting we consider. Our first result states the optimality of LAI and provides a closed form characterization of its optimal cost.

Theorem 3.1. *Lazy Adaptive Interpolation* LAI *is* ***online optimal*** *in the stochastic setting* (2.2) *with total cost*

$$\begin{aligned}
&\mathbb{E}[Cost_{\text{LAI}}[1,T]] \\
&\quad = \sum_{t=1}^{T} \frac{1}{2} \mathbb{E}\left[(v_t - v_{t-1})^T (I - C_t)(v_t - v_{t-1})\right] \\
&\quad \leq \sum_{t=1}^{T} \frac{1}{2} \mathbb{E}\left[(v_t - v_{t-1})^T (I - C_L)(v_t - v_{t-1})\right]
\end{aligned}$$

where $C_L = \frac{A+2I-\sqrt{A^2+4A}}{2}$ *and* $v_0 = x_0$.

Notice that the optimal cost in the martingale setting, described above, grows with the horizon as $\mathcal{O}(T)$.

The fact that $\{C_t\}_t$ depends solely on A and the horizon T unveils an unexpected aspect of the optimal online action: its insensitivity to the specifics of the stochastic environment. The proof of 3.1 is presented in detail in Appendix B.1.

For the martingale setting, we analyze the complete value function to extract its closed form expression. This helps us characterize both the optimal algorithm and its expected cost. Looking deeper into the *derivative* of the value function, we get the following online optimal algorithm:

Theorem 3.2. *For the general case where $\{v_t\}_{t=1}^T$ is any stochastic process, the optimal action at round t is given by*

$$\begin{aligned}
x_t &= C_t x_{t-1} + (I - C_t) v_t \\
&\quad + \sum_{s=t+1}^{T} \left(\prod_{q=t}^{s-1} C_q \right) (I - C_s) \mathbb{E}[v_s - v_{s-1}|\mathcal{F}_t]
\end{aligned}$$

where $\{C_t\}_{t=1}^T$ are the coefficients defined in LAI.

The proof is presented in Appendix C. The most important aspect is the derivative of the value function, which takes a simple yet useful form, stated in Theorem C.1, which result holds for general convex hitting costs and switching costs.

The optimality claims in Theorems 3.2 and 3.1 are agnostic to distributions (beyond the martingale structure for the latter). This implies that the optimality is robust to two extreme stochastic scenarios: (i) **distribution shifts** and, (ii) **heavy tails** (subject to finite variance).

Remark 3.3. *Note that the LAI algorithm is the **unique** online-optimal algorithm in this setting. This is because any online-optimal action has to satisfy the value function expression at every round (Bellman, 1954), which has a unique solution due to its quadratic structure.*

Within the online optimization literature, it is common practice to employ the optimal static action in hindsight as a benchmark to evaluate online algorithms. However, the following result illustrates that such a benchmark is not appropriate within our framework. This highlights the necessity to design an algorithm that adapts dynamically in time, in order to obtain good performance guarantees.

Theorem 3.4. *Consider the stochastic setting (2.2) and $f_t(x) = \frac{1}{2}\|x - v_t\|_2^2$ with the class of static algorithms where the action is $\underline{x} \in \mathbb{R}^d$ for all rounds. When the increments to the minimizers, that is, $\{v_t - v_{t-1}\}_t$, have the same variance σ^2, the cost of such an algorithm, denoted by $Cost_{\underline{x}}[1,T]$, satisfies*

$$\mathbb{E}\left[\inf_{\underline{x}} Cost_{\underline{x}}[1,T]\right] = \Theta\left(T^2\right) \tag{3.1}$$

Theorem 3.4 is proved in Appendix D.1. This result has two significant implications. **First**, in contrast with Theorem 3.1, it highlights the poor performance of the entire class of static algorithms in the stochastic SOQO framework. **Second**, it emphasizes the need for a dynamic benchmark when evaluating online algorithms; thereby leading us to use LAI as the benchmark in this work.

3.2 Stochastic Analysis of ROBD and Fixed Interpolation Algorithms

Having characterized the stochastic optimal algorithm, our attention shifts to analyzing the performance of the adversarial optimal algorithm, ROBD, in the *martingale* setting. As previously highlighted, a common concern is that the performance of adversarial algorithms may potentially be sub-par in stochastic settings. The following result demonstrates that ROBD indeed has markedly poorer performance than the stochastic optimal algorithm LAI in the stochastic setting.

Theorem 3.5. *Consider the stochastic setting (2.2). Additionally, assume that the increments $(v_t - v_{t-1})$ have the same covariance matrix Σ. Then for any $A \neq \lambda I$ (which occurs for $d \geq 2$), the regret of* ROBD *is lower bounded as*

$$Regret_{\text{ROBD}}[1,T] = \begin{cases} \Omega(T) & \text{if } A \text{ and } \Sigma \text{ do NOT have same eigenvectors} \\ \Omega(T) & \text{otherwise if } (A - \lambda_{\min}^A I)\Sigma \neq 0_{d\times d} \end{cases}$$

The most important takeaway here is the large class of covariance matrices for which the above negative result holds, revealing that ROBD has linear regret even in very simple stochastic settings. To induce such regret for ROBD, all that is needed is positive variance along one of the eigenvectors of A that does not correspond to the minimum eigenvalue. The primary factor that causes ROBD to have linear regret is the wide gap between the matrix C_{ROBD} and the matrix sequence of LAI, that is, $\{C_t\}_{t=1}^T$. In fact, this problem persists in a large class of algorithms, of which ROBD is a part.

Definition 3.6. Consider a symmetric matrix C with the same set of eigenvectors as A such that $C \prec I$. The class of Fixed Interpolation (FI) algorithms makes decisions at time t according to $x_t = Cx_{t-1} + (I - C)v_t, \forall\, t$.

It is straightforward to see that ROBD is an FI algorithm, owing to $C_{\text{ROBD}} \prec I$ and having the same set of eigenvectors as A. Although FI algorithms feature simplicity of implementation, the following result demonstrates their poor stochastic performance and, consequently, the linear regret of ROBD.

Theorem 3.7. *Consider the stochastic setting (2.2). Additionally, assume the increments $(v_t - v_{t-1})$ have the same covariance matrix Σ. Then for any $C \neq C_L = \frac{A+2I-\sqrt{A^2+4A}}{2}$,*

$$Regret_{\text{FI}}[1,T] = \begin{cases} \Omega(T) & \text{if } A \text{ and } \Sigma \text{ do NOT have same eigenvectors} \\ \Omega(T) & \text{otherwise if } (C - C_L)\Sigma \neq 0_{d\times d} \end{cases}$$

Here, the importance of the recursion followed by $\{C_t\}_{t=1}^T$, in the LAI algorithm, is highlighted in the context of fixed interpolation algorithms. Unless the matrix C is C_L, one can show that a "non-vanishing gap" exists between C and $\{C_t\}_{t=1}^T$, which becomes the driving factor for the linear regret. We present the proof of the above theorem along with the exact expression of the lower bound, which quantifies the factors responsible for the linear regret, in Theorem B.5 in Section B.2. Note that the matrix C_L, in fact, has a vanishing gap to $\{C_t\}_{t=1}^T$ and we revisit this important property in upcoming subsections.

3.3 Adversarial Analysis of Lazy Adaptive Interpolation Algorithms

We now present the other half of the contrast between the stochastic optimal and adversarially optimal algorithms. Specifically, in this section, we bound the performance of the stochastic optimal policy LAI in the adversarial setting.

Theorem 3.8. *In the adversarial setting, the competitive ratio of the* LAI *algorithm for hitting cost* $f_t(x) = \frac{1}{2}(x - v_t)^T A(x - v_t)$ *and switching cost* $c(x_t, x_{t-1}) = \frac{1}{2}\|x_t - x_{t-1}\|_2^2$ *satisfies* $CR_{\text{LAI}} \leq 1 + \frac{1}{\lambda^A_{\min}}$.

The above theorem is proved in Appendix E.1. Contrasting the adversarial performance of LAI with that of ROBD, the cost ratio between Lazy Adaptive Interpolation and ROBD can become unbounded as $\lambda^A_{\min}$ shrinks, in the following manner

$$\frac{\text{Cost}_{\text{LAI}}[1,T]}{\text{Cost}_{\text{ROBD}}[1,T]} \leq \frac{1 + \frac{1}{\lambda^A_{\min}}}{1 + \frac{1}{2}\left(\sqrt{1 + \frac{4}{\lambda^A_{\min}}} - 1\right)}$$

Although the aforementioned observation serves as an upper bound, we note that the suboptimal performance of LAI relative to ROBD is further evident in our numerical experiments. The sub-optimality of LAI can be primarily attributed to the fact that it relies only on the matrix C_T, which further dictates the entire matrix sequence $\{C_t\}_{t=1}^T$ of LAI.

Note that the uniqueness of the online optimal algorithm establishes that any algorithm aiming for optimality in the stochastic setting behaves poorly in the adversarial setting.

We analyze the LAI algorithm in the adversarial context by developing a framework that can furnish adversarial guarantees for a large of class of static and dynamic algorithms within a broader family of cost functions. In that context, we briefly introduce some terminology associated with this framework, frequently encountered in the convex optimization literature.

Definition 3.9. *The Bregman divergence between two vectors in* $\mathbb{R}^d$*,* y *and* x*, with respect to* $h(\cdot)$*. is defined as* $D_h(y||x) := h(y) - h(x) - \langle \nabla h(x), y - x\rangle$*. The function* $h(x)$ *is* α*-strongly convex and* β*-smooth if and only if*

$$\frac{\alpha}{2}\|y - x\|_2^2 \leq D_h(y||x) \leq \frac{\beta}{2}\|y - x\|_2^2 \;\forall\, y, x \tag{3.2}$$

We now present a bound on the competitive ratio of a general set of policies, including LAI, for m-strongly convex hitting costs $f_t(x)$ and switching costs as $D_h(x_t||x_{t-1})$, where $h(\cdot)$ is α-strongly convex and β-smooth.

Theorem 3.10. *Any online algorithm* ALG *that can be expressed in the following form, where* $g_t(\cdot)$ *is* α'_t*−strongly convex and* β'_t*−smooth for* $0 \leq \alpha'_t \leq \beta_t$*,*

$$x_t = \underset{x}{\text{argmin}}\; f_t(x) + D_h(x||x_{t-1}) + D_{g_t}(x||v_t) \;\forall\, t.$$

Algorithm 2 Lazy Adaptive Interpolation (γ)

Input: A $(= PD_AP^T)$, where D_A is a diagonal matrix, $\gamma \in [0, 1]$

Initialize: $\tilde{\lambda}_i^T = \left(1 + \lambda_i^A + \frac{\lambda_i^A}{2}\left(\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\right)^{-1}$ and $\tilde{D}_t = diag\left(\tilde{\lambda}_1^t, \ldots, \tilde{\lambda}_d^t\right)$. Define $\tilde{C}_T = P\tilde{D}_TP^T$ and $\tilde{C}_t^{-1} = 2I + A - \tilde{C}_{t+1}\;\forall\; t \in \{1, \ldots, T-1\}$

for $t = 1, 2, \ldots, T$ **do**
  $x_t \leftarrow \tilde{C}_t x_{t-1} + (I - \tilde{C}_t)v_t$
end for

has a competitive ratio in the adversarial setting upper-bounded as

$$CR_{\text{ALG}} \leq 1 + \max\left\{\frac{\max_t \beta'_t}{m}, \frac{\beta^2/\alpha}{\min_t \alpha'_{t-1} + m}\right\}$$

The significance of this result lies in its applicability to a broad spectrum of algorithms, all while preserving the optimal competitive ratio (Goel et al., 2019) for squared switching costs, that is, $c(x_t, x_{t-1}) = \frac{1}{2}\|x_t - x_{t-1}\|_2^2$. The proof, presented in Section B.6, follows a potential function technique, a popular approach in the adversarial online algorithms literature. The cornerstone of the proof is a potential function that is tailored to the algorithm considered here, yielding a competitive ratio that is specific to the $\{g_t(\cdot)\}_t$.

3.4 A Near-Optimal Algorithm for Stochastic and Adversarial SOQO

In the previous two subsections, we demonstrated the shortcomings of the adversarial optimal algorithm in stochastic settings and the sub-optimal performance of the stochastic optimal algorithm in adversarial environments. Our aim now is *to design an algorithm that has adversarial performance similar to that of* ROBD *without sacrificing stochastic performance.*

Our proposed algorithm builds on LAI, and is presented in Algorithm 2. LAI(γ)'s design incorporates a parameter γ, such that when $\gamma = 0$, it behaves as Lazy Adaptive Interpolation, and when $\gamma = 1$, it becomes a fixed interpolation algorithm with $C = C_L$, which achieves a near-optimal competitive ratio. The expression in $\tilde{C}_t$, specifically $\frac{\lambda_i^A}{2}\left(\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)$, is chosen to replicate the structure of ROBD, that is, $\frac{\lambda^A_{\min}}{2}\left(\left(1 + \frac{4}{\lambda^A_{\min}}\right)^{\frac{1}{2}} - 1\right)$, while being close to LAI.

Our choice of $\tilde{C}_t$ is guided by two key observations from previous subsections. First, we recognize the significance of the recursion in LAI's matrix sequence, which, as we will

see, is instrumental in achieving near-optimal stochastic performance. Second, we take into account the dependence of ROBD on $\lambda^A_{\min}$ in Definition 2.1, and its effect on adversarial performance.

Our main result below characterizes LAI(γ)'s performance in both adversarial and stochastic settings across the entire spectrum of γ values.

Theorem 3.11. *The following two results hold for* LAI(γ) *in the stochastic and adversarial settings:*

(i) *In the stochastic setting* (2.2) *with increments having identical covariance* Σ, LAI(γ) *has* ***constant regret*** *with respect to* LAI *where* $\sigma^2 = tr(\Sigma)$.

$$\textit{Regret}_{\text{LAI}(\gamma)}[1,T] \leq \frac{\sigma^2}{4}\left(\frac{\left(1+\frac{4}{\lambda^A_{\min}}\right)^{\frac{\gamma}{2}}-1}{\lambda^A_{\min}+2}\right).$$

(ii) *In the adversarial setting,* LAI(γ)*'s competitive ratio,* $CR_{\text{LAI}(\gamma)}$*, is upper bounded by*

$$1+\max\left\{\frac{1}{2}\left(\sqrt{\kappa(A)^2+\frac{4\kappa(A)}{\lambda^A_{\min}}}-\kappa(A)\right),\ \frac{2}{\lambda^A_{\min}}\left(\left(1+\frac{4}{\lambda^A_{\min}}\right)^{\frac{\gamma}{2}}+1\right)^{-1}\right\}$$

where $\kappa(A)$ *is the condition number of* A.

The guarantee on adversarial performance, proved in Appendix B.7, results from this specific choice of $\tilde{C}_T$, coupled with the competitive analysis framework discussed in Theorem 3.10. The proof of stochastic performance leverages the structure of the recursion $\tilde{C}_t^{-1} = 2I + A - \tilde{C}_{t+1}$ to show constant regret for any interpolation algorithm satisfying it. We show this in detail in Appendix B.4.

Recall the importance of the case of $\lambda^A_i \ll 1$ discussed in Remark 2.2. We study $\text{CR}_{\text{LAI}(\gamma)}$ in this regime and contrast it with the adversarial optimal ROBD (in terms of dependence of the competitive ratio on λ^A_i). Below, we present our ***best-of-both-worlds*** result, with the proof in Appendix B.7.

Corollary 3.12. *For the* $\lambda^A_{\max} \ll 1$ *regime in the adversarial setting, the* LAI(γ) *algorithm with* $\gamma = 1$ *achieves a competitive ratio of*

$$CR_{\text{LAI}(1)} \leq 1+\sqrt{\frac{\kappa(A)}{\lambda^A_{\min}}}$$

while simultaneously achieving a ***constant regret*** *in the stochastic setting* (2.2).

Comparing the above result with ROBD's competitive ratio of $O(1/\sqrt{\lambda^A_{\min}})$ in the $\lambda^A_{\min} \ll 1$ regime highlights that LAI(1) achieves a near-optimal competitive ratio in the adversarial setting, for well-behaved hitting costs, while simultaneously having near-optimal performance in the martingale setting.

4 Numerical Experiments

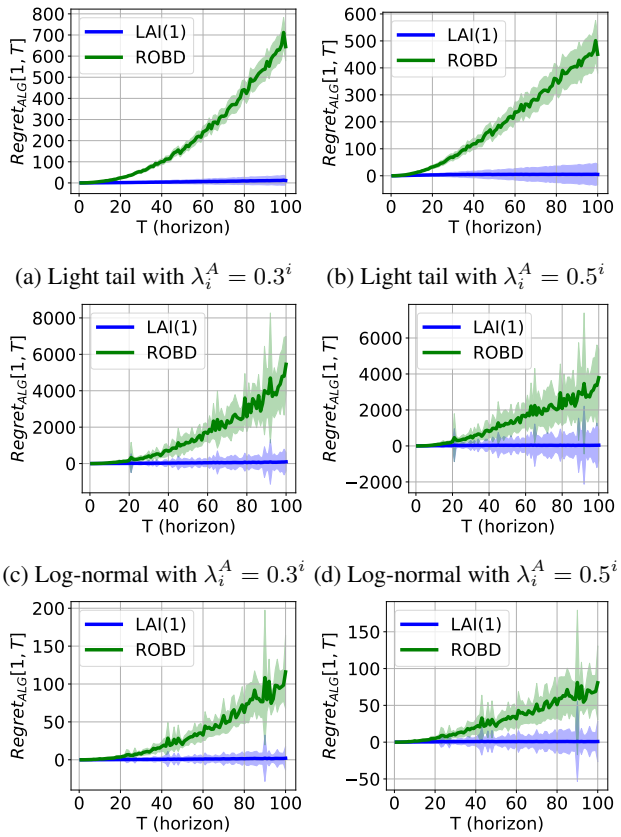
To further explore SOQO in stochastic and adversarial settings, we conduct empirical experiments to evaluate the performance of our algorithms in a range of environments. Our experiments fall into two main categories: first, purely stochastic experiments where we provide empirical evidence for two primary claims – the inferior performance of ROBD in comparison to LAI(γ) and LAI, and the robustness of our algorithms to distribution shifts and variations in tail behavior (light or heavy). Second, we delve into mixed adversarial-stochastic experiments that combine elements of both adversarial and stochastic scenarios, examining how our algorithms perform in this hybrid setting. Although we offer a summarized overview of our experimental setup here, detailed procedures can be found in Appendix G.

In all of our experiments, we maintain a consistent action space $\mathbb{R}^{10}$ and employ the matrix A with eigenvalues selected from one of three sequences: $\{0.3^i\}_{i=0}^9$, $\{0.45^i\}_{i=0}^9$, or $\{0.5^i\}_{i=0}^9$. The sequence of minimizers, denoted $\{v_t\}_t$, is adjusted according to the specific type of experiment under consideration. In both stochastic and adversarial environments, we compare the ROBD algorithm to the LAI(γ) algorithm with $\gamma = 1$. The specific value of γ is chosen to emphasize that even the adversarial extreme of LAI(γ) demonstrates exceptional stochastic performance.

4.1 Experiments in purely stochastic environments

Our stochastic experiments are categorized into three distinct groups. The first category encompasses light tail distributions, featuring five zero-mean distributions: uniform, normal, Laplace, logistic, and Gumbel. In this particular experiment, we introduce a new distribution every $T/5$ rounds, underscoring the robustness of our results to distribution shifts. The following two categories focus on heavy-tail distributions, specifically log-normal and Pareto distributions. Across all these experiments, the minimizer sequence adheres to a martingale structure, where the increments exhibit uniform variance and are drawn from the respective distribution. Our analysis involves plotting the expected total regret (relative to LAI) for both $\{0.3^i\}_{i=0}^9$ and $\{0.5^i\}_{i=0}^9$ over various horizon values $T \in \{1, 2, \ldots, 100\}$. We present the results as the sample mean derived from $N = 1000$ runs, accompanied by the 95th percentile for added clarity.

The trends in Fig. 1 validate our claims in Theorem 3.5

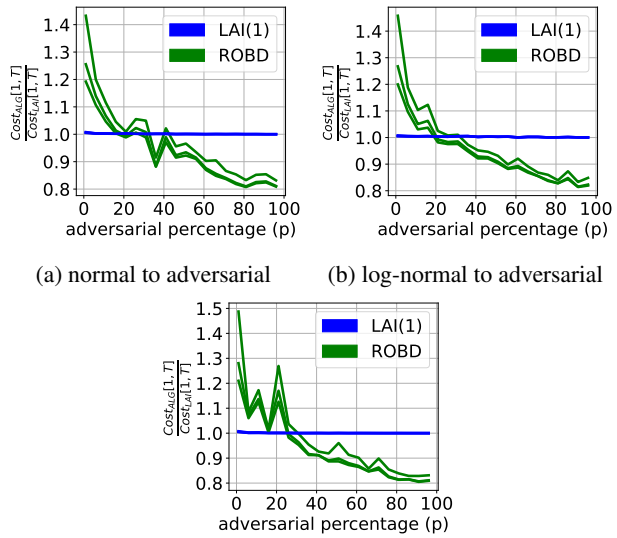


(a) Light tail with $\lambda_i^A = 0.3^i$ (b) Light tail with $\lambda_i^A = 0.5^i$

(c) Log-normal with $\lambda_i^A = 0.3^i$ (d) Log-normal with $\lambda_i^A = 0.5^i$

(e) Pareto with $\lambda_i^A = 0.3^i$ (f) Pareto with $\lambda_i^A = 0.5^i$

Figure 1: Regret of LAI(1) and ROBD for martingale minimizers with light and heavy tails



(a) normal to adversarial (b) log-normal to adversarial

(c) pareto to adversarial

Figure 2: Behavior of LAI(1) and ROBD for a mixed sequence of minimizers. In each figure, the topmost plot corresponds to $\{0.3^i\}_{i=0}^9$, the middle one represents $\{0.45^i\}_{i=0}^9$, and the bottom-most plot pertains to $\{0.5^i\}_{i=0}^9$.

and Theorem 3.11 (i) regarding ROBD's linear regret and LAI(1)'s constant regret. In fact, we observe that the regret of LAI(1) is virtually zero in contrast to that of ROBD, demonstrating the superiority of LAI(γ) in practice. The insensitivity of LAI(γ) to the form of the distribution is further highlighted by its consistent near-optimal stochastic performance in all simulated stochastic settings. In particular, we would like to stress the stability demonstrated by LAI(γ) under shifting distributions and in heavy-tailed stochastic environments. The negligible regret shown by LAI(1) establishes the superiority of LAI(γ) over ROBD in any distribution with finite variance.

4.2 Experiments in stochastic and adversarial environments

In this series of experiments, we introduce adversarial minimizers into a martingale minimizer sequence, with the extent of adversarial influence determined by the parameter known as the *adversarial percentage* denoted (p). This parameter spans from 0 (indicating a fully stochastic scenario) to 100 (representing a fully adversarial scenario). To facilitate meaningful comparisons across stochastic and adversarial environments, we calculated the ratio of an online algorithm's total cost to that of LAI. This normalization technique helps account for the differences in the orders of magnitude between the total costs in stochastic and adversarial settings, allowing us to evaluate the relative performance of various algorithms effectively.

In these experiments, we explore scenarios involving three different sequences of eigenvalues: $\{0.3^i\}_{i=0}^9$, $\{0.45^i\}_{i=0}^9$, and $\{0.5^i\}_{i=0}^9$. The key observation is that, on the adversarial end, the relative performance of LAI(1) and ROBD is consistent across different A matrices. However, it becomes evident that the stochastic performance of ROBD deteriorates significantly when a smaller A matrix is used.

As we gradually intensify the adversarial characteristics of the environment, we notice a relatively smooth shift from LAI(1) to ROBD in terms of identifying the "superior algorithm." In fact, until a certain threshold of adversarial influence, approximately around 20%, LAI(1) surpasses ROBD in performance. This intriguing observation prompts a more in-depth analysis of SOQO in a stochastic environment with adversarial contamination.

5 Concluding Remarks

This research broadens the horizons of smoothed online optimization in two unexplored dimensions: the examination

of SOQO within stochastic contexts and the concept of "best of both worlds" algorithms in this domain. These directions hold immense potential, especially the stochastic analysis beyond quadratic costs. Within the framework discussed in this work, two extensions are especially intriguing: (i) multi-agents systems with costs coupling the agents and (ii) learning the A matrix.

Acknowledgements

This work was partially supported by the NSF grants CIF-2113027 and CPS-2240982 and the Georgia Tech ARC-ACO Fellowship.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

References

Abbasi-Yadkori, Y. and Szepesvári, C. Regret bounds for the adaptive control of linear quadratic systems. volume 19, pp. 1–26. PMLR, 1 2011. URL https://proceedings.mlr.press/v19/abbasi-yadkori11a.html.

Andrew, L., Barman, S., Ligett, K., Lin, M., Meyerson, A., Roytman, A., and Wierman, A. A tale of two metrics: Simultaneous bounds on competitiveness and regret. volume 30, pp. 741–763. PMLR, 7 2013.

Antoniadis, A., Barcelo, N., Nugent, M., Pruhs, K., Schewior, K., and Scquizzato, M. Chasing Convex Bodies and Functions. In Kranakis, E., Navarro, G., and Chávez, E. (eds.), *LATIN 2016: Theoretical Informatics*, pp. 68–81, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg. ISBN 978-3-662-49529-2.

Antonios, K. A. and Schewior. A tight lower bound for online convex optimization with switching costs. pp. 164–175. Springer International Publishing, 2018. ISBN 978-3-319-89441-6.

Badiei, M., Li, N., and Wierman, A. Online convex optimization with ramp constraints. pp. 6730–6736, 2015. doi: 10.1109/CDC.2015.7403279.

Bansal, N., Gupta, A., Krishnaswamy, R., Pruhs, K., Schewior, K., and Stein, C. A 2-competitive algorithm for online convex optimization with switching costs. volume 40, pp. 96–109. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015. ISBN 978-3-939897-89-7. doi: 10.4230/LIPIcs.APPROX-RANDOM.2015.96. Keywords: Stochastic, Scheduling.

Bellman, R. The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60(6):503–515, 1954.

Bellman, R. Dynamic programming. *Science*, 153:34–37, 1966. doi: 10.1126/science.153.3731.34.

Bubeck, S. and Slivkins, A. The Best of Both Worlds: Stochastic and Adversarial Bandits. In Mannor, S., Srebro, N., and Williamson, R. C. (eds.), *Proceedings of the 25th Annual Conference on Learning Theory*, volume 23 of *Proceedings of Machine Learning Research*, pp. 42.1–42.23, Edinburgh, Scotland, 2012. PMLR. URL https://proceedings.mlr.press/v23/bubeck12b.html.

Chen, N., Agarwal, A., Wierman, A., Barman, S., and Andrew, L. L. H. Online convex optimization using predictions. pp. 191–204. Association for Computing Machinery, 2015. ISBN 9781450334860. doi: 10.1145/2745844.2745854.

Chen, N., Comden, J., Liu, Z., Gandhi, A., and Wierman, A. Using predictions in online optimization: Looking forward with an eye on the past. pp. 193–206. Association for Computing Machinery, 2016. ISBN 9781450342667. doi: 10.1145/2896377.2901464.

Chen, N., Goel, G., and Wierman, A. Smoothed online convex optimization in high dimensions via online balanced descent. volume 75, pp. 1574–1594. PMLR, 7 2018.

Chen, S., Tu, W.-W., Zhao, P., and Zhang, L. Optimistic online mirror descent for bridging stochastic and adversarial online convex optimization. *arXiv preprint arXiv:2302.04552*, 2023.

Christianson, N., Handina, T., and Wierman, A. Chasing convex bodies and functions with black-box advice. volume 178, pp. 867–908. PMLR, 7 2022.

Christianson, N., Shen, J., and Wierman, A. Optimal robustness-consistency tradeoffs for learning-augmented metrical task systems. volume 206, pp. 9377–9399. PMLR, 7 2023.

Cohen, A., Koren, T., and Mansour, Y. Learning linear-quadratic regulators efficiently with only t regret. volume 97, pp. 1300–1309. PMLR, 1 2019. URL https://proceedings.mlr.press/v97/cohen19b.html.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. Regret bounds for robust adaptive control of the linear quadratic regulator. volume 31. Curran Associates, Inc.,

2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/file/0ae3f79a30234b6c45a6f7d298ba1310-Paper.pdf.

dos Santos Coelho, L. and Lee, C.-S. Solving economic load dispatch problems in power systems using chaotic and gaussian particle swarm optimization approaches. *International Journal of Electrical Power & Energy Systems*, 30:297–307, 2008. ISSN 0142-0615. doi: https://doi.org/10.1016/j.ijepes.2007.08.001. URL https://www.sciencedirect.com/science/article/pii/S0142061507001032.

Gan, L., Topcu, U., and Low, S. H. Optimal decentralized protocol for electric vehicle charging. *IEEE Transactions on Power Systems*, 28:940–951, 2013. doi: 10.1109/TPWRS.2012.2210288.

Gandhi, A., Gupta, V., Harchol-Balter, M., and Kozuch, M. A. Optimality analysis of energy-performance tradeoff for server farm management. *Perform. Eval.*, 67(11): 1155–1171, 2010.

Goel, G. and Wierman, A. An online algorithm for smoothed regression and lqr control. volume 89, pp. 2504–2513. PMLR, 7 2019.

Goel, G., Chen, N., and Wierman, A. Thinking fast and slow: Optimization decomposition across timescales. *SIGMETRICS Perform. Eval. Rev.*, 45:27–29, 10 2017. ISSN 0163-5999. doi: 10.1145/3152042.3152052.

Goel, G., Lin, Y., Sun, H., and Wierman, A. Beyond online balanced descent: An optimal algorithm for smoothed online optimization. *Advances in Neural Information Processing Systems*, 32, 2019.

Jennings, O. B., Mandelbaum, A., Massey, W. A., and Whitt, W. Server staffing to meet time-varying demand. *Management Science*, 42:1383–1394, 1996. ISSN 00251909, 15265501. URL http://www.jstor.org/stable/2634372.

Joseph, V. and de Veciana, G. Jointly optimizing multi-user rate adaptation for video transport over wireless systems: Mean-fairness-variability tradeoffs. pp. 567–575, 2012. doi: 10.1109/INFCOM.2012.6195799.

Kim, S.-J. and Giannakis, G. B. An online convex optimization approach to real-time energy pricing for demand response. *IEEE Transactions on Smart Grid*, 8:2784–2793, 2017. doi: 10.1109/TSG.2016.2539948.

Kim, T., Yue, Y., Taylor, S., and Matthews, I. A decision tree framework for spatiotemporal sequence prediction. pp. 577–586. Association for Computing Machinery, 2015. ISBN 9781450336642. doi: 10.1145/2783258.2783356.

Koutsopoulos, I. and Tassiulas, L. Control and optimization meet the smart power grid: scheduling of power demands for optimal energy management. pp. 41–50. Association for Computing Machinery, 2011. ISBN 9781450313131. doi: 10.1145/2318716.2318723. URL https://doi.org/10.1145/2318716.2318723.

Kuhn, F., Wattenhofer, R., and Zollinger, A. Worst-case optimal and average-case efficient geometric ad-hoc routing. pp. 267–278. Association for Computing Machinery, 2003. ISBN 1581136846. doi: 10.1145/778415.778447. URL https://doi.org/10.1145/778415.778447.

Lale, S., Azizzadenesheli, K., Hassibi, B., and Anandkumar, A. Explore more and improve regret in linear quadratic regulators. *ArXiv*, abs/2007.12291, 2020. URL https://api.semanticscholar.org/CorpusID:220768718.

Li, T., Yang, R., Qu, G., Shi, G., Yu, C., Wierman, A., and Low, S. Robustness and consistency in linear quadratic control with predictions. *arXiv preprint arXiv:2106.09659*, 2021.

Li, T., Yang, R., Qu, G., Shi, G., Yu, C., Wierman, A., and Low, S. Robustness and consistency in linear quadratic control with untrusted predictions. *Proc. ACM Meas. Anal. Comput. Syst.*, 6, 2 2022. doi: 10.1145/3508038. URL https://doi.org/10.1145/3508038.

Li, Y., Qu, G., and Li, N. Using predictions in online optimization with switching costs: A fast algorithm and a fundamental limit. pp. 3008–3013, 2018. doi: 10.23919/ACC.2018.8431296.

Lin, M., Wierman, A., Andrew, L. L. H., and Thereska, E. Dynamic right-sizing for power-proportional data centers. pp. 1098–1106, 2011. doi: 10.1109/INFCOM.2011.5934885.

Lin, M., Liu, Z., Wierman, A., and Andrew, L. L. H. Online algorithms for geographical load balancing. pp. 1–10, 2012. doi: 10.1109/IGCC.2012.6322266.

Lin, Y., Hu, Y., Shi, G., Sun, H., Qu, G., and Wierman, A. Perturbation-based regret analysis of predictive control in linear time varying systems. volume 34, pp. 5174–5185. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/298f587406c914fad5373bb689300433-Paper.pdf.

Lu, L., Tu, J., Chau, C.-K., Chen, M., and Lin, X. Online energy generation scheduling for microgrids with

intermittent energy sources and co-generation. pp. 53–66. Association for Computing Machinery, 2013a. ISBN 9781450319003. doi: 10.1145/2465529.2465551.

Lu, T., Chen, M., and Andrew, L. L. H. Simple and effective dynamic provisioning for power-proportional data centers. *IEEE Transactions on Parallel and Distributed Systems*, 24:1161–1171, 2013b. doi: 10.1109/TPDS.2012.241.

Matsuo, H. Cyclic sequencing problems in the two-machine permutation flow shop: Complexity, worst-case, and average-case analysis. *Naval Research Logistics (NRL)*, 37:679–694, 10 1990. ISSN 0894-069X. doi: https://doi.org/10.1002/1520-6750(199010)37:5⟨679::AID-NAV3220370507⟩3.0.CO;2-Q. URL https://doi.org/10.1002/1520-6750(199010)37:5<679::AID-NAV3220370507>3.0.CO2-Q.

Mei, S., Wang, Y., and Sun, Z. Robust economic dispatch considering renewable generation. pp. 1–5, 2011. doi: 10.1109/ISGT-Asia.2011.6167096.

Narayanaswamy, B., Garg, V. K., and Jayram, T. S. Online optimization for the smart (micro) grid. pp. 1–10, 2012. doi: 10.1145/2208828.2208847.

Robey, A., Chamon, L., Pappas, G. J., and Hassani, H. Probabilistically robust learning: Balancing average and worst-case performance. volume 162, pp. 18667–18686. PMLR, 9 2022. URL https://proceedings.mlr.press/v162/robey22a.html.

Rutten, D., Christianson, N., Mukherjee, D., and Wierman, A. Smoothed online optimization with unreliable predictions. *Proc. ACM Meas. Anal. Comput. Syst.*, 7, 3 2023. doi: 10.1145/3579442.

Sachs, S., Hadiji, H., van Erven, T., and Guzmán, C. Between stochastic and adversarial online convex optimization: Improved regret bounds via smoothness. volume 35, pp. 691–702. Curran Associates, Inc., 2022.

Sachs, S., Hadiji, H., van Erven, T., and Guzman, C. Accelerated rates between stochastic and adversarial online convex optimization. *arXiv preprint arXiv:2303.03272*, 2023.

Shin, S., Lin, Y., Qu, G., Wierman, A., and Anitescu, M. Near-optimal distributed linear-quadratic regulator for networked systems. *SIAM Journal on Control and Optimization*, 61:1113–1135, 2023. doi: 10.1137/22M1489836. URL https://doi.org/10.1137/22M1489836.

Sriskandarajah, C. and Sethi, S. P. Scheduling algorithms for flexible flowshops: Worst and average case performance. *European Journal of Operational Research*, 43:143–160, 1989. ISSN 0377-2217. doi: https://doi.org/10.1016/0377-2217(89)90208-7. URL https://www.sciencedirect.com/science/article/pii/0377221789902087.

Tang, Y., Zheng, Y., and Li, N. Analysis of the optimization landscape of linear quadratic gaussian (lqg) control. volume 144, pp. 599–610. PMLR, 1 2021. URL https://proceedings.mlr.press/v144/tang21a.html.

Urgaonkar, R., Urgaonkar, B., Neely, M. J., and Sivasubramaniam, A. Optimal power cost management using stored energy in data centers. pp. 221–232. Association for Computing Machinery, 2011. ISBN 9781450308144. doi: 10.1145/1993744.1993766. URL https://doi.org/10.1145/1993744.1993766.

Vera, A., Banerjee, S., and Gurvich, I. Online allocation and pricing: Constant regret via bellman inequalities. *Operations Research*, 69:821–840, 2021. doi: 10.1287/opre.2020.2061.

Wang, C. and Shahidehpour, S. M. Effects of ramp-rate limits on unit commitment and economic dispatch. *IEEE Transactions on Power Systems*, 8:1341–1350, 1993. doi: 10.1109/59.260859.

Wang, H., Huang, J., Lin, X., and Mohsenian-Rad, H. Exploring smart grid and data center interactions for electric power load balancing. *SIGMETRICS Perform. Eval. Rev.*, 41:89–94, 1 2014. ISSN 0163-5999. doi: 10.1145/2567529.2567556.

Zanini, F., Atienza, D., Benini, L., and Micheli, G. D. Multicore thermal management with model predictive control. pp. 711–714, 2009. doi: 10.1109/ECCTD.2009.5275073.

Zanini, F., Atienza, D., Micheli, G. D., and Boyd, S. P. Online convex optimization-based algorithm for thermal management of mpsocs. pp. 203–208. Association for Computing Machinery, 2010. ISBN 9781450300124. doi: 10.1145/1785481.1785532.

Zhang, L., Jiang, W., Lu, S., and Yang, T. Revisiting smoothed online learning. *Advances in Neural Information Processing Systems*, 34:13599–13612, 2021.

Zhang, L., Jiang, W., Yi, J., and Yang, T. Smoothed online convex optimization based on discounted-normal-predictor. volume 35, pp. 4928–4942. Curran Associates, Inc., 2022.

A Detailed Literature Review

A.1 Stochastic Smoothed Online Optimization

It is noteworthy that the literature in smoothed online optimization focuses on achieving strong performance in adversarial environments, without addressing the possibility of a stochastic environment, let alone seeking best-of-both-worlds algorithms. In the LQR control literature, there have been some works that consider noisy inputs (Chen et al., 2015; 2016). However, their noise model and dynamics are very specific, while our set-up considers stochasticity in more generality. In other problems involving dynamical systems, stochastic inputs have been considered, with prominent example being staff/inventory management (Gandhi et al., 2010; Jennings et al., 1996) and power systems (Urgaonkar et al., 2011; Koutsopoulos & Tassiulas, 2011).

In particular, within the realm of multi-arm bandits (MAB), (Bubeck & Slivkins, 2012) presents an algorithm that provides robust performance guarantees for both Independent and Identically Distributed (IID) reward sequences as well as adversarial reward sequences. In the context of Online Convex Optimization (OCO) literature, recent attention (Chen et al., 2023; Sachs et al., 2022; 2023) has been directed towards bridging the gap between an IID environment, which is relevant to stochastic (stationary) optimization, and an adversarial environment, representing the traditional OCO problem. However, it is important to note that these works *do not address switching costs*, and the assumptions regarding hitting costs and action spaces differ between OCO and ours.

A.2 Adversarial Smoothed Online Optimization

The body of research on smoothed online optimization is predominantly centered around adversarial results and can be broadly divided into two distinct communities. The first community examines it through the lens of Metrical Task Systems and employs the *competitive ratio* as the performance metric. Works of (Lin et al., 2011; Bansal et al., 2015) gave the first results considering the one-dimensional case. Further research (Chen et al., 2018) established that assumptions stronger than convexity, like α-polyhedrality or $m-$strong convexity are required for meaningful results. (Zhang et al., 2021) proved for α-polyhedral hitting costs and ℓ_1-norm switching costs that following the minimizer has a very good competitive ratio of $\left(\max\left\{1, \frac{2}{\alpha}\right\}\right)$. (Goel & Wierman, 2019) gave an order-optimal algorithm ROBD for $m-$strongly convex hitting costs and ℓ_2-norm switching costs with a competitive ratio of $1 + \mathcal{O}\left(\frac{1}{\sqrt{m}}\right)$. The analytical tools present in these works, like potential function analysis and receding horizon control, focus on guaranteeing worst-case performance, and do not take advantage of structure in the hitting costs.

Apart from the competitive ratio, regret with respect to the dynamic, hindsight optimal sequence of decisions has been another important performance metric, popular in the OCO literature. In the adversarial setting, (Andrew et al., 2013) showed that for any online algorithm, it is impossible to simultaneously achieve a finite competitive ratio and a sublinear regret. Subsequent research aimed to establish a dynamic regret that scales in relation to the path-length, denoted as L_T, of the optimal action sequence. For scenarios involving α-polyhedral convex hitting costs and ℓ_2-norm switching costs, (Chen et al., 2018) demonstrated a dynamic regret of $\mathcal{O}\left(\sqrt{TL_T}\right)$ for a modified version of OBD. In cases with strongly convex hitting costs and squared ℓ_2-norm switching costs, (Li et al., 2018) established a lower bound of $\mathcal{O}\left(L_T\right)$ on the dynamic regret. This bound was closely matched by (Goel et al., 2019). Finally, (Goel et al., 2019) demonstrated that ROBD achieves a dynamic regret of $\mathcal{O}\left(\sqrt{TL_T}\right)$ in these settings. The problem set-up in these results is different than the MTS-style ones, and there is very little on their connection (Andrew et al., 2013).

It is worth noting that these dynamic regret results rely on the strong assumption that the action space has a bounded diameter. Since we are dealing with metric switching costs, the absence of a boundedness assumption makes the problem much more difficult. Considering we have a stochastic setting in this work, the absence of such assumptions generalizes our results to any distribution over $\mathbb{R}^d$ with finite second moment (light or heavy tail), something one cannot achieve in the OCO setting.

B Main Proofs

In this section, we shed light on the analysis techniques employed in this work, through detailed proofs of some of our main contributions. These include the stochastic optimality of LAI, the poor performance of ROBD algorithm, the near-optimality of LAI(γ) in both stochastic and adversarial settings, and our general adversarial analysis framework.

B.1 Proof of Theorem 3.1

To prove these results, we show that LAI is a dynamic programming solution in our stochastic setting (2.2); thus proving it to be an optimal online algorithm. The first step is to show the optimality of the action at round T, given by the following lemma.

Lemma B.1. *The online optimal action at round T is*

$$x_T^{\text{LAI}} = C_T x_{T-1}^{\text{LAI}} + (I - C_T)\, v_T$$

and, consequently, the value function (2.1) for round T is

$$V_T\left(x_{T-1}^{\text{LAI}}, v_T\right) = \frac{1}{2}\left(x_{T-1}^{\text{LAI}} - v_T\right)^T (I - C_T)\left(x_{T-1}^{\text{LAI}} - v_T\right)$$

At round T, the optimal decision is

$$x_T = \operatorname*{argmin}_x \frac{1}{2}(x - v_T)^T A(x - v_T) + \frac{1}{2}\|x - x_{T-1}^{\text{LAI}}\|_2^2. \tag{B.1}$$

Proof. Differentiating the objective,

$$A(x - v_T) + (x - x_{T-1}^{\text{LAI}}) = 0 \tag{B.2}$$

gives the online optimal action at round T as

$$\begin{aligned} x_T^{\text{LAI}} &= (A+I)^{-1} x_{T-1}^{\text{LAI}} + (A+I)^{-1} A v_T && \text{(B.3)}\\ &= (A+I)^{-1} x_{T-1}^{\text{LAI}} + (I - (A+I)^{-1}) v_T && \text{(B.4)}\\ &= C_T x_{T-1}^{\text{LAI}} + (I - C_T) v_T. && \text{(B.5)} \end{aligned}$$

Putting x_T^{LAI} back into the current cost gives the value function at T,

$$\begin{aligned} V_T(x_{T-1}^{\text{LAI}}, v_T) =& \frac{1}{2}(x_T - v_T)^T A(x_T - v_T) + \frac{1}{2}\|x_T - x_{T-1}\|_2^2 && \text{(B.6)}\\ =& \frac{1}{2}(x_{T-1}^{\text{LAI}} - v_T)^T \left((A+I)^{-1} A (A+I)^{-1}\right)(x_{T-1}^{\text{LAI}} - v_T) \\ &+ \frac{1}{2}(x_{T-1}^{\text{LAI}} - v_T)^T \left((A+I)^{-1} A^2 (A+I)^{-1}\right)(x_{T-1}^{\text{LAI}} - v_T) && \text{(B.7)}\\ =& \frac{1}{2}(x_{T-1}^{\text{LAI}} - v_T)^T \left((A+I)^{-1} A\right)(x_{T-1}^{\text{LAI}} - v_T) && \text{(B.8)}\\ =& \frac{1}{2}(x_{T-1}^{\text{LAI}} - v_T)^T (I - C_T)(x_{T-1}^{\text{LAI}} - v_T) && \text{(B.9)} \end{aligned}$$

□

We next present the main element of the proof of LAI's optimality in the form of the following proposition, from which Theorem 3.1 directly follows.

Proposition B.2. *Consider hitting costs $\frac{1}{2}(x - v_t)^T A(x - v_t)$ and switching cost is $\frac{1}{2}\|x_t - x_{t-1}\|_2^2$ with the sequence of minimizers being a martingale, that is (2.2). The exact value function at $T - t$ is*

$$\begin{aligned} V_{T-t}(x_{T-t-1}^{\text{LAI}}, v_{T-t}) =& \frac{1}{2}(x_{T-t-1}^{\text{LAI}} - v_{T-t})^T (I - C_{T-t})(x_{T-t-1}^{\text{LAI}} - v_{T-t}) \\ &+ \sum_{s=0}^{t-1} \frac{1}{2}\mathbb{E}\left[(v_{T-s} - v_{T-s-1})^T (I - C_{T-s})(v_{T-s} - v_{T-s-1}) | \mathcal{F}_{T-t}\right] \end{aligned}$$

Proof. The proof follows via an induction argument. The value function at round T, according to Lemma B.1, is

$$V_T(x_{T-1}^{\text{LAI}}, v_T) = \frac{1}{2}(x_{T-1}^{\text{LAI}} - v_T)^T (I - C_T)(x_{T-1}^{\text{LAI}} - v_T). \tag{B.10}$$

The proposition is, therefore, true for $t = 0$ (sum from 0 to -1 is zero). Assuming the proposition for some $t \geq 0$, we have the optimal action at round $T - t - 1$ as

$$\begin{aligned} x_{T-t-1}^{\text{LAI}} = \operatorname*{argmin}_x \Big\{ & \frac{1}{2}(x - v_{T-t-1})^T A(x - v_{T-t-1}) + \frac{1}{2}\|x - x_{T-t-2}^{\text{LAI}}\|_2^2 \\ & + \mathbb{E}[V_{T-t}(x, v_{T-t})|\mathcal{F}_{T-t-1}] \Big\} \end{aligned} \tag{B.11}$$

$$\begin{aligned} = \operatorname*{argmin}_x \Big\{ & \frac{1}{2}(x - v_{T-t-1})^T A(x - v_{T-t-1}) + \frac{1}{2}\|x - x_{T-t-2}^{\text{LAI}}\|_2^2 \\ & + \frac{1}{2}\mathbb{E}\left[(x - v_{T-t})^T (I - C_{T-t})(x - v_{T-t})|\mathcal{F}_{T-t-1}\right] \\ & + \sum_{s=0}^{t-1} \frac{1}{2}\mathbb{E}\left[(v_{T-s} - v_{T-s-1})^T (I - C_{T-s})(v_{T-s} - v_{T-s-1})|\mathcal{F}_{T-t-1}\right] \Big\} \end{aligned} \tag{B.12}$$

$$\begin{aligned} = \operatorname*{argmin}_x \Big\{ & \frac{1}{2}(x - v_{T-t-1})^T (A + I - C_{T-t})(x - v_{T-t-1}) + \frac{1}{2}\|x - x_{T-t-2}^{\text{LAI}}\|_2^2 \\ & + (x - v_{T-t-1})^T (I - C_{T-t})\mathbb{E}[v_{T-t-1} - v_{T-t}|\mathcal{F}_{T-t-1}] \\ & + \frac{1}{2}\mathbb{E}[(v_{T-t} - v_{T-t-1})^T (I - C_{T-t})(v_{T-t} - v_{T-t-1})|\mathcal{F}_{T-t-1}] \\ & + \sum_{s=0}^{t-1} \frac{1}{2}\mathbb{E}\left[(v_{T-s} - v_{T-s-1})^T (I - C_{T-s})(v_{T-s} - v_{T-s-1})|\mathcal{F}_{T-t-1}\right] \Big\} \end{aligned} \tag{B.13}$$

$$\begin{aligned} = \operatorname*{argmin}_x \Big\{ & \frac{1}{2}(x - v_{T-t-1})^T (A + I - C_{T-t})(x - v_{T-t-1}) + \frac{1}{2}\|x - x_{T-t-2}^{\text{LAI}}\|_2^2 \\ & + \sum_{s=0}^{t} \frac{1}{2}\mathbb{E}\left[(v_{T-s} - v_{T-s-1})^T (I - C_{T-s})(v_{T-s} - v_{T-s-1})|\mathcal{F}_{T-t-1}\right] \Big\} \end{aligned} \tag{B.14}$$

Differentiating the objective of the optimization problem above, we get

$$(A + I - C_{T-t})(x - v_{T-t-1}) + (x - x_{T-t-2}^{\text{LAI}}) = 0 \tag{B.15}$$

giving us the optimal action at round $T - t - 1$ as

$$x_{T-t-1}^{\text{LAI}} = (2I + A - C_{T-t})^{-1} x_{T-t-2}^{\text{LAI}} + \left(I - (2I + A - C_{T-t})^{-1}\right) v_{T-t-1} \tag{B.16}$$

$$= C_{T-t-1} x_{T-t-2}^{\text{LAI}} + (I - C_{T-t-1})\, v_{T-t-1} \tag{B.17}$$

where the invertibility of $(2I + A - C_{T-t})$ is easy to prove and is provided in the appendix. Consequently, the value function at round $T - t - 1$ as

$$\begin{aligned} & V_{T-t-1}(x_{T-t-2}^{\text{LAI}}, v_{T-t-1}) \\ & = \frac{1}{2}(x_{T-t-1}^{\text{LAI}} - v_{T-t-1})^T A(x_{T-t-1}^{\text{LAI}} - v_{T-t-1}) + \frac{1}{2}\|x_{T-t-1}^{\text{LAI}} - x_{T-t-2}^{\text{LAI}}\|_2^2 \\ & \quad + \mathbb{E}[V_{T-t}(x_{T-t-1}^{\text{LAI}}, v_{T-t})|\mathcal{F}_{T-t-1}] \end{aligned} \tag{B.18}$$

$$\begin{aligned} & = \frac{1}{2}(x_{T-t-2}^{\text{LAI}} - v_{T-t-1})^T C_{T-t-1}(C_{T-t-1}^{-1} - I) C_{T-t-1}(x_{T-t-2}^{\text{LAI}} - v_{T-t-1}) \\ & \quad + \frac{1}{2}(x_{T-t-2}^{\text{LAI}} - v_{T-t-1})^T (I - C_{T-t-1})^2 (x_{T-t-2}^{\text{LAI}} - v_{T-t-1}) \\ & \quad + \sum_{s=0}^{t} \frac{1}{2}\mathbb{E}\left[(v_{T-s} - v_{T-s-1})^T (I - C_{T-s})(v_{T-s} - v_{T-s-1})|\mathcal{F}_{T-t-1}\right] \end{aligned} \tag{B.19}$$

$$
\begin{aligned}
&=\frac{1}{2}(x^{\text{LAI}}_{T-t-2}-v_{T-t-1})^T(I-C_{T-t-1})(x^{\text{LAI}}_{T-t-2}-v_{T-t-1})\\
&\quad+\sum_{s=0}^{t}\frac{1}{2}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-C_{T-s})(v_{T-s}-v_{T-s-1})|\mathcal{F}_{T-t-1}\right]
\end{aligned}
\tag{B.20}
$$

proving the proposition through induction. The cost of LAI algorithm, which is also the online optimal cost, is

$$
\mathbb{E}\left[\text{Cost}_{\text{LAI}}[1,T]\right]=\mathbb{E}[V_1(x_0,v_1)|\mathcal{F}_0]=\sum_{t=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_t-v_{t-1})^T(I-C_t)(v_t-v_{t-1})|\mathcal{F}_0\right]
\tag{B.21}
$$

where $\mathcal{F}_0$ is the trivial sigma field and $v_0=x_0$. The upper bound in terms of C_L is a consequence of the following important observation.

Lemma B.3. *The matrix sequence in* LAI *algorithm, that is, $\{C_t\}_t$ and C_L are related as*

$$
C_L\prec C_1\prec C_2\prec\ldots\prec C_{T-1}\prec C_T.
$$

The proof follows from the recursion satisfied by $\{C_t\}_t$, that is, $C_{t-1}^{-1}=2I+A-C_t$ and the fact that C_L is the stationary point of this recursion. We prove this lemma, along with other interesting properties of $\{C_t\}_{t=1}^T$, in Corollary D.5. □

B.2 Proof of Theorem 3.7

We split the proof into four parts, one of which will present the exact characterization of the lower bound. The first step is the cost of the FI algorithm, as the stated in the proposition below.

Proposition B.4. *The following is the expected cost of the* FI *algorithm, with matrix C for martingale minimizers* (2.2) *with increments having same covariance matrix, Σ,*

$$
\mathbb{E}[\textit{Cost}_{\text{FI}}[1,T]]\geq\sum_{t=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_t-v_{t-1})^T\left(I-C^2\right)^{-1}(AC^2+(I-C)^2)(v_t-v_{t-1})\right]-\left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2}
$$

The proof of this involves an induction argument to get the cost of the last t rounds and is presented right after the theorem proof in Appendix B.3 to focus on the regret lower bound. The next step is to compare it to the cost of the LAI algorithm to get the following lower bound to the regret.

Theorem B.5. *Consider the hitting costs, switching costs and stochastic setting* (2.2) *from the problem set-up with (v_t-v_{t-1}) having covariance matrix Σ. The following is the regret of the* FI *algorithm,*

$$
\mathbb{E}[\textit{Regret}_{\text{FI}}[1,T]]\geq\frac{T}{2}\sum_{j\notin Z}\left(w\left(\lambda^A_j,\lambda^C_j\right)-\left(1-\lambda^L_j\right)\right)(e_j^TP^T\Sigma Pe_j)-\left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2}.
$$

where $w(\alpha,c)=\frac{\alpha c^2+(1-c)^2}{1-c^2}$, $Z=\{i\in\{1,\ldots,d\}:\lambda^C_i=\lambda^L_i\}$, P is the modal matrix of A and, $\lambda^A_j,\lambda^C_j,\lambda^L_j$ are the jth eigenvalues of A,C,C_L respectively.

Proof. We can write $A=PD_AP^T$, $C_L=PD_LP^T$ and $C=PDP^T$, where P is the modal matrix of A and the middle matrix in each being the diagonal matrix of respective eigenvalues.

$$
\begin{aligned}
&\mathbb{E}[\text{Cost}_{\text{FI}}[1,T]]-\mathbb{E}[\text{Cost}_{\text{LAI}}[1,T]]\\
&\geq\sum_{t=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_t-v_{t-1})^T\left(I-C^2\right)^{-1}(AC^2+(I-C)^2)(v_t-v_{t-1})\right]\\
&\quad-\left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2}-\sum_{t=1}^{T}\mathbb{E}[(v_t-v_{t-1})^T(1-C_L)(v_t-v_{t-1})]
\end{aligned}
\tag{B.22}
$$

$$\geq \sum_{t=1}^{T} \frac{1}{2}\mathbb{E}\left[(v_t - v_{t-1})^T P \left(\left(I - D^2\right)^{-1}\left(D_A D^2 + (I-D)^2\right) - (I - D_L)\right) P^T (v_t - v_{t-1})\right] - \left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2} \tag{B.23}$$

Observe that we can write

$$\left(I - D^2\right)^{-1}\left(D_A D^2 + (I-D)^2\right) - (I - D_L) \tag{B.24}$$

$$= diag\left(w\left(\lambda^A_1, \lambda^C_1\right) - \left(1 - \lambda^L_1\right), \ldots, w\left(\lambda^A_d, \lambda^C_d\right) - \left(1 - \lambda^L_d\right)\right) \tag{B.25}$$

We define the set of indices $Z = \{i \in \{1, \ldots, d\} : \lambda^C_i = \lambda^L_i\}$. The following lemma shows that $\left(w\left(\lambda^A_j, \lambda^C_j\right) - \left(1 - \lambda^L_j\right)\right) > 0 \; \forall \, j \notin Z$.

Lemma B.6. *The minimum of $w(\alpha, c)$ has the following properties*

$$\operatorname*{argmin}_{c\in[0,1]} w(\alpha, c) = \frac{\alpha + 2 - \sqrt{\alpha^2 + 4\alpha}}{2} = \frac{2}{\alpha + 2 + \sqrt{\alpha^2 + 4\alpha}} = c_L \;\; \textit{and} \;\; \min_{c\in[0,1]} w(\alpha, c) = 1 - c_L.$$

Now, using Lemma B.6, $w\left(\lambda^A_i, \lambda^C_i\right) - \left(1 - \lambda^L_i\right) > 0$ if $i \notin Z$ and $w\left(\lambda^A_i, \lambda^C_i\right) - \left(1 - \lambda^L_i\right) = 0$ if $i \in Z$. There has to be i such that $i \notin Z$. Otherwise, C and C_L will have all eigenvalues and their corresponding eigenvectors same, making them equal. Denoting $P^T(v_t - v_{t-1}) = y_t$, the gap of FI to LAI is

$$Regret_{\text{FI}}[1, T] \geq \sum_{t=1}^{T} \frac{1}{2}\mathbb{E}\left[y_t^T diag\left(w\left(\lambda^A_1, \lambda^C_1\right) - \left(1 - \lambda^L_1\right), \ldots, w\left(\lambda^A_d, \lambda^C_d\right) - \left(1 - \lambda^L_d\right)\right) y_t\right] - \left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2} \tag{B.26}$$

$$= \sum_{t=1}^{T} \frac{1}{2}\sum_{j\notin Z}\left(w\left(\lambda^A_j, \lambda^C_j\right) - \left(1 - \lambda^L_j\right)\right)\mathbb{E}(y_t)^2_j - \left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2} \tag{B.27}$$

$$= \frac{T}{2}\sum_{j\notin Z}\left(w\left(\lambda^A_j, \lambda^C_j\right) - \left(1 - \lambda^L_j\right)\right)\mathbb{E}(y)^2_j - \left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2} \tag{B.28}$$

where

$$\mathbb{E}(y_t)^2_j = \mathbb{E}[(e_j^T y_t)^2] \tag{B.29}$$

$$= \mathbb{E}[e_j^T y_t y_t^T e_j] \tag{B.30}$$

$$= \mathbb{E}[e_j^T P^T (v_t - v_{t-1})(v_t - v_{t-1})^T P e_j] \tag{B.31}$$

$$= e_j^T P^T \Sigma P e_j. \tag{B.32}$$

Therefore,

$$\mathbb{E}[\text{Cost}_{\text{FI}}[1, T]] - \mathbb{E}[\text{Cost}_{\text{LAI}}[1, T]] \geq \frac{T}{2}\sum_{j\notin Z}\left(w\left(\lambda^A_j, \lambda^C_j\right) - \left(1 - \lambda^L_j\right)\right)\left(e_j^T P^T \Sigma P e_j\right) - \left(\frac{\left(\lambda^C_{\max}\right)^2}{1-\left(\lambda^C_{\max}\right)^2}\right)\frac{\sigma^2}{2}. \tag{B.33}$$

□

Proof of Theorem 3.7. Lastly, we show $(e_j^T P^T \Sigma P e_j) > 0$ for both cases of Σ. Let's consider the second case where A and Σ have same eigenvectors. We have $\Sigma = P D_\Sigma P^T$ giving us

$$\mathbb{E}(y)^2_j = e_j^T P^T \Sigma P e_j = D_\Sigma(j, j) \tag{B.34}$$

and

$$Regret_{\text{FI}}[1,T] \geq \frac{T}{2}\sum_{j\notin Z}\left(w\left(\lambda_j^A,\lambda_j^C\right)-\left(1-\lambda_j^L\right)\right)D_\Sigma(j,j)-\left(\frac{\left(\lambda_{\max}^C\right)^2}{1-\left(\lambda_{\max}^C\right)^2}\right)\frac{\sigma^2}{2}. \tag{B.35}$$

Now, $(D-D_L)D_\Sigma = 0_{d\times d} \implies (C_{\text{ROBD}}-C_L)\Sigma = 0_{d\times d}$. Therefore, the condition in the second case, that is, $(C-C_L)\Sigma \neq 0_{d\times d}$ implies, $\bar{D}=(D-D_L)D_\Sigma \neq 0_{d\times d}$, meaning, $\bar{D}(j,j)=0$ if $j\in Z$ and $\bar{D}(j,j)=\left(\lambda_j^C-\lambda_j^L\right)D_\Sigma(j,j)$ if $j\notin Z$. For $\bar{D}\neq 0$, there has to be some $j\notin Z$ such that $D_\Sigma(j,j)>0$ proving that FI has $\Omega(T)$ regret.

For the case where A and Σ different eigenvectors, Lemma B.7 below shows that $\mathbb{E}(y)_j^2>0\,\forall\, j$, giving $\Omega(T)$ regret.

Lemma B.7. *Consider $A=PD_AP^T$, where D_A is a diagonal matrix and a random vector $u_t\in\mathbb{R}^d$ with covariance matrix Σ. Suppose Σ has modal matrix $Q\neq P$. Then P^Tu_t satisfies*

$$\mathbb{E}\left[e_j^T\left(P^Tu_tu_t^TP\right)e_j\right]=e_j^TP^T\Sigma Pe_j>0\,\forall\, j,t$$

Proof. We can write $\Sigma=QD_\Sigma Q^T$ where Q is the matrix having eigenvectors of Σ as columns and D_Σ is the diagonal matrix of eigenvalues of Σ. We will have $Q\neq P$ and

$$e_j^TP^T\Sigma Pe_j = e_j^TP^TQD_\Sigma Q^TPe_j \tag{B.36}$$

$$= \sum_{k=1}^{d} D_\Sigma(k,k)(Q^TPe_j)_k^2 \tag{B.37}$$

Now, notice that Q^TPe_j is the decomposition of the jth eigenvector in P into the components along the eigenbasis defined by Q. Since, P and Q are different eigenbases, Pe_j cannot be orthogonal to any eigenvectors in Q, that is any row of Q^T. Therefore each element of Q^TPe_j is non-zero, giving us

$$(Q^TPe_j)_k^2>0\,\forall\, k. \tag{B.38}$$

There has to be at least one positive diagonal element in D_Σ, otherwise $\Sigma=0$ (as $\Sigma\succcurlyeq 0$). This proves

$$e_j^TP^T\Sigma Pe_j>0\,\forall\, j \tag{B.39}$$

□

B.3 Proof of Proposition B.4

We calculate the cost of the following iterative algorithm,

$$x_t = Cx_{t-1}+(I-C)v_t \tag{B.40}$$

where $C\prec I$. The following claim holds.

Claim: The cost from round $T-t$ to T is

$$\begin{aligned}&\mathbb{E}[\text{Cost}_{\text{FI}}[T-t,T]]\\ =&\frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t})^T(AC^2+(I-C)^2)\left(\sum_{i=0}^{t}C^{2i}\right)(x_{T-t-1}-v_{T-t})\right]\\ &+\sum_{s=1}^{t}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(I-C^{2s})(v_{T-s+1}-v_{T-s})\right]\end{aligned} \tag{B.41}$$

We prove this claim by induction. First, we look at $t=0$ or cost at round T.

$$\begin{aligned}\mathbb{E}[\text{Cost}_{\text{FI}}[T,T]] =&\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^TCAC(x_{T-1}-v_T)\right]\\ &+\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T(I-C)(I-C)(x_{T-1}-v_T)\right]\end{aligned} \tag{B.42}$$

$$= \frac{1}{2}\mathbb{E}\left[(x_{T-1} - v_T)^T (AC^2 + (I-C)^2)(x_{T-1} - v_T)\right] \tag{B.43}$$

Now, $A, C, (I-C)$ are all real symmetric matrices with the **same modal matrix** P. This means their **product is commutative**, which is used in the second inequality here.

Therefore, the claim holds for $t = 0$ ($C^0 = I$ and sum from 1 to 0 is zero). Further, we take $0^0 = 1$. Assuming the claim holds for some $t \geq 0$, we have

$$\begin{aligned}
&\mathbb{E}[\text{Cost}_{\text{FI}}[T-t, T]] \\
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-1} - v_{T-t})^T (AC^2 + (I-C)^2)\left(\sum_{i=0}^{t} C^{2i}\right)(x_{T-t-1} - v_{T-t})\right] \\
&+ \sum_{s=1}^{t} \frac{1}{2}\mathbb{E}\left[(v_{T-s+1} - v_{T-s})^T (I-C^2)^{-1}(AC^2 + (I-C)^2)(I-C^{2s})(v_{T-s+1} - v_{T-s})\right]
\end{aligned} \tag{B.44}$$

$$\begin{aligned}
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-1} - v_{T-t-1})^T (AC^2 + (I-C)^2)\left(\sum_{i=0}^{t} C^{2i}\right)(x_{T-t-1} - v_{T-t-1})\right] \\
&+ \mathbb{E}\left[(x_{T-t-1} - v_{T-t-1})^T (AC^2 + (I-C)^2)\left(\sum_{i=0}^{t} C^{2i}\right)\mathbb{E}[v_{T-t} - v_{T-t-1}|\mathcal{F}_{T-t-1}]\right] \\
&+ \frac{1}{2}\mathbb{E}\left[(v_{T-t} - v_{T-t-1})^T (AC^2 + (I-C)^2)\left(\sum_{i=0}^{t} C^{2i}\right)(v_{T-t} - v_{T-t-1})\right] \\
&+ \sum_{s=1}^{t} \frac{1}{2}\mathbb{E}\left[(v_{T-s+1} - v_{T-s})^T (I-C^2)^{-1}(AC^2 + (I-C)^2)(I-C^{2s})(v_{T-s+1} - v_{T-s})\right]
\end{aligned} \tag{B.45}$$

$$\begin{aligned}
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-1} - v_{T-t-1})^T (AC^2 + (I-C)^2)\left(\sum_{i=0}^{t} C^{2i}\right)(x_{T-t-1} - v_{T-t-1})\right] \\
&+ \frac{1}{2}\mathbb{E}\left[(v_{T-t} - v_{T-t-1})^T (AC^2 + (I-C)^2)(I-C^2)^{-1}\left(I - C^{2(t+1)}\right)(v_{T-t} - v_{T-t-1})\right] \\
&+ \sum_{s=1}^{t} \frac{1}{2}\mathbb{E}\left[(v_{T-s+1} - v_{T-s})^T (I-C^2)^{-1}(AC^2 + (I-C)^2)(I-C^{2s})(v_{T-s+1} - v_{T-s})\right]
\end{aligned} \tag{B.46}$$

$$\begin{aligned}
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-1} - v_{T-t-1})^T (AC^2 + (I-C)^2)\left(\sum_{i=0}^{t} C^{2i}\right)(x_{T-t-1} - v_{T-t-1})\right] \\
&+ \sum_{s=1}^{t+1} \frac{1}{2}\mathbb{E}\left[(v_{T-s+1} - v_{T-s})^T (I-C^2)^{-1}(AC^2 + (I-C)^2)(I-C^{2s})(v_{T-s+1} - v_{T-s})\right]
\end{aligned} \tag{B.47}$$

The cost from round $T-t-1$ to T, therefore, is

$$\begin{aligned}
&\mathbb{E}[\text{Cost}_{\text{FI}}[T-t-1, T]] \\
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-1} - v_{T-t-1})^T A(x_{T-t-1} - v_{T-t-1})\right] + \frac{1}{2}\mathbb{E}\left[\|x_{T-t-1} - x_{T-t-2}\|_2^2\right] \\
&+ \mathbb{E}[C_{\text{FI}}[T-t, T]]
\end{aligned} \tag{B.48}$$

$$\begin{aligned}
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-1} - v_{T-t-1})^T \left((AC^2 + (I-C)^2)\left(\sum_{i=0}^{t} C^{2i}\right) + A\right)(x_{T-t-1} - v_{T-t-1})\right] \\
&+ \frac{1}{2}\mathbb{E}\left[\|x_{T-t-1} - x_{T-t-2}\|_2^2\right] \\
&+ \sum_{s=1}^{t+1} \frac{1}{2}\mathbb{E}\left[(v_{T-s+1} - v_{T-s})^T (I-C^2)^{-1}(AC^2 + (I-C)^2)(I-C^{2s})(v_{T-s+1} - v_{T-s})\right]
\end{aligned} \tag{B.49}$$

$$
\begin{aligned}
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-2}-v_{T-t-1})^T C\left((AC^2+(I-C)^2)\left(\sum_{i=0}^{t}C^{2i}\right)+A\right)C(x_{T-t-2}-v_{T-t-1})\right]\\
&+\frac{1}{2}\mathbb{E}\left[(x_{T-t-2}-v_{T-t-1})^T(I-C)^2(x_{T-t-2}-v_{T-t-1})\right]\\
&+\sum_{s=1}^{t+1}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(I-C^{2s})(v_{T-s+1}-v_{T-s})\right]
\end{aligned}
\tag{B.50}
$$

$$
\begin{aligned}
=&\frac{1}{2}\mathbb{E}\Bigg[(x_{T-t-2}-v_{T-t-1})^T\left((AC^2+(I-C)^2)\left(\sum_{i=1}^{t+1}C^{2i}\right)+AC^2+(I-C)^2\right)\dots\\
&(x_{T-t-2}-v_{T-t-1})\Bigg]\\
&+\sum_{s=1}^{t+1}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(I-C^{2s})(v_{T-s+1}-v_{T-s})\right]
\end{aligned}
\tag{B.51}
$$

$$
\begin{aligned}
=&\frac{1}{2}\mathbb{E}\left[(x_{T-t-2}-v_{T-t-1})^T(AC^2+(I-C)^2)\left(\sum_{i=0}^{t+1}C^{2i}\right)(x_{T-t-2}-v_{T-t-1})\right]\\
&+\sum_{s=1}^{t+1}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(I-C^{2s})(v_{T-s+1}-v_{T-s})\right]
\end{aligned}
\tag{B.52}
$$

Now, to complete the proof, the total cost is

$$
\begin{aligned}
&\mathbb{E}[\text{Cost}_{\text{FI}}[1,T]]\\
=&\frac{1}{2}\mathbb{E}\left[(x_0-v_1)^T(I-C^2)^{-1}(AC^2+(I-C)^2)(I-C^{2T})(x_0-v_1)\right]\\
&+\sum_{s=1}^{T-1}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(I-C^{2s})(v_{T-s+1}-v_{T-s})\right]
\end{aligned}
\tag{B.53}
$$

$$
\begin{aligned}
=&\sum_{s=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(v_{T-s+1}-v_{T-s})\right]\\
&-\sum_{s=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T C^{2s}(v_{T-s+1}-v_{T-s})\right]
\end{aligned}
\tag{B.54}
$$

$$
\begin{aligned}
\geq&\sum_{s=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(v_{T-s+1}-v_{T-s})\right]\\
&-\sum_{s=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T\left(\lambda_{\max}^C\right)^{2s}I(v_{T-s+1}-v_{T-s})\right]
\end{aligned}
\tag{B.55}
$$

$$
\begin{aligned}
\geq&\sum_{s=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_{T-s+1}-v_{T-s})^T(I-C^2)^{-1}(AC^2+(I-C)^2)(v_{T-s+1}-v_{T-s})\right]\\
&-\left(\frac{\left(\lambda_{\max}^C\right)^2}{1-\left(\lambda_{\max}^C\right)^2}\right)\frac{\sigma^2}{2}
\end{aligned}
\tag{B.56}
$$

$$
\begin{aligned}
\geq&\sum_{t=1}^{T}\frac{1}{2}\mathbb{E}\left[(v_t-v_{t-1})^T\left(I-C^2\right)^{-1}(AC^2+(I-C)^2)(v_t-v_{t-1})\right]\\
&-\left(\frac{\left(\lambda_{\max}^C\right)^2}{1-\left(\lambda_{\max}^C\right)^2}\right)\frac{\sigma^2}{2}
\end{aligned}
\tag{B.57}
$$

□

B.4 Proof of Theorem 3.11 (i)

The analysis of the upper bound on the regret of $\text{LAI}(\gamma)$ involves two key steps. The first is an upper bound on the cost of $\text{LAI}(\gamma)$, as stated in the following proposition.

Proposition B.8. *The total cost of* $\text{LAI}(\gamma)$ *for martingale minimizers* (2.2) *is upper bounded as*

$$\mathbb{E}[Cost_{\text{LAI}(\gamma)}[1,T]] \leq \frac{1}{2}\sum_{t=1}^{T}\mathbb{E}\left[(v_t - v_{t-1})^T(I-\tilde{C}_t)(v_t - v_{t-1})\right]. \tag{B.58}$$

The proof of this proposition uses an induction argument by cleverly employing the recursion followed by $\{\tilde{C}_t\}_t$ and is proved right after the theorem proof in Appendix B.5. The second step uses the following key property exhibited by eigenvalues of $\{\tilde{C}_t\}_t$ of $\text{LAI}(\gamma)$ and $\{C_t\}_t$ of LAI.

Lemma B.9. *The ith eigenvalue of* $\text{LAI}(\gamma)$*'s $\tilde{C}_t$, that is $\tilde{\lambda}_i^t$, and the ith eigenvalue of* LAI*'s C_t, that is, λ_i^t, are related in the following manner*

$$\left(\lambda_i^t - \tilde{\lambda}_i^t\right) \leq \frac{\lambda_i^A}{2}\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\left(\lambda_i^T\right)^{2(T-t+1)}$$

The proof of this lemma uses the recursion followed by both $\{\tilde{C}_t\}_t$ and $\{\tilde{C}_t\}_t$ and is presented in Appendix F.1. We now proceed to upper bound the regret of $\text{LAI}(\gamma)$ algorithm.

Proof of Theorem 3.11. We write each C_t and $\tilde{C}_t$ as PD_tP^T and $P\tilde{D}_tP^T$ where D_t and $\tilde{D}_t$ are the eigenvalue diagonal matrices of C_t and $\tilde{C}_t$ respectively. Therefore, regret of $\text{LAI}(\gamma)$ is **upper-bounded** as

$$\mathbb{E}[Cost_{\text{LAI}(\gamma)}[1,T]] - \mathbb{E}[\text{Cost}_{\text{LAI}}[1,T]] \leq \frac{1}{2}\sum_{t=1}^{T}\mathbb{E}\left[(v_t - v_{t-1})^T(C_t-\tilde{C}_t)(v_t - v_{t-1})\right] \tag{B.59}$$

$$= \frac{1}{2}\sum_{t=1}^{T}\mathbb{E}\left[(v_t - v_{t-1})^T P(D_t-\tilde{D}_t)P^T(v_t - v_{t-1})\right] \leq \frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{d}\left(\lambda_i^t - \tilde{\lambda}_i^t\right)\mathbb{E}(y_t)_i^2 \tag{B.60}$$

$$= \frac{1}{2}\sum_{i=1}^{d}\mathbb{E}(y)_i^2\sum_{t=1}^{T}\left(\lambda_i^t - \tilde{\lambda}_i^t\right) \leq \frac{1}{2}\sum_{i=1}^{d}\mathbb{E}(y)_i^2\sum_{t=1}^{T}\frac{\lambda_i^A}{2}\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\left(\lambda_i^T\right)^{2(T-t+1)} \tag{B.61}$$

$$\leq \frac{1}{2}\sum_{i=1}^{d}\mathbb{E}(y)_i^2\frac{\lambda_i^A}{2}\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\sum_{t=1}^{\infty}\left(\lambda_i^T\right)^{2t} = \frac{1}{2}\sum_{i=1}^{d}\mathbb{E}(y)_i^2\frac{\lambda_i^A}{2}\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\frac{\left(\lambda_i^T\right)^2}{1-\left(\lambda_i^T\right)^2} \tag{B.62}$$

$$= \frac{1}{2}\sum_{i=1}^{d}\mathbb{E}(y)_i^2\frac{\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)}{2\left(\lambda_i^A+2\right)} \tag{B.63}$$

$$\tag{B.64}$$

where $y_t = P^T(v_t - v_{t-1})$. Note that $\mathbb{E}(y_t)_i^2 = \mathbb{E}(y)_i^2$ is time-invariant as all $(v_t - v_{t-1})$ have same covariance Σ. The function $\left(\left(1+\frac{4}{q}\right)^{\frac{\gamma}{2}} - 1\right)/(2(q+2))$ is always decreasing in q for $\gamma \in [0,1]$. Therefore,

$$\text{Regret}_{\text{LAI}(\gamma)}[1,T] \leq \frac{1}{4}\frac{\left(\left(1+\frac{4}{\lambda_{\min}^A}\right)^{\frac{\gamma}{2}} - 1\right)}{\left(\lambda_{\min}^A+2\right)}\sum_{i=1}^{d}\mathbb{E}(y)_i^2 = \frac{\left(\left(1+\frac{4}{\lambda_{\min}^A}\right)^{\frac{\gamma}{2}} - 1\right)}{\left(\lambda_{\min}^A+2\right)}\frac{\sigma^2}{4} \tag{B.65}$$

where the last step uses $\sum_{i=1}^{d}\mathbb{E}(y)_i^2 = \sigma^2 = tr(\Sigma)$ which is stated and proved as Lemma F.1. □

B.5 Proof of Proposition B.8

Claim: The cost from round $T-t$ to T for martingale minimizers is upper bound by

$$\begin{aligned}\mathbb{E}[\text{Cost}_{\text{LAI}(\gamma)}[T-t,T]] \leq &\frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t})^T(I-\tilde{C}_{T-t})(x_{T-t-1}-v_{T-t})\right] \\ &+\frac{1}{2}\sum_{s=0}^{t-1}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right]\end{aligned} \tag{B.66}$$

For $t=0$, it is the expected cost of the final round, which is

$$\mathbb{E}\left[\text{Cost}_{\text{LAI}(\gamma)}[T,T]\right] = \frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T\left(A(\tilde{C}_T)^2+(I-\tilde{C}_T)^2\right)(x_{T-1}-v_T)\right] \tag{B.67}$$

$$\begin{aligned}=&\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T(I-\tilde{C}_T)(x_{T-1}-v_T)\right] \\ &+\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T\left(A(\tilde{C}_T)^2+(I-\tilde{C}_T)^2-(I-\tilde{C}_T)\right)(x_{T-1}-v_T)\right]\end{aligned} \tag{B.68}$$

$$\begin{aligned}=&\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T(I-\tilde{C}_T)(x_{T-1}-v_T)\right] \\ &+\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T\left(A(\tilde{C}_T)^2-\tilde{C}_T(I-\tilde{C}_T)\right)(x_{T-1}-v_T)\right]\end{aligned} \tag{B.69}$$

$$\begin{aligned}=&\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T(I-\tilde{C}_T)(x_{T-1}-v_T)\right] \\ &+\frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T\left(\tilde{C}_T(A+I)(\tilde{C}_T-(A+I)^{-1})\right)(x_{T-1}-v_T)\right]\end{aligned} \tag{B.70}$$

Now, observe two things. First that $\tilde{\lambda}_i^T < \frac{1}{\lambda_i^A+1}$ and second that $\tilde{C}_T$ has the same set of eigenvectors as A. Therefore, $\tilde{C}_T(A+I)(\tilde{C}_T-(A+I)^{-1})$ will be negative definite as $\tilde{C}_T \prec (A+I)^{-1}$, $\tilde{C}_T \succ 0$ and $(A+I) \succ I$. Therefore,

$$\mathbb{E}[\text{Cost}_{\text{LAI}(\gamma)}[T,T]] \leq \frac{1}{2}\mathbb{E}\left[(x_{T-1}-v_T)^T(I-\tilde{C}_T)(x_{T-1}-v_T)\right] \tag{B.71}$$

Therefore, the claim holds for $t=0$ as the sum from -1 to 0 is zero. Assuming the claim holds for some $t \geq 0$,

$$\begin{aligned}\mathbb{E}[\text{Cost}_{\text{LAI}(\gamma)}[T-t,T]] \leq &\frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t})^T(I-\tilde{C}_{T-t})(x_{T-t-1}-v_{T-t})\right] \\ &+\frac{1}{2}\sum_{s=0}^{t-1}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(1-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right]\end{aligned} \tag{B.72}$$

$$\begin{aligned}\leq &\frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t-1})^T(I-\tilde{C}_{T-t})(x_{T-t-1}-v_{T-t-1})\right] \\ &+\mathbb{E}\left[(x_{T-t-1}-v_{T-t-1})^T(I-\tilde{C}_{T-t})(v_{T-t-1}-v_{T-t})\right] \\ &+\frac{1}{2}\mathbb{E}\left[(v_{T-t-1}-v_{T-t})^T(I-\tilde{C}_{T-t})(v_{T-t-1}-v_{T-t})\right] \\ &+\frac{1}{2}\sum_{s=0}^{t-1}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right]\end{aligned} \tag{B.73}$$

$$\begin{aligned}\leq &\frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t-1})^T(I-\tilde{C}_{T-t})(x_{T-t-1}-v_{T-t-1})\right] \\ &+\mathbb{E}\left[(x_{T-t-1}-v_{T-t-1})^T(I-\tilde{C}_{T-t})\mathbb{E}\left[v_{T-t-1}-v_{T-t}|\mathcal{F}_{T-t-1}\right]\right] \\ &+\frac{1}{2}\sum_{s=0}^{t}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right]\end{aligned} \tag{B.74}$$

$$\leq \frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t-1})^T(I-\tilde{C}_{T-t})(x_{T-t-1}-v_{T-t-1})\right] + \frac{1}{2}\sum_{s=0}^{t}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right] \tag{B.75}$$

Therefore, the cost from $T-t-1$ to T is

$$\mathbb{E}[Cost_{\text{LAI}(\gamma)}[T-t-1,T]]$$

$$\leq \frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t-1})^T A(x_{T-t-1}-v_{T-t-1})\right] + \frac{1}{2}\mathbb{E}\|x_{T-t-1}-x_{T-t-2}\|_2^2 + \mathbb{E}[\tilde{C}_{\text{LAI}(\gamma)}[T-t,T]] \tag{B.76}$$

$$\leq \frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-v_{T-t-1})^T(A+I-\tilde{C}_{T-t})(x_{T-t-1}-v_{T-t-1})\right] + \frac{1}{2}\mathbb{E}\left[(x_{T-t-1}-x_{T-t-2})^T(x_{T-t-1}-x_{T-t-2})\right] + \frac{1}{2}\sum_{s=0}^{t}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right] \tag{B.77}$$

$$\leq \frac{1}{2}\mathbb{E}[(x_{T-t-2}-v_{T-t-1})^T\tilde{C}_{T-t-1}(\tilde{C}_{T-t-1}^{-1}-I)\tilde{C}_{T-t-1}(x_{T-t-2}-v_{T-t-1})] + \frac{1}{2}\mathbb{E}\left[(v_{T-t-1}-x_{T-t-2})^T(I-\tilde{C}_{T-t-1})^2(v_{T-t-1}-x_{T-t-2})\right] + \frac{1}{2}\sum_{s=0}^{t}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right] \tag{B.78}$$

$$\leq \frac{1}{2}\mathbb{E}[(x_{T-t-2}-v_{T-t-1})^T\left(\tilde{C}_{T-t-1}(I-\tilde{C}_{T-t-1})+(I-\tilde{C}_{T-t-1})^2\right)(x_{T-t-2}-v_{T-t-1})] + \frac{1}{2}\sum_{s=0}^{t}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right] \tag{B.79}$$

$$\leq \frac{1}{2}\mathbb{E}[(x_{T-t-2}-v_{T-t-1})^T(\tilde{C}_{T-t-1}+I-\tilde{C}_{T-t-1})(I-\tilde{C}_{T-t-1})(x_{T-t-2}-v_{T-t-1})] + \frac{1}{2}\sum_{s=0}^{t}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right] \tag{B.80}$$

$$\leq \frac{1}{2}\mathbb{E}[(x_{T-t-2}-v_{T-t-1})^T(I-\tilde{C}_{T-t-1})(x_{T-t-2}-v_{T-t-1})] + \frac{1}{2}\sum_{s=0}^{t}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right] \tag{B.81}$$

Now that our claim is proven, we can get the total cost of the algorithm by putting $t=T-1$

$$\mathbb{E}[\text{Cost}_{\text{LAI}(\gamma)}[1,T]] \leq \frac{1}{2}\mathbb{E}\left[(x_0-v_1)^T(I-\tilde{C}_1)(x_0-v_1)\right] + \frac{1}{2}\sum_{s=0}^{T-2}\mathbb{E}\left[(v_{T-s}-v_{T-s-1})^T(I-\tilde{C}_{T-s})(v_{T-s}-v_{T-s-1})\right] \tag{B.82}$$

$$= \frac{1}{2}\mathbb{E}\left[(x_0-v_1)^T(I-\tilde{C}_1)(x_0-v_1)\right] + \frac{1}{2}\sum_{t=2}^{T}\mathbb{E}\left[(v_t-v_{t-1})^T(I-\tilde{C}_t)(v_t-v_{t-1})\right] \tag{B.83}$$

□

B.6 Proof of Theorem 3.10

The proof technique used here is inspired from ROBD's competitive analysis in (Goel et al., 2019). The key aspect of our proof is the potential function that depends on $g_t(\cdot)$ of the online algorithm ALG in question. To recap, the online algorithm ALG is

$$x_t = \underset{x}{\operatorname{argmin}}\, f_t(x) + D_h(x||x_{t-1}) + D_{g_t}(x||v_t) \tag{B.84}$$

where f_t is m-strongly convex and $h(\cdot)$ is α-strongly convex and β-smooth. Recall that $g_t(\cdot)$ is $\alpha'-$strongly convex and $\beta'-$smooth. This means

$$\nabla f_t(x_t) + (\nabla h(x_t) - \nabla h(x_{t-1})) + (\nabla g_t(x_t) - \nabla g_t(v_t)) = 0. \tag{B.85}$$

We denote the hindsight optimal action sequence as $\{x_t^*\}_t$. Using strong convexity of f_t

$$\begin{aligned} f_t(x_t^*) \geq & f_t(x_t) + \langle \nabla f_t(x_t), x_t^* - x_t \rangle + \frac{m}{2} \|x_t^* - x_t\|_2^2 && \text{(B.86)} \\ = & f_t(x_t) - \langle \nabla h(x_{t-1}) - \nabla h(x_t), x_t - x_t^* \rangle && \text{(B.87)} \\ & - \langle \nabla g_t(v_t) - \nabla g_t(x_t), x_t - x_t^* \rangle + \frac{m}{2} \|x_t^* - x_t\|_2^2 . && \text{(B.88)} \end{aligned}$$

Lemma B.10 (Generalized Triangle Inequality)**.** *For any three points $x, y, x \in \mathbb{R}^n$, the following holds for Bregman divergence $D_h(\cdot||\cdot)$ defined using a strongly convex function $h(\cdot)$*

$$\langle \nabla h(y) - \nabla h(z), z - x \rangle = D_h(x||y) - D_h(x||z) - D_h(z||y)$$

Using the above lemma for $D_h\left(\cdot||\cdot\right)$ and $D_{g_t}\left(\cdot||\cdot\right)$, we have

$$\langle \nabla h(x_{t-1}) - \nabla h(x_t), x_t - x_t^* \rangle = D_h(x_t^*||x_{t-1}) - D_h(x_t^*||x_t) - D_h(x_t||x_{t-1}) \tag{B.89}$$

and

$$\langle \nabla g_t(v_t) - \nabla g_t(x_t), x_t - x_t^* \rangle = D_{g_t}(x_t^*||v_t) - D_{g_t}(x_t^*||x_t) - D_{g_t}(x_t||v_t). \tag{B.90}$$

Substituting the two above identities into inequality (B.88), we get

$$\begin{aligned} & f_t(x_t) + D_h(x_t||x_{t-1}) + (D_h(x_t^*||x_t) + D_{g_t}(x_t^*||x_t)) + \frac{m}{2} \|x_t^* - x_t\|_2^2 + D_{g_t}(x_t||v_t) && \text{(B.91)} \\ & \leq f_t(x_t^*) + D_{g_t}(x_t^*||v_t) + D_h(x_t^*||x_{t-1}) && \text{(B.92)} \end{aligned}$$

It follows that

$$\begin{aligned} & f_t(x_t) + D_h(x_t||x_{t-1}) + (D_h(x_t^*||x_t) + D_{g_t}(x_t^*||x_t)) + \frac{m}{2} \|x_t^* - x_t\|_2^2 \\ & \leq f_t(x_t^*) + D_{g_t}(x_t^*||v_t) + D_h(x_t^*||x_{t-1}) \end{aligned} \tag{B.93}$$

Now, define the potential function $\phi(x_t, x_t^*) = (D_h(x_t^*||x_t) + D_{g_t}(x_t^*||x_t)) + \frac{m}{2} \|x_t^* - x_t\|_2^2$ and the potential difference as $\Delta\phi = \phi(x_t, x_t^*) - \phi(x_{t-1}, x_{t-1}^*)$. Subtracting $\phi(x_{t-1}, x_{t-1}^*)$ from both sides of the above inequality, we get

$$\begin{aligned} & f_t(x_t) + D_h(x_t||x_{t-1}) + \Delta\phi \\ & \leq \underbrace{(f_t(x_t^*) + D_{g_t}(x_t^*||v_t))}_{X_1} + \underbrace{D_h(x_t^*||x_{t-1}) - (D_h(x_{t-1}^*||x_{t-1}) + D_{g_{t-1}}(x_{t-1}^*||x_{t-1})) - \frac{m}{2} \left\|x_{t-1}^* - x_{t-1}\right\|_2^2}_{X_2} \end{aligned} \tag{B.94}$$

We now analyze X_2

$$D_h(x_t^*||x_{t-1}) - \left(D_h(x_{t-1}^*||x_{t-1}) + D_{g_{t-1}}(x_{t-1}^*||x_{t-1})\right) - \frac{m}{2} \left\|x_{t-1}^* - x_{t-1}\right\|_2^2 \tag{B.95}$$

$$= (D_h(x_t^*||x_{t-1}) - D_h(x_{t-1}^*||x_{t-1})) - D_{g_{t-1}}(x_{t-1}^*||x_{t-1}) - \frac{m}{2}\left\|x_{t-1}^* - x_{t-1}\right\|_2^2 \tag{B.96}$$

$$= D_h(x_t^*||x_{t-1}^*) + \langle \nabla h(x_{t-1}) - \nabla h(x_{t-1}^*), x_{t-1}^* - x_t^* \rangle - D_{g_{t-1}}(x_{t-1}^*||x_{t-1}) - \frac{m}{2}\left\|x_{t-1}^* - x_{t-1}\right\|_2^2 \tag{B.97}$$

$$\leq D_h(x_t^*||x_{t-1}^*) + \|\nabla h(x_{t-1}) - \nabla h(x_{t-1}^*)\|_2 \|x_{t-1}^* - x_t^*\|_2 - D_{g_{t-1}}(x_{t-1}^*||x_{t-1}) - \frac{m}{2}\left\|x_{t-1}^* - x_{t-1}\right\|_2^2 \tag{B.98}$$

$$\leq D_h(x_t^*||x_{t-1}^*) + \beta\|x_{t-1} - x_{t-1}^*\|_2 \|x_{t-1}^* - x_t^*\|_2 - D_{g_{t-1}}(x_{t-1}^*||x_{t-1}) - \frac{m}{2}\left\|x_{t-1}^* - x_{t-1}\right\|_2^2 \tag{B.99}$$

$$\begin{aligned} &\leq D_h(x_t^*||x_{t-1}^*) + \frac{1}{2}\left(\frac{\beta^2}{\alpha'_{t-1} + m}\right)\|x_{t-1}^* - x_t^*\|_2^2 + \left(\frac{\alpha' + m}{2}\right)\|x_{t-1} - x_{t-1}^*\|_2^2 \\ &\quad - D_{g_{t-1}}(x_{t-1}^*||x_{t-1}) - \frac{m}{2}\left\|x_{t-1}^* - x_{t-1}\right\|_2^2 \end{aligned} \tag{B.100}$$

$$= D_h(x_t^*||x_{t-1}^*) + \frac{1}{2}\left(\frac{\beta^2}{\alpha'_{t-1} + m}\right)\|x_{t-1}^* - x_t^*\|_2^2 + \left(\frac{\alpha'_{t-1}}{2}\|x_{t-1} - x_{t-1}^*\|_2^2 - D_{g_{t-1}}(x_{t-1}^*||x_{t-1})\right) \tag{B.101}$$

$$\leq D_h(x_t^*||x_{t-1}^*) + \frac{1}{2}\left(\frac{\beta^2}{\alpha'_{t-1} + m}\right)\|x_{t-1}^* - x_t^*\|_2^2 \leq \left(1 + \frac{\beta^2/\alpha}{\alpha'_{t-1} + m}\right) D_h(x_t^*||x_{t-1}^*) \tag{B.102}$$

Therefore,

$$f_t(x_t) + D_h(x_t||x_{t-1}) + \Delta\phi \leq f_t(x_t^*) + D_{g_t}(x_t^*||v_t) + \left(1 + \frac{\beta^2/\alpha}{\alpha'_{t-1} + m}\right) D_h(x_t^*||x_{t-1}^*) \tag{B.103}$$

$$\leq f_t(x_t^*) + \frac{\beta'_t}{2}\|x_t^* - v_t\|^2 + \left(1 + \frac{\beta^2/\alpha}{\alpha'_{t-1} + m}\right) D_h(x_t^*||x_{t-1}^*) \tag{B.104}$$

$$= \left(1 + \frac{\beta'_t}{m}\right) f_t(x_t^*) + \left(1 + \frac{\beta^2/\alpha}{\alpha'_{t-1} + m}\right) D_h(x_t^*||x_{t-1}^*) \tag{B.105}$$

$$\leq \left(1 + \frac{\max_t \beta'_t}{m}\right) f_t(x_t^*) + \left(1 + \frac{\beta^2/\alpha}{\min_t \alpha'_{t-1} + m}\right) D_h(x_t^*||x_{t-1}^*) \tag{B.106}$$

□

B.7 Proofs of Theorem 3.11 (ii) and Corollary 3.12

To bound the competitive ratio of LAI(γ), we use Theorem 3.10 by writing LAI(γ) in the form:

$$A(x_t - v_t) + (x_t - x_{t-1}) + (\tilde{C}_t^{-1} - I - A)(x_t - v_t) = 0 \tag{B.107}$$

or in the form of an optimization problem as

$$x_t = \operatorname*{argmin}_x \frac{1}{2}(x_t - v_t)^T A(x_t - v_t) + \frac{1}{2}\|x_t - x_{t-1}\|_2^2 + \frac{1}{2}(x_t - v_t)^T(\tilde{C}_t^{-1} - I - A)(x_t - v_t). \tag{B.108}$$

Therefore, $g_t(x) = \frac{1}{2}x^T(\tilde{C}_t^{-1} - I - A)x$. Now, the eigenvalues satisfy the recursion $\frac{1}{\tilde{\lambda}_i^t} = 2 + \lambda_i^A - \tilde{\lambda}_i^{t+1}$ as result of $\tilde{C}_t^{-1} = 2I + A - \tilde{C}_{t+1}$ and $\tilde{C}_t$ having same eigenvectors as A. The above property, along with $\tilde{\lambda}_i^T \geq \lambda_i^L$ (as $\gamma \in [0, 1]$), gives $\tilde{\lambda}_i^t \leq \tilde{\lambda}_i^{t+1}\ \forall\ i$ and $C_L \preccurlyeq \tilde{C}_1 \preccurlyeq \ldots \preccurlyeq \tilde{C}_T$. The proof of this fact is straightforward and is done in Corollary D.5. Consequently,

$$C_L^{-1} - (I + A) \succcurlyeq \tilde{C}_1^{-1} - (I + A) \succcurlyeq \ldots \succcurlyeq \tilde{C}_T^{-1} - (I + A) \tag{B.109}$$

or

$$\frac{1}{\lambda_i^L} - 1 - \lambda_i^A \geq \frac{1}{\tilde{\lambda}_i^1} - 1 - \lambda_i^A \geq \ldots \geq \frac{1}{\tilde{\lambda}_i^T} - 1 - \lambda_i^A. \tag{B.110}$$

This means

$$\max_t \beta'_t = \max_i \left\{\frac{1}{\lambda_i^L} - 1 - \lambda_i^A\right\} = \frac{\sqrt{(\lambda_{\max}^A)^2 + 4\lambda_{\max}^A} - \lambda_{\max}^A}{2} = \frac{\lambda_{\min}^A}{2}\left(\sqrt{\kappa(A)^2 + \frac{4\kappa(A)}{\lambda_{\min}^A}} - \kappa(A)\right) \tag{B.111}$$

$$\min_t \alpha'_{t-1} = \min_i \left\{ \frac{1}{\tilde{\lambda}_i^T} - 1 - \lambda_i^A \right\} = \min_i \left\{ \frac{\lambda_i^A}{2} \left(\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1 \right) \right\} \tag{B.112}$$

Since, $\frac{y}{2}\left(\left(1+\frac{4}{y}\right)^{\frac{\gamma}{2}} - 1\right)$ is a non-decreasing function in $y \geq 0$ and any $\gamma \in [0,1]$, we have

$$\min_t \alpha'_{t-1} = \frac{\lambda_{\min}^A}{2} \left(\left(1 + \frac{4}{\lambda_{\min}^A}\right)^{\frac{\gamma}{2}} - 1 \right) \tag{B.113}$$

Using Theorem 3.10,

$$\text{CR}_{\text{LAI}(\gamma)} = \max \left\{ 1 + \frac{1}{2}\left(\sqrt{\kappa(A)^2 + \frac{4\kappa(A)}{\lambda_{\min}^A}} - \kappa(A) \right), 1 + \frac{1}{\frac{\lambda_{\min}^A}{2}\left(\left(1 + \frac{4}{\lambda_{\min}^A}\right)^{\frac{\gamma}{2}} + 1 \right)} \right\} \tag{B.114}$$

□

Corollary B.11. *In the adversarial setting where $\lambda_{\max}^A \ll 1$, $\text{LAI}(\gamma)$ has a competitive ratio of*

$$CR_{\text{LAI}(\gamma)} \leq 1 + \max \left\{ \sqrt{\frac{\kappa(A)}{\lambda_{\min}^A}}, \frac{2^{1-\gamma}}{\left(\lambda_{\min}^A\right)^{1-\frac{\gamma}{2}}} \right\} \ \forall\, \gamma \in (0,1].$$

Proof. With $\lambda_{\max}^A \ll 1 \implies \frac{1}{\lambda_{\min}^A} \gg \kappa(A) > 1$ and $\gamma \in (0,1]$

$$\text{CR}_{\text{LAI}(\gamma)} = 1 + \max \left\{ \sqrt{\frac{\kappa(A)}{\lambda_{\min}^A}}, \frac{1}{\frac{\lambda_{\min}^A}{2} \cdot \left(\frac{4}{\lambda_{\min}^A}\right)^{\frac{\gamma}{2}}} \right\} = 1 + \max \left\{ \sqrt{\frac{\kappa(A)}{\lambda_{\min}^A}}, \frac{2^{1-\gamma}}{\left(\lambda_{\min}^A\right)^{1-\frac{\gamma}{2}}} \right\}. \tag{B.115}$$

For $\gamma = 1$, this translates to

$$\text{CR}_{\text{LAI}(1)} = 1 + \max \left\{ \sqrt{\frac{\kappa(A)}{\lambda_{\min}^A}}, \frac{1}{\left(\lambda_{\min}^A\right)^{\frac{1}{2}}} \right\} = 1 + \sqrt{\frac{\kappa(A)}{\lambda_{\min}^A}} \tag{B.116}$$

□

C General Value function derivative and Proof of Theorem 3.2

Before the proof of the theorem, we present an important property of the value function that is applicable to a much broader class of convex hitting costs and switching costs.

Theorem C.1. *Consider stochastic convex hitting costs $f_t(x)$ and switching costs $D_h(x_t||x_{t-1})$ where $h(\cdot)$ is a convex function and $D_h(x_t||x_{t-1})$ is the Bregman divergence and the optimization problem to find the online optimal policy*

$$\operatorname*{argmin}_{\{x_t\}_{t=1}^T : x_t \in \mathcal{F}_t} \mathbb{E}\left[\sum_{t=1}^T f_t(x_t) + D_h(x_t||x_{t-1}) \right].$$

The value function at round t, $V_t(x_{t-1}, v_t)$, satisfies the following property

$$\nabla_{x_{t-1}} V(x_{t-1}, v_t) = \nabla^2 h(x_{t-1})(x_{t-1} - x_t)$$

Proof. First, it is clear that the value function at round t is the optimal cost starting from round t, means it depends only in the starting state, that is, (x_{t-1}, v_t). Further it is obtained by solving the following optimization problem where $\mathbb{E}[V_{t+1}(x, v_{t+1})|\mathcal{F}_t]$ is a function of v_t because of the expectation with respect to $\mathcal{F}_t$.

$$V_{(x_{t-1}, v_t)} = \operatorname*{argmin}_x f_t(x) + D_h(x||x_{t-1}) + \mathbb{E}[V_{t+1}(x, v_{t+1})|\mathcal{F}_t] \tag{C.1}$$

$$= \operatorname*{argmin}_x f_t(x) + (h(x) - h(x_{t-1}) - \nabla h(x_{t-1})^T (x - x_{t-1})) + \mathbb{E}[V_{t+1}(x, v_{t+1})|\mathcal{F}_t] \tag{C.2}$$

Therefore, we get the optimal action x_t by solving the following differential equation

$$\nabla_x f_t(x) + (\nabla_x h(x) - \nabla h(x_{t-1})) + \nabla_x \mathbb{E}[V(x, v_{t+1})|\mathcal{F}_t] = 0. \tag{C.3}$$

Solving this gives x_t as an operator on x_{t-1} and v_t. We define the Jacobian of x_t with respect to x_{t-1} as $\left[\frac{\delta(x_t)_j}{\delta(x_{t-1})_i}\right]_{i,j} = J \in \mathbb{R}^{d\times d}$. The optimal cost-to-go at round t is

$$V(x_{t-1}, v_t) = f_t(x_t) + D_h(x_t, x_{t-1}) + \mathbb{E}[V(x_t, v_{t+1})|\mathcal{F}_t]. \tag{C.4}$$

Therefore,

$$\begin{aligned}
&\nabla_{x_{t-1}} V(x_{t-1}, v_t) \\
=&J\nabla_{x_t} f_t(x_t) + J\nabla_{x_t} h(x_t) - \nabla h(x_{t-1}) && \text{(C.5)}\\
&- (J - I)\nabla h(x_{t-1}) - \nabla^2 h(x_{t-1})(x_t - x_{t-1}) + J\nabla_{x_t}\mathbb{E}[V(x_t, v_{t+1})|\mathcal{F}_t] \\
=&J(\nabla_{x_t} f_t(x_t) + \nabla_{x_t} h(x_t) - \nabla_{x_t} h(x_{t-1}) + \nabla_{x_t}\mathbb{E}[V(x_t, v_{t+1})|\mathcal{F}_t]) - \nabla^2 h(x_{t-1})(x_t - x_{t-1}) && \text{(C.6)}\\
=&- \nabla^2 h(x_{t-1})(x_t - x_{t-1}) && \text{(C.7)}
\end{aligned}$$

□

We use this in the proof of Theorem 3.2 where $\nabla^2 h(x_{t-1}) = I$ as our switching cost is $\frac{1}{2}\|x_t - x_{t-1}\|_2^2$.

Proof of Theorem 3.2. Irrespective of the stochastic process $\{v_t\}_{t=1}^T$, in the final round we know the current minimizer v_T and therefore the optimal action is given by Lemma B.1 as $x_T = C_T x_{T-1} + (I - C_T)v_T$. Therefore, the value function derivative at round T is

$$\nabla_{x_{T-1}} V(x_{T-1}, v_T) = x_{T-1} - x_T = (I - C_T)x_{T-1} - (I - C_T)v_T \tag{C.8}$$

The optimal action at round $T-1$ is obtained by solving

$$A(x_{T-1} - v_{T-1}) + (x_{T-1} - x_{T-2}) + \mathbb{E}[\nabla_{x_{T-1}} V(x_{T-1}, v_T)|\mathcal{F}_{T-1}] = 0 \tag{C.9}$$

which is

$$(2I + A - C_T)x_{T-1} = x_{T-2} + (I + A - C_T)v_{T-1} + (I - C_T)\mathbb{E}[v_T - v_{T-1}|\mathcal{F}_{T-1}] \tag{C.10}$$

$$C_{T-1}^{-1} x_{T-1} = x_{T-2} + (C_{T-1}^{-1} - I)v_{T-1} + (I - C_T)\mathbb{E}[v_T - v_{T-1}|\mathcal{F}_{T-1}] \tag{C.11}$$

which gives

$$x_{T-1} = C_{T-1}x_{T-2} + (I - C_{T-1})v_{T-1} + C_{T-1}(I - C_T)\mathbb{E}[v_T - v_{T-1}|\mathcal{F}_{T-1}] \tag{C.12}$$

proving the theorem for $t = T-1$ and $t = T$ (sum from $T+1$ to T is zero). Suppose the result holds for some $t \in \{1, \dots, T\}$ then the optimal action at round $t-1$ is a solution of

$$A(x_{t-1} - v_{t-1}) + (x_{t-1} - x_{t-2}) + \mathbb{E}[x_{t-1} - x_t|\mathcal{F}_{t-1}] = 0 \tag{C.13}$$

or

$$\begin{aligned}
(2I + A - C_t)x_{t-1} =&x_{t-2} + Av_{t-1} + (I - C_t)\mathbb{E}[v_t|\mathcal{F}_{t-1}] \\
&+ \sum_{s=t+1}^{T}\left(\prod_{q=t}^{s-1} C_q\right)(I - C_s)\mathbb{E}[v_s - v_{s-1}|\mathcal{F}_{t-1}] && \text{(C.14)}\\
=&x_{t-2} + (I + A - C_t)v_{t-1} + (I - C_t)\mathbb{E}[v_t - v_{t-1}|\mathcal{F}_{t-1}] \\
&+ \sum_{s=t+1}^{T}\left(\prod_{q=t}^{s-1} C_q\right)(I - C_s)\mathbb{E}[v_s - v_{s-1}|\mathcal{F}_{t-1}] && \text{(C.15)}
\end{aligned}$$

$$
\begin{aligned}
&= x_{t-2} + (C_{t-1}^{-1} - I)v_{t-1} + (I - C_t)\mathbb{E}[v_t - v_{t-1}|\mathcal{F}_{t-1}] \\
&\quad + \sum_{s=t+1}^{T} \left(\prod_{q=t}^{s-1} C_q\right)(I - C_s)\mathbb{E}[v_s - v_{s-1}|\mathcal{F}_{t-1}]
\end{aligned} \tag{C.16}
$$

which gives

$$
\begin{aligned}
x_{t-1} =& C_{t-1}x_{t-2} + (I - C_{t-1})v_{t-1} + C_{t-1}(I - C_t)\mathbb{E}[v_t - v_{t-1}|\mathcal{F}_{t-1}] \\
&+ \sum_{s=t+1}^{T} C_{t-1}\left(\prod_{q=t}^{s-1} C_q\right)(I - C_s)\mathbb{E}[v_s - v_{s-1}|\mathcal{F}_{t-1}]
\end{aligned} \tag{C.17}
$$

$$
= C_{t-1}x_{t-2} + (I - C_{t-1})v_{t-1} + \sum_{s=t}^{T}\left(\prod_{q=t-1}^{s-1} C_q\right)(I - C_s)\mathbb{E}[v_s - v_{s-1}|\mathcal{F}_{t-1}] \tag{C.18}
$$

proving Theorem 3.2 by induction. □

D Additional proofs in the stochastic environment

D.1 Proof of Theorem 3.4

We want the hindsight static optimal for $f_t(x) = \frac{\lambda}{2}\|x - v_t\|_2^2$ and switching cost $\frac{1}{2}\|x_t - x_{t-1}\|_2^2$ with the minimizers evolve as martingale. We write the minimizer at t as a sum of the initial point x_0 and martingale differences $u_k = v_k - v_{k-1}$ with $v_0 = x_0$, described below

$$
v_t = x_0 + \sum_{k=1}^{t} u_k. \tag{D.1}
$$

The martingale differences u_k are assumed to have a variance of at least $\sigma^2 > 0$. The hindsight optimal minimizing the total cost over the horizon is

$$
x^* = \operatorname*{argmin}_x \frac{1}{2}\|x - x_0\|_2^2 + \sum_{t=1}^{T}\frac{\lambda}{2}\|x - v_t\|_2^2. \tag{D.2}
$$

The optimal solution will satisfy the following expression, obtained by differentiating the objective function,

$$
(1 + \lambda T)x^* = x_0 + \lambda\sum_{t=1}^{T} v_t \tag{D.3}
$$

$$
\tag{D.4}
$$

which gives

$$
\begin{aligned}
x^* &= \left(\frac{x_0}{1+\lambda T}\right) + \left(\frac{\lambda T}{1+\lambda T}\right)\left(\frac{\sum_{t=1}^{T} v_t}{T}\right) \\
&= \left(\frac{x_0}{1+\lambda T}\right) + \left(\frac{\lambda T}{1+\lambda T}\right)\left(x_0 + \frac{\sum_{t=1}^{T}\sum_{k=1}^{t} u_k}{T}\right) \\
&= x_0 + \left(\frac{\lambda T}{1+\lambda T}\right)\left(\frac{\sum_{t=1}^{T}\sum_{k=1}^{t} u_k}{T}\right) \\
&= x_0 + \left(\frac{\lambda T}{1+\lambda T}\right)\left(\frac{\sum_{k=1}^{T} u_k \sum_{t=k}^{T} 1}{T}\right) \\
&= x_0 + \left(\frac{\lambda T}{1+\lambda T}\right)\left(\frac{\sum_{k=1}^{T}(T-k+1)u_k}{T}\right)
\end{aligned}
$$

$$= x_0 + \sum_{k=1}^{T} \frac{\lambda(T-k+1)u_k}{1+\lambda T}$$

This implies

$$x^* - x_0 = \sum_{k=1}^{T} \frac{\lambda(T-k+1)u_k}{1+\lambda T} \tag{D.5}$$

$$x^* - v_t = \sum_{k=1}^{T} \frac{\lambda(T-k+1)u_k}{1+\lambda T} - \sum_{k=1}^{t} u_k. \tag{D.6}$$

We can simplify $x^* - v_t$ to a more usable expression,

$$x^* - v_t = \sum_{k=1}^{T} \frac{\lambda(T-k+1)u_k}{1+\lambda T} - \sum_{k=1}^{t} u_k \tag{D.7}$$

$$= \sum_{k=t+1}^{T} \frac{\lambda(T-k+1)u_k}{1+\lambda T} + \sum_{k=1}^{t} \left(\frac{\lambda(T-k+1)}{1+\lambda T} - 1 \right) u_k \tag{D.8}$$

$$= \sum_{k=t+1}^{T} \left(\frac{\lambda(T-k+1)}{1+\lambda T} \right) u_k - \sum_{k=1}^{t} \left(\frac{\lambda(k-1)+1}{1+\lambda T} \right) u_k \tag{D.9}$$

$$\tag{D.10}$$

Therefore, the total expected cost of using x^* is

$$\mathbb{E}[\text{Cost}_{x^*}[1,T]] - \frac{1}{2}\mathbb{E}\left\|x^* - x_0\right\|_2^2 \tag{D.11}$$

$$= \sum_{t=1}^{T} \frac{\lambda}{2}\mathbb{E}\left\|x^* - v_t\right\|_2^2 \tag{D.12}$$

$$= \sum_{t=1}^{T} \frac{\lambda}{2}\mathbb{E}\left\| \sum_{k=t+1}^{T} \left(\frac{\lambda(T-k+1)}{1+\lambda T} \right) u_k - \sum_{k=1}^{t} \left(\frac{\lambda(k-1)+1}{1+\lambda T} \right) u_k \right\|_2^2 \tag{D.13}$$

$$= \sum_{t=1}^{T} \frac{\lambda}{2(1+\lambda T)^2}\mathbb{E}\left\| \sum_{k=t+1}^{T} \lambda(T-k+1)u_k - \sum_{k=1}^{t} (\lambda(k-1)+1)u_k \right\|_2^2 \tag{D.14}$$

$$\tag{D.15}$$

Now consider $i > j$ for the following,

$$\mathbb{E}[u_j^T u_i] = \mathbb{E}[\mathbb{E}[u_j^T u_i | \mathcal{F}_{i-1}]] \tag{D.16}$$

$$= \mathbb{E}[u_j^T \mathbb{E}[u_i | \mathcal{F}_{i-1}]] \tag{D.17}$$

$$= \mathbb{E}[u_j^T 0] = 0. \tag{D.18}$$

This gives us

$$\mathbb{E}[\text{Cost}_{x^*}[1,T]] - \frac{1}{2}\mathbb{E}\left\|x^* - x_0\right\|_2^2 \tag{D.19}$$

$$= \frac{\lambda}{2(1+\lambda T)^2} \sum_{t=1}^{T} \left(\sum_{k=t+1}^{T} \lambda^2 (T-k+1)^2 \sigma^2 + \sum_{k=1}^{t} (\lambda(k-1)+1)^2 \sigma^2 \right) \tag{D.20}$$

$$= \frac{\lambda}{2(1+\lambda T)^2} \left(\sum_{t=1}^{T} \sum_{k=t+1}^{T} \lambda^2 (T-k+1)^2 \sigma^2 + \sum_{t=1}^{T} \sum_{k=1}^{t} (\lambda(k-1)+1)^2 \sigma^2 \right) \tag{D.21}$$

$$= \frac{\lambda}{2(1+\lambda T)^2}\left(\sum_{k=2}^{T}\lambda^2(T-k+1)^2\sigma^2\sum_{t=1}^{k-1}1+\sum_{k=1}^{T}(\lambda(k-1)+1)^2\sigma^2\sum_{t=k}^{T}1\right) \tag{D.22}$$

$$= \frac{\lambda}{2(1+\lambda T)^2}\left(\sum_{k=2}^{T}(k-1)\lambda^2(T-k+1)^2\sigma^2+\sum_{k=1}^{T}(T-k+1)(\lambda(k-1)+1)^2\sigma^2\right) \tag{D.23}$$

$$= \frac{\lambda\sigma^2}{2(1+\lambda T)^2}\sum_{k=2}^{T}(T-k+1)\left[\lambda^2(k-1)(T-k+1)+(\lambda(k-1)+1)^2\right]+\frac{T\lambda\sigma^2}{2(1+\lambda T)^2} \tag{D.24}$$

$$= \frac{\lambda\sigma^2}{2(1+\lambda T)^2}\sum_{k=2}^{T}(T-k+1)\left[\lambda^2(k-1)T-\lambda^2(k-1)^2+(\lambda(k-1)+1)^2\right]+\frac{T\lambda\sigma^2}{2(1+\lambda T)^2} \tag{D.25}$$

$$= \frac{\lambda\sigma^2}{2(1+\lambda T)^2}\sum_{k=2}^{T}(T-k+1)\left[\lambda^2(k-1)T+2\lambda(k-1)+1\right]+\frac{T\lambda\sigma^2}{2(1+\lambda T)^2} \tag{D.26}$$

$$= \frac{\lambda\sigma^2}{2(1+\lambda T)^2}\sum_{k=2}^{T}(T-k+1)\left[\lambda(k-1)(\lambda T+2)+1\right]+\frac{T\lambda\sigma^2}{2(1+\lambda T)^2} \tag{D.27}$$

$$= \frac{\sigma^2\lambda^2(\lambda T+2)}{2(1+\lambda T)^2}\sum_{k=2}^{T}(T-k+1)(k-1)+\frac{T\lambda\sigma^2}{2(1+\lambda T)^2}+\frac{\lambda\sigma^2}{2(1+\lambda T)^2}\sum_{k=2}^{T}(T-k+1) \tag{D.28}$$

$$= \frac{\sigma^2\lambda^2(\lambda T+2)}{2(1+\lambda T)^2}\left(T\sum_{k=2}^{T}(k-1)-\sum_{k=2}^{T}(k-1)^2\right)+\frac{\lambda\sigma^2}{2(1+\lambda T)^2}\left(\frac{T(T+1)}{2}\right) \tag{D.29}$$

$$= \frac{\sigma^2\lambda^2(\lambda T+2)}{2(1+\lambda T)^2}\left(\frac{T^2(T-1)}{2}-\frac{(T-1)T(2T-1)}{6}\right)+\frac{\lambda\sigma^2}{2(1+\lambda T)^2}\left(\frac{T(T+1)}{2}\right) \tag{D.30}$$

$$= \frac{\sigma^2\lambda^2(\lambda T+2)}{2(1+\lambda T)^2}\left(\frac{(T-1)T(T+1)}{6}\right)+\frac{\lambda\sigma^2}{2(1+\lambda T)^2}\left(\frac{T(T+1)}{2}\right) \tag{D.31}$$

$$= \frac{\sigma^2\lambda^2}{12}\left(\frac{(\lambda T+2)(T-1)T(T+1)}{(1+\lambda T)^2}\right)+\frac{\lambda\sigma^2}{2(1+\lambda T)^2}\left(\frac{T(T+1)}{2}\right) \tag{D.32}$$

$$= \Theta(T^2) \tag{D.33}$$

□

D.2 Properties of $\{C_t\}_t$

Lemma D.1. *For any $t \in \{1,\dots,T\}$, the matrix C_t satisfies the following properties:*

(a) *C_t is invertible*

(b) *C_t has the same set of eigenvectors as A.*

(c) *Eigenvalues of C_t lie in $(0,1)$, and hence, C_t is positive definite.*

Proof. We prove this result using induction. Because A is a positive definite matrix, its eigenvalues are greater than 1. Consequently, $(I+A)$ is invertible and eigenvalues of $C_T = (I+A)^{-1}$ are in $(0,1)$. Further, $(I+A)^{-1}$ has the same set of eigenvectors as A. Therefore, the lemma holds for $t = T$. Now assume the claim for $t \le T$. From the recursion relation, we have

$$C_{t-1}^{-1} = 2I + A - C_t \tag{D.34}$$

Now, $2I + A - C_t$ will have the same eigenvectors as A. Further, eigenvalues of $2I + A - C_t$ are $\{2+\lambda_i^A - \lambda_i^t\}_t$ and greater than 1 as $\{\lambda_i^t\}_t$, the eigenvalues of C_t, are in $(0,1)$. Therefore, $2I + A - C_t$ is invertible and so is C_{t-1}. Consequently, C_{t-1} has same eigenvectors as that of A and the eigenvalues of C_{t-1} are in $(0,1)$. □

D.3 Asymptotic behavior of $\{C_t\}_{t=1}^T$

Lemma D.2. *For any horizon $T > 0$, consider the sequence of numbers $\{c_t\}_{t=1}^T$ satisfying*

$$\frac{1}{c_t} = 2 + \lambda - c_{t+1}$$

with $c_T = \frac{1}{\lambda+1}$ for some $\lambda > 0$. Such a sequence of numbers has the following properties

1. *For $A = \lambda I$, the* LAI *algorithm will have*
$$C_t = c_t I \; \forall \, t \in \{1, \ldots, T\}$$
2. *The evolution of c_t is*
$$(c_t - c_L) = (c_{t+1} - c_L) c_t c_L$$
3. *$\{c_t\}_{t=1}^T$ is an increasing sequence that is*
$$c_L < c_1 < \ldots < c_T$$
4. *$\{c_t\}$ satisfy*
$$(c_t - c_L) \leq (1 - c_L) \left(c_T\right)^{2(T-t+1)}$$
5. *The behavior of c_1 as horizon goes to infinity is*
$$\lim_{T \to \infty} c_1 = c_L$$

where $c_L = \frac{\lambda+2-\sqrt{\lambda^2+4\lambda}}{2} = \frac{2}{\lambda+2+\sqrt{\lambda^2+4\lambda}} \in (0, 1)$ for $\lambda > 0$.

Proof. For the part 1 of the lemma, we prove it by induction. First observe that

$$C_T = (A + I)^{-1} \tag{D.35}$$
$$= \left(\frac{1}{\lambda + 1}\right) I \tag{D.36}$$

for $\lambda > 0$. This means part (i) holds for $t = T$. Assuming the claim for some $t \leq T$, that is, $C_t = c_t I$

$$C_{t-1} = (2I + A - C_t)^{-1} \tag{D.37}$$
$$= ((1 + \lambda - c_t) I)^{-1} \tag{D.38}$$
$$= \left(\frac{1}{2 + \lambda - c_t}\right) I \tag{D.39}$$
$$= c_{t-1} I \tag{D.40}$$

Now, observe that

$$c_L + \frac{1}{c_L} = \frac{\lambda + 2 - \sqrt{\lambda^2 + 4\lambda}}{2} + \frac{\lambda + 2 + \sqrt{\lambda^2 + 4\lambda}}{2} \tag{D.41}$$
$$= \lambda + 2 \tag{D.42}$$

and to recap $\{c_t\}_{t=1}^T$ satisfy the recursion of Lemma D.2, that is

$$\frac{1}{c_t} = 2 + \lambda - c_{t+1} \tag{D.43}$$

and $c_T = \frac{1}{\lambda+1}$. Therefore,

$$\frac{1}{c_L} - \frac{1}{c_t} = c_{t+1} - c_L \tag{D.44}$$

which gives

$$(c_t - c_L) = (c_{t+1} - c_L)c_t c_L \tag{D.45}$$

proving part 2 of the lemma. Now based on the recursion (D.43),

$$c_{t+1} \in (0,1) \implies c_t \in (0,1) \tag{D.46}$$

and $c_T = \frac{1}{\lambda+1} \in (0,1)$, which inductively proves that $c_t \in (0,1)$ for all t. Now, observe that $c_L \in (0,1)$ for $\lambda > 0$ and hence,

$$\begin{aligned} (c_t - c_L) &< (c_{t+1} - c_L) \\ c_t &< c_{t+1} \end{aligned} \tag{D.47}$$

Also notice that because of the recursion (D.44),

$$c_{t+1} > c_L \implies c_t > c_L$$

and we know

$$c_L = \frac{2}{(\lambda+1) + (1+\sqrt{\lambda^2+4\lambda})} < \frac{1}{\lambda+1} = c_T \; \forall \; \lambda > 0 \tag{D.48}$$

which inductively proves that $c_t > c_L$ for all t completing the proof of part 3 of the lemma. Observe that

$$(c_t - c_L) = (c_{t+1} - c_L)c_t c_L \tag{D.49}$$

$$\leq (c_{t+1} - c_L)c_T^2 \tag{D.50}$$

$$\leq (c_T - c_L)c_T^{2(T-t)} \tag{D.51}$$

Also, note that

$$\frac{1}{c_T} = 2 + \lambda - 1 \implies (c_T - c_L) = (1 - c_L)c_T c_L \tag{D.52}$$

and therefore,

$$(c_t - c_L) \leq (1 - c_L)c_T^{2(T-t+1)} \tag{D.53}$$

proving part 4 of the lemma. Finally because $c_T < 1$ for $\lambda > 0$,

$$\lim_{T\to\infty} c_1 \leq \lim_{T\to\infty} (c_T - c_L)c_T^{2T} = 0 \tag{D.54}$$

proving part 5 of the lemma. □

Definition D.3. The square root of a real positive definite matrix M is the unique positive definite matrix $M^{1/2}$ such that

$$M^{1/2}\left(M^{1/2}\right)^T = \left(M^{1/2}\right)^T M^{1/2} = M.$$

Since, M can be written as PD_MP^T where D_M is the diagonal matrix of eigenvalues of M and P is the matrix having eigenvectors of M as columns (also called the *modal matrix* of A),

$$M^{1/2} = PD_M^{1/2}P^T,$$

and $D_M^{1/2}$ is the diagonal matrix of square root of eigenvalues of M.

Definition D.4. Define C_L as

$$C_L := \frac{A + 2I - \sqrt{A^2 + 4A}}{2}.$$

Then C_L will have the same set of eigenvectors as A, with eigenvalues $\left\{\lambda_i^L\right\}_{i=1}^d$ where

$$\lambda_i^L = \frac{\lambda_i^A + 2 - \sqrt{\left(\lambda_i^A\right)^2 + 4\lambda_i^A}}{2} = \frac{2}{\lambda_i^A + 2 + \sqrt{\left(\lambda_i^A\right)^2 + 4\lambda_i^A}}.$$

Note that λ_i^L is a decreasing function of λ_i^A and therefore,

$$\lambda_{\min}^L = \frac{\lambda_{\max}^A + 2 - \sqrt{\left(\lambda_{\max}^A\right)^2 + 4\lambda_{\max}^A}}{2} \tag{D.55}$$

$$\lambda_{\max}^L = \frac{\lambda_{\min}^A + 2 - \sqrt{\left(\lambda_{\min}^A\right)^2 + 4\lambda_{\min}^A}}{2} \tag{D.56}$$

Corollary D.5. *For any horizon $T > 0$, consider a matrix sequence $\{M_t\}_{t=1}^T$ satisfying*

$$M_t^{-1} = 2I + A - M_{t+1} \; \forall \, t \in \{1, \ldots, T-1\} \tag{D.57}$$

with $C_L \prec M_T \prec I$ and having eigenvectors same as A. Denoting the eigenvalues of $\{M_t\}_{t=1}^T$ corresponding to the ith eigenvector as $\{\rho_i^t\}_{t=1}^T$, they satisfy the following properties:

1. *The ith eigenvalue of the matrices $\{M_t\}_{t=1}^T$ satisfy*

$$\frac{1}{\rho_i^t} = 2 + \lambda_i^A - \rho_i^{t+1} \; \forall \, t \in \{1, \ldots, T-1\}$$

2. *The ith eigenvalue of the matrices $\{M_t\}_{t=1}^T$ satisfy*

$$\left(\rho_i^t - \lambda_i^L\right) = \left(\rho_i^{t+1} - \lambda_i^L\right) \rho_i^t \lambda_i^L$$

3. *The following order holds for any horizon T*

$$C_L \prec M_1 \prec \ldots \prec M_T$$

4. *The ith eigenvalues of $\{M_t\}_{t=1}^T$ show the following behavior*

$$\left(\rho_i^t - \lambda_i^L\right) \leq \left(\rho_i^T - \lambda_i^L\right) \left(\rho_i^T\right)^{2(T-t)}$$

5. *The asymptotic behavior of $\{M_t\}_{t=1}^T$ is*

$$\lim_{T \to \infty} z^T (M_1 - C_L) z = 0 \; \forall \, z \in \mathbb{R}^d.$$

Proof. Consider the following matrix

$$C_L = \frac{A + 2I - \sqrt{A^2 + 4A}}{2} \tag{D.58}$$

where $\sqrt{A^2 + 4A}$ is the unique positive definite square root of the positive definite matrix $(A^2 + 4A)$. We know that the square root has to be $PD^{1/2}P^T$ where D is the diagonal matrix of eigenvalues of $(A^2 + 4A)$ and P is the matrix having the

corresponding eigenvectors of (A^2+4A) as columns. Therefore, $\sqrt{A^2+4A}$ has the same set of eigenvectors as (A^2+4A). This means that C_L has the same set of eigenvectors as A. This means that

$$\lambda_i^L = \frac{\lambda_i^A + 2 - \sqrt{\left(\lambda_i^A\right)^2 + 4\lambda_i^A}}{2} = \frac{2}{\lambda_i^A + 2 + \sqrt{\left(\lambda_i^A\right)^2 + 4\lambda_i^A}} < \rho_i^T \tag{D.59}$$

where λ_i^L is the ith eigenvalue of C_L and λ_i^A is the ith eigenvalue of A. Now, consider the recursion followed by $\{C_t\}_{t=1}^T$,

$$M_t^{-1} = 2I + A - M_{t+1} \tag{D.60}$$

and we know that each one of $\{M_t\}_{t=1}^T$ has same the set of eigenvectors as A. This gives us the following recursion followed by the ith eigenvalue of $\{C_t\}_{t=1}^T$

$$\frac{1}{\rho_i^t} = 2 + \lambda_i^A - \rho_i^{t+1} \tag{D.61}$$

Now, from the value of λ_i^L, we know

$$\lambda_i^L + \frac{1}{\lambda_i^L} = 2 + \lambda_i^A. \tag{D.62}$$

which when substituted in (D.61) gives

$$\left(\rho_i^t - \lambda_i^L\right) = \left(\rho_i^{t+1} - \lambda_i^L\right)\rho_i^t\lambda_i^L \tag{D.63}$$

Now observe that $\{\rho_i^t\}_{t=1}^T$ follows recursion (D.43) with $\lambda = \lambda_i^A$. Further, λ_i^L is c_L with $\lambda = \lambda_i^A$. Therefore, applying the properties in Lemma D.2, we get that

$$\lambda_i^L < \rho_i^1 < \ldots < \rho_i^T \;\forall\, i \in \{1, \ldots, d\} \tag{D.64}$$

and since, C_L and each of $\{M_t\}_t$ have the same set of eigenvectors,

$$C_L \prec M_1 \prec M_2 \prec \ldots \prec M_T. \tag{D.65}$$

Further,

$$(\rho_i^t - \lambda_i^L) \leq (\rho_i^{t+1} - \lambda_i^L)\cdot(\rho_i^T)^2 \tag{D.66}$$

$$\leq (\rho_i^T - \lambda_i^L)\cdot(\rho_i^T)^{2(T-t)} \tag{D.67}$$

Since $\rho_i^T < 1$ for every i, we have

$$\lim_{T\to\infty}\left(\rho_i^1 - \lambda_i^L\right) \leq \lim_{T\to\infty}(\rho_i^T - \lambda_i^L)\cdot(\rho_i^T)^{2(T-1)} = 0 \;\forall\, i \in \{1, \ldots, d\}. \tag{D.68}$$

Now, consider M_1 and C_L. We know $M_1 = PD_1P^T$ and PD_LP^T where $D_1 = diag\left(\left(\rho_i^1\right)_{i=1}^d\right)$, $D_L = diag\left(\left(\lambda_i^L\right)_{i=1}^d\right)$ and P is the modal matrix of A.

$$z^T(M_1 - C_L)z = z^TP(M_1 - D_L)P^Tz \tag{D.69}$$

$$= \begin{bmatrix} k_1 & \ldots & k_d \end{bmatrix} \begin{bmatrix} \rho_1^1 - \rho_1^L & 0 & \ldots & 0 \\ 0 & \rho_2^1 - \rho_2^L & \ldots & 0 \\ \vdots & \ddots & \ldots & \vdots \\ 0 & \ldots & \ldots & \rho_d^1 - \lambda_d^L \end{bmatrix} \begin{bmatrix} k_1 \\ \vdots \\ k_d \end{bmatrix} \tag{D.70}$$

$$= \sum_{i=1}^d \left(\rho_i^1 - \lambda_i^L\right) k_i^2 \tag{D.71}$$

for any $z \in \mathbb{R}^d$. Therefore, in the case of infinite horizon,

$$\lim_{T\to\infty} z^T(M_1 - C_L)z = 0 \;\forall\, z \in \mathbb{R}^d \tag{D.72}$$

□

D.4 Proof of Theorem 3.5

The ROBD solution for hitting costs $f_t(x) = \frac{1}{2}(x - v_t)^T A(x - v_t)$ and switching costs $c(x_t, x_{t-1}) = \frac{1}{2}\|x_t - x_{t-1}\|_2^2$ is

$$x_t = \operatorname*{argmin}_x \frac{1}{2}(x - v_t)^T A(x - v_t) + \frac{\mu_1}{2}\|x - x_{t-1}\|_2^2 + \frac{\mu_2}{2}\|x - v_t\|_2^2. \tag{D.73}$$

Equating the derivative of the objective to zero

$$(A + \mu_1 I + \mu_2 I)x = \mu_1 x_{t-1} + (A + \mu_2)v_t. \tag{D.74}$$

Setting $\mu_1 = 1$, we get

$$x_t = (A + I + \mu_2 I)^{-1} x_{t-1} + (I - (A + I + \mu_2 I)^{-1})v_t \tag{D.75}$$
$$= C_{\text{ROBD}} x_{t-1} + (I - C_{\text{ROBD}})v_t \tag{D.76}$$

with μ_2 as

$$\mu_2 = \frac{\lambda_{\min}^A}{2}\left(\sqrt{1 + \frac{4}{\lambda_{\min}^A}} - 1\right). \tag{D.77}$$

Now, lets have a look at eigenvalues of C_{ROBD}

$$C_{\text{ROBD}} = (A + I + \mu_2 I)^{-1} \tag{D.78}$$
$$= P(D_A + I + \mu_2 I)^{-1} P^T \tag{D.79}$$

which gives us

$$\lambda_i^{C_{\text{ROBD}}} = \frac{1}{\lambda_i^A + 1 + \frac{\sqrt{(\lambda_{\min}^A)^2 + 4\lambda_{\min}^A} - \lambda_{\min}^A}{2}} \tag{D.80}$$

Therefore, $A \neq \lambda I$ implies there exists i such that $\lambda_i^A \neq \lambda_{\min}^A$, giving us $C_{\text{ROBD}} \neq C_L$. Now,

$$\lambda_i^{C_{\text{ROBD}}} = \lambda_i^L \iff \lambda_i^A = \lambda_{\min}^A \tag{D.81}$$

and therefore, we can replace the condition $(C_{\text{ROBD}} - C_L)\Sigma \neq 0_{d\times d}$ and

$$Z = \{i \in \{1, \dots, d\} : \lambda_i^{C_{\text{ROBD}}} = \lambda_i^L\}$$

with $(A - \lambda_{\min}^A I)\Sigma \neq 0_{d\times d}$ and

$$Z = \{i \in \{1, \dots, d\} : \lambda_i^A = \lambda_{\min}^A\}$$

which proves the theorem through Theorem B.5.

□

D.5 Linear regret for Follow the Minimizer (FTM) algorithm

Lemma D.6. *For $f_t(x) = \frac{1}{2}(x - v_T)^T A(x - v_t)$ and $c(x_t, x_{t-1}) = \frac{1}{2}\|x_t - x_{t-1}\|_2^2$ with minimizers $\{v_t\}_t$ being a martingale, the* FTM *algorithm has regret with respect to* LAI *as*

$$\textit{Regret}_{\text{FTM}}[1, T] \geq \left(\frac{\lambda_{\min}^L \sigma^2}{2}\right) T$$

Proof. The cost of FTM algorithm for $f_t(x) = \frac{1}{2}(x - v_t)^T A(x - v_t)$ will be

$$\mathbb{E}[\text{Cost}_{\text{FTM}}[1, T]] = \mathbb{E}\left[0 + \frac{1}{2}\|x_0 - v_1\|_2^2 + \sum_{t=2}^{T}\left(0 + \frac{1}{2}\|v_t - v_{t-1}\|_2^2\right)\right] \tag{D.82}$$

and, therefore, the gap to LAI will be

$$\begin{aligned}
\mathbb{E}[\text{Cost}_{\text{FTM}}[1,T]] - \mathbb{E}[\text{Cost}_{\text{LAI}}[1,T]] =& \mathbb{E}\left[\frac{1}{2}(x_0 - v_1)^T(x_0 - v_1)\middle|\mathcal{F}_0\right] \\
&+ \sum_{t=2}^{T} \mathbb{E}\left[\frac{1}{2}(v_t - v_{t-1})^T(v_t - v_{t-1})\middle|\mathcal{F}_0\right] \\
&- \mathbb{E}\left[\frac{1}{2}(x_0 - v_1)^T(I - C_1)(x_0 - v_1)\middle|\mathcal{F}_0\right] \\
&- \sum_{t=2}^{T} \mathbb{E}\left[\frac{1}{2}(v_t - v_{t-1})^T(I - C_t)(v_t - v_{t-1})\middle|\mathcal{F}_0\right] && \text{(D.83)} \\
=& \mathbb{E}\left[\frac{1}{2}(x_0 - v_1)^T C_1(x_0 - v_1)\middle|\mathcal{F}_0\right] \\
&+ \sum_{t=2}^{T} \mathbb{E}\left[\frac{1}{2}(v_t - v_{t-1})^T C_t(v_t - v_{t-1})\middle|\mathcal{F}_0\right] && \text{(D.84)} \\
\geq& \mathbb{E}\left[\frac{1}{2}(x_0 - v_1)^T C_L(x_0 - v_1)\middle|\mathcal{F}_0\right] \\
&+ \sum_{t=2}^{T} \mathbb{E}\left[\frac{1}{2}(v_t - v_{t-1})^T C_L(v_t - v_{t-1})\middle|\mathcal{F}_0\right] && \text{(D.85)} \\
\geq& \mathbb{E}\left[\frac{1}{2}(x_0 - v_1)^T \lambda_{\min}^L I(x_0 - v_1)\middle|\mathcal{F}_0\right] \\
&+ \sum_{t=2}^{T} \mathbb{E}\left[\frac{1}{2}(v_t - v_{t-1})^T \lambda_{\min}^L I(v_t - v_{t-1})\middle|\mathcal{F}_0\right] && \text{(D.86)} \\
=& \left(\frac{\lambda_{\min}^L \sigma^2}{2}\right) T && \text{(D.87)}
\end{aligned}$$

□

E Additional proofs for LAI's adversarial performance

E.1 Proof of Theorem 3.8

The LAI algorithm is

$$x_t = C_t x_{t-1} + (I - C_t) v_t \tag{E.1}$$

which means

$$A(x_t - v_t) + (x_t - x_{t-1}) + (C_t^{-1} - A - I)(x_t - v_t) = 0. \tag{E.2}$$

The above equation can be written as

$$x_t = \operatorname*{argmin}_x \frac{1}{2}(x_t - v_t)^T A(x_t - v_t) + \frac{1}{2}\|x_t - x_{t-1}\|_2^2 + \frac{1}{2}(x_t - v_t)^T(C_t^{-1} - (I + A))(x_t - v_t). \tag{E.3}$$

This means $m = \lambda_{\min}^A$ and $\alpha = \beta = 1$. Remember,

$$C_L^{-1} \succ C_1^{-1} \succ \ldots \succ C_T^{-1} \tag{E.4}$$

which means

$$C_L^{-1} - (I + A) \succ C_1^{-1} - (I + A) \succ \ldots \succ C_T^{-1} - (I + A) = 0 \tag{E.5}$$

Now, α'_t will be the smallest eigenvalue of $C_t^{-1} - (I + A)$ and β'_t will be the largest eigenvalue of $C_t^{-1} - (I + A)$. Therefore, $\min_t \alpha'_t = \alpha'_T = 0$ and $\max_t \beta'_t = \beta'_1$. Now, the largest eigenvalue of $C_1^{-1} - (I + A)$ is less than that of of $C_L^{-1} - (I + A)$, that is

$$\beta'_1 < \max_i \left\{ \frac{1}{\lambda_i^L} - (1 + \lambda_i^A) \right\} \tag{E.6}$$

$$= \max_i \frac{\sqrt{\left(\lambda_i^A\right)^2 + 4\lambda_i^A} - \lambda_i^A}{2} \tag{E.7}$$

$$= \frac{\sqrt{\left(\lambda_{\max}^A\right)^2 + 4\lambda_{\max}^A} - \lambda_{\max}^A}{2} \tag{E.8}$$

This gives us the competitive ratio of dynamic programming for any horizon T as

$$\text{CR}_{\text{LAI}} = \max \left\{ 1 + \frac{\sqrt{\left(\lambda_{\max}^A\right)^2 + 4\lambda_{\max}^A} - \lambda_{\max}^A}{2\lambda_{\min}^A}, 1 + \frac{1}{\lambda_{\min}^A} \right\} \tag{E.9}$$

$$\tag{E.10}$$

Now the $(\sqrt{y^2 + 4y} - y) < 2$ for all $y \geq 0$, which means the second term is always larger than the first and

$$\text{CR}_{\text{LAI}} = 1 + \frac{1}{\lambda_{\min}^A} \tag{E.11}$$

□

F Additional proofs regarding LAI(γ)

F.1 Proof of Lemma B.9

As a consequence of $\{C_t\}_t$ and $\{\tilde{C}_t\}$ having eigenvectors same as that of A and satisfying

$$C_t^{-1} = 2I + A - C_{t+1} \tag{F.1}$$

$$\tilde{C}_t^{-1} = 2I + A - \tilde{C}_{t+1}, \tag{F.2}$$

the eigenvalues of C_t and $\tilde{C}_t$ follow the recursion

$$\frac{1}{\lambda_i^t} = 2 + \lambda_i^A - \lambda_i^{t+1} \tag{F.3}$$

$$\frac{1}{\tilde{\lambda}_i^t} = 2 + \lambda_i^A - \tilde{\lambda}_i^{t+1} \tag{F.4}$$

for $t \in \{1, \ldots, T-1\}$. This means

$$\frac{1}{\tilde{\lambda}_i^t} - \frac{1}{\lambda_i^t} = \lambda_i^{t+1} - \tilde{\lambda}_i^{t+1} \tag{F.5}$$

or

$$\left(\lambda_i^t - \tilde{\lambda}_i^t\right) = \left(\lambda_i^{t+1} - \tilde{\lambda}_i^{t+1}\right) \lambda_i^t \tilde{\lambda}_i^t \tag{F.6}$$

$$\leq \left(\lambda_i^{t+1} - \tilde{\lambda}_i^{t+1}\right) \lambda_i^T \tilde{\lambda}_i^T \tag{F.7}$$

$$\leq \left(\lambda_i^{t+1} - \tilde{\lambda}_i^{t+1}\right) \left(\lambda_i^T\right)^2 \tag{F.8}$$

$$\leq \left(\lambda_i^T - \tilde{\lambda}_i^T\right) \left(\lambda_i^T\right)^{2(T-t)} \tag{F.9}$$

Further,

$$\frac{1}{\lambda_i^T} = 2 + \lambda_i^A - 1 \tag{F.10}$$

$$\frac{1}{\tilde{\lambda}_i^T} = 2 + \lambda_i^A - \left(1 - \frac{\lambda_i^A}{2}\left(\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\right) \tag{F.11}$$

meaning

$$\lambda_i^T - \tilde{\lambda}_i^T = \frac{\lambda_i^A}{2}\left(\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\lambda_i^T\tilde{\lambda}_i^T \tag{F.12}$$

$$\leq \frac{\lambda_i^A}{2}\left(\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\left(\lambda_i^T\right)^2. \tag{F.13}$$

This gives

$$\left(\lambda_i^t - \tilde{\lambda}_i^t\right) \leq \frac{\lambda_i^A}{2}\left(\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\left(\lambda_i^T\right)^{2(T-t+1)} \tag{F.14}$$

□

Lemma F.1. *Consider $y_s = P^T(v_{T-s+1} - v_{T-s})$ for $s \in [1, T]$. For increments $(v_{T-s+1} - v_{T-s})$ having same covariance matrix Σ, we have*

$$\sum_{i=1}^{d} \mathbb{E}[(y_s)_i^2] = \sigma^2$$

where $(y_s)_i$ is the ith coordinate of y and $\sigma^2 = Tr(\Sigma)$.

Proof. We take $y_s = P^T(v_{T-s+1} - v_{T-s})$. If all increments $(v_{T-s+1} - v_{T-s})$ have same covariance matrix Σ, we will have

$$\mathbb{E}(y_s)_i^2 = \mathbb{E}[e_i^T P^T (v_{T-s+1} - v_{T-s})(v_{T-s+1} - v_{T-s})^T P e_i] \tag{F.15}$$

$$= e_i^T P^T \Sigma P e_i \tag{F.16}$$

$$= \mathbb{E}(y)_i^2. \tag{F.17}$$

Therefore,

$$\sum_{i=1}^{d} \mathbb{E}(y)_i^2 = \sum_{i=1}^{d} \mathbb{E}[(v_{T-s+1} - v_{T-s})^T P e_i e_i^T P^T (v_{T-s+1} - v_{T-s})] \tag{F.18}$$

$$= \mathbb{E}[(v_{T-s+1} - v_{T-s})^T P \left(\sum_{i=1}^{d} e_i e_i^T\right) P^T (v_{T-s+1} - v_{T-s})] \tag{F.19}$$

$$= \mathbb{E}[(v_{T-s+1} - v_{T-s})^T P I P^T (v_{T-s+1} - v_{T-s})] \tag{F.20}$$

$$= \mathbb{E}\|(v_{T-s+1} - v_{T-s})\|_2^2 \tag{F.21}$$

$$= \sigma^2 \tag{F.22}$$

□

Corollary F.2. *For the relaxed assumption that increments only have same variance σ^2 (and not necessarily the entire covariance matrix),* LAI(γ) *has the following constant regret,*

$$\textit{Regret}_{\text{LAI}(\gamma)}[1, T] \leq \frac{\sigma^2}{4}\sum_{i=1}^{d}\left(\frac{\left(1 + \frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1}{\lambda_i^A + 2}\right)$$

Proof. For the looser assumption (that the increments only have same variance and not the entire covariance matrix)

$$\mathbb{E}[Cost_{\text{LAI}(\gamma)}[1,T]] - \mathbb{E}[\text{Cost}_{\text{LAI}}[1,T]]$$

$$\leq \frac{1}{2}\sum_{t=1}^{T}\mathbb{E}\left[(v_t - v_{t-1})^T P(D_t - \tilde{D}_t)P^T(v_t - v_{t-1})\right] \tag{F.23}$$

$$\leq \frac{1}{2}\sum_{t=1}^{T}\mathbb{E}\left[(v_t - v_{t-1})^T P\left(\sum_{i=1}^{d}\lambda_i^t - \tilde{\lambda}_i^t\right) IP^T(v_t - v_{t-1})\right] \tag{F.24}$$

$$= \sum_{t=1}^{T}\left(\sum_{i=1}^{d}\lambda_i^t - \tilde{\lambda}_i^t\right)\frac{\sigma^2}{2} \tag{F.25}$$

$$= \frac{\sigma^2}{2}\sum_{i=1}^{d}\frac{\lambda_i^A}{2}\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\sum_{t=1}^{T}\left(\lambda_i^T\right)^{2(T-t+1)} \tag{F.26}$$

$$\leq \frac{\sigma^2}{2}\sum_{i=1}^{d}\frac{\lambda_i^A}{2}\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)\sum_{t=1}^{\infty}\left(\lambda_i^T\right)^{2t} \tag{F.27}$$

$$= \frac{\sigma^2}{4}\sum_{i=1}^{d}\left(\frac{\lambda_i^A\left(\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1\right)}{\left(\lambda_i^A+1\right)^2 - 1}\right) \tag{F.28}$$

$$= \frac{\sigma^2}{4}\sum_{i=1}^{d}\left(\frac{\left(1+\frac{4}{\lambda_i^A}\right)^{\frac{\gamma}{2}} - 1}{\lambda_i^A+2}\right) \tag{F.29}$$

□

G Details of Numerical Experiments

G.1 Purely-stochastic environment

In this subsection, we explain how the minimizer sequence $\{v_t\}_t$ is developed for the experiments. First, the minimizer is written as a sum of random variables $\{u_t\}_t$,

$$v_t = \sum_{s=1}^{t} u_s \tag{G.1}$$

which can be thought of as increments to the minimizers. The different distributions discussed are with respect to the the increments $\{u_t\}_t$. We build $\{u_t\}_t$ such that they are not necessarily independent of each other. The process for any subset of the horizon $\tau \subseteq \{1,\ldots,T\}$ involves the following steps:

1. Choosing a common distribution $F_\tau(\cdot)$ and sampling z_t

$$z_t \sim F_\tau \;\forall\, t \in \tau \tag{G.2}$$

2. Choosing a positive definite covariance matrix $\Sigma \in \mathbb{R}^{10|\tau|\times 10|\tau|}$ and performing its Choleksy decomposition to get L, such that $LL^T = \Sigma$

3. Concatenating $\{z_t\}_{t\in\tau}$ into a single vector $z \in R^{10|\tau|}$ and transforming it into

$$u = Lz \tag{G.3}$$

and splitting it back into $\{u_t\}_{t\in\tau}$

For the light tail case, we want to exhibit the robustness to distribution shift. We, therefore, split the horizon into five subsets $\{\tau_i\}_{i=0}^4$ where $\tau_i = \{(iT/5)+1, \ldots, (i+1)T/5\}$. For each subset we create the increments u_t according to the above mentioned process, each subset pertaining to a different form of distribution.

To show that our results hold for heavy tail distributions, we choose the pareto distribution (Type 2, that is, Lomax) and the log-normal distributions as our candidates. Since these distributions are one sided, we build the z_t vector, in the process above, by first sampling $d = 10$ elements from the one dimensional distribution and then multiply a $p = 0.5$ Bernoulli random variable to each element, making z_t a zero-mean heavy tail random vector.

G.2 Mixed Environment

Our next set of experiments deal with an environment which is partly stochastic and partly adversarial. We do this by putting adversarial minimizers in an otherwise martingale sequence of minimizers. The level of adversarial infiltration ranges from zero to hundred, with zero being martingale minimizers and hundred being completely adversarial minimizers. We perform this in the following way:

1. First, we generate an IID sequence of increments $\{u_t \sim F : t \in \{1, \ldots, T\}\}$ and define a martingale minimizers as $v_t = \sum_{s=1}^t u_s$. Here, the underlying distribution F can be light or heavy tailed.
2. Next, we decide the percentage (p) of rounds we want to be adversarial and choose that many rounds from $\{1, \ldots, T\}$ uniformly at random. The set of adversarial rounds τ is fixed for the $N = 1000$ sample runs.
3. During the runs, we replace $\{v_t\}_{t\in\tau}$ with $\{\tilde{v}_t\}_{t\in\tau}$ which are generated adversarially. Note that $\{\tilde{v}_t\}_{t\in\tau}$ is fixed for all sample runs of a given adversarial percentage (p).

For light tail, we choose the normal distribution and for heavy tails we take the two-sided pareto and log-normal distributions. The adversarial percentage (p) varies from 0 to 100 and we consider a horizon of $T = 100$. The scale of total cost for an adversarial environment can be different than that of a stochastic environment. To mitigate this issue, we plot the ratio of the total cost of an online algorithm ALG to that of LAI, that is, $\frac{\mathbb{E}[\text{Cost}_{\text{ALG}}[1,T]]}{\mathbb{E}[\text{Cost}_{\text{LAI}}[1,T]]}$. Note that as the environment becomes more adversarial than stochastic, the expectation has no effect.